# SEARCH FOR THIRD GENERATION SCALAR LEPTOQUARKS IN THE $\mathrm{t}\tau$ CHANNEL WITH THE CMS DETECTOR

HALIL SAKA

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF PHYSICS
ADVISOR: CHRISTOPHER G. TULLY

JUNE, 2015

# Abstract

A search for pair production of third generation scalar leptoquarks decaying to a top quark and a $\tau$ lepton is presented using 19.5 fb$^{-1}$ of proton-proton collision data at $\sqrt{s} = 8$ TeV collected by the CMS detector at the LHC. The search is performed using events that contain a same-sign muon and hadronically decaying $\tau$ lepton pair, and two or more jets. The observations are found to be consistent with the standard model predictions. Assuming that all leptoquarks decay to a top quark and a $\tau$ lepton, the existence of pair produced, charge $-1/3$, third generation leptoquarks up to a mass of 550 GeV is excluded at 95% confidence level. This result constitutes the first direct limit for leptoquarks decaying into a top quark and a $\tau$ lepton, as well as the first CMS result in the same-sign muon - hadronically decaying $\tau$ lepton final state at 8 TeV. The result is also directly applicable to pair produced bottom squarks decaying via the R-parity violating coupling $\lambda'_{333}$.

# Acknowledgements

« *Time interval is a strange and contradictory matter in the mind. It would be reasonable to suppose that a routine time or an eventless time would seem interminable. It should be so, but it is not. It is the dull eventless times that have no duration whatever. A time splashed with interest, wounded with tragedy, crevassed with joy - that's the time that seems long in the memory. And this is right when you think about it. Eventlessness has no posts to drape duration on. From nothing to nothing is no time at all.* »

– *John Steinbeck, East of Eden*

My time as a graduate student definitely seems long in the memory. And I am indebted to many individuals who, by their support, encouragement, and camaraderie, have helped this period to be eventful, and surely filled with interest.

First and foremost, I would like to thank my advisor, Christopher Tully, who has been a very patient and accommodating mentor throughout the years. His guidance and intuition has always proved to be invaluable, and I have always deemed myself extremely fortunate to be able to collaborate with him on matters of physics and life.

The lion's share of what I managed to acquire along the way is due to my friends and colleagues. I am truly glad to have met the Princeton Physics' incoming year of 2008, and in particular my housemates Guilherme Pimentel, Blake Sherwin, and David McGady, who made life in Jadwin and Princeton better than it would have otherwise been. Friends and colleagues at CERN have always been very helpful and collaborative. They have made the transition to the land of wine-and-cheese a joyful one. I consider myself extremely lucky to have had Andrzej Zuranski and Edward Laird as housemates, as well as the friendship of Xiaohang Quan, Edmund Berry, Benjamin Kreis, Davide Gerbaudo, Joshua Hardenbrook, Michael Mooney, Paul Lujan, Juliana Froggatt, and Tatiana Medvedeva. I will always remain in their debt, both academically and personally.

In a collaboration of thousands, a number of names do deserve special mention. Richard Kellogg has been the single most interesting experimental physicist I have had the pleasure of working with. I can only wish to be so lucky to have his very methodical and insightful approach to problems in CMS as well as in life. Within the CMS HCAL collaboration, I am much obliged to all the members of the HCAL operations teams of 2010-13, and in particular to Dragoslav Lazic, German Martinez, Adriaan Heering, Tullio Grassi, and Andrew Baden for their assistance all along. In the CMS EXOTICA domain, I will always remember the support and contributions of Paolo Rumerio, Ketino Kaadze, and John Paul Chou. I am wholeheartedly thankful to Francesco Santanastasio and James Hirschauer for their time and effort that have guided the work that eventually came to be this thesis.

Last, but not least, I am truly grateful for the sacrifices as well as the unconditional, loving support of my parents, Nazife and Kamil Saka, my grandmother, Ayten Avkan, my brother, Kamuran Saka, and my best friend and collaborator in life, Mine Kanol. Their presence is a gift that I could not possibly do without.

# Introduction

« *Some physicists even speak of sub-particles, of which the nucleons are supposed to be made. I am referring to those hypothetical particles that carry the awkward name of "quarks". Now these quarks may not exist. In some ways I, personally, wish they do not, because if so, the whole story would be repeated over again and the real answer to the questions of "why" and "'what" - why are there elementary particles and what really makes them fundamental - would then be postponed and probably would not be found in my lifetime.*[1] »

*– Victor F. Weisskopf, Director-General of CERN, 1961-66*

The concept of scientific reductionism has led to a deeper understanding of the natural phenomena around us. A particular path of this line of thought culminated in a collective theory of fundamental particles and their interactions, known as the standard model of particle physics, guided by a set of symmetry constraints and expressed in the language of a quantum field theory. The standard model has evolved and successfully withstood the test of time since early 1960s, and all of its predicted constituent particles have now been discovered with the latest experimental observation of a Higgs-like scalar boson.

Notwithstanding its triumphant history, the standard model of particle physics suffers from a few important notorious shortcomings, including but not limited to the omission of gravity, as well as the absence of dark matter and dark energy. There are also issues regarding the standard model parameters needing to assume very specific values in order for theory to be in agreement with empirical data, such as the fine-tuning problem in the Higgs sector, the CP problem in the strong sector, and the quantization of charges. The near-unification of the fundamental forces (gauge-coupling unification) also hints at a single underlying force at some higher energy scale, which manifests itself as the three fundamental gauge interactions we experience today, namely the strong, weak, and electromagnetic interactions.

A prevailing interpretation of the standard model regards it as a low energy limit of a more fundamental and elegant framework. Although the inclusion of gravitational

[1] Excerpt from "Nuclear Structure And Modern Research", *Phys. Today* **20**, 5, 23 (1967).

force might necessitate a whole paradigm shift in terms of the mathematical language and abstraction required to describe the natural phenomena, the grand unified and supersymmetric schemes constitute two of the most attractive extensions of the standard model within the field theoretical framework that could address a number of these aforementioned shortcomings and beyond. Their theoretical appeal is furthered by their predictive power as well as immensely rich phenomenology, which renders them indispensable in shaping the experimental efforts in particle physics.

A peculiar feature that the standard model does not offer an explanation for is the ordering and distinction among its fermions, the leptons and quarks. When each of the six leptons and six quarks are viewed in terms of the weak interaction, three distinct, mass-ordered families of fermions arise, each with a pair of leptons and a pair of quarks. Yet, there is no known force carrier in the standard model to directly mediate a transition between a lepton and a quark. This very suggestive symmetry can be conveniently realized in grand unified theories, via a set of new leptoquark bosons carrying both lepton and baryon quantum numbers, and leptoquark-like interactions also arise in a variety supersymmetric and other models. Thus, searches for leptoquarks or leptoquark-like particles prove to be attractive probes of physics beyond the standard model. Any empirical hint to this effect is guaranteed to lead to many more decades of experimental and theoretical work in order to yield a yet more fundamental description of nature, possibly keeping Weisskopf's questions of "what" and "why" unanswered for a while longer whilst bringing us closer to the actual answers.

A leading, modern day approach of experimental research in particle physics is to scrutinize the debris emerging from high energy hadron collisions in a controlled environment, and perform a statistical analysis in search of new structures and patterns that could lead to a higher-order symmetrization of physical laws. To this end, the Large Hadron Collider (LHC) at the European Organization for Nuclear Research (CERN) has been designed and built to collide protons and heavy ions at the highest energies available, and is instrumented with four major detectors, the CMS, ATLAS, ALICE, and LHCb detectors. The LHC scientific program was launched in 2008, eventually reaching to the new frontier of 7 TeV center-of-mass energy in proton-proton collisions in 2010, and 8 TeV in 2012.

This thesis presents a search for scalar leptoquarks that decay into a top quark and a

$\tau$ lepton, conducted using the proton-proton collision data collected at a center-of-mass energy of 8 TeV by the CMS detector at the CERN LHC during 2012. The search targets events with a same-sign muon and a hadronically decaying $\tau$ lepton final state, and uses a novel implementation of a data-driven technique for the estimation of backgrounds due to misidentified leptons. This constitutes the first CMS search in this final state, as well as the first direct search in this heaviest leptoquark decay channel.

The presentation of this dissertation is organized as follows. Chapter 1 lays out the governing symmetries and principles of the standard model, and how the leptoquarks and leptoquark-like particles fit into some of the proposed extensions. Chapter 2 details an effective leptoquark model, the so-called minimal Buchmüller-Rückl-Wyler framework, which facilitates the analysis of leptoquark models in a standard model respecting fashion. Chapters 3 and 4 discuss the infrastructure used to accelerate and collide the protons, and the detector components as well as the methods and algorithms used to disentangle the resulting particle signatures with an emphasis on those used in this analysis. Chapters 5 motivates and presents the event selection requirements used in the search for the scalar leptoquarks. Chapters 6 describes the composition and properties of the standard model events satisfying the given selection requirements, as well as the techniques used in the estimation of these background contributions. Chapter 7 presents the search results together with the implications on the leptoquark model under consideration. A summary of the findings in the context of leptoquarks and supersymmetric scenarios is provided in Chapter 8.

# Contents

# Chapter 1

# The Standard Model and Beyond

## 1.1 The Standard Model of Particle Physics

The standard model (SM) [1] of particle physics is a quantum field theory that describes the interplay between the elementary particles and the forces of nature via the electromagnetic, weak, and strong interactions, with the notable exception of gravity. This section aims to provide a brief summary and motivation for the structure of the SM, and unless explicitly stated otherwise, the discussion is based on Refs. [2, 3, 4, 5, 6, 7].

After the advent of special relativity and quantum mechanics as successful descriptions of nature, quantum field theory emerged as a reconciliation of the two as a relativistic quantum theory for the description of matter particles at energies comparable to their masses. A quantum field theory is expressed in terms of a *Lagrangian* from which the evolution and interaction of the fields can be deduced. The form of the SM Lagrangian is dictated by a set of fundamental symmetries nature is believed to hold, which can be listed as those pertaining to the principle of relativity and those to the internal degrees of freedom the fields possess. The principle of relativity is built into the theory by requiring invariance under the transformations of the *Poincaré group*, ISO(3,1), the 10 generators of which correspond to those of the translations, rotations and boosts.[1] The subgroup of rotations and boosts is known as the *Lorentz group*, SO(3,1). The study of representations of the Lorentz group reveals that the Lie algebra of SO(3,1) can be decomposed into two SU(2) algebras, $\mathrm{SO}(3,1) \simeq \mathrm{SU}(2) \otimes \mathrm{SU}(2)$, such that distinct irreducible representations can be built as a direct product of two SU(2) groups commonly labelled as $(j_-, j_+)$,

[1] Improper transformations include charge conjugation, parity transformation and time reversal, and the SM is invariant under the combined CPT transformation only.

where the Lorentz rotation generators are given as $\vec{J} = \vec{j}_- + \vec{j}_+$. Thus, the fields in the SM Lagrangian are characterized by their intrinsic angular momentum (spin) quantum numbers, namely the scalar (spin-0), spinor (spin-1/2) and vector (spin-1) fields, corresponding to $(0, 0)$, $(1/2, 0)$ and $(0, 1/2)$, and $(1/2, 1/2)$ representations, respectively. The two spin-1/2 representations refer to the left $(1/2, 0)$ and right $(0, 1/2)$ chiral states which are linked by a parity transformation, and they also constitute helicity eigenstates (i.e. $\hat{p} \cdot \vec{J} = \mp\frac{1}{2}$, where $\hat{p}$ denotes the momentum unit vector) in the massless particle limit.

Given that stability of matter necessitates the fundamental particles composing it to be of spin-1/2 (fermionic) nature [8], the free-particle SM Lagrangian can be written as a sum of Poincaré invariant kinetic terms of the fermionic matter fields. The known matter content of the universe is presented in Table 1.1, where six leptons and six quarks are presented in three generations from left to right in order of increasing mass.

| | | | |
|---|---|---|---|
| Leptons | Electron ($e^-$)<br>Electron neutrino ($\nu_e$) | Muon ($\mu^-$)<br>Muon neutrino ($\nu_\mu$) | Tau ($\tau^-$)<br>Tau neutrino ($\nu_\tau$) |
| Quarks | Up quark ($u^{+\frac{2}{3}}$)<br>Down quark ($d^{-\frac{1}{3}}$) | Charm quark ($c^{+\frac{2}{3}}$)<br>Strange quark ($s^{-\frac{1}{3}}$) | Top quark ($t^{+\frac{2}{3}}$)<br>Bottom quark ($b^{-\frac{1}{3}}$) |

Table 1.1: Fermions of the standard model, where the superscripts denote the electric charge in units of proton charge, $|e|$.

Historically inspired by the approximate SU(2) isospin and SU(3) flavor symmetries, variations in the masses and electrical charges of these fundamental particles can initially be ignored to define certain exact internal symmetries on their respective fields that would result in transformations of these fields among themselves. Imposing the invariance of the Lagrangian under local transformations of these symmetries yields a class of quantum field theories known as *Yang-Mills* theories. The requirement of local gauge invariance of the Lagrangian necessitates the introduction of the covariant derivatives to replace the ordinary derivatives in the kinetic terms for each symmetry group, where a set of vector *gauge fields* is defined for each generator of the symmetry group in order to be able to maintain a meaningful definition of the space-time derivatives while preserving their Lorentz transformation properties. This *gauge principle* then predicts the existence of new spin-1 particles which would mediate transformations within these symmetry groups,

and also introduces new interaction terms to the Lagrangian among these gauge bosons themselves.

Guided by a century of experimental observations, a particular form of this picture proves to reproduce the full spectrum of known fundamental interactions. Namely, SM is the Yang-Mills theory of $\mathrm{SU(3)} \otimes \mathrm{SU(2)} \otimes \mathrm{U(1)}$ gauge group, where all fundamental particles are assigned to particular representations of these groups as given in Table 1.2. The left-chirality leptons and quarks are grouped into SU(2) weak doublets, whereas quarks of each flavor (u, d, c, s, t, b) are given as SU(3) color triplets (red, green, blue). The U(1) hypercharge assignments are chosen to reproduce the electrical charges of particles. Particle states represented as singlets under a given symmetry group remain unchanged by the associated transformations, thus, do not partake in the corresponding interactions. The parity violating nature of weak interactions, and the absence of leptons in the strong (color) interactions are therefore built into the theory by this choice of representation assignments. Recalling that the U(N) and SU(N) groups have $\mathrm{N}^2$ and $\mathrm{N}^2 - 1$ generators, respectively, the SM symmetry group implies the existence of 12 gauge bosons which are the photon, $\mathrm{Z}^0$ and $\mathrm{W}^{\pm}$ bosons, and the 8 gluons.

| SM fermions | $\ell_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}$ | $q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix}$ | $\nu_R$ | $e_R$ | $u_R$ | $d_R$ |
|---|---|---|---|---|---|---|
| $\mathrm{SU(3)}_C \otimes \mathrm{SU(2)}_L \otimes \mathrm{U(1)}_Y$ representation | (**1**,**2**,-1) | (**3**,**2**,1/3) | (**1**,**1**,0) | (**1**,**1**,-2) | (**3**,**1**,4/3) | (**3**,**1**,-2/3) |

Table 1.2: Representations of SM fermions. The charged and neutral leptons are denoted as $e$ and $\nu$, respectively, whereas the up-type (u, c, t) and down-type (d, s, b) quarks are represented by the generic labels, $u$ and $d$, respectively. The $q_L$ and $\ell_L$ labels denote the doublet representations of SU(2). Subscripts $L$ and $R$ refer to the left- and right-chiral representations of the Lorentz group as discussed above. Particles are labelled by the dimensions of the corresponding SU(3) and SU(2) representations, as well as the hypercharge, Y. The electromagnetic charge, $Q_{\mathrm{EM}}$, is given as $Q_{\mathrm{EM}} = \mathrm{T}^3 + \frac{\mathrm{Y}}{2}$ where $\mathrm{T}^3$ is the third eigenvalue component of the SU(2) representation. The gauge singlet right-chirality neutrinos may be considered absent in the SM in its simplest form [9].

Since the combination of Lorentz and gauge invariance requirements prevents the addition of explicit mass terms to fermions and gauge bosons, such a theory would contradict the experimental observations in a two-fold way: the massless and degenerate fermion states in the same symmetry group are not massless and degenerate in nature, and not all forces have an infinite range as would be implied by massless force carriers, the gauge bosons. In the SM, these problems are circumvented by the introduction of

a new doublet of complex scalar field belonging to the (**1**,**2**,1) representation, the Higgs field, which enables the construction of new *Yukawa interaction* (spinor-spinor-scalar) terms as well as the Higgs field kinetic term to the Lagrangian obeying the Lorentz and gauge symmetries. By construction, the Higgs potential has a non-zero minimum value. Thus the Higgs doublet field acquires a non-zero vacuum expectation value via the *spontaneous symmetry breaking* (SSB) mechanism, turning the Yukawa terms into fermionic mass terms, whereas the covariant derivatives in the Higgs field kinetic term give rise to the masses of $\mathrm{Z}^0$ and $\mathrm{W}^{\pm}$ gauge bosons. This so-called *Higgs mechanism* also predicts the existence of a scalar, charge neutral particle, the Higgs boson. Thus the electroweak symmetry group $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ is broken into the $\mathrm{U}(1)_{\mathrm{EM}}$ of quantum electrodynamics (QED) and the Higgs SSB reduces the SM gauge symmetry group to $\mathrm{SU}(3)_C \otimes \mathrm{U}(1)_{\mathrm{EM}}$. The interaction corresponding to the symmetry group $\mathrm{SU}(3)_C$, quantum chromodynamics (QCD), remains intact after the Higgs SSB. The discovery of a new boson with a mass of about 125 GeV was announced by the ATLAS and CMS experiments on July 4, 2012 [10, 11], and studies thus far reveal the new particle to be consistent with a SM Higgs boson [12, 13].

Although the fundamental principles of Poincaré and gauge invariance together with the Higgs mechanism dictate the structure of the SM Lagrangian to a great extent, another condition, *renormalizability*, also limits the forms of allowed terms. Renormalization is the systematic procedure of removing the divergences due to loop-corrections resulting in a field theory, and it is a desirable property since it ensures that predictions are independent of the energy scales involved without introducing a cutoff scale above which the theory at hand looses its validity. Working in the so-called natural units ($\hbar = c = 1$) and 4 space-time dimensions, it can be shown that Lagrangians with interaction terms whose coupling constants have a negative mass dimension are non-renormalizable. Since the Lagrangian density has a mass dimension of 4, the scalar and vector fields have a mass dimension of 1, and the spinor fields have a mass dimension of 3/2, the requirement of renormalizability explains the absence of 4-point fermion interaction terms as well as higher order terms in the SM Lagrangian. Since the SM predictions are observed to be sufficient up to the experimentally probed energy scales, there is no need to introduce additional non-renormalizable interactions. If the SM is seen as an *effective field theory*, a low energy limit of a more fundamental theory, the effects of

non-renormalizable interactions are expected to hint at new physics beyond the SM as higher energy regimes are reached, and the LHC provides one such environment.

## 1.2 Looking Beyond

After decades of precision tests, the SM emerges as a highly predictive and successful theory [14], albeit with some reservations [15]. A number of SM parameters should assume very specific values or patterns, unmotivated by theory but necessitated by observation, which may hint at a more fundamental underlying structure. Examples of these are the gauge hierarchy (fine-tuning) problem arising from the low Higgs boson mass and existence of fermion families with specific lepton-quark electric charge relations (and their quantization) needed for anomaly cancellations. Moreover, contrary to empirical evidence, the SM fails to deliver a viable dark matter candidate, cannot account for the matter-antimatter asymmetry in the universe, and in its simplest form, does not accommodate the existence of massive neutrinos. Grand unified theories (GUT) [16], supersymmetric models (SUSY) [17, 18, 19], technicolor (TC) [20], and various other SM extensions aim to address different subsets of these issues. Although absent in the SM, new scalar or vector particles that can couple to both leptons and quarks, leptoquarks (LQs), naturally arise in models where leptons and quarks are treated on the same footing [21].

In GUTs, vector LQs arise among the gauge bosons of the proposed unification symmetry group which mediate transitions between leptons and quarks (such as SU(4), SU(5), SU(10), SU(15), and $E_6$ models), whereas scalar LQ-like states emerge in the symmetry breaking sector as color triplet states alongside with the SM Higgs doublet. Reference [22] provides an illustration within the SU(5) framework. In order to avoid strict experimental bounds on flavor-changing-neutral-currents (FCNCs) and proton decay, the simplest unification models such as Pati-Salam SU(4) [23] and Georgi-Glashow SU(5) [24] necessitate the masses of the vector LQs to be around the unification scale, $\mathcal{O}(10^{14}$ - $10^{16})$ GeV [25]. In the scalar sector, the predictions, and hence the requirements, on the particle content are highly model-specific. In those with possible non-zero couplings to bi-quark states, scalar LQs suffer strict constraints due to proton decay limits [26], whereas other unification models such as SU(15) [27] and extended SU(5) [28] could favor the existence of these new scalar particles around the TeV scale. Moreover,

since scalar LQ states reside in the same GUT multiplet as the Higgs doublet, it could also be plausible to expect these to have Higgs-like enhanced Yukawa couplings to the fermions of the third generation [29].

In SUSY models, particularly the Minimal Supersymmetric Standard Model (MSSM), R-parity ($R_p = (-1)^{3(\mathrm{B-L})+2\mathrm{S}}$, where B, L, and S denote the baryon, lepton, and spin quantum numbers, respectively) is usually treated as an exact multiplicative symmetry such that the $R_p$-odd supersymmetric particles cannot purely decay into their $R_p$-even SM counterparts, and thus the lightest supersymmetric particle emerges as a stable dark-matter candidate [30]. However, extending the MSSM to include all renormalizable interactions terms, it is possible to have $R_p$ violating (RPV) interaction terms in the so-called RPV SUSY models [31], where Yukawa-type couplings enable scalar fermion (sfermion) two-body decays into pairs of SM fermions via baryon and lepton number violating (BNV and LNV, respectively) processes. Although scenarios with simultaneous sizable contributions from both BNV and LNV terms face stringent limits due to proton decay, it is possible to satisfy experimental constraints whilst keeping sfermions masses accessible by limiting significant RPV contributions to either BNV or LNV interactions. In the RPV SUSY scenarios where LNV interactions dominate the phenomenology, the scalar quarks (squarks) can have interactions identical to those of scalar LQs, and if SUSY naturalness is assumed [32], their masses could be in the $\mathcal{O}(10^2$ - $10^3)$ GeV range.

TC models introduce a new, QCD-like technicolor gauge interaction which becomes strong around the electroweak symmetry breaking energy scale, and techni-charged fermion-antifermion pairs form bound states acting as a composite scalar Higgs-like state which gives mass to the SM gauge bosons [21]. In extended technicolor (ETC) theories, additional techni-fermion families and gauge interactions between SM-fermions and techni-fermions are proposed; and if the fermion-technifermion symmetry is assumed to be broken at very high energies, ETC models predict effective four-fermion interaction terms which produce the fermion mass terms after the formation of technimeson condensates. Furthermore, ETC theories may predict LQ-like technimeson states consisting of techniquark and antitechni-lepton pairs, which can decay into ordinary lepton-quark pairs via the ETC gauge boson. Compositeness/substructure models [33] also introduce similar bound states of lepton and baryon number carrying constituent particles which can behave as effective LQs.

# Chapter 2

# Leptoquarks in the Standard Model

## 2.1 The mBRW Model

The Buchmüller-Rückl-Wyler (BRW) framework [34] prescribes a model-independent approach to classify and study the production and decay of LQs. Assuming renormalizable and $\mathrm{SU}(3)_C \otimes \mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ gauge invariant interactions with the SM fermions and bosons, LQs can be scalar or vector particles, and each lepton-quark-leptoquark Yukawa coupling is denoted by a dimensionless $\lambda$. Thus, the most general Lagrangian of lepton-quark-leptoquark interactions in the BRW model can be written as the sum of seven scalar ($\mathcal{L}_{\mathrm{S}}$) and seven vector ($\mathcal{L}_{\mathrm{V}}$) couplings preserving the L and B quantum numbers, and is given as [21, 35]:

$$\begin{aligned}\mathcal{L}_{\mathrm{S}} \;=\{\; & (\lambda_L^{S_1}\; \bar{q}_L^c \epsilon \ell_L + \lambda_R^{S_1}\; \bar{u}_R^c e_R)\; S_1^\dagger \;+\; \lambda_R^{\tilde{S}_1}\; \bar{d}_R^c e_R\; \tilde{S}_1^\dagger \;+ \\ & (\lambda_L^{S_2}\; \bar{u}_R \ell_L + \lambda_R^{S_2}\; \bar{q}_L \epsilon e_R)\; S_2^\dagger \;+\; \lambda_L^{\tilde{S}_2}\; \bar{d}_R \ell_L\; \tilde{S}_2^\dagger \;+ \\ & \lambda_L^{S_3}\; \bar{q}_L^c \epsilon \vec{\tau} \ell_L \cdot \vec{S}_3^\dagger \;\} \;+\; h.c.\end{aligned} \tag{2.1a}$$

$$\begin{aligned}\mathcal{L}_{\mathrm{V}} \;=\{\; & (\lambda_L^{V_1}\; \bar{q}_L \gamma_\mu \ell_L + \lambda_R^{V_1}\; \bar{d}_R \gamma_\mu e_R)\; V_1^{\mu\dagger} \;+\; \lambda_R^{\tilde{V}_1}\; \bar{u}_R \gamma_\mu e_R\; \tilde{V}_1^{\mu\dagger} \;+ \\ & (\lambda_L^{V_2}\; \bar{d}_R^c \gamma_\mu \ell_L + \lambda_R^{V_2}\; \bar{q}_L^c \gamma_\mu e_R)\; V_2^{\mu\dagger} \;+\; \lambda_L^{\tilde{V}_2}\; \bar{u}_R^c \gamma_\mu \ell_L\; \tilde{V}_2^{\mu\dagger} \;+ \\ & \lambda_L^{V_3}\; \bar{q}_L^c \epsilon \vec{\tau} \ell_L \cdot \vec{V}_3^\dagger \;\} \;+\; h.c.\end{aligned} \tag{2.1b}$$

where $e_R$, $\ell_L$, $u_R$, and $q_L$ denote the SM leptons and quarks as defined in Table 1.2, $S$ and $V$ denote the scalar and vector LQs, respectively, labelled by the dimension of their $\mathrm{SU}(2)_L$ representation, and Yukawa couplings are labelled by the subscripts L and R referring to the chirality of the corresponding LQ as given in the superscripts. The Dirac gamma matrices are denoted by $\gamma_\mu$ for $\mu = 0, 1, 2, 3$, Pauli matrices are denoted by

$\tau_i$ for $i = 1, 2, 3$, and $\epsilon = i\tau_2$ provides the antisymmetric $\mathrm{SU}(2)_L$ contraction. Fermion spinors are taken to be in the chiral basis such that $\bar{f}_{R,L} = (P_{R,L} f)^\dagger \gamma^0$ and $f^c_L = (P_L f)^c$ where $P_{R,L} = \frac{1 \pm \gamma^5}{2}$ are the chirality projection operators and charge conjugation is defined as $f^c = C f^* = -i\gamma^2 f^*$. $\mathrm{SU}(3)_C$ and $\mathrm{SU}(2)_L$ incides are suppressed for clarity and a sum over all fermion generations is implied for each term. In the most general form, the Yukawa couplings carry two generation indices, $\lambda^{ij}$, such that an $i^{th}$ generation quark is coupled to an $j^{th}$ generation lepton via the LQ vertex.

Although the magnitudes of the Yukawa couplings described in the BRW model are free parameters, certain constraints have to be imposed in order to assure compliance with the current experimental data, forming the minimal BRW model, mBRW [36]. In the presence of sizable contributions from flavor-changing couplings or simultaneous left and right handed couplings, LQs have to satisfy stringent constraints in order not to be in contradiction with certain low energy measurements [21, 37]. Bounds on certain processes, such as $\mathrm{K} \to \mu\mathrm{e}$, suppress LQs to simultaneously couple to first and second generation SM fermions. Similarly, in order to prevent large contributions to the helicity suppressed $\pi^+ \to \mathrm{e}^+\nu_\mathrm{e}$ decays, LQs are required to satisfy $M_\mathrm{LQ} > |\lambda_L \lambda_R|^{1/2}$ 200 TeV, where $M_\mathrm{LQ}$ is the LQ mass. This essentially means that if LQs are to be within the currently accessible mass range, either their couplings have to be both suppressed which could lead to long-lived states, or the LQ couplings have to be chiral. Thus, in the mBRW effective model, LQs are assumed to have chiral couplings with a single generation of SM fermions at a time, and they are labelled as such. Mass degeneracy is assumed within each of the $\mathrm{SU}(2)_L$ and $\mathrm{SU}(3)_C$ multiplets.

It must be noted that the diagonality of LQ Yukawa couplings cannot be completely achieved in realistic models due to the non-trivial CKM matrix [38, 39]. In particular, LQs diagonally coupled to the left-handed up-type quarks will have CKM-suppressed but non-zero, non-diagonal interactions with the down-type quarks and vice-versa. Nevertheless, the diagonality assumption is retained as a simplification in the mBRW framework. The LQ states and the associated quantum numbers as allowed in the mBRW are summarized in Table 2.1 [35, 40].

| Leptoquark | Spin | F | $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ representation | $Q_{EM}$ | Coupling | | Decay mode |
|---|---|---|---|---|---|---|---|
| $S_1$ | 0 | 2 | (**3**,**1**,−2/3) | −1/3 | $\lambda_{L,R}(u, e_{L,R})$, | $-\lambda_L(d, \nu_e)$ | $t\tau^-$, $b\nu$ |
| $\tilde{S}_1$ | 0 | 2 | (**3**,**1**,−8/3) | −4/3 | $\lambda_R(d, e_R)$ | | $b\tau^-$ |
| $S_2$ | 0 | 0 | ($\overline{\mathbf{3}}$,**2**,−7/3) | −2/3<br>−5/3 | $\lambda_L(u, \nu_e)$,<br>$\lambda_{L,R}(u, e_{L,R})$ | $\lambda_R(d, e_R)$ | $\bar{t}\nu$ , $b\tau^+$<br>$\bar{t}\tau^-$ |
| $\tilde{S}_2$ | 0 | 0 | ($\overline{\mathbf{3}}$,**2**,−1/3) | +1/3<br>−2/3 | $\lambda_L(d, \nu_e)$<br>$\lambda_L(d, e_L)$ | | $\bar{b}\nu$<br>$\bar{b}\tau^-$ |
| $S_3$ | 0 | 2 | (**3**,**3**,−2/3) | +2/3<br>−1/3<br>−4/3 | $\sqrt{2}\lambda_L(u, \nu_e)$<br>$-\lambda_L(u, e_L)$,<br>$-\sqrt{2}\lambda_L(d, e_L)$ | $-\lambda_L(d, \nu_e)$ | $t\nu$<br>$t\tau^-$, $b\nu$<br>$b\tau^-$ |
| $V_1$ | 1 | 0 | ($\overline{\mathbf{3}}$,**1**,−4/3) | −2/3 | $\lambda_{L,R}(d, e_{L,R})$, | $\lambda_L(u, \nu_e)$ | $\bar{b}\tau^-$, $\bar{t}\nu$ |
| $\tilde{V}_1$ | 1 | 0 | ($\overline{\mathbf{3}}$,**1**,−10/3) | −5/3 | $\lambda_R(u, e_R)$ | | $\bar{t}\tau^-$ |
| $V_2$ | 1 | 2 | (**3**,**2**,−5/3) | −1/3<br>−4/3 | $\lambda_L(d, \nu_e)$,<br>$\lambda_{L,R}(d, e_{L,R})$ | $\lambda_R(u, e_R)$ | $b\nu$ , $t\tau^-$<br>$b\tau^-$ |
| $\tilde{V}_2$ | 1 | 2 | (**3**,**2**,+1/3) | +2/3<br>−1/3 | $\lambda_L(u, \nu_e)$<br>$\lambda_L(u, e_L)$ | | $t\nu$<br>$t\tau^-$ |
| $V_3$ | 1 | 0 | ($\overline{\mathbf{3}}$,**3**,−4/3) | +1/3<br>−2/3<br>−5/3 | $\sqrt{2}\lambda_L(d, \nu_e)$<br>$-\lambda_L(d, e_L)$,<br>$\sqrt{2}\lambda_L(u, e_L)$ | $\lambda_L(u, \nu_e)$ | $\bar{b}\nu$<br>$\bar{b}\tau^-$, $\bar{t}\nu$<br>$\bar{t}\tau^-$ |

Table 2.1: Scalar and vector LQs as defined in the mBRW model. Representations under the SM gauge group are labelled by the dimensions of $SU(3)_C$ and $SU(2)_L$ representations, and the $U(1)_Y$ hypercharge Y, respectively. Fermion number is defined as $F = 3B + L$, and electric charge, $Q_{EM} = T^3 + Y/2$, is in units of proton charge, $|e|$, where $T^3$ is the third eigenvalue component of the SU(2) representation. Decay mode is provided assuming third generation LQs only.

## 2.2 Production and Decay of Leptoquarks

In addition to the lepton-quark-leptoquark interactions given in Eq. (2.1), the mBRW model also includes interactions among the LQs and the SM gauge bosons. Scalar LQ - gauge boson couplings are completely defined by the SM gauge invariance assumption, whereas vector LQs might have additional undetermined anomalous couplings to gauge bosons which render their studies model dependent. The discussion is limited to scalar LQs with an emphasis on third generation couplings to a top quark in what follows. Discussions on the production of vector LQs at hadron colliders [41] as well as the current direct and indirect limits on scalar and vector leptoquarks [21, 39, 42, 43] can be found elsewhere.

In hadron colliders, LQs can be produced singly via the unknown Yukawa couplings discussed in Section 2.1, or in pairs via the gauge boson couplings (in addition to the Yukawa interactions) such that leptoquark-gluon interactions dominate the pair production cross section through gluon-gluon fusion ($gg \to LQ\overline{LQ}$) or quark-antiquark annihilation ($q\bar{q} \to LQ\overline{LQ}$) processes [44]. The corresponding leading order Feynman diagrams for the pair production processes are given in Fig. 2.1.

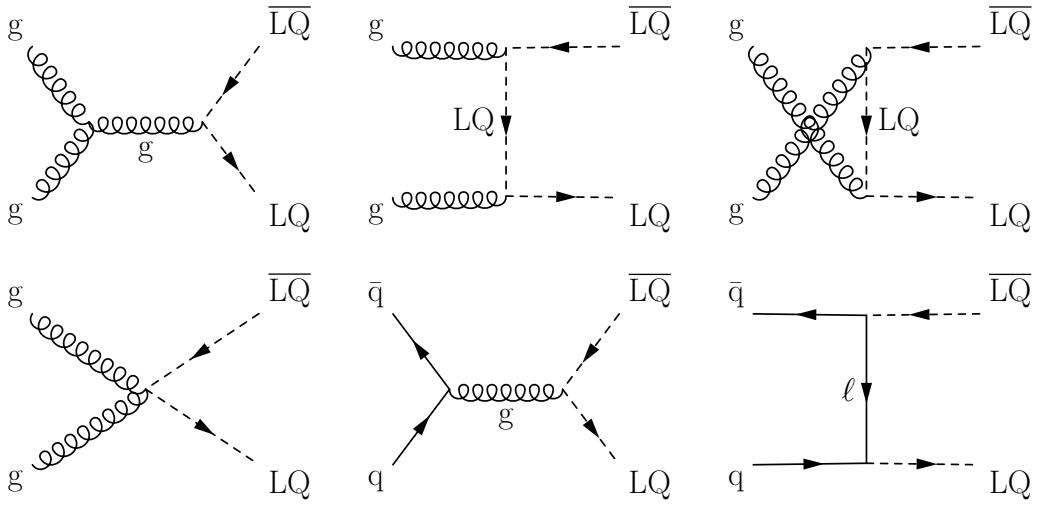


Figure 2.1: Leading order Feynman diagrams for LQ pair production at the LHC. The $gg \to LQ\overline{LQ}$ processes given in the upper diagrams and the lower left diagram as well as the $q\bar{q} \to LQ\overline{LQ}$ process depicted in lower center diagram are entirely determined by the $SU(3)_C$ gauge invariance and are proportional to $\alpha_s^2$. The contribution due to the unknown Yukawa couplings where a quark-antiquark pair interacts through the exchange of a lepton of the same generation is given in the lower right diagram, and is proportional to $\lambda^2$.

Since the top quark content of the proton is negligible at the current LHC ener-

gies [45], contributions depending on the unknown Yukawa couplings as depicted in the lower right diagram in Fig. 2.1 vanish, essentially rendering the pair production cross section of third generation scalar leptoquarks exactly calculable at a given mass hypothesis. The leading order pair production cross sections are expressed as [46, 47]:

$$\sigma^{\mathrm{gg}\to\mathrm{LQ}\overline{\mathrm{LQ}}} = \frac{\pi\alpha_s^2}{96\hat{s}}\left\{\beta(41-31\beta^2) - (17-18\beta^2+\beta^4)\log\left|\frac{1+\beta}{1-\beta}\right|\right\} \tag{2.2a}$$

$$\sigma^{\mathrm{q}\bar{\mathrm{q}}\to\mathrm{LQ}\overline{\mathrm{LQ}}} = \frac{2\pi\alpha_s^2\beta^3}{27\hat{s}} \tag{2.2b}$$

where $\alpha_s$ denotes the strong coupling constant, $\sqrt{\hat{s}}$ is the parton center-of-mass energy, and $\beta = (1-4M_{\mathrm{LQ}}^2/\hat{s})^{1/2}$ is the leptoquark velocity. The next-to-leading-order (NLO) calculation is found to stabilize the pair production cross section with respect to the unphysical factorization/normalization scales; thus NLO cross sections at $\sqrt{s} = 8$ TeV are calculated as prescribed in Ref. [46] and the resulting values for scalar leptoquark masses of 200 - 1000 GeV are presented in Table 2.2.

| $M_{\mathrm{LQ}}$ (GeV) | $\sigma_{\mathrm{NLO}}$ (pb) | $\delta_{\mathrm{PDF}}$ (pb) | $\delta_\mu$ (pb) |
|---|---|---|---|
| 200 | 17.4 | 1.24 | 2.35 |
| 250 | 5.26 | $4.87\cdot10^{-1}$ | $7.00\cdot10^{-1}$ |
| 300 | 1.89 | $2.14\cdot10^{-1}$ | $2.50\cdot10^{-1}$ |
| 350 | $7.69\cdot10^{-1}$ | $1.02\cdot10^{-1}$ | $1.02\cdot10^{-1}$ |
| 400 | $3.42\cdot10^{-1}$ | $5.19\cdot10^{-2}$ | $4.50\cdot10^{-2}$ |
| 450 | $1.63\cdot10^{-1}$ | $2.78\cdot10^{-2}$ | $2.15\cdot10^{-2}$ |
| 500 | $8.21\cdot10^{-2}$ | $1.55\cdot10^{-2}$ | $1.09\cdot10^{-2}$ |
| 550 | $4.31\cdot10^{-2}$ | $8.93\cdot10^{-3}$ | $5.75\cdot10^{-3}$ |
| 600 | $2.35\cdot10^{-2}$ | $5.30\cdot10^{-3}$ | $3.20\cdot10^{-3}$ |
| 650 | $1.32\cdot10^{-2}$ | $3.22\cdot10^{-3}$ | $1.80\cdot10^{-3}$ |
| 700 | $7.61\cdot10^{-3}$ | $2.00\cdot10^{-3}$ | $1.05\cdot10^{-3}$ |
| 750 | $4.48\cdot10^{-3}$ | $1.26\cdot10^{-3}$ | $6.25\cdot10^{-4}$ |
| 800 | $2.69\cdot10^{-3}$ | $8.10\cdot10^{-4}$ | $3.80\cdot10^{-4}$ |
| 850 | $1.64\cdot10^{-3}$ | $5.27\cdot10^{-4}$ | $2.35\cdot10^{-4}$ |
| 900 | $1.01\cdot10^{-3}$ | $3.47\cdot10^{-4}$ | $1.47\cdot10^{-4}$ |
| 950 | $6.34\cdot10^{-4}$ | $2.31\cdot10^{-4}$ | $9.40\cdot10^{-5}$ |
| 1000 | $4.01\cdot10^{-4}$ | $1.55\cdot10^{-4}$ | $6.05\cdot10^{-5}$ |

Table 2.2: LQ pair production cross sections at $\sqrt{s} = 8$ TeV.

