Université catholique de Louvain
Secteur des Sciences et Technologies
Institut de Recherche en Mathématique et Physique
Center for Cosmology, Particle Physics and Phenomenology

Development of a new tracker for the CMS upgrade phase 2 and study of the HL-LHC physics reach

Doctoral dissertation presented by
Martin Delcourt
in fulfilment of the requirements for the degree of Doctor in Sciences

Thesis support committee

Prof. Eduardo Cortina Gil	UCLouvain, Belgium
Prof. Christophe Delaere (Supervisor)	UCLouvain, Belgium
Prof. Vincent Lemaitre (Chair)	UCLouvain, Belgium
Prof. Steven Lowette	VUB, Belgium
Dr. Andrea Venturi	INFN, Italy

September, 2020

Remerciements

J'aimerais commencer ce document par remercier tous ceux qui auront rendu ces six dernières années si riches et qui auront participé à façonner à leur manière ce travail.

Avant toute chose, je voudrais remercier Christophe pour m'avoir encadré, conseillé, soutenu et motivé depuis maintenant sept ans. Bien que m'incitant à devenir indépendant, ta porte était toujours ouverte pour moi et tes conseils avisés, ta compréhension et ta bonne humeur m'ont donné l'énergie pour venir à bout de ce doctorat.

I want to express all my gratitude to Eduardo, Vincent, Steven and Andrea for their participation in my thesis committee. Your interesting questions during the private defence and comments on the draft of my thesis were precious to help improving the quality of this document.

Participating to test beam activities and to online operations at Point 5 allowed me to meet many amazing colleagues from all around the world. In particular, I would like to thank Georg, Sarah, Stefano and Tom with whom I spent countless hours in control rooms, for your incredible work and perpetual good mood.

En travaillant toutes ces années au cyclotron, j'ai eu l'occasion de rencontrer de nombreux collègues qui ont permis de rendre cette thèse aussi agréable qu'enrichissante. Tout d'abord, merci à Laurent, Maxime, Valentin, Meriem, Domicien et Khawla pour avoir partagé mon bureau avec autant de bonne humeur que de patience. Merci également à Alex, Brieuc, Christophe, Christophe, Cristian, Disrael, Florian, Jérôme, Julien, Miguel, Olivier, Pieter, Sébastien, Sébastien, Sébastien et tant d'autres pour tous ces temps de midi passés ensemble à parler de tout et de rien, et pour avoir été des victimes de *frag* très coopérantes. De manière générale, je tiens à remercier tous mes collègues du CP3 et de l'UCLouvain pour une magnifique ambiance au travail.

Rien de tout ceci n'aurait non plus été possible sans le soutien sans faille de ma famille. Un énorme merci pour tout ce que vous avez fait pour moi durant toutes ces années, pour avoir toujours attisé ma curiosité, ainsi que pour votre rôle prépondérant dans mon comité de relecture. Merci également à tous mes *vrais* amis pour tous ces bons moments passés ensemble, pour votre soutien et vos encouragements.

Dernièrement, je voudrais remercier l'UCLouvain elle-même pour m'avoir financé à travers un poste d'assistant et avoir rendu ce travail possible. Merci également aux si nombreux étudiants qu'il m'a été permis de rencontrer au cours d'innombrables séances d'exercices et de laboratoires. Bien que chronophages, ces tâches d'enseignement étaient extrêmement enrichissantes, tant sur le plan intellectuel que humain.

Un énorme merci à tous!

Contents

Introduction

The aim of particle physics is to understand what the fundamental components of nature are and how they interact with each other. The best answer we currently have to this question is given by the standard model of particle physics. This theory, developed during the second half of the last century, provides a coherent description of high-energy physics processes and has been hugely successful in providing experimental predictions.

Among the long list of predictions realised to-date by the standard model, the most significant is arguably that of the discovery of the Higgs boson in 2012. A new scalar field was theorised in the nineteen-sixties to explain the origin of mass within the standard model. From this postulate, the existence of a new particle, the Higgs boson, was predicted and its discovery half a century later provided the last cornerstone needed for the standard model to become fully consistent.

Despite huge successes in predicting the behaviour of high-energy processes, the standard model suffers from major shortcomings. Faced with an inability to explain gravity, dark matter and matter/anti-matter asymmetry and some internal theoretical issues, the particle physics community is now looking for new answers beyond the standard model.

Probing new physics phenomena will require experimental input through the study of new theorised processes or the precise measurement of standard model parameters enabling potentially anomalous values to be highlighted. On the path leading towards a better understanding of particle physics, an in-depth precision study of the Higgs boson and its properties is crucial. One measurement of interest which will guide the reader through this document is the combined production of two Higgs bosons allowing for a direct handle on the Higgs self-coupling parameter of the standard model, and as such, a glimpse of the shape of the scalar field potential around its vacuum expectation value.

This relentless work of characterising the properties of the Higgs boson is currently being undertaken at the Large Hadron Collider, where proton collisions at a centre of mass energy of $\sqrt{s}$ =13 TeV are being recorded by dedicated detectors. These experiments provide a continuous improvement to the understanding of the standard model, with new predictions being validated and stronger constraints on its parameters and on new theories being reached every year.

Amid these tremendous achievements, some processes of interest remain too weak to be detected with the current installations, and would lie beyond grasp were nothing to be done. This is, for example, the case of the Higgs boson pair production which would not be expected to be discovered within the lifetime of the present machine.

To maximise the physics reach of the accelerator, it was decided to subject the installations to a major upgrade, thus allowing a strong increase in instantaneous luminosity, leading to a tenfold increase in overall collisions produced. Such a dramatic change will bring major challenges to the experiments recording these collisions and upgrades are required if they are to maintain their outstanding performance.

This document will explore the upgrade of the CMS detector, and in particular its silicon strip sub-detector, centred on the work performed by the author while completing his PhD thesis. The upgrade process involves three major components : a new detector design is required, prototypes based on this design have to be produced and validated, and the resulting detector performance needs to be simulated to assess the validity of the upgrade and the expected physics reach.

While these activities are presented sequentially in this document, they are fundamentally intertwined. The design choices are refined based on results from early prototypes, and simulation output is used to define the detector layout and requirements.

In this context, the author played a significant role in the validation of prototypes for the upgrade of the outer tracker of the CMS detector. The author has actively participated in beam tests, both during the operations and in the related data analysis. Furthermore, he contributed to the system integration efforts with a particular interest in the offline data processing steps and the compatibility of the different data formats involved. This aspect of his work was featured in two publications [1,2], and used as input to the tracker technical design report [3]. The author also published a conference report summarising the CMS tracker upgrade [4].

The author further contributed to the upgrade of the detector through the realisation of a prospective analysis. The Higgs boson pair production in the $H \rightarrow b\overline{b}, H \rightarrow WW \rightarrow l^+ \nu l^- \nu$ final state was studied. This work, described in an internal analysis note [5], was combined with other Higgs boson pair production studies as the object of a public analysis summary [6] and published within the CERN Yellow Report on Higgs physics at the HL-LHC and HE-LHC [7].

The present document is structured as follows : a brief introduction of the standard model will be outlined in chapter 1, followed by a presentation of the current accelerator complex, detector design, data analysis techniques and the reach of the current HH analysis in chapter 2. The upgrade of the detector will then be described in chapter 3, and the author's contribution within the prototype testing campaign will be outlined in chapter 4 with technical details on the read-out chain in appendix A. The prospective Higgs boson pair production analyses are described in the final chapter 5.

Chapter

1. The standard model of particle physics

The goal of particle physics is to understand what the universe is fundamentally made of, and how its components interact with each other. While the work reported in this document is strongly focused on experimental and detector issues, a brief overview of the theoretical aspects is crucial to fully understand the context and the goals of this research.

This section will briefly describe the current understanding of particle physics, as described by the standard model. Its particle content will be outlined, followed by the Lagrangian density describing the behaviour and interactions of the particles. A short elaboration will follow of predictions and how event generators bridge the gap between theory and experiment. Finally, a few issues of the standard model will be indicated and the importance of precision measurements to pave the way for future physics will be presented.

This chapter is mostly based on Peskin & Schroeder's introduction on Quantum Field Theory [8], Mandl & Shaw's book on Quantum Field Theory [9], and Smith's Introduction to the Standard Model lectures [10].

1.1. Fundamental particles

Over the last century, or more specifically between 1897 and 2000, twelve fundamental types, or flavours, of matter particles of spin $\frac{1}{2}$ - or fermions - were discovered. These particles were found to be interacting with each other through three fundamental forces, the strong interaction, weak interaction and electromagnetism. They were also found to be impacted by gravitational fields.

These twelve particles can be split into three generations with similar properties. In each generation, two quarks display a charge under the strong interaction, or colour, and have a $\frac{2}{3}$ and $-\frac{1}{3}$ electric charge. From the two remaining particles, or leptons, one exhibits no electric charge, while the other has a -1 charge. An anti-particle, with the same mass but opposite charges under the different interactions corresponds to each of these particles.

Following this order, the three generations are the following : up (u) and down (d) quarks, electron neutrino (ν_e) and electron (e), charm (c) and strange (s) quarks, muon neutrino (ν_μ) and muon (μ), and finally top (t) and bottom (b) quarks, tau neutrino (ν_τ) and tau (τ). A summary of these particles can be found on table 1.1.

These particles were shown to be interacting with each other through the exchange of spin 1 gauge bosons. The photon (γ) carries the electromagnetic force, the eight gluons (g) the strong force, and the W and Z bosons mediate the weak force (see table 1.2).

Through the action of the $W^{\pm}$ boson, the weak force was shown to couple to pairs of leptons or of quarks leading to a source of flavour violation in the standard model. This explains why particles from the second and third generations tend to decay into their first generation counterparts, and why the neutron is unstable. Furthermore, this interaction was shown to only involve left-handed particles, leading to the following decomposition.

$$L^1 = \begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix} \qquad L^2 = \begin{pmatrix} \nu_{\mu L} \\ \mu_L \end{pmatrix} \qquad L^3 = \begin{pmatrix} \nu_{\tau L} \\ \tau_L \end{pmatrix} \tag{1.1.}$$

$$Q^1 = \begin{pmatrix} u_L \\ d_L \end{pmatrix} \qquad Q^2 = \begin{pmatrix} c_L \\ s_L \end{pmatrix} \qquad Q^3 = \begin{pmatrix} t_L \\ b_L \end{pmatrix} \tag{1.2.}$$

$$E^{1,2,3} = e_R, \mu_R, \tau_R \quad U^{1,2,3} = u_R, c_R, t_R \quad D^{1,2,3} = d_R, s_R, b_R \tag{1.3.}$$

The hypothetical right-handed neutrinos were omitted from this description.

Electric charge	Quarks		
$+\frac{2}{3}$	up (u) $2.16^{+0.49}_{-0.26}$ MeV	charm (c) 1.27±0.02 GeV	top (t) 172.9±0.4 GeV
$-\frac{1}{3}$	down (d) $4.67^{+0.48}_{-0.17}$ MeV	strange (s) 93^{+11}_{-5} MeV	bottom (b) $4.18^{+0.03}_{-0.02}$ GeV
	Leptons		
0	electron neutrino (ν_e) $\sum m_\nu < 0.152$ eV	muon neutrino (ν_μ)	tau neutrino (ν_τ)
-1	electron (e) 511 keV	muon (μ) 106 MeV	tau (τ) 1.78 GeV

Table 1.1. List of fermions in the standard model and their mass and electric charge. The limit on the sum of neutrino masses was taken from [11], all other values are taken from the world average provided by the particle data group [12].

Particle	Gluons (g)	$W^{\pm}$ bosons	Z boson	Photon (γ)
Mass	0	80.4 GeV	91.2 GeV	$< 1\times10^{-18}$ eV
Interaction	Strong force	Weak force		Electromagnetic force

Table 1.2. List of gauge bosons in the standard model, their corresponding force and their respective mass. These latter values were taken from the world measurement average [12], with the exception of the gluon where the theoretical mass was quoted.

1.2. Standard model Lagrangian

The Dirac Lagrangian density, describing the behaviour of free massive fermions can be made to encompass interactions by imposing local gauge symmetry through the following transformation :

$$\mathcal{L}_{\text{Dirac}} = \overline{\psi} i \not{\partial} \psi - m \overline{\psi} \psi \rightarrow \mathcal{L} = \overline{\psi} i \not{D} \psi - m \overline{\psi} \psi \tag{1.4.}$$

where ψ is a fermionic field of mass m, and ∂_μ is promoted to the covariant derivative $D_\mu = \partial_\mu + igA^a_\mu T_a$ with A^a_μ the gauge field of coupling g and T_a are the generators of the symmetry imposed.

In quantum electrodynamics (QED) for example, imposing a local $U(1)$ symmetry allows the $D_\mu = \partial_\mu + igA_\mu$ term to be added to the Lagrangian density that will be kept invariant under $\psi \rightarrow \psi' = e^{i\lambda(x)}\psi$ and $A_\mu \rightarrow A'_\mu = A_\mu - \frac{1}{g}\partial_\mu \lambda(x)$ transformations. Furthermore, it can be shown that the field strength tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is also invariant under this transformation and the interaction term $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$ can thus be added to the Lagrangian density leading to the following expression.

$$\mathcal{L}_{\text{QED}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \overline{\psi} i \gamma^\mu (\partial_\mu + igA_\mu)\psi - m\overline{\psi}\psi \tag{1.5.}$$

The standard model Lagrangian is based on the same gauge invariance principle, requiring a local invariance on $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$.

The first $SU(3)_C$ symmetry describes quantum chromodynamics (QCD), explaining the behaviour of the strong force. The eight generators of this group, given by the Gell-Mann matrices λ_a with $a = 1, \cdots 8$, lead to the existence of eight gauge bosons G^a_μ. The QCD Lagrangian density can thus be written as :

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4}G_{a,\mu\nu}G^{\mu\nu}_a + \overline{\psi}_i \left(i \not{D}^{ij} \right) \psi_j \tag{1.6.}$$

where

$$G_{a,\mu\nu} = \partial_\mu G_{a,\nu} - \partial_\nu G_{a,\mu} - g_s f_{abc} G_{b,\mu} G_{c,\nu} \,, \qquad D^{ij}_\mu = \partial_\mu - i g_s G_{a,\mu} \frac{\lambda^{ij}_a}{2} \tag{1.7.}$$

and ψ_i is a quark of colour i, and f_{abc} is the structure constant of the $SU(3)$ group.

The weak and electromagnetic forces originate from the $SU(2)_L \otimes U(1)_Y$ groups, where L refers to left-handed fermions and Y to the weak-hypercharge. The corresponding fields are the three W^i_μ $(i = 1,2,3)$ and B_μ fields respectively. The structure of their groups lead to the following field strength tensors :

$$W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu + g\epsilon^{ijk} W^j_\mu W^k_\nu \,, \qquad B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{1.8.}$$

where ϵ is the fully antisymmetric structure constant of the $SU(2)$ group. Similarly to QCD, the interaction with fermions will arise from a covariant derivative as

$$\mathcal{L}_{EW} \supset \overline{\psi}^I \left(i \not{D} \right) \psi^I + \overline{\psi}^J \left(i \not{D}' \right) \psi^J \,, \qquad \text{with } \psi^I = L^i, Q^i \text{ and } \psi^J = E^i, U^i, D^i. \tag{1.9.}$$

with the two covariant derivatives acting differently on left- and right-handed fermions :

$$D_\mu = \partial_\mu - ig\frac{\sigma^i}{2}W^i_\mu - ig'\frac{Y}{2}B_\mu \quad (1.10.)$$

$$D'_\mu = \partial_\mu - ig'YB_\mu \quad (1.11.)$$

where g and g' are two coupling constants, σ^i are the Pauli matrices and Y is the weak hypercharge.

At this stage, all particles described are massless. Adding a $m\overline{\psi}\psi$ term to the Lagrangian density, as was done in equation 1.4 would not be allowed since it is not invariant under $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$.

To allow for fermions and gauge boson masses and to recover $U(1)_{\text{em}}$, an additional scalar doublet of hypercharge $Y = 1$ and no colour, ϕ, with potential $V(\phi^\dagger\phi)$ is introduced,

$$\phi = \begin{pmatrix} \varphi^+ \\ \varphi^0 \end{pmatrix} \quad \longrightarrow \quad \mathcal{L}_{\text{scalar}} = D_\mu\phi^\dagger D^\mu\phi - V(\phi^\dagger\phi) \quad (1.12.)$$

$$\text{where } V(\phi^\dagger\phi) = \mu^2(\phi^\dagger\phi) + \lambda(\phi^\dagger\phi)^2 \quad (1.13.)$$

is the most general renormalizable potential respecting gauge symmetry. By requiring that $\mu^2 < 0$, the minimum of the potential is displaced to $\langle 0|\phi^\dagger\phi|0\rangle = \frac{\mu^2}{\lambda} \equiv v^2$ where v is the vacuum expectation value of this scalar field.

Imposing the conservation of electric charge leads to a null vacuum value for the charged field $\langle 0|\phi^+|0\rangle = 0$, and the ground state of the field can be chosen to be $\frac{1}{\sqrt{2}}\begin{pmatrix} 0 & v \end{pmatrix}^T$. By expanding around this minimum through a polar parametrisation, then applying an $SU(2)_L$ rotation, the following field in the unitary gauge is obtained.

$$\phi(x) = \frac{1}{\sqrt{2}}\exp\left(i\sigma^i \cdot \xi^i(x)\right)\begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \rightarrow \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \quad (1.14.)$$

The scalar Lagrangian density can then be explicitly developed as :

$$\mathcal{L}_{\text{scalar}} = D_\mu\phi^\dagger D^\mu\phi - \mu^2(\phi^\dagger\phi) - \lambda(\phi^\dagger\phi)^2 \quad (1.15.)$$

$$= \frac{v^2}{8}\left(g^2W^i_\mu W^{i,\mu} + g'^2B_\mu B^\mu - 2g'gB_\mu W^{3,\mu}\right)\left(1 + \frac{h}{v}\right)^2$$
$$+\frac{1}{2}\left(\partial_\mu h\partial^\mu h\right) - \lambda v^2h^2 - \lambda vh^3 - \frac{\lambda}{4}h^4 - \frac{\lambda v^4}{4}. \quad (1.16.)$$

The first line of equation 1.16 shows that the $W^{1,2,3}$ and B gauge bosons gain a mass term and that the B field mixes with W^3. The physical bosons can then be extracted from the gauge bosons :

$$W^\pm_\mu = \frac{1}{\sqrt{2}}\left(W^1_\mu \mp iW^2_\mu\right) \quad (1.17.)$$

$$A_\mu = \cos\theta_W B_\mu + \sin\theta_W W^3_\mu \quad (1.18.)$$

$$Z_\mu = -\sin\theta_W B_\mu + \cos\theta_W W^3_\mu \tag{1.19.}$$

where $\cos\theta_W = g/\sqrt{g^2+g'^2}$, leading to respective masses of $m_W = \frac{gv}{2}$, $m_A = 0$ and $m_Z = m_W \sin\theta_W$.

Finally, adding this field allows the following Yukawa terms to be added to the Lagrangian density

$$\mathcal{L}_\mathrm{Y} = -\overline{U}^I Y_u^{IJ}\left(Q^J \times \phi\right) - \overline{D}^I Y_d^{IJ}\left(Q^J \times H^c\right) - \overline{E}^I Y_e^{IJ}\left(L^J \times H^c\right) + h.c. \tag{1.20.}$$

$$\rightarrow -\frac{v}{\sqrt{2}}\left(\overline{u}^I_R Y_u^{IJ} u^J_L + \overline{d}^I_R Y_d^{IJ} d^j_L + \overline{e}^I_R Y_e^{IJ} e^J_L\right)\left(1+\frac{h}{v}\right) + h.c. \tag{1.21.}$$

where $I, J = 1,2,3$ and the latest expression is obtained by moving to the unitary gauge. Fermion masses originate from this term in the Lagrangian density, with an added coupling to the h field, proportional to the fermion mass. These terms replace the traditional $m\overline{\psi}\psi$ in the Dirac Lagrangian density (see equation 1.4) without breaking gauge symmetry.

In a compact form, the standard model Lagrangian density can finally be written as :

$$\mathcal{L}_{SM} = \mathcal{L}_{gauge} + \mathcal{L}_{fermions} + \mathcal{L}_{scalar} + \mathcal{L}_{Yukawa} \tag{1.22.}$$

$$= -\frac{1}{4}G^a_{\mu\nu}G^{a,\mu\nu} - \frac{1}{4}W_{\mu\nu}W^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \tag{1.23.}$$

$$+\overline{\psi}^I\left(i\not{D}\right)\psi^I + D_\mu\phi^\dagger D^\mu\phi - V(\phi^\dagger\phi) + \overline{\psi}^I Y^{IJ}\psi^J\phi + h.c..$$

We have shown how adding a scalar field provides masses to otherwise massless gauge bosons and as a by-product, it also provides an origin to fermion masses. Writing the Lagrangian density in the unitary gauge leads to the emergence of the Higgs-field (h) in the following expression, and predicts the existence of a corresponding massive particle, the Higgs boson.

$$\mathcal{L}_\mathrm{scalar} \supset \frac{1}{2}\left(\partial_\mu h \partial^\mu h\right) - \lambda v^2 h^2 - \lambda v h^3 - \frac{\lambda}{4}h^4 - \frac{\lambda v^4}{4} \tag{1.24.}$$

where $v^2 \equiv \frac{\mu^2}{\lambda}$. The second term in this equation provides a mass term to the Higgs field with $m_h = \frac{1}{2}\lambda v^2$, and the following terms imply Higgs trilinear ($\lambda v h^3$) and quartic ($\frac{\lambda}{4}h^4$) couplings.

The vacuum expectation value can be measured from the masses of gauge fields and was found to be of about 256 GeV [12]. The discovery in 2012 of the Higgs boson [13,14] with a mass of 125.10(14) GeV [12,15–17] allowed to fix the remaining free parameter.

The coupling of the Higgs boson to the W and Z bosons, top and bottom quarks, and τ and μ leptons were measured [18,19] to be compatible with the theoretical expectations as shown on figure 1.1, thus confirming the role of the scalar field in yielding a mass to fundamental particles.

While the Higgs quartic coupling is far from being within our grasp, the trilinear

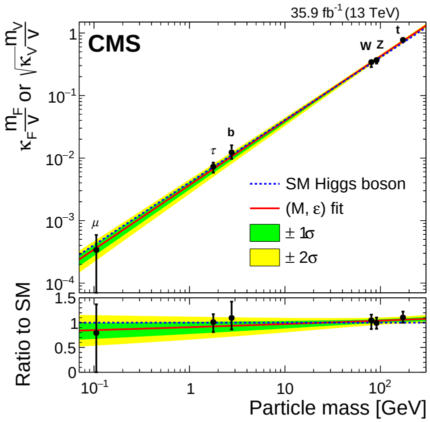


Figure 1.1. Measurement of the coupling of the Higgs boson with top and bottom quarks, W and Z bosons and τ and μ leptons performed by the cms collaboration [18]. The agreement with the theory confirms the role of the scalar field in providing a mass to fundamental particles.

coupling $\lambda^{\mathrm{SM}}_{\mathrm{HHH}} = \frac{m_H^2}{2v}$ is expected to be close to being measurable. A direct measurement of Higgs boson pair production (also know as di-Higgs) would allow to constrain the shape of the potential close to its minimum.

1.3. Predictions in particle physics

Measurements in high-energy physics are often performed through scattering experiments. Two particles a and b of momentum p_a and p_b in an initial state $\langle i|$ are made to collide and produce a final state $|f\rangle$ of particles with momentum $\left\{p_f\right\}$. If the interacting Hamiltonian H is known, the probability of transitioning between these two states can be written as

$$\mathcal{P}\left(i \to f\right) = \lim_{t\to\infty} \left|\left\langle f \left| e^{-iH2t} \right| i \right\rangle\right|^2 \equiv \left|\left\langle f \left|S\right| i \right\rangle\right|^2 . \tag{1.25.}$$

From the S matrix, the part due to interactions can be extracted as $S = 1 + iT$, and we can define the invariant matrix element $\mathcal{M}$ as :

$$\left\langle f \left|iT\right| i \right\rangle = (2\pi)^4 \, \delta^{(4)}\left(p_a + p_b - \sum p_f\right) \mathcal{M}\left(p_a, p_b \to \left\{p_f\right\}\right) \tag{1.26.}$$

with the energy and momentum conservation requirements being factored out. Since in collider experiments the impact parameter in the initial state is not known, it is common to integrate it out and consider the microscopic cross-section :

$$\begin{aligned} d\sigma \quad = \quad & \frac{1}{2E_a 2E_b \left|v_a - v_b\right|} \left(\prod_f \frac{d^3 p_f}{(2\pi)^3} \frac{1}{2E_f} \right) \left| \mathcal{M}\left(p_a, p_b \to \left\{p_f\right\}\right)\right|^2 \\ & \times (2\pi)^4 \delta^{(4)}\left(p_a + p_b - \sum p_f\right) \end{aligned} \tag{1.27.}$$

where the first fraction comes from the integral over the impact parameter in the laboratory reference frame, and the product on the d^3p_f expresses the final state. All the interactions, or the dynamics of the system are enclosed in the matrix element $\mathcal{M}$.

To compute this latter value, the definition of the $\mathcal{S}$ matrix can be written as a series known as the Dyson expansion :

$$\mathcal{S} = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int \cdots \int d^4x_1 \cdots d^4x_n T \left\{ \mathcal{H}_I(x_1) \cdots \mathcal{H}_I(x_n) \right\} \tag{1.28.}$$

where T denotes time-ordering, and $\mathcal{H}_I$ the interaction Hamiltonian density. Using Wick's theorem, this expression can be re-written as a sum of Feynman propagators that can be individually computed and represented as a Feynman diagram. With each term of this series adding an additional interaction, if the coupling constant is small enough ($\alpha = g^2/4\pi \ll 1$) the cross section can be computed perturbatively by truncating the series at a given order.

$$d\sigma = d\sigma_{LO} \times \left(1 + \frac{\alpha}{2\pi} d\sigma_1 + \left(\frac{\alpha}{2\pi} \right)^2 d\sigma_2 + \cdots \right) \tag{1.29.}$$

In this last equation, LO designates the leading order, and the cross section computed using the subsequent terms will be referred to as next-to-leading-order (NLO), next-to-next-to-leading order ($NNLO$), and so on. As an example, the leading order Feynman diagrams for the $gg \rightarrow$ hh process are shown on figure 1.2. At higher orders, additional partons in the final state may arrise to avoid infrared divergences.

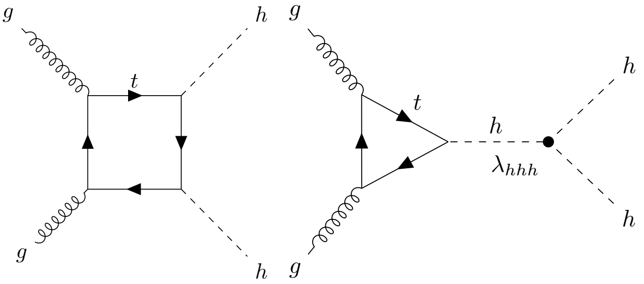


Figure 1.2. Dominant Feynman diagrams involved in the gg → hh process.

In hadron colliders, the particles made to collide are not fundamental, but can be seen as complex assemblies of quarks and gluons. Computing the matrix element for such objects can be very challenging, especially given the non-perturbative nature of QCD at low energies. Fortunately, the very different energy and time scales between the hard process involving high momentum transfers between partons (gluons and quarks) and lower momentum transfers, or soft processes occurring in the hadrons enables the two contributions to be factorised.

The hadronic parton content is encoded in the parton distribution function (p.d.f.). These probability distributions are assumed to be universal and are obtained from collider measurements. They are expressed as the probability density of finding a parton of flavour a carrying the fraction x of the hadron total longitudinal momentum. Here, a can refer to either a gluon or any possible quark, even outside of the hadron's valence quarks.

The microscopic cross-section is therefore to be integrated over the p.d.f.

$$d\sigma_{hh\to X} = \sum_{a,b=g,q_i} \int dx_a dx_b f_a(x_a) f_b(x_b) d\sigma_{a,b\to X}(x_a, x_b) \qquad (1.30.)$$

where X refers to a desired final state.

No detector can record the immediate output of the hard-process, causing any short-lived particle to decay and hadronization processes to take place. The same factorisation arguments can then be made to facilitate the computation of the final state, allowing these steps to be computed separately from the hard process.

In parallel to this simple picture, any other process providing a similar final state has to be considered, including possible additional radiation from the incoming or outgoing particles, known as initial state and final state radiations respectively, possible multiple parton interaction, the behaviour of the remaining particle content of the initial hadrons, and the evolution and decay of the outgoing particles. A summary picture representing the complexity of hadron collisions is shown on figure 1.3.

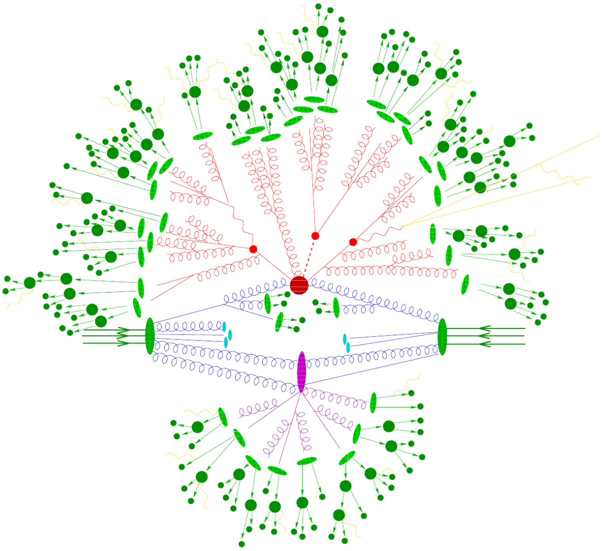


Figure 1.3. Summary picture of the complexity of event generation in hadron collisions. The red part of this diagram represents the hard process, surrounded by hadronization in dark green and the emission of soft radiation in yellow. The purple part shows a secondary parton interaction, while the teal represents the underlying event [20].

Given the complexity of the task at hand, an analytical integration of the complete expression is not possible. To bridge the gap between the theoretical Lagrangian and experimental observables, event generators are used.

Dedicated softwares, such as MadGraph5_amc@nlo [21] or powheg [22], are tasked to convert either the standard model Lagrangian or any alternative model into a set of Feynman rules, and to compute the corresponding matrix element of a given process at a fixed order. These matrix elements can then be used to compute a cross-section, or to generate events using Monte-Carlo techniques [23]. Hadronization, initial and final state radiation and particle decay can then be simulated on these events.

The matrix element event generation and evolution of outgoing particles can either be simulated by a unique software, or more commonly by separate tools that use a common event interface. As such, while the `pythia` [24] generator is able to perform the whole generation, it is common to use MadGraph5_amc@nlo for the hard process and `pythia` for the hadronization.

The predictions provided by the standard model are being thoroughly tested by a wide range of experiments. As an example, a summary of standard model cross section predictions and corresponding measurements by the CMS collaboration is shown on figure 1.4. A remarkable agreement can be seen spanning nine orders of magnitude.

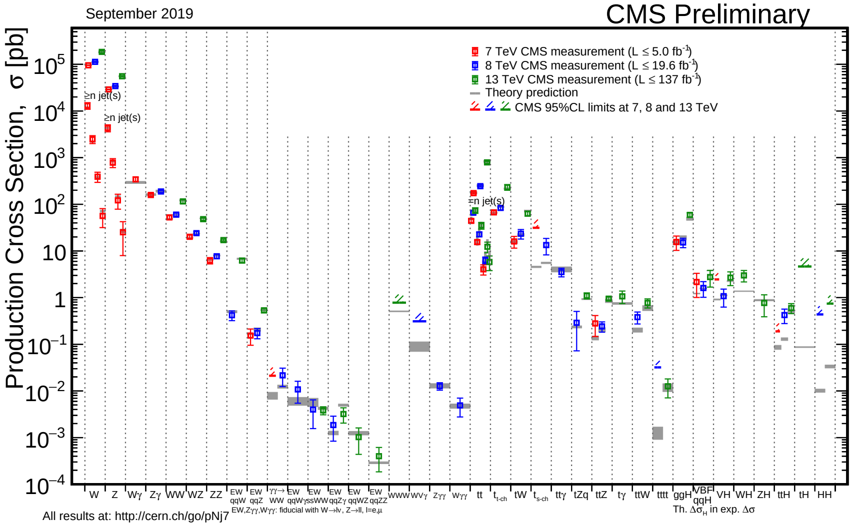


Figure 1.4. Summary plot of standard model theoretical cross sections and corresponding measurements or limits observed by the CMS collaboration [25].

Throughout the years, the standard model has proven to be hugely successful in predicting the existence of new particles and phenomenon, and in providing a wide range of precise and experimentally validated predictions.

1.4. Issues in the standard model

With all its merits and successes over the years, some remaining questions in particle physics indicate that the standard model alone cannot be the answer. A few striking examples will be outlined here, but this selection is in no way exhaustive nor is it intended to reflect a general consensus on the biggest issues ahead within the theoretical community.

- **Quantum gravity.** With all its predictive power, there are some phenomena that cannot be explained by the standard model. A first striking example, is the inability to include the gravitational interaction within the model. The traditional approach of trying to explain this force through the addition of a hypothetical boson, the graviton, would lead to a non-renormalizable theory, meaning that predictions of observable quantities would diverge.
- **Dark matter.** Astronomical observations and cosmological models have shown that about 85% of the matter content of the universe should be made of some additional matter that does not interact electromagnetically (hence the name) and that is not accounted for in the standard model.
- **Matter-antimatter asymmetry.** Astronomical observations indicate that antimatter is a very rare occurrence in the universe compared to ordinary matter. While matter-antimatter asymmetry exists within the standard model, it is too small and unable to explain the observed asymmetry if one assumes a symmetric production in the early universe.
- **Strong CP problem.** Setting aside unexplained observations, some theoretical issues remain in the standard model. The theory is built based on gauge symmetry. Any expressions that respect gauge symmetry and the hypotheses of the standard model are added to the Lagrangian. One of these terms should lead to some CP symmetry violation through the strong force, but no such asymmetry could be observed in nature. While this can be fixed by assigning a parameter θ_{CP} describing the strength of this violation, this parameter is often considered to be unnaturally small and in need of some additional explanation.
- **Hierarchy problem.** When renormalizing the mass of the Higgs boson to take into account quantum loop corrections, quadratic contributions in the regularisation parameter Λ are expected. Since nothing prevents this parameter to extend up to the Plank scale, the bare mass of the Higgs boson would have to be incredibly fine-tuned to yield a physical mass of 125 GeV. This problem, known as the hierarchy problem hints at some additional phenomenon being needed to explain the low Higgs mass.

Numerous new models have been proposed to address these issues, known as theories beyond the standard model (BSM). New experimental inputs are crucial to cast some light on which theories could best extend the standard model.

Two different approaches can be used to discover or set limits on these models. The first, quite obviously, is through direct observation of some phenomenon or particle. A second indirect approach would be to probe phenomena at energy scales higher than currently available in collision experiments through the precise measurement of observable quantities. The existence of new particles or new couplings could have some impact on standard model couplings through loop corrections. It is therefore crucial to measure them as precisely as possible. The di-Higgs production in hadron colliders provides once again a good example.

As shown on figure 1.2, two concurrent processes co-exist, through either Higgs self-coupling or through a quark loop. The occurrences of these processes are intrinsically small, and a matrix element computation shows that they will interfere destructively, leading to an even fainter overall production cross section. While it renders its observation quite challenging, this process remains an ideal candidate to probe new physics.

The resonant production of two Higgs bosons is predicted by a number of extensions to the standard model. In the two-Higgs doublet model [26], as the name suggests, an additional complex $SU(2)$ doublet is added, allowing the existence of a new scalar H that can decay into two Higgs bosons if sufficiently heavy. Another example comes from the Randall-Sundrum models [27] based on a warped extra dimension, leading to the possibility of either a spin-0 (radion) or spin-2 (graviton) state decaying into two Higgs bosons.

In the non-resonant approach, the impact of processes involving an energy scale Λ greater than what is available in the collider is parameterised in effective field theories (EFT). A historical example is given by Fermi's interaction describing beta decay that was later understood as a low-energy EFT corresponding to the weak force.

There are many ways to understand and parameterise the standard model as an effective field theory, depending on underlying assumptions and the desired interpretation of the data. A few examples are given in references [28–31]. In this document, the κ-framework [32] parameterization of the standard model will be used.

The idea behind this framework is to assume the structure of the standard model while allowing the strength of the coupling constants to vary. For a given process j, a coupling modifier κ_j is defined as $\kappa_j^2 = \sigma_j / \sigma_j^{\mathrm{SM}}$. Interpreting measurements in terms of these κ parameters allows to verify if the observations are compatible with a standard model expectation of $\kappa_j = 1$. This framework also provides an understanding of how constrained the standard model parameters are. If discrepancies are found with respect to the standard model, this framework enables a quantification but other approaches are required in order to provide an interpretation in terms of higher energy processes.

This work will explore how strongly the Higgs boson self-coupling modifier κ_λ can be constrained at the LHC in the coming decades. A measurement of this parameter is of particular interest both for its rich potential to shed some light on new physics processes, and also because it is the only way of directly observing the shape of the scalar field around its minimum value.

Chapter

2. Experimental high energy physics with the CMS detector

In this chapter, the way the standard model is being experimentally investigated will be described, through the example of Higgs boson pair production (di-Higgs) measurements at the CMS experiment, and in particular its $b\overline{b}\ \ell\ \nu\ \ell\ \nu$ decay channel.

While many experiments allow different aspects of the current theory to be probed, we will focus here on the ongoing activities at the CMS experiment located at the CERN facilities. A brief overview of the facilities and in particular of the large hadron collider will be provided. Then, after a broad description of the CMS experiment, data acquisition and reconstruction techniques, common analysis techniques will be presented. Finally, di-Higgs measurements at this experiment will be outlined, underlining both the reach and limitation of the current facilities.

2.1. CERN and the Large Hadron Collider

After the second world war, the scale of experiments in high energy physics and the wish to restore science in Europe lead to the conclusion that a new international laboratory was needed. A council for European nuclear research was founded (Conseil Européen pour la Recherche Nucléaire, CERN), uniting physicists throughout twelve European countries. A few years later, in 1955, the first stone of this laboratory was laid in the town of Meyrin, near Geneva, on the border between Switzerland and France.

While initially focusing on nuclear physics, this laboratory gradually evolved with the theoretical progress in high energy physics and moved towards particle physics. Although both activities are still being researched at CERN, the majority of the laboratory is now working on particle physics.

During the early eighties, while the 27 km lepton accelerator LEP was being built at CERN, it became increasingly clear that a high energy hadron collider would eventually become the next discovery machine in high energy particle physics. In 1984, at the International Committee for Future Accelerators seminar, the projects of a Large Hadron Collider [33] (LHC) in the LEP tunnel and of an American 87 km Superconducting Super Collider were presented. While the LEP ultimately collided electrons and positrons up to a center of mass energy of 209 GeV, the two hadron colliders projects were envisionned to allow for proton collisions at 14 and 40 GeV respectively. Although being initially approved, the Superconducting Super Collider project was eventually abandoned.

A decade later, in December 1994, the LHC project proposed a year earlier in the so-called "white book" [34] was approved by the CERN council. The layout proposed one year later in the conceptual design [35] shown on figure 2.1 included two high-luminosity experiments, the Compact Muon Solenoid (CMS) and A Large Toroidal Lhc ApparatuS

(ATLAS), as well as A Large Ion Collider Experiment (ALICE) dedicated to lead collisions, and a low luminosity experiment (LHCb).

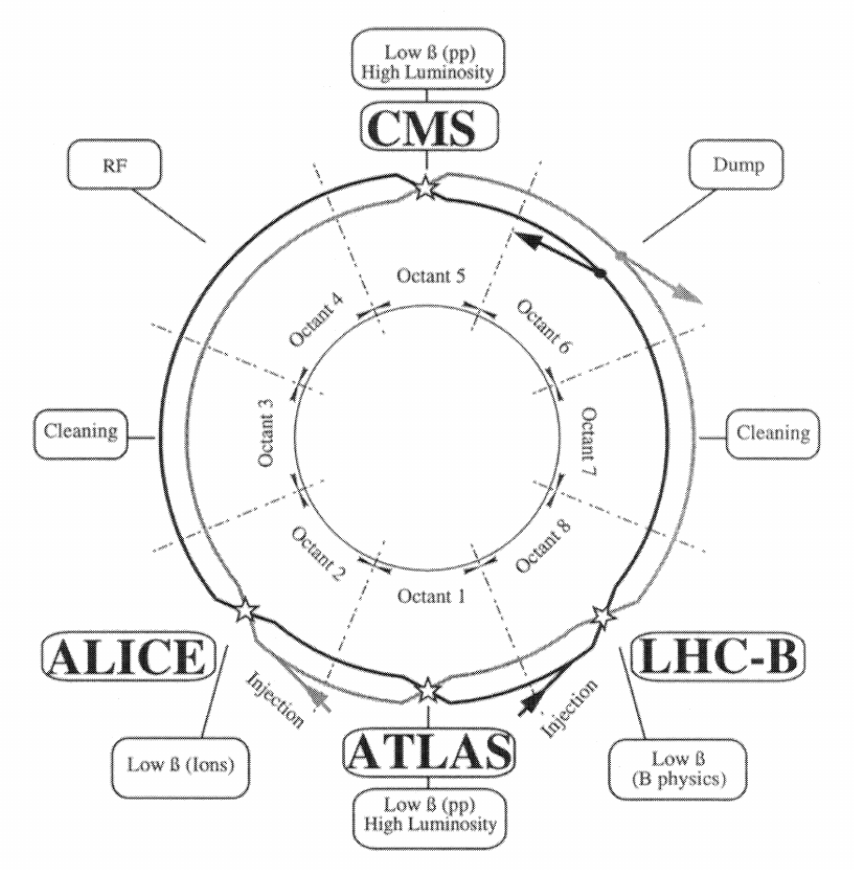


Figure 2.1. | Schematics for a large hadron collider proposal, as described in the conceptual design [35] published in 1995.

The machine was ultimately built, provided its first collisions end of 2009, and reached the world record centre of mass energy of 7 TeV in 2010. It has since risen up to 13 TeV.

While the LHC will be the main focus of this document given the high energy requirement of the process being studied, a wide range of other experimental conditions are made available at CERN through its accelerator complex, shown on figure 2.2.

2.1.1. The LHC injectors

No single accelerator would be able to accelerate protons up to a centre of mass energy of 13 TeV alone. While the LHC provides the final energy increase and beam tuning, a series of smaller accelerators are needed to provide proton and ion beams of the required energy and parameters to the LHC. These are referred to as injectors.

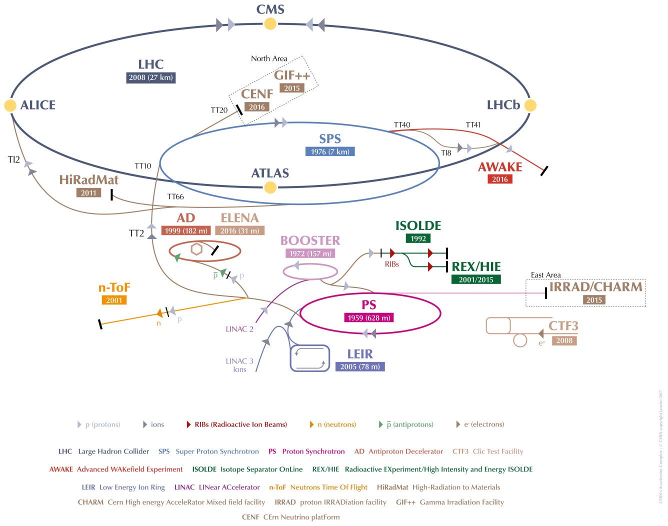


Figure 2.2. | Schematic view of the CERN accelerator complex, as installed between 2016 and 2018 [36].

The first acceleration step is provided by the Linear Accelerator 2 (LINAC 2). A dihydrogen source is connected to a duoplasmotron where the H_2 gas is converted into plasma from which a proton beam is extracted by a charged grid. This continuous 90 keV proton beam is injected into a radio-frequency quadrupole bunching the protons into packets (called bunches), focusing the beam, and accelerating these packets to 750 keV. Then the beam is injected into an Alvarez Drift-Tube Linac accelerating the beam to 50 MeV.

Linear acceleration becomes increasingly impractical and costly with rising energy, therefore the rest of the acceleration is performed by synchrotrons. In these machines, the beams are maintained on a circular path through the use of dipole magnets, and a single accelerating cavity is required. With each rotation in the machine, the beam will gain energy and the current in the dipole magnets will be raised accordingly. The transverse properties of the beam are assured by the usage of quadrupole magnets. Passing through these, the beam is focused along a transverse direction, and defocused along the other. Creating a lattice with dipoles and drift sections and also quadrupoles in an alternating focusing direction, provides the basic tool needed to focus a beam on a circular trajectory.

The first circular accelerator to have been built at CERN is the proton synchrotron (PS). This 628 m long machine was for a brief period of time the highest energy accelerator in the world, but has since been used as an injector for other larger machines.

Directly injecting protons from the LINAC 2 would severely limit the number of bunches that can be accelerated in the PS and the overall beam intensity. An intermediate accelerator made of four superposed 157 m long rings (1/4 of the length of the PS) was built. This accelerator is named the proton synchrotron booster (PSB). When used for LHC filling, each ring accelerates one high-intensity bunch at a time to an energy of 1.4 GeV. A first injection of four bunches is followed by a second injection of two bunches 1.2 s later. These six bunches and an additional empty gap are evenly spaced in the PS.

Alongside its role of accelerating the beam, the PS also has to provide the beam structure required by the LHC. Firstly, at the injection energy of 1.4 GeV, a process known as RF gymnastics is performed to split every bunch into three. Then, the beam is accelerated to the extraction energy of 25.1 GeV before a second step of RF gymnastic splitting every bunch by a factor four. Finally, a rotation in the $\Delta E, \Delta t$ phase space is performed to shorten the bunches from 11 ns to 4 ns. This whole procedure is performed in 3.6 seconds.

After leaving the PS, proton bunches are injected to a 6.9 km machine located in an underground tunnel across the French-Swiss border, the super proton synchrotron (SPS). When filling the LHC, a machine operation cycle, called super-cycle, is as follows: between two and four batches of 72 bunches are injected with a 3.2 s interval between each batch. The beam is then accelerated until it reaches an energy of 450 GeV. At that point, the machine gets prepared for extraction, finely tunes itself with respect to the LHC in order to inject the bunches at the correct position and reshape the bunches by longitudinally compressing them and cutting their tails. After the extraction, the 21.6 s super-cycle is ended by ramping down the magnets to get ready for the next injection.

This initial injection scheme, that could in theory provide two beams of up to 2808 bunches, was changed in 2016 to favour less but more intense bunches. This is performed through the Batch Compression, Merging and Splitting (BCMS) scheme. The PS is filled with two batches of four packets from the PSB that are placed on harmonic 9 (8 bunches + 1 empty bunch). The beam energy is raised to 2.5 GeV and, at that point, the harmonic is raised to 14, thus compressing the batch of 8 bunches. Bunches are then merged by moving to harmonic 7, leading to four bunches of twice the initial intensity. The usual triple splitting, energy increase and quadruple splitting are then performed, leading to 48 bunches produced per cycle.

2.1.2. The Large Hadron Collider

The Large Hadron Collider is the final accelerator and collider of the chain. Lodged in a 27 km long tunnel, approximately 100 m under the French-Swiss border, it is the biggest and most energetic collider in the world to date. The base elements of this machine are the main dipole magnets. Two beam pipes are placed inside Nb-Ti superconducting dipoles, cooled down at 1.9 K by super-fluid helium. Vertical magnetic fields in opposite directions of up to 7.7 T maintain two beams travelling in opposite directions on their respective trajectories. The proton beams are focused by 392 quadrupoles, and are accelerated to an energy of 6.5 TeV by eight superconducting radio-frequency cavities each. The lattice is also completed by 688 sextupoles and 168 octupoles corrector magnets.

The beams are made to collide in four different locations, as initially foreseen and shown on figure 2.1. Two of these interaction points, one located near the Swiss town of Meyrin, the other next to the French village of Cessy, are dedicated to high luminosity experiments. Around these, the two beam pipes merge into one, and the beams are brought closer together via dipoles located approximately at 140 m and 75 m on each side of the interaction point. Then, quadrupole magnet triplets located at 23 m of the interaction point provides the final focusing to ensure the correct beam properties.

A cleaning system is required in order to continuously remove beam halo that would cause superconducting magnet quenches and high backgrounds in the experimental area if left unchecked. This is done through a collimation system, primarily located in the insertion point 3 and 7 regions. Furthermore, these collimators serve as a passive protection system against abnormal beam losses. The system was design to sustain an energy deposit of up to 2 MJ over 200 ns in the event of major equipment failure.

Since the LHC beams can have a total energy of up to 350 MJ, a reliable and safe way of discarding the beam, called the beam dump, is required. Two 7 m long water-cooled carbon cylinders of 700 mm diameter and enclosed within 750 tons of steel and concrete shielding are used as absorption material. In the event of a beam dump, fifteen magnets with a rise-time of 2.85 µs are fired, providing the horizontal kick needed to steer the beam towards the absorber. On its way towards the absorber, the beam will then be deviated by magnets generating sinusoidal fields. The beam will therefore adopt a trajectory in the absorber shaped as the letter e, optimising the deposited energy spread. A 3 µs gap is left empty in the LHC beam structure to accommodate for the kicker magnet rising time. In the unlikely event of an asynchronous beam-dump procedure, additional shielding material is placed between the accelerator and the absorber to protect downstream equipment.

A typical LHC sequence will consist of the following. A first, low intensity beam, named the probe beam, will be injected in the machine to perform measurements and checks on the different accelerator sub-systems. If a safe and stable circulating beam is established, it will be dumped and the physics beam will then be injected. During the later years of exploitation, each beam would comprise up to 2556 bunches. Once the injection from the SPS is completed, the radio-frequency cavities will progressively increase the beam energy from the initial 450 GeV up to 6500 GeV in roughly twenty minutes. Once this maximum energy is reached, the beam is focused at the interaction points in a stage called squeezing. It takes about 15 minutes to obtain the required parameters and an extra 15 minutes to bring the beams together during the adjust period.

Once the beams are colliding in a safe fashion, stable beams are declared, allowing the experiments to record the collisions produced. While this state could in theory be kept for a long period of time (beams were stable for 37h10 in November of 2017), it is not in the best interest of data-taking due to the decrease of the beam intensity from the proton-proton collisions. In practice, this state is kept on average for 8 hours before the beams are manually or automatically dumped and the procedure is restarted.

2.1.3. Luminosity and pile-up

The most important characteristic of a collider, besides its centre of mass energy, is probably its luminosity. The instantaneous luminosity of a collider is defined as the production rate of a given interaction divided by the cross section of this process. For a Gaussian beam, the luminosity can be expressed as

$$\mathcal{L} = \frac{N_b^2 n_b f_{\mathrm{rev}} \gamma}{4\pi \epsilon_n \beta^*} F \tag{2.1.}$$

with N_b referring to the number of charges per bunch ($\sim 1.1 \cdot 10^{11}$), n_b the number of bunches circulating in each beam (2544), f_{rev} the revolution frequency (11.2455 kHz), γ the relativistic parameter (6930), ϵ_n the normalised transverse beam emittance ($\sim 2\,\mu\mathrm{m\,rad}$) and β^* the beam β function evaluated at the interaction point ($\sim 0.3\,\mathrm{m}$). The values quoted in parentheses correspond to typical beam parameters at the start of fills during the end of 2018.

Since the beams do not collide head-on but at a slight angle, an additional geometrical factor F defined as follows has to be considered.

$$F = \frac{1}{\sqrt{1 + \left(\frac{\theta_c \sigma_z}{2\sigma^*}\right)^2}} \tag{2.2.}$$

In this equation, θ_c is the total crossing angle ($260\,\mu\mathrm{rad}$), σ_z the RMS bunch length ($\sim 70\,\mathrm{mm}$) and $\sigma^* = \sqrt{\epsilon_n \beta^* / \gamma}$ the transverse RMS beam size at the interaction point ($\sim 15\,\mu\mathrm{m}$). This geometrical factor leads to a $\sim 15\%$ decrease in luminosity.

