"Search for anomalous Higgs boson production in association with single top quarks using the CMS detector"

Popov, Andrey

ABSTRACT

This thesis describes a search for production of a Higgs boson in association with a single top quark, conducted in the CMS experiment at the CERN Large Hadron Collider (LHC). This process offers a unique way to study certain aspects of the interaction between the Higgs boson and the top quark. The search targets the anomalous case of a negative top quark Yukawa coupling constant, which leads to a significant enhancement of the tH production cross section with respect to the expectation in the standard model. The study exploits data collected in proton-proton collisions at a centre-of-mass energy of 8 TeV, corresponding an accumulated luminosity of about 20/fb. In the considered final state the top quark decays producing a muon or an electron, and the Higgs boson decays into a pair of b quarks. The event selection takes advantage of multiple b-quark jets as well as a jet in the forward region in pseudorapidity. The selected sample of events is largely dominated by the top quark pair production, and artificial neural networks are applied in order to identify signal-like events. A special attention is given to construction of input variables in the multijet final state. Observed data are found to be compatible with the background-only expectation, and an upper limit on the tH production cross section is set. The observed limit at the 95% confidence level is 1.8 pb, which exceeds the expected cross section of the anomalous tH production by a factor of 7.6. Data that will be collected by the LHC in the years to come will allow to rule out (or discover) the considered anomalou...

CITE THIS VERSION

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

Search for anomalous Higgs boson production in association with single top quarks using the CMS detector

Doctoral dissertation presented by

Andrey Popov

in fulfilment of the requirements for the degree of Doctor in Sciences

Prof. A. Giammanco (adviser)	UCL, Belgium
Prof. V. Lemaître (chairman)	UCL, Belgium
Dr. R. Chierici	IPNL, France
Prof. B. Clerbaux	ULB, Belgium
Prof. F. Maltoni	UCL, Belgium

October 2015

Contents

Introduction

The standard model (SM) of electroweak and strong interactions was formulated by mid-1970s. It has proven to be an extraordinarily successful theory, which has been consistently explaining every result obtained in experimental studies of subatomic particles ever since. The only crucial element of the theory that used to lack experimental confirmation until recently was the Higgs boson. In 2012 it was finally discovered by the ATLAS and CMS collaborations [1, 2], ending an almost fifty-year-long quest. After the discovery, it is important to investigate properties of the new particle in order to determine whether it is indeed the Higgs boson of the standard model or if there are deviations from the expectations, providing hints for physics beyond the standard model (BSM).

A property of the Higgs boson that is of a particular interest is the strength of its interaction with the top quark. In the standard model, elementary particles obtain the mass through the interaction with the Higgs boson. The exceptionally large mass of the top quark then suggests that there might a special relation between the two particles. In fact, the SM Yukawa coupling constant of the top quark y_t is enigmatically very close to unity, while coupling constants of the remaining fermions are of the order of 10^{-2} or less. The value of y_t can be measured in the $t\bar{t}H$ production, whose cross section is proportional to y_t^2. The study in $t\bar{t}H$ can be complemented by investigating the tH production, where the Higgs boson is created in association with a single top quark instead of a pair. This process possesses a unique sensitivity to certain aspects of the Yukawa interaction. In particular, it allows to determine the sign of y_t, which is defined with respect to the coupling between the Higgs and weak bosons.

This thesis is devoted to a search for the tH production with $H \to b\bar{b}$ and semileptonic decays of the top quark. In the standard model, this process has a very small cross section and cannot be observed with available data.

However, its production rate increases by more than an order of magnitude if $y_t = -1$. The search targets the case with the enhanced cross section, but a large $t\bar{t}$ background makes this a challenging task nonetheless. In order to improve the sensitivity, methods of multivariate analysis (MVA) are exploited.

Data used in this search have been recorded with the CMS detector in 2012 in pp collisions at a centre-of-mass energy $\sqrt{s} = 8\,\text{TeV}$. It corresponds to an integrated luminosity of about $20\,\text{fb}^{-1}$. In addition, $5\,\text{fb}^{-1}$ of 7 TeV data recorded in 2011 are available, but these data have not been utilized in the search because the resulting moderate increase in the integrated luminosity would not boost the sensitivity significantly. No attempt has been made to analyse first 13 TeV data since the integrated luminosity recorded at the time of writing, about $40\,\text{pb}^{-1}$, is insufficient for this search.

This thesis is organized as follows. Chapter 1 briefly introduces the standard model, focusing on the top quark and the Higgs boson. The tH production is discussed, and experimental results relevant to this search are summarized. In Chapter 2 the Large Hadron Collider (LHC) and the CMS detector are described. Chapter 3 explains how individual particles are reconstructed from detector signals. It also addresses reconstruction of derived objects and quantities such as jets and missing E_T and lists identification criteria applied in the search. The reconstructed objects are exploited to define the event selection, which is documented in Chapter 4. Simulated data used in the search are also discussed there. Chapter 5 describes the multivariate analysis, providing also a brief introduction to artificial neural networks. Finally, the statistical inference is addressed in Chapter 6. It elaborates on the method applied to set an upper limit on the tHq cross section, discusses considered uncertainties, and reports the results. In addition, constraints obtained in complementary searches in other Higgs boson decay channels and the overall combination are presented.

The search documented in this thesis was published as Ref. [3].

Chapter 1

Higgs boson and top quark

In this thesis the interaction between the Higgs boson and the top quark is studied. Its theoretical description is provided by the standard model, although in the following a special attention is given to possible BSM effects. Some of them can be probed in the associated tH production.

1.1 Standard model

The standard model of electroweak [4–6] and strong [7–10] interactions describes all phenomena of the microscopic world. It does so using only a small set of fundamental particles. The interactions between them are defined by internal symmetries, as prescribed by the framework of gauge theories. The symmetry of the standard model is spontaneously broken by the ground state, in the event of which fundamental particles acquire masses.

1.1.1 Particle content

The majority of common matter, despite of the great diversity of its forms, is constructed with just a few elementary particles. Atoms consist of negatively charged electrons orbiting massive positively charged nuclei, which are composed of protons and neutrons, collectively referred to as nucleons. Nucleons are not elementary but consist of three valence quarks of two flavours: up and down, or u and d. The composition of a proton and a neutron is uud and

udd respectively. The u and d quarks have non-integral electric charges $+2/3$ and $-1/3$, which gives integral charges to nucleons. Similar to electrons, quarks have spin $1/2$. In β decays another spin-$1/2$ elementary particle is produced, the electron antineutrino $\bar{\nu}_e$. The four particles u, d, e^-, and ν_e, together with the corresponding antiparticles, form the first generation of fermions, and they are the only elementary fermions that contribute to common matter.

For some mysterious reason, two more fermion generations exist in nature. The additional particles differ in mass but in other respects participate in interactions in the same way as analogous fermions of the first generation. The second generation consists of charm and strange quarks, muon, and muon neutrino, and the third generation includes top and bottom quarks, τ lepton, and τ neutrino. There is a pronounced mass hierarchy in which particles of each successive generation (with a potential exception for neutrinos) are significantly more massive than their counterparts in the previous generation. The top quark, discovered in 1995 by the CDF and DØ collaborations [11, 12], is the heaviest elementary particle and has a mass of about $173\,\mathrm{GeV}/c^2$, which is similar to the mass of a tungsten atom. The large masses cause particles of higher generations to decay to the first generation, which is why they do not contribute to common matter. The presence of the forth generation populated with even heavier fermions is strongly disfavoured after the discovery of the Higgs boson [13].

Quarks can form a great variety of composite particles called hadrons. On the other hand, free isolated quarks do not exist due to a phenomenon known as colour confinement: an attempt to extract a single quark from a hadron results in spontaneous creation of quark–antiquark pairs, which then combine with the remnant of the hadron and the quark to form new hadrons. Main types of hadron structure that are predicted by the quark model and have been observed experimentally, are pairs of quarks and antiquarks (mesons) and triplets of quarks or antiquarks (baryons). Recently, a pentaquark, a state consisting of four quarks and an antiquark, has been observed by the LHCb collaboration [14]. Hadrons can contain valence quarks of any flavour except for the top quark. Due to its large mass, the top quark has a very short lifetime and decays via electroweak interaction before a bound state can be formed.

Interactions between the fermions are mediated by several types of vector bosons: photons, $W^\pm$ and Z^0 bosons, and gluons, which are responsible for the electromagnetic, weak, and strong interactions, respectively. Although all fermions (except for, maybe, one flavour of neutrino) are massive and thus participate also in gravitational interaction, it has no impact at experimentally accessible energies. Photons are massless, allowing the electromagnetic forces

to act at large distances. On the other hand, the W and Z bosons are heavy, with masses of about 80.4 and 91.2 GeV/c^2, and this limits the range of the weak interaction. Gluons are also massless, but they carry colour charge and thus are subject to the confinement, which constrains the strong force to the range of about the size of a nucleon.

Elementary fermions and weak bosons obtain masses dynamically via interaction with the Higgs field, as will be described below. It should be noted, however, that this mechanism accounts for only a few per cent of mass of common matter, while the dominant contribution to the mass of nucleons is provided by the strong interaction. The discovery of the excitation of the Higgs field, the Higgs boson with a mass of about 125 GeV/c^2, has completed the list of particles of the standard model.

1.1.2 Standard model as a gauge theory

The standard model is a gauge quantum field theory based on the local symmetry group $G_{\mathrm{SM}} = SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$. All its fields are arranged into representations of these subgroups, and the Lagrangian density $\mathcal{L}$ is constructed to be invariant under G_{SM}, which consequently makes the action $S = \int \mathcal{L}\, \mathrm{d}^4x$ invariant. The $SU(3)_C$ subgroup describes the strong interaction, while the $SU(2)_L \otimes U(1)_Y$ symmetry describes the electroweak interaction. Conserved charges related to these symmetries by the Noether theorem are colour, weak isospin, and weak hypercharge. The $SU(2)_L \otimes U(1)_Y$ symmetry is, however, spontaneously broken down to an $U(1)_Q$ symmetry, which corresponds to electromagnetic interactions and conserved electric charge.

The $SU(3)_C$ part of the theory describes quarks with fundamental representations of the group, thus introducing three components for the quark field of each flavour, which correspond to the three colours. Since the symmetry group has eight generators, there are eight corresponding gauge bosons, the gluon fields $G^a_\mu(x)$, $a = 1, \ldots, 8$, which are described in the adjoint representation. The Lagrangian of the $SU(3)_C$ sector is

$$\mathcal{L}_{\mathrm{QCD}} = \bar{q}_i \left(i \not{D}_{ij} - \delta_{ij} m_q \right) q_j - \frac{1}{4} G^a_{\mu\nu} G^{\mu\nu}_a, \tag{1.1}$$

where i and j are indices of the fundamental representations of the $SU(3)_C$ group, summation over all repeated indices is implied, and the Feynman slash notation is used, $\not{D} \equiv \gamma_\mu D^\mu$. In the first term in Eq. (1.1), a summation over quark flavours is also implied. The covariant derivative is defined as

$$D_\mu = \partial_\mu - i g_s \frac{\lambda^a}{2} G^a_\mu, \tag{1.2}$$

where g_s is the coupling constant of the strong interaction, and λ^a are the Gell–Mann matrices. The $SU(3)$ group is non-abelian, and thus the gluon field strength tensor $G^a_{\mu\nu}$ contains a term quadratic in field:

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_s f^{abc} G^b_\mu G^c_\nu, \tag{1.3}$$

where f^{abc} are the structure constants of the group.

Transformations corresponding to the $SU(2)_L \otimes U(1)_Y$ symmetry act differently on left- and right-handed chiral states

$$\psi_L = \frac{1-\gamma_5}{2}\psi, \quad \psi_R = \frac{1+\gamma_5}{2}\psi. \tag{1.4}$$

The left-handed fermions are organized into $SU(2)_L$ doublets

$$Q^m_L = \begin{pmatrix} u^m_L \\ d^m_L \end{pmatrix}, \quad L^m_L = \begin{pmatrix} \nu^m_L \\ e^m_L \end{pmatrix}, \tag{1.5}$$

which are the fundamental representation of the subgroup. Here index $m = 1, \ldots, 3$ denotes the generation. Differently, the right-handed fermions ν^m_R, e^m_R, u^m_R, d^m_R are not affected by the $SU(2)_L$ transformations, and thus they fall into the trivial representation of this subgroup. The $U(1)_Y$ group acts on all fermions except for the right-handed neutrinos, whose $U(1)_Y$ charge is zero. Electroweak and electric charges of all fermions are shown in Table 1.1. Three gauge bosons W^i_μ, $i = 1, \ldots, 3$, correspond to the $SU(2)_L$ symmetry. They are described in the adjoint representation. The $U(1)_Y$ group adds a forth field B_μ.

Table 1.1: Fermion charges under $SU(2)_L$, $U(1)_Y$, and $U(1)_Q$ groups: electric charge Q, weak isospin T_3, and weak hypercharge Y_W, respectively. The three charges are related as $Q = T_3 + Y_\mathrm{W}/2$. Fermions are split into left- and right-handed chiral states. Each line refers to all three generations.

Fermions	Q	T_3	Y_W
ν_L	0	$+1/2$	-1
e_L	-1	$-1/2$	-1
u_L	$+2/3$	$+1/2$	$+1/3$
d_L	$-1/3$	$-1/2$	$+1/3$
ν_R	0	0	0
e_R	-1	0	-2
u_R	$+2/3$	0	$+4/3$
d_R	$-1/3$	0	$-2/3$

The Lagrangian of the electroweak sector reads as

$$\mathcal{L}_{\mathrm{EW}} = i\bar{Q}_L \not{D} Q_L + i\bar{u}_R \not{D} u_R + i\bar{d}_R \not{D} d_R + i\bar{L}_L \not{D} L_L + i\bar{e}_R \not{D} e_R - \frac{1}{4} W^i_{\mu\nu} W_i^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} + \mathcal{L}_{\mathrm{Higgs}} + \mathcal{L}_{\mathrm{Yukawa}}, \quad (1.6)$$

where a summation over the generations is implied. The covariant derivative is defined as

$$D_\mu = \partial_\mu - ig \frac{\sigma^i}{2} W^i_\mu - ig' \frac{Y}{2} B_\mu, \quad (1.7)$$

where g and g' are the coupling constants for groups $SU(2)_L$ and $U(1)_Y$, σ^i are the Pauli matrices, and Y is the weak hypercharge of the field on which the derivative acts. In case of $SU(2)_L$ singlets, the second term in Eq. (1.7) is omitted. The field strength tensors are given by

$$W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu + g\epsilon^{ijk} W^j_\mu W^k_\nu, \quad (1.8)$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu, \quad (1.9)$$

where ϵ^{abc} is the Levi–Civita symbol, which is the structure constant of the $SU(2)$ group.

The Higgs field is described by the Lagrangian

$$\mathcal{L}_{\mathrm{Higgs}} = (D_\mu \Phi)^\dagger D^\mu \Phi - V(\Phi^\dagger \Phi), \quad (1.10)$$

where $\Phi = \left(\phi^+, \phi^0\right)^T$ is an $SU(2)_L$ doublet of complex scalar fields with a weak hypercharge $Y_H = 1$. The covariant derivative follows the definition (1.7). The potential of the Higgs field is chosen in the form

$$V(\Phi^\dagger \Phi) = \mu^2 \Phi^\dagger \Phi + \lambda \left(\Phi^\dagger \Phi\right)^2, \quad (1.11)$$

and μ^2 and $\lambda > 0$ are its parameters.

The last term in the Lagrangian (1.6) describes the Yukawa interaction between the fermions and the Higgs field. It reads as

$$\mathcal{L}_{\mathrm{Yukawa}} = -Y^u_{mn} \bar{Q}_{Lm} \tilde{\Phi} u_{Rn} - Y^d_{mn} \bar{Q}_{Lm} \Phi d_{Rn} - Y^e_{mn} \bar{L}_{Lm} \Phi e_{Rn} + \mathrm{h.\,c.} \quad (1.12)$$

with $\tilde{\Phi} = i\sigma_2 \Phi^*$. Here $Y^{u,d,e}_{mn}$ are Yukawa coupling constants for up- and down-type quarks and charged (down-type) leptons, and indices m and n enumerate generations. As will be shown later, these interactions are responsible for generation of fermion masses. Since at least two flavours of neutrino are massive, the corresponding term $-Y^\nu_{mn} L_{Lm} \tilde{\Phi} \nu_{Rn}$ can be added to the Lagrangian. However, neutrino masses are very small, and they are neglected in the following discussion by setting $Y^\nu_{mn} = 0$.

1.1.3 Electroweak symmetry breaking

Although from experimental results elementary fermions are known to be massive, the electroweak Lagrangian (1.6) does not contain the corresponding mass terms. In fact, a term $-m_f(\bar{f}_L f_R + \bar{f}_R f_L)$, which would assign a mass m_f to the fermion f, is not invariant under the $SU(2)_L$ transformations and thus cannot be added explicitly. The four gauge electroweak bosons are also seemingly massless, while the weak interaction must be mediated by massive particles in order to have a short range. There apparent contradictions are resolved by the Brout–Englert–Higgs mechanism [15–20]. It describes the spontaneous breaking of the $SU(2)_L \otimes U(1)_Y$ symmetry down to $U(1)_Q$ and generates masses for fermions and weak bosons dynamically, as discussed below.

The ground state of the theory depends on the shape of the Higgs potential (1.11). If $\mu^2 > 0$, the vacuum expectation $\langle 0|\Phi|0\rangle$ is zero. On the other hand, if $\mu^2 < 0$, $|\Phi| = 0$ is the point of an unstable local maximum of the potential, while the minimum is reached with

$$\Phi^\dagger \Phi = \frac{v^2}{2}, \quad v = \sqrt{-\frac{\mu^2}{\lambda}}. \tag{1.13}$$

Without loss of generality, the ground state can be chosen so that

$$\langle 0|\Phi|0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \tag{1.14}$$

Here, the vacuum expectation of the electrically charged upper component of the Higgs doublet is zero. Although the Lagrangian is invariant under $SU(2)_L \otimes U(1)_Y$ transformations, the chosen ground state does not share this symmetry. In fact, it breaks all generators of the group except for the combination $1/2 \cdot \sigma^3 + Y_H/2 \cdot \mathbb{1}_2$, which corresponds to the electric charge. The $U(1)_Q$ symmetry is therefore respected by the ground state (1.14).

With the unitary gauge, perturbations of the scalar field around the chosen vacuum can be parameterized as

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix}, \tag{1.15}$$

where H is a Hermitian field. If this is substituted into the Lagrangian (1.10), terms that do not include gauge bosons read as

$$\mathcal{L}_{\text{Higgs}} \supset \frac{1}{2}\partial_\mu H \partial^\mu H - \lambda v^2 H^2 - \lambda v H^3 - \frac{\lambda}{4} H^4 + \frac{\lambda}{4} v^4, \tag{1.16}$$

which reveals the mass of the Higgs boson $m_H = \sqrt{2\lambda v^2}$. The terms quadratic in v provide masses to weak bosons:

$$\mathcal{L}_{\text{Higgs}} \supset \frac{v^2}{8}\left(g^2|W_\mu^1 - iW_\mu^2|^2 + (gW_\mu^3 - g'B_\mu)^2\right) = \frac{g^2v^2}{8}W_\mu^+W^{+\mu} + \frac{g^2v^2}{8}W_\mu^-W^{-\mu} + \frac{(g^2+g'^2)v^2}{8}Z_\mu Z^\mu + 0\cdot A_\mu A^\mu, \quad (1.17)$$

where the mass eigenstates are defined as

$$W_\mu^\pm \equiv \frac{W_\mu^1 \mp iW_\mu^2}{\sqrt{2}} \quad (1.18)$$

$$A_\mu \equiv \frac{g'W_\mu^3 + gB_\mu}{\sqrt{g^2+g'^2}}, \quad Z_\mu \equiv \frac{gW_\mu^3 - g'B_\mu}{\sqrt{g^2+g'^2}}. \quad (1.19)$$

These eigenstates correspond to the physical W and Z bosons with masses $m_W = gv/2$ and $m_Z = v\sqrt{g^2+g'^2}/2$, as well as the massless photon A_μ.

The Yukawa block with the parameterization (1.15) translates into

$$\mathcal{L}_{\text{Yukawa}} = -Y_{mn}^u \bar{u}_{Lm}u_{Rn}\frac{v+H}{\sqrt{2}} - Y_{mn}^d \bar{d}_{Lm}d_{Rn}\frac{v+H}{\sqrt{2}} - Y_{mn}^e \bar{e}_{Lm}e_{Rn}\frac{v+H}{\sqrt{2}} + \text{h. c.} \quad (1.20)$$

Terms proportional to v give masses to the fermions, while the remaining ones are responsible for interactions with the Higgs boson. In general, the mass matrices $v/\sqrt{2}\cdot Y$ are not diagonal, which means that the fermion fields considered so far are not mass eigenstates. The mass matrices can be diagonalized with the help of unitary matrices $V_{L,R}^u$:

$$-\frac{v}{\sqrt{2}}\bar{u}_L Y^u u_R + \text{h. c.} = -\frac{v}{\sqrt{2}}\bar{u}_L V_L^{u\dagger}\left(V_L^u Y^u V_R^{u\dagger}\right)V_R^u u_R + \text{h. c.} = -\frac{v}{\sqrt{2}}\bar{u}_L V_L^{u\dagger} Y_D^u V_R^u u_R + \text{h. c.}, \quad (1.21)$$

where $Y_D^u = \text{diag}(m_u, m_c, m_t)$, and a similar operation can be performed for down-type quarks and charged leptons. The transformation matrices $V_{L,R}^{u,d,e}$ can be absorbed into definitions of fermion fields thus translating them into the mass eigenstates. Since neutrino mass terms are ignored, there is no need to reproduce their mass eigenstates, and then it is convenient to transform these fields in the same way as charged leptons. After the change of the basis, the transformation matrices emerge in the charged current part of the Lagrangian (1.6):

$$\mathcal{L}_{\text{c. c.}} = -\frac{g}{\sqrt{2}}\left(\bar{u}_m\gamma^\mu\frac{1-\gamma_5}{2}V_{mn}^{\text{CKM}}d_n + \bar{\nu}_m\gamma^\mu\frac{1-\gamma_5}{2}e_m\right)W_\mu^+ + \text{h. c.}, \quad (1.22)$$

although only in the quark sector thanks to the chosen rotation for neutrinos. Here $V^{\mathrm{CKM}} = V_L^u V_L^{d\dagger}$ is the Cabibbo–Kobayashi–Maskawa mixing matrix.

As can be seen from Eq. (1.21), the resulting masses of fermions read as

$$m_f = \frac{y_f v}{\sqrt{2}}, \tag{1.23}$$

where y_f is the appropriate component of the diagonalized Yukawa matrix Y_D. Values of the Yukawa coupling constants are among the free parameters of the standard model. Mysteriously, only top quark coupling constant is of the natural scale, $y_t \sim 1$, whereas constants for other fermions are of the order of 10^{-2} or smaller. Even more strangely, the top quark coupling constant is very close to unity: $y_t = 0.995 \pm 0.005$ [21]. The standard model offers no explanation for these experimental observations.

1.2 Associated tH production

The large mass of the top quark and, more specifically, the fact that its Yukawa coupling constant is very close to unity, suggests that the quark might play a special role in the electroweak symmetry breaking. Because of this, it is very important to study the interaction between the Higgs boson and the top quark in great detail. This thesis is focused on the production of Higgs boson in association with a single top quark. Although the $t\bar{t}H$ process offers the most straightforward way to study the interaction between the two particles, the tH production provides a unique possibility to investigate certain aspects of this interaction at tree level, which otherwise are only accessible through loops.

1.2.1 Theory

At the leading order, there are three modes of the tH production [22], which result in different final states: tHq, where q denotes a quark of the first two generations, tHW, and tHb. Equivalently, they can be classified into the t-channel, the s-channel, and tW-associated production, as it is usually done for production of single top quarks and shown in Fig. 1.1. Here the most attention is given to the t-channel, or tHq, production, which has the largest cross section of the three [22, 23]. The dominant Feynman diagrams contributing to the process are shown in Fig. 1.2. Although the Higgs boson can also be attached to the b-quark line, the corresponding amplitude is suppressed by the small value of the Yukawa coupling of b quark.

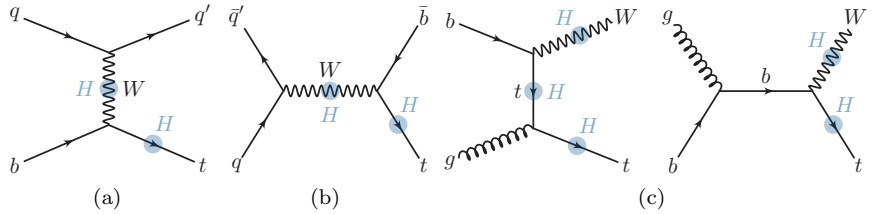


Figure 1.1: Representative Feynman diagrams for production of single top quarks in the t channel (a), the s channel (b), and in association with a W boson (c). To create corresponding diagrams for the tH production, a Higgs boson can be attached to lines marked with blue circles.

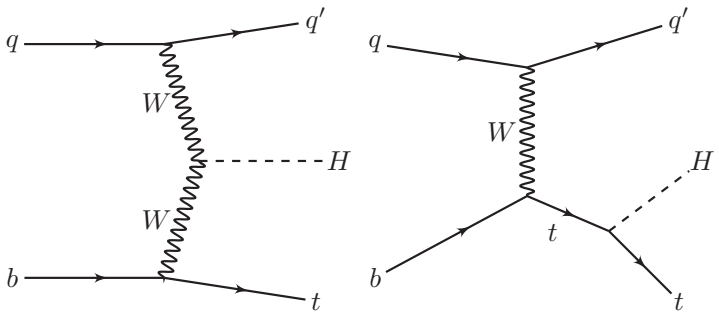


Figure 1.2: Dominant Feynman diagrams for the tH production in the t channel.

In the standard model there is a large destructive interference between the two diagrams shown in Fig. 1.2 [24, 25]. The resulting partial cancellation of the two amplitudes makes the tHq production an interesting process to probe for a deviation of couplings of the Higgs boson from their SM expectations since this can potentially disturb the cancellation and cause a significant increase of the cross section. In a simple yet generic case, the deviations of the couplings to the top quark and the W boson can be parameterized with real-valued factors κ_t, κ_W that scale the coupling strength with respect to the SM expectation. Without loss of generality, $\kappa_W \geqslant 0$ is assumed. The SM case is reproduced with $\kappa_t = \kappa_W = 1$, while setting one of the factors to zero excludes the contribution from the corresponding diagram. In Ref. [26], the total amplitude for the central hard scattering $Wb \to tH$ is derived in the high-energy regime $s, -t, -u \gg m_t^2, m_H^2, m_W^2$, where $s = (p_W + p_b)^2$, $t = (p_W - p_H)^2$, and $u = (p_W - p_t)^2$ are

the Mandelstam variables of the central process. The amplitude is found to be

$$\mathcal{A} = \frac{g}{\sqrt{2}} \left((\kappa_t - \kappa_W) \frac{m_t \sqrt{s}}{m_W v} \cdot A(t/s, \phi) + \left(\kappa_W \frac{2 m_W}{v} \frac{s}{t} + (2\kappa_t - \kappa_W) \frac{m_t^2}{m_W v} \right) \cdot B(t/s, \phi) \right), \quad (1.24)$$

where ϕ is the azimuthal angle with respect to the direction of the W boson and explicit expressions for matrices A and B are given in Ref. [26]. If $\kappa_t = \kappa_W$, the first summand vanishes, and the amplitude is constant in the $s \to +\infty$ limit (where the ratio s/t, which is related to the angle of scattering, is kept finite). Therefore, the cross section decreases as $1/s$. On the other hand, if $\kappa_t \neq \kappa_W$, the amplitude grows as $\sqrt{s}$ leading to a constant cross section for large values of s, thereby providing an enhancement. Although the first summand must be eliminated to respect the unitarity constraints [24], with $\kappa_t \neq \kappa_W$ the perturbative unitarity is not lost until some large enough scale is reached. In Ref. [26] it is estimated as

$$\Lambda = 12\pi\sqrt{2} \frac{v^2}{m_t |\kappa_t - \kappa_W|}, \quad (1.25)$$

which gives $\Lambda \sim 10\,\text{TeV}$ for $\kappa_W = -\kappa_t = 1$. This is a high scale even for 13 TeV pp collisions at the LHC. Indeed, if the corresponding matrix element is convoluted with the parton distribution functions (PDF), the fraction of tHq events where the invariant mass of the tH system exceeds 1 TeV is negligible [26]. This observation allows to consider the case of $\kappa_t \neq \kappa_W$ as an effective theory under the LHC conditions.

The $\kappa_W = -\kappa_t = 1$ case mentioned above is of a special interest. Although absolute values of the coupling strengths can be measured in HW and $t\bar{t}H$ production, these processes have no sensitivity to the relative sign of the couplings. On the other hand, flipping of the sign causes a spectacular enhancement of the tHq cross section. For pp collisions at $\sqrt{s} = 8\,\text{TeV}$, it increases from about 18 fb in the SM case to 235 fb [26], i. e. by a factor of 13, significantly exceeding the $t\bar{t}H$ cross section of about 130 fb [27]. Admitting that the SM rate is way too low to be observed with the data recorded so far by the LHC, the large enhancement for the flipped sign potentially allows to observe or exclude the anomalous production, as has been suggested in a number of phenomenological papers [23, 25, 26, 28–30]. An experimental study of this possibility is the subject of this thesis.

It should be noted that the tH production is not the only process that is sensitive to the relative sign of the coupling constants. Through a similar

interference effect, it affects $gg \to ZH$ [31], $gg \to HH$ [26], and the branching fraction of the $H \to \gamma\gamma$ decay [23]. In the SM, the cross section of $gg \to ZH$ is about 20 fb, and it is enhanced by a factor of 5 [31]. However, this process is more difficult to study than the tHq production, in part, because of the irreducible SM ZH background with a cross section of about 400 fb [31]. The $gg \to HH$ production with an SM cross section of about 10 fb [32] is even more challenging. The decay $H \to \gamma\gamma$ predominantly occurs through a loop with a top quark or a W boson. Its branching fraction is modified with a scale factor κ_γ^2, which can be approximately parameterized in the following way [27, Eq. (113)]:

$$\kappa_\gamma^2 \approx 0.07 \cdot \kappa_t^2 + 1.61 \cdot \kappa_W^2 - 0.68 \cdot \kappa_t \kappa_W . \tag{1.26}$$

As can be seen, the flipped sign increases the branching fraction by a factor of about 2.4, which can be observed by measuring the ratio $\mathcal{B}(H \to \gamma\gamma)/\mathcal{B}(H \to VV^*)$, where $V = W, Z$, in inclusive Higgs boson production.

Although the last approach is the easiest of the three from the experimental point of view, in Eq. (1.26) a strong assumption is made that there is no contribution from new particles in the $H \to \gamma\gamma$ loop. If this assumption is lifted, the dependence of $\mathcal{B}(H \to \gamma\gamma)$ on κ_t and κ_W is largely unknown. Likewise, the $gg \to ZH$ and HH production occurs through loops, and therefore an interpretation of results obtained with these processes would suffer from the same deep problem. In fact, the tH production is the only process available experimentally that allows to probe the κ_t sign at tree level. Another process with the tree-level sensitivity is the HWW production [33], but because of its small cross section it can only be studied in future, when a very large amount of data is available.

So far, only $\mathcal{CP}$-conserving couplings have been considered. However, a pseudoscalar component can be introduced to the interaction between the top quark and the Higgs boson:

$$\mathcal{L}_t = -\frac{m_t}{v} \left(\kappa_t \, \bar{t}t + i\tilde{\kappa}_t \, \bar{t}\gamma_5 t\right) H. \tag{1.27}$$

This can be viewed as a generalization of the $\kappa_V = 1$, $\kappa_t = \pm 1$ cases discussed above, and the SM is reproduced by setting $\kappa_t = 1$, $\tilde{\kappa}_t = 0$. The interference between the two diagrams shown in Fig. 1.2 is also sensitive to the pseudoscalar coupling scale factor $\tilde{\kappa}_t$ and, in particular, to the $\mathcal{CP}$-violating phase $\xi_t = \arctan(\tilde{\kappa}_t/\kappa_t)$ [25,34]. Studies of the $\mathcal{CP}$-violating top quark Yukawa coupling in the tHq process are discussed in Refs. [35,36]. They can also be complemented with studies of angular distributions in $t\bar{t}H$ events [35, 37, 38].

1.2.2 Experimental results

Because of the Higgs boson involved, experimental constraints relevant to the tH production can be set virtually only by the ATLAS [39] and CMS [40] experiments. The CMS collaboration has performed searches for the tHq production with $\kappa_t = -1$ in multiple decay channels of the Higgs boson [3, 41–43]. The search in the $H \to b\bar{b}$ channel is the topic of this thesis, and results obtained in the other channels will also be discussed. On the other hand, the ATLAS collaboration has not attempted direct searches for this process but instead included it as a background in a $t\bar{t}H$, $H \to \gamma\gamma$ search [44].

Both collaborations have performed global combinations of their measurements of properties of the Higgs boson [45, 46], although in the case of CMS it does not include the direct searches for tHq. The simplest model considered for the combinations contains only two free parameters, κ_f and κ_V, which simultaneously rescale coupling constants of the Higgs boson to all fermions or all vector bosons, respectively. Both ATLAS and CMS strongly favour $\kappa_f \sim 1$ and $\kappa_V \sim 1$, and although there is a second local maximum of the likelihood at $\kappa_f \sim -1$ and $\kappa_V \sim 1$, this point is excluded at more than 95% confidence level (CL). It should be noted, however, that the sensitivity to the sign of κ_f comes exclusively from the $H \to \gamma\gamma$ decay channel, which means that the $\kappa_f = -1$ case is disfavoured because because the naive enhancement of the branching fraction $\mathcal{B}(H \to \gamma\gamma)$ by a factor of 2.4 is not compatible with data. This observation is confirmed by the measurements of the double ratio

$$\rho_{\gamma\gamma/WW^*} = \frac{\mathcal{B}(H \to \gamma\gamma)/\mathcal{B}(H \to WW^*)}{[\mathcal{B}(H \to \gamma\gamma)/\mathcal{B}(H \to WW^*)]_{\mathrm{SM}}}, \tag{1.28}$$

which incorporate little to no model dependence. Their values are $\rho_{\gamma\gamma/WW^*} = 0.97^{+0.32}_{-0.25}$ and $1.21^{+0.41}_{-0.31}$ for ATLAS and CMS respectively, positively excluding $\mathcal{B}(H \to \gamma\gamma)/\left[\mathcal{B}(H \to \gamma\gamma)\right]_{\mathrm{SM}} \approx 2.4$ if the branching fraction of the $H \to WW^*$ decay is not modified.

The apparent exclusion of the $\kappa_t = -1$ case based on the value of $\mathcal{B}(H \to \gamma\gamma)$ only holds if the parameterization (1.26) is correct, which relies on the assumption that no new particles potentially running in the $H \to \gamma\gamma$ loop modify the branching fraction. The results change when the $H\gamma\gamma$ effective coupling, as well as Hgg and $HZ\gamma$ ones, are controlled by independent parameters in the fit, therefore removing any assumptions on their parameterizations in terms of couplings of the Higgs boson to SM particles. The CMS result of $\kappa_t = 1.60^{+0.34}_{-0.32}$ has been derived imposing $\kappa_t \geqslant 0$ and thus does not provide any information on the sign. On the other hand, ATLAS allows negative values for κ_t and reports

$$\kappa_t \in [-1.12, -1.00] \cup [0.93, 1.60], \tag{1.29}$$

assuming that the Higgs boson decays only to SM particles, $\Gamma_{\text{BSM}} = 0$. This result is further illustrated by the profile likelihood ratio as a function of κ_t, which is shown in Fig. 1.3. The orange curve corresponds to a fit when all loop-induced couplings of the Higgs boson are fully defined by modificators of its tree-level couplings to SM particles. It is evident that the $\kappa_t = -1$ case is excluded at a more than 99.7% CL. The blue curve shows the results in the case when all loop-induced couplings, including $ggZH$, can be varied independently in the fit. In this generalized model the sensitivity to the sign emerges exclusively from the tH contribution in the $t\bar{t}H$, $H \to \gamma\gamma$ search, and the $\kappa_t = -1$ case is only disfavoured at about 68% CL.

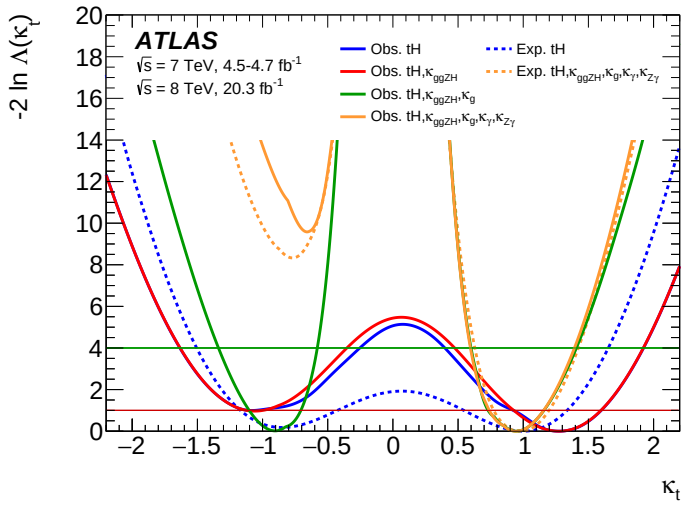


Figure 1.3: Experimental constraints on κ_t from a global fit to ATLAS data [46]. The red (green) horizontal line indicates the cut-off value of the profile likelihood ratio corresponding to a 68% (95%) CL interval for κ_t.

The strength of the HWW coupling is also measured in the global fit. The following values are found using the same generalized model with $\Gamma_{\text{BSM}} = 0$ as discussed above:

$$\kappa_W = 0.92^{+0.14}_{-0.15} \quad \text{(ATLAS)}, \qquad \kappa_V = 0.96^{+0.14}_{-0.15} \quad \text{(CMS)}, \tag{1.30}$$

where the CMS fit assumes $\kappa_W = \kappa_Z$. These results are predictably more precise than the measurements of the absolute value of κ_t. Their relatively small

uncertainties imply that future searches for tH production can take $\kappa_W = 1$ as a good approximation, and it will be more important to vary $|\kappa_t|$.

There are no direct constraints on the $\mathcal{CP}$-violating top quark Yukawa coupling. According to studies of angular distributions in decays $H \to VV^*$ [41, 47], the Higgs boson is a $\mathcal{CP}$-even scalar. Although the pure pseudoscalar hypothesis is strongly disfavoured, a substantial admixture of the $\mathcal{CP}$-odd state is still allowed. Moreover, in a general case the amount of $\mathcal{CP}$ violation may differ between decays to weak bosons and fermions. In Ref. [48] a global combination of ATLAS and CMS results has been performed to set an upper limit $|\tilde{\kappa}_t| < 0.4$ by exploiting the contribution of the top quark to the Hgg and $H\gamma\gamma$ loop-induced couplings. However, the most stringent constraint is currently imposed by the upper limit on the electric dipole moment (EDM) of the electron, $|d_e| < 8.7 \cdot 10^{-27}\, e \cdot \text{cm}$ [49], which translates to $|\tilde{\kappa}| < 0.01$ [48]. At the same time, it should be noted that the interpretation of the EDM results assumes the SM interaction between the Higgs boson and the electron and can also be modified in extensions of the SM. Thus, a direct search for $\mathcal{CP}$ violation in top quark Yukawa coupling should be performed in future.

Chapter 2

Experimental setup

This search is performed with experimental data recorded by the CMS detector [40], which is installed at the Large Hadron Collider (LHC) [50]. The facility was built by the European organization for nuclear research (CERN) between 1998 and 2008 near Geneva, Switzerland.

2.1 Large Hadron Collider

The LHC is currently the largest and the most powerful particle collider. It is hosted in the 26.7 km underground tunnel previously occupied by the LEP collider, which lies between 45 to 170 m below the surface.

2.1.1 Design and nominal operating parameters

The LHC is a synchrotron designed to collide protons at a centre-of-mass energy of $\sqrt{s} = 14\,\text{TeV}$ and provide a luminosity of the order of $10^{34}\,\text{Hz/cm}^2$. It is also capable of colliding fully stripped lead ions $^{208}\text{Pb}^{82+}$ at an energy of $\sqrt{s_{NN}} = 2.76\,\text{TeV}$ per nucleon, but this operation regime is of no interest for this search and will not be discussed further in the text. The machine and its design parameters are briefly described below following Ref. [50].

The unprecedentedly high target luminosity demands large beam currents and thus cannot be achieved with proton-antiproton beams, which could be accommodated within a single beam pipe. The LHC uses two adjacent beam pipes,

in which protons or lead ions circulate in opposite directions. The LHC ring consists of eight circular arcs and eight straight sections. The straight sections are about 0.5 km long, as dictated by the pre-existing tunnel, and accommodate collision points, beam injection and extraction facilities, collimation systems, and radio-frequency systems to accelerate particles. There are four collision points, where the beams can be crossed. In the two high-luminosity points, located at diametrically opposite straight sections, the general-purpose ATLAS [39] and CMS experiments are installed. The other two collision points are occupied by the LHCb [51] and ALICE [52] experiments, which are designed to study b physics and heavy ions respectively.

The LHC relies on approximately 9 000 magnets of about 50 different types. Central to its operation are 1 232 dipole magnets that keep the beams on their quasi-circular orbits and whose bending power limits the maximal achievable centre-of-mass energy. Superconducting niobium-titanium coils are used to provide the nominal magnetic field of 8.3 T, with an ultimate limit of 9.0 T. The corresponding nominal current is 11.9 kA. The magnets are maintained at a temperature of 1.9 K. At this temperature the liquid helium used as coolant is superfluid and thus provides an exceptionally high thermal conductivity, which allows to refrigerate large structures efficiently. The magnets follow a twin-bore design. The two bores, each housing a beam pipe, have separate coils but share common mechanical structure, cryostat, and iron yoke and thus are magnetically coupled. This choice was motivated by the lack of space in the LEP tunnel, which would not allow to install two independent proton rings. Each dipole magnet, embedded into a cryostat, has a cylindrical shape with a length of about 16 m and a diameter of 1 m and weights about 35 t. In addition to the dipoles, 392 twin-bore superconducting quadrupole magnets focus the beams, and a number of multipole magnets introduce fine corrections to the magnetic optics of the LHC. Other groups of magnets are utilized for beam injection and extraction and instrument the straight sections with collision points, where the two beams are guided into a single beam pipe and squeezed to provide the desired luminosity. In particular, at the two high-luminosity collision points the beams are squeezed by quadrupole triples, decreasing their radius from about 1 mm to 17 μm. The total energy stored by the LHC magnets is about 10 GJ, mostly contained in the dipoles.

Protons are accelerated by superconducting radio-frequency cavity systems. The cavities are made of copper, with niobium sputtering, and are maintained at a temperature of 4.5 K. There are eight cavities per beam, each capable of delivering an accelerating voltage of 2 MV. They operate at a frequency of about 400.8 MHz, which corresponds to a harmonic number of 35 640 with respect to the revolution frequency of 11 245 Hz. The total energy contained in

a single beam under nominal conditions is about 360 MJ, and it can be reached in about 20 minutes.

Since the LHC is a synchrotron, it is not capable of accelerating particles starting from zero energy. Instead, it uses an injector chain consisting of other CERN accelerators, which have been upgraded to meet challenges of this task. The injection can be done in various ways, resulting in different LHC bunch structures. Here the standard filling scheme is described [53]. Protons, which are obtained by ionizing hydrogen atoms, are first accelerated to an energy of 50 MeV in a linear accelerator Linac 2. They then are injected into the Proton Synchrotron Booster (PSB), which increases the energy to 1.4 GeV. In the next accelerator in the chain, the Proton Synchrotron (PS), the LHC bunch structure starts to be formed. The PSB can accelerate up to four bunches at a time. The PS captures six bunches from two consecutive cycles of the PSB. It splits each bunch into three and then accelerates them to an energy of 25 GeV. Afterwards, the bunches are split in two twice, thus producing 72 bunches from the six PSB ones. At this point the LHC bunch spacing of 24.97 ns is formed. Finally, bunches in the PS are shortened to 4 ns. Depending on the step of the LHC filling procedure, bunches from two, three, or four PS cycles are accumulated in the Super Proton Synchrotron (SPS). They are accelerated to an energy of 450 GeV and shortened to 1.7 ns. Twelve cycles of the SPS are used to fill the LHC. With the nominal bunch spacing, there are 3 564 bunch places along the LHC ring. However, only 2 808 are filled because of the time needed for operation of injection magnets. They are organized into 39 groups of 72 bunches. Within each group the bunch spacing is 24.97 ns, but the groups are separated by larger gaps. The largest continuous gap is 119 bunches in length (3 μs), and it is used for beam dump, providing a safe time window to power on the extraction magnet. After the LHC is filled, protons are accelerated to the target energy, during which bunches are further shortened to 1 ns (at the 4 σ level).

The minimal LHC filling time is about 20 minutes. Taking into account also the time needed to reach the nominal LHC energy and to ramp the magnets down to 450 GeV after a beam dump, the minimal total turnaround time is about 1 hour. This is reasonably shorter than the design luminosity lifetime of about 15 hours, which is the time starting from the beginning of a fill after which the luminosity decreases by a factor of e.