For a two-body scalar leptoquark decay to a charged lepton - quark pair, $\ell$q, the tree-level decay width is given as [34]:

$$\Gamma_{\ell,\mathrm{q}} = \frac{\lambda^2 M_{\mathrm{LQ}}}{16\pi} F(m_\ell, m_{\mathrm{q}}, M_{\mathrm{LQ}}) \tag{2.3}$$

where $F(m_\ell, m_{\mathrm{q}}, M_{\mathrm{LQ}}) = (1-d_\ell-d_{\mathrm{q}})\sqrt{(1+d_\ell-d_{\mathrm{q}})^2-4d_\ell}$ is the suppression factor

due to non-zero lepton and quark masses with $d_{\ell,\mathrm{q}} = m^2_{\ell,\mathrm{q}}/M^2_{\mathrm{LQ}}$ [48], and $m_\ell$ and $m_\mathrm{q}$ denote the masses of leptoquark decay products. A similar relation holds for LQ decays to a neutral lepton - quark pair, $\nu\mathrm{q}'$. The suppression factor, $F$, is essentially equal to unity for leptoquarks coupling to first or second generation fermions, whereas it has a nontrivial contribution in the case of a third generation leptoquark coupling to a top quark. The leptoquark branching fractions into $\ell\mathrm{q}$ and $\nu\mathrm{q}'$ are usually denoted as $\beta$ and $1-\beta$, respectively, and $\beta$ is given as [47]:

$$\beta = \frac{\Gamma_{\ell,\mathrm{q}}}{\Gamma_{\ell,\mathrm{q}} + \Gamma_{\nu,\mathrm{q}'}}. \tag{2.4}$$

Third generation leptoquarks with couplings to top quark have the weakest limits from existing low enegry measurements. The most stringent bounds are obtained by studying the radiative corrections to Z physics and constraining any contributions due to scalar leptoquarks [49, 50]. Working in natural units such that the electromagnetic coupling constant is given as $g_e = \sqrt{4\pi\alpha} \sim 0.3$ where $\alpha$ denotes the fine-structure constant, the leptoquark Yukawa couplings are required to satisfy $|\lambda| \lesssim 0.5$ (1) for leptoquark masses of $\sim 200$ (700) GeV.

For leptoquarks masses of $\mathcal{O}(10^2$ - $10^3)$ GeV and $\lambda \sim g_e$, the corresponding decay widths are in the range of 0.2 - 2 GeV, and hence, leptoquark decay products are expected to form narrow resonances in the lepton-quark invariant mass spectrum. The corresponding decay length is given as $d_o = \tau|\vec{p}|/M_{\mathrm{LQ}}$ where $\tau = 1/\Gamma$ denotes the leptoquark lifetime, $\vec{p}$ is the leptoquark momentum vector in the lab frame, and $M_{\mathrm{LQ}}$ is the leptoquark mass. Neglecting suppression factor effects and assuming $|\vec{p}|/M_{\mathrm{LQ}} \sim 10$ at the LHC, it can be calculated that $d_o \leq 10$ $\mu$m for $\lambda \geq 10^{-5}$ resulting in prompt leptoquark decays within the CMS detector, whilst highly suppressed values of the Yukawa couplings could produce long-lived leptoquarks ($d_o \geq 10$ cm for $\lambda \leq 10^{-7}$) that decay at a displaced vertex. If the leptoquark total width is much less than the QCD confinement scale, $\Gamma_{\mathrm{LQ}} \ll \Lambda_{\mathrm{QCD}}$ where $\Lambda_{\mathrm{QCD}} \sim 200$ MeV, leptoquarks may hadronize to form new exotic bound color-singlet states (LQ-$\bar{\mathrm{q}}$, LQ-$\overline{\mathrm{LQ}}$, LQ-q-q) [51, 52, 53] before eventually decaying into lepton-quark pairs. For the purposes of this analysis, leptoquarks are assumed to undergo prompt decays to a lepton-quark pair of the same generation as denoted in Table 2.1.

# Chapter 3

# The LHC and the CMS Detector

## 3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [54, 55] is a superconducting hadron accelerator and collider, built with the aim of exploring the Higgs and beyond the standard model (BSM) physics at the TeV scale. The LHC resides in CERN's former Large Electron-Positron (LEP) Collider tunnel across the French-Swiss border, with a circumference of about 27 km at an approximate depth of 100 m, and serves to four major particle detectors: A Large Ion Collider Experiment (ALICE) [56], A Toroidal LHC Apparatus (ATLAS) [57], Compact Muon Solenoid (CMS) [58] detector, and Large Hadron Collider beauty (LHCb) [59] detector.

Protons, produced from gaseous hydrogen and accelerated through a chain of linear and cyclic accelerators, Linear Accelerator 2 (Linac2), Proton Synchrotron Booster (PSB), Proton Synchrotron (PS), and Super Proton Synchrotron (SPS), to energies of 50 MeV, 1.4 GeV, 25 GeV, and 450 GeV, respectively, are injected into the LHC which are then further accelerated and finally brought into collision at the four interaction points [54]. A schematic of the LHC injector chain is provided in Fig. 3.1 (left). The LHC is also capable of accelerating, storing, and colliding lead ions, and its heavy-ion program is discussed elsewhere [60, 61].

The LHC is not a perfect circle, but is rather composed of eight arcs and eight straight sections labelled as eight octants as given in Fig. 3.1 (right). The arcs are instrumented with twin-bore superconducting dipole magnets operated at $\leq 2$ K using superfluid helium and provide fields up to 9 T for the necessary bending strength. The

straight sections serve as experimental or utility insertions. The four major experiments reside in octants 1, 2, 5, and 8 whereas collimation systems are located in octants 3 and 7. The remaining insertions in octants 3 and 7 house the kicker magnets for beam dump, and the RF systems for each beam line are located in octant 4. The RF system operates at 400.8 MHz, and is instrumental in capturing, squeezing, further accelerating, and storing the injected beams for collisions. The LHC beam pipes are maintained at a pressure of about $10^{-11}$ mbar (equivalent to $10^{13}$ $H_2$ $m^{-3}$) around the room temperature sections near the detectors in order to achieve the target beam lifetime of 100 hours and reduce background to the experiments [54].

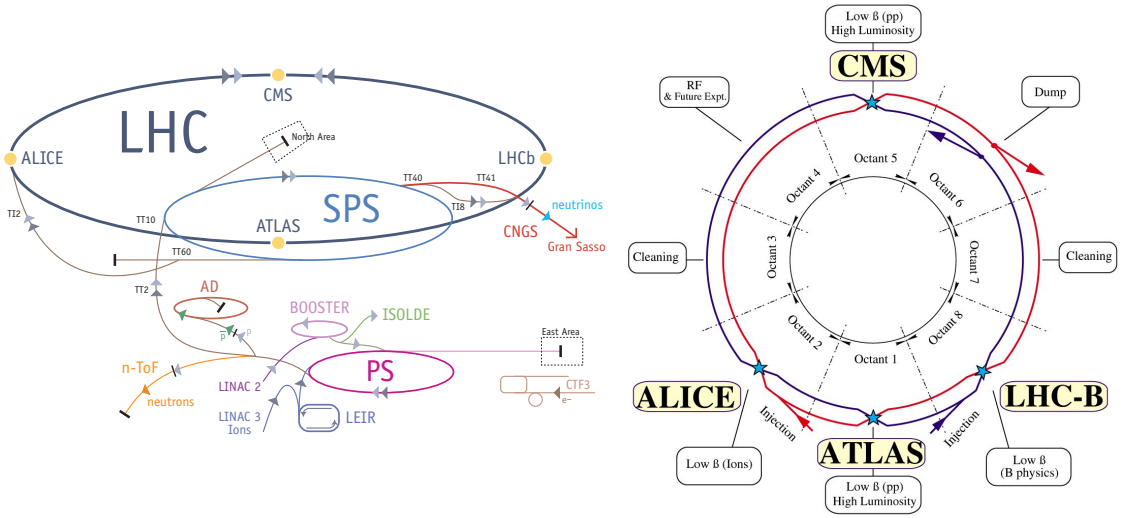


Figure 3.1: Schematic views of the LHC injector chain (left) [62] and the collider ring (right) [54].

The production rate of any given process at a collider, $\frac{dN_p}{dt}$, is directly proportional to the process cross section, $\sigma_p$, and the instantaneous luminosity, $\mathcal{L}$:

$$\frac{dN_p}{dt} = \sigma_p \mathcal{L} \, . \tag{3.1}$$

Assuming round, Gaussian beams, the instantaneous collider luminosity can be expressed as in Eq. (3.2) [54]:

$$\mathcal{L} = \frac{N^2 k_b f}{4\pi \sigma_x^* \sigma_y^*} F = \frac{N^2 k_b f \gamma}{4\pi \epsilon_n \beta^*} F \, , \tag{3.2}$$

where $N$ is the number of particles per bunch, $k_b$ is the number of colliding bunches, $f$ is the revolution frequency, $\gamma$ is the relativistic factor, $\sigma_{x,y}^*$ are the transversal beam sizes in x and y directions, $\epsilon_n$ is the normalized emittance, $\beta^*$ is the value of the beta function, and $F$ is the geometrical reduction factor due to the non-zero beam-beam crossing angle. The superscript asterix denotes values taken at the CMS interaction point.

At ultra-relativistic speeds, particles take about 89 $\mu$s to circulate the LHC, and the ring can nominally accommodate 2808 proton bunches with a spacing of 25 ns. In 2012, the LHC operated with 1374 bunches spaced at 50 ns (1368 colliding bunches at CMS) and each bunch contained up to $1.7 \times 10^{11}$ protons at an energy of 4 TeV. The optimized 2012 proton-proton (pp) collision parameters for $\epsilon_n$, $\beta^*$ and F at CMS are 2.5 $\mu$m, 0.6 m, and 0.8, respectively, yielding a peak instantaneous luminosity of $7.7 \times 10^{33}$ cm$^{-2}$s$^{-1}$ (7.7 nb$^{-1}$s$^{-1}$) [63, 64, 65, 66].

During the 8 TeV run in 2012, the LHC has delivered an integrated luminosity of 23.3 fb$^{-1}$ pp collisions data to the ATLAS and CMS experiments, of which 21.8 fb$^{-1}$ have been recorded by the CMS detector and 19.7 fb$^{-1}$ have been certified for physics analyses [67]. Under the 8 TeV operation conditions, the average number of additional primary vertices in the collision of the two beams in the same proton bunch crossing (pileup) was observed to be 20, with tails extending as high as 40 [68]. The time-evolution of the total integrated delivered and recorded luminosities, as well as the distribution of pileup interactions during the 8 TeV run are illustrated in Figure 3.2.

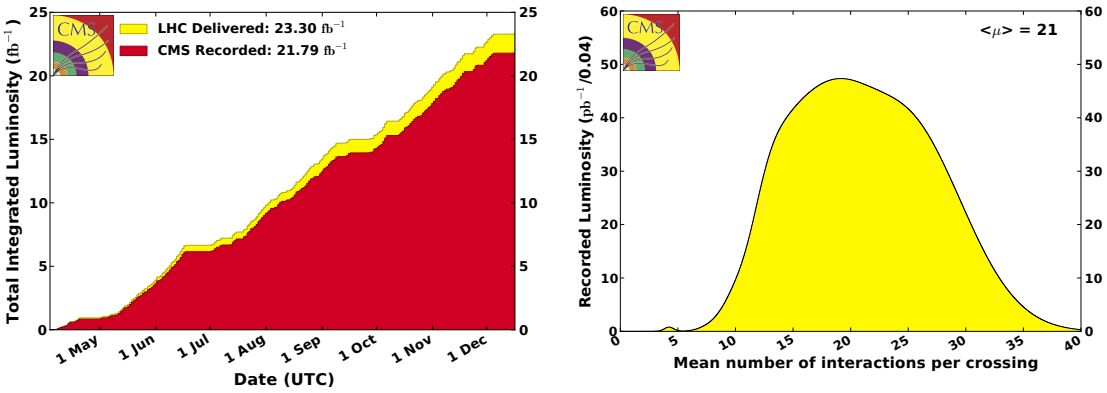


Figure 3.2: Integrated luminosity versus day delivered to (yellow), and recorded by the CMS detector (red) evolution (left), and the distribution of mean number of interactions per bunch crossing (right) during stable beams and for pp collisions at 8 TeV in 2012 [68].

## 3.2 The Compact Muon Solenoid

CMS is one of the two general purpose detectors, alongside with ATLAS, located at the CERN LHC to investigate physics at the TeV scale in pp and heavy-ion collisions. In order to meet the goals of the LHC physics program, CMS is designed to provide good lepton momentum resolution as well as good dielectron, diphoton and dimuon

mass resolution ($\sim 1\%$ at 100 GeV). Efficient $\tau$ lepton and b quark triggering and offline tagging relies on the fine charged particle momentum resolution of the inner tracker, whereas large hermetic coverage of hadron calorimeter enables good missing transverse energy and dijet mass resolution capabilities. The CMS detector is instrumented to fulfill these goals in a high occupancy and high radiation environment [58].

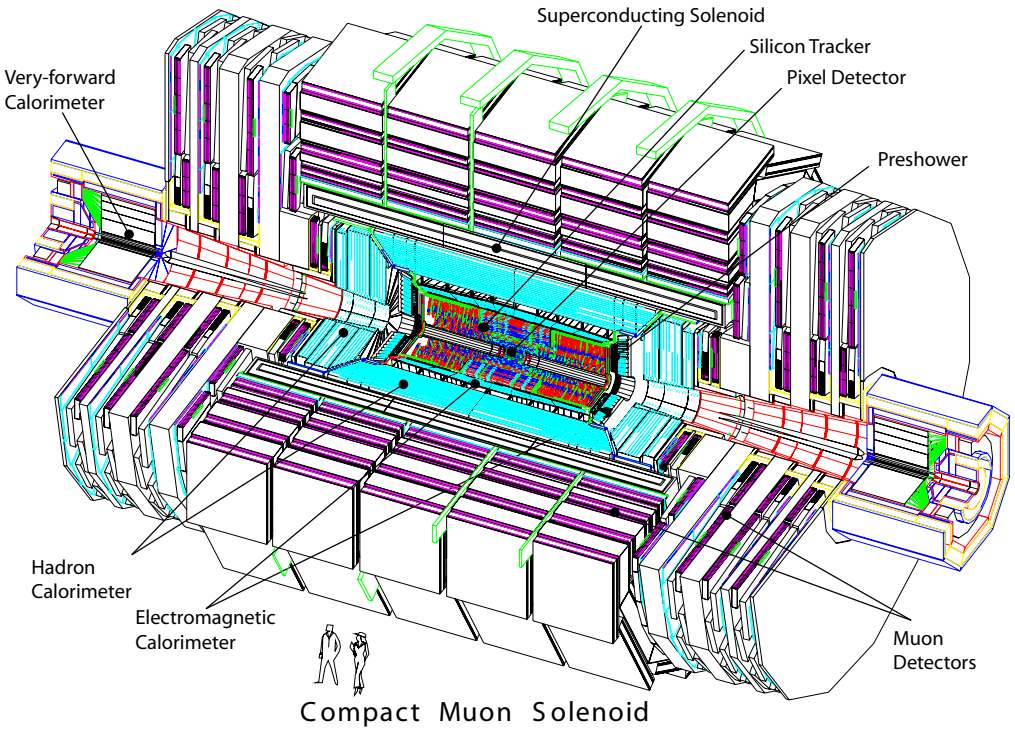


Figure 3.3: Schematic view of the CMS detector [58].

CMS consists of a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter, a brass-scintillator hadron calorimeter, a superconducting solenoid magnet and gas-ionization muon chambers; and thus is able to reconstruct particles in a wide range of transverse momentum $p_{\mathrm{T}}$, and pseudo-rapidity $\eta$, where $\eta = -\ln \tan(\theta/2)$ and $\theta$ is the polar angle measured from the z-axis in a right-handed coordinate system with positive x-axis pointing radially inward towards the center of the LHC and positive z-axis pointing in the counter-clockwise beam direction. The azimuthal angle, $\phi$, is measured in the x-y plane [58]. A schematic view of the CMS detector is illustrated in Fig. 3.3.

### 3.2.1 Superconducting Magnet

The central feature of CMS is a superconducting solenoid magnet [69] which surrounds the inner tracking system as well as the electromagnetic and hadron calorimeters. The magnet stretches approximately 6 m in diameter and 12.5 m in length, and consists of niobium-titanium coils cooled to liquid helium temperatures ($\sim$ 4.5K). In its superconducting state, the solenoid is capable of carrying currents of about 19 kA and producing magnetic fields up to 4 T. The magnet is supplemented by an iron return yoke of 5 barrel wheels and 2 endcaps with 3 disks each, which are instrumented with four stations of the muon system, to confine the magnetic flux to the volume of the detector. Schematic views of the solenoid coil and the yoke are given in Fig. 3.4.

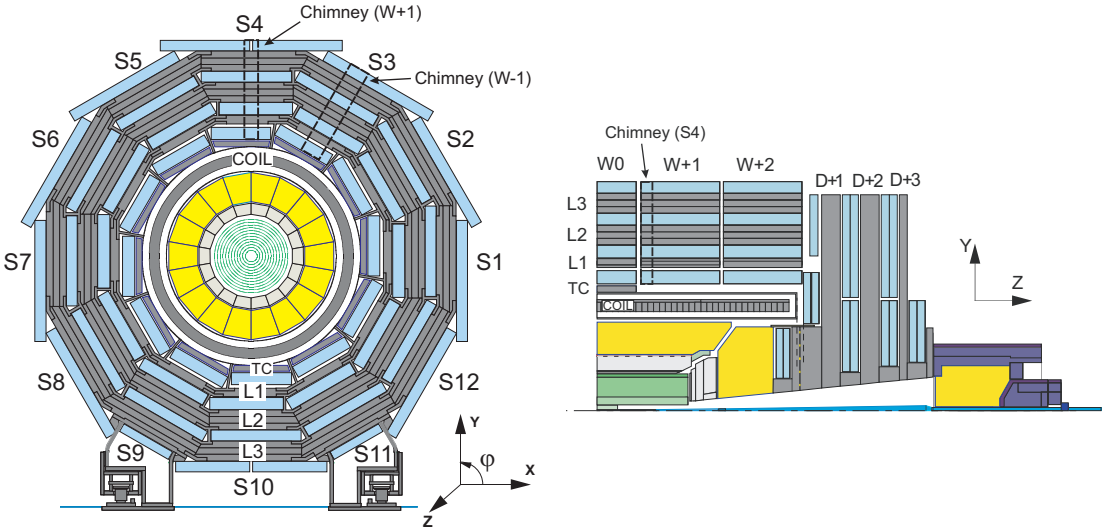


Figure 3.4: Schematic views of the CMS detector with transverse (left) and longitudinal (right) slices at z = 0 [70]. Magnet coil and the yoke layers are in gray; yoke barrel layers are labelled as L1, L2, and L3, whereas endcap disks are labelled as D+1, D+2 and D+3. Muon stations (light blue) are visible in between the yoke layers both in the barrel and endcap regions.

The uniform high magnetic field provided by the solenoid over the entire body of CMS is crucial in achieving the target muon momentum resolution of about 1% at 100 GeV, up to $|\eta| < 2.4$. Despite the design value of 4 T, the central magnetic flux density was reduced to 3.8 T prior to the start of the physics program in order to increase the longevity of the solenoid [70].

### 3.2.2 Tracking System

The tracking system [71, 72] of the CMS detector is built to provide precise and efficient measurement and reconstruction of charged particle trajectories and secondary vertices emerging from pp collisions in a high particle multiplicity and high radiation environment [58]. The tracking system consists of a silicon pixel detector around the interaction point and an outer silicon strip tracker, as illustrated in Fig. 3.5, and is exposed to particle fluxes of about 3 kHz/mm$^2$ at the outer layers of the strip tracker and up to 1 MHz/mm$^2$ at the first layer of the pixel detector at each collision of proton bunches. The choice of silicon sensor technology is driven by the demands of fast response, high granularity tracker channels needed to achieve an occupancy rate of $\lesssim 1\%$, combined with the desired radiation hardness for an expected lifetime of 10 years.

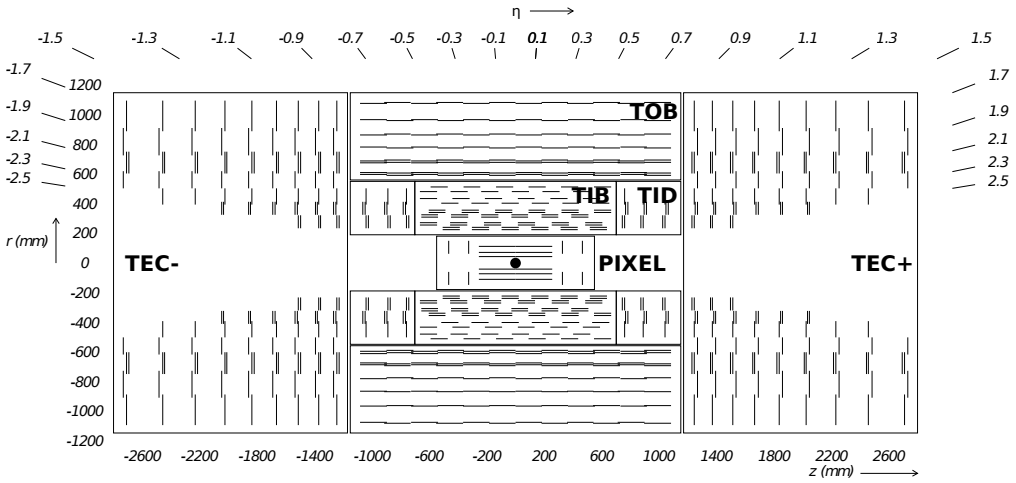


Figure 3.5: Schematic cross section of the CMS inner tracking system [73]. Each line represents a detector module, whereas double lines indicate double-sided modules which deliver stereo hits. The four main subsystems of the silicon strip tracker are the four-layer Tracker Inner Barrel (TIB), the six-layer Tracker Outer Barrel (TOB), the three-disk Tracker Inner Disks (TID) and the nine-disk Tracker End Caps (TEC).

The pixel detector is about 97 cm in length and 30 cm in diameter, consisting of 2 endcap disks on each side and 3 barrel layers the innermost of which is 4.4 cm away from the beam line, and is instrumented with approximately 66 million detector channels [74], whereas the silicon strip tracker stretches 5.5 m in length and 2.4 m in diameter around the pixel detector, with a total of 9.3 million detector channels [73]. With pixel and strip cell sizes of about $100 \times 150$ $\mu$m$^2$ and 10 cm $\times$ 80 $\mu$m (up to 25 cm $\times$ 180 $\mu$m in the outer strip detector), respectively, the CMS inner tracking system achieves transverse and longitudinal impact parameter resolutions of 10 $\mu$m and 20 - 70 $\mu$m, respectively, as

well as a momentum resolution of 1 - 7% for particles with a transverse momentum of 100 GeV and $|\eta| < 2.4$ [75].

Due to its very low occupancy and operational stability, CMS has also utilized the pixel detector in precision luminosity measurement by a pixel cluster counting technique as discussed elsewhere [76].

### 3.2.3 Electromagnetic Calorimeter

The CMS electromagnetic calorimeter (ECAL) [77] is a near hermetic, homogeneous calorimeter aiming accurate detection of electrons and photons, and is composed of 75848 lead tungstate crystals in the barrel and two endcap regions with an additional silicon strip preshower detector in front of the each endcap unit. A schematic transverse view of a quarter of the ECAL is given in Fig. 3.6.

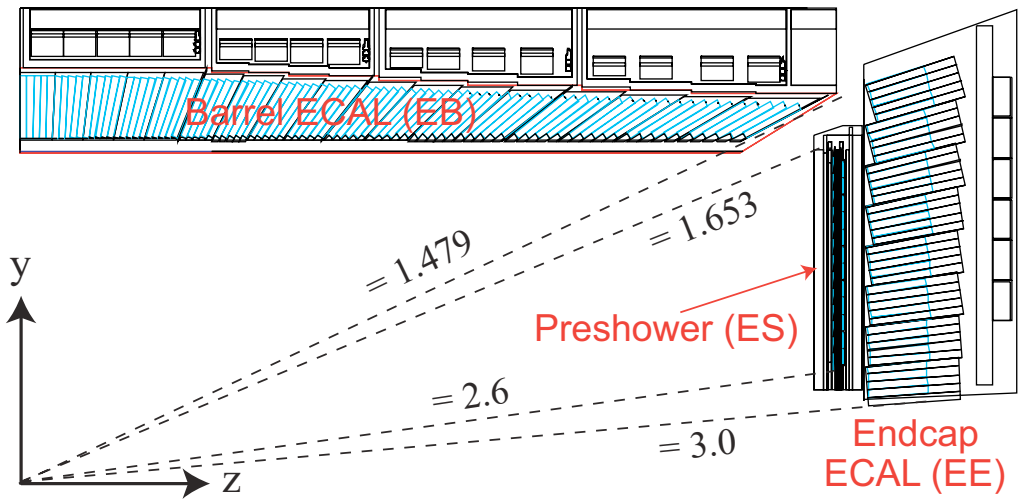


Figure 3.6: Transverse view of the CMS electromagnetic calorimeter [75], where individual crystals in the barrel (EB) and endcap (EE) regions, as well as the silicon detectors in the preshower (ES) unit are denoted in light blue. Detector boundaries are provided in $\eta$.

Crystal front-end faces and lengths measure $2.2\times2.2$ cm$^2$ by 23 cm and $2.86\times2.86$ cm$^2$ by 22 cm, and are instrumented by avalanche photodiodes (APDs) and vacuum phototriodes (VPTs) [78] in the barrel and endcap regions, respectively. Lead tungstate crystals have a radiation length ($X_0$) of 0.89 cm, a Molière radius of 2.2 cm, and a scintillation light decay time of $< 25$ ns, enabling a radiation resistant, high-granularity, fast readout. ECAL crystals yield scintillation light with wavelengths around 425 nm for impinging electrons and photons, which is then collected by the APD and VPT photodetectors.

Both the light yield of crystals and the gain of the photodetectors are highly temperature sensitive, and thus, ECAL is nominally operated at a steady temperature of 18°C, where each APD and VPT collect an average of 4.5 photoelectrons per MeV of energy deposition within the crystal, with a gain of about 50 and 10, respectively [58, 79].

The preshower detector is a sampling calorimeter consisting of 2 layers of lead instrumented with 2 orthogonal layers of silicon strips of 2 mm width, and stretches 20 cm along the z-direction in front of both ECAL endcap units corresponding to about 3 radiation lengths. It is instrumental in increasing the granularity of the ECAL endcap regions, and thus resolving the diphoton hits originating from the $\pi^0 \to \gamma\gamma$ process which could otherwise be misreconstructed as single photon hits in the detector [58].

These instrumentation choices for the ECAL detector, largely driven by the requirement of efficient detection of the two-photon decays of the Higgs boson, result in a barrel ($|\eta| < 1.479$) electron energy resolution as given in Eq. (3.3) [80, 81].

$$\frac{\sigma_{\mathrm{E}}}{\mathrm{E}} = \frac{2.8\%}{\sqrt{\mathrm{E(GeV)}}} \oplus \frac{12\%}{\mathrm{E(GeV)}} \oplus 0.3\% \tag{3.3}$$

Together with the precise ECAL calibration maintained during the $\sqrt{s} = 7$ TeV pp data taking period in 2010 and 2011, the electron energy resolution is reported to be $< 2\%$ for $|\eta| < 0.8$ (2 - 5% for $|\eta| \geq 0.8$) in $\mathrm{Z} \to \mathrm{ee}$ events, and the photon energy resolution in 125 GeV Higgs decays, $\mathrm{H} \to \gamma\gamma$, is found vary in the range of 1.1 - 2.6% for $|\eta| < 0.8$ (2.2 - 5.0% for $|\eta| \geq 0.8$) [82].

### 3.2.4 Hadron Calorimeter

The hadron calorimeter (HCAL) [83] is primarily designed to measure hadron jets as well as neutrinos and other exotic particles which may escape detection resulting in apparent missing transverse energy. HCAL is a hermetic, non-compensating, sampling calorimeter [81, 84] consisting of the barrel (HB), endcap (HE) and forward (HF) detectors where each half detector is composed of 18 identical wedges with an azimuthal extend of 20°, as well as the outside detector (HO) composed of 5 rings and 12 azimuthal sectors outside the CMS magnet solenoid. The longitudinal schematic view of HCAL is illustrated in Fig. 3.7.

HB and HE wedges consist of 17 and 18 active layers interlaid with brass absorber plates (with the exception of steel innermost and outermost layers), respectively, and

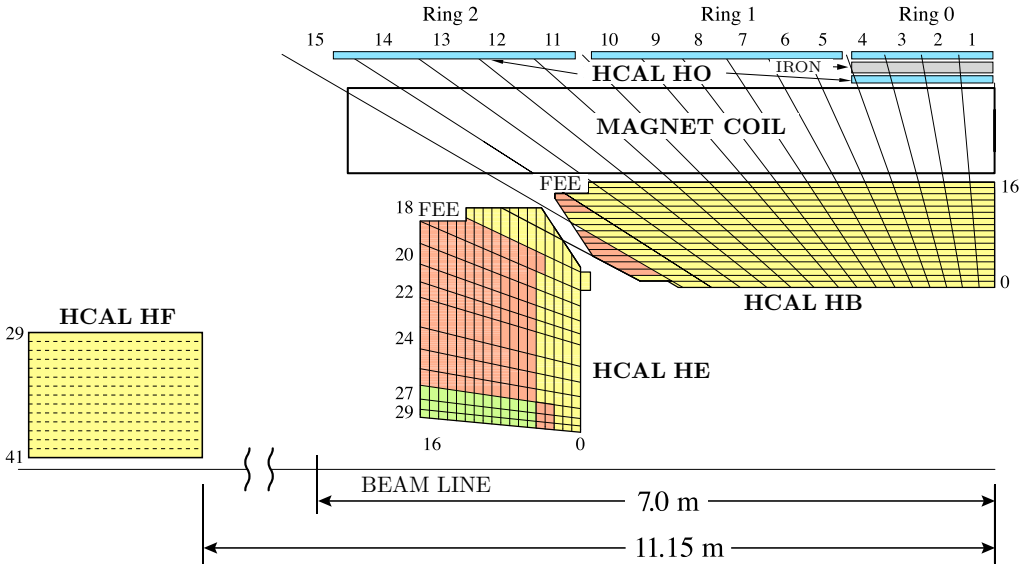


Figure 3.7: Longitudinal view of the CMS hadron calorimeter [85], displaying the hadron barrel (HB, $|\eta| < 1.3$), endcap (HE, $1.3 \leq |\eta| < 3.0$), outer (HO, $|\eta| < 1.3$), and forward (HF, $3.0 \leq |\eta| < 5.0$) detectors, whereas solid lines extending outwards from the interaction point on the right-bottom corner denote HCAL tower segments and the coloring scheme follows the longitudinal segmentation of each HCAL tower. HB and HE are located on the outer side of the tracker and the electromagnetic calorimeter (not pictured), HO sits on the outer side of the magnet solenoid, and HF surrounds the beam line at z = ±11.15 m.

in each active layer, individual scintillator plates are instrumented as a single megatile stretching across the whole longitudinal length of the wedge. Scintillator plates are embedded with wavelength shifting (WLS) fibers carrying the scintillation light out to optical decoding units (ODUs) where optical signals are combined into HCAL towers. Most of the HB detector consists of single towers of 17 layers extending the entire transverse length of the detector, whereas HE towers are read out as 1, 2 or 3 longitudinal channels. Locations of HB wedges and the structure of a single HB wedge are provided in Fig. 3.8.

HB and HE towers have a granularity of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087(5^\circ)$ in the barrel region ($|\eta| < 1.6$) and about $\Delta\eta \times \Delta\phi = 0.17 \times 0.17(9.7^\circ)$ in the endcap ($|\eta| \geq 1.6$). The total HCAL+ECAL material budget (excluding HO and HF) provides a minimum of about 7 interaction lengths ($\lambda_0$) around $\eta = 0$ in the barrel, increasing up to 10 in the endcap. The outer calorimeter complements the HB detector by detecting and measuring any late showering jets, leaking beyond the HB which is at about 4.3% level for 300 GeV pions. HO utilizes the passive body of the CMS magnet as its absorber layer in rings ±1 and ±2, whereas the innermost ring is supplemented by a second iron

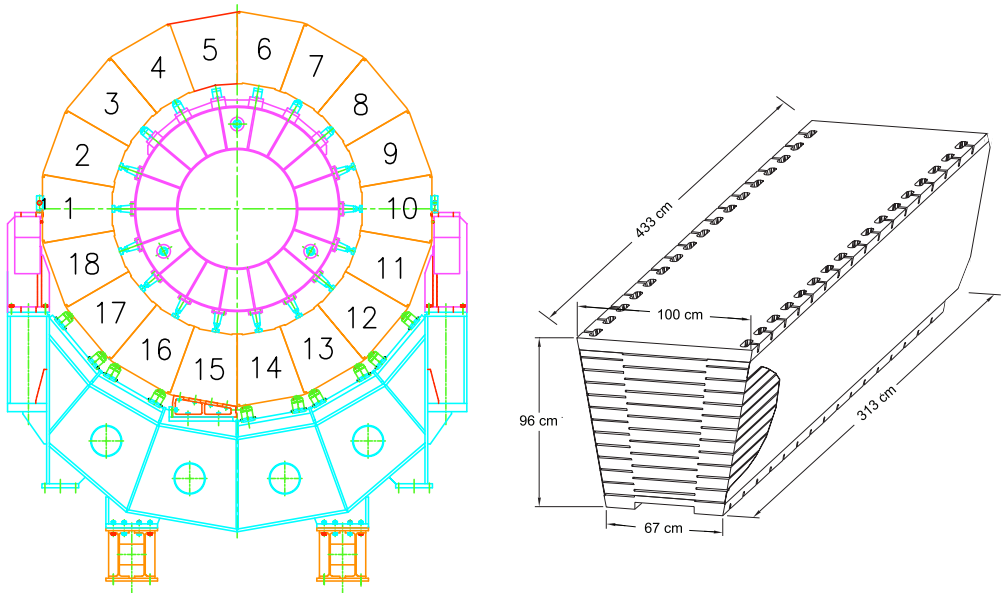


Figure 3.8: Schematic view and numbering of HB wedges in the $+x$ direction (left) [58] and the isometric projection of an HB wedge (right) [86].

absorber layer in between the 2 scintillators. HO increases the total minimum depth of the barrel calorimeter system to 11.8 interaction lengths, with the exception of barrel-endcap boundary region [58].

Optical signals from HB, HE and HO detectors are digitized by the hybrid photodiodes (HPDs) [87] located in read-out boxes installed on the calorimeter, and hence are subject to magnetic fields up to 3.8 T. HPDs consist of a photocathode kept at a bias voltage of 8 - 10 kV relative to the silicon photodiode located 3.3 mm underneath, and have 19 hexagonal pixels, 18 of which are instrumented to receive optical signals coming from HCAL tower segments. HPDs are magnetic field resistant, provide a gain of 2000 with a quantum efficiency of 20% and a response time of 100 ps [88]. Schematic views of megatile trays with the scintillator plates and WLS fibers, as well as the structure of the HPD are presented in Fig. 3.9.

The forward calorimeter is a cylindrical structure with steel absorber layers around the beam line with an outer radius and length of 130 cm and 165 cm (about 10 interaction lengths), respectively, and sits on both sides of the interaction point at a distance of 11.15 m. Unlike HB, HE, and HO, HF is subject to extremely high particle fluxes and radiation levels, which drives the choice of radiation-hard quartz fibers as the active component that yield Cherenkov light when impinged by particles, and hence it is mostly sensitive to the electromagnetic component of the particle showers, and largely insensitive

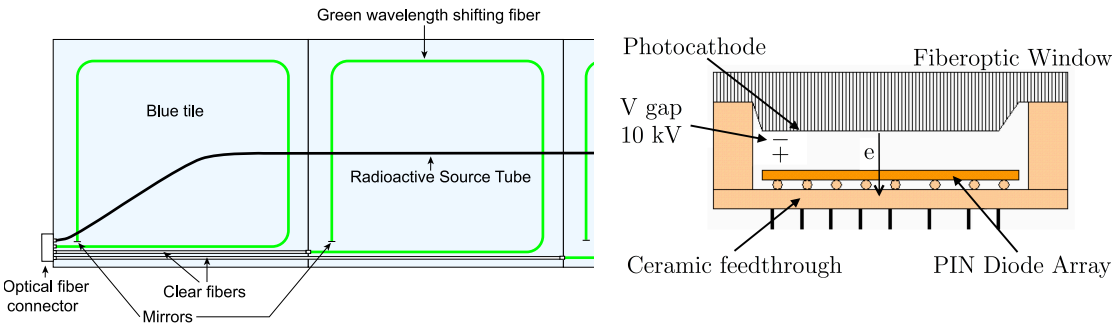


Figure 3.9: Closeup view of a typical megatile (2 of 16 scintillator plates are fully visible) (left) [86], and schematic view of the HPD structure (right) reproduced from Refs. [89, 90].

to neutrons and low energy particles emerging from decays of activated radionucleids [58]. As it is located outside the CMS magnet, traditional photomultiplier tubes (PMTs) are used to convert the light yield into electrical signals, which operate at a bias voltage of 2 kV and provide quantum efficiency and timing values similar to those of the HPDs while delivering a higher gain of about $10^6$ [88].

Performance of the CMS calorimetry system has been tested in dedicated test beam studies. Using a function of the form $\sigma_E/E = a/\sqrt{E(GeV)} \oplus b$, the calorimetry resolution parameters, [a, b], for the combined HCAL and ECAL system are found to be about [110.7%, 7.3%], [84.7%, 7.4%] when corrected for the non-compensating detector response, in the combined barrel and endcap region [85, 91]. Inclusion of the outer calorimeter reduces the constant b-term to 6.6% in the HO ring-1 region [92]. In the forward region, these parameters are measured as [198%, 9%] and [280%, 11%] for the electromagnetic and hadronic components, respectively [93].

### 3.2.5 Muon System

The CMS muon system [94] aims to provide efficient muon identification as well as good momentum measurement and trigger capabilities with the aid of the high-field magnet and the flux return yoke. The muon system is composed of three different particle detectors which rely on gas ionization for particle detection, the drift tubes (DTs), the resistive plate chambers (RPCs), and the cathode strip chambers (CSCs), interspersed among the magnet yoke layers in the barrel and endcap regions. Figure 3.10 illustrates the location and pseudo-rapidity coverage of the three different muon detectors in the CMS.