The beam parameters changing with respect to time, this quantity is also referred to as the instantaneous luminosity ($\mathcal{L}_{\mathrm{inst}}$). Its integral with respect to time is the integrated luminosity $\mathcal{L}_{\mathrm{int}} = \int \mathcal{L}_{\mathrm{inst}} dt$.

Some of the parameters involved in the definition of luminosity are difficult to measure precisely. In order to provide a real-time measurement of the luminosity, an alternative method is used. The experiments at the interaction points have identified and implemented luminometers, measurements of quantities that have a rate proportional to the instantaneous luminosity.

$$R = \sigma_{\mathrm{vis}} \mathcal{L}_{\mathrm{inst}} \tag{2.3.}$$

These measurements are provided continuously during operations to provide real time feedback to monitor the machine performance. The luminosity measurement is also crucial during data analysis as it is needed to either compute a cross section or to set an upper limit on a process production rate.

The proportionality factor σ_{vis} in equation 2.3, the visible cross-section, has to be precisely calibrated if an accurate measurement of luminosity is to be provided.

The luminosity calibration is performed during a special procedure named the Van der Meer scan [37,38]. The rate of the luminometers are measured while the beam positions with respect to each other are changed. This allows a direct measurement of the beam parameters and thus access to the absolute luminosity.

When the accelerator was designed, the envisaged instantaneous luminosity was $\mathcal{L}_{\mathrm{nominal}} = 1 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. Since its launch ten years ago, the performance and reliability of the machine has been constantly improving. As a result of the tremendous efforts of the accelerator department, this luminosity goal was not only reached, but surpassed. The machine is now able to deliver luminosities up to twice this nominal value, with filling schemes involving less colliding bunches, leading to even higher collision multiplicities during bunch crossings.

Since the rate of collisions is proportional to the luminosity while the frequency of beam crossings is constant, the number of collisions per beam crossing will vary with luminosity. This quantity is referred to as pile-up. While physics analyses gain in sensitivity when integrated luminosity is increased, an increased pile-up leads to difficult experimental conditions and can decrease reconstruction performance. The pile-up distribution recorded by the CMS experiment is shown on figure 2.3.

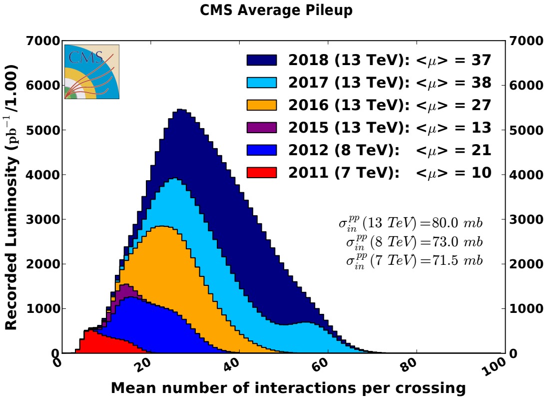


Figure 2.3. | Stacked distributions of the average number of collisions per bunch crossing from 2011 to 2018 [39].

For this reason, the maximum luminosity delivered to the experiments can be throttled through luminosity levelling. By continuously tuning either the separation between the beams, their crossing angle, or the β^* parameters, the luminosity is maintained at the maximum acceptable level during the beginning of the fill. The bunches are then slowly being depleted and the machine gradually returns to its maximum capabilities.

Data taking periods at the LHC are organised in three-to-four-year-long runs, with collisions happening between the spring and the end of the year. The winter period is dedicated to technical stops. Between runs, long shutdowns allow the machine and the experiments to undergo major maintenance and upgrades.

The first LHC run occurring between 2010 and 2012 at a centre of mass energy of 7 and 8 TeV allowed the CMS detector to collect nearly 30 fb^{-1}. After the first long shutdown period, the energy was raised to 13 TeV, and CMS recorded 113.6 fb^{-1} during the run 2 period between 2015 and 2018. A second long shutdown to improve the machine luminosity, increase the centre of mass energy to 14 TeV, and allow detectors to perform upgrades is currently ongoing.

The luminosity recorded per year by the CMS detector is shown on figure 2.4.

2.2. The Compact Muon Solenoid experiment

The Compact Muon Solenoid detector is one of the two high-luminosity, general purpose detectors installed at the LHC. The detector is operated by a collaboration of more than

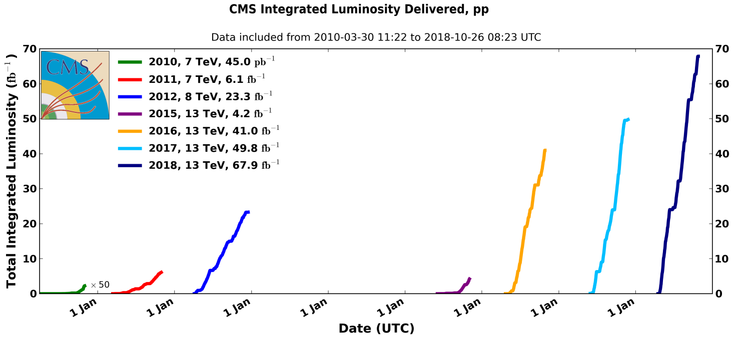


Figure 2.4. Progression of the yearly integrated luminosity from 2011 to 2018 [39].

5500 members at the time of writing, including 28 from the UCLouvain university.

In this section, a brief overview of the detector will be presented, followed by a more detailed description of the different sub-systems. Data acquisition systems and reconstruction methods will then be described and typical analysis methods will be shown. An analysis example will be provided by an early run-2 effort to measure di-Higgs production.

2.2.1. Overview

The CMS detector is a 21 m long cylinder with a diameter of 15 m, weighing about 14000 t. The detector is located in an experimental cavern a hundred meters underground near the small French town of Cessy. The read-out system and some detector services (powering, gas, etc.) are located in a second, isolated, cavern called the service cavern.

Two coordinate frames are used, as shown in figure 2.5. A right-handed Cartesian system is defined as follows. The z axis is aligned with the beam and pointing in a anti-clockwise direction, the y axis is vertical and in the upward direction, and the x axis is horizontal and points towards the centre of the LHC. The plane formed by the x and y directions is referred to as the transverse plane.

To exploit the symmetry of the experiment, spherical coordinates are often used. In this system, a point is defined through its distance to the centre of the experiment, its azimutal angle in the transverse plane with respect to the x axis, ϕ, and its angle in the $z-y$ plane with respect to the z axis, θ. The pseudorapidity is defined as $\eta = -\log\left[\tan\left(\frac{\theta}{2}\right)\right]$, and the radius or angular distance between two objects is given by $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$.

An overview of the CMS detector is shown on figure 2.6.

The central device around which the rest of the experiment was designed and built is a superconducting solenoid. A current over 18000 A circulates through this magnet to produce a 3.8 T magnetic field aligned in the z direction. The solenoid is 13 m long and has a diameter of 6 m. Coils of niobium-titanium enclosed in aluminium are placed within a cold-box cooling the magnet to 5 K [42].

Particles produced by the collisions occurring in the centre of the detector travel outwards

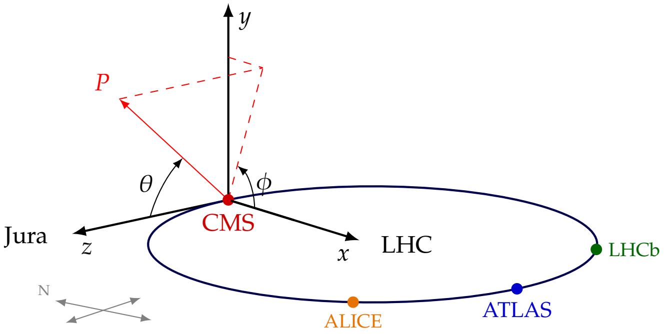


Figure 2.5. | CMS coordinate system definition. Taken from [40].

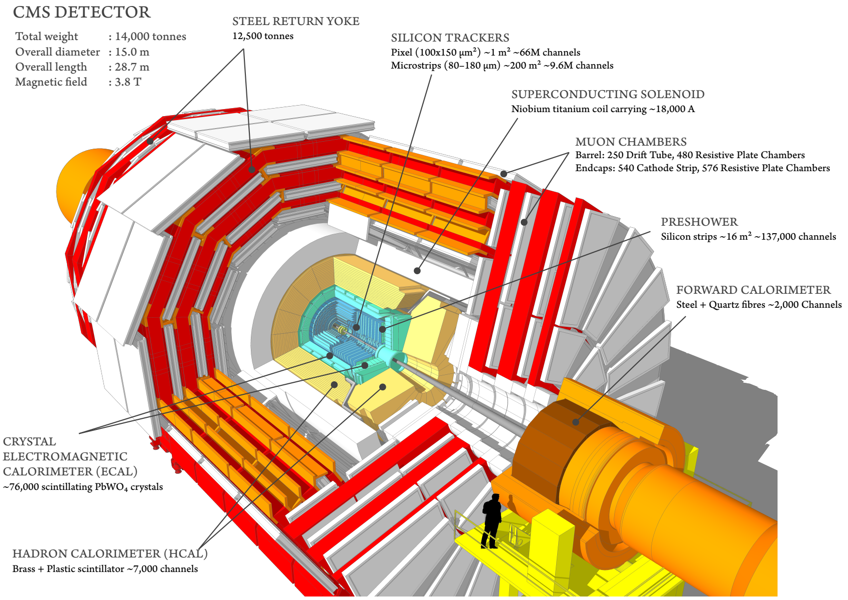


Figure 2.6. | Overview of the CMS detector [41].

and cross its different layers.

Firstly, a silicon tracker measures the path of charged particles. Given the intense magnetic field present, they will be curved in the transverse plane, and the measurement of this curvature is a proxy for the charge and transverse momentum of the particle.

Then, an electromagnetic calorimeter made of lead tungstate crystals will stop and measure the energy of electrons and photons. Heavier particles will have most of their energy remaining after exiting this detector and enter the hadronic calorimeter. This sampling calorimeter uses brass as absorption layers and plastic scintillators as sensitive material.

Gas detectors are located outside of the solenoid to precisely identify and measure the momentum of muons exiting the detector. Multiple flavours of gas detectors serving different purposes and accommodating different geometries are placed within layers of iron guiding the magnetic field.

A bunch crossing rate of 40 MHz and a raw event size of the order of the megabyte leads to the overwhelming total data rate of the order of 50 GBs^{-1}. Given that such a tremendous data rate is impossible to either store or even read-out, a two-staged trigger system has been developed to select events of interest. The first-level trigger receives low resolution information of some sub-detectors and selects which events to read-out. The second-level, or high-level trigger, receives the full detector data of these selected events, performs a partial event reconstruction and selects events to be stored. These two systems reduce the event rate from 40 MHz to about 1 kHz.

2.2.2. The silicon tracker

The innermost sub-system of the CMS experiment is a silicon tracker comprised of two components, a pixel detector surrounded by a silicon strip tracker (SST). A sketch of one quarter of the detector is shown on figure 2.7.

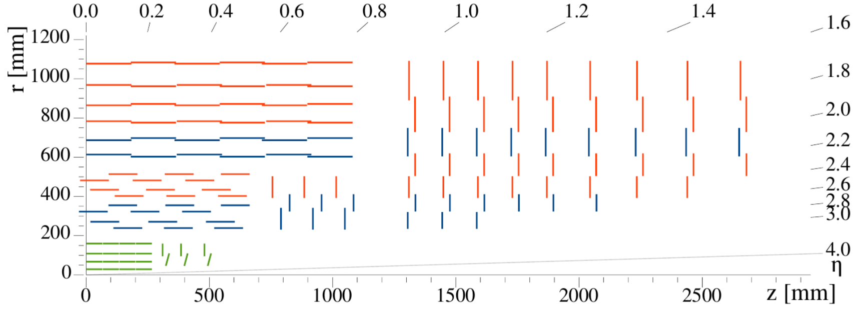


Figure 2.7. | Sketch of one quarter of the silicon tracker in the $r-z$ plane. The phase-1 pixel detector modules are shown in green, the silicon strip stereo modules in blue, and the single sided silicon strip modules in red. [3].

2.2.2.1. The pixel detector

Being the innermost detector of the experiment, the pixel detector was designed to provide high-resolution tracking in an experimentally harsh region characterised by high track multiplicity and heavy irradiation.

The pixel detector in place at the time of writing was installed in 2017 to replace the original detector to maintain the overall performance of the CMS tracking system with rising instantaneous luminosities.

Referred to as the CMS Phase 1 upgrade pixel system, this detector consists of four barrel layers of radii between 3 and 16 cm and of three disks on each side. The detector is populated by pixel modules comprising in a rectangular pixelated silicon sensor bump bonded to two rows of eight read-out chips. These are in turn linked to a high-density interconnect hosting control and readout electronics.

The silicon $16.2 \times 64.8\,\mathrm{mm}^2$ sensors are of type n^+-in-n and segmented into $100 \times 150\,\mu\mathrm{m}^2$ pixels. When a charged particle passes through the silicon, the signal generated in the silicon is collected by the read-out chip and, if superior to a programmable threshold set to about $1800\mathrm{e}^-$ ($2600\mathrm{e}^-$ in the innermost barrel layer), is stored in an internal buffer. If the event is to be read out, this hit is digitised by the read-out chip and sent to a control circuit aggregating data from multiple read-out chips and providing the interface with the data acquisition system.

The detector has shown a hit efficiency of the order of or above 99% for instantaneous luminosities up to $1.8 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, with the exception of the innermost layer dropping to 97% at high luminosities. The hit resolution is of 10 µm and 30 µm in the transverse and longitudinal directions respectively [43].

2.2.2.2. The silicon strip tracker

Moving away from the interaction point, the track multiplicity decreases and a precise z coordinate measurement becomes less important. Pixelated detectors are no longer required and a strip detector provides the remaining inner tracking information.

The Silicon Strip Tracker (SST) is comprised of 15148 modules, for a total active area of $198\,\mathrm{m}^2$ split into 9.3 million channels. The 5 m long and 2.5 m diameter detector, shown on figure 2.7, is split into four different regions. The barrel of the detector has ten layers, split into four layers in a short inner barrel (TIB), and six layers in an outer barrel (TOB). The remaining space in the barrel is filled by three disks on each side, forming the tracker inner disks (TID). On each side of the barrel, a tracker endcap (TEC) formed of nine disks completes the coverage of the detector.

The sensitive material is p-on-n single sided silicon micro-strip sensors, of pitch between 85 µm and 205 µm. Depending on the position in the detector, multiple sensor geometries co-exist. In the inner barrel and the four most central rings of the disks, single 320 µm thick sensors from 6 inches wafers are used. In the outer part of the detector, strips span the length of two 500 µm thick sensors. Four layers of the barrel and three rings of the endcaps (shown in blue on figure 2.7) are populated by stereo modules. In these modules, a second module tilted by a 100 mrad angle is added. The difference in strips fired in the two modules provides an improved z-direction measurement in the barrel and R measurement in the endcaps.

The signal produced in the silicon is processed by the analogue pipeline voltage (APV25) chip [44]. Each silicon strip is covered by an electrically isolated aluminium strip, forming a capacitor. They are in turn wire-bonded either to a second sensor in the outer regions of the tracker, or to a pitch adapter itself connected to the APV25 chips. As shown on figure 2.8, the first stage of this ASIC is to transform the charge produced into a potential difference, through its pre-amplifier. Since the chip was designed to work with both sensor polarities, a programmable inverter will then make sure the signal is positive before entering the shaper. The output of this 50 ns CR-RC circuit is then sampled at the 40 MHz LHC frequency and stored in the central element of the chip, a 128 strip wide and 192 element deep analogue pipeline, hence the chip name.

Depending on the expected particle rate, three different read-out modes implemented in the Analog Pulse Shape Processor (APSP) are possible. Two pointers cycle through

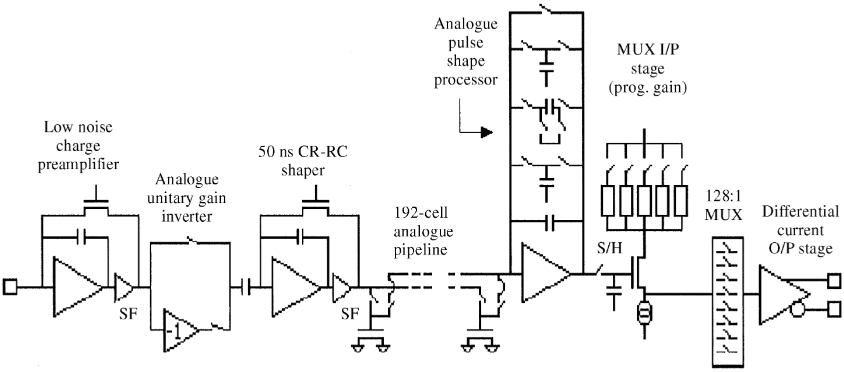


Figure 2.8. | Schematics of the APV25 read-out chip. The signal induced in the strips is amplified and shaped before being stored in an analogue pipeline. Events to be read out are then further transformed and serialised before being sent optically to the data acquisition system [44].

the chip buffer at the same speed: a write pointer storing the current charge followed at a programmable latency by the read pointer. In the *peak* read-out mode, upon receiving a Level-1 trigger accept signal, a single pipeline voltage will be read out. The delay between the read and write pointer is tuned to read out the peak of the CR-RC signal. While this simple read-out scheme provides the best charge measurement when particle rates are low, at higher intensities contamination from previous bunch crossings could have a significant impact. In the *deconvolution* mode, three consecutive voltages are read out corresponding to the required bunch crossing, 25 ns and 50 ns earlier. These three voltages are sent through a finite impulse response network to extract the charge measurement from the expected out-of-time contributions. A third, *multi* mode is also implemented to provide three consecutive *peak* mode measurements for calibration purposes [45].

The resulting analogue strip-by-strip charge measurement to be read out will then be multiplexed, amplified by a programmable gain and sent optically to the data acquisition system alongside some digital information (header, address, error bit, synchronisation bit).

The data coming from the strip tracker is read out by Front-End Drivers (FED) [46]. These VME64 modules receive and process data from up to 96 fibres, each corresponding to two APV25 chips. The analogue data is firstly digitised by a 10 bit ADC, the pedestal corrections are applied and the common mode is removed on an event-by-event basis. A clusterizing algorithm is then run on an FPGA and in normal data-taking conditions, the FED will send a list of clusters containing the fired strip addresses and corresponding charges with an eight bit precision to the central data acquisition system.

To ensure a uniform charge measurement, two corrections are to be applied. Since the charge collected by the detector is transferred analogically through optic fibres to the counting room where it is digitised, any change in fibre transparency will impact the recorded charge. To compensate this effect, light pulses of pre-determined amplitudes are sent through the fibres and the measured output is used to provide the first calibration constant.

Additional detector effects can modify the charge deposited in the silicon and reconstructed by the APV25s. A second calibration constant compensates these effects and is determined by collecting the charge deposited by minimally ionising particles and dividing it by path length, for every APV25 of the detector. The resulting information is used to fit a Landau distribution, from which the most probable value is extracted. Comparing the output of this value to expectations yields the second calibration constant, also named *particle gain*. The periodic production of these particle gains was the responsibility of the author of this document during the 2015 to 2018 data taking period.

After ten years of data taking, the CMS strip tracker performance remains outstanding despite first signs of detector ageing and luminosity induced inefficiencies.

The signal to noise ratio is decreasing with total fluence as expected. A current most probable value around 16.5 for thin sensors and 22 for thick sensors (see figure 2.9), alongside a slow and understood decreasing trend, indicate the S/N should be sufficient for quality data-taking for the detector's foreseen lifetime. These high signal-to-noise ratios allow for an improved spatial resolution with respect to a binary read-out due to charge sharing between neighbouring strips [47].

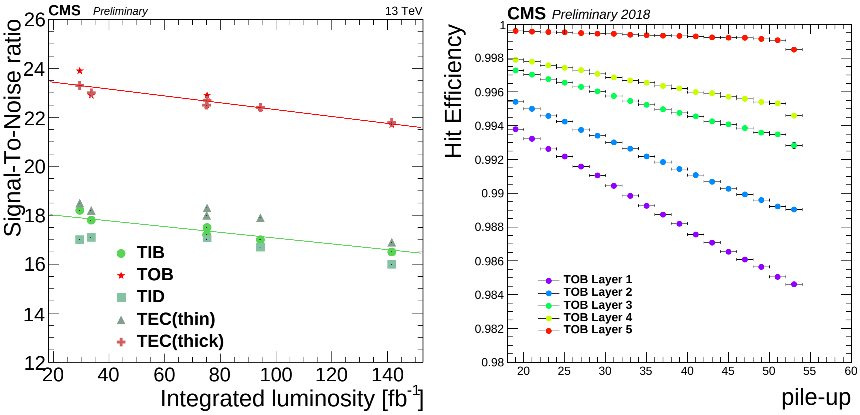


Figure 2.9. | The evolution of the run-2 signal-to-noise ratio in the different detector partitions as a function of integrated luminosity is shown on the left. The figures on the right show the dynamic inefficiencies as a function of the average collision multiplicity, as seen during the 2018 data-taking period [48].

The single hit efficiency remains above 99% for the TIB, TID and TEC and above 98% in the TOB [48]. As shown on the right-hand side of figure 2.9, a significant dynamic inefficiency appears in some layers of the TOB at high luminosity (or pile-up) levels. This quasi-linear effect was understood as originating from Highly Ionising Particles (HIP). Occasional nuclear collisions of hadrons in the silicon sensors produce secondary particles depositing a high amount of energy that can correspond to hundreds of minimum ionising particles. In turn, this energy deposit will saturate the read-out chip, and depress its baseline, rendering the chip blind. The chip gradually returns to its normal baseline, but remains inefficient in the meanwhile [49]. This effect being directly proportional to the volume of silicon read-out by an APV25 but also to the particle flux (thus to R^{-2}), the most sensitive partition is the first layer of the TOB, as observed.

A crucial aspect of radiation hardness is the ability to fully deplete the silicon. The tracker not yet reaching its type inversion point, the full depletion voltage is currently decreasing and is not an experimental issue. Besides changing the depletion voltage, radiation damage also increases leakage current. Some stereo modules with damaged cooling contacts or on a closed cooling loop cannot be cooled effectively and have suffered from a positive temperature-current feedback loop known as thermal runaway. In 2017, this effect lead to some modules coming close to the current limit of the power supplies. While reducing the operational temperature from $-15\,^{\circ}\mathrm{C}$ to $-20\,^{\circ}\mathrm{C}$ allowed to power these modules through run-2, it is expected that, even if the temperature is further reduced to $-25\,^{\circ}\mathrm{C}$, more than 170 modules could suffer thermal runaway at an integrated luminosity of $400\,\mathrm{fb}^{-1}$.

The ten years that the tracker detector has been operating confirm that the detector performance is and should continue to be outstanding for the years to come. The expected instantaneous luminosity and radiation impact on the detector also indicate that greatly increasing the luminosity or running way above a $500\,\mathrm{fb}^{-1}$ integrated luminosity would have a dramatic impact on performance.

2.2.3. The electromagnetic calorimeters

The electromagnetic calorimeter (ECAL), shown on figure 2.10 is composed of a barrel (EB) covering the pseudorapidity region below $|\eta| = 1.479$ and two endcaps (EE) reaching up to $|\eta| = 3.0$.

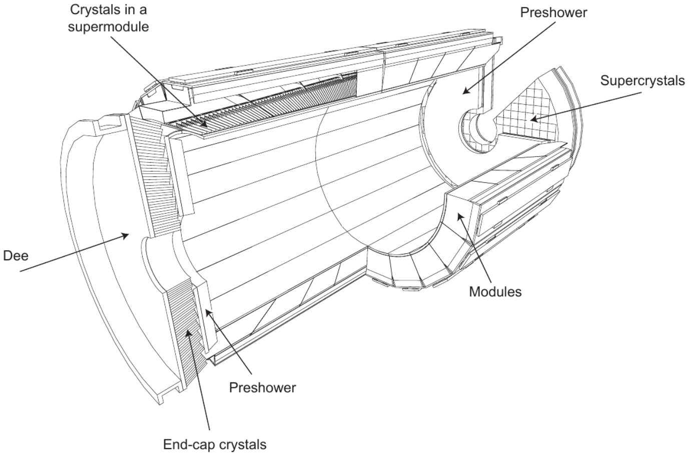


Figure 2.10. Sketch of the electromagnetic calorimeter [50].

Lead-tungstate ($PbWO_4$) was chosen as the sensitive material for this detector due to its high density, short radiation length and small Molière radius, making it ideal for a compact and high granularity detector. Furthermore, it has a fast scintillation decay time, allowing 80% of the light to be produced within 25 ns. The light output is relatively low with 4.5 photo-electrons per MeV of deposited energy being collected by the photo-detectors. The temperature is maintained at precisely 18 °C with a light output decreasing by 2.1% every degree.

In the barrel region, the 61200 crystals have a rectangular 26×26 mm^2 section and a length of 230 mm. The crystals are placed facing the interaction point at a slight angle in order to avoid loss in coverage due to the interface between two crystals. Their granularity corresponds approximately to 0.0174×0.0174 in η,ϕ and their depth to 25.8 radiation lengths (X_0). The photo-detectors used in the EB are pairs of avalanche photo-diodes. The endcaps are filled by 7324 crystals of 28.62×28.62 mm^2 section and 220 mm depth (24.7 X_0) arranged in a x,y grid and read-out by Vacuum Photo-Triodes.

The signal from five photo-detectors are pre-amplified, shaped and amplified with three different gains (1, 6 and 12) in a very front-end card (VFE). These three signals are then digitised in parallel by 12-bit ADCs and the highest non-saturated value with the corresponding ADC gain is sent to the front-end card. There, the signals from five VFE cards are stored in a 256-word deep buffer awaiting a Level-1 accept signal. If such a signal is received, the signal from ten consecutive bunch crossings are sent to the off-detector read-out system where a multiple pulse fit will extract the signal contribution from out-of-time pile-up. In parallel, the energy sum from the 25 crystals connected alongside a "fine-grain" bit describing the lateral profile of the shower is sent to the Level-1 trigger system every bunch crossing. The on-detector electronics, optical transmitters and other services are located on motherboards. In the EB, these boards also provide bias stabilisation to the sensitive avalanche photodiodes.

To improve the angular resolution in the EE and help resolve the signal from high energy neutral pions decaying to two close photons, the endcap is complemented by a preshower detector (ES). The preshower is a small sampling calorimeter with two absorber layers (2 and 1 X_0) coupled to a silicon strip sensor each.

While the intrinsic luminosity of the crystals is radiation hard, their transparency can be affected and a continuous calibration system compensates the induced signal loss. This is achieved through the injection of two laser pulses during the LHC abort gap. The detector response to these pulses allows the crystal response to scintillation light to be measured and corrections to be applied promptly. The crystal luminosity, shown on figure 2.11, is greatly affected in the forward region while the barrel shows a better resilience.

2.2.4. The hadronic calorimeters

The hadronic calorimeter (HCAL) is a sampling calorimeter aimed at measuring the energy of high energy hadrons. It has four different components, as shown on figure 2.12. A barrel (HB) is located in the region between the EB and the solenoid, covering $|\eta| < 1.3$. The limitation in space and therefore of absorber material lead to the addition of a second small calorimeter outside the solenoid (HO) to catch any tail in energy deposition. The hadron calorimeter endcap (HE) covers the $1.3 < |\eta| < 3.0$ and a forward endcap (HF), placed 11.2 m from the interaction point extends the overall coverage up to $|\eta| = 5.2$.

In the HB, a 40 mm steel front plate, eight 50.5 mm and six 56.5 mm brass layers and

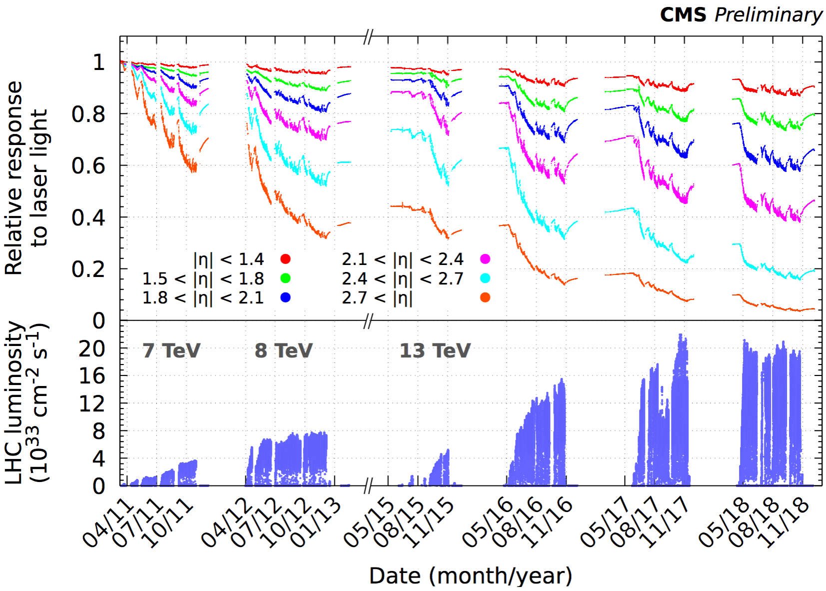


Figure 2.11. | Evolution of the ECAL crystal transparency as a function of time [51].

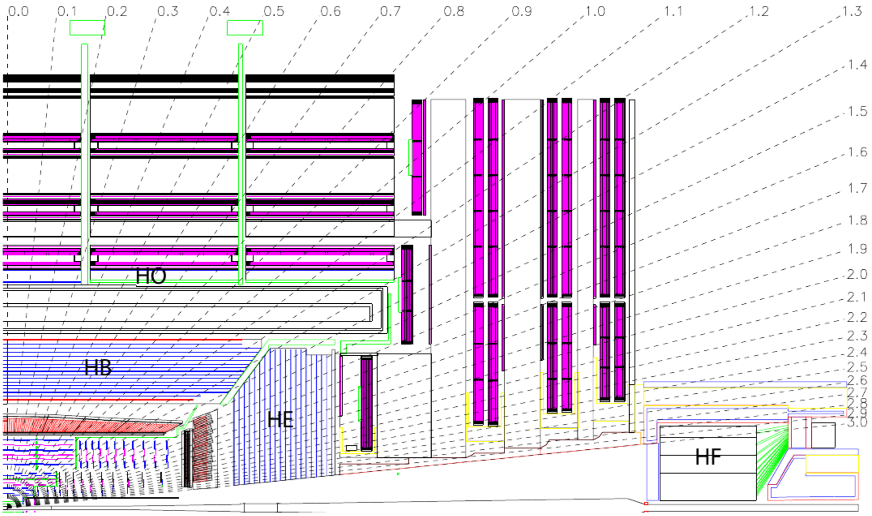


Figure 2.12. | Layer of one quarter of the HCAL system in the $r-z$ plane [50].

a 75 mm steel back plate are used as absorbers. Given the cylindrical geometry of the detector, the constant absorber radius corresponds to an absorber thickness 5.82 interaction lengths (λ_I) at $\eta = 0$ and 10.6 λ_I at $|\eta| = 1.3$. The presence of the ECAL adds another $\sim 1.1\ \lambda_I$ of material. The hadronic showers are sampled by 70000 plastic scintillator tiles and the light produced is collected by wavelength shifting fibres connected to hybrid photodiodes (HPD). In these detectors, the electrons converted from the scintillation light are accelerated by an 8000 V difference of potential onto a pixelated silicon photodiode, thus providing an internal gain of approximately 2000. The scintillators have a $(\Delta\eta,\Delta\phi)$ granularity of (0.087,0.087) and mostly no longitudinal segmentation.

The HE uses the same scintillating and absorber material. The angular granularity is of $(\Delta\eta,\Delta\phi)$ = (0.087,0.087) for $|\eta| < 1.6$ and approximately (0.17,0.17) for higher pseudorapidities. Unlike most of the barrel, the HE has a longitudinal segmentation. The initial detector had two or three longitudinal read-outs upgraded to up to seven in 2017, allowing for more precise calibration to mitigate radiation effects in this high fluence region. While initially read out by the same HPD as for the barrel, these were eventually replaced by Silicon PhotoMultiplier (SiPM) detectors providing better radiation tolerance.

The HO is comprised of a layer of scintillator tiles which match the HB granularity using the solenoid volume and the first iron layer of the magnetic return yoke as extra absorption material. Given the angular dependence of the total material depth, a second HO layer is added in the central region of the barrel. The total depth of the calorimeter is thus never less than 11.8 λ_I except in the barrel-endcap transition region.

In the very forward region covered by the HF, huge particle fluxes render the environment very hostile for conventional calorimetry. Traditional scintillators not being adequate, quartz fibre detectors were chosen for their outstanding radiation tolerance. These fibres are placed in a cylindrical grooved 165 cm long steel absorber with a radius of 135 cm. Half the optical fibre span the whole length of the absorber while the other half do not cover the first 22 cm. The separate read-out of both sets of fibres allows to discriminate between showers originating from an electron or a photon. The Cherenkov light mostly produced by the electromagnetic component of these showers emitted in the fibres is read out by conventional photo-multiplier tubes with an angular segmentation of 0.175×0.175 in $(\Delta\eta,\Delta\phi)$.

Throughout the whole detector, ultra-violet laser flashes are fed into the HCAL sensitive material to emulate the light produced by a charged particle and provide a crucial calibration of possible transparency degradation caused by radiation damage. Radioactive sources can also be inserted into steel tubes inside the detector to further calibrate the system.

The analogue signal produced by the photo-detectors are read out by front-end cards that integrate the signal and digitise them into a 7-bit non-linear scale. This signal, alongside some monitoring information is then sent to the service cavern and stored in a pipeline. In parallel, the detector information is linearized, the longitudinal contributions are summed and transformed into transverse energy through a look-up table. This trigger primitive is then compressed and sent to the Level-1 trigger system.

2.2.5. The muon chambers

Three dedicated gas detector sub-systems are located outside the solenoid to accurately identify and characterise muons, the drift tubes (DT), cathode strip chambers (CSC) and resistive plate chambers (RPC).

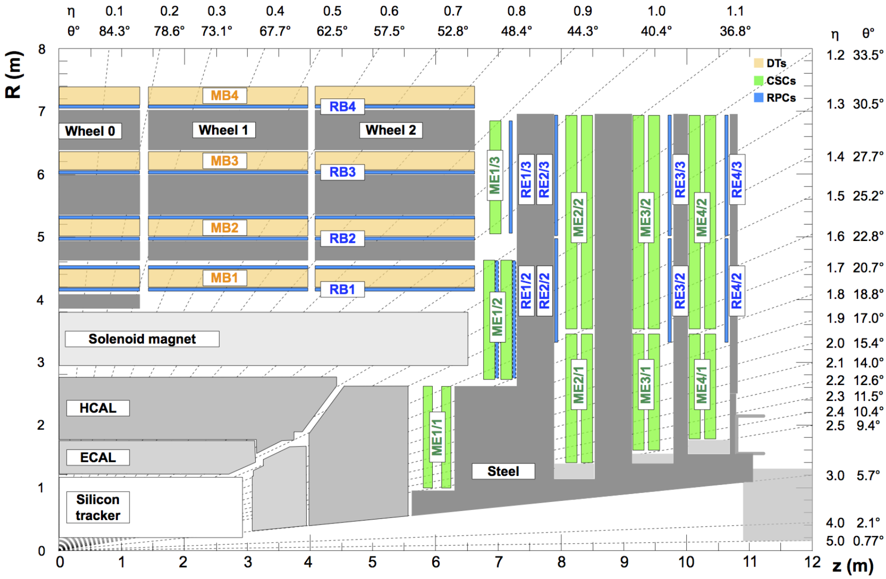


Figure 2.13. Layout of one quarter of the muon systems in the $r-z$ plane [52].

In the central region covered by the barrel where the magnetic field is uniform along the beam direction, the muon rate is low and neutron-induced background is reduced, standard rectangular drift cells are used to measure the muon momentum. A 50 µm gold-plated stainless steel wire is placed in a 42 mm×13 mm×2-3 m rectangular CO_2-Ar filled chamber under a shaped electric field. These chambers are arranged into superlayers (SL), and a drift-tube station is composed of two SL aligned with the beam axis, and one SL in the perpendicular direction. The outermost stations of the drift-tube system only have the two SL aligned with the beam axis. Four barrel layers cover the pseudorapidity region $|\eta| < 1.2$, as shown of figure 2.13.

The signal produced by the drift tube is processed by an on-detector front-end chip that pre-amplifies and shapes the signal, compares it to a threshold and sends it to read-out boards in a minicrate located on the SL. The read-out electronics digitise and precisely time incoming signals, aggregate signals from multiple detectors and send them to electronic boards located on balconies next to the detector to be optically sent to the service cavern. In parallel, trigger boards in the minicrates will create individual SL track segments, correlate multiple segments in the same chamber into tracks and the best two tracks will be sent to the trigger system.

The endcap region is populated by 4 wheels of cathode strip chamber detectors covering $0.9 < |\eta| < 2.4$. The CSC are trapezoidal multiwire proportional chambers consisting of 6 tangential anode wire planes providing coarse R coordinate measurement interleaved with 6 radial cathode strip planes allowing for a precise ϕ measurement. The chambers are filled with a Ar-CO_2-CF_4 mixture.

The charge collected by the wires is processed by 12 to 42 anode front-end boards that amplify the signal before feeding it to a constant fraction discriminator. The resulting hits are sent to an anode local charge board aggregating the data and storing it in a FIFO awaiting a Level-1 accept signal. In parallel, an FPGA compares the hits from all anode planes with pre-defined patterns, and the matching patterns are transmitted to the trigger system.

The signal induced in cathode strips is processed by up to five cathode front-end boards where it is amplified, shaped and sent down two separate paths. In the first path, the signal peaking in about 150 ns with no tail is sampled every 50 ns and stored in an analogue memory. Upon receiving a Level-1 accept signal, 8 to 16 consecutive samples are read, digitised and sent to the read-out system. The second path taken by the signal will send it to a comparator network identifying the muon position in half-strip units by locating the maximum charge deposited within strip triplets. This information is sent to a trigger motherboard that, like the anode front-end board, searches for track patterns to be used as trigger primitives.

In the ME1/1 station, the cathod strip front-end boards were replaced in 2014 by digital cathod front-end boards [53] where the detector signal is amplified and digitised into a programmable gate array. This upgrade allows the detector to handle much higher particle rates.

The last detectors are the resistive plate chambers complementing the DT and CSC in the $|\eta| < 1.9$ region. The primary goal of these chambers is to provide precise timing information for the muon trigger.

The basic RPC modules consist of two 2 mm gaps filled with a $C_2H_2F_4$, iC_4H_{10} and SF_6 mixture operated in avalanche mode with common read-out strips. The strips are aligned with the beam axis in the barrel and the CSC strips in the endcaps.

The signal induced in the strips is read out by two to four front-end boards that provide amplification and zero-crossing discrimination offering a time-walk below the nanosecond. This discriminator produces a 100 ns long pulse masking any possible after-pulse. The front-end boards are in turn connected to link-boards that transmit the data optically to the service cavern.

A summary of the CMS muon subsystems is shown in table 2.1

2.2.6. Trigger and data acquisition systems

Given the overwhelming collision rate and that the vast majority of these events only involve well-understood QCD processes, a rate reduction from 40 MHz down to about 1 kHz is required. This is achieved through a two-stage trigger system. The first stage, called Level-1, hardware based, reduces the rate to about 100 kHz allowing a second stage, the high-level trigger (HLT), to use commercial CPUs to further reduce the data rate to the desired levels.

2.2.6.1. Level-1 trigger

As described in the previous sections, only summary information from a range of subdetectors (mainly the calorimeters and muon systems), called trigger primitives, are provided to the Level-1 trigger system, while the full data is stored in local buffers. The trigger then has to provide an accept signal in a fixed 3.8 μs latency for events to

Muon subsystem	DT	CSC	RPC
$\|\eta\|$ coverage	0.0-1.2	0.9-2.4	0.0-1.9
Number of stations	4	4	4
Number of chambers	250	540	Barrel : 480 Endcap : 576
Number of layers/chamber	$R-\phi:8;\, z:4$	6	2 in RB1 and RB2 1 elsewhere
Number of readout channels	172 000	Strips: 266112 Anode ch.: 210 816	Barrel: 68 136 Endcap: 55296
Percentage of active channels	98.4%	99.0%	98.3%

Table 2.1. | Properties and parameters of the CMS muon subsystems during the 2016 data collection period. [52]

be read-out. To achieve this challenging task, the first level trigger has the following architecture, as illustrated on figure 2.14 and described in a technical design report [54].

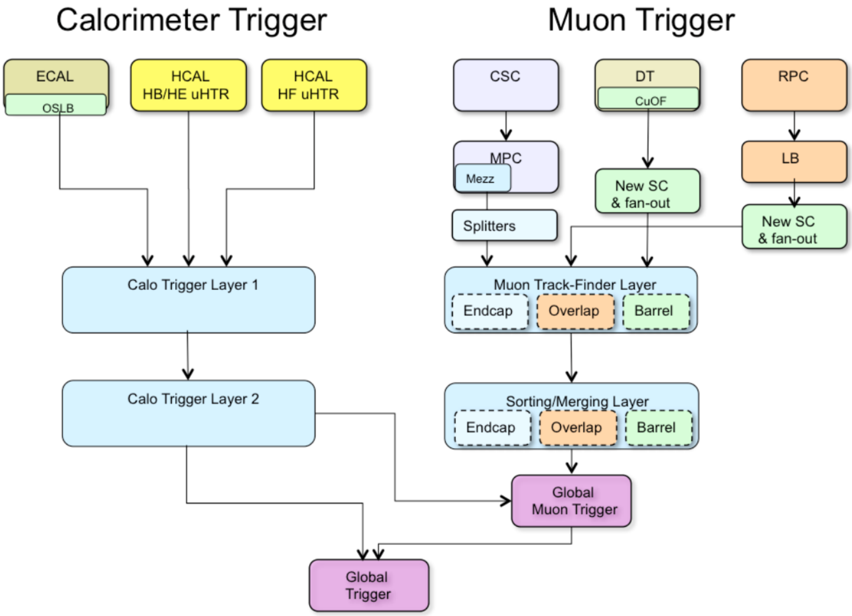


Figure 2.14. | Schematic representation of the Level-1 trigger conceptual organisation. Trigger primitives from the detector, on top, are gradually processed by different algorithmic layers to provide a final trigger decision [54].

Energy sums from adjacent detectors forming trigger towers in the ECAL or the HCAL are sent to a first trigger layer performing energy calibration. A time-multiplexed system then allows all the calibrated values from a given bunch crossing to be sent to a single layer-2 card performing object reconstruction.

Jet candidates are built as 9 by 9 trigger towers around a local energy maximum (the seed). Pile-up contributions are estimated from the four adjacent 3 by 9 rectangular regions. The three lowest transverse energy contributions from the four are removed from the jet candidate. Then, a final jet energy correction is applied based on its transverse energy and pseudorapidity.

Level-1 e/γ candidates originate from a local energy maximum in the ECAL that then dynamically add neighbouring trigger towers up to a maximum two by five size in η,ϕ. After a shape selection is applied, the corresponding hadronic energy deposit is computed and a cut on the energy ratio is applied. Finally, an isolation criteria is applied. The isolation energy in the 6×9 η/ϕ region around the candidate is summed, excluding the e/γ itself, and a dynamic cut on this value depending on a pile-up estimator is applied. A similar method is used in the hadronic calorimeter to create τ candidates with different cuts applied and the possibility of merging two neighbouring clusters.

The last calorimeter object reconstructed is the missing transverse energy defined as the negative sum of all energy deposits above threshold up to $|\eta| = 5$. The threshold is both η and pile-up dependent.

The trigger primitives from the muon detectors are sent to three different systems depending on their geometrical position. The $|\eta| < 0.83$ is covered by the barrel muon track finder (BMTF), the $0.83 < |\eta| < 1.23$ by the overlap muon track finder(OMTF) and the $1.24 < |\eta| < 2.4$ by the endcap muon track finder, (EMTF).

In the OMTF and EMTF, muon candidates are produced by comparing the trigger primitives gathered with patterns derived offline and stored in look-up tables. In the barrel region covered by the BMTF, a road-search algorithm is used to produce the candidates. Muons from the three systems are sent to a global muon trigger to remove duplicates and apply quality cuts on the tracks produced.

All trigger objects are then sent to a global trigger that has a programmable list of rules to apply before generating a Level-1 accept signal. A wide variety of rules coexist, ranging from simple objects (muon candidates above 22 GeV for example), to more advanced combinations such as asking for three muons with transverse momentum above 5, 3.5 and 2.5 GeV including two opposite sign muons with p_{T} above 5 and 2.5 GeV and with an invariant mass between 5 and 17 GeV. Some rules are dedicated to calibration studies and monitoring, such as first collision in a bunch train or purely random triggers.

Since the total trigger accept rate is limited, a random selection will be applied for events passing certain rules. This is known as prescaling, and the proportion of events passing a given rule that are read out can range from 1 (all events selected) for strict rules, to one every hundred thousandth for very lax rules and calibration streams.

2.2.6.2. Data acquisition and High-Level trigger

If an event is selected by the Level-1 trigger, the full detector is read out and data is sent to the data acquisition system (DAQ) to build events and apply a final selection. This architecture choice represents a major change with the traditional three levels system widely used at the time. When the experiment was designed, major and constant advances in the technological field and crucially in processing power and network speed lead to the decision of having a two-step decision system. This allows the High-Level trigger to access the full detector granularity and the whole detector coverage, which would not be possible in the second level of a three stage trigger. It also allows for an improved flexibility for new trigger algorithms.

A conceptual sketch of the DAQ system is shown on figure 2.15.

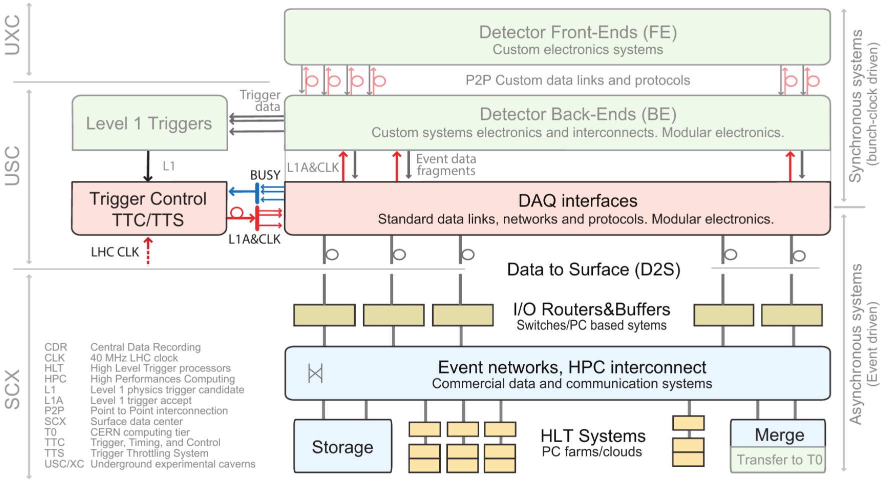


Figure 2.15. | Schematic representation of the data acquisition and high-level trigger systems. Data fragments from the detector systems (in grey) are sent to DAQ interfaces aggregating the data and transferring it through standard protocol to the surface data centre where full events are built, filtered and stored temporarily [55].

For accepted events, the Level-1 accept signal is sent to custom detector front-end boards that will propagate the signal back to the detector electronics keeping the collision information. This information is then read-out by front-end driver cards (FED) and sent synchronously via a custom protocol to Front-End Read Out Link cards (FEROL). These cards will assemble data from a small number of FEDs, build an event fragment and transmit the data downstream using standard TCP/IP protocol on 10 gigabit optical links. This step marks a double conceptual separation, from synchronous data-stream to an event-by-event basis on the one side, and from custom protocols and hardware to commercial data and communication systems on the other side.

The fragments are then sent to an event-builder interconnect network. Data from multiple FEROLs is sent to read-out units hosted in commercial computers that will create super-fragments. Then, these super-fragments are sent through an infiniband interconnect network to a farm of builder-unit computers. All super-fragments from a given event are collected by a single builder -unit (BU) that will generate the full event and store it on a 256 GB RAM disk.

Since the data rate is still too high to keep all these events, several filter units (FU) are attached to every builder unit to further reduce the rate down to about 1 kHz. Given the event rates involved and the number of filter units available to run this selection mechanism known as high-level trigger, the decision whether to keep or discard an event has to be taken on average within a few hundred micro-seconds. As this time constraint is too short for many reconstruction techniques, the HLT uses a trigger path method. A path consists of sequential reconstruction steps needed to provide the final decision, such as unpacking the data, constructing energy deposits or detector tracks or reconstructing physics objects. The system was designed to start with fast steps and to exit as soon as a decision would lead to discard an event. This allows to take only 50 ms for the majority of events, while having up to a few seconds to apply complicated algorithms that require tracking information.

Events that pass this final selection are stored in a local data-centre that is connected to the central CERN computing network that will take care of long-term storage or data reconstruction.

2.3. Data reconstruction and enhancement techniques

The system described above allows raw data to be permanently stored to disk or to tape. In order to be able to draw conclusions on the nature of high-energy physics from these raw events, multiple steps are required. High-level objects have to be reconstructed, the behaviour of the detector is to be simulated and dedicated physics analyses need to process this information to enhance the sensitivity to a phenomenon of interest. This section will give a brief overview of these different steps, with a focus on the particular objects and methods used in the analysis at hand.

2.3.1. Object reconstruction

Events produced by the detector are stored in a raw format containing collections of detector level objects. This raw format is impractical for most data analyses, therefore these collections will be converted into physics objects centrally by the collaboration. This process, know as reconstruction, can be done multiple times with different levels of precision and sets of calibration constants.

In this section we will describe object reconstruction using the particle-flow method [56]. This algorithm combines tracks and calorimeter deposits (described in section 2.3.1.1) throughout the detector into particle candidates that are then used for physics analysis. The similar methods used in the online HLT object reconstruction will not be discussed in this document.

2.3.1.1. Tracking and calorimetry

Signals produced in the pixel and strip detectors are transformed into hits through a local reconstruction process. These hits in local coordinate systems are then converted to global coordinates. This step requires the precise knowledge of the position and orientation of every detector module; information which is obtained through an alignment procedure [57].

Transforming these hit collections into tracks is a computationally hard task, especially under high particle multiplicities, that is performed by the Combinatorial Track Finder (CTF). An iterative tracking method is implemented to produce tracks starting by the easiest to identify and remove the corresponding hits to simplify the task of future iterations. Every iteration has the same four steps. A seed is generated, usually from triplets or quadruplets of hits in the inner pixel, then a Kalman filter extrapolates the tracks to the outer layers; the track parameters are fitted and a track is generated if it passes some quality criteria.

In the first tracking iterations, seeding is performed by forming hit pairs compatible with a track originating from the interaction point in adjacent layers and wheels of the pixel detector. Then, these hit pairs are compared and hit triplets or quadruplets are formed. In subsequent iteration steps, the impact parameter, transverse momentum and impact parameter requirements are modified and the strip tracker or even muon chambers can be used for seeding.

Track parameters and corresponding uncertainties are extracted from these seeds and used by a Kalman-Filter based track finding module. Each step of this module takes the track parameters from the previous step and extrapolates the position to the next detection layer. Compatible hits are then added to the trajectory and its parameters are updated. The first iteration of this algorithm assumes a uniform magnetic field and no energy loss (leading to a perfect helix trajectory), while subsequent iterations will take into account the material budget when propagating the track. The track finding algorithm will stop if too many hits are missing, the transverse track momentum goes under a preset threshold, discarding the candidate, or if it reaches the last tracker layer. This search is then performed in reverse, from the outermost layers back towards the interaction point to add possible missing hits. Since this method can produce multiple tracks from the hits generated by a single charged particle, a trajectory cleaner is applied to remove duplicates.

The track finding step produces a set of hits and an estimate of track parameters that are then fitted by a Kalman-Filter and smoothed, providing the final track candidate. Final quality criteria based on the track chi-squared, number of hits, track uncertainties and number of layers intercepted are then applied.

The impact of the different tracking iterations on simulated $t\bar{t}$ events are shown on figure 2.16. Data analysis using the tag-and-probe method on both Z and J/ψ to $\mu^{\pm}\ \mu^{\mp}$ yield an inner-tracker only efficiency of above 97% [58, 59].

