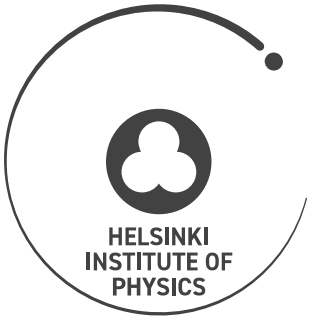

HIP Internal Report Series
HIP-2022-04

Measurement of the Top Quark Mass with the CMS Experiment

Hannu Siikonen

Doctoral dissertation, to be presented for public discussion with the permission of the Faculty of Science of the University of Helsinki, in the Auditorium A129 at Chemicum, A.I. Virtasen aukio 1, Helsinki, on the 22nd of November 2022 at 13 o'clock.

Helsinki Institute of Physics
Faculty of Science
University of Helsinki
Finland

Helsinki 2022

Supervisor
Associate Professor Mikko Voutilainen, University of Helsinki, Finland

Pre-examiners
Professor Sara Strandberg, Stockholm University
Univ.-Prof. Dr. Andre Hoang, Universität Wien

Opponent
Senior lecturer, Royal Society university research fellow Mark Owen, University of Glasgow, United Kingdom

Custos
Associate Professor Mikko Voutilainen, University of Helsinki, Finland

The Faculty of Science uses the Ouriginal system (plagiarism recognition) to examine all doctoral dissertations.

Contact information

Helsinki Institute of Physics
P.O. Box 68 (Gustaf Hällströmin katu 2a)
FI-00014 University of Helsinki
Finland

ISSN 1455-0563 (print)
ISSN 2814-9459 (online)
ISBN 978-951-51-8411-5 (paper)
ISBN 978-951-51-8412-2 (pdf)
HIP Internal Report Series, published electronically at
http://ethesis.helsinki.fi
Helsinki 2022
Unigrafia

Acknowledgements

I started working at the Helsinki Institute of Physics (HIP) CMS Programme as a summer student in 2014. With a recently finished BSc degree from Aalto, the summer at CERN offered a fresh take on the career options of a physicist. At CERN, fundamental theory was closely connected to practice, and a door to studies in topics forbidden at Aalto – such as particle physics and general relativity – were opened. In comparison, the theoretically most involved options at Aalto appeared less inspiring at the time: fusion research is advancing extremely slowly, while quantum & low temperature physics is less fundamental than particle physics.

Now, short of a decade later my PhD thesis is finished and an era is closing its end. My time at HIP has included various sources of turbulence in world politics, including the numerous consequences of the COVID-19 pandemic. In the midst of all the unrest, the research work has provided a solid unchanging factor in my life. **Three years of my PhD work were funded by Vilho, Yrjö ja Kalle Väisälän rahasto,** for which I am extremely grateful. I also want to thank **Suomen Kulttuurirahasto (Kauko ja Anita Lehtisen rahasto) for funding a short of a year** near the end of my PhD studies. In addition, thanks are due to **Waldemar von Frenckells stiftelse for funding the majority of the travels during my PhD work**, including an eye-opening journey to Fermilab. I also want to thank the **University of Helsinki and HIP** for the necessary supplementary funding.

Lastly, I want to thank the taxpayer reading this book for funding our educational system and (maybe unknowingly) the research at CERN. Even if CERN is focused at fundamental research, the research often ends up widely benefiting industry and society. Cutting-edge detector hardware is developed and tested at the CERN facilities, and a portion of this work is done in Finland. Furthermore, the PhD training produces experts in detector hardware, data-analysis and computing – many of which end up working outside academia. For the politically inclined reader I want to emphasize that these benefits are only reached through a long-term stabil-

ity in financing. In the recent years both university studies and research in Finland have been under a pressure to increase efficiency. Most of the actions to reach such goals have been carried out through economic metrics, which adapt poorly to the academic world. For instance, the more time a researcher needs to spend on funding applications, the less they have psychological safety and time to perform the actual research. True innovation cannot be forced and it is not tied to the clock: it only arises as a side product when sufficient resources are provided to the correct people.

Personal acknowledgements are due to a number of people. I want to thank Assoc. Prof. **Mikko Voutilainen** for providing ideas, funding, guidance and feedback throughout the years. Thanks are also in place for Dr. **Juska Pekkanen**, who took a great role in first introducing me to the work at CMS. A special note of thank you is directed to Dr. **Hartmut Stadie** for providing access to the earlier versions of the top quark mass analysis code, and computing resources at the DESY computers.

My great gratitude is also aimed to all the people involved in the process of creating this thesis. Most importantly these include Senior Lecturer **Mark Owen**, who has accepted to take the time and effort of being my opponent at the public defence of the thesis. Invaluable work was also done by Professor **Sara Strandberg** and Univ.-Prof. Dr. **Andre Hoang** in the pre-examination and commenting of this text. Crucial feedback was also provided by Drs. **Tapio Lampén** and **Sami Lehti**.

Many thanks are directed to the colleagues that have passed through our office along the years. A special mention is directed to Dr. **Toni Mäkelä**, whose earlier work is tied with the contents of this thesis, and with whom it was very fruitful to collaborate. Also worthy of a separate mention is our present PhD student, **Mikael Myllymäki**, for whom it has been a pleasure to introduce the tools used in our analysis framework, and whose work has taken our ability to analyze the parameter values used in the CMS simulations onto a new level. Further thanks are in place for **Joona, Jaana, Santeri, Terhi, Laura, Mikko and Kimmo**.

I am thankful to all my friends from Aalto, who inspired my studies, and with whom I have shared many interesting travels. Many of them would deserve an individual mention, but to limit this listing into two pages, I only have room to bring up **Alex and Vera-Maria**. Their family is always welcoming and provides pleasant company. A final special mention is directed to the members of our informal reading circle, who have functioned as an important social vent amidst all the recent turmoils.

Great indebtedness is aimed to my parents **Timo** and **Maikku**, who have lovingly supported me in my studies and in my work without strongly questioning my choices. The most important mention is due to my beloved **Aliisa**, who has supported and understood me even in the most frustrating and stressful moments of the construction process of this thesis. I also want to thank **Eero** for bringing many smiles into our daily life, and for helping me with the final proof-reading of this text.

Abstract

The standard model of particle physics describes an astonishing number of phenomena. Yet at the same time it is incomplete: it does not describe e.g. gravitation. Finding explicit weaknesses in the standard model predictions has proved to be difficult, and hence precision measurements are currently one of the most promising methods towards this goal. One of the most intriguing precision measurements is that of the top quark mass (m_t), which is connected for instance to the question about the meta-stability of the universe.

This thesis strives multilaterally towards a more precise measurement and interpretation of the top quark mass. The work begins with efforts towards a more precise jet calibration at the CMS. Then, the possible weaknesses of a DØ m_t analysis are reviewed. Finally, a m_t measurement at the CMS is constructed for the legacy 2017–2018 datasets.

The jet energy corrections are the most important experimental factor in the uncertainties of the top quark mass. Hence, they are closely linked with the m_t measurement. The work on jets in this thesis aims for an exceptionally precise jet energy calibration for the CMS Run 2 legacy datasets. The author has made several important contributions towards the jet energy corrections in the Run 2 legacy reconstruction.

The re-assessment of a DØ top quark mass measurement is performed outside of the CMS and DØ affiliations. The DØ top quark mass value is an important outlier in the top quark mass world combination, and a better understanding of the reasons behind this is desirable. In an earlier study it was shown that there are possible discrepancies in the flavor-dependent jet energy corrections at DØ. In this thesis we demonstrate that these discrepancies (if they can be confirmed) shift the DØ top quark mass measurement to a value that is more in line with the other major measurements from CMS, ATLAS and CDF.

The work culminates in the design and validation of the first direct CMS lepton+jets m_t measurement on the 2017–2018 datasets. The analysis is executed using a new profile likelihood method, where the collected data can constrain systematic uncertainties *in situ*. Agreement between data and simulation is verified within the systematic uncertainties using control plots. The impact of an extensive set of systematic uncertainties on the m_t measurement is assessed using simulations. Also the full effects of limited statistics in simulations are demonstrated using toy experiments. It is confirmed that the limiting systematic uncertainty in the current m_t measurements is the modeling of b quark jets. This challenge can be encountered either by enhancing the b jet energy corrections or by performing the measurement on a larger amount of data. In profile likelihood analyses, the latter is also a valid approach.

Author's Contributions

Chapters 4, 5 and 6 contain the main results of this thesis. Chapter 4 is an introduction into jet physics containing both theory, and results produced by the author. The author has worked extensively on the production of the SMPJ (Standard Model Physics: Jets) ntuples, and their analysis with the **jetphys** [1] software tool. The SMPJ ntuples are a predecessor of the compact NANOAOD data format, currently used at CMS. These ntuples have been used in multiple groups for the analysis of dijet and multijet events.

From the analysis with the jetphys tool the author has iteratively derived results e.g. on jet energy fractions and jet veto maps during the Run 2 data analysis. The results on jet energy fractions have been published earlier in Ref. [2]. The purpose of all this work has been the derivation of the Run 2 legacy dataset jet energy corrections. Some of these results are presented in chapter 4, where the figures produced by the author are denoted with a **Work in Progress** tag. These have not been published earlier, but have been endorsed by the CMS Jets and Missing Energy Physics Object Group to be presented in this thesis. In addition to these contributions, Chapter 4 presents the author's contributions on the development of the common CMS software release (CMSSW).

Most of the CMS jet studies presented in chapter 4 have been performed for full 2016–2018. In contrast, the top quark mass studies done for the CMS experiment in chapter 6 consider only the 2017–2018 datasets. To promote clarity, the results in chapter 4 are also mainly presented for 2017–2018.

Chapter 5 is fully original work by the author, performed outside of the CMS and DØ affiliations. These studies are based on previous results derived by Dr. Toni Mäkelä [3], hinting that the DØ Run IIb flavor-dependent jet energy corrections might have been incorrectly calibrated. This sparked interest in the related DØ lepton+jets top quark mass measurement, which is a clear outlier among the most important m_t measurements.

To further study the subject, the author designed a method for inspecting how the recalibration of DØ's flavored jet energy corrections would affect the top quark mass measurement. The results were first presented in a pre-print [4], and later discussed at the virtual TOP2020 conference [5]. The pre-print was not pushed for publication, but instead most of it has been integrated into this monograph.

In chapter 6 a new and more precise m_t measurement is pursued for the CMS collaboration. Before this work the main analysis code repository was developed at the University of Hamburg. The author has single-handedly and heavily updated this code [6] and its CMSSW integration [7] to meet the needs of today. The final steps of the new analysis have been implemented by the author alone using a profile-likelihood analysis based on the Combine tool. Similarly as in chapter 4, the figures produced by the author are indicated with a **Work in Progress** tag. These have been endorsed by the CMS Top Quark Physics Analysis Group to be presented in this thesis.

Contents

Acronyms and Symbols

Anti-k_T	Jet clustering algorithm
AOD	Analysis Object Data
ATLAS	A Toroidal LHC ApparatuS (a multipurpose experiment at the LHC)
b jet	Jet originating from a b quark
CERN	European Laboratory for Particle Physics (Conseil Européen pour la Recherche Nucléaire)
CHS	Charged Hadron Substraction
CMS	Compact Muon Solenoid (a multipurpose experiment at the LHC)
CMSSW	CMS SoftWare framework
CR	Color Reconnection
Data/data	Data gathered by the CMS detector (not simulated), in contrast to simulated data
ECAL	Electromagnetic Calorimeter
ERD	Early Resonance Decay
FSR	Final State Radiation
FNAL	Fermi National Accelerator Lab
GWS	Glashow-Weinberg-Salam unification theory for electroweak interactions
HCAL	Hadronic Calorimeter
HEP	High Energy Physics
HLT	High-Level Trigger
ID	Identification (Condition)
ISR	Initial State Radiation
JEC	Jet Energy Correction
JER	Jet Energy Resolution
JES	Jet Energy Scale
Jet	A spray of particles originating from an energetic parton
l jet	Jet originating from a light quark (udsc)
LEP	Large Electron-Positron Collider

LHC	Large Hadron Collider
LO	Leading Order
LN	Level N (e.g. Level 1)
MC	Monte Carlo, often refers to simulations in general
ME	Matrix Element
MPI	Multi-Parton Interactions
MS	Minimal Substraction
N^{n}LO	(Next-to-)^{n}Leading Order
OOC	Out-Of-Cone
Parton	A quark or a gluon
PDF	Parton Distribution Function
PF	Particle Flow (algorithm)
PV	Primary Vertex
pp	proton-proton
PS	Parton Showers
Pileup	Particle contamination from adjacent pp collisions
QCD	Quantum Chromodynamics
QED	Quantum Electrodynamics
QFT	Quantum Field Theory
SM	Standard Model
SV	Secondary Vertex
TF	Transfer Function
UE	Underlying Event
WP	Working Point

$p_T, \vec{p}_T$	Transverse momentum
$p_{\mathrm{T}}^{\mathrm{miss}}, \vec{p}_{\mathrm{T}}^{\mathrm{miss}}$	Missing transverse momentum
ϕ	Azimuthal angle (in the transverse plane)
η	Pseudo-rapidity
y	Rapidity

R	Jet radius in the $\eta - \phi$ plane
ΔR	Directional distance between two objects in the $\eta - \phi$ plane
σ	Cross-section
$\mathcal{L}$	Integrated Luminosity

Notation and conventions

The four-momenta of particles are studied in a right-handed coordinate system, where the x-axis points towards the LHC ring center, y-axis points towards the sky and z-axis is parallel to the proton beam axis. The CMS detector is cylindrically symmetric, and hence variables are often viewed in cylindrical coordinates. Most importantly, (four-)momentum is often expressed in terms of p_T, η, ϕ and m instead of p_x, p_y, p_z and E. Here, p_T is the transverse ($x-y$ plane) momentum and ϕ is the corresponding azimuthal angle. The variable m is mass and η is a transformation of the polar angle θ, as explained in chapter 3.

Physical quantities are expressed in terms of natural units, where $c = \hbar = 1$. In consequence, e.g. mass and energy share the same units. These are expressed in multiples of electron volts, e.g. GeV.

1 Introduction

The operation of the Large Hadron Collider (LHC) at CERN (European Laboratory for Particle Physics) consists of high-energy collision data-taking periods (Runs) and hardware upgrade and maintenance periods (Long Shutdowns). Run 1 (2009–2013) of the LHC provided the long-waited experimental discovery of the Higgs boson [8, 9].

The second Run of the LHC commenced with equally high expectations, and it ranged from 2015 to 2018. The most important differences between Run 1 and Run 2 were the increase in the proton-proton collision energy from 7–8 TeV to 13 TeV, the increase in the total integrated luminosity[1] from 27 fb^{-1} to 151 fb^{-1} [10] and the increased amount of adjacent proton-proton collisions (pileup) in each proton bunch crossing from 10–21 to 27–38 [10].

So far one of the most notable Run 2 findings has been the lack of evidence for low energy **supersymmetry**, which many anticipated to be observed at the Run 2 collision energies. Consequently, more attention has shifted towards precision measurements. Some of the currently most interesting precision-measurement results are the gradually increasing hints of lepton-universality breaking [11–13].

Precision measurements are demanding and time-consuming. However, they can point out small discrepancies that have been previously overlooked. In the current state of particle physics – where making new discoveries is increasingly difficult – these methods are becoming more and more important. It is certain that there are yet discoveries to be made in particle physics: the greatest challenge is having the correct foresight on where to look.

This thesis deals with one of the most crucial precision measurements: that of the top quark mass (m_t). The measurement is performed using the 100 fb^{-1} of data collected by the Compact Muon Solenoid (CMS) detector during the years 2017–2018.

[1] A measure of the number of observed collision events.

1.1 The Motivation for This Study

The top quark is the most massive elementary particle discovered. It is the only quark that is more massive than the W boson, giving it a special relationship to the weak interaction. In consequence, the top quark decays very quickly[2] into a W boson and a quark. Hence, it is unable to form bound states with other quarks.

Many predictions and interpretations are highly dependent on the exact value of the top quark mass. Some of the most interesting predictions require a significantly higher measurement precision than currently provided to reach a sufficient level of precision. One of the most intriguing results of this kind is the interpretation of the metastability of the universe. Within the measurement uncertainty of the top quark mass, the electroweak vacuum could be stable or metastable [15].

Achieving a higher measurement precision in itself is a great motivation. The value of m_t is a fundamental parameter of the standard model of particle physics. and the lowest feasible measurement precision[3] of $\mathcal{O}(200\,\text{MeV})$ [16] ($0.06\,\% - 0.11\,\%$ w.r.t. the top quark mass) is still unreached. Reaching this level of precision with the current data is feasible, and doing so would likely shift the focus towards the interpretation of the results. Such studies include for instance the impact of varying the parton shower models, which play an important part in the top quark decay modeling.

1.2 Previous Studies

Top mass measurements start from the discovery at the Tevatron collider at Fermi National Accelerator Lab (FNAL) in 1995 [17, 18]. Since then, both the Tevatron CDF and DØ collaborations and the CERN ATLAS and CMS collaborations have performed numerous top quark mass measurements. In Fig. 1.1 some of the most prominent top quark mass measurements are presented. It is noteworthy that the ATLAS, CDF and CMS measurements agree quite well, whereas DØ produces a distinctively different result. With the studies of this thesis, new insight is given also on the world combinations.

The most precise CERN measurements rely on a number of separate uncorrelated observables (referred to as **dimensions**). At the CMS the most precise Run 1 and early Run 2 analyses are based on two dimensions: an event-wise best fit value of m_t, and the hadronic W boson mass, reconstructed from two jets [19, 20]. The latter dimension effectively performs *in-situ* tuning of jet energy scales. The ATLAS Run 1 analyses have taken a similar approach, adding a third dimension[4] for separately tuning the b jet energy scale [21].

Finally, the recent CMS Run 2 analysis of Ref. [22] added two more dimensions,

[2] In approximately 4×10^{-25} s [14].

[3] Determined by the features of QCD.

[4] Ratio between the p_T sums of two b jets and the two quark jets from the W boson.

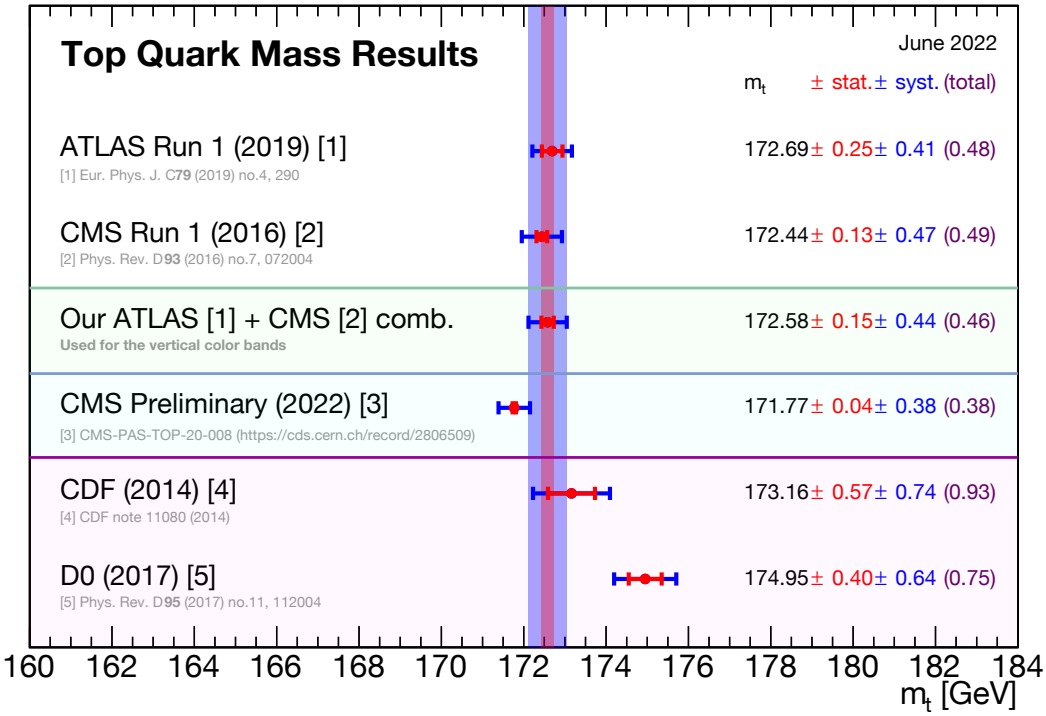

Figure 1.1: Important recent m_t results. Our private combination of the ATLAS (2019) and CMS (2016) results is used for the vertical error bands.

ending up with a 5D analysis. The analysis adopted the new technology of **profiling** the systematic uncertainties into the likelihood function used in the top quark mass fit. As a result, no separate jet energy scale parameters are required, but the nuisance parameters of the systematic uncertainties provide a more general method for achieving the same effect.

The newest and most precise CMS result on the 2016 lepton+jets dataset (presented in Fig. 1.1) is still preliminary [22]. The same 2016 lepton+jets dataset has been analyzed earlier using a different method, yielding $m_t = 172.25 \pm 0.63\,\text{GeV}$ [20]. The new result is somewhat lower than the earlier results, and it will be interesting to see if this trend continues when other similar measurements with the new likelihood techniques are released.

The methods used in the m_t measurement of this thesis are similar to those in Ref. [22], but a 3 times larger dataset is analyzed: 100 fb^{-1} in 2017–2018 vs. 36 fb^{-1} in 2016 [10].

1.3 The Structure of This Thesis

In chapter 2 we go through a selection of the most important theoretical concepts in particle physics. An emphasis is given to understanding what is really being measured, when a quark mass is being determined.

Chapter 3 introduces the CMS detector and its operation at the LHC. In addition, the ways in which the CMS collision events are simulated are presented intertwined with the experimental layout. As the author of this thesis has been heavily involved with simulational studies, this is a natural path to take – in the end, simulations present our deepest understanding of the way in which nature works.

The three following chapters contain a mixture of theory/simulations and results by the author. This is a natural approach, as for the topics of these chapters theory and practice are closely connected.

In chapter 4 the nature of jets[5] at the particle detectors are discussed in detail. Moreover, the chapter explores the nature of Jet Energy Corrections (JECs) on the data measured with the CMS detector. The end of the chapter presents some of the author's work towards better JEC calibration at the CMS. This includes studies on jet energy fractions, work on adding the $\text{W} \to qq'$ channel into the JEC global fit and derivation of veto jet maps that are based on divergent jet counts. The veto maps are used to exclude poorly calibrated regions from the analysis. The results have been endorsed by the CMS collaboration to be presented in this thesis.

Chapter 5 presents private work done independent of the CMS collaboration. A theoretical method is developed to study the possible effect of errors in JEC calibrations in a previous DØ top quark mass measurement. Later on, the method

[5] Jets are collimated particle showers produced in high-energy particle collisions.

is applied to the DØ data. The hypothesis of the issues in the DØ JECs is based on a previous Master's thesis [3].

This chapter also functions as a historical review: it gives an understanding on how the top quark mass measurements have been performed earlier. The results have been previously published as a pre-print [4], written alone by the author of this thesis. As the pre-print was written only to initiate scientific discussion on the topic and no official publication was intended to be made, most of the text is integrated into this thesis.

Chapter 6 introduces the methods used in the CMS top quark mass measurement. Moreover, a comparison between data and simulation is done using a number of control plots. A review and assessment of the expected systematic uncertainties is performed. The effect of statistical uncertainties in simulation on the systematic uncertainties are further studied with toy experiments. Also these results have been endorsed by the CMS collaboration to be presented in this thesis.

Finally, chapter 7 provides a summary and an outlook for the future of m_t measurements. Moreover, supplementary material is provided in the appendices.

2 Theoretical Background

This chapter gives a brief introduction to the theoretical aspects behind top quark mass (m_t) measurements. The topics include the Standard Model (SM) and the definitions of mass. Moreover, the necessary steps for observing the top quark at the LHC are clarified. The text is generally based on textbook information from the references [14, 23, 24]. Specific references in this chapter are given only for information not given in these books and for topics that cannot be considered to be common knowledge in the field.

2.1 The Foundation of Particle Physics

In his daily life, the layman becomes accustomed to gravity and the various forms of electromagnetic interactions. The basic formulations of these two interactions are quite easy to understand and are usually taught in mandatory physics classes. In contrast, the remaining two fundamental interactions are often skimmed over and portrayed as complicated and poorly understood. Hence, it might surprise the layman that this view is incorrect. The strong and the weak interactions are understood in many aspects equally well as the electromagnetic interactions. This viewpoint is motivated by Quantum Mechanics, which plays an important role in modern physics.

2.1.1 Quantum Field Theories: A Historical Review

All the best theories in particle physics rely on the same fundamental principles: quantization of relativistic fields. This short expression combines three important concepts: Quantum Mechanics, Special Relativity and Field Theory. The implication of **Quantum Field Theories** (QFTs) is that nature is built of elementary particles. These are the **field quanta** that dwell in their respective fields. The strong, weak and electromagnetic fields and interactions are all described by QFTs.

In comparison, Quantum Mechanics has not been successfully accommodated to the celebrated theory of General Relativity, which is thus not a QFT. On the other hand, General Relativity is special in the way that it takes the concept of relativity further than the QFTs. General Relativity is a classical field theory that accurately describes the curvature of space and time.

Of the three fundamental concepts, it is perhaps easiest to agree with Special Relativity. It imposes that physical laws should remain the same in all inertial coordinate systems. Historically, Einstein combined this postulate with the experimental observation that the speed of light is constant. In consequence, he found the transformation laws between two inertial coordinate systems for the space and time coordinates. This was a remarkable discovery, as the electromagnetic fields also follow these **Lorentz transformations**. Electromagnetism is described by the Maxwell equations, which were scrutinized as early as in the 19th century. Special Relativity came into being only in the beginning of the 20th century. An important consequence of Special Relativity is the equation between the speed of light c, the rest mass m_0, energy E and momentum $\vec{p}$:

$$E^2 = \left([m_0 c]^2 + |\mathbf{p}|^2\right)^2 c^2. \tag{2.1}$$

In contrast to Special Relativity, it is more difficult to come up with natural easy-to-understand arguments that would fully explain Quantum Mechanics. The essence of Quantum Mechanics is that at the sub-microscopical level most physical observables are **quantized**, i.e. they can be described by integer values. In the modern world this is common knowledge, as almost every adult knows that all matter consists of discrete atoms rather than some sorts of continuous media. A less commonly-known fact is that the quantized states can be fully described by complex-valued wave functions. The square of the absolute value of a wave function describes the probability to observe a particle at a given position. Einstein remained fundamentally unhappy with this concept, as is implied in his famous letter to Max Born (1926) [25]:

> *Die Quantenmechanik ist sehr achtunggebietend. Aber eine innere Stimme sagt mir, daß das noch nicht der wahre Jakob ist. Die Theorie liefert viel, aber dem Geheimnis des Alten bringt sie uns kaum näher. Jedenfalls bin ich überzeugt, daß der nicht würfelt.*
>
> Quantum mechanics is certainly imposing. But an inner voice tells me that it is not yet the real thing. The theory says a lot, but does not really bring us any closer to the secret of the "old one." I, at any rate, am convinced that He does not throw dice.

The lack of intuitivity in Quantum Mechanics is common knowledge in the modern world, but originally also field theories faced similar criticism. At the time, Newton was perplexed by the philosophical implications of his theory of gravitation (1692) [26]:

> It is inconceivable that inanimate Matter should, without the Mediation of something else, which is not material, operate upon, and affect other matter without mutual Contact... That Gravity should be innate, inherent and essential to Matter, so that one body may act upon another at a distance thro' a Vacuum, without the Mediation of any thing else, by and through which their Action and Force may be conveyed from one to another, is to me so great an Absurdity that I believe no Man who has in philosophical Matters a competent Faculty of thinking can ever fall into it. Gravity must be caused by an Agent acting constantly according to certain laws; but whether this Agent be material or immaterial, I have left to the Consideration of my readers.

The modern QFTs have found these Newton's Agents: they are the field quanta of the interaction fields. The Agents of the electromagnetic fields are, for instance, **photons**. Furthermore, the creation of an electromagnetic potential via photon interactions is well understood. The behavior of particles and fields can be expressed in the three-dimensional (3D) position space or the corresponding 3D momentum space. In the 3D momentum space, photons are described by inverse momentum terms of the form $1/|\vec{p}|^2$ [14]. For a charged point-particle emitting photons, this can be Fourier-transformed into a radial $1/r$-potential in the 3D position space.

It would seem that Newton's wishes have been fulfilled, but in reality the philosophical issues have only been pushed further. In terms of the physical features, the electromagnetic field is somewhat analogous to a lattice of atoms or ions. Lattice vibrations in matter are also quantized and referred to as **phonons**, which behave analogously to photons [14]. While phonons are propagated in physical matter, photons seem to be propagated in emptiness. This lattice of emptiness is called a field. Giving the seemingly **void** a name does not imply that we understand what it is. In the 19th century, various mechanical models for describing fields were entertained – unsuccessfully. The frustration with fields was famously put into words by Lord Kelvin (1884) [27]:

> I can never satisfy myself until I can make a mechanical model of a thing. If I can make a mechanical model I can understand it. As long as I cannot make a mechanical model all the way through I cannot understand, and that is why I cannot get the electromagnetic theory. I firmly believe in an electromagnetic theory of light, and that when we understand electricity and magnetism and light we shall see them ail together as parts of the whole. But I want to understand light as well as I can, without introducing things that we understand even less of.

This summarizes the status of modern physics. **A growing number of phenomena can be described, but perpetually fewer things are fundamentally understood.** Progress in theoretical physics does not seem to bring more understanding, but only more tools, and more questions on why these tools seem to work.

This could explain why some topics in modern physics – such as Quantum Mechanics and General Relativity – are also quite popular in metaphysics. Nevertheless, **if physics would have followed Lord Kelvin's philosophy of not introducing less-understood postulates, many great discoveries would not have been made**. Describing the things that are not intuitively or philosophically understood is a great challenge for science communication, but not for the progress of science itself.

2.1.2 Field Explanation for Interactions and Matter

What follows from the QFTs is that deep down, everything is quantized – all interactions and matter. It takes another leap to understand how and why the elementary particles dwelling in their corresponding fields make up the existing world. In brief, the elementary particles making up matter are fermions, and the elementary particles responsible for interactions are bosons.

In mathematical terms, fermion field operators A and B follow anti-commutation relations $\mathbf{AB} + \mathbf{BA} = \mathbf{0}$, whereas boson field operators C and D follow commutation relations $\mathbf{CD} - \mathbf{DC} = \mathbf{0}$. From these relations it follows that fermion wave functions are antisymmetric and boson wave functions symmetric. For the fermions, the antisymmetry implies the **Pauli exclusion principle**, which states that each energy state can be occupied only by a single fermion.

Practically all energy states are spatially extended, and typically the exact same spatial structure can be shared by two spin states. Thus it follows from the Pauli exclusion principle that a group of fermions is forced to produce a spatially extended structure. In contrast, one could fit arbitrarily many bosons in a spatially limited zone.

All the experimentally observed elementary particles are shown in Fig. 2.1. The fields of the elementary particles have complicated interactions with each other, which are also indicated in the figure. Some of the bosons can interact with each other, but the most important interactions are found between bosons and fermions. The interaction between two fermion elementary particles always requires a mediator boson. For instance the interaction between two electrons is most likely carried by photons, which vibrate in the electromagnetic field. An electron may interact also through the weak interaction, where the W and Z bosons have a similar role as photons in electromagnetism.

The strong interaction excludes completely the leptonic fermions (including electrons) and only considers quarks. It is mediated by **gluons**, which also interact with each other. The **Higgs boson** has somewhat analogous modes of interaction as the Z boson but the Higgs field is not primarily considered an interaction field. It has a more important role, as it allows the mass terms for particles to emerge. Moreover, the Higgs boson is a **scalar boson**, in contrast to the other elementary particle bosons, which are **vector bosons**. Here, vector refers to a particle with the absolute spin of unity, and scalar to a particle with the absolute spin of zero. Fermion

spin can attain fractional values such as 1/2, 3/2 and 5/2, but not integer values.

In summary, the feasible field interactions are the most important pieces of knowledge related to each elementary particle. The world would be drastically different, if for instance electrons would not interact with photons.

As General Relativity is not a QFT, there is no certainty about the hypothetical gravitational field quanta, **gravitons**. In particle physics gravitational effects are very small, which means that finding a particle theory for gravity is difficult, but the lack of it does not disturb other measurements. Quantum theories are important at short length scales, whereas the predictions of General Relativity operate at very long distances. Extending this concept, there is a distinct hierarchy of the important length scales for all the fundamental forces:

- Gravity dominates very long length scales, ranging from the surface gravity of a planet to the interactions between galaxies.
- Electromagnetism rules at short to intermediate length scales, e.g. from the interactions between protons and electrons to the contact interaction between a finger and a keyboard.
- The strong and weak interactions become important at short and very short length scales, i.e. at atomic and sub-atomic distances.

This explains why General Relativity is mainly a very important tool for the cosmologists. The particle physicist concentrates on the remaining interactions, which

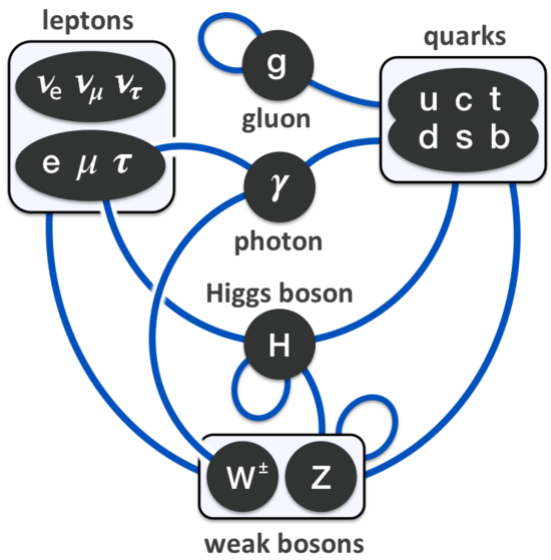

Figure 2.1: The known elementary particles and their mutual interactions (blue lines) [28].

are important at short length scales and large energies. Consequently, the most notable experiments in modern particle physics operate under the term **High-Energy Physics** (HEP).

2.1.3 The Standard Model

The theoretical understanding needed in the experiments of particle physics is extensive, even if General Relativity is excluded. All of the best current theories have been collected under the brand name **Standard Model**. Notably, the SM gives a satisfactory theoretical explanation for almost all of the observed phenomena. The elementary particles and QFTs corresponding to Fig. 2.1 are the main pillars of the SM, implying that the SM does not describe gravity at all. The SM is exactly what its name states: a commonly accepted minimal model that describes most of the observed physical phenomena. It is more of a descriptive collection of theories than an elegant theory of everything. These theories are glued together with experimentally measured parameters, such as the masses of elementary particles.

Experimental particle physics is in a perpetual search for physics beyond the SM. When science takes small steps forwards, the small new pieces of information can be incremented to the SM. There are, for instance, many currently unexplained symmetries within the SM - starting from that between the numbers of quark and lepton generations - as depicted in Fig. 2.1. Recently, there have been indications of small violations in the universality of leptons [11–13]. One possible explanation for this are the leptoquarks, which would also provide an explanation for the quark-lepton symmetry observed in Fig. 2.1.

In the SM, the electric and the weak interaction have been unified according to the Glashow-Weinberg-Salam (GWS) theory [29–31]. The theory begins with three internally symmetric fields W_i, $i = 1, 2, 3$ and another field, B. The W_i fields interact only with fermions of left-handed chirality. Handedness and spin are closely tied to each other, and they are perfectly correlated at the limit of massless particles. The W_i fields follow a $\mathrm{SU(2)_L}$-symmetry and the B field a U(1)-symmetry. The symmetry groups are not handled more closely here, as this is only a qualitative review.

At this stage the theory requires the mediator-bosons W_i and B to be massless. Moreover, the Higgs field ϕ with a non-zero expectation value is introduced. Through the interaction with the Higgs field, mass terms emerge for the W_i and the B fields. Owing to a mathematical coincidence, the W_3 field is aligned with the B field. Moreover, it is possible to rotate the linear combination of the W_3 and B fields so that one massive and one massless sum field emerge. These are the mass eigenstates of the fields, which are preferred by nature. The former is the Z boson field and the latter the electromagnetic field.

The charged $\mathrm{W}^{\pm}$ bosons emerge from the linear combinations of W_1 and W_2. The splitting of the original W_i fields to the Z and $\mathrm{W}^{\pm}$ fields is referred to as spontaneous symmetry breaking, which occurred as the universe cooled-off below a

given temperature. The symmetry-breaking triggered the Higgs mechanism, which gave the Z and $W^{\pm}$ bosons masses.

In summary, the GWS theory and the Higgs mechanism together make a connection between the massless photons, the massive vector bosons $W^{\pm}$, Z and the Higgs field scalar boson H. The Z boson otherwise behaves similarly to the photon, but its $1/r$ potential attains a multiplicative factor of $\exp(-M_z r)$[1], where M_z is the Z boson mass. The weakness of the weak interactions is due to the exponential term, which is non-negligible only at short length scales. In contrast to the Z boson, the $W^{\pm}$ bosons carry electromagnetic charge. This feature allows the coupling between charged leptons and neutrinos, and up and down type quarks.

The Higgs boson is the only known scalar boson elementary particle. Due to the left-handed chirality of the weak interactions, the elementary fermions in the SM cannot have natural mass terms. However, the non-zero expectation value of the Higgs field dynamically produces the mass terms for the fermions through Yukawa-couplings, making the SM internally consistent. The explanation of mass terms through the Higgs field indicates that the coupling of each fermion with the Higgs field is proportional to the mass of the fermion. Even so, the interaction between the Higgs field and fermions tends to be weak, as the Higgs boson is even more massive than the weak vector bosons.

The final component of the SM is Quantum Chromodynamics (QCD), which describes the strong force. QCD involves an unbroken SU(3) symmetry, which means that the strong force is carried by 8 internally symmetric types of gluons. As the strong symmetry group is not broken through the Higgs mechanism, the gluons remain massless. The strong force is somewhat similar to the electromagnetic force, but instead of the electric charge there are three different charges. The **Chromo** in QCD reflects the fact that the three different charges have been colourfully nicknamed as red, green and blue (rgb).

In their basic theoretical layout, the QFTs of the three fundamental forces are quite similar. However, diverse physical features emerge for them as a result of small initial differences. An important division line comes from the fact that the electroweak unification and Higgs mechanism creates a link between all other components of the SM, but QCD stands out as a separate island. Further differences between QCD and quantum electrodynamics (QED) are discussed in the next chapter.

Despite the SM being very successful, making predictions using it is not trivial. The dynamics of strong interactions, combined with the weak and electromagnetic interactions are highly complicated. Strong interactions can be thoroughly simulated only in non-dynamic setups – for instance in the case of a static proton. In general, such stable or semi-stable combinations of quarks (and gluons) as the proton are called **hadrons**. For most dynamical processes involving strong interactions only

[1] Here and from here on we use **natural units**, where both mass and energy have the same unit as inverse length.

phenomenological models exist. Moreover, these models are typically riddled with numerous tuning parameters. Hence, the toolbox of modern particle physics includes a great deal of **engineering** and simulations in addition to elegant theories.

The main message to give about the SM is that it is a very successful piece of theory and engineering, and very incomplete at the same time. These two facts are not in disagreement, but they underline the fact that the SM has strict limitations. The limitations start from the lacking SM-integration of gravitation, and continue to the unexplained quark-lepton symmetry. Furthermore, e.g. the experimentally confirmed non-zero neutrino masses have not yet been explained in the SM. Massive neutrinos are bound to have additional consequences, such as the possible existence of right-handed neutrinos. Right-handed neutrinos would not interact with any of the known vector bosons, and hence they could offer at least a partial explanation for dark matter. The list of the SM shortcomings continues, and each particle physicist is likely to have their own favorites.

2.2 Interactions and Renormalization

A thorough definition of mass is imperative, as this thesis is focused on the determination of the mass of an elementary particle. To understand all aspects of mass, a detour through the theory of interactions and renormalization is necessary. As described above, the Higgs field has a central role in the generation of fermion masses. Moreover, the fermion masses are also linked to the bosonic self-interactions of the fermions in question. However, these mass contributions are unfortunately found to be infinite with the current point-particle formalism.

Renormalization is an intricate process for fixing any such divergences in the mathematical theory of particle physics. Completely understanding and mastering renormalization is a tedious task mostly suited for specialists. Nevertheless, it is not possible to understand the physical definitions of mass without some knowledge in renormalization.

To proceed, a more explicit definition of the term **interaction** is necessary. The forces mediated by elementary vector bosons are often referred to as interactions. Nevertheless, these bosons are also described by fields that can interact with other fields. This confusion arises from the fact that the different fermion fields are not known to interact with each other without the mediation of vector boson fields. To summarize, the full interaction between two fermion fields consists of two or more interactions between fermion and boson fields. Thus, the boson fields could be more descriptively titled **interaction mediator** fields. From here onwards the word **interaction** will be used for both purposes, with the meaning indicated by the context.

2.2.1 Interactions Between Fields

In the following, the interaction between two fields will be mathematically described. The properties of Quantum Mechanical wave functions are accessed through operators and state vectors that express the state of the field. For instance the Hamiltonian $\hat{H}$ functions as a generic energy operator. A normalized state vector is denoted by $|A\rangle$, so that the dot product between vectors yields $\langle A|A\rangle = 1$. The energy of the state $|A\rangle$ is found by accessing the operator: $E_A = \langle A| \hat{H} |A\rangle$.

Let us define an initial state vector as $|in\rangle_{t=t_{in}}$ and a final state vector as $|out\rangle_{t=t_{out}}$ with $t_{out} > t_{in}$. We are fundamentally interested in the transition probability from a given initial state *in* to a given final state *out*. For this purpose we define an operator $U(t, t_0)$ that describes the time development of a vector state:

$$|\psi(t)\rangle = U(t, t_0) |\psi(t_0)\rangle . \tag{2.2}$$

The transition probability can now be written as

$$P = \left| {}_{t=t_{out}}\langle out| U(t_{out}, t_{in}) |in\rangle_{t=t_{in}} \right|^2 . \tag{2.3}$$

Finding a form of the operator U is equal to solving the Quantum Mechanical behavior in time. There are two main approaches for this: the Schrödinger picture described by the Schrödinger equation, and the Heisenberg picture described by an operator-based version of the Schrödinger equation. In the Schrödinger picture, time development is built into the state vectors, and the operators remain constant. Correspondingly in the Heisenberg picture, the state vectors remain constant and the time development is embedded in the operators.

The interaction picture combines these two approaches. Starting in the Schrödinger picture, the full Hamiltonian $\hat{H}$ of the interacting fields can be split into two parts: $\hat{H} = \hat{H}_0 + \hat{H}_{i0}$. The term $\hat{H}_0$ describes the generic state of the field, as $\hat{H}_{i0}$ describes the interactions of the field. In the interaction picture, the time-dependence of $\hat{H}_0$ is treated in the Heisenberg picture and that of $\hat{H}_i$ in the Schrödinger picture. We find

$$\hat{H}_i = e^{iH_0 t/\hbar} \hat{H}_{i0} e^{-iH_0 t/\hbar} \tag{2.4}$$

and

$$U(t, t_0) = T \left\{ \exp\left(-\frac{i}{\hbar} \int_{t_0}^{t} \hat{H}_i(t') dt' \right) \right\}, \tag{2.5}$$

where T is the time ordering operator. Furthermore, the time scales of Quantum Mechanics are so small that we can express the initial state to occur at $t_{in} = -\infty$ and the final state at $t_{out} = \infty$. Eq. (2.3) now reads:

$$P = \lim_{T\to\infty} |\langle out| U(-T, T) |in\rangle|^2 = |\langle out| U(-\infty, \infty) |in\rangle|^2 . \tag{2.6}$$

Here, the term $U(-\infty, \infty)$ is conventionally called the $\mathcal{S}$-matrix, for which the exponential series can be expanded:

$$\mathcal{S} = 1 + \sum_{n=1}^{\infty} \frac{1}{n!} T \left\{ \left(-\frac{i}{\hbar} \int_{t \in \mathcal{R}} \hat{H}_i(t') dt' \right)^i \right\} = 1 + i\mathcal{T}. \tag{2.7}$$

If the interaction is relatively small, the series will converge, and the leading terms become the most important ones. In the above expression the interactions are embodied in $\mathcal{T}$, as the unitary term indicates an interaction with no changes. In the common notation,

$$\langle out | \, i\mathcal{T} \, | in \rangle \propto i\mathcal{M}_{in,out}, \tag{2.8}$$

where $\mathcal{M}_{in,out}$ is the **matrix element** (ME) for the transition between the states *in* and *out*. Putting it all together, the transition probability from an initial state $\langle in|$ to a final state $\langle out|$ is

$$P(in, out) \propto |\mathcal{M}_{in,out}|^2. \tag{2.9}$$

Hence by determining the ME of a certain field interaction process, one will find the probability of the said process.

The described perturbative approach is used in situations where the exponential sum of Eq. (2.5) converges. Each term in the sum corresponds to a certain mode of interaction, which can be graphically expressed with **Feynman diagrams**. When the sum converges, the first term is referred to as the leading order (LO) contribution, and it usually contains the majority of the interaction.

The left Feynman diagram in Fig. 2.2 depicts an example LO interaction, where two electrons interact through photon transfer. In the diagram, time passes from left to right, electrons[2] are solid straight lines with an arrow and photons are wobbly lines. Furthermore, the right-hand diagram in Fig. 2.2 visualizes a related next-to-LO (NLO) interaction: the two-photon-exchange, which is much more rare than the LO process. In these LO and NLO processes the initial and final states are identical and indistinguishable. It becomes apparent that when the perturbative approach converges, the higher order terms are both growingly complex and less important.

There are also many scenarios where the perturbative sum does not converge. Furthermore, QFTs would not be at all compatible with perturbation theory without the mathematical tricks introduced in renormalization. As an alternative, a more rigorous derivation of interactions between quantum fields is due to Richard Feynman and his path integral formalism. Here, the probability for a particle to propagate from point a to point b is a squared sum of $e^{iS/\hbar}$ over all possible paths from a to b. The term S is the action of a path, which can be determined from the related Lagrangian. Only the paths that deviate slightly from the classical path contribute notably to the sum. This approach is more arduous than the perturbation theory, but does not suffer from the perturbative divergences. Moreover, the

[2] and fermions in general

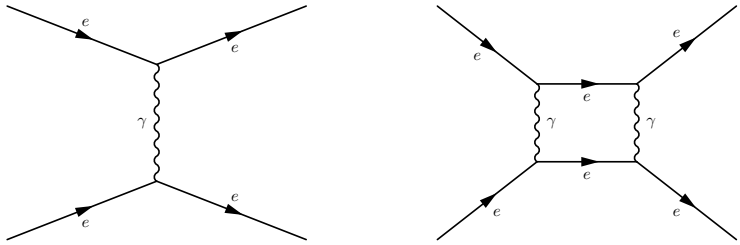

Figure 2.2: Example Feynman diagrams for the interaction between two electrons. LO (left) and NLO (right).

approach confirms many of the conventional perturbation theory results, validating the perturbative approximation.

2.2.2 Renormalization

The perturbative sum of Eq. (2.5) provides an infinite number of terms, meaning that a full interaction is not contained within the LO and NLO terms. Furthermore, some NLO (or NNLO, NNNLO, ...) interactions are indistinguishable from the LO interaction, as they have the same initial and final states. An example of this is depicted in Fig. 2.2.

It would be beneficial to automatically include the indistinguishable higher-order contributions into the LO term. This task is confronted in **renormalization** theory. The result of renormalization is highly dependent on the nature of the relevant interaction. In the definition of quark masses, the most important interaction is QCD. Renormalization for QCD is highly involved, and hence QED renormalization is handled first.

Quantum Electrodynamics

Renormalization effectively inverts the way in which the models in particle physics are built. Ideally, physical observables would be derived from the bottom up using fundamental principles. However, a more successful approach is found by starting at the top with the measured values of physical observables. The numerical values of these observables (such as elementary particle masses) are then given as parameters to the renormalized QFTs. Hence, the necessity of renormalization brings the theories of particle physics closer to the models used in engineering.

For the electric potential at the distance r, renormalization adds a first-order correction term on top of the usual $1/r$ form [24]:

$$V(r) = -\frac{\alpha}{r}\left(1 + \frac{\alpha}{4\sqrt{\pi}}\frac{e^{-2mr}}{(mr)^{3/2}} + \dots\right), \tag{2.10}$$

where $\alpha \approx 1/137$. The extended Coulomb potential presented in Eq. (2.10) is referred to as the *Uehling potential*. Notably, this expression is equal to the conventional $1/r$ form at long distances, but at shorter distances the electric field appears to be stronger. This is due to vacuum polarization, i.e. the constant creation and annihilation of electron-positron pairs in the vacuum.

The result of Eq. (2.10) can be alternatively interpreted as a slight r-dependence of the electric charge or the electric coupling. For experiments, it is thus necessary to define a **plane of renormalization**, where the physical observables are defined. For QED this is most conveniently found at $r \to \infty$, where the exponential second order term in Eq. (2.10) can be neglected. For energies, this limit corresponds to $E \to 0$. This convergence at long distances and small energies has allowed measuring the QED parameters with exceeding precision.

For large energy transfers $\sqrt{q^2}$, the electromagnetic coupling α runs as [23]

$$\alpha_{eff}(q^2) = \frac{\alpha}{1 - \frac{\alpha}{3\pi}\left\{\log\left(\frac{q^2}{m^2}\right) - \frac{5}{3}\right\}}. \tag{2.11}$$

Here m is the particle mass and the energy transfer is considered to be large when $|q^2| >> m^2$. Eq. (2.11) carries the most important implication of renormalization. The couplings between two fields are not constant, but they run with energy (or distance) - as opposed to classical physics, where the constant masses or charges define a constant coupling strength.

A common result in solid-state physics is the acquisition of an effective mass for a particle propagated in a conducting medium. For instance the effective mass of an electron can be strictly different from the free electron mass. This effect is due to electromagnetic interactions. Solid-state physics have many analogies with particle physics, and thus one could suspect that the electromagnetic interactions also have an effect on the free electron mass. It is indeed found that through renormalization, the electron mass m can be expressed as

$$m = m_0 + \delta m, \tag{2.12}$$

where m_0 is the electron mass before the impact of the electric fields and δm is the difference brought by the fields. The term δm is the self energy of the electron in the electromagnetic field. Unfortunately, [23]

$$\delta m = \lim_{\Lambda \to \infty} \frac{3\alpha}{4\pi} m_0 \log\left(\frac{\Lambda^2}{m_0^2}\right) = \infty \tag{2.13}$$

for point particles. Here, Λ is the QED renormalization cutoff scale. That is, a large auxiliary mass term, which needs to be significantly larger than any of the physically meaningful masses and energies in question.

The expressions of Eqs. (2.12,2.13) present the fundamental trade-off of renormalization. If m_0 is fixed to a finite value, the physical mass m becomes infinite. If

m is fixed to the experimentally measured value, the value of m_0 goes to $-\infty$ and becomes meaningless. Only in the latter case is it possible to construct a physically meaningful model, implying that renormalization is bound to losing insight into any underlying fundamental parameters, such as m_0.

Quantum Chromodynamics

As QFTs, QED and QCD resemble each other. Nevertheless, renormalization gives QCD properties that are opposite to QED. It is found that at the level of one renormalization loop, the strong coupling constant runs as [23]

$$\alpha_s(q^2) = \frac{\alpha_s \mu^2}{1 + \alpha_s(\mu^2)\beta_0 \log\left(q^2/\mu^2\right)} = \frac{1}{\beta_0 \log\left(q^2/\Lambda^2_{QCD}\right)}, \tag{2.14}$$

where β_0 is a constant, μ is a chosen scale where α_s is initially determined. The scale Λ_{QCD} is defined as

$$\log \Lambda^2_{QCD} = \log \mu^2 - 1/(\beta_0 \alpha_s(\mu^2)). \tag{2.15}$$

Eq. (2.14) signifies that contrary to the electromagnetic coupling constant, α_s diverges at small energies q^2 or long distances. In QED, it was possible to choose $\mu^2 = 0$. As α_s diverges when $q^2 \to 0$, there is no such natural choice of μ^2 here. The usual convention is to measure α_s at the Z boson mass scale, $\mu^2 = M_z^2$.

From Eq. (2.14) it is observed that $\alpha_s(q^2)$ does not only diverge as $q^2 \to 0$, but the divergence occurs already as $q^2 \to \Lambda_{QCD}$. Thus, QCD becomes absolutely non-perturbative at the scale Λ_{QCD}. This presents the fundamental difference between QCD and QED: even though the two share a similar QFT formalism, the perturbative approach is more practical for renormalized QED. Convergence at long distances and small energies in QED facilitates experimental setups, in contrast to the corresponding divergences in QCD.

Path integrals are a viable tool for probing the non-perturbative regime of QCD. No exact theory has been able to utilize path integrals in this sense, but a discrete numerical **lattice QCD** model has been able to formulate a non-perturbative version of QCD. This model is used to derive physical estimates directly from the QCD theory. However, mostly static models are used, as dynamical systems of elementary particles subject to QCD (quarks and gluons) bring up many complications.

In terms of the distance parameter r, lattice QCD simulations have shown that the gluon-mediated potential between two quarks has a linear r component:

$$V(r) = -\frac{1}{r} + \sigma r. \tag{2.16}$$

This indicates that the strong interaction resembles the electromagnetic interaction at short distances, but at long distances the linear term becomes dominating. If the linear term would always be present, the universe would not be stable. Fortunately,

the term disappears for **colorless** combinations of quarks and gluons. In a colorless state the counts of a color charge and its anti-color charge, or the counts of all three color charges must be equal.

By Eq. (2.16), when a quark is separated from a colorless state, a linear potential grows between it and the remaining quark. If the separation grows enough, there is sufficiently energy to form a quark anti-quark pair. Thus, completely separating a quark (or a gluon) from a colorless state is forbidden in QCD. This is referred to as the confinement of the color charge.

As a consequence of the confinement, bare quarks and gluons cannot be observed, in contrast to elementary particles only subject to QED. QCD is thus the most definitive interaction for quarks, which are subject to both QED and QCD. Furthermore, different types of quarks and gluons cannot be distinguished from each other with the same accuracy as e.g. electrons can be distinguished from muons. This motivates a common nickname for the quarks and gluons: **partons**.

The linear potential of Eq. (2.16) has been used successfully for decades as a phenomenological model in particle physics simulations. The most successful Monte Carlo (MC) simulation program for high energy collisions, PYTHIA employs this **Lund String model** [32–34]. In PYTHIA, the inter-quark potentials are modelled by strings, which can break when they become sufficiently energetic. In the model, an energetic parton donates its energy to strings, which can finally break and form additional partons. The string model faces issues when the number of quarks grows, as the number of quark-quark interactions grows as $O(N^2)$. Moreover, it is non-trivial to consider which strings are screened off by the color charges of other quarks.

Due to the divergence of α_s at long distances and small energies, it is not possible to define such properties as the mass at infinite distances. Therefore, the concept of quark mass is complicated. Theoretically speaking, it is not necessary to worry about the exact quark mass definition, as free quarks are not allowed. As long as the infinite terms are cancelled out in renormalization, one is at liberty to choose the renormalization scale, at which quark masses are defined. This freedom has brought up several **renormalization schemes**. Quark masses are scheme dependent, but by definition physical observations in the different schemes must remain the same. An early popular renormalization scheme is the Minimal Substraction (MS) scheme, where only the divergent terms were required to cancel out. A later modification of this, the $\overline{\text{MS}}$ scheme, is one of the currently popular schemes.

2.3 The Significance of the Top Quark Mass

In this section, the importance of m_t is reviewed. This review begins by the definitions and interpretations of **mass** in physics. Then, the alternative definitions of the m_t are studied. Finally, the impacts of the experimental uncertainty in the measured m_t values are discussed.

2.3.1 The Concept of Mass

Even though similar concepts were developed earlier, the first solid presentation of an **equivalence principle** came with Newton's laws and theory of gravity. The indication of this principle is that gravity and motion (inertia) are governed by the same exact thing: the mass. Newton did not have a good explanation for the origin of mass, so it simply acted as a parameter.

The next great leaps in the understanding of mass were brought by Einstein's theories in Special Relativity and General Relativity. A combination of the concepts of Special Relativity and the equivalence principle lead to General Relativity, where mass and energy are tied to the curvature of space. However, already the theory of Special Relativity brought about the famous equation, $\mathbf{E = mc^2}$, implying that mass and energy are the same thing. Here, m is the **relativistic mass**, which is proportional to the total energy of a particle or a system.

The concept of a relativistic mass is no longer favored in particle physics, and the use of the rest mass (m_0) is now preferred. In consequence, the famous $E = mc^2$ formula is transformed into a format that is closely tied to Eq. (2.1):

$$E = \frac{m_0 c^2}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{2.17}$$

This expression has the great advantage that for particles, atoms etc. m_0 is a constant. For a system at rest we have $E = m_0 c^2$, which is the connection between the rest mass and the rest energy of a particle or system. Generally, the terms *mass* and *rest mass* are now used interchangeably, and the subscript zero can be omitted.

Einstein also provided us with a simple method for calculating the rest mass. By applying Eq. (2.1), the rest mass of a particle or system can be reconstructed if the momentum and energy are measured. This equation is perhaps the single most valuable tool in the toolbox of a particle physicist.

A simple example for the use of rest mass can be found in hydrogen, the bound state of a proton and an electron. Hydrogen mass is smaller than the sum of proton and electron masses by the potential energy of the system. A more intriguing example is found in the proton. A proton consists of two u quarks and a d quark, for which the combined sum of masses is less than 10 MeV. In comparison, the complete mass of a proton is approximately 1000 MeV. Therefore, a majority of the proton mass appears to be missing. The key for this mystery is the summed-up energy of the gluons and temporary quark anti-quark pairs, which keep the three **valence quarks** u, u and d together. A comparison between the proton and the hydrogen atom gives a great example of the differences between the QED and QCD: a binding QED potential reduces mass, whereas a binding QCD potential increases it.

Any particle interacting through the field of a bosonic elementary particle has a self energy within that field – e.g. the electron has a self energy in the electromagnetic field. This potential energy should impact the rest mass of an elementary particle

similarly as in the case of hydrogren or a proton. For point particles, the semi-classical self energy within a $1/r$-potential[3] is infinite. This is in line with the infinite renormalized contribution of QED to fermion masses, found in Eq. (2.13).

The infinite mass terms found through renormalization or semi-classical self energy calculations are frequently ignored, as they are not fully understood. Nonetheless, ignoring a contribution because it is infinite is not completely justified. The infinite terms could imply that the masses of the elementary particle fermions are closely tied to the bosonic fields through which they interact. This link would also explain the smallness of the masses of neutrinos, which are only subject to the weak interaction and gravitation.

Remarkably, if elementary particles are allowed to have a small but finite size, an upper limit for Λ is found, and Eq. (2.13) becomes finite. The same result is valid for semi-classical self energy calculations. Hence at very short length scales an approach differing from the point particle formalism could be fruitful. If this was the case, ignoring the terms that are infinite in the point particle formalism would be explicitly erroneous.

In the Higgs boson era, it is commonly expressed that the Higgs boson creates the masses of the elementary particles. For the vector bosons of the weak interaction this is more accurate, but for fermions this can be misleading. The non-zero vacuum expectation value of the Higgs field **allows** the presence of mass terms in the fermion equations of motion, but it does not **explain why** the terms are there. Consequently, the fermion masses act as arbitrary parameters, describing the couplings with the Higgs field. In a theory more complete than the current SM, a deeper explanation for the magnitudes of the masses could be provided.

2.3.2 Interpretation of the Top Quark Mass

In the preceding section the nature of quark masses was explored through renormalization. In particle physics, the rest mass constructed through Eq. (2.1) is frequently referred to as **pole mass**. For fermions not subject to QCD the pole mass is well defined, but for quarks this is not the case. This follows from the lack of a natural fixed renormalization scale in QCD and hence quark masses must be handled e.g. within the $\overline{\text{MS}}$ scheme.

Compared to the other quarks the top quark mass is exceptional since the experimentally observed m_t-resonance behaves in an analogous manner with resonances originating from well-defined pole masses. In other words, something resembling a pole mass resonance can be reconstructed from the top quark decay products. This is the starting point for the most precise m_t studies: even if theory does not favor a pole-mass treatment of the top quark, experimental data and evidence leads to reconstructing the mass of the top quark using its decay products. The current most precise m_t measurements performed with this approach typically yield a mass

[3] That is, Coulomb's potential equated with $E = mc^2$.

value around 172-173 GeV (with the exception of DØ) [19, 20, 35, 36]. Also the m_t analyses presented in this thesis follow the same conventions.

One of the secrets behind the pole-like behavior of m_t is in the brevity of the top quark lifetime. The value of m_t is approximately as large as the mass of a single ytterbium atom, or equivalently over twice as large as the W boson mass. This makes the top quark more massive than all the other known elementary particles. Consequently, the weak interaction is not *weak* for the top quark. A rapid weak decay makes the top quark lifetime extremely short, around 4×10^{-25} s [14]. This means that even if a top quark is traveling at a speed close to the speed of light, it will decay after advancing approximately 1.2×10^{-16}m. This is a great contrast to the semi-stable particles observed at the CMS detector, which must be able to travel several millimeters before decaying.

Interpreting the extremely short top quark lifetime through the Heisenberg uncertainty principle, QCD interactions below the order of 1 GeV should be effectively suppressed for the top quark. On the other hand, the perturbative QCD instabilities mainly arise around and below the order of Λ_{QCD} (< 1 GeV). Hence, using the Heisenberg interpretation, the top quark escapes the ordinary full QCD confinement by decaying extremely briefly. The full explanation of the non-perturbative QCD features of the top quark is more complicated, but the uncertainty principle offers one simple method for understanding the special features of the top quark. Another explanation for the top quark dynamics is pointing out that the top quark mass is simply in a much higher energy scale than the typical non-perturbative QCD effects, making the relative magnitude of QCD effects much smaller than for the other quarks.

In any case, the above offers some explanation on why the top quark mass can be conveniently treated similarly to a pole in contrast to the masses of the other quarks. In consequence, the use of the $\overline{\text{MS}}$ mass is disfavored in top quark studies, as the $\overline{\text{MS}}$ m_t value is strongly shifted from the kinematically favored one. The pole mass approach is not free of obstacles, either: in the vicinity of 1 GeV, (self-)interactions can introduce instabilities. The optimal solution for this is the use of the MSR scheme mass with the R parameter fixed approximately at 1 GeV. In practical terms, MSR can be viewed as an interpolation between the pole mass and $\overline{\text{MS}}$ mass schemes, with $R \to 0$ corresponding to the pole mass. The choice $R \propto 1$ GeV picks up the positive features of the pole mass scheme and cuts off the sub-GeV instabilities. Relating the effective top quark pole mass to the MSR scheme mass requires intensive work from theorists. [37–39]

The other consequence of the short lifetime of the top quark is the fact that it is not known to form bound states. Quark bindings require time-scales that are accessible only in the non-perturbative (sub-GeV) regime of QCD. To summarize, the behavior of the top quark is much more explained by the weak than the strong interaction, in contrast to the other quarks.

An additional complication with the top quark mass is the way in which it is modeled in simulations. Originally, experimentalists tended to think that the top

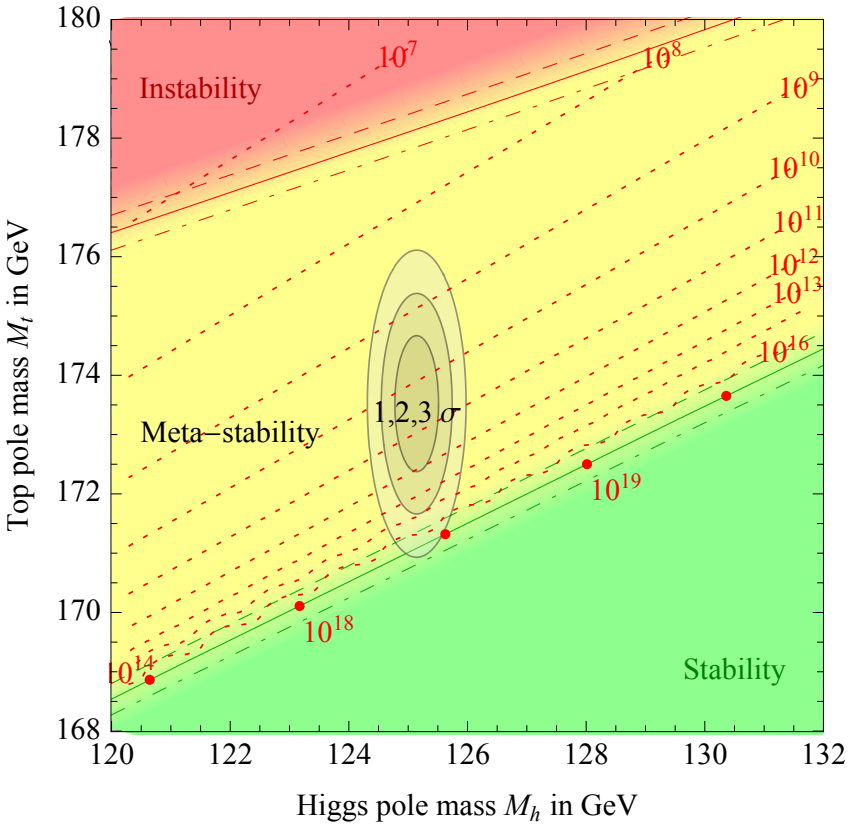

Figure 2.3: Metastability vs. m_t and Higgs boson mass [15].

quark mass parameter in the simulations can be interpreted as something resembling the effective top quark pole mass. Following this reasoning, the simulated mass parameter (m_t^{MC}) has been used for the top quark mass extraction. This interpretation has turned out to be slightly incorrect, and a correction for the m_t^{MC} value is necessary. Currently at slightly below a 0.5 GeV uncertainty, m_t^{MC} appears to be consistent with the MSR mass [38, 39].

The work on relating the different mass definitions with each other has progressed well in recent years. CMS is yet to make studies on the relationship between MSR and m_t^{MC}, which was recently accomplished by ATLAS [39].

2.3.3 The Fate of the Universe

In the paper [15] by Buttazzo et al. the concept of the electroweak vacuum metastability was introduced. The concept was developed further in [40]. The most important conclusion of the paper was that within the experimental error limits of the top quark mass, it is unknown whether the universe is stable or metastable. Also for instance the Higgs boson mass is connected to the same result. This is depicted

in Fig. 2.3.

