

Top quark mass measurement and color effects at the LHC

Dissertation zur Erlangung des Doktorgrades
an der Fakultät für
Mathematik, Informatik und Naturwissenschaften
Fachbereich Physik
der Universität Hamburg

vorgelegt von

Nataliia Kovalchuk

aus Kiew, Ukraine

Hamburg

2018



Gutachter der Dissertation:	Prof. Dr. Peter Schleper Prof. Dr. Elisabetta Gallo
Gutachter der Disputation:	Prof. Dr. Peter Schleper Prof. Dr. Elisabetta Gallo Prof. Dr. Robin Santra Prof. Dr. Gudrid Moortgat-Pick Dr. Christian Sander
Datum der Disputation:	29.03.2018
Vorsitzender des Prüfungsausschusses:	Prof. Dr. Robin Santra
Vorsitzender des Promotionsausschusses:	Prof. Dr. Wolfgang Hansen
Leiter des Fachbereichs Physik:	Prof. Dr. Michael Potthoff
Dekan der Fakultät für Mathematik, Informatik und Naturwissenschaften:	Prof. Dr. Heinrich Graener

Abstract

The top quark, the heaviest fundamental particle discovered to date, is one of the most peculiar particles that were discovered so far. It plays a crucial role in consistency checks of the Standard Model and in searches for new physics, e.g., supersymmetry, composite Higgs, and many other exotic models. In this thesis, an important property of the top quark is measured: the mass. This analysis is based on the data recorded at a center-of-mass energy of 13 TeV in 2016 with the CMS detector at the CERN LHC, and corresponds to an integrated luminosity of 35.9 fb^{-1} . The mass of the top quark is measured using the top quark pair event candidate, which corresponds to events with one muon or electron and at least four jets. The corresponding decay products are used in a kinematic fit to perform the jet quark assignment, increase the fraction of correctly reconstructed top quarks and to improve the mass resolution. Using the ideogram method the top quark mass is measured simultaneously with the jet scale factor (JSF), constrained by the jets arising from the W boson decay. The estimated result is calibrated with samples simulated with a next-to-leading order matrix element generator matched to the parton shower. The top quark mass is measured to be $m_t = 172.25 \pm 0.08 \text{ (stat+JSF)} \pm 0.62 \text{ (syst) GeV}$. The results are tested for possible kinematic dependence by performing measurements of the top quark mass in different phase space regions.

The residual data-to-simulation calibration of the energy of the jets is also estimated from dijet events with data collected at center-of-mass energy of 13 TeV in 2015 with the CMS detector corresponding to an integrated luminosity of 2.1 fb^{-1} . The corrections are performed using selected back-to-back dijet events by the MPF and dijet balance methods and are found to differ from unity by less than 3% in the barrel region and up to 17% in the endcap and forward regions of the detector. This result was used in the top mass measurement.

Additionally, a tune for a possible color reconnection between the different multiple parton interaction systems is performed for the recently available QCD-inspired model. The predictions obtained by this tune are able to describe underlying event observables as well as measured charged particle multiplicities. The results are used for a more reliable estimation of the corresponding systematic uncertainty.

Zusammenfassung

Das Top-Quark, das schwerste bekannte Elementarteilchen, ist eines der eindrucksvollsten Teilchen, das bislang entdeckt wurde. Es spielt eine entscheidende Rolle für Konsistenztests innerhalb des Standardmodells und für Suchen nach neuer Physik, wie z.B. Supersymmetrie, Composite Higgs- und anderen exotischen Modellen. In dieser Arbeit wird eine wichtige Eigenschaft des Top-Quarks gemessen: die Masse. Die Analyse basiert auf dem Datensatz, der bei einer Schwerpunktsenergie von 13 TeV im Jahr 2016 mit dem CMS Detektor am LHC aufgezeichnet wurde und einer integrierten Luminosität von 35.9 fb^{-1} entspricht. Die Masse des Top-Quarks wird in Ereignissen mit einem Myon oder einem Elektron und mindestens vier Jets gemessen. Die entsprechenden Zerfallsprodukte werden in einem kinematischen Fit verwendet, um Jets zu den Quarks zuzuordnen und um den Anteil von korrekt rekonstruierten Top-Quarks zu erhöhen. Die Masse des Top-Quarks wird gleichzeitig mit den Jet-Skalenfaktoren (JSF) mit Hilfe der Ideogramm-Methode gemessen, indem Bedingungen an die aus dem Zerfall des W-Bosons entstandenen Jets gestellt werden. Die Messung wird mit Datensätzen kalibriert, die mit einem Generator simuliert sind, der ein Matrix-Element in nächst-führender Ordnung mit dem Parton-Schauer kombiniert. Die gemessene Masse des Top-Quarks beträgt $172.25 \pm 0.08 \text{ (stat+JSF)} \pm 0.62 \text{ (syst) GeV}$. Mögliche kinematische Abhängigkeiten werden überprüft, indem differentielle Messungen der Masse des Top-Quarks in verschiedenen Phasenräumen durchgeführt werden.

Kalibrierungen der Jetenergie in Daten und Simulation werden unter Auswertung eines Datensatzes bestimmt, der bei einer Schwerpunktsenergie von 13 TeV im Jahr 2015 mit dem CMS Detektor aufgezeichnet wurde und einer integrierten Luminosität von 2.1 fb^{-1} entspricht. Die Korrekturfaktoren werden in Ereignissen mit zwei selektierten Jets und mit Hilfe der MPF- und Zweijet-Balance-Methoden berechnet. Im zentralen Detektorbereich weichen die Korrekturfaktoren um weniger als 3% und im äußeren Detektorbereich um maximal 17% von Eins ab.

Eine Anpassung für eine mögliche Farb-Wiederverbindung zwischen verschiedenen Parton-Interaktionssystemen wird auf Basis eines seit kurzem verfügbaren QCD-inspirierten Modells durchgeführt. Die erhaltenen Vorhersagen der Anpassung ermöglichen eine Beschreibung von zugrunde liegenden Ereignisobservablen sowie gemessener geladener Teilchenmultiplizitäten. Mit diesen Ergebnissen wird eine zuverlässige Abschätzung der entsprechenden systematischen Unsicherheit ermöglicht.

Contents

1. Introduction	1
2. Theoretical overview	5
2.1. The Standard Model of elementary particles	5
2.1.1. Gauge symmetries	7
2.1.2. Electromagnetic interaction	8
2.1.3. Weak interaction	9
2.1.4. Electroweak unification	11
2.1.5. Strong interaction	13
2.1.6. The Higgs mechanism	14
2.1.7. Perturbation theory and renormalization	18
2.1.8. Precision test of the Standard Model	20
2.1.9. Beyond the Standard Model	20
2.2. Top quark in the standard model	21
2.2.1. Top quark production and decay	23
2.2.2. Top quark mass definitions	26
3. Experimental setup	29
3.1. The Large Hadron Collider	29
3.2. Description of the CMS experiment	32
3.2.1. Tracking system	35
3.2.2. Electromagnetic calorimeter	37
3.2.3. Hadronic calorimeter	38
3.2.4. Solenoid magnet	40
3.2.5. Muon system	42
3.2.6. Trigger system	44
4. Simulation of hadronic collision	47
4.1. Proton structure	48
4.2. Hard scattering and matrix element	50
4.3. Parton shower	51
4.4. Matrix element - parton shower matching	52
4.5. Hadronization and decay	52
4.6. Underlying event and pileup	54

4.7. Event generators	55
5. Tuning of the color reconnection model in Pythia8	59
5.1. Color reconnection models in Pythia 8	60
5.2. Underlying event measurements	62
5.3. Tuning strategy	63
5.4. Results and tune validation	67
5.5. Summary	73
6. Object reconstruction and identification	75
6.1. The CMS Particle Flow algorithm	75
6.2. Primary vertex reconstruction	78
6.3. Muon selection	79
6.4. Electron selection	82
6.5. Jet reconstruction and calibration	85
6.6. Missing transverse momentum	98
7. Determination of residual jet energy correction	101
7.1. Dijet event topology	102
7.2. Jet response determination	103
7.2.1. The dijet p_T balance method	103
7.2.2. Missing transverse energy projection fraction method	104
7.3. Dijet event selection and reconstruction	105
7.3.1. Data and simulation samples	105
7.3.2. Trigger efficiency study	106
7.3.3. Dijet event selection	108
7.4. Event reweighting in simulation	110
7.5. L2 residual correction estimation	111
7.5.1. Radiation correction of the methods	112
7.5.2. Relative L2 residual determination	114
7.5.3. L2 residual correction in the forward detector region	115
7.5.4. Validation of the pileup simulation	117
7.6. L2 residual correction uncertainties	118
7.7. Results	122
7.8. Summary	122
8. Top quark mass measurement	125
8.1. Measurement strategy	126
8.2. Data samples	127
8.3. Lepton+jets event selection and reconstruction	133
8.3.1. Online selection	133
8.3.2. Preselection	133

8.3.3.	Baseline selection	137
8.3.4.	Kinematic reconstruction of the top quark mass	142
8.3.5.	Final selection	146
8.4.	Top quark mass extraction	148
8.4.1.	Ideogram method	149
8.4.2.	Method implementation	151
8.4.3.	Method and statistical uncertainty validation	154
8.5.	Systematic uncertainties	161
8.5.1.	Experimental uncertainties	162
8.5.2.	Modeling uncertainties	164
8.5.3.	Discussion of the systematic uncertainties	168
8.6.	Results	173
8.7.	Comparison with other measurements	176
8.8.	Theoretical interpretation of the result	179
8.9.	Summary and outlook	180
9.	Measured top quark mass as a function of kinematic observables	181
9.1.	Methodology	181
9.2.	Method calibration	182
9.3.	Method implementation	182
9.4.	Results	183
9.5.	Summary	190
10.	Conclusions	191
A.	Appendix	195
A.1.	Technicalities	195
A.2.	L2 Residual jet energy correction	197
A.3.	Data-MC comparison for the top quark mass measurement	204
A.4.	Permutation yields	210
A.5.	Templates for the μ/e +jets channel	211
A.6.	Details on systematic uncertainties	225
A.7.	Color reconnection	237
A.8.	Measured top quark mass as a function of kinematic observables . . .	239
	Bibliography	249

1. Introduction

Some of the fundamental physics questions about the structure of the Universe are still waiting to be answered. The field of particle physics aims to probe and to describe the most fundamental objects and the interactions between them. Great progress has been made in this field since 1970, when the most precise and successful theory of elementary particles was developed. This theory is referred to as the Standard Model (SM) of particle physics, and its validity has been tested in numerous measurements. The Higgs boson, the last missing piece of the SM, was discovered in 2012 at the Large Hadron Collider (LHC) [1]. Its discovery was the last milestone in a long series of measurements of the SM.

A particular role in the SM is played by the top quark, the heaviest elementary particle known, discovered in 1995 at the Tevatron [2, 3]. The large mass of the top quark m_t causes its extremely short lifetime ($\approx 5 \cdot 10^{-25}$ s), which inhibits the formation of bound states and prevents the creation of top-flavored hadrons. These properties provide unique possibilities for studying the top quark by reconstructing its decay products, in particular, its mass m_t . The top quark mass is a fundamental parameter in Quantum Chromodynamics (QCD) and affects the theoretical predictions of particle production cross sections. A precise top quark mass measurement is necessary to explore the Higgs boson properties as well as to search for physics beyond the SM. Due to the large value of the top quark mass, it influences the Higgs and massive gauge bosons masses through loop corrections, and on the extrapolation of the SM to high-energy scales. Moreover, since the Higgs boson mass has been measured very precisely, the importance of an accurate m_t estimate has become crucial in the study of the stability of the electroweak vacuum.

Previous measurements of m_t have been performed employing different methods and using different $t\bar{t}$ decay channels, e.g., all-jet, lepton+jet and dilepton, where jets are the experimental signatures of quarks and gluons. The first direct measurements of the top quark mass were performed at the Tevatron at center-of-mass energy of 1.96 TeV by the D0 and CDF collaborations. The latest results from the D0 collaboration [4] in the lepton+jets $t\bar{t}$ decay channel yields $m_t = 174.98 \pm 0.76$ GeV. The latest m_t measurement from CDF collaboration in lepton+jet decay channel yields $m_t = 172.85 \pm 1.11$ GeV [5]. The most precise m_t measurement was performed at the LHC at center-of-mass energy of 7 and 8 TeV by the CMS collaboration [6] in the lepton+jets channel, $m_t = 172.35 \pm 0.51$ [7]. A measurement in the same channel from the ATLAS collaboration yields $m_t = 172.08 \pm 0.91(\text{syst})$ GeV [8]. The

world combination is $m_t = 173.34 \pm 0.52 \text{ GeV}$ and combines $t\bar{t}$ events in $p\bar{p}$ and pp collisions [9].

This thesis presents the first top quark mass measurement using the full data sample collected in 2016 with the CMS detector at a center-of-mass energy of 13 TeV [10], corresponding to an integrated luminosity of $35.9 \pm 0.9 \text{ fb}^{-1}$ [11]. The measurement is performed using the ideogram method [12]. With respect to the measurement in [7], new Monte Carlo (MC) event generators have been incorporated and a larger amount of data has been collected at a higher center-of-mass energy of 13 TeV. The MC samples are based on new next-to-leading order matrix elements and new parton shower implementation. Together with the newly available studies of the underlying event and color reconnection, it is possible to achieve better signal and background modeling, and thereby reduce the theoretical uncertainties in the measurement of m_t .

The analysis is performed in the lepton+jets final state where one hadronically decaying W boson from a top quark and another W boson decaying into a charged lepton and the corresponding neutrino, are expected from the signal process. This lepton+jets decay channel corresponds to a large branching fraction and the presence of a charged lepton reduces the background from multijet processes. The kinematic information from the decay products of the hadronically decaying top quark are used to reconstruct m_t . In order to increase the fraction of the correct reconstructed top quarks, a kinematic fit of the decay products with the $t\bar{t}$ hypothesis is applied, but without assuming a value for m_t . The top quark mass is measured simultaneously with the jet energy scale factor (JSF). The JSF is used to reduce one of the dominant sources of systematic uncertainty, stemming from the standard CMS jet energy corrections [13]. In order to account for the possible biases in the measurement, m_t is additionally studied in bins of kinematic observables. The presented measurement was released by the CMS collaboration [10] as a preliminary public result.

In particular the top quark mass measurement, but also numerous other measurements and searches for potential new physics rely on precise measurements of quark and gluon jet properties. In the CMS collaboration the jet energy is calibrated using a sequential chain. Within the scope of this thesis, residual differences in the jet response between data and simulation are measured using proton-proton collision data recorded with the CMS detector in 2015, corresponding to an integrated luminosity of 2.1 fb^{-1} . The correction scale factors are determined as a function of the pseudorapidity using a data-driven method and dijet event topologies. The obtained data-to-simulation residual jet energy corrections are used in all publications of the CMS collaboration.

The m_t measurement as well as other precise measurements of the SM are significantly affected by the modeling of the systematic uncertainty, in particular possible reconnections of the color flow. A better understanding of the color reconnection of the partons belonging to different multiple parton interactions occurring within

the same collision has been gained just recently, and is summarized in [14, 15]. In the context of this thesis, a tune based on the QCD-inspired color reconnection model [14] is developed for the PYTHIA 8.219 [16, 17] MC event generator. It is derived using the underlying event data measured at a center-of-mass energy of 7 TeV and 13 TeV with the CMS and ATLAS detectors. The obtained parameters for the QCD-inspired color reconnection tune are available as supplementary material in [10]. The MC samples with the corresponding tune are produced and used for the investigation of the color reconnection systematic uncertainty.

The thesis is organized as follows. In Chapter 2 the basics of the SM of particle physics are discussed, with an emphasis on the top quark properties. An overview of the LHC and the CMS detector is given in Chapter 3. Chapter 4 describes the simulation of the hadronic collisions by the MC event generators. Chapter 5 presents the color reconnection tune. Chapter 6 describes the reconstruction and identification of the physics objects. The determination of residual jet energy corrections is given in Chapter 7. The selection, methodology, results, and the corresponding uncertainties of the top quark mass measurement are discussed in Chapter 8. A study of the top quark mass as a function of various kinematic observables is presented in Chapter 9. Finally, a summary is given in Chapter 10.

2. Theoretical overview

Particle physics explores nature at small distances and describes the properties of the fundamental particles and their interactions. It successfully accumulates the knowledge about the matter constituents and their properties over almost one century. Nowadays, at the time the most powerful and complex experimental facilities ever built are utilized, many studies, like those described in this thesis, aim to expand the understanding of the laws of the nature. In the followings a brief theoretical overview of particle physics relevant to this study is presented.

2.1. The Standard Model of elementary particles

The Standard Model of particle physics (SM) [18] is the most elaborated and successful theory of fundamental physics. It comprises the matter particles and their interactions. The matter particles, i.e., fermions (spin = 1/2), consist of six quark and six leptons, while their interactions are carried through several force fields mediated via gauge vector bosons (spin = 1), which are quanta of the gauge fields. The only one spin-zero particle in the SM is the scalar Higgs boson, which is responsible for the mass of the other particles. The different generations of quarks and leptons can be referred to as “flavor”. An overview of the content of all fundamental particles of the SM is shown in Fig. 2.1.

The quarks and leptons can be grouped into three generations, each containing a pair, according to the masses and charges:

- *up-type quarks*: up (u), charm (c) and top (t) quarks $Q = +\frac{2}{3}$,
- *down-type quarks*: down (d), strange (s) and bottom (b) quarks with $Q = -\frac{1}{3}$,
- *leptons*: electron (e), muon (μ) and tau (τ) with $Q = -1$,
- *neutrinos*: electron- (ν_e), muon- (ν_μ) and tau- (ν_τ) neutrinos with $Q = 0$,

where Q is the electrical charge. Each matter particle has its antipartner, i.e., *antiparticle*, which has identical properties besides its charges, thus $Q_{\text{antiparticle}} = -1 \cdot Q_{\text{particle}}$. However, this statement should be verified for neutrinos, since they might be *Majorana particles*¹. Each fermion has two degrees of freedom, i.e., with spin up or down, or can be either left or right handed. In addition, each quark

¹The Majorana particles is a fermion that is its own antiparticle.

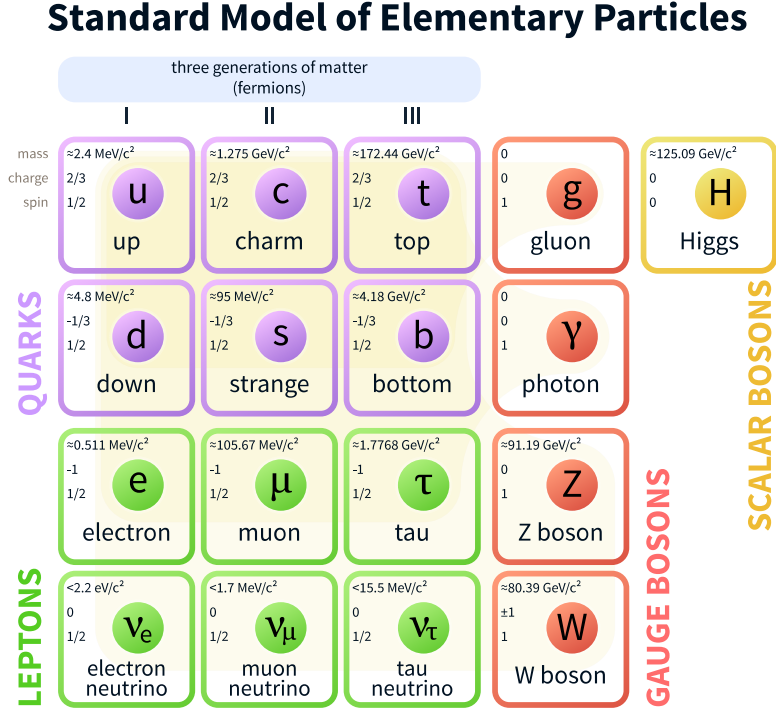


Figure 2.1.: The particle of the Standard Model with three generation of fermions, gauge bosons and the Higgs boson. In the corners of the boxes a mass (or limit for neutrinos), an electric charge and a spin are given [19].

carries one of three *color charges*, namely *red* (r), *green* (g), or *blue* (b), while each antiquark has its anticolor, namely *antired* ($\bar{r}$), *antigreen* ($\bar{g}$), or *antiblue* ($\bar{b}$). The first generation of fermions build up the ordinary matter, i.e., proton² ($p = uud$) and neutron ($n = udd$) and together with the electron are the building blocks of the atoms - components of the visible matter.

The SM distinguishes the following bosons

- 8 vector (spin = 1) gluons (g),
- 4 vector (spin = 1) electroweak bosons: photon (γ), and Z^0 , $W^\pm$ bosons,
- 1 scalar (spin = 0) Higgs boson,

where only $W^\pm$ have non-zero electrical charge, $Q = \pm 1$. The gauge bosons are the gluons and the electroweak bosons, where Z^0 and $W^\pm$ bosons are *massive* and have three degrees of freedom (polarizations), while gluons and photons are *massless* with two degrees of freedom (polarizations). In addition, gluons carry color charges and can be distinguished as a color octet or color singlet states, i.e., $3 \oplus 3 = 8 \oplus 1$ with the neutral singlet $(r\bar{r} + g\bar{g} + b\bar{b})/\sqrt{3}$. The gauge bosons are used as mediators of

²At the moment protons are the only stable baryons known.

the three fundamental interaction. Thus the *strong* interaction between quarks and gluons is mediated by gluons, *electromagnetic* interaction between charged particles is mediated by photons, and *weak* charged-current and neutral-current interactions are mediated by the heavy $W^\pm$ and Z^0 bosons, respectively. Gravity (General Relativity) is not included in the SM, since gravity is still non-renormalizable and the renormalizable theory of quantum gravity has still to be developed.

In order to finalize the description of the elements in the SM, the Higgs boson is introduced. It was the final piece added to the SM and discovered at the Large Hadron Collider (LHC) by the ATLAS and CMS collaborations in 2012 [1, 20]. The Higgs boson is a spin-zero particle, that is introduced by the Higgs mechanism. More detail on the interaction forces and the Higgs mechanism is given in the followings.

2.1.1. Gauge symmetries

The building blocks of the SM are the quantum fields introduced by the Quantum Field Theory (QFT), which describe the particles and their interactions. Within this formalism, fermions are represented by quantised spinor fields (Ψ) and gauge bosons by quantised vector fields (F). In QFT, each fermion and interaction is described by a term in a scalar function of the corresponding field, i.e., the Lagrangian density³ ($\mathcal{L}$). Thus, the interactions in the SM can be described, if the solutions for $\mathcal{L}$ are found to be invariant under local gauge transformations⁴ U of the fermion spinors Ψ

$$\Psi \rightarrow \Psi' = U\Psi. \quad (2.1)$$

Each of three types of interactions described in the SM obeys a set of rules, which are implemented by the underlying non-abelian Lie-group⁵ $SU(N)$ ¹ in the gauge formalism of the Yang-Mills theory [21]. In general, symmetries play a very important role in the Quantum Field Theory. The two major types of symmetries that can be distinguished in the SM are *local* and *global*. The group of local symmetries denotes the invariance of the Lagrangian and observables with respect to transformations of coordinates and (or) fields if the transformations act differently at a separate point in space-time. In global symmetries the transformations act in the same way at every point. An important result is provided by Noether's theorem, which states that if action is invariant with respect to some group transformation symmetry, this give rise to conserved quantities (constants of motion). An example can be the equation of motion for a free particle that does not interact with other particles and thus it is independent on its space-time coordinates. Thus, the energy and momentum of this particle are the conserved quantities.

³The Lagrangian density will be referred to as the Lagrangian.

⁴The unitary local transformations U .

⁵The Lie-groups represent the theory of continuous symmetry of the objects and structures.

The Standard Model is based on a product of gauge symmetry groups

$$SU(3)_C \times SU(2)_L \times U(1)_Y$$

where each gauge group describes one of three fundamental interaction included in the SM. The dimension the gauge group N is used to calculate the number of the symmetry group generators via $N^2 - 1$. Thus, the strong interaction of quarks and gluons is invariant under the transformations of $SU(3)_C$ group and is mediated via $3^2 - 1 = 8$ gauge bosons, in this case gluons. The $SU(3)_C$ describes Quantum Chromodynamics (QCD), where C stands for the color charge. The electromagnetic (EM) and weak interactions can be unified within the theory of electroweak interaction that is a gauge theory and is invariant under $SU(2)_L \times U(1)_Y$ transformations. The dimension of the $SU(2)_L$ group implies that there are $2^2 - 1 = 3$ gauge bosons mediating the weak interaction, namely $W^\pm$ and Z^0 bosons. The simplest $U(1)_Y$ group has only one gauge boson⁶ mediating the electromagnetic interaction, i.e., γ . The $SU(N)$ group is non-abelian⁷ and reflects the possible self interaction of the gauge bosons. $SU(2)_L \times U(1)_Y$ describes Quantum Electrodynamics (QED), where L stands for the “left-handed”, which indicates that the mediators of the weak interaction ($W^\pm$ and Z^0 bosons) only couple to fermions with left-handed (negative) chiralities, and Y denotes the hypercharge, as discussed in Section 2.1.4.

2.1.2. Electromagnetic interaction

In QED the electromagnetic interaction describes the interaction between electrically charged particles via exchange of a virtual photon⁸. Thus, the electromagnetic field is the field that couples to charged particles by satisfying the conditions of gauge invariance of the fermionic field $\mathcal{L}$ under the unitary transformations of $U(1)_{\text{EM}}$. The symmetric abelian group $U(1)$ provides charge conservation in QED.

The simplest example of electromagnetic interaction is electron-positron scattering. In order to describe this interactions the free Lagrangians for electron and photons should be used, namely

$$\mathcal{L}_0(\Psi) = i\bar{\Psi}\gamma_\mu\partial^\mu\Psi - m\bar{\Psi}\Psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \quad (2.2)$$

that is invariant under the global $U(1)$ transformations

$$\Psi(x) \rightarrow \exp(ie\theta)\Psi(x), \quad \bar{\Psi}(x) \rightarrow \exp(-ie\theta)\bar{\Psi}(x), \quad A_\mu \rightarrow A_\mu, \quad (2.3)$$

where θ is a global constant angle, γ_μ is a Dirac matrix, A_μ is the Dirac spinor of the photon, $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is a electromagnetic field strength.

⁶For $U(1)_Y$ the number of the symmetry group generators is calculated via N^2 .

⁷The group is abelian if all elements of a group commute with each another, $a \cdot b = b \cdot a$.

⁸Since photons is massless and neutral, they do not directly interact with each other.

Moreover, $F_{\mu\nu}$ is invariant under a local transformations: $A_\mu(x) \rightarrow A_\mu(x) + \partial_\mu \chi(x)$. Thus, the global constant θ can be substituted by a local function $\chi(x)$ that can vary from one space-time point to another. Therefore for fermions this transformation looks like:

$$\Psi(x) \rightarrow \exp(ie\chi(x))\Psi(x), \quad \bar{\Psi}(x) \rightarrow \exp(-ie\chi(x))\bar{\Psi}(x). \quad (2.4)$$

In order to get the fermion Lagrangian to be also invariant, the covariant derivative is introduced:

$$\partial_\mu \rightarrow D_\mu, \quad D_\mu \Psi \equiv (\partial_\mu - ieA_\mu)\Psi, \quad D_\mu \bar{\Psi} \equiv (\partial_\mu + ieA_\mu)\bar{\Psi}. \quad (2.5)$$

Afterwards the QED Lagrangian can be constructed:

$$\mathcal{L}_{\text{QED}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\bar{\Psi}\gamma_\mu D^\mu \Psi - m\bar{\Psi}\Psi, \quad (2.6)$$

and can be then rewritten as:

$$\mathcal{L}_{\text{QED}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\bar{\Psi}\gamma_\mu \partial^\mu \Psi - m\bar{\Psi}\Psi + e\bar{\Psi}\gamma_\mu \Psi A^\mu, \quad (2.7)$$

where $\mathcal{L}_{\text{int}} = e\bar{\Psi}\gamma_\mu \Psi A^\mu$ describes the interactions of the electron and positron with photons.

The interaction strength between the charged particles and photons is specified by the electromagnetic coupling⁹ α_{EM} :

$$\alpha_{\text{EM}} = \frac{e^2}{4\pi\hbar c} \approx \frac{1}{137}, \quad (2.8)$$

where e electron charge and its value is correct at the zero momentum transfer. However, in the field of the electric charge, the electron positron pair from the photon becomes polarized, known as vacuum polarization. This effect can change the effective charge, and leads to the change of α . Therefore, interaction strength α depends on the momentum transfer Q^2 , that becomes higher at higher scales, the term of running of $\alpha(Q^2)$ is used. At the scale of the Z boson mass ($Q = m_Z = 91.2 \text{ GeV}$) α is slightly bigger [22]

$$\alpha(m_Z^2) \approx \frac{1}{128.957}. \quad (2.9)$$

2.1.3. Weak interaction

The weak interactions are described by $SU(2)_L$ and involve all types of fermions, and it is mediated via exchange by the electroweak gauge bosons ($W^\pm$ and Z^0 bosons).

⁹The natural units are used, where the speed of light c and reduced Planck constant $\hbar$ are equal to 1.

The large masses of the current gauge bosons $m_{W^\pm} = 80.4 \text{ GeV}$ and $m_{Z^0} = 90.2 \text{ GeV}$ reflect into its short lifetime about 10^{-24} s that naturally restricts the range of the weak interaction that is about 10^{-18} m . The exchange of charged gauge bosons $W^\pm$ allow the transformation of the weak eigenstate partners into each other, thus the flavor mixing of the interacting particles is possible. While the exchange of neutral gauge boson Z^0 does not change the flavor of the particles.

The nuclear β decay is the most prominent effect of the weak interaction. First suggestion to describe β decay $n \rightarrow e^- + p + \bar{\nu}_e$ was done by Enrico Fermi in 1933, that describes interactions in the form of product of two vector current ($V \times V$), that was inspired by EM where electric charge e is replaced by a new coupling constant G_F that is a Fermi coupling constant:

$$\mathcal{L}_{\text{int}} = G_F \bar{\Psi}_p \gamma_\mu \Psi_n \cdot \bar{\Psi}_e \gamma^\mu \Psi_{\nu_e} + h.c. \quad (2.10)$$

Later in 1957 this model was modified. The axial-vector together with vector current was used, in order to get a parity violating interaction. This means that weak interactions can be describes in form of product of two vector minus axial-vector $((V-A) \times (V-A))$:

$$\mathcal{L}_{\text{int}} = G_F \bar{\Psi}_p \gamma_\mu (1 - \gamma^5) \Psi_n \cdot \bar{\Psi}_e \gamma^\mu (1 - \gamma^5) \Psi_{\nu_e} + h.c. \quad (2.11)$$

The Fermi coupling constant [23] can be given as:

$$\frac{G_F}{(\hbar c)^3} = \frac{\sqrt{2}}{8} \frac{g^2}{m_W^2}, \quad (2.12)$$

where g is the coupling constant of the weak interaction.

The special characteristic of the weak interaction is its chirality, i.e., only the left-handed fermions contribute, which leads to another set of modification to describe differential distribution of β decays. The modified form of Fermi Lagrangian includes three generations of fermions:

$$\mathcal{L}_{\text{Fermi}} = G_F (\bar{e}_L, \bar{\mu}_L, \bar{\tau}_L) \gamma_\mu \begin{pmatrix} \nu_{e,L} \\ \nu_{\mu,L} \\ \nu_{\tau,L} \end{pmatrix} \cdot (\bar{u}'_L, \bar{c}'_L, \bar{t}'_L) V_u^\dagger \gamma_\mu V_d \begin{pmatrix} d'_L \\ s'_L \\ b'_L \end{pmatrix} + \dots, \quad (2.13)$$

where quarks q' and q are the eigenstates of strong and weak interactions, respectively. Matrices V_u and V_d describe quark mixing:

$$\begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} = V_d \times \begin{pmatrix} d'_L \\ s'_L \\ b'_L \end{pmatrix}. \quad (2.14)$$

The probability of flavor-changing is given by the amplitudes of the couplings in the

Cabibbo-Kobayashi-Maskawa (CKM) matrix [18]

$$V_u^\dagger V_d \equiv V_{\text{CKM}} = \begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.974 & 0.225 & 0.004 \\ 0.220 & 0.995 & 0.041 \\ 0.008 & 0.040 & 0.999 \end{pmatrix}, \quad (2.15)$$

where V_{CKM} is unitary ($V_i^\dagger V_i = V_i V_i^\dagger = 1$) and the $|V_{i,j}|^2$ corresponds to the probability of a i flavor quark to transit to the j flavor quark. This probability is needed in order to keep the number of quarks to be conserved. The matrix-elements are measured experimentally for each specific weak decays and on the diagonal of the V_{CKM} they are close to unity.

A similar mixing was observed also between neutrinos, their mass eigenstates are not the same as their flavor eigenstates, this leads to the neutrino oscillations. These oscillations happen between different generations, and are a proof of their non-zero mass. Therefore, in the SM the neutrino masses is introduced as an extension.

2.1.4. Electroweak unification

The electromagnetic and weak interactions, where the weak interactions is described by Fermi theory, which is not a gauge theory, need to be unified in order to provide a gauge symmetry. Therefore, by taking the gauge group $SU(2)_L \times U(1)_Y$ both of them are introduced as gauge ones. The $U(1)$ is used to provide intersections between fermions and massless vector field B_μ , while $SU(2)$ transformation requires the introduction of three W_μ^i fields with $i = 1, 2, 3$.

In order to account for parity violation that happens in the weak decay, a different behavior of left- and right-handed components of the fermions under $SU(2)_L$ is assumed:

$$I_j \Psi_L = \frac{1}{2} \tau_j \Psi_L, \equiv I_j \Psi_R = 0, \quad (2.16)$$

where τ_j are 2×2 Pauli matrices and I_j weak isospin. Thus, left-handed fermions are doublets

$$\begin{pmatrix} u \\ d' \end{pmatrix}_L, \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L + 2 \text{ generations}, \quad (2.17)$$

while right-handed singlets

$$u_R, d_R, e_R, (\nu_{e,R}) + 2 \text{ generations}. \quad (2.18)$$

Similarly to the EM case, having the Lagrangian for the fermions

$$\mathcal{L}_{\text{Fermion}}(\Psi) = i\bar{\Psi}\gamma_\mu\partial^\mu\Psi - m\bar{\Psi}\Psi, \quad (2.19)$$

transformations applied. Thus, in case of $SU(2)_L \times U(1)_Y$ the local phase transfor-

mations

$$\begin{pmatrix} \Psi_{L\uparrow} \\ \Psi_{L\downarrow} \end{pmatrix} \rightarrow \begin{pmatrix} \Psi'_{L\uparrow} \\ \Psi'_{L\downarrow} \end{pmatrix} = \exp\left(\frac{1}{2}ig'Y\chi(x)\right) \begin{pmatrix} \Psi_{L\uparrow} \\ \Psi_{L\downarrow} \end{pmatrix} \quad (2.20)$$

$$\Psi_R \rightarrow \Psi'_R = \exp\left(\frac{1}{2}ig'Y\chi(x)\right) \Psi_R, \quad (2.21)$$

and local isospin transformations

$$\begin{pmatrix} \Psi_{L\uparrow} \\ \Psi_{L\downarrow} \end{pmatrix} \rightarrow \begin{pmatrix} \Psi'_{L\uparrow} \\ \Psi'_{L\downarrow} \end{pmatrix} = \exp\left(\frac{g}{2}i\tau^i\lambda_i(x)\right) \begin{pmatrix} \Psi_{L\uparrow} \\ \Psi_{L\downarrow} \end{pmatrix} \quad (2.22)$$

are applied. Where $\chi(x)$ and $\tau^i(x)$ are any scalar functions, g and I_j , and g' and $\frac{Y}{2}$ are the couplings and group generators for $SU(2)_L$ and $U(1)_Y$, respectively.

In order this Lagrangian to be invariant under $SU(2)_L \times U(1)_Y$ transformation, ∂_μ has to be replaced by the covariant derivative similarly to the EM case, but here D_μ is given as:

$$\partial_\mu \rightarrow D_\mu, \quad D_\mu \equiv \partial_\mu + ig\tau^j W_{\mu,j} + ig'\frac{Y}{2}B_\mu. \quad (2.23)$$

Since $SU(2)_L \times U(1)_Y$ group contains both abelian and non-abelian groups simultaneously, to any of $SU(2)_L \times U(1)_Y$ transformation the following transformations have to be applied:

$$B_\mu \rightarrow B'_\mu = B_\mu - \partial_\mu\chi(x), \quad (2.24)$$

$$W_\mu^i \rightarrow W_\mu'^i = W_\mu^i - \partial_\mu\tau^i(x) - ge_{ijk}\lambda^j W_\mu^k, \quad (2.25)$$

where $i, j, k = 1, 2, 3$ and e_{ijk} denotes the antisymmetric epsilon tensor.

Defining the strength tensors for each of these gauge fields, one for the abelian and one for the nonabelian Lie group:

$$B_{\mu\nu} \equiv \partial_\mu B_\nu - \partial_\nu B_\mu, \quad W_{\mu\nu}^i \equiv \partial_\mu W_\nu^i - \partial_\nu W_\mu^i + ge_{ijk}W_\mu^j W_\nu^k. \quad (2.26)$$

The Lagrangian can be written as:

$$\mathcal{L}_{EW} = -\frac{1}{4}W_{\mu\nu}^i W^{i\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}. \quad (2.27)$$

By Gell-Mann-Nishijima Y is related to the Q via its linear combination with weak isospin quantum number I_3 :

$$Y = 2(Q - I_3). \quad (2.28)$$

The weak isospin for the neutrinos and left-handed up-type quarks is equal to $I_3 = \frac{1}{2}$, while leptons and left-handed down-type quarks is equal to $I_3 = -\frac{1}{2}$. The right-handed particles have $I_3 = 0$.

The fields of the gauge bosons of the gauge group $SU(2)_L \times U(1)_Y$ can be given as:

$$W_\mu^\pm = \frac{W_\mu^1 \pm iW_\mu^2}{\sqrt{2}}, \quad W_\mu^0 = W_\mu^3 \quad B_\mu. \quad (2.29)$$

The W_μ^0 and B_μ couples to neutrinos, and thus cannot represent EM field. Instead, they have the same quantum numbers and can be mixed:

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{pmatrix} \times \begin{pmatrix} B_\mu \\ W_\mu^0 \end{pmatrix}, \quad (2.30)$$

where θ_W denotes the weak mixing angle, Z_μ is a term responsible for the neutrino coupling and A_μ for EM field. The weak mixing angle, Weinberg angle, is given by:

$$\cos \theta_W = \frac{g}{\sqrt{g^2 + g'^2}}, \quad \sin \theta_W = \frac{g'}{\sqrt{g^2 + g'^2}} \quad (2.31)$$

and a free parameter in the SM. It is measured experimentally, e.g., $\sin^2 \theta_W(Q) = 0.2397$ at $Q^2 = 0.026 \text{ GeV}^2$ [24].

2.1.5. Strong interaction

The strong interactions between quarks and gluons is described by a non-abelian QCD theory, that is based on the local color transformations from $SU(3)_C$ symmetry group. It is not a chiral theory for massless quarks. The gauge bosons that are mediators in this interaction, are color-charged massless gluons with a field G_μ^a with $a = 1, \dots, 8$ that is similar to the fields W_μ^i . However, the gluon coupling to the quarks and gluon self-coupling limits the effective range of the hard interaction only to 10^{-15} m .

In order to get the QCD Lagrangian, the three component color vector

$$\Psi_C(x) = \begin{pmatrix} \Psi_{\text{red}}(x) \\ \Psi_{\text{green}}(x) \\ \Psi_{\text{blue}}(x) \end{pmatrix} \quad (2.32)$$

has to be used in the Lagrangian for the fermion fields:

$$\mathcal{L}_{\text{Fermion}}(\Psi_C) = i\bar{\Psi}_C \gamma_\mu \partial^\mu \Psi_C - m\bar{\Psi}_C \Psi_C \quad (2.33)$$

to account for the possible color states of the quarks. The color transformations are given by:

$$\Psi_C \rightarrow \Psi'_C = \exp\left(\frac{1}{2}ig_S\lambda^k\beta_k(x)\right)\Psi_C, \quad (2.34)$$

where g_S is a strong coupling, which is a free parameter in SM, $\beta_k(x)$ denotes any scalar functions with $k = 1, \dots, 8$ and λ^k are generators of local color transformations.

In order to get the fermionic Lagrangian to be invariant under $SU(3)_C$ transformation the covariant derivative is used instead of ∂_μ :

$$\partial_\mu \rightarrow D_\mu, \quad D_\mu \equiv \partial_\mu + i \frac{g_S}{2} \lambda^k G_{k\mu}. \quad (2.35)$$

Thus, the QCD Lagrangian is given as:

$$\mathcal{L}_{\text{QCD}} = \sum_C \bar{\Psi}_C^i (i \gamma_\mu D^\mu - m_C)_{ij} \Psi_C^j - \frac{1}{4} F_a^{\mu\nu} F_{\mu\nu}^a, \quad (2.36)$$

and the gluon field strength tensor is given as:

$$F_a^{\mu\nu} = \partial^\mu G_a^\nu - \partial^\nu G_a^\mu + g_S f^{abc} G_b^\mu G_c^\nu, \quad (2.37)$$

where G_a with $a = 1, \dots, 8$ are the gluons fields and f^{abc} is the $SU(3)_C$ structure constant.

The strong interaction is not of infinite reach, since the color charge of quarks changes when gluon is emitted and the effective coupling constant α_S decreases with decreasing of the distance. Thus, α_S is large for low energies and small for high energies as shown in Fig. 2.2, which means that at infinite energy the coupling vanishes and the free quarks can then occur in color charged states, this phenomena is known as *asymptotic freedom*. However, the potential energy between these separated color charged states becomes so large that due to the *color confinement* new quark-antiquark pairs are created. The similar situation happens at the non-perturbation scale above $\Lambda_{\text{QCD}} = 200 \text{ MeV}$ where quarks and gluons condense to the color neutral objects, i.e., hadrons. The composition of quark-antiquark is called meson, of three quarks - baryons. Therefore, due to the color confinement the single quarks and gluons are not observed.

2.1.6. The Higgs mechanism

The electroweak gauge bosons are required to be massless according to the electroweak Lagrangian under $SU(2)_L \otimes U(1)_Y$ transformation. Additionally, this transformation requires fermions of each isospin doublet to have the equal masses. Nevertheless, it is experimentally observed that all electroweak gauge bosons except of photon are massive: $m_{W^\pm} = 80.385 \pm 0.015$ and $m_{Z^0} = 91.1876 \pm 0.0021 \text{ GeV}$ [18]. The fermions mass spectrum comprises the mass of neutrinos of 0.1 eV, electron of 511 keV, u and d quarks of a few MeV, s quark and muon of about 100 MeV, c and b quarks and tau of a few GeV and top quark mass is about $m_t = 173.34 \pm 0.27(\text{stat}) \pm 0.71(\text{syst}) \text{ GeV}$ [9]. In order to account for these masses, an explicit mass term must be added. However, this is forbidden by gauge symmetry conservation. Therefore, in order to generate the corresponding mass the electroweak symmetry must be spontaneously broken, this is done via *spontaneous symmetry breaking*, the Higgs-Englert-Brout

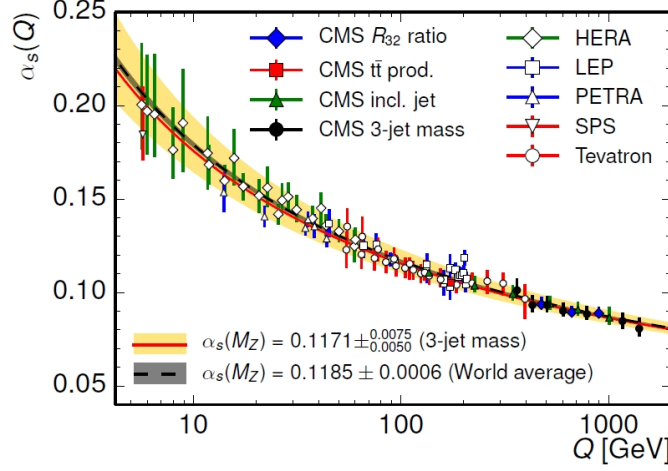


Figure 2.2.: The value of the strong coupling α_s as a function of the energy scale. The error bars on the data points correspond to the total uncertainty [25].

mechanism [26, 27]

According to this mechanism, an additional complex scalar field has to be defined. Thus, having a charged scalar field and a vector field A_μ , their interaction:

$$\mathcal{L} = \partial_\mu \phi^* \partial^\mu \phi - V(\phi) - \frac{1}{4} F_{\mu\nu}^2 + ie(\phi^* \partial_\mu \phi - \partial_\mu \phi^* \phi) A_\mu + e^2 A_\mu A_\nu \phi^* \phi, \quad (2.38)$$

can be considered. Which, in case the potential $V(\phi) \equiv V(\phi^* \phi)$, is invariant under the local $U(1)$ gauge transformation.

$$\phi \rightarrow \exp(ie\chi(x))\phi, \quad \phi^* \rightarrow \exp(-ie\chi(x))\phi^*, \quad A_\mu \rightarrow A_\mu + \partial_\mu \chi(x), \quad (2.39)$$

where $\chi(x)$ is a local function.

Assuming polar coordinates $\phi \equiv \sigma(x)\exp(i\theta(x))$ and $\phi^* \equiv \sigma(x)\exp(-i\theta(x))$ the Lagrangian can be written as:

$$\mathcal{L} = \partial_\mu \sigma \partial^\mu \sigma - V(\phi^* \phi) - \frac{1}{4} F_{\mu\nu}^2 + e^2 \sigma^2 (A_\mu + \frac{1}{e} \partial_\mu \theta)(A_\mu + \frac{1}{e} \partial_\mu \theta), \quad (2.40)$$

where changing $A_\mu + \frac{1}{e} \partial_\mu \theta \rightarrow B_\mu$ can be performed, and thus field strength becomes $F_{\mu\nu}(A) = F_{\mu\nu}(B)$.

It was suggested to take a scalar potential:

$$V(\phi^* \phi) = \mu^2 \phi^* \phi + \lambda (\phi^* \phi)^2, \quad (2.41)$$

that is known as a Higgs potential. It has a mass term and a self-coupling term, and is satisfying the invariance under $SU(2)_L \times U(1)_Y$ transformation.

The values of the factors λ and m are only experimentally determined. The

vacuum stability requires $\lambda > 0$, while depending on the sign of μ^2 two regimes for the vacuum are defined, as shown in Fig. 2.3. They is an unique vacuum state in case $\mu^2 > 0$ (high energy) with minimum of the potential at zero, and degenerate vacuum state in case $\mu^2 < 0$ (low energy) with minimum in a set of points.

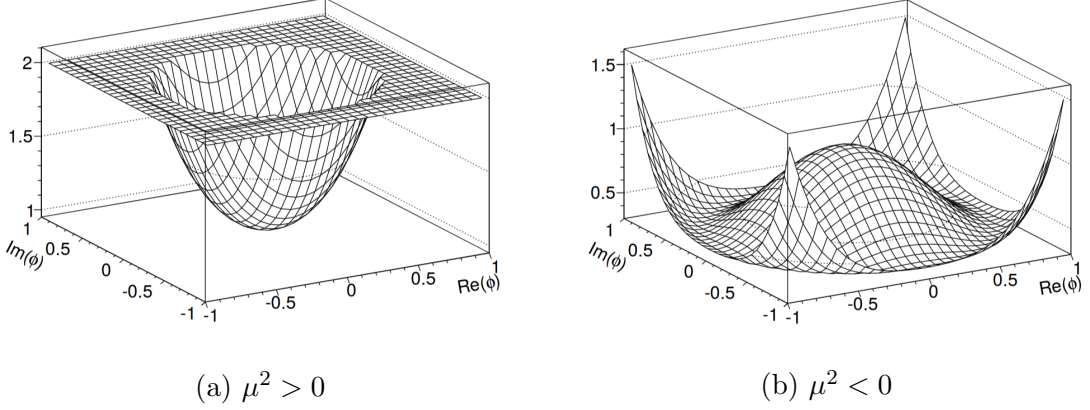


Figure 2.3.: Higgs potential $V(\phi^*\phi)$ for $\lambda > 0$ and different signs of μ^2 [28].

Since the polar coordinates are used, the potential is given as $V(\phi^*\phi) = V(\sigma^2)$, where $\theta(x)$ is the rotational symmetry of the potential. Thus, by taking the derivative of the current potential:

$$\frac{dV(\sigma)}{d\sigma} = 0, \quad (2.42)$$

the two critical points can be found, namely local maximum $\sigma = 0$ and the global minimum $\sigma_0 = \sqrt{-\frac{\mu^2}{2\lambda}}$. By shifting this system to the global minimum:

$$\sigma(x) \rightarrow h(x) + \sigma_0, \quad (2.43)$$

the Lagrangian can be written as:

$$\mathcal{L} = \partial_\mu h \partial^\mu h + e^2 h^2 B_\mu B^\mu + 2e^2 \sigma_0 h B_\mu B^\mu + e^2 \sigma_0^2 B_\mu B^\mu - V(h) - \frac{1}{4} F_{\mu\nu}^2, \quad (2.44)$$

where field B_μ has the mass:

$$m_B^2 = 2e^2 \sigma_0^2 = -\frac{e^2 \mu^2}{\lambda} > 0. \quad (2.45)$$

Thus, μ^2 in the Higgs potential must be negative for symmetry to be broken, which leads to non-zero vacuum expectation value, $v \equiv \sigma_0$, of $\sigma(x)$,

$$v \equiv \langle 0 | \sigma | 0 \rangle = \sqrt{-\frac{\mu^2}{2\lambda}} \neq 0. \quad (2.46)$$

Therefore the rotational symmetry of the Higgs potential $U(1) = O(2)$ is broken spontaneously.

The potential can be now written using the vacuum expectation value:

$$\begin{aligned} V(h) &= \lambda(h+v)^4 + \mu^2(h+v)^2 \\ &= \lambda h^4 + 4\lambda v h^3 + h^2(6\lambda v^2 + \mu^2) + h(4\lambda v^3 + 2\mu^2 v) + \lambda v^4 + \mu^2 v^2, \end{aligned} \quad (2.47)$$

where due to $-2\lambda v^2 = \mu^2$ we can replace $6\lambda v^2 + \mu^2 \rightarrow -2m_h^2$ and $4\lambda v^3 + 2\mu^2 v \rightarrow 0$. Thus, the scalar field h has its own mass term.

On the other hand, the requirement for the complex scalar fields to be gauge invariant assumes an $SU(2)$ isospin doublet with $I_3 = \frac{1}{2}$ and $Y = 1$

$$\Phi = \begin{pmatrix} \Phi^+ \\ \Phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \Phi_1 + i\Phi_2 \\ \Phi_3 + i\Phi_4 \end{pmatrix}, \quad \Phi^\dagger = (\Phi^{+*} \Phi^{0*}). \quad (2.48)$$

The Lagrangian for the Higgs doublet can be written as:

$$\mathcal{L}_{Higgs} = (D_\mu \Phi)^\dagger (D^\mu \Phi) - m^2 \Phi^\dagger \Phi - \lambda (\Phi^\dagger \Phi)^2, \quad (2.49)$$

where $D_\mu \Phi \equiv \partial_\mu \Phi + igW_\mu^a \frac{\tau^a}{2} \Phi + \frac{i}{2} g' B_\mu \Phi$.

Expanding the Higgs field Φ around the vacuum solution in Eq. 2.46 one obtains:

$$\Phi = \exp\left(\frac{i}{2} \tau_j \theta^j\right) \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix}, \quad (2.50)$$

where $j = 1, 2, 3$ and τ_j denotes the Pauli matrices, θ^j and H represents four real scalar fields. The θ^j is a *Nambu-Goldstone* boson [29] with a zero mass, while the Higgs boson described by H is massive with $m_H = \sqrt{2}m$.

The local invariance of the Lagrangian $\mathcal{L}_{EW+H} = \mathcal{L}_{EW} + \mathcal{L}_{Higgs}$ under $SU(2)_L$ allows to rotate fields into states with $\theta^j = 0$ in order to acquire the masses of massive electroweak bosons:

$$m_{W^\pm} = \frac{1}{2}vg, \quad (2.51)$$

$$m_{Z^0} = \frac{m_{W^\pm}}{\cos \theta_W} \quad (2.52)$$

The fermion masses are generated via Yukawa coupling in terms of the form:

$$\mathcal{L}_f = -\frac{g_f v}{\sqrt{2}} (\bar{f}_L f_R + \bar{f}_R f_L) + \frac{g_f h}{\sqrt{2}} (\bar{f}_L f_R + \bar{f}_R f_L), \quad (2.53)$$

where the fermion mass is introduced in the first term, while in the second term its coupling to the Higgs field. Thus, the mass of the leptons and quarks can be given by substitution of the Higgs doublet by its ground state:

$$m_f = \frac{v}{\sqrt{2}} g_f \quad (2.54)$$

2.1.7. Perturbation theory and renormalization

The total Lagrangian of the Standard Model can be defined by using the Lagrangians found for each of the gauge symmetry groups:

$$\mathcal{L}_{SM} = \sum_{\text{flavor}} \bar{\Psi}(i\gamma_\mu D^\mu)\Psi + (D_\mu\Phi)^\dagger(D^\mu\Phi) - V(\Phi) - \frac{1}{4}F_{\mu\nu}^a F_a^{\mu\nu} + \mathcal{L}_{\text{Yukawa}}, \quad (2.55)$$

where the sum runs over all quarks and leptons, $\mathcal{L}_{\text{Yukawa}}$ denotes a mass generating Yukawa term, $F_{\mu\nu}^a$ is a gauge field tensors and D^μ is a covariant derivative:

$$D_\mu \equiv \partial_\mu + i\frac{g}{2}\tau^j W_{j,\mu} + i\frac{g'}{2}B_\mu + i\frac{g_S}{2}\lambda^k G_{k\mu}, \quad (2.56)$$

where $k = 1, \dots, 8$ and $j = 1, 2, 3$. Typically, the behaviour of a physical system can be describe by the equations of motion that can be derived by requiring the action A to be stationary in time t :

$$\delta A = \delta\left(\int dt \mathcal{L}\right) = 0. \quad (2.57)$$

The field equations from the SM Lagrangian take into account the interactions between the objects, therefore the analytical solutions cannot be found. However, it is possible to expand these solutions in perturbation series by orders of the coupling constants. In particular, the transition probability amplitude from an initial state Ψ_i into a final state Ψ_f is given by the matrix element:

$$\mathcal{M} = \langle \Psi_i | S | \Psi_f \rangle = \int d^3x \Psi_f^\dagger S \Psi_i, \quad (2.58)$$

where the amplitude of the scattering process S is defined as:

$$S = \sum_{n=k}^{\infty} \alpha^{n/2} C_n, \quad (2.59)$$

with $\alpha = g^2/4\pi$, where g is the coupling constant of the corresponding interaction and $n = k$ denotes the order, thus the case $n = 2$ represents the first non-trivial order named the “leading order” (LO), $n = 3$ - “next-to-leading order” NLO and so on. According to the Feynman rules [30] this matrix element is expanded in orders of coupling constant:

$$\mathcal{M}(g_S) = \alpha_S^k \mathcal{M}_0 + \alpha_S^{k+1} \mathcal{M}_1 + \alpha_S^{k+2} \mathcal{M}_2 + \dots, \quad (2.60)$$

where the actual probability in this case is given as the square of the given amplitude $\mathcal{M}$. An example for the each term in the expansion of the matrix element is shown in Fig 2.4.

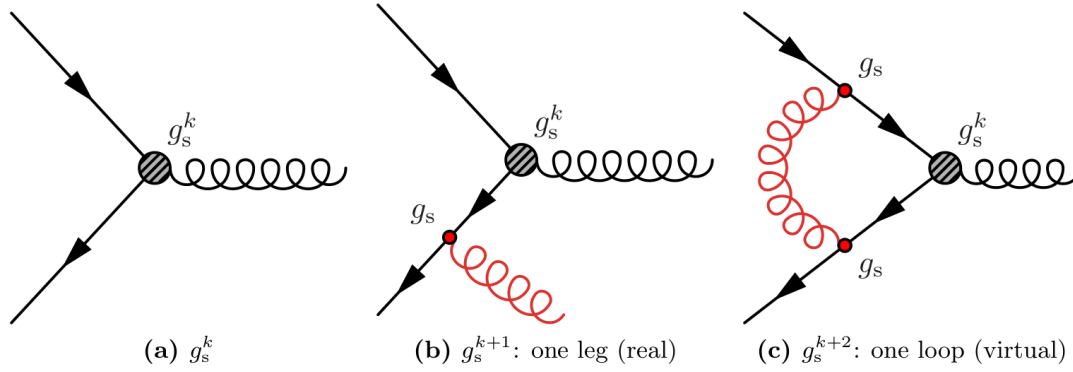


Figure 2.4.: Feynman diagrams [30] contributing to different terms in the expansion of the matrix element [31]. The orders of the strong coupling strength parameter g_s with no additional radiation (a), with radiation of the real gluon (b) and a virtual gluon (c).

Some processes may involve the self-energy loop correction for the gluon propagator. In case the momentum of the particle in the loop approaches infinity, *ultraviolet divergencies* may occur. In order to correct for this divergency the procedures of *regularization* and *renormalization* are utilized. During the regularization step a momentum cut-off Λ_{UV} is introduced that splits divergent integrals into finit and infinit terms. During the renormalization procedure counter-terms are used, which are generated by redefining the “bare” coupling constants and masses by the “effective” ones. The counter-terms replace the dependence of the results from the cut-off parameter to the renormalization scale μ_R [32]. Finally, the defined effective parameters depend on the μ_R and are referred to as running constants. In case all divergent integrals can be described in this way, the theory is called renormalizable.

There are different renormalization schemes for the QCD calculation with corresponding renormalization constant. The most popular one is the modified minimal subtraction $\overline{MS}$ scheme [33,34]. However, it does not have the prescription of choosing μ_R . Nevertheless, the predictions calculated at all orders of α_S do not depend on its choice, since a change of μ_R is compensated by a change of $\alpha_S(\mu_R)$. The expression for the $\alpha_S(\mu_R)$ at 1-loop order is given as:

$$\alpha_S(\mu_R) = \frac{\alpha_S(\mu_{R,0})}{1 + \frac{\alpha_S(\mu_{R,0})}{12\pi} (33 - 2n_f) \ln \frac{\mu_R}{\mu_{R,0}}}, \quad (2.61)$$

where $\mu_R = Q$ denotes the energy scale, n_f corresponds to the number of active quarks at this energy. In case of the SM $n_f = 6$ and $33 - 2n_f > 0$, and thus the coupling of QCD decreases with increasing μ_R . Hence, the property of the asymptotic freedom appears in the limit $\mu_R \rightarrow \infty$:

$$\alpha_S(\mu_R)|_{\mu_R \rightarrow \infty} \rightarrow 0. \quad (2.62)$$

2.1.8. Precision test of the Standard Model

The SM includes a number of precision variables that can be used in order to test the theoretical model. These variables are the decay width and the mass of the massive gauge bosons W and Z, and the top quark mass. By given the ranges to these variables, the mass of the Higgs boson can be constrained as well as the various SM extensions.

The Higgs mass, where its resolution is mainly driven by the top quark mass, is very important. An extremely small mass variation of the Higgs boson has large cosmological consequences [35]. Thus, very accurate measurement of the top quark mass can hold crucial cosmological information. The precise measuring of the top quark mass is the main topic of this thesis and is described in Chapter 8.

2.1.9. Beyond the Standard Model

The SM is a consistent perturbatively renormalizable theory that depends on a finite number of parameters that can be determined by a comparison with the experimental data. However, it can describe only about 4-5 % of the ordinary matter, while Standard Model of Cosmology predicts that there is another 23 % and 72 % of dark matter and dark energy, respectively. The existing of the dark matter was observed and is supposed to some extend to interact gravitationally and flushes the galaxies. Furthermore, this cosmological theory predicts the equal amounts of baryonic matter and antimatter after the Big Bang, however, the absence of antimatter is observed, which is nine order of magnitude less than matter. This asymmetry of matter and antimatter is observed by cosmic rays experiments that report only a small antimatter fraction. In the SM this effect cannot be sufficiently described.

As was already mentioned, the SM described three fundamental forces among four. The Gravity is not described by the SM, since its effect is only visible at high energies around the Planck scale $M_{\text{Pl}} \approx 10^{19} \text{ GeV}$. Moreover, this force is non-renormalizable and cannot be included to the SM.

Additionally, the neutrino in the SM are massless, however the neutrino oscillations were observed, i.e., a change of the neutrino flavor over time, a proof that neutrinos have mass. One more issue is the introduced previously cut-off Λ_{UV} that is used to restrict the loop integral. It can be interpreted as a lower energy scale at which a new physics can be estimated. A natural assumption for Λ_{UV} would be again around the Planck scale M_{Pl} . However the mass of the Higgs boson is around 125 GeV as well as the masses of the gauge bosons that are at the scale of $\sim 100 \text{ GeV}$. This means that in the interval of $\sim 100 \text{ GeV}$ to M_{Pl} no new particles exists, which leads to the *hierarchy problem*.

There are many theories beyond the SM trying to solve the mentioned issues, namely supersymmetry (SUSY), Standard Model of Cosmolog, Grand Unification

theory (GUT)¹⁰ and others.

2.2. Top quark in the standard model

In the SM the Higgs boson couples to each massive fermions via a Yukawa term λ_f and to the massive gauge bosons. These particles affect the mass of the Higgs boson m_H via loop corrections:

$$m_H = m_H^{bare} + \Delta m_H^{ferm.} + \Delta m_H^{gauge} + \Delta m_H^{Higgs} + \dots, \quad (2.63)$$

that are schematically shown in Fig. 2.5.

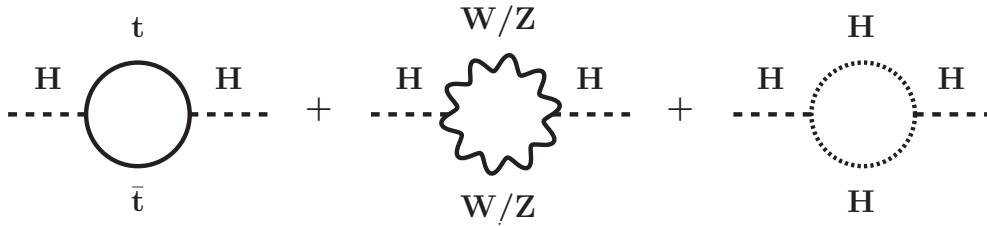


Figure 2.5.: Feynman diagram corresponding to the one-loop corrections to the Higgs bosons mass due to the fermion (top quark), gauge boson and the Higgs boson.

Even though m_H depends on the masses of all other massive particles in the SM, the most important contributions arise from the top quark and W boson, with quadratic dependency on the mass of the top quark m_t and logarithmic dependency on the W boson mass m_W . The dependency of the Higgs mass on the m_t and m_W was used to constrain the Higgs mass before its discovery as shown in Fig. 2.6 (left).

However, the largest correction arises from the top quark, the heaviest fundamental fermion known in the SM with the mass of about 172.5 GeV, while the second most heaviest fermion is the bottom quark that has a mass of about 4.2 GeV. The one-loop radiative corrections affecting the bare mass of the Higgs mass:

$$\Delta m_H^2 = -\frac{|\lambda_t|^2}{8\pi^2} \Lambda_{UV}^2 + \dots, \quad (2.64)$$

where momentum cut-off Λ_{UV}^2 denotes upper (ultraviolet) momentum cut-off that restricts the loop integral as described in Section 2.1.7.

Since the top quark Yukawa coupling is around unity:

$$\lambda_t = \frac{\sqrt{2}m_t}{v} \simeq 1, \quad (2.65)$$

¹⁰The energy at the GUT scale is around $\Lambda_{UV} \approx 10^{16}$ GeV.

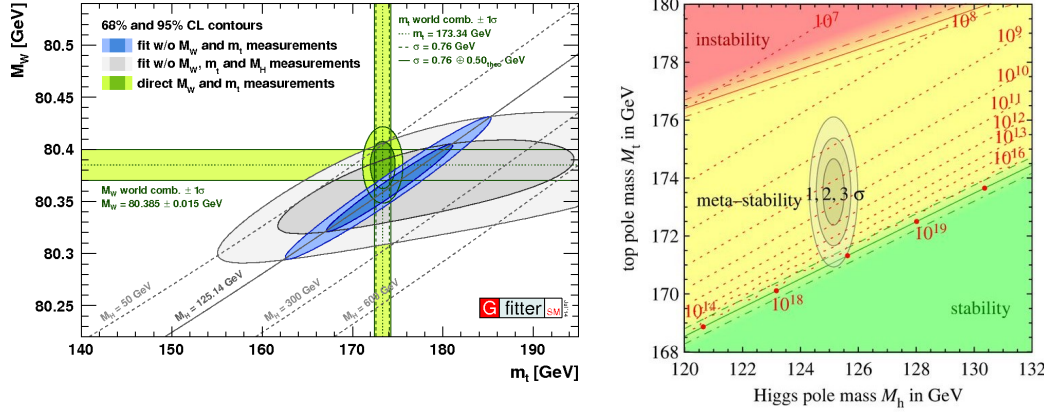


Figure 2.6.: The distribution indirect determination of the m_t with and without inclusion of the Higgs results compared to the electroweak fit in the $m_t - m_W$ plane [36] (left). The vacuum stability domains in the SM for the observed values of m_t and m_H [37] (right).

where the vacuum expectation $v \approx 246$ GeV [18], this means that corrections from the top quark are very big and are of the order of the Λ_{UV} . Hence taking into account that at the energies of the LHC the Higgs boson is produced mainly by gluon-gluon fusion through the top quark loop [38] as shown in Fig. 2.7, the precise measurements of the m_t is a very crucial parameter used for the Higgs boson mass calculation. Moreover, the mass of the top quark and Higgs boson place the SM

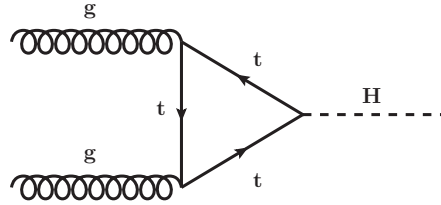


Figure 2.7.: Feynman diagram corresponding to the Higgs boson produced via the gluon fusion process.

vacuum at the border between stability and metastability as shown in Fig. 2.6 (right). Since after the discovery of the Higgs boson, the experimental uncertainty on its mass decreased substantially and the role of the precision of the top quark mass becomes more appropriate to express the stability condition. This is due to the radiative correction from the top quark that modifies the Higgs boson self coupling such that it could achieve negative values at a large energy scale, that could lead to an unstable electroweak vacuum [39,40]. The estimation of this instability scale may require new physics and rely strongly by the precision of m_t and its interpretation in a theoretical framework. Thus, the precise measurement of the top quark mass is necessary to probe the consistency of the SM and to constrain the models of the

new physics models beyond the SM.

The W and Z bosons mass gets also radiative correction from top quark and the Higgs boson, where m_t enters quadratically and m_H logarithmically. The one-loop corrections to their masses are shown in Fig. 2.8.

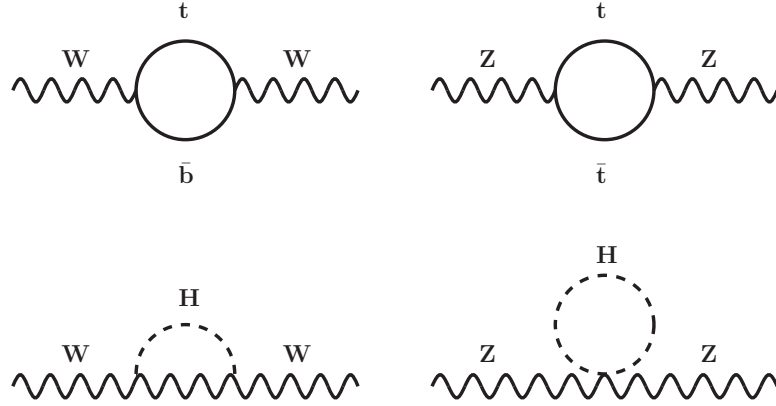


Figure 2.8.: Feynman diagram corresponding to the one-loop corrections to the W and Z bosons masses due to the top quark (upper row) and the Higgs boson (bottom row).

2.2.1. Top quark production and decay

Even before the top quark discovery, predictions for its mass were done with the electroweak fit to the available combined electroweak data. For example, already in 1987 the observation of the $B_d^0 \bar{B}_d^0$ flavor mixing in the process $e^+e^- \rightarrow \Upsilon(4S) \rightarrow B_d^0 \bar{B}_d^0$ [41] was possible, which is mediated via a weak interaction and a virtual top quark loop. The theoretical predictions had very large uncertainties, therefore the top quark was discovered only in 1995 at the Fermilab Tevatron [2, 3].

At LHC a top quark is mainly produced together with its antipartner, a $t\bar{t}$ pair, through the gluon-gluon fusion or quark-antiquark annihilation, which is mediated via strong interaction as shown in Fig. 2.9. At the LHC at the center-of-mass energy of 13 TeV the contribution of the gluon-gluon fusion is about 90%, as discussed in Section 4.1. LHC is also known as a top factory, where at the center-of-mass energy of 8 TeV in 2012 about 5 million and at 13 TeV in 2016 about 30 million of $t\bar{t}$ pairs were produced at each collisions point.

The top quark can be also produced as single top as shown in Fig. 2.10.

The single top quark production is mediated via electroweak interaction and has a production rate three times smaller than $t\bar{t}$, as shown at Tevatron [42] and LHC [43, 44]. Thus, the total production cross-section at center-of-mass energy of 13 TeV

- for $t\bar{t}$ events is about 831.76 pb [45],

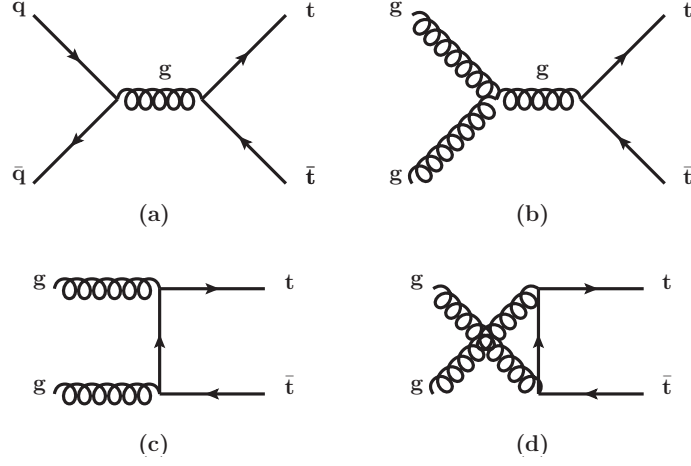


Figure 2.9.: Tree-level diagrams of $t\bar{t}$ production via $q\bar{q}$ annihilation (a) and gg fusion in s-channel (b), t-channel (c) and u-channel (d).

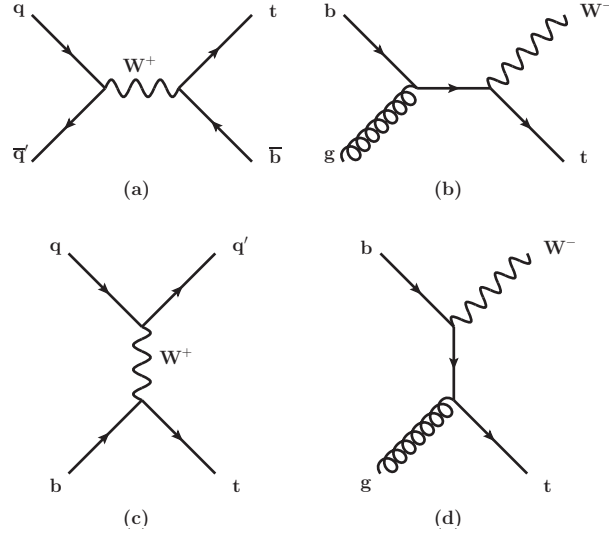


Figure 2.10.: Tree-level diagrams for the electroweak single top quark production in s-channel (a), t-channel (b) and tW-channel (c and d).

- for the single top quark and top antiquark in the t-channel is about 136.02 pb and 80.95 pb [46]
- for the single top quark and top antiquark in the tW-channel is about 35.85 pb and 35.85 pb [47]
- for the single top quark and top antiquark in the s-channel is about 7.20 pb and 4.16 pb [47].

The summary of the measurements of the total $t\bar{t}$ and single top production cross section measured by the CMS collaboration are shown in Fig. 2.11.

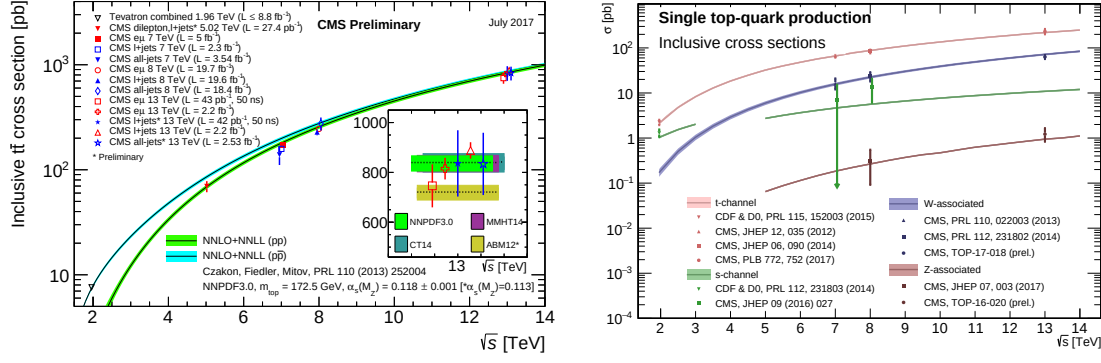


Figure 2.11.: Top quark pair and single top production cross section summary of most precise CMS measurements in comparison with the theory calculation at NNLO+NNLL accuracy [48].

The very large mass and the short life time (ca. $\tau_t = 5 \times 10^{-25}$ s) of the top quark results in its unique property to decay before the hadronization process take place, where the hadronization timescale $\tau_{\text{had}} \approx 1/\Lambda_{QCD} = 3 \times 10^{-24}$ s. The decay width of the top quark depends on its mass and can be calculated via $\Gamma_t \propto m_t^3 = 1/\tau_t$, which is predicted to be $\Gamma_t = 1.33$ GeV for $m_t = 172.5$ [49]. Thus, the large mass of the top quark leads to a large value of its decay width.

Nearly exclusively, the top quark decays into a W boson and a b quark, i.e., $|V_{tb}| > 0.999$, as it can be seen in Eq. 2.15, while other decays, namely into W boson and d or s quark, are highly suppressed in the CKM mechanism and corresponds to $|V_{td}| > 0.008$ and $|V_{ts}| > 0.04$, respectively. The W boson from the top quark can decay either leptonically ($e\nu_e$, $\mu\nu_\mu$ or $\tau\nu_\tau$) with probability 1/9 for each of the decay mode, and hadronically $q\bar{q}'$, namely $u\bar{d}'$ and $c\bar{s}'$, with probability 3/9 for each case. The branching fractions for the $t\bar{t}$ pair is given are Tab. 2.1.

Table 2.1.: Decay channels with corresponding branching fractions for $t\bar{t}$ decays.

		W ⁺ decay			
W ⁻ decay	B ($t\bar{t} \rightarrow b\bar{b} + X$)	$e^+\nu_e$	$\mu^+\nu_\mu$	$\tau^+\nu_\tau$	$q\bar{q}'$
	$e^-\bar{\nu}_e$	1/81	1/81	1/81	6/81
	$\mu^-\bar{\nu}_\mu$	1/81	1/81	1/81	6/81
	$\tau^-\bar{\nu}_\tau$	1/81	1/81	1/81	6/81
	$q\bar{q}'$	6/81	6/81	6/81	36/81

It is possible to distinguished between three type of $t\bar{t}$ decay *all-jets*, *lepton+jets* and *dilepton* decay channels:

All-jets: $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow (q\bar{q}b)(q'\bar{q}'\bar{b})$,

Lepton+jets: $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow (q\bar{q}b)(l\bar{\nu}_l\bar{b})$,

Dilepton: $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow (\bar{l}\nu_l b)(l'\bar{\nu}_{l'}\bar{b})$.

The LO Feynman diagram of these processes are shown in Fig. 2.12.

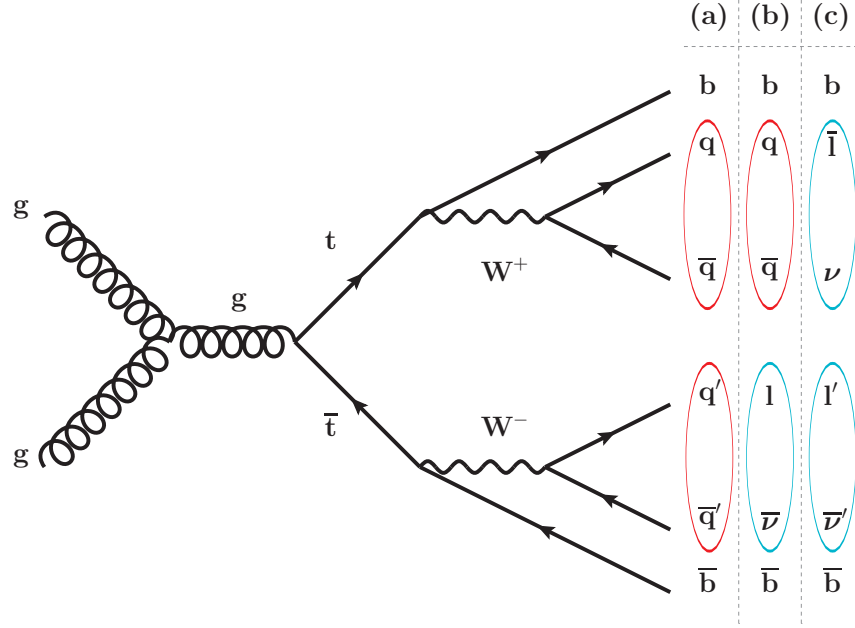


Figure 2.12.: LO Feynman diagram for a all-jet (a), lepto+jets (b) and dilepton (c) $t\bar{t}$ decay.

2.2.2. Top quark mass definitions

The top quark is a color-charged particle and thus the definition of its mass cannot be formulated analog to the mass of a color-neutral particle. According to the Higgs mechanism the mass parameter of each quark is introduced in the QCD Lagrangian, and according to Eq. 2.54 it is proportional to the Yukawa coupling of the quark to the Higgs boson. Moreover, as was already discussed in the perturbation theory the loop correction arise order by order in the particle propagator, hence, the massive quark propagator has a form of:

$$D(p) = \frac{i}{p + m_0 - \Sigma(p)}, \quad (2.66)$$

where m_0 denotes a renormalized mass and $\Sigma(p)$ corresponds to the renormalized one-particle irreducible quark self-energy. The top quark mass parameter m_0 has to be defined within the renormalization scheme [18]:

Pole mass scheme: is defined in perturbative QCD and prescribes subtraction of the divergent correction in the mass of the top quark in a way that the pole of the quark propagator is fixed order by order.

$\overline{\text{MS}}$ mass scheme: prescribes subtraction of the $1/\epsilon$ pole in the divergent correction

in the top quark mass, which results in a shift of the position of the pole in the top quark propagator at all orders in perturbation theory.

Thus, it is possible to have a perturbation expression of the pole mass m_{pole} in terms of the $\overline{\text{MS}}$ mass $m_{\overline{\text{MS}}}$ which known up to four loops [50]:

$$m_{\text{pole}} = m_{\overline{\text{MS}}} \left(1 + 0.4244\alpha_S + 0.8345\alpha_S^2 + 2.375\alpha_S^3 + (8.49 \pm 0.25)\alpha_S^4 \right) = m_{\overline{\text{MS}}} + 7.557 + 1.617 + 0.501 + 0.195 \pm 0.005 \text{ GeV} \quad (2.67)$$

for $m_{\overline{\text{MS}}} = 163.643 \text{ GeV}$ and $\alpha_S \equiv \alpha_S^6(m_{\overline{\text{MS}}}) = 0.1088$. The pole mass and $\overline{\text{MS}}$ mass start to differ at the NLO calculation level. However, the pole mass scheme is applicable only at the perturbative theory, while at the non-perturbation scale above $\Lambda_{\text{QCD}} = 200 \text{ MeV}$ the pole mass carries a renormalon ambiguity, called mass infrared renormalos [51].

In order to avoid this issue, the following scheme is introduced:

MSR mass scheme: is defined using the self-energy diagrams that are used in the pole mass scheme, but the contribution is possible only till the loop momenta is of the order larger than a scale R that typically is chosen to be 1 GeV .

As discussed in [52], the position of the pole mass is expressed by:

$$m_{\text{pole}} = m_{\overline{\text{MS}}} + N \int_0^{m_{\overline{\text{MS}}}} dl \alpha_S + \mathcal{O}(\alpha_S^2) = m_{\overline{\text{MS}}} + N \int_0^{m_{\overline{\text{MS}}}} dl \alpha_S(l), \quad (2.68)$$

where l denotes the scale of the loop momentum, $\mathcal{O}(\alpha_S)$ is a correction to the position of the pole in the quark propagator and N denotes the normalization factor. Since it is assumed that ultraviolet divergences are solved by renormalization, an upper cutoff is utilized, chosen to be $\mu_R = m_{\overline{\text{MS}}}$. The MSR mass can be define as:

$$m_{\text{MSR}}(R) = m_{\overline{\text{MS}}} + N \int_R^{m_{\overline{\text{MS}}}} dl \alpha_S(l), \quad (2.69)$$

or equivalently

$$m_{\text{pole}} = m_{\text{MSR}}(R) + N \int_0^R dl \alpha_S(l). \quad (2.70)$$

Thus, it is possible to interpolate between these mass schemes:

$$m_{\text{MSR}}(R) \xrightarrow{R \rightarrow 0} m_{\text{pole}} \\ m_{\text{MSR}}(R = m_{\overline{\text{MS}}}(m_{\overline{\text{MS}}})) = m_{\overline{\text{MS}}}(m_{\overline{\text{MS}}}). \quad (2.71)$$

3. Experimental setup

The measurements presented in this thesis are based on the data collected with the Compact Muon Solenoid (CMS) detector [6] at the Large Hadron Collider (LHC) [53] during the period 2015-2016, where collisions between two proton beams of 6.5 TeV energy each were recorded. The total certified integrated luminosity was 2.1 fb^{-1} and 35.9 fb^{-1} for a center-of-mass energy of 13 TeV in 2015 and in 2016, respectively.

3.1. The Large Hadron Collider

The Large Hadron Collider (LHC) is the world largest running collider and can provide proton-proton, proton-Pb¹, Pb-Pb and Xe-Xe collisions. It is composed of two accelerator rings, that contain the counter-rotating particle beams colliding at four interaction points (IP) and is the latest upgrade to the CERN² accelerator complex presented in Figure 3.1. The LHC with a circumference of 26.7 km is installed in a tunnel at 100 m underground, which previously housed LEP³, the world highest energy electron-positron collider, until its shutdown in 2000.

The LHC hosts four main particle detectors located at its interaction points, which are designed for different physics programs. The two general purpose, high luminosity experiments ATLAS⁴ [54] and CMS [6] are dedicated to precise measurements of SM quantities and searches of physics beyond the SM. Another two experiments, requiring optimization for low transverse momentum physics, ALICE⁵ [55] and LHCb⁶ [56], are designed mainly for heavy-ion and heavy-flavor physics, respectively. The design of the ATLAS detector allows the reconstructing jets with an excellent resolution and emphasizes on very good flavor tagging, particle identification, and standalone muon measurement. The CMS detector excels for a very precise muon track momentum and electron (photon) energy resolution measurements, as well as flavor tagging. A highly hermetic structure of both detectors allows an accurate determination of the missing transverse momentum. The ALICE detector is especially efficient during the track reconstruction and particle identification, which is necessary for the analysis of the busy heavy-ion environment occurring during

¹(fully stripped lead ions $^{286}\text{Pb}^{+82}$ accelerating to an energy of 2.76 TeV/nucleon)

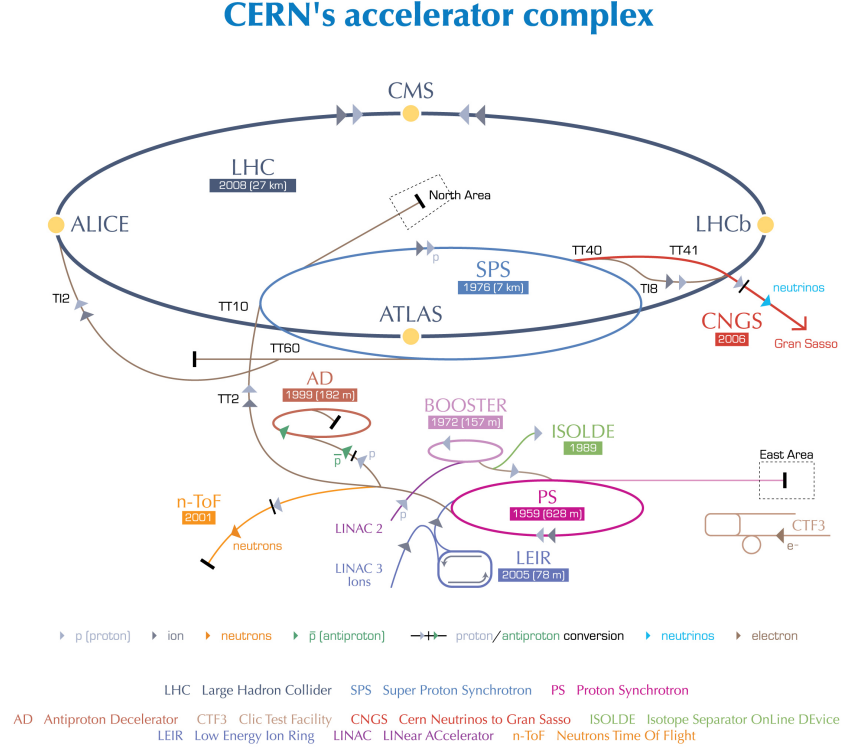
²the European Organization for Nuclear Research, near Geneva

³Large Electron-Positron Collider

⁴A Toroidal LHC Apparatus

⁵A Large Ion Collider Experiment

⁶Large Hadron Collider beauty experiment



European Organization for Nuclear Research | Organisation européenne pour la recherche nucléaire

© CERN 2008

Figure 3.1.: Schematic demonstration of CERN's accelerator complex [57].

heavy-ion collisions. Thus, this detector aims to study the physics of strongly interacting matter at extreme energy densities. The LHCb detector is a single arm forward spectrometer with a trigger for fully hadronic b-hadron decays and has very good low-momentum track resolution and particle identification. It is designed mainly to study b-quark physics.

The main aim of the LHC physics program, an exploration of the proton-proton interactions at novel energies, will be discussed in this thesis, while the other important part, study of the ion-ion and ion-proton collisions, can be found elsewhere [58]. Currently, the LHC machine is designed to operate at pp center-of-mass energy of up to 14 TeV, being the most powerful particle collider ever built. The LHC draws its proton beams from a chain of pre-accelerators that are used to accelerate the proton bunches to an energy of 450 GeV before their injection into the LHC ring. The protons are produced by ionizing Hydrogen and, having a kinetic energy of about 100 keV, are directed into the linear accelerator, Linac 2, where accelerating radio-frequency cavities are installed that bring the protons to an energy of 50 MeV. The beam is then passed to the Proton Synchrotron Booster, where the proton energy is increased to 1.4 GeV, followed by the Proton Synchrotron (PS), accelerating protons to 25 GeV. The last pre-accelerator in the chain is the Super Proton Synchrotron (SPS), that brings the protons to an energy of 450 GeV and passes them into the

LHC rings via two tunnels. In the LHC each beam is accelerated using 16 superconducting cavities. Here the radio frequency cavities, with an operational frequency of 400 MHz, give an acceleration of 485 keV per turn, thus in about 20 minutes the proton energy is increased to 6.5 TeV.

The upper bound of the proton beam energy achieved at the LHC depends on the tunnel length and the achievable dipole magnetic field, that is needed to bend the beam paths. The maximum field strength of the 1232 LHC super-conducting dipole magnets operated at 1.9 K is 8.33 T, that is restricted by the not optimal LEP tunnel structure. The former LEP tunnel, actual LHC tunnel, has eight straight sections, that were used for the multiple accelerations per orbit to compensate the energy loss due to the synchrotron radiation of the electrons and positrons. Thus, the bend of the proton beam at the curved sections of the tunnel is tighter than a perfectly circular ring. Therefore, the maximum achievable energy for each LHC beam is 7 TeV. Having the aim to increase the energy of each beam up to 50 TeV, a new hypothetical future hadron collider the Future Circular Collider (FCC) with an accelerator ring of 100 km may be build as an extension to the CERN accelerator complex. This collider with 100 TeV center-of-mass energy could be the next phase of collider based exploration of the fundamental physics.

Besides energy, one of the most important quantities in collider physics is the *instantaneous luminosity* $\mathcal{L}$ of the colliding beams, which, integrated over time, provides information about the size of the recorded dataset. The instantaneous luminosity is expressed in units of $\text{s}^{-1} \text{cm}^{-2}$ and is calculated as a function of the LHC beam parameters as

$$\mathcal{L} = f_{\text{rev}} \frac{n_{\text{b}} \cdot N_{\text{p}}^2}{4\pi\sigma_x^* \sigma_y^*} \cdot F(\theta_c, \sigma_{x,y}^*, \sigma_z), \quad (3.1)$$

where each ring can accommodate up to $n_{\text{b}} = 2808$ bunches with proton bunch densities of $N_{\text{p}} = 1.14 \cdot 10^{11}$, the revolution (collision) frequency is $f_{\text{rev}} = 11245.5 \text{ Hz}$ and the transverse beam sizes at the interaction point is $\sigma_{x,y}^*$. Additionally, a geometrical correction factor F is introduced to explain the reduction of the instantaneous luminosity due to the crossing angle θ_c between the beams at the collision point. The *integrated luminosity* $\mathcal{L}_{\text{int}} = \int \mathcal{L}(t) dt$ characterizes the statistical precision of any measurement. The expected number of events for a given process with the cross section σ , can be calculated as $N_{\text{ev}} = \sigma \cdot \mathcal{L}_{\text{int}}$.

The first operation of the LHC machine, Run 1 period, was from March 2010 until February 2013. During 2010 - 2011, proton beams collided at $\sqrt{s} = 7 \text{ TeV}$ and during 2012 at $\sqrt{s} = 8 \text{ TeV}$. In 2015, the LHC started the second data taking, Run 2 period, with an increased energy of the proton beams of 6.5 GeV, which leads to a center-of-mass collision energy of $\sqrt{s} = 13 \text{ TeV}$. In this thesis, measurements based on the data collected during 2015 and 2016 are presented. Fig. 3.2 shows the cumulative curves for the integrated luminosity recorded at that time.

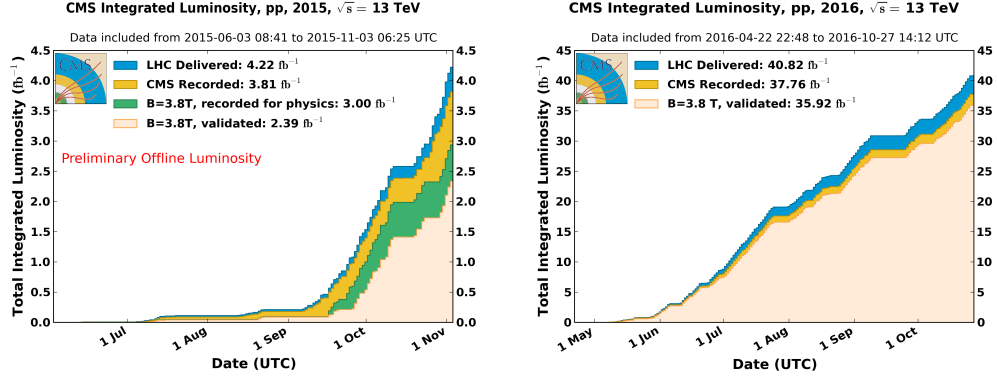


Figure 3.2.: Cumulative curves for the integrated luminosity delivered by LHC (azure), recorded by CMS (orange) and certified as good for physics analysis during stable beams (light orange) [59]. Additionally, for 2015 data the luminosity recorded by CMS with full magnetic field (green) is shown. The luminosity validated for physics analysis corresponds to data recorded with all detector sub-systems and reconstructed physics objects showing a good performance.

3.2. Description of the CMS experiment

The Compact Muon Solenoid (CMS) detector is a 4π hermetic, multi-purpose particle detector built with the main goal of discovering phenomena of physics beyond the standard model (BSM) as well as studying precisely the standard model (SM) parameters. The CMS detector utilizes a powerful superconducting solenoid with an internal diameter of 6 m, which provides a highly uniform magnetic field of $B = 3.8$ T used to bend the paths of charged particles from pp collisions. The size of the CMS detector is 21.6 m in the axial direction and 14.6 m in diameter with a weight of 14,000 t. It is about two times smaller in the volume and two times heavier than the similar purpose ATLAS detector that is 46 m long, 25 m wide and weights 7000 t. A real view inside the CMS detector is shown in Fig.3.3.

The detector consists of many cylindrical detection layers, nested around the beam axis at the interaction point. Each layer exploits different properties of particles to identify and measure energy and momentum of each of them, as discussed in the Chapter 6. The biggest components of the CMS experiment are the tracker system, that is responsible for an accurate momentum measurement, an electromagnetic calorimeter, that is designed to detect and measure electrons and photons, a hadronic calorimeter, that is constructed to entirely surround the collision and to measure strongly interacting particles, the strong magnet and the muon system, which is needed to accurately measure very high momentum particles, namely muons. All mentioned subsystems can be grouped into a barrel section and two endcaps. The layout of the CMS detector is shown in Fig. 3.4. A detailed description of the CMS detector can be found elsewhere [6]. A brief outline of the components that are most

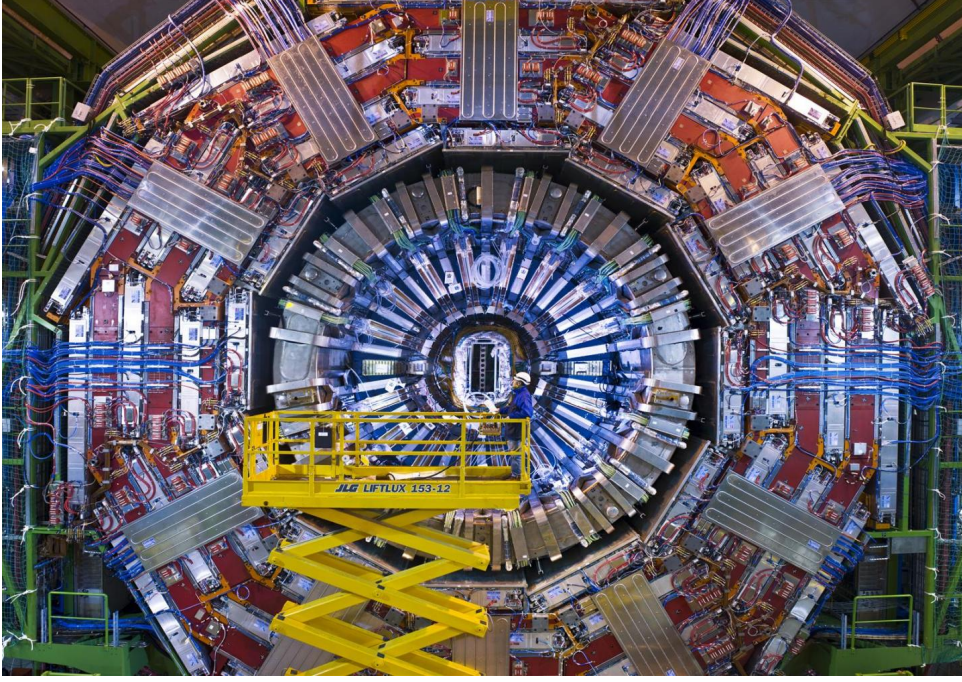


Figure 3.3.: Opened CMS detector showing the inner detector components [60].

relevant for this analysis is given in the followings.

Coordinate system and reconstruction of kinematic variables

The CMS coordinate system is a right-handed Cartesian system with the coordinate origin at the nominal beam-crossing point. The z axis points in the counter-clockwise rotating beam direction, the x axis points towards the center of the LHC ring and y axis is pointing perpendicular to the plane of the LHC ring. In the transverse x - y plane the azimuthal angle $\phi = \arctan(y/x)$ is measured with respect to the x axis, where $\phi = 0$, thus, parallel to the positive y axis $\phi = \pi/2$. The azimuthal angle together with the radial distance, $r = \sqrt{x^2 + y^2}$, exploits the cylindrical symmetry of the detector. The polar angle θ is measured from the positive z axis in the y - z plane and is defined as $\tan\theta = r/z$. A kinematic quantity using the z direction is usually expressed by the *pseudorapidity* η , which characterizes the boost of the particles and is defined as

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right], \quad (3.2)$$

where $\eta = 0$ perpendicular to the beam direction. For massless particles or in the ultra-relativistic limit the pseudorapidity approximately converges with the rapidity y . The difference in rapidity is Lorentz invariant under a boost along the longitudinal axis, therefore it is used together with the azimuthal angle to define an angular

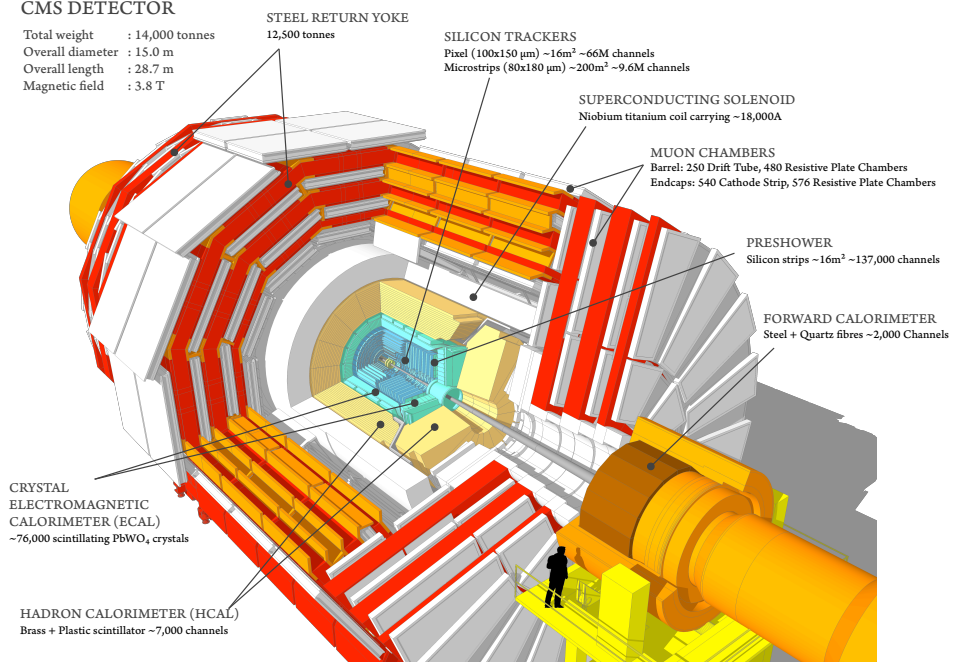


Figure 3.4.: Sectional view of the CMS detector. Main components of the detector are marked with basic information about them [61].

distance ΔR between two objects i and j ,

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}, \quad (3.3)$$

where $\Delta\eta = \eta_i - \eta_j$ and $\Delta\phi = \phi_i - \phi_j$. Due to the fact that in pp collisions the momentum along the beam axis is unknown, while many of the "newborn" objects escape through the beam pipe, carry out a very small amount of transverse momentum, momentum conservation can be applied in the transverse plane. Therefore, the *transverse momentum* p_T , is an extremely useful characteristic of particles and is expressed as the momentum vector projection on the x - y plane perpendicular to the beam axis:

$$p_T = \sqrt{p_x^2 + p_y^2}. \quad (3.4)$$

The presence of particles that mostly do not interact with the CMS detector cause a momentum imbalance of the total p_T sum of all particles, therefore the *missing transverse energy* is defined as the magnitude of the negative vector sum of the transverse momenta of all N particles in a collision:

$$\cancel{E}_T = \left| - \sum_{i=1}^N \vec{p}_{T,i} \right|. \quad (3.5)$$

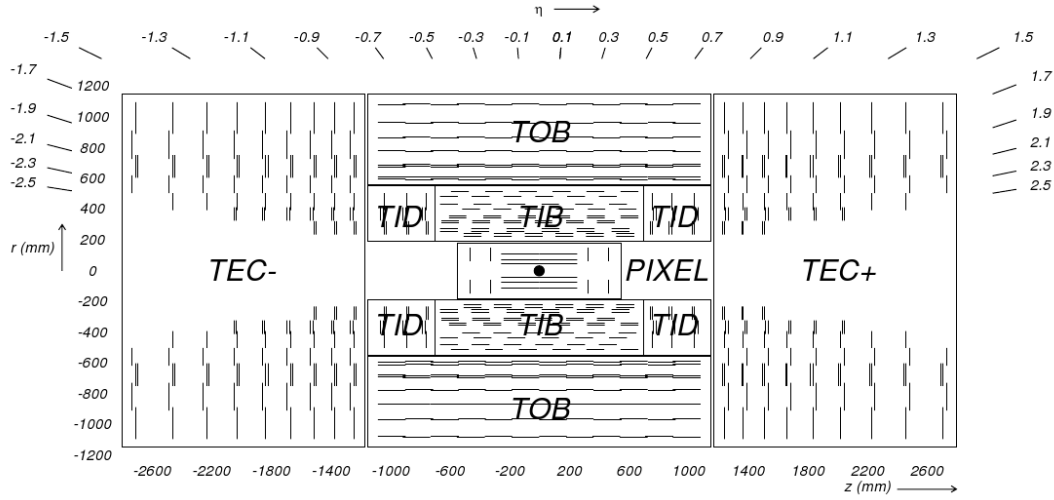


Figure 3.5.: View of the CMS tracker in the rz -plane [6]. Each line in the strip tracker represents a silicon strip detector, whereas lines in the pixel detector represent ladders and petals on which the detectors are mounted in the barrel and endcaps, respectively.

