VNIVERSITAT DE VALÈNCIA

DEPARTAMENT DE FÍSICA ATÒMICA, MOLECULAR I NUCLEAR

PROGRAMA DE DOCTORAT EN FÍSICA

TESI DOCTORAL

Precision measurements in boosted top quarks in the ATLAS detector

Javier Alberto Aparisi Pozo

Sota la supervisió de

Marcel Vos i Santiago González de la Hoz

València, Juny 2022

Dr. Santiago González de la Hoz, Catedràtic d'Universitat del Departament de Física Atómica, Molecular i Nuclear de la Universitat de València.

Dr. Marcel Vos, Investigador Científic del Consell Superior d'Investigacions Científiques.

CERTIFIQUEN que aquesta memòria, *Precision measurements in boosted top quarks in the ATLAS detector*, resumeix el treball de recerca realitzat, sota la seua direcció i tutela al Departament de Física Atòmica, Molecular i Nuclear de la Universitat de València, per Don Javier Alberto Aparisi Pozo i constitueix la seua Tesi per optar al Títol de Doctor per la Universitat de València després de completar els estudis en el Doctorat en Física.

I per a que així conste, en compliment de la legistlació vigent, firmen el present certificat a València, a 10 de juny de 2022.

Dr. Santiago Gonzalez de la Hoz Dr. Marcel Vos

Els estudis recollits en la present tesi doctoral s'han realitzat en el Insititut de Física Corpuscular (IFIC) a València (Espanya).

El IFIC és un centre mixt del Consell Superior d'Investigacions Científiques (CSIC) i de la Universitat de València (UV). Amb al voltant de 280 membres i un reconegut impacte internacional, és un dels centres més importants de Física d'Altes Energies a Espanya i posseeix l'acreditació Severo Ochoa, el qual el distingeix com a Centre d'Excel·lència. Aquesta acreditació reconeix l'excel·lència i les contribucions científiques que fan els centres i les unitats a nivell nacional i internacional, el seu impacte empresarial i social i la seua capacitat per atraure talent. El treball desenvolupat en aquesta tesi ha sigut financiat amb l'ajuda SEV-2014-0398-17-1 per a la formació de personal investigador.

Contents

1 The Standard Model of particle physics **1**

1.1 Introduction 2

1.2 The composition 5

1.3 The structure 8

1.3.1 Quantum electrodynamics 8

1.3.2 Weak interaction and unification 9

1.3.3 Spontaneous symmetry breaking 14

1.3.4 Quantum chromodynamics 17

1.4 Calculability of the theory 19

1.4.1 Ultraviolet divergencies and running 22

1.4.2 Renormalization schemes 25

1.4.3 Infrared divergencies 28

1.5 QCD phenomenology 30

1.5.1 Short- and long-distance domains 30

1.5.2 Non-perturbative effects 32

1.6 Physics at hadron colliders 33

1.6.1 Factorization theorem . . . 35
1.6.2 Fixed-order calculations . . . 38
1.6.3 Resummation of logarithmic terms . . . 39
1.6.4 Multi-purpose Monte Carlo generators . . . 40
1.7 Effective field theories . . . 46

2 Top quark physics **51**
2.1 Mass and width . . . 52
2.2 Standard Model consistency and tests . . . 53
2.3 Top quark properties at hadron colliders . . . 57
2.3.1 Single and top pair production . . . 57
2.3.2 $t\bar{t}$ asymmetry . . . 59
2.3.3 Four-quark operators in effective field theory . . . 63
2.3.4 Top mass measurements . . . 64
2.4 Top quark mass interpretation . . . 70

3 The ATLAS experiment at the Large Hadron Collider **75**
3.1 The Large Hadron Collider . . . 76
3.1.1 Collider coordinates and kinematic variables . . . 80
3.2 The ATLAS detector . . . 82
3.2.1 Inner detector . . . 84
3.2.2 Calorimeter system . . . 86
3.2.3 Muon spectrometer . . . 88
3.3 The ATLAS data management . . . 89

3.3.1 Trigger system . . . 89
3.3.2 Data storage . . . 90
3.3.3 Object reconstruction . . . 92

4 Jet physics and boosted objects **95**

4.1 Introduction . . . 96
4.2 Jet definition and identification . . . 98
4.2.1 Jets in the boosted regime . . . 100
4.2.2 Jet constituents . . . 102
4.2.3 Jet tagging . . . 103
4.3 Soft-component mitigation techniques . . . 104
4.4 Jet calibration . . . 107
4.4.1 The Forward-Folding method . . . 110
4.4.2 In situ calibration from $t\bar{t}$ events . . . 112

5 Interpretation of the top quark mass in ATLAS Monte Carlo generators **123**

5.1 Introduction . . . 124
5.2 NLL factorization for boosted top quark jets . . . 125
5.3 ATLAS Monte Carlo generators . . . 132
5.4 Methodology . . . 136
5.4.1 Generation of theoretical predictions and simulation . . . 136
5.4.2 Event selection and jet reconstruction . . . 138
5.4.3 The jet mass in high-p_{T} top decays . . . 139

5.4.4 Template-fitting procedure . . . 141
5.4.5 Theory and simulation compatibility . . . 143
5.5 Uncertainties . . . 148
5.5.1 Theory uncertainties . . . 148
5.5.2 Method uncertainties . . . 151
5.5.3 Underlying event and color reconnection modelling . . . 153
5.5.4 Summary of uncertainties . . . 156
5.6 Results . . . 157
5.6.1 Featuring the top Monte Carlo mass . . . 158
5.6.2 Calibration of the top Monte Carlo mass . . . 163
5.6.3 Stability of the results . . . 169
5.7 Discussion of the results . . . 172

6 SMEFT bounds from the $t\bar{t}$ charge asymmetry at the LHC 179

6.1 Introduction . . . 179
6.2 Analysis strategy . . . 182
6.2.1 Event selection and reconstruction . . . 182
6.2.2 Unfolding to parton level . . . 187
6.2.3 Results . . . 188
6.3 SMEFT interpretation . . . 191
6.3.1 Methodology . . . 193
6.3.2 Limits on dimension-six operators . . . 196

7 Determination of the running bottom quark mass from Higgs

boson decay rates **203**

7.1 Introduction 204

7.2 Low-energy measurements 204

7.3 Bottom quark mass from Higgs decay rates 207

7.3.1 Dependence of Higgs decay rates on the bottom quark mass 208

7.3.2 Numerical predictions and theory uncertainties 209

7.3.3 Extraction of the bottom quark mass 211

7.4 The running of the bottom quark mass 212

7.4.1 The anomalous mass dimension 212

7.4.2 Testing the running hypothesis 215

7.5 Discussion and prospects 217

8 Summary and conclusions **219**

A SMEFT bounds from differential $m_{t\bar{t}}$ measurements of A_C **229**

Resum **237**

Acronyms **263**

List of Publications **267**

List of Figures **269**

List of Tables **289**

Bibliography **293**

Chapter 1

The Standard Model of particle physics

The idea that matter consists of smaller, limited and primary building blocks dates back to the 6th century BC. It was not until the 19th century when empirical observations were used to shed some light on this: John Dalton and his contemporaries, through his work on stoichiometry, believed that atoms were the fundamental particles of nature. Shortly after, the observed cathode rays radiation by J.J. Thompson suggested that, in fact, the atom was a conglomerate of even smaller particles: a negatively charged particle called electron, and a positively charged yet unknown substance. In the early 1900s, most physicists believed that physics was complete with the description given by classical mechanics, thermodynamics and the Maxwell theory of electromagnetism:

> "There is nothing new to be discovered in physics now. All that remains is more and more precise measurements". Lord Kelvin, 1900.

However, nothing was known about the structure of atoms and nuclei, among others. Nobody anticipated the revolution that particle physics would undergo in

the next 100 years. Two open questions, the unexplained energy distribution in blackbody radiation and the Michelson-Morley experiment, a failed attempt to detect the existence of the luminiferous aether[1], led to two of the most important theories formulated in the last century: quantum mechanics and general relativity.

The *Standard Model* (SM) was the culmination of the revolution in particle physics, incorporating quantum mechanics and relativity in a *Quantum Field Theory* (QFT). All elementary particles and the fundamental interactions (excluding gravity) discovered so far are described. The SM has been successfully tested at an impressive level of accuracy over the decades and provides our best fundamental understanding of the phenomenology of particle physics.

In this chapter, a brief overview of the SM is given. We shall restrict ourselves to the main pieces needed to understand the theoretical basis of the work presented in this thesis. After a general introduction of the content and dynamics (or interactions) of the SM, we will see how to exploit its predictive power with a special focus on the context of modern particle colliders, which represent the ideal laboratory to break through the limits of the current knowledge.

1.1 Introduction

In the SM, elementary particles adopt the interpretation of dimension-less particles described by fields, *i.e.* operator functions on the spacetime employed to compute probability distributions. The quantum states of the fields are determined by a set of quantum numbers. Regarding interactions, only three of the four fundamental forces known are accounted for: the strong, weak and electromagnetic interactions. The gravitational force has not been successfully

[1] A hypothesized medium permeating space that was supposed to be the carrier of light waves. The experiment compared the speed of light in perpendicular directions in an attempt to detect the relative motion of matter through the stationary luminiferous aether.

described in terms of a quantum field theory, so a full-fledged theory of everything remains unknown. Particles and their dynamics are encoded in the Lagrangian density $\mathcal{L}$, a scalar function which solutions (equations of motion) are found by requiring the action S to be stationary:

$$S = \int d^x \mathcal{L}(x) \rightarrow \delta S = 0 \ .$$

The free Lagrangian of the SM accounts for non-interacting particles, and the solutions can be obtained analytically: the propagators. However, we know that matter changes, so particles do interact. The interactions between particles are closely related to the concept of symmetries in QFTs. We say that a given transformation of the fields is a symmetry of the physical system when the Lagrangian density remains unaltered after the transformation. In particular, the Poincaré group represents a set of symmetric transformations (Lorentz transformations, rotations and translations) in the Minkowski spacetime. In quantum mechanics, a continuous symmetry group is represented by unitary transformations U on the states of the system. We shall restrict this discussion to the dimension one unitary abelian group $U(N = 1)$ and the special unitary non-Abelian group $SU(N)$ (with $N = 2, 3$).

In the SM the phase of the fields ϕ is an arbitrary choice. Hence the Lagrangian is required to be invariant under phase transformations of the fields. This is carried out by means of a unitary transformation of the form:

$$\phi(x) \xrightarrow{U(N)} \phi'(x) = U(N)\phi(x) = e^{iT_a\theta_a}\phi(x) \ .$$

The transformation $U(N)$ depends on the product of the $N^2 - 1$ generators T_a of the fundamental representation of the symmetry group and the corresponding real parameters, θ_a. If θ_a are constant, this transformation gives rise to a global

symmetry of the system, whilst the symmetry is said to be local for space-dependent parameters $\theta = \theta(x)$.

The free SM Lagrangian is invariant under global transformations of the fields, but it is no longer invariant under local transformations. The gauge principle specifies a procedure to make it invariant by introducing additional terms in the Lagrangian. These additional terms, the gauge fields, couple to the free fields of the theory, allowing the associated particles to talk to each other through the exchange of a quantum excitation of the gauge field. Formally, the interaction between a gauge field and a matter field can be seen as a rotation in the space of the associated symmetry group. The nature of the symmetry determines the features of the gauge fields and, therefore, the corresponding interaction. For instance, the number of gauge fields for given interaction corresponds to the number of generators of the fundamental representation of the symmetry group. Furthermore, they will self-interact if the symmetry group is non-abelian. In this case, the strength of the interaction will become stronger as the involved particles are pulled apart and weaker as they get close by.

As we will see in the following sections, the choice of the symmetries to describe the three interactions mentioned above did not come out of the blue: it was rather driven by the experimental knowledge available when developing the SM.

1.2 The composition

The SM Lagrangian distinguishes two types of particles of different nature: fermions and bosons. Fermions are 1/2-spin particles described by Dirac fields and hence obey Fermi-Dirac statistics. This means that identical fermions at the same energy cannot occupy the same quantum state. In contrast, bosons are mediators of the interactions represented by integer spin gauge fields. They obey Bose-Einstein statistics instead so an infinite number of identical bosons can share the same quantum state and energy.

Fermions comprise two kinds of particles in turn, quarks (q) and leptons (l), which are organized in a three-fold family structure:

$$\begin{bmatrix} \nu_e & u \\ e^- & d' \end{bmatrix}, \begin{bmatrix} \nu_\mu & c \\ \mu^- & s' \end{bmatrix}, \begin{bmatrix} \nu_\tau & t \\ \tau^- & b' \end{bmatrix}. \tag{1.1}$$

There is no deep rationale behind this organization: it simply permits to explain the behaviour of the fermionic matter as observed. Six different types or flavours of leptons and quarks are organized in pairs of up- and down-type fermions, for each family. They are distinguished by their associated quantum numbers and mass. The down-type leptons are the *electron* (e), the *muon* (μ) and the *tau* (τ), electrically charged. The associated up-type partners are the *electron neutrino* (ν_e), the *muon neutrino* (ν_μ) and the *tau neutrino* (ν_τ), all them electrically neutral. Charged-leptons feel the electromagnetic and weak forces, whilst neutrinos only participate in the weak one. In contrast, quarks feel all fundamental forces including the strong interaction. The up-type quarks are *up (u)*, *charm (c)* and *top (t)* with positive fractional charge. The *down (d)*, *strange (c)* and *bottom (b)* quarks are the down-partners with negative fractional charge. Only the lightest fermions, belonging to the first family, constitute stable matter,

and the more massive fermions of the second and third families eventually decay into their lighter partners.

For each fermion, an antifermion exists which shares the same mass but has opposite charge. On the one hand, charged-leptons are explicitly represented by the electric charge. For example, the antiparticles of the electron e^- and muon μ^- are the positron e^+ and antimuon μ^+, respectively. On the other hand, the antiparticles of quarks and neutrinos are typically denoted by writing $\bar{q}$ and $\bar{\nu}_l$.

Not all fermions (or antifermions) are experimentally observable. The electron was found in 1897 [1], followed by the second family lepton, the muon in 1937 [2]. Their respective neutrinos were found later, ν_e in 1956 [3] and ν_μ in 1962 [4]. The third lepton family was opened with the discovery of τ in 1975 [5] and completed with ν_τ in 2000 [6]. Quarks are only found in composite states called hadrons, which are made of an even and odd number of quarks: mesons $M = q\bar{q}$ and baryons $B = qqq$. Their lifetime is large enough to be observed experimentally, but only protons ($p = uud\bar{}$) and neutrons ($n = udd\bar{}$) form stable atoms together with electrons. Deep inelastic scattering involving electrons and protons confirmed the existence of the first generation of quarks [7–9]. The first particle containing an s quark was found in kaons (mesons with an s and/or $\bar{s}$ quark) in 1947 [10], but its existence was postulated after the discovery of u and d. The c quark was found in 1974 [11, 12] with the discovery of the J/ψ resonance (a meson made of $c\bar{c}$). The existence of the third family of quarks, b and t, introduced to explain the observation of processes violating the CP-symmetry [13], was confirmed in 1977 [14] and 1995 [15, 16], respectively.

Concerning the bosonic content of the SM, the fundamental interactions of the theory are mediated by gauge bosons. *Photons* (γ) and *gluons* (g) are the electromagnetic and strong force-carriers, respectively, both predicted to be massless and electrically neutral. The existence of photons was finally confirmed

in 1923 [17], and gluons were directly observed in 1978 [18, 19]. The weak interaction is mediated by three massive bosons instead: the charged $W^{\pm}$ and neutral Z, experimentally detected in 1983 [20–23]. A different fundamental interaction is mediated by the *Higgs* boson (H), related to the mechanism thanks to which the massive particles of the SM acquire their mass. The discovery of a particle compatible with H was reported in 2012 [24, 25].

Masses and electric charges of the elementary SM particles are collected in Table 1.1.

Table 1.1: Summary of the masses and electric charges Q_f of fermions and bosons in the Standard Model. For neutrinos masses, photon mass and electric charge, upper limits are given. For the top quark mass, the world average from cross-section measurements is reported. All values are extracted from the Particle Data Group [26].

Particle	Mass [GeV]	Q_f [e]
Electron, e	$0.109989461 \pm 0.0000000031$	-1
Muon, μ	105.683745 ± 0.0000024	-1
Tau, τ	1776.84 ± 0.12	-1
Electron neutrino, ν_e	$< 1.1 \times 10^{-9}$	0
Muon neutrino, ν_μ	$< 0.19 \times 10^{-6}$	0
Tau neutrino, ν_ν	$< 18.2 \times 10^{-6}$	0
Up, u	$(2.2\pm^{0.5}_{0.4}) \times 10^{-3}$	2/3
Charm, c	$1.275\pm^{0.025}_{0.035}$	2/3
Top, t	$172. \pm 0.7$	2/3
Down, d	$(4.7\pm^{0.5}_{0.3}) \times 10^{-3}$	$-1/3$
Strange, s	$(95\pm^{9}_{3}) \times 10^{-3}$	$-1/3$
Bottom, b	$4.18\pm^{0.04}_{0.03}$	$-1/3$
Photon, γ	$< 1 \times 10^{-18}$	$< 1 \times 10^{-35}$
Gluon, g	0	0
$W^{\pm}$	80.379 ± 0.012	± 1
Z	91.1876 ± 0.0021	0
Higgs, H	125.18 ± 0.16	0

1.3 The structure

1.3.1 Quantum electrodynamics

Quantum ElectroDynamics (QED) was the first gauge QFT that successfully described one of the fundamental interactions of nature: electromagnetism, involving charged fermions and light. Fully developed in the late 1940s, this theory is based on the $U(1)_Q$ symmetry group, in which the local transformations can be written as $U(x) = e^{iQ\beta(x)}$ where Q is the generator of the symmetry. The algebra of the group satisfies $[Q_i, Q_j] = if_{ijk}Q_k = 0$. In other words, the generator conmutes ($f_{ijk} = 0$) so the symmetry group is abelian. This has a fundamental implications: the gauge boson associated to this symmetry, the photon, does not self-interact[2], making electromagnetism a large-range force. The conserved quantity is the electric charge Q_f, and the coupling constant of the interaction is e which is usually reformulated in terms of the fine-structure constant, $\alpha_{\text{QED}} = e^2/(4\pi)$. The QED Lagrangian takes the following form:

$$\mathcal{L} = i\overline{\psi}(x)\gamma^\mu D_\mu \psi(x) - m\overline{\psi}(x)\psi(x) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \ , \tag{1.2}$$

where $D_\mu = \partial_\mu + ieQA_\mu(x)$ stands for the covariant derivative[3]and A_μ is the associated field of the gauge boson, the photon γ. One can identify the next terms: the kinetic term $i\bar{\psi}(x)\partial_\mu\psi(x)$ of free-propagating fermions and the corresponding mass term $m\bar{\psi}(x)\psi(x)$, the interaction term between fermions and photons $-eQ\bar{\psi}(x)\gamma^\mu A_\mu\psi(x)$ and finally the kinetic term of the photon field $\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$, with $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ as the photon field strength tensor. We can use the conventions of Feynman diagrams to illustrate two basic QED processes. Figure 1.1 represents

[2]This only holds at tree level in perturbation theory. Self-interacting terms are allowed in virtual corrections, but they are highly suppressed.

[3]The covariant derivative is defined in order to cancel out the extra terms that arise from the local transformations of the fields in the Lagrangian.

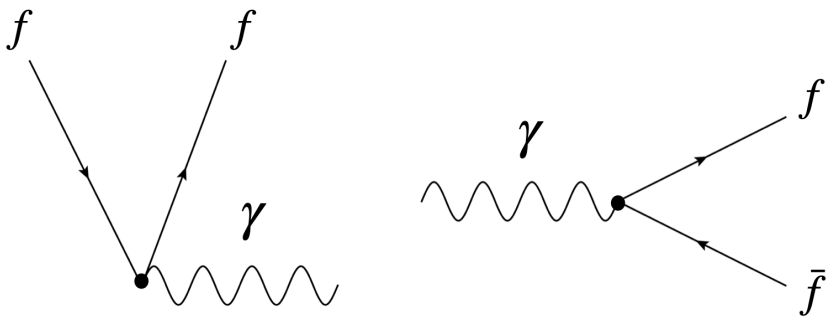


Figure 1.1: Feynman diagrams of two basic interaction processes in the QED Lagrangian.

the emission and absorption of a photon together with the scattering of a fermion f and with a fermion-antifermion $f\bar{f}$ annihilation in the leftmost and rightmost diagrams, respectively.

QED reached an unprecedented success. As mentioned before, the photon was predicted to be massless, in good agreement with the experimental limits: $m_\gamma < 1 \times 10^{-18}$ GeV [26]. The measurement of the anomalous moment of the electron a_e[4] yielded $a_e^{\text{exp}} = 1\ 159\ 652\ 188(\pm 4) \times 10^{-12}$, successfully predicted by QED to eleven significant figures [27]. The fine-structure constant prediction $\alpha_{\text{QED}}^{-1} = 137.035\ 999\ 070(\pm 98)$ [28] is in impressive agreement with the most precise experimental measurement $\alpha_{\text{exp}}^{-1} = 137.035\ 999\ 084(\pm 21)$ [29].

1.3.2 Weak interaction and unification

The description of the weak interaction was a very difficult task. There was a large amount of information about the dynamics underlying flavour-changing processes provided by low-energy experiments. Probably the best known example is the nuclear β-decay process, where a neutron decays into a proton, electron and electron antineutrino: $n \to pe\overline{\nu_e}$.

[4]The intrinsic moment of the electron μ_e depends on a g-factor which reads $g = 2(1 + a_e)$, where the anomalous moment a_e accounts for electron's quantum corrections.

The four-fermion theory, although being perfectly good as a lowest-order approximation, was plagued by unremovable infinities when going to higher orders in perturbative calculations (this formalism is revisited in Section 1.4). Surprisingly, despite the success of the QED theory, it took a few years before physicists began to use the Yang-Mills theory (gauge theory based in $SU(2)$ symmetry) to describe weak interactions, after several failed attempts at describing the strong force. One needs to bear in mind that in 1956 many sacred symmetries of space and time, parity (P) and parity-time (PT) conservation, were found to be violated by the weak interactions [30–33], even the charge-parity (CP) symmetry (1964) was not an exact symmetry [34]. So understanding symmetries was far from trivial. It was in 1957 – 1958 when physicists realized that weak interactions were actually a mixture of vector and axial interactions [35]. This led to the intermediate vector boson theories, where the lepton fields enter the interaction only in combinations of:

$$J^{\mu}(x) = \overline{\psi}^{l}(x)\gamma^{\mu}(1-\gamma_5)\psi^{\nu_l}(x) \ , \tag{1.3}$$

which can be written as the difference of the vector and axial currents:

$$\begin{aligned} J^{\mu}_{V}(x) &= \overline{\psi}^{l}(x)\gamma^{\mu}\psi^{\nu_l}(x) \ , \\ J^{\mu}_{A}(x) &= \overline{\psi}^{l}(x)\gamma^{\mu}\gamma_5\psi^{\nu_l}(x) \ . \end{aligned} \tag{1.4}$$

It turns out that this interaction does not conserve the lepton number as it was originally defined in QED, $N(l) = N(l^-) - N(l^+)$. But it does preserve a modified lepton number defined as $N(l) = N(l^-) - N(l^+) + N(\nu_l) - N(\bar{\nu}_l)$. This is in agreement with any experiment, where the lepton number conservation is found to hold for all processes.

A crucial feature of the weak interactions that determines profoundly the

structure of the SM is related to the chirality of the theory. In a chiral theory, the *Right-Handed* (RH) and *Left-Handed* (LH) components of the fields $\psi_{L,R}(x)$ decouple and can be treated separately. They are defined as $\psi_{L,R}(x) = P_{L,R}\psi(x)$ where $P_{L,R} = 1/2\,(1 \mp \gamma_5)$ are the projector operators. It easily follows that Eq. 1.3 can be written as:

$$J^{\mu}(x) = 2\overline{\psi}^{l}_{L}(x)\gamma^{\mu}\psi^{\nu_l}_{L}(x)\ , \tag{1.5}$$

and the interaction term would become (in analogy to the QED case):

$$\mathcal{L}_{\text{weak}} = g_W 2\overline{\psi}^{l}_{L}(x)\gamma^{\mu}\psi^{\nu_l}_{L}(x)W_{\mu}(x) + h.c.\ , \tag{1.6}$$

where the field W_μ describes W particles and g_W is a dimensionless coupling constant. This means that only the LH (RH) fields are involved for the fermions (antifermions) that participate in the weak interaction. However, there is no rationale to justify the structure of the weak force described here. It was around 1967 when Weinberg [36] and Salam [37] derived independently the unified *ElectroWeak* (EW) theory based on the$SU(2)_L \times U(1)_Y$ gauge group. Above the EW scale, ~ 246 GeV, nature exhibits a higher degree of symmetry: the weak and the electromagnetic forces become indistinguishable and are unified in the so-called EW force.

The transformations of $SU(2)_L$ symmetry have the form $U(x) = e^{iT_i\theta_i(x)}$, where the three generators are the so-called weak isospin operators $T_i = \sigma_i/2$ $(i = 1, 2, 3)$ and σ_i stands for the 2×2 Pauli's matrices. The symmetry generators do not conmute ($f_{ijk} \neq 0$), so the theory is non-abelian and therefore the gauge bosons W^i_μ participate in the interaction they mediate. This means that they self-interact. By considering the extended gauge group $SU(2)_L \times U(1)_Y$, we include the generator of the space given by the $U(1)_Y$ symmetry, the weak

hypercharge Y[5], which gives rise to an additional gauge field B_μ. The conserved quantities of these symmetries are the third component of the weak isospin T_3 and the hypercharge Y (resulting from acting the corresponding generators on the fermionic fields). The four bosons described remain massless, and the coupling constants of the two interactions they mediate are g_W and g_Y.

In order to get back the QED interaction term, we first rewrite:

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) \ , \\ Z_\mu &= \cos(\theta_W)W_\mu^3 - \sin(\theta_W)B_\mu \ , \\ A_\mu &= \sin(\theta_W)W_\mu^3 + \cos(\theta_W)B_\mu \ , \end{aligned} \tag{1.7}$$

where θ_W is the electroweak mixing angle under which B_μ and W_μ^3 are rotated to give rise to the Z_μ and A_μ bosons. We now demand that A_μ matches the electromagnetic field of Eq. 1.2. This leads to the following relations between coupling constants:

$$g_W \sin(\theta_W) = g_Y \cos(\theta_W) = e \ , \tag{1.8}$$

which implies the unification of the electromagnetic and weak forces. Furthermore, the electric charge operator Q is written as a linear combination of the T_3 and Y operators:

$$Q = T_3 + Y \ . \tag{1.9}$$

The four gauge fields defined in Eq. 1.7 have associated the physical bosons of the electroweak force. On the one hand, $W^\pm$ bosons mediate transitions involving up-type fermions and their corresponding down-type partners of given family, as shown in the leftmost diagram of Figure 1.2. They are called charged-current or flavour-changing interactions. In the other hand, the neutral-current interactions have the Z and γ bosons as mediators in reactions where fermions of the same

[5]Weak hypercharge is used instead of Q because it conmutes with the $SU(2)$ generators.

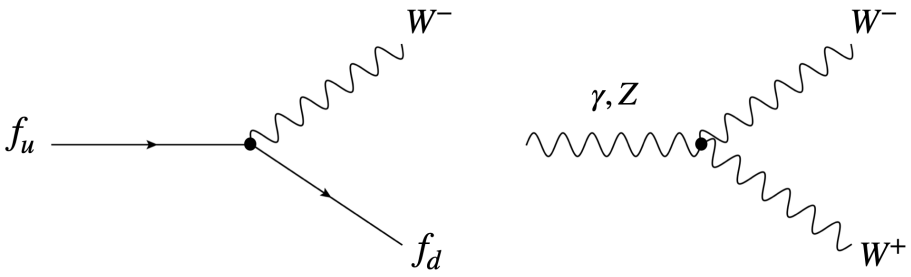


Figure 1.2: Feynman representation of the charged-current interactions and the third-point interaction vertex present in the electroweak Lagrangian. Same third- and fourth-point vertices describe self-interacting bosons, but restricted to the weak ones since the photon does not self-interact.

type are involved instead. The interaction vertices correspond to those sketched in Figure 1.1 . As mentioned above, a characteristic feature of the electroweak interaction (and of any $SU(N)$-based gauge theory) is that the associated bosons couple. An example is the third-point interaction vertex, represented in the rightmost diagram of Figure 1.2.

Naturally, the structure of the interaction fixed by the $SU(2)_L \times U(1)_Y$ symmetry gives back that of Eq. 1.6. However, it is important to notice the difference between the physical, so-called weak interaction and that of the $SU(2)_L$ symmetry: only LH (RH) fermions (antifermions) couple with the $W^{\pm}$ charged weak bosons, since they result from the combination of $SU(2)_L$ gauge bosons ($W_\mu^{1,2}$). The RH (LH) fermions (antifermions) do interact with the Z boson, because it includes the gauge B_μ boson of the $U(1)_Y$ symmetry. Similarly, the electromagnetic field couples to both LH and RH fermions. But, given that neutrinos do not feel the electromagnetic force (they are electrically neutral, $Y = Q_f = 0$), the RH neutrinos do not participate in any SM interaction. All in all, fermions that transform under $SU(2)_L$ rotations are arranged into isospin doublets, and those that do not are represented as isospin singlets instead. For

each fermion family, the fermionic fields are organized as follows:

$$l_L = \begin{pmatrix} \nu_l \\ l^- \end{pmatrix}_L , q_L = \begin{pmatrix} u \\ d' \end{pmatrix}_L , \; l_R, \; \nu_R, \; u_R, \; d'_R \; . \tag{1.10}$$

We have added RH neutrinos for completeness.

1.3.3 Spontaneous symmetry breaking

The development of the electroweak theory in terms of a gauge symmetry was closely related to the *Spontaneous Symmetry Sreaking* (SSB). Experimental evidence suggested that the weak bosons and charged-leptons were massive. However, the gauge invariance forbids mass terms in the EW Lagrangian, so they must acquire their mass in an alternative way: the Higgs mechanism, based on the Goldstone theorem [38, 39] and proposed independently by R. Brout and F. Englert [40] and P. Higgs [41], and further developed by G. Guralnik, C.R. Hagen and T. Kibble [42, 43]. Let's introduce a doublet scalar field ($\phi = [\phi^+(x), \phi^-(x)]$) by means of the complex potential:

$$V(\phi) = -\mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \; , \tag{1.11}$$

where μ^2 and λ are two real parameters. Since it is invariant under $SU(2)_L \times U(1)_Y$ rotations, it can be inserted as a new piece $\mathcal{L}_s$ in the EW Lagrangian without breaking the local symmetry:

$$\mathcal{L}_s = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi) + \sum_f \left(\bar{\psi}_L^f Y^f \phi \psi_R^f + h.c. \right) \; , \tag{1.12}$$

where the f index runs over all fermions. The first term contains the covariant derivative and encodes the interaction between the weak bosons and ϕ. The

interaction with fermions is embodied in the last term, proportional to the matrix Y^f. This potential determines the features of the electroweak vacuum: it will remain stable provided $\lambda > 0$, and if it also holds that $\mu^2 < 0$ the ground-state of the theory is no longer unique and does not correspond to a null *vacuum expectation value* (v.e.v). Rather, the ground-state fields are degenerated with a v.e.v. of $\frac{1}{2}(0, v)$, being $v = \sqrt{\mu^2/\lambda}$. For $v \neq 0$ the local symmetry of the vacuum is broken $SU(2)_T \times U(1)_Y \rightarrow U(1)_Q$[6]. According to the Goldstone theorem, a massless scalar boson appears for each of the three broken weak generators. With a convenient $SU(2)$ transformation[7], one can rewrite the field ϕ as fluctuations around this minimum, $\frac{1}{\sqrt{2}}(0, v + H(x))$ with a massive scalar field $H(x)$. The other three degrees of freedom, the Goldstone bosons, are absorbed by the vector bosons of the weak interaction, adding a longitudinal polarization and allowing them to be massive.

The Higgs potential of Eq. 1.11 can be written in terms of the Higgs field as follows:

$$V(\phi) = -\mu^2 H^2 + \lambda v H^3 + \frac{\lambda}{4} H^4 + \text{const.} \ . \tag{1.13}$$

The kinetic term gives the mass of the Higgs field:

$$m_H = \sqrt{-2\mu^2} = \sqrt{2\lambda} v \ , \tag{1.14}$$

whilst the other terms represent the Higgs self-interactions, with a strength proportional to λ. The last term present in Eq. 1.12 gives rise to the explicit

[6] Only $U(1)_Q$ remains in order to keep the photon massless.

[7] All minimums are indeed connected by $SU(2)$ transformations. So after the SSB, one can choose a unitary gauge transformation $G = e(-i\frac{\sigma_i}{2}\theta_i(x))$ to apply over $\phi = e(i\frac{\sigma_i}{2}\theta^i(x))\frac{1}{\sqrt{2}}(0, v + H(x))$.

interaction between fermions and the Higgs field, rearranged in the Yukawa sector:

$$\mathcal{L}_Y = \sum_f \frac{(v+H)}{\sqrt{2}} \left(y_f \bar{\psi}_L^f \psi_R^f \right) \ . \tag{1.15}$$

The fermion masses are encoded in the Yukawa couplings y_f, and can be expressed as:

$$m_f = \frac{v y_f}{\sqrt{2}} \ , \tag{1.16}$$

being fundamental parameters of the theory which must be determined experimentally. It is important to remark that, thanks to the specific Yukawa and hence mass that any fermion of the SM has, we are no longer blind to fermion flavours, as can be noted from Eq. 1.8 where electroweak bosons couple equally to all SM fermions. Expanding the first term of Eq. 1.12, mass terms involving weak bosons appear as well, so they become massive. Their mass take the form of:

$$m_W = \frac{g_W \nu}{2}, \quad m_Z = \frac{m_W}{\cos(\theta_W)} \ , \tag{1.17}$$

in terms of which the weak mixing angle can be expressed:

$$\sin^2(\theta_W) = 1 - \frac{m_W^2}{m_Z^2} \ . \tag{1.18}$$

As the fermionic mass generation, the quark mixing finds its origin in the Higgs mechanism. Normally, one defines the specific flavour of quarks as eigenstates of the quark mass matrix. Prior to the spontaneous symmetry breaking, quarks are massless so the mass matrix is diagonal. But after quarks become massive, the mass eigenstates of quarks do not coincide with the flavour eigenstates. The *Cabibbo-Kobayashi-Maskawa* (CKM) matrix [44, 45] relates flavour-eigenstates

(d', s', b') to physical mass-eigenstates (d, s, b) of down-type quarks:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = V_{\text{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix} . \tag{1.19}$$

The CKM matrix can be parameterized in terms of three mixing angles and one phase, the only source of CP–violation in the quark sector. Taking this into account, the charged-current weak interaction can be expressed as:

$$J_q^\mu = \sum_{i,j} \bar{q}_L^{u_i} \gamma^\mu V_{i,j} q_L^{d_j} , \tag{1.20}$$

where indices i, j run independently over the three quark families and V_{ij} represents the corresponding CKM matrix element. Thanks to this the top quark, for example, can decay not only into its isospin doublet partner, the b quark, via the exchange of a charged weak boson but also into s and u quarks.

1.3.4 Quantum chromodynamics

This section presents quarks and the strong interaction. Between 1950 and 1960, the development of new particle accelerators and detection techniques enabled the discovery of many stable hadrons and hadron resonances, such as the four Δ resonances (baryons made up of u and d quarks), the six hyperons (baryons containing one or more s quark) and the four kaon mesons. This spurred physicists on to the creation of an organized classification of particles. At this time, there was no general consensus about the existence of elementary constituents of the observed hadrons. The quark model proposed in 1964 by Gell-Mann and Zweig independently [46, 47], tried to make some sense of the growing fauna of hadrons. An experiment done at SLAC in 1968 [7–9] found that electrons were sometimes

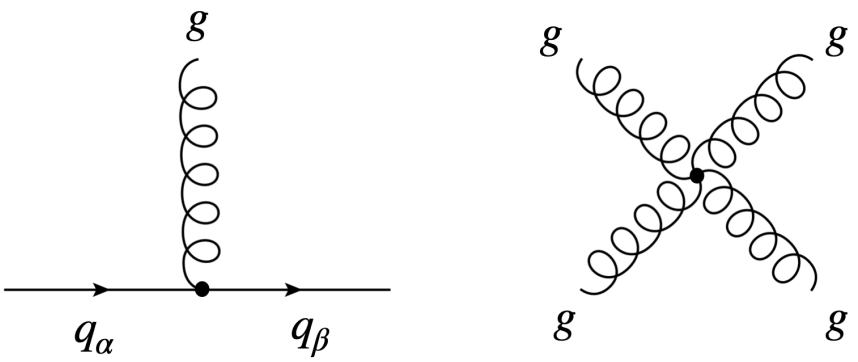


Figure 1.3: Feynman diagrams of two basic interaction vertices of the QCD Lagrangian: on the left, gluon emission by quarks and on the right, quartic gluon self-interaction.

scattered from nucleons at large angles, which was interpreted by Bjorken and Feynman [48, 49] as a prove of that protons and neutrons consisted of point particles, which were identified with Gell-Man and Zweig's quarks. The great mystery was why no one ever observed quarks. A possible explanation, called confinement hypothesis, was that the asymptotic states of the theory (solutions of the physical system) should be singlets of a new quantum number, which implies the non-observability of free quarks. Yang and Mills found eventually the symmetry group that governs *Quantum ChromoDynamics* (QCD).

The symmetry group for QCD is the non-abelian $SU(3)_C$. The transformation under this space is given by $U(x) = e^{i\lambda_a\theta_j(x)}$ with λ $(a = 1, .., 8)$ representing the 3×3 Gell-Mann matrices. Eight massless gauge bosons, called gluons, are hence associated to this symmetry. Color is the conserved charge in the strong interaction, and three different states are possible: green, blue and red color. The constant coupling g_S is defined as $\alpha_S = g_S^2/(4\pi)$. The QCD Lagrangian has the following form:

$$\mathcal{L}_{\text{QCD}} = \sum_q \bar{\psi}_q \left(i\gamma^\mu D_\mu - m_q\right) \psi_q - \frac{1}{4} G_a^{\mu\nu} G^a_{\mu\nu} \ , \tag{1.21}$$

where q runs over the six quark flavours and $G_a^{\mu\nu}$ represents the gluon field

Table 1.2: $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$ quantum numbers for left- and right-handed charged-leptons (l^-), neutrinos (ν) and up- and down-type quarks (u, d), denoted with L and R subscripts respectively, of any of the three fermion families. Leptons have 0 color number N_c, since they do not feel the strong interaction. For quarks N_c represents the number of possible color states (or the dimension of the $SU(3)$ representation). Right-handed fermions have $T_3 = 0$ ($Q_f = Y$), because they do not interact with the $SU(2)_L$ gauge bosons. Left-handed fermions with same Y and opposite T_3 form $SU(2)_L$ doublets.

Fermion	$SU(3)_c$	$SU(2)_L$	$U(1)_Y$
l_L^-	0	−1/2	−1/2
ν_L	0	1/2	−1/2
l_R^-	0	0	−1
ν_R	0	0	0
u_L	3	1/2	1/6
d_L	3	−1/2	1/6
u_R	3	0	2/3
d_R	3	0	−1/3

strength. Gluons also self-interact introducing three-point and four-point interaction terms in the QCD Lagrangian. The rightmost Feynman diagram of Figure 1.3 shows the four-point gluon self-interaction vertex, whilst the leftmost one represents the basic structure of a quark-gluon interaction vertex, where a quark changes its color charge ($\alpha \rightarrow \beta$) by emitting a gluon. Only quarks feel the strong force, so they are grouped into color triplets according to the three-dimensional space generated by $SU(3)$ transformations. A summary of the quantum numbers of fermions is available in Table 1.2.

1.4 Calculability of the theory

With these two ingredients, particles and their interactions, one is in principle ready to start making predictions. In this sense, theory calculations are concerned with observables related to the properties of physical particles in asymptotic states, and are given in terms of probability because of the statistical nature of

quantum mechanics. However, the analytical solutions of an interacting theory cannot be obtained.

For scattering processes a suitable formalism has been derived assuming that, right before and after the interaction occurs, the incoming and outcoming particles behave as free particles[8]. The interaction is treated as a perturbation of the solutions associated to free states. The Lagrangian can thus be expressed as $\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_I$, where the first term of the right-hand side equation contains the free fields and the second term embodies the interaction factors.

The probability transition P associated to given scattering process with a set of initial states $|i\rangle$ and final states $|f\rangle$ is related to the *Matrix Element* (ME) S_{if}:

$$P \sim |S_{if}|^2, \quad S_{if} = \langle f| S |i\rangle \ , \tag{1.22}$$

where S stands for the S-matrix which encodes the dynamics of the theory. If the coupling constant c of the underlying interaction is sufficiently small, the S-matrix can be expressed as a power series in the coupling constant:

$$S = \sum_n S^{(n)}(c^n) \ , \tag{1.23}$$

where the n^{th}-order perturbation term $S^{(n)}$ contains the product of n interaction factors $\mathcal{L}_I$. Let's consider the QED interaction and the positron-electron annihilation $e^+e^- \to \mu^+ - \mu^-$. Since a single interaction factor is linear in absorption and creation operators, in order to destroy the two particles present in $|i\rangle$ and create the final-state particles in $|f\rangle$, one should go to second-order term ($n = 2$)

[8] In other words, only interactions at very short scales are considered. We shall see that this assumption leads to divergencies when long-distance effects show up in perturbation theory.

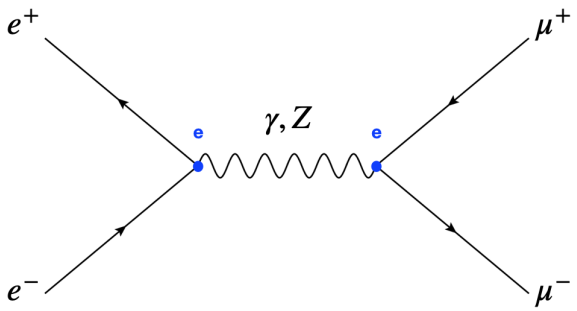


Figure 1.4: Electron-positron annihilation, via Z or γ bosons, with a muon-antimuon pair in the final state.

which contains two interaction factors:

$$S^{(2)} \sim -e^2 N[(\overline{\phi}\gamma^\alpha A_\alpha \phi)_{x_1}(\overline{\phi}\gamma^\beta A_\beta \phi)_{x_2}] \; . \tag{1.24}$$

We note that this first term with physical meaning corresponds to $e^2 \sim \alpha_{\text{QED}}$ and the associated Feynman diagram is depicted in Figure 1.4. As the lowest-level solution in perturbation theory, it is considered the *Leading Order* (LO) or tree-level contribution. When expanding Eq. 1.24, terms that destroy (create) particles in the initial (final) state appear: external lines in Feynman language. There are also terms that correspond to intermediate/virtual propagating particles, *i.e.* particles created and subsequently reabsorbed represented by internal lines in Feynman diagrams. The energy-momentum conservation at the interaction vertices determines the four-momenta of the intermediate particles.

The *Next-to-Leading Order* (NLO) term in perturbation theory goes with α^2_{QED}. It generally encodes corrections of the order α and represents closed loops or real emissions in Feynman diagrams. Figure 1.5 shows the self-energy NLO corrections to the fermion mass via γ loops. The next order correction (NNLO) is also shown. Such virtual, quantum corrections modify the properties of the bare particles as presented in the SM Lagrangian, given by the LO term. Physical, real particles are hence dressed by a cloud of virtual particles. This means that the

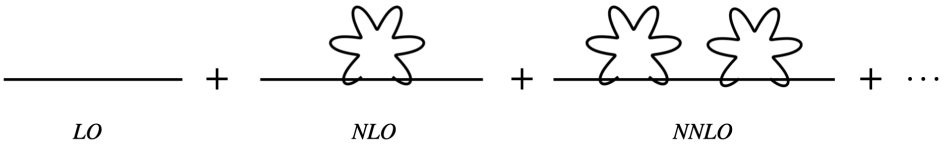


Figure 1.5: Higher-order quantum corrections to the fermion mass through γ loops.

bare parameters of the theory have no physical meaning and cannot be directly related to the experimentally observed particles.

To conclude this section, let's see that an observable $\mathcal{O}$ computed at fixed-order in perturbation theory can be expressed schematically as follows:

$$\mathcal{O} = \mathcal{O}_0 + \alpha\mathcal{O}_1 + \alpha^2\mathcal{O}_2 + ... + \alpha^n\mathcal{O}_n \ , \tag{1.25}$$

where $\mathcal{O}_0$ represents the LO term, $\mathcal{O}_1$ the NLO correction and $\mathcal{O}_n$ the n^{th}-order correction. An observable of special interest in high-energy physics is the cross-section σ which measures the probability that a specific process takes place, and is expressed in units of area. When σ is specified as the differential limit of a function of some final-state variable, it is called differential cross-section. When the cross-section is integrated over all possible final states, it is called total or inclusive cross-section instead.

1.4.1 Ultraviolet divergencies and running

Calculations in perturbation theory are plagued by divergencies that can spoil the validity of the perturbative expansion. Nevertheless, the theory is internally consistent, since it is possible to understand the origin of the divergencies and remove them by adequate procedures.

In particular, the *UltraViolet* (UV) divergencies arise when virtual particles in closed loops beyond LO in perturbation theory have unbounded energy. One should then sum over all possibles values of momentum in order to calculate the total probability transition. Mathematically, this translates into Feynman integrals that diverge when the virtual particle momentum becomes infinity.

A suitable procedure to deal with such divergences is called renormalization, carried out in two steps:

1. Regularization: it consists of rendering infinite integrals finite by parameterizing the divergences. Any regularization technique is required to preserve the local symmetry of the theory. Dimensional regularization fulfills this condition and relies on the observation that Feynman integrals would be finite if the spacetime dimension was less than 4 and that the results would be analytic functions of the dimension. Therefore, the integral dimension is modified as follows:

$$d = 4 \rightarrow d' = d - 2\epsilon \ .$$

   Divergencies are revealed as those terms with an $1/\epsilon$ factor. Furthermore, a renormalization parameter μ_R is introduced in order to keep the constant coupling of the interaction involved dimensionless when taking the limit $\epsilon \rightarrow 0$.

2. Renormalization: this step originates from the recognition that the fields of an interacting theory do not correspond to those in a non-interacting theory, where the perturbative treatment actually starts from. Interaction modifies the properties of the particles and so the predictions of the theory must be expressed in terms of physical properties. This is done allowing the unphysical, bare parameters of the Lagrangian, such as masses and

couplings, to absorb the divergences located in the previous step. There is not an unique way to do this, given rise to the so-called *Renormalization Schemes* (RS). In general, the renormalized parameters will depend on the μ_R scale.

After this procedure, there are no longer loop divergencies so predictions become possible. However, at all orders in perturbation theory, the calculations of physical observables must be independent of such arbitrary choices (μ_R and RS). We must impose the observable $\mathcal{O}$ to be invariant under changes of μ_R:

$$\begin{gathered}\mu_R^2 \frac{d}{d\mu_R^2} \mathcal{O}\left(Q^2/\mu_R^2, \alpha(\mu_R^2), m(\mu_R^2)\right) = \\ \left(\mu_R^2 \frac{\partial}{\partial \mu_R^2} + \beta(\alpha)\frac{\partial}{\partial \alpha} - \gamma(\alpha) m \frac{\partial}{\partial m}\right)\mathcal{O} = 0 \ ,\end{gathered} \tag{1.26}$$

where Q is the energy of the process. This equation is known as *Renormalization Group Equation* (RGE). We note that two further dimensionless functions arise: the anomalous mass dimension γ and the β function, which can be perturbately calculated as a power series of α. The β function encodes the evolution of the coupling constant with the μ_R scale:

$$\beta(\alpha) = \mu_R^2 \frac{\partial \alpha}{\partial \mu_R^2} = -\sum_{i=0} \left(\beta_i \alpha^{i+2}\right) \ . \tag{1.27}$$

The anomalous mass dimension γ encodes the evolution of the mass parameter instead:

$$\gamma(\alpha) = -\mu_R \frac{1}{m}\frac{\partial m}{\partial \mu_R^2} = \sum_{k=0} \left(\gamma_k \alpha^{k+1}\right) \ . \tag{1.28}$$

The solution of these two equations brings up the explicit relation or *running* of the parameters with the μ_R scale at given order in perturbation theory. Finally, they are renormalized to a reference scale μ_0 where highly accurate measurements can be performed. For example, the running of the coupling constant at LO when

$\mu_R >> \mu_0$ reads as:

$$\alpha(\mu_R^2) = \frac{\alpha\left(\mu_0^2\right)}{1 + \beta_0 \alpha\left(\mu_o^2\right) \log\left(\mu_R^2/\mu_0^2\right)} \ , \tag{1.29}$$

where γ_0 and β_0 stands for the first coefficient of the anomalous mass dimension and the β function, respectively, and $\alpha(\mu_0^2)$ is the coupling constant evaluated at the reference scale. The behaviour of the coupling constant with the energy is determined ultimately by the interaction underlying through the sign of the β function. In the QED case, taking $\alpha = \alpha_{\text{QED}}/\pi$, we have at LO $\beta_0 = -1/3$. In contrast, for QCD and $\alpha = \alpha_S/\pi$ it turns out to be $\beta_0 = (11N_c - 2N_f)/12$, where N_c and N_f stands for the number of quark colors and active flavours[9], respectively. Simply put, this implies that the QED force gets weaker at large scales or low energies ($1/\alpha(\mu \approx 0) \sim 137$) and grows logarithmically at small scales or high energies ($1/\alpha(m_W) \sim 128$). In QCD it is the other way around. This feature is commonly known as screening and anti-screening, respectively.

1.4.2 Renormalization schemes

Among a broad range of RSs available, the *modified Minimal Subtraction* ($\overline{MS}$) scheme is commonly used to renormalize the strong coupling. Here, one can interpret the renormalization scale μ_R as the momentum scale above which all QCD corrections to the gluon interactions are included in the bare parameter. Quark and leptons masses have to be renormalized too. Let's restrict ourselves to some RSs of special interest to motivate further discussions in the following chapters, in the context of quark masses renormalization.

After dimensional regularization, the bare quark propagator $S_q^0(p)$ receives

[9]For a given scale energy Q, quarks with masses $m_q^2 \ll Q^2$ are integrated out. The remaining quarks are deemed to be active.

self-energy corrections at NLO from gluon loops [50];

$$S_q^0(p) = \frac{i}{p - m_q^0} \to S_q(p,\mu) = \frac{i}{p - m_q^0 - \sum(p, m_q^0, \mu_R)} \ , \tag{1.30}$$

where p is the quark four-momentum, m_q^0 indicates the bare unrenormalized quark mass and $\sum(p, m_q^0, \mu)$ displays the dominant contribution in the resonance limit $(p^2 \to m_q^2)$:

$$\sum(p, m_q^0, \mu_R) \sim m_q^0 \left(\frac{\alpha_S \mu_R)}{\pi}\right)\left[\frac{1}{\epsilon} + \ln\left(4\pi e^{-\gamma_E}\right) + A^{\text{fin}}(m_q^0/\mu_R)\right] + \dots \ . \tag{1.31}$$

We see that the divergent term $1/\epsilon$ appears explicitly, along with a constant factor exponentiating the Euler constant γ_E and a finite term $A^{\text{fin}}(m_q^0/\mu_R)$ related to quantum fluctuations of soft momenta. In the $\overline{MS}$ and on-shell (pole) mass schemes, the renormalized quark masses absorb different terms:

$$\begin{aligned} m_q^{\overline{MS}}(\mu_R) &= m_q^0 \left\{1 + \left(\frac{\alpha_S(\mu_R)}{\pi}\right)\left[\frac{1}{\epsilon} + \ln\left(4\pi e^{-\gamma_E}\right)\right]\right\} + \dots \ , \\ m_q^{\text{pole}} &= m_q^0 \left\{1 + \left(\frac{\alpha_S(\mu_R)}{\pi}\right)\left[\frac{1}{\epsilon} + \ln\left(4\pi e^{-\gamma_E}\right) + A^{\text{fin}}(m_q^0/\mu_R)\right]\right\} + \dots \ . \end{aligned} \tag{1.32}$$

As we can observe, the $\overline{MS}$ mass is scale-dependent as the QCD coupling, and analogously it can be understood as the momentum scale above which self-energy quantum corrections are included into $m_q^{\overline{MS}}$. This means that the $\overline{MS}$ mass does not receive contributions from low-energy processes, since μ_R acts as a shield. This sort of mass definitions are known as short-distance masses. In the $\overline{MS}$ scheme, the renormalization scale is set to the physical scale of the process that determines the sensitivity of the calculated observable to the quark mass, $\mu_R \gtrsim m_q$. This generally yields a good behaviour of the perturbative QCD expansion. On the other hand, the renormalized mass in the on-shell RS m_q^{pole} includes self-energy corrections from all scales, so that has no dependence with

the renormalization scale. The renormalized mass is hence chosen to be the pole of the propagator, $p - m_q^{\text{pole}}$ by requiring $\sum(p^2 = (m_q^{\text{pole}})^2) = 0$. In this way, the pole mass is closer to the intuitive rest mass[10] associated to quarks treated as real and external particles. The pole and $\overline{MS}$ masses are renormalized quantities of the same bare parameter, so they are connected through:

$$m_q^{\text{pole}} - m_q^{\overline{MS}}(\mu_R) = \frac{4}{3}\left(\frac{\alpha_S(\mu_R)}{\pi}\right) m_q^{\overline{MS}}(\mu_R) + O(\alpha_S^2) \ , \tag{1.33}$$

which is known up to N^{4}LO in QCD [51, 52]. Notice that two RSs are equivalent in the sense that the predicted observables are independent from such choice.

The pole mass exhibits some unphysical features due to its definition: is based on the fact that QCD can resolve virtual and real perturbative corrections down to arbitrarily small scales. This leads to a linear-infrared sensitivity, called the renormalon problem [53]. The point is that, in general, perturbative series are asymptotic series, so terms that decrease at higher orders may eventually adopt divergent patterns [54]. Renormalons are associated to terms that show such divergent patterns in high-order calculations. Partonic calculations in the pole mass RS suffers from infrared renormalons that originate from virtual non-self-energy corrections that are soft and left uncancelled [54]. In order to avoid such large corrections, the asymptotic serie is truncated at the order where they are minimal. The pole mass renormalon ambiguity is related to the uncertainty of doing this, and it has been quantified to be of the order Λ_{QCD} [55–57]. This seems natural because, as we shall see in Section 1.5, quarks are not the proper degrees of freedom at this level. Obviously, the renormalon ambiguity vanishes when considering a short-distance mass at an appropriate μ_R.

Another RS of interest for the work presented in this thesis is the MSR scheme.

[10] Quantity which does not depend on the overall motion of the system, *i.e.* is the same in all frames of reference connected by Lorentz transformations.

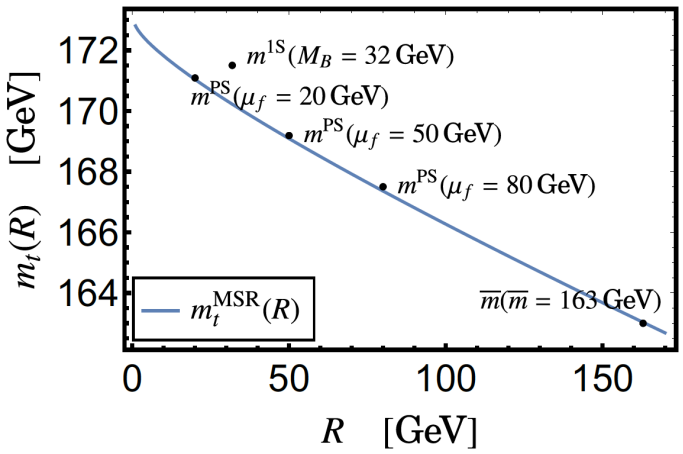


Figure 1.6: Evolution of the MSR mass (blue line) with the energy scale R. Further low-energy short-distance masses at several energy scales are displayed: m^{PS}, m^{1S} and $\overline{m}$ ($m^{\overline{MS}}$ in this work). Image taken from Ref. [50].

It provides a short-distance mass that depends on an R scale above which self-energy corrections are absorbed:

$$\begin{aligned} m_q^{\text{MSR}}(R) = \; & m_q^0 \left\{ 1 + \left(\frac{\alpha_S(\mu_R)}{\pi} \right) \left[\frac{1}{\epsilon} + \ln\left(4\pi e^{-\gamma_E}\right) + A^{\text{fin}}(m_q^0/\mu_R) \right] \right\} \\ & - R \left(\frac{\alpha_S(\mu_R)}{\pi} \right) A^{\text{fin}}(1) + \ldots \,, \end{aligned} \tag{1.34}$$

so unresolved self-energy QCD corrections below R are left to cancel other quantum corrections from scales below R. The MSR mass can be seen as a natural interpolation between the pole ($R \to 0$) and the $\overline{MS}$ masses ($R \to m_q^{\overline{MS}}$). It can be related to the $\overline{MS}$ mass with a precision of 30 MeV, and it is numerically close to other low-scale short-distance masses (m^{PS}, m^{1S}) when probed at their intrinsic energy scales, as shown in Figure 1.6.

1.4.3 Infrared divergencies

Feynman diagrams also contain a second kind of divergencies, this time related not only to closed loops but also to real emissions in the initial and final states. *InfraRed* (IR) divergences, featured by soft and collinear radiation, originate when

the momenta and the splitting angle of massless particles tend to 0, respectively. They can be seen as a degeneracy of states occurring in the soft and collinear limit, since we cannot possibly distinguish soft emissions and collinear splittings from situations where these emissions and splittings are absent. This signals that these divergences may cancel in properly averaged quantities, *i.e.* one needs to take into account all indistinguishable states. The *Kinoshita-Lee-Nauenberg* (KLN) theorem [58, 59] postulates that IR divergences from virtual corrections and real emissions cancel out at each order of perturbation theory. In light of this, two possible solutions show up to save the predictive power of the theory:

1. Calculate fully inclusive observables, in order to benefit from the KLN theorem, which contains all possible final states. An example would be the inclusive cross-section σ.

2. For non-fully inclusive observables, an option is regularize IR divergencies in closed loops and use observables that are not affected by the soft and collinear effects present in the final state. This property is called *InfraRed and Collinear* (IRC) safety and it essentially allows to apply phase-space restrictions in the measurement of an IRC observable.

With the formalism introduced so far, perturbative computations of inclusive or IRC observables in given renormalization scheme for scattering processes are possible. However, the validity of the perturbation theory breaks down when describing low-energy QCD processes, which are crucial in the context of hadron colliders. We shall see the nature of this issue and some solutions to deal with it in the following section.

1.5 QCD phenomenology

Understanding the behaviour of the α_S coupling, *i.e.* the strength of the strong force over the complete energy range is crucial to describe properly the hadronic interactions at both long- and short-distance scales. The short-distance domain allows to draw a real picture of what is going on in high-energy hadronic collisions, the fragment of an energetic parton into further partons as propagating through the spacetime (known as showers), and also to probe theories *Beyond the SM* (BSM) which aim to unify strong and electroweak interactions. The long-distance domain is essential to explain how partons undergo a transition to hadrons (process called hadronization) and the fragmentation of soft partons as well. In a high-energy experiment where the initial states are featured by colliding partons, such as the LHC (largely introduced in Section 3.1), and the final states comprise hadrons, we cannot settle for a partial description of the physical processes: these two worlds need to come together in a coherent and rigorous way in order to profit as much as possible from the information collected by the experiments.

1.5.1 Short- and long-distance domains

As introduced in Section 1.4.1, the behaviour of the running coupling α_S with the energy scale Q (let us identify the normalization scale with the energy at which given process is probed, $\mu_R = Q$) is completely governed by the β function, as shown in Eq. 1.29. Gauge theories based on non-abelian symmetry group in four-dimension spacetime, as QCD, are characterized by a running coupling that decreases in the short-distance regime and increases when going to long-distance domains [60]. In the short-distance regime, when Q gets large values, quarks behave as free particles: they barely interact. In contrast, the interaction becomes very intense as two quarks are pulled apart. These two features are

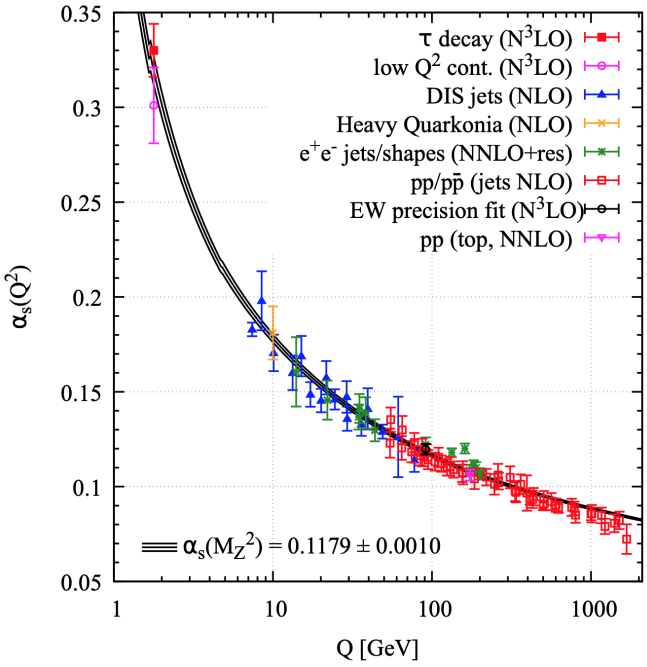


Figure 1.7: Summary of measurements of α_S as a function of the energy scale Q. The corresponding order in QCD perturbation theory used to derive α_S is indicated withinbrackets (where res. stands for resummation, a further technique needed to make predictions covered in Section 1.6.3). Image taken from Ref. [26].

widely known as asymptotic freedom and confinement, respectively and the scale at which they diverge is referred to as hadronization scale, Λ_{QCD}. They ultimately explain why free quarks are not been observed. The trend of α_S with the energy scale is depicted in Figure 1.7, where the value extracted from data measured by several experiments at a given order in perturbation theory is summarized.

Beyond Λ_{QCD}, Eq. 1.29 diverges: this does not mean that α_S gets infinite, but it becomes sufficiently large for perturbation theory to break down. This can be interpreted as, in this regime, quarks and gluons are not the appropriate degrees of freedom. The mass scale of the lightest hadron is precisely about 200 MeV, of the order of Λ_{QCD}. α_S gets so strong as two quarks are pulled apart that creating a $q\bar{q}$ pair from the vacuum becomes energetically favorable, resulting in two mesons. On the other hand, among the consequences of asymptotic freedom is that perturbation theory converges better (faster) at high energies, since the

coupling constant decreases.

High-energy quarks and gluons tend to undergo successive branchings at small angles, producing a series of collimated quarks and gluons. This shower of partons continues generating branchings until the non-perturbative regime is reached, typically of the order of 1 GeV or Λ_{QCD}. At this stage, QCD confines the resulting low-energy partons into hadrons. Therefore, such high-energy partons appear in the final state as a collimated spray of low-energy particles, widely known as jets. Jets play a crucial role in the phenomenology of QCD, since they turn out to be very useful to handle such complex configurations and serve as a proxy of the originating partons. One could also define IRC safe jets so differential predictions or restricted phase-spaces analysis involving jets become possible.

1.5.2 Non-perturbative effects

All physical phenomena that cannot be described in terms of calculations within perturbation theory, such as the hadronization process, are called non-perturbative effects and eventually accounted for as power corrections of Λ_{QCD}/Q to the perturbative result. Predictions of fully inclusive observables do not substantially change, since the non-perturbative effects occur long after the hard scattering takes place. Less inclusive observables, such as jet and rates, may suffer somewhat from non-perturbative effects, but fixed-order predictions are still adequate. Observables that depend on the distributions of individual hadrons, as the jet mass, must take them into account [26].

In practice, although we do not know the structure of the low-energy QCD solutions, we do have some knowledge of the properties that such solution must have. In other words, we can see how these non-perturbative effects may change the final states for a broad set of observables and develop techniques

to describe them phenomenologically (see Section 1.6.4 for an introduction on these techniques). Finally, for sufficiently inclusive observables, non-perturbative effects can be incorporated using a shape function F_{np} which is convolved with the perturbative distribution. In the case of a differential cross-section as a function of the X variable, this would read as follows:

$$\frac{d\sigma_{\text{had}}}{dX} = \int d\epsilon \frac{d\sigma_{\text{pert}}(\epsilon)}{dX} F_{\text{np}}(\epsilon) \ . \tag{1.35}$$

The shape function is a non-perturbative object, so it cannot be computed from first-principles. However, its functional form can be obtained in an effective field theory framework (see Section 1.7 for a brief introduction on this topic) and be expressed in terms of a few hadronic parameters. A common model is:

$$F_{\text{np}}(\epsilon) = \frac{4\epsilon}{\Omega^2} e^{-2\epsilon/\Omega} \ , \tag{1.36}$$

where Ω is a mass scale parameter on the order of Λ_{QCD}. The shape function can be developed in moments, with higher moments suppressed by powers of Λ_{QCD}/X^2. The leading contribution is a shift of the distribution of the observable X and is determined by a universal non-perturbative parameter (which has to be fitted) multiplied by a calculable, observable dependent parameter [61].

1.6 Physics at hadron colliders

Particle colliders are the perfect laboratory to probe physics described by the SM, since a wide range of processes and observables are accessible in many different kinematic regimes. This allows to exploit the predictive power of the theory and also permits to test its limitations. In this context, it is far from trivial to provide a fully-consistent prediction of the hard-scatter at the TeV scale down

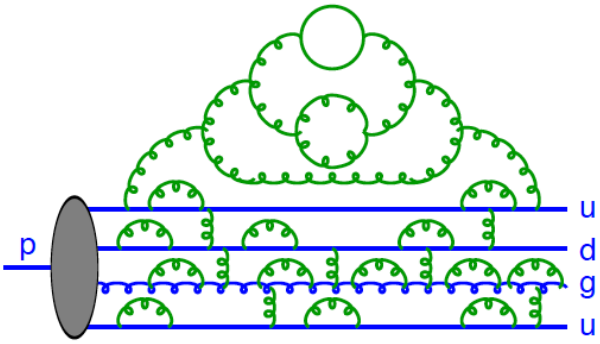


Figure 1.8: Representation of the quantum fluctuations undergoing inside a proton. Figure taken from Ref. [63].

to the formation of stable hadrons hitting the detector at the MeV scale from first-principles. Furthermore, an additional complication arises when considering hadrons in the initial state: their formation and structure are purely long-distance physics, and so intractable in perturbative QCD.

It is still possible to provide a description of the internal structure of high-energy hadrons, schematically displayed in Figure 8.9. In the parton model, hadrons are regarded as a collection of quarks, antiquarks and gluons which relative composition is time-dependent. From a probabilistic point of view, there are some quarks that are more likely found to carry a large fraction of the total hadron momentum. These quarks are identified as the valence quarks, represented with straight lines. The rest of quarks and gluons, shown with green lines, are seen as quantum fluctuations that carry smaller fractions of the hadron momentum. They constitute the so-called parton-sea. In this section, we shall see how a probabilistic treatment of the internal hadron structure enables a full description of the hard-scatter together with the perturbative cross-section by means of a factorization theorem [62]. The impact of further sub-processes arising before and after the hard-scatter is covered too, and finally we will introduce multi-purpose Monte Carlo generators as an essential tool to provide QCD-like predictions in this context.

1.6.1 Factorization theorem

The principle of factorization is the main core of all theoretical calculations of hard-scatter processes involving hadrons in the initial state. Basically, the factorization theorem states that short- and long-distance physics contributing to given process can be separated, up to corrections suppressed by powers of the relevant scale of the process [64]. Short-distance physics are related to the scattering of two energetic partons. Long-distance physics have to do with the internal hadron dynamics described above. This separation is possible because interactions inside the hadron occur over timescales of $1/\Lambda_{\mathrm{QCD}}$, whilst the collision of two partons from different hadrons happens over a much shorter timescale, $1/Q << 1/\Lambda_{\mathrm{QCD}}$. Therefore, partonic quantum fluctuations are frozen in some sense, and the hard probe only sees a snapshot of the internal hadron structure.

This treatment allows to formulate the cross-section of two colliding hadrons as the convolution of two terms: a non-perturbative, yet universal *Parton Density Function* (PDF) for the long-distance regime and a process-dependent perturbative partonic cross-section encoding the short-distance physics. For a proton-proton collision of the type $pp \to ab + X$ (where X stands for any further particle present in the final state) at a center-of-mass energy $\sqrt{s}$, σ can be written as:

$$\sigma_{pp\to ab+X} = \sum_{i,j} \int_0^1 dx_1 dx_2 f_{i,p}(x_1, \mu_f^2) f_{j,p}(x_2, \mu_f^2) \sigma^{\mathrm{pert}}_{pp\to ab+X}(x_1, x_2, s, \mu_f^2) \ , \quad (1.37)$$

where $\sigma^{\mathrm{pert}}_{pp\to ab+X}$ corresponds to the perturbative partonic cross-section, μ_f is the factorization scale that represents the resolution with which the hadron is probed and x_i $(i = 1, 2)$ are the momentum fractions of the total hadron momentum carried by the colliding partons. They are given by the PDFs $f_{k,p}(x_k, \mu_f^2)$, which represents the probability density of finding a parton of type k (any

quark/antiquark or gluon) carrying the momentum fraction x_k when a proton is probed at μ_f.

Although PDFs cannot be calculated in perturbation theory, it is possible to determine their evolution with μ_f perturbately by requiring the Eq. 1.37 be independent of μ_f at any order. Analogously to the RGE that arises in the renormalization process, this requirement gives rise to the DGLAP equations [65–67]:

$$\mu^2 \frac{d}{d\mu^2} f_a(x,\mu^2) = \sum_b P_{ab}(x,\alpha_S(\mu)) \otimes f_b(x,\mu^2) \ , \tag{1.38}$$

where $P_{ab}(z)$ represents the called splitting functions that encode the probability for a parton a to emit a parton b carrying a fraction z of its longitudinal momentum. The solutions of these equations allow to evolve the PDFs from a reference scale μ_0 to a desired scale μ. The DGLAP equations predict the splitting of the initial parton into further partons through $P_{ab}(z)$, which reduces the incoming partonic energy in the hard-scatter.

PDFs can be determined by fitting a parametric form[11] of $f(x,\mu^2)$ to data collected from deep inelastic lepton-nucleon scattering and related hard-scattering processes initiated by nucleons (protons and neutrons). Since the PDFs are universal and therefore process independent, data from different experiments can be combined to constraint the same PDFs in different kinematical regions in x. Figure 1.9 shows the PDFs associated to unpolarized protons fitted by the NNPDF collaboration at two different factorization scales: $\mu_f^2 = 10$ and 10^4 GeV^2. We see that the probabilities of finding up and down quarks with a very large momentum fraction are the highest in both cases, as they are valence quarks. On the contrary, other quarks and gluons exhibit larger probabilities for small x.

[11] Most of the groups devoted to provide such constraints use input PDFs of the form $xf = x^a(...)(1-b)^b$ with 14-28 free parameters in total. The NNPDF collaboration, in contrast, uses an unbiased modeling tool based on neural networks with more than 30 parameters.

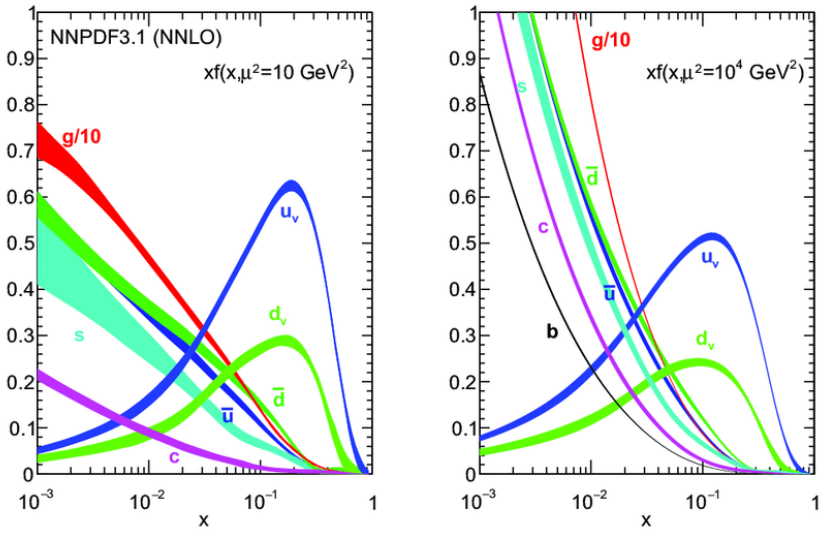


Figure 1.9: Constrained Parton Density Functions for unpolarized protons at two different factorization scales, including in the global fit data from LHC experiments. Image taken from Ref.[68].

The choice of the scale μ_f deserves some discussion. The total partonic cross-section of Eq. 1.37 is independent of μ_f when calculated to all orders in QCD. However, it does depend at any given finite order in perturbation theory. This dependence is usually significant at lower orders, so the choice of the value of μ_f is most important [69] and must respond to physical reasons. Since it appears in both long- and short-distance terms, it cannot take significantly larger values than the scale of the process Q: otherwise, the PDFs would include effects in shorter timescales. It is also introduced beyond LO in order to remove IR divergencies arising from hard collinear initial-state radiation. So it cannot take too lower values than Q either. It is fully process dependent and usually set to $\mu_f \sim Q$. An uncertainty to this choice is commonly assigned, where μ_f is varied over some interval, such as $Q/2 \lesssim \mu_r \gtrsim 2Q$.

1.6.2 Fixed-order calculations

The partonic cross-section entering in the factorization formula can be computed at fixed-order in perturbation theory and expressed as follows:

$$\begin{aligned}\sigma^{\text{pert}}_{pp\to ab+X}(p_1,p_2;Q,\{Q_1,...\};\mu_f^2) = \alpha_S^k(\mu_R^2)\{&\sigma^{(\text{LO})}(p_1,p_2;Q,\{Q_1,...\}) \\ &+ \alpha_S(\mu_R^2)\sigma^{(\text{NLO})}(p_1,p_2;Q,\{Q_1,...\};\mu_R^2;\mu_f^2) \\ &+ \alpha_S^2(\mu_R^2)\sigma^{(\text{NNLO})}(p_1,p_2;Q,\{Q_1,...\};\mu_R^2;\mu_f^2)+...\}\ ,\end{aligned} \tag{1.39}$$

where we have made explicit the dependence of the perturbative total cross-section with the momenta of the colliding partons $p_{1,2} = x_{1,2}p$ and all the kinematic scales present in the process under study, $Q, \{Q_1, ...\}$. The LO term $\sigma^{(\text{LO})}$ gives only an estimate of the order of magnitude of $\sigma^{\text{pert}}_{pp\to ab+X}$, because α_S is not unambiguously defined at this order (this is true provided the higher order terms have a sizable impact). A reliable estimate of $\sigma^{\text{pert}}_{pp\to ab+X}$ would require the inclusion of the NLO term, at least. Regarding the higher-order terms, each n^{th} term in the curly bracket contains logarithmic contributions of the type $(\alpha_S(\mu_R^2)\ln(Q/\mu_R))^n$. So if one sets the value of μ_R too low, such contributions would get very large and therefore the validity of the perturbative expansion breaks down. In order to avoid this, the value of the renormalization scale should satisfy $\mu_R \sim Q$.

In principle, any inclusive or IRC safe observable can be computed by means of Eq. 1.39. However, the logarithmic terms due to a very different Q and μ_R scales are not the only ones that may appear in this kind of calculations. Indeed, there are logarithmically enhanced terms of several natures arising in special regions of the phase-space that can spoil the perturbative convergence. Resummation calculations comprises a set of techniques tailored to identify, order and include them in a systematic and coherent way. Eventually, resummation calculations

have to be matched to fixed-order ones in order to avoid double-counting.

1.6.3 Resummation of logarithmic terms

When the different kinematical scales involved in the process are not of the same order of magnitude, for example $Q >> Q_1$, terms beyond LO in Eq. 1.39 may contain double- and single-logarithmic contributions of the type $(\alpha_S L^2)^n$ and $(\alpha_S L)^n$ where $L = \ln(Q_1/Q) >> 1$. This type of terms also appear when evolving the PDFs from a low input scale μ_0 to the hard-scatter scale Q. They can be systematically resummed, at any order in perturbation theory, in the evolved $f(x, Q)$ through the DGLAP equation. Threshold logarithms, which go like $L = \ln(1 - x)$ are also important when the partonic final state carries a large fraction x (with $x \to 1$) of the available hard-scatter energy [70].

Another source of logarithmically enhanced terms has to do with unbalanced virtual and real emissions in the soft and collinear limit. This happens when tight constraints are placed on the final state, which can affect real emissions but not loop contributions. This specially affects the emission of soft gluons. As a result, double-logarithmic terms appear at each order of α_S.

The fixed-order contribution at LO of Eq. 1.39 now reads:

$$\sigma^{\text{pert}}_{pp\to ab+X} = \alpha_S^k(\mu_R^2)\left(\sigma^{\text{rem}} + \sigma^{\text{res}}\right) \ , \tag{1.40}$$

where σ^{res} embodies the all-order resummations:

$$\begin{aligned}\sigma^{\text{res}} \sim & \sigma^{(\text{LO})} \sum_{n=1}^{\infty} \alpha_S^n(\mu_R^2)\left(C_{2n}^{(n)} L^{2n} + C_{2n-1}^{(n)} L^{2n-1} + C_{2n-2}^{(n)} L^{2n-2} + ...\right) \\ = & \sigma^{(\text{LO})} \sum_{n=1}^{\infty} \sum_{k=1}^{2n} \alpha_S^n(\mu_R^2) C_k^{(n)} L^k \ .\end{aligned} \tag{1.41}$$

A *Leading Log* (LL) resummation means that one accounts for all terms with

$k = 2n$, the *Next-to-Leading Log* (NLL) includes additionally all terms with $k = 2n - 1$, etc. The remainder term σ^{rem} has no logarithmic enhancements, and comes from matching resummed to the fixed-order calculations:

$$\sigma^{\text{rem}} = \sigma^{(\text{LO})} - \left[\sigma^{\text{res}}\right]^{(\text{LO})} \ , \tag{1.42}$$

where the LO part $\sigma^{(\text{LO})}$ is subtracted by the corresponding truncated resummed term $[\sigma^{\text{res}}]^{(\text{LO})}$. Of course, this can be extended to all orders in α_S (NLO, NNLO and so forth). This represents an improved perturbative calculation because: firstly, it saves the predictive power of the theory and, secondly, it allows the perturbative result to reach regions where double-logarithmic terms become large ($\alpha_S L \sim 1$).

It is important to remark that further logarithmic terms exist besides the ones described here. The dominant enhancements that given calculation must account for are determined by the process and phase-space under study.

1.6.4 Multi-purpose Monte Carlo generators

Experimental data needs to be interpreted in the context of a theoretical model, such as the SM. In order to do this, precise predictions of the detector output are required. As we have seen, the description of the evolution of the partonic final states and the subsequent hadronization process is not available in perturbation theory. In order to compare partonic cross-sections to experimental data, effects originating from detection, hadronization and the parton evolution have to be subtracted first. Alternatively, the predictions may be folded to include these effects. In either case, these corrections are possible thanks to the general-purpose *Monte Carlo* (MC) event generators.

MCs are software tools that implement fixed-order calculations of the hard-

scatter and include treatments of the *Parton Shower* (PS), hadronization and detection processes. They use physically-inspired phenomenological models where perturbative QCD breaks down. Despite this, MCs provide detailed simulations of high-energy collisions to a good approximation. These simulations can be used to bring data to that level where purely fixed-order calculations are available, as mentioned above, and also to analyze directly uncorrected experimental data in the context of SM measurements or searches of new particles. MC simulations rely on two key points:

1. Factorization can be used to model physics occurring at different scales: in a stochastic process, one only needs to know what happened in the previous length scale to model the behaviour of the next one.

2. The expected value of a function can be approximated by computing the average value of the function by sampling many times from the underlying probabilistic distribution[12].

The first step of a MC generator is the ME calculation of the partonic scattering under study at fixed-order in perturbation theory. The radiation emitted by the resulting colored partons is simulated with the PS generator. The first challenge is related to the ME+PS interface: the ME emissions of hard and wide angle radiation are preferred, whilst soft and collinear radiation is properly captured by the parton showers. At NLO, however, there is a phase space overlap between the first real emission in the ME and the hardest emission from the parton shower. In order to avoid double counting, a matching procedure is carried out consisting on regulate one of the two emissions: the MadGraph5_aMC@NLO [71, 72] generator, for example, removes the emission in the ME calculation whilst the Powheg [73, 74] generator, on the contrary,

[12]This is why the number of MC simulated events required is quite large and must exceed, in any case, the number of measured events so the averaging is useful.

introduces a resummation scale h_{damp} which regulates the emission in the PS instead. Regardless the matching procedure adopted, the fixed-order accuracy is preserved.