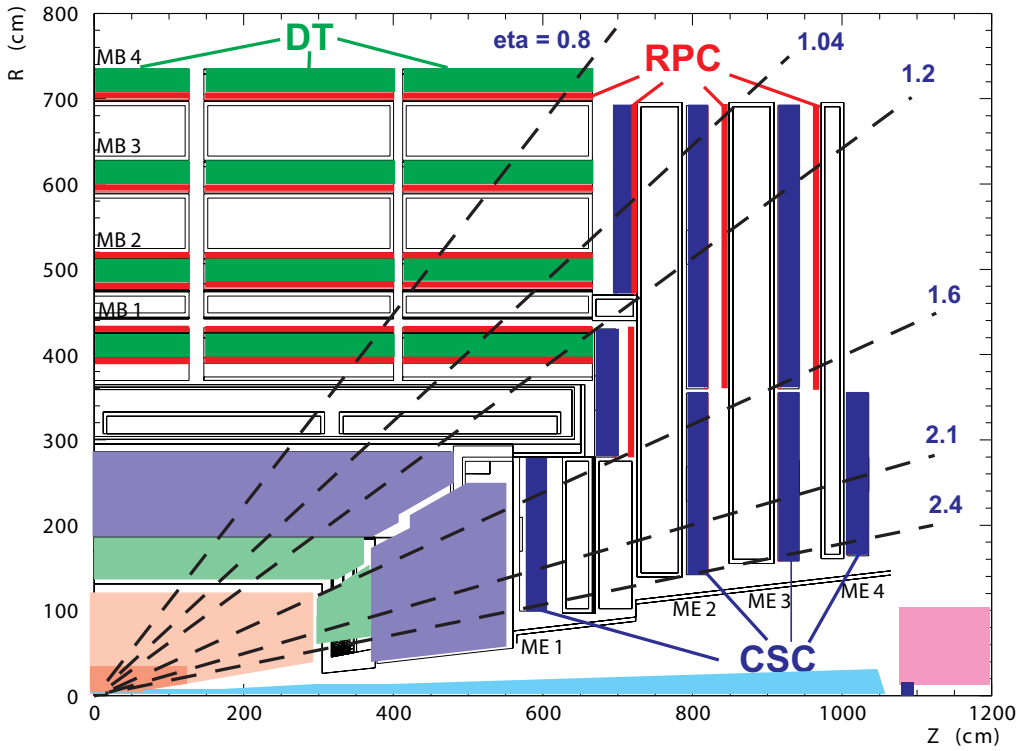


Figure 3.10: Transverse view of the CMS muon system [75].

The drift tube system is composed of 4 layers of concentric cylinders in the $|\eta| < 1.2$ region, and consists of a total of 250 DT chambers, each of which measures approximately 3 m $\times$ 2.5 m. Each DT chamber is composed of 12 layers of drift cells, organized as 3 super-layers (SL) with 4 drift cell layers per SL. There are 2 SLs in the $\phi$-direction, and 1 SL in the z-direction. Each drift cell extends the entire length of the chamber with a transverse size of $42 \times 13$ mm$^2$ and corresponds to an active detector channel, yielding a total of 172000 DT channels. A drift cell operates with an anode wire kept at +3.6 kV which runs through the gaseous internal volume of argon and carbon dioxide, whereas the surrounding strips and cathodes are kept at +1.8 kV and −1.8 kV, respectively, resulting in a gain of $10^5$. The DT chambers provide a spatial resolution of 77 - 123 $\mu$m in r-$\phi$ and 140 - 393 $\mu$m in r-z planes [58, 75, 95].

The CSC system covers $0.9 < |\eta| < 2.4$ in the two endcap regions, overlapping with the DT system. Individual cathode strip chambers are trapezoidal multiwire proportional chambers made up of six gas gap layers of mostly argon and carbon dioxide, and extend 1.7 or 3.4 m in the radial direction, covering an angle of 10° or 20° degrees. Each chamber layer is instrumented with up to 80 radially outward cathode strips subtending an angle of 2.2 - 4.7 mrad in the $\phi$ direction, as well as anode wires running around the

beam line at a fixed radius with $\sim$ 3.2 mm spacing. The anode wires are read out as bundles covering approximately 50 mm widths. The entire CSC system consists of 220000 strip and 183000 wire bundle channels, providing a combined 6-layer r-$\phi$ resolution of $\sim$ 80 $\mu$m. At the operation voltage of 3.6 kV, CSC chambers yield a gain of about $7 \times 10^4$ [58, 75, 95].

The RPCs are resistive gaseous parallel plate detectors accompanying all four DT layers in the barrel and three of the CSC layers in the endcap regions with $|\eta| < 1.6$. Each layer of a basic double gap RPC module is formed by two bakelite resistive plates coated with conductive graphite and insulating foil on the outside. The resistive plates are separated by a 2 mm gap of mostly freon gas, and operate at 9.6 kV in avalanche mode. Each of these double gap modules are read out with common pick-up strips sandwiched by the gas gap layers, and there are up to 96 strips per module. RPC strips lie along the beam axis in the barrel region, and radially in the endcap region. The RPC system has 480 barrel and 432 endcap double-gap chambers, yielding a total of 109000 channels. Although the RPC system provides a relatively coarse spatial resolution of 0.8 - 1.3 cm, it has an extremely fast response time of $\lesssim$ 3 ns, and thus is crucial in the triggering and bunch crossing identification of the muon tracks [58, 75, 95].

The CMS muon system achieves a $p_{\mathrm{T}}$ resolution of $\lesssim$ 10% and 10 - 15% in the barrel and endcap regions, respectively, for muons with 15 GeV $< p_{\mathrm{T}} <$ 100 GeV and $|\eta| < 2.4$. The corresponding muon-system-only trigger efficiencies are found to be $> 90\%$ [96].

### 3.2.6 Trigger System

At the center-of-mass energy of 8 TeV, the total inelastic cross section for pp collisions is around 74 mb [97, 98, 99], yielding an average of about 25 interactions within the CMS detector per each bunch crossing at the peak instantaneous luminosities in 2012. In comparison, the expected event yields for scalar leptoquarks at 8 TeV are suppressed by about 12 orders of magnitude, whereas various SM and BSM processes are suppressed by 6-to-15 orders of magnitude at the LHC design center-of-mass energy of 14 TeV as illustrated in Figure 3.11. Moreover, since each CMS event can be up to around 1 MB in size, recording each bunch crossing with 50 (25) ns intervals yields a currently infeasible data production rate of 20 (40) TB per second [100]. In order to cope with the data acquisition, storage, and processing issues, CMS employs a two-tier trigger system to

select the interesting hard-collision processes at a manageable rate.

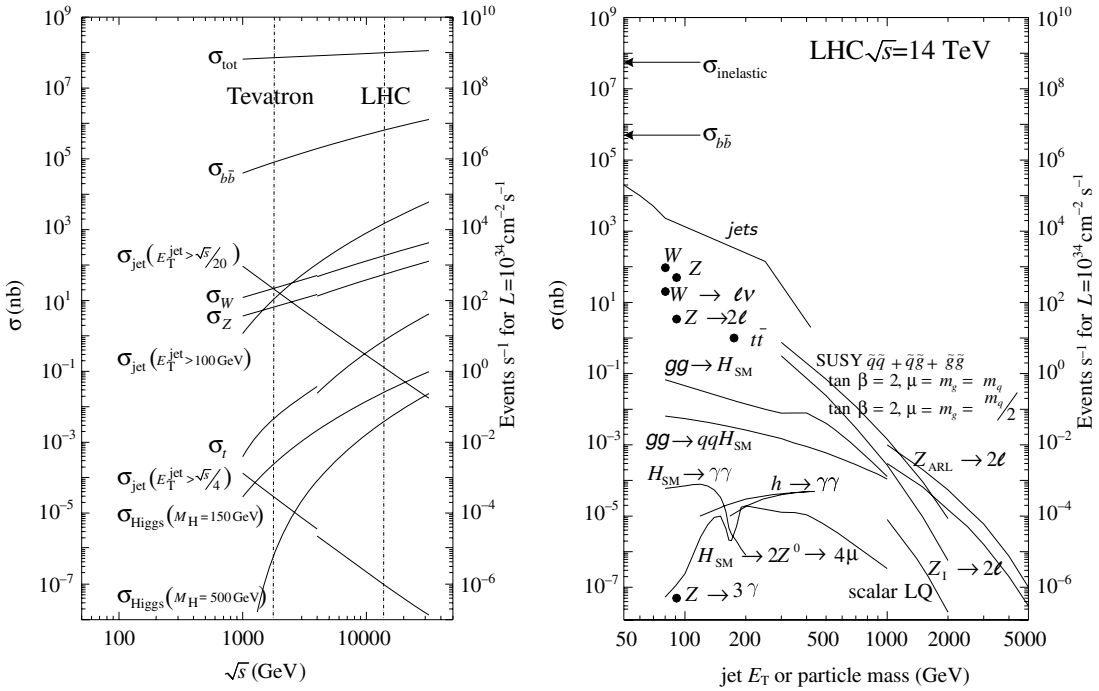


Figure 3.11: Cross-sections and event rates as a function of the center-of-mass energy for $p\bar{p}$ (Tevatron) and pp (LHC) collisions (left), and as a function of the produced particle mass or the highest jet transverse energy for $\sqrt{s} = 14$ TeV for pp collisions (right) [101].

The first layer of the CMS trigger is the Level-1 (L1) system composed of custom electronic components which admits inputs from ECAL and HCAL calorimeters, the muon system, and the beam monitoring detectors. At each bunch crossing, the calorimeters produce separate ECAL and HCAL *trigger primitives* based on calorimeter deposits, which are then processed in the Regional Calorimeter Trigger (RCT) and sent to the Global Calorimeter Trigger (GCT). GCT sorts electron, photon, and jet candidates (including jets due to hadronically decaying $\tau$ leptons) and calculates global quantities like missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) and scalar sum of jet transverse energies ($H_{\mathrm{T}}$) which are fed into the CMS Global Trigger (GT). Similarly, individual muon subsystems provide local trigger information to the Global Muon Trigger (GMT) system where muon candidates are sorted and duplicates are removed before being sent to the GT. GMT also admits information from the RCT on calorimeter deposits around the muon candidates enabling the use of isolated muon triggers. GT receives all the trigger objetcs from GCT and RCT systems, applies programmable topological cuts and different energy thresholds, and issues the final L1 decision. Additionally, GT can issue *technical*

*triggers* upon direct signal from subdetectors for special purposes such as calibration. The GT output rate is designed to be around 100 kHz. The L1 trigger chain latency is $\lesssim 3.2$ $\mu$s, and upon receiving the positive L1 decision, data kept in the front-end buffers (pipeline) pertaining to the specific bunch crossing are synchronously passed onto the Data Acquisition System (DAQ) for further analysis [58, 102, 103]. The single muon L1 triggers are measured to have $\gtrsim 95\%$ efficiency during the 8 TeV run in 2012 [104].

The CMS High Level Trigger (HLT) is a software based system operating in a single processor farm of commercially available PCs, where the full event information is analyzed via a predetermined set of algorithms with programmable structures and thresholds known as *trigger paths*, constituting a *trigger menu*. HLT improves the reconstruction and filtering processes by also having access to the tracker information which enables the reconstruction of complex objects such as hadronically decaying $\tau$ lepton candidates, jets originating from b quarks, and displaced vertices. The tracker information is used in the muon HLTs to improve the momentum and isolation measurements, where the latter is corrected for the pileup interactions in the selected event. Events satisfying any one of the HLT trigger paths are passed on to the local disks at CERN, at a rate of $\mathcal{O}(100)$ Hz. During the course of the 8 TeV run in 2012, the HLT trigger menus have been frequently updated to maintain a manageable data readout and storage rate with respect to the increasing pileup conditions. The efficiencies of isolated single muon HLTs, used in the analysis presented in this dissertation, have been measured to be about 90% [105, 106].

# Chapter 4

# Reconstruction and Identification of Physics Objects

Particles impinging on the layers of the CMS detector leave distinct signatures in various detector channels via strong and electromagnetic interactions, which enable reconstruction algorithms to attempt the identification of the true origin of these electronic signals. All charged particles produce signals in the silicon strips and pixels as they traverse the tracker. Electrons and photons deposit most of their energy via electromagnetic showers in ECAL, whereas charged/neutral hadron showers mostly take place in the HCAL. Muons, acting as minimum ionizing particles, generally leave the detector without showering. An illustration of various particle interactions within the CMS detector is given in Fig. 4.1.

The CMS experiment utilizes various reconstruction techniques to interpret the raw digital information coming from the detector channels and produce the objets with the associated energy, position and charge measurements to be used in the physics analyses. Each of the CMS subdetectors, the tracker, the calorimeters, and the muon system, are capable of independently reconstructing and delivering physics object candidates. Additionally, a holistic approach, namely the particle flow (PF) event reconstruction, is also adopted in the CMS experiment, where information from all subdetectors is combined in an attempt to individually and coherently identify and reconstruct all particles originating from the pp collisions in a given event. Composite objects, such as jets, hadronically decaying $\tau$ leptons, and missing transverse energy, are then built using these PF particles.

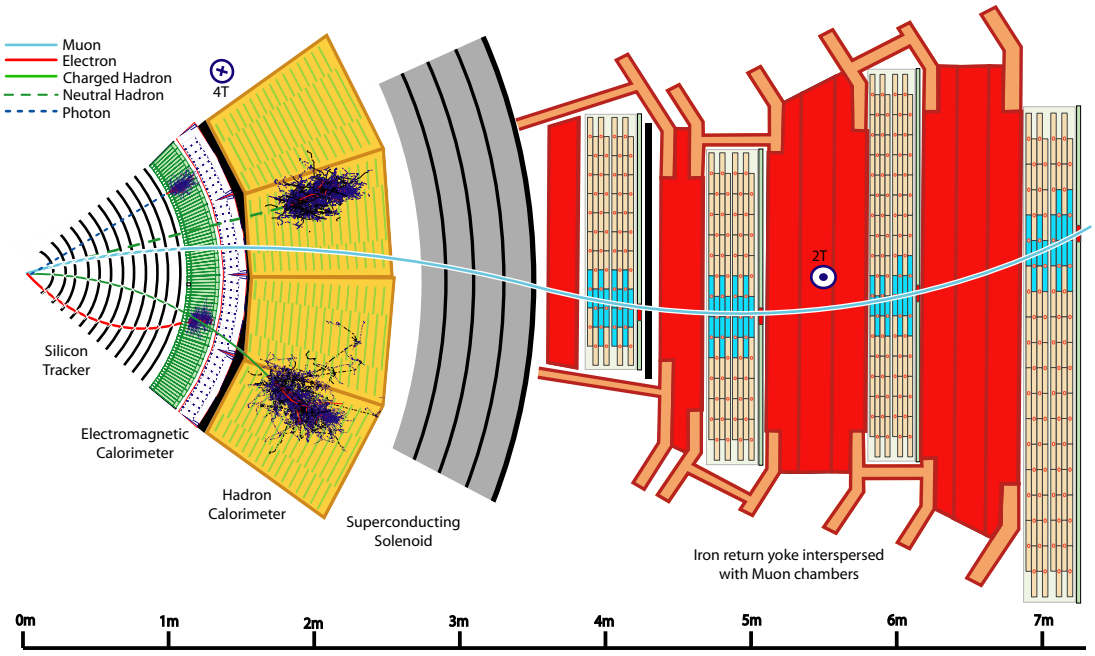


Figure 4.1: Illustration of the transverse slice of the CMS detector with superimposed particle signatures. Reproduced from Ref. [107].

## 4.1 Tracks

Charge particle tracks constitute the most basic and fundamental objects in reconstructing the collision events, contributing to the subsequent reconstruction of electron, muon and hadron candidates as well as the determination of primary interaction vertices and displaced vertices. The CMS tracking software aims to deliever efficienct reconstruction of particle tracks in the $p_{\mathrm{T}}$ range of 0.1 - 1000 GeV in a high occupancy environment with a low misidentification rate. The track reconstruction consists of 5 steps: hit reconstruction, seed generation, pattern recognition, ambiguity resolution, and final track fit [75, 108, 109].

Zero-suppressed signals (digis) produced in pixel and strip tracker layers are the building blocks of track reconstruction algorithm. A local reconstruction step forms the hits by clustering adjacent digis in pixel and strip channels that satisfy a predetermined significant signal-to-noise ratio threshold, from which the initial cluster position estimates are obtained. The pixel detector is able to deliver track and primary vertex candidates solely using the pixel hits [110], and these pixel-tracks are used to estimate the approximate beam spot (luminous region of the pp collisions) position and size via the beam spot fitter [111] prior to full tracking.

Using up to 3 hits together with the beam spot constraint and assuming helical tra-

jectories in a quasi-uniform magnetic field, it is possible to create initial track trajectory and error estimates, and this approach is used to generate the track seeds which constitute the starting point for the subsequent pattern recognition step. The Combinatorial Track Finder (CTF) software processes the available hits starting from the seed layer searching outwards, and attempts to improve the trajectory and momentum measurement by adding more hits to the reconstructed tracks, eventually utilizing the full tracker information. The CTF targets high momentum tracks in the vicinity of the interaction region with at least 2 pixel hits first, and hits belonging to these tracks are not considered for subsequent iterations, which reduces the event complexity and aids to reconstruct softer or more displaced tracks. Particle trajectories are corrected for possible inhomogeneities in the magnetic field as well as the effects of energy loss due to ionization and multiple Coulomb scattering [75, 109].

After the iterative pattern recognition step, any hit or seed ambiguities the track candidates have are resolved using the information on shared hit fraction to prevent double counting. The resulting track candidates are fitted again using a least-squares approach with relaxed seed constraints to eliminate any biases that might arise from the initial coarse trajectory estimates [75, 109].

The reconstructed tracks together with the pixel-based beam spot measurement are used to reconstruct the primary vertices in the event, including those due to pileup collisions. For vertex reconstruction, tracks are required to satisfy certain quality requirements on their transverse impact parameter, the number of strip and pixel hits, and the normalized $\chi^2$ of their trajectory fit, and then clustered along the beam axis and fitted with the adaptive vertex fitter [112].

## 4.2 Electrons

Electron identification builds on the track reconstruction described in Section 4.1, and relies on matching trajectories in the inner tracking system to energy deposits in the ECAL crystals. Although ECAL is poised to initiate and measure the electromagnetic showers, early bremsstrahlung radiation caused by interactions with the silicon tracker layers ($0.35 < X_0 < 1.4$ for $|\eta| < 2.5$) drives the electron reconstruction and measurement strategies [113, 114, 115].

As electrons traverse the tracker outwards, their trajectories curve in the $\phi$ direction

due to the solenoidal magnetic field along the beam line and spread the bremsstrahlung photons in the ECAL in a $p_{\mathrm{T}}$ dependent way. Each of the impinging electrons/photons yield energy deposits in multiple ECAL crystals, usually well constrained within a $5 \times 5$ lattice ($\simeq 97\%$). Thus, a combination of clusters (super-clusters) in a local $\eta$ - $\phi$ window is needed for the recovery of the bremsstrahlung energy for a given incident electron. This is achieved by two different algorithms in the ECAL. In the endcap region, the modified island (multi5x5) algorithm [75, 115] is used to produce clusters of individual showers which are then further combined along $\phi$-segments to yield super-clusters. In the barrel region, an inherent super-clustering algorithm, the hybrid algorithm [75, 115] is used which dynamically seeks bremsstrahlung deposits along the $\phi$ direction and combines them to the primary shower in blocks of 3-to-5 crystals (in the $\eta$ direction). In each event, the ECAL crystal channels are read out using a Selective Readout (SR) zero-suppression algorithm [116], and crystals with energy deposits above a predetermined energy threshold and corresponding to local maxima are used as seeds in both clustering algorithms [117].

Effects of bremsstrahlung radiation also have to be considered in reconstructing electron trajectories in the tracker. In addition to ionization losses dominant for all charged particles, the emitted bremsstrahlung photons carry away a significant fraction of the initial electron energy, exceeding 70% for about 35% of the electrons [113]. These photons can also convert to electron - positron pairs, yielding secondary tracks, and thus further complicate the tracking procedure. Moreover, fluctuations of the energy losses due to bremsstrahlung radiation are highly non-Gaussian, rendering the CTF software unsuited for reconstructing electron trajectories. Instead, a dedicated Gaussian Sum Filter (GSF) [118] algorithm is used in the track fit of suspected electron candidates. Pixel tracks consistent with a position and momentum projection from an ECAL super-cluster (*ECAL-driven seeding*) are considered for the GSF track reconstruction [115].

Although the ECAL-driven seeding and super-clustering reconstruction defined above provides efficient reconstruction of isolated electrons with relatively high $p_{\mathrm{T}}$, a complementary PF electron reconstruction approach is also utilized that especially enhances the coverage for low $p_{\mathrm{T}}$ ($< 10$ GeV) or non-isolated electron candidates. In addition to the ECAL-driven seeding, PF electron reconstruction also utilizes the *tracker-driven seeding*, where CTF tracks with projections matching to an ECAL PF cluster and with

missing tracker hits or poor $\chi^2$ values are considered for the GSF fit [115]. PF clustering and PF electron reconstruction are described in Section 4.4.

The final momentum determination of reconstructed electron candidates utilizes both the tracker and ECAL based measurements. The tracker provides a more precise momentum estimate for low $p_{\mathrm{T}}$ ($< 15$ GeV) electrons but is more liable to underestimates due to bremsstrahlung losses. Thus, a combination of the tracker momentum and super-cluster energy measurements are used depending on the estimated $p_{\mathrm{T}}$ and bremsstrahlung energy fraction of a given electron candidate [113, 115].

## 4.3 Muons

The CMS detector is able to reconstruct muon candidates both using the tracker and the muon detectors. Based on the detector systems utilized in muon reconstruction, three types are defined: the *standalone muon* candidates, the *tracker muon* candidates, and the *global muon* candidates.

Standalone muon reconstruction is mainly based on the DT and CSC systems, with additional information provided by the RPC system in the barrel-endcap overlap region. Track segments obtained by *local reconstruction* [119] in the DT and CDC chambers constitute the fundamental building blocks. In each DT chamber, triplets of hits in the SLs are used as seeds and are grown into linear segments with additional hits, consistent with the beam spot location and size. In CSC, 2-dimensional hits are constructed combining wire-bundle and strip measurements in each chamber layer. A series of at least four hits consistent with a linear fit with a maximum r-$\phi$ spread of 1 cm are retained as segments. Unlike the DT and CSC systems, local RPC reconstruction yields hit locations instead of segments. These are calculated as the geometric centers of the signal carrying strip clusters in the RPC detector plane. Following the local reconstruction step, the innermost DT and CSC segments are treated as seeds for the standalone muon trajectory reconstruction. This step is performed with the Kalman filter technique [120], taking into account the effects of energy losses in the material, multiple scattering, and non-homogenous magnetic fields. The reconstruction is first performed inside-out, and then outside-in with the added inclusion of the beam spot constraint, yielding the final standalone muon candidates [75].

In contrast to the standalone muon construction, the tracker muons are constructed

with minimal use of the muon system. A tracker track obtained as defined in Section 4.1 is considered a tracker muon if it has $p_{\mathrm{T}} > 0.5$ GeV and $p > 2.5$ GeV, and the projection of the tracker track to the muon system yields a match to a DT or CSC segment [96]. Tracker muon reconstruction is especially efficient for muons with $p_{\mathrm{T}} \lesssim 5$ GeV that may not produce multiple segments in the muon system.

Global muons are constructed by the combination of tracker tracks obtained as defined in Section 4.1 with the standalone muon tracks obtained in the muon systems. The matching is performed by projecting the standalone muon trajectories to the outer tracker surface, and the global muon track is obtained via a re-fit of these hits with the Kalman filter technique. The global re-fit is also repeated excluding the hits with high $\chi^2$ values, hits in high occupancy muon chambers, or only including the innermost muon segments in order to detect any muon bremsstrahlung or other significant energy losses which are especially important for TeV range muons. The global muon reconstruction achieves a better $p_{\mathrm{T}}$ resolution in comparison to tracker-only measurements for muons with $p_{\mathrm{T}} \gtrsim 200$ GeV [75, 96].

## 4.4 Particle Flow

The particle flow (PF) approach is an attempt to classify all particles emerging from the pp collisions, namely electrons, muons, photons, and charged/neutral hadrons. Via coherent use of all subdetectors, overlapping energy deposits of charged and neutral particles are resolved, different particle types are optimally identified, and a self-consistent collection of all particles is created with the associated four-vector and charge measurements. This resultant PF particle collection provides the building blocks for more composite objects such as jets, missing transverse energy, and hadronically decaying $\tau$ leptons. The PF algorithm is designed to maximally exploit the high efficiency tracking and high granularity calorimetry performance (especially ECAL) of the CMS detector, and consists of 3 major steps [121].

The first step is the reconstruction of *elements*, charged particle tracks and clusters of calorimeter deposits, which provide the input to the PF. Tracks are obtained as described in Section 4.1, and the clusters are built by a dedicated PF clustering algorithm tuned to achieve high detection efficiency, including low-energy particles and separation of close energy deposits. The clustering procedure starts by the detection

of seeds, which correspond to calorimeter channels with local energy maxima, and the adjacent channels sharing at least a side with those already in the cluster are then added in, forming the *topological clusters*. Energy deposits per calorimeter channel have to be above predetermined thresholds both in ECAL and HCAL in order to suppress contributions due to electronic noise fluctuations. The final cluster positions and energies are iteratively determined by reweighting the individual channel contributions according to the channel-cluster distance [121].

The reconstruction of PF *blocks* constitutes the second step in the PF event reconstruction. Since a single particle traversing the detector can give rise to multiple PF elements in the form of tracks and clusters, these need to be connected via a link algorithm to form blocks in order to eliminate double-counting within the PF collection. Each track is extrapolated from the last measured hit in the tracker to the ECAL and HCAL detectors, at depths compatible with an electron and a hadron shower profile, respectively, and clusters which are found to include the projected track within their boundaries are linked to the track. The ECAL preshower detector is also considered in the endcap region. Individual ECAL (including the preshower detector) and HCAL clusters are also linked to each other if the position of a cluster in the more granular detector is found to be within the boundary of a cluster in the less granular detector. Furthermore, tracks in the tracker are also linked to the segments or tracks in the muon system forming the global and tracker muon candidates as described in Section 4.3. Therefore, each global or tracker muon also corresponds to a PF block [96]. A resultant block usually consists of $\leq 3$ elements, and the block quality is defined by the compatibility of its constituent elements, either as the $\eta$-$\phi$ distance (track-to-cluster and cluster-to-cluster links) or the global fit $\chi^2$ (track-to-track links). The limited size of each PF block renders the PF particle identification performance almost independent of event complexity [121].

The third step of the PF algorithm is the reconstruction and identification of individual PF particles based on the PF blocks. These are detailed below in the context of muons, electrons, charged/neutral hadrons, and photons, whereas composite objects such as jets, missing transverse energy, and hadronically decaying $\tau$ leptons build up on the final PF particle collection.

### 4.4.1 PF Muons

Each global muon object whose combined momentum is compatible with that determined solely in the tracker measurement within three standard deviations and each tracker muon object is considered a candidate for particle-flow muon identification [121]. For a given PF muon candidate, a variety of selection requirements are imposed depending on the candidate isolation. If a given muon candidate has the scalar $p_{\mathrm{T}}$ sum of the neighboring tracks within a cone of $\Delta R < 0.3$ (where $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$) less than 10% of the muon $p_{\mathrm{T}}$, no additional PF selection criteria are applied and the object is identified as a PF muon. Remaining non-isolated PF muon candidates are re-evaluated with a loosened number of muon track hits requirement and a compatibility of these muon tracks with the calorimeter deposits and segments in the muon system. The combination of isolated and non-isolated as well as tracker and global muon candidates yields a high PF muon identification efficiency over a wide $p_{\mathrm{T}}$ range which is essential in the subsequent calculation of PF jet energies as well as PF $E_{\mathrm{T}}^{\mathrm{miss}}$ [96, 122].

### 4.4.2 PF Electrons

PF electron reconstruction starts with electron track candidates seeded by the ECAL-driven or tracker-driven algorithms as described in Section 4.2 and reconstructed by the GSF fit. In contrast to the ECAL-driven reconstruction defined earlier, bremsstrahlung energy losses are recovered by individually detecting and linking these ECAL deposits as PF clusters to the main GSF track. This is achieved by linearly and tangentially extrapolating the GSF track to the ECAL at each tracker layer it traverses, and adding matching PF clusters to the electron energy. ECAL PF clusters that match to the electron super-cluster but are not linked by the PF algorithm are also recovered for the ECAL-driven PF electron candidates. Converted bremsstrahlung photons are also considered by a dedicated Kalman filter-based algorithm where displaced electron tracks are associated with PF clusters. A loose selection is performed on the PF electron candidates based on the tracker and PF cluster observables such as track kinematics, track quality and shower shape, yielding the final PF electron collection with lowered misidentification probability whilst maintaining efficiency [115, 121, 122].

### 4.4.3 PF Charged/Neutral Hadrons and Photons

The reconstruction and identification of charged/neutral hadrons and photons primarily rely on the comparison of the track momentum with the calorimeter energy measurements in a given PF block. This step starts with the removal of the tracks and PF clusters corresponding to the already identified muons and electrons from the PF blocks, and the correction of the energy measurements of muon-linked calorimeter clusters for any muon deposits. Each of the remaining tracks is then required to have a relative $p_{\mathrm{T}}$ uncertainty smaller than the relative calorimetric energy resolution in order to reduce the contributions due to misidentified tracks. Furthermore, PF cluster energies are calibrated to compensate for the non-linear calorimeter response, such that a comparison between the track and calorimeter measurements can be made [121, 123].

If a given track is linked to multiple calorimeter clusters, only the closest HCAL cluster is kept, whereas links to multiple ECAL clusters are allowed as long as the total calorimeter energy does not exceed the track momentum. In the case of a calorimetric energy excess for a track linked to multiple clusters, the links to ECAL clusters are progressively disabled starting with those with the smallest energy until the condition above is satisfied. Since a given cluster could be linked to multiple tracks as well, the sum of individual track momenta and the sum of individual cluster energies are used for comparison in each PF block. If the total tracker momentum exceeds the total cluster energy in a given PF block by over 3 standard deviations, a relaxed search for additional muons and fake tracks is performed in order to reduce the number of tracks in the block, and hence the expected calorimetric energy. The remaining tracks are then identified as PF charged hadrons, and their momenta are assigned using an uncertainty weighted combination of tracker and calorimeter measurements under the charged pion mass hypothesis [121].

If the total calorimetric energy is found to be significantly larger than that of the total track momentum, and if the excess is also larger than the ECAL cluster energy, a PF photon with the entire ECAL cluster energy as well as a neutral hadron with the remaining excess energy is created. Otherwise, just a PF photon with the remaining excess energy is created. Photons are given precedence in accounting for an excess in the calorimeter energy measurements since photons are expected to constitute about 25% of a jet's energy, whereas the neutral hadron component is expected to be about 3%.

Lastly, any remaining unlinked ECAL and HCAL clusters are identified as PF photons and PF neutral hadrons, respectively [121].

### 4.4.4 PF Jets and $E_{\mathrm{T}}^{\mathbf{miss}}$

The quarks (with the exception of the top quark) and gluons originating from a hard interaction undergo a fragmentation and hadronization process in a timescale of $\sim 10^{-23}$s, and produce a collection of hadronic bound states which share the initial energy and momentum via a process known as *QCD jet production*. Fragmentation denotes the low $p_{\mathrm{T}}$ development of the jet where the initial parton momentum is shared by the splitting of partons, whereas hadronization refers to the subsequent formation of color-neutral hadrons. Charged and neutral pions, whose reconstruction and identification methods are detailed in Section 4.4.3, constitute the majority of particles produced in jets with smaller contributions coming from heavier mesons and baryons. Jet algorithms provide a mapping between the partons, hadrons, and the final detector signatures such as charged particle tracks and calorimeter towers. Various reconstruction methods and recombination schemes are available to estimate the energy and momentum of the initial strongly interacting particle via a careful assessment of the detector signatures of the produced hadron collection [3, 124].

A desirable property of jet algorithms is the infrared and collinear (IRC) safety. IRC-safe algorithms provide a stable set of reconstructed jets even if a given event is modified by a collinear splitting of a hard parton or by the addition of a soft emission, and are important in yielding experimental measurements comparable to fixed-order perturbative QCD calculations [124]. Traditionally, iterative and fixed cone algorithms are known to suffer from IRC-unsafety issues [124], whereas the CMS experiment utilizes various IRC-safe algorithms such as the $k_{\mathrm{T}}$ [125, 126], Cambridge/Aachen [127], SISCone [128], and anti-$k_{\mathrm{T}}$ [129] algorithms with various jet radius sizes [130].

The anti-$k_{\mathrm{T}}$ jet algorithm is an IRC-safe, clustering based, sequential recombination algorithm. In a given collection of input objects, the clustering procedure computes two distance measures for each object $i$, namely the distances between $i^{\mathrm{th}}$ and $j^{\mathrm{th}}$ objects as well as the $i^{\mathrm{th}}$ object and the beam line. These are given in Eq. (4.1) below as $d_{ij}$ and $d_{iB}$, respectively, where $k_{\mathrm{T}}$ is the object transverse momentum, $\eta$ is the object

pseudo-rapidity, $\phi$ is the object azimuthal angle, and $R$ is the predefined jet radius:

$$d_{ij} = \min\left\{1/k_{\mathrm{T},i}^2,\ 1/k_{\mathrm{T},j}^2\right\} \frac{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}{R^2} \tag{4.1a}$$

$$d_{iB} = 1/k_{\mathrm{T},i}^2 \tag{4.1b}$$

For the given object $i$, if there is another object $j$ such that $d_{ij} < d_{iB}$, the $(i, j)$ pair is replaced in the input collection by the new combined object obtained via the energy recombination scheme, where individual four-vectors are added. Otherwise this object $i$ is declared as a jet and removed from further consideration. These steps are then repeated until all the objects in the input collection are exhausted. As the soft particles tend to cluster with the hard ones before clustering within themselves, the anti-$k_{\mathrm{T}}$ algorithm produces soft-resilient jets with well defined circular areas of radius $R$ [3, 129].

The CMS experiment can deliver four distinct input object collections with which jets can be reconstructed that differ in the ways individual subdetector contributions are combined; namely the PF jets, calorimeter jets, jet-plus-track jets, and track jets [131]. The analysis presented here uses jets reconstructed by the anti-$k_{\mathrm{T}}$ algorithm with a cone size of $R = 0.5$ with the PF objects as the input collection, also referred to as *AK5 PF jets*. A number of jet energy corrections are applied to the resulting jet object collection in order to account for the effects of pileup, energy response variations in $\eta$ and $p_{\mathrm{T}}$, as well as jet energy discrepancies in data versus simulation [132].

The electrically-neutral color-singlet particles, such as neutrinos, escape the detector without causing any energy deposits, but nevertheless can be detected by the momentum imbalance they induce in the transverse plane to the beam axis. This imbalance can be calculated vectorially, and its magnitude is denoted as the missing transverse energy, $E_{\mathrm{T}}^{\mathrm{miss}} = |-\sum \vec{p}_{\mathrm{T}}|$. Although various input object collections can be used for the calculation of $E_{\mathrm{T}}^{\mathrm{miss}}$, the PF collection is used in this analysis, yielding the quantity PF $E_{\mathrm{T}}^{\mathrm{miss}}$ [133].

The determination of the $E_{\mathrm{T}}^{\mathrm{miss}}$ at a given event is sensitive to a number of factors including the inefficiencies and thresholds in the tracker and calorimeter measurements, the non-linear response of the CMS calorimeter system, contributions due to pileup energy deposits, as well as the imperfections in the detector alignment and calibration. Therefore, PF $E_{\mathrm{T}}^{\mathrm{miss}}$ measurements are also corrected for these effects to yield an unbiased estimate of the true missing transverse energy [132, 133].

### 4.4.5 PF Tau leptons

As the heaviest of the charged leptons with a mass of about 1.777 GeV and a mean lifetime of $2.91 \times 10^{-13}$ s [43], $\tau$ leptons can undergo a variety of decay processes producing hadrons in about two-thirds of all instances, which can then be reconstructed and identified at the CMS detector [122, 134, 135]. Table 4.1 summarizes the dominant $\tau$ decay modes and the corresponding experimental signatures, whereas further details can be found in Ref. [43].

| Decay mode | Resonance | Branching ratio (%) | Reconstruction |
|---|---|---|---|
| $\tau^- \to e^- \bar{\nu}_e \nu_\tau$ | | 17.8 | $e^-$ |
| $\tau^- \to \mu^- \bar{\nu}_e \nu_\tau$ | | 17.4 | $\mu^-$ |
| $\tau^- \to h^- \nu_\tau$ | $\pi^-(140)$ | 11.5 | $\tau_h^-$ (1-prong) |
| $\tau^- \to h^- \pi^0 \nu_\tau$ | $\rho^-(770)$ | 26.0 | $\tau_h^-$ (1-prong) |
| $\tau^- \to h^- \pi^0 \pi^0 \nu_\tau$ | $a_1^-(1260)$ | 10.8 | $\tau_h^-$ (1-prong) |
| $\tau^- \to h^- h^+ h^- \nu_\tau$ | $a_1^-(1260)$ | 9.8 | $\tau_h^-$ (3-prong) |
| $\tau^- \to h^- h^+ h^- \pi^0 \nu_\tau$ | | 4.8 | Not targeted$^\dagger$ |
| | $^\dagger$Not explicitly targeted due to contamination by jet $\to \tau_h$ fakes. | | |
| All hadronic modes | | 64.8 | |

Table 4.1: The dominant $\tau$ lepton decay modes and branching fractions, and the corresponding reconstructed signatures at the CMS detector. Charged hadrons are denoted by $h^-$, which typically stands for a $\pi^-$ or $K^-$, and the hadronically decaying $\tau$ leptons are denoted as $\tau_h$. All $\tau$ lepton decay modes have an accompanying $E_T^{miss}$ component in the detector reconstruction due to the neutrinos in the decay chain. $\tau^+$ decays follow the charge-conjugated table [134, 135].

The hadronic decay modes involve one or three charged hadrons, typically a $\pi^\pm$ or a $K^\pm$, which are often accompanied by one or two $\pi^0$s. The charged hadrons are long-enough lived to traverse the CMS tracking system and subsequently shower in the HCAL subdetector, whereas the neutral pions almost immediately decay into two photons, $\pi^0 \to \gamma\gamma$, which are detected in the ECAL. The $\pi^0$ decays could also give rise to pairs of electrons via photon conversions. Therefore, the distinguishing experimental signature of the hadronically decaying $\tau$ lepton, $\tau_h$, is a collimated jet of few charged tracks accompanied by ECAL and HCAL deposits [134, 135].

In CMS, the individual $\tau_h$ decay products, charged hadrons and photons (including conversions to electrons), are reconstructed and identified by the PF algorithm as described above and they are subsequently clustered into jets by the anti-$k_T$ algorithm. The AK5 PF jet collection provides the input object collection to the Hadron

plus Strips (HPS) algorithm, which then performs the reconstruction and identification steps. Using the AK5 PF jet as the seed, the HPS algorithm first attempts to reconstruct the $\pi^0$ candidates in an iterative procedure. To begin with, a strip extending up to $0.05 \times 0.20$ in the $\eta$-$\phi$ plane is centered on the most energetic photon/electron PF object with $p_{\mathrm{T}} > 0.5$ GeV. The neighboring photon/electron candidates within the strip boundaries with $p_{\mathrm{T}} > 0.5$ GeV are then combined forming a new four-vector on which the strip is re-centered. These steps are repeated until all the photon/electron candidates that could be associated with the strip is exhausted. If the resultant strip satisfies $p_{\mathrm{T}} > 2.5$ GeV, it is considered for the combination with charged hadron candidates within the same jet under various $\tau$ lepton hadronic decay mode assumptions.

The HPS algorithm considers the single charged hadron, one charged hadron plus one strip, one charged hadron plus two strips, and three charged hadrons topologies, and the charged hadrons are required to satisfy one of the intermediate resonance masses as listed in Table 4.1. Additionally, all charged hadrons and strips are required to be in a signal cone of $\Delta R = 3$ GeV$/p_{\mathrm{T}}$ where $p_{\mathrm{T}}$ is the resultant $\tau_{\mathrm{h}}$ momentum. The signal cone is centered on the momentum axis of the $\tau_{\mathrm{h}}$ candidate and further constrained to be between $\Delta R = 0.05$ and $\Delta R = 0.1$. In cases where the reconstructed hadron and strip objects are found to be in agreement with multiple decay mode hypotheses and the signal cone requirement, the assumption yielding the highest $p_{\mathrm{T}}$ $\tau_{\mathrm{h}}$ candidate is given precedence [134, 135].