A similar approach is applied to the muon chambers. In the DT, each cell will reconstruct the position transverse to its wire by measuring the electron drift time in the chamber. In the CSC, the hit position is reconstructed by combining strip and wire information. Since both DT and CSC consist of groups of stacked layers of detectors, quasi-straight segments can be reconstructed. These are then used as the seed for a Kalman Filter based algorithm that aggregates the information of the DT, CSC and RPC to produce standalone muon tracks.

In an inside-out fashion, tracks from the inner tracker passing transverse momentum thresholds are extrapolated to the muon chambers. If at least one muon-chamber segment is compatible with the extrapolated track, it will be reconstructed as a tracker muon track. The opposite outside-in approach is also applied by extrapolating standalone muon tracks and inner tracker tracks to a common plane. If matching conditions are met, a common fit is performed on all hits along the trajectory resulting in a global

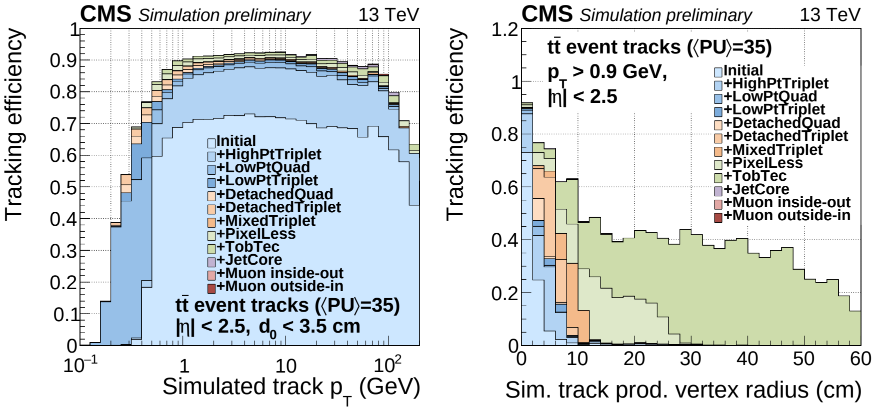


Figure 2.16. The stacked tracking efficiency from different reconstruction steps in simulated $t\bar{t}$ events is shown as a function of transverse track momentum (left) and of production vertex radius (right) [60].

muon track.

Energy deposits in the calorimeters are reconstructed as clusters independently for every sub-detector. The algorithm starts by finding energy deposits that are above a given threshold and its neighbours. These seeds are then grown into topological clusters by adding any neighbouring cell above a set energy threshold (usually twice the noise). Multiple seeds are usually included into a single topological cluster.

Clusters are then generated from these topological clusters through a Gaussian mixture model. In this model, each seed i yields a cluster that has a Gaussian shape of fixed width depending on the sub-detector, of amplitude A_i and of position $\vec{\mu}_i$. The initial parameters are taken to be the energy and position of the seed. Then, for every cell j of the topological cluster, the fraction f_{ij} defined as the expected relative energy contribution from the cluster i is computed. Finally, the energy deposits in the topological clusters are fitted yielding the amplitude A_i and position μ_i of all clusters.

Using the combined information from the ECAL and the inner tracker allows for more precise electron tracking. Energetic electrons in the presence of the tracker material have a high probability of radiating a substantial amount of energy, rendering the Kalman-Filter based algorithm less effective. This property is used to disentangle electrons and charged hadrons in the tracker. Radiating electrons will tend to have less hits collected on its trajectory, or, if soft photons are emitted, have a higher χ^2 than expected. These selected tracks are re-processed using a Gaussian-Sum Filter (GSF), similar to a Kalman-Filter with the exception that it allows for multiple energy losses along the trajectory. The track has to pass a final discriminant based on the track quality, distance to ECAL clusters as well as a comparison with the corresponding Kalman-Filter track in order to produce a GSF track.

The particle flow (PF) algorithm will then link together objects of different type that can share a common origin into PF blocks. An inner track is linked to a calorimeter cluster if the extrapolated position of the track in the ECAL or HCAL is compatible with the centre of the cluster. Furthermore, to collect possible radiation energy emitted by electrons in the tracker, ECAL clusters tangent to a GSF track are included. Calorimeter clusters from different detectors can be linked together, and tracks originating from a possible secondary vertex also.

For every PF block, a sequence of particle flow identification is run, starting with muons, then electrons and isolated photons. Finally, the remaining photon alongside charged and neutral hadrons are reconstructed. At each step, the PF elements used to reconstruct a particle are removed from their PF block.

2.3.1.2. Muon reconstruction

The particle flow algorithm starts the muon identification procedure by looking at isolated muons. The transverse energy and momentum of any track or calorimeter deposit within a cone of radius ΔR with respect to a global or tracker muon is summed and if the total is inferior to 10% of the candidate p_{T}, it is kept and the corresponding objects are removed from the PF block. If this isolation criteria is not met, the candidate has to fulfil additional conditions to minimise misidentification of punch-through hadron. This is done by requiring additional hits in the inner tracker as well as the muon stations and calorimeter deposits compatible with expectations from a muon. Finally, if this last selection has failed, a muon can still be reconstructed if either the track reconstructed by the muon chambers or the inner tracker is of high-quality. In the case of a good inner track, an additional constraint on the shape of the calorimeter clusters is applied to differentiate them from hadrons.

Four application-specific fitters can be applied to determine the muon momentum. For muons with p_{T} below 200 GeV, the *inner tracker fit* method uses only the information from the silicon tracker to compute the momentum. For higher energies, the *Tracker-Plus-First-Muon-Station fit* combines this information with the hits for the first muon station, providing the best information from the muon chambers. In case of showers within a muon station, many hits will be produced in that detector and the so-called *picky fit* method will only consider the hit most compatible with the overall track. Finally, the *dynamic-truncation fit* allows for significant muon energy losses by performing an inside-out refit of the trajectory while allowing hits from the nearest segment if compatible.

To provide the best and most commonly used momentum measurement, the Tune-P algorithm selects the best momentum measurement from the four available based on goodness of fit and uncertainty information.

The muons produced by the particle flow algorithm are then subjected to a series of additional selections. This allows to define multiple categories of objects with varying efficiency and purity that can be used for physics analyses. In particular, two standard (loose & tight) working-points are defined.

Loose muon identification is defined as to encompass any prompt muon originating from the primary vertex, as well as from the decay of hadrons in the detector. Selecting a particle-flow reconstructed muon and requiring that it is also reconstructed as a tracker or global muon is sufficient to allow for a high-efficiency definition with reasonably good purity.

Tight muons are aimed at selecting muons originating from the primary vertex only and suppressing any punch-through hadrons reaching the muon chambers. On top of the basic loose muon requirements, at least six hits in the inner tracker are required, including at least one from the pixel detector. The muon has to be reconstructed as both a tracker and global muon, pass χ^2, transverse and longitudinal impact parameter cuts and have a matching segment in at least two muon stations.

Finally, isolation criteria can be derived by summing the transverse momentum of all reconstructed particle flow objects within a $\Delta R = 0.4$ radius. If this value, after correcting for pile-up contributions, is smaller than pre-set values, the muon passes the given isolation criteria. Two PF based isolation working points are commonly used, loose or tight, that are defined to reach a 95% and 98% efficiency respectively.

Applying both tight identification and isolation criteria lead to an overall efficiency of about 96%, with a p_{T} and η dependence shown on figure 2.17.

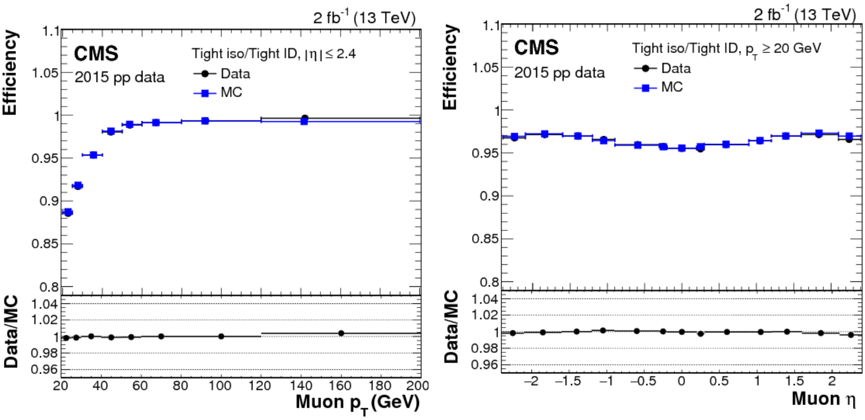


Figure 2.17. | Efficiency of the tight PF isolation on top of the tight identification working point as a function of muon p_{T} and η computed using the tag and probe method [52].

2.3.1.3. Electron and Photon reconstruction

The amount of material in the CMS tracker leads to important electron bremsstrahlung as well as photon conversion to electron-positron pairs. It is therefore natural that both objects are considered together when reconstructing events.

Particle-flow electrons are reconstructed from GSF tracks pointing towards an ECAL cluster that is not linked with more than two tracks. Additional ECAL clusters aligned with tangents of the tracks are also considered to have an estimate of bremsstrahlung and a better energy resolution. Photon candidates are built from energetic ECAL clusters that are not linked with tracks.

Both electron and photon candidates are required to have relatively small HCAL deposits around the electromagnetic cluster. If this condition is met, the object ECAL energy will be corrected through an analytical function depending on the pseudorapidity of the cluster and the measured energy.

Finally, electrons and photon candidates are submitted to selection criteria. Up to fourteen electron related variables, including energies, electromagnetic energy shape and track quality, are combined in boosted decision trees (see in section 2.3.2.2) providing the final particle-flow selection. Photon candidates are selected if they are isolated from other tracks and if energy deposits in the calorimeters are compatible with the photon hypotheses.

Non isolated photons cannot be reconstructed using this method which requires an absence of track around the ECAL cluster. These objects are thus recovered later when hadrons are built.

While many identification methods co-exist for electrons, the two most used are either the cut-based method or the multivariate analysis method (MVA). The identification methods are based on quantities like energy deposition shape, momentum to energy comparison, isolation variables and track to energy deposit compatibility, either through rectangular cuts optimised on $t\bar{t}$ event simulation, or combined in boosted decision trees trained on Drell-Yan+jets simulations in the multivariate method. An overview of the performance of both methods is shown on figure 2.18. These methods were further validated through the tag-and-probe method using Z decay electrons.

Photons also have cut-based and mva-based identifications. Since no track information is available, the identification is based on the shape of the energy deposit, the ratio between electromagnetic and hadronic energy deposit and different isolation parameters. The overall performance is shown on figure 2.18, as extracted from simulations of γ + jets events and validated on data using the tag-and-probe method on electrons from the Z decay that were reconstructed as photons.

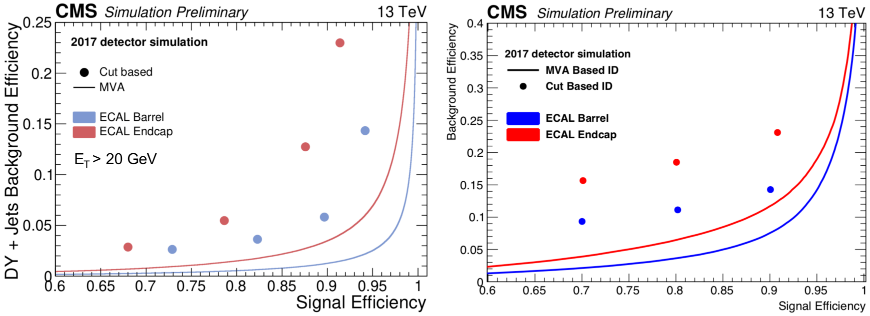


Figure 2.18. | Comparative performance of the cut based and mva based identification criteria compared to DY+jets backgrounds. The left plot describes the electron efficiency [61] and the right plot shows the photon performance [62].

2.3.1.4. Jet reconstruction

Once the objects related to leptons and isolated photons are removed from a particle-flow block, hadrons are extracted from the remaining items.

Within the tracker acceptance, any HCAL clusters not linked to any track will be reconstructed as a neutral hadron, while ECAL energy deposits will be assigned to photons. The remaining energy deposits are then identified as charged hadrons. After calibrating the energy of the candidate, if an abnormal amount of energy has been deposited in the electromagnetic calorimeter, it will be assigned to a photon.

To best capture the behaviour of particles generating these showers, an anti-k_T algorithm [63] is then applied on PF objects, usually using a radius parameter R = 0.4, to produce jets. Given the high number of simultaneous collisions, pile-up mitigation techniques are primordial to remove contributions from the underlying event when the jet is built. One of the possible methods, used in this work, is the Pile-Up Per Particle Identification (PUPPI) technique [64].

Vertices are created from the intersection of charged particle tracks in the detector, and the leading vertex is defined as the vertex with the highest sum of squared transverse momentum (p_T^2) of physics objects attached to it.

The PUPPI algorithm will then assign a so-called shape value to every particle-flow particle. This shape is defined as the logarithm of the sum of the transverse momentum divided by radius of any other particle within a fixed radius cone, excluding co-linear particles. When tracking information is available, only charged particles originating from the primary vertex are considered in the sum.

Since the p_T spectrum of particles within a jet is different from the rapidly falling p_T of the underlying event, the shape of these objects is very different. A weight is then assigned to every particle based on the distance of its shape with respect to the expected shape of pile-up. The weight is defined to be of 1 for jet-like particles, and 0 if they are similar to particles from the underlying event.

The four-momentum of every particle flow object is then re-scaled by this weight before applying the anti-k_T clusterization.

This method has shown to provide a significant increase in jet energy and pseudorapidity resolution. Compared to the common Charge Hadron Subtraction (CHS) method [56], PUPPI has shown to be more effective with increasing pile-up.

Finally, given nonlinearities in the energy measurements of calorimeters, and a non-uniform detector response as a function of pseudorapidity, a jet energy correction depending on p_T and η is applied to obtain the correct jet energy scale.

In a collider experiment, a jet can be produced from multiple objects. It can be crucial for an analysis, and in particular Higgs boson pair production studies, to be able to specifically identify jets originating from b-quarks (b-jets) or hadronically decaying τ leptons.

In order to identify b-jets, the properties of hadrons containing b-quarks is exploited. With a typical lifetime of about 1.5 ps, they can travel up to a centimetre in the detector before decaying, leading to displaced vertices with respect to the primary vertex. Precisely reconstructing this secondary vertex is thus key to identifying heavy jets.

One of the most common ways of identifying b-jets is through the Combined Secondary Vertex [66] methods. Multi variate methods use information from the jets and secondary vertices (displacement, total momentum, number of tracks, reconstructed mass, ...) and combine then into powerful discriminants. This technique assigns a score to every jet, between zero for light jets and one for jets most likely to originate from a b quark. A

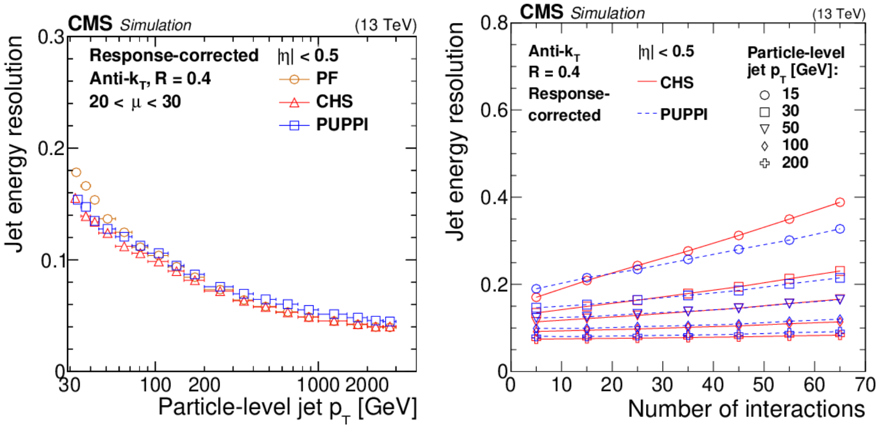


Figure 2.19. | Comparison of the jet energy resolution obtained through simulation of jets built with the PF, CHS and PUPPI methods. It is shown as a function of transverse momentum on the left, and of number of interaction on the right [65].

cut on this value is then performed when performing a physics analysis to select the desired b-jet efficiency and background rejection rate.

These discriminators evolved with the advances of machine learning techniques, using a deeper architecture and allowing additional charged particle inputs in a first time (deepCSV), then using a more complex architecture and allowing for many more inputs (16 properties of up to 25 charged objects and 6 properties of up to 25 neutral objects alongside 17 properties of up to four secondary vertices). This latest discriminant, deepFlavour, is currently the most sensitive and widespread b-tagger used by the CMS collaboration.

Usual cuts are defined from the probability of identifying light jets as b-jets. The Loose, Medium and Tight working points represent a 10%, 1% and 0.1% misidentification probability respectively. Using the Deep-CSV tagger, these working points correspond roughly to a 85%, 70% and 50% b-jet identification probability [67].

Similar techniques are used to identify hadronically decaying τ leptons or c-jets.

2.3.1.5. Missing transverse energy

Many physics analyses, including the one under consideration here, involve the production of neutrinos. Having an all but null probability of interacting with the detector, they cannot be directly seen. On the other hand, since the incoming particle beams are essentially longitudinal, the overall momentum in the transverse plane is expected to be negligible. A precise measurement of the transverse momentum imbalance, referred to as missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) is crucial to identify and quantify invisible objects in collisions.

One way of measuring the missing transverse energy is to perform the negative sum of all p_{T} reweighted objects using the PUPPI technique.

$$E_{\mathrm{T}}^{\mathrm{miss}} = -\sum \vec{p}_{\mathrm{T}}^{\,\mathrm{PUPPI}} \tag{2.4.}$$

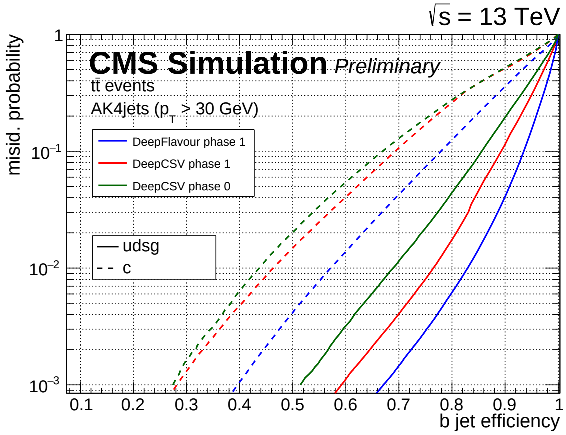


Figure 2.20. | Simulated background rejection efficiency of the different taggers as a function of b-jet efficiency before. The DeepCSV phase 0, in green, was simulated using the original cms pixel detector that has since been replaced [68].

To improve the resolution of the missing transverse energy, jet energy corrections are then applied. The energy of every jet with a transverse momentum above 15 GeV and less than 90% of its energy collected by the ECAL is corrected, after removing muons.

$$E_{\mathrm{T}}^{\text{miss, corr}} = E_{\mathrm{T}}^{\text{miss}} - \sum_{\text{jets}} \left(\vec{p}_{\mathrm{T}}^{\,\text{corr}} - \vec{p}_{\mathrm{T}}\right) \qquad (2.5.)$$

Finally, with such a complicated object, anomalous detector signals have to be suppressed. One striking example comes from the so-called spikes occurring in the ECAL detector when a particle has a nuclear interaction with the photo-detectors, causing an abnormally high energy deposit. If left unchecked, these relatively rare events could significantly increase the occurrences of high missing energy events. These signals, but also anomalous HCAL, HF deposits or particles originating from the beam halo are removed based on the specific signature they leave in the detector.

Missing transverse energy can be studied through the leptonic Z decay. The di-lepton system can be precisely reconstructed and the transverse Z recoil compared to the missing transverse energy of the event after excluding the two leptons. The $E_{\mathrm{T}}^{\text{miss}}$ resolution as a function of number of vertices is shown on figure 2.21, decomposed in the direction longitudinal and perpendicular to the reconstructed Z boson. Expected $E_{\mathrm{T}}^{\text{miss}}$ from simulation was subtracted.

2.3.2. Analysis strategy

With the collision data recorded and fully reconstructed, a wide variety of physics analyses can be performed.

Most analyses performed in CMS share a common methodology and structure. A specific physics process or range of physics processes are selected and compared to simulated events obtained by emulating the detector response on the output of the event

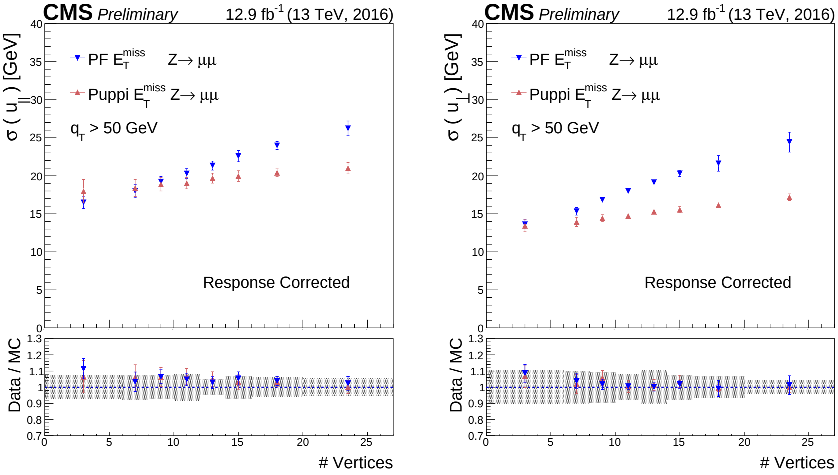


Figure 2.21. Missing transverse energy resolution obtained through muonic Z decays, as a function of the number of primary vertices. The plot on the left shows the missing transverse energy resolution in the direction transverse to the Z boson, and perpendicular to the Z boson on the right [69].

generators described earlier. The same simulated information is also used to compute efficiencies and acceptances, crucial to interpret the data collected.

A data analysis can consist either of the measurement of standard model properties, or searches of new processes outside the standard model.

To enhance the reach of searches or the precision of measurements, data enhancement techniques that can range from the simplest selections up to the usage of advanced machine learning techniques are used. Finally, statistical tests are performed to allow for quantitative conclusions.

2.3.2.1. Detector simulation

In order to convert the particles from event generators into a detector response, the CMS collaboration developed a simulation software based on the GEANT4 toolkit [70] that fully simulates the propagation and energy deposit of particles within the detector and transforms them into a detector response on which the reconstruction software can be run. This method is known as full simulation.

A precise geometry of the CMS detector, that not only includes active material, but also any additional material such as electronics, cooling pipes or support structure, is used to assess interaction probabilities. Based on these distributions and the measured magnetic field, the software will propagate particles through the material, assigning interactions and energy deposits using a Monte Carlo technique.

The energy deposited in the active material is then converted into detector hits. Simulated minimum-bias events are used to account for pile-up effects. Events are selected according to the desired Poisson pile-up distribution, and their hits are superimposed with the hits from the event generators. This is performed both for collisions originating from the same (in-time pile-up) or different bunch crossings (out-of-time pile-up).

The electronics are also simulated, based on noise and detector response models, leading to a simulated raw file that can be fed to the same reconstruction software as collision data.

This full simulation method has shown over the years to be a reliable, precise and ever-evolving tool. The main draw back from this technique is that even with many improvements, it remains a CPU-intensive task. While it is not a significant issue for signal hypotheses simulations, heavily suppressed backgrounds that have a high cross-section represent a major challenge to this method. To overcome this limitation, some data-driven techniques can be applied on physics analyses.

In the case of analyses that cannot afford to use extensive computing ressources, such as prospective analyses, additional fast simulation tools were developed.

One example of these tools, that is used later in this document, is the DELPHES package [71]. In order to gain time, the detector response for different objects under different pile-up situations was parameterised as a function of transverse momentum and pseudorapidity.

This parametric simulation is tuned on full-simulation of the desired detector and allows for fast, but less reliable production of samples from generator data.

2.3.2.2. Data enhancement

Depending on the final state considered, any physics data analysis will start by defining the desired object topology. A basic event pre-selection will be applied based on the different objects present in the final state of the signal process. These cuts allow to reduce significantly the amount of simulated and collision data to be considered and processed, while preserving the signal efficiency.

This pre-selection is usually not sufficient to ensure optimal sensitivity, and a major challenge in data analysis is to optimise the signal purity while not hindering its efficiency. Traditional methods involve reconstructing higher level, physics based variables (such as invariant masses, transverse masses, angular variables, and so on) and defining a signal region through cuts on these variables.

When performing searches, the phase space outside of this signal region is then used to ensure good data to simulation agreement, while keeping the signal region blinded to avoid any bias. When a good agreement is found and the systematic uncertainties are well understood, the blinded region is opened and a statistical analysis can be performed.

In more recent years, the advances in the field of machine learning allowed more powerful tools to be used to optimise signal sensitivity. Two of these methods will be featured at length in this document: boosted decision trees (BDT) and artificial neural networks (ANN).

The first method described here, the boosted decision trees, descends directly from the traditional selection cut methods. Given a series of input variables and two datasets (one for signal, and one for background), supervised machine learning algorithms can easily deduce the ideal sets of cuts to optimise any given function. A common way of proceeding is to maximise background rejection at a given signal efficiency through the application of a genetic algorithm [72], for example as implemented in the GARCON software [73].

While this method already provides an improvement with respect to previous methods, it will always result in a rectangular region of the phase space being accepted. Any correlation existing between the variables will therefore be ignored. To allow any correlation to be added, conditional cuts need to be applied, transforming a flat set of cuts into a decision tree, moving from a rectangular phase space acceptance region into the a sum of rectangles. A simple example highlighting the interest of decision trees is shown on figure 2.22.

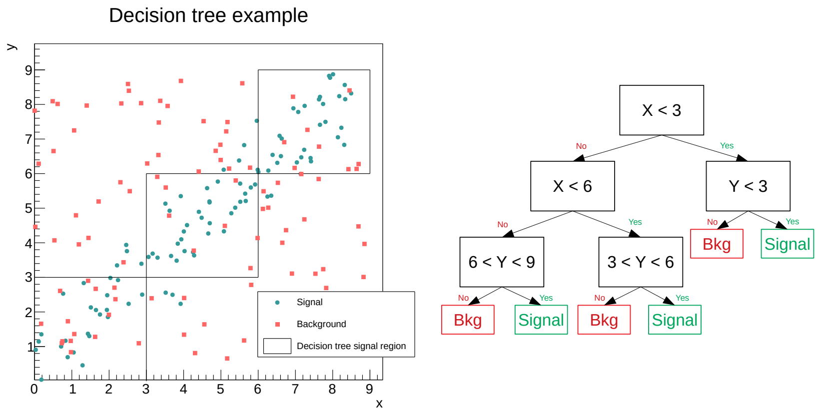


Figure 2.22. | Illustration of a simple decision tree on a uniform background and a diagonal signal (see on the left). While rectangular cuts can only provide small changes in signal to noise ratio, the simple decision tree displayed yields a 84% signal efficiency for a 67% background rejection efficiency.

A decision tree can be procedurally generated, or grown iteratively using a supervised machine learning algorithm. Starting from the head node, cuts will be defined to optimise a given metric. A commonly used metric is the Giny Index given by the product $p \cdot (1-p)$, where p is the purity defined as the ratio between signal events and total number of events in a node. This index will be maximal in a fully mixed node ($p = 0.5$), and minimal when containing only signal or background events.

For every node in the graph, the most sensitive variable and cuts with respect to the desired metric will be selected, until a node either contains only background or signal events, or if the number of events in this node is below a certain threshold.

These automatically generated decision trees, while yielding good performance and having a clear and easy interpretation are very susceptible to statistical fluctuations. A small change in the training sample can change the optimal variables upon which to cut, and completely change the decision tree. Furthermore, they are prone to over-fitting, which is the practice of learning statistical fluctuations of the input dataset.

To circumvent these issues, a common practice is to train multiple decision trees on a given dataset. Random subsets of the data are used to train each individual tree and the resulting object is called a random forest. Given an input event, a score based on the number of trees that classified it as signal will be returned.

Building on this method, boosted decision trees also train an ensemble of multiple trees, this time by reweighting the input data through a boosting algorithm. Assuming we have an input dataset containing N events, the algorithm starts by assigning an initial uniform weight $w_{i,0} = \frac{1}{N}$ to every event. A weak tree h_0 is trained using these weights and a signal classification is given a +1 value, while the background is assigned a −1.

Then, the algorithm will operate recursively. Given a tree h_t, the error rate ϵ_t is computed as the sum of misclassified weights and all misclassified events see their weight multiplied by a common factor :

$$\alpha_t = \beta \log\left(\frac{1-\epsilon_t}{\epsilon_t}\right) \tag{2.6.}$$

$$w_{i,t+1} = \begin{cases} e^{\alpha_t} w_{i,t} & \text{if i misclassified} \\ w_{i,t} & \text{otherwise} \end{cases} \tag{2.7.}$$

where β is a configurable parameter. The whole sample is then renormalised to have a total weight of 1, and a new tree, h_{t+1} is trained. Given an ensemble of T decision trees, the output of the boosted decision tree will be :

$$y_{\text{Boost}}(x) = \frac{1}{\sum_{i=1}^{T} \alpha_i} \sum_{i=1}^{T} \alpha_i \cdot h_i(x) \tag{2.8.}$$

This particular boosting algorithm is an example of an adaptive boost, or adaboost. It allows to convert weak learners into a more powerful discriminator. It is widely used in the high-energy physics community, either to generate different regions based on the BDT score, or as a sensitive variable used for statistical analysis. It is also used in object reconstruction to optimise background rejection for a required selection efficiency, as stated earlier.

The second main class of machine learning based classifiers used in this document are artificial neural networks, and more specifically multi-layer perceptrons (MLP).

While nodes of a decision tree act on a single variable, the basic unit of an MLP, a perceptron, will act on a linear combination of inputs. Each individual perceptron of index i and of input variables $\left\{x_j\right\}$ can be described by input weights, $w_{i,j}$, a bias b_i and an activation function φ. The output of the perceptron will be given by :

$$y_i = \varphi\left(b_i + \sum_j w_{i,j} \cdot x_j\right). \tag{2.9.}$$

Different activation functions can be used, from the traditional step and logistic functions to the Rectified Linear Unit function[1] or Scaled Exponential Linear Unit function [2], more common in recent networks.

---

[1] $ReLU(x) = \max(0,x)$

[2] $\text{selu}(x) = \lambda x$ if $x \geq 0$ and $\lambda \cdot (\alpha e^{-x} - \alpha)$ if $x < 0$

These neurons are then arranged into a number of hidden layers and an output layer. For signal/background classification, a single output neuron is required. For a given layer, each neuron is connected to every neuron from the previous layer, and the inputs of the first layer are connected to the event variables. An example of neural network is shown on figure 2.23.

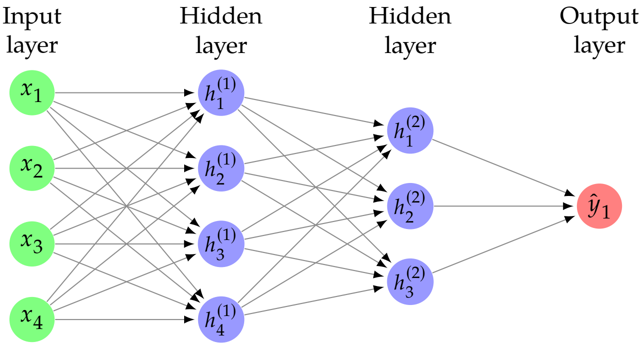


Figure 2.23. Example of multi-layer perceptron with four input parameters, two hidden layers of four and three neurons each, and a single output neuron.

In this feed-forward network structure, the connections between the nodes do not form cycles and the input information will travel through the network unidirectionally. The MLP can be seen as a function $y_{mlp} : \mathbb{R}^n \to \mathbb{R}^m$ with n the number of input variables and m the number of output neurons (here, $m = 1$).

The different neuron weights ($\vec{w}$) are free parameters and the process of finding the optimal weights to minimise a given loss function is called training. A common training method is to use the mean square error on a training sample as loss function :

$$E(\vec{x}_1 \cdots \vec{x}_N | \vec{w}) = \sum_{i=1}^{N} \left(\vec{y}_{mlp}(\vec{x}_i) - \hat{y}_i \right)^2 \tag{2.10.}$$

where $\hat{y}_i$ is the desired output of the network for the event $\vec{x}_i$.

Using differentiable neuron activation functions, the weights can then be fixed using the maximum gradient descent. At every training step, the weights are updated by a small shift in the direction with the steepest E decrease. This process is know as back-propagation.

While artificial neural networks are more difficult to interpret, they have the advantage of being able to learn complex and non-linear relationships between input variables.

2.3.2.3. Statistical analysis

In searches, once the data is collected, signal and background expectations are simulated or modelled, and the sensitivity enhancement strategy is in place, a statistical analysis based on the likelihood ratio method can take place [74].

From the signal and background models, expectations in the different regions of the final discriminant, usually bins in multiple histograms, can be computed as a function of signal model parameters θ. With d_i being the number of observed events, s_i the expected number of signal events and b_i the expected number of background events, the total likelihood is given by the product in every region i of the probability of measuring d_i given the expectations, which is given by the Poisson distribution.

$$\mathcal{L}_{s+b}(\theta) = \prod_i p(d_i|s_i(\theta), b_i) = \prod_i \frac{e^{-(s_i(\theta)+b_i)}\left(s_i(\theta) + b_i\right)^{d_i}}{d_i!} \tag{2.11.}$$

The background likelihood is given by the same probability with $s_i = 0$ for all i. The likelihood ratio can then be defined as the ratio between both likelihoods and used to compute limits in the frequentist approach.

$$\mathrm{X}_{\mathrm{obs}} \equiv \frac{\mathcal{L}_{s+b}(\theta)}{\mathcal{L}_b} \sim \mathrm{X}_{s+b} \tag{2.12.}$$

$$\mathrm{CL}_{s+b}(\theta) \equiv p(\mathrm{X}_{s+b} < \mathrm{X}_{\mathrm{obs}}) \tag{2.13.}$$

Since the distribution of X_{s+b} is known, the confidence level of the signal plus background $\mathrm{CL}_{s+b}(\theta)$ can be computed to infer exclusion limits on the parameter θ. In this approach, measurements due to unfortunate statistical fluctuations below the expected background could lead to the exclusion of any signal hypotheses even at strong confidence levels.

To circumvent this issue, the confidence of the background itself is computed, and the Modified Frequentist confidence level CL_s is computed as the ratio of signal plus background to background confidence levels.

$$\mathrm{CL}_b \equiv p(\mathrm{X}_b < \mathrm{X}_{\mathrm{obs}}) \tag{2.14.}$$

$$\mathrm{CL}_s(\theta) \equiv \frac{\mathrm{CL}_{s+b}(\theta)}{\mathrm{CL}_b} \tag{2.15.}$$

Computing this last ratio as a function of the theoretical θ parameter and comparing it to the expected distribution allows the signal exclusion region at a given confidence level to be computed.

In the previous discussion, a perfect knowledge of the signal and background yields was assumed, which is usually not the case. Any theory uncertainty or detector mismodeling will impact the expected probabilities, and they have to be taken into account.

These systematic uncertainties are usually modelled and constrained through dedicated measurements, and allowed to vary based on this model during the statistical analysis.

In the event of an excess of data with respect to the background only hypotheses, the significance of this excess is computed from the p-value obtained from the likelihood ratio. The p-value is converted into a significance Z as the corresponding threshold of a one-sided Gaussian distribution.

$$p = \int_Z^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx \tag{2.16.}$$

2.4. Higgs boson pair production measurements

The process under consideration in this document, the production of $\mathrm{HH} \to \mathrm{b\bar{b}}VV \to \mathrm{b\bar{b}}l\nu l\nu$, where V denotes either a W or a Z boson, has been previously studied using run-2 data [75,76]. A brief sketch of the non-resonant CMS analysis will be given here[1].

The basic event selection, aimed at reducing the amount of data to process while keeping a high signal efficiency is the following. Events are required to have two opposite sign leptons, with leading/sub-leading transverse momentum above 25 and 15 GeV, 20 and 10 GeV, 25 and 10 GeV and 25 and 15 GeV in the $\mathrm{e}^{\pm}\,\mathrm{e}^{\mp}$, $\mu^{\pm}\,\mu^{\mp}$, $\mathrm{e}^{\pm}\,\mu^{\mp}$ and $\mu^{\pm}\,\mathrm{e}^{\mp}$ categories respectively. These values were tuned to the corresponding high-level trigger cuts.

Electrons are required to be in the $|\eta| < 2.5$ region, with a relative isolation within a radius of 0.3 to be below 0.04. Muons need to fall in the $|\eta| < 2.4$ region and have a relative isolation within $\Delta R = 0.4$ below 0.15.

Jets with a transverse momentum above 20 GeV, $|\eta| < 2.4$ and with a minimum radius with respect to selected leptons $\Delta R_{j,l}$ bigger than 0.3. If more than two jets meet these requirements, the two jets with the highest b-tagging score are selected. These jets are considered as b-tagged if they pass the medium working point of the algorithm.

Finally, the invariant mass of the di-lepton system is required to be above 12 GeV to avoid quarkonia resonances and to be below $m_{ll} < m_{\mathrm{Z}} - 15\,\mathrm{GeV}$ to remove overwhelming contributions from the Z peak and the high Drell-Yan and $t\bar{t}$ tails.

The signal and background contributions are estimated through a full simulation. In the $\mathrm{e}^{\pm}\,\mathrm{e}^{\mp}$ and $\mu^{\pm}\,\mu^{\mp}$ channels, a huge Drell-Yan contribution is expected to be highly suppressed once the b-tagging criteria is applied. The simulation statistics at the time were insufficient to appropriately model this contribution and a data driven approach was used.

The idea of this data-driven estimation is to reweight events with two untagged jets depending on the corresponding b-tagging efficiency. While it is relatively easy to compute this latest value as a function of kinematic variables, namely p_{T} and η, the b-tagging efficiency crucially depends on the flavour of the jet, and more specifically if it is a light, c-jet or b-jet. Since the true nature of the jet cannot be known, an average flavour contribution is computed based on simulation. This value depending strongly on the jet kinematics, a boosted decision tree was trained to distinguish DY + b-jets or c-jets from other processes, and the average fraction of events with jets of the given flavours is computed as a function of this decision tree score.

Using this selection and background modelling, a good agreement between data and Monte-Carlo was found. To optimise sensitivity, a neural network was trained on eight kinematic variables : the invariant mass of the di-lepton system m_{ll}, the angular separation between the two leptons and between the two jets, ΔR_{ll}, ΔR_{jj}, the di-lepton and di-jet transverse momentum p_{T}^{ll}, p_{T}^{jj}, the difference in azimutal angle between the di-jet and di-lepton systems $\Delta\phi(ll,jj)$, the minimum angular separation between leptons and jets $\min\left(\Delta R_{jl}\right)$ and the transverse mass given by

---

[1] The run 2 analysis features both resonant and non-resonant results, but this document will only focus on non-resonant Higgs boson pair production.

$$m_T = \sqrt{2 p_{\mathrm{T}}^{ll} p_{\mathrm{T}}^{\mathrm{miss}} \left[1 - \cos \Delta\phi \left(ll, p_{\mathrm{T}}^{\mathrm{miss}}\right)\right]}. \tag{2.17.}$$

In addition to these variables, the theoretical higgs self-coupling and higgs-top coupling modifier parameters, κ_λ and κ_t, are added as inputs to the network. The network can thus be trained simultaneously on multiple signal models and infer their behaviour as a function of these parameters. It could be shown that this parametric network has performances comparable to dedicated networks on data points on which they were trained, and that it can interpolate well on points upon which it wasn't trained.

It could be shown by separating the data in different categories in m_{jj} that a good data to simulation agreement in the output of the network was found in the signal depleted categories, as shown on figure 2.24.

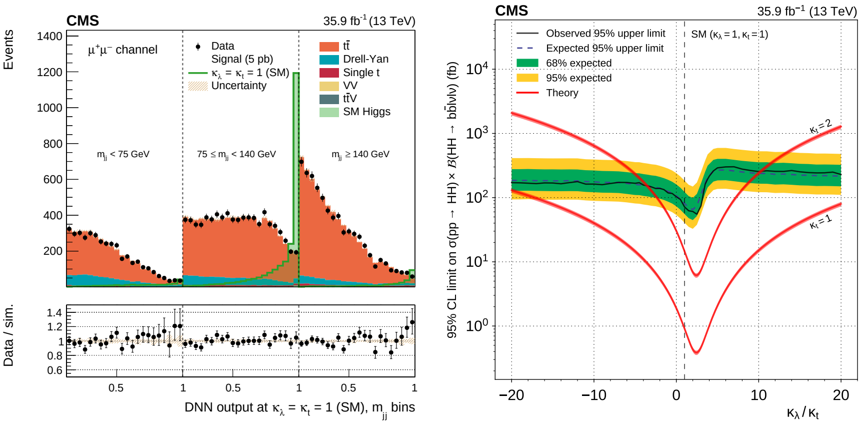


Figure 2.24. | The output of the parameterised neural network for data and simulation evaluated with $\kappa_\lambda = \kappa_t = 1$ is shown on the left plot, binned in m_{jj}. The corresponding expected and observed limits is shown to the right [75].

A statistical analysis on the output of this network evaluated for different κ_λ and κ_t values was performed. Upper limits on the production cross section were set in the modified frequentist approach at the confidence level of 95%, as shown on the right of figure 2.24. In the standard model hypothesis ($\kappa_\lambda = \kappa_t = 1$), the observed limit on the $\mathrm{HH} \rightarrow \mathrm{b\bar{b}}VV \rightarrow \mathrm{b\bar{b}}l\nu l\nu$ cross section times branching fraction was found to be of 72fb, in agreement with the expected limit of 81^{+47}_{-28}fb. This translates into an upper limit of 79 times the SM predictions when considering signal theoretical uncertainties. The full analysis description and results can be found here [75,77].

Other dedicated analyses were performed to cover different final states [79–81], and their results were combined, providing a more stringent limit [78]. In the standard model hypothesis, the observed limit on the HH production cross section was found to be of 22.2 times the standard model. Fixing all the other effective field theory parameters to their standard model values, the κ_λ parameter is constrained to the $-11.8 < \kappa_\lambda < 18.8$ region at the 95% confidence level.

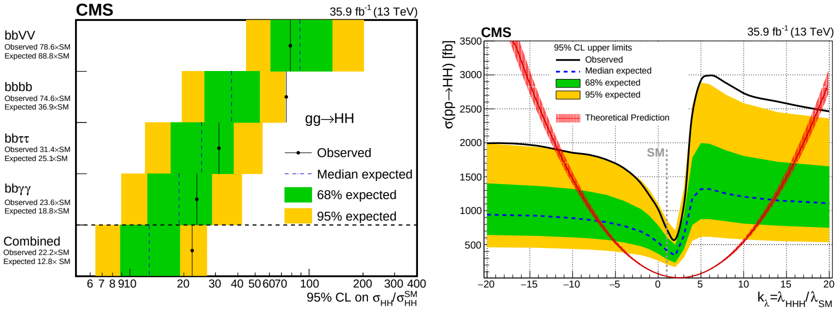


Figure 2.25. | The individual limits on the di-Higgs production cross sections in the standard model hypothesis and their combination is shown on the left. The plot to right shows the expected and observed combined limit as a function of κ_λ and the corresponding theoretical cross section [78].

Similar analyses and their combination [82] were performed by the ATLAS collaboration, yielding an observed (expected) limit of 6.7 (10.4) times the standard model prediction.

These analyses were performed on data collected in 2016, corresponding to an integrated luminosity of about 35 fb^{-1}. New versions of these analyses are being performed on the full run 2 data, corresponding to an integrated luminosity of about 150 fb^{-1}.

While an increase in luminosity as well as some improvements in the analysis strategies and better constraints on the systematic uncertainties will undoubtedly lead to better limits in the near future, the small value of the expected di-Higgs production cross section will render any direct discovery at the LHC impossible.

This measurement, among many others, would require an integrated luminosity that cannot be achieved with the current machine in a timely fashion. An instantaneous luminosity upgrade of the LHC would therefore be required, and the different detectors collecting the data would have to be upgraded accordingly to maintain physics performances under the new data taking conditions.

3. Chapter

The CMS experiment at the HL-LHC

The previous section has shown through the example of the Higgs boson pair production, that despite the outstanding performance of the accelerator complex and detectors, some physics analyses remain statistically limited. This is true for many measurements looking for slight deviation with respect to the standard model.

In order to extend the accelerator's discovery and measurement potential, it was decided to undertake an upgrade to increase the total delivered luminosity by an order of magnitude. This upgrade into a high-luminosity accelerator will have a significant impact on the different detectors which would otherwise suffer greatly if left untouched.

This section will provide a brief overview of the accelerator upgrade, followed by an outline of the main upgrades of the cms detector, with a special emphasis on the outer tracker.

3.1. The High-Luminosity LHC

In order to fully exploit the physics potential of the large hadron collider, an upgrade of the machine is foreseen. The target is to deliver a levelled $5\times10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ instantaneous luminosity, leading to an integrated luminosity of $250\,\mathrm{fb}^{-1}$ each operating year for a total integrated luminosity target of $3000\,\mathrm{fb}^{-1}$ [83].

To achieve such a goal, both the LHC itself, and its injectors will have to be upgraded. While the former will be installed during the long shutdown 3 planned for 2025, the LHC injectors' upgrade (LIU) is currently being performed. A timeline of the LHC and related upgrades is shown on figure 3.1.

The LIU [85, 86] will impact every accelerator in the LHC injection scheme, with a goal of providing a 2.5 times brighter beam with twice the intensity. As discussed in section 2.1.1, the first stage running until 2018 was the LINAC 2. After 40 years of service, this accelerator is being replaced by the LINAC 4 [87]. Unlike its predecessor which accelerated protons, the LINAC 4 accelerates H^- ions up to an increased energy of 160 MeV. This allows to inject the PSB through charge exchange rather than regular multi-turn betatron stacking, leading to a higher injection efficiency and providing a handle on the transverse distribution of protons.

Apart from the necessary transfer line modifications, the PSB will undergo a power supply and radio frequency upgrade, leading to an output energy increased to 2 GeV. This allows an improved brightness in the PS, and an upgrade of its radio frequency cavities ensures a higher beam intensity.

Finally, the radio frequency cavities of the SPS were also upgraded, together with some impedance mitigation work and upgraded beam dump and protection systems.

In the LHC, the biggest limitation comes from the insertion magnets surrounding the

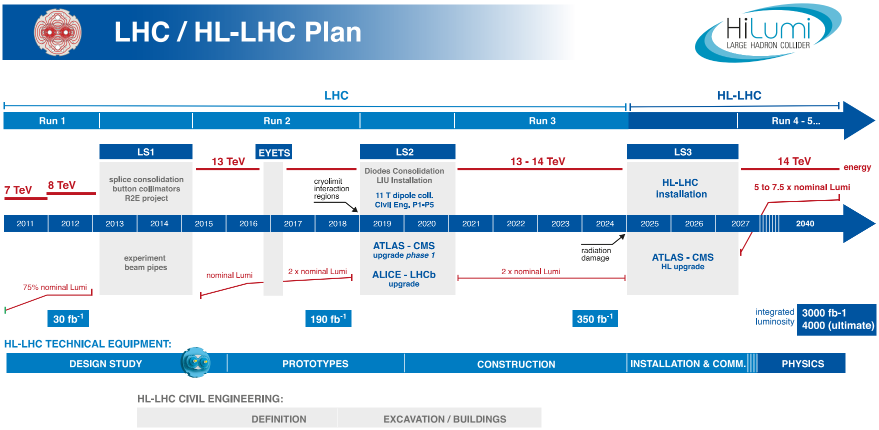


Figure 3.1. | Overseen timeline of LHC and HL-LHC operations. The LHC injector upgrade is currently under way, while the installation of HL-LHC components will occur during the long shutdown 3 [84].

interaction point. High instantaneous luminosities will cause high energy deposits in the quadrupole triplets that have to be evacuated by the cryogenic system. Furthermore, high integrated luminosities will induce radiation damage to these magnets. If the total integrated luminosity delivered were to be significantly above 300 fb^{-1}, the dose received by these magnets could induce risks of electric breakdown, causing significant downtime.

While a reduction in the β^* parameter is crucial in increasing luminosity, as outlined in equation 2.1, it also induces adverse effects. The reduction of the beam size at the interaction point will generate an increased beam angular spread, as a result of Liouville's theorem. In order to maintain beam separation outside of the interaction region, an increase in crossing angle is therefore required.

These two changes will also adversely impact the geometric parameter F defined in equation 2.2. The decrease in σ^* and the increase in crossing angle would both lead to a lower luminosity. As a way of mitigating this geometrical effect, the beam packets are to be rotated so as to collide head on, as seen on figure 3.2. This can be done through the introduction of pairs of superconducting crab-cavities each side of the interaction point inducing a transverse electric field.

Figure 3.2. | Sketch of bunches colliding in a regular accelerator on the left, and with the introduction of crab-cavities to the right [83].

The quadrupole triplets will be replaced by longer Nb_3Sn quadrupoles offering the wider aperture and gradient needed to increase β^*. The cryogenic systems will be replaced to allow a lower operational temperature and to allow the higher heat evacuation which is needed to remove additional energy deposited by beam fragments. Beam shielding is also to be improved.

The space available for the insertion magnets is fixed by interface with the experiment on the one side, and the unchanged LHC lattice on the other side. The extended length of the quadrupoles and the addition of crab-cavities, using an extra 10 m will be compensated by the replacement of the separation/recombination dipoles for higher gradient superconducting magnets.

Besides the insertion magnets, most of the LHC ring will be kept untouched during this upgrade. Some changes are nevertheless required to various subsystems. The collimators will have to be modified to protect the machine under the increased beam intensity while reducing their impact on the accelerator impedance. Beam monitoring and protection systems will also have to be updated, transfer lines will require adapting to the new operating conditions, cryogenic plants and segmentation will need to be improved and power supplies upgraded.

3.2. The CMS upgrade for the HL-LHC

The HL-LHC upgrade will bring great measurement opportunities for the CMS experiment, but also harsh data taking conditions.

The higher instantaneous luminosity will lead to higher pile-up levels and therefore higher particle densities. The upgraded detector is thus required to have a higher granularity to reduce occupancy, an increased bandwidth to accommodate higher data rates and improved Level-1 and high-level trigger capacities to maintain their ability to efficiently select events of interest while not compromising physics potential.

The total integrated luminosity will lead to an unprecedented accumulated radiation dose in the detector, in particular in the innermost and forward regions as shown on figure 3.3. Many sub-systems will thus need to be replaced, with a particular attention to radiation hardness.

In order to maintain the outstanding performance of the detector, a major upgrade (known as Phase-2 upgrade) will be necessary. Its design is well advanced and described in multiple Technical Design Reports [3, 89–92].

This section will start with the upgrade of the trigger and data acquisition systems that had a significant impact on other sub-systems. Then, the upgrade of the different sub-detectors will be described, starting in the innermost region with the pixel detector, then going outwards with the outer tracker, the new minimum ionising particle timing detector and the barrel electromagnetic and hadronic calorimeters. The replacement of the calorimeter endcap will then be discussed before finishing with the muon systems.

3.2.1. Trigger and data acquisition

The first systems affected by the increase in luminosity, driving many upgrade requirements for other sub-detectors, are the trigger and data acquisition systems.

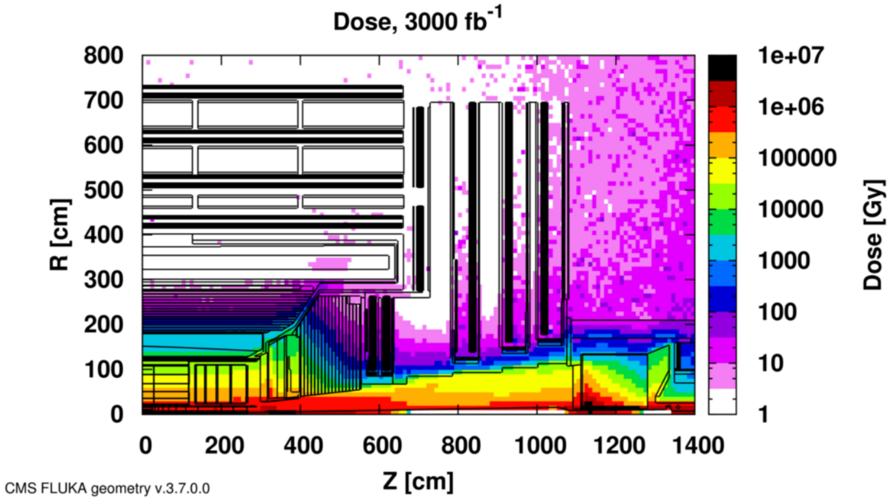


Figure 3.3. Absorbed dose in the CMS cavern at an integrated luminosity of 3000 fb^{-1} [88].