The Higgs boson mass has been measured with better precision, leaving m_t as the most important source of uncertainty in the metastability-question. Many of the most accurate measurements of elementary particle masses are based on the measurement of decay products, explaining the relatively worse measurement accuracy of m_t. The Higgs boson can decay e.g. to a pair of muons, which can be accurately measured. In contrast, the top quark decay always produces quarks, due to which conducting an accurate direct measurement is demanding.

The metastability of the universe could imply that the current physical laws will not remain static during very long timescales. An alternative explanation for the apparent metastability could be a yet unknown elementary particle with a mass around the metastability energy scale. The currently most intriguing LHC Run 2 results have been found in precision b quark physics [11], adding to the almost simultaneously released FNAL results [12]. These findings could indicate the existence of a new elementary particle and a new interaction. Moreover, they emphasize the role of precision measurements in contemporary particle physics. Hence, the metastability and the search for signatures of new physics are highly relevant questions.

3 Data Collection and Simulation at the CMS

CERN is currently the largest particle physics research center in the world, hosting multiple experiments and collaborations. The CERN site is physically located at the rural outskirts of Geneva, placed directly on the border of France and Switzerland. CERN was originally established as a European research center and most of the full member states are still European. However, also countries from other continents are now running for membership (e.g. India and Pakistan). Moreover, numerous countries outside Europe have co-operation agreements with CERN. Consequently, most of the work for CERN is performed around the globe in numerous institutions. In contrast, the main function of the physical CERN site has become providing a location for the experimental apparatus and a place where particle physicists from around the world can gather to work together. [41]

3.1 Experimental Setup

The experimental setups in particle physics are typically dichotomous: they consist of a particle acceleration system and a particle detector. The size of modern accelerators is typically of the order of kilometers, and hence it is economical to place multiple detector experiments along them. The majority of this thesis is focused on the data gathered by the CMS experiment. Hence in the following, the technical details of the LHC accelerator and the CMS detector apparatus are reviewed. Where no other references are given, this chapter is based on the detailed LHC and CMS technical reports of Refs. [42, 43].

To reduce clutter, the term **data** will from here on be used to refer to the (real) CMS detector data. At times the word *data* is also used in its original meaning, which can be understood from the context. The counterpart for data is simulated data – or more briefly simulation or MC (referring to Monte Carlo simulation).

3.1.1 The CERN Site and the LHC

Particle accelerators are commonly based on the same basic principles: charged particles can be accelerated and guided using electromagnetic fields. The two most common topologies for particle colliders are linear and circular. In the modern accelerators particles travel very close to the speed of light. Bending the trajectories of such particles requires strong magnets and causes energy losses through *brehmsstrahlung*, which is a challenge for the circular accelerators. Linear accelerators are free from this limitation, but they are limited by the length of the apparatus. In a circular accelerator a charged particle can travel an almost arbitrary amount of loops, and the acceleration process can hence be more prolonged.

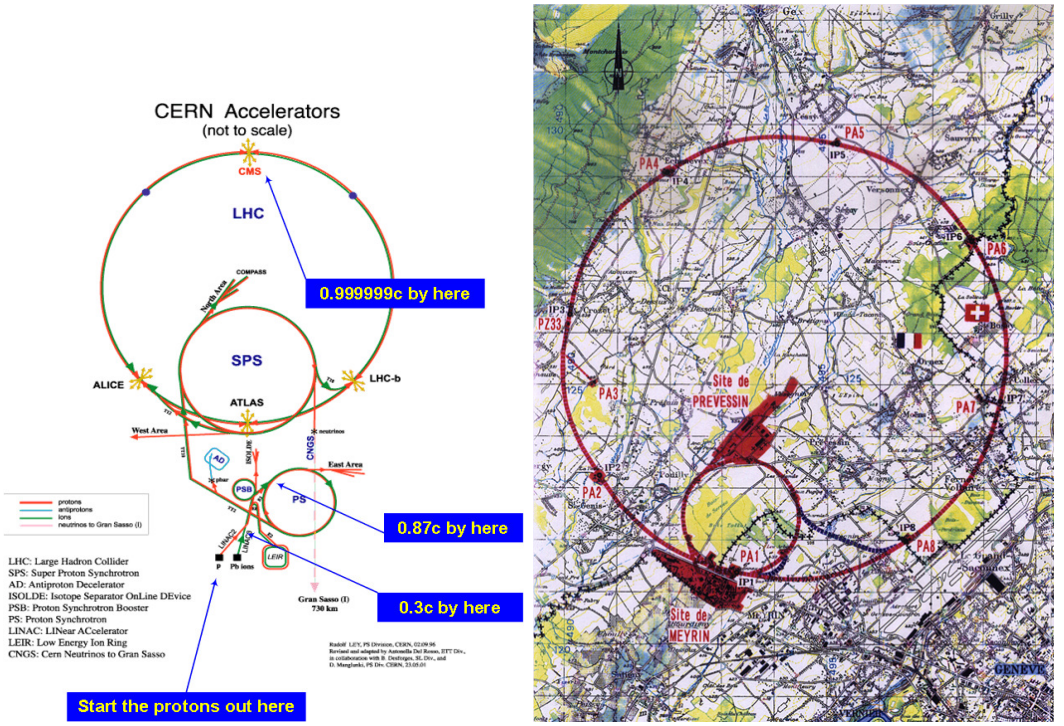

Figure 3.1: The LHC in a schematic figure (left) [44] and in a map illustration from 1998 (right) [45].

The most powerful current accelerators are circular, including the LHC. However, there have been plans also for new linear colliders, such as the ILC [46]. The LHC is currently the most prominent accelerator. It was built into the tunnel of the Large Electron Positron collider (LEP) after LEP stopped operating in 2000. One of the reasons for the LHC being circular is that the LEP was circular. It started operating in 2008, with its first and foremost mission being the discovery of the Higgs boson. This goal was met in 2012, after which the machine entered the first long shutdown period. The LHC operated again during the years 2015–2018, after

which the second long shutdown period began. Now in 2022, the shutdown period has ended and Run 3 has successfully started [47].

Circular accelerators require one or two parallel vacuum pipes depending on the accelerated particle types. If a particle and its antiparticle are being accelerated, one pipe can suffice. This follows directly from the properties of electric fields: opposite charges are accelerated in opposite directions. Otherwise two pipes are necessary, with crossing points placed at the detector sites. Thus being a proton-proton (pp), proton-lead and lead-lead accelerator, the LHC consists of two pipes with an inner diameter of 50 mm. Acceleration in the pipes is handled at the radio frequency cavities, placed at a single site along the LHC circle, with electric fields of 5.5 MV/m. The particle beams are guided kept together with powerful magnets. [42]

Fig. 3.1 depicts the 27 km LHC tunnel and the major experiments along it. The LHC is not exactly a circle, as it slightly resembles an octahedron. Four major particle detector experiments reside at the beam crossing sites along the LHC. Of these, CMS and ATLAS are the only multipurpose experiments, and they are also the largest two CERN collaborations. Two experiments making similar measurements reduces the effects of possible systematical errors in major findings. The other two major experiments are built for special purposes: the LHCb experiment is dedicated to b quark physics and ALICE studies lead-ion collisions.

The LHC circle is only the last step in the acceleration process of the proton pairs. The acceleration begins at the Linear Accelerators (LINACs) A and B and continues stepwise to the Proton Synchrotron Booster (PSB), the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS). At each step, the protons are accelerated to a higher energy, closer to the speed of light. Finally after the SPS, they are fed to the LHC pipes. These steps are depicted in the left-hand side image of Fig. 3.1. [44]

The LHC accelerates protons in sequential bunches. Hard pp collisions occur when the bunches collide at the interaction point, where the two proton beams cross. Collisions occur when the proton beams are focused tightly by electromagnetic lensing to the interaction point. During Run 2 of the LHC, a pp collision energy of $\sqrt{s} = 13$ TeV was utilized. As a reference, the Run 1 energies were $\sqrt{s} = 7$ TeV and 8 TeV, and the Run 3 energy is $\sqrt{s} = 13.6$ TeV [48].

Multiple collisions occur at each bunch crossing. During LHC Run 2 this number ranged approximately from 10 to over 70, as depicted in Fig. 3.2. From the perspective of a single pp collision, the other simultaneously present collisions are referred to as pileup.

Most pileup comes from the same bunch crossing, but also crossings adjacent in time can interfere. Currently, bunch crossings occur every 25 nanoseconds [49]. Achieving more frequent crossings is challenging, as the pileup from the adjacent crossings becomes too dominant. The collider experiments make their best attempts to clean the data from pileup, but typically all pileup contamination cannot be removed. The greatest pileup challenge is posed by neutral particles, for which a precise path reconstruction is non-trivial.

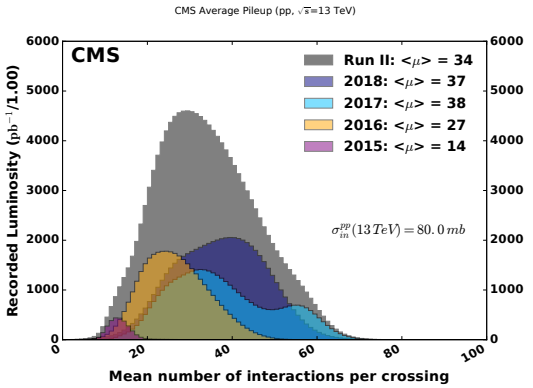

Figure 3.2: Pileup distributions during the LHC Run 2 [10].

3.1.2 The CMS Detector

The CMS detector has been designed to catch signals from most of the particles depicted in the SM. It has a cylindrical onion-like structure where different layers serve different purposes. The detector structure is not entirely static due to the wearing out of components, which are usually replaced during long shutdown periods. In addition to one-to-one component replacements, also upgrades to the detector are constantly being performed. For instance, between 2016 and 2017 the important Phase 1 upgrade was carried out at the CMS [50]. This included most importantly hardware enhancements for the pixel detector, which are discussed in more detail below.

The position within the cylindrical detector is defined by an azimuthal angle ϕ and by a polar angle θ. The **poles** ($\theta = 0$ and $\theta = \pi$) correspond to the original proton beam directions. In experimental particle physics, the theoretically best-motivated variable for describing the θ direction is **rapidity** (y). In high-energy pp collisions, observed particle counts per differential changes in rapidity are approximately constant. For massless particles rapidity is equal to **pseudo-rapidity**, defined as

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right). \tag{3.1}$$

Pseudo-rapidity has a simple one-to-one correspondence with θ, whereas rapidity is in addition dependent on the particle mass. There is often notable uncertainty related to the masses of the reconstructed particles, and hence pseudo-rapidity is commonly preferred over rapidity.

Detector calibrations are linked to the detector geometry, and hence also the mass/energy-independent mapping between η and θ becomes more important than the theoretical advantages of y. The approximation $\eta \approx y$ is reasonably good, and thus a somewhat even population of particles is found in the $\eta - \phi$ plane. This plane is in many cases convenient for studying collision event statistics.

Fig. 3.3 displays the layers of the CMS in a cross-section image depicted in the plane transverse to the proton beam. To add further understanding, Fig. 3.4 illustrates the longitudinal detector cross-section. In Fig. 3.3 a simple understanding is also given on the interactions of the detector layers with different types of particles.

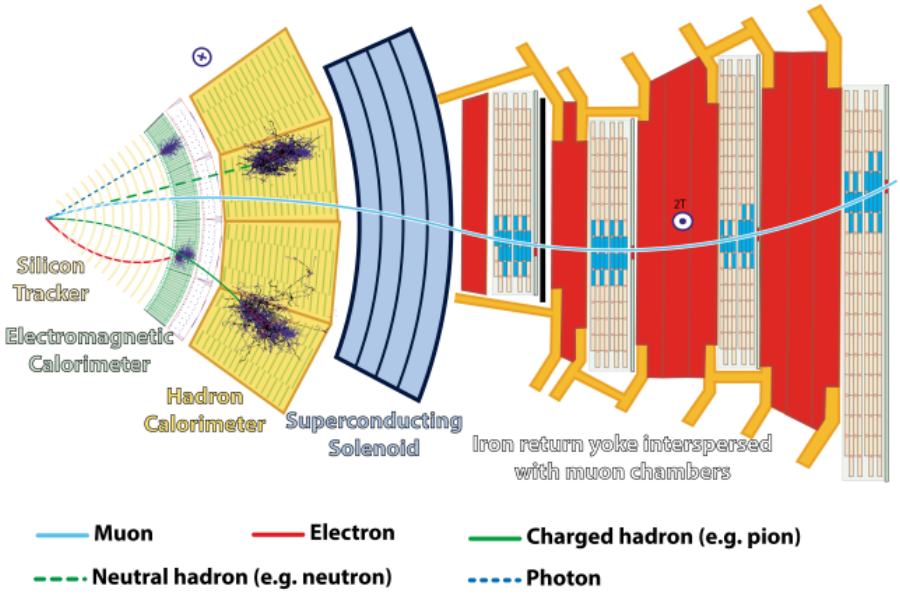

Figure 3.3: A cross-sectional image of the CMS in the transverse direction, with the proton-proton collision occurring at the left edge. Here all the subdetectors are displayed, as well as the interactions of different particle types with the detector. [51]

Following the cylindrical symmetry of the detector, also the momentum vectors are studied in cylindrical coordinates. Generally, the 3D momentum consists of a transverse 2D component $\vec{p}_T$, and a perpendicular axial component p_z. The collision events and the CMS detector are both symmetrical in the ϕ-direction, which makes p_T a more informative variable than its components p_x and p_y.

In the following, a brief tour of the layers will be given from the beam axis outwards. The description is given mostly for the **barrel** region of the CMS detector, i.e. $|\eta| < 1.3$. In the complementary region $|\eta| > 1.566$ one finds the **endcaps** [53], leaving the barrel-endcap **transition region** in between (i.e. roughly $1.3 < |\eta| < 1.566$). In the endcaps the layers follow a similar logic as in the barrel, but in the transition region the detector coverage is more limited due to the cylindrical geometry and cabling.

The beam pipe containing the vacuum where the proton bunches collide is located in the center of the cylinder. The innermost sector of the actual detector is the silicon **tracker** (or **tracking detector**), which consists of multiple sub-detectors. The innermost part of the tracker is the **pixel detector**, which is made up of layers, each of which excels in determining the 2D $\eta - \phi$ positions of charged particles.

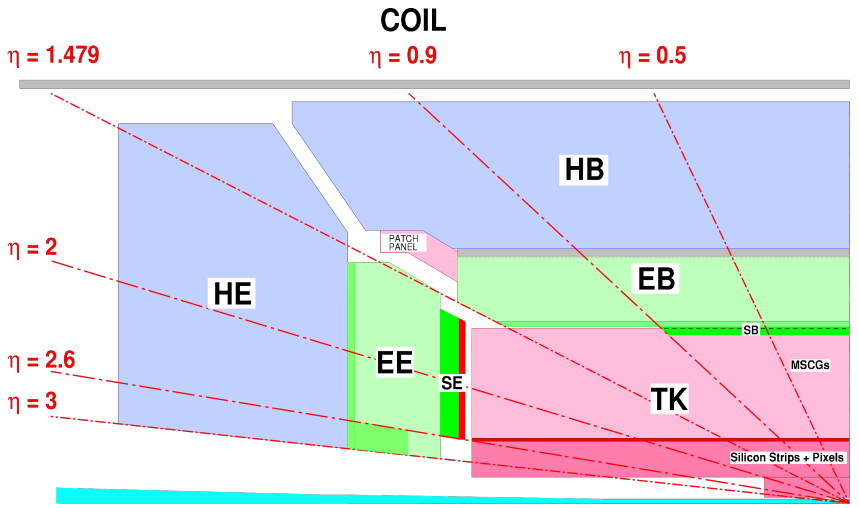

Figure 3.4: A cross-section of the CMS detector [52]. The x-axis is the cylindrical axis and protons collide in the lower right corner. EB/EE = ECAL Barrel/Endcap, HB/HE = HCAL Barrel/Endcap, COIL = Magnet, TK = Tracker. Muon chambers are outside of the image.

Combining the information from the pixel layers a 3D track can be reconstructed. The individual pixels in the pixel sensors have dimensions of 150 μm by 150 μm [54].

The pixel detector originally consisted of 3 silicon layers, but a 4-layer redesign was installed in the Phase 1 upgrade at the end of 2016 [50]. The Phase 1 upgrade imposed a gap between the 2016 and 2017-2018 Run 2 datasets, where the detector features were significantly changed. This is seen e.g. in electron reconstruction, which is reliable up to $|\eta|< 2.5$ in 2017-2018 as compared to $|\eta|< 2.4$ in 2016.

After the Phase 1 upgrade, the barrel region pixel layers range from 29 mm to 160 mm from the proton beam axis. This is a great contrast to the status before the Phase 1 upgrade, with the pixel layers at 44-102 mm. These length scales are also used in the simulations: particles that can travel more than 10 mm at the speed of light before decaying are considered stable or semi-stable. An even more drastic change occurred in the forward-region ($|\eta|> 1.5$) of the pixel detector: now there are four disks instead of two. [50]

The pixel detector is followed by the **silicon strip detector**. The strip detector is divided in the inner and outer sections, which span to 635 mm and 1200 mm from the proton beam axis, correspondingly. The inner strip detector has 5 layers and the outer strip detector 6 layers in the barrel region [54].

Each layer of silicon strips is able to measure charged particle positions only in one dimension. The strips in different layers are positioned at varying angles, so that more complete information on the charged particle positions is obtained. The combined pixel and strip information is used to reconstruct the 3D **tracks** of

charged particles. The detection of charged particle tracks is limited approximately to $|\eta| < 2.5$ by the spatial coverage of the strip tracker.

The majority of electrons and photons are stopped in the Electromagnetic CALorimeter (ECAL), which surrounds the tracker. It is built from lead tungstate ($PbWO_4$) and consists of towers with approximate front face dimensions of 22 mm by 22 mm. In the $\eta - \phi$ plane this corresponds to 0.0175×0.0175 in the barrel region, but this increases gradually towards 0.05×0.05 at higher $|\eta|$ values. This makes the $\eta - \phi$ resolution of the ECAL more coarse than that of the tracking detectors. [55]

The aforementioned conservative transition region range of $1.3 < |\eta| < 1.566$ is reduced to $1.4442 < |\eta| <= 1.566$ [53] when only the ECAL is considered. This is a cut commonly used for electrons and photons, as they are poorly reconstructed in this region.

The Hadronic CALorimeter (HCAL) envelops the ECAL with the aim of measuring the energy of the hadrons. It is built from alternating layers of 5 cm thick copper plates and 4 mm thick plastic scintillators in the barrel region [56]. It has the $\eta - \phi$ granularity of 0.087×0.087, so that one HCAL tower corresponds to 5x5 ECAL towers [56].

Beyond the HCAL resides the superconducting solenoid magnet, from which the third letter of the abbreviation CMS is derived. The magnetic field is the strongest inside the magnet (4 T), but it is also notable (1.7 T) in the detector section following the magnet [57]. The magnetic field is confined by a **steel yoke**, which makes up a majority of the 14,000 tonne weight of the CMS.

The final and only layer encircling the magnet are the muon chambers, which detect muons through gas ionization. There are three different technologies used in the muon chambers: drift tubes, cathode strip chambers and resistive plate chambers. The chambers are not designed to stop the muons, but to measure their trajectories, completing the information from the tracker. Moreover, they help in distinguishing muons from charged particles that do not reach the muon chambers – such as electrons. [58]

3.2 Data Flow: from the Detector to Physics

Measuring and interpreting information from a pp collision is a complex task, which begins at the detector hardware. The detector components collect information of passing particles and translate it into electric signals. An ideal detector would catch all particle information without losses, but this is not achievable in practice. Moreover, the wearing out of ECAL and HCAL decreases the detector's ability to measure energy deposits of particles. This must be compensated for by utilizing time-dependent calibrations. The reconstruction of a full high energy particle collision is dependent on the calibration of all the subdetectors.

A fully interpreted collision event consists of objects with physical meaning, such as solitary photons and electrons. At the CMS, the interpretation task is mostly

handled with the Particle Flow (PF) algorithm [59–61]. Most CMS physics analyses are based on the PF reconstruction of the detector signals.

Producing a detailed understanding of particle collisions is only made possible by the work of numerous scientists. Few of them have comprehensive knowledge in all the involved steps, but having a holistic view of the entire route from the detector to PF objects is important to all analyzers. Only so, the general features of the processed data can be correctly understood. In the following subsections, the steps required for reaching the PF interpretation are clarified.

3.2.1 The Hard Collision

Detector signal interpretation commences at having prior knowledge on the pp collision dynamics. The nature of a collision is reflected directly in the momentum directions of the outgoing particles. If particle content is mainly found in the direction of the original proton beam, the interaction between the protons was likely superficial, leaving the parton content of the protons unchanged.

On the contrary, when decay products appear in the transverse direction, a **hard collision** between the protons has likely occurred. This is motivated by the fact that the protons have originally no momentum in the transverse direction. In a hard collision there are significant interactions between the partons, causing the colliding protons to breach. Only such proton-breaching collisions can introduce significant transverse momentum components. Thus in the CMS data, high p_T objects act as signals of interesting physical processes.

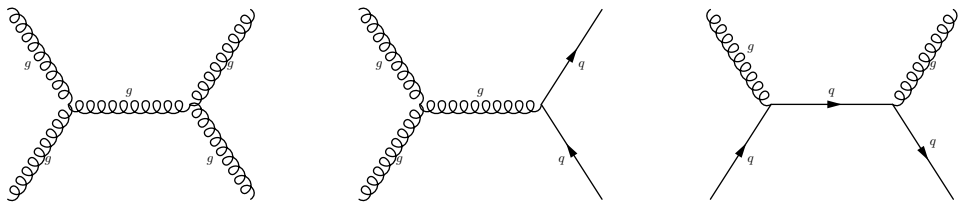

Figure 3.5: Frequent hard processes at the CMS: gluon pair to gluon pair (left), gluon pair to quark pair (middle), gluon and quark to gluon and quark (right).

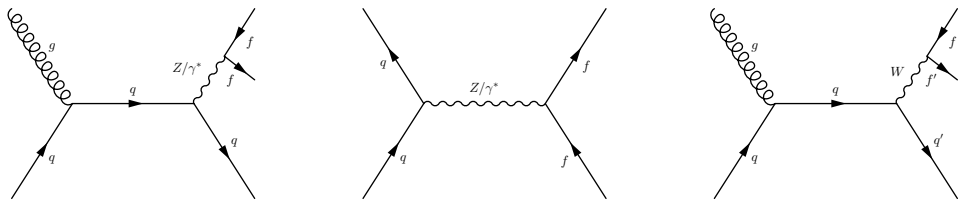

Figure 3.6: Boson-production hard processes at the CMS: quark and gluon to quark and Z/γ (left), quark pair annihilation to fermion pair production (middle), quark and gluon to quark' and W (right).

A single hard collision consists of numerous interactions between particles. In simulations, these are ranked either according to the interaction energy or the involved p_T. The highest-profile interaction is assigned the title **hard process**. Even if this concept originates from simulations, it is useful also in the interpretation of data. The hard process determines the particle content observed in the detector.

A hard process can consist for instance of two light quarks colliding and producing two new light quarks. In this thesis such a collision event with only QCD dynamics is referred to as a **Multijet** event. Even if Multijet-like topologies can be produced in other than pure QCD processes (such as fully hadronic $t\bar{t}$), the QCD-based cross-sections are vastly larger at the LHC compared to any of these alternative processes.

At the LHC, the two final state quarks can also be top quarks instead of light quarks. Energetically this is made possible by the high collision energies of the LHC, which correspond to over 10^4 proton masses[1] or over 75 top quark masses[2]. In comparison, the top quark mass corresponds only to around 1.8×10^3 proton masses.

Frequent hard process initial states observed at the LHC include collisions between gluons, and those between a quark and a gluon, while hard processes between two quarks are less frequent. Common examples of these kinds of hard processes leading into the production of quarks and gluons are presented in Fig. 3.5. The collision between a quark and a gluon can lead to the production of a (virtual) photon or a Z or W boson. Such example hard processes are depicted with Feynman Diagrams in Fig. 3.6.

3.2.2 Hadronization and Jets

Energetic quarks tend to decay and radiate generating more quarks and gluons. This process can continue until the particle energy reaches the order of 1 GeV – corresponding to proton mass and Λ_{QCD}. As the partons are no longer multiplied, hadronization occurs. This refers to a spontaneous process where partons are self-organized, forming stable or semi-stable hadrons.

The most common hadrons observed in pp collisions are made up of two quarks (mesons) or three quarks (baryons). The mesons and baryons observed at the detector consist mostly of the three light quark flavors: u, d and s. Top, bottom and charm quarks all decay before they are able to reach the detector [62], which is approximately 10 mm apart from the proton beam [50].

Hadrons can remain stable or decay producing other hadrons, photons, electrons and muons. Even though the original partons produced in the hard process are split into a multitude of hadrons and other particles, these multitudes retain collimated shapes. The more energetic the original parton was, the stronger collimation is observed. According to the conservation of energy and momentum, the collimated

[1] The proton mass is equal to energy of 0.938 GeV [62].
[2] 172–173 GeV [62].

shower of particles continues travelling in the same direction as the original parton. Such a particle shower is referred to as a **jet**.

In summary, even if free partons cannot be observed, the momentum imprint of high energy partons manifests as jets, which are distinct and measurable. The reconstruction of jets from observed particles is a complex task, which is considered in Ch. 4.

3.2.3 Triggers

The CMS – as all modern particle detectors – is able to store only a fraction of all observed collision events. For this purpose the CMS experiment utilizes triggers that determine which events are stored and which are not. The triggering is performed in two phases: first at the Level 1 (L1) hardware-based trigger and then at the High-Level Trigger (HLT), which is software-based. The L1 trigger processes events at an input rate of 40 MHz and an output rate of 100 kHz. At the HLT the final output rate is decreased to less than 1 kHz. [63]

The triggers operate early in the data processing chain to relieve the CMS servers from being flooded with data. They are able to quickly sort the recorded collision events according to specific signatures. For instance, classification according to the count of solitary leptons or photons above a given p_T threshold can be made. A single trigger is rarely able to select only a single hard process and some hard processes need to be analyzed using several triggers. This must be taken into account later in the analysis.

Sometimes a process takes too much bandwidth from the maximum possible rate to store the events, and the rate must be artificially lowered without tightening the object selection. In the CMS terminology, a dataset where less than 100% of the events are stored, is called **prescaled**. For instance, a prescale of 1000 means that the data acquisition system saves only one such collision event in a thousand. Depending on which trigger paths are activated, collision events are saved at different rates. This allows storing also rare events with decent statistics.

3.2.4 Vertices

The interaction point where two partons collide is called the Primary Vertex (PV). By matching the momentum vectors of particles associated with the PV, it is possible to separate the particles coming from the hard collision from those coming from pileup. An exception is made by neutral particles, for which the momentum direction measurements are not accurate.

Some unstable particles – most importantly b and c quarks and tau leptons – tend to decay shortly after leaving the PV by emitting a W boson [62]. Such a decay point can be distinguished from the PV, and is referred to as a Secondary Vertex (SV). The SVs serve as secondary origins for particles observed at the detector. One pp collision contains exactly one PV, but there can be zero or multiple SVs.

Utilizing the SV information and some supplementary properties, it is possible to efficiently distinguish e.g. energetic b quark jets from jets with different parton-level origins. This has become an important task at the modern collider experiments, and is referred to as **b-tagging**. There are similar efforts with c and tau tagging, but currently b tagging tends to be the most widely used and successful SV application.

3.2.5 Particle Taxonomy and the Particle Flow Algorithm

The structure of the CMS creates a distinct hierarchy for detecting particles produced at a pp collision. This structure is closely followed by the PF algorithm. The algorithm produces multiple hypothetical interpretations of a single event and the particle interpretations are labeled as PF candidates. The CMS is able to differentiate five main PF candidate categories: photons, electrons, muons and charged and neutral hadrons. The same detector deposits can have alternative PF candidate interpretations e.g. as a photon or electron. The analysts must make further identification and selections to treat the alternative hypotheses. In the following, the measurement of charged particles is first reviewed, followed by the measurement of neutral particles (photons and hadrons), and lastly the indirect estimation of neutrinos momentum is discussed.

The tracker measures the tracks of charged particles, which are bent by the magnetic field. This bending assists in momentum and charge measurements. As a rule, the charged particles are observed in one subdetector in addition to the tracker. Electrons are completely halted and measured in the ECAL, whereas HCAL captures charged hadrons. Furthermore, the flight of the muons continues through the whole detector, and they are also observed at the muon chambers.

Also photons can produce tracks despite being neutral particles. Photons tend to decay into electron-positron pairs roughly 50% of the time within the tracker [64, 65]. In the ECAL, the behavior of electrons and photons is exactly the same: they appear as an electromagnetic cascade. Photons are split into electron-positron pairs, while electrons and positrons emit photons. For these reasons, the PF algorithm is sometimes unable to distinguish electrons from photons.

In summary, the CMS detector performs well at observing charged leptons. Muon information combined from the tracker and muon chambers is highly accurate. The pure muon signal is mainly contaminated by leptonically decaying taus. The electron momentum is measured in the tracker and the ECAL yielding accurate numbers, but the momentum measurement and identification are more difficult than with muons. Furthermore, the PF electron candidates suffer from photon contamination in addition to the leptonic tau decays.

ECAL energy not associated to tracks is interpreted as photons. Analogously, unaccounted HCAL energy is interpreted as neutral hadrons. Thus, Particle Flow must distribute ECAL energy mainly between electrons and photons - and HCAL energy between charged and neutral hadrons. The HCAL is much more coarse than the ECAL, making the hadron calibration and interpretation a complicated task.

From the low directional resolution of the HCAL it also follows that matching neutral hadrons with PVs or SVs is relatively inaccurate. This makes neutral hadrons the particle category most prone to pileup contamination.

The neutrinos remain undetected by the CMS, and they can only be observed indirectly through **missing transverse momentum** ($\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$)[3]. The quantity $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is a 2D vector in the transverse plane [90]. Following from the conservation of momentum, the transverse momentum sum should be zero, and a deviation from this can indicate the presence of neutrinos.

Large $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ values can also indicate the presence of other poorly observable particles, such as dark matter. Only when a single energetic unobserved particle is present, the $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ 2D-vector gives accurate information on its transverse momentum. If multiple unobserved particles are present, the value of $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ becomes a vector sum of their transverse momenta, which is difficult to divide between the individual particles.

The PF algorithm relates momentum and energy using the known masses for the particle categories. For photons, electrons and muons, this relationship between momentum and energy is accurate. For hadrons the situation is more complicated, as the CMS detector cannot differentiate most hadron types from each other. For this reason all hadrons are assigned with the masses of the most common hadron types. For charged hadrons this means pions and for neutral hadrons eta mesons [59–61, 66].

3.2.6 Detector Data as an Inversion Problem

While the CMS detector captures a snapshot of particles produced in high energy collisions, the fundamental target of many physics analyses is in interpreting the underlying hard processes. The CMS detector can be compared to a microscope that takes images of the hard process, but this analogy is somewhat misleading. A microscope utilizes e.g. visible light or electrons to probe the surface layers of a static target. In contrast at the CMS hadrons are used to breach hadrons, the decay products of which are measured. There is no similar probe-target system present as in conventional microscopy.

A more fitting analogy for the CMS detector would be computer tomography (CT). In CT, a static target is illuminated and imaged with x-rays from multiple directions. This introduces an inversion problem, where a 2D or 3D image is reconstructed from the ensemble of snapshots using mathematical and computational methods. Even this analogy suffers from the fact that CT images are taken of a static target, whereas the CMS detector records dynamic collision events. [67]

As the target of the CMS detector is dynamic, gaining understanding of the hard processes necessitates recording numerous pp collisions, which can be interpreted only in a statistical sense. The inversion from a detector-level image to the hard

[3] Also called **MET** in older reference material, but the use of this term is now discouraged.

processes is made possible by simulations. The simulations are built on the best current knowledge in theoretical and experimental particle physics. In simulations the hard processes are known, and hence understanding of the hard processes in data can be found through comparisons with simulations.

To find the relevant hard processes behind a data dataset, it is often necessary to find a set of simulations that describes the data well enough. This set can also include signals of hypothetical new particles, as was the case with the Higgs boson discovery at the CMS and ATLAS. In some cases also data-driven methods are applicable beside simulations.

The simulations involve steps where only approximate or phenomenological models are available. Hence, the comparison between simulations and data is meaningful only after heavy tuning and calibrations. Otherwise, a mismatch in the comparison can be an indication of insufficient calibrations, or unaccounted error sources.

With a selected set of calibrations, it is possible to enhance the match between simulations and data without biasing other measurements. These are focused on well understood SM processes, where statistics are high. The resulting machinery is very effective, but heavy to operate.

The use of simulations necessitates considering systematical modeling uncertainties as a part of the full set of uncertainties in any measurement. These can be related e.g. to the parameter choices and tuning uncertainties. The objective of tuning and calibrations is reaching a sufficient level of agreement between data and simulation. Considering all the relevant modeling uncertainties appropriately can be challenging, however. At the worst e.g. overfitting can occur in the tuning and calibrations, possibly erasing weak signals of new physics.

3.3 Computation and Simulation

Historically, the CERN sites have served a dual role: providing a site for the experimental apparatus, and offering office space for the people analyzing the data. High-speed internet connections have changed this image drastically: it is now possible to perform real-time analysis anywhere around the globe. Following this development, most of the CERN analyses are now performed outside the physical CERN site.

Data centers and computing facilities have become an important part of the experimental setup in the internet era. Some of these are located at the CERN site, but most are distributed around the world. The first layer of many modern analyses is often performed utilizing the infrastructure of the worldwide LHC Computing Grid [68]. The Grid allows a dataset of interest to be stored and analyzed at any computing site employed by CERN. The datasets can be in size around 10-100 terabytes, and typically multiple copies of each of them exist. They are implemented as ROOT tree structures [69]. The first analysis layer produces secondary datasets, which are maximally in the size order of 100 gigabytes and suitable for further analysis at local machines.

With the tremendous dataset sizes, heavy computing has become a central element in the CERN analyses. In the following, a closer view of the relevant simulation software and simulation techniques is given. The simulation overview is mainly based on the comprehensive Pythia 6 manual [32]. Some details are also taken from the more brief Pythia 8 [33, 34] and Herwig 7 [70] manuals. Finally, the software and dataset workflows at the CMS are also described.

3.3.1 Simulated Particle Collisions

When involved, QCD tends to complicate high-energy particle collisions. If charged leptons and their anti-particles[4] are forced to collide, the most likely hard processes involve the annihilation between the two and photoproduction. Consequently, the final state of the hard process can consist of lepton or quark pairs. In the former case, the collision event observed at the detector is relatively clean, and in the best case no partons are produced either in the initial or the final state of the hard process. In the latter case, the collision can be much more complex, as even a few energetic partons can multiply producing numerous jets. In pp colliders the collisions are further complicated by the parton content from the initial states of the hard processes. That is, significant surplus interactions between partons are possible before the hard process, possibly leading to jet production. These extra jets can add combinatoric challenges to the data analysis.

Also for simulations, the proper treatment of QCD interactions is more involved in contrast to QED. The simulation process commonly begins with a selected hard process, which must be connected to the initial state of two colliding protons, and to the final state of **semi-stable** [5] hadrons, photons and charged leptons arriving to the detector. The simulations can also keep track of the neutrinos, which are not observed at the detector.

The hard process is produced using a mathematical model of the corresponding Matrix Element (ME), given in Eq. (2.8). The evaluation of the ME is the theoretically most solid part of the simulations, whereas the consequent layers involve approximate methods. This does not guarantee that even the ME accurately describes data, as can occur when the ME is evaluated at Leading Order.

Simulated datasets are naturally grouped according to their MEs. Moreover, the ME evaluation is relatively quick in the CMS simulations, and well-optimized compared to the QCD evolution and hadronization processes. Hence, some selections and kinematic cuts can be conveniently performed at the level of a hard process, saving much computation time.

For connecting the hard process with the pp collision and a recorded collision event, the main tools are Initial and Final State Radiation (ISR and FSR). These are together referred to as Parton Showers (PS). In both of the radiation types, a parton

[4] E.g. electrons and positrons.

[5] Stable in the time frame of arriving from the PV to the detector, travelling close to the speed of light.

emits a gluon. FSR is strictly defined as showers that are pointed forwards in time, while ISR allows backwards evolution. Thus, the main purpose of ISR is to connect the hard process to the initial state of two colliding protons *a posteriori*. FSR on the other hand connects the final state of the hard process to detector signals.

It must be highlighted that in the context of simulations, the uses of the terms ISR and FSR have somewhat shifted from their theoretical origin. FSR refers to TimeShower and ISR to SpaceShower in Pythia 8 [33]. Theoretically the radiation occurs in the initial or final state of the ME, for ISR and FSR correspondingly. In contrast in simulation the ISR/FSR branchings can also occur elsewhere and successively. That is, e.g. a gluon that was emitted by an ISR branching in the initial state of the ME can experience an FSR branching. Through numerous ISR and FSR branchings most of the collision event is produced at the parton-level, starting from the somewhat artificial hard process.

On top of ISR and FSR, the simulation software typically has Multi-Parton Interactions (MPI) included. These take care of possible interactions of the proton remnants in a pp collision. The term MPI is used by the simulation software authors, whereas the experimental side often uses the term Underlying Event (UE), which mostly consists of MPI.

Moreover, when comparisons are made between data and simulation, the simulation of pileup interactions is also necessary. Both UE and pileup contaminate the observation of the hard process with additional particle content. At the CMS, pileup simulations are refined in an iterative manner: in the first iterations, a flat distribution of the number of pileup vertices is produced. In the consequent versions, pileup distributions closer to those observed at the CMS detector are manufactured. To make the match close to perfect, the simulated pileup profiles are finally re-weighted to match data.

In the current simulation software, the hard process MEs are variably available at NNLO, NLO and LO accuracy. As a general rule the highest-order accuracy available is preferred. The most favored NLO hard process generators are powheg [71, 72] and Madgraph5_aMCatNLO [73]. The Parton Showers and MPI steps are implemented only in a few software packages. These have internal methods for producing the hard process, but can also be connected with an externally produced hard process.

The Parton Shower generators currently found in use are Pythia [33, 34], Herwig [70] and Sherpa [74]. Due to a historical burden and the wide options available in Pythia, most studies are performed exclusively with Pythia, and only some additional checks are possibly made with Herwig, or rarely with Sherpa. This is not ideal, and a wider use of the different Parton Shower generators would be beneficial. The software models have differences e.g. in their phenomenological hadronization models. Moreover, bugs are occasionally found in the software implementations.

To summarize, much uncertainty is related to the simulation process. Understanding all potential systematical uncertainties would require a wide and regulated usage of different software tools. This is limited by time and workforce: it takes

much effort to produce results even with a single simulation software tool. Duplicating all the steps with another software package can easily duplicate the required time resources.

3.3.2 Software and Datasets at the CMS

The pp collision simulation software can be used as such for standalone studies. However, a complete analysis requires a full simulation, including detector simulation and the simulation of pileup. The detector simulation parameters are produced at the CMS with careful calibrations that utilize e.g. lasers and pion beams [75, 76]. Moreover, the degradation of the detector components over time is also parametrized. In the big picture, the detector simulation is performed with the GEANT 4 software package [77].

The end products of a full simulation chain are distributions of physical observables. When these distributions agree reasonably between data and simulation, it is meaningful to assume that the **simulation truth** also describes data. The simulation truth includes both the fractions and types of the underlying hard processes, and the true energies and momenta of the particles arriving to the detector. This information can be used in various calibrations, e.g. for jets.

Even with a sufficiently calibrated detector simulation, discrepancies can exist between data and simulation. This can be alleviated by tuning the PS generators. The tuning is typically focused on the UE component of the interaction, meaning mainly MPI and optionally ISR settings. Additionally, the tuning can affect e.g. the ISR and FSR calibration. The UE tunes are calibrated by focusing on certain archetypical events. As the UE tuning involves a multitude of parameters, the tuning process always carries the danger of over-fitting. In the course of this thesis, the CMS collaboration has been utilizing the CP5 (CMS Pythia 5) Pythia 8(.205) tune [78] and the CH3 (CMS Herwig 3) Herwig 7 tune [79].

Both Pythia 8 and Herwig 7 are mainly used for PSs, but also occasionally for ME generation. The production of the MEs varies process by process. For the $t\bar{t}$ and single top processes, the ME is generated at NLO using POWHEG V2 [72]. On the other hand, the W+jets and DY+jets MEs are produced at LO using MadGraph5_aMCatNLO 2.2.2 [80]. Finally, the MEs for Multijet and diboson production are generated at LO using the Pythia 8 built-in ME generator. In the CP5 tune, NNPDF31_NNLO [81] is used to extrapolate the Parton Distribution Function (PDF) parametrization with the value $\alpha_s(M_Z) = 0.118$ of the strong coupling constant.

To streamline the use of the numerous involved software tools, the CMS collaboration has collected a number of relevant software tools under a single release: the CMS SoftWare framework (CMSSW) [66]. Working with the CMSSW allows building a quick interface to the already installed software packages. Moreover, CMSSW is compatible with the analysis of large datasets stored around the world, which have either been centrally simulated or recorded by the CMS detector. The CMSSW runs

are split into batch jobs that are processed on the Grid. Single job threads are run locally at the server sites where the datasets are stored.

Essentially no analyses at the CMS are run on top of raw collision data. A general data format aimed for the analysers are the Analysis Object Data (AOD) files, in which the raw signals have been interpreted as particles. With time passing, the AOD files became very large, motivating a compressed MINIAOD format. Both AOD and MINIAOD datasets need to be interfaced with CMSSW – AOD requiring more configuration steps. Working with MINIAOD is complicated due to the steep learning curve of CMSSW. This is one of the motivations for the ongoing adaptation of NANOAOD, where the need for using CMSSW is discarded.

NANOAOD is a beneficial format for analyses that do not require heavy computation on the Grid. The intended use of NANOAOD is in copying subsets of the stored event tree structures onto local machines. In AOD and MINIAOD analyses, a CMSSW analysis is run over the Grid, producing so-called **ntuples** – analogous with the NANOAOD trees. The two weaknesses of NANOAD include Grid computation and the limited and compact dataset sizes. When heavy computation on the Grid is necessary, NANOAOD does not offer benefits over MINIAOD. Moreover, due to the target of compactness there is less information included on the NANOAOD level, limiting many analyses. In this work only MINIAOD datasets are utilized, motivated by the computational requirements and also the freedom provided by MINIAOD. The samples used in this work are listed in Appendix A.

3.4 Combination of Simulated Samples

This thesis deals with the analysis of two different particle physics processes: Multijet production, and top quark pair production. In the former case, relatively loose event cuts are placed and a majority of the events originate from simple QCD processes that produce quarks and gluons. In this case, comparison between data and simulation is effortless. In contrast when $\mathrm{t\bar{t}}$ production is considered, strong phase-space cuts are imposed. Despite this, multiple non-$\mathrm{t\bar{t}}$ background processes need to be considered when comparing simulation with data.

The m_t measurement relies on the comparison between simulation and data, and hence it is essential to scale all the involved signal and background (simulation) processes in the right proportions. In this section, the correct methodology for comparison between data and simulation is reviewed with this motivation. First, the concepts **luminosity** and **cross-section** are introduced for the estimation of the rates of different processes. Then, it is shown how simulated samples from different productions are combined and matched with recorded data. Finally, the most important top quark production modes are introduced, and the choice of the **semileptonic** decay channel in this work is motivated.

3.4.1 Luminosity in Data and Simulation

At a given collision energy $\sqrt{s}$, each pp collision at the LHC follows the same probability distribution. It is possible to calculate theoretically motivated estimates e.g. for the observation probability of a single energetic photon or two energetic muons. With these calculations, the estimation of event statistics becomes highly efficient. For a single process **Proc**, the calculations yield a number referred to as a cross-section (σ_{Proc}). Using this number, the number of pp collision events ending up in a final state **Proc**, is [82]

$$N_{\text{Proc}} = \mathcal{L}_{\text{Proc}} \times \sigma_{\text{Proc}}. \tag{3.2}$$

Here, $\mathcal{L}$ stands for the integrated luminosity. It is a standard measure of the amount of data measured in a given time interval.

Luminosity is measured by luminometers in the inverse units of cross-sections, i.e. inverse barns (b^{-1}). For all unprescaled processes in data the luminosity is the same, so that $\mathcal{L}_{\text{Proc}} = \mathcal{L}_{\text{Data}}$. During the whole Run 2, the CMS recorded 146.9 fb^{-1} of pp collisions. The focus of this work is in the data collected during 2017-2018, making up 108.65 fb^{-1} [10].

Strictly speaking Eq. (3.2) holds as such only for data and simulations with unitary weights. Especially in simulations with an NLO ME, the event weights are often non-unitary and occasionally negative. Then, N_{Proc} and $\mathcal{L}_{\text{Proc}}$ in Eq. (3.2) must be replaced by the corresponding effective values, $N^{\text{Eff.}}_{\text{Proc}}$ and $\mathcal{L}^{\text{Eff.}}_{\text{Proc}}$. In simulation, the luminosity can only be found by inverting Eq. (3.2), leading to

$$\mathcal{L}^{\text{Eff.}}_{\text{Proc}} = \frac{N^{\text{Eff.}}_{\text{Proc}}}{\sigma_{\text{Proc}}}. \tag{3.3}$$

The value of $N_{\text{Eff.}}$ can be computed using the event weights w_{evt}. Considering Poisson statistics for the sum of weights, the following formula for $N_{\text{Eff.}}$ is found:

$$\frac{\sigma_{\text{Poisson}}}{\mu_{\text{Poisson}}} = \frac{\sqrt{\sum_{\text{evt}} w^2_{\text{evt}}}}{\sum_{\text{evt}} w_{\text{evt}}} = \frac{1}{\sqrt{N^{\text{Eff.}}_{\text{Proc}}}}, \tag{3.4}$$

using which Eq. (3.3) becomes

$$\mathcal{L}^{\text{Eff.}}_{\text{Proc}} = \frac{\left(\sum_{\text{evt}} w_{\text{evt}}\right)^2}{\sigma_{\text{Proc}} \sum_{\text{evt}} w^2_{\text{evt}}} = \frac{N_{\text{Tot.}}}{\sigma_{\text{Proc}}} \times \frac{\langle w_{\text{evt}} \rangle^2_{\text{evt}}}{\left\langle w^2_{\text{evt}} \right\rangle_{\text{evt}}}. \tag{3.5}$$

It can be quickly confirmed that for constant weights Eq. (3.5) agrees with Eq. (3.3) with $N_{\text{Tot.}} = N^{\text{Eff.}}_{\text{Proc}}$.

To promote robustness in the MINIAOD analyses, $\sum_{\text{evt}} w_{\text{evt}}$, $\sum_{\text{evt}} w^2_{\text{evt}}$ and the total event count $N_{\text{Tot.}}$ are collected into histograms. These are (and must be) filled before any cuts or filters are applied to the simulated datasets. In the process care needs to be taken: each process corresponding to a given value of σ must be gathered exactly into a single summary histogram, from where the values can be extracted.

3.4.2 Cross-sections

At a given collision energy, the cross-sections are constant – and they generally tend to grow as $\sqrt{s}$ gets higher. The probabilities of different processes grow with $\sqrt{s}$ at different rates, making this view more complicated. Nevertheless, this is the short motivation for building new particle colliders capable of higher collision energies: rare processes may become statistically visible.

In general, the cross-sections are calculated from the corresponding MEs. In Eq. (2.9) it was found that the event probability is related to $|\mathcal{M}|^2$. Reviewing Eq. (3.2), it can be concluded that the observation probability is also proportional to σ. Thus the following must hold:

$$\sigma \propto |\mathcal{M}|^2. \tag{3.6}$$

It is not possible to make an event selection that only selects events from the **signal** process. For instance, in top quark pair measurements, single top quark production often becomes an irreducible **background**. When many types of process are present, Eq. (3.6) expands to:

$$\sigma_{\text{tot}} \propto \left| \sum_i \mathcal{M}_i \right|^2 \approx \sum_i |\mathcal{M}_i|^2 \tag{3.7}$$

The approximation presented in Eq. (3.7) is non-trivial: there is no general guarantee that the cross terms of different processes would average to zero. However, this approximation usually holds reasonably well and allows a very convenient trick. Using the results of Eq. (3.7) and Eq. (3.6), the total cross-section consisting of a signal and background processes can be given:

$$\sigma_{\text{tot}} \approx \sigma_{\text{signal}} + \sum_{\text{bkg}} \sigma_{\text{bkg}} \tag{3.8}$$

This result can be interpreted so that with a certain event selection and a certain integrated luminosity $\mathcal{L}$, the total set of collision events is found simply by summing up the different event types. That is, quantum interference (i.e. the cross terms) does not contribute significantly to the end result.

As a final remark on the assumptions on interference, this is in general not a topic to be taken lightly. In the single top simulation community there is currently high interest in the interference between $\mathrm{t\bar{t}}$ processes and some single top production processes. Handling this is a smaller worry for $\mathrm{t\bar{t}}$ analyses than the single top analyses, as $\mathrm{t\bar{t}}$ production has higher cross-sections. At the time of writing of this thesis, the solutions were still a work in progress.

At the level of Eq. (3.8) the remaining issue is that background processes can have much larger cross-sections than the signal process. This is the case e.g. for the Multijet background, when compared to a $\mathrm{t\bar{t}}$ signal. To treat this issue, event

selections can be placed e.g. on muon p_T values. The selections constrain the **phase space**, and make the observed/effective cross-sections smaller. The event selection is optimized to maximize the signal-to-background ratio.

3.4.3 Weighting of the Simulated Samples

In simulation studies, there is typically a variety of samples available. These should be matched with an experimental dataset carrying the integrated luminosity $\mathcal{L}_{\text{Data}}$. A generic weight formula for a single event in a simulated sample can be expressed as

$$w_{\text{evt}}^{\text{tot}} = \frac{w_{\text{evt}}}{A} \frac{\mathcal{L}_{\text{Data}}}{\mathcal{L}_{\text{Proc}}^{\text{Eff.}}} \tag{3.9}$$

The normalization constant A considers the possibility of non-unitary w_{evt} values, so that

$$\sum_{\text{evt}} \frac{w_{\text{evt}}}{A} = N_{\text{Proc}}^{\text{Eff.}}. \tag{3.10}$$

In LO samples the typical weight choice has been $w_{\text{evt}} = 1.$, leading to $A = 1$. In NLO samples $w_{\text{evt}} \in \{\pm 1\}$ or $w_{\text{evt}} \in \{\pm C\}$ (C = constant) are frequent options. It should be noted that while w_{evt} values can be negative, the sum of weights and A should remain positive. The fewer negative weights, the better: Eq. (3.5) shows how the effective luminosity in simulation degrades when the fraction of negative weights increases.

Plugging Eq. (3.10) and Eq. (3.3) into Eq. (3.9) one finds the master formula

$$w_{\text{evt}}^{\text{tot}} = \frac{w_{\text{evt}} \cancel{N_{\text{Proc}}^{\text{Eff.}}}}{\sum_{\text{evt}} w_{\text{evt}}} \frac{\sigma \mathcal{L}_{\text{Data}}}{\cancel{N_{\text{Proc}}^{\text{Eff.}}}} = \frac{\sigma \mathcal{L}^{\text{Data}} w_{\text{evt}}}{\sum_{\text{evt}} w_{\text{evt}}}. \tag{3.11}$$

Hence, $N_{\text{Eff.}}$ cancels out in the total event weights and is only required for estimating the effective luminosity in Eq. (3.3).

3.5 Top Quark Production

For top quark studies, $\text{t}\bar{\text{t}}$ production processes have been historically the most prominent ones. This is easily explained referring to statistics. At $\sqrt{s} = 13$ TeV, top quark pair production has a cross-section around 832 pb, whereas the production of a single top quark has a combined cross-section around 300 pb [83, 84]. For the years 2017–2018, this corresponds to 90 million recorded top quark pair events and 33 million recorded single top events. Thus, the amount of recorded single top events is only a third of the recorded top quark pair production events. A majority of the top quark pairs observed at the LHC are produced via gluon fusion. In pp interactions, two energetic gluons can be fused into one, which further decays into a top quark pair.

3.5.1 Top Quark Pair Decays

As the top quark decays dominantly through the weak interaction, its decay products are a quark and a W boson. As a further curiosity, the associated quark is a b quark with a 99.9% probability [62]. The W boson decays into a pair of quarks with an approximate probability of 2/3, and else into a charged lepton and a neutrino [62]. This complicates the measurement of m_t using the direct decay products of the top quark. These inevitably include at least one quark, and possibly also neutrinos. Only electrons, muons and photons can be measured with a high precision by the particle detectors.

In summary, nearly 100% of the time a top quark pair decays producing two b quarks and two W bosons. The top quark pair events can be further classified into three categories according to the leptonic or hadronic decays of the W bosons. These are the fully hadronic events (no charged leptons), the semileptonic events (1 charged lepton) and the dileptonic events (2 charged leptons). Statistically speaking, the fully hadronic and semileptonic channels take approximately 4/9 each of the full top quark pair cross-section, and the dileptonic channel is left only with 1/9 [62]. Thus the semileptonic and fully hadronic channels are statistically most interesting, with a $\sigma \approx 370$ pb each.

3.5.2 The Semileptonic Decay Channel

In the fully hadronic decay channel, the common event topology includes two b quarks and two pairs of W-decay quarks. Furthermore, the high-energy collisions tend to include energetic gluons, which produce additional jets. The modern particle detectors are able to distinguish jets originating from a b quark relatively well from other jet types. However, making a distinction between lighter quark jets and gluon jets is more complicated. This leads to high combinatorics in the fully hadronic decay channel – reducing the signal fraction in reconstructed events.

The semileptonic or **lepton+jets** collision events are combinatorically easier to handle. First of all, the single energetic charged lepton originating from a W boson functions as a strong indicator for these events in the CMS data. The collision event consists of two b quark jets and only two light quark jets originating from a W boson. The reconstructed W boson mass can be utilized as an indicator for the success of a top quark pair event hypothesis, e.g. employing **kinematic fitting**. A kinematic fit is utilized in the CMS analysis, and will be described in chapter 6. The greatest weakness of the lepton+jets channel is the neutrino, as its momentum can only be estimated. The estimation is performed employing the leptonic W boson resonance, the missing transverse momentum and kinematic fitting methods.

Following from similar reasons as were presented above, the lepton+jets channel of top quark pair production has historically been the most prominent one in m_t measurements. This channel is found in two places within this thesis: first in the studies of the DØ measurements and then in our own CMS studies. In Fig. 3.7

the most likely topologies for $t\bar{t}$ production in the lepton+jets decay channel are presented. The processes found by interchanging the hadronic and leptonic decay branches are equally likely as the ones presented here.

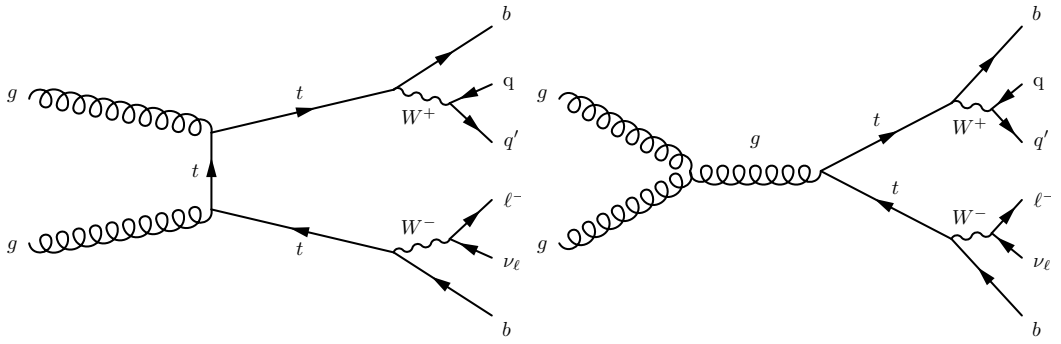

Figure 3.7: Example Feynman diagrams for a semileptonic $t\bar{t}$ decay.

4 Jet Calibration and Reconstruction at the CMS

Jets are produced in high-energy particle collisions, and they collectively form momentum-imprints of the original partons. If the laws of QCD were even slightly different, the end product of a HEP collision could as well be an extremely complex porridge of hadrons and leptons. Consequently, numerous HEP measurements rely on the ability to measure jets. This is a non-trivial task, as QCD presents exceedingly complex dynamics into proton-proton collisions.

The composition of a jet varies stochastically, but in general jets consist of hadrons, photons and charged leptons. In theoretical and simulation studies, also neutrinos can be considered as a part of the jets, to capture the full momentum of the original parton. The more energetic the original parton was, the more collimated the resulting jet is.

Ideally, jets would give as accurate a picture of the original parton produced in a hard process as is the case for energetic electrons and muons observed at the detector. Since jets are composite objects, this is in practice infeasible. The particle contents of jets vary notably, and different particles have different detector responses, In consequence, jets require numerous calibrations and corrections, and even after this there is a level of uncertainty left.

This chapter introduces some important features of jets, and also some jet-related contributions by the author are presented. These contributions have been generally performed for the full Run 2 (2016–2018). However, as the main m_t measurement in Ch. 6 only focuses on the 2017–2018 datasets, the same demarcation is performed for the presentation of jet studies.

4.1 Jet Definitions

Below, the available jet definitions are characterized, and the current jet definition choices at the CMS are explained.

4.1.1 Jet Algorithms

There is no single theoretical or algorithmic definition for jets. Hence, a qualitative experimentally motivated definition is the best starting point. Etymologically, the term **jet** refers to the French word meaning literally **a throw**.

In Particle Physics, the notion of a *jet* can also be understood as a possible reference to the fountain **Jet d'Eau** in Geneve, close to CERN. This is a massive fountain that sprays water up to 140 meters, as depicted in Fig. 4.1. The jets observed at particle colliders have an analogous nature to this water spray - with the distinction that a particle jet only lasts for a few fractions of a second. In a particle jet, the water droplets of a water spray are replaced by various kinds particles. The analogy can be appreciated by viewing the detector image of jets in Fig. 4.2.

A human is typically able to identify the single jets by hand from a detector image in the $\eta - \phi$ plane. However, manual jet clustering is infeasible, as jets in millions of collision events need to be processed. Thus, an algorithmic approach must suffice.

Defining a satisfactory jet algorithm is complicated, as it is difficult to generalize all the possible exceptions in a format that computers understand. Therefore, it is necessary to define the jet features that are prioritized in the algorithm. With this approach, robustness is one of the first and foremost features. Robustness means that small changes in jet features (e.g. particle numbers and energies) should not dramatically change the interpretation of the jets in the event. The most important jet algorithm robustness categories are **infrared** and **collinear safety**. Infrared safety refers to robustness under small variations in particle energies. Furthermore, collinear safety refers to robustness under the emission of extra gluons approximately in the jet axis direction. Secondary requirements for jet algorithms typically include e.g. jet shape regularity, referring to somewhat round shapes in the $\eta - \phi$ plane.

Most jet clustering algorithms operate in the $\eta - \phi$ plane. This is motivated by the fact that the $\eta - \phi$ plane particle density is somewhat constant in HEP collisions. Typical algorithms process a list of **pseudo-jets**, and at each step of the algorithm decide which two pseudo-jets should be combined into a single one. At the first step of the algorithm, the list of pseudo-jets equals all of the observed particles. Consequently, after the **ending condition** is reached, the remaining pseudo-jets are interpreted as the final jets. The currently most successful jet algorithms compare pseudo-jets with an $\eta - \phi$ distance metric defined as

$$d_{ij} = \min\left(p_{T,i}^{2p}, p_{T,j}^{2p}\right)\frac{\Delta R_{ij}}{R}, \tag{4.1}$$

$$d_{iB} = p_{T,i}^{2p}. \tag{4.2}$$

Here, i and j stand for the indices of pseudo-jets. The $\eta - \phi$ plane distance between pseudo-jets i and j is defined as

$$\Delta R_{ij} = \sqrt{\left(\eta_i - \eta_j\right)^2 + \left(\phi_i - \phi_j\right)^2} \tag{4.3}$$

Figure 4.1: The fountain Jet d'Eau in Geneve, in the current location since 1891 [85].

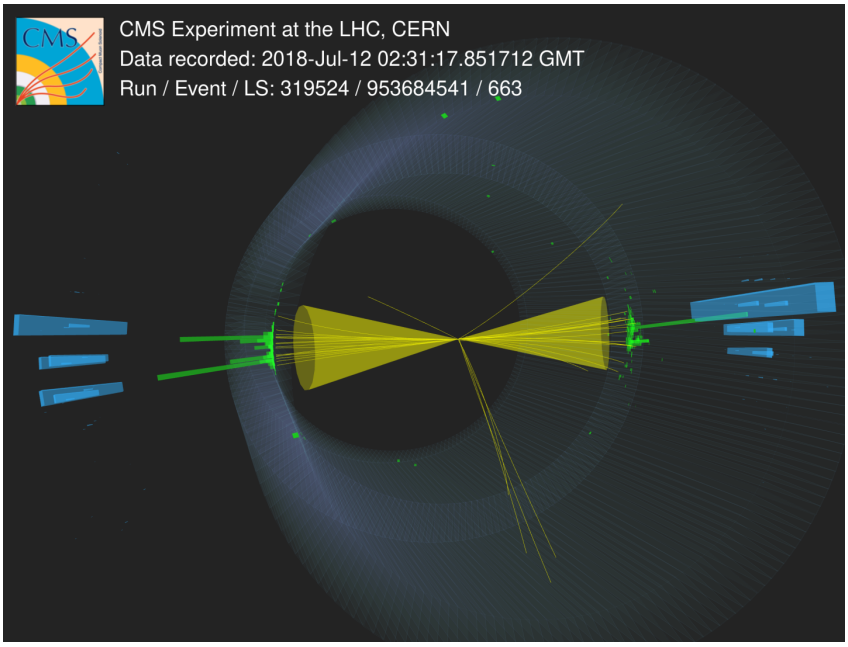

Figure 4.2: An image depicting jet production at the CMS [86]. Yellow lines visualize tracks of charged particles. The towers on the rim of the circle indicate calorimeter energy deposits. The cones signify the shapes of the jets.

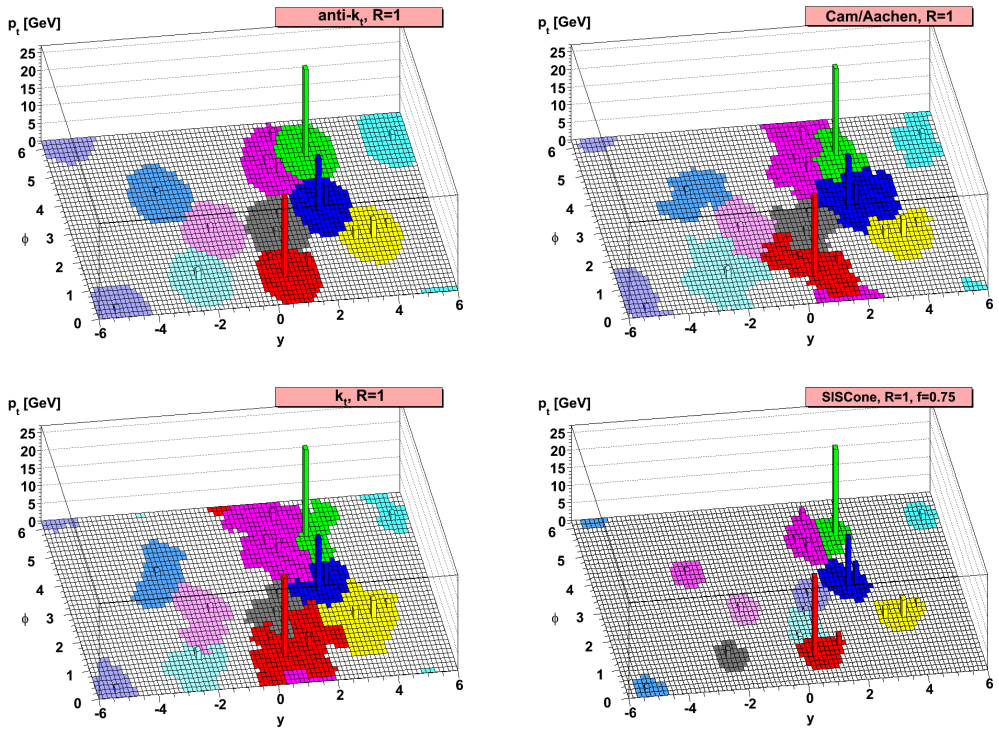

Figure 4.3: Comparison between Jet Algorithms [87]. Each colored shape is a jet in the y-ϕ plane. The z-axis denotes the p_T within one histogram cell, indicating that most of the jet energies here are concentrated within a single cell.

Moreover, the parameter R is the radius parameter of the jets. The integer value of p defines the name and behavior of the jet algorithm. The value $p = 0$ corresponds to the Cambridge-Aachen algorithm, $p = 1$ to the k_T algorithm, and $p = -1$ to the anti-k_T algorithm. The features of these and the SISCone algorithm can be compared using the famous comparison of Fig. 4.3. Currently, the anti-k_T algorithm is the most favored one – partially since it often finds continuous and round shapes for the jets. This regularizes the behavior of jet areas, which is important while correcting the energies of jets. As depicted in Fig. 4.3, even the anti-k_T algorithm is not flawless: one of the jets has the shape resembling a half-moon. Here, rapidity (y) is used instead of the common choice η, but this does not change the conclusions.

The algorithm given by Eqs. (4.1,4.2) reconstructs the measures d_{ij} and d_{iB} at each step of the algorithm. Then, the minimal value of these will be considered. If the minimal d_{ij}/d_{iB} value is of the d_{ij} type, the two pseudo-jets i and j will be merged. If the value is of the d_{iB} type, the pseudo-jet i will no longer be considered for merging, and it will be found on the list of final jets. The ending condition is

reached when the list of pseudo-jets considered for merging is empty. The d_{iB} condition and the R parameter together make sure that the radius of a jet corresponds maximally roughly to the parameter R.

4.1.2 Jets in the CMS Experiment

In the CMS experiment, the anti-k_T algorithm with the radius $R = 0.5$ was favored during Run 1. In Run 2, the main jet radius was reduced to $R = 0.4$. Moreover, the value $R = 0.8$ is used as a secondary **wide-jet** radius. In the CMS terminology, the jets clustered using PF candidates are shortened to AK4PF, AK5PF and AK8PF jets. This expression carries the information of the origin (PF candidates), the algorithm and jet radius (AK4 = anti-k_T, $R = 0.4$).