For a physics analysis, the most important properties of an accelerator are its collision energy and instantaneous luminosity. The latter can be calculated as

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

where N_b is the number of protons in one bunch, n_b is the number of colliding bunches, f_{rev} is the revolution frequency, ϵ_n is the normalized transverse beam emittance, β^* is the amplitude function at the collision point, and $\gamma_r = E_{\text{beam}}/m_p$ is the relativistic γ factor. Here the expression $\pi\epsilon_n\beta^*/\gamma_r$ gives the area of the transverse beam cross section at the interaction point. In the nominal LHC configuration $\epsilon_n = 3.75 \cdot 10^{-6}\,\text{m} \cdot \text{rad}$, $N_b = 1.15 \cdot 10^{11}$ protons, and $f_{\text{rev}} = 11\,245\,\text{Hz}$. The number of colliding bunches and the amplitude function depend on the collision point. At the two high-luminosity points the beams are squeezed to $\beta^* = 55\,\text{cm}$ and all bunches collide, i. e. $n_b = 2\,808$. In the vicinity of each collision point both beams share the same beam pipe of about 130 m in length. If the beams were collinear, this would result in about 30 parasitic collisions along the common beam pipe. In order to prevent this, the beams are crossed at an angle $\theta_c \approx 300\,\mu\text{rad}$. In Eq. (2.1) the resulting decrease of the luminosity is described by the geometric factor

$$F = 1 \Bigg/ \sqrt{1 + \left(\frac{\theta_c \sigma_z}{2\, r_{\text{beam}}}\right)^2}\,, \tag{2.2}$$

where σ_z is the bunch length and $r_{\text{beam}} = \sqrt{\epsilon_n\beta^*/\gamma_r}$ is the beam radius at the interaction point.

If the luminosity is high enough, several *pp* interactions can occur within the same bunch crossing. Their average number can be found as

$$\mu = \frac{\sigma_{\text{tot}} L}{n_b f_{\text{rev}}}, \tag{2.3}$$

where σ_{tot} is the total cross section of the *pp* inelastic scattering. For the nominal LHC conditions $\mu \approx 20$. This effect, known as "pile-up", increases the number of particles recorded in a single event and degrades performance of the event reconstruction. In addition, if a detector does not provide a fast enough response, its signals can be polluted by *pp* interactions that happened in the previous or following bunch crossings (so-called, "out-of-time pile-up").

2.1.2 Operation

Commissioning of the LHC with beams was started in 2008, but it was soon followed by a severe incident [54]. A faulty electric connection between a dipole and a quadrupole magnets produced an electric arc, which punctured the helium distribution line and provoked a destructive release of a large amount of helium into the insulation vacuum of the cryostat. The release also caused a significant mechanical damage, affecting 51 main magnets (dipoles and quadrupoles).

Recovery from the incident delayed the start of the LHC by one year and led to a revision of its operating parameters. First collisions for physics analyses, which were delivered to the experiments in 2010, were produced at a centre-of-mass energy of $\sqrt{s} = 7\,\text{TeV}$, factor two lower than the design value. The luminosity was also low, reaching a peak value of about $0.2 \cdot 10^{33}\,\text{Hz/cm}^2$ only in the end of year (see Fig. 2.1), which resulted in a total integrated luminosity delivered to the CMS experiment of about $44\,\text{pb}^{-1}$ [55]. However, in 2011 the luminosity was increased significantly, reaching a maximum of $4.0 \cdot 10^{33}\,\text{Hz/cm}^2$, and about $6\,\text{fb}^{-1}$ of data were delivered to the CMS. In 2012 the collision energy was raised to 8 TeV. During this year, the LHC operated with a bunch spacing of 50 ns and about 1 380 bunches per beam [56], most of which collided at the high-luminosity points. At these points the beams were squeezed to $\beta^* = 60\,\text{cm}$, providing a maximal luminosity of $7.7 \cdot 10^{33}\,\text{Hz/cm}^2$ and delivering about $23\,\text{fb}^{-1}$ of data. The mean number of pp interactions in a bunch crossing was about 21 on the average but varied with the luminosity (see its distribution in Fig. 2.2). The period 2010 – 2013, referred to as the Run I, was followed by the first long shutdown of the LHC (LS1), during which the machine was prepared for operation with a higher collision energy and detectors of the LHC experiments were upgraded.

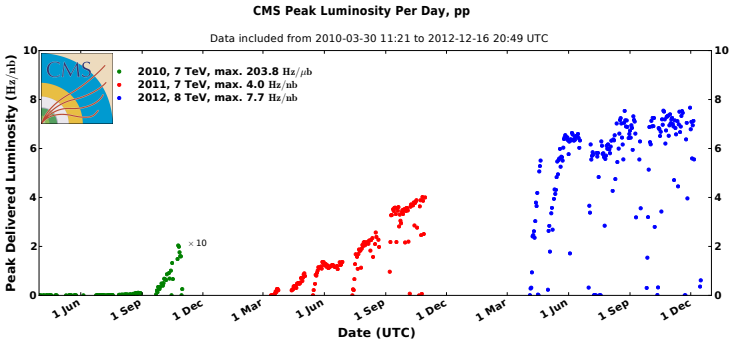


Figure 2.1: Per-day peak instantaneous luminosity at the CMS collision point during Run I [55].

In 2015 the LHC Run II was started. The centre-of-mass energy was boosted to $\sqrt{s} = 13\,\text{TeV}$. During the initial period, the LHC operated at a bunch spacing of 50 ns, delivering about $0.1\,\text{fb}^{-1}$ of data under these conditions, and then it was switched to the nominal 25 ns bunch spacing. By beginning of September 2015, the instantaneous luminosity at 25 ns bunch spacing peaked at about $0.9 \cdot 10^{33}\,\text{Hz/cm}^2$, while it is expected that a value of $1.8 \cdot 10^{34}\,\text{Hz/cm}^2$ might be reached by the end of the Run II in 2018 [57].

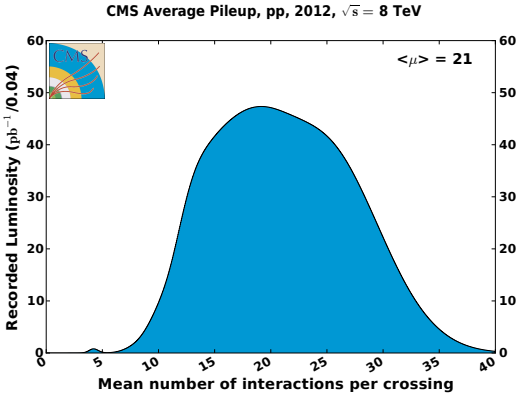


Figure 2.2: Distribution of the mean number of pp interaction per bunch crossing, as recorded at the CMS collision point in 2012 [55].

2.2 CMS detector

The Compact Muon Solenoid (CMS) is one of the two general-purpose detectors installed at the high-luminosity LHC collision points. It is located in an underground cavern in Cessy, France, occupying LHC Point 5. The detector has an approximately cylindrical shape, with a length of 21.6 m and a diameter of 14.6 m. The total weight of the structure is about 12 500 t. The detector is described in detail in Ref. [40], and its design performance is discussed in Refs. [58, 59].

The coordinate system adopted by the experiment has its origin in the nominal collision point inside the detector. The x axis points radially inwards, towards the centre of the LHC ring, while the y axis aims vertically up. The z axis, therefore, points towards the Jura mountains, in the direction of the anticlockwise beam. A cylindrical coordinate system is also used. The radial distance is calculated as $r = \sqrt{x^2 + y^2}$, and the azimuthal angle ϕ is measured with respect to the x axis in the (x, y) plane. The pseudorapidity is defined as $\eta = -\ln\tan(\theta/2)$, where the polar angle θ is measured with respect to the z axis.

Active elements of the detector are organized into a layered structure, as shown in Fig. 2.3. Closest to the interaction point is the inner tracking system, which allows to identify trajectories of electrically charged particles. It is surrounded by calorimeters, which measure energies of photons, electrons, and hadrons via total absorption. The outermost system consists of gaseous detectors and

registers muons, which are the only detectable particles capable of penetrating the calorimeters and steel absorbers.

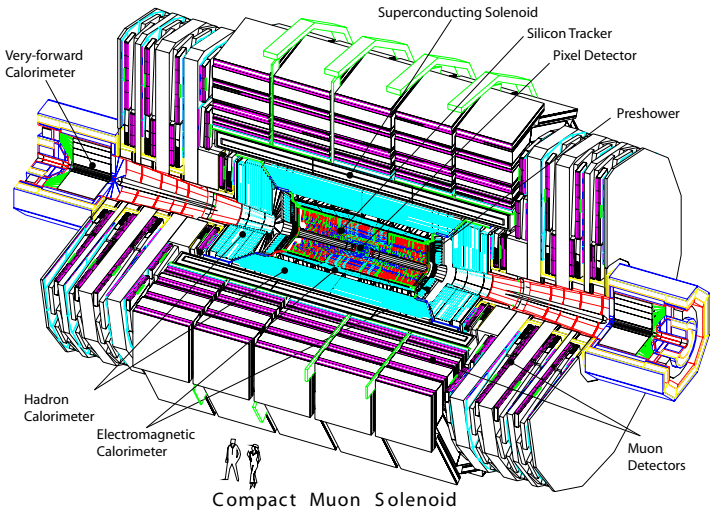


Figure 2.3: View of the CMS detector. Main elements are labelled.

The inner tracker and the calorimeters are accommodated inside a superconducting solenoid, which provides the magnetic field required to measure momentum of a particle based on the curvature of its trajectory. It has a length of 13 m and an inner diameter of 6 m. The niobium-titanium coil operates at a temperature of 4.5 K, which is maintained by liquid helium refrigerant. The magnet is designed to produce a highly homogeneous axial field inside the solenoid, with a magnetic flux density of 4.0 T in the centre. The nominal electric current required for this is 19.1 kA, and the stored energy is 2.6 GJ. However, in order to increase the longevity of the magnet, it is operated with a current lower than the nominal, resulting a field of 3.8 T [60]. The magnetic flux is returned through a 10 000 t yoke, which accounts for the most of the weight of the detector and provides mechanical support to the whole structure.

Operational regime of the LHC presents a number of challenges to be met by the CMS experiment. The short time between bunch crossings requires a fast operation of the trigger and read-out systems and a good time resolution of detectors. With the large number of particles stemming from pile-up interactions, a high granularity is needed to provide a sufficiently low occupancy. In addition, detectors and front-end electrons must be capable of operating under radiation levels caused by the high rate of particles coming from the interaction region.

2.2.1 Inner tracking system

The inner tracking system provides measurements of trajectories of charged particles originating from the interaction point. This information can be utilized to measure momentum of a particle from the curvature of the trajectory and to reconstruct primary and secondary vertices, which is important to recognize pile-up interactions and identify jets stemming from b quarks.

The tracking system is based entirely on silicon technology, which meets the requirements of high granularity, speed, and radiation hardness. The tracker has a length of 5.8 m and an outer diameter of 2.5 m. It is cooled to about +4°C in order to increase the longevity under the harsh radiation conditions, in particular, to protect against a potential damage that can be caused by an increased leakage current. The tracker consists of several components, as shown in Fig. 2.4.

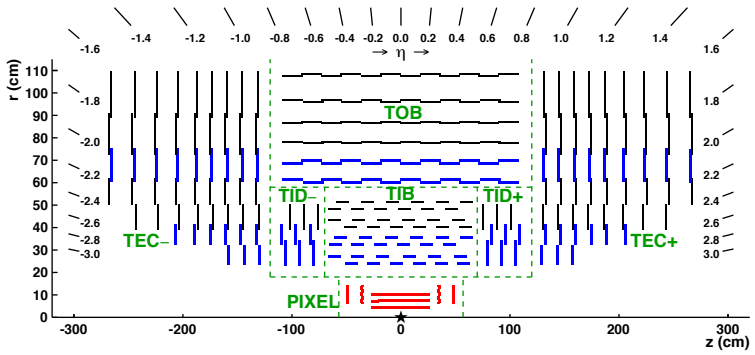


Figure 2.4: Half view of the inner tracking system [61]. Pixel detectors are shown in red. Blue lines mark back-to-back strip modules, which can measure the coordinate along the strip direction.

The innermost component is the pixel detector, which faces a very high flux of particles because of its location close to the interaction point. It consists of three cylindrical layers with radii of 4.4, 7.3, and 10.2 cm and a length of 53 cm, which are completed on each side by two endcap rings at $z = \pm 34.5$ and ± 46.5 cm with an inner (outer) radius of 6 cm (15 cm). The pixel cell size is $100 \times 150\,\mu\text{m}^2$. The detector has a total active area of about $1\,\text{m}^2$, resulting in $66 \cdot 10^6$ read-out channels. The occupancy under the nominal LHC conditions is of the order of 10^{-4} per pixel and bunch crossing.

The pixel detector is surrounded by the silicon strip tracker. Similar to the pixel detector, it includes elements of two types: cylindrical layers and annuli

perpendicular to the beam axis. Strips in the cylindrical layers a parallel to the beam axis, while in the annuli they follow the radial direction. The elements can be organized into four groups shown in Fig. 2.4. Tracker inner barrel and discs (TIB and TID) instrument radial distances from 20 to 55 cm and include four cylindrical layers and three annuli at each end. In the TIB the strip pitch varies between 80 and 120 μm, while in the TID the mean pitch ranges from about 100 to 140 μm. The TIB and TID are environed by the tracker outer barrel (TOB), which extends up to a radius of 116 cm and consists of six barrel layers with a strip pitch between 120 and 180 μm approximately. The region $124 < |z| < 282$ cm is instrumented with the tracker endcaps (TEC). Each of the two TEC is composed of nine annuli carrying up to seven rings with the mean strip pitch varying from about 100 to 180 μm. The strip length ranges from 10 cm in the innermost layers to 25 cm in the outer region. The coordinate along the strip direction can be measured by pairs of back-to-back detector modules rotated by 100 mrad, which are marked in Fig. 2.4. The silicon strip tracker has an active area of about 200 m^2 and $9.3 \cdot 10^6$ individual strips. The typical occupancy varies between 1% and $2 - 3$% depending on the region.

Both the pixel detector and the strip tracker provide a coverage up to $|\eta| \approx 2.5$. Each track within the acceptance has three high-precision hits in the pixel detector and at least about nine hits in the strip tracker, of which at least about four are delivered by the back-to-back modules. For tracks with a transverse momentum $p_\mathrm{T} \sim 100\,\mathrm{GeV}/c$ this allows a p_T resolution of $1-2$% up to $|\eta| \approx 1.6$, beyond which the resolution degrades because of the reduced level arm. The material budget of the inner tracker is about $0.4\,X_0$ at $|\eta| \approx 0$, where X_0 is the radiation length, and reaches a maximum of $2.0\,X_0$ at $|\eta| \approx 1.4$.

2.2.2 Electromagnetic calorimeter

The inner tracker is surrounded by the electromagnetic calorimeter (ECAL). It absorbs electrons and photons but does not contain enough material to stop typical hadrons.

The ECAL is made of lead tungstate ($PbWO_4$) crystals. This material has a high density ($8.3\,\mathrm{g/cm}^3$), short radiation length (0.89 cm), and small Molière radius (2.2 cm), which allows to construct a compact calorimeter with a high granularity. It also has a short scintillation decay time, delivering about 80% of the light yield within 25 ns. The light output, however, strongly depends on the temperature. To overcome this difficulty, the calorimeter is maintained at a temperature of $(18.00 \pm 0.05)^\circ$C.

The ECAL is partitioned into a barrel and two endcaps, as shown in Fig. 2.5. The barrel part covers pseudorapidity range of $|\eta| < 1.479$ and consists of 61 200 crystals mounted in a quasi-projective geometry. The inner radius of the barrel is 129 cm. Each crystal has a truncated pyramidal shape, with a cross section of $22 \times 22\,\mathrm{mm}^2$ at the front face and $26 \times 26\,\mathrm{mm}^2$ at the rear face; the length is 230 mm, which corresponds to $25.8\,X_0$. The granularity in the (η, ϕ) plane is approximately 0.0174×0.0174. The scintillation light is registered with the help of avalanche photodiodes mounted on the rear faces of the crystals.

The endcaps, whose front elements are located at coordinates $z = \pm 315.4$ cm, cover the rapidity range $1.479 < |\eta| < 3.0$. Each endcap includes 7 324 crystals, which also arranged in a quasi-projective geometry. The crystals have a cross section of about $30 \times 30\,\mathrm{mm}^2$ and a length of 220 mm ($24.7\,X_0$). The light is detected with vacuum phototriodes, which are better suited for the axial magnetic field and the larger level of radiation.

As measured in a test beam, the energy resolution for electrons with an energy of 20 GeV is about 1%, and it improves to below 0.5% for energies above 100 GeV [58].

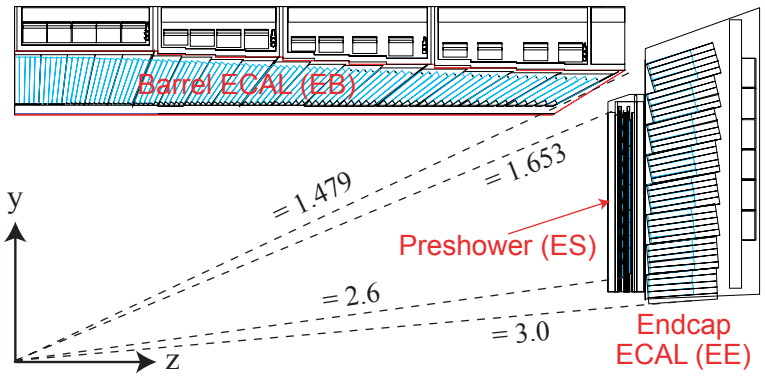


Figure 2.5: The electromagnetic calorimeter and the preshower.

In front of each endcap a preshower detector (PS) is installed, as shown in Fig. 2.5. Its aim is to distinguish between photons and neutral pions decaying into pairs of closely spaced photons. It also allows to improve position measurement for electrons and photons. The preshower covers a region $1.653 < |\eta| < 2.6$. It does not instrument the barrel because in the central region neutral pions have a smaller energy for the same transverse momentum, and the calorimeter granularity is sufficient to resolve the two photons. The preshower has thickness of 20 cm and consists of two layers of lead, each of which is followed by a silicon strip detector. At $|\eta| \approx 1.653$ the thickness of

the two lead radiators measures as 2 and 1 X_0 respectively, and most photons start showering before they reach the second sensitive layer. Strips in the both layers have a pitch of 1.9 mm. They are oriented orthogonally, which allows a two-dimensional measurement of the position. Similar to the silicon tracker, the preshower is cooled to below -5°C.

2.2.3 Hadronic calorimeters

Hadrons are registered with the help of several detector components, which are shown in yellow in Fig. 2.6. The barrel part (HB) is installed between the ECAL and the superconducting solenoid. It is complemented by two endcaps (HE). Additional hadron outer detectors (HO) are mounted in the barrel around the solenoid in order to catch potential tails of hadronic showers. The forward region is equipped with two dedicated calorimeters (HF).

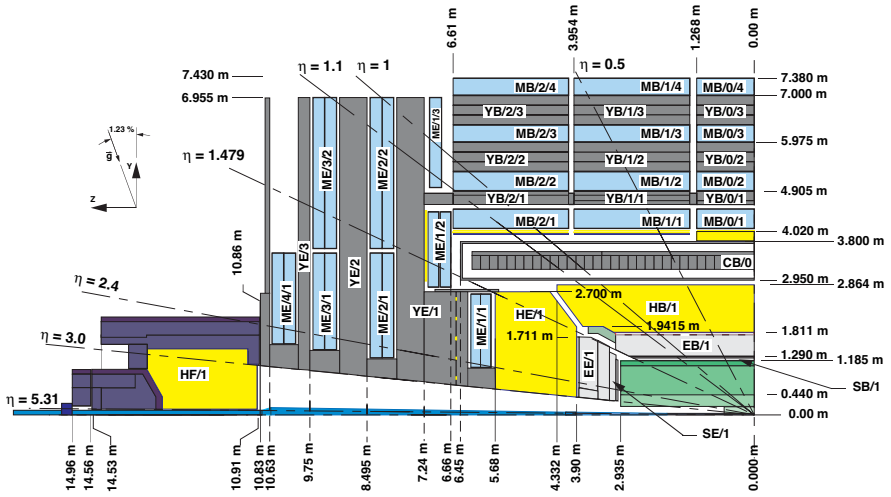


Figure 2.6: A quarter view of the CMS detector in the (y, z) plane [62].

The HB is a sampling calorimeter covering the region $|\eta| \lesssim 1.3$. It occupies a radial distance from 177 to 295 cm and consists of 36 identical azimuthal wedges, which are constructed out of flat absorber plates parallel to the beam axis. Each wedge is further segmented into four sectors in the azimuthal angle. The innermost and outermost absorber plates, with a thickness of 40 and 75 mm respectively, are made of stainless steel in order to provide an additional structural support. The intermediate absorber plates, 14 in total, are made of

brass and have a thickness of 50.5 or 56.5 mm. This translates into a total absorber thickness of 5.8 nuclear interaction lengths λ_I at $\eta = 0$, which increases to $10.6\,\lambda_I$ at $|\eta| = 1.3$. Between the absorber plates, plastic scintillating tiles are inserted, about 70 000 in total. They are organized into 17 sensitive cylindrical layers, which are divided into 32 segments along the z direction. The resulting segmentation in the (η, ϕ) plane is approximately 0.087×0.087. The scintillation light is collected by wavelength-shifting fibres and then registered with the help of hybrid photodiodes. Tiles within the same projective tower are read out collectively except for the four belts of towers closest to the HE (two belts on each side), in which the read-out is done independently for two longitudinal segments in each tower.

The HE covers the range $1.3 \lesssim |\eta| < 3.0$. It follows the sampling design and consists of 17 layers of brass absorber, each 79 mm thick, and 18 layers of plastic scintillator, which are built of $2 \times 10\,458$ tiles. The combined thickness of the HE and EE is about $10\,\lambda_I$. The granularity in the (η, ϕ) plane is 0.087×0.087 for $|\eta| < 1.6$ and approximately 0.17×0.17 in the more forward region. The scintillation light is collected and registered in the same way as in the HB. Three belts of projective towers closest to the beam pipe have three longitudinal segments that are read out independently, which allows to apply an additional calibration in order to recover from the radiation damage. Most of remaining towers are divided into two longitudinal segments.

In the central region the thickness of the HB is sometimes not sufficient to provide the full containment for hadronic showers. Their tails can be registered by an additional scintillator detector, the HO, which instruments the region $|\eta| < 1.3$. It utilizes the magnet as an additional absorber with an effective thickness of $1.4\,\lambda_I$ at $\eta = 0$. The scintillator tiles are mounted at a radial distance of 407 cm, directly in front of the innermost layer of muon detectors. In the region $|z| < 127$ cm the HO includes a second layer of tiles at a radial distance of 382 cm. A 19.5 cm thick steel absorber is inserted between the two HO layers, extending the combined thickness of the ECAL, HB, and HO to a minimum of $11.8\,\lambda_I$, except for the transition region between the barrel and the endcaps. The granularity of the HO in the (η, ϕ) plane is 0.087×0.087, provided by 2 730 tiles. Same as in the HB and HE, the read-out is performed with the help of wavelength-shifting fibres and hybrid photodiodes.

The HF instruments the region $3.0 < |\eta| < 5.2$. Because of its forward location, the detector experiences a large flux of particles and thus must be able to operate under very harsh radiation conditions, which has driven its design. Each of the two detectors has a cylindrical shape with an outer radius of 130 cm and a hole for the beam pipe with a radius of 12.5 cm. The front face of the HF is

located 11.2 m away from the interaction point. The detector exploits a 165 cm thick steel absorber, which corresponds to about $10\,\lambda_\mathrm{I}$. It is penetrated by quartz fibres parallel to the beam axis. Showers developing inside the absorber are detected with the help of Cherenkov light emitted when particles of the shower pass through the fibres. The HF is thus mostly sensitive to the electromagnetic component of showers. The fibres form a rectangular grid in the (x, y) plane with a step size of 5 mm. They are bundled to provide 0.175×0.175 segmentation in the (η, ϕ) plane. Only half of the fibres run through the full depth of the calorimeter, and the other half start at a depth of 22 cm from the front face of the detector. The two sets are read out separately and allow to distinguish between electromagnetic and hadronic showers since the former ones deposit a large fraction of their energy within first 22 cm of the absorber, while hadronic showers produce a more uniform deposition. The Cherenkov light from the quartz fibres is detected with the help of photomultipliers, which are protected from the radiation by dedicated shielding and connected to the fibres with the help of light guides.

The relative energy resolution of the HB, as measured with a test beam of charged pions, is $\sigma/E = 120\%/\sqrt{E} \oplus 9.5\%$, where the pion energy E is measured in GeV [63]. The resolution of the HE is $153\%/\sqrt{E} \oplus 6.3\%$ [64]. The presence of the ECAL causes a visible degradation of the resolution, and in order to obtain realistic results the both measurements were performed with a sector of the ECAL installed in front of the HCAL wedge. The HF resolution for pions is $198\%/\sqrt{E} \oplus 9\%$ [65]. Since there is no dedicated electromagnetic calorimeter in the region $|\eta| > 3$ and instead the HF attempts to discriminate between hadronic and electromagnetic showers based on the signatures read from the short and long fibres, it is also important to know the HF performance with electrons. The corresponding resolution is $280\%/\sqrt{E} \oplus 11\%$.

2.2.4 Muon system

The muon system identifies muons and provides means to trigger on them. It also improves momentum resolution for muons with $p_\mathrm{T} \gtrsim 1\,\mathrm{TeV}$, complementing measurements from the inner tracker. The system is mounted inside the return yoke of the magnet, representing the outermost detector layer of the CMS. Three different types of gaseous detectors are utilized, providing a total of $25\,000\,\mathrm{m}^2$ of sensitive planes.

In the barrel region the muon rate is low, the magnetic field is mostly constrained within the steel yoke, and the neutron-induced background is small. These conditions allow to utilize drift tube chambers (DT), which instrument

the region $|\eta| < 1.2$. They are assembled from rectangular drift cells, which have a cross section of $13 \times 42\,\mathrm{mm}^2$ and a length of about 2.4 m. In the centre of each cell an anode wire with a diameter of $50\,\mu$m is stretched. The chambers are organized into four concentric cylinders called stations, as shown in Fig. 2.7. Each station is divided in twelve azimuthal sectors and five wheels along the z direction. DT chambers of the three inner stations consist of three superlayers, each made of four layers of drift cells staggered by half a cell. Two of the superlayers have the wires aligned along the beam axis to bring the precision of a measurement in the (r, ϕ) plane to about $100\,\mu$m. They are separated along the radial direction as much as possible to improve the angular resolution. Between them the third superlayer is inserted, whose wires are oriented orthogonally to provide a measurement of the z coordinate. This third superlayer is missing in chambers of the outermost station. The DT system contains about 172 000 sensitive wires in total.

In the endcap regions the muon rates and background levels are high, and magnetic field is high and non-uniform. Cathode strip chambers (CSC) are used there to cover the region $0.9 < |\eta| < 2.4$. They are multi-wire proportional chambers consisting of six gas gaps, each one instrumented with cathode strips and anode wires. Cathode strips run radially and provide a ϕ measurement, while information read out from the anode wires is used to measure the pseudorapidity. The pitch varies with the radial distance, and its typical value is of the order of 1 cm; the anode wire spacing is about 3 mm. There are four CSC stations in each endcap, as shown in Fig. 2.8. The total active area is about $5\,000\,\mathrm{m}^2$. There are about 220 000 cathode strip read-out channels and 180 000 anode wire channels. Spacial resolution in the (r, ϕ) plane varies between $75\,\mu$m in the inner part of the first station to $150\,\mu$m in other chambers.

The rapidity range $|\eta| < 1.6$ is additionally instrumented by resistive plate chambers (RPC). Each chamber is a parallel-plate gaseous detector consisting of two continuous anodes and a shared cathode segmented into strips. RPC provide a coarser position measurement than DT or CSC but have an excellent time resolution of about 1 ns, which allows an unambiguous assignment of a muon to the bunch crossing. In total six layers of RPC arranged into four stations are installed in the barrel, accompanied by three layers in each endcap, as shown in Fig. 2.8. In the barrel strips are aligned along the beam axis, in the endcaps while they follow the radial direction. In both cases the pitch varies corresponding to a fixed azimuthal angle of $^5/_{16}^\circ$. The strip length is about 80 or 120 cm in the barrel, depending on the station, and ranges from about 25 to 80 cm in the endcaps. The total surface area is about $3\,000\,\mathrm{m}^2$, with 10^5 read-out channels.

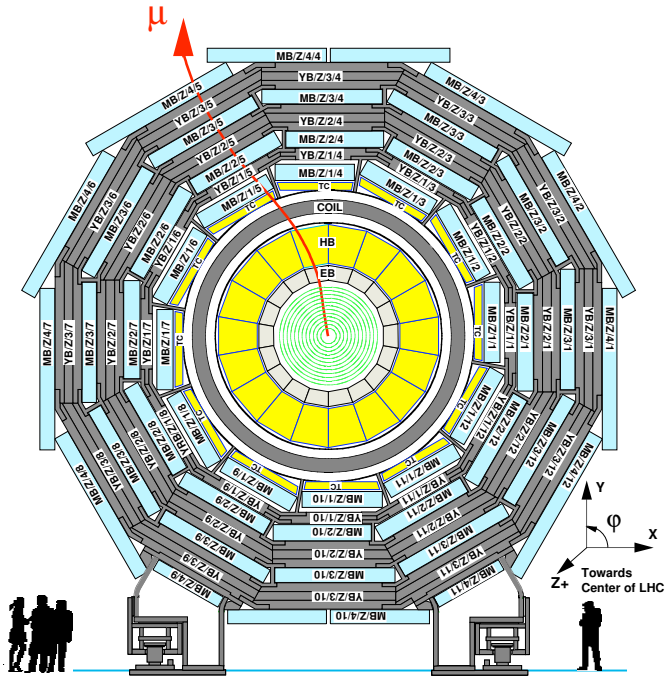


Figure 2.7: Cross section of the CMS detector in the (r, ϕ) plane showing the DT system.

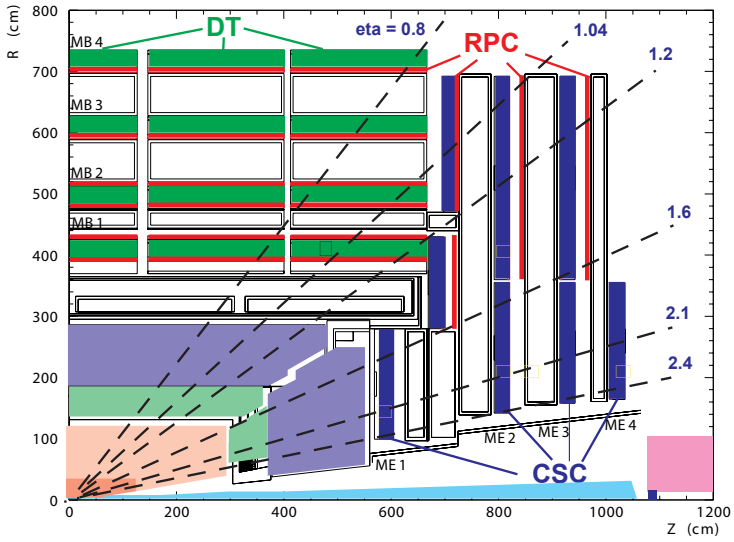


Figure 2.8: A quarter view of the CMS detector showing muon systems.

The muon system provides a muon identification efficiency of 95% to 99%, depending on the pseudorapidity and transverse momentum. The large amount of material in the calorimeters, the solenoid, and the return yoke, nearly eliminates the possibility of punch-through particles reaching the system. The design momentum resolution, as measured by the muon system alone, is of the order of 10% for $p_{\mathrm{T}} \lesssim 200\,\mathrm{GeV}/c$ and increases to 15 – 40% for $p_{\mathrm{T}} \sim 1\,\mathrm{TeV}/c$, depending on the pseudorapidity.

During the long shutdown of 2013 – 2015, a planned upgrade was performed in the CMS muon system. The forth CSC station in each endcap was completed by installing ME/4/2 chambers (see Fig. 2.6). In addition, coverage of the RPC system was extended to $|\eta| < 2.1$. These detectors were foreseen in the original CMS design, but their presence was not required during the Run I as the LHC operated at a luminosity lower than the nominal.

2.2.5 Trigger system

The CMS detector has been designed to operate with a bunch crossing interval of 25 ns, which corresponds to a (peak) collision rate of 40 MHz. It is not possible to store and process this amount of data, and thus a prompt decision should be taken on whether a particular event is to be stored for further analyses or discarded. This is done with the help of the trigger system. The CMS collaboration has adopted a two-level trigger system, consisting of the level 1 trigger (L1T) and the high-level trigger (HLT). L1T exploits custom-designed, largely programmable electronics. It bases its decision on information from calorimeters and muon system but does not exploit the inner tracker. It analyses every bunch crossing and decides whether to accept the event within 3.2 μs. The output rate is limited to 100 kHz by the time needed to read out full detector information. In order to provide a safety margin, the limit for the expected maximal output rate is set to 90 kHz. Events that pass the L1T are subject to a refined selection performed by the HLT. The HLT is a software system implemented in a farm of about 10^4 commercial processors. The average processing time is limited to 200 ms, but some rare events can be analysed for about 1 s. By the end of LHC Run I, the average HLT output rate has reached about 1 kHz, only 400 Hz of which were reconstructed promptly. The remaining accepted events were stored without immediate processing and reconstructed later, in between or after data-taking periods, when computing resources were available (so-called "data parking"). After the LS1, the maximum output rate was increased to 1 kHz, excluding the bandwidth for potential data parking. The HLT has access to full detector information. It exploits simplified versions

of the same event reconstruction algorithms as executed offline. These algorithms will be discussed in the next chapter, and this section focuses on the L1T.

The calorimeter part of the L1T starts by constructing trigger towers that sum transverse energy from electromagnetic and hadronic calorimeters. In the barrel region the tower size in the (η, ϕ) plane is 0.087×0.087, which corresponds to 5×5 ECAL crystals and a single HCAL cell. It increases to 0.35×0.087 in the most forward regions of the endcaps and to 0.5×0.35 in HF. The regional calorimeter trigger (RCT) identifies electron/photon candidates and calculates transverse energy sums in trigger regions, which are defined as groups of 4×4 or single trigger towers in ECAL+HCAL or HF respectively. It also provides information about calorimeter deposits to calculate isolation of muons. In each trigger region with $|\eta| < 2.5$ one isolated and one non-isolated e/γ candidate with greatest E_{T} is found. A tower with the largest energy deposit is identified, to which the energy of a broad-side neighbour tower with largest E_{T} is added. A candidate is required to pass a selection on the lateral profile of the shower in the ECAL, namely that the shower is contained within a block of 2×5 crystals. An additional requirement on the ratio between energy deposits in the ECAL and HCAL is applied. An isolated candidate must, in addition to it, pass a selection on the sum of ECAL E_{T} depositions in the eight surrounding trigger towers. The surrounding towers must also satisfy the requirements on the lateral shower shape and the ratio of ECAL and HCAL depositions. On top of this, at least one quiet corner of five trigger towers surrounding the hit tower is required. Four isolated and four non-isolated candidates with greatest E_{T} are identified per each of 18 RCT crates, each of which covers a region of $\Delta\eta \times \Delta\phi = 5.0 \times 0.7$. RCT also sums the transverse energy in every region and calculates τ veto bits, which are set if the active trigger towers occupy a continuous region of a size larger than 2×2 trigger towers since τ jets are typically narrower than quark or gluon jets.

The global calorimeter trigger (GCT) produces jets by clustering trigger regions, whose energies have been calculated by RCT. A jet is defined as a block of 3×3 trigger regions, which translates into 12×12 trigger towers in the central region ($|\eta| < 3$) and 3×3 towers in the HF. If none of the nine trigger regions fail the τ veto, the jet is pronounced a τ candidate. The jets are ordered in E_{T} and four leading ones in each category of central, forward, and τ jets are selected. GCT also calculates energy sums such as H_{T}, the scalar sum of E_{T} of all trigger regions, and missing transverse energy $\not{E}_{\mathrm{T}}$, or the imbalance in the vectorial sum of $\vec{E}_{\mathrm{T}}$, as well as jet multiplicities for several p_{T} thresholds. It receives the e/γ candidates and isolation bits for muons from RCT and propagates this information further.

All three muon systems contribute to the L1T, complementing each other. In each DT chamber, track segments in the (r, ϕ) plane are constructed from hits in the two ϕ superlayers. These track segments, together with hit positions in the z superlayers, are exploited by DT track finder (DTTF) to reconstruct full tracks. This is done independently in the (r, ϕ) and (r, z) planes. In the (r, ϕ) plane a source track segment is extrapolated to the next station based on a pre-calculated trajectory originating at the nominal interaction point. If a compatible segment is found, it is linked to the source segment. This allows to reconstruct the full projection of the trajectory. Tracks in the (r, z) plane are reconstructed directly from the hit pattern in the plane. The two projections are then matched in order to derive the full three-dimensional trajectory.

Each CSC chamber contains multiple layers of cathode strips and anode wires. Track segments are first reconstructed independently by cathode and anode electronics, and then the two projections are combined into three-dimensional track segments. Track segments from different chambers are joined into complete tracks by CSC track finder (CSCTF). It identifies pairwise combination of segments that are compatible with the hypothesis of a single track extrapolated from the interaction point. In the region $|\eta| \sim 1$ the track reconstruction is performed by utilizing information from both CSC and DT systems.

The RPC trigger does not attempt to reconstruct full tracks. Instead, it searches for coincident hits, profiting from its excellent time resolution, and measures the bending in the (r, ϕ) plane by comparing strip signals in projective trigger towers against predefined hit patters.

Information from the three muon trigger systems is combined by the global muon trigger (GMT). It receives up to four muon candidates from DT, CSC, barrel and endcap RPC systems, each. From GCT it also receives information about energy deposits in each calorimeter region of size $\Delta\eta \times \Delta\phi = 0.35 \times 0.35$, which is used to assess isolation of muon candidates. In addition to it, GCT provides a bit of compatibility with minimum ionization particle. DT and CSC candidates are matched to barrel and endcap RPC candidates respectively based on spacial coordinates, and if the match succeeds, their kinematic parameters are merged. Duplicate muons between the barrel and the endcaps are removed.

Trigger objects and global event properties, such as H_{T} or $\not{E}_{\mathrm{T}}$, constructed by GCT and GMT are forwarded to the global trigger (GT), which takes the decision to accept or reject the event. It supports up to 128 algorithmic trigger rules. The most basic algorithms apply a selection on transverse energy or momentum of a single trigger object, but they can also be more complicated, involving multiple objects and applying topological selection such as asking for

an η window. In addition, up to 64 technical triggers are supported. They are based on trigger signal received directly from subdetectors and used for purposes of calibration or commissioning. If needed, individual triggers can be masked, in case of which they do not affect the final GT decision. In addition, triggers can be prescaled so that only every k^{th} positive decision of a trigger with the prescale factor k is taken into account in the GT. It is typical to define several sets of prescales, or prescale columns, for the same trigger menu in order to allow it to be used for a range of instantaneous luminosities. Triggers with low thresholds can be prescaled at higher luminosities in order to respect the limitation on the maximal output rate. The prescale column can be changed without pausing the data taking. It is though guaranteed to be fixed during a *luminosity section*, a time period equal to 2^{18} LHC orbits, or approximately 23.3 s. The GT accepts an event if it has satisfied at least one trigger rule, taking into account the mask and the prescales. If this happens, the event is further scrutinized by the HLT.

Chapter 3

Event reconstruction

Particles of various types are produced in high-energy pp collisions, but only few live long enough to reach the detector. They are electrons, photons, muons, some hadrons (predominantly, pions, kaons, protons, and neutrons), and neutrinos. Neutrions cannot be registered because of their weak interaction with matter, while others generate response in different detector systems of the CMS. Charged particles produce signals in the inner tracker. Electrons and photons are absorbed in the ECAL. Hadrons traverse the ECAL and deposit most of their energy in the HCAL. Finally, muons can penetrate all layers of the CMS and reach the gaseous chambers. Signals obtained from different detector systems are combined and analysed to reconstruct the particles, deducing their types and measuring properties. The reconstructed particles are then exploited to test hypotheses about the pp collision, such as assessing its compatibility with the tHq production.

This chapter discusses the algorithms adopted to reconstruct particles and derived objects such as jets. The reconstruction is done in the same way for recorded collisions and their simulation. In addition, similar procedures are applied for the fast event reconstruction run in the HLT. In this case the algorithms are tuned to achieve a greater speed while keeping a sufficient reconstruction efficiency, but they usually follow the same general scheme as their counterparts in the offline reconstruction. The HLT reconstruction is thus not discussed here, and its details can be found in Refs. [66, 67].

3.1 Tracking

Reconstruction of tracks of charged particles is a crucial task. It allows a precise measurement of particle momenta via magnetic spectrometry. Tracks are exploited to reconstruct primary vertices, and a good vertexing capability is needed in order to mitigate the effect of pile-up by recognizing tracks that originate from additional pp collisions. Tracking is also vital to identification of jets that stem from hadronization of b quarks. At the same time, track reconstruction is a challenging task at the LHC. Under the nominal conditions, of the order of 10^3 charged particles can be produced in a bunch crossing, originating from an average of 20 pp interactions. This makes tracking a difficult combinatorial problem, especially because an accurate measurement of jet momenta requires reconstruction of particles with transverse momenta well below $1\,\mathrm{GeV}/c$. Furthermore, the reconstruction should be fast enough in order to run at the HLT level.

Algorithm deployed for reconstruction of tracks in the inner tracking system is summarized below. It has been evolving over time, but the general concept remained unchanged. Ref. [61] and references therein provide a detailed description of this algorithm as well as related vertex reconstruction and discuss their performance. A general overview of approaches used for track and vertex reconstruction can be found in Ref. [68].

3.1.1 Track parameterization

A helix track can be described by five parameters. It is customary to specify its position with respect to a reference point $\boldsymbol{r}_O = (x_O, y_O, z_O)$, such as reconstructed primary vertex. The perigee parameterization [69], which is adopted here, relies on the point of closest approach, in the transverse plane, of the track to the reference point, $\boldsymbol{r}_P = (x_P, y_P, z_P)$. The track can be parameterized as $(d_0, d_z, \phi, \theta, \kappa)$. Here $d_z = z_O - z_P$ is the longitudinal impact parameter, ϕ and θ are the azimuthal and polar angles of the direction of the particle momentum $\boldsymbol{p}$ at the point $\boldsymbol{r}_P$. The signed transverse impact parameter d_0 (see Fig. 3.1) has an absolute value $|\boldsymbol{d}|$, where vector $\boldsymbol{d}$ is the projection of $\boldsymbol{r}_P - \boldsymbol{r}_O$ on the transverse plane. By convention, d_0 is positive if $\boldsymbol{d} \times \boldsymbol{p}_\mathrm{T} \uparrow\uparrow \boldsymbol{z}$ and negative otherwise. The last parameter κ is the track curvature multiplied by the electric charge of the particle.

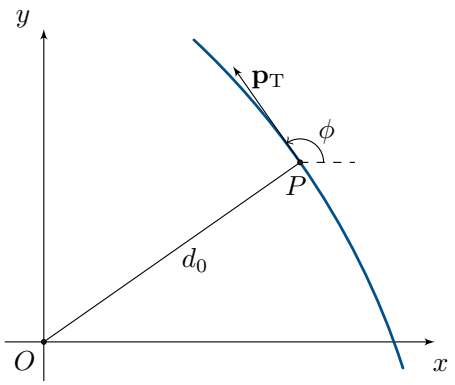


Figure 3.1: Perigee track parameterization in the transverse plane. The track is shown by the blue line. O and P are projections of the reference point and the point of closest approach on the transverse plane. In this case the transverse impact parameters d_0 is positive.

3.1.2 Hit reconstruction

Before the track reconstruction is started, zero-suppressed signals in individual pixel and strip modules are clustered to produce *hits*. In case of strip modules, coordinate of a hit is found as the charge-weighted average of positions of strips included in the cluster. Although in pixel modules the clustering is performed in two dimensions, position of a hit along each of the two module axes is determined from a one-dimensional projection of the corresponding cluster to the given axis. Two different algorithms are used depending on the step of the track reconstruction.

In the simpler first-pass version, which also performs the calculation faster, the geometrical centre of the projected cluster is used as the initial approximation. It is then improved by taking into account the ratio between the charges in the two boundary pixels of the projected cluster as well as the expected width of the projection. The expected width is calculated based on the incidence angle of the track relative to the plane of the detector. If the track is not known, it is approximated by a straight line drawn from the geometric centre of the inner tracker.

The more advanced algorithm, so-called template-based reconstruction, utilizes the full shape of the projected charge distribution. The coordinate of the hit is determined by fitting the distribution with templates derived from detailed simulation of a pixel module. The templates are parameterized by the incidence

angle of the track. They also account for the radiation damage, allowing to partly recover from the corresponding degradation of performance.

Coordinates of both pixel and strip hits are corrected for the Lorentz drift of the charges within the body of the silicon detectors. In addition to the positions, their uncertainties are also estimated. The positions are then translated to a global coordinate system, and the uncertainties are updated taking into account possible misalignment of the tracker modules.

3.1.3 Iterative tracking

Reconstructed hits are used as the basis to find tracks of individual particles. This is a twofold task, which includes identification of hits produced by the same particle and determination of parameters for its trajectory. The tracking algorithm adopted by the CMS is referred to as the combinatorial track finder (CTF), which is an adaptation of the combinatorial Kalman filter [70]. The Kalman filter, first applied in the context of the high-energy physics in Ref. [71], is advantageous because it addresses the both aspects of the track reconstruction. The CTF is applied iteratively. In early iterations it reconstructs tracks that are easiest to find, e. g. those with large transverse momenta and produced in the vicinity of the interaction region. Then hits associated with tracks are removed, and subsequent iterations attempt to find more difficult tracks, profiting from the reduced combinatorial complexity. Each iteration includes four steps, which are discussed below one by one.