In order to achieve optimal performance in the final system, it is foreseen that the detector readout electronics and DAQ will be upgraded to allow a read-out rate of up to 750 kHz. An upgrade of the Level-1 trigger will therefore be needed to allow such a data reduction without compromising the physics potential of the experiment.

While the current trigger system only gathers information from the ECAL, HCAL and muon chambers (DT,CSC and RPC), the upgraded system will also include central tracking information. This will enable a new correlator trigger to match tracks with calorimeter and muon information, yielding numerous improvements in performance.

Given the high number of channels in the outer tracker, data reduction at the front-end level will be necessary. This will be done through local p_T measurement and the detector will only send information to the trigger system for hits compatible with tracks with a transverse momentum above 2 GeV. While a data reduction of an order of magnitude will be achieved by applying such a selection, an average of 200 tracks will still have to be provided to the central trigger within 5 µs, as will be detailed in section 3.2.2.4.

These tracks will then be used as standalone trigger primitives, or combined with calorimeter energy deposits or muon tracks to provide objects as close as possible to offline selection or physics analysis objects. The full granularity of the ECAL detector will therefore have to be made available to implement efficient track-matching and isolation evaluation in a high pile-up environment.

This increase in available information, along with an expected improvement in commercial FPGA performance will allow complex algorithms to be implemented, such as primary vertex reconstruction, displaced muon triggers and heavy stable charged particle identification. It is also foreseen that PF and PUPPI algorithms could be implemented to allow triggering on complex physical objects, including missing transverse energy and scalar summed p_T.

The complexity of the upgraded hardware trigger, while offering a clear performance improvement, will also require an increase in processing time. The detector readout electronics will therefore be upgraded to allow a maximum trigger latency of 12.5 µs, corresponding to 500 LHC bunch crossings.

Studies based on a simplified trigger menu show that key physics signals will safely be covered within the maximum rate constrain of 750 kHz [93].

Given the higher granularity requirement of the Phase-2 upgrade, an increase of event size from 2.0 to 7.4 MB is expected, leading the data acquisition system to read out the detector at an unprecedented data rate of up to 50 $\mathrm{Tbits^{-1}}$. To achieve this, each BE module will asynchronously transfer its accepted data to a DAQ concentrator, where it will be gathered before being transferred to the surface through standard protocols into the Event Network. This high performance switched network of commercial processors will assemble different fragments into a complete event. Once the event is built, it will be stored in a buffer awaiting the availability of a processor of the HLT that will perform a fast reconstruction and a final event selection.

Under 200 PU, it is expected that the HLT computing power required will be nearly 20 times higher than the current system. With an increased HLT accept rate of 7.5 kHz, the storage throughput should reach 61 $\mathrm{GBs^{-1}}$, compared to the current 2.5 $\mathrm{GBs^{-1}}$.

3.2.2. The Silicon tracker upgrade

The two innermost sub-detectors of the CMS experiment, the silicon pixel and strip detectors, will both require replacement due to a total accumulated dose way above their designed tolerance. As well as the improved radiation tolerance, the upgraded detectors need to remain efficient under an increased multiplicity, and the silicon outer tracker will be party to the Level-1 trigger decisions. The detector acceptance is also to be increased to allow effient tracking up to $|\eta| = 4$.

This subsection will describe the design of these new detectors and was adapted from a publication following a presentation by the author at the 14th Pisa Meeting on Advanced Detectors [4]. The author also presented these designs at the 39th International Conference on High Energy Physics [94].

3.2.2.1. Inner Tracker

Being so close to the beam pipe and the given maximal expected total fluence of $2.3\times10^{16}\mathrm{n_{eq}cm^{-2}}$, the radiation tolerance will be a key element for the inner tracker.

Two different silicon sensor designs are being considered, both n-in-p. They are either traditional thin planar sensors, with a thickness between 100 and 150 µm, or thicker 3 dimensional (3D) sensors. While 3D sensors have a higher intrinsic radiation tolerance due to a shorter charge collection path, the production process is more expensive and thus they are only considered for the first layers of the detector.

The sensors will be segmented into pixels six times smaller than for the current detector, with either squares of 50 µm ×50 µm or rectangles of 100µm×25µm aligned in the z direction for the barrel and r direction for the endcaps. Both geometries give similar tracking performance with a trade-off between primary vertex discrimination and p_T resolution. In high η regions, rectangular pixels would reduce the cluster size and allow for a higher charge collected per hit. As a result, the bandwidth consumption would be lowered while the signal to noise ratio would be increased.

The initial studies have been carried out for both geometries ; the results shown here will assume rectangular pixels.

The Pixel Read Out Chip (PROC) will need to meet multiple requirements, including low electronic noise, high radiation hardness and the ability to cope with hit rates approaching 3 GHz cm^{-2}. Given the similar requirements for the inner tracker of both the CMS and ATLAS detectors, initial design, prototyping and validation efforts were combined within the RD53 collaboration [95]. A first read-out chip, RD53A, was developed and validated and is used to characterise module prototypes.

Now entering the later stages of development, different flavours of the pixel chips are arising for the ATLAS and CMS collaborations, with the RD53B-ATLAS and C-ROC versions respectively.

3.2.2.2. Outer Tracker

The ability to participate in the Level-1 trigger [89] is one of the driving principles for the Outer Tracker. As previously stated, bandwidth limitations coupled with the necessity of sending out data at a 40MHz rate will impose data reduction at the detector level. This is achieved through "p_T modules" correlating hits from two sensors placed a few millimetres apart (see Figure 3.4). Tracks from the charged particles will be curved in the transverse plane by the 3.8T magnetic field with the bending depending on the p_T of the particle. Assuming the track originated from the primary vertex, the displacement of hits between the two sensors gives an indication of the transverse momentum. A "stub" is produced by correlating hits within a programmable window, and a data reduction of roughly an order of magnitude can be obtained for a window corresponding to a 2GeV cut. The stubs are read out for every bunch crossing but the full hit information will only be sent upon receiving an accept signal from the triggering system, at a 750kHz rate. Both data streams will be aggregated and sent asynchronously, with roughly 80% of the bandwidth used by the stub data (see Figure 3.5).

Two different types of module populate the outer tracker, both with 200µm thick n-in-p sensors. For the outer layers, the modules have two aligned strip sensors (2S). Both the

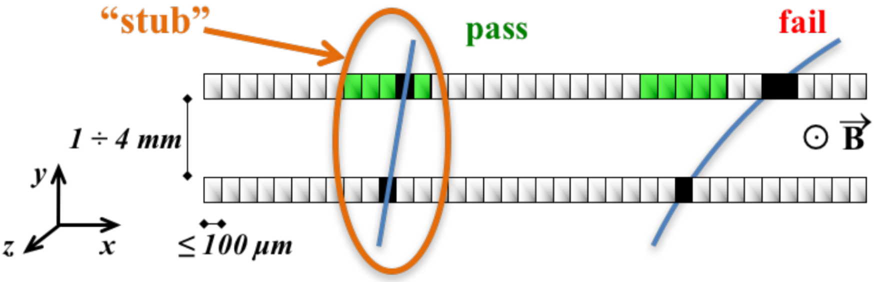


Figure 3.4. Illustration of the p_T module concept. Correlation of hits in the two close sensors allows to reject low p_T particles. The channels in green represent the programmable acceptance window.

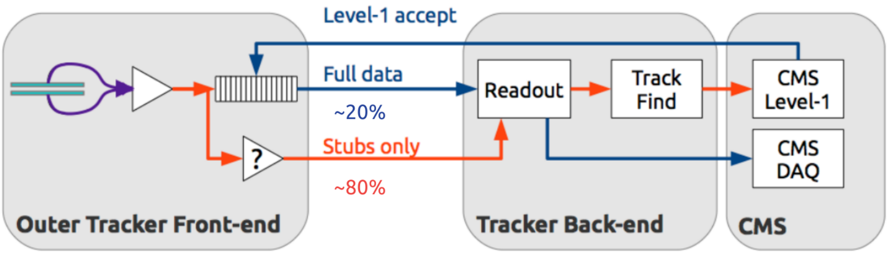


Figure 3.5. | Sketch of the read-out structure. Stubs (in orange) are read-out at 40MHz, sent to the track finding system and the tracks are then sent to the CMS Level-1 trigger. For selected events (at 750 kHz), a Level-1 accept signal is sent and the full data is read out from the detector.

sensors will be roughly 10 cm × 10 cm in size, with two rows of 1016 strips with 90µm pitch and 5 cm in length. The inner regions will be populated by modules with a macro pixel and a strip sensor (PS). The top sensor consists of two rows of 960 2.5 cm length × 100 µm pitch strips and the bottom sensor of 960 rows of 32 macro pixels of 100 µm pitch and 1.5 mm length, for a total size of roughly 5 cm×10 cm.

Three different read-out chips are used to process the signal generated by the passage of a charged particle. In PS modules, the strips are connected to eight Short-Strip ASICs (SSA) and the macro-pixels are bump-bonded to Marco-Pixel ASICs (MPA). The 2S modules use CMS binary Chips (CBC). The binary data generated by these read-out chips is aggregated by Concentrator Integrated Circuits (CIC) and sent to the trigger and data acquisition systems through a combination of a low power gigabit transceiver (lpGBT) chips [96] serialising the data which is sent to the off-detector electronics by a VTRx+ optoelectronic transceiver [97]. A more detailed description of 2S modules will be offered in section 4.1.1.

3.2.2.3. Layout and expected performance

A baseline tracker layout can be seen in Figure 3.6, with a central "barrel" region completed by two "endcaps".

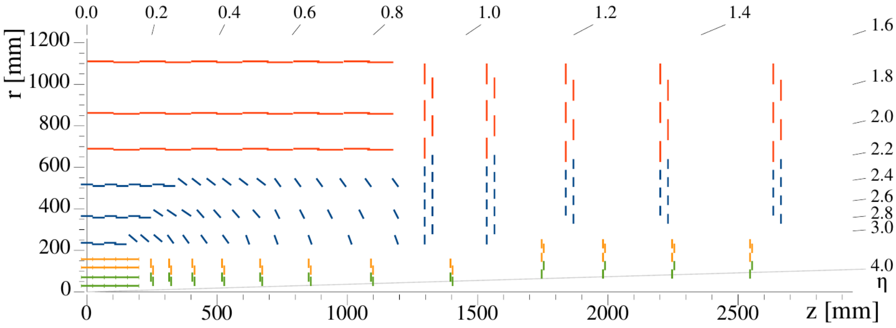


Figure 3.6. | Sketch of one quarter of the CMS phase 2 tracker in $r - z$ view. The green and yellow lines correspond to inner tracker modules with one and two read-out chips, while the blue and red lines represent the outer tracker PS and 2S modules.

The inner tracker has four barrel layers and the endcaps have eight small and four large double disks on each extremity of the barrel. This design will allow a geometrical coverage of up to $|\eta| \sim 4$ (with respect to a current coverage of up to $|\eta| \sim 2.5$), which, besides the physics gain from the forward acceptance, also mitigates pileup, particularly in the endcap calorimeters.

The outer tracker is composed of six barrel layers and five endcap double disks on each extremity of the barrel. The outer part ($r > 60$cm) is populated with 2S modules and the inner part has the PS modules. The first three layers of the barrel are tilted in order to keep the sensors as perpendicular to the tracks originating from $z = 0$ as possible. Although mechanically challenging, this geometry maximises the probability for a track to cross both sensors in the regions read by the same front end hybrid and therefore enhances stub finding efficiency.

The number of layers and wheels and their positioning as well as the service routing has been optimised in order to maintain a robust tracking while also minimising the material budget. Using this geometry and prototyping results, full Monte Carlo studies have been performed in order to estimate the detector tracking performance. The resolution on transverse momentum and impact parameter as a function of pseudorapidity for single muons is shown in Figure 3.7. The new geometry results in an improvement on both parameters for the whole η range.

In order to evaluate the track reconstruction performance, $t\bar{t}$ events were generated with superimposed minimum-bias events. Figure 3.8 shows the efficiency and fake rate for the tracks from the $t\bar{t}$ events under two pile-up scenarios. In both cases, the tracking efficiency is around 90% and the fake rate is kept under 2% for the 140 pileup scenario. The tracker geometry is still being optimised and the efficiency dip around $|\eta| \sim 1.2$ is being addressed.

3.2.2.4. Trigger primitives

Tracking at the Level-1 trigger is a challenging task. With a fixed 12.5µs trigger latency, tracks have to be produced within 5µs after the collision. Two different FPGA-based approaches are being considered to meet these time constraints.

In the first approach [98], pairs of stubs in adjacent layers/disks are combined to form a "tracklet". This tracklet is extrapolated to the rest of the tracker and matched stubs are then collected (see Figure 3.9). A track is then formed by performing a linearised χ^2 fit. Since multiple tracklets can originate from a single track, a filtering process will remove duplicates.

The second approach is based on Hough transform. Assuming particles of $p_T > 2$GeV originate from the primary vertex, any given stub in the detector can be produced by a track that respects a specific linear relation between its parameters [99]. By combining multiple stubs, a track is detected at the intersection of these lines in the parameter space (see Figure 3.10). The track parameters are then fitted through a Kalman filter before going through a duplicate removal process.

Hardware demonstrators were set up for both approaches and were fed with simulated $t\bar{t}$ events for different pile-up scenarios. In both cases, tracks were produced within the required timing constraints with very similar results. The tracking efficiency and resolu-

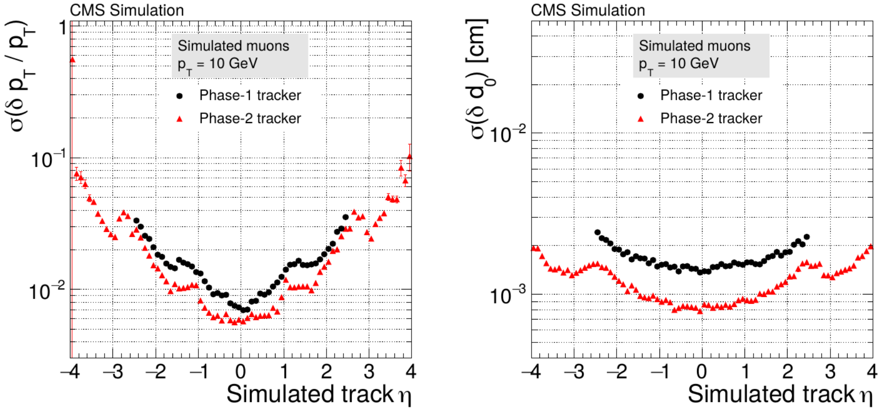


Figure 3.7. | Comparison of the resolution on the transverse momentum (left) and impact parameter (right) of isolated muons as a function of pseudo-rapidity between the phase-1 and phase-2 upgrade tracker.

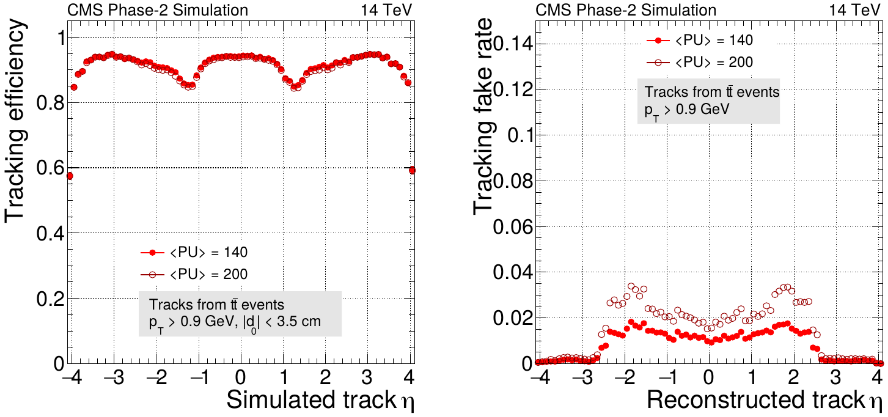


Figure 3.8. | Tracking efficiency (left) and fake rate (right) as a function of pseudorapidity for tracks originating from $t\bar{t}$ events with 140 and 200 superimposed minimum bias events.

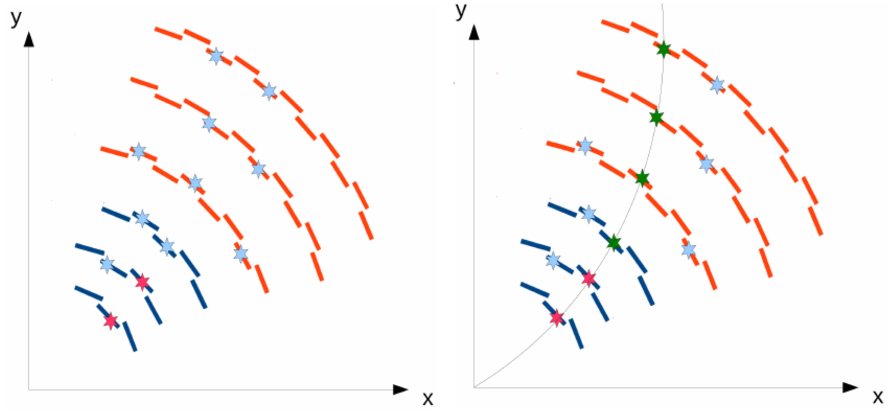


Figure 3.9. Illustration of the tracklet concept [98]. Here, layers one and two of the barrel are used to form a tracklet and stubs from the next layers are added by extrapolation.

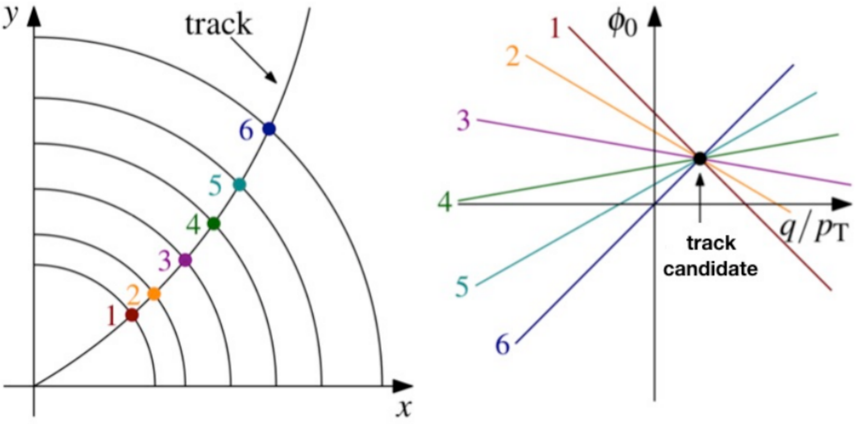


Figure 3.10. Illustration of the track finding algorithm through Hough transform [99]. Stubs in the geometrical space are converted to lines in the parameter space. The intersection of these lines identifies a track and its parameters.

tion on transverse momentum is shown on Figure 3.11. An efficiency exceeding 90% for high transverse momentum tracks even under high pile-up has been achieved, with a p_T resolution between two and six percent depending on the particle pseudorapidity.

3.2.3. The MIP timing detector

At the HL-LHC, the assignment of tracks to the primary vertex will be complicated by the increase in pile-up. Studies have shown that the optimal longitudinal track impact parameter (z_0) selection for primary vertex reconstruction is of roughly a millimetre to catch secondary interactions, decays and tracks in the resolution tail. Given the bunch length, it is expected that the line density will reach values of up to $2\,\mathrm{mm}^{-1}$, leading to a high risk of misidentified tracks, vertex merging, and overall loss in reconstruction performance.

To avoid this issue, a new-subdetector is being considered to add precise timing information to the reconstructed tracks : the Minimum-Ionising Particle Timing detector(MTD). Given the longitudinal spread of the beam, the collisions are not simultaneous, but are distributed over time with an RMS of 180 ps to 200 ps, mostly uncorrelated with track position. Timing information will help assign tracks to the correct vertex, with a starting resolution of 30 ps to 40 ps, increasing to 50 ps to 60 ps by the end of HL-LHC due to radiation damage.

This will be achieved by locating the MTD between the silicon tracker and the calorimeters. Given the very different conditions in the barrel and endcap regions, two different detectors, relying on different technological choices, were designed. They are referred to as the Barrel and Endcap Timing Layers (BTL and ETL respectively).

Given the low particle fluxes in the barrel region, a segmentation of 1 cm to 2 cm is enough to keep the occupancy at the percent level. This will be achieved through LYSO:CE scintillator rods, of 5.7 cm length in the ϕ direction, and 3 mm width in the z direction. These fast scintillators will be coupled to SiPMs on both sides to mitigate biases due to the light travelling time in the crystal.

The ASICs connected to the SiPMs, named TOFHIR (Time-of-flight, High Rate), will perform the time measurements. They will strongly amplify the signal to provide a sharp rise-time. A timing measurement will be provided by a discrimination of the leading edge followed by a time-to-digital converter. A pulse amplitude measurement corrects the time walk of the signal.

Read-out units will be made of three columns of 128 rods of scintillator, leading to 768 SiPMs, read out by 24 ASICs and connected to a concentrator card. Data transmission to the back-end upon Level-1 accept, precise clock delivery, control commands and monitoring will be performed by a (lpGBT).

The radiation and occupancy levels in the endcap region have guided the technological choice towards silicon sensors with a segmentation of 1.3 mm×1.3 mm. The timing precision needed cannot be reached with a single sensor, and a two disk system will provide two timing measurements per track leading to the desired resolution. The signal will be internally amplified, based on the Low Gain Avalanche Detector design (LGAD).

A module of 16×32 LGAD pads will be read out by two ASICs, amplifying the signal further. The rising and falling edges of the discriminated signal will then be measured

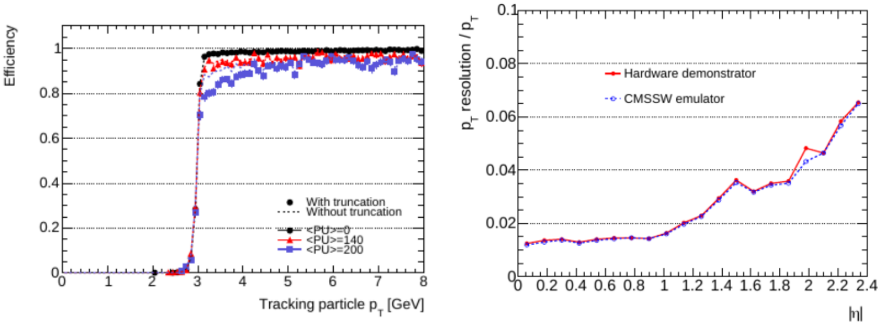


Figure 3.11. | Reconstruction efficiency for tracks from $t\bar{t}$ events as a function of transverse momentum under different pile-up assumptions in the tracklet demonstrators on the left. Transverse momentum resolution as a function pseudorapidity in the Hough transform approach is presented on the right.

to provide a time of arrival for each particle, corrected for time walk using the time-over-threshold method.

The mechanical design of the ETL will allow interventions during extended technical stops for possible repairs.

In order to minimise the time jitter due to electrical noise and dark current, highly temperature dependant, both BTL and ETL will be maintained between $-20°$ to $-30°$.

These design choices were validated through extensive simulations [90]. It has been shown that a 30 ps to 40 ps resolution will reduce vertex merging from 15% to 1%, and reduce the wrong track association by a factor two. Among the many improvements to reconstruction this would provide, some are particularly relevant to the study at hand. The b-jet identification and lepton isolation measurements will be improved. Global event variables sensitive to pile-up will also be improved significantly, including the missing transverse momentum.

3.2.4. Upgrade of the ECAL barrel

Due to its position in the detector, the ECAL barrel should not be too heavily impacted by the increase in luminosity. Although the crystals will suffer a loss in transparency, and the Avalanche Photo-Diode (APD) will have an increased dark current, steps can be taken to mitigate the need to replace the sensors.

Since the APD dark current scales exponentially with temperature [100], the baseline operating temperature will be decreased from the current 18°C to 9°C. As energy and timing resolution crucially depend on the signal to noise ratio, this change is expected to have a clear impact on physics object performance. Changes in the cooling plant serving the detector will be required, alongside a complete replacement of the coolant distribution pipes. The cooling system within the supermodules can however be kept.

The Very Front-End (VFE) cards, located on the detector and performing pulse amplification, shaping and digitisation, will be replaced to yield optimal noise performance and to help discriminate anomalous APD signals (termed "spikes"). The key feature of the new VFE is a reduced signal shaping time, leading to multiple performance

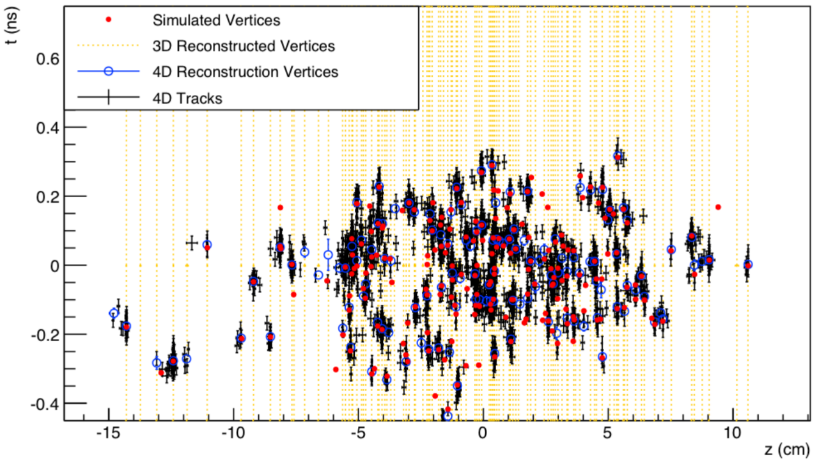


Figure 3.12. | Simulated and reconstructed tracks under 200PU. The yellow lines indicate vertices reconstructed without timing information, and the blue circles with timing information. [90]

enhancements. The APD leakage current contribution being proportional to the square root of the shaping time, a decrease in noise is expected. This allows in turn a reduction of out-of-time pile-up contribution and a better signal arrival time measurement. Finally, a shorter shaping time will increase the shape difference between signal and spikes, facilitating their tagging.

The read-out architecture will be changed, with upgraded FE cards sending single crystal data sampled at 160 MHz to the back end electronics for processing. This change requires an increase in data transfer rate of an order of magnitude which is made possible by recent developments in high speed radiation tolerant optical links (lpGBT and Versatile Link plus).

The back-end electronics being located in the service cavern, the processors will not have to be radiation tolerant and commercial FPGAs can therefore be used. These processors will need sufficient computing power to generate clock and control signals, perform the conversion of digitised pulses to transverse energy, basic energy clustering and generation of trigger primitives.

Multiple studies have been performed to estimate the impact of these upgrades on the overall performance. The amplitude resolution was shown to be sufficient to effectively suppress out-of-time pile-up using the MultiFit [101] algorithm[1] under the HL-LHC operating conditions.

Anomalous signals produced by occasional hadronic interaction with the APD (spikes) are localised high energy signals, often satisfying the conditions to trigger on isolated electrons or photons. Since spikes do not have a scintillation component, their time profile is different from signal (see Fig 3.13). The improved timing resolution and new spike rejection algorithms implemented in the BE boards will provide the necessary

---

[1] In this method, the profile of the signal is fitted alongside up to nine time-shifted contributions. This allows to extract in-time contribution from out-of-time pile-up.

rejection and the rate of spikes for a 20 GeV threshold will be negligible compared to the 0.5 MHz to 0.75 MHz Level-1 bandwidth allocated to the Level-1 trigger.

Finally, precise pulse timing measurement will improve primary vertex identification in $\gamma\gamma$ pairs, and therefore the opening angle between the photons. An assumed resolution of 30 ps will yield a $\simeq 10\%$ improvement in the mass resolution of a pair of photons originating from a Higgs. This will be of particular importance for the $HH \rightarrow b\bar{b}\gamma\gamma$ search described later.

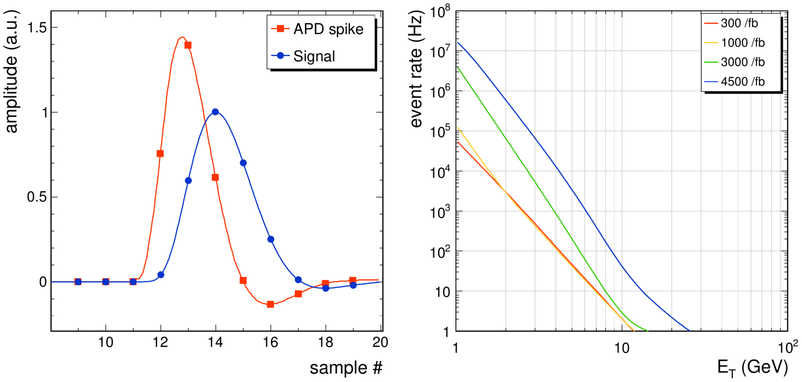


Figure 3.13. | Pulse shape from signal and spike on the left, sampled at 160 MHz. Expected event rate from spikes, assuming an average pile-up of 200 and 2800 colliding bunches on the right. [92]

3.2.5. Upgrade of the HCAL barrel

The HCAL barrel has three main elements which can be considered separately for the upgrade.

The first element is the brass absorber and the active material is comprised of plastic scintillators and optic fibers. The photosensors (initially HPDs), and related front-end, on-detector electronics, constitute the second element. The last item to consider is the back-end electronics, located in the service cavern.

Degradation to the performance of the HCAL detector due to radiation damage has been observed and studied since 2010. To mitigate this loss in performance, a sector of the HCAL endcap was upgraded during the 2016/17 year-end technical stop by replacing the HPDs by silicon photo-multipliers (SiPM). Data taken in 2017 have shown that most of the signal loss observed could be recovered using SiPMs, thus allowing to seperate the observed effects on the active material and on the photosensors.

It could therefore be concluded that the ageing due to radiation damage on the barrel active material will have a negligible impact on data taking at the HL-LHC, and they will therefore not be subjected to an upgrade. The photosensors are currently being replaced by SiPMs and, in the barrel, will continue to used for the HL-LHC period.

3.2.6. The high granularity calorimeter endcaps

The Calorimeter Endcaps (CE) were designed to be efficient up to an integrated luminosity of 500 fb^{-1}. Studies have shown that this sub-detector would suffer a significant loss in performance at higher integrated luminosities, and will therefore have to be replaced.

The upgraded CE will have to meet many specific requirements, alongside the radiation tolerance and Level-1 trigger compatibility needs common to all sub-detectors.

The lateral granularity of the detector will be improved. This will allow better pile-up energy removal, two-jet separation, position measurement and energy calibration using minimum ionising particles. The detector will therefore have to be dense to preserve the lateral compactness of showers and profit from the improved granularity.

A fine longitudinal segmentation will improve the energy resolution of electromagnetic showers, allow for pattern recognition algorithms and help with pile-up discrimination. Finally, a 100 ps timing resolution will further help reduce pile-up effects.

To accommodate these requirements, the following design (Fig. 3.14) is considered. The active material will be silicon for the highest fluence regions, and plastic scintillators read out by SiPM in less active regions. To mitigate radiation effects on the detectors and read-out chips, the whole CE volume will be cooled down to −30°C.

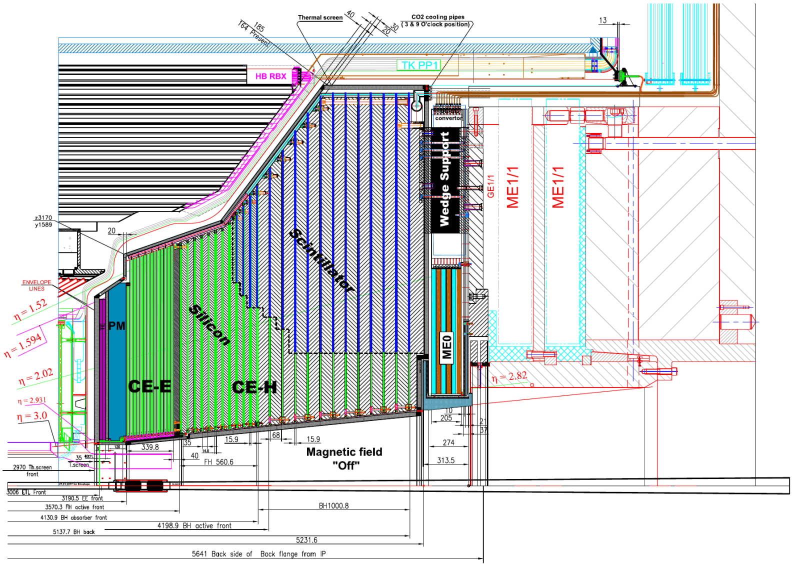


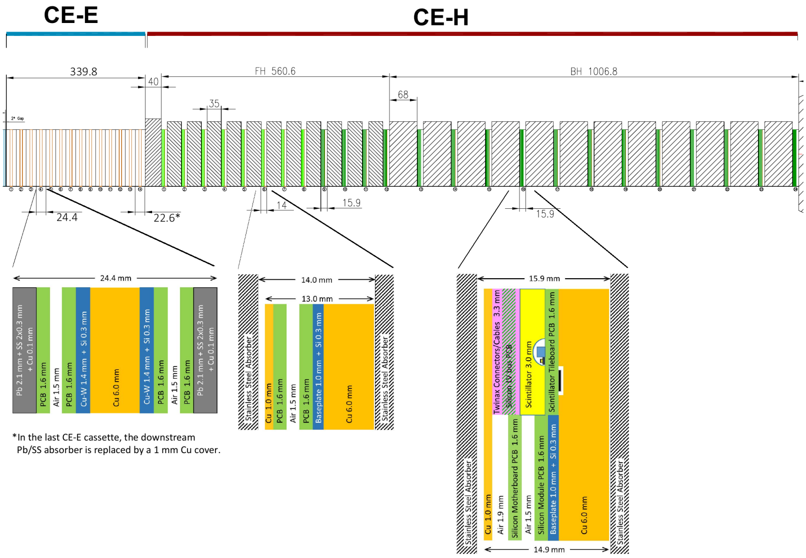
Figure 3.14. | Layout of half of an upgraded calorimeter endcaps [91].

The silicon module design results from the industrial silicon wafers being 8 inch in diameter. To optimally use the wafers and therefore minimise costs, the sensors will be hexagonal. The sensors will come in three different flavours, either 200 µm or 300 µm thick sensors segmented into either 192 1.18 cm^2cells, or 100 µm thick sensors segmented into 432 0.52 cm^2 cells. The thinner detectors will be used in the higher fluence regions for optimal radiation hardness while thicker sensors in the outer regions will provide a better signal to noise ratio.

The sensors will be fixed to a PCB, and the hexagonal cells will be connected to read-out chips via wire-bonding through holes in the PCB. A gold plated Kapton layer will provide backplane biasing, and the module will be completed by a metallic baseplate to provide adequate cooling.

Square plastic scintillators (either polyvinyltoluene or polystrene-based) arranged in a $R-\phi$ grid will pave low fluence regions of the detector. The geometry of the detector will cause the tiles to be of very different sizes, varying from $4\,\mathrm{cm}^2$ to $32\,\mathrm{cm}^2$. A dome will be left empty in the centre of the 3 mm thick detectors to accommodate a SiPM. The scintillator and SiPM will be mounted on a PCB hosting the front end read-out chips.

Absorber layers will be placed between both silicon and scintillator modules. The thickness and composition will differ between electromagnetic (EC-E) and hadronic (EC-H) sections, as shown on Fig. 3.15.



*In the last CE-E cassette, the downstream Pb/SS absorber is replaced by a 1 mm Cu cover.

Figure 3.15. | Longitudinal layout of the upgraded calorimeter endcaps [91].

The electronics associated with both silicon and scintillator detectors will be split on two boards. While the read-out chips pre-amplifying and processing the signal will be attached directly to the detectors, a motherboard connected to multiple modules will host a concentrator and various services.

To address the challenges of a high luminosity environment, the read-out chips will have to meet strict requirements. The thin silicon sensors being the most challenging detector to read out, the silicon read-out chip will be described here. While not fully defined at the time of writing, the scintillator read-out chips should be very similar in conception to its silicon counterpart.

The calibration of the system will rely on its ability to measure MIPs, imposing both a low noise and a very high dynamic range in the energy measurement. Two methods with overlapping ranges are used depending on the energy deposited. For small energy deposits, with a charge under 100 fC to 150 fC, an ADC will measure the pulse height.

For higher deposits, with a charge above 50 fC, a TDC measuring the time-over-threshold will provide the signal amplitude. This requires a fast signal shaping time to measure high energy deposits while minimising OOT contributions. Both methods are required to have a high linearity on their respective energy range.

A second TDC will provide a 100 ps resolution for pulses corresponding to at least three MIPs, allowing for pile-up rejection. The digital part of the read-out chip will store the three measurements in a buffer awaiting Level-1 accept signal. At that point, the data will be zero-suppressed and sent to the concentrator chip located on the motherboard.

Simultaneously, ADC and Time-over-Threshold information will be combined and calibrated to offer a single charge measurement for triggering purposes. This information is sent for every bunch crossing to the concentrator through a dedicated line.

As for many other detector-level chips, a special care has to be taken to ensure radiation tolerance and low power consumption.

The motherboards connected to multiple modules will perform multiple functions. The concentrator ASIC will aggregate format and serialise the data sent to the service cavern upon Level-1 accept through lpGBT links. This chip will also select the trigger sums of interest and store them in a FIFO. They will then be transferred to the triggering system through a dedicated optical line.

The motherboard will also distribute clock signals, bunch crossing counter, reset and control signals to the FE. Finally, this board will also distribute power to the modules.

Multiple studies have been conducted based on this design to evaluate its impact on physics. It has been shown that the selected longitudinal and transverse granularity, combined with timing measurements, allow for efficient particle-flow calorimetry preserving, and in some selected areas, improving the physics potential of the detector under HL-LHC conditions.

3.2.7. Upgrade of the muon detectors

The muon detectors play a crucial role in CMS, with many sensitive signatures including one or more muon in their final state. To maintain the excellent detector performance, an upgrade of the current system is necessary to meet the numerous requirements of HL-LHC data-taking.

This upgrade will comprise two major components. The first will be to update the current sub-systems to be compatible with the increased Level-1 latency and read-out rates, to validate 30 years of longevity and to become compliant with new fluorine gas regulations. The second component will consist of improving the detector's capabilities in the forward region, thus requiring new sub-systems. The new sub-systems will allow to extend the acceptance up to $|\eta| = 2.4$ to match the extended tracker coverage, improve the resolution and redundancy in the high pseudorapidity region, and allow for standalone triggering in order to be sensitive to long-lived particles.

3.2.7.1. Upgrade of the existing muon detectors

The three current muon detectors (Drift Tubes, Cathod Strip Chambers and Resistive Plate Chambers) will be upgraded during LS3.

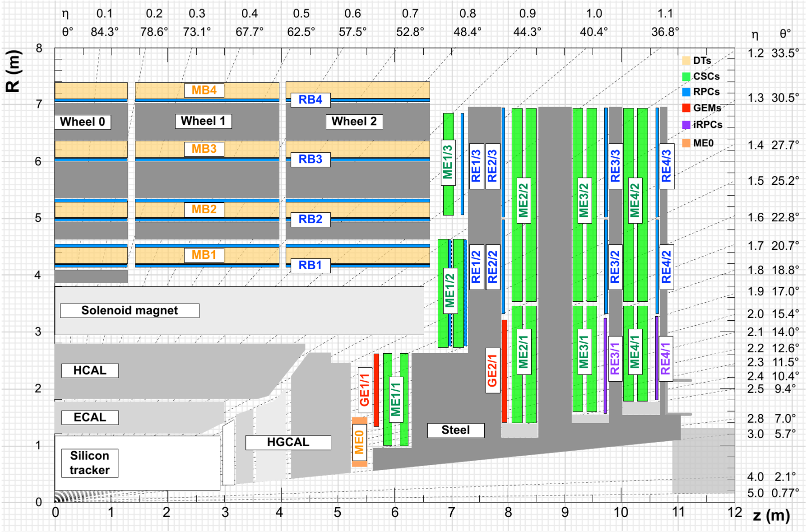


Figure 3.16. | R-z sketch of the CMS muon detectors, including the phase-1 (GE1/1) and phase-2 (ME0, GE2/1, RE3/1, RE4/1) upgrades.

Irradiation tests on the Drift Tube chambers have shown that some ageing is expected in the outermost layers due to ambient radiation. Wire deposits at the end of Phase-2 are expected to decrease the performance of the chamber by reducing the gas range, hit efficiency and breakdown voltage, and increase the background hit rate. This could lead to hit efficiencies as low as 50% in some chambers if nothing is done. Since the hit multiplicity per track is very high in this region of the detector (more than 40), the global effect on the track reconstruction will nevertheless stay limited with efficiencies remaining above 90(95) % in the barrel (overlap) region. Furthermore, mitigation measures are being explored. An increase in High-Voltage, lowering of Front-End threshold or an increase in gas flux could help recover the lost efficiency.

Moreover, the read-out electronics will be replaced and re-designed. The current FE boards, MiC1, which process the signal from the chambers, produce trigger primitives and read out the data, will become obsolete due to the increase in radiation levels and trigger changes. In the upgraded architecture, an increase in bandwidth will allow the trigger primitive generation to be located in the service cavern according multiple advantages. The trigger logic will be hosted by commercial FPGAs increasing the computing power and allowing for more complex algorithms using information from detectors located in different stations, leading to a better efficiency and timing resolution and thus improved noise rejection, bunch crossing identification and spatial resolution. Accessibility will also be greatly improved, leading to an easier maintenance, and on-detector power consumption will be reduced, leading to reduced cooling requirements.

Longevity tests on the CSCs have shown that no deterioration of the key chambers parameters is expected. While the chambers are expected to be used until the end of HL-LHC, studies are ongoing to find alternatives to the gas mixture currently used to limit or avoid the usage of CF4. An upgrade of the read-out system is needed due to both radiation hardness concerns and an increase in trigger latency and output bandwidth.

In the detector, with the exception of the ME1/1 ring, the current strip front-end boards (CFEB) store the analogue signal in capacitor arrays which will then be digitised if a trigger signal is received. In parallel, fast trigger information is sent to the trigger motherboards. The upgraded system will use digital cathode front-end boards (DCFEBs) equipped with flash ADCs and powerful FPGAs able to process and store up to 700 events, leading to negligible dead time in all innermost rings of the CSC. The Anode Local Charge Track (ALCT) boards reading out the wires will be replaced to host more powerful FPGAs able to keep up with the increased data rate. The increase in bandwidth requirements will be matched by the introduction of optical links between front-end boards and trigger and data motherboards that will be upgraded accordingly.

Similarly, the Resistive Plate Chambers were shown to be able to operate until the end of Phase-2 with an upgraded read-out system. The technology used in the current electronic boards is outdated and spare parts can no longer be found on the market. Furthermore, replacing the link system will enable the exploitation of the full 1.5 ns resolution of the chambers (against steps of 25 ns at the moment), leading to an improved background rejection, the possibility of HSCP triggering and facilitated synchronisation.

3.2.7.2. Improving muon measurements in the very forward region

Three new subsystems are envisaged to extend the muon system acceptance up to $|\eta| = 2.4$ to match the extended tracker coverage, improve the resolution and redundancy in the high pseudorapidity region, and to allow for standalone triggering. Two different Gas Electron Multiplier (GEM) flavours will be installed in the first two wheels and improved Resistive Plate Chambers (iRPC) will complete the system in the last two wheels.

The GEM technology is widely used in other physics experiments (with the notable examples of COMPASS [102], LHCb [103], KLOE-2 [104] and BESIII [105]) and has shown a high level of reliability for hit rates up to the order of a few $\mathrm{MHz\,cm^{-2}}$ and no ageing problems have been reported. These properties make it easily adequate for the high fluence regions of the CMS muon system and the technology has therefore been chosen for two new detectors.

A first GEM sub-detector (GE1/1) is currently being installed in the first wheel of the muon endcap. A ring of two layers of triple GEM detectors will improve the lever arm in the first station, leading to an improved measurement of the standalone track bending measurement, and therefore significantly reducing the Level-1 rates by rejecting mismeasured tracks [53].

The phase 2 upgrade of the muon detector will add a second GEM subdetector located in the second endcap wheel, extending the GEM system to $|\eta| = 2.4$ and allowing for a second precise local bending measurement. The enhanced redundancy in this challenging region will allow robust track reconstruction as of Level-1, sharply reducing the number of tracks with an overestimated momentum and therefore lowering the trigger rate at a given threshold. Finally, the two direction measurements will allow the

Level-1 trigger to measure the momentum for muons from displaced vertices, where the vertex constraint cannot be exploited.

This new detector will be segmented in chambers each with a pair of triple GEM detectors, covering 20° in ϕ and $|\eta|$ between 1.6 and 2.4. A binary read out of the signal production by strips aligned in the R direction will yield a resolution of the ϕ coordinate from 210 µm to 410 µm and a crude R measurement. The read-out system will allow a timing resolution below 7.5ns as a result of a constant fraction discriminator and support a read-out latency up to 25 µs. Trigger signals will be sent to both the EMTF directly and the CSC motherboard in order to improve its precision. Tests have been performed using gamma rays and neutrons and no ageing effect is expected.

A final muon GEM detector, the ME0 detector, will be installed behind the new HGCAL calorimeter, covering the $|\eta|$ region between 2.0 and 2.8. The main purpose of this sub-system will be to provide a trigger signal and additional offline hits in the very forward region. Six layers of triple-GEM detectors will improve the muon momentum measurement by adding up to six hits per track and improving the lever arm, leading to a suppression of mismeasured muons triggering at Level-1 of an order of magnitude. It will also offer a full GEM + CSC system in the forward region covered by the extended tracker, leading to an efficient muon tagging and an improved muon triggering both at Level-1 and at the HLT.

The final part of the muon system upgrade is the addition of improved RPC (iRPC) in endcap wheels two and three. The design of this chamber is similar to that of the current detector, with thinner electrodes and a narrower gas gap. This allows a smaller high voltage and a lowered charge production in a discharge, thus crucially reducing the recovery time of the chamber and detector ageing effects. The reduction in gas gain will be compensated by an increased signal amplification in improved front-end electronics. Both sides of the strips will be read by these new low noise electronics to provide a measurement of the coordinate along the strip.

The addition of the iRPCs will enhance the local muon resolution in the endcap wheels two and three by increasing the lever arm, improve the CSC time resolution by a factor two allowing for better background rejection and HSCP reconstruction. The coordinate measurement along the strips will help to resolve ambiguities in the endcap trigger and the improved redundancy will allow for robust track reconstruction in this difficult region.

3.2.8. Summary of the CMS upgrade for the HL-LHC

To maintain the detector's performance under the new data taking conditions and extend the physics reach of the large hadron collider, every sub-detector is to be upgraded. A summary of the changes described is outlined in table 3.1.

The task of the Level-1 trigger system can only be accomplished if the event description is improved by a higher overall granularity and the inclusion of tracking information. Furthermore, the trigger latency and accept rate are to be increased, requiring an adaptation of the sub-detectors' read-out systems.

A new pixel detector will be installed with thinner sensors segmented in smaller pixels. The read-out chips are designed to sustain unprecedented particle rates. The strip tracker

Detector	Main challenges	Main changes
Pixel	Fluence	Thin radiation tolerant sensors
	Particle rate	Improved granularity and high rate read-out chips
	Extending acceptance	New layout extending $\|\eta\|$ coverage to 4
Strip	Fluence	Thin radiation tolerant sensors
	Particle rate	Improved granularity and binary read-out
	L1 contribution	Data reduction at the detector level
MTD	Pile-up reduction	New timing layer detector
EB	Noise mitigation	New VFE cards and lower operational temperature
	L1 requirements	New BE boards
HB	L1 requirements	New BE boards
CE	Fluence	Use of silicon sensors in affected regions
	Pile-up reduction	Improved granularity and timing resolution
Muon	L1 requirements	New FE and BE boards
	Extending $\|\eta\|$ coverage	New GEM and improved RPC detectors

Table 3.1. | Summary of the CMS upgrade phase-2

will be fully replaced by a new detector with an improved granularity, a binary read-out and built-in triggering capabilities. A timing layer detector will be added between the tracker and calorimeters to provide a time-based pile-up reduction.

While the ECAL and HCAL will mainly be affected by changes to their read-out electronics, the calorimeter endcaps will be replaced. The new endcap will use silicon sensors in the highest fluence regions to improve radiation tolerance and an improved spatial and timing resolution will allow to mitigate pile-up induced effect.

Similarly, the muon system will be mostly affected by an electronics upgrade and the forward region will be reinforce by additional GEM stations and improved RPC detectors extending the coverage of the muon system up to $|\eta| = 2.4$.

The combination of all of these individual improvements provides a complete upgrade design that is expected to be up to the challenge presented by the high-luminosity LHC. Despite the unprecedented particle rate and radiation fluence, this proposed design should at least maintain the detector performance at the outstanding levels observed during the LHC era.

Chapter

4. Characterization of the cms binary chips and their read-out systems

In parallel to the design of the different upgrades, a long incremental effort in prototyping and validation is being carried out, both guided by and providing a feed-back to the design choices.

Even by restricting this document to the description of the outer-tracker upgrade, it is not possible to honour the wide variety of activities which were performed to ensure the feasibility of the project. The work undertaken with regard to the sensors, service and front-end hybrids, mechanical structure, PS modules and stub processing, for instance, will not be mentioned.

In this section, a broad description of the prototyping and in-beam testing of the cms binary read-out chips in the 2013 to 2018 time period will be elaborated, centred on the contribution of the author. The outcome of these tests has resulted in two publications [1, 2] and it was featured in the tracker upgrade technical design report [3].

As will be discussed in greater detail in this section, the author provided a significant contribution to this project. He actively participated in multiple beam tests in the preparation of the detector read-out chain and online data taking activities. His contributions also extended to the offline processing of the data collected, through both a crucial role in the production of the final data files required, and a contribution to beam test data analysis.

4.1. CBC2 design and test campaign

As discussed at length in the previous chapter, the tracker upgrade is crucial to maintain the outstanding performance of the CMS detector in the high luminosity era. The path towards this new detector necessitates multiple prototypes and thorough testing and development campaigns. This section will summarise briefly a series of test beams and related development centred on an iteration of the outer tracker 2S module read-out chip, the CMS Binary Chip version 2 (CBC2).

The first version of the read-out chip was created in the second half of 2010 and was then validated in beam. It had a reduced set of functionalities, being connected to a single sensor and lacking stub logic. The second version of the chip, discussed in this section, is connected to two sensors and hosts the stub building logic, but lacks the stub read-out capabilities. The third version of the chip, discussed in the next section, while not the final version of the chip, has the full set of functionalities required for the final setup.

Although the sensors and read-out chips underwent dedicated, standalone testing, this document will only discuss the validation of complete modules. These tests serve two main purposes : firstly the design choices can be validated and the driving principles can be proved to be working ; secondly, they serve as a catalyst for software and firmware development, integration exercise and allow for in-depth feedback on the full design.

Date	Facilities	Primary objective	Module type	Sensor
11/2013	DESY	First CBC2 module test and validation	2 CBC	n-in-p
06/2015	SPS	Module test with an irradiated sensor	2 CBC	n-in-p
11/2015	SPS	Full-size module test	8 CBC	n-in-p
05/2016	SPS	Module test with an irradiated sensor	2 CBC	n-in-p
10/2016	SPS	Precision measurements	2 CBC	n-in-p
11/2016	CRC	Heavy ionising particle measurements	1 CBC*	p-in-n

* 1 CBC working in a 2-CBC mini-module.

Table 4.1. | Summary table of CBC2 beam tests where the author offered a significant contribution. While multiple prototypes were usually tested during these experiments, only the prototype corresponding to the primary objective is described.

This section will start with a description of the prototypes being tested. The test setup and DAQ chain will then be described before delving into the procedure of qualifying a prototype and ending with the final conclusions which could be drawn.