During Run 2, CMS has also introduced jet grooming through Charged-Hadron Substraction (CHS) [88]. This method removes charged pileup particles from jets. Removing neutral pileup particles would be more complicated, as determining their vertex of origin is difficult. The resulting CHS jets, which are mainly considered in this thesis, are titled AK4PFchs jets. In Run 3 CMS plans to use the PUPPI algorithm [89] for its primary jet collections. PUPPI employs advanced weighting techniques for the reduction of pileup effects. This will help in dealing with the increased pileup conditions in Run 3.

As the LHC is a proton-proton collider, all collision events contain jets. Beside jets, a collision event can contain isolated electrons, muons or photons: these are all likely to be originated in the hard process. Tau leptons are so short-lived that only their decay products are observed. High-energy neutrinos remain unobserved, but their momenta can be estimated through the missing transverse momentum.

An event that does not contain any isolated charged leptons or photons is generally characterized as a **Multijet** event. However, most of the events produced at the LHC are generic QCD-induced Multijet events. Hence, these can be studied even without imposing exclusion conditions e.g. on charged leptons or photons.

The number of Multijet events falls steeply as a function of the highest jet p_T in the event. If all Multijet events were recorded, this would increase the amount of events processed and stored at the CMS by several orders of magnitude. Most of these events would be those containing only low-p_T jets, and provide relatively little physical interest value. This kind of resource allocation is infeasible, as the CMS data acquisition system (and servers) would not be able to handle this amount of data throughput. Consequently, the HLT jet triggers at the CMS are used to keep the rates of stored jets at different p_T values acceptable. Most of the triggers specialized on Multijet events are thus **prescaled**. This allows storing only a small fraction of the events with lowest jet p_T values.

A listing of the relevant single-jet triggers at the CMS 2017–2018 runs is given in Table 4.1. These are the triggers used for the jet studies of this work. The PF jet triggers operate on the jet p_T values constructed at the HLT level. The HLT reconstruction is performed quickly, and hence the final calibrated jet p_T values

differ somewhat from the HLT ones. At the HLT, timing is crucial. In statistical terms, the relationship between HLT and the final reconstruction is a convolution. Most analyses are required to inspect AK4PFchs p_T values where the trigger is fully active. The active regime is indicated by the trigger **turn-on** points, given in Table 4.1.

Table 4.1: HLT single-jet triggers used at the CMS during Run 2.

Trigger name	HLT p_T min (GeV)	AK4PFchs turn-on p_T (GeV)
HLT_ZeroBias	0	0
HLT_PFJet40	40	49
HLT_PFJet60	60	84
HLT_PFJet80	80	114
HLT_PFJet140	140	196
HLT_PFJet200	200	272
HLT_PFJet260	260	330
HLT_PFJet320	320	395
HLT_PFJet400	400	468
HLT_PFJet450	450	548
HLT_PFJet500	500	686

When any trigger **fires**, the collision event is stored. If a jet with its HLT p_T exceeding the given minimum exists in the event the trigger in question *fires* – excluding the effect of a possible prescale. Hence, if a jet with a very high HLT p_T exists in a collision event, all the triggers up to the given p_T value would *fire* without the prescales. Considering the effect of the prescales, it is unlikely that the lower-p_T triggers are active, and usually only one or two of the PFJet triggers fire at the time. As a universal trend the prescale values grow exponentially as the HLT p_T cutoff decreases.

For the lowest p_T range, the ZeroBias trigger is used. It saves all events, regardless of the p_T values, but it is kept active only at short intervals during data collection. In contrast, the HLT_PFJet triggers look at all the events, saving constantly a fraction of the events according to their given prescales and p_T minima.

4.1.3 Distinction Between Jets and Isolated Particles

As a convenient generalization, electrons, muons, photons, tau leptons and jets are the categories of **physics objects** into which the detector signals are grouped. Some detector signals can be interpreted as multiple physics objects, as is the case between electrons and photons. Each analysis needs to consider how to treat such possible ambiguities and overlaps.

A jet is a collection of particles limited in time (a single collision) and space ($\eta - \phi$ plane). Besides these characteristics, there are no other *a priori* limitations

for jets. This means that most of the other physics objects can also be interpreted as jets: e.g. an electron or a photon producing an electromagnetic cascade in the ECAL is a jet. At the CMS, other physics objects are often found to be included in the inclusive jet collections. The final decision of excluding objects from the jet collection is left for the analyst.

A jet enthusiast could even say that all the physics object categories found in reconstruction are actually jet categories: electron(-jet)s, photon(-jet)s, ... and (miscellaneous) jets. The last category contains what is conventionally understood to be **jets** - and in this category the energy reconstruction has the greatest uncertainties. Thus, there is a dual motivation for the physics object categories: both the physical interpretation and energy reconstruction vary greatly by category.

For distinguishing charged leptons and photons from jets, CMS offers carefully calibrated isolation and identification (ID) criteria. Moreover, ID criteria are also calibrated for jets, so that jet-centered analyses can exclude atypical jets. While treating complicated events and applying ID filters for all physics objects, it is possible that some objects will be excluded from all categories.

These misidentified jets can still play their part in the full event content, and should not be forgotten. As an example, this can be the case when the analysis focuses on the jets with highest p_T values in an event. If a misidentified jet has a higher p_T value than the accepted jets and this misidentified jet is discarded from the analysis, it becomes more likely that the collision event is misinterpreted. An exception is made when such a jet originates from pileup, and does not contribute to the interpretation of the collision at the PV.

4.2 Jet Classification and Features

Below, some in-depth information on jet features is given. The theoretical explanation relies generally on Ref. [87] and the parts on jets at the CMS on Ref. [91]. Furthermore, some of the explanations are based on the author's long experience with jet-related results from simulations and data.

4.2.1 Jet Categorization

Jets originate mostly from quarks or gluons, and thus they are dominated by hadronic content. This rule is violated only by hadronically decaying tau leptons, which produce very collimated jets. In many analyses there is a need to determine the parton type from which a jet has originated.

The most important distinction is that of b quark jets versus all other parton jet types, as b jets are important indicators for many interesting collision event types. Luckily, b jets are characterized by a distinct scale factor – following from the fact that the b quark decays soon after leaving the PV. The greatest trouble for b jet tagging (b tagging) is confusion with c quark jets and occasionally with hadronic tau jets. Both of these also produce a SV. [92]

For b/c-quark jets, the SV denotes the point where the b/c-quark hadron decays into hadrons, photons and leptons containing only lighter quark flavours. When a quark changes flavour, a W boson is also produced. Roughly one third of the time, the W boson decays into a charged lepton and a neutrino [62]. These b-decays are referred to as **semileptonic**. For an average b/c quark jet it follows that a large portion of the jet energy fractions is carried by charged leptons, accompanied by unobservable neutrinos. The undetected presence of neutrinos distorts the relationship between the jet and the quark energies. B quarks decay mainly producing c quarks and only subsequently lighter quark flavors. This indicates that an average b jet includes a greater fraction of leptons than the average c-jet, as b quarks have two opportunities for semileptonic decays.

Finally, quark-gluon tagging is an important and evolving tool in jet categorizations [93, 94]. Compared to a high-energy quark, a high-energy gluon produces more secondary gluons, and hence the number of particles in gluon jets are taken to be higher in average than in quark jets. Moreover, the gluon jets are typically wider than quark jets. There is a statistical distinction between quark and gluon jets, but for a single jet making the distinction is more difficult. If this could be done with as good a precision as b-tagging, many analyses would benefit from the results.

4.2.2 Simulated Jet Energy Composition

Even if isolated charged leptons and photons are conventionally not considered to be jets, jets can contain non-isolated leptons and photons. In average, photons populate fractions ranging from 30% to 40% of the jet energy and particle content. Most of these photons originate from π^0 mesons decaying into a photon pair. The greatest average jet energy fraction – up to 60% – comes from charged hadrons. Moreover, around 10% typically comes from neutral hadrons. Finally, only a few percent of average jet energy is found as electrons and muons. A similar fraction of jet energy is found in the form of neutrinos, which cannot be included into the reconstruction-level jets at the CMS.

With the above characterization, it is apparent that a jet is defined most of all by the diversity of particles. This motivates the study of jet energy fractions. This can be done as a function of different jet variables, but the most common choice is to study jet energy fractions as a function of jet p_T. Fig. 4.4 (left) displays the behavior of the jet energy fractions in a simulation without detector reconstruction and pileup. Such simulated results are generally referred to as **generator-level** results/simulations, - in contrast to **detector-level** results (or measurement), where the detector effects are included.

At generator-level, the fractions of various particle types do not significantly change as a function of jet p_T. The same holds also for other jet variables, such as η. Reasons for the slight p_T dependent variations in Fig. 4.4 (left) include e.g. the varying rates of different hard processes at different p_T values, as depicted in Fig. 4.4 (right). Jets originating from quarks tend to produce somewhat different

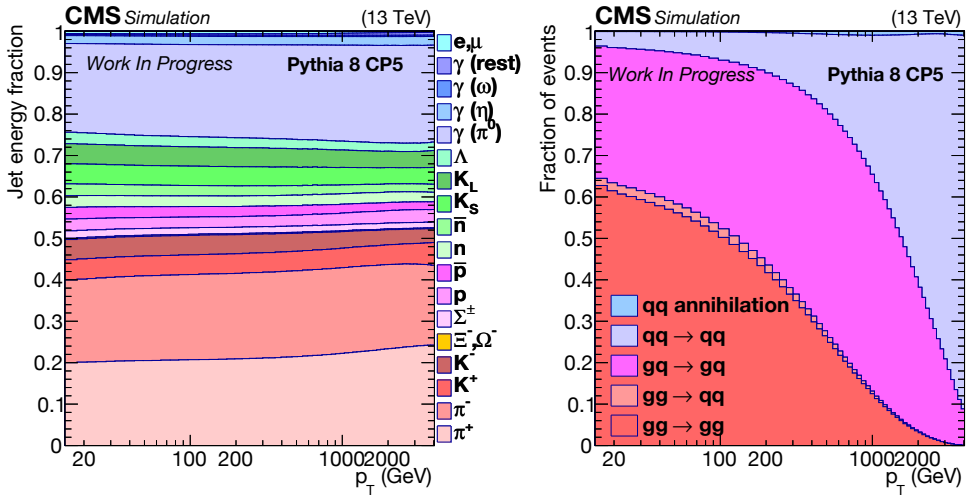

Figure 4.4: Jet energy fractions as a function of jet p_T for Pythia 8 in $|\eta|< 1.3$ (left), and rates of different hard process types at different p_T values (right).

particle content than those originating from gluons.

The weakness of almost all generator-level studies is the inability to compare data with simulation. This is also the case for jet energy fractions: comparison studies can only be made for detector-level measurements. Hence, the *ground-truth* for both data and simulation must be extrapolated from simulation-only generator studies.

4.2.3 Measured Jet Energy Composition

The fact that the jet energy fractions display stable behavior on the generator-level makes them an excellent precision tool and indicator for studying the performance of detector and jet calibration. The features of the detector are directly conveyed to the detector-level jet energy fractions. Also mismatches in calibration are often seen as jet energy fraction disagreement between data and simulation. The level of agreement between reconstruction-level jet energy fractions in simulations and data signifies how well the data is understood.

The cylindrical detector structure has a major impact on the η-dependence of the measured jet energy fractions. In this thesis we focus on the $|\eta|< 1.3$ zone, where the jet calibration is most accurate and detector artefacts are limited. Fig. 4.5 presents the 2017-2018 jet energy fractions. In these figures the fractions of PF candidate categories are presented as a function of the jet p_T. Moreover, the charged pileup[1] is plotted on top of the charged hadron fraction. The figures are produced in Multijet events with the triggers given in Table 4.1. The markers stand for data and the histograms for simulation.

[1] Removed by the CHS algorithm.

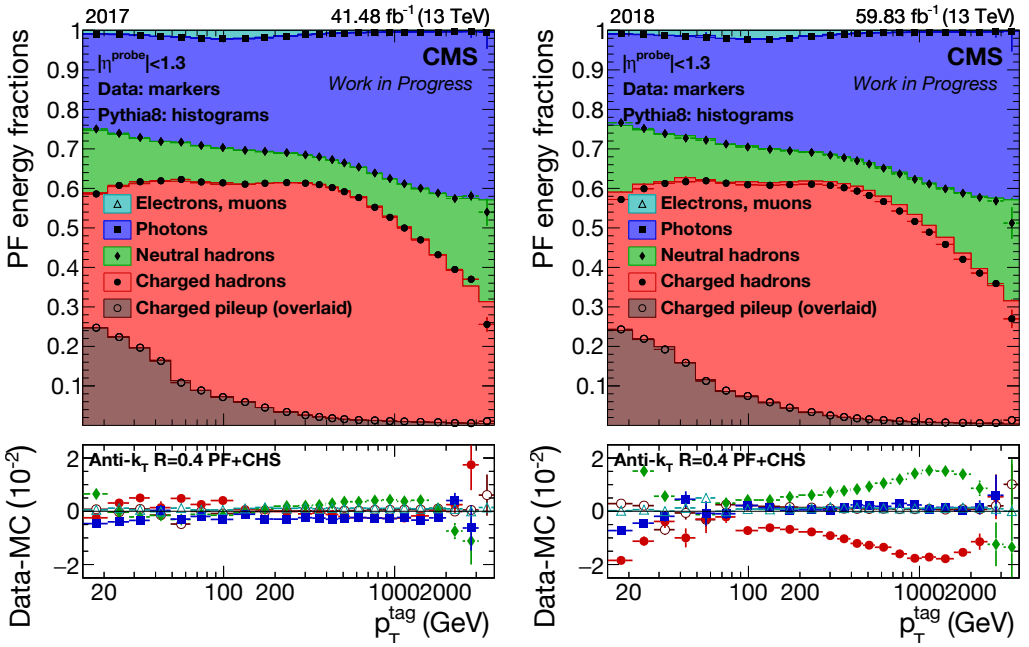

Figure 4.5: 2017 (left) and 2018 (right) PF jet energy fractions at $|\eta| < 1.3$ in Multijet events as a function of p_T.

The jet energy fraction results presented in Fig. 4.5 have been perpetually updated as the calibration process has proceeded. The analysis leading to these figures is openly available under the **jetphys** repository [1]. It is observed that the charged pileup that has been removed by the CHS algorithm obeys a strong p_T dependence. This follows from the fact that the amount of charged pileup is not dependent on the jet p_T, and hence the relative fraction of pileup is larger at lower p_T.

By comparing Fig. 4.5 with Fig. 4.4 differences in the p_T dependence are observed. The fraction of neutral hadrons and photons grows and that of charged hadrons is diminished at high p_T for the PF jet energy fractions. This behavior follows from the internal logic of the PF algorithm. Calorimeter energy not associated with tracks is reconstructed as photons (ECAL) and neutral hadrons (HCAL). As the p_T of a high-energy charged particle grows, it becomes more difficult to determine whether the particle track bends in a magnetic field.

In Figs. 4.6 and 4.7 similar jet energy fractions are presented in semileptonic $\mathrm{t\bar{t}}$ events. In Fig. 4.6 the jets in question are most likely light quark jets from a W boson decay. The light quark and b (quark) jets are distinguished by b-tagging. Analogously, Fig. 4.7 presents results for b jets. In all these figures the p_T-axis is scaled to cover a shorter range than in the Multijet events. The $\mathrm{t\bar{t}}$ production processes are unable to reach as high jet p_T values as the Multijet production.

Through a comparison between Figs. 4.6 and 4.7 the contribution of semileptonic

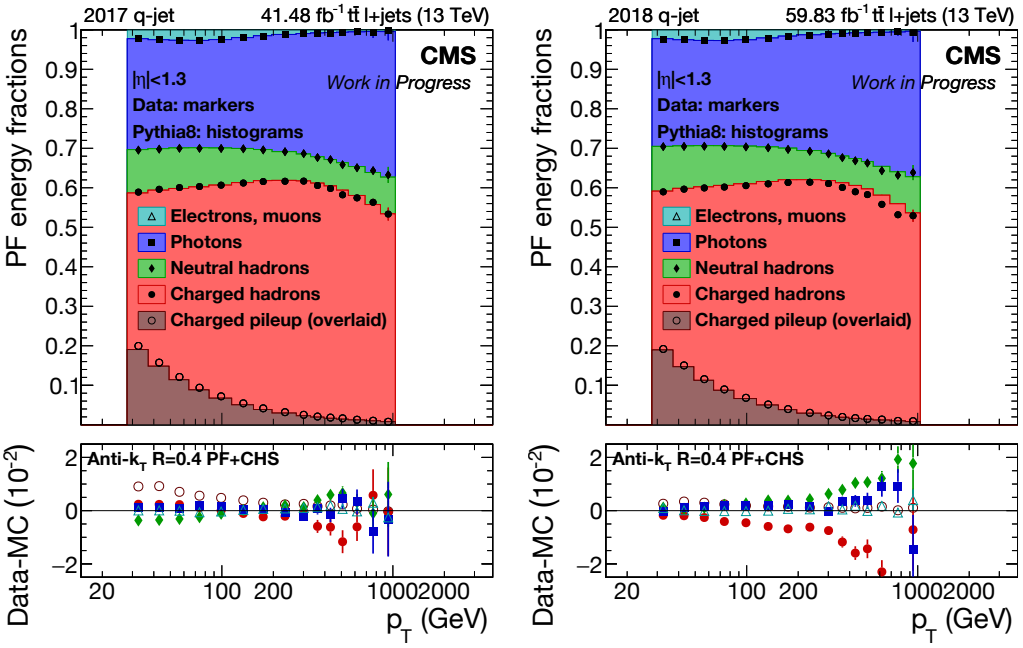

Figure 4.6: 2017–2018 light quark PF jet energy fractions at $|\eta|< 1.3$ in $t\bar{t}$ events as a function of p_T.

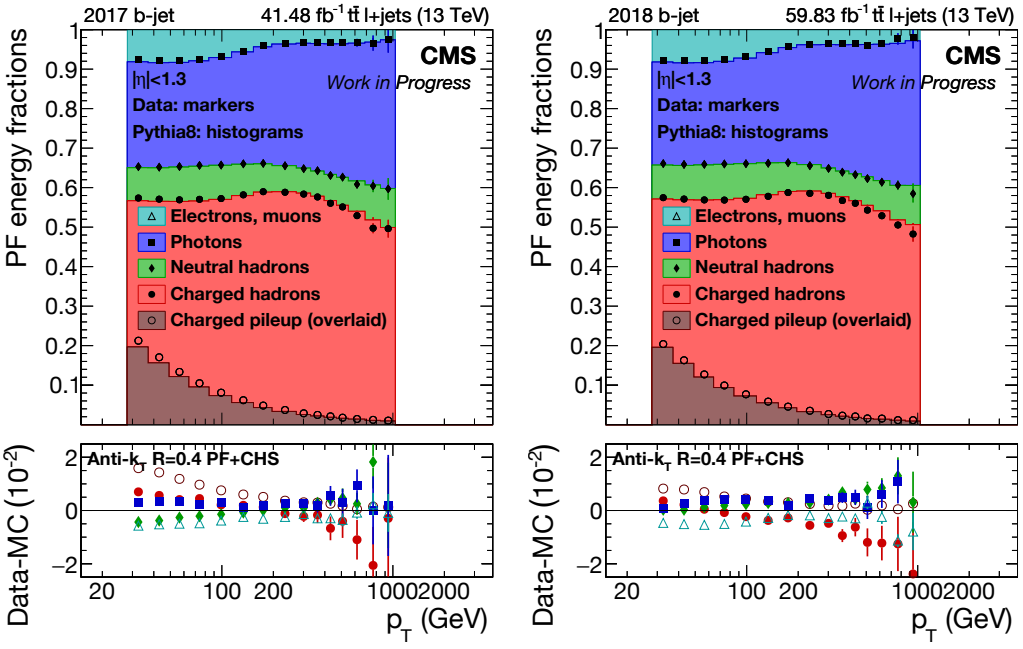

Figure 4.7: 2017–2018 b quark PF jet energy fractions at $|\eta|< 1.3$ in $t\bar{t}$ events as a function of p_T.

b quark decays can be observed. As most of the leptonic content in an **average** b jet originates from the semileptonically decaying third of b jets, the leptonic content of an average semileptonic b jet is far greater. This presents a great challenge for JEC calibration: in the semileptonic b jets the detector response is notably different than in hadronic b jets. Moreover, the semileptonic b jets lose a significant fraction of their energy to neutrinos. The value of $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is too inaccurate for efficiently estimating the momenta of these semileptonic b jet neutrinos, as it is most useful when only a single energetic neutrino is present in an event. For instance, in the semileptonic $\mathrm{t\bar{t}}$ decay channel there is by definition at least one energetic neutrino present, so utilizing $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ for b jet neutrinos is out of the question. The b jet neutrinos are an acute challenge for m_t measurements, where the connection between b jets and b-quarks plays an important role.

With simulation studies, it can be shown that the presence of a muon within a b jet almost always indicates the presence of a neutrino. However, the presence of electrons is not a similarly pure neutrino-indicator, as electrons are found also e.g. in π^0 decays. Counting the number of electrons in the jet would also not help, since it is not guaranteed that both of the π^0 electrons end up in the jet. In future b jet calibrations, the distinction between b jets with and without neutrinos could be utilized again. This has been done before by the DØ collaboration, where b jets with and without muons were treated separately.

4.3 Jet Calibration

The most important task in jet calibration is finding the Jet Energy Corrections (JECs). The objective of the JECs is correcting the measured energy/p_T of an average jet to match the **generator-level truth**. These corrections are meaningful for averages with good statistics: for a single jet the relationship between the truth and corrected p_T typically follows a Gaussian-like resolution. These resolutions can be reduced by further specialization of JECs, e.g. by treating light quark and b quark jets separately.

In the following, the most relevant definitions for JEC calibration at the CMS are given following Refs. [2, 91, 95, 96]. In addition to the main components of the JEC measurement, subsidiary measurements are also important for the calibration process. These include the PF jet energy compositions presented above, which are currently also being experimented as a direct component of the JEC calibration. Additional important subsidiary results are the jet veto zones studied below, to which the author of this thesis has made major contributions.

4.3.1 Jet Energy Corrections at the CMS

The definition of the reference target (generator truth) for JECs is non-trivial. Many analyses would benefit from the target being on the parton-level, but the ability to

perform a transition from jets to partons is limited both by theory and practice. One of the related issues is the definition of a parton level: the simulation tools operate using phenomenological models, which aim to produce physically meaningful results mainly on the hadron level. This indicates that the physical accuracy of the parton level developments could be relatively poor.

There are also other obstacles for a full jet-parton connection, such as the treatment of neutrinos. It would be useful to correct for the missing neutrinos, but in practice this is a difficult task to execute. Thus, before any jet corrections can be made, a robust reference target must be defined.

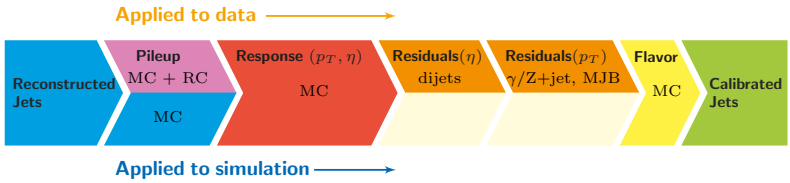

Figure 4.8: JEC workflow at the CMS [91].

At the CMS (and many other collaborations) the reference target for JECs are the **particle jets**. A particle jet is a jet clustered from the particles observed at the detector, but excluding detector simulation and calibration. In other words, a jet defined at the **generator-level**. This is a compromise that provides robustness, but loses the full connection to partons. The two most important components that are lost from a full jet-parton connection are the neutrinos, and the particles ending up Out-Of-Cone (OOC) with the $R = 0.4$ jet radius definition. The OOC component becomes greater as p_T decreases, as higher p_T means a higher level of jet collimation. To give a rough understanding of the relevant energy scales for AK4PFchs jets, at $p_T > 50\,\text{GeV}$ the OOC component is relatively small (a few percents) and at $p_T < 30\,\text{GeV}$ the OOC component becomes significant (more than 5 percents) [97].

The correction procedure relies heavily on the detector simulation and simulated events. On top of the corrections derived from simulation, residual corrections must be applied to close the remaining gap between data and simulation. Fig. 4.8 presents the general JEC workflow applied during Runs 1 and 2 at the CMS. It is slightly outdated on the newest methods applied for the residual JECs, as explained later. The JEC levels are applied layer by layer on the reconstructed (raw) jet energy. The steps are numbered from left to right starting from Level 1 - and the calibrations can be openly accessed in the Repository [98]. Most JEC levels depend only on jet p_T and η, but also adding ϕ dependence has been considered in the newest JEC developments. In the following, the JEC levels are inspected closer.

Offset Corrections

If the PF algorithm would be able to accurately reconstruct all particles and their origins (in space and time), JECs could be replaced by particle-level corrections. The

origin is important in determining which particles originate from the hard processes and which from pileup. Neutral particles are an obstacle for this ideal: their exact origin and number is difficult to determine. This implies that jets will always be contaminated by neutral pileup particles with the current detector design.

To correct for the presence of pileup, it is beneficial to consider the whole ensemble of particles within a jet. From this ensemble the pileup contamination can be reduced in a statistically meaningful way. For the AK4PFchs jets this task is handled by the **Offset corrections**. The PUPPI jets will be used primarily in Run 3, and they will greatly reduce the need for offset corrections, as pileup reduction is built into the algorithm.

In the AK4PFchs jets considered in this thesis, the offset corrections remove mostly neutral pileup contributions, but also some residual amounts of ambiguous charged pileup that is not removed by the CHS algorithm. In short, the background energy density ρ in the $\eta - \phi$ plane is measured in each event. The AK4PFchs jets have a relatively round shape, and their exact area in the $\eta - \phi$ plane can be determined. By multiplying the jet area with the energy density, the pileup contribution in an average jet is found and can be subtracted. Also the UE produces a similar but smaller effect. Corrections for UE contributions have been planned for the final round of Run 2 JECs.

Both simulations and measured detector data are used for tuning the offset corrections. The offset corrections are expressed as a function of p_T with parameters binned in the granularity of HCAL towers on the η-axis.

MC Truth Corrections

The **MC truth corrections** are fully based on simulations. These employ a full detector simulation, which is based on measured detector responses. The corrections are based on the particle jet responses defined as

$$R_{\text{ptcl}}\left(\langle p_T \rangle, \eta\right) = \frac{\langle p_T \rangle}{\left\langle p_{T,\text{ptcl}} \right\rangle}\left[p_{T,\text{ptcl}}, \eta\right], \tag{4.4}$$

where ptcl refers to the particle jets that do not take into account detector simulation. The angular brackets refer to the average value over jets at limited p_T and η ranges. The MC truth corrections are naturally prone to any shortcomings in the simulation steps.

The responses given in Eq. (4.4) are fully generic and only depend on the p_T and η of the jet. Ideally, the calibration would be performed particle-by-particle, or at least considering the distribution of jet energy fractions in each jet. It is obvious that the response of a jet with a dominant neutral hadron content is different from e.g. a jet with large contributions from charged leptons. A particle(-type)-wise calibration might be feasible in the future, as the algorithm development progresses. However, moving from the collective approach that considers only whole jets at the time to the one that considers particle categories will require a substantial algorithm re-design.

Residual Corrections Between Data and Simulation

The **Residual Corrections** are a finishing touch for the JECs, where the jets in data are scaled to match simulations. If the detector simulation and collision modelling were perfect, these corrections would be equal to unity. The relative corrections are found by studying the jet responses R. Conventionally, two alternative response definitions are used:

$$R_{\text{jet},p_T} = \frac{p_T^{\text{probe}}}{p_T^{\text{tag}}} \tag{4.5}$$

$$R_{\text{jet,MPF}} = 1 + \frac{\vec{p}_{\text{T}}^{\text{miss}} \cdot \vec{p}_{T,\text{tag}}}{p_{T,\text{tag}}^2} \tag{4.6}$$

The former variant is a simple scale between p_T's, referred to as p_T balance. The latter method is referred to as MPF, i.e. Missing transverse momentum Projection Fraction. MPF is motivated mathematically by the fact that the sum of the $\vec{p}_T$ values of the objects in an event is $\vec{0}$, if neutrinos do not play a part. A *tag* object is measured at a higher precision, while the residual JEC calibrations for the *probe* jet kinematics are being performed. The distinction between the tag and the probe is discussed more later.

For the derivation of the residual corrections, the following kind of ratios between responses R in data and simulation are used:

$$\frac{R_{\text{Data}}}{R_{\text{MC}}}. \tag{4.7}$$

These are used in an involved manner to make the responses in data and simulation agree.

In the usual workflow, dijet results are first used to derive residual corrections with respect to η. Dijet events consist mainly of two jets, so that the rest of the jets in the event have significantly lower p_T values. One jet must be in the barrel region of the detector ($|\eta| < 1.3$) and it is used as the tag, while the other functions as the probe. The impact of any additional jets is extrapolated to zero, so that only the tag and the probe, originating from the hard process remain. At the limit of the extrapolation, the tag and the probe should be situated back-to-back in the transverse plane (i.e. at opposite ϕ values).

Consequently, a **Global Fit** is performed to find the complete residual corrections. Originally, the global fit brought corrections derived in photon+jet and Z+jet events on top of the dijet results. These channels are used to derive the p_T dependence of the residual corrections, ranging from around 30 GeV up to 1000 GeV. Later on, more complicated methods have been added to the global fit.

The photon+jet and Z ($\rightarrow \mu\mu/ee$) + jet events are analogous to dijet events and they can be studied in a similar manner as presented above. Now, the photon or charged lepton pair functions as a tag. At the upper p_T limit of around 1000 GeV the

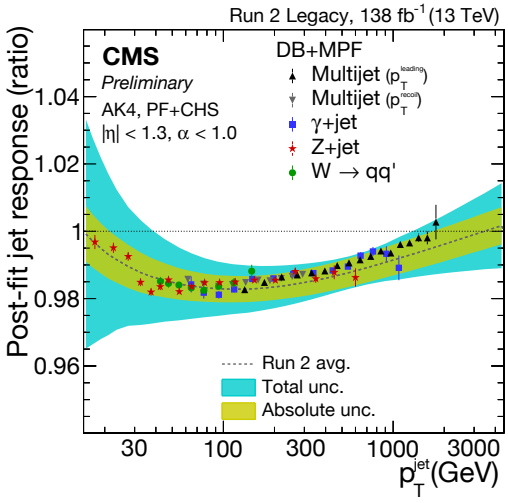

Figure 4.9: The L2L3 Residual JEC Global Fit for 2016-2018. [99].

statistical significance of these measurement channels becomes limited, and hence the *multijet balance* method was introduced. Here, a group of lower-p_T jets are used to balance a single high-p_T jet. Consequently, the JEC derivation stretches above 1000 GeV. The current p_T ranges and measurement channels in the global fit of the residual corrections can be viewed in Fig. 4.9. Here, a great agreement between the measurement channels is displayed.

The remaining channel in Fig. 4.9 not already explained is the hadronic W$\rightarrow qq'$ channel from semileptonic $t\bar{t}$ production events. The thesis author has contributed to the development of this method, and it is currently in a more mature state. It takes advantage of the great amount of jet energy scale information that can be provided by the W resonance. Earlier, this has only been used for the *in situ* jet energy scale tuning in m_t measurements, but now it has been brought into the global JEC calibration. Furthermore, also the use of jet energy fractions in the global fit has been tested out, but this approach has not reached equal maturity.

To summarize, the JEC chain in simulations consists of the Offset and MC truth corrections. For data, the additional step of the Residual corrections is required. On top of these corrections, it is possible to apply also e.g. **flavor corrections**. These make a distinction between b jets and light quark jets. While the studies of this thesis were performed, the flavor corrections for the 2017–2018 datasets were still a work in progress.

Jet Energy Resolutions

A final category of necessary jet energy scales is the smearing of Jet Energy Resolutions (JERs), which is applied on simulation using scale factors. The JERs are a feature of jets, which do not fully match between data and simulation. The term

smearing follows from the fact that resolutions in data are typically slightly higher than those in simulation. The resolutions depend on the energy density ρ and jet η, and are parametrized as a function of jet p_T. These are measured separately both for data and simulations.

To correct for the JER mismatch between data and simulations, purely η-dependent scale factors are provided for the simulations. Employing the scale factors, the preferred way of making JERs match between data and simulations is applying the following jet energy scale:

$$c_{\text{JER}} = 1 + (s_{\text{JER}} - 1)\,\frac{p_T - p_{T,\text{ptcl}}}{p_T}. \tag{4.8}$$

Here, s_{JER} is the scale factor and p_T and $p_{T,\text{ptcl}}$ refer to the jet and particle jet transverse momenta. This is the **scaling method** for JER smearing, which is dependent on a direct match with a particle jet.

In this work the **hybrid method** for JER smearing is employed, where a secondary approach is taken when no direct match is available. This fallback approach is the **stochastic method**, where the smearing jet energy scale is found to be

$$c_{\text{JER}} = \mathcal{N}\left(1, \sigma_{\text{JER}}\sqrt{\max\left(0, s_{\text{JER}}^2 - 1\right)}\right). \tag{4.9}$$

Here $\mathcal{N}$ refers to the normal distribution and σ_{JER} is the JER. The stochastic method is only able to increase the JER, but typically this is the required case.

The factor c_{JER} is not considered to be a correction on the jet energy. However, the JER energy scales have a factual impact on the jet energies and hence they are analogous to the JECs. This implies that such scales can lead to migration over cut limits and bin edges. This is not statistically optimal e.g. when systematic uncertainty variations are performed for the JER scale factors, Changing scales on jet energy can change the set of jets and events that are analyzed. In the future, a weight-based JER scale factor method would be preferable.

Additional Remarks and $\vec{p}_{\text{T}}^{\,\text{miss}}$

As the momenta of the jets are scaled, the effect of the scales must be conveyed properly to the value of $\vec{p}_{\text{T}}^{\,\text{miss}}$. This requirement also includes the JER smearing scales in simulation. The same reasoning also holds for electron, photon and muon energy scale corrections, but the effect of these is generally much smaller than that of jets.

The uncertainties in all of the energy scales are uncertainties for the value of $\vec{p}_{\text{T}}^{\,\text{miss}}$, meaning that the $p_{\text{T}}^{\text{miss}}$ resolution is much poorer than that of jets and other physical objects. This is an important conclusion for collisions where a single energetic neutrino is produced, such as the semileptonic $\text{t}\bar{\text{t}}$ events. The transverse

momentum of the neutrino can be estimated from the value of $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$, but there are great uncertainties involved in this value.

For the uncertainties in $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ one must also consider the **unclustered energy**. PF Jets are not reconstructed below a few GeV, meaning that in a single collision event a portion of the particles are not included in the final set of jets. Also the momentum uncertainties of these particles need to be considered in the full $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ uncertainty.

As $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is a composite object, accidental bugs are occasionally introduced in the $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ calculation routines. The author of this thesis has helped to elegantly solve a known bug in the propagation of JER smearing to $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ in the central CMSSW distribution [100]. Furthermore, the thesis author has single-handedly found and solved an indexing bug in the JEC application in the central CMSSW [101]. The latter bug resulted into some JEC levels not being correctly applied in the simulations in certain conditions.

4.3.2 Jet Veto Maps

In precision measurements it is necessary to keep track of detector areas with divergent behavior. Such measurements include the JEC calibration and also e.g. m_t measurements. If a divergent $\eta - \phi$ sector in the detector is included while deriving the JECs, this will bias the JECs around the whole detector. It is better to derive the JECs with the best-functioning parts of the detector and find a separate solution for the divergent zones. Special $\eta - \phi$ dependent JECs for the divergent detector zones can be derived *a posteriori*, or alternatively these zones can simply be given greater systematic JEC errors. The most extreme strategy is to completely remove the jets in such zones from the analysis. This approach is taken in the m_t measurement of this thesis, where extreme precision is desired. For the 2017 and 2018 datasets, this means approximately an acceptable 15% reduction in statistics.

The author of this thesis has made heavy contributions both into the development of the jet veto strategies and the derivation of the jet veto maps. The maps have been applied in the legacy dataset JEC calibration and now in the m_t analysis of this thesis.

Theoretical Definition

Divergent behavior is best found with a statistical approach. The approach can be based on the cylindrical ϕ-symmetry of the detector, studying a fixed η slice at the time.

A physically motivated $\eta - \phi$ division is found in the granularity of the calorimeters. The most meaningful granularity is that of the HCAL towers, which is also used for the JECs. A single HCAL tower matches exactly 5×5 ECAL towers. In the central η zones of the detector, the HCAL towers form patches of

$\Delta\eta \times \Delta\phi \approx 0.087 \times 0.087$ [56]. As the jet radius is 0.4, a single jet can span over multiple towers. Still, the majority of the jet energy usually lands into the HCAL tower matching the jet axis.

The original algorithm for the jet veto zones was designed by the thesis author's supervisor, Asst. Prof. Mikko Voutilainen. The full implementation for it is available in Ref. [1] and the analysis is performed in the Multijet event topology. It considers all the symmetric HCAL towers in a single η block at the time. Within the given η block, statistically divergent jet counts can be distinguished. The detector components are constantly degrading from radiation damage, which is in turn compensated by calibrations. The jet counts are an excellent tool for probing out bad calibrations within the detector.

In the first iteration of the algorithm, the jet counts and their sample deviations are computed:

$$\mu_\eta^{trg} = \frac{\sum_{i=1}^{N_\phi} N_{jet,\eta}^{i,trg}}{N_\phi}, \tag{4.10}$$

$$s_\eta^{trg} = \sqrt{\frac{\sum_{i=1}^{N_\phi} \left(N_{jet,\eta}^{i,trg} - \mu\right)^2}{N_\phi - 1}}. \tag{4.11}$$

Here, $N_\phi = 72$ is the count of symmetric HCAL towers found at fixed η in the ϕ direction. The tag trg stands for the HLT jet trigger, given in Table 4.1. That is, this step of the analysis is performed for all the single-jet triggers separately. For each trigger, only jets above the current trigger turn-on and below the trigger turn-on of the next HLT_PFJet trigger are considered.

A second iteration is performed, excluding any HCAL towers that differ by over $3s_\eta^{trg}$ from the mean μ_η^{trg}. This step intends to trim long tails of the distributions from the analysis. Three standard deviations mark a sweet spot for this analysis, as it covers 99.7% of the statistics and excludes tails. The selection makes the total count of considered towers a variable, $N_\phi^{trg}(\eta) \leq 72$.

The values of μ_η^{trg} and s_η^{trg} after the second iteration are used to find trigger-wise excesses or lacks of jets. A **cold zone** has less than $\mu_\eta^{trg} - 3s_\eta^{trg}$ jets, whereas a **hot zone** has more than $\mu_\eta^{trg} + 3s_\eta^{trg}$ jets. A hot or cold zone is not a great problem in itself, but if it is not modeled well in the simulations, it should be taken seriously. Historically, the hot zones tended to be more of a problem, but in the 2017-2018 runs also the cold zones started to have issues. The current jet veto maps are constructed in combination of both hot and cold zones. With the choice of a 3σ limit and an assumption of Gaussian fluctuations, only one hot/cold HCAL tower in 300 can be explained by statistical noise.

For a single trigger, the hot and cold zones can indicate issues also with the HLT instead of the subdetectors. To exclude such effects, the analysis combines the information from as many triggers as possible. The high-p_T triggers are active only

at lower $|\eta|$ values, so this combination needs to proceed stepwise as a function of $|\eta|$. To consider a trigger for a given η strip of HCAL towers, at least 14 of the 72 symmetric ϕ towers need to contain at least a single jet. To help the trigger combination the following function is defined:

$$f(x) = \begin{cases} 0, & \text{if } x < 3 \\ x, & \text{else} \end{cases} \tag{4.12}$$

To determine the hot zones within a single $\eta - \phi$ bin, the $3\sigma+$ fluctuations are summed up from the different triggers:

$$\Delta_{tot}^{\eta,\phi} = \frac{1}{N_{Trg}^{Avail}} \sum_{trg=1}^{N_{Trg}^{Avail}} f\left(\frac{\left(N_{jet,\eta}^{i_\phi,trg} - \mu_\eta^{trg} \right)}{s_\eta^{trg}} \right) \tag{4.13}$$

The value of N_{Trg}^{Avail} indicates the count of triggers that provide sufficient statistics at the current fixed η value. The greater N_{Trg}^{Avail} is, the more universal is the issue behind a divergent $\Delta_{tot}^{\eta,\phi}$ value. Algorithmically it is necessary to set a threshold σ_{tot}^{thresh} for the vetos, so that an HCAL tower is vetoed when $\Delta_{tot}^{\eta,\phi} > \sigma_{tot}^{thresh}$. If less triggers are available, the combined statistics are worse and a higher total variation should be required. The threshold values are determined as a function of N_{Trg}^{Avail}:

$$\sigma_{tot}^{thresh} = \begin{cases} 4, & \text{if } N_{Trg}^{Avail} = 1 \\ 3, & \text{if } N_{Trg}^{Avail} = 2, 3 \\ 2, & \text{if } N_{Trg}^{Avail} = 4, 5, 6 \\ 1.5, & \text{if } N_{Trg}^{Avail} = 7, 8 \\ 1, & \text{else,} \end{cases} \tag{4.14}$$

that is, for a single trigger we require 4σ, for 2 or 3 triggers in average 3σ per trigger, and so on. Looking at the variety of triggers, these have been empirically found to be sensible threshold choices. To determine the cold zones, an expression similar to Eq. (4.13) is used. The only exception is that the arguments of the function f are multiplied by -1.

The theoretical contribution of the thesis author was in replacing the jet count by an analogous sum of weights:

$$N_{jet,\eta}^{i,trg} \to \sum_{j \in jets} w_{j,\eta}^{i,trg}. \tag{4.15}$$

The original analysis corresponds to $w_{j,\eta}^{i,trg} = 1$. Using these weights, it is possible to separate the above process for the different PF candidate types. This is done by setting the value of $w_{j,\eta}^{i,trg}$ for each jet equal to the different PF jet energy fractions.

That is, the algorithm is run separately for electrons, muons, photons, charged hadrons and neutral hadrons.

This relatively simple variation of the original method opens up a view to separate detector components. When problematic behaviour is seen most of all for neutral and charged hadrons, the problem is most likely with the HCAL. If problematic behavior is centered around photons, electrons and charged hadrons, the issue probably lies within ECAL. This analysis method has been remarkably successful for the Run 2 legacy reconstructions, which were mainly produced during 2020-2021.

Results

The Run 2 legacy reconstructions have been essential for this thesis, as the previous reconstructions have not provided sufficient precision. With the new reconstruction, also a new batch of JECs was derived. For these JECs, the new jet veto strategies were fully employed – providing one more layer of added precision.

In Fig. 4.10 the jet veto mappings for the 2017–2018 datasets are displayed. In all the legacy runs, the greatest problems have been encountered with the HCAL. In 2017 the HCAL wedge HEP17 (HCAL Endcap Plus, wedge 17 – in the lower right corner) became for the most part highlighted by the jet veto analysis. For completeness the remaining parts of this wedge were manually added to the 2017 veto map. In 2018 major issues with the HCAL wedges HEM15-16 (HCAL Endcap Minus, wedges 15–16 – in the lower left corner) were detected. The HEM issues were already previously known, but locating them displayed the power of the analysis. Furthermore, in the latter parts (C and D) of legacy 2018 a problematic area was found in the central barrel region.

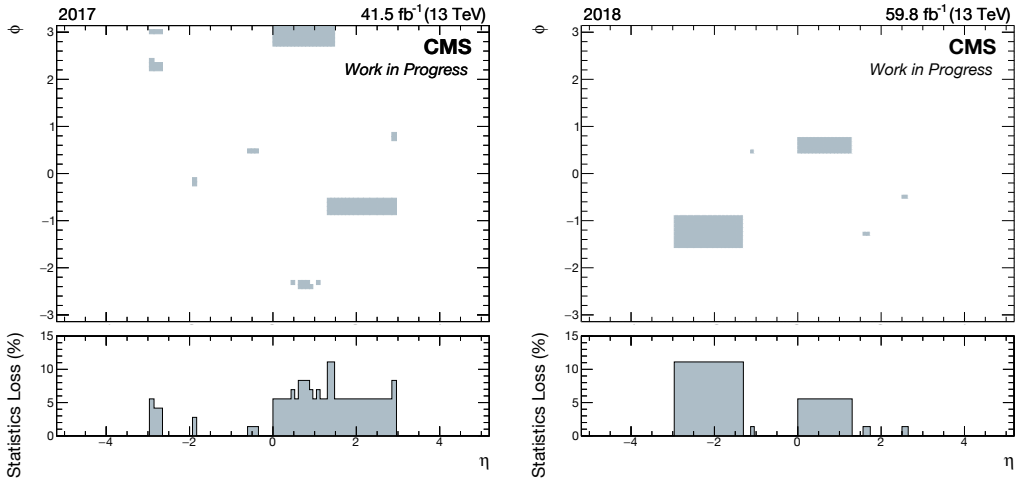

Figure 4.10: The jet veto mapping for the 2017–2018 legacy datasets.

5 Re-interpretation of the DØ Top Mass Measurement

The studies below are based on work that is done privately, outside of the CMS and DØ affiliations. A majority of this analysis has been previously published in the arXiv pre-print [4]. The pre-print was provided to support conversation around the topic with the DØ authors and other involved people before this thesis was given out.

The motivation for these studies arises from the recalibration of the DØ flavor-dependent JECs, presented in Ref. [3]. The recalibration was driven by the observation that the corrections for DØ Run IIb differ notably from those of Run IIa. The significant conclusion of the numerical analysis of Ref. [3] was that the Run IIb corrections should have resembled those of Run IIa more.

The recalibrated JECs have an influence on the reconstructed top quarks through the jets that are produced in top quark decays. Hence, a change in the flavor-dependent JECs implies the need to revise the top quark mass measurement, performed by DØ in the lepton+jets channel [36]. The DØ top quark mass average is almost completely determined by this measurement [102], and hence it has a direct impact on the top quark mass world average, which was most recently officially computed in Ref. [35].

A complete revision of the DØ top quark mass measurement is unlikely to occur. In this chapter, the top quark mass shift resulting from a JEC recalibration is demonstrated using a more lightweight methodology. Based on Ref. [4], these results were discussed during 2020, leading to the TOP2020 conference and the proceedings in Ref. [5].

Much interest in the topic was sparked, but the final reaction from DØ was limited to a short 3-page letter [103]. The counter-arguments for our results were placed in three categories:

1. **The numbers and figures in the DØ analysis notes (ANs) cannot be**

trusted.

2. **There are η-extrapolations in the results of Ref. [3], so an even deeper study without extrapolations would be needed to turn the heads of the DØ authors.**

3. **It is claimed that *There were no interactions between the authors of Refs. [1, 4] and the primary DØ authors concerning the details in the internal notes.***

Without going into much more detail, these arguments were easily countered [5]:

- If the numbers and figures in the ANs cannot be trusted, how can the end results be trusted, as they are reviewed based on the ANs?

- The η-extrapolations provided in the studies of Ref. [3] were physically meaningful, and their expected systematic uncertainty is smaller than the great magnitude of the found m_t shift.

- Especially in the areas without extrapolation the qualitative results of Ref. [3] hold: DØ does not address the great tension between the official Run IIa and Run IIb flavor-dependent JEC results, which we **did not** observe using the DØ source material for the central η-region, which carries the most accurate JEC reconstruction.

- The magnitude of the mass shift found below becomes valid already by simply assuming that both Run IIb and Run IIa flavor JEC results should have been generally in line with Run IIa.

- While the results of Ref. [3] were derived, deep interaction with the DØ authors was necessary for gaining sufficient reference material: the arguments of a lack in communication seem to be highly polarized.

- Our main request has been for DØ to open up a small section of their b jet calibration to settle whether or not the conclusions in Ref. [3] were valid.

- If these kind of results are not enough to spark a greater activity from large collaborations like the DØ, are the self-corrective ideals of science properly fulfilled?

Until DØ agrees to fully open its b jet calibrations, it is meaningless to continue these studies. The end result derived below is a level of strong doubt on the m_t measurement, but these results cannot be confirmed or refuted without further DØ involvement. If the potential weaknesses in the DØ Run IIb flavor JEC calibration were confirmed, the results of this section could be taken as a direct estimate of the arising impact on the DØ m_t measurement.

Against this background, this topic is best suited to be directly integrated into this thesis, without pursuing a separate publication of the pre-print [4]. The original arXiv text was completely written by the author of this thesis, so no conflict of interest exists. Here, some of the sections in Ref. [4] are simplified and shortened, and hence the arXiv reference functions as the analysis note containing full technical details. The review of the DØ analysis strategy also provides a great historical reference for the CMS m_t measurement of Section 6.

5.1 The DØ lepton+jets Top Mass Measurement

Separate flavor dependent JECs (F_{corr}) are applied for light quark jets and b jets. The DØ lepton+jets top quark mass measurement relies on the ME method, where the top quark mass is evaluated on the level of the ME. On this level there are two W-boson resonances in addition to the two top quark resonances. Hence, the ME method makes efficient use of the known W-resonances. In the following, the ME method is analyzed in detail, based on the DØ paper of Ref. [36].

5.1.1 The Matrix Element Method

The ME method makes a connection from the measured detector data down to the ME-level, from where the top quark mass is extracted. It assumes that the measured momentum of the charged lepton is close to that at the ME-level, while the neutrino momentum is determined indirectly. The transverse component of the neutrino momentum is determined from the transverse momentum of the $\mathrm{t\bar{t}}$ system. In contrast, the z-component is found by employing the leptonic W resonance, and the W boson mass.

As the lepton momenta are thus determined, significant freedom is left in the non-trivial connection between quarks and jets. The energies of reconstructed jets (E_j) are mapped to the corresponding ME-level quarks (E_q) using **Transfer Functions** (TF). These operate as probability distributions projected from E_j to E_q. We have chosen to use the subscripts q and j to increase clarity, in contrast to the corresponding DØ notation y and x. Assuming a semi-Gaussian shape, the Transfer Functions can be thought to consist of a smearing resolution and a jet-to-parton correction factor. Technically speaking, the jet-to-parton correction should be treated as a separate source of JEC-related systematic errors. However, it is only assumed that this will be accounted for in the assessment of method errors.

As the masses of quarks are known, the quark four-momentum is determined solely by the energy and the direction of the quark. The DØ measurement assumes that the quark directions are equal to those of the corresponding jets. Hence an energy-based Transfer Function is sufficient for determining the quark properties.

The Transfer Functions are calibrated using simulations. We denote the func-

tional form of a Transfer Function by $\mathrm{TF}_{\mathrm{era}}(E_j, E_q)$, where

$$\int_R \mathrm{TF}_{\mathrm{era}}(E_j, E_q)\, dE_j \;=\; 1 \tag{5.1}$$

and era $\in$ {IIa,IIb1,IIb2,IIb34} refers to the DØ Run II eras. It is important to note that the DØ jet energy E_j excludes neutrinos and muons. Hence, the Transfer Functions for different types of quarks are quite different, as for instance b jets are enriched in neutrinos and muons.

The Transfer Functions act as weights while connecting the measured final state to the feasible ME-level states. This limits the allowed phase space of the ME-level states, and permits an integral over them. Such an integral gives a probabilistic view of the ME-state for each event candidate. The related event probability (P_{evt}) is dependent on the underlying ME-level top quark mass (m_t).

To improve the measurement accuracy, DØ introduces a slack-parameter K_{JES}, which is used for an in-situ jet energy calibration:

$$E'_j = \frac{E_j}{K_{JES}}. \tag{5.2}$$

The calibration is applied by replacing the measured jet energy E_j with the interpreted jet energy E'_j. This is performed exclusively in the top quark mass measurement, and does not affect the calibration of the Transfer Functions. As Eq. (5.2) shifts the difference between the jet and quark energies, the event probability P_{evt} is dependent on K_{JES} in addition to m_t.

The measured collision dataset corresponds to a set of P_{evt} values. A combined likelihood function is constructed from these in the $m_t - K_{\mathrm{JES}}$ plane:

$$\mathcal{L}(m_t, K_{\mathrm{JES}}) \propto \prod_{i \in \mathrm{Data}} P^i_{evt}(m_t, K_{\mathrm{JES}}). \tag{5.3}$$

DØ interprets this likelihood as an unnormalized probability distribution. It attains approximately a 2D-Gaussian form as a function of m_t and K_{JES}.

In the likelihood $\mathcal{L}(m_t, K_{\mathrm{JES}})$ there are two parameters, but the underlying ME has three degrees of freedom: two top resonances and the hadronic W boson resonance. The leptonic W resonance is completely consumed by neutrino reconstruction. The top resonances are dependent on m_t, but also on K_{JES}. On the other hand, the hadronic W resonance depends only on K_{JES}.

The global maximum likelihood value of $\mathcal{L}(m_t, K_{\mathrm{JES}})$ must be such that the hadronic W boson data agrees best with the W mass world average. On the quark level, this argument implies that the average hadronic diquark mass equals the W boson mass:

$$\left\langle m_{q_1 q_2} \right\rangle = m_W. \tag{5.4}$$

As the hadronic W resonance depends only on K_{JES}, this means that the maximum likelihood value of K_{JES} and the hadronic W resonance are directly connected.

Hence the optimal value of K_{JES} can be fixed without referring to the top quark mass resonances.

At a given value of K_{JES}, the relationship between the likelihood function and the two top quark mass resonances is purely determined by m_t. If separate values m_t^{had} and m_t^{lep} were used for the hadronic and leptonic resonances in the ME, it is likely that the global maximum likelihood value would be reached at separate m_t values: $m_t^{had} \neq m_t^{lep}$. Therefore, the full m_t dependence of the likelihood function is determined by the combination of the two resonances.

Assuming that the 2D-Gaussian hypothesis is accurate, the probability distribution can be projected onto the m_t (K_{JES}) axis to obtain a 1D distribution. It retains the correct maximum likelihood position of the full 2D likelihood. For m_t, the projected likelihood is

$$\mathcal{L}(m_t) = \int_{\mathbb{R}} dK_{\mathrm{JES}} \mathcal{L}(m_t, K_{\mathrm{JES}}). \tag{5.5}$$

Treating the likelihood $\mathcal{L}(m_t)$ as an unscaled probability, we find the basic formula for mass extraction at the DØ:

$$m_t^{fit} = \langle m_t \rangle_0 = \frac{\int_{\mathbb{R}} dm_t m_t \mathcal{L}(m_t)}{\int_{\mathbb{R}} dm_t \mathcal{L}(m_t)}. \tag{5.6}$$

The Eqs. (5.5,5.6) are valid also for K_{JES} by interchanging it with m_t. Thus, the ME method combines the theoretical ME with the information extracted from simulations. The latter is completely compressed into the Transfer Functions.

To perform a numerical integration of $\mathcal{L}(m_t, K_{\mathrm{JES}})$, a grid of (m_t, K_{JES}) values is used. If the grid is sufficiently wide and dense, the whole $\mathbb{R}^2$ is effectively covered by the integral. Each event-wise probability P_{evt}^i needs to be evaluated at each grid point.

5.1.2 Calibration of the Matrix Element Method

Like most top quark mass measurements, the DØ measurement utilizes simulations and the underlying generator-level top quark mass (m_t^{gen}). Eq. (5.6) is based on the assumption that the theoretical ME integral with a top quark pole mass agrees directly with the simulation-based Transfer Functions. DØ studies this assumption by assuming that m_t^{gen} is equal to the ME-level pole mass. Thus, the m_t^{fit} value extracted from simulation results is expected to be equal to m_t^{gen}. To find the true connection between m_t^{fit} and m_t^{gen}, DØ employs simulated pseudo-experiments, in which multiple values of m_t^{gen} are used. The connection is parametrized as a linear mapping:

$$g : m_t^{gen} \rightarrow m_t^{fit}. \tag{5.7}$$

DØ denotes the mapping parameters as

$$m_t^{fit} = g(m_t^{gen}) = \mathcal{S}\,(m_t^{gen} - 172.5) + \mathcal{O} + 172.5, \tag{5.8}$$

where $\mathcal{S}$ stands for scale and $\mathcal{O}$ for offset. Typically $\mathcal{S}$ differs slightly from unity and $\mathcal{O}$ is a small non-zero number. This shows the necessity of the calibration.

The DØ m_t extraction procedure goes as follows: the likelihood function is calculated from the measured data, and from this m_t^{fit} is computed. Using the inverse mapping g^{-1} of Eq. (5.7), m_t^{fit} can be mapped to a value analogous to m_t^{gen}. For clarity, this result for data is denoted by m_t^{calib} instead of m_t^{gen}:

$$m_t^{\text{calib}} = g^{-1}(m_t^{fit}). \tag{5.9}$$

Thus, the DØ analysis chain assumes that m_t^{gen} and m_t^{calib} correspond to the top pole mass. In truth, a subtle difference appears to exist between the pole mass and m_t^{gen} [38].

For the calibration of K_{JES}, similar steps are taken. The statement of Eq. (5.2) can be multiplied by a *generator-level* energy scale, K_{JES}^{gen}. Ideally, K_{JES}^{fit} should be equal to K_{JES}^{gen} to keep the hadronic W resonance at a constant position. By applying multiple K_{JES}^{gen} gen values in the simulated pseudo-experiments, a linear mapping similar to that in Eq. (5.7) is found:

$$f : K_{\text{JES}}^{gen} \rightarrow K_{\text{JES}}^{fit}. \tag{5.10}$$

For this function, the following parametrization is used:

$$K_{\text{JES}}^{\text{fit}} = f(K_{\text{JES}}^{gen}) = 1 + \mathcal{S}_K(K_{\text{JES}}^{gen} - 1) + \mathcal{O}_K. \tag{5.11}$$

The calibration procedure of K_{JES} has no practical impact on the measured m_t value, as it is simply a change of variables on the K_{JES} axis. Nevertheless, DØ does not give a complete picture of the effects that are compensated by this mapping. One could argue that the main importance is in reducing any disagreement between the simulations and the ME integral. In the DØ measurements, $f(1)$ can differ notably from the value 1, which implies that the mapping f may be used to compensate errors in the shifts imposed by the Transfer Functions.

It is not necessary to calculate m_t^{calib} and $K_{\text{JES}}^{\text{calib}}$ manually using the inverse functions of Eqs. (5.7,5.10). It can be shown (see Appendix B) that by replacing m_t by $g(m_t)$ and K_{JES} by $f(K_{\text{JES}})$ in the likelihood functions, the fits automatically yield correct results.

However, the DØ papers show some ambiguity on this subject, as is further discussed in Appendix B. DØ indicates that instead of the functions f and g, their inverse functions are inserted into the likelihood. The author of this thesis has been in correspondence with DØ authors about the subject, but the investigation has been inconclusive. Correcting such a misplacement would lower the DØ top quark mass value by 0.1 GeV and make the fit χ^2 values more sensible. As this is a matter of fine-tuning, we will not concentrate on it outside Appendix B.

5.1.3 Steps Between the Matrix Element and Detector Data

To understand the DØ measurement, we need to understand the relationship between jets and quarks. The full difference between E_j and E_q in simulation stands as

$$E_j = F_{\text{Corr}} \otimes \text{JEC} \otimes \text{Detector Sim.} \otimes \text{Hadronization} \otimes \text{Showering} \otimes E_q, \quad (5.12)$$

where we have taken into account Jet Energy Corrections (JECs) and all simulation steps. As a reminder, F_{Corr} stands for the flavor dependent JECs. The symbol $\otimes$ underlines that some of the steps are more complicated than simple multiplications, but can be approximated as such. For instance, the full complexity of Showering and Hadronization can be appreciated by reviewing these subjects in the Pythia 6 manual [32]. For data, the corresponding equation reads as

$$E_j = \text{JEC} \otimes \text{Detector} \otimes \text{Physics} \otimes E_q. \quad (5.13)$$

In general, the step Physics $\otimes E_q$ is only partially understood. To gain further understanding of it, the Transfer Functions calibrated on simulations are used. In terms of a Transfer Function, the transition between jet and quark energies reads as

$$E_q = \text{TF} \otimes E_j, \quad (5.14)$$

Eq. (5.14) can be inverted to produce

$$E_j = \text{TF}^{-1} \otimes E_q. \quad (5.15)$$

The Transfer Functions are calibrated purely using simulations. Hence, comparing Eqs. (5.12) and (5.15), we have

$$\text{TF}^{-1} = F_{\text{Corr}} \otimes \text{JEC} \otimes \text{Detector Sim.} \otimes \text{Hadronization} \otimes \text{Showering}. \quad (5.16)$$

Eq. (5.16) is valid for both data and simulation, as it is the written-out form of the Transfer Functions. The steps within it are interpreted as those of the Transfer Function calibration simulated sample. This underlines the fact that for data, we do not know what E_q is for certain. The analysis is based on the assumption that after all the calibrations, the quark level in data corresponds to that in simulation.

In the top quark mass measurement, jet energies E_j are replaced with E'_j from Eq. (5.2) using the calibration from Eq. (5.10), so that Eq. (5.15) reaches the form

$$E_j = f\left(K_{\text{JES}}\right) \otimes \text{TF}^{-1} \otimes E_q. \quad (5.17)$$

By utilizing Eq. (5.16), Eq. (5.17) can be expressed as

$$E_j = f\left(K_{\text{JES}}\right) \otimes F_{\text{Corr}} \otimes \text{JEC} \otimes \text{Detector Sim.} \otimes E_j^{gen}, \quad (5.18)$$

where we have made the substitution

$$E_j^{gen} = \text{Hadronization} \otimes \text{Showering} \otimes E_q. \quad (5.19)$$

E_j^{gen} is the **generator-level** jet energy. By using the same simulation settings as DØ, E_j^{gen} values analogous to those of DØ can be produced.

5.2 Phenomenological Propagation of Changes in F_{Corr} to Top Mass

In this section, the impact of a F_{Corr} recalibration on the lepton+jets top quark mass measurement is shown on a phenomenological level. The tag DØ is used for the original F_{Corr} values, and the recalibrated ones are expressed without a tag.

5.2.1 F_{Corr} and the Interpreted Quark Energy

As was shown in Eq. (5.16), the values of F_{Corr} are propagated to the corresponding Transfer Functions. According to Eq. (5.17), we can write the interpreted E_q for a given fixed measurement E_j:

$$E_q \equiv E_q\left(f\left(K_{\mathrm{JES}}\right)\right) = \frac{\mathrm{TF} \otimes E_j}{f\left(K_{\mathrm{JES}}\right)}. \tag{5.20}$$

That is, by fixing E_j, E_q becomes a function of $f\left(K_{\mathrm{JES}}\right)$. Compared to the more approximate $\otimes$ steps, the division by $f\left(K_{\mathrm{JES}}\right)$ above is exact.

Considering Eq. (5.16) with the original and the new F_{Corr} values, and equating the terms independent of F_{Corr}, we find how F_{Corr} transforms the Transfer Functions:

$$F_{\mathrm{Corr}} \otimes \mathrm{TF} = F_{\mathrm{Corr}}^{\mathrm{D\O}} \otimes \mathrm{TF}^{\mathrm{D\O}}, \tag{5.21}$$

from which it follows that

$$\mathrm{TF}^{\mathrm{D\O}} = \left(\frac{F_{\mathrm{Corr}}}{F_{\mathrm{Corr}}^{\mathrm{D\O}}}\right) \otimes \mathrm{TF}. \tag{5.22}$$

Combining Eqs. (5.20,5.22) gives a relationship between the quark energy interpretations:

$$\begin{aligned} E_q^{\mathrm{D\O}} &= \frac{\mathrm{TF}^{\mathrm{D\O}} \otimes E_j}{f^{\mathrm{D\O}}\left(K_{\mathrm{JES}}^{\mathrm{D\O}}\right)} \\ &= \frac{f\left(K_{\mathrm{JES}}\right)}{f^{\mathrm{D\O}}\left(K_{\mathrm{JES}}^{\mathrm{D\O}}\right)} \otimes \left(\frac{F_{\mathrm{Corr}}}{F_{\mathrm{Corr}}^{\mathrm{D\O}}}\right) \otimes \frac{\mathrm{TF} \otimes E_j}{f\left(K_{\mathrm{JES}}\right)} \\ &= \frac{f\left(K_{\mathrm{JES}}\right)}{f^{\mathrm{D\O}}\left(K_{\mathrm{JES}}^{\mathrm{D\O}}\right)} \otimes \left(\frac{F_{\mathrm{Corr}}}{F_{\mathrm{Corr}}^{\mathrm{D\O}}}\right) \otimes E_q. \end{aligned} \tag{5.23}$$

This indicates that for a fixed jet energy measurement E_j, a change in the F_{Corr} values leads to a change in the interpreted quark energy. To fully understand this result, it is necessary to understand how the change in F_{Corr} is reflected in $f\left(K_{\mathrm{JES}}\right)$.

5.2.2 F_{Corr} and K_{JES} Scaling

In Eq. (5.4) it was argued that the maximum likelihood solution extracted by DØ makes the hadronic diquark system agree in average with the W boson mass. This mechanism sets the value of K_{JES}, and it can be used to find the transformation law for $f\left(K_{\mathrm{JES}}\right)$. The invariant mass of the two zero-mass light quarks can be expressed as

$$m_{q_1 q_2} = \sqrt{2E_{q_1}E_{q_2}(1-\cos\phi_{12})}, \tag{5.24}$$

where E_{q_1} and E_{q_2} are the quark energies and ϕ_{12} is the opening angle between the quark directions. Employing Eq. (5.23), and considering Eq. (5.24) for the interpreted quark energies with the DØ and the recalibrated F_{Corr} values, it is found

$$m^{\mathrm{D\O}}_{q_1 q_2} = \frac{f\left(K_{\mathrm{JES}}\right)}{f^{\mathrm{D\O}}\left(K^{\mathrm{D\O}}_{\mathrm{JES}}\right)} \otimes \left(\frac{F^{lq}_{\mathrm{Corr}}}{F^{lq,\mathrm{D\O}}_{\mathrm{Corr}}}\right) \otimes m_{q_1 q_2}. \tag{5.25}$$

In this expression, we have fused the average F^{lq}_{Corr} value of the two light quarks into a single factor. The following abbreviations for the light quark and b quark F_{Corr} values were adopted:

$$F^{lq}_{\mathrm{Corr}} \equiv F_{\mathrm{Corr,u/d/s/c\ quark}}, \tag{5.26}$$

$$F^{b}_{\mathrm{Corr}} \equiv F_{\mathrm{Corr,b\ quark}}. \tag{5.27}$$

Averaging over Eq. (5.25) and applying Eq. (5.4) for both the DØ and the recalibrated results, it is found

$$m_W = \frac{f\left(K_{\mathrm{JES}}\right)}{f^{\mathrm{D\O}}\left(K^{\mathrm{D\O}}_{\mathrm{JES}}\right)} \otimes \left(\frac{F^{lq}_{\mathrm{Corr}}}{F^{lq,\mathrm{D\O}}_{\mathrm{Corr}}}\right) \otimes m_W. \tag{5.28}$$

Here, the F^{lq}_{Corr} factors are interpreted as average values. Canceling out the m_W terms, Eq. (5.28) implies

$$\frac{f\left(K_{\mathrm{JES}}\right)}{f^{\mathrm{D\O}}\left(K^{\mathrm{D\O}}_{\mathrm{JES}}\right)} = \frac{F^{lq,\mathrm{D\O}}_{\mathrm{Corr}}}{F^{lq}_{\mathrm{Corr}}}. \tag{5.29}$$

Hence, the light quark F_{Corr} values have a scaling effect on the K_{JES} terms.