## 4.5 Object Isolation and Discrimination Variables

In physics analyses, the sought after lepton objects often originate from Z or W boson decays that are produced in the hard interaction (including those coming from a top quark decay), and thus, found in isolated regions of the detector away from any substantial hadronic activity that also originates from the same hard interaction with the notable exception of very boosted topologies.

In order to utilize this aspect of lepton signatures, certain isolation variables have been defined that relies on the assessment of detector activity within an isolation cone of a given size around the object under consideration. Although these variables are separately tuned for electrons, muons, and $\tau$ leptons, a common challenge to all these isolation schemes is the elimination of the inevitable contributions due to the pileup collisions at

the LHC. Charged hadron pileup contributions can be separated by the superb tracking and vertexing capabilities of the CMS detector and the PF event reconstruction, whereas different approaches are adopted in dealing with the neutral particle contributions due to pileup as these are not detected in the tracking system and cannot be associated with any vertex. Additionally, $\tau_{\mathrm{h}}$ identification algorithms utilize the tracker, calorimeter and muon system information in the vicinity of the $\tau$ lepton candidate under consideration to eliminate contributions due to misidentified electrons and muons.

The subset of these isolation and $\mathrm{e}/\mu \to \tau_{\mathrm{h}}$ rejection variables as well as the pileup subtraction schemes used in the analysis presented here are described below.

### 4.5.1 Electron Isolation Variables

A combined relative isolation with effective-area pileup corrections is used for the electron objects as defined in Eq. (4.2a). The charged hadrons, neutral hadrons, and photons are denoted as $h^{\pm}$, $h^{0}$, and $\gamma$, respectively, where $h^{\pm}$s are required to be associated with the vertex from which the electron candidate originates. All summations of $p_{\mathrm{T}}$s are performed over PF objects of the given type within a cone of size $\Delta R < 0.4$ centered on the electron object [115].

The effective-area pileup correction is calculated using the average energy density in a given event denoted as $\rho$, and the effective area $A_{\mathrm{eff}}$ of the isolation cone in terms of its catchment of neutral hadrons and photons coming from pileup. The $\rho$ parameter is calculated via the FastJet technique [136, 137, 138] and is defined as the median of the energy density distribution for particles within the area of any jet in the event with $p_{\mathrm{T}} > 3$ GeV and $|\eta| < 2.5$ as reconstructed by the $k_{\mathrm{T}}$ jet algorithm with a cone size of $R = 0.6$. Both the neutral particle (PF neutral hadrons plus photons) contributions within the given isolation cone and the average energy density $\rho$ can be defined by first-degree polynomials as a function of vertex multiplicity in the event. The effective area is then defined as the ratio of the slopes of these two linear functions, and is given in Eq. (4.2b) in bins of electron $|\eta|$ [115].

$$I_{\mathrm{e}} = \sum_{\Delta R<0.4} p_{\mathrm{T}}^{h^{\pm}} + \max\left\{ \sum_{\Delta R<0.4} p_{\mathrm{T}}^{h^{0}} + \sum_{\Delta R<0.4} p_{\mathrm{T}}^{\gamma} - \rho A_{\mathrm{eff}},\ 0 \right\} \tag{4.2a}$$

$$A_{\mathrm{eff}} = \begin{cases} 0.13 & \text{if} \quad |\eta| < 1.0 \\ 0.14 & \text{if} \quad 1.0 \leq |\eta| < 1.479 \\ 0.07 & \text{if} \quad 1.479 \leq |\eta| < 2.0 \\ 0.09 & \text{if} \quad 2.0 \leq |\eta| < 2.2 \\ 0.11 & \text{if} \quad 2.2 \leq |\eta| < 2.4 \\ 0.14 & \text{if} \quad 2.4 \leq |\eta| \end{cases} \tag{4.2b}$$

Two working points (WPs) are defined for the electron isolation. The *loose* combined relative PF isolation is given as $I_{\mathrm{e}}/p_{\mathrm{T}}^{\mathrm{e}} < 15\%$ for electrons with $|\eta| < 1.442$ over the whole $p_{\mathrm{T}}$ range, whereas it is $I_{\mathrm{e}}/p_{\mathrm{T}}^{\mathrm{e}} < 15(10)\%$ for electrons with $|\eta| > 1.442$ and $p_{\mathrm{T}} > 20$ GeV ($p_{\mathrm{T}} < 20$ GeV). Similarly, the *tight* combined relative PF isolation is $I_{\mathrm{e}}/p_{\mathrm{T}}^{\mathrm{e}} < 10\%$ for electrons with $|\eta| < 1.442$ over the whole $p_{\mathrm{T}}$ range, whereas it is $I_{\mathrm{e}}/p_{\mathrm{T}}^{\mathrm{e}} < 10(7)\%$ for electrons with $|\eta| > 1.442$ and $p_{\mathrm{T}} > 20$ GeV ($p_{\mathrm{T}} < 20$ GeV) [139].

### 4.5.2 Muon Isolation Variables

Muon objects also utilize the combined relative isolation, but with the $\Delta\beta$ corrections instead of the effective-area corrections, as defined in Eq. (4.3a). The charged hadrons, neutral hadrons, and photons are denoted as $h^{\pm}$, $h^{0}$, and $\gamma$, respectively, where $h^{\pm}$s are required to be associated with the vertex from which the muon candidate originates. The $\Delta\beta$ correction given in Eq. (4.3b) is calculated using the charged hadrons that are not associated with vertex of the muon candidate under consideration, and this is denoted as PU. All summations of $p_{\mathrm{T}}$s are performed over PF objects of the given type within a cone of size $\Delta R < 0.4$ centered on the muon object [140].

$$I_{\mu} = \sum_{\Delta R<0.4} p_{\mathrm{T}}^{h^{\pm}} + \max\left\{ \sum_{\Delta R<0.4} p_{\mathrm{T}}^{h^{0}} + \sum_{\Delta R<0.4} p_{\mathrm{T}}^{\gamma} - \Delta\beta,\ 0 \right\} \tag{4.3a}$$

$$\Delta\beta = \frac{1}{2} \cdot \sum_{\Delta R<0.4} p_{\mathrm{T}}^{\mathrm{PU}} \tag{4.3b}$$

The factor of 0.5 corresponds to the approximate ratio of neutral to charged particles as measured in jets. Two WPs are defined, such that the *loose* combined relative PF isolation is given as $I_{\mu}/p_{\mathrm{T}}^{\mu} < 20\%$ and the *tight* combined relative PF isolation is given as $I_{\mu}/p_{\mathrm{T}}^{\mu} < 12\%$ [140].

### 4.5.3 Tau lepton Isolation and Discrimination Variables

The decay-mode based reconstruction yields a $\tau_h$ object collection heavily contaminated by misidentified (fake) candidates originating from QCD quark and gluon jets. Additionally, muons and electrons are also known to produce fake contributions especially to the single charged hadron $\tau$ decay modes. Thus, a number of isolation as well as anti-$\mu$ and anti-e rejection discriminators are designed to suppress these contributions [134, 135].

A pileup corrected isolation requirement is used to suppress the jet $\rightarrow \tau_h$ fakes. The isolation quantity is defined as the scalar $p_T$ sum of all PF charged hadrons and photons with $p_T > 0.5$ GeV within an isolation cone of size $\Delta R = 0.5$ centered around the $\tau_h$ candidate. The PF charged hadrons are required to be within 2 mm along the beam line and 0.3 mm in the transverse plane of the primary vertex the $\tau_h$ candidate is associated with. The PF photon contributions to the isolation are corrected via the $\Delta\beta$ method, where the pileup charged hadron contributions within an extended cone of size $\Delta R < 0.8$ are used to estimate the pileup photon contributions within the isolation cone [134, 135].

$$I_\tau = \sum_{\Delta R<0.5} p_T^{\text{charged}}(|d_z| < 2 \text{ mm}, |d_{xy}| < 0.3 \text{ mm}) + \max\left\{\sum_{\Delta R<0.5} p_T^{\gamma} - \Delta\beta,\ 0\right\} \tag{4.4a}$$

$$\Delta\beta = 0.4576 \cdot \sum_{\Delta R<0.8} p_T^{\text{charged}}(|d_z| > 2 \text{ mm}) \tag{4.4b}$$

The final isolation quantity $I_\tau$ is defined in Eq. (4.4a) and the $\Delta\beta$ quantity is given in Eq. (4.4b). For both calculations, individual summations are performed around the $\tau_h$ candidate within the stated cone size whilst excluding the PF charged hadrons and PF photons used in the reconstruction of the $\tau_h$. The numerical constant used in the $\Delta\beta$ description is tuned to render $\tau_h$ isolation insensitive to pileup [134, 135]. Various WPs have been defined with the $\Delta\beta$ corrected isolation criteria such as *vloose*, *loose*, *medium*, and *tight* WPs, corresponding to an upper $I_\tau$ threshold of 3 GeV, 2 GeV, 1 GeV, and 0.8 GeV, respectively. An alternate set of *loose*, *medium*, and *tight* WPs is also available where a looser (3 instead of 8) number of track hits requirement is applied for the PF charged hadron candidates in the isolation cone [141].

The electron rejection is performed via an MVA-based discriminator. Various parameters relating to the ECAL and HCAL shower shapes as well as the bremsstrahlung

energy deposits along the leading $p_{\mathrm{T}}$ track and overall particle multiplicity are inputs to the MVA algorithm. Five WPs, *vloose*, *loose*, *medium*, *tight*, and *vtight*, are defined for the electron-rejection discriminant as different cuts on the MVA output [135, 141].

The muon rejection is performed via a cut-based discriminator. In the *loose* WP, no track segments are allowed in more than one muon station within a cone of size $\Delta R = 0.5$ centered on the $\tau_{\mathrm{h}}$ candidate, and the combined ECAL and HCAL energy deposit associated to the leading $p_{\mathrm{T}}$ track of the $\tau_{\mathrm{h}}$ candidate is required to be more than 20% of the track momentum. In the *tight* WP, the $\tau_{\mathrm{h}}$ candidates are required to satisfy the *loose* WP and have no hits in the CSC, DT or RPC chambers located in the two outermost muon stations within a cone of size $\Delta R = 0.5$ centered on the $\tau_{\mathrm{h}}$ candidate [135, 141].

## 4.6 Object Selection Requirements

Certain quality criteria are applied to the physics object collections reconstructed in the CMS detector to suppress the fake contributions. These object selection requirements are summarized below.

### 4.6.1 Electron Selection Requirements

The *CMS cut-based loose electron selection* is used for the reconstructed electron objects [115, 139]. The electrons are required to have a $p_{\mathrm{T}} > 15$ GeV and $|\eta| < 2.5$, and are vetoed in the ECAL barrel-endcap transition region given as $1.442 < |\eta| < 1.566$. To obtain isolated candidates, the loose combined relative PF isolation as defined in Section 4.5.1 is used for all electron objects. The complete selection requirements are provided in Table 4.2.

The differences between the ECAL super-cluster and GSF track $\eta$ and $\phi$ measurements are denoted as $\Delta\eta_{\mathrm{in}}$ and $\Delta\phi_{\mathrm{in}}$, respectively. The $\sigma_{\mathrm{i\eta i\eta}}$ parameter is a weighted sum over $5 \times 5$ ECAL crystals, quantifying the lateral extent of the electron shower profile in the $\eta$ direction. The sum of the HCAL tower energies within a cone of $\Delta R = 0.15$ around the electron object is denoted as $H$. The electron super-cluster energy is given as $E_{\mathrm{SC}}$, and the track momentum at the point of closest approach to the primary vertex is given as $p_{\mathrm{in}}$. The transverse and longitudinal distances to the primary vertex at the point of closest approach in the transverse plane are denoted as $|d_{\mathrm{XY}}|$ and $|d_{\mathrm{Z}}|$, respectively.

| Electron parameters | Criteria | |
|---|---|---|
| | Barrel ($\lvert\eta\rvert < 1.442$) | Endcap ($\lvert\eta\rvert > 1.566$) |
| $\Delta\eta_{\text{in}}$ | $< 0.007$ | $< 0.009$ |
| $\Delta\phi_{\text{in}}$ | $< 0.15$ rad | $< 0.10$ rad |
| $\sigma_{i\eta i\eta}$ | $< 0.01$ | $< 0.03$ |
| $H/E_{\text{SC}}$ | $< 0.12$ | $< 0.10$ |
| $\lvert 1/E_{\text{SC}} - 1/p_{\text{in}}\rvert$ | $< 0.05$/GeV | $< 0.05$/GeV |
| $\lvert d_{\text{XY}}\rvert$ | $< 0.02$ cm | $< 0.02$ cm |
| $\lvert d_{\text{Z}}\rvert$ | $< 0.2$ cm | $< 0.2$ cm |
| Conversion fit probability | $< 10^{-6}$ | $< 10^{-6}$ |
| Missing hits | $\leq 1$ | $\leq 1$ |
| $I_{\text{e}}/p_{\text{T}}^{\text{e}}$ | $< 15\%$ | $< 15\%$ if $p_{\text{T}}^{\text{e}} > 20$ GeV<br>$< 10\%$ if $p_{\text{T}}^{\text{e}} < 20$ GeV |

Table 4.2: Electron selection criteria as used in this analysis. Access methods to these variables within the CMSSW framework are provided in Ref. [142].

The rejection of secondary electrons coming from photon conversions in the tracker is achieved by the assessment of the $\chi^2$ probability that the electron track is compatible with a conversion vertex, as well as the number of missing hits in the innermost layer of the tracker [115, 139].

### 4.6.2 Muon Selection Requirements: Loose and Tight

Two sets of muon selections are defined for the purposes of this analysis, denoted as *loose* and *tight* selections, which are differentiated by the isolation requirements on the muon object. The loose selection is equal to the so-called *CMS tight muon selection* [140]. The tight selection includes this loose selection, and additionally requires the muon candidate to satisfy the tight combined relative PF isolation as defined in Section 4.5.2. The details of these selections are provided in Table 4.3. All muons are required to have $p_{\text{T}} > 25$ and $|\eta| < 2.1$.

The number of degrees of freedom in the global muon track fit is denoted as $N_{\text{dof}}$, whereas the transverse and longitudinal distances to the primary vertex at the point of closest approach in the transverse plane are denoted as $|d_{\text{XY}}|$ and $|d_{\text{Z}}|$, respectively.

### 4.6.3 Tau lepton Selection Requirements: Loose and Tight

Similar to the muon selections described above, two $\tau_{\text{h}}$ selections are defined. These are denoted as the *loose* and *tight* selections, and are differentiated by the isolation require-

| Muon parameters | Criteria |
|---|---|
| Loose selection: | |
| Is global muon? | Yes |
| Is particle flow muon? | Yes |
| Global muon track fit $\chi^2/N_{\text{dof}}$ | $< 10$ |
| $\lvert d_{\text{XY}}\rvert$ | $< 0.2$ cm |
| $\lvert d_{\text{Z}}\rvert$ | $< 0.5$ cm |
| Number of segments in muon stations | $> 1$ |
| Number of hits in the pixel detector | $> 0$ |
| Number of hits in tracker layers | $> 5$ |
| Number of hits in muon chambers included in the global muon track fit | $> 0$ |
| Tight selection (in addition to the loose selection): | |
| $I_\mu/p_{\text{T}}^{\mu}$ | $< 10\%$ |

Table 4.3: Loose and tight muon selection criteria as used in this analysis. Access methods to these variables within the CMSSW framework are provided in Ref. [142].

ments on the $\tau$ object. Each $\tau$ lepton selection refer to a collection of predefined boolean $\tau$ object discriminators as described in Section 4.5.3, which are cut-based decisions based on either simple reconstructed objects and detector measurements or various MVA outputs [135, 141]. The PF $\tau$ lepton discriminator collections corresponding to the loose and tight selections are detailed in Table 4.4. All $\tau_{\text{h}}$ objects are required to have $p_{\text{T}} > 20$ GeV and $|\eta| < 2.1$. Additionally, $\tau_{\text{h}}$ objects overlapping ($\Delta R < 0.3$) with objects satisfying the electron or muon (loose) selection criteria described above are discarded.

| PF $\tau_{\text{h}}$ discriminators | Bool value |
|---|---|
| Loose selection: | |
| ByDecayModeFinding | Yes |
| ByMVA3TightElectronRejection | Yes |
| ByTightMuonRejection2 | Yes |
| ByVLooseCombinedIsolationDBSumPtCorr | Yes |
| Tight selection (in addition to the loose selection): | |
| ByTightCombinedIsolationDBSumPtCorr3Hits | Yes |

Table 4.4: Loose and tight $\tau_{\text{h}}$ selection criteria as used in this analysis. Access methods to these variables within the CMSSW framework are provided in Ref. [142].

### 4.6.4 Jet Selection Requirements

The jet selection is based on the AK5 PF jet collection as described in Section 4.4.4, and corresponds to the so-called *CMS loose jet selection criteria* [143] as detailed in Table 4.5. Jets are required to have $p_{\mathrm{T}} > 40$ GeV and $|\eta| < 3.0$, and those overlapping ($\Delta R < 0.5$) with the objects satisfying the electron, muon (loose), or $\tau_{\mathrm{h}}$ (loose) selection criteria described above are discarded.

| PF jet parameters | Criteria |
|---|---|
| Neutral hadron fraction | $< 0.99$ |
| Neutral EM fraction | $< 0.99$ |
| Number of constituents | $> 1$ |
| For jets with $\|\eta\| > 2.4$, additionally: | |
| Charged hadron fraction | $> 0$ |
| Charged multiplicity | $> 0$ |
| Charged EM fraction | $< 0.99$ |

Table 4.5: Jet selection criteria as used in this analysis. Access methods to these variables within the CMSSW framework are provided in Ref. [142].

## 4.7 Custom Variables

The following quantities are constructed using the physics objects described earlier.

- $\Delta R$ is the separation between two objects in the $\eta$-$\phi$ space, given as $\sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.
- $S_{\mathrm{T}}$ is the scalar $p_{\mathrm{T}}$ sum of all objects in the event, including muons, $\tau$ leptons, electrons, jets and $E_{\mathrm{T}}^{\mathrm{miss}}$. It provides a measure of the mass scale of the particles produced in the collision.
- $H_{\mathrm{T}}$ is the scalar $p_{\mathrm{T}}$ sum of jets in the event.
- $M(\ell_1, \ell_2)$ is the invariant mass of the given leptons. It is used in selecting resonant decays of Z bosons to muon or $\tau$ lepton pairs.
- $M_{\mathrm{T}}(\ell, E_{\mathrm{T}}^{\mathrm{miss}})$ is the transverse mass of the given lepton with the $E_{\mathrm{T}}^{\mathrm{miss}}$ present in the event. It is instrumental in the definition of various control regions.

$$M_{\mathrm{T}}(\ell, E_{\mathrm{T}}^{\mathrm{miss}}) = \sqrt{2\ p_{\mathrm{T}}^{\ell}\ E_{\mathrm{T}}^{\mathrm{miss}}(1 - \cos(\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \ell)))} \tag{4.5}$$

- $\Delta p_\zeta(\ell_1, \ell_2)$ is a measure of the collinearity of the two given lepton candidates with the $E_{\mathrm{T}}^{\mathrm{miss}}$ in the event, in the context of a two body decay where each of the decay products subsequently decays into visible and invisible particles. $\Delta p_\zeta$ is given in Eq. (4.6) where $\hat{\zeta}$ is the unit transverse vector bisecting $\vec{p}_{\mathrm{T}}^{\,\ell_1}$ and $\vec{p}_{\mathrm{T}}^{\,\ell_2}$. This parameter is instrumental in the selection of Z boson decays to $\tau$ lepton pairs.

$$\Delta p_\zeta = p_\zeta - 1.5\ p_\zeta^{vis} \tag{4.6a}$$

$$p_\zeta = (\vec{p}_{\mathrm{T}}^{\,\ell_1} + \vec{p}_{\mathrm{T}}^{\,\ell_2} + \vec{E}_{\mathrm{T}}^{\mathrm{miss}}) \cdot \hat{\zeta} \tag{4.6b}$$

$$p_\zeta^{vis} = (\vec{p}_{\mathrm{T}}^{\,\ell_1} + \vec{p}_{\mathrm{T}}^{\,\ell_2}) \cdot \hat{\zeta} \tag{4.6c}$$

- $\Delta R(\ell, j)_{\mathrm{min}}$ is the $\Delta R$ between the given muon or $\tau_{\mathrm{h}}$ candidate and the nearest AKF PF jet object (selected from the raw PF jet collection with no selection applied) present in the event with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 4.0$. Since both muon and $\tau$ lepton candidates are PF objects themselves, and therefore could be present in the PF jet collection, the associated muon-jet or $\tau$-jet has to be removed in order to prevent double-counting and artificially small $\Delta R$ values. If the lepton candidate is a $\tau_{\mathrm{h}}$, the associated $\tau$-jet is described as the PF jet within $\Delta R < 0.2$ of the $\tau$ candidate. In case of a muon, the associated muon-jet is either a PF jet within $\Delta R < 0.05$ of the muon, or a PF jet with muon energy fraction higher than 70% and within $\Delta R < 0.5$ of the muon. Additionally, all other PF jets which are within $\Delta R < 0.5$ of the muon candidate under consideration are corrected by a vectorial subtraction of the muon contribution. $\Delta R(\ell, j)_{\mathrm{min}}$ cuts are employed in lepton selection criteria to reduce the process dependency of lepton prompt and fake probabilities, by eliminating those with nearby jets which are likely to spoil the object isolation.

- $\widetilde{|\eta|}$ is the pseudo-rapidity defined as $\widetilde{|\eta|} = -\ln\tan(\bar{\theta}/2)$, where $\bar{\theta}$ is the average absolute polar angle of all electrons, muons, and hadronically decaying $\tau$ leptons in an event as measured from the beam axis in the lab frame, and is used as a measure of the event centrality. The $\mathrm{LQ}_3$ search is conducted in two channels of $\widetilde{|\eta|}$ in order to achieve higher signal sensitivity.

# Chapter 5

# Search for Leptoquarks

## 5.1 Analysis Strategy

The $LQ_3\overline{LQ_3}$ decays give rise to a variety of final states ranging from all-leptonic to multijet signatures due to the presence of top quark and $\tau$ lepton pairs. Same-sign $\mu\tau_h$ pairs, which are expected to originate from same-sign $W\tau$ pairs, are sought in this analysis. Such $LQ_3\overline{LQ_3}$ final states with same-sign $\mu\tau_h$ have a total branching fraction of 18.76%, where the visible constituents are either $\mu^{\pm}\tau_h^{\pm}\ell^{\mp}4j$ or $\mu^{\pm}\tau_h^{\pm}\ell^{\mp}\ell^{\mp}2j$ ($\ell$ stands for all three lepton generations, and $j$ for a quark jet). Figure 5.1 illustrates one of the $LQ_3\overline{LQ_3}$ decay chains with a same-sign $\mu\tau_h$ signature.

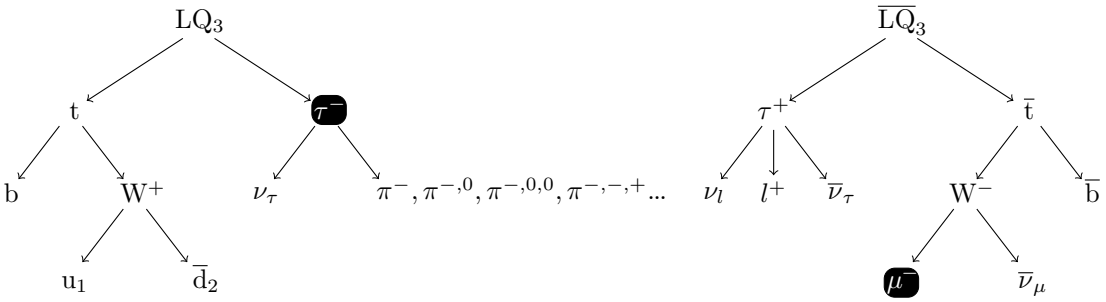


Figure 5.1: One of the $LQ_3\overline{LQ_3}$ decay chains with same-sign $\mu\tau_h$ signature. $l^{\pm}$ denotes either an electron or a muon, whereas u(d) is a generic label for up(down) type quarks.

The final state for this analysis is chosen as $\mu^{\pm}\tau_h^{\pm}$ and at least two jets. This analysis constitutes the first CMS search exclusively targeting the $\mu^{\pm}\tau_h^{\pm}$ signature at 8 TeV, as well as the first ever direct search for third generation leptoquarks in the $LQ_3 \to t\tau^-$ channel. Previous searches at hadron colliders have targeted LQs decaying into leptons and quarks of the first and second generations [144, 145, 146], or the third generation in

the $LQ_3 \to b\nu$ and $LQ_3 \to b\tau^-$ channels [48, 147, 148, 149, 150].

Since hadronic $\tau$ lepton and semi-leptonic W decays have intrinsic missing energy, a resonant search is not readily possible without extensive use of the simulated signal samples and assumptions on the decay kinematics. Instead, a less model dependent counting experiment is performed where an excess over the SM background contributions is sought. A selection with same-sign $\mu\tau_h$ requirement suppresses irreducible SM background contributions while mostly suffering from contributions due to misidentified lepton candidates. These misidentified lepton backgrounds are estimated using a data-driven method, whereas the irreducible contributions are estimated using the simulated samples. These aspects of the expected background composition drive the event selection strategy described below. A detailed description of the data and simulated samples used in this analysis is provided in Appendix A.

## 5.2 Event Selection

A same-sign $\mu\tau_h$ pair and at least two jets are required in each event. The muon and hadronic $\tau$ lepton candidates are required to satisfy the loose selection criteria as given in Sections 4.6.2 and 4.6.3, respectively. The jet objects must satisfy the selection requirements described in Section 4.6.4. If the event contains more than one pair of same-sign $\mu\tau_h$ candidates, the pair with the largest scalar $p_T$ sum is selected. The muon and hadronic $\tau$ lepton objects in this pair are tagged as $\mu_{LT}$ and $\tau_{LT}$, and are required to satisfy their respective tight selections as defined in Sections 4.6.2 and 4.6.3. A baseline selection, detailed in Table 5.1, is defined where events are required to have $\tau_{LT}$ $p_T > 35$ GeV and $S_T > 400$ GeV in order to reject bulk of the SM background contributions. The signal region is the subset of events within the final selection where both $\mu_{TL}$ and $\tau_{TL}$ satisfy their respective tight selection requirements.

Simulated samples normalized to the total integrated CMS luminosity are used to study the behavior of dominant backgrounds and optimize the event selection criteria for each leptoquark mass hypothesis. In order to exploit a feature of the signal model which places the $LQ_3$ contribution dominantly in the central region as shown in Fig. 5.2, the search is split into central ($\widetilde{|\eta|} < 0.9$) and forward ($\widetilde{|\eta|} \geq 0.9$) channels.

| | |
|---|---|
| Trigger = `HLT_Mu24_eta2p1` | if $25 < \mu_{\mathrm{LT}}\ p_{\mathrm{T}} \leq 45$ GeV |
| Trigger = `HLT_IsoMu24_eta2p1` | if $25 < \mu_{\mathrm{LT}}\ p_{\mathrm{T}} \leq 45$ GeV and $\mu_{\mathrm{LT}}$ passes tight selection |
| Trigger = `HLT_Mu40_eta2p1` | if $\mu_{\mathrm{LT}}\ p_{\mathrm{T}} > 45$ GeV |
| $N_\mu \geq 1$ | (loose selection) |
| $N_\tau \geq 1$ | (loose selection) |
| A same-sign muon - $\tau$ lepton pair: $\mu^{\pm}_{\mathrm{LT}}\tau^{\pm}_{\mathrm{LT}}$ | |
| $\tau_{\mathrm{LT}}\ p_{\mathrm{T}} > 35$ GeV | |
| $\Delta R(\mu_{\mathrm{LT}}, \mu_{\mathrm{HLT}}) < 0.15$ | (trigger matching) |
| $S_{\mathrm{T}} > 400$ GeV | |
| $N_{\mathrm{j}} \geq 2$ | |
| Veto if a $\mu^{\mp}\mu^{\pm}$ pair is present with $\lvert M(\mu^-, \mu^+) - M_{\mathrm{Z}}\rvert / M_{\mathrm{Z}} < 10\%$ | |

Table 5.1: Non-optimized baseline selection requirements. The muon candidate produced by the HLT algorithm satisfying any of the listed triggers is denoted as $\mu_{\mathrm{HLT}}$. Mass of the Z boson is given as $M_{\mathrm{Z}}$.

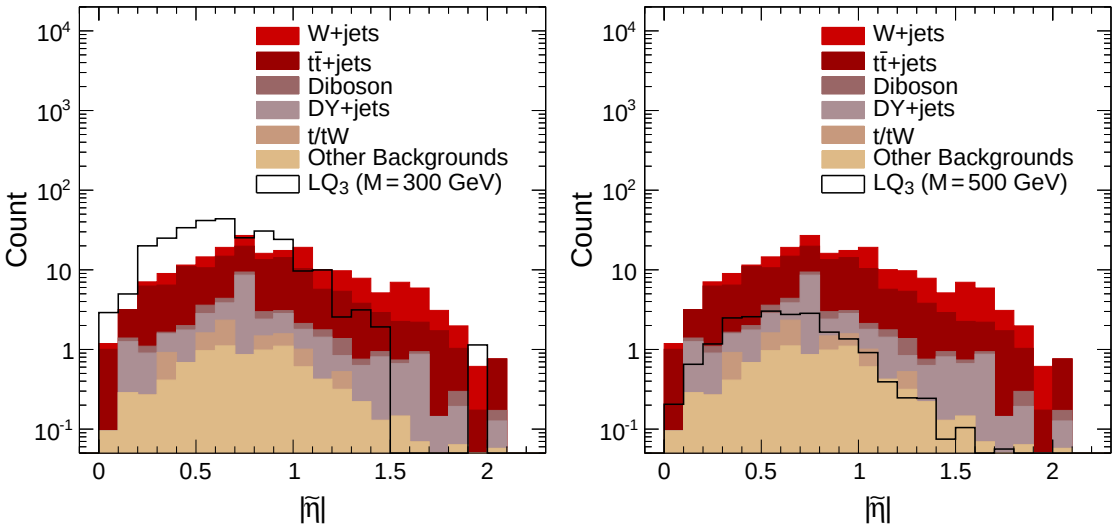


Figure 5.2: The distributions of simulated SM background and $\mathrm{LQ}_3$ signal events for 300 GeV (left) and 500 GeV (right) $\mathrm{LQ}_3$ mass hypotheses normalized to 19.5 fb$^{-1}$ as a function of $|\widetilde{\eta}|$ using the baseline final selection given in Table 5.1. Reproduced from Ref. [142].

Furthermore, leptoquark mass dependent optimizations are carried out in the central channel on the 2D $\tau_{\mathrm{LT}}\ p_{\mathrm{T}}$ - $S_{\mathrm{T}}$ plane, since these parameters are observed to have high signal-to-background discrimination power as illustrated in Fig. 5.3. The optimizations, as exemplified in Fig. 5.4, are accomplished by maximizing the figure of merit given in Eq. (5.1) [151] where $\varepsilon$ is the signal efficiency and $B$ is the number of background events.

$$\chi(\tau_{\mathrm{LT}}\ p_{\mathrm{T}},\ S_{\mathrm{T}}) = \frac{\varepsilon(\tau_{\mathrm{LT}}\ p_{\mathrm{T}},\ S_{\mathrm{T}})}{1 + \sqrt{B(\tau_{\mathrm{LT}}\ p_{\mathrm{T}},\ S_{\mathrm{T}})}} \tag{5.1}$$

The $(\tau_{\mathrm{LT}}\ p_{\mathrm{T}},\ S_{\mathrm{T}})$ values for the optimized selections and the corresponding efficiencies as a function of the $\mathrm{LQ}_3$ mass are presented in Table 5.2.

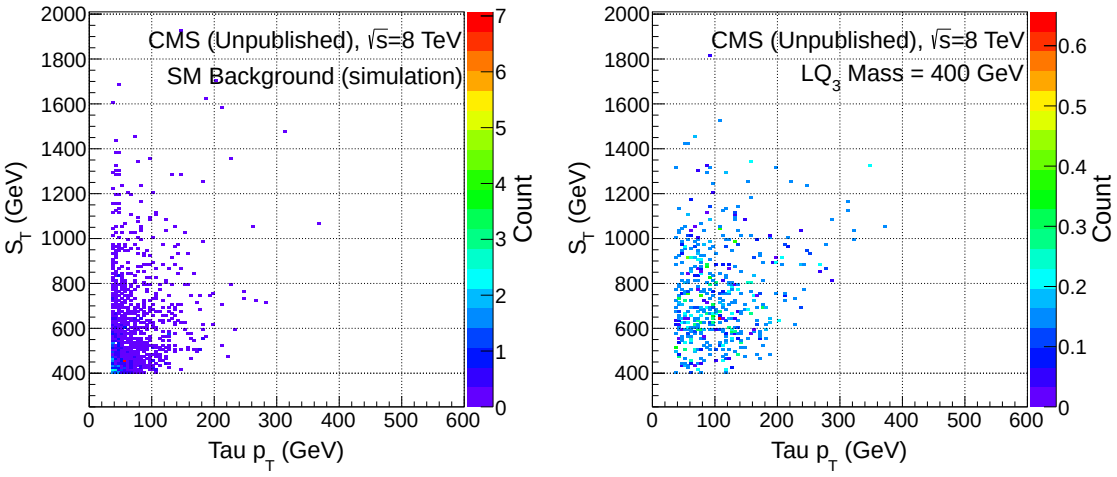


Figure 5.3: 2D distribution of SM background (left) and $LQ_3$ signal (right) events in the $\tau_{LT}$ $p_T$ - $S_T$ plane, using the baseline selection with $S_T > 400$ GeV, $N_j \geq 2$ and a same-sign $\mu\tau_h$ pair. The signal distribution is shown for $M_{LQ_3} = 400$ GeV [152].

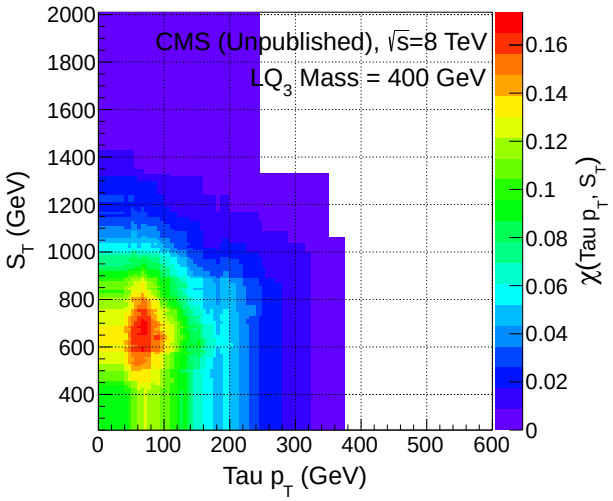


Figure 5.4: 2D distribution of the figure of merit, $\chi = \epsilon/(1+\sqrt{B})$, in the $\tau_{LT}$ $p_T$ - $S_T$ plane for $M_{LQ_3} = 400$ GeV, calculated for each point $(S_T^*, p_T^*)$ using the distributions in Fig. 5.3 with a rectangular cut of $S_T > S_T^*$ and $p_T > p_T^*$ [152].

## 5.3 Signal-depleted Selection

A background dominated sample is created by imposing an upper limit on the jet multiplicity, $N_j$. This yields a signal-depleted selection orthogonal to the signal region, which is instrumental in commissioning the normalization factors for the simulated samples, as well as pileup and trigger weights. Signal-depleted selection requirements are listed in Table 5.3, and plots displaying data - MC agreement in $\tau_{LT}$ $p_T$ and $S_T$ variables are presented in Fig. 5.5. The agreement is found to be 20%, and is assigned as the normal-

| $M_{LQ_3}$ (GeV) | $\tau_{LT}$ $p_T$ (GeV) | $S_T$ (GeV) | $N_{Bkg}^{pp,SS}$ ± stat. | $N_{Bkg}^{pp,OS}$ ± stat. | Total $N_{Bkg}$ ± stat. | $N_{LQ_3}$ ± stat. | $\epsilon_{LQ_3}$ (%) |
|---|---|---|---|---|---|---|---|
| | | | Central channel: $\widetilde{\lvert\eta\rvert} < 0.9$ | | | | |
| 200 | 35 | 410 | 6.04±0.53 | 2.47±0.82 | 101.04±8.01 | 52.61±20.55 | 0.08 |
| 250 | 35 | 410 | 6.04±0.53 | 2.47±0.82 | 101.04±8.01 | 251.88±24.41 | 1.31 |
| 300 | 50 | 470 | 3.39±0.40 | 0.82±0.32 | 28.82±2.67 | 153.47±11.08 | 2.22 |
| 350 | 50 | 490 | 3.22±0.39 | 0.82±0.32 | 25.63±2.57 | 92.44±5.56 | 3.29 |
| 400 | 65 | 680 | 0.91±0.20 | 0.0047±0.0034 | 2.11±0.45 | 28.36±2.07 | 2.27 |
| 450 | 65 | 700 | 0.78±0.18 | 0.0034±0.0032 | 1.57±0.35 | 17.27±1.10 | 2.90 |
| 500 | 65 | 770 | 0.47±0.15 | 0.00025±0.00021 | 0.93±0.25 | 9.76±0.59 | 3.25 |
| 550 | 65 | 800 | 0.38±0.14 | 0.00025±0.00021 | 0.61±0.18 | 6.13±0.34 | 3.89 |
| 600 | 65 | 850 | 0.20±0.08 | 0.00025±0.00021 | 0.41±0.14 | 3.61±0.19 | 4.20 |
| 650 | 65 | 850 | 0.20±0.08 | 0.00025±0.00021 | 0.41±0.14 | 2.19±0.11 | 4.54 |
| 700 | 85 | 850 | 0.12±0.07 | 0.00025±0.00021 | 0.27±0.12 | 1.28±0.07 | 4.60 |
| 750 | 85 | 850 | 0.12±0.07 | 0.00025±0.00021 | 0.27±0.12 | 0.82±0.04 | 5.01 |
| 800 | 85 | 850 | 0.12±0.07 | 0.00025±0.00021 | 0.27±0.12 | 0.51±0.03 | 5.19 |
| | | | Forward channel: $\widetilde{\lvert\eta\rvert} \geq 0.9$ | | | | |
| 200 | 35 | 410 | 3.03±0.36 | 1.17±0.37 | 81.50±5.34 | – | – |
| 250 | 35 | 410 | 3.03±0.36 | 1.17±0.37 | 81.50±5.34 | 50.21±10.53 | 0.26 |
| 300 | 50 | 470 | 1.46±0.24 | 0.31±0.19 | 25.08±2.47 | 33.42±5.23 | 0.48 |
| 350 | 50 | 490 | 1.37±0.23 | 0.31±0.19 | 19.94±2.10 | 18.45±2.51 | 0.66 |
| 400 | 65 | 680 | 0.55±0.15 | 0.16±0.13 | 3.46±0.77 | 6.11±0.95 | 0.49 |
| 450 | 65 | 700 | 0.55±0.15 | 0.16±0.13 | 3.05±0.73 | 3.84±0.54 | 0.64 |
| 500 | 65 | 770 | 0.50±0.14 | 0.03±0.03 | 2.18±0.53 | 1.61±0.24 | 0.54 |
| 550 | 65 | 800 | 0.42±0.13 | 0.0±0.0 | 2.07±0.53 | 1.15±0.15 | 0.73 |
| 600 | 65 | 850 | 0.27±0.10 | 0.0±0.0 | 1.59±0.46 | 0.56±0.08 | 0.65 |
| 650 | 65 | 850 | 0.27±0.10 | 0.0±0.0 | 1.59±0.46 | 0.29±0.04 | 0.60 |
| 700 | 85 | 850 | 0.14±0.06 | 0.0±0.0 | 1.18±0.42 | 0.18±0.03 | 0.65 |
| 750 | 85 | 850 | 0.14±0.06 | 0.0±0.0 | 1.18±0.42 | 0.13±0.02 | 0.79 |
| 800 | 85 | 850 | 0.14±0.06 | 0.0±0.0 | 1.18±0.42 | 0.08±0.01 | 0.81 |

Table 5.2: Optimization results as conducted in MC, using the baseline final selection given in Table 5.1. $N_{Bkg}^{pp,SS}$ is the irreducible same-sign prompt-prompt background, whereas $N_{Bkg}^{pp,OS}$ is the prompt-prompt contribution due to $\tau$ lepton charge mismeasurement. It can be seen that events with fake leptons constitute the major background, especially in selections optimized for $M_{LQ_3} < 400$ GeV. The expected signal efficiency is denoted as $\epsilon_{LQ_3}$, where the denominator is defined as the total number of expected $LQ_3$ signal events at 8 TeV with a same-sign $\mu\tau_h$ pair and the numerator is defined as the number of expected signal events, $N_{LQ_3}$, both calculated with the given set of selection requirements.

ization systematic uncertainty for the non-rare SM contributions in the signal region. It must be noted that the fake lepton contributions are estimated via data-driven methods in the signal selection.

| | | |
|---|---|---|
| Trigger = `HLT_IsoMu24_eta2p1` | | |
| | $N_\tau \geq 1$ | (loose selection) |
| | $N_\mu \geq 1$ | (loose selection) |
| | A same-sign muon - $\tau$ lepton pair: $\mu_{\mathrm{LT}}$ - $\tau_{\mathrm{LT}}$ | |
| | $\tau_{\mathrm{LT}}$ and $\mu_{\mathrm{LT}}$ satisfies respective tight selections | |
| | $\Delta R(\mu_{\mathrm{LT}}, \mu_{\mathrm{HLT}}) < 0.15$ | (trigger matching) |
| | $N_{\mathrm{j}} \leq 1$ | (to suppress $\mathrm{LQ}_3$ signal events) |
| | $M_{\mathrm{T}}(\mu_{LT}, E_{\mathrm{T}}^{\mathrm{miss}}) > 40$ GeV | (to suppress QCD multijet contributions) |

Table 5.3: Signal-depleted sample selection requirements. The muon candidate produced by the HLT algorithm satisfying the trigger requirement is denoted as $\mu_{\mathrm{HLT}}$.