At the beginning of each iteration, initial track candidates, or seeds, are generated. They are constructed using only three-dimensional hits, either from the pixel detector or the back-to-back strip modules. The pixel hits are reconstructed using the first-pass algorithm. The five track parameters can be extracted using three such hits or two hits and an additional constraint on the point of origin of the track. For this constraint, the centre of the beam spot or primary vertices reconstructed with a simplified algorithm using only pixel hits, are utilized; their reconstruction will be described in Section 3.2. The seeds are subjected to the filtering based on the estimated transverse momentum and compatibility with the imposed point of origin. It can also be checked that charge distributions of the hits agree with the expectation based on the track parameters, especially the incidence angles. The specific filtering criteria and requirements on the types of seeds (pixel or strip hits and their number) vary with the iteration of the algorithm. At late iterations seeds contain only hits from strip modules, and this allows to find tracks produced outside of the pixel detector volume.

At the second step, the Kalman filter is exploited to find tracks. The filter starts with a coarse estimate of the track parameters provided by the seed and improves it by propagating the track to successive detector layers, adding hits from there, and updating the parameters accordingly. Parameters of a track candidate estimated at the current layer are used to identify which adjacent layers of the detector can be intersected by extrapolation of the trajectory, taking into account current uncertainties. The extrapolation is performed analytically, assuming a homogeneous magnetic field and ignoring possible Coulomb scattering and energy losses. In the found layer, all silicon modules compatible with the trajectory are identified. Starting from this point, a more accurate propagator is used, which accounts for an increased uncertainty due to the multiple scattering and adjusts the momentum for the expected mean energy loss given by the Bethe–Bloch equation. All hits in these modules that are (loosely) compatible with the track candidate are selected. If there are none, a ghost hit can be added to the track in order to represent the possibility that a particle produced no detectable ionization in the module. Then new track candidates are formed from each original candidate by adding exactly one of the compatible hits (which might be a ghost). The new hits are used to update the track parameters by combining the information from the added hits with the extrapolated trajectory of the original track candidate. The procedure is repeated until the track candidate leaves the tracker volume, receives too many ghost hits, or its p_{T} drops below a threshold. In addition, track candidates that have the lowest compatibility with the included hits are removed at each iteration. When the search for hits in the outward direction identifies a sufficient number of them, all or some of the seed hits are removed from the track, and an inwards search is initiated. This allows to recover hits in the inner region. Finally, duplicate reconstructed tracks, which are identified based on the fraction of shared hits, are removed.

At the third step, the trajectory is refitted using a Kalman filter and smoother. The pixel hits are reconstructed with the template-based algorithm. The filter starts from the innermost hit, using the track parameters estimated by running the filter on few innermost hits. The filter processes hits in the outward direction. It is followed by a second run of a Kalman filter, in the inward direction, which starts with the obtained parameters of the track. Optimal track parameters at the position of each hit are then evaluated by averaging of the estimates by the two filters. In this step the propagation accounts for the inhomogeneity of the magnetic field, solving the underlying differential equation numerically using the fourth-order Runge–Kutta method. This is especially important for the region $|\eta| > 1$, where the inhomogeneity is greatest. Finally, tracks are cleaned against spurious hits that show a poor compatibility with the

rest of the track, and refitted again if needed, repeating this until no outliers are found.

The above steps produce a significant fraction of fake tracks, which do not correspond to any real charged particles. In order to remove them, the tracks are subjected to a quality selection. The requirements address the number of layers with hits, number of three-dimensional hits, number of layers without hits, the χ^2 fit quality of the track, the compatibility with the centre of the beam spot and at least one pixel vertex. The specific criteria depend on the iteration number.

The four steps presented above are repeated for each iteration of the CTF algorithm, resulting several track collections. The collections from different iterations are then merged, and duplicate tracks are removed. Efficiency of reconstruction of isolated muon tracks with $p_{\mathrm{T}} > 1\,\mathrm{GeV}/c$ is above 99%. On the other hand, up to 20%, depending on the pseudorapidity, of tracks of isolated pions with $p_{\mathrm{T}} \gtrsim 1\,\mathrm{GeV}/c$ are not reconstructed as a result of nuclear interactions in the material of the tracker; the inefficiency increases further in case of pions with smaller transverse momenta.

3.2 Primary vertices and beam spot

Using reconstructed tracks, points at which individual *pp* interactions occurred can be identified [61]. Reconstruction of primary vertices is carried out in two steps. First, tracks are clustered to find groups that are likely to originate from the same vertex. Then in each group a fit is performed to determine the vertex position. Reconstructed vertices as well as tracks alone can be used to deduce position and size of the *beam spot*, or the luminous region in which collisions take place.

3.2.1 Track clustering

Tracks exploited in reconstruction of primary vertices are subjected to an additional selection. They are required to contain a sufficient number of hits, in total and in the pixel system alone, and comply with a requirement on the minimal χ^2 fit quality. In addition, tracks with large transverse impact parameters with respect to the centre of the beam spot are excluded. On the other hand, no selection on the transverse momentum is applied in order to allow reconstruction of vertices that lack high-p_{T} tracks.

Selected tracks are clustered on the basis of z coordinates of the points of their closest approach to the centre of the beam spot. The procedure is described in detail in Refs. [61,72]. It exploits the deterministic annealing algorithm [73] to find the number of clusters N_c and their positions $\hat{z}_k$, $k = 1, \ldots, N_c$. The coordinate z_i of track i is used together with its uncertainty σ_i. Prior assessment of whether the track can originate from a primary vertex, which can be done based on such properties as the transverse impact parameter, is incorporated by the parameter $p_i \in [0, 1]$. It is interpreted as a probability to admit the track to the clustering. Given that track i is admitted, its assignment to cluster k is treated as a random event with an unknown conditional probability p_{ik}, which satisfies the normalization condition

$$\sum_k p_{ik} = 1 \quad \forall i. \tag{3.1}$$

The level of compatibility between the tracks and suggested cluster positions is quantified with the distortion function

$$E = \sum_i \sum_k p_i p_{ik} E_{ik}, \quad E_{ik} = \frac{(z_i - \hat{z}_k)^2}{\sigma_i^2}, \tag{3.2}$$

where indices i and k refer to tracks and clusters respectively. A straightforward minimization of the distorsion would introduce an independent cluster for each track so that $E = 0$. Instead, it is minimized while keeping the Shannon entropy of the system

$$S = -\sum_i (1 - p_i) \ln(1 - p_i) - \sum_i \sum_k p_i p_{ik} \ln(p_i p_{ik}) \tag{3.3}$$

at a constant level. This constrained optimization task can be equivalently formulated as minimization of the Lagrange function

$$F = E - T(S - S_0), \tag{3.4}$$

where T is the Lagrangian multiplier and S_0 is the chosen level of entropy, whose specific value, however, is not important for the following discussion.

Similarly to the well-known simulated annealing algorithm [74, 75], an analogy with a thermodynamical system can be built. In this case the distortion function E, the entropy S, and the parameter T are identified with the internal energy of the system, its thermodynamical entropy (defined up to an arbitrary additive constant), and the temperature. The Lagrange function F plays a role of Helmholtz free energy. If the temperature and volume of the system are kept fixed, its equilibrium state is given by the minimum of the Helmholtz free energy.

Minimization of F with respect to the association probabilities p_{ik} while imposing the normalization constraint (3.1), yields the Boltzmann distribution:

$$p_{ik} = \frac{\exp\left(-E_{ik}/T\right)}{\sum_{k'} \exp\left(-E_{ik'}/T\right)}. \tag{3.5}$$

Minimization with respect to $\hat{z}_k$ gives the equation for positions of the clusters:

$$\hat{z}_k = \frac{\sum_i p_i p_{ik} z_i/\sigma_i^2}{\sum_i p_i p_{ik}/\sigma_i^2}. \tag{3.6}$$

Eqs. (3.5), (3.6) can be solved iteratively for a given temperature T. The optimal number of clusters to be imposed in these equations is not known a priori. It is convenient to assume that the number is arbitrarily large, but positions of some clusters are identical. In the following the groups of coinciding clusters are treated as effective clusters. In this basis, Eq. (3.5) translates into

$$p_{iq} = \frac{w_q \exp\left(-E_{iq}/T\right)}{\sum_{q'} w_{q'} \exp\left(-E_{iq'}/T\right)}, \tag{3.7}$$

where w_q is the fraction of all clusters that have identical position $\hat{z}_q$ and thus form effective cluster q. Optimal relative weights of effective clusters are calculated as

$$w_q = \frac{\sum_i p_i p_{iq}}{\sum_i p_i}, \tag{3.8}$$

and their positions are still given by Eq. (3.6) if indices k, k' are understood to enumerate effective clusters.

The algorithm starts at a large temperature. In the limit $T \to +\infty$ all assignment probabilities p_{iq} are equal, and there is only a single effective cluster at

$$\hat{z}_1^{(0)}|_{T\to+\infty} = \frac{\sum_i p_i z_i/\sigma_i^2}{\sum_i p_i/\sigma_i^2} \tag{3.9}$$

with a weight $w_1^{(0)} = 1$. The temperature is gradually decreased, performing the annealing. When it falls below the critical value

$$T_q^* = 2 \sum_i \frac{p_i p_{iq}}{\sigma_i^2} \left(\frac{z_i - \hat{z}_q}{\sigma_i}\right)^2 \Bigg/ \sum_i \frac{p_i p_{iq}}{\sigma_i^2}\ , \tag{3.10}$$

it becomes advantageous to split the effective cluster in two. This event is analogous to a phase transition in a thermodynamical system. The splitting is forced, setting positions and weights of the daughter effective clusters to

$$\hat{z}_1^{(1)} = \hat{z}_1^{(0)} - \delta, \quad \hat{z}_2^{(1)} = \hat{z}_1^{(0)} + \delta, \tag{3.11}$$

$$w_1^{(1)} = w_2^{(1)} = w_1^{(0)}/2, \tag{3.12}$$

where δ is some small separation. As annealing continues, positions and weights of the clusters are adjusted by solving Eqs. (3.6)–(3.8). The algorithm then proceeds recursively, splitting daughter effective clusters when the temperature drops below the corresponding critical values (3.10).

If continued down to $T = 0$, the algorithm would generate an independent cluster for each track. To achieve a compromise between the ability to resolve closely spaced vertices and the risk of assigning tracks originating from the same vertex to multiple clusters, the algorithm is stopped at a temperature $T_{\min} = 4$, which is an empirical threshold. At this point the same track can be assigned to several clusters with comparable probabilities. To perform a strict assignment, the annealing is continued down to $T = 1$, but clusters are not allowed to split further. The assignment probabilities are modified by including an additional term in the denominator:

$$\tilde{p}_{iq} = \frac{w_q \exp\left(-E_{iq}/T\right)}{\exp\left(-E_0/T\right) + \sum_{q'} w_{q'} \exp\left(-E_{iq'}/T\right)}. \tag{3.13}$$

This breaks the normalization condition (3.1), allowing to downweight tracks that are poorly compatible with any cluster. In addition, each cluster is required to contain at least two tracks that are incompatible with all other clusters. Otherwise, the cluster is removed, and its tracks get reassigned. When the annealing finishes at $T = 1$, final clusters are constructed by keeping assignments with probabilities $\tilde{p}_{iq} > 1/2$. It is possible that some tracks are not included into any cluster.

3.2.2 Vertex fitting

Tracks included into each cluster constructed at the previous step are assumed to originate from the same primary vertex. The unknown three-dimensional vertex position $\boldsymbol{v}$ is found with the help of the adaptive vertex fitting algorithm [76]. The central element of the algorithm is minimization of the loss function

$$E(\boldsymbol{v}) = \sum_i w_i \chi_i^2(\boldsymbol{v}), \quad \chi_i = \frac{d_i}{\sigma_i}, \tag{3.14}$$

where d_i is the approximate distance from track i to the vertex position, σ_i is its uncertainty, and w_i is the weight, which will be discussed below. In order to simplify calculation of the distance d_i, tracks are approximated by straight lines in the vicinity of the current estimated vertex position.

The cluster can mistakenly include tracks that originate from other vertices, as well as misreconstructed tracks. Under these conditions, a straightforward

minimization of $\sum_i \chi_i^2$, which represents the method of least squares, may not provide a reliable estimate of the vertex position. Instead, a robust estimate in the presence of outliers is built by performing an annealing with dynamically adjusted track weights

$$w_i = 1 \Big/ \left(1 + \exp\left(\frac{\chi_i^2 - \chi_c^2}{2T}\right)\right), \tag{3.15}$$

where empirical constant $\chi_c = 3$ defines a threshold at which $w = 1/2$, and T is the temperature controlling the annealing. Outlying tracks, which have large values of χ^2, can be suppressed by small weights. The exact amount of suppression depends on the temperature. The algorithm starts from a high temperature, at which all tracks have similar weights, and then the temperature is gradually decreased, allowing to downweight the outliers. During the cooling, the loss function (3.14) is minimized iteratively. The algorithm stops when $T = 1$ is reached. After the annealing is finished, weights typically have values close to either 1 or 0. Tracks with weights smaller than $1/2$ are interpreted as outliers, while others are said to be compatible with the vertex.

The quality of the least-squares vertex fit with n tracks is usually quantified with the minimal value of the loss function $\sum_i \chi_i^2$, which follows the χ^2 distribution with $2n - 3$ degrees of freedom. The latter is not true for the loss function (3.14) because its probabilistic interpretation is altered by the track weights. An effective number of degrees of freedom

$$\tilde{n}_d = 2 \sum_i w_i - 3 \tag{3.16}$$

is used instead to characterise the quality of the fit. In this study, a vertex is considered reliably reconstructed if $\tilde{n}_d > 4$, which roughly corresponds to a requirement of at least four tracks compatible with the vertex. In addition, primary vertices are required to be contained in a cylinder $r < 2\,\text{cm}$, $|z| < 24\,\text{cm}$.

Because of the presence of overlapping pp collisions, several primary vertices are typically reconstructed in a single event. They are sorted in the decreasing order in the variable

$$\Sigma_{p_\text{T}^2} = \sum_i \big(\max(p_{\text{T}i} - \sigma_{p_\text{T}i}, 0)\big)^2, \tag{3.17}$$

where $p_{\text{T}i}$ is the transverse momentum of track i, $\sigma_{p_\text{T}i}$ is its uncertainty, and the sum runs over all tracks assigned to the vertex. The vertex with the largest value of $\Sigma_{p_\text{T}^2}$ is identified with the principal pp collision, while others are attributed to the pile-up.

3.2.3 Pixel primary vertices

As described in Section 3.1, a simplified vertex reconstruction is needed to build tracks. It relies only on hits in the pixel detector and runs much faster than the full reconstruction. The algorithm starts from pixel triplets formed in the same manner as described in the context of track reconstruction. The triplets are fitted to build pixel tracks, to which a selection on transverse momentum is applied. The tracks are clustered based on z coordinates of points of their closest approach to the centre of the beam spot with the help of the gap algorithm. The clustering algorithm scans an ordered list of the track coordinates $\{z_i\}$, starting from the smallest value, and finds all pairs of consecutive coordinates (z_i, z_{i+1}) such that the separation $z_{i+1} - z_i$ is larger than the predefined threshold $\Delta z_{\max}$. Cluster boundaries are put between the coordinates in each pair. The resulting consecutive blocks of coordinates define the track clusters. For each cluster that contains at least two tracks, the vertex position is determined with the adaptive vertex fitting algorithm described above.

3.2.4 Beam spot

The shape of the beam spot can approximately be described by a biaxial ellipsoid, which is typically displaced slightly from the nominal centre of the inner tracker and may be tilted with respect to the detector axes. Unlike other reconstructed quantities, parameters of the beam spot are determined by averaging over many events. The shortest period of time considered for the averaging is given by the length of a luminosity section.

The most straightforward way to measure parameters of the beam spot is to analyse the three-dimensional distribution of reconstructed primary vertices. An alternative approach [77] exploits the correlation between the transverse impact parameter d_0 and the azimuthal angle ϕ of a track at the point of the closest approach to the expected centre of the beam spot. If the transverse size of the beam spot is neglected, the two parameters be related by the equation

$$d_0(\phi, z^*) = \left(x_{\mathrm{BS}} + (z^* - z_{\mathrm{BS}}) \tan\theta_{\mathrm{BS}}^{(x)}\right) \sin\phi - \left(y_{\mathrm{BS}} + (z^* - z_{\mathrm{BS}}) \tan\theta_{\mathrm{BS}}^{(y)}\right) \cos\phi, \quad (3.18)$$

where z^* is the coordinate of the point of closest approach, $(x_{\mathrm{BS}}, y_{\mathrm{BS}}, z_{\mathrm{BS}})$ are coordinates of the actual centre of the beam spot, and angles $\theta_{\mathrm{BS}}^{(x)}$, $\theta_{\mathrm{BS}}^{(y)}$ define the orientation of the main axis of the beam spot. The position of the beam

spot in the transverse plane and the two tilt angles are found by minimizing

$$\chi^2 = \sum_i \left(\frac{d_{0i} - d_0(\phi_i, z_i^*)}{\sigma_i} \right)^2 , \tag{3.19}$$

where the sum runs over tracks considered for the calculation, and the uncertainty σ_i includes the uncertainty of d_{0i} and the transverse size of the beam spot. This method cannot determine z_{BS} and the size of the beam spot, which are instead deduced from the distribution of primary vertices and used as inputs to the algorithm.

It is also possible to determine the transverse size of the beam spot, σ_x and σ_y, using tracks [61]. A displacement of a primary vertex with respect to the centre of the beam spot due to its finite size introduces a correlation between transverse impact parameters of a pair of tracks that originate from the vertex:

$$\left\langle d_0^{(1)} d_0^{(2)} \right\rangle = \frac{\sigma_x^2 + \sigma_y^2}{2} \cos\left(\phi_1 - \phi_2\right) - \frac{\sigma_x^2 - \sigma_y^2}{2} \cos\left(\phi_1 + \phi_2\right) . \tag{3.20}$$

Here ϕ_1 and ϕ_2 are azimuthal angles of the two tracks at the points of their closest approach to the centre of the beam spot, whose position must be known. Although the method relies on reconstruction of primary vertices, it is not sensitive to the resolution of determination of vertex position.

3.3 Muon reconstruction

Muon reconstruction uses information from both the inner tracker and the muon system. The latter one provides means for muon identification and improves momentum resolution for high-p_{T} muons. The trajectory in the inner tracker is reconstructed as described in Section 3.1. In addition to it, an independent reconstruction is performed from signals in the muon system alone, producing so-called *stand-alone muon track* [58].

The stand-alone reconstruction starts by building track segments or single hits at the level of individual chambers of the muon system [78]. In the RPC detectors activated strips are clustered to produce hits, whose positions are determined as the centres of mass of the clusters. However, more elaborate procedures are applied in case of the DT and CSC.

In the DT chambers, the (r, ϕ) and (r, z) projections are first considered independently. Drift distances are calculated in activated cells, taking into account the magnetic field and the incidence angle, which is approximated using the direction to the nominal interaction point. Both possible hits that originate from

the right-left ambiguity are considered for every cell. The algorithm selects a pair of hits in two different layers, starting from the most separated ones. The pair must be compatible with a track pointing to the nominal interaction point. Hits compatible with the straight track segment defined by the pair are searched for in all layers, and a linear fit is performed. The procedure is repeated to find other segment candidates in the given projection. Then candidates with a poor χ^2 fit quality as well as duplicates are removed. Positions of hits used for each candidate are then improved by taking into account the incidence angle of the reconstructed track segment, and the linear fit is redone with the updated hits. Finally, the two projections are combined, keeping all possible combinations.

In the CSC detectors, hits are reconstructed independently in each of the six layers. First, strips are clustered, and azimuthal coordinates of future hits are found by fitting distributions of collected charge in the clusters. Radial coordinates are given by positions of read-out groups of anode wires, and the hits are reconstructed by considering all possible combinations from the two projections that agree in timing. Similar to the case of the DT, track segments are approximated by straight lines, and their reconstruction starts by considering pairs of hits from the first and the last layers that are roughly compatible with the direction to the beam spot. Compatible hits from the intermediate layers are added to the segment candidate, and a linear fit is performed. The track segment is accepted if it contains at least four hits and the fit quality is decent. In this case its hits are masked out, and the procedure is repeated.

After the detector-level reconstruction is completed, stand-alone muon tracks are built combining information from individual chambers with the help of the Kalman filter [71]. The procedure starts from the innermost layer of the muon system. When the track candidate is propagated to the next station, the algorithm searches for compatible measurements and uses them to improve the current estimate of the track parameters. The propagation is done taking into account the energy loss in the material, the effect of multiple scattering, and the inhomogeneous magnetic field. The DT track segments and RPC hits are exploited in the Kalman filter directly. On the other hand, individual CSC hits are used instead of the reconstructed track segments because the segments have been reconstructed assuming no magnetic field, which is a poor approximation in the endcap region. The procedure is iterated until the outermost layer of the muon system is reached. Then a backward Kalman filter is applied to calculate the track parameters at the innermost layer and extrapolate the track to the central region of the detector, where a requirement of compatibility with the first primary vertex is imposed.

Two approaches are used to combine the stand-alone muon track reconstruction and the iterative tracking [79]. With the *global muon reconstruction*, for each stand-alone track a matching CTF track is found by propagating the both tracks onto a common surface. The hits of the both tracks are then refitted using the Kalman filter. If the momentum calculated in the global fit is compatible with the measurement from the inner tracker alone and in both cases $p_{\mathrm{T}} > 200\,\mathrm{GeV}/c$, the muon is assigned the momentum from the global fit. Otherwise, the momentum is set according to the measurement in the inner tracker only.

In the second approach, the *tracker muon reconstruction*, all CTF tracks with a sufficiently large transverse momentum are propagated to the muon system, taking into account the energy loss, the multiple scattering, and the inhomogeneous magnetic field. If at least one DT or CSC track segment matches the reconstructed track, it is qualified as a tracker muon. Because of the looser requirement on activity in the muon system, this approach is more efficient than the global reconstruction for muon with a small momentum $p_{\mathrm{T}} \lesssim 5\,\mathrm{GeV}/c$.

About 99% of muons produced in pp collisions within the geometrical acceptance of the muon system and having sufficiently large momentum are reconstructed as either global or tracker muons, and often both. Fig. 3.2 shows momentum resolution measured in a sample of $Z \to \mu^+\mu^-$ events. There are also specialized algorithms to reconstruct cosmic or high-p_{T} muon, but they are out of the scope of this thesis. Performance of the muon reconstruction is discussed in detail in Ref. [79].

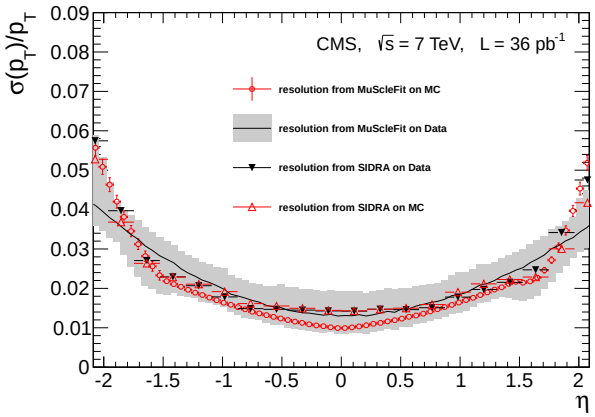


Figure 3.2: Muon momentum resolution, as determined by two complementary methods [79].

3.4 Electron reconstruction

Electrons are reconstructed by combining trajectory measurements obtained from the inner tracking system and energy depositions in the ECAL. The main challenge for the reconstruction is the bremsstrahlung, which an electron undergoes while traversing material of the tracker. On average, an electron radiates from about 30% to 90% of its total energy, depending on the pseudorapidity, before reaching the ECAL [80]. A special algorithm [80] has been developed to ensure an efficient reconstruction under these conditions. It combines two complementary approaches. First one starts from an ECAL cluster [81] and delivers best performance for high-p_{T} isolated electrons. However, its efficiency degrades for electrons with $p_{\mathrm{T}} \lesssim 10\,\mathrm{GeV}/c$ since their tracks have large curvature and bremsstrahlung photons can thus be distributed over a large area of the ECAL. In this case it is more efficient to start the electron reconstruction from track candidates [80,82]. This second approach was initially developed in the context of the particle-flow (PF) reconstruction [82,83], which is the subject of the next section. It is capable of reconstructing electrons with a transverse momentum down to $2\,\mathrm{GeV}/c$, even if they are surrounded by hadronic activity. The two approaches, ECAL- and tracker-seeded, are tightly entangled. They are described below.

3.4.1 Clustering of ECAL energy depositions

The ECAL-seeded reconstruction starts by clustering energy depositions in the ECAL. For an accurate measurement, it is important to collect the energy of bremsstrahlung photons, which mainly spreads along the ϕ direction due to the bending of the electron trajectory in the magnetic field. Different clustering algorithms are applied in the barrel and the endcap regions [84].

In the barrel, the so-called "hybrid" algorithm is used. It starts from a seed crystal with the largest transverse energy E_{T}. A window of 5×35 crystals, or approximately 0.09×0.6, in $\eta \times \phi$ centred at the seed is constructed. Inside the window, all 5×1 arrays of crystals in $\eta \times \phi$ with sufficient total energy depositions are selected, and adjacent arrays are clustered together. The set of all clusters in the 5×35 window that satisfy a certain selection on the energy, is called a supercluster (SC).

A different algorithm, the "multi-5×5", is applied in the ECAL endcaps. It uses as seeds crystals with an energy larger than those of their four side neighbours (the "Swiss cross" pattern). A cluster of 5×5 crystals is built around every seed. The clusters can partly overlap, but energy of each crystal can only be

attributed to a single cluster. The clusters are iterated in a decreasing order in the transverse energy of their seeds. Around each cluster with a sufficiently large total transverse energy, a 0.14×0.6 window in $\eta \times \phi$ is created, and an SC is formed out of all clusters that fall into the window and have not been assigned to a different SC at a previous iteration. Then the energy deposited in the preshower detector is added. This is done by projecting the energy-weighted positions of all clusters in an SC to the two sensitive layers of the PS and constructing a rectangular window in $\eta \times \phi$ large enough to accommodate projections of all clusters, with some margin added. The energy recorded in the window is attributed to the SC.

The tracker-seeded reconstruction also profits from an ECAL clustering, but of a different type. The seeds are chosen as crystals with locally maximal energy depositions and then clusters are formed by adding adjacent crystals whose energy depositions are not compatible with electronic noise. More details about the algorithm will be given in Section 3.5.1. In the following, the resulting clusters are referenced under the name of PF clusters.

3.4.2 Tracking

The standard track reconstruction described in Section 3.1 is not optimal in case of electrons. The Kalman filter (KF) it exploits assumes Gaussian fluctuations in the measurements, while the energy losses due to bremsstrahlung are highly non-Gaussian. In the Bethe–Heitler model [85] they can be described by a probability density function

$$p(z) = \frac{(-\ln z)^{t/\ln 2 - 1}}{\Gamma(t/\ln 2)}, \tag{3.21}$$

where z is the fraction of energy remaining after the electron has travelled in the material a path of length t, measured in units of radiation length. A more accurate reconstruction is achieved with the help of the Gaussian sum filter (GSF) algorithm [86], which models the observation error with a mixture of several Gaussian components and this way can approximate the distribution (3.21). Unlike the KF, the algorithm operates with multiple state vectors and corresponding covariance matrices, but in other aspects it is similar to the KF and follows the same alternating sequence of propagation and update steps. Indeed, it can be understood as several KF running in parallel [87], each corresponding to an individual Gaussian component. The track parameters can be estimated either from a weighted average of all components or a single most probable component. The second approach has been chosen for the reconstruction [81].

Before the GSF track reconstruction can be performed, track seeds are constructed. The seeds consist of two or three hits in the pixel detector or the TEC. In case of the seeds with two hits, pixel vertices are exploited to provide the missing constraint needed to define the track candidate. Two complementary approaches are followed.

The ECAL-seeded reconstruction exploits the SC constructed as described above to find seeds. SC are selected by imposing requirements on the minimal transverse energy and the maximal value of H/E_{SC}, where E_{SC} is the energy of the SC and H is the sum of energy depositions in the HCAL around the direction of the electron. For each selected SC, its energy and position are used to construct an approximate electron trajectory that originates at the nominal interaction point. This is done assuming a perfect helix, and both electric charge hypotheses are considered. The points of intersections of the approximate trajectories with innermost layers of the tracker designate potential positions of the seed hits. They are checked for coincidence with hits of the standard tracker seeds. The first hits are searched for using wide windows in $z \times \phi$ or $z \times r$, which also depend on the value of E_{SC}. When the first hits are matched, this information is used to refine the helices, and the windows for the second hits are tightened. Finally, a tracker seed is selected if its first two hits are matched with the predictions from an SC.

A different algorithm is used with the tracker-seeded reconstruction. The KF track is accurate when an electron undergoes little to no bremsstrahlung. The algorithm profits from this by attempting to match the standard CTF tracks to PF clusters. The matching is done based on the track position propagated to the ECAL and electron energies measured from the track curvature and the PF cluster. The tracks that fail this condition are analysed further. If the bremsstrahlung radiation is significant, the KF cannot follow the true trajectory efficiently and stops collecting hits early or produces a track with a low χ^2 fit quality. Tracks with such properties are refitted with the GSF algorithm. Then a multivariate analysis (MVA) is performed to select candidate electron tracks based on the χ^2 fit quality of the KF and GSF tracks and geometrical and energy matching between the tracker and ECAL measurements. Seeds of the tracks selected by either of the two methods are exploited in the subsequent electron reconstruction. Compared to the ECAL-seeded reconstruction, this approach is more efficient in case of low-p_{T} or non-isolated electrons, as well as in the transition region between the barrel and the endcaps.

Seeds identified by either the ECAL- or tracker-seeded reconstruction are then utilized to construct electron tracks. First, hits contributing to the track are found. Similar to the CTF tracking, this is done with the help of the KF. It

starts from an estimate of the track parameters obtained from the seed and then repeatedly propagates the trajectory to the next detector layer, searching for a compatible hit and adding it to the track candidate. The algorithm accounts for the mean energy losses given by the Bethe–Heitler equation (3.21), but their distribution is not modelled correctly. Instead, relatively loose criteria are applied to identify compatible hits in order to maintain a good efficiency. If multiple compatible hits are found, several track candidates are created and developed. To avoid following on hits from converted bremsstrahlung photons, tracks with missing hits are disfavoured.

Once all hits of a track candidate have been collected, the trajectory is fitted with the GSF. It provides accurate estimates of electron momentum both at the point of the closest approach to the centre of the beam spot, p_{in}, and on the surface of the ECAL, p_{out}. Then the fraction of energy lost due to bremsstrahlung can be quantified as $f_{\text{brem}} = (p_{\text{in}} - p_{\text{out}})/p_{\text{in}}$.

3.4.3 Track-cluster association

After GSF tracks are reconstructed, they are associated to ECAL SC to produce electron candidates. ECAL-seeded tracks are paired with the corresponding hybrid or multi-5×5 SC, requiring certain loose selection based on geometrical matching. On the other hand, a more elaborate procedure is applied in case of tracker-seeded candidates.

Tracker-seeded candidates are associated with PF SC, which are constructed as follows. First, the GSF track is propagated to the ECAL surface, and the matching PF cluster is included in the SC. Then for each layer of the tracker, a straight line tangent to the track is followed to the ECAL, approximating the path of a potential bremsstrahlung photon. If a matching PF cluster is found, it is added to the SC. This procedure allows to recover the most of bremsstrahlung radiation, but converted photons can be missed. They are found with the help of a dedicated algorithm that uses displaced CTF tracks and profits from an MVA pattern recognition. If photons reconstructed from the conversion are compatible with the electron track, they are associated with the SC. Finally, if the GSF track has at the same time been seeded by an ECAL SC, all PF clusters that are geometrically matched to the seed SC, are also added to the PF SC.

The compatibility between a GSF track and the associated PF SC is assessed with an MVA classifier, which combines information on properties of the track and the SC, as well as the quality of geometrical and kinematic matching between the two. A weak selection is applied on the minimal level of compatibility.

It can happen that two nearby GSF tracks are associated with the same SC, especially if a bremsstrahlung photon carries a significant fractions of the initial energy of the electron and undergoes a conversion. In this case a single track is selected, based on the number of missing hits, the E_{SC}/p ratio, and the type of the seed of the track (ECAL-seeded tracks are preferred to tracker-seeded ones).

The overall efficiency of reconstruction of electron candidates in $Z \to e^+e^-$ events is about 93%. It includes building of SC, reconstruction of GSF tracks, and establishing the association between the two.

3.4.4 Estimation of momentum

The electron momentum is calculated as a weighted average of the GSF track momentum p and the energy of the associated SC E_{SC}. The relative weight in the combination is found using an MVA regression, which bases its decision on the values of p and E_{SC}, along with their estimated uncertainties, and additionally exploits the bremsstrahlung class of the electron. The classification of the bremsstrahlung patter is based on the f_{brem} fraction and the ECAL footprint, mostly, the number of clusters in the SC. Three main classes include electrons with little radiation, those with a large amount of bremsstrahlung, which, however, is emitted in a single step, and electrons that produce a massive bremsstrahlung, essentially starting showering inside the tracker volume.

A number of corrections affect the momentum measurement. First, the energy response of individual ECAL crystals is calibrated before the clustering, including a correction for the temporary degradation caused by irradiation. Using an MVA regression, the SC energy as a whole is corrected for the leakage of the energy in the gaps between the crystals, in the HCAL, or outside of the SC window, and also for a contribution from pile-up. Finally, the combined momentum is adjusted using samples of $Z, J/\psi, \Upsilon \to e^+e^-$ events.

For electrons from Z boson decay, the resulting momentum resolution varies from 1.7% to 4.5%, depending on the bremsstrahlung pattern and pseudorapidity. As can be seen from Fig. 3.3, above an energy of about 35 GeV, the measurement is dominated by the ECAL response, but for $E \lesssim 15\,\mathrm{GeV}$ the track momentum is more precise than the ECAL measurement.

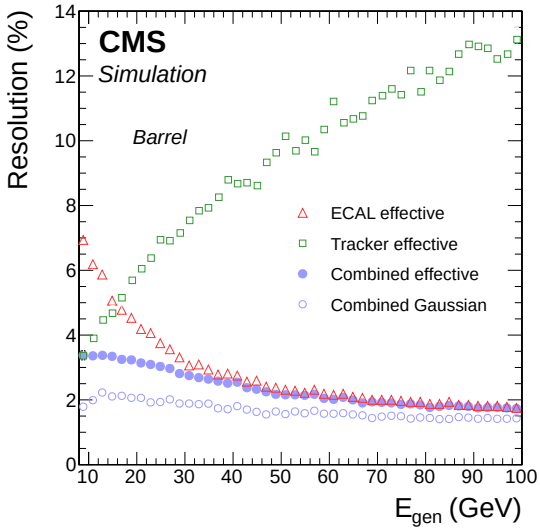


Figure 3.3: Electron momentum resolution after combining p and E_{SC} measurements, compared to estimates based on track momentum and SC energy alone [80]. Open circles correspond to the Gaussian core of the distribution.

3.4.5 Determination of charge

Electric charge of an electron can be deduced from the curvature of its GSF track. However, this judgement can be impaired in a case when the electron emits a photon and it undergoes conversion. In order to determine the charge in a robust way, the method using the GSF track is complemented by two independent measurements. The first one exploits the curvature of the KF track, which is matched to the GSF one by requiring that at least one hit in the innermost region of the tracker is shared between the two tracks. The second method deduces the charge from the position, in the (r, ϕ) plane, of the first hit of the GSF track with respect to the line connecting the beam spot and the SC. The charge of the electron is then chosen as the one shared by at least two of the three measurements. The probability of charge misidentification by this method can be measured from a sample of events with same-sign electron pairs whose invariant mass is close to the mass of the Z boson.

3.5 Particle-flow reconstruction

The particle-flow (PF) approach [82, 83] attempts to reconstruct each stable particle originating in an event individually. It combines information from all relevant detector systems to achieve the best performance. At the same time, it avoids potential double counting from assigning the same signals to multiple particle candidates. The CMS detector is particularly well-suited for the PF reconstruction. Thanks to its excellent tracking capabilities, trajectories of particles with transverse momentum down to $150\,\mathrm{MeV}/c$ can be reconstructed. The high granularity of the ECAL allows to resolve energy depositions of individual particles in jets with $p_\mathrm{T} \sim 10^2\,\mathrm{GeV}/c$. Although the HCAL has a much coarser granularity, its energy resolution is nevertheless sufficient to deduce the present of neutral hadrons by comparing energy depositions in the HCAL with momenta of associated tracks.

In the central pseudorapidity region, the PF reconstruction results in a list of muons, electrons, photons, charged and neutral hadrons. Muons and electrons are reconstructed independently as described in the previous sections and used as inputs to the PF algorithm. This section is thus focused on reconstruction of charged and neutral hadrons and photons.

3.5.1 PF blocks

Basic input elements considered in the PF-based reconstruction are CTF tracks and calorimeter clusters. They are grouped by building an association between tracks and clusters or clusters in different calorimeters. The association is done with the help of a linking algorithm, which constructs possible pairs of elements and quantifies the quality of the link in each pair. All elements linked together, directly or via other elements, are grouped into a *PF block*, which supposedly aggregates all parts of the footprint of a single particle in the detector.

The clustering of energy deposits is performed independently in each subdetector: EB, EE, HB, HE, and the two layers of the PS. The same algorithm is used in all cases, and it was already briefly discussed in the context of electron reconstruction. First, seeds are identified as cells with locally maximal energy depositions, above certain thresholds. Then clusters are formed by iteratively aggregating all cells that have a common side with a cell already included in the cluster and have an energy larger than a threshold. Same cell can be shared among several clusters. In such case, its energy is divided between the clusters according to the distances to their centres. On the other hand, in the HF each calorimeter cell is treated as an independent cluster.

The linking between a track and a calorimeter cluster is done as follows. The track is extrapolated from its last identified hit in the tracker to the two layers of the PS, to the ECAL, as a depth corresponding to the expected maximum of a typical longitudinal electron shower profile, or to the HCAL, at a depth of one nuclear interaction length. The track is then linked to a cluster if the extrapolated position is within the boundaries of the cluster, possibly with some margin to account for the gaps between cells or modules, uncertainty of the position of the shower maximum, and effects of multiple scattering. The link quality is calculated as the angular distance between the extrapolated track position and the cluster position.

A link between two calorimeter clusters is established if the position of the cluster in the more granular detector (PS or ECAL) falls within the cluster envelope in the less granular one (ECAL or HCAL). Similarly to the linking between a track and a cluster, an additional margin can be introduced for the matching. The quality of the link is given by the angular distance between positions of the two clusters.

3.5.2 PF algorithm

Individual particles are reconstructed from PF blocks. But before this can be done, elements of the blocks that correspond to reconstructed muons and electrons, are identified and removed, including the small calorimeter depositions of muons. Remaining CTF tracks are subjected to further selection, excluding those whose relative uncertainty on p_{T} is larger than the expected uncertainty for charged hadrons from calorimeters alone.

Although all reconstructed electrons are removed from PF blocks, an additional selection is applied in case of muons. The selection is hierarchical and consists of three steps. First, isolated global muons are identified and accepted. A muon is considered isolated for the purpose of PF reconstruction if the sum of p_{T} of tracks and E_{T} of calorimeter depositions in a cone of size $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.3$ around the muon is smaller than 10% of the muon p_{T}. Muons that do not meet this requirement are analysed further. If a muon satisfies a selection on the minimal number of hits and the compatibility of energy depositions in the calorimeters with expectations, it is also accepted. Finally, if some of remaining muons passes a relaxed selection on the number of hits, its track satisfies an additional requirement on the compatibility with hits in the muon stations, and the ratio between the calorimeter energy depositions and the track momentum is inconsistent with the hypothesis of a charged

hadron, the muon is accepted as well. Muons that do not satisfy any of the three variants of the selection, are not excluded from PF blocks.

Remaining PF blocks give rise to charged hadrons, photons, and neutral hadrons. More rarely additional muons can also be reconstructed. In order to achieve the best energy resolution, calorimeter depositions are calibrated, and these corrections depend on the target particle type. Energies of photons are deduced solely from the ECAL, and the calibration accounts for the energy loss due to the clustering thresholds. A more involved correction is applied in case of hadrons. Their energies are calculates as linear combinations of energy depositions in the ECAL and HCAL. The combination coefficients depend on the pseudorapidity of the HCAL cluster and the total uncalibrated energy of the hadron candidate. Same calibration is applied to charged and neutral hadrons. Both hadron and photon calibration are derived from simulation and validated with collision data [88].

A single PF block can contain multiple tracks, ECAL or HCAL clusters, with an elaborate link structure. It is simplified before the further processing, which might result in a splitting of the block. If a track is linked to several HCAL clusters, only the link to the closest one is kept. On the other hand, multiple tracks are allowed to be linked to a single HCAL cluster because of the coarse HCAL granularity. Also only link of a track to the closest ECAL cluster is kept if there are several ones. There might be additional ECAL clusters that originate from early showers of charged hadrons, in which case they should be linked to the corresponding tracks to avoid double counting of the energy of the hadrons, or overlapping photons, and then the links should be removed to allow reconstruction of the photons. To choose between the two possibilities, the ECAL clusters linked to tracks are first ordered according to their distance to the closest track. The list is then scanned, and the clusters are accepted one by one until the total calibrated energy of the corresponding HCAL cluster, if any, and all accepted ECAL clusters is smaller than the sum of momenta of the tracks. Links between the tracks and the accepted ECAL clusters are kept, and links to the remaining clusters in the list are removed.

In rare cases, the total calorimeter energy, calibrated under the hadron hypothesis, is smaller than the total track momentum in the block by a large amount. If it happens, an attempt to attribute the excess momentum to a muon is undertaken. The muon selection described above is loosened, and the additional muons are checked against the tracks in the PF block. If a match is found, it is added to the reconstructed muons. If, on the other hand, the excess cannot be explained by muons, the tracks are removed from the block in the decreasing order of their absolute p_{T} uncertainty, until all tracks with

the uncertainty above a certain threshold have been removed or the total track momentum is about to become smaller than the calorimeter energy.

Each of the remaining tracks in the PF block gives rise to a charged hadron candidate, whose momentum and energy are obtained from the track momentum under the $\pi^{\pm}$ mass hypothesis. If the calibrated calorimeter energy is compatible with the sum of momenta of all associated tracks within the uncertainties, it is used to refine the momentum measurements. This correction is especially important for high-p_{T} particles or in the forward region, where the track momentum is measured with a larger uncertainty.

Neutral particles are then reconstructed. Their presence in the vicinity of a track (or multiple tracks linked to the same HCAL cluster) is indicated by an excess of the calorimeter energy calibrated for hadrons, with respect to the track momentum. In this case a photon is formed from the ECAL energy deposition, with the dedicated calibration applied. If this is not sufficient to fully account for the excess, the remaining part of it is attributed to a neutral hadron. Photons are preferred to neutral hadrons in the ECAL because, on the average, about 25% of the jet energy is carried by photons, while neutral hadrons deposit in the ECAL only about 3% of the jet energy. All remaining ECAL and HCAL clusters are not linked to any track. They are interpreted as photons and neutral hadrons respectively.

A special albeit much simpler approach is followed to reconstruct particles in the HF. The HF lies outside of the tracker coverage, and thus no information about the charge of particles is available. Two types of candidates are reconstructed from the clusters: hadrons and electromagnetic particles, which combine photons and electrons. The discrimination between the two is performed based on signals from the long and short read-out fibres.

3.6 Physics objects

PF candidates are not used directly to define an event selection or construct observables. Additional identification criteria are applied to muons and electrons in order to suppress leptons that do not stem from the hard interaction, as well as misreconstructed leptons. Properties of quarks and gluons at the high energy scale are approximated with the help of jets, which are constructed by clustering of numerous PF candidates. Finally, all PF candidates are exploited to find the missing transverse energy $\vec{\not{E}}_{\mathrm{T}}$. Reconstruction of these physics objects is described below. In this search τ leptons and photons are not used directly, and thus they are not addressed here.

Performance of reconstruction of physics objects can be degraded by the pile-up. To mitigate its impact, some PF candidates are excluded from consideration via a procedure called *charged hadron subtraction* (CHS) [89]. Charged hadrons are associated with primary vertices, depending on the clusters in which their tracks have been included during the vertex reconstruction described in Section 3.2.1. Hadrons associated with any vertex except for the first one, are assumed to originate from overlapping pp collisions and thus can be removed. At the same time, hadrons that are not associated with any vertex, are kept.

Although the CHS is effective for charged hadrons, this technique cannot remove neutral component of the pile-up nor it is applicable in the forward region. In this case the average contribution from the pile-up, charged and neutral components together, can be found using the mean angular density of transverse momentum of pile-up particles. The density can be estimated on the event-by-event basis [90] as

$$\rho = \text{median}\left[\frac{p_{\mathrm{T}}^{(j)}}{A^{(j)}}\right], \tag{3.22}$$

which is calculated using jets clustered from all PF candidates according to the k_{T} algorithm [91, 92] with a distance parameter of 0.6. Here $p_{\mathrm{T}}^{(j)}$ and $A^{(j)}$ are transverse momentum of jet j and its geometric area respectively. Under typical conditions, the density given by the Eq. (3.22) is insensitive to jets from the hard interaction.