3.2.1. Tracking system

Particles produced in the proton-proton collision first enter the inner most part of the CMS detector, the inner tracker, where charged particles interacting with the material leave electrical signals (hits). The hits are used for a precise and efficient particle trajectory (track) reconstruction. Tracks are later used to find the collision particles origin (vertex), sign of its charge and its transverse momentum. In order to determine the charge assigned to the particle, the direction of its track bending in a magnetic field is measured, The transverse momentum of the charged particle is estimated with the banding radius for trajectory:

$$p_T \left[\frac{\text{GeV}}{c} \right] = 0.2998 \cdot r[\text{m}] \cdot B[\text{T}]. \quad (3.6)$$

The CMS tracker system covers the pseudorapidity range of $|\eta| < 2.5$ and consists of silicon pixel and silicon strip detector modules. A sketch of the CMS tracking system is shown in Fig. 3.5, where a fine-grained structure is important to provide a pure and efficient charged-particle trajectory reconstruction in a high particle multiplicity environment.

Pixel Detector The silicon pixel tracker is the closest detector to the interaction region and effectively covers a pseudorapidity range of $|\eta| < 2.5$. It consists of three 53 cm long cylindrical barrel layers (BPIX) and two endcap discs (FPIX) on each side, as demonstrated in Fig. 3.6. The radii of the barrel cylinders are 4.4 cm, 7.3 cm and 10.2 cm, while endcap discs are located at longitudinal distances of $z = \pm 34$ cm and ± 46.5 cm. All these parts, constructed from

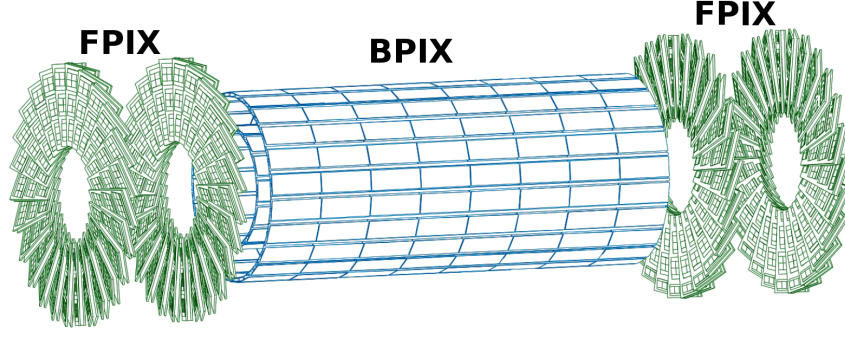


Figure 3.6.: Detailed layout of the pixel tracker [6]. The barrel part (BPIX) and the forward disks (FPIX).

66 million *hybrid silicon pixels* of $100 \times 150 \mu\text{m}^2$ size, have been arranged in 1440 modules and cover an area of about 1 m^2 . Thus, a size of each pixel leads to a spatial resolution of around $10 \mu\text{m}$ in the $r - \phi$ plane and of about $20 \mu\text{m}$ in the z direction. The high granularity of the pixel detector provides about 99% hit-reconstruction efficiency, which is crucial for the secondary vertices reconstruction. After the 2016 data taking it was replaced by a new pixel detector.

Strip Detector The strip tracker has a radius of 1.1 m and is placed in the remaining tracker volume. It covers the same pseudorapidity range as the pixel detector. The increasing distance to the interaction point, where the expected particle flux is smaller, allows the use of the silicon strip sensor "cells", which are less-granular as the pixel detector components. In total around 11.4 million strip sensors, which are grouped into 15 148 strip modules, are installed in the strip detector. The strip tracker consists of an inner part, the *Tracker Inner Barrel and Discs* (TIB, TIDs), and an outer part, the *Tracker Outer Barrel and Endcaps* (TOB, TECs). Since the strip size depends on the distance to the interaction point, in the radial region between 20 cm and 55 cm (TIB and TIDs) strip sensors with a size of $10 \text{ cm} \times 80 \mu\text{m}$ and a thickness of $320 \mu\text{m}$ are used, while in the region with higher radii up to 116 cm (TOB and TEC), the strip sensors have a size of $25 \text{ cm} \times 180 \mu\text{m}$ and thickness of $500 \mu\text{m}$. The strip tracker provides a very good resolution in $r - \phi$ for the barrel and $z - \phi$ for the endcap region because of its geometry. The inner part, that has four barrel layers from TIB and three discs on the each side from the TID, provides up to four $r - \phi$ track measurements. Depending on the size of the strip sensors, the TIB reaches the resulting resolution of $23\text{-}34 \mu\text{m}$ in $r - \phi$ and of $23 \mu\text{m}$ in z . The outer part gives six $r - \phi$ measurements from the TOB, with the resolution of $35\text{-}52 \mu\text{m}$ in $r - \phi$ and $52 \mu\text{m}$ in z , while the TEC, covering the region of $22.5 \text{ cm} < r < 113.5 \text{ cm}$ and $124 \text{ cm} < |z| < 282 \text{ cm}$, has nine discs

in each endcap and provides up to nine ϕ measurements. Additionally, since the strip detector provides less information on the orthogonal direction, more than a third of the strip tracker modules are double-layered and are shown as a double line in Fig. 3.5. The second layer is mounted back-to-back with a tilt of 100 *mrad*, allowing a stereo measurement.

3.2.2. Electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) [62] is a homogeneous calorimeter that surrounds the inner tracking system. It is designed to stop particles in an absorber and to measure the amount of kinetic energy deposited, where as the absorber material lead-tungstate (PbWO_4) crystals are used. These act not only as absorbers, but also as active detector material (scintillators). The corresponding scintillation process is very fast and about 80% of the scintillation light is emitted within 25 ns, which is in good agreement with the LHC bunch crossing time. The main idea of the ECAL is to measure energy deposition of the particles interacting mainly electromagnetically, e.g., electrons and photons. However, another important characteristic of the measured particles is the shape of the developed shower in the calorimeter that appears when a particle is passing through the calorimeter and interacts with it. The shower shape is described by the *Molière radius*, the average radius of a cylinder containing approximately 90% of the energy of the electromagnetic shower, which shows the width of the shower, and a *radiation length* (X_0), is a density-weighted depth of material where a traversing electron reduces its initial energy E_0 to $\frac{1}{e}E_0$ due to radiation, which points to the depth of the electromagnetic shower. Thus, the main reasons to choose the lead-tungstate crystals of PbWO_4 was their high density of 8.3g/cm³, the short radiation length X_0 (0.89 cm) and the small Molière radius of 2.2 cm. These conditions provide a good resolution of the energy measurement of electromagnetic showers. Fig. 3.7 demonstrates the main components of the ECAL: a barrel ECAL (EB), that covers a pseudorapidity range of $|\eta| < 1.479$, an endcap ECAL (EE) on both sides of the EB, that extends the range to $|\eta| < 3$ and a preshower ECAL (ES) located in front of EE calorimeter discs that covers the pseudorapidity range of $1.653 < |\eta| < 2.6$. The total crystal mass in the EB is 67.4 t, while in EE it is 22.9 t.

ECAL Barrel (EB) The EB calorimeter, within a radius of 1.3 m, is made of 61 200 crystals with a high granularity. The crystal mounted closest to the tracker, at front face, have a cross section of $22 \times 22 \text{ mm}^2$ or roughly 0.0174×0.0174 in (η, ϕ) , while at the rear face it is $26 \times 26 \text{ mm}^2$. The length of a crystal is 230 mm, corresponding to $25.6 X_0$. Thus, the electromagnetic showers have no energy leakage beyond the ECAL.

ECAL Endcap (EE) The ECAL endcap discs consist of 7 324 crystals, with the size

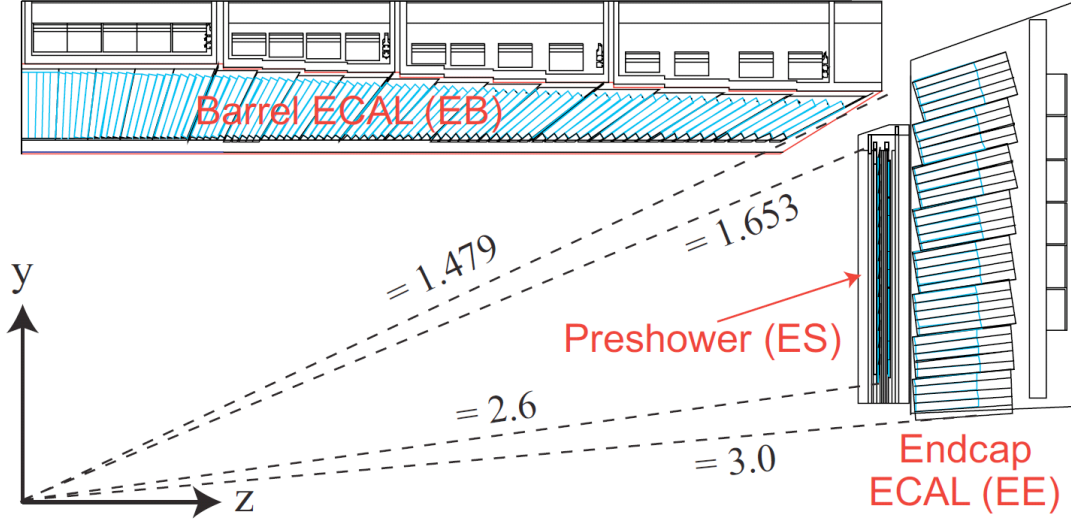


Figure 3.7.: Geometric view of one quarter of the ECAL [63].

of $28.62 \times 28.62 \text{ mm}^2$ in the front face and $30 \times 30 \text{ mm}^2$ in the back. The length of the crystal is 220 mm, corresponding to $24.7 X_0$.

ECAL Preshower (ES) The ES calorimeter, located in the endcap region, aims to provide a handle to distinguish γ and π^0 , where π^0 is detected by its decay into two collimated photons. The ES consists of two thin lead absorbers and one silicon sensor plane, thus its total thickness is 20 cm, that corresponds to $3 X_0$.

To characterize the ECAL performance, the relative energy resolution is parameterized as

$$\left(\frac{\sigma(E)}{E}\right)^2 = C^2 + \left(\frac{S}{\sqrt{E[\text{GeV}]}}\right)^2 + \left(\frac{N}{E[\text{GeV}]}\right)^2, \quad (3.7)$$

where the constants C , S , and N describe calibration errors, sampling fluctuation and noise contributions due to electronics, respectively. The barrel ECAL energy resolution for electrons was measured in a test beam [64] and is determined to be 0.4% for a 100 GeV electron, with the parameters of $C = 0.3\%$, $S = 2.8\%$ and $N = 12\%$.

3.2.3. Hadronic calorimeter

The hadronic calorimeter (HCAL) [65] is designed to measure the energy of hadronic jets, made of quarks and gluons. The particles that reach the HCAL calorimeter are mainly heavier than the ones detected in the ECAL and cause broader and deeper showers compared to electromagnetic ones. Therefore, a huge amount of material is required. One of the important quantities, which characterises the depth of hadronic showers, is the *nuclear interaction length* (λ_I), which is the mean distance that a

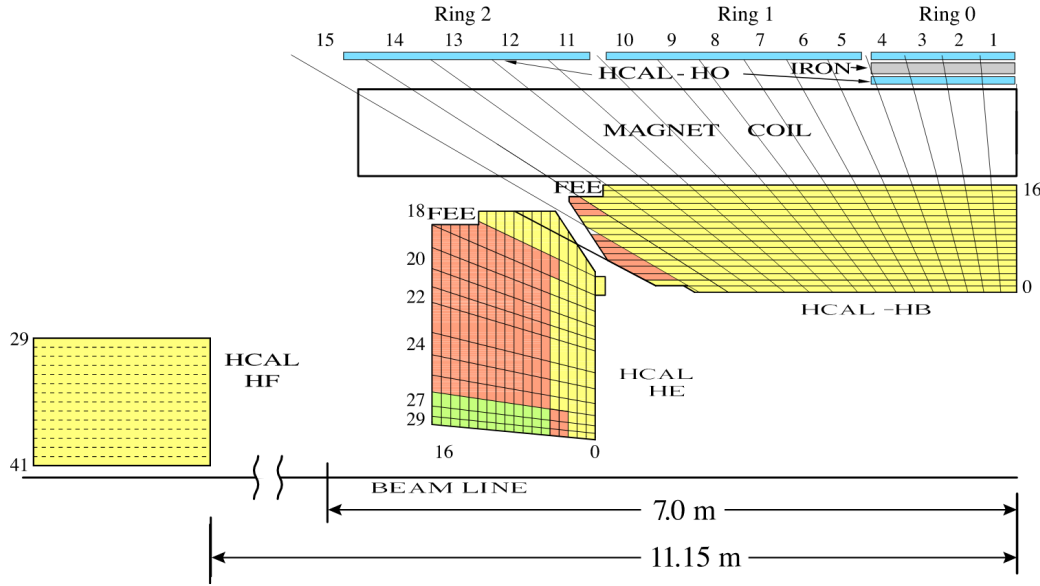


Figure 3.8.: The CMS HCAL detector (quarter slice). “FEE” indicates the locations of the Front End Electronics for HB and HE. The signals of the tower segments with the same color are added optically, to provide the HCAL “longitudinal” segmentation. HB, HE and HF are built of 36 identical azimuthal wedges ($\Delta\phi = 20$ degrees) [66].

travelled hadron travels before undergoing an inelastic nuclear interaction. The HCAL is hermetic and covers the pseudorapidity range up to $|\eta| < 5.2$, which is very useful for the missing transverse energy determination. Unlike the ECAL, the HCAL is a sampling calorimeter, where absorber, i.e. layers of brass or steel, and active detector materials, i.e. tiles of plastic scintillator, are separated. The two subdetectors, namely the HCAL barrel (HB) and the HCAL endcaps (HE), are mounted inside the magnet system, that is presented in Section 3.2.4, while the two others, the HCAL outer (HO) and the HCAL forward (HF), are placed outside the magnet. These components are shown in Fig. 3.8.

HCAL Barrel (HB) The HB calorimeter covers the pseudorapidity range up to $|\eta| < 1.3$ and spans the detector at radii of 1.77 m to 2.95 m. The HB utilizes 16 layers of absorber plates, the first and the last layers are made of steel, while all others are made of non-magnetic brass, and 17 layers of plastic scintillator tiles. Those calorimeter layers are segmented in “towers”. The HB has a fine granularity with the size of the each tower of 0.087×0.087 in (η, ϕ) , where one tower matches 5×5 ECAL crystal. The effective depth for the HB depends on the pseudorapidity, thus at $|\eta| = 0$ it is equal to $5.8 \lambda_I$, while for $|\eta| = 1.3 - 10 \lambda_I$.

HCAL Endcap (HE) The HE calorimeter extends the pseudorapidity range up to $1.3 < |\eta| < 3$ and is mounted on the endcap iron yokes. A similar approach

in terms of the material compared to the HB is taken for the endcap HCAL, that consists of 17 layers of absorber and 18 scintillator layers. However, unlike the HB the HE, starting from $|\eta| \geq 1.6$, has a lower granularity, thus a larger crystal size of 0.17×0.17 in (η, ϕ) . It has an effective depth of about $10 \lambda_I$.

HCAL Outer (HO) The HO calorimeter is an additional system of two layers of plastic scintillator and an iron absorber in the middle that is placed in the outer barrel region and used as a tail-catcher, in order to cover all hadronic showers. The HO covers the pseudorapidity range of $0.4 < |\eta| < 1.26$ and is placed outside the solenoid, where the whole magnet coil is used as an additional absorber material. The HO calorimeter extends the effective depth at $|\eta| = 0$ up to $11.8 \lambda_I$.

HCAL Forward (HF) The HF calorimeter is a radiation hard calorimeter where two modules of 900 towers supplement the detector coverage up to $2.9 < |\eta| < 5.2$. It consist of a steel absorber, that has along the z axis a size of 1.65 m, and quartz-fiber scintillators which are interfaced with photomultiplier tubes (PMT). The quartz-fiber material was chosen due to its radiation hardness, which is required by the Cherenkov⁷ based technology for the charged particles detection. The PMT are used to convert the radiated Cherenkov light of the traversing charged particles into electric signal. The HF towers have two read-out channels at different depths, that are used to separate the electromagnetic and hadronic shower component. Its hadronic effective depth is $10 \lambda_I$.

The relative energy resolution of the CMS HCAL is worse than the ECAL one, since the charged hadrons start to deposit the part of their energy already in the ECAL and the dead absorber material causes worse resolution. Therefore, the relative energy resolution of pion measured during the test beam [67] corresponds to

$$\left(\frac{\sigma(E)}{E}\right)^2 = C^2 + \left(\frac{S}{\sqrt{E[\text{GeV}]}}\right)^2, \quad (3.8)$$

where the constant C , which describes calibration errors, is equal 9%, the stochastic term S , which corresponds to the sampling fluctuation, is equal 110% and the contribution from the noise is around 200 MeV.

3.2.4. Solenoid magnet

The solenoid magnet is the central element of the detector layout, with the main purpose to provide a strong magnetic field to bend the path of charged particles, where the bending power is used for precise momentum calculation and the electric

⁷Cherenkov radiation is the electromagnetic radiation emitted by the charged particles passing through a dielectric medium.

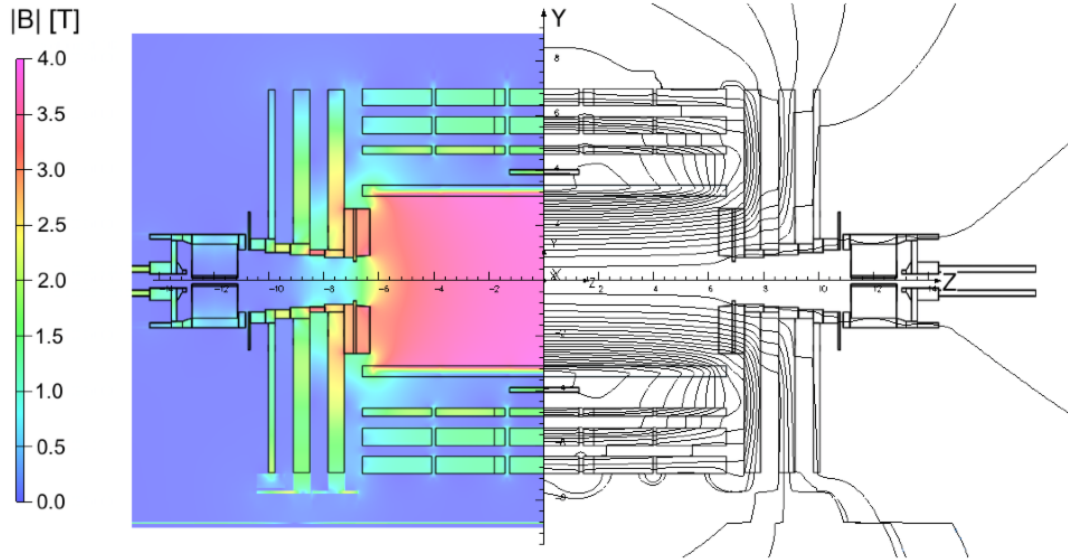


Figure 3.9.: Value of $|B|$ (left) and field lines (right) predicted on a longitudinal section of the CMS detector, for the underground model at a central magnetic flux density of 3.8 T [68]. Each field line represents a magnetic flux increment of 6 Wb.

charged assignment. The magnet hosts the tracking and calorimeter systems and provides within its volume a uniform magnetic field in the z direction up to 4 T. The magnetic flux is returned using a massive iron yoke, which also helps to generate a strong 2 T magnetic field inside the muon system. The schematic demonstration of the magnetic field strength and lines of the CMS detector is shown in Fig. 3.9.

The strong magnetic field of the CMS superconducting solenoid allows for a two times smaller detector volume compared to the ATLAS detector, that is using mainly a system of toroidal magnets. The schematic illustration of the magnet system with the magnetic field configurations for both the CMS and ATLAS experiments is shown in Fig. 3.10. The main advantages of the solenoid is a simple structure and the large homogeneous magnetic field provided inside the magnet, thus the CMS inner tracker is placed in an uniform magnetic field, while the toroid magnet has a complex structure and provides a non-uniform magnetic field. However, the solenoid magnet has a weak opposite field in the return yoke.

The CMS magnet with the total weight of 220 t has a length of 12.5 m and a diameter of the free bore of 6.3 m and is composed of 50 km NbTi superconducting cable with a cross section of $64 \times 22 \text{ mm}^2$. The liquid helium cryostat is used to cool down the system to the operating temperature of 4.6 K. The CMS solenoid is operating at $B = 3.8 \text{ T}$ with a current of 19.5 kA, which allows a storage of 2.6 GJ total energy.

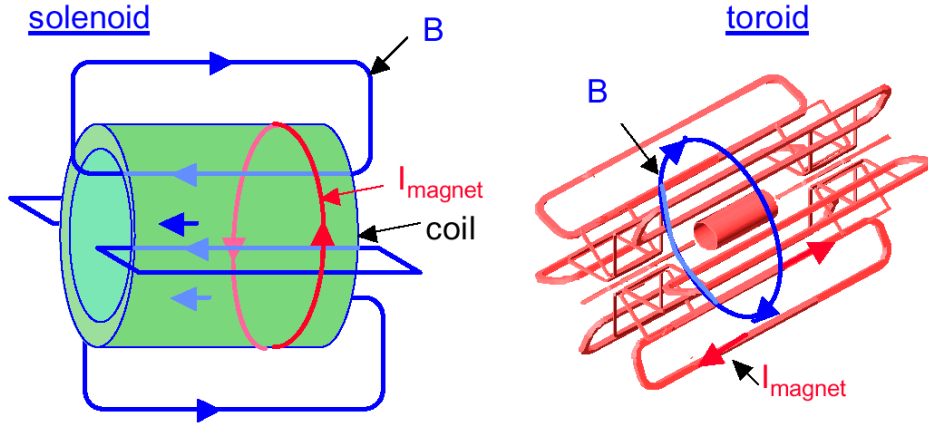


Figure 3.10.: Magnetic field configuration for solenoid (left) and toroid (right) magnet system [69].

3.2.5. Muon system

The muon system is one of the most important parts of the CMS experiment, which is designed to identify muons with high efficiency. Since muons are 200 times heavier than electrons and do not interact strongly, they are passing the tracker system and calorimeters mostly without energy loss. Therefore, to measure high momentum particles, such as muons, the muon system together with a very strong magnet is needed. The muon detector surrounds the magnet with a steel flux-return yoke [6], that provides the field flux return as well as it does not allow particles other than muons to pass to this system. The CMS muon system consists of 25 000 m² of detection planes. Barrel drift tubes (DT) cover the pseudorapidity range of $|\eta| < 1.2$, endcap cathode strip chambers (CSC) cover the pseudorapidity range of $0.9 < |\eta| < 2.5$, and resistive plate chambers (RPC) cover the pseudorapidity range of $|\eta| < 1.6$, as demonstrated in Fig. 3.11. The DT and the CSC detectors are used to measure the precise position of the muons and thus their momentum, while the RPC chambers are used to provide the quick information for the Level-1 trigger.

Drift tubes (DT) The drift tubes are designed to detect the muons in the barrel part of the detector with the uniform magnetic field, where the low rate of the muons and small amount of the neutron-induced background are expected. Each tube contains a stretched wire within a gas volume and the tubes are staggered by half a tube and arranged into SuperLayers (SL). An ionizing

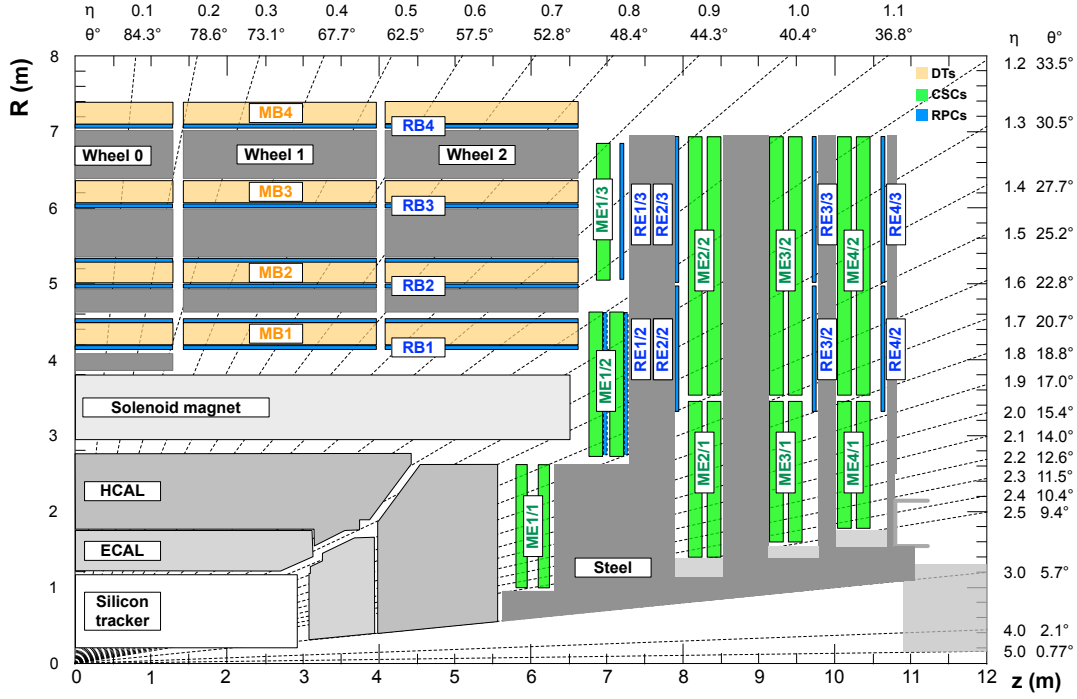


Figure 3.11.: Layout of one quadrant of the CMS detector [70]. Muon stations are placed in the steel return yoke of the magnet (gray). The four DT stations in the barrel (yellow), the four CSC disks in the endcap (green), and the RPCs in both barrel and endcap regions (blue) are shown.

particle passing through the volume of the tube liberates the electrons of the gas atoms, which are then registered along the wire. The structure of the DT improves the efficiency of the muon detection and provides a good time resolution of the measurement. The DT is build of detector sectors, each containing four stations, the group of the SL, interleaving with the layers of the yoke plates. The measurement of the muon coordinate in the r - ϕ plane is provided with the eight drift tubes placed in the first three stations, while in the z coordinate the track is detected by the four drift tubes. The part of the z measurement is missing in the last section. The spatial and angular resolution of each DT station is $100 \mu\text{m}$ and 1 mrad , respectively.

Cathode strip chambers (CSC) The cathode strip chambers are placed in the endcap region, where a strong and inhomogeneous magnetic field is present. Thus, the detector in this region should be very fast and radiation-hard, which requires finer segmented structure of the chambers than in the DT detector case. Therefore, the CSC's chambers have a multiwire structure, where each cathode plane is segmented into strips running across wires. In total 486 cathode strip trapezoidal chambers are used. Unlike the DT the CSC's spatial and angular resolution is $100\text{-}200 \mu\text{m}$ and 10 mrad , respectively, that is requested due to the large background contamination and high muon rate. In total, the CSC

detector contains of four stations, that are placed between the yoke plates, where positively-charged anode wires placed perpendicular in the chambers read out the time of the muon crossing, as well as a its η position, while the negatively-charged copper cathode strips provide the information on the r - ϕ coordinates.

Resistive plate chambers (RPC) The resistive plate chambers are fast gaseous detectors. Their information is used independently of the DT and the CSC on the first level of the muon triggering process, providing an independent and fast trigger with a sharp p_T threshold. The RPCs have a good spatial resolution and a time resolution of 1 ns. In the barrel part six RPCs are placed. To provide good measurement of the low momentum muons four RPCs are attached from inner and outer side of the first two DT stations, while a single layer of RPCs is mounted in front of the drift chambers for the outer DT stations. In the endcaps the RPCs are placed in the four yoke stations parallel to the CSCs. The RPCs are double-gap chambers, consisting of two parallel plates, an anode and a cathode, and are separated by a gas volume. The read-out strips are installed between the two gaps and oriented along the z direction.

The described muon systems provide a very efficient muon identification with an efficiency above 95% [71] for the muons at the pseudorapidity range of $|\eta| < 2.4$, with the transverse momentum larger than a few GeV.

3.2.6. Trigger system

The aim of the CMS trigger system is to select only the most interesting hard-scattering events that are about one out of a thousand, while during the LHC operation the collisions are delivered every 25 ns, which corresponds to the bunch-crossing rate of 31.6 MHz, that is about 600 millions collision per second. Therefore, the efficient event triggering, that avoids the loss of an important data, is extremely important, which is also crucial during the luminosity rise, when the number of the proton interactions also increase. A two-stage approach is done with a level-1 trigger (L1T) and then with a high-level trigger (HLT).

Level-1 trigger (L1T) The level-1 trigger (L1T) is the first stage, which performs rough estimations of the event parameters and based on loose requirements takes a very quick decision for selecting events. The level-1 trigger system uses the calorimeter and muon systems at the reduced granularity and thus is composed of two *paths*: a *calorimeter trigger* and a *muon trigger*, as shown schematically in Fig. 3.12. The calorimeter trigger uses inputs from the ECAL and the HCAL, while the muon trigger uses information from the muon tracker subsystems such as the DT, CSS, and RPC. The L1T calculates the total

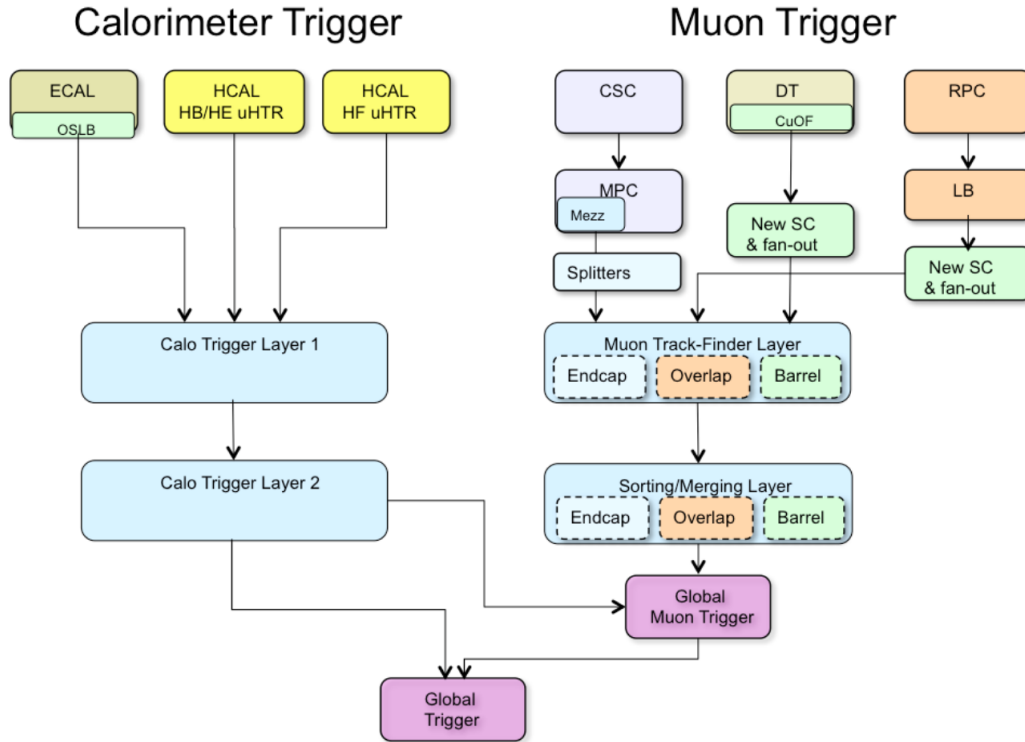


Figure 3.12.: Data flow for the overall trigger upgrade [72].

energy of the event and scans it for high-energetic particles like muons, using information from the muon trigger, and electrons, photons, jets, utilizing input from the calorimeter trigger. Jets are signatures of hard scattering events with hadronic final states, as described in detail in Section 6.5. Information from the two trigger paths is combined in the global trigger (GT) processor, that applies kinematic and topological selections to the events and decides within $1\,\mu\text{s}$ whether an event is accepted. The L1T output rate is less than 100 kHz.

High-level trigger (HLT) The high-level trigger (HLT) is the second level system, that uses the fully built data events, which are accepted by the L1T. It is based on the CMS software, which reconstructs the data with higher granularity. Various trigger logics (*trigger paths*) are implemented in the HLT. Those trigger paths are created to cover a broad range of physics signatures used in many analyses. The HLT output rate is a few hundred Hz.

4. Simulation of hadronic collision

In order to simulate a proton-proton collision event, or generally a hadronic collision, standard Monte Carlo (MC) event generators, which are based on a factorized approach as demonstrated in Fig. 4.1, are used. This is necessary since colorless hadrons have a composite parton (quarks and gluons) structure. They are inherently non-perturbative objects, thus perturbative theory is not applicable to the full evolution of high-energy hadron-hadron collisions. The factorized approach relies

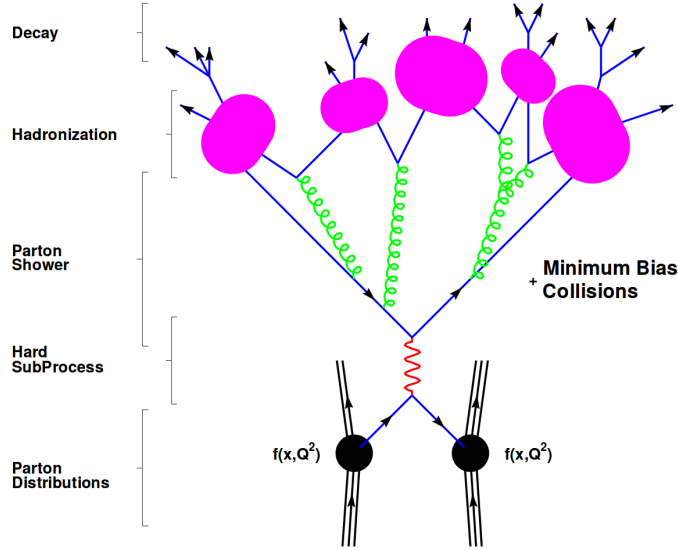


Figure 4.1.: The sketch of the basic structure of event generation in proton-proton collisions [73].

on the *factorization theorem* [74], which allows to separate the “hard” scattering, where two initial partons give rise to an interaction at a large exchanged transverse momentum p_T , and the underlying event (UE) process. The “hard” scattering is characterized by a large scale (small distance), where the running strong coupling constant α_S is small, allowing to apply a perturbative approach to the calculation of the matrix element¹. On the other hand, the UE represents the whole remaining activity of a hadronic collision that occurs at lower, non-perturbative scales, where α_S becomes large. These processes are described by phenomenological and theoretical models [75], which will be discussed later in this chapter. Hence, the hard

¹The matrix element defines the hard scattering in the cross section.

scattering is accompanied by soft interactions, which results in hadronic collision being a complex multiparticle processes.

4.1. Proton structure

Protons are not elementary particles and are made out of valence (up, up, down) quarks², and sea (anti-)quarks³ and gluons. Therefore, the probability that determines which parton from the proton will interact with a fraction x of the proton's longitudinal momentum, is determined by the *parton distribution function* (PDF)⁴ [77] $f_i(x, Q^2)$ for a parton of flavor i and the momentum transfer squared Q^2 . The PDFs are evaluated using a certain factorization scale μ_F , which separates the partonic cross section, where only the hard scattering part is considered, from the non-perturbative regime of the PDFs⁵. A common choice for μ_F is on the order of the hard scattering scale [78] Q^2 so that $\mu_F = \mu_R = Q^2$, where μ_R is the renormalization scale.

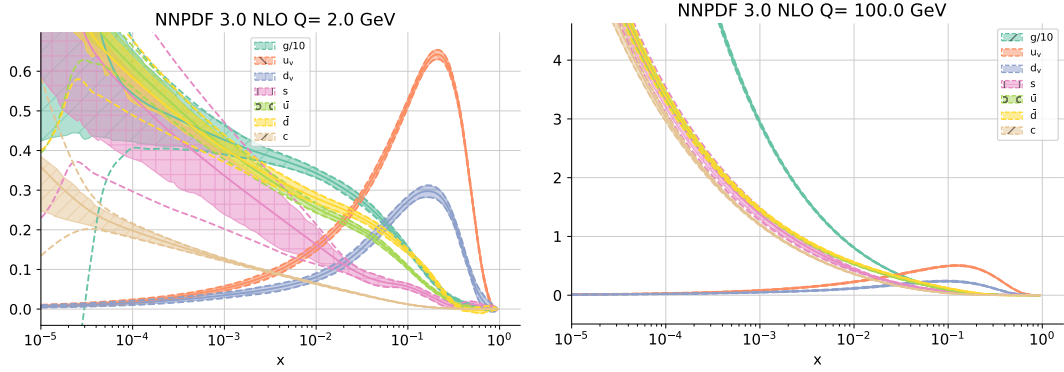


Figure 4.2.: The NNPDF parton distribution functions for $\alpha_S = 0.118$ at NLO for the low (left) and high (right) Q^2 energy scale as a function of x [79].

The PDFs are extracted by using data from different experiments and processes, mainly ep collisions, where the results are going down to the very low scales, e.g., the results from the HERA collider with Q^2 up to 1 GeV^2 . With the PDFs being available at an initial scale it is possible to evaluate its evolution to higher Q^2 values by using the DGLAP equations [80–82]. Figure 4.2 shows an example for the evaluation PDFs at higher energy scales. The figure shows the NNPDF3.0 NLO

²The valence quarks of the proton define its quantum numbers.

³The valence quarks constantly radiate and absorb gluons, which may split into virtual $q\bar{q}$ pairs. These virtual quarks are called sea-quarks.

⁴First described in *parton model* by Richard Feynman in 1969 [76], where hadrons are described as bound states of point-like constituents, e.g., partons.

⁵The PDFs are inherently non-perturbative in nature, and are not calculable a priori.

PDF set [83]⁶ for each parton flavor evaluated at $Q = 2 \text{ GeV}$ and $Q = 100 \text{ GeV}$. It is observed, that at low scales the valence quarks carry a large portion of the proton momentum, while at high scales the gluons carry an increasing fraction of the proton momentum. The evaluation of the PDF sets at higher energy scales is important for the LHC physics program, as demonstrated in Fig. 4.3, where the accessible values of x and Q^2 for the LHC at $\sqrt{s} = 13 \text{ TeV}$ are shown. Thus, to produce top quark pairs, a minimal momentum fraction of $x = \sqrt{x_1 x_2} \geq 2m_t/\sqrt{s} \approx 0.027$ is needed, while the minimal energy needed for $t\bar{t}$ pair production is $\sqrt{\hat{s}} = 2m_t \approx 345 \text{ GeV}$, with the partonic center-of-mass energy is $\sqrt{\hat{s}} = \sqrt{s x_1 x_2}$.

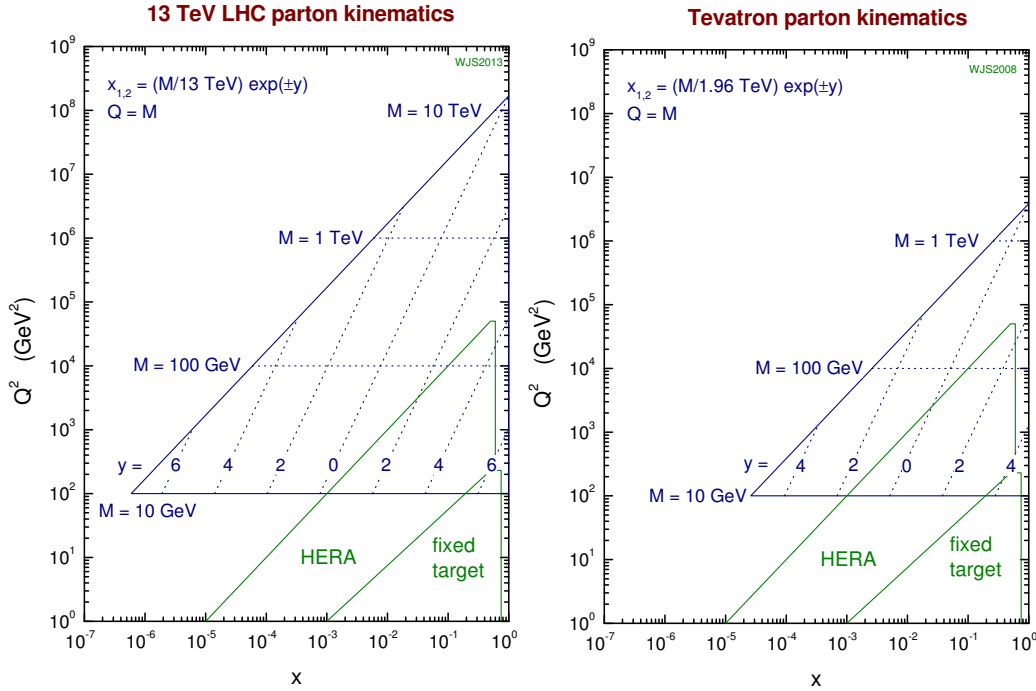


Figure 4.3.: The accessible x and Q^2 values at fixed-target experiments, e.i., HERA and LHC (left) or Tevatron (right) [84]. M corresponds to the invariant mass and y to the rapidity.

The most popular PDF sets used at the LHC are NNPDF [83], CTEQ [85], MSTW [86], HERAPDF [87], ABKM [88] and GJR [89], which extract PDFs from the measurements from different experiments, like SLAC [90], ZEUS and H1 [91], CDF [92], CMS [93] etc.

⁶ The NNPDF3.0 NLO PDF set is used to evaluate the systematic uncertainty in the top quark mass discussed in Chapter 8

4.2. Hard scattering and matrix element

The probability of two hadrons with four-momenta P_1, P_2 to interact and give rise to a hard $p_1 p_2 \rightarrow X$ process is given by the corresponding cross section σ :

$$\sigma_{p_1 p_2 \rightarrow X} = \sum_{i,j} \int dx_1 f_{i,p_1}(x_1, Q^2) \int dx_2 f_{j,p_2}(x_2, Q^2) \hat{\sigma}_{ij \rightarrow X}(x_1 P_1, x_2 P_2, \alpha_S, \mu_F, \mu_R), \quad (4.1)$$

which is factorized into a “hard” (high- Q^2) part $\hat{\sigma}_{ij \rightarrow X}(x_1 P_1, x_2 P_2, \alpha_S, \mu_F, \mu_R)$, e.g., the partonic cross section for the process $p_1 p_2 \rightarrow X$, and “soft” (low- Q^2) part, given by the parton distribution functions $f_{i,p_1/p_2}(x_{1/2}, Q^2)$ for colliding proton beams 1 and 2. The transverse momentum of the incoming partons, $x_1 P_1$ and $x_2 P_2$, at high- Q^2 scales, e.g., involved in hard interaction, is negligible, due to the *collinear approximation*. Therefore, the parton level cross section $\hat{\sigma}_{ij \rightarrow X}(\alpha_S, \mu_F, \mu_R)$ depends mainly on the strong coupling constant, and the factorization and renormalization scales⁷. It is calculated using the corresponding *matrix element (ME)*, which represents the probability amplitude $\mathcal{M}$ for the transition from the initial state into a final state X , e.g., $t\bar{t}$. Considering an arbitrary final state F , which can include explicitly all orders, the hard scattering cross section is given by:

$$\sigma_F = \sum_{k=0}^{N_{\text{legs}}} d\Phi_{X+k} \left| \sum_{l=0}^{N_{\text{loops}}} \mathcal{M}_{X+k}^{(l)} \right|, \quad (4.2)$$

where $d\Phi_{X+k}$ is the phase space for the final state X with k additional outgoing partons (*legs*), and $\mathcal{M}_{X+k}^{(l)}$ is the ME for the final state X with k legs and an additional l loops. A leg is a real correction to the cross section and the loop is a virtual correction, e.g., it enters in the calculations, but it does not produce a new particle. Thus, by choosing k and l , the different final states F are defined:

- $k = 0, l = 0$: X production at the leading order (LO) $2 \rightarrow 2$ process,
- $k = n, l = 0$: X production with additional quarks and gluons in the final state⁸ at the LO $2 \rightarrow 3$ process,
- $k + l \leq n$: X production at the next ^{n} -to-leading order ($N^n\text{LO}$), process like NLO $2 \rightarrow 2$, when $k + l = 1$.

Increasing the number of the legs and loops reduces the dependence on μ_F and μ_R in the calculation.

In order to show the impact of the NLO correction to the LO cross section, a *K-factor* can be used, which is defined as the ratio of the NLO to LO cross section

⁷In case of the $t\bar{t}$ cross section, the mass of top quark is used as factorization and renormalization scale: $\mu_F = \mu_R = m_t$

⁸It is possible for the current method to calculate the case with $k = 5$.

of the given process. This factor depends on several variables used at LO and NLO, like the PDF, or the factorization and renormalization scales. However, have it does not include information about a shape change of the studied observable, which can accrue due to the NLO correction. Thus, this K-factor does not provide sufficient description of the impact of the NLO corrections on the LO cross section, but it is still useful to calculate interesting processes at LHC [94].

Within the scope of this thesis the LO $2 \rightarrow 2+n$ Matrix element MADGRAPH5 [95], and the NLO $2 \rightarrow 2$ MEs from the aMC@NLO v2.2.2 [95] and POWHEG V2 [96–99] generators are used.

4.3. Parton shower

The parton shower concept is introduced to approximate higher-order real emission corrections to the hard scattering, which is possible until a certain minimum value for the hadronization scale ($Q_{\min} \sim 1 \text{ GeV}$) is reached. At this scale perturbation theory does not work anymore and the shower evolution is stopped. The shower evolves with the radiation of gluons from quarks⁹, where these gluons can further emit gluons or split into $q\bar{q}$ pairs. There are four possible QCD splittings at LO: $q \rightarrow qg$, $q \rightarrow gq$, and $g \rightarrow q\bar{q}$ and $g \rightarrow gg$, which are mostly soft, because at this stage Q decreases, and thus α_S increases. This *evolution* is described by effective parametrizations using the DGLAP equations [80], which give the probability of emissions a parton produced at the highest energy scale $Q_{\max}$. The probability $P_a(t)$ of not splitting between two scales t_0 and t is defined as:

$$P_a(t) = \Gamma_a(t) e^{(-\int_{t_0}^t dt' \Gamma_a(t'))}, \quad (4.3)$$

where $e^{(-\int_{t_0}^t dt' \Gamma_a(t'))}$ is the Sudakov form factor [100, 101], that gives the probability of a parton a with an initial scale t_0 to develop without showering until next scale t , and $\Gamma_a(t)$ describes the parton evolution.

Not only the outgoing parton of the hard scattering can radiate gluons, which is called *final state radiation (FSR)*, but also the initial state partons can radiate gluons, which is referred to as *initial state radiation (ISR)*. The ISR is handled within the *backwards evolution* scheme, since the incoming parton momenta is fixed by the hard process. This is possible by using the evolved PDFs in the Sudakov factor, which suppresses the backwards evolution if no matching “mother” partons are present in the PDF.

In this analysis PYTHIA 8.219 [16, 17] and HERWIG++ 2.4 [102] are used for the parton shower. In PYTHIA, parton radiation is ordered by decreasing momentum

⁹After the hard scattering happens, the outgoing partons have a large momentum. However, partons are colored objects and this leads to QCD radiation, namely gluons.

$Q^2 = p_T^2$, where p_T is the transverse momentum of the emitting parton, while in HERWIG++ the parton emission is angular-ordered $Q^2 = 2E_a^2(1 - \cos\theta)$, where θ is an angle between the radiated gluon and the parent parton with energy E_a .

4.4. Matrix element - parton shower matching

The calculation done by the matrix element can already include the radiation of hard partons. Therefore, it is necessary to avoid a possible double counting of equivalent phase-space configurations, when using the parton shower, and to compute the terms that are not included by the ME. Thus, the double counting between ME and PS can be avoided by using a phase space cut. This uses the PS below the cutoff, which is the minimum value of hadronization scale $Q_{\min}$, and the ME above the cutoff, which is the chosen factorization scale. This is possible by using a *matching algorithm* that merges samples with different multiplicities.

The most popular approaches used at CMS are the multi-jet matching MLM [103] and the next-to-leading order jet matching FxFx [104]. The MLM matching is based on a shower veto, i.e., *truncated shower*, which rejects events that contain parton emissions from the PS that are already calculated by the ME. The PS can develop additional partons only after the ME is below the cutoff of ~ 20 GeV. The FxFx matching scheme is similar to the MLM matching, but it uses a next-to-leading order ME, thus not only the real correction are taken into account, but also the virtual ones. Therefore, applying the same order PS gives the correct distribution, but with NLO leading-logarithmic (NLO+LL) accuracy, which yields negative weights for those phase-space regions where the exact matrix element is smaller than the shower approximation. Thus, special weights need to be applied in this case to the PS.

Summarizing, MLM matching is used with LO ME generators like MADGRAPH, while FxFx matching is applied to the NLO ME generators like aMC@NLO. However, the NLO ME POWHEG v2 does not use any additional matching schemes, since it contains the hardest emission and positive virtual corrections, thus negative event weights do not occur. This generator gives the lowest jet p_T scale to the PS generator to further develop the event.

4.5. Hadronization and decay

The evolution of the parton shower stops at the minimum value of the hadronization scale, on the order of 1 GeV, where the perturbation theory does not work, and the strong coupling is large $\alpha_S \approx 1$. Below this scale the final state partons start to form color neutral hadrons via the *Lund* or *string fragmentation* model [105] implemented in PYTHIA 6.4 [16] or the *cluster fragmentation* model [106,107] used by HERWIG++ 2.4.

The string fragmentation models are based on the assumption of a uniform color potential, which only depends on the distance between $q\bar{q}$ pairs in the parton shower. Thus, with the spatial separation of quark and antiquark, the potential linearly increases, leading to the formation of color strings between them, while gluons are ignored and treated as "kinks". At a certain distance, when the potential energy is on the order of hadron masses, the production of a new $q\bar{q}$ pair is energetically favored. This new pair is produced by a tunneling process, with the tunneling probability:

$$P \propto \exp\left(-\frac{\pi m_{q,\perp}^2}{k}\right) = \exp\left(-\frac{\pi p_{q,\perp}^2}{k}\right) \exp\left(-\frac{\pi m_q^2}{k}\right), \quad (4.4)$$

where $k \approx 1 \text{ GeV/fm}$ is the color string tension, $p_{q,\perp}^2$ the non-perturbative transverse momentum of the quark (around 250 MeV^2), and m is the mass of new hadron. This implies the suppression of heavy quark production, since $u\bar{u} : d\bar{d} : s\bar{s} : c\bar{c} \approx 1 : 1 : 0.3 : 10^{-11}$. The string evolution is regulated by the string-fragmentation function [105]:

$$f(f) \propto \frac{1}{z}(1-z)^{p_1} \exp\left(\frac{-p_2 m_{q,\perp}^2}{z}\right), \quad (4.5)$$

where z points into the direction of the color string, and $p_1 = 0.3$ and $p_2 = 0.58 \text{ GeV}^{-2}$ are the Lund parameters. The evolution continues until the string potential energy can not produce any more hadrons. A schematic example of color field lines between the $q\bar{q}$ pair and the motion and the breakup of string system are shown in Fig. 4.4.

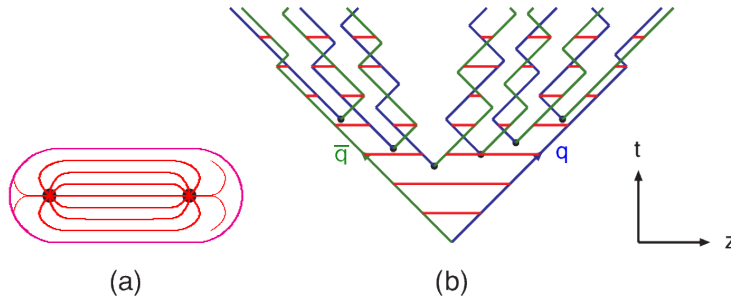


Figure 4.4.: A graphical representation of the color field lines between quark and antiquark (a). The motion and breakup of a string system in longitudinal direction over time (b) [108].

The cluster fragmentation model introduces the intermediate step of forming clusters of quarks and antiquarks with mass scales of a few GeV. While gluons that remain at the end of the parton shower evolution are split into $q\bar{q}$ pairs, assuming the local compensation of color. The formed colorless clusters of partons with the mass higher than a certain threshold will decay directly to two lighter clusters or to a pair of hadrons via a quasi-two-body decay [109], when they are too light. In the

Fig. 4.5 the clusters are shown as light green ellipses.

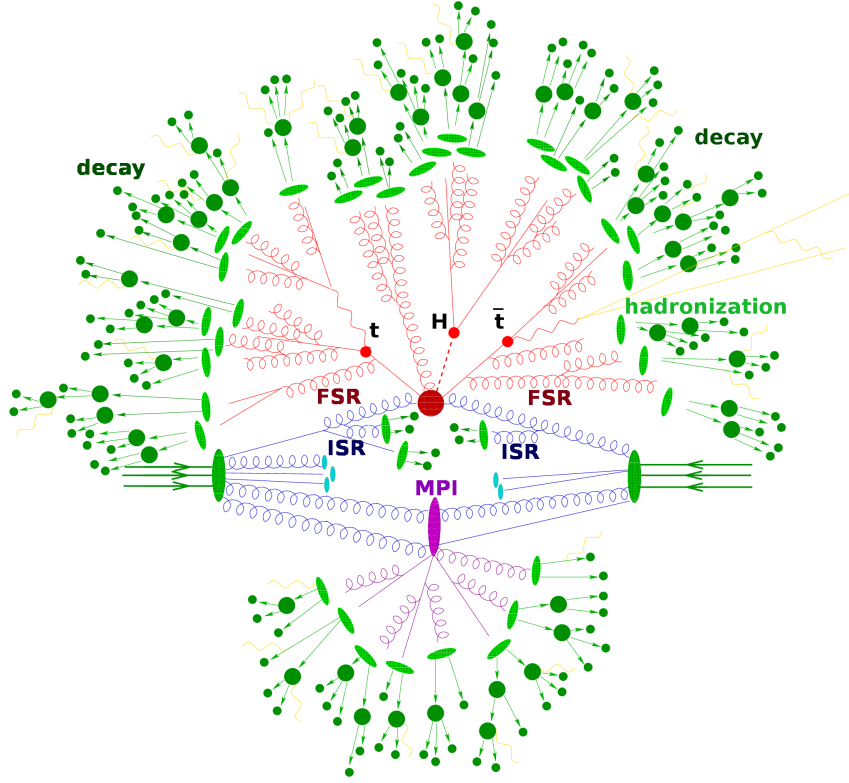


Figure 4.5.: Sketch of a $t\bar{t}H$ event produced in a proton-proton collision. The red marker in the center represents the hard collision, which is surrounded by a tree-like structure representing Bremsstrahlung as simulated by parton showers. The purple marker indicates a secondary hard scattering event. Parton-to-hadron transitions are represented by light green markers, dark green markers indicate hadron decays, while yellow lines signal soft photon radiation. This is modified version of the figure from [108].

The formed hadrons have a very short lifetime and decay further within a typical timescale of $c\tau \gtrsim 10$ nm. In MC, the processes of hadronization and decay are simulated at the generator level.

4.6. Underlying event and pileup

A underlying event activity, typically gluon scattering processes $gg \rightarrow gg$, occurring in addition to the hard scattering at lower scales constitutes the underlying event (UE). It consists of several components, such as multiple parton interactions (MPI), beam-beam remnants (BBR), initial- and final-state radiation (ISR and FSR). These processes can produce a large number of colored partons that, similarly to partons from the hard scattering, are later re-arranged into colorless hadrons, according to the description of the hadronization process given in Section 4.5. Thus, the UE

yields much higher multiplicity in the event, therefore this effect must be taken into account by the MC. This is usually done by using appropriate models that aim to describe UE and return the values of the sensitive parameters, which are grouped in so called “tune”. In the CMS the main UE tunes, used during MC productions, are CUETP8M2T4 [110] and EE5C [111]. The *pileup* effect, additional pp interactions that overlap with the event of interest, is also simulated in the multi-purpose MC generators, i.e., PYTHIA 8.219 [16, 17]. More details on the pileup are given in Section 6.2.

The MPIs are found to be especially relevant for the underlying event and minimum-bias (pileup) processes. Several MPIs may occur within the same collision event, which are represented as additional 2-to-2 parton-parton interactions, as shown in Fig. 4.5. Moreover, the contribution from the MPI increases with increasing collision energy, which is due to the increase of the parton density at small x .

In the final state, the colored partons from MPI may be connected by color strings. These strings can overlap with the color strings originating from the partons of the hard scattering interaction as well as with color fields from other MPIs. Therefore, the formation of the corresponding color strings is regulated by the *color reconnection* (CR) mechanism, which models the total potential energy between QCD color charges [112]. The CR mechanism allows changes in the color configuration of the event before the hadronization process. A good understanding and description of CR in MC is unavoidable, in order to describe the increase of the mean transverse momentum $\langle p_{\perp} \rangle$ as a function of multiplicity, as shown in Fig. 5.1. It was recently discovered that the modeling of the CR is one of the dominant systematic uncertainty source for many physics analyses. Therefore, PYTHIA 8.219, which was used to simulate the samples used for the precise top quark measurement in this thesis, was tuned to the new CR models, in order to give a consistent description of the CR in MC. The results of the MC CR tune study is documented in Chapter 5.

4.7. Event generators

Many MC generators are designed to simulate hadronic collision events. In the followings, the main properties of the generators used in the studies presented in this thesis, will be described.

MadGraph5 [95, 115]: is a LO $2 \rightarrow 2 + n$ matrix element generator, where n denotes the number of additional partons. Since MADGRAPH is time-consuming, events are usually generated with $n \leq 4$. MADGRAPH does not include the parton shower or underlying event calculation. Therefore, MADGRAPH is interfaced with PYTHIA through the MLM [103] matching procedure. It can be also interfaced with HERWIG++.

amc@nlo v2.2.2 [95, 104]: is a NLO $2 \rightarrow 3$ ME generator that provides precise

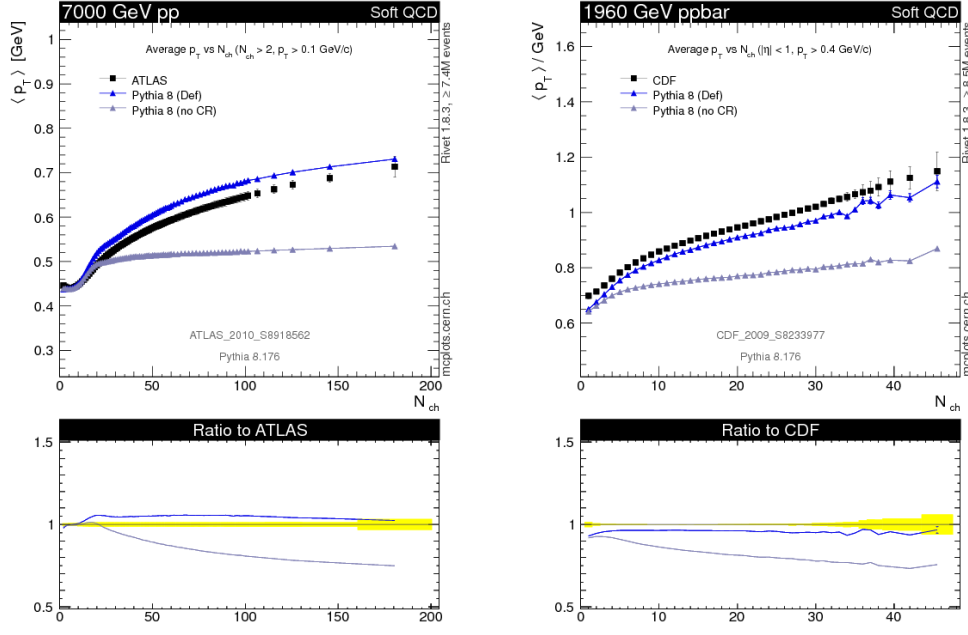


Figure 4.6.: Measurements $\langle p_{\perp} \rangle$ of single charged hadrons vs. number of charged particles N_{ch} in minimum-bias events at 7000 GeV [113] (left) and 1960 GeV [114] (right) compared to PYTHIA 8.175, with and without color reconnections switched on.

cross section estimation. Since no parton shower is included, the FxFx [104] matching procedure allows to use PYTHIA for the simulation of the UN, PS and hadronization.

powheg V2 [96–99]: is a NLO $2 \rightarrow 2$ ME with up to one additional hard partons. It is based on POWHEG box [98]. POWHEG does not include UE and PS simulation, therefore is interfaced with PYTHIA or HERWIG++, without the help of an additional matching algorithm. This NLO ME generator are expected to show better modelling of the signal and background processes.

pythia 8.219 [16, 17]: is a multi-purpose event generator. The hard scattering is calculated at LO, i.e., with an LO $2 \rightarrow 2$ ME. The ISR and FSR is covered by a p_T -ordered parton shower that follows [80]. The hadronization is performed via the Lund string model. The UE tune used to regulate the MPI and CR contributions is CUETP8M2T4 [116, 117].

herwig++ 2.4 [102]: is a multi-purpose event generator with a LO $2 \rightarrow 2$ ME. The parton shower uses angular-ordered parton emission, e.g., the radiation is generated coherently inside parton splitting cones. The hadronization is performed via a cluster model. The EE5C tune [111] is used for the UE.

Matrix-element corrections to the top quark decay are not applied in the HERWIG++ sample.

The parton distribution function (PDF) set NNPDF3.0 NLO ($\alpha_s = 0.118$) [83] is used as the default parametrization of the PDF in all simulations.

Detector simulation

In order to propagate the simulated particles through the detectors, a full simulation of the CMS detector based on GEANT4 [118–120] is used. This is referred as *full simulation* that aims to calculate the corresponding hardware responses by modeling the sensitive detectors, as well as the inert infrastructure of the detector. In addition to the particle interactions in simulation of the transversed material, the behavior of the electronics that is used to readout the signals is also modeled.

5. Tuning of the color reconnection model in Pythia8

The multiple parton interaction (MPI) models describe parton-parton interactions that occur in addition to the hard interaction in an event. These additional interactions are an unavoidable consequence of hadronic collisions at high energies. Moreover, an increase of the collision energy at the LHC leads to an increase in the MPI activity.

Several parameters in the MC generators can influence the amount of simulated MPIs. The main parameters are the parton distribution function (PDF), the overlap of the matter distributions of the two colliding hadrons, as well as the cut-off p_{T0} parameter. The latter parameter is used to determine the energy dependence of the simulated MPI contribution and is described in the PYTHIA 8 simulation by the equation:

$$p_{T0}(\sqrt{s}) = p_{T0}^{\text{ref}} (\sqrt{s}/\sqrt{s_0})^\epsilon, \quad (5.1)$$

where p_{T0} is the parameter used at a given center-of-mass energy scale $\sqrt{s}$ and p_{T0}^{ref} is the p_{T0} at a reference energy $\sqrt{s_0}$. The parameter ϵ is the energy rescaling space, that describes the energy dependence. Moreover, in the model the QCD 2-to-2 cross section is regularized by the parameter p_{T0} , in order to control its divergent behavior for $p_T \rightarrow 0$.

As mentioned in Section 4.6, the colored partons from the MPI can be connected via color strings. According to the different models for the *color reconnection* (CR) mechanism, it is possible to exchange color information between partons from the MPI as well as the partons from the hard scattering process after the parton shower by reconnection of the color strings. Typically the total string length is minimized. In addition to the string length minimization, random reconnections can also be forced. An example for color reconnections in an $t\bar{t}$ event is shown in Fig. 5.1.

One of the major reasons to study the CR mechanism is its ability to reproduce the increase of average transverse momentum of the charged particles detected in the final states of generic hadronic collisions, as a function of the charged particle multiplicity, as shown Section 4.6. The largest remaining tension with respect to the data is observed in the soft region ($p_T \sim 2$ to 5 GeV), where CR effects are expected to be more relevant. This experimental observation can be described by connecting the color fields of the partons that are produced by the MPIs and the hard scattering process [122]. By introducing correlations between the various partonic systems, it

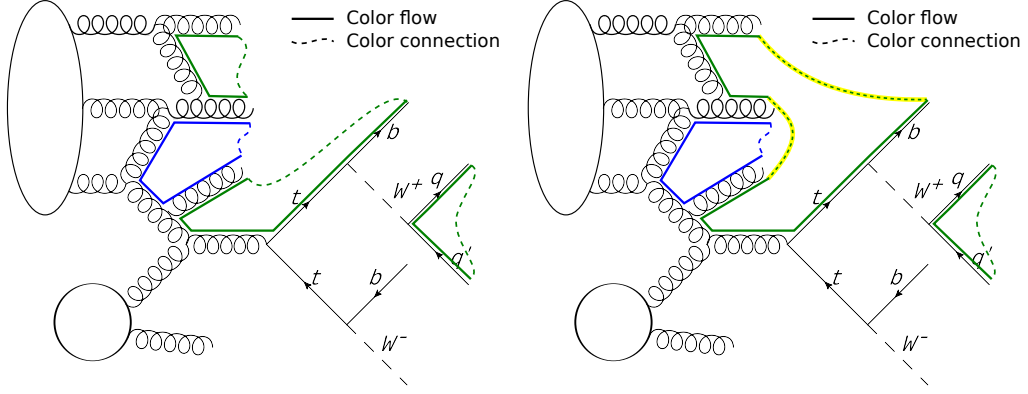


Figure 5.1.: An example of color reconnection in $t\bar{t}$ production [121].

is possible to reduce the number of produced particles in the final state, resulting in a larger average transverse momentum than in a scenario without CR.

However, color reconnection effects are also important for other processes occurring at larger scales in proton-proton collisions. Thus, the CR modeling is one of the dominant sources of systematic uncertainties in the top quark mass measurement [15]. It is claimed that the uncertainty in the top quark mass due to the color reconnection, which is estimated by taking the difference in the predictions of a model with and a model without CR, might represent an underestimation of the assessed uncertainty [15]. Therefore, a better method for estimating the CR modeling uncertainty is suggested, which considers a variety of CR models.

Different phenomenological models have been developed and included in the MC generator since publication of the Run 1 results, e.g., in the PYTHIA 8 MC event generator. Their general idea is to construct a color potential depending on the string length of the color fields, where the string length is minimized in order to create the partonic configuration. This potential is successively applied in the hadronization process. Both the MPI processes and CR mechanism are not completely determined by first principles and include adjustable parameters which can be optimized for a better event modelling. A specified set of these parameters that has been adjusted to better fit certain aspects of the data is referred to as a “tune”.

In this chapter, the results for the tune performed with the *QCD-inspired* [14] CR model are presented. In combination with already existing tunes, that use the *MPI-based* [123] CR model, this set of tunes allows to quantify the performance with respect to several observables, to deepen our understanding of the CR mechanism, and to evaluate uncertainties due to CR effects.

5.1. Color reconnection models in Pythia 8

Recent versions of PYTHIA 8 implement three CR models, which can be tested in proton-proton collisions. These models, referred to as the “MPI-based”, “QCD-

inspired” and “gluon-move” CR models, are briefly described in the following.

MPI-based model [123]: A very simple CR model implemented in the PYTHIA 8, which gives each parton pair a probability to reconnect with a harder system defined as:

$$P_{\text{rec}}(p_T) = \frac{p_{T\text{Rec}}^2}{(p_{T\text{Rec}}^2 + p_T^2)} \text{ with } p_{T\text{Rec}} = R \cdot p_{T0} \quad (5.2)$$

where R is the reconnecting range parameter and p_{T0} is the quantity which regularizes the partonic cross section to avoid divergence at low p_T (Eq. 5.1). According to this formula, MPI systems at high p_T would tend to escape out of the interaction point without being color reconnected to the hard scattering system. Color fields originating from low- p_T MPI systems would be more likely to exchange color. After deciding which systems are connected to the hard scattering products, the configuration which minimizes the string length is chosen. The MPI-based model introduces only one free parameter, R in Eq. 5.2, which can be tuned.

QCD-inspired model [14]: A CR model implemented in PYTHIA that takes into account the QCD color rules in addition to the minimization of the string length. The model constructs all pairs of dipoles that are allowed to reconnect by QCD color rules and uses a simple space-time picture to causally connect the produce strings. Additionally, it allows the creation of junction structures [124], that allow the color combinations represented as Y-shaped “color junction” as shown in Fig. 5.2.

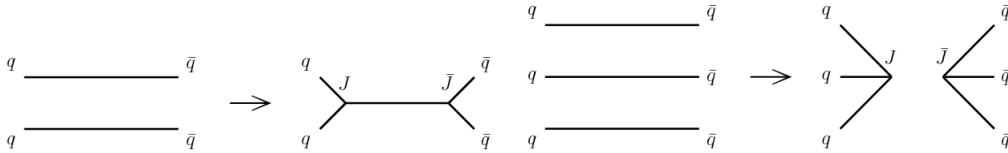


Figure 5.2.: An example of the formation of a connected junction-antijunction-pair (left), and the formation of a junction and antijunction, which are not directly connected color reconnection (right) [14].

The junction structures offer the possibility to include higher-order color connection effects. Junctions are not simply dipoles, but they can be directly produced by three or four dipoles. This is repeated until no further allowed reconnection reduces the total string length λ , given by

$$\lambda = \sum_{\text{dipoles}} \log(1 + \sqrt{2}E/m_0), \quad (5.3)$$

where E corresponds to the energies defined in the dipole or also in the junction rest frame and m_0 is a constant with dimensions of energy of order Λ_{QCD} . The QCD-inspired model then takes the leading-color strings and transforms them into different color configurations based on three principles: first, the SU(3) color rules from QCD determine if two strings are color-compatible; second, a simplistic space-time picture defines the causal contact between the strings; third and finally, the evaluation of λ decides which color configuration is favoured. Detailed information on the QCD-inspired model is presented at [14]. Due to the close relationship between junctions and baryons, the new model predicts an enhanced baryon production. It was shown to be able to consistently describe baryon/meson production ratios, measured both at the LEP and LHC experiments [14], which were not well predicted by tunes using the MPI-based CR model. The free parameters of the QCD-inspired CR model are the lower limit for the mass of the particles, which are not allowed to be color-connected, and the parameters that define of the causal contact between color strings.

Gluon-move model [125]: A CR model in which some or all gluons are selected, along with all color connected pairs of partons. An iterative process in which the change in λ is calculated for each final-state gluon attached to a string. The change in λ results from moving the gluon to a string piece belonging to a different pair of partons. The gluon movement that gives the largest negative λ change is applied. In this scheme, quarks would not be reconnected, since they would always remain in the same position without any color exchange. To improve this model, the flip mechanism was included, which also gives the quarks the possibility to be reconnected. Junctions are allowed to take part in the flipping step as well, but no considerable differences are expected due to the limitation in the junction formation in this model. The main free parameters of the gluon-move model are the lower limit of the string length, that is allowed for color connection, the fraction of gluons allowed to move or flip, and the lower limit of the allowed reduction on the string lengths.

In this thesis, the new CR tune based on the QCD-inspired model is developed.

5.2. Underlying event measurements

One of the most common ways of investigating and constraining MPI and CR effects is to study measurements of the number and of the energy of charged particles produced in a generic proton-proton collision. These types of measurement are sensitive to underlying event (UE) contributions and have been performed at various collision energies by different collaborations [122, 126–128]. The measurements typically include the multiplicity and the sum of p_T of the charged particles as a function of

the p_T of the leading charged particle in the event. The charged particle multiplicity and p_T are easily determined with a tracking system and can be measured even at very low p_T .

These studies usually distinguish between different regions in the transverse plane to the beam direction, as defined by the direction of the leading charged particle. The relative azimuthal distance $\Delta\phi$ is defined between the selected charged particles and the leading charged particle.

- The “toward” region, $|\Delta\phi| < \pi/3$, mainly includes the products of the hard scattering.
- The two “transverse” regions, $\pi/3 < |\Delta\phi| < 2\pi/3$, contain the products of MPI and are affected by contributions from ISR and FSR.
- The “away” region, $|\Delta\phi| > 2\pi/3$, populated by recoiling objects related mostly to the hard scattering process.

More recent measurements [126, 128] subdivide the transverse region into a “TransMIN” and a “TransMAX” region, which are defined to be the regions with the smaller and larger number of particles of the two transverse regions. This is done in order to better disentangle contributions from MPI, ISR and FSR. A sketch of the different regions is shown in Figure 5.3.

5.3. Tuning strategy

In this thesis a new CMS tune using the QCD-inspired model implemented in the PYTHIA 8.219 event generator is presented. The corresponding tune is named CUETP8M2-CR1 in this work. It is obtained by simultaneously constraining the parameters governing the contributions of the MPI and of the QCD-inspired CR model. As baseline, the tune for the nominal MPI-based model is used, officially known as CUETP8M2T4 [110] and named CUETP8M2 in this work. The MPI-based CR tune uses the NNPDF30_lo_as_0130 [83] PDF set, with the hadronization parameters documented in [116], a negative exponential distribution function for modelling the overlap between the two colliding protons, and the strong coupling $\alpha_S^{\text{ISR}} = 0.1108$ for the simulation of the initial-state shower. This value was optimized to describe high jet multiplicity scenarios in top-antitop events [110], when using a next-to-leading (NLO) matrix element simulated by the POWHEG event generator [98] and the matched PS from PYTHIA 8.

As mentioned before, by providing tunes with different CR models, one has a consistent way of evaluating systematic uncertainties due to color reconnection effects for many analyses. The studied tune with the QCD-inspired CR model is

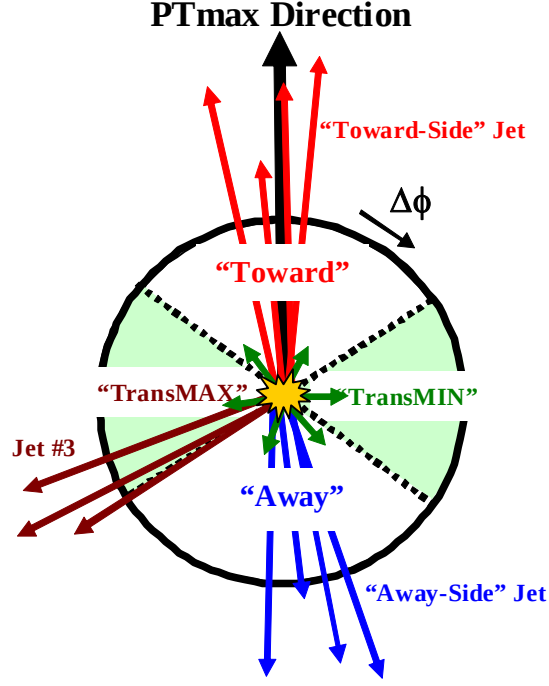


Figure 5.3.: Illustration of the topology of a hadron-hadron collision in which a “hard” parton-parton collision has occurred [129]. The “toward” region contains the “toward-side” jet, while the “away” region contains the “away-side” jet on average. For events with large initial or final-state radiation, the transMAX region contains the third jet, while both the transMAX and transMIN regions receive contributions from the MPI and beam-beam remnants. Thus, the transMIN region is very sensitive to MPI.

obtained by using the UE data measured at 7 TeV [14, 15, 125], as well as the measurements at 13 TeV [126, 130]. In previous tunes, only the measurements at 7 TeV were considered.

The strategy for obtaining the QCD-inspired tune is similar to the one used for deriving the nominal CUETP8M2 since in both cases the same observables, which are sensitive to soft and semi-hard MPI, are considered for constraining the parameters. These observables are:

- the charged particle multiplicity and average p_T sum as a function of the leading charged-particle transverse momentum,
- p_T^{max} , measured by CMS at 13 TeV in TransMIN and TransMAX [126],
- the charged particle multiplicity as a function of pseudorapidity η , measured by CMS at 13 TeV [130].

The region between $0.5 < p_T^{max} < 3$ GeV is excluded from the fit as it was done for the CUETP8M2 tune, since it is affected by contributions of diffractive processes.

Therefore, the free parameters of the corresponding processes are also not considered in the tuning procedure.

The MPI-related parameters, that are fitted for the QCD-inspired tune, are listed below:

pT0Ref is the parameter that regularizes the partonic QCD cross section and sets the lower scale of the MPI contribution;

expPow is the parameter in the overlap function, $\exp(-b^{\text{expPow}})$, which describes the convolution of the matter distributions of the two incoming hadrons.

The tuneable parameters of the QCD-inspired CR model that are considered in the fit, are:

m0 variable used in the string length measure λ , which determines whether a possible reconnection is actually favoured;

junctionCorrection multiplicative correction for the junction formation, applied to the m0 parameter;

timeDilationPar parameter that describes the time dilation and disallows color reconnection between strings that are not in causal contact.

For the QCD-inspired tune, the parameter values proposed in [14] for the hadronization are used as the initial values. The parameters of the fit as well as the ranges for each of them are shown in Tab. 5.1.

Table 5.1.: Multi Parton Interaction and color reconnection parameter ranges used in the fitting procedure.

UE parameters	min	max
MultipartonInteractions:pT0Ref	1.0	3.0
MultipartonInteractions:expPow	1.3	4.0
QCD-inspired model		
ColourReconnection:m0	0.2	4.5
ColourReconnection:junctionCorrection	0.05	9.0
ColourReconnection:timeDilationPar	0	90

The fits are performed using the PROFESSOR 1.4.0 tuning software¹ [132] and RIVET 2.4.0² [134]. Around 150 different choices of parameters are taken into account to construct a random scan in the parameter space, where for each set

¹The Professor tuning system parametrizes a generator's response to parameter shifts on a bin-by-bin basis [131].

²The RIVET analysis toolkit is a library of experimental analyses, as well as a tool for the calculation of physical observables from an event record. The generators interfaced to RIVET produce Monte Carlo datasets [133] which are used by Professor to perform the tune.

of parameter choices, 2 million events were generated. This is done in order to obtain in each considered bin a smaller statistical uncertainty for the MC predictions compared to the one from data. The MC to data agreement is checked by using bin-by-bin envelopes, which help to control the generator runs. An example of the uncertainties for the predictions from PROFESSOR for pseudorapidity η of charged hadrons and transverse N_{chg} density versus p_T of the leading track distributions are shown in Fig. 5.4, which demonstrates that different MC predictions describe the data well.

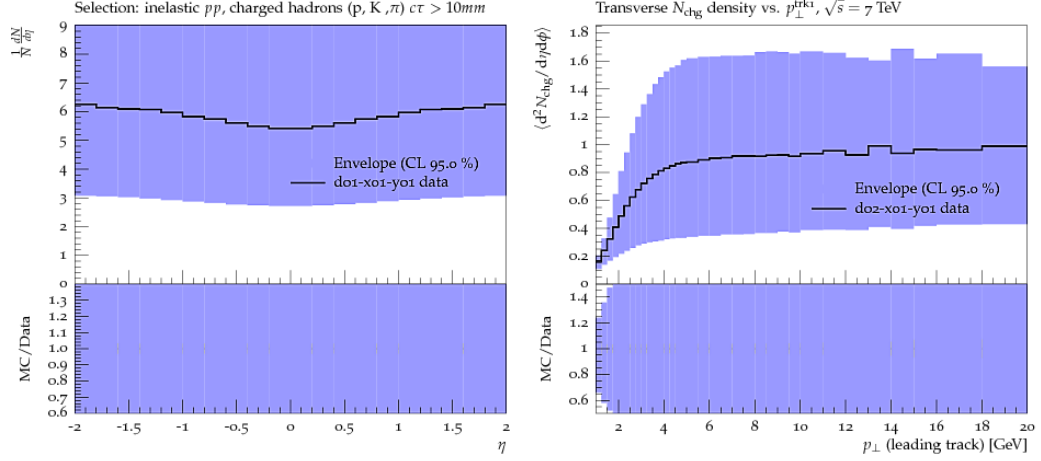


Figure 5.4.: An example of the uncertainties (blue area) for the predictions from PROFESSOR for pseudorapidity η of charged hadrons at 13 TeV (left) and transverse N_{chg} density vs. $p_T^{\text{leadingTrk.}}$ at 7 TeV (right).

After building the scan in the parameter space, the values in each bin b of the used observables are interpolated by PROFESSOR in so called eigentunes as a function of the vector of the parameters, $\vec{p}$. This interpolation is done by a set of third-order³ polynomial functions $f^b(\vec{p})$, that provide a good compromise between fit performance and the required number of simulated samples. After the $f^b(\vec{p})$ functions are defined, the most probable parameter value is found by minimizing the χ^2 function:

$$\chi^2(\vec{p}) = \sum_O w_0 \sum_{b \in O} \frac{(f^b(\vec{p}) - R_b)^2}{\Delta_b^2}, \quad (5.4)$$

where O are the sector bins for each observable used in the fit, the data value for each bin b is given by R_b , and the corresponding total bin uncertainty is Δ_b . The statistical uncertainty in the MC samples is neglected. The extracted parameter values for the CUETP8M2-CR1 fits to the data at different center-of-mass energies as well as the values for the goodness of the fit are shown in Tab. 5.2.

³The degree of the polynomial does not influence the tune result.

Table 5.2.: The parameters obtained during the tuning of the PYTHIA 8.226 event generator. The QCD-inspired color reconnection model is used and the results are compared to the parameters of the tune CUETP8M2 [110], which uses the MPI-based color reconnection model. Parameters marked with * were kept fixed while fitting. For the QCD-inspired color reconnection model, the default values implemented in PYTHIA 8.226 are indicated in parentheses, where applicable. The goodness of fit divided by the number of degrees of freedom is also shown. This information is given as supplementary material in [10].

Parameters	MPI-based CUETP8M2	QCD-inspired CUETP8M2-CR1
MultipartonInteractions: pT0Ref	2.20	2.17
MultipartonInteractions: expPow	1.60	1.31
MultipartonInteractions: ecmRef	7000	7000*
MultipartonInteractions: ecmPow	0.25	0.25*
ColourReconnection: range	6.59	-
ColourReconnection: junctionCorrection	-	0.12 (1.20)
ColourReconnection: timeDilationPar	-	15.9 (0.18)
ColourReconnection: m0	-	1.2 (0.3)
PDF set	NNPDF30_LO [83]	NNPDF30_LO [83]
SpaceShower:alphaSvalue	0.1108*	0.1108*
Goodness of fit/dof	1.89 [110]	1.06

5.4. Results and tune validation

According to the obtained parameters values for the CUETP8M2-CR1 tune that are shown in Tab. 5.2, an increase of the mass and the time dilation parameters is observed with respect to the default tune. However, the parameter for the junction correction becomes very close to 0. Changes in parameter values are also obtained for both the **PT0REF** and **EXPPOW** parameter in the new tune CUETP8M2-CR1 with respect to CUETP8M2, while using the same PDF set. The goodness of fit value is found to be close to 1, demonstrating how well observables that are sensitive to UE contributions can be described with the considered CR model.

In the following, a set of validation plots is shown for observables measured at a center-of-mass energy of 7 and 13 TeV. In each figure, black points denotes the corresponding data results, which are compared to predictions obtained from the PYTHIA 8 event generator with the CUETP8M2 tune (red), CUETP8M2-CR1 tune (blue), and the default tune (green). The default tune is obtained by using the parameter values proposed by the authors [14] of the QCD-inspired CR model:

- **JUNCTIONCORRECTION** = 1.2,
- **TIMEDILATIONPAR** = 0.18,
- **M0** = 2.15.

The panel under each figure shows the ratio between each prediction and the measurement.

The developed QCD-inspired tune show a good agreement in general with the measurements at 13 TeV. The UE observables, comprising the charged particle multiplicity and average p_T sum in the TransMIN and TransMAX regions measured by CMS at 13 TeV [135] are shown in Fig. 5.5 and are well described by all considered tunes in the plateau region. The discrepancy in the rising part of the spectrum, in the region $p_T^{max} < 3$ GeV, arises due to diffractive contributions. However, the optimization of these components is out of scope for this study.

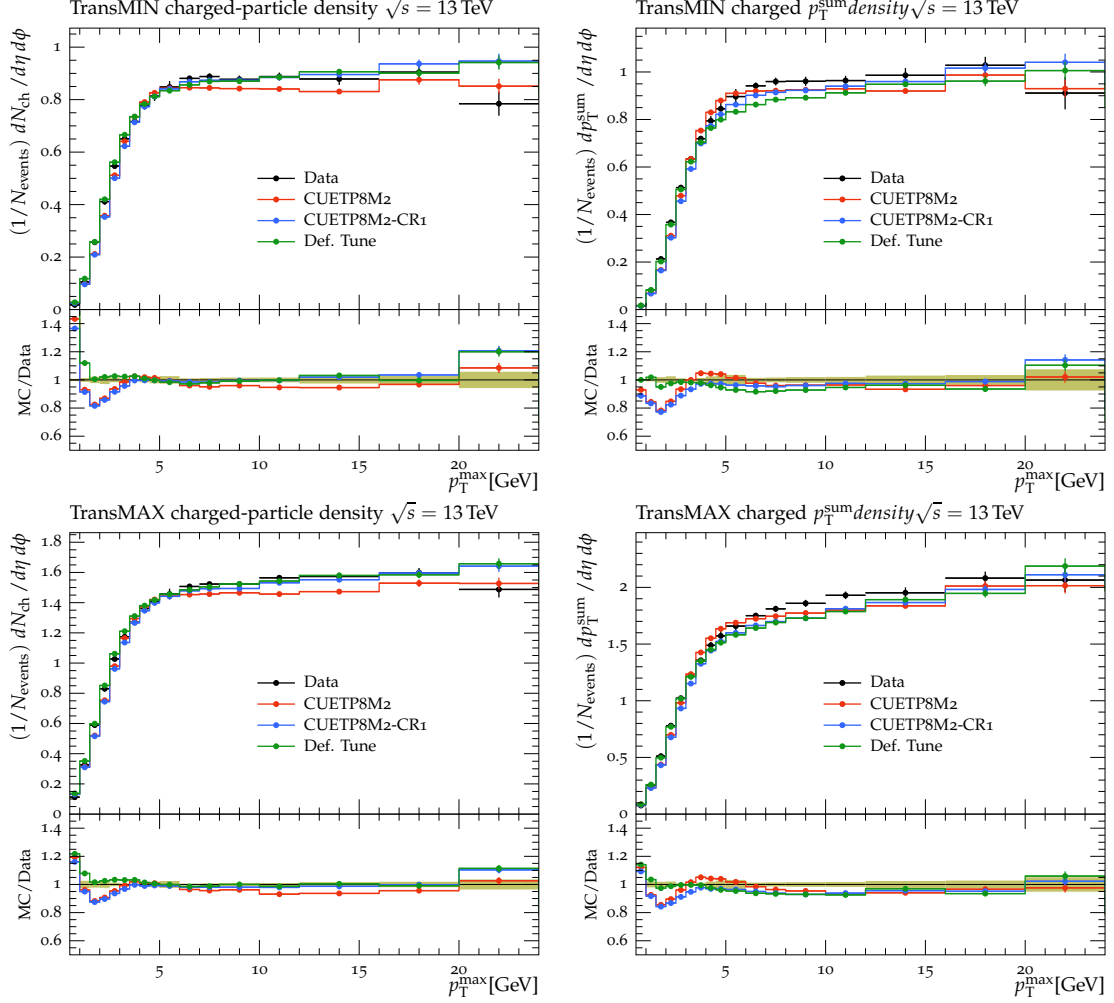


Figure 5.5.: Charged particle multiplicity (left) and p_T^{sum} (right) in the TransMIN (up) and TransMAX (down) regions, measured by CMS at 13 TeV [135]. The measurements are compared to predictions from the tunes CUETP8M2, CUETP8M2-CR1 and default tune. The bottom panels show the ratios between each Monte Carlo prediction and the data.

Observables that are sensitive to the softer part of the MPI spectrum, such as the charged particle multiplicity as a function of pseudorapidity in the inclusive selection are also very well described by predictions using the new CMS tunes, as shown in the comparison with data measured by CMS at 13 TeV [130,135] and shown

in Fig. 5.6 (left). Observables that are measured at large pseudorapidities are also well reproduced by the CMS tunes, as can be seen in the forward energy flow in two different selections, measured by CMS at 13 TeV [136], and presented in Fig. 5.6 (right).

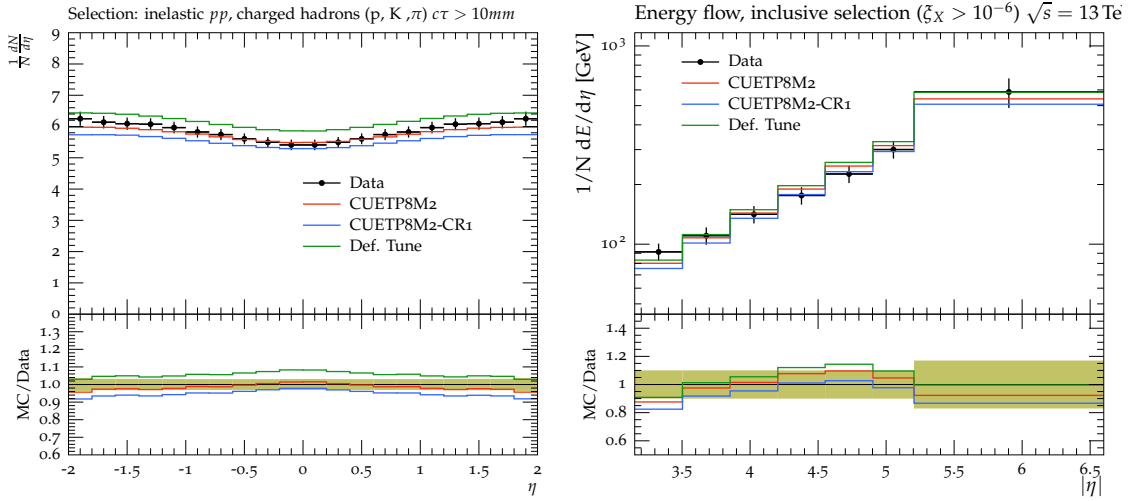


Figure 5.6.: Charged particle multiplicity as a function of pseudorapidity after the inclusive selection [130,135] (left) and forward energy flow as a function of pseudorapidity in minimum-bias events [136] (right) measured by CMS at 13 TeV. The measurements are compared to predictions from the tunes CUETP8M2, CUETP8M2-CR1 and the default tune. The bottom panels show the ratios between each Monte Carlo prediction and the data.

Measurements at 7 TeV are also very well described when applying the new CMS tunes. Figure 5.7 shows the multiplicity and transverse momentum sum of the charged particles as a function of p_T^{max} as measured by ATLAS at 7 TeV [127]. A similar level of agreement as observed at 13 TeV is achieved at 7 TeV with PYTHIA 8 using the new tunes.

Figure 5.8 shows the inclusive charged particle multiplicity and the multiplicity of strange particles, i.e., Λ baryons as a function of pseudorapidity, measured at 7 TeV by CMS [137,138]. The inclusive charged particle multiplicity is well reproduced by the predictions. The strange particle spectra are underestimated by the predictions of all new tunes and no significant improvement is observed between predictions using CUETP8M2 and the new CR model tune. In [14], it was shown that new color reconnection models might be beneficial for the description of ratios of strange particle multiplicity ratio observables, e.g., the production cross section $\sigma(\Lambda)$ divided by $\sigma(K_s^0)$. Figure 5.9 shows that this is not true for the multiplicity itself, but it has been shown that this might be fixed by using different hadronization models [139,140].

A crucial test for the performance of reliable UE tunes, and of the CR simula-

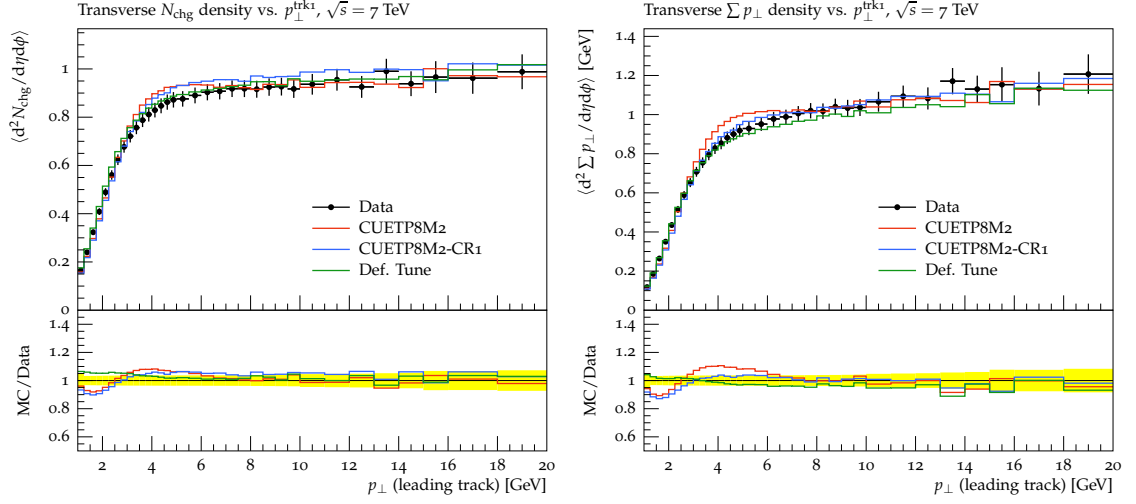


Figure 5.7.: Charged particle multiplicity (left) and $p_{\text{T}}^{\text{sum}}$ (right) in the transverse region, measured by ATLAS at 7 TeV [127] as a function of p_{T} the leading charged particle. The measurements are compared to the predictions from the tunes CUETP8M2, CUETP8M2-CR1 and the default tune. The lower panels show the ratios between each Monte Carlo prediction and the data.

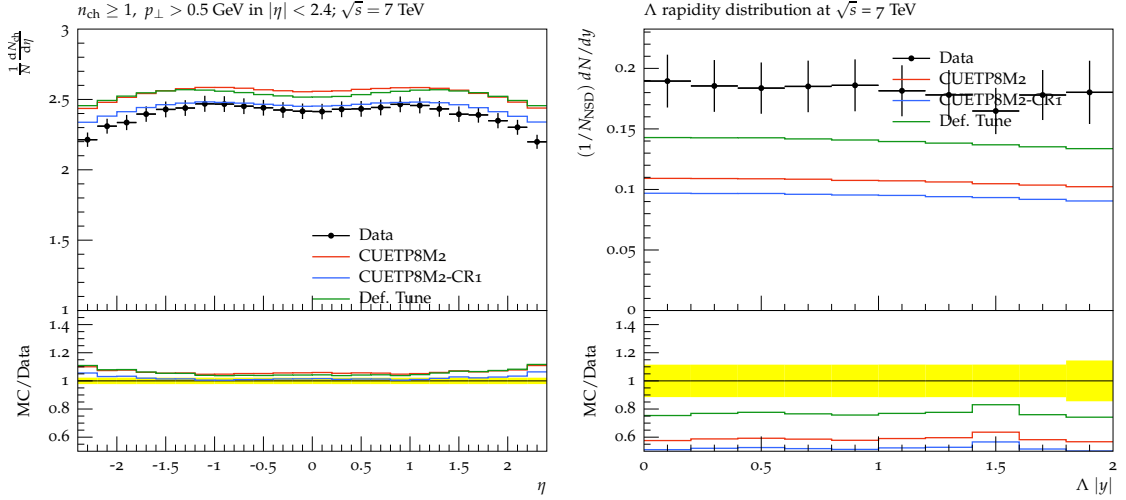


Figure 5.8.: Charged particle multiplicity as a function of pseudorapidity, for inclusive particles (left) and Λ baryons (right) measured by CMS at 7 TeV [137, 138]. The measurements are compared to predictions from the tunes CUETP8M2, CUETP8M2-CR1 and the default tune. The bottom panels show the ratios between each Monte Carlo prediction and the data.

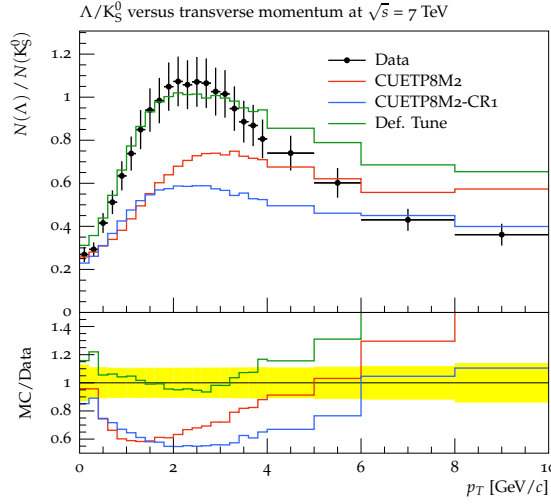


Figure 5.9.: Charged particle multiplicity as a function of p_T , for the ratio between the number of observed Λ baryons and the number of K_s^0 mesons measured by CMS at 7 TeV [138]. The measurements are compared to predictions from the tunes CUETP8M2, CUETP8M2-CR1 and default tune. The bottom panels show the ratios between each Monte Carlo prediction and the data.

tion in particular, is the description of the average transverse momentum of charged particles as a function of the charged particle multiplicity. Comparisons to these observables, measured by CMS at 7 TeV [141], are shown in Figure 5.10. These distributions demonstrate the crucial effect of the CR model and their tuning implementation in the MC. All CR tunes provide a good description of the measurements, confirming the accuracy of the parameters obtained for the new CR models. Similar observables measured at 13 TeV by the ATLAS experiment [128] are shown in Figure A.29.