4.1.1. 2S modules

A 2S module, as shown on figure 4.1, is a fully independent unit of two sensors, two front-end hybrids and a service hybrid. The sensors are each segmented in two rows of 1016 strips of 90µm pitch and 5 cm in length. Each row of strips will be read-out independently by eight CBC chips, each connected to 127 strips from the top and bottom sensors. The trigger and DAQ data of these chips will be aggregated by a Concentrator Integrated Circuit (CIC), and transmitted optically to the service cavern by lpGBT chips [96] that serialise the data which is then sent by the VTRx+ optoelectronic transceiver [97]. The module is powered through a pair of low-voltage cables connected to a dc-dc converter providing the required voltages to the different chips. A high-voltage connector is also necessary to bias the sensors.

Two different sensor spacing are envisaged for 2S modules. Most of them will have a 1.8 mm spacing, with the exception of modules on the innermost 2S ring of the three last endcap dees, and the two innermost 2S rings of the last endcap dee, that will have a spacing of 4 mm to improve stub p_T resolution.

Some components were not available at the time of the testing campaign, and the modules therefore lack some functionalities. The prototypes tested, while varying in size and polarity, always comprised a set of one to sixteen CBC2 chips bump-bonded to a front-end hybrid and wire-bonded to one or two sensors. The service hybrid is then connected to an interface board providing the multiple low voltages needed by the read-out chip and to a set of electrical data links. Two different modules are shown on figure 4.2.

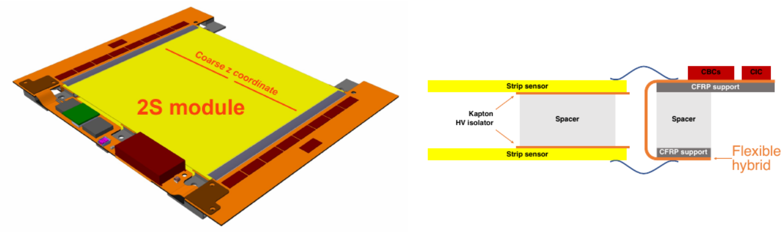


Figure 4.1. | The 2S module of the Outer Tracker. Two front-end hybrids (orange) are located each side of the sensors (yellow). A service hybrid hosting DC-DC converters and optical links is connected to the two front-end hybrids. [3].

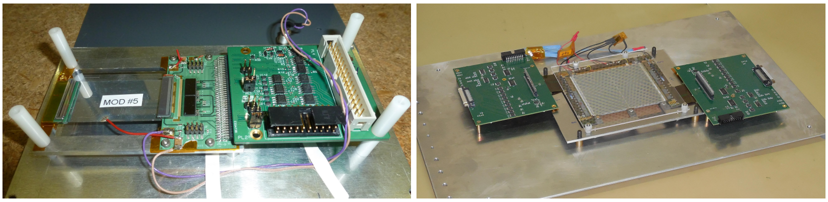


Figure 4.2. | Two different modules tested during this campaign. On the left, a picture of a 2 CBC2 minimodule and its interface board is shown. On the right, a full-size 16 CBC2 with two interface boards is displayed.

4.1.2. The CMS Binary Chip v2

As the CBC2 chip has a central role in data taking, a brief description of it is necessary in order to understand the beam test campaign. The chip can be seen as having two separate parts : an analogue part shaping the signal coming from the sensors and creating a binary hit/no-hit detect value, and a digital part processing these binary signals and hosting the stub-creation logic.

The analogue part of the chip, as displayed on figure 4.3, starts with a pre-amplifier wire-bonded to aluminium strips AC coupled to the sensors. The pre-amplifier receives the signal originating either from electrons or from holes (depending on the polarity of the sensor). The feedback resistors can be selected depending on the sensor polarity. For n-in-p sensors reading electrons, a single $200\,\mathrm{k\Omega}$ resistor is used, while a T-network yielding the same time constant can be selected for p-in-n sensors. This distinction is necessary in order to accommodate leakage currents of the order of $1\,\mu\mathrm{A}$ without saturating the preamplifier.

The preamplifier is then linked to the postamplifier through a capacitive coupling, thus removing any DC component resulting from the leakage current. The quiescent point of this amplifier can be set changing the $\mathrm{V}_{\mathrm{PLUS}}$ value, common to the whole chip. Finally, its output is connected to a resistor linked to a current source and to the input of the comparator. This current source, controllable channel-by-channel, allows any channel variation in comparator threshold to be corrected. The resulting DC offset voltage is referred to as $\mathrm{V}_{\mathrm{dda}}$.

The final hit/no-hit binary signal is provided by a 4-bit hysteresis comparator, of globally

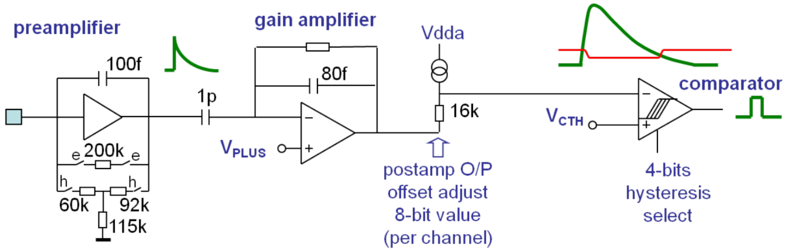


Figure 4.3. | Schematic description of the analogue part of the CBC2 read-out chip [106]. The signal induced on the strips is shaped by a preamplifier depending on the sensor polarity before being amplified. An offset is then applied independently on each channel and the signal is then converted to a binary value by a comparator.

set V_{CTH} threshold. A hit-detection logic is then required to synchronise the output of the comparator to the LHC clock. In this version of the chip, two different modes can be selected. In the first mode, a synchronised version of the output of the comparator is produced, meaning that a hit will be detected whenever the comparator output is true at the rising edge of the 25ns clock. In the second mode, a hit will be produced with a duration of 25 nanoseconds whenever the threshold is exceeded. This means that any short pulse over threshold will result in a hit, and also that longer signals will be shortened to produce a hit only in the first bunch crossing. The hit-detection logic can also be disabled and no hits would then be produced by the chip.

The digital part of the chip is shown on figure 4.4. Two different data-streams coexist, one for the full data and the other to produce trigger data.

The full-data stream is straightforward with the output of the hit detection logic directly sent to a 256 depth pipeline buffer, corresponding to a maximum latency of 6.4 µs. If a Level-1 trigger signal is received by the chip, the corresponding event will be sent to a 32-word deep data buffer by the pipeline controller before being serially read out.

The trigger-data stream starts with the same hits being sent to a channel mask. This first step allows for any noisy channel to be excluded. The different hits are then sent to the cluster width discriminator (CWD). This step, if enabled, aggregates hits from adjacent firing strips from a sensor into a cluster. If the cluster is wider than a programmable size, it will be discarded. If not, the central position of the cluster (rounded down in this version of the chip) will be sent downstream. If the CWD is disabled, every hit will be used.

These clusters are then shifted by an optional offset between -3 and +3 strips to correct for a shift in ϕ between the sensors due to geometrical lateral displacement within a module. This shift can be set independently for the two halves of the CBC2, resulting in 16 different regions within a module. Then, the resulting clusters are correlated between the top and the bottom sensors. If their absolute difference is smaller than a programmable window (between 1 and 8), a stub is produced.

In this version of the chip, the stubs cannot be read out directly. An OR logic is applied on the output of the combinatorial correlation, and a single stub-bit can then be read

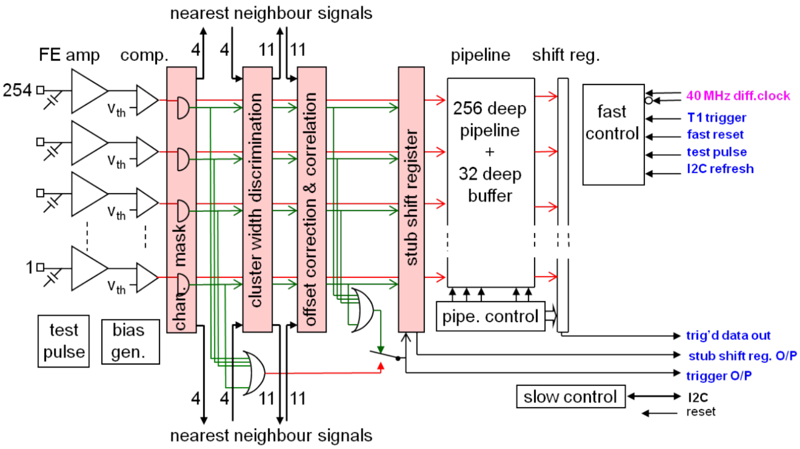


Figure 4.4. | Schematic of the binary logic part of the CBC2 read-out chip [106]. The red lines represent the data path, while the lines in green display the stub bit production path.

out. The chip can also be configured in order to read out an OR on the output of the hit-mask in place of the stub-bit.

The numerous programmable chip parameters can be set through the slow-control interface. This interface uses the standard I^{2}C protocol to write the desired parameter values in eight bit registers. Four dedicated lines allow for fast control of the chip by issuing trigger, fast reset, test pulse or I^{2}C refresh commands.

To avoid loss of efficiency at the edges of a read-out chip, the different CBC2 of a module are connected to their neighbours via fifteen lines each. Four lines are used by the CWD to get the last two channels from both top and bottom sensors, three are needed for the offset correction and the last eight allow for an overlap of the coincidence window on two chips.

4.1.3. A typical test beam setup

Since this section will present the results from a total of seven different beam-tests, a typical beam-test will be described here and any relevant specificities will be mentioned when discussing the results.

A typical geometrical layout is shown on figure 4.5. A beam of protons, electrons, muons or pions impinges on a module (denoted as the Device Under Test, or DUT) located on a rotation stage either within or after a telescope. This telescope, composed of multiple pixel detectors, allows track building to be used as a reference for efficiency measurements. Since the telescopes used have a long time integration period, the tracks produced are required to be correlated with an additional hit in a pixel detector clocked at 40 MHz. This latter detector is denoted as the reference plane. Additional fast scintillators placed in coincidence are used to provide a common trigger signal.

Additional auxiliary equipment is also required but not shown on figure 4.5. Low and high-voltage power supply units provide read-out chip powering and sensor bias

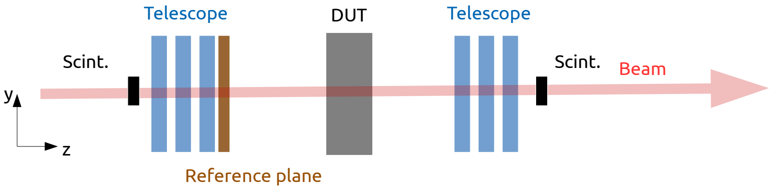


Figure 4.5. | Geometrical layout of a typical CBC2 test beam. The incoming particles will travel through scintillators used to trigger the setup, telescope planes used to build tracks to the module being tested, a reference plane used to identify in-time tracks and the device being tested.

respectively. Thermal regulation of the prototypes may be required, especially when working with irradiated sensors. A cooling element is thus often added either by the addition of a thermoelectric cooling element or by the use of a chiller unit circulating cooled dry gas. The same also applies to the telescope detectors.

A simplified overview of the data acquisition system is shown on figure 4.6. Correlated signals from the scintillators are sent to a remotely controllable trigger logic unit that ensures the synchronisation of all data streams. The trigger signal is then sent to either the telescope and reference plane or to the module DAQ. The module is read out through either a single read-out board or a double FE/BE system fulfilling a series of tasks. They provide a 40 MHz clock to the module and continuously read out stub information and store it in a cyclic buffer. They also provide an interface to propagate I^2C commands to configure the CBC2 and the fast signals. When a trigger signal is received, a TDC assigns a phase to the event and the trigger signal is propagated to the read-out chip. The corresponding event is then read out, combined with the appropriate stub and TDC phase, and stored locally awaiting to be sent to the acquisition computer through a gigabit Ethernet port.

The raw event, with some extra calibration information is then dumped to disk on a network attached storage device. The telescope and reference plane information, processed by a separate computer, will be stored on the same device awaiting offline processing.

Since compliance to the global CMS DAQ is required, the raw files are then processed by a miniature version of the XDAQ [107] based P5 acquisition software referred to as the playback chain. The output of this chain is then sent to the unpacker that converts the raw binary files into CMSSW object collections. To facilitate the data analysis, these files are then converted into plain ROOT [108] files by the ntuple producer, which are then merged with the track collections from the telescope and reference plane.

In parallel, the raw files are read by a DQM program providing basic distributions to verify the integrity of the data being taken.

The UCLouvain university played a central role in the data-taking chain, with the development and maintenance of the unpacker, a strong contribution in the playback chain and the ntuple producer as well as a coordination role between the different steps and associated data formats. A detailed description can be found in the additional content appendix A.

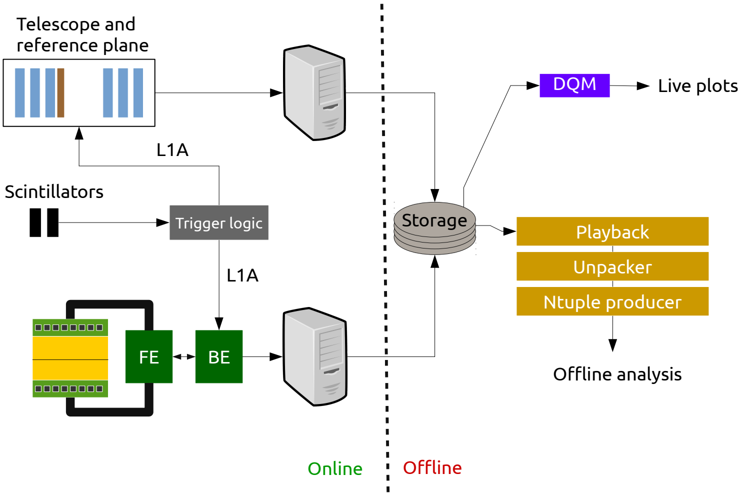


Figure 4.6. | Simplified overview of the data acquisition system used during the different beam-tests.

The author was strongly involved in these activities, making sure that the unpacker would run during these beam tests. As such, he actively debugged the code, underlined issues and suggested the fixes required to run. He also participated in the deployment of the playback chain used during beam tests and made sure the files saved to disk were processed in a timely fashion. He was also amongst the main contributors of the ntuple producer and often responsible for the batch production of the final files used for data analysis. Finally, in the preparation of the later beam tests, the author centralised the information related to the different data formats and served as a contact person to ensure compatibility between the different data processing steps.

4.1.4. Calibrating and characterising a module

At the beginning of every beam test, a full calibration of the system is required.

As described in section 4.1.2, the hit/no-hit binary output of the comparator depends on two global parameters, V_{CTH} and V_{PLUS} and an independent channel-by-channel offset to allow for a uniform pedestal within the chip. The calibration procedure for these 256 hexadecimal parameters is done in three consecutive steps.

The first step is to set a first set of calibration constants for both V_{CTH} and the offsets. These target values are chosen to provide the best possible linearity of the comparator threshold and to avoid any observed discrepancies in the comparator output. They are usually chosen to be of 0x78 and 0x50 respectively.

Then, a V_{PLUS} swipe is performed and the parameter is fixed to ensure an average 50% hit occupancy. Since this is done without any beam, the target V_{CTH} will be the average chip pedestal.

Finally, the offsets are scanned and also fixed to ensure a 50% channel occupancy. To minimise interference between neighbouring channels, this procedure is performed separately in eight groups of evenly separated channels with all other channels being disabled. An illustration of this offset tuning on a later iteration of the chip is shown on figure 4.7.

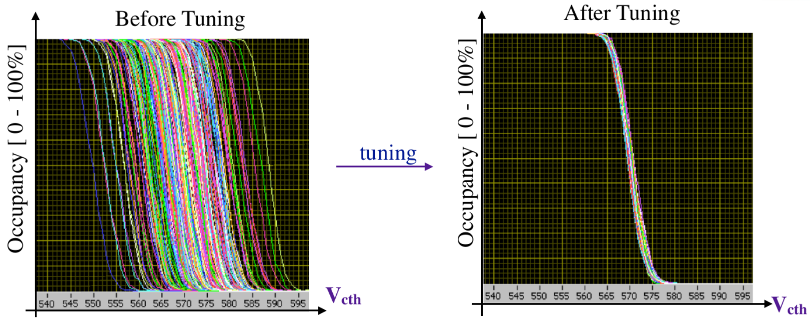


Figure 4.7. | Illustration of the effect of the offset tuning on a CBC3 [109].

While the calibration software allows for a full scan of these parameters, saving the full histograms to be fitted, a speedier approach is usually taken to significantly decrease the calibration time. The sweeping parameter is set to 0x00 or 0xFF (depending on the signal polarity) and all the bits are then flipped one after the other from most to least significant. At every step, if the occupancy is higher than 50%, the bit is flipped back to its original value. This bisection procedure allows, if the chip is behaving correctly, to get the same results in a much faster way.

With these chip parameters being set, pedestal, noise and common mode can then be measured using the same V_{CTH} sweep method. The nominal working point is usually set around five times the average noise above the pedestal value to be as efficient as possible with a low noise contamination.

Two latency parameters will then have to be set. The data (stub) latency, is the time difference between the read and write pointer in the CBC2 (read-out board stub buffer), in units of clock cycle. To set these values correctly, the parameter is scanned and set to the value corresponding to maximum hit or stub occupancy. While the data latency will always be positive by construction, the stub latency can be either positive or negative depending on the read-out architecture (cables and number of read-out boards).

A physical interpretation of the latency parameters can be reached by following the signals time propagation, starting at the instant when an incoming beam particle crosses a trigger scintillator. A schematic view of these different signals and propagation time is shown on figure 4.9.

On the first hand, the trigger signal will take a $t_1 + t_2$ time to be propagated to the internal CBC2 hit buffer, where t_1 is the time needed for the trigger signal to be generated and sent to the read-out electronics and t_2 the time for this signal to be propagated further to the read-out buffer. On the other hand, the hits corresponding to this trigger are expected to be stored in this same buffer after a $t_3 + t_4$ time needed for the particle to travel from the scintillators to the silicon sensors and for the corresponding hits to be

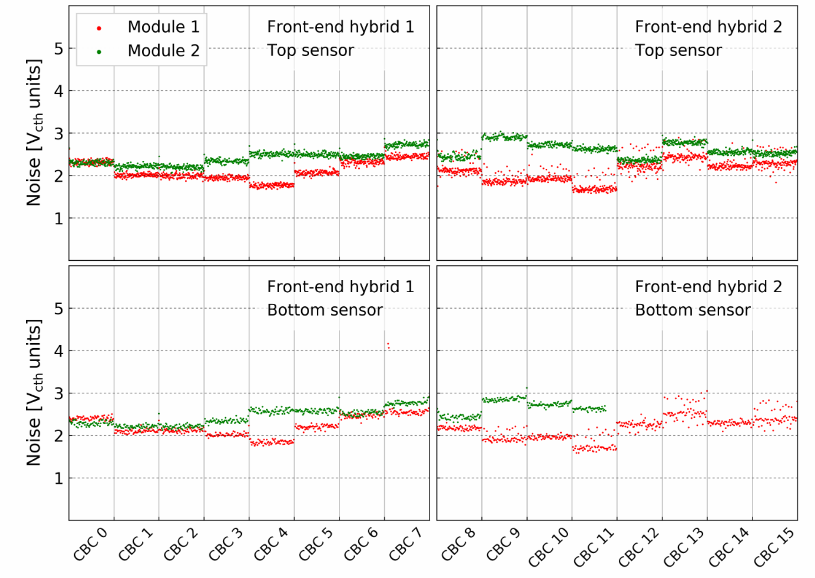


Figure 4.8. Noise measurement of two full-size CBC2 modules operating at room temperature and biased at 250 V and 400 V respectively [3].

generated and propagated to the buffer respectively. Calibrating the data latency aims at finding the difference $t_1 + t_2 - (t_3 + t_4)$.

Similarly, the trigger latency will be tuned to the time difference $t_1 + t_6 - (t_3 + t_5)$ where t_6 is the time for the trigger signal to be propagated within the read-out electronics and t_5 the time required for the stub bits from this event to be generated and propagated to the stub buffer.

From this description, it is clear that the data latency will always be positive but the same cannot be said for the trigger latency.

Once these calibration parameters are set (see summary table 4.2), the system is then ready to take data. A measurement typically consists of a scan of either a physical parameter (beam intensity, bending angle, tilt angle, bias voltage, and so on) or an internal parameter of the chip (such as threshold, or stub logic parameters) with data taking split into different runs, each with a constant set of parameters. A fast analysis of a first V_{CTH} scan allows a fine tuning of this parameter and a calibration cross-check.

Many different observables can be considered when analysing the data. Cluster, hit and stub efficiency and occupancy are usually the most relevant, but to highlight some specific behaviour, additional quantities are considered, such as cluster size, displacement between hits of the same or different sensors, hit residuals, reconstructed stub bit mismatch, to name a few.

The author of this document played an important role in the detector calibration and data taking operations, alongside his roles as shifter in the first time, then shift leader in the later beam tests. His involvement in fast data processing lead him to perform

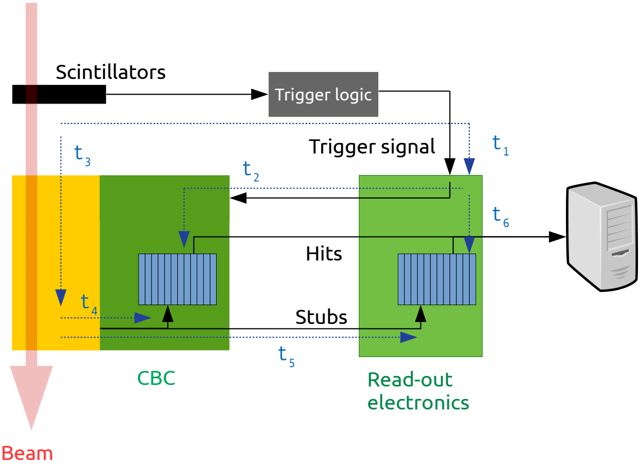


Figure 4.9. | Graphical representation of the hit, stub and trigger propagation time and their impact on latency calibration.

simplified data analysis in real time, allowing for a clear understanding of the behaviour and performances of the prototype during operations. This allowed for the discovery of numerous data taking issues and participated in shaping the measurement program.

4.1.5. Results

The legacy of this campaign has been very rich, and a comprehensive and detailed picture cannot be drawn in this document. Most of the efforts during the beam tests and data analysis were aimed at the development of the system itself. Both the firmware installed on the front-end boards and the software used to calibrate, record and process the data, have greatly improved following these beam tests. Errors arising from the prototype itself and the read-out chain have been relentlessly sought out leading to smoother and more stable data taking over time.

Given the importance of stub production in the upgraded system, a significant task of these beam tests has been to prove that this iteration of the chip was able to correctly produce stub-bits, paving the way for full stub information production in the next iteration.

Part of the effort consisted of validating the internal logic of the chip. Since for each event both the stub bits and the corresponding strip data is stored, it has been possible to emulate the stub-bit production and compare both data streams. While significant initial differences were observed (up to 7% of the stub-bits were wrongly assigned), the origin of this phenomenon was later understood to have been caused by some communication issues between the front-end and acquisition PC. Once this effect had been accounted for, both streams were found to be compatible up to the per million level. While this tiny

Parameter	Calibration target	Typical value
V_{PLUS}	50% CBC2 hit occupancy at V_{CTH} and offset target values	~0x75 (~500 mV)
Offsets	50% strip hit occupancy at target V_{CTH}	0x50 (~ 100 mV)
V_{CTH}	~5 times average noise above pedestal	0x72 ± 0xF (~ 0.5 - 0.6 V)
Trigger latency	Maximise hit occupancy in beam	~ 0x3 to 0xF
L1 latency	Maximise stub occupancy in beam	~ 0x0 to 0xF

Table 4.2. | Summary table of parameters to be calibrated before taking data, the observable used to calibrate these parameters and typical resulting values.

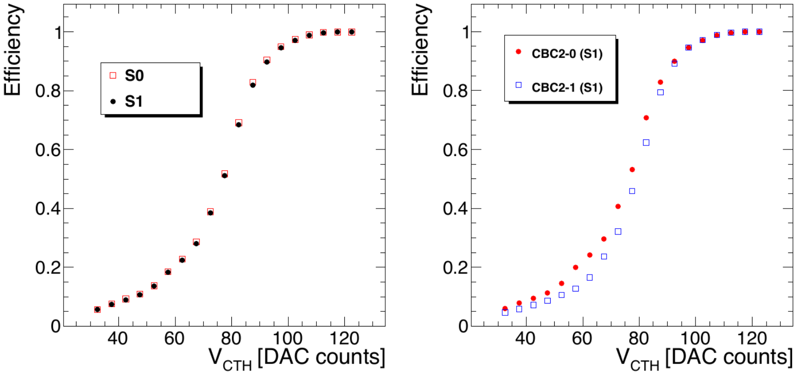


Figure 4.10. | Detection efficiency averaged over sensors and CBC2 as a function of comparator threshold for a perpendicular positron beam [110].

difference was not further investigated nor thoroughly explained, the impact of such a discrepancy can safely be considered as negligible in the final system.

Using tracks from the telescope matched to an in-time cluster in the reference plane, it is possible, after alignment, to extrapolate tracks to the module being tested. Comparing extrapolated tracks and hits in the device under test coupled with the production of a stub-bit, the stub-production efficiency can be measured, using a Bayesian method with flat prior assumption when required by the limited statistics.

In October 2015, a full-size prototype was tested in a secondary SPS beam line at the CERN north area. This module with two $10 \times 10\,\mathrm{cm}^2$ sensors featuring 2032 channels each were submitted to a 120 GeV pion beam. A uniform stub production efficiency of about 97.5% for impinging particles perpendicular to the sensor (therefore emulating high p_T particles in the final system) is shown on figure 4.11. Since the incoming beam, and more importantly the acceptance of the trigger, is much smaller than the full-size detector, this plot shows a composite image of multiple runs with the module being moved with respect to the beam. A lack of statistics on a particular spot of the detector explains the absence of measurements in part of the second CBC2.

While a second row of eight CBC2 was connected to the second half of the sensor, it was not possible to perform the same analysis due to issues in front-end firmware that have since been resolved.

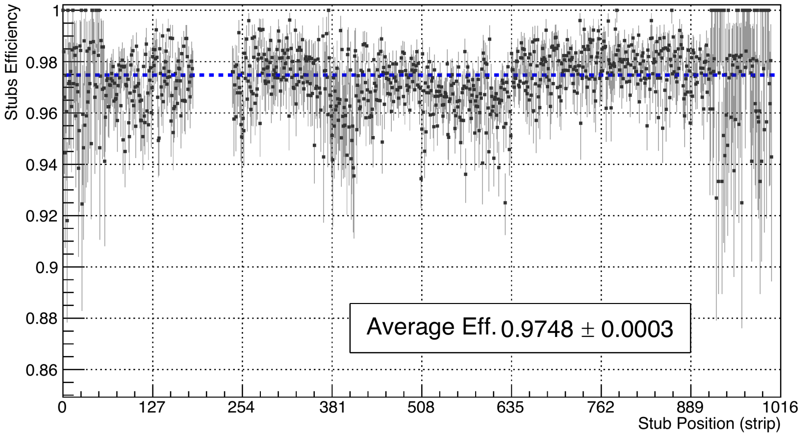


Figure 4.11. Stub reconstruction efficiency of an unirradiated 2S full-size module as a function of the stub position, for incoming particles perpendicular to the sensor. The vertical lines indicate the 127 strips belonging to each individual CBC2.

As radiation hardness plays such a crucial role in any HL-LHC detector design, a number of sensors and service hybrids were tested after irradiation. In particular, two similar mini-modules were compared. The unirradiated reference is referred to as CNM, after the name of its sensors' manufacturer. This module, with a sensor spacing of 2.75 mm is biased at −250 V. The second module had two baby-sensors (covering only one CBC2) produced by Hamamatsu Photonics K.K. (HPK). The sensors were irradiated to a fluence of $6 \cdot 10^{14} \mathrm{n}_{\mathrm{eq}}/\mathrm{cm}^2$, corresponding to twice the maximum irradiation expected for 2S modules of $3 \cdot 10^{14} \mathrm{n}_{\mathrm{eq}}/\mathrm{cm}$. These sensors were then bonded to unirradiated front-end hybrids.

Since the electric performance of these sensors was covered by dedicated measurements, the beam test focused on analysing the impact of radiation damage on Level-1 momentum discrimination. To do so, the prototypes were rotated within the beam to emulate particles of different transverse momentum. The stub reconstruction efficiency as a function of angle could then be measured, as seen on the left of figure 4.12. This angle can be converted into a transverse momentum in the final detector. Assuming the particles originated from a primary vertex and had a constant curvature from a uniform 3.8T magnetic field and negligible scattering and energy loss, the incoming angle with respect to the normal of a module placed at a radius R is given by

$$\sin(|\alpha|) \simeq \frac{0.57 \cdot R[\mathrm{m}]}{p_T[\mathrm{GeV/c}]}. \qquad (4.1.)$$

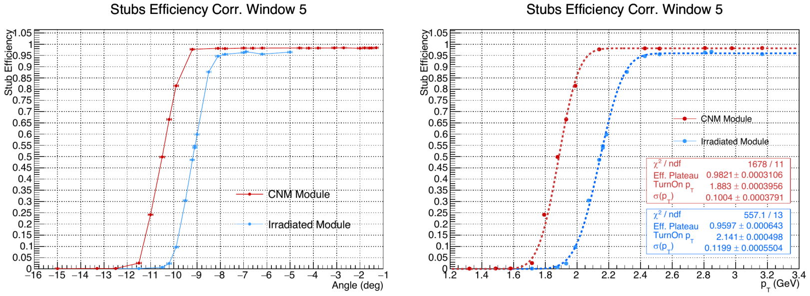


Figure 4.12. | Stub production efficiency for an unirradiated and an irradiated sensor as a function of the beam impinging angle (on the left) and as a function of the corresponding ideal particle transverse momentum if the module is placed at a radius of 60 cm (on the right).

By inverting this equation, the rotation angle can be converted into particle transverse momentum, as shown on the right of figure 4.12 for a module at a radius of 60 cm. These results were fitted with an s-curve. While this plot should not be over-interpreted given the very crude simplifications, some conclusions can be drawn.

Firstly, this result shows that the p_T module concept is indeed validated. While the efficiency plateau for low angles (high p_T) drops from 98% down to 96% after irradiation, it stays well within the requirements for Level-1 track trigger, as expected from sensor qualification and efficiency measurements.

The width of the fitted turn-on curve is of the order of 5% of the threshold, for both modules. Although this number cannot be directly interpreted as a p_T resolution given that the impact of the incoming beam divergence is non negligible and that the conversion from angle to momentum assumes no particle energy loss in the tracker, it is sharp enough to hint towards an efficient selection.

The threshold difference for the two modules is explained by a difference in sensor spacing. The sensors of the CNM modules were distant by roughly 2.6 mm compared to 3.4 mm for the irradiated module.

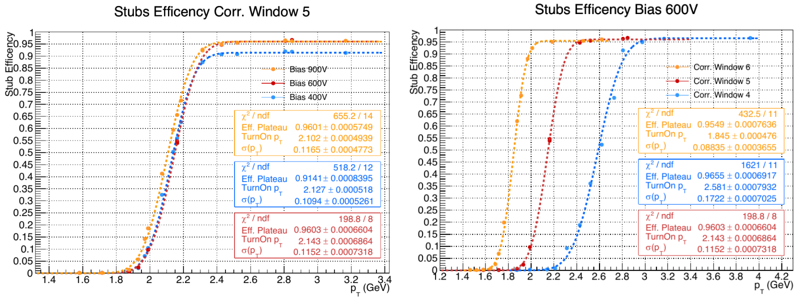


Figure 4.13. | Stub production efficiency as a function of ideal particle transverse momentum. The effect of sensor bias voltage is shown on the left and the effect of a change in the programmable correlation window is shown on the right.

Since the correlation window is programmable, the effect of the module position in the detector can be accounted for to obtain a uniform transverse momentum discrimination threshold (p_T^th). Assuming the same ideal particle trajectories described above, and single cluster hits, the corresponding correlation window width is given by

$$w = \frac{s}{p} \tan\left(\arcsin\left(\frac{0.57R\,[\mathrm{m}]}{p_\mathrm{T}^\mathrm{th}\,[\mathrm{GeV/c}]}\right)\right) - \frac{1}{2} \tag{4.2.}$$

where s is the sensor spacing and p the strip pitch. During beam tests, the sensor spacing is constant and a change in correlation window results in a change in p_T threshold. Assuming a small incident angle, the previous equation can be inverted as

$$p_\mathrm{T}^\mathrm{th}\,[\mathrm{GeV/c}] \sim \frac{s}{p}\frac{0.57R\,[\mathrm{m}]}{w+\frac{1}{2}}. \tag{4.3.}$$

The impact of this parameter on the turn-on curve is shown on figure 4.13, with an expected increase in threshold as the correlation window is tightened. With a 100 µm pitch and a 2.6 mm spacing, the expected p_T thresholds at a radius of 60 cm are 2.6, 2.1 and 1.8 GeV/c for correlation window widths of 4, 5 and 6 respectively. These estimated values are confirmed by the fitted s-curves.

A challenging aspect when dealing with an irradiated sensor is to maintain the ability of depleting it. Damage to the silicon lattice results in an increase in dark current and a change in depletion voltage due to an increase in acceptor levels. Standalone studies that had shown that the sensors should be adequate even after radiation were confirmed during the test beam, as shown on the left of figure 4.13. The effect of bias voltage on the plateau efficiency can clearly be seen, and applying a 600 V difference of potential is enough to ensure a fully efficient module, as expected from sensor measurements [3].

The author was very active in the prototype data analysis aspects through technical support related to the usage of the data files used, personal analysis contributions, the overall reflection surrounding the characterisation of the prototypes and some involvement in the redaction of publications.

As can be seen on this small sub-selection of plots, this test beam campaign has shown that after thorough development and debugging, the CBC2 prototypes and read-out chain were up to the challenges faced for the phase-2 upgrade of the CMS tracker. While these results were sufficient to validate this iteration of the read-out chip and pave the way for future prototypes, multiple smaller scale beam tests were organised to further explore the detector's properties. One such experiment, organised in Louvain, will be the topic of the next section.

4.1.6. The Louvain low energy proton beam test

In 2016, with increasing luminosity being reached by the LHC, the attention of the CMS tracker community was focused on luminosity impact on the detector, in particular in the case of heavily-ionising particle events. While the upgraded detector design should naturally mitigate the impact of such events, a beam test was organised at the Light Ion Facility of the Centre de Ressources du Cyclotron (CRC) in Louvain-La-Neuve in order to have an experimental cross-check.

At these facilities, the proton beam is accelerated to a fixed energy of 65 MeV by a

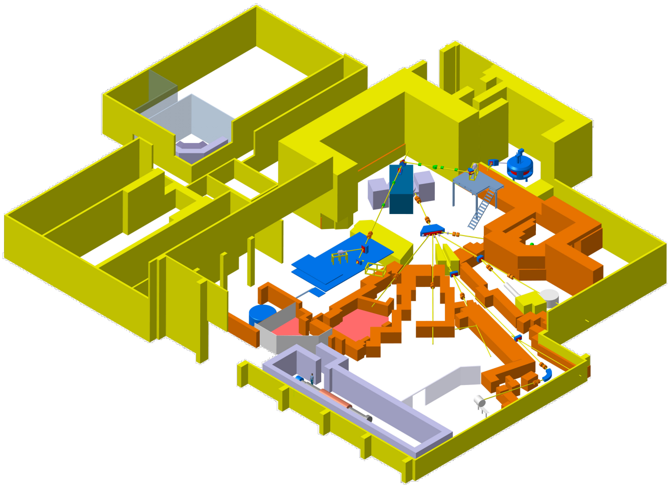


Figure 4.14. | Layout of the UCLouvain beam facilities. The CYCLONE120 cyclotron can be seen on the top-left and the light ion beam area and counting room used for the beam test are coloured in red [111].

cyclotron. The beam is delivered to the experimental area shown on figure 4.15. It will first pass through a lead diffuser to increase the beam divergence from the initial 0.8 cm. After a 2.5 m drift, the beam will exit the vacuum chamber through a 50 µm steel window. In the experimental area, degradors of various depth can be added to decrease the beam energy. After passing a rectangular collimator, incoming particles will pass through the device under test surrounded by two thin scintillators that provide the trigger signal.

The device under test during this experiment was a 2S mini-module with two unirradiated p-in-n sensors. One of the CBC2 was faulty and was therefore disabled during the beam test. While globally similar to other beam tests, the control system and data acquisition had some particularities.

Firstly, given the high radiation environment in the experimental area, as many functionalities as possible were moved to the counting room, resulting in a system split into two, as shown on figure 4.16.

The device under test and trigger detectors were powered and biased by remote controlled low and high voltage sources in the experimental area. While the data from the 2S module could be sent directly to the control room, the analogue signal from the photomultipliers both before and after being pre-amplified were too faint and short to be transported by 10 m LEMO cables and therefore had to be digitised directly using discriminators.

These discriminated signals were then sent to the control room to be converted into a trigger signal. Two separated data paths could be used, either using basic discrimination

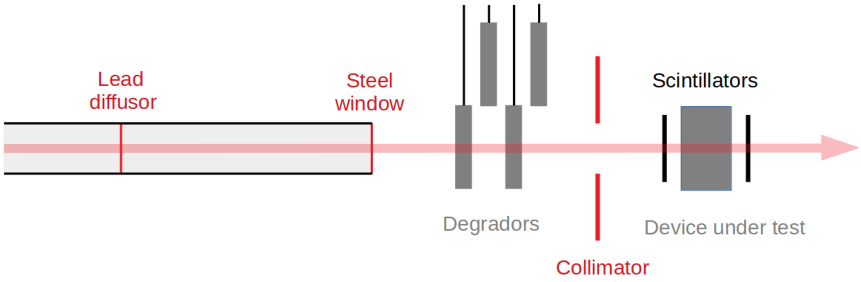


Figure 4.15. | Sketch of the experimental area layout at the CRC facilities. The proton beam is diffused by a thin lead foil and their energy is reduced by any combination of four degradors. The beam is then collimated before impinging on the triggering scintillators and the device under test.

and coincidence logic implemented in various NIM cards or a single controllable VME card[1]. Both these trigger signals could be prescaled to throttle the trigger rate to a given value and could be further throttled by busy systems. Two timing and trigger control cards[2] were also used to select the desired trigger source, start and stop the runs and encode the trigger signal into a 25ns clock sent to the DUT DAQ. Furthermore, scintillator counts, trigger rates, voltages and currents were among the variables that were continuously monitored and stored offline as environment data.

The 2S module was read-out directly by a single GLIB card [116] in the control room connected to the acquisition computer. The unpacking and processing of the data was done offline.

In parallel to the primary goal of building up expertise in order to eventually construct a tracker endcap in Belgium, this beam test enabled original measurements to be recorded. Given the nature of the beam used, it was possible to test the module with very ionising particles and over a wide range of incoming fluxes.

As in the case of other beam tests of the campaign, one of the first steps after calibration was to perform a threshold scan. While the ideal variable to consider is the module efficiency, this beam line is not equipped with a tracking device and other observables had to be used.

Two different values can be used as a proxy for efficiency. Assuming a low sensor noise and that both sensors are independent, the correlation between the two sensors can be used. For a given hit in the reference sensor, a window is defined on the sensor under test. Using this correlation method, the efficiency can be defined as the ratio of matched hits with expected hit. This is shown in red on the plot to the left of figure 4.17. The second method that can be used as a proxy for efficiency, disregarding the assumption that the sensors are independent, is the proportion of non-empty events. While the normalisation of the plot is to be corrected by a geometrical factor taking into account the ratio between detector and trigger acceptance, this correction is independent of the

---

[1] CAEN MOD. V812, 16 channel constant fraction discriminator [112]

[2] TTC VX led encoder [113] and TTC VI mark II VMEbus interface [114]. Their usage in the context of the LHC is described here [115].

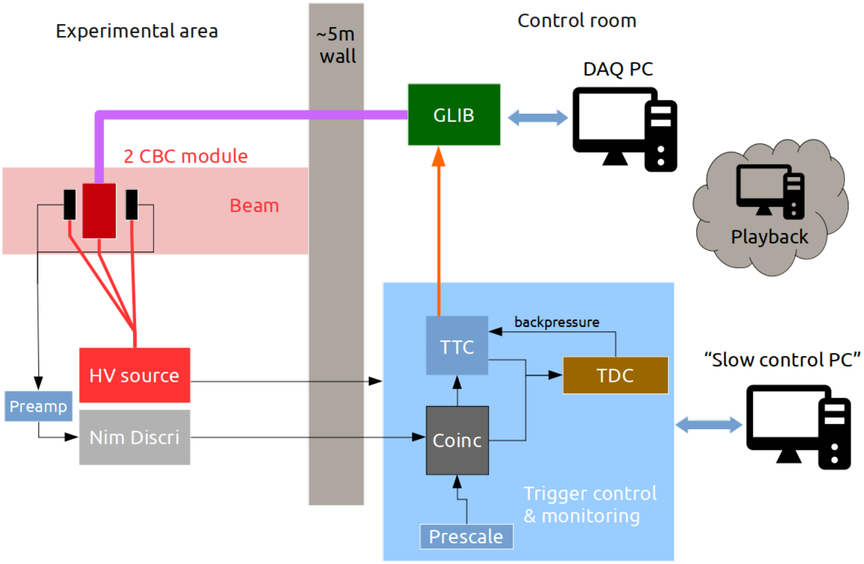


Figure 4.16. Simplified overview of the control and data acquisition system.

threshold used. This second efficiency behaviour estimation is shown in blue on the same plot.

By comparing the two trends, it is clear that both measurements have a very different behaviour. While no definitive answer can be given for this discrepancy, after considering multiple possibilities and making comparisons to simulations, a possible explanation could be some cross-talk between adjacent channels on the read-out chip. This behaviour, however, would be unexpected since the chip had been tested by injecting a significant charge into the analogue circuit and yielded much lower cross-talk levels. However, the conditions were different during the beam test with the read-out chip being used under extreme circumstances. Further measurements would be necessary in order to resolve this tension, for example by performing the same test but with a high beam crossing angle. While this discrepancy remains unresolved, it is not considered to be of significant effect in the final setup.

The impact of instantaneous luminosity on the prototype was then tested. The beam flux was varied over multiple orders of magnitude with a fixed trigger rate, and the response of the detector was monitored. The plot on the right of figure 4.17 shows three variables of interest. The average number of clusters is shown to be increasing once the flux rises above $\sim 2 \cdot 10^5 \mathrm{Hz cm}^{-2}$. Looking specifically at the 1-strip clusters shows that they are the main contribution to this additional signal, while staying relatively small at lower fluxes. This could be explained by some out of time signal where only the strip with the highest energy deposit is above threshold.

Finally, the sensor leakage current is shown on the same plot. The low flux value is constant around 4.9 µA before rising proportionally with particle rate, probably due to the signal induced in the silicon. Both increase in leakage current and in cluster

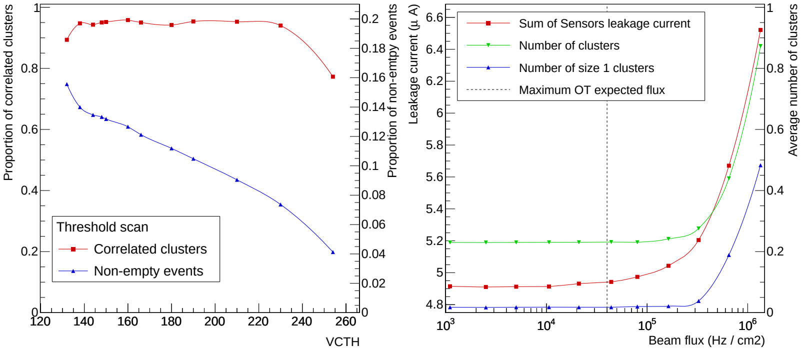


Figure 4.17. | Threshold scan on the left showing the difference in behaviour between the proportion of correlated clusters and the proportion of non-empty events in a given sensor. The impact of incoming proton fluxes on data-taking is shown on the right with an increase in leakage current (in red) and in number of clusters (in green), mostly coming from clusters of a single strip (in blue).

occupancy seem to be occurring at particle rates significantly higher than the maximum levels expected at CMS.

While this result seems to validate the rate tolerance of the prototype, the very specific conditions under which this test was performed must be kept in mind. The beam structure is very different from the bunched 25 ns beam at the LHC, and the trigger used in this system could very well be susceptible to particle rate.

Given the context of the time, a measurement of interest was to check the behaviour of the prototype under very ionising particles. In other terms, another goal was to measure the response of the detector as a function of energy of the incoming beam.

This was achieved by using different combinations of the available degradors, and measuring the proportion of non-empty events on either sensor 1 or sensor 2. This is shown on the left of the figure 4.18 in red. This result could not be interpreted at first glance and a Monte-Carlo simulation of the setup was therefore required to try to draw some conclusions. A Geant4 simulation performed in collaboration with a master's student [117] is shown in blue on the same plot. While the general behaviour was very similar between data and Monte-Carlo, two parameters were fine-tuned to obtain the result shown here. A scale-factor was added to match data and simulation for low absorber thickness and a shift of 2.5 mm of absorber depth was also introduced to match the peak and compensate possible mismodeling of any material composition or depth placed upstream of the sensors.

After this fine-tuning, the simulation reproduces nicely the features observed in the data. Furthermore, it gives an estimate of the energy of the incoming protons, as shown on the right of figure 4.18. The points correspond to the most probable value of the energy while the area designates the full width at half maximum. The levelling around 5 MeV before a sudden drop comes from the requirement that the proton triggers the downstream scintillator to record the event.

The increase in number of recorded events at an absorber thickness of 20 mm corre-

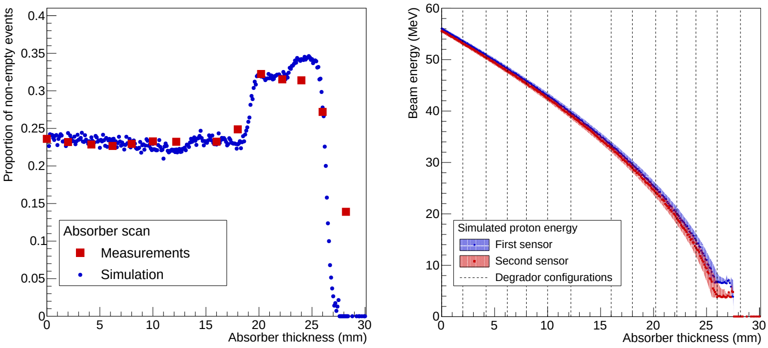


Figure 4.18. | The impact of the degrador depth on the number of events with at least one cluster on any sensor is compared with simulation on the left. On the right, the simulated proton energy on the two sensors as a function of degrador depth is shown.

sponding to a beam energy of roughly 25 MeV could therefore be understood not as a detector effect, but linked to the condition under which the data has been recorded. Indeed, the module was not placed directly in the beam area, but in a protective and light-shielding case. The case was built out of an aluminium box, with two windows carved out on both sides of the sensor, covered in tape. Given that the collimated beam is much bigger than both the sensor and the carved windows, the material budget is highly non-uniform.

With a reduced beam energy, a large fraction of the incoming protons will be stopped in the aluminium case, upstream of the second scintillator. This will result in effectively focusing the beam for recorded events on the sensor. For extremely low energies, the silicon itself will be enough to significantly reduce the energy of the protons and prevent them of triggering the read-out system, hence the final drop in non-empty events.

Simulating the full setup allowed understanding the observed behaviour of the recorded data, has also proven impossible to precisely measure the impact of low energy particles on the prototype. These results are nevertheless showing that no clear degradation can be observed.

All in all, this beam test was a success. On the one hand, the comparatively small team of about ten scientists involved was able to set up a fully functional system, bringing valuable expertise to the Belgian phase 2 tracker community. On the other hand, it was possible to provide first results on such a prototype under unique operating conditions that, even if not quantitative, bring in some valuable qualitative information, and a fine addition to the overall CBC2 measurement campaign.

The author had a predominant role in this beam test. Besides the usual data taking and processing involvement, he contributed significantly to the organisation of the beam test, the development of the trigger system and the slow control interface. Finally, he was the main contributor to the offline data analysis efforts.

4.2. CBC3 design and test campaign

Building on the success of a CBC2 prototype design that was shown to appropriately meet the challenges presented by the high-luminosity LHC, a third version of the chip was designed to provide all the features required to operate in the final system.

In this section, the specificities of this later version will be outlined and the performance measured in-beam will be discussed.

4.2.1. Changes with respect to version 2

The third version of the read-out chip is very similar to the second version, so only the most relevant differences will be described here. A more detailed description can be found in the user manual [118].

The analogue part of this chip is mostly untouched with respect to the previous version. Two changes, however, are important to mention here. The first is that, while the CBC2 was aimed at testing, the CBC3 was envisaged as a fully functional chip. With that optic in mind, the holes read-out mode was removed given that the choice of sensor polarity had been made.

The second change in the analogue part of the chip is an improvement in the threshold granularity. The newer chip has a comparator threshold encoded over 10bits, compared with 8 bits for the older one.

It is in the digital part of the chip that most of the changes occurred. Firstly, a different hit detection logic step was installed at the output of the comparator where any of four different logics can be used. The fixed pulse width method will produce a hit during 25 nanoseconds following any over-threshold signal longer than about 300 ps. In the 40 MHz sampled mode, a hit is produced if the comparator signal is over-threshold during the rising edge of the clock. A logic OR of the two previous methods is available. In HIP-suppressed output mode, the number of consecutive bunch crossings where the sampled mode produces a hit will be counted and a veto signal is produced if it goes above a programmable limit. This veto can then be applied either to the OR logic or the sampled mode to provide the last hit detection logic.

Besides the hit detection improvement, radical changes can be seen in the stub production logic. An option was added to select which sensor will act as a seed layer and as a correlation layer. This was needed given the asymmetrical nature of stub production and the module arrangement in the tracker.

Unlike the CBC2 that only produced stub-bits, this version of the chip will provide the stub information to the data acquisition system. While the logic to produce the stubs is unchanged, the chip now computes an eight bit address and a five bit bend for every stub, corresponding respectively to the position of the seeding cluster and the difference of position between seeding and correlated clusters, in half-strip units.

Given the limited bandwidth available, the five bit bend is then reduced to four bits. This is done through mapping in a programmable Look-Up Table. While the precision is inevitably diminished when the correlation window is wide, it allows for the full half-strip precision for the 2S modules located in the innermost regions.

This information needing to be produced synchronously, only a fixed amount of stubs can be read-out per bunch crossing. This limit has been set to three, meaning that however busy an event is for the read-out chip, only the first three stubs sorted by address will be produced.

The final change to be mentioned is the addition of a Delay Locked Loop. Both the read-out chip and the LHC will share the same 40 MHz clock, but since the particles from the interaction point will take some time to reach the detector, and the length of the optic fibres transporting this clock varies from module to module, it is necessary to be able to finely delay the clock in the read-out chip. This enables an in-time signal in the same clock phase to be obtained, thus preventing any timing inefficiencies. This phase shifting can be done with the granularity of one nanosecond.

4.2.2. CBC3 mini-module beam test

The first beam measurements on a double sensor CBC3 mini-module occurred at the Fermilab test-beam facility in November of 2017. A bunched beam of 120 GeV protons was used to characterise a CBC3 module placed in the centre of six CMS pixel detectors providing tracking information.

The device under test and tracking detectors were given a clock signal slightly different from the usual 40 MHz allowing synchronisation with the incoming beam. A downstream scintillator provided the trigger signal.