3.6.1 Muons

In this search global muons with $p_{\mathrm{T}} > 10\,\mathrm{GeV}/c$ and $|\eta| < 2.5$ are considered. To be selected, a muon must additionally meet the following set of identification requirements, which are based on the tight selection documented in Ref. [79]. It is required to pass the PF identification, described in Section 3.5.2. The goodness-of-fit of the global track must satisfy $\chi^2/\text{n.d.f.} < 10$, and the track must include at least one hit in a muon chamber. The corresponding tracker track is required to be matched to track segments in at least two muon stations, which implies that the muon is reconstructed also as a tracker muon. The transverse and longitudinal impact parameters of the tracker track with respect to the first primary vertex must satisfy $|d_0| < 2\,\mathrm{mm}$ and $|d_z| < 5\,\mathrm{mm}$, respectively. Finally, the muon is required to have hits in at least six tracker layers, including at least one hit in the pixel detector. Collective efficiency of these identification criteria is reported in Fig. 3.4.

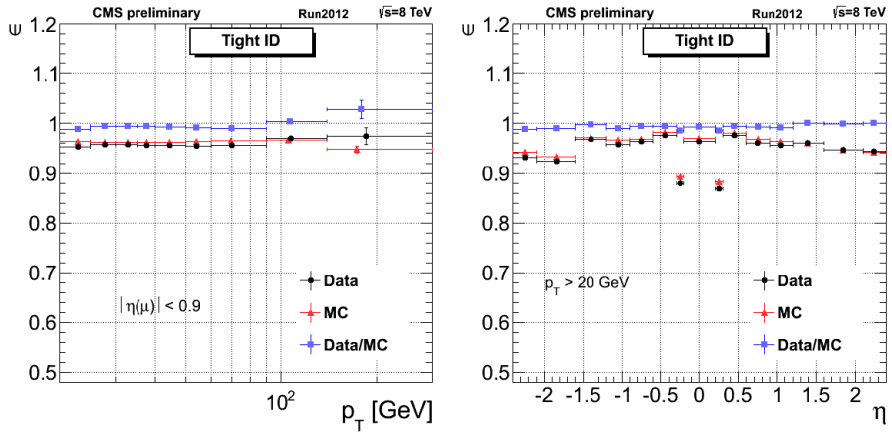


Figure 3.4: Efficiency of tight identification criteria applied to muons [93]. The drop at $|\eta| \sim 0.25$ is caused by a gap between DT wheels.

In order to suppress muons that stem from decays of hadrons in jets, the muon is also required to be isolated, satisfying $\mathcal{I}_{\Delta\beta} < 0.12$, where the relative $\Delta\beta$-corrected isolation is defined as

$$\mathcal{I}_{\Delta\beta} = \frac{1}{p_{\mathrm{T}}} \left(I_{h^{\pm}} + \max\left(I_{h^0} + I_{\gamma} - \beta \cdot I_{h^{\pm}}^{\mathrm{PU}}, 0 \right) \right). \tag{3.23}$$

Here p_{T} is the transverse momentum of the muon and $I_{h^{\pm},\, h^0,\, \gamma}$ are sums of transverse momenta of charged and neutral hadrons and photons, in a cone of size of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.4$ around the muon. $I_{h^{\pm}}^{\mathrm{PU}}$ is the contribution from charged pile-up hadrons, which are identified by the CHS procedure. The unknown contribution from neutral pile-up particles is assumed to be proportional to $I_{h^{\pm}}^{\mathrm{PU}}$, which allows to correct the total contribution of neutral PF candidates $I_{h^0} + I_{\gamma}$. The proportionality factor $\beta = 0.5$ is motivated by the expected ratio of numbers of neutral and charged pions [94]. It is validated by requiring that the corrected contribution from neutral PF candidates is stable with respect to the number of overlapping pp collision. Efficiency of the isolation requirement is shown in Fig. 3.5.

A muon that meets the above requirements is referred to as a "tight" muon. For the purpose of the event selection, also a relaxed identification is introduced, defining a "loose" muon. It is only required to pass the PF identification and satisty $p_{\mathrm{T}} > 10\,\mathrm{GeV}/c$, $|\eta| < 2.5$, and $\mathcal{I}_{\Delta\beta} < 0.2$.

This search exploits a single-muon trigger, whose requirements will be described in Section 4.4. Its efficiency is shown in Fig. 3.6. The drop at $|\eta| \sim 0.25$ is a

result of a gap between DT wheels. The asymmetry in the $|\eta| > 2.1$ region is caused by non-functional CSC chambers.

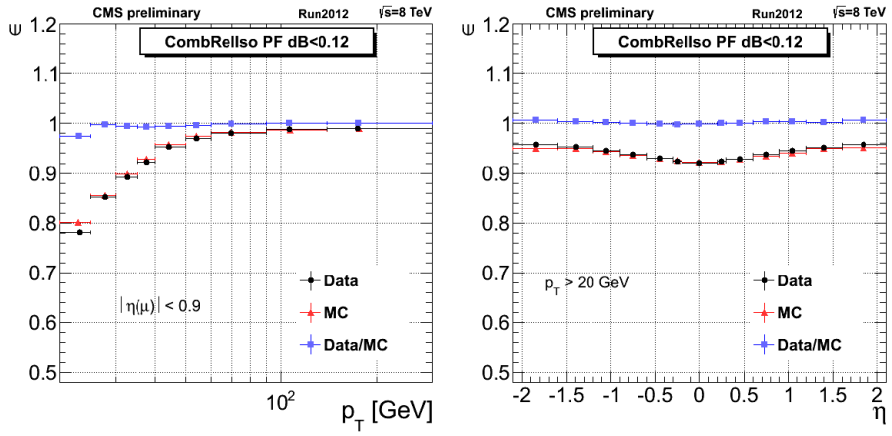


Figure 3.5: Efficiency of the isolation requirement $\mathcal{I}_{\Delta\beta} < 0.12$ [93].

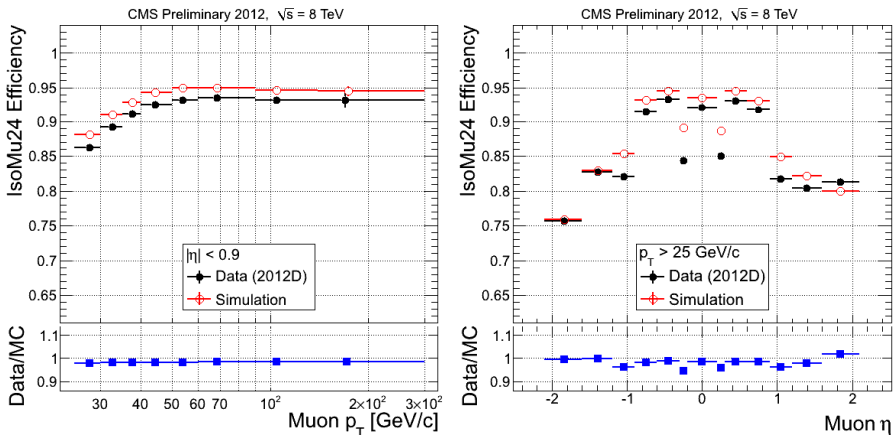


Figure 3.6: Efficiency of the muon trigger exploited in this search [93].

3.6.2 Electrons

An electron is considered if it satisfies the kinematic selection $p_{\mathrm{T}} > 20\,\mathrm{GeV}/c$ and $|\eta| < 2.5$. It is further required to pass an MVA identification [80], which combines three groups of observables. The first group represents information

from the tracker. Example observables of this kind are f_{brem} and χ^2 fit quality of the GSF track and the associated KF track. The second group characterizes the calorimeter footprint through such properties as the geometrical size of the SC along the η and ϕ directions or the variable $R_9 = E_{3\times3}/E_{\text{SC}}$, where $E_{3\times3}$ is the energy deposited in a 3×3 matrix of ECAL crystals centred at the highest-E_{T} crystal of the SC. The lateral spread of the shower is described by the variable $\sigma_{\eta\eta}$ defined as

$$\sigma^2_{\eta\eta} = s^2_\eta \sum_i (\eta_i - \bar{\eta}_{5\times5})^2 \, w_i \Big/ \sum_j w_j \,, \quad w_i = \ln \frac{E_i}{E_{5\times5}} + \text{const}, \tag{3.24}$$

as well as $\sigma_{\phi\phi}$, which is calculated in a similar way. The sum runs over crystals in a 5×5 matrix centred at the highest-E_{T} crystal, E_i is the energy of the crystal i, $E_{5\times5}$ is the total energy deposited in the 5×5 matrix, and $\bar{\eta}_{5\times5}$ is its energy-weighted mean position in the η projection. To make the observable insensitive to the varying size of the gap between crystals, the coordinates η_i and $\bar{\eta}_{5\times5}$ are expressed in units of crystals, and the average crystal size s_η is included to correct the overall scale. The third group of input variables includes observables that quantify the quality of matching between the track and the SC. The geometric matching is described by differences $\Delta\eta$ and $\Delta\phi$ between the track direction and the position of the SC, while the compatibility in energy is assess through such variables as E_{SC}/p and $1/E_{\text{SC}} - 1/p$, where p is the track momentum.

The MVA identification was constructed for electrons that pass a loose selection intended mimic requirements applied in the HLT. An electron is thus also required to pass this selection, which includes such observables as the number of missing hits in the GSF track, $\sigma_{\eta\eta}$, H/E_{SC}, and several versions of isolation. In addition, the transition region between the EB and EE, $1.4442 < |\eta_{\text{SC}}| < 1.5660$, is excluded.

The electron must not originate from a photon conversion. To suppress this contribution, no missing hits are allowed in the electron track. Futhermore, conversion vertices are reconstructed from pairs of electron tracks, exploiting the distinctive topology with the two tracks being nearly collinear at the vertex. If the electron is identified with a conversion vertex, it is rejected. Impact parameters of the electron with respect to the primary vertex are also used to exclude conversions.

Finally, the electron is required to be isolated, $\mathcal{I}_\rho < 0.1$, using the ρ-corrected isolation

$$\mathcal{I}_\rho = \frac{1}{p_{\text{T}}} \left(I_{h^\pm} + \max\left(I_{h^0} + I_\gamma - \rho \cdot A_{\text{eff}}, 0 \right) \right) \tag{3.25}$$

with a cone of size $\Delta R = 0.3$. Here ρ is the angular p_{T} density (3.22). The effective area A_{eff}, which depends on the pseudorapidity of the electron, is adjusted to compensate for the neutral component of the pile-up.

Similarly to the case of muons, the requirements above define a "tight" electron. They are complemented with a "loose" selection that accepts electrons with $p_{\mathrm{T}} > 20\,\mathrm{GeV}/c$, $|\eta| < 2.5$, and $\mathcal{I}_\rho < 0.15$.

Efficiency of the full tight electron selection described above is shown in Fig. 3.7. In this search also a single-electron trigger is utilized. Its requirements will be

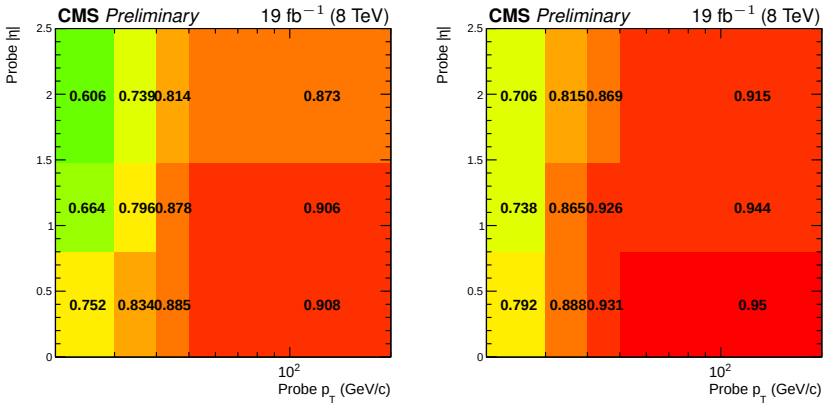


Figure 3.7: Efficiency of the tight electron selection in data (left) and simulation (right) [95].

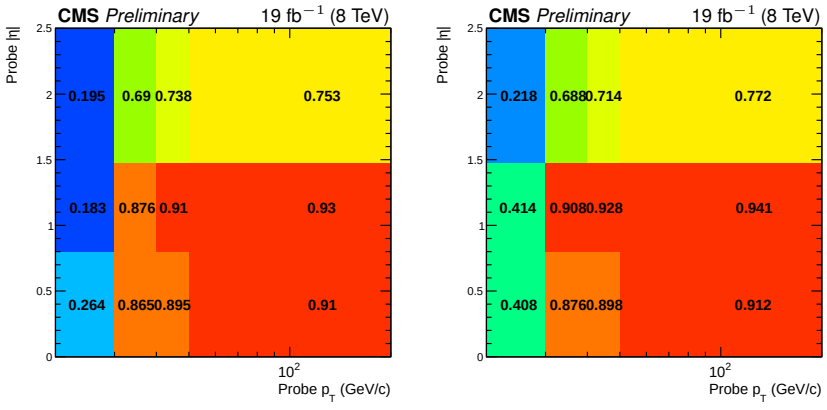


Figure 3.8: Efficiency of the electron trigger exploited in this search as measured in data (left) and obtained from simulation (right) [95].

summarized in Section 4.4, and the corresponding efficiency is provided in Fig. 3.8.

3.6.3 Jets

Jets are built from PF candidates, after removing charged hadrons identified by the CHS procedure. To ensure that particles reconstructed as leptons are not accounted for the second time as jet constituents, loose muons and electrons are also removed. The remaining particles are clustered using the anti-k_{T} algorithm [96] with a distance parameter of 0.5.

Produced jets are checked against loose identification requirements [97]. In order to be accepted, a jet must consist of at least two PF candidates. If the jet lies in the central region $|\eta| < 2.4$, it must contain at least one charged constituent, and a non-zero fraction of the jet energy is required to be attributed to charged hadrons. Finally, the jet is rejected if more than 99% of its energy is carried solely by particles of one of the following classes: neutral hadrons, photons, or electrons. These identification criteria are met by about 99% of physics jets.

There is a systematic difference between energies of reconstructed and corresponding particle-level jets, as illustrated by Fig. 3.9. Energies of reconstructed jets are corrected [98,99] in order to account for the contribution from pile-up, non-linearity of the calorimeter response, a slight detector mismodelling, and other effects. The corrections are applied sequentially, in three steps. They are

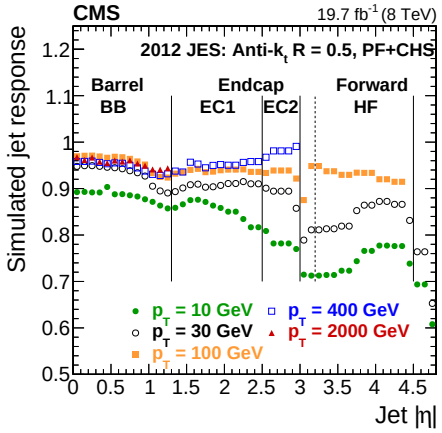


Figure 3.9: Jet response as obtainted from simulation [98].

multiplicative and rescale the energy of a jet and the magnitude of its momentum simultaneously, while the direction of the momentum is kept unchanged.

Jet energy receives some contribution from overlapping pp collisions. The charged component is partly removed by the CHS, and the remaining energy is subtracted by the first correction of the jet energy scale (JES), which is derived using so-called hybrid jet area method. The jet four-momentum is rescaled by a factor

$$C_{\text{hybrid}}(p_{\text{T}}, \eta, A, \rho) = 1 - \Big(\rho_0(\eta) + \rho \cdot \beta(\eta) \cdot \big(1 + \gamma(\eta) \cdot \log(p_{\text{T}})\big)\Big) \cdot \frac{A}{p_{\text{T}}}, \quad (3.26)$$

where p_{T} and η are uncorrected transverse momentum and pseudorapidity of the jet, A is its geometric area, and ρ is the angular p_{T} density. Essentially, this correction subtracts an amount of ρA from the jet p_{T}, although a number of amendments are introduced to achieve the best performance. Parameters $\rho_0(\eta)$, $\beta(\eta)$, and $\gamma(\eta)$ are derived from simulation, by comparing jets in the same dijet events reconstructed with and without the pile-up overlay. With a pile-up of 20 interactions, $C_{\text{hybrid}} \approx 0.85$ for central jets with $p_{\text{T}} = 30\,\text{GeV}/c$, and the correction factor approaches 1 as the jet p_{T} increases.

When applied to recorded events, the pile-up correction is slightly modified by adding a scale factor to the expression in parentheses in Eq. (3.26). This scale factor accounts for a difference between expected and observed contribution from pile-up.

At the second step, a simulated-response calibration is applied. It is derived purely from simulation and corrects the energy of a reconstructed jet such that on average is equals the energy of the corresponding particle-level jet. Particle-level jets are clustered, applying the same algorithm as for reconstructed jets, from all stable simulated particles that stem from the nominal pp collision in a bunch crossing, except for neutrinos. They are matched to the closest reconstructed jets within a cone of size $\Delta R = 0.25$. The correction factor is derived as a function of the jet transverse momentum (with the pile-up correction applied) and pseudorapidity. For jets with $p_{\text{T}} = 30\,\text{GeV}/c$, this correction factor is about 1.1 in the barrel and increased up to 1.3 in the forward region.

Lastly, residual corrections are applied to recorded data in order to account for small differences between the data and simulation. The corrections are derived in two steps. First, the residual dependence of the jet response on the pseudorapidity is measured in dijet events. In selected events, one jet is restricted to the barrel region $|\eta| < 1.3$, where the response is stable. The second jet scans the whole pseudorapidity range. By studying the balance between the two jets,

the response as a function of the pseudorapidity is measured relative to the average response in the barrel. This is done in several bins in jet p_{T}. After the derived intermediate correction is applied, an absolute p_{T} calibration is performed using jets in the barrel. Samples of $Z \to \mu^+\mu^- + \text{jet}$, $Z \to e^+e^- + \text{jet}$, $\gamma + \text{jet}$, and multijet events are exploited. In the first three samples, the jet is balanced against the Z boson or the photon, and the calibration profits from a good energy resolution of the reconstructed vector boson. In multijet events, a high-p_{T} jet in the barrel region is balanced against a system of two or more jets with lower transverse momenta. This allows to access larger values of p_{T} than can be done with $Z/\gamma + \text{jet}$ events. The p_{T} corrections are derived with a global fit combining all four samples. The overall residual corrections are less than 3% in the barrel, less than 10% in the endcaps, and about 10% in the HF, with an exception of the transition region $3.0 \lesssim |\eta| \lesssim 3.2$ where the corrections reach 20%.

Uncertainties of the full JES corrections are shown in Fig. 3.10. Typical values are below 1% in the central region. For most of jets considered in this search the uncertainties do not exceed 2%.

The JES correction addresses the average jet response but not the jet energy resolution (JER), which characterizes the variation of the response. In simulation, the resolution is derived from the spread of the distribution of $p_{\mathrm{T}}/p_{\mathrm{T}}^{\mathrm{gen}}$, where $p_{\mathrm{T}}^{\mathrm{gen}}$ is the transverse momentum of a particle-level jet and p_{T} is the corrected transverse momentum of the matched reconstructed jet, with the matching done as described above. As shown in Fig. 3.11 (a), the resolution in simulation is about 20% for jets with $p_{\mathrm{T}} = 30\,\mathrm{GeV}/c$ and a pile-up of 20 interactions. The JER is also measured in recorded data, using dijet and γ+jet events. Similarly to the JES calibration, the measurements profit from the balance conditions between the two jets or the photon and jet. The resolution is observed to be worse than expected from simulation, by up to 10% in the barrel and up to 40% elsewhere (see Fig. 3.11 (b)). This discrepancy is eliminated by smearing reconstructed jets in simulation. Their four-momenta are rescaled by a factor

$$C_{\mathrm{JER}} = \frac{\max\left(p_{\mathrm{T}}^{\mathrm{gen}} + s(\eta)\cdot(p_{\mathrm{T}} - p_{\mathrm{T}}^{\mathrm{gen}}), 0\right)}{p_{\mathrm{T}}}, \tag{3.27}$$

where p_{T} is the corrected transverse momentum of the reconstructed jet and $s(\eta)$ is the ratio of the resolution in data and simulation, measured as a function of the jet pseudorapidity.

This search exploits central jets ($|\eta| < 2.4$) with $p_{\mathrm{T}} > 20\,\mathrm{GeV}/c$ and forward jets ($2.4 < |\eta| < 4.7$) with $p_{\mathrm{T}} > 40\,\mathrm{GeV}/c$. The higher threshold in the forward region is motivated by a mismodelling of the detector response observed there.

In addition, pseudorapidities of forward jets are modified in order to mask the mismodelling. This problem will be discussed in detail in Section 4.3.

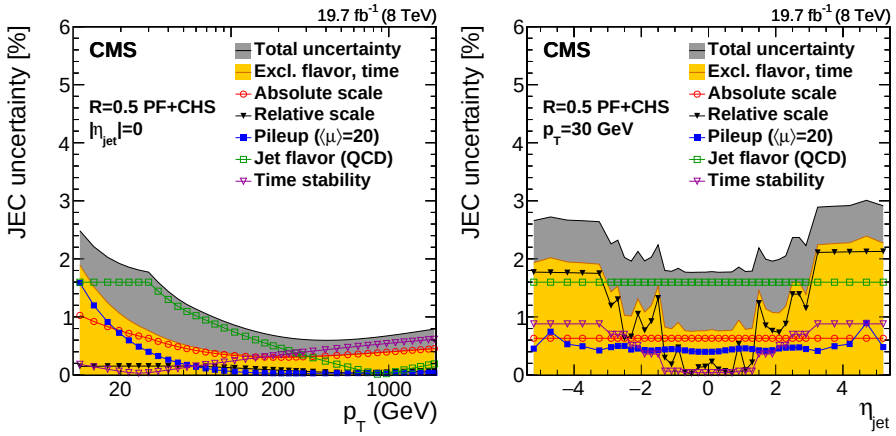
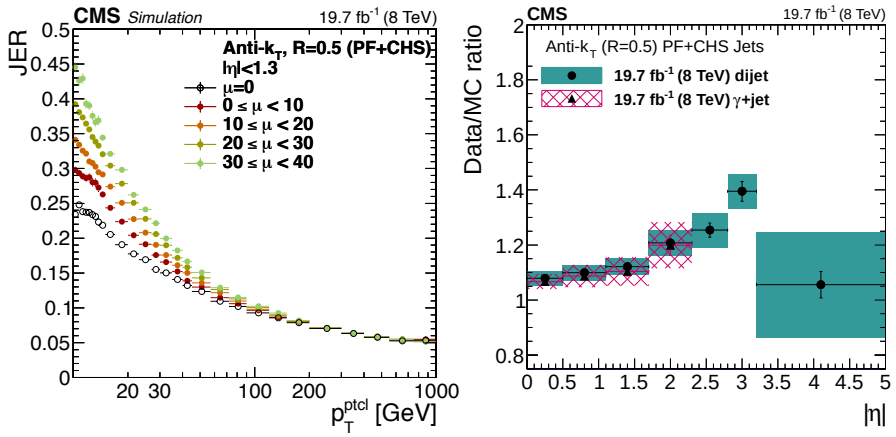


Figure 3.10: Uncertainties of JES corrections [100]. The ones considered in this search are shown by yellow band.



(a) (b)

Figure 3.11: Left: JER as derived from simulation for different bins in the number of overlapping *pp* collisions. Right: Difference between JER in data and simulation [98].

3.6.4 Identification of *b*-quark jets

The signature of the signal process considered in this search includes b quarks. It is thus important to identify reconstructed jets that stem from fragmentation of b quarks. This identification [101] is performed based on tracks associated with a jet. Due to the tracker coverage, the algorithm is restricted to jets with $|\eta| < 2.4$.

A subset of CTF tracks is exploited to deliver the b-tagging decision about a given jet. A track is associated with the jet if it falls within a cone of size $\Delta R = 0.3$ around the jet axis and its spatial distance from the axis at the point of closest approach is sufficiently small. The track is required to satisfy a selection that identifies well-reconstructed high-purity tracks. It must have $p_{\mathrm{T}} > 1\,\mathrm{GeV}/c$, contain enough hits in the whole tracker and the pixel detector alone, and pass a selection on the χ^2 fit quality. Based on the impact parameters, the track must also be loosely compatible with the first primary vertex. In addition, the decay length, which is defined as the distance from the primary vertex to the point of closest approach of the track to the jet axis, must be smaller than 5 cm.

The b-tagging can be done with tracks alone, by judging the presence of B hadrons from larger impact parameters of the tracks of their decay products. However, its performance can be improved if secondary vertices corresponding to B hadrons are reconstructed and their properties are analysed. When this is done, a tighter selection is applied to tracks in order to reduce the combinatorial complexity. Secondary vertices are reconstructed with the help of the same adaptive vertex fitting algorithm as used to reconstruct primary vertices, profiting from its robustness in the presence of outliers. The reconstruction is performed in an iterative manner. At each iteration, the algorithm finds a vertex and ranks tracks based on their compatibility with the vertex by assigning them weights (3.15). All tracks that receive weights larger than $1/2$, are removed from subsequent iterations. The procedure is repeated until no new vertex candidates are found. Reconstructed secondary vertices are then subjected to a selection. A secondary vertex is rejected if it shared more than 65% of tracks with the primary vertex or statistical significance of the radial distance between the two vertices is not larger enough. An additional selection on the maximal radial distance to the primary vertex and the mass of the secondary vertex is applied to suppress vertices formed by decays of long-lived mesons, such as K^0, or interactions of particles with the detector material. Finally, the flight direction of the B hadron candidate, which is approximated by a straight line drawn from the primary to the secondary vertex, must lie within a cone of size $\Delta R = 0.5$ around the jet axis.

Properties of the reconstructed secondary vertices and track-based information are utilized in the multivariate combined secondary vertex algorithm (CSV), which is exploited in this search. It uses such characteristics of the vertices as the significance of the flight distance, vertex mass, number of tracks associated with the vertex, the ratio of the total momentum of these tracks to all tracks in the jet, and others. If no secondary vertex is found for a jet, the algorithm attempts to construct a "pseudovertex" by grouping tracks with impact parameters of a large significance. Some of the input variables, such as the vertex mass, can still be calculated in this case. If even the "pseudovertex" cannot be reconstructed, the algorithm restricts itself to purely track-based observables, such as number of tracks associated with the jet and significances of three-dimensional impact parameters of the tracks.

Several working points are available for the CSV algorithm. In this search the tight one is exploited. Fig. 3.12 shows the corresponding b-tagging efficiencies measured in simulation in this analysis. Typical probability to accept a b-quark jet is about 55%, while the probability to incorrectly accept a light-flavour jet, which stems from an u, d, or s quark or a gluon, is of the order of 0.1%.

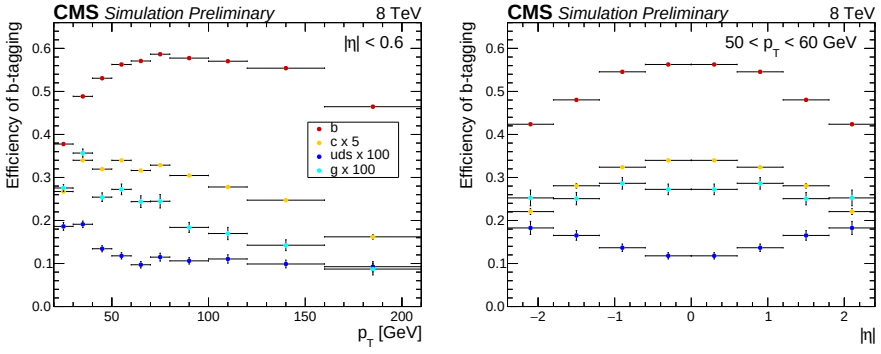


Figure 3.12: Efficiencies of b-tagging as determined from simulation.

3.6.5 Missing E_T

The final state of the signal process considered in this search includes a semileptonically decaying top quark, $t \to b\ell\nu$. Although the neutrino produced in this decay cannot be detected, its presence can still be inferred by implying momentum conservation in the plane transverse to the beam axis. The potential imbalance of measured momentum is described by the missing transverse en-

ergy $\not{E}_{\mathrm{T}}$. It is found as

$$\vec{\not{E}}_{\mathrm{T}}^{\,\mathrm{uncorr.}} = -\sum_i \vec{p}_{\mathrm{T}}(i), \tag{3.28}$$

where $\vec{p}_{\mathrm{T}}(i)$ is the vectorial transverse momentum of PF candidate i, and the sum runs over all PF candidates, including those that have been identified by the CHS procedure. In order to better approximate the true missing momentum, several additive corrections are applied to this quantity [100].

First correction profits from the improved description of the detector response obtained through the JES calibration. For instance, it allows to account for the non-linearity of the calorimeter response for jets. The correction is calculated as

$$\vec{C}_{\mathrm{JES}} = -\sum_j \left(\vec{p}_{\mathrm{T}}^{\,\mathrm{full}}(j) - \vec{p}_{\mathrm{T}}^{\,\mathrm{PU}}(j)\right), \tag{3.29}$$

where $\vec{p}_{\mathrm{T}}^{\,\mathrm{full}}(j)$ and $\vec{p}_{\mathrm{T}}^{\,\mathrm{PU}}(j)$ are vectorial transverse momenta of jet j with full and only pile-up corrections applied, respectively, and the sum runs over all jets with $p_{\mathrm{T}}^{\mathrm{full}} > 10\,\mathrm{GeV}/c$. The threshold is motivated by the domain of validity of the JES corrections. Essentially, this $\not{E}_{\mathrm{T}}$ correction removes from the sum (3.28) PF candidates clustered into jets and replaces them with the jet momenta, while keeping the contribution from pile-up.

Contribution to the sum (3.28) from pile-up candidates can be split into three components:

$$\vec{\not{E}}_{\mathrm{T}}^{\,\mathrm{PU}} = \vec{\not{E}}_{\mathrm{T}}^{\,\mathrm{CHS}} + \vec{\not{E}}_{\mathrm{T}}^{\,\mathrm{no\ PV}} + \vec{\not{E}}_{\mathrm{T}}^{\,\mathrm{neutral}}, \tag{3.30}$$

which represent, respectively, particles identified by the CHS procedure, remaining charged pile-up particles that are not associated to any primary vertex and thus cannot be recognized by the CHS, and neutral pile-up particles. The probability to produce neutrinos in an inelastic pp interaction is small, and therefore in ideal conditions $\not{E}_{\mathrm{T}}^{\mathrm{PU}} = 0$. However, an apparent momentum misbalance is caused by detector response to neutral particles, whose energies are measured with an inferior precision compared to charged ones. To profit from the precise measurement of the CHS component, the momentum balance is imposed in the overlapping pp collisions, and the remaining two components of the $\not{E}_{\mathrm{T}}$ are estimated as

$$\vec{\not{E}}_{\mathrm{T}}^{\,\mathrm{neutral}} + \vec{\not{E}}_{\mathrm{T}}^{\,\mathrm{no\ PV}} = -R \cdot \vec{\not{E}}_{\mathrm{T}}^{\,\mathrm{CHS}}, \tag{3.31}$$

where the proportionality factor R accounts for the detector response. It is derived from simulation and parameterized as a function of $\not{E}_{\mathrm{T}}^{\mathrm{CHS}}$. Then the full estimated pile-up contribution is removed from the total $\not{E}_{\mathrm{T}}$ with a correction

$$\vec{C}_{\mathrm{PU}} = -\left(1 - R\left(\not{E}_{\mathrm{T}}^{\mathrm{CHS}}\right)\right) \cdot \vec{\not{E}}_{\mathrm{T}}^{\,\mathrm{CHS}}. \tag{3.32}$$

Although a uniform distribution is expected for the azimuthal angle of $\vec{\not{E}}_{\mathrm{T}}$, some modulation has been observed in both recorded data and simulation. It is caused by imperfect detector alignment, inefficiencies in individual detector modules and cells, a residual ϕ dependence in calibrations, and a shift of the beam spot from the nominal position. The ϕ asymmetry is more pronounced for larger pile-up. It is corrected by a summand

$$\vec{C}_{\phi} = -\left(\vec{c}_0 + \vec{c}_v \cdot n_{\mathrm{PV}}\right), \tag{3.33}$$

where n_{PV} is the number of reconstructed primary vertices, which serves as an estimate of the amount of pile-up, and parameters $\vec{c}_0$ and $\vec{c}_v$ are determined from samples of $Z \to \mu^+\mu^-$ events. This correction is different between data and simulation.

Finally, the corrected missing E_{T} is calculated as

$$\vec{\not{E}}_{\mathrm{T}} = \vec{\not{E}}_{\mathrm{T}}^{\,\mathrm{uncorr.}} + \vec{C}_{\mathrm{JES}} + \vec{C}_{\mathrm{PU}} + \vec{C}_{\phi}. \tag{3.34}$$

In Ref. [100] $\not{E}_{\mathrm{T}}$ resolution is studied in $Z(\to \ell\ell)+$jets and $\gamma+$jets events. The measurements profit from an accurate determination of momenta of Z bosons and photons and the absence of genuine $\not{E}_{\mathrm{T}}$, balancing the vector boson against the hadronic recoil, which gives the dominant contribution to the $\not{E}_{\mathrm{T}}$ resolution. Effectively, $\vec{\not{E}}_{\mathrm{T}}$ is recalculated from PF candidates of the hadronic recoil and compared to the transverse momentum of the vector boson $\vec{q}_{\mathrm{T}}$. Resolution is calculated independently for parallel and perpendicular components of the hadronic recoil, $u_{\parallel}$ and $u_{\perp}$, which are defined in Fig. 3.13. Obtained results are presented in Fig. 3.14 as functions of the vector boson momentum q_{T}.

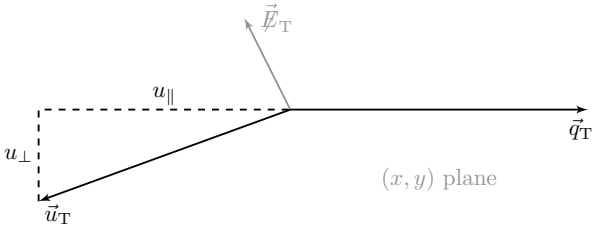


Figure 3.13: Transverse momentum balance in a $Z/\gamma+$jets event. Here $\vec{q}_{\mathrm{T}}$ and $\vec{u}_{\mathrm{T}}$ are momenta of the vector boson and the hadronic recoil.

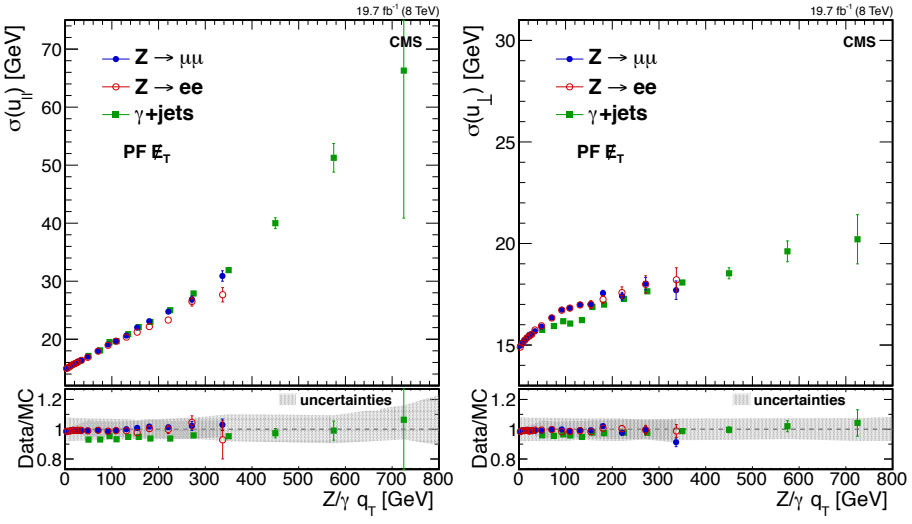


Figure 3.14: $\not\!\!E_{\mathrm{T}}$ resolution curves for parallel (left) and perpendicular (right) components of the hadronic recoil [100].

3.7 Reconstruction of neutrino

This search targets events that contain single semileptonically decaying top quarks. A special procedure is needed to reconstruct the top quarks since their decay products include undetectable neutrinos. Only events with single charged leptons are considered. The lepton is identified with the one from the $t \to b\ell\nu$ decay, and it is assumed that the accompanying neutrino provides the dominant contribution to the measured $\not\!\!E_{\mathrm{T}}$. Although the transverse momentum of the neutrino can be approximated as $\vec{p}_{\mathrm{T}}(\nu) \approx \vec{\not\!\!E}_{\mathrm{T}}$, the z component of its momentum is unknown. It is determined by imposing the W boson mass constraint

$$(E(\ell) + E(\nu))^2 - (\vec{p}(\ell) + \vec{p}(\nu))^2 = m_W^2, \tag{3.35}$$

following an approach described in Ref. [102]. This requirement leads to a quadratic equation for the z component $p_z(\nu)$, whose solutions are

$$p_z(\nu) = \frac{\Lambda \cdot p_z(\ell)}{p_{\mathrm{T}}^2(\ell)} \pm \sqrt{\frac{\Lambda^2 \cdot p_z^2(\ell)}{p_{\mathrm{T}}^4(\ell)} - \frac{E^2(\ell) \cdot p_{\mathrm{T}}^2(\nu) - \Lambda^2}{p_{\mathrm{T}}^2(\ell)}}, \tag{3.36}$$

where

$$\Lambda = \frac{m_W^2}{2} + \vec{p}_{\mathrm{T}}(\ell) \cdot \vec{p}_{\mathrm{T}}(\nu). \tag{3.37}$$

The solutions are first evaluated taking $\vec{p}_{\mathrm{T}}(\nu) = \vec{\not{E}}_{\mathrm{T}}$. If the radicand is positive, two real-valued solutions are present, and the one that is smaller in absolute value is chosen. This preference is motivated by a comparison to the true momentum of the neutrino from top quark decays performed in simulated events. However, in some cases the quadratic equation does not have real-valued solutions for $\vec{p}_{\mathrm{T}}(\nu) = \vec{\not{E}}_{\mathrm{T}}$. As can be shown, this is equivalent to the relation $m_{\mathrm{T}}(W) > m_W$, where the transverse W boson mass $m_{\mathrm{T}}(W)$ is defined as

$$m_{\mathrm{T}}^2(W) \equiv (p_{\mathrm{T}}(\ell) + p_{\mathrm{T}}(\nu))^2 - (\vec{p}_{\mathrm{T}}(\ell) + \vec{p}_{\mathrm{T}}(\nu))^2 = 2\, p_{\mathrm{T}}(\ell)\, p_{\mathrm{T}}(\nu) \cdot (1 - \cos\Delta\phi(\ell, \nu)). \quad (3.38)$$

If momenta of the charged lepton and the neutrino were known exactly and the W boson was on its mass shell, the condition $m_{\mathrm{T}}(W) \leqslant m_W$ would hold. The main reason it can be broken is the experimental resolution of $\not{E}_{\mathrm{T}}$. To find $p_z(\nu)$ under these circumstances, $\vec{p}_{\mathrm{T}}(\nu)$ is allowed to depart from $\vec{\not{E}}_{\mathrm{T}}$. It is modified to make the radicand in Eq. (3.36) zero while minimizing the Euclidean distance $|\vec{p}_{\mathrm{T}}(\nu) - \vec{\not{E}}_{\mathrm{T}}|$. The z component of the neutrino momentum is then calculated as $p_z(\nu) = \Lambda \cdot p_z(\ell)/p_{\mathrm{T}}^2(\ell)$, setting $\vec{p}_{\mathrm{T}}(\nu)$ in Eq. (3.37) to the found value.

For the purpose of reconstruction of the semileptonically decaying top quark, $\vec{p}_{\mathrm{T}}(\nu)$ is always set to the measured $\vec{\not{E}}_{\mathrm{T}}$, regardless of whether the solutions (3.36) are real-valued or not.

Chapter 4

Datasets and event selection

The final state of the signal process considered in this search includes a top quark and a Higgs boson. The search targets the $H \to b\bar{b}$ decay, which has the largest branching ratio of about 58% for a Higgs boson with a mass about $125\,\mathrm{GeV}/c^2$ [27]. In order to suppress the abundant multijet background, only semileptonic decays of the top quark are considered. Moreover, no attempt to reconstruct τ leptons is made, and only final states with a muon or an electron are selected. Fig. 4.1 shows an example Feynman diagram with all decays. The muon or electron from the top quark decay is complemented by neutrino, which creates a moderate $\not{E}_{\mathrm{T}}$. There are four b quarks, three of which stem from decays of the Higgs boson or the top quark. The forth b quark sometimes has a small transverse momentum or falls outside of the tracker coverage, and thus it is not always registered as a b-quark jet. The quark q' recoils against the W boson and has a prominent non-central distribution in pseudorapidity [22], which is a feature shared with the t-channel production of single top quarks [103]. Finally, additional jets can be produced in initial- and final-state radiation. This experimental signature of signal events defines the event selection and determines the list of most important background processes.

4.1 Simulation

Properties of signal and background processes are predicted with the help of Monte Carlo simulation (MC). This section discusses the simulation in general and provides specific details about the datasets exploited in this search.

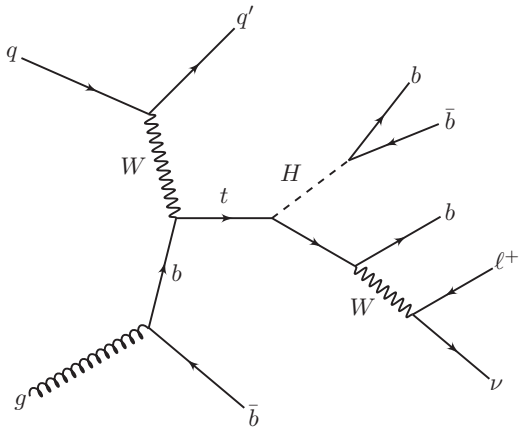


Figure 4.1: Example Feynman diagram for tHq production with all decays shown.

4.1.1 Event generation and detector simulation

First, an event stemming from the process of interest is generated. Starting from the pp collision, it is developed until stable and long-lived particles are produced, which can reach the detectors. A number of methods are exploited in the event generation, and a review can be found in Ref. [21, Chapter 40]. The event generation is followed by simulation of the detector response to the produced particles. The whole sequence is briefly summarized below.

The event generation is typically factorized into several steps, which are characterized by different energy scales. It starts by considering the collision of individual partons, or the hard interaction, at a scale $\mu \gg \Lambda_{\mathrm{QCD}}$, where $\Lambda_{\mathrm{QCD}} \sim 200\,\mathrm{MeV}$ is the characteristic QCD scale. By calculating Feynman diagrams, the collision can be described up to a fixed order in the perturbative expansion in the QCD or electroweak coupling constants. The matrix element (ME) created from amplitudes corresponding to the Feynman diagrams, fully describes distributions of particles in the final state for a given initial state. To produce the final distribution, the matrix element is convoluted with the proton parton distribution functions (PDF) [21, Chapter 19]. The resulting multidimensional distribution is then sampled with a Monte Carlo method to generate the desire number of parton-level events.

Partons in the generated hard-scattering events undergo repeated $1 \to 2$ branchings, producing parton showers (PS). The branching probabilities are given by the Dokshitzer-–Gribov—Lipatov—Altarelli-–Parisi (DGLAP) equations [104–

106], which are derived in the collinear approximation and thus do not describe large-angle emissions accurately. As a final-state shower evolves, the associated energy scale decreases until a cut-off value $\mu_0 \sim 1\,\text{GeV}$ is reached and the shower is stopped to avoid entering in the non-perturbative regime. On the other hand, initial-state showers start from a low scale and then reach the scale of the hard scattering. But since their outcome, which is the initial state of the matrix element, is known, it it more practical to model them in the decreasing order in the scale by reversing the DGLAP evolution equations.

At a scale $\mu \sim \Lambda_{\text{QCD}}$ the confinement takes place, and the coloured particles produced in the parton showers form colourless hadrons. Since the perturbation theory breaks down at low energies, it cannot be deployed to describe the hadronization from first principles. Instead, several empirical approaches exist, of which the most widely used is the Lund string model [107, 108]. The model assumes that the potential of the colour field produced by a pair of colour-connected quarks depends linearly on the distance between them. This is effectively described by a string with a tension $\kappa \sim 1\,\text{GeV/fm}$ that connects the two quarks. As the quarks move apart, the string stretches, and when it has deposited a sufficient energy, it can break into two strings producing a quark-antiquark pair. Each of the resulting strings can break in its turn. The produced pairs of quarks then hadronize into mesons. Also baryons can be formed if a pair of diquarks is created at a string break. In a more complex configuration when a pair of quarks is indirectly connected in colour via one or more intermediate gluons, the string starts from one quark and ends at the second one, while passing by the gluons. Each gluon is thus connected to two segments of the string, forming a transverse kink. To first approximation, each segment fragments in a way similar to the $q\bar{q}$ string.

At the last step of event generation, primary hadrons produced in the hadronization are decayed according to branching ratios measured in experiment. Angular distributions of their decay products are often described in a simplified way. The final list of particles includes only stable ones and those that live long enough to reach the detector, such as pions and kaons.

Detector response to the generated particles is simulated with the help of the Geant 4 program [109, 110]. The particles are propagated through the detector, using a detailed description of its geometry. Interactions with the detector material are simulated taking into account an exhaustive list of contributing processes. It includes but is not limited to ionization, multiple scattering, bremsstrahlung, photon conversion, absorption of hadrons, and inelastic hadron–nucleus scattering. Simulation of additional pp collisions is overlaid with primary signals to model the effect of pile-up. To take the out-of-time

pile-up into account, additional collisions are admixed not only in the nominal bunch crossing but also in several bunch crossings before and after the nominal one. Based on the signals in sensitive elements of the detector, electronic read-out is simulated, and this information is used to perform the event reconstruction, which was described in the previous chapter.