Additional observables sensitive to the used CR model are given in Appendix A.7.

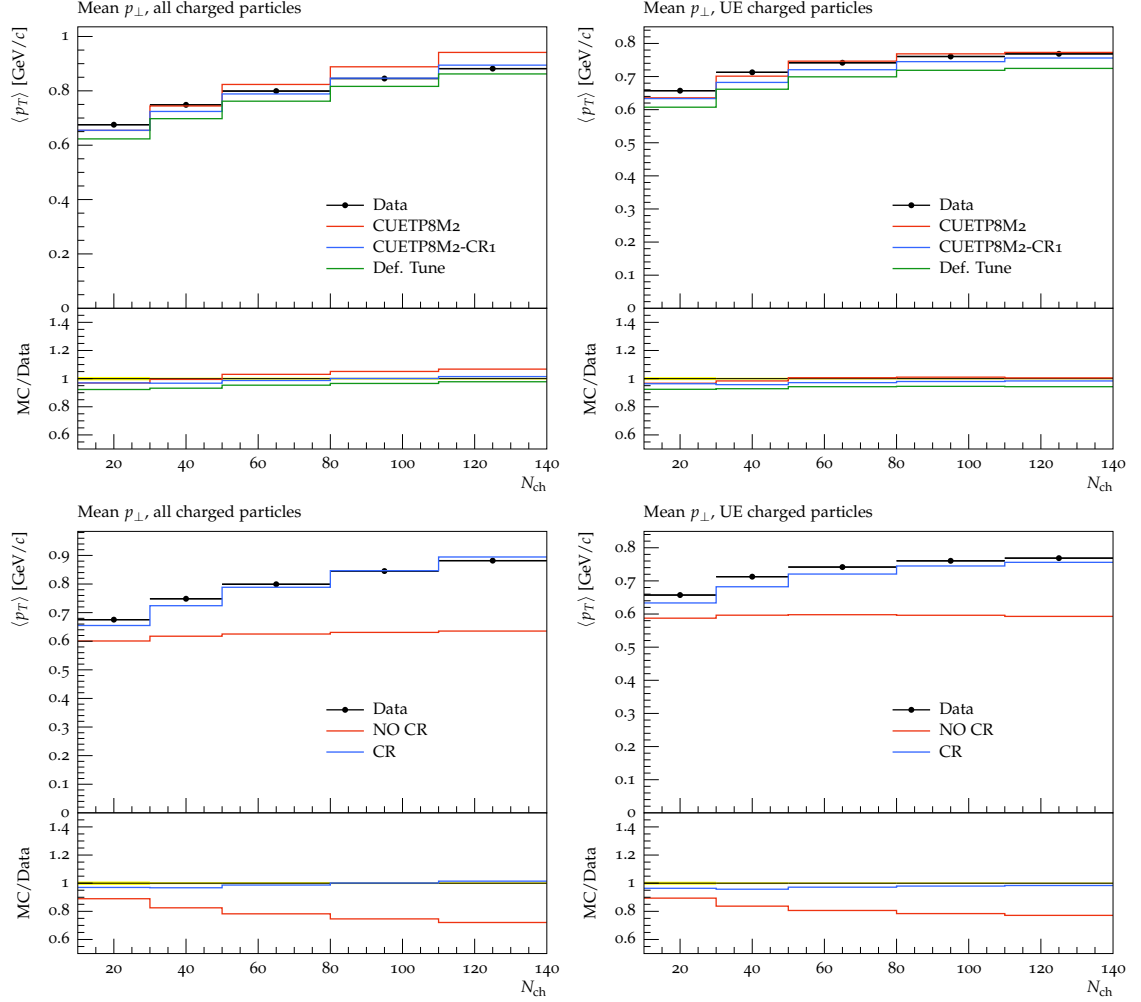


Figure 5.10.: Average p_T of the charged particles measured by CMS at 7 TeV [141] for all charged particles (left column) and for UE charged particles (right column). The corresponding observables are shown with (upper row) and without (lower row) CR model. The measurements are compared to predictions from the tunes CUETP8M2, CUETP8M2-CR1 and the default tune. The lower row shows in blue the results for the CUETP8M2-CR1, while in red the results without CR tune applied are shown. The bottom panels show the ratios between each Monte Carlo prediction and the data.

5.5. Summary

This study presents a development of a new tune for the PYTHIA 8 Monte Carlo event generator based on CMS and ATLAS data, which uses the QCD-inspired color reconnection (CR) model [14]. This new tune uses the NNPDF 3.0 PDF fitted with $\alpha_S(m_Z) = 0.130$ and is called CUETP8M2-CR1. MC predictions applying this tune are able to describe underlying event observables as well as charged particle multiplicities measured at 7 and 13 TeV, as shown in Fig. 5.11. It is found that none of the color reconnection models is able to predict the strange-particle yields measured at 7 TeV. The corresponding parameters for the developed QCD-inspired CR tune are available as supplementary materials in [10] where it is called “Pythia v8.226 QCD inspired”.

The derived new CR tune together with one for the “gluon move” CR model [125], shown in Fig. 5.11, allow a more reliable estimation of the color reconnection uncertainty by comparing the results for each of the color reconnection models. Its implementation for top quark mass measurements is discussed in Section 8.5.2.

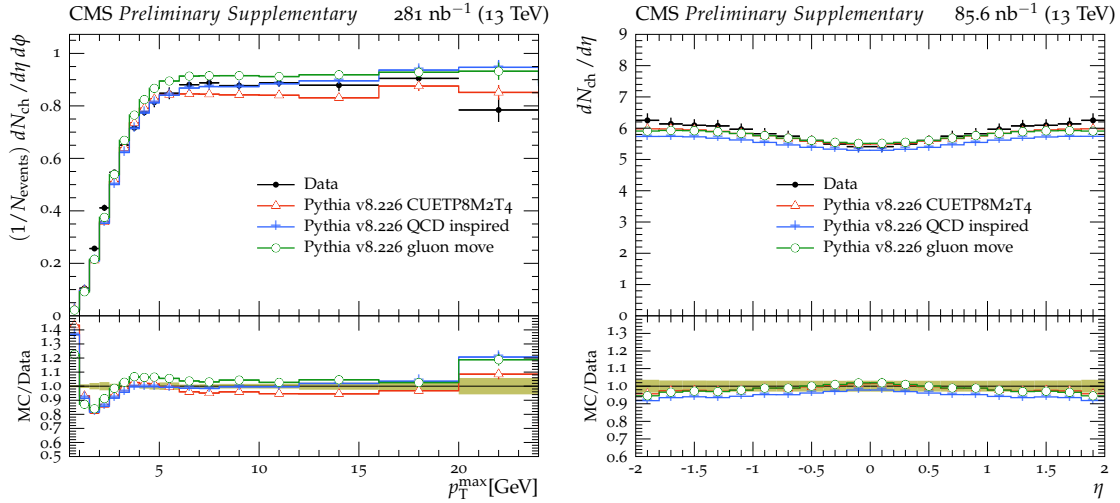


Figure 5.11.: CMS data [126] at $\sqrt{s} = 13$ TeV for charged particle densities with $p_T > 0.5$ GeV and $|\eta| < 0.8$ in the TransMIN region defined by the leading charged particle, as a function of the transverse momentum of the leading charged-particle p_T^{max} (left). CMS data [130] at $\sqrt{s} = 13$ TeV for the charged-particle pseudorapidity distribution, $dN_{\text{ch}}/d\eta$, in inelastic proton-proton collisions (right). The data are compared to predictions by PYTHIA 8.226 and Tune CUETP8M2T4, and predictions by PYTHIA 8.226 with the two tunes using the QCD-inspired and gluon-move color reconnection models, respectively. The ratios of MC predictions to data are given below each panel [10].

6. Object reconstruction and identification

In this chapter, the reconstruction of basic physics objects, which are produced during the high-energy proton-proton collisions and recorded by the CMS detector, is described. These objects correspond to single visible particles, such as leptons, or an ensemble of hadrons named jets. In order to reconstruct these particles, it is important to record precisely the information about their momentum, energy and trajectory, as it is shown in Chapter 3. This information is used afterwards by the *Particle Flow* [142] algorithm (PF), which is designed to identify individually each produced particle and hence helps reconstructing the collision event.

The reconstruction of the physics objects, namely muons, electrons and jets, that are used in the selection of the μ/e +jets events to perform the top quark mass measurement, as well as the dijets events to provide the residual correction, are also discussed in this chapter.

6.1. The CMS Particle Flow algorithm

The Particle Flow algorithm is a successful and efficient tool to identify visible particles produced in proton-proton collisions at the LHC and recorded with the CMS detector. The physics objects reconstructed by the PF algorithm, *particle flow candidates*, are used to explore the physics process studied in many analyses. First, the readout information from the different sub-detectors is translated into a description of the physical event, as shown schematically in Fig. 6.1. The PF algorithm attempts to use the characteristic signatures that various particles leave in the different parts of the CMS detector, where, e.g., charged particles, due to the high magnetic field and the large tracker cylindrical volume, are well separated. Figure 6.2 illustrates the signals produced in the detector by the different visible particles passing through such as muons, electrons, charged hadrons (e.g., pions), neutral hadrons (e.g., neutrons), and photons. The combination of sub-detector signatures is used by the PF method to determine the type of the particle. The algorithm starts with the reconstruction of charged objects, using track information from the inner tracking detector and the calorimeter cluster positions. In the following, the reconstruction of each category of particles is described.

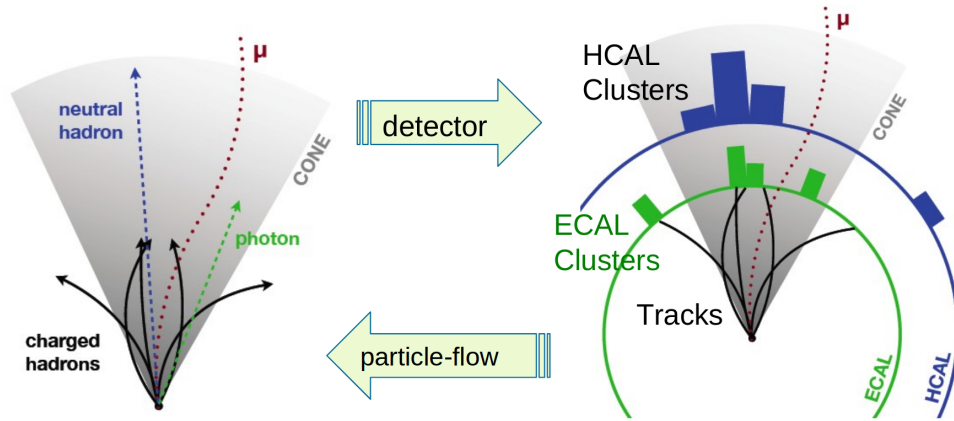


Figure 6.1.: Components of a “traditional” particle physics experiment. Each particle type has its own signature in the detector [143], which is exploited by the Particle Flow algorithm.

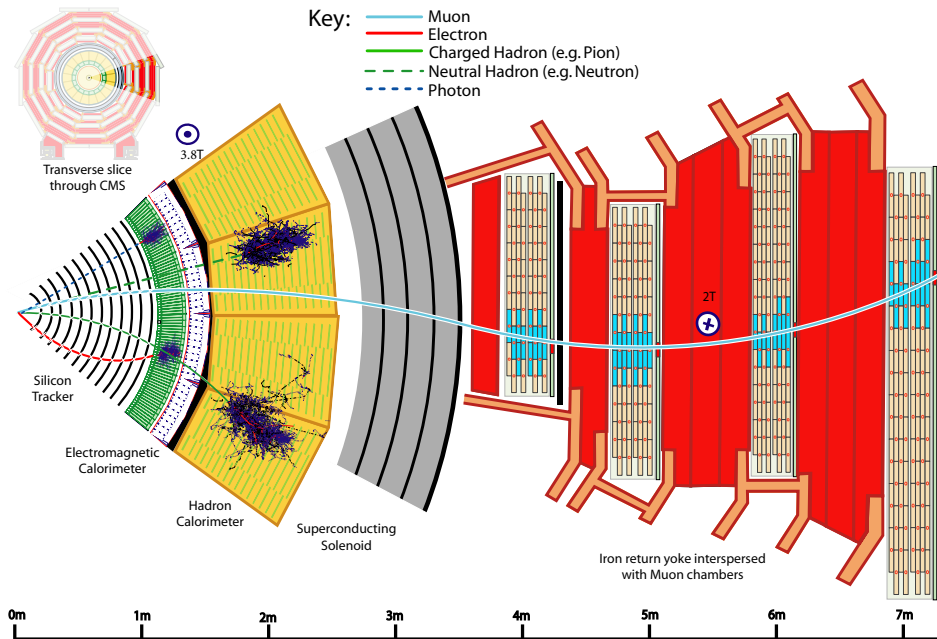


Figure 6.2.: A sketch of the specific particle interactions in a transverse slice of the CMS detector, from the beam interaction region to the muon detector [142].

Muons (μ): To ensure that, the tracks from muons do not bias the charged hadrons identification, the muons are reconstructed first, using the hit information in the inner and outer (muon) tracker systems. A negligible muon energy is expected to be deposited in the calorimeters. The reconstructed muons are then grouped into two categories [71] *standalone muons*, where information only from segments and hits in the muon chambers are used for the identification, and *global muons*, where information from the tracker and the muon detectors is used. Thus, a global muon track is fitted by the Kalman-filter (KF) technique [144], that combines the hits from standalone muon track and the expected muon tracker track.

Electrons (e): The electrons leave hits in the tracker system and are stopped in the ECAL depositing all their energy. Ideally, electron showers deposit 97% of their energy in the 5×5 ECAL crystals. However, the presence of the tracker in front of the ECAL calorimeter results in a frequent energy loss through bremsstrahlung, photon emission from the electron interacting with the tracker material. Additionally, the azimuthal bending of the electron in the strong magnetic field, leads to a spread of its energy in ϕ -direction before reaching the calorimeter. Therefore, a *Gaussian Sum Filter* [145] (GSF), which is an extended KF [146] charged particle tracking technique, is used. The GSF algorithm groups clusters from the electron and corresponding possible bremsstrahlung photons into one *supercluster* (SC), which is measured around the electron direction in a relatively small region of η and a larger region of ϕ . Additionally, to perform a better electron identification and momentum estimation, the fraction of its energy lost due to the bremsstrahlung, f_{brem} , is calculated. To define f_{brem} the momentum at the point of closest approach to the interaction point, p_{in} , and the momentum extrapolated to the surface of the ECAL from the track at the exit of the tracker, p_{out} , are used: $f_{\text{brem}} = (p_{\text{in}} - p_{\text{out}})/p_{\text{in}}$.

Charged hadrons: After the tracks used for the muon and electron identification are removed, the remaining tracks are extrapolated using mainly the HCAL calorimeter information, namely the boundaries of one of the several calorimeter clusters, where the tracks are matched to the boundaries of these clusters. This set of tracks and associated cluster(s) are used to compose a charged hadron, for which the momentum calculated from the tracking information is compared to the energy estimated from the calorimeter. If momentum and energy are compatible, the measurements from both detector components are combined.

Photons and neutral hadrons: After all tracks have been associated to charged particles, the reconstruction of photons and neutral hadrons is possible by only looking into the energy deposit in the ECAL and in the HCAL, where

the photons produce the signal in ECAL only, while the neutral hadrons in both ECAL and HCAL. Additionally, neutral hadrons can be identified if the momentum of the set of tracks calculated from the tracking information and energy of the associated cluster(s) estimated from the calorimeter are not compatible.

Once all the deposits are assigned to the corresponding particle flow candidates, the determination of jets, i.e. ensembles of particles produced during partons hadronisation, and missing transverse energy, i.e. the negative vectorial sum of all PF particles, is performed. The jets, reconstructed by clustering the four-momenta of the PF candidates as described in Section 6.5, are named *particle flow (PF) jets* and consist mostly of charged hadrons (65 %) and photons (25 %) recorded by the high precision tracker and the ECAL, the remaining part comes from neutral hadrons (10 %) that are measured in the HCAL. The performance of jet clustering using the PF algorithm [142] shows a very great improvement for the jet momentum (about 95 % of the reconstructed p_T) and resolution (improvement at low p_T) with respect to calorimeter jets, where only information from calorimeter towers that consist of one or more HCAL cells and corresponding ECAL crystals are used. The *particle flow (PF) missing transverse energy*, E_T^{miss} , is defined as the negative sum off all four momenta of the reconstructed PF candidates ($\vec{E}_T^{\text{miss}} = | - \sum_{i=1}^{N_{\text{PF candidates}}} \vec{p}_{T,i} |$) and is used to describe any weakly interacting neutral particle, i.e. neutrino, that does not deposit any energy in the detector.

6.2. Primary vertex reconstruction

For each event, several primary vertex candidates are reconstructed because of the spread of the protons in the proton bunches. To find a vertex candidate, the tracks of the particles found by the PF algorithm that appear to originate from the same vertex are clustered by a *deterministic annealing* (DA) algorithm [147]. At CMS the DA algorithm uses the distance of the tracks to the interaction point in z direction to find as many vertex candidates as possible without splitting true vertices. The selected vertex candidates with at least two associated tracks are further analyzed through the *adaptive vertex reconstruction* (AVR) algorithm [148] based on the *adaptive vertex fitter* (AVF) [149], which assigns weights (or probabilities) w_i to each track i [150] to be a part of a certain vertex candidate. A minimal weight of 0.5 is requested for each track-vertex connection. The quality of the vertex fit is defined by the the number of degrees of freedom n_{dof} ,

$$n_{\text{dof}} = -3 + 2 \sum_{i=1}^{N_{\text{tr}}} w_i, \quad (6.1)$$

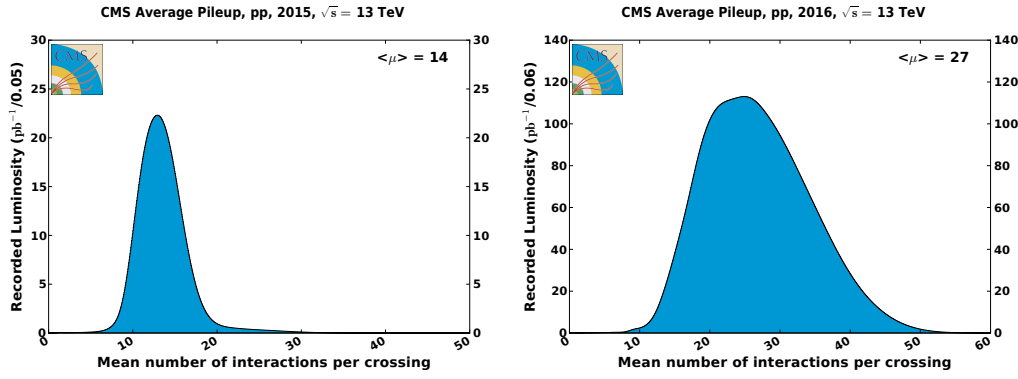


Figure 6.3.: Mean number of interactions per bunch crossing (pileup) for the 2015 (left) and 2016 (right) pp run at 13 TeV [59].

where N_{tr} is a number of the tracks assigned to the vertex candidate.

The cluster with the the highest p_{T}^2 sum of the associated tracks and $n_{\text{dof}} > 4$ is selected as a primary vertex (PV). A minimum of four tracks have to be assigned to the PV and the position of the PV is required to be within 24 cm in longitudinal and 2 cm in transverse direction from the nominal interaction point (beam spot). Other reconstructed vertices can correspond to secondary vertices or to pileup. The *pileup* (PU) describes an additional soft activity with p_{T} of a few GeV in the event that does not originate from the hard scattering process. The *in-time pileup* and *out-of-time pileup* can be distinguished. The in-time pileup characterized the additional pp interactions happen within the same bunch crossing, while out-of-time pileup describes the pp interactions from the previous bunch crossing, which arises due to the detector technology. The distributions of the number of pileup interactions during the 2015 and the 2016 LHC run period, are shown in Fig. 6.3. The numbers of the proton-proton interactions increased and thus the pileup is also increased. This is due to the increase in instantaneous luminosity.

In order to correct for the pileup contribution in the event, charged hadrons coming from pileup can be discard by the *charged hadron subtraction* (CHS) [151] method. However, CHS does not correct for whole contributions from the PU, more detail on the pileup corrections is given in Section 6.5.

6.3. Muon selection

In this section general selection properties for the signal muon in the $t\bar{t}$ muon+jets final state will be discussed. The muons, reconstructed using the PF algorithm described in Section 6.1, have to proceed an additional set of selection criteria. For the current analysis, muons need to meet *tight identification (ID)* and *tight isolation* requirements [152]. The tight ID criterias are:

- the muon has been reconstructed by the PF algorithm and measured in both

tracker and muon system, i.e. is a global muon.

- each muon candidate has to have valid hits in at least 5 tracker layers, at least 1 pixel layer, to ensure a good track quality and suppress muons from in flight decays.
- at least 1 muon chamber hit has to be detected in the muon system.
- the global muon track fit has to have $\chi^2/\text{ndf} < 10$, to suppress both muons in flight decay and hadronic punch-through, where the hadron shower remnants penetrating through the calorimeters can reach the muon system.
- in the tracker system, the a transverse impact parameter $|d_{xy}|$ of the muon track has to be less than 0.2 cm with regard to the beam spot in order to suppress cosmic muons.
- the longitudinal impact parameter with respect to the primary vertex of the corresponding muon track has to be $|d_z| < 0.5$ cm, in order to suppress muons from in flight decays and tracks from pileup.

Isolation

In addition to the muon identification requirements, isolation criteria are applied. This ensures that the selected muon is a decay product of, e.g., a W boson, and is not produced from in flight decays or during heavy-flavor (charm and beauty hadron) decays. In this analysis *particle-flow relative isolation* [71] $I_{\text{PF}}^{\text{rel}}$ is used, where the isolation of the muon is based on the p_T sum of all particle flow neutral hadrons (h^0), photons (γ) and charged hadrons ($h^\pm$) associated to the PV as well as to pileup vertices reconstructed within a cone of $\Delta R \leq 0.4$ around the muon momentum:

$$I_{\text{PF}}^{\text{rel}} = \frac{1}{p_{\text{T}}^{\text{lepton}}} \cdot \sum [p_{\text{T}}^{h^\pm \text{ from PV}} + \max(p_{\text{T}}^{h^0} + p_{\text{T}}^{\gamma} - \beta \cdot p_{\text{T}}^{h^\pm \text{ from PU}}, 0)], \quad (6.2)$$

where the pileup contribution is corrected by the $\delta\beta$ -correction method $\beta \cdot p_{\text{T}}^{\text{ch. had from PU}}$, with β scale factor of 0.5. Each muon candidate has to fulfill a tight isolation, which corresponds to the $I_{\text{PF}}^{\text{rel}} < 0.15$. The corresponding isolation efficiencies as a function of pseudorapidity η as well as transverse momentum p_T are shown for data and MC in Fig. 6.4. The MC scale factors [153] obtained from the comparison of the efficiencies in data and MC are found to be slightly different from 1, which illustrate a very good simulation of the muon candidates.

Additional muon veto

In order to suppress the SM background contributions, the events with additional isolated muons (*loose ID* [152]) with the following properties:

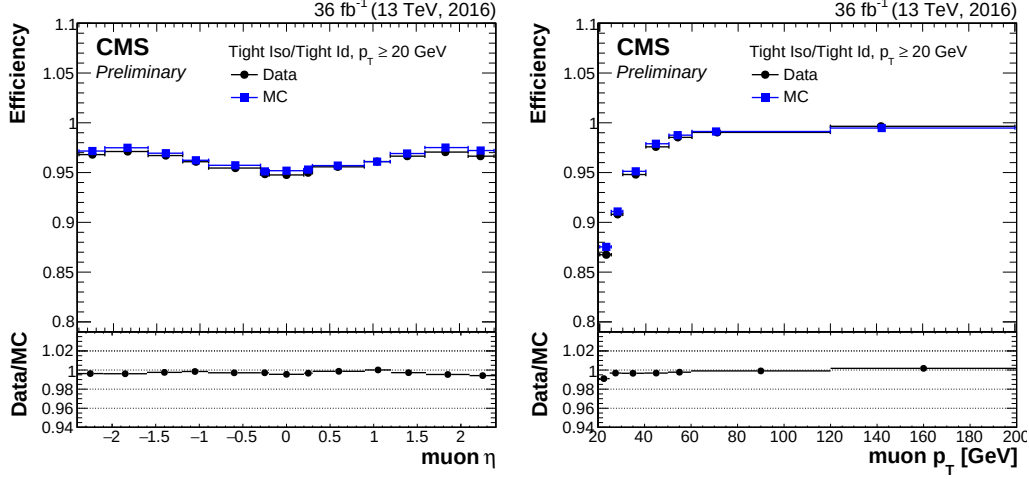


Figure 6.4.: The muon tight isolation efficiency vs η (left) and p_T (right) for MC and data and the corresponding scale factor [153]. The black markers denote data and blue markers denote MC.

- the muon candidate is reconstructed with the *Global* or *Tracker* [144] muon algorithm;
- the muon has minimum transverse momentum of $p_T > 15$ GeV and an absolute value of the pseudorapidity $|\eta| < 2.4$;
- the particle-flow relative isolation of the muon is $I_{PF}^{rel} < 0.25$;

are rejected.

Single-muon trigger efficiency

Events are selected online with the single-muon triggers¹. They use information from the inner tracker and muon detectors. The online-reconstructed transverse momentum of the muon candidate is required to be greater than the pre-defined trigger threshold of 24 GeV. The trigger efficiencies are determined via the tag-and-probe method [154] with $Z \rightarrow \mu^+ \mu^-$ events.

The tag-and-probe method allows an extraction of muon tracking efficiency directly from decays of known resonances. The *tag* muon is required to fire the HLT_IsoMu24_v* or HLT_IsoTkMu24_v* and fulfill the tight selection criteria listed in Section 6.3 without applying the cut on the longitudinal distance to the primary vertex, but with $p_T > 24$ GeV and $|\eta| < 2.4$. The *probe* muon is then required to be oppositely-charged with PF isolation $\delta\beta/p_T < 0.15$ and fulfill tight ID selection. The ΔR between tag and probe muons is required to be larger than 0.3 in order to avoid effects from the correlation between two muons. The invariant mass of

¹ HLT_IsoMu24_v* or HLT_IsoTkMu24_v*

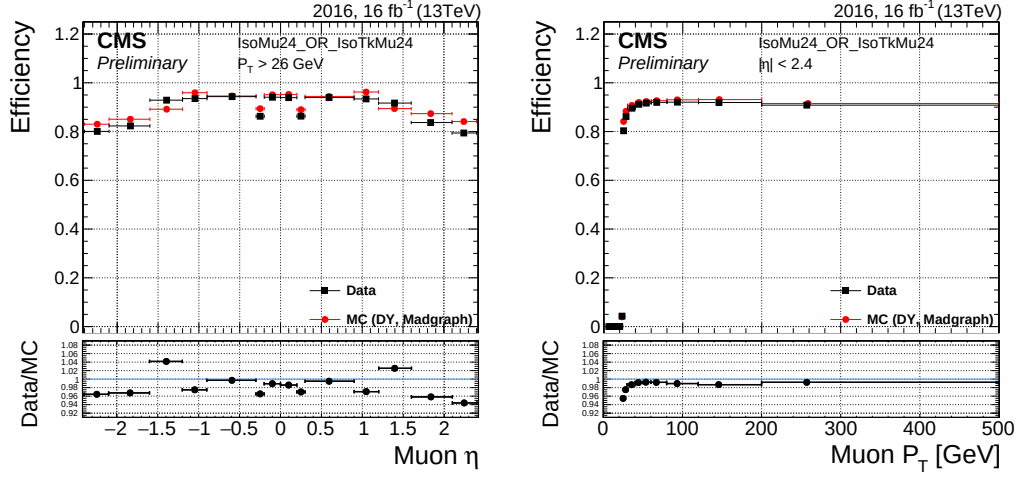


Figure 6.5.: Trigger efficiency as a function of η (left) and p_T (right) for the OR combination of the HLT_IsoMu24 and HLT_IsoTkMu24 with respect to the offline reconstructed muon passing identification and isolation requirements [155].

the muon pair has to be within a window of 15 GeV around the Z boson mass. The efficiencies are determined by applying selection criteria and by analyzing the trigger information to the *probe* muons. They are presented in Fig.6.5 as a function of pseudorapidity η and a transverse momentum p_T . The presented muon scale factors are applied as event weight depending on the muon transverse momentum and pseudorapidity.

6.4. Electron selection

In case of the electron+jets $t\bar{t}$ final state the presence of exactly one reconstructed electron is expected. Therefore in this analysis, the reconstructed particle flow electron candidate driven by association of a GSF track and a cluster in the ECAL, as described in Section 6.1, are used. In addition to the particle flow requirements, a stand-alone approach [156] is exploited, resulting in a better performance. In the next stage, the *cut-based* algorithm is applied, where electrons need to fulfill *tight identification (ID)* and *tight isolation* requirements. The algorithm selection criteria for the barrel, $|\eta_{SC}| \leq 1.479$, and endcap, $|\eta_{SC}| > 1.479$, regions listed in the followings.

Conversion rejection:

- each electron candidate has to pass a conversion veto².

²Since, photons have a very large probability to convert into e^+e^- , the probability of the electron object to originate from the reconstruction of the photon conversion electron is calculated. Thus, only events, where the electron candidate is considered as highly unlikely to originate from a conversion, are selected.

- the electron track trajectory has to have not more than two missing hits in different layers of the inner tracker system.

Electron ID conditions:

- for each electron candidate, the ratio of the HCAL tower energy behind the ECAL supercluster over the energy of the ECAL supercluster, H/E , has to be < 0.0414 (barrel), 0.0641 (endcap)
- the measure of the η dispersion of the ECAL supercluster, $\sigma_{\eta\eta}$, has to be less than 0.00998 (barrel), and 0.0292 (endcap), where $\sigma_{\eta\eta}^2 = \sum_i^{5 \times 5} w_i (\eta_i - \bar{\eta}_{5 \times 5})^2 / \sum_i^{5 \times 5} w_i$, with i running over the crystals, $\bar{\eta}_{5 \times 5}$ is the mean of the 5×5 crystal, and $w_i = \ln(E_i/E_{5 \times 5})$
- the absolute difference in pseudorapidity η between the ECAL supercluster position and the extrapolated position in the ECAL using the track parameters at the interaction vertex, $|\Delta\eta_{in}| = |\eta_{SC} - \eta^{\text{extrap}}|$, has to be less than 0.00308 (barrel) and 0.00605 (endcap).
- the absolute difference in ϕ between the ECAL supercluster position and the track ϕ extrapolated from the innermost track position to the position of closest approach in the ECAL supercluster, $|\Delta\phi_{in}| = |\phi_{SC} - \phi^{\text{extrap}}|$, has to be less than 0.0816 (barrel) and 0.0394 (endcap).
- the discrepancy of the energy measured in the ECAL, E_{SC} , and the momentum, p , measured in the tracker, of the electron candidate, $|\frac{1}{E_{SC}} - \frac{1}{p}|$, has to be less than 0.0129 (barrel) and 0.0129 (endcap). This helps to distinguish an electron-induced signal from a case of accidentally overlapping tracks and ECAL clusters.

Isolation requirements:

- the particle flow relative isolation of the electron candidate calculated within a cone of $\Delta R \leq 0.3$ around the electron momentum and corrected for the pileup contamination is given by

$$I_{\text{PF}}^{\text{rel,corr}} = \frac{1}{p_{\text{T}}^e} \cdot \sum [p_{\text{T}}^{h^\pm \text{ from PV}} + \max(p_{\text{T}}^{h^0} + p_{\text{T}}^\gamma - \rho A_{\text{eff}}, 0)] \quad (6.3)$$

and has to be less than 0.0588 (barrel) and 0.0571 (endcap). Here, the pileup correction is implemented through the ρ -effective area correction, $\rho \cdot A_{\text{eff}}$, where ρ is the average pileup energy density in the $\phi - \eta$ plane and A_{eff} is the effective area calculated in bins of pseudorapidity using the 90%-efficient contour approach [157].

In additions, electron candidates that passed the tight ID and isolation conditions listed previously have to fulfill the impact parameter requirements and an additional relative isolation requirement. The electron track has to fulfill the transverse impact parameter with regard to the beam spot of $d_{xy} < 0.05$ cm in the barrel region and $d_{xy} < 0.1$ cm in the endcap region, and the longitudinal impact parameter with respect to the primary vertex $d_z < 0.1$ cm in the barrel region and $d_z < 0.2$ cm in the endcap region. In case the electron supercluster's pseudorapidity is reconstructed in the transition region between ECAL barrel and endcap calorimeters, $1.4442 < |\eta_{SC}| < 1.5660$, the corresponding electron candidate is excluded from the analysis because of the very inefficient electron identification [158].

Additional electron veto

Since, in the $t\bar{t}$ electron+jets decay channel, exactly one isolated electron has to be selected, a veto for additional electrons is applied, which helps to suppress the background contribution from multi-lepton final states. Thus, an event fails the selection if additional isolated electrons with the following properties are reconstructed:

- a photon conversion veto is passed.
- electrons with $p_T > 20$ GeV, $|\eta| < 2.4$ are selected.
- the electron passes the cut-based electron ID with loose working poin:
 - $H/E < 0.0414$ (barrel), 0.0641 (endcap),
 - $\sigma_{\eta\eta} < 0.011$ (barrel), 0.0314 (endcap),
 - $|\Delta\eta_{in}| < 0.00477$ (barrel), 0.00868 (endcap),
 - $|\Delta\phi_{in}| < 0.222$ (barrel), 0.213 (endcap),
 - $|\frac{1}{E_{SC}} - \frac{1}{p}| < 0.241$ (barrel), 0.14 (endcap),
 - $I_{PF}^{rel,corr} < 0.0994$ (barrel), 0.107 (endcap).

Electron trigger efficiency

In order to select events with electron candidates the single-electron trigger³ paths are used. They select events with at least one electron with a transverse momentum larger than 32 GeV and an absolute value of pseudorapidity less than 2.1. The electron trigger efficiency is calculated via the tag-and-probe method in $Z \rightarrow e^+e^-$ events. The *tag* electron is required to fire the trigger and fulfill the following tight selection criteria: $p_T > 40$ GeV, $|\eta_{SC}| < 2.1$, tight cut-based ID and impact parameter cuts. The *probe* electron is required have $p_T > 25$ GeV, $|\eta_{sc}| < 2.1$, tight cut-based ID and impact parameter cuts. The invariant mass of the electron pair

³HLT_Ele32_eta2p1_WPTight_Gsf_v*

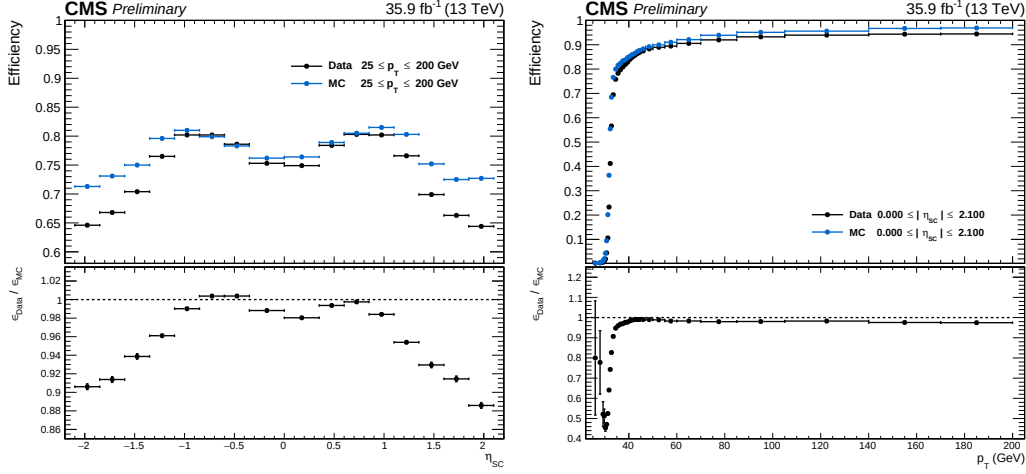


Figure 6.6.: Electron efficiencies and the corresponding scale factors as a function of the reconstructed electron pseudorapidity, η , (left), transverse momentum, p_T , (right) [160]. The results are shown for the HLT_Ele32_eta2p1_WPTight the single electron trigger. The plotted uncertainties are statistical only.

has to be in the range of 60 to 120 GeV [159]. Efficiencies are determined by applying selection criteria and the trigger decisions to the *probe* electrons. Figure.6.6 shows the efficiencies of the single electron trigger as a function of electron pseudorapidity and transverse momentum. To reach the same trigger efficiency in data and MC, scale factors (SF) are derived as shown on the lower plots of Fig.6.6, where the small deviations of SF in bins of $|\eta|$ are observed. The presented electron scale factors are applied as event weight per transverse momentum and pseudorapidity.

6.5. Jet reconstruction and calibration

As mentioned in Section 2, the strong interactions happening at high-energies are described in terms of the phase space of the quarks and gluons, often called partons. They carry color charge, thus cannot remain free due to color confinement that leads to fragmentation and hadronization of colored partons. As a result a collimated flux of color-neutral particles is produced, which is experimentally observed and is referred to as a *jet*.

Clusters of energy deposits in the calorimeters are used in the jet-finding algorithms. However, an imperfect detector response and the presence of particles originating from pileup interactions can bias the measured jet energy. Thus, a well defined jet clustering algorithm, and well measured constituents of the jet are required for the jet reconstruction. The main properties of jet algorithms are *infrared*- and *collinear-safety* (*IRC*) [161]. An algorithm is infrared-safe if the emission of a soft (infrared) gluon does not change the jets. It is collinear-safe when collinear

splitting, the splitting of a hard particle into two, does not change jets. Both IRC safety requirements are important, since the usage of IRC unsafe algorithms will result in altering the number and contents of the jets. Additionally, the size of the jet cone determines the sensitivity of the jet to soft radiation, the larger the jet radius, the more amount of pileup and underlying event will be captured by the jet. Another important point in the jet reconstruction are the constituents of the jet. In CMS particle-flow (PF) candidates, reconstructed with the PF algorithm, are used for the *PF jet* clustering and allow to reconstruct the jet constituents with a precision of around 90%, as mentioned in Section 6.1, making the reconstructed jet energy close to its true value. Additionally, charged hadrons associated with pileup vertices can be subtracted from the PF jets by using the *Charged Hadron Subtraction* (CHS) algorithm [151], which helps to remove pileup effects on jets, these jets are named *PF+CHS jets*. Taking into account the not perfect detector response and the efficiency of the jet clustering, additional jet energy corrections are estimated for the measured jet energy. Furthermore, identification criteria of the jets allow the identification of the jets flavor, e.g., the b-tag algorithm distinguishes the jets originating from b-hadrons [162].

Jet clustering algorithms

Two main types of jet clustering algorithms are distinguished: *cone algorithms* and *sequential clustering algorithms*. The cone algorithms can be split into a three most known technique among them: an iterative cone with progressive removal (IC-PR)(iterative version of [163]), an iterative cone with split-merge procedure (IC-SM) [164] and a seedless infrared-safe cone (SIScone) [165]. The most used sequential clustering algorithms are the k_T [166], anti- k_T - [167] and Cambridge-Aachen-(CA) [168,169] clustering algorithms. The cone algorithms define the jet as a dominant direction of energy flow, that is described in terms of a stable iterative cone (IC) with the fixed radius, R , in the η - ϕ plane. However, it generally suffers from IRC unsafety (except for SIScone). Therefore, CMS employs for the jet reconstruction the sequential clustering algorithms that are all IRC safe. They are implemented in a transverse momentum plane, assuming that particles have a small differences in p_T within the jet, that also results in a fluctuation of their area in η - ϕ plane.

In the sequential clustering algorithms two distance variables are introduced: the distance d_{ij} between two particles i and j carrying the transverse momentum $p_{T,i}$ and $p_{T,j}$:

$$d_{ij} = \min(p_{T,i}^{2b}, p_{T,j}^{2b}) \frac{\Delta R_{ij}^2}{R^2} \quad (6.4)$$

and the distance d_{iB} between the beam axis and the detected particle:

$$d_{iB} = p_{T,i}^{2b}, \quad (6.5)$$

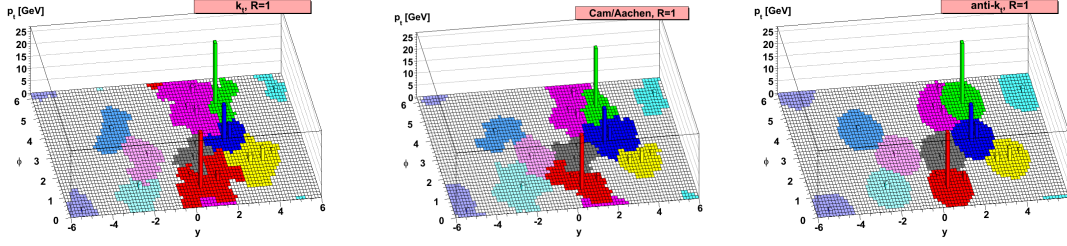


Figure 6.7.: A sample parton-level event (generated with HERWIG [170]), together with many random soft “ghosts”, clustered with tree different jets algorithms: k_T , Cambridge-Aachen, and anti- k_T illustrating the “active” catchment areas of the resulting hard jets [167].

where $\Delta R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$ is the distance in η - ϕ space, R is a radius parameter related to the area of the jet, usually between 0.4 - 0.8. The parameter b is a power of the energy relative to the geometrical scale, which regulate the behavior of the algorithm, $b = 0$ for the Cambridge-Aachen algorithm, $b = 1$ and $b = -1$ for the k_T and anti- k_T algorithm, respectively. The clustering algorithms start by finding two particles with the smallest distance parameter and merging them into one jet by adding their four-momenta. To do this, the minimum of the entire set of d_{ij}, d_{iB} are explored. Then, if the d_{iB} is the minimum, the object i is considered as a final jet and is excluded from the further clustering procedure. If the minimum is d_{ij} , the particles i and j are clustered together and treated as a single particle in the next iteration. The process is repeated until all objects are assigned to a jet.

Depending on the analysis a different choice for the algorithm can be made. The k_T algorithm first clusters the soft particles and is susceptible to the underlying event and pileup interactions. It is also good for the studying of its subjets. The anti- k_T algorithm first clusters the hard particles with high p_T . On the one hand, it is not so susceptible to the underlying event and pileup, on the other hand it has poor declustering, therefore, is not the best choice for study the jet substructure. Between the two discussed methods, one can place the Cambridge-Aachen algorithm, which is independent of the momentum and is not much susceptible to the underlying event and pileup contamination. Its advantage is the purely spatial character of the distance variables, that allows to study the jet substructure. An example of the usage of these main algorithms at CMS is shown in Fig. 6.7, which illustrates the shapes of jets with $R = 1$ in the η - ϕ plane.

In this thesis, jets reconstructed by the anti- k_T algorithm that is part of the family of the sequential clustering algorithms, with a radius parameter $R = 0.4$, namely AK4 jets, are used.

Signal jet selection

The reconstructed particle flow jets corrected for the pileup contamination by using the CHS method have to fulfill the next selection step. The jet candidates with $|\eta| \leq 2.4$ are required to pass the *loose identification requirements for the PF jets* [171] listed in the following;

- the energy fraction from the neutral hadrons in the total jet energy has to be smaller than 99%.
- the energy fraction from the neutral electromagnetic (EM) deposits in the total jet energy has to be smaller than 99%.
- the number of constituents of the jet has to be greater than one.
- the charged particle multiplicity of the jet has to be greater than zero .
- the energy fraction from charged hadrons in the total jet energy is required to be present.
- the energy fraction from electromagnetic (EM) deposits in the total jet energy must be smaller than 99%.

Jet energy correction

The measured energy of reconstructed jets does not naturally agree with the energy of the jets formed after the hadronization stage of the partons from the hard scattering, denoted as particle level jets. The differences in the measured energy or in the transverse momentum arise mainly due to the non-linear detector response. However, pileup, additional soft QCD radiation, or noise from the electronics in the detector can also influence the reconstructed energy of the jets. Therefore, it is important to correct the energy of the reconstructed jets to the particle level jets, since there are many physics analyses relying on the precision of the jet energy measurements.

In order to correct the reconstructed PF jet energy, an appropriate Jet Energy Correction (JEC) is needed. In the CMS collaboration the JEC is based on a factorized approach [13] as schematically shown in Fig 6.8. The corresponding corrections are mainly estimated by looking at the size of the shift of the *jet energy response*, which is calculated by measuring the ratio of the p_T of the reconstructed PF jet to the p_T of particle level jet. In the following, each step of the factorized approach for the JEC estimation will be described.

L1 pileup offset correction aims to correct for pileup effects. Since the pileup contribution in the event increases with increasing instantaneous luminosity, the

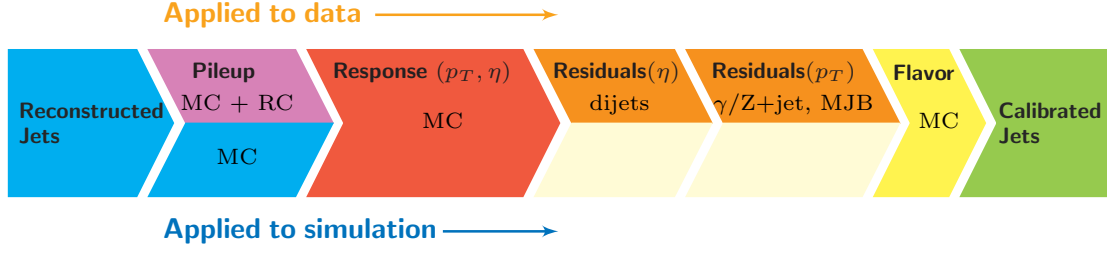


Figure 6.8.: Consecutive stages of JEC, for data and MC simulation. All corrections marked with MC are derived from simulation studies, RC stands for random cone, and MJB refers to the analysis of multijet events [13].

L1 offset correction is crucial in the LHC environment. Even though CHS⁴ which mitigates pileup effects, has been applied to the PF jets, there are still contributions from neutral particles and charged hadrons, which were not unambiguously linked to pileup vertices. These can influence the jet energy response. Moreover, the out-of-time pileup⁵ may also contribute to the jet reconstruction. In addition, instead of CHS the PUPPI⁶ algorithm [172] can be applied to the PF jets, which is an alternative way to mitigate pileup effects. It calculates the probability of each PF candidate in the event to originate from a pileup vertex and construct corresponding weights. The L1 offset turned out to be small and uniform versus PU for the PUPPI jets, therefore dedicated correction for the pileup offset is not applied to the PUPPI PF jets.

The offset energy density ρ [173,174], the sum of UE, pileup and noise, is used to correct for the pileup contribution. It can be defined in each event as the median of the energies calculated in a grid of $\eta - \phi$ cells. In Run 1 it was evaluated similarly by using the jet $p_{T,i}$ divided by the corresponding jet area⁷ A_i , $\rho = \text{median}(p_{T,i}/A_i)$ [175]. The median is chosen instead of mean for the calculation of ρ so the high p_T jets from the hard scattering do not bias the result.

The η - and p_T - dependent scale factors for the particle-level offset are calculated from MC the using same reconstructed events with and without pileup overlay. The reconstructed jets are then matched between these samples and the average difference in their p_T is calculated and denoted as the particle-level offset, and applied to both MC and data.

In order to correct for the remaining offset differences in data, additional scale

⁴Charged hadrons that arise from pileup vertices are not taken into account during CHS jets clustering, as described in Section 6.2.

⁵Out-of-time pileup arises from proton collisions before or after the collision of interest.

⁶**P**ileUp **P**er **P**article **I**dentification

⁷The *jet area* is computed by adding soft particles randomly distributed in the $(\eta - \phi)$ -plane (ghosts), while the *active jet area* denotes the fraction of the ghost particles clustered with the hard jet components.

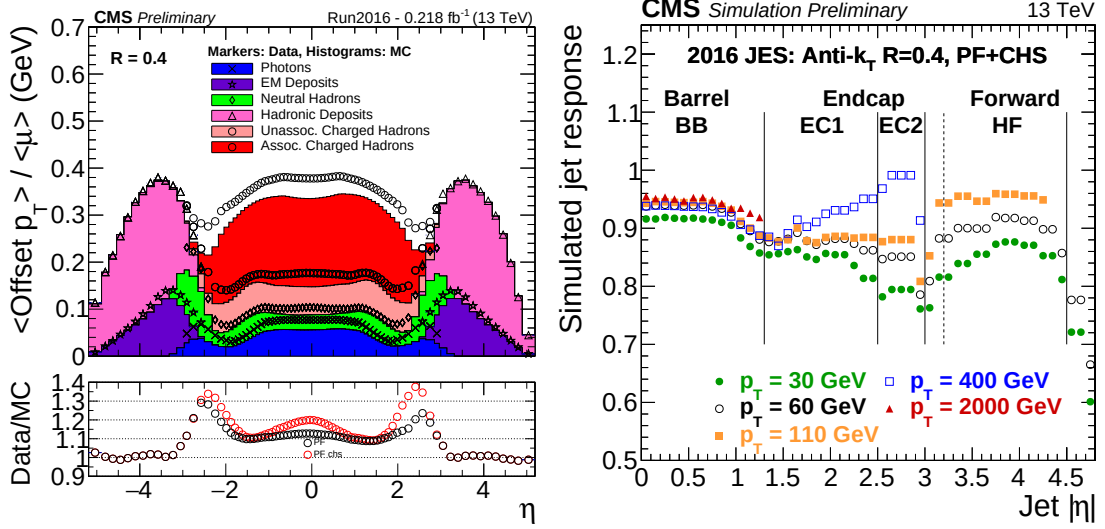


Figure 6.9.: Data/MC comparison for the average offset per additional pileup interaction. μ denotes an additional pileup interaction(left). Simulated jet response in different regions of the detector (right) [177]. The results are estimated for AK4 PF jets.

factors are derived using the random cone (RC) method [176]. The zero-bias data⁸ and simulation are used that do not include the events with the hard interactions. Thus, the p_T offset due to pileup is estimated as a function of η and ρ using the average of the sum of the p_T of the PF candidates in the random cones. Finally, the pileup corrections are calculated by taking the ratio of the average p_T offset in data and in MC. The resulting offset p_T corrections as a function of jet pseudorapidity and the corresponding data to MC ratio are given in Fig. 6.9 (left).

Simulated L2L3 Response corrections aim to correct the energy of the reconstructed jets, which have been corrected for pileup offset, to the energy of the particle level jets. It can be further subdivided into the *L2 Relative* and *L3 Absolute* corrections, which represent the η and p_T - dependent jet response corrections, respectively. The response values are derived from QCD multijet events, where the PYTHIA 8.219 generator [16,17] with the CUETP8M1 [116] UE tune and full GEANT4 [120] detector simulation was used to generate the events used for the estimation of the correction.

In order to compute the response, the particle level jets are matched with the reconstructed jets within a distance parameter of 0.2, half of the jet cone size used for the jets reconstruction is 0.4. Afterwards, this pair of jets is used to calculate the response, the ratio between reconstructed and generated p_T of the jet, $p_T^{\text{reco}}/p_T^{\text{gen}}$, in bins of η_T^{gen} and p_T^{gen} . The obtained simulated jet response

⁸The zero-bias data was selected by random triggers in the presence of bunch crossing.

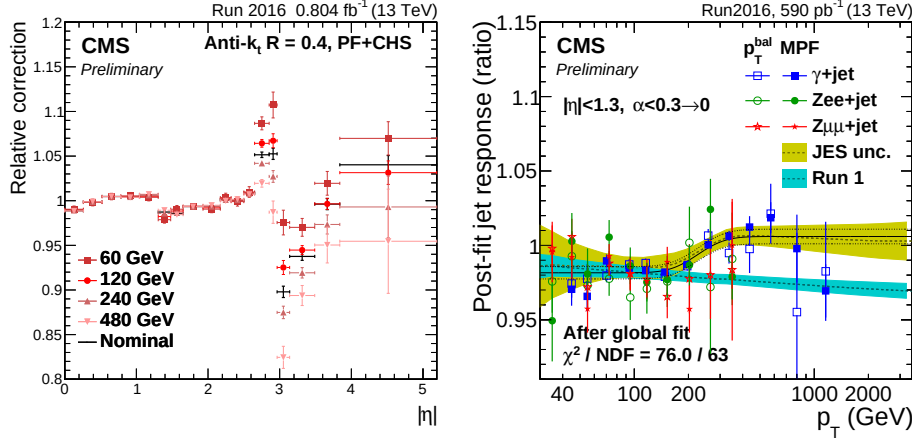


Figure 6.10.: Data/MC scale factors for the jet response η -dependence, evaluated on dijet events (left). Data/MC scale factors for the jet response p_T -dependence, evaluated from photon+jet, $Z(\mu\mu)$ +jet and $Z(ee)$ +jet events, combined in a global fit (right) [177]. The results are estimated for the AK4 PF+CHS jets.

for different detector parts is shown in Fig. 6.9 (right), where a stronger p_T dependence in the endcap and forward regions is observed. These corrections are applied on MC simulation and on data.

The L2L3 response estimation relies only on simulated events, which ensure no biases in the obtained corrections that are natural in the data-based methods due to the imperfect detector resolution. Additionally, MC-based methods have better phase space access compared to the data-based ones, which lacks measurements in the very forward η , as well as low and high p_T , regions.

L2L3Res Residual corrections for data are derived in order to correct for remaining residual data-MC discrepancies. The *L2 residual* corrections are estimated using dijet events, while *L3 residual* corrections are derived from a combination of Z/γ +jet and multijet events. Both corrections are based on exploiting the transverse momentum balance at the hard scattering level. For both cases the reference object is selected to be in the barrel region with $|\eta| < 1.3$ and in case of the L3Res correction, the p_T is evaluated in the range from 30 GeV to 1 TeV. Additionally, the L2Res corrections are derived in the forward region, as shown in Fig. 6.10 (left).

In order to estimate the L2Res and L3Res corrections, the difference in the jet energy response, measured by the p_T balance or MPF⁹ methods [176], is evaluated. Both method first select the reference object and use its $p_{T, \text{ref}}$. In the p_T balanced method the $p_{T, \text{jet}}$ of the second reconstructed jet is then

⁹Missing transverse momentum projection fraction

directly compared to the reference one, so that the response is estimated like:

$$R_{\text{jet}, p_{\text{T}}^{\text{bal}}} = \frac{p_{\text{T}, \text{jet}}}{p_{\text{T}, \text{ref}}}. \quad (6.6)$$

While in the MPF method the whole hadronic recoil is taken into account and the response is computed as:

$$R_{\text{jet}, \text{MPF}} = 1 + \frac{\vec{E}_{\text{T}}^{\text{miss}} \cdot \vec{p}_{\text{T}, \text{ref}}}{(p_{\text{T}, \text{ref}})^2}, \quad (6.7)$$

where $\vec{E}_{\text{T}}^{\text{miss}}$ corresponds to the missing transverse momentum that may occur due to a mismeasurement of the jet momenta. For the high jet p_{T} extrapolation multijet events are used in the *multijet balance* (MJB) method [178], where a high- p_{T} jet in a barrel region is balanced by a recoil system, which can consist of more than two low- p_{T} jets.

Both the p_{T} balance and MPF methods are used in the L2Res and L3Res corrections. The data/MC scale factors for the jet response are shown in Fig. 6.10. The data/MC L3Res response values as a function of the absolute jet p_{T} , obtained for different decay channels, are fitted by a global fit. The global fit function is a fixed erf-log function with a fixed mean value of 225 GeV and width 0.4. Figure 6.10 (right) represents the data to MC scale factors arising from the L3Res correction after applying the global fit. The observed kink around 150 GeV is due to issues with reconstruction in the beginning of Run 2. All obtained response mean values are found to be close to unity with an uncertainty close to 3%.

The jet energy correction is one of the dominant systematic uncertainties in many physics analyses as well as in the top quark mass measurement presented in this thesis. Therefore, within the scope of this thesis the L2Res corrections is discussed in details in Chapter 7 to illustrate the most challenging residual jet energy correction measurement.

Jet energy scale uncertainties

The obtained jet energy corrections sensitive to different uncertainty sources, like modeling of the parton shower and UE, as well as the detector response. The potential biases in the methods used for the JEC estimation can also arise in the JEC uncertainties. These systematic sources are categorized into groups: pileup, absolute energy scale vs. p_{T} , relative energy scale vs. η , extrapolation in p_{T} , jet flavor and time stability. Figure 6.12 shows the contribution from the mentioned sources to the combined JEC uncertainty as a function of the jet p_{T} and η computed for the 2015 Run 2 data.

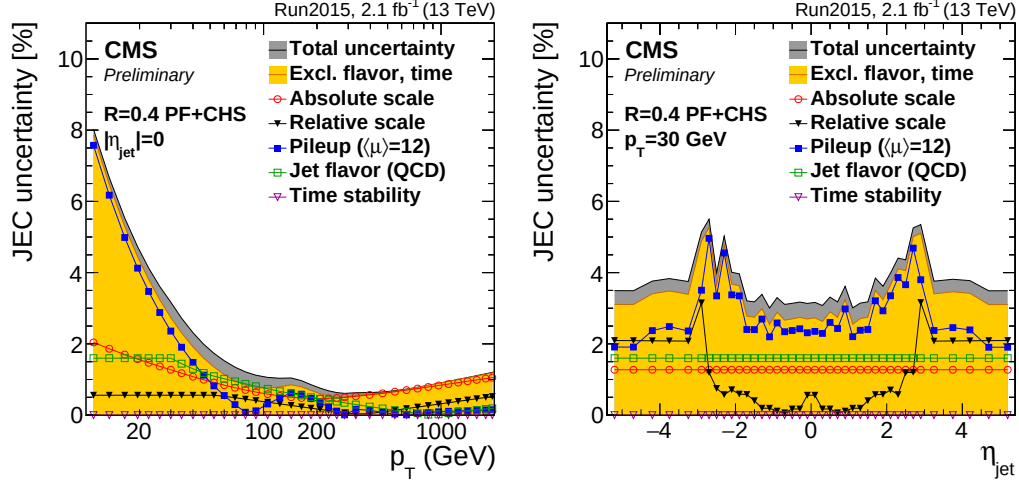


Figure 6.11.: Jet Energy Correction (JEC) Uncertainties for 2015 Run 2 data [177]. The results are estimated for AK4 PF+CHS jets.

For the pileup offset, the most relevant contribution comes from the observed data to MC differences in the pileup offset estimation with the random cone method. The pileup uncertainty is dominant below a jet p_T of 50 GeV as demonstrated in Fig. 6.12 (left). The uncertainty from the absolute energy scale has a dominant contribution in the $|\eta|$ region below 3, while for the $|\eta|$ region above 3, the uncertainty due to the relative energy scale is dominant. The jet flavor uncertainty aims to provide the uncertainty separately for different jet flavors. They are computed by using the dijet QCD samples that are rich in gluon jets, while the $Z/\gamma + \text{jet}$ samples are enriched with quark jets. The jets are reconstructed using the PYTHIA and HERWIG++ generators and the largest observed difference in the response from both of them is taken as uncertainty. The jet flavor is defined by matching the particle jet to the nearest parton from the hard scattering at the generator level within half of the jet cone radius size. PYTHIA and HERWIG++ agrees on the quark flavor response, while a significant difference in the gluon response modeling is found [177]. The UE tunes, e.g., CUETP8M2T4 [116, 117] for PYTHIA and EE5C [111] for HERWIG++, aim to cover data-to-MC differences in the jet fragmentation or in the variables that are used for the quark tagging, i.e., b tagging. The flavor correction for the b quark and the corresponding uncertainty are checked with $Z + b$ events. The jet flavor corrections have not yet been considered at $\sqrt{s} = 13$ TeV. The uncertainty due to the time stability is found to be negligible.

The described uncertainty due to the jet flavor is the dominant one in the top quark mass measurement.

Jet energy resolution

After the JEC is applied to the event, the difference in jet energy resolution (JER) between data and simulation can be measured. The estimation of the JER is based on the dijet balance method, where the momentum conservation in the transverse plane is used.

Since the p_T of the reconstructed and corresponding particle level jets measured in the same η region is slightly different, their resolution values are also expected to be different. Therefore, the widths of the asymmetry distributions are used to derive the JER, which is known as the *dijet asymmetry* method [179]. Thus, the transverse momenta of the two leading jets, $p_{T,1}$ and $p_{T,2}$ are measured and used to obtain the asymmetry distribution:

$$A(p_{T,1}, p_{T,2}) = \frac{p_{T,1} - p_{T,2}}{p_{T,1} + p_{T,2}}. \quad (6.8)$$

The jet- p_T resolution $\sigma(p_T)$ is calculated following the relation:

$$\frac{\sigma(p_T)}{\langle p_T \rangle} = \sqrt{2} \sigma_A. \quad (6.9)$$

The obtained JER results for the 2015 13 TeV data and MC are compared and shown in Fig. 6.12. Figure 6.12 (left) illustrates, that the resolution of the jets measured in data is worse than in MC by about 10% in the barrel and 20% in the forward region. Figure 6.12 (right) demonstrates that the resolution of jets with p_T above 100 GeV is stable against pileup.

Since the JER is worse in data than in MC, the resolution of the reconstructed jets from simulation samples should be smeared using the JER data to MC scale factors. The JER scale factors used in the top quark mass measurement are given in Tab. 6.1, and the way they are applied to the jets is explain in Section 7.4.

Table 6.1.: JER Scaling factors and Uncertainty derived for data recorded at 13 TeV in 2016 with the CMS detector, (Spring16_25nsV10) [180].

Barrel $ \eta $	0.0 -0.5	0.5-0.8	0.8-1.1	1.1-1.3		
data/MC	1.109	1.138	1.114	1.123		
JER SF	± 0.008	± 0.013	± 0.013	± 0.024		
Endcap $ \eta $	1.3-1.7	1.7 - 1.9	1.9-2.1	2.1 - 2.3	2.3 - 2.5	2.5-2.8
data/MC	1.084	1.082	1.140	1.067	1.177	1.364
JER SF	± 0.011	± 0.035	± 0.047	± 0.053	± 0.041	± 0.039
Forward $ \eta $	2.8-3.0	3.0-3.2	3.2-5.0			
data/MC	1.857	1.328	1.16			
JER SF	± 0.071	± 0.022	± 0.029			

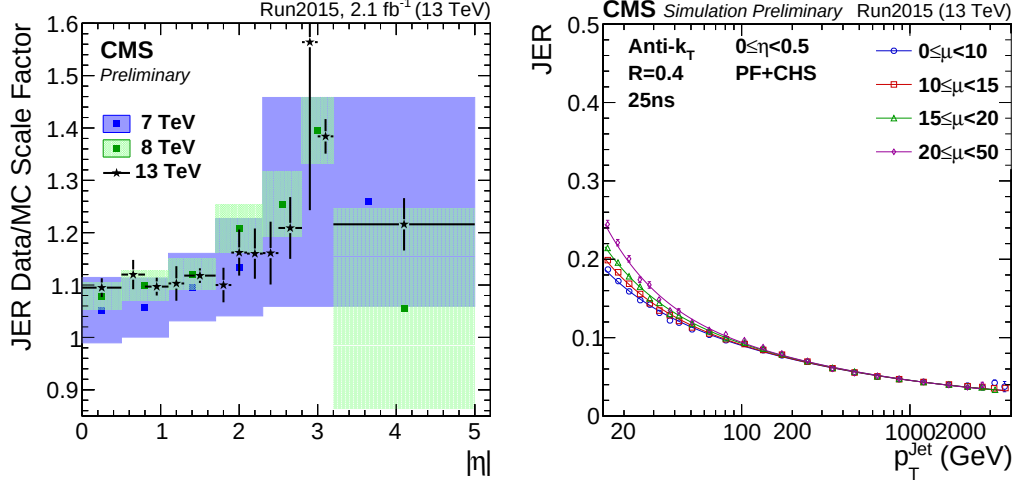


Figure 6.12.: Data/MC scale factors for the jet energy resolution evaluated with the asymmetry method on dijet events (left). Relative jet resolution on simulation, defined as the width (evaluated as the sigma of a Gaussian fit) of the ratio between reconstructed and generated jet p_T , as a function of p_T , for central jets with ($|\eta| < 0.5$) [177]. The results are estimated for the AK4 PF+CHS jets.

Tagging b jets

To perform a precise top quark mass measurement, an accurate identification (*tagging*) of heavy flavor jets is very important. This helps to suppress the background events from QCD processes as well as to improve the $t\bar{t}$ event reconstruction. For the $t\bar{t}$ events selection, where each top quark decays to a W boson and b quark, a *b-tagging* technique to identify the jet arising from b quark hadronization is used. The unique properties of the weakly decaying bottom hadrons, i.e., a high mass $m_B \simeq 5.3$ GeV and relatively long lifetime $c\tau_B \approx 500 \mu\text{m}$, results in a measurable secondary vertex (SV) of the decay products, as shown in Fig. 6.13(left). In contrast to the AVR algorithm, which is used to reconstruct the primary vertices, the secondary vertex candidates are reconstructed by the *inclusive vertex finder (IVF)* [183] algorithm, which operates on the collection of all available tracks in the event, i.e., without requiring jets. A set of looser selection criteria are additionally applied to the selected tracks with IVF such as:

- the number of hits in the silicon tracker is required to be ≥ 8
- the transverse momentum of the tracks has to be ≥ 0.8 GeV
- the longitudinal impact parameter d_z has to be < 0.3 cm.

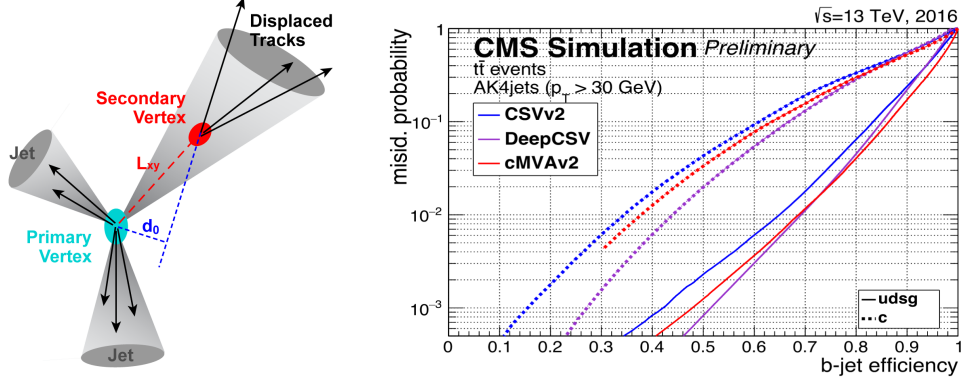


Figure 6.13.: Tracks coming from b-hadron decays have larger IP (so they are displaced) and come from a SV that can be resolved with a high resolution tracker [181]. The L_{xy} is the distance between primary and secondary vertices, d_0 is the impact parameter for a single track (left).

Performance of the b jet identification efficiency algorithms demonstrating the probability for non-b jets to be misidentified as b jet as a function of the efficiency to correctly identified b jets [182]. The curves are obtained on simulated $t\bar{t}$ events using jets within tracker acceptance with $p_T > 30$ GeV, b jets from gluon splitting to a pair of b quarks are considered as b jets (right). The lines shown are for CSVv2, DeepCSV, and cMVA2.

Among these tracks, the displaced tracks are identified as seed tracks when they have an impact parameter significance¹⁰ of at least 1.2 and an impact parameter of at least $50\mu\text{m}$. The resulting seed tracks and nearby tracks are clustered together and fitted by the *adaptive vertex fitter* [149], in order to identify the secondary vertex candidates.

The information about the secondary vertices associated to a jet, the displaced tracks and the presence of a lepton, coming from the semi-leptonic decays of B hadrons, is utilized by the various CMS b-taggers. The most efficient three tagging methods used in the Run 2 are listed below.

CSVv2 The *Combined Secondary Vertex* version 2 (CSVv2) [184] is based on the Run 1 CSV-tagger described in [162]. It uses a multivariate technique, which combines the information about SV candidates and displaced tracks, to evaluate a probability for the jet to originate from a b quark.

DeepCSV The DeepCSV [182] is a new tagger based on the same set of observables used by the CSVv2 algorithm. It is based on a deep neural network.

cMVA2 The *combined multivariate algorithm* version 2 (cMVA2) [184] uses a

¹⁰The impact parameter significance, $\text{IP}/\sigma(\text{IP})$, - an individual track's impact parameter significance, used to account resolution effects. The $\sigma(\text{IP})$ corresponds to the uncertainty of the IP, which varies with the number of hits

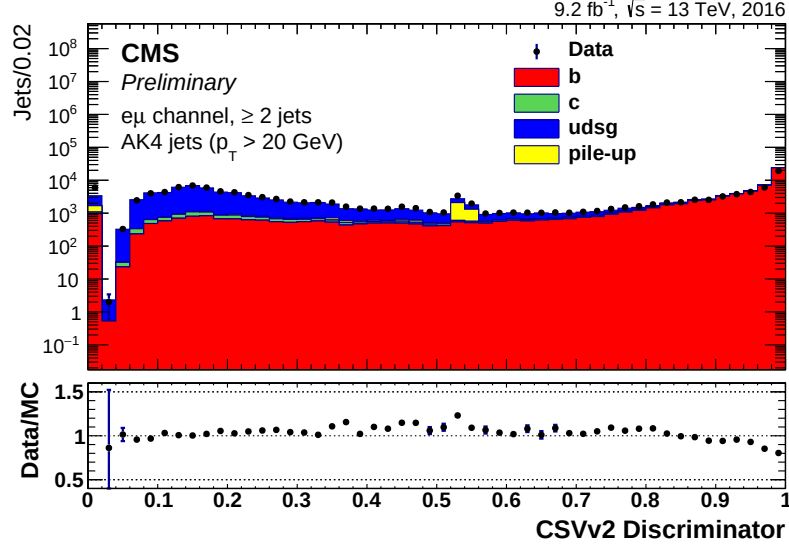


Figure 6.14.: The CSVv2 discriminators for AK4 jets in events fulfilling the top quark pair event selection [187]. The markers correspond to the data. The stacked, colored histograms indicate the contributions of different jet flavors in the simulation. Jets originating from pileup collisions are defined as jets for which no matching generated jet is found with a transverse momentum exceeding 8 GeV. The small "bump" around CSVv2 discriminator values of 0.55 is due to jets with a low transverse momentum and a single displaced track from pileup collisions. The last bins of the histograms contain all entries above the histogram range. The underflow bin is included in the first bin.

Boosted Decision Tree (BDT) [185] that takes as input the different algorithm outputs of CSVv2 and Jet Probability (JP) [162] methods as well as the information from the soft leptons inside the jet.

The performance of the discussed b-tagging algorithms obtained on simulated $t\bar{t}$ events is shown in Fig. 6.13(right). It demonstrates the probability for jets to be misidentified as b jets as a function of the efficiency to identify b jets. The mistag rates of 10%, 1% and 0.1% [186] define the loose, medium and tight working points (WP), respectively. For each WP a corresponding discriminator value is calculated, where the discriminator is the output of the b-tagging algorithms, which is computed from tracks associated with the jets and is sensitive to the flavor content of the jet. The Fig. 6.14 demonstrates the CSVv2 discriminator for AK4 jets, illustrating the contributions of different jet flavors. The performance of the taggers at different WPs is evaluated for data and for MC. Figure 6.15 shows the p_T -dependent misidentification probability of the CSVv2L tagging algorithm in data and MC, where the loose working point is used.

In the top quark mass measurement analysis, the *medium* WP of the CSVv2 tagger is used, with a corresponding discriminator cut of ≥ 0.8484 , which allows a

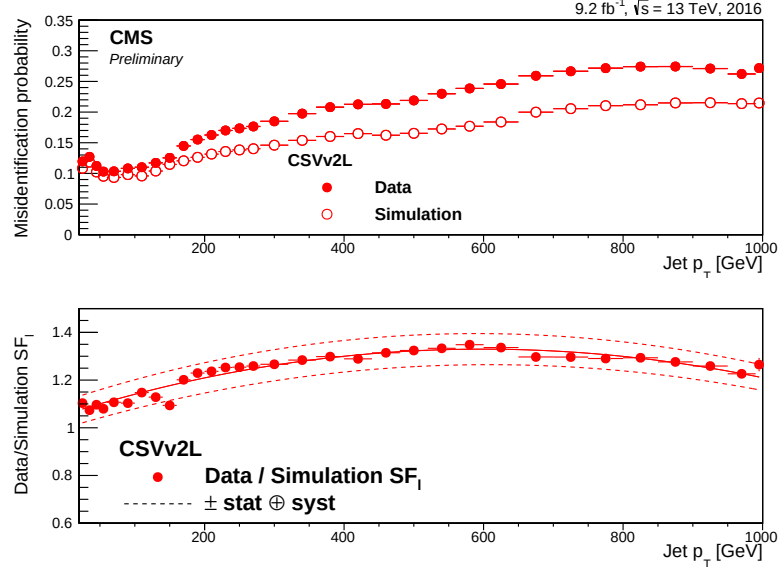


Figure 6.15.: The top panel shows the misidentification (mistag) probability for light jets in data (solid markers) and simulation (open markers) for the CSVv2L tagger as a function of the jet p_T [187]. In the bottom panel the data-to-MC ratio or scale factor for the misidentification probability is shown.

mistag rate of 1% and thus, the b-tagging efficiency is about 70%.

6.6. Missing transverse momentum

The presence of an momentum imbalance in the transverse plane, due to the existence of the weakly interacting neutral particles that are produced in the proton-proton collisions, like neutrinos from W boson decays, is estimated with the *missing transverse momentum* or *energy*. A precise missing transverse momentum reconstruction is important for the lepton+jets $t\bar{t}$ event selection. Furthermore, it may also give impact on new physics discoveries by providing evidence for potential non-SM particles, that are invisible and can carry a considerable amount of momentum. As was mentioned in Section 6.1 the missing transverse momentum is estimated assuming momentum conservation in the transverse plane and thus can be calculated by taking the negative vectorial sum of all four momenta of the reconstructed particle flow candidates:

$$\vec{E}_T^{\text{miss}} = - \sum_{i=1}^{N_{\text{PF candidates}}} \vec{p}_{T,i}. \quad (6.10)$$

The missing transverse momentum, similar to the jets, has to be corrected for the detector misalignment and the presence of pileup and underlying event. This is performed by the *type-I correction* [188, 189], which uses the four momentum of

the PF candidates, that are clustered to calibrated jets. The type-I PF- $\cancel{E}_T$, $\cancel{E}_T^{\text{corr}}$, is estimated as:

$$\vec{\cancel{E}}_T^{\text{corr}} = - \sum_{i=1}^{N_{\text{PF candidates}}} \vec{p}_{T,i} - \sum_{j=1}^{N_{\text{jets}}} (\vec{p}_{T,j}^{\text{corr}} - \vec{p}_{T,j}^{\text{raw}}) \quad (6.11)$$

where $\vec{p}_T^{\text{corr}}$ and $\vec{p}_T^{\text{raw}}$ are the jet transverse momenta with and without jet energy corrections applied. The pileup corrections are excluded in the jet energy corrections. For the type-I correction the jets that have less than 90% of their energy deposited in the ECAL and $p_T > 15 \text{ GeV}$ are considered [190].

7. Determination of residual jet energy correction

Jets, formed by the collimated spray of colorless hadrons produced during pp collisions, are very important event signatures. They are widely used by many physics analyses. An accurate measurement of the jet energy is crucial for the precise measurements of standard model (SM) properties as well as physics beyond the SM. An example is the measurement of the top quark mass, a SM parameter that depends strongly on the precision of the measured jet energy, as discussed in depth in Section 8.5. Additionally, searches for new physics based on the final states with several jets and missing transverse energy can also profit from accurate jet energy scale correction. Due to the limited detector resolution, mismeasurement of jet energies can happen, which can lead to a large momentum imbalance and thus large missing transverse energy.

The factorized approach used for the jet energy correction (JEC) at CMS is described in Section 6.5. In this chapter, the first η -dependent L2 relative residual corrections measurement at a center-of-mass energy of 13 TeV at CMS is presented. The full dataset recorded in 2015 with the CMS detector is analyzed, corresponding to an integrated luminosity of $\mathcal{L} = 2.11 \text{ fb}^{-1}$. The residual corrections are based on a data-driven method that compares the jet energy responses from data and MC. The L2 residual correction is one part of the chain that is used to compute the final JEC. It is determined from the dijet event samples after correcting the jet energy for pileup and for the difference between reconstructed and generated jet energy in simulation. Dijet events, due to their large production rate, are well suited for estimating the response corrections of jets over a wide p_T -range relative to the reference jet in the barrel part of the detector. Thus, the corrections can be derived with high statistics even at high jet momenta.

The determination of the relative residual corrections is documented in [13, 151]. The properties of dijet events (such as the p_T balance in back-to-back azimuthal topology) are utilized by the p_T balance [176] and the MPF [176] methods, in order to derive scale factors to correct for the data to MC differences.

The results for the L2 residual correction presented in this chapter were performed by the author of this thesis together with Marc Stöver [191] and are documented in [192].

7.1. Dijet event topology

The dijet event topology has a perfect back-to-back structure with two jets balanced in p_T . One of the two leading jets is required to be reconstructed in the barrel region of the calorimeter and is called the barrel jet, while the other jet can be at arbitrary η and is called the probe jet. In order to ensure that in the transverse plane both leading jets point into opposite directions, they are required to be azimuthally separated by $\Delta\phi_{\text{dijet}}$ as shown in Fig. 7.1. In the azimuthal plane $\Delta\phi_{\text{dijet}}$ is defined by:

$$\Delta\phi_{\text{dijet}} = \min(|\phi_1 - \phi_2|, 2\pi - |\phi_2 - \phi_1|), \quad (7.1)$$

where ϕ_1 and ϕ_2 corresponds to the azimuthal angle of the first and second leading jet, respectively.

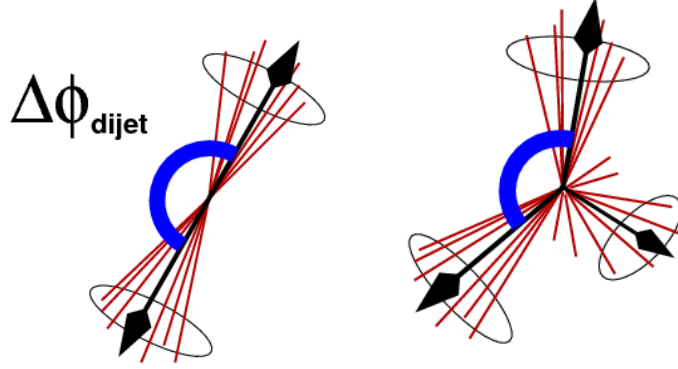


Figure 7.1.: A sketch of the angle $\Delta\phi_{\text{dijet}}$ in the azimuthal plane in a dijet event (left) and in an event with radiation outside the dijet system (right). [193].

Requiring exactly a 2-to-2 process reduces the statistics drastically and biases towards specific hadronisation topologies, and so an additional third jet is allowed in the event. However, it is important to control the transverse momentum of the third jet with respect to the two leading jets. This is achieved by requiring an upper bound of about 0.2 on the relative transverse momentum of the third jet:

$$\alpha = \frac{p_T^{\text{jet3}}}{p_T^{\text{ave}}}, \quad (7.2)$$

where $p_T^{\text{ave}} = (p_T^{\text{barrel}} + p_T^{\text{probe}})/2$ is the average transverse momentum of the dijet system.

Similar calculations of the response value of the jet energy is performed with respect to other reference objects that typically has a good energy resolution, mostly photon and di-lepton. However, these samples have much lower statistics. However, since both jets in the dijet event have a similar, poor resolution, their usage as reference objects may lead to bias in the relative response measurement. This

may occur because a p_T bin of the reconstructed jet contains not only p_T from the particle level jets that lies in the same bin, but also some other jet p_T outside the bin, whose response can fluctuate. Therefore, in order to reduce bias from the jet energy resolution (JER) in the relative response measurement, the measurements are performed in bins of the average p_T of the dijet system p_T^{ave} . The bias due to ISR and FSR, as well as bias in the ratio variables, is canceled when using symmetric transverse momentum bins.

7.2. Jet response determination

The transverse momentum of the reconstructed jet p_T^{reco} is not equal to the transverse momentum of the corresponding jet at the particle level p_T^{gen} . The difference can be quantified by the p_T jet response $\mathcal{R}$,

$$\mathcal{R} = \frac{p_T^{\text{reco}}}{p_T^{\text{gen}}}, \quad (7.3)$$

which depends on the η and p_T of a given jet.

The measurement of the jet response is performed using a simulated QCD multijet sample. However, the p_T of generated jets is not available directly in real data. Therefore, one possibility is to use the asymmetry $\mathcal{A}$ in order to measure the jet energy response in data and in MC. The asymmetry $\mathcal{A}$ is estimated for events with at least two jets from the p_T of the leading jets:

$$\mathcal{A} = \frac{p_T^{\text{probe}} - p_T^{\text{barrel}}}{p_T^{\text{probe}} + p_T^{\text{barrel}}} = \frac{p_T^{\text{probe}} - p_T^{\text{barrel}}}{2p_T^{\text{ave}}}. \quad (7.4)$$

7.2.1. The dijet p_T balance method

The average asymmetry value $\langle \mathcal{A} \rangle$ or balanced quantity $\mathcal{B}$ can be used to determine the jet response using the dijet p_T balance method [176, 194–196] that is based on transverse momentum conservation :

$$\mathcal{R}_{\text{rel}}^{p_T} = \frac{1 + \langle \mathcal{A} \rangle}{1 - \langle \mathcal{A} \rangle} \equiv \frac{2 + \langle \mathcal{B} \rangle}{2 - \langle \mathcal{B} \rangle}, \quad (7.5)$$

where $\mathcal{B}$ is calculated as:

$$\mathcal{B} = \frac{p_T^{\text{probe}} - p_T^{\text{barrel}}}{p_T^{\text{ave}}}. \quad (7.6)$$

The defined relative response of the probe jet with respect to the barrel jet given in Eq. 7.5 is mathematically equivalent in the limit of narrow p_T^{ave} bins to the equation:

$$\mathcal{R}_{\text{rel}}^{p_T} = \frac{\langle p_T^{\text{probe}} \rangle}{\langle p_T^{\text{barrel}} \rangle}. \quad (7.7)$$

Hence for a single event, the dijet p_T balance relative response is given by the momentum ratio of the probe jet over barrel jet, $p_T^{\text{probe}}/p_T^{\text{barrel}}$. The p_T balance method is also used for the jet p_T resolution measurement [176].

Even though the measurement is performed in bins of p_T^{ave} , the resolution bias is not completely canceled on average. Therefore, the use of the dijet p_T balance method warrants a study of the resolution bias effect.

7.2.2. Missing transverse energy projection fraction method

The missing transverse energy projection fraction (MPF) method [195–197] is designed to consider the whole hadronic activity response, while in the dijet p_T balance method the jet response is computed by only using the p_T of the measured jet relative to the p_T of the reference jet.

At the particle level, the reference barrel jet is perfectly balanced in the transverse plane with its hadronic recoil:

$$\vec{p}_{T, \text{true}}^{\text{barrel}} + \vec{p}_{T, \text{true}}^{\text{recoil}} = 0 \quad (7.8)$$

At the reconstructed level, the detector responses to the barrel jet $\mathcal{R}_{\text{barrel}}$ and the hadronic recoil $\mathcal{R}_{\text{recoil}}$ differ from unity. Therefore, Eq. 7.8 should be rewritten as:

$$\mathcal{R}_{\text{barrel}} \cdot \vec{p}_{T, \text{true}}^{\text{barrel}} + \mathcal{R}_{\text{recoil}} \cdot \vec{p}_{T, \text{true}}^{\text{recoil}} = -\vec{E}_T^{\text{miss}}. \quad (7.9)$$

Since a dijet event does not have an intrinsic E_T^{miss} , the missing transverse momentum used in Eq. 7.9 corresponds to the possible missmeasurement of the jet and hadronic recoil. Moreover, $\vec{E}_T^{\text{miss}}$ accounts for possible additional soft activity in the event. In order to obtain an estimator for the MPF response and the probe jet, the projection of $\vec{E}_T^{\text{miss}}$ along the barrel jet axis is used as shown in Fig. 7.2.

The MPF response or $\mathcal{R}_{\text{recoil}}$ is then evaluated using Eq. 7.8 and Eq. 7.9:

$$\mathcal{R}_{\text{recoil}} = \mathcal{R}_{\text{barrel}} + \frac{\vec{E}_T^{\text{miss}} \cdot \vec{p}_{T, \text{true}}^{\text{barrel}}}{(\vec{p}_{T, \text{true}}^{\text{barrel}})^2} \equiv \mathcal{R}_{\text{MPF}}. \quad (7.10)$$

In order to extract the jet energy response from the MPF response $\mathcal{R}_{\text{MPF}}$, extra steps are needed. If the additional particles in the event are not clustered into the probe jet, but their response is similar to that of the probe jet, the approximation that $\mathcal{R}_{\text{probe}} = \mathcal{R}_{\text{recoil}}$ can be made. This is also the case if the additional particles are in the direction perpendicular to the barrel jet axis. Naturally, $\mathcal{R}_{\text{probe}}$ is equal to $\mathcal{R}_{\text{recoil}}$ if most of the recoil, usually coming from ISR and FSR, is clustered in the probe jet or reduced as much as possible. In order to reduce additional activity, the selection requirement for α given in Eq. 7.2 is applied. Moreover, since the central part of the detector is corrected by the L3 residual correction at a later stage, the response of the jet in the barrel region $\mathcal{R}_{\text{barrel}}$ is set to be equal one, and

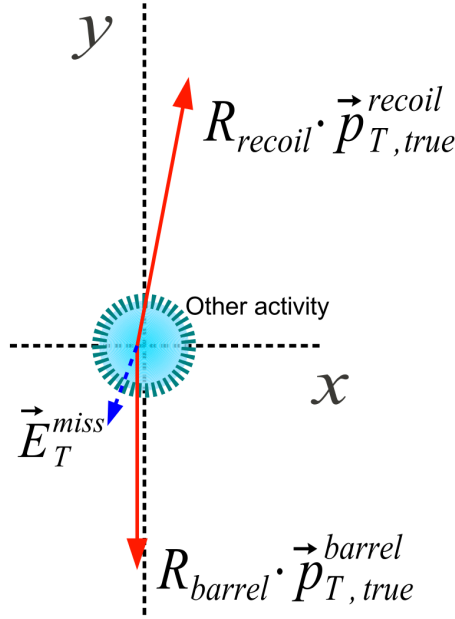


Figure 7.2.: A sketch of the projection of $\vec{E}_T^{\text{miss}}$ along the reference jet axis, used by the MPF method [198].

thus $\vec{p}_T^{\text{barrel}} = \vec{p}_{T, \text{true}}^{\text{barrel}}$.

Having these requirements fulfilled, Eq. 7.10 can be rewritten in terms of the probe jet response:

$$\mathcal{R}_{\text{probe}} \approx \mathcal{R}_{\text{MPF}} = 1 + \frac{\vec{E}_T^{\text{miss}} \cdot \vec{p}_T^{\text{barrel}}}{(\vec{p}_T^{\text{barrel}})^2}. \quad (7.11)$$

By utilizing $\vec{E}_T^{\text{miss}}$, the MPF method is less sensitive to the exact modeling of the radiation in the simulation, as well as to the resolution bias. Therefore, the MPF method is used in CMS as the main method to measure the jet energy response and the dijet p_T balance method is only used for the systematic uncertainty estimation.

7.3. Dijet event selection and reconstruction

In this section the datasets used for the residual JEC calculation are listed. The study of the trigger efficiency and optimization of a new trigger threshold, as well as the selection of the dijet events are presented.

7.3.1. Data and simulation samples

The L2 residual correction presented in this study is obtained using the full dataset¹ collected with the CMS detector in 2015, which corresponds to an integrated lumi-

¹Cert_246908-260627_13TeV_PromptReco_Collisions15_25ns_JSON.txt

osity of 2.1 fb^{-1} . The “JetHT” data samples are used, which contain events with jets in the HCAL forward (HF) region. The data used for the analysis are taken from the prompt reconstruction stream.

Simulated QCD event samples are used throughout the study. The MC samples are generated with the multipurpose PYTHIA 8.219 generator [16, 17] using the CUETP8M1 UE tune [116]. The simulation samples are processed through the CMS detector simulation based on GEANT4 [118, 119]. The events in the nominal MC sample are generated with a flat p_T spectrum. This is done in order to ensure efficient coverage of the large kinematic phase space with small statistical uncertainties².

The used data and MC samples are listed in Tab 7.1. These are used during the L2 residual correction measurement.

Table 7.1.: The data set with corresponding integrated luminosity and simulated MC samples with the corresponding number of events.

Data samples	$\mathcal{L}_{\text{int}}$ (pb^{-1})
/JetHT/Run2015C_25ns-05Oct2015-v1/MINIAOD	16.3
/JetHT/Run2015D-PromptReco-v3/MINIAOD	552.7
/JetHT/Run2015D-PromptReco-v4/MINIAOD	1521.0
Simulated MC sample	N_{events}
/QCD_Pt-15to7000 TuneCUETP8M1 Flat 13TeV pythia8 with RunIISpring15DR74-Asympt25nsRaw	9858314

The corresponding data and MC events are corrected for the pileup offset, and p_T and $|\eta|$ differences between reconstructed and generated jets using the official CMS corrections: Summer15_25nsV6 [199].

7.3.2. Trigger efficiency study

The events are selected online with HLT triggers that are based on the p_T^{ave} of two leading jets, which provide an unbiased dataset for the analysis. These dijet triggers are available with large prescale³ factors, so the number of data events in the endcap and HF regions is limited. Therefore, large η bins are used during the residual JEC estimation in the forward regions. Additionally, the corresponding trigger selects events with a certain threshold value on the p_T^{ave} of two leading jets.

In order to optimize the trigger selection, a trigger efficiency study is performed. The trigger is considered to be efficient if it reaches 99% of the efficiency plateau. The following sets of triggers are studied HLT_DiPFJetAve* and HLT_DiPFJetAve*_HFJEC, where the last one is used to increase the number of events with jets in the HF region.

²The p_T spectrum is reweighted later in the analysis.

³The trigger path is prescaled due to the high instantaneous luminosity, and thus only a fraction of all events satisfying the trigger thresholds are recorded.

The efficiency for each trigger path, except the one with the lowest p_T^{ave} threshold, is estimated by

$$e(B) = \frac{N(\text{passing B})}{N(\text{passing A})}, \quad (7.12)$$

where $N(\text{passing A})$ and $N(\text{passing B})$ are the number of events that passed the trigger path with lower and higher threshold. The efficiency distributions for the available trigger path are shown in Fig. 7.3. To restrict the analysis to events

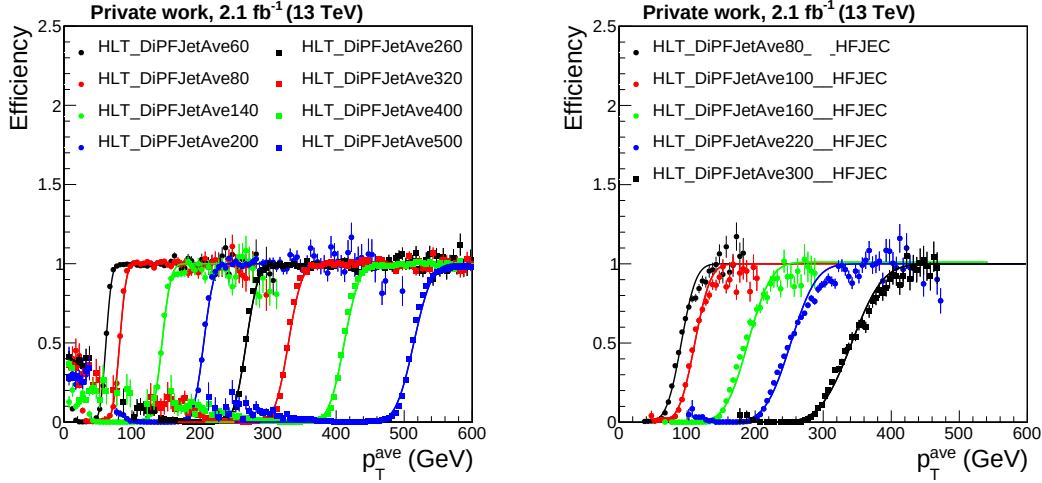


Figure 7.3.: Trigger efficiency turn on curves determined for the dijet nominal (left) and HF (right) trigger path.

where the triggers are fully efficient, new threshold values are obtained by fitting the distribution of the trigger efficiency. The function that is used to parameterize the dijet trigger efficiency is the smooth function [200]:

$$f(p_T^{\text{ave}}) = \frac{1}{2}N \left(1 + \text{erf} \left(\frac{p_T^{\text{ave}} - p_0}{\sqrt{2} \cdot p_1} \right) \right), \quad (7.13)$$

with the error function⁴ erf, where parameters p_0 and p_1 are the p_T values at which the efficiency reaches 50% and the slope of the curve, respectively. N characterizes the maximum efficiency in the plateau region. Thus, the plots are normalized by scaling the efficiency distribution with $1/N$, since the absolute value of N depends on the prescale of the trigger. The maximum efficiency for each trigger path is in the plateau region. The new threshold is estimated at the value where the efficiency function equals 99%. The new trigger threshold versus the generated trigger threshold is shown in Fig. 7.4. The obtained values are fitted with a straight line and the threshold values for the lowest p_T^{ave} triggers are obtained by a linear extrapolation.

The measured dijet trigger thresholds are summarized in Tab. 7.2.

⁴ $\text{erf}(x) = \frac{1}{\sqrt{\pi}} \int_{-x}^x e^{-t^2} dt$

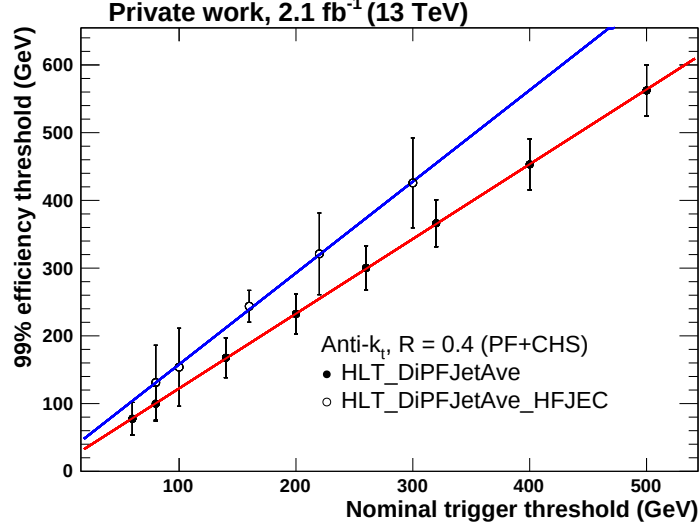


Figure 7.4.: Trigger thresholds that correspond to an efficiency of 99% versus the generated trigger threshold and the linear extrapolation for the triggers with the lowest p_T^{ave} .

Table 7.2.: Trigger thresholds of dijet nominal and HF p_T^{ave} triggers.

Trigger	Threshold (GeV)	HF Trigger	Threshold (GeV)
HLT_DiPFJetAve40	56	HLT_DiPFJetAve60_HFJEC	77
HLT_DiPFJetAve60	78	HLT_DiPFJetAve80_HFJEC	131
HLT_DiPFJetAve80	100	HLT_DiPFJetAve100_HFJEC	154
HLT_DiPFJetAve140	168	HLT_DiPFJetAve160_HFJEC	244
HLT_DiPFJetAve200	232	HLT_DiPFJetAve220_HFJEC	321
HLT_DiPFJetAve260	300	HLT_DiPFJetAve300_HFJEC	426
HLT_DiPFJetAve320	366		
HLT_DiPFJetAve400	453		
HLT_DiPFJetAve500	562		

7.3.3. Dijet event selection

For the dijet event selection only PF candidates after charged hadron subtraction (CHS), as described in Section 6, are used. Jets are clustered with the anti- k_T [167] algorithm with the radius parameter $R = 0.4$. At this stage, the energy of the jet is corrected for pileup effects with the L1 offset correction, and for differences in the jet energy response of the reconstructed and generated jets.

In order to select dijet events, the following requirements have to be fulfilled:

- The event selection requires at least two jets to be reconstructed.
- All selected jets have to pass the p_T^{ave} trigger thresholds.
- At least one of the two leading jets serves as a reference and is reconstructed in

the barrel region $|\eta| < 1.3$, i.e., the barrel jet, while the other jet has arbitrary η , i.e., the probe jet.

- Two leading jets are required to be back-to-back, with an azimuthal separation of $\Delta\varphi_{\text{dijet}} > 2.7$
- The dijet asymmetry values between the two leading jets has to be $|A| < 0.7$, in order to suppress contributions from mismeasured jets.
- In order to control the additional jet activity in the event, the ratio of the p_T of the additional most energetic jet to the p_T^{ave} of the leading jets, i.e., α value, has to be less than 0.2.

The analysis is performed in bins of p_T^{ave} and η of the probe jet, in order to account for the resolution dependence in p_T and η . The p_T^{ave} is defined by the trigger thresholds, as shown in Tab. 7.3 together with the corresponding number of selected data events in each bin.

Table 7.3.: Bins of p_T^{ave} used in the analysis with the required trigger to be fired, as well as the corresponding number of selected data events for the full η region.

bin Nr	p_T^{ave} range	fired trigger	number of events
1	56 – 78	HLT_DiPFJetAve40	160123
2	78 – 100	HLT_DiPFJetAve60	304608
3	100 – 168	HLT_DiPFJetAve60	214865
4	168 – 232	HLT_DiPFJetAve140	76943
5	232 – 300	HLT_DiPFJetAve200	322417
6	300 – 366	HLT_DiPFJetAve260	135542
7	366 – 453	HLT_DiPFJetAve320	368976
8	453 – 562	HLT_DiPFJetAve400	484838

The jet η bins are given in Tab. 7.4.

Table 7.4.: Bins of $|\eta|$ used to perform the residual response.

bin Nr	$ \eta $ range	bin Nr	$ \eta $ range
1	0 - 0.261	9	2.172 - 2.322
2	0.261 - 0.522	10	2.322 - 2.5
3	0.522 - 0.783	11	2.5- 2.65
4	0.783 - 1.044	12	2.65 - 2.853
5	1.044 - 1.305	13	2.853 - 2.964
6	1.305 - 1.653	14	2.964 - 3.139
7	1.653 - 1.93	15	3.139 - 3.489
8	1.93 - 2.172	16	3.489 - 5.191

A comparison between the data and MC is shown in Section 7.5 after applying the event reweighting in the simulation.

7.4. Event reweighting in simulation

The simulated events used in this study are generated with a flat p_T spectrum. Therefore it is necessary to reweight the simulated p_T^{ave} distribution in order to get the p_T spectrum predicted by PYTHIA. The p_T resolution of jets in the simulation is naturally different in data than in MC. Additionally, the pileup in data depends directly on the instantaneous luminosity, which is accounted for the treatment of the MC samples.

Pileup reweighting

Even though the MC events simulated from PYTHIA contain additional proton-proton interactions in each bunch crossing, i.e., pileup (PU), they should be reweighted to the pileup from the data. In data the measured PU distribution is obtained using the BRIL Work Suite software [201] using the best-fit value for the total pp cross section of 69 mb [175]. An example PU calculation is given in Appendix A.1. Figure 7.5 (left) shows the PU distributions in data and MC. The MC distribution is reweighted to describe the obtained data. The weights are normalized in order not to change the total number of simulated events. The pileup reweighting applied to the simulated events naturally influences the distribution of the primary vertices, as shown in Fig. 7.5 (right).