All parton showers describe the stochastic evolution of the colored partons with a model inspired by the DGLAP equations, since they allow to evolve radiating particles from one scale to another. Starting from a initial configuration with n partons and cross-section σ_n, in the soft and collinear limit[13], the branching j that certain parton i undergoes modifies σ_n as follows [75]:

$$d\sigma_{n+1} = \sigma_n \sum_{i,j} \frac{\alpha_S}{2\pi} \frac{dk^2}{k^2} P_{ji}^T(z) dz \ , \tag{1.43}$$

where $P_{ji}^T(z)$ are the splitting kernels related to the DGLAP splitting functions, z is the momentum fraction carried by the emitted parton and k is a variable proportional to the relative transverse momentum of the emitted parton with respect to the parent parton. The parton shower may be developed by adding successive parton branchings one at a time. But not all branchings should be allowed, since we cannot resolve particles beyond Λ_{QCD}: at this level, they form hadrons. This leads to define a infrared cut-off parton shower scale $Q_0 >> \Lambda_{\text{QCD}}$ which prevents unresolved particles and potential soft and collinear divergencies occurring when $k \to 0$. It is introduced by means of the so-called Sudakov form factor $\Delta(Q^2, q^2)$ [76, 77]:

$$\begin{aligned} \Delta(Q^2, q^2) &= \exp\left[-\sum_f \int_{q^2}^{Q^2} \frac{dk^2}{k^2} \frac{\alpha_S}{2\pi} \int_{Q_0^2/k^2}^{1-Q_0^2/k^2} P_{ji}^T(z) dz \right] \\ &\sim \exp\left[-C_F \frac{\alpha_S}{2\pi} \log^2\left(\frac{Q^2}{Q_0^2}\right) \right] \ , \end{aligned} \tag{1.44}$$

which determines the unconditional survival probability for a parton not to

[13] In this limit, it holds that $\frac{d\theta^2}{\theta^2} = \frac{dk^2}{k^2}$ with θ as the opening angle between parent and children and k the children's momentum with respect to the parent's direction.

undergo a branching process between two scales Q^2 and q^2. The scale Q is indeed an ordering variable that can be arbitrarily chosen[14]. The q^2 value is randomly sampled from the underlying probability distribution of the previous branching, given by the derivative of $\Delta(Q^2, q^2)$. The Sudakov factor brings up three possible scenarios:

- The parton evolves from Q^2 to q^2 without undergoing a branching splitting, provided $q^2 >> Q_0^2$.
- The parton evolves from Q^2 to q^2 undergoing a branching splitting with $k^2 = q^2$ following the splitting kernel form $P_{ji}^T(z)$, provided $q^2 >> Q_0^2$.
- The parton shower stops for the parton because the generated scale q^2 is lower than Q_0^2.

This probability is embedded into Eq. 1.43 so the PS dynamics is essentially defined. From the second line of Eq. 1.44 we see that this Sudakov factor resums the largest logarithmic enhancements at α_S order, which is the reason why PS is said to be at LL accuracy. The PS treatment can be applied not only to final state partons but also to the initial state ones, with further modifications as including the PDFs. The extra radiation produced before and after the hard-scatter is referred to as *Initial-* and *Final-State Radiation* (ISR/FSR), respectively. Such radiation populates typically with soft particles the final state, but it may also give rise to hard emissions that alter the phase space configuration with, for instance, hard jets.

After the PS does its job, all remaining colored partons in the final state are combined into colorless hadrons. The hadron formation is described with physically-inspired phenomenological hadronization models. They are

[14] PYTHIA generator [78] uses a p_T-ordered shower [79] and HERWIG++[80] uses angular ordering instead[81].

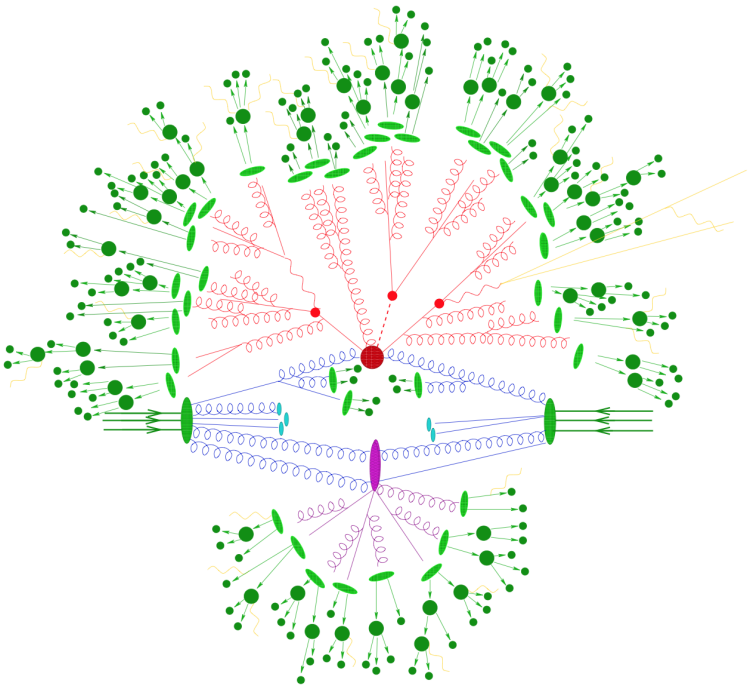


Figure 1.10: Schematic representation of a hadron-hadron collision as simulated by a Monte Carlo generator: the incident hadrons are represented through three valence quarks as straight green lines. The primary hard parton collision, big red blow dominating the center of the figure, gives three smaller blows as a result. Each of them is surrounded by a tree-like structure representing the parton shower. Parton-to-hadron transitions can be identified as light green blobs, while dark green blobs stand for hadron decay. A secondary collision with its corresponding dynamics is indicated by the purple blow. Figure taken from Ref. [82].

parametrized in such a way that can be tuned to data. The Lund string model [83, 84], used by the PYTHIA [78] generator, is based on the observation that the quark-antiquark potential rises linearly with the distance between quarks in a meson system. So $q\bar{q}$ pairs are grouped into strings with an energy given by such potential: if the energy is large enough, the string breaks into two further strings complemented by another $q\bar{q}$ pair that pops out from the vacuum. If not, the hadron is deemed to be stable. During the hadronization process, unstable particles and hadrons decay. There are specific MC generators devoted to simulate certain types of decays, as the EvtGen[85], especially designed for B- and D-hadron decays.

Besides ISR/FSR and hadronization, there are further processes which any MC generator must include in order to provide a more realistic picture. The fact is that there might be particle production not associated to the leading parton-parton process under study, originating from secondary interactions involving partons that belong to the same colliding hadrons. The color configuration or topology of the partons present in the colliding hadrons may cause non-trivial changes in the evolution of the subsequent cascade of particles, which typically populate the final state with a uniform background of soft particles, known as *Underlying Event* (UE). The multiplicity (*i.e.* number of particles) of the final state is a feature thay may depend of the initial color configuration of the colliding system, so several *Color Reconnection* (CR) models are developed to study their influence on final-state partons and kinematics [86, 87]. UE effects are commonly described with *Multi-Parton Interactions* models in the MC generators. Figure 1.10 displays a schematic representation of the main parts forming the structure of a simulated MC event. Finally, the interaction of all stable hadrons present in the final state with the dectector must be included too in the MC simulation chain, which requires a precise knowledge of the detection setup used: detector

geometry and sub-systems, inactive materials, etc..

1.7 Effective field theories

We have seen that, in the context of particle colliders, fixed-order and resummed predictions of observables in perturbation theory require a renormalized quantum field theory and the factorization theorem, thanks to which the short- and long-distance physics contributing to given process can be safely factorized. Now, we are in place to introduce a very powerful tool that allows to simplify the treatment of a complex system where several scales are involved.

Effective Field Theories (EFTs) are QFTs that provide a description of a physical system considering only the relevant information that determines the dynamics known at the characteristic scale. A key point in the EFTs formulation is that physics effects originating from very different scales are weekly coupled. The description can be therefore expanded by incorporating contributions coming from a much higher scale. In this sense, EFTs can be understood as a low-energy limit of a higher-energy yet unknown theory that solves a physics problem with the degrees of freedom relevant at the scale of such problem [88]. An illustrative example is QED: it offers a valid description when probed at energies much lower than the unification scale, $m_e << m_W$.

EFTs incorporate the usual ingredients to compute an experimental observable with some finite uncertainty in perturbation theory: degrees of freedom (fields content), symmetries (fixing the interactions) and some renormalization process. The terms in the Lagrangian are explicitly organized according to a series expansion around given small parameter δ, called power counting parameter. Calculations can therefore be done at some order in δ, determining in this way their precision and turning the number of free parameters

in the theory finite. The δ parameter is usually identified with the ratio of two energy scales Q/Λ, where Q represents low-energy physics which underlying theory is well known and Λ signals the UV regime where the formal theoretical description is out-of-reach.

The *Soft and Collinear Effective field Theory* (SCET), for example, is an EFT devoted to describe energetic QCD processes where the final states have small invariant mass compared to the center-of-mass energy of the collision, as in jet production in high-energy proton-proton collisions. The expansion parameters are actually three: Λ_{QCD}/Q, M_J/Q and $\alpha_S(Q)/(4\pi)$ where Q represents the center-of-mass energy of the hard-scatter and M_J the invariant mass of a jet in the final state.

The EFT extension of the SM, known as SMEFT, is of special interest for this thesis. In the SMEFT, the SM Lagrangian represents the low-energy or long-distance physics at the EW scale Q. The relevant degrees of freedom are the usual SM fields which mass m is given by the EW vacuum expectation value, as reflected in Eq. 1.15. Short-distance effects originating at a much higher-energy scale Λ are added as a power series of $1/\Lambda$. The SMEFT Lagrangian reads as follows:

$$\mathcal{L}_{\mathrm{SMEFT}} = \sum_{D\geq 4} \frac{\mathcal{L}^{(D)}}{\Lambda^{D-4}} = \mathcal{L}^{(4)}_{\mathrm{SM}} + \frac{\mathcal{L}^{(5)}}{\Lambda} + \frac{\mathcal{L}^{(6)}}{\Lambda^2} + \dots \ . \tag{1.45}$$

The SM Lagrangian corresponds to the first term of the expansion (the lowest-energy one) having dimension $D = 4$ in energy. The next terms in the expansion are higher-dimensional ($D > 4$) and are formulated as the product of two factors: operators $O^{(D)}$ representing the high-energy effects and Wilson coefficients $C^{(D)}$ that regulate the strength of such effects. In this way, Eq. 1.45 takes the form:

$$\mathcal{L}_{\mathrm{SMEFT}} = \mathcal{L}^{(4)}_{\mathrm{SM}} + \sum_{D>4,i} \frac{C^{(D)}_i}{\Lambda^{D-4}} O^{(D)}_i \ , \tag{1.46}$$

where i runs over all operators allowed for given dimension D. The $O^{(D)}$ operators involve the same components as those of the SM: light fields (fermions and bosons) and derivatives. At energies below Λ, the behaviour of the different operators are classified according to their dimensionality. On the one hand, those with $D > 4$ are said irrelevant in the sense that their effects are suppressed by powers of $1/\Lambda^{D-4}$ and thus are small at low energies. The dominant contributions are expected to come from the first terms of the expansion. On the other, operators with $D < 4$ are called relevant since they are not suppressed at all, and those with $D = 4$ are deemed marginal because quantum effects may make them fall on either side of relevance and irrelevance.

In general, all operators are required to respect the underlying symmetries and conserve the quantum numbers of the SM[15]. None of the dimension-five operators entering in $\mathcal{L}^{(5)}$ fulfill the two conditions. For the next higher-order term $\mathcal{L}^{(6)}$, it is possible to derive a complete basis of independent dimension-six operators that satisfy these requirements and fully characterize the first term $1/\Lambda^2$ of the expansion and account partially for the contribution proportional to $1/\Lambda^4$:

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \frac{1}{\Lambda^2}\sum_i C_i O_i + \frac{1}{\Lambda^4}\sum_{i,k} C_j C_k O_{jk} + O(1/\Lambda^6)\ . \tag{1.47}$$

In this equation, terms proportional to $1/\Lambda^2$ represent the interference between SM amplitudes with those featured by one dimension-six operator. Analogously, terms proportional to $1/\Lambda^4$ stem from the square of amplitudes involving one insertion of the same dimension-six operator or from amplitudes involving two insertions of different dimension-six operators. The Wilson coefficients $C_i^{(D)}$ accompanying such operators represent the strength with which the unknown physics contained in the corresponding operators $O_i^{(D)}$ couple to the SM. Their

[15] In SMEFT. In other EFTs, terms that violate some symmetry (as chirality or the lepton number) are introduced on purpose to interpret such violations in light of physics beyond the Standard Model.

values are not fixed from first-principles. Experimental results are needed to constrain the possible values of $C_i^{(D)}$ according to the observed deviations from the SM predictions.

The SMEFT can be used to interpret particle collider data in a nearly model-independent way. The only assumptions inherent are that Λ is higher than the scale directly probed by the experiments and that the new physics respects the SM symmetries and conservation of quantum numbers. It is a powerful tool to search for subtle patterns that could reflect the presence of higher-scale new physics, whilst the consistency with all other measurements where the SM is successful is kept.

Chapter 2

Top quark physics

The top quark was discovered by the DØ and CDF collaborations at Fermilab's collider Tevatron in 1995 from proton-antiproton ($p\bar{p}$) collisions data, confirming the three-generation structure of the SM and allowing the exploration of the top quark sector. The top quark was found to be the most massive of the elementary particles discovered at that time, with a mass m_t [15, 16]:

$$\begin{aligned} m_t &= 199^{+19}_{-21}\ (\text{stat})\ \pm\ 22\ (\text{syst})\ \text{GeV} \qquad (\text{D}\emptyset)\ , \\ m_t &= 176 \pm\ 8\ (\text{stat})\ \pm\ 10\ (\text{syst})\ \text{GeV} \qquad (\text{CDF})\ . \end{aligned} \tag{2.1}$$

Since these first measurements, the determination of the top quark mass has reached sub-% precision. The most recent combination of the Tevatron results yields $m_t = 174.30 \pm 0.65$ GeV [89], with an impressive relative precision of 0.37%. The top quark is 35 times heavier than its weak-isospin partner the b quark and, after the discovery of the Higgs boson and the determination of its mass $m_H = 125.10 \pm 0.14$ GeV [26], the top quark remains the heaviest SM particle. Due to that, the top quark influences many quantitative and conceptual aspects within the SM and many extensions thereof. Together with the Tevatron, the LHC collider (introduced in Chapter 3) assumes a major role in characterizing

the properties of the top quark from proton-proton (pp) collisions data. In this chapter, devoted to a description of the aspects related to the top quark physics that are most relevant for this thesis, we will focus on the charge asymmetry A_{C} and the top quark mass in the context of hadron colliders. The former can be used to confront SM predictions and to check the presence of physics BSM. The latter is related to a more subtle question regarding the interpretation of the top quark mass measurements performed with different methods.

2.1 Mass and width

The large top quark mass determines profoundly its phenomenology. The probability of a heavy particle to decay into lighter particles is encoded in the width decay Γ. As discussed in Section 1.3.3, quarks mixing permits the top quark to decay into all down-type quarks, and not only into its weak isospin partner, via charged-current interactions. The total top width results from adding up the partial width of all allowed channels:

$$\Gamma_t = \sum_q \Gamma(t \to Wq), \quad q = b, s, d \ , \tag{2.2}$$

where $\Gamma(t \to Wq)$ is proportional to the corresponding CKM matrix element $|V_{tq}|^2$. In the SM, $\Gamma_t \propto m_t^3$. The latest combination of the most precise ATLAS (see Section 3.2), CMS and DØ measurements yield [26]:

$$\Gamma_t = 1.42^{+0.19}_{-0.15} \text{ GeV} \ , \tag{2.3}$$

in agreement with the NLO SM prediction of 1.35 GeV [26]. The lifetime is inversely proportional to the width decay, $\tau \sim 1/\Gamma$. With a lifetime of $\tau_t \sim 10^{-25}$ s, the top quark becomes one of the shortest-lived particles in the SM, which means

that it decays very rapidly. It occurs in such a short time scale that the top quark does not hadronize, since confinement operates over larger time scales, $\tau_{\mathrm{QCD}} \sim 1/\Lambda_{\mathrm{QCD}} \sim 10^{-24}$–$10^{-23}$ s. The study of the top quark properties gives direct access to the understanding of free-quark features. Effects of the decay width can be isolated when considering heavy hadrons that decay after hadronization, for instance.

The branching ratio $\mathcal{B}$ is defined as the probability of a given particle to decay into a single channel over all possible channels, and can be expressed as:

$$\mathcal{B}(t \to Wq) = \frac{\Gamma(t \to Wq)}{\Gamma_t} = \frac{|V_{tq}|^2}{|V_{tb}|^2 + |V_{ts}|^2 + |V_{td}|^2} \ . \tag{2.4}$$

The $|V_{tb}|$ element is expected to be the largest one since the others are off-diagonal. It has been found that the top decays into Wb in the $|V_{tb}|^2 \sim 98.8\%$ of the cases [26], so $|V_{tb}| >> |V_{ts}|, |V_{td}|$.

2.2 Standard Model consistency and tests

One of the reasons why the top quark has captured the attention of the physics community comes from its special influence on many aspects of the SM. As we saw in the previous chapter, many parameters of the EW sector are connected to each other. According to Eqs. 1.14 and 1.17 v, λ and μ parameters fix the mass of the Higgs and weak bosons. In turn, the weak angle depends on m_W and m_Z, Eq. 1.18. One could find many relations between observables and the free parameters of the theory. Some observables can be precisely measured and calculated, if accounting properly for quantum corrections beyond LO in perturbation theory. The top quark plays an important role here owing to its large mass, inducing sizeable corrections through loop diagrams.

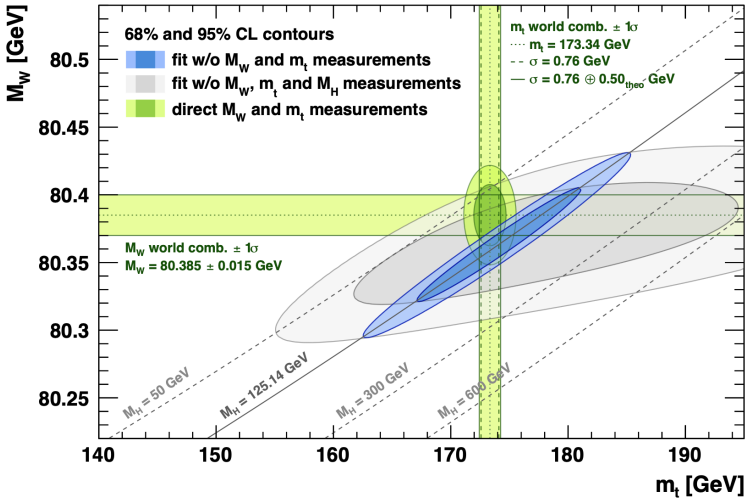


Figure 2.1: Comparison of world combined m_W and m_t results with contours at 68% ad 95% Confidence Level computed from scans of such parameters, excluding (gray) and including (blue) m_H in the fit. Figure from Ref. [90].

In this sense, global EW fits test the internal consistency of the SM by fixing some free parameters of the theory from very precise experimental EW measurements ($\sin^2(\theta_W), \alpha_{\text{QED}}, m_H$, etc.) and making accurate predictions as a function of the remaining free parameters. Such predictions are then confronted with the experimental direct measurements. Figure 2.1 shows an example of this exercise for the masses of the W boson, the Higgs boson and the top quark. On the one hand, the experimental world combination of the direct measurements of m_W and m_t are represented separately by the horizontal and vertical lines, respectively, with their associated total uncertainty band. The green ellipse stands for the correlation between the two. On the other hand, a scan of the *Confidence Level* (CL) predictions of m_W and m_t (without information of the direct measurements) through global EW fits are shown in two possible cases: including (blue) or not (gray) the direct m_H measurement in the global fit. We see that both indirect determinations are in agreement with the direct measurements within the current uncertainties, which gives support to the internal consistency of the SM. It is important to note that when the direct measurement of m_H is not

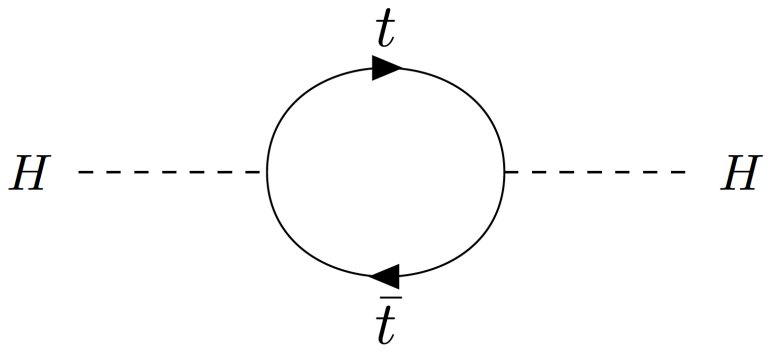


Figure 2.2: One-loop correction to the Higgs propagator involving top quarks.

included in the global fit, the indetermination on its value leads to an additional uncertainty on the predicted m_W and m_t and bigger ellipses that impoverished the quality of the test. Analogously, more precise measurements of the top quark mass would enable more stringent tests by reducing the size of the experimental green ellipse.

The Higgs boson mass is a very illustrative example of the large corrections induced by top quark loops. The NLO Feynman diagram is depicted in Figure 2.2. Such corrections are of the order of the squared top quark mass m_t^2. The physically observable Higgs mass would be:

$$m^2_{H,\text{phys}} = m^2_{H,\text{bare}} + \mathcal{O}(m_t^2) \ . \tag{2.5}$$

The quantum corrections $\mathcal{O}(m_t^2)$ are large and, in order to match the experimentally measured $m_{H,\text{phys}}$, the value of the bare Higgs mass has to be carefully adjusted. This is known as *fine-tuning*, characteristic of unnatural theories where the quantum corrections are larger than the observed quantities. If the heaviest state of the theory lies in much higher scales, loop corrections would bring $m_{H,\text{bare}}$ (of the order of the EW scale) up there, ruining the hierarchy between large and small scales. A more complex Higgs structure than that introduced in the SM through SSB would help to keep the Higgs boson naturally

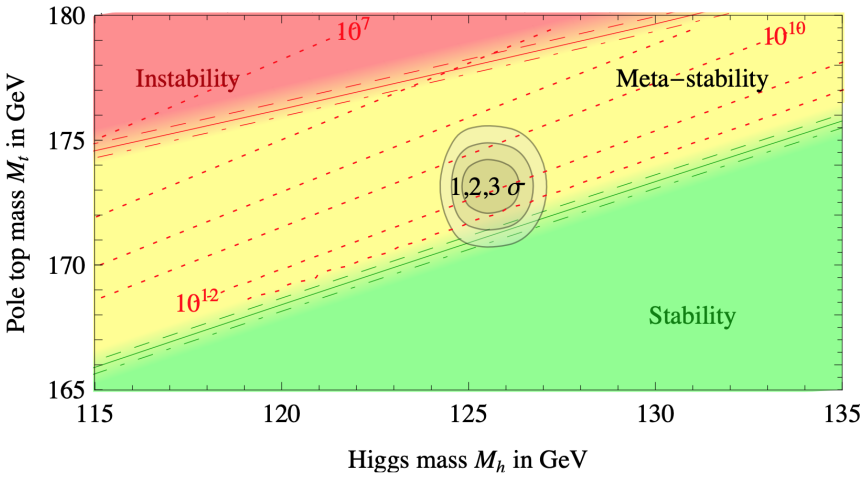


Figure 2.3: Stability of the EW vacuum as a function of the measured m_t and m_H with 1,2 and 3σ CL contours. Figure taken from Ref. [91].

light. One of the most popular SM extensions, supersymmetry, postulates the existence of supersymmetric fermion (boson) partners for each boson (fermion) particle present in the SM model, balancing the loop corrections in a natural way and making fine-tuning unnecessary.

The top quark mass also induces large corrections in the quartic coupling λ of the Higgs field through closed loops. The value of λ determines the stability of the EW vacuum. Therefore, it ultimately depends on the measured top quark and Higgs boson masses. The stability of the EW vacuum as a function of these two parameters is shown in Figure 2.3. The measured m_H falls right below the value that would guarantee the vacuum stability, ~ 130 GeV. With the current m_H and m_t measurements the unstable scenario can be confidently ruled out, where the stable and pseudo-stable hypotheses are still plausible within the current precision, limited by the uncertainties in the top quark mass determination.

From the discussion of this section we conclude that experimental measurements of the top quark mass are required to be as precise as possible. We shall see in the following the current methods employed to measure m_t, with special focus on their meaning and interpretation.

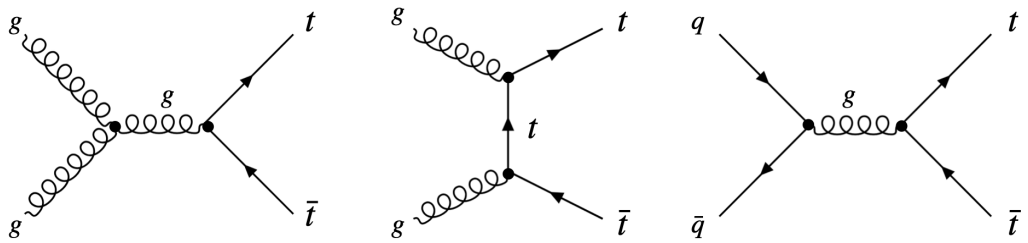


Figure 2.4: Leading Order Feynman diagrams representing the dominant production modes of $t\bar{t}$ pairs in hadron collisions. Leftmost and center diagrams refer to $s-$ and $t-$channel in gluon-gluon fusion. The rightmost diagram stands for the $q\bar{q}$ annihilation process.

2.3 Top quark properties at hadron colliders

2.3.1 Single and top pair production

At hadron colliders, top quarks are produced predominantly in pairs at LO in QCD through $gg \to t\bar{t}$ and $q\bar{q} \to t\bar{t}$ processes, schematically represented on Figure 2.4. The contributing cross-section from each process depends on the center-of-mass-energy and the colliding hadrons. In the LHC, where currently protons are collided at $\sqrt{s}$ = 13 TeV, the contributions are about 90% and 10% for gluon-gluon fusion and $q\bar{q}$ annihilation, respectively. On the contrary, at the Tevatron the dominant process is the $q\bar{q}$ annihilation. The most precise calculation of the inclusive $t\bar{t}$ cross-section is done at NNLO+NNLL in QCD [92]. Predictions of differential cross-sections are also available at NNLO in QCD [93]. With this formal precision, calculations must account for EW corrections since they become relevant [94].

Top quarks can also be produced singly in flavour-changing processes, probing charged-current weak interactions. At the LHC energies, the main contributions come from the exchange of a virtual W boson and the so-called associated productions with a real W boson in the final state. The inclusive $t\bar{t}$ and single-top quark production cross-sections have been measured at the Tevatron and LHC

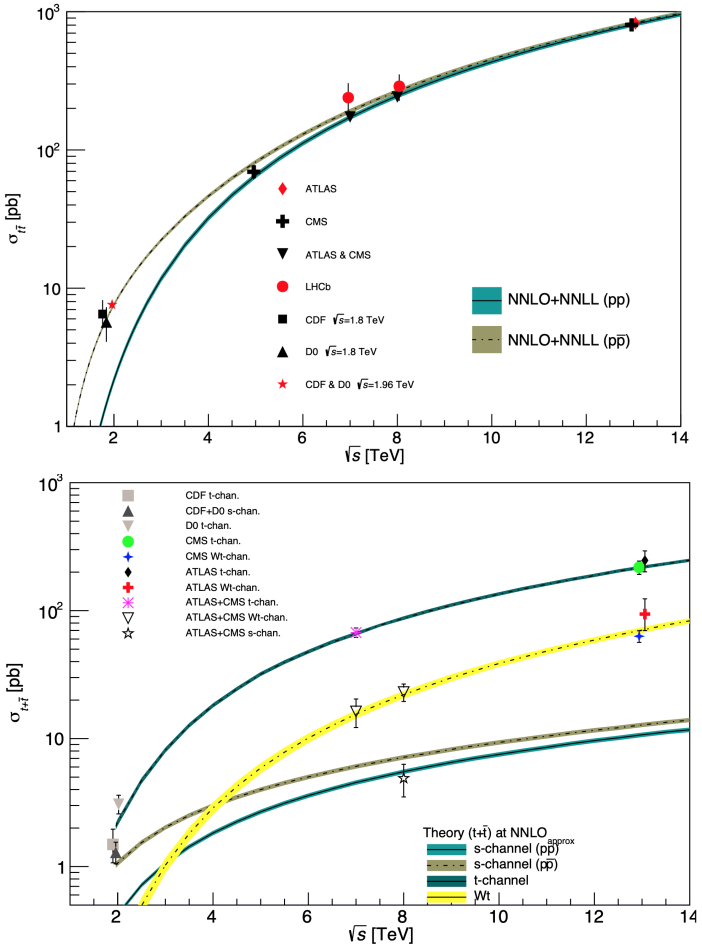


Figure 2.5: Comparison of measured and predicted total production cross-section of $t\bar{t}$ (upper panel) and single-top quarks (lower panel) as a function of the center-of-mass energy $\sqrt{s}$. Figure from Ref. [26].

colliders spanning several orders of magnitude, finding an excellent agreement with the most precise SM predictions as shown in the upper and lower panels of Figure 2.5, respectively.

As described in the previous section, the top quark decays mainly to Wb. In turn, the W^+ (W^-) boson may decay leptonically as $\bar{l}\nu_l$ ($l\bar{\nu}_l$) or hadronically into $q\bar{q}'$ ($q'\bar{q}$) pairs with branching ratios of $\sim$ 33% and $\sim$ 67%, respectively. The τ lepton decays before detection though, with $\mathcal{B}(\tau \to \text{hadrons})$ 66% and $\mathcal{B}(\tau \to l_{e,\mu}\nu_{e,\mu}\nu_\tau)$ 33%. From an experimental point of view, it is very convenient to classify events containing $t\bar{t}$ pairs according to the number of leptons in the

final state: they can be detected precisely and used to suppress large backgrounds coming from soft QCD processes. The all-hadronic channel has no leptons in the final state and represents the 55% of total events. The semileptonic channel has exactly one electron or muon and accounts for 38% of the cases, whilst the dileptonic channel collects the remaining 7%.

In this thesis, several aspects related to top quark physics are studied from $t\bar{t}$ events in the three channels, so events containing single-tops will represent a source of background.

2.3.2 $t\bar{t}$ asymmetry

In the SM, the lowest-order contribution to $t\bar{t}$-production is symmetric under charge conjugation. However, an asymmetry between tops and antitops arises due to the interference with higher-order diagrams in the $q\bar{q}$ and qg initial states, which do not respect such symmetry. As a consequence, top (antitop) quarks are preferentially emitted in the direction of the incoming quark (antiquark). This motivates the definition of two observables based on the observable angular variable called rapidity:

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right) . \tag{2.6}$$

The large the value of this quantity, the close the particle travels to the beam line. In $p\bar{p}$ collisions, the dominant contribution to the asymmetry originates from $q\bar{q}$ annihilation, since the largest fraction of high-energetic quarks (antiquarks) is carried in the proton (antiproton) beam. The rapidity distribution of the produced $t\bar{t}$ pairs exhibits an excess of tops (antitops) populating that region pointed by the incoming proton (antiproton) beam, as illustrated in the leftmost panel of Figure 2.6. This motivates the definition of the forward-backward

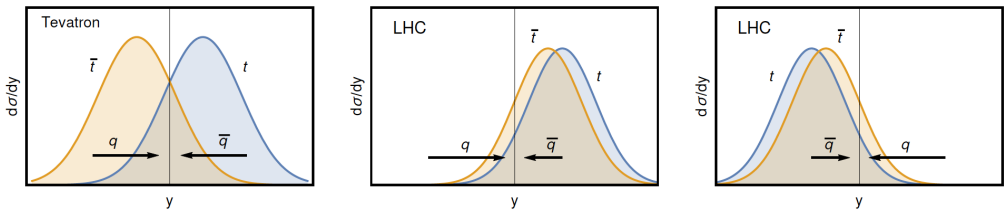


Figure 2.6: Rapidity distributions of top and antitop quarks produced in pairs at the Tevatron (left) and the LHC (centre, right). Figure taken from Ref. [99].

asymmetry A_{FB} in the $t\bar{t}$ rest-frame:

$$A_{\mathrm{FB}} = \frac{N(\Delta y > 0) - N(\Delta y < 0)}{N(\Delta y > 0) + N(\Delta y < 0)} \ , \tag{2.7}$$

where N is the number of events for which the difference between the rapidity of the top-antitop pair $\Delta y \equiv y_t - y_{\bar{t}}$ is greater or smaller than zero. The DØ and CDF experiments reported the first measurement of A_{FB} [95, 96], resulting in a combined value of $A_{\mathrm{FB}} = 0.124 + 0.025$ [97]. The SM prediction at NNLO in QCD + NLO EW corrections, $A_{\mathrm{FB}}^{\mathrm{SM}} = 0.095 + 0.007$ [98], is compatible within one standard deviation.

In contrast, the rapidity distributions of top and antitop quarks in pp collisions do not exhibit this asymmetry due to the symmetric nature of the initial state, as illustrated in the central and rightmost panels of Figure 2.6. Therefore, A_{FB} is no longer useful. In pp collisions the dominant gluon-gluon process, symmetric under charge conjugation at all orders, partially washes out the asymmetry originating from $q\bar{q}$ and qg initial states. However, the absolute rapidity of the top/antitop depends on the momentum carried by the incoming parton. According to Figure 1.9, valence quarks carry larger fractions of the hadron momentum than the sea antiquarks. Hence a small asymmetry arises between top quarks produced, on average, with large absolute rapidity and antitops preferentially populating the central region (see again Figure 2.6). We can take advantage of this feature by

formulating a central-forward asymmetry referred to as charge asymmetry A_{C}:

$$A_{\mathrm{C}} = \frac{N(\Delta|y| > 0) - N(\Delta|y| < 0)}{N(\Delta|y| > 0) + N(\Delta|y| < 0)} \ , \tag{2.8}$$

with $|y| = |y_t| - |y_{\bar{t}}|$. The inclusive SM prediction is one order of magnitude smaller than that expected in the Tevatron: $A_{\mathrm{C}} = 0.0111 \pm 0.0004$ at $\sqrt{s} = 8$ TeV [100]. The ATLAS and CMS collaborations reported $A_{\mathrm{C}} = 0.009 \pm 0.005$ (stat + syst) [101] and $A_{\mathrm{C}} = 0.011 \pm 0.013$ (stat + syst) [102] respectively, in good agreement with the expected value.

The dilution caused by gg-initiated processes is enhanced at high energies, but it can be partially suppressed by selecting events with large rapidity or $t\bar{t}$ invariant mass, $m_{t\bar{t}}$ [99]. The ATLAS collaboration provided differential A_{C} measurements as a function of the $m_{t\bar{t}}$ at $\sqrt{s} = 8$ TeV [101]: the largest asymmetry was found to be $A_{\mathrm{C}} = 0.068 \pm 0.044$ for events where $m_{t\bar{t}} > 900$ GeV. A dedicated analysis was developed to optimize this study in a very boosted regime [103].

It is important to stress that, although they are different, the A_{FB} and A_{C} observables are originated by the same underlying mechanism, so they are correlated. In addition, any deviation of the experimentally measured values from the SM expectation may hint to physics BSM. Figure 2.7 shows how several SM extensions affect the inclusive A_{C} and A_{FB}.

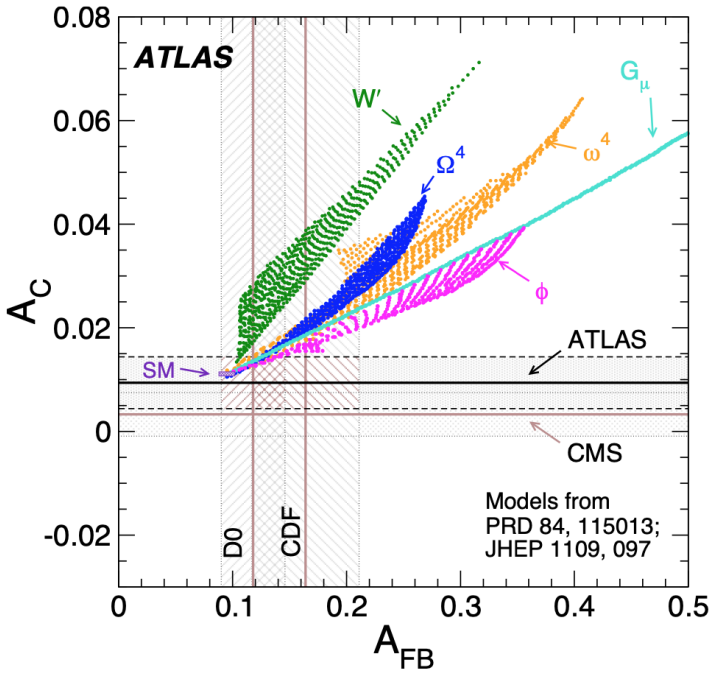


Figure 2.7: Comparison of measured inclusive charge asymmetries at the LHC (ATLAS and CMS in vertical lines) and the Tevatron (DØ and CDF, horizontal lines) to Standard Model predictions. Experimental values are represented with the associated uncertainty band at 68% Confidence Level. Several models incorporating beyond Standard Model effects are also included: W', heavy axigluon (G_μ), scalar isodoublet (ϕ), color-triplet scalar (ω^4) and color-sextet scalar (Ω^4). Figure from Ref. [101].

2.3.3 Four-quark operators in effective field theory

In the context of effective field theories introduced in Section 1.7, we consider dimension-six effective operators that may affect the $t\bar{t}$ production at tree-level and beyond in hadron colliders and therefore explain any deviation between the SM predictions and the experimental observations. Paying special attention to those processes that may impact the A_{C}, we limit ourselves to operators involving two quarks in the initial and final state, respectively, referred to as top four-quark operators. The operator O_{tG} modifies the $t\bar{t}G$ vertex and can have a sizeable contribution too. In the Warsaw basis [104], a complete list of effective operators relevant in this thesis is given in Eq. 2.9, where the four-quark ones are organized following their chiral structure:

$$
\begin{aligned}
O_{tG} &= \left(\overline{Q}\sigma^{\mu\nu}\lambda^A t\right)\phi G^A_{\mu\nu} \ , \\
\overline{L}L\overline{L}L: & \\
O_{Qq}^{(8,1)} &= \left(\overline{Q}\gamma_\mu\lambda^A Q\right)\left(\overline{q}\gamma^\mu\lambda^A q\right) \ , \\
O_{Qq}^{(1,1)} &= \left(\overline{Q}\gamma_\mu Q\right)\left(\overline{q}\gamma^\mu q\right) \ , \\
O_{Qq}^{(8,3)} &= \left(\overline{Q}\gamma_\mu\tau^I\lambda^A Q\right)\left(\overline{q}\gamma^\mu\tau^I\lambda^A q\right) \ , \\
O_{Qq}^{(1,3)} &= \left(\overline{Q}\gamma_\mu\tau^I Q\right)\left(\overline{q}\gamma^\mu\tau^I q\right) \ , \\
\overline{R}R\overline{R}R: & \\
O_{tu}^{(8)} &= \left(\bar{t}\gamma_\mu\lambda^A t\right)\left(\overline{u}\gamma^\mu\lambda^A u\right) \ , \\
O_{tu}^{(1)} &= \left(\bar{t}\gamma_\mu u\right)\left(\overline{u}\gamma^\mu t\right) \ , \\
O_{td}^{(8)} &= \left(\bar{t}\gamma_\mu\lambda^A t\right)\left(\overline{d}\gamma^\mu\lambda^A d\right) \ , \\
O_{td}^{(1)} &= \left(\bar{t}\gamma_\mu d\right)\left(\overline{d}\gamma^\mu t\right) \ , \\
\overline{L}L\overline{R}R: & \\
O_{Qu}^{(8)} &= \left(\overline{Q}\gamma_\mu\lambda^A Q\right)\left(\overline{u}\gamma^\mu\lambda^A u\right) \ ,
\end{aligned}
\tag{2.9}
$$

$$O_{Qd}^{(8)} = \left(\overline{Q}\gamma_\mu\lambda^A Q\right)\left(\overline{d}\gamma^\mu\lambda^A d\right) \ ,$$
$$O_{Qt}^{(8)} = \left(\overline{Q}\gamma_\mu\lambda^A Q\right)\left(\overline{t}\gamma^\mu\lambda^A t\right) \ ,$$

$\overline{L}R\overline{R}L$:

$$O_{Qu}^{(1)} = \left(\overline{Q}u\right)\left(\overline{u}Q\right) \ ,$$
$$O_{Qd}^{(1)} = \left(\overline{Q}d\right)\left(\overline{d}Q\right) \ ,$$
$$O_{Qt}^{(1)} = \left(\overline{Q}t\right)\left(\overline{t}Q\right) \ ,$$

where q and Q stands for the weak left-handed doublets of the first two and the third generation, respectively. u, d represent the weak right-handed singlets of the first and second generation, and t is the top right-handed singlet. τ^I are the Pauli matrices and λ^A the Gell-Mann matrices. All these operators contribute to the $u\bar{u}, d\bar{d} \to t\bar{t}$ reaction. The Feynman diagrams including the modified interaction vertex induced by these operators are depicted in Figure 2.8; whilst diagram (a) represents the SM contribution to the total amplitude, diagrams (b) and (c) are corrections to such amplitud due to the presence of the O_{tG} operator and any four-fermion operator, respectively. A review of the role played by these operators in top quark studies at hadron colliders is given in Refs. [105, 106]. Limits on the size of their possible contribution have been derived from a wide set of processes and observables through global fits [107–113]. In this thesis, the operators listed in Eq. 2.9 are constrained with inclusive and differential measurements of the A_C from $\sqrt{s} = 13$ TeV data collected by the ATLAS experiment at the LHC.

2.3.4 Top mass measurements

The top quark mass is the parameter measured with highest precision in the QCD sector. The story begins at the Tevatron after its discovery, with the final combined measurement yielding $m_t = 174.30 \pm 0.65$ GeV [89]. At the LHC era,

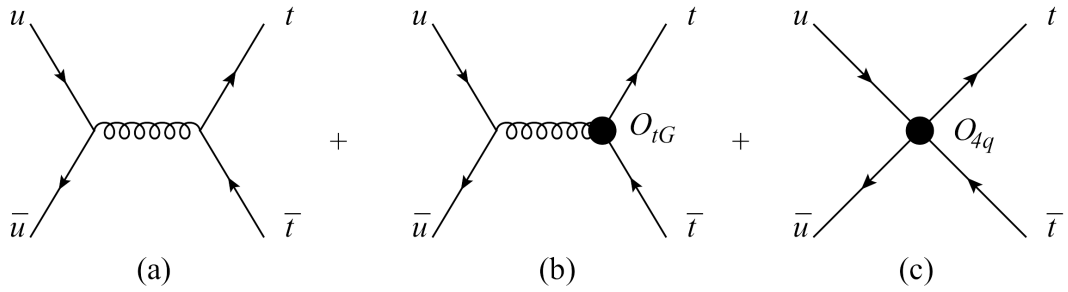


Figure 2.8: Feynman diagrams for $u\bar{u} \to t\bar{t}$ process. (a) diagram represents the Standard Model amplitude, and (b) and (c) the modified interaction vertex due to the insertion of dimension-six operators affecting the $t\bar{t}G$ coupling and four-fermion processes, respectively. Figure taken from Ref. [114].

a major effort is devoted towards the determination of the top quark mass as precisely as possible. Any strategy tailored to the extraction of the top quark mass heavily relies in the choice of an observable with high sensitivity to it. In this sense, we can distinguish two methods based on observables with mass sensitivities originating from different physics effects.

On the one hand, the direct measurements rely on observables built directly from top quark decay products that exhibit what is commonly called kinematic mass sensitivity. The top mass sensitivity of such observables translates into resonance and endpoint structures [50], as can be noted from the reconstructed top mass m_t^{reco} distribution shown in the left-hand plot of Figure 2.9. This sensitivity is initiated by hard reactions involving large scales $\gtrsim m_t$ and enhanced by dynamical QCD and EW effects produced at energy scales $\ll m_t$. The highest mass sensitivity comes from the location of the resonance structure, tied to m_t, but its shape and exact location/width arise from low-energy QCD ($\mathcal{O}(\Gamma_t)$) and EW effects, also limited by the experimental resolution. In this context, NLO-matched MCs are suitable thanks to their ability to provide a complete description of the whole process. Data and MC predictions can be compared directly at reconstruction or detector level in different ways. In the template method, real data is fitted with MC templates where the value of the top quark mass in the

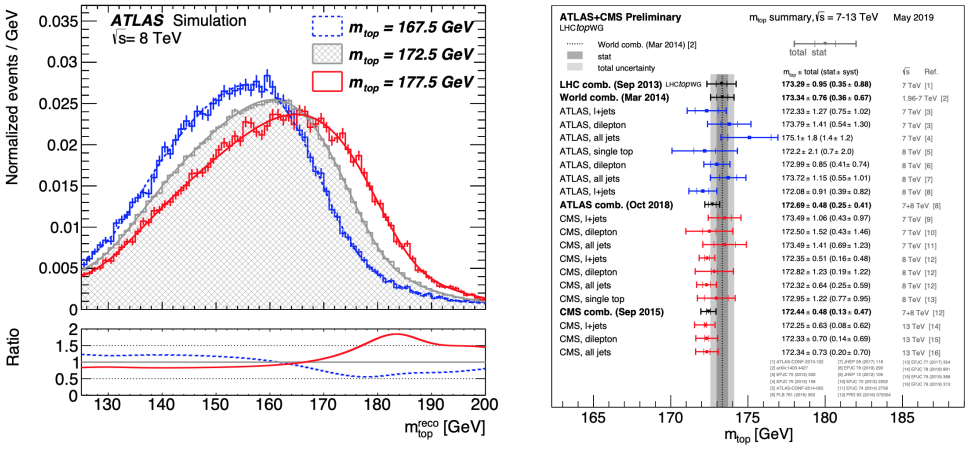


Figure 2.9: On the left, sensitivity of the reconstructed top quark mass from hadronic top decays (semileptonic channel) to three m_t^{MC} values in simulated events. On the right, summary of ATLAS and CMS direct top quark mass measurements. Figures from Refs. [115] and [116].

MC generator (m_t^{MC}) is varied. In the ideogram and matrix element methods, a likelihood assuming a $t\bar{t}$ production hypothesis for the whole final state is tested event-by-event in real data. The top quark MC mass is given by the best-fit or the highest cumulative likelihood. The right-hand panel of Figure 2.9 summarizes the state-of-the-art top quark mass direct measurements at the LHC. The latest combination performed by ATLAS and CMS are:

$$\begin{aligned} m_t^{\mathrm{MC}} &= 172.26 \pm 0.61 \text{ GeV} \quad \text{(ATLAS)} \ , \\ m_t^{\mathrm{MC}} &= 172.69 \pm 0.48 \text{ GeV} \quad \text{(CMS)} \ , \end{aligned} \tag{2.10}$$

where an impressive relative uncertainty around 0.30% is achieved. Thanks to the good precision of the MC generators and the high top mass sensitivity of the observables used, the direct measurements represent the most precise top quark mass determinations.

The total uncertainty quoted in the direct measurements includes the detector and statistical experimental uncertainties. It also accounts for uncertainties that cover some aspects of the MC implementation, such as the parton shower and

the hadronization modelling, generally assessed by using MC generators with different models. The ISR/FSR and UE treatment can be studied with variations of the MC tuning parameters. The uncertainty on the ME computations is taken into account too, usually evaluated by varying the renormalization and factorization scales in given ME calculation and/or using an alternative ME calculation. However, some limitations of the MC generators that may have a sizeable impact on the direct top quark measurements are not considered. For instance, the partonic description of the sharp structures proper to kinematic mass sensitive observables lacks of subleading QCD precision (NLL) in the parton shower (typically computed at LL), when such threshold structures are governed by soft and collinear radiation. This uncertainty cannot be estimated by comparing two different parton showers with the same LL accuracy. In addition, the LO treatment given to the top decay describes boosted tops, and many direct measurements profit from low-momenta tops. Whilst the impact of such limitations have not been quantified yet, the *Narrow Width Approximation* (NWA)[1] adopted by most of the current MC generators[2] has been studied in Ref. [117], showing that a direct top mass measurement can be affected at 0.5 – 1 GeV level.

The limitations of the MC generators, specially related to the treatment of the soft and collinear radiation and the non-perturbative effects, obscures the interpretation of the top MC quark mass in terms of a field-theoretical mass scheme. This is extensively discussed in Section 2.4.

On the other hand, the indirect measurements profit from inclusive and differential $t\bar{t}$ cross-sections that have an indirect mass sensitivity, *i.e.* originating

[1] In NWA, finite-lifetime effects that give rise to large logarithms of the type Γ_t/m_t are neglected.

[2] HERWIG 7 uses NWA and PYTHIA 8 employs a NWA supplemented with a Breit-Wigner-like distribution.

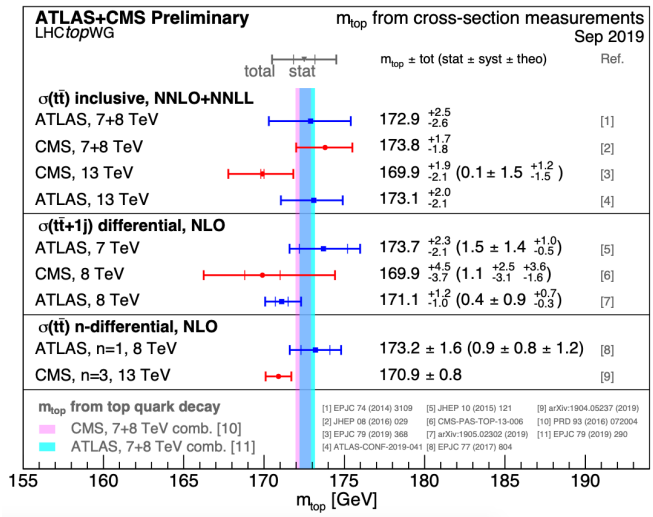


Figure 2.10: Summary of top quark pole mass determinations performed by the ATLAS and CMS experiments from inclusive and differential cross-sections. Vertical bands correspond to m_t^{MC} averages from $\sqrt{s}$ = 7 and 8 TeV data. Figure from Ref. [116].

from the hard scatter only. So the energy scale of the QCD dynamics generating the mass sensitivity is $\gtrsim m_t$. This enables the description from first-principles of parton-level cross-sections through resummed and fixed-order calculations[3]. In this context, the interpretation of the top quark mass as a field-theoretical mass becomes more clear. ATLAS and CMS have used NNLO+NNLL predictions of the inclusive $\sigma_{t\bar{t}}$ in the on-shell scheme to determine the top quark pole mass at several center-of-mass energies, as summarized in Figure 2.10. The latest results with data collected at $\sqrt{s}$ = 13 TeV are:

$$\begin{aligned} m_t^{\mathrm{pole}} &= 173.1^{+2.0}_{-2.1}\ \mathrm{GeV} \quad (\mathrm{ATLAS})\ , \\ m_t^{\mathrm{pole}} &= 169.9^{+1.9}_{-2.1}\ \mathrm{GeV} \quad (\mathrm{CMS})\ . \end{aligned} \tag{2.11}$$

The larger uncertainties in comparison to the direct methods result from uncertainties in the normalization of the cross-section and from the relatively

[3]For proper differential cross-section observables that exhibit a low sensitivity to parton shower or low-energy effects, NLO-matched MC computations at parton could be used as well. Here, the clearer interpretation of the extracted top mass is retained since the MC descriptions at parton level are nearly independent from non-perturbative effects.

weak top mass sensitivity: relative uncertainties of the order of 5% on the measured $\sigma_{t\bar{t}}$ translates into uncertainties of 1% on the extracted m_t^{pole} [92] (*i.e.* > 1 GeV). Improvements may come from different fronts: increasing the formal accuracy of the prediction, switching to a proper mass scheme in this regime as the $\overline{MS}$ mass (which offers better perturbative convergence and hence smaller missing higher-order uncertainties) and using differential measurements (first proposed in Ref. [118]). The ATLAS collaboration used for the first time a normalised differential cross-section as a function of the $t\bar{t}$+1-jet system invariant mass, where the top mass sensitivity is further enhanced by the emission of a hard-gluon, to analyze data at $\sqrt{s}$ = 7 and 8 TeV. The top pole mass resulted $m_t^{\text{pole}} = 171.1^{+1.2}_{-1.0}$ GeV [119], the most precise measurement at that time. The CMS collaboration improved this result by extracting simultaneously the α_S and m_t^{pole} parameters from triple-differential cross-sections[4] [120]. The pole mass was found to be $m_t^{\text{pole}} = 170.9 \pm 0.8$ GeV, which poses some tension with the world averages of m_t^{MC} and m_t^{pole}. This remarkable improvement partly comes from the inclusion of the $t\bar{t}$ invariant mass, which exhibits a high kinematic mass dependence in the threshold region ($m_{t\bar{t}} \approx 2m_t$). The increased sensitivity also leads to new subtleties in the interpretation.

The top quark mass extracted in indirect measurements has a small dependence on the assumed top MC mass. In order to bring the measured data to the level where fixed-order and resummed calculations are available, a correction procedure must be followed. Measured data is corrected by phase-space acceptances and detector efficiencies resulting from the event selection and detection process, becoming observable- and process-dependent. Such corrections are derived from MC simulations, where a value for top MC mass must be taken.

[4] Double- and triple-differential cross-sections where studied, but the best result came from the latter as a function of the number of jets in the final state N_{jet}, $t\bar{t}$ invariant mass $m_{t\bar{t}}$ and $t\bar{t}$ rapidity $y_{t\bar{t}}$.

The dependence may be quite mildly though, and special efforts are devoted to reduce its impact on the measured top mass.

2.4 Top quark mass interpretation

The interpretation of the top MC quark mass is related to the identification of m_t^{MC} with a well-defined renormalized top quark mass. The m_t^{MC} has often been identified with the pole mass, because the MC generators treat the top quark as a resonant particle, $p_{\mathrm{T}}^2 = (m_t^{\mathrm{MC}})^2$, which coincides with the physical notion attached to the top quark in the on-shell scheme. We have seen that indeed the current direct and indirect top quark mass measurements yield compatible results, although the uncertainty quoted in the indirect measurements is significantly larger. However, to what precision can the top MC mass be identified with the Lagrangian mass parameter m_t in a theoretically well-defined renormalization scheme? Naturally, it heavily relies on the quality of the perturbative and non-perturbative treatments given in the MC generators. A concrete proposal to understand this relation quantitatively was put forward by the authors of Ref. [121]. The top MC mass is related to a renormalized mass in a process called MC mass calibration, where a combined measurement of the pole and MC top mass using the inclusive cross-section was performed. This relation was found to be $m_t^{\mathrm{pole}} - m_t^{\mathrm{MC}} < 2$ GeV. The following discussion, mainly based on Refs. [50, 122], is devoted to make this relation more precise.

The most general relation between the top MC mass and a renormalized top mass can be expressed as follows:

$$m_t^{\mathrm{MC}} = m_t^{\mathrm{norm}} + \Delta_{\mathrm{MC}}^{\mathrm{pert}} + \Delta_{\mathrm{MC}}^{\text{non-pert}} + \Delta^{\mathrm{MC}} \ . \tag{2.12}$$

The three Δ terms are in principle MC dependent and may have different sign. $\Delta_{\rm MC}^{\rm pert}$ embodies perturbative corrections related to the PS and the top decay, $\Delta_{\rm MC}^{\text{non-pert}}$ stands for non-perturbative effects originating from hadronization which may alter the evolution of the MC mass and $\Delta^{\rm MC}$ encodes systematic uncertainties inherent to the MC setup that can affect indirectly (*i.e.* not related to the top quark dynamics themselves) the meaning of $m_t^{\rm MC}$, such as MPI and the CR model. In an ideal scenario where the MC generators provided QCD-based hadronization models and consistent parton-level calculations (reaching NLL precision), the only term that would really matter is $\Delta_{\rm MC}^{\rm pert}$. Only in this case, $m_t^{\rm MC}$ would be observable-independent.

The $\Delta_{\rm MC}^{\rm pert}$ term can be characterized rigorously. As described in Section 1.6.4, the state-of-the-art MC generators include calculations of the final-state partons at NLO within perturbation theory. To a good approximation, the hard-scattering ME only affect the overall top production, and does not alter the meaning of the top MC mass. At this level, $m_t^{\rm MC}$ is the mass in the top propagator before decaying. The parton shower starts from here, and computes the top parton decay and successive splittings into partonic states with lower virtualities. The top decay and soft-collinear radiation is described at LO and LL, respectively. Therefore, top self-energy corrections which arise beyond LO are not explicitly accounted for, so they are deemed to be absorbed by $m_t^{\rm MC}$. The PS terminates at the shower cutoff $Q_0 \sim 1$ GeV. As a consequence, the $m_t^{\rm MC}$ does not absorb any correction below Q_0. That is why the pole mass is not strictly speaking the MC mass: $m_t^{\rm pole}$ includes self-interaction corrections from all scales. Furthermore, $m_t^{\rm MC}$ does not receive non-perturbative contributions characteristic of this region ($< Q_0$).

As far as perturbative self-energy corrections are concerned, a short-distance mass such as the MSR mass that only incorporates self-energy corrections above

$R = Q_0$ is conceptually closer to the $m_t^{\rm MC}$ as defined here. In fact, a dedicated study of the *Coherent Branching* formalism (CB), *i.e.* the theoretical basis of the angular-ordered parton showers (implemented in HERWIG, for instance), demonstrated from first-principles that [123]:

- A NLL resummation in the PS is enough to control the top mass scheme with NLO precision.

- For vanishing shower cutoffs ($Q_0 = 0$) the top MC mass parameter at NLL agrees with the top pole mass at NLO. This no longer applies when $Q_0 \gtrsim \Lambda_{\rm QCD}$, required to avoid infinite multiplicities in the PS, among others.

- The effect of the Q_0 in the pole propagator is shifting away the top MC mass from the top pole mass by a term linear in Q_0. Such shifted mass, called CB mass $m_t^{\rm CB}$, is a short-distance mass which relation with other renormalized masses can be therefore calculated perturbately:

$$\begin{aligned} m_t^{\rm MSR}(Q_0) - m_t^{\rm CB}(Q_0) &= 120 \pm 70 \text{ MeV} , \\ m_t^{\rm pole} - m_t^{\rm CB}(Q_0) &= 480 \pm 260 \text{ MeV} . \end{aligned} \tag{2.13}$$

The origin of the non-perturbative term $\Delta_{\rm MC}^{\rm non\text{-}pert}$ of Eq. 2.12 can be understood as follows: after reaching Q_0, partons resulting from the PS are hadronized according to some phenomenological-inspired model, which is tuned to describe data the best but cannot be related rigorously to first-principles QCD. Hence, the evolution of the top MC mass is definitely out of (formal) control. This fact represents the main reason why the $m_t^{\rm MC}$ cannot be systematically connected to a renormalized mass. In addition, as explained in the previous section, the exact details on the position and shape of any observable with kinematic mass sensitivity are governed by physical processes occurring at energy scales below Q_0. So the final

form of such observable, eventually used to infer $m_t^{\rm MC}$, is not shaped by any QCD calculation.

Whilst the size of the $\Delta_{\rm MC}^{\rm pert}$ term for the top pole (MSR) mass has been estimated to be of the order of 500 (100) MeV according to Eq. 2.13, the individual size of the other two terms remains unknown. The total sum of the three Δ terms was reported in Ref. [124]. Here, a precise calibration of the top MC mass in terms of the pole and MSR($R = 1$ GeV) mass was determined by fitting NNLL + NLO predictions for 2–jettiness distributions to pseudo-data generated with Pythia 8.2 in $e^+e^- \to t\bar{t}$ processes at particle level:

$$\begin{aligned} m_t^{\rm MC} &= m_t^{\rm pole} + (0.57 \pm 0.29)\ \text{GeV}\ , \\ m_t^{\rm MC} &= m_t^{\rm MSR}(R = 1\ \text{GeV}) + (0.18 \pm 0.23)\ \text{GeV}\ . \end{aligned} \tag{2.14}$$

The sum of the three Δ terms for the top pole and MSR mass is found to be slightly greater than the perturbative term only. The disentanglement of the interplay between these three terms is crucial in order to find out the potential observable-dependence hidden in the $\Delta_{\rm MC}^{\text{non-pert}}$ and $\Delta^{\rm MC}$ terms. These results support the statement that the top MC mass cannot be identified with the top pole mass with an uncertainty, in the best cases, below 0.5 GeV. The $m_t^{\rm MC}$ is rather closely related to the top MSR mass at the appropriate scale. In light of this, the interpretation problem is confirmed to be at the level of 0.5 – 1 GeV, of the size of the current direct measurements precision. Such results benefit from the high precision predictions available in e^+e^- and its clean environment. At hadron colliders, the presence of ISR, UE and CR effects may complicate things significantly. Since the way the MC generators handle such effects may further shape the evolution and meaning of the top MC mass, it is convenient to study the top MC and pole mass relation in the context where the current direct measurements are performed.

Chapter 3

The ATLAS experiment at the Large Hadron Collider

Microscopes allow to zoom in in certain region of the space and study the properties of nature at the characteristic distance scale. In particle physics, the microscopes are particle accelerators and colliders and the distance scale probed depends on the center-of-mass-energy $\sqrt{s}$ of the incident particles. In this sense, the largest microscope ever built is the *Large Hadron Collider* (LHC), where the highest energies are reached and the smallest distance scales are explored. The LHC accelerates and collides protons at $\sqrt{s}$ = 13 TeV, so the experiments installed at the interaction points study the interactions that happen at distance scales as small as 10^{-20} m. This chapter provides a description of how protons are accelerated and collided in the LHC and how *A Toroidal LHC ApparatuS* (ATLAS) is able to capture the outcomes of the pp collisions to trace-back the initiating process.

3.1 The Large Hadron Collider

The LHC is located at the *Conseil Europeen pour la Reserche Nucleaire* (CERN), which sits astride the Franco-Swiss border near Geneva. Founded in 1954, the CERN laboratory is the result of the cooperation between 23 member states and the joint work of 17500 people from around the world. The CERN's main goal is to provide the accelerator complex and other infrastructures (as a computing resources to analyze, distribute and store real and simulated data) required to perform high-energy physics experiments. The LHC is assembled in a 27 km tunnel where the *Large Electron Positron* collider (LEP) was hosted, 100 meters underground. It is the last and largest accelerator of a chain of accelerators that increases the energy of protons in successive steps, as shown in Figure 3.1.

It first started up on September 2008 with some inaugural tests. The first LHC operational run period (Run 1) spanned from 2009 to 2013 and reached a center-of-mass energy of 8 TeV. After a two-year break for upgrades, known as *Long Shutdown 1* (LS1), it restarted the operations on April 2015 with a nominal energy of 13 TeV. This second LHC operation run period (Run 2) extended up to 2018, and was followed by a four-year break called the *Long Shutdown 2* (LS2). The LHC became operational again on April 2022 with a maximum energy of 13.6 TeV (Run 3) and it is expected to last until 2026. The LHC's successor, the *High-Luminosity LHC* (HL-LHC) whose objective is to increase the total number of interaccions produced by a factor of 10 beyond the LHC's design value, is schedule to come into operation at the beginning of 2029.

The journey of protons through the CERN accelerator complex starts at the hydrogen container. After being stripped of the electrons, protons are injected in a linear accelerator, called *Linac*, which increases the energy up to 50 MeV. This is followed by the *Proton Synchroton Booster* (PSB), a small circular accelerator

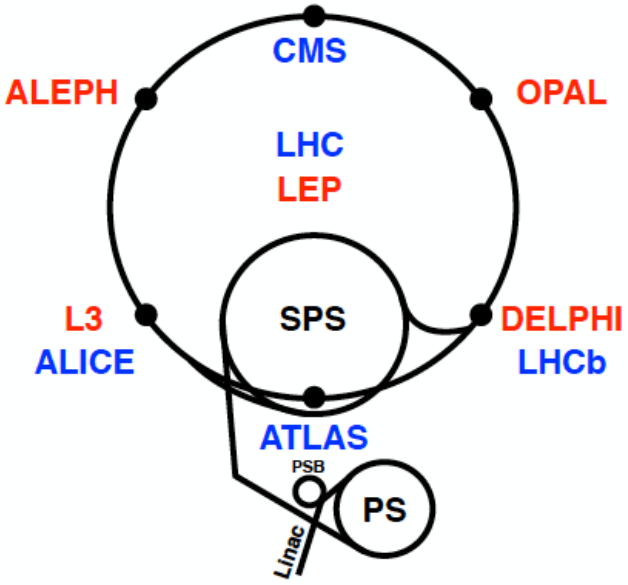


Figure 3.1: Sketch of the CERN accelerator complex with the experimental detectors placed at each interaction point in the LHC (blue) and LEP (red) era.

(25 m in radius) that speeds the protons up to energies of 1.4 GeV. A larger synchrotron, the *Proton Synchroton* (PS) with a radius of 100 m, increases the energy to 25 GeV. The PS accelerator has played an important role on its own right; many experiments have profited from the beams provided by the PS, enabling great discoveries such as the weak neutral currents in 1970 [125]. After this, protons are passed to the *Super Proton Synchroton* (SPS) accelerator, reaching energies of 450 GeV. Analogously, the SPS facilitated the discovery of the W and Z bosons by the UA1 [126, 127] and UA2 collaborations [128, 129] [1]. In the last stage, protons are injected into the LHC where the energy is ramped up to 4 TeV in Run 1 and 6.5 TeV in Run 2.

During this process protons are accelerated by means of *Radio Frequency* (RF) cavities, where a longitudinal oscillating electric field is carefully adjusted in such a way that some protons always see the field in the accelerating direction.

[1] In the 1980s, the SPS was operated as a proton-antiproton collider (the $Sp\bar{p}S$) and created the W and Z bosons that were discovered by UA1 and UA2.

Those protons that are not synchronized with the RF cavity (because arrive slightly before or after) oscillate longitudinally around the synchronized ones. This implies that protons are not uniformly spread along the beam but grouped into approximately 2.800 bunches, containing around $N_b = 10^{11}$ protons each. For protons traveling at nearly the speed of light, the bunches are separated about 25 ns so the bunch crossing frequency is of $1/25$ ns^{-1} = 40 MHz.

Protons are steered around the LHC thanks to 1238 superconducting dipole magnets. The proton's energy depends on its mass, the magnet strength B and the accelerator radius R through:

$$E[\text{GeV}] = \sqrt{4\pi\alpha} \times \left(\frac{B}{T}\right)\left(\frac{R}{m_p}\right) \approx 800 \times \left(\frac{B}{T}\right) \ . \tag{3.1}$$

This means that for protons with mass $m_p \sim 1$ GeV traveling at $E = 6.5$ TeV in an accelerator of radius of $R \approx 4$ km, the dipole magnets need to be powered to ~ 8 T (six orders of magnitude bigger that the intensity of the geomagnetic field). Furthermore, 392 quadrupole magnets with similar intensity are employed to keep the beams focused and collimated. 96 tones of super-fluid helium-4 are required to cool the magnets down at their operating temperature of 1.9 K. The cooling of the magnets is one of the main limitations in the number of protons per bunch; accelerating charged particles radiate and the emitted power per particle is proportional to $P \propto (B/T)^2(\text{m/GeV})^{-4}$. Although it is highly suppressed for protons, it results in a significant amount of synchrotron radiation when considering bunches with N_b protons.

Bunches of protons are carried in two separate beams traveling in opposite directions. These two beams eventually cross in 4 points along the accelerator ring, hosting each one an experimental detector: ATLAS [130], CMS (*Compact Muon Solenoid*) [131], ALICE (*A Large Ion Collider Experiment*) [132] and LHCb

(*Large Hadron Collider beauty*) [133]. In order to increase the probability of colliding two protons, quadrupole magnets are used to squeeze the beams down. In the laboratory frame, the beam/bunch size becomes of the order of $\sigma \sim 10^{-5}$ m, which is significantly larger than the size of the proton, $\sigma_p \sim 10^{-15}$ m. The average number of interactions at each bunch crossing is given by $\mu \sim N_b^2 \sigma_p^2 / \sigma^2 \sim 100$. The actual number of collisions per second is referred to as the instantaneous luminosity $\mathcal{L}$, given in full by:

$$\mathcal{L} = \frac{1}{4} \frac{N_b^{(1)} N_b^{(2)}}{\sigma_x \sigma_y} f \ , \tag{3.2}$$

where $N_b^{(1,2)}$ stands for the number of protons in the two colliding bunches, $f \sim 40 \cdot 10^6$ Hz is the frequency of the bunch crossing and σ_x and σ_y correspond to the beam widths in the transverse plane, assuming that the beam cross-section can be described in terms of a bidimensional Gaussian function. The instantaneous luminosity has increased progressively during the LHC operation and so the mean number of interactions per bunch crossing, as illustrated in the left-hand plot of Figure 3.2 for the Run 2. This leads to some experimental complications, such as damages in the closest systems of the detector to the interaction point due to the synchrotron radiation. The *pile-up* (PU), that is the simultaneous detection of particles coming from different collisions, may complicate the isolation and the correct reconstruction of each collision/event. There are two different sources of PU, depending on whether such particle production involves protons in the same (in-time) or different (out-of-time) bunch crossing. The number of these collisions is stochastic and modeled in the MC simulations to match the observed in data.

In order to get the total amount of collisions delivered in certain period of time, the integrated luminosity L is used instead, which simply reads as $L = \int_{t1}^{t2} dt \mathcal{L}$.

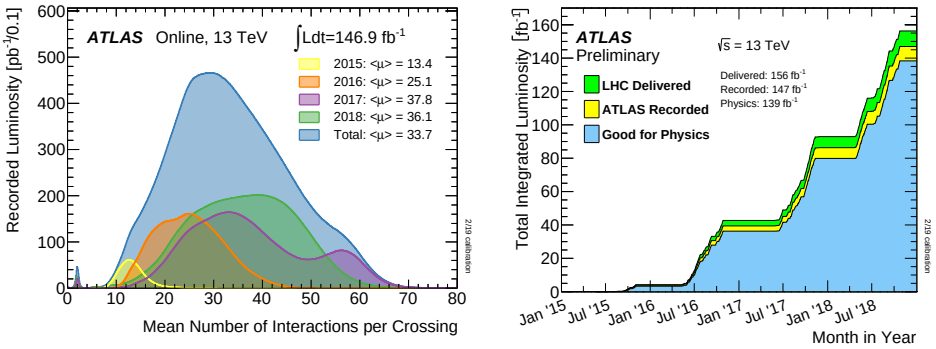


Figure 3.2: On the left: mean number of interactions per bunch crossing during Run 2 for several periods with different characteristic instantaneous luminosity. On the right: cumulative luminosity versus time delivered to ATLAS (green), recorded by ATLAS (yellow), and certified to be good quality data (blue) during stable beams for pp collisions in Run 2. Images from Ref. [134].

As can be noted from the right-hand panel of Figure 3.2, most of the delivered luminosity is recorded by the ATLAS detector. In turn, a major portion of the recorded luminosity is exploited in physics analysis. The number of events observed for given process $pp \to X$ can be written in terms of the integrated luminosity and the corresponding cross-section as:

$$N = \epsilon \cdot \sigma(pp \to X) \cdot L \ , \tag{3.3}$$

where $\epsilon < 1$ encodes a factor correction due to the detector acceptance and constraints on the phase space where the measurement is performed. In order to study rare processes with small cross-sections or populate with enough events extreme regions of the phase space, the integrated luminosity has to be maximized.

3.1.1 Collider coordinates and kinematic variables

Detectors placed at particle colliders typically use a special kind of cylindrical coordinates. It is important to see how the different kinematic variables are

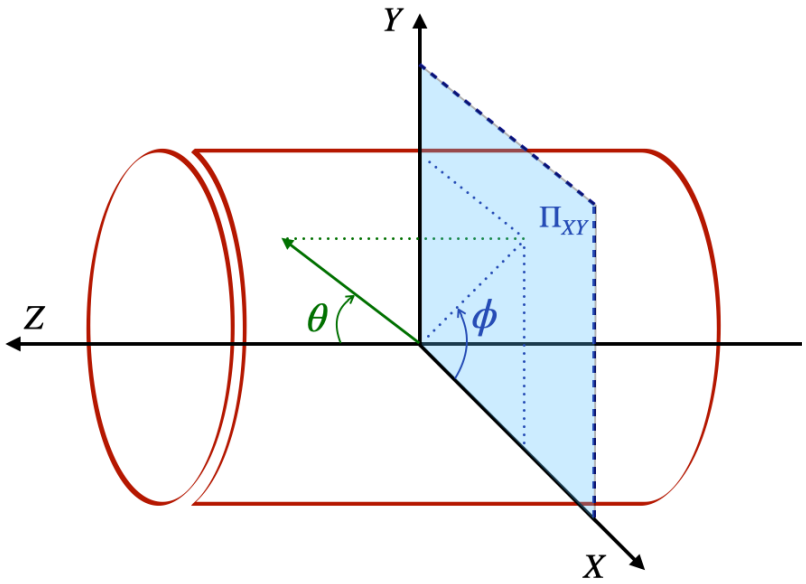


Figure 3.3: Cylindrical coordinates system typically used to describe particles position in detectors at circular colliders.

defined in this context and the physical reason behind this choice.

The schematic representation of a symmetric detector (such as ATLAS and CMS) is shown in Figure 3.3. The origin of the coordinate system coincides with the interaction point where the two beams cross. The $z-$axis is along the beam (longitudinal) direction, whilst the positive $x-$axis points to the center of the accelerator ring. The positive $y-$axis points upwards. The transverse plane to the $z-$axis is given by Π_{xy}. In this plane, the total (transverse) energy must be null since the momentum of the incoming particles only has longitudinal component. The polar angle θ and the azimuthal angle ϕ are measured from and around the $z-$axis, respectively.

The particle's position with momentum p within the detector can be expressed with two interchangeably sets of coordinates: (p_x, p_y, p_z) and $(\eta, \phi, p_{\mathrm{T}})$. The former corresponds to the cartesian coordinate system. The latter, which relates cylindrical coordinates, make use of the pseudorapidity variable η along with the azimuthal angle and the component of the momentum in the transverse plane

$p_T = p_x^2 + p_y^2$. In particular, the pseudorapidity can be calculated as:

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right] = \frac{1}{2}\ln\left(\frac{|\vec{p}| + p_z}{|\vec{p}| - p_z}\right) . \tag{3.4}$$

Particles with $\eta = \pm\infty$ travel along the beam, and those with $\eta = 0$ are contained in the transverse plane. In the relativistic limit of $p \gg m$, the pseudorapidity and rapidity (Eq. 2.6) describe the same quantity, $\eta = y$. These coordinates are of special interest because the difference in pseudorapidity $\Delta\eta$ between two massless particles (or massive particles in the relativistic limit) is invariant under Lorentz boosts along the z-axis. The distance between two particles can be expressed in terms of the distance metric variable:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} , \tag{3.5}$$

which is invariant under longitudinal boosts. At hadron colliders where the longitudinal component of the incoming particle momentum p_z is unknown, the invariance of this quantity becomes crucial.

3.2 The ATLAS detector

The ATLAS detector is a general-purpose machine of 7000 tonnes that measures 44 m in length and 25 m in diameter. It is the biggest detector at CERN and its construction took almost 10 years. It is hosted in a cavern 100 m below ground as the LHC ring, which provides a natural shield to cosmic rays.

It was designed to measure the momentum and properties of hundreds of particles formed in the pp collisions that occur at each bunch crossing. There are two ways to observe these particles. In the one hand, trackers make use of external magnetic fields to bend the trajectory of charged particles. The altered

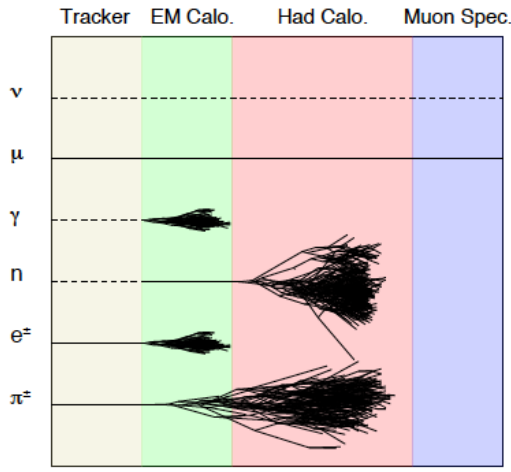


Figure 3.4: Illustration of the signals left by different particles as they pass through several detectors. Dashed lines represent particles that leave no trace in the sub-detector, whilst horizontal solid lines stand for particles that only loose a small fraction of their energy. Figure from Ref. [135].

trajectory, which provides information of the particle's energy and charge, can be reconstructed with high precision from measurements of several positions. The solid state and gaseous detectors must have excellent position sensitivity and present as little material as possible to avoid affecting the trajectory. On the other hand, the calorimeters are designed to stop most of the particles coming from the collision, forcing them to deposite all of their energy within the detector. The energy lost in a calorimeter occurs via a cascade of collisions where particles can interact electromagnetically or strongly, leaving a characteristic shower pattern within the detector. Modern sampling calorimeters essentially count the number of charged particles that forms in the shower. There are particles that are not stopped by the calorimeters, such as muons. They can be identified by placing an additional tracking detectors beyond the calorimeters. In Figure 3.4, a schematic representation of the detector signals left by different particles is presented. As we can see, the only (known) particle that scapes completely the system detection formed by trackers and calorimeters is the neutrino ν.

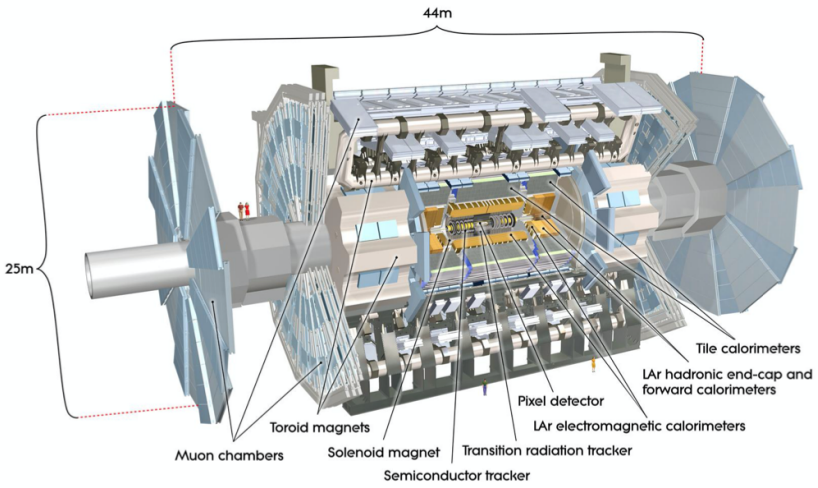


Figure 3.5: The ATLAS detector and its sub-detectors. Image from Ref. [136].

The ATLAS detector consists of several sub-detectors with cylindrical symmetry. Each sub-detector comprises a cylindrical barrel and two end-caps. All sub-detectors are centered around the interaction point and placed according to the detection logic explained above. A diagram of the ATLAS detector and its sub-systems is displayed in Figure 3.5.

3.2.1 Inner detector

The *Inner Detector* (ID) is the innermost ATLAS layer and consists of a series of tracking detectors with different technologies. It is immersed in a solenoidal magnetic field of 2 T parallel to the beam line which bends the trajectory of the charged particles. The structure of the ID is illustrated in Figure 3.6.

The *PIXel* detector (PIX) was originally the closest system to the beam pipe, occupying the region 5 – 15 cm of radius away from the interaction point. 92 million semiconductor silicon pixels are grouped in modules, a rectangular active device approximately 6 cm by 2 cm with 46080 pixels. 1736 modules are arranged in three cylindrical layers (B-layer, layer-1 and layer-2) and 288 are distributed in three disks placed at each end-cap. The original cylindrical layers are composed

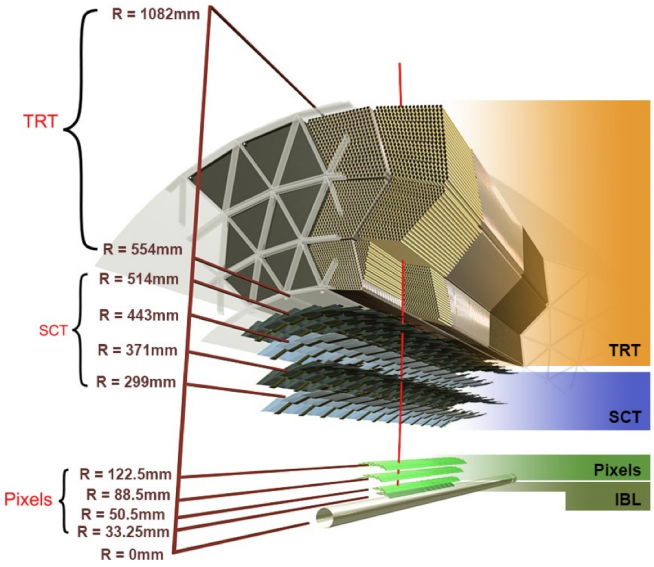


Figure 3.6: The ATLAS inner detector. Image from Ref. [137].

of 250 μm thick planar sensors with a 50×400 μm^2 surface area. Due to the high luminosity and radiation conditions envisaged for the Run 2, in 2014 an *Insertable B-Layer* (IBL) was placed between the B-layer and the beam pipe at only 33.2 mm of the interaction point, made of smaller and thinner planar sensors in the barrel and 3D sensors at high $|\eta|$. This allowed to protect the PIX B-layer and minimize the loss in tracking efficiency. With this design, the pixel detector can provide a precise measurement of the decay vertex position.