4.1.2 Signal process

Out of the three modes of the tH production introduced in Section 1.2.1, only the dominant t-channel process (see Fig. 1.2) is considered. MadGraph 5.1.5.11 [111] is exploited to generate signal events $pp \to tH\bar{b}j$, with subsequent decays $H \to b\bar{b}$ and $t \to b\ell\nu$, $\ell = e, \mu, \tau$. Here the $\bar{t}Hbj$ final state is also implied, and j denotes any quark except for b and t. At NLO there is an interference with the s-channel process. To remove it, Feynman diagrams with W-boson resonances are excluded from the definition of the signal. The four-flavour scheme is used, i. e. b quarks do not contribute to the proton structure functions. Compared to the five-flavour scheme, it is known to reproduce exclusive observables better [34, 112].

Masses are set as follows: $m_H = 125.0$, $m_t = 172.5$, $m_b = 4.8\,\mathrm{GeV}/c^2$. Model parameters are further modified by flipping the sign of the top quark Yukawa coupling. The factorization and renormalization scales are fixed to 100 GeV, or approximately $(m_H + m_t)/3$, as suggested by authors of Ref. [26] in a private communication.

Parton-level events produced by the MadGraph generator are showered and hadronized using Pythia 6.4.26 [113] with the $Z2^*$ tune [114]. Decays of τ leptons are handled by the Tauola 2.7 package [115], which improves description of angular distributions of decay products. The PDF set used is CTEQ6L1 [116].

The dataset is normalized to a cross section of 234^{+5}_{-0} fb [26], excluding branching ratios of $H \to b\bar{b}$ and $t \to b\ell\nu$ decays. It is calculated at the NLO precision in α_s. The uncertainty corresponds to a variation of the renormalization/factorization scale.

In addition to the target tHq production with $y_t = -1$, the corresponding SM process has been studied. Fig. 4.2 provides several example distributions obtained with samples of parton-level events. It includes transverse momenta and pseudorapidities of the top quark, the Higgs boson, and the recoil quark q' (figures (a) to (d)), the invariant mass of the tH system (e) and the cosine of the angle between the lepton from the decay $t \to b\ell\nu$ and the recoil quark,

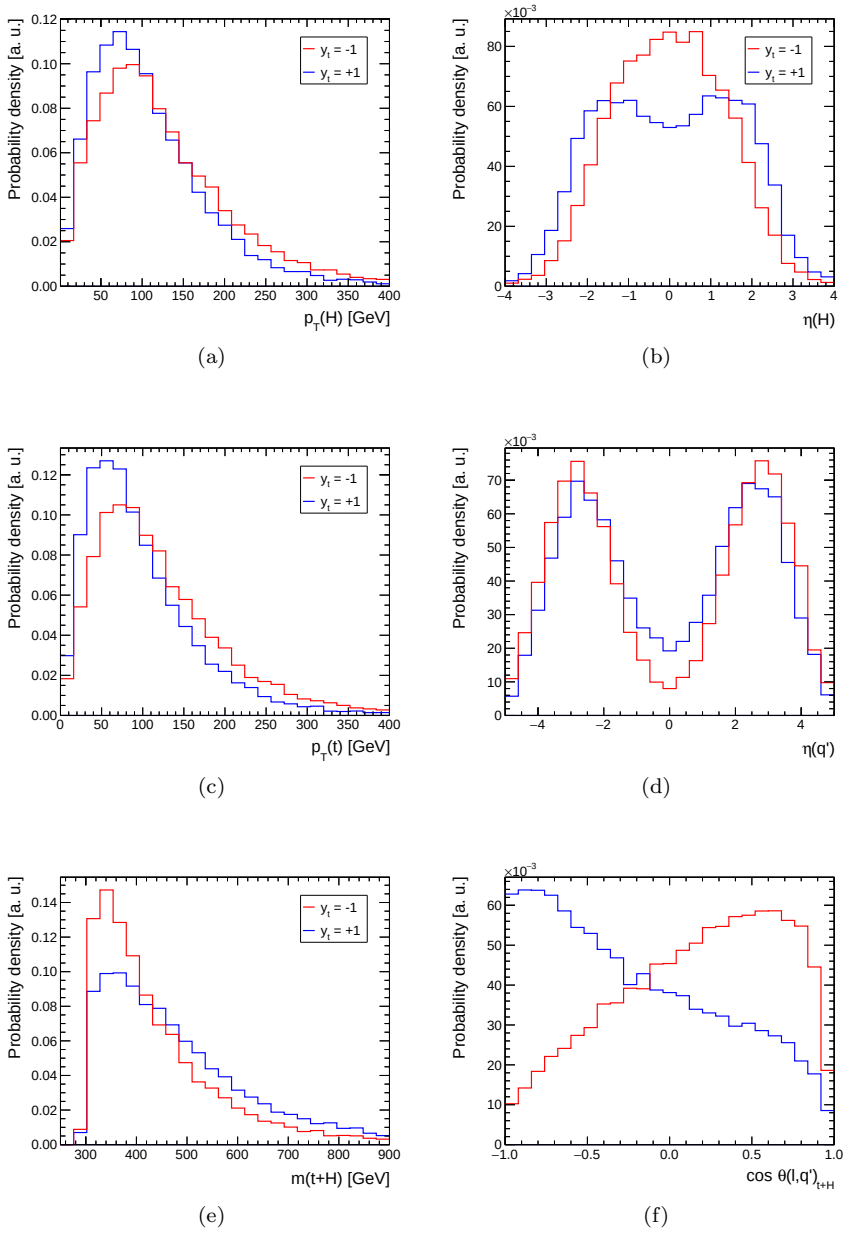


Figure 4.2: Example parton-level distributions for tHq production with $y_t = \pm 1$.

calculated in the rest frame of the tH system, (f). As can be seen, the distributions differ between the two cases, sometimes drastically. For this reason, it has been decided to focus the search on the $y_t = -1$ case exclusively and not to attempt to reinterpret it for other values of the Yukawa coupling.

4.1.3 Top quark pair production

As will be shown in Section 4.4, $t\bar{t}$ is by far the dominant background. A sample of $pp \to t\bar{t}$ events, with up to three additional quarks or gluons, is generated using MadGraph 5.1.4.8. The renormalization and factorization scales are adjusted dynamically, as will be discussed below. Parton showering and hadronization are modelled in the same way as for the signal process. The cross section is $246^{+6}_{-8}\,(\text{scale}) \pm 6\,(\text{PDF})\,\text{pb}$ [117], calculated at the NNLO+NNLL accuracy.

Inclusion of additional partons in the final state of the matrix element allows to describe large-angle emissions accurately. On the other hand, soft and small-angle emissions lead to divergences in the ME calculation and thus must be described through parton showers. In general, the choice between the two approaches is ambiguous. Same n-jet configuration can be produced from an appropriate n-parton ME final state through soft and collinear radiation or from a final state with $n-1$ partons where a hard emission creates an additional jet. This double counting is avoided with the help of the k_T-jet MLM prescription [118, 119], which also sets renormalization scales of QCD vertices in the ME and PS in a consistent manner.

The infrared and collinear divergences in the ME are excluded by requiring that the k_T distance [91] between each pair of partons in the final state, as well as p_T of each of them, is larger than a threshold $q_\text{cut}^\text{ME} = 20\,\text{GeV}$. Then the final state is clustered using the k_T algorithm in an attempt to reconstruct the effective history of emissions. Only clusterings that are compatible with Feynman diagrams included in the ME, are allowed, and this is done before the top quarks are decayed. The clustering proceeds until the central $2 \to 2$ scattering is recovered. Although typically the two top quarks are included in the two different final branches of the central scattering, they can also be clustered in the same branch if the $t\bar{t}$ system is boosted. Profiting from the recovered emission history, the renormalization scale μ_R in each QCD vertex is set to the k_T measure of that vertex. The factorization scale μ_F is chosen as the k_T measure of the central $2 \to 2$ scattering.

At the next step, the event is showered, setting μ_F as the starting scale of the shower. Created partons, excluding those that originate from the top quarks,

are clustered into jets using the k_{T} algorithm with a maximal allowed separation of $q_{\mathrm{cut}} = 40\,\mathrm{GeV}$. Then the n partons in the ME final state are iterated in the decreasing order in p_{T}. At each iteration the current parton is matched to the closest jet in the k_{T} metric, with a maximal allowed separation of q_{cut}. This jet is not considered for matching in the subsequent iterations. If some of the ME partons cannot be matched to a jet, the event is rejected and thus does not enter the final dataset. This can happen, for instance, if two ME partons are not separated enough and produce a common jet after the PS. Such event should be rejected because the same jet configuration can be reproduced from a final state with $n-1$ partons. In addition to the requirement that every ME parton is matched to a jet, no unmatched jets are allowed in events with $n < 3$. On the other hand, for $n = 3$, which is the highest multiplicity considered in the sample, additional jets are allowed as long as their k_{T} separation from the closest ME parton is smaller than the minimal separation between ME partons in the event. After this filtering, the simulated sample provides a consistent description of both hard and soft emissions.

Because of the multiple b quarks in the final state of the signal process, it is useful to classify $t\bar{t}$ events based on the number of additional b-quark jets. This is done using the same approach as the one applied in Ref. [120] in the $H \to b\bar{b}$ decay channel. The procedure operates with b and c quarks found in the event after the parton shower. Products of $t \to bW$ and $W \to cs$ decays are removed from the list of partons. Each parton is matched to the closest reconstructed in the ΔR metric, with a maximal distance of 0.5. A jet is then identified as a b- or c-quark jet if there is at least one quark of the given flavour matched to it, with the precedence given to b quarks. Finally, the event is classified as $t\bar{t}b\bar{b}$, $t\bar{t}b$, or $t\bar{t}c(\bar{c})$ if it contains at least two b-quark jets, exactly one b-quark jet, or at least one c-quark jet, respectively. These three classes are collectively referred to as the $t\bar{t}+\mathrm{HF}$ category. All remaining events are attributed to the $t\bar{t}+\mathrm{LF}$ category.

4.1.4 Minor backgrounds

In addition to the tHq and $t\bar{t}$ production, a number of backgrounds with smaller impacts are considered in this search. The list of processes, together with event generators used to model the hard scattering, is provided in Table 4.1, which includes also tHq and $t\bar{t}$ for the sake of completeness. Predominantly, LO generators are exploited, although subprocesses with additional quarks or gluons in the final state of the hard scattering are included in some cases. The only exception is single top quark production, which is modelled at the

Table 4.1: Processes simulated in this study. LO cross sections are calculated directly by the event generator. When the cross section is calculated by the CMS collaboration, only a reference to the tool used is provided.

Process	Generator	Cross section
tHq	MadGraph	234 fb NLO [26]
$t\bar{t}$	MadGraph	246 pb NNLO+NNLL [117]
$t\bar{t}H$	Pythia 6	130 fb NLO [27]
t, t-channel	POWHEG	87 pb approx. NNLO [125]
t, tW-channel	POWHEG	22.2 pb approx. NNLO [125]
t, s-channel	POWHEG	5.6 pb approx. NNLO [125]
$W \to \ell\nu$	MadGraph	37.5 nb NNLO [126]
$Z/\gamma^* \to \ell^+\ell^-$	MadGraph	3.5 nb NNLO [126]
WW	Pythia 6	55 pb NLO [127]
WZ	Pythia 6	13 pb LO
ZZ	Pythia 6	5.2 pb LO

NLO accuracy using the POWHEG framework [121–124]. Regardless of the generator used for the hard scattering, parton showers and hadronization are modelled with the help of Pythia.

In some auxiliary studies, the processes listed in Table 4.1 are accompanied by simulation of non-prompt production of muons or electrons. This component receives contributions from multijet events, where the leptons can be produced in decays of hadrons, and production of prompt photons, which can then undergo conversion in the detector material creating electrons. The two contributions are generated with Pythia and MadGraph respectively. In order to save computing resources, a filtering is applied to multijet events, rejecting those that are unlikely to contain well-identified muons or electrons after the event reconstruction.

4.2 Corrections to simulation

Generally, MC simulation needs to be tuned in order to describe recorded data precisely. In part, the required adjustments are incorporated into low-level calibrations and definitions of physics objects, an example of which is the JER smearing discussed in Section 3.6.3. This section describes several remaining corrections that are implemented as per-event weights.

The pile-up profile imposed in simulated samples does not agree exactly with the one observed in data. It is deduced from measurements of average instantaneous luminosity in each luminosity section [128], taking into account the (effective) total cross section of inelastic pp scattering. Simulated events are then weighted to reproduce the data distribution in the number of additional pp interactions.

There are some differences between data and simulation in efficiencies of lepton identification and trigger selection, which were presented in Section 3.6. The reported efficiencies were measured in samples of $Z \to \ell^+\ell^-$ events [93, 95] using the tag-and-probe method [129], and they are parameterized by lepton transverse momentum and pseudorapidity. To correct for the difference in efficiencies, each simulated event is assigned a weight equal to the ratio of the efficiencies in data and simulation, evaluated with p_T and η of the only tight lepton found in the event.

From measurements of the differential cross section of $t\bar{t}$ production [130, 131] it is known that the adopted approach to generate $t\bar{t}$ events leads to a harder top quark p_T spectrum than observed in data. This mismodelling is fixed by assigning each $t\bar{t}$ event an additional weight

$$w_{\mathrm{top}\,p_\mathrm{T}} = \sqrt{e^{a+b\cdot p_\mathrm{T}(t_1)} \cdot e^{a+b\cdot p_\mathrm{T}(t_2)}}, \tag{4.1}$$

where $p_\mathrm{T}(t_{1,2})$ are transverse momenta of the two top quarks, as read from the generator content of the event, and a and b are parameters obtained from the differential measurements.

Performance of the b-tagging algorithm is observed to be slightly inferior to what is expected from simulation [101]. This problem is addressed by the last correction, which exploits ratios of b-tagging efficiencies in data and simulation, or b-tagging scale factors, measured in Ref. [132] and shown in Fig. 4.3. For a given simulated event, the probability to reproduce the observed b-tagging configuration is

$$\mathcal{P}_\mathrm{MC} = \prod_{i\in\mathrm{tagged}} \epsilon_i \cdot \prod_{j\notin\mathrm{tagged}} (1-\epsilon_j), \tag{4.2}$$

where ϵ_i is the b-tagging efficiency for jet i, as measured in simulation, and the first (second) product is calculated over b-tagged (not b-tagged) jets. The probability to reproduce the same b-tagging configuration with efficiencies as in data reads as

$$\mathcal{P}_\mathrm{data} = \prod_{i\in\mathrm{tagged}} s_i\epsilon_i \cdot \prod_{j\notin\mathrm{tagged}} (1-s_j\epsilon_j), \tag{4.3}$$

where s_i is the scale factor for jet i, which depends on jet p_T, $|\eta|$, and flavour. The flavour of a jet is determined by PS partons that are found in a cone of

size 0.3 around the jet axis, giving precedence to heavier partons. The event is then assigned a weight

$$w_{b\,\mathrm{tag}} = \mathcal{P}_{\mathrm{data}}/\mathcal{P}_{\mathrm{MC}} = \prod_{i \in \mathrm{tagged}} s_i \cdot \prod_{j \notin \mathrm{tagged}} \frac{1 - s_j \epsilon_j}{1 - \epsilon_j}. \tag{4.4}$$

The b-tagging efficiencies that enter this formula are determined from simulation after the full event selection but b-tagging requirement is applied. They are parameterized by jet p_{T}, $|\eta|$, and flavour and derived independently for each process with a significant contribution.

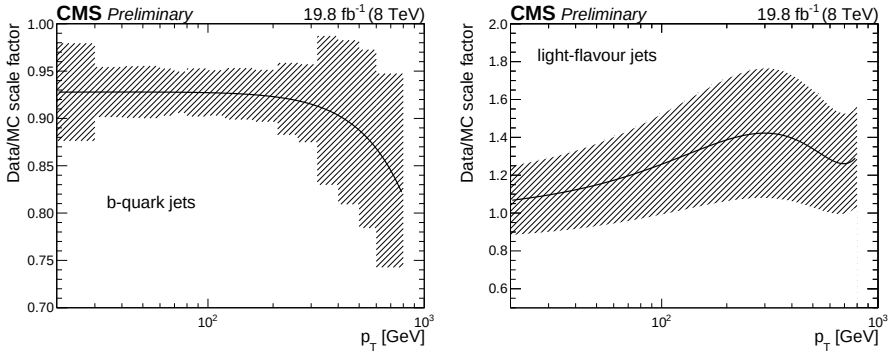
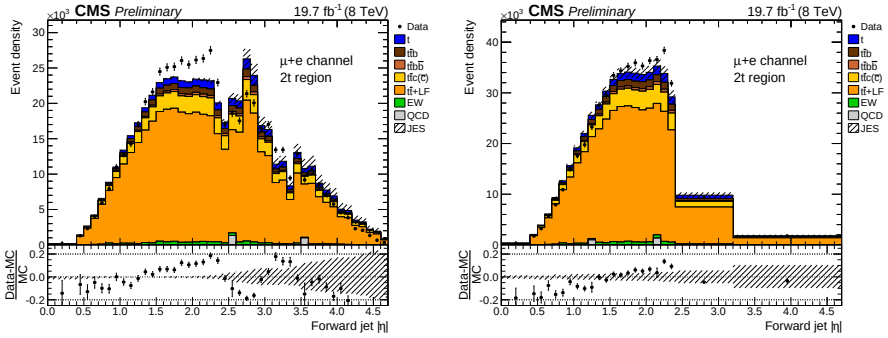


Figure 4.3: Difference in b-tagging efficiencies between data and simulation. The scale factors are derived in a single bin in jet pseudorapidity.

4.3 Mismodelling of jet pseudorapidity

An evident mismodelling of the jet distribution in pseudorapidity has been observed in this search. Fig. 4.4 (a) demonstrates a discrepancy in the $2t$ region, which is a control region dominated by $t\bar{t}$ events and will be defined in the next section. The pseudorapidity of the most forward jet in an event, shown in the figure, is an especially important observable in this search because the presence of a forward jet is a distinctive feature of the sought-for signal.

The main features of the discrepancy are a prominent depression in the data-to-simulation ratio at $|\eta| \sim 2.7$, a steep slope in the $|\eta| \gtrsim 3$ region, and a smaller slope in the central region. The same behaviour is observed also in two other control regions that are dominated by the Drell–Yan process and W boson production, and therefore the problem can be attributed to either



(a) (b)

Figure 4.4: Mismodelling of the pseudorapidity of the most forward jet in an event in the $2t$ region. In the left-hand plot a non-standard jet definition is used, loosening the p_{T} selection to $p_{\mathrm{T}} > 20\,\mathrm{GeV}/c$ throughout the pseudorapidity range. In the right-hand plot the variable jet p_{T} threshold introduced in Section 3.6.3 is used, and a wide binning is applied in the forward region. The hashed area displays the impact of JES systematic uncertainty. Simulation is normalized to observed event yield.

detector simulation or event reconstruction. As shown in Fig. 4.4 (a), the disagreement is not covered by systematic uncertainties in jet energy scale. A further investigation has revealed that it does not vanish when the jet pile-up ID [133] is applied or the CHS is disabled in the reconstruction.

The behaviour in the $|\eta| \gtrsim 3$ region can be understood as an effect of binning adopted for the residual jet energy corrections. They suffer from statistical uncertainties in the forward region, and because of this the corrections are derived in a single bin for jets with $|\eta| > 3.14$. As a result, the data-to-simulation ratio in this domain is corrected on average only, and there is no guarantee that details of the distribution within the region are reproduced accurately.

The depression at $|\eta| \sim 2.7$ is likely to be connected to a bias in jet pseudorapidity. As shown in Fig. 4.5, there is a systematic difference between the reconstructed jet η and the pseudorapidity of the corresponding particle-level jet. This results in a migration of jets with particle-level $|\eta| \sim 2.5$ or 3.1 to reconstructed $|\eta| \sim 2.7$, creating a local excess in the jet distribution. A qualitatively similar, albeit smaller, migration has been observed in a limited production of improved simulation that profits from a better description of

detector conditions and fixes a known problem in the simulation of the electromagnetic calorimeter [134]. Junging from a more accurate description of the data provided by this simulation, the η migration is assumed to be responsible for the depression.

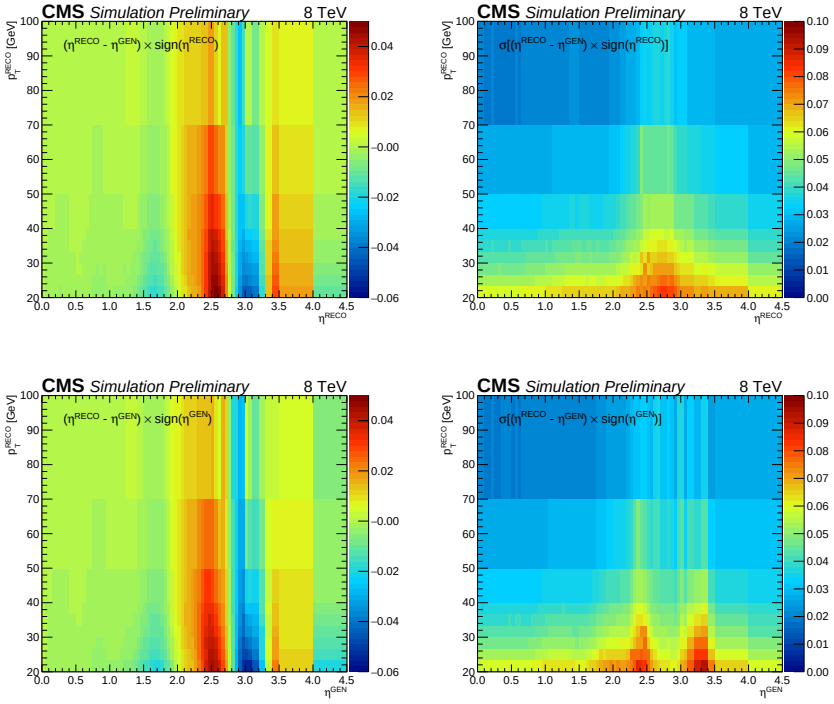


Figure 4.5: Difference between pseudorapidities of a reconstructed jet and the particle-level jet matched to it (left-hand plots) and the standard deviation of the difference (right-hand plots). In the upper row, the quantities are plotted as functions of the reconstructed jet p_T and η while in the lower row η of the matched particle-level jet is used instead. The matching is performed as in Ref. [133].

Although the jet η bias can be measured in simulation, studying it in data is a complicated task, and for this reason no attempt to correct for the bias has been made. However, Fig. 4.5 reveals that the jet migration is mostly restricted to the region $2.4 \lesssim |\eta| \lesssim 3.2$, altering details of the distribution in η inside the region but not affecting the overall normalization in the area. To make the analysis insensitive to the migration, the whole region is treated as a single bin in pseudorapidity, allowing the analysis to access the overall number of jets

in the region but masking mismodelled details of their distribution. Similarly, jets with $|\eta| > 3.2$ are covered by another single bin in order to reflect the parameterization of jet energy corrections. This also accounts for the smaller migration of jets with $|\eta| \sim 3.4$ towards larger $|\eta|$ values.

As will be discussed in detail later, the analysis exploits MVA techniques to separate the signal from backgrounds, and jet pseudorapidity as well as whole jet four-momenta are used to construct input variables. In order to emulate the wide bins, the pseudorapidity of forward jets is modified as follows:

$$\eta \mapsto \eta' = \begin{cases} \eta & \text{if } |\eta| < 2.4, \\ 2.8 \cdot \text{sign}(\eta) & \text{if } 2.4 < |\eta| < 3.2, \\ 3.5 \cdot \text{sign}(\eta) & \text{if } |\eta| > 3.2. \end{cases} \tag{4.5}$$

This transformation is applied to all jets in both data and simulation, before any observable is calculated, which makes the MVA insensitive to the mismodelling of the bias and the response in the most forward region. It alters the meaning of the jet pseudorapidity turning it into an effective quantity rather than a physical observable, but this does not jeopardize validity of the measurement since for an MVA-based analysis it is important to have all its inputs modelled properly whereas they might not have a strict physical meaning. The transformation could also be viewed as an emulation of a hypothetical detector, whose calorimeters have no granularity within $2.4 < |\eta| < 3.2$ and $|\eta| > 3.2$ regions.

The mismodelling and the jet η bias are more pronounced for jets with smaller transverse momenta. In order to exclude the most affected domain, only jets with $p_{\text{T}} > 40\,\text{GeV}/c$ are considered hereafter in the $|\eta| > 2.4$ region. As shown in Fig. 4.4 (b), the higher p_{T} threshold together with the wide binning in pseudorapidity successfully protect against the mismodelling in the forward region. In the central region this observable is modelled within a 10% precision for the bulk of events, which, however, is larger than the uncertainty of jet energy scale. As will be discussed in Section 6.2.4, this uncertainty can be doubled without any significant impact on results of the search.

4.4 Event selection

Semileptonic decay of the top quark in the targeted experimental signature is the most prominent feature that allows to discriminate signal events from the overwhelming multijet background, which is therefore exploited in the trigger selection. Depending on the flavour of the lepton, events are attributed to

either muon or electron channel. In the muon channel, an event is accepted by the trigger if it contains a muon with $p_{\mathrm{T}} > 24\,\mathrm{GeV}/c$ and $|\eta| < 2.1$, while in the electron channel the presence of an electron with $p_{\mathrm{T}} > 27\,\mathrm{GeV}/c$ and $|\eta| < 3$ is required. In both cases the lepton must additionally meet a set of identification requirements resembling those introduced in Section 3.6. The two triggers were not prescaled during the whole period of 2012 data taking and accumulated an integrated luminosity of $19.7 \pm 0.5\,\mathrm{fb}^{-1}$ each. Their decisions are also emulated for simulated events, which are required to pass the trigger selection.

The remaining event selection is described in terms of physics objects defined in Section 3.6. It starts by requiring that an event contains exactly one muon (electron) that meets the tight selection and additionally has $p_{\mathrm{T}} > 26\,\mathrm{GeV}/c$ ($30\,\mathrm{GeV}/c$). In order to suppress contribution from the Drell–Yan process and other backgrounds where multiple prompt leptons are produced, the event is rejected if it contains additional loose leptons of any flavour.

The final state of the signal process includes four b quarks (see Fig. 4.1), one of which might escape detection, as it was mentioned above. Therefore, the event selection requires the presence of either three or four b-tagged jets. The two possibilities are considered independently and eventually define two different selection bins, which are hereafter referred to as $3t$ and $4t$ regions. Occasionally, a version of the event selection with exactly two b-tagged jets is also used as a $t\bar{t}$ control region. In this $2t$ region events are additionally required to contain at least three jets with $|\eta| < 2.4$ (they include the two b-tagged jets), which is needed in order to calculate some of observables that will be discussed in the next chapter. Furthermore, in all the three regions an event must contain at least one jet that fails the b-tagging requirement, which allows to accommodate for the recoil jet and possible radiation. As can be seen from Fig. 4.6, it is important that jets with relatively small transverse momentum, down to $20\,\mathrm{GeV}/c$, are considered in the central region. Although both tHq and semileptonic $t\bar{t}$ events usually contain high-p_{T} jets, the low jet p_{T} thresholds are essential to detect all jet decay products of the Higgs boson or the top quarks.

The small contribution from multijet background that survives the selection up to this point, is removed efficiently with a requirement $\not{E}_{\mathrm{T}} > 35\,(45)\,\mathrm{GeV}$ in the muon (electron) channel. The tighter selection in the electron channel is motivated by the generally larger rate of non-prompt electrons. In addition, an event is required to contain at least four jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}/c$. This suppresses the multijet background even further and also reduces significantly a contribution from W boson production.

Finally, rare anomalous or problematic events are removed. They can demon-

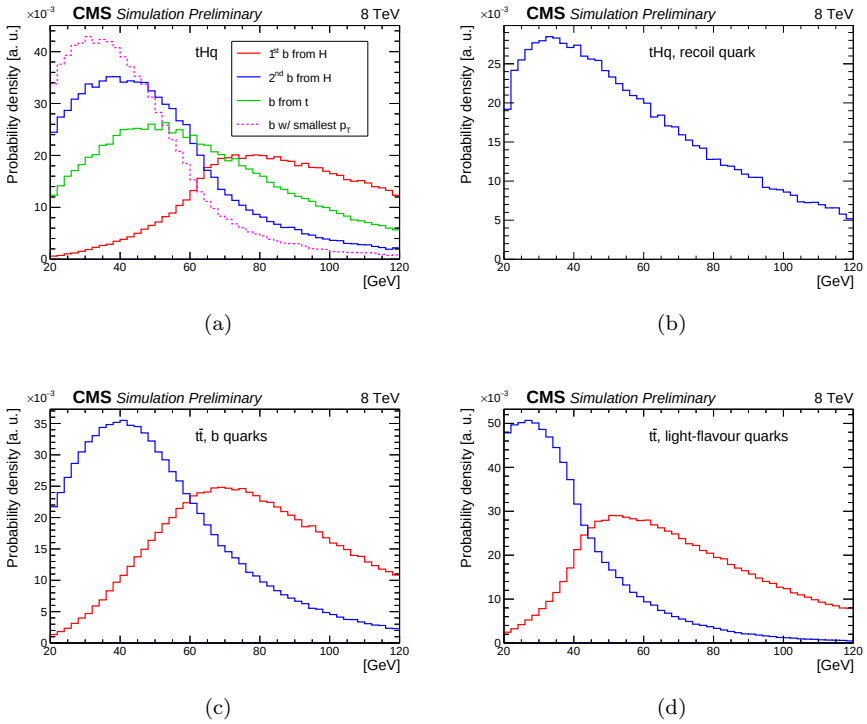


Figure 4.6: Distributions of parton transverse momenta in tHq (a, b) and semileptonic $t\bar{t}$ events (c, d) after the lepton step of the event selection: (a) b quarks from decays $H \to b\bar{b}$ and $t \to b\ell\nu$ and the softest b quark out of the three, (b) recoil quark q', (c) leading and subleading b quarks from decays of top quarks in $t\bar{t}$, (d) leading and subleading light-flavour quarks from the hadronic decay of a top quark.

strate a number of types of abnormal behaviour: no primary vertices that satisfy the quality requirements given in Section 3.2.2, anomalous noise in the HCAL, unnaturally large apparent energy depositions in few problematic cells of the ECAL, aborted track reconstruction, coherent noise in the strip tracker, a spectacular mismatch between the number of tracks and activity in the ECAL, a jet overlapping with a masked problematic region in the ECAL, a muon originating from beam halo, and other signatures.

Most important requirements of the event selection are summarized in Table 4.2. Event yields after this selection are reported in Tables 4.3, 4.4, and Figs. 4.7, 4.8 show distributions of several basic observables. In the tables

Table 4.2: Summary of the event selection to define signal regions.

Muon channel	Electron channel
Exactly one tight muon with $p_{\mathrm{T}} > 26\,\mathrm{GeV}/c$ and $\lvert\eta\rvert < 2.1$	Exactly one tight electron with $p_{\mathrm{T}} > 30\,\mathrm{GeV}/c$ and $\lvert\eta\rvert < 2.5$
No additional loose leptons of any flavour	
$\not{E}_{\mathrm{T}} > 35\,\mathrm{GeV}$	$\not{E}_{\mathrm{T}} > 45\,\mathrm{GeV}$
Three or four b-tagged jets with $p_{\mathrm{T}} > 20\,\mathrm{GeV}/c$	
At least one non-b-tagged jet with $p_{\mathrm{T}} > 20\,\mathrm{GeV}/c$ if $\lvert\eta\rvert < 2.4$ and $p_{\mathrm{T}} > 40\,\mathrm{GeV}/c$ otherwise	
At least four jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}/c$	

and figures, the "electroweak" category combines W, Z/γ^*, WW, WZ, and ZZ production. As can be seen, the selected regions are dominated by the $t\bar{t}$ background, with only a small contribution from the signal.

There is a pronounced $\sim 30\%$ difference between observed and expected yields in the signal regions, which is attributed to an underestimation of the contribution from the $t\bar{t}$+HF components. As described in Section 4.1.3, these processes are modelled at the leading order, as a part of the inclusive $t\bar{t}$ sample. In the context of the $t\bar{t}b\bar{b}$ production, it is known that cross sections of these processes can be very sensitive to the choice of the $\mu_{\mathrm{R,F}}$ scale and can receive significant NLO k-factors unless a special choice of the scale and kinematic selection is made [135]. Experimental studies [136–138] indicate that their contribution might be underestimated in the exploited simulation, reporting enhancement of some tens per cent (with large uncertainties). To account for this, additional systematic uncertainties are assigned to the normalization of $t\bar{t}+\mathrm{HF}$ components, as will be discussed in the dedicated section. Observed normalization factors will be reported in Section 6.3.1.

There is only a limited number of events in the $4t$ region in data, which is a consequence of the usage of the tight working point of the b-tagging algorithm. This can potentially impair the sensitivity of the search through the increased statistical uncertainties. A looser working point has been checked by performing the whole analysis, up to the calculation of the expected upper limit on the signal cross section, with the modified definition of a b-tagged jet. This resulted in a weaker limit, thereby justifying the choice of the tight working point.

Contribution from the multijet background in the signal regions, as estimated

Table 4.3: Observed and expected event yields after the selection in the $3t$ region. Only statistical uncertainty is shown for simulation.

Process	Muon channel	Electron channel
Data	1514	1028
$t\bar{t}$	1058±5	718±4
Single top	39±3	27±3
Electroweak	17^{+7}_{-5}	11±7
$t\bar{t}H$	12.87±0.17	9.35±0.15
Total background	1128±9	767±10
tHq, $y_t = -1$	7.54±0.03	5.15±0.02
S/B ratio	0.7%	0.7%

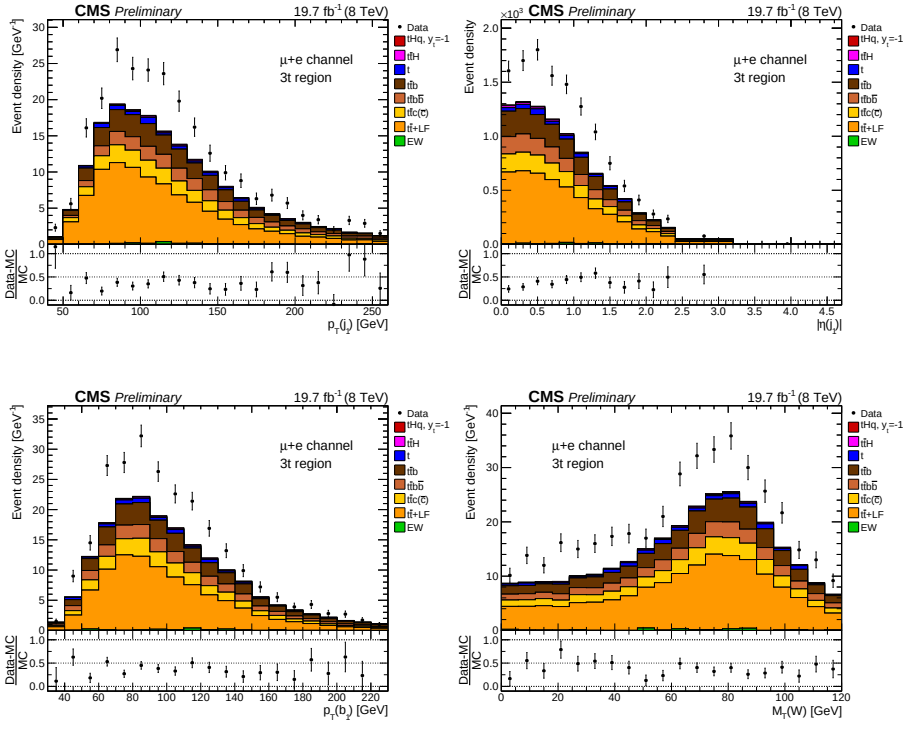


Figure 4.7: Distributions of example basic observables in the $3t$ region. Simulation is normalized to the integrated luminosity.

Table 4.4: Observed and expected event yields after the selection in the $4t$ region. Only statistical uncertainty is shown for simulation.

Process	Muon channel	Electron channel
Data	48	32
$t\bar{t}$	29.1±0.8	19.8±0.7
Single top	$1.1^{+0.8}_{-0.6}$	1.2±1.0
Electroweak	4^{+6}_{-4}	5^{+6}_{-4}
$t\bar{t}H$	1.72±0.06	1.43±0.05
Total background	37^{+6}_{-4}	29^{+7}_{-4}
tHq, $y_t = -1$	0.835±0.010	0.580±0.009
S/B ratio	2.3%	2.0%

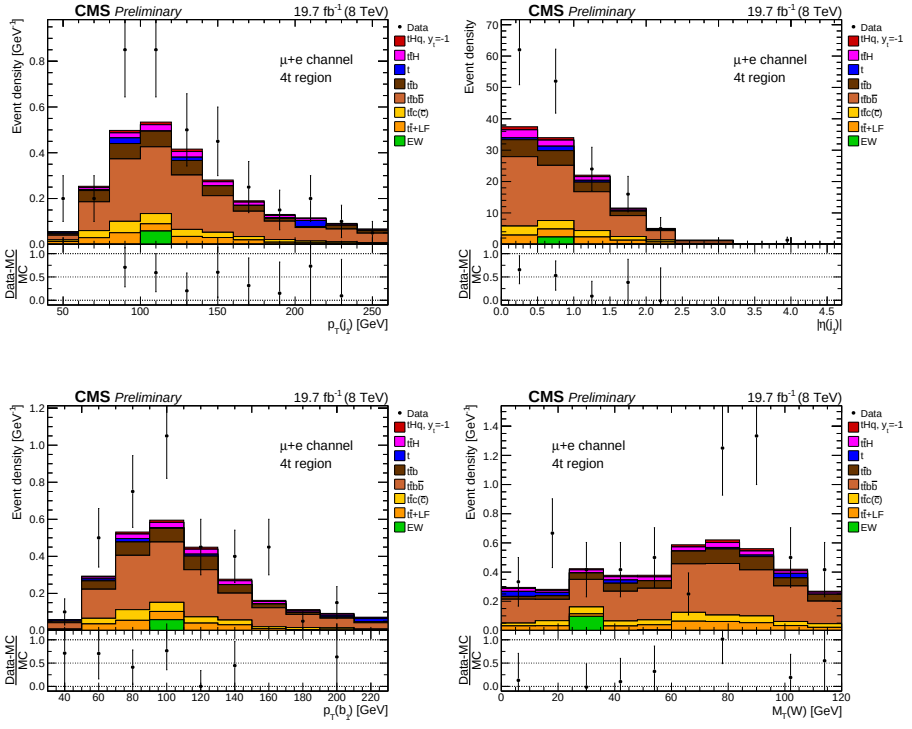


Figure 4.8: Distributions of example basic observables in the $4t$ region. Simulation is normalized to the integrated luminosity.

with simulation, is negligible. This result is additionally verified with a data-driven approach based on the so-called ABCD method (see, for instance, Ref. [139] for a discussion). The cross-check is performed independently for each lepton channel and b-tag multiplicity. For each signal region, here denoted A, three additional control regions B, C, D are defined, respectively, by requiring that the lepton fails the tight selection while passing the loose one, inverting the $\not\!\!E_{\mathrm{T}}$ selection, or the combination of both. Each of the three regions is relatively enriched in the multijet background, and the numbers of events $N_{B,\,C,\,D}$ stemming from it are found by subtracting from the observed yields expectations for all remaining processes. The number of multijet events in the signal region A is then estimated as

$$N_A = \frac{N_B \cdot N_C}{N_D}. \tag{4.6}$$

This estimation assumes factorization of the distribution of multijet events in the four regions, which is verified in simulation. The data-driven approach confirms that the contribution from this process is negligible, and it is ignored henceforth.

Another minor background that is neglected in this search is the tZq production. It occurs through a t-channel exchange of a W boson, similarly to the production of single top quarks in the t channel (see Fig. 1.1 (a)). The Z boson can be radiated from any of the four quark lines as well as from the W boson. With a subsequent $Z \to b\bar{b}$ decay, this process mimics closely the signal signature. It also has a similar inclusive cross section of about 240 fb (NLO) [140]. However, since the branching ratio of the $Z \to b\bar{b}$ decay is only about 15% [21], which is significantly smaller than $\mathcal{B}(H \to b\bar{b}) \approx 58\%$, the contribution from the tZq production to the event yields is expected to be factor 4 smaller than the one from the signal. This process can thus be safely ignored, compared to other backgrounds.

Chapter 5

Multivariate analysis

The small signal-to-background ratio after the event selection makes it necessary to apply methods of multivariate analysis (MVA) in order to additionally discriminate between the tHq production with $y_t = -1$ and its backgrounds. A number of approaches are used to address similar problems in the field of high energy physics [141]. In this search, artificial neural networks (NN) have been adopted. The ultimate result of the MVA discrimination is that each event is assigned a real-valued number that quantifies its compatibility with the signal hypothesis as opposed to the background one. This information is directly exploited in the subsequent statistical inference, without defining any additional event selection.

The two signal regions, $3t$ and $4t$, differ in background composition as well as event distributions in kinematic and b-tagging observables. This suggests that construction of a dedicated neural network for each region might be required in order to deliver the maximal discrimination power. It has been found, however, that the improvement with respect to using a single neural network for both regions is negligible. For this reason, the two regions are considered together throughout this chapter.

5.1 Artificial neural networks

Artificial neural networks are one of the oldest methods of non-linear multivariate analysis. Over their seventy-year history, a great diversity of approaches has

been developed to address various problems, ranging from function approximation to visualization of multidimensional data. In the following neural networks are briefly discussed in the context of a classification problem, while a comprehensive overview can be found in Ref. [142]. An implementation provided by the TMVA package [143] has been used in this search.

The central element of neural networks is an artificial neuron. Historically proposed to model biological neural systems [144], a neuron calculates a *response* y to a given vector of *inputs* $\boldsymbol{x}$:

$$y = h\left(\sum_{i=1}^{n} w_i x_i + w_0\right). \tag{5.1}$$

Here *weights* $\boldsymbol{w}$ are parameters of the neuron, and h is the *activation function*. Typically, a smoothed step function is used for activation, and in this search the hyperbolic tangent has been chosen.

A network is formed by connecting several neurons together, such that responses of some of the neurons are passed as inputs to others. The connection can be performed in different ways. The most commonly used architecture, which is also utilized in this search, is the *multilayer perceptron* with a single hidden layer. The network is organized into three layers, and outputs of neurons in a given layer are connected to inputs of all neuron in the next layer, as shown in Fig. 5.1. The *input layer* (leftmost in the figure) serves as the entry point of the network and trivially propagates its inputs to neurons of the *hidden layer*. Its neurons, in their turn, calculate responses following Eq. (5.1) and feed them as inputs to neurons of the *output layer* (rightmost in the figure), whose responses represent decision of the whole network. In this search neural networks with a single neuron in the output layer are considered.

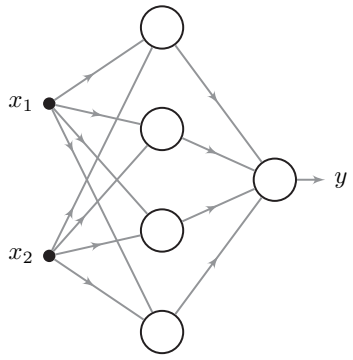


Figure 5.1: Multilayer perceptron with a single hidden layer.

A neural network of such architecture can be viewed as a mapping $\mathbb{R}^n \mapsto \mathbb{R}$, with the weights being the parameters of the mapping. The number of neurons in the hidden layer controls its potential complexity. According to the universal approximation theorem [145–147], a multilayer perceptron with a single hidden layer can approximate any continuous function on a compact subset of $\mathbb{R}^n$, provided that a sufficient number of hidden neurons is used. Of a particular interest in high energy physics are classification problems, where one attempts to attribute an event to the signal or background process based on a set of observables $\boldsymbol{x}$. The problem is typically addressed by constructing a neural network that, for the given inputs $\boldsymbol{x}$, yields a response close to 1 for signal-like events and 0 for background-like ones.

Neural networks are constructed by the means of *supervised learning*, utilizing a set of example events with known classification decisions (the *training set*). Technically, the construction, or the *training*, is performed by minimizing a *loss function*. A typical choice of the loss function, which is also adopted in this search, is the mean squared error

$$E(\boldsymbol{w}) = \frac{1}{N} \sum_{i=1}^{N} \big(y(\boldsymbol{x}_i; \boldsymbol{w}) - \hat{y}_i\big)^2, \tag{5.2}$$

where $y(\boldsymbol{x}_i; \boldsymbol{w})$ is the NN response for event i and $\hat{y}_i$ is the desired classification decision, here chosen to be 1 (0) for signal (background) events. Another commonly used loss function is the cross-entropy $1/N \sum_i (-\hat{y}_i \ln y_i - (1 - \hat{y}_i) \ln(1 - y_i))$, which can be obtained by maximization of the likelihood of correct classification of the training set. In both cases the response of a trained neural network approximates the purity $n_s(\boldsymbol{x})/(n_s(\boldsymbol{x}) + n_b(\boldsymbol{x}))$, where $n_{s,b}(\boldsymbol{x})$ are concentrations of signal and background events at point $\boldsymbol{x}$.