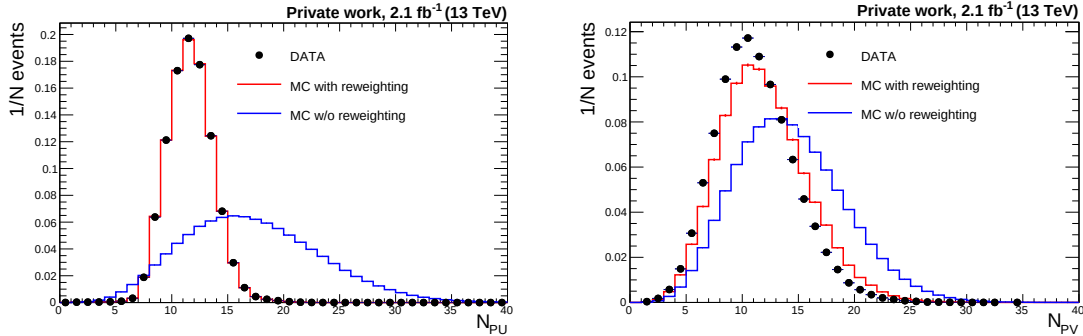


Figure 7.5.: Number of pileup N_{PU} (left) and primary vertices N_{PV} (right) in data and in simulation before and after the PU reweighting. The distributions are normalized to unity.

Jet energy resolution correction

Methods used to calculate the relative response are sensitive to differences in the JER, therefore it is important to have the same jet resolution for the simulation and data. Jets in the simulation need to be smeared to match the measured resolution in data. First, it is necessary to ensure that jets are well-matched at the detector and reconstruction levels. This is done by matching jets within the radius parameter $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} < 0.3$.

Thus, the transverse momenta $p_{T,\text{reco}}^{\text{jet}}$ of the reconstructed jets are smeared using the data/MC JER scale factors C_{JER} listed in Tab. 7.5 in the relation:

$$p_T^{\text{smeared}} = \max\left[0, p_{T,\text{gen}}^{\text{jet}} + C_{\text{JER}}(p_{T,\text{reco}}^{\text{jet}} - p_{T,\text{gen}}^{\text{jet}})\right], \quad (7.14)$$

where $p_{T,\text{gen}}^{\text{jet}}$ and $p_{T,\text{reco}}^{\text{jet}}$ are the transverse momentum of the jet at the generator (detector) level and smeared p_T of the jet at reconstructed level. The missing transverse momentum is corrected accordingly.

Table 7.5.: JER Scaling factors for 13 TeV data recorded in 2015 with the CMS detector [180].

$ \eta $	0.0-0.8	0.8-1.3	1.3-1.9	1.9-2.5	2.5-3.0	3.0-3.2	3.2-5.0
data/MC	1.061	1.088	1.106	1.126	1.343	1.303	1.320
JER SF	± 0.023	± 0.029	± 0.030	± 0.094	± 0.123	± 0.111	± 0.286

Dijet p_T^{ave} event reweighting

In each p_T^{ave} region, which are defined by the trigger thresholds and listed in Tab. 7.3 the number of simulated events are reweighted to match the number of events in data. This procedure is needed in order to take into account the trigger prescale factors used during data taking. The reweighting is applied for events with $\alpha < 0.2$. The MC to data comparison for p_T^{ave} is shown in Fig. 7.6. The data is well described by the MC simulations after event weighting is applied.

7.5. L2 residual correction estimation

After applying the reweighting procedure, the MC simulations are found to describe the data well. This was tested with the discussed methods for the jet relative residual determination according to Eq. 7.7 and Eq. 7.11 given for the dijet p_T balance and MPF methods, respectively. Examples of the reconstructed dijet asymmetry in the barrel region used by the dijet p_T balance method and MPF response are shown in Fig. 7.7.

Since the residual JEC is based on a data-driven method, it is effected by detector-related biases like p_T resolution and physics properties such as the steeply falling p_T spectrum that has a larger number of incoming jets with smaller true p_T . The ISR and FSR bias for the jet response and will be discussed in the following. The p_T resolution and p_T spectrum of the p_T^{ave} were discussed in Section 7.4.

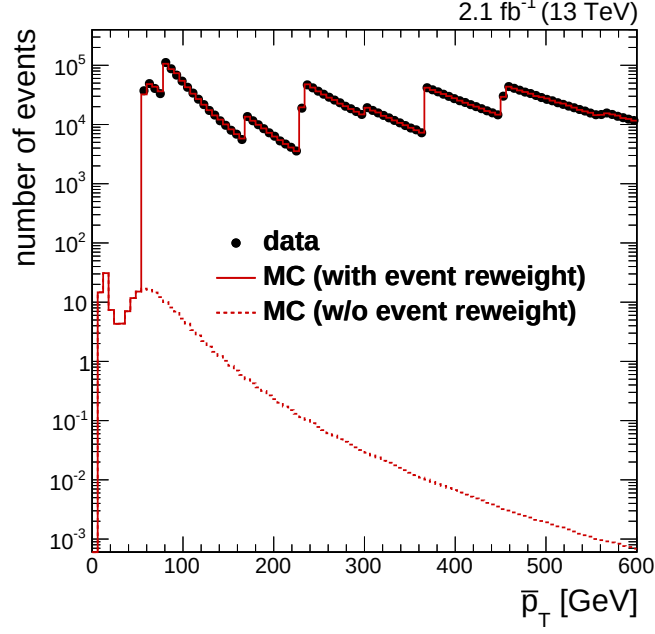


Figure 7.6.: The average transverse momentum of the two leading jets $p_T^{\text{ave}} = \bar{p}_T$ after and before p_T smearing, PU and p_T event reweighting.

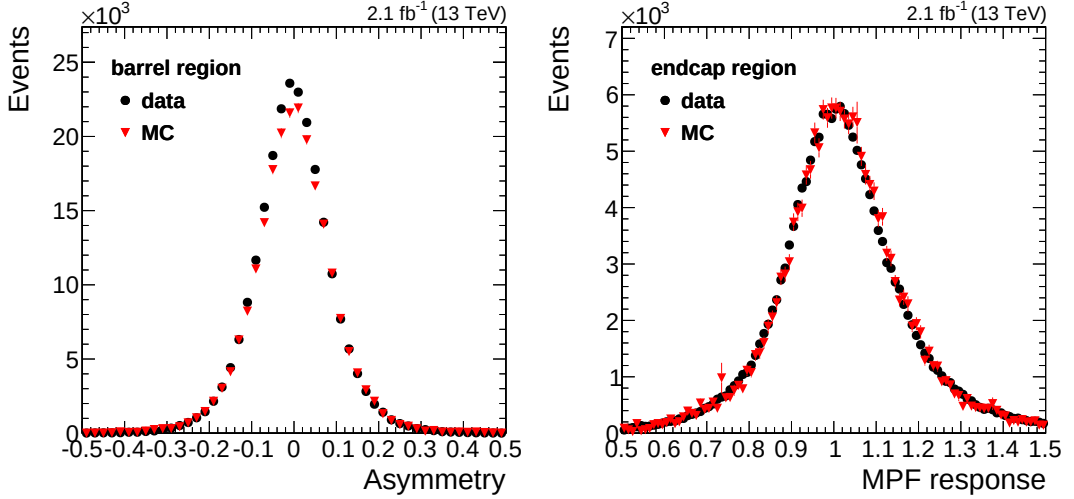


Figure 7.7.: The asymmetry distribution in the barrel region (left) and the MPF response distribution in the endcap region (right).

7.5.1. Radiation correction of the methods

A bias due to ISR and FSR are possible in events with one central and one forward jet. This additional activity depends on the η regions in the detector. Therefore, the measured values of the residual response depend on the additional jet activity α . Hence, in order to reduce the radiation bias, the α value is required to be small. In the case of the relative transverse momentum of the third jet $\alpha = 0$, only two leading jets are studied, but this requirement drastically reduces the number of events. Therefore, the bias due to ISR and FSR in the jet response is estimated as

a function of the suppression of additional jet activity in bins of η .

The correction factors k_{FSR} for the ISR + FSR (for brevity FSR) are obtained by calculating the data/MC double ratio for a linearly extrapolated $\alpha \rightarrow 0$ and $\alpha < 0.2$ as following:

$$k_{\text{FSR}}(\alpha = 0.2) = \frac{\left(\frac{\mathcal{R}^{\text{MC}}}{\mathcal{R}^{\text{data}}} \right)_{\alpha \rightarrow 0}}{\left(\frac{\mathcal{R}^{\text{MC}}}{\mathcal{R}^{\text{data}}} \right)_{\alpha < 0.2}}, \quad (7.15)$$

where $\mathcal{R}$ is the residual response calculated separately for both the dijet p_T balance and MPF methods.

In order to perform the linear extrapolation of α towards 0, events with α up to 0.4 are considered. For each α requirement, the MC-to-data ratio of the response value is calculated using both methods. The obtained ratio values are fitted with a liner function as shown in Fig. 7.8 in the barrel and endcap regions. The ratios for each $|\eta|$ bin are given in Appendix A.2.

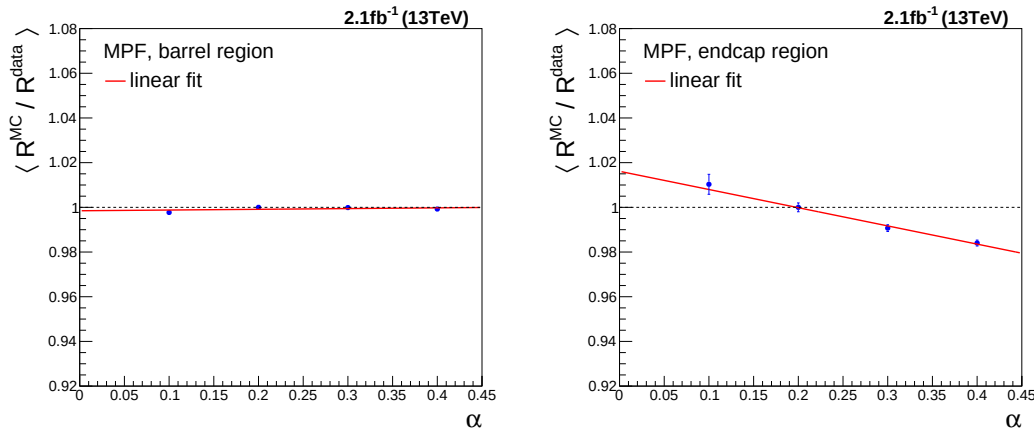


Figure 7.8.: The MC/data ratio of the MPF residual response as a function of the additional jet activity α and the linear extrapolation to $\alpha \rightarrow 0$ of the measured ratios. The results are shown for the barrel and the endcap region.

The values obtained by extrapolating $\alpha \rightarrow 0$ are then used in Eq. 7.15 to calculate the k_{FSR} factor for each $|\eta|$ region. The estimated k_{FSR} factors in each $|\eta|$ bin are shown in Fig. 7.9. The k_{FSR} values in the barrel region are found to be close to unity, while in the forward region deviations of up to 5% for the dijet p_T balance and 2% for the MPF are visible. Even though, by requiring $\alpha < 0.2$ minimizes the resolution-bias effect. The correction factor k_{FSR} for the ISR + FSR shows that the MPF method is less sensitive to additional radiation, due to its definition.

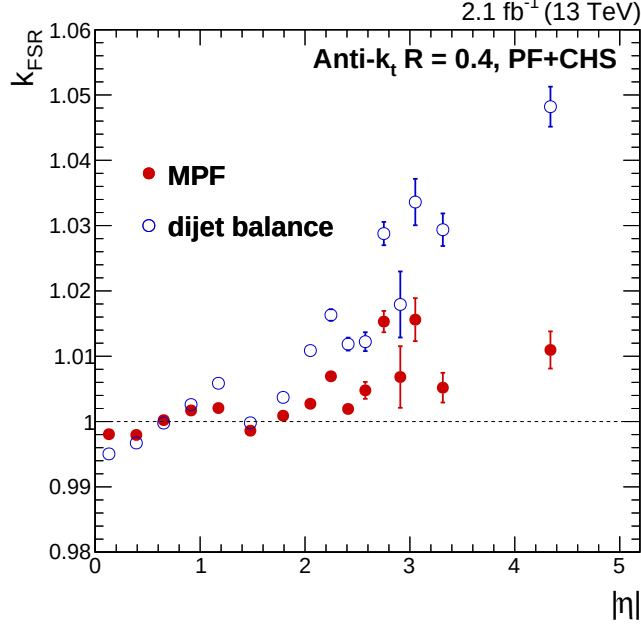


Figure 7.9.: The correction factor k_{FSR} for the ISR + FSR calculated separately for the dijet p_T balance and MPF methods as a function of the probe jet pseudorapidity.

7.5.2. Relative L2 residual determination

The L2 relative residual is estimated for the dijet p_T balance and MPF and corrected for ISR + FSR biases by the correction factor k_{FSR} :

$$\text{L2Res} = \left(\frac{\mathcal{R}^{\text{MC}}}{\mathcal{R}^{\text{data}}} \right)_{\alpha < 0.2} \cdot k_{\text{FSR}}. \quad (7.16)$$

In order to correct the result of the relative L2 residual for possible p_T dependence, a *relative correction* is applied. The relative correction can be estimated by parameterizing the MC-to-data ratio of the response with a log-linear function which depends on p_T^{ave} as:

$$f_{\log\text{Lin}}(p_T^{\text{ave}}) = p_0 + p_1 \log(p_T^{\text{ave}}). \quad (7.17)$$

As a cross check, the MC-to-data response ratio was fitted with a constant function. This parametrization is performed in each bin of $|\eta|$. An example of the response ratio distribution as a function of the average transverse momentum of the dijet system p_T^{ave} is shown for the barrel and endcap regions in Fig. 7.10. The corresponding distribution for each $|\eta|$ bin is given in Appendix A.2. The relative correction is performed separately for the dijet p_T balance and the MPF method. In the barrel region, the p_T dependence of the MC-to-data response ratio is found to be nearly flat, while in the endcap and forward region a significant dependence is observed.

The final L2 relative residual corrections account for the η and p_T dependence and

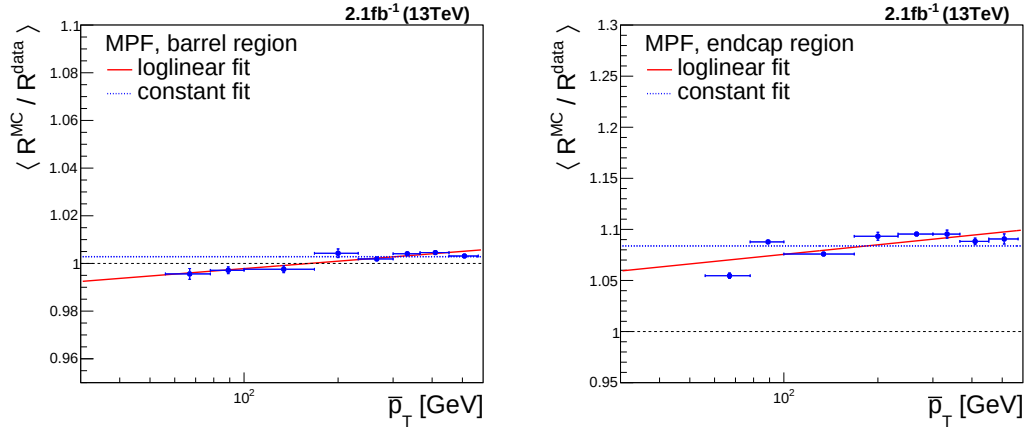


Figure 7.10.: The MC/data ratio of the MPF residual response as a function of an average transverse momentum of the dijet system p_T^{ave} and its parametrization with log-linear and constant functions. The results are shown for the barrel and the endcap regions.

are estimated using the k_{FSR} factors and parameters of the constant or log-linear fit used for the p_T dependence extrapolation:

$$\text{L2Res} = k_{\text{FSR}} \cdot p_{\text{const}}, \quad (7.18)$$

or

$$\text{L2Res} = k_{\text{FSR}} \cdot (p_0 + (p_1 \cdot \log(p_T^{\text{ave}}))), \quad (7.19)$$

where p_T^{ave} corresponds to the mean p_T^{ave} in the corresponding $|\eta|$ bin.

Figure 7.11 illustrates the final residual correction computed with the dijet p_T balance and MPF method as a function of pseudorapidity. Possible biases due to the jet p_T resolution are taken into account using the smearing of the simulated jet p_T by the JER scale factors. A correction for additional jet activity is done using the k_{FSR} factors and the p_T dependence of the result is extrapolated parametrizing the response values in bins of η in a log-linear fit. The results from both methods agree well after the ISR+FSR correction. Some small differences are found in the forward region, due to the small number of events in data in this region, and in the border region between the endcap and the forward calorimeters. Hence, in the barrel region, the resulting relative resolution correction is found to differ from unity by around 2%, while in the forward up to 20%. The MPF method is less sensitive to ISR and FSR biases, and therefore its results are used as the central method for the L2 residual correction.

7.5.3. L2 residual correction in the forward detector region

The nominal triggers used to select dijet events in the forward region in data have large prescales during the data taking. Therefore, in order to get a better precision

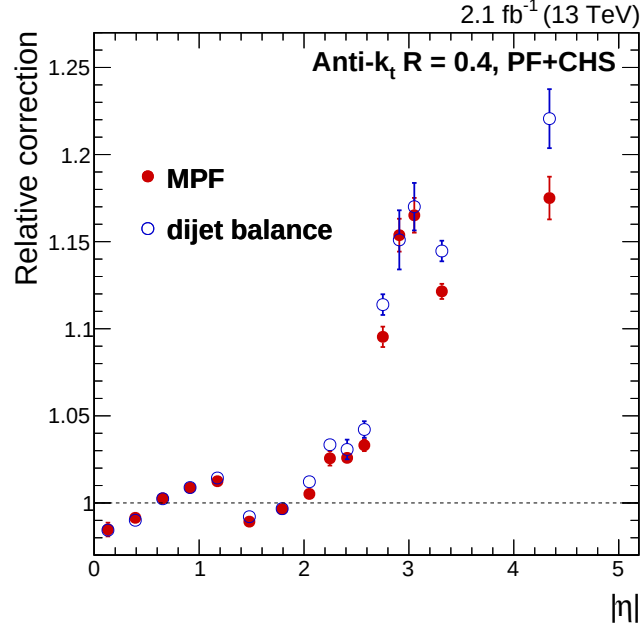


Figure 7.11.: The final residual correction computed with the dijet p_T balance and MPF method as a function of the probe jet pseudorapidity. The results are corrected for the JER bias, additional ISR + FSR activity, as well as p_T dependence.

of the obtained L2 residual correction, as well as to increase the number of events in data in the HF region, events fired with a special “HF” `HLT_DiPFJetAve*_HFJEC` triggers have been included in the analysis. Thus, the last three bins in $|\eta|$ are studied, using the additional events selected by the dijet HF triggers. In Tab. 7.6, event yields selected in each p_T^{ave} bin for the highest η region are shown. The p_T^{ave} bins boundaries are defined by the nominal and HF trigger threshold obtained as shown in Section 7.3.2. As expected, dijet HF triggers select more events in the forward region than the nominal ones. All selected events are used in each step of the described methods. The corresponding distributions, including the fit functions used for the k_{FSR} factors and p_T dependence extrapolation, are given in Appendix A.2. The effect of adding the events selected with the dijet HF triggers is shown in Fig. 7.12 in comparison to the final L2 residual correction shown previously.

The statistical precision in the forward region is found to be better, while in the narrow endcap-forward boundary region, the statistical uncertainty is larger. This is due to the fact that the statistical uncertainty is estimated by Gaussian error propagation of the statistical uncertainties in the fit parameters, obtained in the calculation of the L2Res correction. Since no large improvement was found by adding the dijet HF triggers, this analysis is performed using only the nominal dijet triggers.

Table 7.6.: The numbers of selected events in different bins of p_T^{ave} with corresponding trigger for the highest η -bins.

p_T^{ave} range	HLT trigger	number of events		
		$2.964 < \eta < 3.139$	$3.139 < \eta < 3.489$	$3.489 < \eta < 5.191$
56 – 78	HLT_DiPFJetAve40	2714	5724	5301
78 – 100	HLT_DiPFJetAve60	4733	9417	6925
100 – 168	HLT_DiPFJetAve60	2869	5307	2670
168 – 232	HLT_DiPFJetAve140	652	961	147
232 – 300	HLT_DiPFJetAve200	1517	1348	55
300 – 366	HLT_DiPFJetAve260	240	122	1
366 – 453	HLT_DiPFJetAve320	152	26	0
453 – 562	HLT_DiPFJetAve400	21	2	0
77 – 131	HLT_DiPFJetAve60_HFJEC	4217	7860	5548
131 – 154	HLT_DiPFJetAve80_HFJEC	587	1012	335
154 – 244	HLT_DiPFJetAve100_HFJEC	2329	3336	743
244 – 321	HLT_DiPFJetAve160_HFJEC	6008	4582	144
321 – 426	HLT_DiPFJetAve220_HFJEC	3352	1223	2

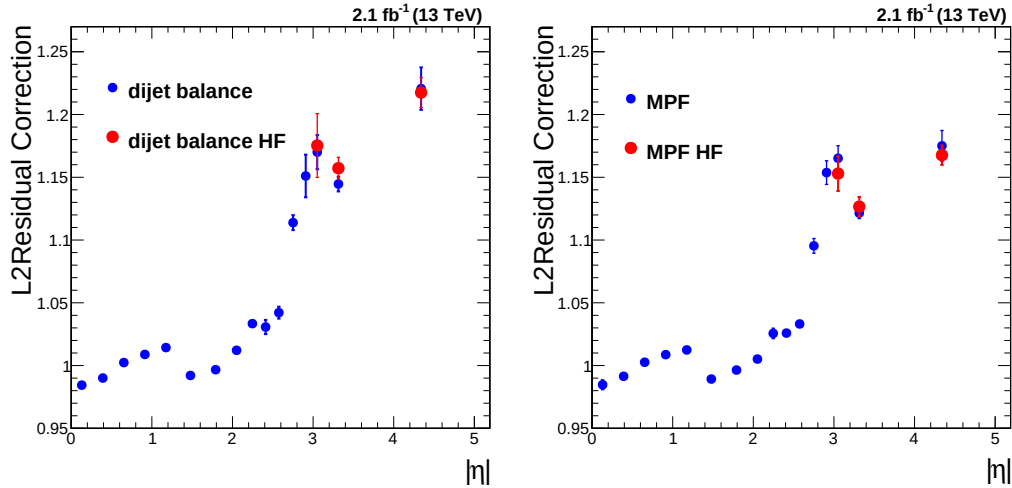


Figure 7.12.: The final residual correction calculated with the dijet p_T balance (left) and MPF (right) methods for events selected with nominal dijet and dijet HF triggers.

7.5.4. Validation of the pileup simulation

In order to test the pileup simulation, a validation test based on the dependence of the L2Res corrections on the number of reconstructed vertices is performed. This test provides a cross-check to the previous jet energy correction steps, especially to the pileup offset correction, which is designed to correct for the pileup influence on the jet energy response.

The validation is performed by comparing the response obtained with the dijet p_T balance and MPF methods to the number of primary vertices in data and MC. The comparison is made in each $|\eta|$ bin, as shown in Appendix A.2; an examples for one $|\eta|$ bin in the barrel and endcap region are shown in Fig. 7.13. No dependence

of the relative response on the number of primary vertices is observed, in both data and MC simulation.

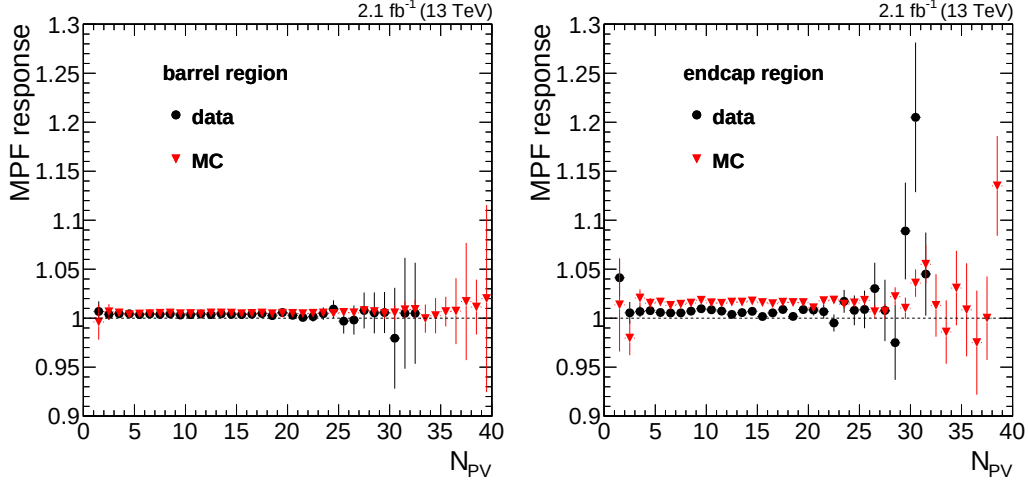


Figure 7.13.: Relative response calculated with the MPF method in data and MC as a function of the number of primary vertices for the $0.261 < |\eta| < 0.522$ (barrel region) and $1.93 < |\eta| < 2.172$ (endcap region).

7.6. L2 residual correction uncertainties

Since the L2 residual corrections are based on data, several biases are expected to come from the detector response and unaccounted for contributions to the dijet event, like the ISR and FSR. Although these effects were corrected for during the method implementation, it is still necessary to account for a possible systematic uncertainty. Thus, by comparing the nominal result with the result obtained for the corresponding systematic variation, a measure of the systemic uncertainty can be estimated. The statistical and systematical uncertainties are discussed in the following.

Statistical uncertainties

Since the parameters obtained by the fitting procedure are used for the L2 residual corrections calculation, the statistical uncertainty for each of them is taken into account during the estimation of the statistical uncertainty. The statistical uncertainty for L2 residual correction in each $|\eta|$ bin is estimated by a Gaussian error propagation of the statistical uncertainties in the fit parameters from the fit used to obtain the k_{FSR} factors, as well as from the p_T -dependent fit. The statistical uncertainties in the barrel region are found to be negligible. In the endcap region they correspond to 0.5%, and in the forward region the statistical uncertainty is about 2%.

Systematic uncertainties

Pileup: Simulated events are reweighted according to their pileup in order to match the data, as described in Section 7.3.3. The estimation of the number of pileup events in data is based on the total inelastic pp cross section of 69 mb. The uncertainty due to pileup is estimated by changing the cross section from the 69 mb to 80 mb, the first recommended value [202].

The impact of the pileup variation on the L2 residual is shown in Fig. 7.14 for the dijet p_T balance and MPF methods. The uncertainty due to pileup is found to be negligible in the barrel region, while for the dijet p_T balance method in the forward region a certain discrepancy is visible. Since, the final L3 residual correction is derived with the MPF method, the pileup uncertainty accounts for a negligible portion of the total systematic uncertainty.

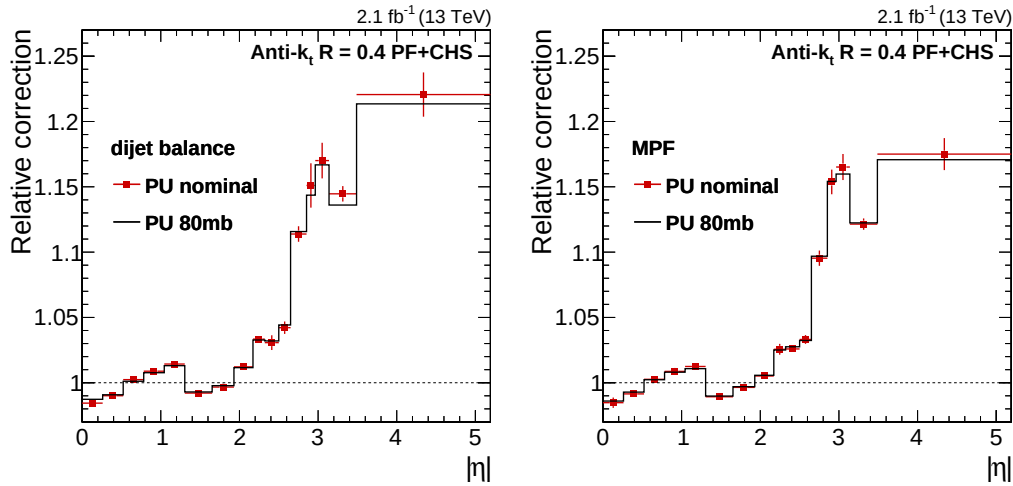


Figure 7.14.: The total L2 residual correction calculated with the dijet p_T balance (left) and MPF (right) methods. Results for two different pileup profiles are shown, where the total inelastic pp cross sections of 69 mb (red) and 80 mb (black) are used.

Jet p_T resolution: The jet p_T resolution (JER) of jets in data is different from simulation. The p_T of the reconstructed jets in the MC events are smeared to match the data, as discussed in Section 7.4. The JER depends on η and p_T , and is varied in simulation within it ± 1 standard deviation with respect to the resolution in data listed in Tab. 7.5.

Figure 7.15 shows the effect of the JER variation on the L2 residual correction factors. Minor changes in the result are observed in the barrel region, while for $|\eta| > 2.5$ an uncertainty of 3% for dijet p_T balance and 1% for MPF is found.

Time dependence: A set of nine data taking periods with approximately comparable integrated luminosity of about 220 pb $^{-1}$ are used in order to derive

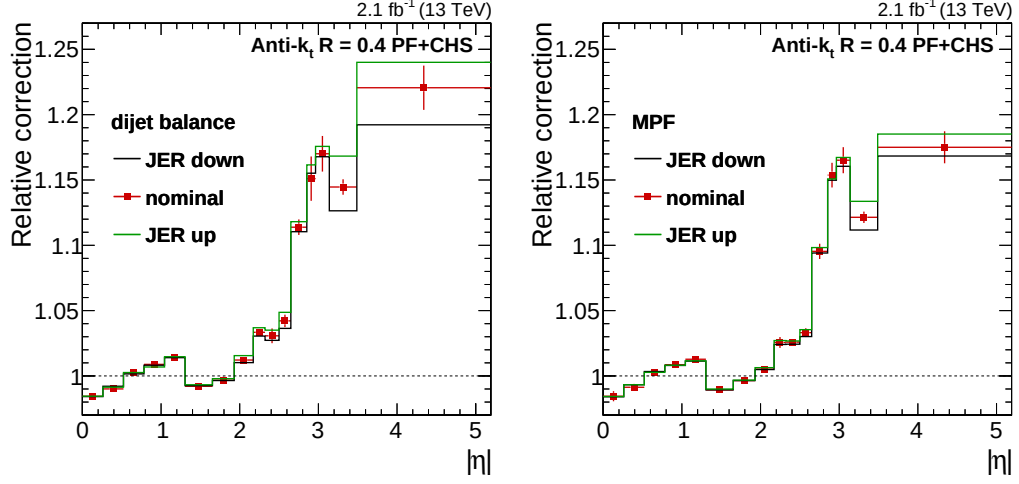


Figure 7.15.: The total L2Res correction calculated with the dijet p_T balance (left) and MPF (right) methods. Results are given for the central JER smearing (red) and for the up (green) and down (black) JER variation.

the uncertainty due to the time dependence of run conditions and detector functionality. More details on the corresponding data sets are given in Appendix A.2. A variation is expected to arise from changes in the energy response in the ECAL and HCAL, due to their cumulative radiation damage at high rapidities. The corresponding uncertainty is calculated by taking the root mean square of the L2 residual results obtained for nine data sets, which are shown in Fig. 7.16. As expected, there is almost no deviation in the barrel region, while for the region with $|\eta| > 2.5$, an uncertainty of 2 to 3% is found.

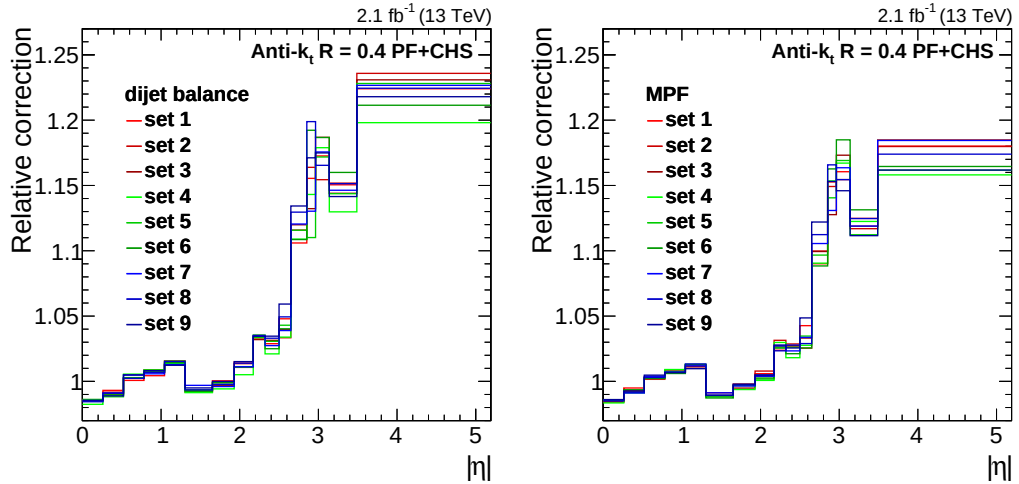


Figure 7.16.: The total L2Res correction calculated with the dijet p_T balance (left) and MPF (right) methods for nine data sets with average integrated luminosity of about 220 pb^{-1} each.

p_T dependence: The final L2 residual results are calculated according to Eq. 7.19, where the log-linear fit is used to extrapolate the p_T dependence of the result. Additionally, the p_T dependence can be estimated by using a constant fit according to Eq. 7.18. Thus, in order to account for possible uncertainties arising due to the choice of the p_T extrapolation fit function, half of the difference between the corrections derived with the log-linear and constant fit is used as a systematic uncertainty, which is found to be less than 1.5%.

Additionally, the average transverse momentum of dijet systems $p_T^{\text{ave}} \equiv \bar{p}_T$ used in the log-linear fit is varied by a factor of 2 up and down. Figure 7.17 shows that only in the endcap region some small deviation is visible and the corresponding systematic uncertainty is about 1%.

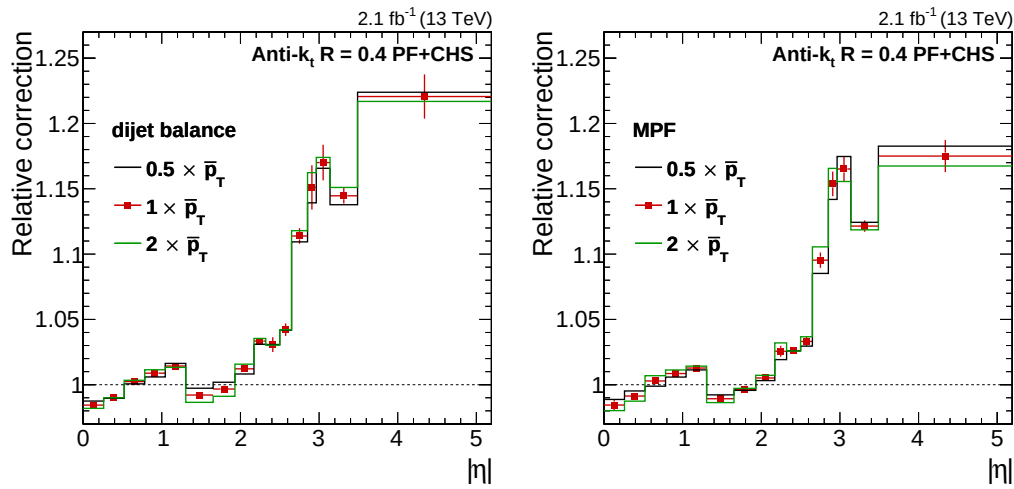


Figure 7.17.: The total L2 residual correction calculated with the dijet p_T balance (left) and MPF (right) methods and the variation by factor of 2 of the average transverse momentum used in Eq. 7.17.

Relative Balance: The results for L2 residual corrections obtained by the MPF and dijet p_T balance methods have good overall agreement. However, in the forward region, differences up to 2% are found. The difference in the final residual correction between both methods is covered by the relative balance systematic uncertainty, which is obtained as the bin-by-bin ratio of the two results. The corresponding uncertainty is used to calculate the final systematic uncertainty in the analysis. In previous studies at 8 TeV, this uncertainty was not used and the dijet method was considered for the cross check.

7.7. Results

The final L2 residual jet energy corrections measured with the MPF method are shown in Fig. 7.18. Statistical uncertainties of the result are on the order of 1% in the barrel region and 2% in the forward region. Systematic uncertainties are found to be on the order of 1% in the barrel region, while for the endcap and forward regions of the detector, the corresponding uncertainties increase to up to 5%. The dominant uncertainty source is given by the time dependence.

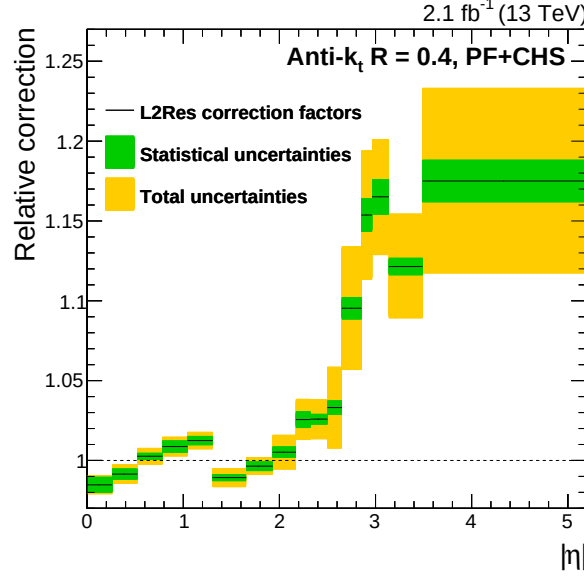


Figure 7.18.: The final L2 residual correction factors for the data collected at 13 TeV with CMS detector in 2015. The results are measured with MPF method. The green band corresponds to the statistic uncertainties of the results, while the orange band represents the systematic uncertainty.

7.8. Summary

The derivation of the L2 residual jet energy corrections are an essential part in the determination of the overall JEC and are successfully calculated in this study. The measurement is based on a data-driven method, that compares the jet energy response in data and MC, where the response is calculated with the dijet p_T balance and MPF methods. Biases that may occur due to the modeling of the jet p_T resolution, ISR and FSR and p_T distributions of the jets, were taken into account.

The L2 residual jet energy corrections were derived constantly during the 2015 data taking period and are available at [199]. The obtained JEC is provided to the CMS collaboration and used by many CMS analyses based on the data recorded in 2015.

The obtained residual corrections are in a good agreement with 8 TeV results [13]. The central value of the L2 residual correction scale factors in the last $|\eta|$ bin is slightly different from the one obtained at 8 TeV, but results agree within their uncertainty. In order to improve the obtained results, the analysis was performed using p_T binned simulated samples that have higher statistic. Moreover, the correction factor k_{FSR} was parameterized in terms of $|\eta|$ with the function given in [176]. The final results for the L2 residual JEC⁵ performed with 13 TeV data are summarized in [177].

⁵Note that presented L2 residual correction were performed by the author of this thesis together with Marc Stöver [191].

8. Top quark mass measurement

The precise measurement of the top quark mass m_t is a fundamental task of the LHC physics program. The accuracy of the m_t measurement has a direct impact on the limits set on new physics models, as well as on the standard model (SM) Higgs boson coupling to the top quark. The most accurate value of the top quark mass $m_t = 172.35 \pm 0.16$ (stat) ± 0.48 (syst) GeV [7] was measured by the CMS collaboration [6] using pp collision Run 1 LHC data at a center-of-mass energy $\sqrt{s} = 8$ TeV. The combined measurement of m_t from the ATLAS, CMS, CDF, and D0 experiments yield a world average mass of $m_t = 173.34 \pm 0.27$ (stat) ± 0.71 (syst) GeV [9]. The theoretical interpretations of the experimentally measured m_t value within different field-theoretical mass schemes are presented in [50, 203–207] and is discussed in Section 8.8.

With respect to the published CMS Run 1 results, new Monte Carlo (MC) event generators have been integrated and a larger amount of data has been collected at a higher center-of-mass energy, $\sqrt{s} = 13$ TeV. The MC is based on new next-to-leading order (NLO) matrix element (ME) calculations and a new parton shower (PS) implementation. Together with the newly available studies of the underlying event (UE) and the color reconnection (CR), the new samples are expected to provide better signal and background modeling, and thus give a more reliable estimation of the dominant theoretical uncertainties of the m_t measurement.

The top quark is the only quark that decays before it hadronizes. Therefore, it is possible to reconstruct the mass peak very well by measuring its decay products. This is implemented in the *direct methods* and gives the most accurate value of m_t . *Alternative methods* measure m_t based on other observables sensitive to the top quark mass, i.e., the B hadron decay length [208], and allow a reduction in the dependence on the detector calibration and to test the stability of the results. The latest results on the top quark mass measurements at CMS with direct and alternative methods are summarized in [209].

In this work, the most precise top quark mass measurement using the full data sample collected in 2016 by the CMS detector at a center-of-mass energy of 13 TeV is presented. The analyzed data corresponds to an integrated luminosity of 35.9 ± 0.9 fb⁻¹ [11]. The direct method of the m_t measurement is performed by analyzing top quark pair $t\bar{t}$ production with lepton+jets final states. The investigated quantity is measured with the *ideogram method* [12] that is implemented and evaluated by utilizing simulated events. The corresponding MC samples use refined NLO ME as

well as an optimized simulation of the jet activity in $t\bar{t}$ events based on the Run 1 data and implemented with new event generation parameters (event tune). Thus, the new MC generators allow a more refined estimation of the systematic modeling uncertainties, like color reconnection, initial- and final-state parton showers, and matrix element calculation, while the analysis techniques and final state are the same as in [7]. The large data set used in the current analysis results in a very small statistical error of the measurement and gives the opportunity to check various dependencies of m_t on the kinematic quantities with high statistical accuracy.

In this chapter, the analysis strategy is presented. The data samples are introduced, followed by the description of the selection of the muon+jets and electron+jets events. Afterwards, the reconstruction of the $t\bar{t}$ events is discussed. The implementation of the kinematic fit of the decay products to a $t\bar{t}$ hypothesis, which is utilized in order to improve the resolution of the reconstructed quantities, is also shown. Finally, the estimation of the top quark mass with an ideogram method and their respective systematic uncertainties are presented for the muon+jets, electron+jets, and combined lepton+jets channel.

8.1. Measurement strategy

The goal of this analysis is to provide the most accurate measurement of the top quark mass. This implies also having the best understanding of the systematic uncertainties sources. The analysis is performed in the lepton+jets final state, which has a high production rate and a relatively low background contamination. Even though the full reconstruction of the event is not possible due to the presence of the neutrino from the leptonically decaying top quark, the existence of the lepton improves the $t\bar{t}$ event selection. The hadronically decaying top quark can be fully reconstructed and is used for the top quark mass measurement, since the invariant mass of the corresponding jets is the most sensitive observable for this study. However, the presence of the four jets, two light jets from a W boson decay and two other jets associated to the b quarks, demands a proper jet-parton assignment. Thus, for the simulated $t\bar{t}$ events three possible parton-jet permutation categories can be defined using parton-level information. Only the leading four jets are considered.

$t\bar{t}$ correct The *correct permutations* (cp) refer to hypotheses with the correct match of the generated partons to the selected jets, by looking at the distance $\Delta R_{ij} = \sqrt{(\Delta\eta_{ij})^2 + (\Delta\phi_{ij})^2}$ between each quark i from the $t\bar{t}$ decay and each leading jets j such that $\Delta R_{ij} < 0.4$. In this case two jets are marked as light quark jets and are associated with the b tagged jet arising from the same top quark decay.

$t\bar{t}$ wrong The *wrong permutations* (wp) refer to hypotheses where each selected jet comes from the $t\bar{t}$ event, but the parton-jet association is performed wrong,

e.g., the b jets are swapped between the two top quarks.

$t\bar{t}$ unmatched The *unmatched permutations* (un) refer to hypotheses in which at least one quark from the $t\bar{t}$ pair decay can not be unambiguously matched to any of the four jets. This situation may occur, when one highly energetic jet from the underlying event was accidentally selected.

Of these categories, the proper parton-jet assignment results in the best resolution of the kinematic properties of the reconstructed objects and derived quantities, and thus the measurement of the top quark mass itself. The next step of the analysis is an application of the event-by-event kinematic fit to each $t\bar{t}$ hypothesis, discussed in details in Section 8.3.4. The fit is performed by applying additional constraints on the measured objects and minimizing a χ^2 function. The χ^2 is built from the four-vectors and the resolutions of the measured $t\bar{t}$ final-state observables, by varying the measured quantities within their resolution. This procedure improves the estimation of the compatibility of the reconstructed event with the $t\bar{t}$ hypothesis, resulting in an improvement in the resolution of the kinematic observables. The last step of the measurement strategy is to determine the most probable top quark mass by using the ideogram method, discussed in detail in Section 8.4.1. The ideogram method utilizes the fitted observables as well as the predicted permutation fractions for each event.

8.2. Data samples

Collision data

For the presented measurement, proton-proton (pp) collision data collected by the CMS detector at a center-of-mass energy of 13 TeV during Run 2 of the LHC are used. According to the LHC running conditions, the data collection period was separated into seven run periods. The corresponding single muon and single electron data samples used in this analysis are presented in Tab. 8.1 with the corresponding event numbers and luminosity estimates. The dataset¹ is certified by the DQM team [63, 211] and corresponds to an integrated luminosity of $\mathcal{L} = 35.9 \text{ fb}^{-1}$.

For each set of recorded data, the number of additional pp interactions (*pileup*) is calculated using the BRIL Work Suite software [201]. The best-fit value for the total pp cross section to get the measured PU distribution in data corresponds to 69.2 mb [175] with an uncertainty of 4.6%. Fig. 8.1 shows the calculated distribution of pileup events for the utilized data, and is given by the black solid line.

¹Cert_271036-284044_13TeV_23Sep2016ReReco_Collisions16_JSON.txt [210]

Table 8.1.: Single lepton datasets for Run 2 data at 13 TeV with a corresponding number of events (N_{events}) and integrated luminosity ($\mathcal{L}_{\text{int}}$).

Dataset name	N_{events}	$\mathcal{L}_{\text{int}}, \text{fb}^{-1}$
/SingleMuon/Run2016B-03Feb2017_ver2-v2/MINIAOD	158,145,722	5.784
/SingleMuon/Run2016C-03Feb2017-v1/MINIAOD	67,441,308	2.573
/SingleMuon/Run2016D-03Feb2017-v1/MINIAOD	98,017,996	4.248
/SingleMuon/Run2016E-03Feb2017-v1/MINIAOD	90,963,495	4.008
/SingleMuon/Run2016F-03Feb2017-v1/MINIAOD	65,489,554	3.102
/SingleMuon/Run2016G-03Feb2017-v1/MINIAOD	14,9916,849	7.541
/SingleMuon/Run2016H-03Feb2017_ver2-v1/MINIAOD	169,642,135	8.391
/SingleMuon/Run2016H-03Feb2017_ver3-v1/MINIAOD	4,393,029	0.215
$\Sigma(\text{Single Muon Data})$	804,010,088	35.862
/SingleElectron/Run2016B-03Feb2017_ver2-v2/MINIAOD	246,440,440	5.788
/SingleElectron/Run2016C-03Feb2017-v1/MINIAOD	97,259,854	2.573
/SingleElectron/Run2016D-03Feb2017-v1/MINIAOD	148,167,727	4.248
/SingleElectron/Run2016E-03Feb2017-v1/MINIAOD	117,321,545	4.009
/SingleElectron/Run2016F-03Feb2017-v1/MINIAOD	70,593,532	3.102
/SingleElectron/Run2016G-03Feb2017-v1/MINIAOD	153,330,123	7.540
/SingleElectron/Run2016H-03Feb2017_ver2-v1/MINIAOD	125,826,667	8.390
/SingleElectron/Run2016H-03Feb2017_ver3-v1/MINIAOD	3,191,585	0.215
$\Sigma(\text{Single Electron Data})$	962,131,473	35.866

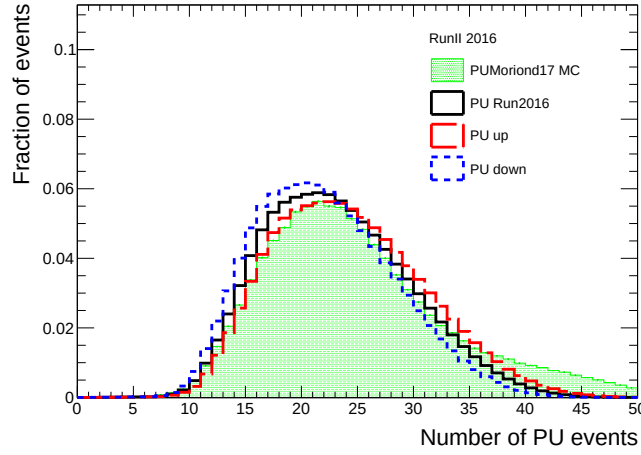


Figure 8.1.: Number of expected pileup interactions in the data collected during the full 2016 run period of Run 2 (solid black line). Number of pileup events in PUMoriond17 MC scenario (green filled area) and reweighted pileup distribution in simulation with up and down uncertainty variation (red and blue lines). The histograms are normalised.

Simulated Monte Carlo events

The understanding of the physics of the studied process is essential for the m_t measurement. Therefore, Monte Carlo (MC) events are simulated and used to develop the analysis techniques. The simulation of the $t\bar{t}$ events, together with corresponding MC generators, is presented in Chapter 4. In this section, the simulated samples are described.

Simulated signal $t\bar{t}$ events were generated with the POWHEG V2 matrix element generator [96–99] and the PYTHIA 8.219 parton-shower generator [16, 17] using the new CUETP8M2T4 underlying event (UN) tune [110] for seven different top quark mass values: 166.5, 169.5, 171.5, 172.5, 173.5, 175.5, and 178.5 GeV. The nominal input top quark mass is set to 172.5 GeV. All these samples are used for the top mass extraction method discussed in Section 8.4.1. All signal PYTHIA 8 samples are generated with the option “early resonance decay” off (erdOFF), which allows the color reconnection to take place before the top quark decay. The generator cross section is calculated using the parametrisation for the mass dependence of the top pair cross section proposed in [45]. Thus, the expected next-to-next-to-leading order (NNLO) $t\bar{t}$ cross section at center-of-mass energy of 13 TeV for the top quark mass of 172.5 GeV is equal to $831.76^{+19.77}_{-29.20}$ (scale) ± 35.06 (PDF+ α_s) pb. The parton distribution function (PDF) set NNPDF3.0 [83] at NLO ($\alpha_s = 0.118$) is used as the default parametrization of the PDF in the signal $t\bar{t}$ simulations. The MC samples include effects of additional overlapping minimum-bias events (pileup). The PUMoriond17 [202] pileup scenario is used in simulations. The corresponding distribution is shown in Fig. 8.1 as a green filled area.

Several background sources are relevant in this analysis, and are also simulated. The diboson [212] (WW, ZZ and WZ) and QCD multijet background events are generated using PYTHIA 8.219. The Drell-Yan and single top s-channel events are simulated with AMC@NLO v2.2.2 [95], additionally for the Drell-Yan case the FxFx merging [104] at next-to-leading order (NLO) level is applied. The MADGRAPH5 generator is used for the W+jets events at leading order (LO) with MLM matching [103]. The background events from tW [213] are generated using POWHEG v2 [214], while the single top t-channel background is simulated using POWHEG v2 and MadSpin [215, 216]. In all background simulations, the PYTHIA 8.219 parton-shower generator with the CUETP8M1 UN tune [116] is used, with the PDF set NNPDF2.3 [217] LO ($\alpha_s = 0.1365$). All events are further processed through a full simulation of the CMS detector based on GEANT4 [118, 119]. The MC campaign corresponds to the RUNIISUMMER16 version.

The list of signal and background samples used in the analysis are shown in Tab. 8.2 and 8.3. The $t\bar{t}$ MC samples used for the systematic variations are given in Tab. 8.4 and are discussed in Section 8.5.

Table 8.2.: List of of signal and background $t\bar{t}$ samples used in this analysis with the corresponding number of events (N_{events}) and cross sections (σ).

Dataset name	N_{events}	σ (pb)
$t\bar{t}$ signal		
POWHEG + PYTHIA 8 with CUETP8M2T4 tune and erdOFF		
TT central $m_t = 172,5$ GeV	78006311	832
TT $m_t = 166,5$ GeV	19380254	982
TT $m_t = 169,5$ GeV (ext1, ext2)	58542590	903
TT $m_t = 171,5$ GeV	19578812	855
TT $m_t = 173,5$ GeV	19419050	809
TT $m_t = 175,5$ GeV (ext1, ext2)	59384660	767
TT $m_t = 178,5$ GeV	16377176	708
Background		
Single top (PYTHIA 8 with CUETP8M1)		
ST_tW_top_5f_inclusiveDecays POWHEG	6952830	35,6
ST_tW_antitop_5f_inclusiveDecays POWHEG	6933094	35,6
ST_t-channel_top_4f_inclusiveDecays POWHEG	67240808	44.33
ST_t-channel_antitop_4f_inclusiveDecays POWHEG	38811017	26.38
ST_s-channel_4f_leptonDecays aMC@NLO	1000000	3.36
V+jets (PYTHIA 8 with CUETP8M1)		
WJetsToLNu MADGRAPH MLM (ext1)	86731806	61526.7
DYJetsToLL_M_50 aMC@NLO FFX	122055388	6025.2
DYJetsToLL_M_10to50 aMC@NLO FFX (ext1)	60626344	22635.1
Diboson (PYTHIA 8 with CUETP8M1)		
WW (ext1)	30694166	115.
WZ (ext1)	3995828	47.1
ZZ (ext1)	1988098	16.5

Table 8.3.: List of QCD multijet MC samples used for the 13 TeV top quark mass analysis.

Dataset name	Events	σ (pb)	filter eff-cy
QCD multijet (PYTHIA 8 with CUETP8M1)			
QCD Pt_15to20 MuEnrichedPt5	4141251	1273190000	0.003
QCD Pt_20to30 MuEnrichedPt5	31475157	558528000	0.0053
QCD Pt_30to50 MuEnrichedPt5	29954815	139803000	0.01182
QCD Pt_50to80 MuEnrichedPt5	19806915	19222500	0.02276
QCD Pt_80to120 MuEnrichedPt5 (ext1)	23584215	2758420	0.03844
QCD Pt_120to170 MuEnrichedPt5	8042721	469797	0.05362
QCD Pt_170to300 MuEnrichedPt5 (ext1)	17350231	117989	0.07335
QCD Pt_300to470 MuEnrichedPt5 (ext1, ext2)	48995686	7820.25	0.10196
QCD Pt_470to600 MuEnrichedPt5 (ext1, ext2)	19362943	645.528	0.12242
QCD Pt_600to800 MuEnrichedPt5 (ext1)	9981311	187.109	0.13412
QCD Pt_800to1000 MuEnrichedPt5 (ext1, ext2)	19767439	32.3486	0.14552
QCD Pt_1000toInf MuEnrichedPt5 (ext1)	13599938	10.4305	0.15544
QCD Pt_20to30 EMEnriched	9218954	557600000	0.0096
QCD Pt_30to50 EMEnriched (ext1)	11498579	136000000	0.073
QCD Pt_50to80 EMEnriched (ext1)	45811241	19800000	0.146
QCD Pt_80to120 EMEnriched (ext1)	77695287	2800000	0.125
QCD Pt_120to170 EMEnriched (ext1)	77771316	477000	0.132
QCD Pt_170to300 EMEnriched	11540163	114000	0.165
QCD Pt_300toInf EMEnriched	7373633	9000	0.15

Table 8.4.: List of the signal $t\bar{t}$ samples used for the systematic variations in the top quark mass analysis. See Section 8.5 for explanations.

Dataset name	Number of events	σ (pb)
$t\bar{t}$ systematic variations (POWHEG + PYTHIA 8 with TuneCUETP8M2T4) for underlying event		
TT scaleup (ext1)	58953660	832
TT scaledown (ext1)	58338240	832
for color reconnection		
TT erdON (ext1)	59882210	832
TT QCDbasedCRTune erdON (ext1)	59620206	832
TT GluonMoveCRTune erdON	56168970	832
TT GluonMoveCRTune erdOFF	59037234	832
for shower scales		
TT fsrup (ext1)	59230899	832
TT fsrdown (ext1)	59306906	832
TT isrup	59033604	832
TT isrdown (ext1)	58999580	832
for ME-PS matching		
TT hdampup (ext1)	58858606	832
TT hdampdown (ext1)	58163976	832
for ME generator		
TTJets TuneCUETP8M2T4 aMC@NLO FxFx	43561608	832
other $t\bar{t}$ MC generators		
TTJets TuneCUETP8M1 MADGRAPH MLM	10139950	832
TT TuneEE5C HERWIG++ (ext2, ext3)	59174465	832

8.3. Lepton+jets event selection and reconstruction

The event selection aims to obtain a clean sample of lepton+jets $t\bar{t}$ events based on the decay topology $t\bar{t} \rightarrow WbWb \rightarrow (l\nu_l b)(q\bar{q}b)$, when one of the two W bosons decays to one muon or electron and the corresponding neutrino, while the other W boson decays into two light flavor quarks, i.e., fully hadronically. Consequently, the final state consists of four jets, a muon or electron, and an undetected neutrino. It is noted, that events where a muon or electron from a τ lepton decay passes the selection criterias can also be included.

To reconstruct the final state particle candidates, the particle flow (PF) algorithm [142] is used, which combines the information from all CMS sub-detectors to identify and reconstruct individual objects. The reconstructed objects like muons, electrons, charged and neutral hadrons are going to be used in the event selection. The PF algorithm is discussed in more detail in Chapter 6. The selection strategy consists of the following steps:

- an online selection - triggering of the events during the data taking,
- a preselection - one isolated lepton and four jets are selected,
- a baseline - two jets must be associated to the hadronization of the b quarks,
- a final selection when additionally the kinematic fit to the $t\bar{t}$ hypothesis is used.

8.3.1. Online selection

The online selection is designed to reject uninteresting low-momentum events, thus reducing the CPU time and storage requirements for the online reconstruction. More details on the online event selection are listed in Section 7.3.2. The events are required to pass the single-muon HLT trigger with a minimum threshold on the transverse momentum (p_T) of an isolated muon of 24 GeV or the single-electron HLT trigger with a p_T threshold for isolated electrons of 32 GeV. Within the CMS experiment the trigger paths are denoted as `HLT_IsoMu24_v*` or `HLT_IsoTkMu24_v*` for the muon+jets case and `HLT_Ele32_eta2p1_WPTight_Gsf_v*` for the electron+jets one. The resulting trigger efficiency, determined by the tag-and-probe method, and the derived data-to-MC scale factors (the ratios of the efficiencies) are discussed in Section 6.3 and 6.4. The obtained trigger scale factors are used during the systematic uncertainty estimation (see Section 8.5) to account for any remaining data-to-MC differences.

8.3.2. Preselection

In the lepton+jets event topology, exactly one lepton assigned to the primary vertex and at least four jets are required. The leptons reconstructed by the particle flow

(PF) algorithm have to pass the isolation and identification requirements listed in Chapter 6. The charged hadron subtracted PF jets clustered by the anti- k_T algorithm with radius parameter of $R = 0.4$ are used. Additionally, jets are corrected using jet energy correction (JEC)² [199] and jet energy resolution (JER)³ [180] scaling factors, according to Section 6.5. Hence, the missing transverse energy from the undetected neutrino is corrected and so-called type-1 PF E_T^{miss} is used.

Events are required to have exactly one isolated muon with transverse momentum $p_T > 26$ GeV and pseudorapidity $|\eta| < 2.4$ or one isolated electron with $p_T > 34$ GeV and $|\eta| < 2.1$, which corresponds to the efficient region of the triggers. In addition, at least four jets with $p_T > 30$ GeV and $|\eta| < 2.4$ are selected, that have a distance parameter $\Delta R(l, \text{jet}) = \sqrt{(\eta_l - \eta_{\text{jet}})^2 + (\phi_l - \phi_{\text{jet}})^2}$ between muon (electron) and each jet larger than 0.4 (0.3). In order to ensure that no additional lepton is present, a veto on any additional muons or electrons is applied, according to the procedure written in Section 6.3 and 6.4. This helps to suppress the background contributions from $t\bar{t}$ dilepton decays as well as from Z+jets processes. No b-tagging is used at the preselection stage.

Comparing the distributions of the kinematic variables of the reconstructed objects, the effect of the applied selection criteria at each selection stage, as well as the data-to-MC difference, can be seen. In this work, the combined lepton+jets final state distributions are presented, while the individual cases for muon+jets or electron+jets final state are shown in Appendix A.3. For each data-to-MC comparison, the simulation is normalized to the luminosity by scaling the MC contribution to the in-situ measured total inclusive cross section. Additionally to the statistic limitations of the simulation samples, uncertainties from the signal and background processes cross sections are included. Thus, the uncertainties in the used cross-sections are 6% for $t\bar{t}$ signal, 30% for W+jets, 10% for single top, 100% for QCD multijet, and 10% for diboson and Z+jets background. In order to remove events with a very large weight in the simulation, the combined weight, which includes the pileup reweighting, lepton isolation and identification SF and trigger weights, is required to be less than 20 that reject less than 1% of the events. All possible permutations in the signal sample, e.g. $t\bar{t}$ correct, $t\bar{t}$ wrong and $t\bar{t}$ unmatched, as defined in Section 8.1, are considered. Additionally, the data-to-MC ratios with their corresponding uncertainties are presented.

Figure 8.2 shows the transverse momentum, pseudorapidity, lepton flavor, and the azimuthal angle distributions of the lepton observables, after the preselection criteria are applied. At this stage, a large contamination from background events is observed, with the largest fraction coming from W+jets and QCD multijet processes. The reconstructed data are well described by the simulation. More leptons are detected in the barrel region of the detector than in the forward one, which leads to a specific

²Summer16_23Sep2016AllV4_DATA, Summer16_23Sep2016V4_MC

³Spring16_25nsV10

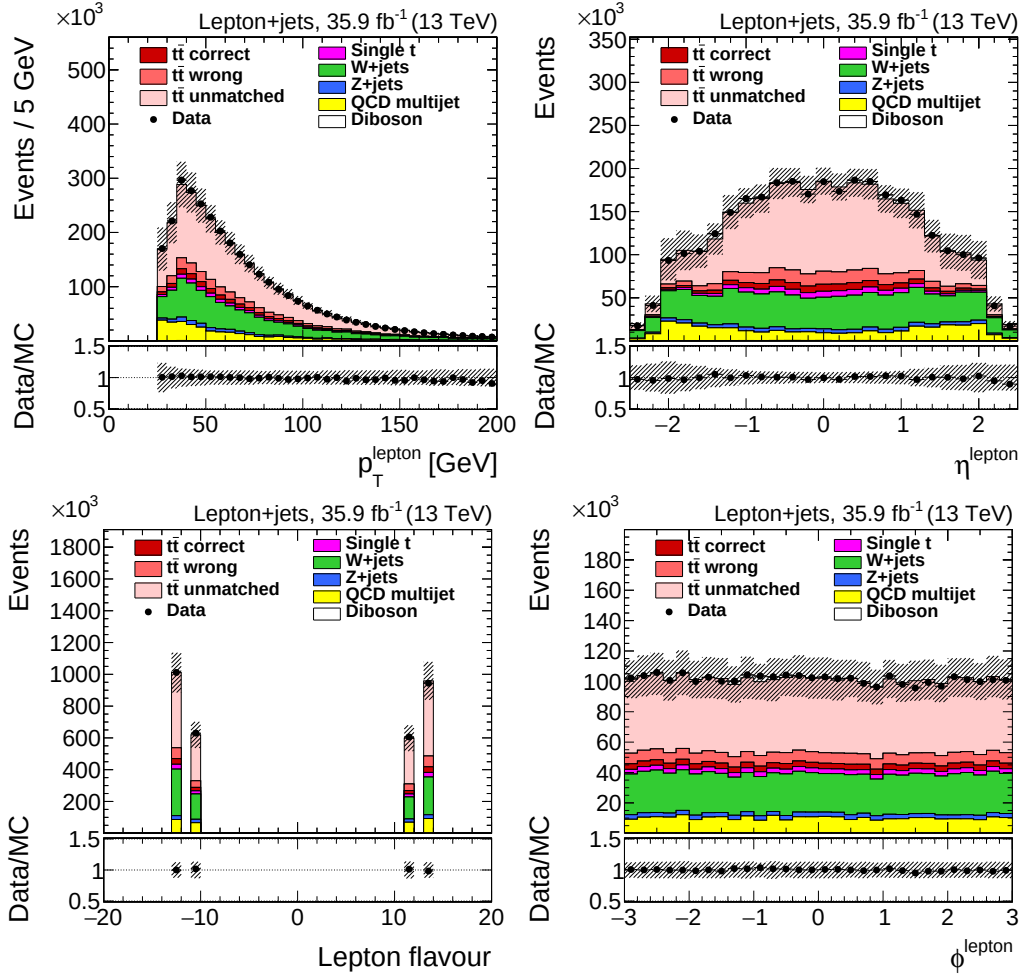


Figure 8.2.: Reconstructed lepton p_T (upper left) and η (upper right), together with the lepton flavor (lower left) and the azimuthal angle ϕ (lower right) for the lepton+jets channel after the preselection. The lepton flavor are -13, 13, -11, 11 for μ^+ , μ^- , e^+ , e^- , respectively. The plots are normalized to the luminosity. The hatched area indicates uncertainties from MC statistics and production cross sections.

shape of the pseudorapidity distribution. However the background processes, e.g., W +jets, show a flat behavior. This arises due to the charge asymmetry nature of proton-proton collisions at the LHC, which leads to the asymmetry between the pseudorapidity distributions of the charged leptons from W^+ and W^- bosons [218, 219] and to the smaller production rate of W^- compared to W^+ bosons. At larger rapidities, an enhancement in the W^+ distribution is observed, which is explained by the harder x in the valence sea quark scattering of two up quarks in the proton, since up quarks carry more proton momentum fraction than down quarks [220, 221].

The jet multiplicity, the missing transverse energy, as well as the distance between the selected lepton and the nearest jet in the $\eta - \phi$ space, are presented in Fig. 8.3. Additionally, the number of b-tagged jets among four leading jets that are going to be used in the next selection step are shown. Here the b-jets are identified by the

CSVv2 b-tagging algorithm [162]. More details on the b-tag approaches are given in Section 6.5. Overall, a good MC to data agreement is observed. However, for low

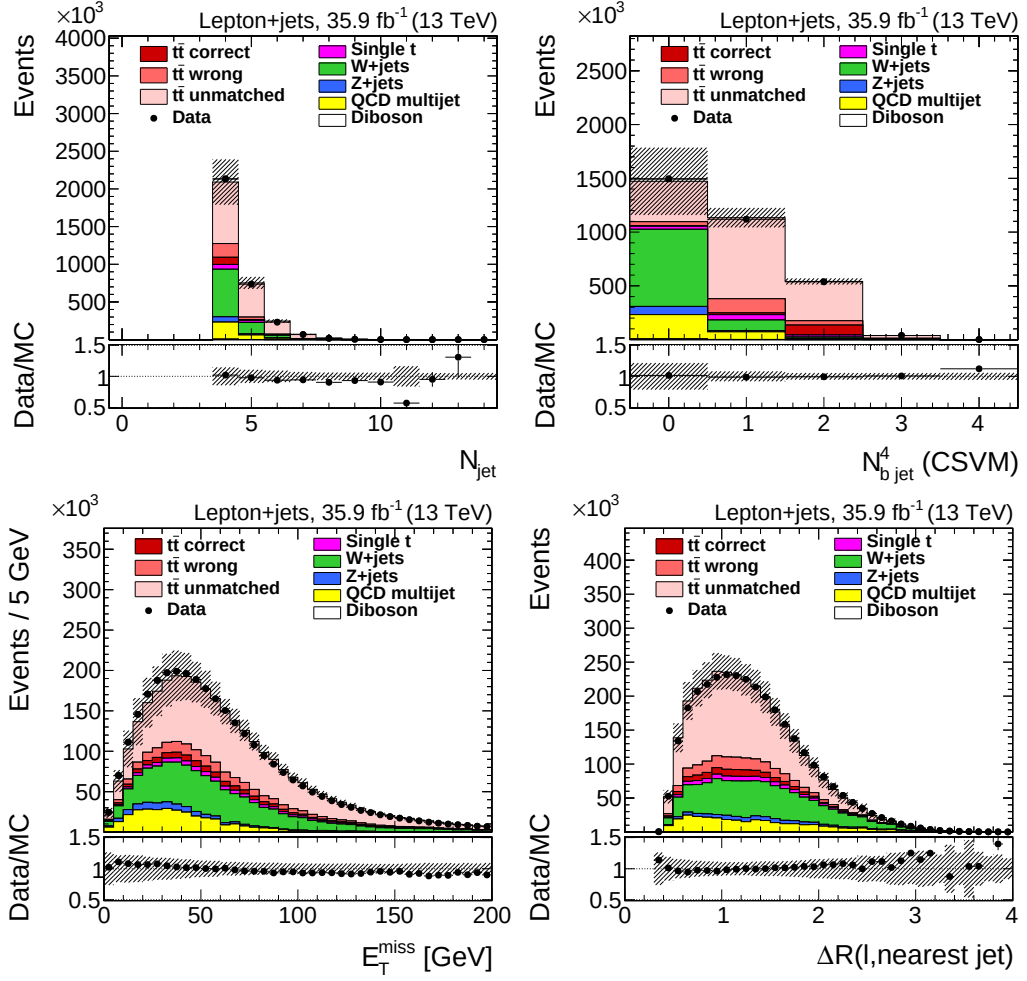


Figure 8.3.: Jet multiplicity for jets with $p_{\text{T}} > 30$ GeV (upper left), the number b-tags among the four leading jets with $p_{\text{T}} > 30$ GeV (upper right), the missing transverse energy (lower left), and minimum distance between selected lepton and four selected jets (lower right) for the lepton+jets channel after the preselection. The simulations are normalized to the integrated luminosity of the data sample.

values of $\Delta R(l, \text{nearest jet})$, a difference between data and MC is observed. This discrepancy arises from the electron+jets decay channel, as shown in Appendix A.3, presumably because of bremsstrahlung effects during the electron interaction with the tracker. The apparent differences in the multiplicity of the jets is due to the limited statistics in the MC.

The total number of events after the preselection step, as shown in Tab. 8.5 – 8.7 in the column “preselection”, is 3 189 837 in the lepton+jets channel. 1 954 682 events come from the muon+jets and 1 235 155 events come from the electron+jets channel. The total fraction of $t\bar{t}$ signal events, as estimated from the simulation, is above 58%. The largest background contamination relative to data comes from

W+jets and QCD multijet processes, 26% and 9.6%, respectively.

Table 8.5.: Event yields and fractions for signal and background processes for the lepton+jets channel. The uncertainties arise from MC statistics and cross-section. For the $t\bar{t}$ -dominated “baseline” and “final” selections, the uncertainties also contain contributions from jet energy scale and resolution, shower scale variation, pileup, ME-PS matching, underlying event, b fragmentation, B semi-leptonic BR and b tag SF variations.

	preselection (1 ℓ , 4 jets, $\Delta R(\ell, \text{jet})$)			baseline selection (+ 2b)			final selection (+ $P_{\text{gof}} > 0.2$)		
	events	$\pm$	[%]	events	$\pm$	[%]	events	$\pm$	[%]
$t\bar{t}$ +jets	1856854	81803	58.3	628900	26329	92.4	155297	5117	95.7
Z+jets	89643	9168	2.8	2388	386	0.4	326	109	0.2
W+jets	828184	248629	26.0	14558	4531	2.1	1171	489	0.7
Single top	97625	9766	3.1	23229	2326	3.4	3996	403	2.5
QCD multijet	306592	306802	9.6	11253	11426	1.7	1517	1743	0.9
Diboson	8090	815	0.3	335	39.8	0.0	44	9	0.0
MC total	3186988	403504	100	680663	29152	100	162351	5443	100
Data	3189837	3549	–	669109	1623	–	161496	797	–

Table 8.6.: Event yields and fractions for signal and background processes for the muon+jets channel. The uncertainties arise from MC statistics and cross-section. For the $t\bar{t}$ -dominated “baseline” and “final” selections, the uncertainties also contain contributions from jet energy scale and resolution, shower scale variation, pileup, ME-PS matching, underlying event, b fragmentation, B semi-leptonic BR and b tag SF variations.

	preselection (1 ℓ , 4 jets, $\Delta R(\ell, \text{jet})$)			baseline selection (+ 2b)			final selection (+ $P_{\text{gof}} > 0.2$)		
	events	$\pm$	[%]	events	$\pm$	[%]	events	$\pm$	[%]
$t\bar{t}$ +jets	1153278	50811	58.5	390443	16346	92.2	97684	3210	95.6
Z+jets	47493	4870	2.4	1156	208	0.3	146	59	0.1
W+jets	530197	159146	26.9	9450	2918	2.2	801	314	0.8
Single top	59486	5950	3.0	14186	1420	3.4	2490	251	2.4
QCD multijet	174519	174640	8.9	7990	8102	1.9	998	1125	1.0
Diboson	4844	487	0.2	195	23	0.0	24	5	0.0
MC total	1969817	241801	100	423420	18531	100	102143	3426	100
Data	1954682	1977	–	416288	912	–	101992	452	–

8.3.3. Baseline selection

The baseline selection aims to exclude contamination from the QCD related processes, e.g., QCD multijet events. Therefore, the b-tag algorithm, which helps to identify jets from b quarks, as described in Section 6.5, is used. For the $t\bar{t}$ event selection two b-jets among the leading jets, defined by the *Combined Secondary Vertex version 2* (CSVv2) algorithm [162], are required. The CSVv2 combines the information about the number of reconstructed secondary vertex candidates, number of tracks in the jet and ΔR between the secondary vertex flight direction and

Table 8.7.: Event yields and fractions for signal and background processes for the electron+jets channel. The uncertainties arise from MC statistics and cross-section. For the $t\bar{t}$ -dominated “baseline” and “final” selections, the uncertainties also contain contributions from jet energy scale and resolution, shower scale variation, pileup, ME-PS matching, underlying event, b fragmentation, B semi-leptonic BR and b tag SF variations.

	preselection (1 ℓ , 4 jets, $\Delta R(\ell, \text{jet})$)			baseline selection (+ 2b)			final selection (+ $P_{\text{gof}} > 0.2$)		
	events	$\pm$	[%]	events	$\pm$	[%]	events	$\pm$	[%]
$t\bar{t}$ +jets	703576	30992	57.8	238457	9983	92.7	57612	1907	95.7
Z+jets	42150	4298	3.5	1232	178	0.5	180	49	0.3
W+jets	297987	89482	24.5	5108	1613	2.0	371	174.7	0.6
Single top	38139	3816	3.1	9043	906	3.5	1506	152	2.5
QCD multijet	132074	132161	10.9	3263	3324	1.3	518	617	0.9
Diboson	3246	328	0.3	141	17.1	0.1	21	4	0.0
MC total	1217172	162688	100	257244	10685	100	60208	2019	100
Data	1235155	1571.6	–	252821	711	–	59504	345	–

the jet axis in the $\eta - \phi$ space. The *medium* working point that corresponds to the discriminator cut of 0.8484 and denotes a mistag rate for light quark of 1%, where light-flavor jets are misidentified as b-jets. For this working point, the CSVv2M algorithm has a b-tagging efficiency of approximately 69% [184].

Since the b-tagging efficiency in data and simulation is different, a weight to the selected MC events is applied in order to predict the correct event yields in data. In the current analysis an event reweighting using scale factors and MC b-tagging efficiencies [222] is used. This procedure is safe for the migration of events from one b-tag multiplicity bin to the other. The event weight is calculated as a ratio of probabilities for a given configuration of jets in data and MC event simulation, $\omega = P(\text{data})/P(\text{MC})$. The probabilities are defined as:

$$P(\text{data}) = \prod_{i=\text{tagged}} \text{SF}_i \varepsilon_i \prod_{j=\text{not tagged}} (1 - \text{SF}_j \varepsilon_j), \quad (8.1)$$

$$P(\text{MC}) = \prod_{i=\text{tagged}} \varepsilon_i \prod_{j=\text{not tagged}} (1 - \varepsilon_j), \quad (8.2)$$

where i and j go through each of the four leading jets, scale factors (SFs) are defined as the ratio of the data to MC b-tagging efficiencies, and ε_i is the MC b-tagging efficiency. Both variables are η , p_T and jet-tag flavor dependent. Thus, for each flavor of the jets the b-tagged SFs are available. An example for the SFs used in the $t\bar{t}$ event selection to correct the contribution of the light jets is presented in Fig. 6.14.

To check the efficiency of the baseline selection, the same lepton and jet observables as for the preselection are shown in Fig. 8.4 and 8.5.

A good agreement between the data and MC distributions is found. As expected,

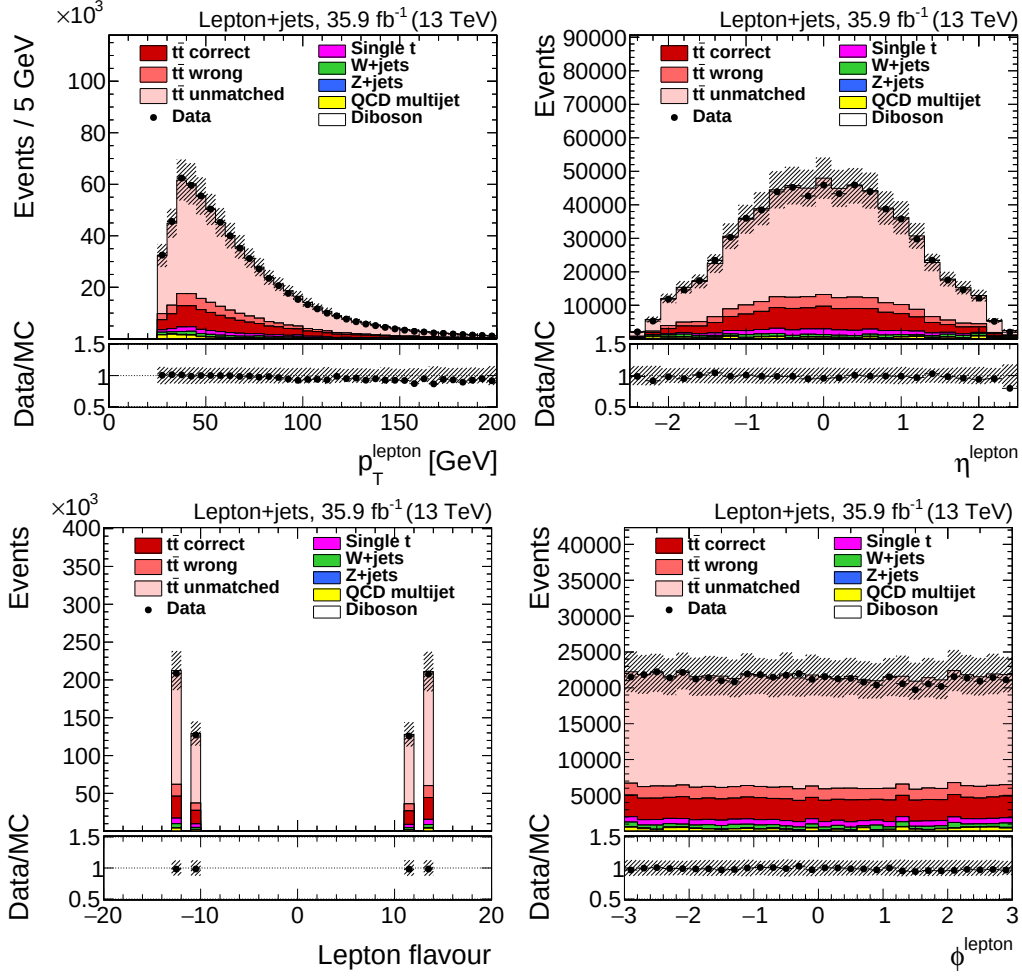


Figure 8.4.: The reconstructed lepton p_T (upper left) and η (upper right), together with the lepton flavor (lower left) and the azimuthal angle ϕ (lower right) for the lepton+jets channel after the baseline selection. The lepton flavor are -13, 13, -11, 11 for μ^+ , μ^- , e^+ , e^- , respectively. The plots are normalized to luminosity. The hatched area indicates uncertainties from MC statistics, cross-section, pile-up, jet energy scale and resolution, $\mu_{R,F}$, and top- p_T reweighting.

the presence of the background events is very much reduced, and thus a clean $t\bar{t}$ sample is obtained. In contrast to the preselection plots, the uncertainty bands take into account not only MC statistics and cross section uncertainties, but also the systematic uncertainties for shower scale variation, pileup, jet energy scale and resolution, ME-PS matching, underlying event, b fragmentation, B semi-leptonic BR and b tag SF variations, which are described in detail in Section 8.5. The requirements of two b-tagged jets suppresses the events where zero or one b-tagged jets are found, as demonstrated by the distribution of the multiplicity of the b-tagged jets among four leading jets distribution. However, the situation when a fifth jet is b-tagged is still possible, and some events are shown in the second bin of the $N_{b \text{ jet}}^4$ distribution. Additionally, the transverse momenta of the first six leading jets are

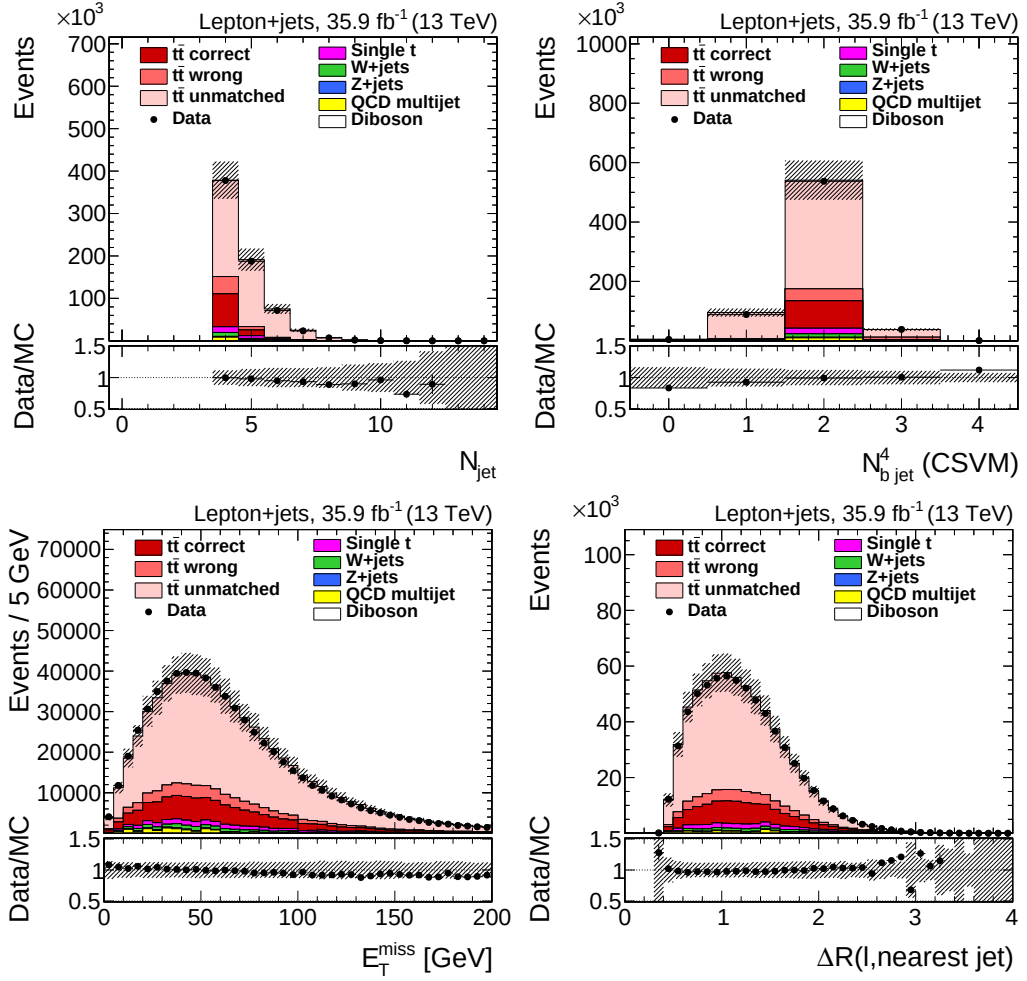


Figure 8.5.: The jet multiplicity, for jets with $p_T > 30$ GeV (upper left) and the number b-tags among four leading jets with $p_T > 30$ GeV (upper right), the missing transverse energy (lower left) and minimum distance between selected lepton and four selected jets (lower right) for the lepton+jets channel after the baseline selection. The simulations are normalized to the integrated luminosity of the data sample.

shown in Fig. 8.6, that appears to be slightly softer in data than in simulation, which may be due to the modelling of the top quark p_T that propagates to its final-states.

Once the baseline event selection and correction of MC simulations are performed, the event yields in data and in MC can be estimated, as shown in Tab. 8.5 to 8.7 in the column “baseline selection”. In the lepton+jets channel 669 109 data $t\bar{t}$ candidate events are selected, where 416 288 come from muon+jets and 252 821 from electron+jets channels. The fractions of the individual signal and background processes contributions are changed significantly. Already at this stage a very clean sample of $t\bar{t}$ events is found. From simulation, the event composition is expected to be 92.4% $t\bar{t}$, 3.4% single top quark, 2.1% W+jets, and 1.7% QCD multijet events. Other processes like diboson production and Z+jets background are found to be

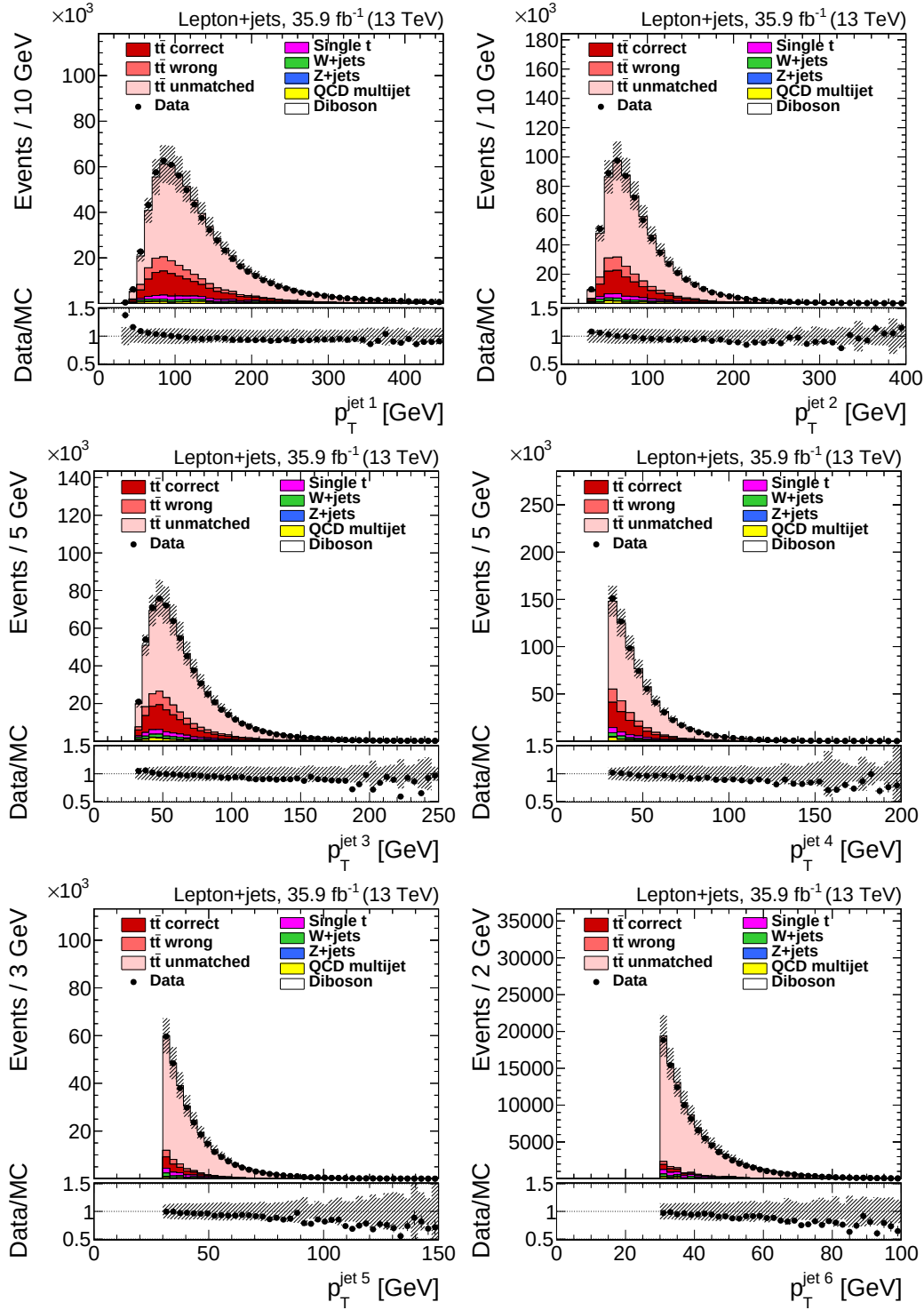


Figure 8.6.: The reconstructed transverse momenta of the six leading jets for the lepton+jets channel after the baseline selection. The simulations are normalized to the integrated luminosity of the data sample.

negligible.

8.3.4. Kinematic reconstruction of the top quark mass

The invariant mass of two untagged jets yield the hadronically decaying W boson mass m_W^{reco} , while the mass m_t^{reco} is computed from the untagged jets and each of the two b-tagged jets. Both masses are shown in Fig. 8.7 and the mass for the correct permutation has a Gaussian shape, while the masses for the wrong and unmatched permutations have a worse resolution, showing a contribution at higher masses that is visible as a tail. Thus, at this stage the fraction of the correct permutation in the reconstructed top quark mass for the lepton+jets channel is found to be only 14.9%, as is shown in Tab. 8.8. The corresponding permutation fractions for the muon+jets and electron+jets channels are presented in Appendix A.4.

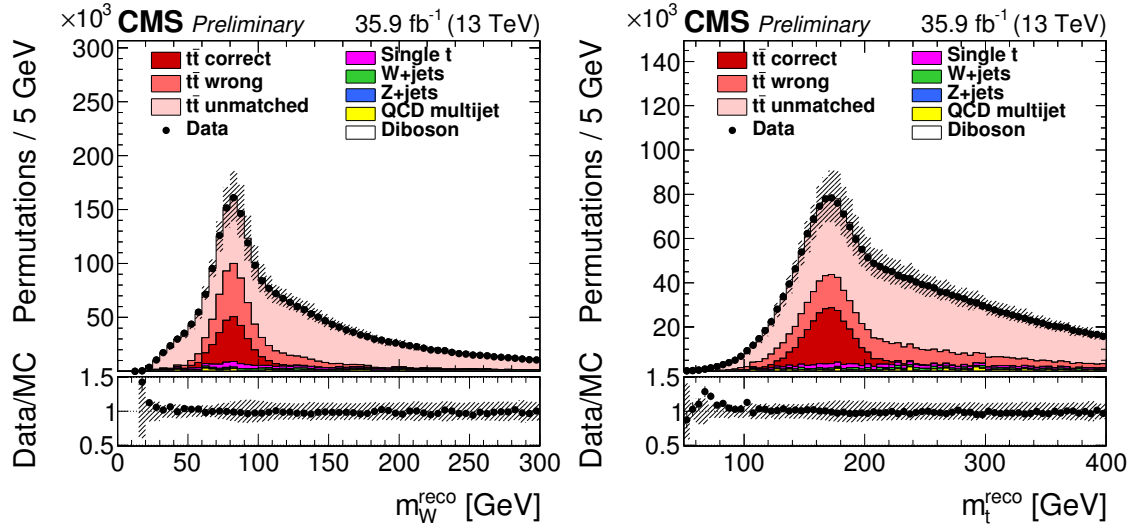


Figure 8.7.: The invariant mass m_W^{reco} of the two untagged jets (left) and invariant mass m_t^{reco} of the two untagged jets and one of the b-tagged jets (right) after the b tagging requirement for lepton+jets channel [10]. The vertical bars show the statistical uncertainty and the hatched bands show the statistical and systematic uncertainties considered in Section 8.5 added in quadrature. The bottom panels show the ratio of the yields between the collision data and the simulation. The simulations are normalized to the integrated luminosity of the data sample.

In order to obtain a better reconstruction of the mass distribution, a *kinematic fit* [223] is applied to the events passing baseline selection. The fit performs a compatibility check of the selected events with a $t\bar{t}$ hypothesis, and thus improves the resolution of reconstructed quantities.

The kinematic fit uses the four-vectors and the resolutions of the final-state objects as an input parameters. Thus, for the lepton+jets in the $t\bar{t}$ decay, one lepton (μ or e), four leading jets (two b-jets and two light jets) and neutrino are considered.

Table 8.8.: Permutation fractions for different for the lepton+jets channel after the baseline, final and effective selection.

	baseline	final $P_{gof} > 0.2$	effective $\times P_{gof}$
$t\bar{t}$ correct	14.9 %	45.7 %	48.0 %
$t\bar{t}$ wrong	8.0 %	15.6 %	16.4 %
$t\bar{t}$ unmatched	77.1 %	38.7 %	35.6 %

Therefore, to constrain the kinematic equations associated to the signal topology, the presence of the 24 or 18 measurements, i.e., the three momentum of each object⁴, is required. Here the leptons, neutrinos and light quarks are assumed to be massless, while the mass of the b quark⁵ is set to 4.7 GeV. Moreover, to represent the transverse momentum of the neutrino, the missing transverse energy is required. However, missing transverse energy is highly correlated with other energy measurements, since it is calculated from the the measured properties of all other objects. Instead, the x and y component of the transverse momentum of the $t\bar{t}$, defined by the vector $\vec{k}_T$

$$\vec{k}_T = \vec{E}_T^{\text{miss}} + \vec{p}_T^{\text{lep}} + \sum_{i=1}^{4 \text{ jets}} \vec{p}_{T,i}^{\text{jet}} \quad (8.3)$$

is used as a fit variables. The longitudinal component of the missing momentum is calculated by applying a kinematic constraint on the masses of the decaying particles. The invariant mass of the lepton-neutrino pair as well as one of quark-antiquark pair is fixed to the mass of W boson⁶, $m_W = 80.4 \text{ GeV}$. Furthermore, the top quark and top antiquark are required to have equal masses. Having the W boson mass and top quark mass constraints, the neutrino longitudinal momentum p_z^ν , can be obtained as shown in [224]. Moreover, for each b-jet assignment, a corresponding value of p_z^ν is calculated, which doubles the number of permutations in the event.

The defined vectors are used by the HITFIT [224] fitting package, which constructs a χ^2 function for each parton-jet permutation,

$$\chi^2 = (\mathbf{x}^f - \mathbf{x}^m)^T \mathbf{R} (\mathbf{x}^f - \mathbf{x}^m). \quad (8.4)$$

The vector of measured and fitted observables is given by $\mathbf{x}^m$ and $\mathbf{x}^f$, respectively. These observables are the kinematic properties, e.g., transverse momentum p_T , pseudorapidity η and azimuthal angle ϕ , of the input objects, like lepton, neutrino, b jets and light jets. The inverse error matrix $\mathbf{R}$ that contains the resolutions of the

⁴If the mass-energy relation is accounted for.

⁵The energy component of the four-vector is set to the desired mass, while the momentum components of the b-jet are taken as a measurement. Since the b-jet mass is much smaller than the jet energies, that are required to be larger than 30 GeV, this has no large influence on the final result

⁶The uncertainty in the W boson mass of 0.02% [18] is negligible for this analysis.

observables is given by

$$\mathbf{R} = \begin{pmatrix} \sigma_{\mu, p_T}^{-2} & 0 & \cdots \\ 0 & \sigma_{\mu, \eta}^{-2} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}, \quad (8.5)$$

and considers all resolutions to be Gaussian. All errors are assumed to be uncorrelated.

The HITFIT package fits events to a $t\bar{t}$ hypothesis by minimizing a χ^2 to the constraints, namely $m_{W_{\text{lep}}} = m_{W_{\text{had}}} = 80.4 \text{ GeV}$ and $m_t = m_{\bar{t}}$, by varying the measured observables within their resolution. To perform this, the method of Lagrange multipliers is used [225], which adds the Lagrange multiplier for the constants to the χ^2 function

$$\chi^2 = (\mathbf{x}^f - \mathbf{x}^m)^T \mathbf{R} (\mathbf{x}^f - \mathbf{x}^m) + 2\boldsymbol{\lambda}^T \mathbf{C}. \quad (8.6)$$

The vector of the Lagrange multipliers is given by $\boldsymbol{\lambda}(\lambda_1, \lambda_3, \lambda_3)$ and the vector of constants is given by $\mathbf{C}(C_1, C_2, C_3)$, where the components are given by

$$C_1 = \frac{(m_{l\nu}^{\text{reco}} - m_W)^2}{\Gamma_W^2}, \quad (8.7)$$

$$C_2 = \frac{(m_{q\bar{q}'}^{\text{reco}} - m_W)^2}{\Gamma_W^2}, \quad (8.8)$$

$$C_3 = \frac{(m_{bq\bar{q}'}^{\text{reco}} - m_{b\nu}^{\text{reco}})^2}{\Gamma_t^2}. \quad (8.9)$$

For the W boson and top quark, the finite decay width $\Gamma = 1$ is neglected. The minimization is performed with a given starting step size and continues until the constraint sum does not decrease anymore after several iterations. Thus, the case when the constraints are ideally fulfilled, e.g., when $C_1 = 0$, and the observables of the input objects have minimal changes, corresponds to the fit minimum and hence the solution is reached. In this case, HITFIT returns for each permutation the fitted observables and their kinematic properties. The quality of the fit is given by the χ^2 value and a goodness of fit probability (gof)⁷ is evaluated via $P_{\text{gof}} = \exp(-\frac{1}{2}\chi^2)$.

Figure 8.8 proves that the mass of the top quark after the kinematic fit has an improved resolution, although the fractions of wrong and unmatched permutations are still large. The fitted W boson mass is always equal to 80.4 GeV, since its mass is constrained in the kinematic fit. To increase the fraction of correct permutation, one can look at the distribution for the goodness-of-fit probability for the kinematic fit P_{gof} . Since the events with correct jet assignment will naturally fulfill the kinematic constraints, the obtained probability P_{gof} will be flat. In contrast to this, the events with wrong and unmatched permutations have rather poor fit quality and χ^2 probability showing up as the sharp peak at zero, as is illustrated in Fig. 8.9 for the

⁷In case of lepton+jets channel P_{gof} has two degrees of freedom

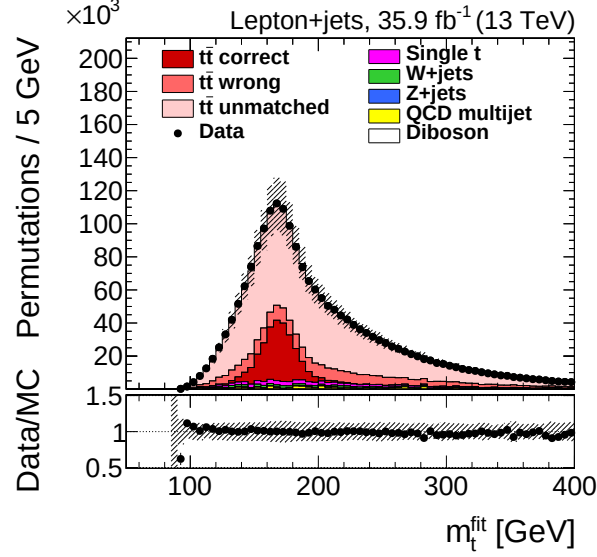


Figure 8.8.: The fitted top quark masses m_t^{fit} for the lepton+jets channel after the baseline selection. The vertical bars show the statistical uncertainty and the hatched bands show the statistical and systematic uncertainties considered in Section 8.5 added in quadrature.

kinematic fit. Therefore, by weighting the permutation with P_{gof} , the high-mass tail for the m_t^{fit} distribution is reduced. However, a more efficient way to increase the fraction of events with correctly assigned jets is to introduce the minimum of the P_{gof} , even though the statistics will be reduced.