5.2.3 The Full Impact of F_{Corr} on Quarks

Utilizing Eqs. (5.29,5.23), the full impact of a change in F_{Corr} on the interpreted quark energies can be written:

$$E^{\mathrm{D\O}}_{q} = \left(\frac{F^{lq,\mathrm{D\O}}_{\mathrm{Corr}}}{F^{lq}_{\mathrm{Corr}}}\right) \otimes \left(\frac{F_{\mathrm{Corr}}}{F^{\mathrm{D\O}}_{\mathrm{Corr}}}\right) \otimes E_q. \tag{5.30}$$

The result of Eq. (5.30) can be separately applied for b quarks and light quarks:

$$E_{lq}^{\mathrm{D\O}} = \cancel{\left(\frac{F_{\mathrm{Corr}}^{lq,\mathrm{D\O}}}{F_{\mathrm{Corr}}^{lq}}\right)} \otimes \cancel{\left(\frac{F_{\mathrm{Corr}}^{lq}}{F_{\mathrm{Corr}}^{lq,\mathrm{D\O}}}\right)} \otimes E_{lq} = E_{lq} \tag{5.31}$$

$$E_{b}^{\mathrm{D\O}} = \left(\frac{F_{\mathrm{Corr}}^{lq,\mathrm{D\O}}}{F_{\mathrm{Corr}}^{lq}}\right) \otimes \left(\frac{F_{\mathrm{Corr}}^{b}}{F_{\mathrm{Corr}}^{b,\mathrm{D\O}}}\right) \otimes E_{b}. \tag{5.32}$$

Thus, the interpreted light quark energies of a fixed measurement are the same on the average level, even if the light quark F_{Corr} calibration changes. This is natural, as the invariant mass of the light quark pair is forced to be equal with the W boson mass. All the changes in the F_{Corr} values are hence conveyed to the interpretation of b quarks. In our studies the trend is that $F_{\mathrm{Corr}}^{lq} < F_{\mathrm{Corr}}^{lq,\mathrm{D\O}}$ and $F_{\mathrm{Corr}}^{b} > F_{\mathrm{Corr}}^{b,\mathrm{D\O}}$. Therefore, both of the coefficients in Eq. (5.32) increase the reconstructed b quark energy.

The reconstructed b quark energies have a great impact on the top quark measurement for both the hadronic and the leptonic top quark mass resonances. On the average, approximately half of the energy of a decaying top quark goes to a b quark. For a reconstructed top quark, a systematically amplified b quark energy shows as a systematically amplified top quark mass. Thus, we have a reason to believe that the top quark mass value produced with the original DØ F_{Corr} calibration was systematically shifted upwards. For the accurate numerical evaluation of this shift, an involved analysis is necessary.

5.3 Analysis Strategy

In the previous section it was phenomenologically confirmed that the a recalibration of the F_{Corr} values leads to a shift in the measured top quark mass value. In this section, a method is designed for the numerical evaluation of this shift.

First, we aim to replicate what the DØ data looked like before the original DØ F_{Corr} values were applied. Then, the recalibrated F_{Corr} values are applied. As a result, the DØ top quark mass estimate is shifted downwards, as the interpreted b quark energies become lower according to Eq. (5.32).

5.3.1 A Fully F_{Corr} Dependent Pseudo-Experiment

It is not feasible to replicate all the steps of the DØ measurement, as given in Eq. (5.18). Thus, the DØ data cannot be replicated on the level of the jet energies E_j. Nevertheless, we will attempt to capture the essential F_{Corr} features using a more compact analysis chain. To reach this goal, it is essential to reconstruct jet energies on a level where the same F_{Corr} behavior is observable, as for the full E_j energies. The terms independent of F_{Corr} can then be left out.

By Eq. (5.29) we know that variations of F_{Corr}^{lq} are propagated into the K_{JES} term. To account for this, the $f\left(K_{\text{JES}}\right)$ term can be factorized into two parts:

$$f\left(K_{\text{JES}}\right) = f\left(K_{\text{JES}}\right)_0 \otimes K_{\text{JES}}^{\text{Res}}. \tag{5.33}$$

All the F_{Corr} dependence is pressed into the residual $K_{\text{JES}}^{\text{Res}}$ term, while $f\left(K_{\text{JES}}\right)_0$ holds all the other dependencies. Combining these developments, Eq. (5.18) transforms to

$$E_j = K_{\text{JES}}^{\text{Res}} \otimes f\left(K_{\text{JES}}\right)_0 \otimes F_{\text{Corr}} \otimes \text{JEC} \otimes \text{Detector Sim.} \otimes E_j^{gen}. \tag{5.34}$$

Here, the term $f\left(K_{\text{JES}}\right)_0 \otimes \text{JEC} \otimes \text{Detector Sim.}$ is in average canceled out. In other words, the JECs cancel out the average detector response. For a single event, this collective term has a smearing effect. As we do not have access to the DØ JEC nor Detector simulation, and the relationship between $f\left(K_{\text{JES}}\right)_0$ and $K_{\text{JES}}^{\text{Res}}$ is not known, we are encouraged to make the inversion

$$E_j'' = \left(f\left(K_{\text{JES}}\right)_0 \otimes \text{JEC} \otimes \text{Detector Sim.}\right)^{-1} \otimes E_j = K_{\text{JES}}^{\text{Res}} \otimes F_{\text{Corr}} \otimes E_j^{gen}. \tag{5.35}$$

The result of Eq. (5.35) is used in the main analyses. It is the minimal jet energy definition that takes into account all F_{Corr}-related effects. Using the E_j''-level jet energies instead of the E_j ones is sufficient, as we are interested only in the F_{Corr}-dependent m_t shift. To capture the full F_{Corr}-dependence in E_j'', it was crucial not to cancel out the K_{JES} dependence completely. Through $K_{\text{JES}}^{\text{Res}}$, any changes in the F_{Corr}^{lq} values can be compensated in the desired way.

5.3.2 Resonance Position Estimators

The DØ likelihood function of Eq. (5.3) focuses on mass resonances. To emulate its behavior, we will need to evaluate the positions of the invariant mass resonances on the level of E_j'' energies. For this purpose, **resonance position estimators** are designed. For the three mass resonances of interest, these are denoted by $\hat{m}_W$, $\hat{m}_t^{had}$ and $\hat{m}_t^{lep}$.

The most complicated resonance position estimator is an intricate fit on the resonance mass distribution (**fit**). Generally, the fits are made using generalizations of Gaussian functions – e.g. Voigtian functions – and the resonance position is evaluated from the fit parameters. Moreover, estimators that pick the maximum (**max**) and median (**med**) value from the mass histogram are used. In addition, two different integral-based estimators are used. One takes the mean value based on the full x-range of the mass histogram (**ave**). The other uses a more limited mass window for averaging (**itg**).

The same estimator is always used for estimating all three resonance positions. The m_t shift is computed as a combination of the estimators, considering their relative errors.

The shift studies have been originally performed separately with all the five estimators. The *maximum* method turns out to be the most unstable one, producing magnitudes greater error terms than the other methods. Moreover, the method *fit* faces instabilities in the leptonic branch. Considering the error estimates, the less stable values of the *maximum* and *fit* methods agree with the others and have a small impact on the combined result. To promote clarity, only the three most stable methods (max, med, itg) are presented and used in this thesis, while the also the remaining two methods are found in Ref. [4].

5.3.3 Analysis Workflow

In Eq. (5.35) $K_{\mathrm{JES}}^{\mathrm{Res}}$ refers to the corresponding maximum likelihood value. Nonetheless, while evaluating the $\hat{m}$ values, it is convenient to interpret $K_{\mathrm{JES}}^{\mathrm{Res}}$ as a free parameter. This allows fixing the maximum likelihood value at a later point. In a pseudo-experiment based on a simulation, the value of m_t^{gen} is also a free parameter. By evaluating the $\hat{m}$ values on a grid of m_t^{gen} and $K_{\mathrm{JES}}^{\mathrm{Res}}$ values, dependencies between the $\hat{m}$s and the underlying parameters are found. We expect that $\hat{m}_W$ is sensitive to $K_{\mathrm{JES}}^{\mathrm{Res}}$, and that the $\hat{m}_t$s are sensitive to both $K_{\mathrm{JES}}^{\mathrm{Res}}$ and m_t^{gen}. This is a direct analogy to the DØ likelihood function of Eq. (5.3).

To make numerical analysis more convenient, $\hat{m}_W$ is translated into a K_{JES} estimator:

$$\hat{K}_{\mathrm{JES}} = \frac{\hat{m}_W}{80.4\,\mathrm{GeV}}. \tag{5.36}$$

The approximate W boson mass (80.4 GeV) [62] is used as a convenient normalization factor here, but the following equations could equally well be constructed without any normalization. It is sensible to expect that the dependence between $\hat{K}_{\mathrm{JES}}$ and $K_{\mathrm{JES}}^{\mathrm{Res}}$ is linear. This can be expressed in the functional form

$$\hat{K}_{\mathrm{JES}}\left(K_{\mathrm{JES}}^{\mathrm{Res}}\right) = \frac{\hat{m}_W\left(K_{\mathrm{JES}}^{\mathrm{Res}}\right)}{80.4\ \mathrm{GeV}} = q_0 + q_1 K_{\mathrm{JES}}^{\mathrm{Res}}. \tag{5.37}$$

In analogy to Eq. (5.37), we expect that the $\hat{m}_t$s depend linearly on m_t^{gen}. Moreover, the $K_{\mathrm{JES}}^{\mathrm{Res}}$ dependence is estimated using a linear term. The resulting 2D function reads as

$$\hat{m}_t\left(m_t^{gen}, K_{\mathrm{JES}}^{\mathrm{Res}}\right) = p_0 + p_1 \times m_t^{gen} + p_2 \times K_{\mathrm{JES}}^{\mathrm{Res}}. \tag{5.38}$$

The parameters q_i and p_i of Eqs. (5.37,5.38) are calibrated on a grid of m_t^{gen} and $K_{\mathrm{JES}}^{\mathrm{Res}}$ values. For $\hat{K}_{\mathrm{JES}}$ a degeneracy is present, as the position of the W boson resonance does not depend on m_t^{gen}. This degeneracy improves the $K_{\mathrm{JES}}^{\mathrm{Res}}$ fit quality, as all the data points are used in the fit.

To proceed with the measurement, the maximum likelihood value of $K_{\mathrm{JES}}^{\mathrm{Res}}$ is determined. In Eq. (5.4) we concluded that the maximum likelihood value of K_{JES} is determined by forcing the average invariant diquark mass to be equal with m_W.

Moreover, we known by Eq. (5.31) that for light quarks the interpreted E_q of a measurement E_j is constant, independent of the F_{Corr} values. That is, the relationship between the dijet and diquark resonance positions is constant. The same argument can be made after swapping the E_j energies to E_j'' energies: the maximum likelihood value of $\hat{m}_W$ should be constant, independent of the F_{Corr} values. The argument may be written as

$$\hat{m}_W^{ML} = \mathcal{A} \times 80.4 \text{ GeV}, \tag{5.39}$$

where $\mathcal{A}$ is a scale-parameter close to unity. As the E_j'' energies are abstract, there is no distinct correct choice for $\mathcal{A}$. This contrasts to the E_j energies, where a physically motivated value for $\mathcal{A}$ exists.

By employing Eqs. (5.37,5.39), the maximum likelihood value of $K_{\text{JES}}^{\text{Res}}$ is found:

$$K_{\text{JES}}^{\text{Res},ML} = \frac{\frac{\mathcal{A}\times 80.4 \text{ GeV}}{80.4 \text{ GeV}} - q_0}{q_1} = \frac{\mathcal{A} - q_0}{q_1}. \tag{5.40}$$

Placing this value into Eq. (5.38) leaves a pure linear dependence between $\hat{m}_t$ and m_t^{gen}. In this fashion, the relationship between the m_t^{gen} and $\hat{m}_t^{had}$ or $\hat{m}_t^{lep}$ can be evaluated using both the DØ F_{Corr} values and the recalibrated ones.

In the analysis chain we first produce a $\hat{m}_t$ value by employing the DØ F_{corr} and m_t^{calib} values. The former are used for scaling jets, and m_t^{gen} is set to be equal with the latter. This $\hat{m}_t$ is re-interpreted using the new F_{corr} values. The resulting m_t^{gen} is interpreted as the m_t^{calib} value that DØ should have found. The procedure must be performed for both the hadronic and leptonic $\hat{m}_t$s, and the total m_t^{calib} shift is found through a combination of the results.

Within the DØ analysis, it is possible that the hadronic and leptonic resonances prefer different m_t^{calib} values. The likelihood function of Eq. (5.3) makes a compromise between these two information sources. The relative importance of the two m_t resonances in this process is unknown, and we express it using the parameter $\alpha \in [0, 1]$:

$$m_t^{\text{calib}} = \alpha \times m_{t,had}^{\text{calib}} + (1 - \alpha) \times m_{t,lep}^{\text{calib}} \tag{5.41}$$

The m_t^{calib} shifts need to be evaluated separately for the electron and muon measurement channels. *A priori* these should be almost identical, but in the DØ measurements, the phase spaces and m_t^{calib} values differ. Moreover, separate m_t^{calib} and F_{corr} values are employed for all the four Run II eras: IIa, IIb1, IIb2 and IIb34. The electron and muon channels need to be considered in all the four Run II eras. Hence the combined m_t^{calib} consists of a total of 8 measurements, each of which involves the shifts of the hadronic and the leptonic m_t resonances.

5.3.4 Equivalence to the Phenomenological Approach

To show that the present approach leads to correct results, it is necessary to understand Eq. (5.40), i.e. the constants q_0, q_1 and $\mathcal{A}$. Intuitively, $q_0 \approx 0$ would seem

to be the most natural choice in Eq. (5.37). However, a non-zero q_0 value can be motivated by the features of some resonance position estimators. The value $\hat{m}$ may consists of a sum of the signal (Sgn, fraction β) and the background (Bkg, fraction $1-\beta$):

$$\hat{m} = \beta \times \hat{m}^{\mathrm{Sgn}} + (1-\beta) \times \hat{m}^{\mathrm{Bkg}}. \tag{5.42}$$

The presence of a background term is obvious especially for the integral-based estimators, ave and itg. In a histogram, the invariant mass extracted from some event candidates can be close to the resonance position accidentally. An integral over a mass window in the histogram also considers these background events. For the itg estimator the integration window is typically balanced so that the value of q_0 is close to zero. For the ave estimator the value of q_0 is out of control, as the histogram widths are not varied.

The background term is expected to be a constant that conveys no information about the resonance position. We can apply Eq. (5.35) for the signal component of $\hat{m}_W$:

$$\hat{m}_W^{\mathrm{Sgn}} \left(K_{\mathrm{JES}}^{\mathrm{Res}} \right) = K_{\mathrm{JES}}^{\mathrm{Res}} \otimes F_{\mathrm{Corr}}^{lq} \otimes \hat{m}_W^{\mathrm{Sgn},gen}. \tag{5.43}$$

The term $\hat{m}_W^{\mathrm{Sgn},gen}$ is a constant, independent of the F_{Corr} and $K_{\mathrm{JES}}^{\mathrm{Res}}$ values. We can insert the result of Eq. (5.42) into Eq. (5.37), and express the signal component using Eq. (5.43):

$$\begin{aligned} \hat{K}_{\mathrm{JES}} \left(K_{\mathrm{JES}}^{\mathrm{Res}} \right) &= \frac{\beta \times \hat{m}_W^{\mathrm{Sgn}}}{80.4\ \mathrm{GeV}} + \frac{(1-\beta) \times \hat{m}_W^{\mathrm{Bkg}}}{80.4\ \mathrm{GeV}} \\ &= \frac{\beta \times F_{\mathrm{Corr}}^{lq} \otimes \hat{m}_W^{\mathrm{Sgn},gen}}{80.4\ \mathrm{GeV}} \otimes K_{\mathrm{JES}}^{\mathrm{Res}} + \frac{(1-\beta) \times \hat{m}_W^{\mathrm{Bkg}}}{80.4\ \mathrm{GeV}}. \end{aligned}$$

By comparing this result with the parametrization of Eq. (5.37), we note that

$$q_0 = \quad (1-\beta) \times \frac{\hat{m}_W^{\mathrm{Bkg}}}{80.4\ \mathrm{GeV}} \quad = (1-\beta) \times \hat{K}_{\mathrm{JES}}^{\mathrm{Bkg}}, \tag{5.44}$$

$$q_1 = \quad \beta \times F_{\mathrm{Corr}}^{lq} \otimes \frac{\hat{m}_W^{\mathrm{Sgn},gen}}{80.4\ \mathrm{GeV}} \quad = \beta \times F_{\mathrm{Corr}}^{lq} \otimes \hat{K}_{\mathrm{JES}}^{\mathrm{Sgn},gen}. \tag{5.45}$$

Thus, an explicit form was found for Eq. (5.37). For a certain estimator $\hat{m}$, the fraction β is expected to be constant. Moreover, the background term $\hat{K}_{\mathrm{JES}}^{\mathrm{Bkg}}$ remains constant, making q_0 constant for the given estimator. On the other hand, q_1 depends only on F_{Corr}^{lq}, as the generator-level mass term is constant.

By referring to Eq. (5.45), Eq. (5.40) turns into

$$K_{\mathrm{JES}}^{\mathrm{Res},ML} = \frac{\mathcal{A} - q_0}{\beta \times F_{\mathrm{Corr}}^{lq} \otimes \hat{K}_{\mathrm{JES}}^{\mathrm{Sgn},gen}} = \frac{\mathcal{B}}{F_{\mathrm{Corr}}^{lq}}, \tag{5.46}$$

where

$$\mathcal{B} = \frac{\mathcal{A} - q_0}{\beta \times \hat{K}_{\mathrm{JES}}^{\mathrm{Sgn},gen}} \tag{5.47}$$

is a constant. Now, applying the $K_{\mathrm{JES}}^{\mathrm{Res}}$ value from Eq. (5.46) on Eq. (5.35), we find the light quark and b quark jet energies:

$$E''_{j,lq} = \frac{\mathcal{B}}{\cancel{F_{\mathrm{Corr}}^{lq}}} \otimes \cancel{F_{\mathrm{Corr}}^{lq}} \otimes E_j^{gen} = \mathcal{B} \otimes E_j^{gen}, \tag{5.48}$$

$$E''_{j,b} = \frac{\mathcal{B}}{F_{\mathrm{Corr}}^{lq}} \otimes F_{\mathrm{Corr}}^{b} \otimes E_j^{gen} = \mathcal{B} \otimes \frac{F_{\mathrm{Corr}}^{b}}{F_{\mathrm{Corr}}^{lq}} \otimes E_j^{gen}. \tag{5.49}$$

These expressions are valid for both the recalibrated and DØ F_{Corr} values in all DØ Run II eras. To shift the DØ mass results, we interpret that the E''_j energies in a single era are the same with both F_{Corr} values. However, the relationship to E_j^{gen} can differ. By equating the E''_j energies, we find

$$\mathcal{B} \otimes E_{j,lq}^{gen,\mathrm{D\emptyset}} = \mathcal{B} \otimes E_{j,lq}^{gen}, \tag{5.50}$$

$$\mathcal{B} \otimes \left(\frac{F_{\mathrm{Corr}}^{b,\mathrm{D\emptyset}}}{F_{\mathrm{Corr}}^{lq,\mathrm{D\emptyset}}} \right) \otimes E_{j,b}^{gen,\mathrm{D\emptyset}} = \mathcal{B} \otimes \left(\frac{F_{\mathrm{Corr}}^{b}}{F_{\mathrm{Corr}}^{lq}} \right) \otimes E_{j,b}^{gen}, \tag{5.51}$$

By referring to Eq. (5.19) and noting that the hadronization and showering steps do not depend on F_{Corr}, we can make a direct inversion from E_j^{gen} to quark energies. Moving all the scaling constants to the right-hand-side and performing the inversion, we find

$$E_{lq}^{\mathrm{D\emptyset}} = E_{lq}, \tag{5.52}$$

$$E_{b}^{\mathrm{D\emptyset}} = \left(\frac{F_{\mathrm{Corr}}^{lq,\mathrm{D\emptyset}}}{F_{\mathrm{Corr}}^{lq}} \right) \otimes \left(\frac{F_{\mathrm{Corr}}^{b}}{F_{\mathrm{Corr}}^{b,\mathrm{D\emptyset}}} \right) \otimes E_b. \tag{5.53}$$

This is exactly the result of the phenomenological proof of Eqs. (5.31,5.32), that is, the presented numerical approach is equivalent to the phenomenological one.

5.3.5 Evaluation of the Maximum Likelihood Value of $K_{\mathrm{JES}}^{\mathrm{Res}}$

The coefficient $\mathcal{A}$ remains yet unknown, and therefore it is not possible to evaluate Eq. (5.40) numerically. While deriving Eqs. (5.52,5.53), the constant $\mathcal{B}$ was completely canceled out. All the $\mathcal{A}$-dependence was given in this term. This implies that the value of $\mathcal{A}$ only sets a reference scale, and it can be chosen quite freely. To make the reference scale behave flexibly with different $\hat{m}$ evaluators, we choose

$$\mathcal{A} = \frac{\hat{m}_W^{gen}}{80.4\ \mathrm{GeV}}. \tag{5.54}$$

This sets the maximum likelihood value of the resonance position estimator to be equal with the corresponding generator-level value. By this choice, Eq. (5.40) for $K_{\mathrm{JES}}^{\mathrm{Res}}$ determination reads

$$K_{\mathrm{JES}}^{\mathrm{Res},ML} = \frac{\frac{\hat{m}_W^{gen}}{80.4\ \mathrm{GeV}} - q_0}{q_1} = \frac{\hat{K}_{\mathrm{JES}}^{gen} - q_0}{q_1}. \tag{5.55}$$

To evaluate Eq. (5.55), only the values of q_0, q_1 and $\hat{K}_{\mathrm{JES}}^{gen}$ need to be determined.

5.3.6 Evaluation of the Combination Parameter α

Determining the value of α requires turning to the DØ measurements. The three likelihood plots in Figs. 23 and 25 in Ref. [36] show a falling linear maximum likelihood slope of m_t as a function of K_{JES}. We can understand this behavior by fixing the resonance position estimate $\hat{m}_t$ in Eq. (5.38) to a constant value. In other words, by looking at such pairs of $(m_t^{gen}, K_{\mathrm{JES}}^{\mathrm{Res}})$ values that could correspond to a given measurement, before fixing the maximum likelihood value of $K_{\mathrm{JES}}^{\mathrm{Res}}$:

$$m_t^{gen} = -\frac{p_0 + \hat{m}_t}{p_1} - \frac{p_2}{p_1} \times K_{\mathrm{JES}}^{\mathrm{Res}}. \tag{5.56}$$

This result displays directly the falling linear slope of the likelihood plots. It must be noted that $K_{\mathrm{JES}}^{\mathrm{Res}}$ has the same characteristic behavior as plain K_{JES} in the DØ measurement. Hence, even if Eq. (5.38) is designed for the E_j'' level energies, the slope of Eq. (5.56) is equally valid to the E_j level energies.

The result of Eq. (5.56) can be written separately for the hadronic and the leptonic mass resonances, and the results will differ. In contrast, the plots in Ref. [36] show the combined effect of both of the resonances. We can make the difference between the resonances more concrete by looking at the slope of the hadronic resonance position:

$$\frac{dm_{t,had}^{gen}}{dK_{\mathrm{JES}}} = -\frac{p_2}{p_1} \approx -m_{t,had}^{gen}. \tag{5.57}$$

This is due to the fact that the hadronic peak consists purely of jets, and the reconstructed energy of each of these is proportional to K_{JES}. On the other hand, for the leptonic resonance position only the b jet and the neutrino reconstruction carry K_{JES}-dependence. Hence, the slope of the leptonic resonance is not as steep as that of the hadronic one.

To extract the value of α, we need to combine the hadronic and the leptonic slopes. The combined slope should be equal to the one found in the DØ likelihood plots. Therefore, we need to take the weighted vectorial sum of the normalized hadronic and leptonic slopes, and set this to be equal with the DØ likelihood slope after normalization. Mathematically speaking, this vectorial version of Eq. (5.41) reads as

$$\text{Normalize}\left\{\alpha \times \hat{v}_{had} + (1-\alpha) \times \hat{v}_{lep}\right\} = \alpha_1 \times \hat{v}_{had} + \alpha_2 \times \hat{v}_{lep} = \hat{v}_{\mathrm{D\emptyset}}, \tag{5.58}$$

where $\hat{v}$s are unit vectors. We can take the dot product of Eq. (5.58) with $\hat{v}_{\mathrm{D\emptyset}}$ and a unit vector perpendicular to it ($\hat{v}_{\mathrm{D\emptyset}}^{\perp}$):

$$\alpha_1 \times \hat{v}_{had} \cdot \hat{v}_{\mathrm{D\emptyset}} + \alpha_2 \times \hat{v}_{lep} \cdot \hat{v}_{\mathrm{D\emptyset}} = 1 \tag{5.59}$$

$$\alpha_1 \times \hat{v}_{had} \cdot \hat{v}_{\mathrm{D\emptyset}}^{\perp} + \alpha_2 \times \hat{v}_{lep} \cdot \hat{v}_{\mathrm{D\emptyset}}^{\perp} = 0 \tag{5.60}$$

Solving for α_2 in the latter equation, the former equation yields

$$\alpha_1 = \left(\hat{v}_{had} \cdot \hat{v}_{\mathrm{D\O}} - \hat{v}_{lep} \cdot \hat{v}_{\mathrm{D\O}} \left[\frac{\hat{v}_{had} \cdot \hat{v}_{\mathrm{D\O}}^{\perp}}{\hat{v}_{lep} \cdot \hat{v}_{\mathrm{D\O}}^{\perp}} \right] \right)^{-1} \tag{5.61}$$

It turns out that the coefficient p_2 is generally much larger than p_1 in Eq. (5.38), as was discussed in the hadronic example case of Eq. (5.57). Hence $\hat{v}_{had} \cdot \hat{v}_{\mathrm{D\O}} \approx 1$ and $\hat{v}_{lep} \cdot \hat{v}_{\mathrm{D\O}} \approx 1$. This leads to $\alpha_1 \approx \alpha$ and $\alpha_2 \approx 1 - \alpha$, and Eq. (5.61) reduces to

$$\alpha \approx \alpha_1 \approx \left(1 - \frac{\hat{v}_{had} \cdot \hat{v}_{\mathrm{D\O}}^{\perp}}{\hat{v}_{lep} \cdot \hat{v}_{\mathrm{D\O}}^{\perp}} \right)^{-1} . \tag{5.62}$$

The result implies that the value of α is determined by the small vectorial deviations in the direction perpendicular to $\hat{v}_{\mathrm{D\O}}$. The F_{Corr} values generally have little impact on the value of α, and hence it is calculated as an average from the DØ and recalibrated F_{Corr} scenarios.

It is likely that the optimal hadronic and leptonic m_t^{calib} values were not equal in the DØ measurement, but it is not possible to determine their values. The least biased choice is to set

$$m_t^{\mathrm{D\O}} \equiv m_t^{\mathrm{calib,D\O}} = m_{t,had}^{\mathrm{calib,D\O}} = m_{t,lep}^{\mathrm{calib,D\O}}. \tag{5.63}$$

Thus, the parameter α is only present while producing the combined shifted top quark mass value. With the choice of Eq. (5.63), Eq. (5.41) implies

$$\begin{aligned} m_t^{\mathrm{calib}} &= \alpha \times m_{t,had}^{\mathrm{calib}} + (1-\alpha) \times m_{t,lep}^{\mathrm{calib}} \\ &= m_t^{\mathrm{D\O}} + \alpha \times \Delta m_{t,had}^{\mathrm{calib}} \left(m_t^{\mathrm{D\O}} \right) + (1-\alpha) \times \Delta m_{t,lep}^{\mathrm{calib}} \left(m_t^{\mathrm{D\O}} \right). \end{aligned} \tag{5.64}$$

Here we have expressed the shifted m_t^{calib} values as the difference to their origin, $m_t^{\mathrm{D\O}}$. If the original values of $m_{t,had}^{\mathrm{calib,D\O}}$ and $m_{t,lep}^{\mathrm{calib,D\O}}$ were known, the corresponding result would be

$$\begin{aligned} m_t^{\mathrm{calib}} &= \alpha \times m_{t,had}^{\mathrm{D\O}} + (1-\alpha) \times m_{t,lep}^{\mathrm{D\O}} \\ &\quad + \alpha \times \Delta m_{t,had}^{\mathrm{calib}} \left(m_{t,had}^{\mathrm{D\O}} \right) + (1-\alpha) \times \Delta m_{t,lep}^{\mathrm{calib}} \left(m_{t,lep}^{\mathrm{D\O}} \right). \end{aligned} \tag{5.65}$$

Referring to Eq. (5.41), the first two terms can be identified as $m_t^{\mathrm{D\O}}$. Furthermore, we may assume that slight variations in the point of origin $m_t^{\mathrm{calib,D\O}}$ have little effect on the shifts $\Delta m_t^{\mathrm{calib}}$. Following from this equality, the above equation becomes equal to the statements of Eq. (5.64). Therefore, the choice of Eq. (5.63) produces the same shifted m_t^{calib} value as the original values of $m_{t,had}^{\mathrm{calib,D\O}}$ and $m_{t,lep}^{\mathrm{calib}}$.

5.3.7 Analysis Setup and Error analysis

For the main measurement, we use the event generator Pythia 6 [32] with the same tuning parameters as DØ used for the lepton+jets measurement of Ref. [36]. Hence, all the generator-related error sources and magnitudes are estimated to be the same as DØ presented. Moreover, we will assume that the F_{corr}-related errors are the same as DØ claimed. This is sensible, as the F_{corr} recalibration of Ref. [3] demonstrates a systematic unaccounted shift in the m_t value. In summary, this indicates that the original DØ error estimates should be applicable for the shifted m_t. On top of this error, we must add the **method error**, related to the presented m_t shifting method. The total method error is found by considering the individual steps of the mass shifting procedure.

According to Eqs. (5.37,5.38), the error of the re-interpreted m_t^{calib} reads as

$$\delta m_t^{\text{calib}} = \delta m_t^{gen} = \frac{\delta\hat{m}_t + p_2 \times \delta K_{\text{JES}}^{\text{Res}}}{p_1} = \frac{\delta\hat{m}_t}{p_1} + \frac{p_2}{p_1} \times \frac{\delta\hat{K}_{\text{JES}}}{q_1}. \tag{5.66}$$

In these error estimates we generally take the fit coefficients as they are, and interpret all fit errors using the variables $\hat{m}_t$ and $\hat{K}_{\text{JES}}$. For the mean values produced by linear fits, the error estimates are found as shown in Appendix C. We note that the term $\delta\hat{m}_t$ must carry the fit error of $\hat{m}_t$, and the total error of the value $\hat{m}_t^{\text{DØ}}$:

$$\delta\hat{m}_t = \delta\hat{m}_t^{\text{Fit}} + \delta\hat{m}_t^{\text{Tot,DØ}} = \delta\hat{m}_t^{\text{Fit}} + \delta\hat{m}_t^{\text{Fit,DØ}} + \delta\hat{m}_t^{\text{DØ}} \tag{5.67}$$

The term $\delta\hat{m}_t^{\text{DØ}}$ can be calculated inverting Eq. (5.66):

$$\delta\hat{m}_t^{\text{DØ}} = -p_2^{\text{DØ}} \times \frac{\delta\hat{K}_{\text{JES}}^{\text{DØ}}}{q_1^{\text{DØ}}} + p_1^{\text{DØ}} \times \delta m_t^{\text{DØ}}. \tag{5.68}$$

Putting Eqs. (5.66,5.67,5.68) together, we find

$$\begin{aligned}\delta m_t^{\text{calib}} = &\frac{1}{p_1}\left(\delta\hat{m}_t^{\text{Fit}} + \delta\hat{m}_t^{\text{Fit,DØ}} + \frac{p_2 \times \delta\hat{K}_{\text{JES}}}{q_1} - \frac{p_2^{\text{DØ}} \times \delta\hat{K}_{\text{JES}}^{\text{DØ}}}{q_1^{\text{DØ}}}\right) \\ &+ \frac{p_1^{\text{DØ}}}{p_1} \times \delta m_t^{\text{DØ}} = \delta m_t^{\text{Stat}} + \delta m_t^{\text{DØ}}.\end{aligned} \tag{5.69}$$

The values of p_1 vary very little between the DØ parametrization and the new one, and any differences between them are mostly statistical fluctuations. Thus, the term $\frac{p_1^{\text{DØ}}}{p_1}$ is set to unity and the last term in Eq. (5.69) becomes equal to the original DØ error. The remaining term is the total statistical error related to the method, consisting of four separate terms. Numerically, the separate error terms in Eq. (5.69) are summed in quadrature.

Care must be taken while combining the central values and errors in the four DØ Run II eras and the electron and muon channels. The mass shift is performed for

all the eight single measurements from the four eras and two channels, considering both the hadronic and leptonic resonances. The combination of the hadronic and the leptonic channels is postponed as much as possible, as this is the most volatile step of the whole procedure. This follows from the notable uncertainty in the parameter α.

The values of α and $\delta\alpha$ are available for the combined electron channel, combined muon channel and the total combination. The α values for the electron and muon channels are used separately, and hence the electron-muon combination must be run after hadron-lepton combination. The differences between the electron and muon αs and the total combination α are used as an estimate of the systematic error in α. The errors related to α determination are complicated, starting from the accuracy of the linear approximation of Eq. (5.41). Also if the calibration error presented in Appendix B is present, an additional error in the α calibration will result. The volatile nature of the α parameter requires a conservative error analysis.

By variational calculus, the total hadron-lepton combination error is:

$$\delta m_{t,comb}^{\text{calib}} = \alpha \times \delta m_{t,had}^{\text{calib}} + (1-\alpha) \times \delta m_{t,lep}^{\text{calib}} + \delta\alpha \times \left(m_{t,had}^{\text{calib}} - m_{t,lep}^{\text{calib}} \right). \tag{5.70}$$

Considering Eq. (5.69), this can be written as

$$\delta m_{t,comb}^{\text{calib}} = \delta m_t^{\text{DØ}} + \alpha \times \delta m_{t,had}^{\text{Stat}} + (1-\alpha) \times \delta m_{t,lep}^{\text{Stat}} + \delta\alpha \times \left(m_{t,had}^{\text{calib}} - m_{t,lep}^{\text{calib}} \right). \tag{5.71}$$

The first term is the transferred (statistical) DØ error, and the next two terms are the statistical method errors of the two resonances. The last term is based on the uncertainty $\delta\alpha$, which does not behave in a similar manner as the statistical error terms. Furthermore, the $\delta\alpha$ term turns out to be the dominating method-related error.

To sum up, the various combination steps are executed in the following order:

1. Combine the (three) methods in all four eras in the electron/muon channels according to the method errors
2. Combine the F_{Corr} sets 1 and 2 according to the method errors
3. Combine the hadronic and the leptonic resonances according to Eq. (5.71)
4. Perform a Run 2 combination for the electron and muon channels separately according to the full errors (DØ + method)
5. Combine the electron and muon channels according to the method errors
6. Combine Pythia 6 and Herwig 7 F_{Corr} sets according to the method errors

In step 3, the errors of Eq. (5.71) are used as weights, excluding the $\delta\alpha$ term. This is motivated by the fact that the error based on $\delta\alpha$ is systematical in its nature, and is not reduced in a combination similarly as the statistical errors. The $\delta\alpha$ term for the combination is found as the weighted average of the corresponding electron and muon terms.

5.4 Measurements

For the measurements, we use three Pythia 6 lepton+jets samples with 10^7 events in each, at the three parameter values $m_t^{gen} = \{172.5, 173.7, 175.0\}$. The mass values are selected so that these are approximately evenly spaced and so that most of the m_t^{calib} values can be interpolated between these. Furthermore, the five values $K_{\text{JES}}^{\text{Res}} = \{1.000, 1.005, 1.010, 1.015, 1.020\}$ are used for the calibrations. In the analyses, we use the DØ Tune A and the kinematic selections presented in Ref. [36]. Moreover, we use the DØ jet cone with FastJet [104] to cluster the particles into jets.

In the analyses, we apply F_{Corr} values derived with two separate simulation software packages: Pythia 6 and Herwig 7 [70, 105]. Furthermore, two separate sets of F_{Corr} parametrizations are used to probe systematical errors. The separating factor between the two parametrizations is a subtle non-trivial choice within the F_{Corr} calibration process. The choice seems to have no practical effect on the calibration quality, and hence it provides a method for assessing systematical F_{Corr} errors. The F_{Corr} results with a less DØ-like choice is denoted with the index 1, and the ones with a more DØ-like choice with the index 2. It turns out that there is little difference between these. Considering the two generators (Pythia 6, Herwig 7) and the two parametrizations, there are in total 4 separate F_{Corr} parameter sets. All these were provided by the author of Ref. [3].

To make the explanation more clear, only figures on the Pythia 6 F_{Corr} set 1 are presented. The sets 1 and 2 are almost identical, so no information is lost here. Between the Pythia 6 and Herwig 7 results there are some differences, with larger shifts observed for Herwig 7. However, the shifts occur in similar patterns, so not much information is lost. For a full reference with all the shifts and channels the reader should refer to [4].

The logistical chain of the analysis is the following:

1. The event information including jets and leptons is saved into a ROOT [69] file using the software handle from Ref. [106].

2. These are further analyzed with the software from Ref. [107].

3. The latter software package applies the F_{Corr} values for the different DØ eras and separates the electron and muon channels.

4. This software tool uses the simulation truth to distinguish the various resonances.

5. It also emulates the reconstruction of a neutrino from the missing transverse momentum and the leptonic W boson resonance.

6. The results are saved mainly into mass histograms, which are turned into resonance position estimators $\hat{m}$ in the next step.

7. Evaluating the values of these estimators involves e.g. fitting and taking mean values of the histograms.

8. Finally, the linear dependence between the $\hat{m}$s and $K_{\mathrm{JES}}^{\mathrm{Res}}$ and m_t^{gen} is fitted using the method of least squares.

9. Based on these fits and the DØ measurements, the shifted top quark mass values and their errors are found.

In this section, the individual steps of the measurement are performed. First, the values of the maximum likelihood $K_{\mathrm{JES}}^{\mathrm{Res}}$ is determined. Then, the hadronic and leptonic top quark mass shifts are derived separately for all estimators, and for the 4 F_{Corr} sets. The results of the five resonance position estimators are combined according to their method errors for both the hadronic and leptonic resonances. In the following step, the combination parameter α is determined. Using the values of α, the results for the hadronic and leptonic channels are combined. Finally, the results based on the four F_{Corr} parameter sets are compared and combined.

5.4.1 Maximum Likelihood $K_{\mathrm{JES}}^{\mathrm{Res}}$

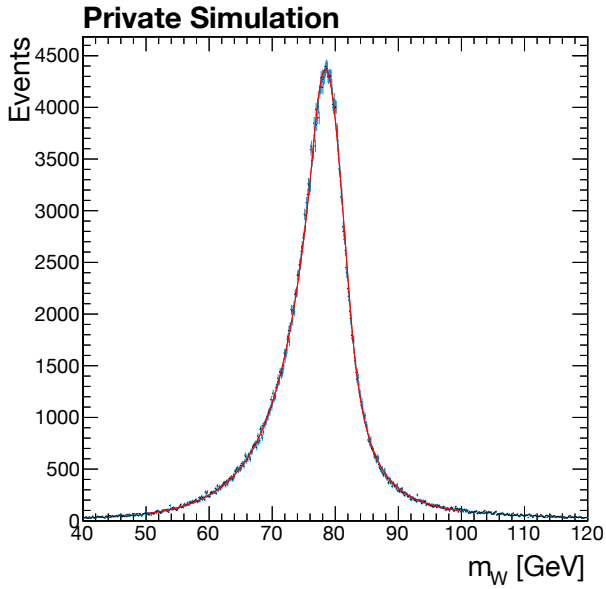

Figure 5.1: Generator-level (no F_{Corr} nor $K_{\mathrm{JES}}^{\mathrm{Res}}$) hadronic W resonances in the muon channel at $m_t^{gen} = 175.0$ GeV.

The maximum likelihood $K_{\mathrm{JES}}^{\mathrm{Res}}$ extraction procedure is based on hadronic W boson resonance histograms. These are evaluated using all the resonance position estima-

tors. In Fig. 5.1 an example distribution is shown. It is not meaningful to provide all of these, as there are:

- 4 runs,
- times 3 m_t^{gen} values,
- times 5 $K_{\mathrm{JES}}^{\mathrm{Res}}$ values,
- times two channels (electron/muon),
- times two F_{Corr} versions (DØ and recalibrated),
- times two sets of F_{Corr} parametrization,
- times two generators (Pythia 6, Herwig 7) in F_{Corr} determination.

Also counting the generator-level resonance histograms at the three m_t^{gen} values, this makes a total of 963 histograms.

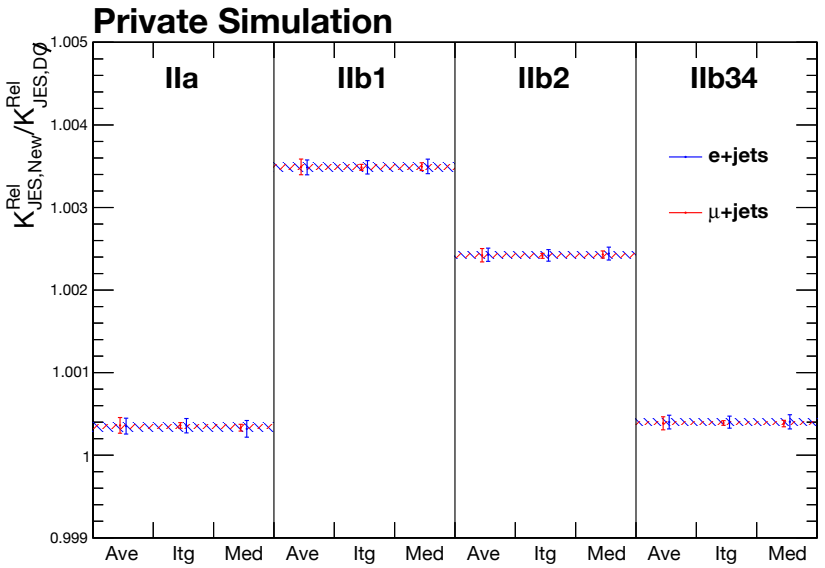

Figure 5.2: Relative maximum likelihood values of $K_{\mathrm{JES}}^{\mathrm{Res}}$ with Pythia 6 (F_{Corr} set 1) with the three best methods in the DØ Run 2 eras: itg (integral), ave (average) and med (median). The method combinations are displayed as hatched boxes.

The exact values of $K_{\mathrm{JES}}^{\mathrm{Res}}$ depend on the resonance position estimator type, and are not informative. This is due to the freedom of choice of the $K_{\mathrm{JES}}^{\mathrm{Res}}$ reference scale, $\mathcal{A}$. However, the relative changes in $K_{\mathrm{JES}}^{\mathrm{Res}}$ provide useful information.

In Fig. 5.2 the fraction of the maximum likelihood $K_{\mathrm{JES}}^{\mathrm{Res}}$ obtained with new F_{Corr} values and that obtained with the DØ ones is shown for Pythia 6 F_{Corr} set 1. Here, a fraction close to unity implies that there has been very little change in F_{Corr}^{lq} and a large deviation from unity implies the opposite. The DØ eras are depicted in order from left to right. The chosen three estimation methods display a good mutual agreement. A range of example cases has shown that especially the ave and med methods exhibit a great precision. Era IIa shows little difference between the original DØ F_{Corr}s and the re-calibrated ones and the same is true for IIb34. In era IIb1 the $K_{\mathrm{JES}}^{\mathrm{Res}}$ scaling effect is the greatest, and in IIb2 almost as great.

5.4.2 Combination of Methods

Examples of the hadronic and leptonic resonances are shown in Fig. 5.3. The resonance positions are evaluated from these using the various resonance position estimators. Fig. 5.4 presents the combination of results from various estimators. Here, a similar notation as in Fig. 5.2 is used. The shifted m_t values are normalized with the original DØ result. Results are displayed separately from the hadronic and leptonic branches.

To summarize, the Run IIa results are almost stagnant, and the hadronic and leptonic results agree. For the other eras a more notable shift is observed and there is typically a 0.5 GeV shift between the hadronic and leptonic branches. The results with the various resonance position estimators are very stable.

Table 5.1 shows the shift results for the different F_{Corr} sets for Pythia 6 and Herwig 7 in all of Run II. It can be confirmed that similar behavior is observed to Herwig 7 and Pythia 6 F_{Corr} set 2, as was presented for Pythia 6 F_{Corr} set 1 in Fig. 5.4. It is also observed that the shifted m_t values for Herwig 7 tend to be 0.5 – 1 GeV lower than the ones for Pythia 6. This difference is discussed more widely later.

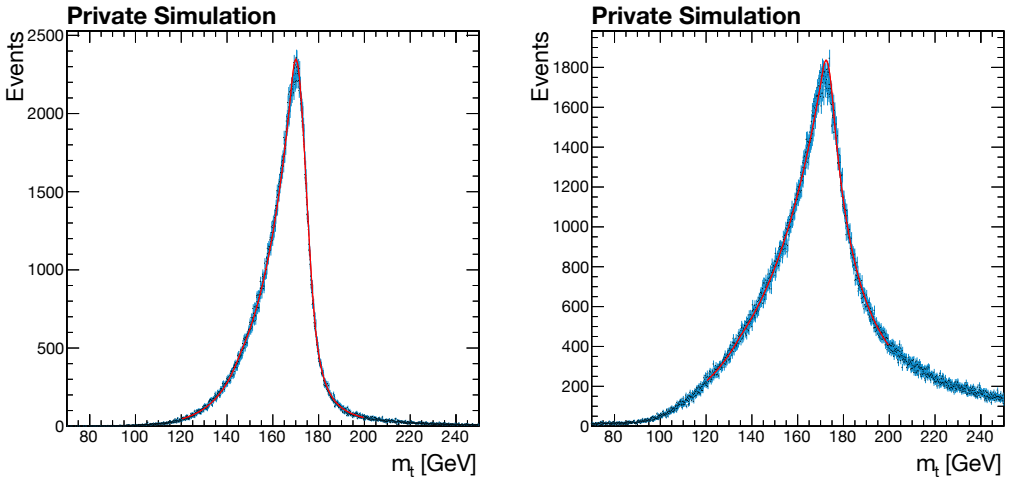

Figure 5.3: Generator-level (no F_{Corr} nor $K_{\text{JES}}^{\text{Res}}$) for the hadronic (left) and leptonic (right) top resonances in the muon channel at $m_t^{gen} = 175.0$ GeV.

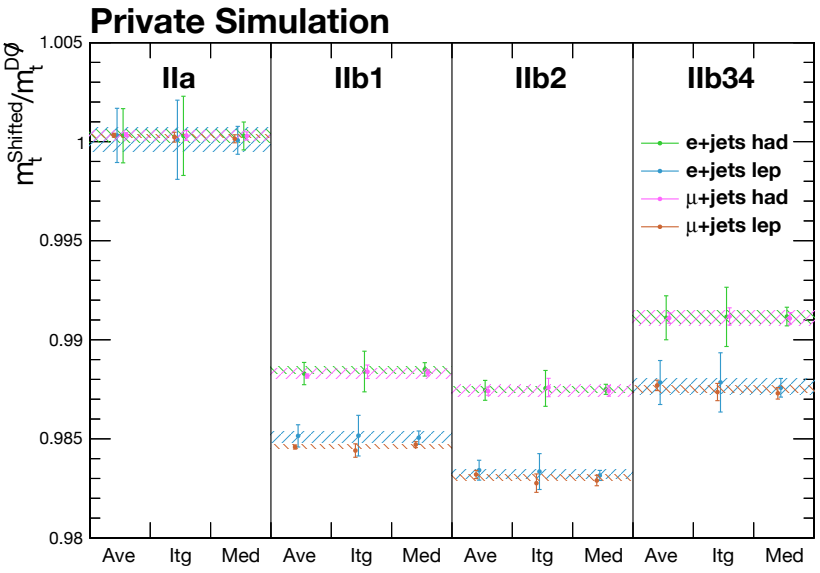

Figure 5.4: The original and shifted m_t values with Pythia 6 (F_{Corr} set 1) with the three best methods in the DØ Run II eras: itg (integral), ave (average) and med (median). The method combinations are displayed as hatched boxes.

Table 5.1: Shifted DØ Run II lepton+jets m_t values with their statistical method errors in the hadronic and leptonic channels. Combination over the methods itg (integral), ave (average) and med (median).

Channel	DØ Run IIa	DØ Run IIb1	DØ Run IIb2	DØ Run IIb34	DØ Run II
Had Pythia 6 1 e	177.70 ± 0.06	172.53 ± 0.03	171.27 ± 0.03	175.05 ± 0.06	173.91 ± 0.03
Had Pythia 6 2 e	177.73 ± 0.06	172.57 ± 0.03	171.32 ± 0.03	175.08 ± 0.06	173.95 ± 0.03
Had Herwig 7 1 e	176.83 ± 0.05	171.53 ± 0.02	169.96 ± 0.02	173.80 ± 0.04	172.71 ± 0.02
Had Herwig 7 2 e	176.82 ± 0.05	171.54 ± 0.02	169.97 ± 0.02	173.80 ± 0.04	172.72 ± 0.02
Lep Pythia 6 1 e	177.67 ± 0.11	171.94 ± 0.05	170.53 ± 0.04	174.43 ± 0.07	173.31 ± 0.04
Lep Pythia 6 2 e	177.72 ± 0.11	172.02 ± 0.05	170.62 ± 0.04	174.50 ± 0.07	173.39 ± 0.04
Lep Herwig 7 1 e	176.52 ± 0.09	170.66 ± 0.04	168.90 ± 0.03	172.87 ± 0.05	171.81 ± 0.03
Lep Herwig 7 2 e	176.52 ± 0.09	170.68 ± 0.04	168.91 ± 0.03	172.88 ± 0.05	171.82 ± 0.03
Had Pythia 6 1 μ	172.72 ± 0.04	171.73 ± 0.04	172.27 ± 0.05	173.24 ± 0.07	172.70 ± 0.04
Had Pythia 6 2 μ	172.75 ± 0.04	171.77 ± 0.04	172.31 ± 0.05	173.27 ± 0.07	172.74 ± 0.04
Had Herwig 7 1 μ	171.86 ± 0.04	170.72 ± 0.04	170.95 ± 0.04	171.98 ± 0.05	171.49 ± 0.03
Had Herwig 7 2 μ	171.85 ± 0.04	170.73 ± 0.04	170.95 ± 0.04	171.98 ± 0.05	171.49 ± 0.03
Lep Pythia 6 1 μ	172.72 ± 0.01	171.09 ± 0.01	171.50 ± 0.03	172.60 ± 0.03	172.08 ± 0.02
Lep Pythia 6 2 μ	172.76 ± 0.01	171.17 ± 0.01	171.59 ± 0.03	172.67 ± 0.03	172.15 ± 0.02
Lep Herwig 7 1 μ	171.53 ± 0.01	169.77 ± 0.01	169.75 ± 0.02	170.98 ± 0.02	170.50 ± 0.01
Lep Herwig 7 2 μ	171.52 ± 0.01	169.79 ± 0.01	169.76 ± 0.02	170.99 ± 0.02	170.51 ± 0.01

The differences between the F_{Corr} sets consist of both statistical and systematical errors. In the final m_t combination calculation, the two F_{Corr} sets are combined in a simple statistical manner. This reduces the systematic F_{Corr} error in an unfavorable manner. To counter this issue, separate Run II combination values from Table 5.1 are used to estimate the systematic uncertainty component. For the hadronic branches, the separation between the F_{Corr} sets is maximally 0.04 GeV for Pythia 6 and 0.01 GeV for Herwig 7. In the leptonic branch the corresponding values are 0.08 GeV for Pythia 6 and 0.01 GeV for Herwig 7. To make a conservative estimate of the systematic F_{Corr} error component, the maximal difference of 0.08 GeV is added to the final method error.

5.4.3 The Combination Parameter α

For α extraction, we compare the likelihood slopes found in the DØ paper [36], and those produced using the resonance position estimators. The calibration process of α is very sensitive, and hence the effect of any errors is magnified. For the lowest quality resonance position estimators the results are thus very poor. To obtain a reliable estimate of the $K_{\text{JES}} - m_t^{\text{calib}}$ slope, we use the two estimators with the best quality: ave and med. The final α value is considered as an average of these two. Moreover, we average over all the four F_{Corr} sets to find generally applicable values for α. The preference is to use common central values for α in all the four measurements, and explain the variations observed here as a larger systematical α error.

Table 5.2: Measured values of α using the *average* resonance estimator.

Measurement	α_e	α_μ	α_{comb}	$f_e \times \alpha_e + (1 - f_e) \times \alpha_\mu$	$\delta\alpha_{comb}$
Pythia 6 1 New	0.7532	0.5882	0.7017	0.6770	0.0247
Pythia 6 2 New	0.7521	0.5861	0.7003	0.6755	0.0248
Herwig 7 1 New	0.7460	0.5562	0.6851	0.6584	0.0267
Herwig 7 2 New	0.7460	0.5560	0.6850	0.6583	0.0268
Pythia 6 1 DØ	0.7561	0.5951	0.7063	0.6818	0.0246
Pythia 6 2 DØ	0.7561	0.5952	0.7064	0.6818	0.0246
Herwig 7 1 DØ	0.7558	0.5946	0.7060	0.6814	0.0246
Herwig 7 2 DØ	0.7558	0.5946	0.7060	0.6814	0.0246

Table 5.3: Measured values of α using the *median* resonance estimator.

Measurement	α_e	α_μ	α_{comb}	$f_e \times \alpha_e + (1 - f_e) \times \alpha_\mu$	$\delta\alpha_{comb}$
Pythia 6 1 New	0.8131	0.6504	0.7612	0.7380	0.02317
Pythia 6 2 New	0.8120	0.6486	0.7598	0.7366	0.02318
Herwig 7 1 New	0.8058	0.6292	0.7490	0.7243	0.02469
Herwig 7 2 New	0.8054	0.6293	0.7488	0.7241	0.02471
Pythia 6 1 DØ	0.8209	0.6492	0.7641	0.7417	0.02245
Pythia 6 2 DØ	0.8208	0.6493	0.7641	0.7417	0.02242
Herwig 7 1 DØ	0.8211	0.6483	0.7638	0.7413	0.02246
Herwig 7 2 DØ	0.8211	0.6483	0.7638	0.7413	0.02246

Table 5.4: Combined α values and error terms.

Measurement	$\alpha_e \pm \delta\alpha_e$	$\alpha_\mu \pm \delta\alpha_\mu$	$\delta\alpha_{comb}$	$\delta\alpha_e^{tot}$	$\delta\alpha_\mu^{tot}$
ave	0.7526 ± 0.0044	0.5833 ± 0.0171	0.0252	0.0255	0.0304
med	0.8150 ± 0.0069	0.6441 ± 0.0092	0.0232	0.0242	0.0250
combination	0.7838 ± 0.0327	0.6137 ± 0.0341	0.0242	0.0408	0.0418

The results for Pythia 6 and Herwig 7 with the average estimator are given in Table 5.2 and the ones with the median estimator in Table 5.3. Here, the DØ results act as a stabilizing agent, as the values of these are almost constant. The small numerical fluctuations are explained by the fact that for each F_{Corr} set, also the interpretation of jet p_T within the F_{Corr}s is re-parametrized. The statistical combination of the electron and muon αs is performed using $f_e = 0.538339$, obtained from the statistical errors of DØ. The two combined αs differ by a factor $\delta\alpha_{comb}$, which is interpreted as a systematical error source.

The merged results with their sample variances are given in Table 5.4. For further error estimation, these plain sample variances are used. They do not suppress systematical differences between the F_{Corr} sets, in contrast to the error of the mean.

The most notable systematical difference can be noted between the ave and med estimators, but there also seems to exist some systematical difference e.g. between Pythia 6 and Herwig 7.

The total errors are found by combining the sample variances with the average $\delta\alpha_{comb}$ terms in quadrature. To make a conservative estimation of the total error, these ($\delta\alpha_e^{tot}$ and $\delta\alpha_\mu^{tot}$) are finally multiplied by the factor of 2. As a result, the electron and muon channels yield the approximate values 0.78 ± 0.08 and 0.61 ± 0.08. The difference between these can potentially be explained by the differences in phase spaces, as electrons are considered in $|\eta| < 1.1$ and muons in $|\eta| < 2.0$ [36].

5.4.4 Combination of the Hadronic and Leptonic Results

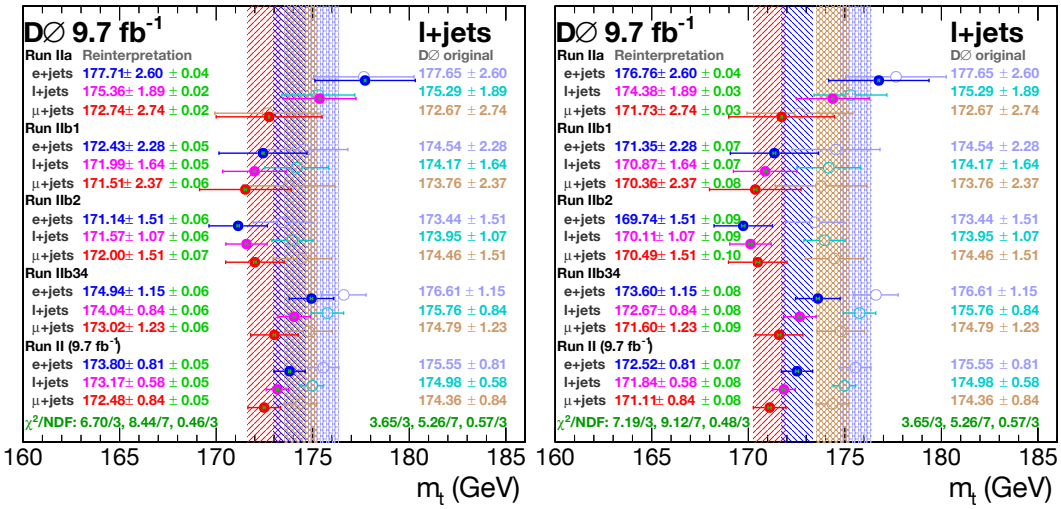

Figure 5.5: Shifted (hadronic + leptonic top quark combination) top quark mass values for the Pythia 6 (left) and Herwig 7 (right) F_{Corr} set 1 + 2 combinations. Shifted results: closed circles. Original DØ results: open circles. The background color bars display the Run II combination values with $\pm\sigma$ errors.

The combination of the hadronic and leptonic channels is displayed in Fig. 5.5 for both Pythia 6 and Herwig 7 F_{Corr}s. The shifted results are displayed in the left side of the figures while the DØ ones are placed on the right side. The method errors are generally small, and displayed in a green font.

The Herwig 7 F_{Corr} calibration displays a notably greater shift compared to Pythia 6, as was already noted in Table 5.1. Notably, also the Run IIa results are somewhat shifted. This could indicate the presence of a significant systematic uncertainty related to the chosen simulation methods. The differences also present two alternative interpretations for this analysis:

- A conservative interpretation (Pythia 6 F_{Corr} only).

- An extension for the estimation of systematic uncertainties at DØ (combination between the Pythia 6 and Herwig 7 F_{Corr} sets).

Both the separate and combined results based on Pythia 6 and Herwig 7 F_{Corr} values are given for this reason.

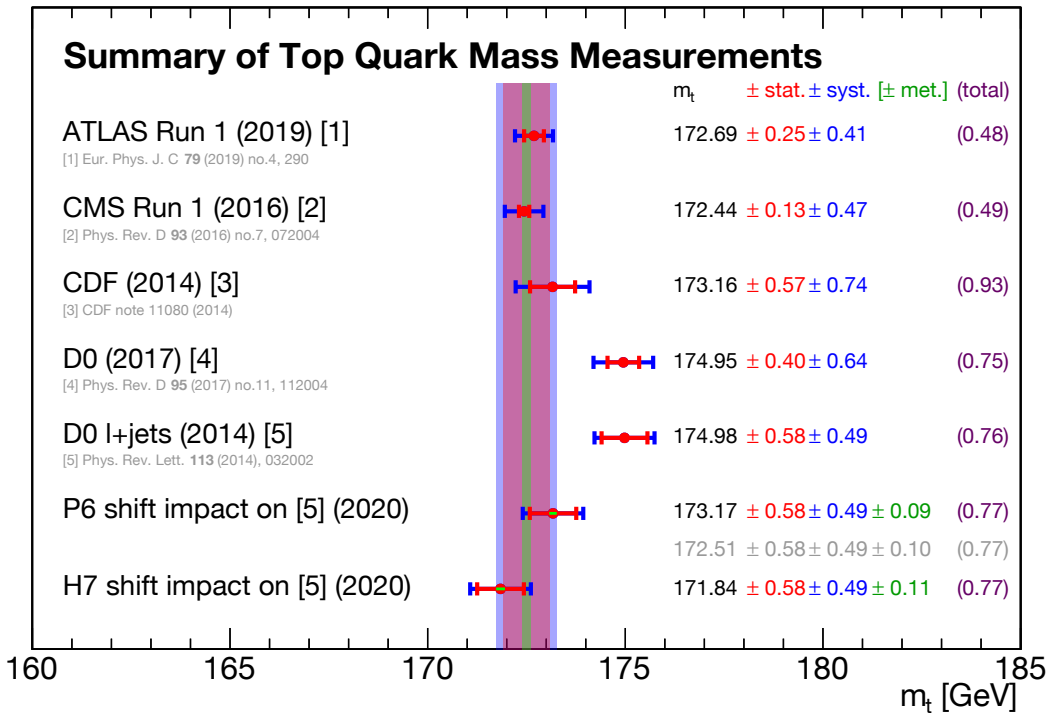

Figure 5.6: A comparison of the shifted top quark mass values with recent measurements.

In Fig. 5.6 the Pythia 6 and Herwig 7 results are compared to recent notable top quark mass measurements. For DØ, the total average and lepton+jets measurement are displayed separately to underline the fact that the lepton+jets measurement dominates the average. The combination of the method error and systematic F_{Corr} error is given in green font. The Pythia 6 and Herwig 7 results combined according to their method errors are shown in the background color bars and gray text. A combination utilizing the method errors gives a slight preference towards the Pythia 6 results.

5.4.5 Discussion

We have derived values for the top quark mass shift necessitated by the re-calibration of the DØ F_{Corr} values. This was done with a variety of resonance position esti-

mators, F_{Corr} values derived using both Pythia 6 and Herwig 7, and two separate derivations for the F_{Corr} values. The small differences between the F_{Corr} sets demonstrate the stability of the method, and of small systematic errors in F_{Corr} determination.

A notable difference between the Pythia 6 and Herwig 7 F_{Corr} based results was observed. This needs to be considered as a significant systematical error, as the two generators employ different heuristics for the collision dynamics and jet production. This kind of error estimation is lacking from the DØ publication. The two results are not on a completely equal footing, as there is no DØ tune available for Herwig 7. Thus, the Herwig 7-based values might be less reliable than the Pythia 6 ones, and one could give more weight to Pythia 6 than Herwig 7 in the combination. Even if a higher weight was given for the Pythia 6 results, the observed total shift remains around $-2\,\text{GeV}$.

The results obtained with the five different resonance position estimators agree remarkably well. However, in this analysis only the following three were used: the *average*, the *median* and the *integral* based estimators. These are the most stable estimators, in contrast to the *maximum* and the *fit* estimators. The *maximum* method is relatively unstable, also presenting high error estimates. Moreover, the *fit* method tends to exhibit instabilities in the leptonic branch, caused by the difficulties of automating a robust fit. Including or not including the *maximum* and *fit* methods has little impact on the full analysis, as they receive lower weights according to their higher errors in the combination of methods. Here, these methods were excluded most of all to promote clarity. The analysis and full control plots for all the five estimators can be found in Ref. [4].

In Fig. 5.6 the final results were displayed. The DØ average differs significantly from that given by the CDF, ATLAS and CMS collaborations. In contrast, the shifted lepton+jets result is in a good agreement with the other collaborations. Shifting the DØ lepton+jets result shifts the whole DØ average.

It should be underlined that a downward shift of m_t is a direct consequence of the F_{Corr} recalibration. The phenomenological proof of Section 5.2 shows that the DØ measurements should have produced a lower m_t value. From the magnitude of the change in F_{Corr} in the recalibration, we knew that this error is in the order of GeVs. Thus, the motivation of this study was in quantitatively propagating the inevitable consequences of the F_{Corr} recalibration to the m_t measurement.

Using the methods presented in this text, the resulting top quark mass shift should be reproducible both phenomenologically, and with numerical evaluation. For the sake of openness, the complete source code of the measurement is provided in Refs. [106, 107]. The source code includes all the numerical parameter values. The found numerical values are completely based on the series of linear fits made on top of the resonance histograms, produced using simulations.

5.5 Summary

This study was motivated by the observation that the DØ calibration of F_{Corr} JECs may include significant bias. A suggested re-calibration of the corrections has a great impact on the top quark mass measurements. To demonstrate a shift caused by the change in F_{Corr} values, an intricate method was designed.

First, the steps between the quark and the jet level in the DØ measurements were dissected analytically. Using the obtained results, it was demonstrated phenomenologically how the changes in F_{Corr} are conveyed to the top quark mass values. Then, a numerical simulation-driven method for estimating the shift in the m_t values was designed. Finally, this method was applied, producing estimates for the m_t value that DØ should have obtained considering the F_{Corr} recalibration.