Typical dimensionality of the $\boldsymbol{w}$ space is $\mathcal{O}(10^2 - 10^3)$, and thus minimization of the loss function is a non-trivial problem. Historically, it was first successfully solved by the *(error) backpropagation* algorithm [148]. This is an iterative minimization algorithm based on the gradient descent. At each iteration, called *epoch*, the gradient $\boldsymbol{\nabla} E$ is calculated analytically, profiting from the knowledge of the neuron function (5.1). Then the weights are updated as

$$\boldsymbol{w} \to \boldsymbol{w}' = \boldsymbol{w} - \eta \boldsymbol{\nabla} E, \tag{5.3}$$

where the *learning rate* $\eta > 0$ is a parameter of the algorithm, which is usually adjusted during the training. In this search the gradient descent is replaced by a more advanced Broyden—Fletcher-–Goldfarb—Shanno (BFGS) algorithm. This is a quasi-Newton method, which makes use of a numerical approximation of the Hessian matrix to provide a faster convergence.

The training continues until a minimum of $E(\boldsymbol{w})$ is reached (so-called *exhaustive training*). However, since a neural network has a large number of free parameters, it can potentially describe noise in the training data rather then the desired underlying relationships. This would result in a poor generalization ability and, therefore, a larger value of the loss function when calculated on new data that have not been included in the training. This problem of *overfitting* is common to all methods based on the supervised learning. In this search it is addressed in a robust way, by splitting available example events into two sets. Only the first one, the training set, is utilized to construct the neural network, while the second *testing set* is used to measure its performance. As a consequence of the law of large numbers, an estimation obtained with the testing set converges to the true value even if the overfitting has occurred. Distributions of the NN response calculated from the testing set are exploited in the subsequent statistical analysis.

Although neural networks are capable of approximating very complex functions, in practice a preparatory transformation of input variables is often helpful to achieve best performance. In this search observables whose distributions have long tails, such as transverse momenta and reconstructed masses, are replaced by their logarithms. In addition, all input variables are linearly transformed to lie within the range $[-1, 1]$. After this transformation all input variables have similar typical scales.

5.2 Jet combinatorics

In order to maximize the performance of the NN discrimination between the tHq production and backgrounds, an effort should be put into construction of input variables that provide a good separation between the two classes. Because of the multijet final state considered in this search and the involved combinatorics, it is not trivial to define observables that exploit efficiently the presence of resonances such as Higgs bosons and top quarks. To address this problem, origin of reconstructed jets is hypothesized in an attempt to identify whether they stem from decay of a given resonance. Various kinematic and other properties (e. g. b-tagging information) are exploited to deduce the likely origin of each jet. Knowing it allows to reconstruct the resonances and define dedicated input variables, such as the transverse momentum of the Higgs boson, for example. The procedure is described in detail later in this section.

The signal region is dominated by the $t\bar{t}$ background, largely with $t\bar{t} \to \ell +$ jets decays. It is thus advantageous to exploit also observables that utilize

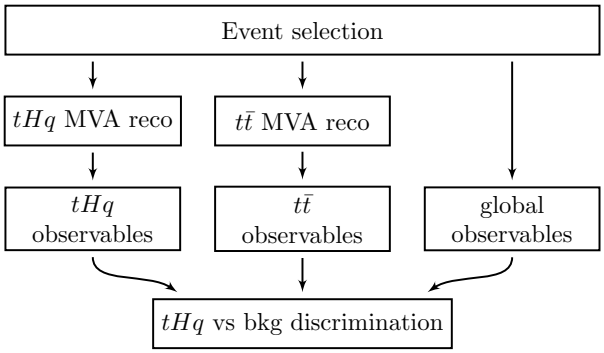


Figure 5.2: Groups of input variables for the neural network for event classification.

information on resonances in this process, an example being the mass of the hadronically decaying top quark. Similarly to what is done in case of the signal process, a dedicated procedure is applied to find likely origins of jets in a semileptonic $t\bar{t}$ event and in this way reconstruct the two top quarks.

In every event that passes the selection summarized in Table 4.2, likely origins of jets are deduced under the tHq and semileptonic $t\bar{t}$ hypotheses in parallel, as shown in the scheme in Fig. 5.2. In each case a set of observables intrinsic to the assumed hypothesis is constructed. These observables, which will be discussed in the next section, are used as input variables for the neural network to discriminate between signal and background events. The procedures for the jet assignment are described in detail below.

5.2.1 Determination of likely jet origins

The jet assignment under the both hypotheses, tHq and semileptonic $t\bar{t}$, is performed following the same general scheme. It is discussed in detail for the case of the signal hypothesis, and the $t\bar{t}$ hypothesis is briefly addressed later, highlighting the differences.

The goal of the procedure is to identify reconstructed jets that correspond to the four quarks in the final state $tHq \to 3bq\ell\nu$. A signal event passing the selection contains 5.4 reconstructed jets on the average. The procedure considers all possible ways to choose four jets and assign them to the four quarks in the final state. One particular way represents a potential *interpretation* of the event. In a simulated signal event it is possible to determine whether a given interpretation is correct or wrong. If all jets in the interpretation are matched

to appropriate quarks within a cone of size $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.3$, the interpretation is considered correct. Otherwise, if the angular distance ΔR from at least one quark to the jet assigned to it is larger than the threshold 0.3, the interpretation is wrong. By construction, one reconstructed jet can be assigned to one quark at maximum. Because not all jets are reconstructed, jets stemming from different quarks can merge, final-state radiation can alter direction of a jet, and for other reasons, not every event has a correct interpretation.

To simplify the involved jet combinatorics, a few simple constraints are imposed during the construction of potential event interpretations. The jets from decays of the top quark and the Higgs boson are required to be central ($|\eta| < 2.4$). In addition, only jets that are not b-tagged can be assigned to the recoil quark. These requirements are met by $\sim 99\%$ of correct interpretations, while they limit the number of wrong interpretations. Still, a signal event has about 60 interpretations on the average.

Each interpretation can be described with a set of observables sensitive to the choice of the four jets. Several such observables are combined in a dedicated neural network, which is constructed to distinguish between correct and wrong interpretations. Every tHq event used for training gives one entry to each of the signal and background categories: the correct interpretation and one randomly chosen wrong interpretation. If an event does not have a correct interpretation, it is excluded from training.

A number of observables have been considered as potential input variables for the neural network. The candidates, which included basic kinematic properties of the reconstructed Higgs boson and top quark, various angular correlations, b-tagging information, and other details, were ranked by their separation power. The separation power, interpreted here as the divergence between distributions of correct and wrong interpretations, can be quantified in various ways [149]. Several definitions were utilized, ranging from the unsophisticated area of the intersection of the two probability density functions to the Jensen–Shannon divergence, which is related to the Shannon entropy. Finally, 12 input variables have been chosen based on the ranking and physics value. The first group of observables is related to the reconstructed Higgs boson: mass, ΔR distance between the two jets stemming from its decay, number of b-tagged jets among the two, transverse momentum of the jet with smallest p_{T}, pseudorapidity of the most forward jet from the decay. The second group is constructed from identified decay products of the top quark: invariant mass of the b-quark jet and the lepton, number of b-tagged jets (zero or one), ΔR distance between the b-quark jet and the reconstructed W boson, product of electric charges of

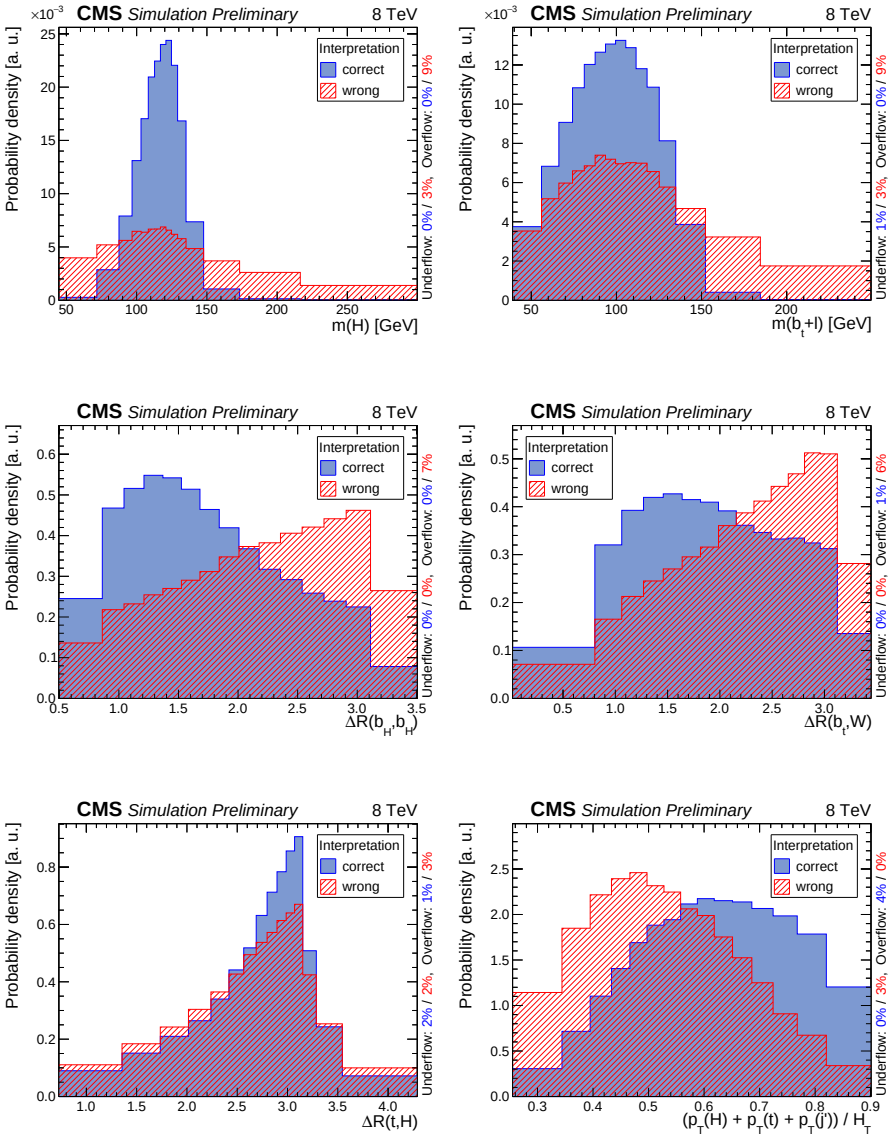


Figure 5.3: Distributions of correct and wrong event interpretations in input variables for the jet assignment under the tHq hypothesis. Part 1.

the jet and the lepton. The jet charge [150] is defined as

$$Q = \frac{1}{p_{\mathrm{T}}} \sum_i Q(i)\, p_{\mathrm{T}}(i), \tag{5.4}$$

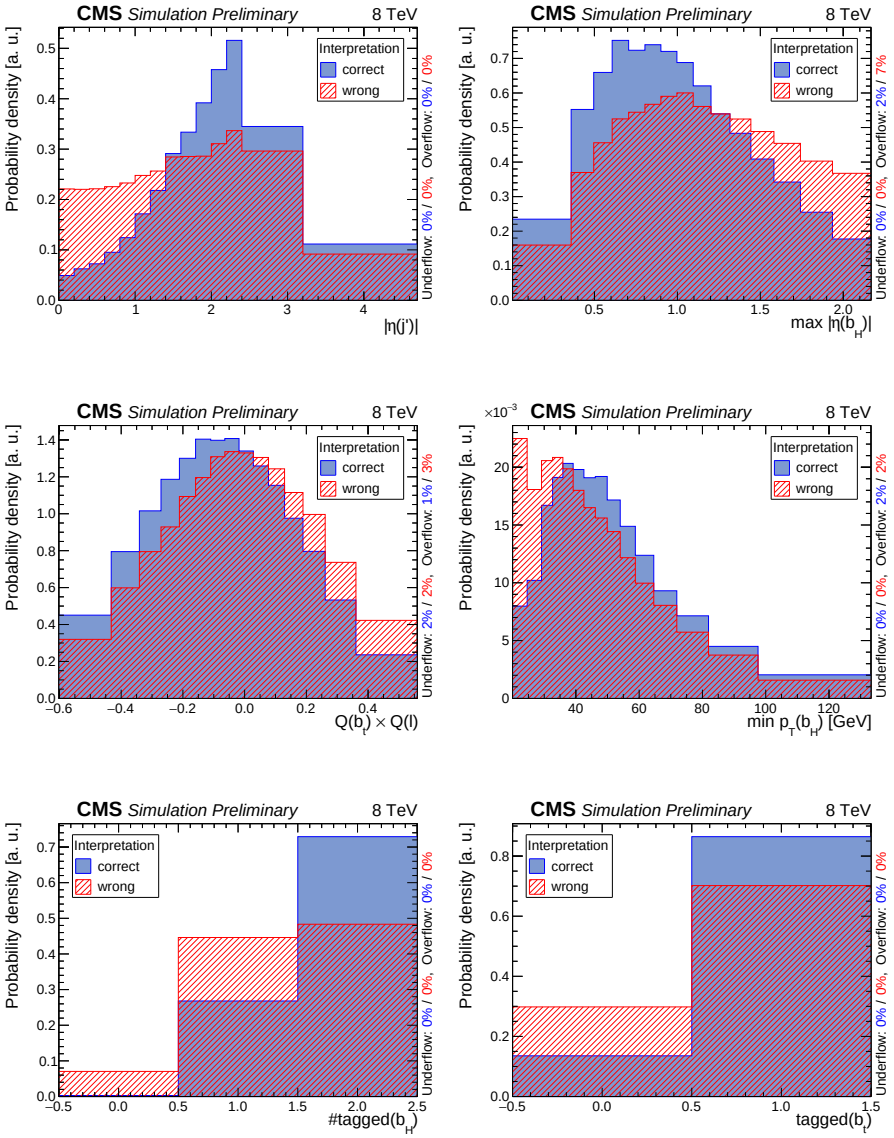


Figure 5.4: Distributions of correct and wrong event interpretations in input variables for the jet assignment under the tHq hypothesis. Part 2.

where $Q(i)$ and $p_{\mathrm{T}}(i)$ are the electric charge and transverse momentum of the i^{th} PF candidate clustered into the jet, and p_{T} is the jet transverse momentum. Correlations between the Higgs boson and the top quark are also considered.

They are represented by the ΔR distance between the two reconstructed objects and the relative H_{T}, defined as $(p_{\mathrm{T}}(t) + p_{\mathrm{T}}(H) + p_{\mathrm{T}}(j'))/H_{\mathrm{T}}$, where j' is the recoil jet and H_{T} is the sum of $\not{E}_{\mathrm{T}}$ and transverse momenta of the lepton and all reconstructed jets. The last observable used is the pseudorapidity of the recoil jet. Distributions of correct and wrong interpretations in these input variables are shown in Figs. 5.3, 5.4.

The neural network is constructed as described in the previous section. In order to cope with the dimensionality of the problem, 30 neurons are included in the hidden layer. Distributions of the NN response and the corresponding receiver operating characteristic (ROC) are provided in Fig. 5.5. A good performance is evident, with no signs of overfitting.

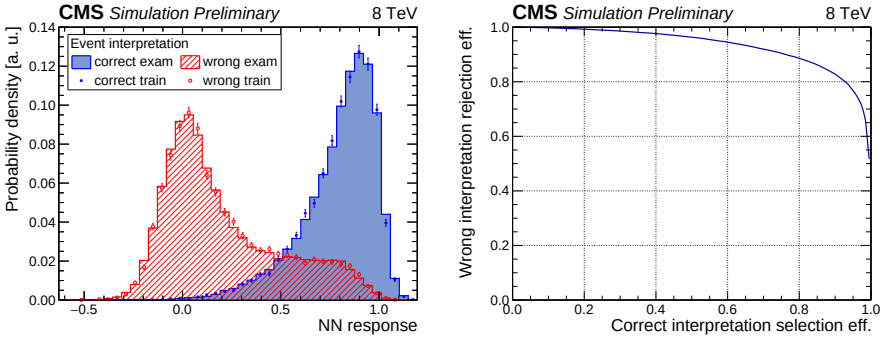


Figure 5.5: Distributions of correct and wrong event interpretations for the jet assignment neural network under the tHq hypothesis (left) and the corresponding ROC curve (right).

The procedure of the jet assignment under the semileptonic $t\bar{t}$ hypothesis aims at identifying reconstructed jets that correspond to the four quarks in the final state $t\bar{t} \to 2b2q\ell\nu$. Correct and wrong interpretations are defined similar to what is done for the signal process. The number of wrong interpretations is limited by a requirement that only central b-tagged jets can be assigned to the two b quarks. With an average of 5.7 reconstruced jets in a semileptonic $t\bar{t}$ event passing the selection, this results in a mean number of about 35 interpretations per event.

Correct and wrong interpretations are distinguished by an additional neural network. Its input variables are selected following the same procedure as described above. In total, 13 observables are used. Most of them refer to the hadronically decaying top quark: transverse momentum and pseudorapidity,

mass of the reconstructed W boson, difference between the masses of the top quark and the W boson, ΔR distance between the two light-flavour jets, ΔR

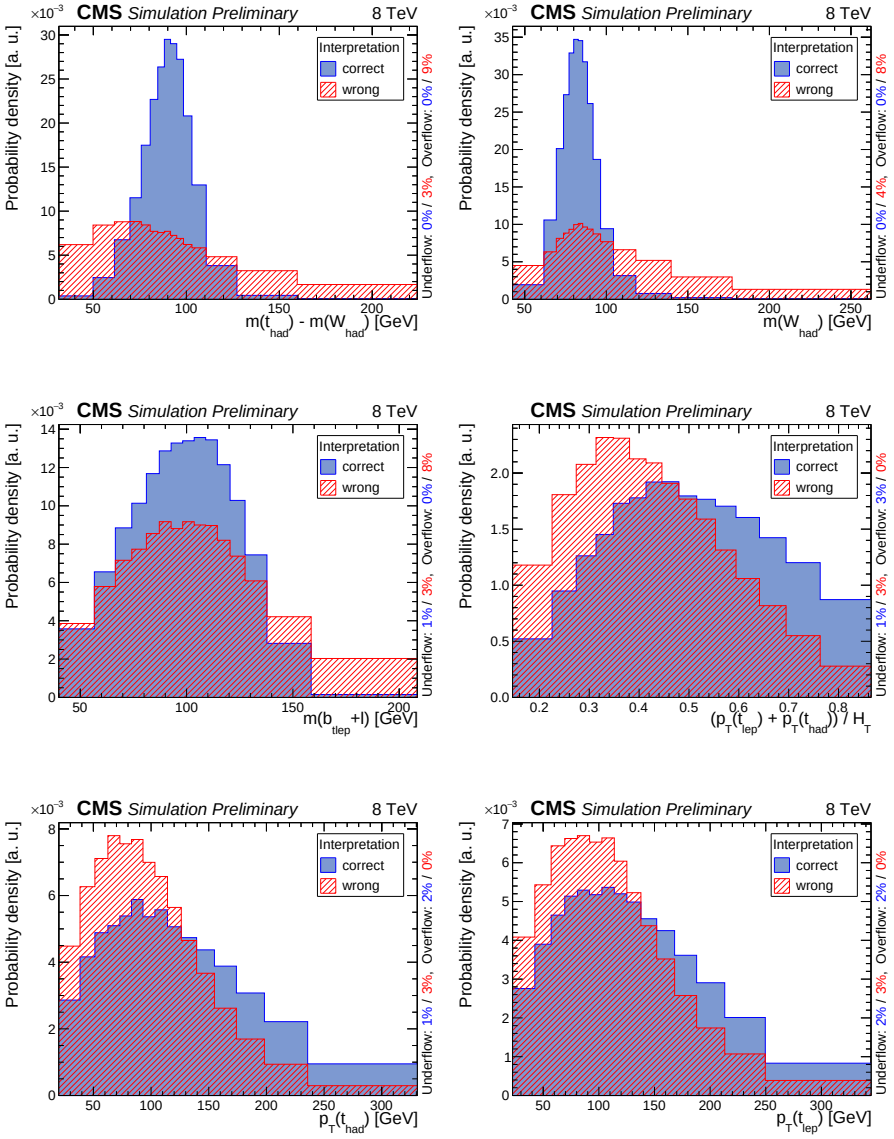


Figure 5.6: Distributions of correct and wrong event interpretations in input variables for jet assignment under the semileptonic $t\bar{t}$ hypothesis. Part 1.

distance between the b-quark jet and the W boson, number of b-tagged jets among the two light-flavour ones, sum of electric charges of the two light-flavour jets, multiplied by the charge of the lepton. Other observables are related to the leptonically decaying top quark: transverse momentum, invariant mass of the b-quark jet and the lepton, ΔR distance between the jet and the reconstructed W boson. Finally, there are observables that combine properties of both resonances: relative H_{T}, defined as $(p_{\mathrm{T}}(t_{\mathrm{had}}) + p_{\mathrm{T}}(t_{\mathrm{lep}}))/H_{\mathrm{T}}$, and difference of electric charges of the two b-quark jets, multiplied by the charge of the lepton. Distributions of correct and wrong interpretations in these variables are shown in Figs. 5.6 – 5.8.

The neural network constructed to identify correct interpretations of semileptonic $t\bar{t}$ events also contains 30 neurons in the hidden layer. Distributions of its response and the corresponding ROC curve are provided in Fig. 5.9.

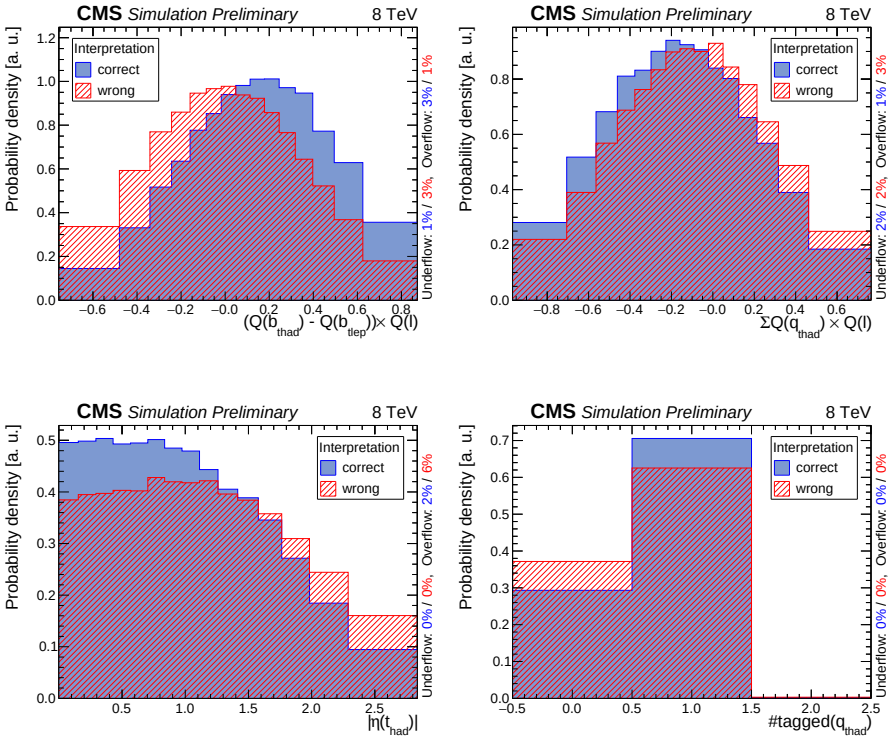


Figure 5.7: Distributions of correct and wrong event interpretations in input variables for jet assignment under the semileptonic $t\bar{t}$ hypothesis. Part 2.

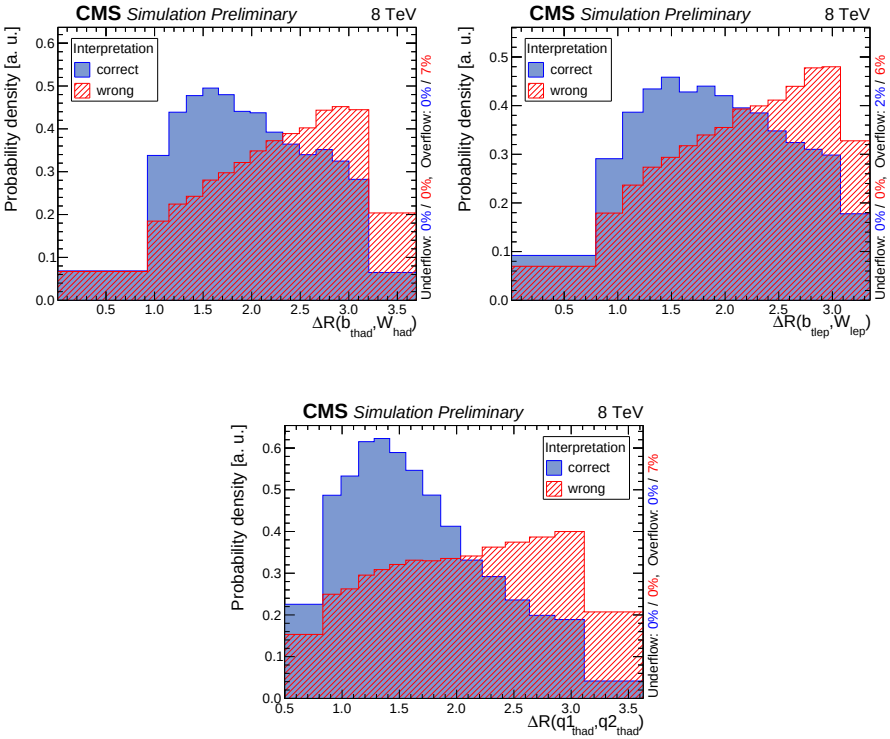


Figure 5.8: Distributions of correct and wrong event interpretations in input variables for jet assignment under the semileptonic $t\bar{t}$ hypothesis. Part 3.

The constructed neural networks are exploited to perform the jet assignment in events that pass the selection. For the hypothesis in question, all possible interpretations that meet the requirements on b-tagging and centrality, are constructed. The NN response is calculated for each interpretation, and the one with the largest response is accepted. The chosen interpretation unambiguously identifies decay products of the resonances.

It should be emphasized that the developed procedures do not assess the overall level of compatibility of an event with the assumed hypothesis. Instead, they find the most likely interpretation of the event even if the compatibility is low.

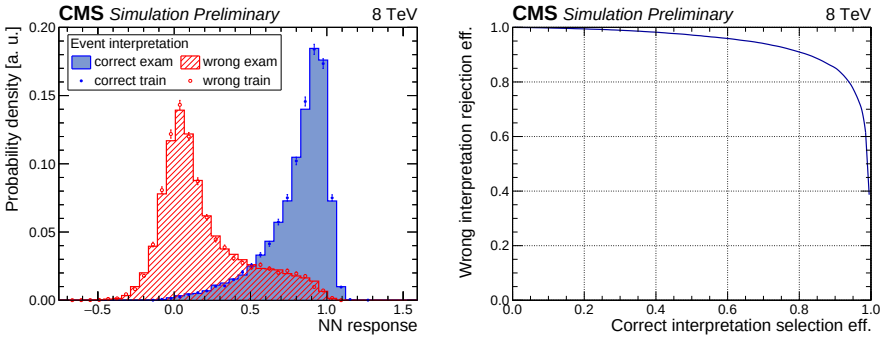


Figure 5.9: Distributions of correct and wrong event interpretations for the jet assignment neural network under the $t\bar{t}$ hypothesis (left) and the corresponding ROC curve (right).

5.2.2 Efficiency of the jet assignment

The ROC curves shown in Figs. 5.5, 5.9 illustrate the performance of the discrimination between correct and wrong event interpretations, which, however, is not related directly to the efficiency of the jet assignment. The latter is quantified by the fraction of cases in which the predicted jet origin agrees with the true one. The true origin of a reconstructed jet is defined in the following way: all quarks considered in the hypothesis in question are checked, and the one with the smallest ΔR distance from the jet is chosen. However, if this distance is larger than 0.3, the jet is said not to originate from any of the considered quarks.

The resulting efficiencies are reported in Tables 5.1, 5.2. Three versions are provided, which differ in what events are utilized to calculate the efficiency: all events passing the selection (type I in the tables); only events in which the quark in question has a matching reconstructed jet (type II); only events that have correct interpretations (type III). Definitions of types I and III can also be used for groups of quarks, such as decay products of the Higgs boson.

In the ideal case when origins of all reconstructed jets are determined correctly, the type III efficiencies are 100%. At the same time, type I efficiencies can be significantly smaller because a quark does not necessarily gives birth to a single reconstructed jet not overlapping with jets from other quarks. In particular, the low type I efficiency in case of the $t\bar{t}$ hypothesis is a consequence of a soft spectrum of the subleading light-flavour jet (see Fig. 4.6 (d)), whose transverse

momentum is often smaller than the jet p_{T} threshold. It should also be noted that in reality the 100% type III efficiency cannot be reached since, due to the quantum nature of the problem, it is impossible to determine with certainty, for instance, whether a b quark stems from the decay of the Higgs boson or the top quark.

Table 5.1: Efficiency of the jet assignment under the tHq hypothesis.

Object(s)	Efficiency, %		
	I	II	III
b from $t \to b\ell\nu$	57	61	66
at least one b from $H \to b\bar{b}$	86	—	92
both b from $H \to b\bar{b}$	51	—	65
recoil quark	52	78	79
all four quarks	22	—	44

Table 5.2: Efficiency of the jet assignment under the semileptonic $t\bar{t}$ hypothesis.

Object(s)	Efficiency, %		
	I	II	III
b from $t \to b\ell\nu$	64	70	66
b from $t \to bq\bar{q}'$	58	64	68
at least one q from $W \to q\bar{q}'$	63	—	90
both q from $W \to q\bar{q}'$	11	—	57
all quarks from $t \to bq\bar{q}'$	9	—	47
all four quarks	6	—	37

As can be seen from the Tables 5.1, 5.2, the probability to identify correctly all jets in the final state is larger than the product of probabilities to reconstruct each individual object (the top quark, the Higgs boson, and the recoil quark in case of the tHq hypothesis and the two top quarks for semileptonic $t\bar{t}$). The underlying correlation is a result of a competition for jets. If, for instance, a jet has been assigned to the top quark, it cannot be utilized to reconstruct the Higgs boson. A drawback of this procedure, here referred to as the global jet assignment, is that an incorrect reconstruction of one object might compromise the others. In order to investigate this aspect, an alternative procedure for jet assignment has been developed, in which each individual object is reconstructed independently, with the help of a dedicated neural network. The same jet can potentially be associated with more than one object. In Fig. 5.10 the two approaches are compared in terms of the ROC curve of the neural network for event classification, which will be discussed in the next section. As can

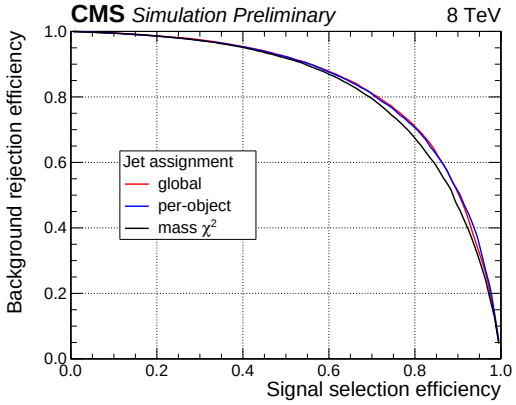


Figure 5.10: Performance of the event classification neural network with different approaches for the jet assignment.

be seen, they yield similar performance. The global jet assignment has been chosen because of its simpler implementation.

For a reference, the jet assignment has also been performed by choosing event interpretations that minimizes the difference between masses of the reconstructed resonances and their expectations. In case of the tHq hypothesis the loss function is defined as

$$\chi^2_{\text{mass}} = \frac{(m_t - \hat{m}_t)^2}{\hat{\sigma}_t^2} + \frac{(m_H - \hat{m}_H)^2}{\hat{\sigma}_H^2}, \tag{5.5}$$

where $m_{t,H}$ are masses of the top quark and the Higgs boson as reconstructed in the current event interpretation, and $\hat{m}_{t,H}$ and $\hat{\sigma}_{t,H}$ are the expected masses and their standard deviations, as derived using correct interpretations. The recoil jet is chosen as the most forward non-b-tagged jet. The jet assignment under the semileptonic $t\bar{t}$ hypothesis is performed in a similar way, building the loss function from the masses of the two top quarks. Fig. 5.10 reveals that this simple approach results in a smaller discrimination power compared to the MVA jet assignment.

5.3 Event classification

Every event in the $3t$ and $4t$ regions is probed for compatibility with the signal hypothesis as opposed to the background one. This is done with the help of an

additional neural network, which exploits three groups of input variables, as illustrated in Fig. 5.2. Observables in the first two groups are constructed by utilizing the two jet assignment procedures, which are applied in parallel and help to highlight peculiar features of the tHq and semileptonic $t\bar{t}$ processes. The third group contains observables that do not rely on the jet assignment.

Input variables for the event classification neural network have been chosen by refining a tentative list of about 60 observables, which reflected basic kinematic properties of reconstructed resonances and their decay products, angular correlations between the objects, b-tagging decisions, electric charges, and global event characteristics. In particular, all input variables exploited for the jet assignment were included in the list. The tentative list was reduced to about 20 observables following the same approach of ranking based on the separation power as described in Section 5.2.1. It was then optimized further while keeping the performance of the event classification neural network approximately constant. The optimization was performed in a recursive manner, removing one observable at a time. At each step, all possible candidate sets of $n-1$ input variables were constructed by discarding one of the observables, and the neural network was retrained with every new candidate set. The set that resulted in the smallest degradation of the performance of the neural network, quantified by the area under the ROC curve and other measures, was accepted. The procedure was repeated until the degradation became significant. The resulting set of input variables is optimal in the sense that removing any of them would cause a visible deterioration of the performance.

The final list contains 8 input variables. The largest group consists of observables that rely on the jet assignment under the tHq hypothesis: transverse momentum of the reconstructed Higgs boson, number of b-tagged jets among its decay products, pseudorapidity and transverse momentum of the recoil jet. Among observables defined under the semileptonic $t\bar{t}$ hypothesis, only those related to the hadronically decaying top quark have been selected: mass, ΔR distance between the light-flavour jets stemming from its decay, and the number of b-tagged jets among them. The only selected observable that does not exploit the jet assignment is the electric charge of the lepton. Distributions of signal and background events in the input variables are shown in Figs. 5.11, 5.12. It is worth noting that the mass of the reconstructed Higgs boson, also shown in the Fig. 5.11 for the sake of completeness, does not provide a sufficient separation and for this reason has not been exploited in the neural network. The explanation is that the multijet final state of $t\bar{t}$ events that dominate the signal region often allows finding a pair of jets that are compatible with properties of the decay products of the Higgs boson. On the other hand, the mass of the hadronically decaying top quark is a strong discriminator because it

is constructed from a triplet of jets and as such is more robust against the combinatorics.

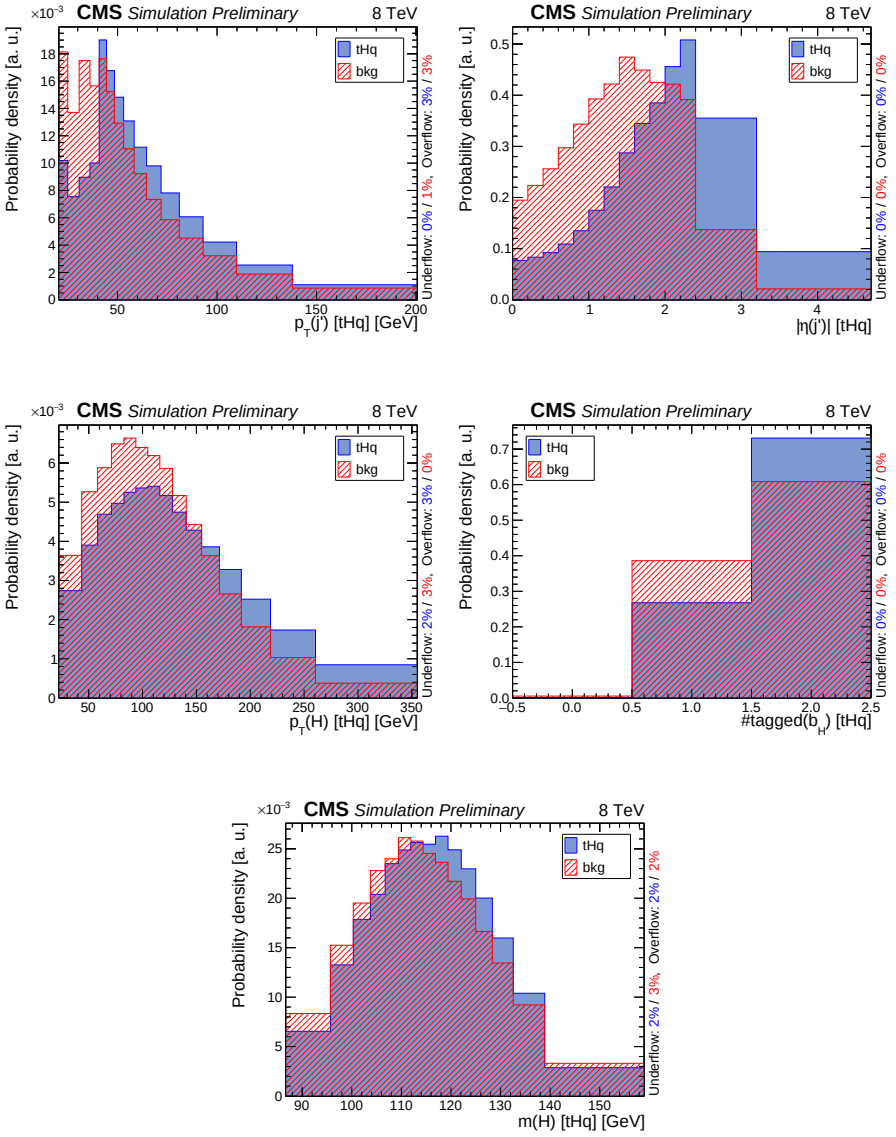


Figure 5.11: Distributions of signal and background events in the input variables for the event classification NN. Part 1. Although the Higgs boson mass is not used as an input, it is shown in the figure for completeness.

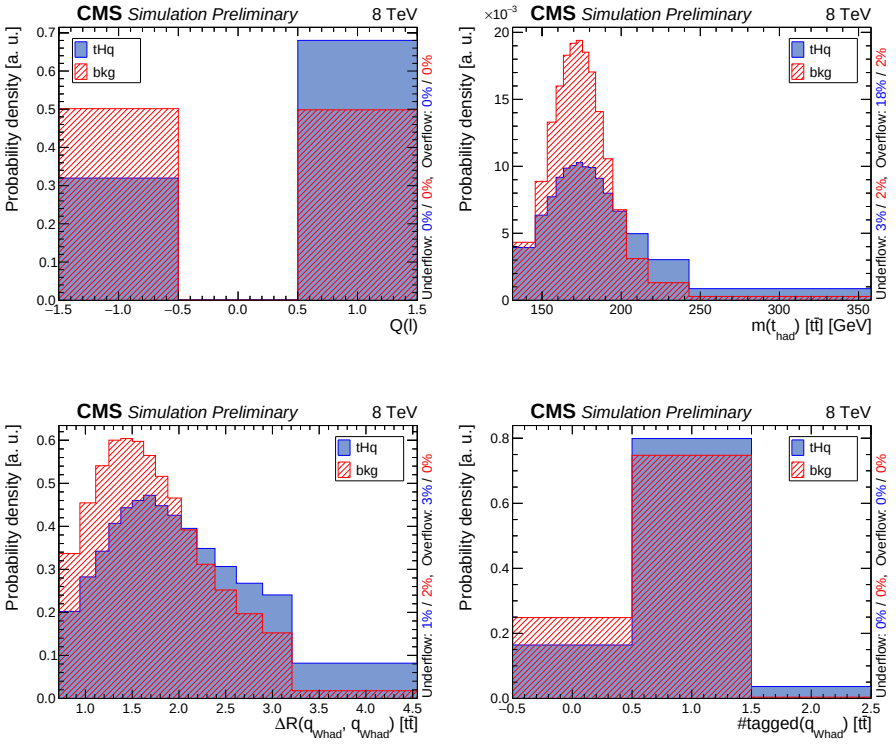


Figure 5.12: Distributions of signal and background events in the input variables for the event classification NN. Part 2.

It has been verified that all input variables are described by simulation at a satisfactory level. The corresponding distributions in all four signal regions are shown in Figs. 5.13, 5.14. In all these figures MC expectations are rescaled to the observed event yield in order to facilitate comparison of the shapes, while the normalization is subject to significant systematic uncertainties, which will be discussed in the dedicated section. In fact, one of the discriminative observables chosen initially, the sphericity, has been discarded because of a pronounced mismodelling.

The event classification neural network has been trained with about $4 \cdot 10^4$ events in the signal and background categories each. The background category is populated by a mixture of semileptonic $t\bar{t}$, dileptonic $t\bar{t}$, and $t\bar{t}H$ events, and the three considered processes are normalized to the corresponding cross sections. On the other hand, the signal process is rescaled in such a way

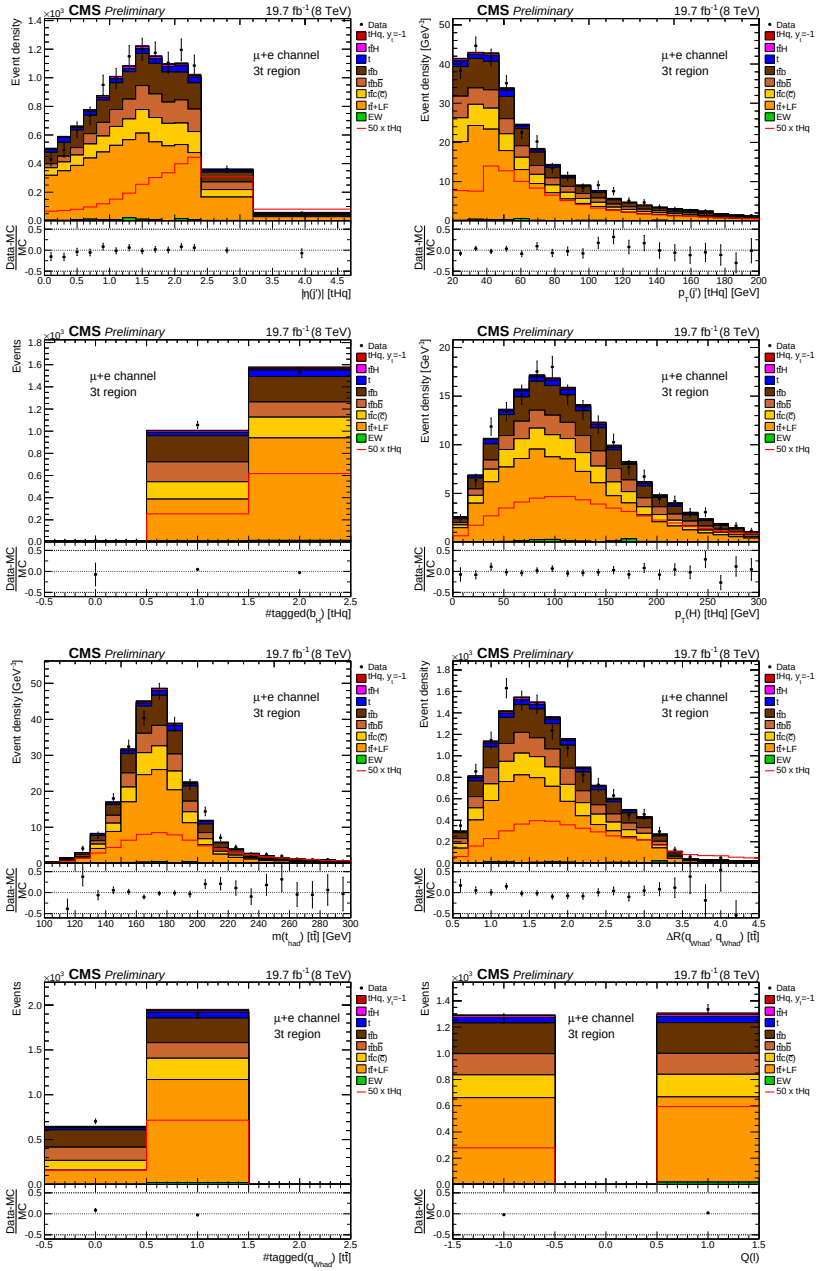


Figure 5.13: Distributions of input variables for the event classification NN in the 3t region, sum of muon and electron channels. Simulation is normalized to the observed event yield to facilitate comparison of shapes.