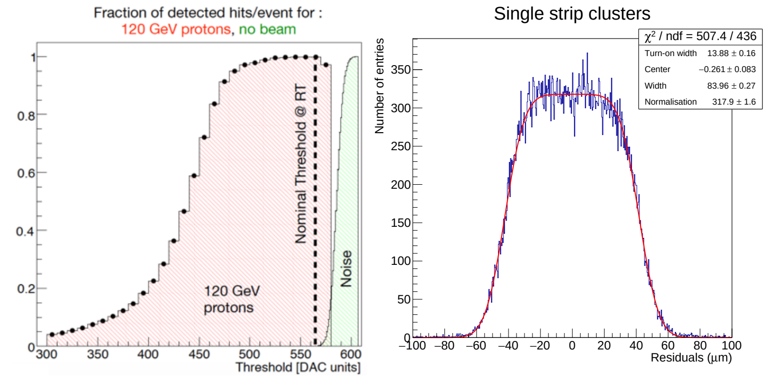


Figure 4.19. The figure on the left [119] shows the fraction of events with a detected hit as a function of comparator threshold either with beam (red), or with no beam and using random triggers (green). The nominal thresold corresponds roughly to 4000e$^-$ above the pedestal. The figure on the right shows the difference in extrapolated track position on the device under test and the centre of detected single hit clusters. The distribution is fitted with a square function convoluted with a Gaussian.

A procedure similar to the one described previously in section 4.1.4 was used to calibrate the prototype and the nominal operation parameters could be determined in order to minimise noise while keeping a high efficiency, as shown on figure 4.19.

Tracks provided by the telescope system [120] were aligned with the device under test by minimising the residuals after application of a translation and three rotation matrices. The resulting residuals for single hits in the sensors are shown on the right of figure 4.19

and are well described by a step function convoluted with a Gaussian to account for the telescope resolution. The pitch measured is compatible with expectations.

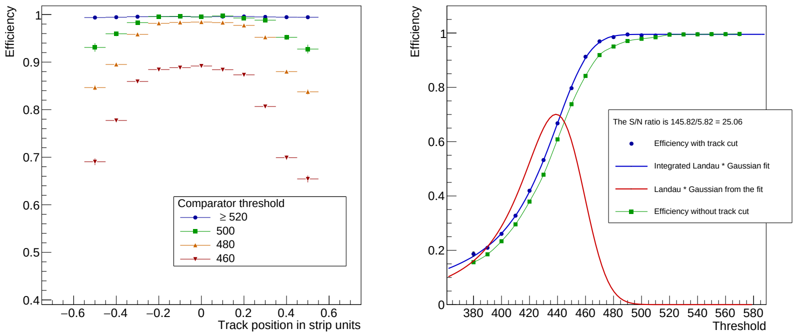


Figure 4.20. | The figure on the left shows the hit detection efficiency for different comparator thresholds as a function of the difference between the extrapolated track position and the nearest strip centre, converted into strip pitch units. The figure on the right displays the hit efficiency as a function of threshold. The green squares are measured without cuts on the track parameters, while only the centre of each strip is considered in blue.

This alignment allowed for a precise efficiency measurement throughout the detector. The plot on the left of figure 4.20 shows the hit efficiency as a function of the difference in position between the aligned track and the closest strip implant for different threshold values. At higher thresholds, the hit efficiency is stable along the strip above 99.5%. When the threshold is decreased (therefore requiring a higher signal to produce a hit), the efficiency gradually decrease, starting with the region furthest from the strip centre. This can be easily explained by an increase in charge sharing between two adjacent strips.

Since the detector has a binary readout, the charge deposited on the sensor isn't directly recorded. In order to measure this parameter, a threshold scan has to be performed. Since all strips are measured independently, the cluster charge cannot be reconstructed and charge sharing would bias our measurement. To mitigate this effect, a cut on the track position with respect to the closest strip implant is applied, selecting only the innermost 10% of the strip.

Using the simplest model of a minimum ionising particle in a detector, the expected charge deposition distribution in the sensor would be described by a Landau function. Given stochastic effects in the silicon, charge collection efficiency and electronic noise, this Landau is expected to be convoluted with a Gaussian function with a mean of zero.

Since we do not measure the charge, but the fraction of events with a signal induced above threshold, assuming a linear comparator threshold and pulse width as a function of charge, this simplified model would predict an efficiency with the following distribution

$$\varepsilon(\mathrm{VCTH}) = N \int_0^{\mathrm{VCTH}} \int_{-\infty}^{+\infty} \mathrm{Landau}(x+y;\mu,c)\cdot\mathrm{Gaus}(y;\mu=0,\sigma)dydx \tag{4.4.}$$

where N is a normalisation parameter, μ and c are the location and scale parameter of the Landau and σ is the standard deviation of the Gaussian function used in the convolution.

The plot on the right of figure 4.20 shows the measured efficiency in the innermost 10% of the strips fitted with this function and the corresponding convoluted Landau function describing the charge measured. From this latter distribution,the most probable value is extracted. The pedestal and noise, both of which measured during the calibration of the detector are then used to yield a signal-to-noise ratio[1] of $S/N = 146/5.8 \sim 25$. The Landau scale paramater c and Gaussian standard deviation σ obtained from the fit are of 10 and 12 V_{CTH} units respectively, and in line with the overall expectations for these detectors.

This result is only valid as long as the particles arrive close enough to the strip implant. The effect of charge sharing can clearly be seen by relaxing the track parameter requirements.

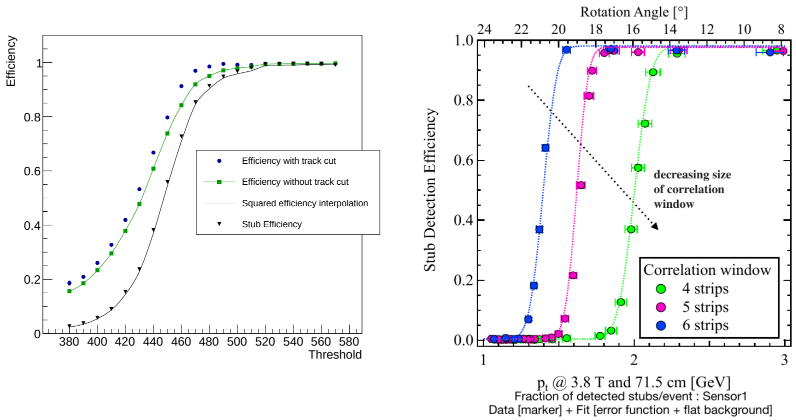


Figure 4.21. | The figure on the left compares the strip efficiency with the stub production efficiency as a function of comparator threshold. On the right [121], the stub production efficiency is shown as a function of corresponding particle angle for different correlation window widths.

In this third version of the chip, it was possible for the first time to read-out and verify the behaviour of stubs. During electrical tests before the beam test, an anomaly was discovered. It has been shown that some of the data lines linking neighbouring CBC3 chips necessary to produce stubs from tracks hitting both sensors in different chips were inverted. This issue in the chip, leading to a loss in stub efficiency for the corresponding channels, was easily identified and was corrected in a future iteration of the chip [122].

By excluding these strips, it has been shown that the behaviour of the stub production and look-up table works as planned. The stub efficiency as a function of comparator threshold for tracks perpendicular to the sensors is shown on the left of figure 4.21. This efficiency, expected for high-momentum particles, follows the single sensor efficiency

---

[1] With the numerator and denominator expressed in arbitrary V_{CTH} units.

squared, as would be assumed for independent sensors, and arrives just below 99.5% for high threshold values (corresponding to a small charge cut).

In varying the beam angle with respect to the sensor, thus emulating different transverse momentum, the stub turn-on efficiency could be measured, as shown on the right of figure 4.21.

This result, combined with the verification that the correct stubs were indeed produced, proves that the prototype is indeed able to discriminate tracks based on local transverse momentum, thus validating the design choice for the upgrade.

4.3. Towards the final detector

Since the beam test in late 2017, the last one in which the author provided a major contribution, a number of significant milestones have been passed. The same mini-module was tested again in May 2018 after an irradiation to a fluence of $2 \cdot 10^{14} \mathrm{n_{eq}/cm^2}$, then again in December of the same year after being submitted to an additional $2.4 \cdot 10^{14} \mathrm{n_{eq}/cm^2}$. In both cases, the mini-module was shown to be capable of being appropriately biased and to operate efficiently, as expected.

Significant advances also occurred in prototyping, with the minor issues found in the CBC3 being corrected in the latest CBC3.1 version of the chip.

In addition, two versions of the Concentrator Integrated Circuit (CIC) have been produced since then. In the context of 2S modules, the goal of this chip, as briefly mentioned in section 4.1.1, is to aggregate the data from eight CBC chips. More specifically, the chips are designed to gather the data from the CBCs, synchronise and sparsify it and build full events. These events are then buffered and serialised, ready to be sent optically to the read-out system. In parallel, this chip is also in charge of gathering stubs and selecting the candidates with highest bend to be read out.

A first version of this chip [123] was produced at the end of 2018 followed by a second version [124] one year later that offered a better radiation tolerance through memory triplication and the possibility of increasing the chip frequency that is required for the read out of some high-occupancy PS modules. These chips were confirmed to be working as expected through test-bench measurements, and a single-event upset test at the UCLouvain Heavy Ion Facilities is planned for the near future[1].

In parallel, significant developments on the service hybrid side in the last few years lead to the production of fully functional prototypes. Hosting low voltage and high voltage input cable pairs, DC-DC converters, a VTRx+ optoelectronic transceiver, a gigabit transceiver chip and a slow control adapter, this hybrid combined with two 8CBC3.1 hybrids with CIC1 mezzanines enabled the production of full 2S module prototypes in 2019. A photograph of such a module can be seen on figure 4.22.

The ability to operate such prototypes in real conditions was proven in a beam test performed at DESY in November 2019 . The analysis of this beam test data is currently ongoing and preliminary results seem to indicate that the prototype performance is in line with expectations.

---

[1] This test is currently planned for the end of June if the evolution of the coronavirus pandemic allows it.

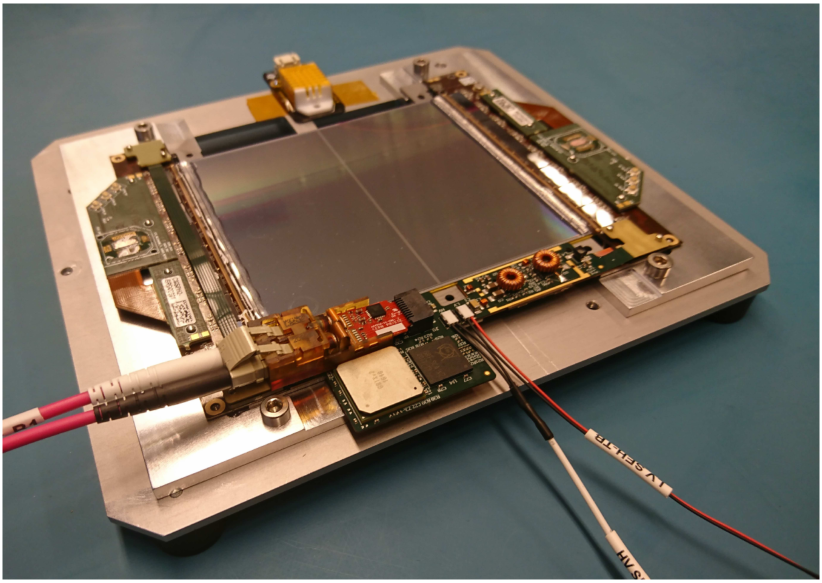


Figure 4.22. | Fully functional 2S module prototype tested in the DESY November 2019 beam test [125].

In parallel to the major advances in detector design, prototyping and validation described in these two chapters, the collaboration is also preparing itself to construct this detector. While private contractors that would provide the basic components are being identified, the tasks of testing and assembly will fall on the collaboration institutes.

In this context, the UCLouvain university will play a central role in the assembly of one of the tracker endcaps. A clean room and the relevant services are being set up in order to integrate modules, mechanical support structures and services into twelve dees. According to the baseline design, the dees will need to be heated to 60° to cure the glue and tested electrically and optically both at room temperature and cooled down to about −35°. Finally, the dees will be assembled into a fully integrated endcap ready to be sent to CERN. The preparation of these installations is well advanced and the institute is readying itself for a pre-production planned to start in 2021.

Chapter

5. Double Higgs discovery opportunities at the HL-LHC

In the previous chapters, the design choices for the upgrade of the CMS detector were described, and experimental results and proof of concept for the upgrade of the outer tracker were presented. Similar work has been performed by the collaboration on other sub-detectors, and it also seems to indicate that the upgrade plans are feasible and that the prototypes live up to expectations.

It is one thing to show that prototypes do meet the requirements linked with the design of the new detector, but it is another to show that the detector design as a whole will be able to cope with the harsh conditions of the upgrade and thus help extend understanding of particle physics.

To validate these choices, for CMS and also for other detectors, a series of prospective analyses or projections were performed for a multitude of physics cases, simulating the upgraded detector and computing expected limits or significance. These studies were compiled in the so-called HL-LHC Yellow-Report [126] document.

In the present document, the particular case of the Higgs boson pair production (or di-Higgs, for short) measurement with the CMS detector as documented here [6] will be discussed, with an emphasis on the $b\bar{b}l\nu l\nu$ final state. This channel was studied by a team from the UCLouvain university, with the author of this thesis as main developper. While the core framework and analysis strategy used were inherited from the published run 2 analysis, the author adapted the analysis to process the simulated samples and its strategy to be in line with the HL-LHC operation assumptions.

5.1. Higgs boson pair production at the HL-LHC

As stated in chapter 1, the direct measurement of di-Higgs production is a crucial, although challenging, test of the standard model and is sensitive to numerous theories beyond the standard model.

The dominant di-Higgs production mechanism at $\sqrt{s} = 14\,\text{TeV}$ is gluon fusion and the increase in center of mass energy is expected to increase the total cross-section to $36.69^{+2.1\%}_{-4.9\%}\text{fb}$. This value was computed at the next-to-next-to-leading order (NNLO) of the perturbative quantum chromodynamics (QCD) calculation, including next-to-next-to-leading-logarithm (NNLL) corrections and finite top quark mass effects [127]. Given the scale of this cross-section, an important integrated luminosity will be required. Even with the LHC and detector upgrades, this measurement should remain challenging and multiple analyses focusing on different Higgs boson decay channels will be needed.

Since most Higgs bosons decay into a pair of b quarks, the four final states expected to be the most sensitive to the standard model based on previous experience are $\text{b}\bar{\text{b}}\gamma\gamma$, $\text{b}\bar{\text{b}}\tau\tau$,

$b\bar{b}b\bar{b}$ and $b\bar{b}WW$ ($WW \to \ell\nu\ell\nu$ with $\ell,\ell = e,\mu$). The very rare but clean $b\bar{b}ZZ(4\ell)$ with $\ell = e,\mu$ will also be considered. Since none of these analyses alone are expected to be sensitive to the standard model, they are all designed to be orthogonal to allow for a crucial combination.

With the exception of $b\bar{b}ZZ(4\ell)$, the four other final states were covered by dedicated run 2 analyses. The design of the prospective analyses relies on the expertise gained from previous work and is quite similar. Notable differences between the two versions of these searches are the absence of full simulation for background processes, the absence of data-driven background modelling and in some cases, the extended use of machine learning functionalities.

5.1.1. Samples

The signal and background hard processes, fragmentation and hadronization were simulated using Monte-Carlo event generators.

Di-Higgs signals were generated using MadGraph5_amc@nlo at leading order, in the five considered final states. The branching fractions shown in table 5.1 were computed assuming the standard model and were used to normalise the generated samples. For anomalous values of λ_{HHH}, the samples were reweighted on an event-by-event basis as a function of the parton level di-Higgs invariant mass and the cosine of the angle between either boson and the beamline in the di-Higgs reference system. The full procedure is detailed in Ref [128].

Top quark backgrounds are simulated using powheg [129–132]. The $t\bar{t}$ samples are normalised to a theoretical cross section at $\sqrt{s} = 14$ TeV of $984.50^{+23.21}_{-34.69}(\text{scale})\,^{+41.31}_{-41.31}(\text{PDF}+\alpha_S)\,^{+27.14}_{-26.29}(\text{mass})$ fb^{-1} computed at NNLO+NNLL [133]. The cross section of single top quark production in the tW channel was computed at NNLO precision to be of $84.4 \pm 2.0(\text{scale})^{+3.00}_{-4.80}(\text{PDF})$ fb^{-1} [134, 135].

Drell-Yan lepton pair production through the $Z/\gamma^* \to \ell\ell$ and $W \to \ell\nu$ processes are generated at the LO precision using MadGraph5_amc@nlo. This production is separated in samples binned by dilepton invariant mass and sum of the transverse parton momentum emitted to increase the number of events passing the selection criteria of the different analyses. The samples are normalised to the generator cross section.

Channel	$\mathcal{B}$ [%]	Number of events
$b\bar{b}b\bar{b}$	33.6	37000
$b\bar{b}\tau\tau$	7.3	8000
$b\bar{b}WW(\ell\nu\ell\nu)$	1.7	1830
$b\bar{b}\gamma\gamma$	0.26	290
$b\bar{b}ZZ(\ell\ell\ell\ell)$	0.015	17

Table 5.1. The five decay channels described in this document, with the corresponding branching fraction and number of events produced at a centre of mass energy of $\sqrt{s} = 14$ TeV and an integrated luminosity of 3000 fb^{-1}. The symbol ℓ represents an electron or a muon. In the $b\bar{b}WW$ channel, the leptons in the final state can originate from the decay of a tau lepton.

QCD samples for multijet backgrounds were produced at leading-order accuracy using MadGraph5_amc@nlo and normalised by the generator cross section scaled by the ratio between run 2 data and simulation at $\sqrt{s}$ =13 TeV using the same generator. To increase the signal yield in the $b\bar{b}b\bar{b}$ final state, samples were produced requiring at least one b quark produced at parton level. The samples are binned in scalar sum of the transverse momentum at matrix element and weighed accordingly. An additional inclusive sample was produced requiring parton pairs with an invariant mass above 1 TeV in order to describe the highly boosted region of the phase space.

Lastly, single Higgs boson backgrounds were generated from gluon fusion ($gg \to H$), vector boson fusion (VBF), in association with a top quark pair ($t\bar{t}H$) and with vector bosons (VH). Different generators were used depending on the Higgs boson decay channel. The MadGraph5_amc@nlo generator is used for Higgs bosons decaying into $ZZ^* \to \ell\ell\ell\ell$ and into two photons for the $gg \to H$, VBF and VH processes. The powheg generator is used for all other processes.

These samples are normalised to their respective standard model cross section as recommended in the handbook of LHC Higgs cross sections [136]. The $gg \to H$ cross-section is computed at next-to-next-to-next-to-leading order (N^3LO) in QCD and at next-to-leading order in electroweak corrections. The VBF and WH production cross sections are computed at NNLO and NLO accuracies in pertubative QCD and EW corrections respectively. The $t\bar{t}H$ cross section is computed at NLO QCD and NLO EW accuracy while the $qq \to ZH$ cross section is computed at NNLO QCD and NLO EW and the $\gamma \to ZH$ contribution is computed at NLO QCD with NLL effects.

The design of the detector upgrade and experimental results on different prototypes were used to elaborate a model of the detector using the Geant4 software [70]. This enabled the reconstruction performance and detector efficiency to be studied under the high-luminosity conditions. These results were used to tune a Delphes model [71] of the detector.

The Monte-Carlo generated samples were superimposed with an average of 200 minimum bias collisions to account for pile-up effects and interfaced by the tuned Delphes model to produce the final data simulation used in these different analyses.

5.1.2. Systematic uncertainties

To perform prospective analyses, some assumptions are required both on theoretical and on experimental uncertainties.

It is assumed that progress will be made in the coming years in the modelisation of all signal and background processes. The uncertainties on all simulated cross-sections considered here are scaled by a factor $1/2$ with respect to the current theoretical uncertainties. Furthermore, while the sample sizes used here are limited, it is assumed that the statistical uncertainty on these samples will become negligible when the whole dataset is analysed. The statistical uncertainty on the Monte-Carlo simulations is therefore not taken into account.

Following the publication of the technical design reports for the phase-2 upgrade of the different sub-detectors, the expected performance of the physics objects with the upgraded CMS detector could be derived from full simulation, as described in the following detector note [137]. This work, enabled the following assumptions on the experimental systematic uncertainties to be made.

The jet energy scale uncertainty was found to be of 0.2% to 2% depending on the source, and was propagated to each jet individually before applying any selection. The impact of these changes was also propagated to the missing transverse energy and other high-level variables. The b tagging efficiency was derived as a function of η and p_{T}. For jets with $p_{\mathrm{T}} < 300$ GeV, an uncertainty of 1% was found, increasing to 2 - 6 % for larger transverse momentums.

The uncertainty in electron identification and isolation efficiency were taken to be of 2.5% and 0.5% for particles below and above 20 GeV respectively. The equivalent uncertainty for muons was 0.5% and 5% for hadronically decaying τ leptons. The uncertainty considered on photon identification efficiency is of 1%, on energy scale of 0.5% and on resolution of 5%.

While it is difficult to predict what the exact triggering criterion will be at the HL-LHC, it is assumed that the phase 2 upgrade will maintain thresholds at the current level or below, with an efficiency similar to that of the offline reconstruction. Triggers are therefore assumed to be 100% efficient in the selected phase-space. This hypothesis was confirmed to be reasonable by Level-1 trigger menu studies featured in the recent Level-1 trigger technical design report [138].

Finally, a 1% uncertainty on the total luminosity, fully correlated in all channels, is considered.

5.2. HH → $\mathrm{b\bar{b}b\bar{b}}$ channel

With more than one third of the di-Higgs decays, the $\mathrm{b\bar{b}b\bar{b}}$ final state has the highest branching fraction among HH final states. On the other hand, this fully hadronic final state suffers from a large multijet QCD background that makes it experimentally challenging.

As for most of the work described here, and following the example set earlier with a Run 2 analysis, the strategy for this analysis will be to use machine learning techniques on events passing selection criteria to build a discriminator and use its output to perform a shape-based statistical analysis. Given the absence of data, limits and significance as a function of anomalous coupling will be computed under the no-signal and standard model hypotheses respectively.

The analysis described below assumes that the four jets from the $HH \rightarrow \mathrm{b\bar{b}b\bar{b}}$ can be reconstructed independently and is referred to as the "resolved" topology. Given the mass of the Higgs boson, the overwhelming majority of standard model events fall in this category. A separate boosted event topology to improve sensitivity to beyond the standard model searches was performed by relaxing this criteria, but this will not be discussed in this document.

The event selection requires at least four jets of transverse momentum $p_{\mathrm{T}} > 45$ GeV, pseudo-rapidity $|\eta| < 3.5$ and the medium working point b-tagging conditions to be satisfied. If more of these objects are available, the four jets with the highest transverse momentum are selected. This leads to a signal preselection efficiency of about 7%.

These jets are then combined into Higgs boson candidates (H_1,H_2) so as to minimise the absolute difference in invariant masses. Events are also required to pass the following selection, passed by 55% of signal events and rejecting about 85% of the multijet background.

$$\sqrt{\left(m_{\mathrm{H}_1} - 120\ \mathrm{GeV}\right)^2 + \left(m_{\mathrm{H}_2} - 120\ \mathrm{GeV}\right)^2} < 40\ \mathrm{GeV} \tag{5.1.}$$

After applying these cuts, the overall signal yield is expected to reach 1370 events for a background of $1.1 \cdot 10^7$ events. To enhance sensitivity, a boosted decision tree (BDT) was trained on the 19 following kinematic variables.

- The invariant masses and the transverse momentum of the two reconstructed Higgs boson candidates
- The invariant mass of the di-Higgs system m_{HH}
- The reduced invariant mass of the di-Higgs system $M_{\mathrm{HH}} = m_{\mathrm{HH}} - (m_{\mathrm{H}_1} - 125\ \mathrm{GeV}) - (m_{\mathrm{H}_2} - 125\ \mathrm{GeV})$, mitigating detector resolution effects for dijets originating from Higgs bosons
- The minimum and maximum $\Delta\eta$, $\Delta\phi$ and $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ between jets
- The $\Delta\eta, \Delta\phi$ and $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ between jets forming both Higgs boson candidates H_1 and H_2
- The cosine of the angle between the direction formed by one of the Higgs and the beamline axis in the di-Higgs rest frame

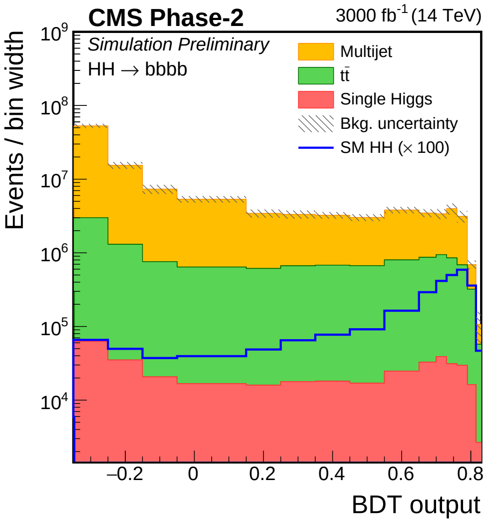


Figure 5.1. | Output of the boosted decision tree trained on 19 kinematic variables of events passing the $\mathrm{HH} \to \mathrm{b\bar{b}b\bar{b}}$ selection criteria [6].

The output of this discriminant, shown on figure 5.1 is used to look for excess signal in the high output region. Assuming there is no signal, a limit of twice the standard model at the 95% confidence level is expected for an integrated luminosity of 3000 fb^{-1}. If the signal predicted by the standard model is assumed, this translates into a significance of 1.0 standard deviation.

5.3. HH → $\mathrm{b\overline{b}}\tau\tau$ channel

The second most probable final state is with one of the Higgs bosons decaying into a pair of b quarks and the second into a pair of τ leptons. The τ leptons will then decay either hadronically into light hadrons and a neutrino, or leptonically into an electron or muon and neutrinos. While the preferred decay channel of a τ lepton is hadronic (roughly two thirds of the time), the leptonic decays will be of particular interest in this analysis since muons and electrons are well reconstructed by the detector and provide a strong handle on the QCD background.

The backgrounds considered in this analysis are top quark pair production, single Higgs, Drell-Yan to dileptons and jets, single top production alongside a W or a Z, di-vector-boson and W+jets.

While the acceptance to multijet background is expected to be very small, the high cross section of this process could yield a significant contribution to the final state. This analysis did not attempt to model this background and it could be addressed in future versions. This background is typically derived from data when available, as was the case for the run 2 analysis [139].

The presence of irreducible background such as $t\bar{t} \to \mathrm{b\overline{b}W^+W^-}$ and $\mathrm{Z}\gamma^* + \mathrm{b\overline{b}} \to \tau^+\tau^- + \mathrm{b\overline{b}}$, along with jets misidentified as hadronically decaying τ in multijet events render this final state very challenging and mandates advanced data analysis techniques to maximise the sensitivity to a signal that cannot be fully reconstructed given the neutrinos in the final state.

The analysis strategy is, after object preselection, to separate the final state depending on the $\tau^+\tau^-$ decay channel and train deep neural networks using basic, reconstructed and global variables to provide the final discriminant.

The channels considered here will involve at least one hadronic tau decay (τ_h), corresponding to roughly 88% of all tau lepton pairs. The $\ell\tau_h$ decay modes, where ℓ denotes either a muon or an electron requires exactly one and only one lepton and at least one reconstructed τ_h of opposite sign. In the $\tau_h\tau_h$ channel, at least two opposite sign τ_h are required with no additional leptons. If multiple τ_h pairs are possible, the highest p_T pair is selected.

All three decay channels also require a pair of b-tagged jets. The detailed values of the cuts considered are displayed in Table 5.2.

From these selected events, a total of 52 features are used to separate signal and backgrounds.

The first 27 variables are basic features of the reconstructed event. The 4-momenta in Cartesian coordinates (p_x, p_y, p_z, E), the norm of the 3-momenta (p) and the masses of the four final state objects are used. The missing transverse momentum is described by its two components (p_x^miss, p_y^miss) and its norm (p_T^{miss}).

Then 21 reconstructed features are computed. As for the primary objects, the 4-momenta

Lepton	Min. p_T [GeV]	Max. $\lvert\eta\rvert$	Max. iso [GeV]
Primary μ	23	2.1	0.15
Primary e	27	2.1	0.1
Veto e/μ	10	2.4	0.3
Hadronic tau	**Min. p_T [GeV]**	**Max. $\lvert\eta\rvert$**	
$\ell\tau_h$ $(\ell = e,\mu)$	20	2.3	
$\tau_h\tau_h$	45	2.1	
Jets	**Min. p_T [GeV]**	**Max. $\lvert\eta\rvert$**	
b-tagged jets	30	2.4	

Table 5.2. Kinematic requirements of electrons, muons, hadronically decay taus and b-tagged jets used in the selection. The additional lepton veto is also displayed.

in Cartesian coordinates (p_x, p_y, p_z, E), the norm of the 3-momenta (p) and the invariant masses of the Higgs candidates coming from the two b-jets (H_{bb}) and the two τ $(H_{\tau\tau})$ and the di-Higgs system (HH) are calculated. The stransverse mass of the system (M_{T2}) and the transverse masses of each τ as defined in Eq. 5.2 are also computed.

$$m_T = \sqrt{2p_{T,\tau} \times p_\mathrm{T}^\mathrm{miss} \times \left(1 - \cos\Delta\phi_{\tau,p_\mathrm{T}^\mathrm{miss}}\right)}. \tag{5.2.}$$

Finally, four global features are used to characterise these events. These are the scalar sum of the muon, tau, b-jet and missing transverse momentum (s_T) on the one side, and the total number of jets, the number of b-jets and the number of τ_h on the other side.

Six of these 52 variables are shown in figure 5.2.

Using simulated samples and the variables defined here, two discriminants, identical in structure, are then built. Each of them are trained on half of the simulated events and then used for the final analysis on the half not trained upon.

Ten fully connected deep-neural networks consisting of three layers of 100 neurons with SELU activation functions were trained on ten folds of the data. These ten networks are then weighted and combined to build an ensemble classifier. These were implemented using Keras with Tensorflow as a back-end.

A simultaneous fit is performed on the output of the classifiers in the three different channels. With the assumed systematic uncertainties, an upper limit on the HH production cross section times branching fraction of 1.4 times the standard model is obtained. This corresponds to a significance of 1.4 standard deviations if the standard model is assumed.

5.4. HH → b$\bar{\text{b}}\gamma\gamma$ channel

The final state with one of the Higgs bosons decaying to a pair of b quarks, and the other into a pair of photons will now be considered.

While the branching fraction of this decay is tiny, the two photons provide a very clean

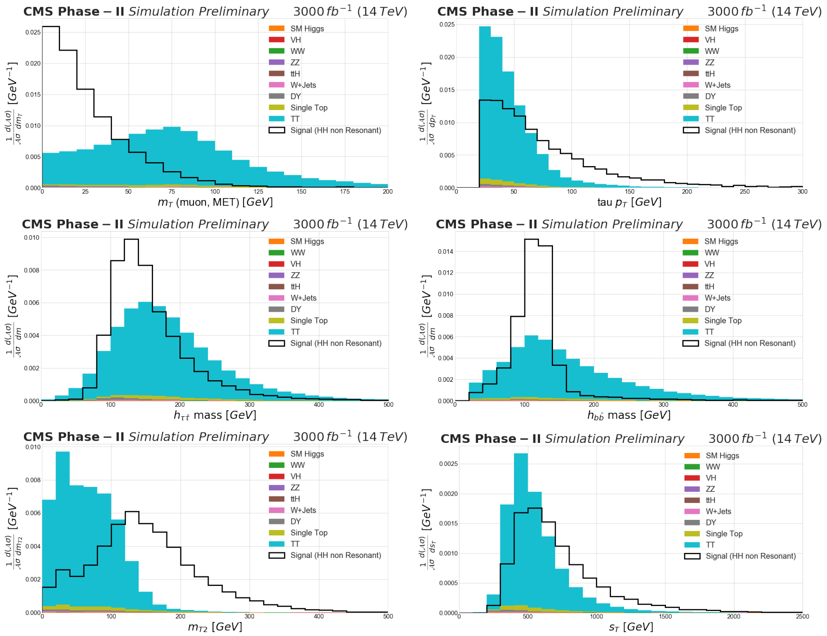


Figure 5.2. | Distribution for signal and background processes of six of the features used as input of an ANN in the τ_μ, τ_h category. From the top left to the bottom right, the distributions displayed are the transverse mass of the muonically decaying tau, as defined in equation 5.2, the transverse momentum of the τ_h, the invariant mass of the $H_{\tau\tau}$ and $H_{b\bar{b}}$ candidates, the system stransverse mass M_{T2} and the scalar sum of the muon, b-jets, τ_h and missing transverse momentum (s_T) [6].

final state and this channel will provide the highest sensitivity to the standard model. The main backgrounds come from the continuous di-photon spectrum, single photon and a misidentified jet, and single Higgs boson production.

The analysis strategy is to use machine learning techniques to mitigate irreducible background contributions, then perform a combined fit on the di-photon and di-jet invariant mass.

The object selection is as follows. Photons and b-tagged jets passing their respective loose identification criteria are considered. While the upgraded detector has a higher angular acceptance, the signal being more central than the background, the objects are required to have a pseudo-rapidity $|\eta|$ smaller than 2.5. The pair of photons of highest transverse momentum and the pair of jets satisfying the highest b-tagging criteria, or highest p_{T} in case of ambiguity are then selected. In addition, these objects have to pass a series kinematic cuts described in table 5.3.

Given the analysis strategy, it is crucial to suppress the Higgs boson production in association with a pair of top-quarks that forms the main resonant background process. Unfortunately, the resolution of the detector is insufficient to exploit the resonant nature of the signal b quark pair production, as clearly visible on figure 5.3. To perform such a task, a boosted decision tree is trained to discriminate the signal from the $t\bar{t}$H

Photon selection
At least 2 photons (loose WP)
$p_T > 30$ GeV and $\lvert\eta\rvert < 1.44$ or $1.57 < \lvert\eta\rvert < 2.5$
100 GeV $< m_{\gamma\gamma} <$ 180 GeV
$p_T/m_{\gamma\gamma} > 1/3$ (leading γ), $1/4$ (sub-leading γ)
Jet selection
At least 2 b-tagged jets (loose WP)
$p_T > 25$ GeV and $\lvert\eta\rvert < 2.5$
80 GeV $< m_{jj} <$ 190 GeV
$\min(\Delta R_{\gamma j}) > 0.4$

Table 5.3. | Photon and jet kinematic selections in the $b\bar{b}\gamma\gamma$ final state.

background.

This classifier is given the total number of leptons, of jets and of medium b-tagged jets as input. For all selected objects, the p_T/M ratio is provided. The angles on the transverse plane between the missing momentum and the jets is also considered. Finally, the angles between the directions of the di-photon object, of the leading jet and of the leading selected photon and the direction of the beam axis in the $\gamma\gamma jj$, jj and $\gamma\gamma$ rest frames respectively will complete the variables provided to the discriminant.

A selection is then applied to the output of this boosted decision tree to reject approximately 75% of the $t\bar{t}$H contamination for a 90% signal efficiency.

To enhance the sensitivity of this analysis, the events are classified into six different categories before performing a global fit.

The first classification is obtained by applying a selection on the M_{HH} variable defined in equation 5.3, similar to what was done in the $b\bar{b}b\bar{b}$ channel. This variable gives access to the invariant mass of the system with a mitigation of resolution effects by assuming that the di-photon and di-jet systems originate from Higgs bosons.

$$M_{HH} = m_{HH} - m_{\gamma\gamma} - m_{jj} + 250 \text{ GeV} \tag{5.3.}$$

Three M_{HH} categories are defined. The low-mass region corresponds to $250 < M_{HH} <$ 350 GeV, medium-mass to $350 < M_{HH} <$ 380 GeV and high-mass to $M_{HH} >$ 480 GeV.

A final classification is then performed based on a second multi-variate discriminant. A boosted decision tree is trained to discriminate the signal from the sum of background processes with the exception of $t\bar{t}$H, already suppressed by the first dedicated discriminant. The input variables of this discriminant are the following :

- The minimum angular distance between the selected jets and photons
- The angles between the di-photon object, leading jet or leading photon and the beam axis in the $\gamma\gamma jj$, jj and $\gamma\gamma$ rest-frame respectively
- The p_T/M ratio for the selected photons and jets

- The number of jets passing the loose, medium and tight b-tagging criteria
- The b-tagging score of the two selected jets
- The two relative photon resolutions σ_E/E
- The angle between the two selected jets and the missing momentum in the transverse plane

Events with a low BDT score are discarded, removing roughly 90% of the background with a 85% signal efficiency, and the remaining events are split relative to their score into high and medium purity categories (HP and MP). The definition of these two categories was optimised in the three M_X regions in order to maximise sensitivity to the standard model. It has been verified that this classification does not alter the $m_{\gamma\gamma}$ and m_{jj} shapes.

In the region considered for the fit (shown in figure 5.3), a total of 40 di-Higgs events are expected with 190 and 3600 resonant and non-resonant background events respectively. While the high mass categories are the most sensitive to the standard model, the lower mass categories allows an enhanced sensitivity to variations of the Higgs boson self-coupling.

The signal and background are modelled by fitting the simulated samples. The non-resonant background is well modelled by an exponentially falling distribution while the resonant signal and backgrounds are modelled by a Crystal Ball and the sum of a Crystal Ball and a Gaussian distribution in $m_{\gamma\gamma}$ and m_{jj} respectively.

Pseudo-data are drawn from these models and a combined fit on these yield a 95% confidence level limit of 1.09 times the standard model, corresponding to a significance of 1.83 standard deviations.

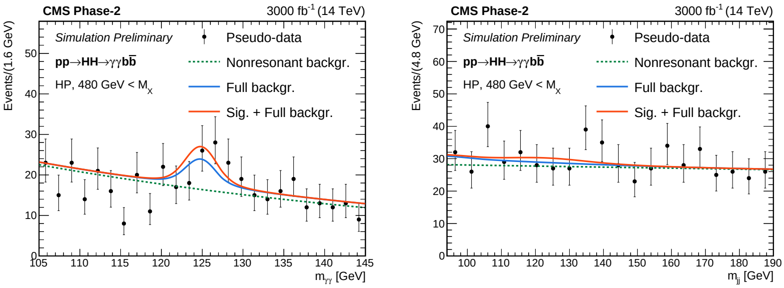
5.5. HH → $\mathrm{b\bar{b}ZZ}$ channel

The unprecedented luminosity offered by the HL-LHC will open new measurement opportunities for very low cross-section phenomena. This analysis, where the process considered is $\mathrm{HH} \rightarrow \mathrm{b\bar{b}}\mathrm{Z}(\ell^+\ell^-)\mathrm{Z}(\ell^+\ell^-)$ with $\ell^\pm = \mathrm{e}^\pm, \mu^\pm$, explores the possibilities offered by a very clean final state with a tiny branching fraction.

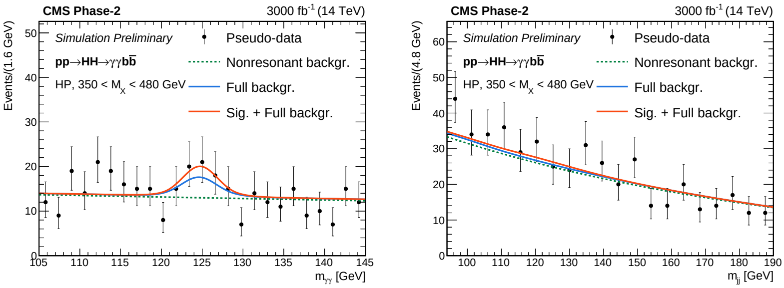
Given the fully resolved final state and the expected yields, the analysis strategy is to select events compatible with the topology and to perform a cut and count analysis.

Events are selected if they have two pairs of opposite sign muons or electrons with invariant masses compatible with a pair of Z bosons originating from a Higgs boson decay. The pair of leptons with an invariant mass closest to the Z mass is noted Z_1, and Z_2 denotes the remaining pair of leptons with the highest transverse momentum. Furthermore, they are required to have a pair of b-tagged jets with a corrected invariant mass, assuming a 20% resolution improvement, compatible with a Higgs boson. No more than three jets passing the medium b-tagging working point are allowed. The full list of cuts is detailed in table 5.4.

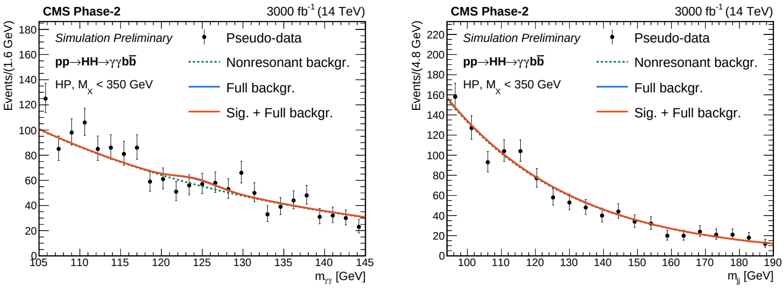
With this selection, a total of 1.0 standard model signal events are expected for an integrated luminosity of 3000 fb^{-1}, split in three different channels ($\mathrm{b\bar{b}}$ 4e, $\mathrm{b\bar{b}}$ 4μ and $\mathrm{b\bar{b}}$ 2e 2μ). The background processes, mainly $t\bar{t}$ H (Z Z), $t\bar{t}$ Z, g g H and Z H are expected to yield a total of 6.8 events. The four lepton invariant mass distributions for the signal and background are shown on figure 5.4. Top quark pairs and Drell-Yan



(a) $m_{\gamma\gamma}$, high mass category **(b)** m_{jj}, high mass category



(c) $m_{\gamma\gamma}$, medium mass category **(d)** m_{jj}, medium mass category



(e) $m_{\gamma\gamma}$, low mass category **(f)** m_{jj}, low mass category

Figure 5.3. Di-photon and di-jet invariant mass distribution for the non-resonant background, total background, signal and pseudo-data assuming the standard model production. The distributions shown are split into low, medium and high mass categories, and fulfil the high purity requirement [6].

in association with jet productions are reducible backgrounds that are expected to be highly suppressed but have a huge production cross-section. The available Monte-Carlo samples and the parametric detector simulation used in this work could not reliably predict the final contribution of these processes to this analysis. It is therefore assumed that some dedicated algorithms and improvements in data reconstruction techniques will be available at the high-luminosity LHC, enabling this background to represent only a negligible contamination.

A statistical analysis of these results split in the three channels gives a predicted upper

	Lepton selection
Muon selection	Loose WP identification and $p_\mathrm{T} > 5$ GeV
	$\lvert\eta\rvert < 2.8$ and isolation < 0.7
Electron selection	Medium WP identification and $p_\mathrm{T} > 7$ GeV
	$\lvert\eta\rvert < 2.8$ and isolation < 0.7
	Z candidate selection
Z candidates	Pair of opposite sign, same flavour leptons
	$\Delta R_{\ell^+\ell^-} > 0.02$
Z_1 candidate	50 GeV $< m_{\ell^+\ell^-} <$ 100 GeV
Z_2 candidate	12 GeV $< m_{\ell^+\ell^-} <$ 60 GeV
	Jet selection
b-jets	2 or 3 b-tagged jets at medium working point
Corrected invariant mass	90 GeV $< m_{\mathrm{b\bar{b}}} <$ 150 GeV
	Event selection
(Sub)Leading lepton	$p_\mathrm{T} > 20$ (10) GeV
Selected leptons	120 GeV $< m_{4\ell} <$ 130 GeV
Missing transverse energy	MET $<$ 150 GeV
Reconstructed Higgs bosons	$\Delta R_\mathrm{HH} \geq 2.0$

Table 5.4. Summary table of object and event definitions and selection cuts in the $\mathrm{HH} \to \mathrm{b\bar{b}ZZ}$ channel.

limit of 6.6 times the standard model at the 95% confidence level, corresponding to a significance of 0.37 standard deviations.

5.6. HH → $\mathrm{b\bar{b}WW}$ channel

5.6.1. Introduction

The last channel considered is the search for HH production in the $H \to b\bar{b}, H \to WW \to l\nu l\nu$ final state. This analysis will be described in greater detail since it was mostly realised by the author of this document. While most of the analysis strategy and the core framework used for the analysis originated from the run 2 analysis, the author adapted them to suit the HL-LHC prospective analysis needs and different data formats. He also performed all the analysis steps and the required developments to arrive to the final result described here. Finally, he helped at the redaction of the documentation and scientific publication that followed.

This final state is not the most favoured in terms of cross section times branching fraction, but it can contribute in a sizeable way when the results from the different decay channels are combined.

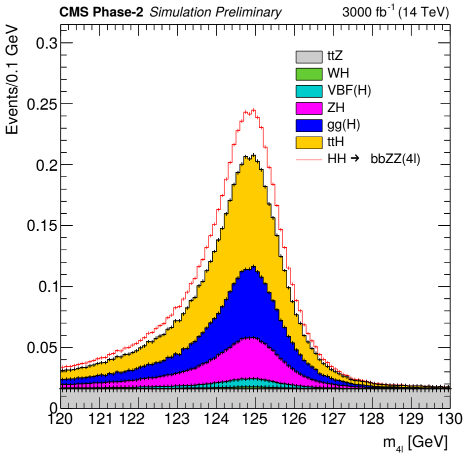


Figure 5.4. | Invariant mass distribution of the four lepton system for signal and background simulated events. A total of 1.0 signal event and 6.8 background events are expected to pass this selection [6].

The dominant and subdominant background processes are the $t\bar{t}$ production in the fully leptonic decay mode, and the Drell-Yan plus jets production, both consisting of exactly the same objects as the signal in the final state.

In order to separate HH events from these backgrounds, different kinematic variables are combined in an artificial Neural Network (NN) discriminant .

5.6.2. Signal samples

The simulated samples corresponding to 3000 fb^{-1} of proton-proton collisions at $\sqrt{s} = 14$ TeV described in section 5.1.1 were generated and analysed.

From the signal samples that describe events at leading order of gluon fusion production of two SM-Higgs bosons with a mass of 125 GeV, one of the Higgs bosons is required to decay into a pair of b quarks, while the other is required to decay into a pair of weak bosons and subsequently to final states containing two leptons and two neutrinos.

This implies that the signal samples contain both $\mathrm{H} \rightarrow \mathrm{Z}(\ell\ell)\mathrm{Z}(\nu\nu)$ and $\mathrm{H} \rightarrow \mathrm{W}(\ell\nu)\mathrm{W}(\ell\nu)$ decay legs, but further kinematic selections will mostly suppress the ZZ contribution hence this analysis being commonly referred to as $\mathrm{b\bar{b}}WW$. The SM branching ratios are assumed.

The leptons considered for the decay process are electrons and muons as well as tau leptons decaying into electrons and muons, and the three neutrinos flavours are considered. The branching fraction is of 1.73 % [140].

To obtain the desired equivalent sample as a function of Higgs self-coupling coupling modifier in the κ-framework [32] from the SM sample, the events are then reweighted at the matrix element level using the distributions of di-Higgs invariant mass (m_{HH})

and cosine of the polar angle ($\cos\theta_{\mathrm{HH}}^{\mathrm{CS}}$) of one Higgs in the Collins-Soper frame [141]. This method is justified by the fact that, at leading order and parton level, the signal can be characterised by these two variables only.

The full list of background and signal samples used in the analysis, along with their assumed cross sections are shown in Table 5.5.

Process	Generator	Cross section [pb]
$t\bar{t}$	POWHEG	854.4
$t\bar{t}h$	MADGRAPH5_AMC@NLO	0.6113
Drell-Yan	MADGRAPH5_AMC@NLO	22461
$\mathrm{H} \rightarrow WW \rightarrow l\nu l\nu$	MADGRAPH5_AMC@NLO	0.01495
H H	MADGRAPH5_AMC@NLO	0.00063

Table 5.5. Background and signal simulated samples used in the analysis and their cross sections in pb. The standard model predicted cross-section times branching fraction is quoted for the HH process.

5.6.3. Event selection and background predictions

The dominant background in this analysis is the irreducible contribution from $t\bar{t}$ events that also decay into a pair of W bosons and b quarks. The second significant background comes from Drell-Yan processes. While this background can be highly suppressed thanks to b-tagging requirements and kinematic cuts, the overwhelming cross section results in a sizeable and difficult to model contribution to the final state. Finally, two additional contributions from single Higgs boson processes ($t\bar{t}\mathrm{H}$ and $\mathrm{H} \rightarrow \mathrm{WW}$) were considered. These were not expected to be the next leading contributors, nor sizeable in any way, but were nevertheless included as a cross-check given that other final states start to be susceptible to single Higgs boson contamination at the HL-LHC.

Events are selected assuming they pass the expected di-lepton trigger requirements. As for the other analyses, it is assumed that the transverse momentum cuts will be similar to that of the run 2, and that the online trigger efficiency will match the offline efficiency, thus leading to a 100% efficiency on selected events.

The cleaning cuts are identical to their run 2 counterparts, by requiring two oppositely charged leptons ($e^+e^-, \mu^+\mu^-, e^\pm\mu^\mp$), in which the leptons are required to have a p_T greater than 25 GeV and 15 GeV for ee events, 20 GeV and 10 GeV for $\mu\mu$ events, 25 GeV and 15 GeV for μe events, 25 GeV and 10 GeV for $e\mu$ events, for the higher and lower p_T lepton, respectively.

Electrons and muons in the pseudo-rapidity range $|\eta| < 2.8$ are considered. Electrons in the $1.444 < |\eta| < 1.5666$ overlap region are rejected. A dilepton mass requirement of $m_{\ell\ell} > 12$ GeV is applied to all flavour combinations in order to suppress lepton onia resonances.

Jets are required to have $p_\mathrm{T} > 20$ GeV, $|\eta| < 2.8$, and be separated from identified leptons by a distance of $\sqrt{\Delta\phi^2 + \Delta\eta^2} = \Delta R > 0.3$.

To identify jets originating from b quarks, a combined multivariate algorithm is used for fully simulated data. In these Delphes samples, the output of this algorithm is emulated and jets are considered as b-tagged if they pass the medium working point of the emulation, which provides around 70% efficiency with a mistag rate of less than 1% [142].

A summary of the object definitions and selection cuts is shown in Table 5.6.

Object	Definition	Selection cuts
leading (sub-leading) e	Isolation ($\|\eta\| > 1.5666$) < 0.559 (0.853)	$p_T > 25(15)$ GeV
		$\|\eta\| < 1.444$ or $1.5666 < \|\eta\| < 2.8$
leading (sub-leading) μ	Loose ID	$p_T > 20(10)$ GeV, $\|\eta\| < 2.8$
	Isolation < 0.25	$p_T > 25$ GeV for μe events
jets	PUPPI Jet	$p_T > 20$ GeV, $\|\eta\| < 2.8 \Delta R_{lj} > 0.3$
b-jets	PUPPI medium MTD WP	

Table 5.6. Object definitions and selection cuts in the HH $\rightarrow b\bar{b}WW$ channel.

Among all possible di-jet combinations fulfilling the previous criteria the two jets with the highest combined transverse momentum are selected.

Several distributions were checked to ensure the understanding of the selected objects in simulation. As an illustration, figure 5.5 shows the $m_{\ell\ell}$ distribution for simulated events requiring two opposite sign leptons and two jets in all possible lepton flavour combinations. The same distributions requiring two opposite sign leptons, two b-tagged jets, and $m_{\ell\ell} < 76$ GeV are shown in Fig. 5.6.

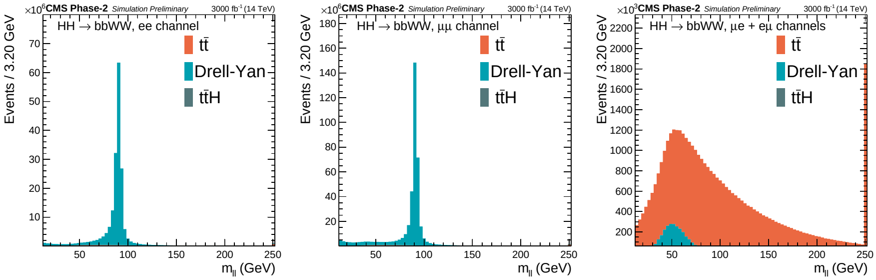


Figure 5.5. The $m_{\ell\ell}$ distribution for simulated events after requiring two leptons, and two jets. The three different lepton flavour combinations (e^+e^- (left), $\mu^+\mu^-$ (centre), $e^\pm\mu^\mp$ (right)) are shown.

After the final object selection consisting of two opposite sign leptons and two b-tagged jets, a cut on $m_Z - m_{\ell\ell} > 15$ GeV is applied to remove the resonant Z peak and the high-$m_{\ell\ell}$ tail of the Drell-Yan+jets and $t\bar{t}$ background process. Several control distributions after this cut are shown in Fig. 5.7.