In the evaluation of the top quark mass shift, two slightly differing F_{Corr} parameter sets were utilized. Moreover, the shift results were derived separately with F_{Corr} values based on Pythia 6 and on Herwig 7. The resulting shifted m_t values differed significantly, but within error bars. A simple combination of the Pythia 6 and Herwig 7 results yielded a shifted top quark mass value of 172.51 GeV. The result of Fig. 5.6 compares the impact of the DØ lepton+jets m_t shift to other important m_t results. Here, it is demonstrated that the F_{Corr} recalibration seems to lead to a significantly better agreement with other measurements.

The fact that the m_t value shifted by the F_{Corr} recalibration agrees with other collaborations is intriguing. *A priori*, there are two peculiarities in the DØ measurement: the differences between the Run IIa and Run IIb F_{Corr} values and the high extracted m_t value with a relatively small error. We have demonstrated that these features could share the same origin. Future top quark mass world combinations should consider these results profoundly.

6 Top Quark Mass Measurement at the CMS

In this chapter the analysis strategy adapted for the $t\bar{t}$ lepton+jets m_t measurement on the 2017–2018 legacy datasets is presented. Similar methodology is utilized here as in the recently released preliminary m_t measurement on the legacy 2016 dataset [22]. As there are considerable similarities and partially shared code between this and the legacy 2016 analysis, the author of this thesis has also contributed into the development and bug tracking of the latter one.

There are also important differences between the analyses. For instance, the implementation of the likelihood approach used for determining the value of m_t in this work is completely new. The results presented here are endorsed by the CMS collaboration to be presented in this thesis.

All $t\bar{t}$ lepton+jets top quark mass measurements at the CMS are built on top of the same basic principles as the DØ measurement. The measurement is direct, implying that a kinematic reconstruction of the top quark is made based on its decay products. **All direct measurements rely heavily on the comparison between simulations and data, which is also their greatest weakness.**

As the value of m_t is discovered through a comparison with simulation, it is referred to as the Monte Carlo (MC) top quark mass (m_t^{MC}). The value of m_t^{MC} can differ by up to $\mathcal{O}(1\,\text{GeV})$ from the theoretical pole mass [38] and it is consistent with the MSR (1 GeV) scheme mass at a $\mathcal{O}(0.5\,\text{GeV})$ uncertainty [39]. In the interpretation of these results, the differences between m_t definitions should be considered.

6.1 Event Selection

The analysis selects semileptonic $t\bar{t}$ events, which contain a single lepton (electron or muon) and four jets in the final state. This allows triggering with single lepton triggers. A full listing of the simulated samples used in the analysis is provided in Appendix A. The most important backgrounds are the dileptonic $t\bar{t}$ and single top

tW channel processes.

6.1.1 Triggers and Lepton Selection

Table 6.1 summarizes the trigger paths used in this analysis: these are the single (charged) lepton triggers relevant in each analysis channel. These triggers have the lowest available p_T thresholds in each data-taking period, while still being **unprescaled**. That is, all events passing the triggers in data have been stored. The electron triggers in 2017 and 2018 are essentially the same, except that on the technical level an additional L1 trigger condition is required in 2017. The muon triggers in 2017 and 2018 have different HLT p_T thresholds.

Table 6.1: The listing of trigger paths used in this m_t analysis: lowest unprescaled single electron/muon trigger in the 2017 and 2018 trigger menus. Reco p_T refers to the AK4PFchs jet p_T, which is given in contrast to the HLT p_T estimate. The 2017 electron trigger is used in association to the additional L1 condition L1SingleEGOr.

Channel	Trigger	HLT p_T threshold [GeV]	Reco p_T cut [GeV]
Electron 2017	HLT_Ele32_WPTight_Gsf_L1DoubleEG	32	35
Electron 2018	HLT_Ele32_WPTight_Gsf	32	35
Muon 2017	HLT_IsoMu27	27	29
Muon 2018	HLT_IsoMu24	24	26

After the trigger is applied, it is required that there is exactly one signal lepton. Events with any additional leptons are vetoed. The signal lepton type (electron or muon) must match the trigger, whereas both electrons and muons are checked for veto in both channels. The lepton cuts are designed consistently, so that the signal lepton selection is a subset of the veto lepton selection. The lepton p_T cuts are designed to be on the plateau region of the trigger turn-on curves, and the values of these are presented in the rightmost column of Table 6.1. All veto leptons are required to have $p_T > 15\,\text{GeV}$.

Further features, cuts, required scale factors and corrections presented for leptons are summarized in Table 6.2. The objective of the scale factors is to correct for the differences in efficiencies and acceptances in data and simulation. Muons require scale factors for isolation, identification (ID) and trigger [108]. On the other hand, electrons need scale factors for reconstruction, ID and trigger: the isolation condition is included in the ID [64]. In muon reconstruction there is less uncertainty and hence no corresponding scale factors are required. The electron trigger scale factors were not centrally provided, and they were derived privately.

Furthermore, whereas all electron energy corrections are included within the PF object, muons require separate Rochester corrections [109]. On the other hand for electrons the impact parameter cuts are not considered in the Tight ID, so they are performed separately. Within the barrel the recommended cuts include $d_0 < 0.05\,\text{cm}$ and $d_z < 0.1\,\text{cm}$ and in the endcap, $d_0 < 0.1\,\text{cm}$ and $d_z < 0.2\,\text{cm}$. Moreover, events

Table 6.2: Lepton selection definitions, excluding the lower limit of p_T, which is $p_T >$ 15 GeV for veto leptons and for signal leptons given in Table 6.1. Settings and scale factors required by lepton identification (ID), isolation, trigger and reconstruction (Reco) are presented.

Lepton category	Signal electron	Veto electron	Signal Muon	Veto Muon
Max. $\|\eta\|$	2.5	2.5	2.4	2.4
ID	Tight	Loose	Tight	Loose
ID version	cutbasedElectronID-Fall17-94X-V2		CutBased	CutBased
Isolation	Within ID	Within ID	PFIsoTight	PFIsoLoose
Extra η cuts	EE/EB transition	-	-	-
Impact Parameter Cut	Applied manually	-	Within ID	-
Energy corrections	pat::Electron	pat::Electron	Rochester	Rochester
Trigger scale factor	Yes	-	Yes	-
Reco scale factor	Yes	-	-	-
ID scale factor	Yes	-	Yes	-
Isolation scale factor	Within ID scale factor	-	Yes	-

where the signal electron supercluster η (η_{sc}) is in the ECAL barrel-endcap transition region ($1.4442 < |\eta_{sc}| <= 1.566$) are cut out from the electron dataset.

As a novel strategy introduced in this analysis, events with high-energy signal leptons are vetoed out from the analysis. Charged lepton reconstruction works best at a relatively low p_T, while high-energy leptons are more suitable for exotic studies. As this is a precision measurement, it is best to focus on the regime where leptons are measured most accurately. Hence events with signal leptons above 300 GeV are vetoed. The lepton p_T spectrum is steeply falling, and thus this cut affects under 1% of the available statistics.

6.1.2 Jet Selection

In the semileptonic $t\bar{t}$ event hypothesis a b quark jet (b jet) originates from each of the two decaying top quarks as depicted in Fig. 3.7. Furthermore, it is assumed that the W boson from the hadronically decaying top produces exactly two light quark jets (l jets). To summarize, the baseline condition for jets is the existence of minimally four signal jets. Additional jets are often produced by ISR, FSR and pileup, adding challenges for the selection of the correct four jets.

For the **signal jet collection** strict baseline cuts are imposed: only AK4PFchs jets with $p_T > 30$ GeV and $|\eta| < 2.5$ are considered. Any jets overlapping with the (only) charged lepton in the event are removed from this jet collection. Taking into account higher $|\eta|$ values than 2.5 is not profitable for two reasons:

1. At $|\eta| > 2.5$, b-tagging becomes challenging as this is outside the tracker acceptance region

2. Considering $|\eta| > 2.5$ lowers the success rate in choosing the correct W jets as

this region contains a higher fraction of ISR, pileup and UE jets than $|\eta| \leq 2.5$

The $p_T > 30\,\text{GeV}$ cut is motivated both by experimental factors and the historical reference of this analysis [19, 20, 22]. Using a significantly lower p_T threshold would be harmful, as descending from 30 GeV the amount of pileup jets quickly increases. Moreover, the jet energy resolutions relative to p_T increase and the JEC precisions decrease at low p_T. The direct JEC calibration from Z+jet events is applicable above 30 GeV. Furthermore, the amount of background events especially from the Multijet topology would significantly increase with a lower jet p_T threshold. On the other hand, increasing the selection significantly from 30 GeV would notably reduce the statistics without introducing significant benefits. The magnitude $p_T \approx 30\,\text{GeV}$ is currently a sweet spot where precision and good statistics meet.

To strengthen the signal-to-background ratio, it is imposed that exactly two of the signal jets must be b jets. These two jets are interpreted as top quark decay products. The b jet status is checked using the DeepJet [110] Medium Working Point (WP). DeepJet is a b jet tagging tool, based on Deep Neural Networks.

The medium WP carries a 1% mistag rate for uds quark and gluon jets [92, 111–113]. The loose and tight WPs conventionally experience respective mistag rates of 10% and 0.1% (for udsg jets). The mistag rates for c quark jets are notably higher, e.g. approximately 16% for the medium WP. Jets with the b tag discriminant value above the medium WP are interpreted as b jets and others as l jets. The two signal l jets with the highest p_T values are interpreted as the W jets. The two b jets and the two leading l jets are labelled **analysis jets**.

Only the 8 signal jets with the highest p_T values are checked for b-tagging status. This is a significant change to the previous analyses – e.g. the legacy 2016 analysis [22], where only the 4 leading jets are considered. Also the scenario of 6 leading jets has been tested, but the difference to 8 leading jets was insignificant. The choice of using the 8 signal jets leading in p_T for b-tagging is more widely handled in Subsection 6.1.3. Briefly, the main arguments are the following:

- With a mistag rate of 1% it is better to limit the number of jets considered for b-tagging from above: b jets are more likely to have a high p_T value, whereas the lower p_T jets are more likely mistagged
- Each jet considered for b-tagging accumulates the b tag scale factor weight of each event, increasing b-tagging uncertainties
- A significant fraction of b jets in $\mathrm{t\bar{t}}$ events are found outside the four leading jets

Events in which one of the four analysis jets does not fulfill the jet ID [114] are vetoed. The jet ID is a collection of quality criteria for jets, making sure that e.g. leptons are not handled as jets. This cut affects a tiny fraction of all events. The identification procedure is aggregated with the jet veto maps. These have been developed for precision studies by the author of this thesis, as presented at the end

of chapter 4. The jet vetos reduce the amount of data notably: by around 15% both in 2017 and 2018. The gain is notable, nonetheless, as e.g. the problematic zones HEP17 in 2017 and HEM15-16 in 2018 are removed. The veto maps are applied symmetrically in data and in simulation, and hence their use does not involve systematic uncertainties. It should be noted that the HEP17 and HEM15-16 removals cause asymmetries in the η/ϕ-dependent jet figures, which are presented later.

Continuing with the same new strategy as with leptons, an extra veto is imposed for events including jets above 1000 GeV. This is the region where e.g. b-tagging is not functioning optimally. Moreover, the hypothesis of four analysis jets becomes less efficient, as the top quarks are boosted and some of the four jets can become merged. The removed statistics are negligible, less than 1%. Finally, as this is an analysis involved with a prompt neutrino producing non-zero values of $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$, the $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ filters [90, 115] are used to veto events. The effect of this veto is very small, also less than 1%.

In summary, the jet selection is:

- Consider jets with $p_T > 30\,\mathrm{GeV}$ and $|\eta| < 2.5$
- Pick maximally 8 jets with the highest p_T values for the following steps, motivated by b-tagging
- There must be at least 4 jets
- There must be exactly two b jets at the DeepJet medium WP
- These two are interpreted as originating directly from the top quark decays
- The two l jets (at DeepJet medium WP) with the highest p_T values are interpreted as $\mathrm{W} \to qq'$ jets
- Veto events with jets above 1000 GeV
- Veto events where one of the four analysis jets do not pass the jet ID or are located in the jet veto zones.

6.1.3 Commentary on B Jet Selection

In addition to lepton triggers, reconstruction, ID and isolation, also jets for which b-tagging is applied require scale factors. The scenario with jets is more complicated than the one with leptons, as there can be up to 8 b-tagged jets in contrast to the single signal lepton. The b-tagging scale factor for fixed WPs can be expressed as [92]:

$$w = \frac{\mathrm{P(Data)}}{\mathrm{P(MC)}} = \left(\prod_{i=\text{tagged}} \mathrm{SF}_i \right) \times \left(\prod_{j=\text{not tagged}} \frac{1 - \mathrm{SF}_j \epsilon_j}{1 - \epsilon_j} \right). \tag{6.1}$$

Here, SF is the relevant scale factor and ϵ refers to the b-tagging efficiency observed in the simulated samples, measured as a function of p_T and η. The efficiencies are measured with the full selection of this analysis, excluding the cut on the number of b jets. The efficiencies and the scale factors are given separately for b jets, c-jets and light quark jets according to the simulation truth. With maximally 8 jets considered for b-tagging, Eq. (6.1) carries exactly 2 terms in the tagged category and 2 to 6 terms in the not tagged category for selected events. In typical signal events, it is extremely rare (in the order of a few permilles) for more than 8 jets at $p_T > 30\,\text{GeV}$ to be present in the LHC Run 2 conditions.

The selection of the b-tagging WP is well optimized, and the usage of the loose WP is ruled out, since this would notably increase the background from mistagged b jets. In contrast, the distinction between the tight and the medium WPs is not as strict. The tagging efficiency of the medium WP is around 74–80%, and that of the tight WP around 56-62% [112, 113]. For the selection of two b jets in the signal events, this indicates a selection efficiency of 55-64% or 31-38%, respectively for the medium and tight WPs.

Thus, swapping from the medium to the tight WP is expected to reduce the signal event yield roughly by 40%. The statistical uncertainties in the current measurements are small [22], but in the currently used profile likelihood methods statistics are also used to constrain systematics, and a 40% decrease in the event yield would have significant consequences. On the other hand, this would lead to a significant reduction of the backgrounds that are not b-enriched. However, the events in the main backgrounds $\mathrm{t\bar{t}}$ dileptonic and single top are enriched in b jets, diminishing the value of the smaller mistag rate.

While using the medium WP, the number of jets considered for b-tagging affects the theoretical count of background events passing the b jet selection. Considering the mistag rates and all jet permutations, the total rate of finding exactly two (mistagged) b jets for the 1% mistag rate is around 1 in 1000. In contrast, for the tight WP and the 0.1% mistag rate the corresponding number is 1 in 100,000, which can be safely ignored. If there were always 8 signal jets available, there would be 80% more potential background events passing in the 8-jet scheme vs. if only 6 jets were used for b-tagging. Compared to the original 4-jet scheme used in the legacy 2016 analysis, the potential increase is 350%.

Experimentally, these pessimistic numbers do not come to pass. The spectrum of the number of signal jets falls exponentially, as seen in Fig. 6.1. Here, the number of signal jets ($p_T > 30\,\text{GeV}$ and $|\eta| < 2.5$) is presented. Therefore, the difference between using 6 and 8 signal jets for b-tagging is very small, and a slight preference towards 8 jets is observed. The 8-jet scheme gives a slightly better signal-to-background ratio than the 6-jet scheme, indicating that the dynamics of the full selection process are more complex than simply indicated mistag rates.

To conclude, it is necessary to underline that this *status quo* of b-tagging can notably change in Run 3 and/or HL-LHC conditions. If the count of pileup or ISR jets increases significantly, high jet multiplicities can become more common,

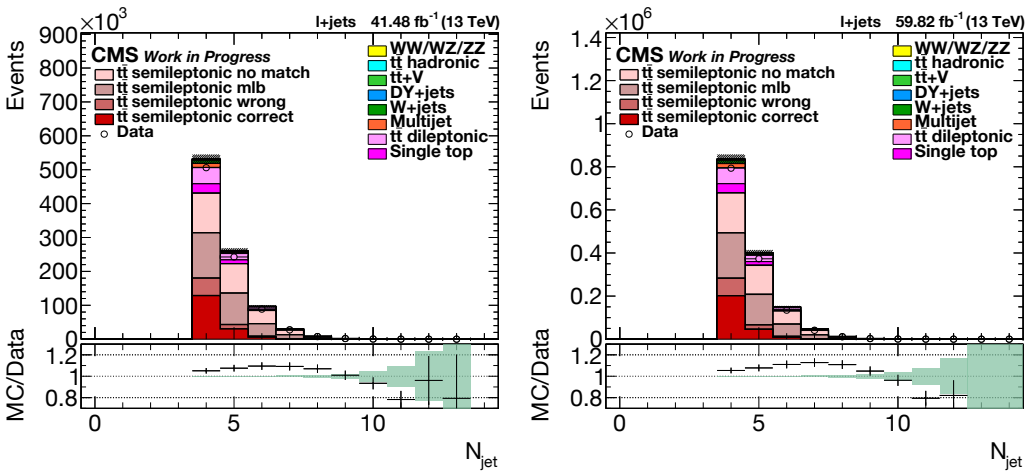

Figure 6.1: Number of jets after the baseline selection, $mu17 + e17$ (left), $mu18 + e18$ (right). Simulation to data ratio error bars consist of statistical and scaling uncertainty in simulation, whereas statistical uncertainty in data is displayed as a green band.

potentially changing the above reasoning. The optimal amount of jets for b-tagging is linked to the current detector conditions.

6.1.4 Commentary on Light Quark Jet Selection

Even if the b-tagging efficiency was close to 100%, a semileptonic event hypothesis could fail in numerous other ways. First of all, the W boson decay can produce a b jet, confusing the b jet selection condition. Additionally in around 50% of the cases the W boson decays producing a charm jet, which is by any b-tagging algorithm more easily mixed up with b jets than the other light quark jet types.

Also the "W boson to 2 jets" decay hypothesis has notable caveats. The W boson can decay either into a single boosted/wide jet, or into two asymmetric jets, one of which is left under the p_T threshold of 30 GeV. There is also a notable proportion of events[1] where the correct W jet pair is found outside the two leading light quark signal jets.

In summary, b jet selection is thoroughly optimized using DeepJet, whereas the p_T-ordered W$\rightarrow qq'$ jet selection has many weaknesses. Numerous alternative methods for this selection have been tested out, including the use of quark-gluon tagging and the DeepJet discriminant values. The best results in terms of increased statistics and the signal-background ratio are found by first focusing on the 3 or 4 leading light quark jets, and then selecting the best pairing among them.

[1] Up to around 50% of the cases with W jets as the leading l jets.

By a good margin, the most promising method has been the selection of the l jet pair closest to the W mass shell (within the four leading light quark jets). The related increase in signal statistics is found to be roughly 30–40%. Without further cuts, this increases the fraction of correct permutations from 17% to 22%, but in $P_{\text{gof}} \geq 0.2$ the fraction decreases from around 47% to 42%. The goodness-of-fit probability cut $P_{\text{gof}} \geq 0.2$ provides the best signal region found in the kinematic fit, defined later in Eq. (6.5). These numbers underline that a mass shell selection does not function optimally with the kinematic fit. The behaviour follows from using a similar mass shell condition in the pre-kinematic fit jet selection as is found in the kinematic fit. This reduces the value of mass shell condition in the fit, and more background events are hence selected e.g. in the $P_{\text{gof}} \geq 0.2$ regime.

If the W mass shell selection method is adapted in future, further sanity checks are necessary. This method can cause trouble with the use of the m_W resonance in the *in-situ* tuning of the JEC and FSR constraints. A minimal test would include a comparison between the best and second best m_W^{reco} jet pair candidates in data and simulation. If there is only little difference between competing m_W^{reco} candidates, the selection method is unlikely to be robust.

A more stable W-jet selection is found by choosing the two light quark jets closest in ΔR (within the four leading light quark jets). The increase in the event yield is slightly less drastic than above: 20%. Without further cuts, this increases the fraction of correct permutations from 17% to 20%, and in the best goodness-of-fit region $P_{\text{gof}} \geq 0.2$ the fraction remains at 47%. Thus, this method is likely more stable and less biased than the mass shell selection presented above. However, also a ΔR selection can have an impact on the phase-space, and further sanity checks are necessary if this kind of selection is implemented. In the present analysis the default "leading p_T selection" of W jets is preferred, and the testing of improved W selections is left for the next iterations.

6.1.5 Event Weights Applied on Simulated Samples

In the above the lepton and b jet selection scale factors were presented to adjust the event acceptances (and thus the event yields) between simulation and data. Jet energy resolution scale factors were presented in chapter 4, but these are applied as jet energy scales instead of event weights. These three are the most important scale factor categories, but there are yet two additional modes of event scaling required for simulations: pileup reweighting and the L1 Prefiring weights. In both of these methods an event-wise weight is computed based on the event topology. Neither of these scales are exactly scale factors, but their usage resembles that of the scale factors.

Pileup

The simulated pileup distribution – i.e. the number of pileup vertices in simulations – mimics that observed in data. However, the profiles in data and simulation seldom are a perfect match. This necessitates further weighting for simulations, performed year-by-year.

The pileup distributions are weighted to match data with the minimum-bias cross-section $\sigma = 69.2\,\text{mb}$ [116]. Minimum-bias events are the most common and uneventful collisions that occur at each bunch crossing. This is in contrast to the more rare and interesting events selected using the HLT conditions. The pileup target profile in data is fetched from the CMS luminosity database. The simulated pileup profile is essentially the same for most simulated samples in a single run year.

L1 Prefiring of ECAL and Muons

In the run years 2016 and 2017, the propagation of the gradual timing shift of the ECAL to the L1 trigger primitives was not performed successfully [117]. As a result, a notable fraction of high $|\eta|$ and high p_T trigger primitives were associated to the previous bunch crossing. Since the L1 trigger does not allow firing in two consecutive bunch crossings, this has caused self-vetoing for some events. For muons, a similar but smaller effect was present during Run 2, stretching up to 2018 [118].

As the prefiring effects were not simulated, additional weights are applied to correct for the effect. In 2018, the full effect is magnitudes smaller than in 2017 (or 2016), as it consists only of muon prefiring. The ECAL prefiring mainly affects objects in $2.5 < |\eta| < 3.0$ above 100–200 GeV, so these issues have a lesser impact in the present analysis.

6.1.6 Control Plots for the Baseline Selection

In this subsection event kinematics after the baseline selection are presented. Furthermore, the split of signal events into categories is explained. Diverging from the choice made in the historical reference [19, 20, 22], ratio plots are shown in the form simulation/data and not vice versa. This is motivated by the simple fact that statistics in data are good, and systematic uncertainties are considered exclusively for simulation. In simulations there is in general minimally a 5% cross-section uncertainty that needs to be incremented to statistical uncertainties. This uncertainty in the theoretical cross-sections follows mainly from the uncertainty in e.g. factorization and renormalization scales.

New Categorization of $t\bar{t}$ Events

A small piece of information from the kinematic fit is included in the plots of this section in the form of the jet matching status within the semileptonic $t\bar{t}$ signal events. The categorization does not affect the shapes observed in these plots, but

it can provide physical insight. The status has been historically divided into three categories: Correct, Wrong, and Unmatched Permutations (CP, WP and UN, respectively) [19, 20, 22].

In the category **correct**, all jets are correctly matched, whereas in the category **wrong** the analysis jets contain the correct four jets, but they have been incorrectly assigned. With the great performance of DeepJet, this mainly means event hypotheses where the hadronic and the leptonic b jets have gotten interchanged. Finally, the category **unmatched** has historically been reserved for all other event hypotheses, including those with incomplete simulation truth information and those with incorrect jet matches. Moreover, dileptonic and all-hadronic $t\bar{t}$ events have also been included.

In this analysis, the dileptonic and all-hadronic events are separated from the **unmatched** category. Moreover, the separate category of $t\bar{t}$ +V (V=W/Z) is appended. With the pythia settings used at the CMS weak radiation is not properly modeled, necessitating the handling of this relatively small background separately. The bosons of the weak interaction decay producing leptons or jets, distorting the base selection used in this analysis.

The remnants of the unmatched category are further split into two: **mlb** and the remaining UN events (**no match**). The category **mlb** refers to cases, where the leptonic b jet has been correctly matched. This category is motivated by the usage of the charged lepton + b jet invariant mass, introduced newly in the legacy 2016 analysis [22]. This experimental variable is defined as

$$m_{lb} = \sqrt{\left(p_{lepton} + p_b^{leptonic}\right)^2}. \tag{6.2}$$

In the **mlb** category, it is likely (but not certain) that also the hadronic b jet is correct. The most frequent reason for an event to drop from the CP category into **mlb** is incorrect matching of the W jets. In practice, the differences between **mlb** and **no match** are analogous to those between CP and WP.

Distributions of Kinematic Variables

The jet number spectrum was already presented above in Fig. 6.1. In simulation, an excess of events at the lower jet counts (<9) appears to exist, whereas data has more events at the high jet counts (>9). This seems to indicate that ISR/FSR and/or matching/merging is not optimal in CP5 for $t\bar{t}$ events. Generally, simulation shows an excess of approximately 5% to the event yield in data, and this excess grows to 10% with the $P_{\text{gof}} \geq 0.2$ selection presented in Section 6.2. This level of disagreement can be explained by parton showers and factorization/renormalization scales, which are not included in the uncertainty bars in the figures.

Fig. 6.2 displays the lepton p_T distribution in the studied channels on a logarithmic y-axis scale. The discrete step at low p_T is due to different electron and

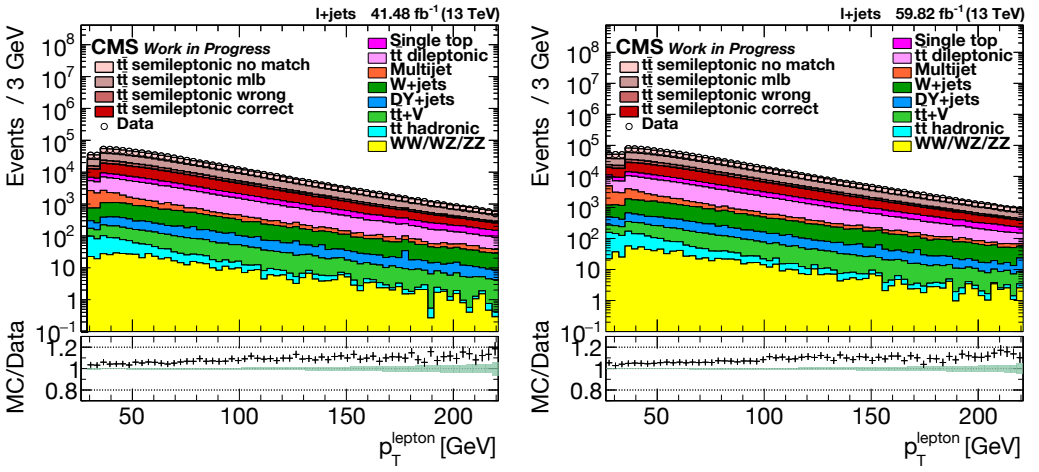

Figure 6.2: Lepton p_T after the baseline selection on a logarithmic y-scale, $mu17$ + $e17$ (left), $mu18$ + $e18$ (right). A discrete step is observed near the left edge, explained by the different p_T thresholds of electrons and muons. Simulation to data ratio error bars consist of statistical and scaling uncertainty in simulation, whereas statistical uncertainty in data is displayed as a green band.

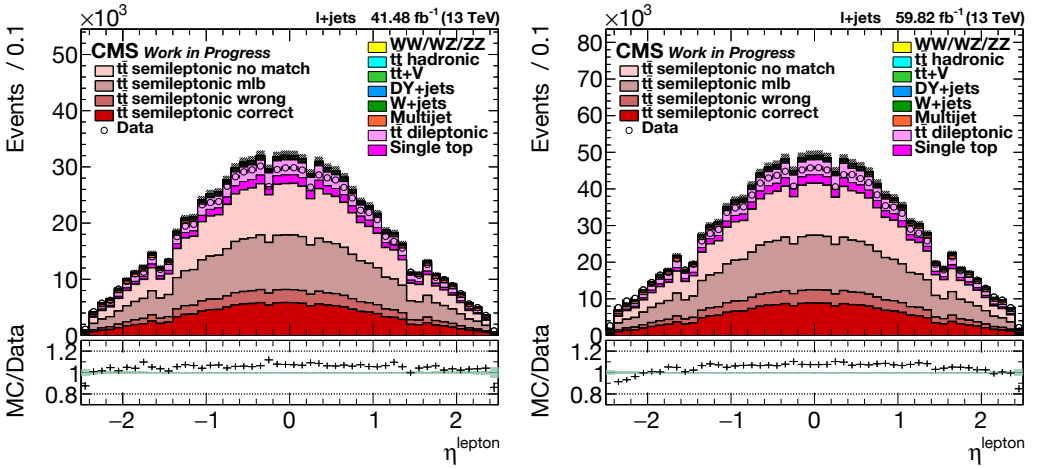

Figure 6.3: Lepton η after the baseline selection on a linear y-scale, $mu17$ + $e17$ (left), $mu18$ + $e18$ (right). Simulation to data ratio error bars consist of statistical and scaling uncertainty in simulation, whereas statistical uncertainty in data is displayed as a green band.

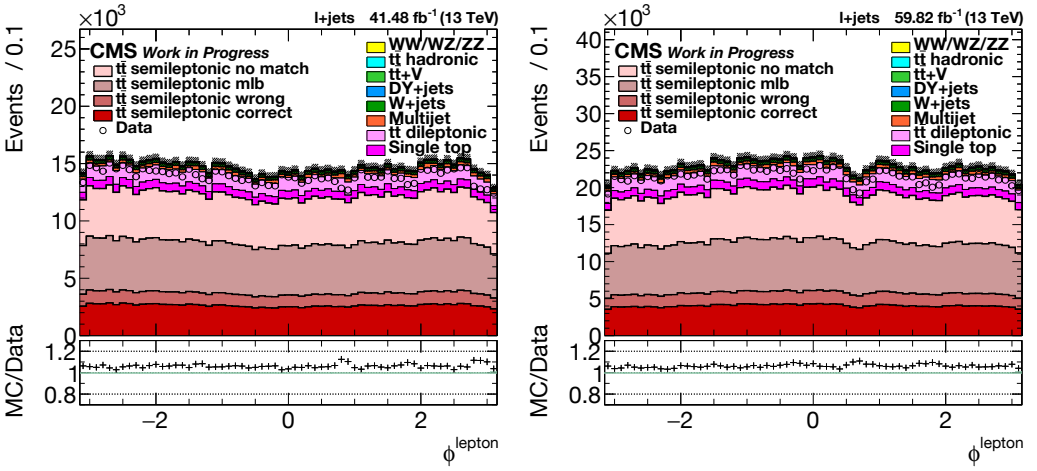

Figure 6.4: Lepton ϕ after the baseline selection on a linear y-scale, $mu17 + e17$ (left), $mu18 + e18$ (right). Simulation to data ratio error bars consist of statistical and scaling uncertainty in simulation, whereas statistical uncertainty in data is displayed as a green band.

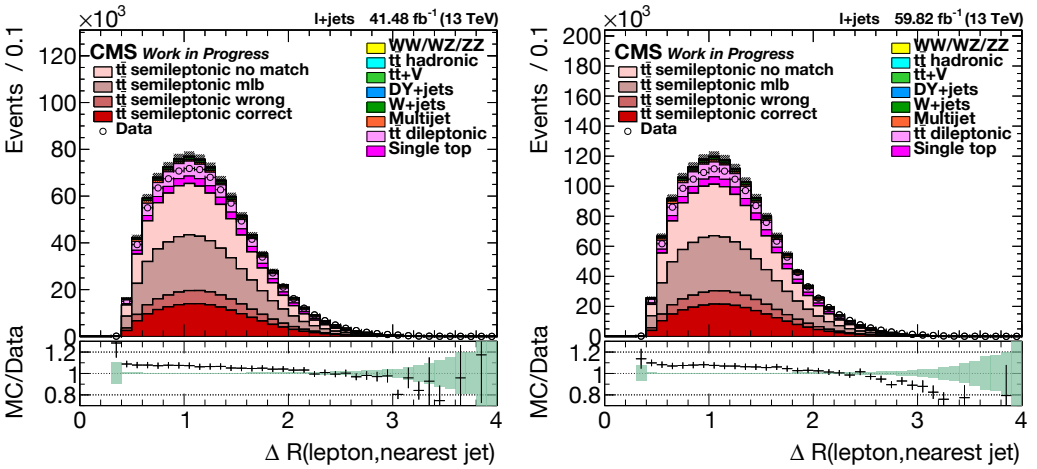

Figure 6.5: ΔR between the lepton and the nearest jet after the baseline selection on a linear y-scale, $mu17 + e17$ (left), $mu18 + e18$ (right). Simulation to data ratio error bars consist of statistical and scaling uncertainty in simulation, whereas statistical uncertainty in data is displayed as a green band.

muon selection p_T thresholds. The increase of the difference between data and simulation with lepton p_T likely originates from the limitations of top quark modeling at NLO [119–123]. It can be observed that many of the signal and background processes offer a somewhat unchanged fraction of the statistics as a function of the lepton p_T, Some of the backgrounds (mainly Multijet and fully hadronic $t\bar{t}$) are concentrated at low lepton p_T, indicating that the events include a fake lepton. In Fig. 6.3 and Fig. 6.4 distributions are given for the lepton η and ϕ, correspondingly. In the former plots, the electron removal in the barrel-endcap transition region can be easily observed. In the ϕ-dependent plot some fluctuations are seen, and these are modeled reasonably well in simulation. The largest non-flat trends originate from the electron channel and are likely explained by detector effects.

In Fig. 6.5, the ΔR distance between the charged lepton and the nearest jet is shown. The statistics are correctly located above the $\Delta R = 0.4$ radius of the jets. The distribution is peaked near $\Delta R = 1$, which is for the most part explained by the b jets of the leptonically decaying tops. If the original top quark p_T is more than 175 GeV, the distances between the decay products are likely to obey $\Delta R < 2.0$. Simulation appears to prefer slightly closer distances between the lepton and the closest jet compared to data, producing a slope. This could be explained by a number of systematic uncertainties.

Fig. 6.6 presents the value of $p_{\mathrm{T}}^{\mathrm{miss}}$ at a linear y-axis. In simulation, somewhat higher $p_{\mathrm{T}}^{\mathrm{miss}}$ values are preferred than in data. From the simulation to data ratio it is observed that the slope is similar as for charged leptons in Fig. 6.2. This is expected, as $p_{\mathrm{T}}^{\mathrm{miss}}$ is most likely dominated by the neutrino p_T. The neutrinos and charged leptons originate from the same W boson, so a similar behavior can be expected. The greatest exception is that the $p_{\mathrm{T}}^{\mathrm{miss}}$ spectrum continues to zero, since there is the selection on this observable.

Finally, distributions of basic jet variables are given. In all the remaining plots the y-axis scale is linear, and similar formatting rules are followed as in the previous plots. Fig. 6.7 gives the p_T distribution of the l jets from $\mathrm{W}\rightarrow qq'$. In a similar fashion, Fig. 6.8 displays b jet p_T distribution. Both of these p_T spectra show similar features, with the b jet p_T values being somewhat higher than those of the l jets. The p_T slopes share similarities with Fig. 6.2 and Fig. 6.6, and the same argument of top p_T mismodeling can be used here. However, there are also differences: the mismatch between data and simulation is maximal around $p_T = 100$ GeV, whereas for the leptons it continues growing. This could be explained by modeling exclusive to jets, such as the parton showers.

The η plots are given only for the ($\mathrm{W}\rightarrow qq'$) l jets in Fig. 6.9 and the b jets in Fig. 6.10. Here, **the impact of the jet veto maps presented in chapter 4 is clearly observed as a yield asymmetry around $\eta = 0$**. There are minor differences between the η-dependence of l jets and b jets, which could be explained for instance by the uncertainties in b-tagging. The data–simulation mismatch tends to be lower at $|\eta| > 2.0$, which could be explained e.g. by simulation tuning uncertainties

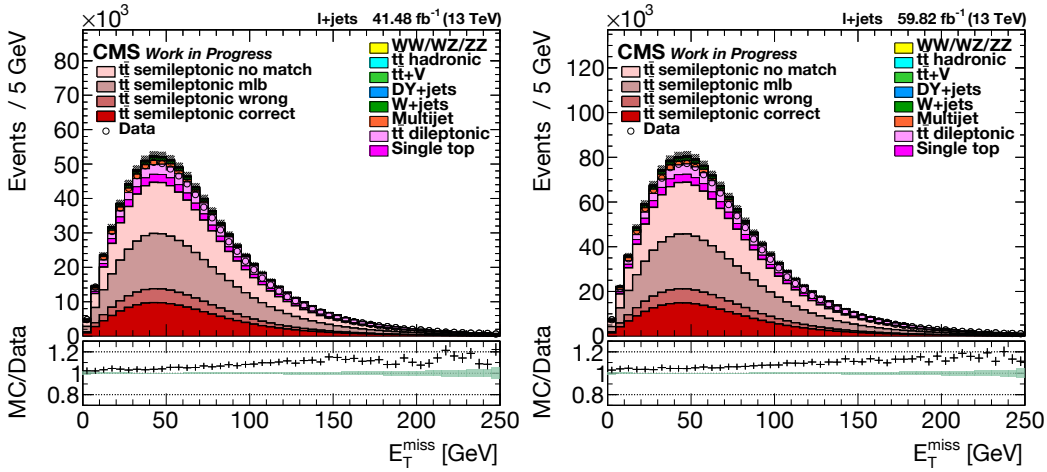

Figure 6.6: $p_{\mathrm{T}}^{\mathrm{miss}}$ after the baseline selection on a linear y-scale, $mu17 + e17$ (left), $mu18 + e18$ (right). Simulation to data ratio error bars consist of statistical and scaling uncertainty in simulation, whereas statistical uncertainty in data is displayed as a green band.

or by the limitations of NLO $\mathrm{t\bar{t}}$ compared to NNLO.

For ϕ, the l jet and b jet plots are given in Fig. 6.11 for the ($\mathrm{W}\rightarrow qq'$) l jets and in Fig. 6.12 for the b jets. Also here the jet veto zone effects can be well observed e.g. around $\phi = 3$ for the 2018 datasets. These jet veto jumps in the even counts are fully expected and not a cause of concern: the effects are the same in data and simulation.

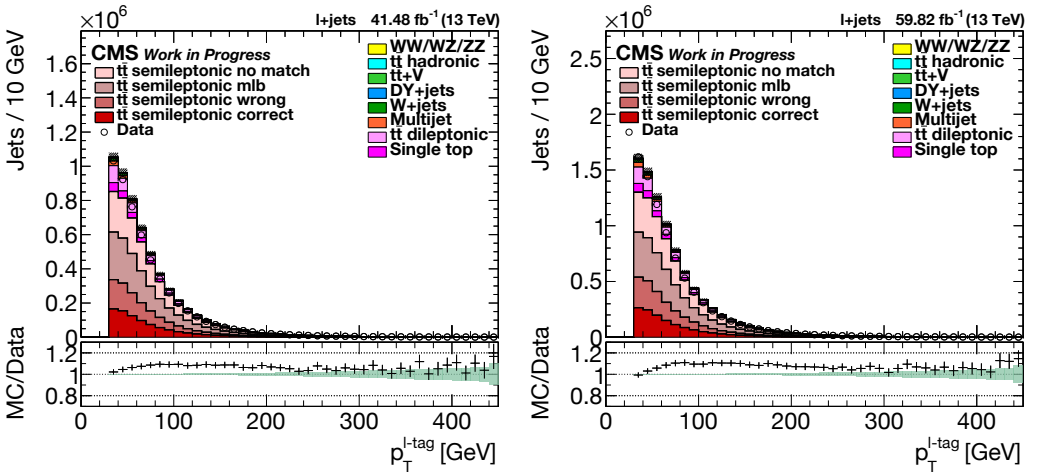

Figure 6.7: l jet (from W→ qq') p_T after the baseline selection on a linear y-scale, $mu17 + e17$ (left), $mu18 + e18$ (right). Simulation to data ratio error bars consist of statistical and scaling uncertainty in simulation, whereas statistical uncertainty in data is displayed as a green band.

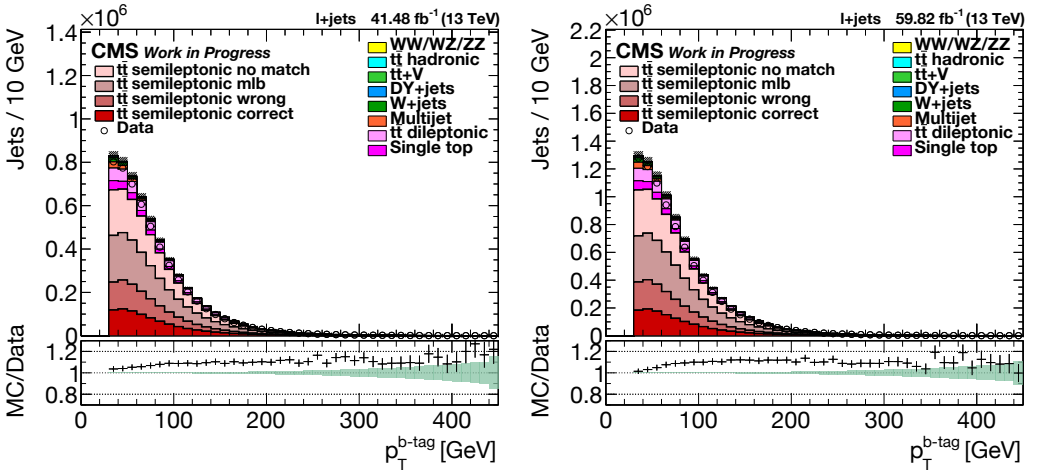

Figure 6.8: b jet p_T after the baseline selection on a linear y-scale, $mu17 + e17$ (left), $mu18 + e18$ (right). Simulation to data ratio error bars consist of statistical and scaling uncertainty in simulation, whereas statistical uncertainty in data is displayed as a green band.

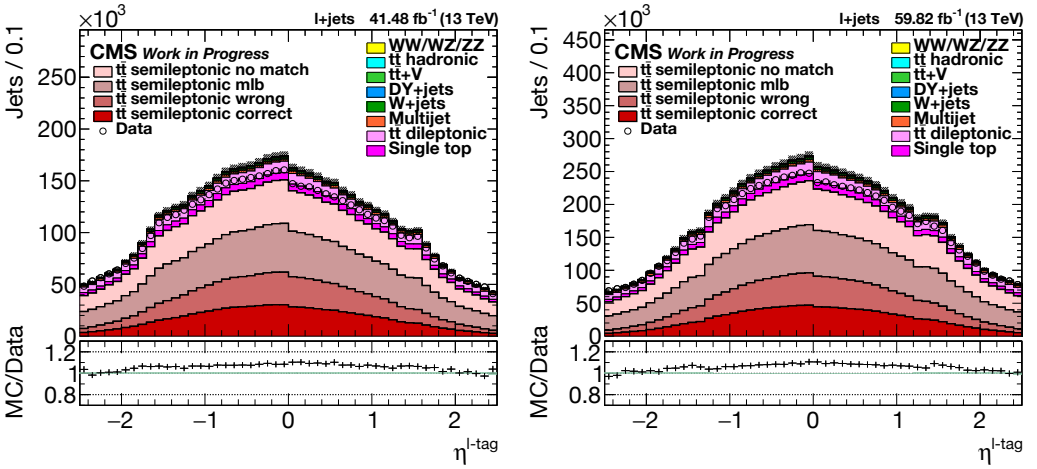

Figure 6.9: l jet (from W→ qq') η after the baseline selection on a linear y-scale, $mu17 + e17$ (left), $mu18 + e18$ (right). A $\pm\eta$ asymmetry is observed, caused by the jet veto zones. Simulation to data ratio error bars consist of statistical and scaling uncertainty in simulation, whereas statistical uncertainty in data is displayed as a green band.

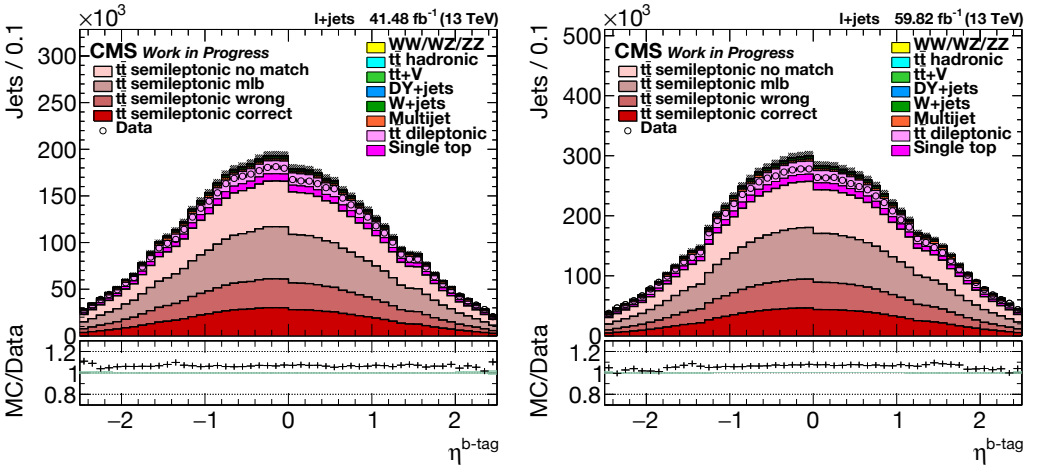

Figure 6.10: b jet η after the baseline selection on a linear y-scale, $mu17 + e17$ (left), $mu18 + e18$ (right). A $\pm\eta$ asymmetry is observed, caused by the jet veto zones. Simulation to data ratio error bars consist of statistical and scaling uncertainty in simulation, whereas statistical uncertainty in data is displayed as a green band.

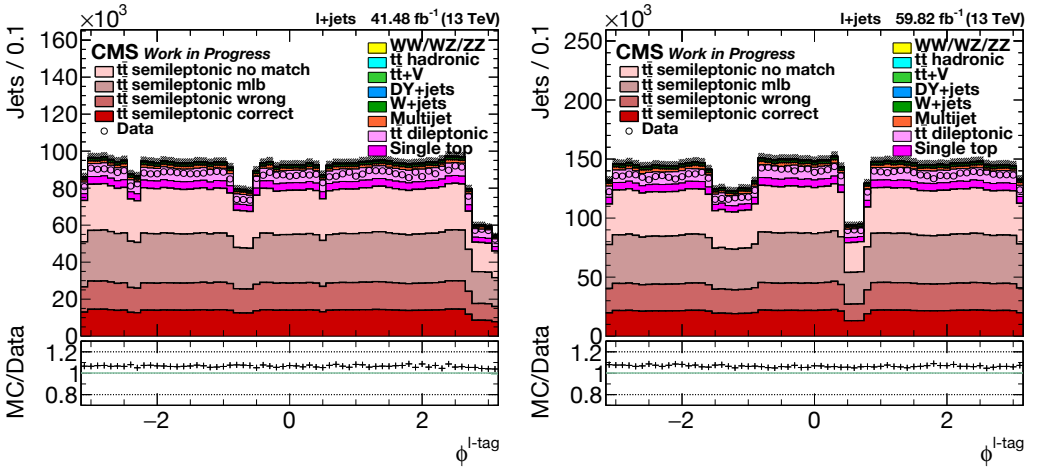

Figure 6.11: l jet (from W→ qq') ϕ after the baseline selection on a linear y-scale, $mu17 + e17$ (left), $mu18 + e18$ (right). Abrupt steps are caused by the jet veto zones. Simulation to data ratio error bars consist of statistical and scaling uncertainty in simulation, whereas statistical uncertainty in data is displayed as a green band.

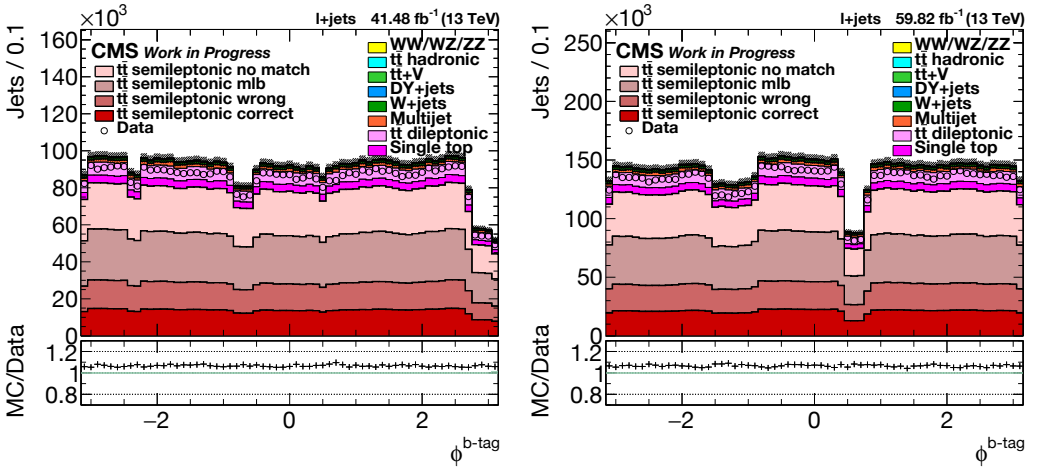

Figure 6.12: b jet ϕ after the baseline selection on a linear y-scale, $mu17 + e17$ (left), $mu18 + e18$ (right). Abrupt steps are caused by the jet veto zones. Simulation to data ratio error bars consist of statistical and scaling uncertainty in simulation, whereas statistical uncertainty in data is displayed as a green band.

6.2 Kinematic Fitting

The HitFit kinematic fit for the semileptonic $t\bar{t}$ channel was originally developed for a DØ thesis in 1995 [124][2]. The study found an impressive result for the time with only $50\,\mathrm{pb}^{-1}$ of data at 1.8 TeV: $m_t = 199\,\mathrm{GeV}$ (both stat and syst. uncertainties around 20 GeV). HitFit is algorithmically specific for the lepton+jets event topology, and for instance the all-hadronic analysis must be performed with different software tools. The original Fortran code was rewritten in C++ for Tevatron Run II, and ported into CMSSW for CMS Run 1 purposes. The object resolutions required in HitFit were first updated to 13 TeV (CMS Run 2) by the author of this thesis, as explained later in this section.

6.2.1 Fitting Procedure

HitFit performs a fit yielding the best estimate of the top quark mass (m_t^{fit}) in each event. The fit employs six physical objects following from the top quark pair decay: four jets, a charged lepton and a neutrino. For the four jets and the charged lepton fitting is performed in three dimensions: p_T, η and ϕ. Three dimensions suffices, as HitFit is intended to make a best estimate of the parton level based on the measured jets. On the parton level, the fourth dimension is made redundant by the fixed mass values. Mass is set to 0 GeV for the charged leptons and the light quarks and to the approximate value of 4.7 GeV for b quarks.

In the fit, explicit neutrino kinematics are replaced by the quantity $\vec{k}_T$ defined as a sum of the $\vec{p}_T$s of the signal objects:

$$\vec{k}_T = \vec{p}_T^{\,\nu} + \vec{p}_T^{\,lep} + \vec{p}_T^{\,b1} + \vec{p}_T^{\,b2} + \vec{p}_T^{\,q1} + \vec{p}_T^{\,q2}. \tag{6.3}$$

At the first step of the fit, $\vec{p}_T^{\,\nu} = \vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$. Fitting is performed on the x and y components of $\vec{k}_T$, the expectation values of which are centered at zero. To be more exact, the $\vec{k}_T$ quantity only covers the transverse neutrino momentum, and the neutrino z component is deduced separately, as is explained later.

There are three eventwise conditions employed in the fit:

1. **Light quark pair mass = m_W = 80.4 GeV**
2. **Neutrino + charged lepton mass = m_W = 80.4 GeV**
3. **Leptonic and hadronic top quark mass values are equal = m_t^{fit}**

Here the m_W and m_t decay widths are omitted. The l jet m_W condition scales away JEC variations and makes m_t^{fit} independent of the reconstruction-level value of

[2] Note that despite of HitFit being developed at DØ, the most recent DØ analysis analyzed in chapter 5 relies on different tools and techniques with a similar purpose [36].

the "W to quark jet pair" mass (m_W^{reco}). The leptonic W mass condition is consumed by the otherwise unknown z-component of the neutrino momentum.

The m_t^{fit} condition also allows the main backgrounds[3] to introduce m_t dependence into the results. Even if there is only one genuine top quark match, this will impact the value of m_t^{fit}, which requires the hadronic and leptonic m_t values to be the same. The genuine m_t resonance (constant) will attract the fake m_t resonance (varying event-by-event) in a non-trivial manner.

Even if the fit is designed for reaching the parton level, there are numerous assumptions and approximations on the way. Hence m_t^{fit} is treated only as an enhanced experimental observable, and not the true value of m_t. The corresponding reconstruction level variable, m_t^{reco}, is less peaked and more convolved. m_t^{fit} is preferred over m_t^{reco} due to the automatic decorrelation between m_t^{fit} and m_W^{reco} that also allows the simultaneous use of these separate variables [19, 20].

The differences between the two m_t variables are seen by comparing Fig. 6.13 with Fig. 6.14. For the *correct* permutations the observed peaks are almost Gaussian, whereas the other permutations have long tails directed towards high m_t values. The Figures 6.13 and 6.14 also indicate that the m_t value is likely to be lower in data than in the default simulation ($m_t = 172.5\,\text{GeV}$), unless some of the systematic uncertainties explain the differences in the m_t peak positions.

The quantity minimized in the fit is

$$\chi^2 = (\vec{x} - \vec{x}_m)^T \, \mathcal{G} \, (\vec{x} - \vec{x}_m) \,, \tag{6.4}$$

where the matrix $\mathcal{G}$ is the error matrix containing the inverse squared resolutions between the jet and parton levels. Additionally, a set of constraints $\vec{F}(\vec{x}) = \vec{0}$ is imposed on the minimization process. Here, the vector $\vec{x}_m$ contains the measured observables[4] and $\vec{x}$ the fitted ones, mimicking parton level. Most importantly, the vector $\vec{F}(\vec{x})$ contains the mass shell conditions.

There are two degrees of freedom in the fit, following from the 26 constraints and 24 quantities (four-momenta of six particles) present in it. Six constraints are found in the particle masses and $3 \times 5 + 2$ in the kinematic properties of the particles and $\vec{k}_T$. Finally, $2 + 1$ constraints are found in the W boson and m_t^{fit} constraints. The χ^2 probability for two degrees of freedom is used as a measure of the goodness of fit:

$$P_{\text{gof}} = 2 \times P_{\chi^2}(k = 2) = \exp\left(-\frac{1}{2}\chi^2\right). \tag{6.5}$$

This measure is used conventionally to select the signal region. Here, the value of the original χ^2 probability density is multiplied by a factor of two for convenience, so that when $\chi^2 \to 0$, $P_{\text{gof}} \to 1$. The same cut $P_{\text{gof}} \geq 0.2$ is used for the main signal selection as in the earlier CMS lepton+jets analyses [19, 20, 22]. The same optimization and arguments for the current selection are valid as in the earlier

[3] Single top tW channel and dileptonic $\mathrm{t\bar{t}}$.

[4] p_T, η and ϕ values for jets and the charged lepton, as well as k_x and k_y.

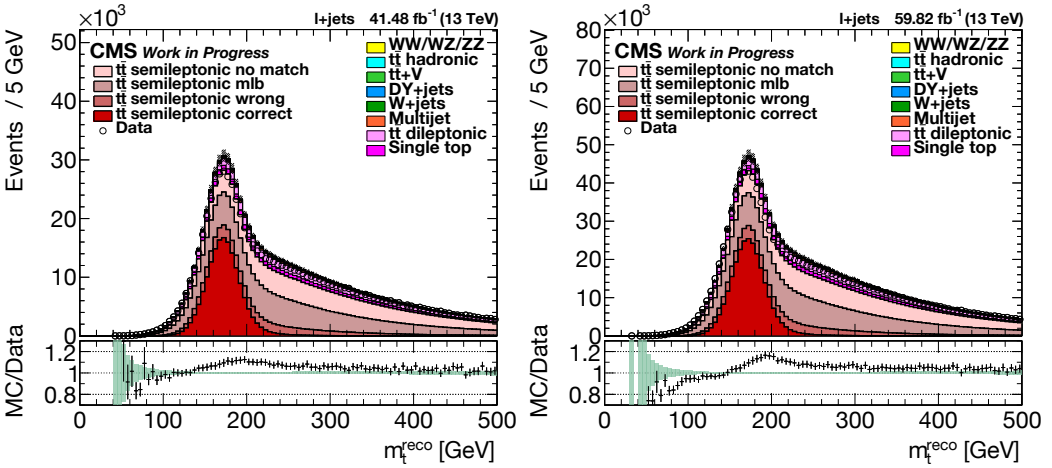

Figure 6.13: m_t^{reco} without P_{gof} cuts, $mu17 + e17$ (left), $mu18 + e18$ (right). Simulation to data ratio error bars consist of statistical and scaling uncertainty in simulation, whereas statistical uncertainty in data is displayed as a green band.

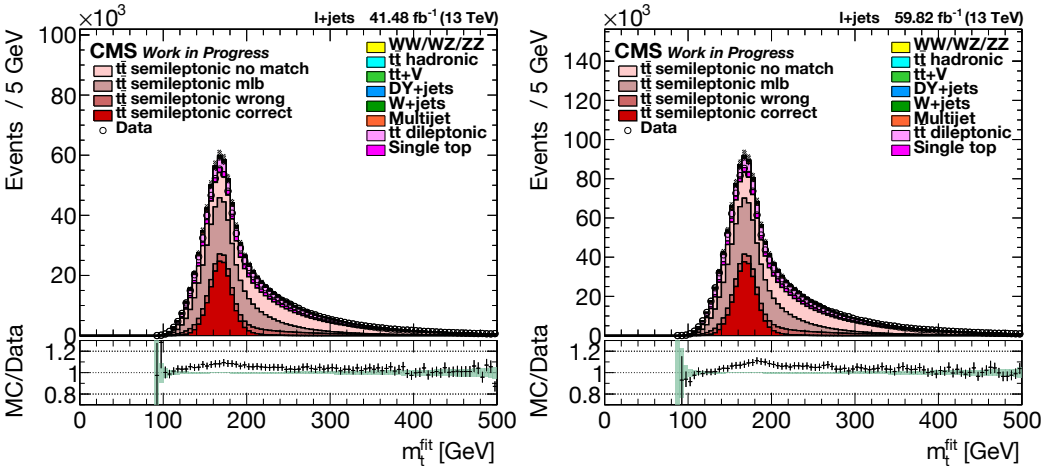

Figure 6.14: m_t^{fit}without P_{gof} cuts, $mu17 + e17$ (left), $mu18 + e18$ (right). Simulation to data ratio error bars consist of statistical and scaling uncertainty in simulation, whereas statistical uncertainty in data is displayed as a green band.

Run 2 results. The main motivation for this kind of a cut is selecting the events where the events are very close to a possible $\mathrm{t\bar{t}}$ hypothesis. One of the most typical reasons for small P_{gof} values in semileptonic $\mathrm{t\bar{t}}$ events is the location of the l jet ($\mathrm{W}\rightarrow qq'$) candidates far from the W boson mass shell. This cut might as well be e.g. $P_{\mathrm{gof}} \geq 0.15$, but as there is no pressing motivation to change it, the historical reference value is used.

As an update to the old $P_{\mathrm{gof}} \geq 0.2$ condition, the $P_{\mathrm{gof}} < 0.2$ domain is also used in the measurement. In the spirit of the legacy 2016 analysis [22], this is motivated by the m_{lb} variable defined in Eq. (6.2). It receives good m_t dependence and statistics in $P_{\mathrm{gof}} < 0.2$ as well. This follows from the baseline selection, which requires b jets and exactly one charged lepton to be present. Against this background, the additional P_{gof} selection mostly excels in filtering out the events with good $\mathrm{W}\rightarrow qq'$ jet matches. Already at the baseline level it is likely that the charged lepton and b jets are correct, especially considering that these can also be correctly matched in some of the top quark production backgrounds.

Each event receives four fits maximally: there are two possible pairings of the b jet with the hadronic W, and in each case maximally two possible initial values for the neutrino z component. In the earlier analyses employing the Ideogram likelihood method, these were treated as four separate permutations, which were then weighted and summed up [20].

The jet permutations are more fundamental than the ones corresponding to the neutrino solutions. The two initial values for the neutrino can end up with the same result, and/or there can exist only a single neutrino solution. In this work, only the permutation with the greatest P_{gof} value is used in a maximum likelihood sense. Thus, double-counting issues with the neutrino permutations are avoided. This strategy has also been adopted in the new legacy 2016 analysis [22].

6.2.2 Object Resolutions in the Fit

The kinematic fit relies heavily on Eq. (6.4), and hence the values in $\mathcal{G}$. It is conventionally assumed that $\mathcal{G}$ is diagonal, so that the diagonal entries are the statistical resolutions of p_T, η and ϕ and k_x/k_y between the parton- and jet-levels. Here, the effective parton level is defined by the parton information provided by the event generator software. For the charged leptons, these resolutions are much smaller than for jets, as the parton level equivalents of charged leptons do not drastically differ from the reconstructed ones. On the other hand, for the k_x/k_y values the resolutions are much larger than for jets, as $\vec{k}_T$ is a composite object that contains all the analysis jets.

The object resolutions were heavily developed during Run 1, but knowledge on them was lost before arriving to Run 2. The previous Run 2 analyses [20, 22] were performed with the Run 1 object resolutions. The author of this thesis made the first Run 2 update on these resolutions by resurrecting Run 1 code developed by the University of Ghent authors Volker Adler and Sven Dildick. This version of the

code is not publicly available any more.

The new version of the resolution code is found in the repository [6], which only CMS members can currently access. This will hopefully be made public in the future to promote scientific openness. The old resolution files are found in the public CMSSW repository [66], whereas the new resolution files have been added into the gitlab analysis repository [6].

In itself, an update from Run 1 to Run 2 resolutions was motivated by the major increase in collision energies and changes in software and calibration. Ideally, the JECs, simulation software versions and tuning should be aligned in the resolution derivation and the target analysis. Moreover, the central CMSSW Run 1 resolutions were partially broken due to rounding errors. These included problems in the electron resolutions [125] and problems in the light quark jet resolutions [126]. These issues were located by the author of this thesis, and consequently also fixed in the legacy 2016 analysis [22].

The resolutions are found through a statistical η and p_T binned handling of the differences between reconstruction and parton level. The p_T binning generally varies by object and for jets, η-binning follows HCAL towers. For each of the observables $X \in \{p_T, \eta, \phi, k_x, k_y\}$, a simple difference metric is constructed:

$$\Delta_X = X^{\text{parton}} - X^{\text{reconstruction}}. \tag{6.6}$$

The Δ_X distributions are fit into a simple Gaussian-like distribution, yielding a central value μ and a resolution σ.

The values of μ for jet p_Ts can be interpreted as the final JEC level: the **corrections to the parton level**. This final jet correction level consists of out-of-cone corrections and corrections for the missing neutrinos. Using the parton JECs together with the resolution σ one ends up with similar full effects as with the DØ TFs. As a reminder, the TFs include both a resolution and a translation. The parton JECs were experimented with during Run 1, but it was observed that they made little difference in the fits. This follows from the general fact that $|\mu| \ll \sigma$ for most observables, including p_T.

The only notable exception to the small $|\mu|$ rule is the b jet p_T. Especially at low p_T the Gaussian distribution becomes asymmetric, with a long tail. This follows from the neutrinos produced in semileptonic b quark decays that are not included in the reconstructed b jets. Here, the peak position and resolution fit well, but the non-Gaussian tail is not considered.

At the worst, the offset position of the peak $|\mu|$ is in the order of 0.5σ. Thus, even in the worst case the shift imposed by the parton JECs is relatively small compared to the resolution. It should also be underlined that the b quark decays are highly dichotomous: 2/3 of the decays are fully hadronic (not semileptonic) [62] and there is no significant offset and tail caused by the missing neutrinos. Hence, leaving out the parton JECs is a conscious decision to prefer the fits of b quark jets decaying in a fully hadronic manner. Ideally, the semileptonic status of a jet could be predicted

from its charged lepton content. This would lead into a separate handling of the semileptonic and hadronic b jet decays in a similar manner as was done at DØ.

In summary, we only focus on the resolutions in this work, as has been done in HitFit analyses since Run 1. In each η-bin, the dependence of the resolutions on p_T is fit according to

$$\sigma_X\left(p_T\right) = \sqrt{C_X^2 p_T^{2\alpha} + R_X^2 p_T^{\alpha} + N_X^2}. \tag{6.7}$$

C_X, R_X and N_X are the parameters determined in the fit for each of the variables $X \in \{p_T, \eta, \phi, k_x, k_y\}$. The parameter α varies by observable:

- For $X = p_T$, $\alpha = 1$ is the most frequent choice.
- The muons are an exception with $X = p_T^{-1}$ and $\alpha = -1$.
- For $X = \eta$ and ϕ, $\alpha = -1$ without exceptions.
- For the k_x/k_y vector components $\alpha = -1$, but these have a more complicated dependence than simply that on k_x/k_y.

Most of these relations are motivated by the fact that as p_T increases, the p_T resolution must also increase. In contrast, the η/ϕ readings get more precise, as p_T grows.

A major critique on many $\mathrm{t\bar{t}}$ analyses and the kinematic fit can be based on the heavy usage of the simulated parton level. Few top analyses are able to function without a simulation truth tagging between the partons formed in a top quark decay and the reconstructed jets. The kinematic fit takes things one step further: the momentum resolutions between the tag partons and jets are computed. The parton level critique is based on two main arguments:

1. **The parton level is not well defined and handled only with phenomenological models in simulations.**
2. **Depending on the simulation software, the parton momenta might not be fully updated into the reconstruction-level coordinate system, and e.g. momentum reshuffling between particles can occur.**

These points of critique are not incorrect, but they also do not express the whole truth. In Pythia 8 the matching between jets and partons works remarkably well, and the jet-parton matching quality is also quite good in Herwig 7. Some differences can be observed between these two generator tools, with better matching properties being found in Pythia 8.

The resolution values extracted from Pythia 8 must be interpreted as a sufficient approximation of the parton level. This is a part of the reasoning of using m_t^{fit} simply as another experimental observable. It also helps that in a top quark production process, the nature of QCD is highly in the perturbative QCD regime.

In this regime, the phenomenological QCD models work best, in contrast to the non-perturbative regime, where ambiguities are more frequent. Finally: it is most important that the magnitudes of the resolutions are correct, as this is a χ^2 fit. The exact resolution values do not have a great impact on the fit results.