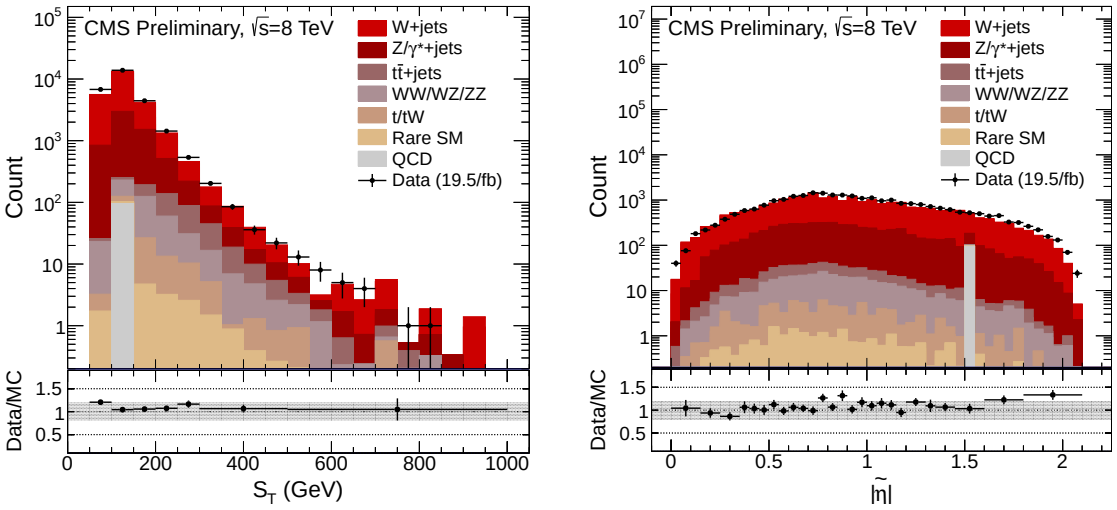


Figure 5.5: Comparison between data and simulation in the $S_{\mathrm{T}}$ (left) and $|\widetilde{\eta}|$ (right) distributions using the signal-depleted selection of events. Rare SM contributions are from processes listed in Table A.3. The gray region in the ratio plot represents the $\pm 20\%$ band [153].

# Chapter 6

# Backgrounds

Within the framework of this analysis, prompt leptons are considered as those that come from the decays of W bosons, Z bosons or $\tau$ leptons, and are usually well isolated, whereas misidentified (fake) leptons either originate from semi-leptonic heavy flavor decays within jets or are simply misreconstructed genuine jets, and in both cases are generally not isolated. The same-sign dilepton requirement in the final selection yields a background particularly consisting of events with misidentified leptons (especially jets misreconstructed as hadronically decaying $\tau$ leptons). Additionally, there are less significant contributions from events with lepton charge mismeasurements and from SM processes with prompt same-sign dileptons. Primary contributors to each of these backgrounds are given in Table 6.1.

| Background | Primary processes | Method |
|---|---|---|
| Single mis-ID | Leptonic DY($Z \to \mu\mu$)+jets and W($\to \mu\nu$)+jets, $t\bar{t}$+jets | Data-driven |
| Double mis-ID | QCD multijet, other all-hadronic processes | Data-driven |
| Charge mis-ID | Leptonic $t\bar{t}$+jets, DY($Z \to \tau\tau \to \mu\tau_h$)+jets, tW | MC |
| Irreducible | WZ, ZZ, $t\bar{t}W$, $t\bar{t}Z$, $W^{\pm}W^{\pm}qq$ | MC |

Table 6.1: Primary background sources and the associated background estimation techniques.

Single-misidentified and double-misidentified lepton contributions are estimated via a matrix method (loose-to-tight extrapolation method) [154] which is described in the following section. Charge-misidentification contributions mainly come from $\tau$ lepton charge mismeasurements and is predicted directly from the simulated samples alongside with contributions due to processes with prompt same-sign dileptons.

## 6.1 Backgrounds due to Misidentified Leptons

### 6.1.1 Matrix Method: Description

Loose-to-tight extrapolation method (LTEM) is a data-driven background estimation method, which relies on the assumption that the probabilities with which prompt and fake leptons satisfy a tight lepton selection given that they satisfy a loose lepton selection, $p$ and $f$ respectively, can be universally described as a function of the lepton and event topology dependent parameters. This assumption allows the measurement of these probabilities in signal-depleted control regions, and their application to the signal region.

For a given set of selection requirements, four combinations are defined based on the selection quality of the chosen same-sign dilepton pair. Events in which both leptons satisfy the tight selection requirements are classified as TT events, whereas those with both leptons failing the tight selection while satisfying the loose selection requirements are classified as LL events. Similarly, events with only one lepton candidate satisfying the tight selection and with the other lepton satisfying the loose selection but failing the tight selection requirements are labeled as TL or LT events. The TT combination constitute the signal region, whereas TL, LT, and TT combinations are used in the estimation of backgrounds.

These events can also be categorized on the basis of true lepton origins, forming the double-fake, double-prompt and single-fake combinations, denoted as FF, PP, PF and FP, respectively. These two categorizations are represented in Fig. 6.1 with the corresponding number of events in each combination, such that $N_{\mathrm{LL}}+N_{\mathrm{LT}}+N_{\mathrm{TL}}+N_{\mathrm{TT}} = N_{\mathrm{FF}} + N_{\mathrm{FP}} + N_{\mathrm{PF}} + N_{\mathrm{PP}}$. The LTEM provides a handle in estimating the sizes of these origin-based subsets that are of actual interest to the analysis, by using the selection-based subsets that can be measured.

The conversion from selection-based categorization to origin-based one can be represented as a multiplication by a transformation matrix as given in Eq. (6.1) where $\widehat{f_i} = 1 - f_i$ and $\widehat{p_i} = 1 - p_i$. The subscripts refer to lepton-1 and lepton-2 as used in Fig. 6.1.

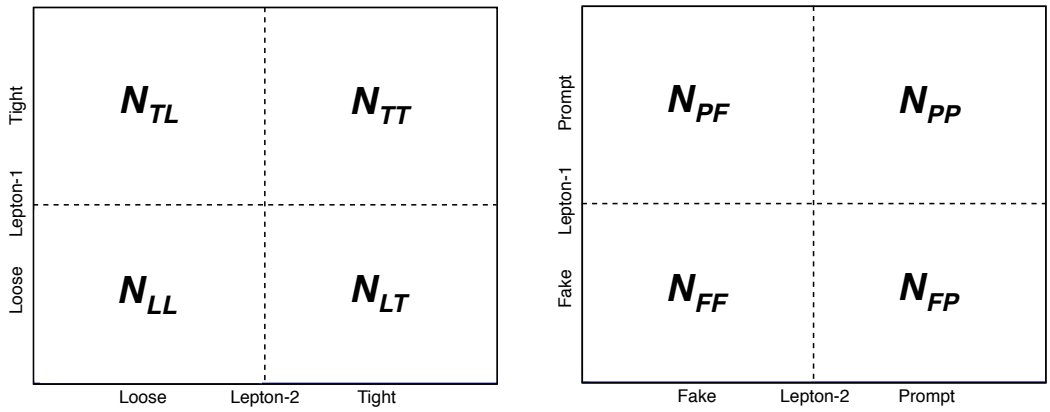


Figure 6.1: Categorizations of the dilepton events based on the selection (left) and origin (right) properties of the dilepton pair.

$$\begin{pmatrix} N_{\mathrm{LL}} \\ N_{\mathrm{LT}} \\ N_{\mathrm{TL}} \\ N_{\mathrm{TT}} \end{pmatrix} = \begin{pmatrix} \widehat{f_1}\cdot\widehat{f_2} & \widehat{f_1}\cdot\widehat{p_2} & \widehat{p_1}\cdot\widehat{f_2} & \widehat{p_1}\cdot\widehat{p_2} \\ \widehat{f_1}\cdot f_2 & \widehat{f_1}\cdot p_2 & \widehat{p_1}\cdot f_2 & \widehat{p_1}\cdot p_2 \\ f_1\cdot\widehat{f_2} & f_1\cdot\widehat{p_2} & p_1\cdot\widehat{f_2} & p_1\cdot\widehat{p_2} \\ f_1\cdot f_2 & f_1\cdot p_2 & p_1\cdot f_2 & p_1\cdot p_2 \end{pmatrix} \begin{pmatrix} N_{\mathrm{FF}} \\ N_{\mathrm{FP}} \\ N_{\mathrm{PF}} \\ N_{\mathrm{PP}} \end{pmatrix} \tag{6.1}$$

Assuming $f_i \neq p_i$, this transformation matrix can be inverted for the calculation of origin-based contributions solely based on the experimentally measurable quantities as presented in Eq. (6.2). The total number of events with fake leptons in the TT combination is the sum of single- and double-fake contributions, and can be expressed as in Eq. (6.3) where $N_{\mathrm{FF}}$, $N_{\mathrm{FP}}$, and $N_{\mathrm{PF}}$ are derived quantities.

$$\begin{pmatrix} N_{\mathrm{FF}} \\ N_{\mathrm{FP}} \\ N_{\mathrm{PF}} \\ N_{\mathrm{PP}} \end{pmatrix} = \frac{1}{(p_1 - f_1)(p_2 - f_2)} \begin{pmatrix} p_1\cdot p_2 & -p_1\cdot\widehat{p_2} & -\widehat{p_1}\cdot p_2 & \widehat{p_1}\cdot\widehat{p_2} \\ -p_1\cdot f_2 & p_1\cdot\widehat{f_2} & \widehat{p_1}\cdot f_2 & -\widehat{p_1}\cdot\widehat{f_2} \\ -f_1\cdot p_2 & f_1\cdot\widehat{p_2} & \widehat{f_1}\cdot p_2 & -\widehat{f_1}\cdot\widehat{p_2} \\ f_1\cdot f_2 & -f_1\cdot\widehat{f_2} & -\widehat{f_1}\cdot f_2 & \widehat{f_1}\cdot\widehat{f_2} \end{pmatrix} \begin{pmatrix} N_{\mathrm{LL}} \\ N_{\mathrm{LT}} \\ N_{\mathrm{TL}} \\ N_{\mathrm{TT}} \end{pmatrix} \tag{6.2}$$

$$N_{\mathrm{TT}}^{\text{Fake Bkg}} = f_1\ f_2\ N_{\mathrm{FF}} + f_1\ p_2\ N_{\mathrm{FP}} + p_1\ f_2\ N_{\mathrm{PF}} \tag{6.3}$$

If the prompt and fake probabilities assume constant values, Eqs. (6.2) and (6.3) can be used in their present form for the estimation of fake lepton containing backgrounds given the $p_i$ and $f_i$ measurements as well as the distribution of events in the selection-based categorization. Equivalently, these can be read as a recipe for applying individual weights to each selected event based on its selection-based classification and as a function of the properties of the dilepton pair (such as $p_{\mathrm{T}}$ and $\eta$) chosen for the LTEM in order to calculate the event's contribution to single- and double-fake lepton categories. Since

the prompt and fake probabilities for the leptons of interest are found to be variables, the latter approach is adopted in this analysis.

### 6.1.2 Matrix Method: Implementation

Isolation is a powerful discriminant between prompt and fake leptons, and primarily relies on the fact that fakes, especially those originating from within jets, have more activity in the immediate vicinity than prompt leptons do. Given that the only difference between loose and tight muon and $\tau$ lepton selections as described in Section 4.6 is isolation, measurement of the prompt and fake probabilities amounts to describing the distribution of average activity in the detector around the chosen lepton object.

Parametrization of prompt and fake probabilities only in terms of lepton-specific variables, such as $p_{\mathrm{T}}$ and $|\eta|$, is essentially assigning an average probability for each fake or prompt lepton to satisfy the isolation requirement given its energy and location in the detector. This approach is valid as long as the measurements of these probabilities and the final analysis itself are conducted in similar topologies. However, given that both the signal model under consideration and $\mathrm{t\bar{t}}$+jets events, which constitute a major background for this analysis, have particularly high hadronic activity unlike the DY+jets or W+jets processes commonly used for these studies, it is beneficial to utilize a parametrization that would minimize the process dependence and any resulting systematic bias.

The total energy of all the activity around the lepton, $\mathcal{I}$, can be described as the sum of three mostly independent components. The primary one is the contribution from the lepton origin, $\mathcal{I}_{\mathrm{O}}$. This quantity is essentially zero for prompt leptons, whereas for fakes it mostly represents the contribution of the jet from which the fake lepton originates. The second term, $\mathcal{I}_{\mathrm{J}}$, is due to the presence of other jets in the event and is a function of the event topology. The overall occupancy of the detector due to pileup interactions as well as contributions from underlying event activity can be thought as a third term, $\mathcal{I}_{\mathrm{PU}}$. Hence:

$$\mathcal{I} = \mathcal{I}_{\mathrm{O}} + \mathcal{I}_{\mathrm{J}} + \mathcal{I}_{\mathrm{PU}} \ . \tag{6.4}$$

If an exact description of $\mathcal{I}$ as given in Eq. (6.4) is used to parametrize prompt and fake probabilities, these quantities will be identical since $\mathcal{I}$ will fully describe the contents of the isolation cone of the lepton candidate; i.e. if the isolation requirement is an absolute cut for the total energy of particles in an isolation cone around the lepton

candidate ($\mathcal{I}_{\text{threshold}}$), any lepton with $\mathcal{I} < \mathcal{I}_{\text{threshold}}$ will satisfy the tight lepton selection (and vice versa) regardless of the lepton being a prompt or a fake one. However, since the LTEM relies on this difference, any description of the isolation quantity around a lepton object should ignore the $\mathcal{I}_{\text{O}}$ term and focus on the external terms. Since lepton selections are designed to be pileup safe, the $\mathcal{I}_{\text{PU}}$ term can be neglected, and the external contributions to the lepton isolation can be approximated as a function of $\mathcal{I}_{\text{J}}$.

A detailed description of $\mathcal{I}_{\text{J}}$ would include a proximity weighted scalar sum of all jet energies around the reconstructed lepton candidate as a function of the lepton isolation cone and the jet shapes in the $\eta$-$\phi$ plane. However, due to difficulties in verifying the details of jet modeling in simulated samples, a simpler approach is adopted in this analysis, where the prompt and fake probabilities for muons and $\tau$ leptons are parametrized in terms of their distance to the nearest external jet, $\Delta R(\ell, j)_{\text{min}}$. This is illustrated in Fig. 6.2.

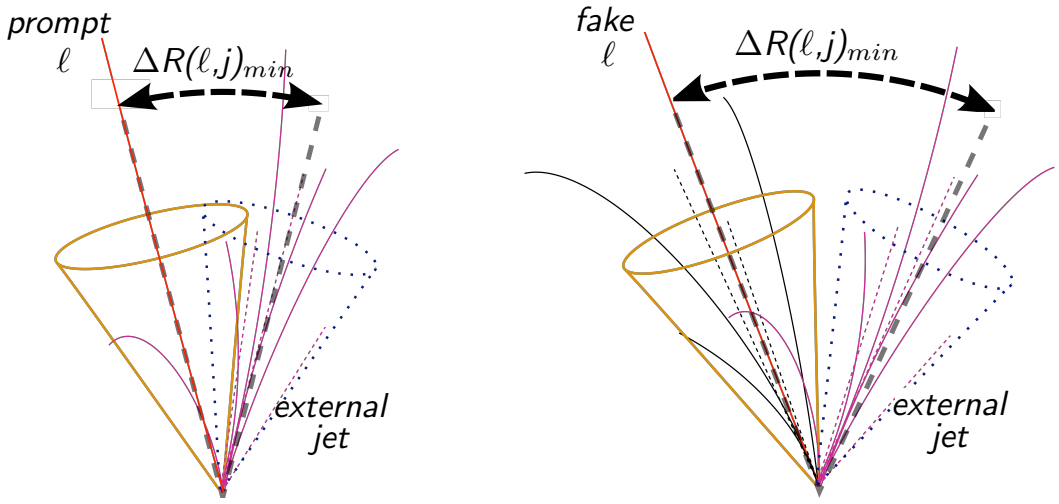


Figure 6.2: Schematic view of contributions to the lepton isolation cone by a nearby external jet for a prompt (left) and a fake (right) lepton. The lepton trajectories are denoted by solid red lines and the lepton isolation cones are denoted by solid orange cones. The dotted blue cones represent the nearby external jet. Black and magenta lines represent the particle trajectories belonging to the jet from which the fake lepton originates and the nearby external jet, respectively.

### 6.1.3 Measurement of the Prompt and Misidentification Probabilities

Determination of the prompt and fake probabilities in data and simulated samples amounts to computing the probability of probe objects passing the tight selection criteria given that they satisfy the loose selection criteria. In data, choosing a lepton object to

conduct these measurements relies on creating a sample enriched in a particular process where a tag-and-probe [155] method is used, and contributions of the undesired lepton type (fakes if probes are taken to be prompt, and vice versa) are subtracted using simulated samples. In contrast, any process can be studied in MC due to the availability of generator level information for selecting a prompt or a fake probe object, and these probabilities are simply the ratio of tight probes over the loose ones. These probabilities can then be described as:

$$\begin{aligned} p^{\text{Data}} &= (N^{\text{Data}}_{\text{tight}} - N^{\text{MC}}_{\text{tight fakes}})/(N^{\text{Data}}_{\text{loose}} - N^{\text{MC}}_{\text{loose fakes}}), \\ f^{\text{Data}} &= (N^{\text{Data}}_{\text{tight}} - N^{\text{MC}}_{\text{tight prompts}})/(N^{\text{Data}}_{\text{loose}} - N^{\text{MC}}_{\text{loose prompts}}), \\ p^{\text{MC}} &= N^{\text{MC}}_{\text{tight prompts}}/N^{\text{MC}}_{\text{loose prompts}}, \\ f^{\text{MC}} &= N^{\text{MC}}_{\text{tight fakes}}/N^{\text{MC}}_{\text{loose fakes}}. \end{aligned} \tag{6.5}$$

A normalization uncertainty of 20% is applied on the subtracted MC predictions as supported by the general level of data - MC agreement in the signal-depleted selection plots given in Fig. 5.5, and in the prompt or fake probe enriched selection plots presented below. These systematic uncertainties are propagated to the $p^{\text{Data}}$ and $f^{\text{Data}}$ calculations given above.

The muon and $\tau$ lepton prompt probabilities are measured in $\text{Z}(\rightarrow \mu\mu)$+jets and $\text{Z}(\rightarrow \tau\tau \rightarrow \mu\tau_{\text{h}})$+jets enhanced data events, respectively, and in simulated $\text{t}\bar{\text{t}}$+jets, W+jets, and $\text{LQ}_3$ events where the probe muon or hadronically decaying $\tau$ lepton is chosen by requiring a match ($\Delta R \leq 0.15$) between the reconstructed candidate and the corresponding generator level object of the same kind. For the $\tau$ lepton fake probability measurements, a $\text{W}(\rightarrow \mu\nu)$+jets enriched dataset with an additional $\tau_{\text{h}}$ candidate is used, whereas a QCD multijet enhanced dataset with a single muon candidate is used for the muon fake probability measurements. In simulated samples, the $\tau$ lepton fake probability measurement is conducted in W+jets, $\text{t}\bar{\text{t}}$+jets, and $\text{LQ}_3$ samples, and the muon fake probability measurement is conducted in QCD multijet, $\text{t}\bar{\text{t}}$+jets, and $\text{LQ}_3$ samples where the probe muon or $\tau$ lepton is chosen by requiring a mismatch ($\Delta R > 0.15$) between the reconstructed candidate and the closest generator level object of the same kind. These simulated processes are chosen to reflect varying levels of hadronic activity such that variations of the prompt and fake probabilities can be studied.

Each $\tau$ lepton prompt and fake probability ($p_\tau$ and $f_\tau$), and muon prompt probability ($p_\mu$) measurement is conducted as a function of the associated object $p_{\text{T}}$ in two regions

of $|\eta|$ and three regions of $\Delta R(\ell, j)_{\text{min}}$, whereas muon fake probability ($f_\mu$) is measured in two regions of $|\eta|$ and two regions of $S_{\text{T}}$. Additionally, $S_{\text{T}}$ dependent corrections are considered for the $\tau$ lepton prompt and fake probability measurements, which are described by a linear slope parameter ($c_{S_{\text{T}}}$) for events with $S_{\text{T}} > 400$ GeV. These are based on measurements in simulated $\text{t}\bar{\text{t}}$+jets samples since the corresponding data samples do not have enough events in the high $S_{\text{T}}$ region. In contrast, no $S_{\text{T}}$ dependent corrections are applied to muon prompt probability measurements as they are dominated by process dependent variations.

A summary of the data and MC samples, as well as the parametrizations used in measuring the muon and $\tau$ lepton prompt and fake probabilities is given in Table 6.2. The details of the prompt and fake muon and $\tau$ lepton enriched selections are presented in Tables 6.3 and 6.4. Figures 6.3 - 6.6 illustrate these probability measurements in data and simulated samples for selected $\Delta R(\ell, j)_{\text{min}}$, $S_{\text{T}}$, and object $|\eta|$ bins.

| | Data | MC | | | Parametrization | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | $p_{\text{T}}$ | $\lvert\eta\rvert$ | $\Delta R(\ell, j)_{\text{min}}$ | $S_{\text{T}}$ |
| $p_\tau$ | $\text{Z}(\rightarrow \tau\tau \rightarrow \mu\tau_{\text{h}})$+jets | W+jets | $\text{t}\bar{\text{t}}$+jets | $\text{LQ}_3$ | ✓ | ✓ | ✓ | ✓ |
| $f_\tau$ | $\text{W}(\rightarrow \mu\nu)$+jets | W+jets | $\text{t}\bar{\text{t}}$+jets | $\text{LQ}_3$ | ✓ | ✓ | ✓ | ✓ |
| $p_\mu$ | $\text{Z}(\rightarrow \mu\mu)$+jets | W+jets | $\text{t}\bar{\text{t}}$+jets | $\text{LQ}_3$ | ✓ | ✓ | ✓ | × |
| $f_\mu$ | QCD multijet | QCD multijet | $\text{t}\bar{\text{t}}$+jets | $\text{LQ}_3$ | ✓ | ✓ | × | ✓ |

Table 6.2: Samples and parametrizations used to measure the muon and $\tau$ lepton prompt and fake probabilities. $S_{\text{T}}$ based binning is preferred for the muon fake probability measurement instead of $\Delta R(\ell, j)_{\text{min}}$ based binning, since the signal events mostly populate the high $S_{\text{T}}$ region and no strong dependence is observed as a function of $\Delta R(\ell, j)_{\text{min}}$ due to the large uncertainties on the muon fake probability measurements.

For a given set of prompt or fake probabilities in a given parameter space as measured in data and in various MC samples, a single global value needs to be constructed. Firstly, a mean MC value is defined as the average of the systematically highest probability yielding process, $r_{\text{max}}$, and the systematically lowest one, $r_{\text{min}}$. This mean MC value, $\overline{\text{MC}}$, and the associated systematic uncertainty, $\Delta$D, for each prompt and fake probabilities can be expressed as:

$$\begin{aligned} \overline{\text{MC}}(r_{\text{max}},\ r_{\text{min}}) &= \frac{r_{\text{max}} + r_{\text{min}}}{2}, \\ \Delta\text{D}(r_{\text{max}},\ r_{\text{min}}) &= \frac{(r_{\text{max}} + \sigma_{r_{\text{max}}}) - (r_{\text{min}} - \sigma_{r_{\text{min}}})}{2}, \end{aligned} \tag{6.6}$$

where individual uncertainties on $r_{\text{max}}$ and $r_{\text{min}}$ are denoted as $\sigma_{r_{\text{max}}}$ and $\sigma_{r_{\text{min}}}$, respectively.

Secondly, a correction factor to account for any systematic difference between data and MC is applied. This factor is defined as the ratio of the data measurement over the comparable probability measurement in MC:

$$\gamma_{p_\tau} = \frac{p_\tau^{\text{Data, Z+jets}}}{p_\tau^{\text{MC, W+jets}}}, \qquad \gamma_{f_\tau} = \frac{f_\tau^{\text{Data, W+jets}}}{f_\tau^{\text{MC, W+jets}}}, \qquad \gamma_{p_\mu} = \frac{p_\mu^{\text{Data, Z+jets}}}{p_\mu^{\text{MC, W+jets}}}. \tag{6.7}$$

Thus, the resultant muon and $\tau$ lepton prompt and fake probabilities can be expressed as in Eq. (6.8), where $\overline{\text{MC}}$s (for $p_\tau$, $f_\tau$ and $p_\mu$) and the $S_\text{T}$ based corrections (for $p_\tau$ and $f_\tau$) are explicitly shown for clarity.

$$\begin{aligned}
p_\tau &= \gamma_{p_\tau} \left( \frac{p_\tau^{\text{MC, W+jets}} + p_\tau^{\text{MC, t}\bar{\text{t}}\text{+jets}}}{2} - \max(0,\ S_\text{T} - 400\ \text{GeV}) |c_{S_\text{T}}| \right) \\
f_\tau &= \gamma_{f_\tau} \left( \frac{f_\tau^{\text{MC, W+jets}} + f_\tau^{\text{MC, t}\bar{\text{t}}\text{+jets}}}{2} - \max(0,\ S_\text{T} - 400\ \text{GeV}) |c_{S_\text{T}}| \right) \\
p_\mu &= \gamma_{p_\mu} \left( \frac{p_\mu^{\text{MC, W+jets}} + p_\mu^{\text{MC, LQ}_3}}{2} \right) \\
f_\mu &= f_\mu^{\text{Data, QCD}}
\end{aligned} \tag{6.8}$$

It must be noted that the $\gamma$ correction factors and the probabilities, $f_i$s and $p_i$s, are functions of some combination of $\Delta R(\ell, j)_\text{min}$, $S_\text{T}$, object $p_T$, and object $|\eta|$ parameters. Since the relative uncertainties on the data based muon fake probability measurements are $\mathcal{O}(100\%)$, only the data based $f_\mu$ values are used in the analysis. The MC measurements are solely treated as a cross-check and no data-MC correction factor is applied.

The resultant muon prompt probabilities vary from $(70 \pm 3)\%$ to $(95 \pm 3)\%$ for low and high $p_\text{T}$ muons, whereas $\tau$ lepton prompt probabilities are around $(60 \pm 6)\%$. The muon and $\tau$ lepton misidentification probabilities are measured to be about $(1 \pm 1)\%$ and $(14 \pm 2)\%$, respectively. The complete lists of muon and $\tau$ lepton prompt and fake probability measurements in data and MC are provided in Tables 6.5 and 6.6. The $S_\text{T}$ based corrections to $\tau$ lepton prompt and fake probabilities are given in Table 6.7.

| |
|---|
| Prompt muon selection in data ($Z(\rightarrow \mu\mu)$+jets) : |
| Trigger = `HLT_IsoMu24_eta2p1` |
| $N_\mu = 2$ (loose selection), $Q_{\mu_1} = -1$, $Q_{\mu_2} = +1$ |
| $\Delta R(\mu_1, \mu_{\mathrm{HLT}}) < 0.15$ (trigger matching), $\mu_1$ passes tight selection. |
| $\lvert(M(\mu_1, \mu_2) - M_{\mathrm{Z}})/M_{\mathrm{Z}}\rvert < 10\%$ |
| $\mu_2$ is the probe muon. |
| $N_{\mathrm{e}} = 0$ |
| $N_{\mathrm{j}} \geq 1$ |
| $E_{\mathrm{T}}^{\mathrm{miss}} < 50$ GeV |
| Prompt muon selection in MC : |
| $N_\mu \geq 1$ (loose selection) |
| Leading $p_{\mathrm{T}}$ muon is the probe muon. |
| $S_{\mathrm{T}} > 100$ GeV |
| $\Delta R(\mu, \mu_{\mathrm{gen}}) \leq 0.15$ |
| Fake muon selection in data (QCD multijet) : |
| Trigger = `HLT_Mu40_eta2p1` |
| $N_\mu = 1$ (loose selection with $p_{\mathrm{T}} > 45$) |
| Veto event if there is a second loose muon with $p_{\mathrm{T}} > 25$ |
| $\Delta R(\mu, \mu_{\mathrm{HLT}}) < 0.15$ (trigger matching) |
| $M_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}}) < 20$ GeV |
| $\Delta R(\mu, j)_{\mathrm{max}} > \pi$ |
| $N_\tau = 0$ (loose selection) |
| $N_{\mathrm{e}} = 0$ |
| $N_{\mathrm{j}} \geq 1$ |
| $H_{\mathrm{T}} > 90$ GeV |
| $E_{\mathrm{T}}^{\mathrm{miss}} < 60$ GeV |
| Fake muon selection in MC : |
| $N_\mu \geq 1$ (loose selection) |
| Leading $p_{\mathrm{T}}$ muon is the probe muon. |
| $S_{\mathrm{T}} > 100$ GeV |
| $\Delta R(\mu, \mu_{\mathrm{gen}}) > 0.15$ |

Table 6.3: Selection requirements for prompt and fake muon enhanced samples in data and MC. The muon candidate produced by the HLT algorithm satisfying the `HLT_IsoMu24_eta2p1` or `HLT_Mu40_eta2p1` trigger is denoted as $\mu_{\mathrm{HLT}}$. The generator level muons in MC are denoted as $\mu_{\mathrm{gen}}$.

| |
|---|
| Prompt $\tau_{\mathrm{h}}$ selection in data ($\mathrm{Z}(\rightarrow \tau\tau \rightarrow \mu\tau)$+jets) : |
| Trigger = `HLT_IsoMu24_eta2p1` |
| $N_{\mu} = 1$ (tight selection) |
| $N_{\tau} > 0$ (loose selection), leading $p_{\mathrm{T}}$ $\tau$ lepton is the probe. |
| Opposite sign muon - probe $\tau$ lepton pair. |
| $\Delta R(\mu, \mu_{\mathrm{HLT}}) < 0.15$ (trigger matching) |
| $\lvert M(\mu, \tau_{\mathrm{h}}) - 60 \rvert <$ 20 GeV |
| $M_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}}) <$ 20 GeV |
| $\Delta p_{\zeta} > -20$ GeV |
| $N_{\mathrm{e}} = 0$ |
| $0 < N_{\mathrm{j}} < 3$ |
| $S_{\mathrm{T}} > 150$ GeV |
| Prompt $\tau_{\mathrm{h}}$ selection in MC : |
| $N_{\tau} > 0$ (loose selection), leading $p_{\mathrm{T}}$ $\tau$ lepton is the probe. |
| $\Delta R(\tau, \tau_{\mathrm{gen}}) \leq 0.15$ |
| $S_{\mathrm{T}} > 100$ GeV |
| Fake $\tau_{\mathrm{h}}$ selection in data ($\mathrm{W}(\rightarrow \mu\nu)$+jets) : |
| Trigger = `HLT_IsoMu24_eta2p1` |
| Same-sign tight $\mu$ - loose $\tau$ lepton |
| No additional loose muons or loose $\tau$ leptons. |
| $\Delta R(\mu, \mu_{\mathrm{HLT}}) < 0.15$ (trigger matching) |
| $N_{\mathrm{e}} = 0$ |
| $N_{\mathrm{j}} < 2$ |
| $S_{\mathrm{T}} > 100$ GeV |
| $M(\mu, E_{\mathrm{T}}^{\mathrm{miss}}) > 50$ GeV |
| $\Delta p_{\zeta} < -10$ GeV |
| Fake $\tau_{\mathrm{h}}$ selection in MC : |
| $N_{\mu}$ =1 (loose selection) |
| $N_{\tau}$ =1 (loose selection) |
| Same-sign $\mu\tau_{\mathrm{h}}$ pair |
| $\Delta R(\tau, \tau_{\mathrm{gen}}) > 0.15$ |
| $S_{\mathrm{T}} > 100$ GeV |

Table 6.4: Selection requirements for prompt and fake $\tau$ lepton enhanced samples in data and MC. The muon candidate produced by the HLT algorithm satisfying the `HLT_IsoMu24_eta2p1` trigger is denoted as $\mu_{\mathrm{HLT}}$. The generator level $\tau$ leptons in MC are denoted as $\tau_{\mathrm{gen}}$.

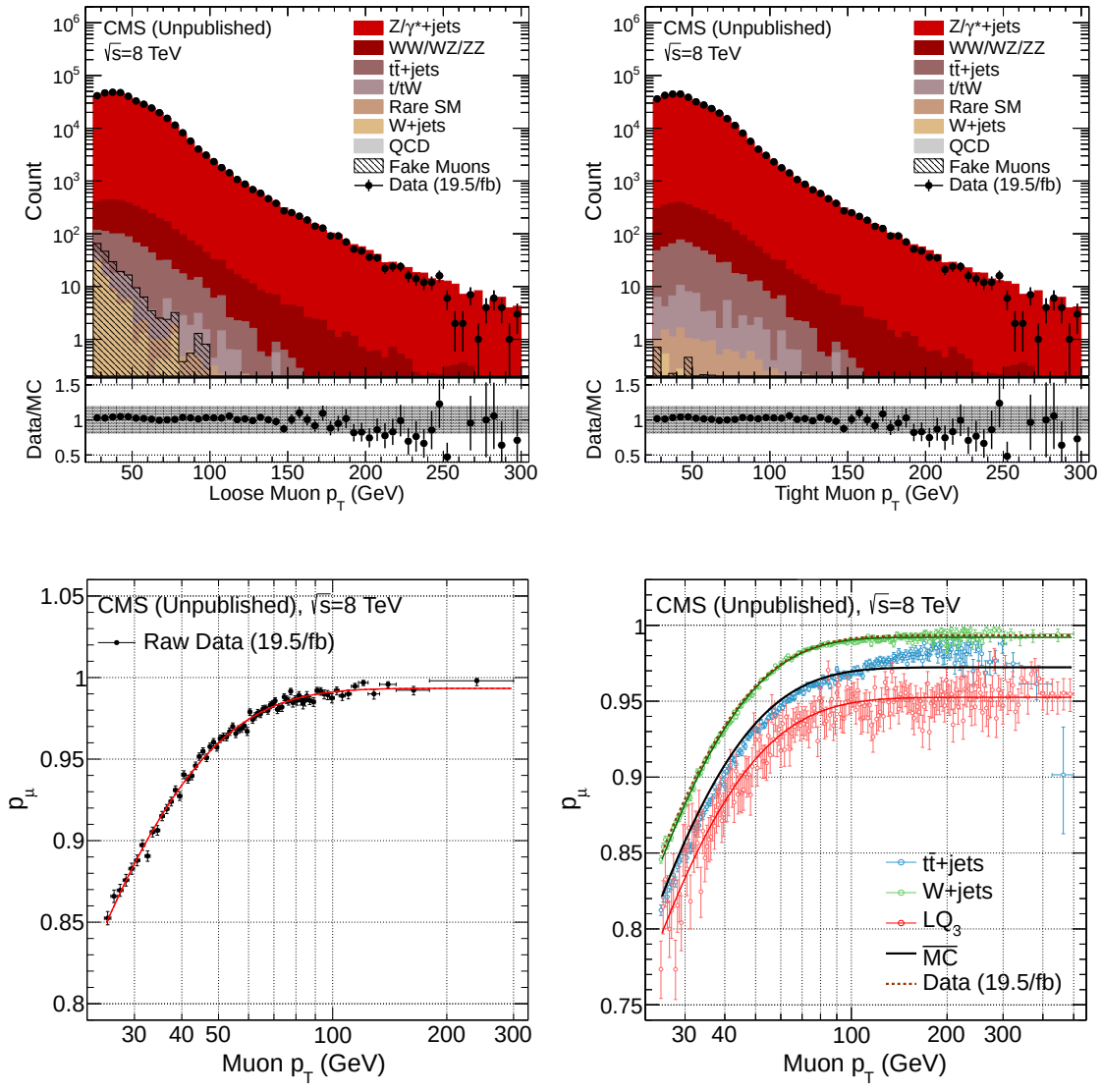


Figure 6.3: Measurement of muon prompt probabilities in the $\Delta R(\mu, j)_{\text{min}} \geq 1.0$, $|\eta| < 1.5$ region in a muon enhanced Z($\rightarrow \mu\mu$)+jets sample in data and in W($\rightarrow \mu\nu$)+jets, $\text{t}\bar{\text{t}}$+jets, and $\text{LQ}_3$ samples in MC using the selection requirements given in Table 6.3. The muon prompt probability in data (lower left) is measured as a function of the muon $p_{\text{T}}$, which is obtained as the ratio of the tight muon (upper right) and loose muon (upper left) distributions after subtracting the fake muon contributions as estimated in MC. The spread of measurements in various MC samples (lower right) are used to assess the process-dependent systematic uncertainties on the muon prompt probabilities.