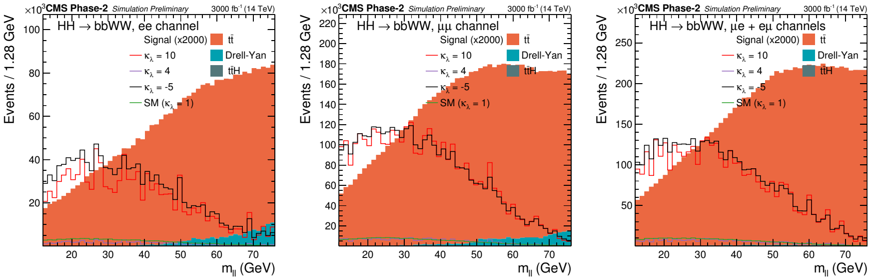


Figure 5.6. The $m_{\ell\ell}$ distribution for simulated events after requiring two leptons, two b-tagged jets, and $m_{\ell\ell} < 76$ GeV. The three different lepton flavour combinations (e^+e^- (left), $\mu^+\mu^-$ (centre), $e^\pm\mu^\mp$ (right)) are shown.

A summary of the signal and background yields after all the selection requirements described in this section is shown in table 5.7.

	ElEl, llbb, mll cut	MuEl, llbb, mll cut	MuMu, llbb, mll cut
Signal samples			
κ_λ = -5	1282000 ± 22000	4091000 ± 49000	3605000 ± 48000
κ_λ = 10	1101000 ± 35000	3878000 ± 85000	3592000 ± 85000
κ_λ = 4	67500 ± 3200	255700 ± 7900	240400 ± 7900
SM (κ_λ = 1)	103730 ± 890	298400 ± 1500	242900 ± 1400
SM samples			
$t\bar{t}$	2646300 ± 5100	8336000 ± 9000	685230 ± 8200
$t\bar{t}H$	5350 ± 50	12787 ± 78	8252 ± 63
Drell-Yan	224900 ± 7800	30700 ± 1700	610000 ± 14000
Higgs to WW	6.6 ± 2.0	17.6 ± 2.7	8.2 ± 2.2
Total ± (stat.)	2876500 ± 9300	8380000 ± 9200	7470200 ± 16600

Table 5.7. Expected signal and background yields requiring two leptons, two b-tagged jets, and $\mathrm{m_Z} - \mathrm{m_{ll}} > 15$ GeV, at an integrated luminosity of 3000 fb^{-1}. The three different lepton flavour combinations (e^+e^- (left), $e^\pm\mu^\mp$ (centre), $\mu^+\mu^-$ (right)) are shown.

5.6.4. Signal extraction

An artificial neural network (ANN) discriminant, based on the TMVA library [72], is used to improve the signal-to-background separation. In a phase space dominated by $t\bar{t}$ production, the NN utilises information related to object kinematics. The variables provided as input to the NN exploit the presence of two Higgs bosons decaying into two b-jets on the one hand, and two leptons and two neutrinos on the other, which results in different kinematics for the di-lepton and di-jet systems between signal and

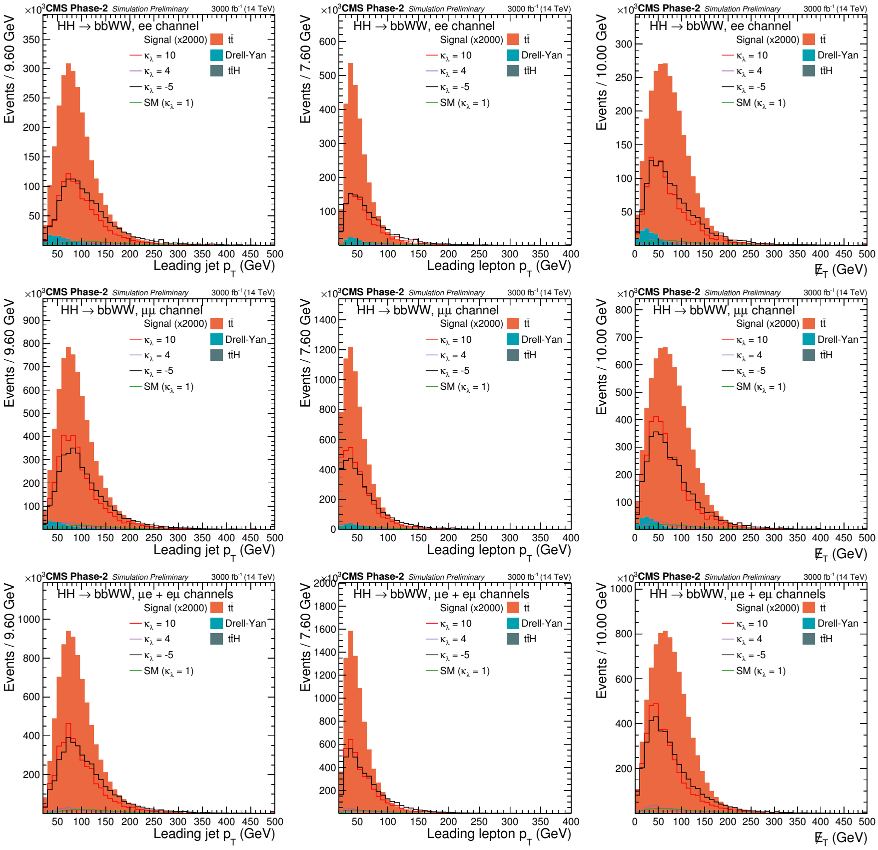


Figure 5.7. The leading jet p_{T}, leading lepton p_{T}, and $/\mathrm{E}_{\mathrm{T}}$ distributions for simulated events after requiring two leptons, two b-tagged jets, and $\mathrm{m}_{\mathrm{Z}} - \mathrm{m}_{\mathrm{ll}} > 15$ GeV. e^+e^- events are shown in the upper row, $\mu^+\mu^-$ in the middle row, and $e^\pm\mu^\mp$ in the lower row.

background processes. The set of variables used as input is: $m_{\ell\ell}$, m_{jj}, $\Delta R_{\ell\ell}$, ΔR_{jj}, $\Delta\phi_{\ell\ell,\mathrm{jj}}$, defined as the $\Delta\phi$ between the di-jet and the di-lepton systems, $p_{\mathrm{T}}^{\ell\ell}$, $p_{\mathrm{T}}^{\mathrm{jj}}$, $\min\left(\Delta R_{\mathrm{j},\ell}\right)$, and M_{T}, defined as $\mathrm{M}_{\mathrm{T}} = \sqrt{2 p_{\mathrm{T}}^{\ell\ell} E_{\mathrm{T}}^{\mathrm{miss}} (1 - \cos(\Delta\phi(\ell\ell, E_{\mathrm{T}}^{\mathrm{miss}})))}$.

The input variables distributions are shown in figures 5.8, 5.9 and 5.10.

Given the high level of uncertainty surrounding the many assumptions in this analysis, the decision was taken not to attempt to optimise the architecture or training of the neural network. This was done in order to prevent over-fitting simulation artefacts while not having the ability to compare with data in control regions. Furthermore,the advances in machine learning which will undoubtedly arise in the many years of data taking ahead, would probably render moot any optimisation done at this point.

The structure is thus very basic, and an MLP is used with two hidden layers of nine and eight perceptrons each, using sigmoid activation functions. A schematic display of the

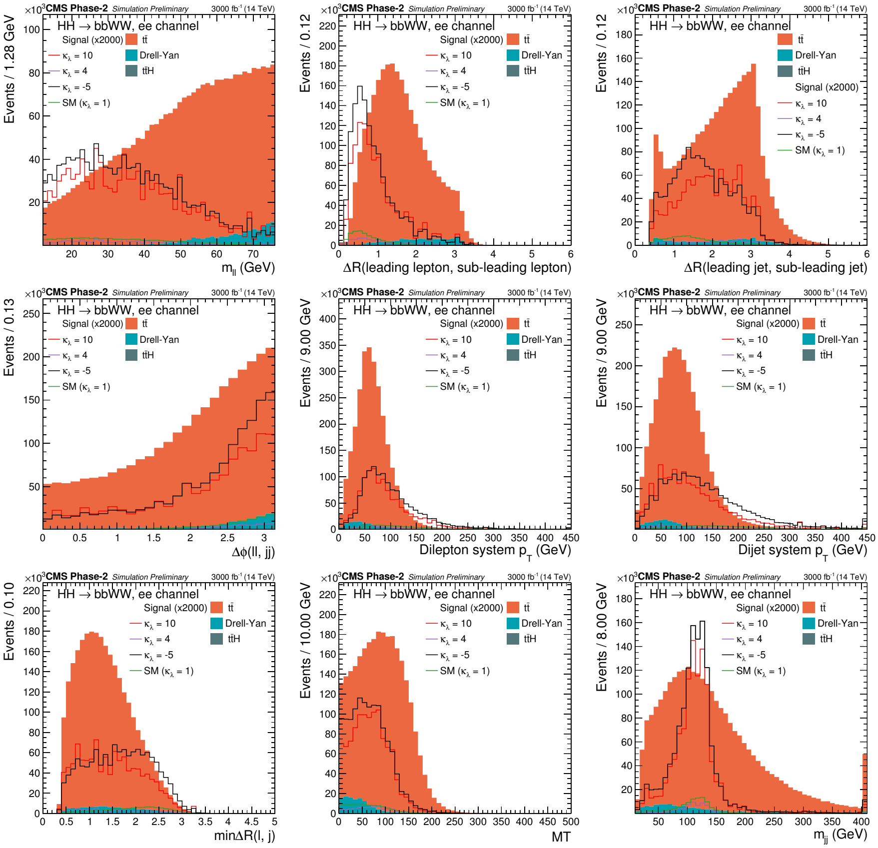


Figure 5.8. The m_{ll}, ΔR_{ll}, ΔR_{jj}, $\Delta\phi_{ll,jj}$, p_T^{ll}, p_T^{jj}, min$\Delta R_{j,l}$, MT and m_{jj} distributions for simulated events after requiring all selection cuts described in Section 5.6.3. The whole set of distributions is used as input of the NN discriminant. The e^+e^- channel only is shown.

network can be seen on figure 5.11.

The NN discriminant output in figure 5.12 shows that the DY contribution cannot be reliably estimated due to statistically limited samples. Since it was assumed that Monte-Carlo statistics would not be a limiting factor at the end of the HL-LHC data taking, the background was fitted with an arbitrary third degree polynomial function split in the different channels. Using the best-fit, fake events were generated as shown in figure 5.12.

Finally, it can also be shown that the discrimination power of the neural network works best for the standard model point. While this can be partially explained by the fact that it was trained on standard model events, the same neural network architecture was trained on different signal hypotheses, yielding no significant improvements, and even a worsening at some times. A possible explanation can be given by the very different shapes of the signal distributions for different κ_λ values and, as can be seen on figures 5.8,

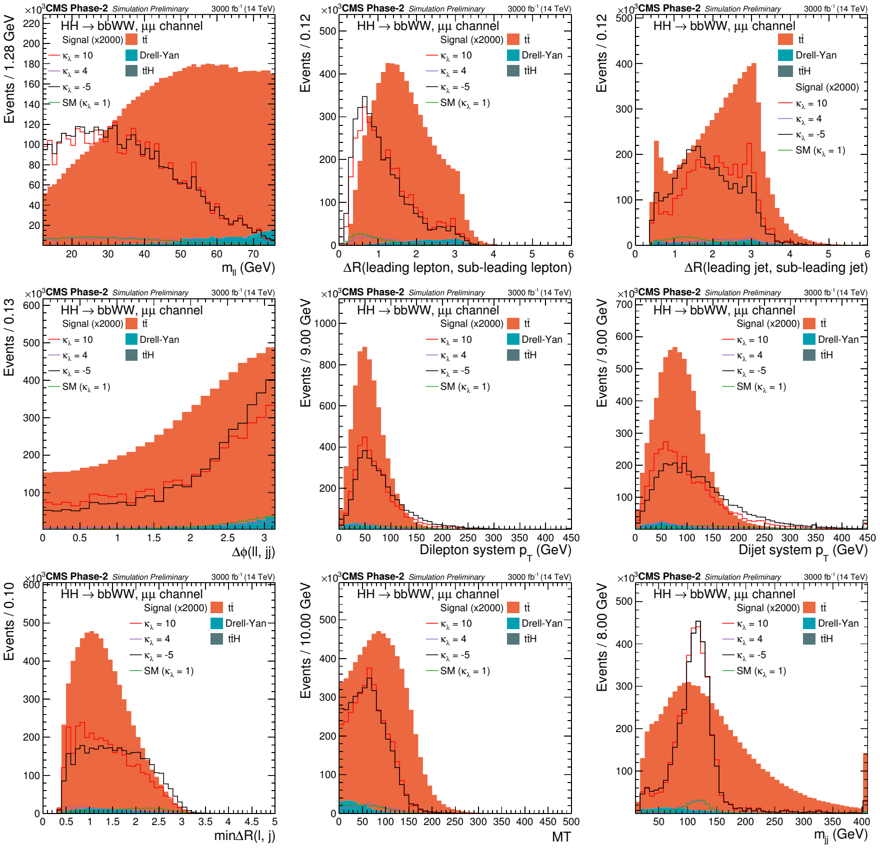


Figure 5.9. The $\mathrm{m_{ll}}$, ΔR_{ll}, ΔR_{jj}, $\Delta\phi_{ll,jj}$, p_T^{ll}, p_T^{jj}, $\min\Delta R_{j,l}$, MT and m_{jj} distributions for simulated events after requiring all selection cuts described in Section 5.6.3. The whole set of distributions is used as input of the NN discriminant. The $\mu^+\mu^-$ channel only is shown.

5.9 and 5.10, the standard model stands out the most from the background distributions.

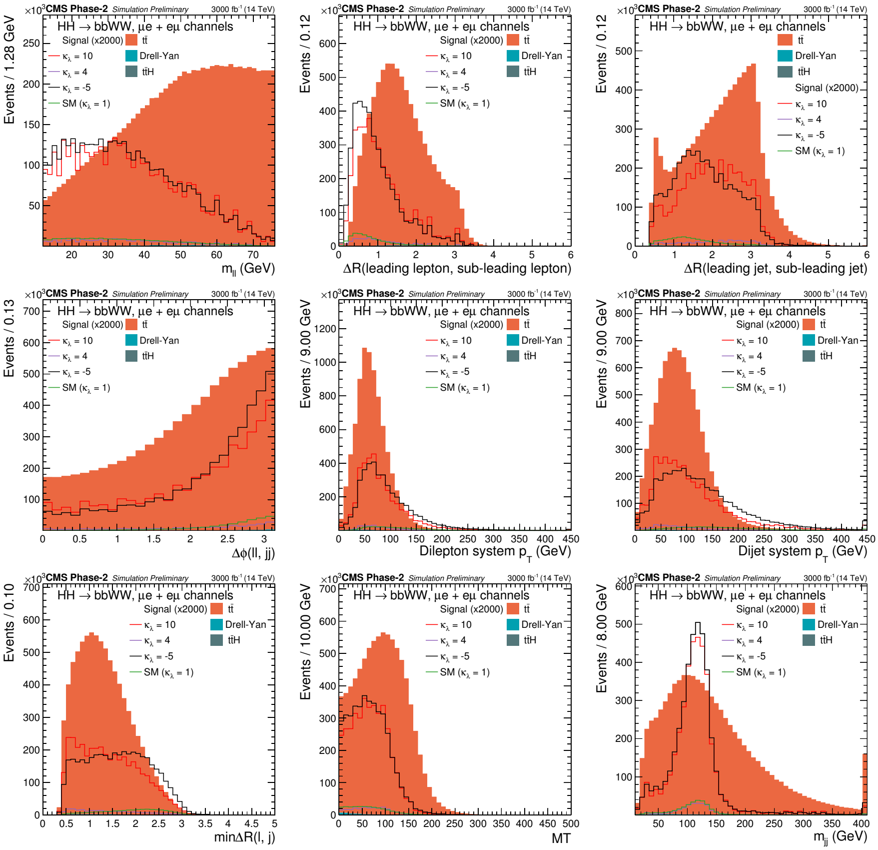


Figure 5.10. | The $\mathrm{m_{ll}}$, ΔR_{ll}, ΔR_{jj}, $\Delta\phi_{ll,jj}$, p_T^{ll}, p_T^{jj}, $\min\Delta R_{j,l}$, MT and m_{jj} distributions for simulated events after requiring all selection cuts described in Section 5.6.3. The whole set of distributions is used as input of the NN discriminant. The $e^{\pm}\mu^{\mp}$ channel only is shown.

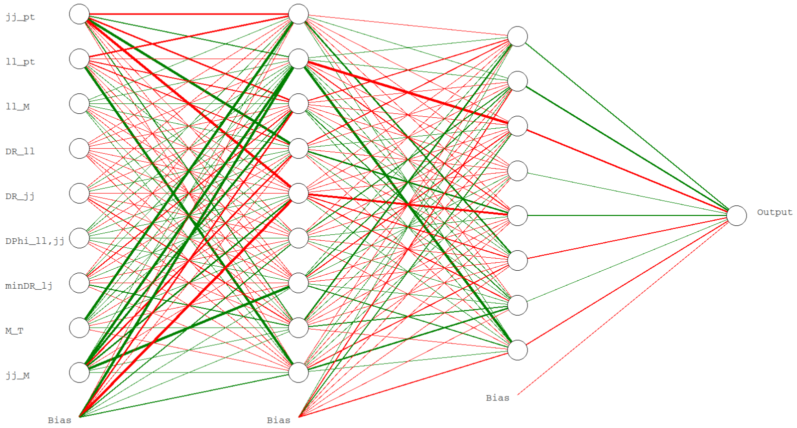


Figure 5.11. | Schematic representation of the neural network used to separate $t\bar{t}$ events from the signal. The width of the lines linking perceptrons gives an indication of the absolute value of the corresponding weight, while the green/red colours indicate a positive/negative sign respectively.

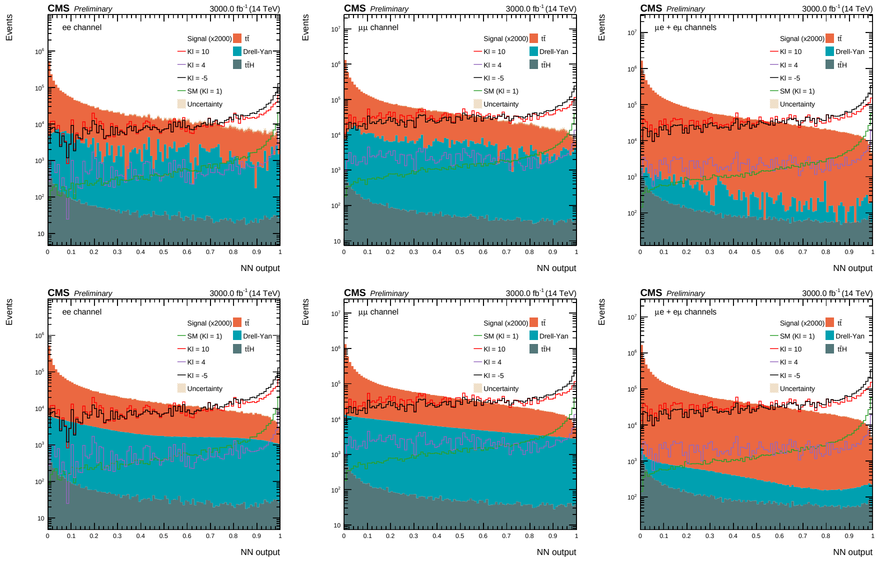


Figure 5.12. | The output of the NN after selection cuts, evaluated in the e^+e^- (left), $\mu^+\mu^-$ (middle), $e^\pm\mu^\mp$ (right) channels on a logarithmic scale. The top row shows the distributions using all Monte-Carlo events while bottom uses DY entries from the fitted distributions.

5.6.5. Systematic uncertainties

The following systematic uncertainties on both shapes and yields are being considered in this prospective analysis.

- **Lepton identification and isolation**: A 0.5% uncertainty on muon isolation and identification efficiency and a 1% (2.5%) uncertainty for electrons above (below) 20 GeV is considered
- **Jet energy scale and resolution**: Four jet energy scale uncertainties are being considered, with their effects being propagated on the MET. The energy scale of all jets was modified by 0.2%, 0.5%, 2% and 0.75% respectively to account for uncertainties on absolute scale, relative scale, Pile-Up and flavour.
- **b-tagging**: The b-tag and mistag rate uncertainties are determined as a function of the jet p_{T}, η, and flavour. Effect on the analysis is estimated by varying these corrections respectively by one standard deviation and split between light and heavy (c,b) jets.
- **Integrated luminosity**: A 1% uncertainty is assumed.
- **$t\bar{t}$ modeling**: No shape uncertainty of $t\bar{t}$ is considered. An arbitrary 2.5% uncertainty on the cross-section is considered, corresponding to half of the value impacting run 2 analysis.
- **Drell-Yan modeling**: While the fitting method used to generate DY distributions has a high uncertainty (of the order of 30%), it is assumed that the DY modeling uncertainty will be small at the end of HL-LHC.

The trigger efficiency uncertainty and MC statistics are assumed to be negligible, and the Pile-Up uncertainties have been included when estimating the other sources of uncertainty.

Since no shape uncertainty is considered to model the $t\bar{t}$ background and that the total number of events for small NN output values is expected to reach tens of millions, there exists a high risk of over-constraining any shape dependent systematic uncertainty. In order to mitigate this effect, the signal-depleted region with NN output < 0.5 will be excluded when running the limits.

5.6.6. Results

To optimise the statistical power of the analysis, a multi-dimensional shape fit was performed on the NN output shown on figure 5.12 for NN output values > 0.5 split into e^+e^-, $\mu^+\mu^-$ and $e^\pm\mu^\mp$ using the *asymptotic* CL_s approach described above, as implemented in the official tool developed by the CMS Higgs combination group [143].

The results are expressed as upper limits on the ratio of the cross section times branching fraction of the process $HH \to b\bar{b}VV \to b\bar{b}ll\nu\nu$ as a function of κ_λ as shown on Fig 5.13.

For the SM point ($\kappa_\lambda = 1$) and assuming no signal is present, the expected upper limit at the 95% confidence level corresponds to 3.5 times the standard model prediction when systematic uncertainties are considered and drops to 3.3 times the standard model if only statistical uncertainties are considered. If the standard model signal is assumed, this translates into a significance of 0.56 standard deviations.

The impact of a misrepresentation of the Drell-Yan output was evaluated by scaling its

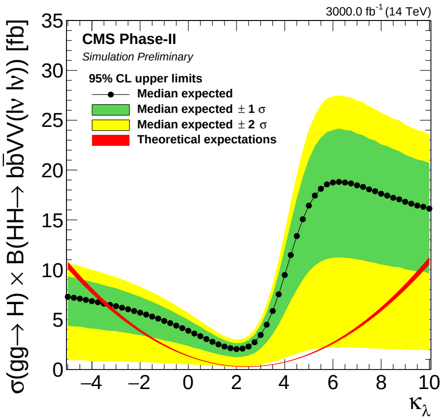


Figure 5.13. Expected 95% CL limit on the cross section times branching fraction of two Higgs bosons to $b\bar{b}VV(ll\nu\nu)$ as a function of κ_λ.

expected yield by a factor of 0.5 or of 2, resulting in changes on the limits and significance of less than 5%. Given the scope of the analysis and the other assumptions made, it is safe to assume that the conclusions drawn will not be affected by such a misrepresentation.

5.7. Combination

As described in the previous sections, no analysis alone is expected to have a high enough sensitivity to provide evidence of the standard model di-Higgs production. In order to provide a better sensitivity, these five analyses were designed to have little overlap in phase space by applying mutually exclusive selections and with a minimal overlap through the use of strict identification criterion. Those channels could thus be combined statistically. The systematic uncertainties linked to the common object definitions and sample simulations are considered to be correlated between the different channels.

The combined upper limit at the 95% confidence level computed amounts to 0.77 times the standard model prediction, with a corresponding significance of 2.6σ. A summary of the significance and limits in the different channels and their combination is shown in table 5.8. While the b$\bar{\mathrm{b}}$WW$(\ell\nu\ell\nu)$ channel is not the most sensitive, this result shows that it does contribute in a sizeable way to the overall measurement.

Different signal samples could be produced as a function of $\kappa_\lambda = \lambda_{\mathrm{HHH}}/\lambda^{\mathrm{SM}}_{\mathrm{HHH}}$, where $\lambda^{\mathrm{SM}}_{\mathrm{HHH}}$ is the standard model prediction using the reweighting method described in section 5.1.1. The limits at the 95% confidence level assuming that no signal is present for every channel and the combination of all channels are shown on figure 5.14. The overall shape of this limit is a result of the change in kinematics of the signal with κ_λ. In the b$\bar{\mathrm{b}}$WW channel, this effect is enhanced by the neural network classifier efficiency being particularly sensitive to κ_λ variations, as mentioned earlier.

Assuming that the HH production cross section is as predicted by the standard model, prospects for the measurement of the λ_{HHH} could be derived. The likelihood scans in

Channel	Significance [σ]		95% CL limit on $\mu = \sigma_{\mathrm{HH}}/\sigma_{\mathrm{HH}}^{\mathrm{SM}}$	
	Stat. + syst.	Stat. only	Stat. + syst.	Stat. only
$\mathrm{b\bar{b}b\bar{b}}$	0.95	1.2	2.1	1.6
$\mathrm{b\bar{b}}\tau\tau$	1.4	1.6	1.4	1.3
$\mathrm{b\bar{b}WW}(\ell\nu\ell\nu)$	0.56	0.59	3.5	3.3
$\mathrm{b\bar{b}}\gamma\gamma$	1.8	1.8	1.1	1.1
$\mathrm{b\bar{b}ZZ}(\ell\ell\ell\ell)$	0.37	0.37	6.6	6.5
Combined	2.6	2.8	0.77	0.71

Table 5.8. | Summary table of the individual channels and combined expected significance and limits, with and without systematical uncertainties being considered [6].

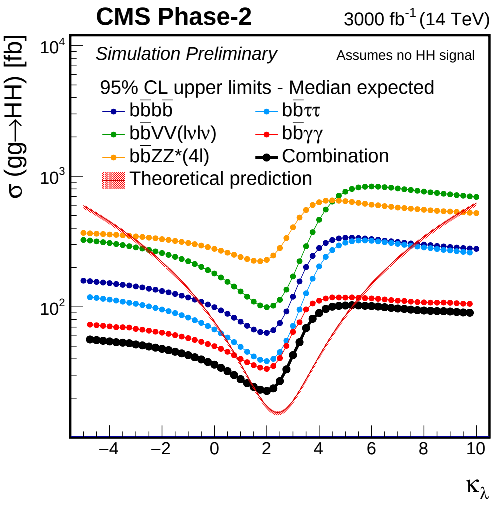


Figure 5.14. | Expected 95% CL upper limit on di-Higgs production cross section as a function of the κ_λ parameter in the different channels and combined, assuming no di-Higgs signal.

the different channels and their combination are shown on figure 5.15. The expected κ_λ confidence intervals are [0.35,1.9] and [−0.18,3.6] at the 68% and 95% CL. respectively.

A double-minimum structure of the significance scans can be seen in the $\mathrm{b\bar{b}b\bar{b}}$, $\mathrm{b\bar{b}}\tau\tau$, $\mathrm{b\bar{b}WW}$ and $\mathrm{b\bar{b}ZZ}$ channels. This shape can be explained by the combination of the loss of sensitivity due to the change in HH kinematics (as shown in 5.14) on the one side, and the quadratic increase in cross-section of the HH signal around a $\kappa_\lambda \approx 2.45$ on the other. The combination of these two effects lead to similar yields for two distinct values in κ_λ.

In order to remove the degeneracy, the analyses need to be sensitive to the differential m_{HH} distribution. While it is somewhat the case in $\mathrm{b\bar{b}b\bar{b}}$ and $\mathrm{b\bar{b}}\tau\tau$, with the BDT and NN being able to differentiate both values to some extent and give a higher significance minimum for $\kappa_\lambda \neq 1$, the $\mathrm{b\bar{b}}\gamma\gamma$ analysis provides the highest handle by far. This latest channel can achieve this through a dedicated classifier for lower m_{HH} values and by providing a good acceptance and purity in this region.

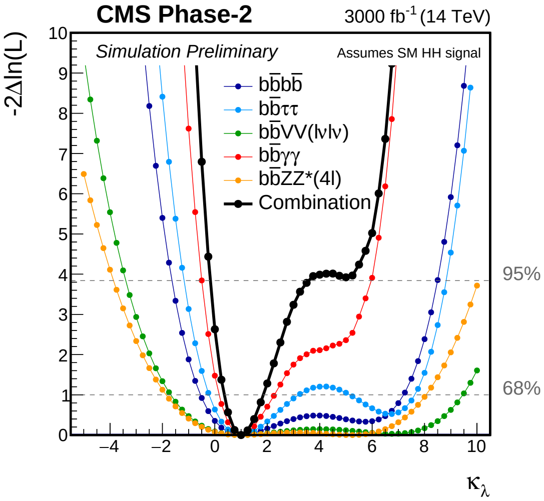


Figure 5.15. | Expected log-likelihood profile as a function of κ_λ of the different channels and combined, assuming the standard model di-Higgs production.

5.8. Outlook

With an expected significance of 2.6 standard deviations on the standard model Higgs production, a first interpretation could cast some doubt on the ability of the upgraded detector to discover this process. Fortunately, a wider view can paint a much more optimistic picture.

While only CMS results have been described in this document, the ATLAS collaboration performed similar studies. The three most sensitive channels were considered : $\mathrm{b\bar{b}b\bar{b}}$, $\mathrm{b\bar{b}}\tau\tau$ and $\mathrm{b\bar{b}}\gamma\gamma$. Extrapolations of the run 2 $\mathrm{b\bar{b}b\bar{b}}$ and $\mathrm{b\bar{b}}\tau\tau$ were performed with some reasonable assumptions on the expected detector performance, as described in a publication [144]. A prospective analysis was performed on the $\mathrm{b\bar{b}}\gamma\gamma$ channel with a parametric simulation of the upgraded detector.

The results from the CMS and ATLAS collaborations were then combined. The luminosity and theoretical uncertainties are the only systematic uncertainties that are expected to be correlated between the two experiments, and their impact on the measurements were shown to be relatively modest. This justified a simple combination with a significance added in quadrature and a summed log-likelihood. Finally, since the $\mathrm{b\bar{b}WW}$ and $\mathrm{b\bar{b}ZZ}$ channels were not studied by the ATLAS collaboration, it was assumed that similar results to those of the CMS collaboration would eventually be obtained.

The individual channels alongside their combination are shown in table 5.9. A final expected significance of 4.0σ is obtained (4.5σ when ignoring systematic uncertainties)

	Significance [σ]			
Channel	Stat. uncertainty only		Stat. and Syst. uncertainties	
	CMS	ATLAS	CMS	ATLAS
$b\bar{b}b\bar{b}$	1.2	1.4	0.95	0.61
$b\bar{b}\tau\tau$	1.6	2.5	1.4	2.1
$b\bar{b}WW(\ell\nu\ell\nu)$	0.59	-	0.56	-
$b\bar{b}\gamma\gamma$	1.8	2.1	1.8	2.0
$b\bar{b}ZZ(\ell\ell\ell\ell)$	0.37	-	0.37	-
Combined	2.8	3.5	2.6	3.0
	Combined		Combined	
	4.5		4.0	

Table 5.9. | Summary table of the individual channels and combined expected significance, with and without systematical uncertainties being considered for both CMS and ATLAS experiments [126].

when assuming the standard model signal. The 68% confidence interval on κ_λ shrinks to [0.52,1.5]. The contribution from the different channels is shown on figure 5.16.

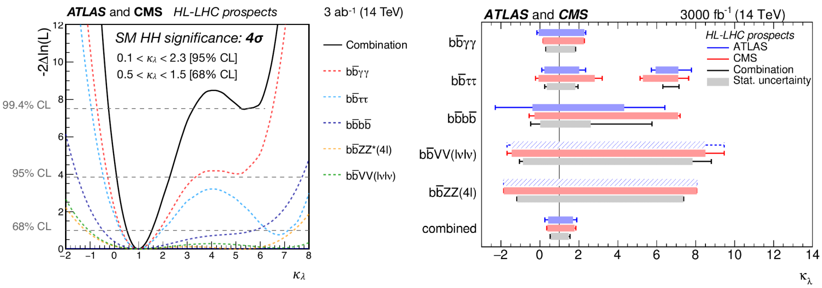


Figure 5.16. | The expected log-likelihood as a function of κ_λ for the different channels combined between CMS and ATLAS and the total combination assuming the standard model di-Higgs signal is shown on the left. The expected 95% confidence level measurements on the κ_λ from the individual channels and their combination, assuming the standard model di-Higgs production is shown on the right [126].

While this final expected significance still falls short of the 5σ threshold that would be required to claim a discovery, these studies were all done using current day data-analysis techniques. Some progress is to be expected in the coming twenty years, thus improving the physics reach of the experiment.

Some improvements have already occurred in data analysis techniques. Taking the example of the $b\bar{b}WW$ channel, new kinematic variables were derived for this channel to boost the sensitivity of the neural networks. As described in a recent publication [145], two variables that could play a role are the topness and higgsness. These new variables are defined as the sum of weighted distances in invariant mass between different high-level objects and their most probable value if they originated from top pair or di-Higgs production.

If the model-independent requirement is dropped and the standard model is assumed, a more stringent limit on κ_λ and indirect information on the di-Higgs self-coupling can be obtained through single Higgs boson measurements. Significant changes in production cross-sections will indeed occur through loop corrections and differential distributions would be impacted, especially in the $t\,\bar{t}\,H$ channel.

All things considered, it seems fair to say that the di-Higgs production discovery and λ_{HHH} measurements are well within grasp, with an upgraded machine and detectors that should rise up to the challenge.

While Higgs boson pair production is just one of the many exciting measurements that will be made feasible, it can also be seen as an indicator of the physics potential of the high-luminosity LHC and the suitability of the detector upgrades.

As clearly visible in table 5.9, this measurement will still be statistically limited at an integrated luminosity of $3000\,\mathrm{fb}^{-1}$. Keeping a detector that can efficiently record signal events is crucial to this measurement. As such, as shown on figure 5.8 for example, keeping low trigger thresholds is fundamental and it confirms how necessary the efforts to improve the Level-1 trigger are, for example by allowing the inclusion of tracking information. In addition, the systematic uncertainties being under control with events simulated with a pile-up of 200 validate further how the design choices of the detector upgrade are up to the challenge presented by the HL-LHC.

While the physics potential of the LHC is expected to be significantly enhanced by this upgrade, seeking precision measurements on the Higgs self-coupling or percent level measurements on the Higgs coupling to other particles would require totally new machines.

In the short term, the physics program of the HL-LHC could be complemented by a linear lepton collider. One candidate is the International Linear Collider which may be built in Japan over the next couple of decades. This 20 km long machine would accelerate electrons and positrons to a centre of mass energy of up to 250 GeV, allowing to reach the peak of the $e^-e^+ \rightarrow Zh$ production cross section. This clean process would allow to precisely measure the coupling of the Higgs boson to the Z boson independently on the reconstruction of the Higgs boson decay products, allowing in turn for strong constraints on the Higgs boson to invisible decay ratio. Combined with the results of the high-luminosity LHC, an uncertainty of about or less than one percent can be expected on the couplings of the Higgs boson to Z and W bosons, bottom and charm quarks, as well as τ leptons, photons and gluons.

Looking at longer time scales spanning a whole lifetime, Future Circular Colliders (FCC) at CERN are being designed [146–148]. The project envisions a 100 km tunnel hosting an electron-positron collider in a first instance. This accelerator would deliver collisions for over a decade at centre of mass energies ranging from 88 GeV to 365 GeV. This would enable, with the highest possible statistics, the study of the Z, $W\,W$, $Z\,h$ and $t\,\bar{t}$ productions to unprecedented precision levels. Ultimately, this lepton collider would be replaced by a hadron collider. This machine would allow the collision of protons to a centre of mass energy of 100 TeV, extending greatly the reach of BSM searches. In the context of di-Higgs production, this machine is expected to be able to bring a precision of 5% on the Higgs trilinear coupling to the community in 70 years.

Conclusion

The existence of the final building block of the standard model was confirmed by the discovery of the Higgs boson by the CMS and the ATLAS detectors in 2012. This discovery ultimately crowned the standard model as the most successful explanation of particle physics ever to be achieved, but this theory still falls short of providing a unified description of the fundamental interactions.

In an attempt to further understand nature's behaviour at high energies, the standard model is being studied in detail in the hope of finding hints of what may lie ahead. The new scalar sector which was recently unlocked is of particular interest in that respect, and the study of the Higgs boson and its interactions is one of the highest priorities of the community.

With all their achievements over the years, the LHC accelerator complex and related experiments have proven to be an invaluable source of experimental input. To fully exploit the physics potential of these machines, it is planned to upgrade the LHC into a high-luminosity collider, thus increasing the total integrated luminosity delivered by a factor of ten.

This document presented the upgrade design of the CMS detector which should maintain its outstanding performances over the coming decades. In the midst of the replacement of the calorimeter endcaps, an extension of the muon systems in the forward region, the introduction of a new timing detector and the overall upgrade in read-out electronics, the development of a new tracker was shown to be crucial to meet the detector's design goals.

A new outer-tracker based on the binary read-out of pairs of silicon sensors was shown to offer a high efficiency and radiation tolerance while providing the so-important tracking information at the Level-1 trigger system.

The prototyping and testing phase of the outer-tracker 2S modules, where the author's work represented a significant contribution, was presented. An in-beam test campaign on two generations of prototypes has been described in previous chapters, providing important feedback on the design choices and allowing for an integration of the different software and hardware components of the read-out system. An in-depth analysis of beam test data validated the feasibility and performances of the proposed new detector, paving the way for the construction phase.

The overall design of the upgraded CMS experiment was scrutinised in the final chapter, where its physics reach was addressed. A prospective analysis of the standard model Higgs boson pair production, with a particular focus on the $\mathrm{b\bar{b}WW}$ channel analysed by the author, was presented. A final expected significance falling just short of the symbolic 5σ threshold indicates that even this elusive process could very well be within grasp. The clear possibility of this major discovery, together with the perspective of vast improvements in the measurement of a multitude of standard model parameters provide a clear sign of an adequate upgrade design and underlines once again the high importance of the HL-LHC project as a whole.

A new era of particle physics is dawning. The scientific issues paradoxically raised by the triumph of the standard model present among the most difficult questions the field has had to answer. An in-depth study of the Higgs boson with this upgrade and future facilities is crucial if these questions are to be answered.

A. Appendix

Beam test read-out chain

A more in-depth view of the offline processing chain and the related data formats used during beam tests is presented in this additional chapter, expanding the broad description offered in section 4.1.3.

A.1. Overview

A summary sketch of the read-out system used during single 2S module beam tests is shown on figure 1.1.

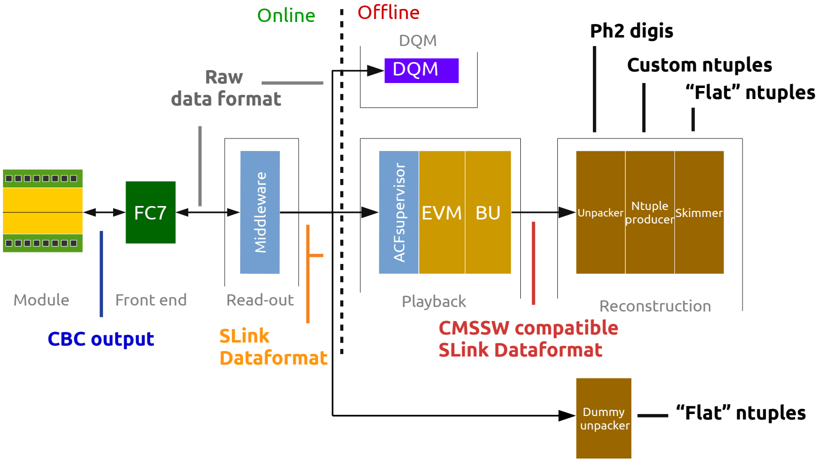


Figure 1.1. Summary view of the different steps involved in test beam data acquisition and offline processing, with the corresponding data format used.

A prototype is connected to one or more front-end cards. In this illustration, an FC7 card [149] is shown, but GLIB cards [116] can also be used. As mentioned in section 4.1.3, this card performs a series of tasks. It distributes the clock and trigger signals (both either self-generated or from an external source) to the detector, continuously gathers stub information and when a trigger is generated, reads out the corresponding event from the module. The data from the detector is combined with the corresponding trigger information and additional monitoring data before being sent downstream to the data acquisition system. In parallel to these tasks, the front-end boards are also responsible for the slow control of the prototype through commands sent using the I^2C protocol, including the configuration of the CMS binary chips.

A data acquisition computer connected to one of the front-end cards controls the system and reads out the data. This is usually done through the help of the CMS Tracker Phase2 Acquisition and Control Framework (Ph2_ACF) [150]. Also known as middelware, this software allows for an easy configuration of the whole setup, implements calibration procedures and allows both standalone and integrated data acquisition.

For an enhanced flexibility, the output of the middleware is usually directly saved to disk during beam tests. This output can be parsed in real time by a data quality monitoring system to produce fast feedback through a series of histograms stored on a remotely accessible web page.

Once a run is completed, it is then sent to the playback chain. The role of this step is to ensure the required CMSSW compatibility. Files stored to disk are parsed by a XDAQ application controlling the middleware named the ACFsupervisor [151]. This application is then linked to an event manager and builder unit, as would be the case in the final system.

These CMSSW compliant data files still contain the same information that was produced by the middleware, and the unpacker is then tasked of converting these files into collections that are easily accessible within the CMSSW framework. This would then in turn allow reconstruction algorithms to be run, either within the high-level trigger or for offline data processing.

In the context of beam test data analysis, it is usually preferable to simplify the output data format, allowing to drop the heavy CMSSW dependencies. This is done through the usage of an ntuple producer that converts the CMSSW collections into high-level `C++` object collections describing hits, clusters, stubs and additional monitoring information. A final simplification step can then be performed by a so-called skimmer that selects information of interest and stores it in plain, generic `C++` format. A test beam data analysis would then only require the usage of the ROOT [108] framework used to store the files, with no additional external dependency.

The multitude of consecutive steps in this chain, while justified in the context of the final system, is redundant if the only goal is to produce the final skimmed files. With that in mind, a light version of the unpacker, tailored for a particular test beam architecture was designed to run directly on the output of the middleware and produce the final light-weight data files. With its very limited scope, this so-called dummy unpacker allows for fast batch data processing, while the much slower full chain requires human interactions for each run in order to launch the ACFsupervisor.

A.2. Data formats

As shown on figure 1.1, a multitude of different data formats serving different purposes coexist during a beam test. In this section, a snapshot of the data formats used during the November 2017 CBC3 beam test will be described.

A.2.1. Cbc output

The first format in the chain is fixed by the CBC read-out chips. During the November 2017 beam test, this data was directly sent to an FC7 board. In later beam tests and in the final system these chips are read out by the concentrator integrated circuits. A graphical representation of the CBC3 output data streams is shown on figure 1.2.

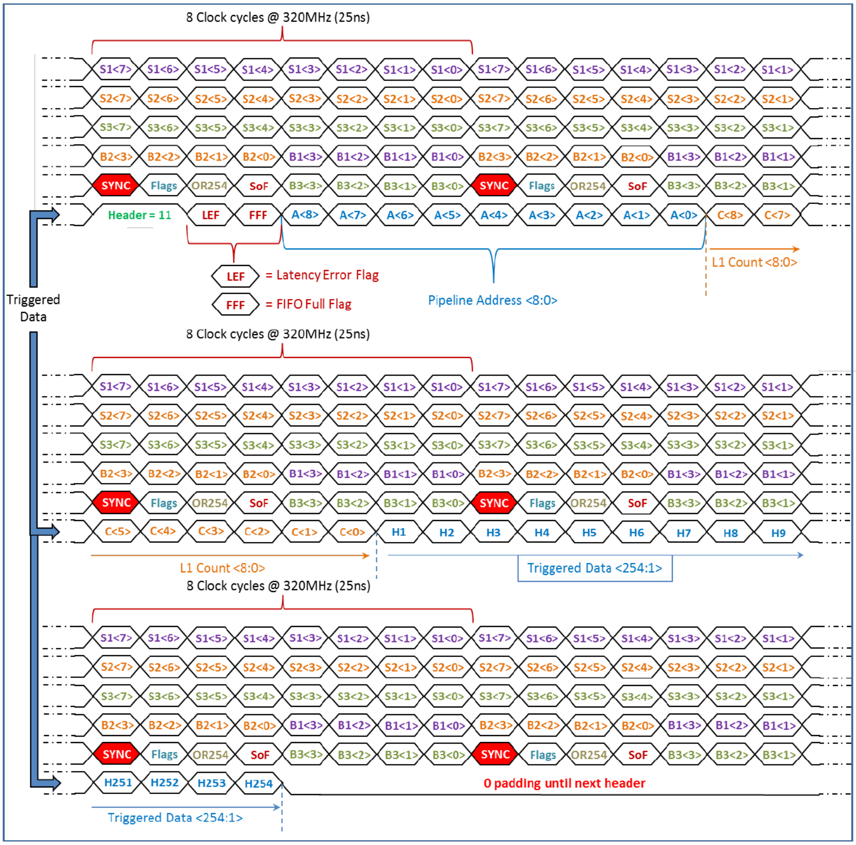


Figure 1.2. | Summary view of the CBC3 output data format [118].

A total of six lines clocked at 320 MHz are used to read out the data, five lines dedicated to continuous stub collection and the sixth to triggered data. In the CBC3 chip, a stub consists of an address and a bend code. The stub address is defined as the position of centre of the seeding cluster in half-strip units, between 2 and 255[1]. The bend code encodes the difference in position between the seeding and correlated cluster as a number between -7 and 7, as configured in an internal look-up table. With a total of eight bits for a stub address and 4 bits for a bend, transmitting the three stubs with smallest absolute bend each bunch crossing over five 320 MHz lines leaves four bits available.

---

[1] The first strip of a CBC is labelled as strip #1

These four bits are filled by a synchronisation bit always set to one, an error flag, an OR254 bit and a stub overflow bit. The error flag in the stub stream is set to one if either the latency error or the buffer overflow errors are raised. The latency error is raised if the difference between the read and write pointer of the internal data pipeline is different from the value stored in the I^2C latency register. A buffer overflow error is raised if the CBC data read-out FIFO capable of storing up to 32 events awaiting to be read out is overflown. These two error bits are also made available in the triggered data stream and in an I^2C register. The OR254 bit is set to one if any of the 254 strips of the CBC are fired, and the stub overflow bit indicates if more than three stubs were produced that clock cycle.

The last line is dedicated to triggered data. If an event is available in the data read-out FIFO, it will be transmitted on this line as soon as it is free and synchronised with a synchronisation bit. The data packet will start with a header of two bits set to one, followed by the latency error and buffer overflow flags. Then, nine bits encoding the corresponding internal pipeline address, and an additional nine bits with the trigger counter are sent. Finally, 254 bits corresponding to the 254 channels are read out. Since data packets are only sent starting at the synchronisation bit, the four last bits of 33 clock cycle long 264 bit data packet are padded with zeroes. If the output FIFO is empty, this line will transmit eight zeroes.

A.2.2. SLink data format

The front-end cards gathering data from multiple CBC chips will combine this information into data packages with the possible addition of external information such as a time phase between the trigger signal and the previous clock. The output data format of the front end can change depending on the setup and firmware version, and is commonly referred to as raw data format.

The middleware receives these data packages and store them to disk. These files are then used to produce DQM histograms, and contains all the necessary information to analyse the data during a beam test.

In the final system, a more robust and scalable data format to transmit collision data from the read-out electronics to the surface through the data concentrator cards will be required. A preliminary version of this data format is implemented in the Phase-2 tracker data format [152], also known as SLink data format. An example of an event in this data format is shown in table 1.1.

Within the CMS data acquisition system, all data packages are formed by a global header, followed by a detector payload, and ended by a global trailer [153]. A schematic view of these is shown on figure 1.3.

The global header is defined as, from most significant to least significant bit, a 4 bit beginning of event fragment identifier (BOE), a 4 bit event type identifier (Evt_ty), a 24 bit Level-1 trigger counter (LV1_id), a 12 bit bunch crossing id counter (BX_id), a 12 bit source identifier word (Source_id), 4 bits to identify the header and trailer data format version (FOV), one bit (H) set to one if another global header is following, and three reserved bits (set to zero).

1	`5100 0002 5500 3300`	Header
2	`26c0 0000 0000 0200`	Tk. header
3	`0000 0000 0000 0001`	
4	`39e0 239e 0200 0000`	
5	`0003 0000 0000 0000`	Payload
6	`0000 0000 0000 0000`	
7	`0000 0000 0000 0000`	
8	`0000 000c 0000 0000`	
9	`0000 0000 0000 0000`	
10	`0000 0000 0000 0000`	
11	`0000 0000 0000 0000`	
12	`0000 0000 0000 0000`	
13	`0000 0000 0000 0000`	
14	`0833 8000 0000 0000`	Stubs
15	`0000 0000 0000 000b`	Cond. data
16	`0000 4f01 0000 0026`	
17	`0000 5001 0000 0002`	
18	`0000 0101 0000 0023`	
19	`0000 0001 0000 0040`	
20	`0000 0c01 0000 0000`	
21	`0000 1201 0000 005e`	
22	`0000 1401 0000 0000`	
23	`0000 1301 0000 0000`	
24	`0000 0004 0000 0000`	
25	`0010 0005 0000 012c`	
26	`ff00 0003 0000 0000`	
27	`a000 001b 031e 0070`	Trailer

Table 1.1. Event example from run 132 of the November 2017 beam test, written in the SLink data format.

In our example, the header `5100 0002 5500 3300` translates as :

BOE = 5, signifying it is the first event fragment
Evt_ty = 1, used to designate physics triggers
LV1_id = 2, this event was the second trigger of a run
BX_id = `0x550`, this bunch crossing is the 1360th of the lhc orbit
Source_id = `0x33`, the fragment was produced by fed#51
FOV = 0, this is the first version of the data format
H =0, this is the last header word.

The global trailer, in the same most significant to least significant bit order, is made of four bits identifying the end of event fragment (EOE), 4 reserved bits, 24 bits containing the event length (Evt_lgth), 16 bits used for cyclic redundancy checks (CRC), a transmission error bit (C), a fed number error bit (F), 2 reserved bits, 4 event status bits (Evt_stat), 4 bits from the trigger throttling system (TTS), a bit set to one if another trailer is following (T), a bit indicating if the slink hardware changed the CRC value (R), and two reserved bits.

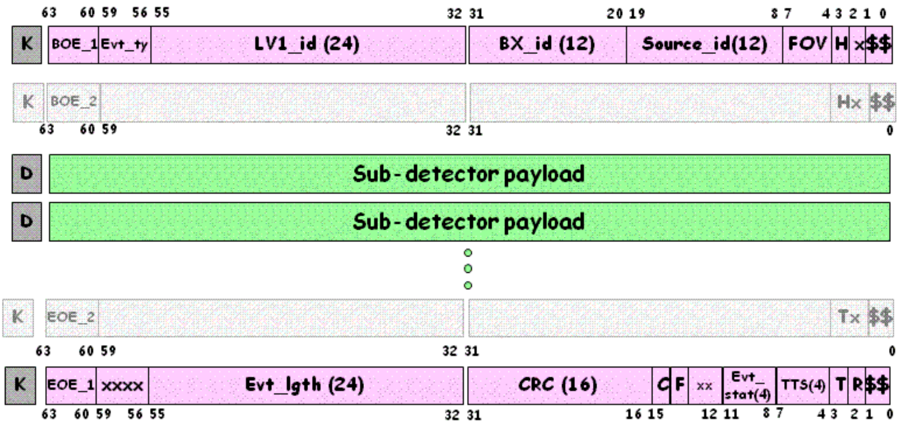


Figure 1.3. Global CMS DAQ header and trailer definitions [153].