6.2.3 HitFit Software Updates

Besides the resolution file update, performance and logic updates have been applied to HitFit by the thesis author. As the CMSSW release/backport cycle is slow, they have not been added to the central CMSSW version at the time of writing, but can be found in the public repository of Ref. [7]. In the following, an important logic update is clarified. The performance upgrades are more abstract and not further explained. The interested viewer can review them in GitHub.

In the central CMSSW version of HitFit, for the neutrino/$\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ resolution the observed resolution of $p_{\mathrm{T}}^{\mathrm{miss}}$ as a function of $p_{\mathrm{T}}^{\mathrm{miss}}$ is being used. This parametrization is applied on the x and y components of $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ as a function of their values. Neither of these partial approaches are physically meaningful, as $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is a vector sum collected over the whole detector. The absolute value $p_{\mathrm{T}}^{\mathrm{miss}}$ or the values of its x and y components are analogous to a result of random walk, and cannot be meaningfully used to calculate a resolution. As a result, the functional dependence of this resolution has been observed to be mostly flat, i.e. this resolution has been effectively taken to be a constant. Even if the underlying logic error is grave, the consequences of it have been marginalized by this semi-constant value of $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ resolution.

In the upgrades provided in Ref. [7] a physically better motivated approach is taken. If the $p_{\mathrm{T}}^{\mathrm{miss}}$ resolution is dependent on any physical observable, it is the total transverse energy sum – i.e. $\sum E_t$, which is typically found paired with $p_{\mathrm{T}}^{\mathrm{miss}}$. The important difference between $p_{\mathrm{T}}^{\mathrm{miss}}$ and $\sum E_t$ is that the former is a vector sum with a zero-centered expectation value, whereas the latter is a scalar sum that always attains positive values. As $p_{\mathrm{T}}^{\mathrm{miss}}$ is a composite object consisting of all vector objects in the event, $\sum E_t$ becomes a good estimator for its uncertainties. In the upgraded code, the x and y components of $\sum E_t$ are estimated, and the resolutions of the x and y components of $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ are taken to depend on these values.

6.2.4 Fit Probability Distributions

The kinematic fit is utilized for two purposes: for the values of m_t^{fit} and P_{gof}. P_{gof} on its part has two uses: the signal event cut and for the selection of the jet permutation that most probably corresponds to a semileptonic $\mathrm{t\bar{t}}$ event. In this subsection, distributions of χ^2 and P_{gof} are studied.

Fig. 6.15 displays the grand picture of χ^2 dependence of the different fit categories. Following Eq. (6.5), the transformation between χ^2 and P_{gof} is one-to-one

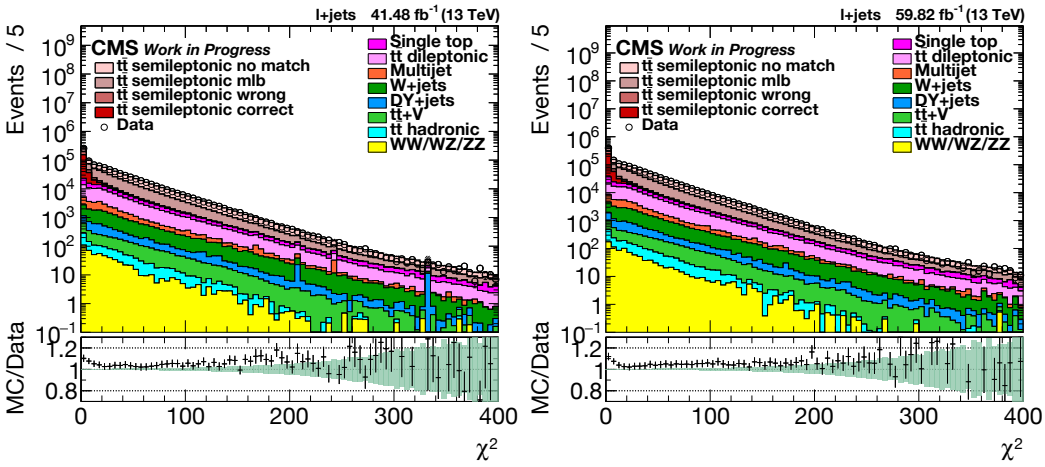

Figure 6.15: The kinematic fit χ^2 dependence on a logarithmic y-scale, $mu17 + e17$ (left), $mu18 + e18$ (right). Simulation to data ratio error bars consist of statistical and scaling uncertainty in simulation, whereas statistical uncertainty in data is displayed as a green band.

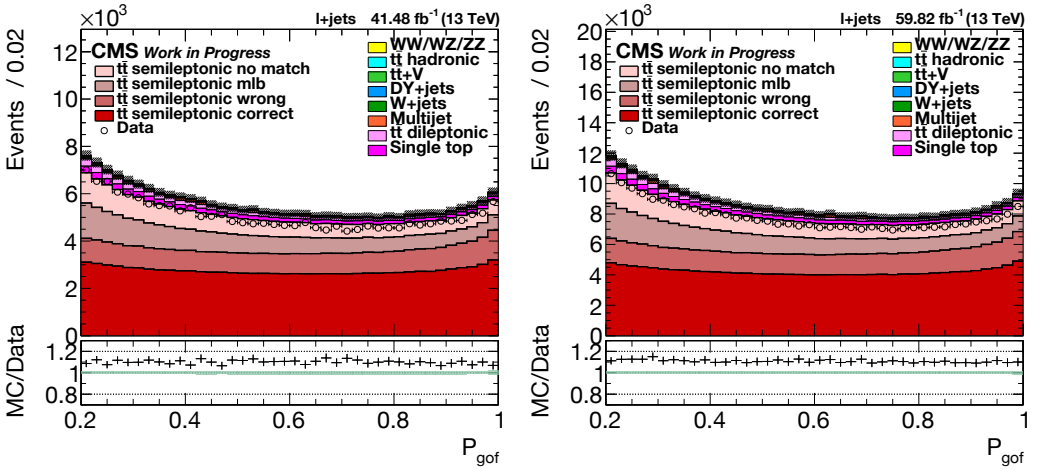

Figure 6.16: The kinematic fit P_{gof} dependence on a linear y-scale, $mu17 + e17$ (left), $mu18 + e18$ (right). Simulation to data ratio error bars consist of statistical and scaling uncertainty in simulation, whereas statistical uncertainty in data is displayed as a green band.

(bijection). As an important difference, changes in χ^2 and P_{gof} occur in the opposite directions: when χ^2 grows, P_{gof} becomes smaller. It can be observed that the *correct* signal permutations are enriched in this region. This topic is discussed in more detail at the end of the next section.

In Fig. 6.16, the P_{gof} dependencies are given in $P_{\text{gof}} \geq 0.2$ (i.e. $\chi^2 \leq 3.2189$). The long tail at $\chi^2 > 3.2189$ is packed at $P_{\text{gof}} \in [0, 0.2]$, excluded from the figures. This tail region is better displayed in the χ^2-dependent figures, and including it in the P_{gof} figures would change the y-axis scaling so that the interesting $P_{\text{gof}} \geq 0.2$ region would become poorly visible. It is easy to note that there is a constant offset of +10% in event yields in simulation as compared to data. From Fig. 6.15 it can be confirmed that this offset arises in the region $\chi^2 < 20$, whereas at greater χ^2 values there is better correspondence between data and simulation. At the current level of knowledge this effect between simulation and data, can be explained by the systematic uncertainty nuisance parameters of the analysis.

6.2.5 Impact of the P_{gof} Cuts

To provide a full picture of event yields before and after P_{gof} cuts, the event yields are given in Table 6.3 and Table 6.4 for the muon channels and in Table 6.5 and Table 6.6 for the electron channels. Similar trends are followed in all the channels.

The yields in the electron channels are notably smaller than in the muon channels for two reasons: the electron p_T thresholds are higher than those of muons and the ECAL barrel-endcap transition region is removed from the electron data. It can now also be confirmed that the data-simulation discrepancy is 5% before the $P_{\text{gof}} \geq 0.2$ cut and 10% after it. There are some variations between the channels and the years, which is taken into account when combining them.

Moreover, it is confirmed that the $P_{\text{gof}} \geq 0.2$ cut increases the relative fraction of *correct* permutations from less than 20% to almost 50%. Furthermore, the combined relative fraction of *correct* and *mlb* permutations in $P_{\text{gof}} < 0.2$ is around 43%, indicating that the charged lepton – b jet pairing is correct significantly often also in this phase-space region. In $P_{\text{gof}} \geq 0.2$ the *mlb* category only contains 16% of all events.

Fig. 6.17 shows the lepton p_T distribution in $P_{\text{gof}} \geq 0.2$, and Fig. 6.18 displays the same lepton p_T distributions without P_{gof} selection as Fig. 6.2 for comparison. Here, the drastic reduction of backgrounds and increase of *correct* permutations by the $P_{\text{gof}} \geq 0.2$ selection is observed. Neither of the main backgrounds dileptonic $\mathrm{t\bar{t}}$ and single top vanish in the $P_{\text{gof}} \geq 0.2$ cut, and the latter becomes the most important background. These backgrounds do not get reduced at high lepton p_T, which can be estimated to follow from the leptonically decaying top quarks.

In Fig. 6.19 the impact of the $P_{\text{gof}} \geq 0.2$ cut on m_t^{fit} is shown in a similar manner. The m_t signal features observed earlier in Figs. 6.13 and 6.14 are strengthened and the m_t resonance is now less wide. The data–simulation disagreement in the m_t^{fit} peak position appears to be stronger for the 2018 dataset. In Fig. 6.20, only the

post-cut m_W^{reco} distributions are shown. The peaked distributions are not Breit-Wigner like nor Gaussian. Simulation appears to prefer higher m_W^{reco} values than data, which can follow from discrepancies in e.g. FSR or JEC calibration.

Most of the CMS Run 1 and Run 2 measurements have been performed only using the m_t^{fit} and m_W^{reco} variables [19, 20]. Following the example of the legacy 2016 analysis [22], also other variables are studied in addition to them. The variable R_{bq} is designed for studying the b jet response, and it was originally introduced by ATLAS [21]. It is defined as the p_T ratio of the two signal b jets to the two signal quark jets:

$$R_{bq} = \frac{p_T^{b1} + p_T^{b2}}{p_T^{q1} + p_T^{q2}}. \tag{6.8}$$

Fig. 6.21 displays the distributions for R_{bq}^{reco}. These are peaked at unity, with a long tail on the right side. This is expected for a fraction between the in average more energetic b jets and the l jets. No significant trends between data and simulation are observed aside of a constant scale difference in event yields.

Above, it was motivated that the variable m_{lb} (defined in Eq. (6.2)) receives a large fraction of correct pairings in $P_{\text{gof}} < 0.2$. In $P_{\text{gof}} \geq 0.2$, however, m_{lb} cannot be utilized as such, since it is partially correlated with m_t^{fit}. The variable m_t^{fit} contains information from both of the m_t resonances with a focus on the hadronic one, so it can be argued that in addition to m_t^{fit} there is unused information worth one m_t resonance within each event. A reduced m_{lb} variable is defined for this task:

$$m_{lb}^{red} = \frac{m_{lb}}{m_t^{fit}}. \tag{6.9}$$

This m_{lb}^{red} variable is mostly decorrelated from m_t^{fit}. Fig. 6.22 and Fig. 6.23 display the distributions of $m_{lb}^{red,reco}$ ($P_{\text{gof}} \geq 0.2$) and m_{lb}^{reco} ($P_{\text{gof}} < 0.2$), correspondingly. The $m_{lb}^{red,reco}$ distribution is highly peaked between 0 and 1, and similarly as for R_{bq}^{reco} no significant trends are observed. The m_{lb}^{reco} distributions are peaked slightly above 100 GeV, and display a long tail on their right side. In the peak region, a slight preference for higher values is observed in simulation compared to data.

Thus in analogy to the legacy 2016 analysis [22], the five distributions m_t^{fit}, m_W, R_{bq}, m_{lb}^{red} (all four in $P_{\text{gof}} \geq 0.2$) and m_{lb} (in $P_{\text{gof}} < 0.2$) are studied and used in the main analysis of this work. The analysis structure is more closely explained in section 6.4. Some of these variables (m_t^{fit} and m_{lb}) function mainly as probes for m_t. On the other hand, m_W excels in *in-situ* light quark JEC and FSR calibration, and R_{bq} in b quark JEC *in-situ* calibration. Finally, m_{lb}^{red} can help in constraining various non-trivial systematic uncertainties.

Table 6.3: Legacy 2017 muon event yields. Simulation uncertainties include statistical and cross-section uncertainties. For data, the Poisson uncertainty is given. Fractions presented w.r.t. $t\bar{t}$ only (left) and full simulation (right).

	Full P_{gof}			$P_{\text{gof}} \geq 0.2$			$P_{\text{gof}} < 0.2$		
Sample	Events [k]	Fraction of Sim. [%]		Events [k]	Fraction of Sim. [%]		Events [k]	Fraction of Sim. [%]	
$t\bar{t}$ total	516.10 ± 0.14	100.00	91.61	130.93 ± 0.07	100.00	95.31	385.17 ± 0.12	100.00	90.41
Correct	100.03 ± 0.06	19.38	17.75	66.82 ± 0.05	51.03	48.64	33.21 ± 0.04	8.62	7.80
Wrong	41.40 ± 0.04	8.02	7.35	22.64 ± 0.03	17.29	16.48	18.76 ± 0.03	4.87	4.40
Mlb	168.24 ± 0.08	32.60	29.86	19.71 ± 0.03	15.05	14.34	148.53 ± 0.08	38.56	34.87
No match	158.81 ± 0.08	30.77	28.19	17.81 ± 0.03	13.60	12.96	141.00 ± 0.07	36.61	33.10
Leptonic	45.70 ± 0.04	8.85	8.11	3.70 ± 0.01	2.83	2.69	41.99 ± 0.04	10.90	9.86
Hadronic	0.52 ± 0.01	0.10	0.09	0.10 ± 0.00	0.08	0.07	0.42 ± 0.01	0.11	0.10
W/Z	1.41 ± 0.01	0.27	0.25	0.16 ± 0.00	0.12	0.11	1.25 ± 0.01	0.33	0.29
Single top	26.75 ± 0.05		4.75	4.32 ± 0.02		3.14	22.43 ± 0.04		5.27
Multijet	9.55 ± 0.28		1.69	1.23 ± 0.12		0.90	8.32 ± 0.25		1.95
W+jets	8.71 ± 0.06		1.55	0.67 ± 0.02		0.49	8.04 ± 0.06		1.89
DY+jets	1.93 ± 0.06		0.34	0.19 ± 0.01		0.14	1.74 ± 0.06		0.41
WW/WZ/ZZ	0.35 ± 0.01		0.06	0.05 ± 0.00		0.03	0.31 ± 0.01		0.07
Simulation total	563.39 ± 0.33		100.00	137.38 ± 0.14		100.00	426.01 ± 0.30		100.00
Data	537.85 ± 0.73		95.47	126.69 ± 0.36		92.22	411.15 ± 0.64		96.51

Table 6.4: Legacy 2018 muon event yields. Simulation uncertainties include statistical and cross-section uncertainties. For data, the Poisson uncertainty is given. Fractions presented w.r.t. $t\bar{t}$ only (left) and full simulation (right).

	Full P_{gof}			$P_{\text{gof}} \geq 0.2$			$P_{\text{gof}} < 0.2$		
Sample	Events [k]	Fraction of Sim. [%]		Events [k]	Fraction of Sim. [%]		Events [k]	Fraction of Sim. [%]	
$t\bar{t}$ total	820.49 ± 0.19	100.00	91.51	208.01 ± 0.09	100.00	95.24	612.48 ± 0.16	100.00	90.31
Correct	157.19 ± 0.08	19.16	17.53	104.89 ± 0.07	50.42	48.02	52.30 ± 0.05	8.54	7.71
Wrong	67.11 ± 0.05	8.18	7.48	36.56 ± 0.04	17.57	16.74	30.55 ± 0.04	4.99	4.50
Mlb	265.56 ± 0.11	32.37	29.62	31.41 ± 0.04	15.10	14.38	234.15 ± 0.10	38.23	34.52
No match	254.95 ± 0.11	31.07	28.43	28.81 ± 0.04	13.85	13.19	226.13 ± 0.10	36.92	33.34
Leptonic	72.49 ± 0.05	8.83	8.08	5.91 ± 0.01	2.84	2.71	66.57 ± 0.05	10.87	9.82
Hadronic	0.93 ± 0.01	0.11	0.10	0.18 ± 0.00	0.08	0.08	0.75 ± 0.01	0.12	0.11
W/Z	2.27 ± 0.01	0.28	0.25	0.25 ± 0.00	0.12	0.12	2.01 ± 0.01	0.33	0.30
Single top	42.18 ± 0.06		4.70	6.88 ± 0.03		3.15	35.30 ± 0.05		5.20
Multijet	15.91 ± 0.48		1.77	1.96 ± 0.24		0.90	13.95 ± 0.41		2.06
W+jets	14.36 ± 0.09		1.60	1.14 ± 0.03		0.52	13.22 ± 0.09		1.95
DY+jets	3.11 ± 0.08		0.35	0.34 ± 0.03		0.16	2.77 ± 0.07		0.41
WW/WZ/ZZ	0.57 ± 0.02		0.06	0.07 ± 0.01		0.03	0.50 ± 0.01		0.07
Simulation total	896.63 ± 0.53		100.00	218.41 ± 0.27		100.00	678.22 ± 0.46		100.00
Data	843.63 ± 0.92		94.09	197.90 ± 0.44		90.61	645.73 ± 0.80		95.21

Table 6.5: Legacy 2017 electron event yields. Simulation uncertainties include statistical and cross-section uncertainties. For data, the Poisson uncertainty is given. Fractions presented w.r.t. $t\bar{t}$ only (left) and full simulation (right).

	Full P_{gof}			$P_{\text{gof}} \geq 0.2$			$P_{\text{gof}} < 0.2$		
Sample	Events [k]	Fraction of Sim. [%]		Events [k]	Fraction of Sim. [%]		Events [k]	Fraction of Sim. [%]	
$t\bar{t}$ total	334.91 ± 0.11	100.00	91.77	82.97 ± 0.06	100.00	95.58	251.94 ± 0.10	100.00	90.58
Correct	65.27 ± 0.05	19.49	17.88	43.27 ± 0.04	52.15	49.85	22.00 ± 0.03	8.73	7.91
Wrong	26.11 ± 0.03	7.80	7.15	14.04 ± 0.02	16.92	16.17	12.07 ± 0.02	4.79	4.34
Mlb	109.93 ± 0.06	32.82	30.12	12.35 ± 0.02	14.89	14.23	97.58 ± 0.06	38.73	35.08
No match	102.64 ± 0.06	30.65	28.12	10.98 ± 0.02	13.23	12.65	91.66 ± 0.06	36.38	32.95
Leptonic	29.78 ± 0.03	8.89	8.16	2.19 ± 0.01	2.64	2.52	27.59 ± 0.03	10.95	9.92
Hadronic	0.22 ± 0.00	0.07	0.06	0.04 ± 0.00	0.04	0.04	0.19 ± 0.00	0.07	0.07
W/Z	0.96 ± 0.01	0.29	0.26	0.10 ± 0.00	0.12	0.12	0.85 ± 0.01	0.34	0.31
Single top	16.81 ± 0.04		4.61	2.68 ± 0.02		3.09	14.13 ± 0.03		5.08
Multijet	5.60 ± 0.15		1.53	0.53 ± 0.05		0.61	5.07 ± 0.14		1.82
W+jets	5.72 ± 0.05		1.57	0.43 ± 0.02		0.49	5.29 ± 0.05		1.90
DY+jets	1.66 ± 0.04		0.46	0.17 ± 0.02		0.20	1.49 ± 0.04		0.53
WW/WZ/ZZ	0.25 ± 0.01		0.07	0.03 ± 0.00		0.03	0.22 ± 0.01		0.08
Simulation total	364.94 ± 0.20		100.00	86.81 ± 0.08		100.00	278.13 ± 0.18		100.00
Data	336.97 ± 0.58		92.34	77.10 ± 0.28		88.82	259.87 ± 0.51		93.43

Table 6.6: Legacy 2018 electron event yields. Simulation uncertainties include statistical and cross-section uncertainties. For data, the Poisson uncertainty is given. Fractions presented w.r.t. $t\bar{t}$ only (left) and full simulation (right).

	Full P_{gof}			$P_{\text{gof}} \geq 0.2$			$P_{\text{gof}} < 0.2$		
Sample	Events [k]	Fraction of Sim. [%]		Events [k]	Fraction of Sim. [%]		Events [k]	Fraction of Sim. [%]	
$t\bar{t}$ total	499.84 ± 0.14	100.00	91.64	123.14 ± 0.07	100.00	95.61	376.69 ± 0.12	100.00	90.42
Correct	96.70 ± 0.06	19.35	17.73	63.95 ± 0.05	51.93	49.65	32.76 ± 0.04	8.70	7.86
Wrong	39.50 ± 0.04	7.90	7.24	20.97 ± 0.03	17.03	16.28	18.53 ± 0.03	4.92	4.45
Mlb	162.96 ± 0.08	32.60	29.88	18.32 ± 0.03	14.87	14.22	144.64 ± 0.08	38.40	34.72
No match	154.50 ± 0.08	30.91	28.33	16.47 ± 0.03	13.38	12.79	138.03 ± 0.08	36.64	33.13
Leptonic	44.36 ± 0.04	8.87	8.13	3.23 ± 0.01	2.62	2.51	41.13 ± 0.04	10.92	9.87
Hadronic	0.37 ± 0.00	0.07	0.07	0.06 ± 0.00	0.05	0.04	0.31 ± 0.00	0.08	0.07
W/Z	1.45 ± 0.01	0.29	0.27	0.15 ± 0.00	0.13	0.12	1.30 ± 0.01	0.34	0.31
Single top	25.17 ± 0.05		4.61	3.95 ± 0.02		3.07	21.22 ± 0.04		5.09
Multijet	8.09 ± 0.19		1.48	0.71 ± 0.06		0.55	7.38 ± 0.18		1.77
W+jets	9.30 ± 0.07		1.71	0.71 ± 0.03		0.55	8.60 ± 0.07		2.06
DY+jets	2.62 ± 0.05		0.48	0.24 ± 0.01		0.19	2.38 ± 0.05		0.57
WW/WZ/ZZ	0.39 ± 0.01		0.07	0.05 ± 0.00		0.04	0.34 ± 0.01		0.08
Simulation total	545.41 ± 0.25		100.00	128.80 ± 0.10		100.00	416.61 ± 0.23		100.00
Data	513.89 ± 0.72		94.22	115.09 ± 0.34		89.36	398.80 ± 0.63		95.72

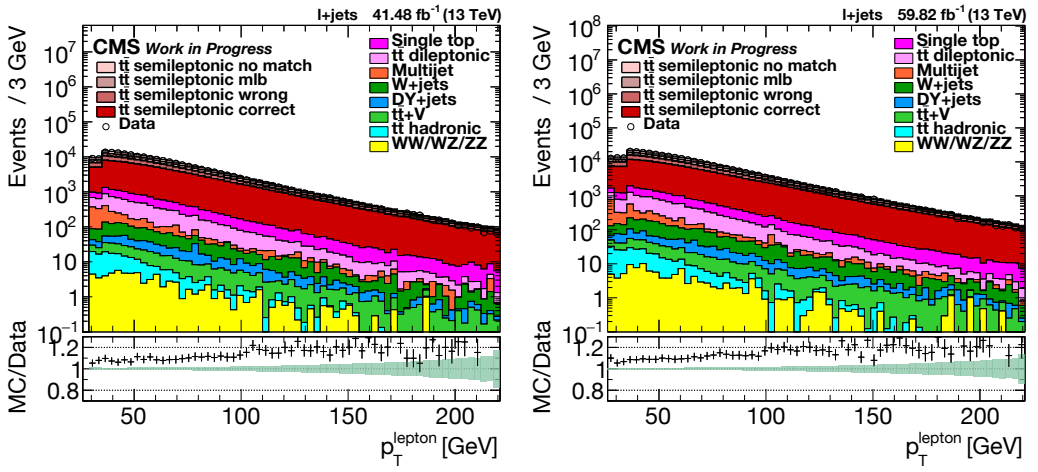

Figure 6.17: Lepton p_T in $P_{\text{gof}} \geq 0.2$, $mu17 + e17$ (left), $mu18 + e18$ (right). Simulation to data ratio error bars consist of statistical and scaling uncertainty in simulation, whereas statistical uncertainty in data is displayed as a green band.

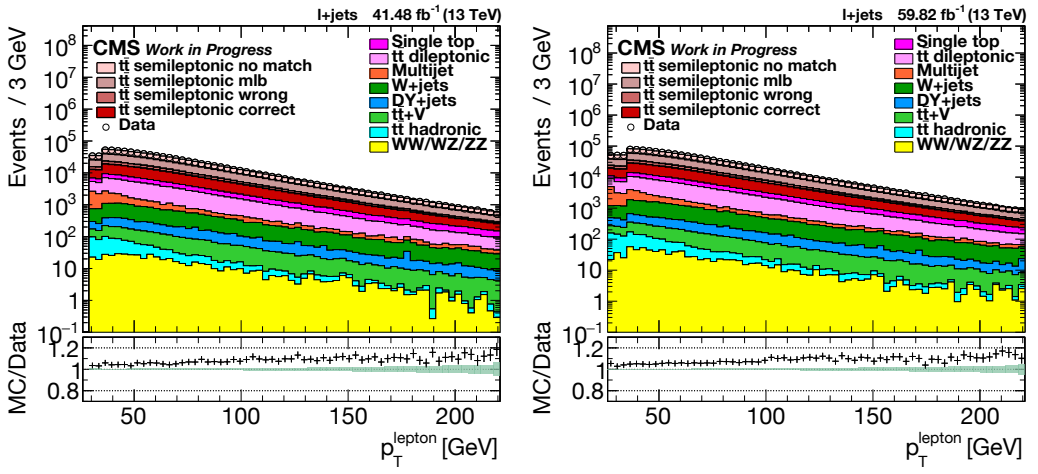

Figure 6.18: Lepton p_T – same as Fig. 6.2 shown for comparison, $mu17 + e17$ (left), $mu18 + e18$ (right). Simulation to data ratio error bars consist of statistical and scaling uncertainty in simulation, whereas statistical uncertainty in data is displayed as a green band.

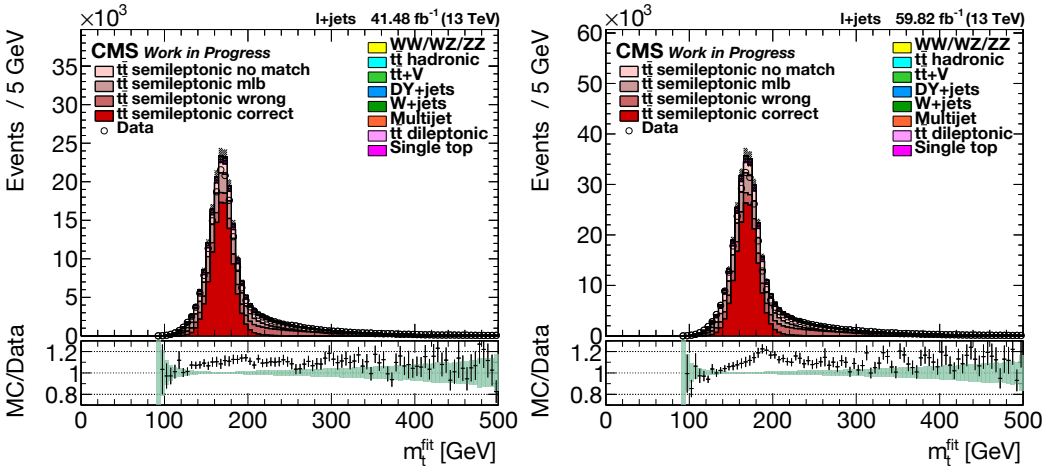

Figure 6.19: m_t^{fit} in $P_{\text{gof}} \geq 0.2$, $mu17 + e17$ (left), $mu18 + e18$ (right). Simulation to data ratio error bars consist of statistical and scaling uncertainty in simulation, whereas statistical uncertainty in data is displayed as a green band.

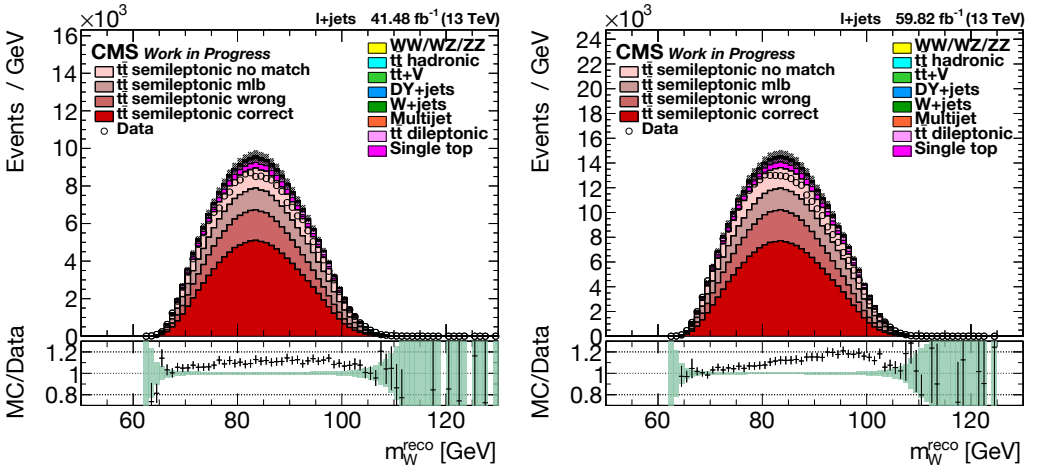

Figure 6.20: m_W^{reco} in $P_{\text{gof}} \geq 0.2$, $mu17 + e17$ (left), $mu18 + e18$ (right). Simulation to data ratio error bars consist of statistical and scaling uncertainty in simulation, whereas statistical uncertainty in data is displayed as a green band.

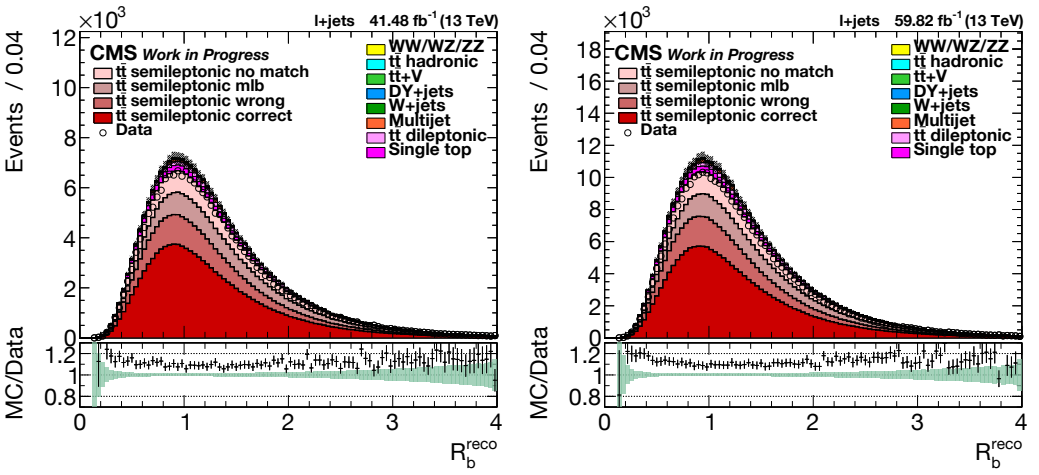

Figure 6.21: R_{bq}^{reco} in $P_{\text{gof}} \geq 0.2$, $mu17 + e17$ (left), $mu18 + e18$ (right). Simulation to data ratio error bars consist of statistical and scaling uncertainty in simulation, whereas statistical uncertainty in data is displayed as a green band.

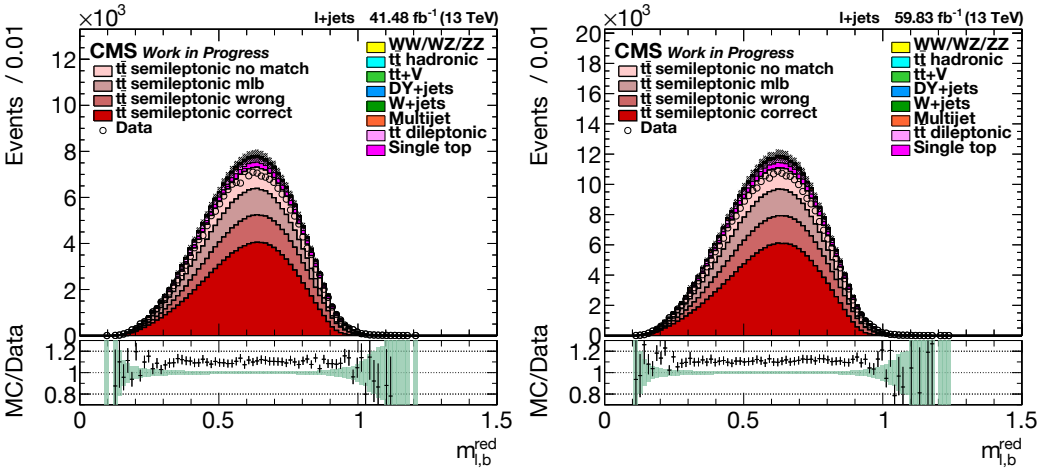

Figure 6.22: $m_{lb}^{red,reco}$ in $P_{\text{gof}} \geq 0.2$, $mu17 + e17$ (left), $mu18 + e18$ (right). Simulation to data ratio error bars consist of statistical and scaling uncertainty in simulation, whereas statistical uncertainty in data is displayed as a green band.

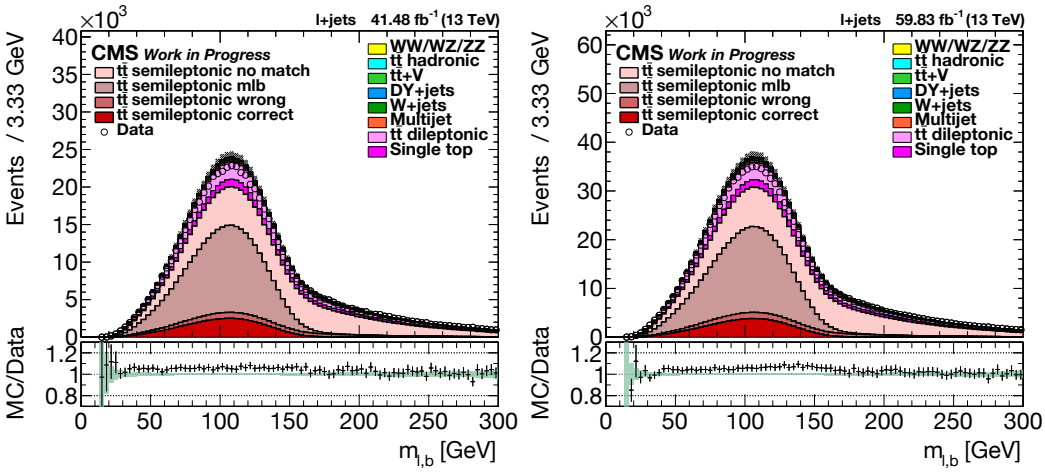

Figure 6.23: m_{lb}^{reco} in $P_{\text{gof}} < 0.2$, $mu17 + e17$ (left), $mu18 + e18$ (right). Simulation to data ratio error bars consist of statistical and scaling uncertainty in simulation, whereas statistical uncertainty in data is displayed as a green band.

6.3 Systematic Uncertainties

In this section, the relevant uncertainties considered in the present analysis are reviewed. The recommendations followed here on systematic uncertainties are mainly distributed internally at the CMS, and hence in the following only limited references can be provided. The recommendations follow a common pattern through the whole Run 2, and the year-wise systematic uncertainties can be mutually correlated or uncorrelated. The uncertainties are split into those with an experimental nature, and those that originate in modeling/theory.

The systematic uncertainties are generally viewed as one standard deviation up and down variations of various uncertain aspects in calibration and modeling. These variations are parametrized by continuous **nuisance parameters** θ. The normalization is such that $\theta = -1$ corresponds to the one standard deviation down variation and $\theta = +1$ to the one standard deviation up variation, while $\theta = 0$ is valid for the nominal/central simulated samples. The systematic nuisances are designed to express the fact that the nominal simulation might not completely agree with data in some aspects, and e.g. the nuisance parameter(s) for the event yield normalization could differ from zero in data.

6.3.1 Experimental Uncertainties

Experimental uncertainties consist mainly of energy scale variations and scale factors. Of these, the jet energy scale variations are the most important in this work.

Jet Energy Corrections

One sigma up and down variations are performed for the standard set of Run 2 JEC uncertainties, available in [98]. An exception is made by the single flavor JEC uncertainty (FlavorQCD), which is replaced by the full set of flavor-dependent variations. This includes e.g. separate b jet JEC uncertainties. The flavor variations are reviewed under modeling uncertainties. The full recommended JEC set (without FlavorQCD) includes 26 JEC variations, dependent on the jet p_T and η. These include e.g. FSR and pileup variations. Between the 2017 and 2018 datasets, a simple correlation scheme is used, i.e. JEC nuisances are treated either as fully correlated or fully uncorrelated between the years.

Jet Energy Resolution

Variations of one standard deviation up and down are applied on Jet Energy Resolution (JER) smearing scale factors. JER smearing scales are not considered to be jet energy corrections, even if the variation of the JER scale factors has a similar effect as the variation of JECs. The variations are performed with the **hybrid method** for JER smearing, which combines the scaling method and the stochastic method, as explained in Subsection 4.3.1. The seed used for stochastic smearing is uniquely deterministic for each event, so that the variations remain correlated. The JER nuisance is considered to be fully uncorrelated between years.

Missing Transverse Momentum ($\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$)

All changes in jet and lepton energy scales – including also the effect of jet smearing - are propagated directly into the value of $p_{\mathrm{T}}^{\mathrm{miss}}$ and the associated $\sum E_t$. This is not reported separately, and the $p_{\mathrm{T}}^{\mathrm{miss}}$ effects are included in these energy scaling systematic uncertainties.

In addition, a specific $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ uncertainty with a one standard deviation up and down variation for the unclustered energy is applied. The nuisance variations have an impact on both the value of $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ and the associated $\sum E_t$. Of these, the value of $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is used as a starting point for the neutrino solution by HitFit, and $\sum E_t$ is utilized for determining $\vec{k}_T$ resolution. This uncertainty is treated as fully uncorrelated along the years.

B-tagging Scale Factors

DeepJet B-tagging nuisances are split into a correlated and an uncorrelated component for b/c-jets and separately for light jets. Thus, there are two correlated nuisances between the run years, and two uncorrelated nuisances per each run year.

Pileup

A $\pm 4.6\%$ uncertainty is assigned to the 69.2 mb minimum-bias cross-section, as described in Ref. [116]. This imposes shifts to the target distribution of pileup vertices observed in data. The pileup uncertainty is considered fully correlated over the years.

Electron Energy Scales

Generally, electron energy scales are expected to contain a systematic and a statistical uncertainty component. However, in the Run 2 legacy production there is only one generic energy scale with one standard deviation up and down variations available, and it is taken to be fully uncorrelated across the years.

Variations of one standard deviation are also applied on the electron energy resolution. This systematic uncertainty is considered to be fully correlated between the years.

Muon Energy Scales

Muon energy scale uncertainties are provided by the Rochester group [109]. These consist of the (fully uncorrelated) statistical one standard deviation up and down variations and four separate systematic uncertainty variations (fully correlated between years).

Electron Scale Factors

For the electron reconstruction and ID scale factors, one standard deviation up and down weight variations are applied on the full event weight. The uncertainties are taken to be fully correlated over the years.

The scale factors are derived in the DY topology from the relevant efficiencies. An additional 1 % scale factor uncertainty is required for electrons, based on the efficiency extrapolation uncertainty from the DY topology into the $\mathrm{t\bar{t}}$ topology. A large proportion of this uncertainty originates from isolation properties, and hence the uncertainty is summed quadratically to the ID scale factor variations (that also contain isolation).

The uncertainties of the privately produced trigger scale factors are found in the trigger scale factor derivation process. These include variations e.g. in fit function types and p_T thresholds. The trigger nuisances are taken to be fully uncorrelated over the years.

Muon Scale Factors

The one standard deviation up and down weight variations for muon Trigger, ID and Isolation are taken into account. For ID and Isolation, a statistical and a systematic

uncertainty component is given. The systematic uncertainty component is to be taken fully correlated and the statistical as fully uncorrelated between the years.

The scale factors are derived in the DY topology from the relevant efficiencies. An additional 0.5 % scale factor uncertainty is required for muons, based on the efficiency extrapolation uncertainty from the DY topology into the $t\bar{t}$ topology. A large proportion of this uncertainty originates from isolation properties, so the uncertainty is appended quadratically to the Isolation scale factor variations.

For the triggers, only a single systematic one standard deviation up and down scale factor variation is considered. Different triggers are used in 2017 and 2018, so they are fully uncorrelated.

L1 ECAL + Muon Prefiring

Variations of one standard deviation up and down are applied with respect to the L1 prefire [117, 118] central weight. The uncertainties are treated as uncorrelated between the two years, since the uncertainty in 2018 only considers muon prefiring. The muon prefire effects are several orders of magnitude smaller than those related to ECAL.

HEP17 (2017) and HEM15-16 (2018)

The jet veto maps are used to veto events where one of the four analysis jets is in a problematic detector region. This is done in data and simulation in a similar manner, so no additional systematic uncertainties are required. The Run 2 legacy JEC derivation follows these same vetos, making the JECs optimal for an analysis following the veto maps. The total loss induced by the veto maps is maximally 15 % of the final collision event statistics, which is motivated for a precision measurement.

Following Poisson statistics, the relative statistical uncertainties are generally proportional to $1/\sqrt{N}$, where N is the number of events. Thus, the statistical impact of the veto cut is only 8 %. The statistical uncertainties play a relatively small role in the current m_t measurements, whereas including the poorly reconstructed detector regions could be more troublesome for the measurement.

Luminosity

Table 6.7: Luminosity uncertainty correlation table for 2017–2018.

Source	2017 [%]	2018 [%]
Uncorrelated 2017	2.0	-
Uncorrelated 2018	-	1.5
Correlated 2017-18	0.6	0.2
Correlated 2016-18	0.9	2.0

The luminosity uncertainties (in percents) are given in Table 6.7. This includes a correlation scheme between the two years. The presented uncorrelated components are used as simple event yield scale uncertainties for each of the years. The correlated components on the other hand are utilized as common event yield scale uncertainties for the two years.

Simulation Normalization

Theoretical and generator-based cross-sections are used for scaling the weights of the simulated samples - depending on whether or not a theoretical estimate is available. Both the theoretical and generator-based uncertainties are typically less than 5% of the whole cross section. A much greater level of uncertainty is presented for the samples with large cross-sections, of which only a small proportion pass the full selection of four prompt jets and a charged lepton. For such minor backgrounds, the selection efficiencies can notably vary between data and simulation, inflating the scale uncertainties. This can be expected e.g. for W+Jets and DY+Jets, where CP5 Pythia 8 tune event yields for high jet counts are known to be unstable [78].

- The three main $t\bar{t}$ channels are given a common uncertainty of $+4.84\,\%$ and $-5.49\,\%$ based on the theoretical combination of α_S, PDF and scaling uncertainties
- The $t\bar{t}$ dileptonic channel is given an extra $5\,\%$ uncertainty for selection efficiency, as this is one of the two most important backgrounds (the same could be done for the $t\bar{t}$ hadronic channel, but it has such a small event yield that the effect is negligible)
- Single top: NLO samples and theoretical predictions $\rightarrow$ multiply the prediction by two to account for the uncertainties in selection efficiency; this leads to $\pm 6\,\%$ for tW and $\pm 8\,\%$ for the other channels
- W+Jets, DY+Jets, $t\bar{t}$ +W/Z and WW/WZ/ZZ: $\pm 10\,\%$ uncertainties based on the simulated cross-section uncertainty and an additional selection efficiency uncertainty
- Multijet: $^{+100\,\%}_{-50\,\%}$ uncertainty, separate nuisances for the electron and muon channels.

Separating the Multijet nuisance into the electron and muon channels is motivated by the fact that different simulations are used for these backgrounds. The modeling in electron-enriched Multijet samples is distinct from that in muon-enriched Multijet samples: muons mainly come from semileptonic quark decays and pair production, while Multijet event electrons are in addition found in fake electron signals. This issue remains even considering the tight electron ID condition. Fake electrons are mostly found at low p_T, so the higher p_T cut for electrons than for muons helps in dealing with these challenges.

Similar simulation samples are used in both years and therefore all of the scaling uncertainties for a single sample are considered correlated across the years.

6.3.2 Modeling Uncertainties

Modeling uncertainties are either fully theoretical, or related to the implementation of the simulations. In the latter category, uncertainties arise either from parametrizations, or the assumptions behind simulation models. As a rule, these nuisances are fully correlated between 2017 and 2018, as mostly the same simulation settings are used in both datasets. The few exceptions to this rule are mentioned separately.

The modeling variations are implemented employing a variety of methods:

- Global variations in event weights,
- Event-wise variations in event weights,
- Jet energy scale variations,
- Separate signal samples with separate settings.

Currently, the methods utilizing event weight variations are highly preferred. These provide full correlation of statistics between the central sample and variations – except for the slight statistical noise in the event weights. That is, there is a small degree of statistical uncertainty in the values of the variation weights.

In this analysis, the dominant modeling uncertainties are those in jet energy scales and FSR variations. The FSR scales have a significant impact on the jet energy scales - but in addition they also vary e.g. the degree of collimation in an average jet. The FSR uncertainties in most of the historical m_t measurements have not been as important as in the present analysis. This follows from the fact that the full FSR split into flavors has become available only recently.

Flavor-dependent JECs

The flavor-dependent JECs have been estimated during Run 1 through a comparison between Pythia 6 Z2 tune and Herwig++2.3. These have been sanity-checked also for Run 2 use. The differences found in the results are based on the modeling differences between Pythia and Herwig in uds/c/b quark and gluon jet responses. The flavor categories are FlavorPureBottom (only b), FlavorPureQuark (only uds), FlavorPureCharm (only c) and FlavorPureGluon (only gluons). The one standard deviation up and down variations are applied in simulation according to the simulation truth information on the jet flavor. In this category, the b jet uncertainties turn out to be the most significant ones.

Parton Showers (ISR/FSR)

The current Pythia 8 versions provide a fully flavor-split scheme for Initial State Radiation (ISR) and Final State Radiation (FSR) [127, 128]. Instead of a single ISR/FSR up/down variation, separate variations for b/t quark gluon radiation, udsc quark gluon radiation and gluon to quark/gluon pair processes can be performed. For each ISR/FSR variation mode, mutually decorrelated variations both in shower renormalization scales and a non-singular term (CNS) are performed.

In this analysis, ISR is a minor systematic uncertainty and hence only the combined variation is utilized. For FSR, the fully split variation scheme is employed. ISR and FSR renormalization scales are varied by factors of 0.5/2.0, which are interpreted as one standard deviation up and down variations. These variations are analogous to variations in the strong coupling constant α_S for the various types of ISR and FSR. The scale down variation corresponds to an α_S up variation. For CNS, the one standard deviation up and down variations are interpreted to correspond to the CNS parameter values $+2/-2$. The meaning of these variations is further explained in Ref. [128]. The CNS variations turn out to be less important than the α_S (shower renormalization scale) variations.

B Fragmentation

B Fragmentation tuning is performed by fitting the x_B parameter in simulation to measurements from the LEP experiments DELPHI [129] and ALEPH [130] to the current simulation tune. This parameter estimates the fraction of energy carried by a b hadron in a b jet.

The fragmentation fits have been performed both for the Bowler-Lund (BL) and the Peterson fragmentation models. For both of these models, a central value and up and down variations, corresponding to LEP results are given. To not end up double-counting the LEP uncertainties, the relative variations of BL and the central value of Peterson fragmentation are used as nuisances. All the nuisances are applied as weights.

Semileptonic B hadron decays

The branching ratio uncertainty in B hadron semi-leptonic decays is considered according to [62]. In semi-leptonic decays a fraction of the b quark momentum is lost to a neutrino or neutrinos. If the fraction of semi-leptonic b quark decays changes, the average relationship between a b jet and a b quark changes – importantly affecting this measurement. The uncertainty is applied through event weights.

Parton Distribution Functions

The central Parton Distribution Function (PDF) is NNPDF31_NNLO [81]. Systematic uncertainty variations to it are performed using event weights. The default

α_S up and down variations provided within this PDF are applied. Systematic uncertainties are also provided for the alternative PDF choices, CT14NNLO [131] and MMHT2014NNLO68CL [132].

The PDF eigenvectors are not an important systematic uncertainty in this analysis, and hence the differences presented by them are summed up quadratically, providing only a single up and down weight variation in each event.

Factorization and Renormalization Scales

The factorization and renormalization scales μ_F and μ_R are varied by a scale factor of 0.5 and 2.0. The variations are applied using event weights, and considered as two separate nuisances.

Matrix Element – Parton Shower Matching Scale

The $\mathrm{t\bar{t}}$ signal sample MEs are produced with POWHEG, which are then matched to the Pythia 8 PS production. The scale parameter h_{damp} controls the matching between POWHEG and Pythia 8 through the damping function $h_{\mathrm{damp}}^2/\left(h_{\mathrm{damp}}^2 + p_T^2\right)$. To evaluate the uncertainty in this process, the value of h_{damp} is varied from $1.379m_t$ by $+0.926m_t$ and $-0.5052m_t$ [78]. These variations correspond approximately to 150 GeV and 400 GeV.

Underlying Event

There is a degree of uncertainty in the CMS UE Tune parameters [78]. To evaluate this, UE Tune up and down variations are provided in separate $\mathrm{t\bar{t}}$ signal samples.

Color Reconnection

There is no certainty of the universally most correct Color Reconnection (CR) model [133]. In this analysis, a **QCDBased** and a **GluonMove** model is compared to the default model in Pythia 8 [134]. These variations are provided in separate $\mathrm{t\bar{t}}$ signal samples, indicating fully uncorrelated statistical noise. Both of the alternative CR models are one-sided systematic uncertainties. This is in contrast to the more typical two-sided uncertainties with one standard deviation up and down variations.

Early Resonance Decays

In the default simulation settings used at CMS, Early Resonance Decays (ERD) are turned off. A separately simulated $\mathrm{t\bar{t}}$ signal sample with ERD on is utilized to evaluate this uncertainty. ERD variation is a one-sided systematic uncertainty.

Top p_T Spectrum

NNLO effects are relevant for the $t\bar{t}$ p_T spectrum [119], meaning that the NLO simulation is insufficient to catch the correct spectrum shape. Hence, a weight variation, based on CMS data [120–123], is applied for correcting the $t\bar{t}$ spectrum. All $t\bar{t}$ events receive a variation weight based on the parton level top p_T values ($p_T^{gen,t}$):

$$\exp\left(0.0615 - 0.0005 \times \frac{p_T^{gen,t1} + p_T^{gen,t2}}{2}\right). \tag{6.10}$$

Instead of using the weights for the central simulation, they are used as a systematic uncertainty. Since the uncertainties are implemented through nuisance parameters, the likelihood fit is able to determine the best p_T spectrum itself.

6.4 Top Mass Fit

The value of the top quark mass is extracted using a likelihood fit, as has been done in most of the direct m_t measurements at the LHC. In the first Run 2 m_t measurement the fit was based on the ideogram method [20], which employs the m_t^{fit} and m_W^{reco} distributions. These are by definition decorrelated, as $m_W = 80.4\,\text{GeV}$ for m_t^{fit}. Motivated by this decorrelation, the distributions of these two observables can be interpreted as two separate analyses. Following the terminology of the previous analyses [20, 22], each analysis observable (m_t^{fit}, m_W^{reco} etc.) is referred to as a **dimension**. E.g. by combining the m_t^{fit} and m_W^{reco} dimensions, one ends up with a 2D analysis.

In the Ideogram method, the likelihood function was constructed from intricate parametrized fits. This technique was prone to overfitting and fit instability. The ATLAS Run 1 measurements employed similar techniques in a 3D fit [21], where the third dimension R_{bq} was introduced for the *in situ* measurement of the b jet energy scale. As the definition of R_{bq} is a ratio of p_Ts, it is decorrelated from jet scales. Furthermore, the jet p_T values scale uniformly according to the parent particle mass, so the jet p_T ratio R_{bq} is not strongly dependent on m_t.

In the legacy 2016 analysis [22], the concept of a profile likelihood fit was first presented in the lepton+jets m_t analysis. Here, m_t^{fit} still employs a parametrized fit function, but the other dimensions are modeled as binned probability distributions. On top of the two dimensions used in the previous CMS analyses, R_{bq} was included as a natural addition. Moreover, the analysis variables m_{lb} and m_{lb}^{red} were introduced into the semileptonic $t\bar{t}$ event topology. Earlier, the m_{lb} variable has been in use in dileptonic $t\bar{t}$ analyses [135]. Pure m_{lb} is correlated to m_t^{fit}, so only m_{lb}^{red} can be utilized in $P_{gof} \geq 0.2$.

Pure m_{lb} is a sensible choice for the selection $P_{gof} < 0.2$, which no previous measurements have used. This is also demonstrated in Fig. 6.23: the low P_{gof} values are mostly caused by the inaccurate reconstruction of m_W^{reco}, leaving the

leptonic branch unaffected. In addition to the correctly reconstructed lepton+b combinations from the semileptonic $t\bar{t}$ signal, correct m_{lb} reconstructions are also found in the dileptonic and single top backgrounds.

6.4.1 A Binned Profile Likelihood Fit

In this iteration of the analysis, the utilization of profile likelihoods is continued. This is motivated by the ever increasing amount of systematic uncertainties, which the older techniques are no longer able to handle.

The same five analysis dimensions are used as in the legacy 2016 analysis: m_t^{fit}, m_W^{reco}, m_{lb}^{red} and R_{bq} in $P_{\text{gof}} \geq 0.2$, and m_{lb} in $P_{\text{gof}} < 0.2$. The choice of these has already been tested and motivated in the lepton+jets topology.

Despite using similar methods, this analysis diverges slightly from the analysis strategy of the legacy 2016 analysis, and utilizes the Combine tool [136]. The main motivation for this is the way in which Combine facilitates run year combinations. Performing the combination as a part of the analysis on a commonly accepted tool saves a lot of trouble in the latter parts of the analysis. Moreover, future measurements can be appended to the present one via Combine in a simple manner.

Combine leads the analyst to making a few analysis choices in a different manner compared to the legacy 2016 analysis:

- **All the dimensions are handled with bins**, including m_t^{fit} that is handled with a fit function in the legacy 2016 analysis.

- Compared to a parametrized fit function, the binned strategy is better at taking into account insufficient numbers of events e.g. in the Multijet simulation: a fit function can ignore single noisy bins, while the binned strategy correctly propagates all bin content (Poisson) uncertainties to the analysis.

- **The legacy 2016 analysis interprets the dimensions as normalized probability distributions**. Hence the full event count in simulation is scaled to match that in data, also for systematic uncertainty variations.

- In contrast, **Combine also considers the differences in the event yields** between data and simulation. This includes systematic uncertainty variations, for which the even yields can change notably from the default simulation following from the changes in cut acceptances.

The differences in event yield handling can turn out to be important, as this is directly connected to cut acceptances. The requirement of four 30 GeV jets (two of which are b jets) AND exactly one charged lepton is a strong phase space constraint even in the $t\bar{t}$ topology. The FSR variations e.g. keep the total event yields (before cuts) the same, but the cut acceptances vary notably. This information is lost, if the number of simulated events is always normalized to match data.

A practical example on the scaling effects in the case of the b jet FSR and energy scales is given in subsection 6.4.6. It is shown that if it were not for the differences in event yield scales, the b jet FSR, b jet JEC and semileptonic b fragmentation would have nearly interchangeable variations. Without the scaling information, it can be difficult for the fit to choose, which of the systematic nuisances is the most appropriate one for explaining an observed variation.

6.4.2 Selections on P_{gof}

A small change in the analysis dimensions w.r.t. the legacy 2016 analysis is also considered. This is about setting a lower limit cut on P_{gof}/χ^2 for the m_{lb} variable in $P_{\text{gof}} < 0.2$. If only the events in $0.001 < P_{\text{gof}} < 0.2$ are included, this variable collects a similar amount of data as the other dimensions. Such a cut was initially used in the analysis to make all the dimensions equal and to focus on the phase-space with the best signal yield.

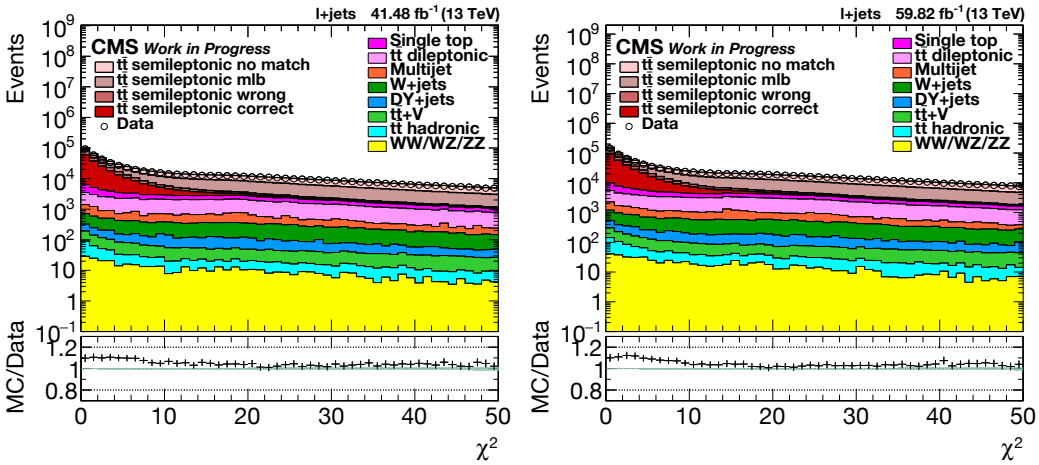

Figure 6.24: The kinematic fit χ^2 dependence on a logarithmic y-scale. $mu17 + e17$ (left), $mu18 + e18$ (right). A zoomed-in version of Fig. 6.15.

Another point of view on the P_{gof} cuts is provided by the notable χ^2 dependence in the event yields, first depicted in Fig. 6.15 and Fig. 6.16. To further highlight the observed effects, an intermediate χ^2 range is displayed in Fig. 6.24. The signal region $P_{\text{gof}} \geq 0.2$ ($\chi^2 \leq 3.2189$) stands on a pedestal, where the simulated event yield is approximately 10 % higher than that in data. If the region $P_{\text{gof}} < 0.2$ was excluded from the analysis, a likelihood fit would explain this discrepancy only using simple scale variations. In contrast, if $P_{\text{gof}} < 0.2$ is used as an anchor for the χ^2 dependence of the yields, the fit will be more likely to find the correct nuisance parameters for explaining the χ^2 dependence.

With the above arguments, it is better to rather include more events for m_{lb}

in $P_{\text{gof}} < 0.2$. The full $P_{\text{gof}} < 0.2$ phase-space contains more than three times the number of events in any of the other dimensions. As a remnant of the lower P_{gof} limit for the exclusion of spurious events, a sanity check lower limit cut is placed at $P_{\text{gof}} = 10^{-50}$ ($\chi^2 \approx 230$). This is a sensible choice that only removes an almost negligible fraction of events, as can be quickly reviewed in Fig. 6.15. As an end result, the number of events received from $P_{\text{gof}} \geq 0.2$ compared to that in $P_{\text{gof}} < 0.2$ is in a good balance of 4 : 3.

6.4.3 Mathematical Formulation

The binned profile likelihood analysis operates on a relatively simple likelihood function:

$$\mathcal{L}(\vec{n}) = \prod_{i \in \text{bins}} \mathcal{P}\left(n_i | \sum_{j \in \text{samples}} \left(1 + \kappa_j\right)^{\eta_j} \times \nu_i^j \left(\vec{\theta}, m_t\right) \right) \times \prod_{k \in \text{nuisances}} \mathcal{G}(\theta_k) \times \prod_{j \in \text{samples}} \mathcal{G}(\eta_j) \,. \tag{6.11}$$

Here, n_i is the number of events observed in data in the bin i and ν_i^j is the expected number of events in this bin for the simulated sample j. The probability function $\mathcal{P}(n|\lambda)$ is the Poisson probability for an observed number of events (n) to match the expected number of events (λ). The expected number of events is a sum of ν_i^js over all signal and background samples j. The variables θ_k stand for the systematic uncertainty nuisance parameters. Furthermore, the parameters κ_j are the scaling uncertainties of each simulated sample.

The event yield ν_i^j in the simulated sample j can depend on the systematic uncertainty variations $\vec{\theta}$ and the top quark mass m_t. The active nuisances $\vec{\theta}$ and the dependence on m_t vary from sample to sample (e.g. the Multijet background is not dependent on m_t). The majority of the nuisances (and the m_t dependence) are focused only on the $\mathrm{t\bar{t}}$ samples and the single top background, as the majority of the event yield is explained by them.

The nuisances θ_k are normalized so that $\theta_k = 0$ is the central expectation, while $\theta_k = -1$ and $\theta_k = +1$ are the down and up variations, correspondingly. Furthermore, the nuisances θ_k are constrained by normalized ($\mu = 0$ and $\sigma = 1$) Gaussian functions $\mathcal{G}(\theta_k)$. This is the general method for including systematic uncertainties into a likelihood: it is explicitly stated that the up and down variations shift the likelihood value by one σ.

The dependence on θ_ks is found through vertical bin morphing. This means that the histogram bins must be so wide that the histogram does not significantly move on the x-axis for the variations $\theta_k = \pm 1$. Horizontal movement would necessitate horizontal morphing, which would introduce unnecessary complications. In practice, vertical morphing means that histograms are produced both for the central simulation samples, and their corresponding systematic uncertainty variations. The

parameter θ_k determines the point of interpolation for the nuisance k. In the grand scheme, the variations introduced by all the nuisances θ_k are summed up to provide the current value for the expected event yield in the current bin.

The scaling uncertainties κ_j are minimally based on the cross-section uncertainty, but they can also include e.g. an additional estimated uncertainty on the cut acceptance of a certain signal/background. These uncertainties were more closely discussed in the previous section. The scale uncertainty variations are modeled using the nuisance parameter η_j, which is analogous to the other nuisances θ_k. Similar common nuisance scales arise also from the uncertainty in luminosity, but for simplicity this is not separately shown in Eq. (6.11).

6.4.4 Analysis Histogram Binning

As the analysis utilizes no fit functions, the only remaining choice is that on the binning: the free parameters in the model are the positions of bin edges. To find a robust binning, a review of the well-studied choices and motivations provided in the legacy 2016 analysis is performed:

- The same number of bins is used in all dimensions
- Binning resulting in approximately equal statistics in each bin is used
- Bin count selection: the main motivation is that too many bins leads to numerical instability and too few bins loses the main dependencies
- An even number of bins is utilized: previously six, now eight
- No major differences between the tested bin counts is observed.

It is easy to agree with the first point. This puts the dependencies in all of the dimensions on the same line, and strong counter-arguments for this are hard to come up with.

The question about equistatistical binning is more complicated. The best argument for this question is found through the systematic uncertainty variation samples. It is easy to divide the nuisances into three categories according to the level of correlation of their statistical uncertainties with the central (nominal) simulation. This has been done e.g. in Ref. [137], and we continue in a similar manner:

- **Category 1**: Variations based on weights: statistical uncertainties almost 100% correlated with the nominal simulation, excluding statistical noise introduced by the weights
- **Category 2**: Object energy scale variations:
  - Statistical uncertainties mostly correlated with the nominal simulation
  - This holds best if the energy scale variations are not very large

  - Differences are caused by bin-migration and changes in the cut acceptances

- **Category 3**: Separate simulation samples

  - Statistical uncertainties completely uncorrelated with the nominal sample(s)
  - There are only a few systematic uncertainties like this: hdamp, UE tune, Color Reconnection variations and ERD
  - The situation can only be helped by producing a greater number of simulated events

The question about correlations in the statistical fluctuations between the nominal simulation and the nuisance samples is central for Combine. The Combine tool is able to consider the statistical uncertainties in the nominal simulation using the Barlow-Beeston method [138]. However, the uncorrelated statistical fluctuations in the systematic uncertainty variation samples are not treated properly.

To democratize this issue in a single analysis, the best option is to use a binning where statistics are approximately the same in all bins. If the opposite choice is made, some low-statistics bins can become useless for the category 3 nuisances (dominated by noise). The same reasoning can also hold for some of the most unstable category 2 nuisances. Thus, we will continue with the strategy of equivalent bin statistics, after giving it a more thorough motivation than before. Since there are differences in the kinematic cuts between channels, binning must be derived separately for each year and measurement channel.

The final open question is the number of bins. The legacy 2016 analysis motivation for the low limit for the bin count also works for us: with less than 6 bins per dimension, the ability of the histograms to depict the desired dependencies becomes weak.

The upper bin count limit can be strongly motivated by the vertical histogram morphing, which does not work well if the variation histograms suffer big horizontal shifts. The variations used for m_t dependence are $\pm 1\,\text{GeV}$, meaning that the bin width in the m_t^{fit} histogram should be notably larger than 1 GeV. This limit forces the bin count to be less than 20.

6.4.5 Fine-tuning of the Number of Bins

Continuing with the legacy 2016 analysis motivation, the issue with statistics becomes more pronounced at bin counts close to 20. At the lower limit, a slightly larger bin count than in the legacy 2016 analysis is considered for two reasons. First of all, the simulated statistics in the legacy 2017 and 2018 datasets are larger than in the legacy 2016 datasets, meaning that issues with statistics come up with larger bin counts. Secondly, the dependence on m_t^{fit} is now completely binned, while it was

previously given in a parametrized form. It is imperative that the m_t dependence is caught with sufficiently high granularity, even if it is only depicted as bins.

The above reasoning motivated a closer study of 9-12 bins, i.e. two even-numbered and two odd-numbered bin counts per dimension. A comprehensive study of a number of variations for these bin counts was made by duplicating the variation figures in subsection 6.4.6 and Appendix D for each bin count. Based on the results, there was little difference between the various binnings. Most importantly, the variations with lower bin counts are able to show most of the trends in the higher bin count histograms. A small exception was found in the 9-bin case, where a few corners and kinks are not completely described.