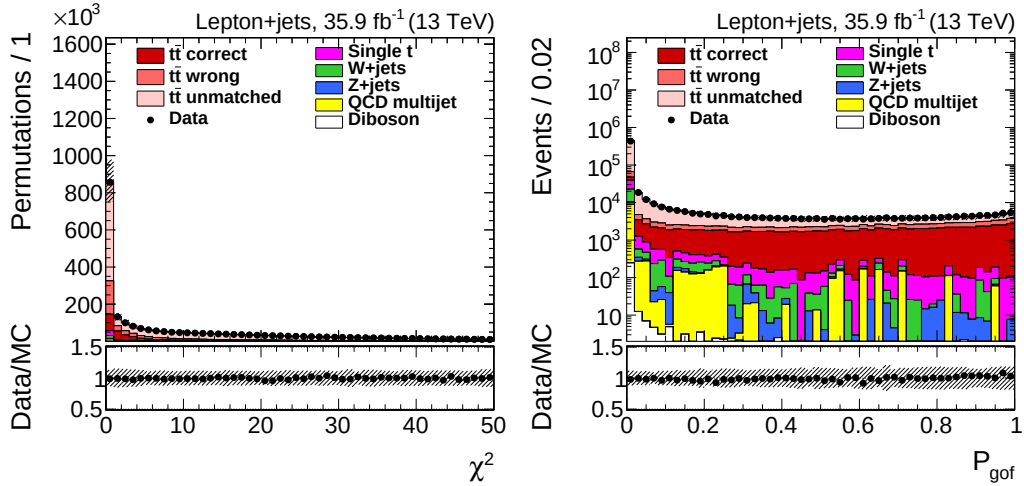


Figure 8.9.: The χ^2 of the kinematic fit (left) and the goodness-of-fit probability P_{gof} (right) for the lepton+jets channel after the baseline selection. The number of permutations in simulation is normalized to the number observed in data.

In order, to find the optimal value for the P_{gof} minimum, a study of the signal fraction and signal significance for the different values of the goodness-of-fit probability is performed. Events with correct permutations are denoted as signal S , and the

events with wrong and unmatched permutations are denoted to be background B . Thus, $S/(S+B)$ and $S/\sqrt{B}$ characterize the signal fraction and signal significance⁸ in the selected events, respectively, and are shown in Fig. 8.10. The minimum cut

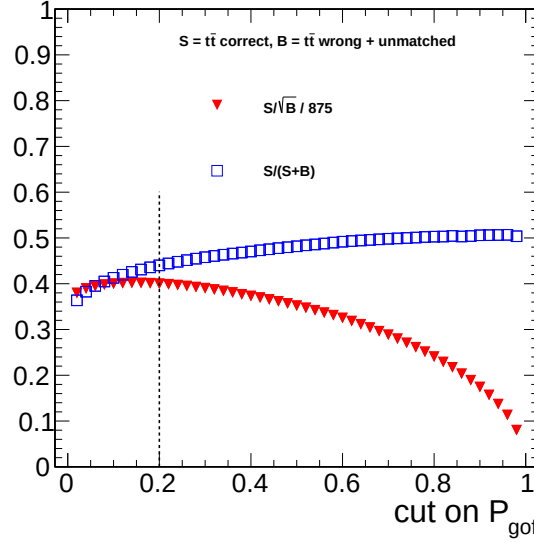


Figure 8.10.: Signal-to-background performance for a cut on P_{gof} . The signal significance $S/\sqrt{B}$ is scaled by $1/875$ to be shown together with the signal fraction $S/(S+B)$.

on P_{gof} should be optimized for the best signal significance as well as for the larger signal fraction. The given distribution of the signal significance shows a shallow optimum and falls for larger P_{gof} , while the signal fraction increases constantly. Thus, the minimum value for the P_{gof} cut should be optimal for $S/\sqrt{B}$. The results of the two requirements for the goodness-of-fit probability is studied: $P_{gof} > 0.15$ increases the fraction for the $t\bar{t}$ correct permutation to 43.8%, while for $P_{gof} > 0.2$, 44.9% of events with correct parton-jet assignments are found.

8.3.5. Final selection

To perform the final selection, exactly two b-tagged jets are required among the four leading jets. These b-tagged jets are used as candidates for the b quarks in the $t\bar{t}$ hypothesis in the kinematic fit. Two untagged jets are used as candidates for the light quarks from the W boson decays.

According to the study of signal significance, the goodness-of-fit probability is required to be at least 0.2, which increases the fraction of correct permutations estimated from simulation from 14.9% to 44.9% as shown in Tab. 8.8 the column “final”. The weighting of the permutation by their P_{gof} values yields an *effective* fraction of correct permutations of 48% for the mass extraction.

⁸signal significance $\approx \frac{S}{\sqrt{S+B}}$, if $S \ll B$ then $S+B$ can be approximated by B

Since two parton-jet assignments per event and two solutions for the z-component of the neutrino momentum are possible, the number of permutations after applying the kinematic fit will increase. It was found that the selection yields 155 297 events for the mass measurement, containing 419 565 permutations. While, the number of reconstructed data events in the lepton+jets channel correspond to 161 496, where 101 992 come from muon+jets and 59 504 from electron+jets channels, as shown in Tab. 8.5 – 8.7 in the column “final selection”. From simulation, the event composition is expected to be 95.7% $t\bar{t}$, 2.5% single top quark and all other background processes are found to be negligible.

The reconstructed W boson mass m_W^{reco} and the top-quark mass from the kinematic fit m_t^{fit} , after the P_{gof} selection, but without P_{gof} weight are shown in Fig. 8.11 for all possible permutations. These two observables are used in the mass extraction.

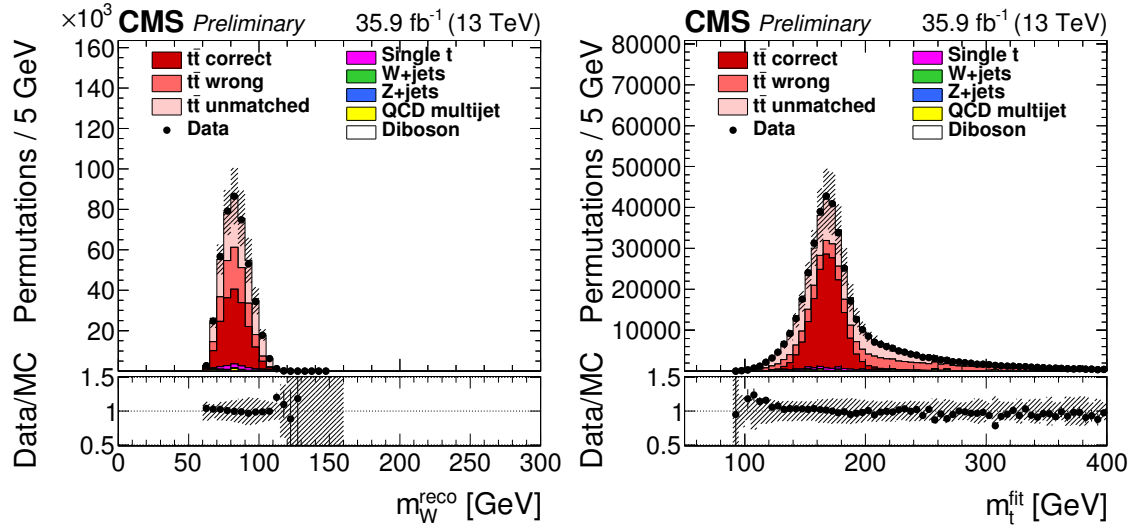


Figure 8.11.: The reconstructed W boson masses m_W^{reco} (left) and fitted top quark masses m_t^{fit} (right) after the goodness-of-fit selection for the lepton+jets channel [10]. The simulations are normalized to the integrated luminosity of the data sample.

8.4. Top quark mass extraction

The top quark mass can be extracted with various methods. The most popular are the template, the matrix element, the ideogram, and the analytical matrix-weighting methods.

Template method [226] uses template distributions for variables sensitive to different top quark mass input values, derived from simulated events. The template distributions are fitted with a corresponding fit functions and the fit parameters that have mass dependence are used to define the *likelihood* function with m_t as the only free parameter. Finally, the likelihood fit is applied to data, and the most likely top quark mass is measured by looking into the negative logarithm of the likelihood function relative to its maximum.

Matrix element method [227] aims to minimize the statistical uncertainty by estimating the probability densities for a selected event to be a genuine $t\bar{t}$ processes. The probability is estimated from the predictions of the differential cross-sections of the studied processes and the detector resolutions. Thus, the core of this probability is the matrix element that is used to give the differential cross-sections of the corresponding process. The probability is calculated as a function of the unknown parameter values to be measured and can be used to determine several unknown quantities simultaneously. The matrix element technique uses full event kinematic information and allows signal-like events to contribute more to the final result. Finally, the most probable values of the unknown quantities are those where the probability is maximal.

Ideogram method [12] combines the template and matrix element techniques. It constructs a two-dimensional event-by-event likelihood function with a jet energy scale factor (JSF) left as a free parameter in addition to the top quark mass itself. The jet energy scale factor (JSF) is aiming to correct for uncertainty coming from jet energy correction (JEC) which is one of the leading systematic uncertainties. The event likelihood function takes into account all possible permutations per event, and applies the goodness-of-fit probability from the kinematic fit. The ideogram method is used in this analysis and is described in detail in Section 8.4.1.

Analytical matrix-weighting technique The analytical matrix-weighting technique is (AMWT) [228] used in the dilepton channel, due to the presence of two neutrinos among the decay products of $t\bar{t}$. The method scans the possible m_t values in small increments, sharing among the two neutrinos the momentum imbalance, resulting from the multiple possible solutions per event, where for each expected result a weight is assigned based on the probability of observing a charged lepton of energy E in a top frame [229]. Solutions with the highest

average weight are named m_t^{AMWT} . Thus, in the AMWT method the top quark mass is determined by comparing the m_t^{AMWT} distribution to simulations for different m_t hypotheses.

8.4.1. Ideogram method

The ideogram method is one of the most popular techniques used for the direct measurement of the top quark mass [7, 121, 230, 231] as well as W boson mass [12]. The method makes use of MC-derived templates, where event-by-event likelihoods are constructed to reflect the ambiguity in the events arising from the different possible jet permutations.

In this analysis, two free parameters are used in the joint likelihood fit: the top quark mass and a jet energy scale factor (JSF), where the latter aims to reduce the uncertainty coming from the jet energy correction, one of the largest systematic source in the top quark mass measurements. Another advantage of the JSF measuring in situ is that it is obtained with the same sample that is used for the top quark mass measurement and thus removes the sample dependence of the JEC scale. In lepton+jets $t\bar{t}$ events the jet energy correction scale is calibrated using the hadronically decaying W boson, therefore the reconstructed W boson mass m_W^{reco} , before it is constrained by the kinematic fit; the scale is then used as an input parameter to the fit. Another input parameter is the fitted top quark mass m_t^{fit} obtained after applying the constrained kinematic fit, described previously, which is used to measure the most probable top quark mass m_t . The distribution of m_t^{fit} and m_W^{reco} are extracted from the available MC simulations of seven different m_t and five different JSF values.

In order to perform the ideogram method, the probability $P(m_t, \text{JSF}|\text{sample})$ for measuring the top quark mass, m_t , and jet energy scale factor, JSF, given an observed sample, is calculated according to Bayes' theorem:

$$P(m_t, \text{JSF}|\text{sample}) = P(\text{sample}|m_t, \text{JSF}) \times \frac{P(m_t) P(\text{JSF})}{P(\text{sample})}. \quad (8.10)$$

Here the $P(\text{sample}|m_t, \text{JSF})$ is a probability to observe the sample given m_t and JSF, and $P(m_t)$, $P(\text{JSF})$ and $P(\text{sample})$ present the prior knowledge of the m_t , JSF and the experimental outcome, respectively. The prior for the top quark mass $P(m_t)$ and the sample $P(\text{sample})$ is taken to be flat in m_t and data, and the resulting $P(m_t)$ and $P(\text{sample})$ yield a constant factor in the total probability. The most likely m_t and JSF, given a certain sample, are those where the probability in Eq. 8.10 is maximal or in terms of a "likelihood"⁹ the maximal value of $-2\ln[\mathcal{L}(m_t, \text{JSF}|\text{sample})]$.

First the m_t^{fit} and m_W^{reco} distributions are utilized to build probability density functions P_j , which are derived separately for different permutations, where j rep-

⁹Denoting an unnormalized probability a "likelihood" $\mathcal{L}(m_t, \text{JSF}|\text{sample})$

resents the correct (cp), wrong (wp) or unmatched (un) permutation types. Since the light quark jets in m_t^{fit} are constrained to the W boson mass $m_W = 80.4 \text{ GeV}$ by the kinematic fit, m_t^{fit} and m_W^{reco} have a low correlation coefficient of 0.02. Therefore, two probability density functions can be estimated, $P_j(m_t^{\text{fit}}|m_t, \text{JSF})$ and $P_j(m_W^{\text{reco}}|m_t, \text{JSF})$, which depend on m_t and JSF and are used to calculate the two-dimensional permutation probability density:

$$P(m_t^{\text{fit}}, m_W^{\text{reco}}|m_t, \text{JSF}) = \sum_j f_j P_j(m_t^{\text{fit}}|m_t, \text{JSF}) \times P_j(m_W^{\text{reco}}|m_t, \text{JSF}), \quad (8.11)$$

where f_j is the relative fraction of each permutation calculated for the simulated sample which was generated with $m_t^{\text{gen}} = 127.5 \text{ GeV}$. By taking the sum of the defined $P(m_t^{\text{fit}}, m_W^{\text{reco}}|m_t, \text{JSF})$, where each permutation is weighted by the goodness-of-fit probability P_{gof} , defined in Section 8.3.4, a single-event likelihood is constructed:

$$\mathcal{L}(\text{event}|m_t, \text{JSF}) = \sum_{i=1}^n P_{\text{gof}}(i) P(m_t^{\text{fit}}, m_W^{\text{reco}}|m_t, \text{JSF}), \quad (8.12)$$

where i runs over n permutations in each event. By definition, individual event likelihoods are named *ideograms*. By taking the product over all single-event likelihood the likelihood, for the whole sample can be calculated:

$$\mathcal{L}(\text{sample}|m_t, \text{JSF}) = \prod_{\text{events}} (\mathcal{L}(\text{event}|m_t, \text{JSF}))^{w_{\text{event}}}, \quad (8.13)$$

where, additionally, an event weight $w_{\text{event}} = c \sum_{i=1}^n P_{\text{gof}}(i)$ is introduced to reduce the impact of events without correct permutations. The constant c normalizes the average w_{event} to 1 and is defined as $c = N_{\text{events}} / \sum_{\text{events}} w_{\text{event}}$. Finally, by using the prior knowledge for the JSF in the likelihood, the $P(m_t, \text{JSF}|\text{sample})$ can be defined as:

$$P(m_t, \text{JSF}|\text{sample}) = \mathcal{L}(\text{sample}|m_t, \text{JSF}) \times P(\text{JSF}). \quad (8.14)$$

Unlike the similar analysis performed in the all-jets channel [232], the contribution from background is not included in the final expression for the lepton+jets case. Since the fraction of background events relative to the data corresponds to the 4.3%, which is found to be negligible at this stage of the analysis, the background can be taken into account during the calibration of the method, as described in Sections 8.4.3.

A different choice for the JSF prior $P(\text{JSF})$ in the likelihood fit, results in several final measures of the m_t and JSF. Thus, if the JSF is fixed to unity the $P_j(m_W^{\text{reco}}|m_t, \text{JSF})$ can be approximated as constant, since they hardly depend on m_t , as mentioned before. Hence, only the m_t^{fit} observable is fitted, and this approach is called *1D method*. The approach without including the prior knowledge

on JSF is called *2D method* and $P(\text{JSF})$ yields a constant factor. Finally, in the “*hybrid*” method the prior knowledge from the jet energy correction, as discussed in Section 6.5, is taken into account. Therefore, the prior $P(\text{JSF})$ is centered at 1.0 and has a Gaussian shape with a width $\sigma_{\text{prior}} = \delta\text{JSF}_{\text{stat}}^{2\text{D}} \sqrt{1/w_{\text{hyb}} - 1} = 0.0012$, where $w_{\text{hyb}} = 0.3$ is the relative weight assigned to the prior knowledge on the JSF derived from the uncertainties in the jet energy corrections [13]. However, the value of w_{hyb} for the current analysis is found to be lower than in the Run 1 analysis [7] where the prior JSF knowledge contributes to 50% of the information. This is due to the larger jet energy uncertainties at the beginning of Run 2. The statistical uncertainty in the JSF expected for the 2D method is additionally utilized by $\delta\text{JSF}_{\text{stat}}^{2\text{D}} = 0.0008$. Thus, minimizing Eq. 8.14 yields the measurement of the parameters, m_t and JSF.

8.4.2. Method implementation

The ideogram method is individually applied in the muon+jets and electron+jets decay channels and the final result is given by multiplying the likelihood for both cases, e.g., $\mathcal{L}_{l+\text{jets}} = \mathcal{L}_{\mu+\text{jets}} \times \mathcal{L}_{e+\text{jets}}$. The mass distribution (either m_t^{fit} or m_W^{reco}) used by the likelihood function is derived for a particular generated top quark mass¹⁰ and jet energy scale factor¹¹ and is referred to as a *template*. Thus, for any given top quark mass and jet energy scale, the probability density function for m_t^{fit} or m_W^{reco} is specified by fitting a functional form to the corresponding template.

The fitted top quark mass and reconstructed W boson mass distributions of correct (cp), wrong (wp), unmatched (un) permutations, reconstructed for each subsample and each decay channel, are fitted by the corresponding fit function. The top quark mass of the correct permutations $m_{t,\text{cp}}^{\text{fit}}$ is fitted with a Voigt functions consisting of the Breit-Wigner function convoluted with the Gaussian resolution function:

$$V(m; \mu, \sigma, \Gamma) = \int dm' BW(m'; \mu, \Gamma) \times G(m; m', \sigma), \quad (8.15)$$

where μ corresponds to the mean value of the distribution, σ characterizes the Gaussian resolution and Γ describes the width. The Breit-Wigner function represents the expected distribution of the average invariant mass of the top quark and is described as

$$BW(m; \mu, \Gamma) = \frac{1}{2\pi} \cdot \frac{\Gamma}{((m - \mu)^2) + \Gamma^2/4}, \quad (8.16)$$

where the width is set to a value of 2 GeV, the standard model value of the top quark decay width. The Gaussian function describes the experimental finite resolution and is given as

$$G(m; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{(m - \mu)^2}{2\sigma^2}\right). \quad (8.17)$$

¹⁰ $m_{t,\text{gen}}$ of 166.5, 169.5, 171.5, 172.5, 173.5, 175.5 and 178.5 GeV

¹¹JSF of 0.96, 0.98, 1.00, 1.02, 1.04

The top quark mass of the wrong $m_{t,wp}^{\text{fit}}$ and unmatched $m_{t,un}^{\text{fit}}$ permutations is fitted by a Crystal Ball function, which is described by a Gaussian that from one side is cut off and then replaced with a power-law function, to account for the larger high mass tail:

$$CB(m; \mu, \sigma, \alpha, n) = N \cdot \begin{cases} \exp\left(-\frac{(m-\mu)}{\sigma}\right)^{-n}, & \text{for } \frac{m-\mu}{\sigma} > -\alpha \\ A \cdot \left(B - \frac{(m-\mu)^2}{2\sigma^2}\right), & \text{for } \frac{m-\mu}{\sigma} \leq -\alpha \end{cases} \quad (8.18)$$

with

$$A = \left(\frac{n}{|\alpha|}\right)^n \cdot \exp\left(-\frac{|\alpha|^2}{2}\right), \quad B = \frac{n}{|\alpha|} - |\alpha|. \quad (8.19)$$

Finally, the mass of the W boson $m_{W,i}^{\text{reco}}$ for all permutations $i = cp, wp, un$ is fitted by the asymmetric Gauss function:

$$AG(m; \mu, \sigma_1, \sigma_2) = \begin{cases} \frac{1}{\sigma_1 \sqrt{2\pi}} \cdot \exp\left(-\frac{(m-\mu)^2}{2\sigma_1^2}\right), & \text{for } \frac{m-\mu}{\sigma_1} < 0 \\ \frac{1}{\sigma_2 \sqrt{2\pi}} \cdot \exp\left(-\frac{(m-\mu)^2}{2\sigma_2^2}\right), & \text{for } \frac{m-\mu}{\sigma_1} \geq 0, \end{cases} \quad (8.20)$$

where μ is a mean value and σ_1 and σ_2 are resolutions. Examples for the template shapes of the fitted top mass and the reconstructed W boson mass from each event category in the muon+jets channel are shown in Fig. 8.12 and Fig. 8.13, respectively. The same distributions for the electron+jets decay channel are presented in Appendix A.5.

The distributions show the influence of the $m_{t,\text{gen}}$ and JSF variation to the reconstructed distribution shape. However, the observable of m_W^{reco} has mostly no dependence on the generator top quark mass because m_W^{reco} is distorted by the analysis selection and the weighting from the fit probability of the kinematic fit, which requires the mass of two light jets to be equal 80.4 GeV.

The fitted parameters for correct, wrong and unmatched permutations show a linear behavior as a function of the top quark mass and JSF. Therefore, the next step is to determine the dependence of each free parameter of the fit on the input $m_{t,\text{gen}}$ and JSF. It is possible to parametrize each free parameter from the template fit using the relation

$$f(m_t, JSF) = p_0 + p_1\left(\frac{m_t}{\text{GeV}} - 172.5\right) + p_2(JSF - 1) + p_3\left(\frac{m_t}{\text{GeV}} - 172.5\right)(JSF - 1), \quad (8.21)$$

where parameter p_0 denotes the mean value, p_1 corresponds to the slope with respect to the m_t , p_2 refers to slope with respect to the JSF, and p_3 gives the correlation strength between m_t and JSF. To obtain all these parameters, all free parameters from the template fits, namely parameters for the Gaussian mean and widths, are fitted first as a linear function of $(JSF - 1)$ for fixed $m_{t,\text{gen}}$. Figure A.25 shows

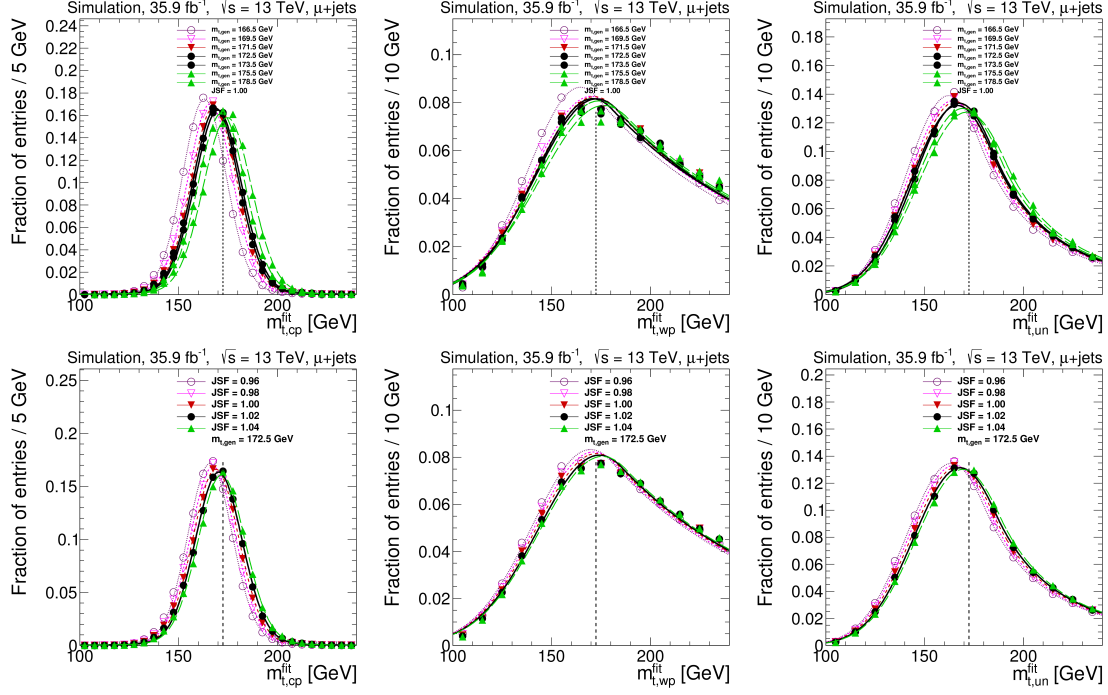


Figure 8.12.: Simulated m_t^{fit} distributions of correct (left column), wrong (middle column), and unmatched (right column) $t\bar{t}$ permutations, for seven generated masses $m_{t,\text{gen}}$ with $\text{JSF} = 1$ in the upper row and for five jet scale factors with $m_{t,\text{gen}} = 172.5$ GeV in the lower row for muon+jets events. The vertical dashed line corresponds to $m_t^{\text{fit}} = 172.5$ GeV.

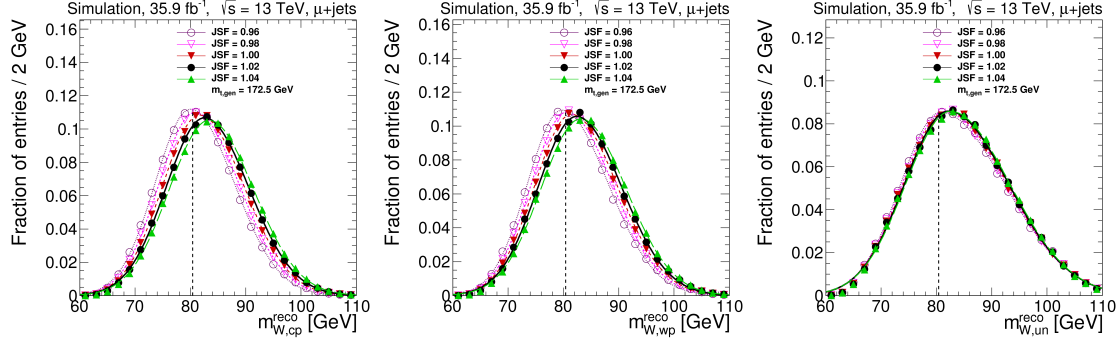


Figure 8.13.: Simulated m_W^{reco} distributions are shown for correct (left), wrong (middle), and unmatched (right) $t\bar{t}$ permutations for five jet scale factors with $m_{t,\text{gen}} = 172.5$ GeV for muon+jets events. The vertical dashed lines indicate the accepted value of the W boson mass of 80.4 GeV.

the parameters obtained for the m_t^{fit} and m_W^{reco} for the correct permutations in the muon+jets channel, the results for wrong and unmatched permutations as well as for the electron+jet channel are given in Appendix A.5. The dependence of the mean and resolution on the JSF are visible. As expected, not much dependence for different input values for $m_{t,\text{gen}}$ on m_W^{reco} is visible, due to the kinematic fit used in the analysis. The parameters obtained in terms of the JSF are fitted linearly as

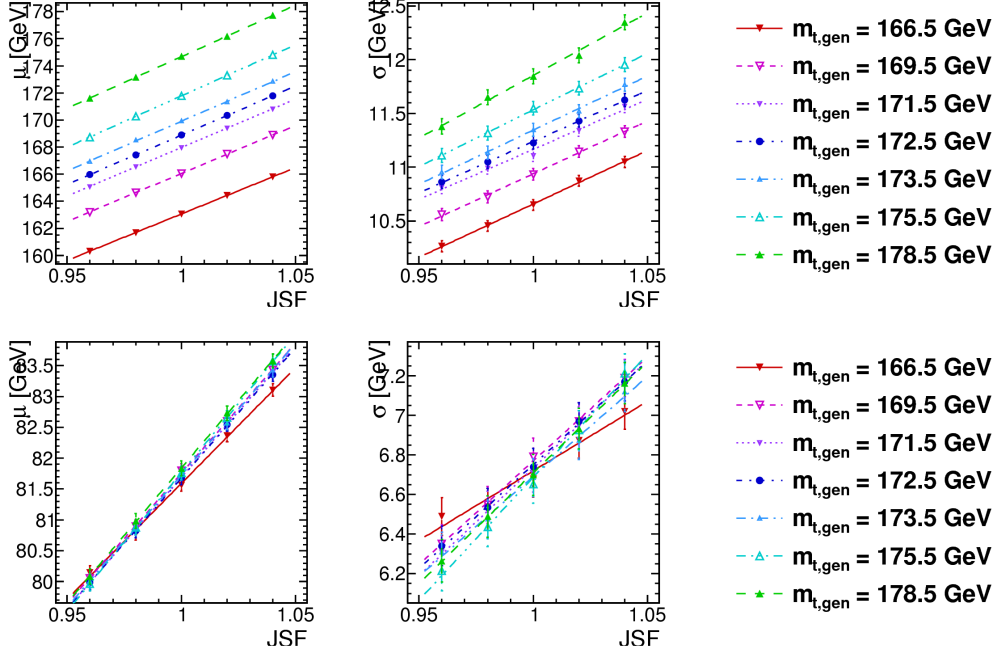


Figure 8.14.: The Gaussian mean (left) and resolution (right) of the m_t^{fit} (upper row) and m_W^{reco} (lower row) distributions of correct $t\bar{t}$ permutations, for seven generated masses $m_{t,\text{gen}}$ and five JSF in the muon+jets events.

a function of $(\frac{m_t}{\text{GeV}} - 172.5)$. The result of the linear fitting is shown in Fig. 8.15 and Fig. 8.16, where the terms “const” denotes the central mass or its width, and “slope” refers to the dependency on the JSF found in the previous step. The results of the fitting procedure in the case of the wrong and unmatched permutations in the electron+jets channel are given in Appendix A.5.

The parameters obtained from the second set of linear fits shown in Fig. 8.15 and Fig. 8.16 correspond to the free parameters in Eq. 8.21 and are shown in Tab. 8.9 and 8.10 for m_t^{fit} and m_W^{reco} for all possible permutations in muon+jets and electron+jets channel, respectively.

By defining each free parameter of the template fit as a function of $m_{t,\text{gen}}$ and JSF, the probability density for m_t^{fit} or m_W^{reco} is determined, i.e., $P_{cp}(m_{t,i}^{\text{fit}}|m_t, JSF) \equiv V(m_{t,i}^{\text{fit}}; \mu(m_t, JSF), \sigma(m_t, JSF), \Gamma \approx 2 \text{ GeV})$, where i denotes the contributions cp, wp, or un. Together with the permutation fractions f_j , listed in Tab. 8.11 for both muon and electron channels, the parameters are used to construct a joint likelihood function. Finally, the most likely top quark mass and JSF is obtained by minimizing the negative logarithm of the likelihood $-2 \ln \mathcal{L}(m_t, JSF|\text{sample})$.

8.4.3. Method and statistical uncertainty validation

Due to neglecting the background contamination in the ideogram method as well as the presence of some simplifications, e.g., fixing the values for the permutation fraction, a possible bias may arise in the estimate of m_t . Therefore, the method is

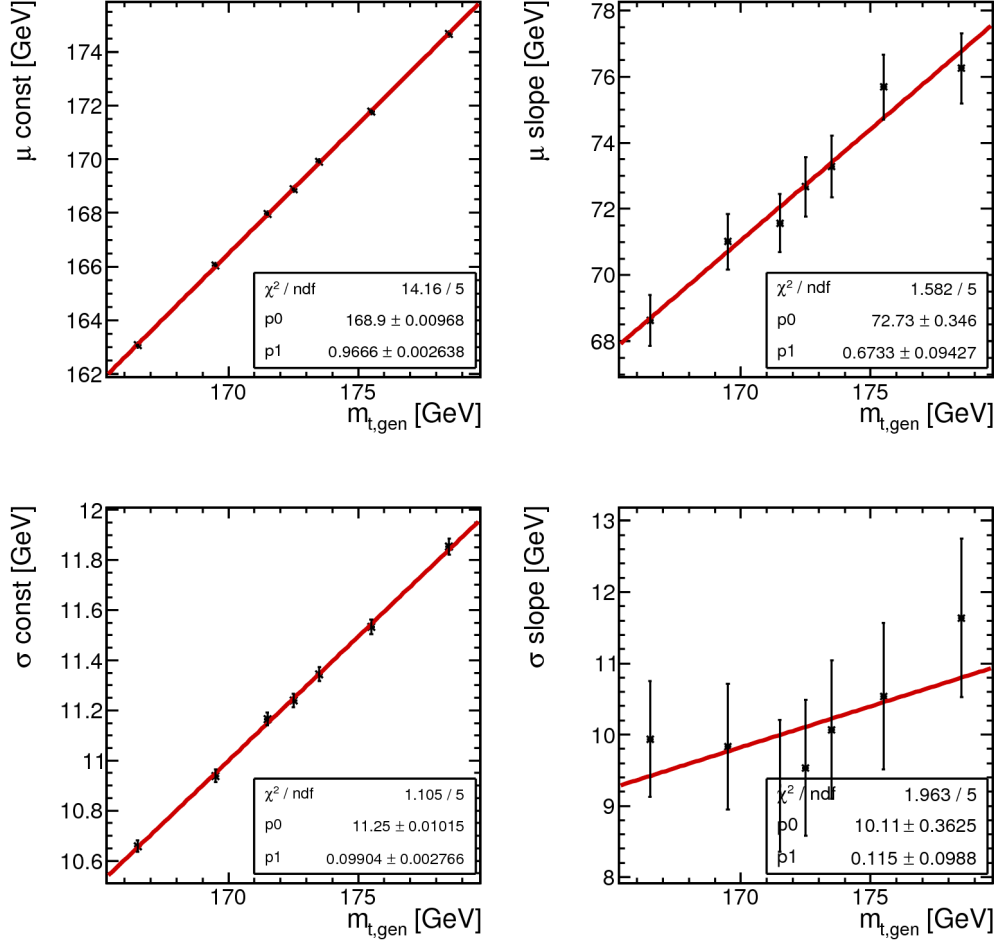


Figure 8.15.: The Gaussian mean (upper left) and resolution (lower left) and their dependency on JSF (upper and lower right) of the m_t^{fit} distributions of correct $t\bar{t}$ permutations, for seven generated masses $m_{t,\text{gen}}$ in the muon+jets events.

calibrated for each combination of seven $m_{t,\text{gen}}$ values and the five JSF scales using Monte Carlo simulation. Both the bias on the measured mass and the correctness of the estimation of the statistical uncertainty are studied by applying an ensemble test, in which 10 000 pseudo-experiments, separately for the muon and electron channels¹², are generated, each matching the size of the observed data¹³ sample. The background processes used in the analysis are included in the pseudo-experiments. The calibration is done only for the 2D approach, but is valid for the 1D and hybrid approaches.

First, for each $m_{t,\text{gen}}$ -JSF combination, the corresponding set of pseudo-experiments is constructed. For each such set, the most probable top quark mass $m_{t,\text{extr}}$ and jet

¹²The number of selected events in data in the muon+jets channel is 101992, while in electron+jets channel it corresponds to 59504 events

¹³The number of events in each pseudo-experiment is given by a Poisson distribution around the number of events observed in data for the individual or combined channels.

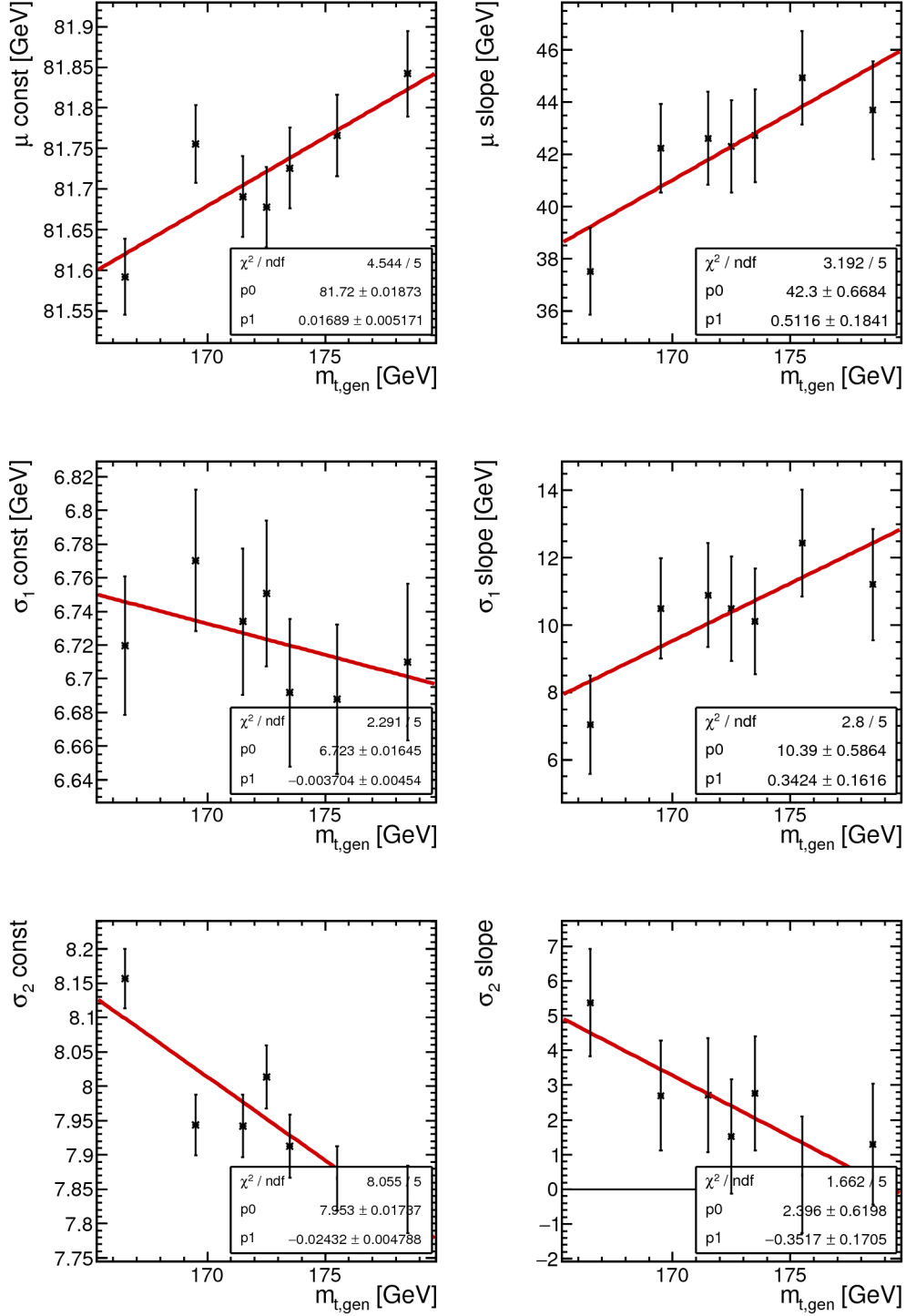


Figure 8.16.: The Gaussian mean (upper left) and its widths (middle and lower left) and their respective dependency on JSF (upper, middle and lower right) of the m_W^{reco} distributions of correct $t\bar{t}$ permutations, for seven generated masses $m_{t,gen}$ in the muon+jets channel.

Table 8.9.: Parameters of the fit functions $V(m; \mu, \sigma, \Gamma \approx 2 \text{ GeV})$ for $m_{t,cp}^{\text{fit}}$, $CB(m; \mu, \sigma, \alpha, n)$ for $m_{t,cp}^{\text{fit}}$ and $m_{t,un}^{\text{fit}}$ and $AG(m; \mu, \sigma_1, \sigma_2)$ for m_W^{reco} for the muon+jets channel. The parameter p_0 denotes the mean value, p_1 is the slope with respect to the m_t , p_2 is the slope with respect to the JSF, and p_3 gives the correlation strength between m_t and JSF.

Distribution	Permutation	Parameter	p_0	p_1	p_2	p_3
m_t^{fit}	correct	μ	168.903	0.966648	72.7285	0.673338
		σ	11.2474	0.0990394	10.1094	0.114957
	wrong	μ	171.372	0.951018	85.962	0.747902
		σ	29.904	0.463822	31.7932	0.291482
		α	0.352363	0.00413678	0.276496	0.00285401
	unmatched	μ	166.662	0.795324	73.9572	0.730175
		σ	22.9724	0.267643	14.1756	-0.296404
		α	0.766839	0.00699443	0.415425	-0.0218869
m_W^{reco}	correct	μ	81.7215	0.0168941	42.2974	0.51165
		σ_1	6.72345	-0.00370375	10.386	0.342442
		σ_2	7.95257	-0.0243178	2.39579	-0.351668
	wrong	μ	81.4659	0.0508487	41.8046	0.387282
		σ_1	6.66013	0.0174422	11.1738	0.376439
		σ_2	8.29277	8.29277	-0.992615	-0.0444368
	unmatched	μ	81.5526	0.0114018	21.9529	0.390227
		σ_1	7.60138	-0.00556115	8.78498	0.224483
		σ_2	11.2402	-0.0213487	-11.6002	-0.216464

scale factor JSF_{extr} are extracted, as shown in Fig. 8.17 for the muon+jets channel. For every pseudo-experiment set, the top quark mass and JSF are fitted by a

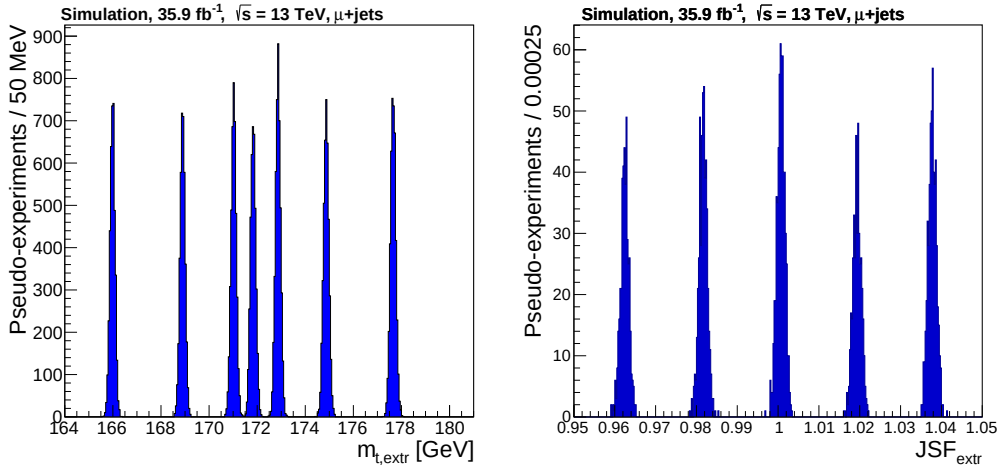


Figure 8.17.: The extracted most probable top quark mass $m_{t,\text{extr}}$ for the different $m_{t,\text{gen}}$ and nominal $\text{JSF} = 1$ (left), and extracted jet scale factor JSF_{extr} for different JSF and nominal $m_{t,\text{gen}} = 127.5$ (right) for the muon+jets channel. The pseudo-experiments before the calibration in the muon+jets channel were used.

Table 8.10.: Parameters of the fit functions $V(m; \mu, \sigma, \Gamma \approx 2 \text{ GeV})$ for $m_{t,cp}^{\text{fit}}$, $CB(m; \mu, \sigma, \alpha, n)$ for $m_{t,cp}^{\text{fit}}$ and $m_{t,un}^{\text{fit}}$ and $AG(m; \mu, \sigma_1, \sigma_2)$ for m_W^{reco} for the electron+jets channel. The parameter p_0 denotes the mean value, p_1 is the slope with respect to the m_t , p_2 is the slope with respect to the JSF, and p_3 gives the correlation strength between m_t and JSF.

Distribution	Permutation	Parameter	p_0	p_1	p_2	p_3
m_t^{fit}	correct	μ	169.016	0.951853	75.6938	0.810477
		σ	11.1851	0.0977989	10.1927	0.0122348
	wrong	μ	172.74	0.693914	90.9727	3.67504
		σ	30.7825	0.168547	34.1595	2.33937
		α	0.337052	0.000273241	0.210567	0.0417988
	unmatched	μ	166.643	0.81974	77.3164	0.788619
		σ	22.6301	0.271993	14.2216	-0.376767
		α	0.759073	0.00830221	0.277106	0.00247072
m_W^{reco}	correct	μ	81.8772	0.0426048	44.7667	-0.227307
		σ_1	6.73424	0.021648	12.6037	-0.0950635
		σ_2	7.8487	-0.0257919	0.17861	0.227245
	wrong	μ	81.7721	0.0194891	44.6181	-0.617021
		σ_1	6.75697	0.0184562	14.1412	-0.403639
		σ_2	8.05543	-0.0130362	-2.98919	0.450587
	unmatched	μ	81.6445	0.0256832	17.3373	-0.0105077
		σ_1	7.62231	0.0100573	6.37459	0.0360975
		σ_2	11.2211	-0.0134243	-8.22603	-0.0621621

Table 8.11.: Relative permutation fractions f_j for the muon+jets and electron+jets channels.

	$\mu + \text{jets}$	$e + \text{jets}$
f_{cp}	0.451	0.468
f_{wp}	0.157	0.154
f_{un}	0.393	0.377

Gaussian function that is used to extract the biases:

$$\text{mass bias} = \langle m_{t,\text{extr}} - m_{t,\text{gen}} \rangle, \quad (8.22)$$

$$\text{JSF bias} = \langle \text{JSF}_{\text{extr}} - \text{JSF} \rangle, \quad (8.23)$$

where the mean value and widths of the Gaussian refer to the exacted value and its uncertainty for the current sample of the pseudo-experiments. Figure 8.18 shows small deviations in the top quark mass and in the JSF as a function of $m_{t,\text{gen}}$ calculated for each set of pseudo-experiments. For each generated JSF, the biases are fitted with a linear function. The parameters of the fits are used to derive calibration factors for the extracted values, similarly to the template parametrization, using Eq. 8.21. The calibration parameters depend linearly on m_t and JSF, and on the product of the two. They are listed in Tab. 8.12. Applying this calibration to

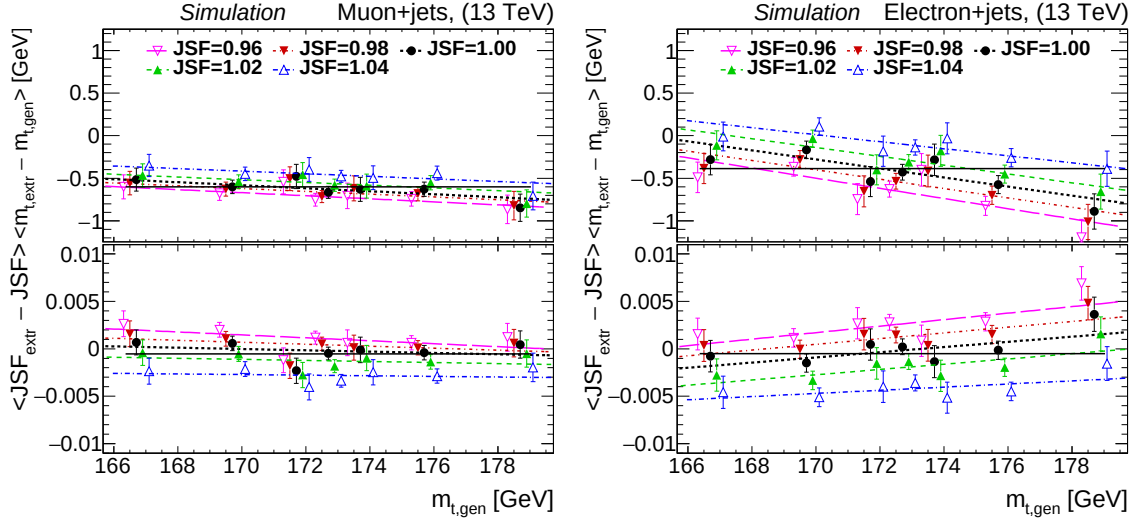


Figure 8.18.: The bias values of m_t and JSF, before the calibration, as a function of different generated $m_{t,\text{gen}}$ and five values of JSF for the muon channel (left) and the electron channel (right). The colored dashed lines correspond to straight line fits, the black solid line corresponds to a constant fit to all calibration points. The error bars indicate the statistical uncertainty in the mean difference.

Table 8.12.: The calibration parameters used to correct the biases in top quark mass and JSF in the muon+jets and electron+jets decay channels. The parameter p_0 denotes the mean value, p_1 is the slope with respect to the m_t , p_2 is the slope with respect to the JSF, and p_3 gives the correlation strength between m_t and JSF.

Channel	Parameter	p_0	p_1	p_2	p_3
μ +jets	m_t	-6.14649×10^{-1}	-1.62321×10^{-2}	3.16208	2.96164×10^{-2}
	JSF	-6.37079×10^{-4}	-9.25061×10^{-5}	-4.82111×10^{-2}	1.51420×10^{-3}
e +jets	m_t	-4.20965×10^{-1}	-5.56753×10^{-2}	7.40424	2.29851×10^{-1}
	JSF	-4.25533×10^{-4}	3.03672×10^{-4}	-9.15393×10^{-2}	-2.03800×10^{-3}

the extracted values of $m_{t,\text{extr}}$ and JSF_{extr} , all biases in the measured values are negligible. A final validation using pseudo-experiments is shown in Fig. 8.19, which shows that the calibrated top quark mass $m_{t,\text{cal}}$ and jet scale factor JSF_{cal} agree with the input values of m_t and JSF, indicating that the biases are close to zero.

In order to validate the estimated statistical uncertainty of the method the *pull* distribution for every pseudo-experiment is studied. The pull is defined as

$$\text{mass pull} = \frac{m_{t,\text{cal}} - m_{t,\text{gen}}}{\sigma(m_{t,\text{cal}})}, \quad (8.24)$$

$$\text{JSF pull} = \frac{\text{JSF}_{\text{cal}} - \text{JSF}}{\sigma(\text{JSF}_{\text{cal}})}, \quad (8.25)$$

where $\sigma(m_{t,\text{cal}})$ and $\sigma(\text{JSF}_{\text{cal}})$ is the statistical uncertainty of an individual $m_{t,\text{cal}}$ and

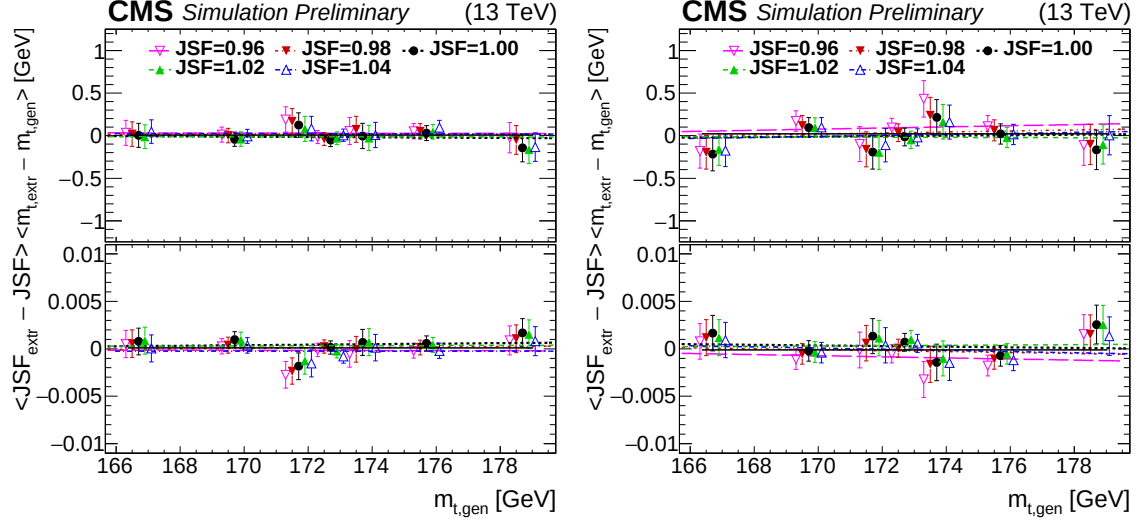


Figure 8.19.: Mean difference between the calibrated and generated values of m_t and JSF as a function of different generated $m_{t, \text{gen}}$ and five values of JSF for the muon channel (left) and the electron channel (right) [10]. The colored dashed lines correspond to straight line fits, the black solid line corresponds to a constant fit to all calibration points. The error bars indicate the statistical uncertainty in the mean difference.

JSF_{cal} for a pseudo-experiment at $m_{t, \text{gen}}$ and JSF, respectively. The pull distribution for all calibrated pseudo-experiments is fitted with a Gaussian function in order to extract the width, which is called the *pull width*. The deviation of the pull width from unity gives the correction for the statistical uncertainties of the method and is found to have a size of 5% as shown in Fig. 8.20. Therefore, the pull values are varied within 5% for different top quark masses and JSF values. However, this effect is negligible as the expected statistical uncertainty in m_t in the 2D approach is only 0.11 GeV and 0.15 GeV in the muon+jets and electron+jets channels, respectively. This is much smaller than the systematic error.

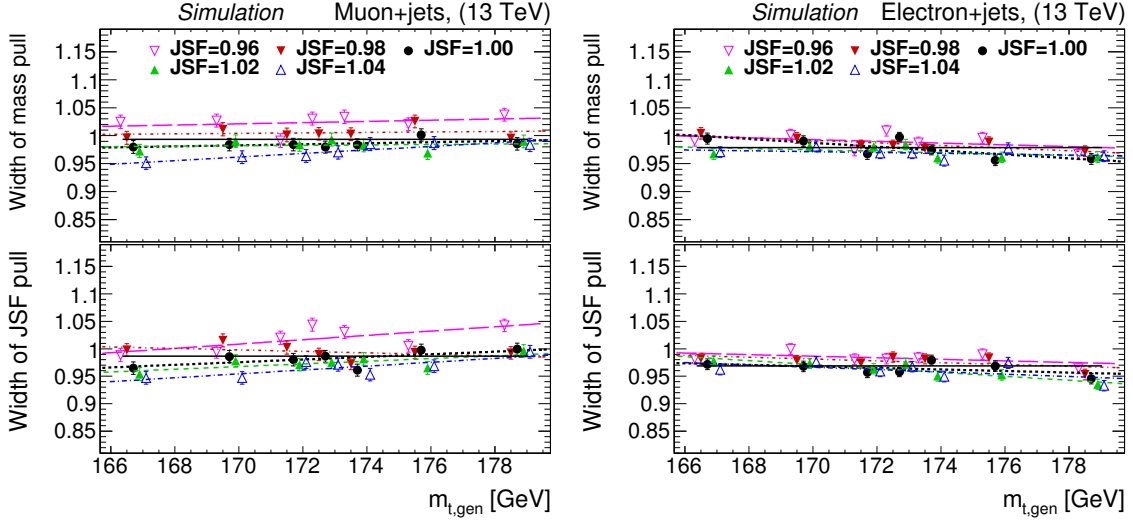


Figure 8.20.: The pull widths of the m_t and JSF as a function of different generated $m_{t,gen}$ and five values of JSF for the muon channel (left) and the electron channel (right). The colored dashed lines correspond to straight line fits. The black solid line corresponds to a constant fit to all calibration points assuming no dependence on m_t or JSF.

8.5. Systematic uncertainties

The ideogram method used for the top quark mass measurement as well as its calibration rely on Monte Carlo simulation. Therefore, in order to avoid any bias between data and simulation, at least 10 000 pseudo-experiments for each systematic variation are produced. Each pseudo-experiment corresponds to an integrated luminosity of 35.9 fb^{-1} . The shift in the mean value of the Gaussian fit to the ensemble of the top quark mass or JSF compared to the default value gives the size of the corresponding systematic uncertainty. Both positive and negative signs of the shifts are possible, where the positive ones appear if a variation to a lower value leads to a higher extracted m_t , and vice versa. To determine the impact of each uncertainty, up and down variations are performed, changing the parameters by ± 1 standard deviation, where the largest observed shift on the result is taken as the corresponding systematic uncertainty.

Three different sets of systematic uncertainties for several systematic sources are calculated, using 2D, 1D and hybrid approaches, as mentioned in Section 8.4.1, and are shown in Tab. 8.19 for lepton+jets, Tab. 8.20 for muon+jets, and Tab. 8.21 for electron+jets channels. In the 2D approach, no prior knowledge about jet energy correction is assumed (that is, a flat prior is used), and both m_t and JSF are evaluated. In the 1D approach, the JSF is fixed to unity in Eq. 8.14 by $P(\text{JSF}) = \delta(1)$, and only m_t results are shown. In order to combine 1D results with 2D results, a relative weight is introduced and used in the hybrid approach. The relative weight

of the 1D approach is calculated by

$$w_{\text{hyb}}^{\text{1D}} = (\sigma_{\text{JSF}}^{\text{2D}})^2 / \left((\sigma_{\text{JSF}}^{\text{2D}})^2 + (\sigma_{\text{JSF}}^{\text{c}})^2 \right), \quad (8.26)$$

where $\sigma_{\text{JSF}}^{\text{2D}}$ is the statistical uncertainty obtained for the JSF in the 2D approach and $\sigma_{\text{JSF}}^{\text{c}}$ corresponds to the Gaussian width that is used to define the prior knowledge of JSF, i.e., $P(\text{JSF})$ is chosen to be a Gaussian constrained with $\text{JSF} = 1 \pm \sigma_{\text{JSF}}^{\text{c}}$. Thus, by choosing the $\sigma_{\text{JSF}}^{\text{c}}$, it helps to extrapolate between the 1D and 2D approaches, where $\sigma_{\text{JSF}}^{\text{c}} \rightarrow \infty$ selects the 2D method and $\sigma_{\text{JSF}}^{\text{c}} \rightarrow 0$ selects the 1D method. In this work the relative weight $w_{\text{hyb}}^{\text{1D}} = 0.3$ is used in the hybrid approach, as explained in more detail in Section 8.6.

The relevant systematic uncertainties considered for this analysis and the methods used to evaluate them, are described in the following. The three main category groups of systematic uncertainties sources are the experimental uncertainties, modeling of the hadronization of the hard scattering process and of the non-perturbative QCD contributions. The $t\bar{t}$ signal samples used for the systematic uncertainties estimation are listed in Tab. 8.4. More detailed information about the exact extracted mass for each sample and all signs of the shifts of the systematic variations are given in the Appendix A.6. Shifts shown in brackets are shown as an overview over different uncertainty categories and are not used in the total uncertainty calculation. The total systematic uncertainty in the top quark mass and JSF measurement is done by adding all contributions in quadrature.

8.5.1. Experimental uncertainties

Systematic uncertainties arise from non perfect detector response and from the assumptions employed by the method to infer the measured values.

Method calibration: To account for the non-linear effects of the method calibration, the quadratic sum of statistical uncertainty and residual biases obtained after the calibration is used as a systematic uncertainty.

Jet energy correction: The jet energy correction (JEC) is one of the most dominant sources of uncertainty in the current measurement. The effect arises due to the direct relation between m_t , JSF, and the jet energy scale. Thus, the limited ability to reconstruct the jet-dependent kinematic variables is taking into account by estimating the p_{T} - and η -dependency of jet energy uncertainties. It is implemented by up and down scaling the energies of all jets using the data/MC scale factors [177] obtained during the JEC estimation. More details on the procedure for the jet energy correction determination can be seen in Section 6.5. The JEC uncertainty includes several individual sources. They can be splitted into multiple categories, as shown in [233] for

the Summer16_23Sep2016V4 uncertainties set. The *InterCalibration*, *MPFInSitu* and *Uncorrelated* groups are used in the analysis, each of them covers the different correlation groups of ATLAS/CMS uncertainties, that makes it easier to combine this result with other experiments. The *InterCalibration* is an η -dependent correction that stems from the FSR modeling. The *MPFInSitu* is a correction for the differences in the E_T^{miss} projection on the jet versus FSR in the data and simulation. The *Uncorrelated* correction corresponds to the difference between a p_T -dependent and flat residual correction for data only.

Jet energy resolution: The jet energy resolution (JER) is taken into account in addition to uncertainties in the reconstructed jet energy, by smearing the energy of jets in the MC events to match the data. The JER depends on η and p_T and is modified in the simulation by varying it with ± 1 standard deviations with respect to the resolution in data [180].

b tagging: The b tagging uncertainty is the uncertainty in the misidentification rates and the efficiency of the identification of the b-jets using the CSVv2M tagging algorithm [187]. It is estimated as the deviation between the nominal value of m_t or JSF and that computed after adjusting the b-tagging scale factors by their reported uncertainties.

Pileup: The number of pileup events are considered and added to the simulated hard scattering events, by reweighing the simulated pileup events to the data, as shown in Fig. 8.1. The uncertainty due to the pileup is estimated by varying the inelastic pp cross section by $\pm 4.6\%$ for all MC simulated samples, where the inelastic pp cross section is determined from the number of reconstructed vertices in $Z \rightarrow \mu^+\mu^-$ events as 69.2 ± 3.2 mb and is used together with the instantaneous luminosity to determine the number of pp interactions per bunch crossing.

Non- $t\bar{t}$ background: The sensitivity of the measurement on the choice of non- $t\bar{t}$ background model is evaluated by varying cross-section predictions used for the normalization of the non- $t\bar{t}$ background samples. Thus, the normalization of the W+jets background samples is varied by $\pm 30\%$, and for the Z+jets, the di-boson samples, and for the single top samples by $\pm 10\%$, and for the QCD-multijet samples by $\pm 100\%$. The luminosity that is used for the background normalization, is centrally measured by CMS using Van der Meer scans and has a corresponding uncertainty of 2.5% [11], which is found to be negligible, compared to the theory cross-section predictions variations.

The lepton modelling uncertainties together with the uncertainties in the missing transverse energy¹⁴ are found to be negligible and are not used for the total systematic calculation. The lepton modelling uncertainties include the uncertainties of the

¹⁴All changes in the energies for the lepton or jets are propagated to the MET.

lepton identification efficiency, reconstruction efficiency, trigger efficiency and the lepton isolation efficiency. The missing transverse energy uncertainty is evaluated by scaling the energy of the PF candidates that are not clustered into a jet with $p_T > 10$ GeV, 10% up and down.

8.5.2. Modeling uncertainties

The impact of the theoretical assumptions in the modeling of the hard scattering process and of the non-perturbative QCD is discussed in the following. The corresponding systematic uncertainty sources are estimated by using the dedicated $t\bar{t}$ samples, which are produced by varying the parameters within one model, POWHEG v2 + PYTHIA8 with the CUETP8M2T4 tune.

Modeling of hadronization

The uncertainty in the jet response for each flavor as well as the impact of b quark fragmentation hardness is evaluated.

Flavor jet energy response: The largest uncertainty in this analysis arises due to the flavor JEC. Since the difference in the response ($p_T / p_{T,\text{particle-level}}$) for different jet flavors comes mostly from the variations in jet fragmentation energy, the Lund string fragmentation implemented in PYTHIA 6.4 [16] is compared to the cluster fragmentation of HERWIG++ 2.4 [102] for the flavor JEC uncertainty evaluation. This is performed by varying the jet energies according to the response differences observed between PYTHIA 6.4 and HERWIG++ [13] models with respect to the Z+jets flavor mixture. PYTHIA 6.4 and HERWIG++ generators rely on a large set of tuning parameters that allow to modify the individual fragmentation of jets initiated from gluons, light quarks, and b quarks. Therefore, the uncertainty evaluation is done individually for each jet flavor, namely jets initiated from gluons, light quarks, charm quarks, and b quarks. These uncertainties are added linearly.

b jet modeling: The b jet modeling uncertainty includes three components that are implemented using event weights. The first two stem from the non-perturbative b quark fragmentation function, that accounts the momentum transfer from the b quark to the B hadron, which can be measured from data. The most sensitive observable for the b fragmentation function is $x_B = p_T(B)/p_T(\text{bjet})$ that was measured by the ALEPH [234] and DELPHI [235] collaborations and is used to vary the fragmentation into b hadrons in simulation. Thus, two scenarios of the fragmentation functions, Bowler-Lund and Peterson, are used and tuned in PYTHIA 8 to the x_B data. The difference of the results according to the ones given by these two approaches is quoted as an uncertainty. Similarly the semileptonic b hadron branching fraction may change

the b jet energy response significantly. Therefore, the semileptonic branching ratio of B hadrons is varied by -0.45% and $+0.77\%$, which is the range given by the measurements of the B^0/B^+ decays and their uncertainties [18]. These uncertainties are added in quadrature to the total systematic uncertainty.

Modeling of the hard scattering process

Parton distribution functions: The PDF4LHC [236] prescriptions are used to calculate the uncertainty due to the choice of the parton distribution functions (PDFs). The updated NNPDF3.0 PDF sets [83] for $\alpha_S(m_Z^2) = 0.118$ at NLO, which include more recent data on gauge boson and inclusive jets production from ATLAS, CMS and LHCb, results from H1 and ZEUS on HERA inclusive structure function from Run2 data, and W+charm and top quark pair production, are used in the generation of simulated events. Since NNPDF sets are based on 100 MC replicas, the PDF uncertainties are evaluated by making a distribution of the measured observable, i.e m_t , for each replica. The largest deviation is considered as the PDF uncertainty. In addition, the value of α_S used by the NNPDF30 PDF set was varied by ± 0.001 . Both uncertainties are added in quadrature.

Renormalization and factorization scales: The uncertainty given by the amount of radiation outside the top quark decays in the simulated events is estimated by changing the choice of the renormalization (μ_R) and factorization (μ_F) scales in POWHEG, which are set to $\mu_R = \mu_F = Q^2$ with the scale of the hard processes $Q^2 = m_t^2 + \sum p_T^2$, where the p_T^2 of all additional final-state partons in the ME calculations is summed. The μ_R and μ_F scales are varied by factors of 0.5 and 2. The systematic uncertainty is given by the largest difference with respect to the nominal μ_R and μ_F scales.

Matrix element/parton shower matching: The matching of the matrix elements (ME) in POWHEG to the PYTHIA parton showers (PS) is done by a dedicated factor $h_{\text{damp}} = 1.58^{+0.66}_{-0.59}$ [110] used in POWHEG. h_{damp} regulates the damping of the real emissions when matching to the PS in the NLO calculation. h_{damp} is varied within the considered uncertainties and corresponding additional $t\bar{t}$ samples are produced, which are used to perform the uncertainty in the ME/PS matching.

ISR PS scale: To account for an additional hard parton from initial state radiation (ISR) for $t\bar{t}$ events the corresponding parton shower scale value in PYTHIA is scaled up by 2 and down by 0.5. The uncertainty in the ISR is estimated with the dedicated samples produced for the up and down scale variation. These samples are used to estimate the effect of the choice of ISR α_S .

FSR PS scale: Dedicated samples for the up and down variation of the parton shower scale value used for the simulation of final state radiation in PYTHIA are used to derive the uncertainty in the FSR modeling. The parton shower scale value [116] is scaled up by $\sqrt{2}$ and down by $1/\sqrt{2}$. This affects the fragmentation and hadronization of the jets initiated by the matrix-element calculation and the possible emission of extra jets. Additionally, it was found that the jet energy response of the light quarks in the varied FSR samples is different by $\pm 1.2\%$ compared to the nominal sample. This response difference would be removed by corresponding residual jet energy corrections derived based on γ/Z +jet simulations with the same parton shower scale. Therefore, the momentum of all jets in the dedicated FSR samples are scaled to match the energy response for light quark induced jets in the default samples.

ME-generator: The effect of the different matrix element generators is estimated by using MC samples with MG5_aMC@NLO generator and the FxFx matching [104] instead of the POWHEG v2 generator used as default.

Top p_T : Recent calculations suggest that the p_T spectrum of the individual top quarks is strongly affected by next-to-next-to-leading-order effects [237]. Moreover, the top p_T spectrum in data is found to be softer than predicted by the simulation. Therefore, the top p_T in simulation is varied to match the distribution measured by CMS [238, 239] and its effect on the measurement is quoted as a systematic uncertainty.

Modeling of non-perturbative QCD

Underlying event: The PYTHIA parameters are tuned to the measurements of the underlying event [110, 116] that accounts for non-perturbative QCD effects. Thus, the uncertainty in the modeling of the underlying event is estimated by varying the parameters of the tune within their uncertainties.

Color reconnection modeling: To account for color reconnection (CR) effects occurring during the multiple parton interactions, two new alternative models of color reconnection are used. The “QCD inspired” [14] model that describes the string formation beyond leading color, and the “gluon move” [15] model that can move the gluons from one string to another string such that the total color string is as short as possible. More details on the models are given in Section 5.1. These two models are tuned to measurements of the underlying event [110, 116], as shown in Section 5.1. The new PYTHIA 8 samples for the corresponding CR tunes are produced and compared to default samples with the “MPI based” CR model. The largest observed shift is considered as CR systematic uncertainty.

Early resonance decays: Additionally, the effects of color reconnection on the top decay products can be used in PYTHIA 8 by turning on the early resonance decays (ERD), while in the default sample the ERD are turned off. All possible shifts in the CR and ERD in lepton+jets, muon+jets and electron+jets channels are listed in Tab. 8.13, 8.14 and 8.15, respectively.

Table 8.13.: Observed shifts for different color reconnection scenarios for the lepton+jets data set. The default results (def.) are compared to the various CR and ERD results. Bold values are used for the color reconnection systematic.

	2D approach		1D approach	Hybrid	
	δm_t^{2D} (GeV)	δJSF^{2D}	δm_t^{1D} (GeV)	δm_t^{hyb} (GeV)	$\delta \text{JSF}^{\text{hyb}}$
def. to "gluon move" (both ERD on)	0.34	-0.001	0.23	0.31	-0.001
def. to "QCD inspired" (both ERD on)	-0.11	-0.001	-0.19	-0.13	-0.001
def. ERD off to on	-0.22	0.008	0.42	-0.03	0.005
"gluon move" ERD on and off	-0.24	-0.006	-0.72	-0.38	-0.004
def. to "gluon move" (both ERD off)	-0.11	0.001	-0.06	-0.10	0.000
"QCD inspired" to "gluon move" (both ERD on)	0.45	-0.000	0.42	0.44	<0.001

Table 8.14.: Observed shifts for different color reconnection scenarios for the muon+jets data set. The default results (def.) are compared to the various CR and ERD results. Bold values are used for the color reconnection systematic.

	2D approach		1D approach	Hybrid	
	δm_t^{2D} (GeV)	δJSF^{2D}	δm_t^{1D} (GeV)	δm_t^{hyb} (GeV)	$\delta \text{JSF}^{\text{hyb}}$
def. to "gluon move" (both ERD on)	0.29	-0.001	0.21	0.27	-0.001
def. to "QCD inspired" (both ERD on)	-0.15	-0.001	-0.22	-0.17	-0.001
def. ERD off to on	-0.15	0.008	0.47	0.04	0.005
"gluon move" ERD on and off	-0.31	-0.006	-0.76	-0.45	-0.004
def. to "gluon move" (both ERD off)	-0.17	0.001	-0.07	-0.14	0.001
"QCD inspired" to "gluon move" (both ERD on)	0.44	-0.000	0.43	0.44	<0.001

Table 8.15.: Observed shifts for different color reconnection scenarios for the electron+jets data set. The default results (def.) are compared to the various CR and ERD results. Bold values are used for the color reconnection systematic.

	2D approach		1D approach	Hybrid	
	δm_t^{2D} (GeV)	δJSF^{2D}	δm_t^{1D} (GeV)	δm_t^{hyb} (GeV)	$\delta \text{JSF}^{\text{hyb}}$
def. to "gluon move" (both ERD on)	0.42	-0.002	0.26	0.37	-0.001
def. to "QCD inspired" (both ERD on)	-0.03	-0.001	-0.14	-0.06	-0.001
def. ERD off to def. ERD on	-0.34	0.008	0.34	-0.14	0.006
"gluon move" ERD on and off	-0.08	-0.007	-0.65	-0.25	-0.005
def. to "gluon move" (both ERD off)	0.00	-0.001	-0.05	-0.02	0.000
"QCD inspired" to "gluon move" (both ERD on)	0.45	-0.001	0.40	0.43	<0.001

8.5.3. Discussion of the systematic uncertainties

The total expected uncertainty in the top quark mass measurement, incorporating statistical and systematic effects, is found to be 0.62 GeV in lepton+jets, 0.64 GeV in muon+jets and 0.69 GeV in electron+jets events and is fully dominated by the systematic uncertainty for the hybrid method.

The usage of the JSF as a nuisance parameter in the fit and its constraint by the m_W^{reco} observable helps to reduce not only the uncertainties coming from the jet energy corrections, but also the modeling uncertainties, as can be seen from the summary tables. Thus, the remaining systematic uncertainties are dominated by effects that cannot be compensated through the simultaneous estimation of m_t and JSF. For example, the uncertainties in the flavor-dependent jet energy corrections cannot be reduced by using the JSF, since it is an overall energy scale factor derived mainly on light flavor jets and applied to all other jets. Other large uncertainties arise from the jet energy corrections, ME generator, color reconnection modeling, and early resonance decays, where the m_t^{fit} observable is effected sometimes differently than m_W^{reco} .

Even though the uncertainty in the color reconnection (CR) is found to be larger than that in Run 1, the better understanding of the CR of the partons belonging to the different multiple parton interactions (MPI) occurring within the same collision are now summarized in [14] and [15]. Hence, the usage of the various CR models allows more reliable estimation of CR effects on the measurement. While in Run 1, due to the lack of the studies on the CR, the CR uncertainties were treated by only observing the effect of an underlying event tune without CR.

In addition to the default POWHEG v2 + PYTHIA8 samples, the $t\bar{t}$ sample provided with aMC@NLO is now used in the analysis. The results obtained in the lepton+jets, muon+jets and electron+jets channels for the setups with the aMC@NLO generator and the FxFx matching [104], aMC@NLO with the MLM matching [103] (both interfaced to PYTHIA8), and POWHEG v2 interfaced to HERWIG++ [102] (v2.7.1) with the tune EE5C [111], are summarize in Tab. 8.16, 8.17 and 8.18, respectively. They show rather large shifts. This effect is already confirmed by another analysis [110], which found that the alternative generator setups do not describe the data well, as shown in Chapter 9. Therefore, only the aMC@NLO FxFx + PYTHIA8 is used as an alternative generator in the evaluation of the ME generator uncertainty, as shown before. A very large shift for POWHEG v2 + HERWIG++ was observed, which is due to the fact that matrix-element corrections to the top decay are not applied in the HERWIG++ sample.

In order to study the data agreement and the kinematic dependence of these generator setups and the color reconnection models, an additional study is performed as reported in Sec. 9.1.

Table 8.16.: Observed shifts for different ME-generators systematic for the lepton+jets data set.

	2D approach		1D approach	Hybrid	
	δm_t^{2D} (GeV)	δJSF^{2D}	δm_t^{1D} (GeV)	δm_t^{hyb} (GeV)	$\delta \text{JSF}^{\text{hyb}}$
PYTHIA 8 to madgraphMLM	0.82	0.000	0.80	0.82	0.000
PYTHIA 8 to powheg herwigpp	-4.39	0.014	-3.26	-4.06	0.010
PYTHIA 8 to aMCatNLOFFFX	0.19	0.001	0.29	0.22	0.001

Table 8.17.: Observed shifts for different ME-generators systematic for the muon+jets data set.

	2D approach		1D approach	Hybrid	
	δm_t^{2D} (GeV)	δJSF^{2D}	δm_t^{1D} (GeV)	δm_t^{hyb} (GeV)	$\delta \text{JSF}^{\text{hyb}}$
PYTHIA 8 to madgraphMLM	0.68	0.001	0.77	0.70	0.001
PYTHIA 8 to powheg herwigpp	-4.37	0.014	-3.31	-4.06	0.010
PYTHIA 8 to aMCatNLOFFFX	0.28	-0.001	0.23	0.26	-0.000

Table 8.18.: Observed shifts for different ME-generators systematic for the electron+jets data set.

	2D approach		1D approach	Hybrid	
	δm_t^{2D} (GeV)	δJSF^{2D}	δm_t^{1D} (GeV)	δm_t^{hyb} (GeV)	$\delta \text{JSF}^{\text{hyb}}$
PYTHIA 8 to madgraphMLM	1.09	-0.003	0.86	1.02	-0.002
PYTHIA 8 to powheg herwigpp	-4.42	0.015	-3.18	-4.05	0.010
PYTHIA 8 to aMCatNLOFFFX	-0.01	0.005	0.38	0.11	0.003

Table 8.19.: List of systematic uncertainties for the fit to the lepton+jets data set.

	2D approach		1D approach	Hybrid	
	δm_t^{2D} (GeV)	δJSF^{2D}	δm_t^{1D} (GeV)	δm_t^{hyb} (GeV)	$\delta \text{JSF}^{\text{hyb}}$
Experimental uncertainties					
Method calibration	0.05	<0.001	0.05	0.05	<0.001
Jet energy corrections (quad. sum)	0.13	0.002	0.83	0.18	0.003
– JEC: InterCalibration	(0.02)	(<0.001)	(0.16)	(0.04)	(<0.001)
– JEC: MPFInSitu	(0.01)	(<0.001)	(0.23)	(0.07)	(<0.001)
– JEC: Uncorrelated	(0.13)	(0.002)	(0.78)	(0.16)	(0.003)
Jet energy resolution	0.08	0.001	0.04	0.04	0.001
b tagging	0.03	<0.001	0.01	0.03	<0.001
Pileup	0.08	0.001	0.02	0.05	0.001
Non- $t\bar{t}$ background	0.04	0.001	0.02	0.02	0.001
Modeling of hadronization					
JEC: Flavor (linear sum)	0.42	0.001	0.31	0.39	<0.001
– light quarks (uds)	(+0.10)	(-0.001)	(-0.01)	(+0.06)	(-0.001)
– charm	(+0.02)	(<0.001)	(-0.01)	(+0.01)	(<0.001)
– bottom	(-0.32)	(<0.001)	(-0.31)	(-0.32)	(<0.001)
– gluon	(-0.22)	(+0.003)	(+0.02)	(-0.15)	(+0.002)
b-jet modeling (quad. sum)	0.13	0.001	0.09	0.12	<0.001
– b fragmentation Bowler-Lund	(0.07)	(< 0.001)	(0.01)	(0.05)	(<0.001)
– b fragmentation Peterson	(0.04)	(< 0.001)	(0.05)	(0.04)	(<0.001)
– semileptonic B hadron decays	(0.11)	(< 0.001)	(0.08)	(0.10)	(<0.001)
Hard scattering and radiation					
PDF	0.02	0.001	0.02	0.02	<0.001
Ren. and fact. scale	0.02	0.001	0.02	0.01	<0.001
ME/PS matching threshold	0.08	0.001	0.03	0.05	0.001
ME generator	0.19	0.001	0.29	0.22	0.001
ISR PS scale	0.07	0.001	0.10	0.06	<0.001
FSR PS scale	0.24	0.004	0.22	0.13	0.003
Top quark transverse momentum	0.02	0.001	0.06	0.01	0.001
Modeling of non-perturbative QCD					
Underlying event	0.10	0.001	0.01	0.07	0.001
Early resonance decays	0.22	0.008	0.42	0.03	0.005
Color reconnection modeling	0.34	0.001	0.23	0.31	0.001
Total systematic	0.71	0.010	1.09	0.62	0.008
Statistical (expected)	0.09	0.001	0.05	0.07	0.001
Total (expected)	0.72	0.010	1.09	0.62	0.008

Table 8.20.: List of systematic uncertainties for the fit to the muon+jets data set.

	2D approach		1D approach	Hybrid	
	δm_t^{2D}	δJSF^{2D}	δm_t^{1D}	δm_t^{hyb}	$\delta \text{JSF}^{\text{hyb}}$
Experimental uncertainties					
Method calibration	0.05	<0.001	0.05	0.05	
Jet energy corrections (quad. sum)	0.11	0.002	0.81	0.19	0.003
– JEC: InterCalibration	(0.03)	(<0.001)	(0.15)	(0.04)	(<0.001)
– JEC: MPFInSitu	(0.01)	(<0.001)	(0.23)	(0.07)	(0.001)
– JEC: Uncorrelated	(0.11)	(0.002)	(0.76)	(0.17)	(0.003)
Jet energy resolution	0.08	0.001	0.04	0.04	0.001
b tagging	0.03	0.000	0.01	0.02	0.00
Pileup	0.08	0.001	0.02	0.05	0.001
Non- $t\bar{t}$ background	0.03	0.001	0.08	0.00	0.001
Modeling of hadronization					
JEC: Flavor (linear sum)	0.43	0.001	0.34	0.40	0.001
– light quarks (uds)	(-0.9)	(-0.001)	(+0.02)	(+0.08)	(-0.001)
– charm	(-0.02)	(<0.001)	(+0.01)	(-0.01)	(<0.001)
– bottom	(+0.31)	(<0.001)	(+0.31)	(-0.30)	(<0.001)
– gluon	(+0.23)	(+0.003)	(<0.01)	(-0.17)	(0.002)
b-jet modeling (quad. sum)	0.08	0.001	0.07	0.07	<0.001
– b fragmentation Bowler-Lund	(0.06)	(0.001)	(0.01)	(0.04)	(<0.001)
– b fragmentation Peterson	(< 0.01)	(< 0.001)	(0.04)	(0.01)	(<0.001)
– semi-leptonic B hadron decays	(0.06)	(< 0.001)	(0.06)	(0.06)	(<0.001)
Hard scattering and radiation					
PDF	0.03	0.000	0.03	0.03	0.00
Ren. and fact. scale	0.02	0.001	0.02	0.02	0.001
ME/PS matching threshold	0.09	0.001	0.02	0.07	0.001
ME generator	0.28	0.001	0.23	0.26	<0.001
ISR PS scale	0.10	< 0.001	0.13	0.11	<0.001
FSR PS scale	0.32	0.005	0.28	0.20	0.004
Top quark transverse momentum	<0.01	0.001	0.06	0.02	0.001
Modeling of non-perturbative QCD					
Underlying event	0.06	0.001	0.08	0.04	0.001
Early resonance decays	0.15	0.008	0.47	0.04	0.005
Color reconnection modeling	0.29	0.001	0.22	0.27	0.001
Total systematic	0.73	0.010	1.16	0.64	0.008
Statistical (expected)	0.11	0.001	0.07	0.08	0.001
Total (expected)	0.74	0.010	1.16	0.64	0.008

Table 8.21.: List of systematic uncertainties for the fit to the electron+jets data set.

	2D approach		1D approach	Hybrid	
	δm_t^{2D}	δJSF^{2D}	δm_t^{1D}	δm_t^{hyb}	$\delta \text{JSF}^{\text{hyb}}$
Experimental uncertainties					
Method calibration	0.05	<0.001	0.05	0.05	
Jet energy corrections (quad. sum)	0.15	0.002	0.87	0.17	0.003
– JEC: InterCalibration	(0.03)	(<0.001)	(0.17)	(0.05)	(0.001)
– JEC: MPFInSitu	(0.01)	(<0.001)	(0.24)	(0.07)	(0.001)
– JEC: Uncorrelated	(0.15)	(0.003)	(0.82)	(0.15)	(0.003)
Jet energy resolution	0.08	0.001	0.05	0.04	0.001
b tagging	0.04	0.000	0.02	0.03	0.00
Pileup	0.07	0.001	0.04	0.04	0.001
Non- $t\bar{t}$ background	0.06	<0.001	0.08	0.07	<0.001
Modeling of hadronization					
JEC: Flavor (linear sum)	0.40	0.001	0.28	0.37	0.001
– light quarks (uds)	(+0.11)	(-0.001)	(-0.02)	(+0.07)	(-0.001)
– charm	(+0.03)	(<0.001)	(-0.01)	(+0.01)	(<0.001)
– bottom	(-0.33)	(<0.001)	(-0.31)	(-0.32)	(<0.001)
– gluon	(-0.21)	(0.003)	(+0.06)	(-0.13)	(0.002)
b-jet modeling (quad. sum)	0.24	0.001	0.12	0.20	<0.001
– b fragmentation Bowler-Lund	(0.07)	(0.001)	(0.01)	(0.05)	(<0.001)
– b fragmentation Peterson	(0.12)	(0.001)	(0.07)	(0.10)	(<0.001)
– semi-leptonic B hadron decays	(0.20)	(0.001)	(0.10)	(0.17)	(0.001)
Hard scattering and radiation					
PDF	0.03	0.000	0.01	0.03	0.00
Ren. and fact. scale	0.03	0.001	0.02	0.02	0.001
ME/PS matching threshold	0.06	0.001	0.07	0.03	0.001
ME generator	0.01	0.005	0.38	0.11	0.003
ISR PS scale	0.15	0.002	0.05	0.10	0.001
FSR PS scale	0.09	0.003	0.19	0.01	0.002
Top quark transverse momentum	0.03	0.001	0.06	<0.01	0.001
Modeling of non-perturbative QCD					
Underlying event	0.30	0.002	0.11	0.24	0.002
Early resonance decays	0.34	0.008	0.34	0.14	0.006
Color reconnection modeling	0.42	0.002	0.26	0.37	0.001
Total systematic	0.82	0.011	1.12	0.67	0.009
Statistical (expected)	0.15	0.001	0.09	0.13	0.001
Total (expected)	0.84	0.011	1.12	0.69	0.009

8.6. Results

The top quark mass m_t and jet energy scale JSF are measured in this thesis using the ideogram method, which is applied to the LHC pp collision data collected by CMS at a center of mass energy of 13 TeV. The corresponding integrated luminosity is 35.9 fb^{-1} .

Considering the performed calibration of the utilized method, obtained using simulated events, the measured quantities with the corresponding statistical and systematical uncertainties are given. The statistical uncertainty is calculated applying the condition $-2\Delta\ln\mathcal{L}(m_t, \text{JSF}|\text{sample}) = 1$. The results are performed for three approaches in the combine lepton+jets decay channel as well as individually in the muon+jets and electron+jets cases.

2D approach results: Using the 161 496 events in the lepton+jets decay channel, the 2D approach yields

$$\begin{aligned} m_t^{2\text{D}} &= 172.40 \pm 0.09 \text{ (stat+JSF)} \pm 0.71 \text{ (syst)} \text{ GeV}, \\ \text{JSF} &= 0.994 \pm 0.001 \text{ (stat)} \pm 0.010 \text{ (syst)}. \end{aligned}$$

When the same method is used with the 101 992 events in the muon+jets decay channel, the approach yields

$$\begin{aligned} m_t^{2\text{D}} &= 172.44 \pm 0.11 \text{ (stat+JSF)} \pm 0.73 \text{ (syst)} \text{ GeV}, \\ \text{JSF} &= 0.995 \pm 0.001 \text{ (stat)} \pm 0.010 \text{ (syst)}, \end{aligned}$$

and for the 59 504 events in the electron+jets decay channel, the 2D approach gives

$$\begin{aligned} m_t^{2\text{D}} &= 172.32 \pm 0.16 \text{ (stat+JSF)} \pm 0.84 \text{ (syst)} \text{ GeV}, \\ \text{JSF} &= 0.993 \pm 0.001 \text{ (stat)} \pm 0.011 \text{ (syst)}. \end{aligned}$$

The 2D likelihood obtained from data is shown in Fig. 8.21 (left) together with the statistical uncertainty contours demonstrating the one, two and three σ intervals. The expected statistical uncertainty in the top quark mass from the pseudo-experiments is presented in Fig. 8.21 (right) and are compared to the one measured in data. Additionally, the obtained value of the JSF is compared to the jet energy scale obtained from events with a Z boson and energetic photon (as shown in Section 6.5), and agrees within its uncertainties. This provides an independent cross-check of the JEC.

1D approach results: Since the default JEC scale from CMS is used in the 1D method $\text{JSF} = 1$ (with systematics evaluated by $\pm 1\sigma$ shifts), only the top

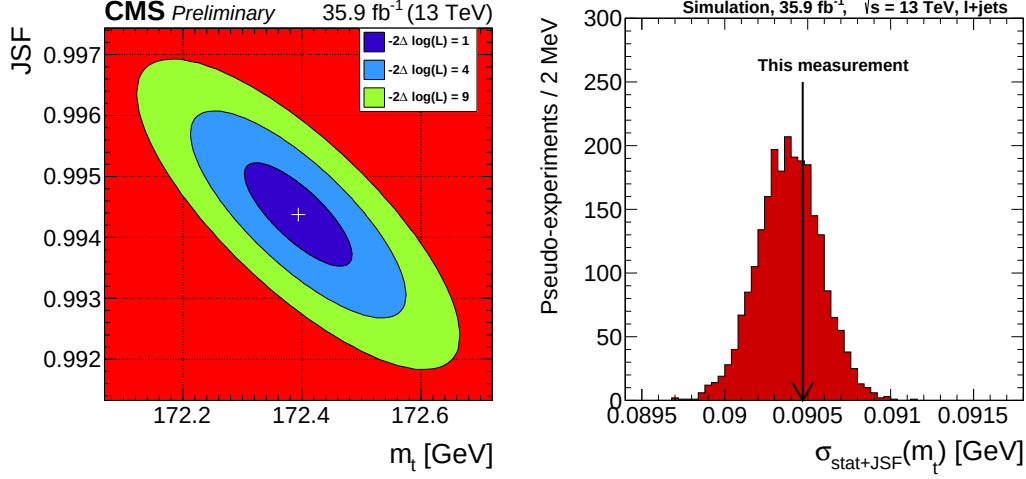


Figure 8.21.: The 2D likelihood ($-2\Delta \log(\mathcal{L})$) measured in the lepton+jets final state (left). The ellipses correspond to statistical uncertainties in m_t and JSF of one, two, and three standard deviations [10]. The statistical uncertainty calculated using 10 000 pseudo experiments and compared to the uncertainty in the data measurement (right).

quark mass values are obtained. Using the 161 496 events in the lepton+jets decay channel for this method yields

$$m_t^{\text{1D}} = 171.93 \pm 0.06 \text{ (stat)} \pm 1.09 \text{ (syst)} \text{ GeV.}$$

Using the 101 992 events in the muon+jets decay channel, the 1D method yields

$$m_t^{\text{1D}} = 172.06 \pm 0.07 \text{ (stat)} \pm 1.16 \text{ (syst)} \text{ GeV,}$$

and using the 59 504 events in the electron+jets decay channel, this approach gives

$$m_t^{\text{1D}} = 171.71 \pm 0.09 \text{ (stat)} \pm 1.12 \text{ (syst)} \text{ GeV.}$$

The results obtained with the 1D method have smaller statistical uncertainties compared to the 2D one. This is due to the fact that in the 2D approach m_t and JSF are measured simultaneously and the statistical uncertainty in m_t combines the statistical uncertainty coming from both the m_t and JSF components. As expected, the total systematical uncertainty is found to be larger in the 1D than in the 2D method, owing to the jet energy corrections uncertainty.

Hybrid approach results: Finally, 1D and 2D results are combined by the hybrid method, which assumes the prior $P(\text{JSF})$ to be a Gaussian centered at 1.0 with

a width $\sigma_{\text{prior}} = \delta\text{JSF}_{\text{stat}}^{2\text{D}} \sqrt{1/w_{\text{hyb}} - 1} = 0.0012$, where $\delta\text{JSF}_{\text{stat}}^{2\text{D}}$ is the statistical uncertainty obtained with the 2D method, and $w_{\text{hyb}}^{1\text{D}}$, given in Eq. 8.26, is the relative weight associated to the 1D approach. The analysis is performed for several hybrid weights as shown in Tab. 8.22. The presented total¹⁵ systematic uncertainty in m_t in Tab. 8.22 does not include color reconnection, early resonance decays and renormalization and factorization scale uncertainties, since they are envelope-based. As expected, the minimum of the JEC uncertainty is found to be at a lower value of $w_{\text{hyb}}^{1\text{D}}$ which is close to the 2D result, however, the total systematic uncertainty has a minimum for a slightly larger value of the hybrid weight. Therefore, in order to optimize the final result, $w_{\text{hyb}}^{1\text{D}} = 0.3$ was chosen, which leads to the smallest total systematic uncertainty.

Table 8.22.: The scan of the $w_{\text{hyb}}^{1\text{D}}$ hybrid weights used in the hybrid approach together with the total systematic uncertainty (middle row) and the JEC uncertainty (last row). In case of total systematic uncertainty the scan was performed without color reconnection, early resonance decays and renormalization and factorization scale uncertainties, since they are envelope-based.

w_{hyb}	0.2	0.25	0.3	0.4	0.5
$\delta m_t^{\text{hyb}} (\text{total})(\text{GeV})$	0.51	0.51	0.50	0.51	0.54
$\delta m_t^{\text{hyb}} (\text{JEC only}) (\text{GeV})$	0.10	0.14	0.19	0.27	0.36

The hybrid approach used on all 161 496 events in the lepton+jets decay channel yields the results:

$$\begin{aligned} m_t^{\text{hyb}} &= 172.25 \pm 0.08 (\text{stat+JSF}) \pm 0.62 (\text{syst}) \text{ GeV}, \\ \text{JSF}^{\text{hyb}} &= 0.996 \pm 0.001 (\text{stat}) \pm 0.008 (\text{syst}), \end{aligned}$$

while using the 101 992 events in the muon+jets decay channel in the hybrid method yields result of:

$$\begin{aligned} m_t^{\text{hyb}} &= 172.33 \pm 0.10 (\text{stat+JSF}) \pm 0.64 (\text{syst}) \text{ GeV}, \\ \text{JSF}^{\text{hyb}} &= 0.997 \pm 0.001 (\text{stat}) \pm 0.008 (\text{syst}). \end{aligned}$$

Finally, the hybrid approach used on the 59 504 electron+jets events yields

$$\begin{aligned} m_t^{\text{hyb}} &= 172.13 \pm 0.14 (\text{stat+JSF}) \pm 0.69 (\text{syst}) \text{ GeV}, \\ \text{JSF}^{\text{hyb}} &= 0.995 \pm 0.001 (\text{stat}) \pm 0.0009 (\text{syst}). \end{aligned}$$

¹⁵For the total systematic uncertainty, the scan was performed without color reconnection, early resonance decays and renormalization and factorization scale uncertainties, since they are envelope-based.

The results given by hybrid approach benefits from a smaller statistical uncertainty compared to the 2D results as well as a smaller total systematic uncertainty compared to the 1D and 2D approaches.

Similarly to the 2D result, the hybrid likelihood obtained from data is shown in Fig. 8.22.

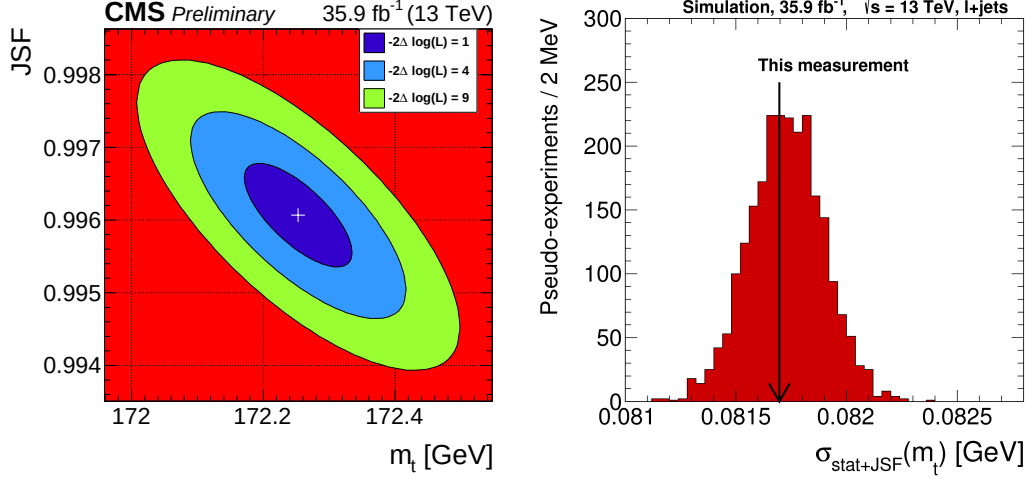


Figure 8.22.: The hybrid likelihood ($-2\Delta \log(\mathcal{L})$) measured for the lepton+jets final state with JSF constrained using the $\delta m_t^{\text{hyb}} = 0.3$ (left). The ellipses correspond to contours of $-2\Delta \log(\mathcal{L}) = 1(4, 9)$ allowing the construction of the one (two, three) σ statistical intervals of m_t [10]. The statistical uncertainty calculated using 10 000 pseudo experiments and compared to the uncertainty of the data measurement (right).

8.7. Comparison with other measurements

The presented hybrid result in this thesis is compared with previous measurements of the top quark mass. Figure 8.23 shows the current result together with CMS published results from the direct m_t measurement separately for three decay channel, as well as the legacy measurement. In addition, the Tevatron combination and World combination are also shown. More details on the m_t measurements performed recently by other experiments is given in the followings.

One of the challenges of this analysis was: if with new NLO ME generators and a better understanding of the modeling uncertainties, a better precision in the m_t measurement can be achieved in comparison to the most accurate measurement performed by CMS using Run 1 data [7] with $m_t = 172.35 \pm 0.16$ (stat+JSF) ± 0.48 (syst) GeV. The new result shows slightly worse precision arising from the modeling of the color reconnection (CR) effect, where the CR uncertainty is estimated using recently available CR models and is found to be 3 times larger than previously determined one.

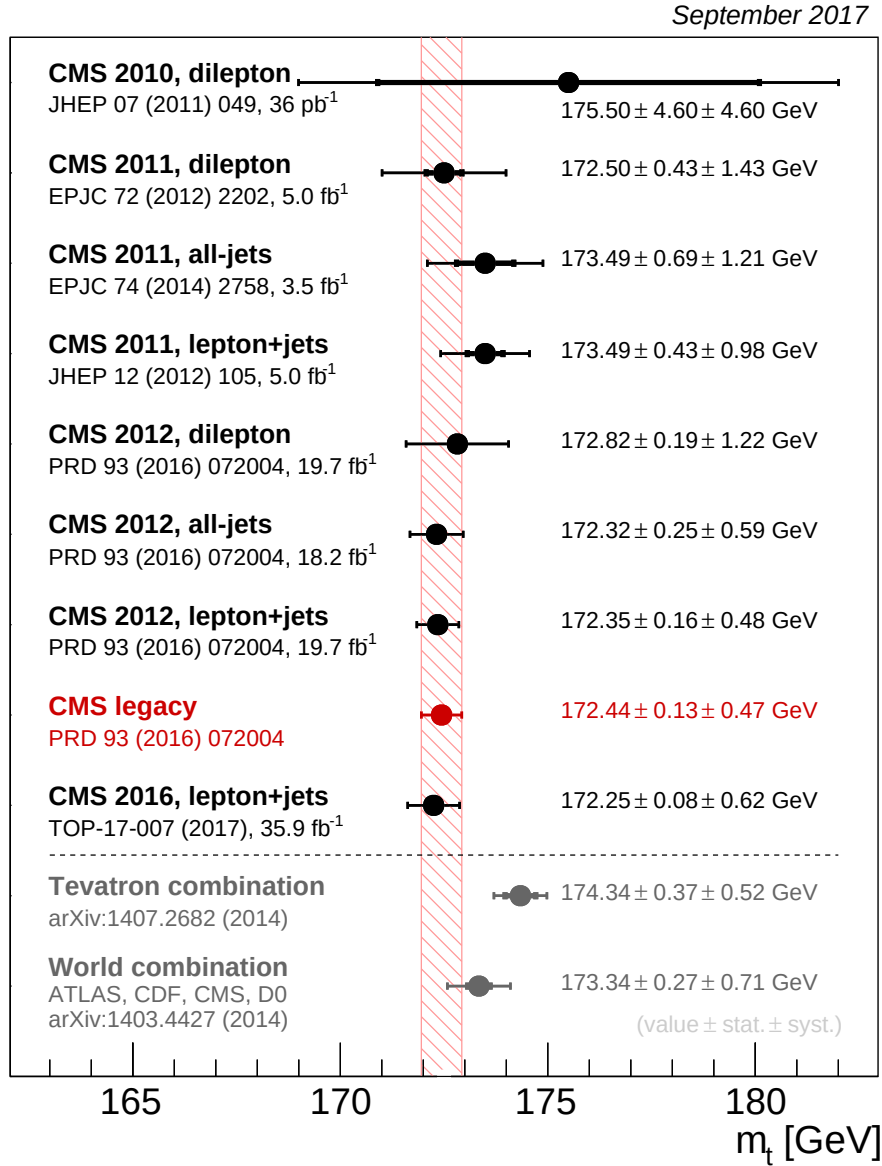


Figure 8.23.: The presented hybrid result in this thesis is compared with previous measurements of the top quark mass [48]. The Run 1 CMS standard measurements are summarized in the CMS legacy result. The current result is shown above the solid line.