Beyond the PIX system, the *Semi Conductor Tracker* (SCT) is deployed at a distance of $30 - 50$ cm. It is composed of 4 concentric double layers of silicon microstrip in the barrel and 18 planar end-caps disks, with 6 million read-out channels stripping every 80 μm on the silicon. This enables the determination of the charged particles position with an accuracy of 17 μm per layer in the direction transverse to the strips, and about 600 μm in the $z-$axis.

Surrounding these two silicon-based detectors is the *Transition Radiation Tracker* (TRT), between 50−100 cm. This region is filled with $350{\cdot}10^3$ straw tubes of radius 2 mm. Each of those tubes, filled with a gas mixture of $Xe/CO_2/O_2$,

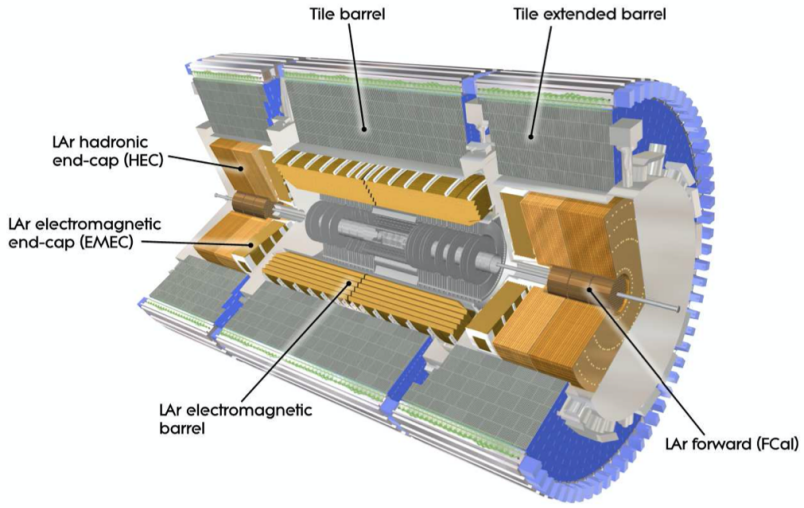


Figure 3.7: The ATLAS calorimeter system. Image from Ref. [138].

has a diameter of 0.03 mm and a gold-plated tungsten wire in the center. With this layout, charged particles leave ionization energy in the tubes, collected by the inner wire. The TRT helps to identify the incident particle, *i.e.* if it is an electron or pion.

3.2.2 Calorimeter system

The ID is surrounded by the ATLAS electromagnetic and hadronic calorimeters. While the two sampling calorimeters have different technologies, they share the same structure: a passive absorber material that stops the hitting particle interleaved with layers of an active medium that measures the energy. The different components of the ATLAS calorimeter system is displayed in Figure 3.7.

The *Electromagnetic CALorimeter* (ECAL) is specially designed to measure cascades of particles initiated by electrons and photons, absorbed by layers of metal such as tungsten, copper or lead. The resulting cascade of particles, which occurs via bremsstrahlung and e^+e^- creation, ionise a *Liquid-Argon* medium (LAr) producing an electromagnetic current registered by copper electrodes. The central

ECAL is of an accordion-geometry design and consists of two half-barrels of 6.5 m long and 53 cm thick with 55000 channels each, providing together full ϕ coverage and partial $|\eta|$ coverage up to 1.4. The EM end-caps, which have thickness 60 cm and radius 2 m, span the range $1.4 < |\eta| < 3.2$. There is a transition gap between the central barrels and the end-caps ($1.37 < |\eta| < 1.52$) which contains a relatively large amount of inactive material. Over the region devoted to precision measurements ($|\eta| < 2.47$, excluding the transition regions), the EM calorimeter features three layers with variable longitudinal segmentation. The first layer consists of strips finely segmented in η, offering an excellent discrimination between photons and $\pi^0 \to \gamma\gamma$ decays. The second layer collects most of the energy and the third layer provides measurements of energy deposited in the tails of the shower, used to correct for leakages.

The *Hadronic CALorimeter* (HCAL) surrounds the ECAL and measures the energy of hadronic particles, such as pions and kaons, that do not deposit all of their energy in the ECAL. The absorbent material of the HCAL barrel, which provides a coverage of $|\eta| < 1.7$, is formed by 14 mm layers of steel, and the sampling layer consists of plastic scintillating plates 3 cm thick, called tiles, which emit light as the particles pass through. The HCAL end-caps use layers of copper and LAr, so they can measure hadronic and electromagnetic showers in the range $1.5 < |\eta| < 3.2$. This coverage is further extended by means of the *Forward Calorimeter* (FCal), made of tungsten/copper and LAr. The FCal allows to measure hadron particles emitted very close to the beam pipe, about $3.2 < |\eta| < 4.9$.

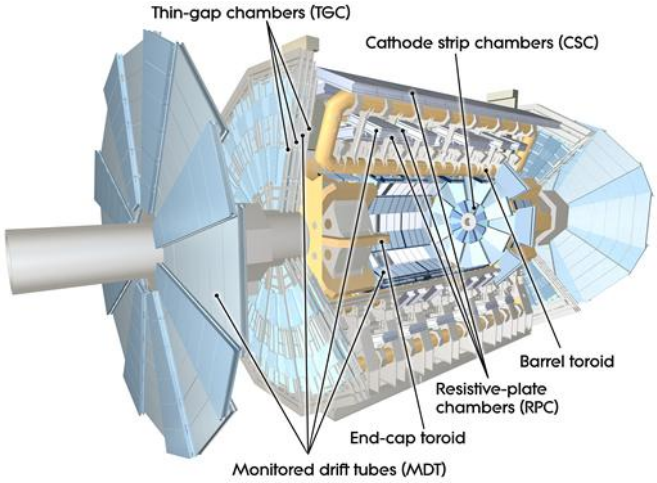


Figure 3.8: Components of the muon spectrometer in the ATLAS experiment. Image from Ref. [139].

3.2.3 Muon spectrometer

Muons and neutrinos are the only expected particles that are not stopped by the calorimeters. Muons can be measured with a detection system placed beyond the calorimeters, but neutrinos cannot because they barely interact. The *Muon Spectrometer* (MS) extends from ~ 4.24 m to ~ 11 m, providing a detecting arm that enables very precise measurements. Analogously to the ID, the muon spectrometer is immersed in a magnetic field that bends the particle's trajectory; three systems of toroidal magnets provide a field strength of 4 T in the barrel and end-caps regions. Four different technologies are used for triggering events and detect muons, as shown in Figure 3.8.

The *Resistive Plate Chambers* (RPC) and *Thin Gap Chambers* (TGC) are dedicated to triggering with roughly $1 - 4$ ns time resolution, and covers the region given by $|\eta| < 1$ and $1 < |\eta| < 2.7$, respectively. The former, organized in three layers with 596 chambers and $380 \cdot 10^3$ channels, consists of parallel plate capacitors filled with gas and separated radially, which allows for a raw but fast momentum measurement. The latter, which counts with 192 chambers and $440 \cdot 10^3$ channels in three layers too, have a finer granularity in order to cope

with the high-multiplicity of the forward region.

On the other hand, the *Monitored Drift Tubes* (MDT) and *Cathode Strip Chambers* (CSC) are devoted to provide high-precision tracking information by measuring the muon's momentum. The MDT covers the central region of the detector and have a 35 μm resolution along the $z-$axis, whilst the CSC covers the forward region of $2.4 < |\eta| < 2.7$ and can measure both the radial and z position with 5 μm and 40 μm resolutions, respectively. It is very important for resolving tracking ambiguities in the forward region due to the higher particle flux.

3.3 The ATLAS data management

3.3.1 Trigger system

As we saw previously, the nominal pp interaction rate is of 40 MHz (1/25 ns) at the LHC. Other that $pp \to pp$ collisions, the total inelastic cross-section is about 78 mb [140]. On the one hand, the cross-section of interesting hard-scatter events, such as $pp \to t\bar{t}$, is of the order of 900 pb at 13 TeV (see Figure 2.5). This means that only one bunch crossing in 10^7 gives potentially interesting events. On the other hand, if each event uses $\mathcal{O}(1)$ MB of disk space, the readout of all events produced would yield 100 TB/s. It is physically undesirable and unfeasible to store all detected events at this rate.

These problems are circumvented by keeping only potentially interesting events. This is the task of the ATLAS trigger system, which is organized in three levels called *Level-1* (L1), *Level-2* (L2) and *Event Filter* (EF). The L1 trigger is hardware based, and uses a sub-set of information given by the calorimeters and the muon detector. By crude approximations to object reconstructions, it decides if an event is either recorded or rejected in only 2.5 μs, reducing the total rate

from 40 MHz to 70 kHz (Run 1) or 100 kHz (Run 2). The L2 trigger is software based instead, so it needs a few miliseconds to decide the fate of the passed event. The L2 uses fine-grained data from all sub-detectors and offline-like algorithms to analyze certain regions of interest already preselected by the L1. In this way, the data acquisition rate is further reduced to 1000 Hz in the Run 1. Finally, the EF trigger is also software based and performs a detailed analysis of the whole event in a few seconds. At the end, only 200 events per second were recorded for offline analysis during Run 1. Every data stream (jets, muons...) has a different trigger chain, and an event can fire multiple triggers. Only one is required to record the event.

During the LS1, the L1 trigger was upgraded to better control the more complex conditions in terms of pile-up and fake-muon rate (non-muon particles wrongly identified as muons). Furthermore, the L2 and EF triggers were merged into a *High Level Trigger* (HLT), in order to optimize the resource sharing and simplify the hardware and software setup. The physics output rate of the HLT during an ATLAS data-taking run is on average of 1200 Hz.

3.3.2 Data storage

The large amount of events recorded in the ATLAS experiment must be stored and distributed for further processing and analysis. It is not possible to handle data streams of $\mathcal{O}$(PB/hour) in a single computing site, so the *Worldwide LHC Computing Grid* project (WLCG) [141] was created to share such task between many more computing centers around the world. It is supported by other grid projects, such as the *LHC Computing Grid* (LCG), and consists of a global collaboration of around 170 computing centers in more than 40 countries, resulting in a complex and distributed computing infrastructure with almost 1

million of computing cores and an exabyte of storage [142].

Originally this computing infrastructure was structured according to a tiered hierarchy, where different classes of computing centers called *Tiers* play certain role according to the specific resources provided. The Tier-0 center is the largest one and is hosted at the CERN Data Center. It stores raw data with a very basic reconstruction as it is collected by the LHC detectors. A redundant copy of such raw data is distributed in the 13 existing Tier-1 centers, spread worldwide. With a direct connection to the Tier-0 center, Tier-1s are also devoted to manage permanently stored data (simulated and processed data) and provide computational power for data reprocessing and analysis. Each Tier-1 facility had associated in turn a set of Tier-2 centers, 155 in total, which were close-by in terms of network connectivity and formed the so-called cloud. Tier-2s provide storage and computational resources for MC event simulation and end-user analysis. Those Tier-2 centers with significant storage and excellent network connections are designated as *Nucleus*, passing the job production to smaller Tier-2s [143]. Finally, end-users can access the data at the local computing facilities Tier-3, which are outside the LCG project. All Tier-N centers that are members of the WLCG are responsible for funding the resources required to fulfill their declared role.

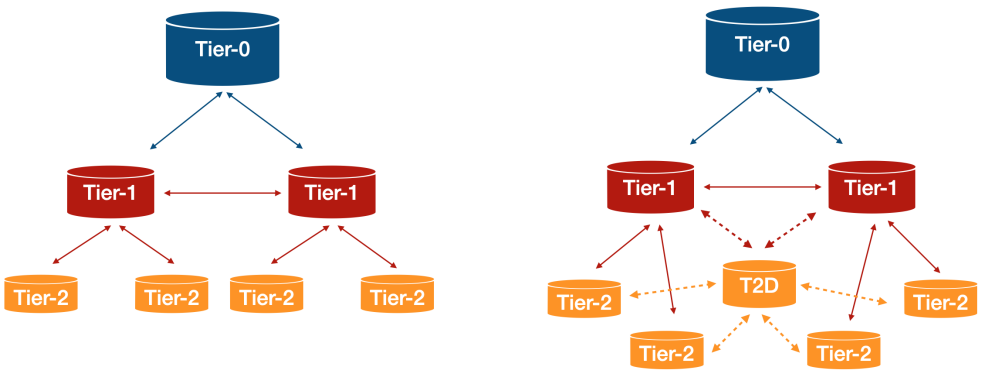


Figure 3.9: Evolution of the ATLAS Computing Model from a hierarchical (left) to a mess (right) mode.

The increased collision rate, detector readout and computing demand for MC simulations envisaged in the Run 2 challenged the original computing model. A more dynamic workload sharing was needed to satisfy these requirements, motivating the evolution from the hierarchical structure to a mess model, shown in Figure 3.9. In the new structure, Tier-2s with enough network connection can link directly with other Tier-2 centers, called *Tier-2 Directly* (T2D). T2Ds can also transfer information with more than one Tier-1 center. In general, a more efficient usage of the disk storage and CPU resources is achieved.

The *ATLAS Computing Model* (ACM) [144] specifies the implementation and the performance of the software resources needed for the data management. It embraces the WLCG paradigm allowing a high level of decentralisation and sharing of computing resources. The main requirement of a computing model is to provide all ATLAS members with swift access to all reconstructed data for analysis and appropriate access to raw data for organised monitoring, alignment and calibration activities.

3.3.3 Object reconstruction

The information stored at each event, eventually used by end-users for physics analysis, refers to the physical objects reconstructed from the detector signals deposited by the incident particles. There is a limited number of particles that can be measured and hence a limited number of reconstructed objects; electrons, muons, photons and hadrons. Let's summarize in the following lines the main aspects related to each object reconstruction:

- Electrons [145, 146] are reconstructed from showers produced in the ECAL that are associated to tracks of reconstructed charged particles in the ID. They are identified by means of a likelihood-based identification, where

probability density functions for electron-candidates and backgrounds built from simulation are used. The main electron misidentification comes from $\gamma \to e^+e^-$ processes and signals left by charged hadrons. Several categories of electrons can be used for analysis, according to the level of isolation (isolated prompt leptons or in busy environments) and the requirements during the reconstruction.

- Muons [147] can be reconstructed from information provided by the MS only, from a combination of MS and ID measurements and from associating a minimum-ionizing particle energy deposit in the calorimeter with tracks in the ID. Muon identification is performed by applying quality requirements that suppress backgrounds, mainly from pions and kaons decays, while selecting prompt muons with high efficiency and/or guaranteeing a robust momentum measurement. Several inclusive identification categories with tighter or looser requirements are available, in order to address the specific needs of the physics analysis: loose, medium and tight. The medium identification criterium provides the default selection for muons in ATLAS, developed to minimize systematic uncertainties associated to the muon reconstruction and calibration.

- Photons [148] reconstruction relies on measurements of electromagnetic showers in the ECAL. If the ECAL energy cluster has no associated track in the ID, the photon is considered unconverted. If it does have associated two oppositely-charged tracks compatible with electrons, the photon is categorized as converted and deemed to be reconstructed from an e^+e^- pair. The shape and properties of the associated electromagnetic shower allow to separate between prompt and background photons, being the latter typically accompanied by surrounding hadronic activity. Isolation criteria plays an important role here.

- Hadronic jets are mainly identified in the HCAL and the ECAL (~ 1/3 of jets are made of photons), but their reconstruction can be highly improved by considering tracking information. There is not only one way to reconstruct jets, as it depends on the jet definition adopted. Since jets play a central role in the work presented in this thesis, an extensive overview of jet definition and reconstruction is given in Chapter 4.

- Missing Energy E_T^{miss} [149] is related to the unbalanced momentum in the transverse plane of the collision, initially attached to the presence of neutrinos. Two different sources contribute to the visible transverse energy: hard objects properly reconstructed and calibrated (photons, electrons, jets, etc.) and soft signals associated to the hard-scatter vertex but not with such hard objects. The contributing objects need to be reconstructed from mutually exclusive detector signals. Misreconstructed and/or miscalibrated hard objects, the pile-up (which affects both hard and soft objects) and omitted soft signals can impoverish the reconstructed E_T^{miss}.

After reconstruction, all objects are calibrated in order to correct for detection effects that may distort the measured four-momenta. In the next chapter, the full calibration chain followed to properly characterize the detector response and resolution to large-R jets in the ATLAS detector is revisited.

Chapter 4

Jet physics and boosted objects

Jets are ubiquitous objects in any experiment placed at the LHC laboratory. Resulting from the successive splittings of partons into further partons, jets populate most (if not all) the events recorded by the ATLAS detector. Many physics analysis rely on how well the response of the detector to jets is understood. This response is not unique though, since there are many ways to define and build jets. In this chapter we review the main concept of what is a jet and how it is defined. We shall see that jets can be formed with information provided by different detector sub-systems, which ultimately determines the jet performance, and the calibration sequence that they undergo before being used in any physics application. We will focus on the jet mass calibration of big jets in the boosted regime, since they play a central role in the work presented in this thesis.

4.1 Introduction

The observable properties of jets, encoded in the radiation pattern inside and around the jet, are related to the quantum properties of the initiating partons (quarks and gluons). The most basic quantum property is the parton three-momentum. For most of the analysis using jets, this quantum property is the only relevant one. For instance, the average p_{T} for given jet is the same as the average p_{T} of the associated parton within 5 – 10% [150]. These kind of jet properties are independent from the structure of the radiation pattern within the jet, also known as jet substructure. However, the recognition and analysis of the jet substructure is essential to associate jets with the initiating partons.

In contrast, there are some other jet properties intimately related to the jet substructure, such as the jet mass. The mass of jets generated by the hadronic decays of genuinely heavy particles, such as top quarks and $W/Z/H$ bosons, is highly correlated with the corresponding parton mass. In the case of jets initiated by light quarks and gluons, the associated mass is unrelated to the on-shell quark and gluon masses but encodes information about the color charge carried by the initiating particles. The jet mass observable has been widely used in many studies involving the Higgs boson [151–155], W/Z bosons [156–163] and top quarks [164–171]. It also appears in measurements of the strong force [172–174] and in searches of physics BSM [175–180].

Jets can be studied from both real and simulated events. When considering real data, jets can be only built from the electric signals collected by the detectors in very busy environments. The precise measurement of the jet constituents energy and location is challenging due to the limited detector resolution and several sources of diffuse noise, which complicate the measurement and the later analysis and interpretation. This soft energy background is expected to spread

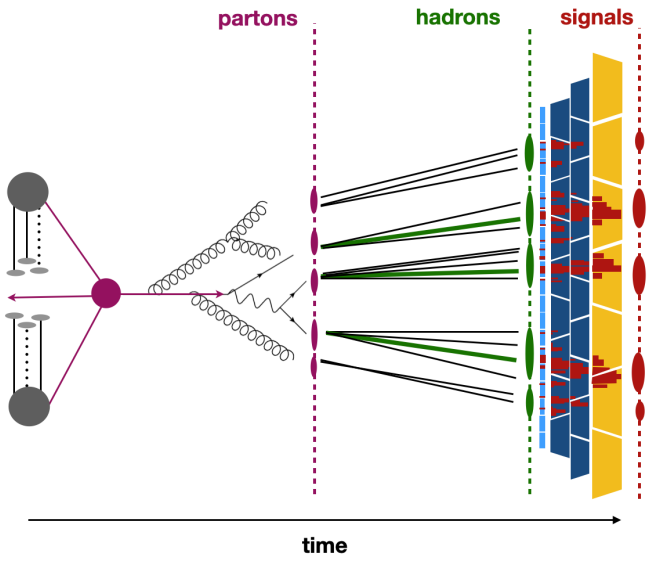


Figure 4.1: Simplified representation of the different stages at which jets can be built: partons (purple), stable hadrons (green) and electric signals (red) are the jets constituents at parton, particle and detector level, respectively.

fairly uniformly throughout the detector, so jets with larger catchment areas are proportionally more affected. Such undesirable effects can be mitigated with dedicated techniques. In simulation it is possible to build jets at several levels, as illustrated in Figure 4.1. The parton level and particle level refer to jets fed with the outcoming partons of the hard-scatter on the one hand and, on the other, with the stable hadrons resulting from hadronization. Jets at detector level can also be built from simulated detector electric signals. The advantage of simulated jets is that they can be traced back to the associated initiating parton, enabling the development of strategies to identify or tag jets according to their substructure.

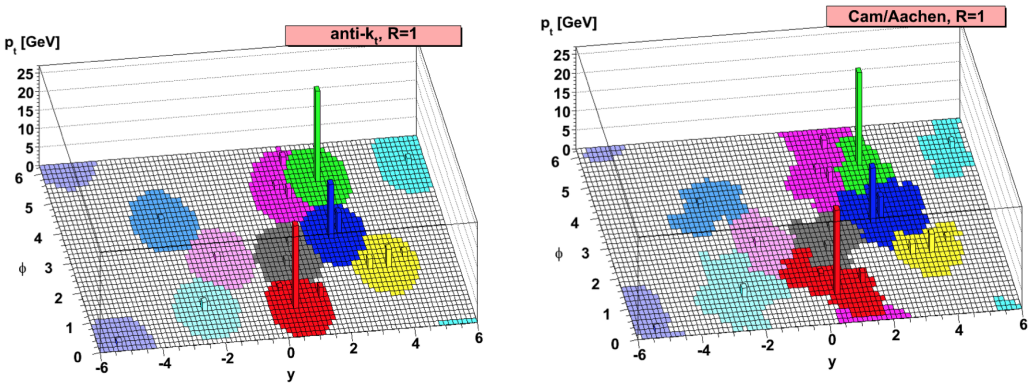
From an experimental point of view, the hadronic final states appear to be collimated in a few directions within the detector. However, the simple concept of what a jet is meant to represent is not enough to actually identify jets in given event. In order to group all entities (partons, hadrons or electric signals in the detector) available into jets, further information is required. As we shall see in the following section, this information is fully supplied by the jet definition.

4.2 Jet definition and identification

The jet definition comprises the recombination algorithm and the combination scheme. The former provides a set of rules for grouping particles into jets. The latter defines how the kinematic properties of given jet is derived from its constituents. In the following, the *E-scheme* is adopted where the four-vector of the constituents are added in the usual metric. It is important to understand that any jet definition will give rise a jet from a single hard isolated particle, which typically determines one of these directions around which the spray of particles are collimated within the detector. However, the treatment of two hard particles that are close by, the soft radiation from the jet and the pile-up heavily relies on the jet definition chosen. Therefore, there is not a single optimal way to build jets; it is up to the analyzer to find the jet definition that best fulfills the physics requirements of his/her particular study.

Concerning the jet recombination or algorithm, we will restrict ourselves to those that are suitable for jet substructure studies. To this end, sequential algorithms are preferred. This kind of jet algorithms are based on the fact that, from a perturbative QCD viewpoint, jets are the result of successive parton branchings. Therefore, they try to invert the branching history of the initiating parton by sequentially recombining two particles into one. Most of the sequential recombination algorithms used in the LHC context belong to the family of the longitudinally invariant k_t algorithms [181], where the recombination is done according to certain distance criterium. It considers two distance measures: the distance d_{ij} between two particles i and j and the distance d_{iB} between the particle i and the beam (B), defined as:

$$d_{ij} = \min(p_{\mathrm{T},i}^{2n}, p_{\mathrm{T},j}^{2n}) \times \frac{\Delta R_{i,j}^2}{R^2} , \tag{4.1}$$



(a) (b)

Figure 4.2: Jet shapes and catchment areas for two illustrative jet clustering algorithms. Image taken from Ref. [182].

$$d_{iB} = p_{\mathrm{T}}^{2n} , \tag{4.2}$$

where $\Delta R_{i,j}^2 = \sqrt{\Delta y_{i,j}^2 + \Delta \phi_{i,j}^2}$ represents the distance between the two particles, with Δy and $\Delta \phi$ the differences in rapidities and azimuthal angles, respectively. The radius or resolution parameter R represents the maximum distance between two constituents in the $y - \phi$ plane to be considered as part of the same jet. So the clustering proceed by taking the smallest value of the two distances from the initial list of available objects. If $d_{i,j}$ is smaller, i and j are recombined (using the combination scheme) into a new object k which is added to the list again. If it is d_{iB}, i becomes a jet and it is removed from the inputs to loop over. Clustering continuous until no more particles are left.

Eq. 4.2 includes also the parameter n which sets the relative power of the energy versus the geometrical scales: if $n = -1$ the clustering algorithm is known as anti-k_t [182] where hard particles are first clustered and soft particles are aggregated around it. This gives rise to jets with regular shapes and soft-radiation resilience, particularly useful to better control the jet calibration. The k_t algorithm [182, 183] corresponds to $n = 1$ and proceeds from the softest to the hardest emission. Here, the k_t distance measures are closely related to the

structure of divergences in QCD emissions: this clustering algorithm attempts an approximate inversion of QCD branching process. In contrast, the *Cambridge-Aechen* (C/A) algorithm [184, 185] corresponds to $n = 0$, leading to an energy-independent and purely angular-ordered clustering: it gives a natural handle of the structure of the jet, preserving its angular ordering. Figures 4.2(a) and 4.2(b) illustrate the jet shape for anti-k_t and C/A algorithms, respectively. As can be noticed, the former exhibits a more regular shape than the latter.

The XCone is a jet algorithm based on minimizing the event-shape N-jettiness[1] and which profits from some developments of the jet-shape N-subjettiness[2]. The N-jettiness provides a measure of the degree to which hadrons in the final state are aligned along N jet axes or the beam direction. The minimization of this quantity allows to identify N jet directions. In this way, XCone defines an exclusive jet algorithm, which means that it returns a fixed number of jets. This is specially convenient in applications where one knows the number of expected jets in advance. Similarly to the anti-k_t algorithm, XCone gives nearly conical jets for well-separated jets and it is able to resolve jets even in boosted regimes where they may partially overlap, as we will see in the next section.

4.2.1 Jets in the boosted regime

At the center-of-mass energy at which the LHC operates, heavy unstable particles may be produced with large Lorentz-boost $\beta = v/c$. In the laboratory frame, the relative angle that the decay products form with the direction of the boost is

[1] Traditionally, event-shapes are observables used to describe the geometry of the outcome of hard collisions at high-energy experiments. For an event with at least N energetic jets, the N-jettiness observable is designed to provide an inclusive measure of how N-jet-like the event looks. It allows to veto additional undesired jets in very busy environments, specially useful to clean up events and discriminate signal and background [186].

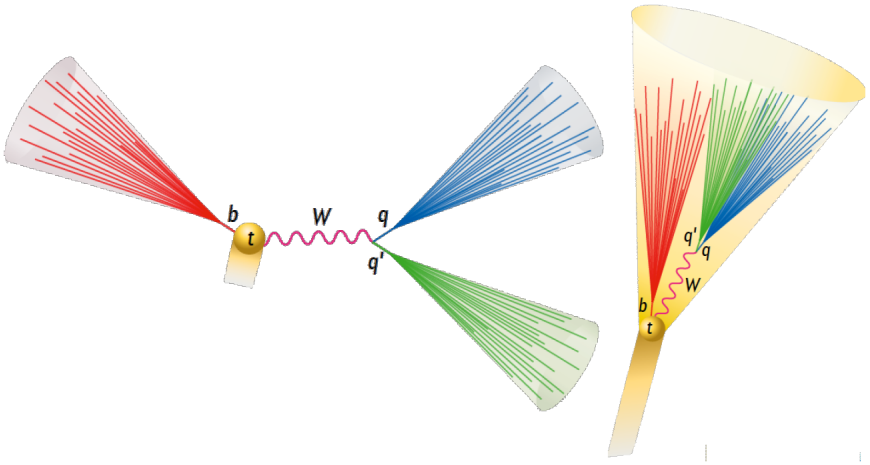
[2] Similarly to event-shapes, jet-shape variables are designed to describe the geometry of the jet substructure. Given N sub-jet axes within a jet, N-subjettiness can tell how N-prong-like the jet looks. It is very powerful to identify and tag jets originating from heavy and light partons, which exhibit a very different jet substructure [187, 188].

given by:

$$\cos(\theta') = \frac{\cos(\theta) + \beta}{1 + \beta\cos(\theta)} \ . \tag{4.3}$$

For particles with large boost $\beta \to 1$, it holds that $\cos(\theta') \to 1$, so the decay products are collimated in the direction of the boosted initiating parton. A high-energy particle with mass m and transverse momentum p_{T} yields a collimated spray of decay products mostly falling into a cone with an angular size of $2m/p_{\mathrm{T}}$. For sufficiently boosted heavy particles, the resulting cascade of particles can be fully captured in a single jet if its catchment area is big enough. This is achieved by setting the jet radius $R \sim 1$, and these big jets are commonly referred to as large-R jets. The small-R jets at the LHC use $R = 0.4$ and are meant to cluster jets initiated by light quark and gluons.

The topology of the top quark decay products offers a good example to illustrate the applicability of large-R jets in boosted regimes. As depicted in Figure 4.3, the decay products of a top quark produced with an energy close to its on-shell mass are nearly emitted back-to-back, so they can be reconstructed as three well-separated jets. The W boson is also produced with low momenta so its hadronic decay products are well angularly separated too. This allows to resolve the three small-R jets describing each top decay sub-product. On the contrary, the decay products are boosted in the direction of the top quark if the top quark is produced with $p_{\mathrm{T}} \gg m_t$. It is no longer possible to accurately isolate and resolve in small-R jets the cascade of particles initiated by each top decay sub-product in a boosted regime. A large-R jet is more appropriate in this case, treating the top quark decay as a whole. The hadronic decay of heavy particles leads to as many sub-jets or prongs as there were primary decay products. In this way, a W-initiated jet will more likely show a two-pronged substructure, whilst a top-initiated jet is characterized by a three-pronged one.



(a) Rest (b) $p_T >> m$

Figure 4.3: Sketch of the hadronic final state topology involving top quarks in the resolved (a) and boosted (b) regime.

4.2.2 Jet constituents

At detector level, jets can be reconstructed from a wide variety of entities. ATLAS has been historically using calorimeter cells as standard inputs for the nominal jets, exploiting in such way the exceptional energy resolution of the calorimeters [189]. The lateral and longitudinal segmentation of the calorimeters allows the combination of calorimeter cells into three-dimensional, massless, topological clusters. Such cells are required to have a ratio of energy measured to the expected average energy due to the noise in that cell (cell significance) above certain level:

$$\zeta_{\text{cell}}^{\text{EM}} = \frac{E_{\text{cell}}^{\text{EM}}}{\sigma_{\text{noise,cell}}^{\text{EM}}}, \tag{4.4}$$

where $\sigma_{\text{noise}} = \sqrt{\left(\sigma_{\text{noise}}^{\text{electronic}}\right)^2 + \left(\sigma_{\text{noise}}^{\text{pile-up}}\right)^2}$. It is important to remark that this topological clustering algorithm is not specially designed to associate energy deposits to the initial hitting particle. It is rather meant to separate continuous energy showers of different nature. Therefore, topoclusters by themselves are not the most suitable objects if one is interested in jet substructure.

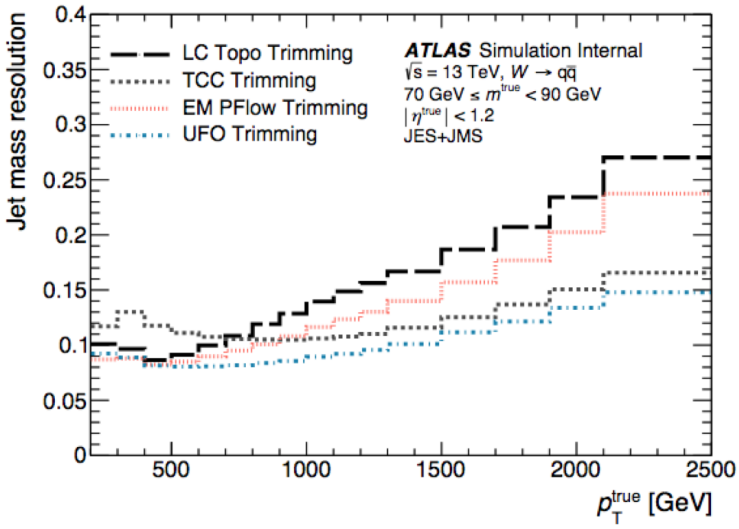


Figure 4.4: Jet mass resolution for W jets as a function of the large-R truth jet p_{T} for jet building blocks.

As objects become more collimated in highly-boosted regimes, the calorimeters lack the angular resolution and granularity to resolve two close-by sub-jets. To tackle this problem, ATLAS has developed new algorithms that make the most of the complementary information provided by both calorimeters and inner detectors. The key point is that the high granularity and spacial resolution of the ID allow to split and merge the topoclusters according to the associated tracks. One of the most promising objects are the so-called *Unified-Flow Objects* (UFOs) [190]. Jets built with UFO objects show the best and more stable performance in a wide p_{T} range, as shown in Figure 4.4.

4.2.3 Jet tagging

To fully exploit the information encoded in the reconstructed jets and hence in the final states they populate, it is very important to accurately identify the jet initiating particle. In the ATLAS experiment, this is done with the so-called tagging algorithms, which typically rely on jet substructure observables. The identification of jets containing the hadronic decay of heavy particles, such as

boosted W/Z bosons and top quarks, and B hadrons against jets initiated by C hadrons, light quarks or gluons is of major importance in order to reject events produced by background processes and improve the sensitivity in searches for physics BSM [191, 192].

After the jet reconstruction, where the primary goal is to most accurately reconstruct the interesting energy flows in signal jets while suppressing the contributions from UE and PU, the information carried by the jet constituents is used in a number of methods to finally distinguish between signal and background jets. The b-jet identification, for instance, is carried out in a two-stage approach; low-level algorithms are first used to reconstruct the characteristic features of b-jets, followed by high-level algorithms that take as input the result of the low-level b-tagging and combine them with a multivariate classifier. The performance of a b-tagging algorithm is characterized by the probability of tagging a b-jet (efficiency) and the probability of wrongly identifying a c-jet or light-flavour jet as a b-jet.

4.3 Soft-component mitigation techniques

The soft-radiation generated in the PS and the UE are sources of soft radiation present at any event recorded with the ATLAS detector. They contribute to the pile-up problem and affect the observed jet properties, altering the radiation pattern within the jet and hence biasing and smearing the jet scale and resolution. The jet mass is specially affected by these soft effects, as illustrated for hadronically-decaying boosted W bosons in Figure 4.5. Soft radiation included in the jet clustering may contribute to the final jet mass at the same level as particles carrying large fractions of the total jet p_{T}.

Since this soft energy background is spread fairly uniformly throughout the

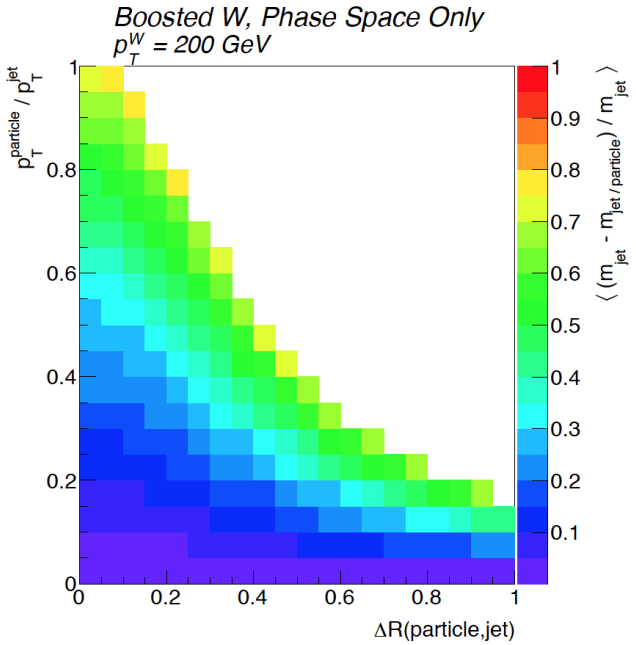


Figure 4.5: Impact in the total jet mass of the jet constituents as a function of the angular distance between the jet axis and jet constituent ΔR and the p_{T} fraction carried by the constituent. Image taken from Ref. [135].

detector, the contribution is proportional to the jet catchment area. Hence, large-R jets are more affected than small-R jets. Several techniques can be applied to remove such undesirable radiation. A very first action is taken during the constituent reconstruction, where those entities that do not pass certain quality criteria are discarded. The ATLAS experiment has already tested constituent-level techniques with encouraging results in terms of pile-up resilience, such as the *Constituent Subtraction* (CS) [193] and *Soft-Killer* (SK) methods [194].

In this work, we have used techniques that operate at jet level, *i.e.* after jet reconstruction. They are commonly known as grooming techniques, and usually depend on several parameters that are carefully adjusted to optimize the jet performance, so the resulting jets can be used in a wide variety of physics applications. Let's see the two grooming techniques of interest for this thesis:

- Trimming [195]: sub-jets of radius R_{sub} are built from the jet constituents by means of the k_t algorithm. After that, those sub-jets with a p_{T} fraction smaller than the f_{cut} of the jet p_{T} are removed. This procedure is depicted in Figure 4.6. In ATLAS, the default values of R_{sub} and f_{cut} are 0.2 and 5%, respectively. The trimming technique is specially useful to provide an accurate description of the total jet energy, but it is less suitable for jet substructure studies.

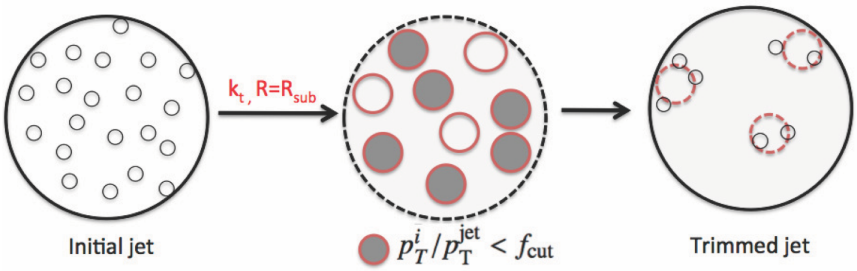


Figure 4.6: Diagram depicting the jet trimming procedure. Figure taken from Ref. [195].

.

- *Soft-Drop* (SD) [196]: this declustering technique removes soft and wide-angle radiation from a jet. Two parameters are required: z_{cut} and β. The former is a soft threshold and the latter affects the amount of soft-collinear components into the jet (the larger values β takes, the more soft-collinear radiation is maintained, having with $\beta = 0$ the limit of a purely cut on soft radiation). Basically, this algorithm starts by reclustering the jet constituents of jets defined with the anti-k_t algorithm and radius R_0 using the C/A algorithm, yielding an angular ordered clustering tree. Next, consider the sub-jets by undoing the last step of the clustering tree, *i.e.* break the proto-jet J into two sub-jets. The SD algorithm removes the softer sub-jet unless:

$$\frac{\min(p_{\text{T},1}, p_{\text{T},2})}{p_{\text{T},1} + p_{\text{T},2}} > z_{\text{cut}} \left(\frac{\Delta R_{12}}{R_0} \right)^{\beta} , \tag{4.5}$$

where ΔR_{12} is the distance between the two sub-jets in the $\eta - \phi$ plane and $p_{\mathrm{T},i}$ are the transverse momenta of the sub-jets. If they pass this condition, the proto-jet j becomes the SD jet. If not, it takes the branch with larger p_{T} as the new proto-jet and iterate the procedure until the condition is hold. The final groomed jet radius R_g is therefore defined by the last R_{12} distance. One of the most interesting features of this technique is that it can be implemented in first-principles calculations for jet-related observables.

4.4 Jet calibration

Due to a number of effects, the reconstructed jet properties still does not correspond to that of the particle-level jet. The jet response of the ATLAS detector is defined as the ratio of the reconstructed jet 4−vector to the corresponding particle-level jet one. The mean of the jet response is the jet scale and the standard deviation is the jet resolution. In this way, the energy component of the jet response gives rises to the *Jet Energy Scale* and *Resolution* (JES/JER). This allows to correct the reconstructed jet energy in a process called jet calibration.

This calibration process can also be applied to the jet mass, in terms of the *Jet Mass Scale* and *Resolution* (JMS/JMR). The calibration procedure that the standard large-R jets undergo in the ATLAS experiment comprises several stages, as sketched in Figure 4.7.

The first correction of the jet calibration chain is carried out at constituent level before the jet building [189]. Signals on the calorimeters have inefficiencies reflected in the energy scale of the topoclusters (such as signal losses due to clustering, energy lost in inactive material or different response to hadrons and photons/electrons). The *Local Cell Weighting* (LCW) scheme aims to bring

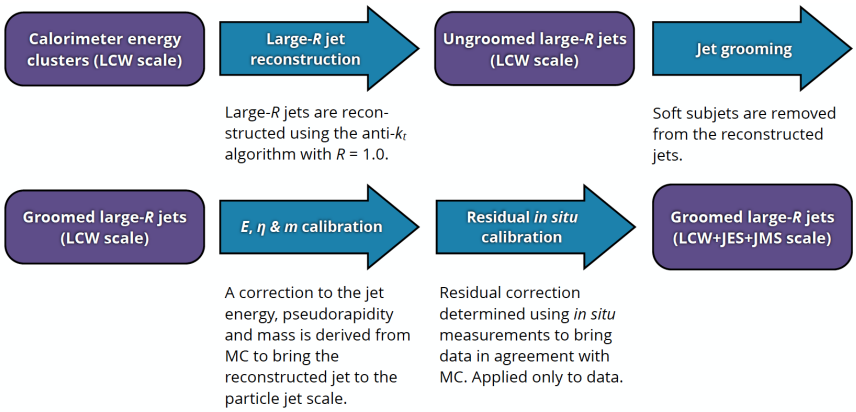


Figure 4.7: The large-R jet calibration and reconstruction chain. Image taken from Ref. [197].

the topoclusters formed in the hadronic calorimeter to the hadronic scale by addressing such inefficiencies through MC-based corrections.

The second correction is also derived from MC simulations. This MC-based calibration consists on bringing the scale of the reconstructed jets to the scale of the matched particle-level jets for given observable. The jet rapidity, energy and mass are sequentially corrected, in this order. The average jet mass response, for instance, defined as $\mathcal{R}_m = m_{\text{reco}}/m_{\text{truth}}$ is not centered around one due to the detection process, as shown in Figure 4.8. The measured jet mass is jet-by-jet corrected by the $\mathcal{C}_m = 1/\mathcal{R}_m$ factor. The resulting average mass response peaks in one, represented by the red points. In this step, the jet response is corrected, to the extent that the Monte Carlo simulation describes the response. This calibration is applied to jets from both simulated and real events.

Any difference between measured and simulated jets are accounted for in a direct data-to-MC comparison. The in situ calibration allows to check the closure of the MC-based calibration in data and correct for those effects that escape the MC-based calibration. Similarly to the MC-based calibration, the jet $|\eta|$, energy and mass are sequentially calibrated. The $|\eta|$-intercalibration extends the well measured jet energy response in the central region of the calorimeter ($|\eta_{det}| < 0.8$)

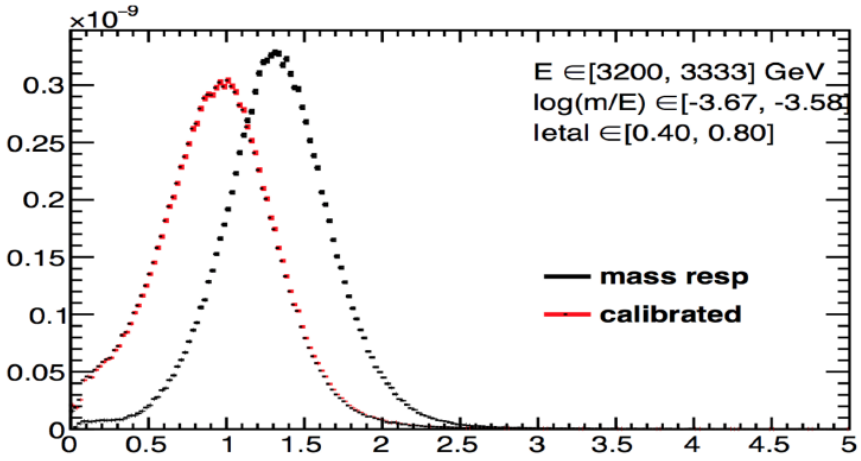


Figure 4.8: Monte Carlo based calibration of the average large-R jet mass.

to the forward region ($0.8 < |\eta_{det}| < 2.5$), so the energy response is uniform across the $|\eta|$ range. The energy calibration relies on events where the large-R jet is balanced against a reference system whose energy is well-measured. Examples are photons, Z bosons with leptonic decay, or systems formed by several small-R jets. The p_{T} ratio between the large-R jet and the reference system yields a narrow peak close to one. The ratio of the peak position determined in MC simulation and data yields the in situ JES/JER correction.

The in situ jet mass calibration, the very last step of the large-R jet calibration chain, deserves a more detailed discussion. Two methods are currently used: the *Forward-Folding* (FF) and the R_{trk}. In the R_{trk} method, tracks left by charged particles in the ID are used to build the so-called track jets, which are geometrically matched with calorimeter-only based jets. In this way, the ATLAS detector provides two independent measurements of the properties of particle-level jets. The point is that the difference between track jets and the associated particle-level jets built only with charged-particles is small compared to the resolution of the calorimeter-only based jets. However, when considering full particle-level jets, the fluctuation of the charge-to-neutral ratio in track jets

enhances the mass resolution, making the difference between the two no longer small compared to the mass resolution of the calorimeter-only based jets. This implies that one cannot use track jets by themselves to check the closure of the jet mass calibration in data, but enables the possibility of using the deviation of the ratio of track and calorimeter jet masses:

$$R_{\text{trk}} = \left\langle \frac{m_{\text{calo}}}{m_{\text{track}}} \right\rangle , \tag{4.6}$$

between data and MC, $R_{\text{trk}}^{\text{data}}/R_{\text{trk}}^{\text{MC}}$. Any deviation from unity of this double ratio is taken as a mismodelling of the jet mass response and, hence, assigned as an uncertainty to the JMS.It is important to emphasize at this point that both, the R_{trk} and FF techniques, provide measurements of the relative (data-to-MC) JMS, and not of the absolute value of the JMS.

The R_{trk} method has been successfully used to calibrate large-R jets built with topoclusters in the Run 2, allowing to characterize the JMS of jets with p_{T} up to 3 TeV. However, the next generation of jets are made of constituents that benefits from tracking information (such as UFOs), so the R_{trk} approach is no longer valid. The FF technique discussed in the next section therefore becomes of paramount importance.

4.4.1 The Forward-Folding method

The calibration of some objects such as electrons [198] and muons [199] in the ATLAS detector relies on the fact that the underlying line-shapes are well-known. This allows to use specific parametric functions to describe the detector response, which have the advantage of a clear and direct interpretation of the fitted parameters. However, these techniques cannot be applied in the calibration of the large-R jet mass originating from hadronic decays due the PS and the jet

clustering. The jet mass response is a complex, non-gaussian transformation, which can hardly be modelled analytically. Furthermore, the jet mass response may depend on certain particle-level features, such as the jet multiplicity, which varies across different MC generators. All in all, any technique developed to calibrate the jet mass response should not assume a specific form of the response and disentangle detection from physics effects in the simulation.

The Forward-Folding technique incorporates these two features successfully. Firstly, it takes the jet response from the MC simulation, so no assumption is needed about its specific analytic form. Secondly, for given particle-level spectrum, the FF method provides detector-level distributions where the nominal detector jet mass response is modified, enabling the simultaneous determination of the jet mass scale and resolution in data with respect to MC. It only needs sharp distributions that exhibit a good sensitivity to the jet mass, such as the reconstructed jet mass from hadronically-decaying top quarks and W bosons in $t\bar{t}$ events. It has been already used to calibrate jets from $\sqrt{s}$ = 8 and 13 TeV data [197, 200, 201]. A detailed description of this method was given in Ref. [202], which also collects the main results obtained from the full Run 2 dataset and the extension of its applicability from $t\bar{t}$ events to W/Z+jet topology, where the large-R jet recoils against a highly-boosted vector boson.

The FF method proceeds as follows: consider $\mathcal{O}^{\text{reco}}$, $\mathcal{O}^{\text{true}}$ and $\mathcal{R}\left(\mathcal{O}^{\text{reco}}, \mathcal{O}^{\text{true}}\right)$ as the detector-level, particle-level and detector response distributions for given observable $\mathcal{O}$ extracted from simulation, respectively. The detector-level distribution for a given scale s and resolution r is generated by folding the particle-level information, event-by-event, according to:

$$\mathcal{O}^{\text{reco}}_{\text{folded}} = s \times \mathcal{O}^{\text{reco}} + \left(\mathcal{O}^{\text{reco}} - \mathcal{O}^{\text{true}} \times \mathcal{R}\left(\mathcal{O}^{\text{true}}, \mathcal{O}^{\text{true}}\right)\right) \times \left(r - s\right) \ . \tag{4.7}$$

By construction, the mean and the standard deviation of the folded distribution can be understood as the original ones but scaled by a factor s and r respectively:

$$< \mathcal{O}^{\text{reco}}_{\text{folded}} >= s \times < \mathcal{O}^{\text{reco}} >, \quad \sigma\left(\mathcal{O}^{\text{reco}}_{\text{folded}}\right) = r \times \sigma\left(\mathcal{O}^{\text{reco}}\right) \ . \tag{4.8}$$

For $s = r = 1$ the nominal jet mass response remains unaltered, and for $s = 0$ and $r = 1$ one gets back $\mathcal{O}^{\text{true}}$. The aim of this procedure is to find the modified jet mass response in MC that best describes data. To do so, a two-dimensional $\chi 2$ fit is performed where data and the alternative detector-level distributions are compared, assuming that the histogram bins are uncorrelated. The 2D-$\chi 2$ minimisation points to the template that best describes data, and the associated s and r values are therefore identified as the relative JMS = $\text{JMS}^{\text{data}}/\text{JMS}^{\text{MC}}$ and JMR = $\text{JMR}^{\text{data}}/\text{JMR}^{\text{MC}}$:

$$(r, s) = \arg\min_{(r',s')} \sum_{i=1}^{n_{\text{bins}}} \left(\frac{h_i^{\text{sim}}(r', s') - h_i^{\text{data}}}{\sigma_i} \right)^2 \ , \tag{4.9}$$

where h_i^{sim} and h_i^{data} represent the i^{th} bin of the $\mathcal{O}^{\text{reco}}_{\text{folded}}$ histogram in simulation and the $\mathcal{O}^{\text{reco}}$ distribution in data, respectively. Each histogram has n_{bins}. The uncertainty σ_i only includes data and MC statistical uncertainties.

4.4.2 In situ calibration from $t\bar{t}$ events

In principle, the FF method can be applied to any observable which particle- and detector-level information are accessible, so the associated detector response can be estimated. The observable is further required to has a resonance-like structure to properly constraint the r and s parameters. It has been used for mass peaks, because the reconstructed jet mass distribution of jets that capture the full decay of heavy particles exhibits a peak centered around the on-shell mass

of the initiating particle. The position and width of this peak are related to the jet mass scale and resolution, respectively.

A high purity sample of large-R jets containing hadronically-decaying W bosons from top quark pairs can be obtained by applying an event selection based on the standard $t\bar{t}$ resonance searches of the ATLAS experiment [203]. The semileptonic final state is selected where one W boson decays hadronically and the opposite W boson decays into lepton and neutrino. In this analysis, only muons are accepted in the leptonic side of the $t\bar{t}$ system. This eases the task of reducing the contribution of the QCD multijet background and helps to identify the final state since muons leave a clean signature in the muon spectrometer. The complete event selection applied is summarized in Table 4.1. An isolated, high-p_{T}

Table 4.1: Summary of the event selection for $t\bar{t}$ events in the μ+jets channel. Specific selections enhancing jets originating from W bosons or top quarks are also described. Table taken from Ref. [197].

Object	Selection	Description
Muon (μ)	Single-muon triggers	Trigger
	$p_{\mathrm{T}}^{\mu} > 27$ GeV, $\lvert\eta\rvert < 2.5$	Preselection
	`tight` muon identification	Identification
	$\Delta R_{\mu,\ \text{small-}R\ \text{jet}} > 0.4$	Isolation
E_T^{miss}	$E_T^{\mathrm{miss}} > 20$ GeV , $m_{\mathrm{T}} > 60$ GeV	Leptonic W decays
small-R Jets (j)	$p_{\mathrm{T}} > 25$ GeV, $\lvert\eta\rvert < 2.5$	Jet selection
	at least one jet with $\Delta R(j,\mu) < 1.5$	Boosted semileptonic top
	at least one b-tagged jet	Flavour tagging
large-R Jets (J)	$p_{\mathrm{T}} > 200$ GeV, $\lvert\eta\rvert < 2$	Preselection
	$\Delta R_{\mathrm{J,closest\ j\ to}\ \mu} > 2$	Opposite hemispheres
W-initiated Jets	$p_{\mathrm{T}} \in [200, 350]$ GeV	W selection
	$\Delta R_{b,\mathrm{J}} > 1$	
Top-initiated Jets	$p_{\mathrm{T}} \in [350, 1000]$ GeV	Top selection
	$\Delta R_{b,\mathrm{J}} < 1$	

central muon, along with a significant missing transverse momentum, are required to identify semileptonically decaying top quarks. Single-muon triggers requiring an isolated energy deposit in the muon spectrometer greater than 26 GeV are used to reject most of the events from QCD multijet background processes. Muons

with $p_{\mathrm{T}} > 27$ GeV and $|\eta| < 2.5$ which satisfy tight quality criteria [199] and are separated from small-R jets by $\Delta R > 0.4$ are considered. To enhance the selection of events where the muon and E_T^{miss}originate from a W boson the transverse mass m_{T}, defined as $m_{\mathrm{T}} = \sqrt{2p_{\mathrm{T}}^{\mu} E_T^{\mathrm{miss}}(1 - \cos\left(\Delta\phi_{\mu, E_T^{\mathrm{miss}}}\right))}$ is required to be greater than 60 GeV and $E_T^{\mathrm{miss}} > 20$ GeV. One small-R jet is required to satisfy an angular distance to the muon $\Delta R(j, \mu) < 1.5$, in order to identify the jet from the semileptonic top quark decay. Events with at least one b-tagged jet are required to further enhance the $t\bar{t}$ signal over other background processes. Finally, at least one large-R jet with a transverse momentum greater than 200 GeV is required to exist, lying in the opposite hemisphere of the semileptonic top quark decay.

Those events that pass the fiducial selection are further categorised according to the angular distance between the closest b-tagged small-R jet and the selected hadronic-like large-R jet, $\Delta R_{b,\mathrm{jet}}$. In the top selection, $\Delta R_{b,\mathrm{jet}} < 1$ so the large-R jet is deemed to capture the whole top quark decay and the mass distribution is expected to be peaked around the top quark mass. In the W selection, it holds that $\Delta R_{b,\mathrm{jet}} > 1$ so the b-jet lies outside the large-R jet and the resonance is expected to be centered at the W boson mass. With this classification, the jet mass response can be calibrated in two separate mass windows, 50 GeV $< m_J <$ 120 GeV and 120 GeV $< m_J <$ 300 GeV, which also allows to check any potential dependence of the detector performance on the jet topology due the presence of the b fragmentation. Furthermore, to enhance the probability of the large-R jet to capture the whole decay of a top quark, the top selection requires the selected large-R jet to have a p_{T} between 350 GeV $< p_{\mathrm{T}} <$ 1000 GeV. Signal jets containing the hadronic W boson decay are required to fulfill 200 GeV $< p_{\mathrm{T}} <$ 350 GeV. These p_{T} ranges are in turn binned according to the statistics available in order to characterize the mass response with the highest granularity possible.

Two different jet definitions are used in this study. On the one hand, jets

built exclusively from topoclusters calibrated to the hadronic scale with the anti-k_t algorithm and groomed with the trimming technique (R_{sub} = 0.2 and f_{cut} = 5%) constitute the nominal large-R jet collection. These jets fully rely in the information provided by the calorimeters, and the associated mass m_{calo} is given by:

$$m_{\text{calo}} = \sqrt{\left(\sum_{i\epsilon J}^{\text{topo.}} E_i\right)^2 - \left(\sum_{i\epsilon J}^{\text{topo.}} \vec{p}_i\right)^2} \ , \tag{4.10}$$

where E_i is the energy of the topocluster i and $\vec{p}_i$ is a vector with magnitud E_i and direction $(\phi, \eta)_i$. A complementary jet mass is obtained by considering tracking information. The *Track-Assisted* (TA) mass is the result of reweighting the track jet mass associated to the calorimeter jet mass by the charge-to-neutral fraction:

$$m_{\text{TA}} = m_{\text{track}} \times \frac{p_{\text{T}}^{\text{calo}}}{p_{\text{T}}^{\text{track}}} \ . \tag{4.11}$$

On the other hand, jets can be made by clustering calibrated small-R jets, built from topoclusters calibrated to the electromagnetic scale, into large-R jets with the anti-k_t algorithm [204]. They are called *ReClustered* (RC) jets and the associated mass m_{RC}, simply resulting from the combination of the individual jet four-vectors, take full advantage of the precise calibration of the small-R jets.

The event selection and jet reconstruction laid out above is applied not only to real data but to simulated events as well. The nominal MC simulation consists of the NLO in QCD Powheg Box (v2) [73, 74, 205] as event generator of the hard scatter, interfaced to Pythia (v8.2.1.0) [206] for fragmentation and hadronization processes. The h_{damp} parameter is set to $1.5 \cdot m_t$ in the event generator. Non-perturbative QCD effects are modelled using a set of tuned parameters called *A*14 tune [207]. Two sources of systematic uncertainties are accounted for in the extraction of the relative JMS and JMR, which are estimated by varying the nominal simulation setup and repeating the fit to data. First, the theoretical

modeling uncertainty on the particle-level jet mass distribution. It includes MadGraph5_aMC [208] as alternative ME generator, HERWIG (v2.7.1) [80, 209] for an alternative shower and hadronization models and finally variations on the ISR/FSR, derived with different values of the normalization and factorization scales, the h_{damp} parameter and variations of the $A14$ tune. Second, experimental uncertainties that may alter the event selection and bias the jet reconstruction. It comprises a set of variations where uncertainties on the electron, muon, photon and small-R jet reconstruction are properly propagated. The total uncertainty is obtained by adding in quadrature the statistical, detector and modelling components.

The jet mass distributions for events passing the W and top selection are depicted in the left and right columns of Figure 4.9, respectively. Data is represented by black markers with the associated statistical uncertainty error bar. The simulation is shown with solid lines and the $t\bar{t}$ signal is displayed on top of all contributing sources of background. The envelope of the systematic uncertainties is represented by the dashed band. As can be observed from the lower panels where the data-to-MC ratio is computed bin-by-bin, there is a good agreement between the two in the peak region. A small correction is therefore expected. The simulated mass response, extracted from the nominal MC setup, is profiled in terms of two sensitive variables: the reconstructed large-R jet p_{T} and the mass of the associated large-R jet at particle level. One could profile the response function in more variables as long as it does not lead to significant statistical fluctuations. Figure 4.10(a) shows the average mass response for calorimeter-only trimmed jets and, as can be noted, it is fairly uniform and near unity in the p_{T} and mass ranges considered in this analysis. Some extreme kinematic regions exhibit a response far from 1, but only a small fraction of the total selected jets fall in here.

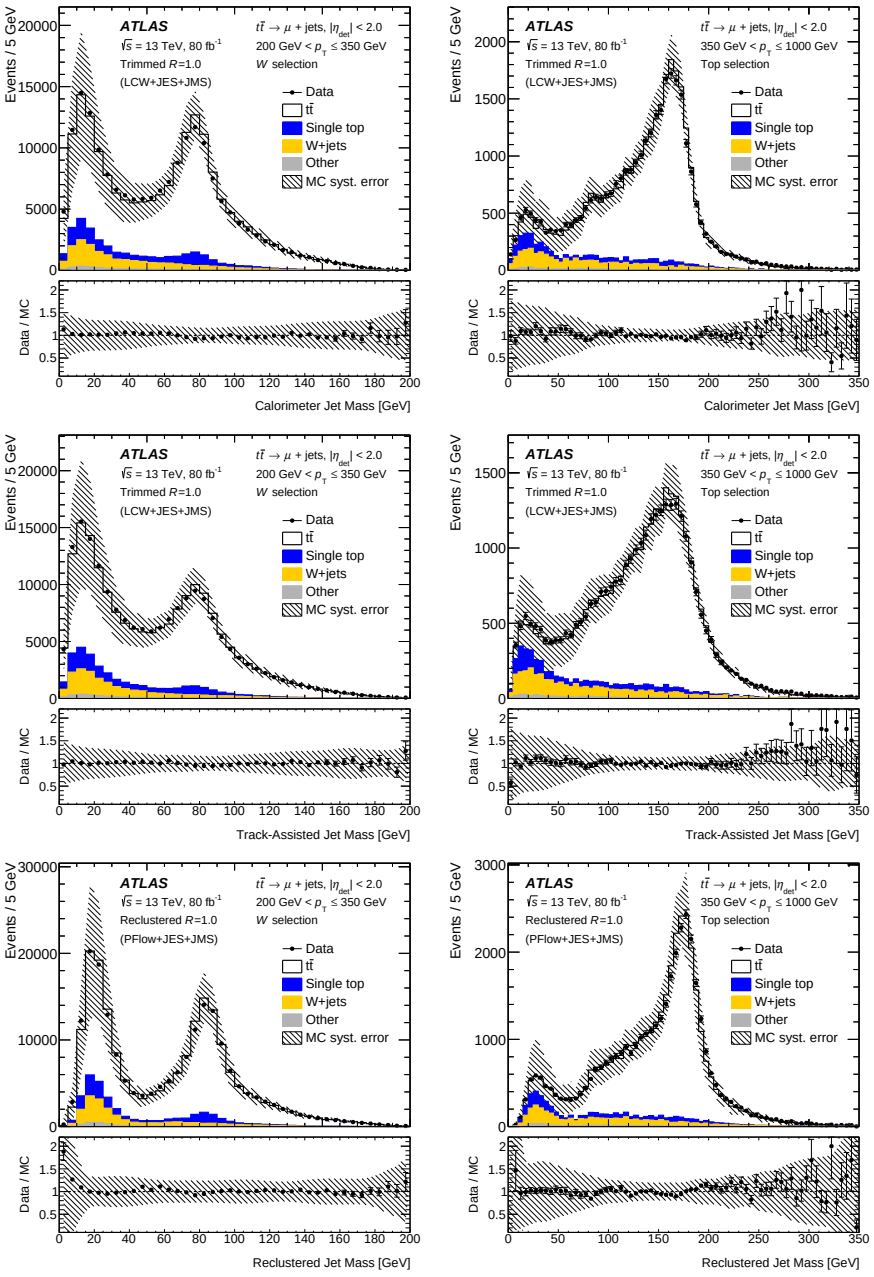
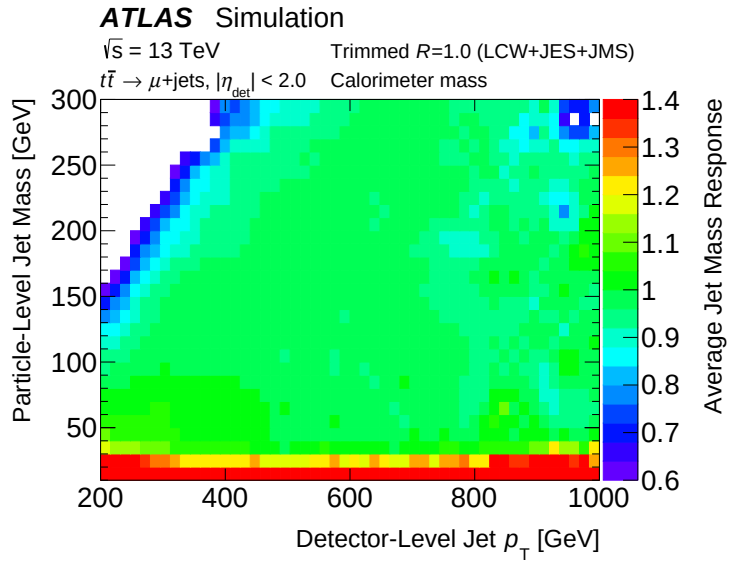
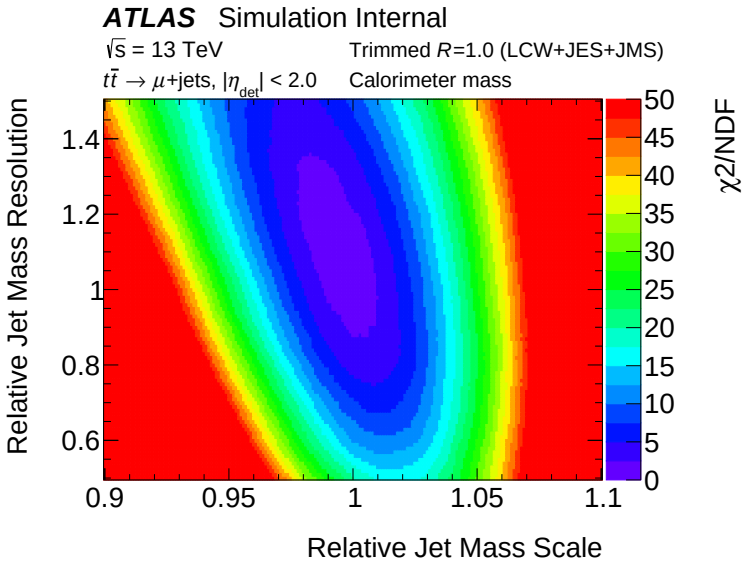


Figure 4.9: The trimmed jet m_{calo}, trimmed jet m_{TA} and reclustered jet mass distributions obtained from events passing the (left) W selection and (right) top quark selection from $t\bar{t}$ events in the first, second and third row, respectively. The "Other" category refers to diboson and Z+jets events. Figure taken from Ref. [197].



(a)



(b)

Figure 4.10: On the top: average mass response obtained from the nominal Monte Carlo simulation of calorimeter-only trimmed jets (figure taken from Ref. [197]). On the bottom: illustrative two-dimensional $\chi 2$ fit of the Monte Carlo templates generated via the Forward-Folding method to data collected by the ALTAS detector.

As explained previously, the alternative MC templates are obtained by modifying the average jet mass response according to Eq. 4.7 and compared to data following Eq. 4.9, yielding a two-dimensional $\chi 2$ distribution. The two dimensions represent the relative JMS and JMR and, as can be observed from Figure 4.10(b), the tilt of the associated ellipse certainly signal a correlation between the two. The 2D-$\chi 2$ distribution is further marginalized onto each dimension and minimised in order to find the MC template with the lowest $\chi 2$ value. The quality of the fit is found to be good overall. The associated $\chi 2$ are summarized in Table 4.2.

Table 4.2: Best-fit $\chi 2$ values associated to the MC-to-data comparisons.

jet p_{T} [GeV]	$\chi 2_{\mathrm{min}}$		
	m_{calo}	m_{TA}	m_{RC}
$[200, 225]$	1.63	0.46	1.73
$[225, 250]$	0.99	0.84	2.32
$[250, 275]$	1.10	1.00	1.64
$[275, 300]$	0.46	0.59	1.22
$[300, 325]$	0.25	0.51	1.17
$[325, 350]$	0.53	0.79	0.64
$[350, 400]$	0.76	0.71	0.82
$[400, 500]$	1.00	0.70	0.61
$[500, 600]$	0.94	0.51	0.83
$[600, 1000]$	0.54	0.60	1.02

The associated relative JMS and JMR values are depicted in Figure 4.11. The upper panel shows the relative JMS and the lower panel the relative JMR. Black and red points corresponds to W- and top-induced jets in $t\bar{t}$ events. Each point has two vertical bars associated: the inner stands for the statistical error only and the outer accounts for the total uncertainty, *i.e.* statistical and systematic. The number of events decreases quickly as the jet p_{T} grows. Consequently, the statistical uncertainty increases in the higher p_{T} bins as can be noted. The leading source of uncertainty in most of the regions is coming from the MC modelling, as reported in Figures 4.12. The modelling of hadronization and showering processes arises as the dominant source of uncertainty. The total uncertainty is found to be of the order of 1% for the relative JMS and of 15% for the relative JMR, which represents a significant improvement with respect to the first iteration (2.4% and 18%) [200].

From Figure 4.12 we conclude that the relative JMS and JMR are compatible with one and roughly constant across the p_{T} range within uncertainties. Since most of the deviations from unity are covered by the total uncertainty, the modelling description of the detector response from simulation is quite accurate. Small differences are found in the extracted JMS/JMR for the three jet mass definitions considered. The calibration factors for RC jets are in general closer to unity. This may be expected because the small-R jets they are built from are already calibrated. Furthermore, the calorimeter-only based jet mass shows a better performance than that of the TA jet mass.

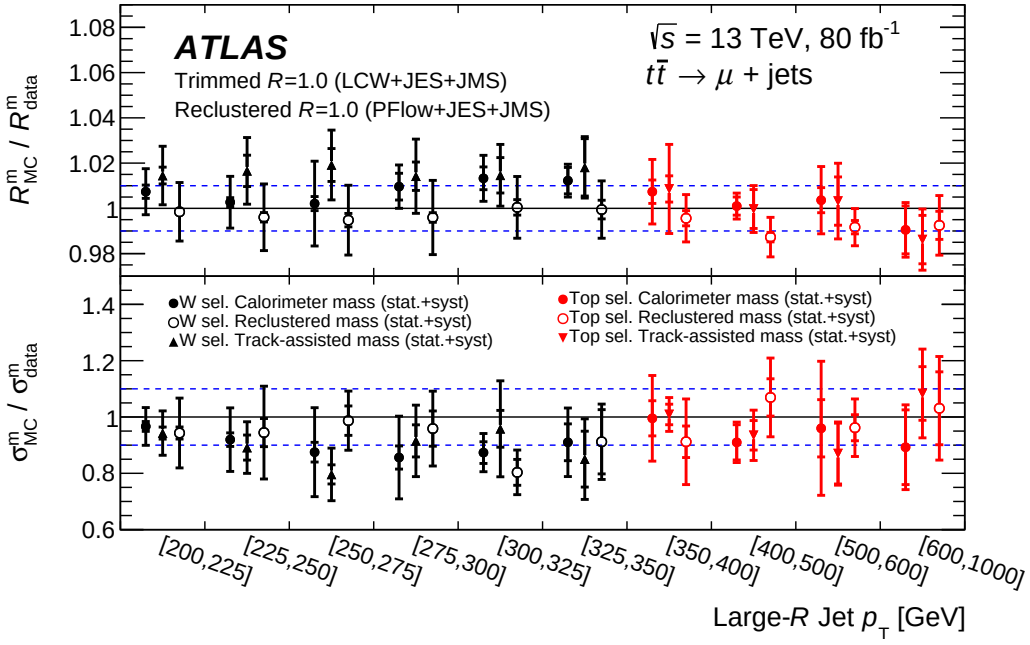
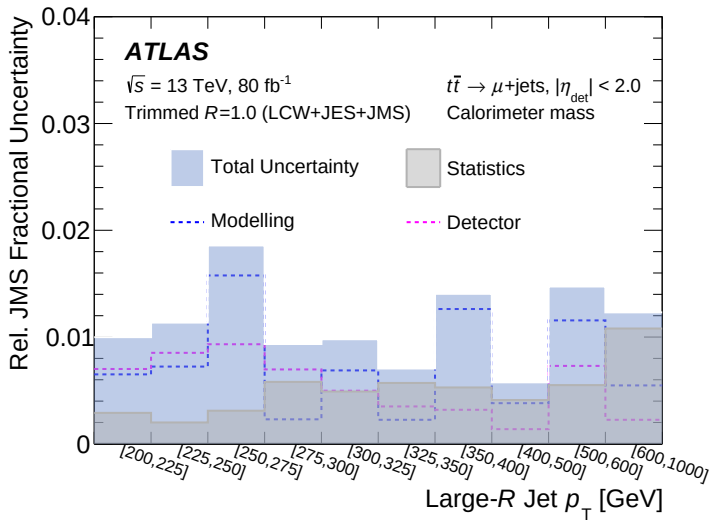
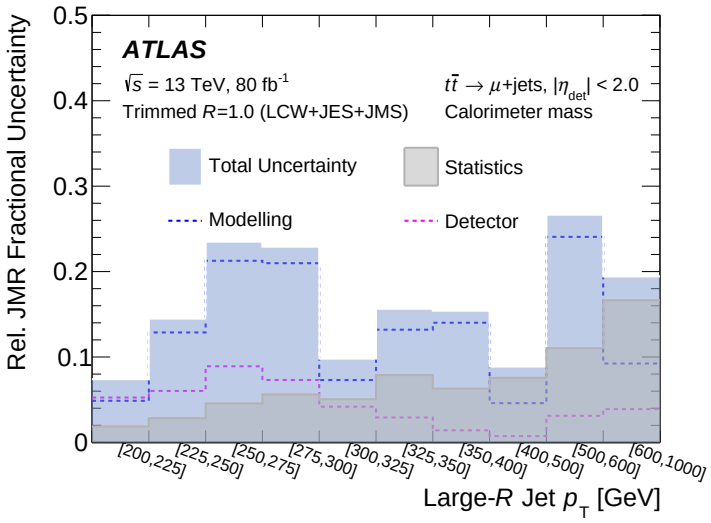


Figure 4.11: The relative Jet Mass Scale and Resolution for several jet mass definitions as a function of the jet p_{T}. Black and red markers represent results for W-like and top-like jets. The inner error bar stands for the statistical uncertainty and the outer error bar includes the statistical and systematic uncertainty. Dashed blue lines, corresponding to ±1% for the JMS and ±10% for the JMR, are drawn for reference. Figure taken from Ref. [197].



(a)



(b)

Figure 4.12: Breakdown of the total uncertainty associated to the relative Jet Mass Scale in (a) and Jet Mass Resolution in (b) for calorimeter-only trimmed jets in $t\bar{t}$ events. Similar trends are found for Track-Assisted and Reclustered jet masses. The estimation of the Monte Carlo modelling uncertainties is limited in accuracy by the statistics available in the Monte Carlo samples used. Figure taken from Ref. [197].

Chapter 5

Interpretation of the top quark mass in ATLAS Monte Carlo generators

In this chapter, a preliminary study of the top Monte Carlo mass m_t^{MC} in ATLAS generators is presented. We derive a quantitative relation between the mass parameter in MC generators by comparing their predictions to a first-principles calculation. This is carried out at particle level, so the effects induced by the treatment of non-perturbative QCD aspects in the MC simulation are directly tested. The observable used to perform this study is the invariant mass of lightly-groomed large-R jets initiated by hadronically decaying top quarks. This observable is especially convenient because it shows a kinematic sensitivity to the top quark mass, as the observables typically used in direct measurements of m_t. The interpretation of m_t^{MC} in the simulation will be studied for several MC configurations and for closely-related observables.

5.1 Introduction

In the previous chapter, we have seen that the jet substructure refers to a set of tools devoted to exploit information from the radiation pattern inside jets. Such radiation pattern is naturally shaped by the initiating parton. If the initiating parton is produced sufficiently boosted and the jet size is big enough, the jet and parton masses are closely related. Let's consider a large-R that captures all the decay products of a top quark. The associated jet mass is given by:

$$M_J^2 = \left(\sum_{i \in J} p_i^\mu\right)^2 \simeq m_t^2 + \Gamma_t m_t + \dots \,, \tag{5.1}$$

where p_i^μ is the four-momentum of the jet constituents. We explicitly see that the kinematic sensitivity of the jet mass with the initiating particle originates in the reconstruction of its decay products. The jet mass is mainly determined by the top quark mass m_t and width Γ_t. The mass distribution of top-initiated jets exhibits a resonant structure, and the highest sensitivity is in the peak region where top quarks are produced close to its mass shell ($M_J^2 - m_t^2 \sim m_t\Gamma_t$). Therefore, the location of the peak can be used to infer m_t. As already explained in Section 2.3.4, the sensitivity carried in resonant-like structures is typically exploited in top quark mass direct measurements.

The top quark production and the jet formation span energy scales of several orders of magnitude: the momentum transfer in the hard-scatter Q and the top mass and width, having $Q \gg m_t \gg \Gamma_t \gtrsim \Lambda_{\mathrm{QCD}}$. The mass sensitivity of this kind of observables comes from the hard-scatter and energetic radiation at scales $\gtrsim m_t$ on the one hand, and from soft-collinear radiation off of top quark and its decay products at scales $\ll m_t$ on the other. The sensitivity generated in the soft sector can be partially washed out in hadron collisions due to uncorrelated soft emissions being captured by the jet, such as ISR and UE. This is schematically

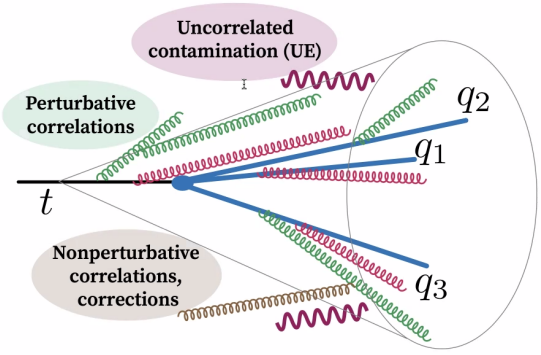


Figure 5.1: Sketch of the different sources of radiation entering in the catchment area of a top-initiated large-R jet.

represented in Figure 5.1. The usage of observables with high resilience to such uncorrelated and undesirable soft radiation becomes mandatory.

The development of a theory benchmark to describe observables with kinematic sensitivity is extremely challenging. In fact, the only prediction currently available involves boosted tops, and will be presented in the next section. The direct measurements, performed typically with non-boosted tops, rely on the prediction given by the MC simulations.

5.2 NLL factorization for boosted top quark jets

For the first time, the authors of Ref. [210] presented a first-principles description of the production and decay of top quarks in pp collisions, including the hadronization of the top decay products. It was based on the previous work of Ref. [124], where a factorization theorem to describe jet observables (2-jettiness) with high sensitivity to the top quark mass in $e^+e^- \to t\bar{t}$ processes was proposed. In the work of Ref. [210], the formalism derived enables the determination of a renormalized top quark mass from measurements of the jet mass M_J associated to those jets initiated by hadronically-decaying top quarks at particle level. The Pythia v8.2 MC generator was used to simulate the jet mass distribution and

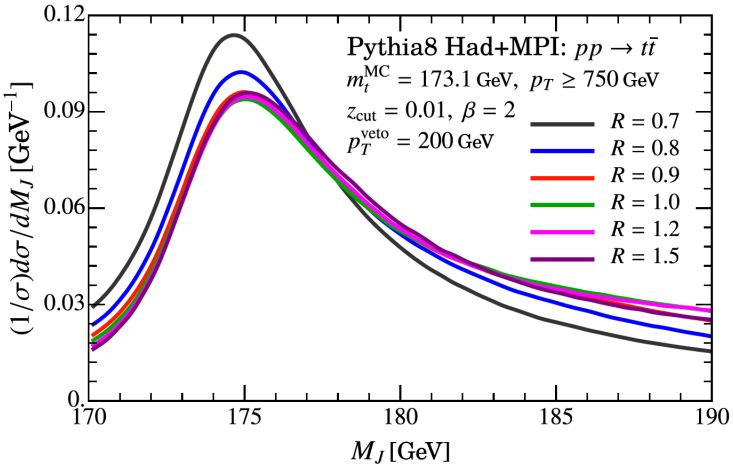


Figure 5.2: Dependence of the jet mass distribution with the jet radius R. Figure from Ref. [210].

probe the theory prediction proposed.

A key assumption of this model is the inclusive treatment of the top quark decays. This becomes possible by considering boosted tops that decay entirely in a big jet. The dependence of the simulated jet mass spectrum with the jet size is displayed in Figure 5.2. We see that a jet radius of $R \simeq 0.9$ is enough to capture the decay products of boosted tops, since the jet mass shape does not change significantly for larger values of R. Thus, the jet radius $R = 1$ is taken as reference. This boosted regime is accesible by requiring the top-initiated jets to have a transverse momentum p_{T} above 750 GeV.