When it comes to even vs. odd binning, there was no strong *a priori* preference. However, it turned out that for odd-numbered binning, the central bin tends to be wasted e.g. for the m_t dependence in m_t^{fit} and quark jet JEC dependence in m_W^{reco}. That is, the up/down variations in the central bin only differ a little from the nominal simulation (and each other). This kind of a problem easily comes up with distributions that have a central peak that is concentrated in a single bin. It is deduced that it is better to use even-numbered binning, where the central peak is distributed between two bins, so that important variations are also seen in the central bin.

Moreover, the poorest quality category 3 nuisances support the use of 10 bins over 12 bins, in order to conserve the small trends that can be observed. As was also mentioned in the legacy 2016 analysis, there is no absolutely strong preference for the exact choice on bin counts, but here an accumulation of small factors is hinting towards the use of **10 bins**.

As a result of the binning choice, each dimension in each channel requires a total of 9 parameters, corresponding to the bin edges. The under- and overflow bins are transferred to the first and last bin, meaning that the upper and lower bin edges do not have an impact on the analysis. The lower edge of the first bin and the upper edge of the last bin are used purely for bookkeeping and potential visualization purposes.

6.4.6 Binning Example for μ17

In Figs. 6.25 and 6.26 the relative variations for the most important systematic uncertainties and the $\pm 1\,\text{GeV}$ m_t variations are presented. This is motivated by the structure of the analysis: if things are not working properly in the histograms, the likelihood fit will end up in problems. Here, the five analysis dimensions are presented from left to right in the histograms in the order m_t^{fit}, m_W, m_{lb}^{red}, R_{bq} (all four in $P_{\text{gof}} \geq 0.2$) and m_{lb} ($P_{\text{gof}} < 0.2$). Binning is performed as presented in the previous subsection: 10 statistically equivalent bins per dimension. The up (red) and down (blue) variations are normalized by the bin yields in nominal (green). Statistical error bars for m_{lb} are smaller than in the other dimensions since there are more than three times the statistics w.r.t. the other bins.

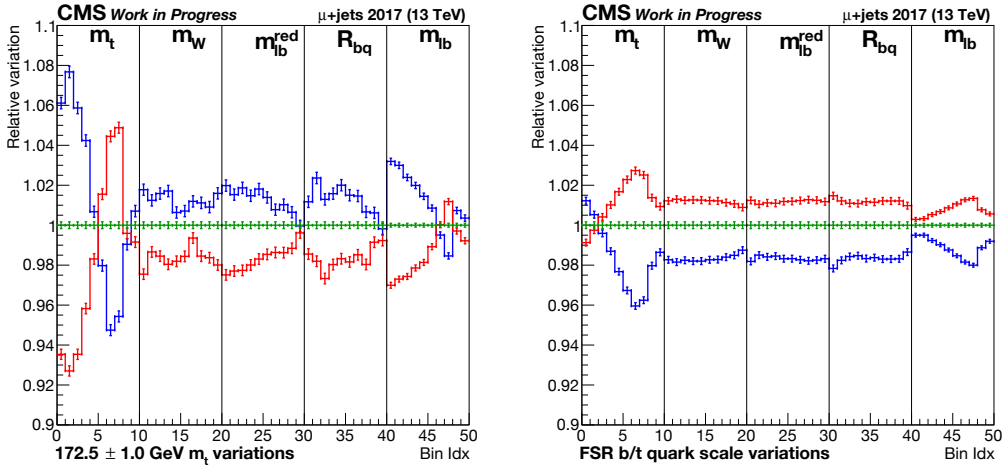

Figure 6.25: Variations of m_t (left) and b/t quark FSR (right) in the five dimensions m_t^{fit}, m_W^{reco}, $m_{lb}^{red,reco}$, R_{bq}^{reco} ($P_{\text{gof}} \geq 0.2$) and m_{lb}^{reco} ($P_{\text{gof}} < 0.2$).

Except for an offset in the total bin yields, the m_t variations and b quark FSR variations in Fig. 6.25 are remarkably similar. For m_t^{fit} and m_{lb} the trends are practically the same, if the yield scaling offset is ignored. In the three central dimensions slight differences are seen.

When the trends are otherwise similar, the yield scaling offsets become highly important for the fit, whereas the magnitudes of bin yield variations are of less interest. The m_t variations have approximately a thrice larger impact on bin yields than those in the b FSR variations. If two variations have the same trends and the same yield scaling offsets, the fit can interpret them as interchangeable.

In the m_t variations the yields are connected to the m_t dependence of the $\mathrm{t\bar{t}}$ cross-section, while in the FSR variations a change in α_S^{FSR} changes cut acceptances. The Pythia 8 weight variations are executed so that when no cuts are imposed, the full event yield remains the same. These two mechanisms present notably different bin yield offsets to the m_t and b FSR variations, allowing Combine to easily distinguish these variations.

Since m_t is a category 3 variable[5] and FSR belongs to category 1[6], the m_t variations include more statistical noise. For m_t the trends in all dimensions are larger than statistical uncertainties, so that belonging into category 3 does not do much harm.

The b jet (flavor) JEC variations and semileptonic b fragmentation variations in Fig. 6.26 also display similar features as depicted in Fig. 6.25. The b jet JEC has a very similar yield scaling offset as the m_t variations, with up and down variations

[5]Statistics in variations are 100 % uncorrelated from nominal.
[6]Statistics in variations are almost 100 % correlated to nominal.

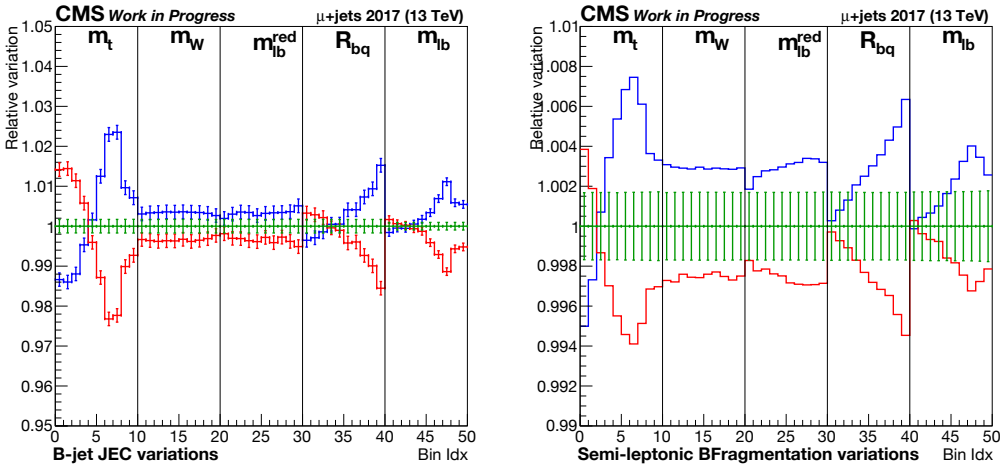

Figure 6.26: Variations of b jet JEC (left) and semileptonic B fragmentation (right) in the five dimensions m_t^{fit}, m_W^{reco}, $m_{lb}^{red,reco}$, R_{bq}^{reco} ($P_{\text{gof}} \geq 0.2$) and m_{lb}^{reco} ($P_{\text{gof}} < 0.2$).

into opposing directions. For a distinction between these variations, the R_{bq} variable has a rescuing effect: for m_t this is flat or slightly concave, whereas for b jet JEC the R_{bq} variations attain a steep slope form. The semileptonic b fragmentation variations display the same sort of trends as seen in b jet JECs, but there is a scaling offset present. Most importantly, the semileptonic variations are relatively small.

To summarize, excluding the yield offset effects and differences in R_{bq}, the following variations resemble each other:

- m_t up (down)
- b/t quark FSR scale up (down) $\equiv$ b/t quark α_S^{FSR} down (up)
- b jet JEC down (up)
- semileptonic b Fragmentation down (up) $\equiv$ less (more) semileptonic b hadron decays

Further examples of bin variations are presented in Appendix D.

6.4.7 Limited Statistics in Simulation

As was underlined above, category 3 nuisances can suffer from the lack of statistics. This issue is the largest when the trends explained by the nuisance variations are smaller than the statistical uncertainties. The impact of the lack of statistical precision was studied using 5000 toy experiments. In each toy experiment, the

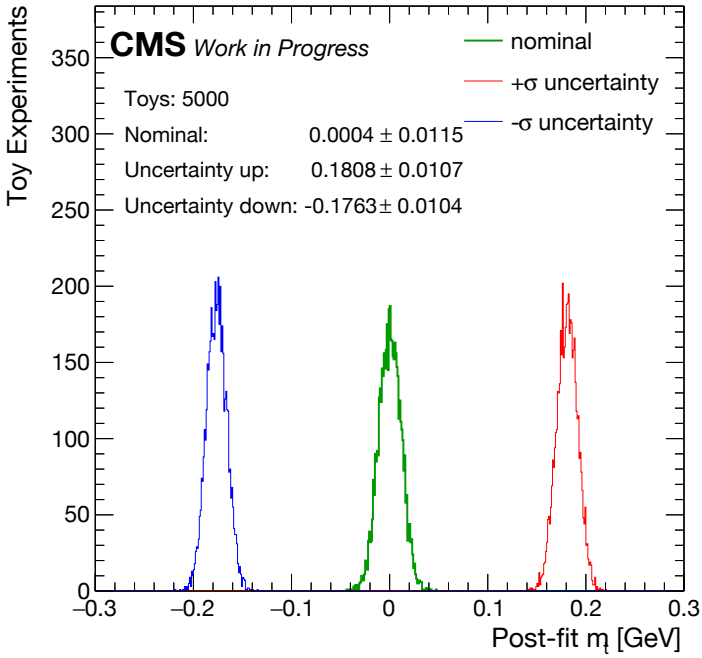

Figure 6.27: Toy experiments for statistical variations: effect on the measurement of m_t.

bin contents in the analysis histograms are fluctuated according to their Poisson-uncertainties, while a fit is made against the nominal (unfluctuated) simulation.

The collective effect on the nominal simulation and systematic uncertainty variations is displayed in Fig. 6.27. The interpretation of the statistical distributions of the uncertainty limits can be made so that **the correct uncertainties are somewhere within these distributions**. This can be better understood visually: to consider uncertainty on uncertainty, one should minimally take the error with one σ error included. This addition is linear (correlated) in its nature, and not quadratic. Thus, generic full uncertainty estimate interpreted from these results would be $\pm(|\mu_{\text{up/down}}|+\sigma_{\text{up/down}}) \rightarrow \mathcal{O}(\pm 0.19\,\text{GeV})$. A more conservative estimate would include 3σs, i.e. $\pm(|\mu_{\text{up/down}}|+3\sigma_{\text{up/down}}) \rightarrow \mathcal{O}(\pm 0.21\,\text{GeV})$.

The results and the methodology are further discussed in Appendix E. The implementation of the toy studies was adapted from the ideas presented in Ref [137]. In the studies of the potentially problematic category 3 nuisances it turned out that limited statistics generally have a small impact on the full results.

6.5 Results

The most important nuisance parameters and their impacts in the full 2017–2018 5D analysis are shown in Fig. 6.28. Figs. 6.29-6.32 present the results in the four analysis channels separately. These impact plots are found using a simulation-only Asimov fit, but they should provide a good estimate for the results that will eventually be found in data.

It is useful to make comparisons between the four analysis channels and the combination. Notably, three b jet nuisances are dominating in the full 2017–2018 combination, including b/t quark FSR, b jet JECs and semileptonic b fragmentation. This is in line with the studies in Section 6.4.6, where it was noted that these nuisances seem to be closely connected. The situation is further complicated by the fact that these b jet nuisances also appear to be somewhat interlinked with the m_t dependence.

It is also good to consider which uncertainties are constrained in the combination. Strong constraints are presented for the different FSR flavors and also for the flavor-dependent JEC uncertainties. Nonetheless, very little constraint is presented on the semileptonic b fragmentation and the correlated b-tagging uncertainties. As the amount of data is increased, it could turn out that these will eventually be the limiting uncertainties in this kind of a m_t measurement. More studies are required on the topic, however.

The current uncertainty forecast of $\mathcal{O}(0.2\,\text{GeV})$ for the combined 5D analysis is well in line with the results of the toy studies in Fig. 6.27. When the analysis is further tuned and finally applied on data, the full uncertainty estimated with Combine can typically change at least by $\mathcal{O}(\pm 0.05\,\text{GeV})$ for this kind of a measurement. Furthermore, if notable pulls are observed in data, the uncertainty estimate could

change even more.

This is demonstrated in the legacy 2016 analysis, where the treatment with a fully split FSR uncertainty scheme yields $\mathbf{m_t = 171.77 \pm 0.38\,GeV}$. However, if b/t quark and light quark FSR uncertainty variations are correlated, the measurement produces $\mathbf{m_t = 172.14 \pm 0.31\,GeV}$ [22]. The question is connected to a notable pull observed in the quark FSR uncertainty, most likely determined by the m_W^{reco} resonance.

In addition to the legacy 2016 analysis, also other recent top quark analyses (e.g. that in Ref. [139]) have hinted that the FSR α_S value in the CP5 tune [78] could be clearly off. As this result has come up in multiple recent results, it would not be surprising to see similar behavior in the present analysis.

However, only the legacy 2016 analysis has so far employed a split scheme of FSR uncertainties. Hence, **it is too early to estimate whether the true FSR discrepancy holds for the combined FSR variations, or only for the light quark FSR**, as presented in the legacy 2016 results. Profile likelihood methods function best when the pulls of nuisance parameters are less than plus/minus one standard deviation, and therefore wider proof on the subject is required.

Considering the 3 times greater luminosity in this analysis compared to the legacy 2016 analysis, a notably smaller uncertainty can be expected. In both the analyses, statistical uncertainties are negligible. However, when systematic uncertainties are profiled into likelihood functions, the statistics in the measurement are able to constrain the systematic uncertainties through the nuisance parameters.

If the leading systematic uncertainties were scaled purely with the amount of data, they would be proportional to the square root of luminosity. A back-of-the-envelope estimate thus leads to the magnitude estimate $0.38/\sqrt{3}\,\mathrm{GeV} \approx 0.22\,\mathrm{GeV}$ (or for the correlated FSR $0.31/\sqrt{3}\,\mathrm{GeV} \approx 0.18,\mathrm{GeV}$) for the projection of the legacy 2016 result to 100 fb^{-1}. This is a naive, yet useful comparison. The full 5D uncertainty in Fig. 6.28 is dominated only by a few b jet uncertainties. Hence, the scaling of the total uncertainty with integrated luminosity depends on how much these systematic uncertainties can be constrained by data.

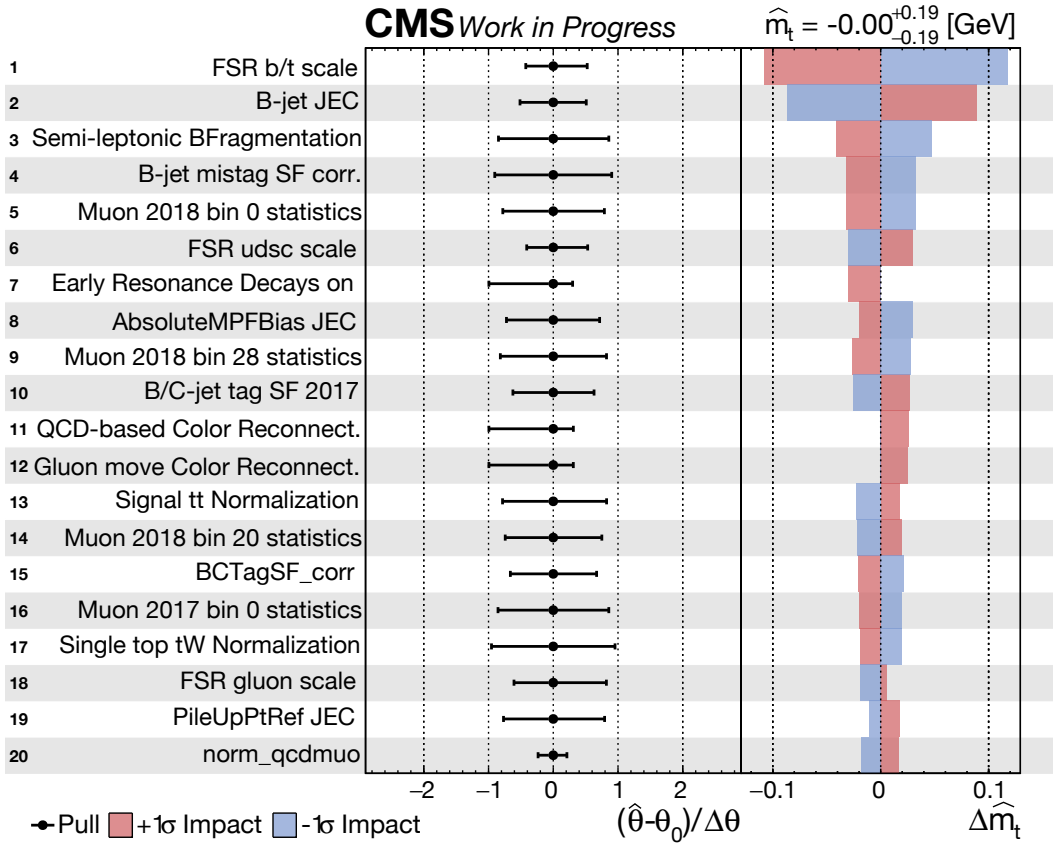

Figure 6.28: The most important impacts in the 5D analysis, combined $\mu 17$ + $\mu 18$ + $e17$ + $e18$ channels.

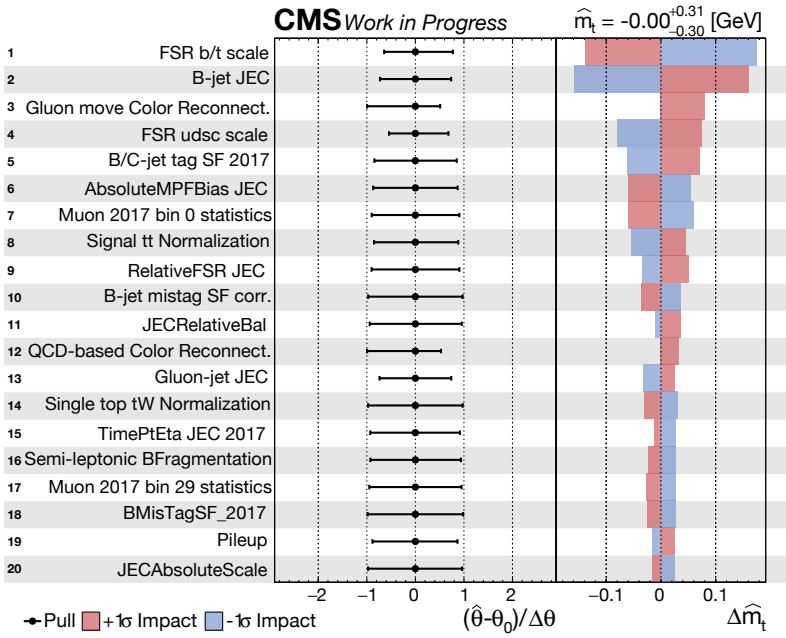

Figure 6.29: The most important impacts in the 5D analysis, μ17 channel.

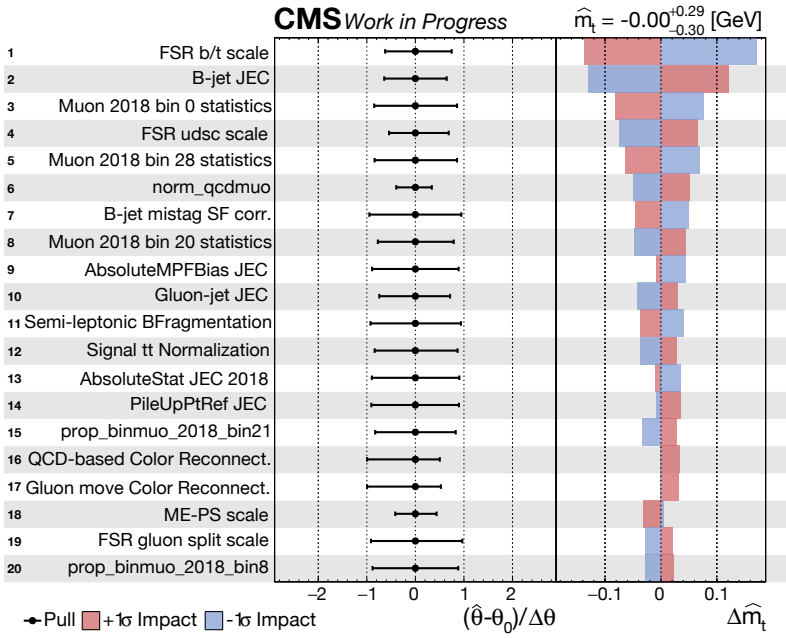

Figure 6.30: The most important impacts in the 5D analysis, μ18 channel.

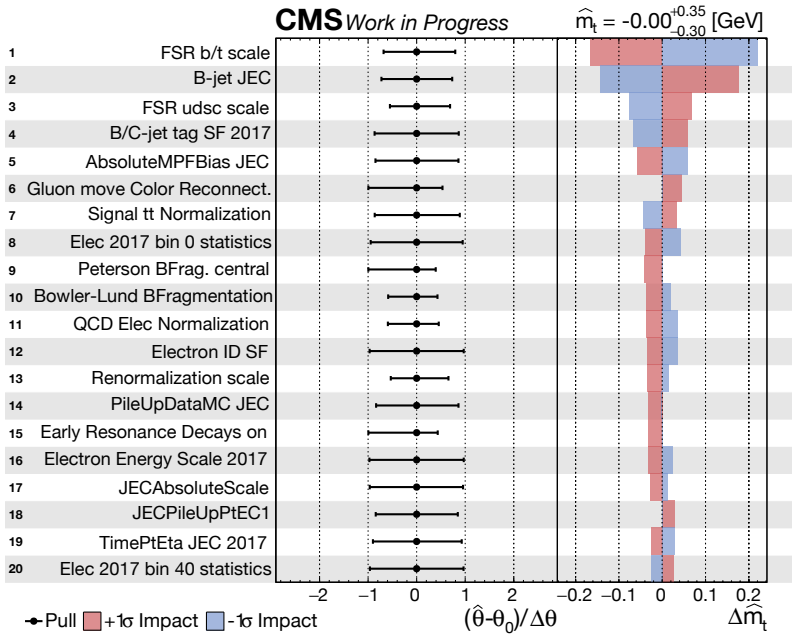

Figure 6.31: The most important impacts in the 5D analysis, $e17$ channel.

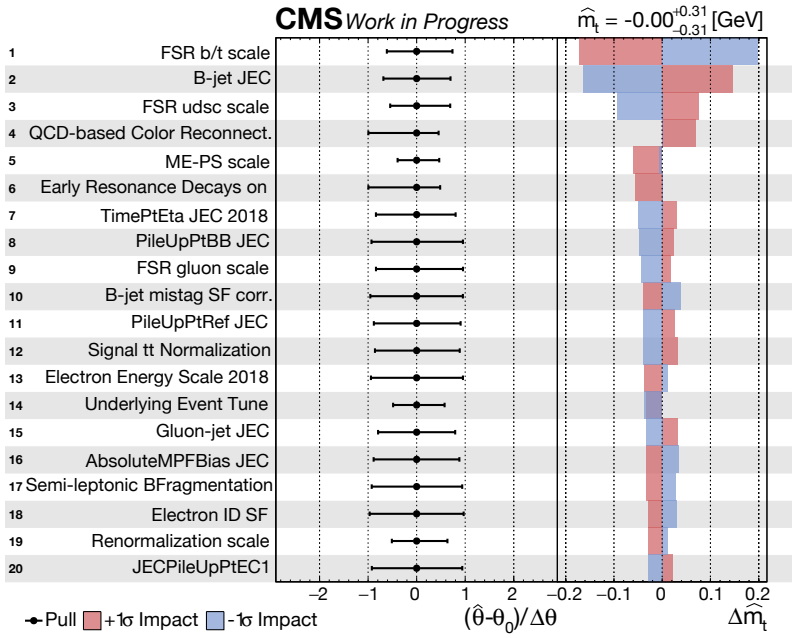

Figure 6.32: The most important impacts in the 5D analysis, $e18$ channel.

7 Summary & Outlook

This thesis strives for a better understanding and a more precise measurement of the top quark mass (m_t). This is an important and ambitious goal, as m_t is a key parameter in the standard model of particle physics. With a relatively small improvement in the measurement precision of m_t, it is feasible to reach new understanding of nature.

Currently the most precise m_t measurements are strongly tied to jet energy corrections, since the top quark decay always produces quarks. These measurements are *direct*, i.e. they are based on the reconstruction of m_t from the top quark decay products. Energetic quarks are always observed as jets at the detector, making the jet modeling and calibration a crucial part of the analysis.

Against this background, the author has worked intensively on tasks enhancing jet calibration. A major target has been the improvement of jet energy corrections in the Run 2 legacy reconstruction, which is the latest and greatest interpretation of the CMS Run 2 data.

A separate study with the same target of a better m_t interpretation was made on the DØ lepton+jets m_t precision measurement. This measurement produced a m_t value notably differing from the other most precise single measurements. Thus, it is important to consider if a physical reason behind this tension could be found.

An earlier study in the Helsinki group located a weak point in the DØ analysis: the flavor-dependent JECs. It appears that the calibration of the flavor JECs has dramatically changed between DØ Run IIa and Run IIb for no apparent reason. By re-implementing a part of the DØ flavor JEC calibration, it was shown in the study that the Run IIb results should have been more like those in Run IIa [3].

In this thesis it is demonstrated that this kind of a change in the flavor JECs shifts the DØ m_t measurement significantly closer to the other most precise measurements. The magnitude of the observed shift in m_t provided compelling evidence for re-opening the examination of the flavor JECs.

The culmination of this thesis is the construction of a lepton+jets m_t analysis

on the CMS legacy 2017–2018 data. This analysis involves much legacy code, and equal amounts of necessary updates. In the end, a completely novel profile likelihood approach with Combine was developed.

The Combine implementation greatly benefited from the similar profile likelihood method presented in the legacy 2016 analysis [22]. An implementation differing from that in the legacy 2016 analysis was chosen to better support the combination of results from different data-taking years. The Combine tool has built-in abilities for making these kinds of combinations.

Even if the Combine implementation was originally motivated by other factors, this choice resulted in a number of additional new analysis decisions. To mention one, Combine is able to conserve the dependence on event yields, whereas the legacy 2016 analysis normalizes all changes in yields away (e.g. from systematic uncertainties).

The preliminary results of the legacy 2016 analysis [22] have been made publicly available, indicating $\mathbf{m_t = 171.77 \pm 0.38\,GeV}$. In an earlier measurement on essentially the same dataset it was found $\mathbf{m_t = 172.25 \pm 0.63\,GeV}$ [20], while the CMS Run 1 combination yielded $\mathbf{m_t = 172.44 \pm 0.48\,GeV}$ [19]. Finally, the m_t analysis of this thesis forecasts an uncertainty on the order of $\mathcal{O}(\pm\mathbf{0.20\,GeV})$ for 2017–2018.

In the legacy 2016 analysis the number of analyzed events was somewhat increased by fine-tuning the selections, but this is not enough to explain the differences in the uncertainties of the two measurements on the 2016 data. Even in the earlier 2016 analysis the statistical uncertainty was not a limiting factor [20]. Thus, the improvements must be explained by the new approaches taken in the legacy 2016 analysis: measurement dimensions have been increased from 2 to 5, and the systematic uncertainties are studied through profile likelihoods.

The profile likelihood approach is revolutionary in the way it benefits from added statistics. When the uncertainty is parametrized as a part of the likelihood analysis, an increase in the number of events can constrain the systematic uncertainties. This is also seen in the forecasts of this thesis. B jet modeling is one of the leading uncertainties, and strong constraints for b jet JEC and b(/t) quark FSR uncertainty are mainly found in the 100 fb^{-1} electron + muon combination. The constraints in the individual measurement channels are less substantial.

An interesting novel feature first presented in the legacy 2016 analysis are the split FSR uncertainties, where different quark/gluon flavors are handled separately. In contrast, the earlier 2016 analysis only utilizes a merged FSR uncertainty variation. Also other settings in the simulation were notably changed between the analyses, as the Pythia 8 tune was changed to CP5 [78].

Consequently, a strong pull in data is observed only for the quark FSR uncertainty – and not for the b/t quark FSR. By correlating the light quark and b quark uncertainties, the legacy 2016 measurement finds an alternative value, $\mathbf{m_t = 172.14 \pm 0.31\,GeV}$ [22]. This is more in line with the previous measurements, and explains the apparent disagreement. This also underlines how strongly

the analysis is dependent on FSR.

The legacy 2016 m_t measurement is the first one among the most precise measurements, where a split scheme of FSR uncertainties has been employed. Only based on this result, it is difficult to say whether a correlated or uncorrelated treatment of the FSR flavors is more natural. In the legacy 2016 analysis the hadronic W resonance is an efficient probe for the light quark FSR, but an equally good measure for the b/t quark FSR might not be found.

Future measurements with split FSR uncertainties will provide new insight into this question, also including the m_t measurement of this thesis. Interesting differences to the legacy 2016 analysis might arise, as for instance the treatment of event yields is more advanced in Combine. With the event selections of the present m_t analysis, FSR variations are known to have a notable impact on observed event yields.

The developments with profile likelihood measurements also bring added excitement towards Run 3 results. The $\mathrm{t\bar{t}}$ cross section will grow with the increased collision energy [140], and the potential for integrated luminosity is around 200 fb^{-1} [48].

When the Run 2 and Run 3 datasets are combined, three times more data will be available than in the present analysis. This is a similar increase as between this and the legacy 2016 analysis. The sheer addition of collision event statistics holds a great potential for constraining the m_t measurement notably more, even if no further advances are made in b jet modeling. The High-Luminosity LHC will have an even more drastic impact, as the amount of integrated luminosity is expected to grow by a factor of 10 [141].

Based on the above considerations, the vacuum metastability question [15] is bound to obtain a more precise interpretation in the upcoming years. The LHC is running again, and the profile likelihood methods bring fresh development potential to m_t measurements. The theoretical m_t precision limit of $\mathcal{O}\,(200\,\mathrm{MeV})$ could soon be within reach.

It is evident that the current standard model of particle physics is not a comprehensive theory of everything. Still it seems that only precision measurements might be able to uncover the yet hidden secrets. In this process, the precision measurement of m_t plays a great role. Time will tell, whether the results of this thesis are used in the evaluation of the question of metastability, or some completely new interpretations.

A Datasets Employed in this Work

In this thesis we are mainly involved with jet production and top quark pair production. To promote readability, for simulations the full dataset path in the format /[Dataset]*/*Summer20UL[YR]*/MINIAODSIM is not given. Instead, we only list the [Dataset] identifier paired up with event counts and cross-sections.

Cross-sections for the simulated samples are mostly derived on the NLO and NNLO level, and their values are extracted from Refs. [83, 119, 142–149]. An exception is made by the binned background samples (binning in HT or Pt), for which only rougher σ-estimates are available - usually estimated in parallel to the simulation.

A.1 Datasets for Jet Studies

In the calibration process of jets, various event types are studied with the base assumption that the behavior of jets is more or less universal. For the specific jet studies presented in this thesis, we use generic Multijet events, which is the process with the greatest statistics. Such events provide excellent material for generic jet studies, as there are rarely any *special* objects – such as energetic photons or muons – within these events. The samples used for Multijet data are found in Table A.1, while the simulated Multijet datasets are found in Table A.2.

Table A.1: 2017-2018 Multijet data samples

Run	ZeroBias dataset path	JetHT dataset path
2017B	/ZeroBias/Run2017B-UL2017_MiniAODv2-v1/MINIAOD	/JetHT/Run2017B-UL2017_MiniAODv2-v1/MINIAOD
2017C	/ZeroBias/Run2017C-UL2017_MiniAODv2-v1/MINIAOD	/JetHT/Run2017C-UL2017_MiniAODv2-v1/MINIAOD
2107D	/ZeroBias/Run2017D-UL2017_MiniAODv2-v1/MINIAOD	/JetHT/Run2017D-UL2017_MiniAODv2-v1/MINIAOD
2017E	/ZeroBias/Run2017E-UL2017_MiniAODv2-v1/MINIAOD	/JetHT/Run2017E-UL2017_MiniAODv2-v1/MINIAOD
2017F	/ZeroBias/Run2017F-UL2017_MiniAODv2-v1/MINIAOD	/JetHT/Run2017F-UL2017_MiniAODv2-v1/MINIAOD
2018A	/ZeroBias/Run2018A-UL2018_MiniAODv2-v1/MINIAOD	/JetHT/Run2018A-UL2018_MiniAODv2-v1/MINIAOD
2018B	/ZeroBias/Run2018B-UL2018_MiniAODv2-v1/MINIAOD	/JetHT/Run2018B-UL2018_MiniAODv2-v1/MINIAOD
2018C	/ZeroBias/Run2018C-UL2018_MiniAODv2-v1/MINIAOD	/JetHT/Run2018C-UL2018_MiniAODv2-v1/MINIAOD
2018D	/ZeroBias/Run2018D-UL2018_MiniAODv2-v1/MINIAOD	/JetHT/Run2018D-UL2018_MiniAODv2-v2/MINIAOD

Table A.2: 2017-2018 Multijet MC files

QCD_Pt-15to7000_TuneCP5_Flat2018_13TeV_pythia8
SingleNeutrino

A.2 Datasets for Top Quark Studies

In a short summary, the relevant data in the m_t measurement consists mainly of $t\bar{t}$ and single top production. Additional backgrounds include qcd/Multijet, w+jets, dy+jets and diboson production. This/These simulated sample(s) are referred to as the central sample(s). This is in contrast to the variation samples used for m_t dependence and studies on systematic errors (nuisances).

A.2.1 CMS Data

Table A.3: Single lepton datasets for Run 2 (2017–2018) legacy data at 13 TeV.

Dataset	Event count	$\mathcal{L}^{\text{Data}}$ [fb^{-1}]
/SingleMuon/Run2017B-UL2017_MiniAODv2-v1/MINIAOD	136 300 266	4.8
/SingleMuon/Run2017C-UL2017_MiniAODv2-v1/MINIAOD	165 652 756	9.6
/SingleMuon/Run2017D-UL2017_MiniAODv2-v1/MINIAOD	70 361 660	4.2
/SingleMuon/Run2017E-UL2017_MiniAODv2-v1/MINIAOD	154 618 774	9.3
/SingleMuon/Run2017F-UL2017_MiniAODv2-v1/MINIAOD	242 140 980	13.5
/SingleMuon/Run2018A-UL2018_MiniAODv2-v3/MINIAOD	241 613 524	14.1
/SingleMuon/Run2018B-UL2018_MiniAODv2-v2/MINIAOD	119 918 017	7.1
/SingleMuon/Run2018C-UL2018_MiniAODv2-v2/MINIAOD	109 986 009	6.9
/SingleMuon/Run2018D-UL2018_MiniAODv2-v3/MINIAOD	513 909 894	29.2
$\sum$ Single Muon Data 2017	769 074 436	41.4
$\sum$ Single Muon Data 2018	985 427 444	57.3
$\sum$ Single Muon Data	1 754 501 880	101.3
/SingleElectron/Run2017B-UL2017_MiniAODv2-v1/MINIAOD	60 537 490	4.8
/SingleElectron/Run2017C-UL2017_MiniAODv2-v1/MINIAOD	136 637 888	9.6
/SingleElectron/Run2017D-UL2017_MiniAODv2-v1/MINIAOD	51 526 521	4.2
/SingleElectron/Run2017E-UL2017_MiniAODv2-v1/MINIAOD	102 122 055	9.3
/SingleElectron/Run2017F-UL2017_MiniAODv2-v1/MINIAOD	128 467 223	13.5
/EGamma/Run2018A-UL2018_MiniAODv2-v1/MINIAOD	339 013 231	14.1
/EGamma/Run2018B-UL2018_MiniAODv2-v1/MINIAOD	153 792 795	7.1
/EGamma/Run2018C-UL2018_MiniAODv2-v1/MINIAOD	147 827 904	6.9
/EGamma/Run2018D-UL2018_MiniAODv2-v2/MINIAOD	752 524 583	29.2
$\sum$ Single Electron Data 2017	479 291 177	41.4
$\sum$ Single Electron Data 2018	1 393 158 513	57.3
$\sum$ Single Electron Data	1 872 449 690	101.3

The datasamples for legacy 2017–2018 are given in Table A.3 together with their luminosities ($\mathcal{L}^{\text{Data}}$). These correspond to a total of 101.3 fb^{-1} data. The event counts between different datasets are not fully comparable, as the set of included

triggers varies. The most notable difference occurs between 2017 and 2018, as the electron dataset changes from SingleElectron to the broader set EGamma.

A.2.2 Simulated Top Quark Production

Top quark production processes, listed in Table A.4, make up a majority of the data after event selections. The samples are generated with the default value of $m_t = 172.5\,\text{GeV}$. The utilized $t\bar{t}$ samples sum up to the full $t\bar{t}$ cross section. In contrast, for the single top samples the fully hadronic decay modes are often excluded, as they pass the final selection extremely rarely.

Table A.4: List of top quark production baseline samples.

Dataset	Events (2017)	Events (2018)	σ [pb]	$\mathcal{L}_{\text{eff}}^{2017}$ [fb^{-1}]	$\mathcal{L}_{\text{eff}}^{2018}$ [fb^{-1}]
$t\bar{t}$ samples: Tune CP5 (POWHEG +PYTHIA NLO)			831.8		
TTToSemiLeptonic	355.33 M	478.98 M	365.5	948.5	1 282.5
TTTo2L2Nu	106.72 M	146.06 M	88.3	1 185.0	1 573.3
TTToHadronic	235.72 M	343.25 M	378.0	613.5	864.5
TTWJetsToLNu	7.46 M	10.52 M	0.2	10 138.5	14 296.9
TTWJetsToQQ	0.66 M	0.97 M	0.4	450.6	663.8
TTZToLLNuNu_M-10	14.20 M	19.61 M	0.2	14 113.8	19 479.4
TTZToQQ	13.98 M	19.82 M	0.5	6 748.8	9 584.1
single top samples (not inclusive): Tune CP5 (POWHEG +PYTHIA NLO)			263.7		
ST_tW_top_5f_NoFullyHadronicDecays	8.51 M	11.27 M	21.7	383.7	570.8
ST_tW_antitop_5f_NoFullyHadronicDecays	8.43 M	11.02 M	21.7	383.9	543.1
ST_t-channel_top_4f_inclusiveDecays	129.90 M	178.76 M	136.0	815.5	1 151.4
ST_t-channel_antitop_4f_inclusiveDecays	69.92 M	95.83 M	81.0	760.3	1 047.6
ST_s-channel_4f_leptonDecays	13.88 M	19.37 M	3.4	1 730.1	2 402.0

The $t\bar{t}$ + V processes are considered separately, as weak showering is not turned on in the CMS PYTHIA settings. The weak boson adds extra leptons or extra jets into the event topology, which changes acceptances in all three $t\bar{t}$ channels. Technically, $t\bar{t}$ +V is a constituent of the full $t\bar{t}$ cross-section. However, it constitutes only around 2 permille of it, which is much less than the cross-section uncertainty.

A.2.3 Simulated Vector Boson Production

In Table A.5 a listing of important HT-binned backgrounds is given, corresponding to the DY+jets and W+jets channels. The W+jets channel may pass the final selections when the W decays leptonically. Furthermore, the DY+jets channel is occasionally able to pass through all the selections when a pair of leptons is produced, and one of them is lost.

Samples for both of these backgrounds exist also without the HT binning, but they provide insufficient statistics. The low limit of the HT-binning at 70 GeV is suitable for the event selection. The 4 jets and the lepton sum up at least to HT $\approx 150\,\text{GeV}$ on the reconstruction level, as HT refers to a simple p_T sum.

Diboson production processes stand as the smallest component of simulated backgrounds. The relevant samples are presented in Table A.6.

Table A.5: List of HT-binned background samples.

Dataset	Events (2017)	Events (2018)	σ [pb]	$\mathcal{L}_{\text{eff}}^{2017}$ [fb^{-1}]	$\mathcal{L}_{\text{eff}}^{2018}$ [fb^{-1}]
W+Jets samples: Tune CP5 (MadGraph-MLM+pythia LO)					
WJetsToLNu_HT-70To100	44.74 M	66.57 M	1 272.0	34.8	51.8
WJetsToLNu_HT-100To200	47.42 M	51.54 M	1 254.0	37.8	41.1
WJetsToLNu_HT-200To400	42.60 M	58.33 M	336.2	126.7	173.5
WJetsToLNu_HT-400To600	5.47 M	7.44 M	45.2	114.5	164.5
WJetsToLNu_HT-600To800	5.55 M	7.72 M	11.0	504.1	701.7
WJetsToLNu_HT-800To1200	5.09 M	7.31 M	4.9	1 033.4	1 484.1
WJetsToLNu_HT-1200To2500	4.96 M	6.48 M	1.2	4 213.5	5 606.8
WJetsToLNu_HT-2500ToInf	1.19 M	2.10 M	0.03	45 221.2	80 001.8
DY+jets samples: Tune CP5 (MadGraph-MLM+pythia LO)					
DYJetsToLL_M-10to50	68.48 M	99.29 M	18 610.0	3.7	5.3
DYJetsToLL_M-50_HT-70to100	12.21 M	17.00 M	159.0	76.7	106.9
DYJetsToLL_M-50_HT-100to200	18.96 M	26.20 M	159.5	118.8	164.2
DYJetsToLL_M-50_HT-200to400	12.51 M	18.46 M	43.6	286.8	401.7
DYJetsToLL_M-50_HT-400to600	5.54 M	8.91 M	5.9	935.1	1 502.7
DYJetsToLL_M-50_HT-600to800	5.28 M	7.04 M	1.4	3 664.8	4 885.1
DYJetsToLL_M-50_HT-800to1200	4.51 M	6.68 M	0.6	6 980.0	10 342.5
DYJetsToLL_M-50_HT-1200to2500	4.80 M	6.17 M	0.2	29 203.5	40 441.1
DYJetsToLL_M-50_HT-2500toInf	1.48 M	1.98 M	0.003	435 307.9	552 632.9

Table A.6: List of Diboson background samples.

Dataset	Events (2017)	Events (2018)	σ [pb]	$\mathcal{L}_{\text{eff}}^{2017}$ [fb^{-1}]	$\mathcal{L}_{\text{eff}}^{2018}$ [fb^{-1}]
Diboson samples: Tune CP5 (pythia LO)					
WW	15.63 M	15.68 M	118.7	131.7	129.7
WZ	7.89 M	7.94 M	45.0	175.3	176.4
ZZ	2.71 M	3.90 M	16.9	160.0	208.5

A.2.4 Simulated QCD Multijet Production

The QCD Multijet background samples are presented in Table A.7. The p_T ranges in the sample names correspond to the employed $\hat{p}_T$ ranges. In pythia LO events, $\hat{p}_T$ is the p_T value of the two outgoing particles in the ME. Samples with lower $\hat{p}_T$ values are required in the muon channel, as the muon p_T cut and HLT p_T threshold are lower than that of electrons (explained in detail in Subsection 6.1.1).

As presented in Eq. (3.9), $\mathcal{L}_{\text{eff}}$ cannot be much smaller than $\mathcal{L}_{\text{Data}}$, or the event weights become too high. High-weighted events inflate the statistical errors in histogram bins and reduce the sensitivity of the analysis. As the full Multijet production cross-section is very large, a sufficient luminosity requires an overwhelming amount of simulated events. Hence, QCD samples enriched in electrons and muons are utilized, allowing higher effective luminosities at lower event numbers.

In the electron channel, both EMEnriched and bcToE samples are required, as they are by design mutually orthogonal. The bcToE samples have been considered in the Run1 lepton+jets m_t analyses [19], but they have been so far excluded in all Run2 analyses [20, 22]. The added contribution of the bcToE samples in the signal region ranges around 75-100% of that of the EMEnriched samples. The tag

Table A.7: List of QCD background samples.

Dataset	Events (2017)	Events (2018)	σ [pb]	$\mathcal{L}_{\text{eff}}^{2017}$ [fb^{-1}]	$\mathcal{L}_{\text{eff}}^{2018}$ [fb^{-1}]	$\mathcal{L}_{\text{emu}}^{2017}$ [fb^{-1}]	$\mathcal{L}_{\text{emu}}^{2018}$ [fb^{-1}]
QCD Muon Enriched samples: Tune CP5 (PYTHIA LO)							
QCD_Pt-20To30_MuEnrichedPt5	64.34 M	60.64 M	3 987 854.9	0.025	0.024	0.645	0.635
QCD_Pt-30To50_MuEnrichedPt5	58.36 M	58.63 M	1 368 000.0	0.043	0.043	1.145	0.366
QCD_Pt-50To80_MuEnrichedPt5	40.38 M	40.02 M	377 800.0	0.107	0.105	1.927	1.941
QCD_Pt-80To120_MuEnrichedPt5	45.98 M	45.57 M	88 490.0	0.516	0.509	10.446	11.231
QCD_Pt-120To170_MuEnrichedPt5	39.39 M	39.11 M	21 190.0	1.852	1.846	36.141	37.835
QCD_Pt-170To300_MuEnrichedPt5	73.07 M	71.93 M	7 039.0	10.373	10.147	157.025	164.137
QCD_Pt-300To470_MuEnrichedPt5	58.69 M	58.95 M	620.5	94.501	94.111	1 243.632	1 246.059
QCD_Pt-470To600_MuEnrichedPt5	39.49 M	38.45 M	59.0	668.301	636.769	8 685.195	7 552.609
QCD_Pt-600To800_MuEnrichedPt5	39.32 M	38.41 M	18.2	2 140.579	2 095.289	22 567.194	23 346.524
QCD_Pt-800To1000_MuEnrichedPt5	78.17 M	78.94 M	3.3	23 817.427	24 068.812	247 465.689	231 835.238
QCD_Pt-1000_MuEnrichedPt5	27.48 M	27.43 M	1.1	25 289.228	25 196.572	220 193.399	223 171.767
QCD EM Enriched samples: Tune CP5 (PYTHIA LO)							
QCD_Pt-30to50_EMEnriched	8.78 M	8.57 M	6 418 000.0	0.001	0.001	0.878	0.891
QCD_Pt-50to80_EMEnriched	10.21 M	10.52 M	1 987 000.0	0.005	0.005	2.528	2.648
QCD_Pt-80to120_EMEnriched	9.62 M	9.47 M	366 100.0	0.026	0.026	12.650	12.820
QCD_Pt-120to170_EMEnriched	9.90 M	9.68 M	66 520.0	0.146	0.145	50.113	48.982
QCD_Pt-170to300_EMEnriched	3.68 M	3.71 M	16 550.0	0.222	0.224	51.161	68.844
QCD_Pt-300toInf_EMEnriched	2.21 M	2.22 M	1 099.0	2.015	2.016	262.584	366.849
QCD bcToE (b/c to electron) samples: Tune CP5 (PYTHIA LO)							
QCD_Pt-30to80_bcToE	15.24 M	15.36 M	362 300 000	0.042	0.042	21.129	2.571
QCD_Pt-80to170_bcToE	15.57 M	15.19 M	33 700 000	0.436	0.449	98.169	61.121
QCD_Pt-170to250_bcToE	15.50 M	15.74 M	2 125 000	7.174	7.340	1 123.094	813.421
QCD_Pt-250toInf_bcToE	15.56 M	15.77 M	562 500	27.483	28.007	2 516.810	2 082.436

bcToE refers to prompt electron production in association to b and c quarks, which is particularly important for the present semileptonic event selection, which requires a prompt lepton and two b-jets.

The QCD cross-sections and $\mathcal{L}_{\text{eff}}$ values depend strongly on the $\hat{p}_T$ ranges of the samples. Low $\mathcal{L}_{\text{eff}}$ values are an issue especially at low $\hat{p}_T$ values, meaning that even after charged lepton enrichment the Multijet background remains problematic. At $\hat{p}_T$ from 20 to 30 GeV very few events pass the event selection in the muon channel, and in the electron channel none, so the lower $\hat{p}_T$ values cause little issues. The small $\mathcal{L}_{\text{eff}}$ values are most troublesome at $\hat{p}_T$ between 30 and 120 GeV.

In order to produce higher QCD luminosities without unreasonably high sample sizes, the selections of this analysis would require enrichment of b-jets in addition to charged lepton enrichment. B-jet enriched QCD samples do exist, but no QCD productions with both b-jet and charged lepton enrichment are available. Moreover, the b-jet enriched samples are unable to catch the full QCD phase-space of this analysis. B-jet enrichment is performed mainly on the ME level, and it cannot account for light quark and gluon jets that are mistagged as b-jets.

As an alternative to the simulated samples with low statistics, data-driven QCD estimation can be considered. In the present analysis, QCD is only a minor background, ranging around $0.5 - 2.0\,\%$ of all the events in data. Furthermore, the measurement of m_t is highly dependent on the complete distributions and shapes measured in data. Hence, data-driven methods would risk biasing the m_t measurement more than they would help with the QCD evaluation.

With no other reasonable alternatives left, b-jet enrichment is simulated by the method of **b tag emulation**, where the b tag cuts are loosened. From the set of available jets, 0-2 pseudo-b-jets are randomly picked, so that the total number of

b-jets becomes 2. Finally, the Multijet event yield is scaled to match the one with full b tag cuts applied. For the emulation, multiple different strategies have been considered. The one presented above is the least problematic and most robust one.

The emulation procedure increases the effective luminosities of the QCD samples by a factor ranging from 10 to 40. The resulting sample luminosities are listed in the rightmost columns of Table A.7 as $\mathcal{L}_{\text{emu}}$.

For the muon samples, emulation is only required at $\hat{p}_T$ below 300 GeV. In this region, the $\mathcal{L}_{\text{emu}}$ values and the resulting event weights become reasonable. In the electron samples, relatively low statistics are currently available. As a temporary solution, this is solved using a boosted emulation method that also considers the electron HLT and ID conditions. When more statistics become available, these samples will be phased out to the normal emulation method, focusing only on b tags.

A.2.5 Simulated Top Mass Variations

Table A.8: List of m_{t}-variation samples.

Dataset	Events (2017)	Events (2018)	σ [pb]	$\mathcal{L}_{\text{eff}}^{2017}$ [fb^{-1}]	$\mathcal{L}_{\text{eff}}^{2018}$ [fb^{-1}]
$\text{t}\bar{\text{t}}$ samples: Tune CP5 (POWHEG +PYTHIA NLO)					
TTToSemiLeptonic_mtop171p5	132.66 M	195.25 M	375.6	353.2	519.8
TTToSemiLeptonic_mtop173p5	140.28 M	199.85 M	355.6	394.5	562.0
TTTo2L2Nu_mtop171p5	41.40 M	59.85 M	90.8	456.0	659.0
TTTo2L2Nu_mtop173p5	42.74 M	59.99 M	86.0	497.2	697.9
single top samples: tune CP5 (POWHEG +PYTHIA NLO)					
ST_tW_top_5f_NoFullyHadronicDecays_mtop1715	3.32 M	4.84 M	22.0	151.0	219.9
ST_tW_top_5f_NoFullyHadronicDecays_mtop1735	3.39 M	4.81 M	21.4	158.3	225.5
ST_tW_antitop_5f_NoFullyHadronicDecays_mtop1715	3.25 M	4.72 M	22.0	147.5	214.4
ST_tW_antitop_5f_NoFullyHadronicDecays_mtop1735	3.45 M	4.79 M	21.4	161.3	223.9
ST_t-channel_top_4f_inclusiveDecays_mtop1715	55.27 M	73.93 M	137.1	403.2	539.2
ST_t-channel_top_4f_inclusiveDecays_mtop1735	54.05 M	74.68 M	135.0	400.5	553.4
ST_t-channel_antitop_4f_inclusiveDecays_mtop1715	27.00 M	37.15 M	81.6	330.8	455.1
ST_t-channel_antitop_4f_inclusiveDecays_mtop1735	27.34 M	35.99 M	80.3	340.6	448.3

Even if the analysis is designed for the semileptonic $\text{t}\bar{\text{t}}$ channel, the m_{t}-dependent backgrounds can introduce notable m_{t}-dependence into the final results. This is a precision measurement of m_t, all the significant backgrounds with m_t-dependence need to be considered. Only the $\text{t}\bar{\text{t}}$ all-hadronic, $\text{t}\bar{\text{t}}$ + V and single top s-channel backgrounds are insignificant enough to be safely ignored in their impact to m_{t}.

As the Combine Tool performs interpolation between histograms, only the m_{t} variations 171.5 GeV and 173.5 GeV are utilized in this analysis. These samples are presented in Table A.8. Considering the historical record of m_{t} measurements, it is unlikely that the legacy 2017–2018 measurement will produce a value outside of the range $171.5 - 173.5$ GeV. In the rare occurrence that this would happen, Combine would be forced to extrapolate and the obtained m_{t} value could become inaccurate. In this scenario the m_{t} variation samples can be effortlessly swapped into a looser range of 169.5 GeV and 175.5 GeV.

A.2.6 Simulated Systematic Variations

Table A.9: List of systematic variation samples for the signal.

Dataset	Events (2017)	Events (2018)	σ [pb]	$\mathcal{L}_{\text{eff}}^{2017}$ [fb^{-1}]	$\mathcal{L}_{\text{eff}}^{2018}$ [fb^{-1}]
$t\bar{t}$ samples: Tune CP5 (POWHEG +PYTHIA NLO)					
TTToSemiLeptonic_hdampUP	132.98 M	199.40 M	365.5	363.9	545.6
TTToSemiLeptonic_hdampDOWN	134.32 M	193.21 M	365.5	367.5	528.7
TTToSemiLeptonic_TuneCP5up	139.11 M	199.46 M	365.5	380.6	545.8
TTToSemiLeptonic_TuneCP5down	135.03 M	190.09 M	365.5	369.5	520.1
TTToSemiLeptonic_TuneCP5_erdON	127.44 M	198.19 M	365.5	348.7	542.3
TTToSemiLeptonic_TuneCP5CR1	141.66 M	199.33 M	365.5	387.6	545.4
TTToSemiLeptonic_TuneCP5CR2	141.92 M	195.10 M	365.5	388.3	533.9

In the current analysis, most systematic variations are executed either using weights or energy scale variations on the central simulated samples. This brings the great benefit of statistical correlation between the central and variated samples. However, some systematics are still produced using separate, completely uncorrelated simulated samples. For the semileptonic $t\bar{t}$ signal, these samples are presented in Table A.9. Besides the question of non-correlated statistics, these systematics suffer from being normalized to the same cross-section with the central simulated signal. The weight-based variations, such as FSR, are able to display physically motivated changes in the total event counts.

Optionally, the same variations can be performed also for the leptonic $t\bar{t}$ channel and single top tW. For the nuisances, considering these is not as crucial as for the mass dependence. Both of these background samples make up 2-3 % of the data after final selections, so the majority of the nuisance behavior is caught by the semileptonic signal. The single top t-channel can be safely ignored, as it makes up only around 0.5 % of the data after selections.

B

DØ Likelihood Calibration Ambiguity

In the 2011 [150] and 2015 [36] DØ measurements, the analysis chain of the lepton+jets top quark mass measurement is slightly altered from its previous incarnations. The 2011 paper [150] explanation for Eqs. (21,22) is the following:

"The likelihoods ... are calibrated by replacing m_t and K_{JES} by parameters fitted to the response plots of Sec. VI:

$$\begin{aligned} m_t^{calib} &= \frac{(m_t - 172.5\ GeV) - p_0^{m_t}}{p_1^{m_t}} + 172.5\ GeV, \\ k_{JES}^{calib} &= \frac{(k_{JES} - 1) - p_0^{K_{JES}}}{p_1^{K_{JES}}} + 1.\text{“} \end{aligned}$$

Correspondingly, the 2015 paper [36] explanation for Eqs. (35,36) proceeds as:

"The linear calibrations are applied through the likelihoods ... by transforming m_t and K_{JES} parameters according to

$$\begin{aligned} m_t \to m_t' &= \frac{(m_t - 172.5\ GeV) - o_{m_t}}{s_{m_t}} + 172.5\ GeV, \\ k_{JES} \to k_{JES}' &= \frac{(k_{JES} - 1) - o_{K_{JES}}}{s_{K_{JES}}} + 1.\text{“} \end{aligned}$$

The replacement suggested by both papers is of the form $m_t \to g^{-1}(m_t)$ and $K_{\text{JES}} \to f^{-1}(K_{\text{JES}})$, following the equivalent notation of Eq. (5.8) and Eq. (5.11).

We will show using m_t, where this transformation leads (analogous derivation is valid for K_{JES}). Starting by making the substitution $m_t \to g^{-1}(m_t)$ for the

likelihoods in Eq. (5.6):

$$m_t^{\mathrm{D\O}} = \langle m_t \rangle_{\mathrm{Transformed}} = \frac{\int_{\mathbb{R}} dm_t m_t \mathcal{L}(g^{-1}(m_t))}{\int_{\mathbb{R}} dm_t \mathcal{L}(g^{-1}(m_t))}. \tag{B.1}$$

Remembering the parametrization Eq. (5.8) of Eq. (5.7), the inverse function attains the parametrized form

$$m_t^{gen} = g^{-1}(m_t^{fit}) = \frac{\left(m_t^{fit} - 172.5\right) - \mathcal{O}}{\mathcal{S}} + 172.5. \tag{B.2}$$

Let us perform a change of variables: $m_t' = g^{-1}(m_t)$, i.e. $m_t = g(m_t')$. By Eq. (B.2), we find the transformation

$$dm_t = (dg(m_t')/dm_t')\ dm_t' = \mathcal{S}\ dm_t'. \tag{B.3}$$

After the transformation, the integration limits are still over the whole $\mathbb{R}$. Eq. (B.1) now reads

$$\begin{aligned}
\langle m_t \rangle_{\mathrm{Transformed}} &= \frac{\int_{\mathbb{R}} \mathcal{S}\ dm_t' g(m_t') \mathcal{L}(m_t')}{\int_{\mathbb{R}} \mathcal{S}\ dm_t' \mathcal{L}(m_t'))} \\
&= \frac{\int_{\mathbb{R}} dm_t' \left(\mathcal{S}\left(m_t' - 172.5\right) + \mathcal{O} + 172.5\right) \mathcal{L}(m_t')}{\int_{\mathbb{R}} dm_t' \mathcal{L}(m_t'))} \\
&= \mathcal{S} \frac{\int_{\mathbb{R}} dm_t' m_t' \mathcal{L}(m_t')}{\int_{\mathbb{R}} dm_t' \mathcal{L}(m_t'))} + (-\mathcal{S}\ 172.5 + \mathcal{O} + 172.5) \frac{\cancel{\int_{\mathbb{R}} dm_t' \mathcal{L}(m_t')}}{\cancel{\int_{\mathbb{R}} dm_t' \mathcal{L}(m_t'))}} \\
&= \mathcal{S}\left(\langle m_t \rangle_0 - 172.5\right) + \mathcal{O} + 172.5 = f(\langle m_t \rangle_0),
\end{aligned}$$

where we have again utilized Eqs. (5.6,5.8). The found result is

$$\langle m_t \rangle_{transformed} = g(\langle m_t \rangle_0) = g(m_t^{fit}). \tag{B.4}$$

This is not the desired result from Eq. (5.9): we should have obtained $g^{-1}(m_t^{fit})$ and not $g(m_t^{fit})$. By interchanging g and g^{-1} in the equation chain above, it turns out that the substitution $m_t \to g(m_t)$ into $\mathcal{L}(m_t)$ produces the correct m_t^{calib}. Now, utilizing Eqs. (5.9) and (B.4), we notice that

$$m_t^{\mathrm{calib}} = g^{-1}(m_t^{fit}) = g^{-1}(g^{-1}(m_t^{D0})). \tag{B.5}$$

Continuing with Eqs. (B.2) and (B.5), we find

$$m_t^{\mathrm{calib}} = \frac{\left(\frac{\left(m_t^{D0} - 172.5\right) - \mathcal{O}}{\mathcal{S}} + \cancel{172.5} - \cancel{172.5}\right) - \mathcal{O}}{\mathcal{S}} + 172.5. \tag{B.6}$$

That is,

$$m_t^{\text{calib}} = \frac{\left(m_t^{D0} - 172.5\right) - \mathcal{O}\,(1+\mathcal{S})}{\mathcal{S}^2} + 172.5. \tag{B.7}$$

For σ's we can derive similar statements, except that the translational terms are dropped out:

$$\sigma_t^{\text{calib}} = \sigma_t^{D0}/\mathcal{S}^2 \tag{B.8}$$

In summary, if the 2011 and 2015 DØ lepton+jets measurements have performed the calibration step as they state in the papers, we find explicit corrections that should be performed to the measured values. The scale $\mathcal{S}$ is typically around 0.89 - 0.97, so the variances (statistical errors) found by DØ should be scaled slightly upwards. In addition, the measured m_t values attain non-trivial translational terms.

Applying the transformations presented in Eqs. (B.7) and (B.8), we find for the 2011 paper [150]

$$m_t = 174.94 \pm 1.49 \rightarrow 174.126 \pm 1.55$$

and or the 2015 paper [36], we find similarly

$$\begin{aligned} m_t^e &= 175.55 \pm 0.81 \rightarrow 175.88 \pm 0.97 \\ m_t^\mu &= 174.36 \pm 0.84 \rightarrow 174.02 \pm 0.87 \\ m_t^{tot} &= 174.98 \pm 0.58 \rightarrow 174.87 \pm 0.65 \end{aligned}$$

The changes are subtle, yet notable. As an additional note: in the transformed values it has also been taken into account that in the electronic channel for IIb34 (according to Fig. (15) in [36]) the pull around $m_t = 175$ GeV is actually approximately 1.4, instead of the stated 1.16.

As a final interesting detail, the χ^2 values of combining the four RunII eras (NDF = 3) are viewed. For the electron and muon channels separately, these are found to transform as follows:

$$\begin{aligned} \chi^2/NDF_e &= 1.22 \rightarrow 3.71 \\ \chi^2/NDF_\mu &= 0.19 \rightarrow 1.02 \end{aligned}$$

So if the presented kind of a transformation was required, the χ^2/NDF for electrons moves from reasonable to slightly unreasonable (too high). Correspondingly, the muonic value would transform from unreasonable (too low) to reasonable. This could suggest that the electron channel errors are underestimated.

C Error Estimates of the Mean in Linear Regression

For $\hat{K}_{\text{JES}}$, the one-parameter linear regression model of Eq. (5.37) is used. Correspondingly, for $\hat{m}_t$ we use the two-parameter linear regression of Eq. (5.38). The predictions produced by the fits are interpreted as the best estimates for the mean values of fits. For the one-parameter regression of $\hat{K}_{\text{JES}}$ based on N data points, this estimate reads [151]

$$\begin{aligned}\delta\hat{K}_{\text{JES}} = t_{\alpha/2}^{N-2} \times & \sqrt{\frac{1}{N-2}\sum_{i}^{N}\left(q_0 + q_1 \times K_{\text{JES}}^{\text{Res},i} - \hat{K}_{\text{JES}}^{i}\right)^2} \\ & \times \sqrt{\frac{1}{N} + \frac{\left(K_{\text{JES}}^{\text{Res}} - \overline{K}_{\text{JES}}^{\text{Res}}\right)^2}{\sum_{i=1}^{N}\left(K_{\text{JES}}^{\text{Res},i} - \overline{K}_{\text{JES}}^{\text{Res}}\right)^2}},\end{aligned} \tag{C.1}$$

where the summations are taken over the data points used for the regression. By $\overline{K}_{\text{JES}}^{\text{Res}}$ we refer to the average $K_{\text{JES}}^{\text{Res}}$ value of the regression dataset. The first square root term is the unbiased estimator of the regression error variance, and the second a $K_{\text{JES}}^{\text{Res}}$-dependent factor. Moreover, $t_{\alpha/2}^{N-2}$ is the Student t score value with $N-2$ degrees of freedom. Here, we select $\alpha \approx 68\%$ to retain a correspondence to a one sigma deviation. With this choice, at the the infinite N limit where the Student t distribution corresponds to a normal distribution,

$$\lim_{N\to\infty} t_{\alpha/2}^{N-2} = 1. \tag{C.2}$$

For the two parameter $\hat{m}_t$ fit the corresponding expression becomes more complicated. It stands as [151]

$$\delta\hat{m}_t = t_{\alpha/2}^{N-3} \times \sqrt{\frac{1}{N-3}\sum_i^N \left(p_0 + p_1 \times m_t^{gen,i} + p_2 \times K_{\mathrm{JES}}^{\mathrm{Res},i} - \hat{m}_t^i\right)^2} \times \sqrt{\vec{z}^T \left(\mathbf{X}^T\mathbf{X}\right)^{-1} \vec{z}}. \tag{C.3}$$

This is a generalization of Eq. (C.1), where the degrees of freedom have been reduced by one, due to the one additional fit parameter. The latter square root term conveys dependence on m_t^{gen} and $K_{\mathrm{JES}}^{\mathrm{Res}}$. The column vector term is defined as

$$\vec{z} = [1, m_t^{gen}, K_{\mathrm{JES}}^{\mathrm{Res}}]^T. \tag{C.4}$$

The measurement matrix $\mathbf{X}$ holds the data points used for regression in its N rows. The ith row x_i of $\mathbf{X}$ consists of the row vector

$$x_i = [1, m_t^{gen,i}, K_{\mathrm{JES}}^{\mathrm{Res},i}]. \tag{C.5}$$

D Systematic variations for analysis histograms

This appendix is dedicated for variations of the nuisance parameters on the histogram level, carrying on from subsection 6.4.6. In the histograms, the central values are presented by a green line, while the up and down variations of the current systematic are given in red and blue, correspondingly.

The plots are given exclusively for the $\mu 17$ channel, as the differences between the channels are insignificant. A binning with 10 bins in each dimension is used. We focus on relative variations, normalized by the nominal simulation. This is motivated at the end of this appendix. As a reminder, the four leftmost dimensions are located in $P_{\text{gof}} \geq 0.2$, while the rightmost dimension has higher statistics and is located in $P_{\text{gof}} < 0.2$.

D.1 Further Variations on B Jet Fragmentation

Fig. D.1 displays the central Peterson b fragmentation weight and the relative Bowler-Lund b fragmentation variations. These trends are somewhat different to those presented in Figs. 6.25,6.26, considering other b quark/jet variations.