Results on the m_t in the dilepton at CMS at 8 TeV [7] are based on analytical matrix-weighting technique [228] that provides a larger systematic uncertainty and yields $m_t = 172.82 \pm 0.19$ (stat) ± 1.22 (syst) GeV. While the results in all-jets decay channels at CMS in Run 1 [7] that utilizes the same framework of ideogram method provide compatible small uncertainties, thus $m_t = 172.32 \pm 0.25$ (stat) ± 0.59 (syst) GeV. The analysis in this decay channel for 13 TeV is ongoing.

Recently, ATLAS collaboration presented a new result on top quark mass measure-

ment in the lepton+jets decay channel yielding $m_t = 172.08 \pm 0.39$ (stat+JSF+bJSF) ± 0.82 (syst) GeV [8]. For this measurement a three-dimensional template method was used. This determines in addition to the top quark mass and jet energy scale factor, a relative b-to-light-jet energy scale factor (bJSF). To reconstruct bJSF the distribution of the $R_{qb}^{\text{reco}} = (p_T^{b_{\text{had}}} + p_T^{b_{\text{lep}}}) / (p_T^{q_1} + p_T^{q_2})$ was used, where q_1 and q_2 are the light jets from W boson. Such a simultaneous measurement reduces the JES and bJES uncertainties in the top quark mass, where bJES aims to reduce the flavor-dependent JEC uncertainty. However, this cause a larger uncertainty due to b tagging. Concerning the other decay channels, the results in dilepton decay channel at 8 TeV were measured with unbinned max-likelihood and yields $m_t = 172.99 \pm 0.41$ (stat) ± 0.74 (syst) GeV [240]. The top quark mass in all-jets decay channel at 8 TeV was measured with binned min- χ^2 fit which yields $m_t = 172.99 \pm 0.41$ (stat) ± 0.74 (syst) GeV [241].

Even though 23 years passed since the top quark was discovered at Tevatron, there are still many analyses exploiting different decay channels. The latest result on the m_t measurement on the center-of-mass energy of 1.96 TeV in the lepton+jets decay channel is done by the D0 collaboration [4], and utilize a matrix-element method (MEM). The MEM method uses full event kinematic information event-by-event and calculates the likelihood for m_t and k_{JSF} based on LO matrix element for $t\bar{t}$ production, thus by minimizing the obtained likelihood for the whole sample yields $m_t = 174.98 \pm 0.58$ (stat+JSF) ± 0.49 (syst) GeV. This the most precise Tevatron result on m_t . Relatively new results in dilepton decay channel were also performed at D0. They are based on two different methods, namely neutrino weighting template method [228] and matrix element method, both of them apply a JES constraint from lepton+jets to the energy of jets in data. The combination of both results yields $m_t = 173.50 \pm 1.31$ (stat) ± 0.84 (syst) GeV [242].

The CDF collaboration performs the measurement of the m_t in lepton+jet decay channel using the template fit, which is build by the reconstructed W mass and two masses from the both top quark candidates and yields $m_t = 172.85 \pm 0.71$ (stat) ± 0.85 (syst) GeV [5]. The similarly, the template method was used to measure m_t in the events with jets and missing transverse energy, where no identified lepton was required and 4 to 6 jets were selected, this yields $m_t = 173.93 \pm 1.64$ (stat+JSF) ± 0.87 (syst) GeV [243].

The world combination of the top quark mass was performed in [9]. It was based on the results from the ATLAS, CMS, CDF, and D0 experiments and yields $m_t = 173.34 \pm 0.27$ (stat) ± 0.71 (syst) GeV. The LHC pp collision data used for the combination correspond to an integrated luminosity of up to 4.9 fb^{-1} and were recorded at a center-of-mass energy of 7 TeV during Run 1. The Tevatron $p\bar{p}$ data used for the combination correspond to an integrated luminosity of up to 8.7 fb^{-1} and were collected at a center-of-mass energy of 1.96 TeV during Run 2.

Moreover, the SM according to its invariance under CPT transformation pre-

dicts the equality of the mass of the particle and antiparticle. The study [244] provides the latest top quark mass asymmetry $\Delta m = m_t - m_{\bar{t}}$ measurement at a center-of-mass energy of 8 TeV by the CMS collaboration. It measured $\Delta m = -0.15 \pm 0.19$ (stat) ± 0.09 (syst) GeV, which is consistent with the SM expectation.

8.8. Theoretical interpretation of the result

As shown in [39, 245, 246] the SM model appears to be at the border of stability and instability regions of the electroweak vacuum. Thus, with a measured mass of $m_t^{\text{hyb}} = 172.25 \pm 0.08$ (stat+JSF) ± 0.62 (syst) GeV the electroweak vacuum is closer to be stable and is safely below the instability region. For the case m_t higher than 176 GeV the vacuum is in the unstable regime, higher than 171 GeV in the metastable and below this value in the stable region. Thus, according to results on m_t , no indication of new physics below the Planck scale is observed and the stability of the Higgs vacuum should be considered in a cosmological framework when studying, e.g., inflationary processes.

The top quark is a colored object, thus its mass can not be defined in terms of the mass of its decay products, since any final-state particle system can not be explicitly associated with it. However, there is a relation of m_t to the mass of the system of objects from the top decay. Therefore, the distribution of the mass of these systems can be used to measure m_t , where m_t enters as a parameter, as was done in the presented analysis in this thesis. In order to compare the measured mass with a theoretical prediction, it is necessary to find out within which mass schemes discussed in Section 2.2.2 the measured m_t appears as a parameter.

The current parton shower MC generators are accurate at the leading order, which does not allow to distinguish the mass parameter schemes that only differ at the next-to-leading-order (NLO) and beyond, e.g., pole mass and $\overline{\text{MS}}$ mass. This statement is essential in case the perturbative accuracy is concerned. However, in the recent study presented in [203], the estimation of the theoretical uncertainties in the determination of the top quark mass at LHC using NLO+PS generators interfaced with PYTHIA 8.2 was reported. This study aims to estimate if (neglecting experimental uncertainties) the mass peak position can be shifted due to the influence of potentially unknown perturbative and non-perturbative effects. Thus, it was found out that within PYTHIA 8.2 shower studies, which also includes the comparison among the different NLO+PS generators, no significant (about 30 MeV) peak shifts at the particle-truth level were observed. The obtained top quark mass value reported in this paper is very close to the definition of m_t measured in this work. Thus, the results of the measurements presented in this thesis can be interpreted in terms of pole mass.

In the study presented in [207], the procedure to calibrate the measured top quark mass into the well-known mass scheme using a fit of MC predictions to the

hadron level QCD measurement is presented. The predictions are obtained for the variable referred to as 2-Jettiness measured in e^+e^- annihilation. The results were obtained for the m_{pole} and $m_{\text{MSR}}(1\text{ GeV})$ at the next-to-leading logarithmic (NLL) and next-to-next-to-leading logarithmic (NNLL) accuracy to PYTHIA 8.2. According to this study the measured top quark mass value agrees with the pole mass within its uncertainty.

Moreover, the propagator pole of the top quark is at the complex position $\sqrt{p^2} = m_{\text{pole}} - \frac{i}{2}\Gamma_t$, where p denotes the four-momentum and Γ_t decay width of the top quark. In the fixed-order calculation the pole of the propagator gives the peak in the region of the Wb invariant mass distribution when $p^2 \approx m_{\text{pole}}^2$. Since the $m_{\text{pole}} \gg \frac{\Gamma_t}{2}$, this peak can be related to the pole mass.

8.9. Summary and outlook

The obtained results $m_t = 172.25 \pm 0.08 (\text{stat+JSF}) \pm 0.62 (\text{syst}) \text{ GeV}$ is the first top quark mass measurement using the full data sample recorded in 2016 during LHC Run 2 period. In this analysis, for the first time, the new NLO generators were utilized. The result is obtained using the similar approach as the one used for the Run 1 analysis performed at 8 TeV. The higher precision on the statistical uncertainty is due to the larger integrated luminosity and the higher $t\bar{t}$ cross section at $\sqrt{s} = 13 \text{ TeV}$ compared with the Run 1 result. However, due to the more evolved treatment of the modeling uncertainties, i.e., color reconnection, the systematic uncertainties are found to have slightly larger values. Corresponding to the theoretical studies [203, 207] the obtained m_t value can be interpolated in terms of the pole mass scheme. The presented measurement was already released by the CMS collaboration [10] in a public preliminary result.

The accuracy of the top quark measurement depends significantly on the modeling uncertainty. Thus, better understanding and modeling, in particular the non-perturbative effects, will not only allow a more accurate measurement of m_t , but also will help to interpret the measured value in terms of the well-known mass schemes. The difference between the results obtained with PYTHIA and HERWIG generators are found to be dramatically large. Moreover, in the theoretical study mentioned before [203] NLO+PS generators interfaced with both PYTHIA 8.2 and HERWIG 7.1 were used, where the difference between the results was also visible. Hence, more studies in this area are needed. Note that the jet flavor uncertainties for the jet energy correction are estimated using exactly PYTHIA and HERWIG generators. As was already shown the uncertainty in the top quark mass due to the flavor jet energy correction has the biggest impact on the result.

9. Measured top quark mass as a function of kinematic observables

The precise measurements of the top quark mass are effected by the modeling of perturbative and soft QCD effects, as was shown during the systematic uncertainties investigation in Section 8.5. One way to account for the potential limitations of current event generators as well as to validate the provided models is to perform the differential measurement of the investigated quantity as a function of the kinematic properties. This procedure can also facilitate a good cross-check for the possible biases in the measurements.

The differential study requires to have a large dataset in order to consider some possible deviation, thus to check stability of the measured top quark mass. The data collected at a center-of-mass energy of 13 TeV in 2016, corresponding to an integrated luminosity of 35.9 fb^{-1} , aims to maximize the accuracy of the results compared to the previous data from Run 1 [7].

9.1. Methodology

The top quark mass measurement presented in this thesis is very sensitive to the choice of the ME generator as well as to the color reconnection (CR) model. Thus, by measuring m_t as a function of the kinematic properties of the $t\bar{t}$ system the perturbative QCD effects modeled in the different MC generators and new provided CR models are validated. The validation is performed by comparing the results from MC to the measurements on data.

In order to probe for potential effects from the kinematics of jets coming from the top quark decays, the number of jets with $p_T > 30 \text{ GeV}$ (N_{jets}), the p_T and the pseudorapidity of the b jet assigned to the hadronic decay branch ($p_T^{\text{b, had}}$ and $|\eta^{\text{b, had}}|$), the ΔR between the b jets ($\Delta R_{b\bar{b}}$), and the ΔR between the light quark jets ($\Delta R_{q\bar{q}}$) are used. These variables can also probe effects from color reconnection and initial- and final-state radiation, e.g., b quark carry the color charge of top quark and can be color connected to the ISR or FSR, or the light quarks from the W boson that are color connected can provide some effect to the CR. The variables of the transverse momentum of the hadronically decaying top quark ($p_T^{\text{t, had}}$), the invariant mass of the $t\bar{t}$ system ($m_{t\bar{t}}$), the transverse momentum of the $t\bar{t}$ system ($p_T^{\text{t}\bar{t}}$) are also used, since the measured m_t may depend from the kinematic properties of the

top quark.

In order to study the potential effects from the modeling of perturbative QCD data are compared to pseudo-experiments with $t\bar{t}$ samples from different generator setups. Thus, to check two different ME generators, $t\bar{t}$ samples generated with MADGRAPH5_aMC@NLO [95] at LO with MLM [103] matching and at NLO with FxFx [104] matching are used. Both setups are interfaced to PYTHIA8 [16]. Additionally, as the inclusive m_t results from HERWIG++ [102] samples differ very much from the nominal results, it is important to perform the current study using POWHEG v2 [96] interfaced to HERWIG++. In order to test new color reconnection models, the results from POWHEG v2 interfaced to PYTHIA8 samples with the default tune, as well as with the “QCD inspired” model [14], and the “gluon move” model [15] are considered. In addition, the samples with the default tune are simulated with early resonance decays in PYTHIA8. This option was also used during the simulation of the samples with the “QCD inspired” and “gluon move” models.

Data and various MC samples are divided into three to five bins of the previously mentioned kinematic variables. The bins are filled with all permutations which lie within the boundaries of each bin. In order to account for the fact, that some variables may depend on the jet-quark assignment, like reconstructed top quark p_T or η , which cannot be resolved in an unambiguous way, some simplifications are allowed. This will permit for a single event to contribute to multiple bins. The technique, analogi to the inclusive measurements, is used to measure m_t and JSF in each bin of the kinematic variables.

9.2. Method calibration

An additional calibration needs to be applied to the value of m_t and JSF obtained in each bin, to allow these results to be interpreted in terms of its agreement with the corresponding observable from the inclusive measurement. This due to the possibly different sensitivity of the measurement to the different phase-space regions, with regard to the input m_t and JSF. For example the mean value of the m_t^{fit} or m_W^{reco} distributions can be slightly shifted for different kinematic regions. Therefore, a bin-by-bin calibration of the results is required. It is performed for each individual bin analogous to the procedure described in Section 8.4.3 and applied to the values extracted from the pseudo-experiments. After the calibration procedure, the calibrated top quark mass $m_{t,\text{cal}}$ and jet scale factor JSF_{cal} are obtained for each bin of the kinematic variable.

9.3. Method implementation

The results on the phase-space dependence of the measured $m_{t,\text{cal}}$ and JSF_{cal} are given as $m_{t,\text{cal}} - \langle m_t \rangle$ or $\text{JSF}_{\text{cal}} - \langle \text{JSF} \rangle$, where mean values $\langle m_t \rangle$ and $\langle \text{JSF} \rangle$ corre-

sponds to the inclusive result obtained in the corresponding sample. It is performed in this way, since the systematic uncertainties are assumed to be correlated among all bins and by subtracting $\langle m_t \rangle$ any possible constant shift is removed. These correlated shifts on the value of $m_{t,\text{cal}}$ measured in each bin arise due to the nature of systematic uncertainties of the inclusive measurement. Moreover, the different m_t in data, which correspond to the values shown in Section 8.6, and in MC simulations, which correspond to $m_t = 172.5 \text{ GeV}$, may also cause some correlated shifts on measured $m_{t,\text{cal}}$ in bins of kinematic variables. By using this approximation, an underestimation of the uncertainties from the p_T/η -dependent JES may occur.

The statistical uncertainties in the results from data and MC are correlated. This occurs due to the correlation arising from the subtraction of the mean values, since the statistical uncertainties from both the measured value and the inclusive value contribute. The statistical and dominant systematic uncertainties are combined in quadrature separately for each bin. The dominant systematic uncertainties corresponds to the uncertainty coming from JEC (p_T -, η - and flavor-dependent), JER, pileup, b fragmentation, renormalization and factorization scales, ME to PS matching, ISR and FSR parton-shower scales, and the underlying event.

In this chapter, only the results for the top quark mass calculated with the hybrid approach in each bin of the kinematic variable are shown, while the similar results for m_t calculated with the 1D and 2D approach, as well as results for the JSF are shown in Appendix A.8.

The level of agreement between data and MC results for each variable and model is quantified by a cumulative χ^2 that is defined as:

$$\chi^2 = \sum \frac{(m_t(\text{data}) - m_t(\text{MC}))^2}{\delta_{m_t,\text{data}}^2 + \delta_{m_t,\text{MC}}^2}, \quad (9.1)$$

where $\delta_{m_t,\text{data}}$ and $\delta_{m_t,\text{MC}}$ is the total uncertainty in the data and the statistical uncertainty in the model values. As the mean measured m_t is subtracted, the number of degrees of freedom (ndf) for each variable is equal the total number of bins minus one. While the p-values (χ^2 probabilities) is obtained by taking the ratio of χ^2 to ndf.

9.4. Results

The results for the chosen kinematic variables and the calibrated mass difference $m_{t,\text{cal}} - \langle m_t \rangle$ are presented in this section. Series of differential measurements of m_t in data are compared to simulations based on LO (MADGRAPH) and NLO (POWHEG and MC@NLO) matrix element calculation and to simulations with different PS calculation, e.g., PYTHIA8 and HERWIG++, as well as to simulations with alternative models of color reconnections. In the following the results are grouped according to their sensitivity to color reconnection, ISR and FSR, as well as to the kinematics of

the b quark from the hadronically decaying top quark. The biases are well described by the various simulation samples, however the results from POWHEG matched with HERWIG++ do not describe data well.

Color reconnection: The modeling of the color flow can be studied by looking into the color field between the light quarks from the hadronically decaying W boson as well as between the top quark and the color connected spectator partons. The sensitivity of m_t against the opening angle $\Delta R_{q\bar{q}}$ between two light jets from hadronic W in terms of separation in $(\eta - \phi)$ space are shown in Fig. 9.1.

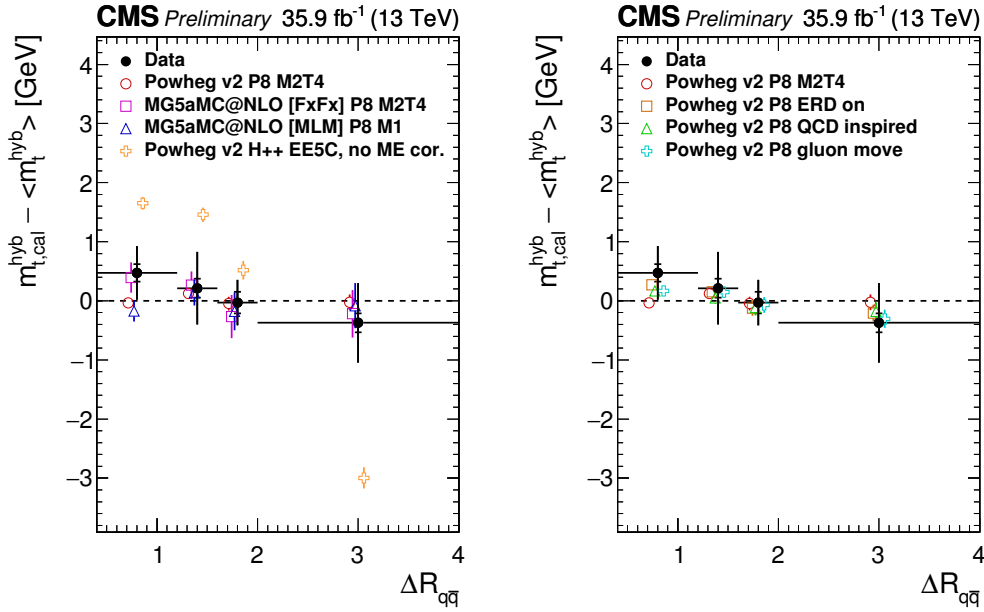


Figure 9.1.: Measurements of m_t as a function of ΔR between the light quark jets ($\Delta R_{q\bar{q}}$) compared to different generator choices (left) and to alternative models of color reconnections (right) [10].

The sensitivity of m_t against transverse momentum and pseudorapidity of the top quarks are shown in Fig. 9.2. No significant dependence is observed. MC results agree with data within the total uncertainty.

Initial and final state radiation: The effect of additional radiation in the event on the measurement of the top quark mass can be estimated by looking into the jet multiplicity as well as by looking into the invariant mass and transverse momentum of the $t\bar{t}$ system. The results are shown in Fig. 9.3 and Fig. 9.4.

Even though the effect of the softer radiation may be excluded in the distribution of the jet multiplicity by rejecting jets with transverse momentum less than 30 GeV, some indications for the small sensitivity are observed. This is especially visible in case of the JSF as a function of increasing jet multiplicity,

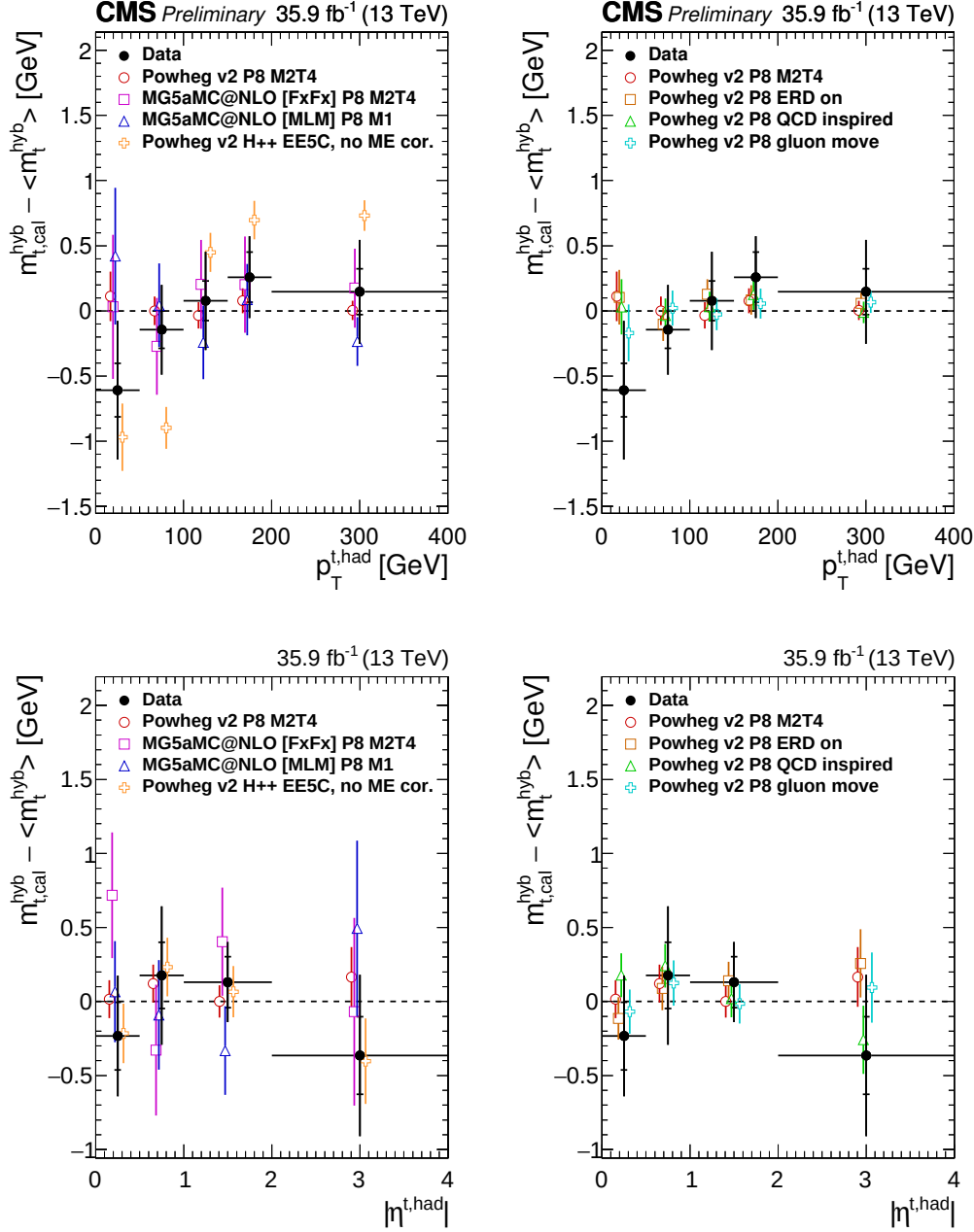


Figure 9.2.: Measurements of m_t as a function of the transverse momentum of the hadronically decaying top quark ($p_T^{t,\text{had}}$) (upper row) and of the pseudorapidity of the hadronically decaying top quark ($\eta^{t,\text{had}}$) (lower row) compared to different generator choices (left) and to alternative models of color reconnections (right) [10].

shown in the lower row of Fig. 9.3. The measurement of the differential top quark mass shows a small dependence on the p_T of the $t\bar{t}$ system and is well described by all MC samples. However, in case of the $t\bar{t}$ mass distribution the results for POWHEG matched with HERWIG++ show a different behavior.

Bottom quark observables: The bottom quarks carry the color charge from the top

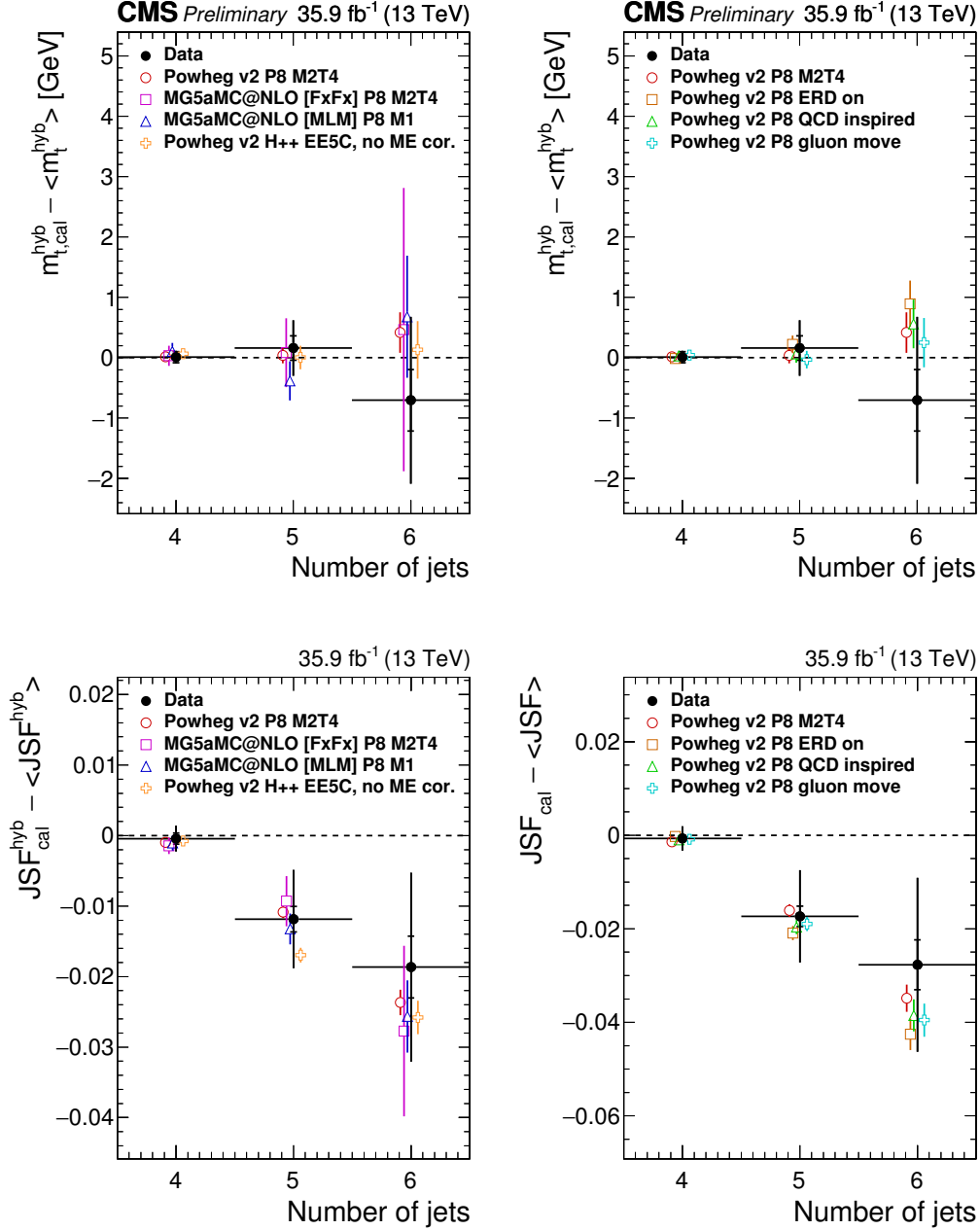


Figure 9.3.: Measurements of m_t (upper row) and JSF (lower row) as a function of the number of jets with $p_T > 30$ GeV compared to different generator choices (left) and to alternative models of color reconnections (right) [10].

quarks, therefore due to the color confinement it must be color connected to the ISR or to the MPI as shown in Section 5.1. Since the b quark is used for the top quark mass calculation, the sensitivity of m_t due to the b quark kinematics is expected. Therefore, the transverse momentum and pseudorapidity of the b jet from the hadronic top quark is studied. In addition, the behavior of the distribution of the angular separation of the b jets from both top quarks are studied on the $(\eta - \phi)$ plane. Figure 9.5 shows the spatial correlations

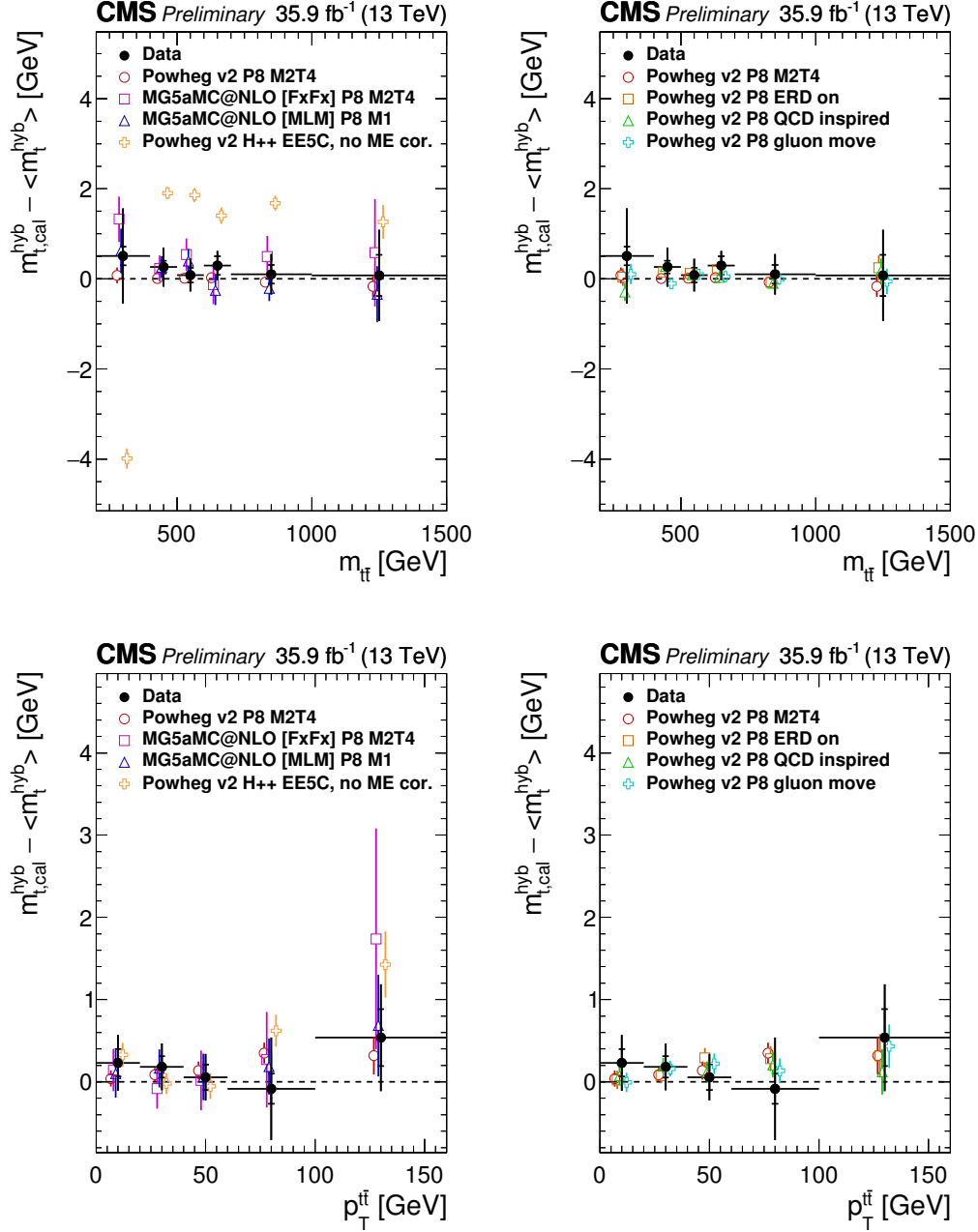


Figure 9.4.: Measurements of m_t as a function of the invariant mass of the $t\bar{t}$ system ($m_{t\bar{t}}$) (upper row) and the transverse momentum of the $t\bar{t}$ system ($p_T^{t\bar{t}}$) (lower row) compared to different generator choices (left) and to alternative models of color reconnections (right) [10].

between two b's from the two top quarks. No significant differences are seen in the behavior of m_t .

The results for the b-jet p_T and η are shown in Fig. 9.6, where some small disagreement between data and MC is visible in the distribution of the b jet pseudorapidity. In case of p_T of the b jet the results from sample of POWHEG matched with HERWIG++ show quite different behavior.

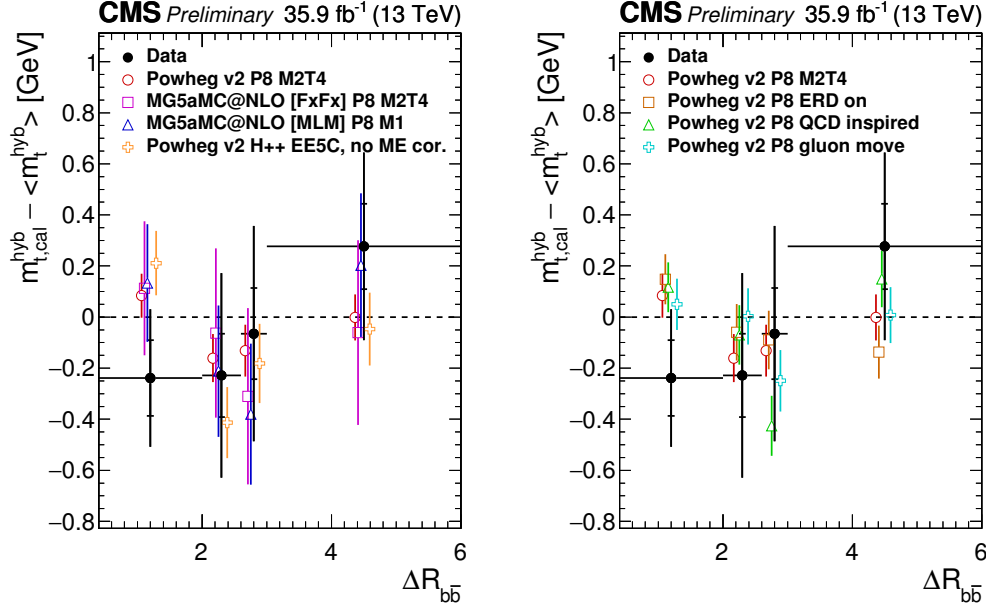


Figure 9.5.: Measurements of m_t as a function of ΔR between the b jets ($\Delta R_{b\bar{b}}$) compared to different generator choices (left) and to alternative models of color reconnections (right) [10].

The simulation to data agreement for each model and each variable is qualified in terms of χ^2 probabilities that are summarize in Tab. 9.1. Within the improved statistical precision of Run 2 data, no significant deviation for m_t is observed.

Table 9.1.: Compatibility of different models with the differential measurement of the top quark mass on data.

Model	χ^2 probability								
	$p_T^{t,\text{had}}$	$\eta^{t,\text{had}}$	$m_{t\bar{t}}$	$p_T^{t\bar{t}}$	N_{jets}	$p_T^{b,\text{had}}$	$ \eta^{b,\text{had}} $	$\Delta R_{b\bar{b}}$	$\Delta R_{q\bar{q}}$
POWHEG P8 M2T4	0.68	0.71	0.93	0.90	0.71	0.98	0.61	0.59	0.68
MG5 P8 [FxFx] M2T4	0.93	0.29	0.80	0.85	0.90	0.72	0.26	0.66	0.97
MG5 P8 [MLM] M1	0.49	0.40	0.79	0.99	0.39	0.97	0.16	0.68	0.57
POWHEG H++ EE5C	0.07	0.99	5×10^{-14}	0.53	0.73	2×10^{-4}	0.55	0.36	8×10^{-6}
POWHEG P8 ERD on	0.75	0.75	0.99	0.83	0.53	0.95	0.64	0.37	0.96
POWHEG P8 QCD inspired	0.80	0.78	0.93	0.94	0.66	0.99	0.71	0.48	0.89
POWHEG P8 gluon move	0.87	0.80	0.93	0.93	0.71	0.93	0.51	0.57	0.92

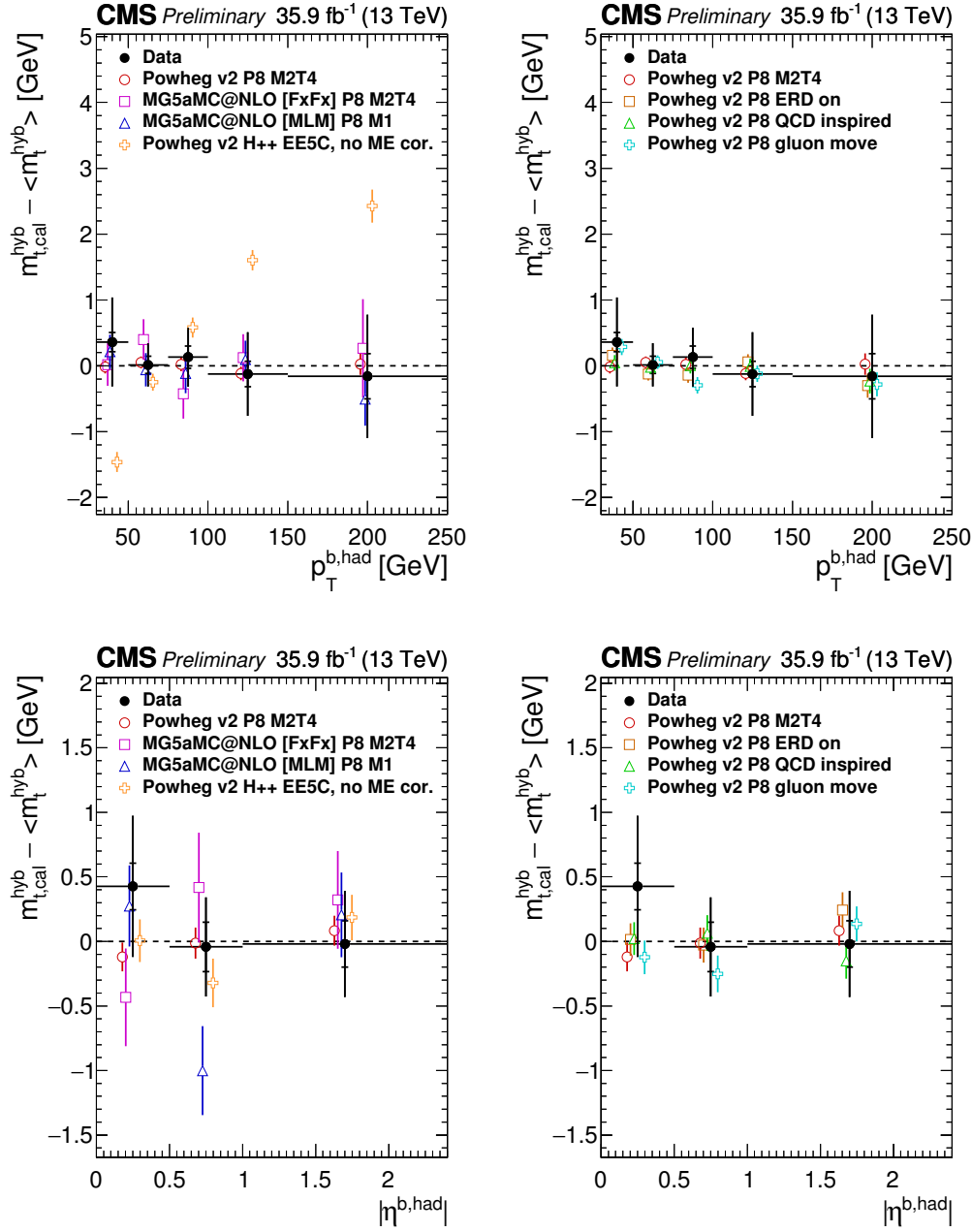


Figure 9.6.: Measurements of m_t as a function of the p_T of the b jet assigned to the hadronic decay branch ($p_T^{\text{b, had}}$) (upper row) and the pseudorapidity of the b jet assigned to the hadronic decay branch ($|\eta^{\text{b, had}}|$) (lower row) compared to different generator choices (left) and to alternative models of color reconnections (right) [10].

9.5. Summary

The measured top quark mass was studied as a function of kinematic observables using the full Run2 data set collected with the CMS detector in 2016, that corresponds to an integrated luminosity of 35.9 fb^{-1} . The simulated samples from a new ME generator as well as MC samples generated with a new CR tunes are used to probe possible biases arising from the color reconnection and ISF/FSR in the top quark mass measurement. No significant deviation between data and the default POWHEG v2 matched with PYTHIA8 generator setup was found, as well as no evidence for a bias in the measurement was found. The results are well described by the various simulation samples, however the results for $\Delta R_{q\bar{q}}$, $m_{t\bar{t}}$ and $p_T^{\text{b, had}}$ from POWHEG matched with HERWIG++ show large deviations. This fact confirms that the matrix element corrections to the top quark decay need to be included in POWHEG matched with HERWIG++ setup.

The uncertainty due to the color reconnection for the inclusive top quark mass measurement is three times large in comparison with Run 1 result [7]. This arise due to the new procedure of its estimation, where the samples with two new CR tunes are used. It was found that the difference between the results with the nominal MPI-based CR model and new gluon-move CR model causes such a large difference, as shown in Tab. 8.13. Therefore, the differential study of m_t could cross check the results from the corresponding MC sample. However, the data sample is still too small to rule out any one of the different models.

10. Conclusions

The precise measurement of the top quark mass, the heaviest known fundamental particle, is crucial for the prediction of the electroweak vacuum stability in the Standard Model (SM). It is also necessary to probe the self consistency of the SM and to constrain physics models beyond the SM at high energies.

Within the scope of this thesis, the first measurement of the top quark mass m_t using the full data sample collected in 2016 with the CMS detector at a center-of-mass energy of 13 TeV is presented. The data corresponds to an integrated luminosity of $35.9 \pm 0.9 \text{ fb}^{-1}$. An extremely pure sample of $t\bar{t}$ events is selected in the lepton+jets decay channel requiring the presence of exactly one muon or electron (including muons and electrons originating from τ lepton decays) and at least four jets in the final state. A b-tagging algorithm is used to reconstruct jets originating from b quarks. In order to enrich the fraction of correctly reconstructed top quarks, the decay products from selected events in the simulated samples are fitted with a kinematic fit to the $t\bar{t}$ hypothesis using the jet-quark assignment and the constraints from the known value of m_W , and the equality of the mass of the two top quarks. An ideogram method is used to measure the most probable value of the top quark mass and the jet energy scale factor (JSF) simultaneously. With this approach, the uncertainties arising from jet energy corrections are reduced, resulting in a more precise measurement of m_t . The measured values of m_t and JSF in a hybrid approach are:

$$\begin{aligned} m_t^{\text{hyb}} &= 172.25 \pm 0.08 \text{ (stat+JSF)} \pm 0.62 \text{ (syst)} \text{ GeV}, \\ \text{JSF}^{\text{hyb}} &= 0.996 \pm 0.001 \text{ (stat)} \pm 0.008 \text{ (syst)}, \end{aligned}$$

with a total uncertainty in m_t of 0.63 GeV, where the hybrid approach combines the information from the default jet energy correction obtained by CMS with the JSF constraint based on the W boson decay. The precision of the measurement is limited by systematic uncertainties, in particular, by the jet flavor energy correction and the modeling of the color reconnection.

In order to obtain the highest possible precision in m_t measurements, a better description of perturbative and non-perturbative effects is necessary. Thus, a new NLO matrix element generator [98] with a new underlying event tune CUETP8M2T4 [116, 117] is used for the first time for the m_t measurement. Moreover, MC generators with various color reconnection (CR) models [14, 123, 125] recently became

available. This allows a more reliable estimation of the color reconnection uncertainty by comparing the results for each of the color reconnection models, since there are no first-principles models to describe CR. In addition, new simulation samples with independent variations of initial and final state radiation (ISR and FSR), which allow the specification of the choice of α_S separately for ISR and FSR, are used.

Given the larger data sample and higher center-of-mass energy compared to the Run 1, one can cross-check the obtained results using well-defined top quark objects for different theoretical models. The results from PYTHIA 8.219 and HERWIG++ 2.4 are compared for consistency, since they are used for the jet flavor uncertainty estimation associated with the jet energy correction, which yields the largest contribution to the total uncertainty of the result. Samples with different matrix element generators are also compared to check for additional sources of variation. In addition, since a new procedure of uncertainty estimation associated with color reconnection effects is recommended [15], the total uncertainty of the m_t measurement is expected to be three times larger compared to the previous Run 1 results. The results regarding the new color reconnection models are also compared. No large deviations of m_t can be observed. Therefore, within the data uncertainties, none of the different models can be ruled out. The presented results using the full 2016 data recorded at the center-of-mass energy of 13 TeV have been released by the CMS collaboration as a public preliminary result [10, 247].

The precise measurement of the jet energy is an important requirement not only for the measurements of the top quark mass, but also for the potential discovery of physics beyond the SM. Thus, in addition to the main topic of this thesis, the residual differences in the jet response between data and simulation are measured. The correction scale factors are obtained for the pp collision data recorded with the CMS detector in 2015, corresponding to an integrated luminosity of 2.1 fb^{-1} . The response values are estimated as a function of pseudorapidity $|\eta|$ using dijet events with a back-to-back topology and applying two methods, namely the MPF and dijet p_T balance methods. The results from both methods are in good agreement and the obtained residual scale factors are consistent with unity within about 3% for $|\eta| < 2.5$ and 17% for $|\eta| > 2.5$. The obtained corrections are included into the chain of other CMS jet energy corrections, e.g., those arising from the pileup contribution or detector simulation, and are applied in all CMS analyses using 2015 data.

The precision of m_t relies directly on a description of the dependence interference between the decays of the top quark during hadronization and multiple parton interactions. Therefore, the modeling of the color reconnection becomes important for precision measurements, where the best way to carry out uncertainty estimation is to study different sets of CR models. Given the fact that new CR models have recently been made available, as well as the fact that more observables and measurements have recently been published, a new underlying event tune has been tested.

In the current thesis, the tuning for the PYTHIA 8.219 generator is performed based on the QCD-inspired color reconnection model [14]. The predictions obtained after applying this tune are able to accurately describe the underlying event observables as well as the charged particle multiplicities measured during the Run 1 and Run 2 LHC periods. Still, the strange particle yields measured at 7 TeV is not described well with presented CR tune, but nor is it with previous tunes. The corresponding tune parameters are available as supplementary materials in [10]. The MC samples with the corresponding tune were produced and used for the investigation of the color reconnection systematic uncertainty.

Since the top quark mass is measured from a fit to measured distributions, it is unavoidably affected by theoretical uncertainties. Therefore, a better understanding of modeling uncertainty is crucial for a reliable estimation of the systematic uncertainties. On the other hand, the better understanding and modeling, in particular the non-perturbative effects, will help to interpret the exact nature of the precise m_t measurement at hadron colliders in terms of the well-known mass schemes, such as the pole mass and $\overline{\text{MS}}$ mass scheme. The relation between the measured m_t and field-theory mass schemes is now actively studied, and according to [204], the measurement of the top quark mass presented in this thesis can be interpreted as the pole mass within ± 500 MeV.

The newly available abundance of data collected with the CMS detector is necessary for the differential study of the m_t as a function of kinematic observables. This will be key in testing new models and will allow the results to be checked for possible biases.

A. Appendix

A.1. Technicalities

Instantaneous luminosity

The instantaneous luminosity is calculated with brilcalc tool version 3.1.1:

```
brilcalc lumi -b "STABLE BEAMS"
```

```
--normtag /afs/cern.ch/user/l/lumipro/public/Normtags/normtag_DATACERT.json  
-i certified_runs.json.
```

The run is certified at the stage of the Data Quality Monitoring (DQM) by passing first the *online* step, where the basic detector properties, like channel qualities, data integrity and timing synchronization, are checked and *offline* step, where the quality of high-level objects, like particle multiplicity and momentum spectra, are verified.

Pileup

The pileup interactions are calculated, by using a BRIL Work Suite software [201] using the following command:

```
pileupCalc.py -Cert_271036-284044_13TeV_23Sep2016ReReco_Collisions16_JSON.txt  
--inputLumiJSON pileup\_JSON.txt --calcMode true  
--minBiasXsec 69400 --maxPileupBin 50 --numPileupBins 500 PileupHist.root.
```

Thus, the evolution of the pileup events, calculated for the various json files presented in Tab. A.1 is shown in Fig. A.1, where the mean number of the interactions per bunch crossing increased from 17 to 22.

Table A.1.: Json files with certified data and an integrated luminosity ($\mathcal{L}_{\text{int}}$)

JSON file (golden):	$\mathcal{L}_{\text{int}}, \text{pb}^{-1}$
Cert_271036-273730_13TeV_PromptReco_Collisions16_JSON.txt	627.623
Cert_271036-274443_13TeV_PromptReco_Collisions16_JSON.txt	2837
Cert_271036-275783_13TeV_PromptReco_Collisions16_JSON.txt	6320.99
Cert_271036-277933_13TeV_PromptReco_Collisions16_JSON.txt	16955.7
Cert_271036-284044_13TeV_23Sep2016ReReco_Collisions16_JSON.txt	35861.7

The MC samples include effects pileup events and are calculated for the PUMoriond17 PU scenario by the command:

```
python generateMCPUHist.py  
PileupCFI=mix_2016_25ns_Moriond17MC_Poisson00TPU_cfi  
output=PUMoriond17-nPUtrue.root.
```

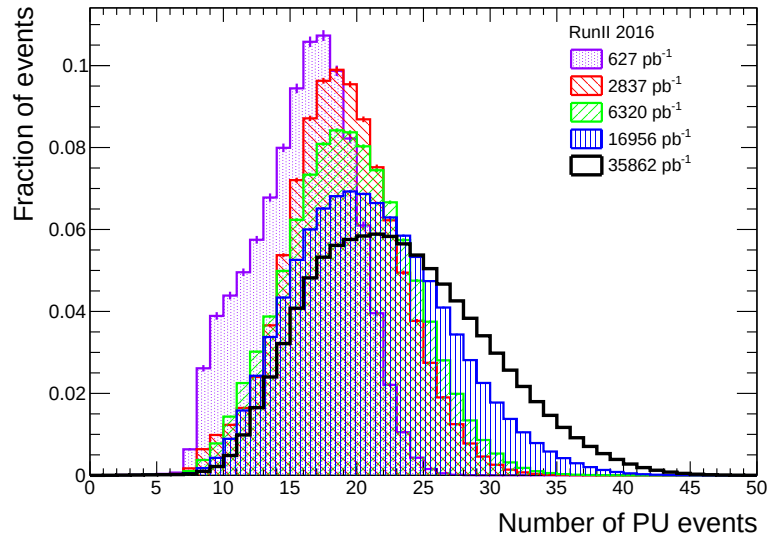


Figure A.1.: Number of expected pileup interactions in data collected in different run periods of Run 2. The histograms are normalised to 1.

Global tags

The following analysis performs the event selection and reconstruction using the CMS software, namely CMSSW framework, with corresponding version: CMSSW 8.0.26 patch1. The MiniAOD format to store the data was use as a new high-level data tier. For detector conditions and object calibrations, the global tags used in the analysis:

80X_dataRun2_2016SeptRepro_v7 (data: eras B-G),

80X_dataRun2_Prompt_v16 (data: era H) and

80X_mcRun2_asymptotic_2016_TracheIV_v6 (simulation).

The luminosity is calculated with brilcalc tool version 3.1.1:

```
brilcalc lumi -b "STABLE BEAMS"
```

```
--normtag /afs/cern.ch/user/l/lumipro/public/Normtags/normtag_DATACERT.json
```

```
-i lumi.json
```

A.2. L2 Residual jet energy correction

In this section supplementary material for the L2 residual jet energy correction at CMS with 2015 data collected at $\sqrt{s} = 13$ TeV is given.

Data sets for the time dependence study

Table A.2.: Run regions and corresponding luminosity used for time dependence study.

	RunNr	L (/pb)
1	254231-257613	210.282
2	257614-258136	232.832
3	258157-258177	237.130
4	258211-258448	229.872
5	258655-258741	214.421
6	258742-259685	225.970
7	259686-260373	219.400
8	260424-260534	266.200
9	260536-260627	258.137

Linear extrapolation $\alpha \rightarrow 0$ for the kFSR

The ratios of the MC to data response values obtained with MPF and dijet p_T balance methods as a function of α , e.i, relative transverse momentum of the third jet, are shown in Fig. A.2 and A.3. This distributions are obtained in several bins of $|\eta|$ of the probe jet and used for the linear extrapolation $\alpha \rightarrow 0$. The extrapolated value is than used in Eq. 7.15.

p_T extrapolations

In order to correct the result for the relative L2 residual corrections for the possible p_T dependence, the MC-to-data ratio of the response calculated with MPF and dijet p_T balance methods are parameterizing by a log-linear or constant functions as shown in Fig. A.4 and A.5.

The MPF response values in data and in MC are shown in Fig. A.6 as a function of a number of primary vertices.

Correction factors for the events passed HF triggers

The α and p_T extrapolation done for events that passed dijet HF HLT trigger path are shown in Fig. A.7 and Fig. A.8, respectively.

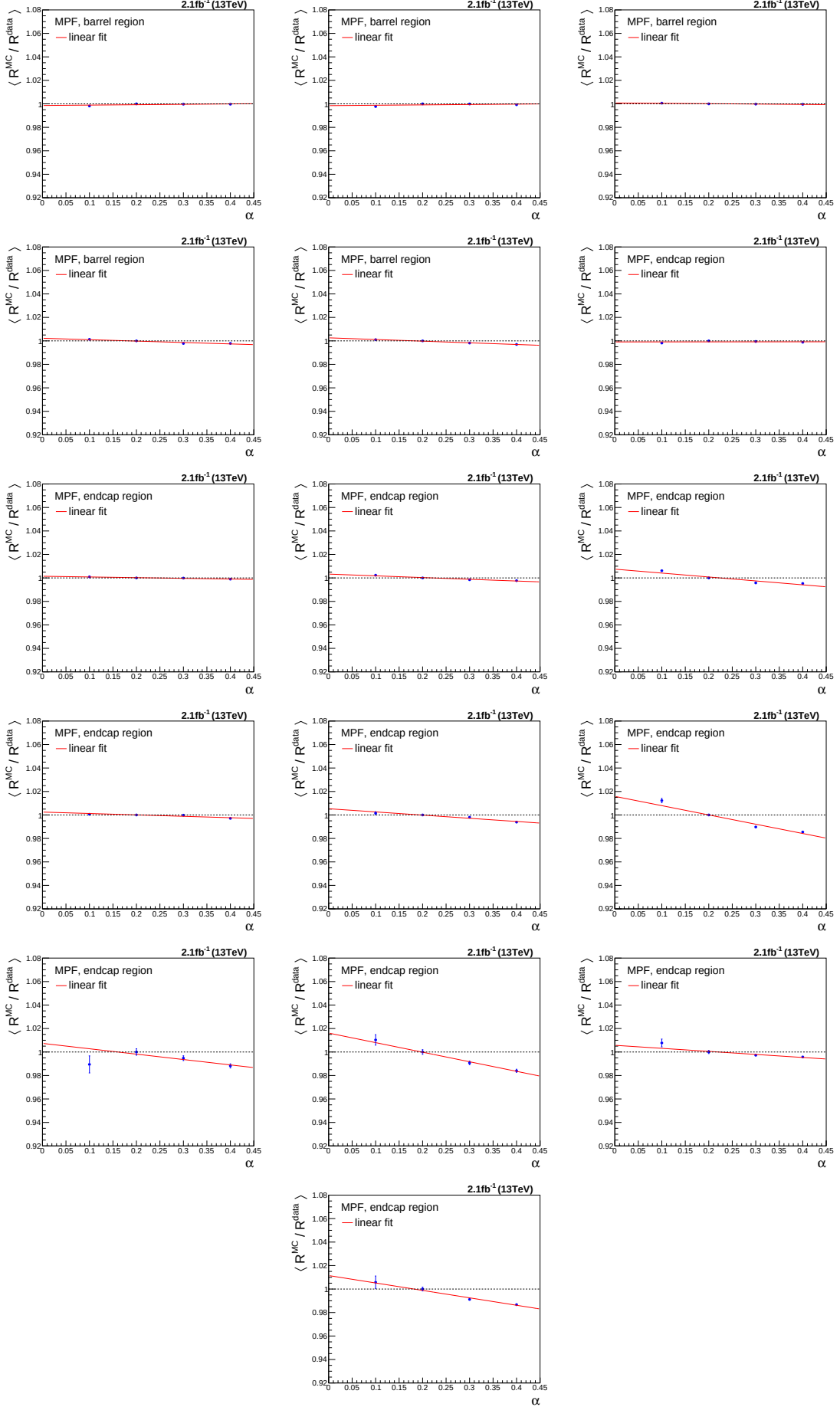


Figure A.2.: Linear kFSR extrapolations in bins of $|\eta|$ over α using the MPF method for the events that passed selection with nominal dijet HLT triggers.

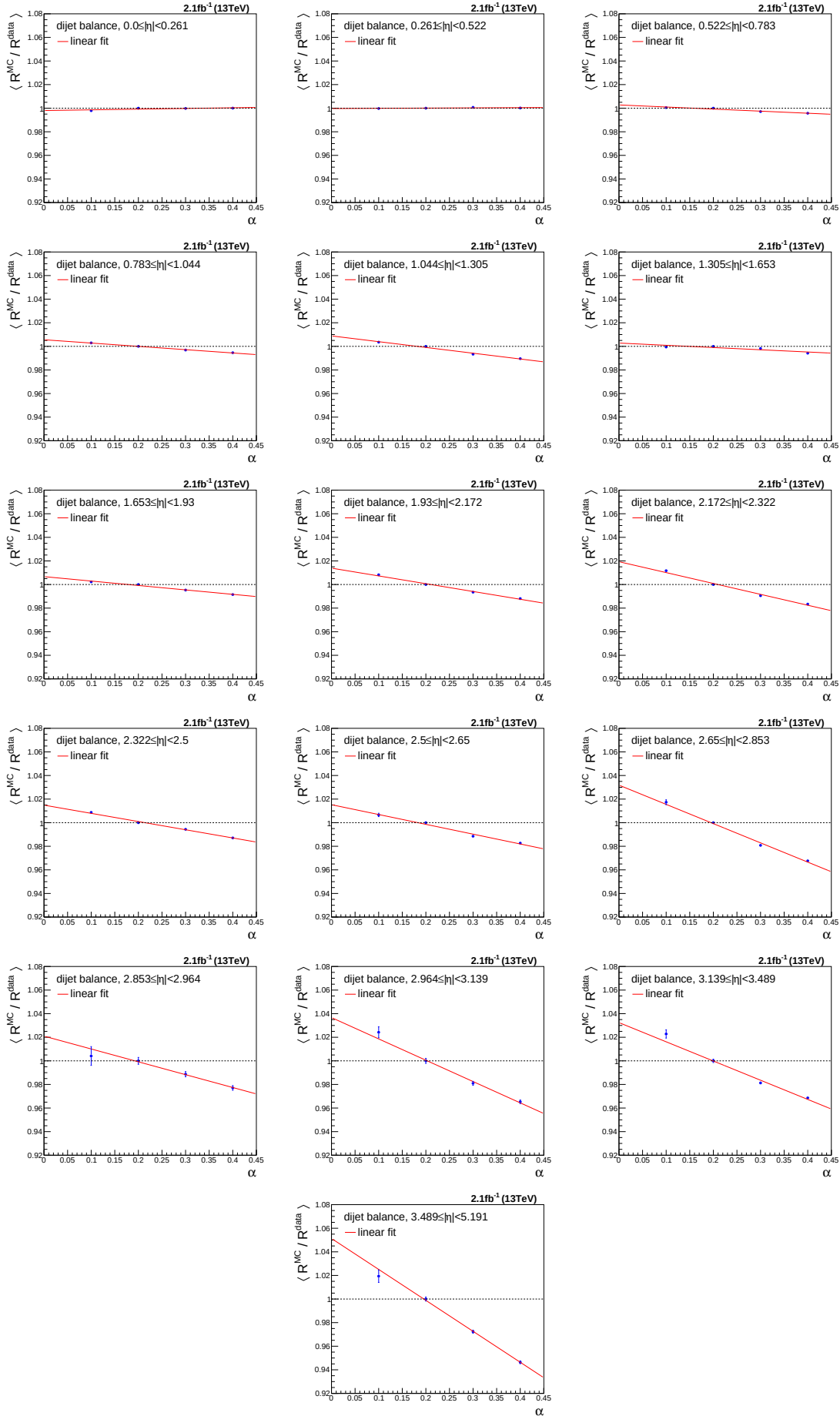


Figure A.3.: Linear kFSR extrapolations in bins of $|\eta|$ over α using the dijet p_T balance method for the events that passed selection with nominal dijet HLT triggers.

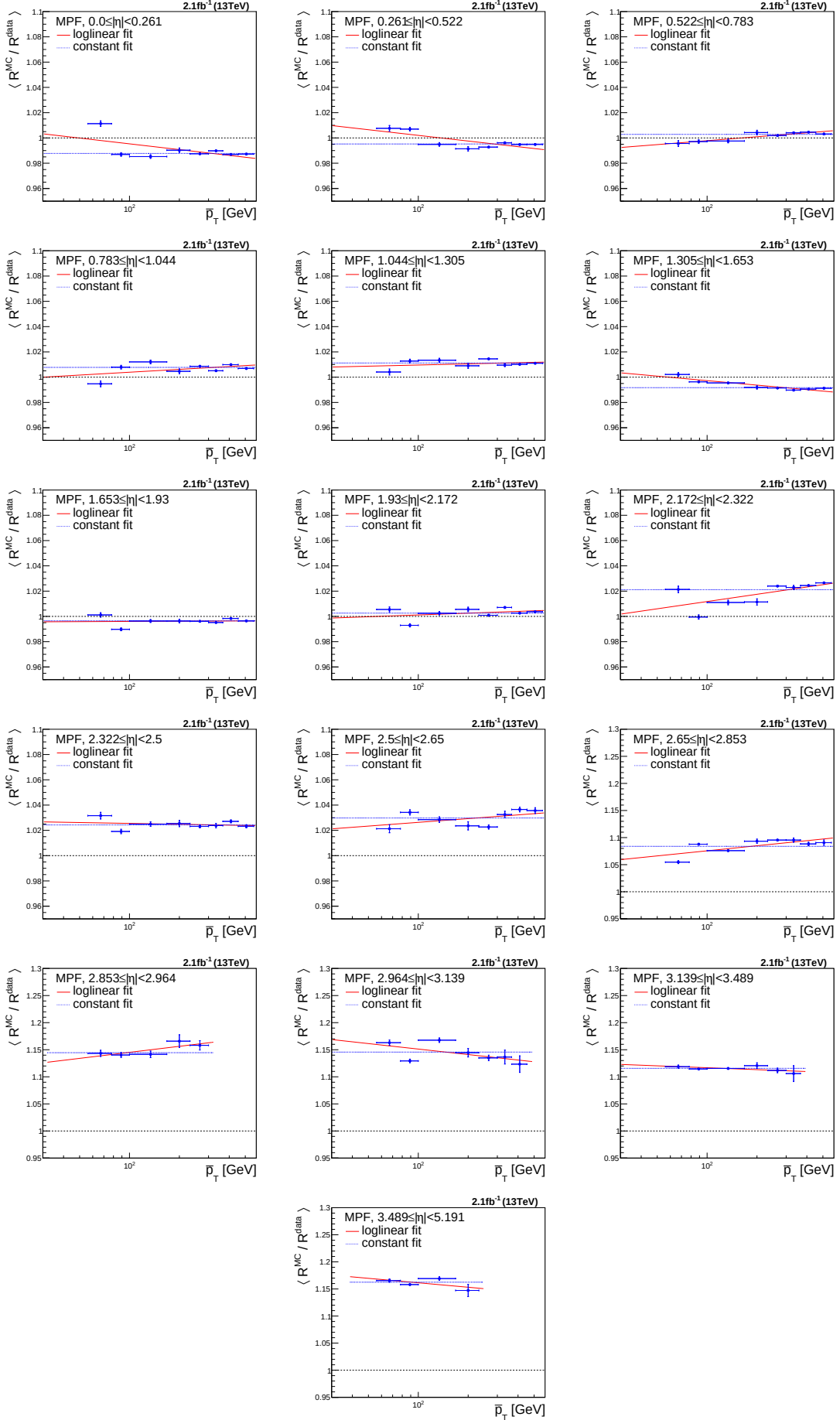


Figure A.4.: Linear and loglinear extrapolations in bins of $|\eta|$ over $\bar{p}_T$ using the MPF method for the events that passed selection with nominal dijet HLT triggers.

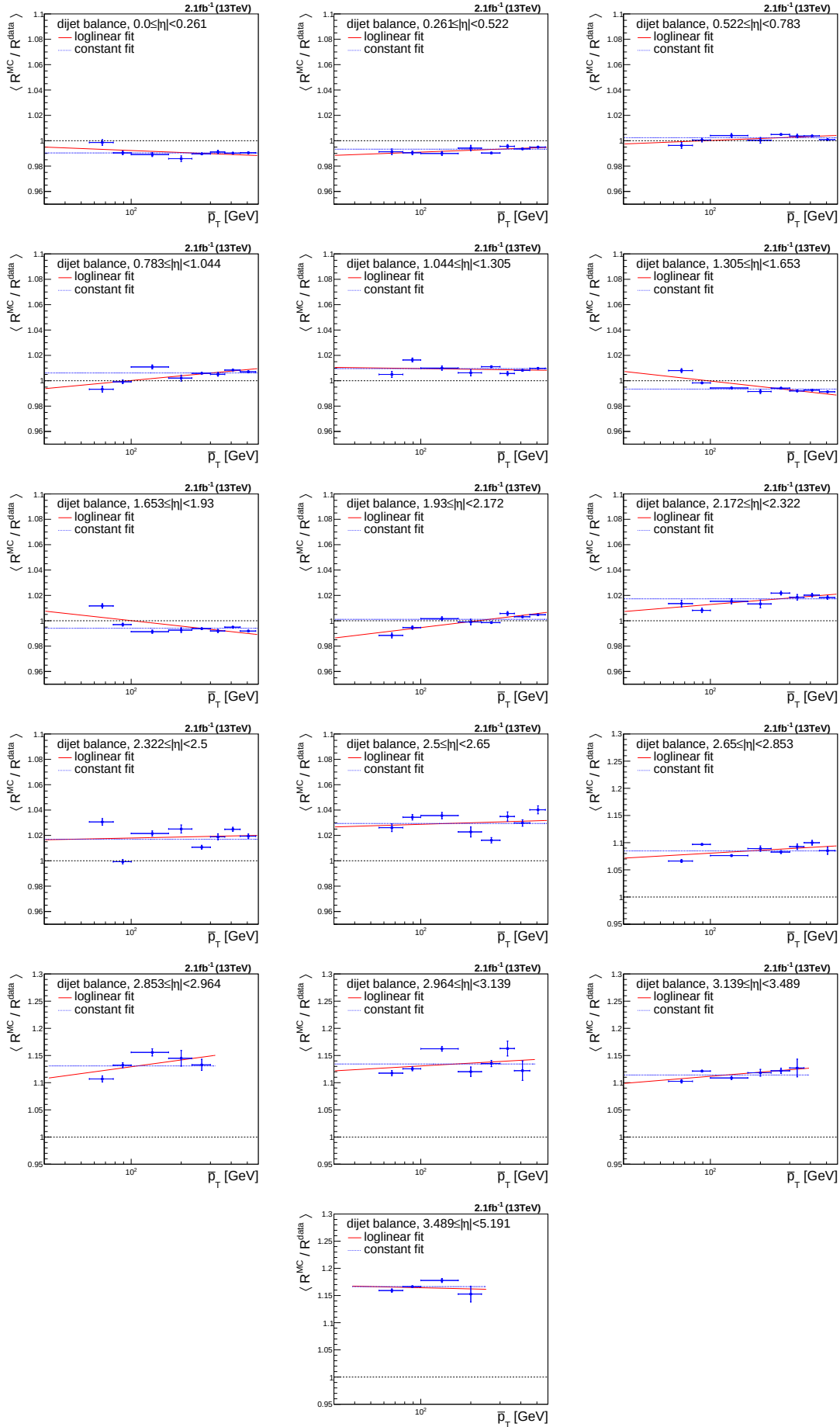


Figure A.5.: Linear and loglinear extrapolations in bins of $|\eta|$ over $\bar{p}_T$ using the dijet p_T balance method for the events that passed selection with nominal dijet HLT triggers.

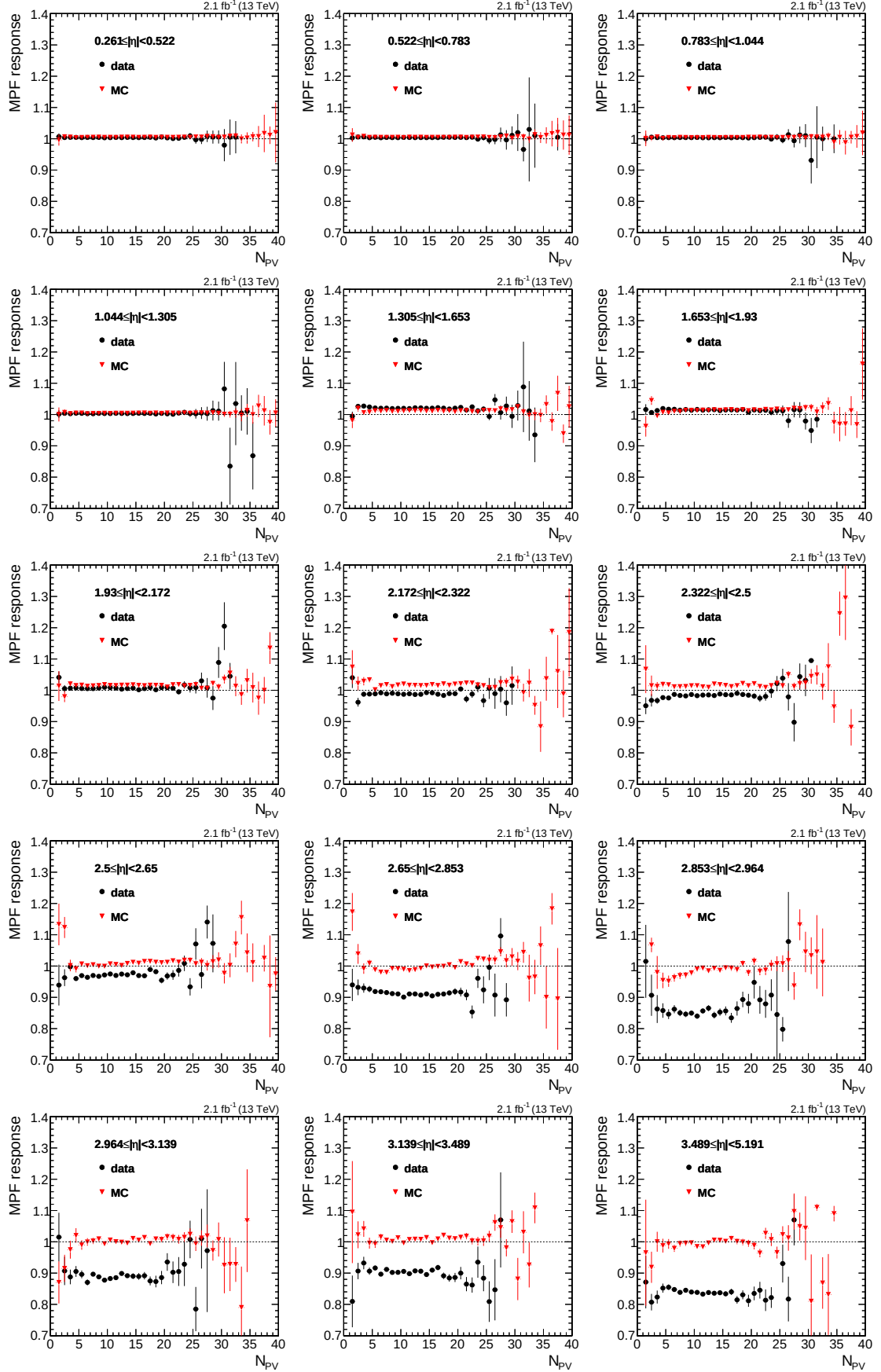


Figure A.6.: Relative response as a function of the number of primary vertices in bins of $|\eta|$ over $\bar{p}_T$ using the dijet p_T balance method for the events that passed selection with nominal dijet HLT triggers.

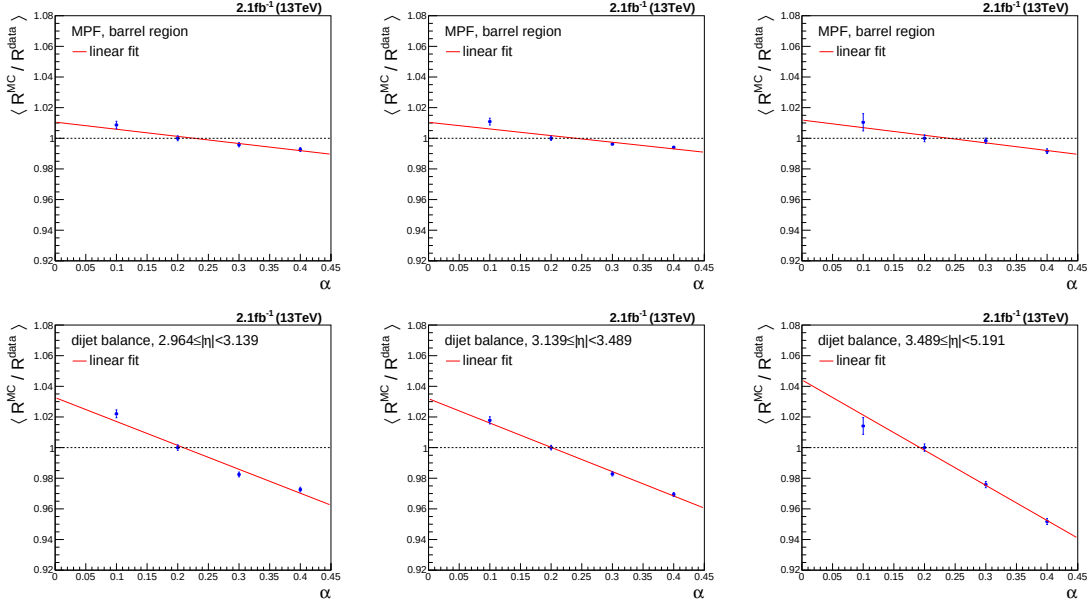


Figure A.7.: Linear kFSR extrapolations in bins of $|\eta|$ at forward region over α using the MPF method (upper row) and dijet p_T balance method (lower row) for the events that passed selection with dijet HF HLT triggers.

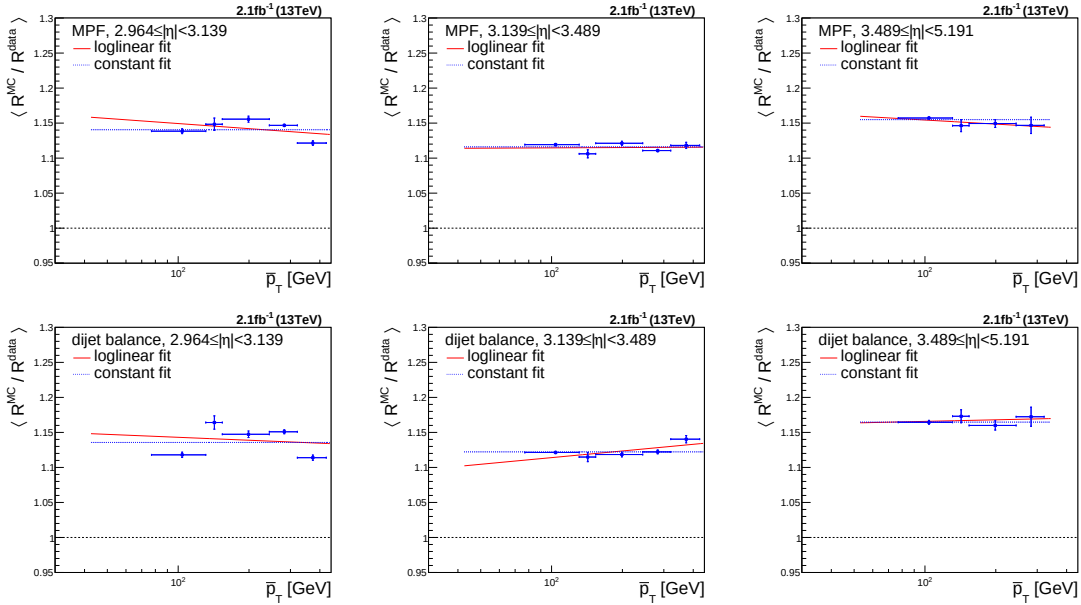


Figure A.8.: Linear and loglinear extrapolations in bins of $|\eta|$ over $\bar{p}_T$ using the MPF (upper row) and dijet p_T balance (lower row) method for the events that passed selection with nominal dijet HLT triggers.

A.3. Data-MC comparison for the top quark mass measurement

In this section the data-MC comparison between the variables reconstructed for muon+jets and electron+jets channels on each selection step is presented.

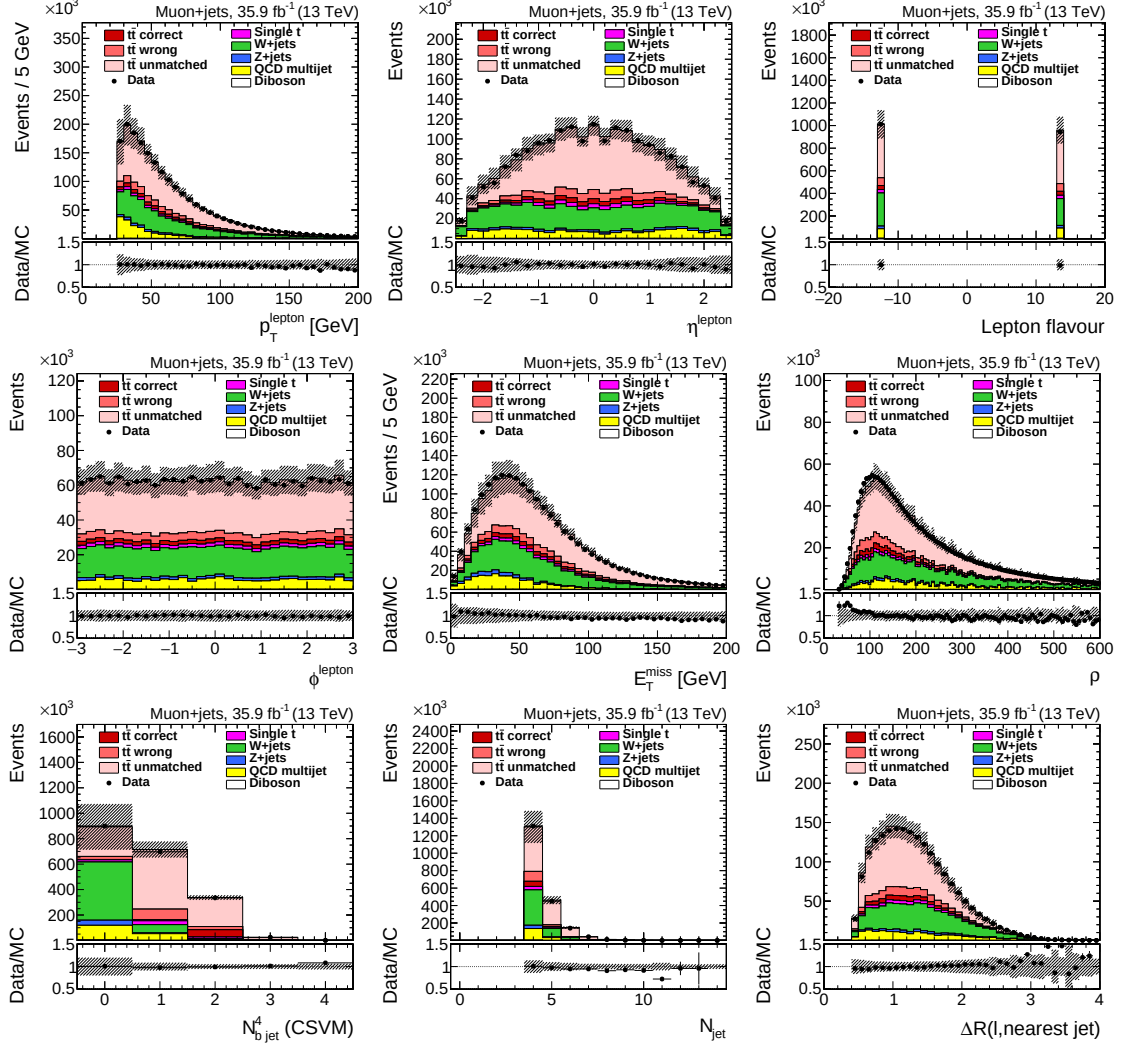


Figure A.9.: Preselection (Muon+jets): Reconstructed lepton p_T (left), η (middle) and lepton flavor $-13 = \mu^+$, $13 = \mu^-$ (right) (upper row). Reconstructed lepton ϕ (left), missing transverse momentum (middle) and energy density ρ (right) (middle row). Number of b-tags among four leading jets with $p_T > 30$ GeV (left), number of jets with $p_T > 30$ GeV (middle) and minimum distance between selected lepton and a jet with $p_T > 30$ GeV (right) (lower row). The plots are normalized to luminosity. The hatched area indicates uncertainties from MC statistics and cross-sections.

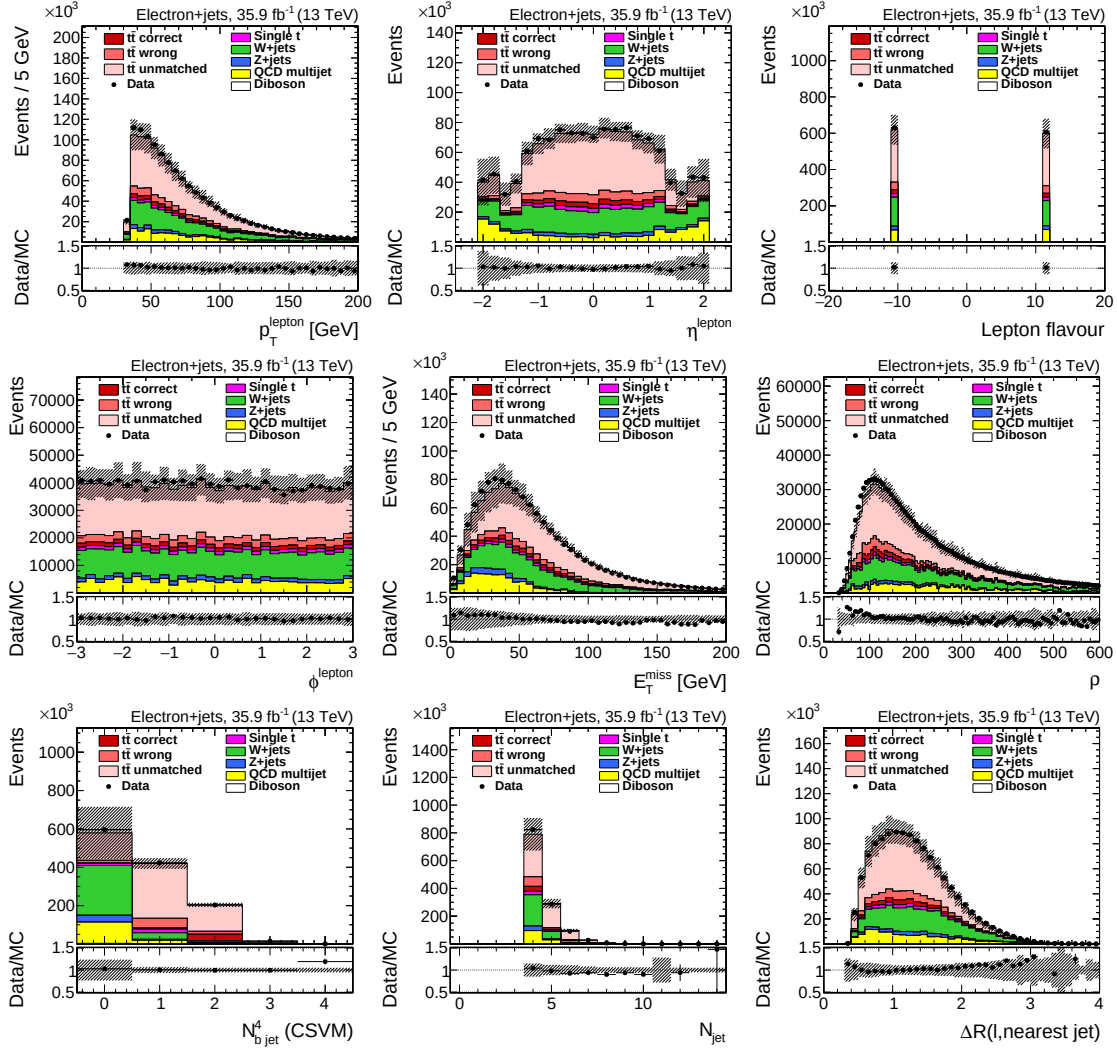


Figure A.10.: Preselection (Electron+jets): Reconstructed lepton p_T (left), η (middle) and lepton flavor $-13 = \mu^+$, $13 = \mu^-$ (right) (upper row). Reconstructed lepton ϕ (left), missing transverse momentum (middle) and energy density ρ (right) (middle row). Number of b-tags among four leading jets with $p_T > 30$ GeV (left), number of jets with $p_T > 30$ GeV (middle) and minimum distance between selected lepton and a jet with $p_T > 30$ GeV (right) (lower row). The plots are normalized to luminosity. The hatched area indicates uncertainties from MC statistics and cross-sections.

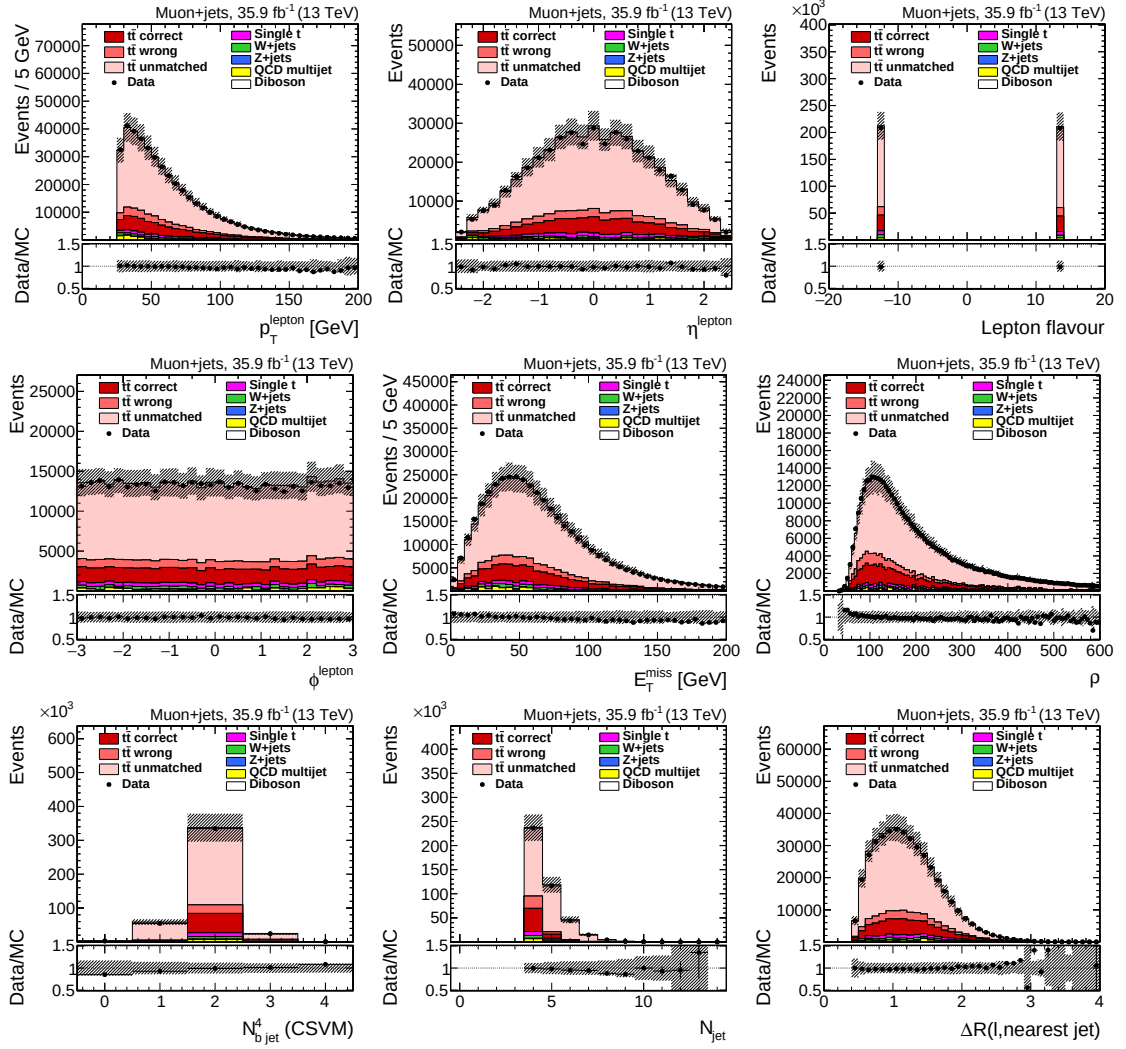


Figure A.11.: Baseline selection (Muon+jets): Reconstructed lepton p_T (left), η (middle) and lepton flavor $-13 = \mu^+$, $13 = \mu^-$ (right) (upper row). Reconstructed lepton ϕ (left), missing transverse momentum (middle) and energy density ρ (right) (middle row). Number of b-tags among four leading jets with $p_T > 30$ GeV (left), number of jets with $p_T > 30$ GeV (middle) and minimum distance between selected lepton and a jet with $p_T > 30$ GeV (right) (lower row). The plots are normalized to luminosity. The hatched area indicates uncertainties from MC statistics and cross-sections.

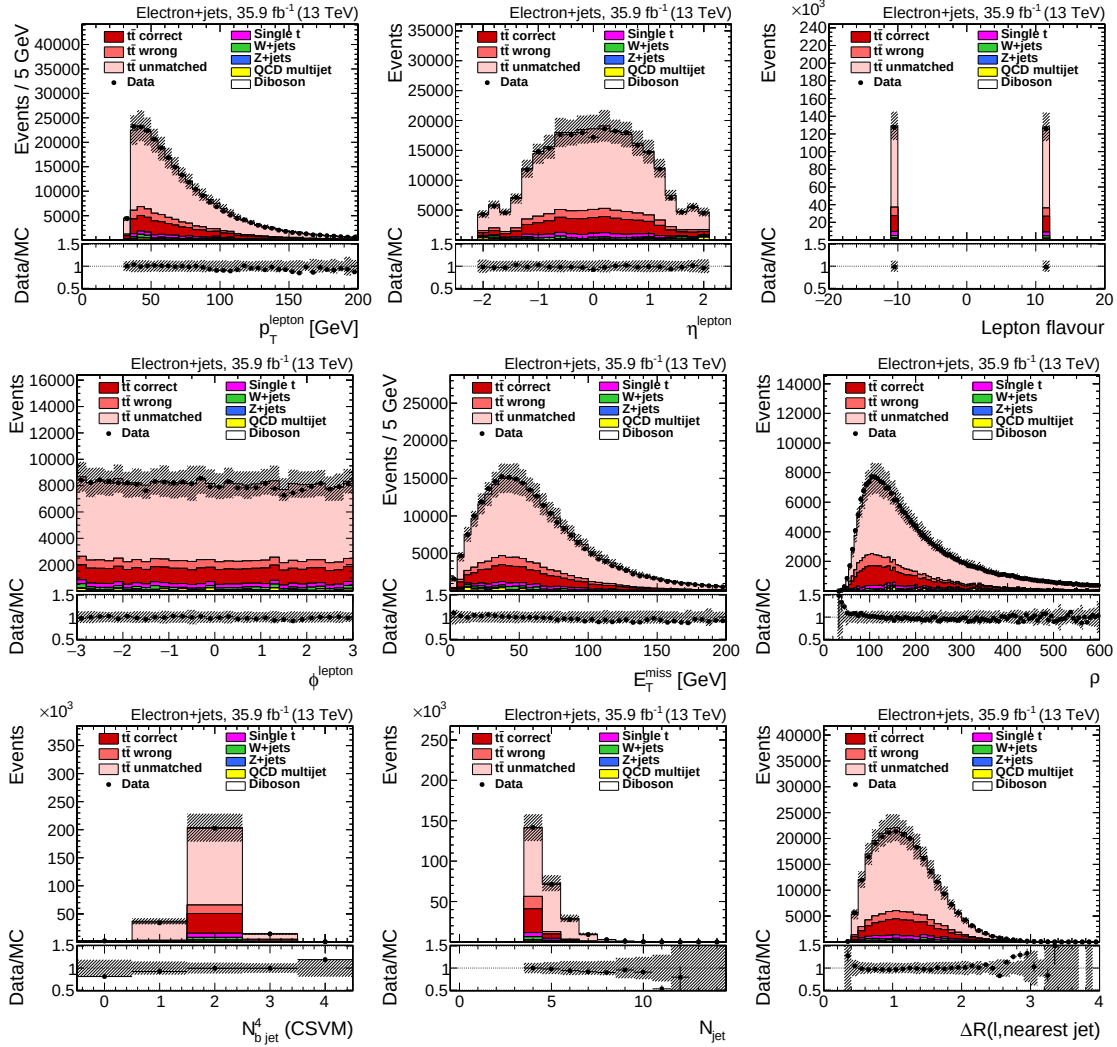


Figure A.12.: Baseline selection (Electron+jets): Reconstructed lepton p_T (left), η (middle) and lepton flavor $-13 = \mu^+$, $13 = \mu^-$ (right) (upper row). Reconstructed lepton ϕ (left), missing transverse momentum (middle) and energy density ρ (right) (middle row). Number of b-tags among four leading jets with $p_T > 30$ GeV (left), number of jets with $p_T > 30$ GeV (middle) and minimum distance between selected lepton and a jet with $p_T > 30$ GeV (right) (lower row). The plots are normalized to luminosity. The hatched area indicates uncertainties from MC statistics and cross-sections.

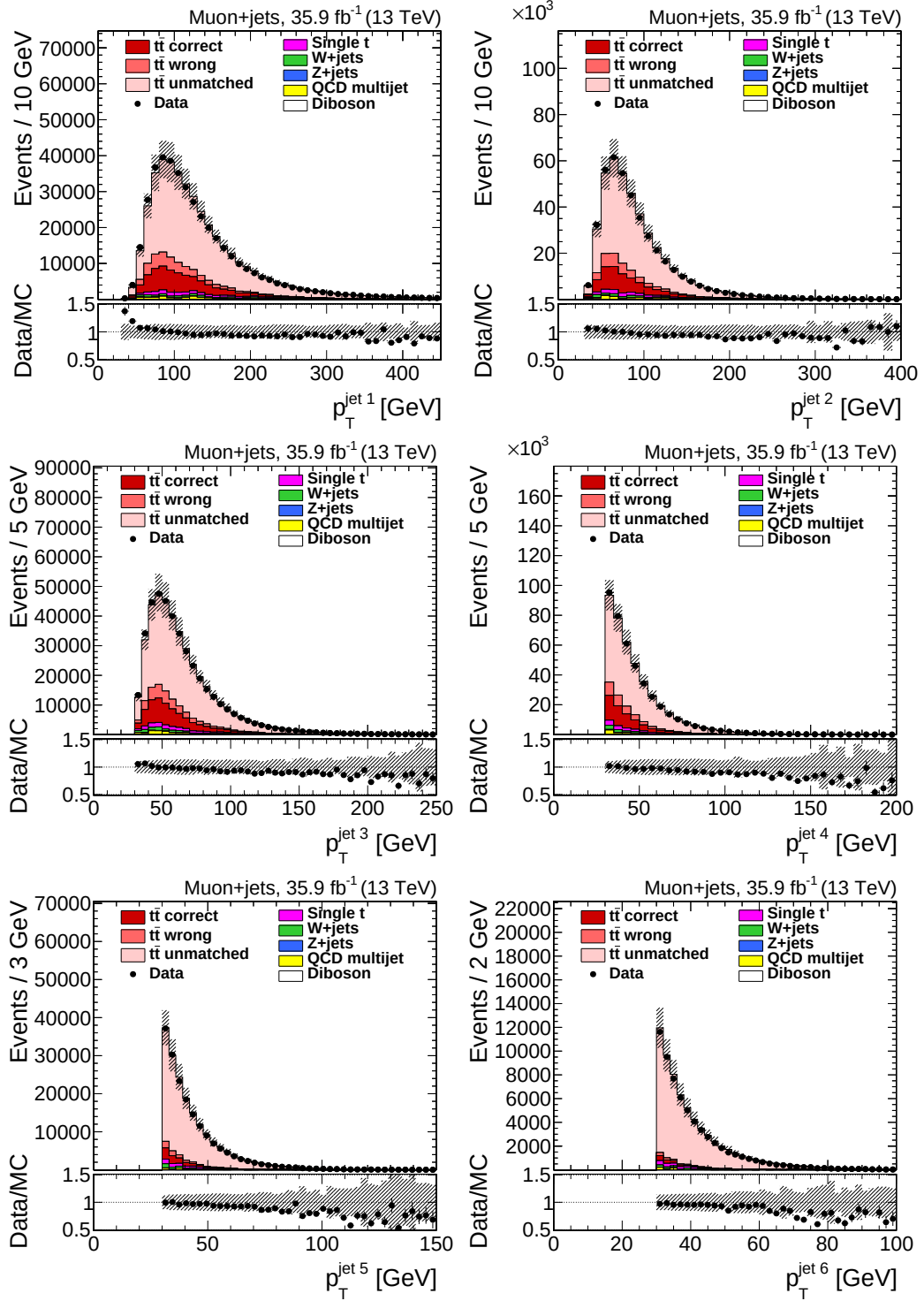


Figure A.13.: Baseline selection (Muon+jets): Reconstructed transverse momenta of the six leading jets.