The effects of uncorrelated soft emissions and hadronization are reduced by applying the SD grooming technique (see Section 4.3). It is required to be light to maintain the inclusive treatment and to retain the kinematic sensitivity to m_t generated in the soft sector. The light-grooming conditions take the form:

$$\begin{aligned} z_{\mathrm{cut}} &\lesssim \frac{\Gamma_t}{h^{2+\beta} m_t} \left(\frac{p_{\mathrm{T}}}{m_t}\right)^{\beta} , \\ z_{\mathrm{cut}}^{\frac{1}{2+\beta}} &\gg \frac{1}{2} \left(\frac{\Gamma_t m_t}{p_{\mathrm{T}}^2}\right)^{\frac{1}{2+\beta}} , \end{aligned} \tag{5.2}$$

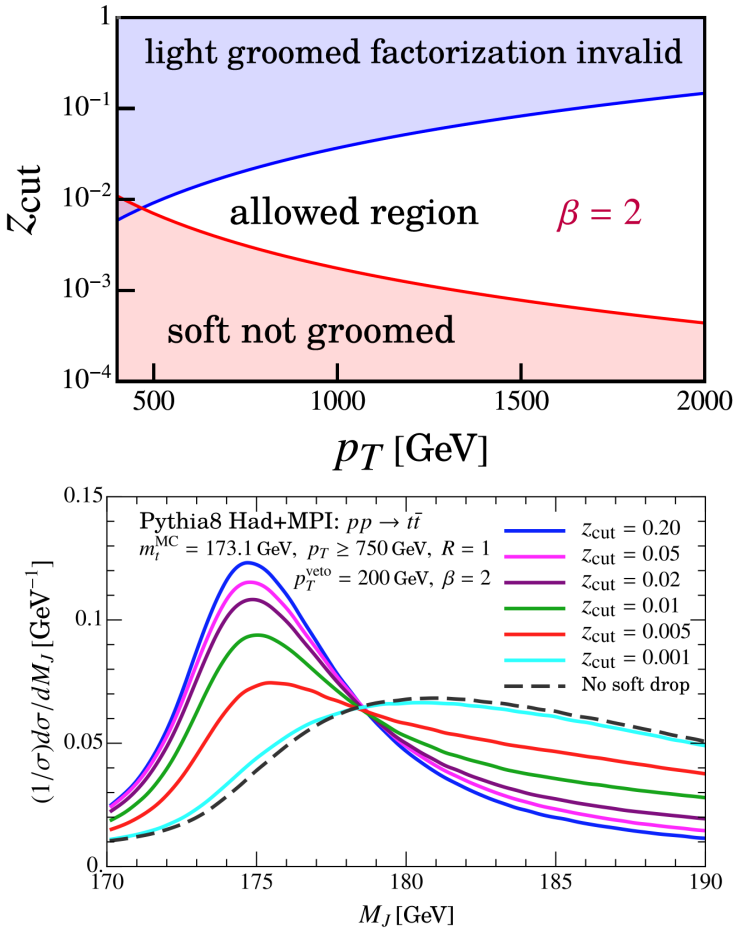


Figure 5.3: Upper panel: allowed region for $z_{\rm cut}$ values as a function of the jet transverse momentum $p_{\rm T}$. Lower panel: dependence of the jet mass distribution with $z_{\rm cut}$ for $\beta = 2$ and jet $p_{\rm T} > 750$ GeV. Figures from Ref. [210].

where the dimensionless function h relates the angles of the top decay products with respect to the jet axis. The allowed $z_{\rm cut}$ values as a function of the jet $p_{\rm T}$ for $\beta = 2$ are shown in the upper panel of Figure 5.3, and their impact on the simulated jet mass distribution is displayed in the lower panel. We see that for $z_{\rm cut} \gtrsim 0.005$ the light grooming is enough to restore the position of the jet mass peak, which remains approximately stable for higher $z_{\rm cut}$ values as expected (harder groomings will not remove physics effects at scales $\gtrsim m_t$). However for $z_{\rm cut} \gtrsim 0.05$ (which invalidates the light-grooming factorization), soft radiation from the top quark is groomed away and yields a narrower jet mass distribution.

Such a complex system involving top production, fragmentation and decay

with radiation at different scales and including hadronization effects has been only described in terms of several EFTs, as depicted in Figure 5.4. The energetic, collinear and soft radiation emitted by the top quark and affected by SD is treated with the SCET and encoded in a soft-collinear function S_c. The top quark fragmentation and decay is described by the *boosted Heavy Quark Effective Theory* (bHQET) and embodied in the jet function J_B, which also accounts for *Ultra-Collinear* (UC) radiation, unaffected by SD. As mentioned above, the top decay is inclusively incorporated with Γ_t and, at tree-level, J_B is just a *Breit-Wigner* (BW) function peaked at the top mass, carrying in this way the main mass sensitivity. The factorization theorem ties all them together. The groomed jet mass differential cross-section, which can be schematically written as:

$$\frac{d\sigma}{dM_J} \sim N(\mu_R, \Phi_J, z_{\text{cut}}, \beta) \otimes J_B(\mu_R, \Gamma_t, ...) \otimes S_c(\mu_R, \beta, ...) \otimes \int dk F_{\text{np}}(k) \ , \quad (5.3)$$

resums at NLL accuracy large logarithms in the partonic cross-section. The normalization factor $N(\mu_R, \Phi_J, z_{\text{cut}}, \beta)$ incorporates the underlying hard process, the PDFs and the radiation groomed away by SD. That is why it depends on z_{cut}, β and some jet kinematic variables ($\Phi_J = p_{\text{T}}, \eta_J$) together with the renormalization scale μ_R.

The effects of hadronization on the groomed jet mass differential cross-section are taken into account in Eq. 5.3 in the form of a shape function F_{np}. Given the nature of the soft radiation and the complex interplay with the SD grooming, the hadronization corrections depend on the kinematic jet variables, the grooming parameters and the kinematic phase space of the top decay products. Intuitively one can see that, if the hadronization corrections originate in the soft sector, they are fully enhanced in ungroomed jets and partially suppressed in groomed jets. Formally, the hadronization corrections are incorporated as the product of two

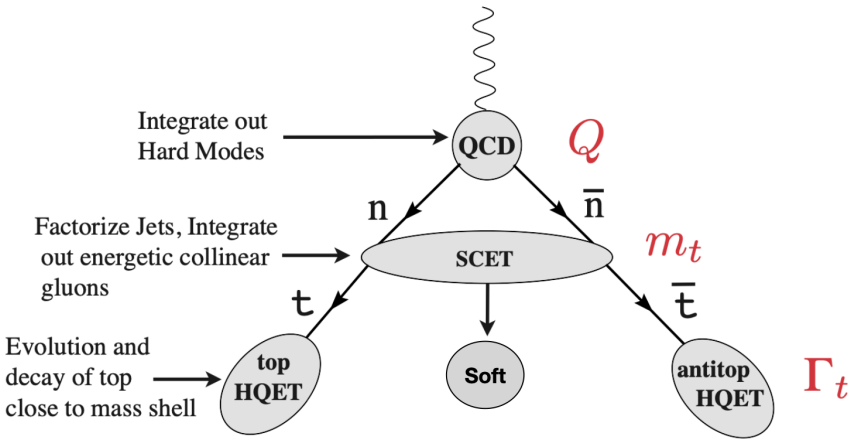


Figure 5.4: Sequence of effective field theories participating in the computation of the invariant jet mass distribution. Modified figure from Ref. [211].

terms: a perturbative coefficient calculated at LL accuracy that carries the whole kinematic dependence $(\Phi_J, z_{\text{cut}}, \beta)$ and a non-perturbative parameter that fixes the shape function F_{np}. The leading power non-perturbative corrections are given by the first moment $\Omega_{1q}^{\circ\circ}$ $(n = 1)$ of the shape function:

$$\Omega_{nq}^{\circ\circ} \equiv \int_0^\infty dk k^n F_{\text{np}}^q(k), \quad x_2 \equiv \frac{\Omega_{2q}^{\circ\circ}}{(\Omega_{1q}^{\circ\circ})^2} - 1 \ . \tag{5.4}$$

The x_2 parameter, which encodes higher-order power corrections, is also needed to fix the form of F_{np}. Interestingly, the same shape function appears in the e^+e^- analysis to describe jets initiated by light quarks. In this sense the dominant hadronization effects are universal, and do not depend on the grooming configuration or jet kinematics unlike the associated perturbative coefficients.

The predicted jet mass distribution can be directly compared to data or used to calibrate the top MC mass, as stated in Ref [210]. In the absence of auxiliary calculations that help to constraint the two hadronization parameters, the three free parameters of the theory, m_t, $\Omega_{1q}^{\circ\circ}$ and x_2 must be extracted simultaneously. m_t determines the position of the jet mass peak, further shifted and broadened by $\Omega_{1q}^{\circ\circ}$ and x_2. The degeneracy between them can be resolved following the example

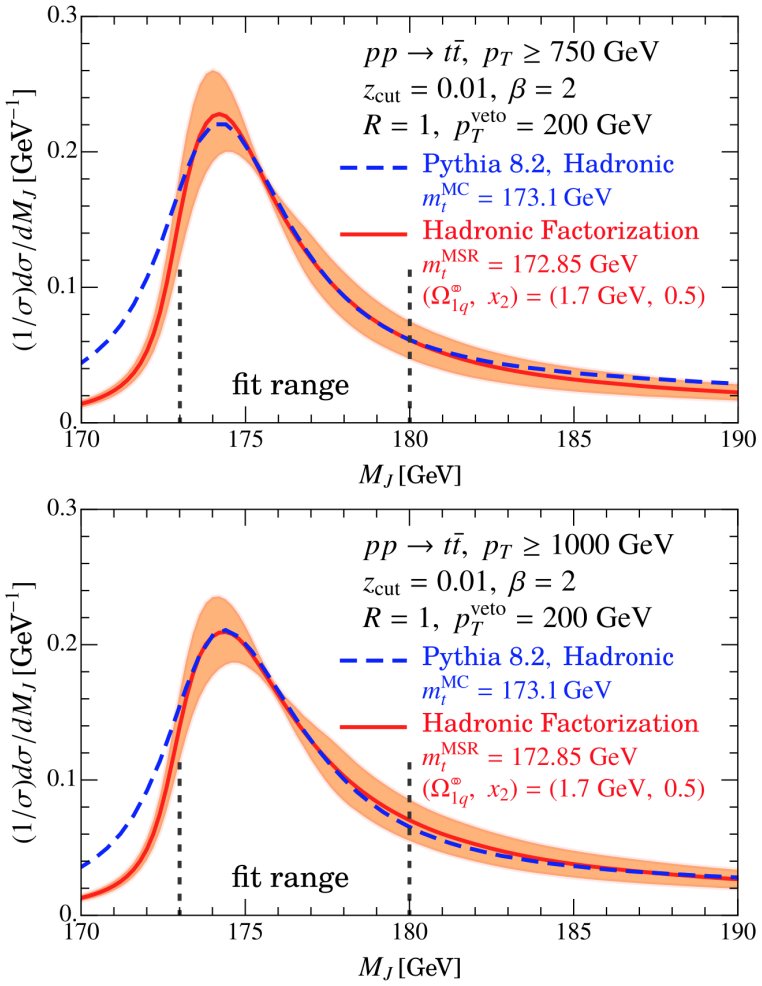


Figure 5.5: Comparison of the simulated jet mass distribution with PYTHIA 8 (in blue) to the NLL prediction (in red) for two different $p_{\rm T}$ ranges: $p_{\rm T} > 750$ GeV on the upper panel and $p_{\rm T} > 1000$ GeV on the lower panel. Figures from Ref. [210].

of Ref. [210]: as the perturbative parameters of the hadronization corrections have a kinematic jet dependence, such degeneracy can be broken by a simultaneous fit to jet mass distributions in different $p_{\rm T}$ intervals. An example of this procedure is shown in Figure 5.5 where the jet mass distributions generated with PYTHIA 8, setting $m_t^{\rm MC} = 173.1$ GeV, are compared to NLL predictions in two $p_{\rm T}$ ranges for lightly-groomed jets with $z_{\rm cut} = 0.01$ and $\beta = 2$.

The normalization of the cross-section is rather poorly determined by predictions at NLL. A normalized shape in the fit range is more robust against higher-order corrections, and hence should be used in comparing with data or

MC. The theory uncertainty due to missing higher-order corrections, represented by the orange band, can be obtained as the envelope of alternative jet mass distributions normalized in the fit range where the relevant scales involved are systematically varied up and down by a factor 2.

The noticeable difference between the NLL prediction and the MC simulation in the left tail of the peak is caused by radiation emitted off of the primary top decay products that is groomed away in the simulation. This produces a migration of jets above the peak to lower jet mass values, enhancing the left tail. The inclusive treatment given in the NLL prediction prevents from this situation. Fortunately, such effect has a very small impact in the peak region, where a good agreement between simulation and theory holds. The fit range is hence limited to the peak region, $M_J \in [173, 180]$ GeV denoted by the vertical dashed lines.

The hadronic factorization performs better as the jet p_{T} increases. The lower limit given by $p_{\mathrm{T}} > 750$ GeV guarantees a valid NLL description though, and at the same time provides an accesible threshold for the experimental analysis, where the available statistics decreases rapidly with the jet p_{T}. The theory-to-MC comparison can be made simultaneously is several p_{T} windows and for two renormalization mass schemes: the pole and the MSR. For the MSR mass at $R = 1$ GeV, the m_t value that best describes MC is found to be $m_t^{\mathrm{MSR}}(R = 1 \text{ GeV}) = 172.85$ GeV, 250 MeV bellow the value of m_t^{MC} used in the simulation.

5.3 ATLAS Monte Carlo generators

The relation between the top MC mass and a renormalized top mass is investigated for several ATLAS official MC samples at particle level. In this study, we decided not to include UE effects in the simulation because they cannot be properly handled by the theory. For this reason, a set of dedicated MC samples was generated with the setting that regules such effects disabled, referred to as MPI setting or model in the following. The MC samples described in the following, although they do not take into account UE effects, they do start from the same hard-scatter calculation as the standard ATLAS MC samples used for top quark mass measurements involving $t\bar{t}$ production.

The nominal ATLAS MC sample for top quark pair production is generated with the NLO ME event generator POWHEG BOX 2 [73, 74, 205] with the NNPDF3.0 NLO PDF set [212]. The factorization and renormalization scales in the ME calculation are set to $\mu_F = \mu_R = \sqrt{m_t^2 + (p_{\mathrm{T},t}^2)}$, where the top quark p_{T} is evaluated before it is allowed to emit radiation [213]. The ME generator is interfaced to PYTHIA (v8.210) [78], and uses the A14 set of tuned parameters [207] for the PS, hadronization and MPI models. The EvtGen afterburner program [85] handles decays of B and C hadrons. The top quark mass parameter is set to 172.5 GeV. The MC scheme described here will be referred to as the 'nominal' sample in the following.

Several samples can be directly obtained by internal reweighting of the nominal MC sample. This allows to consider variations of ±0.001 in the default value of the strong coupling $\alpha_s = 0.118$. Independent variations of a factor 0.5 and 2 in the factorization and normalization scales of the ME calculation are accessible as well. The baseline PDF set can be also replaced by the PDF4LHC15 set [214], which actually contains the statistical combination of three independent PDF

Table 5.1: List of Monte Carlo samples obtained by internal reweighting of the nominal. f_R and f_F represent the multiplicative factors that change the default values of the renormalization and factorization scales in the Matrix Element calculation, $\mu_{R(F)} = f_{R(F)} \cdot \mu^0_{R(F)}$. For the initial- and final-state radiation samples, they are relative to the renormalization and factorization scales that regulate the QCD emissions in the corresponding subprocess.

Sample	PDF set	α_s^{ME}	f_R	f_F	Comments
Nominal	NNPDF23	0.118	1.0	1.0	-
Alt. PDFs	PDF4LHC15	0.118	1.0	1.0	Includes 30 variations
α_s-Up	NNPDF23	0.119	1.0	1.0	Through PDF set
α_s-Dw	NNPDF23	0.117	1.0	1.0	Through PDF set
μ_R-Up	NNPDF23	0.118	2.0	1.0	ME scale
μ_R-Dw	NNPDF23	0.118	0.5	1.0	ME scale
μ_F-Up	NNPDF23	0.118	1.0	2.0	ME scale
μ_F-Dw	NNPDF23	0.118	1.0	0.5	ME scale
ISR, RadHigh	NNPDF23	0.118	0.5	0.5	$h_{\mathrm{damp}} = 3m_t$ and Var3c-Up
ISR, RadLow	NNPDF23	0.118	2.0	2.0	Var3c-Dw
FSR, RadHigh	NNPDF23	0.118	0.5	0.5	-
FSR, RadLow	NNPDF23	0.118	2.0	2.0	-

sets - CT14, MMHT2014 and NNPDF3.0. A total of 30 variations are included in the PDF4LHC15 recommendations, accounting for intra-PDF and inter-PDF uncertainties. The effects of ISR/FSR can be also investigated through internal reweighting. Samples with increased and decreased amount of ISR can be derived by varying the renormalization and factorization scales for QCD emissions in the ISR and the A14 Var3c eigentune. Analogously, for variations in the final-state, the renormalization and factorization scales in the FSR are modified instead. These samples are listed in Table 5.1.

A set of varied samples based on this nominal ATLAS configuration is studied in order to factorize the role played by the internal settings of Pythia 8 in shaping the jet mass distribution. Such variations affect:

- the A14 tune: many MC event generators use PS, hadronization, and MPI models which include parameters whose values may be fixed using fits to experimental measurements. This is done to more accurately generate

events that model collider data. This optimization process is known as 'tuning,' and the resulting set of parameter values are referred to as 'MC tunes.' The nominal Pythia 8 configuration in ATLAS makes use of the A14 tune [207], which was developed from ATLAS measurements of jet substructure and UE observables at $\sqrt{s}$ = 7 and 8 TeV, and is nominally provided for use with the NNPDF23LO PDF set. The fitted parameters in this tune are related to CR (explained in more details in Section 5.5.3), MPI and ISR/FSR and can be grouped into sub-sets called eigentunes, labelled as Var1, Var2, Var3a, Var3b and Var3c. They are varied systematically in order to provide a good coverage of the experimental and modelling uncertainties implicit in the tuning. The Var1 eigentune is devoted to the modelling of UE effects, whilst the others are related to ISR/FSR. The impact of these tune variations on the shape of the jet mass distribution is evaluated.

- the h_{damp} parameter. The value of h_{damp} for these studies was chosen to be equal to $(3/2)m_t$. Events with $h_{\text{damp}} = 3m_t$ are also studied.

- the *Matrix Element Correction* (MEC), related to the matrix-element-to-parton-shower matching, is switched off in a dedicated sample [215]. This disables MEC to the first emission in the PS.

- the *Recoil-To-Colored* (RTC) setting: the way recoils to colored objects and color singlets are treated in the PS may affect the jet mass distribution. To study this effect, two MC samples that differ only in the choice of the "recoil-to-colored" switch in Pythia 8 are considered. By default, this switch is set to "ON" in Pythia 8, but recent studies show that "OFF" may be an equally plausible choice [216].

- the value of the *rFacB* parameter, which controls the longitudinal momentum sharing of B hadrons in the string-based hadronization model used in Pythia 8. Lower values of rFacB give rise to softer B spectra. For this study it is varied from its default value of 0.65 to 1.05, replicating the strategy adopted in the ATLAS $t\bar{t}$ modelling uncertainties prescription [215].

- the EvtGen generator, responsible for handling the decay of B and D hadrons. By default, Pythia 8 is interfaced with EvtGen, so a dedicated sample without EvtGen is considered as well.

Alternative $t\bar{t}$ simulations are used to assess variations in the definition of the top mass among MC programmes. The Powheg Box generator interfaced to Herwig (v7.1.3) [217] with the H7UE tune provides alternative models for the PS, UE and hadronization. The MadGraph5_aMC@NLO (v2.3.3.p1) generator [71, 72] provides an alternative ME calculation. This ME calculation is interfaced to Pythia 8 with the same settings as the nominal Powheg sample, and the same renormalization and factorization scales are also used. These alternatives samples and those enumerated before are listed in Table 5.2.

Finally, MC simulations with the MPI model enabled are also considered to check the impact of the UE effects in the jet mass distribution. To this end, the nominal MC sample is studied together with variations of the default MPI model. This includes variations of the A14 eigentune Var1 and alternative CR models available in Pythia 8.

Table 5.2: List of dedicated Monte Carlo samples with the parameters of interest.

Sample	ME+PS	MEC	h_{damp}	RTC	bFrag	EvtGen	Tune
Nominal	Ph+Pyt8	✓	$(3/2)m_t$	✓	0.65	✓	A14
MEC off	Ph+Pyt8	✗	$(3/2)m_t$	✓	0.65	✓	A14
h_{damp}	Ph+Pyt8	✓	$3m_t$	✓	0.65	✓	A14
RTC off	Ph+Pyt8	✓	$(3/2)m_t$	✗	0.65	✓	A14
rFacB	Ph+Pyt8	✓	$(3/2)m_t$	✓	1.05	✓	A14
EvtGen off	Ph+Pyt8	✓	$(3/2)m_t$	✓	0.65	✗	A14
Var2 Up	Ph+Pyt8	✓	$(3/2)m_t$	✓	0.65	✓	A14/Var2 Up
Var2 Dw	Ph+Pyt8	✓	$(3/2)m_t$	✓	0.65	✓	A14/Var2 Dw
Var3a Up	Ph+Pyt8	✓	$(3/2)m_t$	✓	0.65	✓	A14/Var3a Up
Var3a Dw	Ph+Pyt8	✓	$(3/2)m_t$	✓	0.65	✓	A14/Var3a Dw
Var3b Up	Ph+Pyt8	✓	$(3/2)m_t$	✓	0.65	✓	A14/Var3b Up
Var3b Dw	Ph+Pyt8	✓	$(3/2)m_t$	✓	0.65	✓	A14/Var3b Dw
Var3c Up	Ph+Pyt8	✓	$(3/2)m_t$	✓	0.65	✓	A14/Var3c Up
Var3c Dw	Ph+Pyt8	✓	$(3/2)m_t$	✓	0.65	✓	A14/Var3c Dw
Alt. ME	aMc+Pyt8	✗	$(3/2)m_t$	✓	0.65	✓	A14
Alt. PS	Ph+H7	✓	$(3/2)m_t$	✓	0.65	✓	H7UE

5.4 Methodology

5.4.1 Generation of theoretical predictions and simulation

The theoretical predictions are generated in a fine grid of values of the three free parameters of the theory for the nominal light-grooming configuration ($\beta = 2$ and $z_{\text{cut}} = 0.01$) with the nominal scale set: m_t between 171.0 GeV and 174.0 GeV with 0.05 GeV steps, $\Omega_{1q}^{\circ\circ}$ between 0.1 GeV and 4.0 GeV in 0.1 GeV steps, and x_2 between 0.02 and 1.0 with steps of 0.02. This is done for two renormalization mass schemes: the pole and MSR, taking the latter as the reference prediction in the following. The scale of the top quark mass in the MSR scheme is set to $R = 1$ GeV, according to the discussion of the interpretation of the top MC mass held in Section 2.4. Templates are produced for each p_{T} range considered: 750 GeV $< p_{\text{T}} <$ 1 TeV, 1 TeV $< p_{\text{T}} <$ 1.5 TeV, 1.5 TeV $< p_{\text{T}} <$ 2 TeV and 2 TeV $< p_{\text{T}} <$ 2.5 TeV. This amounts a total of approximately 5 million templates only for the standard configuration.

In order to evaluate the theory uncertainty associated to the scale set used, theory predictions are obtained where the different scales involved are varied according to the prescription adopted in Ref. [210], further discussed in Section 5.5.1. In addition, the code that implements the calculation allows the configuration of various levels of SD grooming, as long as the requirements of light grooming are satisfied. We consider three more configurations: $\{\beta = 2, z_{\text{cut}} = 0.005\}$, $\{\beta = 2, z_{\text{cut}} = 0.02\}$ and $\{\beta = 1, z_{\text{cut}} = 0.01\}$.

Concerning the production of the MC events, we distinguish between two scenarios. In the first scenario, we make use of the ATLAS machinery at our disposal to reprocess the already existing NLO ME simulations, stored in the so-called *Les Houche Events* (LHE). Around 1 billion of LHE events have been generated for $t\bar{t}$ production at NLO with the POWHEG BOX generator, as explained previously. These events store the basic information of the hard-scatter outcomes, such as partons and their four-momenta. Only a small fraction of these events contains tops or antitops with a transverse momentum sufficiently large to populate the phase space where the theory validity holds. The identification and selection of potentially interesting boosted events after the showering and hadronization, as it is usually done, is very suboptimal. In order to save a significant amount of time and CPU resources (and so turning feasible this MC production), a dedicated LHE filter was developed where only events containing tops and antitops with a p_{T} greater than 700 GeV were interfaced with the next MC program. Only the 0.029% of events were selected by the LHE filter, and almost 90% of events after fragmentation and hadronization contained a jet with p_{T} above the selection threshold.

In the second scenario, we generated events with our local implementation of the PYTHIA 8 MC generator, following the setup of the version used in ATLAS. We will refer to this MC production as PYTHIA 8 standalone throughout this

chapter. Firstly, this allowed us to modify many aspects of the MC simulation and study how the jet mass lineshape is affected, covered in Section 5.4.3. Secondly, samples with varied values of the m_t^{MC} (ranging from 171.5 GeV to 173.5 GeV) were produced in order to study the linearity of the calibration presented here. The agreement of the jet mass distributions obtained from the ATLAS MC samples and Pythia 8 standalone is assessed in Section 5.4.5. Thirdly, the LO ME generator of Pythia 8 permits the production of $t\bar{t}$ pairs in given phase space by setting the maximum and minimum invariant p_{T} of the system, enabling in this way a MC calibration at any arbitrary p_{T} scale.

5.4.2 Event selection and jet reconstruction

The jet mass distribution from MC simulation is obtained for particle-level jets by clustering all stable final-state particles produced by the generators (equivalent to Pythia 8 status code 1). By default, these final-state particles are produced by the hadronization algorithm of a MC generator. By disabling the hadronization algorithms (*e.g.* by disabling the 'HadronLevel:all' switch in Pythia 8), the set of final-state particles instead corresponds to a set of partons produced in the hard-scatter and PS; this parton-level picture is used in some studies presented in Sections 5.4.3 and 5.4.4.

The nominal jet reconstruction algorithm used is the XCone algorithm with $\beta = 2$, $\gamma = 2$ and radius parameter $R = 1$, as implemented in FastJet [218]. The anti-k_t clustering algorithm is also studied, with the same radius parameter. The SD grooming with parameter values $\beta = 0, 1$ or 2 and $z_{\mathrm{cut}} = 0.01$ or 0.05 is applied in order to remove soft- and wide-angled contributions to the jet.

A simple event selection is applied to the MC simulated events. Events are required to contain at least one large-R jet with a p_{T} above 750 GeV. This

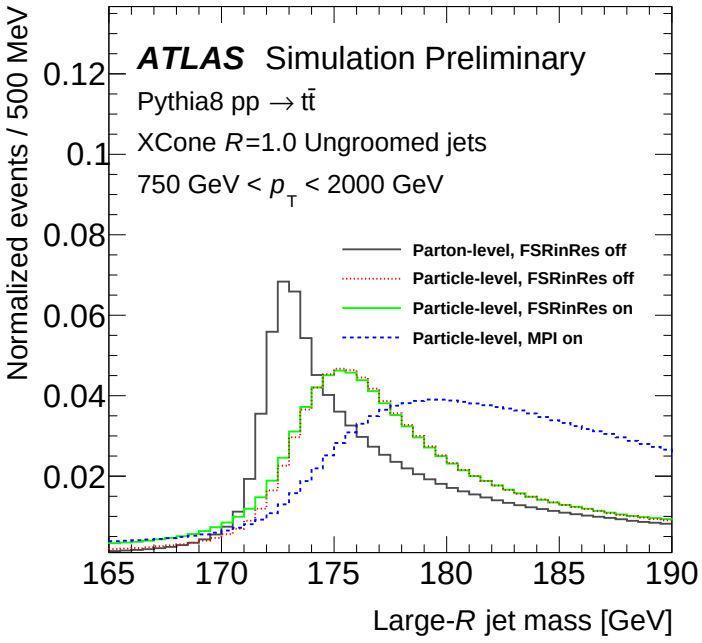
jet is matched to the top/anti-top parton after emitting FSR, but before the top decays, by requiring ΔR(jet,top) < 1.0. In order to avoid pathological configurations where two large-R jets overlap (which cannot be compared to the theory prediction) the leading and subleading large-R jets must be separated by a distance of $\Delta R > 1.0$.

5.4.3 The jet mass in high-p_{T} top decays

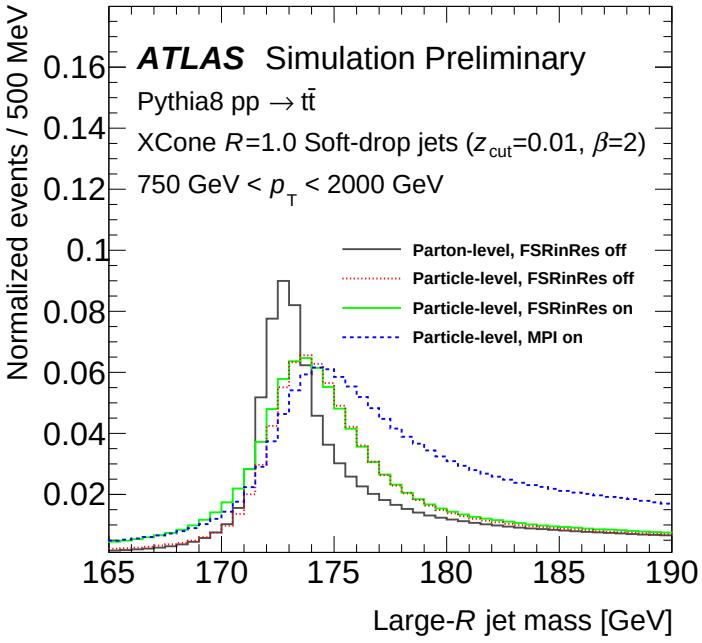
The evolution of the jet mass lineshape from parton level to particle level is considered in this section by sequentially enabling aspects of the Pythia 8 simulation, including FSR from the top quark decay products, the PS, hadronization and the UE.

Figure 5.6 shows the jet mass distribution for XCone jets with R = 1 and ungroomed jet p_{T} between 0.75 – 2 TeV, produced with Pythia 8 standalone and a top MC mass of m_t^{MC} = 172.5 GeV. The results in Figure 5.6(a) correspond to ungroomed jets, and those in Figure 5.6(b) to jets groomed with the SD algorithm configured using the light-grooming settings, β = 2 and z_{cut} = 0.01.

A comparison of the four histograms in each figure shows how the different stages of the MC generator transform the jet mass distribution. The histogram represented by the black line shows the jet mass distribution at parton level after fragmentation, obtained when the MPI model is switched off. The FSR from top decay products is disabled here using the Pythia 8 switch "FSRinResonances" (we will use the notation "FSRinRes off" as a shorthand), which allows FSR from the top quark itself, but prevents top quark decay products from radiating. Hadronization, as well as radiation off of the top quarks and emissions in the PS cause a significant tail at large values of the jet mass. A comparison of Figure 5.6(a) and Figure 5.6(b) shows that grooming reduces sensitivity to



(a)



(b)

Figure 5.6: Large-R jet mass distributions obtained from the reference PYTHIA 8 setup for (a) ungroomed XCone jets with $R = 1$ and (b) the same jets after applying a 'light' soft-drop grooming ($\beta = 2$, $z_{\text{cut}} = 0.01$). Monte Carlo simulations at parton and particle level are shown, where the labels FSR and MPI respectively indicate whether final-state radiation in resonances and multi-parton interactions (*i.e.* underlying event) are activated in the simulation.

hadronization and the UE, improving the top jet mass resolution. Even with the light-grooming settings used here, the high-mass tail is reduced significantly.

The effect of hadronization becomes clear by comparing such configuration at parton and particle level. In the ungroomed case, the distribution is profoundly altered: the top mass peak is smeared and shifted by more than 3.5 GeV. Light grooming reduces the impact of the hadronization, limiting the shift of the peak to approximately 1.5 GeV.

In the case where the FSR in resonance decays is switched on, while the MPI is still off (labelled 'FSRinRes on' in the figure), this effect leaves the peak position unaltered, but has an effect on the low-mass tail, as wide-angle emissions from the top decay products can either escape the catchment area of the jet or be removed by grooming. This effect is therefore most clear for groomed jets, which have an increased sensitivity to the top decay products and their radiation.

When the MPI model is turned on ('MPI on'), the additional radiation that falls on the jet catchment area has the effect of further broadening the distribution and lifting up the high-mass tail. The impact is most pronounced for the ungroomed jets, where the top mass peak is smeared out over a broad mass range between 175 GeV and 195 GeV. Grooming effectively mitigates the impact of UE on the jet mass distribution. The shift of the peak position in the particle-level result when the MPI model is toggled on/off is less than 0.5 GeV.

5.4.4 Template-fitting procedure

The fit strategy followed in this work is based on a template-fit, where data (in this case, MC simulations) is compared to a set of templates generated by varying the values of the parameters they depend on. A $\chi 2$ fit and minimization is carried out to find the values of such parameters that best describe the MC simulation.

For each template, the $\chi 2$ is calculated using the normalized number of events N of the MC and the template, and the MC statistical uncertainty:

$$\chi 2 = \sum_k \left(\frac{N_k^{\mathrm{MC}} - N_k^{\mathrm{tmp}}(\theta_1, ..., \theta_N)}{\sigma_k^{\mathrm{MC}}(\mathrm{stat})} \right)^2 , \tag{5.5}$$

where k runs over the histogram bins and $\theta_1, ..., \theta_N$ represents the N parameters the template depends on.

At this point, we perform two different fits depending on the template used. If the template corresponds to a NLL theory prediction, the fit yields a three-dimensional $\chi 2(m_t, \Omega_{1q}^{\circ\circ}, x_2)$. The total 3D-$\chi 2$ comes from a combination of fits performed on the large-R jet mass distribution in three bins of the ungroomed large-R jet p_{T}, following the approach taken in Ref. [210] in order to disentangle the non-perturbative parameter behaviour from the behaviour of m_t^{MSR}. The following three p_{T} intervals are considered in the determination of the MSR and MC mass relation: 750 GeV $<$ p_{T} $<$ 1 TeV, 1 TeV $<$ p_{T} $<$ 1.5 TeV and 1.5 TeV $<$ p_{T} $<$ 2 TeV.

To obtain the central value and uncertainty for each parameter, the total 3D-$\chi 2$ distribution is marginalized by scanning the values of the other parameters and finding the lowest $\chi 2$ value. This procedure is repeated for all values of the parameter of interest and the resulting one-dimensional marginalized $\chi 2$ distribution is fitted with a second-order polynomial function. The best estimate for a given parameter is taken as the value which minimizes fitted function, and its associated uncertainty is delimited by those values of the parameter that lead to an increase of the $\chi 2$ by 1 unit with respect to the minimum.

In the second scenario, we use MC-based templates as the reference prediction where the value of the top MC mass is varied. The resulting $\chi 2$ will depend only on a single parameter, the m_t^{MC}, so it is calculated in a single p_{T} bin:

0.75 TeV $< p_{\mathrm{T}} <$ 2 TeV. The best-fit value and the associated statistical uncertainty is extracted as explained above for marginalised 1D-$\chi 2$ distributions. The MC-based $\chi 2$ fit is crucial to study the effects induced by the MPI modelling in the jet mass distribution and the impact of several MC settings in those samples where the number of events available is not enough to populate the three p_{T} bins and carry on with the theory fit.

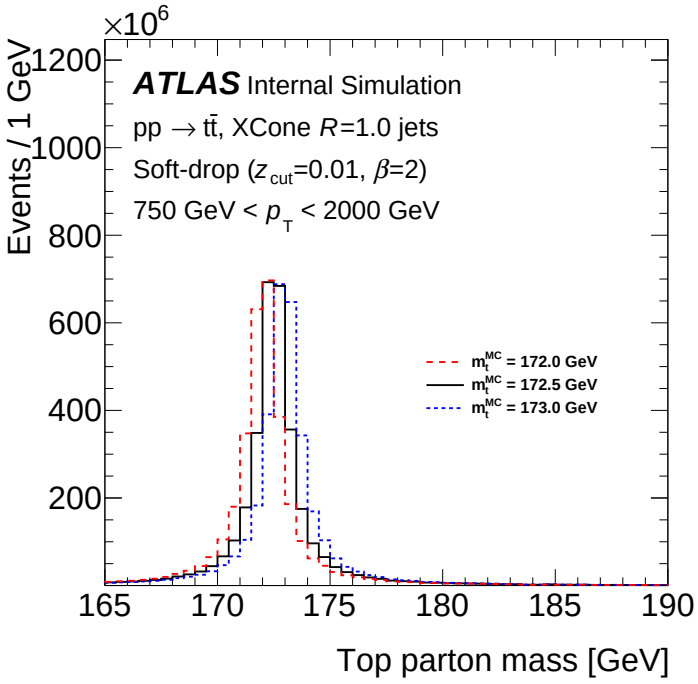
The MC-based templates are generated with a truth-level reweighting. The method assumes that the parton-level top quark mass distribution is well described by a Breit-Wigner curve as the one implemented in the following function:

$$\mathrm{BW}(m_t, \Gamma_t, x) = \frac{2\sqrt{2} m_t \Gamma_t \sqrt{m_t^2 (m_t^2 + \Gamma_t^2)}}{\pi \sqrt{m_t^2 + \sqrt{m_t^2 (m_t^2 + \Gamma_t^2)}}\ ((x^2 - m_t^2)^2 + m_t^2 \Gamma_t^2)} \ , \tag{5.6}$$

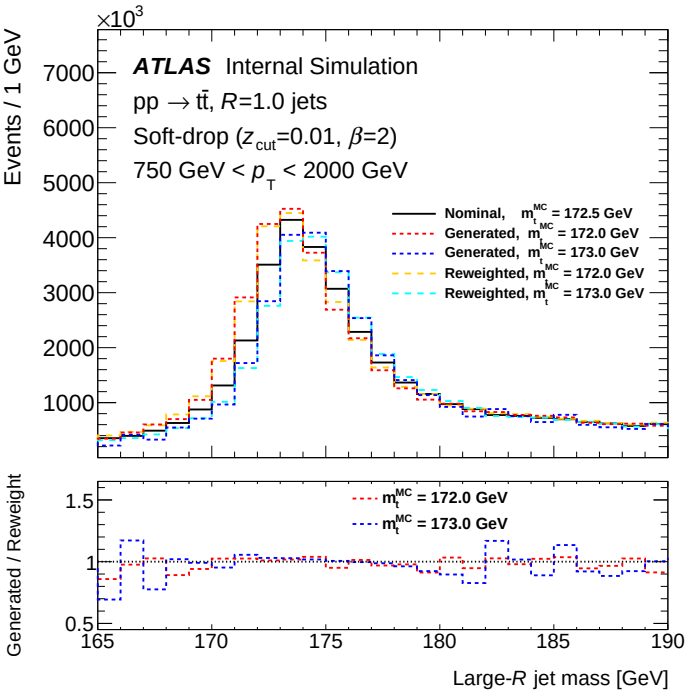
where x represents the top quark mass after radiating and before decaying (which varies event-by-event) and m_t corresponds to m_t^{MC} (fixed for all events in a given sample reweighting). Figure 5.7(a) shows the lineshape of BWs for several values of m_t^{MC}, which ratio is used in Figure 5.7(b) to correct a distribution generated with m_t^{MC} = 172.5 GeV as it was generated with an alternative values of 172.0 and 173.0 GeV. As can be observed, the reweighted distributions obtained from the reference mass are in good agreement with those generated with the corresponding mass (172.0 and 173.0 GeV).

5.4.5 Theory and simulation compatibility

This section is devoted to address the question of the agreement between the best-fit theory calculation and the MC simulation. As already introduced at the end of Section 5.2, there might be some sources of discrepancy between the theory prediction and the MC simulation that must be carefully investigated to



(a)



(b)

Figure 5.7: Breit-Wigner (a) derived for a nominal (172.5 GeV) and alternative (172.0 and 173.0 GeV) Monte Carlo mass values and (b) the corresponding nominal, generated and reweighted truth jet mass distributions.

fairly compare the two. To carry this out, results of the PYTHIA 8 standalone simulation are considered along with the NLL calculation. The top quark mass parameter is set to m_t^{MC} = 172.5 GeV in the MC generator, and the range of groomed large-R jet masses included in these comparisons is always taken to be between 170 – 180 GeV.

In Figure 5.8(a), the comparison is performed at parton level, *i.e.* without FSR in resonance decays, hadronization or MPI in PYTHIA 8, represented with the black histogram. The parton-level NLL calculation, represented with the dashed line, has the top MSR mass as the only free parameter. This parton-level MSR mass is fitted with a $\chi 2$ minimization, yielding a value of m_t^{MSR} = 172.75 GeV. The results are in good agreement; the theory prediction lies on top of the MC generator result over a wide mass range around the peak. Any deviations in the shape are well within the theory uncertainties, represented by the gray band and calculated as explained in Section 5.5.1.

In Figure 5.8(b) the results of the NLL calculation and the MC generator are compared at particle level, but with FSR in resonance decays still disabled. The effect of the hadronization model on the MC generator is to shift and smear the top mass peak. The NLL calculation includes the effect of hadronization in the form of a shape function with two parameters, as described in Section 5.5.1. The top mass and the two additional degrees of freedom are determined in a three-dimensional fit. The best-fit curves again provide an adequate description of the MC generator prediction.

Let's allow now the top decay products to radiate by activating the FSRinRes switch in the MC generator. Some of the radiation is groomed away, leading to changes in the low-mass tail under the top mass peak, as discussed in Section 5.4.3. In Figure 5.8(c) the result is compared to the best-fit result of the calculation. Even with three free parameters, discrepancies arise between the

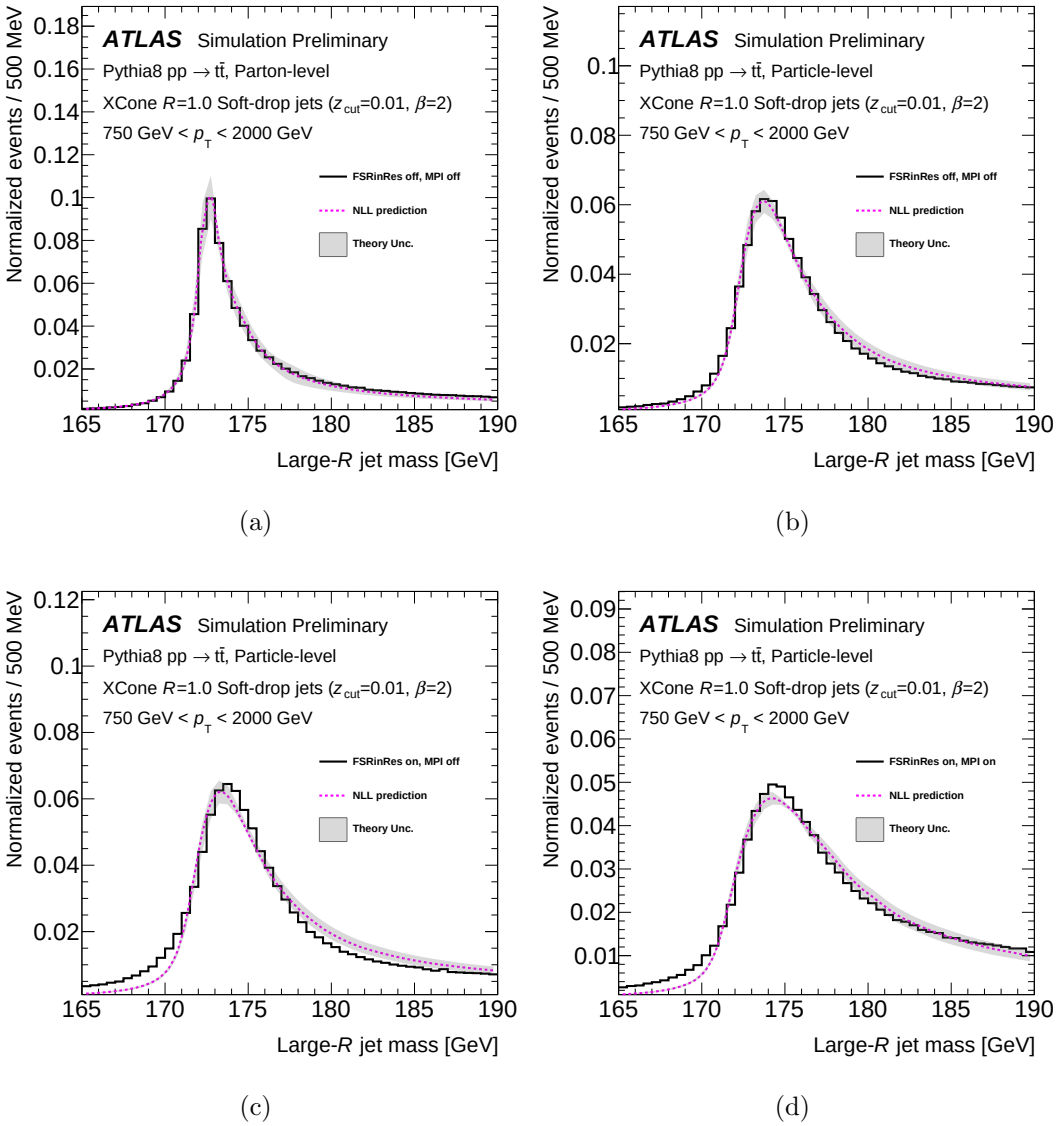


Figure 5.8: Comparison of the NLL prediction of Ref. [210] to several PYTHIA 8 predictions with m_t^{MC} = 172.5 GeV. The distributions correspond to (a) the Monte Carlo and NLL predictions at parton level and with final-state radiation in resonance decays (PYTHIA 8 setting 'FSRinRes') turned off, (b) the particle-level prediction with FSRinRes and underlying event modelling (PYTHIA 8 setting 'MPI') turned off, (c) the particle-level predictions with FSRinRes turned on and MPI switched off, and (d) the particle-level predictions with FSRinRes and MPI turned on. The distributions are normalized and the fit is performed on the interval 170 GeV $< M_J <$ 180 GeV, where M_J is the large-R jet mass. Fits are performed in three bins of the ungroomed large-R jet p_{T}; these figures are shown in a single p_{T} bin for illustrative purposes.

jet mass distributions of the theory and simulation which cause the $\chi 2$ value to increase. These differences arise because the theory treats the top decay products inclusively, so it does not allow for the possibility of radiation from the top decay products to be groomed during the SD procedure. As the relation between the MC mass and the MSR mass is determined with FSR in resonance decays switched on, a careful treatment of the low-mass tail is necessary.

Finally, UE effects in the MC simulation are added in Figure 5.8(d) by switching on the MPI setting. It leads to a more pronounced discrepancy between theory and simulation in the peak region, with the corresponding increase of the $\chi 2$ value. This impoverishment is specially visible when each of the three p_{T} intervals is represented separately. The theory fails to simultaneously describe the three regions consistently. The reason behind this is that the shape function devoted to encode hadronization effects is not originally designed to absorb the UE, so it cannot properly scale with the jet p_{T} and radius. The consequence is an artificial enhancement of the hadronization parameters $\Omega_{1q}^{\circ\circ}$ and x_2. The m_t^{MSR} is affected as well.

The situations described above affect the extracted value of the MSR mass. In those fits of Figure 5.8(a), m_t^{MSR} is largely independent from the mass region considered in the fit. That is no longer true when fitting with FSR in resonance decays and hadronization activated (Figure 5.8(c)). The calculation fails to describe the tail below the top mass peak that is present in the generator prediction. The discrepancy in the low-mass tail limits the theory prediction's range of validity. A fit that includes the low large-R jet mass tail will bias the extracted top mass to lower values. Therefore, the jet mass window where theory and simulation are compared needs to be carefully adjusted in order to provide a reliable relation between the mass parameter in the generator and the m_t^{MSR} in the calculation.

Figure 5.9 shows the dependence of the best-fit $\chi 2$ value on the choice of the lower limit of the fit range, together with the statistical uncertainty. When the fit range starts at very low masses, the discrepant low-mass tail is included, causing the $\chi 2$ value to increase. But if the fit range starts at too high values, the peak region is not included and the statistical uncertainty on the top mass parameter increases very rapidly. The lower limit of the fit range is therefore set to the lowest possible value in the $\chi 2$ plateau. In practice, this is achieved by identifying the lower limit of the fit range with the mass parameter of the MC generator. This choice ensures that the mass peak, which sits 1.5 GeV above the MC mass parameter, is properly included. Consequently, the large-R jet mass range used in the calibration fits is between 172.5 GeV – 180 GeV. An uncertainty associated with the choice of the large-R jet mass range used in the fit is discussed in Section 5.5.2.

5.5 Uncertainties

This section provides a description of the sources of uncertainty on the relation between the MC mass and the MSR mass, for a given observable and MC generator setup. Theory uncertainties are estimated in Section 5.5.1 and uncertainties related to the methodology in Section 5.5.2. The impact of UE is estimated separately, in Section 5.5.3.

5.5.1 Theory uncertainties

The uncertainty on the prediction due to missing higher-order corrections is estimated by varying several scales in the calculation. There are five scale parameters: the hard scale parameter e_{H} (related to the renormalization and

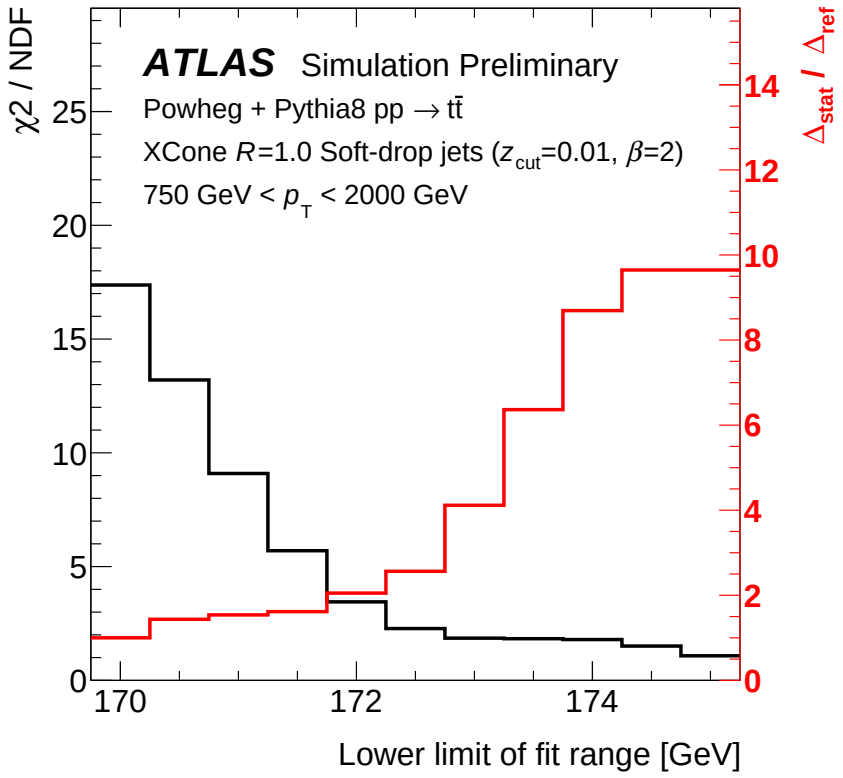


Figure 5.9: The $\chi 2$ value and the statistical uncertainty obtained from a fit with the NLL calculation of Ref. [210] to a Monte Carlo prediction at particle level, as a function of the lower limit of the fit. The Monte Carlo prediction is based on Pythia 8 with a top quark mass of 172.5 GeV. The multi-parton interaction setting is turned off, and the final-state radiation in resonances is turned on in the Monte Carlo.

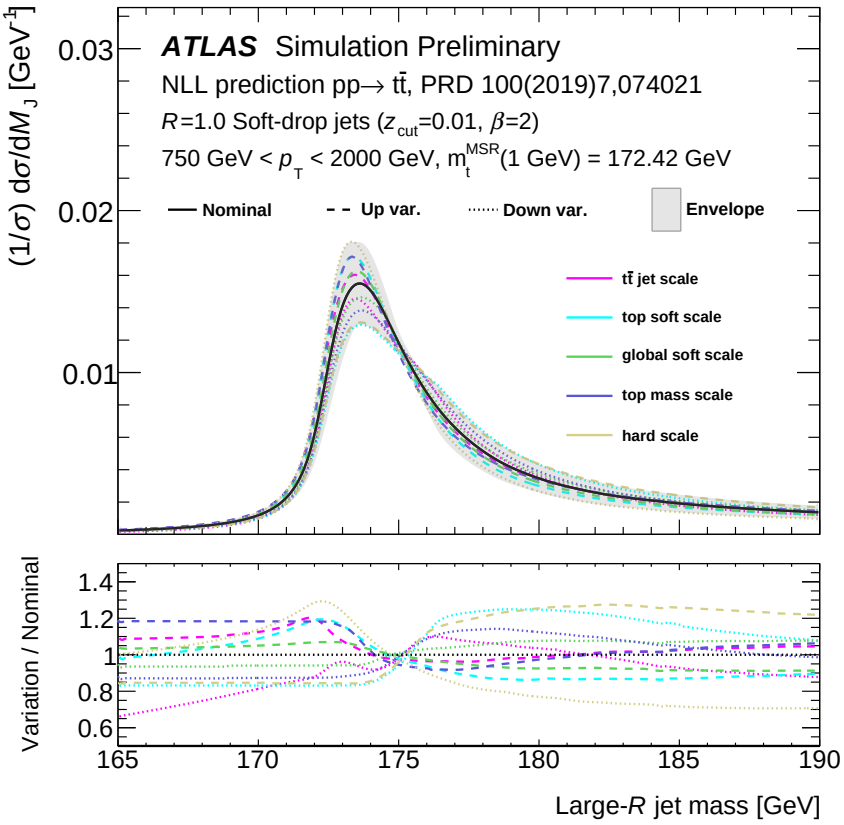


Figure 5.10: NLL prediction of Ref. [210] for the normalized particle-level jet mass distribution of large-R jets containing a boosted hadronic top quark decay (black line). The colored and dashed/dotted lines are obtained by varying the five scales in the calculation. The grey band corresponds to the envelope of all scale variations. All curves are normalized to the same area in the groomed large-R jet mass interval of 172.5 – 180 GeV.

factorization scales), the top mass scale parameter e_{m}, the general soft scale parameter e_{SG}, the top soft scale parameter e_{St} and the top-antitop jet scale parameter e_{Gt}. The five profile functions that govern these scales are varied by factors of 2 and 1/2, as in Ref. [210]. The three parameters m_t^{MSR}, $\Omega_{1q}^{\circ\circ}$ and x_2, are set to the best-fit values obtained in a fit to the nominal ATLAS MC prediction with Powheg + Pythia 8. Figure 5.10 shows the best-fit prediction and the ten scale variations considered.

The grid of theory NLL predictions generated with the default scales is then fit to each of the alternative predictions, with the three parameters freely

floating. In this way, the impact of the theory uncertainty on the mass relation is estimated. Figure 5.11 presents the fit results for the ten scale variations. The total uncertainty is taken as the difference of the fitted mass value and the nominal results. For most scale variations the MSR mass is shifted by less than ±200 MeV. The largest deviations come from variations of the soft and jet scale parameters. The total theory uncertainty is taken as the envelope of the ten scale variations. In the MSR scheme, it amounts to +230 MeV in the positive and −310 MeV in the negative direction. The same exercise is repeated in the on-shell scheme, leading to a slightly smaller variation of the mass: +150 MeV and −250 MeV. This is because the MSR mass dependence on the scale R allows it to assess a source of uncertainty that is not assessed by scale variations for the pole mass.

The size of the scale uncertainty is similar to the result reported for the calculation at NLL accuracy of the 2-jettiness in electron-positron collisions in Ref. [219]. The theory uncertainty is expected to decrease as the formal accuracy of the calculation increases in the future. Ref [219] observes an important reduction of the uncertainty from NLL to NNLL, and from NNLL to N^{3}LL. Assuming a similar convergence in the hadron collider case, the scale uncertainty could be reduced to approximately half of the current value with an NNLL calculation, and to a quarter at N^{3}LL.

5.5.2 Method uncertainties

The fit result depends on certain choices made in the fit method. The most important ones are the choice of the fit range and the kinematic ranges that are considered in the fit. Uncertainties are assigned to cover any potential bias of the mass relation due to these choices.

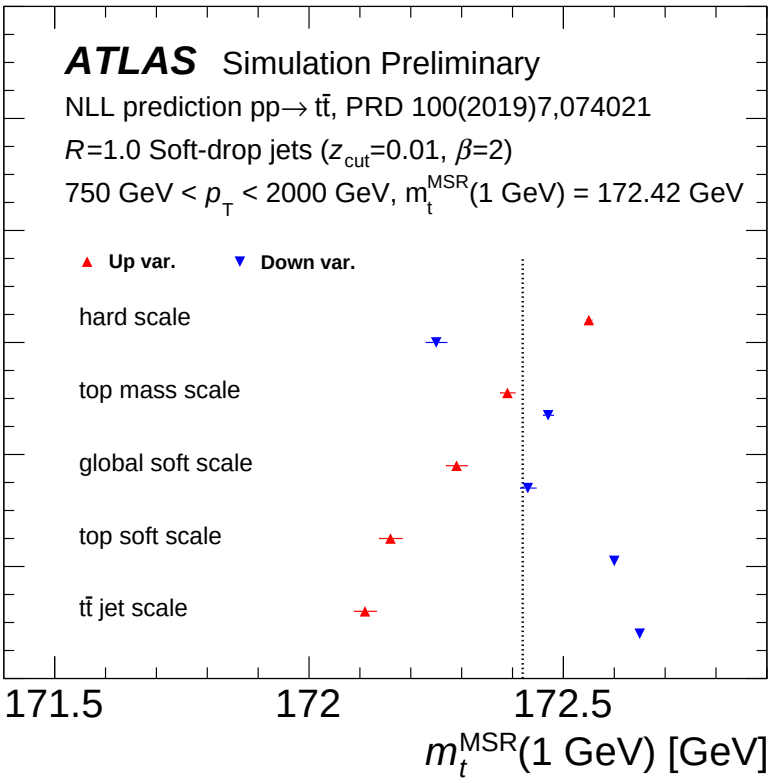


Figure 5.11: Impact of scale variations in the NLL calculation expressed in terms of the top quark MSR mass. The values of m_t^{MSR} are found by fitting the nominal NLL calculation to the predictions obtained with different choices of the five scales. The three parameters of the calculation are floated in the fit. For the calculations with alternative scale choices the parameters are set to the best-fit values for the ATLAS Powheg +Pythia 8 sample: $m_t^{\mathrm{MSR}}(1\ \mathrm{GeV}) = 172.42$ GeV (indicated by a vertical line), $\Omega_{1q}^{\circ\circ} = 1.49$ GeV, $x_2 = 0.52$.

The impact of the choice of the mass range in the fit is evaluated by varying the lower limit, as described in Section 5.5.2. Theory-to-MC comparisons are carried out in two alternative jet mass ranges, beginning at 172.0 GeV or 173.0 GeV. The fit range extends up to 180 GeV in all cases. In this exercise the value of x_2 is limited to ±0.1 around the best-fit result, to avoid excessive instability of the fit. The MSR mass values obtained from the fits with alternative mass ranges are compared to the nominal fit result and the difference is assigned as an uncertainty. This yields an uncertainty of ±170 MeV. Similar variations in the upper edge of the fit range by ±1 GeV result in variations of the result by ±30 MeV.

The impact of the choice of the large-R jet p_{T} intervals included in the fit is evaluated by comparing fits on sub-sets of two p_{T} intervals. The fit is repeated on all permutations of two out of three intervals defined in Section 5.4.4, and compared to the nominal fit, that has all three intervals. The maximal variation, ±80 MeV, is taken as the uncertainty.

These two components are added in quadrature, resulting in a combined methodological uncertainty of 190 MeV.

5.5.3 Underlying event and color reconnection modelling

As the MPI setting is switched off in the samples used to determine the relations between the MC mass and the MSR mass in Section 5.6, the presence of UE effects on the large-R jet mass distribution must be accounted for separately. A robust fit is obtained with the MC-based template-fit method introduced in Section 5.4.4 to alternative samples where variations on the A14 Var1 eigentune and the CR models available in Pythia 8 are considered.

The UE modelling is varied in samples generated with Pythia 8 standalone, with the MPI setting switched on (MPI on). Figure 5.12(a) presents the top

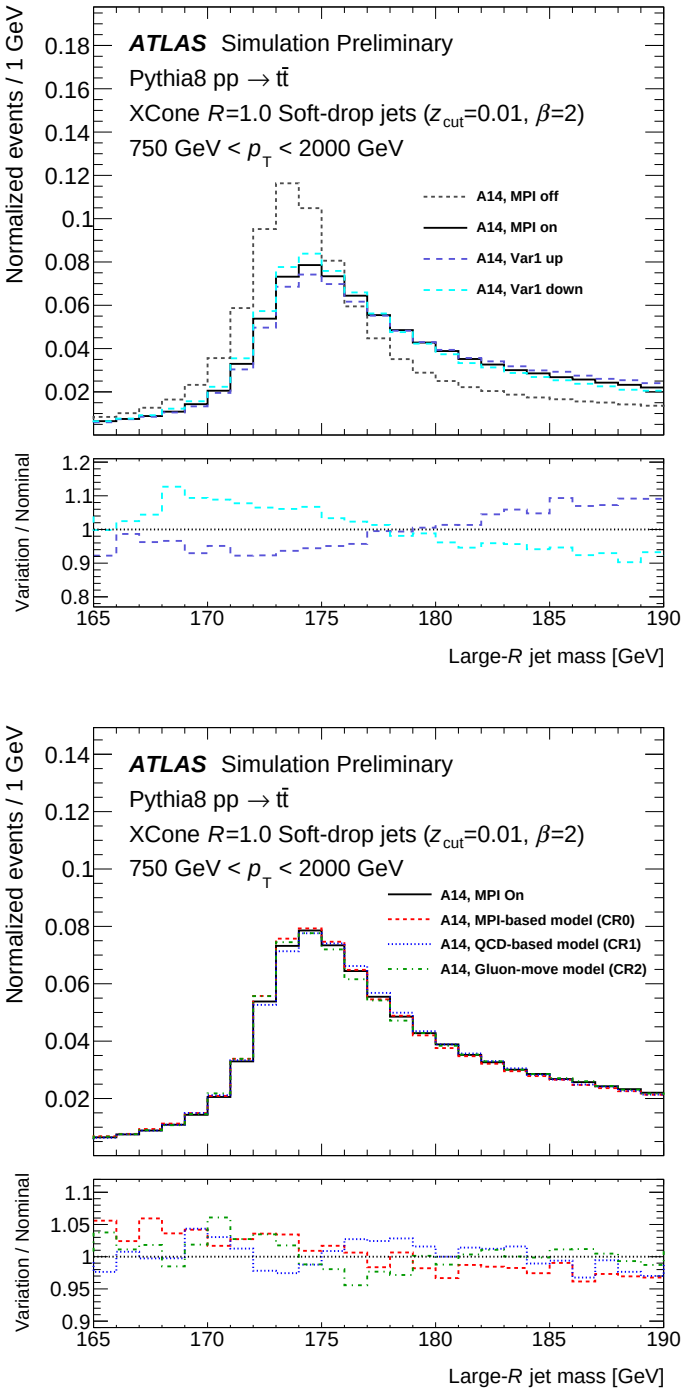


Figure 5.12: Large-R jet mass for boosted, hadronically-decaying top quarks for (left) the Pythia 8 A14 Var1 up and down variations and (right) several color reconnection models available in Pythia 8. For reference, the distribution is also shown for the nominal A14 tune, and with multi-parton interactions disabled in the left panel. Large-R jets are reconstructed using the XCone algorithm and light soft-drop grooming.

jet mass distributions obtained with the standard MC setup and the MPI model enabled, together with two variations of the A14 tune. The sample with the MPI model disabled is presented for reference. The impact of the MPI is a clear shift of the top mass peak to larger values. The distribution is also smeared out considerably. The A14 Var1 'up' and 'down' tune variations [207] change the UE activity by varying the BeamRemnants:reconnectRange parameter in Pythia 8 from 1.73 (up) to 1.69 (down) (the nominal value is 1.71), and the MultipartonInteractions:alphaSvalue from 0.131 (up) to 0.121 (down), (the nominal value for α_s in the MPI model from NNPDF is 0.126). These variations alter the large-R jet mass distribution: the Var1 up mass spectrum is harder and the Var1 down mass spectrum is softer than the nominal A14 distribution.

The MC-based template-fit yields values of m_t^{MC} = 172.641 ± 0.008 GeV and 172.342 ± 0.007 GeV for the A14 Var1 up and down variations, respectively, where the uncertainty corresponds to the MC statistical uncertainty. The difference with respect to the MC mass value used in the nominal template, m_t^{MC} = 172.5 GeV, is taken as an uncertainty, resulting in a symmetrized UE uncertainty of ±150 MeV.

Several alternative CR models available in Pythia 8 are also studied. Hadronization models operate over two partons that are actually color connected. The question of how to connect all partons available for hadronization is very hard to answer, since it fully relies on aspects beyond perturbative QCD. The MC event generators have traditionally used the leading-color approximation to trace the color flow and connect each parton to a single (unique) other parton in the event [220]. The CR models incorporate other possible connections between partons beyond leading-color. In pp collisions, this is specially challenging and important because the initial-state partons are colored and associated to colored beam remnants. The default CR model in Pythia 8 ('CR0') is MPI-based, where the color flow of partons belonging to different MPI systems can be fused.

It is used in the A14 tune, and corresponds to the CR model of the nominal large-R jet mass distributions when the MPI model is enabled. The standard CR0 configuration has been slightly modified within the A14 tune, allowing more reconnections to happen and adjusting some parameters of the standard MPI model. The resulting jet mass distributions are shown in Figure 5.12(b). As can be observed, two different CR schemes are investigated too. The first alternative model ('CR1') is based on QCD color rules and is observed to produce more massive jets, while the second alternative model ('CR2') which is based on the gluon-move scheme produces fewer massive jets. According to the approach adopted by the ALTAS experiment to asses the uncertainty due the CR choice, the default configuration should be directly compared to the two alternative CR models. We found that the associated m_t^{MC} values for the different CR models are 172.52 ± 0.01 GeV and 172.44 ± 0.01 GeV, respectively. The CR uncertainty is therefore −56 MeV and +20 MeV.

The contributions from UE and CR modelling are added in quadrature and symmetrized to yield a total uncertainty of ±155 MeV.

5.5.4 Summary of uncertainties

The uncertainties on the relation between the top quark mass parameter in MC event generators and the MSR mass scheme $m_t^{\mathrm{MSR}}(1\ \mathrm{GeV})$ are summarized in Table 5.3.

These uncertainties apply to the mass relation for a given MC generator setup. Variations of the PS and hadronization model and their parameters are not considered uncertainties. The relation between MC mass parameter and the MSR mass scheme can be determined for each generator setup. The results for several of the ATLAS samples used in direct mass measurements are presented in

Section 5.6, and the impact of a larger set of variations is estimated in MC-to-MC fits.

These uncertainties are strictly valid only for the large-R jet mass that is used to derive the mass relation. The stability of the result under variations of the observables is studied, within the limitations of the NLL calculations, in Section 5.6.3. Additional uncertainties due to the extrapolation to a different observable and selection may be required if the mass relation is used to calibrate direct mass measurement that use a different set of observables.

Table 5.3: Uncertainties on the relation between the top quark mass parameter in Monte Carlo generators and the MSR mass at a scale of 1 GeV.

Source	**size [MeV]**	**Comment**
Theory - higher-order corrections	$+230/-310$	Envelope of NLL scale variations
Fit methodology	± 190	Choice of fit range, p_{T} bins
Underlying Event model	± 155	A14 eigentune variations, CR models
Total	$+350/-400$	

5.6 Results

In this section, the main results of this chapter are presented. The impact of varying aspects of the MC generation scheme on the top jet mass distribution is studied by using alternative generators and samples.

We shall first present the results of this study in terms of shifts in the top MC mass. It will be followed by the determination of the relation between the top mass parameter of the MC generator and the top MSR mass:

$$m_t^{\mathrm{MC}} = m_t^{\mathrm{MSR}}(1\ \mathrm{GeV}) + \Delta_m^{\mathrm{MSR}}.$$

The value of Δ_m^{MSR} is determined for the nominal $pp \to t\bar{t}$ generator setup used in ATLAS top physics analyses, Powheg +Pythia 8, and for several internal

variations of the nominal MC generator. Additional fits are performed with a varied PS model using a sample generated with POWHEG +HERWIG 7. The MPI model is switched off for all MC samples considered in these fits.

5.6.1 Featuring the top Monte Carlo mass

The impact of a complete set of MC variations is obtained using MC-based template-fits. In these fits MC templates for the nominal ATLAS sample, with a floating mass parameter, are compared to alternative MC generator setups.

Pythia 8 standalone with the Monash tune

Ref. [210] studied PYTHIA 8 (v8.240) with the Monash tune. For reference, we determine the shift of the top quark mass parameter in the nominal ATLAS POWHEG + PYTHIA 8 sample. This yields a mass shift of +50 MeV. The slight shift is entirely due to the change of tune: the mass shift reduces to 0 if the A14 tune [207] is used.

Matrix Element variation

The jet mass distribution obtained from events generated with MADGRAPH5_aMC@NLO and showered/hadronized with PYTHIA 8 exhibits a clear shift to lower values with respect to the nominal MC sample, as can be seen in Figure 5.13. The MC-based template-fit confirms this observation and yields a MC mass of 172.38 ± 0.05 (stat) GeV. In order to isolate the effect of the ME-PS matching, the result must be compared to a POWHEG +PYTHIA 8 with the MEC in PYTHIA 8 switched off ('MEC off'). For this sample we find a MC mass of 172.40 ± 0.04 (stat) GeV. Therefore, we conclude that the effect of replacing the nominal ME generator in ATLAS MC samples, POWHEG by the alternative MADGRAPH5_aMC@NLO has a negligible effect in terms of the MC mass, within the uncertainties of the template-fits. The MEC themselves do

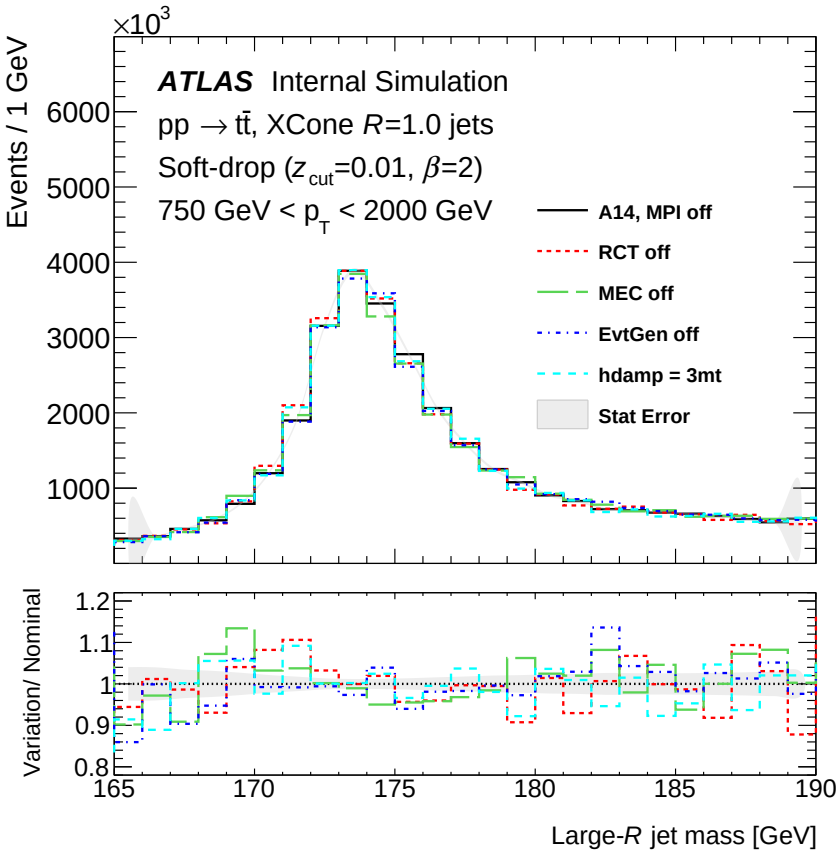


Figure 5.13: Large-R jet mass distribution for Monte Carlo events generated with different varied settings of the standard configuration.

introduce a shift of 100 MeV on the MC mass.

Tune variations

Figure 5.14 compares the large-R jet mass distribution for many of the A14 eigentunes [207]. The Var2, Var3a, Var3b vary parameters related to the emission of additional jets (ISR or FSR), as the α_s value in time-like showers and the reference p_{T} of space-like showers. Var3c modifies the α_s value in space-like showers and affects only the ISR. The shape of the jet mass distribution is observed to change for the tune variations that affect the FSR and is insensitive to Var3c, as expected.

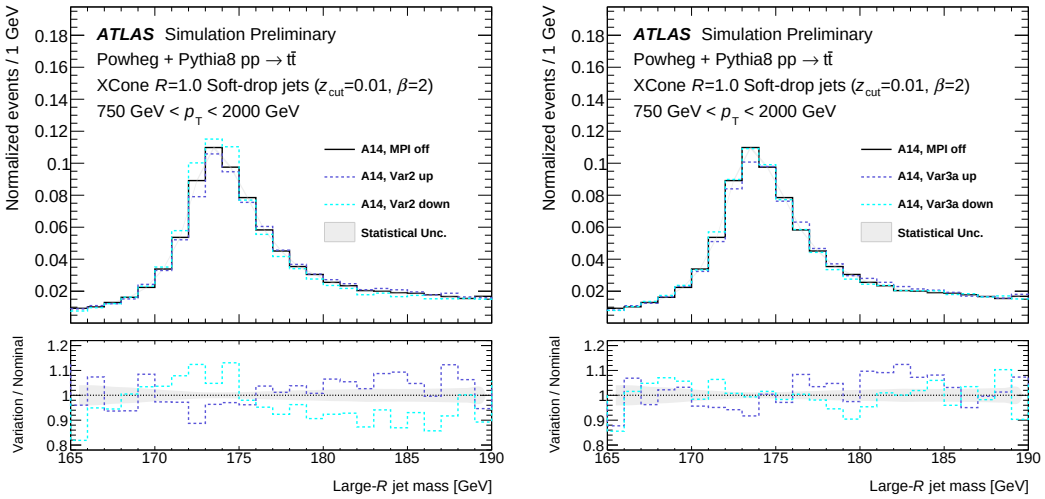
The mass shifts corresponding to each of these variations are summarized in Figure 5.15. The A14 eigentune variations Var2, Var3a and Var3b lead to a

mass shift of up to ±150 MeV and are generally symmetric for the up and down variations.

An advantage of the MC-based template-fit is that it can be performed also in presence of UE. We can therefore compare the results for the eigentunes of Figure 5.15 with those obtained in the corresponding MC samples with the MPI model switched on. For the Var2 and Var3 eigentunes the pattern of mass shifts is very similar for MPI on and MPI off. This is discussed in details in Section 5.7. We can anticipate that most results agree within 20 MeV and all are within 50 MeV. The relation between the MSR mass and MC mass that we derive in a MC sample without UE is therefore a good indication of the evolution of the mass definition in the PS, even if the UE modelling may further modify the mass relation.

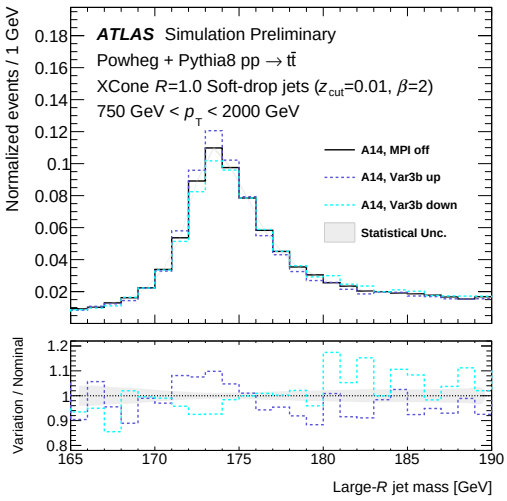
Other Monte Carlo variations

The jet mass distributions of the other alternative $t\bar{t}$ samples listed in Section 5.3 are displayed in Figure 5.13, and the results of the MC-based template-fit are also collected in Figure 5.15. The effects of the EvtGen package and the value of the *rFacB* parameter are found to be negligible, while variations of the RTC setting, MEC and the h_{damp} parameter are of the order of 100 MeV. The effect of these settings on the top jet mass is relatively small compared to its impact in observables used in direct mass measurements, such as the invariant mass of the bottom quark and lepton system m_{bl} in resolved $t\bar{t}$ events. The groomed jet mass is a very inclusive observable, as only very-wide-angle radiation can escape the catchment area of the large-R jet. The effect of these settings in ungroomed jets is found to be even smaller, providing further support for this explanation.

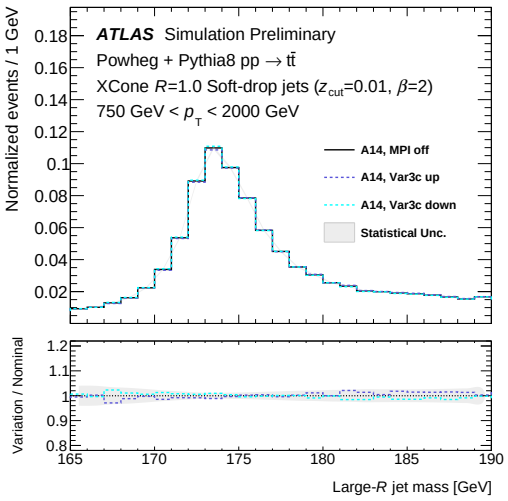


(a)

(b)



(c)



(d)

Figure 5.14: Large-R jet mass distributions for boosted, hadronically-decaying top quarks. Results are shown for $t\bar{t}$ events generated with POWHEG + PYTHIA 8 with the multi-parton interaction model (PYTHIA 8 setting 'MPI') disabled. Several variations of the A14 tune are shown along with the nominal setup, Var2 in panel (a), Var3a in panel (b), Var3b in panel (c) and Var3c in panel (d).

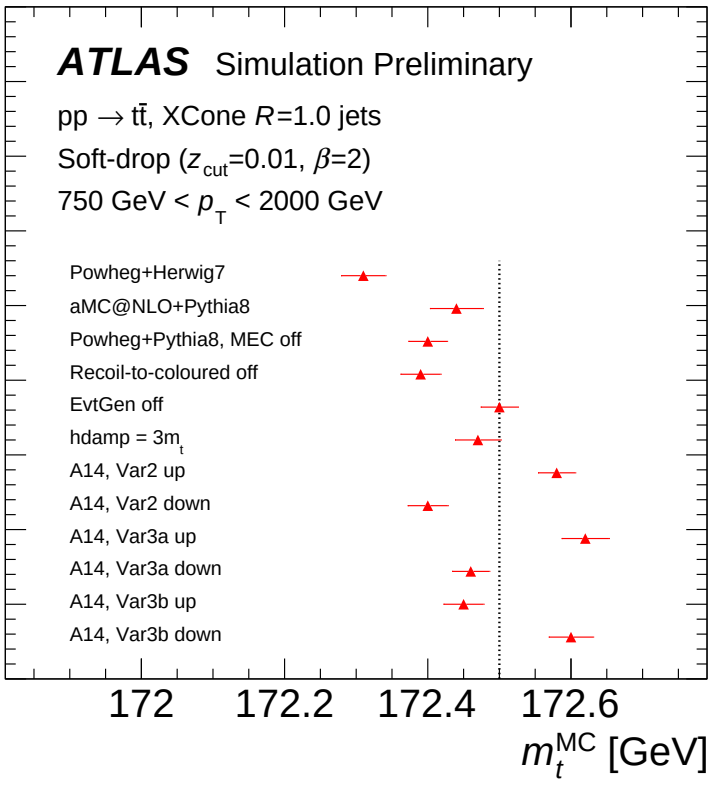


Figure 5.15: Equivalent shifts in the top Monte Carlo mass m_t^{MC} with respect to the nominal Powheg +Pythia 8 Monte Carlo sample. The offsets are found by fitting the nominal MC mass template with a floating MC mass parameter to several alternative choices of the Powheg +Pythia 8 configuration and to the Powheg + Herwig 7 and MadGraph5_aMC@NLO +Pythia 8 samples. The multi-parton interaction model is disabled in these fits.