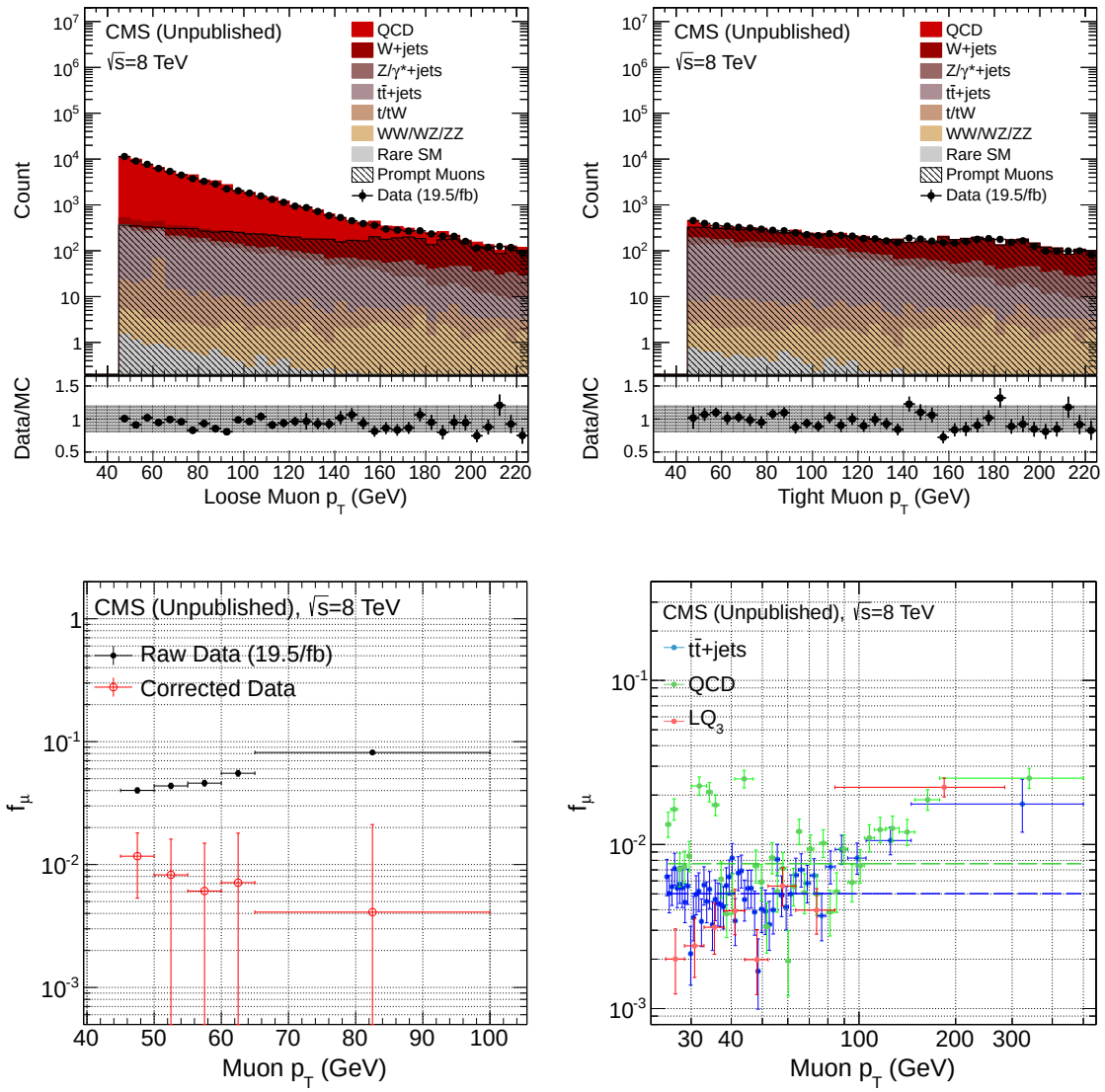


Figure 6.4: Measurement of muon fake probabilities in the $S_{\mathrm{T}} > 400$ GeV, $|\eta| < 1.5$ region in a muon enhanced QCD multijet sample in data and in QCD multijet, $\mathrm{t\bar{t}}$+jets, and $\mathrm{LQ_3}$ samples in MC using the selection requirements given in Table 6.3. The muon fake probability in data (lower left) is measured as a function of the muon $p_{\mathrm{T}}$, which is obtained as the ratio of the tight muon (upper right) and loose muon (upper left) distributions after subtracting the prompt muon contributions as estimated in MC. Since the uncertainties on the data based measurements are $\mathcal{O}(100\%)$, MC measurements (lower right) are treated solely as a cross-check and data based values are used in the analysis.

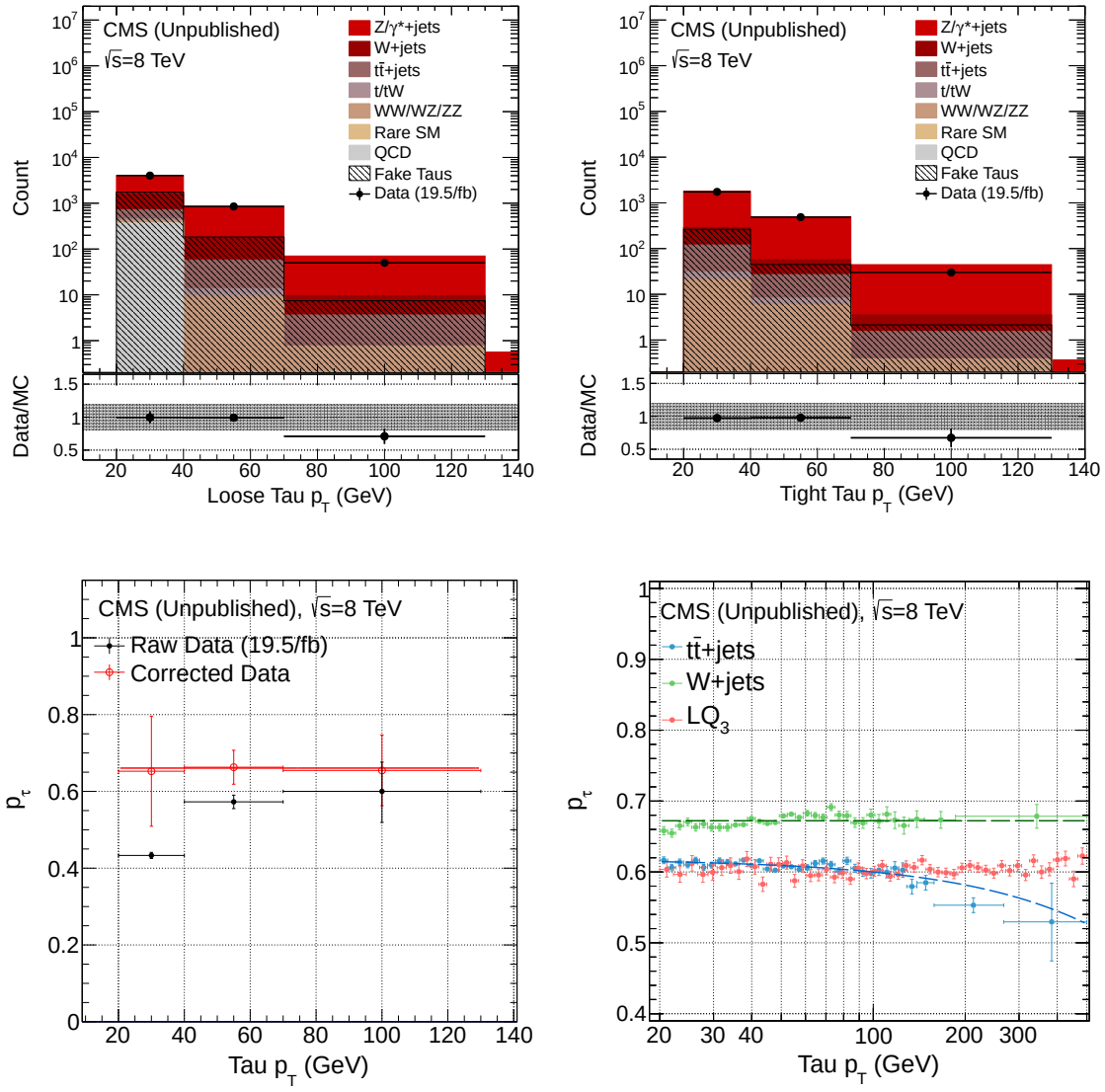


Figure 6.5: Measurement of $\tau$ lepton prompt probabilities in the $\Delta R(\tau, j)_{\text{min}} \geq 1.0$, $|\eta| < 1.5$ region in a $\tau$ lepton enhanced $\text{Z}(\rightarrow \tau\tau \rightarrow \mu\tau)$+jets sample in data and in $\text{W}(\rightarrow \tau\nu)$+jets, $\text{t}\bar{\text{t}}$+jets, and $\text{LQ}_3$ samples in MC using the selection requirements given in Table 6.4. The $\tau$ lepton prompt probability in data (lower left) is measured as a function of the $\tau$ $p_{\text{T}}$, which is obtained as the ratio of the tight $\tau$ (upper right) and loose $\tau$ (upper left) distributions after subtracting the fake $\tau$ contributions as estimated in MC. The spread of measurements in various MC samples (lower right) are used to assess the process-dependent systematic uncertainties on the $\tau$ prompt probabilities.

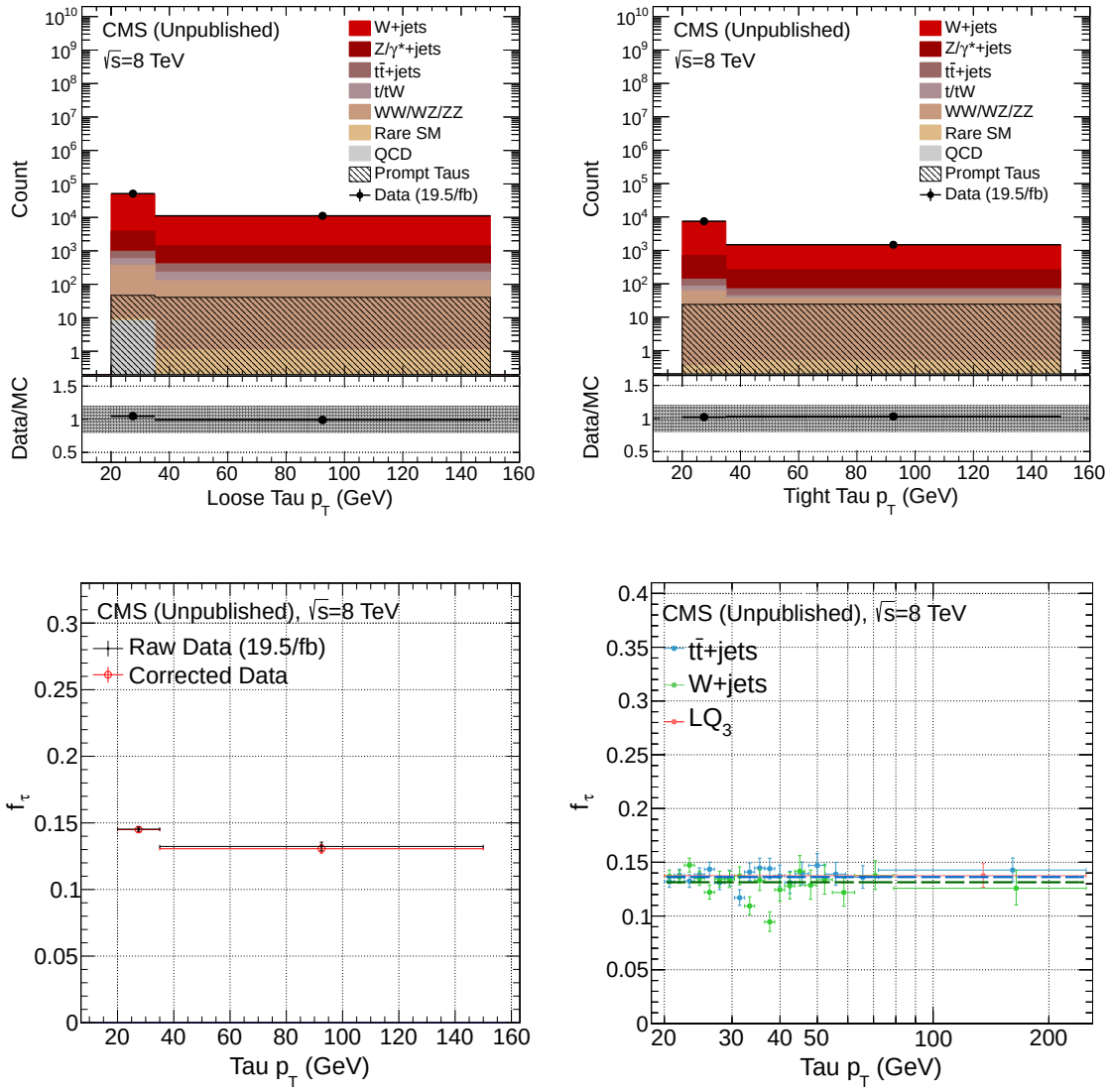


Figure 6.6: Measurement of $\tau$ lepton fake probabilities in the $\Delta R(\tau, j)_{\min} \geq 1.0$, $|\eta| < 1.5$ region in a $\tau$ lepton enhanced W($\rightarrow \mu\nu$)+jets sample in data and in W+jets, $t\bar{t}$+jets, and $LQ_3$ samples in MC using the selection requirements given in Table 6.4. The $\tau$ lepton fake probability in data (lower left) is measured as a function of the $\tau$ $p_T$, which is obtained as the ratio of the tight $\tau$ (upper right) and loose $\tau$ (upper left) distributions after subtracting the prompt $\tau$ contributions as estimated in MC. The spread of measurements in various MC samples (lower right) are used to assess the process-dependent systematic uncertainties on the $\tau$ fake probabilities.

| | | $\|\eta\| < 1.5$ | $\|\eta\| \geq 1.5$ | $\|\eta\| < 1.5$ | $\|\eta\| \geq 1.5$ | $\|\eta\| < 1.5$ | $\|\eta\| \geq 1.5$ |
|---|---|---|---|---|---|---|---|
| | $\Delta R(\mu, j)_{\min}$ | W+jets | | $LQ_3$ | | Data | |
| Prompt | 0.5 - 0.7 | 0.970±0.002 | 0.995±0.003 | 0.922±0.003 | 0.94±0.01 | 0.991±0.005 | 0.987±0.007 |
| | | 0.079±0.002 | 0.107±0.002 | 0.077±0.001 | 0.104±0.007 | 0.086±0.002 | 0.116±0.005 |
| | | $(3.0\pm0.3)\times10^{-4}$ | | | | $(4.7\pm0.8)\times10^{-5}$ | |
| | 0.7 - 1.0 | 0.987±0.001 | 0.992±0.001 | 0.943±0.001 | 0.957±0.004 | 0.987±0.002 | 0.984±0.003 |
| | | 0.122±0.001 | 0.147±0.001 | 0.120±0.003 | 0.161±0.008 | 0.126±0.004 | 0.173±0.006 |
| | | $(1.6\pm0.2)\times10^{-4}$ | | $(6\pm4)\times10^{-5}$ | | $(1.2\pm0.6)\times10^{-4}$ | |
| | $\geq 1.0$ | 0.9920±0.0002 | 0.9932±0.0002 | 0.9526±0.0008 | 0.973±0.002 | 0.9933±0.0008 | 0.9950±0.0007 |
| | | 0.1410±0.0005 | 0.1650±0.0005 | 0.140±0.001 | 0.171±0.005 | 0.144±0.001 | 0.179±0.002 |
| | | $(7.4\pm0.8)\times10^{-5}$ | | | | $(4\pm2)\times10^{-5}$ | |
| | | $t\bar{t}$+jets | | QCD Multijet | | Data | |
| Fake | $100 < S_T \leq 400$ GeV | 0.009±0.009 | 0.03±0.03 | 0.02±0.02 | 0.06±0.06 | 0.010±0.010 | 0.030±0.030 |
| | $S_T > 400$ GeV | 0.005±0.005 | 0.02±0.02 | 0.008±0.008 | 0.02±0.02 | 0.005±0.005 | 0.010±0.010 |

Table 6.5: Prompt and fake probabilities for muons as measured in data and MC. For all prompt probability bins, fit functions of the form $c_0 - e^{-c_1(p_T)^{0.8} - c_2(p_T)^2}$ and $c_0 - e^{-c_1(p_T)^{0.8}}$ are used for $|\eta| < 1.5$ and $|\eta| \geq 1.5$ regions, respectively, where $p_T$s are in units of GeV. The fit parameters, $c_i$, are reported for each prompt probability measurement.

| | $\Delta R(\tau, j)_{\min}$ | $\lvert\eta\rvert < 1.5$ | $\lvert\eta\rvert \geq 1.5$ | $\lvert\eta\rvert < 1.5$ | $\lvert\eta\rvert \geq 1.5$ | $\lvert\eta\rvert < 1.5$ | $\lvert\eta\rvert \geq 1.5$ |
|---|---|---|---|---|---|---|---|
| | | W+jets | | $t\bar{t}$+jets | | Data | |
| Prompt | 0.7 - 0.8 | 0.525±0.005 | 0.559±0.012 | 0.463±0.002<br>$(-3.2\pm0.5)\times10^{-4}$ | 0.507±0.009 | N/A | N/A |
| | 0.8 - 1.0 | 0.598±0.003 | 0.634±0.007 | 0.539±0.002<br>$(-1.8\pm0.4)\times10^{-4}$ | 0.587±0.005 | 0.526±0.061 | N/A |
| | $\geq$ 1.0 | 0.672±0.001 | 0.692±0.002<br>$(1.8\pm0.5)\times10^{-4}$ | 0.618±0.001<br>$(-1.8\pm0.3)\times10^{-4}$ | 0.650±0.002 | 0.661±0.038 | 0.743±0.060 |
| | | $\overline{\text{MC}}$ | | Correction Factor | | $p_\tau$ | |
| | 0.7 - 0.8 | 0.494<br>$-1.6\times10^{-4}$ | 0.533±0.038 | 0.880±0.102 | 1.064±0.086 | 0.435<br>$-1.4\times10^{-4}$ | 0.567±0.061 |
| | 0.8 - 1.0 | 0.569<br>$-0.9\times10^{-4}$ | 0.611±0.030 | 0.880±0.102 | 1.064±0.086 | 0.500<br>$-0.8\times10^{-4}$ | 0.650±0.061 |
| | $\geq$ 1.0 | 0.645<br>$-0.9\times10^{-4}$ | 0.671<br>$0.9\times10^{-4}$ | 0.984±0.057 | 1.064±0.086 | 0.635<br>$-0.9\times10^{-4}$ | 0.714<br>$1.0\times10^{-4}$ |
| | | W+jets | | $t\bar{t}$+jets | | Data | |
| Fake | 0.7 - 0.8 | 0.113±0.011 | 0.154±0.019 | 0.113±0.008 | 0.112±0.010 | 0.143±0.018 | 0.092±0.031 |
| | 0.8 - 1.0 | 0.114±0.014 | 0.173±0.041 | 0.127±0.006 | 0.132±0.011 | 0.132±0.012 | 0.154±0.027 |
| | $\geq$ 1.0 | 0.124±0.004 | 0.169±0.009 | 0.140±0.003 | 0.157±0.007 | 0.131±0.003 | 0.166±0.007 |
| | | $\overline{\text{MC}}$ | | Correction Factor | | $f_\tau$ | |
| | 0.7 - 0.8 | 0.113±0.011 | 0.133±0.040 | 1.265±0.201 | 0.597±0.214 | 0.143±0.027 | 0.079±0.037 |
| | 0.8 - 1.0 | 0.121±0.021 | 0.153±0.061 | 1.158±0.177 | 0.890±0.262 | 0.140±0.032 | 0.136±0.067 |
| | $\geq$ 1.0 | 0.132±0.012 | 0.163±0.015 | 1.056±0.041 | 0.982±0.067 | 0.139±0.014 | 0.160±0.018 |

Table 6.6: Prompt and fake probabilities for $\tau$ leptons as measured in data and MC, neglecting the $S_{\text{T}}$ based corrections. In data bins marked as N/A, the corresponding measurements are discarded as they are dominated by statistical uncertainties and the nearest $\Delta R(\tau, j)_{\min}$ bins are used for the calculation of associated correction factors. In bins where two values are reported, a first order polynomial fit as a function of $\tau$ lepton $p_{\text{T}}$ is used with $p_{\text{T}}$ in units of GeV, and these two values are to be taken as the y-intercept and the slope respectively. For these $\tau$ lepton $p_{\text{T}}$ dependent probabilities, the associated uncertainties (as a function of the $\tau$ lepton $p_{\text{T}}$) are calculated via $\Delta D$ as expressed in Eq. (6.6).

| $\Delta R(\tau, j)_{\min}$ | $c_{S_T}$ for $p_\tau$ | | $c_{S_T}$ for $f_\tau$ | |
|---|---|---|---|---|
| | $\lvert\eta\rvert < 1.5$ | $\lvert\eta\rvert \geq 1.5$ | $\lvert\eta\rvert < 1.5$ | $\lvert\eta\rvert \geq 1.5$ |
| 0.7 - 0.8 | $(11.3\pm1.5) \cdot 10^{-5}$ | $(15.0\pm6.2) \cdot 10^{-5}$ | $(1.5\pm1.5) \cdot 10^{-5}$ | – |
| 0.8 - 1.0 | $(8.7\pm1.1) \cdot 10^{-5}$ | $(4.3\pm2.5) \cdot 10^{-5}$ | $(1.5\pm1.5) \cdot 10^{-5}$ | – |
| $\geq 1.0$ | $(8.0\pm0.7) \cdot 10^{-5}$ | $(4.3\pm2.5) \cdot 10^{-5}$ | $(1.5\pm1.5) \cdot 10^{-5}$ | – |

Table 6.7: $S_T$ based correction parameters for $\tau$ lepton prompt and fake probabilities. Each correction parameter, $c_{S_T}$, is measured by fitting a first order polynomial to the corresponding $t\bar{t}$+jets sample, and has dimensions of $GeV^{-1}$. No correction parameters are reported for the $f_\tau$ values in the $|\eta| \geq 1.5$ bins where no clear $S_T$ dependent behavior is observed.

## 6.2 Irreducible Backgrounds

Background contributions due to lepton charge mismeasurements and irreducible same-sign dilepton processes are estimated directly from the simulated samples. These prompt-prompt contributions are calculated by requiring a match of $\Delta R < 0.15$ between the reconstructed lepton candidate and a generator level object of the same kind without any requirements on the electric charge. The charge misidentification backgrounds are dominated by $\tau_h$ candidates. These backgrounds contribute to 2 - 3% of the total expected backgrounds in selections for $M_{LQ_3} \leq 400$ GeV, and are negligible in those for higher LQ3 masses.

## 6.3 Systematic Uncertainties

### 6.3.1 Systematic Uncertainties for Misidentified Lepton Backgrounds

The uncertainties (systematic and statistical) on the muon and $\tau$ lepton prompt and fake probability measurements constitute the systematic uncertainty affecting the misidentified lepton background estimate. This uncertainty, $\sigma_{LTEM}$, is estimated by independently varying the individual muon and $\tau$ lepton prompt and fake probabilities within their respective uncertainties, and observing the overall change in the misidentified lepton background estimate, $N_{TT}^{\text{Fake Bkg}}$, as given below:

$$\sigma_{LTEM} = \left( \sum_{r_i = f_i, p_i} \left( \frac{N_{TT}^{\text{Fake Bkg}}(r_i + \sigma_{r_i}) - N_{TT}^{\text{Fake Bkg}}(r_i - \sigma_{r_i})}{2} \right)^2 \right)^{\frac{1}{2}}. \quad (6.9)$$

The summation is over all prompt and fake probabilities, $r_i$s, and the associated total uncertainty for each probability is denoted as $\sigma_{r_i}$.

The uncertainties of the $\tau$ lepton fake probabilities dominate the total systematic uncertainty of $N_{\mathrm{TT}}^{\mathrm{Fake\ Bkg}}$, since events with prompt muon - misidentified $\tau_{\mathrm{h}}$ candidates constitute the major background in the signal selection. Correlations among the prompt and fake probabilities have also been studied but are observed to have negligible impact on the final results. The LTEM uncertainties on the misidentified lepton background are assumed to be fully correlated between the two $\widetilde{|\eta|}$ channels. The uncertainties in the background rate of misidentified leptons lie in the range of 21 - 28% in the central channel and 21 - 36% in the forward channel.

### 6.3.2 Systematic Uncertainties for Simulation-based Backgrounds

Since the signal efficiencies as well as the prompt-prompt contributions to the background are estimated from simulated samples, certain sources of systematic uncertainty are considered to account for any biases the simulation might have. The multiplicative uncertainty sources are accounted for by per-object or per-event application of the corresponding uncertainty values. The effects of the non-multiplicative sources are evaluated by observing the overall change in the MC yield for a given event selection when the corresponding values are independently varied within their uncertainty bands.

Normalization uncertainties of 20% are applied for $\mathrm{t\bar{t}}$+jets, DY+jets and diboson processes as observed in the signal-depleted selection plots given in Fig. 5.5. An uncertainty of 30% is applied for other rare SM process following the theoretical uncertainties in the NLO cross sections for processes such as $\mathrm{t\bar{t}W}$, $\mathrm{t\bar{t}Z}$ [156, 157], and triboson [158] production. The luminosity used in the normalization of signal and MC samples has an uncertainty of 2.6% [76].

The uncertainty in the pileup re-weighting of simulated samples is estimated by varying the total inelastic cross section [99] by 5%. Since variations among optimized selection points are found to be small, flat values of 1% and 2.5% are assumed for the central and forward channels, respectively.

Muon identification, isolation, and trigger efficiencies are determined with a tag-and-probe method in $\mathrm{Z}(\rightarrow \mu\mu)$+jets enriched data [159]. The muon reconstruction and isolation uncertainty is about 1% and the single muon trigger matching uncertainty is $\leq 0.5\%$. Muon momentum scale and resolution uncertainties are 0.2% and 0.6%, respectively, with an additional uncertainty of 5%/TeV for muons with $p_{\mathrm{T}} > 200$ GeV [96].

An uncertainty of 6% is applied for each prompt $\tau$ lepton found in an event to account for uncertainties in the efficiency of $\tau$ lepton identification. The uncertainty in the $\tau$ lepton energy is taken into account by varying the energy of all $\tau$ leptons by $\pm 3\%$. Uncertainties induced by the energy resolution of prompt $\tau$ leptons in simulated samples are estimated by changing the resolution by $\pm 10\%$ [160].

Uncertainties in the jet energy resolution [132] are assessed by changing the simulation correction factors within their uncertainties. These correction factors lie between 1.05 and 1.29 depending on jet $\eta$, with corresponding uncertainties varying from 5% to 16%. The $p_{\mathrm{T}}$- and $\eta$-dependent scale factors for the jet energy scale [132] are similarly varied by one standard deviation to obtain the corresponding uncertainties in simulated samples. This corresponds to a 1 - 3% variation of the scale factors.

Electon energy scale uncertainties as a function of electron $\eta$ are also considered, but are observed to have negligible impact on the results.

The jet energy resolution correction uncertainties as well as muon, $\tau$ lepton, electron, and jet energy scale and resolution uncertainties are propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ of the event as shown in Eq. (6.10), where the sum is over all leptons satisfying the selection criteria as given in Section 4.6 and over all PF jets (no jet selection applied) with $p_{\mathrm{T}} > 15$ GeV and $|\eta| < 3$, provided they are not overlapping with the identified lepton objects ($\Delta R > 0.5$).

$$\vec{E}_{\mathrm{T}}^{\mathrm{miss}\,\prime} = \vec{E}_{\mathrm{T}}^{\mathrm{miss}} + \sum_{\ell,j} (\vec{p}_{\mathrm{T}} - \vec{p}\,'_{\mathrm{T}}) \tag{6.10}$$

Signal samples are produced with the CTEQ6L1 [161] parton distribution function (PDF) set and the associated PDF uncertainties in the signal acceptance are estimated using the PDF uncertainty prescription for LHC [162].

All systematic uncertainty sources considered for simulated samples are summarized in Table 6.8. Energy scale and resolution based systematic uncertainties for signal and background estimates as well as the PDF based systematic uncertainties for signal estimates for each $M_{\mathrm{LQ}_3}$ selection are listed in Table 6.9.

## 6.4 Closure Tests

For the verification of the prompt and fake probabilities and the associated uncertainties within LTEM, two closure tests are conducted in a signal-depleted selection of events and in the signal region in MC. The level of agreement between the expected and observed

| | | |
|---|---|---|
| Multiplicative systematic uncertainties: | | |
| Integrated luminosity | 2.6% | (per event) |
| $\tau$ lepton reco/ID/iso | 6% | (per $\tau$ lepton) |
| Muon HLT | 0.5% | (per trigger matched muon) |
| Muon reco/ID/iso | 1% | (per muon) |
| Non-multiplicative systematic uncertainties: | | |
| $\tau$ lepton energy scale | 3% | |
| $\tau$ lepton energy resolution | $\begin{cases} p_{\mathrm{T}}^{\mathrm{reco}\prime} = \max(0, p_{\mathrm{T}}^{\mathrm{gen}} + c_r(p_{\mathrm{T}}^{\mathrm{reco}} - p_{\mathrm{T}}^{\mathrm{gen}})) \\ c_r = 1.0 \pm 0.1 \end{cases}$ | |
| Electron energy scale | $\begin{cases} 0.4\%, & \lvert\eta\rvert < 1.442 \\ 4.1\%, & \lvert\eta\rvert \geq 1.442 \end{cases}$ | |
| Muon momentum scale and resolution | $\begin{cases} 0.63\%, & p_{\mathrm{T}} < 200 \text{ GeV} \\ 0.6 \oplus \frac{0.05 \times p_{\mathrm{T}}}{1000 \text{ GeV}}\%, & p_{\mathrm{T}} \geq 200 \text{ GeV} \end{cases}$ | |
| Jet energy scale | $\lesssim 3\%$ as a function of jet $p_{\mathrm{T}}$ and $\eta$ | |
| Jet energy resolution | $\begin{cases} p_{\mathrm{T}}^{\mathrm{reco}\prime} = \max(0, p_{\mathrm{T}}^{\mathrm{gen}} + c_r(p_{\mathrm{T}}^{\mathrm{reco}} - p_{\mathrm{T}}^{\mathrm{gen}})) \\ c_r = 1.052 \pm 0.063, & \text{if } 0.0 \leq \lvert\eta\rvert < 0.5 \\ c_r = 1.057 \pm 0.057, & \text{if } 0.5 \leq \lvert\eta\rvert < 1.1 \\ c_r = 1.096 \pm 0.065, & \text{if } 1.1 \leq \lvert\eta\rvert < 1.7 \\ c_r = 1.134 \pm 0.094, & \text{if } 1.7 \leq \lvert\eta\rvert < 2.3 \\ c_r = 1.288 \pm 0.200, & \text{if } 2.3 \leq \lvert\eta\rvert < 5.0 \end{cases}$ | |
| Parton distribution function (PDF) | Largest uncertainty envelope as defined by CT10 [163], MSTW2008nlo68cl [164], and NNPDF2.0 [165] PDF sets centered on the mean CT10 value. | |
| Pileup | 5% (on total inelastic cross section) | |

Table 6.8: Major systematic uncertainties affecting the signal efficiency and prompt-prompt background estimate in MC. A correction to the mean value is performed in the case of jet energy resolution in addition to the propagation of the related uncertainty.

| $M_{LQ_3}$ (GeV) | | JER | JER$_{\pm\sigma}$ | JEC$_{\pm\sigma}$ | TES$_{\pm\sigma}$ | EES$_{\pm\sigma}$ | MES$_{\pm\sigma}$ | TER$_{\pm\sigma}$ | PDF |
|---|---|---|---|---|---|---|---|---|---|
| 200 | Background | 0.3 | 0.5 | 4.0 | 4.5 | 0.0 | 0.7 | 0.9 | |
| | Signal | 0.3 | 0.3 | 3.8 | 7.8 | 0.0 | 3.8 | 0.0 | $^{+6.1}_{-1.7}$/ N/A |
| 250 | Background | 0.3 | 0.5 | 4.0 | 4.5 | 0.0 | 0.7 | 0.9 | |
| | Signal | 0.5 | 0.9 | 4.1 | 3.8 | 0.6 | 1.3 | 1.8 | $^{+8.1}_{-2.6}$/$^{+4.0}_{-3.5}$ |
| 300 | Background | 0.4 | 0.7 | 4.2 | 5.4 | 0.0 | 0.5 | 1.3 | |
| | Signal | 1.2 | 2.0 | 2.9 | 4.5 | 0.0 | 1.3 | 1.0 | $^{+5.1}_{-2.1}$/$^{+4.2}_{-3.6}$ |
| 350 | Background | 1.3 | 1.1 | 4.7 | 5.1 | 0.0 | 0.6 | 1.3 | |
| | Signal | 0.9 | 1.9 | 2.6 | 3.1 | 0.1 | 0.8 | 1.2 | $^{+4.7}_{-2.7}$/$^{+8.7}_{-7.4}$ |
| 400 | Background | 0.9 | 1.2 | 6.5 | 5.6 | 0.0 | 0.7 | 1.2 | |
| | Signal | 0.3 | 0.9 | 2.0 | 3.7 | 0.0 | 0.9 | 2.3 | $^{+3.0}_{-2.2}$/$^{+2.7}_{-3.4}$ |
| 450 | Background | 0.6 | 1.2 | 5.8 | 5.8 | 0.0 | 0.4 | 1.1 | |
| | Signal | 0.5 | 0.7 | 2.6 | 3.4 | 0.3 | 0.7 | 0.6 | $^{+3.0}_{-2.6}$/$^{+2.2}_{-4.1}$ |
| 500 | Background | 0.7 | 2.5 | 4.5 | 5.6 | 0.1 | 1.5 | 0.7 | |
| | Signal | 1.4 | 2.0 | 2.7 | 2.9 | 0.0 | 0.4 | 0.9 | $^{+2.9}_{-2.9}$/$^{+6.7}_{-9.2}$ |
| 550 | Background | 0.7 | 2.4 | 6.7 | 8.7 | 0.0 | 0.7 | 1.7 | |
| | Signal | 0.5 | 0.9 | 1.9 | 2.0 | 0.2 | 0.4 | 0.9 | $^{+2.9}_{-4.3}$/$^{+2.4}_{-6.2}$ |
| 600 | Background | 1.4 | 1.0 | 4.9 | 7.2 | 0.0 | 1.0 | 1.7 | |
| | Signal | 0.6 | 0.4 | 1.3 | 2.0 | 0.0 | 0.4 | 1.2 | $^{+1.7}_{-6.0}$/$^{+6.5}_{-9.4}$ |
| 650 | Background | 1.4 | 1.0 | 4.9 | 7.2 | 0.0 | 1.0 | 1.7 | |
| | Signal | 0.2 | 0.7 | 1.2 | 1.5 | 0.0 | 0.1 | 1.1 | $^{+2.4}_{-7.9}$/$^{+2.1}_{-11.9}$ |
| 700 | Background | 0.5 | 0.5 | 5.9 | 6.6 | 0.0 | 0.2 | 1.8 | |
| | Signal | 0.2 | 0.9 | 0.7 | 1.8 | 0.0 | 0.1 | 1.3 | $^{+1.9}_{-10.4}$/$^{+7.0}_{-17.5}$ |
| 750 | Background | 0.5 | 0.5 | 5.9 | 6.6 | 0.0 | 0.2 | 1.8 | |
| | Signal | 0.7 | 0.9 | 1.0 | 2.3 | 0.0 | 0.4 | 1.0 | $^{+1.5}_{-9.9}$/$^{+7.0}_{-17.5}$ |
| 800 | Background | 0.5 | 0.5 | 5.9 | 6.6 | 0.0 | 0.2 | 1.8 | |
| | Signal | 0.4 | 0.7 | 0.4 | 1.8 | 0.0 | 0.4 | 1.1 | $^{+4.5}_{-13.1}$/$^{+10.8}_{-20.8}$ |

Table 6.9: Percent values of the non-multiplicative systematic uncertainties on signal and background yields for the optimized selection criteria at a given $M_{LQ_3}$. Uncertainties reported as single values are valid for both central and forward channels. PDF based systematic uncertainties are computed only for the signal yields, and are reported for central and forward channels, respectively. Labels JER, JEC, TES, EES, MES, and TER refer to jet energy resolution, jet energy correction, $\tau$ lepton energy scale, electron energy scale, muon momentum scale & resolution, and $\tau$ lepton energy resolution uncertainties, respectively. JER refers to the change in the yields due to the jet energy resolution corrections, whereas JER$_{\pm\sigma}$ refers to the change in the yields due to the uncertainties in the corrections.

number of events in a given bin or channel is expressed in terms of the associated Z-score, which is calculated taking into account the total uncertainty on the expected number of events. A unit Z-score, $|\mathrm{Z}| = 1$, refers to a two-tailed 1-standard deviation quantile ($\sim 68\%$) of the normal distribution. The LTEM yields consistent results for the misidentification background when applied to a signal-depleted selection of events in data and to simulated events in the signal region. The expected yields are in agreement with the observations within 1.5 standard deviations in both selections as demonstrated below.

### 6.4.1 Closure Test in Data Events: Signal-depleted Selection

A signal-depleted selection as defined in Table 6.10 is created in data by reverting the jet-plus-lepton multiplicity cut. This region is similar but orthogonal to the data selections used in measuring individual muon and $\tau$ lepton prompt and fake probabilities.

| | |
|---|---|
| Trigger= | $\begin{cases} \texttt{HLT\_Mu24\_eta2p1} & \text{if } 25 < \mu_{\mathrm{LT}}\ p_{\mathrm{T}} \leq 45 \text{ GeV} \\ \texttt{HLT\_IsoMu24\_eta2p1} & \text{if } 25 < \mu_{\mathrm{LT}}\ p_{\mathrm{T}} \leq 45 \text{ GeV}, \mu_{\mathrm{LT}} \text{ passes tight selection} \\ \texttt{HLT\_Mu40\_eta2p1} & \text{if } \mu_{\mathrm{LT}}\ p_{\mathrm{T}} > 45 \text{ GeV} \end{cases}$ |
| | $N_\tau \geq 1$ (loose selection) |
| | $N_\mu \geq 1$ (loose selection) |
| | A same-sign muon - $\tau$ lepton pair: $\mu_{\mathrm{LT}}$ - $\tau_{\mathrm{LT}}$ |
| | $N_{\mathrm{j}} + N_\ell \leq 3,\ \ell : \mathrm{e}, \mu, \tau_{\mathrm{h}}$ (to remove signal events) |
| | $\Delta R(\mu_{\mathrm{LT}}, \mu_{\mathrm{HLT}}) < 0.15$ (trigger matching) |
| | $S_{\mathrm{T}} > 150$ GeV |
| | $\Delta p_\zeta > -10$ GeV |
| | For $N_\mu = 2$, remove event if $\lvert M(\mu^-, \mu^+) - M_{\mathrm{Z}} \rvert / M_{\mathrm{Z}} < 10\%$ |

Table 6.10: Signal-depleted selection requirements in data. The same-sign $\mu_{\mathrm{LT}}$ - $\tau_{\mathrm{LT}}$ pair denotes the dilepton pair selected for the application of LTEM. Mass of the Z boson is given as $M_{\mathrm{Z}}$.

The closure test yields a total expected background of $571 \pm 86$ events for an observation of 556 events, which are consistent within uncertainties. Distributions of $S_{\mathrm{T}}$ and $\tau_{\mathrm{LT}}$ $p_{\mathrm{T}}$ parameters are presented in Fig. 6.7. No statistically significant deviations are observed in the distributions, and there is an overall agreement within two standard deviations of the expected background.