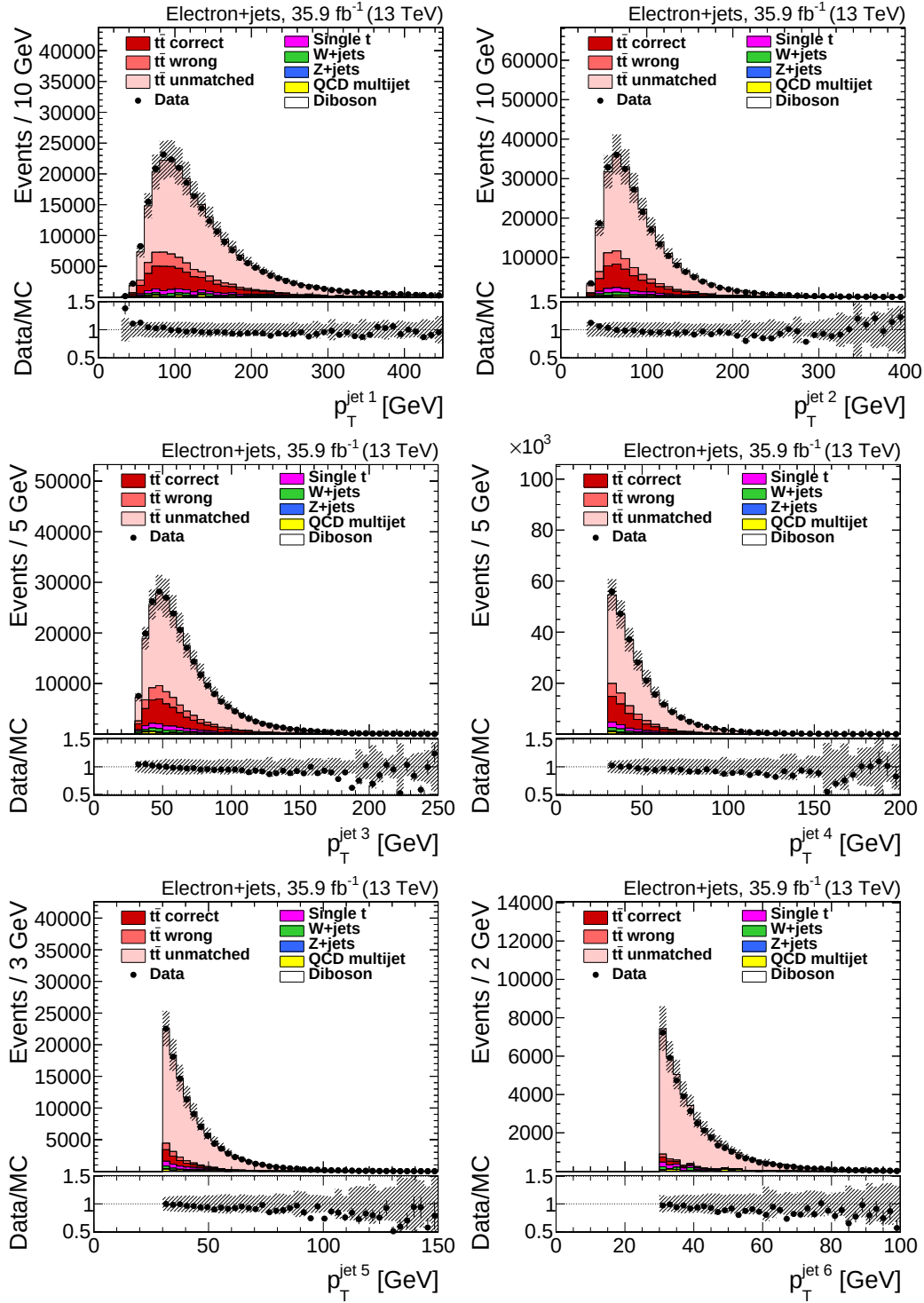


Figure A.14.: Baseline selection (Electron+jets): Reconstructed transverse momenta of the six leading jets.

A.4. Permutation yields

Table A.3.: (Muon+jets): Permutation fractions for different signal classes.

	baseline	final $P_{gof} > 0.2$	effective $\times P_{gof}$
$t\bar{t}$ correct	14.8 %	45.1 %	47.3 %
$t\bar{t}$ wrong	8.1 %	15.7 %	16.6 %
$t\bar{t}$ unmatched	77.1 %	39.3 %	36.2 %

Table A.4.: (Electron+jets): Permutation fractions for different signal classes.

	baseline	final $P_{gof} > 0.2$	effective $\times P_{gof}$
$t\bar{t}$ correct	15.0 %	46.8 %	49.1 %
$t\bar{t}$ wrong	7.9 %	15.4 %	16.2 %
$t\bar{t}$ unmatched	77.1 %	37.7 %	34.7 %

A.5. Templates for the μ/e +jets channel

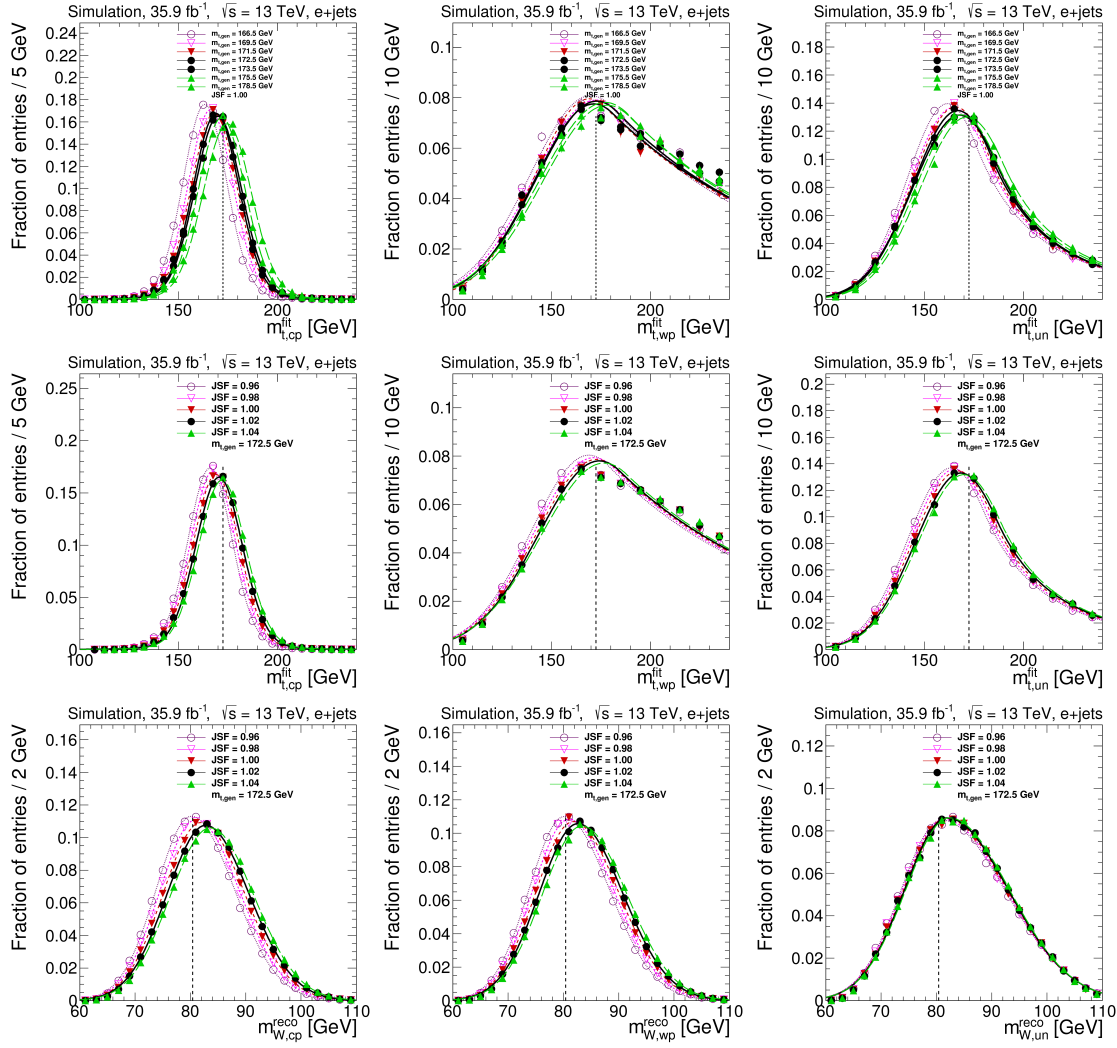


Figure A.15.: Simulated m_t^{fit} distributions of correct (left), wrong (middle), and unmatched (right) $t\bar{t}$ permutations, for seven generated masses $m_{t,\text{gen}}$ with $\text{JSF} = 1$ in the upper row and for five jet scale factors with $m_{t,\text{gen}} = 172.5 \text{ GeV}$ in the middle row for electron+jets events. The vertical dashed line corresponds to $m_t^{\text{fit}} = 172.5 \text{ GeV}$. In the last row the simulated m_W^{reco} distributions are shown for correct (left), wrong (middle), and unmatched (right) $t\bar{t}$ permutations for five jet scale factors with $m_{t,\text{gen}} = 172.5 \text{ GeV}$ for electron+jets events. The vertical dashed lines indicates the accepted value of the W-boson mass of 80.4 GeV.

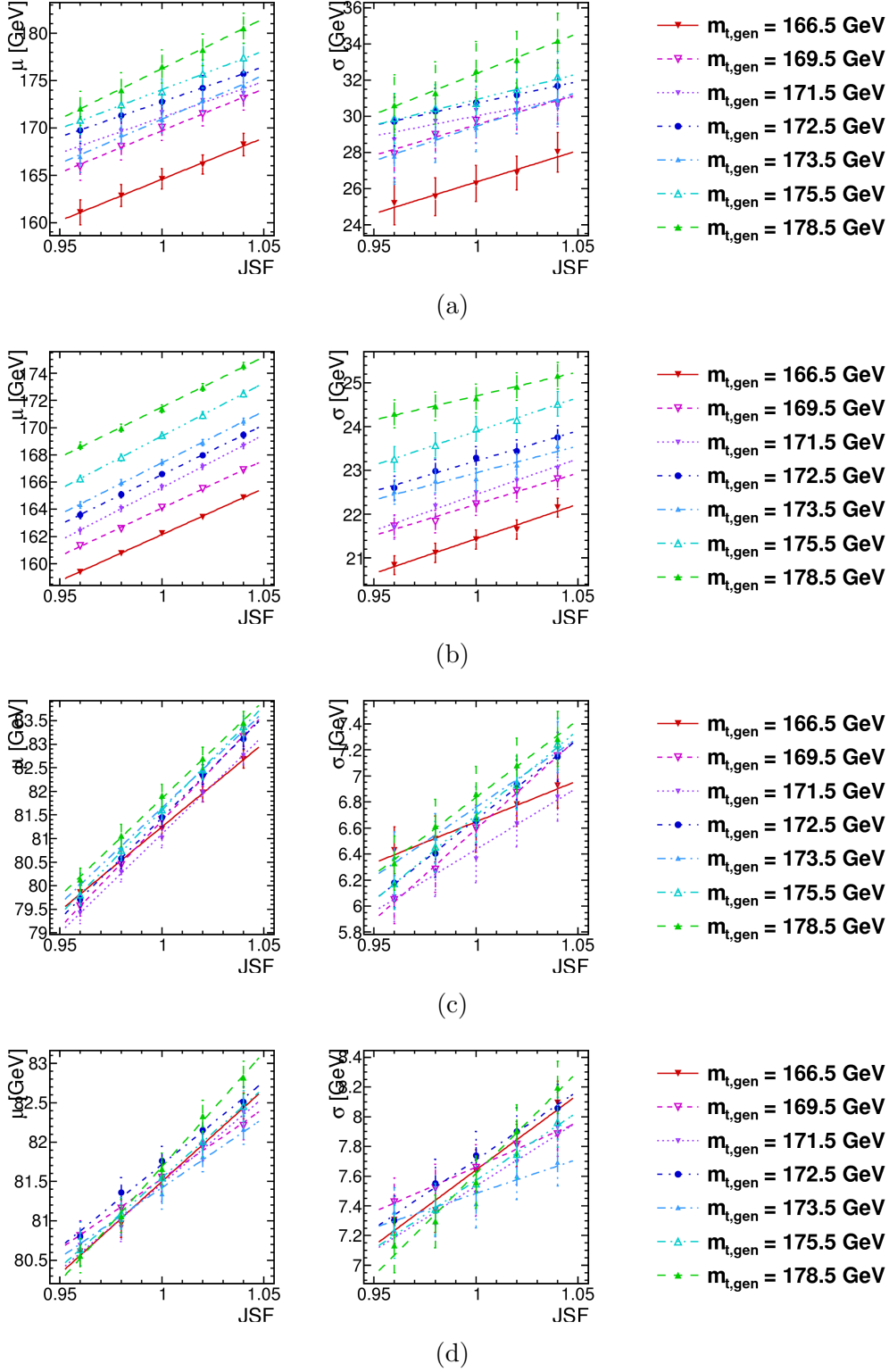


Figure A.16.: (Muon+jets) The Gaussian mean (left) and resolution (right) of m_t^{fit} distributions of wrong (a) and unmatched (b) $t\bar{t}$ permutations and of m_W^{reco} distributions of wrong (c) and unmatched (d) $t\bar{t}$ permutations, for seven generated masses $m_{t,\text{gen}}$ and five JSF.

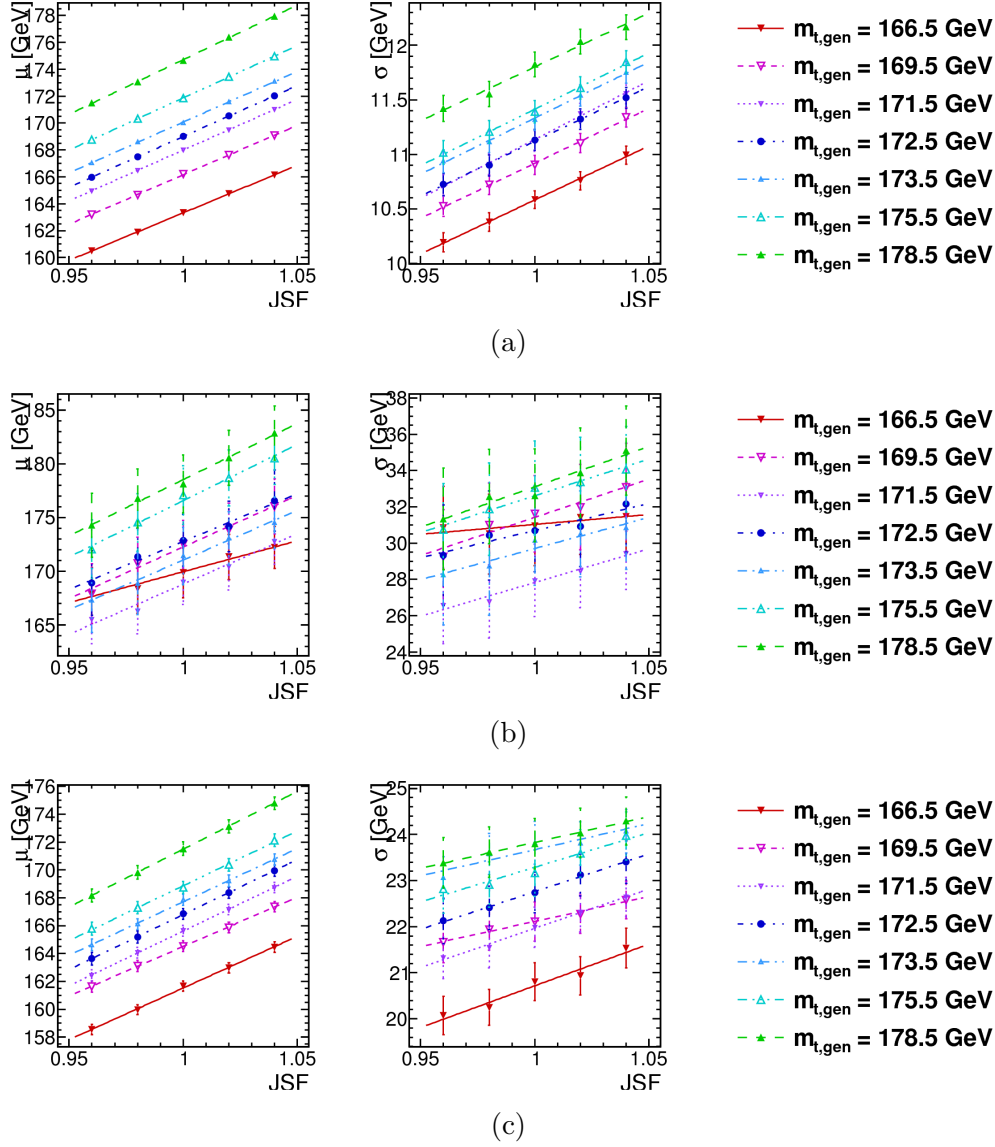


Figure A.17.: (Electron+jets) The Gaussian mean (left) and resolution (right) of the m_t^{fit} distributions of (a) correct, (b) wrong, and (c) unmatched $t\bar{t}$ permutations, for seven generated masses $m_{t,\text{gen}}$ and five JSF.

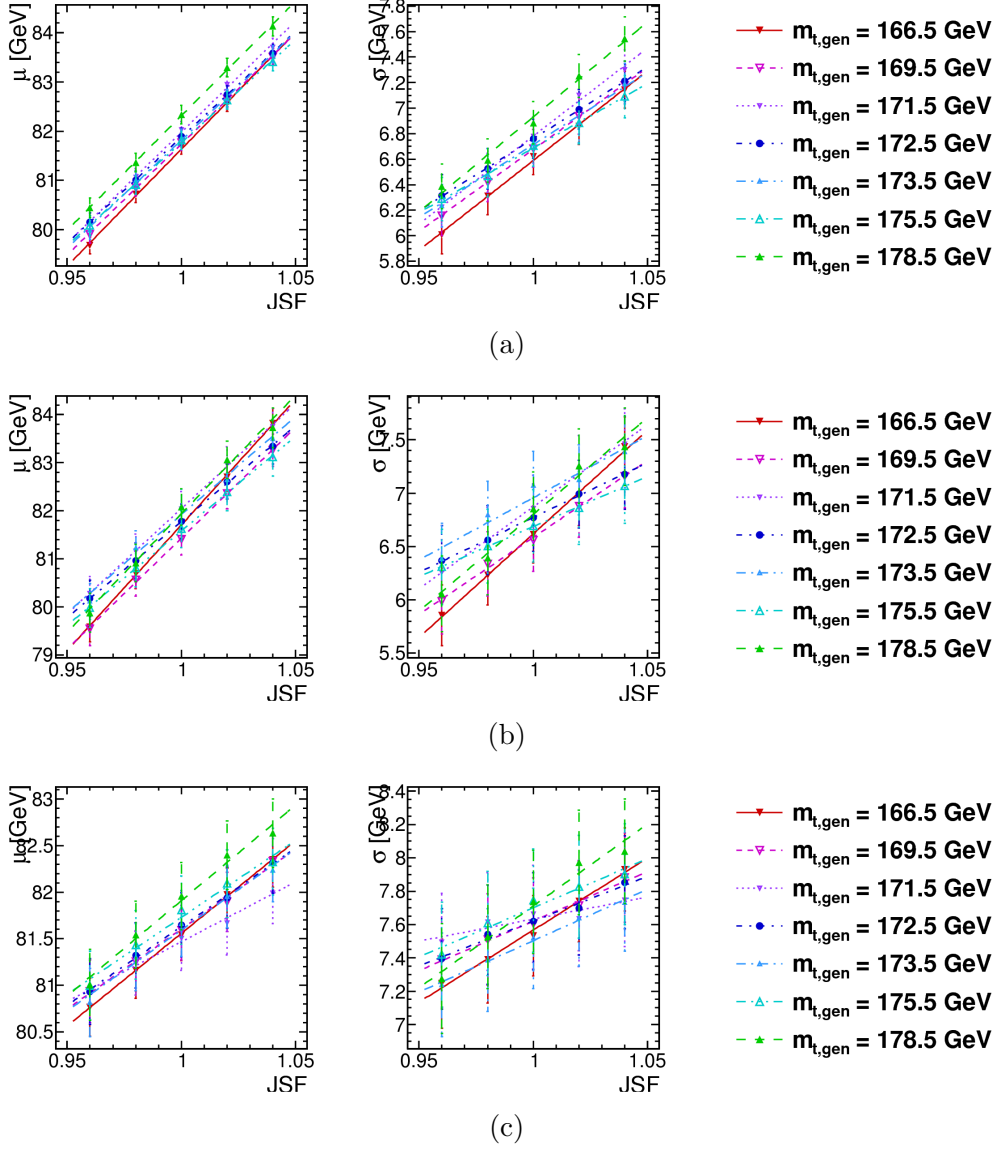


Figure A.18.: (Electron+jets) The Gaussian mean (left) and resolution (right) of the m_W^{reco} distributions of (a) correct, (b) wrong, and (c) unmatched $t\bar{t}$ permutations, for seven generated masses $m_{t,\text{gen}}$ and five JSF.

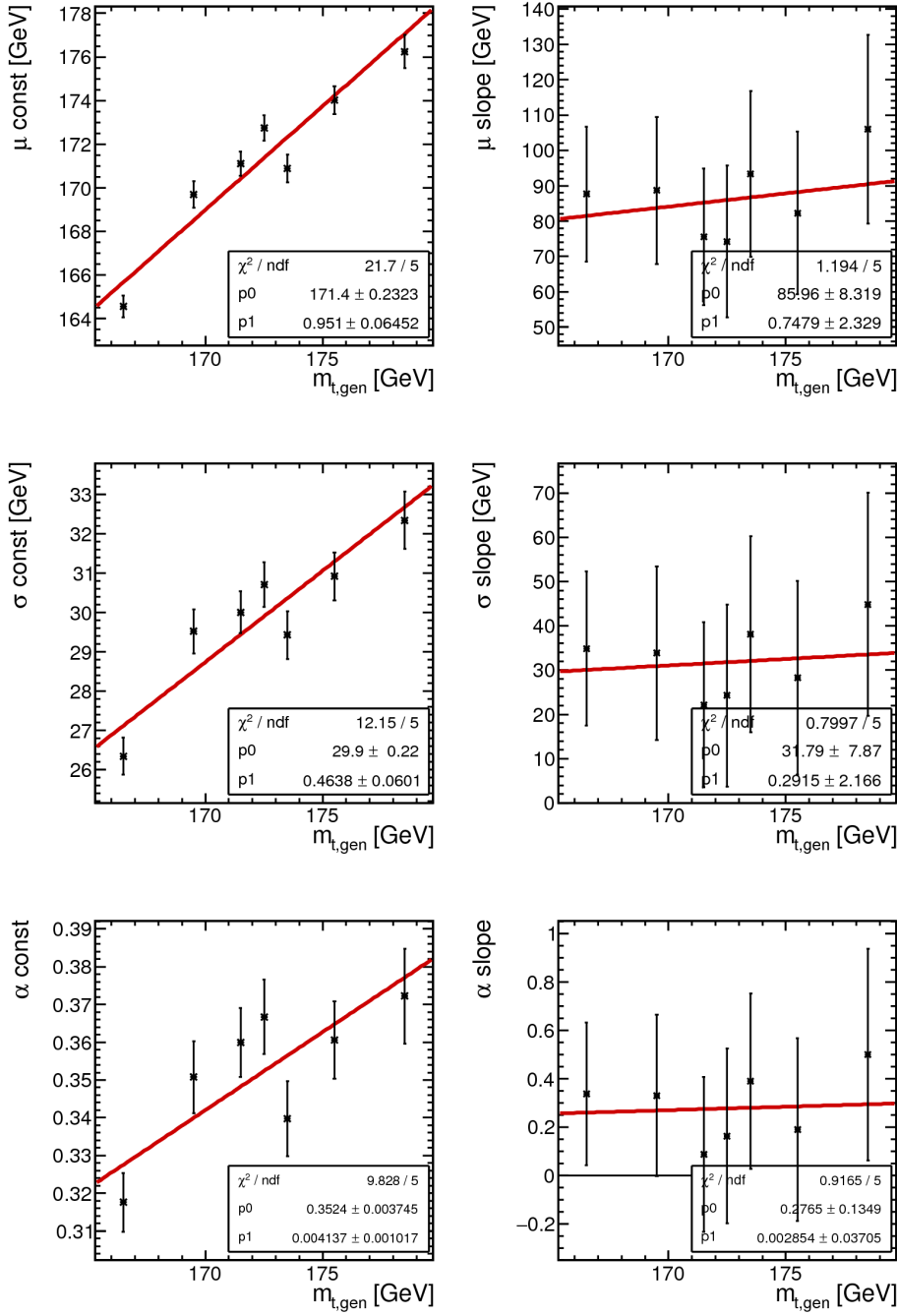


Figure A.19.: (Muon+jets) The Gaussian mean (upper left), resolution (middle left) and exponential threshold (lower left) and their dependency on JSF (upper/middle/lower right) of the m_t^{fit} distributions of wrong $t\bar{t}$ permutations, for seven generated masses $m_{t,gen}$.

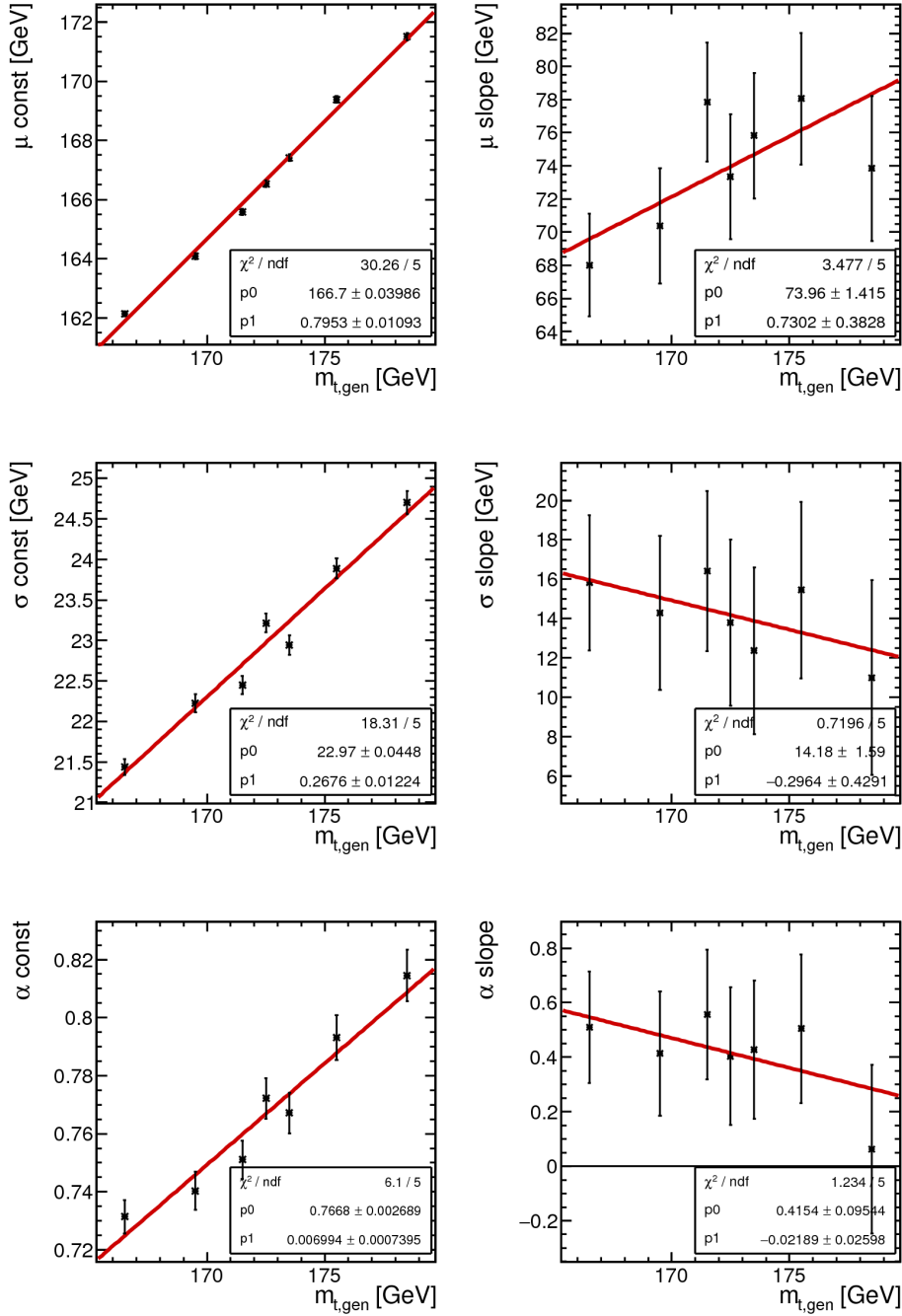


Figure A.20.: (Muon+jets) The Gaussian mean (upper left) and resolution (middle left) and exponential threshold (lower left) and their dependency on JSF (upper/middle/lower right) of the m_t^{fit} distributions of unmatched $t\bar{t}$ permutations, for seven generated masses $m_{t,gen}$.

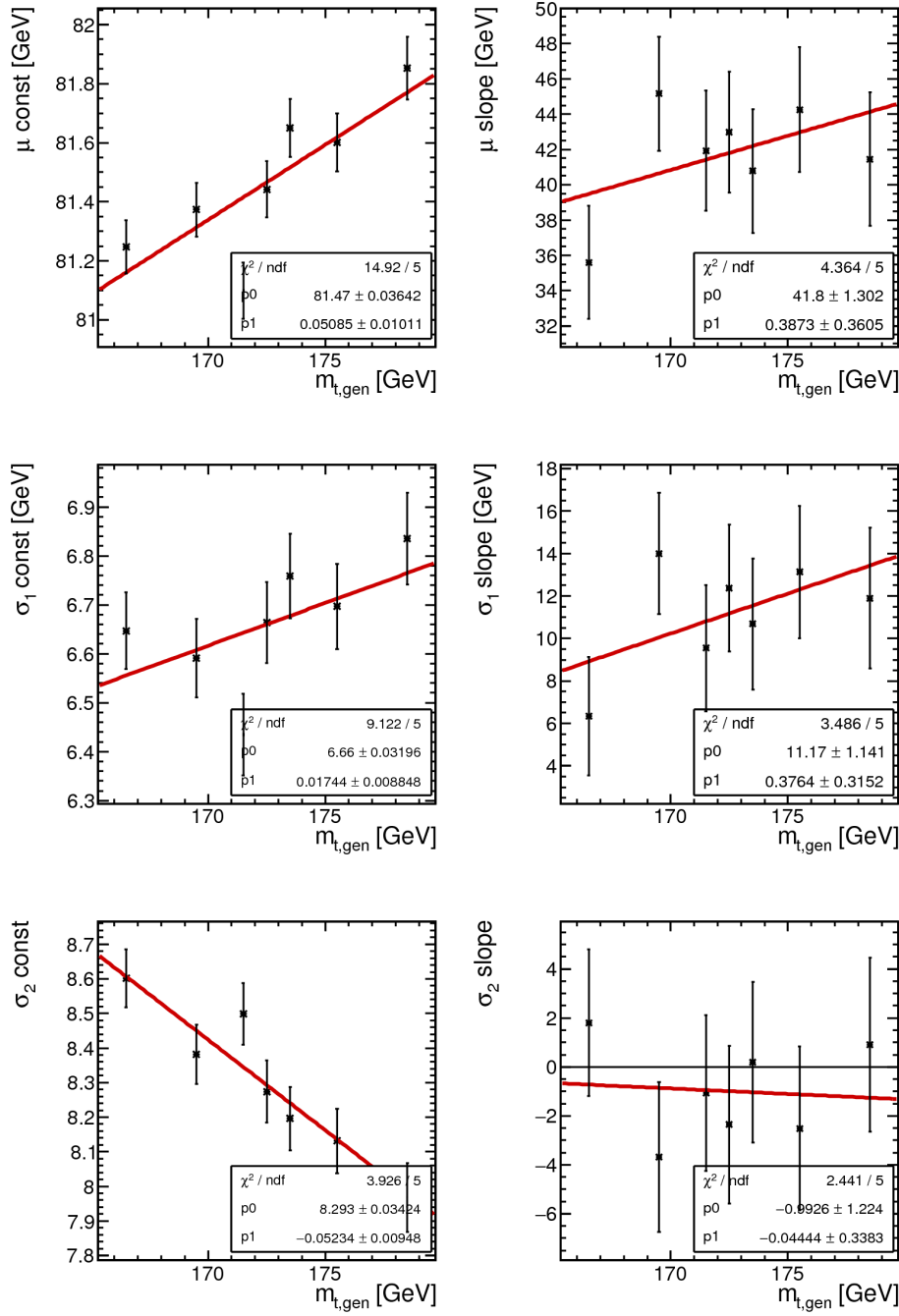


Figure A.21.: (Muon+jets) The Gaussian mean (upper left) and resolutions (middle/lower left) and their dependency on JSF (upper/middle/lower right) of the m_W^{reco} distributions of wrong $t\bar{t}$ permutations, for seven generated masses $m_{t,\text{gen}}$.

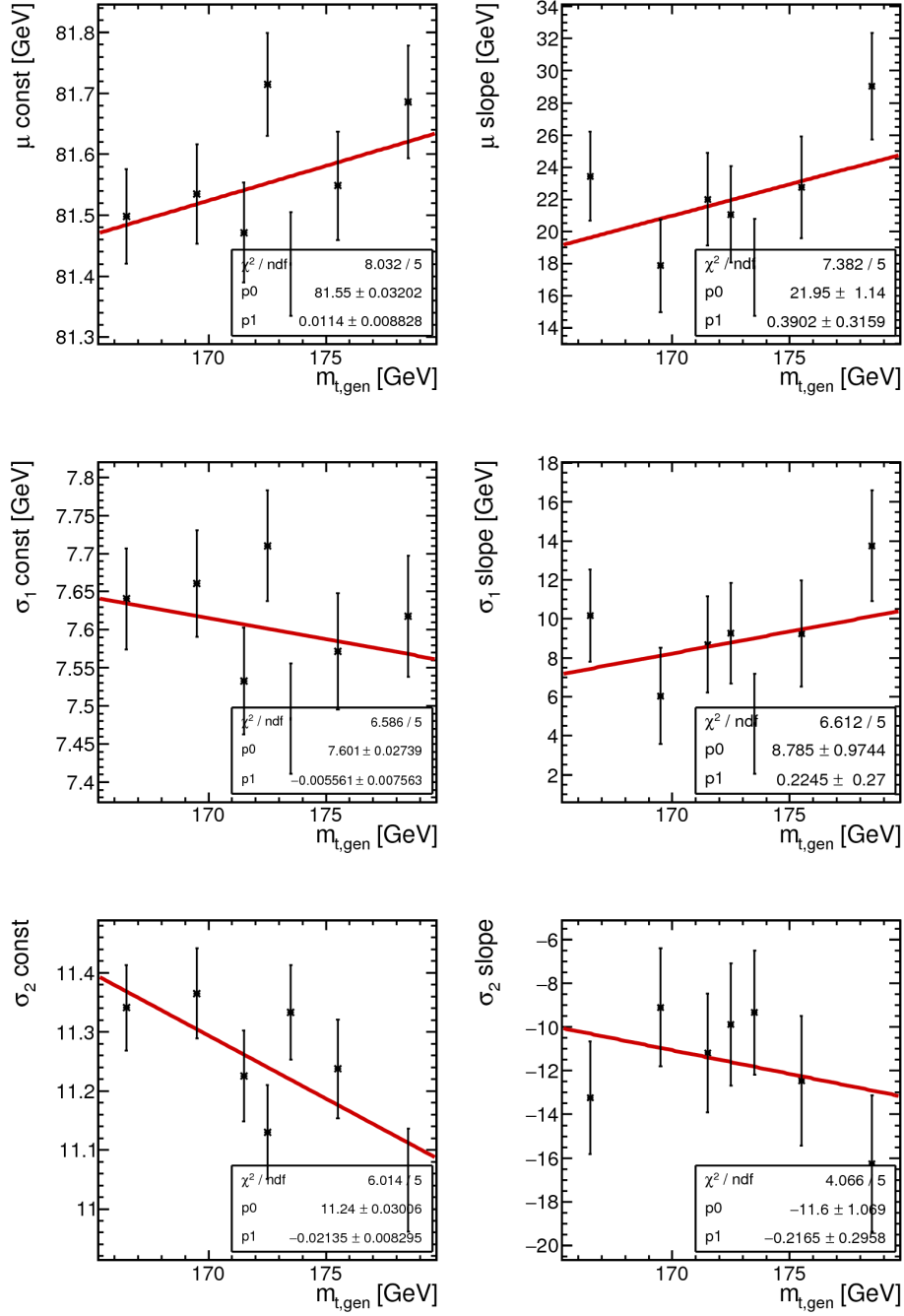


Figure A.22.: (Muon+jets) The Gaussian mean (upper left) and resolutions (middle/lower left) and their dependency on JSF (upper/middle/lower right) of the m_W^{reco} distributions of unmatched $t\bar{t}$ permutations, for seven generated masses $m_{t,\text{gen}}$.

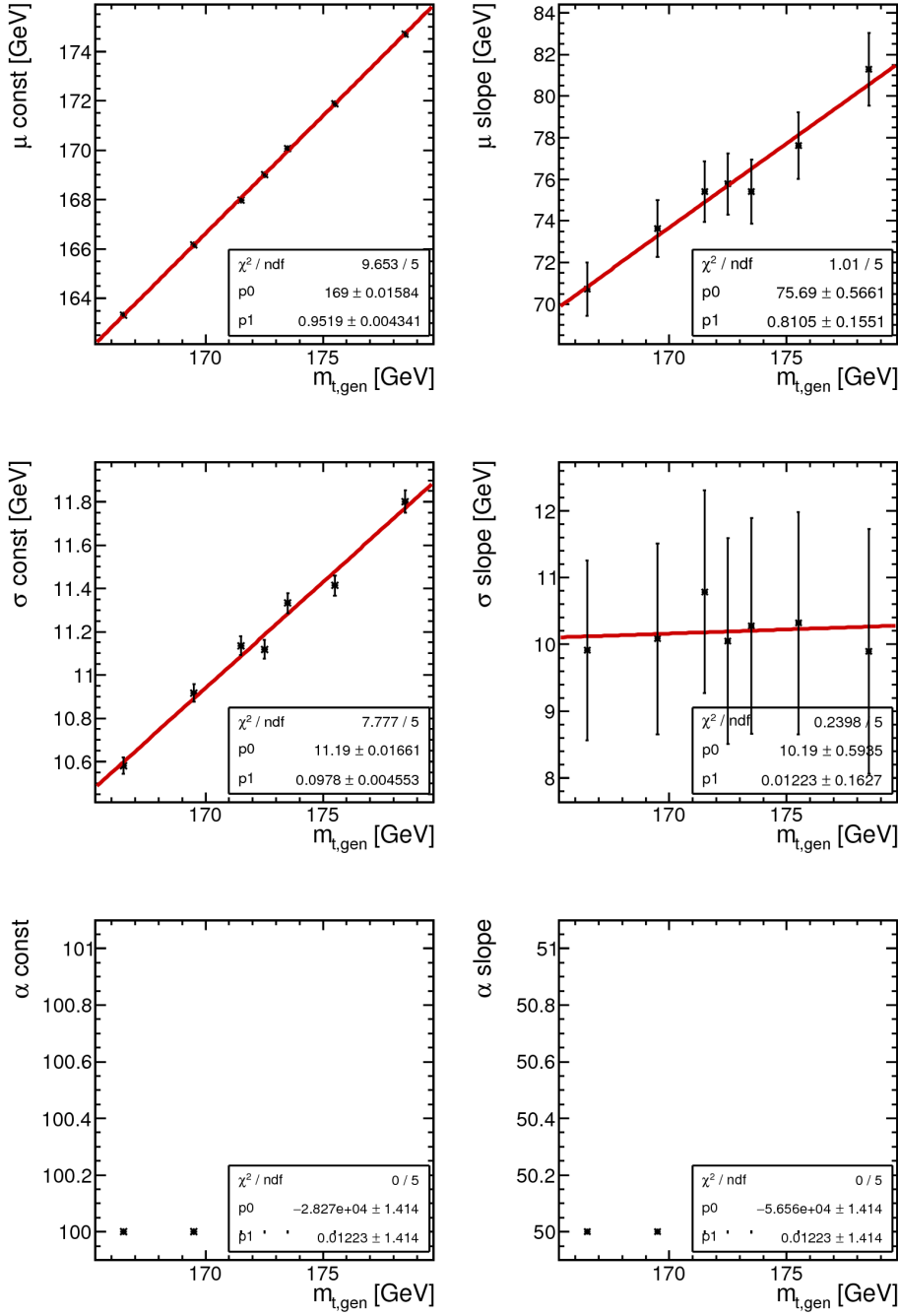


Figure A.23.: (Electron+jets) The Gaussian mean (upper left), resolution (middle left), width (fixed to 2 GeV) and their dependency on JSF (upper/middle right) of the m_t^{fit} distributions of correct $t\bar{t}$ permutations, for seven generated masses $m_{t,\text{gen}}$.

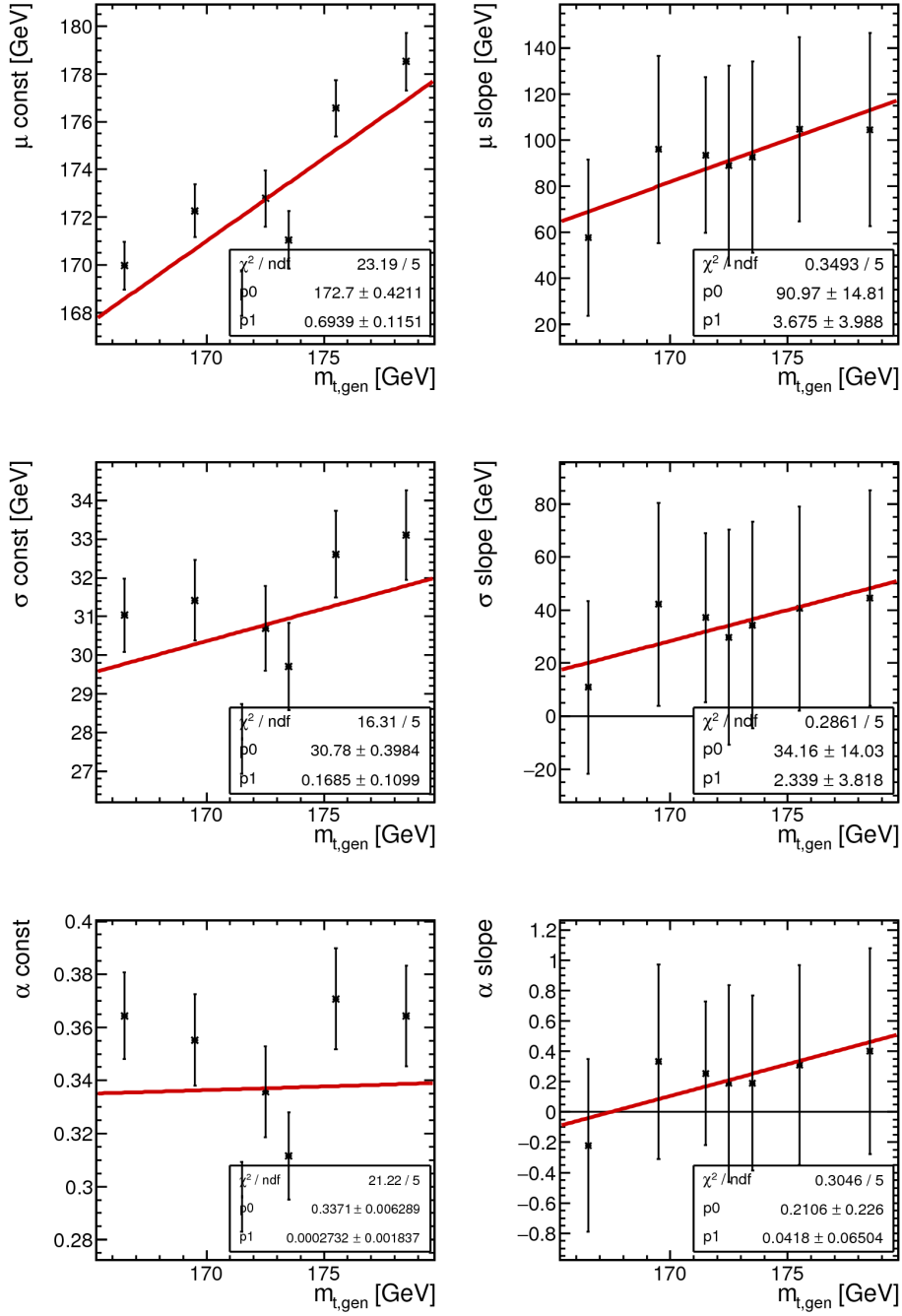


Figure A.24.: (Electron+jets) The Gaussian mean (upper left), resolution (middle left) and exponential threshold (lower left) and their dependency on JSF (upper/middle/lower right) of the m_t^{fit} distributions of wrong $t\bar{t}$ permutations, for seven generated masses $m_{t,gen}$.

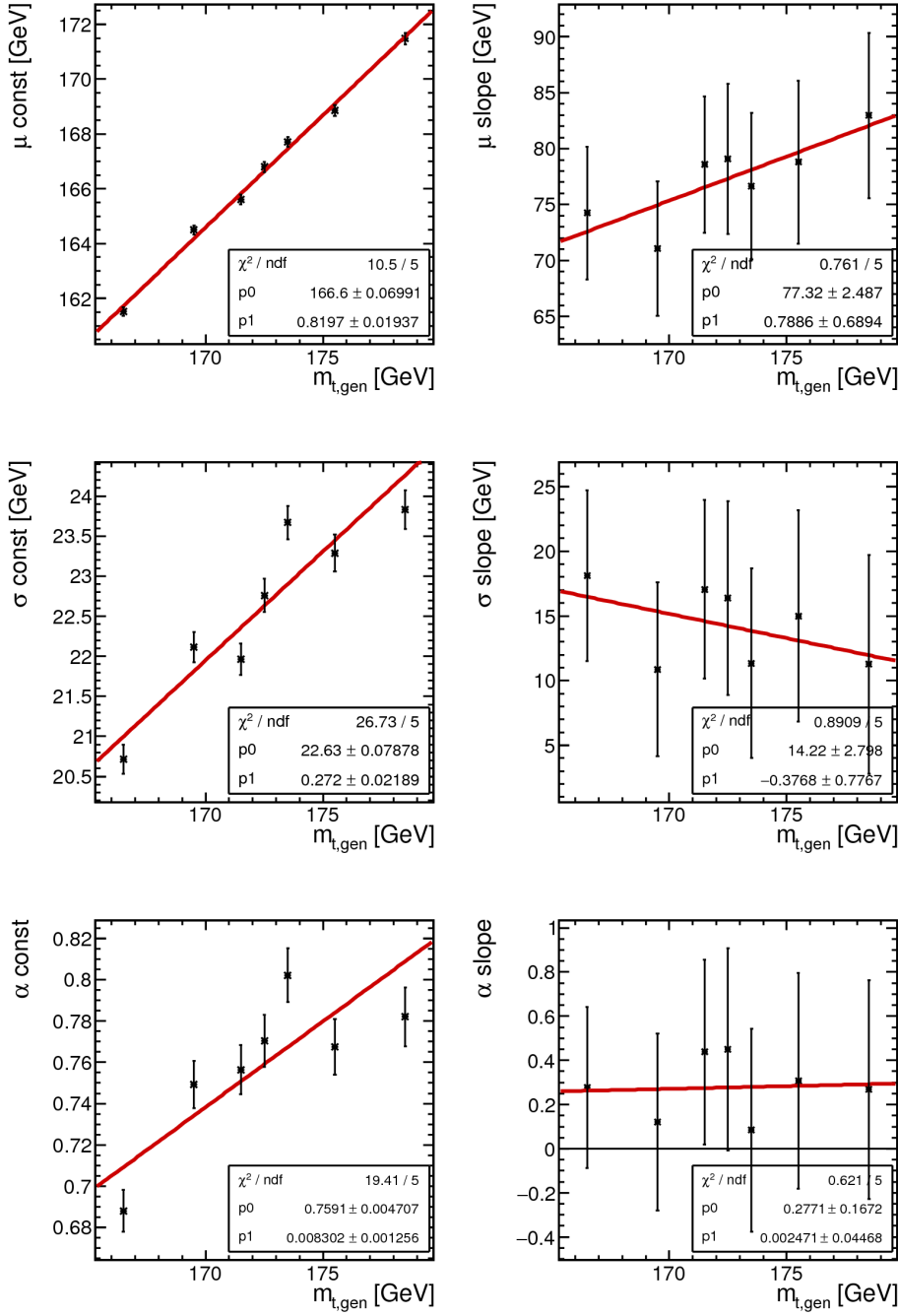


Figure A.25.: (Electron+jets) The Gaussian mean (upper left) and resolution (middle left) and exponential threshold (lower left) and their dependency on JSF (upper/middle/lower right) of the m_t^{fit} distributions of unmatched $t\bar{t}$ permutations, for seven generated masses $m_{t,\text{gen}}$.

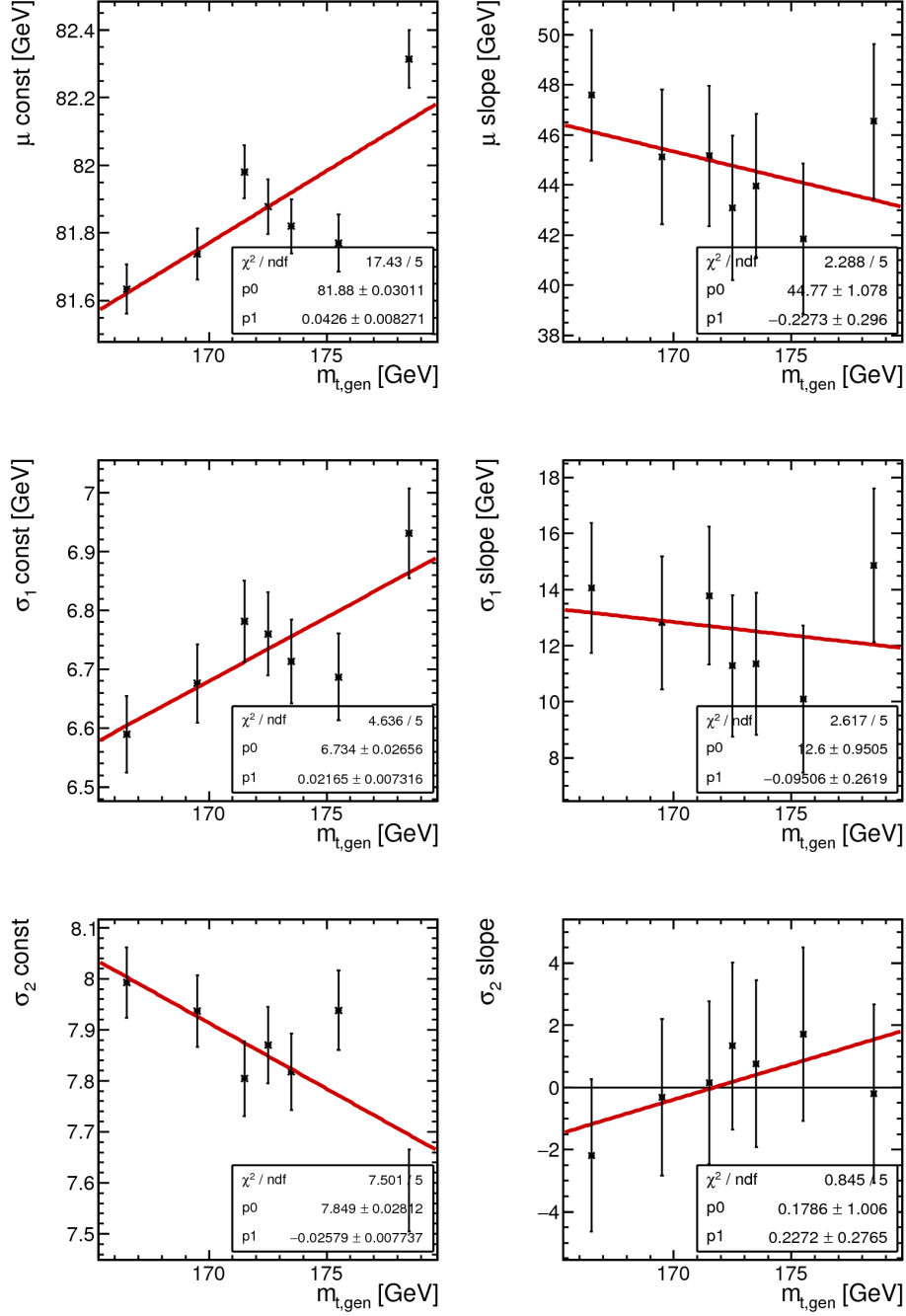


Figure A.26.: (Electron+jets) The Gaussian mean (upper left) and resolutions (middle/lower left) and their dependency on JSF (upper/middle/lower right) of the m_W^{reco} distributions of correct $t\bar{t}$ permutations, for seven generated masses $m_{t,\text{gen}}$.

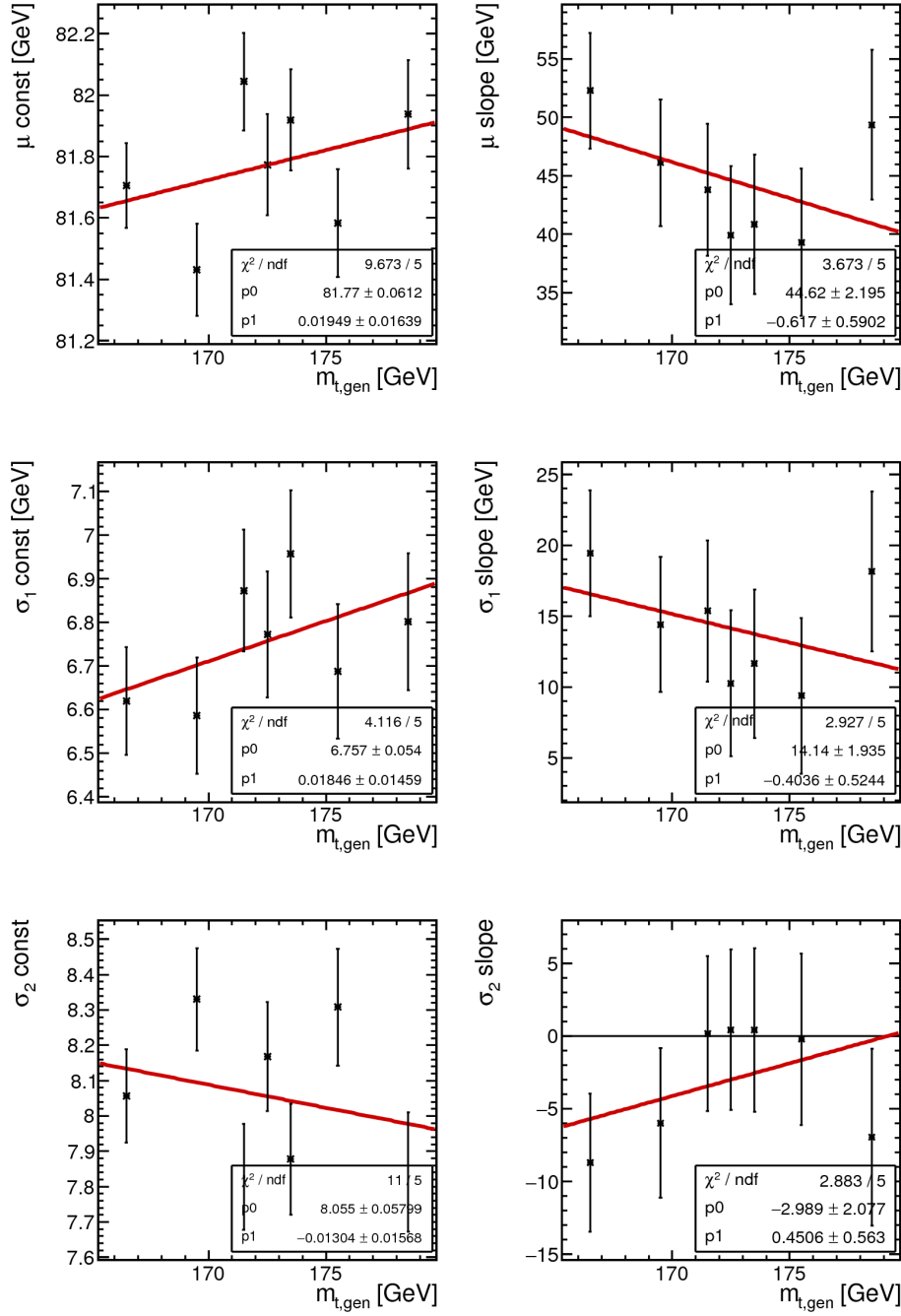


Figure A.27.: (Electron+jets) The Gaussian mean (upper left) and resolutions (middle/lower left) and their dependency on JSF (upper/middle/lower right) of the m_W^{reco} distributions of wrong $t\bar{t}$ permutations, for seven generated masses $m_{t,\text{gen}}$.

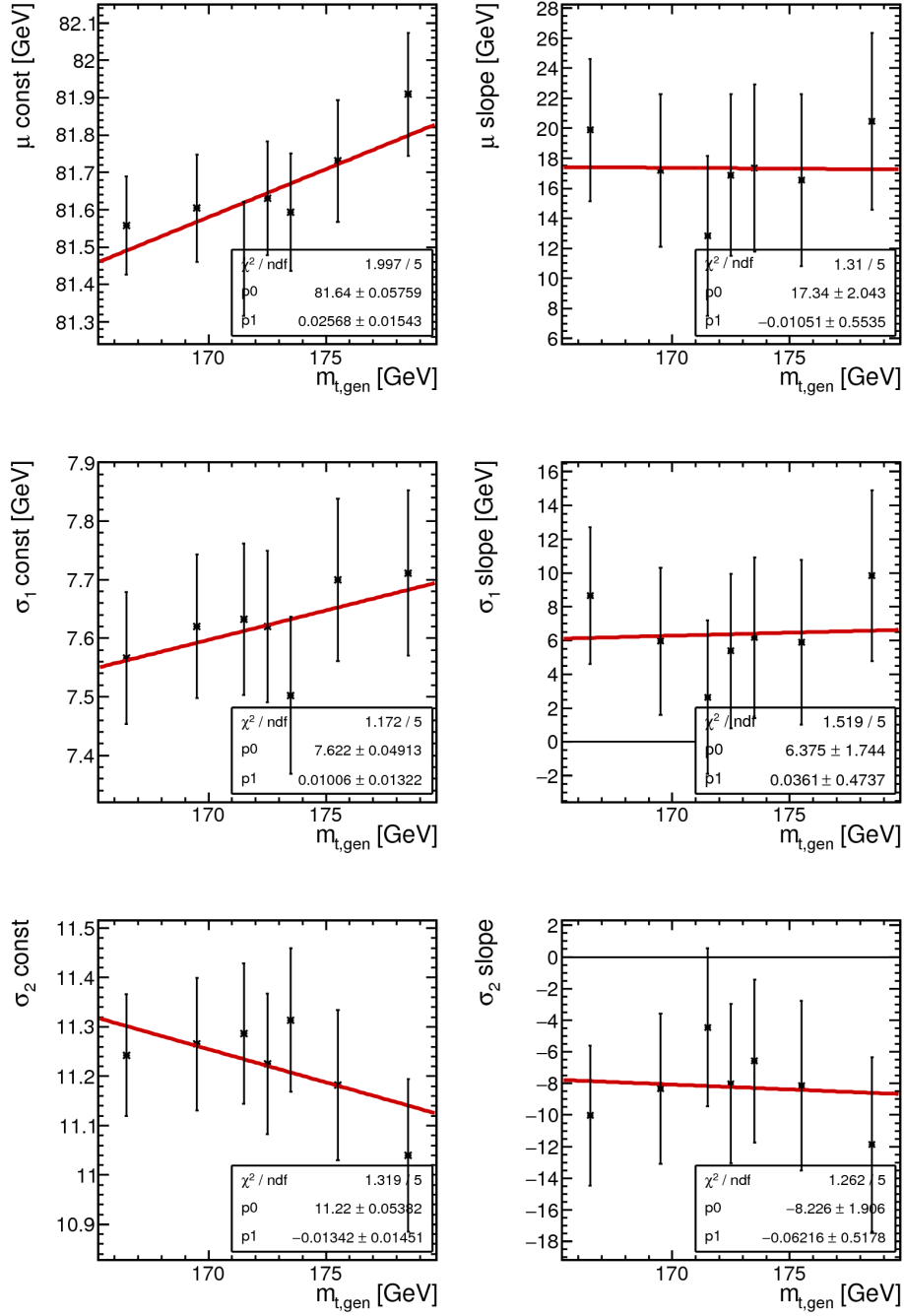


Figure A.28.: (Electron+jets) The Gaussian mean (upper left) and resolutions (middle/lower left) and their dependency on JSF (upper/middle/lower right) of the m_W^{reco} distributions of unmatched $t\bar{t}$ permutations, for seven generated masses $m_{t,\text{gen}}$.

A.6. Details on systematic uncertainties

Signed uncertainties for lepton+jets channel

Table A.5.: Lepton+jets: List of systematic uncertainties for the combined fit to the entire data set, including statistical uncertainties and sign information.

The shift is defined as variation – nominal in simulation.

Uncertainty name	δm_t^{2D}	δJSF^{2D}	δm_t^{1D}	δm_t^{hyb}	$\delta \text{JSF}^{\text{hyb}}$
def. ERD off to ERD on	0.22±0.09	0.008±0.001	0.42±0.05	0.03±0.07	0.005±0.001
- shift	-0.22	+0.008	+0.42	-0.03	+0.005
def. to CR GluonMove (both ERD off)	0.11±0.09	0.001±0.001	0.06±0.05	0.10±0.07	0.000±0.001
- shift	-0.11	+0.001	-0.06	-0.10	+0.000
CR GluonMove ERD on to ERD off	0.24±0.09	0.006±0.001	0.72±0.06	0.38±0.08	0.004±0.001
- shift	-0.24	-0.006	-0.72	-0.38	-0.004
CR QCDbased to GluonMove (both ERD on)	0.45±0.09	0.000±0.001	0.42±0.06	0.44±0.08	0.000±0.001
- shift	+0.45	-0.000	+0.42	+0.44	-0.000
def. to CR GluonMove (both ERD on)	0.34±0.09	0.001±0.001	0.23±0.06	0.31±0.08	0.001±0.001
- shift	+0.34	-0.001	+0.23	+0.31	-0.001
def. to CR QCDbased (both ERD on)	0.11±0.09	0.001±0.001	0.19±0.06	0.13±0.08	0.001±0.001
- shift	-0.11	-0.001	-0.19	-0.13	-0.001
FSR var.	0.25±0.06	0.010±0.001	0.61±0.04	0.06±0.05	0.007±0.000
- up	+0.25	-0.010	-0.61	-0.01	-0.007
- down	-0.12	+0.007	+0.49	+0.06	+0.005
- shift	+0.25	-0.010	-0.61	-0.01	-0.007
FSR var. (with logSF)	0.24±0.06	0.004±0.001	0.22±0.04	0.13±0.05	0.003±0.000
- up	+0.24	-0.004	-0.13	+0.13	-0.003
- down	-0.10	+0.004	+0.22	-0.01	+0.003
- shift	+0.24	-0.004	-0.13	+0.13	-0.003
Factorisation	0.01±0.00	0.000±0.000	0.02±0.00	0.01±0.00	0.000±0.000
- up	-0.00	-0.000	-0.01	-0.01	-0.000
- down	+0.01	+0.000	+0.02	+0.01	+0.000
- shift	-0.00	-0.000	-0.01	-0.01	-0.000
ISR var.	0.07±0.09	0.001±0.001	0.10±0.05	0.06±0.07	0.000±0.001
- up	+0.05	+0.001	+0.10	+0.06	+0.000
- down	+0.07	-0.001	+0.02	+0.05	-0.000
- shift	+0.05	+0.001	+0.10	+0.06	+0.000
Isolation	0.00±0.00	0.000±0.000	0.00±0.00	0.00±0.00	0.000±0.000
- up	+0.00	-0.000	+0.00	+0.00	-0.000
- down	+0.00	-0.000	-0.00	+0.00	-0.000
- shift	+0.00	-0.000	+0.00	+0.00	-0.000
JEC : TotalNoFlavorNoTime	0.13±0.00	0.012±0.000	0.85±0.00	0.19±0.00	0.008±0.000

Continued on next page

Table A.5 – continued from previous page

Uncertainty name	δm_t^{2D}	δJSF^{2D}	δm_t^{1D}	δm_t^{hyb}	$\delta \text{JSF}^{\text{hyb}}$
- up	-0.09	+0.011	+0.85	+0.19	+0.008
- down	+0.13	-0.012	-0.83	-0.17	-0.008
- shift	-0.09	+0.011	+0.85	+0.19	+0.008
JEC : Intercalibration	0.02±0.00	0.002±0.000	0.16±0.00	0.04±0.00	0.001±0.000
- up	-0.01	+0.002	+0.16	+0.04	+0.001
- down	+0.02	-0.002	-0.16	-0.03	-0.001
- shift	-0.01	+0.002	+0.16	+0.04	+0.001
JEC : MPFInSitu	0.01±0.00	0.003±0.000	0.23±0.00	0.07±0.00	0.002±0.000
- up	+0.00	+0.003	+0.23	+0.07	+0.002
- down	+0.01	-0.003	-0.23	-0.06	-0.002
- shift	+0.00	+0.003	+0.23	+0.07	+0.002
JEC : PileUp	0.12±0.00	0.004±0.000	0.26±0.00	0.01±0.00	0.003±0.000
- up	-0.10	+0.004	+0.26	+0.01	+0.003
- down	+0.12	-0.004	-0.25	+0.01	-0.003
- shift	-0.10	+0.004	+0.26	+0.01	+0.003
JEC : Uncorrelated	0.13±0.00	0.011±0.000	0.78±0.00	0.16±0.00	0.008±0.000
- up	-0.10	+0.011	+0.78	+0.16	+0.007
- down	+0.13	-0.011	-0.77	-0.14	-0.008
- shift	-0.10	+0.011	+0.78	+0.16	+0.007
JER	0.08±0.00	0.001±0.000	0.04±0.00	0.04±0.00	0.001±0.000
- up	-0.06	+0.001	+0.04	-0.03	+0.001
- down	+0.08	-0.001	-0.04	+0.04	-0.001
- shift	-0.06	+0.001	+0.04	-0.03	+0.001
Lepton Identification	0.01±0.00	0.000±0.000	0.00±0.00	0.01±0.00	0.000±0.000
- up	+0.01	-0.000	+0.00	+0.01	-0.000
- down	+0.01	-0.000	+0.00	+0.00	-0.000
- shift	+0.01	-0.000	+0.00	+0.01	-0.000
ME-PS matching	0.08±0.00	0.001±0.000	0.03±0.00	0.05±0.00	0.001±0.000
- up	-0.03	+0.001	+0.03	-0.01	+0.000
- down	+0.08	-0.001	-0.02	+0.05	-0.001
- shift	-0.03	+0.001	+0.03	-0.01	+0.000
PDF CT10nlo	0.01±0.00	0.000±0.000	0.02±0.00	0.02±0.00	0.000±0.000
- shift	-0.01	-0.000	-0.02	-0.02	-0.000
PDF MMHT2014nlo68clas118	0.01±0.00	0.000±0.000	0.01±0.00	0.01±0.00	0.000±0.000
- shift	-0.01	+0.000	-0.01	-0.01	+0.000
PDF as	0.00±0.00	0.000±0.000	0.00±0.00	0.00±0.00	0.000±0.000
- up	+0.00	-0.000	+0.00	+0.00	-0.000
- down	+0.00	-0.000	+0.00	+0.00	-0.000
- shift	+0.00	-0.000	+0.00	+0.00	-0.000
Pileup	0.08±0.00	0.001±0.000	0.02±0.00	0.05±0.00	0.001±0.000
- up	-0.07	+0.001	+0.02	-0.04	+0.001
- down	+0.08	-0.001	-0.02	+0.05	-0.001
- shift	-0.07	+0.001	+0.02	-0.04	+0.001
Renormalisation	0.02±0.00	0.000±0.000	0.01±0.00	0.01±0.00	0.000±0.000
- up	+0.02	-0.000	+0.01	+0.01	-0.000
- down	-0.01	+0.000	-0.01	-0.01	+0.000

Continued on next page

Table A.5 – continued from previous page

Uncertainty name	δm_t^{2D}	δJSF^{2D}	δm_t^{1D}	δm_t^{hyb}	$\delta \text{JSF}^{\text{hyb}}$
- shift	+0.02	-0.000	+0.01	+0.01	-0.000
Top-pt reweighting	0.03 ± 0.00	0.002 ± 0.000	0.17 ± 0.00	0.03 ± 0.00	0.002 ± 0.000
- shift	+0.03	-0.002	-0.17	-0.03	-0.002
Top-pt reweighting Run2	0.02 ± 0.00	0.001 ± 0.000	0.06 ± 0.00	0.01 ± 0.00	0.001 ± 0.000
- shift	+0.02	-0.001	-0.06	-0.01	-0.001
Trigger	0.00 ± 0.00	0.000 ± 0.000	0.00 ± 0.00	0.00 ± 0.00	0.000 ± 0.000
- up	-0.00	+0.000	+0.00	-0.00	+0.000
- down	+0.00	-0.000	+0.00	+0.00	-0.000
- shift	-0.00	+0.000	+0.00	-0.00	+0.000
Underlying Event	0.10 ± 0.08	0.001 ± 0.001	0.01 ± 0.05	0.07 ± 0.07	0.001 ± 0.001
- up	-0.10	+0.001	+0.01	-0.07	+0.001
- down	-0.00	+0.000	+0.01	+0.00	+0.000
- shift	-0.10	+0.001	+0.01	-0.07	+0.001
aMC@NLO v2.2.2 FxFx	0.19 ± 0.14	0.001 ± 0.001	0.29 ± 0.08	0.22 ± 0.11	0.001 ± 0.001
- shift	+0.19	+0.001	+0.29	+0.22	+0.001
b Tag	0.03 ± 0.00	0.000 ± 0.000	0.01 ± 0.00	0.02 ± 0.00	0.000 ± 0.000
- up	+0.03	-0.000	+0.01	+0.02	-0.000
- down	-0.02	+0.000	-0.01	-0.02	+0.000
- shift	+0.03	-0.000	+0.01	+0.02	-0.000
b Tag Cjets	0.01 ± 0.00	0.000 ± 0.000	0.00 ± 0.00	0.01 ± 0.00	0.000 ± 0.000
- up	-0.01	+0.000	-0.00	-0.01	+0.000
- down	+0.01	-0.000	+0.00	+0.01	-0.000
- shift	-0.01	+0.000	-0.00	-0.01	+0.000
b frag. Bowler-Lund	0.07 ± 0.00	0.001 ± 0.000	0.01 ± 0.00	0.05 ± 0.00	0.000 ± 0.000
- up	-0.07	+0.001	-0.01	-0.05	+0.000
- down	+0.05	-0.000	+0.01	+0.04	-0.000
- shift	-0.07	+0.001	-0.01	-0.05	+0.000
b frag. Peterson	0.04 ± 0.00	0.000 ± 0.000	0.05 ± 0.00	0.04 ± 0.00	0.000 ± 0.000
- shift	+0.04	+0.000	+0.05	+0.04	+0.000
b misTag	0.01 ± 0.00	0.000 ± 0.000	0.00 ± 0.00	0.01 ± 0.00	0.000 ± 0.000
- up	-0.01	+0.000	-0.00	-0.00	+0.000
- down	+0.01	-0.000	+0.00	+0.01	-0.000
- shift	-0.01	+0.000	-0.00	-0.00	+0.000
calibration	0.04 ± 0.00	0.000 ± 0.000	0.01 ± 0.00	0.03 ± 0.00	0.000 ± 0.000
- shift	+0.04	-0.000	+0.01	+0.03	-0.000
evtgen	0.09 ± 0.00	0.001 ± 0.000	0.14 ± 0.00	0.11 ± 0.00	0.000 ± 0.000
- shift	+0.09	+0.001	+0.14	+0.11	+0.000
flavor bottom	0.32 ± 0.00	0.000 ± 0.000	0.31 ± 0.00	0.32 ± 0.00	0.000 ± 0.000
- up	-0.31	+0.000	-0.31	-0.31	+0.000
- down	+0.32	-0.000	+0.31	+0.32	-0.000
- shift	-0.31	+0.000	-0.31	-0.31	+0.000
flavor charm	0.03 ± 0.00	0.000 ± 0.000	0.01 ± 0.00	0.02 ± 0.00	0.000 ± 0.000
- up	+0.03	-0.000	-0.01	+0.02	-0.000
- down	-0.02	+0.000	+0.01	-0.01	+0.000
- shift	+0.03	-0.000	-0.01	+0.02	-0.000
flavor gluon	0.23 ± 0.00	0.003 ± 0.000	0.02 ± 0.00	0.15 ± 0.00	0.002 ± 0.000

Continued on next page

Table A.5 – continued from previous page

Uncertainty name	δm_t^{2D}	δJSF^{2D}	δm_t^{1D}	δm_t^{hyb}	$\delta \text{JSF}^{\text{hyb}}$
- up	-0.23	+0.003	+0.02	-0.15	+0.002
- down	+0.22	-0.003	-0.02	+0.15	-0.002
- shift	-0.23	+0.003	+0.02	-0.15	+0.002
flavor light	0.12±0.00	0.002±0.000	0.02±0.00	0.08±0.00	0.001±0.000
- up	+0.12	-0.002	-0.01	+0.08	-0.001
- down	-0.10	+0.001	+0.02	-0.06	+0.001
- shift	+0.12	-0.002	-0.01	+0.08	-0.001
MADGRAPH5 MLM	0.82±0.16	0.000±0.002	0.80±0.10	0.82±0.14	0.000±0.001
- shift	+0.82	-0.000	+0.80	+0.82	-0.000
non $t\bar{t}$ BG	0.04±0.00	0.001±0.000	0.02±0.00	0.02±0.00	0.001±0.000
- up	+0.04	-0.001	-0.02	+0.02	-0.001
- down	-0.04	+0.001	+0.02	-0.02	+0.001
- shift	+0.04	-0.001	-0.02	+0.02	-0.001
HERWIG++ 2.4	4.39±0.09	0.014±0.001	3.26±0.06	4.06±0.08	0.010±0.001
- shift	-4.39	+0.014	-3.26	-4.06	+0.010
μ_R and μ_F	0.01±0.00	0.000±0.000	0.01±0.00	0.01±0.00	0.000±0.000
- up	+0.01	-0.000	-0.01	+0.01	-0.000
- down	-0.01	+0.000	+0.00	-0.00	+0.000
- shift	+0.01	-0.000	-0.01	+0.01	-0.000
semi-leptonic BR of B	0.11±0.00	0.000±0.000	0.08±0.00	0.10±0.00	0.000±0.000
- up	+0.05	-0.000	+0.02	+0.04	-0.000
- down	+0.11	-0.000	+0.08	+0.10	-0.000
- shift	+0.05	-0.000	+0.02	+0.04	-0.000
Check FSR	0.48	0.009	0.44	0.26	0.01
JEC quad. sum	0.13	0.011	0.83	0.18	0.01
JEC: Flavor	0.41	0.003	0.31	0.36	0.00
- linear	-0.42	0.001	-0.31	-0.39	0.001
PDF	0.02	0.000	0.02	0.02	0.00
b tagging	0.03	0.000	0.01	0.03	0.00
b-jet modelling	0.13	0.001	0.09	0.12	0.00

Signed uncertainties for muon+jets channel

Table A.6.: Muon+jets: List of systematic uncertainties for the combined fit to the entire data set, including statistical uncertainties and sign information.
The shift is defined as variation – nominal in simulation.

Uncertainty name	δm_t^{2D}	δJSF^{2D}	δm_t^{1D}	δm_t^{hyb}	$\delta \text{JSF}^{\text{hyb}}$
def. ERD off to ERD on	0.15±0.11	0.008±0.001	0.47±0.07	0.04±0.09	0.005±0.001
- shift	-0.15	+0.008	+0.47	+0.04	+0.005
def. to CR GluonMove (both ERD off)	0.17±0.10	0.001±0.001	0.07±0.07	0.14±0.09	0.001±0.001
- shift	-0.17	+0.001	-0.07	-0.14	+0.001
CR GluonMove ERD on to ERD off	0.31±0.11	0.006±0.001	0.76±0.07	0.45±0.10	0.004±0.001

Continued on next page

Table A.6 – continued from previous page

Uncertainty name	δm_t^{2D}	δJSF^{2D}	δm_t^{1D}	δm_t^{hyb}	$\delta \text{JSF}^{\text{hyb}}$
- shift	-0.31	-0.006	-0.76	-0.45	-0.004
CR QCDbased to GluonMove (both ERD on)	0.44±0.12	0.000±0.001	0.43±0.07	0.44±0.10	0.000±0.001
- shift	+0.44	-0.000	+0.43	+0.44	-0.000
def. to CR GluonMove (both ERD on)	0.29±0.11	0.001±0.001	0.21±0.07	0.27±0.10	0.001±0.001
- shift	+0.29	-0.001	+0.21	+0.27	-0.001
def. to CR QCDbased (both ERD on)	0.15±0.11	0.001±0.001	0.22±0.07	0.17±0.09	0.001±0.001
- shift	-0.15	-0.001	-0.22	-0.17	-0.001
FSR var.	0.32±0.08	0.011±0.001	0.55±0.05	0.07±0.06	0.007±0.001
- up	+0.32	-0.011	-0.55	+0.05	-0.007
- down	-0.13	+0.008	+0.54	+0.07	+0.006
- shift	+0.32	-0.011	-0.55	+0.05	-0.007
FSR var. (with logSF)	0.32±0.08	0.005±0.001	0.28±0.05	0.20±0.06	0.004±0.000
- up	+0.32	-0.005	-0.09	+0.20	-0.004
- down	-0.11	+0.005	+0.28	+0.00	+0.003
- shift	+0.32	-0.005	-0.09	+0.20	-0.004
Factorisation	0.01±0.00	0.000±0.000	0.02±0.00	0.01±0.00	0.000±0.000
- up	-0.01	-0.000	-0.02	-0.01	-0.000
- down	+0.01	+0.000	+0.02	+0.01	+0.000
- shift	-0.01	-0.000	-0.02	-0.01	-0.000
ISR var.	0.10±0.11	0.000±0.001	0.13±0.07	0.11±0.09	0.000±0.001
- up	+0.10	+0.000	+0.13	+0.11	+0.000
- down	+0.03	+0.000	+0.05	+0.03	+0.000
- shift	+0.10	+0.000	+0.13	+0.11	+0.000
Isolation	0.00±0.00	0.000±0.000	0.00±0.00	0.00±0.00	0.000±0.000
- up	+0.00	-0.000	+0.00	+0.00	-0.000
- down	+0.00	-0.000	+0.00	+0.00	-0.000
- shift	+0.00	-0.000	+0.00	+0.00	-0.000
JEC : Intercalibration	0.03±0.00	0.002±0.000	0.15±0.00	0.04±0.00	0.002±0.000
- up	-0.01	+0.002	+0.15	+0.04	+0.001
- down	+0.03	-0.002	-0.15	-0.03	-0.002
- shift	-0.01	+0.002	+0.15	+0.04	+0.001
JEC : MPFInSitu	0.01±0.00	0.003±0.000	0.23±0.00	0.07±0.00	0.002±0.000
- up	+0.01	+0.003	+0.23	+0.07	+0.002
- down	+0.01	-0.003	-0.22	-0.06	-0.002
- shift	+0.01	+0.003	+0.23	+0.07	+0.002
JEC : PileUp	0.11±0.00	0.004±0.000	0.26±0.00	0.01±0.00	0.003±0.000
- up	-0.09	+0.004	+0.26	+0.01	+0.003
- down	+0.11	-0.004	-0.24	+0.01	-0.003
- shift	-0.09	+0.004	+0.26	+0.01	+0.003
JEC : Uncorrelated	0.11±0.00	0.011±0.000	0.76±0.00	0.17±0.00	0.007±0.000
- up	-0.08	+0.010	+0.76	+0.17	+0.007
- down	+0.11	-0.011	-0.74	-0.15	-0.007
- shift	-0.08	+0.010	+0.76	+0.17	+0.007

Continued on next page

Table A.6 – continued from previous page

Uncertainty name	δm_t^{2D}	δJSF^{2D}	δm_t^{1D}	δm_t^{hyb}	$\delta \text{JSF}^{\text{hyb}}$
JER	0.08 ± 0.00	0.001 ± 0.000	0.04 ± 0.00	0.04 ± 0.00	0.001 ± 0.000
- up	-0.06	+0.001	+0.04	-0.03	+0.001
- down	+0.08	-0.001	-0.03	+0.04	-0.001
- shift	-0.06	+0.001	+0.04	-0.03	+0.001
Lepton Identification	0.00 ± 0.00	0.000 ± 0.000	0.00 ± 0.00	0.00 ± 0.00	0.000 ± 0.000
- up	+0.00	-0.000	+0.00	+0.00	-0.000
- down	+0.00	-0.000	+0.00	+0.00	-0.000
- shift	+0.00	-0.000	+0.00	+0.00	-0.000
ME-PS matching	0.09 ± 0.00	0.001 ± 0.000	0.02 ± 0.00	0.07 ± 0.00	0.001 ± 0.000
- up	-0.02	+0.001	+0.02	-0.01	+0.000
- down	+0.09	-0.001	+0.01	+0.07	-0.001
- shift	-0.02	+0.001	+0.02	-0.01	+0.000
PDF CT10nlo	0.02 ± 0.00	0.000 ± 0.000	0.02 ± 0.00	0.02 ± 0.00	0.000 ± 0.000
- shift	-0.02	-0.000	-0.02	-0.02	-0.000
PDF MMHT2014nlo68clas118	0.02 ± 0.00	0.000 ± 0.000	0.02 ± 0.00	0.02 ± 0.00	0.000 ± 0.000
- shift	-0.02	-0.000	-0.02	-0.02	-0.000
PDF as	0.01 ± 0.00	0.000 ± 0.000	0.01 ± 0.00	0.01 ± 0.00	0.000 ± 0.000
- up	+0.01	-0.000	+0.01	+0.01	-0.000
- down	+0.00	-0.000	-0.00	+0.00	-0.000
- shift	+0.01	-0.000	+0.01	+0.01	-0.000
Pileup	0.08 ± 0.00	0.001 ± 0.000	0.02 ± 0.00	0.05 ± 0.00	0.001 ± 0.000
- up	-0.08	+0.001	+0.02	-0.05	+0.001
- down	+0.08	-0.001	-0.02	+0.05	-0.001
- shift	-0.08	+0.001	+0.02	-0.05	+0.001
Renormalisation	0.02 ± 0.00	0.000 ± 0.000	0.01 ± 0.00	0.02 ± 0.00	0.000 ± 0.000
- up	+0.02	-0.000	+0.01	+0.02	-0.000
- down	-0.02	+0.000	-0.01	-0.01	+0.000
- shift	+0.02	-0.000	+0.01	+0.02	-0.000
Top-pt reweighting	0.00 ± 0.00	0.002 ± 0.000	0.19 ± 0.00	0.05 ± 0.00	0.002 ± 0.000
- shift	+0.00	-0.002	-0.19	-0.05	-0.002
Top-pt reweighting Run2	0.00 ± 0.00	0.001 ± 0.000	0.06 ± 0.00	0.02 ± 0.00	0.001 ± 0.000
- shift	+0.00	-0.001	-0.06	-0.02	-0.001
Trigger	0.00 ± 0.00	0.000 ± 0.000	0.00 ± 0.00	0.00 ± 0.00	0.000 ± 0.000
- up	+0.00	-0.000	+0.00	+0.00	-0.000
- down	+0.00	-0.000	+0.00	+0.00	-0.000
- shift	+0.00	-0.000	+0.00	+0.00	-0.000
Underlying Event	0.06 ± 0.10	0.001 ± 0.001	0.08 ± 0.06	0.04 ± 0.08	0.001 ± 0.001
- up	+0.01	+0.001	+0.08	+0.03	+0.001
- down	-0.06	+0.001	-0.00	-0.04	+0.000
- shift	+0.01	+0.001	+0.08	+0.03	+0.001
aMC@NLO v2.2.2 FxFx	0.28 ± 0.17	0.001 ± 0.002	0.23 ± 0.10	0.26 ± 0.14	0.000 ± 0.001
- shift	+0.28	-0.001	+0.23	+0.26	-0.000
b Tag	0.02 ± 0.00	0.000 ± 0.000	0.01 ± 0.00	0.02 ± 0.00	0.000 ± 0.000
- up	+0.02	-0.000	+0.01	+0.02	-0.000
- down	-0.02	+0.000	-0.01	-0.02	+0.000
- shift	+0.02	-0.000	+0.01	+0.02	-0.000

Continued on next page

Table A.6 – continued from previous page

Uncertainty name	δm_t^{2D}	δJSF^{2D}	δm_t^{1D}	δm_t^{hyb}	$\delta \text{JSF}^{\text{hyb}}$
b Tag Cjets	0.01±0.00	0.000±0.000	0.00±0.00	0.01±0.00	0.000±0.000
- up	-0.01	+0.000	-0.00	-0.01	+0.000
- down	+0.01	-0.000	+0.00	+0.01	-0.000
- shift	-0.01	+0.000	-0.00	-0.01	+0.000
b frag. Bowler-Lund	0.06±0.00	0.001±0.000	0.01±0.00	0.04±0.00	0.000±0.000
- up	-0.06	+0.001	-0.01	-0.04	+0.000
- down	+0.05	-0.000	+0.01	+0.04	-0.000
- shift	-0.06	+0.001	-0.01	-0.04	+0.000
b frag. Peterson	0.00±0.00	0.000±0.000	0.04±0.00	0.01±0.00	0.000±0.000
- shift	-0.00	+0.000	+0.04	+0.01	+0.000
b misTag	0.01±0.00	0.000±0.000	0.00±0.00	0.01±0.00	0.000±0.000
- up	-0.00	+0.000	-0.00	-0.00	+0.000
- down	+0.01	-0.000	+0.00	+0.01	-0.000
- shift	-0.00	+0.000	-0.00	-0.00	+0.000
calibration	0.05±0.00	0.000±0.000	0.04±0.00	0.05±0.00	0.000±0.000
- shift	+0.05	-0.000	+0.04	+0.05	-0.000
evtgen	0.14±0.00	0.001±0.000	0.22±0.00	0.16±0.00	0.001±0.000
- shift	+0.14	+0.001	+0.22	+0.16	+0.001
flavor bottom	0.31±0.00	0.000±0.000	0.31±0.00	0.31±0.00	0.000±0.000
- up	-0.30	-0.000	-0.31	-0.30	-0.000
- down	+0.31	-0.000	+0.31	+0.31	-0.000
- shift	-0.30	-0.000	-0.31	-0.30	-0.000
flavor charm	0.02±0.00	0.000±0.000	0.01±0.00	0.01±0.00	0.000±0.000
- up	+0.02	-0.000	-0.01	+0.01	-0.000
- down	-0.02	+0.000	+0.01	-0.01	+0.000
- shift	+0.02	-0.000	-0.01	+0.01	-0.000
flavor gluon	0.24±0.00	0.003±0.000	0.01±0.00	0.17±0.00	0.002±0.000
- up	-0.24	+0.003	+0.00	-0.17	+0.002
- down	+0.23	-0.003	+0.01	+0.16	-0.002
- shift	-0.24	+0.003	+0.00	-0.17	+0.002
flavor light	0.12±0.00	0.001±0.000	0.02±0.00	0.08±0.00	0.001±0.000
- up	+0.12	-0.001	-0.00	+0.08	-0.001
- down	-0.09	+0.001	+0.02	-0.06	+0.001
- shift	+0.12	-0.001	-0.00	+0.08	-0.001
MADGRAPH5 MLM	0.68±0.21	0.001±0.002	0.77±0.13	0.70±0.18	0.001±0.001
- shift	+0.68	+0.001	+0.77	+0.70	+0.001
non t $\bar{t}$ BG	0.03±0.00	0.001±0.000	0.08±0.00	0.00±0.00	0.001±0.000
- up	+0.03	-0.001	-0.08	-0.00	-0.001
- down	-0.03	+0.001	+0.08	+0.00	+0.001
- shift	+0.03	-0.001	-0.08	-0.00	-0.001
HERWIG++ 2.4	4.37±0.11	0.014±0.001	3.31±0.07	4.06±0.10	0.010±0.001
- shift	-4.37	+0.014	-3.31	-4.06	+0.010
μ_R and μ_F	0.00±0.00	0.000±0.000	0.02±0.00	0.01±0.00	0.000±0.000
- up	-0.00	-0.000	-0.02	-0.01	-0.000
- down	-0.00	+0.000	+0.01	+0.00	+0.000
- shift	-0.00	-0.000	-0.02	-0.01	-0.000

Continued on next page

Table A.6 – continued from previous page

Uncertainty name	δm_t^{2D}	δJSF^{2D}	δm_t^{1D}	δm_t^{hyb}	$\delta \text{JSF}^{\text{hyb}}$
semi-leptonic BR of B	0.06 ± 0.00	0.000 ± 0.000	0.06 ± 0.00	0.06 ± 0.00	0.000 ± 0.000
- up	+0.01	+0.000	+0.01	+0.01	+0.000
- down	+0.06	+0.000	+0.06	+0.06	+0.000
- shift	+0.01	+0.000	+0.01	+0.01	+0.000
TotalNoFlavorNoTime	0.10 ± 0.00	0.011 ± 0.000	0.82 ± 0.00	0.19 ± 0.00	0.008 ± 0.000
- up	-0.07	+0.011	+0.82	+0.19	+0.008
- down	+0.10	-0.011	-0.81	-0.17	-0.008
- shift	-0.07	+0.011	+0.82	+0.19	+0.008
JEC quad. sum	0.11	0.011	0.81	0.19	0.01
JEC: Flavor	0.41	0.003	0.31	0.36	0.00
- linear	-0.43	0.001	-0.34	-0.40	0.001
PDF	0.03	0.000	0.03	0.03	0.00
b tagging	0.03	0.000	0.01	0.02	0.00
b-jet modelling	0.08	0.001	0.07	0.07	0.00

Signed uncertainties for electron+jets channel

Table A.7.: Electron+jets: List of systematic uncertainties for the combined fit to the entire data set, including statistical uncertainties and sign information. The shift is defined as variation – nominal in simulation.

Uncertainty name	δm_t^{2D}	δJSF^{2D}	δm_t^{1D}	δm_t^{hyb}	$\delta \text{JSF}^{\text{hyb}}$
def. ERD off to ERD on	0.34±0.14	0.008±0.001	0.34±0.09	0.14±0.12	0.006±0.001
- shift	-0.34	+0.008	+0.34	-0.14	+0.006
def. to CR GluonMove (both ERD off)	0.00±0.15	0.001±0.001	0.05±0.09	0.02±0.12	0.000±0.001
- shift	-0.00	-0.001	-0.05	-0.02	-0.000
CR GluonMove ERD on to ERD off	0.08±0.16	0.007±0.001	0.65±0.09	0.25±0.13	0.005±0.001
- shift	-0.08	-0.007	-0.65	-0.25	-0.005
CR QCDbased to GluonMove (both ERD on)	0.45±0.16	0.001±0.001	0.40±0.09	0.43±0.13	0.000±0.001
- shift	+0.45	-0.001	+0.40	+0.43	-0.000
def. to CR GluonMove (both ERD on)	0.42±0.16	0.002±0.001	0.26±0.09	0.37±0.13	0.001±0.001
- shift	+0.42	-0.002	+0.26	+0.37	-0.001
def. to CR QCDbased (both ERD on)	0.03±0.15	0.001±0.001	0.14±0.09	0.06±0.13	0.001±0.001
- shift	-0.03	-0.001	-0.14	-0.06	-0.001
FSR var.	0.11±0.11	0.009±0.001	0.70±0.06	0.14±0.09	0.007±0.001
- up	+0.11	-0.009	-0.70	-0.14	-0.007
- down	-0.08	+0.006	+0.42	+0.07	+0.004
- shift	+0.11	-0.009	-0.70	-0.14	-0.007
FSR var. (with logSF)	0.09±0.11	0.003±0.001	0.19±0.06	0.01±0.09	0.002±0.001
- up	+0.09	-0.003	-0.19	+0.01	-0.002
- down	-0.07	+0.002	+0.13	-0.01	+0.002
- shift	+0.09	-0.003	-0.19	+0.01	-0.002
Factorisation	0.01±0.00	0.000±0.000	0.01±0.00	0.01±0.00	0.000±0.000
- up	+0.01	-0.000	-0.01	+0.01	-0.000
- down	-0.01	+0.000	+0.01	-0.01	+0.000
- shift	+0.01	-0.000	-0.01	+0.01	-0.000
ISR var.	0.15±0.15	0.002±0.001	0.05±0.09	0.10±0.12	0.001±0.001
- up	-0.07	+0.001	+0.05	-0.04	+0.001
- down	+0.15	-0.002	-0.03	+0.10	-0.001
- shift	-0.07	+0.001	+0.05	-0.04	+0.001
Isolation	0.00±0.00	0.000±0.000	0.00±0.00	0.00±0.00	0.000±0.000
- up	-0.00	+0.000	-0.00	-0.00	+0.000
- down	-0.00	-0.000	-0.00	-0.00	-0.000
- shift	-0.00	+0.000	-0.00	-0.00	+0.000
JEC : TotalNoFlavorNoTime	0.16±0.00	0.012±0.000	0.89±0.00	0.18±0.00	0.008±0.000
- up	-0.12	+0.012	+0.89	+0.18	+0.008
- down	+0.16	-0.012	-0.88	-0.16	-0.008
- shift	-0.12	+0.012	+0.89	+0.18	+0.008

Continued on next page

Table A.7 – continued from previous page

Uncertainty name	δm_t^{2D}	δJSF^{2D}	δm_t^{1D}	δm_t^{hyb}	$\delta \text{JSF}^{\text{hyb}}$
JEC : Intercalibration	0.03±0.00	0.002±0.000	0.17±0.00	0.05±0.00	0.001±0.000
- up	-0.03	+0.002	+0.15	+0.03	+0.001
- down	+0.00	-0.002	-0.17	-0.05	-0.001
- shift	-0.03	+0.002	+0.15	+0.03	+0.001
JEC : MPFInSitu	0.01±0.00	0.003±0.000	0.24±0.00	0.07±0.00	0.002±0.000
- up	-0.01	+0.003	+0.24	+0.06	+0.002
- down	+0.01	-0.003	-0.24	-0.07	-0.002
- shift	-0.01	+0.003	+0.24	+0.06	+0.002
JEC : PileUp	0.11±0.00	0.004±0.000	0.28±0.00	0.01±0.00	0.003±0.000
- up	-0.10	+0.004	+0.27	+0.01	+0.003
- down	+0.11	-0.004	-0.28	-0.01	-0.003
- shift	-0.10	+0.004	+0.27	+0.01	+0.003
JEC : Uncorrelated	0.15±0.00	0.011±0.000	0.82±0.00	0.15±0.00	0.008±0.000
- up	-0.13	+0.011	+0.81	+0.15	+0.008
- down	+0.15	-0.011	-0.82	-0.14	-0.008
- shift	-0.13	+0.011	+0.81	+0.15	+0.008
JER	0.08±0.00	0.001±0.000	0.05±0.00	0.04±0.00	0.001±0.000
- up	-0.06	+0.001	+0.05	-0.03	+0.001
- down	+0.08	-0.001	-0.05	+0.04	-0.001
- shift	-0.06	+0.001	+0.05	-0.03	+0.001
Lepton Identification	0.01±0.00	0.000±0.000	0.01±0.00	0.01±0.00	0.000±0.000
- up	-0.00	+0.000	-0.00	-0.00	+0.000
- down	-0.01	+0.000	-0.01	-0.01	+0.000
- shift	-0.00	+0.000	-0.00	-0.00	+0.000
ME-PS matching	0.06±0.00	0.001±0.000	0.07±0.00	0.03±0.00	0.001±0.000
- up	-0.06	+0.001	+0.04	-0.03	+0.001
- down	+0.04	-0.001	-0.07	+0.01	-0.001
- shift	-0.06	+0.001	+0.04	-0.03	+0.001
PDF CT10nlo	0.02±0.00	0.000±0.000	0.01±0.00	0.02±0.00	0.000±0.000
- shift	-0.02	+0.000	-0.01	-0.02	+0.000
PDF MMHT2014nlo68clas118	0.02±0.00	0.000±0.000	0.01±0.00	0.02±0.00	0.000±0.000
- shift	-0.02	+0.000	-0.01	-0.02	+0.000
PDF as	0.00±0.00	0.000±0.000	0.01±0.00	0.00±0.00	0.000±0.000
- up	+0.00	+0.000	+0.00	+0.00	+0.000
- down	-0.00	-0.000	-0.01	-0.00	-0.000
- shift	+0.00	+0.000	+0.00	+0.00	+0.000
Pileup	0.07±0.00	0.001±0.000	0.04±0.00	0.04±0.00	0.001±0.000
- up	-0.07	+0.001	+0.03	-0.04	+0.001
- down	+0.06	-0.001	-0.04	+0.03	-0.001
- shift	-0.07	+0.001	+0.03	-0.04	+0.001
Renormalisation	0.01±0.00	0.000±0.000	0.02±0.00	0.01±0.00	0.000±0.000
- up	+0.01	+0.000	+0.02	+0.01	+0.000
- down	-0.01	-0.000	-0.02	-0.01	-0.000
- shift	+0.01	+0.000	+0.02	+0.01	+0.000
Top-pt reweighting	0.07 ± 0.00	0.002 ± 0.000	0.14 ± 0.00	0.01 ± 0.00	0.002 ± 0.000
- shift	+0.07	-0.002	-0.14	+0.01	-0.002

Continued on next page

Table A.7 – continued from previous page

Uncertainty name	δm_t^{2D}	δJSF^{2D}	δm_t^{1D}	δm_t^{hyb}	$\delta \text{JSF}^{\text{hyb}}$
Top-pt reweighting Run2	0.03 ± 0.00	0.001 ± 0.000	0.06 ± 0.00	0.00 ± 0.00	0.001 ± 0.000
- shift	+0.03	-0.001	-0.06	+0.00	-0.001
Trigger	0.00 ± 0.00	0.000 ± 0.000	0.00 ± 0.00	0.00 ± 0.00	0.000 ± 0.000
- up	-0.00	+0.000	-0.00	-0.00	+0.000
- down	-0.00	-0.000	-0.00	-0.00	-0.000
- shift	-0.00	+0.000	-0.00	-0.00	+0.000
Underlying Event	0.30 ± 0.13	0.002 ± 0.001	0.11 ± 0.08	0.24 ± 0.11	0.002 ± 0.001
- up	-0.30	+0.002	-0.11	-0.24	+0.002
- down	+0.12	-0.001	+0.03	+0.09	-0.001
- shift	-0.30	+0.002	-0.11	-0.24	+0.002
aMC@NLO v2.2.2 FxFx	0.01 ± 0.23	0.005 ± 0.002	0.38 ± 0.13	0.11 ± 0.18	0.003 ± 0.001
- shift	-0.01	+0.005	+0.38	+0.11	+0.003
b Tag	0.03 ± 0.00	0.000 ± 0.000	0.01 ± 0.00	0.03 ± 0.00	0.000 ± 0.000
- up	+0.03	-0.000	+0.01	+0.02	-0.000
- down	-0.03	+0.000	-0.01	-0.03	+0.000
- shift	+0.03	-0.000	+0.01	+0.02	-0.000
b Tag Cjets	0.02 ± 0.00	0.000 ± 0.000	0.01 ± 0.00	0.02 ± 0.00	0.000 ± 0.000
- up	-0.02	+0.000	-0.01	-0.02	+0.000
- down	+0.01	-0.000	+0.00	+0.01	-0.000
- shift	-0.02	+0.000	-0.01	-0.02	+0.000
b frag. Bowler-Lund	0.07 ± 0.00	0.001 ± 0.000	0.01 ± 0.00	0.05 ± 0.00	0.000 ± 0.000
- up	-0.07	+0.001	-0.01	-0.05	+0.000
- down	+0.06	-0.001	+0.01	+0.05	-0.000
- shift	-0.07	+0.001	-0.01	-0.05	+0.000
b frag. Peterson	0.12 ± 0.00	0.001 ± 0.000	0.07 ± 0.00	0.10 ± 0.00	0.000 ± 0.000
- shift	+0.12	-0.001	+0.07	+0.10	-0.000
b misTag	0.01 ± 0.00	0.000 ± 0.000	0.00 ± 0.00	0.00 ± 0.00	0.000 ± 0.000
- up	-0.01	+0.000	-0.00	-0.00	+0.000
- down	+0.00	-0.000	-0.00	+0.00	-0.000
- shift	-0.01	+0.000	-0.00	-0.00	+0.000
calibration	0.01 ± 0.00	0.001 ± 0.000	0.05 ± 0.00	0.01 ± 0.00	0.000 ± 0.000
- shift	+0.01	-0.001	-0.05	-0.01	-0.000
evtgen	0.00 ± 0.00	0.000 ± 0.000	0.01 ± 0.00	0.00 ± 0.00	0.000 ± 0.000
- shift	-0.00	+0.000	+0.01	+0.00	+0.000
flavor bottom	0.33 ± 0.00	0.000 ± 0.000	0.31 ± 0.00	0.32 ± 0.00	0.000 ± 0.000
- up	-0.33	+0.000	-0.31	-0.32	+0.000
- down	+0.33	-0.000	+0.31	+0.32	-0.000
- shift	-0.33	+0.000	-0.31	-0.32	+0.000
flavor charm	0.03 ± 0.00	0.000 ± 0.000	0.01 ± 0.00	0.02 ± 0.00	0.000 ± 0.000
- up	+0.03	-0.000	-0.01	+0.01	-0.000
- down	-0.03	+0.000	+0.00	-0.02	+0.000
- shift	+0.03	-0.000	-0.01	+0.01	-0.000
flavor gluon	0.21 ± 0.00	0.003 ± 0.000	0.06 ± 0.00	0.13 ± 0.00	0.002 ± 0.000
- up	-0.21	+0.003	+0.06	-0.13	+0.002
- down	+0.19	-0.003	-0.06	+0.11	-0.002
- shift	-0.21	+0.003	+0.06	-0.13	+0.002

Continued on next page

Table A.7 – continued from previous page

Uncertainty name	δm_t^{2D}	δJSF^{2D}	δm_t^{1D}	δm_t^{hyb}	$\delta \text{JSF}^{\text{hyb}}$
flavor light	0.12±0.00	0.001±0.000	0.02±0.00	0.08±0.00	0.001±0.000
- up	+0.11	-0.001	-0.02	+0.07	-0.001
- down	-0.12	+0.001	+0.01	-0.08	+0.001
- shift	+0.11	-0.001	-0.02	+0.07	-0.001
MADGRAPH5 MLM	1.09±0.28	0.003±0.003	0.86±0.17	1.02±0.23	0.002±0.002
- shift	+1.09	-0.003	+0.86	+1.02	-0.002
non $t\bar{t}$ BG	0.06±0.00	0.000±0.000	0.08±0.00	0.07±0.00	0.000±0.000
- up	+0.06	+0.000	+0.08	+0.07	+0.000
- down	-0.06	-0.000	-0.08	-0.07	-0.000
- shift	+0.06	+0.000	+0.08	+0.07	+0.000
HERWIG++ 2.4	4.42±0.15	0.015±0.001	3.18±0.10	4.05±0.13	0.010±0.001
- shift	-4.42	+0.015	-3.18	-4.05	+0.010
μ_R and μ_F	0.03±0.00	0.000±0.000	0.02±0.00	0.02±0.00	0.000±0.000
- up	+0.02	-0.000	+0.01	+0.02	-0.000
- down	-0.03	+0.000	-0.02	-0.02	+0.000
- shift	+0.02	-0.000	+0.01	+0.02	-0.000
semi-leptonic BR of B	0.20±0.00	0.001±0.000	0.10±0.00	0.17±0.00	0.001±0.000
- up	+0.13	-0.001	+0.04	+0.10	-0.001
- down	+0.20	-0.001	+0.10	+0.17	-0.001
- shift	+0.13	-0.001	+0.04	+0.10	-0.001
JEC quad. sum	0.15	0.012	0.87	0.17	0.01
JEC: Flavor	0.41	0.003	0.32	0.36	0.00
- linear	-0.40	0.001	-0.28	-0.37	0.001
PDF	0.03	0.000	0.01	0.03	0.00
b tagging	0.04	0.000	0.02	0.03	0.00
b-jet modelling	0.24	0.001	0.12	0.20	0.00

Continued on next page

A.7. Color reconnection

A crucial test for the performance of reliable UE tunes, and of the CR simulation in particular, is the description of the average transverse momentum of the charged particles, as a function of the charged particle multiplicity. Comparisons to these observables, measured by ATLAS at 13 TeV in the transMAX and transMIN regions [128], are shown in Figure A.29.