5.6.2 Calibration of the top Monte Carlo mass

Nominal ATLAS Powheg + Pythia 8 sample

We are now ready to present the main result of this chapter. Figure 5.16 displays the normalized jet mass distributions in the three p_{T} intervals for the nominal ATLAS $t\bar{t}$ sample. The best-fit NLL predictions are compared to the MC simulation in the same figure. The NLL calculation is able to describe the main shape of the MC simulation over the mass range and p_{T} bins included in the fit, well within the theory uncertainty band.

The distribution of the reduced[1] 2D-$\chi 2$ is shown as a function of two of the three fit parameters in Figure 5.17. The minimum $\chi 2$ is 2.3, which is considered adequate for the purpose of this study. We see that the three parameters remain highly correlated even in the fit to three p_{T} bins, as the degeneracy is only lifted partially by the different scaling with p_{T}.

The marginalized results for the MSR mass, $\Omega_{1q}^{\circ\circ}$ and x_2 are given by:

$$\begin{aligned} m_t^{\mathrm{MSR}}(R = 1\ \mathrm{GeV}) &= 172.42 \pm 0.10\ \mathrm{GeV}\ , \\ \Omega_{1q}^{\circ\circ} &= 1.49 \pm 0.03\ \mathrm{GeV}\ , \\ x_2 &= 0.52 \pm 0.09\ , \end{aligned} \tag{5.7}$$

where the associated uncertainty corresponds to the statistical uncertainty due to the limited MC sample. The relation between the MSR mass and the MC mass parameter is hence found to be:

$$m_t^{\mathrm{MC}} = m_t^{\mathrm{MSR}}(1\ \mathrm{GeV}) + 80^{+350}_{-400}\ \mathrm{MeV}, \tag{5.8}$$

where the uncertainty includes a statistical contribution (±100 MeV) and

[1] Normalized to the *Number of Degrees of Freedom (NDF)*.

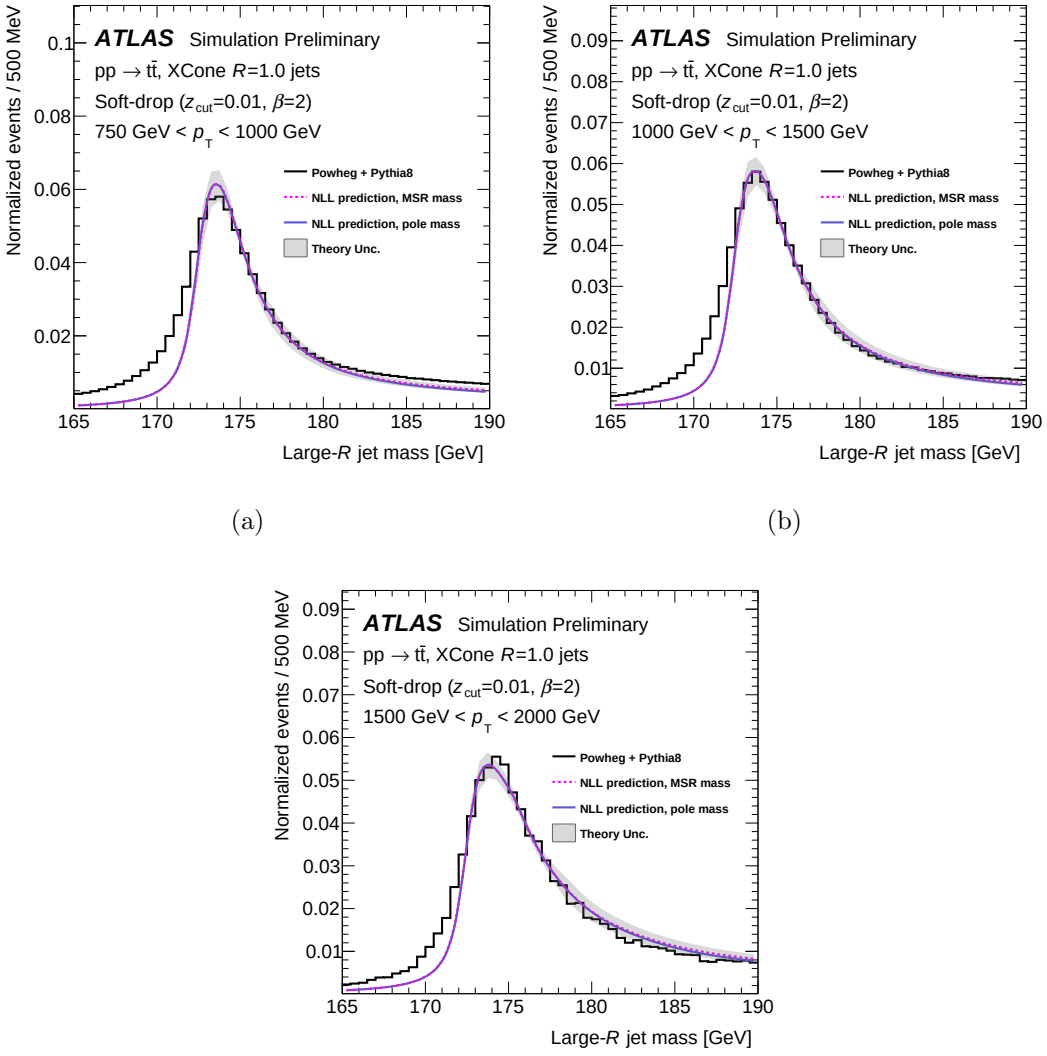


Figure 5.16: The particle-level jet mass distribution of the nominal POWHEG +PYTHIA 8 sample with the multi-parton interaction model disabled (black histogram) and the NLL prediction of Ref. [210] which best describe the Monte Carlo prediction (smooth curves). The NLL calculation is performed in the MSR mass scheme (dashed pink curve) and in the pole mass scheme (purple continuous curve). In both mass schemes, the three parameters of the calculation, the top quark mass, $\Omega_{1q}^{\circ\circ}$ and x_2, are fitted to find the best description of the three p_{T} intervals used in the calibration procedure. The distributions are normalized and the fit is performed on the interval 172.5 GeV $< M_J <$ 180 GeV.

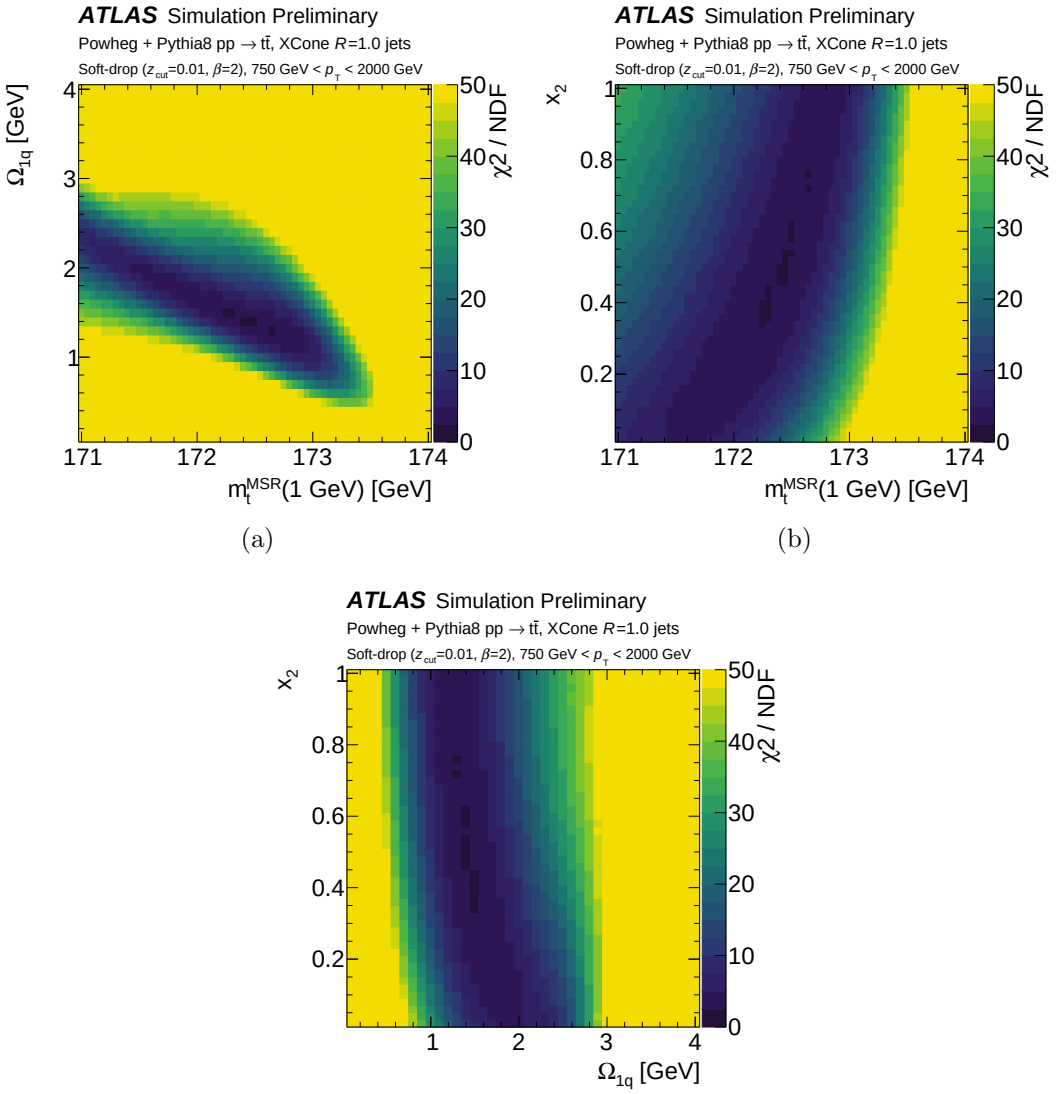


Figure 5.17: The $\chi 2$ results as a function of the MSR mass $m_t^{\mathrm{MSR}}(1\ \mathrm{GeV})$ and the parameters of the shape function $\Omega_{1q}^{\circ\circ}$ and x_2, obtained from the fit with the NLL prediction of Ref. [210] to the particle-level jet mass distribution for boosted top quarks in the nominal ATLAS Powheg +Pythia 8 sample. The results are shown in the two-dimensional plane of (a) MSR mass versus $\Omega_{1q}^{\circ\circ}$, (b) MSR mass versus x_2 and (c) $\Omega_{1q}^{\circ\circ}$ versus x_2. In each case, the 2D distribution is obtained by marginalizing over the third parameter. The three parameters of the calculation are varied to find the best description of the three p_{T} intervals used in the calibration procedure.

systematic contributions due to missing higher-orders in the NLL calculation (+230/ – 310 MeV), due to the uncertainty associated to the fit methodology (±190 MeV), and due to the UE uncertainty (±155 MeV). The MC mass is compatible with the MSR mass, given the uncertainty.

Carrying out fits with a pole mass parameter (as opposed to the scale dependent MSR mass) in the NLL predictions, we find a mass relation that can be compared to the equivalent fits in Ref. [210]:

$$m_t^{\mathrm{MC}} = m_t^{\mathrm{pole}} + 350^{+300}_{-360}\ \mathrm{MeV}, \tag{5.9}$$

where the total uncertainty is calculated in the same way as for Eq. 8.3. The smaller total uncertainty in comparison to MSR result is due to the reduced theory uncertainty, as mentioned in Section 5.5.

It is interesting to remark that the MSR mass is numerically close to the top quark pole mass, within the intrinsic uncertainty of 140 MeV due to the pole mass renormalon ambiguity. Therefore, the pole mass interpretation of the MC mass parameter is validated to the precision that is usually assigned to this identification [50].

Internal variations of the nominal Monte Carlo scheme

To determine the modelling uncertainties of ATLAS top physics analyses, many aspects of the $t\bar{t}$ production and hadronization processes are studied in variations of the nominal sample. These variations are accessible via reweighting of the nominal MC sample. We repeat the fit to the jet mass distribution with the NLL predictions for all variations available.

The results are collected in Figure 5.18. The alternative models include variations of the value of the strong coupling constant α_s, of the renormalization and factorization scales, of the PDFs, of the ISR/FSR and variations of the A14

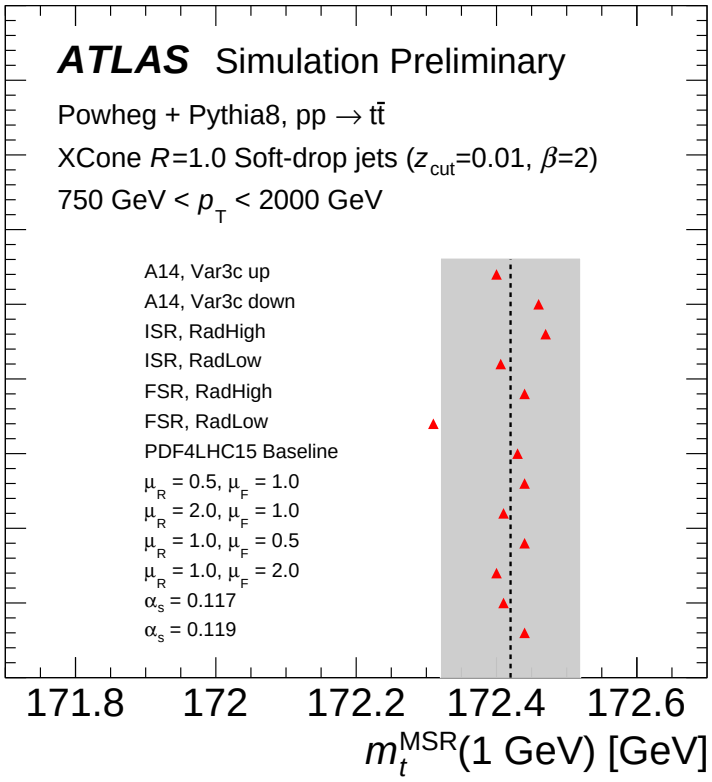


Figure 5.18: The MSR mass extracted from samples where a given aspect of the $t\bar{t}$ production process is altered with respect to the nominal Monte Carlo setup (dashed line). The variations affect the PDF, α_s and the renormalization and factorization scale, as well as the rate of Initial- and Final-State Radiation. The vertical band indicates the uncertainty on the nominal fit value.

Var3c eigentune. Most of these variations have a very small impact on the jet mass distribution, and lead to variations of the best-fit MSR mass of tens of MeV. We therefore conclude that the jet mass of boosted top jets is a robust observable that is relatively insensitive to these aspects of the MC generator.

The main exception is the down variation of the rate of the FSR, where the central mass value is shifted downwards by 110 MeV. The different mass value is accompanied by a higher value of $\Omega_{1q}^{\circ\circ} = 1.6$ GeV and lower value of $x_2 = 0.42$.

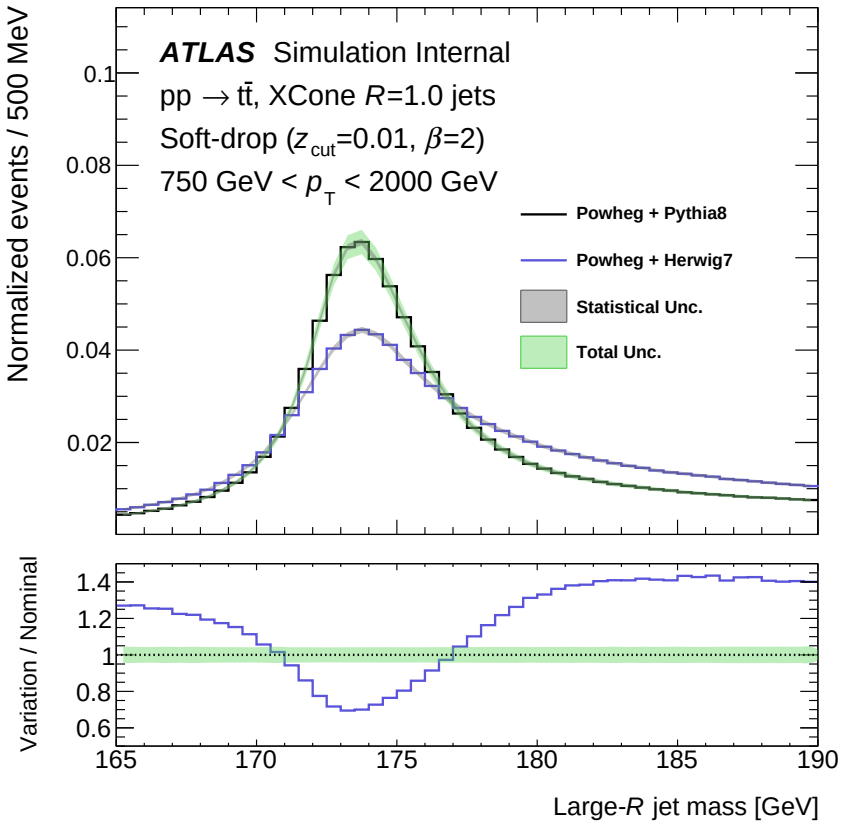


Figure 5.19: Comparison of $t\bar{t}$ events generated with POWHEG and showered with either PYTHIA 8 or HERWIG 7. The total error band accounts for statistical uncertainty, as well as variations of PDFs, α_s and renormalization / factorization scales within the nominal PYTHIA 8 sample. The ratio between nominal and the alternative Monte Carlo sample is provided in the bottom panel.

Powheg + Herwig 7

The POWHEG + HERWIG 7 sample is often used to estimate PS and hadronization uncertainties for physics analyses of top processes. The predictions for the top jet mass distribution of this sample are compared to POWHEG+PYTHIA 8 in Figure 5.19. PYTHIA 8 and HERWIG 7 predict very different jet mass distributions; the latter yields a harder jet mass spectrum. The MSR mass extracted for POWHEG + HERWIG 7 is:

$$
\begin{aligned}
m_t^{\mathrm{MSR}}(R = 1\ \mathrm{GeV}) &= 172.27 \pm 0.09\ \mathrm{GeV}\ , \\
\Omega_{1q}^{\circ\circ} &= 1.90 \pm 0.07\ \mathrm{GeV}\ , \\
x_2 &= 0.98 \pm 0.12\ ,
\end{aligned}
\tag{5.10}
$$

where the uncertainty is due to the limited MC statistics. The two parameters of the shape function absorb the difference between the two jet mass distributions, with significantly higher values for both $\Omega_{1q}^{\circ\circ}$ and x_2. The mass relation for the POWHEG + HERWIG 7 setup is, however, compatible within the statistical uncertainty with that obtained with POWHEG+PYTHIA 8, $m_t^{\mathrm{MSR}}(R = 1\ \mathrm{GeV}) = 172.42 \pm 0.10$ GeV.

5.6.3 Stability of the results

In this study, the relation between the MSR mass and the MC mass parameter is determined for a specific observable in a limited and extreme kinematic region. To study whether the relation maintains its validity beyond the environment it was derived in we investigate the stability of the result when the fit is repeated with a number of related observables that are accessible with the first-principles calculation of Ref. [210].

We vary the user-defined parameters of the SD algorithm to study the effect on the mass relation of the observable used in the fit, within the range of validity of the theory calculation [210]. The analysis is repeated for different sets of grooming parameters which lie in the calculation's region of validity: $\{\beta = 1, z_{\mathrm{cut}} = 0.02\}$, which grooms both soft- and wide-angled radiation more aggressively than the nominal parameters; $\{\beta = 2, z_{\mathrm{cut}} = 0.02\}$, which grooms soft radiation more aggressively but does not alter the SD angular weighting; and $\{\beta = 2,$

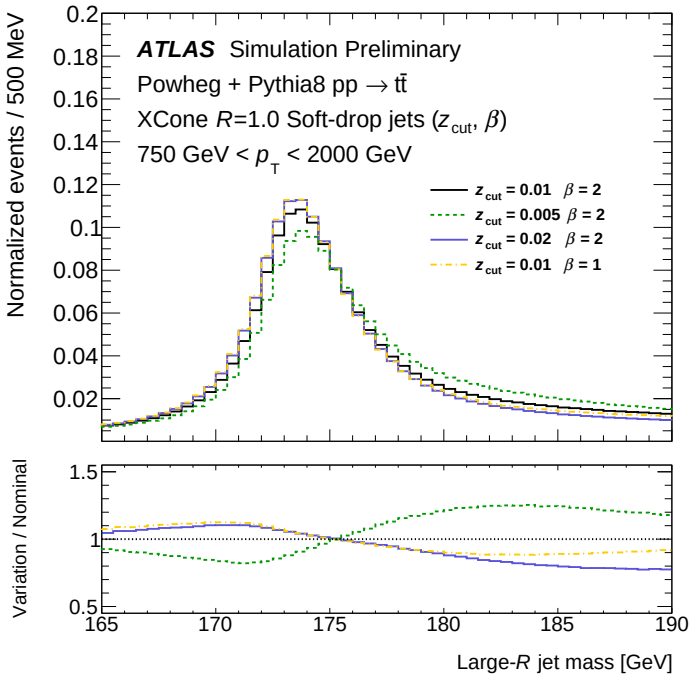
$z_{\text{cut}} = 0.005\}$, which is less aggressive.

Figure 5.20(a) shows how these variations of the SD configuration shape the jet mass distribution: the least-aggressive grooming option ($\beta = 2, z_{\text{cut}} = 0.005$) corresponds to the distribution with the largest mass values, since the fewest components are removed by the grooming procedure. A more aggressive grooming configuration, with larger z_{cut} value reduces pronounced mass tails and shifts the distributions to lower mass values, as do lower values of β.

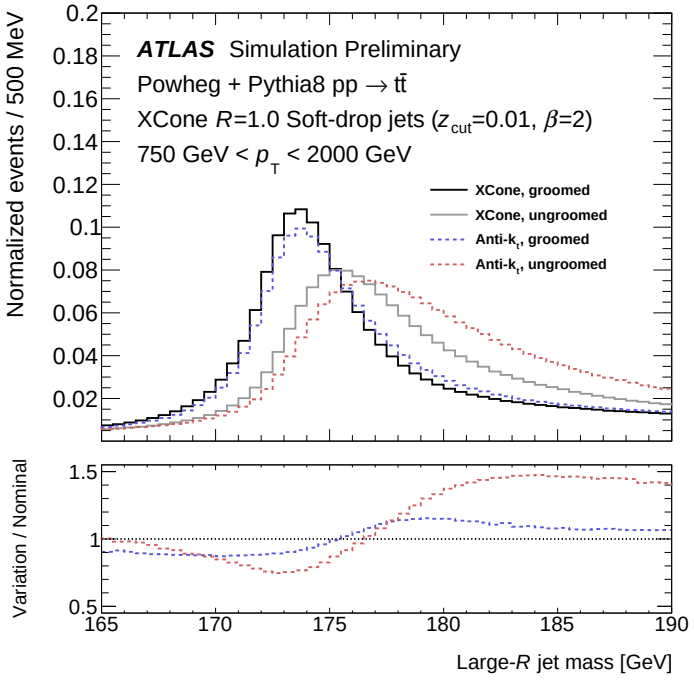
The fit to the NLL calculation is repeated for each of these distributions. For each comparison, the same grooming parameter settings are used in the MC simulation and in the theory calculation. The fit range is adjusted to follow the average jet mass, to avoid introducing second-order effects due to shifts of the top mass peak. The MSR mass is found to be 172.35 ± 0.05 GeV for $\{\beta = 2,\ z_{\text{cut}} = 0.005\}$ (less aggressive grooming) and 172.23 ± 0.04 GeV for $\{\beta = 2,\ z_{\text{cut}} = 0.02\}$ (more aggressive than the nominal). The result for $\{\beta = 1,\ z_{\text{cut}} = 0.01\}$ (more aggressive for soft- and wide-angled radiation) is 172.23 ± 0.04 GeV.

The algorithm used to cluster stable particles into large-R jets also has a non-negligible effect on the jet mass shape, as can be observed in Figure 5.20(b). The jets clustered with the anti-k_t and XCone jet reconstruction have different catchment areas and collect different constituents. The difference is clearly observed prior applying any grooming technique and remains visible with the light grooming applied here. The MSR mass extracted from large-R jets built with the anti-k_t algorithm and the nominal SD grooming parameters is found to be 172.56 ± 0.06 GeV, within 140 MeV of the nominal result.

The maximal variations of the mass relation due to the definition of the mass-sensitive observable are found to be contained within ±200 MeV, within the uncertainties of the procedure. We conclude that this relation between MC mass



(a)



(b)

Figure 5.20: The impact of different grooming configurations and jet clustering algorithms on the top jet mass distribution for (a) large-R XCone jets groomed with several soft-drop configurations, and (b) ungroomed and groomed (nominal soft-drop configuration) large-R jets clustered with the XCone and anti-k_t algorithms.

and MSR mass is stable at the level of ±200 MeV; this conclusion is compatible with the predictions of the underlying theoretical framework.

5.7 Discussion of the results

Interpretation of the main result

The key point of this work is to shed some light on the interpretation from first-principles of the top quark mass as parameter in the MC generator. This requires to study the relation between a renormalized mass and the top MC mass before detection. The purely experimental issues, such as detection inefficiencies and pile-up, may partially distort the sensibility of the m_t^{MC} to the MC underlying configuration. For this reason, experimental issues are left to be assessed in a separate study including real data.

In this respect, a specific MSR-MC mass relation for certain MC is given in full by its configuration: shower cut-off, top decay, hadronization of its decay products, etc.. Any variation from a reference MC setup should not be understood as an uncertainty, but as a different treatment and potential evolution and meaning of the top MC mass that has to be quantified separately. The uncertainties quoted in this result refers purely to the limitations of the underlying theory model, related to missing higher-orders, the phase space considered or the description of radiation off of the top quark decay groomed away as well as the UE contribution.

The result obtained in this work puts some evidence on the close relation between the MSR (at $R = 1$ GeV) and the MC mass, of the order of 80 MeV and finds some tension with respect to the pole mass, 350 MeV, still compatible within uncertainties. It is true that the theory lacks of the sufficient formal precision in

α_s to place the result of the pole mass on solid footing, but it goes in the same direction as the results found in the e^+e^- study at N^{3}LL [219].

The different relations derived for different hadronization models (Pythia 8 and Herwig) confirm the dependence of the evolution of the top MC mass with the treatment of the non-perturbative QCD aspects. This is expected due to the kinematic sensitivity of the observable used. However, the difference of 230 MeV between the two relations is totally covered by the current uncertainty of the method. On the contrary, we found that for given MC generator, the MSR-MC relation is very robust against internal variations of the hard scales, the PDF set, the α_s value and variations on the ISR/FSR.

The potential dependence of the MSR-MC relation on the value of m_t^{MC} itself is also investigated. This is carried out by using Pythia 8 standalone simulations, so samples with several values of m_t^{MC} and unlimited statistics are accesible. The fit procedure is frozen, and only the lower limit of the fit range is adjusted in order to follow the shift of the jet mass peak accordingly (as explained in Section 5.4.5, this is achieved in practice by identifying the lower limit with m_t^{MC}). Figure 5.21 shows the MSR-MC relation as a function of the top MC mass. As can be observed, the trend is clearly linear and the slope is found to be compatible with one within the uncertainties of the fit method.

Uncertainties and UE factorization

The uncertainties considered in this study deserve some further discussion. Concerning the usage of the envelope of the scale variations as a estimate of the missing higher-orders in the theory calculation, it is important to stress that in the theoretical community, both envelopes and quadratic sums are used depending on the underlying observable and prior experience. In addition, in this case the scale variations are not proper independent variations. The variations considered

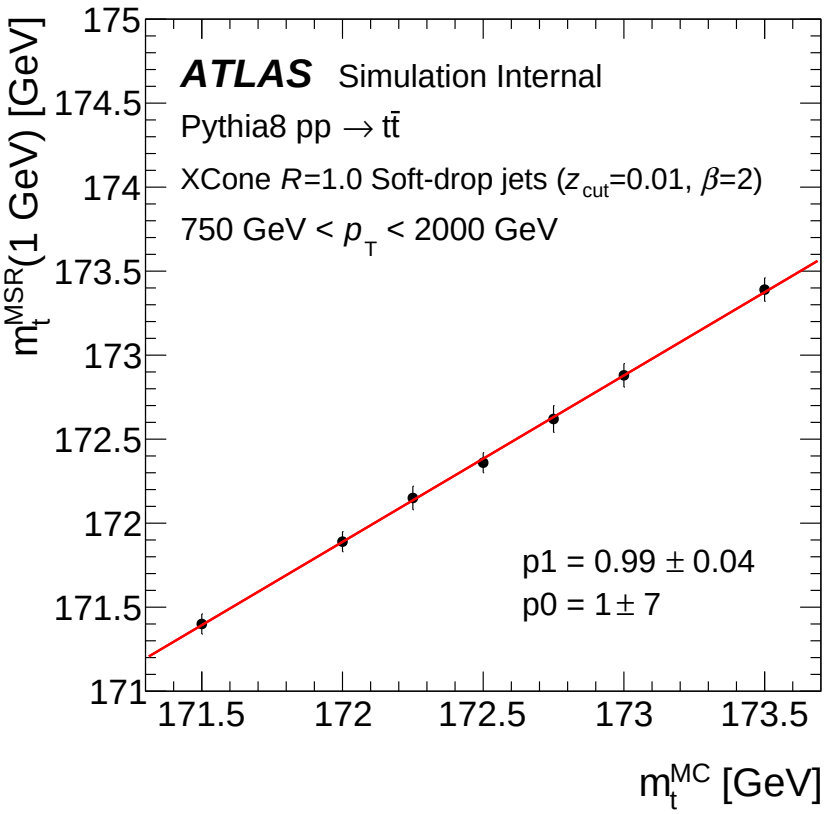


Figure 5.21: Study of the linearity of the MSR and Monte Carlo mass relation based on Pythia 8 standalone simulations. The fit procedure is frozen to the nominal NLL-based template-fit, and only the lower limit of the fit range is adjusted accordingly to the value of the top Monte Carlo mass with which the samples were generated.

are the result of profile functions that vary multiple scales at once. According to the authors of the NLL factorization theorem [210], the envelope is proposed because it provides coverage for the NNLL correction. This is known thanks to the guidance from calculations where higher-orders corrections are available, in particular the e^+e^- calculation in Refs. [124, 219].

The uncertainty assigned to cover the modelling of UE effects is subject of discussion as well. Ideally, the theory model would include the treatment of UE effects in a consistent way, and the MSR-MC mass would reflect the modelling given in the MC simulation. The authors of the NLL theorem are working on the development of a extended theory model that incorporates a dedicated shape function to absorb UE effects. But for now, a dedicated uncertainty needs to be assigned in the absence of a rigorous theoretical treatment.

The estimation of the UE uncertainty includes A14 Var1 eigentune variations, where the α_s value in the MPI model is altered. Such variations are intended to cover the uncertainties in the UE treatment. But the question is whether or nor the A14 tune, developed in 2014 from $\sqrt{s}$ = 7 and 8 TeV data with the best intuitions at that point, is able to describe the UE observed in 13 TeV collisions. In most distributions of interest, A14 is within 5 – 10% of the 13 TeV UE measurements or closer (see Ref [221] and the ongoing effort to constrain CR models with similar observables based on charged-particle multiplicities and momenta). It is unclear, right now, whether the variations of the Var1 eigentune provide complete coverage. This uncertainty should be revisited when better estimations of the UE contributions become available.

One remaining question regarding the UE treatment is whether or not the relation found with the MPI model disabled is still valid to describe situations where it is activated. To sort this out, the jet mass distributions from events generated with the MPI setting switched off and on were compared. In these

Table 5.4: The results of Monte Carlo based template-fits of the nominal jet mass distribution with a floating mass parameter to the A14 eigentune variations and alternative samples generated with Pythia 8 standalone.

		top MC mass [GeV]			
Sample		MPI-off		MPI-on	
Eigentune	Variation	Value	Δm	Value	Δm
Var1	Up	172.50 ± 0.005	0	172.64 ± 0.008	+0.14
	Down	172.50 ± 0.005	0	172.34 ± 0.007	−0.16
Var2	Up	172.64 ± 0.005	0.14	172.68 ± 0.008	0.18
	Down	172.39 ± 0.005	−0.11	172.35 ± 0.007	−0.15
Var3a	Up	172.61 ± 0.005	+0.11	172.59 ± 0.007	+0.09
	Down	172.49 ± 0.005	−0.01	172.51 ± 0.007	+0.01
Var3b	Up	172.41 ± 0.005	−0.09	172.39 ± 0.007	−0.11
	Down	$172,62 \pm 0,006$	+0.12	172.67 ± 0.008	+0.17
Var3c	Up	172.57 ± 0.005	+0.07	172.58 ± 0.007	+0.08
	Down	172.47 ± 0.005	−0.03	172.42 ± 0.007	−0.08
RTC off		172.45 ± 0.005	−0.05	172.46 ± 0.007	−0.04
CR Model 1		172.502 ± 0.005	+0.002	172.52 ± 0.007	+0.02
CR Model 2		172.495 ± 0.006	−0.005	172.44 ± 0.007	−0.06

conditions, the impact of several variations of internal settings present in the Pythia 8 configuration were considered. The shift on the position of the jet mass peak in the nominal sample was evaluated with a MC-based template-fit, and the results are collected in Table 5.4. As we can note, the shift caused by the A14 eigentune variations and setting off the RTC switch is the same within 50 MeV when MPI is off and on. We also see that the Var1 eigentune has no impact when MPI is off, as expected, and so the alternative CR models.

Stability and applicability of the results

A crucial and subtle aspect of the results presented in this chapter is related to the extrapolation of the conclusions drawn from an inclusive jet observable, based on highly boosted, hadronically-decaying top quarks, to any other observable in other kinematic regime.

The key assumption behind the proposal for the calibration of Refs. [124] is,

of course, that the mass relation found for an e^+e^- calculation, or a pp calculation in a different kinematic regime, holds also for the observables used in direct mass measurements. There is evidence that certain aspects of the MC approach lead to genuine changes in the mass definition. These will affect all observables in the same way and will be captured by our calibration, even if performed on a very different observable. The best example of a feature that generically affects the meaning of the MC mass parameter is the PS cut-off. As explained in Section 2.4, the cut-off will alter the mass relation [123].

We have explored the stability of the mass relation by varying aspects of the jet reconstruction algorithm, the SD grooming parameters and the kinematic range. All results are compatible with the assumption of one universal mass relation, within the uncertainties of our method and within the limited range of observables accessible to the calculation. A calculation by the same team has been used to derive a mass relation for e^+e^- observables [219], that is again in good agreement with our result. We are therefore fairly confident that the relation we find is more generally applicable than to just the observable used to derive the relation. It is plausible that the "calibration" proposed in Ref. [124] will bring the direct measurements closer to the field-theoretical mass scheme.

There is some evidence, on the other hand, that other aspects of the MC generators lead to different effects in different types of observables. The example here is the RTC setting in Pythia 8 that regulates the second gluon emission after the top decay. It has a strong ($\mathcal{O}(500$ MeV$)$) effect on m_{bl}, used in direct top mass measurements; a small effect on the lightly-groomed top jet mass; and virtually no effect at all on the ungroomed top jet mass. The more inclusive the observable, the less sensitive to this effect. A calibration based on the more inclusive observable will fail to account for the impact of the RTC setting in the generator on the measurement. A bias due to a mismodelling (*i.e.* a failure of

the PS to reproduce all-order-QCD) in this emission will not be corrected by the calibration. Modelling uncertainties must still be applied to account for these effects that are not covered by the calibration.

We note here, though, that the alternative to applying a calibration of this kind has been to leave the interpretation of direct mass measurements open, stating that it is not too far from the pole mass without quoting an uncertainty. That has led to the addition of an ad-hoc uncertainty (*i.e.* in the EW fit). We claim that the present study increases the precision of the interpretation, and it should be seen in the light of this alternative.

Chapter 6

SMEFT bounds from the $t\bar{t}$ charge asymmetry at the LHC

6.1 Introduction

The first measurements of the forward-backward asymmetry in $t\bar{t}$ events $A_{\rm FB}^{t\bar{t}}$, performed at Tevatron by the CDF and DØ experiments, found a significant discrepancy with respect to the prediction of the SM at NLO in QCD [222, 223]. CDF reported a 3σ discrepancy at high $t\bar{t}$ invariant mass $m_{t\bar{t}} \geq 450$ GeV [224], and also in the inclusive measurement. The DØ experiment confirmed this anomaly [225]. These results initiated a major research activity in two different directions: explaining the discrepancy in terms of BSM physics and improving the prediction within the SM. It also fostered the search of other anomalies in the $t\bar{t}$ production. Authors of Ref. [222] proposed the charge asymmetry $A_{\rm C}$ as candidate to replace $A_{\rm FB}^{t\bar{t}}$ at the LHC.

Although $A_{\rm FB}$ and $A_{\rm C}$ are different observables, the fact that the first measurements of $A_{\rm C}^{t\bar{t}}$ performed by the ATLAS [101, 226] and CMS [227, 228]

experiments with data collected at $\sqrt{s}$ = 7 and 8 TeV were in good agreement with the SM, called into question the anomalies found at Tevatron. This is because the underlying physics mechanism that generates the asymmetry is the same for the two observables, as mentioned in Section 2.3.2.

The significance of the discrepancy between the observed $A_{\mathrm{FB}}^{t\bar{t}}$ and the SM prediction has always critically hinged on the size of missing higher-order corrections [229]. Originally, the theory prediction was at NLO precision in QCD. The inclusion of weak, weak-QCD and QCD-QED corrections increased the predicted $A_{\mathrm{FB}}^{t\bar{t}}$ by 25% with respect to the NLO QCD one [97]. The most accurate calculation, with an approximate N^{3}LO QCD+NLO EW [230] increases the NLO QCD value by a factor 1.3. With the full Tevatron dataset analyzed, CDF reports an excess of 1.7σ over the SM prediction [231] and DØ finds an agreement within 1σ [232].

The ATLAS and CMS collaborations are actively working on providing inclusive and differential measurements with the highest precision possible. So far, all results reported were found to be consistent with the SM predictions. However, the charge asymmetry is still a very powerful observable for two reasons: first, it has a sensitivity to new physics BSM that complements the cross-section [233]. Second, the theory uncertainties are smaller than the experimental ones.

In this chapter, the differences between the latest measurement of the charge asymmetry performed by the ATLAS experiment and the most accurate SM prediction available up to date are interpreted in terms of SMEFT dimension-six operators. To this end, the measurement of the $A_{\mathrm{C}}^{t\bar{t}}$ with 139 fb^{-1} of data at 13 TeV collected by the ATLAS detector is presented. The measurement is made in $t\bar{t}$ events selected in three orthogonal regions: in the boosted and resolved topologies of the semileptonic channel, and in the dileptonic channel. Furthermore, the $A_{\mathrm{C}}^{t\bar{t}}$ is measured inclusively and differentially as a function of the invariant mass of

the $t\bar{t}$ system $m_{t\bar{t}}$, the transverse momentum $p_{\mathrm{T},t\bar{t}}$ and the z-component of the velocity $\beta_{z,t\bar{t}}$, in both channels. In this chapter, we will focus on the inclusive and the $m_{t\bar{t}}$ differential measurements.

The measurement of the $A_{\mathrm{C}}^{t\bar{t}}$ from all available regions are simultaneously combined during the unfolding procedure, based on the *Fully Bayesian Unfolding* (FBU) method [234]. This technique allows to bring the measured $A_{\mathrm{C}}^{t\bar{t}}$ from detector level to parton level by correcting for detection and hadronization effects. At this point, a direct comparison between measurement and first-principles theory predictions becomes possible.

All signal and background processes are modelled using MC simulations, with the exception of fake lepton backgrounds, for which an accurate MC-based estimation is not available. The nominal $t\bar{t}$ sample is generated with the Powheg Box generator, which provides the ME calculation at NLO accuracy in QCD, with the NNPDF3.0NLO PDF set and the h_{damp} parameter set to $1.5 \cdot m_t$. The renormalization and factorization scales are set to $\mu_R = \mu_F = \sqrt{m_t^2 + (p_{\mathrm{T},t}^2)}$. Events are interfaced with Pythia v8.2 to simulate the PS and the hadronization, using the A14 set of tuned parameters and the NNPDF23LO PDF set. The full simulation of the detector response is carried out with the GEANT4 toolkit [235]. The different sources of background, *i.e.* processes involving the production of single tops, QCD V + jets, two bosons (VV), $t\bar{t}V/H$, tWZ and tZ are simulated with the best MC models available.

The systematic uncertainties associated to this measurement affect, on the one hand, the modelling of signal and background processes. This includes the $t\bar{t}$ and single-top PS and hadronization modelling, the $t\bar{t}$ and single-top radiation (ISR/FSR) and the PDF set among others. On the other hand, a set of systematic uncertainties are included to cover experimental aspects, such as the luminosity, background estimation and the event/object selection and reconstruction. In

this last category, the lepton identification, reconstruction and isolation, the jet energy scale/resolution and the missing transverse energy reconstruction play a larger role.

6.2 Analysis strategy

6.2.1 Event selection and reconstruction

Semileptonic channel

In the semileptonic channel, the resolved and boosted topologies share part of the event selection. It consists of a lepton selection (exactly one electron or muon with $p_{\mathrm{T}} > 28$ GeV), a significant $E_T^{\mathrm{miss}} > 30$ GeV with a W boson transverse mass of $M_{\mathrm{T}}^W > 30$ GeV for the electron channel and $E_T^{\mathrm{miss}} + M_{\mathrm{T}}^W > 60$ GeV for the muon channel and, finally, at least one b-tagged jet. The number of b-tagged jets is used to further classify events into 1b-tag-exclusive and 2b-tag-inclusive categories.

The event selection of the resolved topology further requires at least 4 small-R jets with $p_{\mathrm{T}} > 25$ GeV, a boosted veto in order to remove overlaps between events passing the resolved and the boosted topologies and an event reconstruction requirement, related to the efficiency of the parton-to-jet assignments (we will see what this veto means shortly).

In the boosted topology case, the selection also includes at least one small-R jet with $p_{\mathrm{T}} > 25$ GeV close to the lepton candidate ($\Delta R(\mathrm{jet}_{R=0.4}, l) < 1.5$), at least one top-tagged large-R jet with $p_{\mathrm{T}} > 350$ GeV and $|\eta| < 2$ on the opposite hemisphere of the event ($\Delta R(\mathrm{jet}_{R=1.0}, l) > 2.3$ and $\Delta R(\mathrm{jet}_{R=1.0}, \mathrm{jet}_{R=0.4}) > 1.5$) and, finally, a invariant mass of the reconstructed $t\bar{t}$ system of $m_{t\bar{t}} > 500$ GeV.

The reconstruction of the $t\bar{t}$ kinematics is carried out in the resolved regime

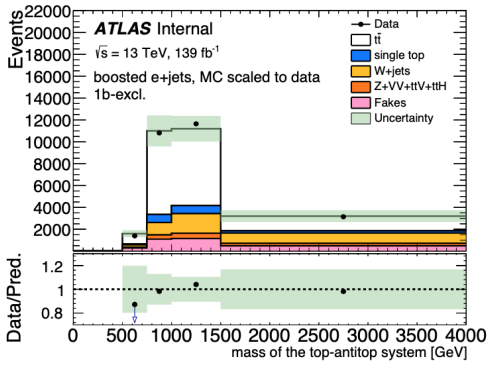
by means of a multivariate *Boosted Decision Tree* (BDT) technique implemented in the TMVA package [236]. The BDT combines information from the kinematic likelihood fitter [237] and the b-tagging information into a single discriminant with value from −1 to 1. The goal is to correctly assign individual selected jets to the corresponding partons from the decaying $t\bar{t}$ pair. So each possible permutation of the parton-to-jet assignments is evaluated, and the one with the highest score of the BDT discriminant is used for the $t\bar{t}$ reconstruction. However, only those events with a BDT score above certain threshold are considered. In this way events with a poor parton-to-jet pairing are discarded. The BDT is trained in MC simulations and the choice of the discriminant value threshold, which turns out to be 0.3, is based on the statistical uncertainty in the differential measurements.

The $t\bar{t}$ kinematic reconstruction in the case of the boosted topology is way simpler. On the one hand, the four-momentum of the large-R jet candidate is taken as the four-momentum of the hadronically-decaying top (antitop) quark. On the other hand, the four-momentum of the leptonically-decaying antitop (top) quark is reconstructed from the selected small-R jet, the lepton candidate and the neutrino four-momentum (derived from constraints of E_T^{miss}, lepton kinematics and the mass of the W boson in the PDG).

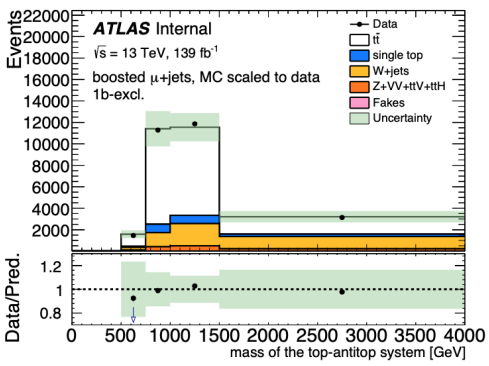
In both cases, the $m_{t\bar{t}}$ is obtained by adding up all four-vectors of the kinematic $t\bar{t}$ decay products. The control plots of the variables used for the asymmetry measurement in the boosted regime are shown in Figure 6.1.

Dileptonic channel

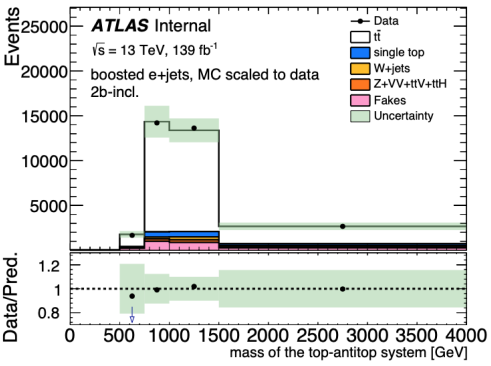
The event signature of the dileptonic channel is two opposite-signed leptons and 2 b-jets. Depending on the flavours of the leptons, events are separated into three categories: ee, $\mu\mu$ or $e\mu$. The event selection requires one charged lepton with $p_{\text{T}} > 28$ GeV and one additionally charged lepton with $p_{\text{T}} > 25$ GeV,



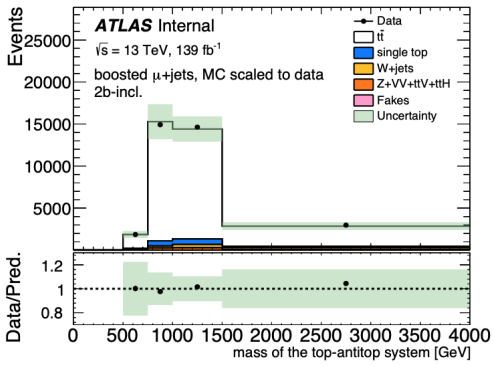
(a) $m_{t\bar{t}}$, boosted e+jets, 1b



(b) $m_{t\bar{t}}$, boosted μ+jets, 1b



(a) $m_{t\bar{t}}$, boosted e+jets, 2b



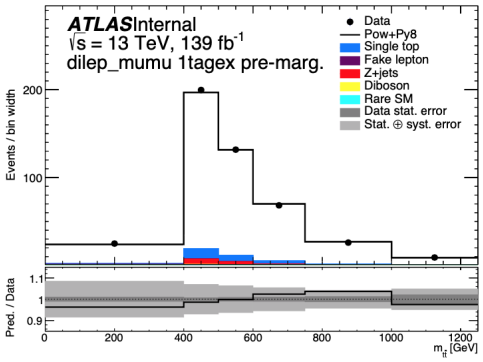
(b) $m_{t\bar{t}}$, boosted μ+jets, 2b

Figure 6.1: Comparison of the observed data and the Monte Carlo prediction (signal and backgrounds) for the differential charge asymmetry measurement as a function of $m_{t\bar{t}}$ in the semileptonic channel and boosted regime. The light green bands represents the sum in quadrature of the statistical and pre-marginalization systematic uncertainties. The Monte Carlo is scaled to data.

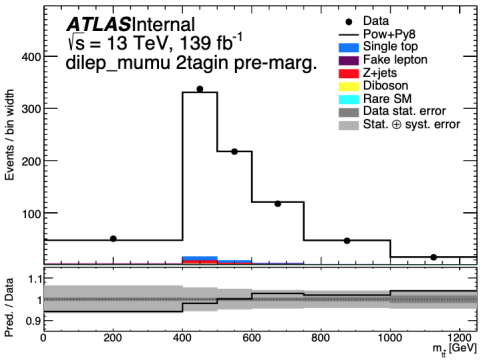
at least two small-R jets with $p_{\mathrm{T}} > 25$ GeV and at least one b-tagged jet. It also imposes several vetoes in the same flavour channels (ee and $\mu\mu$): a veto in the invariant mass of the dilepton system so it is outside the Z mass window ($|m_{ll} - m_Z| > 10$ GeV), a veto of $E_T^{\mathrm{miss}} > 60(30)$ GeV for the $1b$-tag-exclusive ($2b$-tag-inclusive) regions, and finally a Drell-Yan veto where $m_{ll} > 15$ GeV to suppress low mass resonance backgrounds.

The kinematic reconstruction of the $t\bar{t}$ system in the dileptonic channel is specially challenging due to the presence of two unobserved neutrinos. The neutrino weighting technique [238] allows to extract the four-momentum of the two neutrinos from the total E_T^{miss} by considering additional constraints on the top quark mass, the W boson mass and on the pseudorapidities of the two neutrinos. Different values of the neutrinos pseudorapidities are considered and used to provide a reconstructed E_T^{miss}. A weight is introduced in order to quantify the agreement between the reconstructed and observed E_T^{miss}. The highest weight points to the values of the pseudorapidities that are most likely the correct ones.

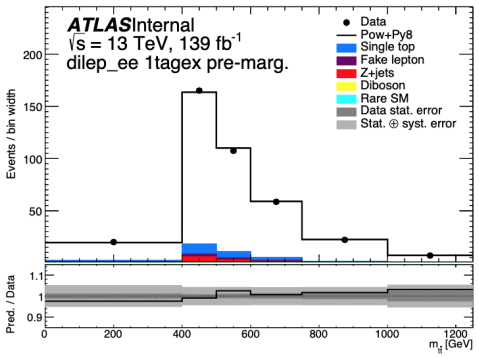
Once the top and antitop quarks are reconstructed from neutrinos, leptons and the small-R jets, the $t\bar{t}$ system is obtained by adding up all four-vectors of the kinematic $t\bar{t}$ decay products. The control plots for the reconstructed $m_{t\bar{t}}$ used in the differential measurements are shown in Figure 6.2.



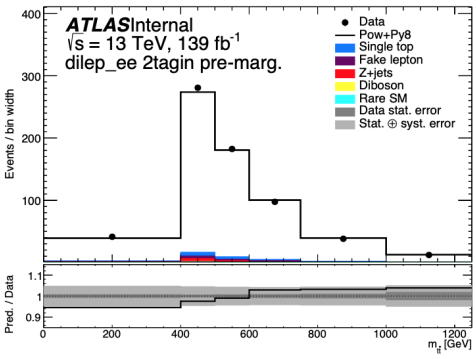
(a) $m_{t\bar{t}}$, $\mu\mu$ channel, 1-b excl.



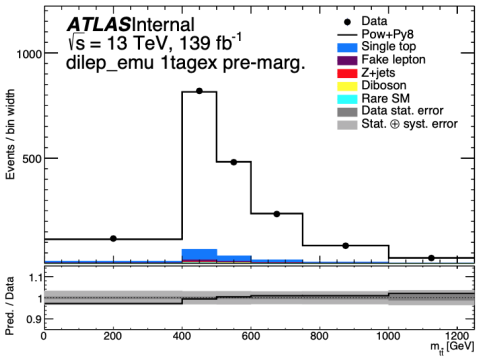
(b) $m_{t\bar{t}}$, $\mu\mu$ channel, 2-b incl.



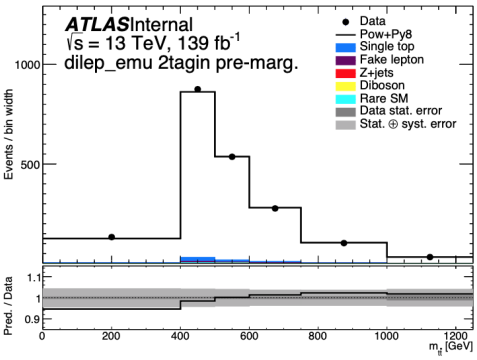
(a) $m_{t\bar{t}}$, ee channel, 1-b excl.



(b) $m_{t\bar{t}}$, ee channel, 2-b incl.



(a) $m_{t\bar{t}}$, $e\mu$ channel, 1-b excl.



(b) $m_{t\bar{t}}$, $e\mu$ channel, 2-b incl.

Figure 6.2: Comparison of the observed data and the Monte Carlo prediction (signal and backgrounds) for the asymmetry $m_{t\bar{t}}$ differential measurement in the dileptonic channel and resolved regime. The dark uncertainty bands in the ratio plots stand for the statistical uncertainty, whilst the light bands accounts for statistical and systematic uncertainties added in quadrature.

6.2.2 Unfolding to parton level

As introduced previously, the parton-level charge asymmetry is estimated from the measured spectra by means of the FBU technique [234]. Simply put, the FBU relies on the Bayesian inference to the unfolding problem: given the data $D \in \mathbb{N}^{N_r}$, we want to know the associated true spectrum $\tilde{T} \in \mathbb{R}^{N_t}$, represented by histograms of dimensions N_r and N_t respectively. The detector response matrix $\mathcal{M} \in \mathbb{R}^{N_r} \times \mathbb{R}^{N_t}$ encodes the information needed for the unfolding: it relates the reconstructed signal distribution D and the true distribution $\tilde{T}$ taking into account the detection efficiency, detector acceptance, selection efficiency and migrations across bins. The detector response matrix is estimated from MC simulations. According to the Bayes theorem, the probability $P(T \mid D, \mathcal{M})$ of obtaining T (formally different from $\tilde{T}$) given D and $\mathcal{M}$ is given in full by:

$$P(T \mid D, \mathcal{M}) \propto \mathcal{L}(D \mid T, \mathcal{M})\, \pi(T) \ , \tag{6.1}$$

where $\mathcal{L}(D \mid T, \mathcal{M})$ is the likelihood function of D for a given T and $\mathcal{M}$. It is based upon the assumption that the data follows Poisson statistics, so it is constructed as the product of Poisson probabilities for each bin of the spectrum. The $\pi(T)$ term of Eq. 6.1 represents the prior probability density function for the true spectrum T, and it is based on what we know about T before the measurement is performed. In this way, by sampling the prior probability distribution, the posterior distribution of the true spectrum can be obtained.

The unfolding procedure allows to incorporate the systematic uncertainties as a nuisance parameters by extending the likelihood term. The nuisance parameters are marginalized by projecting the true distribution over each nuisance parameter dimension. The FBU also exploits the fact that the orthogonal regions where the selected events are categorized have different background contamination.

Through a combination of several regions, the algorithm can marginalize certain nuisance parameters and hence reduce the associated systematic uncertainties. An example is the b-tagging efficiency uncertainty, which is better constrained by unfolding regions where events are classified according to the b-jet multiplicity. Finally, this unfolding technique is able to simultaneously unfold the two channels considered in the analysis, with their respective regions, given rise to a single unfolded distribution. This allows to further constraint the systematic uncertainties.

The potential bias introduced during the unfolding procedure is estimated with a linearity test, where pseudo-data samples with known true distributions are considered. These pseudo-samples are obtained by reweighting the baseline prediction to alternative predictions of $A_{\mathrm{C}}^{t\bar{t}}$. A dedicated uncertainty is assigned to cover it.

6.2.3 Results

The unfolded charge asymmetry for the inclusive and differential measurements is presented in the semileptonic and dileptonic channels, as well as for the combination of the two in Figure 6.3, together with the SM prediction. The results are collected in Table 6.1 with a breakdown of the total uncertainty.

The measured asymmetry is consistent with the NNLO QCD + NLO EW calculations [239]. These calculations are obtained using m_t = 172.5 GeV, with the PDF4LHCLUX17 PDF set. A dynamical renormalization and factorization scale [240] is used, with the nominal value μ_0 chosen as $H_{\mathrm{T}}/4$, where $H_{\mathrm{T}} = \sqrt{m_t^2 + p_{\mathrm{T},t}^2} + \sqrt{m_{\bar{t}}^2 + p_{\mathrm{T},\bar{t}}^2}$. The scale uncertainty band is derived by performing variations of the scale by a factor 2 around the central value μ_0, and indicates the maximum and minimum values that the asymmetry takes under such variations.

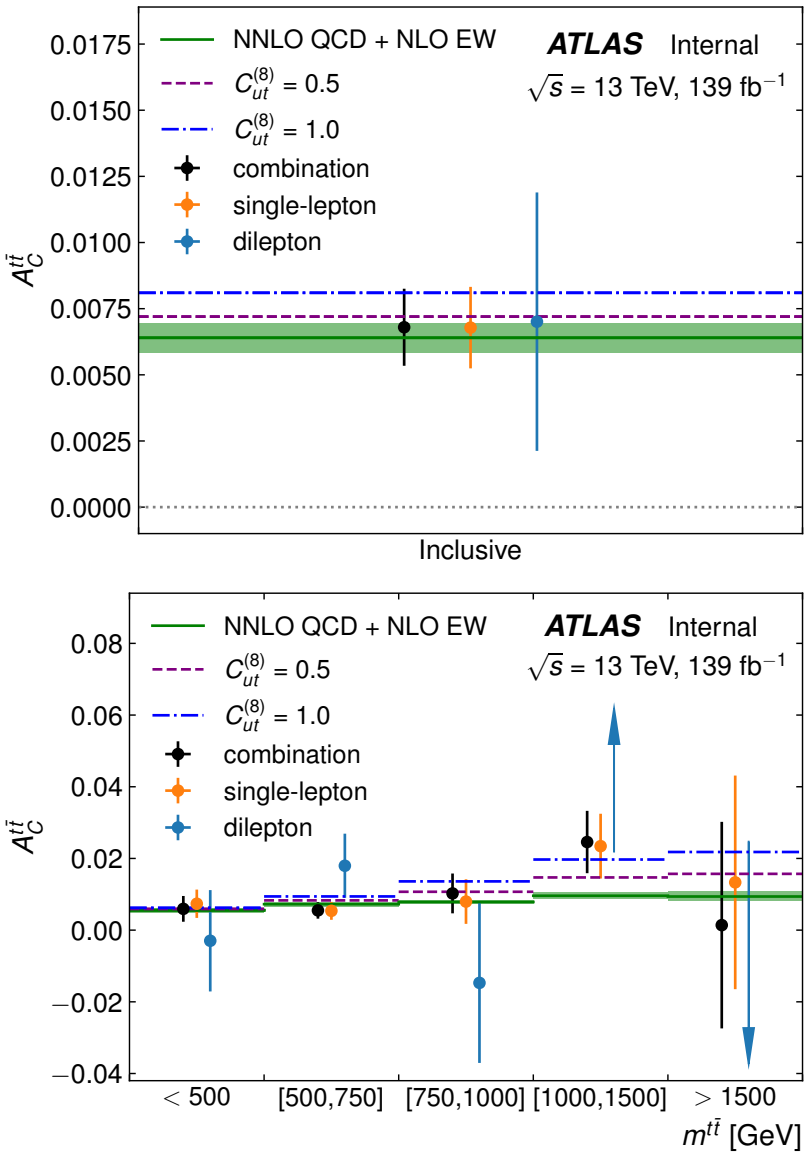


Figure 6.3: The unfolded inclusive and differential charge asymmetries as a function of the $t\bar{t}$ invariant mass on the upper and lower panels, respectively. The vertical bar represents the total uncertainty, accounting for the systematic and statistic uncertainties. The impact of the linear term of the $C_{ut}^{(8)}$ Wilson coefficient on the $A_{\mathrm{C}}^{t\bar{t}}$ prediction for two different values is shown as dashed lines. In the dileptonic channel the last two $m_{t\bar{t}}$ bins have low statistics; the vertical bars representing the total uncertainties are shown, while the $A_{\mathrm{C}}^{t\bar{t}}$ results, which are outside of the visible scale, are indicated by the direction of the associated arrows.

Table 6.1: Results with statistical, systematic and the total uncertainties for the inclusive and $m_{t\bar{t}}$ differential $A_{\mathrm{C}}^{t\bar{t}}$ measurements. The statistical uncertainty is obtained by unfolding the data, but excluding all the nuisance parameters. The systematic uncertainty is evaluated by subtracting in quadrature the statistical uncertainty from the total uncertainty. The Standard Model predictions are calculated at NNLO in QCD and NLO in EW theory.

		Data 139 fb^{-1}				SM prediction
	Channel	$A_{\mathrm{C}}^{t\bar{t}}$	Stat.	Syst.	Total unc.	
Inclusive	Semileptonic	0.0068	0.0011	0.0011	0.0015	$0.0064^{+0.0005}_{-0.0006}$
	Dileptonic	0.0070	0.0034	0.0035	0.0049	
	Combination	0.0068	0.0010	0.0010	0.0015	
$<$ 500 GeV	Semileptonic	0.0074	0.0028	0.0028	0.0039	$0.0055^{+0.0006}_{-0.0006}$
	Dileptonic	-0.0030	0.0114	0.0084	0.0141	
	Combination	0.0059	0.0027	0.0024	0.0036	
500 – 750 GeV	Semileptonic	0.0054	0.0020	0.0015	0.0025	$0.0072^{+0.0006}_{-0.0006}$
	Dileptonic	0.0180	0.0061	0.0066	0.0089	
	Combination	0.0055	0.0019	0.0013	0.0023	
750 – 1000 GeV	Semileptonic	0.0080	0.0048	0.0040	0.0062	$0.0079^{+0.0004}_{-0.0006}$
	Dileptonic	-0.0147	0.0188	0.0120	0.0223	
	Combination	0.0102	0.0046	0.0030	0.0056	
1000 – 1500 GeV	Semileptonic	0.0234	0.0075	0.0050	0.0090	$0.0096^{+0.0009}_{-0.0009}$
	Dileptonic	0.0663	0.0371	0.0244	0.0444	
	Combination	0.0246	0.0074	0.0045	0.0087	
$>$ 1500 GeV	Semileptonic	0.0133	0.0288	0.0076	0.0298	$0.0094^{+0.0015}_{-0.0011}$
	Dileptonic	-0.1313	0.1444	0.0590	0.1560	
	Combination	0.0014	0.0280	0.0068	0.0288	

The MC integration uncertainty, typically smaller than the scale uncertainty, is added in quadrature.

The sensitivity of the $A_{\mathrm{C}}^{t\bar{t}}$ is greater with increasing values of the mass of the $t\bar{t}$ invariant system, as expected. This is caused by the suppression of the gluon fusion initiated processes, which is enhanced at higher energies and has the effect of diluting the asymmetry generated through quark-quark and quark-gluon processes. The sensitivity of the SMEFT contributions also grows with $m_{t\bar{t}}$. This is a feature of the four-fermion operators.

On the other hand, the sensitivity to the asymmetry is lower for the dileptonic channel, mostly caused by the limited branching ratio and the migrations due to difficult reconstruction of the two neutrinos in the final-state. This leads to a modest improvement in the combination, specially relevant at low $m_{t\bar{t}}$ where the systematic uncertainties play a larger role.

In light of these results, we can say that there is a 4.5σ evidence of a very subtle SM effect in the asymmetry, since it is found to be below 1% in all bins considered. The statistical uncertainty remains sizeable, so there is room for improvement even if the systematic uncertainties remain the same. Finally, the theory uncertainty is clearly subdominant, which means that it will not dominate until very deep into the HL-LHC program.

6.3 SMEFT interpretation

In this section, the charge asymmetry measurements are interpreted in the context of the SMEFT. The difference between the measured and predicted $A_{\mathrm{C}}^{t\bar{t}}$ is translated in terms of contributions of dimension-six operators that encode the effect of new physics phenomena at a scale Λ beyond the direct reach of the

experiment.

In the SMEFT interpretation, the following 14 four-fermion dimension-six operators from the Warsaw basis [104] are considered, where eight have a RR and LL chiral structures:

$$
\begin{aligned}
O_{Qq}^{(8,1)} &= \left(\overline{Q}\gamma_\mu\lambda^A Q\right)\left(\overline{q}\gamma^\mu\lambda^A q\right) , & O_{Qq}^{(1,1)} &= \left(\overline{Q}\gamma_\mu Q\right)\left(\overline{q}\gamma^\mu q\right) , \\
O_{Qq}^{(8,3)} &= \left(\overline{Q}\gamma_\mu\tau^I\lambda^A Q\right)\left(\overline{q}\gamma^\mu\tau^I\lambda^A q\right) , & O_{Qq}^{(1,3)} &= \left(\overline{Q}\gamma_\mu\tau^I Q\right)\left(\overline{q}\gamma^\mu\tau^I q\right) , \\
O_{tu}^{(8)} &= \left(\overline{t}\gamma_\mu\lambda^A t\right)\left(\overline{u}\gamma^\mu\lambda^A u\right) , & O_{tu}^{(1)} &= \left(\overline{t}\gamma_\mu u\right)\left(\overline{u}\gamma^\mu t\right) , \\
O_{td}^{(8)} &= \left(\overline{t}\gamma_\mu\lambda^A t\right)\left(\overline{d}\gamma^\mu\lambda^A d\right) , & O_{td}^{(1)} &= \left(\overline{t}\gamma_\mu d\right)\left(\overline{d}\gamma^\mu t\right) ,
\end{aligned}
\tag{6.2}
$$

and six further operators with LR structure:

$$
\begin{aligned}
O_{Qu}^{(8)} &= \left(\overline{Q}\gamma_\mu\lambda^A Q\right)\left(\overline{u}\gamma^\mu\lambda^A u\right) , & O_{Qu}^{(1)} &= \left(\overline{Q}u\right)\left(\overline{u}Q\right) , \\
O_{Qd}^{(8)} &= \left(\overline{Q}\gamma_\mu\lambda^A Q\right)\left(\overline{d}\gamma^\mu\lambda^A d\right) , & O_{Qd}^{(1)} &= \left(\overline{Q}d\right)\left(\overline{d}Q\right) , \\
O_{Qt}^{(8)} &= \left(\overline{Q}\gamma_\mu\lambda^A Q\right)\left(\overline{t}\gamma^\mu\lambda^A t\right) , & O_{Qt}^{(1)} &= \left(\overline{Q}t\right)\left(\overline{t}Q\right) ,
\end{aligned}
\tag{6.3}
$$

where q and Q stands for the weak left-handed doublets of the first two and the third generation, respectively. u, d represent the weak right-handed singlets of the first and second generation, and t is the top right-handed singlet. Color singlet operators are denoted with (1) and color octet operators with (8). The O_{tG} operator is also considered as it affects the $t\bar{t}G$ vertex:

$$
O_{tG} = \left(\overline{Q}\sigma^{\mu\nu}\lambda^A t\right)\phi G_{\mu\nu}^A . \tag{6.4}
$$

6.3.1 Methodology

Limits are derived individually, so only one Wilson coefficient C is different from zero in the theory calculation. The predicted $A_{\mathrm{C}}^{t\bar{t}}$ is written explicitly in terms of the cross-sections coming from both the SM and EFT sectors:

$$A_{\mathrm{C}}^{t\bar{t}} = \frac{(\sigma^{+}_{\mathrm{SM}} - \sigma^{-}_{\mathrm{SM}}) + \left(\sigma^{+}_{\Lambda^{-2}} - \sigma^{-}_{\Lambda^{-2}}\right) \cdot C + \left(\sigma^{+}_{\Lambda^{-4}} - \sigma^{-}_{\Lambda^{-4}}\right) \cdot C^2}{(\sigma^{+}_{\mathrm{SM}} + \sigma^{-}_{\mathrm{SM}}) + \left(\sigma^{+}_{\Lambda^{-2}} + \sigma^{-}_{\Lambda^{-2}}\right) \cdot C + \left(\sigma^{+}_{\Lambda^{-4}} + \sigma^{-}_{\Lambda^{-4}}\right) \cdot C^2} \,, \tag{6.5}$$

where σ^{+} (σ^{-}) terms represent the cross-section of $t\bar{t}$ events with $\Delta y = y_t - y_{\bar{t}} > 0$ ($\Delta y < 0$). In this way, the SM-EFT and EFT-EFT interference terms, proportional to Λ^{-2} and Λ^{-4} respectively, are considered and taken into account in the total normalization. For $C = 0$, the predicted $A_{\mathrm{C}}^{t\bar{t}}$ coincides with the SM prediction.

The contributions to the cross-section from dimension-six operators are obtained numerically using the SMEFTNLO UFO model from Ref. [241] in the MadGraph5_aMC@NLO package [72]. As the name suggests, the model provides NLO accuracy. A full set of parameterizations was provided based on the work of Ref. [242]. For the SM predictions, we adopt the most precise calculations currently available, which correspond to a NNLO QCD + NLO EW calculation introduced in the previous section [239].

To quantify the level of agreement between the measured and predicted $A_{\mathrm{C}}^{t\bar{t}}$, a $\chi 2$ test-statistic is employed. The uncertainty associated to the measurement is considered and added in quadrature to the uncertainty on the prediction of the SM cross-section, which accounts for different choices of the renormalization and factorization scales, the MC integration uncertainty and the choice of an alternative PDF set, NNPDF31LUX. The asymmetric nature of the theory uncertainties are properly treated in the calculation of the $\chi 2$, following the prescription given in Ref. [243]. From the up and down variations of the absolute

uncertainty $\delta^{\text{up,dw}}$, the symmetrized uncertainty δ^S and the asymmetry $\mathcal{A}$ are defined as:

$$\delta^S = \frac{\delta^{\text{up}} + \delta^{\text{dw}}}{2} \ , \quad \mathcal{A} = \frac{\delta^{\text{up}} - \delta^{\text{dw}}}{\delta^{\text{up}} + \delta^{\text{dw}}} \ . \tag{6.6}$$

A sensible form for $\chi 2$ from asymmetric errors is therefore given by:

$$\chi 2 = \left(\frac{v}{\delta^S}\right)^2 \left(1 - 2\mathcal{A}\left(\frac{v}{\delta^S}\right) + 5\mathcal{A}^2\left(\frac{v}{\delta^S}\right)^2\right) \ , \tag{6.7}$$

where v represents the residuals, *i.e.* the difference between the measured and predicted $A_{\text{C}}^{t\bar{t}}$.

Limits are constrained at 68% and 95% CL, which corresponds to the values of C associated to the $\chi 2$ minimum +1 and +4, respectively. They are derived in two different scenarios: considering only terms proportional to Λ^{-2} (linear in C accounting for SM-EFT interferences) and also considering dimension-six operators squared terms of order Λ^{-4} (quadratic in C encoding EFT-EFT interferences).

This procedure is illustrated in Figure 6.4, where the comparison between the ATLAS measurement of the inclusive $A_{\text{C}}^{t\bar{t}}$ and the theory prediction is displayed in the lower panel. The associated $\chi 2$ distributions are shown in the upper left- and right-hand panels for the fits including only Λ^{-2} and $\Lambda^{-2} + \Lambda^{-4}$ terms, respectively. The intersection between the theory prediction and the green (yellow) experimental uncertainty band yields the 68% (95%) CL bounds on the Wilson coefficient. In this case the $C_{ut}^{(8)}$ coefficient, which is taken as a benchmark. The corresponding C values are represented by vertical dashed lines in the $\chi 2$ distributions.

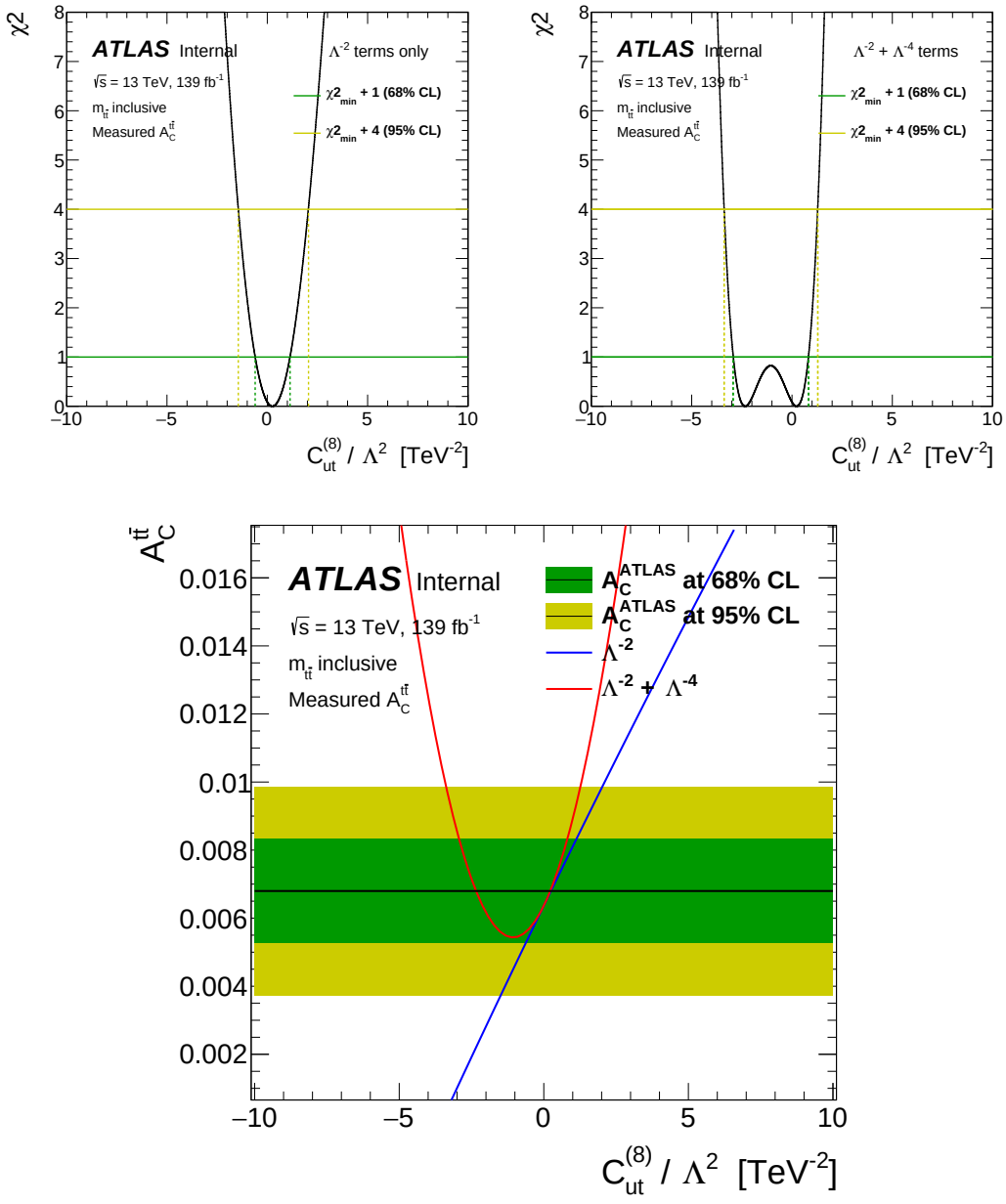


Figure 6.4: $\chi 2$ distribution from which limits on $C_{ut}^{(8)}$ operator are derived when considering (upper-left panel) only SM-EFT terms and (upper-right panel) SM-EFT and EFT-EFT terms in the EFT contribution. Bottom panel shows the comparison between ATLAS and predicted $A_{\mathrm{C}}^{t\bar{t}}$ values, with the 68% and 95% Confidence Level bands drawn in green and yellow, respectively. Inclusive $t\bar{t}$ events.

6.3.2 Limits on dimension-six operators

The procedure described above for the inclusive measurement is repeated in several invariant $t\bar{t}$ mass intervals: $m_{t\bar{t}} \in [0, 500, 750, 1000, 1500, \infty]$ GeV. A simultaneous fit where all the differential measurements are included is performed as well. In this case, the $\chi 2$ calculation:

$$\chi 2 = V^T \times C^{-1} \times V \tag{6.8}$$

incorporates the inverse of the total covariance matrix $C = C^{\text{exp}} + C^{\text{theo}}$, that takes into account bin-to-bin correlations induced by the presence of the systematic uncertainties and the unfolding in the experimental measurement through C^{exp}. It also includes the covariance in the theory prediction C^{theo}. V represents the vector of residuals.

The theory covariance matrix is evaluated from the theory predictions $A_{\text{C}}^{t\bar{t}}$ and the associated up and down uncertainties $\delta^{\text{up,dw}}$ as:

$$C_{ij}^{\text{theo}} = \rho_{ij}^{S} \times \delta_i^{S} \times \delta_j^{S} \ , \tag{6.9}$$

where δ_i^S represents the symmetrized uncertainty of bin i and ρ_{ij}^S encodes the correlation between bins i and j. The correlation can take the values $+1/-1$, depending on the signs of the up and down variations:

$$\rho_{ij}^{S_{\text{up}}} = \text{sgn}\left(\delta_i^{\text{up}}\right) \times \text{sgn}\left(\delta_j^{\text{up}}\right) \ , \tag{6.10}$$

$$\rho_{ij}^{S_{\text{dw}}} = \text{sgn}\left(\delta_i^{\text{dw}}\right) \times \text{sgn}\left(\delta_j^{\text{dw}}\right) \ . \tag{6.11}$$

If $\rho_{ij}^{S_{\text{up}}}$ and $\rho_{ij}^{S_{\text{dw}}}$ coincide, then $\rho_{ij}^{S} = \rho_{ij}^{S_{\text{up}}}$. Otherwise, the sign at each bin is given by $\max\left(|\delta_i^{\text{up}}, \delta_i^{\text{dw}}|\right)$, and the final sign of ρ_{ij}^S is determined by the product of the

two signs.

All in all, the 68% and 95% CL bounds on $C_{ut}^{(8)}/\Lambda^2$ from the measurement of the $t\bar{t}$ charge asymmetry presented in this chapter are shown in Figure 6.5. The different markers correspond to the inclusive $A_{\mathrm{C}}^{t\bar{t}}$ measurement, the different bins of the differential measurement as a function of $m_{t\bar{t}}$, and the combined fit on the differential result. Limits derived from the LHC combination of charge asymmetry measurements at $\sqrt{s}$ = 8 TeV and from the Tevatron combination of the forward-backward asymmetry measurements in 1.96 TeV $p\bar{p}$ collisions are given for reference in the lower panel. The nominal result, obtained with a linear parameterization of $C_{ut}^{(8)}/\Lambda^2$ is shown in blue. The result of a fit that takes into account also the Λ^{-4} term due to the contribution of the dimension-six operators squared is shown in red.

The presence of double wells (two local minima) in the $\chi 2$ distribution, as shown in the upper right panel of Figure 6.4, is possible in fits that include the Λ^{-4} term. It induces an asymmetry in the 68% (95%) CL interval obtained when the maximum value between the two wells is lower than $\chi 2_{\mathrm{min}} + 1(4)$. In these cases, since the size of the interval does not change with the local minimum picked, we only show one of the best-fit points as a reference. When the maximum value of the distribution is larger than $\chi 2_{\mathrm{min}} + 1(4)$, as happens for the $[1000, 1500\}$ GeV $m_{t\bar{t}}$ bin, the 68% (95%) CL interval splits into two separate sub-intervals[1].

We also find that the inclusive measurement, with an uncertainty of 1.5 per mille, yields a tight bound: $-1.44 < C_{ut}^{(8)}/\Lambda^2 < 1.99$ TeV^{-2} at 95% CL in the linear fit. Despite the large dilution by gluon-initiated $t\bar{t}$ production at the LHC at $\sqrt{s}$ = 13 TeV, this bound based on a single measurement improves considerably on the limits derived from the LHC 8 TeV combination [244] and from the Tevatron

[1] A complete set of figures reflecting the trend of the $\chi 2$ distribution, for the linear and quadratic fits, as a function of the $m_{t\bar{t}}$ bin is available in App. A.