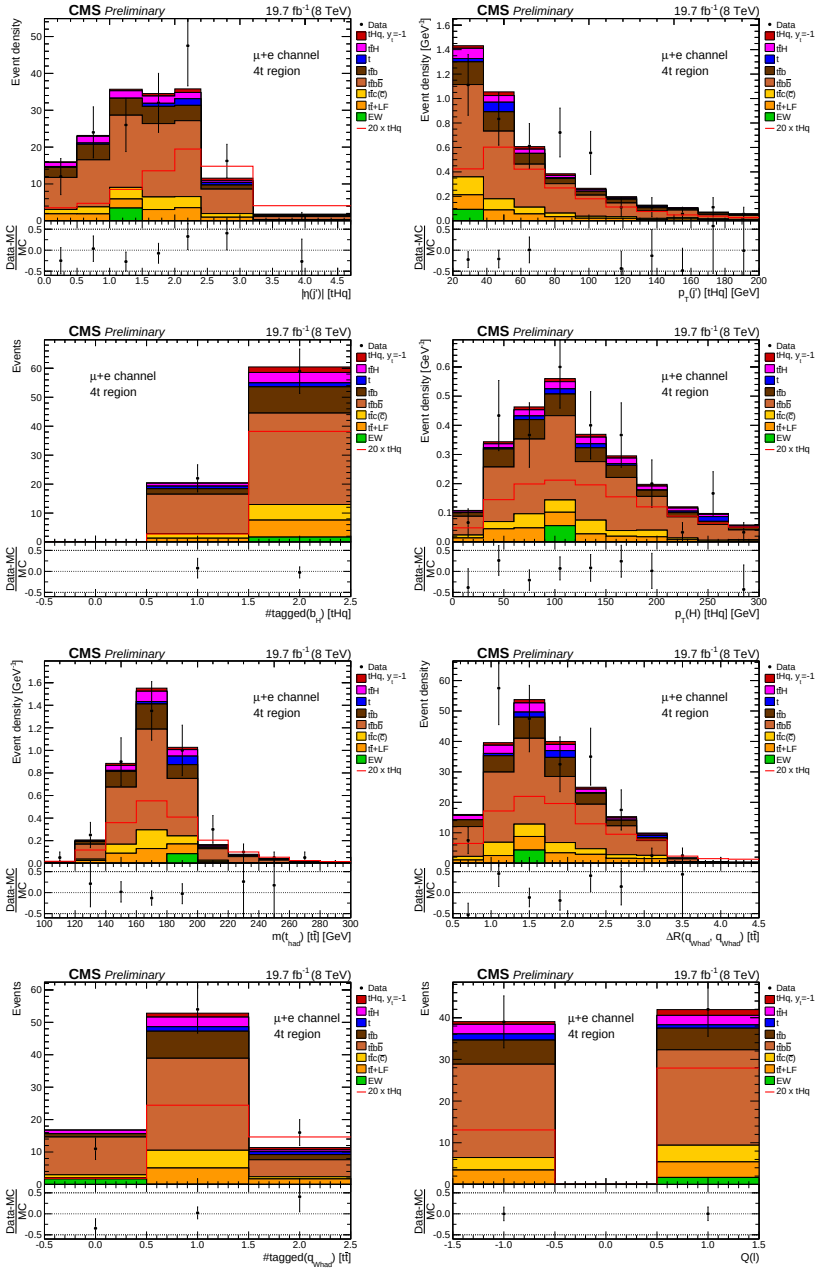


Figure 5.14: Distributions of input variables for the event classification NN in the 4*t* region, sum of muon and electron channels. Simulation is normalized to the observed event yield to facilitate comparison of shapes.

that it matches the total background normalization since this transformation is known to improve the classification performance in the case of a small signal-to-background ratio. Sets of tHq and semileptonic $t\bar{t}$ events utilized in the training overlap with those exploited to construct the jet assignment neural networks. It would also be useful to include t- and tW-channel single top processes, but the available number of simulated events passing the selection ($\mathcal{O}(100)$) makes it impossible. Other minor backgrounds have not been considered for training because of their small impact; in addition, this preserves more simulated events to be used in the statistical inference.

The neural network contains 30 neurons in the hidden layer, which has been found to give the best performance after comparing several configurations with the number of hidden neurons ranging from 5 to 50. Distributions of signal and background events in the response of the constructed neural network, as well as its ROC curve are shown in Fig. 5.15. No signs of overfitting are visible. However, it should be stressed that even if overfitting occurred, it would not compromise validity of results of the search because the subsequent statistical analysis is based solely on events that have not been utilized in the training.

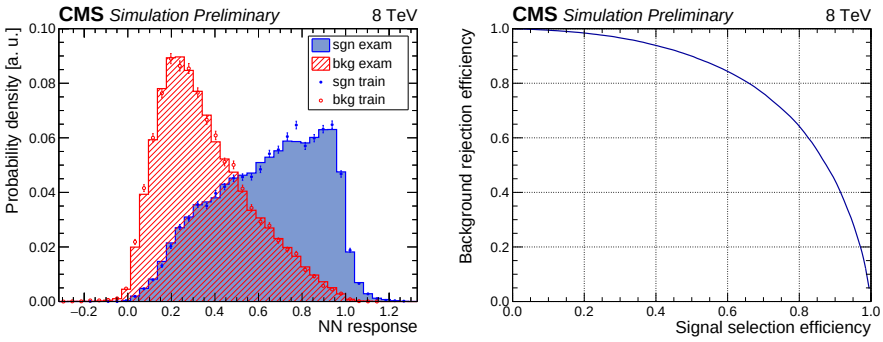


Figure 5.15: Distributions of the event classification NN response for signal and background events (left) and the corresponding ROC curve (right).

Fig. 5.16 demonstrates a good agreement between data and simulation in the response of the event classification neural network in the $t\bar{t}$ control region. Distributions in the signal regions were purposely not reviewed until the statistical analysis was finished. They will be reported in the next chapter. This policy, known as blinding, is applied to protect against potential subjective bias, which occurs when an analysis is unintentionally adjusted to reproduce desired behaviour in the signal region. On the other hand, distributions of input variables

shown in Figs. 5.13 – 5.14 could be checked safely because none of them alone provides sufficient discrimination between signal and background.

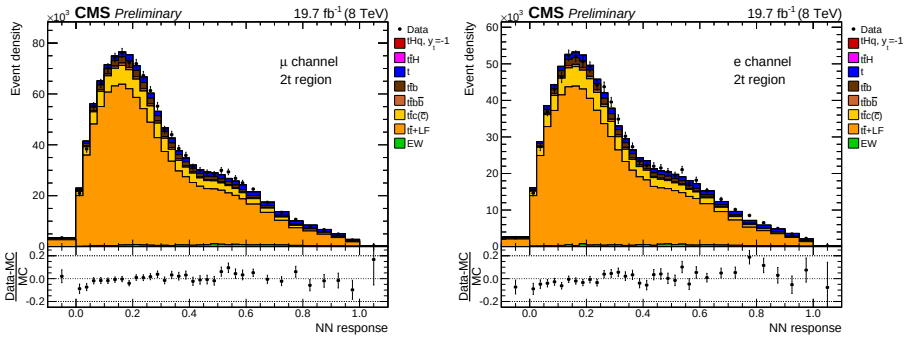


Figure 5.16: Distributions of the event classification NN response in the $2t$ regions. Simulation is normalized to the observed event yield to facilitate comparison of shapes.

Chapter 6

Statistical analysis

In order to search for the signal process, data distributions in response of the neural network introduced in the previous chapter are compared to those deduced from the simulation. The presence of the anomalous tHq production with $y_t = -1$ is probed against the pure SM, where the tHq process is neglected because of its small cross section. Although this search is model-dependent and the expected cross section for the signal process is known accurately, it is still convenient to probe for a range of alternatives with varying signal cross section. Results of the search are reported in the form of upper limits on the production cross section.

This chapter describes the adopted procedure to calculate the upper limits, with a brief introduction to relevant statistical concepts. Systematic uncertainties considered in this search are discussed next. Finally, the upper limits are reported.

6.1 Upper limits

In the field of high-energy physics, the problem of interval estimation is addressed in a number of different ways [21, Chapter 38]. The approach followed in this work has been agreed upon between the ATLAS and CMS collaborations in the context of searches of the Higgs boson [151]. Formulated in the frequentist interpretation of statistics, the method exploits the "LHC-type" test statistic for upper limits and accounts for nuisance parameters with the

help of the profile construction. The CL_s criterion is used to define the critical region of the underlying statistical test.

A pivotal element of the following discussion is the likelihood function, which defines the statistical model used. In this search data are described in the form of histograms, and the likelihood function is therefore built as a product of Poissonian terms corresponding to the bins of the histogram:

$$L(\mu, \nu) = \prod_i \frac{\big(\mu \cdot s_i(\nu) + b_i(\nu)\big)^{n_i}}{n_i!} e^{-(\mu \cdot s_i(\nu) + b_i(\nu))}. \tag{6.1}$$

Here n_i is the number of data events observed in bin i, s_i and b_i are expected numbers of events stemming from the signal process and all backgrounds respectively. It is customary to consider a scale factor $\mu \geqslant 0$ modifying the signal cross section, where $\mu = 1$ corresponds to the nominal value, i. e. the cross section predicted under the $y_t = -1$ model. Expectations s_i and b_i depend on (a vector of) nuisance parameters ν that encode systematic uncertainties.

6.1.1 Hypothesis tests and confidence intervals

Construction of confidence intervals (including one-sided limits as a special case) can be treated as a problem of hypothesis testing. In this case the interval for a parameter of interest of the statistical model is given by the union of all values of the parameter that are not excluded by the test. In the context of putting constraints on the signal strength, the *null hypothesis* is chosen to assume the presence of both signal, with some fixed signal strength μ that is put under the test, and background; in the following it is also referred to as the $s + b$ hypothesis. The *alternative hypothesis*, hereafter also mentioned as the b-only hypothesis for short, assumes that the sample is populated by background events only. It can be constructed as a special case of the $s + b$ hypothesis by taking $\mu = 0$. Although in general both hypotheses contain a number of parameters, which correspond to systematic uncertainties, for the sake of simplicity in the current discussion the hypotheses are considered to be *simple*, i. e. they do not contain any parameters and therefore completely define population distributions.

In order to construct a test, one needs to define a *test statistic*. It is a value calculated from the sample that quantifies its compatibility with the null hypothesis as opposed to the alternative one. The statistic is often chosen to increase monotonically for the increasing (decreasing) compatibility with the $s+b$ (b-only) hypothesis. The concrete statistic exploited in this search will be

introduced later, but a good example is the ratio of the likelihood for the $s+b$ and b-only hypotheses:

$$\lambda(\mu) = \frac{L_{s+b}}{L_b} = \frac{L(\mu)}{L(0)}, \tag{6.2}$$

where the nuisance parameters in the likelihood function (6.1) are ignored because the hypotheses are assumed to be simple.

Definition of a test is completed by choosing the *critical value* of the statistic. In order to perform the test with the results of an experiment, the observed value t of the statistic is calculated and then compared to the critical value t^*. The $s+b$ hypothesis is accepted if $t \geqslant t^*$, and rejected otherwise. Values of the test statistic for which the null hypothesis is rejected (i. e. $t < t^*$) form the *critical region*. Since in practical applications it is not possible to discriminate the two hypotheses with all certainty, the result of the test might be mistaken. The probability to incorrectly reject the $s+b$ hypothesis when it is true, which is to commit a *type I error*, is called the *size* of the test α. When simple hypotheses are considered, the size of the test coincides with its *significance level*. A complementary property is the probability to incorrectly accept the false $s+b$ hypothesis (*type II error*). This probability β is related to the *power* of the test $(1-\beta)$. Fig. 6.1 provides a graphical representation of both error probabilities. Usually the significance level α is fixed to a predefined value (typical choices are 1% or 5%), and one tries to find a test with maximal power while respecting the constraint on the significance level. According to the Neyman–Pearson lemma, the optimal test for simple hypotheses is the one that utilizes the likelihood-ratio statistic (6.2).

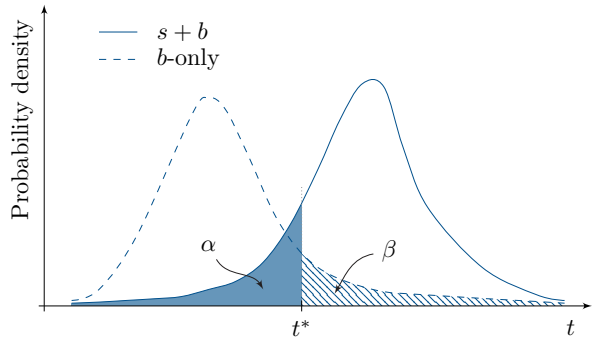


Figure 6.1: An illustration to hypothesis testing. Distributions of the test statistic t under $s+b$ and b-only hypotheses are shown. The solid (hatched) area corresponds to the probability of type I (type II) error.

In practice, finding the critical value of the test statistic for the desired significance level can be a computationally intensive task. An equivalent but sometimes more convenient approach is to calculate the p-value under the $s+b$ hypothesis for the observed test statistic. The *p-value* is defined as the probability to obtain an observation of the same or a larger level of incompatibility with the hypothesis in question when compared to the actual observation. With the chosen direction of ordering of the statistic (larger values indicate better compatibility with the $s+b$ hypothesis), the p-value for an observation t_{obs} is calculated as

$$p_{s+b} = \int_{-\infty}^{t_{\mathrm{obs}}} f\left(t \mid \mu\right) \, \mathrm{d}t, \tag{6.3}$$

where $f(t \mid \mu)$ is the distribution of the test statistic under the $s+b$ hypothesis with the given signal strength μ. If p_{s+b} is larger than the predefined significance level, the $s+b$ hypothesis is accepted; otherwise it is rejected.

The outcome of the test depends on the signal strength μ assumed in the $s+b$ hypothesis. The union of all values of μ for which the $s+b$ hypothesis is not excluded constitutes the *confidence interval* for μ. The interval is said to correspond to a *confidence level* of $1-\alpha$. The confidence interval is a function of the sample and thus varies when the experiment is repeated. The fraction of repeated experiments when the interval includes the true value of μ is called the *coverage probability*. In an ideal case the coverage probability equals the confidence level (which can be understood as the target coverage probability). However, this condition can not always be satisfied, and a real interval might under- or over-cover.

6.1.2 CL_s criterion

When the procedure introduced above is applied to low-sensitivity searches, it might lead to seemingly counter-intuitive results. In an experiment that is not sensitive enough to discriminate the $s+b$ and b-only hypotheses (which occurs with necessity in every search if a sufficiently small signal cross section is probed), the two distributions of the test statistic in Fig. 6.1 are (virtually) indistinguishable. However, the rate of type I error is by construction fixed to the size of the test α and does not depend on the distribution of the test statistic under the b-only hypothesis at all. As a result, even if the experiment possesses no sensitivity to the $s+b$ hypothesis, it rejects the hypothesis with a probability α on a purely random basis.

Although the described behaviour is perfectly valid from the statistical point of view, it is not usually desirable in publication of a physical result, where it is preferable not to claim an exclusion of a hypothesis to which the experiment has no sensitivity. The reason of the controversy is that confidence intervals (and frequentist statistics in general) address the question of the likelihood of observed data under an assumption of a given hypothesis; in contrast to it, what physicists are really interested in is the likelihood of a hypothesis, given the observed data. The latter question is in focus of the Bayesian school of statistics and requires introduction of prior probabilities, which are subjective in nature and thus not always desirable for reporting results.

To overcome the inconvenience in case of a low sensitivity, the CL_s criterion was proposed [152–154]. It exploits a statistic

$$CL_s = \frac{p_{s+b}}{1 - p_b}, \tag{6.4}$$

where p_{s+b} and p_b are p-values under $s + b$ and b-only hypotheses respectively (see Fig. 6.2). Note that since a smaller value of the test statistic t corresponds to a better compatibility with the b-only hypothesis,

$$p_b = \int_{t_{\mathrm{obs}}}^{+\infty} f\left(t \mid 0\right) \mathrm{d}t \tag{6.5}$$

(compare to Eq. (6.3)). The $s + b$ hypothesis is rejected if $CL_s < \alpha$, where α is the desired significance level. As before, the union of all values of the signal strength μ for which the $s + b$ hypothesis is not rejected, forms an allowed interval for μ that formally corresponds to a confidence level $1 - \alpha$.

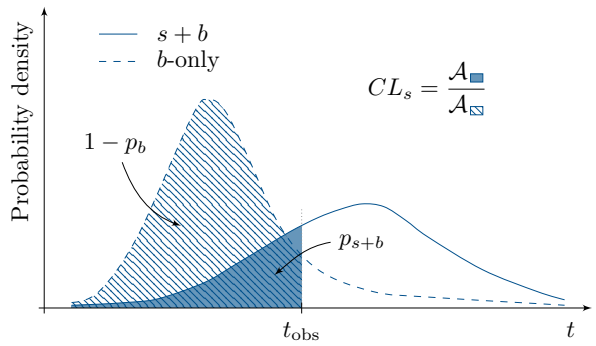


Figure 6.2: Definition of the CL_s statistic. The figure shows distributions of the test statistic t under the $s+b$ and b-only hypotheses and the observed value t_{obs}. The solid (hatched) area equals p_{s+b} $(1 - p_b)$.

As can be seen from Eq. (6.4), $CL_s \geqslant p_{s+b}$, which means that the CL_s criterion leads to a deliberately conservative coverage. Because of this property, resulting intervals do not have same meaning as the frequentist confidence intervals discussed before. Nevertheless, if the test statistic t separates the $s+b$ and b-only hypotheses efficiently, CL_s approaches p_{s+b} for large values of t; if, in addition, $p_b \approx 1$ for the chosen critical value t^*, i. e. t^* is large enough, results of the test based on the p_{s+b} criterion are reproduced. On the other hand, in the problematic case of a low sensitivity $p_{s+b} \approx 1 - p_b$ and hence $CL_s \approx 1$, which prevents the exclusion.

It should be stressed out that the intervals obtained using the CL_s criterion are neither frequentist confidence nor Bayesian credible intervals. The criterion has been developed as an ad hoc procedure and does not have solid foundations in the statistical theory. Nevertheless, it possesses useful practical properties and has been proven functional in a number of searches performed at LEP [155] and Tevatron [156] colliders, and in recent years also at LHC [1, 2].

6.1.3 Nuisance parameters

The likelihood-ratio statistic (6.2) is only proven optimal to test simple hypotheses, whereas in real-life applications statistical models usually contain nuisance parameters. Several approaches have been proposed to incorporate them, and overviews can be found in Refs. [157–159]. The most straightforward way is to perform the full *Neyman construction* over all nuisance parameters using the likelihood-ratio statistic. For every fixed point ν in the allowed space of nuisance parameters an interval X_ν for the signal strength μ is constructed. Since with the fixed nuisances the $s+b$ and b-only hypotheses are effectively simple, the chosen statistic is optimal for each point ν individually. All intervals are then united to ensure a conservative coverage for μ: $X = \bigcup X_\nu$. The resulting interval X contains all values of μ for which the $s+b$ hypothesis is not excluded at least for some point in the ν-space.

Although being very simple conceptually, this approach is usually impractical because it demands large computational resources. Another drawback is that the requirement to ensure at least the nominal coverage for all allowed values of the nuisance parameters can lead to a severe over-coverage for their true values [157].

The latter problem can be mitigated by using as the test statistic the profile likelihood ratio

$$\lambda_{\mathrm{p}}(\mu) = \frac{L(\mu, \hat{\nu}_\mu)}{L(\hat{\mu}, \hat{\nu})}, \tag{6.6}$$

where $\hat{\mu}$ and $\hat{\nu}$ maximize the likelihood L unconditionally and $\hat{\nu}_\mu$ maximizes it for the given μ. Since signal strength is non-negative, it is often required that the likelihood in the denominator is maximized respecting the physical constraint $\hat{\mu} \geqslant 0$. If nuisances are neglected, the ratio λ_p simplifies to the well-known Feldman–Cousins ordering rule [160]. An important property of the statistic (6.6) is that λ_p, when calculated for the given experiment, does not depend on the nuisance parameters. This gives a hope to obtain a *similar test*, which is a test whose critical region is independent from the choice of point in the ν-space. However, the independence is only approximate because, despite being eliminated from the definition of the test statistic, the nuisances still affect its distributions $f(\lambda_\mathrm{p} \mid \mu', \nu)$ under the $s + b$ and b-only hypotheses (with $\mu' = \mu$ and $\mu' = 0$ respectively) and therefore also the p-values (6.3) and (6.5).

The approximate similarity of a test based on the statistic (6.6) also allows to reduce the demand for computational power, as suggested in Ref. [161]. With the full Neyman construction of CL_s intervals, a signal strength μ is excluded if $CL_s(\mu, \nu) < \alpha$ for all allowed values of ν. As can be seen from the definition (6.4),

$$\max_\nu \big(CL_s(\mu, \nu)\big) \leqslant CL_s^*(\mu) \equiv \frac{p_{s+b}(\mu, \tilde{\nu}_\mu)}{1 - p_b(\tilde{\nu}_0)}, \tag{6.7}$$

where

$$\tilde{\nu}_\mu = \arg\max_\nu \big(p_{s+b}(\mu, \nu)\big), \quad \tilde{\nu}_0 = \arg\max \big(p_b(\nu)\big). \tag{6.8}$$

Thus, asking for $CL_s^*(\mu) < \alpha$ provides a conservative coverage. The true values of the nuisance parameters (probably, except for those related to signal only) are, of course, same for the $s + b$ and b-only hypotheses, and having the two potentially different values (6.8) is only an artificial approach to estimate CL_s from above. The very calculation of $\tilde{\nu}_\mu$ and $\tilde{\nu}_0$ is a difficult problem, but since larger p-values indicate better compatibility with the hypothesis in question, they can be approximated by the conditional maximum-likelihood estimates $\hat{\nu}_\mu$ and $\hat{\nu}_0$. If the similarity were exact, it would be sufficient to perform the construction for a single (arbitrary) point in the ν-space; with the approximate similarity choosing the most favourable point for each hypothesis provides an approximation to the full Neyman construction. This approach is discussed under the name of *profile construction* in Refs. [158, 159]. It provides a good scalability with the number of nuisance parameters. Note, however, that the conservative coverage is not guaranteed, and the profile construction can still result in under-coverage if the p-values calculated with nuisances $\hat{\nu}_\mu$ and $\hat{\nu}_0$ depart sufficiently from their maximal values.

The actual test statistic exploited in this search is

$$\tilde{q}(\mu) = -2 \ln \frac{L(\mu, \hat{\nu}_\mu)}{L(\hat{\mu}, \hat{\nu})}, \quad 0 \leqslant \hat{\mu} \leqslant \mu, \tag{6.9}$$

as introduced in Ref. [162]. Together with other related statistics, it is sometimes referenced by the name of "LHC-type" statistic, distinguishing it from definitions adopted in LEP and Tevatron experiments. The argument of the logarithm coincides with the profile likelihood ratio (6.6) except for the constraint on $\hat{\mu}$ imposed in maximisation of the denominator. The first inequality $\hat{\mu} \geqslant 0$ signals that physical values of the signal strength are non-negative. The requirement $\hat{\mu} \leqslant \mu$ indicates that if data are better described by a signal strength larger than μ, it should not be regarded as an evidence against the signal strength μ. This makes the statistic $\tilde{q}$ produce intervals that are not detached from zero and can therefore be reinterpreted to obtain upper limits on the signal strength. The logarithm and the numeric factor, being a monotonous transformation, simply serve a purpose of making analytical calculations more convenient. However, because of the negative sign, larger values of $\tilde{q}$ indicate a worse compatibility with the $s + b$ hypothesis, in contrast to the convension adopted before.

6.1.4 Distribution of test statistic

The last missing components needed to test for a particular signal strength are distributions of the test statistic under the $s + b$ and b-only hypotheses, which are required to calculate the p-values in the definition of CL_s (6.4). They can be constructed with the help of Monte Carlo simulation, but this approach is usually computationally expensive. Alternatively, the distributions can be approximated analytically in the large-sample limit [162], i. e. when contents n_i of all bins in Eq. (6.1) are sufficiently large. According to the Wilks theorem, when certain general conditions are met and the $s + b$ hypothesis is true, the quantity $-2 \ln \lambda_\mathrm{p}(\mu)$ asymptotically follows the χ^2-distribution with one degree of freedom, independently of the nuisance parameters.

This result is generalized in Ref. [163] for the case of a hypothesis with a signal strength μ' that can differ from the one assumed in $\lambda_\mathrm{p}(\mu)$. It is found that

$$-2 \ln \lambda_\mathrm{p}(\mu) \approx \frac{(\mu - \hat{\mu})^2}{\sigma^2}, \quad \hat{\mu} \sim \mathcal{N}(\mu', \sigma^2), \tag{6.10}$$

where, as before, $\hat{\mu}$ denotes the point of global maximum of the likelihood. As can be seen, the Wilks result is reproduced for $\mu' = \mu$, whereas in general case

the statistic $-2\ln\lambda_{\mathrm{p}}(\mu)$ follows the non-central χ^2-distribution with one degree of freedom.

The only unknown parameter in Eq. (6.10) is the standard deviation σ of the normal distribution for $\hat{\mu}$. As demonstrated in Ref. [162], it can be approximated as

$$\sigma^2 \approx \sigma_{\mathrm{A}}^2 = -\frac{(\mu-\mu')^2}{2\cdot\ln\lambda_{\mathrm{A}}(\mu)}, \tag{6.11}$$

where

$$\lambda_{\mathrm{A}}(\mu,\nu) = \frac{L_{\mathrm{A}}(\mu,\hat{\nu}_\mu)}{L_{\mathrm{A}}(\mu',\nu)}, \tag{6.12}$$

and L_{A} is the likelihood function (6.1) evaluated for so-called Asimov sample. It is an artificial sample defined by setting observations n_i in Eq. (6.1) to their expectations for the $s+b$ hypothesis with a signal strength μ':

$$n_{i,\mathrm{A}} = \mu'\cdot s_i + b_i. \tag{6.13}$$

Although values $n_{i,\mathrm{A}}$ can be non-integer, this does not pose a problem since the corresponding terms cancel out in the ratio (6.12). Note that if $\mu'\neq\mu$, the standard deviation σ_{A} reveals sensitivity to nuisance parameters via λ_{A}.

From Eq. (6.10) one can derive analytical expression for the distribution of the statistic $\tilde{q}$. The resulting cumulative distribution reads as [162]

$$F\big(\tilde{q}(\mu)\mid\mu',\nu\big) = \begin{cases} \Phi\left(\sqrt{\tilde{q}(\mu)} - \frac{\mu-\mu'}{\sigma}\right) & \text{for} \quad 0<\tilde{q}(\mu)\leqslant\mu^2/\sigma^2, \\ \Phi\left(\frac{\tilde{q}(\mu)-\left(\mu^2-2\mu\mu'\right)/\sigma^2}{2\mu/\sigma}\right) & \text{for} \quad \tilde{q}(\mu)>\mu^2/\sigma^2, \end{cases} \tag{6.14}$$

where Φ is the cumulative distribution function of the standard normal distribution. Using these results to evaluate the p-values in Eq. (6.7), the CL_s value can be estimated from above as

$$CL_s^*(\mu) = \frac{1-F\big(\tilde{q}(\mu)\mid\mu,\hat{\nu}_\mu\big)}{1-F\big(\tilde{q}(\mu)\mid 0,\hat{\nu}_0\big)}, \tag{6.15}$$

where the nuisance parameters are set according to the profile construction.

If the number of events used to construct the likelihood (6.1) is not large enough, the approximations made are not valid. Under these conditions the asymptotic expression (6.15) is known to be biased, claiming stronger exclusion in the absence of signal [151]. In this case distributions must be derived from Monte Carlo simulation.

6.1.5 Overview

In practice the upper limit is found by solving numerically the equation

$$CL_s(\mu) = \alpha, \tag{6.16}$$

where $1-\alpha$ is the desired confidence level. The central element of this procedure is probing for a particular signal strength μ. This step is summarized below.

First, the likelihood function (6.1) is constructed for the observed data. It is maximized with respect to the nuisance parameters, for the $s + b$ and b-only hypotheses independently, finding their values that will be used for the profile construction:

$$\hat{\nu}_{\mu}^{\text{obs}} = \arg\max_{\nu} \big(L(\mu, \nu)\big), \quad \hat{\nu}_{0}^{\text{obs}} = \arg\max_{\nu} \big(L(0, \nu)\big). \tag{6.17}$$

Observed value of the test statistic $\tilde{q}^{\text{obs}}$ is calculated according to Eq. (6.9), assuming the signal strength under the test.

In the asymptotic regime this is sufficient to calculate the observed value of the CL_s statistic from Eq. (6.15). If the analytical approximation cannot be used, pseudoexperiments are generated under the $s+b$ and b-only hypotheses. They are constructed by sampling the number of observed counts in each bin of the likelihood (6.1) from its expectation according to the Poisson distribution:

$$n_i \sim \text{Pois}\left(\mu' \cdot s_i\big(\hat{\nu}_{\mu'}^{\text{obs}}\big) + b_i\big(\hat{\nu}_{\mu'}^{\text{obs}}\big)\right), \tag{6.18}$$

where $\mu' = \mu$ ($\mu' = 0$) for the $s+b$ (b-only) hypothesis. The test statistic $\tilde{q}(\mu)$ is calculated for each pseudoexperiment, and this way the distributions $f(\tilde{q}(\mu) \mid \mu', \hat{\nu}_{\mu'}^{\text{obs}})$ are constructed. Note that although the nuisance parameters are fixed in the pseudoexperiments, they are allowed to vary in calculation of the test statistic (6.9). The distributions are then exploited to calculate the p-values in Eq. (6.4) and find the value of CL_s.

If $CL_s(\mu) < \alpha$, the upper limit is larger than μ, and vice versa. The procedure summarized above is repeated to probe other values of the signal strength until the solution to Eq. (6.16) is found.

An important sanity check is to calculate the upper limit expected under the b-only hypothesis. This is done by generating b-only pseudoexperiments with $\nu = \hat{\nu}_{0}^{\text{obs}}$ and calculating the upper limit for each of them. The most straightforward way would be to repeat exactly the procedure described above, treating each pseudoexperiment as if it were real data. However, if the distributions of the test statistic are to be constructed with Monte Carlo simulation, this would

require enormous computational resources. The limitation is overcome by reusing the distributions $f(\tilde{q}(\mu) \mid \mu', \hat{\nu}_{\mu'}^{\text{obs}})$ constructed with the nuisances fitted to real data, instead of the current pseudoexperiment. This is an approximation, but it is believed to be suitable [164].

From the constructed distribution of expected upper limits usually the median and $1\,\sigma$ and $2\,\sigma$ bands are reported. An incompatibility between the observed and the expected limit is an indirect evidence that data are not described well by the b-only hypothesis.

A software implementation of the outlined procedure is available in a set of internal CMS tools based on the RooStats project [165]. They are widely used by the collaboration in analyses involving the Higgs boson and have been exploited in this search as well.

6.2 Systematic uncertainties

The statistical model used in this search accounts for statistical uncertainties by construction via the Poissonian terms in the likelihood (6.1). Systematic uncertainties, on the other hand, need to be incorporated explicitly by means of the nuisance parameters. In the simplest case they only affect the overall expected event yield from the signal or (a component of) the background while preserving the shape of the NN discriminator distribution. In general, however, also the shape of the distribution can be altered. In the following the two groups of uncertainties are referred to as rate-only and shape-changing, respectively.

Many systematic uncertainties are continuous by nature, that is they are described by real-valued nuisance parameters. Although the continuous parametrization is trivial to implement for rate-only uncertainties, for shape-changing ones the direct construction of the discriminator distribution for every value of the corresponding nuisance parameter is computationally impractical. Instead, it is customary to construct the distributions for only two variations of each independent nuisance parameter, which are conventionally referred to as "up" and "down" variations. The expected number of events $\lambda^{(i)}$ from the given physics process in bin i of the histogram is then deduced from the nominal value $\lambda_0^{(i)}$ and the two systematic variations $\lambda_+^{(i)}$ and $\lambda_-^{(i)}$. In this search the following parametrization is exploited:

$$\lambda^{(i)}(\nu) = \lambda_0^{(i)} + \left(\frac{\Delta\lambda_+^{(i)} - \Delta\lambda_-^{(i)}}{2} + \left(\frac{\Delta\lambda_+^{(i)} + \Delta\lambda_-^{(i)}}{2} \right) \cdot \theta(\nu) \right) \cdot \nu, \tag{6.19}$$

$$\Delta\lambda_\pm^{(i)} = \lambda_\pm^{(i)} - \lambda_0^{(i)}, \tag{6.20}$$

where the parameter ν, which can be identified with one of the nuisance parameters in Eq. (6.1), is varied simultaneously for all bins of the histogram, and $\theta(\nu)$ is a smooth step function defined as

$$\theta(\nu) = \begin{cases} {}^1\!/_8 \cdot \nu \, \left(\nu^2 \left(3\,\nu^2 - 10\right) + 15\right) & \text{if } |\nu| \leqslant 1, \\ \text{sign}(\nu) & \text{otherwise.} \end{cases} \tag{6.21}$$

The function (6.19) and its first and second derivatives are continuous. It reproduces the nominal and the reference "up" and "down" discriminator distributions for $\nu = 0, \pm 1$ and provides a linear extrapolation for $|\nu| > 1$.

An important property of the parametrization (6.19) is that $\lambda(\nu)$ lies between λ_+ and λ_- for $|\nu| < 1$, provided that λ_0 also falls between the systematic variations. It is known as "vertical" interpolation. Although such approach is adequate for this search, it should be noted that in some situations it is not suitable, most notably to interpolate between two peaking distributions with different positions of the peaks. Other methods have been designed to address this problem [166, 167].

6.2.1 Experimental uncertainties

First, uncertainties of experimental origin are discussed. They reflect the current confidence in the detector performance and can potentially be decreased in future, by achieving more precise calibrations and development of better experimental techniques.

Since this search exploits a tight selection on b-tagging, uncertainties on the b-tagging scale factors are important. Their impact is evaluated by varying the scale factors in Eq. (4.4) within their uncertainties [132]. The scale factors for b- and c-quark jets are varied simultaneously and are controlled by a single nuisance parameter. Since no dedicated measurements for c-quark jets are available, uncertainties for them are set twice as large as for b-quark jets. Scale factors for light-flavour jets are varied independently. The b-tagging uncertainty is thus described by two independent parameters.

Effect of uncertainty in the jet energy scale (JES) calibration is found by rescaling four-momenta of all jets with $p_{\mathrm{T}} > 10\,\mathrm{GeV}/c$ by factors $1 + \alpha_j$, where α_j is varied in the range allowed by JES uncertainties for jet j. Momenta of all jets are rescaled simultaneously, and thus this uncertainty is controlled by a single nuisance parameters. It has been checked, however, that splitting the total uncertainty into several independent sources [98] does not alter the results.

Uncertainties on jet energy resolution (JER) are included by varying the smearing factors in Eq. 3.27 within their uncertainties [98]. It affects all jets with $p_{\mathrm{T}} > 10\,\mathrm{GeV}/c$ for which a matching particle-level jet is present. Variations for all jets are done simultaneously.

Variations of jet momenta affect $\not{E}_{\mathrm{T}}$ through its corrections. In this way uncertainties on JES and JER are accounted in $\not{E}_{\mathrm{T}}$. In addition to them, the unclustered missing E_{T}, which is calculated from jets with $p_{\mathrm{T}} < 10\,\mathrm{GeV}/c$ and PF candidates not clustered in jets, is varied by 10% resulting in an independent uncertainty.

Uncertainties in pile-up originate from uncertainties in the cross section of the total inelastic pp scattering and the luminosity uncertainty. It is conservatively estimated by rescaling the effective total inelastic cross section, which is exploited in the reweighting in Section 4.2, by 5%. Integrated luminosity to which predictions from simulation are normalized, is assigned an uncertainty of 2.6% [128].

Finally, differences between data and simulation in efficiencies of lepton identification and trigger selection are conservatively covered by 2% rate-only uncertainties, independently in the electron and muon channels.

6.2.2 Theoretical uncertainties

Another group of uncertainties is related to theoretical inputs. They represent the level of confidence in modelling of considered physics processes, excluding the simulation of the detector response.

Cross sections used to normalize signal and background processes are varied according to uncertainties provided in the references in Section 4.1. The variations reflect the arbitrariness in the choice of the renormalization and factorization scales, commonly set to same value $\mu_{\mathrm{R,F}}$, and uncertainties in the parton distribution functions (PDF). They are summarized in Table 6.1. Scales $\mu_{\mathrm{R,F}}$ in $t\bar{t}$ and single top production are taken to be fully correlated. Similarly, PDF uncertainties are grouped by the dominant initial state.

An important background is the $t\bar{t}$ production in association with heavy-flavour quarks. As discussed in Section 4.4, it is expected that cross sections of these processes are not predicted accurately by the simulation. To protect against this, additional normalization uncertainties of $^{+50\%}_{-30\%}$ are assigned to the $t\bar{t}b\bar{b}$, $t\bar{t}b$, and $t\bar{t}c(\bar{c})$ processes, independently.

Table 6.1: Uncertainties on cross sections. Uncertainties in same column are varied simultaneously.

Process	$\mu_{R,F}$ scale				PDF		
	t	V	VV	$t\bar{t}H$	gg	$q\bar{q}$	qg
tHq							2%
$t\bar{t}H$				12.5%	9%		
$t\bar{t}$	3%				2.6%		
t	2%						4.6%
W		1.3%				4.8%	
VV			3.5%				

Uncertainties due to the arbitrariness in the renormalization/factorization scales $\mu_{R,F}$ in the signal process and the $t\bar{t}$ background are estimated with the help of dedicated samples. They have been produced by varying the scale, which is dynamic in case of $t\bar{t}$, by a factor of 2 in each direction. These uncertainties affect the acceptance and, potentially, the shape of the distributions in NN response, while variations in the cross sections are accounted with independent nuisance parameters as described above. For this reason the samples with varied $\mu_{R,F}$ scales are normalized to the nominal cross sections. In the $t\bar{t}$ process, the variations for the $t\bar{t}b\bar{b}$, $t\bar{t}b$, and $t\bar{t}c(\bar{c})$ components and also $t\bar{t}$ production without additional heavy-flavour quarks are controlled by independent nuisance parameters.

Dedicated $t\bar{t}$ samples are also exploited to evaluate the effect of the choice of the threshold used to match matrix element and parton shower calculations. In the nominal sample jets are separated with a measure of $k_T > 40\,\text{GeV}$, while in the samples with the upwards and downwards variation the threshold is set to 60 and 30 GeV respectively. Similar to the $\mu_{R,F}$ scale uncertainty, variations in the four subcomponents are done independently.

The last uncertainty accounts for the correction of the p_T spectrum of top quarks in the $t\bar{t}$ production, as described in Section 4.2. In order to quantify its impact, the corresponding reweigthing is either disabled or the event weights are squared, redoubling the effect of the reweighting.

6.2.3 Statistical uncertainties in simulation

The last considered type of systematic uncertainty accounts for statistical fluctuations in MC samples. As a direct consequence of the optimization strategy for the event selection, this uncertainty is typically small for the signal process

but can be significant for backgrounds. The expected number of events in a given bin stemming from a process with a cross section σ is calculated as

$$\lambda = \frac{\sigma\mathcal{L}}{N}\sum_{i=1}^{n} w_i, \tag{6.22}$$

where $\mathcal{L}$ is the intergrated luminosity, N is the total number of events in the sample, n is the number of events, out of N, that pass the selection and fall in the chosen bin, and w_i are their weights. The number n follows the binominal or, in a typical case of $n \ll N$, the Poissonian distribution, which propagates into an uncertainty on the expectation λ.

In this search the problem is addressed by introducing an additional nuisance parameter per bin and per sample, following the approach of Ref. [168]. The expected number of events is allowed to vary around the estimation (6.22) within the uncertainty

$$\Delta\lambda = \frac{\sigma\mathcal{L}}{N}\sqrt{\sum_{i=1}^{n} w_i^2}. \tag{6.23}$$

The extensive approach results in about 550 additional nuisance parameters, to be compared to about 30 parameters to describe all other uncertainties. The problem is simplified by neglecting the uncertainty in bins where $\Delta\lambda/\lambda < 5\%$; however, this removes only ~ 100 parameters. A drastic reduction of the dimensionality could be achieved if only one nuisance parameter were introduced per bin to reflect the combined uncertainty from all contributing samples [169]. This approach, however, has not been implemented in the tool used for the statistical inference.

6.2.4 Supplementary studies

It is important to verify that the statistical model describes the data without pathologies. To do so, a maximum-likelihood fit is performed keeping all parameters of the model floating, and constraints imposed on every nuisance parameter by the data are investigated. For each parameter ν independently a profile likelihood is constructed by maximizing over all remaining parameters θ:

$$L(\nu, \hat{\theta}_\nu) = \max_{\theta} L(\nu, \theta). \tag{6.24}$$

The constraints on ν are found with the usual approach of MINOS [170], solving against ν the equation

$$-2\ln L(\nu, \hat{\theta}_\nu) = -2\ln L(\hat{\nu}, \hat{\theta}) + 1, \tag{6.25}$$

Table 6.2: Impacts of different groups of systematic uncertainties on the upper limit. They are evaluated by excluding one group at a time from the statistical model.

Source	Impact, %
b-tagging, b-, c-jets	< 1
b-tagging, light-flavour jets	< 1
JES	3
JER	< 1
unclustered $\not{E}_{\mathrm{T}}$	1
pile-up	< 1
luminosity	< 1
lepton ID	< 1
cross sections, Q^2 scale	< 1
cross sections, PDF	< 1
$t\bar{t}b\bar{b}$ rate	< 1
$t\bar{t}b$ rate	< 1
$t\bar{t}c(\bar{c})$ rate	1
$\mu_{\mathrm{R,F}}$ scale, tHq and $t\bar{t}$	4
ME/PS matching scale	2
top p_{T} reweighting	2
MC uncertainty	< 1

where $(\hat{\nu}, \hat{\theta})$ is the point of global maximum given by the maximum-likelihood fit.

No pathological constraints have been observed in this investigation. In all cases where a nuisance parameter is constrained tighter than dictated by the nominal uncertainty, it has been understood as a result of a conservative estimation of the uncertainty.

Although this question has no direct relevance to the validity of results of this search, it is interesting to quantify the impact of different sources of systematic uncertainties on the upper limit on the anomalous tHq production. First, the expected limit is calculated using the asymptotic approximation described in Section 6.1.4. To evaluate the impact of a group of uncertainties, the limit is then recalculated having the corresponding nuisance parameters in Eq. (6.9) fixed to the maximum-likelihood fit results. This produces a tighter limit, and the improvement is taken as the quantitative measure of the impact of the group of uncertainties. The results are shown in Table 6.2.

In response to the mismodelling of jet pseudorapidity described in Section 4.3,

the impact of doubling of jet energy scale uncertainty is also investigated. The variation corresponding to this uncertainty source is doubled while all other systematic uncertainties are treated in the standard way. Then the expected limit is calculated using the asymptotic approximation, and a degradation of 0.7% with respect to the standard value is observed.

6.3 Results

In this section the upper limit on the cross section of the anomalous tHq production is reported. The search has been focused on the $H \to b\bar{b}$ decay, but complementary results in other decay channels are also summarized and a combination of all channels is presented.

6.3.1 Results in the $H \to b\bar{b}$ decay channel

The statistical model described above is exploited to set an upper limit on the signal cross section. For this purpose the $3t$ and $4t$ regions are considered independently and are further divided into muon and electron channels. In each of the resulting four regions the NN response is described with the help of a histogram containing 15 equidistant bins. From these histograms the likelihood function (6.1) is constructed.

In order to verify that the statistical model is capable of providing an accurate description of the data, a maximum-likelihood fit is performed. Fig. 6.3 shows the resulting distributions of the NN response, which are constructed from the histograms exploited in the definition of the likelihood function. All simulated processes are normalized to the results of the fit, and systematic uncertainties are constrained accordingly. Of special interest are normalizations of the $t\bar{t}$+HF components. Compared to expectations, they are scaled by factors 1.3 ± 0.2, 1.4 ± 0.2, and 1.2 ± 0.3 for $t\bar{t}b\bar{b}$, $t\bar{t}b$, and $t\bar{t}c(\bar{c})$ respectively. Together with adjustments of other nuisance parameters, this leads to an accurate description of the data.

An alternative way to characterize the compatibility between the data and simulation is presented in Fig. 6.4. It shows all bins from the four histograms in Fig. 6.3, ordered by the signal-to-background ratio (and grouped together for similar values of the ratio). It can be seen that, because of the small contribution from the signal process, the data are compatible with both $s + b$ and b-only hypotheses.

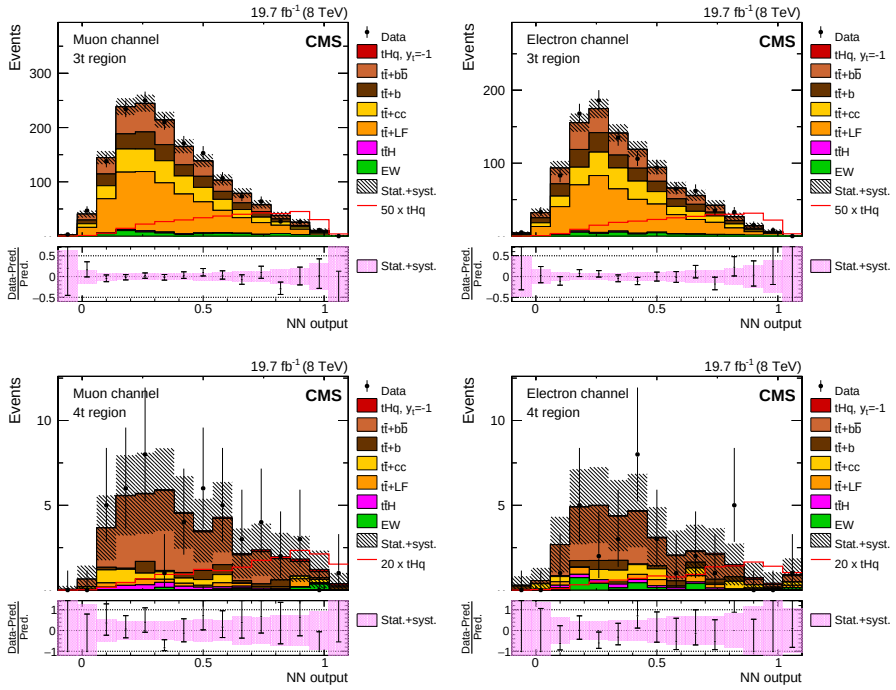


Figure 6.3: Distributions of NN response in the $3t$ and $4t$ regions, muon and electron channels. All processes are normalized to results of the maximum-likelihood fit to data. In these plots the "electroweak" category includes also single top quark production.

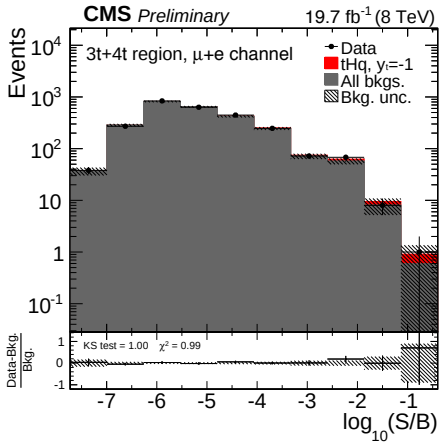


Figure 6.4: Comparison of data and MC expectation as a function of the S/B ratio.