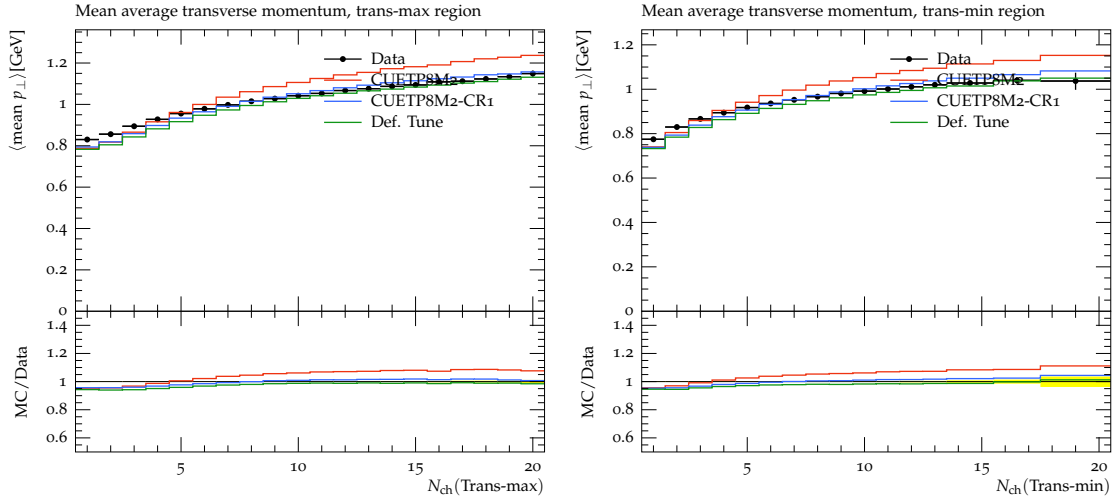


Figure A.29.: Average p_T of the charged particles in the transMIN (left) and transMAX (right) region, measured by ATLAS at 13 TeV [128]. The measurements are compared to predictions from the tunes CUETP8M2, CUETP8M2-CR1 and Def. Tune. The lower panels show the ratios between each Monte Carlo prediction and the data.

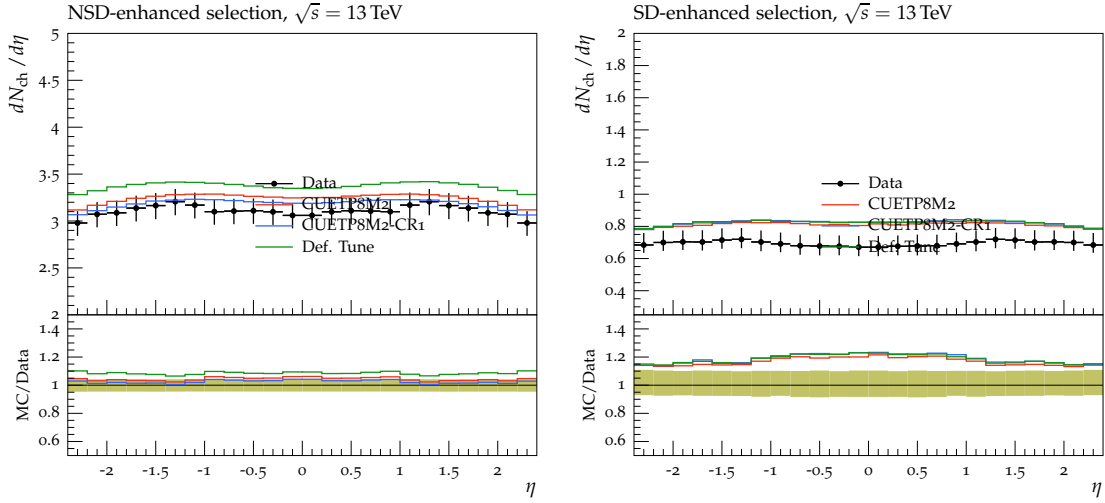


Figure A.30.: Charged particle multiplicity as a function of pseudorapidity after non-single-diffractive (NSD) enhanced (left) and single-diffractive (SD) enhanced (right) selections, measured by CMS at 13 TeV [130, 135]. The measurements are compared to predictions from the tunes CUETP8M2, CUETP8M2-CR1 and Def. Tune. The lower panels show the ratios between each Monte Carlo prediction and the data.

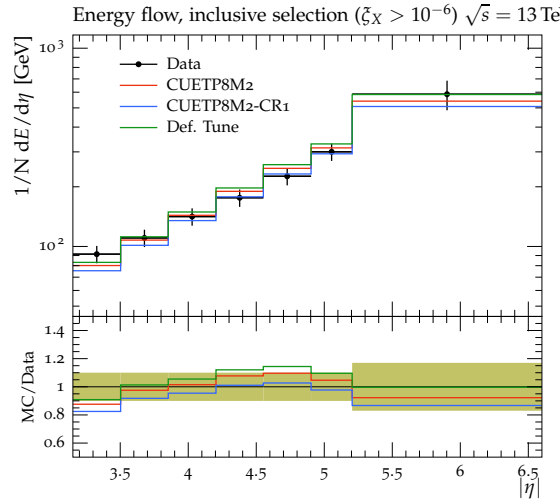


Figure A.31.: Forward energy flow as a function of pseudorapidity measured in events with a presence of a hard dijet system [136] measured by CMS at 13 TeV. The measurements are compared to predictions from the tunes CUETP8M2, CUETP8M2-CR1 and Def. Tune. The lower panels show the ratios between each Monte Carlo prediction and the data.

A.8. Measured top quark mass as a function of kinematic observables

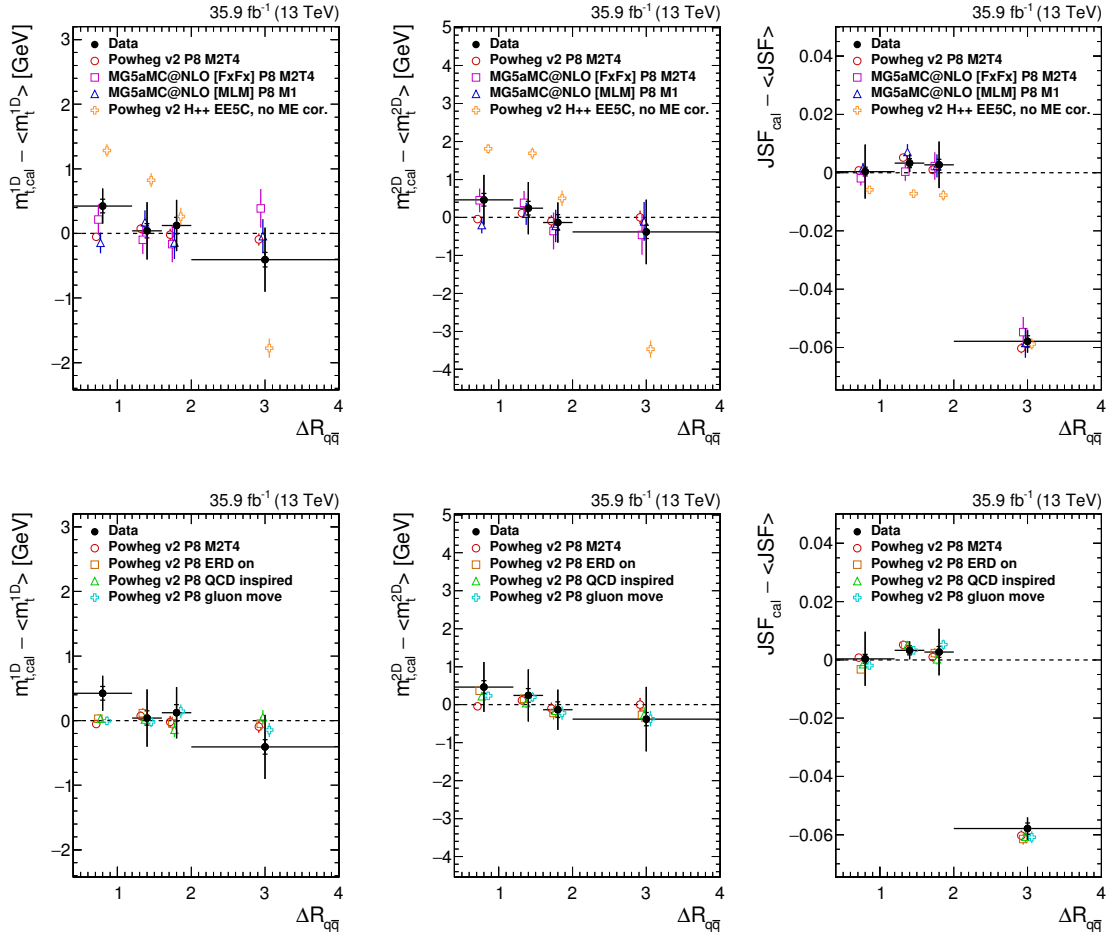


Figure A.32.: Measurements of m_t as a function of the ΔR between the light quark jets ($\Delta R_{q\bar{q}}$) compared to different generator choices (upper row) and to alternative models of color reconnections (lower row).

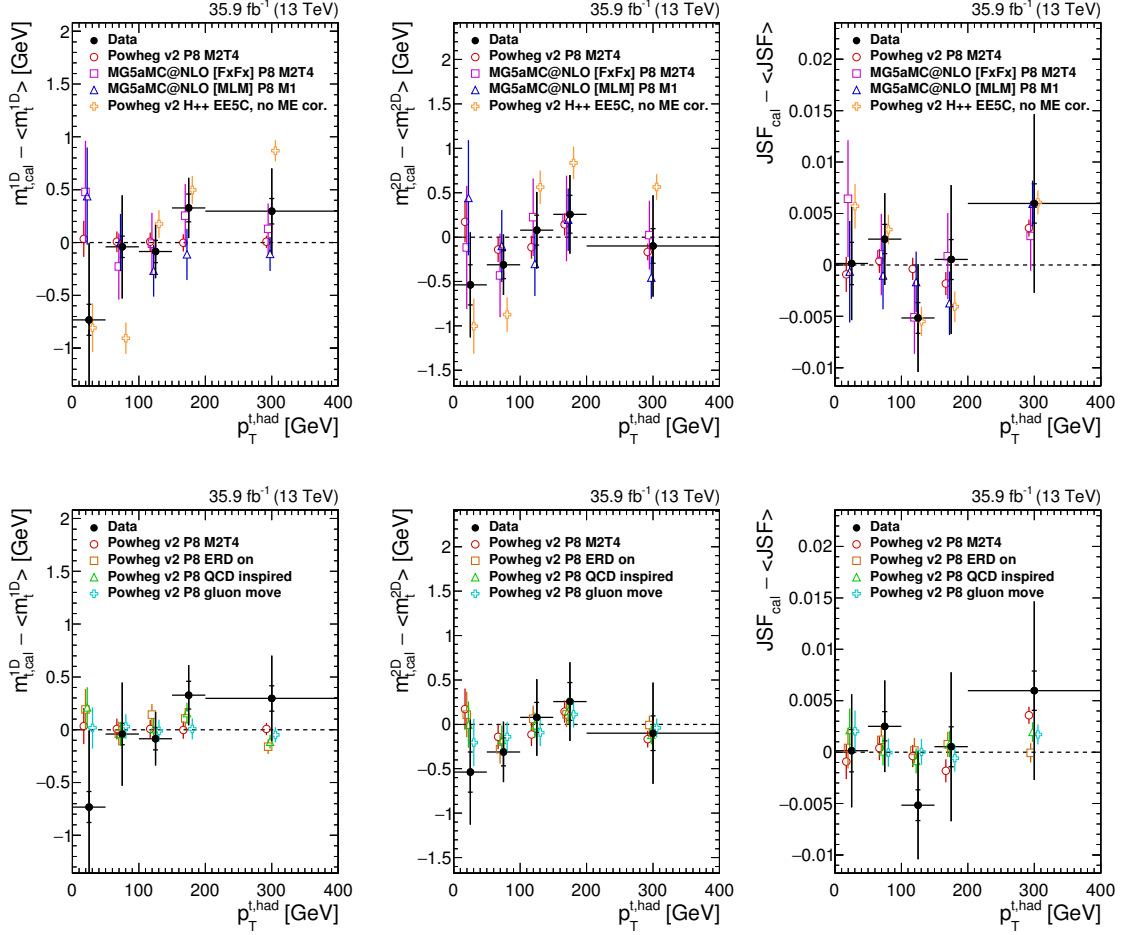


Figure A.33.: Measurements of m_t as a function of the transverse momentum of the hadronically decaying top quark ($p_T^{t, \text{had}}$) compared to different generator choices (upper row) and to alternative models of color reconnections (lower row).

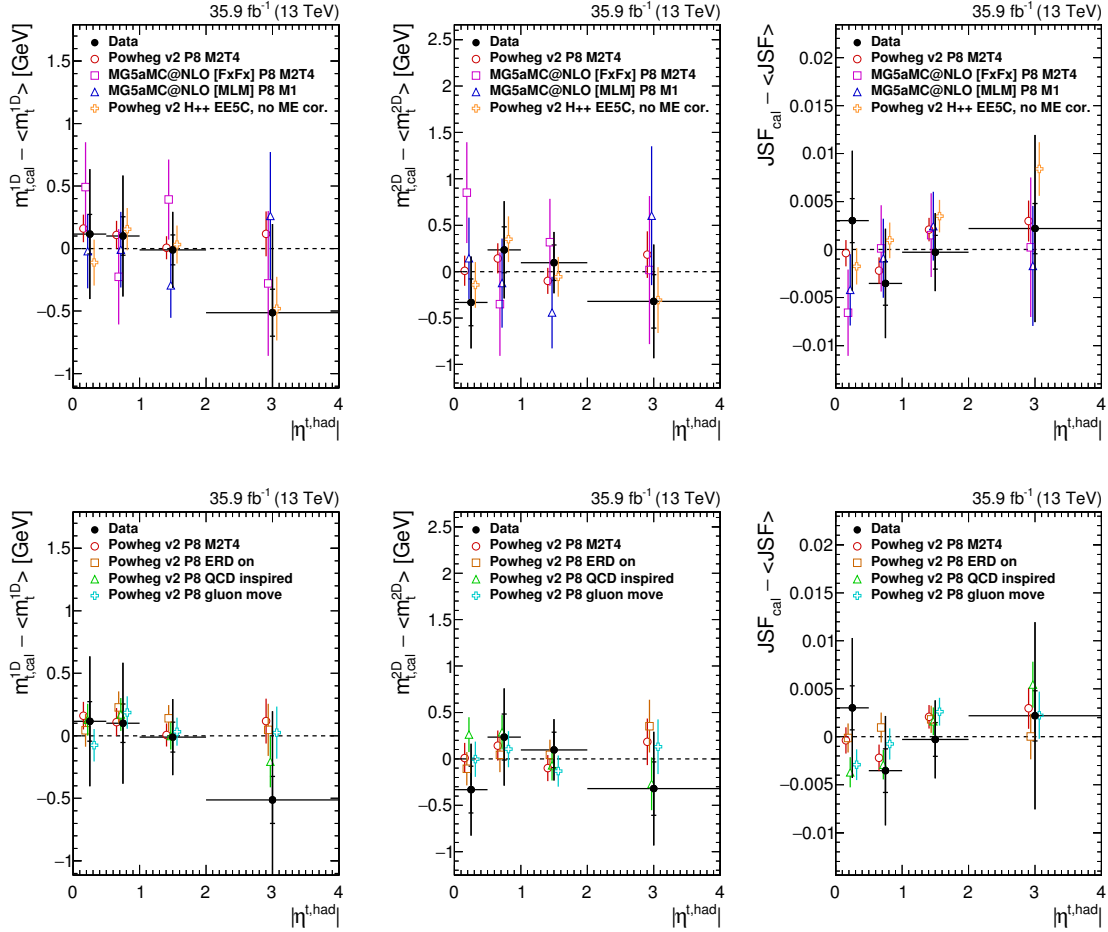


Figure A.34.: Measurements of m_t as a function of the pseudorapidity of the hadronically decaying top quark ($\eta^{t,\text{had}}$) compared to different generator choices (upper row) and to alternative models of color reconnections (lower row).

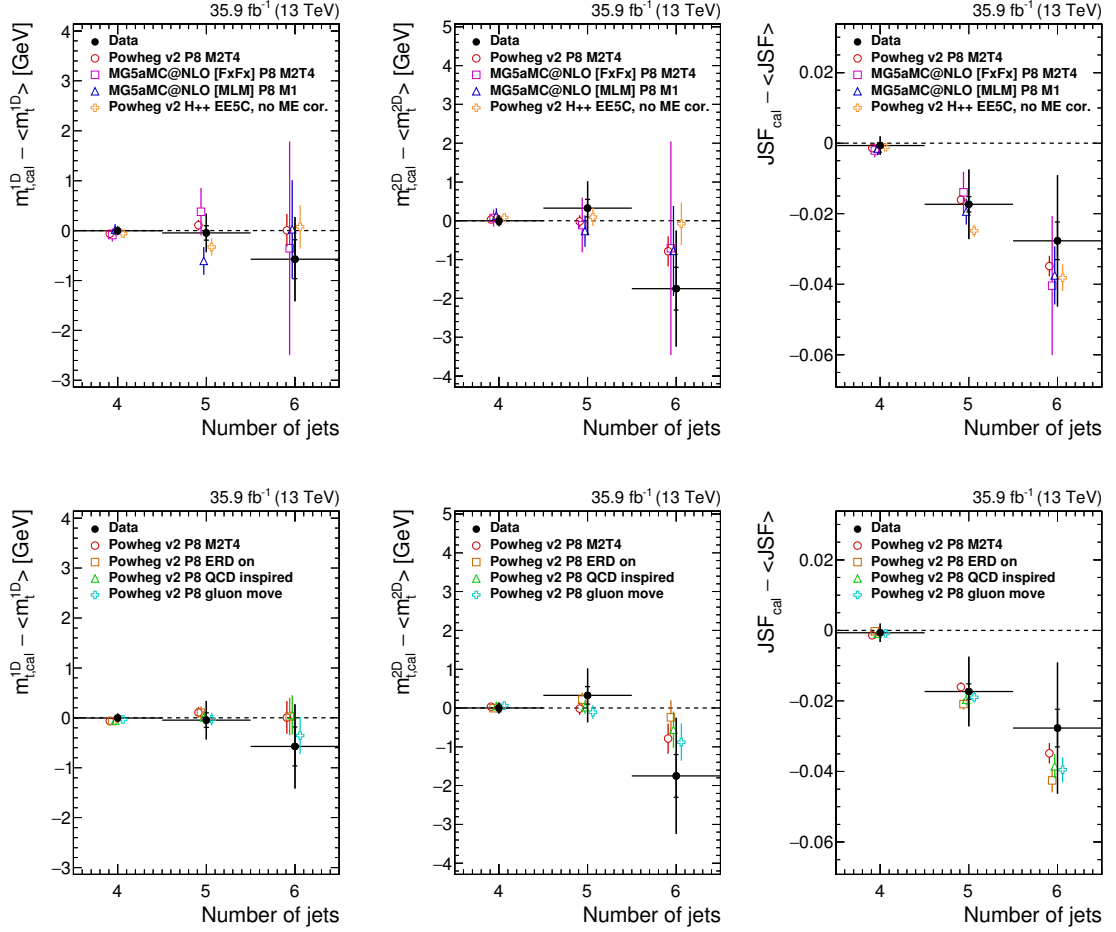


Figure A.35.: Measurements of m_t as a function of the number of jets with $p_T > 30$ GeV compared to different generator choices (upper row) and to alternative models of color reconnections (lower row).

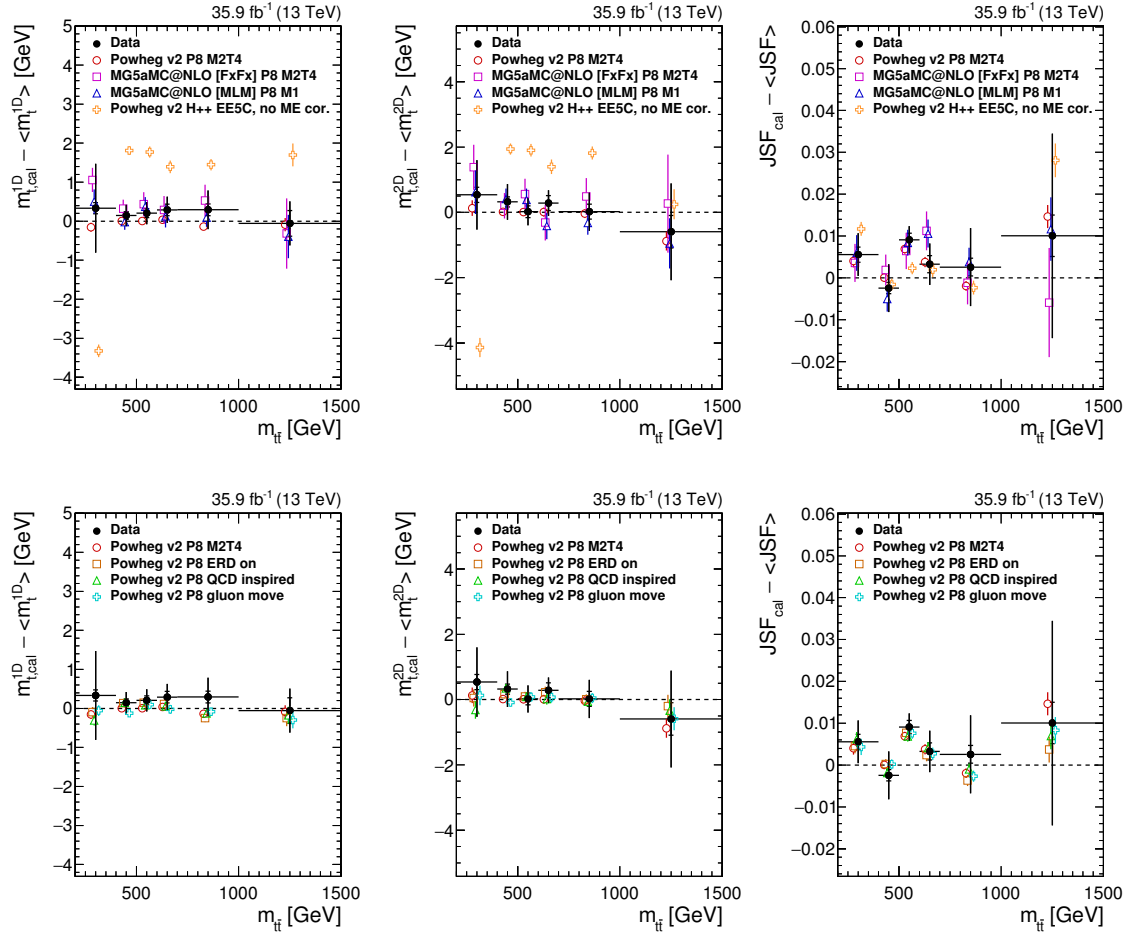


Figure A.36.: Measurements of m_t as a function of the invariant mass of the $t\bar{t}$ system ($m_{t\bar{t}}$) compared to different generator choices (upper row) and to alternative models of color reconnections (lower row).

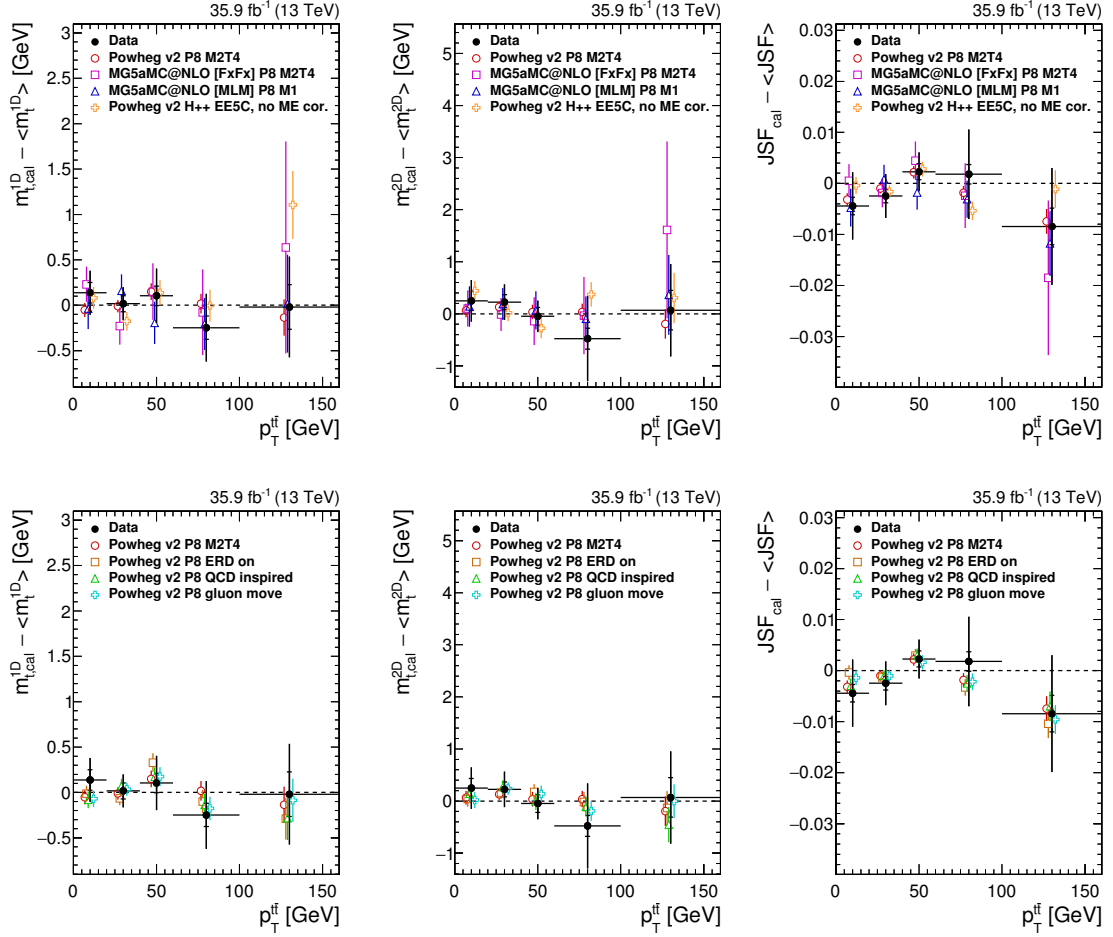


Figure A.37.: Measurements of m_t as a function of the transverse momentum of the $t\bar{t}$ system ($p_T^{t\bar{t}}$) compared to different generator choices (upper row) and to alternative models of color reconnections (lower row).

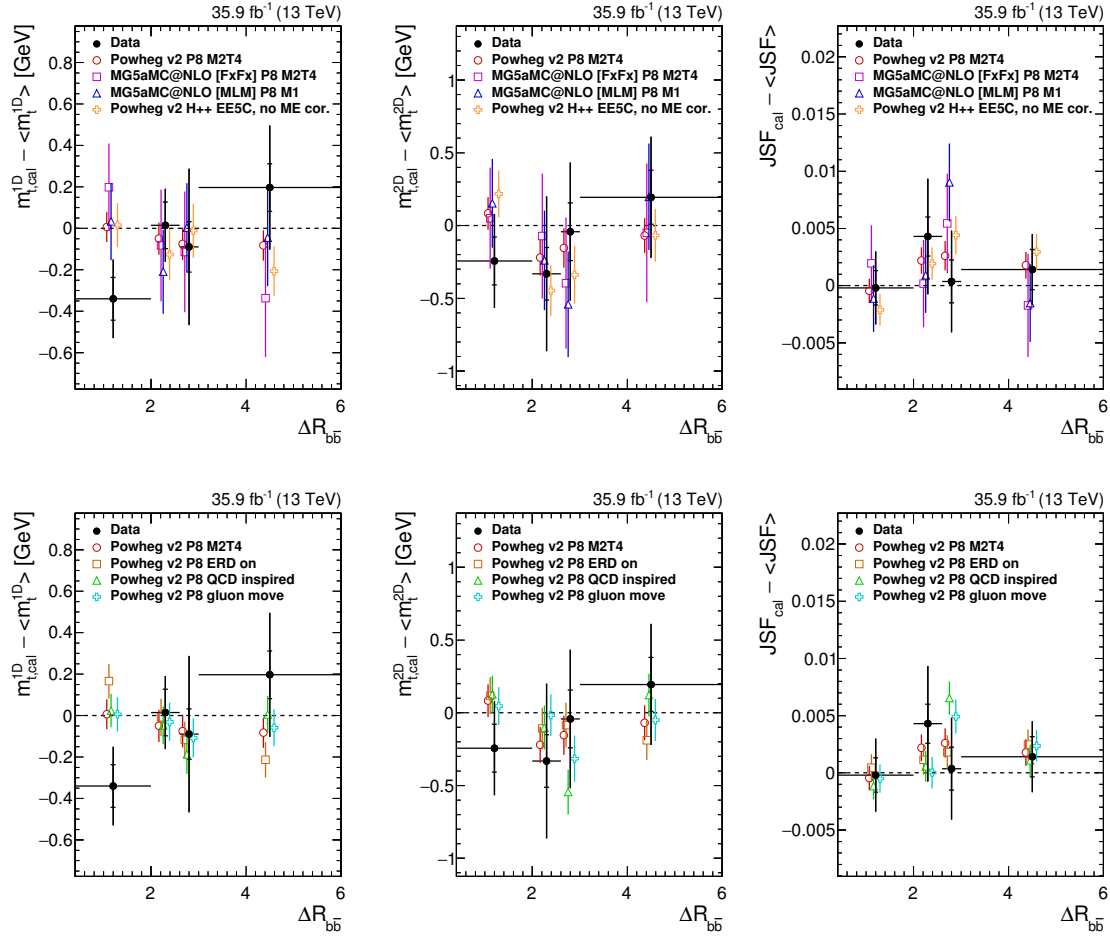


Figure A.38.: Measurements of m_t as a function of the ΔR between the b jets ($\Delta R_{b\bar{b}}$) compared to different generator choices (upper row) and to alternative models of color reconnections (lower row).

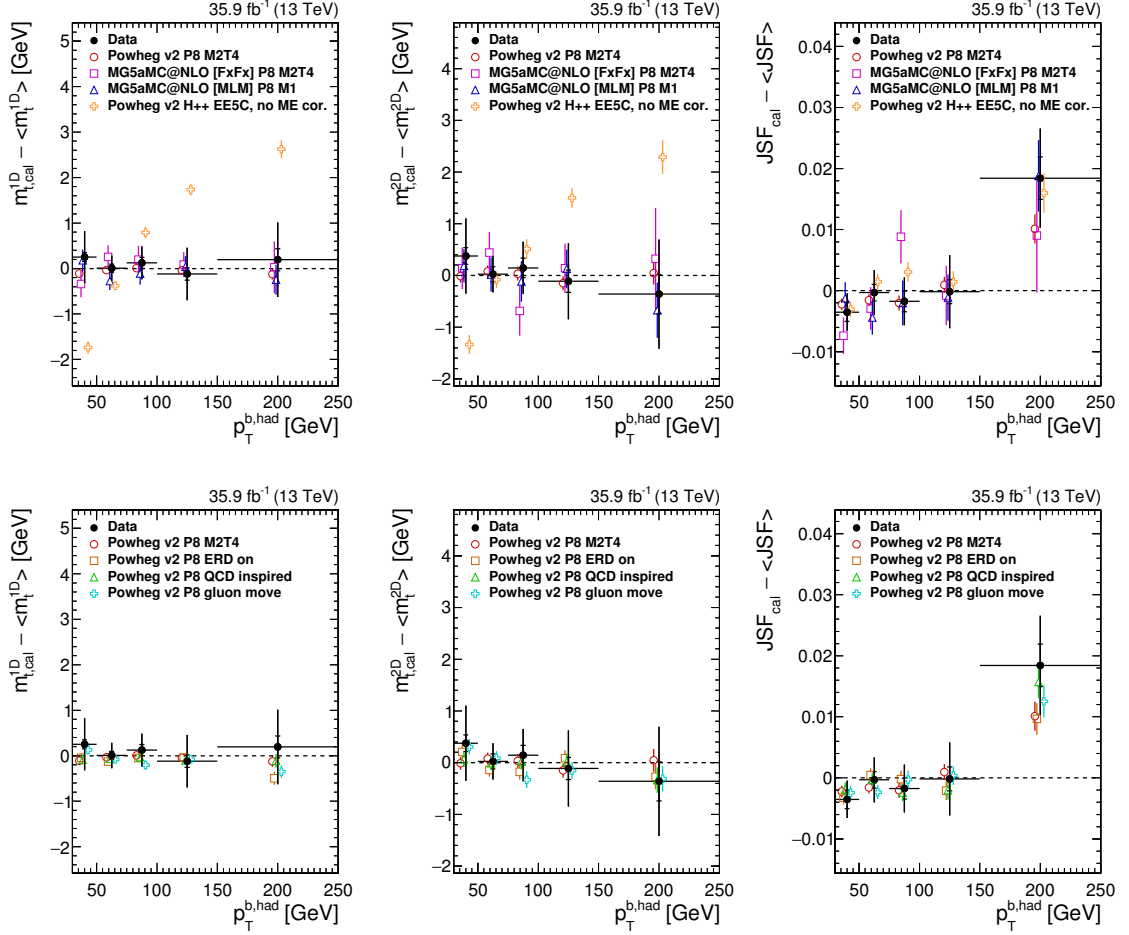


Figure A.39.: Measurements of m_t as a function of the p_T of the b jet assigned to the hadronic decay branch ($p_T^{b, \text{had}}$) compared to different generator choices (upper row) and to alternative models of color reconnections (lower row).

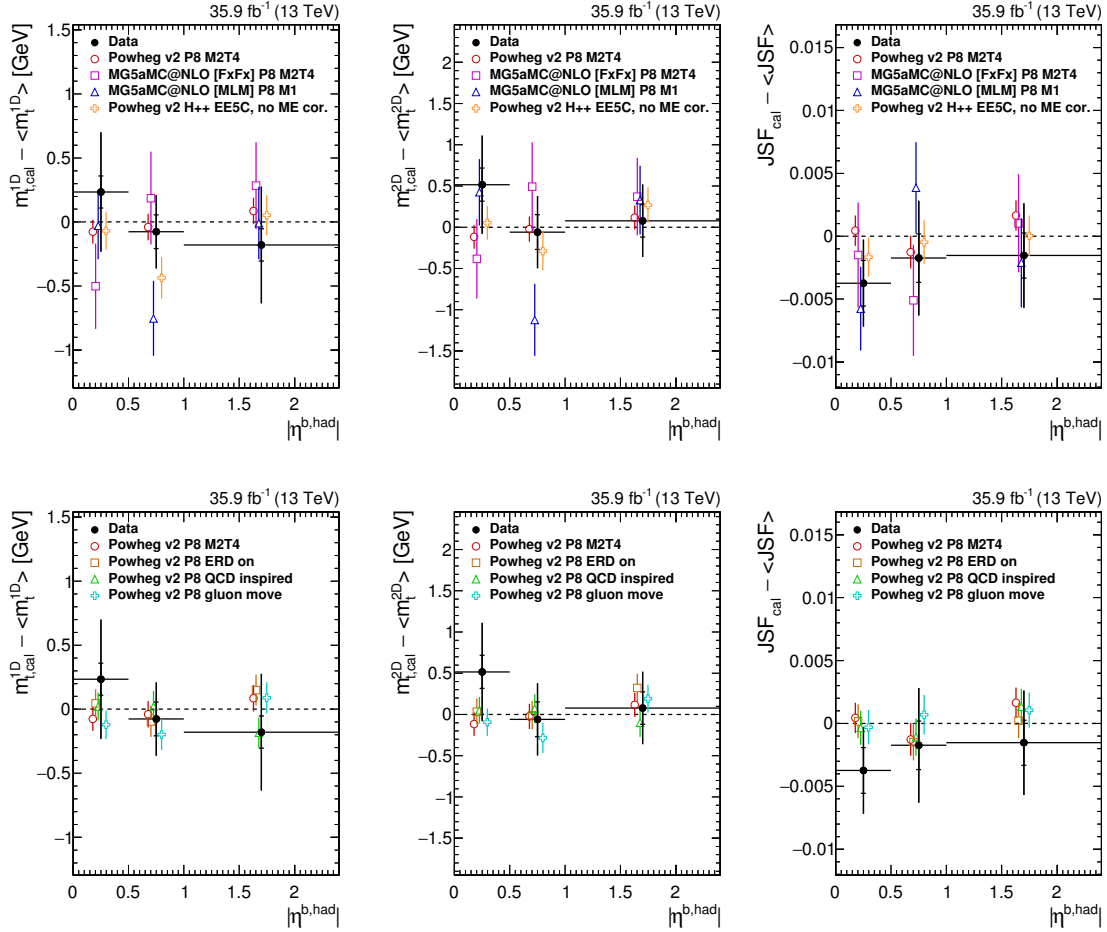


Figure A.40.: Measurements of m_t as a function of the p_T of the pseudorapidity of the b jet assigned to the hadronic decay branch ($|\eta^{b,had}|$) compared to different generator choices (upper row) and to alternative models of color reconnections (lower row).

Bibliography

- [1] CMS Collaboration, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC,” *Phys. Lett. B* 716 (2012) 30–61, [arXiv:1207.7235 \[hep-ex\]](#).
- [2] D0 Collaboration, “Observation of the top quark,” *Phys. Rev. Lett.* 74 (1995) 2632–2637, [arXiv:hep-ex/9503003 \[hep-ex\]](#).
- [3] P. K. Sinervo, “The Discovery of the top quark,” in *Quarks and colliders. Proceedings, 10th Winter Institute, Lake Louise, Canada, February 19-25, 1995*, pp. 316–351. 1995. http://www-cdf.fnal.gov/physics/conf95/cdf3411_top_quark_llwi.ps.
- [4] D0 Collaboration, “Precision measurement of the top-quark mass in lepton+jets final states,” *Phys. Rev. Lett.* 113 (2014) 032002, [arXiv:1405.1756 \[hep-ex\]](#).
- [5] CDF Collaboration, “Precision Top-Quark Mass Measurements at CDF,” *Phys. Rev. Lett.* 109 (2012) 152003, [arXiv:1207.6758 \[hep-ex\]](#).
- [6] CMS Collaboration, “The CMS Experiment at the CERN LHC,” *JINST* 3 (2008) S08004.
- [7] CMS Collaboration, “Measurement of the top quark mass using proton-proton data at $\sqrt{s} = 7$ and 8 TeV,” *Phys. Rev. D* 93 no. 7, (2016) 072004, [arXiv:1509.04044 \[hep-ex\]](#).
- [8] ATLAS Collaboration, “Measurement of the top quark mass in the $t\bar{t}$ lepton+jets channel from $\sqrt{s}=8$ TeV ATLAS data,” ATLAS conference note ATLAS-CONF-2017-071, 2017. <http://cds.cern.ch/record/2285809>.
- [9] ATLAS, CDF, CMS, D0 Collaboration, “First combination of Tevatron and LHC measurements of the top-quark mass,” [arXiv:1403.4427 \[hep-ex\]](#).
- [10] CMS Collaboration, “Measurement of the top quark mass with lepton+jets final states in pp collisions at $\sqrt{s} = 13$ TeV,” CMS Physics Analysis Summary CMS-PAS-TOP-17-007, 2017. <https://cds.cern.ch/record/2284594>.
- [11] CMS Collaboration, “CMS Luminosity Measurements for the 2016 Data Taking Period,” CMS Physics Analysis Summary CMS-PAS-LUM-17-001, 2017. <https://cds.cern.ch/record/2257069>.
- [12] DELPHI Collaboration, “Measurement of the mass and width of the W boson in e^+e^- collisions at $\sqrt{s} = 161 - 209$ GeV,” *Eur. Phys. J. C* 55 (2008) 1, [arXiv:0803.2534 \[hep-ex\]](#).
- [13] CMS Collaboration, “Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV,” *JINST* 12 no. 02, (2017) P02014, [arXiv:1607.03663 \[hep-ex\]](#).
- [14] J. R. Christiansen and P. Z. Skands, “String Formation Beyond Leading Colour,” *JHEP* 08 (2015) 003, [arXiv:1505.01681 \[hep-ph\]](#).
- [15] S. Argyropoulos and T. Sjostrand, “Effects of color reconnection on $t\bar{t}$ final states at the LHC,” *JHEP* 11 (2014) 043, [arXiv:1407.6653 \[hep-ph\]](#).
- [16] T. Sjostrand, S. Mrenna, and P. Z. Skands, “PYTHIA 6.4 Physics and Manual,” *JHEP* 05 (2006) 026, [arXiv:hep-ph/0603175 \[hep-ph\]](#).

- [17] T. Sjostrand, S. Mrenna, and P. Z. Skands, “A Brief Introduction to PYTHIA 8.1,” *Comput. Phys. Commun.* 178 (2008) 852–867, [arXiv:0710.3820 \[hep-ph\]](#).
- [18] Particle Data Group, “Review of Particle Physics,” *Chin. Phys.* C40 no. 10, (2016) 100001.
- [19] Wikipedia, “Standard Model of Elementary Particles,” wikipedia, Accessed: 01/02/2018. https://en.wikipedia.org/wiki/Standard_Model#/media/File:Standard_Model_of_Elementary_Particles.svg.
- [20] 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) 1–29, [arXiv:1207.7214 \[hep-ex\]](#).
- [21] C.-N. Yang and R. L. Mills, “Conservation of Isotopic Spin and Isotopic Gauge Invariance,” *Phys. Rev.* 96 (1954) 191–195.
- [22] F. Jegerlehner, “The Running fine structure constant $\alpha(E)$ via the Adler function,” *Nucl. Phys. Proc. Suppl.* 181-182 (2008) 135–140, [arXiv:0807.4206 \[hep-ph\]](#).
- [23] T. Morii, C. S. Lim, and S. N. Mukherjee, *The physics of the standard model and beyond*. 2004.
- [24] SLAC E158 Collaboration, “Precision measurement of the weak mixing angle in Moller scattering,” *Phys. Rev. Lett.* 95 (2005) 081601, [arXiv:hep-ex/0504049 \[hep-ex\]](#).
- [25] CMS Collaboration, “Measurement of the inclusive 3-jet production differential cross section in protonproton collisions at 7 TeV and determination of the strong coupling constant in the TeV range,” *Eur. Phys. J.* C75 no. 5, (2015) 186, [arXiv:1412.1633 \[hep-ex\]](#).
- [26] F. Englert and R. Brout, “Broken Symmetry and the Mass of Gauge Vector Mesons,” *Phys. Rev. Lett.* 13 (1964) 321–323.
- [27] P. W. Higgs, “Broken Symmetries and the Masses of Gauge Bosons,” *Phys. Rev. Lett.* 13 (1964) 508–509.
- [28] V. Ippolito, *Measurement of the properties of the new particle observed within the search for the Higgs boson in the $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$ channel at ATLAS*. PhD thesis, Rome U., 2013-12-06. <http://inspirehep.net/record/1339804/files/CERN-THESIS-2013-342.pdf>.
- [29] J. Goldstone, A. Salam, and S. Weinberg, “Broken Symmetries,” *Phys. Rev.* 127 (1962) 965–970.
- [30] D. I. Kaiser, “Physics and Feynman’s diagrams,” *Am. Sci.* 93 (2005) 156–165.
- [31] N. Bartosik, *Associated top-quark-pair and b-jet production in the dilepton channel at $\sqrt{s} = 8$ TeV as test of QCD and background $tt + \text{Higgs}$ production*. PhD thesis, U. Hamburg, Dept. Phys., Hamburg, 2015. <http://bib-pubdb1.desy.de/search?cc=Publication+Database&of=hd&p=reportnumber:DESY-THESIS-2015-035>.
- [32] J. C. Collins and D. E. Soper, “The Theorems of Perturbative QCD,” *Ann. Rev. Nucl. Part. Sci.* 37 (1987) 383–409.
- [33] G. ’t Hooft, “Dimensional regularization and the renormalization group,” *Nucl. Phys.* B61 (1973) 455–468.
- [34] S. Weinberg, “New approach to the renormalization group,” *Phys. Rev.* D8 (1973) 3497–3509.

- [35] S. Alekhin, A. Djouadi, and S. Moch, “The top quark and Higgs boson masses and the stability of the electroweak vacuum,” *Phys. Lett.* B716 (2012) 214–219, [arXiv:1207.0980 \[hep-ph\]](#).
- [36] Gfitter Group, “The global electroweak fit at NNLO and prospects for the LHC and ILC,” *Eur. Phys. J.* C74 (2014) 3046, [arXiv:1407.3792 \[hep-ph\]](#).
- [37] G. Degrandi, “The role of the top quark in the stability of the SM Higgs potential,” *Nuovo Cim.* C037 no. 02, (2014) 47–53, [arXiv:1405.6852 \[hep-ph\]](#).
- [38] A. Quadri, “Top quark physics at hadron colliders,” *Eur. Phys. J.* C48 (2006) 835–1000.
- [39] G. Degrandi *et al.*, “Higgs mass and vacuum stability in the Standard Model at NNLO,” *JHEP* 08 (2012) 098, [arXiv:1205.6497 \[hep-ph\]](#).
- [40] F. Bezrukov and M. Shaposhnikov, “Why should we care about the top quark Yukawa coupling?,” *J. Exp. Theor. Phys.* 120 (2015) 335–343, [arXiv:1411.1923 \[hep-ph\]](#). [*Zh. Eksp. Teor. Fiz.*147,389(2015)].
- [41] ARGUS Collaboration, “Observation of B0 - anti-B0 Mixing,” *Phys. Lett.* B192 (1987) 245–252. [*51(1987)*].
- [42] CDF, D0 Collaboration, “Combination of measurements of the top-quark pair production cross section from the Tevatron Collider,” *Phys. Rev.* D89 no. 7, (2014) 072001, [arXiv:1309.7570 \[hep-ex\]](#).
- [43] CMS Collaboration, “Measurement of the t-channel single-top-quark production cross section and of the $|V_{tb}|$ CKM matrix element in pp collisions at $\sqrt{s}=8$ TeV,” *JHEP* 06 (2014) 090, [arXiv:1403.7366 \[hep-ex\]](#).
- [44] ATLAS Collaboration, “Comprehensive measurements of t-channel single top-quark production cross sections at $\sqrt{s}=7$ TeV with the ATLAS detector,” *Phys. Rev.* D90 no. 11, (2014) 112006, [arXiv:1406.7844 \[hep-ex\]](#).
- [45] M. Czakon, P. Fiedler, and A. Mitov, “Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\frac{4}{3})$,” *Phys. Rev. Lett.* 110 (2013) 252004, [arXiv:1303.6254 \[hep-ph\]](#).
- [46] J. M. Campbell, R. Frederix, F. Maltoni, and F. Tramontano, “Next-to-Leading-Order Predictions for t-Channel Single-Top Production at Hadron Colliders,” *Phys. Rev. Lett.* 102 (2009) 182003, [arXiv:0903.0005 \[hep-ph\]](#).
- [47] N. Kidonakis, “Top Quark Production,” in *Proceedings, Helmholtz International Summer School on Physics of Heavy Quarks and Hadrons (HQ 2013): JINR, Dubna, Russia, July 15-28, 2013*, pp. 139–168. 2014. [arXiv:1311.0283 \[hep-ph\]](#).
<http://inspirehep.net/record/1263209/files/arXiv:1311.0283.pdf>.
- [48] CMS Collaboration, “Top quark physics summary figures,” (public), Accessed: 14/02/2018.
<https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsTOPSummaryFigures>.
- [49] M. Jezabek and J. H. Kuhn, “QCD Corrections to Semileptonic Decays of Heavy Quarks,” *Nucl. Phys.* B314 (1989) 1–6.
- [50] P. Marquard *et al.*, “Quark Mass Relations to Four-Loop Order in Perturbative QCD,” *Phys. Rev. Lett.* 114 (2015) 142002, [arXiv:1502.01030 \[hep-ph\]](#).
- [51] M. Beneke, “Renormalons,” *Phys. Rept.* 317 (1999) 1–142, [arXiv:hep-ph/9807443 \[hep-ph\]](#).

- [52] P. Nason, “The Top Mass in Hadronic Collisions,” 2017. [arXiv:1712.02796](https://arxiv.org/abs/1712.02796) [hep-ph].
<http://inspirehep.net/record/1642228/files/arXiv:1712.02796.pdf>.
- [53] L. Evans and P. Bryant, “LHC Machine,” *JINST* 3 (2008) S08001.
- [54] ATLAS Collaboration, “The ATLAS Experiment at the CERN Large Hadron Collider,” *JINST* 3 (2008) S08003.
- [55] ALICE Collaboration, “The ALICE experiment at the CERN LHC,” *JINST* 3 (2008) S08002.
- [56] LHCb Collaboration, “The LHCb Detector at the LHC,” *JINST* 3 (2008) S08005.
- [57] F. Marcastel, “CERN’s Accelerator Complex. La chane des accclateurs du CERN,”
<https://cds.cern.ch/record/1621583>. (general photo).
- [58] O. S. Bruning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock,
“LHC Design Report Vol.1: The LHC Main Ring,” *Technical Design Report* (2004) .
<http://cds.cern.ch/record/782076>.
- [59] CMS Collaboration, “Data quality information,” (public), 2017.
<https://twiki.cern.ch/twiki/bin/view/CMSPublic/DataQuality>.
- [60] CMS Collaboration, “The CMS detector,” (public), 2017.
<https://home.cern/about/experiments/cms>.
- [61] T. Sakuma and T. McCauley, “Detector and event visualization with sketchup at the cms
experiment,” *Journal of Physics: Conference Series* 513 no. 2, (2014) 022032.
<http://stacks.iop.org/1742-6596/513/i=2/a=022032>.
- [62] CMS Collaboration, “The CMS electromagnetic calorimeter project,” *Technical Design
Report* (1997) . <http://cds.cern.ch/record/349375>.
- [63] CMS Collaboration, *Technical Design Report Volume 1: Detector Performance and
Software*. Technical Design Report CMS. CERN, Geneva, 2006.
<https://cds.cern.ch/record/922757>.
- [64] CMS Collaboration, “Energy Calibration and Resolution of the CMS Electromagnetic
Calorimeter in pp Collisions at $\sqrt{s} = 7$ TeV,” *JINST* 8 (2013) P09009, [arXiv:1306.2016](https://arxiv.org/abs/1306.2016)
[hep-ex].
- [65] CMS Collaboration, “The CMS hadron calorimeter project,” *Technical Design Report*
(1997) . <http://cds.cern.ch/record/1614058>.
- [66] CMS Collaboration, “Performance of the CMS Hadron Calorimeter with Cosmic Ray
Muons and LHC Beam Data,” *JINST* 5 (2010) T03012, [arXiv:0911.4991](https://arxiv.org/abs/0911.4991)
[physics.ins-det].
- [67] CMS Collaboration, “The CMS barrel calorimeter response to particle beams from 2-GeV/c
to 350-GeV/c,” *Eur. Phys. J.* C60 (2009) 359–373. [Erratum: *Eur. Phys. J.*C61,353(2009)].
- [68] CMS Collaboration, “Precise Mapping of the Magnetic Field in the CMS Barrel Yoke using
Cosmic Rays,” *JINST* 5 (2010) T03021, [arXiv:0910.5530](https://arxiv.org/abs/0910.5530) [physics.ins-det].
- [69] C. Joram, “Particle Identification,” CERN Summer Student Lectures 2003, Accessed:
01/12/2017.
https://indico.cern.ch/event/414005/attachments/844684/1175186/SSL_PD_V.pdf.
- [70] CMS Collaboration, “CSC local efficiency measurements from run II data,” CMS Detector
Performance Note CMS-DP-2015-052, 2015. <https://cds.cern.ch/record/2104473>.

- [71] CMS Collaboration, “Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV,” *JINST* 7 (2012) P10002, [arXiv:1206.4071 \[physics.ins-det\]](#).
- [72] CMS Collaboration, “CMS Technical Design Report for the Level-1 Trigger Upgrade,” *Technical Design Report* (2013) . <http://cds.cern.ch/record/1261310>.
- [73] M. A. Dobbs *et al.*, “Les Houches guidebook to Monte Carlo generators for hadron collider physics,” [arXiv:hep-ph/0403045 \[hep-ph\]](#). Compiled by the Working Group on Quantum Chromodynamics and the Standard Model.
- [74] S. Drell and T.-M. Yan, “Partons and their Applications at High-Energies,” *Annals Phys.* 66 (1971) 578.
- [75] J. M. Campbell and R. K. Ellis, “MCFM for the Tevatron and the LHC,” *Nucl. Phys. Proc. Suppl.* 205-206 (2010) 10–15, [arXiv:1007.3492 \[hep-ph\]](#).
- [76] R. P. Feynman, “Very high-energy collisions of hadrons,” *Phys. Rev. Lett.* 23 (1969) 1415–1417.
- [77] S. Forte, “Parton distributions at the dawn of the LHC,” *Acta Phys. Polon.* B41 (2010) 2859–2920, [arXiv:1011.5247 \[hep-ph\]](#).
- [78] P. Skands, “Introduction to QCD,” in *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: Searching for New Physics at Small and Large Scales (TASI 2012): Boulder, Colorado, June 4-29, 2012*, pp. 341–420. 2013. [arXiv:1207.2389 \[hep-ph\]](#).
- [79] NNPDF Collaboration, “Parton distributions from high-precision collider data,” *Eur. Phys. J.* C77 no. 10, (2017) 663, [arXiv:1706.00428 \[hep-ph\]](#).
- [80] Y. L. Dokshitzer, “Calculation of the Structure Functions for Deep Inelastic Scattering and e^+e^- Annihilation by Perturbation Theory in Quantum Chromodynamics,” *Sov. Phys. JETP* 46 (1977) 641–653.
- [81] V. N. Gribov and L. N. Lipatov, “Deep inelastic ep scattering in perturbation theory,” *Sov. J. Nucl. Phys.* 15 (1972) 438–450.
- [82] L. N. Lipatov, “The parton model and perturbation theory,” *Sov. J. Nucl. Phys.* 20 (1975) 94–102. [*Yad. Fiz.*20,181(1974)].
- [83] NNPDF Collaboration, “Parton distributions for the LHC Run II,” *JHEP* 04 (2015) 040, [arXiv:1410.8849 \[hep-ph\]](#).
- [84] J. Stirling, “Parton luminosity and cross section plots,” Accessed: 01/01/2018. <http://www.hep.ph.ic.ac.uk/~wstirling/plots/plots.html>.
- [85] H.-L. Lai *et al.*, “Uncertainty induced by QCD coupling in the CTEQ global analysis of parton distributions,” *Phys. Rev.* D82 (2010) 054021, [arXiv:1004.4624 \[hep-ph\]](#).
- [86] R. S. Thorne, A. D. Martin, W. J. Stirling, and G. Watt, “The Effects of combined HERA and recent Tevatron $W \rightarrow l \nu$ charge asymmetry data on the MSTW PDFs,” *PoS DIS2010* (2010) 052, [arXiv:1006.2753 \[hep-ph\]](#).
- [87] A. M. Cooper-Sarkar, “HERAPDF fits including $F_2(\text{charm})$ data,” *PoS DIS2010* (2010) 023, [arXiv:1006.4471 \[hep-ph\]](#).
- [88] S. Alekhin, J. Blumlein, S. Klein, and S. Moch, “Variable-Flavor-Number Scheme in Analysis of Heavy-Quark Electro-Production Data,” in *Proceedings, 17th International Workshop on Deep-Inelastic Scattering and Related Subjects (DIS 2009): Madrid, Spain, April 26-30, 2009*. 2009. [arXiv:0908.3128 \[hep-ph\]](#).

- [89] P. Jimenez-Delgado and E. Reya, “Variable Flavor Number Parton Distributions and Weak Gauge and Higgs Boson Production at Hadron Colliders at NNLO of QCD,” *Phys. Rev. D* 80 (2009) 114011, [arXiv:0909.1711 \[hep-ph\]](#).
- [90] L. W. Whitlow *et al.*, “Precise measurements of the proton and deuteron structure functions from a global analysis of the SLAC deep inelastic electron scattering cross-sections,” *Phys. Lett. B* 282 (1992) 475–482.
- [91] ZEUS, H1 Collaboration, “Combined Measurement and QCD Analysis of the Inclusive $e+p$ Scattering Cross Sections at HERA,” *JHEP* 01 (2010) 109, [arXiv:0911.0884 \[hep-ex\]](#).
- [92] CDF Collaboration, “Direct Measurement of the W Production Charge Asymmetry in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV,” *Phys. Rev. Lett.* 102 (2009) 181801, [arXiv:0901.2169 \[hep-ex\]](#).
- [93] CMS Collaboration, “Measurement of the electron charge asymmetry in inclusive W production in pp collisions at $\sqrt{s} = 7$ TeV,” *Phys. Rev. Lett.* 109 (2012) 111806, [arXiv:1206.2598 \[hep-ex\]](#).
- [94] SM and NLO Multileg Working Group Collaboration, T. Binoth *et al.*, “The SM and NLO Multileg Working Group: Summary report,” in *Physics at TeV colliders. Proceedings, 6th Workshop, dedicated to Thomas Binoth, Les Houches, France, June 8-26, 2009*, pp. 21–189. 2010. [arXiv:1003.1241 \[hep-ph\]](#).
- [95] J. Alwall *et al.*, “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations,” *JHEP* 07 (2014) 079, [arXiv:1405.0301 \[hep-ph\]](#).
- [96] P. Nason, “A new method for combining NLO QCD with shower Monte Carlo algorithms,” *JHEP* 11 (2004) 040, [arXiv:hep-ph/0409146 \[hep-ph\]](#).
- [97] S. Frixione, P. Nason, and C. Oleari, “Matching NLO QCD computations with Parton Shower simulations: the POWHEG method,” *JHEP* 11 (2007) 070, [arXiv:0709.2092 \[hep-ph\]](#).
- [98] S. Alioli, P. Nason, C. Oleari, and E. Re, “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX,” *JHEP* 06 (2010) 043, [arXiv:1002.2581 \[hep-ph\]](#).
- [99] J. M. Campbell, R. K. Ellis, P. Nason, and E. Re, “Top-pair production and decay at NLO matched with parton showers,” *JHEP* 04 (2015) 114, [arXiv:1412.1828 \[hep-ph\]](#).
- [100] V. V. Sudakov, “Vertex parts at very high-energies in quantum electrodynamics,” *Sov. Phys. JETP* 3 (1956) 65–71. [*Zh. Eksp. Teor. Fiz.* 30,87(1956)].
- [101] J. C. Collins, “Sudakov form-factors,” *Adv. Ser. Direct. High Energy Phys.* 5 (1989) 573–614, [arXiv:hep-ph/0312336 \[hep-ph\]](#).
- [102] M. Bahr *et al.*, “Herwig++ physics and manual,” *Eur. Phys. J. C* 58 (2008) 639, [arXiv:0803.0883 \[hep-ph\]](#).
- [103] J. Alwall *et al.*, “Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions,” *Eur. Phys. J. C* 53 (2008) 473, [arXiv:0706.2569 \[hep-ph\]](#).
- [104] R. Frederix and S. Frixione, “Merging meets matching in MC@NLO,” *JHEP* 12 (2012) 061, [arXiv:1209.6215 \[hep-ph\]](#).

- [105] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, “Parton Fragmentation and String Dynamics,” *Phys. Rept.* 97 (1983) 31–145.
- [106] D. Amati and G. Veneziano, “Preconfinement as a Property of Perturbative QCD,” *Phys. Lett.* 83B (1979) 87–92.
- [107] A. Bassetto, M. Ciafaloni, and G. Marchesini, “Color Singlet Distributions and Mass Damping in Perturbative QCD,” *Phys. Lett.* 83B (1979) 207–212.
- [108] S. Hoeche, “Introduction to parton-shower event generators,” in *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders (TASI 2014): Boulder, Colorado, June 2-27, 2014*, pp. 235–295. 2015. [arXiv:1411.4085 \[hep-ph\]](#).
- [109] A. Buckley *et al.*, “General-purpose event generators for LHC physics,” *Phys. Rept.* 504 (2011) 145–233, [arXiv:1101.2599 \[hep-ph\]](#).
- [110] CMS Collaboration, “Investigations of the impact of the parton shower tuning in Pythia 8 in the modelling of $t\bar{t}$ at $\sqrt{s} = 8$ and 13 TeV,” CMS Physics Analysis Summary CMS-PAS-TOP-16-021, 2016. <https://cds.cern.ch/record/2235192>.
- [111] M. H. Seymour and A. Siodmok, “Constraining MPI models using σ_{eff} and recent Tevatron and LHC Underlying Event data,” *JHEP* 10 (2013) 113, [arXiv:1307.5015 \[hep-ph\]](#).
- [112] M. Sandhoff and P. Z. Skands, “Colour annealing - a toy model of colour reconnections,” in *Physics at TeV colliders. Proceedings, Workshop, Les Houches, France, May 2-20, 2005*. 2005. http://lss.fnal.gov/cgi-bin/find_paper.pl?conf-05-518-t.
- [113] ATLAS Collaboration, “Charged-particle multiplicities in pp interactions measured with the ATLAS detector at the LHC,” *New J. Phys.* 13 (2011) 053033, [arXiv:1012.5104 \[hep-ex\]](#).
- [114] CDF Collaboration, “Measurement of Particle Production and Inclusive Differential Cross Sections in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ -TeV,” *Phys. Rev.* D79 (2009) 112005, [arXiv:0904.1098 \[hep-ex\]](#). [Erratum: *Phys. Rev.*D82,119903(2010)].
- [115] J. Alwall *et al.*, “MadGraph 5 : Going Beyond,” *JHEP* 06 (2011) 128, [arXiv:1106.0522 \[hep-ph\]](#).
- [116] P. Skands, S. Carrazza, and J. Rojo, “Tuning PYTHIA 8.1: the Monash 2013 Tune,” *Eur. Phys. J. C* 74 (2014) 3024, [arXiv:1404.5630 \[hep-ph\]](#).
- [117] CMS Collaboration, “Investigations of the impact of the parton shower tuning in Pythia 8 in the modelling of $t\bar{t}$ at $\sqrt{s} = 8$ and 13 TeV,” CMS Physics Analysis Summary CMS-PAS-TOP-16-021, 2016. <https://cds.cern.ch/record/2235192>.
- [118] GEANT4 Collaboration, “GEANT4: A Simulation toolkit,” *Nucl. Instrum. Meth.* A506 (2003) 250–303.
- [119] J. Allison *et al.*, “Geant4 developments and applications,” *IEEE Trans. Nucl. Sci.* 53 (2006) 270.
- [120] J. Allison *et al.*, “Recent developments in G_{EANT4} ,” *Nucl. Instrum. Meth.* A835 (2016) 186–225.
- [121] M. Seidel, *Precise measurement of the top-quark mass at the CMS experiment using the ideogram method*. PhD thesis, U. Hamburg, Dept. Phys., Hamburg, 2015. <https://bib-pubdb1.desy.de/search?cc=Publication+Database&of=hd&p=reportnumber:DESY-THESIS-2015-034>.

- [122] CMS Collaboration, “Event generator tunes obtained from underlying event and multiparton scattering measurements,” *Eur. Phys. J. C* 76 no. 3, (2016) 155, [arXiv:1512.00815 \[hep-ex\]](#).
- [123] T. Sjostrand and M. van Zijl, “A Multiple Interaction Model for the Event Structure in Hadron Collisions,” *Phys. Rev. D* 36 (1987) 2019.
- [124] T. Sjostrand and P. Z. Skands, “Multiple interactions and the structure of beam remnants,” *JHEP* 03 (2004) 053, [arXiv:hep-ph/0402078 \[hep-ph\]](#).
- [125] J. R. Christiansen and T. Sjostrand, “Color reconnection at future e^+e^- colliders,” *Eur. Phys. J. C* 75 no. 9, (2015) 441, [arXiv:1506.09085 \[hep-ph\]](#).
- [126] CMS Collaboration, “Underlying Event Measurements with Leading Particles and Jets in pp collisions at $\sqrt{s} = 13$ TeV,” CMS Physics Analysis Summary CMS-PAS-FSQ-15-007, 2015. <https://cds.cern.ch/record/2104473>.
- [127] ATLAS Collaboration, “Measurement of underlying event characteristics using charged particles in pp collisions at $\sqrt{s} = 900\text{GeV}$ and 7 TeV with the ATLAS detector,” *Phys. Rev. D* 83 (2011) 112001, [arXiv:1012.0791 \[hep-ex\]](#).
- [128] ATLAS Collaboration, “Measurement of charged-particle distributions sensitive to the underlying event in $\sqrt{s} = 13$ TeV proton-proton collisions with the ATLAS detector at the LHC,” *JHEP* 03 (2017) 157, [arXiv:1701.05390 \[hep-ex\]](#).
- [129] CDF Collaboration, “Study of the energy dependence of the underlying event in proton-antiproton collisions,” *Phys. Rev. D* 92 no. 9, (2015) 092009, [arXiv:1508.05340 \[hep-ex\]](#).
- [130] CMS Collaboration, “Pseudorapidity distribution of charged hadrons in proton-proton collisions at $\sqrt{s} = 13$ TeV,” *Phys. Lett. B* 751 (2015) 143–163, [arXiv:1507.05915 \[hep-ex\]](#).
- [131] TASSO Collaboration, “Determination of α_s in First and Second Order QCD From e^+e^- Annihilation Into Hadrons,” *Z. Phys. C* 26 (1984) 157.
- [132] A. Buckley *et al.*, “Systematic event generator tuning for the LHC,” *Eur. Phys. J. C* 65 (2010) 331–357, [arXiv:0907.2973 \[hep-ph\]](#).
- [133] B. M. Waugh *et al.*, “HZTool and Rivet: Toolkit and Framework for the Comparison of Simulated Final States and Data at Colliders,” in *15th International Conference on Computing in High Energy and Nuclear Physics (CHEP 2006) Mumbai, Maharashtra, India, February 13-17, 2006*. 2006. [arXiv:hep-ph/0605034 \[hep-ph\]](#).
- [134] A. Buckley *et al.*, “Rivet user manual,” *Comput. Phys. Commun.* 184 (2013) 2803–2819, [arXiv:1003.0694 \[hep-ph\]](#).
- [135] CMS Collaboration, “Measurement of pseudorapidity distributions of charged particles in proton-proton collisions at $\sqrt{s} = 13$ TeV by the CMS experiment,” CMS Physics Analysis Summary CMS-PAS-FSQ-15-008, 2016. <https://cds.cern.ch/record/2145373>.
- [136] CMS Collaboration, “Measurement of the pseudorapidity dependence of the energy and transverse energy density in pp collisions at $\sqrt{s} = 13$ TeV with CMS,” CMS Physics Analysis Summary CMS-PAS-FSQ-15-006, 2016. <https://cds.cern.ch/record/2146007>.
- [137] CMS Collaboration, “Pseudorapidity distributions of charged particles in pp collisions at $\sqrt{s} = 7$ TeV with at least one central charged particles,” CMS Physics Analysis Summary CMS-PAS-QCD-10-024, 2011. <https://cds.cern.ch/record/1341853>.

- [138] CMS Collaboration, “Measurement of neutral strange particle production in the underlying event in proton-proton collisions at $\sqrt{s} = 7$ TeV,” *Phys. Rev. D* **88** (2013) 052001, [arXiv:1305.6016 \[hep-ex\]](#).
- [139] C. Bierlich, G. Gustafson, and L. Lnnblad, “A shoving model for collectivity in hadronic collisions,” [arXiv:1612.05132 \[hep-ph\]](#).
- [140] C. Bierlich, “Rope Hadronization and Strange Particle Production,” 2017. [arXiv:1710.04464 \[nucl-th\]](#).
- [141] CMS Collaboration, “Jet and underlying event properties as a function of charged-particle multiplicity in protonproton collisions at $\sqrt{s} = 7$ TeV,” *Eur. Phys. J. C* **73** no. 12, (2013) 2674, [arXiv:1310.4554 \[hep-ex\]](#).
- [142] CMS Collaboration, “Particle-flow reconstruction and global event description with the CMS detector,” *JINST* **12** no. 10, (2017) P10003, [arXiv:1706.04965 \[physics.ins-det\]](#).
- [143] C. Lippmann, “Particle identification,” *Nucl. Instrum. Meth. A* **666** (2012) 148–172, [arXiv:1101.3276 \[hep-ex\]](#).
- [144] R. Fruhwirth, “Application of Kalman filtering to track and vertex fitting,” *Nucl. Instrum. Meth. A* **262** (1987) 444–450.
- [145] W. Adam, R. Frhwirth, A. Strandlie, and T. Todor, “Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC,” CMS Note CMS-NOTE-2005-001, 2005. <https://cds.cern.ch/record/815410>.
- [146] CMS Collaboration, “Description and performance of track and primary-vertex reconstruction with the CMS tracker,” *JINST* **9** no. 10, (2014) P10009, [arXiv:1405.6569 \[physics.ins-det\]](#).
- [147] K. Rose, “Deterministic annealing for clustering, compression, classification, regression, and related optimization problems,” in *Proceedings of the IEEE*. Nov 1998. <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=726788>.
- [148] W. Waltenberger, “Adaptive Vertex Reconstruction,” CMS Note CMS-NOTE-2008-033, 2008. <http://cds.cern.ch/record/1166320>.
- [149] R. Fruhwirth, W. Waltenberger, and P. Vanlaer, “Adaptive vertex fitting,” *J. Phys. G* **34** (2007) N343.
- [150] E. Chabanat and N. Estre, “Deterministic annealing for vertex finding at CMS,” in *Computing in high energy physics and nuclear physics. Proceedings, Conference, CHEP’04, Interlaken, Switzerland, September 27-October 1, 2004*, pp. 287–290. 2005. <http://doc.cern.ch/yellowrep/2005/2005-002/p287.pdf>.
- [151] H. Kirschenmann, “Jet performance in CMS,” CMS Physics Analysis Summary CMS-CR-2013-325, 2013. <https://cds.cern.ch/record/1627818>.
- [152] CMS Collaboration, “Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV,” *JINST* **7** (2012) P10002, [arXiv:1206.4071 \[physics.ins-det\]](#).
- [153] CMS Collaboration, “Muon Identification and Isolation efficiency on full 2016 dataset,” CMS Detector Performance Note CMS-DP-2017-007, 2017. <https://cds.cern.ch/record/2257968>.
- [154] CMS Collaboration, “Reference muon ID, isolation and trigger efficiencies for Run 2,” (internal), Accessed: 01/12/2017. https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceEffsRun2#Results_for_CMSSW_7_6_X_dataset.

- [155] CMS Collaboration, “Summary performance Plots for 2016 data,” (public), Accessed: 01/12/2017.
<https://twiki.cern.ch/twiki/bin/view/CMSPublic/HLTplotsSummary2016>.
- [156] S. Baffioni *et al.*, “Electron reconstruction in CMS,” *Eur. Phys. J.* C49 (2007) 1099–1116.
- [157] CMS Collaboration, “PF based isolation,” (internal), Accessed: 01/01/2018.
<https://twiki.cern.ch/twiki/bin/view/CMS/EgammaPFBasedIsolationRun2>.
- [158] CMS Collaboration, “Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $s = 8$ TeV,” *JINST* 10 no. 06, (2015) P06005, [arXiv:1502.02701](https://arxiv.org/abs/1502.02701) [physics.ins-det].
- [159] CMS Collaboration, “Lepton ID, Isolation and Trigger efficiencies and scale factor,” (internal), Accessed: 01/01/2018. <https://twiki.cern.ch/twiki/bin/view/CMS/HWW2016TriggerAndIdIsoScaleFactorsResults>.
- [160] CMS Collaboration, “Top Trigger Run 2,” (internal), Accessed: 01/01/2018.
<https://twiki.cern.ch/twiki/bin/view/CMS/TopTrigger>.
- [161] G. P. Salam, “Towards Jetography,” *Eur.Phys.J.* C67 (2010) 637–686, [arXiv:0906.1833](https://arxiv.org/abs/0906.1833) [hep-ph].
- [162] CMS Collaboration, “Identification of b-quark jets with the CMS experiment,” *JINST* 8 (2013) P04013, [arXiv:1211.4462](https://arxiv.org/abs/1211.4462) [hep-ex].
- [163] UA1 Collaboration, “Hadronic Jet Production at the CERN Proton - anti-Proton Collider,” *Phys. Lett.* 132B (1983) 214.
- [164] G. C. Blazey *et al.*, “Run II jet physics,” in *QCD and weak boson physics in Run II. Proceedings, Batavia, USA, March 4-6, June 3-4, November 4-6, 1999*, pp. 47–77. 2000. [arXiv:hep-ex/0005012](https://arxiv.org/abs/hep-ex/0005012) [hep-ex].
- [165] G. P. Salam and G. Soyez, “A Practical Seedless Infrared-Safe Cone jet algorithm,” *JHEP* 0705 (2007) 086, [arXiv:0704.0292](https://arxiv.org/abs/0704.0292) [hep-ph].
- [166] S. D. Ellis and D. E. Soper, “Successive combination jet algorithm for hadron collisions,” *Phys. Rev.* D48 (1993) 3160–3166, [arXiv:hep-ph/9305266](https://arxiv.org/abs/hep-ph/9305266) [hep-ph].
- [167] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- k_t jet clustering algorithm,” *JHEP* 04 (2008) 063, [arXiv:0802.1189](https://arxiv.org/abs/0802.1189) [hep-ex].
- [168] Y. L. Dokshitzer, G. Leder, S. Moretti, and B. Webber, “Better jet clustering algorithms,” *JHEP* 9708 (1997) 001, [arXiv:hep-ph/9707323](https://arxiv.org/abs/hep-ph/9707323) [hep-ph].
- [169] M. Wobisch and T. Wengler, “Hadronization corrections to jet cross-sections in deep inelastic scattering,” in *Monte Carlo generators for HERA physics. Proceedings, Workshop, Hamburg, Germany, 1998-1999*, pp. 270–279. 1998. [arXiv:hep-ph/9907280](https://arxiv.org/abs/hep-ph/9907280) [hep-ph].
- [170] G. Corcella, I. Knowles, G. Marchesini, S. Moretti, K. Odagiri, *et al.*, “HERWIG 6.5 release note,” [arXiv:hep-ph/0210213](https://arxiv.org/abs/hep-ph/0210213) [hep-ph].
- [171] CMS Collaboration, “Jet Identification for the 13 TeV data Run2016,” (internal), Accessed: 01/12/2017. <https://twiki.cern.ch/twiki/bin/view/CMS/JetID13TeVRun2016>.
- [172] D. Bertolini, P. Harris, M. Low, and N. Tran, “Pileup Per Particle Identification,” *JHEP* 10 (2014) 059, [arXiv:1407.6013](https://arxiv.org/abs/1407.6013) [hep-ph].
- [173] M. Cacciari, G. P. Salam, and G. Soyez, “FastJet user manual,” [arXiv:1111.6097](https://arxiv.org/abs/1111.6097) [hep-ph].

- [174] M. Cacciari and G. P. Salam, “Dispelling the N^3 myth for the k_t jet-finder,” *Phys.Lett.* B641 (2006) 57–61, [arXiv:hep-ph/0512210](#) [hep-ph].
- [175] M. Cacciari and G. P. Salam, “Pileup subtraction using jet areas,” *Phys. Lett.* B659 (2008) 119–126, [arXiv:0707.1378](#) [hep-ph].
- [176] CMS Collaboration, “Determination of jet energy calibration and transverse momentum resolution in CMS,” *JINST* 6 (2011) P11002, [arXiv:1107.4277](#) [physics.ins-det].
- [177] CMS Collaboration, “Jet energy scale and resolution performances with 13TeV data,” CMS Detector Performance Note CMS-DP-2016-020, 2016.
<http://cds.cern.ch/record/2160347>.
- [178] ATLAS Collaboration, “Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV,” *Eur. Phys. J.* C73 no. 3, (2013) 2304, [arXiv:1112.6426](#) [hep-ex].
- [179] CMS Collaboration, “Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS,” *Journal of Instrumentation* 6 (2011) P11002, [arXiv:1107.4277](#) [physics.ins-det].
- [180] CMS Collaboration, “Jet Energy Resolution,” (internal), Accessed: 01/12/2017.
https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetResolution#From_dijets_recommended_numbers.
- [181] D0 Collaboration, “B-Jet Identification,” (public), Accessed: 01/12/2017.
https://www-d0.fnal.gov/Run2Physics/top/singletop_observation/singletop_observation_updated.html.
- [182] CMS Collaboration, “Heavy flavor identification at CMS with deep neural networks,” CMS Detector Performance Note CMS-DP-2017-005, 2017.
<http://cds.cern.ch/record/2255736>.
- [183] CMS Collaboration, “Measurement of $B\bar{B}$ Angular Correlations based on Secondary Vertex Reconstruction at $\sqrt{s} = 7$ TeV,” *JHEP* 03 (2011) 136, [arXiv:1102.3194](#) [hep-ex].
- [184] CMS Collaboration, “Identification of b quark jets at the CMS Experiment in the LHC Run 2,” CMS Physics Analysis Summary CMS-PAS-BTV-15-001, 2016.
<https://cds.cern.ch/record/2138504>.
- [185] F. Pedregosa *et al.*, “Scikit-learn: Machine Learning in Python,” *J. Machine Learning Res.* 12 (2011) 2825–2830, [arXiv:1201.0490](#) [cs.LG].
- [186] CMS Collaboration, “Performance of b tagging at $\sqrt{s} = 8$ tev in multijet, $t\bar{t}$ and boosted topology events,” CMS Physics Analysis Summary CMS-PAS-BTV-13-001, 2013.
<http://cdsweb.cern.ch/record/1581306>.
- [187] CMS Collaboration, “Performance of b-Tagging Algorithms in Proton Collisions at 13 TeV using the 2016 Data,” CMS Detector Performance Note CMS-DP-2016-042, 2016.
<http://cds.cern.ch/record/2202967>.
- [188] CMS Collaboration, “Introduction to missing transverse momentum (MET),” (public), Accessed: 07/12/2017.
<https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookMetAnalysis>.
- [189] CMS Collaboration, “Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s} = 8$ TeV,” *JINST* 10 no. 02, (2015) P02006, [arXiv:1411.0511](#) [physics.ins-det].

- [190] CMS Collaboration, “Performance of missing energy reconstruction in 13 TeV pp collision data using the CMS detector,” CMS Physics Analysis Summary CMS-PAS-JME-16-004, 2016. <http://cds.cern.ch/record/2205284>.
- [191] M. Stöver, *Search for third-generation scalar leptoquarks with the CMS experiment*. PhD thesis, Hamburg U., Hamburg, 2018. http://inspirehep.net/record/1653531/files/PhD_Thesis_Stoever.pdf.
- [192] CMS Collaboration, “Determination of residual jet-energy corrections using dijet events in 13 TeV data,” CMS Analysis Note CMS-AN-15-254, 2015. http://cms.cern.ch/iCMS/jsp/db_notes/noteInfo.jsp?cmsnoteid=CMS%20AN-2015/254. Internal documentation.
- [193] M. Wobisch *et al.*, “A new quantity for studies of dijet azimuthal decorrelations,” *JHEP* 01 (2013) 172, [arXiv:1211.6773](https://arxiv.org/abs/1211.6773) [hep-ph].
- [194] UA2 Collaboration, “Measurement of Production and Properties of Jets at the CERN anti-p p Collider,” *Z. Phys.* C20 (1983) 117–134.
- [195] D0 Collaboration, “Determination of the absolute jet energy scale in the D0 calorimeters,” *Nucl. Instrum. Meth.* A424 (1999) 352–394, [arXiv:hep-ex/9805009](https://arxiv.org/abs/hep-ex/9805009) [hep-ex].
- [196] CDF Collaboration, “Determination of the jet energy scale at the collider detector at Fermilab,” *Nucl. Instrum. Meth.* A566 (2006) 375–412, [arXiv:hep-ex/0510047](https://arxiv.org/abs/hep-ex/0510047) [hep-ex].
- [197] CDF Collaboration, “The Dijet angular distribution in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV,” *Phys. Rev. Lett.* 69 (1992) 2896–2900.
- [198] CMS Collaboration, “Determination of the jet energy scale in CMS,” *J. Phys. Conf. Ser.* 404 (2012) 012013.
- [199] CMS Collaboration, “Recommended Jet Energy Corrections and Uncertainties For Data and MC,” (internal), Accessed: 01/12/2017. <http://twiki.cern.ch/twiki/bin/viewauth/CMS/JECDataMC>.
- [200] D0 Collaboration, “Top trigger efficiency measurements and the top trigger package,” D0 Analysis Note D0 Note 4512, 2004. http://nicadd.niu.edu/~lq3/contrib/D0note4512_Top_Trigger_efficiency_Measurement_and_the_top_trigger_package.pdf.
- [201] CMS Collaboration, “Beam radiation instrumentation and luminosity,” (internal), Accessed: 01/12/2017. <https://cms-service-lumi.web.cern.ch/cms-service-lumi/brilwsdoc.html>.
- [202] CMS Collaboration, “Utilities for Accessing Pileup Information for Data,” (internal), Accessed: 01/12/2017. <https://twiki.cern.ch/twiki/bin/viewauth/CMS/PileupJSONFileforData>.
- [203] S. Ferrario Ravasio, T. Jezo, P. Nason, and C. Oleari, “A Theoretical Study of Top-Mass Measurements at the LHC Using NLO+PS Generators of Increasing Accuracy,” [arXiv:1801.03944](https://arxiv.org/abs/1801.03944) [hep-ph].
- [204] S. Moch *et al.*, “High precision fundamental constants at the TeV scale,” [arXiv:1405.4781](https://arxiv.org/abs/1405.4781) [hep-ph].
- [205] G. Corcella, “Interpretation of the top-quark mass measurements: a theory overview,” in *Proceedings, 8th International Workshop on Top Quark Physics (TOP2015): Ischia, Italy, September 14-18, 2015*, vol. TOP2015, p. 037. 2016. [arXiv:1511.08429](https://arxiv.org/abs/1511.08429) [hep-ph].

- [206] M. Beneke, P. Marquard, P. Nason, and M. Steinhauser, “On the ultimate uncertainty of the top quark pole mass,” *Phys. Lett. B* 775 (2017) 63–70, [arXiv:1605.03609 \[hep-ph\]](#).
- [207] M. Butenschoen *et al.*, “Top Quark Mass Calibration for Monte Carlo Event Generators,” *Phys. Rev. Lett.* 117 no. 23, (2016) 232001, [arXiv:1608.01318 \[hep-ph\]](#).
- [208] C. S. Hill, J. R. Incandela, and J. M. Lamb, “A Method for measurement of the top quark mass using the mean decay length of b hadrons in $t\bar{t}$ events,” *Phys. Rev. D* 71 (2005) 054029, [arXiv:hep-ex/0501043 \[hep-ex\]](#).
- [209] CMS Collaboration, N. Kovalchuk, “Top quark mass measurements with CMS,” in *Proceedings, 25th International Workshop on Deep-Inelastic Scattering and Related Topics (DIS 2017): Birmingham, UK, April 3-7, 2017*, p. 131. 2017. <https://inspirehep.net/record/1647839>.
- [210] CMS Collaboration, “Certified collision data from 2016,” (internal), Accessed: 01/01/2018. <https://cms-service-dqm.web.cern.ch/cms-service-dqm/CAF/certification/Collisions16/13TeV/ReReco/Final/>.
- [211] CMS Collaboration, “CMS DQM graphical user interface,” (internal), Accessed: 01/01/2018. <https://cmsweb.cern.ch/dqm/offline/session/BQyCZR>.
- [212] T. Melia, P. Nason, R. Rontsch, and G. Zanderighi, “W+W-, WZ and ZZ production in the POWHEG BOX,” *JHEP* 11 (2011) 078, [arXiv:1107.5051 \[hep-ph\]](#).
- [213] E. Re, “Single-top Wt-channel production matched with parton showers using the POWHEG method,” *Eur. Phys. J. C* 71 (2011) 1547, [arXiv:1009.2450 \[hep-ph\]](#).
- [214] S. Alioli, P. Nason, C. Oleari, and E. Re, “NLO single-top production matched with shower in POWHEG: s - and t -channel contributions,” *JHEP* 09 (2009) 111, [arXiv:0907.4076 \[hep-ph\]](#).
- [215] S. Frixione, E. Laenen, P. Motylinski, and B. R. Webber, “Angular correlations of lepton pairs from vector boson and top quark decays in Monte Carlo simulations,” *JHEP* 04 (2007) 081, [arXiv:hep-ph/0702198 \[HEP-PH\]](#).
- [216] P. Artoisenet, R. Frederix, O. Mattelaer, and R. Rietkerk, “Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations,” *JHEP* 03 (2013) 015, [arXiv:1212.3460 \[hep-ph\]](#).
- [217] R. D. Ball *et al.*, “Parton distributions with LHC data,” *Nucl. Phys. B* 867 (2013) 244–289, [arXiv:1207.1303 \[hep-ph\]](#).
- [218] E. Richter-Was and Z. Was, “W production at LHC: lepton angular distributions and reference frames for probing hard QCD,” *Eur. Phys. J. C* 77 no. 2, (2017) 111, [arXiv:1609.02536 \[hep-ph\]](#).
- [219] S. Catani, G. Ferrera, and M. Grazzini, “W Boson Production at Hadron Colliders: The Lepton Charge Asymmetry in NNLO QCD,” *JHEP* 05 (2010) 006, [arXiv:1002.3115 \[hep-ph\]](#).
- [220] E. L. Berger, F. Halzen, C. S. Kim, and S. Willenbrock, “WEAK BOSON PRODUCTION AT TEVATRON ENERGIES,” *Phys. Rev. D* 40 (1989) 83. [Erratum: *Phys. Rev. D* 40, 3789 (1989)].
- [221] A. D. Martin, R. G. Roberts, and W. J. Stirling, “Improved Parton Distributions and W, Z Production at p anti-p Colliders,” *Mod. Phys. Lett. A* 4 (1989) 1135.

- [222] CMS Collaboration, “Methods to apply b-tagging efficiency scale factors,” (internal), Accessed: 01/12/2017.
<https://twiki.cern.ch/twiki/bin/viewauth/CMS/BTagSFMethods>.
- [223] D0 Collaboration, “Direct measurement of the top quark mass at D0,” *Phys. Rev. D* 58 (1998) 052001, [arXiv:hep-ex/9801025](#) [hep-ex].
- [224] CMS Collaboration, “Programming documentation of hitfit: A kinematic fitter for top quark-antiquark pair lepton+jets event,” (internal), Accessed: 01/12/2017. http://cms.cern.ch/iCMS/jsp/db_notes/noteInfo.jsp?cmsnoteid=CMS%20AN-2011/171.
- [225] S. S. Snyder, *Measurement of the Top Quark Mass at D0*. PhD thesis, SUNY, Stony Brook, 1995. http://lss.fnal.gov/cgi-bin/find_paper.pl?thesis-1995-27.
- [226] CDF Collaboration, “Top quark mass measurement using the template method in the lepton + jets channel at CDF II,” *Phys. Rev. D* 73 (2006) 032003.
- [227] D0 Collaboration, “Precise measurement of the top quark mass from lepton+jets events at D0,” *Phys. Rev. Lett.* 101 (2008) 182001, [arXiv:0807.2141](#) [hep-ex].
- [228] D0 Collaboration, “Measurement of the top quark mass using dilepton events,” *Phys. Rev. Lett.* 80 (1998) 2063–2068, [arXiv:hep-ex/9706014](#) [hep-ex].
- [229] R. H. Dalitz and G. R. Goldstein, “The Decay and polarization properties of the top quark,” *Phys. Rev. D* 45 (1992) 1531–1543.
- [230] D0 Collaboration, “Measurement of the top quark mass in the lepton + jets channel using the ideogram method,” *Phys. Rev. D* 75 (2007) 092001, [arXiv:hep-ex/0702018](#).
- [231] CDF Collaboration, “Measurement of the top-quark mass in all-hadronic decays in $p\bar{p}$ collisions at CDF II,” *Phys. Rev. Lett.* 98 (2007) 142001, [arXiv:hep-ex/0612026](#) [hep-ex].
- [232] CMS Collaboration, E. Schlieckau, “Measurement of the top-quark mass in all-jets $t\bar{t}$ events in pp collisions at 7 TeV,” in *Proceedings, 6th International Workshop on Top Quark Physics (TOP2013): Durbach, Germany, September 14-19, 2013*, pp. 228–232. 2014.
<https://inspirehep.net/record/1344093/files/P68.pdf>.
- [233] CMS Collaboration, “Jet energy scale uncertainty sources,” (internal), Accessed: 27/12/2017.
<https://twiki.cern.ch/twiki/bin/viewauth/CMS/JECUncertaintySources>.
- [234] ALEPH Collaboration, “Study of the fragmentation of b quarks into B mesons at the Z peak,” *Phys. Lett. B* 512 (2001) 30–48, [arXiv:hep-ex/0106051](#) [hep-ex].
- [235] DELPHI Collaboration, “A study of the b-quark fragmentation function with the DELPHI detector at LEP I and an averaged distribution obtained at the Z Pole,” *Eur. Phys. J. C* 71 (2011) 1557, [arXiv:1102.4748](#) [hep-ex].
- [236] J. Butterworth *et al.*, “PDF4LHC recommendations for LHC Run II,” *J. Phys.* G43 (2016) 023001, [arXiv:1510.03865](#) [hep-ph].
- [237] M. Czakon, D. Heymes, and A. Mitov, “High-precision differential predictions for top-quark pairs at the LHC,” *Phys. Rev. Lett.* 116 (2016) 082003, [arXiv:1511.00549](#) [hep-ph].
- [238] CMS Collaboration, “Measurement of differential cross sections for top quark pair production using the lepton+jets final state in proton-proton collisions at 13 TeV,” *Phys. Rev. D* 95 no. 9, (2017) 092001, [arXiv:1610.04191](#) [hep-ex].

- [239] CMS Collaboration, “Measurement of the differential cross section for $t\bar{t}$ production in the dilepton final state at $\sqrt{s} = 13$ TeV,” CMS Physics Analysis Summary CMS-PAS-TOP-16-011, 2016. <https://cds.cern.ch/record/2140061>.
- [240] ATLAS Collaboration, “Measurement of the top quark mass in the $t\bar{t} \rightarrow$ dilepton channel from $\sqrt{s} = 8$ TeV ATLAS data,” *Phys. Lett. B* 761 (2016) 350–371, [arXiv:1606.02179 \[hep-ex\]](#).
- [241] ATLAS Collaboration, “Top-quark mass measurement in the all-hadronic $t\bar{t}$ decay channel at $\sqrt{s} = 8$ TeV with the ATLAS detector,” *JHEP* 09 (2017) 118, [arXiv:1702.07546 \[hep-ex\]](#).
- [242] D0 Collaboration, “Measurement of the Top Quark Mass Using the Matrix Element Technique in Dilepton Final States,” *Phys. Rev. D* 94 no. 3, (2016) 032004, [arXiv:1606.02814 \[hep-ex\]](#).
- [243] CDF Collaboration, “Top-quark mass measurement in events with jets and missing transverse energy using the full CDF data set,” *Phys. Rev. D* 88 no. 1, (2013) 011101, [arXiv:1305.3339 \[hep-ex\]](#).
- [244] CMS Collaboration, “Measurement of the mass difference between top quark and antiquark in pp collisions at $\sqrt{s} = 8$ TeV,” *Phys. Lett. B* 770 (2017) 50–71, [arXiv:1610.09551 \[hep-ex\]](#).
- [245] D. Buttazzo, G. Degrandi, P. P. Giardinò, G. F. Giudice, F. Sala, A. Salvio, and A. Strumia, “Investigating the near-criticality of the Higgs boson,” *JHEP* 12 (2013) 089, [arXiv:1307.3536 \[hep-ph\]](#).
- [246] A. Andreassen, W. Frost, and M. D. Schwartz, “Scale Invariant Instantons and the Complete Lifetime of the Standard Model,” [arXiv:1707.08124 \[hep-ph\]](#).
- [247] CMS Collaboration, N. Kovalchuk, “Measurement of the top quark mass with lepton+jets final states in pp collisions at $\sqrt{s} = 13$ TeV,” in *Proceedings, 10th International Workshop on Top Quark Physics (TOP 2017): Braga, Portugal, September 17-22, 2017*, p. 131. 2017. <http://inspirehep.net/record/1648557/files/arXiv:1801.05619.pdf>.

Acknowledgements

At this point, I would like to cordially thank people who made direct or indirect contributions to the formation of my thesis. Thank you very much for sharing your experience and knowledge with me, as well as for your encouraging support!

First of all, I would like to express my special gratitude to Prof. Dr. Peter Schleper for the continuous support throughout my PhD study and research, for his advice and encouragement, as well as for providing me an opportunity to be a part of the Hamburg University group and to contribute to the research in the CMS Collaboration.

I am very grateful to my thesis committee: Prof. Dr. Elisabetta Gallo, Prof. Dr. Peter Schleper, Prof. Dr. Gudrid Moortgat-Pick, Dr. Christian Sander, and Prof. Dr. Robin Santra for evaluating my thesis and thesis defense.

My sincere thanks goes to my supervisor Dr. Hartmut Stadie, who guided me during all these years of my PhD work. I thank you for your crucial advice, instructions, and that you always could find a time when I needed your help.

I would also like to extend my thanks to Dr. Paolo Gunnellini for the beneficial experience during the cooperation work on the CR tune performance, and to Dr. Anastasia Karavdina, who took responsibility for the JEC estimation at the time when I needed to start working on my main subject.

A special gratitude goes out to people who find the time to read my thesis and to provide useful comments. Thank you very much Dr. Adrian Perieanu, Dr. Fred Stober, Dr. Samuel Bein, Dr. Marc Stoeber, and Arne Reimers.

I am very happy and proud to be a group member at the University of Hamburg and of the UHH top quark analysis team, where I met wonderful people and learned a lot of new things. Especially, I would like to thank Marek Niedziela, Simon Kurz, and Arne Reimers for creating a nice and warm atmosphere in the office and beyond.

I would like to express my very great appreciation to Prof. Dr. Robert Klanner, Prof. Dr. Volodymyr Aushev, and Prof. Dr. Petro Holod, who helped me take the first steps in my research career.

I cordially thank my Ukrainian colleagues and friends at DESY for your emotional support and help, especially at the beginning of my PhD work. My best friends, Yuliia, Olga S.-R., Olga V., and Inna, thank you very much for your continuous support and for believing in me!

With a special mention to my loving parents, who supported me at each step during all my life, I would like to send my deepest gratitude to them. This accomplishment would not be possible without you. My sincere thanks also goes to my older brother Serhii and his wife Hanna, who

encouraged me and reassured me that I can always count on them. I thank you to my whole family for providing me with continuous encouragement!

Marc, it was indeed not such a bad idea to share the service work with you :P. Thank you for your unfailing support and patience, especially during the writing of my thesis. Thank you for always being on my side, either during tough or tranquil stages of my life. Danke, dass Du immer für mich da bist! 🐶

Eidesstattliche Versicherung / Declaration on oath

Hiermit versichere ich an Eides statt, die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Hilfsmittel und Quellen benutzt zu haben.

Die eingereichte schriftliche Fassung entspricht der auf dem elektronischen Speichermedium.

Die Dissertation wurde in der vorgelegten oder in einer ähnlichen Form nicht schon einmal in einem früheren Promotionsverfahren angenommen oder als ungenügend beurteilt.

Ort, Datum

Nataliia Kovalchuk