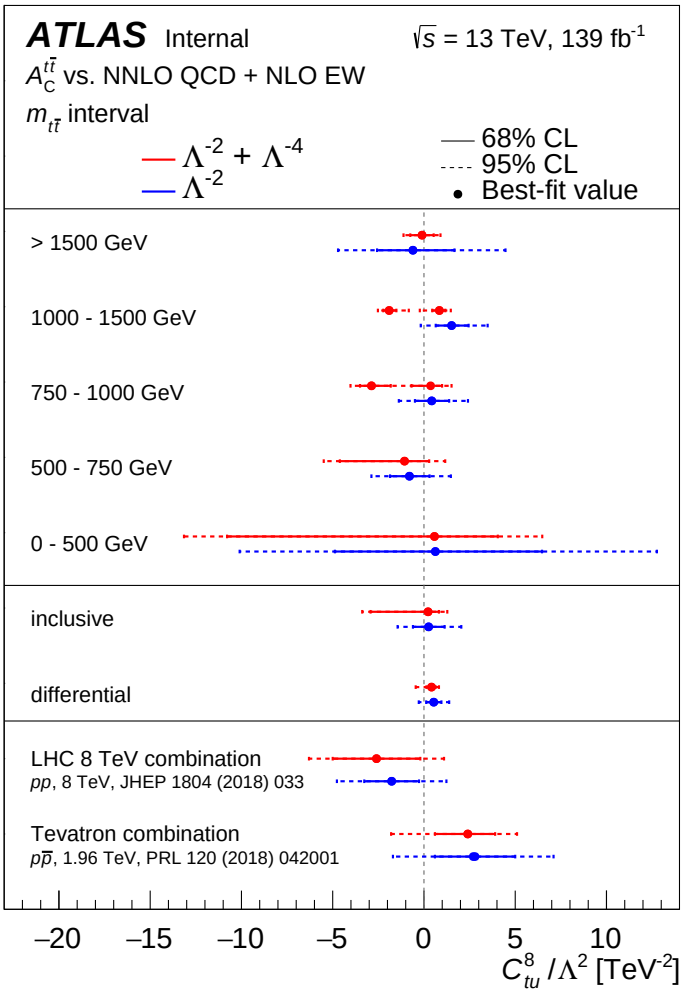


Figure 6.5: Individual 68% and 95% Confidence Level limits on the Wilson coefficient $C_{ut}^{(8)}/\Lambda^2$ in the Standard Model Effective Field Theory. The bounds are derived from the measured charge asymmetry presented in this chapter, combining the semileptonic and dileptonic channels. The theory uncertainty corresponds to the NNLO QCD + NLO EW calculation of Ref. [239]. The impact of dimension-six operators is parameterized at NLO accuracy in QCD using the SMEFTNLO package in MadGraph5_aMC@NLO. Bounds are also shown from the forward-backward asymmetry measurements in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV at the Tevatron and the charge asymmetry measurements in pp collisions at a center-of-mass energy of 8 TeV in LHC Run 1.

combination [245].

The bounds from the differential measurements reflect the interplay between the sensitivity, that increases strongly at higher $m_{t\bar{t}}$, and the uncertainty, that grows from 2 – 3 per mille in the lowest bin to 2.9% in the highest mass bin. The tightest limit from the linear parametrization is obtained in the mass bin from 1 to 1.5 TeV. In the quadratic parametrization, it is found to be in the highest mass bin. In general, the combined fit including quadratic terms to all differential measurements yields the tightest bound: $-0.39 < C_{ut}^{(8)}/\Lambda^2 < 0.78$ TeV^{-2} at 95% CL. It is more than a factor 2 better than the bound derived from the inclusive measurement, thanks to the strongly increasing sensitivity at high $m_{t\bar{t}}$.

The individual 68% and 95% CL bounds on the Wilson coefficients of all dimension-six operators listed above are presented in Figure 6.6. All intervals are also given in Table 6.2. We first note that the asymmetry affecting the 95% CL intervals in the quadratic fits are also present for most of the operators. For the inclusive measurement, the linear fit yields tighter bounds on the Wilson coefficients for octet operators, that enter at tree-level, than for the coefficients of singlet operators that only affect the $t\bar{t}$ production at NLO [246]. This does not apply to the results obtained with the quadratic fit. Regarding the bounds determined with a fit including all differential measurements, we do not find a systematic behaviour depending on the terms included in the fit or the nature of the operator considered, with one exception: the down-type operators are always less constrained than the equivalent up-type, something expected given the proton composition (the chances of colliding up-type quarks are about twice larger). But we do find that the bounds derived through the fit to all differential measurements are generally about a factor two stronger than those derived from the inclusive measurement. Overall, we see that the charge asymmetry is sensitive to four-quark operator coefficients and C_{tg} in the range of $[-2, 2]$ (TeV/Λ^{-2}) at

95% CL (except for $C_{qd}^{(8)}$).

Compared to global fits of the top quark sector [246–248] and fits including top, Higgs and electro-weak data [242, 249] the bounds found in this analysis are of the same order of magnitude as the individual bounds reported on the basis of a much larger data set (that includes, for the more recent fits, a preliminary result for the inclusive charge asymmetry in the semileptonic channel included in this work). Often, the bounds from the differential analysis are significantly better than the global bounds, indicating that the inclusion of these results in future global fits can improve the global result, by disentangling some of the poorly constrained combinations of operator coefficients.

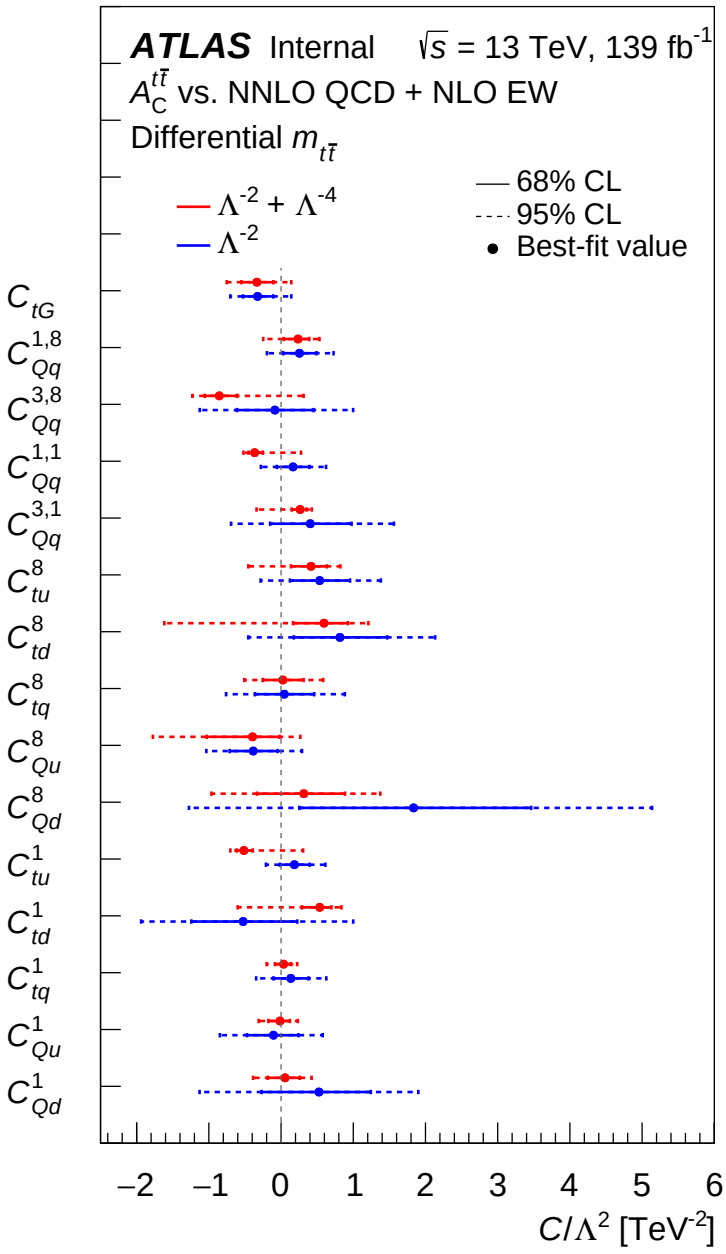


Figure 6.6: Individual 68% and 95% Confidence Level limits on the Wilson coefficients C_i/Λ^2 in the Standard Model Effective Field Theory from the differential charge asymmetry measurements versus $m_{t\bar{t}}$. The bounds are derived from the measured charge asymmetry presented in this chapter, combining the semileptonic and dileptonic channels. The theory uncertainty corresponds to the NNLO QCD + NLO EW calculation of Ref. [239]. The impact of dimension-six operators is parameterized at NLO accuracy in QCD using the SMEFTNLO package in MG5_aMCNLO.

Table 6.2: Individual 68% and 95% Confidence Level bounds on the relevant Wilson coefficients of the Standard Model Effective Field Theory in units of TeV^{-2}. The bounds are derived from the charge asymmetry measurement presented in this chapter, combining the semileptonic and dileptonic channels. The upper block corresponds to the bounds derived from the inclusive measurement, the second block to the bounds from the differential charge asymmetry measurements versus $m_{t\bar{t}}$. In both cases, the experimental uncertainties are accounted for, in the form of the complete covariance matrix that keeps track of correlations between bins for the differential measurement. The theory uncertainty from the NNLO QCD + NLO EW calculation is included by explicitly varying the renormalization and factorization scales, or the parton density functions, in the calculation and registering the variations in the intervals.

Individual bounds (in units of TeV^{-2}) from the inclusive $A_C^{t\bar{t}}$ measurement.				
	linear fit including terms $\propto \Lambda^{-2}$		*quadratic* fit with also $(D6)^2$ terms $\propto \Lambda{-}4$	
operator coefficient	68% CL	95% CL	68% CL	95% CL
C_{tG}/Λ^2	[-0.54,0.37]	[-0.89,1.03]	[-0.56,0.37]	[-0.97,0.99]
C_{qq}^{81}/Λ^2	[-0.32,0.61]	[-0.78,1.10]	[-0.37,0.51]	[-3.47,0.84]
C_{qq}^{83}/Λ^2	[-0.88,1.63]	[-2.10,2.94]	[-1.97,0.90]	[-2.41,1.33]
C_{qq}^{11}/Λ^2	[-1.24,2.29]	[-2.97,4.13]	[-0.60,0.48]	[-0.79,0.67]
C_{qq}^{13}/Λ^2	[-6.74,3.73]	[-12.1,8.99]	[-0.51,0.57]	[-0.70,0.75]
C_{ut}^{8}/Λ^2	[-0.60,1.13]	[-1.44,2.05]	[-2.93,0.82]	[-3.38,1.28]
C_{dt}^{8}/Λ^2	[-0.97,1.80]	[-2.30,3.26]	[-4.34,1.28]	[-5.05,1.99]
C_{qt}^{8}/Λ^2	[-0.96,0.54]	[-1.73,1.29]	[-2.77,0.45]	[-3.23,0.93]
C_{qu}^{8}/Λ^2	[-1.06,0.59]	[-1.90,1.41]	[-3.63,0.51]	[-4.17,1.06]
C_{qd}^{8}/Λ^2	[-3.71,2.08]	[-6.61,5.02]	[-3.36,1.29]	[-4.41,2.35]
C_{ut}^{1}/Λ^2	[-0.81,1.49]	[-1.93,2.69]	[-0.80,0.53]	[-1.02,0.75]
C_{dt}^{1}/Λ^2	[-17.6,32.7]	[-41.8,59.4]	[-0.98,0.96]	[-1.32,1.30]
C_{qt}^{1}/Λ^2	[-1.25,2.31]	[-3.00,4.16]	[-0.34,0.47]	[-0.55,0.68]
C_{qu}^{1}/Λ^2	[-2.40,4.40]	[-5.76,7.92]	[-0.43,0.53]	[-0.68,0.78]
C_{qd}^{1}/Λ^2	[-53.0,88.0]	[-134,152]	[-0.72,0.74]	[-1.10,1.12]
Individual bounds (in units of TeV^{-2}) from the differential $A_C^{t\bar{t}}$ measurement versus $m_{t\bar{t}}$.				
	linear fit including terms $\propto \Lambda^{-2}$		*quadratic* fit with also $(D6)^2$ terms $\propto \Lambda{-}4$	
operator coefficient	68% CL	95% CL	68% CL	95% CL
C_{tG}/Λ^2	[-0.53,-0.11]	[-0.70,0.14]	[-0.55,-0.11]	[-0.75,0.14]
C_{qq}^{81}/Λ^2	[0.03, 0.49]	[-0.20,0.73]	[0.04, 0.39]	[-0.25,0.53]
C_{qq}^{83}/Λ^2	[-0.61, 0.45]	[-1.13,1.00]	[-1.06,-0.61]	[-1.23,0.31]
C_{qq}^{11}/Λ^2	[-0.06, 0.39]	[-0.28,0.62]	[-0.45,-0.25]	[-0.52,0.28]
C_{qq}^{13}/Λ^2	[-0.15, 0.98]	[-0.69,1.56]	[0.15, 0.35]	[-0.34,0.43]
C_{ut}^{8}/Λ^2	[0.12, 0.96]	[-0.28,1.38]	[0.14, 0.63]	[-0.45,0.82]
C_{dt}^{8}/Λ^2	[0.18, 1.47]	[-0.45,2.13]	[0.17, 0.92]	[-1.62,1.21]
C_{qt}^{8}/Λ^2	[-0.36, 0.46]	[-0.76,0.88]	[-0.25, 0.31]	[-0.51,0.58]
C_{qu}^{8}/Λ^2	[-0.71,-0.05]	[-1.03,0.29]	[-1.03,-0.02]	[-1.78,0.27]
C_{qd}^{8}/Λ^2	[0.25, 3.46]	[-1.28,5.14]	[-0.33, 0.89]	[-0.96,1.37]
C_{ut}^{1}/Λ^2	[-0.02, 0.40]	[-0.21,0.61]	[-0.62,-0.39]	[-0.70,0.31]
C_{dt}^{1}/Λ^2	[-1.24, 0.22]	[-1.94,1.00]	[0.29, 0.70]	[-0.60,0.84]
C_{qt}^{1}/Λ^2	[-0.10, 0.38]	[-0.35,0.63]	[-0.08, 0.14]	[-0.20,0.22]
C_{qu}^{1}/Λ^2	[-0.47, 0.24]	[-0.85,0.58]	[-0.18, 0.12]	[-0.31,0.23]
C_{qd}^{1}/Λ^2	[-0.27, 1.24]	[-1.13,1.90]	[-0.19, 0.26]	[-0.39,0.42]

Chapter 7

Determination of the running bottom quark mass from Higgs boson decay rates

In the course of the studies of the top quark mass, a new opportunity to investigate the scale evolution of quark masses presented itself. In this chapter, an extraction of the bottom quark $\overline{MS}$ mass from Higgs decay rates is presented. This extraction yields $m_b(m_H)$, the bottom quark mass at the scale of the Higgs boson mass. A comparison of the result with low-scale determinations, collected in the world average for $m_b(m_b)$, and the measurements of $m_b(m_Z)$ in e^+e^- collisions at the Z-pole, yields the first observation of the scale evolution or running of the bottom quark mass, and is expected to develop into a precision test of this feature of the SM in the next decades. The content of this chapter is mainly extracted from Ref. [250].

7.1 Introduction

As we have extensively discussed in Section 1.4.1, the fundamental parameters of the SM Lagrangian are renormalized in order to properly describe the physical observables as they are measured in the experiments. This includes not only the coupling constant that govern interaction rates at colliders, but also particle masses. As a consequence, they depend on a dimensionful renormalization scale and on the RS adopted. In the following, we will take the $\overline{MS}$ RS as a benchmark.

The dependence of the SM parameters with the energy is governed by the RGE, which can be calculated at 5-loop ($\mathcal{O}(\alpha_S^5)$) precision now [251–253]. It can be tested by performing different measurements at different energy scales. An example is the strong coupling, which dependence with the scale has been experimentally observed over a broad energy range. Not only the strong coupling has been studied, but also quark masses. Low-energy measurements of the bottom quark $\overline{MS}$ mass have been performed in B-factories, reaching a very high precision. It has been also measured at higher energies at LEP, as we shall see in the next section. The new measurement based on $\sqrt{s}$ = 13 TeV data at LHC is presented in Section 7.3.1.

7.2 Low-energy measurements

The lowest-energy determination of the bottom quark mass is provided by the PDG [26] and results from the combination of several measurements. The typical energy of such measurements is relatively low, of the order of the bottom quark production threshold ($\sim m_b$). The most precise extractions rely on the measurement of the mass of bottomonium bound states and the $e^+e^- \to$ hadrons cross-sections as experimental inputs, in combination with QCD sum rules and

perturbative QCD calculations [254–264]. The average includes also inputs from HERA [265] and the Babar and Belle experiments at the B-factories [266, 267].

The world average quoted by the PDG is given in the $\overline{MS}$ scheme at the scale of the bottom mass:

$$m_b(m_b) = 4.18^{+0.03}_{-0.02} \text{ GeV} , \tag{7.1}$$

with an impressive relative uncertainty of 1%. Since the measurements used to determine m_b at higher renormalization scales are different from those included here, we can safely assume that they are fairly independent.

The determination of m_b at much higher energy scales became possible at LEP and SLC, thanks to the fact that event-shapes and jet-rates variables are sensitive to subleading mass effects. In Ref. [268] a practical method to extract m_b from events with hadronic decays of Z bosons (Z-pole data) was proposed. The mass of the bottom quark is therefore measured at the scale of the Z boson mass. Three independent groups completed the required NLO calculations of the related variables to make this extraction method possible [269–273].

In this analysis, we take advantage of the most precise measurements available from Z-pole data, summarized in Table 7.1. The first measurement of this type was performed by the DELPHI collaboration [274, 275], followed by the ALEPH [276] and OPAL [277] collaborations, all of them profiting from LEP data as well. Similar measurements were performed with Z-pole data collected in the SLAC accelerator by the SLD collaboration [278, 279].

All these results are combined with the Convino package [280]. It allows to combine measurements with asymmetric uncertainties and taking into account correlations between the systematic uncertainties of the different measurements considered. On the one hand, theory uncertainties are assumed to be 100% correlated among the measurements, as they all rely on the NLO prediction of

three-jet rates. On the other hand, the correlation between the experimental uncertainties is taken to be 50%. This choice responds to the observation of a correlated hadronization uncertainty, that is found to be of the similar magnitude as the statistical and purely experimental uncertainties in DELPHI and ALEPH. The robustness of the combination is tested by considering varied correlations between experimental uncertainties of ±20% around the default value. The combined m_b (m_Z) value found in both cases is within 40 MeV of the result obtained with the default choice.

We obtain the following nominal value:

$$m_b(m_Z) = 2.82 \pm 0.28 \text{ GeV}. \tag{7.2}$$

Table 7.1: Measurements of the bottom quark $\overline{MS}$ mass at the renormalization scale $\mu = m_Z$, from three-jet rates with bottom quarks in e^+e^- collisions at the Z-pole at LEP and SLD. For ALEPH and DELPHI the hadronization uncertainty is added in quadrature with the experimental uncertainty to yield the total systematic uncertainty.

experiment	$m_b(m_Z)$ [GeV]
ALEPH[276]	3.27 ± 0.22 (stat) ± 0.44 (syst) ± 0.16 (theo)
DELPHI[274, 275]	2.85 ± 0.18 (stat) ± 0.23 (syst) ± 0.12 (theo)
OPAL[277]	2.67 ± 0.03 (stat) $^{+0.29}_{-0.37}$ (syst) ± 0.19 (theo)
SLD[278, 279]	2.56 ± 0.27 (stat) $^{+0.28}_{-0.38}$ (syst) $^{+0.49}_{-1.48}$ (theo)

7.3 Bottom quark mass from Higgs decay rates

Similarly to LEP and SLAC, the LHC has opened the possibility to reach higher energy scales. The discovery of the Higgs boson [24, 25] and the observation of the Higgs boson decay to bottom quarks [281, 282] enable an entirely new measurement of m_b. The analysis of data collected at $\sqrt{s}$ = 7, 8 and 13 TeV in Run 1 and Run 2 has led to the characterization of many of the Higgs boson production modes and decay channels predicted by the SM. The bosonic decay channel is well stablished and precisely measured. But the dominant decay channel is into pairs of bottom quarks, with a branching ratio of 58% (for a Higgs mass m_H = 125 GeV). The measurement of the $H \to b\bar{b}$ branching ratio, combining several production modes (VH, $t\bar{t}H$ and vector-boson fusion[1]), has achieved a precision of approximately 40% in Run 1 [283] and 20% in Run 2 [284–286].

In this study, we make use of the ratio $\mathcal{B}(H \to b\bar{b})/\mathcal{B}(H \to ZZ)$ of the branching ratios to bottom quarks and Z bosons. Both the ATLAS and CMS experiments provide this measurement with respect to the prediction of the SM, which is found to be:

$$\frac{\mathcal{B}(H \to b\bar{b})}{\mathcal{B}(H \to ZZ)} = \frac{\Gamma(H \to b\bar{b})}{\Gamma(H \to ZZ)} = 22.0 \pm 0.5 \ . \tag{7.3}$$

The results reported by ATLAS and CMS are based on 139 fb^{-1} and 35 fb^{-1}

[1]The dominant production mode is the gluon-gluon fusion, but it has a large background from multijet production. The most sensitive production modes are the associated production of H and a vector boson W/Z (VH) , where its leptonic decay enables an efficient triggering and the reduction of the multijet background.

of data at $\sqrt{s} = 13$ TeV [284, 286], respectively:

$$\begin{aligned}\mu^{b\bar{b}}/\mu^{ZZ} &= 0.87^{+0.22}_{-0.17}\ (\text{stat})\ ^{+0.18}_{-0.12}\ (\text{syst}) = 0.87^{+0.28}_{-0.21}\quad (\text{ATLAS}),\\ \mu^{b\bar{b}}/\mu^{ZZ} &= 0.84^{+0.27}_{-0.21}\ (\text{stat})\ ^{+0.26}_{-0.17}\ (\text{syst}) = 0.84^{+0.37}_{-0.27}\quad (\text{CMS}).\end{aligned} \tag{7.4}$$

7.3.1 Dependence of Higgs decay rates on the bottom quark mass

In order to use the results given in Eq. 7.4, we first need to establish the relation between the Higgs decay rate into $b\bar{b}$ and the bottom quark mass itself. In the limit $m_b << m_H$ and assuming standard Yukawa couplings between the Higgs boson and the bottom quark, the partial decay width can be expressed as follows:

$$\Gamma(H \to b\bar{b}) = \frac{3G_F m_H}{4\sqrt{2}\pi}\ m_b(\mu)^2\ \ (1+\delta_{\text{ew}}) \times (1+\delta_{\text{QCD}}+\delta_t+\delta_{\text{mix}})\ , \tag{7.5}$$

where G_F denotes the Fermi constant, δ_{QCD} the QCD corrections related to the scalar correlator, δ_t the QCD corrections due to the interference with $H \to gg$ diagrams that start to contribute at NNLO, δ_{ew} the EW corrections and finally δ_{mix}, the mixed QCD-EW corrections. We see that the decay width has a quadratic dependence on the bottom quark mass and can be precisely predicted. In particular, the QCD corrections δ_{QCD} are known up to N^{4}LO [287–300], the interference term δ_t at NNLO [301–303], the EW corrections δ_{ew} at NLO [304–307] and finally the mixed corrections δ_{mix} at two-loop order [308–313].

The characteristic energy scale of the process $H \to b\bar{b}$ is the Higgs boson mass. Therefore, the measurement of the bottom quark mass from the partial width of the Higgs decay into b-pairs has an associated energy scale given by the Higgs boson mass [287, 288]. It is illustrative to see now the excellent convergence of the perturbative series when computing the partial width at the renormalization

scale $\mu_R = m_H$. We first take $m_b(m_b) = 4.18$ GeV and $\alpha_S(m_Z) = 0.1179$ as inputs, and evolve the bottom quark mass and the strong coupling according to the SM RGE, with 5-loops precision and the number of active flavours n_f equal to 5. We obtain the following result for the leading QCD series:

$$1 + \delta_{\text{QCD}} = 1 + 0.2030 + 0.0374 + 0.0019 - 0.0014 \ . \tag{7.6}$$

If $\mu_R = m_b$ is used instead, we find:

$$1 + \delta_{\text{QCD}} = 1 - 0.5665 + 0.0586 + 0.1475 - 0.1274 \ . \tag{7.7}$$

We observe that the convergence of the QCD perturbative series when considering m_b as the characteristic scale of the process is poorly behaved, leading to large perturbative uncertainties due to powers of the large logarithm $\ln(m_H/m_b)$. These logarithms can be resummed to all orders in Eq. 7.6, which explains its excellent behaviour. These results support the motivation of using the partial Higgs decay rate to b-pairs to perform this measurement.

7.3.2 Numerical predictions and theory uncertainties

The dependence of the ratio of partial widths $\Gamma^{b\bar{b}}/\Gamma^{ZZ}$ with the bottom quark mass is determined as follows: $\Gamma^{b\bar{b}}$ and Γ^{ZZ} are estimated separately and numerically, and the ratio is built for several values of the bottom quark mass at the Higgs mass scale. Next, the distribution of points is parametrized in terms of $m_b(m_H)$ with a polynomial function. The fitted function takes the following form:

$$\frac{\Gamma(H \to b\bar{b})}{\Gamma(H \to ZZ)} = 2.82 \frac{m_b^2}{\text{GeV}^2} - 0.0014 \frac{m_b^4}{\text{GeV}^4} + \mathcal{O}(m_b^6) \ . \tag{7.8}$$

The uncertainty in the fitted mass value due to this parametrization is below the per mille level over the mass range of interest, and variations of the functional form and the fit range considered lead to negligible uncertainties.

The partial decay width into a b-pair $\Gamma^{b\bar{b}}$ is numerically estimated with the HDECAY package [314, 315]. The calculation accounts for N^{4}LO corrections in QCD and NLO EW corrections. The full bottom quark mass effects are also included up to NLO and logarithmic ones up to NNLO. We use m_H = 125.1 GeV and $\alpha_S(m_Z)$ = 0.01179 (according to the PDG world average) throughout this chapter.

In turn, the partial decay width into a Z-pair Γ^{ZZ} is numerically computed with the Prophecy4f package [316, 317] (version 3.0 [318]). This package includes the full QCD and EW NLO corrections to the Higgs boson decay width to four fermions, the interference contributions between different WW/ZZ channels, and all off-shell effects of intermediate W/Z bosons. The partial width $\Gamma(H \to ZZ)$ for our choice of parameters is 0.109 MeV.

The theory uncertainty of the numerical predictions due to missing higher-orders is estimated following the prescription of Ref. [319]. The renormalization scales of the terms entering in the predicted ratio are varied by a factor two and a half. Independent variations of the scales for α_S and m_b give a relative uncertainty of 0.3%. The uncertainty associated to EW corrections beyond NLO is estimated to be of the order of 0.5% in both partial widths [319].

The uncertainties on the parameters entering in the numerical predictions are also included: first, the 0.001 uncertainty on $\alpha_S(m_Z)$ quoted by the PDG is propagated through HDECAY, yielding a shift on the partial widths ratio of 0.2%. Second, the nominal value on the Higgs boson mass is varied by ±240 MeV and the predicted ratio recalculated with Prophecy4f. It leads to a variation on

Γ^{ZZ} of 3%, resulting the dominant uncertainty in the predicted ratio.

Both set of uncertainties are added in quadrature, yielding a total relative theory uncertainty of 4.4% on the $\Gamma^{b\bar{b}}/\Gamma^{ZZ}$ ratio. It translates into an uncertainty of 60 MeV on the measured $m_b(m_H)$, which is negligible in comparison with the experimental uncertainties reported by the ATLAS and CMS analysis.

7.3.3 Extraction of the bottom quark mass

The ATLAS and CMS measurements reported in Eq. 7.4 are plugged into Eq. 7.8, obtaining the following results for $m_b(m_H)$:

$$\begin{aligned} m_b(m_H) &= 2.61^{+0.32}_{-0.27}\ (\text{stat})\ ^{+0.26}_{-0.19}\ (\text{syst})\ \text{GeV} \quad (\text{ATLAS}), \\ m_b(m_H) &= 2.57^{+0.39}_{-0.35}\ (\text{stat})\ ^{+0.37}_{-0.28}\ (\text{syst})\ \text{GeV} \quad (\text{CMS}). \end{aligned} \tag{7.9}$$

Both results are combined with the Convino package, taking into account the correlations between the asymmetric systematic uncertainties. The resulting value of the bottom quark mass:

$$m_b(m_H) = 2.60^{+0.36}_{-0.31}\ \text{GeV}, \tag{7.10}$$

is the main result of this chapter. When evolved to low scales, using the 5-loop RGE implemented in the REvolver package [320], the combined result corresponds to $m_b(m_b) = 3.90^{+0.52}_{-0.47}$ GeV, fully compatible with the PDG world average of $m_b(m_b) = 4.18^{+0.03}_{-0.02}$ GeV.

7.4 The running of the bottom quark mass

We can now compare the measurements of m_b at different scales and confront them with the RGE prediction of the SM. Figure 7.1 shows the individual $m_b(m_H)$ measurements performed by the ATLAS and CMS experiments with open red markers; the combined result presented in this work is indicated with the red star; individual $m_b(m_Z)$ measurements are represented with open blue markers. The PDG world average of $m_b(m_b)$ is added with the green star and its predicted evolution to higher energy scales, according to the RGE, is shown with the solid black line (it is obtained with the REvolver code at 5-loop precision). The dark gray band stands for the theory uncertainty associated to this prediction, with the dominant uncertainties stemming from the parametric uncertainties on $m_b(m_b)$ and α_S [26] (missing higher-order uncertainty, estimated as half the difference between three-loop and five-loop precision, is negligible in comparison).

Within the current precision, we found an overall good agreement between the experimental measurements and the prediction of the SM.

7.4.1 The anomalous mass dimension

In Section 1.4.1 we saw that the evolution of any quark mass m_q with the renormalization scale is encoded in the anomalous dimension function. If we only consider the first term of the expansion in $\alpha(\mu) = \alpha_S(\mu)/\pi$ of Eq. 1.28, we obtain at LL the approximation:

$$\gamma_0 = -\beta_0 \log\left(\frac{m_q(\mu^2)}{m_q(\mu_0^2)}\right) / \log\left(\frac{\alpha_S(\mu^2)}{\alpha_S(\mu_0^2)}\right) , \tag{7.11}$$

where γ_0 and β_0 indicate the first term of the anomalous mass dimension and β function, respectively. In the SM, $\gamma_0 = -1$ and β_0 is given by N_c and N_f. In

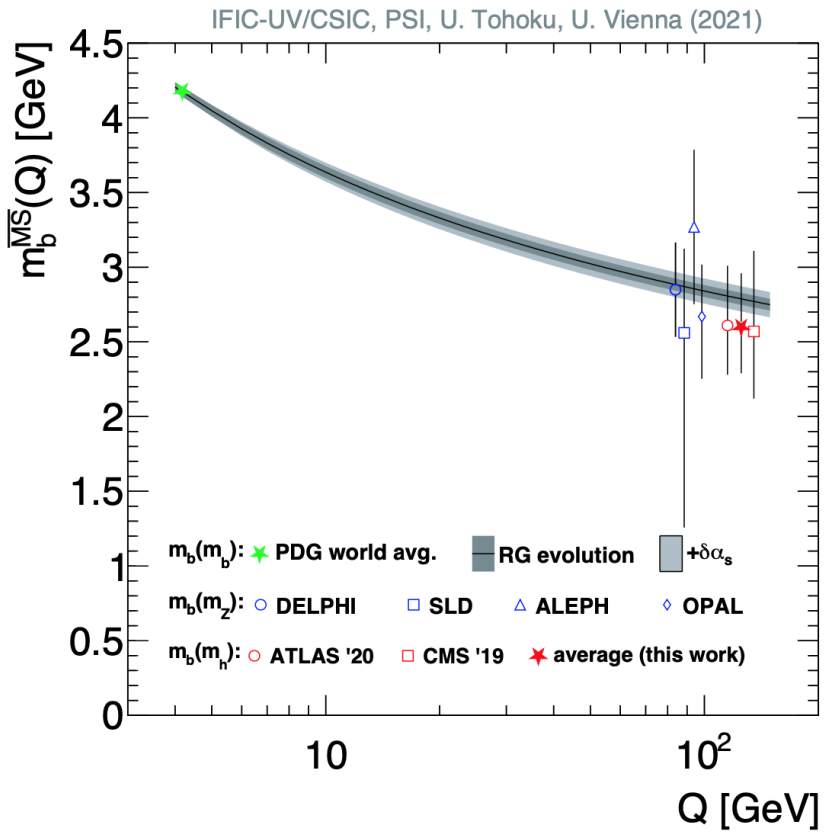


Figure 7.1: The scale evolution of the bottom quark $\overline{MS}$ mass. The measurements include the Particle Data Group world average for $m_b(m_b)$ determined at a typical scale of the bottomonium mass, the measurements of $m_b(m_Z)$ from jet rates at the Z-pole at LEP and SLC and the measurement of $m_b(m_H)$ from Higgs boson branching fractions. The prediction of the evolution of the mass is calculated at five-loop precision with REvolver [320]. The inner dark grey error band includes the effect of missing higher-orders and the parametric uncertainties from $m_b(m_b)$ and α_S from the Particle Data Group averages. The outer band with a lighter shading includes additionally the effect of a ± 0.004 variation of $\alpha_S(m_Z)$.

the five-flavour scheme and taking $N_c = 3$, $\beta_0 = (11N_c - 2N_f)/12 = 23/12$. The uncertainty on the LL approximation is given by the NLL correction, and it is found to be about 13%.

It is possible to use the values of $m_b(m_b)$, $m_b(m_Z)$ and $m_b(m_H)$ given in Eqs. 7.1, 7.2 and 7.10 to determine γ_0 experimentally. Taking as a reference scale $\mu_0 = m_b$ and $\alpha_{S,0} = 0.1179$, we obtain with REvolver the evolution of m_b and α_S to higher scales (m_Z and m_H). A $\chi 2$ fit and minimization is carried out to obtain the following best-fit value for the ratio:

$$\gamma_0/\beta_0 = -0.64 \pm 0.12\ (\text{exp})\ \pm\ 0.08\ (\text{theo})\ \pm 0.03\ (\alpha_S). \tag{7.12}$$

The uncertainty on the value of α_S is evaluated by propagating the experimental uncertainties on $\alpha_S(m_b)$, $\alpha_S(m_Z)$ and $\alpha_S(m_H)$. To reduce the SM bias, a conservative uncertainty of 0.004 is assigned to the high-scale α_S values. This covers the envelope of experimental measurements of α_S at high scale from deep-inelastic scattering and parton distribution function fits as well as EW precision fits based on the pre-averaging quoted in Ref. [321].

Taking $\beta_0 = 23/12$, we find the following value for the anomalous mass dimension:

$$\gamma_0 = -1.23 \pm 0.22\ (\text{exp})\ \pm 0.14\ (\text{theo})\ \pm 0.06\ (\alpha_S)\ , \tag{7.13}$$

in good agreement with the SM result.

It is important to stress that the LL approximation is found to be sufficiently accurate for the current measurement precision. A combined analysis of the evolution of the strong coupling and the bottom quark mass is required though to disentangle the running of α_S and m_b. This may be an interesting diagnostic

tool for new physics effects that impact the scale evolution in different ways.

7.4.2 Testing the running hypothesis

The evolution or running of the bottom quark mass with the energy scale can be further studied. The independent measurements of m_b performed at different scales enable an alternative study where the hypothesis of the SM evolution is tested. This is achieved by modifying the evolution given by the RGE through the following transformation, adapted from Ref. [322]:

$$m(\mu; x, m_b(m_b)) = x\Big[m_b^{\mathrm{RGE}}(\mu, m_b(m_b)) - m_b(m_b)\Big] + m_b(m_b), \tag{7.14}$$

where $m_b^{\mathrm{RGE}}(\mu, m_b(m_b))$ describes the RGE evolution of a reference mass $m_b(\mu_0 = m_b)$ at the scale scale μ within the SM, and x is a multiplicative factor that adjusts the scale dependence, interpolating smoothly between the no-running scenario ($x = 0$) and the beyond the SM ($x > 1$).

The values of $m_b(m_b)$ and the x parameter are mapped to obtain a grid of predicted $m_b(m_Z)$ and $m_b(m_H)$ values. They are compared to the measured $m_b(\mu)$ at the corresponding scales, given in Eqs. 7.1, 7.2 and 7.10, with a $\chi 2$ fit:

$$\chi 2(m_b(m_b), x) = \frac{\sum_{\mu_i}(m_b^{\mathrm{exp}}(\mu_i) - m(\mu_i; x, m_b(m_b)))^2}{\sigma_i^2}\ , \tag{7.15}$$

where $m(\mu_i; x, m_b(m_b)$ is given by Eq. 7.14 and the index i runs over the three energy scales considered. Only the total uncertainty quoted on the experimental results is taken into account in the $\chi 2$ formula. The theory uncertainty of the predicted evolution, even with the more conservative uncertainty on α_S of 0.004 at higher scales, is negligible in comparison.

The resulting two-dimensional $\chi 2$ distribution is marginalized onto each

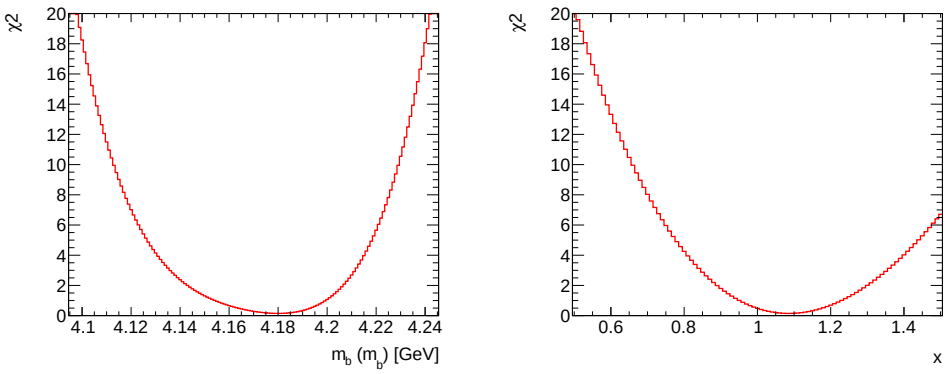


Figure 7.2: The $\chi 2$ of the fit of Eq. 7.14 to the measurements of $m_b(m_b)$, $m_b(m_Z)$ and $m_b(m_H)$, as a function of the reference bottom quark mass $m_b(m_b)$ and the factor x that multiplies the Renormalization Group Equation evolution to higher scale. The factor x interpolates smoothly between the "no-running" scenario ($x = 0$) and beyond the RGE evolution predicted within the Standard Model ($x > 1$).

parameter dimension, resulting in the one-dimension distributions displayed in Figure 7.2. The minimization gives the best-fit values, which are found to be:

$$\begin{aligned} m_b(m_b) &= 4.18^{+0.03}_{-0.02}\ \text{GeV}\ , \\ x &= 1.08 \pm 0.15\ (\text{exp})\ \pm 0.05\ (\alpha_S). \end{aligned} \tag{7.16}$$

The fitted value of the reference mas $m_b(m_b)$ is fully compatible with the PDG world average within the total uncertainty. The fit also yields a value for x compatible with the SM running within 1σ and differs by nearly seven standard deviations from the no-running scenario. A fit without the Higgs data yields $x = 1.03 \pm 0.21$, just below five standard deviations. This result represents the first demonstration of the running of the bottom quark mass.

7.5 Discussion and prospects

The use of the ratio $\mu^{b\bar{b}}/\mu^{ZZ}$ for a determination of the bottom quark mass implies a strong assumption that physics BSM does not affect the Higgs boson coupling to bottom quarks. The procedure followed by the ATLAS and CMS experiments is quite robust against certain new physics effects, such as the contribution of unknown invisible decays to the Higgs width, that cancel out in the $\mu^{b\bar{b}}/\mu^{ZZ}$ ratio. Other assumptions, *e.g.* on the Higgs boson production cross-sections, can be tested to excellent precision in measurements at the LHC and at a future e^+e^- Higgs factory. One can, however, imagine the following scenario: a shift of the bottom quark Yukawa coupling (and none of the other Higgs couplings) with respect to the SM expectation would lead to a bias in the mass measurement. The results in this chapter assume that the bottom quark Yukawa coupling is accurately predicted by the SM.

Future improvements of the Higgs branching fraction measurements are expected to rapidly reduce the uncertainties of this method. Projections for the HL-LHC [323] envisage a measurement of $\mathcal{B}(H \to b\bar{b})$ with a precision of 4.4%, reducing the uncertainty on $m_b(m_H)$ to ±60 MeV.

A future e^+e^- Higgs factory can reach sub-% precision for the Higgs boson couplings [324–326]. The recoil-mass analysis provides direct measurements of Higgs decay rates, with minimal assumptions on the total width and the production rates. The ratio $\mu^{b\bar{b}}/\mu^{WW}$ of the Higgs branching fractions to bottom quarks and W bosons (this decay channel has a much larger branching fraction than $H \to ZZ$) is expected to be precisely measured. The uncertainties of 0.86% for the 250 GeV stage of the *International Linear Collider* (ILC) and 0.47% for the complete 250 + 500 GeV program [327, 328], correspond to an uncertainty on $m_b(m_H)$ of ±12 MeV and ±6 MeV, respectively. Furthermore, future e^+e^-

colliders offer opportunities to improve the precision of $m_b(m_Z)$, either with a dedicated "GigaZ" or "TeraZ" run at the Z-pole or with radiative-return events, and to extend the analysis to $m_b(250\text{ GeV})$ [329].

At this point, the bottom mass determined from Higgs decay rates reaches a competitive precision compared to the current low-energy measurements. It is important to remark that these numbers only account for the QCD evolution without α_S uncertainties (*i.e.* the connection between high-scale and low-scale masses can only be made so precise if the improvements in the Higgs coupling measurements are accompanied by a significant improvement in α_S). With this precision, theory and parametric uncertainties in the mass determination and the evolution must be treated more carefully.

Chapter 8

Summary and conclusions

The work presented in this thesis collects several studies related to top quark precision measurements in the boosted regime. These studies have been carried out with proton-proton collision data at the center-of-mass energy of $\sqrt{s}$ = 13 TeV collected by the ATLAS experiment at the CERN LHC during the Run 2 data-taking period.

The basis for precision measurements in boosted top quark production is a detailed understanding and calibration of the ATLAS detector's jet energy and substructure response. In particular, the mass response of jets initiated by heavy particles can be characterized by studying large-R jets that capture the hadronic decay of W bosons and top quarks in $t\bar{t}$ events. In this thesis, the jet mass response of several large-R jet collections is calibrated in situ by comparing real and Monte Carlo (MC) simulated events. This allows to remove potential differences between the two after the MC-based calibration. We consider calorimeter-only based and track-assisted jets, as well as jets resulting from the reclustering of already calibrated small-R jets. The calibration is derived in terms of the ratio of the Jet Mass Scale (JMS) and Resolution (JMR) between data and MC. The Forward-Folding technique is used, which provides a robust method to

fit the jet mass response in the simulation without assuming any analytic form. A χ2-statistic test is employed in order to find the simulated jet mass response that best describes data. Statistical, detector and modelling sources of uncertainty are taken into account in this analysis.

The results of the in situ calibration using the partial Run 2 data set collected in 2015 – 2017 are summarized in Figure 8.1. The calibration factors (the inverse of the measured relative JMS/JMR) are shown in the upper and lower panels for the jet mass scale and resolution, respectively, for the three jet collections considered in a broad p_{T} range, spanning from 200 MeV up to 1 TeV. W- and top-initiated jets are represented with solid black and red markers, respectively, and the statistical and total uncertainty are given by the inner and outer vertical bars. As can be observed, the calorimeter-based and reclustered jets exhibit the best performance. The total uncertainty is below 1.5% and 20% for the relative JMS and JMR, dominated by the modelling of the parton shower in the MC. This represents the most accurate jet mass calibration obtained up to date. Most of the points are compatible with unity within 1σ, which makes unnecessary the application of an explicit calibration. In those points still far away from one, the difference is added in quadrature to the total uncertainty.

A precise interpretation for the top quark MC mass parameter in terms of a field-theoretical mass scheme is derived using a comparison of MC templates to a calculation in Soft Collinear Effective field Theory at NLL. The theory calculation incorporates the effects of hadronization to the parton-level lineshape by means of a shape function and the constraining of its two free parameters, $\Omega_{1q}^{\circ\circ}$ and x_2. The mass of lightly-groomed soft-drop jets initiated by hadronically-decaying top quarks is used, since it exhibits kinematic sensitivity to the top quark mass. For that reason, it is especially suitable to determine the relation between the top quark mass parameter in the MC generators used by ATLAS and the top mass in

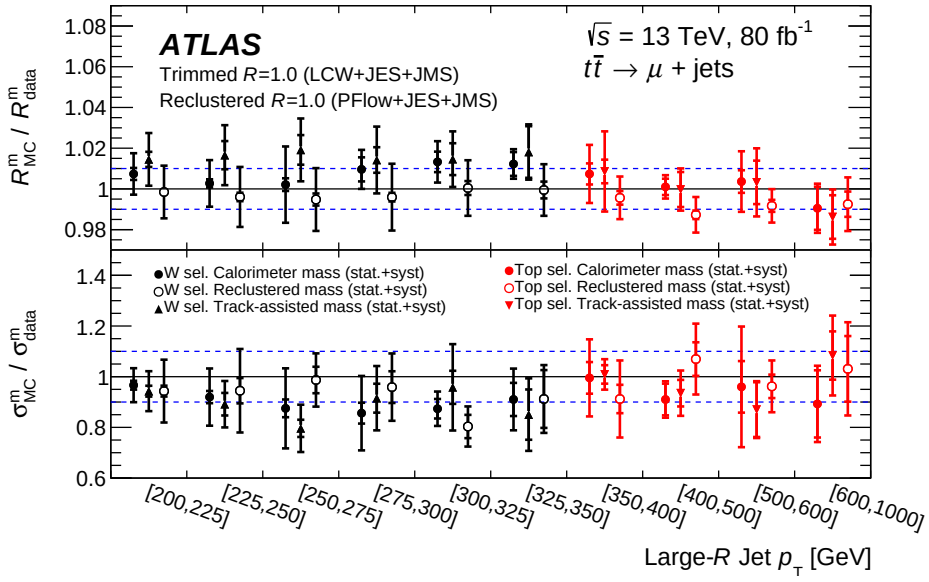


Figure 8.1: The relative Jet Mass Scale and Resolution for several large-R jet mass definitions as a function of the jet p_{T}. Black and red markers represent results for W- and top-initiated jets. The inner error bar stands for the statistical uncertainty and the outer error bar includes the statistical and systematic uncertainty. Dashed blue lines, corresponding to ±1% for the relative jet mass scale and ±10% for the relative jet mass resolution, are drawn for reference.

the MSR(R = 1 GeV) and pole renormalization schemes. The treatment of both perturbative and non-perturbative QCD effects in the MC generator, such as the parton shower and hadronization, is expected to change the shape of the jet mass distribution and, therefore, the relation between the MC and the renormalized mass.

We find the following relations between the MC and the MSR/pole mass for the nominal Powheg + Pythia 8 setup:

$$m_t^{\mathrm{MC}} = m_t^{\mathrm{MSR}}(1\ \mathrm{GeV}) + 80^{+350}_{-400}\ \mathrm{MeV}, \tag{8.1}$$

$$m_t^{\mathrm{MC}} = m_t^{\mathrm{pole}} + 350^{+300}_{-360}\ \mathrm{MeV}, \tag{8.2}$$

where the quoted uncertainties are computed from scale variations on the theory calculation, the choice of the fit range, the p_{T} intervals used and the treatment

of the underlying event effects. The top MC mass is found to be numerically closer to the MSR mass, and it is confirmed to be compatible with the pole mass with an uncertainty below 500 MeV. This result is compatible with the calibration found in e^+e^- collisions [124] and gives support to the size of the interpretation problem addressed in Ref. [54]. Furthermore, it is stable within 200 MeV against alternative closely-related jet definitions, and the variations of the internal Pythia 8 settings considered here do not induce shifts larger than 100 MeV in the relation given in Eq. 8.2. The question regarding the extrapolation of this calibration to other non-related observables and non-boosted regimes, remains open.

The dependence of the result of Eq. 8.2 with the fragmentation and hadronization models used is inspected by replacing Pythia 8 by Herwig 7, yielding $m_t^{\mathrm{MC}} = m_t^{\mathrm{MSR}}(1\ \mathrm{GeV}) + 230 \pm 90\ (stat)$ MeV. Although it is compatible with that derived with Pythia 8, it confirms that the central value of the top MC mass calibration is specific for the MC scheme adopted.

With the interpretation of the measured $t\bar{t}$ charge asymmetry $A_{\mathrm{C}}^{t\bar{t}}$ in the context of the Standard Model Effective Field Theory (SMEFT) we complete the block of top quark studies. The charge asymmetry is measured with 139 fb^{-1} of pp data at $\sqrt{s} = 13$ TeV, collected by the ATLAS detector. The measurement combines events selected in the semileptonic and dileptonic $t\bar{t}$ final state, and is performed inclusively and differentially in the invariant mass of the $t\bar{t}$ system $m_{t\bar{t}}$, among others. In line with previous measurements at $\sqrt{s} = 7$ and 8 TeV and preliminary results at 13 TeV [101, 226], the measured $A_{\mathrm{C}}^{t\bar{t}}$ is in good agreement with the NNLO QCD + NLO EW Standard Model expectation [239].

With these results, 15 dimension-six operators of the SMEFT that can impact the measured $t\bar{t}$ charge asymmetry are constrained. The contribution of the associated Wilson operators is computed at NLO [241] and added

to the SM prediction in two scenarios: including Λ^{-2} terms accounting for interferences between the SM and SMEFT only, and considering also SMEFT-SMEFT interferences, proportional to Λ^{-4}. The measured and predicted $A_{\mathrm{C}}^{t\bar{t}}$ are compared with a $\chi 2$ test, and the 68% and 95% confidence level limits are derived individually for each coefficient from a fit to the inclusive measurement and from a simultaneous fit to all differential $m_{t\bar{t}}$ measurements made. The fit includes experimental and theory uncertainties, and keeps track of the bin-to-bin correlations in the differential case. The derived bounds are summarized in Figure 8.2.

Overall, we observe that the linear fit yields tighter bounds on the Wilson coefficients for octet operators, that enter at tree-level, than for the coefficients of singlet operators that only affect the $t\bar{t}$ production at NLO [246]. The bounds derived from the differential fit are about a factor two stronger than those obtained with the single fit to the inclusive measurement. This is because in the simultaneous fit, we take advantage of all information available without diluting the individual sensitivity of the differential $m_{t\bar{t}}$ bins. The limits obtained in this work are competitive with the individual bounds reported on the basis of a much larger data set [246–248] that includes top, Higgs and EW data [242, 249], and they will contribute to global-fits.

Finally, this thesis presents the first measurement of the bottom quark mass at the renormalization scale of the Higgs boson mass in the $\overline{MS}$ scheme. This result relies on measurements of the Higgs boson decay rates $\mathcal{B}(H \to b\bar{b})/\mathcal{B}(H \to ZZ)$ performed by the ATLAS and CMS experiments, based on 139 fb^{-1} and 35 fb^{-1} of data at $\sqrt{s}$ = 13 TeV [284, 286], respectively. The relation between the Higgs decay rate into $b\bar{b}$ and the bottom quark mass can be obtained analytically assuming $m_b << m_H$ and standard Yukawa couplings between the Higgs boson and the bottom quark. At the renormalization scale of m_H the QCD prediction shows

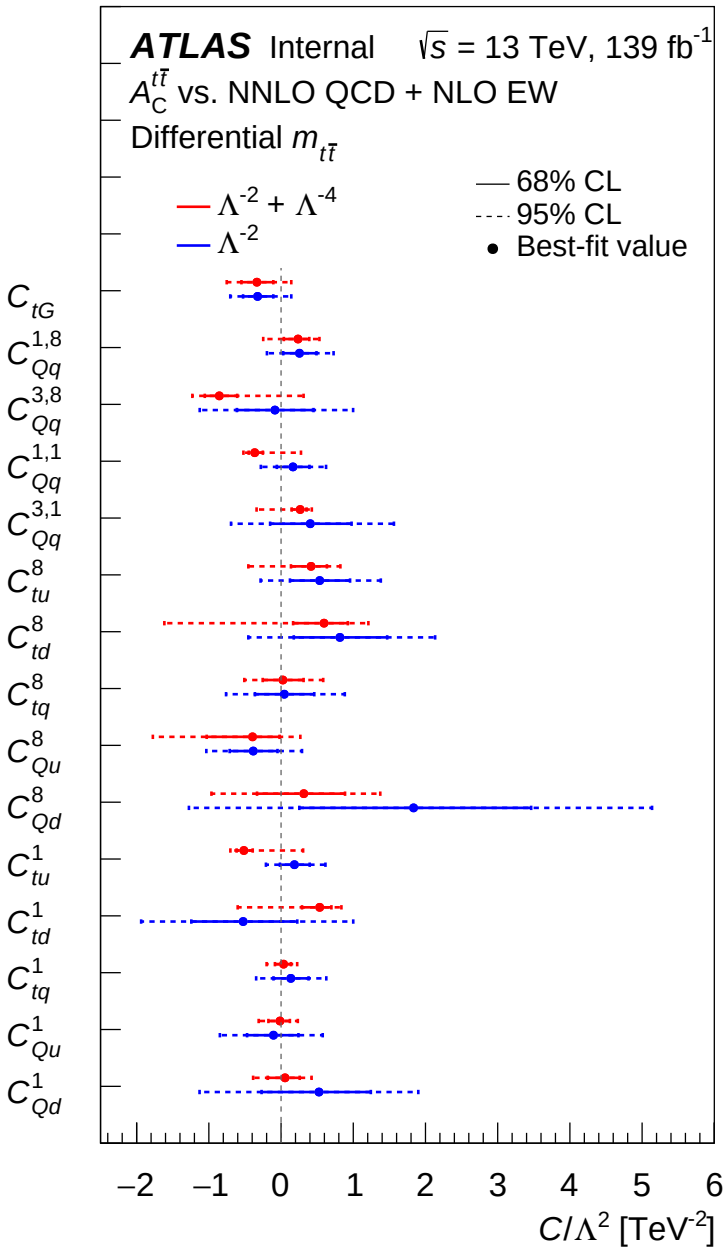


Figure 8.2: Individual 68% and 95% Confidence Level limits on the Wilson coefficients C_i/Λ^2 in the Standard Model Effective Field Theory from the differential charge asymmetry measurements versus $m_{t\bar{t}}$. The bounds are derived from the measured charge asymmetry presented in this chapter, combining the semileptonic and dileptonic channels. The theory uncertainty corresponds to the NNLO QCD + NLO EW calculation of Ref. [239]. The impact of dimension-six operators is parameterized at NLO accuracy in QCD using the SMEFTNLO package in MG5_aMCNLO.

an excellent convergence, which enables a precise numerical calculation. The ratio of partial widths to $b\bar{b}$ and ZZ is parametrized as a function of m_b. The uncertainty on this parametrization accounts for missing higher-orders in the numerical prediction of $\Gamma^{b\bar{b}}$ and Γ^{ZZ}, and on the parametrization itself due to the value of $\alpha_s(m_Z)$ and m_H. The total uncertainty, at the level of 4.4%, translates into an uncertainty of 60 MeV on the measured $m_b(m_H)$, negligible in comparison with the experimental uncertainties.

The $m_b(m_H)$ values associated to the ATLAS and CMS measurements are combined with the Convino package [280], assuming a 50% correlation between the experimental uncertainties and 100% for the theory ones. The combined measurement reads:

$$m_b(m_H) = 2.60^{+0.36}_{-0.31} \text{ GeV} , \tag{8.3}$$

fully compatible with the Particle Data Group (PDG) world average of $m_b(m_b)$ when evolved using the 5-loop renormalization group equation implemented in the REvolver package [320]. This result is used along with previous measurements of the bottom quark mass at the scale of the bottom quark mass and the Z boson mass to confront the scale evolution predicted by the SM, as shown in Figure 8.17. Within the current precision, we find an overall good agreement between the experimental measurements and the prediction of the SM.

The hypothesis of the running is further tested by introducing a parameter x that modifies the SM scale evolution, which corresponds to $x = 1$. In this way, m_b is evolved from $\mu_R = m_b$ to m_Z and m_H considering different initial values and modified scale evolutions. With a $\chi 2$ test, the values of $m_b(m_b)$ and x that best describes the experimental measurements are:

$$\begin{aligned} m_b(m_b) &= 4.18^{+0.03}_{-0.02} \text{ GeV} , \\ x &= 1.08 \pm 0.15 \text{ (exp)} \pm 0.05 \ (\alpha_s). \end{aligned} \tag{8.4}$$

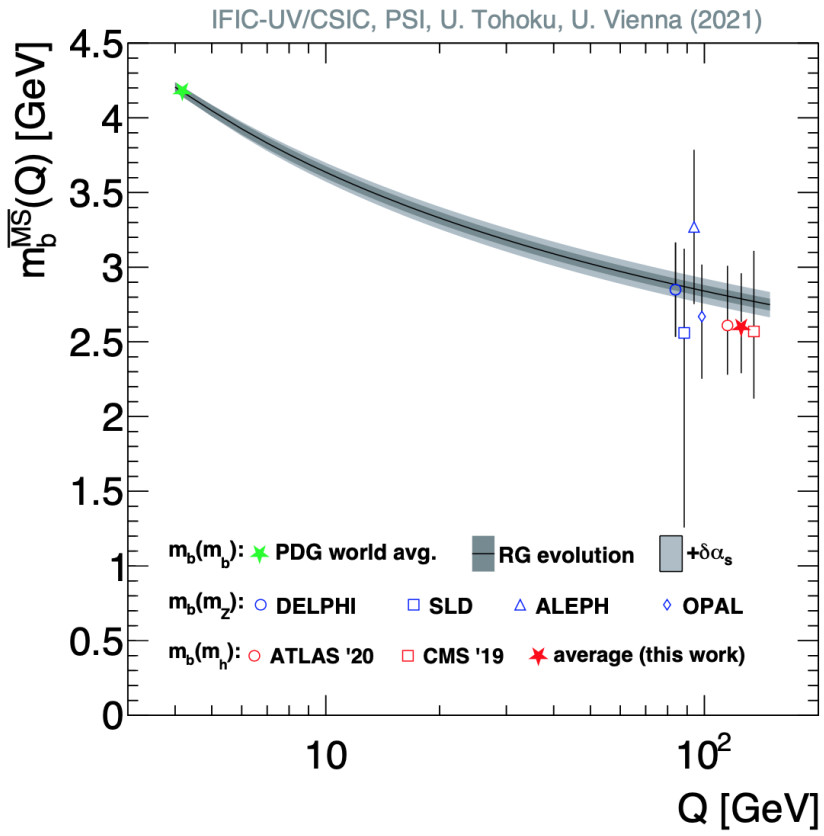


Figure 8.3: The scale evolution of the bottom quark $\overline{MS}$ mass. The measurements include the Particle Data Group world average for $m_b(m_b)$ determined at a typical scale of the bottomonium mass, the measurements of $m_b(m_Z)$ from jet rates at the Z-pole at LEP and SLC and the measurement of $m_b(m_H)$ from Higgs boson branching fractions. The prediction of the evolution of the mass is calculated at five-loop precision with REvolver [320]. The inner dark grey error band includes the effect of missing higher orders and the parametric uncertainties from $m_b(m_b)$ and α_s from the PDG averages. The outer band with a lighter shading includes additionally the effect of a ± 0.004 variation of $\alpha_s(m_Z)$.

The fitted value for the reference $m_b(m_b)$ is fully compatible with the PDG world average within the total uncertainty. The value of x compatible with the SM running within 1σ and differs by nearly seven standard deviations from the no-running scenario. A fit without the Higgs data yields a value of x of 1.03 ± 0.21, right below five standard deviations. Therefore, this result establish the evolution of the bottom quark mass with the energy scale.

Appendix A

SMEFT bounds from differential $m_{t\bar{t}}$ measurements of A_{C}

SMEFT limits on the $C_{ut}^{(8)}$ operator are obtained form differential measurements of the $t\bar{t}$ charge asymmetry as a function of the invariant mass of the $t\bar{t}$ system. Several invariant $t\bar{t}$ mass intervals are considered:

$$m_{t\bar{t}} \in [0, 500, 750, 1000, 1500, \infty]\,\mathrm{GeV}\ .$$

The comparison between the measured and predicted $A_{\mathrm{C}}^{t\bar{t}}$ are shown in the lower panel of Figures A.1 through A.5. The associated $\chi 2$ distributions are displayed in the upper panels, which correspond to fits considering only linear terms in the EFT parametrization proportional to Λ^{-2}, on the one hand, and including also Λ^{-4} quadratic terms, on the other.

In Figure A.6 the individual 68% and 95% CL limits for all dimension-six operators obtained from Asimov and the measured A_C are compared. Here, only Λ^{-2} terms in the EFT parametrization are considered. In Figure A.7, the comparison is done for limits derived including Λ^{-4}.

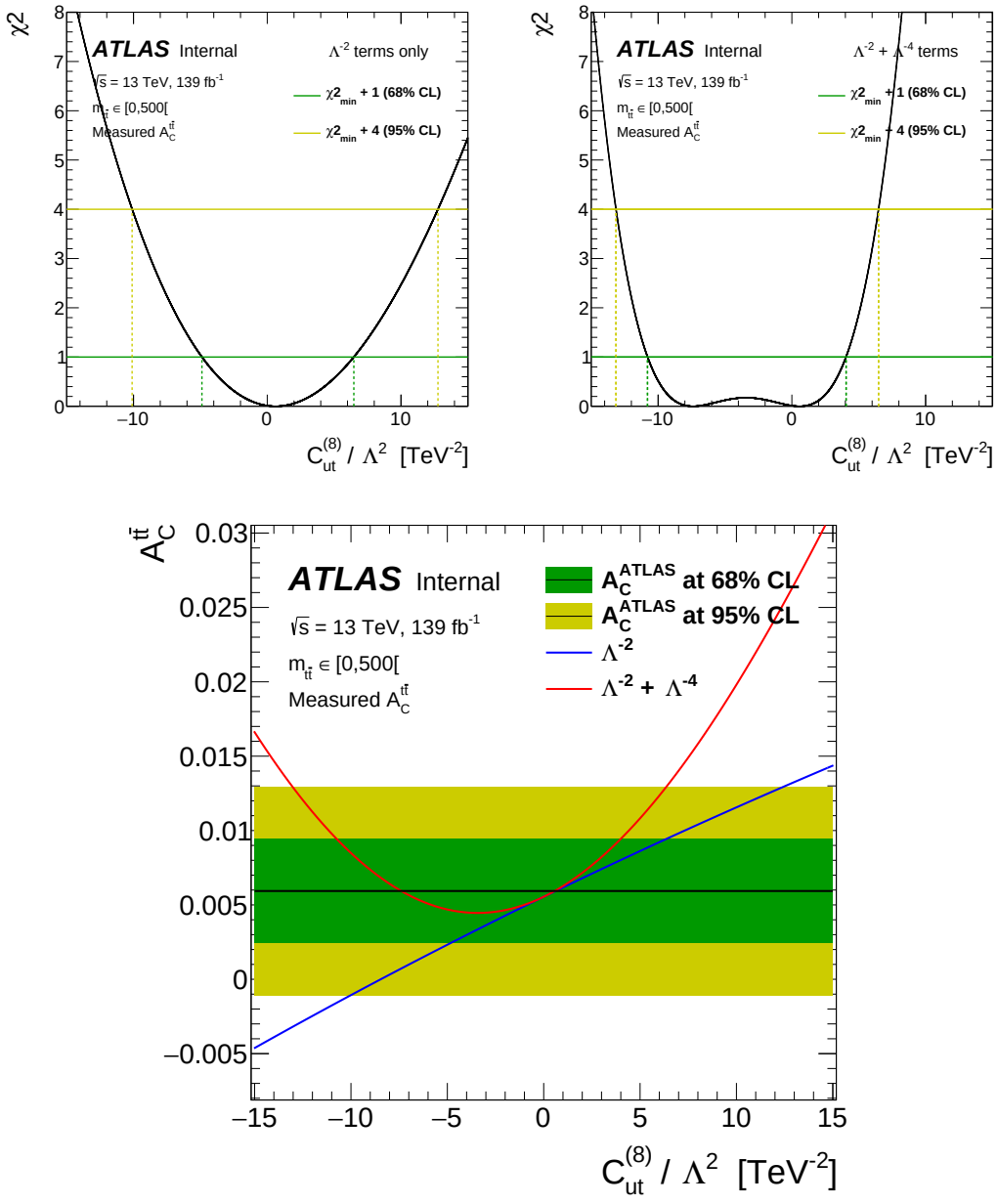


Figure A.1: $\chi 2$ distribution from which limits on $C_{ut}^{(8)}$ operator are derived when considering (upper-left panel) only SM-EFT terms and (upper-right panel) SM-EFT and EFT-EFT terms in the EFT contribution. Bottom panel shows the comparison between ATLAS and predicted $A_{\mathrm{C}}^{t\bar{t}}$ values, with the 68% and 95% Confidence Level bands drawn in green and yellow, respectively. $m_{t\bar{t}} \in [0, 500\}$ GeV.

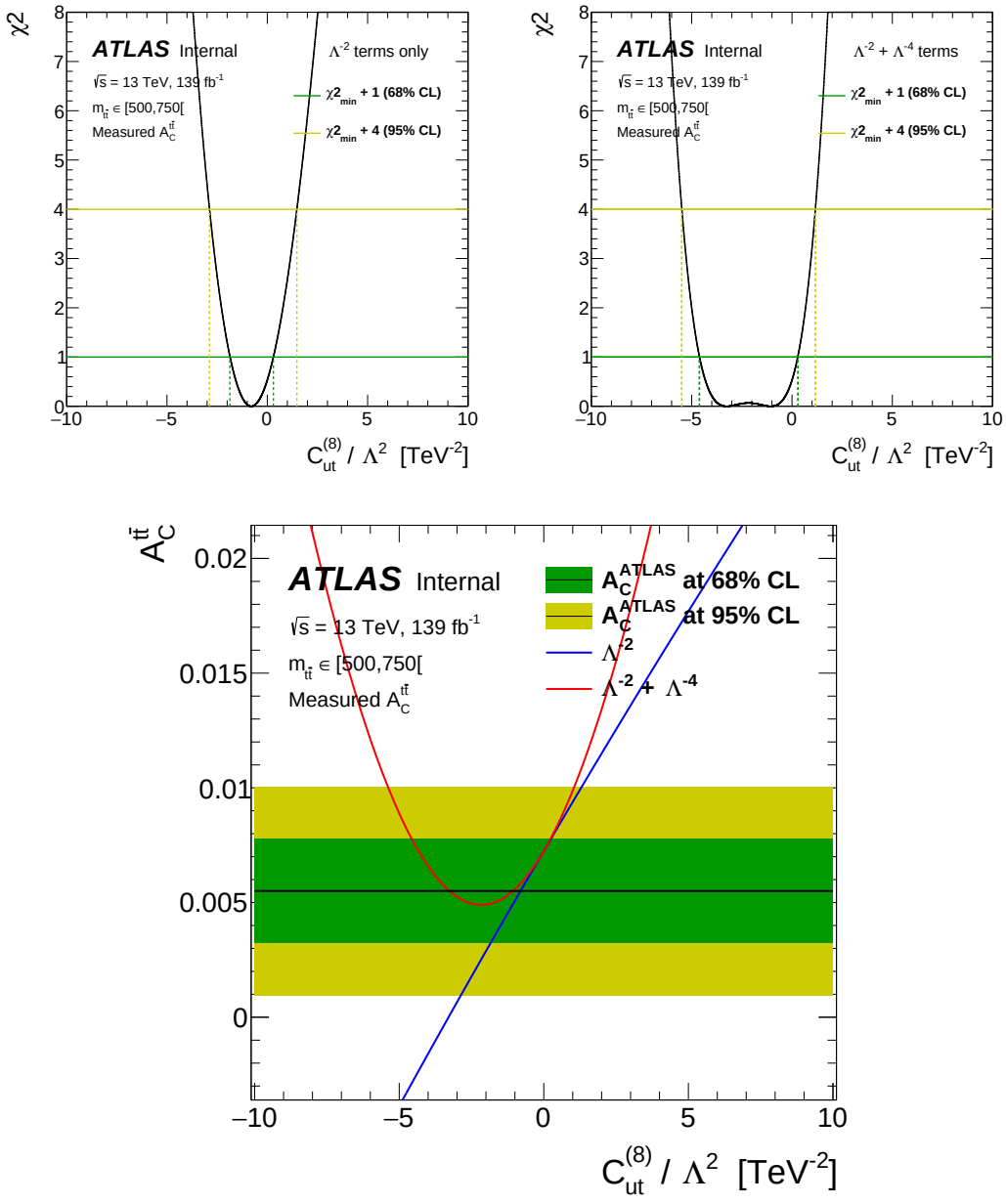


Figure A.2: $\chi 2$ distribution from which limits on $C_{ut}^{(8)}$ operator are derived when considering (upper-left panel) only SM-EFT terms and (upper-right panel) SM-EFT and EFT-EFT terms in the EFT contribution. Bottom panel shows the comparison between ATLAS and predicted $A_{\mathrm{C}}^{t\bar{t}}$ values, with the 68% and 95% Confidence Level bands drawn in green and yellow, respectively. $m_{t\bar{t}} \in [500, 750\}$ GeV.

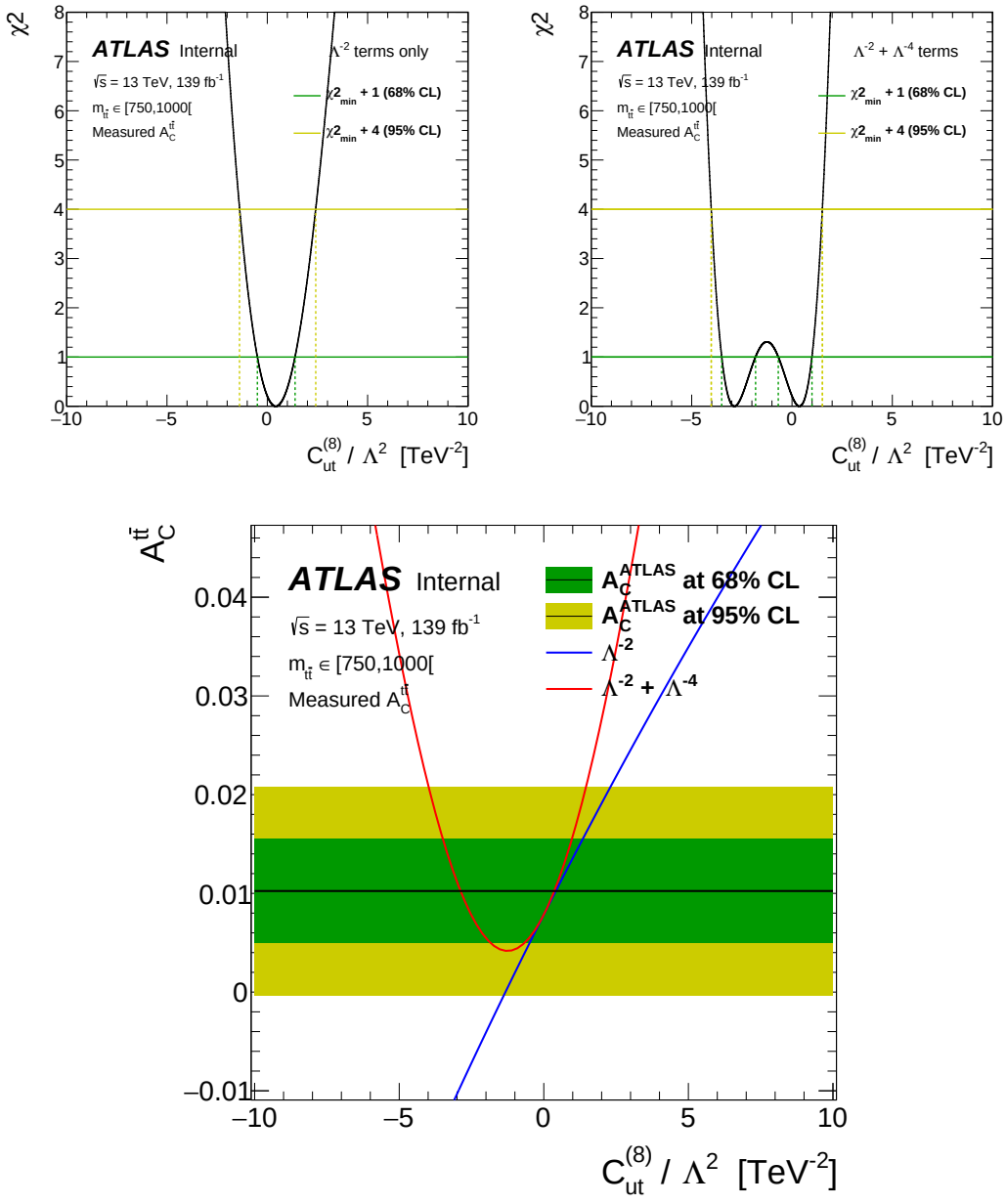


Figure A.3: $\chi 2$ distribution from which limits on $C_{ut}^{(8)}$ operator are derived when considering (upper-left panel) only SM-EFT terms and (upper-right panel) SM-EFT and EFT-EFT terms in the EFT contribution. Bottom panel shows the comparison between ATLAS and predicted $A_{\mathrm{C}}^{t\bar{t}}$ values, with the 68% and 95% Confidence Level bands drawn in green and yellow, respectively. $m_{t\bar{t}} \in [750, 1000\}$ GeV.

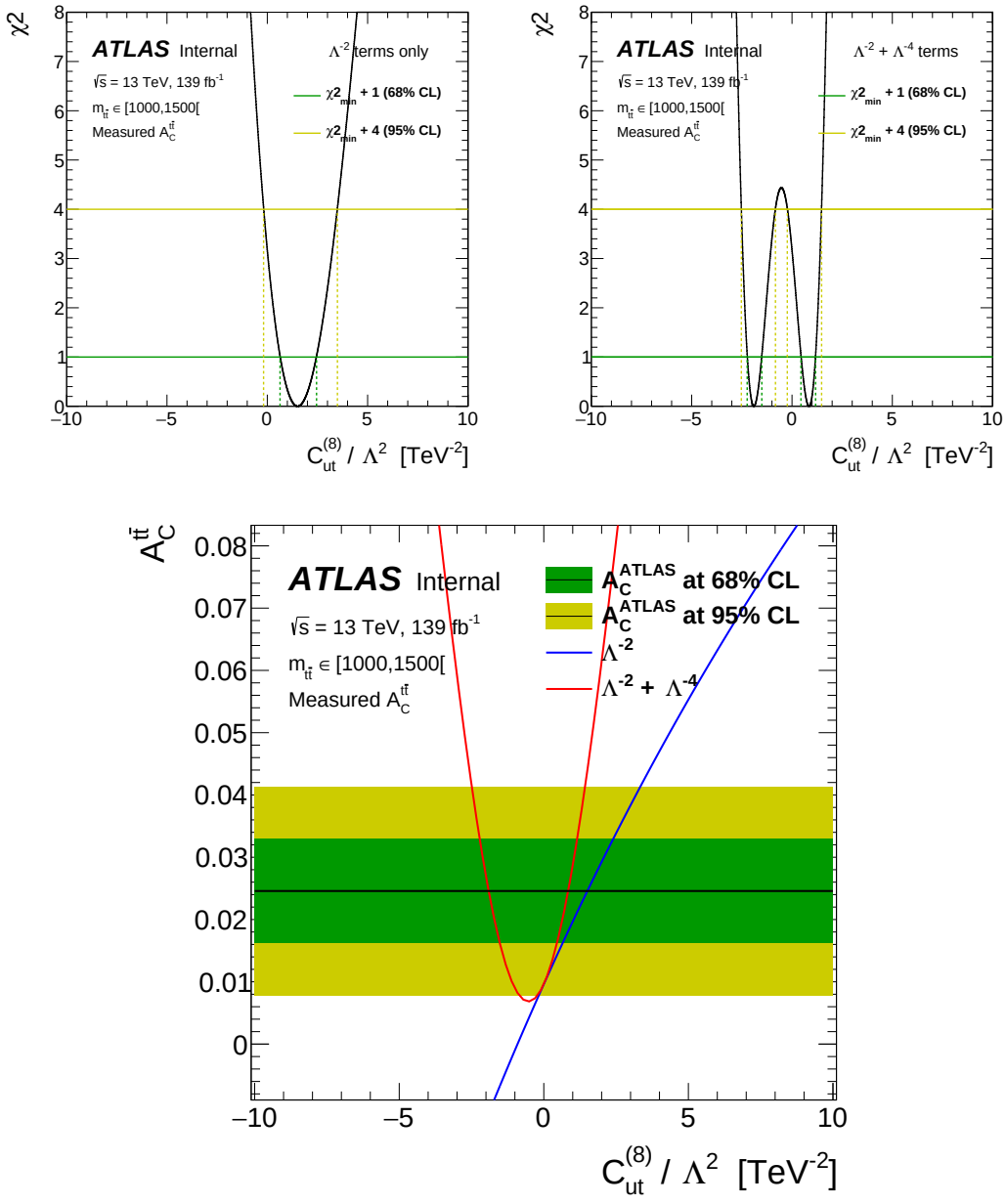


Figure A.4: $\chi 2$ distribution from which limits on $C_{ut}^{(8)}$ operator are derived when considering (upper-left panel) only SM-EFT terms and (upper-right panel) SM-EFT and EFT-EFT terms in the EFT contribution. Bottom panel shows the comparison between ATLAS and predicted $A_{\mathrm{C}}^{t\bar{t}}$ values, with the 68% and 95% Confidence Level bands drawn in green and yellow, respectively. $m_{t\bar{t}} \in [1000, 1500\}$ GeV.

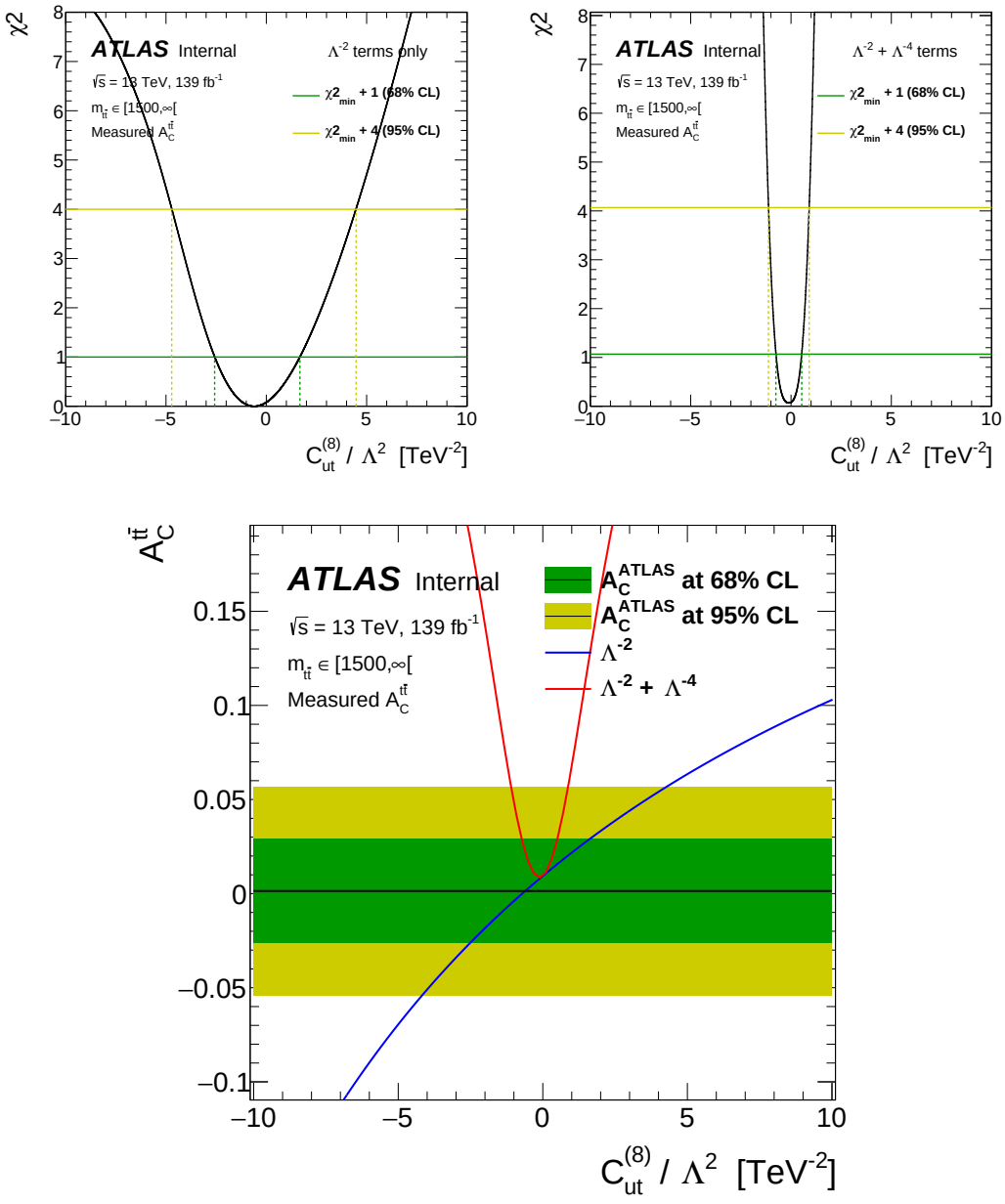


Figure A.5: $\chi 2$ distribution from which limits on $C_{ut}^{(8)}$ operator are derived when considering (upper-left panel) only SM-EFT terms and (upper-right panel) SM-EFT and EFT-EFT terms in the EFT contribution. Bottom panel shows the comparison between ATLAS and predicted $A_{\mathrm{C}}^{t\bar{t}}$ values, with the 68% and 95% Confidence Level bands drawn in green and yellow, respectively. $m_{t\bar{t}} \in [1500, \infty\}$ GeV.