The upper limit is set following procedure documented in Section 6.1. As discussed in Section 6.1.4, the distributions of the test statistic, which are needed to calculate the upper limit, can be constructed by means of MC simulation or using an analytical approximation. Both approaches have been tried and lead to different results, which indicates that conditions for the analytical approximation do not hold. For this reason, the much more computationally intensive MC-based approach is followed to obtain final results. The calculated upper limits at the 95% confidence level are reported in Table 6.3, for the $3t$ and $4t$ regions separately and for their combination. The uncertainties in the expected limits correspond to the 15.9% and 84.1% quantiles, whereas the nominal values are given by the medians. As can be seen, the limit in the $4t$ region is weaker than in the $3t$ one despite the higher purity. This is a consequence of the larger statistical uncertainties. Another interesting observation is that the observed combined limit is weaker than the $3t$ alone. This is understood as a result of the $4t$ limit fluctuating about $2\,\sigma$ above the expectation.

The final result of this search is an upper limit on the cross section of the associated production of single top quarks and Higgs bosons in the t channel, with the kinematics of the $y_t = -1$ case. The found value is 7.6 times the expectation for the flipped-sign model. After substituting the expected cross section, this result translates into an upper limit $\sigma_{tHq}^{y_t=-1} < 1.8\,\text{pb}$ at the 95% confidence level.

Table 6.3: Upper limits on the anomalous tHq production at 95% CL, measured in units of the expected cross section $\sigma_{tHq}^{y_t=-1} \approx 234\,\text{fb}$.

Region	Upper limit	
	Observed	Expected
$3t$, $\mu + e$	7.0	$5.7^{+2.4}_{-1.9}$
$4t$, $\mu + e$	19.8	$10.6^{+4.2}_{-2.8}$
combination	7.6	$5.1^{+2.1}_{-1.4}$

6.3.2 Combination with other decay channels

In addition to the search for the anomalous tHq production with $H \to b\bar{b}$ [3], several other Higgs boson decay channels have been utilized by the CMS collaboration [41–43]. All searches target semileptonic decays of the top quark, while the Higgs boson can decay to a pair of photons, W bosons, or τ leptons. In the final state with $H \to WW$ or $\tau\tau$ at least one of the products is expected

to decay leptonically, and considered signatures are 3ℓ, $\ell^{\pm}\ell^{\pm}$, $\ell^{\pm}\ell^{\pm}\tau_{\mathrm{h}}^{\mp}$, where $\ell = \mu, e$ and τ_{h} denotes a hadronically decaying τ lepton. All these searches adopt data-driven methods to estimate important backgrounds.

Upper limits obtained in each search are shown in Table 6.4. The multilepton ($H \to WW, \tau_\ell\tau_\ell$) channel demonstrates an expected sensitivity similar to the $H \to b\bar{b}$ search. On the other hand, the search in the diphoton channel is more sensitive. The reason for this is not only the very clean final state, but also the fact that the enhancement of the branching ratio $\mathcal{B}(H \to \gamma\gamma)$ due to $y_t = -1$ is treated as an additional evidence in favour of the anomalous tHq production.

Table 6.4: Upper limits on the anomalous tHq production at 95% CL. Searches in other Higgs boson decay channels.

Decay channel	Upper limit	
	Observed	Expected
$H \to \gamma\gamma$	4.1	4.1
$H \to WW, \tau_\ell\tau_\ell$	6.7	$5.0^{+2.1}_{-1.4}$
$H \to \tau_\ell\tau_{\mathrm{h}}$	8.9	$11.0^{+5.8}_{-3.6}$

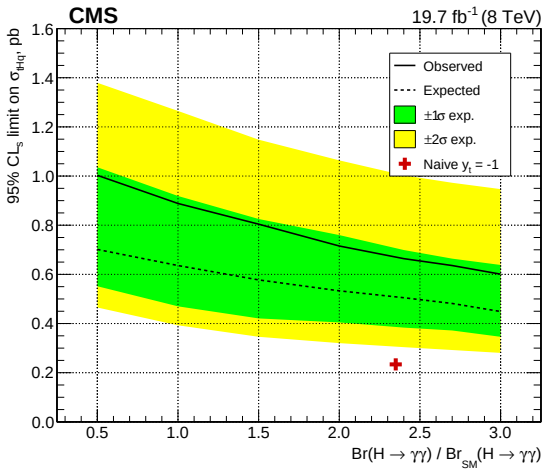


Figure 6.5: Combined limit as a function of the assumed $\mathcal{B}(H \to \gamma\gamma)$.

All CMS searches for the tHq production are combined in Ref. [41]. This requires an assumption on Higgs boson branching ratios. Although they are known in SM [27] and in the $y_t = -1$ model [23], where only $\mathcal{B}(H \to \gamma\gamma)$ is directly affected by the sign of the Yukawa coupling, the $H \to \gamma\gamma$ decay should

be treated with care. In contrast to other considered decays, it occurs through a loop-induced coupling and, as discussed in Section 1.2, can potentially receive significant BSM corrections. To report results in a robust way, no specific assumption on the value of $\mathcal{B}(H \to \gamma\gamma)$ is made, and the combination is performed for a range of possible values of the branching ratio. The results are shown in Fig. 6.5. If the enhancement of factor 2.4 is imposed, as predicted by the pure $y_t = -1$ model, the observed combined upper limit reads as 2.8, with an expectation of $2.2^{+0.8}_{-0.6}$, both expressed in the units of $\sigma_{tHq}^{y_t=-1} \approx 234\,\text{fb}$.

Summary and outlook

A search for the associated tH production in the t channel has been performed in the CMS experiment. It targets the case of an anomalous top quark Yukawa coupling constant $y_t = -1$, where the sign is defined with respect to the coupling between Higgs and W bosons. In the considered final state, the top quark decays producing a muon or an electron, which provides means to suppress the overwhelming multijet background, and the $H \to b\bar{b}$ decay is chosen for its large branching fraction. The event selection takes advantage of multiple b-quark jets in the final state, as well as a jet in the forward region, which is a prominent feature of the t-channel process. Nonetheless, the selected sample of events is dominated by the $t\bar{t}$ background, which has motivated the use of MVA methods to discriminate between the signal process and backgrounds. Construction of efficient input variables in a multijet final state is a challenging task. It is addressed by identifying jets that are likely to originate from fragmentation of decay products of the Higgs boson or the top quarks. Following a novel approach, the identification of jet origin is performed under both tHq and semileptonic $t\bar{t}$ hypotheses in parallel, which allows to construct a dedicated set of input variables for each hypothesis. Both sets, together with some additional observables, are exploited to identify signal-like events. Finally, response of the MVA discriminator is fitted to data to set an upper limit on the tHq cross section. The limit at the 95% confidence level is measured to be $\sigma_{tHq}^{y_t=-1} < 1.8\,\text{pb}$, which is factor 7.6 larger that the expected cross section for the $y_t = -1$ case.

This analysis is complemented by analogous searches that target decays of the Higgs boson to a pair of photons, W bosons, or τ leptons. The search in the $H \to \gamma\gamma$ channel faces the small branching fraction of this decay but, on the other hand, profits from a great reduction of background. Clean experi-

mental signatures are also obtained in the channels with $H \to WW^*/\tau\tau$ and semileptonic decays of the top quark as they can result in final states with two same-sign leptons (muons or electrons) or three leptons. Results of searches for the tHq production with $y_t = -1$ in all considered decay channels were combined. In order to do so, assumptions on the branching fractions had to be made. Probabilities of all considered decays except for $H \to \gamma\gamma$ were set to their SM values because these decays occur at tree level and thus are not expected to be altered by BSM contributions significantly. On the other hand, the branching fraction $\mathcal{B}(H \to \gamma\gamma)$ can potentially be modified by new particles running in the loop. To take this into account, the combined limit was calculated as a function of the assumed value of $\mathcal{B}(H \to \gamma\gamma)$. For the SM branching fraction an upper limit $\sigma_{tHq}^{y_t=-1} < 0.9\,\text{pb}$ is found at the 95% confidence level. This exceeds the expected cross section by a factor of 3.8.

The search in the $H \to b\bar{b}$ channel was performed using about $20\,\text{fb}^{-1}$ of 8 TeV data, and its sensitivity is limited by statistical uncertainties (which equally applies to other decay channels). At $\sqrt{s} = 13\,\text{TeV}$ a fourfold increase of the signal cross section is expected. Although the signal-to-background ratio will not change significantly because the $t\bar{t}$ cross section also increases by a factor of 3.3, the larger production rate will allow to reproduce the 8 TeV result already with $5\,\text{fb}^{-1}$ of 13 TeV data. According to a simple projection [171], the exclusion of the $y_t = -1$ case with the $H \to b\bar{b}$ channel alone can be achieved with an integrated luminosity of $100\,\text{fb}^{-1}$. It should be noted that with this amount of data the sensitivity of the search is entirely limited by systematic uncertainties, which means that if the projection holds, this channel will never allow to observe the much more rare SM tHq production. However, it can be expected that both theoretical and experimental uncertainties decrease in future as a result of improved physics modelling and better understanding of the detector. In addition, there are several ways to advance the search. The most promising option, which was not available in past for technical reasons, is to utilize full jet b-tagging information in construction of input variables, rather than only checking if the jet is b-tagged according to the chosen working point. This will not only provide additional information for the MVA discrimination, but might also allow to loosen the b-tagging requirements in the event selection.

In addition to the test for $y_t = -1$, the tH production allows to examine the interaction between the top quark and the Higgs boson in several other ways, which can become topics of further investigation in future. The presented search can be generalized to derive two-dimensional constraints in the (κ_t, κ_W) plane, where κ_W modifies the strength of the interaction between the Higgs and W bosons, and $\kappa_t \equiv y_t/y_t^{\text{SM}} = y_t$. Although independently $|\kappa_t|$ and $|\kappa_W|$ can be better constrained by studying the $t\bar{t}H$ and HW production,

the tH process can provide complementary information thanks for the involved interference. The interference also allows to access the potential $\mathcal{CP}$-violating phase in the top quark Yukawa coupling. Finally, the tH production can occur through flavour-changing neutral currents (FCNC) involving the Higgs boson, and thus this process can be exploited to probe for FCNC [172, 173], complementing searches in $t\bar{t}$, $t \to Hq$.

Acknowledgements

This thesis would never come to life without direct or indirect contribution from many people.

First of all, I would like to acknowledge the work of hundreds of scientists and engineers who designed and constructed the LHC and the CMS detector. I would also like to emphasize the constant effort put into monitoring of detector and accelerator operation, both by crew in control rooms and from remote. Likewise, the excellent physics performance of the experiment has only been achieved thanks to the hard work of a number people developing ever-better reconstruction algorithms. Although it is not possible to name here all the people involved, their contribution is truly crucial.

I am deeply grateful to my adviser, Andrea Giammanco, for his scientific guidance, patience, and constant availability to help. This search has profited greatly from his valuable suggestions, deep expertise, and feeling of the broad picture. And his friendly and open attitude made working together a genuinely positive experience.

The topic of this search was originally proposed to me by Fabio Maltoni. While conducting the studies, I collaborated with a number of people, and I would like to give special thanks to Benedikt Maier, Christian Böser, Simon Fink, Nils Faltermann, Afiq Anuar, and Dan Knowlton. The paper on combination of all tH searches in CMS would not be possible without the work of Fabrizio Margaroli and Kenneth Bloom. I would also like to thank Matthias Komm for fruitful discussions.

I thank Roberto Chierici, Barbara Clearbaux, Vincent Lemaître, and Fabio Maltoni as well as Christophe Delaere and Jorgen D'Hondt for accepting to be in my dissertation committee and the comité d'accompagnement. I appreciate

a lot their constructive criticism and suggestions, which have helped to shape this thesis.

During my postgraduate study, I was a recipient of the FRIA scholarship funded by the Fonds de la Recherche Scientifique (FNRS). I am grateful to the selection committee for supporting my application.

I would also like to thank Ginette Tabordon for her assistance in administrative procedures, and I appreciate the help of Pavel Demin and Jérôme de Favereau in computing infrastructure.

Finally, I am grateful to my parents for their understanding, concern, and emotional support they have been providing from far away during these three years.

Bibliography

[1] ATLAS collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys. Lett. B* **716** (2012) 1–29, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.

[2] CMS collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett. B* **716** (2012) 30–61, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.

[3] CMS collaboration, "Search for $H \to b\bar{b}$ in association with single top quarks as a test of Higgs boson couplings", CMS PAS HIG-14-015, 2014.

[4] S. L. Glashow, "Partial symmetries of weak interactions", *Nucl. Phys.* **22** (1961) 579–588, doi:10.1016/0029-5582(61)90469-2.

[5] S. Weinberg, "A model of leptons", *Phys. Rev. Lett.* **19** (1967) 1264–1266, doi:10.1103/PhysRevLett.19.1264.

[6] A. Salam, "Weak and electromagnetic interactions", in *8^{th} Nobel symposium*, pp. 367–377. 1968.

[7] M. Gell-Mann, "A schematic model of baryons and mesons", *Phys. Lett.* **8** (1964) 214–215, doi:10.1016/S0031-9163(64)92001-3.

[8] G. Zweig, "An SU_3 model for strong interaction symmetry and its breaking", in *Developments in the quark theory of hadrons*, D. Lichtenberg and S. P. Rosen, eds., volume 1, pp. 22–101. 1964.

[9] H. Fritzsch and M. Gell-Mann, "Current algebra: Quarks and what else?", in *16^{th} international conference on high-energy physics (ICHEP)*, pp. 135–165. 1972. arXiv:hep-ph/0208010.

[10] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, "Advantages of the Color Octet Gluon Picture", *Phys. Lett. B* **47** (1973) 365–368, doi:10.1016/0370-2693(73)90625-4.

[11] CDF collaboration, "Observation of top quark production in $\bar{p}p$ collisions", *Phys. Rev. Lett.* **74** (1995) 2626–2631, doi:10.1103/PhysRevLett.74.2626, arXiv:hep-ex/9503002.

[12] D0 collaboration, "Observation of the top quark", *Phys. Rev. Lett.* **74** (1995) 2632–2637, doi:10.1103/PhysRevLett.74.2632, arXiv:hep-ex/9503003.

[13] O. Eberhardt et al., "Impact of a Higgs boson at a mass of 126 GeV on the standard model with three and four fermion generations", *Phys. Rev. Lett.* **109** (2012) 241802, doi:10.1103/PhysRevLett.109.241802, arXiv:1209.1101.

[14] LHCb collaboration, "Observation of $J/\psi\, p$ resonances consistent with pentaquark states in $\Lambda_b^0 \to J/\psi\, K^- p$ decays", *Phys. Rev. Lett.* **115** (2015), no. 7, 072001, doi:10.1103/PhysRevLett.115.072001, arXiv:1507.03414.

[15] F. Englert and R. Brout, "Broken symmetry and the mass of gauge vector mesons", *Phys. Rev. Lett.* **13** (1964) 321–323, doi:10.1103/PhysRevLett.13.321.

[16] P. W. Higgs, "Broken symmetries, massless particles and gauge fields", *Phys. Lett.* **12** (1964) 132–133, doi:10.1016/0031-9163(64)91136-9.

[17] P. W. Higgs, "Broken symmetries and the masses of gauge bosons", *Phys. Rev. Lett.* **13** (1964) 508–509, doi:10.1103/PhysRevLett.13.508.

[18] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global conservation laws and massless particles", *Phys. Rev. Lett.* **13** (1964) 585–587, doi:10.1103/PhysRevLett.13.585.

[19] P. W. Higgs, "Spontaneous Symmetry Breakdown without Massless Bosons", *Phys. Rev.* **145** (1966) 1156–1163, doi:10.1103/PhysRev.145.1156.

[20] T. W. B. Kibble, "Symmetry breaking in nonAbelian gauge theories", *Phys. Rev.* **155** (1967) 1554–1561, doi:10.1103/PhysRev.155.1554.

[21] Particle Data Group Collaboration, "Review of particle physics", *Chin. Phys. C* **38** (2014) 090001, doi:10.1088/1674-1137/38/9/090001.

[22] F. Maltoni, K. Paul, and S. Willenbrock, "Associated production of Higgs and single top at hadron colliders", *Phys. Rev. D* **64** (2001) 094023, doi:10.1103/PhysRevD.64.094023, arXiv:hep-ph/0106293.

[23] S. Biswas, E. Gabrielli, and B. Mele, "Single top and Higgs associated production as a probe of the Htt coupling sign at the LHC", *JHEP* **01** (2013) 088, doi:10.1007/JHEP01(2013)088, arXiv:1211.0499.

[24] G. Bordes and B. van Eijk, "On the associate production of a neutral intermediate mass Higgs boson with a single top quark at the LHC and SSC", *Phys. Lett. B* **299** (1993) 315–320, doi:10.1016/0370-2693(93)90266-K.

[25] T. M. P. Tait and C.-P. Yuan, "Single top quark production as a window to physics beyond the standard model", *Phys. Rev. D* **63** (2000) 014018, doi:10.1103/PhysRevD.63.014018, arXiv:hep-ph/0007298.

[26] M. Farina et al., "Lifting degeneracies in Higgs couplings using single top production in association with a Higgs boson", *JHEP* **05** (2013) 022, doi:10.1007/JHEP05(2013)022, arXiv:1211.3736.

[27] LHC Higgs Cross Section Working Group, "Handbook of LHC Higgs cross sections: 3. Higgs properties", doi:10.5170/CERN-2013-004, arXiv:1307.1347.

[28] V. Barger, M. McCaskey, and G. Shaughnessy, "Single top and Higgs associated production at the LHC", *Phys. Rev. D* **81** (2010) 034020, doi:10.1103/PhysRevD.81.034020, arXiv:0911.1556.

[29] S. Biswas, E. Gabrielli, F. Margaroli, and B. Mele, "Direct constraints on the top-Higgs coupling from the 8 TeV LHC data", *JHEP* **07** (2013) 073, doi:10.1007/JHEP07(2013)073, arXiv:1304.1822.

[30] J. Chang, K. Cheung, J. S. Lee, and C.-T. Lu, "Probing the top-Yukawa coupling in associated Higgs production with a single top quark", *JHEP* **05** (2014) 062, doi:10.1007/JHEP05(2014)062, arXiv:1403.2053.

[31] B. Hespel, F. Maltoni, and E. Vryonidou, "Higgs and Z boson associated production via gluon fusion in the SM and the 2HDM", *JHEP* **06** (2015) 065, doi:10.1007/JHEP06(2015)065, arXiv:1503.01656.

[32] J. Baglio et al., "The measurement of the Higgs self-coupling at the LHC: theoretical status", *JHEP* **04** (2013) 151, doi:10.1007/JHEP04(2013)151, arXiv:1212.5581.

[33] E. Gabrielli et al., "Anomalous Higgs-boson coupling effects in HW^+W- production at the LHC", *Phys. Rev. D* **89** (2014), no. 5, 053012, doi:10.1103/PhysRevD.89.053012, arXiv:1312.4956.

[34] F. Demartin, F. Maltoni, K. Mawatari, and M. Zaro, "Higgs production in association with a single top quark at the LHC", *Eur. Phys. J. C* **75** (2015), no. 6, 267, doi:10.1140/epjc/s10052-015-3475-9, arXiv:1504.00611.

[35] J. Ellis, D. S. Hwang, K. Sakurai, and M. Takeuchi, "Disentangling Higgs-top couplings in associated production", *JHEP* **04** (2014) 004, doi:10.1007/JHEP04(2014)004, arXiv:1312.5736.

[36] J. Yue, "Enhanced thj signal at the LHC with $h \to \gamma\gamma$ decay and $\mathcal{CP}$-violating top-Higgs coupling", *Phys. Lett. B* **744** (2015) 131–136, doi:10.1016/j.physletb.2015.03.044, arXiv:1410.2701.

[37] F. Boudjema, R. M. Godbole, D. Guadagnoli, and K. A. Mohan, "Laboratory-frame observables for probing the top-Higgs boson interaction", *Phys. Rev. D* **92** (2015), no. 1, 015019, doi:10.1103/PhysRevD.92.015019, arXiv:1501.03157.

[38] S. Khatibi and M. Mohammadi Najafabadi, "Exploring the anomalous Higgs-top couplings", *Phys. Rev. D* **90** (2014), no. 7, 074014, doi:10.1103/PhysRevD.90.074014, arXiv:1409.6553.

[39] ATLAS collaboration, "The ATLAS experiment at the CERN Large hadron collider", *JINST* **3** (2008) S08003, doi:10.1088/1748-0221/3/08/S08003.

[40] CMS collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[41] CMS collaboration, "Search for the associated production of a Higgs boson with a single top quark in proton-proton collisions at $\sqrt{s} = 8\,\text{TeV}$", *submitted to JHEP* (2015) arXiv:1509.08159.

[42] CMS collaboration, "Search for associated production of a single top quark and a Higgs boson in events where the Higgs boson decays to two photons at $\sqrt{s} = 8\,\text{TeV}$", CMS PAS HIG-14-001, 2014.

[43] CMS collaboration, "Search for associated production of a single top quark and a Higgs boson in leptonic channels", CMS PAS HIG-14-026, 2015.

[44] ATLAS collaboration, "Search for $H \to \gamma\gamma$ produced in association with top quarks and constraints on the Yukawa coupling between the top quark and the Higgs boson using data taken at 7 TeV and 8 TeV with the ATLAS detector", *Phys. Lett. B* **740** (2015) 222–242, doi:10.1016/j.physletb.2014.11.049, arXiv:1409.3122.

[45] CMS collaboration, "Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV", *Eur. Phys. J. C* **75** (2015), no. 5, 212, doi:10.1140/epjc/s10052-015-3351-7, arXiv:1412.8662.

[46] ATLAS collaboration, "Measurements of the Higgs boson production and decay rates and coupling strengths using *pp* collision data at $\sqrt{s} = 7$ and 8 TeV in the ATLAS experiment", *submitted to EPJC* (2015) arXiv:1507.04548.

[47] ATLAS collaboration, "Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector", *submitted to EPJC* (2015) arXiv:1506.05669.

[48] J. Brod, U. Haisch, and J. Zupan, "Constraints on CP-violating Higgs couplings to the third generation", *JHEP* **11** (2013) 180, doi:10.1007/JHEP11(2013)180, arXiv:1310.1385.

[49] ACME collaboration, "Order of magnitude smaller limit on the electric dipole moment of the electron", *Science* **343** (2014) 269–272, doi:10.1126/science.1248213, arXiv:1310.7534.

[50] L. Evans and P. Bryant, "LHC machine", *JINST* **3** (2008) S08001, doi:10.1088/1748-0221/3/08/S08001.

[51] LHCb collaboration, "The LHCb detector at the LHC", *JINST* **3** (2008) S08005, doi:10.1088/1748-0221/3/08/S08005.

[52] ALICE collaboration, "The ALICE experiment at the CERN LHC", *JINST* **3** (2008) S08002, doi:10.1088/1748-0221/3/08/S08002.

[53] M. Ferro-Luzzi, W. Herr, and T. Pieloni, "LHC bunch filling schemes for commissioning and initial luminosity optimization", Technical Report LHC-PROJECT-NOTE-415, 2008.

[54] M. Lamont, "LHC: Status and commissioning plans", in *44^{th} rencontres de Moriond on QCD and high energy interactions*, pp. 21–26. 2009. arXiv:0906.0347.

[55] twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults.

[56] M. Lamont, "The first years of LHC operation for luminosity production", in *4^{th} international particle accelerator conference (IPAC 2013)*, p. MOYAB101. 2013.

[57] J. Wenninger, "Strategy for Run 2", in *LHC performance workshop (Chamonix 2014)*. 2014.

[58] CMS collaboration, "CMS physics: Technical design report. Volume I: Detector performance and software", Technical Report CMS-TDR-008-1, 2006.

[59] CMS collaboration, "CMS physics: Technical design report. Volume II: Physics performance", *J. Phys. G* **34** (2007) 995–1579, doi:10.1088/0954-3899/34/6/S01.

[60] CMS collaboration, "Precise mapping of the magnetic field in the CMS barrel yoke using cosmic rays", *JINST* **5** (2010) T03021, doi:10.1088/1748-0221/5/03/T03021, arXiv:0910.5530.

[61] CMS collaboration, "Description and performance of track and primary-vertex reconstruction with the CMS tracker", *JINST* **9** (2014), no. 10, P10009, doi:10.1088/1748-0221/9/10/P10009, arXiv:1405.6569.

[62] CMS collaboration, "Performance of the CMS drift tube chambers with cosmic rays", *JINST* **5** (2010) T03015, doi:10.1088/1748-0221/5/03/T03015, arXiv:0911.4855.

[63] CMS HCAL collaboration, "Energy response and longitudinal shower profiles measured in CMS HCAL and comparison with Geant4", Technical Report CMS-NOTE-2006-143, 2007.

[64] CMS HCAL collaboration, "Design, performance, and calibration of CMS hadron endcap calorimeters", Technical Report CMS-NOTE-2008-010, 2008.

[65] CMS HCAL collaboration, "Design, performance and calibration of the CMS forward calorimeter wedges", *Eur. Phys. J. C* **53** (2008) 139–166, doi:10.1140/epjc/s10052-007-0459-4.

[66] CMS collaboration, "The CMS high level trigger", *Eur. Phys. J. C* **46** (2006) 605–667, doi:10.1140/epjc/s2006-02495-8, arXiv:hep-ex/0512077.

[67] CMS collaboration, "The CMS trigger", *to be submitted to JINST* (2015). CMS-TRG-12-001, in preparation.

[68] A. Strandlie and R. Frühwirth, "Track and vertex reconstruction: From classical to adaptive methods", *Rev. Mod. Phys.* **82** (2010) 1419–1458, doi:10.1103/RevModPhys.82.1419.

[69] P. Billoir and S. Qian, "Fast vertex fitting with a local parametrization of tracks", *Nucl. Instrum. Meth. A* **311** (1992) 139–150, doi:10.1016/0168-9002(92)90859-3.

[70] P. Billoir, "Progressive track recognition with a Kalman like fitting procedure", *Comput. Phys. Commun.* **57** (1989) 390–394, doi:10.1016/0010-4655(89)90249-X.

[71] R. Frühwirth, "Application of Kalman filtering to track and vertex fitting", *Nucl. Instrum. Meth. A* **262** (1987) 444–450, doi:10.1016/0168-9002(87)90887-4.

[72] CMS collaboration, "Offline primary vertex reconstruction with deterministic annealing clustering", CMS Internal Note 2011-014, 2010.

[73] K. Rose, "Deterministic annealing for clustering, compression, classification, regression, and related optimization problems", *Proc. IEEE* **86** (1998) 2210–2239, doi:10.1109/5.726788.

[74] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing", *Science* **220** (1983) 671–680, doi:10.1126/science.220.4598.671.

[75] V. Černý, "Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm", *J. Optimiz. Theory App.* **45** (1985), no. 1, 41–51, doi:10.1007/BF00940812.

[76] R. Frühwirth, W. Waltenberger, and P. Vanlaer, "Adaptive vertex fitting", *J. Phys. G* **34** (2007) N343, doi:10.1088/0954-3899/34/12/N01.

[77] T. Miao, H. Wenzel, F. Yumiceva, and N. Leioatts, "Beam position determination using tracks", Technical Report CMS-NOTE-2007-021, 2007.

[78] G. Bruno et al., "Local reconstruction in the muon detectors", Technical Report CMS-NOTE-2002-043, 2002.

[79] CMS collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7\,\text{TeV}$", *JINST* **7** (2012) P10002, doi:10.1088/1748-0221/7/10/P10002, arXiv:1206.4071.

[80] CMS collaboration, "Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8\,\text{TeV}$", *JINST* **10** (2015), no. 06, P06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.

[81] S. Baffioni et al., "Electron reconstruction in CMS", *Eur. Phys. J. C* **49** (2007) 1099–1116, doi:10.1140/epjc/s10052-006-0175-5.

[82] CMS collaboration, "Particle-flow commissioning with muons and electrons from J/ψ and W events at 7 TeV", CMS PAS PFT-10-003, 2010.

[83] CMS collaboration, "Particle-flow event reconstruction in CMS and performance for jets, taus, and MET", CMS PAS PFT-09-001, 2009.

[84] M. Anderson et al., "Review of clustering algorithms and energy corrections in ECAL", Technical Report CMS-IN-2010-008, 2010.

[85] H. Bethe and W. Heitler, "On the stopping of fast particles and on the creation of positive electrons", *Proc. Roy. Soc. Lond. A* **146** (1934) 83–112, doi:10.1098/rspa.1934.0140.

[86] W. Adam, R. Frühwirth, A. Strandlie, and T. Todorov, "Reconstruction of electrons with the Gaussian sum filter in the CMS tracker at LHC", *J. Phys. G* **31** (2005) N9, doi:10.1088/0954-3899/31/9/N01, arXiv:physics/0306087.

[87] R. Frühwirth, "Track fitting with non-Gaussian noise", *Comput. Phys. Commun.* **100** (1997) 1–16, doi:10.1016/S0010-4655(96)00155-5.

[88] CMS collaboration, "Commissioning of the particle-flow reconstruction in minimum-bias and jet events from pp collisions at 7 TeV", CMS PAS PFT-10-002, 2010.

[89] CMS collaboration, "Pileup removal algorithms", CMS PAS JME-14-001, 2014.

[90] M. Cacciari and G. P. Salam, "Pileup subtraction using jet areas", *Phys. Lett. B* **659** (2008) 119–126, doi:10.1016/j.physletb.2007.09.077, arXiv:0707.1378.

[91] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, "Longitudinally-invariant $k_\perp$ clustering algorithms for hadron-hadron collisions", *Nucl. Phys. B* **406** (1993) 187–224, doi:10.1016/0550-3213(93)90166-M.

[92] S. D. Ellis and D. E. Soper, "Successive combination jet algorithm for hadron collisions", *Phys. Rev. D* **48** (1993) 3160–3166, doi:10.1103/PhysRevD.48.3160, arXiv:hep-ph/9305266.

[93] CMS collaboration, "Single muon efficiencies in 2012 data", Technical Report CMS-DP-2013-009, 2013.

[94] CMS collaboration, "Search for neutral Higgs bosons decaying to τ pairs in pp collisions at $\sqrt{s} = 7\,\text{TeV}$", *Phys. Lett. B* **713** (2012) 68–90, doi:10.1016/j.physletb.2012.05.028, arXiv:1202.4083.

[95] CMS collaboration, "Electron efficiency measurement for top quark physics at $\sqrt{s} = 8\,\text{TeV}$", CMS Analysis Note AN-12-429, 2012.

[96] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[97] CMS collaboration, "Performance of the particle-flow jet identification criteria using proton-proton collisions at $\sqrt{s} = 8\,\text{TeV}$", CMS Analysis Note AN-11-330, 2011.

[98] CMS collaboration, "Jet energy scale and resolution in the CMS experiment", *to be submitted to JINST* (2015). CMS-JME-13-004, in preparation.

[99] CMS collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS", *JINST* **6** (2011) P11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.

[100] CMS collaboration, "Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s} = 8\,\text{TeV}$", *JINST* **10** (2015), no. 02, P02006, doi:10.1088/1748-0221/10/02/P02006, arXiv:1411.0511.

[101] CMS collaboration, "Identification of b-quark jets with the CMS experiment", *JINST* **8** (2013) P04013, doi:10.1088/1748-0221/8/04/P04013, arXiv:1211.4462.

[102] T. Chwalek, "Measurement of W-boson helicity fractions in top quark decays with the CDF II experiment and prospects for an early ttbar cross-section measurement with the CMS experiment". PhD thesis, Karlsruhe U., EKP, 2010.

[103] T. Stelzer, Z. Sullivan, and S. Willenbrock, "Single top quark production at hadron colliders", *Phys. Rev. D* **58** (1998) 094021, doi:10.1103/PhysRevD.58.094021, arXiv:hep-ph/9807340.

[104] V. N. Gribov and L. N. Lipatov, "Deep inelastic ep scattering in perturbation theory", *Sov. J. Nucl. Phys.* **15** (1972) 438–450.

[105] Y. L. Dokshitzer, "Calculation of the structure functions for deep inelastic scattering and e^+e^- annihilation by perturbation theory in quantum chromodynamics", *Sov. Phys. JETP* **46** (1977) 641–653.

[106] G. Altarelli and G. Parisi, "Asymptotic freedom in parton language", *Nucl. Phys. B* **126** (1977) 298, doi:10.1016/0550-3213(77)90384-4.

[107] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, "Parton fragmentation and string dynamics", *Phys. Rept.* **97** (1983) 31–145, doi:10.1016/0370-1573(83)90080-7.

[108] B. Andersson, "The Lund model". Cambridge University Press, 1997.

[109] GEANT4 collaboration, "GEANT4: A simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250–303, doi:10.1016/S0168-9002(03)01368-8.

[110] GEANT4 collaboration, "GEANT4 developments and applications", *IEEE Trans. Nucl. Sci.* **53** (2006) 270, doi:10.1109/TNS.2006.869826.

[111] J. Alwall et al., "MadGraph 5: Going beyond", *JHEP* **06** (2011) 128, doi:10.1007/JHEP06(2011)128, arXiv:1106.0522.

[112] F. Maltoni, G. Ridolfi, and M. Ubiali, "b-initiated processes at the LHC: a reappraisal", *JHEP* **07** (2012) 022, doi:10.1007/JHEP07(2012)022, arXiv:1203.6393. [Erratum: JHEP04(2013)095].

[113] T. Sjöstrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 physics and manual", *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175.

[114] CMS collaboration, "Study of the underlying event at forward rapidity in pp collisions at $\sqrt{s} = 0.9$, 2.76, and 7 TeV", *JHEP* **04** (2013) 072, doi:10.1007/JHEP04(2013)072, arXiv:1302.2394.

[115] N. Davidson et al., "Universal interface of TAUOLA: Technical and physics documentation", *Comput. Phys. Commun.* **183** (2012) 821–843, doi:10.1016/j.cpc.2011.12.009, arXiv:1002.0543.

[116] P. M. Nadolsky et al., "Implications of CTEQ global analysis for collider observables", *Phys. Rev. D* **78** (2008) 013004, doi:10.1103/PhysRevD.78.013004, arXiv:0802.0007.

[117] M. Czakon, P. Fiedler, and A. Mitov, "Total top-quark pair-production cross section at hadron colliders through $\mathcal{O}(\alpha_s^4)$", *Phys. Rev. Lett.* **110** (2013) 252004, doi:10.1103/PhysRevLett.110.252004, arXiv:1303.6254.

[118] J. Alwall et al., "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions", *Eur. Phys. J. C* **53** (2008) 473–500, doi:10.1140/epjc/s10052-007-0490-5, arXiv:0706.2569.

[119] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, "Matching matrix elements and shower evolution for top-quark production in hadronic collisions", *JHEP* **01** (2007) 013, doi:10.1088/1126-6708/2007/01/013, arXiv:hep-ph/0611129.

[120] CMS collaboration, "Search for the associated production of the Higgs boson with a top-quark pair", *JHEP* **09** (2014) 087, doi:10.1007/JHEP09(2014)087, arXiv:1408.1682. [Erratum: JHEP10,106(2014)].

[121] S. Frixione, P. Nason, and C. Oleari, "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.

[122] S. Alioli, P. Nason, C. Oleari, and E. Re, "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.

[123] S. Alioli, P. Nason, C. Oleari, and E. Re, "NLO single-top production matched with shower in POWHEG: s- and t-channel contributions", *JHEP* **09** (2009) 111, doi:10.1088/1126-6708/2009/09/111, arXiv:0907.4076. [Erratum: JHEP02,011(2010)].

[124] E. Re, "Single-top Wt-channel production matched with parton showers using the POWHEG method", *Eur. Phys. J. C* **71** (2011) 1547, doi:10.1140/epjc/s10052-011-1547-z, arXiv:1009.2450.

[125] N. Kidonakis, "Differential and total cross sections for top pair and single top production", in *20th international workshop on deep-inelastic scattering and related subjects (DIS 2012)*, pp. 831–834. 2012. arXiv:1205.3453. doi:10.3204/DESY-PROC-2012-02/251.

[126] K. Melnikov and F. Petriello, "Electroweak gauge boson production at hadron colliders through $\mathcal{O}(\alpha_s^2)$", *Phys. Rev. D* **74** (2006) 114017, doi:10.1103/PhysRevD.74.114017, arXiv:hep-ph/0609070.

[127] J. M. Campbell and R. K. Ellis, "MCFM for the Tevatron and the LHC", *Nucl. Phys. Proc. Suppl.* **205–206** (2010) 10–15, doi:10.1016/j.nuclphysbps.2010.08.011, arXiv:1007.3492.

[128] CMS collaboration, "CMS luminosity based on pixel cluster counting - Summer 2013 update", CMS PAS LUM-13-001, 2013.

[129] CMS collaboration, "Measurements of inclusive W and Z cross sections in pp collisions at $\sqrt{s} = 7\,\text{TeV}$", *JHEP* **01** (2011) 080, doi:10.1007/JHEP01(2011)080, arXiv:1012.2466.

[130] CMS collaboration, "Measurement of differential top-quark pair production cross sections in pp colisions at $\sqrt{s} = 7\,\text{TeV}$", *Eur. Phys. J. C* **73** (2013), no. 3, 2339, doi:10.1140/epjc/s10052-013-2339-4, arXiv:1211.2220.

[131] CMS collaboration, "Measurement of the differential cross section for top quark pair production in pp collisions at $\sqrt{s} = 8\,\text{TeV}$", arXiv:1505.04480. submitted to EPJC.

[132] CMS collaboration, "Performance of b-tagging at $\sqrt{s} = 8\,\text{TeV}$ in multijet, $t\bar{t}$ and boosted topology events", CMS PAS BTV-13-001, 2013.

[133] CMS collaboration, "Pileup jet identification", CMS PAS JME-13-005, 2013.

[134] CMS collaboration, "Tuning of out-of-time pileup simulation with CMS ECAL data", CMS Analysis Note AN-13-330, 2013.

[135] A. Bredenstein, A. Denner, S. Dittmaier, and S. Pozzorini, "NLO QCD corrections to $t\bar{t}b\bar{b}$ production at the LHC: 2. full hadronic results", *JHEP* **03** (2010) 021, doi:10.1007/JHEP03(2010)021, arXiv:1001.4006.

[136] ATLAS collaboration, "Study of heavy-flavor quarks produced in association with top-quark pairs at $\sqrt{s} = 7\,\text{TeV}$ using the ATLAS detector", *Phys. Rev. D* **89** (2014), no. 7, 072012, doi:10.1103/PhysRevD.89.072012, arXiv:1304.6386.

[137] CMS collaboration, "First measurement of the cross section ratio $\sigma(t\bar{t}b\bar{b})/\sigma(t\bar{t}jj)$ in pp collisions as $\sqrt{s} = 7\,\text{TeV}$", CMS PAS TOP-12-024, 2012.

[138] CMS collaboration, "Measurement of the cross section ratio $\sigma_{t\bar{t}b\bar{b}}/\sigma_{t\bar{t}jj}$ in pp collisions at $\sqrt{s} = 8\,\text{TeV}$", *Phys. Lett. B* **746** (2015) 132–153, doi:10.1016/j.physletb.2015.04.060, arXiv:1411.5621.

[139] CMS collaboration, "Data driven methods for QCD background estimation in electroweak muon analysis: the ABCD and the template methods", CMS Analysis Note AN-08-113, 2008.

[140] J. Campbell, R. K. Ellis, and R. Röntsch, "Single top production in association with a Z boson at the LHC", *Phys. Rev. D* **87** (2013) 114006, doi:10.1103/PhysRevD.87.114006, arXiv:1302.3856.

[141] P. C. Bhat, "Multivariate analysis methods in particle physics", *Ann. Rev. Nucl. Part. Sci.* **61** (2011) 281–309, doi:10.1146/annurev.nucl.012809.104427.

[142] S. O. Haykin, "Neural networks and learning machines". Prentice Hall, third edition, 2008.

[143] A. Hocker et al., "TMVA - Toolkit for multivariate data analysis", arXiv:physics/0703039. Homepage: http://tmva.sf.net.

[144] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity", *Bulletin of mathematical biophysics* **5** (1943), no. 4, 115–133, doi:10.1007/BF02478259.

[145] B. Cybenko, "Approximation by superposition of a sigmoidal function", *Mathematics of Control, Signals, and Systems* **2** (1989) 303–314.

[146] K. Funahashi, "On the approximate realization of continuous mapping by neural networks", *Neural Networks* **2** (1989) 183–192.

[147] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators", *Neural Networks* **2** (1989) 359–366.

[148] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors", *Nature* **323** (1986) 533–536.

[149] S.-H. Cha, "Comprehensive survey on distance/similarity measures between probability density functions", *Int. J. Math. Model. Method Appl. Sci.* **1** (2007) 300–307.

[150] D. Krohn, M. D. Schwartz, T. Lin, and W. J. Waalewijn, "Jet charge at the LHC", *Phys. Rev. Lett.* **110** (2013) 212001, doi:10.1103/PhysRevLett.110.212001, arXiv:1209.2421.

[151] ATLAS and CMS collaborations, "Procedure for the LHC Higgs boson search combination in summer 2011", Technical Report ATL-PHYS-PUB-2011-11, CMS-NOTE-2011-005, 2011.

[152] A. L. Read, "Modified frequentist analysis of search results (The CL_s method)", CERN OPEN-2000-205, 2000. doi:10.5170/CERN-2000-005.81.

[153] A. L. Read, "Presentation of search results: The CL_s technique", *J. Phys. G* **28** (2002) 2693–2704, doi:10.1088/0954-3899/28/10/313.

[154] T. Junk, "Confidence level computation for combining searches with small statistics", *Nucl. Instrum. Meth. A* **434** (1999) 435–443, doi:10.1016/S0168-9002(99)00498-2, arXiv:hep-ex/9902006.

[155] ALEPH, DELPHI, L3, OPAL collaborations, "Search for the standard model Higgs boson at LEP", *Phys. Lett. B* **565** (2003) 61–75, doi:10.1016/S0370-2693(03)00614-2, arXiv:hep-ex/0306033.

[156] CDF and D0 collaborations, "Higgs boson studies at the Tevatron", *Phys. Rev. D* **88** (2013), no. 5, 052014, doi:10.1103/PhysRevD.88.052014, arXiv:1303.6346.

[157] R. D. Cousins, "Treatment of nuisance parameters in high energy physics, and possible justifications and improvements in the statistics literature", in *PHYSTAT05: Statistical problems in particle physics, astrophysics and cosmology*, pp. 75–85. Imperial College Press, 2005.

[158] K. S. Cranmer, "Frequentist hypothesis testing with background uncertainty", in *PHYSTAT03: Statistical problems in particle physics, astrophysics and cosmology*. 2003. arXiv:physics/0310108.

[159] K. S. Cranmer, "Statistical challenges for searches for new physics at the LHC", arXiv:physics/0511028.

[160] G. J. Feldman and R. D. Cousins, "A unified approach to the classical statistical analysis of small signals", *Phys. Rev. D* **57** (1998) 3873–3889, doi:10.1103/PhysRevD.57.3873, arXiv:physics/9711021.

[161] G. J. Feldman, "Multiple measurements and parameters in the unified approach", in *Workshop on confidence limits*. Fermilab, 2000.

[162] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J. C* **71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727.

[163] A. Wald, "Tests of statistical hypotheses concerning several parameters when the number of observations is large", *Trans. Amer. Math. Soc.* **54** (1943) 426–482, doi:10.1090/S0002-9947-1943-0012401-3.

[164] G. Petrucciani. Private communication.

[165] L. Moneta et al., "The RooStats project", in *ACAT 2010: 13th international workshop on advanced computing and analysis techniques in physics research*, p. 057. 2010. arXiv:1009.1003. Homepage: https://twiki.cern.ch/twiki/bin/view/RooStats.

[166] A. L. Read, "Linear interpolation of histograms", *Nucl. Instrum. Meth. A* **425** (1999) 357–360, doi:10.1016/S0168-9002(98)01347-3.

[167] M. Baak, S. Gadatsch, R. Harrington, and W. Verkerke, "Interpolation between multi-dimensional histograms using a new non-linear moment morphing method", doi:10.1016/j.nima.2014.10.033, arXiv:1410.7388.

[168] R. J. Barlow and C. Beeston, "Fitting using finite Monte Carlo samples", *Comput. Phys. Commun.* **77** (1993) 219–228, doi:10.1016/0010-4655(93)90005-W.

[169] J. Conway, "Incorporating nuisance parameters in likelihoods for multisource spectra", in *PHYSTAT11: Statistical issues related to discovery claims in search experiments and unfolding*. 2011. arXiv:1103.0354.

[170] F. James and M. Roos, "Minuit: A system for function minimization and analysis of the parameter errors and correlations", *Comput. Phys. Commun.* **10** (1975) 343–367, doi:10.1016/0010-4655(75)90039-9.

[171] N. Faltermann, "Search for Standard Model Higgs boson production in association with a single top quark with the CMS experiment". Diploma thesis, Karlsruhe U., EKP, 2015.

[172] D. Atwood, S. K. Gupta, and A. Soni, "Constraining the flavor changing Higgs couplings to the top-quark at the LHC", *JHEP* **10** (2014) 57, doi:10.1007/JHEP10(2014)057, arXiv:1305.2427.

[173] L. Wu, "Enhancing thj production from top-Higgs FCNC couplings", *JHEP* **02** (2015) 061, doi:10.1007/JHEP02(2015)061, arXiv:1407.6113.