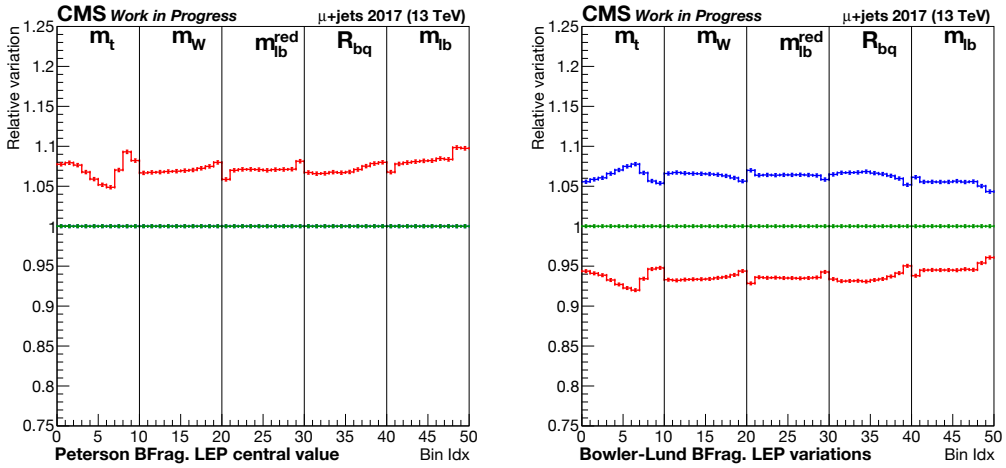

Figure D.1: Central variation of Peterson B Fragmentation (left) and Relative Bowler-Lund B fragmentation variations (right) in the five dimensions m_t^{fit}, m_W^{reco}, $m_{lb}^{red,reco}$, R_{bq}^{reco} ($P_{\text{gof}} \geq 0.2$) and m_{lb}^{reco} ($P_{\text{gof}} < 0.2$).

D.2 Quark Jets

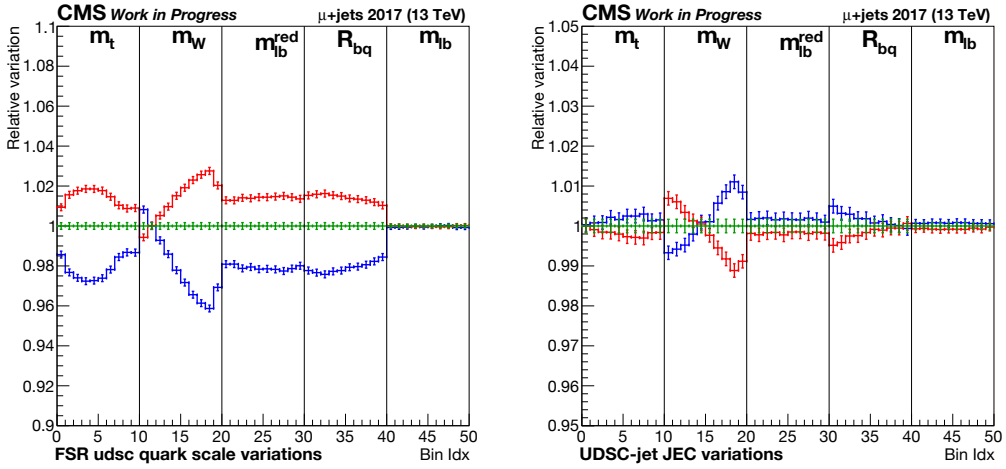

Figure D.2: Variations of quark FSR (left) and quark jet JEC (right) in the five dimensions m_t^{fit}, m_W^{reco}, $m_{lb}^{red,reco}$, R_{bq}^{reco} ($P_{\text{gof}} \geq 0.2$) and m_{lb}^{reco} ($P_{\text{gof}} < 0.2$).

Quark FSR is presented in Fig. D.2 together with quark jet JEC – including a yield scaling offset. Similar symmetries can be observed between these two as was done in the b jet case. The main differences are in the yield scaling offsets and the

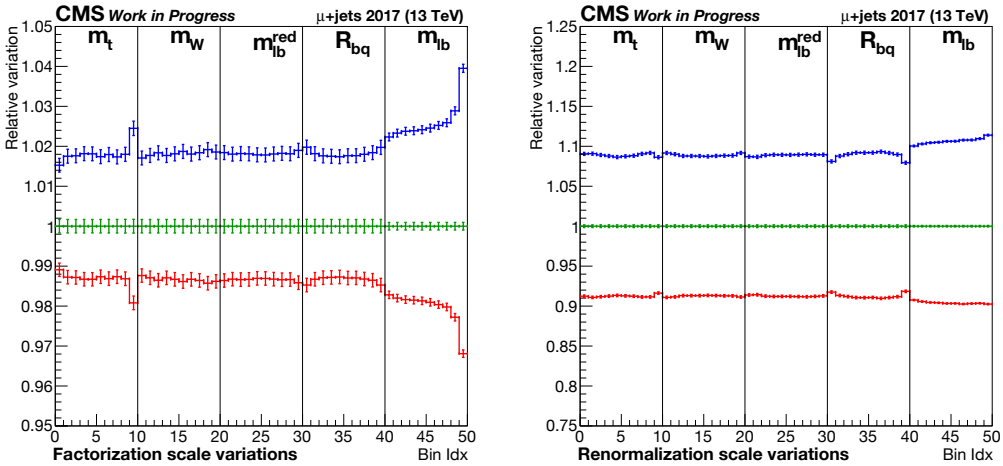

Figure D.3: Variations of Factorization (left) and Renormalization (right) scales in the five dimensions m_t^{fit}, m_W^{reco}, $m_{lb}^{red,reco}$, R_{bq}^{reco} ($P_{\text{gof}} \geq 0.2$) and m_{lb}^{reco} ($P_{\text{gof}} < 0.2$).

behavior of m_t^{fit}. Significant P_{gof} dependence is observed in quark FSR, where the variations in the total yields differ notably between $P_{\text{gof}} < 0.2$ and $P_{\text{gof}} \geq 0.2$.

D.3 Other Weight Variations

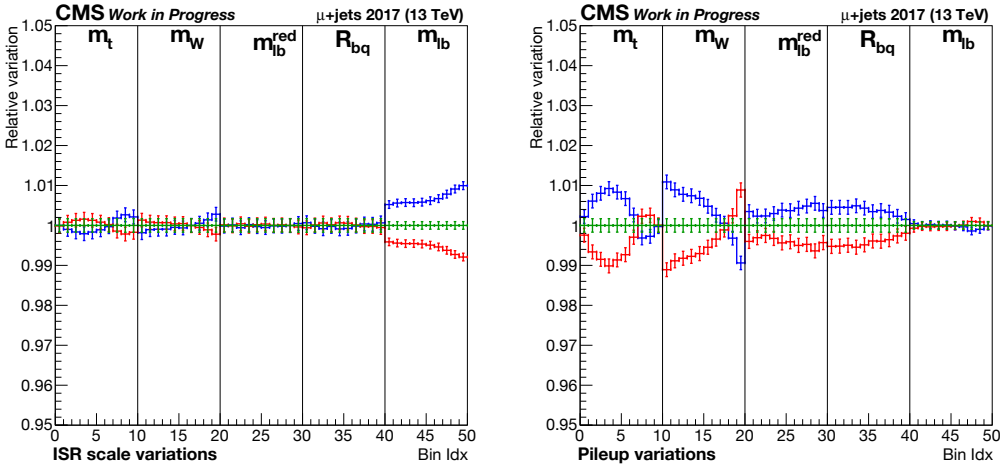

Figure D.4: Variations of ISR (left) and Pileup (right) in the five dimensions m_t^{fit}, m_W^{reco}, $m_{lb}^{red,reco}$, R_{bq}^{reco} ($P_{\text{gof}} \geq 0.2$) and m_{lb}^{reco} ($P_{\text{gof}} < 0.2$).

In Fig. D.3 the factorization and renormalization scale variations are shown.

The renormalization scale thus mostly acts as an additional scaling systematic. For factorization the case is almost the same, excluding slight trends in m_t^{fit} and m_{lb}. In Fig. D.4 the ISR and Pileup dependencies are shown. The ISR dependence is strongest at low P_{gof}, while pileup effects are most important at high P_{gof}. It is also observed that for ISR the yields depend on P_{gof} in a similar manner as for FSR.

D.4 Category 2 Variations

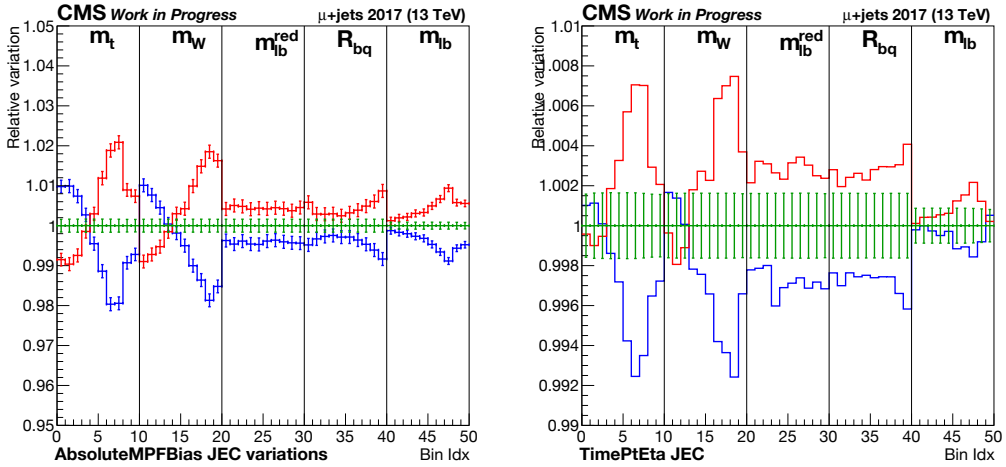

Figure D.5: Variations of JEC AbsoluteMPFBias (left) and TimePtEta (right) in the five dimensions m_t^{fit}, m_W^{reco}, $m_{lb}^{red,reco}$, R_{bq}^{reco} ($P_{\text{gof}} \geq 0.2$) and m_{lb}^{reco} ($P_{\text{gof}} < 0.2$).

In this part, some of the most important category 2 variations[1] are presented. Fig. D.5 displays the AbsoluteMPFBias and TimePtEta JECs, while in Fig. D.6 the PileUpDataMC and RelativeFSR JECs are given. Excluding the yield offsets, there appears to be some redundancy between the JEC variations. This is not unexpected. For the JEC variations it can be estimated that the correlation in statistics between the nominal and variations is good, as the relative fluctuations are much smaller than the displayed statistical errors.

In Fig. D.7 the $|\eta| < 1.93$ JER variations and variations of the unclustered $p_{\text{T}}^{\text{miss}}$ are given. Both of these are completely unique as compared to the variations presented earlier.

[1] Partial correlation of statistics.

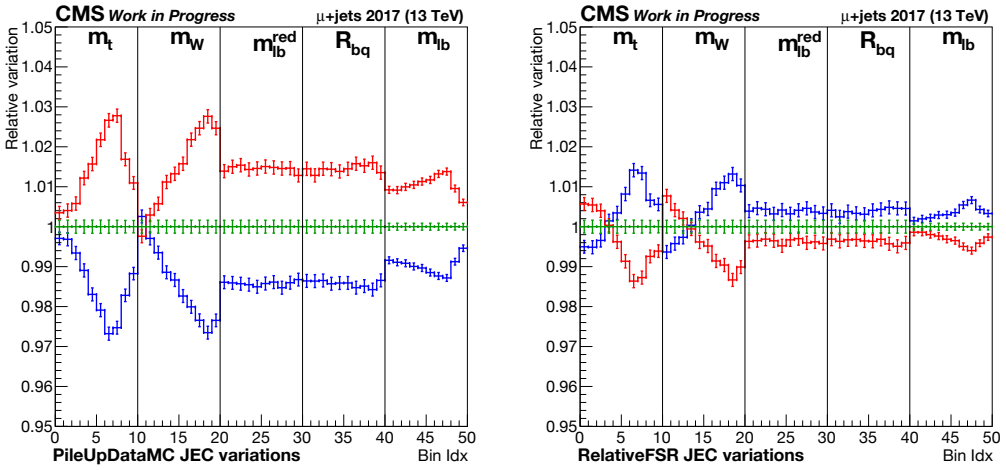

Figure D.6: Variations of JEC PileUpDataMC (left) and JEC RelativeFSR (right) in the five dimensions m_t^{fit}, m_W^{reco}, $m_{lb}^{red,reco}$, R_{bq}^{reco} ($P_{\text{gof}} \geq 0.2$) and m_{lb}^{reco} ($P_{\text{gof}} < 0.2$).

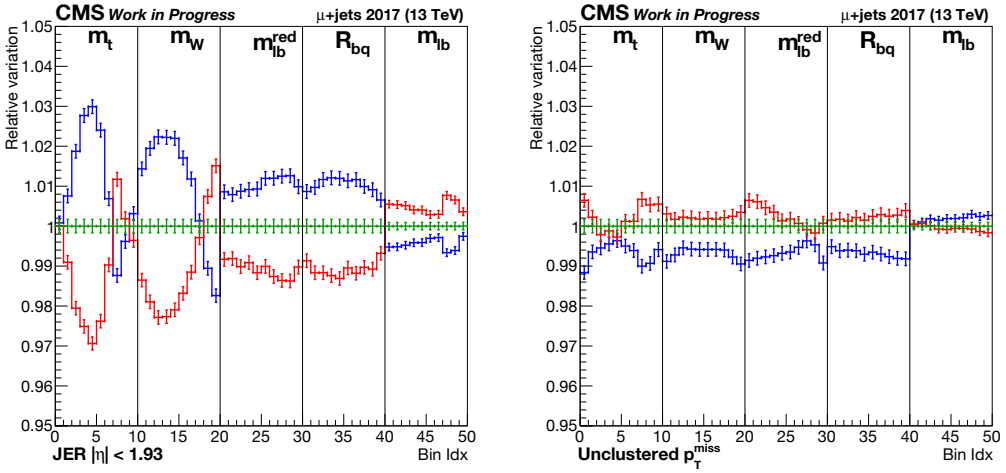

Figure D.7: Variations of the JER $|\eta| < 1.93$ and unclustered $p_{\text{T}}^{\text{miss}}$ in the five dimensions m_t^{fit}, m_W^{reco}, $m_{lb}^{red,reco}$, R_{bq}^{reco} ($P_{\text{gof}} \geq 0.2$) and m_{lb}^{reco} ($P_{\text{gof}} < 0.2$).

D.5 Category 3 Variations

Here, we concentrate lastly on the uncorrelated category 3 variations. In Fig. D.8, the hdamp-dependence is given, accompanied by the ERD-dependence. These both have well-defined and unique trends in the five dimensions. For HDAMP the $P_{\text{gof}} < 0.2$ zone functions as an anchor of the event yield in a similar manner as for ISR and FSR.

Fig. D.9 shows the QCDBased and GluonMove Color Reconnections, where the signal trends are barely distinguishable from noise. These variations are rescued by the combination between all channels in 2017-2018, as these systematics are fully correlated between the years.

Finally, Fig. D.10 variations are given. Here, the variations would be almost indistinguishable from noise, if it were not for the high statistics in m_{lb} – this is also confirmed by Kolmogorov tests of histogram compatibility. The trends are better distinguishable from noise in the even-binned cases – with best results given for the 10-bin case.

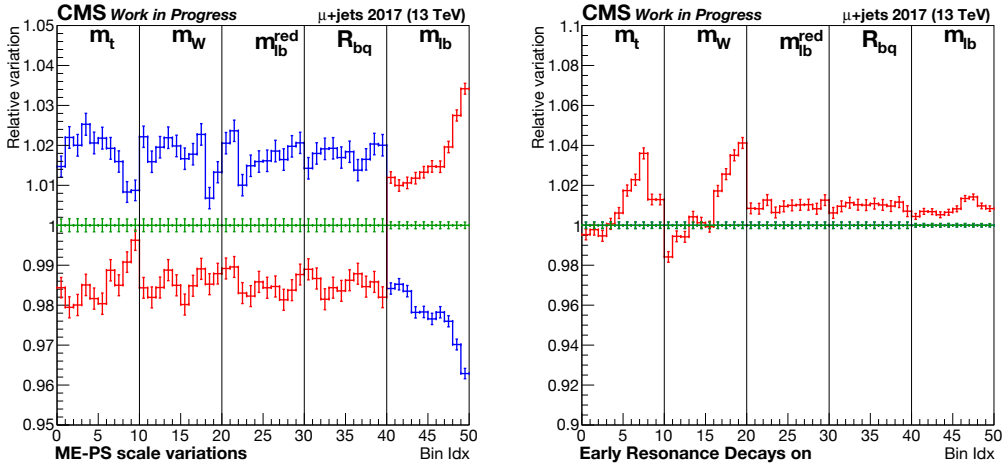

Figure D.8: Variations of HDAMP (left) and ERD (right) in the five dimensions m_t^{fit}, m_W^{reco}, $m_{lb}^{red,reco}$, R_{bq}^{reco} ($P_{\text{gof}} \geq 0.2$) and m_{lb}^{reco} ($P_{\text{gof}} < 0.2$).

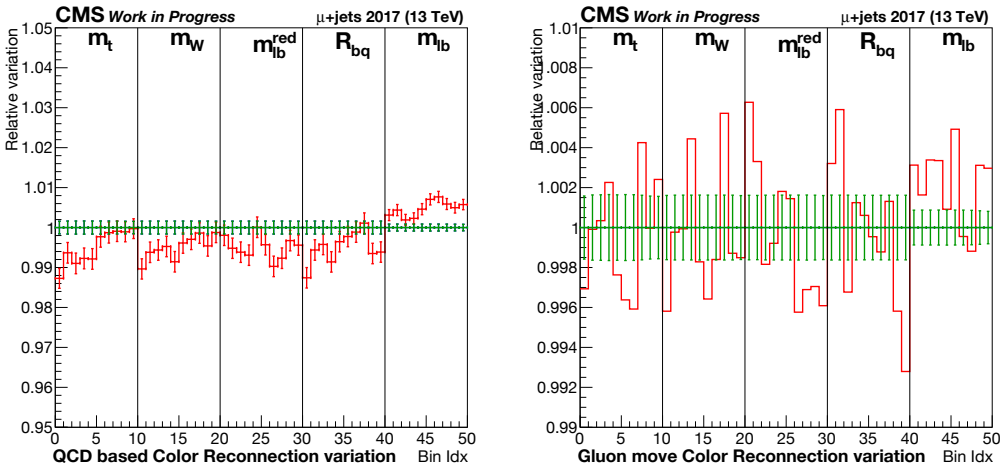

Figure D.9: Variations of CR-QCDBased (left) and GluonMove (right) in the five dimensions m_t^{fit}, m_W^{reco}, $m_{lb}^{red,reco}$, R_{bq}^{reco} ($P_{\text{gof}} \geq 0.2$) and m_{lb}^{reco} ($P_{\text{gof}} < 0.2$).

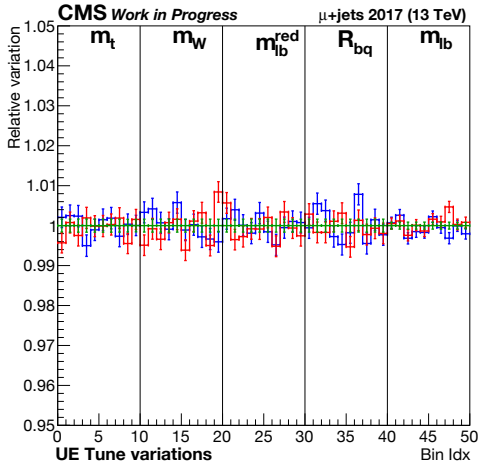

Figure D.10: Variations of UE Tune in the five dimensions m_t^{fit}, m_W^{reco}, $m_{lb}^{red,reco}$, R_{bq}^{reco} ($P_{\text{gof}} \geq 0.2$) and m_{lb}^{reco} ($P_{\text{gof}} < 0.2$).

D.6 Absolute Variations

In Fig. D.11 and Fig. D.12 absolute counts for some variations are given. The m_{lb} statistics have been reduced here so that the m_{lb} variable lands on a similar y-position as the other variables. It is easy to appreciate that especially with the smaller variations, the absolute plots are not that informative. Even in the case of the distinctive trends in m_t, fluctuations become distracting. Small fluctuations in the nominal sample are due to the fine-tuning of the binning variables. Such small differences are statistically insignificant, and not a reason to revisit the bin edge position calibration.

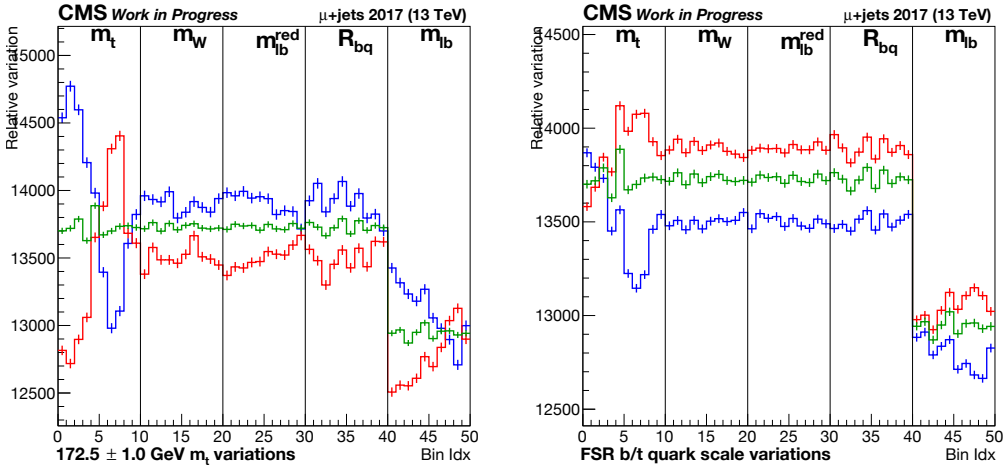

Figure D.11: Absolute variations in the five analysis dimensions (m_t^{fit}, m_W^{reco}, $m_{lb}^{red,reco}$, R_{bq}^{reco} ($P_{\text{gof}} \geq 0.2$) and m_{lb}^{reco} ($P_{\text{gof}} < 0.2$)) for m_t (left) and b-FSR (right). Statistics reduced for m_{lb} in $P_{\text{gof}} < 0.2$ (rightmost plot).

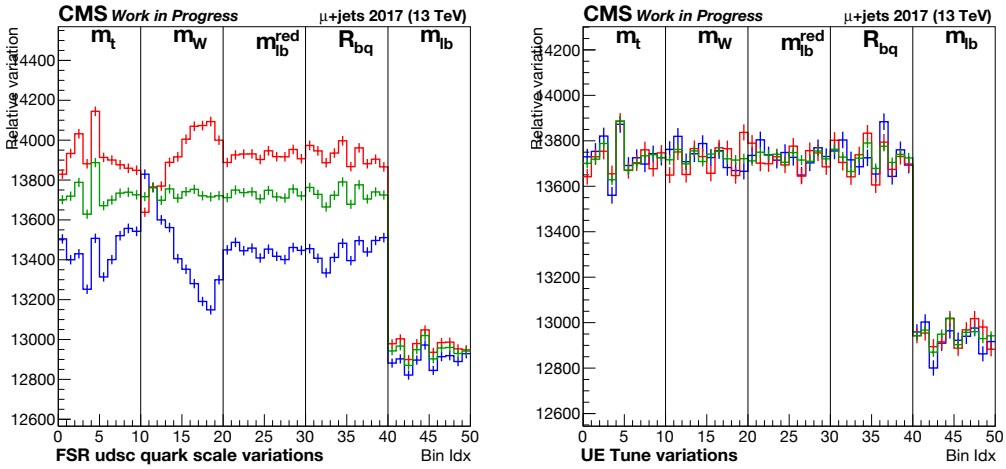

Figure D.12: Absolute variations in the five analysis dimensions (m_t^{fit}, m_W^{reco}, $m_{lb}^{red,reco}$, R_{bq}^{reco} ($P_{\text{gof}} \geq 0.2$) and m_{lb}^{reco} ($P_{\text{gof}} < 0.2$)) for quark FSR (left) and UE tune (right). Statistics reduced for m_{lb} in $P_{\text{gof}} < 0.2$ (rightmost plot).

E Toy Studies on Limited Statistics

This chapter presents a collection of toy studies, mainly on the category 3 variables. The studies were executed by varying the bin contents in the nominal (central) simulation and the category 3 nuisances independently, according to Poisson-distributed statistical uncertainties. The bin contents in all other nuisance categories are varied synchronously with the nominal simulation. As this treatment allows fluctuation in the nominal sample, the Barlow-Beeston uncertainties are turned off. Fluctuating all the nuisances at the same time takes into account the possible correlations between nuisances.

The fit is performed repetitively on the same Asimov data 5000 times to gather sufficient statistics, varying the statistical Poisson fluctuations in each iteration. In each iteration, the analysis histogram bin contents are varied as described above. the values, constraints and impacts of all nuisances are recorded.

Fig. E.1 displays ERD, Fig. E.2 QCD-Based Color Reconnection and Fig. E.3 Gluon-Move Color Reconnection.. Finally, HDAMP results are found in Fig. E.4 and UE in Fig. E.5. All the distributions are Gaussian. The central values and variances of the nominal fit and the $\pm\sigma$ variations are displayed in the figures.

The results generally agree with Asimov. Best fit values are centered at zero and resolutions are small. It can hence be concluded that the limited statistics pose no notable disadvantages.

To give examples on toy experiments for category 1 and 2 nuisances, Fig. E.6 presents the variations on b/t quark FSR and Fig. E.7 those on b jet JEC. The results are in line with Asimov impacts.

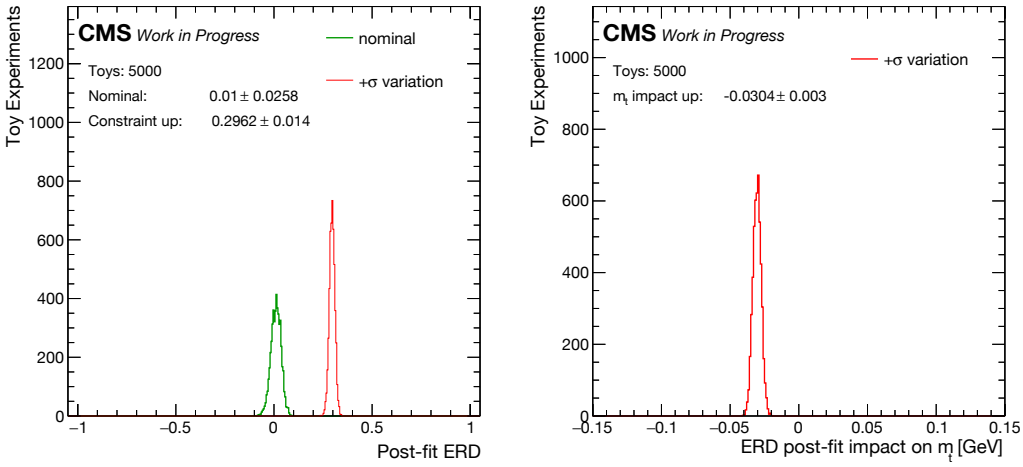

Figure E.1: ERD post-fit constraints (left) and m_t impacts (right).

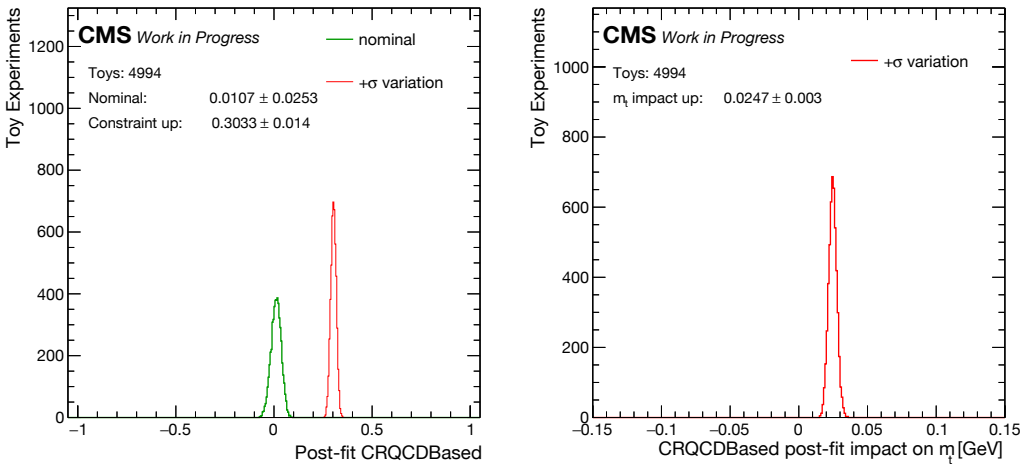

Figure E.2: CR-QCDBased post-fit constraints (left) and m_t impacts (right).

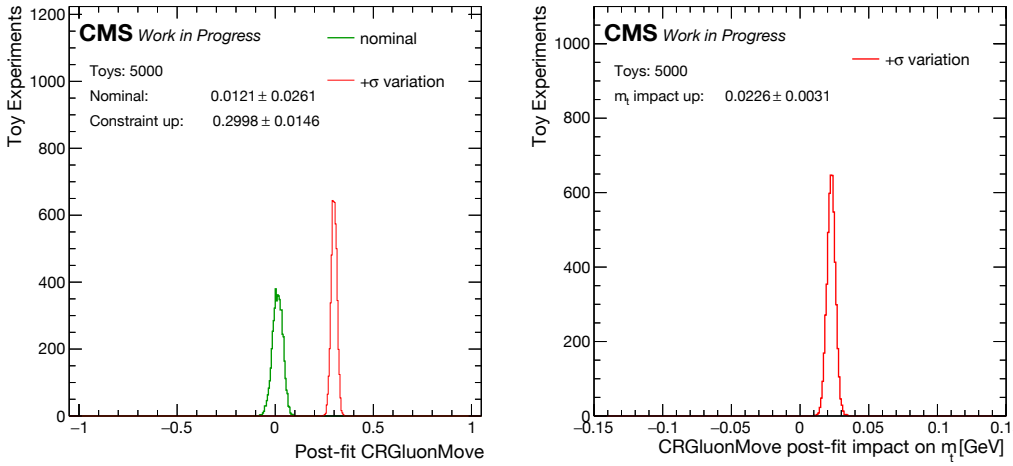

Figure E.3: CR-GluonMove post-fit constraints (left) and m_t impacts (right).

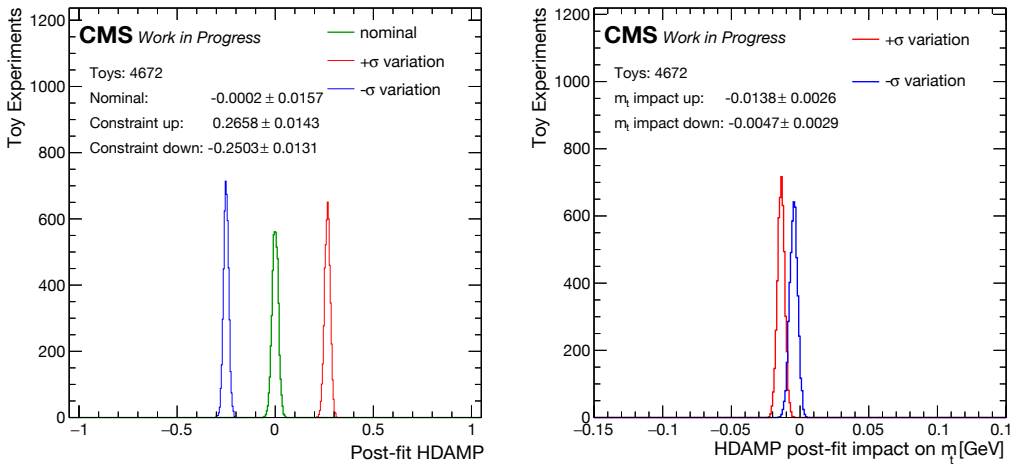

Figure E.4: HDAMP post-fit constraints (left) and m_t impacts (right).

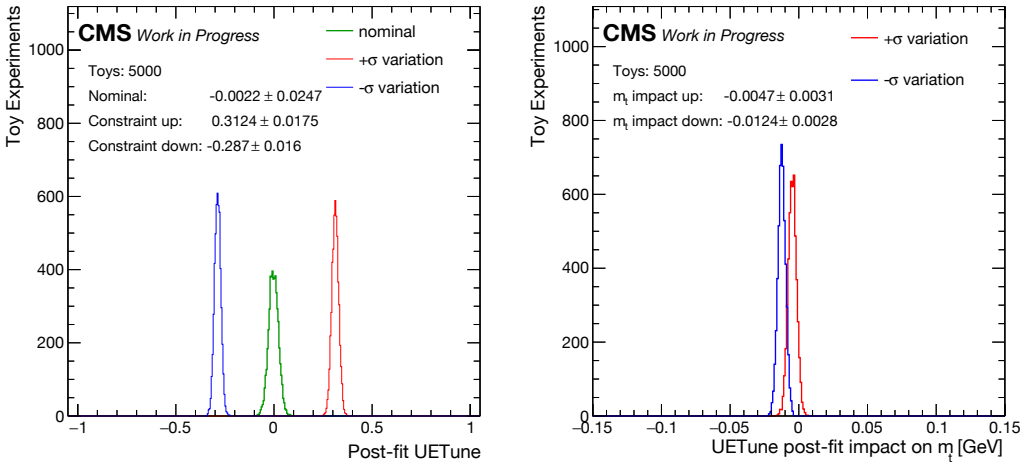

Figure E.5: UE Tune post-fit constraints (left) and m_t impacts (right).

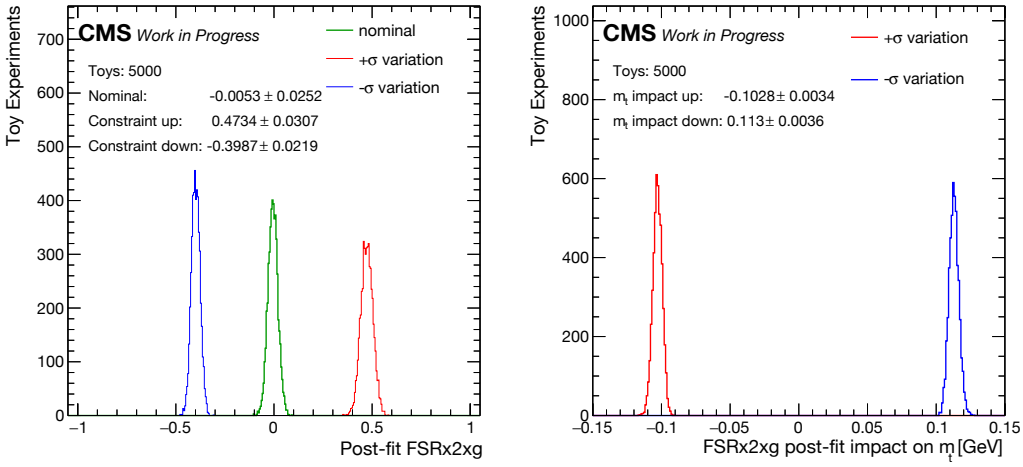

Figure E.6: FSR (b/t quark) post-fit constraints (left) and m_t impacts (right).

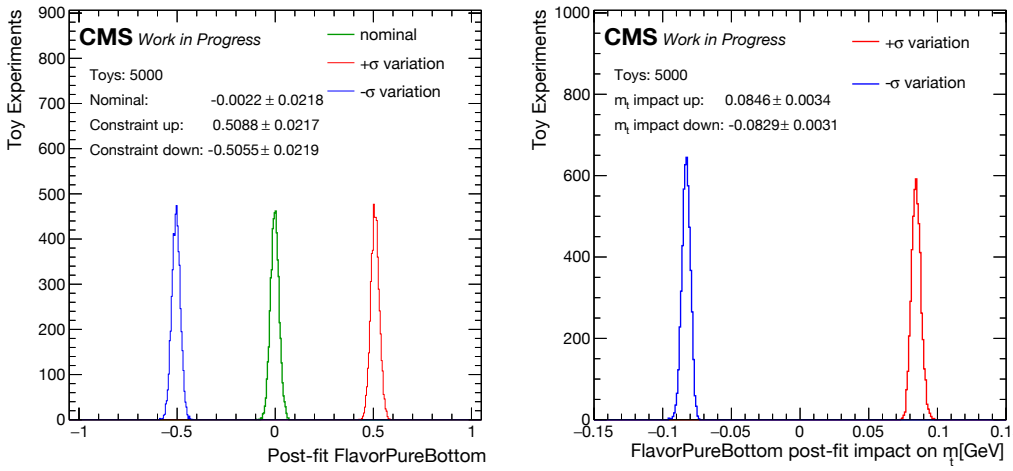

Figure E.7: B jet JEC post-fit constraints (left) and m_t impacts (right).

References

[1] M. Voutilainen, H. Siikonen, and al., "A Collection of Jet Analysis Tools Developed for the CMS". https://github.com/miquork/jetphys, 2010-. [Online; accessed 1-January-2022].

[2] CMS Collaboration, "Jet energy scale and resolution performance with 13 TeV data collected by CMS in 2016-2018", technical report, (Apr, 2020). http://cds.cern.ch/record/2715872.

[3] T. Mäkelä, "Flavour-dependent jet energy corrections and top quark mass", master's thesis (tech.), Aalto University, 2019-06-18. http://urn.fi/URN:NBN:fi:aalto-201906234090.

[4] H. Siikonen, "Top Mass Shift Caused by the Recalibration of Flavor-Dependent Jet Energy Corrections in the D0 Lepton+Jets Top Mass Measurement". https://arxiv.org/abs/2002.06073, 2020. doi:10.48550/ARXIV.2002.06073.

[5] H. Siikonen, T. Mäkelä, and M. Voutilainen, "Top mass shift resulting from the recalibration of flavor-dependent jet energy corrections in the DØ lepton+jets top mass measurement". https://arxiv.org/abs/2103.00572, 2021. doi:10.48550/ARXIV.2103.00572.

[6] CMS Collaboration, "The University of Hamburg Top Quark Mass Analysis Framework", 2022. https://gitlab.cern.ch/stadie/uhh-top-mass/-/tree/leptonjets_run2 [Online, currently accessible to CMS members; accessed 1-May-2022].

[7] H. Siikonen, "Private HitFit upgrade on CMSSW". https://github.com/errai-/cmssw/tree/HitFitUpgrade10629. [Online; accessed 1-May-2022].

[8] S. Chatrchyan et al., "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Physics Letters B* **716** (sep, 2012) 30–61, doi:10.1016/j.physletb.2012.08.021. https://doi.org/10.1016%2Fj.physletb.2012.08.021.

[9] G. Aad et al., "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Physics Letters B* **716** (sep, 2012) 1–29, doi:10.1016/j.physletb.2012.08.020.

[10] CMS Collaboration, "CMS Luminosity – Public Results". https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults. [Online; accessed 25-May-2022.].

[11] LHCb Collaboration et al., "Test of lepton universality in beauty-quark decays". https://arxiv.org/abs/2103.11769, 2021. doi:10.48550/ARXIV.2103.11769.

[12] B. Abi et al., "Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm", *Physical Review Letters* **126** (apr, 2021) doi:10.1103/physrevlett.126.141801.

[13] T. Aoyama et al., "The anomalous magnetic moment of the muon in the Standard Model", *Physics Reports* **887** (dec, 2020) 1–166, doi:10.1016/j.physrep.2020.07.006.

[14] F. Halzen and A. D. Martin, "Quarks and Leptons". John Wiley & sons, 1st edition, 1984.

[15] D. Buttazzo et al., "Investigating the near-criticality of the Higgs boson", *Journal of High Energy Physics* **2013** (dec, 2013) doi:10.1007/jhep12(2013)089.

[16] CMS Collaboration, "Projected improvement of the accuracy of top-quark mass measurements at the upgraded LHC", Technical Report CMS-PAS-FTR-13-017, CERN, Geneva, (2013). https://cds.cern.ch/record/1605627.

[17] F. Abe et al., "Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab", *Physical Review Letters* **74** (apr, 1995) 2626–2631, doi:10.1103/physrevlett.74.2626.

[18] S. Abachi et al., "Observation of the Top Quark", *Physical Review Letters* **74** (apr, 1995) 2632–2637, doi:10.1103/physrevlett.74.2632.

[19] CMS Collaboration, "Measurement of the top quark mass using proton-proton data at $\sqrt{s} = 7$ and 8 TeV", *Phys. Rev. D* **93** (Sep, 2015) 072004. 37 p, doi:10.1103/PhysRevD.93.072004, arXiv:1509.04044. Comments: Submitted to Phys. Rev. D.

[20] A. M. Sirunyan et al., "Measurement of the top quark mass with lepton+jets final states using pp collisions at $\sqrt{s} = 13\,\text{TeV}$", *The European Physical Journal C* **78** (Nov, 2018) doi:10.1140/epjc/s10052-018-6332-9.

[21] "Measurement of the Top Quark Mass from $\sqrt{s} = 7$ TeV ATLAS Data using a 3-dimensional Template Fit", technical report, CERN, Geneva, (May, 2013). All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2013-046.

[22] CMS Collaboration, "A profile likelihood approach to measure the top quark mass in the lepton+jets channel at $\sqrt{s} = 13$ TeV", Technical Report CMS-PAS-TOP-20-008, CERN, Geneva, (2022). https://cds.cern.ch/record/2806509.

[23] I. J. R. Aitchison and A. J. G. Hey, "Gauge Theories in Particle Physics: a practical introduction", volume 1-2. Taylor & Francis group, 4th edition, 2013.

[24] M. E. Peskin and D. V. Schroeder, "An Introduction to Quantum Field Theory". Addison-Wesley Publishing Co., 1st edition, 1995.

[25] A. Einstein and M. Born, "Born-Einstein Letters, 1916-1955: Friendship, Politics and Physics in Uncertain Times (Macmillan Science)". Palgrave Macmillan, 1st edition, 2004.

[26] I. Newton, "Four letters from Sir Isaac Newton to Doctor Bentley. Containing some arguments in proof of a deity.". Printed for R. and J. Dodsley, 1st edition, 1756.

[27] W. Thomson, "Notes of Lectures on Molecular Dynamics and the Wave Theory of Light delivered at the Johns Hopkins University, Baltimore, by Sir William Thomson, stenographic report by A. S. Hathaway.". Johns Hopkins University, Baltimore, 1884.

[28] Eric Drexler via Wikimedia, "Elementary Particle interactions in the Standard Model", 2014. [Online; accessed 11-May-2022].

[29] S. L. Glashow, "The renormalizability of vector meson interactions", *Nuclear Physics* **10** (1959) 107–117.

[30] S. Weinberg, "A Model of Leptons", *Phys. Rev. Lett.* **19** (Nov, 1967) 1264–1266, doi:10.1103/PhysRevLett.19.1264.

[31] A. Salam, "Weak and Electromagnetic Interactions", *Conf. Proc. C* **680519** (1968) 367–377, doi:10.1142/9789812795915$_0$034.

[32] T. Sjostrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual", *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175.

[33] T. Sjostrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1", *Comput. Phys. Commun.* **178** (2008) 852–867, doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820.

[34] T. Sjostrand et al., "An Introduction to PYTHIA 8.2", *Comput. Phys. Commun.* **191** (2015) 159–177, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.

[35] ATLAS Collaboration, CDF Collaboration, CMS Collaboration, D0 Collaboration Collaboration, "First combination of Tevatron and LHC measurements of the top-quark mass", arXiv:1403.4427.

[36] D0 Collaboration, "Precision measurement of the top-quark mass in lepton+jets final states", *Phys. Rev.* **D91** (2015), no. 11, 112003, doi:10.1103/PhysRevD.91.112003, arXiv:1501.07912.

[37] A. H. Hoang et al., "The MSR Mass and the $\mathcal{O}\left(\Lambda_{\rm QCD}\right)$ Renormalon Sum Rule", *Journal of High Energy Physics* **2018** (apr, 2018) doi:10.1007/jhep04(2018)003.

[38] M. Butenschoen et al., "Top Quark Mass Calibration for Monte Carlo Event Generators", *Phys. Rev. Lett.* **117** (2016), no. 23, 232001, doi:10.1103/PhysRevLett.117.232001, arXiv:1608.01318.

[39] ATLAS Collaboration Collaboration, "A precise interpretation for the top quark mass parameter in ATLAS Monte Carlo simulation", technical report, CERN, Geneva, (Jul, 2021). https://cds.cern.ch/record/2777332.

[40] J. R. Espinosa et al., "The cosmological Higgstory of the vacuum instability", *Journal of High Energy Physics* **2015** (sep, 2015) doi:10.1007/jhep09(2015)174.

[41] CERN, "The history of CERN". https://home.cern/about/who-we-are/our-history. [Online; accessed 25-May-2022.].

[42] L. Evans and P. Bryant, "LHC Machine", *Journal of Instrumentation* **3** (aug, 2008) S08001–S08001, doi:10.1088/1748-0221/3/08/s08001.

[43] CMS Collaboration, "The CMS experiment at the CERN LHC. The Compact Muon Solenoid experiment", *JINST* **3** (2008) S08004. 361 p, doi:10.1088/1748-0221/3/08/S08004. Also published by CERN Geneva in 2010.

[44] CERN, "LHC machine outreach". http://lhc-machine-outreach.web.cern.ch/lhc-machine-outreach/lhc_in_pictures.htm, 2015. [Online; accessed 20-December-2021].

[45] J.-L. Caron, "Geographic location of CERN.. Carte geographique de l'implantation du CERN.". http://cds.cern.ch/record/841532, Mar, 1998. AC Collection. Legacy of AC. Pictures from 1992 to 2002.

[46] P. Bambade et al., "The International Linear Collider: A Global Project". https://arxiv.org/abs/1903.01629, 2019. doi:10.48550/ARXIV.1903.01629.

[47] CERN, "Longer term LHC schedule". https://lhc-commissioning.web.cern.ch/schedule/LHC-long-term.htm. [Online; accessed 25-May-2022.].

[48] S. Fartoukh et al., "LHC Configuration and Operational Scenario for Run 3", technical report, CERN, Geneva, (Nov, 2021). https://cds.cern.ch/record/2790409.

[49] G. Rumolo et al., "Expected performance in the injectors at 25 ns without and with Linac4", doi:10.5170/CERN-2014-006.17.

[50] CMS Tracker Group Collaboration, "The CMS Phase-1 Pixel Detector Upgrade", *JINST* **16** (Dec, 2020) P02027. 84 p, doi:10.1088/1748-0221/16/02/P02027, arXiv:2012.14304.

[51] D. Barney, "CMS Detector Slice". https://cds.cern.ch/record/2120661, Jan, 2016. CMS Collection.

[52] CMS Collaboration, "ECAL Technical Design Report (TDR) Figures from Chapter 3". http://cds.cern.ch/record/1327664, Dec, 1997. CMS Collection.

[53] CMS Collaboration, "Electromagnetic physics objects commissioning with first LHC data", technical report, CERN, Geneva, (2010). http://cds.cern.ch/record/1247384.

[54] CMS CollaborationV. Karimäki, et al., "The CMS tracker system project: Technical Design Report". Technical design report. CMS. CERN, Geneva, 1997.

[55] CMS Collaboration, "The CMS electromagnetic calorimeter project: Technical Design Report". Technical design report. CMS. CERN, Geneva, 1997. https://cds.cern.ch/record/349375.

[56] CMS Collaboration, "The CMS hadron calorimeter project: Technical Design Report". Technical design report. CMS. CERN, Geneva, 1997.

[57] CMS Collaboration, "The CMS magnet project: Technical Design Report". Technical design report. CMS. CERN, Geneva, 1997.

[58] CMS Collaboration, "The performance of the CMS muon detector in proton-proton collisions at $\sqrt{s} = 7$ TeV at the LHC", *JINST* **8** (Jun, 2013) P11002. 101 p, doi:10.1088/1748-0221/8/11/P11002, arXiv:1306.6905. Comments: Submitted to JINST.

[59] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET", Technical Report CMS-PAS-PFT-09-001, CERN, 2009. Geneva, (Apr, 2009). https://cds.cern.ch/record/1194487.

[60] CMS Collaboration, "Commissioning of the Particle-Flow Reconstruction in Minimum-Bias and Jet Events from p p Collisions at 7 TeV", CMS Physics Analysis Summary CMS-PAS-PFT-10-002, CERN, (2010). http://cdsweb.cern.ch/record/1279341.

[61] A. Sirunyan et al., "Particle-flow reconstruction and global event description with the CMS detector", *Journal of Instrumentation* **12** (oct, 2017) P10003–P10003, doi:10.1088/1748-0221/12/10/p10003.

[62] Particle Data Group Collaboration, "Review of Particle Physics", *Phys. Rev. D* **98** (Aug, 2018) 030001, doi:10.1103/PhysRevD.98.030001.

[63] V. Khachatryan et al., "The CMS trigger system", *Journal of Instrumentation* **12** (jan, 2017) P01020–P01020, doi:10.1088/1748-0221/12/01/p01020.

[64] CMS Collaboration, "Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC", *Journal of Instrumentation* **16** (may, 2021) P05014, doi:10.1088/1748-0221/16/05/p05014.

[65] "Performance of photon reconstruction and identification with the CMS detector in proton-proton collisions at $\sqrt{s}$= 8 TeV", *Journal of Instrumentation* **10** (aug, 2015) P08010–P08010, doi:10.1088/1748-0221/10/08/p08010.

[66] CMS Collaboration, "CMS Software". https://github.com/cms-sw/cmssw, 2022. [Online; accessed 22-May-2022].

[67] S. Bazrafkan et al., "Deep Learning Based Computed Tomography Whys and Wherefores". https://arxiv.org/abs/1904.03908, 2019. doi:10.48550/ARXIV.1904.03908.

[68] ATLAS Collaboration Collaboration, V. Begy, M.-S. Barisits, and M. Lassnig, "Forecasting Network Throughput of Remote Data Access in

Computing Grids", technical report, CERN, Geneva, (Jun, 2018). https://cds.cern.ch/record/2621616.

[69] R. Brun and F. Rademakers, "ROOT: An object oriented data analysis framework", 1997. doi:10.1016/S0168-9002(97)00048-X.

[70] J. Bellm et al., "Herwig 7.0/Herwig++ 3.0 release note", *Eur. Phys. J.* **C76** (2016), no. 4, 196, doi:10.1140/epjc/s10052-016-4018-8, arXiv:1512.01178.

[71] C. Oleari, "The POWHEG-BOX", *Nucl. Phys. Proc. Suppl.* **205-206** (2010) 36–41, doi:10.1016/j.nuclphysbps.2010.08.016, arXiv:1007.3893.

[72] S. Alioli et al., "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.

[73] 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** (May, 2014) 079. 157 p, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301. Comments: 158 pages, 27 figures; a few references have been added.

[74] T. Gleisberg et al., "Event generation with SHERPA 1.1", *JHEP* **02** (2009) 007, doi:10.1088/1126-6708/2009/02/007, arXiv:0811.4622.

[75] CMS HCAL/ECAL Collaboration, "The CMS Barrel Calorimeter Response to Particle Beams from 2 to 350 GeV/c", technical report, CERN, Geneva, (Jun, 2008). https://cds.cern.ch/record/1166316.

[76] A. Sirunyan et al., "Calibration of the CMS hadron calorimeters using proton-proton collision data at $\sqrt{s} = 13$ TeV", *Journal of Instrumentation* **15** (may, 2020) P05002–P05002, doi:10.1088/1748-0221/15/05/p05002.

[77] GEANT4 Collaboration, "GEANT4: A Simulation toolkit", *Nucl. Instrum. Meth.* **A506** (2003) 250–303, doi:10.1016/S0168-9002(03)01368-8.

[78] CMS Collaboration, "Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements", Technical Report CERN-EP-2019-007, CERN, Geneva, (2018). https://cds.cern.ch/record/2669320.

[79] CMS Collaboration, "Development and validation of HERWIG 7 tunes from CMS underlying-event measurements", *Eur. Phys. J. C* **81** (Nov, 2020) 312. 64 p, doi:10.1140/epjc/s10052-021-08949-5, arXiv:2011.03422. Replaced with the published version. Added the journal reference and the

DOI. All the figures and tables can be found at http://cms-results.web.cern.ch/cms-results/public-results/publications/GEN-19-001 (CMS Public Pages).

[80] J. Alwall et al., "MadGraph 5 : Going Beyond", *JHEP* **06** (2011) 128, doi:10.1007/JHEP06(2011)128, arXiv:1106.0522.

[81] NNPDF Collaboration, "Parton distributions from high-precision collider data", arXiv:1706.00428.

[82] W. Herr and B. Muratori, "Concept of luminosity", doi:10.5170/CERN-2006-002.361.

[83] M. Czakon, P. Fiedler, and A. Mitov, "Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$", *Phys. Rev. Lett.* **110** (2013) 252004, doi:10.1103/PhysRevLett.110.252004, arXiv:1303.6254.

[84] ATLAS-CMS Collaboration, "NLO single-top channel cross sections". https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SingleTopRefXsec. [Online; accessed 25-May-2022.].

[85] Sailko via Wikimedia, "File:Ginevra, jet d'eau, arcobaleno", 2014. https://commons.wikimedia.org/wiki/File:Ginevra,_jet_d%27eau,_arcobaleno.JPG [Online; accessed 11-May-2022].

[86] CMS, "Displays of candidate events in the search for new heavy resonances decaying to dibosons in the all-jets final state in the CMS detector". http://cds.cern.ch/record/2809446, May, 2022. CMS Collection.

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

[88] CMS Collaboration, "Pileup Removal Algorithms", technical report, CERN, Geneva, (2014). https://cds.cern.ch/record/1751454.

[89] D. Bertolini et al., "Pileup per particle identification", *Journal of High Energy Physics* **2014** (oct, 2014) doi:10.1007/jhep10(2014)059.

[90] CMS Collaboration, "Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$ using the CMS detector", *JINST* **14** (2019) P07004, doi:10.1088/1748-0221/14/07/P07004, arXiv:1903.06078.

[91] CMS Collaboration, "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV", *JINST* **12** (2017) P02014, doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.

[92] CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", *JINST* **13** (2018) P05011, doi:10.1088/1748-0221/13/05/P05011, arXiv:1712.07158.

[93] CMS Collaboration, "Performance of quark/gluon discrimination in 8 TeV pp data", Technical Report CMS-PAS-JME-13-002, CERN, Geneva, (2013). https://cds.cern.ch/record/1599732.

[94] CMS Collaboration, "Performance of quark/gluon discrimination in 13 TeV data",. https://cds.cern.ch/record/2234117.

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

[96] CMS Collaboration, "Jet energy scale and resolution measurement with Run 2 Legacy Data Collected by CMS at 13 TeV", technical report, (Sep, 2021). https://cds.cern.ch/record/2792322.

[97] CDF Collaboration, "Generic Jet Scale Corrections for Run II", technical report, (Jan, 2005). https://inspirehep.net/files/796e1e8cd6a4cb5d24a88c93ed84fdf0.

[98] CMS Collaboration, "CMS JEC Database". https://github.com/cms-jet/JECDatabase, 2022. [Online; accessed 22-May-2022].

[99] CMS Collaboration, "CMS Run 2 Ultra Legacy Global Fit of Jet Energy Corrections". https://twiki.cern.ch/twiki/pub/CMSPublic/PhysicsResultsDP2021/GlobalFit_Run2.pdf. [Online; accessed 25-May-2022.].

[100] M. Wassmer and H. Siikonen, "A bug fix for the propagation of jet smearing to MET". https://github.com/cms-sw/cmssw/pull/36797. [Online; accessed 1-May-2022].

[101] H. Siikonen, "A critical bugfix for jet energy corrections". https://github.com/cms-sw/cmssw/pull/36559. [Online; accessed 1-May-2022].

[102] D0 Collaboration, "Combination of D0 measurements of the top quark mass", *Phys. Rev.* **D95** (2017), no. 11, 112004, doi:10.1103/PhysRevD.95.112004, arXiv:1703.06994.

[103] DØ Collaboration, "Statement on the D0 top mass determination in the lepton + jets channel". https://www-d0.fnal.gov/Run2Physics/WWW/results/final/TOP/T14E/D0_top-JES_statement.pdf, 2020. [Online; accessed 25-May-2022. Requires a login e.g. on CERN credentials.].

[104] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual", *Eur. Phys. J.* **C72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

[105] M. Bahr et al., "Herwig++ Physics and Manual", *Eur. Phys. J.* **C58** (2008) 639–707, doi:10.1140/epjc/s10052-008-0798-9, arXiv:0803.0883.

[106] H. Siikonen, "A Multi-Purpose Software Handle for MC Event Generators in High Energy Physics, the *greedy_pythia6* code package". https://gitlab.cern.ch/hsiikone/gen_handle, 2016-. [Online; accessed 13-March-2020].

[107] H. Siikonen, "A Collection of Analysis Tools for MC Generated Proton-Proton Collisions, the *D0mtshift* code package". https://gitlab.cern.ch/hsiikone/gen_analysis/-/tree/master/D0mtshift, 2016-. [Online; accessed 13-March-2020].

[108] CMS Collaboration, "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV", *JINST* **13** (2018) P06015, doi:10.1088/1748-0221/13/06/P06015, arXiv:1804.04528.

[109] A. Bodek et al., "Extracting muon momentum scale corrections for hadron collider experiments", *The European Physical Journal C* **72** (Oct, 2012) doi:10.1140/epjc/s10052-012-2194-8.

[110] Verzetti, Mauro, "Machine learning techniques for jet flavour identification at CMS", *EPJ Web Conf.* **214** (2019) 06010, doi:10.1051/epjconf/201921406010.

[111] E. Bols et al., "Jet Flavour Classification Using DeepJet", *JINST* **15** (2020) P12012, doi:10.1088/1748-0221/15/12/P12012, arXiv:2008.10519.

[112] CMS Collaboration, "Performance of the DeepJet b tagging algorithm using 41.9/fb of data from proton-proton collisions at 13 TeV with Phase 1 CMS detector", CMS Detector Performance Note CMS-DP-2018-058, (2018). http://cds.cern.ch/record/2646773.

[113] CMS Collaboration, "B-tagging performance of the CMS Legacy dataset 2018.", technical report, (Mar, 2021). https://cds.cern.ch/record/2759970.

[114] CMS Collaboration Collaboration, "Jet algorithms performance in 13 TeV data", technical report, CERN, Geneva, (2017). https://cds.cern.ch/record/2256875.

[115] CMS Collaboration Collaboration, "Mitigation of anomalous missing transverse momentum measurements in data collected by CMS at sqrt(s) = 13 TeV during the LHC Run 2",.

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

[117] CMS Collaboration, "Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JINST* **15** (Jun, 2020) P10017. 63 p, doi:10.1088/1748-0221/15/10/P10017, arXiv:2006.10165. Replaced with the published version. Added the journal reference and the DOI. All the figures and tables can be found at http://cms-results.web.cern.ch/cms-results/public-results/publications/TRG-17-001 (CMS Public Pages).

[118] CMS Collaboration, "Performance of the CMS muon trigger system in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JINST* **16** (Feb, 2021) P07001. 52 p, doi:10.1088/1748-0221/16/07/P07001, arXiv:2102.04790. Replaced with the published version. Added the journal reference and the DOI. All the figures and tables can be found at http://cms-results.web.cern.ch/cms-results/public-results/publications/MUO-19-001 (CMS Public Pages).

[119] N. Kidonakis, "Differential and total cross sections for top pair and single top production", in *Proceedings, 20th International Workshop on Deep-Inelastic Scattering and Related Subjects (DIS 2012)*, pp. 831–834. 2012. arXiv:1205.3453. [,831(2012)]. doi:10.3204/DESY-PROC-2012-02/251.

[120] CMS Collaboration, "Measurement of differential top-quark pair production cross sections in the lepton+jets channel in pp collisions at 8 TeV", Technical Report CMS-PAS-TOP-12-027, CERN, Geneva, (2013).

[121] CMS Collaboration, "Measurement of the differential ttbar cross section in the dilepton channel at 8 TeV", Technical Report CMS-PAS-TOP-12-028, CERN, Geneva, (2013).

[122] CMS Collaboration, "Measurement of the inclusive and differential tt production cross sections in lepton + jets final states at 13 TeV", Technical Report CMS-PAS-TOP-16-008, CERN, Geneva, (2016). https://cds.cern.ch/record/2141097.

[123] CMS Collaboration, "Top quark pair differential cross sections in the dilepton channel at 13 TeV", Technical Report CMS-PAS-TOP-16-011, CERN, Geneva, (2016). https://cds.cern.ch/record/2140061.

[124] S. S. Snyder, "Measurement of the Top Quark Mass at D0", (5, 1995) doi:10.2172/1422822.

[125] "Problem with electron resolutions". https://github.com/cms-sw/cmssw/blob/master/TopQuarkAnalysis/TopHitFit/data/resolution/tqafElectronResolution.txt#L37-L63. Accessed: 2020-03-30.

[126] "Problem with light quark jet resolutions". https://github.com/cms-sw/cmssw/blob/master/TopQuarkAnalysis/TopHitFit/data/resolution/tqafUdscJetResolution.txt#L89-L103. Accessed: 2020-03-30.

[127] W. T. Giele, D. A. Kosower, and P. Z. Skands, "Higher-order corrections to timelike jets", *Physical Review D* **84** (Sep, 2011) doi:10.1103/physrevd.84.054003.

[128] S. Mrenna and P. Skands, "Automated parton-shower variations in pythia 8", *Physical Review D* **94** (Oct, 2016) doi:10.1103/physrevd.94.074005.

[129] 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, doi:10.1140/epjc/s10052-011-1557-x, arXiv:1102.4748.

[130] ALEPH Collaboration, "Study of the fragmentation of b quarks into B mesons at the Z peak", *Phys. Lett. B* **512** (2001) 30–48, doi:10.1016/S0370-2693(01)00690-6, arXiv:hep-ex/0106051.

[131] S. Dulat et al., "New parton distribution functions from a global analysis of quantum chromodynamics", *Physical Review D* **93** (Feb, 2016) doi:10.1103/physrevd.93.033006.

[132] L. A. Harland-Lang et al., "Parton distributions in the LHC era: MMHT 2014 PDFs", *The European Physical Journal C* **75** (May, 2015) doi:10.1140/epjc/s10052-015-3397-6.

[133] P. Z. Skands and D. Wicke, "Non-perturbative QCD effects and the top mass at the Tevatron", *Eur. Phys. J. C* **52** (2007) 133, doi:10.1140/epjc/s10052-007-0352-1, arXiv:hep-ph/0703081.

[134] CMS Collaboration, "CMS PYTHIA 8 colour reconnection tunes based on underlying-event data",.

[135] CMS Collaboration, "Measurement of the top quark mass in the dileptonic $t\bar{t}$ decay channel using the mass observables $M_{b\ell}$, M_{T2}, and $M_{b\ell\nu}$ in pp collisions at $\sqrt{s} = 8$ TeV", *Physical Review D* **96** (Aug, 2017) doi:10.1103/physrevd.96.032002.

[136] CMS Collaboration, "The Combine Tool". http://cms-analysis.github.io/HiggsAnalysis-CombinedLimit/. [Online; accessed 1-May-2022].

[137] CMS Collaboration Collaboration, "Measurement of the top quark pole mass using $t\bar{t}$+jet events in the dilepton final state at $\sqrt{s} = 13$ TeV", Technical Report CMS-PAS-TOP-21-008, CERN, Geneva, (2022). https://cds.cern.ch/record/2804936.

[138] R. Barlow and C. Beeston, "Fitting using finite Monte Carlo samples", *Computer Physics Communications* **77** (1993), no. 2, 219 – 228, doi:https://doi.org/10.1016/0010-4655(93)90005-W.

[139] CMS Collaboration Collaboration, "Measurement of the jet mass distribution and top quark mass in hadronic decays of boosted top quarks in pp collisions at $\sqrt{s} = 13$ TeV", Technical Report CMS-PAS-TOP-21-012, CERN, Geneva, (2022). http://cds.cern.ch/record/2809549.

[140] ATLAS-CMS Collaboration, "$t\bar{t}$ cross sections". https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TtbarNNLO. [Online; accessed 25-May-2022.].

[141] O. Aberle et al., "High-Luminosity Large Hadron Collider (HL-LHC): Technical design report". CERN Yellow Reports: Monographs. CERN, Geneva, 2020. https://cds.cern.ch/record/2749422.

[142] N. Kidonakis, "Next-to-next-to-leading soft-gluon corrections for the top quark cross section and transverse momentum distribution", *Phys. Rev. D* **82** (2010) 114030, doi:10.1103/PhysRevD.82.114030, arXiv:1009.4935.

[143] K. Melnikov and F. Petriello, "Electroweak gauge boson production at hadron colliders through $O(\alpha_s^2)$", *Phys. Rev. D* **74** (2006) 114017, doi:10.1103/PhysRevD.74.114017, arXiv:hep-ph/0609070.

[144] J. M. Campbell and R. K. Ellis, "MCFM for the Tevatron and the LHC", *Nucl. Phys. Proc. Suppl.* **205-206** (2010) 10, doi:10.1016/j.nuclphysbps.2010.08.011, arXiv:1007.3492.

[145] J. M. Campbell, J. W. Huston, and W. J. Stirling, "Hard Interactions of Quarks and Gluons: A Primer for LHC Physics", *Rept. Prog. Phys.* **70** (2007) 89, doi:10.1088/0034-4885/70/1/R02, arXiv:hep-ph/0611148.

[146] M. Czakon and A. Mitov, "Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders", *Comput. Phys. Commun.* **185** (2014) 2930, doi:10.1016/j.cpc.2014.06.021, arXiv:1112.5675.

[147] J. Ohnemus, "Order $\alpha^{-}s$ calculations of hadronic $W^{\pm}\gamma$ and $Z\gamma$ production", *Phys. Rev.* **D47** (1993) 940–955, doi:10.1103/PhysRevD.47.940.

[148] U. Baur, T. Han, and J. Ohnemus, "QCD corrections to hadronic $W\gamma$ production with nonstandard $WW\gamma$ couplings", *Phys. Rev.* **D48** (1993) 5140–5161, doi:10.1103/PhysRevD.48.5140, arXiv:hep-ph/9305314.

[149] N. Kidonakis, "Top Quark Production", arXiv:1311.0283.

[150] D0 Collaboration, "Precise measurement of the top-quark mass from lepton+jets events at D0", *Phys. Rev.* **D84** (2011) 032004, doi:10.1103/PhysRevD.84.032004, arXiv:1105.6287.

[151] S. Weisberg, "Applied Linear Regression". John Wiley & Sons, 3rd edition, 2005.