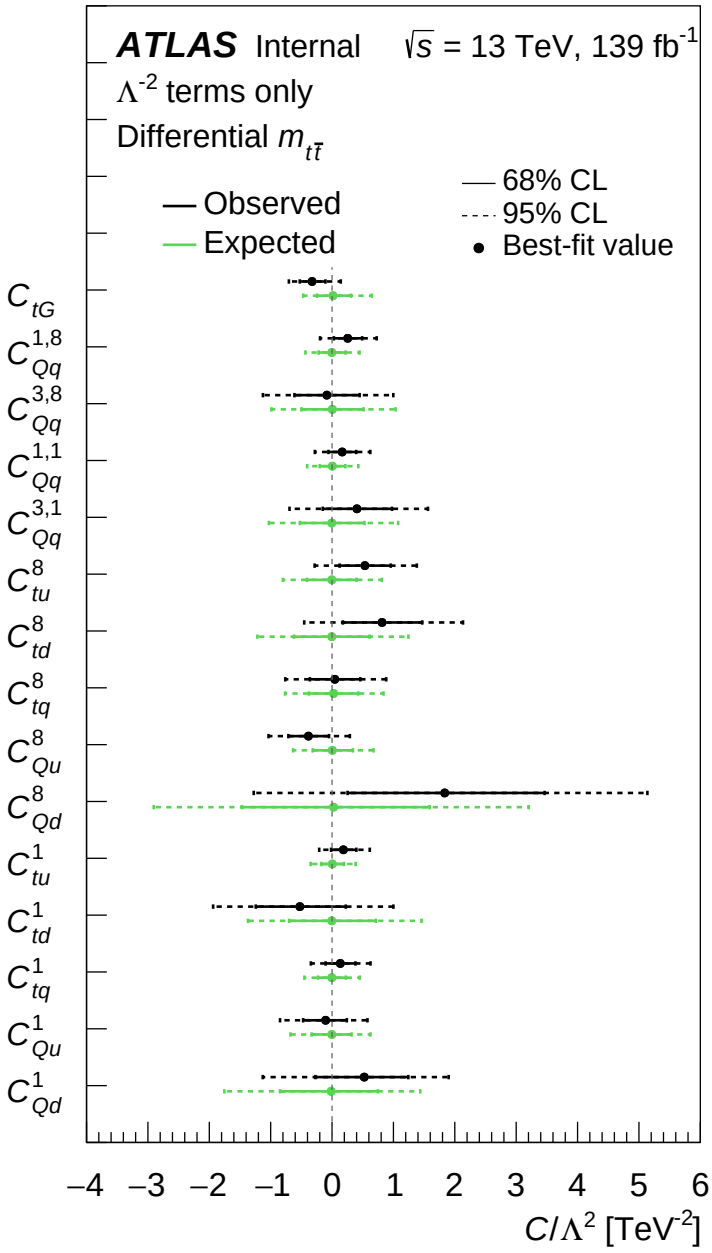


Figure A.6: Comparison of the individual 68% and 95% Confidence Level expected and measured limits on the Wilson coefficients C_i/Λ^2 in the Standard Model Effective Field Theory from the differential charge asymmetry measurements versus $m_{t\bar{t}}$. Only fits including Λ^{-2} terms in the EFT parametrization are shown.

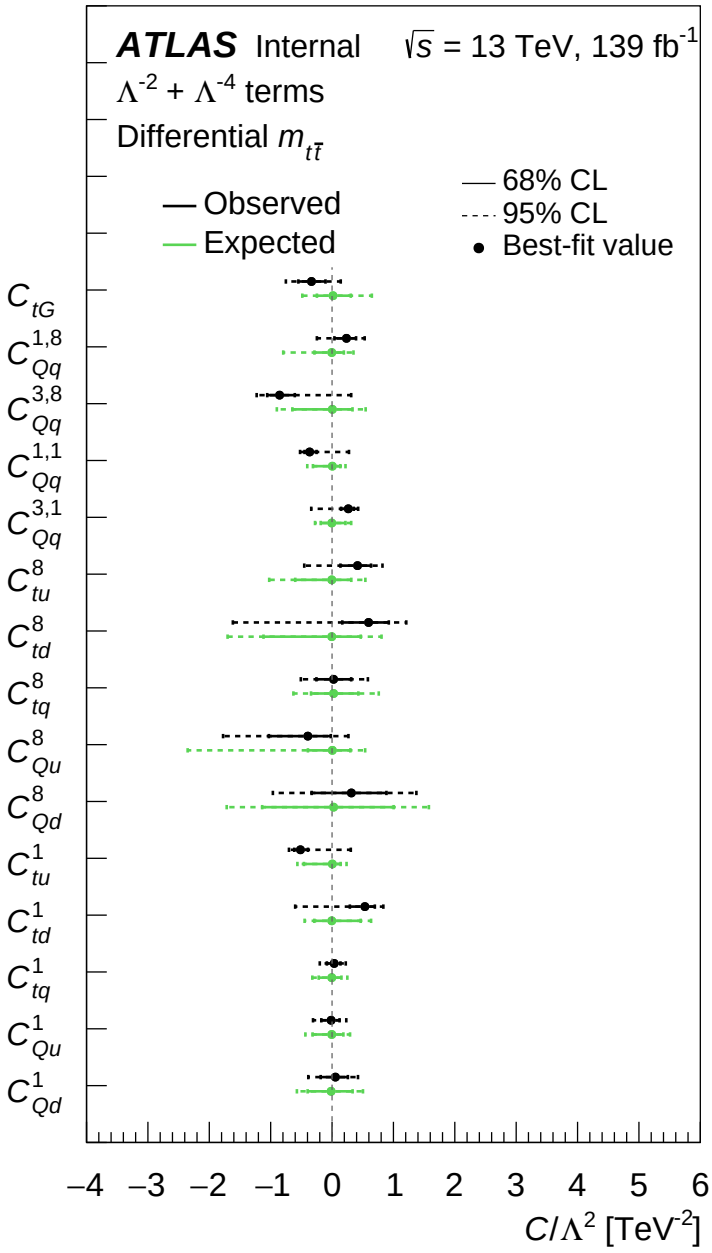


Figure A.7: Comparison of the individual 68% and 95% Confidence Level expected and measured limits on the Wilson coefficients C_i/Λ^2 in the Standard Model Effective Field Theory from the differential charge asymmetry measurements versus $m_{t\bar{t}}$. Only fits including $\Lambda^{-2} + \Lambda^{-4}$ terms in the EFT parametrization are shown.

Resum

La teoria de les partícules fonamentals

Actualment, la major part del coneixement i enteniment que posseïm sobre com funciona el món subatòmic es condensa en l'anomenat Model Estàndar (SM per les seues sigles en anglés). Aquest representa la culminació d'una revolució en el camp de física de partícules que esdevení al començament del segle XX, marcat pel desenvolupament de la mecànica quàntica i la relativitat especial. D'aquesta forma, totes les partícules que formen la matèria que hem pogut observar fins ara, i les interaccions que hi ha entre elles, són descrites amb éxit al SM.

Matemàticament, el SM es perfila com una teoria quàntica de camps en què les partícules elementals adopten la interpretació d'ens puntuals (sense dimensió) descrites per camps, l'estat quàntic dels quals romanen intrinsecament determinats per nombres quàntics, com l'espín. Tanmateix, es diferencien dos tipus de partícules elementals: fermions i bosons. La Figura 8.8 mostra un resum de les partícules fonamentals del SM.

Els bosons tenen espín sencer i, a més a més, són el responsables de transmetre la interacció entre fermions. Formalment, les interaccions sorgeixen com a conseqüència de la invariància física del sistema davant de certes transformacions, és a dir, d'una simetria. El coneixement empíric que se'n obtingué durant les primeres dècades del segle XX sobre l'estructura i la naturalesa de les interaccions

Figure 8.8: Partícules fonamentals del Model Estàndar.

fonamentals, va ajudar a trobar la simetria global que permet incorporar tres de les quatre interaccions o forces fonamentals observades a la natura. Així, la interacció nuclear forta (QCD), que explicava la formació del nucli dels àtoms; la força nuclear feble, que per contra era la responsable de la seua desintegració; i la força electromagnètica, origen del fenómens relacionats amb la llum, són descrits en terms del grup de simetria $SU(3) \times SU(2) \times U(1)$. Els gluons són els mediadors la interacció nuclear forta, mentre que els bosons W, Z i els fotons γ són els missatgers de la força electrofeble (EW, resultat de l'unificació de la força nuclear feble i electromagnètica). El bosó de Higgs, a diferència de la resta, està relacionat amb un aspecte crucial de les partícules elementals: l'origen de la massa. La massa de les partícules és conseqüència de la interacció amb el camp del bosó d'Higgs, mitjançant el intercamvi d'un bosó d'Higgs.

Els fermions, per altra banda, es caracteritzen per tindre un espín fraccionari i per ser el constitüents fonamentals de la matèria. Alhora, els fermions es classifiquen en quarks i leptons. Com mostra la Figura 8.8, n'hi ha un total de sis

quarks i sis leptons organitzats en una estructura de tres famílies o generacions. Els fermions més lleugers pentanyen a la primera família, mentre que els de la tercera són el més massius.

Tots el quarks posseeixen càrrega elèctrica fraccionària i sòn sensibles a totes les forçes citades anteriorment. La natura de la força forta no permet als quarks viure molt de temps sense formar estats lligats de més d'una partícula, anomenats hadrons. Aquest fenòmen es coneix com confinament, i es manifesta tan promte com els quarks se separen prou en l'espai. A diferència d'això, la llibertat asimptòtica té com característica principal que els quarks tot just interaccionen quan es troben prou a prop. Amb tot, els estats lligats de dos i tres quarks són coneguts como mesons i barions, respectivament. Per contra, el leptons no formen estats lligats de més d'una partícula i només tres d'ells estàn carregats elèctricament (electró e, muó μ i tau τ), tenint cadascun un altre leptó neutre, anomenat neutrí ν (neutrí electrònic, muònic i tau). Els neutrins tenen la peculiaritat de ser molt lleugers i increïblemente difícils de detectar als experiments.

La massa de les partícules fonamentals en determina la fenomenologia. D'entrada, els hadrons més estables estan formats pels quarks més lleugers: el *up* i *down*. El protó (uud) i el neutró (udd) formen els nuclis atómics dels àtoms. Cal tenir en compte que, en realitat, tots el hadrons gaudixen d'una estructura interna molt més complexa. Per il·lustrar-ho, posem per cas el protó, representat en la Figura 8.9; seria més precís visuzalitzar-la com un núvol o mar de quarks i gluons en constant interacció, on la probabilitat de trobar un quark *up* o *down* amb una fracció gran de l'energia total del protó (considerats quarks de València) és molt major que la de trobar-se qualsevol altre quark o gluó. En segon lloc, les partícules més masives tendeixen a desintegrar-se a les seues companyes més lleugeres, sent el temps de vida d'aquestes partícules més curta com més gran siga la seua massa.

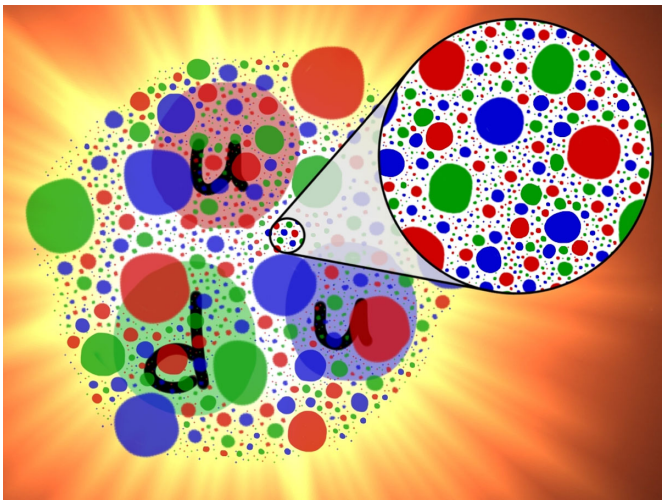


Figure 8.9: Il·lustració de l'estructura interna del protó, on es pot observar els tres quarks de valència (dos *up* i un *down*) i el núvol o mar de quarks i gluons en constant interacció. La grandària de les partícules es representativa de la fracció de l'energia total del protó que tenen.

Així, el quark *top*, el més massiu de tots, té la vida mitjana més curta i sempre es desintegra abans de formar estats lligats amb altres quarks, és a dir, abans d'hadronitzar. L'estudi directe del quark *top* als detectors experimentals permet conèixer les propietats de quarks lliures.

Amb aquests ingredients bàsics, partícules i interaccions, el SM pot predir amb molta precisió com i amb quina freqüència interactuen aquestes partícules, dins d'uns límits. En efecte, la capacitat predictiva de la teoria és limitada per l'aparició d'una sèrie de divergències que tenen com a resultat càlculs sense cap sentit físic. Malgrat això, l'origen d'aquestes divergències és conegut i poden solucionar-se mitjançant processos de renormalizatzació, els quals consisteixen en absorbir en els paràmetres lliures de la teoria, com masses i constants de interacció o acoblament, les divergències. La manera exacta en què això es duu a terme està determinada per l'esquema de renormalització (RS) adoptat. Els observables físics calculats, en canvi, sempre romandran els mateixos. La principal conseqüència d'aquesta condició és la introducció d'una dependència amb una escala d'energia, coneguda com escala de renormalització, en els

paràmetres renormalizats de la teoria. Des d'açi, el valor d'aquests paràmetres deixa de ser una constant.

Tot i que la teoria ha recuperat el seu poder predictiu, encara patix limitacions d'un altre tipus. Particularment, no pot assolir càlculs en què les constants d'interacció són massa grans. Això fa que el SM no pugua oferir una descripció formal i exacta de la formació dels hadrons. En aquest procés, el qual es desenvolupa en règims donats per l'escala d'hadronització $\Lambda_{\mathrm{QCD}} \sim \mathcal{O}(1\ \mathrm{GeV})$, la força forta es torna massa intensa. Els aspectes que no poden tractar-se formalment, són descrits amb models fenomenològics inspirats en l'evidència científica i basats en les anomenades tècniques de Monte Carlo (MC). Aleshores, les mesures experimentals sempre es poden aprofitar i confrontar amb la teoria.

Sens dubte, l´èxit del SM no té precedents. Les seues prediccions han estat corroborades per la majoria d'experiments realitzats fins ara. Tot i això, aquest èxit no és total, ja que hi han observacions experimentals que no poden ser explicades des del SM. Per tant, el SM no pot ser la teoria final. En aquest context, els esforços de la comunitat física d'altes energies, tant experimental com teòrica, es dirigeixen a forçar els límits de validesa del SM i trobar proves que apunten a una nova teoria més general i completa.

Per una banda, aquests esforços es dirigeixen a mesurar el paràmetres fonamentals del SM amb la màxima precisió possible, des de les dades experimentals. Així, es poden realitzar prediccions més precises i desvetllar possibles discrepàncies amb allò que s'ha observat als experiments. En la present tesi doctoral, mesurem la massa del quark *bottom* a una escala d'energia mai no inspeccionada. Encara més, utilitzant mesures prèvies a energies més baixes, comprovarem si l'evolució del valor de la massa d'aquesta partícula correspon amb la predicció del SM. Juntament amb això, ens endinsarem en la discussió de la interpretació de la massa del quark *top*, la qual es pot mesurar servint-se de

tècniques diferents. El problema de la interpretació es relaciona amb el signifcant del la massa mesurada, que de fet depén de la tècnica utilitzada.

Per altra banda, s'investiguen possibles extensions del SM que puguen explicar les discrepàncies existents. Com que n'hi han moltes opcions igualment vàlides i en els experiments no és pràctic comprovar-les totes, generalment s'utilitzen teories efectives de camps (EFT) com intermediàries. Aquestes EFT introdueixen canvis en la predicció del SM servint-se d'un llenguatge al qual qualsevol altra teoria por ser traduïda. Concretament, l'extensió EFT del SM (coneguda com SMEFT) fa ús d'uns termes anomenats operadors de Wilson per a modificar les interaccións tal i com són descrites originalment en el SM. L'intensitat i magnitud d'aquesta modificació està regulada per mitjà dels seus coeficients associats. Aleshores, si apareix alguna discrepància entre una observació experimental i la predicció nominal del SM, és possible interpretar aquesta diferència com l'efecte de la presència d'un operador de Wilson. En aquesta tesi, utilitzarem dades experimentals recollides a l'experiment ATLAS al LHC per posar límits als operadors de Wilson que estiguen relacionats amb la producció i distribució angular de parells quark *top* i *antitop*.

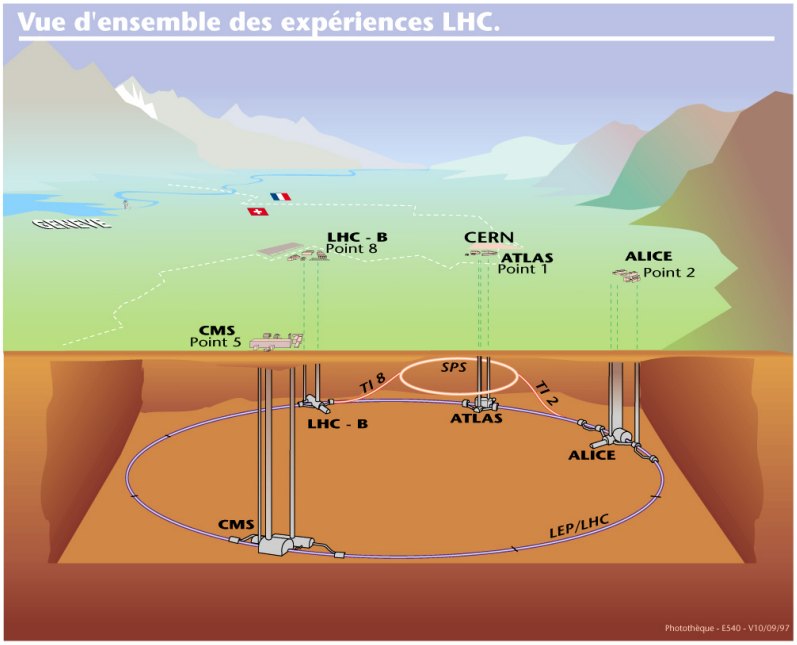


Figure 8.10: Localització del gran col·lisionador d'hadrons, LHC. Es mostren els quatre detectors experimentals principals a cadascun dels punts d'intersecció del feix de protons.

ATLAS i el gran col·lisionador d'hadrons

Una forma d'investigar la naturalesa de les partícules fonamentals i les seues interaccions està en fer col·lisionar partícules altament energètiques i observar el resultat d'aquesta col·lisió. Així, com si es tractés d'un microscopi, com més gran és l'energia de les partícules incidents, més gran és el poder d'inspeccionar regions de l'espai cada cop més menudes. A més, si l'energia de la col·lisió és prou alta, és possible produir partícules molt massives i estudiar-ne directament les propietats, com ara el bosó d'Higgs. El centre europeu per a la investigació de física nuclear i de partícules (CERN), ubicat a la frontera franc-suïssa prop de Ginebra, ha pres el relleu a la carrera per construir els acceleradors de partícules més potents fins ara, amb el gran col·lisionador d'hadrons (LHC) com a resultat. Com il·lustra la Figura 8.10, el LHC està localitzat a 100 m sota el terra i posseeix un perímetre de 27 km, la qual cosa el converteix en el major accelerador i col·lisionador circular de partícules mai construït.

El LHC consta de dos tubs que discorren paral·lels i en què s'injecten feixos de protons prèviament accelerats fins a una energia de 460 MeV. Els 1238 dipols magnètics superconductors disposats al llarg del seu recorregut circular permeten als protons assolir una energia màxima de 6.5 TeV durant el segon període d'operacions, conegut com Run 2 (2013 – 2018). Els dos feixos de protons s'estrenyen i creuen en 4 punts, cosa que afavoreix una taxa alta de col·lisions de protons a una energia al centre de masses de $\sqrt{s} = 13$ TeV. Tots dos aspectes són crucials per als objectius del LHC: assolir una alta energia disponible a la col·lisió per generar partícules molt massives i produir moltes col·lisions. D'aquesta manera, els experiments disposen de suficients dades per analitzar processos "rars" (amb poca probabilitat que succeïxen) d'interès amb prou precisió.

A cadascun dels punts d'intersecció s'instal·la un detector de partícules. El detector ATLAS és el de més dimensió i es va concebre amb l'objectiu d'abastar un rang molt ampli de física: des de l'estudi de les propietats del SM (com a seccions eficaces de producció o constants d'acoblament) fins a la cerca de noves partícules que puguen explicar la matèria fosca. Es va construir de manera simètrica al voltant del punt d'intersecció, de manera que exhibeix una geometria cilíndrica. Alhora, està constituït per una sèrie de sub-detectors que, de dins cap a fora, són el detector intern de traces, els calorímetres electromagnètic i hadrònic i, finalment, la càmera de muons, com es mostra al panell superior de la Figura 8.11. Aquesta disposició permet distingir els tipus de partícules resultants de la col·lisió i mesurar-ne les propietats, com l'energia total, la trajectòria i el moment. El panell inferior de la Figura 8.11 mostra una representació esquemàtica del tipus de senyals que deixen diferents partícules al seu pas per aquests subdetectors. Així, per a cada col·lisió s'obté una "fotografia" del detector on es registren la majoria de les partícules resultants, la reconstrucció i anàlisi de les quals permet esbrinar el procés físic que li va donar origen.

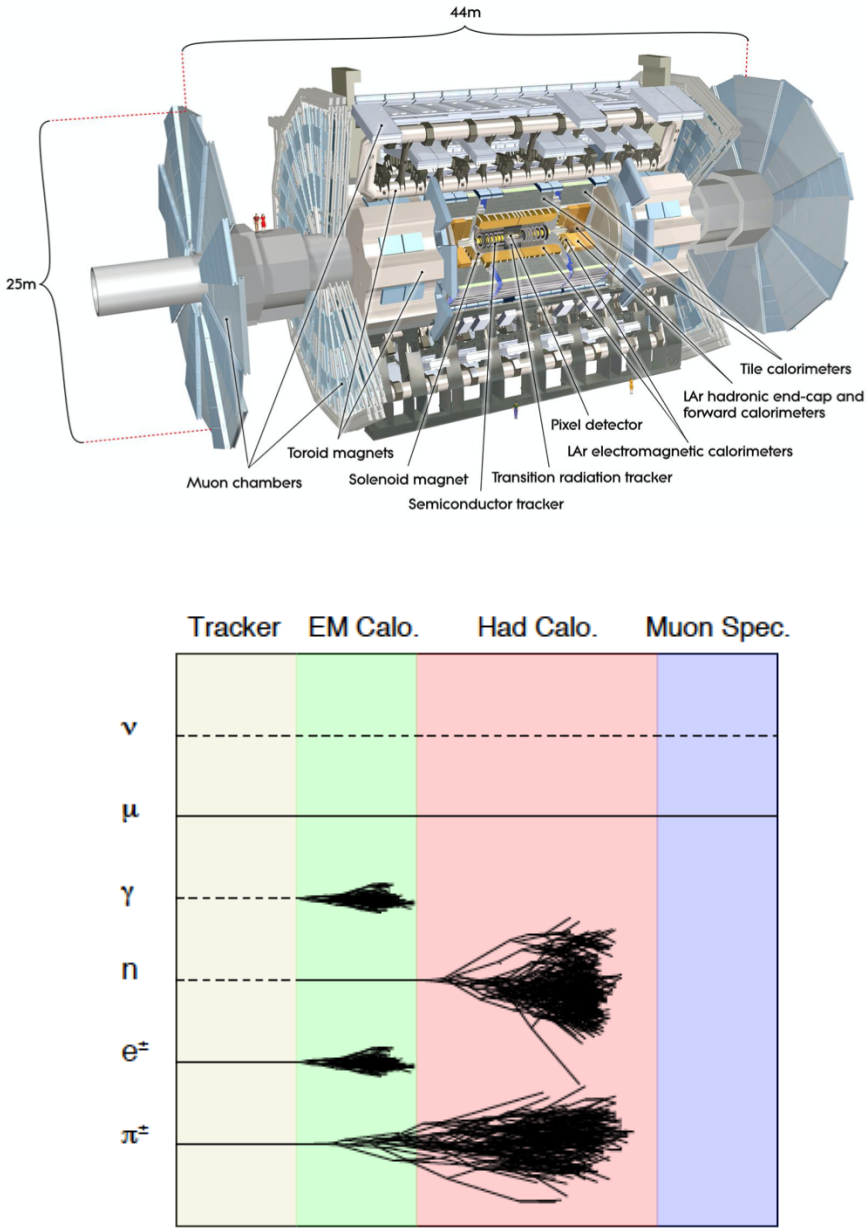


Figure 8.11: Al panell superior, el detector ATLAS i els seus subsistemes [136]. Al panell inferior, il·lustració del dipòsit d'energia de diferents partícules en funció del detector que travessen. Les línies discontínues representen partícules que no deixen energia, és a dir, que no interaccionen.

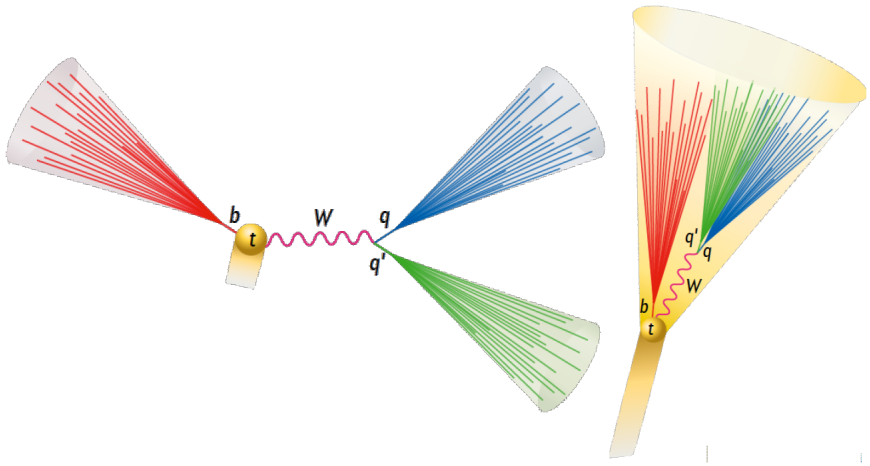
Física del quark *top*

El quark *top*, descobert per les col·laboracions DØ i CDF al col·lisionador Tevatron el 1995, és la partícula fonamental més pesada que es coneix. A causa d'això, s'espera que tingua un paper important en teories BSM, ja que partícules més massives i encara desconegudes podrien interactuar-hi amb especial intensitat. A banda d'això, el quark *top* rep un protagonisme especial en molts processos en el marc del SM, ja que indueix correccions quàntiques grans. Juntament amb la massa del bosó d'Higgs, la massa del quark *top* és un paràmetre fonamental, el coneixement precís del qual permet realitzar tests estrictes d'autoconsistència del SM.

El quark *top* pot produir-se tant en parells *top-antitop* $t\bar{t}$ com individualment, a partir de col·lisions protó-protó al LHC: a 13 TeV, el 90% dels parells $t\bar{t}$ es produeixen via *gluon-fussion* $gg \to t\bar{t}$, i el 10% a través de l'aniquilació de parells quark-antiquark $q\bar{q} \to t\bar{t}$. Com s'ha mencionat anteriorment, el quark *top* es desintegra ràpidament abans de formar estats estables amb altres quarks. En el 99% dels casos, la interacció electrofeble permet al quark *top* desintegrar-se en un bosó W i un quark *bottom*. Aquesta probabilitat s'anomena amplària de desintegració, i la total ve donada per la suma de tots el canals de desintegració possibles:

$$\Gamma_t = \sum_q \Gamma(t \to Wq), \quad \text{on } q = b, s, d \ . \tag{8.1}$$

En funció del canal de desintegració del bosó W, que es pot desintegrar al seu torn en dos quarks o en un leptó carregat i el corresponent antineutrí, els estats finals dels processos en què s'ha generat un parell $t\bar{t}$ es classifiquen en totalment hadrònic, totalment leptònic o semilepònic. Des d'un punt de vista experimental, a causa de la facilitat de detectar i aïllar amb molta precisió electrons i muons al calorímetre electromagnètic i a la cambra de muons, respectivament, els estats



(a) Rest (b) $p_T >> m$

Figure 8.12: Representació esquemàtica de la desintegració d'un quark *top* en dos règims diferents. A l'esquerra, el quark top és produït amb una energia propera a la seua massa, per la qual cosa els seus productes de desintegració s'emeten *back-to-back*. A la dreta, l'energia total és molt més gran que la seua massa, per això s'imprimeix un *boost* o empenta al sistema que fa que els productes de desintegració es concentren en una zona reduïda de l'espai.

finals $t\bar{t}$ que inclouen leptons solen utilitzar-se per a mesures de precisió.

La física relacionada amb el quark *top* es pot estudiar a través dels seus productes de desintegració, ja sigua per reconstruir l'estat final i obtenir seccions de producció eficaç o per mitjà de la reconstrucció parcial dels seus productes de desintegració. Una eina d'interès especial per a aquesta tesi són els anomenats jets. Quan una partícula es desintegra o interacciona amb el detector, genera tota una cascada de partícules de baixa energia. Per associar aquesta cascada de partícules amb la partícula que la va originar, fem servir un algorisme que permet seleccionar i combinar de manera seqüencial la informació de la cascada en una única entitat. Com és habitual, no hi ha una única manera de fer això, i el criteri per seleccionar i combinar la informació disponible dependrà dels interessos de cada anàlisi. Els jets es poden construir tant a partir de senyals elèctrics d'un detector com des dels productes de desintegració d'una partícula, com ara el quark *top*. Si el quark *top* és produït amb prou energia, els seus productes de

desintegració queden limitats a una petita àrea de l'espai fàsic. Si l'àrea del jet és prou gran, aquest podria capturar tots els productes de desintegració, i les propietats del quark *top* quedarien reflectides a les propietats del jet, tal com s'il·lustra al panell dret de la Figura 8.12. Si el quark *top* no té prou energia, el bosó W i el quark *bottom* es produeixen *back-to-back*, de manera que es necessita més d'un jet, generalment més menuts, per descriure'ls separadament.

Límits SMEFT a partir de $A_{\mathrm{C}}^{t\bar{t}}$

El SM prediu una característica molt peculiar relacionada amb la producció de parells *top-antitop* a partir de l'aniquilació de quarks: els quarks *top* solen produir-se preferentment en la direcció del quark inicial, mentre que els antitops ho fan en la direcció de l'antiquark. Encara que l'estat inicial al LHC és simètric (col·lisionen protons), hi ha la possibilitat que els quarks incidents siguen els de valència els quals, com ja hem dit anteriorment, tenen una fracció d'energia més gran que els antiquarks (que formen part del núvol de partícules). Com més energia o moment posseeïx una partícula, més a prop viatjarà respecte a la línia del feix i més gran serà, en valor absolut, la variable angular *rapidity* y. Com a conseqüència, es prediu un excés de *tops* a les regions posterior i anterior del detector ATLAS ($|y| > 0$) i una major població d'antiquarks a la regió central ($i \approx 0$), tal com s'il·lustra a la Figura 8.13. Això motiva la definició de l'observable anomenat asimetria de càrrega, $A_{\mathrm{C}}^{t\bar{t}}$:

$$A_{\mathrm{C}}^{t\bar{t}} = \frac{N(\Delta|y| > 0) - N(\Delta|y| < 0)}{N(\Delta|y| > 0) + N(\Delta|y| < 0)} \ , \tag{8.2}$$

Aquest observable és sensible a física més enllà del SM (BSM), per això s'estudia en el context de SMEFT: tots els operadors que puguen afectar la producció

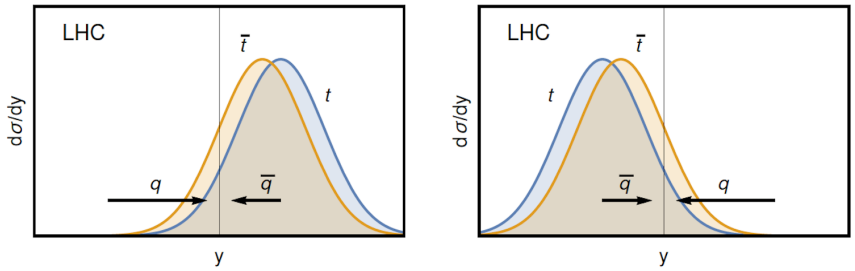


Figure 8.13: Distribució de la variable rapidity y associada a *tops* i *antitops* en producció $t\bar{t}$ a l'experiment ATLAS. Els tops tendeixen a produir-se en regions amb més valor absolut de i, mentre que els *antitops* poblen regions caracteritzades per valors de y propers a 0.

de $t\bar{t}$ i així alterar la $A_{\mathrm{C}}^{t\bar{t}}$, són considerats. D'aquesta manera, l'asimetria de càrrega teòrica es pot escriure com a resultat d'una contribució fixa només del SM, i una altra variable deguda al valor C dels operadors de Wilson considerats: $A_{\mathrm{C}}^{t\bar{t}} = A_{\mathrm{C}}^{t\bar{t}}(\mathrm{SM}) + A_{\mathrm{C}}^{t\bar{t}}(C)$.

Els valors que els operadors de Wilson poden adoptar són acotats gràcies a les mesures experimentals de $A_{\mathrm{C}}^{t\bar{t}}$. En aquesta tesi, hem utilitzat la mesura més recent obtinguda per l'experiment ATLAS, la qual es basa en les dades recollides durant tot el Run 2 a 13 TeV que contenen parells *top-antitop* als estats finals leptònic i semileptònic. La mesura es fa de forma inclusiva i diferencial en la massa del sistema invariant $t\bar{t}$, ja que la dilució induïda per l'alta taxa de *tops* produïts via *gluon fussion* (que no genera asimetria), es redueix per a valors alts de $m_{t\bar{t}}$. Les dades experimentals i la predicció teòrica es comparen mitjançant un test de mínims quadrats, que té en compte les incerteses experimentals, teòriques i les correlacions entre les mesures diferencials quan cal. L'ajust es fa a la mesura inclusiva, a les mesures diferencials per separat i, finalment, a totes les mesures diferencials simultàniament. Aquest darrer mètode d'ajust és el que proporciona els límits més ajustats, ja que aprofita tota la informació disponible alhora que manté la sensibilitat a l'asimetria que cada mesura diferencial té.

Els límits es deriven individualment per a cadascun dels 15 operadors de

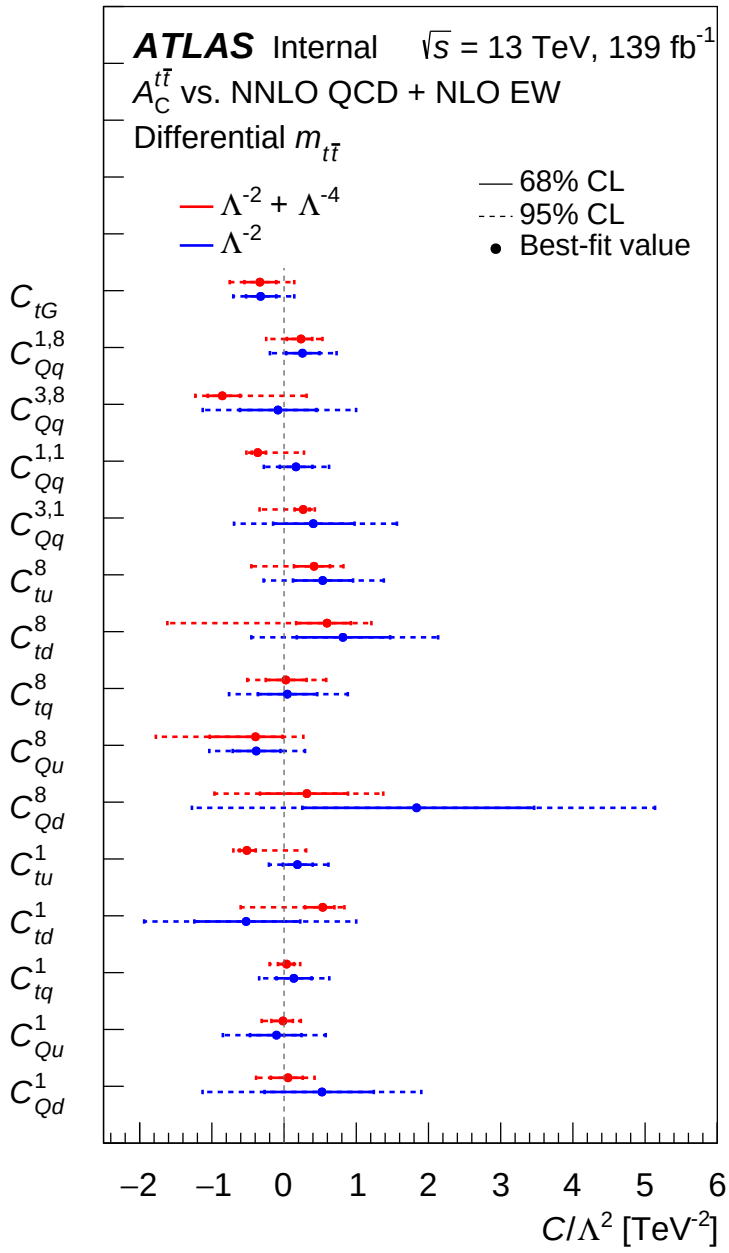


Figure 8.14: Límits individuals obtinguts amb un 68% i 95% de nivell de confiança per als diferents coeficients de Wilson C_i/Λ^2 considerats en aquest treball. Aquests límits es deriven a partir de les mesures diferencials a $m_{t\bar{t}}$ de l'asimetria de càrrega, comparant-les amb les prediccions del SM a NNLO QCD + NLO EW [239]. La contribució dels operadors de Wilson es calculen a NLO usant el paquet SMEFTNLO al generador de Monte Carlo MadGraph5_aMC@NLO [72].

Wilson considerats a partir de la minimització del $\chi 2$, amb un 68% i 95% de nivell de confiança (CL) en dos escenaris diferents: utilitzant una parametrització lineal (Λ^{-2}) i quadràtica ($\Lambda^{-2} + \Lambda^{-4}$) de la contribució dels operadors SMEFT. La Figura 8.14 mostra aquests límits. Com es pot comprovar, tots són compatibles amb 0 (corresponent a la hipòtesi en què no hi ha física BSM) amb un 95% de confiança. Això, però, no descarta la presència de nova física que afecte aquest observable, però fa necessària una reducció de les incerteses, tant experimental com teòrica, per continuar examinant aquesta possibilitat. Comparats amb els límits que podem trobar a la literatura, els derivats en aquest treball són competitius encara provenint d'un conjunt de dades molt més reduït, i contribuiran als ajustaments globals.

Interpretación de la masa Monte Carlo del quark *top*

Com s'ha introduït prèviament, la massa de les partícules descrites en el SM són paràmetres renormalitzats. Cal insistir en que la definició rigorosa d'aquests depèn de l'esquema de renormalització adoptat. Com a paràmetre de la teoria, la massa de qualsevol partícula ha de ser determinada experimentalment. Per això, calen dos ingredients bàsics. El primer, un observable que mostre sensibilitat a la massa del quark *top*. El segon, l'habilitat de poder mesurar la distribució de probabilitat experimentalment i calcular-la teòricament. Pel que fa al primer, assenyalarem que la sensibilitat d'un observable a la massa del quark *top* s'origina principalment en dos sectors diferents. D'una banda, a la pròpia producció de quarks *top* a partir d'un estat inicial particular. Un exemple seria la secció eficaç de producció $t\bar{t}$, que creixeria dràsticament quan l'energia al centre de masses disponible a la col·lisió assolís l'energia mínima per produir un

parell $t\bar{t}$ en repòs, $\sqrt{s} \geq m_{t\bar{t}}$. D'altra banda, la sensibilitat es pot generar en la desintegració del quark *top* i, per tant, en la distribució espacial i energètica dels seus productes de desintegració. Així, els observables que es construeixen des dels productes de desintegració del *top*, com la massa invariant del sistema format pel leptó i el quark *bottom* en estats finals semileptònics o dileptònics, constitueixen observables que posseeixen el que es coneix com a sensibilitat cinemàtica. La distribució d'aquests observables exhibeix una estructura amb una ressonància clara, un pic, la posició i la forma exacta del qual ve determinada per efectes de la interacció forta i electrofeble a escales de l'ordre de Λ_{QCD}.

Depenent de l'observable escollit, es farà servir un mètode o un altre per mesurar la massa del quark *top*. Per a observables la sensibilitat dels quals s'origina en la producció de $t\bar{t}$, hi ha càlculs teòrics rigorosos que poden ser comparats amb les mesures experimentals, sempre que la comparació es faça abans que cap partícula s'haja desintegrat o hadronitzat (conegut com nivell partònic). Aquest mètode es coneix com "indirecte", i dona resultats amb incerteses de l'ordre de 1−2 GeV. Per altra banda, tindrem els mètodes "directes", on s'utilitzen observables que es beneficien de la sensibilitat originada tant en la producció de *tops* com en la seua desintegració. Aquest mètodes no poden utilitzar càlculs únicament basats en QCD. Això és perquè, com es mencionà abans, QCD perd el seu poder predictiu en règims l'energia característica dels quals és de l'ordre de Λ_{QCD}. Al seu lloc s'utilitzen simulacions de MC, les quals ofereixen una descripció aproximada tant del procés d'hadronització com de la detecció de les partícules al detector. Així, les simulacions de MC i les dades experimentals es poden comparar directament, permetent estimar la massa del quark *top* amb una precisió per davall del GeV.

Durant molt de temps s'ha considerat que la definició de la massa del quark *top* com a paràmetre del simulador de MC (massa MC en endavant, m_t^{MC})

és, essencialment, equivalent a la de massa a l'esquema de renormalització *on-shell* o pol. Mentre les mesures directes i indirectes s'han tornat més precises, aquesta identificació ha conduït al conegut com problema d'interpretació de la massa del quark *top*, el qual es pot formular amb la següent pregunta: fins a quin punt la massa pol pot identificar-se amb la massa MC? Des d'un punt de vista estrictament formal, aquesta identificació no es pot fer sistemàticament i analíticament a causa del modelatge dels efectes d'hadronització al MC, principalment, ja que modifiquen la definició de la massa renormalitzada que encara perviu en les primeres etapes de la simulació.

El problema de la interpretació de la massa MC del quark *top* ha rebut una moderada atenció a la literatura, tot i que ha protagonitzat intensos debats a la comunitat científica. A Ref. [121], els autors troben un límit superior entre la massa MC i la massa pol: $m_t^{\text{pole}} - m_t^{\text{MC}} < 2$ GeV. Els autors de Ref. [124, 210] han liderat la investigació sobre aquest tema, desenvolupant diversos càlculs teòrics basats en observables amb sensibilitat cinemàtica. Aquests càlculs inclouen de forma rigorosa efectes d'hadronització a través de paràmetres lliures que s'han d'ajustar a les dades, juntament amb la massa del quark *top* en un esquema de renormalització donat. Això és essencial per traduir en termes d'una massa renormalitzada l'evolució d'un observable a causa del model d'hadronització utilitzat al simulador de MC.

En concret, a Ref. [210] es proposa un càlcul amb precisió NLL que permet predir la distribució de la massa de jets que contenen els productes de desintegració hadrònica de quarks *top*, després d'hadronitzar. Amb aquest càlcul, s'ha estudiat com els diferents generadors de MC de l'experiment ATLAS poden alterar la distribució de massa de jets i com aquesta alteració pot modificar el valor de la massa renormalitzada que se n'extreu. Així, podem escriure aquesta

relació com:

$$m_t^{\mathrm{MC}} = m_t^{\mathrm{MSR}}(1\ \mathrm{GeV}) + \Delta_m^{\mathrm{MSR}}\ ,$$

on es pren com a referència l'esquema de renormalització MSR a 1 GeV, ja que posseeix una sèrie de característiques que el fan conceptualment més proper a la massa MC tal com s'implementa als simuladors.

Per avaluar la grandària del terme Δ_m^{MSR}, que representa la diferència entre la massa MC i la MSR a causa de principalment el modelatge de l'hadronització al MC, es generen prediccions teòriques variant els tres paràmetres lliures disponibles: la massa MSR m_t^{MSR}, $\emptyset 1q$ i x_2. Els dos darrers codifiquen els efectes d'hadronització. Per a desacoblar, parcialment almenys, la correlació entre aquestes tres variables, l'ajust es fa en tres intervals del moment transvers del jet: $[750, 1000, 1500, 2000]$ GeV. Mitjançant un test de mínims quadrats, la distribució de massa de jets teòrica es compara amb la simulació del MC nominal de l'experiment ATLAS, que utilitza el generador POWHEG per a la interacció protó-protó i PYTHIA 8 per a la fragmentació/desintegració del quark *top* i l'hadronització dels seus productes de desintegració. La massa MC del quark *top* es fixa a 172.5 GeV. La Figura 8.15 mostra aquesta comparació per als tres intervals de moment transvers considerats. Com es pot apreciar, la predicció teòrica és compatible amb la distribució de MC dins la banda d'incertesa teòrica, en el rang d'ajust considerat (172.5 – 180.0 GeV).

La minimització del $\chi 2$ proporciona el càlcul teòric que millor descriu les dades de MC. Els paràmetres associats són:

$$m_t^{\mathrm{MSR}}(R = 1\ \mathrm{GeV}) = 172.42 \pm 0.10\ \mathrm{GeV},\quad \Omega_{1q}^{\circ\circ} = 1.49 \pm 0.03\ \mathrm{GeV},\quad x_2 = 0.52 \pm 0.09,$$

sent la incertesa d'origen purament estadístic. Amb això, la relació entre la massa

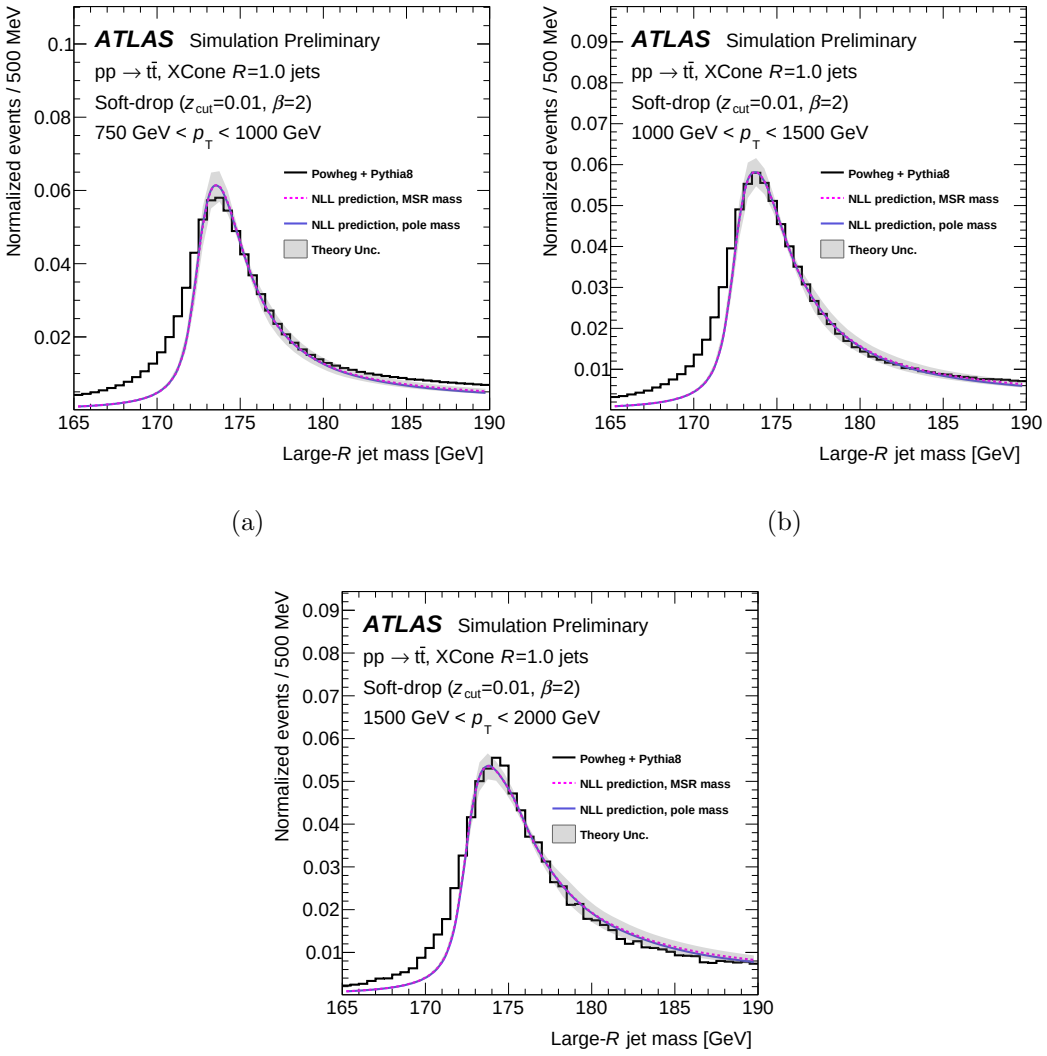


Figure 8.15: Distribució de *top* jets per a la mostra nominal POWHEG +PYTHIA 8 (línia negra) i la predicció NLL de Ref. [210] que millor descriu les dades. El càlcul teòric es realitza a l'esquema de massa MSR (corba rosa puntejada) i a l'esquema de la massa pol (corba porpra contínua). La comparació es fa en el rang 172.5 GeV $< M_J <$ 180 GeV.

MSR i la massa MC és:

$$m_t^{\mathrm{MC}} = m_t^{\mathrm{MSR}}(1\ \mathrm{GeV}) + 80^{+350}_{-400}\ \mathrm{MeV}, \tag{8.3}$$

on la incertesa total té en compte l'error en el càlcul teòric, l'elecció del rang d'ajust, els intervals de moment transvers utilitzats i el modelatge de l'*Underlying Event* (UE). Si es repeteix l'ajust però amb prediccions teòriques per a la massa pol s'obté:

$$m_t^{\mathrm{MC}} = m_t^{\mathrm{pole}} + 350^{+300}_{-360}\ \mathrm{MeV}. \tag{8.4}$$

Amb això, concloem que encara que la massa MC és compatible tant amb la massa MSR com amb el pol dins de les incerteses considerades, està numèricament més a prop de la massa MSR, tal com esperàvem. La relació entre la massa MC i la massa pol és compatible amb l'obtinguda en col·lisions e^+e^- [124]. La relació donada a l'Eq. 8.3 s'ha derivat per a configuracions alternatives del MC nominal que poguessen afectar la distribució de massa de jets. Totes les variacions trobades estan per devall dels 100 MeV. A més, s'ha comprovat que aquesta relació és estable dins de 200 MeV quan s'utilitzen altres definicions de jet similars a la utilitzada com a referència.

Finalment, per il·lustrar el paper que exerceix el model d'hadronització del simulador de MC, s'ha substituït Pythia 8 per Herwig 7, el qual incorpora un model d'hadronització diferent. Tot i que la simulació de la interacció protó-protó és exactament la mateixa, tots dos MC prediuen una distribució de massa de jets molt diferent, com il·lustra la Figura 8.16. La massa MSR que es troba per a Powheg + Herwig 7 és, en aquest cas:

$$m_t^{\mathrm{MSR}}(1\ \mathrm{GeV}) = 172.27 \pm 0.09\ \mathrm{GeV}, \quad \Omega_{1q}^{\circ\circ} = 1.9 \pm 0.07\ \mathrm{GeV}, \quad x_2 = 0.98 \pm 0.12,$$

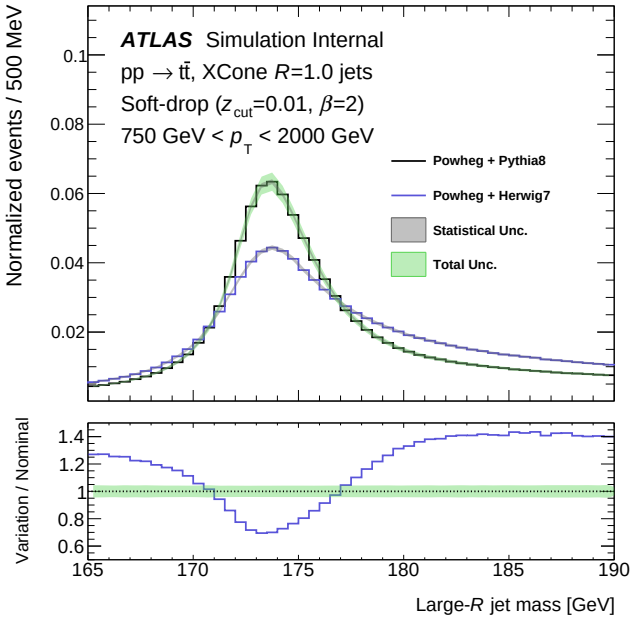


Figure 8.16: Comparació de successos $t\bar{t}$ generats amb POWHEG i processats amb PYTHIA 8 o HERWIG 7. La ràtio entre les dues distribucions es mostra al panell inferior.

la qual és quasi 150 MeV més baixa que per POWHEG +PYTHIA 8, encara que són compatibles dins de l'error. La diferència en la distribució de massa és absorbida majoritàriament pels paràmetres d'hadronització $\Omega_{1q}^{\circ\circ}$ i x_2.

Amb tot, cal dir que aquests resultats són una peça important per resoldre el problema de la interpretació de la massa del quark *top*, ja que ha permès posar el focus en el significat de la massa de MC en mesures basades en dades obtingudes en col·lisions protó-protó. Una major precisió formal en el càlcul teòric que a més incloga efectes de UE permetrà continuar investigant la interpretació de la massa de manera més precisa i exhaustiva.

La massa del quark *bottom* i el bosó d'Higgs

Finalment, en aquesta tesi es presenta la primera mesura del valor de la massa del quark *bottom* a l'escala d'energia de la massa del bosó d'Higgs. Aquest resultat ha permès estudiar més estrictament l'evolució de la massa del quark *bottom* amb l'energia del procés en què es produeix. Amb aquest fí, s'han utilitzat mesures existents de baixa energia de la massa del quark *bottom* a l'escala de la seua massa massa m_b i a l'escala de la massa del bosó Z:

$$m_b(m_b) = 4.18^{+0.03}_{-0.02}\ \text{GeV}\ , \tag{8.5}$$

$$m_b(m_Z) = 2.82 \pm 0.28\ \text{GeV}\ . \tag{8.6}$$

Pel que fa a la mesura a l'escala de la massa del bosó d'Higgs, s'han aprofitat les mesures de l'experiment ATLAS i CMS de la desintegració del bosó d'Higgs en parells $b\bar{b}$ i ZZ:

$$\begin{aligned} \mu^{b\bar{b}}/\mu^{ZZ} &= 0.87^{+0.22}_{-0.17}\ (\text{stat})\ ^{+0.18}_{-0.12}\ (\text{syst}) = 0.87^{+0.28}_{-0.21}\quad (\text{ATLAS}), \\ \mu^{b\bar{b}}/\mu^{ZZ} &= 0.84^{+0.27}_{-0.21}\ (\text{stat})\ ^{+0.26}_{-0.17}\ (\text{syst}) = 0.84^{+0.37}_{-0.27}\quad (\text{CMS}). \end{aligned} \tag{8.7}$$

Com es por observar, s'utilitza la ràtio de l'amplària de destingració del bosó d'Higgs a parells $b\bar{b}$ i ZZ, $\Gamma(H \to b\bar{b})/\Gamma(H \to ZZ)$, per reduïr la incertesa de la mesura experimental. A més a més, el resultat de l'Eq. 8.7 es dona respecte a la predicció del SM ($\mu = \Gamma^{\text{exp}}/\Gamma^{\text{SM}}$).

Ara bé, per saber qual és la massa del quark *bottom* que, amb major probabilitat, donaria aquest resultat, cal en primer lloc parametritzar la dependència de $\mathcal{B}(H \to b\bar{b})/\mathcal{B}(H \to ZZ)$ amb m_b. Això es fa mitjançant càlculs teòrics numèrics, que gaudixen de molt bona precisió pel fet de fer-se a l'escala

d'energia de la massa del bosó d'Higgs. La parametrització dona:

$$\frac{\Gamma(H \to b\bar{b})}{\Gamma(H \to ZZ)} = 2.82\frac{m_b^2}{\text{GeV}^2} - 0.0014\frac{m_b^4}{\text{GeV}^4} + \mathcal{O}(m_b^6) \ . \tag{8.8}$$

En aquesta relació es tenen en compte incerteses en els valors de la massa del bosó d'Higgs i la constant de interacció forta que n'hi ha que prendre per fer els càlculs. També per la pròpia precisió del càlcul. En qualsevol cas, la incertesa d'aquesta parametrització és molt més petita que la incertesa de les mesures experimentals. Amb tot, els valors obtinguts pels experiments ATLAS i CMS es introduïxen en l'Eq. 8.8. La combinació dels dos resultats dóna un valor de:

$$m_b(m_H) = 2.60^{+0.36}_{-0.31} \text{ GeV}, \tag{8.9}$$

Aquest resultat i els de l'Eq. 8.6, es mostren a la Figura 8.17 juntament amb la predicció de l'evolució segons el SM, obtinguda amb el programa REvolver [320]. Trobem un bon acord entre la predicció del SM i les mesures experimentals, dins de les incerteses actuals.

Per comprovar amb més precisió i rigor la hipòtesi de la dependència de la massa del quark *bottom* amb l'escala d'energia o *running*, s'utilitzen prediccions teòriques en què l'evolució predita pel SM és regulada a través del paràmetre x. Així, si $x = 0$ deixa d'haver evolució amb l'energia i per a $x = 1$, l'evolució és exactament l'esperada segons el SM. Usant com a punt de partida diferents valors de $m_b(m_b)$ i mitjançant un ajustament per mínims quadrats, podem trobar quin és el valor de $m_b(m_b)$ i x que millor s'ajusten a les mesures experimentals. Després de minimitzar el $\chi 2$ per a les dues variables en qüestió, trobem:

$$\begin{aligned} m_b(m_b) &= 4.18^{+0.03}_{-0.02} \text{ GeV} \ , \\ x &= 1.08 \pm 0.15 \ (\text{exp}) \ \pm 0.05 \ (\alpha_s). \end{aligned} \tag{8.10}$$

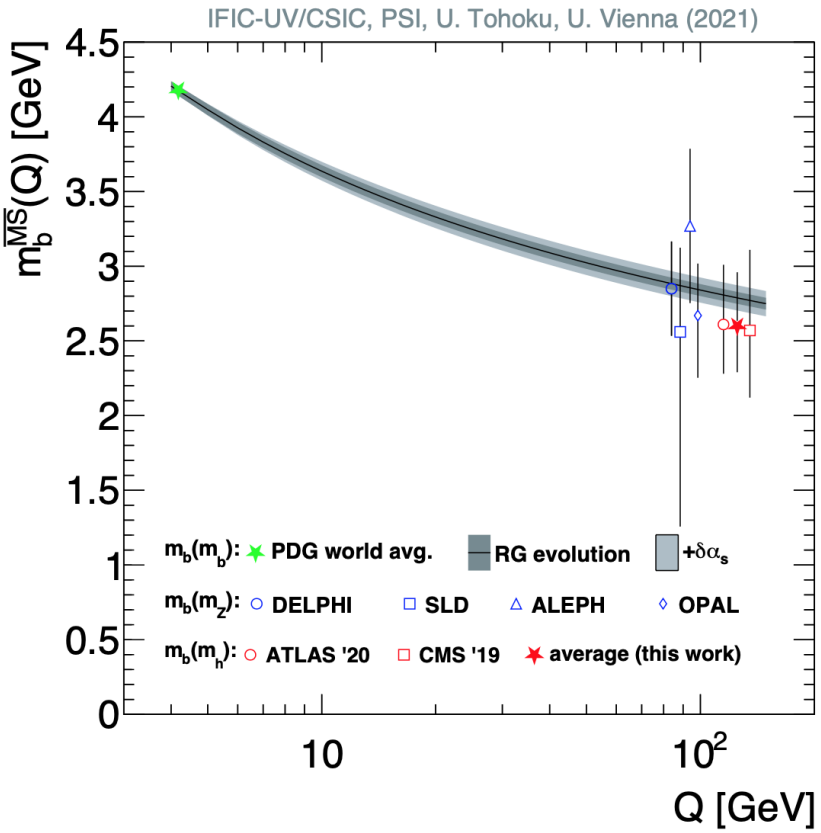


Figure 8.17: Evolució de la massa del quark *bottom* a l'esquema de renormalització $\overline{MS}$ amb l'escala d'energia. Les mesures inclouen la mitjana mundial de la massa a l'escala de la massa del bottomonium $m_b(m_b)$ obtingudes en factories B, la mesura a l'escala de la massa del bosó Z $m_b(m_Z)$ a partir de *jet rates* de LEP i SLC i, finalment, la mesura $m_b(m_H)$ a partir desintegracions del bosó d'Higgs. La predicció de l'evolució segons el SM s'obté amb el programa REvolver [320]. Les bandes representen la incertesa associada a aquesta predicció.

D'una banda, el valor ajustat per a $m_b(m_b)$ és totalment compatible amb el proporcionat a l'Eq. 8.6. D'altra banda, el valor de x és també compatible amb l'evolució del SM dins de la incertesa associada, i descarta l'escenari on la massa no evoluciona amb l'energia amb una significància de 7σ. Un ajustament sense la mesura realitzada en aquest treball donaria una significància just per davall de 5σ, la qual no es consideraria definitiva. Aquest resultat estableix la primera demostració de la dependència i evolució de la massa del quark *bottom* amb l'escala d'energia.

Acronyms

ATLAS	**A** **T**oroidal **L**HC **A**pparatu**S**
BSM	**B**eyond the **S**tandard **M**odel
BW	**B**reit-**W**igner
CERN	**C**onseil **E**uropeen pour la **R**eserche **N**ucleaire
CMS	**C**ompact **M**uon **S**olenoid
CL	**C**onfidence **L**evel
CR	**C**olor **R**econnection
EFT	**E**ffective **F**ield **T**heory
EW	**E**lectro**W**eak
FB	**F**orward-**B**ackward
FBU	**F**ully **B**ayesian **U**nfolding
FF	**F**orward-**F**olding
FSR	**F**inal-**S**tate **R**adiation
HL-LHC	**H**igh-**L**uminosity **LHC**
ID	**I**nner **D**etector
IR	**I**nfra**R**ed
IRC	**I**nfra**R**ed and **C**ollinear
ISR	**I**nital-**S**tate **R**adiation
JER	**J**et **E**nergy **R**esolution
JES	**J**et **E**nergy **S**cale

JMR	**J**et **M**ass **R**esolution
JMS	**J**et **M**ass **S**cale
LHC	**L**arge **H**adron **C**ollider
LH	**L**eft-**H**anded
LL	**L**eading **L**og
LO	**L**eading **O**rder
MC	**M**onte **C**arlo
MPI	**M**ulti-**P**arton **I**nteractions
ME	**M**atrix **E**lement
MEC	**M**atrix **E**lement **C**orrection
MS	**M**uon **S**pectrometer
NLO	**N**ext-to-**L**eading **O**rder
NLL	**N**ext-to-**L**eading **L**og
NNLO	**N**ext-to-**N**ext-to-**L**eading **O**rder
PS	**P**arton **S**hower
PDF	**P**arton **D**ensity **F**unction
QCD	**Q**uantum **C**hromo**D**ynamics
QED	**Q**uantum **E**lectro**D**ynamics
QFT	**Q**uantum **F**ield **T**heory
RC	**R**e**C**lustered
RGE	**R**enormalization **G**roup **E**quations
RH	**R**ight-**H**anded
RS	**R**ernormalization **S**chemes
RTC	**R**ecoil-**T**o-**C**olored
SCET	**S**oft and **C**ollinear **E**ffective field **T**heory
SD	**S**oft-**D**rop
SM	**S**tandard **M**odel

SMEFT	**S**tandard **M**odel **E**ffective **F**ield **T**heory
SSB	**S**pontaneous **S**ymmetry **B**reaking
TA	**T**rack-**A**ssisted
UC	**U**ltra-**C**ollinear
UE	**U**nderlying **E**vent
UV	**U**ltra**V**iolet

pp	proton-proton
$p\bar{p}$	proton-antiproton
$\sqrt{s}$	center-of-mass energy
$\overline{MS}$	modified minimal subtraction renormalization mass scheme
MSR	MSR renormalization mass scheme

List of Publications

[I] ATLAS Collaboration, "In-situ measurements of the ATLAS large-radius jet response in 13 TeV pp collisions," July 2017. **ATLAS-CONF-2017-063**.

Responsible of the large-R jet mass calibration.

[II] ATLAS Collaboration, "Jet reclustering and close-by effects in ATLAS run II," July 2017. **ATLAS-CONF-2017-062**.

Responsible of the large-R jet mass calibration.

[III] ATLAS Collaboration, "Inclusive and differential measurement of the charge asymmetry in $t\bar{t}$ events at 13 TeV with the ATLAS detector", July 2019. **ATLAS-CONF-2019-026**.

Responsible of the charge asymmetry SMEFT interpretation.

[IV] ATLAS Collaboration, "In situ calibration of large-radius jet energy and mass in 13 TeV proton-proton collisions with the ATLAS detector," *Eur. Phys. J. C 79 (2019) 135*, March 2019. **CERN-EP-2018-191**.

Responsible of the large-R jet mass calibration.

[V] ATLAS Collaboration, "Measurement of the ATLAS Detector Jet Mass Response using Forward Folding with 80 fb^{-1} of $\sqrt{s}$ = 13 TeV pp data," July 2020. **ATLAS-CONF-2020-022**.

Contact editor and responsible of the large-R jet mass calibration from $t\bar{t}$ events.

[VI] ATLAS Collaboration, "Inclusive and differential measurement of the charge asymmetry in $t\bar{t}$ events at 13 TeV with the ATLAS detector", 2022. **Under internal approval**.

Responsible of the charge asymmetry SMEFT interpretation.

[VII] ATLAS Collaboration, "Towards a precise interpretation of the top quark mass parameter in ATLAS Monte Carlo samples", July 2021. **ATL-PHYS-PUB-2021-034**.

Contact editor and responsible of the fitting code development and fits.

[VIII] J. Aparisi, J. Fuster, A. Irles, *et al.*, "m_b at m_H: The Running Bottom Quark Mass and the Higgs Boson", *Phys. Rev. Lett. 128, 122001*, March 2022. **arXiv:2110.10202 [hep-ph]**.

Responsible of the bottom quark mass extraction and fits to the mass evolution.

List of Figures

1.1 Feynman diagrams of two basic interaction processes in the QED Lagrangian. 9

1.2 Feynman representation of the charged-current interactions and the third-point interaction vertex present in the electroweak Lagrangian. Same third- and fourth-point vertices describe self-interacting bosons, but restricted to the weak ones since the photon does not self-interact. 13

1.3 Feynman diagrams of two basic interaction vertices of the QCD Lagrangian: on the left, gluon emission by quarks and on the right, quartic gluon self-interaction. 18

1.4 Electron-positron annihilation, via Z or γ bosons, with a muon-antimuon pair in the final state. 21

1.5 Higher-order quantum corrections to the fermion mass through γ loops. 22

1.6 Evolution of the MSR mass (blue line) with the energy scale R. Further low-energy short-distance masses at several energy scales are displayed: m^{PS}, m^{1S} and $\overline{m}$ ($m^{\overline{MS}}$ in this work). Image taken from Ref. [50]. 28

1.7 Summary of measurements of α_S as a function of the energy scale Q. The corresponding order in QCD perturbation theory used to derive α_S is indicated withinbrackets (where res. stands for resummation, a further technique needed to make predictions covered in Section 1.6.3). Image taken from Ref. [26]. 31

1.8 Representation of the quantum fluctuations undergoing inside a proton. Figure taken from Ref. [63]. 34

1.9 Constrained Parton Density Functions for unpolarized protons at two different factorization scales, including in the global fit data from LHC experiments. Image taken from Ref.[68]. 37

1.10 Schematic representation of a hadron-hadron collision as simulated by a Monte Carlo generator: the incident hadrons are represented through three valence quarks as straight green lines. The primary hard parton collision, big red blow dominating the center of the figure, gives three smaller blows as a result. Each of them is surrounded by a tree-like structure representing the parton shower. Parton-to-hadron transitions can be identified as light green blobs, while dark green blobs stand for hadron decay. A secondary collision with its corresponding dynamics is indicated by the purple blow. Figure taken from Ref. [82]. 44

2.1 Comparison of world combined m_W and m_t results with contours at 68% ad 95% Confidence Level computed from scans of such parameters, excluding (gray) and including (blue) m_H in the fit. Figure from Ref. [90]. 54

2.2 One-loop correction to the Higgs propagator involving top quarks. 55

2.3 Stability of the EW vacuum as a function of the measured m_t and m_H with 1,2 and 3σ CL contours. Figure taken from Ref. [91]. . . 56

2.4 Leading Order Feynman diagrams representing the dominant production modes of $t\bar{t}$ pairs in hadron collisions. Leftmost and center diagrams refer to $s-$ and $t-$channel in gluon-gluon fusion. The rightmost diagram stands for the $q\bar{q}$ annihilation process. . . . 57

2.5 Comparison of measured and predicted total production cross-section of $t\bar{t}$ (upper panel) and single-top quarks (lower panel) as a function of the center-of-mass energy $\sqrt{s}$. Figure from Ref. [26]. 58

2.6 Rapidity distributions of top and antitop quarks produced in pairs at the Tevatron (left) and the LHC (centre, right). Figure taken from Ref. [99]. 60

2.7 Comparison of measured inclusive charge asymmetries at the LHC (ATLAS and CMS in vertical lines) and the Tevatron (DØ and CDF, horizontal lines) to Standard Model predictions. Experimental values are represented with the associated uncertainty band at 68% Confidence Level. Several models incorporating beyond Standard Model effects are also included: W', heavy axigluon (G_μ), scalar isodoublet (ϕ), color-triplet scalar (ω^4) and color-sextet scalar (Ω^4). Figure from Ref. [101]. 62

2.8 Feynman diagrams for $u\bar{u} \to t\bar{t}$ process. (a) diagram represents the Standard Model amplitude, and (b) and (c) the modified interaction vertex due to the insertion of dimension-six operators affecting the $t\bar{t}G$ coupling and four-fermion processes, respectively. Figure taken from Ref. [114]. 65

2.9 On the left, sensitivity of the reconstructed top quark mass from hadronic top decays (semileptonic channel) to three m_t^{MC} values in simulated events. On the right, summary of ATLAS and CMS direct top quark mass measurements. Figures from Refs. [115] and [116]. 66

2.10 Summary of top quark pole mass determinations performed by the ATLAS and CMS experiments from inclusive and differential cross-sections. Vertical bands correspond to m_t^{MC} averages from $\sqrt{s}$ = 7 and 8 TeV data. Figure from Ref. [116]. 68

3.1 Sketch of the CERN accelerator complex with the experimental detectors placed at each interaction point in the LHC (blue) and LEP (red) era. 77

3.2 On the left: mean number of interactions per bunch crossing during Run 2 for several periods with different characteristic instantaneous luminosity. On the right: cumulative luminosity versus time delivered to ATLAS (green), recorded by ATLAS (yellow), and certified to be good quality data (blue) during stable beams for pp collisions in Run 2. Images from Ref. [134]. 80

3.3 Cylindrical coordinates system typically used to describe particles position in detectors at circular colliders. 81

3.4 Illustration of the signals left by different particles as they pass through several detectors. Dashed lines represent particles that leave no trace in the sub-detector, whilst horizontal solid lines stand for particles that only loose a small fraction of their energy. Figure from Ref. [135]. 83

3.5 The ATLAS detector and its sub-detectors. Image from Ref. [136]. 84

3.6 The ATLAS inner detector. Image from Ref. [137]. 85

3.7 The ATLAS calorimeter system. Image from Ref. [138]. 86

3.8 Components of the muon spectrometer in the ATLAS experiment. Image from Ref. [139]. 88

3.9 Evolution of the ATLAS Computing Model from a hierarchical (left) to a mess (right) mode. 91

4.1 Simplified representation of the different stages at which jets can be built: partons (purple), stable hadrons (green) and electric signals (red) are the jets constituents at parton, particle and detector level, respectively. 97

4.2 Jet shapes and catchment areas for two illustrative jet clustering algorithms. Image taken from Ref. [182]. 99

4.3 Sketch of the hadronic final state topology involving top quarks in the resolved (a) and boosted (b) regime. 102

4.4 Jet mass resolution for W jets as a function of the large-R truth jet p_{T} for jet building blocks. 103

4.5 Impact in the total jet mass of the jet constituents as a function of the angular distance between the jet axis and jet constituent ΔR and the p_{T} fraction carried by the constituent. Image taken from Ref. [135]. 105

4.6 Diagram depicting the jet trimming procedure. Figure taken from Ref. [195]. 106

4.7 The large-R jet calibration and reconstruction chain. Image taken from Ref. [197]. 108

4.8 Monte Carlo based calibration of the average large-R jet mass. . . 109

4.9 The trimmed jet m_{calo}, trimmed jet m_{TA} and reclustered jet mass distributions obtained from events passing the (left) W selection and (right) top quark selection from $t\bar{t}$ events in the first, second and third row, respectively. The "Other" category refers to diboson and Z+jets events. Figure taken from Ref. [197]. 117

4.10 On the top: average mass response obtained from the nominal Monte Carlo simulation of calorimeter-only trimmed jets (figure taken from Ref. [197]). On the bottom: illustrative two-dimensional $\chi 2$ fit of the Monte Carlo templates generated via the Forward-Folding method to data collected by the ALTAS detector. 118

4.11 The relative Jet Mass Scale and Resolution for several jet mass definitions as a function of the jet p_{T}. Black and red markers represent results for W-like and top-like jets. The inner error bar stands for the statistical uncertainty and the outer error bar includes the statistical and systematic uncertainty. Dashed blue lines, corresponding to ±1% for the JMS and ±10% for the JMR, are drawn for reference. Figure taken from Ref. [197]. 121

4.12 Breakdown of the total uncertainty associated to the relative Jet Mass Scale in (a) and Jet Mass Resolution in (b) for calorimeter-only trimmed jets in $t\bar{t}$ events. Similar trends are found for Track-Assisted and Reclustered jet masses. The estimation of the Monte Carlo modelling uncertainties is limited in accuracy by the statistics available in the Monte Carlo samples used. Figure taken from Ref. [197]. 122

5.1 Sketch of the different sources of radiation entering in the catchment area of a top-initiated large-R jet. 125

5.2 Dependence of the jet mass distribution with the jet radius R. Figure from Ref. [210]. 126

5.3 Upper panel: allowed region for z_{cut} values as a function of the jet transverse momentum p_{T}. Lower panel: dependence of the jet mass distribution with z_{cut} for $\beta = 2$ and jet $p_{\text{T}} > 750$ GeV. Figures from Ref. [210]. 127

5.4 Sequence of effective field theories participating in the computation of the invariant jet mass distribution. Modified figure from Ref. [211]. 129

5.5 Comparison of the simulated jet mass distribution with Pythia 8 (in blue) to the NLL prediction (in red) for two different p_{T} ranges: $p_{\text{T}} > 750$ GeV on the upper panel and $p_{\text{T}} > 1000$ GeV on the lower panel. Figures from Ref. [210]. 130

5.6 Large-R jet mass distributions obtained from the reference Pythia 8 setup for (a) ungroomed XCone jets with $R = 1$ and (b) the same jets after applying a 'light' soft-drop grooming ($\beta = 2$, $z_{\text{cut}} = 0.01$). Monte Carlo simulations at parton and particle level are shown, where the labels FSR and MPI respectively indicate whether final-state radiation in resonances and multi-parton interactions (*i.e.* underlying event) are activated in the simulation. 140

5.7 Breit-Wigner (a) derived for a nominal (172.5 GeV) and alternative (172.0 and 173.0 GeV) Monte Carlo mass values and (b) the corresponding nominal, generated and reweighted truth jet mass distributions. 144

5.8 Comparison of the NLL prediction of Ref. [210] to several PYTHIA 8 predictions with m_t^{MC} = 172.5 GeV. The distributions correspond to (a) the Monte Carlo and NLL predictions at parton level and with final-state radiation in resonance decays (PYTHIA 8 setting 'FSRinRes') turned off, (b) the particle-level prediction with FSRinRes and underlying event modelling (PYTHIA 8 setting 'MPI') turned off, (c) the particle-level predictions with FSRinRes turned on and MPI switched off, and (d) the particle-level predictions with FSRinRes and MPI turned on. The distributions are normalized and the fit is performed on the interval 170 GeV $< M_J <$ 180 GeV, where M_J is the large-R jet mass. Fits are performed in three bins of the ungroomed large-R jet p_{T}; these figures are shown in a single p_{T} bin for illustrative purposes. 146

5.9 The $\chi 2$ value and the statistical uncertainty obtained from a fit with the NLL calculation of Ref. [210] to a Monte Carlo prediction at particle level, as a function of the lower limit of the fit. The Monte Carlo prediction is based on PYTHIA 8 with a top quark mass of 172.5 GeV. The multi-parton interaction setting is turned off, and the final-state radiation in resonances is turned on in the Monte Carlo. 149

5.10 NLL prediction of Ref. [210] for the normalized particle-level jet mass distribution of large-R jets containing a boosted hadronic top quark decay (black line). The colored and dashed/dotted lines are obtained by varying the five scales in the calculation. The grey band corresponds to the envelope of all scale variations. All curves are normalized to the same area in the groomed large-R jet mass interval of 172.5 – 180 GeV. 150

5.11 Impact of scale variations in the NLL calculation expressed in terms of the top quark MSR mass. The values of m_t^{MSR} are found by fitting the nominal NLL calculation to the predictions obtained with different choices of the five scales. The three parameters of the calculation are floated in the fit. For the calculations with alternative scale choices the parameters are set to the best-fit values for the ATLAS POWHEG +PYTHIA 8 sample: $m_t^{\mathrm{MSR}}(1\ \mathrm{GeV}) = 172.42$ GeV (indicated by a vertical line), $\Omega_{1q}^{\circ\circ} = 1.49$ GeV, $x_2 = 0.52$. 152

5.12 Large-R jet mass for boosted, hadronically-decaying top quarks for (left) the PYTHIA 8 A14 Var1 up and down variations and (right) several color reconnection models available in PYTHIA 8. For reference, the distribution is also shown for the nominal A14 tune, and with multi-parton interactions disabled in the left panel. Large-R jets are reconstructed using the XCone algorithm and light soft-drop grooming. 154

5.13 Large-R jet mass distribution for Monte Carlo events generated with different varied settings of the standard configuration. 159

5.14 Large-R jet mass distributions for boosted, hadronically-decaying top quarks. Results are shown for $t\bar{t}$ events generated with POWHEG + PYTHIA 8 with the multi-parton interaction model (PYTHIA 8 setting 'MPI') disabled. Several variations of the A14 tune are shown along with the nominal setup, Var2 in panel (a), Var3a in panel (b), Var3b in panel (c) and Var3c in panel (d). . . . 161

5.15 Equivalent shifts in the top Monte Carlo mass m_t^{MC} with respect to the nominal POWHEG +PYTHIA 8 Monte Carlo sample. The offsets are found by fitting the nominal MC mass template with a floating MC mass parameter to several alternative choices of the POWHEG +PYTHIA 8 configuration and to the POWHEG + HERWIG 7 and MADGRAPH5_aMC@NLO +PYTHIA 8 samples. The multi-parton interaction model is disabled in these fits. 162

5.16 The particle-level jet mass distribution of the nominal POWHEG +PYTHIA 8 sample with the multi-parton interaction model disabled (black histogram) and the NLL prediction of Ref. [210] which best describe the Monte Carlo prediction (smooth curves). The NLL calculation is performed in the MSR mass scheme (dashed pink curve) and in the pole mass scheme (purple continuous curve). In both mass schemes, the three parameters of the calculation, the top quark mass, Ω_{1q}^{∞} and x_2, are fitted to find the best description of the three p_{T} intervals used in the calibration procedure. The distributions are normalized and the fit is performed on the interval 172.5 GeV $< M_J <$ 180 GeV. . . . 164

5.17 The $\chi 2$ results as a function of the MSR mass $m_t^{\mathrm{MSR}}(1\ \mathrm{GeV})$ and the parameters of the shape function $\Omega_{1q}^{\circ\circ}$ and x_2, obtained from the fit with the NLL prediction of Ref. [210] to the particle-level jet mass distribution for boosted top quarks in the nominal ATLAS Powheg +Pythia 8 sample. The results are shown in the two-dimensional plane of (a) MSR mass versus $\Omega_{1q}^{\circ\circ}$, (b) MSR mass versus x_2 and (c) $\Omega_{1q}^{\circ\circ}$ versus x_2. In each case, the 2D distribution is obtained by marginalizing over the third parameter. The three parameters of the calculation are varied to find the best description of the three p_{T} intervals used in the calibration procedure. 165