### 6.4.2 Closure Test in Simulated Events: Signal Selection

The closure test in the signal region is conducted in MC in its entirety for each of the optimized selection criteria given in Table 5.2, with the fake lepton containing backgrounds

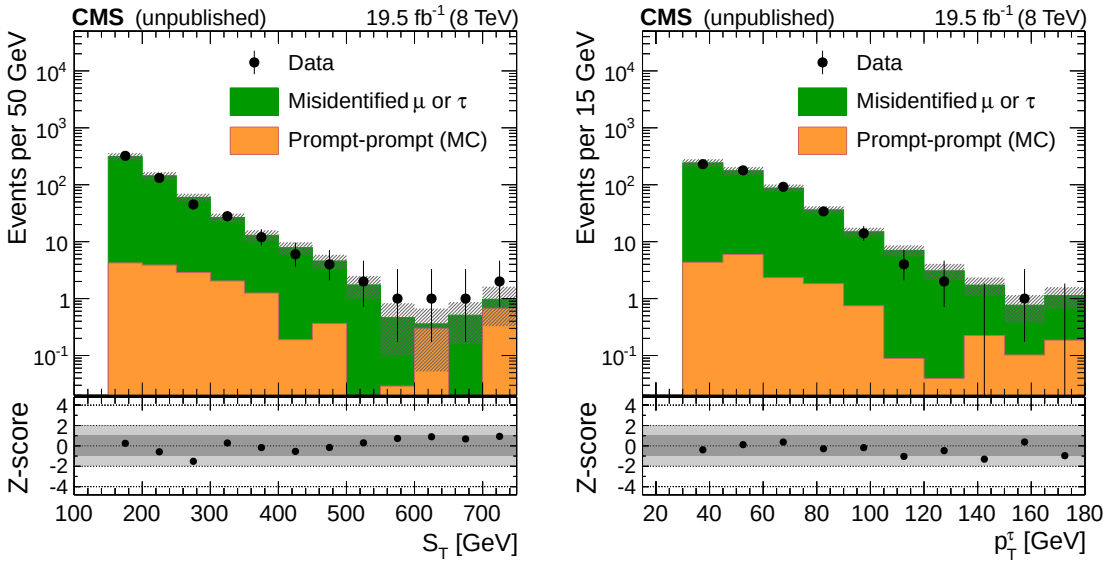


Figure 6.7: Closure test distributions for $S_{\mathrm{T}}$ (left) and $\tau_{\mathrm{LT}}$ $p_{\mathrm{T}}$ (right) parameters using the signal-depleted selection in data events. Each bin of these distributions is treated as an independent measurement, for which the expected number of events, the associated LTEM systematic uncertainties, prompt-prompt MC normalization uncertainties, and the corresponding Z-scores are computed separately. The global LTEM systematic uncertainty is calculated as $\sim$ 15.3%, whereas the associated statistical uncertainty is $\sim$ 2.0%. The prompt-prompt contribution estimated using MC constitutes about 2.8% of the total in this control region with $15.8 \pm 3.4$ events [152].

estimated via the LTEM. Since the entire test is performed in MC, the reported uncertainties are combinations of statistical and the LTEM systematic uncertainties only, with no additional simulation specific systematic uncertainties (normalization, object energy resolution and scale, object efficiencies). For the implementation of the LTEM, no data - MC correction factors are applied to the muon and $\tau$ lepton prompt and fake probabilities. Results are summarized in Table 6.11 for the set of optimized cuts at each $\mathrm{LQ}_3$ mass hypothesis, demonstrating an overall agreement between the expected and observed number of events within uncertainties.

| $M_{LQ_3}$ (GeV) | Total $N_{Bkg}^{Exp}$ $\pm$stat. $\pm$ sys. | $N_{Bkg}^{Obs}$ $\pm$stat. | Z-score | $N_{LQ_3}^{Exp\ Bkg}$ $\pm$stat. $\pm$ sys. | $N_{LQ_3}^{Obs}$ $\pm$stat. |
|---|---|---|---|---|---|
| | Central channel: $\widetilde{\lvert\eta\rvert} < 0.9$ | | | | |
| 200 | 105±3±15 | 101±8 | −0.2 | 2.68±4.39±1.22 | 52.6±20.6 |
| 250 | 105±3±15 | 101±8 | −0.2 | 11.2±4.8±5.0 | 252±24 |
| 300 | 34.9±1.4±4.9 | 28.8±2.7 | −0.8 | −0.97±2.06±3.44 | 154±11.1 |
| 350 | 30.4±1.3±4.3 | 25.6±2.6 | −0.7 | 0.03±1.02±2.06 | 92.4±5.6 |
| 400 | 5.35±0.51±0.81 | 2.11±0.45 | −1.5 | 0.54±0.41±0.74 | 28.4±2.1 |
| 450 | 4.69±0.48±0.72 | 1.57±0.35 | −1.2 | −0.07±0.21±0.52 | 17.3±1.1 |
| 500 | 2.75±0.33±0.43 | 0.93±0.25 | −1.0 | 0.05±0.11±0.27 | 9.76±0.59 |
| 550 | 2.35±0.31±0.39 | 0.61±0.18 | −0.8 | −0.09±0.07±0.24 | 6.13±0.34 |
| 600 | 1.80±0.27±0.32 | 0.41±0.14 | −1.3 | −0.00±0.07±0.10 | 3.61±0.19 |
| 650 | 1.80±0.27±0.32 | 0.41±0.14 | −1.3 | −0.05±0.02±0.12 | 2.19±0.11 |
| 700 | 1.11±0.21±0.20 | 0.27±0.12 | −0.9 | −0.05±0.01±0.05 | 1.28±0.06 |
| 750 | 1.11±0.21±0.20 | 0.27±0.12 | −0.9 | −0.04±0.01±0.04 | 0.82±0.04 |
| 800 | 1.11±0.21±0.20 | 0.27±0.12 | −0.9 | −0.03±0.00±0.01 | 0.51±0.02 |
| | Forward channel: $\widetilde{\lvert\eta\rvert} \geq 0.9$ | | | | |
| 200 | 74.5±3.0±12.4 | 81.5±5.34 | +0.6 | 1.86±1.45±0.23 | – |
| 250 | 74.5±3.0±12.4 | 81.5±5.34 | +0.6 | 6.53±2.42±1.35 | 50.2±10.5 |
| 300 | 23.6±1.4±4.0 | 25.1±2.47 | +0.3 | 0.15±1.04±0.79 | 33.4±5.2 |
| 350 | 20.7±1.3±3.6 | 19.9±2.10 | −0.1 | 0.07±0.46±0.39 | 18.5±2.5 |
| 400 | 3.30±0.48±0.64 | 3.46±0.77 | −0.1 | −0.17±0.18±0.17 | 6.11±0.95 |
| 450 | 3.14±0.47±0.60 | 3.05±0.73 | 0.0 | 0.01±0.10±0.09 | 3.84±0.54 |
| 500 | 1.96±0.32±0.38 | 2.18±0.53 | +0.1 | 0.06±0.05±0.04 | 1.61±0.24 |
| 550 | 1.46±0.28±0.30 | 2.07±0.53 | +0.5 | −0.01±0.03±0.03 | 1.15±0.15 |
| 600 | 0.75±0.21±0.15 | 1.59±0.46 | +1.2 | −0.02±0.01±0.02 | 0.56±0.08 |
| 650 | 0.75±0.21±0.15 | 1.59±0.46 | +1.2 | 0.01±0.01±0.01 | 0.29±0.04 |
| 700 | 0.24±0.13±0.05 | 1.18±0.42 | +1.2 | −0.01±0.00±0.01 | 0.18±0.02 |
| 750 | 0.24±0.13±0.05 | 1.18±0.42 | +1.2 | −0.01±0.00±0.00 | 0.13±0.02 |
| 800 | 0.24±0.13±0.05 | 1.18±0.42 | +1.2 | 0.00±0.00±0.00 | 0.08±0.01 |

Table 6.11: Closure test results as conducted in MC using the optimized signal selection requirements. The reported total expected and observed background events, total $N_{Bkg}^{Exp}$ and $N_{Bkg}^{Obs}$, include prompt-prompt and misidentified lepton background components, but ignore the signal yields. Z-scores are provided to quantify the level of agreement between the total expected and observed background events. Misidentified lepton backgrounds due to the signal events as calculated via the LTEM are denoted as $N_{LQ_3}^{Exp\ Bkg}$, and are consistent with zero within the uncertainties (except for $M_{LQ_3} = 250$ GeV, which is still negligible in comparison to the high signal yield). All reported systematic uncertainties are LTEM only. No expected signal yield for $M_{LQ_3} = 200$ GeV is reported in the forward channel as it is measured to be zero.

# Chapter 7

# Results and Statistical Interpretation

Search results are presented in Table 7.1 for each set of the optimized selection requirements, and in Figs. 7.1 and 7.2 in terms of $S_{\mathrm{T}}$, $\tau_{\mathrm{LT}}$ $p_{\mathrm{T}}$, jet multiplicity, and $|\widetilde{\eta}|$ distributions. Signal distributions for $M_{\mathrm{LQ}_3} = 400$ GeV are superimposed to demonstrate the signal-like behavior. For each set of selection requirements, the observed number of events is found to be in an overall agreement with the SM-only hypothesis within 1.5 standard deviations, and the distributions reveal no statistically significant deviations from the SM expectations.

A limit is set on the pair production cross section of charge $-1/3$ third generation scalar LQs using the $CL_S$ modified frequentist approach [166, 167] at 95% CL, where the counting experiment method is adopted for the statistical analysis. The uncertainties affecting the mean values of expected signal and background events are treated as nuisance parameters modeled by log-normal and gamma distributions for systematic and statistical uncertainties, respectively [168]. Limits are obtained for each of the central and forward channels, as well as a combination channel formed by the statistical combination of these two. For the combination channel, statistical uncertainties on the signal and background yields are assumed to be fully uncorrelated between the two channels, whereas all other systematic uncertainties are assumed to be fully correlated.

The expected and observed limits, as well as the expected signal significances are presented in Fig. 7.3. The exclusion limits worsen as the $\mathrm{LQ}_3$ mass approaches the mass of the top quark and the $\mathrm{LQ}_3$ decay products become softer. At $M_{\mathrm{LQ}_3} = 200$ GeV, more

than 90% of $\tau$ leptons originating from $LQ_3$ decays have $p_T < 60$ GeV, which causes a decrease both in the signal selection efficiency and the discriminating performance of the $S_T$ and $\tau$ lepton $p_T$ spectra. Therefore, no exclusion limits are quoted for masses below 200 GeV. Assuming $\beta = 1$, the observed (expected) limits on the $LQ_3$ mass in the central, forward and combination channels are 545 (576), 454 (435), and 550 (582) GeV, respectively.

| $M_{LQ_3}$ (GeV) | $\tau_{LT}$ $p_T$ (GeV) | $S_T$ (GeV) | $N^{PP}_{Bkg}$ $\pm$stat. | Total $N^{Exp}_{Bkg}$ $\pm$stat. $\pm$ sys. | $N^{Obs}$ | Z-score | $N^{Exp}_{LQ_3}$ $\pm$stat. | $\epsilon_{LQ_3}$ (%) |
|---|---|---|---|---|---|---|---|---|
| | | | Central channel: $\widetilde{\lvert\eta\rvert} < 0.9$ | | | | | |
| 200 | 35 | 410 | $8.5 \pm 1.0$ | $128 \pm 5 \pm 25$ | 105 | $-1.0$ | $53 \pm 21$ | 0.08 |
| 250 | 35 | 410 | $8.5 \pm 1.0$ | $128 \pm 5 \pm 25$ | 105 | $-1.0$ | $252 \pm 24$ | 1.31 |
| 300 | 50 | 470 | $4.2 \pm 0.5$ | $39.9 \pm 2.9 \pm 8.3$ | 27 | $-1.5$ | $153 \pm 11$ | 2.22 |
| 350 | 50 | 490 | $4.0 \pm 0.5$ | $34.6 \pm 2.7 \pm 7.1$ | 25 | $-1.2$ | $92.4 \pm 5.6$ | 3.29 |
| 400 | 65 | 680 | $0.9 \pm 0.2$ | $7.2 \pm 1.2 \pm 1.7$ | 4 | $-1.0$ | $28.4 \pm 2.1$ | 2.27 |
| 450 | 65 | 700 | $0.8 \pm 0.2$ | $6.3 \pm 1.1 \pm 1.6$ | 4 | $-0.8$ | $17.3 \pm 1.1$ | 2.90 |
| 500 | 65 | 770 | $0.5 \pm 0.2$ | $3.2 \pm 0.8 \pm 0.8$ | 4 | $+0.5$ | $9.8 \pm 0.6$ | 3.25 |
| 550 | 65 | 800 | $0.4 \pm 0.1$ | $2.7 \pm 0.8 \pm 0.6$ | 4 | $+0.7$ | $6.1 \pm 0.3$ | 3.89 |
| 600 | 65 | 850 | $0.2 \pm 0.1$ | $1.8 \pm 0.6 \pm 0.4$ | 3 | $+0.9$ | $3.6 \pm 0.2$ | 4.20 |
| 650 | 65 | 850 | $0.2 \pm 0.1$ | $1.8 \pm 0.6 \pm 0.4$ | 3 | $+0.9$ | $2.2 \pm 0.1$ | 4.54 |
| 700 | 85 | 850 | $0.1 \pm 0.1$ | $1.1 \pm 0.5 \pm 0.3$ | 2 | $+0.8$ | $1.3 \pm 0.1$ | 4.60 |
| 750 | 85 | 850 | $0.1 \pm 0.1$ | $1.1 \pm 0.5 \pm 0.3$ | 2 | $+0.8$ | $0.8 \pm 0.1$ | 5.01 |
| 800 | 85 | 850 | $0.1 \pm 0.1$ | $1.1 \pm 0.5 \pm 0.3$ | 2 | $+0.8$ | $0.5 \pm 0.1$ | 5.19 |
| | | | Forward channel: $\widetilde{\lvert\eta\rvert} \geq 0.9$ | | | | | |
| 200 | 35 | 410 | $4.2 \pm 0.5$ | $72 \pm 4 \pm 15$ | 87 | $+1.1$ | – | – |
| 250 | 35 | 410 | $4.2 \pm 0.5$ | $72 \pm 4 \pm 15$ | 87 | $+1.1$ | $50 \pm 11$ | 0.26 |
| 300 | 50 | 470 | $1.8 \pm 0.3$ | $20.3 \pm 2.2 \pm 3.9$ | 23 | $+0.5$ | $33.4 \pm 5.2$ | 0.48 |
| 350 | 50 | 490 | $1.7 \pm 0.3$ | $18.2 \pm 2.0 \pm 3.5$ | 19 | $+0.2$ | $18.5 \pm 2.5$ | 0.66 |
| 400 | 65 | 680 | $0.7 \pm 0.2$ | $2.7 \pm 0.7 \pm 0.6$ | 1 | $-0.9$ | $6.1 \pm 1.0$ | 0.49 |
| 450 | 65 | 700 | $0.7 \pm 0.2$ | $2.3 \pm 0.6 \pm 0.4$ | 1 | $-0.7$ | $3.8 \pm 0.5$ | 0.64 |
| 500 | 65 | 770 | $0.5 \pm 0.1$ | $1.2 \pm 0.4 \pm 0.2$ | 1 | 0.0 | $1.6 \pm 0.2$ | 0.54 |
| 550 | 65 | 800 | $0.4 \pm 0.1$ | $0.9 \pm 0.4 \pm 0.2$ | 1 | $+0.3$ | $1.2 \pm 0.2$ | 0.73 |
| 600 | 65 | 850 | $0.3 \pm 0.1$ | $0.6 \pm 0.3 \pm 0.1$ | 1 | $+0.6$ | $0.6 \pm 0.1$ | 0.65 |
| 650 | 65 | 850 | $0.3 \pm 0.1$ | $0.6 \pm 0.3 \pm 0.1$ | 1 | $+0.6$ | $0.3 \pm 0.1$ | 0.60 |
| 700 | 85 | 850 | $0.1 \pm 0.1$ | $0.4 \pm 0.2 \pm 0.1$ | 0 | $-0.4$ | $0.2 \pm 0.1$ | 0.65 |
| 750 | 85 | 850 | $0.1 \pm 0.1$ | $0.4 \pm 0.2 \pm 0.1$ | 0 | $-0.4$ | $0.1 \pm 0.1$ | 0.79 |
| 800 | 85 | 850 | $0.1 \pm 0.1$ | $0.4 \pm 0.2 \pm 0.1$ | 0 | $-0.4$ | $0.1 \pm 0.1$ | 0.81 |

Table 7.1: Search results in the signal region for each of the $LQ_3$ mass hypotheses. The $\tau$ lepton $p_T$ and $S_T$ columns represent the optimized thresholds defined in Section 5.2. The corresponding expected number of prompt-prompt and total background events, as well as the observed number of data events are listed as $N^{PP}_{Bkg}$, total $N^{Exp}_{Bkg}$, and $N^{Obs}$. The statistical and systematic uncertainties quoted in the expected number of background events are combinations of misidentified lepton and prompt-prompt components. The $\epsilon_{LQ_3}$ is the expected signal efficiency at a given $LQ_3$ mass with respect to the total number of expected $LQ_3$ signal events at $\sqrt{s} = 8$ TeV with a same-sign $\mu\tau_h$ pair. No expected signal efficiency for $M_{LQ_3} = 200$ GeV is reported in the forward channel since the associated yield in the signal sample is measured to be zero.

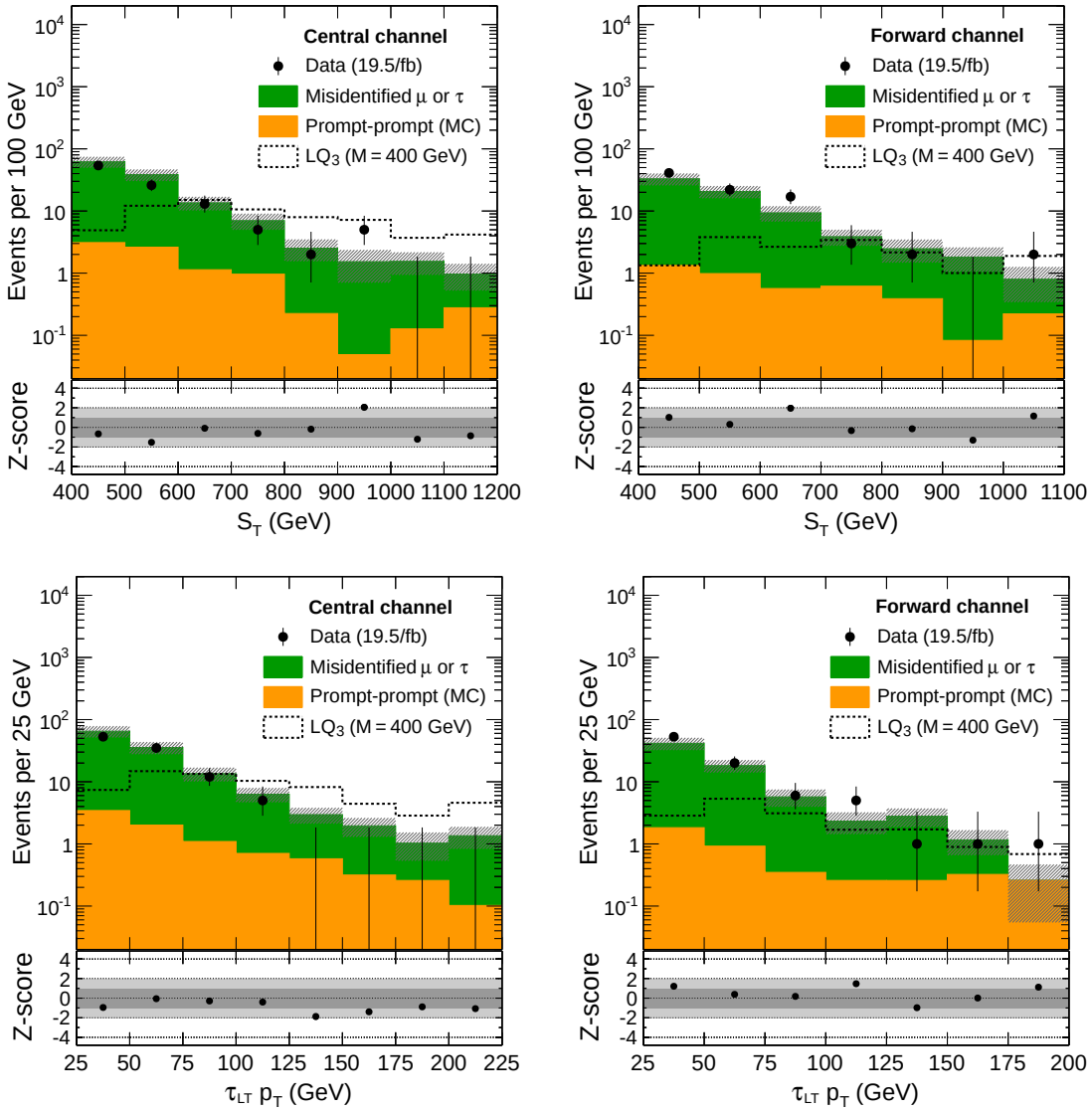


Figure 7.1: The $S_{\mathrm{T}}$ and $\tau$ lepton ($\tau_{\mathrm{LT}}$) $p_{\mathrm{T}}$ distributions for central (left column) and forward (right column) channels, using the optimized selection for $M_{\mathrm{LQ}_3} = 200$ GeV. All other optimized selection criteria yield events that are a subset of this selection. The rightmost bin of each distribution includes overflow and no statistically significant excess is observed in the suppressed bins. The systematic uncertainty for each bin of these distributions is determined independently. Shaded regions represent the total statistical and systematic uncertainty in the background expectation. The Z-score distribution is provided at the bottom of each plot. Reproduced from Ref. [153].

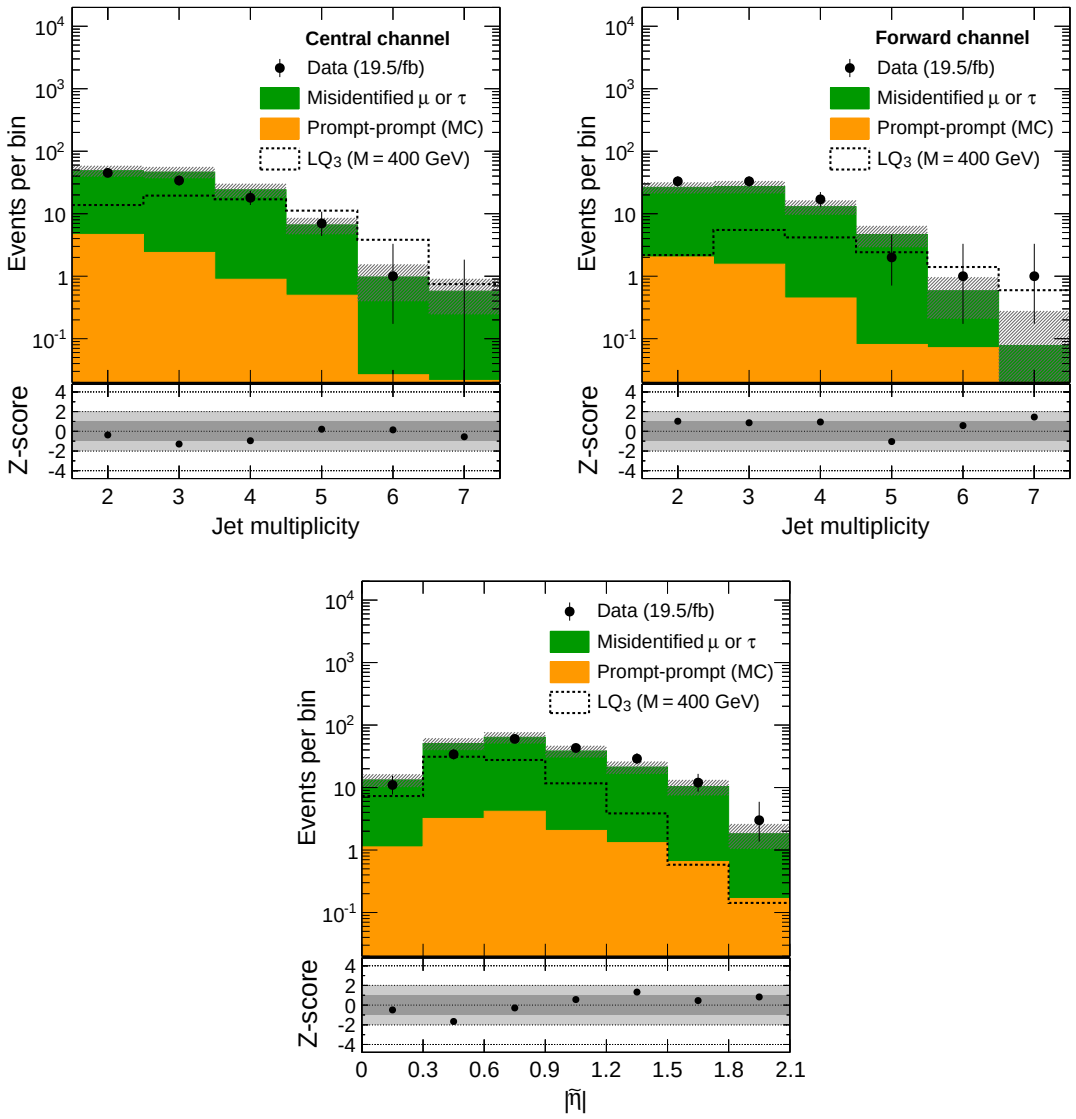


Figure 7.2: The jet multiplicity distributions for central (upper left) and forward (upper right) channels, and $|\widetilde{\eta}|$ (lower) distribution, using the optimized selection for $M_{\mathrm{LQ}_3} = 200$ GeV. All other optimized selection criteria yield events that are a subset of this selection. The rightmost bin of each distribution includes overflow and no statistically significant excess is observed in the suppressed bins. The systematic uncertainty for each bin of these distributions is determined independently. Shaded regions represent the total statistical and systematic uncertainty in the background expectation. The Z-score distribution is provided at the bottom of each plot. Reproduced from Ref. [153].

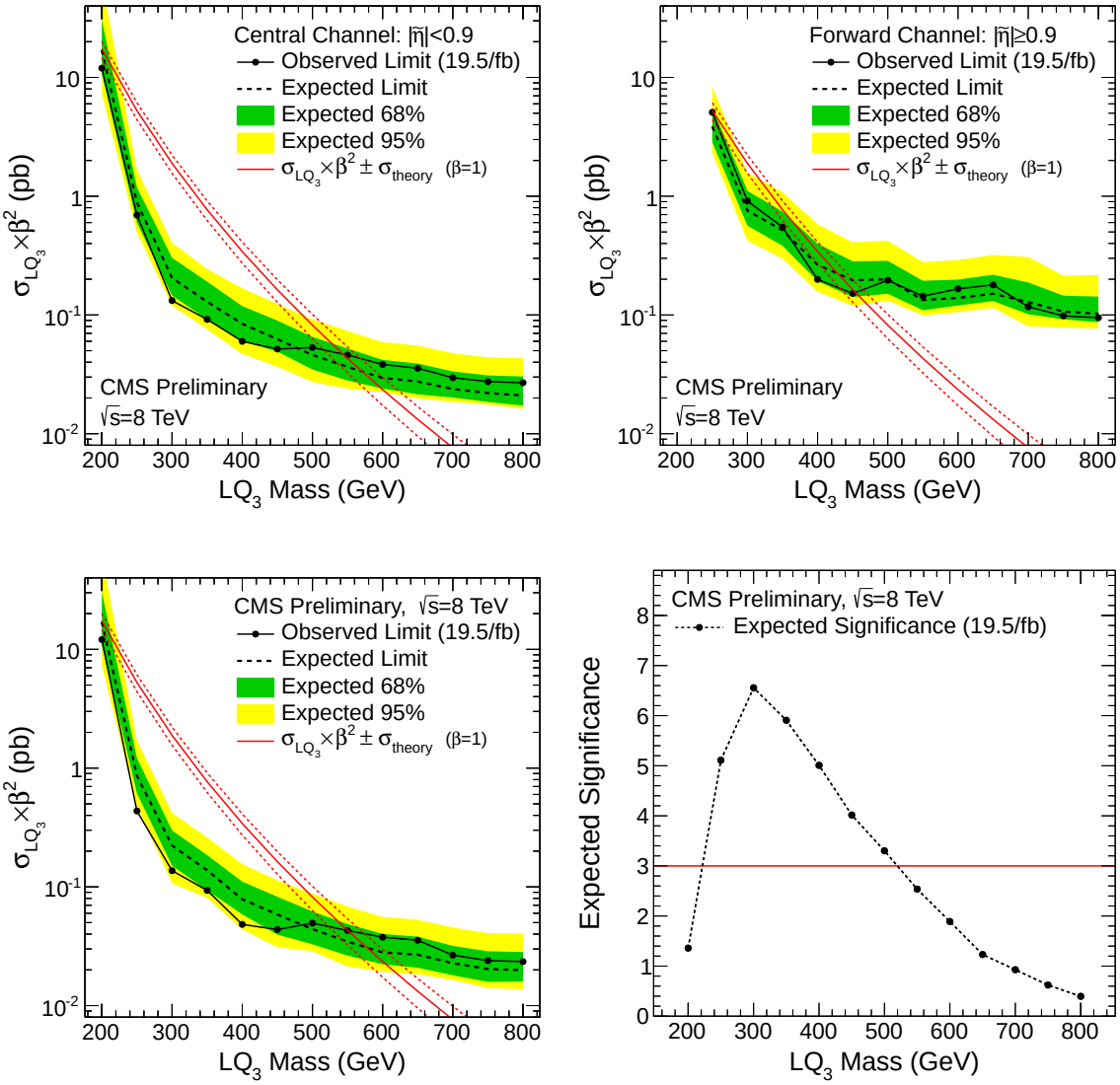


Figure 7.3: The expected and observed exclusion limits at 95% CL on the $LQ_3$ pair production cross section times $\beta^2$ in the central (upper left), forward (upper right), and combination (lower left) channels, as well as the expected significance as computed in the combination channel (lower right). The red-dashed curves represent the theoretical uncertainty on the $LQ_3$ pair production cross section due to the PDF and renormalization/factorization scale uncertainties. As expected, the central channel has greater sensitivity to the $LQ_3$ signal than the forward channel. In the forward channel, $M_{LQ_3} = 200$ GeV point is not included since the associated event selection has no signal efficiency [153].

# Chapter 8

# Conclusions

A search for pair produced, charge $-1/3$, third generation scalar leptoquarks decaying to top quark and $\tau$ lepton pairs has been conducted in the same-sign $\mu\tau_{\mathrm{h}}$ channel with two or more jets, using a proton-proton collisions data sample collected with the CMS detector at $\sqrt{s} = 8$ TeV corresponding to an integrated luminosity of 19.5 $\mathrm{fb}^{-1}$.

Two regions of $|\widetilde{\eta}|$ are considered with progressively tighter requirements on $S_{\mathrm{T}}$ and $\tau$ lepton $p_{\mathrm{T}}$ parameters, optimized for each of the $\mathrm{LQ}_3$ mass hypotheses. Backgrounds due to events with misidentified muons or $\tau$ leptons are estimated using data events via a matrix method, whereas irreducible backgrounds with same-sign $\mu\tau_{\mathrm{h}}$ pairs are estimated using simulated samples.

No statistically significant excess is observed over the SM background expectations. Assuming that all leptoquarks decay to a top quark and a $\tau$ lepton, the pair production of charge $-1/3$, third generation scalar leptoquarks is excluded at 95% CL for masses up to 550 GeV (582 GeV expected). The observed mass limit is also directly applicable to pair produced bottom squarks decaying via the RPV coupling $\lambda'_{333}$. This constitutes the first direct result for leptoquarks decaying in this channel as well as the first 8 TeV CMS result targeting the same-sign $\mu\tau_{\mathrm{h}}$ final state, and has been published as a CMS Physics Analysis Summary [153].

The results presented in this dissertation have also been included in a CMS combination publication [169], covering the same- and opposite-sign $\mathrm{e}\tau_{\mathrm{h}}$ and $\mu\tau_{\mathrm{h}}$ final states. This yields an improved exclusion limit of 685 GeV for leptoquarks exclusively decaying to a top quark and a $\tau$ lepton.

# Appendix A

# Data and Simulated Samples

## A.1 Data Samples

This thesis is based on the data collected by the CMS detector in 2012 during stable pp collisions at a center-of-mass energy of $\sqrt{s} = 8$ TeV, corresponding to an integrated luminosity of 19.5 fb$^{-1}$. The events have been acquired by a combination of isolated and non-isolated single muon HLTs, and stored in the *SingleMu Primary Dataset*. Table A.1 details the datasets, split into four periods labelled as A, B, C, and D, as well as the corresponding *run numbers*, *JSON files*, and *global tags*.

| SingleMu Primary Dataset | Run numbers | JSON† | Global tag |
|---|---|---|---|
| Run2012A-recover-06Aug2012-v1 | 190782-190949 | [1] | `FT_53_V6C_AN3` |
| Run2012A-13Jul2012-v1 | 190456-193621 | [2] | `FT_53_V6C_AN3` |
| Run2012B-13Jul2012-v1 | 193833-196531 | [2] | `FT_53_V6C_AN3` |
| Run2012C-24Aug2012-v1 | 198022-198913 | [3] | `FT53_V10A_AN3` |
| Run2012C-EcalRecover-11Dec2012-v1 | 201191-201191 | [4] | `FT_P_V42C_AN3` |
| Run2012C-PromptReco-v2 | 198934-203002 | [5] | `GR_P_V42_AN3` |
| Run2012D-PromptReco-v1 | 203768-208686 | [5] | `GR_P_V42_AN3` |

† JSON files are located at:
https://cms-service-dqm.web.cern.ch/cms-service-dqm/CAF/certification/Collisions12/8TeV/
[1] `Reprocessing/Cert_190782-190949_8TeV_06Aug2012ReReco_Collisions12_JSON.txt`
[2] `Reprocessing/Cert_190456-196531_8TeV_13Jul2012ReReco_Collisions12_JSON_v2.txt`
[3] `Reprocessing/Cert_198022-198523_8TeV_24Aug2012ReReco_Collisions12_JSON.txt`
[4] `Reprocessing/Cert_201191-201191_8TeV_11Dec2012ReReco-recover_Collisions12_JSON.txt`
[5] `Prompt/Cert_190456-208686_8TeV_PromptReco_Collisions12_JSON.txt`

Table A.1: The list of datasets with corresponding run numbers, JSON files, and global tags.

A run number is a unique identifier referring to the basic continuous data taking unit

(run) of the CMS detector. Each run is split into successive 23.31 s long periods (luminosity sections), which are individually certified for proper detector operation conditions before being fed into physics analyses. These certifications are in the form of dedicated Java Script Object Notation (JSON) files, specifying valid run numbers and luminosity sections, which are used to filter collision data for known detector problems. For a given collection of successive runs, the corresponding global tag [170] specifies the then-present detector conditions including alignment and calibration information to be used during event reconstruction.

Events used in this analysis were required to pass certain event filters to eliminate any possible effects of electronic noise, unfavorable beam conditions and other known detector hardware problems that might affect event reconstruction in an event-by-event basis. All data events have been reconstructed by the CMS software package [171] CMSSW_5_3 series, and were analyzed using CMSSW_5_3_8.

During 7 and 8 TeV pp runs, the CMS HCAL has suffered from single-event-upsets (SEUs) in the Clock and Control Modules (CCMs) triggered by the irradiation of optocouplers which resulted in erroneously issued reset signals [172]. Apart from setting and monitoring the operational configurations of the front-end-electronics, CCMs also provide a periodic clock signal synchronized with the LHC clock to each of the read-out boxes (RBXs) instrumenting each of the 20° wedges in the HCAL barrel and endcap regions. Hence, an SEU of this type in a given CCM results in disabling the data readout from an entire HCAL wedge. Although a vast majority of these incidents were detected and fixed within a time span of 10 s by ad hoc monitoring tools developed during the 8 TeV run, any recorded collision data overlapping with these incidents within the same luminosity sections had to be discarded. An effective cross section of about 0.09 pb has been observed for these so-called *RBX Data Loss* events in the HCAL barrel and endcap regions during the 8 TeV pp run, and they affected a total of 340 $pb^{-1}$ of data with over 2000 incidents as shown in Fig. A.1. Each of these incidents caused at least one luminosity section worth of collision data to be discarded in the subsequent data analyses, as also reflected in the JSON files given in Table A.1.

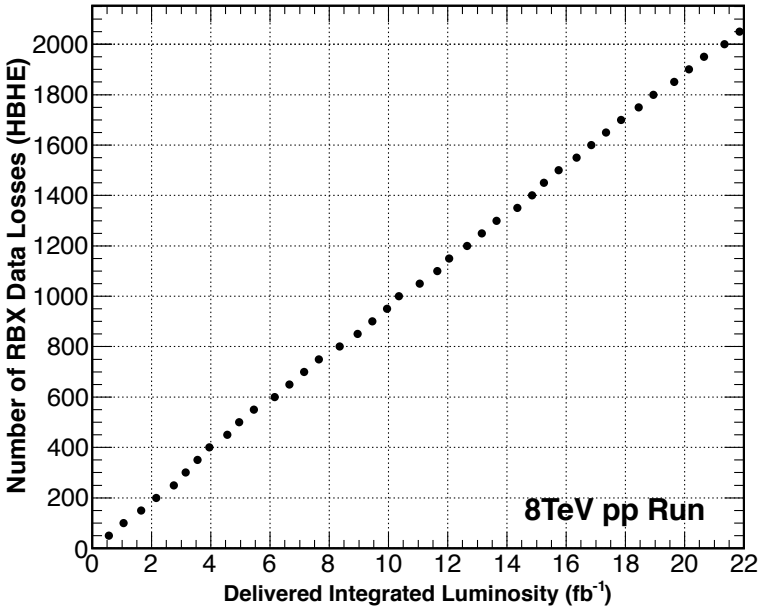


Figure A.1: The evolution of the integrated number of RBX data loss incidents at the CMS HCAL barrel and endcap regions as a function of the delivered integrated luminosity during the 8 TeV pp collisions at the LHC.

## A.2 Simulated Samples

SM background processes as well as the signal samples used in this analysis were centrally produced within the *CMS Summer12-DR53X* Monte Carlo (MC) generation campaign. Single top quark and top quark pair production samples were simulated with POWHEG (v1.0) [173, 174, 175], whereas MADGRAPH (v5.1) [176] was used for the W+jets, Drell-Yan(DY)+jets and $t\bar{t}W/t\bar{t}Z$ processes. Diboson, QCD multijet, and Higgs boson associated production samples were produced with PYTHIA (v6.426) [177]. All other SM samples were produced with MADGRAPH. The $LQ_3 \to t\tau^-$ signal samples were also produced with PYTHIA. The parton shower and hadronization in samples generated with POWHEG or MADGRAPH were performed with PYTHIA, and all $\tau$ lepton decays were simulated via TAUOLA [178]. For MADGRAPH samples, the matching to PYTHIA has been done with the MLM scheme [179]. In all these samples, the response of the CMS detector was simulated in GEANT4 [180], where event digitization and reconstruction were done in CMSSW_5_2 or CMSSW_5_3 series. All simulated samples were analyzed with the global tag `START53_V7G` using CMSSW_5_3_8.

These simulated samples are normalized using next-to-next-to-leading-order (NNLO) (W+jets, DY+jets [181], $t\bar{t}$+jets [182], WH, ZH [183]), approximate NNLO (t, tW [184]), next-to-leading-order (diboson [185], $t\bar{t}W$ [157, 156], $t\bar{t}Z$ [156], $t\bar{t}H$ [186, 187], tribo-

son [158]) or leading-order ($W^{\pm}W^{\pm}qq$, $t\bar{t}WW$, $W\gamma^{*}$, QCD multijet, other rare SM processes with same-sign leptons [176, 177]) cross sections at 8 TeV. The complete list detailing the generators for all the simulated SM and $LQ_3$ samples and the associated cross sections is given in Table A.3.

All MC events are required to satisfy one of the single muon triggers used in this analysis. To compensate for the differences in trigger performance between data and simulated samples, muon $p_{\mathrm{T}}$ and $\eta$ dependent re-weighting is implemented. Trigger efficiency scale factors are measured using a tag-and-probe method [188], and they denote the ratio of the probability of the selected muon candidate to pass the trigger requirements in data over simulated samples. Since one chosen muon candidate is required to match to a muon trigger object reconstructed by the HLT system ($\Delta R < 0.3$) in all events, the trigger reweighting amounts to an application of a simple multiplicative factor as listed in Table A.2.

| Trigger | Run Period | $\lvert\eta\rvert < 0.9$ | $0.9 < \lvert\eta\rvert < 1.2$ | $1.2 < \lvert\eta\rvert < 2.1$ |
|---|---|---|---|---|
| `HLT_IsoMu24_eta2p1` | | | | |
| | 2012A | $0.9560 \pm 0.0008$ | $0.9528 \pm 0.0021$ | $0.9809 \pm 0.0016$ |
| | 2012B | $0.9798 \pm 0.0004$ | $0.9618 \pm 0.0010$ | $0.9814 \pm 0.0008$ |
| | 2012C | $0.9841 \pm 0.0003$ | $0.9688 \pm 0.0009$ | $1.0021 \pm 0.0007$ |
| | 2012D | $0.9815 \pm 0.0003$ | $0.9616 \pm 0.0009$ | $0.9972 \pm 0.0007$ |
| `HLT_Mu40_eta2p1` | | | | |
| | 2012A | $0.9799 \pm 0.0013$ | $0.9621 \pm 0.0042$ | $0.9851 \pm 0.0034$ |
| | 2012B | $0.9773 \pm 0.0007$ | $0.9573 \pm 0.0021$ | $0.9754 \pm 0.0017$ |
| | 2012C | $0.9817 \pm 0.0004$ | $0.9640 \pm 0.0014$ | $0.9973 \pm 0.0011$ |
| | 2012D | $0.9830 \pm 0.0006$ | $0.9604 \pm 0.0019$ | $0.9955 \pm 0.0015$ |
| `HLT_Mu24_eta2p1` (prescaled) | 2012ABCD | $0.9770 \pm 0.0010$ | $0.9629 \pm 0.0020$ | $0.9913 \pm 0.0020$ |

Table A.2: Single muon HLT data-MC scale factors for the 8 TeV run [188].