In our example, the trailer `a000 001b 031e 0070` stands for :

EOE	= `0xa`,	signifying it is the first event fragment
Evt_length	= `0x1b`,	correctly indicating the size of the event is 27
CRC	= `0x31e`	computed using the CRC-16-IBM method
C,F	= `0`,	expected when no transmission hardware is used
Evt_stat	= `0`,	of no particular meaning in this context
TTS	= `7`,	of no particular meaning in this context

The tracker data is located between these header and trailer words, and is split into four parts : a header, the Front-End payload, stub information and condition data.

The header starts with two words containing, in the usual most significant to least significant bit order, 4 data format bits, 2 header format bits, 4 event type bits (2 bits for payload type, 1 bit to indicate if condition data is present and 1 bit to indicate is the data is real or fake), 30 DTC status bits, 16 bits containing the number of CBCs and 72 front-end status bits. Then, the header is completed by CBC status bits that can be of various length and padded with zeroes to a 64 bit word.

In the read-out mode used during the November 2017 beam test, using the data format version 2 and a full-debug header format, 20 CBC status bits are used per chip containing the pipeline and overflow error bits, 9 pipeline address bits and 9 Level-1 counter bits.

In our example, the three tracker header words are `26c0 0000 0000 0200`, `0000 0000 0000 0001` and `39e0 239e 0200 0000`, indicating that :

Data format = `0b10` , second version of the data format
Header format = `0b01` , standing for full debug header
Event type = `0b1011` , unsparsified real data with condition data
DTC status = `0` , no DTC information
Number of CBC = `2` , 2 CBC connected
Front-end status = `1` , the first front-end is connected
CBC0 status = `0x39e02` , split into :
  `0b00` no CBC errors
  `0x1cd` pipeline address
  `0x002` second Level-1 trigger
CBC1 status = `0x39e02` identical to the first CBC.

This tracker header is followed by a payload that is further split into Front-End payload, stub data and condition data.

In the unsparsified read-out mode used during the beam tests described in this document, the Front-End payload starts with a 16 bit word indicating each CBC present in this payload, followed by 254 bits of data for each CBC, each representing a strip. This payload is then padded to 64 bit words with zeroes.

In our example event, the words five to thirteen give this front-end payload. The first sixteen bits are `0b0000000000000011`, indicating that two CBCs are connected. The rest of the payload is given by `0b...00011000...` where the only two bits set to one are the bits #204 and #205 if we start counting at zero. Since the even bits store the bottom sensor information, and the odd bits store the top sensor information, this translates into the strip #102 being fired in both sensors.

The stub information is contained after the tracker payload. For each front-end, the stub information starts with a 6 bit header, the first five used to declare the number of stubs to be read out, and the last bit to differentiate between 2S (0) or PS (1) data. In the case of 2S modules, each stub takes 16 bits, with the first four identifying the CBC, the next eight containing the stub position on the chip, with the first strip being assigned the number 1, in half strip units. The last four bits gives the stub bend code, as defined in the internal chip look-up table.

In our example, the only stub word is `0833 8000 0000 0000`, translated into :

FE header = `0b00001 0`, declaring that one stub is present, originating from a 2S module
Stub #1 = `0x0ce0`, split into :
  `0x0` CBC#0
  `0xce` Stub position = 206, corresponding to the strip 102
  `0x0` Null bend code

Finally, if the corresponding bit in the header is set to one, the tracker payload is concluded by condition data providing additional external information. The first word stores the number of condition data to be read-out, followed by the corresponding number of words, with the 32 first bits defining a key, and the last 32 the corresponding value.

The key parameters follow specific conventions to indicate what the value stored is. The first eight bit are either set to the front-end id or to 0xFF in the case of external data. The four next bits can indicate either the corresponding CBC or sensor, and the twelve next bits can indicate the internal CBC I^{2}C page and register. Finally, the last eight bits store the data type (UID). Some commonly used definitions are the cbc registers (UID=1), the trigger phase (UID=3), the beam angle (UID=4) and bias voltage (UID=5). UID values between `0x80` and `0xFF` are user defined.

In our example, the first word, `0000 0000 0000 000b` indicate that 11 condition data words are in the payload. They are given by the following :

`0000 4f01 0000 0026` → I^{2}C reg. `0x4f` (VCTH [1/2])

`0000 5001 0000 0002` → I^{2}C reg. `0x50` (VCTH [2/2])

`0000 0101 0000 0023` → I^{2}C reg. `0x01` (Trigger latency [2/2])

`0000 0001 0000 0040` → I^{2}C reg. `0x00` (Misc. settings and trigger latency [1/2])

`0000 0c01 0000 0000` → I^{2}C reg. `0x0c` (Misc. settings)

`0000 1201 0000 005e` → I^{2}C reg. `0x12` (Stub logic configuration [1/3])

`0000 1401 0000 0000` → I^{2}C reg. `0x14` (Stub logic configuration [3/3])

`0000 1301 0000 0000` → I^{2}C reg. `0x13` (Stub logic configuration [2/3])

`0000 0004 0000 0000` → Perpendicular beam

`0010 0005 0000 012c` → Bias voltage = 300 V

`ff00 0003 0000 0000` → TDC phase = 0

These condition data can represent a significant fraction of the overall data payload for information that is mostly identical in every event, but they allow for a cross-check of the configuration of the chip and the parameters under which the data was taken. They often play a crucial role in understanding issues in data taking.

A.2.3. Offline data formats

In order to be able to use this information either in a software trigger or for offline analysis, these files are transformed into CMSSW collections. This is the role of the unpacker.

During beam tests, this unpacker converts files stored to disk by the builder unit into collections of `Phase2TrackerDigi`, `Phase2TrackerCluster1D`, `Phase2TrackerStub` and `Phase2TrackerFEDFEDebug`.

The first two collections, `Phase2TrackerDigi` and `Phase2TrackerCluster1D` are used if the SLink data is unsparsified or sparsified respectively to store either hit or cluster collections. As the name suggests, the `Phase2TrackerStub` are used to store stubs that include a central seeding cluster position and a bend. Finally, the `Phase2TrackerFEDFEDebug` collection stores header information, including the individual CBC status words. Alongside these collections, the full payload is stored in a `FEDRawDataCollection` object. Within the CMSSW framework, tools are made available to easily extract any information from these objects and to convert local coordinates into global cms coordinates through the usage of full detector geometries.

In the context of a beam test, the CMSSW machinery is usually not needed and adds unnecessary complexity to the data analysis. To ease development and reduce the computing time involved, these CMSSW collections are converted into standalone `C++` objects. This is performed by the Ntuple producer. In addition, external data, such as telescope tracks, can be added to the file, and some general offline reconstruction can be performed.

In this data format, events are stored in a dedicated class, as defined in figure 1.4. These objects contain the decoded header and condition data information, devices hit (in unsparsified mode), clusters (in sparsified mode) and stub collections. From the hit information, offline clusters and offline stub collections are also produced. Finally, telescope tracks are integrated into these files.

The CBC and hit object are quite straightforward, storing the CBC header debug information and the strip being fired respectively. The cluster objects store the position of the first hit fired and the size of the corresponding cluster, and has a method to return the central position of the cluster, as often needed when analysing the data. In unsparsified mode, only offline reconstructed clusters are available (in the `offlineClusters` collection).

Finally, the stub objects store the position of the seeding cluster, a bend code and bend information. They originate either from the prototype itself (in the `cbcClusters` collection), or are emulated offline from the available clusters and chip parameters (in the `offlineStubs` collection). Since the look-up table is not stored in the condition data, the bend from the CBC stubs are produced using the default look-up table parameters.

This data format offers a nice compromise with the flexibility and power offered by these `C++` objects, while not requiring the CMSSW framework in order to be read. On the other side, the definition of all these objects are still required, and the modularity of this format makes it heavier than strictly necessary for beam tests involving a unique module. Furthermore, much of the condition data and header information can be ignored in most analyses.

A last simplification step is thus available through the usage of the skimmer, producing so called flat-tuples. The events are stored through a series of standard `C++` vectors. The flat-tuple data format used in the November 2017 beam test is shown on figure 1.5.

```
uint run, lumiSection, event, time, unixtime;  General event information

bool isSparisified;                            Decoded event header
uint dataFormatVersion, condData, debugMode,
     readoutMode, dataType, glibStatusCode,
     numberOfCBC;

map <uint, uint> conddatamap;                  Raw condition data map
uint tdcPhase   , HVsettings, DUTangle,        and decoded values
     offset1Reg , offset2Reg, hipRegReg,
     vcth1Reg   , vcth2Reg  , compReg,
     logicReg   , triggerLatencyReg,
     stubLatency, triggerLatency;

map <uint  , vector<cbc>> cbcs;                CBC header info
map <string, vector<hit>> dutHits;             Hit collections
map <string, vector<cluster>> cbcClusters;     Cluster collections
map <string, vector<stub>> cbcStubs;           Stub collections
map <string, vector<cluster>> offlineClusters; Reconstructed clusters
map <string, vector<stub>> offlineStubs;       Reconstructed stubs
vector<Track>  tracks;                         External tracks
```

Figure 1.4. | Event object definition, as used in November 2017 Ntuples. The specific integer sizes and additional class methods are ignored here.

```
vector <int> hit0, hit1;                       Hits in the two sensors
vector <int> stub, bendCode, bend;             Stubs from the DUT
vector <int> stubSim, bendCodeSim, bendSim;    Emulated stubs

uint CbcL1A0, Pipeline0, Error0;               CBC status from the event
uint CbcL1A1, Pipeline1, Error1;               header

uint L1A, TDC;                                 Trigger counter and TDC phase
```

Figure 1.5. | Flat tuple definition, as used in the November 2017 beam test. The specific integer sizes are ignored here.

B. Appendix

Louvain beam test

A few additional details and illustrations of the Louvain beam test described in section 4.1.6 are gathered in this appendix.

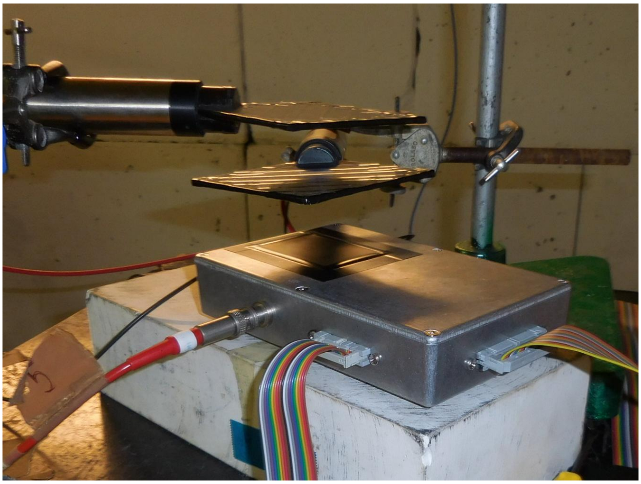


Figure 2.1. | Picture of a module in its protective enclosure being calibrated using cosmic muons during the Louvain 2017 beam test.

A photograph of the module enclosed in his protective box taking data in cosmic mode is shown on figure 2.1. The protective box is clearly visible here.

Holes were carved out of a 188×120×33mm Hammond 1590DD [154] cast aluminium instrument enclosure to allow for the High and Low Voltage power supply, and data connectors. Two additional windows carved above and below the sensors and covered with black wrapping minimise the total material needing to be crossed by the beam.

The trigger system can also be seen in detail. It is comprised of six elements, bought for the occasion and listed in table 2.1.

Qty	Item type	Item description
2	Scintillators	100x100x2 mm thick EJ212 [155] plastic scintillator wrapped in black light tight wrapping
2	Photomultipliers	R9800 Hamamatsu PMT [156]
2	Preamplifiers	CAEN A1424 scintillation preamplifiers [157]

Table 2.1. | Components of the trigger system

Two 100×100×2mm tiles made of EJ212 [155] plastic scintillator were used as active material. The geometry of these scintillators is designed to match the size of a full sensor. The small thickness is crucial when low energy protons are used. The EJ212 material is a common choice for general purpose plastic scintillator tiles, with a fast timing ideal for triggering purposes.

The light generated in a scintillator tile is collected at one of its corners by a R9800 Hamamatsu [156] photomultiplier tube, offering a fast response time, also ideal for a trigger system.

Before being treated by general purpose NIM and VME electronic cards, the signal produced by these photomultipliers are sent to CAEN A1424 scintillation preamplifiers [157], not shown on the picture. As for other components, this preamplifier offers the fast response required.

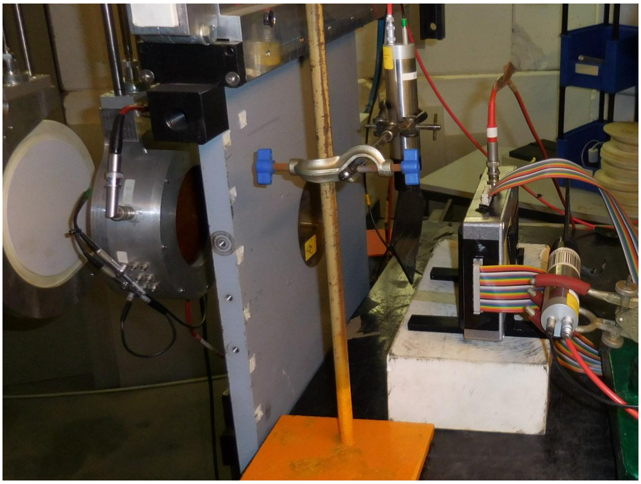


Figure 2.2. | Picture of a module in its protective enclosure ready to take data during the Louvain 2017 beam test.

The same items placed to record data in beam configuration are shown on figure 2.2. The two trigger scintillators are placed on each sides of the device under test in perpendicular directions. To the left of these detectors, the collimator plate with a rectangular opening can be seen. Finally, a wire chamber used to calibrate the beam being delivered and the first degrador disk are also visible on this photograph.

A wider view of the experimental area is shown on figure 2.3. The five degrador disks are visible on the left of this picture, followed by the wire chamber. The laser guide used to align the devices with the beam can also be seen on the collimator.

Figure 2.3. | Picture of the experimental area of the Light Ion Facilities during the Louvain 2017 beam test.

Acronyms

V_{CTH} Voltage Comparator Threshold 86, 89–91, 93, 105

Ph2_ACF CMS Tracker Phase2 Acquisition & Control Framework 140

GEANT4 GEometry ANd Tracking 52, 111

2S 2 Strip modules of the upgraded Outer Tracker 66–68, 84

ADC Analog to Digital Converter 32, 35, 76, 77, 79

ALCT Analog Local Charge Track 79

ALICE A Large Ion Collider Experiment 22

ANN Artificial Neural Network 53, 116, 125

APD Avalanche photodiode 72, 73

APSP Analog Pulse Shape Processor 31

APV25 Analog Pipeline Voltage strip tracker read-out chip 31–34

ASIC Application-Specific Integrated Circuit 31, 67, 71, 77

ATLAS A Toroidal LHC ApparatuS 22, 60, 134, 137

BCMS Alternative bunch shaping scheme in the PS 24

BDT Boosted Decision Tree 53, 55, 113

BE Back-End 65, 88

BMTF Barrel Muon Track Finder 41

BSM Beyond the SM 18, 135

BTL Barrel Timing Layer 71, 72

BU Builder Unit PC 42

CBC CMS Binary Chip 67

CBC2 CMS Binary Chip version 2 83–88, 90, 91, 93, 94, 96, 97, 101, 102

CBC3 CMS Binary Chip version 3 90, 102, 103, 105, 106, 140, 141

CE Calorimeter Endcap 74, 75

CERN European research facility for high-energy physics, located close to Geneva 21–23, 43, 93, 107, 135

CFEB Cathod Front End Board, Board reading out CSC strips 79

CHS Charged-Hadron Subtraction 49, 50

CIC Concentrator Integrated Circuit 67, 84, 106

CMS Compact Muon Solenoid 8, 17, 21, 27–30, 33, 39, 40, 47, 50–52, 58, 63–65, 67, 77–79, 81, 83, 88, 96, 109, 134, 137, 140, 142, 144

CMSSW CMS software 88, 140

CNM Centro Nacional de Microelectrónica 94, 95

CRC Centre de Recherche du Cyclotron 96, 98

CSC Cathode strip chamber 37–40, 44, 64, 78–80

CTF Combinatorial Track Finder 43

CWD Cluster-width Discrimination 86, 87

DAQ Data Acquisition 41, 42, 64, 65, 84, 88, 98, 144

DCFEB Digital Cathod Front End Board, Upgraded board to read out CSC strips 79

DQM Data Quality Monitoring 88

DT Drift tube 37, 39, 40, 44, 64

DUT Device Under Test 87, 98

EB Electromagnetic calorimeter barrel 34, 35

EC-E Calorimeter Endcap - Electromagnetic 76

EC-H Calorimeter Endcap - Hadronic 76

ECAL Electromagnetic calorimeter 34, 36, 40, 41, 45–48, 51, 64, 72, 81

EE Electromagnetic calorimeter endcap 34, 35

EFT Effective Field Theory 19

EMTF Endcap Muon Track Finder 41, 80

ES Electromagnetic calorimeter preshower 35

ETL Endcap Timing Layer 71, 72

FE Front-End 77, 78, 88

FED Front-End Driver 32, 42

FEROL Front-End Readout Optical Link 42

FIFO First In First Out 39, 77, 142

FPGA Field-Programmable Gate Array 32, 39, 64, 68, 73, 78, 79

FU Filter Unit PC 42

GEM Gas Electron Multiplier 79–81

GLIB Gigabit Link Interface Board 98

GSF Gaussian Sum Filter 45–47

HB Hadron calorimeter barrel 35, 37

HCAL Hadron calorimeter 35–37, 40, 45, 47, 48, 51, 64, 74, 81

HE Hadron calorimeter endcap 35, 37

HF Forward hadron calorimeter 35, 37, 51

HGCAL High Granularity Calorimeter 80

HIP Highly Ionizing particle 33, 102

HL-LHC High-Luminosity Large Hadron Collider, upgrade to the LHC to increase its instantenous luminosity by one order of magnitude 62, 63, 71, 73, 74, 77, 78, 94, 109, 112, 118, 122, 138

HLT High-Level Trigger 39, 43, 65, 80

HO Outer hadron calorimeter 35, 37

HPD Hybrid photodiode 37, 74

HPK Hamamatsu Photonics K.K. 94

HSCP Heavy Stable Charged Particles 79, 80

I^{2}C Inter-Integrated Circuit 87, 88, 140, 142, 146

iRPC Improved RPC 79, 80

LGAD Low Gain Avalanche Detector 71

LHC Large Hadron Collider 19, 21, 22, 24, 25, 27, 28, 31, 35, 60–63, 65, 81, 86, 96, 98, 102, 103, 109, 135, 137

LHCb LHCb 79

LINAC 2 Linear Accelerator 2 23, 61

LINAC 4 Linear Accelerator 4 61

LIU LHC injectors upgrade 61

lpGBT Low Power GigaBit Transceiver 67, 71, 73, 77, 84

MIP Minimum Ionizing Particle 76, 77

MLP Multilayer Perceptron 55, 56, 125

MPA Macro-Pixel ASIC 67

MTD MIP timing detector 71

MVA Multivariate Analysis 48

NIM Nuclear Instumentation Module 98

OMTF Overlap Muon Track Finder 41

OOT Out-of-time 77

p.d.f. Parton distribution function 15, 16

P5 Point 5, Location of the CMS detector 88

PCB Printed Circuit Board 75, 76

PF Particle-Flow 45–47, 49, 50, 64

PROC Pixel Read-Out Chip 66

PS Pixel-Strip modules of the upgraded Outer Tracker 67

PS Proton Synchrotron 23, 24, 61, 67, 106

PSB Proton Synchrotron Booster 24, 61

PUPPI Pile Up Per Particle Identification 49, 50, 64

QCD Quantum Chromodynamics 11, 15, 39

QED Quantum Electrodynamics 11

RMS Root Mean Square 71

RPC Resistive plate chamber 37, 39, 40, 44, 64, 80

SiPM Silicon Photomultiplier 71, 74–76

SL Drift tube super layer 38

SPS Super Proton Synchrotron 24, 25, 61, 93

SSA Short Strip ASIC 67

SST Silicon Strip Tracker 30, 31

TDC Time to Digital Converter 77, 88

TEC Tracker End Caps 31, 33

TIB Tracker Inner Barrel 31, 33

TID Tracker Inner Disks 31, 33

TOB Tracker Outer Barrel 31, 33, 34

TOFHIR Time-of-flight, High Rate 71

VBF Vector Boson Fusion 111

VFE Very Front End 35, 72

VME Versa Module Europa bus based electronics card 32, 98

VTRx+ Versatile TRansceiver plus 67, 84, 106

XDAQ Cross daq, CMS online software 88, 140

References

[1] W. Adam et al., "Beam test performance of prototype silicon detectors for the Outer Tracker for the Phase-2 Upgrade of CMS", *Journal of Instrumentation* **15** (mar, 2020) P03014–P03014. `doi:10.1088/1748-0221/15/03/p03014`.

[2] W. Adam et al., "Test beam demonstration of silicon microstrip modules with transverse momentum discrimination for the future CMS tracking detector", *Journal of Instrumentation* **13** (mar, 2018) P03003–P03003. `doi:10.1088/1748-0221/13/03/p03003`.

[3] The CMS Collaboration, "The Phase-2 Upgrade of the CMS Tracker", Technical Report CERN-LHCC-2017-009. CMS-TDR-014, Jun, 2017. `https://cds.cern.ch/record/2272264`.

[4] The CMS Collaboration, "The CMS Tracker Upgrade for the High Luminosity LHC", Technical Report CMS-CR-2018-086, Jun, 2018. `http://cds.cern.ch/record/2650717`.

[5] O. Bondu et al., "Study of nonresonant Higgs boson pair production in the bblnulnu final state for an upgraded CMS detector at the HL-LHC", technical report, December, 2018.

[6] The CMS Collaboration, "Prospects for HH measurements at the HL-LHC", Technical Report CMS-PAS-FTR-18-019, 2018. `https://cds.cern.ch/record/2652549`.

[7] E. M. Cepeda et al., "Higgs physics at the HL-LHC and HE-LHC", *CERN Yellow Reports: Monographs* **7** (2019), no. 0, 221.

[8] Peskin, Michael E and Schroeder, Daniel V, "An introduction to quantum field theory". Westview, Boulder, CO, 1995. `http://cds.cern.ch/record/257493`.

[9] Mandl, Franz and Shaw, Graham G, "Quantum field theory; 2nd ed.". Wiley, New York, NY, 2010. `https://cds.cern.ch/record/1236742`.

[10] Christopher Smith, "Introduction to the Standard Model", 2014. Lecture notes.

[11] S. Roy Choudhury et al., "Updated Bounds on Sum of Neutrino Masses in Various Cosmological Scenarios", *JCAP* **09** (2018) 017. `doi:10.1088/1475-7516/2018/09/017`, `arXiv:1806.10832`.

[12] The Particle Data Group Collaboration, "Review of Particle Physics", *Phys. Rev. D* **98** (Aug, 2018) 030001. `doi:10.1103/PhysRevD.98.030001`.

[13] The CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Physics Letters B* **716** (2012), no. 1, 30 – 61. `doi:https://doi.org/10.1016/j.physletb.2012.08.021`.

[14] G. Aad et al., "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Physics Letters B* **716** (Sep, 2012) 1–29. `doi:10.1016/j.physletb.2012.08.020`.

[15] The CMS Collaboration, “Measurements of properties of the Higgs boson decaying into the four-lepton final state in pp collisions at $\sqrt{s} = 13$ TeV”, *JHEP* **11** (2017) 047. `doi:10.1007/JHEP11(2017)047`, `arXiv:1706.09936`.

[16] The ATLAS Collaboration, “Measurement of the Higgs boson mass in the $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$ channels with $\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector”, *Phys. Lett. B* **784** (2018) 345–366. `doi:10.1016/j.physletb.2018.07.050`, `arXiv:1806.00242`.

[17] The ATLAS and CMS collaborations, “Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments”, *Phys. Rev. Lett.* **114** (2015) 191803. `doi:10.1103/PhysRevLett.114.191803`, `arXiv:1503.07589`.

[18] The CMS Collaboration, “Combined measurements of Higgs boson couplings in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *Eur. Phys. J. C* **79** (Sep, 2018) 421. 67 p. `doi:10.1140/epjc/s10052-019-6909-y`.

[19] The ATLAS Collaboration *Phys. Rev. D* **101** (Jan, 2020) 012002. `doi:10.1103/PhysRevD.101.012002`.

[20] S. Höche, “Introduction to parton-shower event generators”, Technical Report arXiv:1411.4085. SLAC-PUB 16160, Nov, 2014. `https://cds.cern.ch/record/1970716`.

[21] J. Alwall et al., “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”, *Journal of High Energy Physics* **2014** (Jul, 2014). `doi:10.1007/jhep07(2014)079`.

[22] C. Oleari, “The POWHEG BOX”, *Nuclear Physics B - Proceedings Supplements* **205-206** (Aug, 2010) 36–41. `doi:10.1016/j.nuclphysbps.2010.08.016`.

[23] S. Weinzierl, “Introduction to Monte Carlo methods”, 2000. `arXiv:hep-ph/0006269`.

[24] T. Sjöstrand et al., “An introduction to PYTHIA 8.2”, *Computer Physics Communications* **191** (Jun, 2015) 159–177. `doi:10.1016/j.cpc.2015.01.024`.

[25] The CMS Collaboration, “Summaries of CMS cross section measurements”. `https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsCombined`, 2019.

[26] G. Branco et al., “Theory and phenomenology of two-Higgs-doublet models”, *Physics Reports* **516** (2012), no. 1, 1 – 102. `doi:https://doi.org/10.1016/j.physrep.2012.02.002`.

[27] L. Randall et al., “Large Mass Hierarchy from a Small Extra Dimension”, *Phys. Rev. Lett.* **83** (Oct, 1999) 3370–3373. `doi:10.1103/PhysRevLett.83.3370`.

[28] G. Isidori et al., “Higgs form factors in associated production”, *Journal of High Energy Physics* **2014** (Feb, 2014). `doi:10.1007/jhep02(2014)082`.

[29] B. Grzadkowski et al., “Dimension-six terms in the Standard Model Lagrangian”, *Journal of High Energy Physics* **2010** (Oct, 2010). `doi:10.1007/jhep10(2010)085`.

[30] G. N. Remmen et al., “Consistency of the standard model effective field theory”, *Journal of High Energy Physics* **2019** (Dec, 2019). `doi:10.1007/jhep12(2019)032`.

[31] G. Buchalla et al., “A systematic approach to the SILH Lagrangian”, *Nuclear Physics B* **894** (May, 2015) 602–620. `doi:10.1016/j.nuclphysb.2015.03.024`.

[32] T. L. H. C. S. W. Group et al., "Handbook of LHC Higgs Cross Sections: 3. Higgs Properties", 2013. `arXiv:1307.1347`.

[33] G. Brianti et al., "Summary report", *ECFA-CERN Workshop on Large Hadron Collider in the LEP tunnel* (Apr, 1984) 78 p. `doi:10.5170/CERN-1984-010-V-1.1`.

[34] Y. Baconnier et al., "LHC: the Large Hadron Collider accelerator project". CERN, Geneva, 1993. `http://cds.cern.ch/record/257706`.

[35] The LHC Study Group Collaboration, "The Large Hadron Collider: conceptual design", Technical Report CERN-AC-95-05-LHC, Oct, 1995. `http://cds.cern.ch/record/291782`.

[36] E. Mobs, "The CERN accelerator complex. Complexe des accélérateurs du CERN", technical report, Jul, 2016. `https://cds.cern.ch/record/2197559`. General Photo.

[37] S. White et al., "First Luminosity Scans in the LHC", *1st International Particle Accelerator Conference* (Jun, 2010) 3 p.

[38] S. van der Meer, "Calibration of the effective beam height in the ISR", Technical Report CERN-ISR-PO-68-31. ISR-PO-68-31, 1968. `https://cds.cern.ch/record/296752`.

[39] The CMS Collaboration, "CMS Luminosity - Public results", 2019. `https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults`.

[40] I. Neutelings, "example_spherical_coordinates". `https://wiki.physik.uzh.ch/cms/latex:example_spherical_coordinates`, 2017.

[41] The CMS Collaboration, T. Sakuma, "Cutaway diagrams of CMS detector". `http://cds.cern.ch/record/2665537`, May, 2019.

[42] F. Kircher et al., "Status report on the CMS superconducting solenoid for LHC; 1998 ed.", Technical Report CMS-CR-1998-019. DAPNIA-STCM-98-011, Sep, 1998. `http://cds.cern.ch/record/384232`.

[43] The CMS Collaboration, "The CMS Phase-1 pixel detector – experience and lessons learned from two years of operation", Technical Report CMS-CR-2019-019. 07, Feb, 2019. `https://cds.cern.ch/record/2687017`.

[44] M. J. French et al., "Design and results from the APV25, a deep sub-micron CMOS front-end chip for the CMS tracker", *Nucl. Instrum. Methods Phys. Res., A* **466** (2001), no. 2, 359–65. `doi:10.1016/S0168-9002(01)00589-7`.

[45] L. Jones, "APV25-S1: User guide version 2.2". RAL Microelectronics Design Group, Chilton, 2001. `https://cds.cern.ch/record/1069892`.

[46] S. A. Baird et al., "Design of the front-end driver card for CMS silicon microstrip tracker readout", *6th Workshop on Electronics for LHC Experiments* (2000). `doi:10.5170/CERN-2000-010.444`.

[47] The CMS Collaboration, "Operational experience with the Silicon Strip Tracker at the CMS experiment", Technical Report CMS-CR-2019-012, Feb, 2019. `http://cds.cern.ch/record/2689274`.

[48] The CMS Collaboration, "Silicon Strip Tracker Performance results 2018", technical report, Sep, 2018.

[49] W. Adam et al., “The effect of highly ionising particles on the CMS silicon strip tracker”, *Nucl. Instrum. Methods Phys. Res., A* **543** (2005) 463–482. doi:10.1016/j.nima.2004.11.049.

[50] The CMS Collaboration, “The CMS experiment at the CERN LHC. The Compact Muon Solenoid experiment”, *JINST* **3** (2008) S08004. 361 p. doi:10.1088/1748-0221/3/08/S08004.

[51] The CMS Collaboration, “CMS ECAL Response to Laser Light”, technical report, Mar, 2019. http://cds.cern.ch/record/2668200.

[52] The CMS Collaboration, “Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV”, *JINST* **13** (Apr, 2018) P06015. 53 p. doi:10.1088/1748-0221/13/06/P06015.

[53] The CMS Collaboration, “Technical proposal for the upgrade of the CMS detector through 2020”, Technical Report CERN-LHCC-2011-006. LHCC-P-004, Jun, 2011. https://cds.cern.ch/record/1355706.

[54] The CMS Collaboration, “CMS Technical Design Report for the Level-1 Trigger Upgrade”, Technical Report CERN-LHCC-2013-011. CMS-TDR-12, Jun, 2013. https://cds.cern.ch/record/1556311.

[55] The CMS Collaboration, “The Phase-2 Upgrade of the CMS DAQ Interim Technical Design Report”, Technical Report CERN-LHCC-2017-014. CMS-TDR-018, Sep, 2017. https://cds.cern.ch/record/2283193.

[56] CMS Collaboration, “Particle-flow reconstruction and global event description with the CMS detector”, *JINST* **12** (2017), no. 10, P10003. doi:10.1088/1748-0221/12/10/P10003.

[57] The CMS Collaboration, “Alignment of the CMS tracker with LHC and cosmic ray data”, *JINST* **9** (Mar, 2014) P06009. 56 p. doi:10.1088/1748-0221/9/06/P06009. Published in JINST.

[58] The CMS Collaboration, “CMS Tracking POG Performance Plots for 2017 Dataset”. https://twiki.cern.ch/twiki/bin/view/CMSPublic/TrackingPOGResults2017.

[59] The CMS Collaboration, “Description and performance of track and primary-vertex reconstruction with the CMS tracker”, *JINST* **9** (May, 2014) P10009. 80 p. doi:10.1088/1748-0221/9/10/P10009. Comments: Replaced with published version. Added journal reference and DOI.

[60] The CMS Collaboration, “CMS Tracking POG Performance Plots For 2017 with PhaseI pixel detector”. https://twiki.cern.ch/twiki/bin/view/CMSPublic/TrackingPOGPerformance2017MC.

[61] The CMS Collaboration, “Electron and Photon performance in CMS with the full 2017 data sample and additional 2016 highlights for the CALOR 2018 Conference”, technical report, May, 2018. https://cds.cern.ch/record/2320638.

[62] The CMS Collaboration, “Electron and Photon performance in CMS with the data sample collected in 2016-17-18 at 13 TeV for the ICHEP 2018 Conference”, technical report, Jul, 2018. http://cds.cern.ch/record/2629363.

[63] M. Cacciari et al., "The anti-ktjet clustering algorithm", *Journal of High Energy Physics* **2008** (Apr, 2008) 063–063. `doi:10.1088/1126-6708/2008/04/063`.

[64] D. Bertolini et al., "Pileup per particle identification", *Journal of High Energy Physics* **2014** (Oct, 2014). `doi:10.1007/jhep10(2014)059`.

[65] The CMS Collaboration, "Pileup mitigation at CMS in 13 TeV data", Technical Report arXiv:2003.00503. CMS-JME-18-001-003, Mar, 2020. `https://cds.cern.ch/record/2711736`.

[66] C. Weiser, "A Combined Secondary Vertex Based B-Tagging Algorithm in CMS", Technical Report CMS-NOTE-2006-014, Jan, 2006. `http://cds.cern.ch/record/927399`.

[67] The CMS Collaboration, "Jet algorithms performance in 13 TeV data", Technical Report CMS-PAS-JME-16-003, 2017. `https://cds.cern.ch/record/2256875`.

[68] The CMS Collaboration, "Performance of b tagging algorithms in proton-proton collisions at 13 TeV with Phase 1 CMS detector", technical report, Jun, 2018.

[69] The CMS Collaboration, "Performance of missing energy reconstruction in 13 TeV pp collision data using the CMS detector", Technical Report CMS-PAS-JME-16-004, 2016. `https://cds.cern.ch/record/2205284`.

[70] The GEANT4 Collaboration, "GEANT4: A Simulation toolkit", *Nucl. Instrum. Meth.* **A506** (2003) 250–303. `doi:10.1016/S0168-9002(03)01368-8`.

[71] The DELPHES 3 Collaboration, "DELPHES 3, A modular framework for fast simulation of a generic collider experiment", *JHEP* **02** (2014) 057. `doi:10.1007/JHEP02(2014)057`, `arXiv:1307.6346`.

[72] A. Hocker et al., "TMVA - Toolkit for Multivariate Data Analysis", (2007). `arXiv:physics/0703039`.

[73] S. Abdullin et al., "GARCON: Genetic Algorithm for Rectangular Cuts OptimizatioN. User's manual for version 2.0", 2006. `arXiv:hep-ph/0605143`.

[74] T. Junk, "Confidence level computation for combining searches with small statistics", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **434** (Sep, 1999) 435–443. `doi:10.1016/s0168-9002(99)00498-2`.

[75] A. M. Sirunyan et al., "Search for resonant and nonresonant Higgs boson pair production in the bblnulnu final state in proton-proton collisions at sqrt(s) = 13TeV", *Journal of High Energy Physics* **2018** (Jan, 2018). `doi:10.1007/jhep01(2018)054`.

[76] M. Aaboud et al., "Search for Higgs boson pair production in the bbWW decay mode at sqrt(s) = 13 TeV with the ATLAS detector", *Journal of High Energy Physics* **2019** (Apr, 2019). `doi:10.1007/jhep04(2019)092`.

[77] S. Wertz, "Search for Higgs boson pair production in the $\mathrm{b\bar{b}}\ell\nu\ell\nu$ final state with the CMS detector", 2018. `https://cds.cern.ch/record/2632195`.

[78] A. M. Sirunyan et al., "Combination of Searches for Higgs Boson Pair Production in Proton-Proton Collisions at s=13 TeV", *Physical Review Letters* **122** (Mar, 2019). `doi:10.1103/physrevlett.122.121803`.

[79] The CMS Collaboration, "Search for resonant pair production of Higgs bosons decaying to two bottom quark-antiquark pairs in proton-proton collisions at 13 TeV", Technical Report CMS-PAS-HIG-16-002, 2016. `https://cds.cern.ch/record/2141024`.

[80] The CMS Collaboration, "Search for Higgs boson pair production in events with two bottom quarks and two tau leptons in proton-proton collisions at s=13TeV", *Physics Letters B* **778** (2018) 101 – 127. `doi:https://doi.org/10.1016/j.physletb.2018.01.001`.

[81] The CMS Collaboration, "Search for Higgs boson pair production in the ggbb final state in pp collisions at s=13TeV", *Physics Letters B* **788** (2019) 7 – 36. `doi:https://doi.org/10.1016/j.physletb.2018.10.056`.

[82] The ATLAS Collaboration, "Combination of searches for Higgs boson pairs in pp collisions at 13 TeV with the ATLAS experiment.", technical report, 9, 2018.

[83] A. G. et al., "High-Luminosity Large Hadron Collider (HL-LHC): Technical Design Report V. 0.1". CERN Yellow Reports: Monographs. CERN, Geneva, 2017. `https://cds.cern.ch/record/2284929`.

[84] The HL-LHC Collaboration, "HL-LHC project schedule", 2020. `https://project-hl-lhc-industry.web.cern.ch/content/project-schedule`.

[85] H. Damerau et al., "LHC Injectors Upgrade, Technical Design Report, Vol. I: Protons", Technical Report CERN-ACC-2014-0337, Dec, 2014. `http://cds.cern.ch/record/1976692`.

[86] J. Coupard et al., "LHC Injectors Upgrade, Technical Design Report, Vol. II: Ions", Technical Report CERN-ACC-2016-0041, Apr, 2016. `http://cds.cern.ch/record/2153863`.

[87] L. Arnaudon et al., "Linac4 Technical Design Report", Technical Report CERN-AB-2006-084. CARE-Note-2006-022-HIPPI, Dec, 2006. `https://cds.cern.ch/record/1004186`.

[88] D. Contardo et al., "Technical Proposal for the Phase-II Upgrade of the CMS Detector", Technical Report CERN-LHCC-2015-010. LHCC-P-008. CMS-TDR-15-02, Jun, 2015. `https://cds.cern.ch/record/2020886`.

[89] The CMS Collaboration, "The Phase-2 Upgrade of the CMS L1 Trigger Interim Technical Design Report", Technical Report CERN-LHCC-2017-013. CMS-TDR-017, Sep, 2017. `https://cds.cern.ch/record/2283192`.

[90] The CMS Collaboration, "A MIP Timing Detector for the CMS Phase-2 Upgrade", Technical Report CERN-LHCC-2019-003. CMS-TDR-020, Mar, 2019. `https://cds.cern.ch/record/2667167`.

[91] The CMS Collaboration, "The Phase-2 Upgrade of the CMS Endcap Calorimeter", Technical Report CERN-LHCC-2017-023. CMS-TDR-019, Nov, 2017. `https://cds.cern.ch/record/2293646`.

[92] The CMS Collaboration, "The Phase-2 Upgrade of the CMS Barrel Calorimeters", Technical Report CERN-LHCC-2017-011. CMS-TDR-015, Sep, 2017. `https://cds.cern.ch/record/2283187`.

[93] The CMS Collaboration, "The Phase-2 Upgrade of the CMS L1 Trigger Interim Technical Design Report", Technical Report CERN-LHCC-2017-013. CMS-TDR-017, Sep, 2017. `https://cds.cern.ch/record/2283192`.

[94] For the CMS Collaboration, Martin Delcourt, "The CMS Tracker Upgrade for the High Luminosity LHC". `https://indico.cern.ch/event/686555/contributions/2972183/`, Jul, 2018.

[95] J. C. Chistiansen et al., "RD Collaboration Proposal: Development of pixel readout integrated circuits for extreme rate and radiation", Technical Report CERN-LHCC-2013-008. LHCC-P-006, Jun, 2013. `https://cds.cern.ch/record/1553467`.

[96] CERN, "LpGBT specification document". `https://espace.cern.ch/GBT-Project/LpGBT/Specifications/LpGbtxSpecifications.pdf`.

[97] J. Troska et al., "The VTRx+, an optical link module for data transmission at HL-LHC", *PoS* **TWEPP-17** (2017) 048. 5 p. `doi:10.22323/1.313.0048`.

[98] E. Bartz et al., "FPGA-Based Tracklet Approach to Level-1 Track Finding at CMS for the HL-LHC. FPGA-Based Tracklet Approach to Level-1 Track Finding at CMS for the HL-LHC", *EPJ Web Conf.* **150** (Jun, 2017) 00016. 11 p. `doi:10.1051/epjconf/201715000016`.

[99] D. Cieri et al., "An FPGA-based track finder for the L1 trigger of the CMS experiment at the HL-LHC", *PoS* **TWEPP-17** (2017) 131. 5 p. `doi:10.22323/1.313.0131`.

[100] CMS Collaboration Collaboration, "The CMS electromagnetic calorimeter project: Technical Design Report". Technical Design Report CMS. CERN, Geneva, 1997. `https://cds.cern.ch/record/349375`.

[101] A. Z. and, "Performance of the CMS electromagnetic calorimeter during the LHC Run II", *Journal of Physics: Conference Series* **1162** (jan, 2019) 012001. `doi:10.1088/1742-6596/1162/1/012001`.

[102] B. Ketzer et al., "Triple GEM tracking detectors for COMPASS", *Nuclear Science, IEEE Transactions on* **49** (11, 2002) 2403 – 2410. `doi:10.1109/TNS.2002.803891`.

[103] A. Cardini et al., "The Operational Experience of the Triple-GEM Detectors of the LHCb Muon System: Summary of 2 Years of Data Taking", Technical Report LHCb-PROC-2012-060. CERN-LHCb-PROC-2012-060, Nov, 2012. `https://cds.cern.ch/record/1495070`.

[104] W. Krzemien et al., "The KLOE-2 experiment: Overview of recent results", *International Journal of Modern Physics A* **34** (Sep, 2019) 1930012. `doi:10.1142/s0217751x19300126`.

[105] R. Farinelli, "Research and development in cylindrical triple-GEM detector with mTPC readout for the BESIII experiment", 2019. `arXiv:1904.06548`.

[106] D. Braga et al., "CBC2: a microstrip readout ASIC with coincidence logic for trigger primitives at HL-LHC", *Journal of Instrumentation* **7** (oct, 2012) C10003–C10003. `doi:10.1088/1748-0221/7/10/c10003`.

[107] D. S. Luciano Orsini, "CMS Online Software User's Manual", Dec, 2018. `https://edms.cern.ch/document/1001791/3.2`.

[108] Rene Brun and Fons Rademakers, "ROOT - An Object Oriented Data Analysis Framework", *Nucl. Inst. & Meth. in Phys. Res.* **1997** (1996), no. A 389,. `doi:10.1016/S0168-9002(97)00048-X`. See also http://root.cern.ch/.

[109] S. S. E. Nasr-Storey, "Preparing CBC for data taking [Chip calibration]", (Jul, 2017). `https://indico.cern.ch/event/637390/timetable/`. Phase-2 DAQ training workshop.

[110] W. Adam et al., "Test beam demonstration of silicon microstrip modules with transverse momentum discrimination for the future CMS tracking detector", *Journal of Instrumentation* **13** (mar, 2018) P03003–P03003. `doi:10.1088/1748-0221/13/03/p03003`.

[111] Centre de Ressources du Cyclotron Collaboration, "Layout of the CYCLONE120 beam facilities", (2020).

[112] CAEN, "V812 16 Channel Constant Fraction Discriminator". `https://www.caen.it/products/v812/`.

[113] The RD12 Collaboration, P. Gällnö, "TTCvx Technical Description and Users Manual". `http://www.cern.ch/TTC/TTCvxManual1a.pdf`.

[114] The RD12 Collaboration, Ph. Farthoua and P. Gällnö, "TTC-VMEbus INTERFACE TTCvi - MkII". `http://www.cern.ch/TTC/TTCviSpec.pdf`.

[115] The RD12 Collaboration, B. G. Taylor, "Timing distribution at the LHC", Nov, 2002. `https://cds.cern.ch/record/592719`.

[116] P. Vichoudis et al., "The Gigabit Link Interface Board (GLIB), a flexible system for the evaluation and use of GBT-based optical links", *Journal of Instrumentation* **5** (nov, 2010) C11007–C11007. `doi:10.1088/1748-0221/5/11/c11007`.

[117] V. Goffinet, "Study of the silicon detector prototypes for the upgrade of the CMS tracker", (2017).

[118] M. Prydderch, "CBC3 User Manual". `http://www.hep.ph.ic.ac.uk/~dmray/CBC_documentation/CBC3_User_Manual_V1p1.docx`, 2017.

[119] Lorenzo Uplegger, "Overview on the beam test - CBC2/CBC3 data analyses". `https://indico.cern.ch/event/709645/`.

[120] S. Kwan et al., "The pixel tracking telescope at the Fermilab Test Beam Facility", *Nuclear Instruments and Methods in Physics Research. Section A, Accelerators, Spectrometers, Detectors and Associated Equipment* **811** (3, 2016). `doi:10.1016/j.nima.2015.12.003`.

[121] Sarah Seif El Nasr, "Approval CBC3 Performance Results". `https://indico.cern.ch/event/731306/`.

[122] M. Prydderch, "CBC3.1 User Manual". `http://www.hep.ph.ic.ac.uk/ASIC/cbc3.1/CBC3p1_User_Manual_V1p3.pdf`, 2018.

[123] Caponetto, L. and Galbit, G. and Nodari, B. and Scarfi, S. and Viret, S., "CIC1 technical specification". `https://espace.cern.ch/Tracker-Upgrade/Electronics/CIC/Shared%20Documents/Specifications/CIC_specs_v3.pdf`, Dec, 2018.

[124] Caponetto, L. and Galbit, G. and Nodari, B. and Scarfi, S. and Viret, S., "CIC2 technical specification". `https://espace.cern.ch/Tracker-Upgrade/Electronics/CIC/Shared%20Documents/Specifications/CIC2_specs_v1p0.pdf`, Dec, 2018.

[125] C. Dziwok, "Picture of a 2S module prototype", Sep, 2019. `https://indico.cern.ch/event/850565/contributions/3577272/`.

[126] A. Dainese et al., "Report on the Physics at the HL-LHC, and Perspectives for the HE-LHC", Technical Report CERN-2019-007, 2019. `https://cds.cern.ch/record/2703572`.

[127] M. Grazzini et al., "Higgs boson pair production at NNLO with top quark mass effects", *JHEP* **05** (2018) 059. `doi:10.1007/JHEP05(2018)059`, `arXiv:1803.02463`.

[128] A. Carvalho et al., "On the reinterpretation of non-resonant searches for Higgs boson pairs", *arXiv* (2017). `arXiv:1710.08261`.

[129] P. Nason, "A New method for combining NLO QCD with shower Monte Carlo algorithms", *JHEP* **11** (2004) 040. `doi:10.1088/1126-6708/2004/11/040`, `arXiv:hep-ph/0409146`.

[130] S. Frixione et al., "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method", *JHEP* **11** (2007) 070. `doi:10.1088/1126-6708/2007/11/070`, `arXiv:0709.2092`.

[131] S. Alioli et al., "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043. `doi:10.1007/JHEP06(2010)043`, `arXiv:1002.2581`.

[132] S. Frixione et al., "A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction", *JHEP* **09** (2007) 126. `doi:10.1088/1126-6708/2007/09/126`, `arXiv:0707.3088`.

[133] M. Czakon et al., "Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders", *Comput. Phys. Commun.* **185** (2014) 2930. `doi:10.1016/j.cpc.2014.06.021`, `arXiv:1112.5675`.

[134] M. Aliev et al., "HATHOR: HAdronic Top and Heavy quarks crOss section calculatoR", *Comput. Phys. Commun.* **182** (2011) 1034–1046. `doi:10.1016/j.cpc.2010.12.040`, `arXiv:1007.1327`.

[135] P. Kant et al., "HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions", *Comput. Phys. Commun.* **191** (2015) 74–89. `doi:10.1016/j.cpc.2015.02.001`, `arXiv:1406.4403`.

[136] D. de Florian et al., "Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector", CERN Report CERN-2017-002-M, 2016. `arXiv:1610.07922`.

[137] The CMS Collaboration, "Expected performance of the physics objects with the upgraded CMS detector at the HL-LHC", Technical Report CMS-NOTE-2018-006. CERN-CMS-NOTE-2018-006, Dec, 2018. `https://cds.cern.ch/record/2650976`.

[138] The CMS Collaboration, "The Phase-2 Upgrade of the CMS Level-1 Trigger", Technical Report CERN-LHCC-2020-004. CMS-TDR-021, Apr, 2020. `http://cds.cern.ch/record/2714892`.

[139] The CMS Collaboration, "Search for Higgs boson pair production in the $bb\tau\tau$ final state in proton-proton collisions at $\sqrt{(s)} = 8$ TeV", *Phys. Rev. D* **96** (2017), no. 7, 072004. `doi:10.1103/PhysRevD.96.072004`, `arXiv:1707.00350`.

[140] B. Mellado Garcia et al., "CERN Report 4: Part I Standard Model Predictions", technical report, May, 2016. `https://cds.cern.ch/record/2150771`.

[141] J. C. Collins et al., "Angular distribution of dileptons in high-energy hadron collisions", *Phys. Rev. D* **16** (Oct,1977) 2219–2225. `doi:10.1103/PhysRevD.16.2219`.

[142] The CMS Collaboration, "Identification of b-quark jets with the CMS experiment", *JINST* **8** (2013) P04013.

[143] The CMS Collaboration, H. PAG, "Combine tool home page". `https://cms-analysis.github.io/HiggsAnalysis-CombinedLimit/`.

[144] The ATLAS Collaboration, "Measurement prospects of the pair production and self-coupling of the Higgs boson with the ATLAS experiment at the HL-LHC", Technical Report ATL-PHYS-PUB-2018-053, Dec, 2018. `https://cds.cern.ch/record/2652727`.

[145] J. H. Kim et al., "Probing the Triple Higgs Self-Interaction at the Large Hadron Collider", *Physical Review Letters* **122** (Mar, 2019). `doi:10.1103/physrevlett.122.091801`.

[146] Preprint edited by M. Mangano et al, "Future Circular Collider Study. Volume 1: Physics Opportunities. Conceptual Design Report.", *Eur. Phys. J. C.* (Dec, 2018). CERN accelerator reports, CERN-ACC-2018-0056, Geneva.

[147] Preprint edited by M. Benedikt et al, "Future Circular Collider Study. Volume 2: The Lepton Collider (FCC-ee). Conceptual Design Report.", *Eur. Phys. J. ST.* (Dec, 2018). CERN accelerator reports, CERN-ACC-2018-0057, Geneva.

[148] Preprint edited by M. Benedikt et al, "Future Circular Collider Study. Volume 3: The Hadron Collider (FCC-hh). Conceptual Design Report.", *Eur. Phys. J. ST.* (Dec, 2018). CERN accelerator reports, CERN-ACC-2018-0058, Geneva.

[149] P. Vichoudis et al., "FC7 Project Home Page". `https://espace.cern.ch/project-FC7/SitePages/Home.aspx`.

[150] "CMS Tracker Phase2 Acquisition & Control Framework". `https://gitlab.cern.ch/cms_tk_ph2/Ph2_ACF`.

[151] G. Auzinger, "Acquisition & Control Framework Supervisor". `https://github.com/gauzinge/ACFSupervisor`.

[152] C. Delaere, "Phase-2 tracker data format for 2S modules". `https://cms-docdb.cern.ch/cgi-bin/DocDB/ShowDocument?docid=12091`.

[153] "DAQ/FE interfacing guide". `http://cmsdoc.cern.ch/cms/TRIDAS/horizontal/RUWG/DAQ_IF_guide/DAQ_IF_guide.html`.

[154] Hammond manufacturing, "1590DD Enclosure". `http://www.farnell.com/datasheets/1806244.pdf`.

[155] E. TECHNOLOGY, “GENERAL PURPOSE PLASTIC SCINTILLATOR EJ-200, EJ-204, EJ-208, EJ-212”. `https://eljentechnology.com/images/products/data_sheets/EJ-200_EJ-204_EJ-208_EJ-212.pdf`.

[156] Hamamatsu, “R9800 Photomultiplier tube”. `https://www.hamamatsu.com/jp/en/product/type/R9800/index.html`.

[157] CAEN, “A1424 Scintillation Preamplifier”. `https://www.caen.it/products/a1424/`.