In order to match the probability density function describing the pileup interactions per bunch crossing the MC samples were generated with to that observed in data events, an event-by-event reweighting is applied to all simulated SM and signal samples [189].

| | Sample name | Cross section |
|---|---|---|
| W+jets | WJetsToLNu_TuneZ2Star_8TeV-madgraph-tarball* (0 jets skim) | 27783 (pb) |
| | W1JetsToLNu_TuneZ2Star_8TeV-madgraph* | 6663 |
| | W2JetsToLNu_TuneZ2Star_8TeV-madgraph* | 2159 |
| | W3JetsToLNu_TuneZ2Star_8TeV-madgraph* | 640.4 |
| | W4JetsToLNu_TuneZ2Star_8TeV-madgraph* | 264.0 |
| DY+jets | DYJetsToLL_M-50_TuneZ2Star_8TeV-madgraph-tarball† (0 jets skim) | 2534 |
| | DY1JetsToLL_M-50_TuneZ2Star_8TeV-madgraph* | 666.3 |
| | DY2JetsToLL_M-50_TuneZ2Star_8TeV-madgraph* | 215.0 |
| | DY3JetsToLL_M-50_TuneZ2Star_8TeV-madgraph* | 60.7 |
| | DY4JetsToLL_M-50_TuneZ2Star_8TeV-madgraph* | 27.4 |
| | DYJetsToLL_M-10To50_TuneZ2Star_8TeV-madgraph* | 11050 |
| t/tW | T_s-channel_TuneZ2star_8TeV-powheg-tauola* | 3.79 |
| | T_t-channel_TuneZ2star_8TeV-powheg-tauola* | 56.4 |
| | T_tW-channel-DR_TuneZ2star_8TeV-powheg-tauola* | 11.1 |
| | Tbar_s-channel_TuneZ2star_8TeV-powheg-tauola* | 1.76 |
| | Tbar_t-channel_TuneZ2star_8TeV-powheg-tauola* | 30.7 |
| | Tbar_tW-channel-DR_TuneZ2star_8TeV-powheg-tauola* | 11.1 |
| $t\bar{t}$+jets | TTJets_FullLeptMGDecays_8TeV-madgraph* | 26.2 |
| | TTJets_SemiLeptMGDecays_8TeV-madgraph* | 103.7 |
| | TTJets_HadronicMGDecays_8TeV-madgraph* | 104.1 |
| QCD multijet | QCD_Pt-15to20_MuEnrichedPt5_TuneZ2star_8TeV_pythia6* | $2.74 \cdot 10^6$ |
| | QCD_Pt-20to30_MuEnrichedPt5_TuneZ2star_8TeV_pythia6* | $1.87 \cdot 10^6$ |
| | QCD_Pt-30to50_MuEnrichedPt5_TuneZ2star_8TeV_pythia6* | $8.06 \cdot 10^5$ |
| | QCD_Pt-50to80_MuEnrichedPt5_TuneZ2star_8TeV_pythia6* | $1.76 \cdot 10^5$ |
| | QCD_Pt-80to120_MuEnrichedPt5_TuneZ2star_8TeV_pythia6* | $4.04 \cdot 10^4$ |
| | QCD_Pt-120to170_MuEnrichedPt5_TuneZ2star_8TeV_pythia6* | $7.46 \cdot 10^3$ |
| | QCD_Pt-170to300_MuEnrichedPt5_TuneZ2star_8TeV_pythia6* | $2.30 \cdot 10^3$ |
| | QCD_Pt-300to470_MuEnrichedPt5_TuneZ2star_8TeV_pythia6* | $1.52 \cdot 10^2$ |
| | QCD_Pt-470to600_MuEnrichedPt5_TuneZ2star_8TeV_pythia6* | $1.18 \cdot 10^1$ |
| | QCD_Pt-600to800_MuEnrichedPt5_TuneZ2star_8TeV_pythia6* | 2.69 |
| | QCD_Pt-800to1000_MuEnrichedPt5_TuneZ2star_8TeV_pythia6* | $3.69 \cdot 10^{-1}$ |
| | QCD_Pt-1000_MuEnrichedPt5_TuneZ2star_8TeV_pythia6* | $8.49 \cdot 10^{-2}$ |
| Diboson | WW_TuneZ2star_8TeV_pythia6_tauola* | 57.1 |
| | WZ_TuneZ2star_8TeV_pythia6_tauola* | 32.3 |
| | ZZ_TuneZ2star_8TeV_pythia6_tauola* | 17.0 |
| Rare SM | TTGJets_8TeV-madgraph* | 2.166 |
| | TTWJets_8TeV-madgraph* | 0.232 |
| | TTZJets_8TeV-madgraph_v2* | 0.208 |
| | TBZToLL_4F_TuneZ2star_8TeV-madgraph-tauola* | 0.0114 |
| | TTWWJets_8TeV-madgraph* | 0.002 |
| | TTTT_TuneZ2star_8TeV-madgraph-tauola* | 0.00091 |
| | WW_DoubleScattering_8TeV-pythia8* | 0.5879 |
| | WWGJets_8TeV-madgraph_v2* | 0.528 |
| | WGstarToLNu2Mu_TuneZ2star_7TeV-madgraph-tauola* | 1.914 |
| | WGstarToLNu2Tau_TuneZ2star_7TeV-madgraph-tauola* | 0.336 |
| | WpWpqq_8TeV-madgraph* | 0.248 |
| | WmWmqq_8TeV-madgraph* | 0.0889 |
| | WWWJets_8TeV-madgraph* | 0.0822 |
| | WWZNoGstarJets_8TeV-madgraph* | 0.063 |
| | WZZNoGstarJets_8TeV-madgraph* | 0.019 |
| | ZZZNoGstarJets_8TeV-madgraph* | 0.0046 |
| | WH_ZH_TTH_HToWW_M-125_8TeV-pythia6* | 0.260 |
| | WH_ZH_TTH_HToZZ_M-125_8TeV-pythia6* | 0.032 |
| | WH_ZH_TTH_HToTauTau_M-125_8TeV-pythia6-tauola* | 0.078 |
| $LQ_3$ | LQToTTau_M-**X**_TuneZ2star_8TeV_pythia6* (**X**=200,250,300,..,800) | See Table 2.2 |
| | *Summer12_DR53X-PU_S10_START53 †Summer12-PU_S7_START52 | |

Table A.3: Simulated SM and $LQ_3$ samples and cross sections at $\sqrt{s} = 8$ TeV.

# References


[1] G. Altarelli in *Encyclopedia of Mathematical Physics*, ch. Standard Model of Particle Physics, p. 32, Academic Press, 2006, arXiv:hep-ph/0510281, CERN-PH-TH-2005-206.

[2] M. Maggiore, *A Modern Introduction to Quantum Field Theory*, Oxford University Press, 2005.

[3] C. G. Tully, *Elementary Particle Physics in a Nutshell*, Princeton University Press, 2011.

[4] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, Westview Press, 1995.

[5] C. Burgess and G. Moore, *The Standard Model: A Primer*, Cambridge University Press, 2007.

[6] M. D. Schwartz, *Quantum Field Theory and the Standard Model*, Cambridge University Press, 2013.

[7] A. Zee, *Quantum Field Theory in a Nutshell*, Princeton University Press, 2010.

[8] E. H. Lieb and R. Seiringer, *The Stability of Matter in Quantum Mechanics*, Cambridge University Press, 2010.

[9] M. Drewes, "The Phenomenology of Right Handed Neutrinos", *Int. J. Mod. Phys.* **E22** (2013) 1330019, arXiv:1303.6912.

[10] 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.* **B716** (2012) 1, arXiv:1207.7214.

[11] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett.* **B716** (2012) 30, arXiv:1207.7235.

[12] CMS Collaboration, "Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV", *JHEP* **1306** (2013) 081, arXiv:1303.4571.

[13] CMS Collaboration, "Evidence for the direct decay of the 125 GeV Higgs boson to fermions", *Nature Phys.* **10** (2014) 557, arXiv:1401.6527.

[14] ALEPH Collaboration, *et al.*, "Precision Electroweak Measurements and Constraints on the Standard Model", arXiv:1012.2367, CERN-PH-EP-2010-095.

[15] A. Pomarol, "Beyond the Standard Model", arXiv:1202.1391, CERN-2012-001.

[16] P. Langacker, "Grand Unified Theories and Proton Decay", *Phys. Rept.* **72** (1981) 185.

[17] Y. A. Golfand and E. P. Likhtman, "Extension of the algebra of poincare group generators and violation of p invariance", *JETP Lett.* **13** (1971) 323.

[18] D. V. Volkov and V. P. Akulov, "Is the neutrino a goldstone particle?", *Phys. Lett.* **B46** (1973) 109.

[19] J. Wess and B. Zumino, "Supergauge transformations in four-dimensions", *Nucl. Phys.* **B70** (1974) 39.

[20] E. Farhi and L. Susskind, "Technicolor", *Phys. Rept.* **74** (1981) 277.

[21] S. Davidson, D. C. Bailey, and B. A. Campbell, "Model independent constraints on leptoquarks from rare processes", *Z. Phys.* **C61** (1994) 613, arXiv:hep-ph/9309310.

[22] I. Dorsner, S. Fajfer, J. F. Kamenik, and N. Kosnik, "Can scalar leptoquarks explain the f(D(s)) puzzle?", *Phys. Lett.* **B682** (2009) 67, arXiv:0906.5585.

[23] J. C. Pati and A. Salam, "Lepton Number as the Fourth Color", *Phys. Rev.* **D10** (1974) 275.

[24] H. Georgi and S. L. Glashow, "Unity of all elementary-particle forces", *Phys. Rev. Lett.* **32** (1974) 438.

[25] S. S. Gershtein, A. A. Likhoded, and A. I. and Onishchenko, "TeV-scale leptoquarks from GUTs/string/M-theory unification", *Phys. Rept.* **320** (1999) 159.

[26] P. Nath and P. Fileviez Perez, "Proton stability in grand unified theories, in strings and in branes", *Phys. Rept.* **441** (2007) 191, arXiv:hep-ph/0601023.

[27] P. H. Frampton, "Light leptoquarks as possible signature of strong electroweak unification", *Mod. Phys. Lett.* **A7** (1992) 559.

[28] I. Dorsner and P. F. Perez, "Unification without supersymmetry: Neutrino mass, proton decay and light leptoquarks", *Nucl. Phys.* **B723** (2005) 53, arXiv:hep-ph/0504276.

[29] P. Yu. Popov, A. V. Povarov, and A. D. Smirnov, "Fermionic decays of scalar leptoquarks and scalar gluons in the minimal four color symmetry model", *Mod. Phys. Lett.* **A20** (2005) 3003, arXiv:hep-ph/0511149.

[30] S. P. Martin, "A Supersymmetry primer", *Adv. Ser. Direct. High Energy Phys.* **21** (2010) 1, arXiv:hep-ph/9709356.

[31] R. Barbier, C. Berat, M. Besancon, M. Chemtob, A. Deandrea, *et al.*, "R-parity violating supersymmetry", *Phys. Rept.* **420** (2005) 1, arXiv:hep-ph/0406039.

[32] J. L. Feng, "Naturalness and the Status of Supersymmetry", *Ann. Rev. Nucl. Part. Sci.* **63** (2013) 351, arXiv:1302.6587.

[33] B. Gripaios, "Composite Leptoquarks at the LHC", *JHEP* **1002** (2010) 045, arXiv:0910.1789.

[34] W. Buchmüller, R. Rückl, and D. Wyler, "Leptoquarks in Lepton - Quark Collisions", *Phys. Lett.* **B191** (1987) 442. Erratum: *Phys. Lett.* **B448** (1999) 320.

[35] M. Hirsch, H. V. Klapdor-Kleingrothaus, and S. G. Kovalenko, "New low-energy leptoquark interactions", *Phys. Lett.* **B378** (1996) 17, arXiv:hep-ph/9602305.

[36] Y. Sirois in *Hadron Collider Physics 2002*, ch. Searches and Discovery Prospects at HERA, p. 380, Springer Berlin Heidelberg, 2003.

[37] J. L. Hewett and T. G. Rizzo, "Much ado about leptoquarks: A Comprehensive analysis", *Phys. Rev.* **D56** (1997) 5709, arXiv:hep-ph/9703337.

[38] M. Leurer, "New bounds on leptoquarks", *Phys. Rev. Lett.* **71** (1993) 1324, arXiv:hep-ph/9304211.

[39] M. Leurer, "A Comprehensive study of leptoquark bounds", *Phys. Rev.* **D49** (1994) 333, arXiv:hep-ph/9309266.

[40] D. E. Acosta and S. K. Blessing, "Leptoquark searches at HERA and the Tevatron", *Ann. Rev. Nucl. Part. Sci.* **49** (1999) 389.

[41] J. L. Hewett, T. G. Rizzo, S. Pakvasa, H. E. Haber, and A. Pomarol, "Vector leptoquark production at hadron colliders", arXiv:hep-ph/9310361.

[42] M. Leurer, "Bounds on vector leptoquarks", *Phys. Rev.* **D50** (1994) 536, arXiv:hep-ph/9312341.

[43] K. A. Olive *et al.* (Particle Data Group), "Review of Particle Physics", *Chin. Phys.* **C38** (2014) 090001.

[44] J. Blumlein, E. Boos, and A. Kryukov, "Leptoquark pair production in hadronic interactions", *Z. Phys.* **C76** (1997) 137, arXiv:hep-ph/9610408.

[45] O. J. P. Eboli, R. Zukanovich Funchal, and T. L. Lungov, "Multilepton signatures for leptoquarks", *Phys. Rev.* **D59** (1999) 035002, arXiv:hep-ph/9808288.

[46] M. Krämer, T. Plehn, M. Spira, and P. M. Zerwas, "Pair production of scalar leptoquarks at the CERN LHC", *Phys. Rev.* **D71** (2005) 057503.

[47] A. Belyaev, C. Leroy, R. Mehdiyev, and A. Pukhov, "Leptoquark single and pair production at LHC with CalcHEP/CompHEP in the complete model", *JHEP* **0509** (2005) 005, arXiv:hep-ph/0502067.

[48] D0 Collaboration, "Search for third-generation leptoquarks in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.96 TeV", *Phys. Rev. Lett.* **99** (2007) 061801, arXiv:0705.0812.

[49] O. J. P. Eboli, "Bounds on scalar leptoquarks from the LEP data", arXiv:hep-ph/9508342, IFUSP-P-1170.

[50] J. K. Mizukoshi, O. J. P. Eboli, and M. C. Gonzalez-Garcia, "Bounds on scalar leptoquarks from Z physics", *Nucl. Phys.* **B443** (1995) 20, arXiv:hep-ph/9411392.

[51] C. Friberg, E. Norrbin, and T. Sjöstrand, "QCD aspects of leptoquark production at HERA", *Phys. Lett.* **B403** (1997) 329, arXiv:hep-ph/9704214.

[52] D. Bowser-Chao, T. D. Imbo, B. A. King, and E. C. Martell, "Leptomesons, leptoquarkonium and the QCD potential", *Phys. Lett.* **B432** (1998) 167, arXiv:hep-ph/9803348.

[53] V. V. Kiselev, "Perturbative fragmentation of a scalar leptoquark into a heavy leptoquarkonium", *Phys. Atom. Nucl.* **62** (1999) 300, IHEP-97-78.

[54] L. Evans and P. Bryant, "LHC Machine", *JINST* **3** (2008) S08001.

[55] O. Bruning, *et al.*, "LHC Design Report", CERN-2004-003-V-1, CERN, Geneva, 2004.

[56] ALICE Collaboration, "The ALICE experiment at the CERN LHC", *JINST* **3** (2008) S08002.

[57] ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider", *JINST* **3** (2008) S08003.

[58] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004.

[59] LHCb Collaboration, "The LHCb Detector at the LHC", *JINST* **3** (2008) S08005.

[60] P. Steinberg, "First results from the LHC heavy ion program", *New J. Phys.* **14** (2012) 035006.

[61] R. Singh, L. Kumar, P. K. Netrakanti, and B. Mohanty, "Selected Experimental Results from Heavy Ion Collisions at LHC", *Adv. High Energy Phys.* **2013** (2013) 761474, arXiv:1304.2969.

[62] CERN Communication Group, "CERN LHC: the guide", CERN-BROCHURE-2006-003-ENG, CERN, Geneva, 2006.

[63] M. Lamont, "Status of the LHC", *J. Phys. Conf. Ser.* **455** (2013) 012001.

[64] B. Goddard and S. Dubourg, "Proceedings of the 2012 Evian Workshop on LHC Beam Operation", CERN-ATS-2013-045, CERN, Geneva, 2013.

[65] M. Lamont, "The LHC's first long run", *CERN Courier Volume 53, Number 7* (2013) .

[66] F. Zimmermann, "LHC: The Machine", lecture presented at the $40^{th}$ SLAC Summer Institute, 2012.

[67] CMS Collaboration, "Public CMS Data Quality Information", https://twiki.cern.ch/twiki/bin/view/CMSPublic/DataQuality?rev=6. Accessed: September 18, 2014.

[68] CMS Collaboration, "Public CMS Luminosity Information", https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults?rev=101. Accessed: August 12, 2014.

[69] CMS Collaboration, "The CMS Magnet Project: Technical Design Report", CERN-LHCC-97-10, CERN, Geneva, 1997.

[70] CMS Collaboration, "Precise Mapping of the Magnetic Field in the CMS Barrel Yoke using Cosmic Rays", *JINST* **5** (2010) T03021, arXiv:0910.5530.

[71] CMS Collaboration, "The CMS tracker system project: Technical Design Report", CERN-LHCC-98-006, CERN, Geneva, 1997.

[72] CMS Collaboration, "The CMS tracker: addendum to the Technical Design Report", CERN-LHCC-2000-016, CERN, Geneva, 2000.

[73] CMS Collaboration, "Commissioning and Performance of the CMS Silicon Strip Tracker with Cosmic Ray Muons", *JINST* **5** (2010) T03008, arXiv:0911.4996.

[74] CMS Collaboration, "Commissioning and Performance of the CMS Pixel Tracker with Cosmic Ray Muons", *JINST* **5** (2010) T03007, arXiv:0911.5434.

[75] CMS Collaboration, "CMS Physics: Technical Design Report Volume 1: Detector Performance and Software", CERN-LHCC-2006-001, CERN, Geneva, 2006.

[76] CMS Collaboration, "CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update", CMS Physics Analysis Summary CMS-PAS-LUM-13-001, CERN, Geneva, 2013.

[77] CMS Collaboration, "The CMS electromagnetic calorimeter project: Technical Design Report", CERN-LHCC-97-33, CERN, Geneva, 1997.

[78] R. M. Brown, "Avalanche photodiodes and vacuum phototriodes for the CMS electromagnetic calorimeter", *Nucl. Instr. and Meth.* **A572** (2007) 26.

[79] CMS Collaboration, "Performance and Operation of the CMS Electromagnetic Calorimeter", *JINST* **5** (2010) T03010, arXiv:0910.3423.

[80] P. Adzic, R. Alemany-Fernandez, C. B. Almeida, N. M. Almeida, G. Anagnostou, *et al.*, "Energy resolution of the barrel of the CMS electromagnetic calorimeter", *JINST* **2** (2007) P04004.

[81] C. W. Fabjan and F. Gianotti, "Calorimetry for particle physics", *Rev. Mod. Phys.* **75** (2003) 1243.

[82] CMS Collaboration, "Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter in $pp$ Collisions at $\sqrt{s} = 7$ TeV", *JINST* **8** (2013) P09009.

[83] CMS Collaboration, "The CMS hadron calorimeter project: Technical Design Report", CERN-LHCC-97-031, CERN, Geneva, 1997.

[84] R. Wigmans, "Effects of calorimeter peculiarities on the jet energy scale", *J. Phys. Conf. Ser.* **323** (2011) 012001.

[85] CMS Collaboration, "Performance of the CMS Hadron Calorimeter with Cosmic Ray Muons and LHC Beam Data", *JINST* **5** (2010) T03012, arXiv:0911.4991.

[86] CMS HCAL Collaboration, "Design, performance, and calibration of CMS hadron-barrel calorimeter wedges", *Eur. Phys. J.* **C55** (2008) 159.

[87] P. Cushman, A. Heering, and A. Ronzhin, "Custom HPD readout for the CMS HCAL", *Nucl. Instrum. Meth.* **A442** (2000) 289.

[88] A. Heering *et al.*, "Performance of silicon photomultiplers with the CMS HCAL front-end electronics", *Nucl. Instrum. Meth.* **A576** (2007) 341.

[89] P. Cushman and B. Sherwood, "Lifetime studies of the 19-channel hybrid photodiode for the CMS hadronic calorimeter", *Nucl. Instrum. Meth.* **A587** (2008) 250.

[90] J. Freeman, "Innovations for the CMS HCAL", *Int. J. Mod. Phys.* **A25** (2010) 2421.

[91] CMS HCAL/ECAL Collaborations, "The CMS barrel calorimeter response to particle beams from 2-GeV/c to 350-GeV/c", *Eur. Phys. J.* **C60** (2009) 359.

[92] CMS HCAL Collaboration, "Design, performance, and calibration of the CMS Hadron-outer calorimeter", *Eur. Phys. J.* **C57** (2008) 653.

[93] CMS HCAL Collaboration, "Design, performance and calibration of the CMS forward calorimeter wedges", *Eur. Phys. J.* **C53** (2008) 139.

[94] CMS Collaboration, "The CMS muon project: Technical Design Report", CERN-LHCC-97-032, CERN, Geneva, 1997.

[95] CMS Collaboration, "The performance of the CMS muon detector in proton-proton collisions at $\sqrt{s}$ = 7 TeV at the LHC", *JINST* **8** (2013) P11002, arXiv:1306.6905.

[96] CMS Collaboration, "Performance of CMS muon reconstruction in $pp$ collision events at $\sqrt{s} = 7$ TeV", *JINST* **7** (2012) P10002, arXiv:1206.4071.

[97] TOTEM Collaboration, "Luminosity-Independent Measurement of the Proton-Proton Total Cross Section at $\sqrt{s} = 8$ TeV", *Phys. Rev. Lett.* **111** (2013) 012001.

[98] ATLAS Collaboration, "Measurement of the Inelastic Proton-Proton Cross-Section at $\sqrt{s} = 7$ TeV with the ATLAS Detector", *Nature Commun.* **2** (2011) 463, arXiv:1104.0326.

[99] CMS Collaboration, "Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 7$ TeV", *Phys. Lett.* **B722** (2013) 5, arXiv:1210.6718.

[100] G. Bauer, U. Behrens, M. Bowen, J. Branson, S. Bukowiec, *et al.*, "Operational experience with the CMS data acquisition system", *J. Phys. Conf. Ser.* **396** (2012) 012007.

[101] H. P. Beck, "Triggering at high luminosity colliders", *New J. Phys.* **9** (2007) 334, arXiv:0704.2548.

[102] V. O'Dell, "Removing the haystack: The CMS trigger and data acquisition systems", *Mod. Phys. Lett.* **A25** (2010) 749.

[103] CMS Collaboration, "Operation and Performance of the CMS Level-1 Trigger during 7 TeV Collisions", *Phys. Procedia* **37** (2012) 1908.

[104] CMS Collaboration, "Performance of the CMS Level-1 Trigger", *PoS* **ICHEP2012** (2013) 508, arXiv:1302.2469, CERN-CMS-CR-2012-322.

[105] CMS Collaboration, "Commissioning of the CMS High-Level Trigger with Cosmic Rays", *JINST* **5** (2010) T03005, arXiv:0911.4889.

[106] CMS Collaboration, "The CMS High Level Trigger", *PoS* **ICHEP2012** (2013) 505, CMS-CR-2012-355.

[107] D. Barney, "CMS Public Document: 5581-v1", https://cms-docdb.cern.ch/cgi-bin/PublicDocDB/ShowDocument?docid=5581. Reproduced in http://www.ba.infn.it/~pierro/ebook/cms/slice_white.pdf. Accessed: September 19, 2014.

[108] CMS Collaboration, "CMS Tracking Performance Results from early LHC Operation", *Eur. Phys. J.* **C70** (2010) 1165, arXiv:1007.1988.

[109] CMS Collaboration, "Description and performance of track and primary-vertex reconstruction with the CMS tracker", *JINST* **9** (2014) P10009, arXiv:1405.6569.

[110] S. Cucciarelli, M. Konecki, D. Kotlinski, and T. Todorov, "Track reconstruction, primary vertex finding and seed generation with the pixel detector", CMS Note CMS-NOTE-2006-026, CERN, Geneva, 2006.

[111] T. Miao, H. Wenzel, F. Yumiceva, and N. Leioatts, "Beam position determination using tracks", CMS Note CMS-NOTE-2007-021, CERN, Geneva, 2007.

[112] R Fruhwirth, W. Waltenberger, and P. Vanlaer, "Adaptive Vertex Fitting", CMS Note CMS-NOTE-2007-008, CERN, Geneva, 2007.

[113] S. Baffioni, C. Charlot, F. Ferri, D. Futyan, P. Meridiani, *et al.*, "Electron reconstruction in CMS", *Eur. Phys. J.* **C49** (2007) 1099.

[114] CMS Collaboration, "Electron reconstruction and identification at $\sqrt{s} = 7$ TeV", CMS Physics Analysis Summary CMS-PAS-EGM-10-004, CERN, Geneva, 2010.

[115] CMS Collaboration, "Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV", arXiv:1502.02701, CMS-EGM-13-001.

[116] N. Almeida, P. Silva, J.C. Da Silva, M. Husejko, A. Jain, *et al.*, "Data filtering in the readout of the CMS electromagnetic calorimeter", *JINST* **3** (2008) P02011.

[117] E. Meschi, T. Monteiro, C. Seez, and P. Vikas, "Electron Reconstruction in the CMS Electromagnetic Calorimeter", CMS Note CMS-NOTE-2001-034, CERN, Geneva, 2001.

[118] W. Adam, R. Fruhwirth, A. Strandlie, and T. Todorov, "Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC", *J. Phys. G: Nucl. Part. Phys.* **31** (2005) N9.

[119] G. Bruno, P. T. Cox, S. Lacaprara, N. Neumeister, B. Van de Vyver, *et al.*, "Local Reconstruction in the Muon Detectors", CMS Note CMS-NOTE-2002-043, CERN, Geneva, 2002.

[120] R. Fruhwirth, "Application of Kalman filtering to track and vertex fitting", *Nucl. Instrum. Meth.* **A262** (1987) 444.

[121] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET", CMS Physics Analysis Summary CMS-PAS-PFT-09-001, CERN, Geneva, 2009.

[122] CMS Collaboration, "Commissioning of the particle-flow event reconstruction with leptons from J/Psi and W decays at 7 TeV", CMS Physics Analysis Summary CMS-PAS-PFT-10-003, CERN, Geneva, 2010.

[123] J. A. C. Ballin, "Particle Flow at CMS and the ILC", CERN Thesis CERN-THESIS-2010-086, Imperial College, London, 2010.

[124] G. P. Salam, "Towards Jetography", *Eur. Phys. J.* **C67** (2010) 637, arXiv:0906.1833.

[125] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, "Longitudinally invariant $k_\perp$ clustering algorithms for hadron hadron collisions", *Nucl. Phys.* **B406** (1993) 187.

[126] S. D. Ellis and D. E. Soper, "Successive combination jet algorithm for hadron collisions", *Phys. Rev.* **D48** (1993) 3160, arXiv:hep-ph/9305266.

[127] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B.R. Webber, "Better jet clustering algorithms", *JHEP* **9708** (1997) 001, arXiv:hep-ph/9707323.

[128] G. P. Salam and G. Soyez, "A Practical Seedless Infrared-Safe Cone jet algorithm", *JHEP* **0705** (2007) 086, arXiv:0704.0292.

[129] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-$k_\mathrm{T}$ jet clustering algorithm", *JHEP* **0804** (2008) 063, arXiv:0802.1189.

[130] CMS Collaboration, "Jet Algorithms (JA) Group", https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetAlgorithms?rev=159. Accessed: January 13, 2015 (CMS internal content).

[131] CMS Collaboration, "Jet Performance in pp Collisions at $\sqrt{s} = 7$ TeV", CMS Physics Analysis Summary CMS-PAS-JME-10-003, CERN, Geneva, 2010.

[132] CMS Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS", *JINST* **6** (2011) P11002, arXiv:1107.4277.

[133] CMS Collaboration, "Performance of the missing transverse energy reconstruction by the CMS experiment in $\sqrt{s} = 8$ TeV pp data", *JINST* **10** (2015) P02006, arXiv:1411.0511.

[134] CMS Collaboration, "Performance of $\tau$-lepton reconstruction and identification in CMS", *JINST* **7** (2012) P01001, arXiv:1109.6034.

[135] CMS Collaboration, "Tau reconstruction and identification in CMS during LHC run 1", CMS Physics Analysis Summary CMS-PAS-TAU-14-001, CERN, Geneva, 2015.

[136] M. Cacciari and G. P. Salam, "Pileup subtraction using jet areas", *Phys. Lett.* **B659** (2008) 119, arXiv:0707.1378.

[137] M. Cacciari, G. P. Salam, and G. Soyez, "The Catchment Area of Jets", *JHEP* **0804** (2008) 005, arXiv:0802.1188.

[138] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual", *Eur. Phys. J.* **C72** (2012) 1896, arXiv:1111.6097.

[139] CMS Collaboration, "Cut-based Electron Identification", https://twiki.cern.ch/twiki/bin/viewauth/CMS/EgammaCutBasedIdentification?rev=30. Accessed: January 16, 2015 (CMS internal content).

[140] CMS Collaboration, "Baseline muon selections", https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideMuonId?rev=49. Accessed: December 4, 2014.

[141] CMS Collaboration, "Software Guide for Tau Reconstruction", https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuidePFTauID?rev=157. Accessed: January 14, 2015.

[142] H. Saka and C. G. Tully, "Search for Third Generation Scalar Leptoquarks Decaying to Top Quark - Tau Lepton Pairs in pp Collisions", CMS Analysis Note CMS-AN-2012-289, CERN, Geneva, 2014 (CMS internal content).

[143] CMS Collaboration, "Jet Identification", https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetID?rev=75. Accessed: January 20, 2015 (CMS internal content).

[144] ATLAS Collaboration, "Search for first generation scalar leptoquarks in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector", *Phys. Lett.* **B709** (2012) 158, arXiv:1112.4828. Erratum: *Phys. Lett.* **B711** (2012) 442.

[145] ATLAS Collaboration, "Search for second generation scalar leptoquarks in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector", *Eur. Phys. J.* **C72** (2012) 2151, arXiv:1203.3172.

[146] CMS Collaboration, "Search for pair production of first- and second-generation scalar leptoquarks in pp collisions at $\sqrt{s} = 7$ TeV", *Phys. Rev.* **D86** (2012) 052013, arXiv:1207.5406.

[147] CDF Collaboration, "Search for Third Generation Vector Leptoquarks in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV", *Phys. Rev.* **D77** (2008) 091105, arXiv:0706.2832.

[148] ATLAS Collaboration, "Search for third generation scalar leptoquarks in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector", *JHEP* **1306** (2013) 033, arXiv:1303.0526.

[149] CMS Collaboration, "Search for third-generation leptoquarks and scalar bottom quarks in $pp$ collisions at $\sqrt{s} = 7$ TeV", *JHEP* **1212** (2012) 055, arXiv:1210.5627.

[150] CMS Collaboration, "Search for pair production of third-generation scalar leptoquarks and top squarks in proton-proton collisions at $\sqrt{s} = 8$ TeV", *Phys. Lett.* **B739** (2014) 229, arXiv:1408.0806.

[151] G. Punzi, "Sensitivity of searches for new signals and its optimization", *eConf* **C030908** (2003) MODT002, arXiv:physics/0308063.

[152] CMS Collaboration, "CMS Exotica Public Physics Results: Search for third-generation scalar leptoquarks in the $t\tau$ channel in proton-proton collisions at $\sqrt{s} = 8$ TeV (CMS-EXO-14-008)", https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsEXO14008?rev=34. Accessed: April 10, 2015.

[153] CMS Collaboration, "Search for Third Generation Scalar Leptoquarks Decaying to Top Quark - Tau Lepton Pairs in pp Collisions", CMS Physics Analysis Summary CMS-PAS-EXO-12-030, CERN, Geneva, 2014.

[154] CMS Collaboration, "Measurement of the $t\bar{t}$ production cross section in the dilepton channel in pp collisions at $\sqrt{s} = 7$ TeV", *JHEP* **1211** (2012) 067, arXiv:1208.2671.

[155] CMS Collaboration, "Measurement of the Inclusive W and Z Production Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV", *JHEP* **1110** (2011) 132, arXiv:1107.4789.

[156] M. V. Garzelli, A. Kardos, C. G. Papadopoulos, and Z. Trocsanyi, "$t\bar{t}W^{\pm}$ and $t\bar{t}Z$ Hadroproduction at NLO accuracy in QCD with Parton Shower and Hadronization effects", *JHEP* **1211** (2012) 056, arXiv:1208.2665.

[157] J.M. Campbell and R.K. Ellis, "$t\bar{t}W^{\pm}$ production and decay at NLO", *JHEP* **1207** (2012) 052, arXiv:1204.5678.

[158] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, *et al.*, "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP* **1407** (2014) 079, arXiv:1405.0301.

[159] CMS Collaboration, "Single Muon Efficiencies in 2012 Data", CMS Performance Note CMS-DP-2013-009, CERN, Geneva, 2013.

[160] CMS Collaboration, "Tau-ID: Recommendation from the Tau POG ", https://twiki.cern.ch/twiki/bin/view/CMS/TauIDRecommendation?rev=15. Accessed: March 26, 2015 (CMS internal content).

[161] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, Pavel M. Nadolsky, *et al.*, "New generation of parton distributions with uncertainties from global QCD analysis", *JHEP* **0207** (2002) 012, arXiv:hep-ph/0201195.

[162] M. Botje, J. Butterworth, A. Cooper-Sarkar, A. de Roeck, J. Feltesse, *et al.*, "The PDF4LHC Working Group Interim Recommendations", arXiv:1101.0538.

[163] H. L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, *et al.*, "New parton distributions for collider physics", *Phys. Rev.* **D82** (2010) 074024, arXiv:1007.2241.

[164] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, "Parton distributions for the LHC", *Eur. Phys. J.* **C63** (2009) 189, arXiv:0901.0002.

[165] R. D. Ball, L. Del Debbio, S. Forte, A. Guffanti, J. I. Latorre, *et al.*, "A first unbiased global NLO determination of parton distributions and their uncertainties", *Nucl. Phys.* **B838** (2010) 136, arXiv:1002.4407.

[166] T. Junk, "Confidence level computation for combining searches with small statistics", *Nucl. Instrum. Meth.* **A434** (1999) , arXiv:hep-ex/9902006.

[167] A. L. Read, "Presentation of search results: The $CL_S$ technique", *J. Phys.* **G28** (2002) 2693.

[168] ATLAS and CMS Collaborations and the LHC Higgs Combination Group, "Procedure for the LHC Higgs boson search combination in Summer 2011", CMS Note CMS-NOTE-2011-005, CERN, Geneva, 2011.

[169] CMS Collaboration, "Search for third-generation scalar leptoquarks in the t$\tau$ channel in proton-proton collisions at $\sqrt{s} = 8$ TeV", arXiv:1503.09049, CMS-EXO-14-008.

[170] CMS Collaboration, "Global Tags for Conditions Data", https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideFrontierConditions?rev=403. Accessed: January 17, 2015.

[171] CMS Collaboration, "CMS Software", http://cms-sw.github.io/. Accessed: January 17, 2015.

[172] CMS Collaboration, "CMS Technical Design Report for the Phase 1 Upgrade of the Hadron Calorimeter", CERN-LHCC-2012-015, CERN, Geneva, 2012.

[173] 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* **1006** (2010) 043, arXiv:1002.2581.

[174] S. Alioli, P. Nason, C. Oleari, and E. Re, "NLO single-top production matched with shower in POWHEG: s- and t-channel contributions", *JHEP* **0909** (2009) 111, arXiv:0907.4076.

[175] E. Re, "Single-top Wt-channel production matched with parton showers using the POWHEG method", *Eur. Phys. J.* **C71** (2011) 1547, arXiv:1009.2450.

[176] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, "MadGraph 5 : Going Beyond", *JHEP* **1106** (2011) 128, arXiv:1106.0522.

[177] T. Sjöstrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual", *JHEP* **0605** (2006) 026, arXiv:hep-ph/0603175.

[178] Z. Was, "TAUOLA the library for $\tau$ lepton decay, and KKMC/KORALB/KORALZ/... status report", *Nucl. Phys. Proc. Suppl.* **98** (2001) 96, arXiv:hep-ph/0011305.

[179] J. Alwall, S. Hoche, F. Krauss, N. Lavesson, L. Lonnblad, *et al.*, "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions", *Eur. Phys. J.* **C53** (2008) 473, arXiv:0706.2569.

[180] GEANT4 Collaboration, "GEANT4: A Simulation toolkit", *Nucl. Instrum. Meth.* **A506** (2003) 250.

[181] Y. Li and F. Petriello, "Combining QCD and electroweak corrections to dilepton production in FEWZ", *Phys. Rev.* **D86** (2012) 094034, arXiv:1208.5967.

[182] M. Czakon, P. Fiedler, and A. Mitov, "Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$", *Phys. Rev. Lett.* **110** (2013) 252004, arXiv:1303.6254.

[183] O. Brein, R. V. Harlander, and T. J. E. Zirke, "vh@nnlo - Higgs Strahlung at hadron colliders", *Comput. Phys. Commun.* **184** (2013) 998, arXiv:1210.5347.

[184] N. Kidonakis and R. Vogt, "The Theoretical top quark cross section at the Tevatron and the LHC", *Phys. Rev.* **D78** (2008) 074005, arXiv:0805.3844.

[185] J. M. Campbell, R. K. Ellis, and C. Williams, "Vector boson pair production at the LHC", *JHEP* **1107** (2011) 018, arXiv:1105.0020.

[186] R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, R. Pittau, *et al.*, "Scalar and pseudoscalar Higgs production in association with a top-antitop pair", *Phys. Lett.* **B701** (2011) 427, arXiv:1104.5613.

[187] M. V. Garzelli, A. Kardos, C. G. Papadopoulos, and Z. Trocsanyi, "Standard Model Higgs boson production in association with a top anti-top pair at NLO with parton showering", *Europhys. Lett.* **96** (2011) 11001, arXiv:1108.0387.

[188] CMS Collaboration, "Single Muon Efficiencies in 2012 Data", CMS Detector Performance Summary CMS-DP-2013-009, CERN, Geneva, 2013.

[189] CMS Collaboration, "2012 Pileup Information", https://twiki.cern.ch/twiki/bin/viewauth/CMS/PdmVPileUpDescription?rev=22. Accessed: February 26, 2015 (CMS internal content).