5.18 The MSR mass extracted from samples where a given aspect of the $t\bar{t}$ production process is altered with respect to the nominal Monte Carlo setup (dashed line). The variations affect the PDF, α_s and the renormalization and factorization scale, as well as the rate of Initial- and Final-State Radiation. The vertical band indicates the uncertainty on the nominal fit value. 167

5.19 Comparison of $t\bar{t}$ events generated with Powheg and showered with either Pythia 8 or Herwig 7. The total error band accounts for statistical uncertainty, as well as variations of PDFs, α_s and renormalization / factorization scales within the nominal Pythia 8 sample. The ratio between nominal and the alternative Monte Carlo sample is provided in the bottom panel. 168

5.20 The impact of different grooming configurations and jet clustering algorithms on the top jet mass distribution for (a) large-R XCone jets groomed with several soft-drop configurations, and (b) ungroomed and groomed (nominal soft-drop configuration) large-R jets clustered with the XCone and anti-k_t algorithms. 171

5.21 Study of the linearity of the MSR and Monte Carlo mass relation based on PYTHIA 8 standalone simulations. The fit procedure is frozen to the nominal NLL-based template-fit, and only the lower limit of the fit range is adjusted accordingly to the value of the top Monte Carlo mass with which the samples were generated. 174

6.1 Comparison of the observed data and the Monte Carlo prediction (signal and backgrounds) for the differential charge asymmetry measurement as a function of $m_{t\bar{t}}$ in the semileptonic channel and boosted regime. The light green bands represents the sum in quadrature of the statistical and pre-marginalization systematic uncertainties. The Monte Carlo is scaled to data. 184

6.2 Comparison of the observed data and the Monte Carlo prediction (signal and backgrounds) for the asymmetry $m_{t\bar{t}}$ differential measurement in the dileptonic channel and resolved regime. The dark uncertainty bands in the ratio plots stand for the statistical uncertainty, whilst the light bands accounts for statistical and systematic uncertainties added in quadrature. 186

6.3 The unfolded inclusive and differential charge asymmetries as a function of the $t\bar{t}$ invariant mass on the upper and lower panels, respectively. The vertical bar represents the total uncertainty, accounting for the systematic and statistic uncertainties. The impact of the linear term of the $C_{ut}^{(8)}$ Wilson coefficient on the $A_{\mathrm{C}}^{t\bar{t}}$ prediction for two different values is shown as dashed lines. In the dileptonic channel the last two $m_{t\bar{t}}$ bins have low statistics; the vertical bars representing the total uncertainties are shown, while the $A_{\mathrm{C}}^{t\bar{t}}$ results, which are outside of the visible scale, are indicated by the direction of the associated arrows. 189

6.4 $\chi 2$ distribution from which limits on $C_{ut}^{(8)}$ operator are derived when considering (upper-left panel) only SM-EFT terms and (upper-right panel) SM-EFT and EFT-EFT terms in the EFT contribution. Bottom panel shows the comparison between ATLAS and predicted $A_{\mathrm{C}}^{t\bar{t}}$ values, with the 68% and 95% Confidence Level bands drawn in green and yellow, respectively. Inclusive $t\bar{t}$ events. 195

6.5 Individual 68% and 95% Confidence Level limits on the Wilson coefficient $C_{ut}^{(8)}/\Lambda^2$ in the Standard Model Effective Field Theory. The bounds are derived from the measured charge asymmetry presented in this chapter, combining the semileptonic and dileptonic channels. The theory uncertainty corresponds to the NNLO QCD + NLO EW calculation of Ref. [239]. The impact of dimension-six operators is parameterized at NLO accuracy in QCD using the SMEFTNLO package in MadGraph5_aMC@NLO. Bounds are also shown from the forward-backward asymmetry measurements in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.96 TeV at the Tevatron and the charge asymmetry measurements in pp collisions at a center-of-mass energy of 8 TeV in LHC Run 1. 198

6.6 Individual 68% and 95% Confidence Level limits on the Wilson coefficients C_i/Λ^2 in the Standard Model Effective Field Theory from the differential charge asymmetry measurements versus $m_{t\bar{t}}$. The bounds are derived from the measured charge asymmetry presented in this chapter, combining the semileptonic and dileptonic channels. The theory uncertainty corresponds to the NNLO QCD + NLO EW calculation of Ref. [239]. The impact of dimension-six operators is parameterized at NLO accuracy in QCD using the SMEFTNLO package in MG5_aMCNLO. 201

7.1 The scale evolution of the bottom quark $\overline{MS}$ mass. The measurements include the Particle Data Group world average for $m_b(m_b)$ determined at a typical scale of the bottomonium mass, the measurements of $m_b(m_Z)$ from jet rates at the Z-pole at LEP and SLC and the measurement of $m_b(m_H)$ from Higgs boson branching fractions. The prediction of the evolution of the mass is calculated at five-loop precision with REvolver [320]. The inner dark grey error band includes the effect of missing higher-orders and the parametric uncertainties from $m_b(m_b)$ and α_S from the Particle Data Group averages. The outer band with a lighter shading includes additionally the effect of a ±0.004 variation of $\alpha_S(m_Z)$. 213

7.2 The $\chi 2$ of the fit of Eq. 7.14 to the measurements of $m_b(m_b)$, $m_b(m_Z)$ and $m_b(m_H)$, as a function of the reference bottom quark mass $m_b(m_b)$ and the factor x that multiplies the Renormalization Group Equation evolution to higher scale. The factor x interpolates smoothly between the "no-running" scenario (x = 0) and beyond the RGE evolution predicted within the Standard Model ($x > 1$). 216

8.1 The relative Jet Mass Scale and Resolution for several large-R jet mass definitions as a function of the jet p_{T}. Black and red markers represent results for W- and top-initiated jets. The inner error bar stands for the statistical uncertainty and the outer error bar includes the statistical and systematic uncertainty. Dashed blue lines, corresponding to ±1% for the relative jet mass scale and ±10% for the relative jet mass resolution, are drawn for reference. 221

8.2 Individual 68% and 95% Confidence Level limits on the Wilson coefficients C_i/Λ^2 in the Standard Model Effective Field Theory from the differential charge asymmetry measurements versus $m_{t\bar{t}}$. The bounds are derived from the measured charge asymmetry presented in this chapter, combining the semileptonic and dileptonic channels. The theory uncertainty corresponds to the NNLO QCD + NLO EW calculation of Ref. [239]. The impact of dimension-six operators is parameterized at NLO accuracy in QCD using the SMEFTNLO package in MG5_aMCNLO. 224

8.3 The scale evolution of the bottom quark $\overline{MS}$ mass. The measurements include the Particle Data Group world average for $m_b(m_b)$ determined at a typical scale of the bottomonium mass, the measurements of $m_b(m_Z)$ from jet rates at the Z-pole at LEP and SLC and the measurement of $m_b(m_H)$ from Higgs boson branching fractions. The prediction of the evolution of the mass is calculated at five-loop precision with REvolver [320]. The inner dark grey error band includes the effect of missing higher orders and the parametric uncertainties from $m_b(m_b)$ and α_s from the PDG averages. The outer band with a lighter shading includes additionally the effect of a ± 0.004 variation of $\alpha_s(m_Z)$. 226

A.1 $\chi 2$ distribution from which limits on $C_{ut}^{(8)}$ operator are derived when considering (upper-left panel) only SM-EFT terms and (upper-right panel) SM-EFT and EFT-EFT terms in the EFT contribution. Bottom panel shows the comparison between ATLAS and predicted $A_{\mathrm{C}}^{t\bar{t}}$ values, with the 68% and 95% Confidence Level bands drawn in green and yellow, respectively. $m_{t\bar{t}} \in [0, 500\}$ GeV. 230

A.2 $\chi 2$ distribution from which limits on $C_{ut}^{(8)}$ operator are derived when considering (upper-left panel) only SM-EFT terms and (upper-right panel) SM-EFT and EFT-EFT terms in the EFT contribution. Bottom panel shows the comparison between ATLAS and predicted $A_{\mathrm{C}}^{t\bar{t}}$ values, with the 68% and 95% Confidence Level bands drawn in green and yellow, respectively. $m_{t\bar{t}} \in [500, 750\}$ GeV. 231

A.3 $\chi 2$ distribution from which limits on $C_{ut}^{(8)}$ operator are derived when considering (upper-left panel) only SM-EFT terms and (upper-right panel) SM-EFT and EFT-EFT terms in the EFT contribution. Bottom panel shows the comparison between ATLAS and predicted $A_{\mathrm{C}}^{t\bar{t}}$ values, with the 68% and 95% Confidence Level bands drawn in green and yellow, respectively. $m_{t\bar{t}} \in [750, 1000\}$ GeV. 232

A.4 $\chi 2$ distribution from which limits on $C_{ut}^{(8)}$ operator are derived when considering (upper-left panel) only SM-EFT terms and (upper-right panel) SM-EFT and EFT-EFT terms in the EFT contribution. Bottom panel shows the comparison between ATLAS and predicted $A_{\mathrm{C}}^{t\bar{t}}$ values, with the 68% and 95% Confidence Level bands drawn in green and yellow, respectively. $m_{t\bar{t}} \in [1000, 1500\}$ GeV. 233

A.5 $\chi 2$ distribution from which limits on $C_{ut}^{(8)}$ operator are derived when considering (upper-left panel) only SM-EFT terms and (upper-right panel) SM-EFT and EFT-EFT terms in the EFT contribution. Bottom panel shows the comparison between ATLAS and predicted $A_{\mathrm{C}}^{t\bar{t}}$ values, with the 68% and 95% Confidence Level bands drawn in green and yellow, respectively. $m_{t\bar{t}} \in [1500, \infty\}$ GeV. 234

A.6 Comparison of the individual 68% and 95% Confidence Level expected and measured limits on the Wilson coefficients C_i/Λ^2 in the Standard Model Effective Field Theory from the differential charge asymmetry measurements versus $m_{t\bar{t}}$. Only fits including Λ^{-2} terms in the EFT parametrization are shown. 235

A.7 Comparison of the individual 68% and 95% Confidence Level expected and measured limits on the Wilson coefficients C_i/Λ^2 in the Standard Model Effective Field Theory from the differential charge asymmetry measurements versus $m_{t\bar{t}}$. Only fits including $\Lambda^{-2} + \Lambda^{-4}$ terms in the EFT parametrization are shown. 236

8.8 Partícules fonamentals del Model Estàndar. 238

8.9 Il·lustració de l'estructura interna del protó, on es pot observar els tres quarks de valència (dos *up* i un *down*) i el núvol o mar de quarks i gluons en constant interacció. La grandària de les partícules es representativa de la fracció de l'energia total del protó que tenen. 240

8.10 Localització del gran col·lisionador d'hadrons, LHC. Es mostren els quatre detectors experimentals principals a cadascun dels punts d'intersecció del feix de protons. 243

8.11 Al panell superior, el detector ATLAS i els seus subsistemes [136]. Al panell inferior, il·lustració del dipòsit d'energia de diferents partícules en funció del detector que travessen. Les línies discontínues representen partícules que no deixen energia, és a dir, que no interaccionen. 245

8.12 Representació esquemàtica de la desintegració d'un quark *top* en dos règims diferents. A l'esquerra, el quark top és produït amb una energia propera a la seua massa, per la qual cosa els seus productes de desintegració s'emeten *back-to-back*. A la dreta, l'energia total és molt més gran que la seua massa, per això s'imprimeix un *boost* o empenta al sistema que fa que els productes de desintegració es concentren en una zona reduïda de l'espai. 247

8.13 Distribució de la variable rapidity y associada a *tops* i *antitops* en producció $t\bar{t}$ a l'experiment ATLAS. Els tops tendeixen a produir-se en regions amb més valor absolut de i, mentre que els *antitops* poblen regions caracteritzades per valors de y propers a 0. 249

8.14 Límits individuals obtinguts amb un 68% i 95% de nivell de confiança per als diferents coeficients de Wilson C_i/Λ^2 considerats en aquest treball. Aquests límits es deriven a partir de les mesures diferencials a $m_{t\bar{t}}$ de l'asimetria de càrrega, comparant-les amb les prediccions del SM a NNLO QCD + NLO EW [239]. La contribució dels operadors de Wilson es calculen a NLO usant el paquet SMEFTNLO al generador de Monte Carlo MadGraph5_aMC@NLO [72]. 250

8.15 Distribució de *top* jets per a la mostra nominal Powheg +Pythia 8 (línia negra) i la predicció NLL de Ref. [210] que millor descriu les dades. El càlcul teòric es realitza a l'esquema de massa MSR (corba rosa puntejada) i a l'esquema de la massa pol (corba porpra contínua). La comparació es fa en el rang 172.5 GeV $< M_J <$ 180 GeV. 255

8.16 Comparació de successos $t\bar{t}$ generats amb POWHEG i processats amb PYTHIA 8 o HERWIG 7. La ràtio entre les dues distribucions es mostra al panell inferior. 257

8.17 Evolució de la massa del quark *bottom* a l'esquema de renormalització $\overline{MS}$ amb l'escala d'energia. Les mesures inclouen la mitjana mundial de la massa a l'escala de la massa del bottomonium $m_b(m_b)$ obtingudes en factories B, la mesura a l'escala de la massa del bosó Z $m_b(m_Z)$ a partir de *jet rates* de LEP i SLC i, finalment, la mesura $m_b(m_H)$ a partir desintegracions del bosó d'Higgs. La predicció de l'evolució segons el SM s'obté amb el programa REvolver [320]. Les bandes representen la incertesa associada a aquesta predicció. 260

List of Tables

1.1 Summary of the masses and electric charges Q_f of fermions and bosons in the Standard Model. For neutrinos masses, photon mass and electric charge, upper limits are given. For the top quark mass, the world average from cross-section measurements is reported. All values are extracted from the Particle Data Group [26]. 7

1.2 $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$ quantum numbers for left- and right-handed charged-leptons (l^-), neutrinos (ν) and up- and down-type quarks (u, d), denoted with L and R subscripts respectively, of any of the three fermion families. Leptons have 0 color number N_c, since they do not feel the strong interaction. For quarks N_c represents the number of possible color states (or the dimension of the $SU(3)$ representation). Right-handed fermions have T_3 = 0 (Q_f = Y), because they do not interact with the $SU(2)_L$ gauge bosons. Left-handed fermions with same Y and opposite T_3 form $SU(2)_L$ doublets. 19

4.1 Summary of the event selection for $t\bar{t}$ events in the μ+jets channel. Specific selections enhancing jets originating from W bosons or top quarks are also described. Table taken from Ref. [197]. 113

4.2 Best-fit $\chi 2$ values associated to the MC-to-data comparisons. . . . 119

5.1 List of Monte Carlo samples obtained by internal reweighting of the nominal. f_R and f_F represent the multiplicative factors that change the default values of the renormalization and factorization scales in the Matrix Element calculation, $\mu_{R(F)} = f_{R(F)} \cdot \mu^0_{R(F)}$. For the initial- and final-state radiation samples, they are relative to the renormalization and factorization scales that regulate the QCD emissions in the corresponding subprocess. 133

5.2 List of dedicated Monte Carlo samples with the parameters of interest. 136

5.3 Uncertainties on the relation between the top quark mass parameter in Monte Carlo generators and the MSR mass at a scale of 1 GeV. 157

5.4 The results of Monte Carlo based template-fits of the nominal jet mass distribution with a floating mass parameter to the A14 eigentune variations and alternative samples generated with Pythia 8 standalone. 176

6.1 Results with statistical, systematic and the total uncertainties for the inclusive and $m_{t\bar{t}}$ differential $A_{\mathrm{C}}^{t\bar{t}}$ measurements. The statistical uncertainty is obtained by unfolding the data, but excluding all the nuisance parameters. The systematic uncertainty is evaluated by subtracting in quadrature the statistical uncertainty from the total uncertainty. The Standard Model predictions are calculated at NNLO in QCD and NLO in EW theory. 190

6.2 Individual 68% and 95% Confidence Level bounds on the relevant Wilson coefficients of the Standard Model Effective Field Theory in units of TeV^{-2}. The bounds are derived from the charge asymmetry measurement presented in this chapter, combining the semileptonic and dileptonic channels. The upper block corresponds to the bounds derived from the inclusive measurement, the second block to the bounds from the differential charge asymmetry measurements versus $m_{t\bar{t}}$. In both cases, the experimental uncertainties are accounted for, in the form of the complete covariance matrix that keeps track of correlations between bins for the differential measurement. The theory uncertainty from the NNLO QCD + NLO EW calculation is included by explicitly varying the renormalization and factorization scales, or the parton density functions, in the calculation and registering the variations in the intervals. 202

7.1 Measurements of the bottom quark $\overline{MS}$ mass at the renormalization scale $\mu = m_Z$, from three-jet rates with bottom quarks in e^+e^- collisions at the Z-pole at LEP and SLD. For ALEPH and DELPHI the hadronization uncertainty is added in quadrature with the experimental uncertainty to yield the total systematic uncertainty. 206

Bibliography

[1] J. J. T. M. F.R.S., "Xl. cathode rays," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 44, no. 269, pp. 293–316, 1897. DOI: 10.1080/14786449708621070. eprint: https://doi.org/10.1080/14786449708621070. [Online]. Available: https://doi.org/10.1080/14786449708621070.

[2] S. H. Neddermeyer and C. D. Anderson, "Note on the nature of cosmic-ray particles," *Phys. Rev.*, vol. 51, pp. 884–886, 10 1937. DOI: 10.1103/PhysRev.51.884. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.51.884.

[3] F. Reines and C. Cowan, "The neutrino," *Nature*, vol. 178, 4531 1956. DOI: 10.1038/178446a0. [Online]. Available: https://doi.org/10.1038/178446a0.

[4] G. Danby *et al.*, "Observation of high-energy neutrino reactions and the existence of two kinds of neutrinos," *Phys. Rev. Lett.*, vol. 9, pp. 36–44, 1 1962. DOI: 10.1103/PhysRevLett.9.36. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.9.36.

[5] M. L. Perl *et al.*, "Evidence for anomalous lepton production in $e^+ - e^-$ annihilation," *Phys. Rev. Lett.*, vol. 35, pp. 1489–1492, 22 1975. DOI: 10.

1103/PhysRevLett.35.1489. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.35.1489.

[6] K. Kodama *et al.*, "Observation of tau neutrino interactions," *Physics Letters B*, vol. 504, no. 3, 218–224, 2001, ISSN: 0370-2693. DOI: 10.1016/s0370-2693(01)00307-0. [Online]. Available: http://dx.doi.org/10.1016/S0370-2693(01)00307-0.

[7] E. D. Bloom *et al.*, "High-energy inelastic $e-p$ scattering at 6° and 10°," *Phys. Rev. Lett.*, vol. 23, pp. 930–934, 16 1969. DOI: 10.1103/PhysRevLett.23.930. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.23.930.

[8] M. Breidenbach *et al.*, "Observed behavior of highly inelastic electron-proton scattering," *Phys. Rev. Lett.*, vol. 23, pp. 935–939, 16 1969. DOI: 10.1103/PhysRevLett.23.935. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.23.935.

[9] J. I. Friedman and H. W. Kendall, "Deep inelastic electron scattering," *Annual Review of Nuclear Science*, vol. 22, no. 1, pp. 203–254, 1972. DOI: 10.1146/annurev.ns.22.120172.001223. [Online]. Available: https://doi.org/10.1146/annurev.ns.22.120172.001223.

[10] G. D. Rochester and C. C. Butlerd, "Evidence for the existence of new unstable elementary particles," *Nature*, vol. 160, 4077 1947. DOI: 10.1038/160855a0. [Online]. Available: https://doi.org/10.1038/160855a0.

[11] J. E. Augustin *et al.*, "Discovery of a narrow resonance in e^+e^- annihilation," *Phys. Rev. Lett.*, vol. 33, pp. 1406–1408, 23 1974. DOI: 10.1103/PhysRevLett.33.1406. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.33.1406.

[12] J. J. Aubert *et al.*, "Experimental observation of a heavy particle J," *Phys. Rev. Lett.*, vol. 33, pp. 1404–1406, 23 1974. DOI: 10.1103/PhysRevLett.33.1404. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.33.1404.

[13] M. Kobayashi and T. Maskawa, "CP-Violation in the Renormalizable Theory of Weak Interaction," *Progress of Theoretical Physics*, vol. 49, no. 2, pp. 652–657, 1973, ISSN: 0033-068X. DOI: 10.1143/PTP.49.652. [Online]. Available: https://doi.org/10.1143/PTP.49.652.

[14] S. W. Herb *et al.*, "Observation of a dimuon resonance at 9.5 gev in 400-gev proton-nucleus collisions," *Phys. Rev. Lett.*, vol. 39, pp. 252–255, 5 1977. DOI: 10.1103/PhysRevLett.39.252. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.39.252.

[15] D0 Collaboration, "Observation of the top quark," *Phys. Rev. Lett.*, vol. 74, pp. 2632–2637, 1995. DOI: 10.1103/PhysRevLett.74.2632. arXiv: hep-ex/9503003.

[16] CDF Collaboration, "Observation of top quark production in $\bar{p}p$ collisions," *Phys. Rev. Lett.*, vol. 74, pp. 2626–2631, 1995. DOI: 10.1103/PhysRevLett.74.2626. arXiv: hep-ex/9503002.

[17] A. H. Compton, "A quantum theory of the scattering of x-rays by light elements," *Phys. Rev.*, vol. 21, pp. 483–502, 5 1923. DOI: 10.1103/PhysRev.21.483. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.21.483.

[18] C. Berger *et al.*, "Jet analysis of the $\Upsilon(9.46)$ decay into charged hadrons," *Physics Letters B*, vol. 82, no. 3, pp. 449–455, 1979, ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(79)90265-X. [Online].

Available: https://www.sciencedirect.com/science/article/pii/037026937990265X.

[19] C. Berger *et al.*, "Topology of the Υ-decay," *Zeitschrift für Physik C Particles and Fields*, vol. 8, pp. 101–114, 1980.

[20] P. Bagnaia *et al.*, "Evidence for $Z^0 \to e^+e^-$ at the cern pp collider," *Physics Letters B*, vol. 129, no. 1, pp. 130–140, 1983, ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(83)90744-X. [Online]. Available: https://www.sciencedirect.com/science/article/pii/037026938390744X.

[21] M. Banner *et al.*, "Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the cern pp collider," *Physics Letters B*, vol. 122, no. 5, pp. 476–485, 1983, ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(83)91605-2. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0370269383916052.

[22] G. Arnison *et al.*, "Experimental observation of isolated large transverse energy electrons with associated missing energy at s=540 gev," *Physics Letters B*, vol. 122, no. 1, pp. 103–116, 1983, ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(83)91177-2. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0370269383911772.

[23] G. Arnison *et al.*, "Experimental observation of lepton pairs of invariant mass around 95 gev/c2 at the cern sps collider," *Physics Letters B*, vol. 126, no. 5, pp. 398–410, 1983, ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(83)90188-0. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0370269383901880.

[24] ATLAS Collaboration, "Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc," *Physics Letters B*, vol. 716, no. 1, 1–29, 2012, ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.08.020. [Online]. Available: http://dx.doi.org/10.1016/j.physletb.2012.08.020.

[25] CMS Collaboration, "Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc," *Physics Letters B*, vol. 716, no. 1, 30–61, 2012, ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.08.021. [Online]. Available: http://dx.doi.org/10.1016/j.physletb.2012.08.021.

[26] Particle Data Group, "Review of Particle Physics," *PTEP*, vol. 2020, no. 8, p. 083C01, 2020. DOI: 10.1093/ptep/ptaa104.

[27] S. J. Brodsky, "Review of 'Quantum Electrodynamics' by T. Kinoshita, World Scientific," Jun. 1992.

[28] G. Gabrielse, D. Hanneke, T. Kinoshita, M. Nio, and B. Odom, "Erratum: New determination of the fine structure constant from the electron g value and qed [phys. rev. lett. 97, 030802 (2006)]," *Phys. Rev. Lett.*, vol. 99, p. 039 902, 3 2007. DOI: 10.1103/PhysRevLett.99.039902. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.99.039902.

[29] P. J. Mohr, B. N. Taylor, and D. B. Newell, "Codata recommended values of the fundamental physical constants: 2010," *Rev. Mod. Phys.*, vol. 84, pp. 1527–1605, 4 2012. DOI: 10.1103/RevModPhys.84.1527. [Online]. Available: https://link.aps.org/doi/10.1103/RevModPhys.84.1527.

[30] T. D. Lee and C. N. Yang, "Question of parity conservation in weak interactions," *Phys. Rev.*, vol. 104, pp. 254–258, 1 1956. DOI: 10.1103/

PhysRev.104.254. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.104.254.

[31] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, "Experimental test of parity conservation in beta decay," *Phys. Rev.*, vol. 105, pp. 1413–1415, 4 1957. DOI: 10.1103/PhysRev.105.1413. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.105.1413.

[32] R. L. Garwin, L. M. Lederman, and M. Weinrich, "Observations of the failure of conservation of parity and charge conjugation in meson decays: The magnetic moment of the free muon," *Phys. Rev.*, vol. 105, pp. 1415–1417, 4 1957. DOI: 10.1103/PhysRev.105.1415. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.105.1415.

[33] J. I. Friedman and V. L. Telegdi, "Nuclear emulsion evidence for parity nonconservation in the decay chain $\pi^+ - \mu^+ - e^+$," *Phys. Rev.*, vol. 105, pp. 1681–1682, 5 1957. DOI: 10.1103/PhysRev.105.1681.2. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.105.1681.2.

[34] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, "Evidence for the 2π decay of the K_2^0 meson," *Phys. Rev. Lett.*, vol. 13, pp. 138–140, 4 1964. DOI: 10.1103/PhysRevLett.13.138. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.13.138.

[35] E. C. G. Sudarshan and R. E. Marshak, "Chirality invariance and the universal fermi interaction," *Phys. Rev.*, vol. 109, pp. 1860–1862, 5 1958. DOI: 10.1103/PhysRev.109.1860.2. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.109.1860.2.

[36] S. Weinberg, "A model of leptons," *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, 21 1967. DOI: 10.1103/PhysRevLett.19.1264. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.19.1264.

[37] A. Salam, "Weak and Electromagnetic Interactions," *Conf. Proc. C*, vol. 680519, pp. 367–377, 1968. DOI: 10.1142/9789812795915_0034.

[38] Y. Nambu, "Quasi-particles and gauge invariance in the theory of superconductivity," *Phys. Rev.*, vol. 117, pp. 648–663, 3 1960. DOI: 10.1103/PhysRev.117.648. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.117.648.

[39] J. Goldstone, A. Salam, and S. Weinberg, "Broken symmetries," *Phys. Rev.*, vol. 127, pp. 965–970, 3 1962. DOI: 10.1103/PhysRev.127.965. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.127.965.

[40] F. Englert and R. Brout, "Broken symmetry and the mass of gauge vector mesons," *Phys. Rev. Lett.*, vol. 13, pp. 321–323, 9 1964. DOI: 10.1103/PhysRevLett.13.321. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.13.321.

[41] P. W. Higgs, "Spontaneous symmetry breakdown without massless bosons," *Phys. Rev.*, vol. 145, pp. 1156–1163, 4 1966. DOI: 10.1103/PhysRev.145.1156. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.145.1156.

[42] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global conservation laws and massless particles," *Phys. Rev. Lett.*, vol. 13, pp. 585–587, 20 1964. DOI: 10.1103/PhysRevLett.13.585. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.13.585.

[43] T. W. B. Kibble, "Symmetry breaking in non-abelian gauge theories," *Phys. Rev.*, vol. 155, pp. 1554–1561, 5 1967. DOI: 10.1103/PhysRev.155.1554. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.155.1554.

[44] M. Kobayashi and T. Maskawa, "CP Violation in the Renormalizable Theory of Weak Interaction," *Prog. Theor. Phys.*, vol. 49, pp. 652–657, 1973. DOI: 10.1143/PTP.49.652.

[45] N. Cabibbo, "Unitary symmetry and leptonic decays," *Phys. Rev. Lett.*, vol. 10, pp. 531–533, 12 1963. DOI: 10.1103/PhysRevLett.10.531. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.10.531.

[46] M. Gell-Mann, "A Schematic Model of Baryons and Mesons," *Phys. Lett.*, vol. 8, pp. 214–215, 1964. DOI: 10.1016/S0031-9163(64)92001-3.

[47] G. Zweig, "An SU(3) model for strong interaction symmetry and its breaking. Version 1," Jan. 1964.

[48] J. D. Bjorken, "Asymptotic sum rules at infinite momentum," *Phys. Rev.*, vol. 179, pp. 1547–1553, 5 1969. DOI: 10.1103/PhysRev.179.1547. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.179.1547.

[49] R. P. Feynman, "Very high-energy collisions of hadrons," *Phys. Rev. Lett.*, vol. 23, pp. 1415–1417, 24 1969. DOI: 10.1103/PhysRevLett.23.1415. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.23.1415.

[50] A. H. Hoang, "What is the top quark mass?" *Annual Review of Nuclear and Particle Science*, vol. 70, no. 1, 225–255, 2020, ISSN: 1545-4134. DOI:

10.1146/annurev-nucl-101918-023530. [Online]. Available: http://dx.doi.org/10.1146/annurev-nucl-101918-023530.

[51] K. Chetyrkin, "Four-loop renormalization of qcd: Full set of renormalization constants and anomalous dimensions," *Nuclear Physics B*, vol. 710, no. 1-2, 499–510, 2005, ISSN: 0550-3213. DOI: 10.1016/j.nuclphysb.2005.01.011. [Online]. Available: http://dx.doi.org/10.1016/j.nuclphysb.2005.01.011.

[52] M. Czakon, "The four-loop qcd β-function and anomalous dimensions," *Nuclear Physics B*, vol. 710, no. 1-2, 485–498, 2005, ISSN: 0550-3213. DOI: 10.1016/j.nuclphysb.2005.01.012. [Online]. Available: http://dx.doi.org/10.1016/j.nuclphysb.2005.01.012.

[53] A. H. Hoang *et al.*, "The MSR mass and the $\mathcal{O}(\Lambda_{\rm QCD})$ renormalon sum rule," *JHEP*, vol. 04, p. 003, 2018. DOI: 10.1007/JHEP04(2018)003. arXiv: 1704.01580 [hep-ph].

[54] A. H. Hoang, "What is the Top Quark Mass?" *Ann. Rev. Nucl. Part. Sci.*, vol. 70, pp. 225–255, 2020. DOI: 10.1146/annurev-nucl-101918-023530. arXiv: 2004.12915 [hep-ph].

[55] M. Beneke, "Renormalons," *Physics Reports*, vol. 317, no. 1-2, 1–142, 1999, ISSN: 0370-1573. DOI: 10.1016/s0370-1573(98)00130-6. [Online]. Available: http://dx.doi.org/10.1016/S0370-1573(98)00130-6.

[56] D. J. Gross and A. Neveu, "Dynamical symmetry breaking in asymptotically free field theories," *Phys. Rev. D*, vol. 10, pp. 3235–3253, 10 1974. DOI: 10.1103/PhysRevD.10.3235. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.10.3235.

[57] M. Beneke and V. A. Smirnov, "Ultraviolet renormalons in Abelian gauge theories," *Nucl. Phys. B*, vol. 472, pp. 529–590, 1996. DOI: 10.1016/0550-3213(96)00205-2. arXiv: hep-ph/9510437.

[58] T. Kinoshita, "Mass singularities of Feynman amplitudes," *J. Math. Phys.*, vol. 3, pp. 650–677, 1962. DOI: 10.1063/1.1724268.

[59] T. D. Lee and M. Nauenberg, "Degenerate systems and mass singularities," *Phys. Rev.*, vol. 133, B1549–B1562, 6B 1964. DOI: 10.1103/PhysRev.133.B1549. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.133.B1549.

[60] D. J. Gross and F. Wilczek, "Asymptotically free gauge theories. i," *Phys. Rev. D*, vol. 8, pp. 3633–3652, 10 1973. DOI: 10.1103/PhysRevD.8.3633. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.8.3633.

[61] A. J. Larkoski, I. Moult, and B. Nachman, "Jet substructure at the large hadron collider: A review of recent advances in theory and machine learning," *Physics Reports*, vol. 841, 1–63, 2020, ISSN: 0370-1573. DOI: 10.1016/j.physrep.2019.11.001. [Online]. Available: http://dx.doi.org/10.1016/j.physrep.2019.11.001.

[62] J. C. Collins, D. E. Soper, and G. F. Sterman, "Factorization of Hard Processes in QCD," *Adv. Ser. Direct. High Energy Phys.*, vol. 5, pp. 1–91, 1989. DOI: 10.1142/9789814503266_0001. arXiv: hep-ph/0409313.

[63] T. Sjostrand, "Monte Carlo Generators," pp. 51–74, Nov. 2006. arXiv: hep-ph/0611247.

[64] C. W. Bauer, S. Fleming, D. Pirjol, I. Z. Rothstein, and I. W. Stewart, "Hard scattering factorization from effective field theory," *Physical Review D*, vol. 66, no. 1, 2002, ISSN: 1089-4918. DOI: 10.1103/physrevd.66.

014017. [Online]. Available: http://dx.doi.org/10.1103/PhysRevD.66.014017.

[65] Y. L. Dokshitzer, "Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.," *Sov. Phys. JETP*, vol. 46, pp. 641–653, 1977.

[66] G. Altarelli and G. Parisi, "Asymptotic Freedom in Parton Language," *Nucl. Phys. B*, vol. 126, pp. 298–318, 1977. DOI: 10.1016/0550-3213(77)90384-4.

[67] F. D. Aaron *et al.*, "Combined measurement and qcd analysis of the inclusive e ± p scattering cross sections at hera," *Journal of High Energy Physics*, vol. 2010, no. 1, 2010, ISSN: 1029-8479. DOI: 10.1007/jhep01(2010)109. [Online]. Available: http://dx.doi.org/10.1007/JHEP01(2010)109.

[68] L. Del Debbio, "Parton distributions in the LHC era," *EPJ Web Conf.*, vol. 175, M. Della Morte, P. Fritzsch, E. Gámiz Sánchez, and C. Pena Ruano, Eds., p. 01 006, 2018. DOI: 10.1051/epjconf/201817501006.

[69] F. Maltoni, T. McElmurry, R. Putman, and S. Willenbrock, "Choosing the Factorization Scale in Perturbative QCD," Mar. 2007. arXiv: hep-ph/0703156.

[70] S. Catani, "Aspects of QCD, from the Tevatron to the LHC," in *1st Les Houches Workshop on Physics at TeV Colliders*, May 2000, pp. 4–33. arXiv: hep-ph/0005233.

[71] S. Frixione and B. R. Webber, "Matching nlo qcd computations and parton shower simulations," *Journal of High Energy Physics*, vol. 2002, no. 06, 029–029, 2002, ISSN: 1029-8479. DOI: 10.1088/1126-6708/2002/06/029.

[Online]. Available: http://dx.doi.org/10.1088/1126-6708/2002/06/029.

[72] 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*, vol. 07, p. 079, 2014. DOI: 10.1007/JHEP07(2014)079. arXiv: 1405.0301 [hep-ph].

[73] P. Nason, "A new method for combining nlo qcd with shower monte carlo algorithms," *Journal of High Energy Physics*, vol. 2004, no. 11, 040–040, 2004, ISSN: 1029-8479. DOI: 10.1088/1126-6708/2004/11/040. [Online]. Available: http://dx.doi.org/10.1088/1126-6708/2004/11/040.

[74] S. Frixione, P. Nason, and C. Oleari, "Matching nlo qcd computations with parton shower simulations: The powheg method," *Journal of High Energy Physics*, vol. 2007, no. 11, pp. 070–070, 2007. DOI: 10.1088/1126-6708/2007/11/070.

[75] A. Buckley *et al.*, "General-purpose event generators for lhc physics," *Physics Reports*, vol. 504, no. 5, 145–233, 2011, ISSN: 0370-1573. DOI: 10.1016/j.physrep.2011.03.005. [Online]. Available: http://dx.doi.org/10.1016/j.physrep.2011.03.005.

[76] V. V. Sudakov, "Vertex parts at very high-energies in quantum electrodynamics," *Sov. Phys. JETP*, vol. 3, pp. 65–71, 1956.

[77] J. C. Collins, "Sudakov form-factors," *Adv. Ser. Direct. High Energy Phys.*, vol. 5, pp. 573–614, 1989. DOI: 10.1142/9789814503266_0006. arXiv: hep-ph/0312336.

[78] T. Sjöstrand *et al.*, "An introduction to pythia 8.2," *Computer Physics Communications*, vol. 191, 159–177, 2015, ISSN: 0010-4655. DOI: 10.1016/

j.cpc.2015.01.024. [Online]. Available: http://dx.doi.org/10.1016/j.cpc.2015.01.024.

[79] T. Sjöstrand and P. Z. Skands, "Transverse-momentum-ordered showers and interleaved multiple interactions," *The European Physical Journal C*, vol. 39, no. 2, 129–154, 2005, ISSN: 1434-6052. DOI: 10.1140/epjc/s2004-02084-y. [Online]. Available: http://dx.doi.org/10.1140/epjc/s2004-02084-y.

[80] M. Bähr *et al.*, "Herwig++ physics and manual," *The European Physical Journal C*, vol. 58, no. 4, 639–707, 2008, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-008-0798-9. [Online]. Available: http://dx.doi.org/10.1140/epjc/s10052-008-0798-9.

[81] S. Gieseke, P. Stephens, and B. Webber, "New formalism for qcd parton showers," *Journal of High Energy Physics*, vol. 2003, no. 12, 045–045, 2003, ISSN: 1029-8479. DOI: 10.1088/1126-6708/2003/12/045. [Online]. Available: http://dx.doi.org/10.1088/1126-6708/2003/12/045.

[82] S. Höche, "Introduction to parton-shower event generators," 2015. arXiv: 1411.4085 [hep-ph].

[83] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, "Parton Fragmentation and String Dynamics," *Phys. Rept.*, vol. 97, pp. 31–145, 1983. DOI: 10.1016/0370-1573(83)90080-7.

[84] S. Ferreres-Solé and T. Sjöstrand, "The space–time structure of hadronization in the lund model," *The European Physical Journal C*, vol. 78, no. 11, 2018, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-018-6459-8. [Online]. Available: http://dx.doi.org/10.1140/epjc/s10052-018-6459-8.

[85] D. J. Lange, "The EvtGen particle decay simulation package," *Nucl. Instrum. Meth. A*, vol. 462, S. Erhan, P. Schlein, and Y. Rozen, Eds., pp. 152–155, 2001. DOI: 10.1016/S0168-9002(01)00089-4.

[86] S. Argyropoulos and T. Sjöstrand, "Effects of color reconnection on $t\bar{t}$ final states at the lhc," *Journal of High Energy Physics*, vol. 2014, no. 11, 2014, ISSN: 1029-8479. DOI: 10.1007/jhep11(2014)043. [Online]. Available: http://dx.doi.org/10.1007/JHEP11(2014)043.

[87] J. R. Christiansen and P. Z. Skands, "String formation beyond leading colour," *Journal of High Energy Physics*, vol. 2015, no. 8, 2015, ISSN: 1029-8479. DOI: 10.1007/jhep08(2015)003. [Online]. Available: http://dx.doi.org/10.1007/JHEP08(2015)003.

[88] A. V. Manohar, *Introduction to effective field theories*, 2018. arXiv: 1804.05863 [hep-ph].

[89] T. E. W. Group, *Combination of cdf and d0 results on the mass of the top quark using up* 9.7fb^{-1} *at the tevatron*, 2016. arXiv: 1608.01881 [hep-ex].

[90] M. Baak *et al.*, "The global electroweak fit at nnlo and prospects for the lhc and ilc," *The European Physical Journal C*, vol. 74, no. 9, 2014, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-014-3046-5. [Online]. Available: http://dx.doi.org/10.1140/epjc/s10052-014-3046-5.

[91] G. Degrassi *et al.*, "Higgs mass and vacuum stability in the standard model at nnlo," *Journal of High Energy Physics*, vol. 2012, no. 8, 2012, ISSN: 1029-8479. DOI: 10.1007/jhep08(2012)098. [Online]. Available: http://dx.doi.org/10.1007/JHEP08(2012)098.

[92] M. Czakon, P. Fiedler, and A. Mitov, "Total top-quark pair-production cross section at hadron colliders througho(αs4)," *Physical Review Letters*, vol. 110, no. 25, 2013, ISSN: 1079-7114. DOI: 10.1103/physrevlett.

110.252004. [Online]. Available: http://dx.doi.org/10.1103/PhysRevLett.110.252004.

[93] M. Guzzi, K. Lipka, and S.-O. Moch, "Top-quark pair production at hadron colliders: Differential cross section and phenomenological applications with difftop," *Journal of High Energy Physics*, vol. 2015, no. 1, 2015, ISSN: 1029-8479. DOI: 10.1007/jhep01(2015)082. [Online]. Available: http://dx.doi.org/10.1007/JHEP01(2015)082.

[94] J. Kühn, A. Scharf, and P. Uwer, "Weak interaction effects in top-quark pair production at hadron colliders," *The European Physical Journal C*, vol. 51, no. 1, 37–53, 2007, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-007-0275-x. [Online]. Available: http://dx.doi.org/10.1140/epjc/s10052-007-0275-x.

[95] DØ Collaboration, "Measurement of the forward-backward asymmetry in top quark-antiquark production in $p\bar{p}$ collisions using the lepton+jets channel," *Phys. Rev. D*, vol. 90, p. 072011, 7 2014. DOI: 10.1103/PhysRevD.90.072011. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.90.072011.

[96] CDF Collaboration, "Measurement of the top quark forward-backward production asymmetry and its dependence on event kinematic properties," *Phys. Rev. D*, vol. 87, p. 092002, 9 2013. DOI: 10.1103/PhysRevD.87.092002. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.87.092002.

[97] J. Aguilar-Saavedra, D. Amidei, A. Juste, and M. Pérez-Victoria, "Asymmetries in top quark pair production at hadron colliders," *Reviews of Modern Physics*, vol. 87, no. 2, 421–455, 2015, ISSN: 1539-0756. DOI:

10.1103/revmodphys.87.421. [Online]. Available: http://dx.doi.org/10.1103/RevModPhys.87.421.

[98] M. Czakon, P. Fiedler, and A. Mitov, "Resolving the tevatron top quark forward-backward asymmetry puzzle: Fully differential next-to-next-to-leading-order calculation," *Physical Review Letters*, vol. 115, no. 5, 2015, ISSN: 1079-7114. DOI: 10.1103/physrevlett.115.052001. [Online]. Available: http://dx.doi.org/10.1103/PhysRevLett.115.052001.

[99] J. H. Kuehn and G. Rodrigo, *Forward-backward and charge asymmetries at tevatron and the lhc*, 2014. arXiv: 1411.4675.

[100] W. Bernreuther and Z.-G. Si, "Top quark and leptonic charge asymmetries for the tevatron and lhc," *Physical Review D*, vol. 86, no. 3, 2012, ISSN: 1550-2368. DOI: 10.1103/physrevd.86.034026. [Online]. Available: http://dx.doi.org/10.1103/PhysRevD.86.034026.

[101] ATLAS Collaboration, "Measurement of the charge asymmetry in top-quark pair production in the lepton-plus-jets final state in pp collision data at $\sqrt{s}$ = 8 tev with the atlas detector," *The European Physical Journal C*, vol. 76, no. 2, 2016, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-016-3910-6. [Online]. Available: http://dx.doi.org/10.1140/epjc/s10052-016-3910-6.

[102] CMS Collaboration, "Measurements of $t\bar{t}$ charge asymmetry using dilepton final states in pp collisions at $\sqrt{s}$ = 8 TeV," *Phys. Lett. B*, vol. 760, pp. 365–386, 2016. DOI: 10.1016/j.physletb.2016.07.006. arXiv: 1603.06221 [hep-ex].

[103] ATLAS Collaboration, "Measurement of the charge asymmetry in highly boosted top-quark pair production in $\sqrt{s}$ = 8 pp collision data collected by the atlas experiment," *Physics Letters B*, vol. 756, 52–71, 2016, ISSN:

0370-2693. DOI: 10.1016/j.physletb.2016.02.055. [Online]. Available: http://dx.doi.org/10.1016/j.physletb.2016.02.055.

[104] B. Grzadkowski, M. Iskrzynski, M. Misiak, and J. Rosiek, "Dimension-Six Terms in the Standard Model Lagrangian," *JHEP*, vol. 10, p. 085, 2010. DOI: 10.1007/JHEP10(2010)085. arXiv: 1008.4884 [hep-ph].

[105] C. Degrande, J.-M. Gérard, C. Grojean, F. Maltoni, and G. Servant, "Non-resonant new physics in top pair production at hadron colliders," *Journal of High Energy Physics*, vol. 2011, no. 3, 2011, ISSN: 1029-8479. DOI: 10.1007/jhep03(2011)125. [Online]. Available: http://dx.doi.org/10.1007/JHEP03(2011)125.

[106] C. Zhang and S. Willenbrock, "Effective-field-theory approach to top-quark production and decay," *Physical Review D*, vol. 83, no. 3, 2011, ISSN: 1550-2368. DOI: 10.1103/physrevd.83.034006. [Online]. Available: http://dx.doi.org/10.1103/PhysRevD.83.034006.

[107] A. Buckley *et al.*, "Constraining top quark effective theory in the lhc run ii era," *Journal of High Energy Physics*, vol. 2016, no. 4, 1–32, 2016, ISSN: 1029-8479. DOI: 10.1007/jhep04(2016)015. [Online]. Available: http://dx.doi.org/10.1007/JHEP04(2016)015.

[108] C. Degrande, F. Maltoni, K. Mimasu, E. Vryonidou, and C. Zhang, "Single-top associated production with a z or h boson at the lhc: The smeft interpretation," *Journal of High Energy Physics*, vol. 2018, no. 10, 2018, ISSN: 1029-8479. DOI: 10.1007/jhep10(2018)005. [Online]. Available: http://dx.doi.org/10.1007/JHEP10(2018)005.

[109] N. P. Hartland *et al.*, "A monte carlo global analysis of the standard model effective field theory: The top quark sector," *Journal of High Energy Physics*, vol. 2019, no. 4, 2019, ISSN: 1029-8479. DOI: 10.1007/

jhep04(2019)100. [Online]. Available: http://dx.doi.org/10.1007/JHEP04(2019)100.

[110] I. Brivio *et al.*, "O new physics, where art thou? a global search in the top sector," *Journal of High Energy Physics*, vol. 2020, no. 2, 2020, ISSN: 1029-8479. DOI: 10.1007/jhep02(2020)131. [Online]. Available: http://dx.doi.org/10.1007/JHEP02(2020)131.

[111] C. Englert, L. Moore, K. Nordström, and M. Russell, "Giving top quark effective operators a boost," *Phys. Lett. B*, vol. 763, pp. 9–15, 2016. DOI: 10.1016/j.physletb.2016.10.021. arXiv: 1607.04304 [hep-ph].

[112] C. Zhang, "Constraining qqtt operators from four-top production: A case for enhanced eft sensitivity," *Chinese Physics C*, vol. 42, no. 2, p. 023 104, 2018, ISSN: 1674-1137. DOI: 10.1088/1674-1137/42/2/023104. [Online]. Available: http://dx.doi.org/10.1088/1674-1137/42/2/023104.

[113] A. Buckley *et al.*, "Global fit of top quark effective theory to data," *Physical Review D*, vol. 92, no. 9, 2015, ISSN: 1550-2368. DOI: 10.1103/physrevd.92.091501. [Online]. Available: http://dx.doi.org/10.1103/PhysRevD.92.091501.

[114] C. Zhang and S. Willenbrock, "Effective-field-theory approach to top-quark production and decay," *Physical Review D*, vol. 83, no. 3, 2011, ISSN: 1550-2368. DOI: 10.1103/physrevd.83.034006. [Online]. Available: http://dx.doi.org/10.1103/PhysRevD.83.034006.

[115] ATLAS Collaboration, "Measurement of the top quark mass in the $t\bar{t} \to$ lepton+jets channel from $\sqrt{s} = 8$ TeV ATLAS data and combination with previous results," *Eur. Phys. J. C*, vol. 79, no. 4, p. 290, 2019. DOI: 10.1140/epjc/s10052-019-6757-9. arXiv: 1810.01772 [hep-ex].

[116] L. T. W. Group, *Lhctopwg summary plots*, 2018. [Online]. Available: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCTopWGSummaryPlots.

[117] G. Heinrich *et al.*, "Nlo and off-shell effects in top quark mass determinations," *Journal of High Energy Physics*, vol. 2018, no. 7, 2018, ISSN: 1029-8479. DOI: 10.1007/jhep07(2018)129. [Online]. Available: http://dx.doi.org/10.1007/JHEP07(2018)129.

[118] S. Alioli *et al.*, "A new observable to measure the top-quark mass at hadron colliders," *The European Physical Journal C*, vol. 73, no. 5, 2013, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-013-2438-2. [Online]. Available: http://dx.doi.org/10.1140/epjc/s10052-013-2438-2.

[119] ATLAS Collaboration, "Measurement of the top-quark mass in $t\bar{t}$ + 1-jet events collected with the atlas detector in pp collisions at $\sqrt{s}$ = 8 TeV," *Journal of High Energy Physics*, vol. 2019, no. 11, 2019, ISSN: 1029-8479. DOI: 10.1007/jhep11(2019)150. [Online]. Available: http://dx.doi.org/10.1007/JHEP11(2019)150.

[120] CMS Collaboration, "Measurement of $\mathrm{t\bar{t}}$ normalised multi-differential cross sections in pp collisions at $\sqrt{s}$ = 13 TeV, and simultaneous determination of the strong coupling strength, top quark pole mass, and parton distribution functions," *The European Physical Journal C*, vol. 80, no. 7, 2020, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-020-7917-7. [Online]. Available: http://dx.doi.org/10.1140/epjc/s10052-020-7917-7.

[121] J. Kieseler, K. Lipka, and S.-O. Moch, "Calibration of the top-quark monte carlo mass," *Phys. Rev. Lett.*, vol. 116, p. 162 001, 16 2016. DOI: 10.1103/

PhysRevLett.116.162001. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.116.162001.

[122] A. H. Hoang, *The top mass: Interpretation and theoretical uncertainties*, 2014. arXiv: 1412.3649 [hep-ph].

[123] A. H. Hoang, S. Plätzer, and D. Samitz, "On the cutoff dependence of the quark mass parameter in angular ordered parton showers," *Journal of High Energy Physics*, vol. 2018, no. 10, 2018, ISSN: 1029-8479. DOI: 10.1007/jhep10(2018)200. [Online]. Available: http://dx.doi.org/10.1007/JHEP10(2018)200.

[124] M. Butenschoen, B. Dehnadi, A. H. Hoang, V. Mateu, M. Preisser, and I. W. Stewart, "Top quark mass calibration for monte carlo event generators," *Phys. Rev. Lett.*, vol. 117, p. 232 001, 23 2016. DOI: 10.1103/PhysRevLett.117.232001. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.117.232001.

[125] Gargamelle Neutrino Collaboration, "Observation of neutrino-like interactions without muon or electron in the gargamelle neutrino experiment," *Physics Letters B*, vol. 46, no. 1, pp. 138–140, 1973, ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(73)90499-1. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0370269373904991.

[126] UA1 Collaboration, "Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s}$ = 540 gev," *Physics Letters B*, vol. 122, no. 1, pp. 103–116, 1983, ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(83)91177-2. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0370269383911772.

[127] UA1 Collaboration, "Experimental observation of lepton pairs of invariant mass around 95 $GeV/2c^2$ at the cern sps collider," *Physics Letters B*, vol. 126, no. 5, pp. 398–410, 1983, ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(83)90188-0. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0370269383901880.

[128] UA2 Collaboration, "Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the cern pp collider," *Physics Letters B*, vol. 122, no. 5, pp. 476–485, 1983, ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(83)91605-2. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0370269383916052.

[129] UA2 Collaboration, "Evidence for $Z^0 \to e^+e^-$ at the cern pp collider," *Physics Letters B*, vol. 129, no. 1, pp. 130–140, 1983, ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(83)90744-X. [Online]. Available: https://www.sciencedirect.com/science/article/pii/037026938390744X.

[130] ATLAS Collaboration, "The ATLAS experiment at the CERN large hadron collider," *Journal of Instrumentation*, vol. 3, no. 08, S08003–S08003, 2008. DOI: 10.1088/1748-0221/3/08/s08003. [Online]. Available: https://doi.org/10.1088/1748-0221/3/08/s08003.

[131] CMS Collaboration, "The CMS experiment at the CERN LHC," *Journal of Instrumentation*, vol. 3, no. 08, S08004–S08004, 2008. DOI: 10.1088/1748-0221/3/08/s08004. [Online]. Available: https://doi.org/10.1088/1748-0221/3/08/s08004.

[132] ALICE Collaboration, "The ALICE experiment at the CERN LHC," *Journal of Instrumentation*, vol. 3, no. 08, S08002–S08002, 2008. DOI:

10.1088/1748-0221/3/08/s08002. [Online]. Available: https://doi.org/10.1088/1748-0221/3/08/s08002.

[133] LHCb Collaboration, "The LHCb detector at the LHC," *Journal of Instrumentation*, vol. 3, no. 08, S08005–S08005, 2008. DOI: 10.1088/1748-0221/3/08/s08005. [Online]. Available: https://doi.org/10.1088/1748-0221/3/08/s08005.

[134] ATLAS Collaboration, *Luminosity Public Results for Run 2*. [Online]. Available: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/.

[135] B. P. Nachman, "Investigating the quantum properties of jets and the search for a supersymmetric top quark partner with the atlas detector," Ph.D. dissertation, Standford, 2016.

[136] ATLAS Collaboration, "Studies of the performance of the atlas detector using cosmic-ray muons," *The European Physical Journal C*, vol. 71, no. 3, 2011, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-011-1593-6. [Online]. Available: http://dx.doi.org/10.1140/epjc/s10052-011-1593-6.

[137] ATLAS Collaboration, "Track Reconstruction Performance of the ATLAS Inner Detector at $\sqrt{s}$ = 13˜TeV," Jul. 2015.

[138] ATLAS Collaboration, "Topological cell clustering in the atlas calorimeters and its performance in lhc run 1," *The European Physical Journal C*, vol. 77, no. 7, 2017, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-017-5004-5. [Online]. Available: http://dx.doi.org/10.1140/epjc/s10052-017-5004-5.

[139] ATLAS Collaboration, "Commissioning of the atlas muon spectrometer with cosmic rays," *The European Physical Journal C*, vol. 70, no. 3, 875–916, 2010, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-010-1415-2.

[Online]. Available: http://dx.doi.org/10.1140/epjc/s10052-010-1415-2.

[140] ATLAS Collaboration, "Measurement of the inelastic proton-proton cross section at $\sqrt{s}$ = 13 TeV with the atlas detector at the lhc," *Physical Review Letters*, vol. 117, no. 18, 2016, ISSN: 1079-7114. DOI: 10.1103/physrevlett.117.182002. [Online]. Available: http://dx.doi.org/10.1103/PhysRevLett.117.182002.

[141] *World lhc computing grid*. [Online]. Available: https://wlcg.web.cern.ch.

[142] Cameron, David, Forti, Alessandra, Klimentov, Alexei, Pacheco Pages, Andres, and South, David, "Evolution of atlas analysis workflows and tools for the hl-lhc era," *EPJ Web Conf.*, vol. 251, p. 02 002, 2021. DOI: 10.1051/epjconf/202125102002. [Online]. Available: https://doi.org/10.1051/epjconf/202125102002.

[143] S. G. de la Hoz *et al.*, "Computing activities at the Spanish Tier-1 and Tier-2s for the ATLAS experiment towards the LHC Run3 and High-Luminosity periods," *EPJ Web Conf.*, vol. 245, C. Doglioni, D. Kim, G. A. Stewart, L. Silvestris, P. Jackson, and W. Kamleh, Eds., p. 07 027, 2020. DOI: 10.1051/epjconf/202024507027.

[144] ATLAS Collaboration, "The ATLAS Computing Model," Dec. 2004. [Online]. Available: https://cds.cern.ch/record/811058/files/soft-2004-007.pdf.

[145] ATLAS Collaboration, "Electron reconstruction and identification in the atlas experiment using the 2015 and 2016 lhc proton–proton collision data at $\sqrt{s}$ = 13 tev," *The European Physical Journal C*, vol. 79, no. 8, 2019,

ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-019-7140-6. [Online]. Available: http://dx.doi.org/10.1140/epjc/s10052-019-7140-6.

[146] ATLAS Collaboration, "Electron efficiency measurements with the atlas detector using 2012 lhc proton–proton collision data," *The European Physical Journal C*, vol. 77, no. 3, 2017, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-017-4756-2. [Online]. Available: http://dx.doi.org/10.1140/epjc/s10052-017-4756-2.

[147] ATLAS Collaboration, "Muon reconstruction performance of the atlas detector in proton–proton collision data at $\sqrt{s}$ = 13 tev," *The European Physical Journal C*, vol. 76, no. 5, 2016, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-016-4120-y. [Online]. Available: http://dx.doi.org/10.1140/epjc/s10052-016-4120-y.

[148] ATLAS Collaboration, "Photon identification in 2015 ATLAS data," Aug. 2016. [Online]. Available: http://cds.cern.ch/record/2203125.

[149] ATLAS Collaboration, "Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s}$ = 13 TeV," *Eur. Phys. J. C*, vol. 78, no. 11, p. 903, 2018. DOI: 10.1140/epjc/s10052-018-6288-9. arXiv: 1802.08168 [hep-ex].

[150] G. P. Salam, "Towards jetography," *The European Physical Journal C*, vol. 67, no. 3-4, 637–686, 2010, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-010-1314-6. [Online]. Available: http://dx.doi.org/10.1140/epjc/s10052-010-1314-6.

[151] ATLAS Collaboration, *Search for boosted resonances decaying to two b-quarks and produced in association with a jet at* $\sqrt{s}$ = 13 *TeV with the ATLAS detector*, ATLAS-CONF-2018-052, 2018. [Online]. Available: https://cds.cern.ch/record/2649081.

[152] ATLAS Collaboration, "Search for the standard model Higgs boson produced in association with top quarks and decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," *Phys. Rev.*, vol. D97, no. 7, p. 072 016, 2018. DOI: 10.1103/PhysRevD.97.072016. arXiv: 1712.08895 [hep-ex].

[153] ATLAS Collaboration, "Search for pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," *JHEP*, vol. 01, p. 030, 2019. DOI: 10.1007/JHEP01(2019)030. arXiv: 1804.06174 [hep-ex].

[154] CMS Collaboration, "Search for electroweak production of a vector-like quark decaying to a top quark and a Higgs boson using boosted topologies in ful ly hadronic final states," *JHEP*, vol. 04, p. 136, 2017. DOI: 10.1007/JHEP04(2017)136. arXiv: 1612.05336 [hep-ex].

[155] CMS Collaboration, "Inclusive Search for a Highly Boosted Higgs Boson Decaying to a Bottom Quark–Antiquark Pair," *Phys. Rev. Lett.*, vol. 120, p. 071 802, 2018. DOI: 10.1103/PhysRevLett.120.071802. arXiv: 1709.05543 [hep-ex].

[156] ATLAS Collaboration, "Measurement of the jet mass in high transverse momentum $Z(\rightarrow b\bar{b})\gamma$ production at $\sqrt{s} = 13$ TeV using the ATLAS detector," 2019. arXiv: 1907.07093 [hep-ex].

[157] ATLAS Collaboration, "Measurement of jet-substructure observables in top quark, W boson and light jet production in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," *JHEP*, vol. 08, p. 033, 2019. DOI: 10.1007/JHEP08(2019)033. arXiv: 1903.02942 [hep-ex].

[158] ATLAS Collaboration, "Search for the electroweak diboson production in association with a high-mass dijet system in semileptonic final states

in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector," *Phys. Rev.*, vol. D100, no. 3, p. 032 007, 2019. DOI: 10.1103/PhysRevD.100.032007. arXiv: 1905.07714 [hep-ex].

[159] ATLAS Collaboration, "Search for diboson resonances in hadronic final states in 139 fb^{-1} of pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector," *JHEP*, vol. 09, p. 091, 2019. DOI: 10.1007/JHEP09(2019)091. arXiv: 1906.08589 [hep-ex].

[160] ATLAS Collaboration, "Search for dark matter in events with a hadronically decaying vector boson and missing transverse momentum in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector," *JHEP*, vol. 10, p. 180, 2018. DOI: 10.1007/JHEP10(2018)180. arXiv: 1807.11471 [hep-ex].

[161] ATLAS Collaboration, "Search for pair- and single-production of vector-like quarks in final states with at least one Z boson decaying into a pair of electrons or muons in pp collision data collected with the ATLAS detector," *Phys. Rev. D*, vol. 98, p. 112 010, 2018. DOI: 10.1103/PhysRevD.98.112010. arXiv: 1806.10555 [hep-ex].

[162] CMS Collaboration, "A multi-dimensional search for new heavy resonances decaying to boosted WW, WZ, or ZZ boson pairs in the dijet final state at 13 TeV," 2019. arXiv: 1906.05977 [hep-ex].

[163] CMS Collaboration, "Search for massive resonances decaying into pairs of boosted bosons in semi-leptonic final states at $\sqrt{s}$ = 8 TeV," *JHEP*, vol. 08, p. 174, 2014. DOI: 10.1007/JHEP08(2014)174. arXiv: 1405.3447 [hep-ex].

[164] ATLAS Collaboration, "Measurements of top-quark pair differential and double-differential cross-sections in the ℓ+jets channel with pp collisions

at $\sqrt{s}$ = 13 TeV using the ATLAS detector," *Eur. Phys. J. C*, vol. 79, no. 12, p. 1028, 2019. DOI: 10.1140/epjc/s10052-019-7525-6. arXiv: 1908.07305 [hep-ex].

[165] ATLAS Collaboration, "Measurements of $t\bar{t}$ differential cross-sections of highly boosted top quarks decaying to all-hadronic final states in pp collisions at $\sqrt{s}$ = 13 TeV using the ATLAS detector," *Phys. Rev.*, vol. D98, no. 1, p. 012 003, 2018. DOI: 10.1103/PhysRevD.98.012003. arXiv: 1801.02052 [hep-ex].

[166] ATLAS Collaboration, "Search for heavy particles decaying into a top-quark pair in the fully hadronic final state in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector," *Phys. Rev.*, vol. D99, no. 9, p. 092 004, 2019. DOI: 10.1103/PhysRevD.99.092004. arXiv: 1902.10077 [hep-ex].

[167] ATLAS Collaboration, "Search for a scalar partner of the top quark in the jets plus missing transverse momentum final state at $\sqrt{s}$=13 TeV with the ATLAS detector," *JHEP*, vol. 12, p. 085, 2017. DOI: 10.1007/JHEP12(2017)085. arXiv: 1709.04183 [hep-ex].

[168] ATLAS Collaboration, "Search for top-squark pair production in final states with one lepton, jets, and missing transverse momentum using 36 fb $^{-1}$ of $\sqrt{s}$ = 13 TeV pp collision data with the ATLAS detector," *JHEP*, vol. 06, p. 108, 2018. DOI: 10.1007/JHEP06(2018)108. arXiv: 1711.11520 [hep-ex].

[169] CMS Collaboration, "Search for anomalous $t\bar{t}$ production in the highly-boosted all-hadronic final state," *JHEP*, vol. 09, p. 029, 2012. DOI: 10.1007/JHEP09(2012)029. arXiv: 1204.2488 [hep-ex].

[170] CMS Collaboration, "Measurement of the jet mass in highly boosted $t\bar{t}$ events from pp collisions at $\sqrt{s}$ = 8 TeV," *Eur. Phys. J. C*, vol. 77, p. 467,

2017. DOI: 10.1140/epjc/s10052-017-5030-3. arXiv: 1703.06330 [hep-ex].

[171] CMS Collaboration, "Search for $t\bar{t}$ resonances in highly-boosted lepton+jets and fully hadronic final states in proton–proton collisions at $\sqrt{s}$ = 13 TeV," *JHEP*, vol. 07, p. 001, 2017. DOI: 10.1007/JHEP07(2017)001. arXiv: 1704.03366 [hep-ex].

[172] ATLAS Collaboration, "Measurement of soft-drop jet observables in pp collisions with the ATLAS detector at $\sqrt{s}$ = 13 TeV," *Phys. Rev.*, vol. D101, no. 5, p. 052 007, 2020. DOI: 10.1103/PhysRevD.101.052007. arXiv: 1912.09837 [hep-ex].

[173] ATLAS Collaboration, "Measurement of the Soft-Drop Jet Mass in pp Collisions at $\sqrt{s}$ = 13 TeV with the ATLAS Detector," *Phys. Rev. Lett.*, vol. 121, no. 9, p. 092 001, 2018. DOI: 10.1103/PhysRevLett.121.092001. arXiv: 1711.08341 [hep-ex].

[174] CMS Collaboration, "Search for anomalous couplings in boosted $WW/WZ \to \ell\nu q\bar{q}$ production in proton–proton collisions at $\sqrt{s}$ = 8 TeV," *Phys. Lett. B*, vol. 772, p. 21, 2017. DOI: 10.1016/j.physletb.2017.06.009. arXiv: 1703.06095 [hep-ex].

[175] ATLAS Collaboration, "Search for a right-handed gauge boson decaying into a high-momentum heavy neutrino and a charged lepton in pp collisions with the ATLAS detector at $\sqrt{s}$ = 13 TeV," *Phys. Lett.*, vol. B798, p. 134 942, 2019. DOI: 10.1016/j.physletb.2019.134942. arXiv: 1904.12679 [hep-ex].

[176] ATLAS Collaboration, "Search for R-parity-violating supersymmetric particles in multi-jet final states produced in pp collisions at $\sqrt{s}$ = 13 TeV using the ATLAS detector at the LHC," *Phys. Lett. B*, vol. 785, p. 136,

2018. DOI: 10.1016/j.physletb.2018.08.021. arXiv: 1804.03568 [hep-ex].

[177] ATLAS Collaboration, "Search for light resonances decaying to boosted quark pairs and produced in association with a photon or a jet in proton–proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector," *Phys. Lett. B*, vol. 788, p. 316, 2019. DOI: 10.1016/j.physletb.2018.09.062. arXiv: 1801.08769 [hep-ex].

[178] ATLAS Collaboration, "A search for resonances decaying into a Higgs boson and a new particle X in the $XH \to qqbb$ final state with the ATLAS detector," *Phys. Lett.*, vol. B779, pp. 24–45, 2018. DOI: 10.1016/j.physletb.2018.01.042. arXiv: 1709.06783 [hep-ex].

[179] CMS Collaboration, "Search for pair production of vector-like T and B quarks in single-lepton final states using boosted jet substructure techniques at $\sqrt{s}$ = 13 TeV," *JHEP*, vol. 11, p. 085, 2017. DOI: 10.1007/JHEP11(2017)085. arXiv: 1706.03408 [hep-ex].

[180] CMS Collaboration, "Inclusive search for supersymmetry in pp collisions at $\sqrt{s}$ = 13 TeV using razor variables and boosted object identification in zero and one lepton final states," *JHEP*, vol. 03, p. 031, 2019. DOI: 10.1007/JHEP03(2019)031. arXiv: 1812.06302 [hep-ex].

[181] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, "Longitudinally invariant K_t clustering algorithms for hadron hadron collisions," *Nucl. Phys. B*, vol. 406, pp. 187–224, 1993. DOI: 10.1016/0550-3213(93)90166-M.

[182] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-kt jet clustering algorithm," *Journal of High Energy Physics*, vol. 2008, no. 04, pp. 063–

063, 2008. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].

[183] S. D. Ellis and D. E. Soper, "Successive combination jet algorithm for hadron collisions," *Phys. Rev. D*, vol. 48, pp. 3160–3166, 7 1993. DOI: 10.1103/PhysRevD.48.3160. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.48.3160.

[184] Y. Dokshitzer, G. Leder, S. Moretti, and B. Webber, "Better jet clustering algorithms," *Journal of High Energy Physics*, vol. 1997, pp. 001–001, 1997. DOI: 10.1088/1126-6708/1997/08/001. arXiv: hep-ph/9907323 [hep-ph].

[185] M. Wobisch and T. Wengler, "Hadronization corrections to jet cross sections in deep-inelastic scattering," 1999. arXiv: hep-ph/9907280 [hep-ph].

[186] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, "Njettiness: An inclusive event shape to veto jets," *Physical Review Letters*, vol. 105, no. 9, 2010, ISSN: 1079-7114. DOI: 10.1103/physrevlett.105.092002. [Online]. Available: http://dx.doi.org/10.1103/PhysRevLett.105.092002.

[187] J. Thaler and K. Van Tilburg, "Identifying boosted objects with n-subjettiness," *Journal of High Energy Physics*, vol. 2011, no. 3, 2011, ISSN: 1029-8479. DOI: 10.1007/jhep03(2011)015. [Online]. Available: http://dx.doi.org/10.1007/JHEP03(2011)015.

[188] J. Thaler and K. Van Tilburg, "Maximizing boosted top identification by minimizing n-subjettiness," *Journal of High Energy Physics*, vol. 2012, no. 2, 2012, ISSN: 1029-8479. DOI: 10.1007/jhep02(2012)093. [Online]. Available: http://dx.doi.org/10.1007/JHEP02(2012)093.

[189] ATLAS Collaboration, "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1," *Eur. Phys. J. C*, vol. 77, p. 490, 2017. DOI: 10.1140/epjc/s10052-017-5004-5. arXiv: 1603.02934 [hep-ex].

[190] ATLAS Collaboration, "Optimisation of large-radius jets reconstruction for the ATLAS detector in $\sqrt{s}$ = 13 TeV proton–proton collisions," *Eur. Phys. J. C*, p. 65, 2020. arXiv: 2009.04986 [hep-ex].

[191] ATLAS Collaboration, "Performance of top-quark and W-boson tagging with ATLAS in Run 2 of the LHC," *Eur. Phys. J. C*, vol. 79, no. 5, p. 375, 2019. DOI: 10.1140/epjc/s10052-019-6847-8. arXiv: 1808.07858 [hep-ex].

[192] ATLAS Collaboration, "ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s}$ = 13 TeV," *Eur. Phys. J. C*, vol. 79, no. 11, p. 970, 2019. DOI: 10.1140/epjc/s10052-019-7450-8. arXiv: 1907.05120 [hep-ex].

[193] P. Berta, M. Spousta, D. W. Miller, and R. Leitner, "Particle-level pileup subtraction for jets and jet shapes," *Journal of High Energy Physics*, vol. 2014, no. 6, 2014, ISSN: 1029-8479. DOI: 10.1007/jhep06(2014)092. arXiv: 1403.3108 [hep-ex]. [Online]. Available: http://dx.doi.org/10.1007/JHEP06(2014)092.

[194] M. Cacciari, G. P. Salam, and G. Soyez, "Softkiller, a particle-level pileup removal method," *The European Physical Journal C*, vol. 75, no. 2, 2015, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-015-3267-2. arXiv: 1403.0408 [hep-ph]. [Online]. Available: http://dx.doi.org/10.1140/epjc/s10052-015-3267-2.

[195] D. Krohn, J. Thaler, and L.-T. Wang, "Jet trimming," *Journal of High Energy Physics*, vol. 2010, no. 2, 2010, ISSN: 1029-8479. DOI: 10.1007/jhep02(2010)084. arXiv: 0912.1342 [hep-ph]. [Online]. Available: http://dx.doi.org/10.1007/JHEP02(2010)084.

[196] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, "Soft drop," *Journal of High Energy Physics*, vol. 2014, no. 5, 2014, ISSN: 1029-8479. DOI: 10.1007/jhep05(2014)146. arXiv: 1402.2657 [hep-ph]. [Online]. Available: http://dx.doi.org/10.1007/JHEP05(2014)146.

[197] ATLAS Collaboration, "In situ calibration of large-radius jet energy and mass in 13 TeV proton–proton collisions with the ATLAS detector," *Eur. Phys. J. C*, vol. 79, p. 135, 2019. DOI: 10.1140/epjc/s10052-019-6632-8. arXiv: 1807.09477 [hep-ex].

[198] ATLAS Collaboration, "Electron and photon performance measurements with the ATLAS detector using the 2015-2017 LHC proton–proton collision data," 2019. arXiv: 1908.00005 [hep-ex].

[199] ATLAS Collaboration, "Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ = 13 TeV," *Eur. Phys. J. C*, vol. 76, p. 292, 2016. DOI: 10.1140/epjc/s10052-016-4120-y. arXiv: 1603.05598 [hep-ex].

[200] ATLAS Collaboration, "Measurement of large radius jet mass reconstruction performance at $\sqrt{s}$ = 8 TeV using the ATLAS detector," *ATLAS-CONF-2016-008*, 2016. [Online]. Available: https://cds.cern.ch/record/2139642.

[201] ATLAS collaboration, "Jet mass reconstruction with the ATLAS Detector in early Run 2 data," Jul. 2016. [Online]. Available: https://cds.cern.ch/record/2200211.

[202] ATLAS collaboration, "Measurement of the ATLAS Detector Jet Mass Response using Forward Folding with 80 fb^{-1} of $\sqrt{s}$ = 13 TeV pp data," Jul. 2020. [Online]. Available: https://inspirehep.net/literature/1808022.

[203] ATLAS Collaboration, "Search for heavy particles decaying into top-quark pairs using lepton-plus-jets events in proton–proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector," *Eur. Phys. J. C*, vol. 78, p. 565, 2018. DOI: 10.1140/epjc/s10052-018-5995-6. arXiv: 1804.10823 [hep-ex].

[204] ATLAS Collaboration, "Jet reclustering and close-by effects in ATLAS run II," Jul. 2017.

[205] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX," *JHEP*, vol. 06, p. 043, 2010. DOI: 10.1007/JHEP06(2010)043. arXiv: 1002.2581 [hep-ph].

[206] T. Sostrand, S. Mrenna, and P. Skands, "A brief introduction to pythia 8.1," *Computer Physics Communications*, vol. 178, no. 11, pp. 852 –867, 2008, ISSN: 0010-4655. DOI: https://doi.org/10.1016/j.cpc.2008.01.036. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0010465508000441.

[207] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, 2014. [Online]. Available: https://cds.cern.ch/record/1966419.

[208] J. Alwall *et al.*, "The automated computation of tree-level and next-to-leading order differential cross sections, and their matchin g to parton shower simulations," *Journal of High Energy Physics*, vol. 2014, no. 7,

2014, ISSN: 1029-8479. DOI: 10.1007/jhep07(2014)079. [Online]. Available: http://dx.doi.org/10.1007/JHEP07(2014)079.

[209] J. Bellm *et al.*, "Herwig 7.0/herwig++ 3.0 release note," *The European Physical Journal C*, vol. 76, no. 4, 2016, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-016-4018-8. [Online]. Available: http://dx.doi.org/10.1140/epjc/s10052-016-4018-8.

[210] A. H. Hoang, S. Mantry, A. Pathak, and I. W. Stewart, "Extracting a short distance top mass with light grooming," *Physical Review D*, vol. 100, no. 7, 2019, ISSN: 2470-0029. DOI: 10.1103/physrevd.100.074021. [Online]. Available: http://dx.doi.org/10.1103/PhysRevD.100.074021.

[211] S. Fleming, A. H. Hoang, S. Mantry, and I. W. Stewart, "Top jets in the peak region: Factorization analysis with next-to-leading-log resummation," *Phys. Rev. D*, vol. 77, p. 114 003, 11 2008. DOI: 10.1103/PhysRevD.77.114003. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.77.114003.

[212] NNPDF Collaboration, "Parton distributions for the LHC Run II," *JHEP*, vol. 04, p. 040, 2015. DOI: 10.1007/JHEP04(2015)040. arXiv: 1410.8849 [hep-ph].

[213] ATLAS Collaboration, "Studies on top-quark Monte Carlo modelling for Top2016," CERN, Geneva, Tech. Rep. ATL-PHYS-PUB-2016-020, Sep. 2016. [Online]. Available: https://cds.cern.ch/record/2216168.

[214] J. Butterworth *et al.*, "Pdf4lhc recommendations for lhc run ii," *Journal of Physics G: Nuclear and Particle Physics*, vol. 43, no. 2, p. 023 001, 2016, ISSN: 1361-6471. DOI: 10.1088/0954-3899/43/2/023001. [Online]. Available: http://dx.doi.org/10.1088/0954-3899/43/2/023001.

[215] ATLAS Collaboration, *Study of top-quark pair modelling and uncertainties using ATLAS measurements at* $\sqrt{s}$ = 13 *TeV*, ATL-PHYS-PUB-2020-023, 2020. [Online]. Available: https://cds.cern.ch/record/2730443.

[216] H. Brooks and P. Skands, "Coherent showers in decays of colored resonances," *Phys. Rev. D*, vol. 100, no. 7, p. 076 006, 2019. DOI: 10.1103/PhysRevD.100.076006. arXiv: 1907.08980 [hep-ph].

[217] J. Bellm *et al.*, "Herwig 7.0/Herwig++ 3.0 release note," *Eur. Phys. J.*, vol. C76, no. 4, p. 196, 2016. DOI: 10.1140/epjc/s10052-016-4018-8. arXiv: 1512.01178 [hep-ph].

[218] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet user manual," *Eur. Phys. J. C*, vol. 72, p. 1896, 2012. DOI: 10.1140/epjc/s10052-012-1896-2. arXiv: 1111.6097 [hep-ph].

[219] B. Bachu, A. H. Hoang, V. Mateu, A. Pathak, and I. W. Stewart, "Boosted Top Quarks in the Peak Region with N^{3}LL Resummation," Dec. 2020. arXiv: 2012.12304 [hep-ph].

[220] J. R. Christiansen and P. Z. Skands, "String formation beyond leading colour," *Journal of High Energy Physics*, vol. 2015, no. 8, 2015, ISSN: 1029-8479. DOI: 10.1007/jhep08(2015)003. [Online]. Available: http://dx.doi.org/10.1007/JHEP08(2015)003.

[221] ATLAS Collaboration, "Measurement of charged-particle distributions sensitive to the underlying event in $\sqrt{s}$ = 13TeV tev proton-proton collisions with the atlas detector at the lhc," *Journal of High Energy Physics*, vol. 2017, no. 3, 2017, ISSN: 1029-8479. DOI: 10.1007/jhep03(2017)157. [Online]. Available: http://dx.doi.org/10.1007/JHEP03(2017)157.

[222] J. H. Kuhn and G. Rodrigo, "Charge asymmetry in hadroproduction of heavy quarks," *Phys. Rev. Lett.*, vol. 81, pp. 49–52, 1998. DOI: 10.1103/PhysRevLett.81.49. arXiv: hep-ph/9802268.

[223] J. H. Kuhn and G. Rodrigo, "Charge asymmetry of heavy quarks at hadron colliders," *Phys. Rev. D*, vol. 59, p. 054 017, 5 1999. DOI: 10.1103/PhysRevD.59.054017. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.59.054017.

[224] CDF Collaboration, "Evidence for a Mass Dependent Forward-Backward Asymmetry in Top Quark Pair Production," *Phys. Rev. D*, vol. 83, p. 112 003, 2011. DOI: 10.1103/PhysRevD.83.112003. arXiv: 1101.0034 [hep-ex].

[225] DØ Collaboration, "Forward-backward asymmetry in top quark-antiquark production," *Phys. Rev. D*, vol. 84, p. 112 005, 2011. DOI: 10.1103/PhysRevD.84.112005. arXiv: 1107.4995 [hep-ex].

[226] ATLAS Collaboration, "Measurement of the top quark pair production charge asymmetry in proton-proton collisions at $\sqrt{s}$ = 7 tev using the atlas detector," *Journal of High Energy Physics*, vol. 2014, no. 2, 2014, ISSN: 1029-8479. DOI: 10.1007/jhep02(2014)107. [Online]. Available: http://dx.doi.org/10.1007/JHEP02(2014)107.

[227] CMS Collaboration, "Inclusive and Differential Measurements of the $t\bar{t}$ Charge Asymmetry in Proton-Proton Collisions at $\sqrt{s}$ = 7 TeV," *Phys. Lett. B*, vol. 717, pp. 129–150, 2012. DOI: 10.1016/j.physletb.2012.09.028. arXiv: 1207.0065 [hep-ex].

[228] CMS Collaboration, "Inclusive and differential measurements of the $t\bar{t}$ charge asymmetry in pp collisions at $\sqrt{s}$ = 8 tev," *Physics Letters B*, vol. 757, 154–179, 2016, ISSN: 0370-2693. DOI: 10.1016/j.physletb.

2016.03.060. [Online]. Available: http://dx.doi.org/10.1016/j.physletb.2016.03.060.

[229] M. Czakon, P. Fiedler, and A. Mitov, "Resolving the Tevatron Top Quark Forward-Backward Asymmetry Puzzle: Fully Differential Next-to-Next-to-Leading-Order Calculation," *Phys. Rev. Lett.*, vol. 115, no. 5, p. 052 001, 2015. DOI: 10.1103/PhysRevLett.115.052001. arXiv: 1411.3007 [hep-ph].

[230] N. Kidonakis, "The top quark forward-backward asymmetry at approximate N^{3}LO," *Phys. Rev. D*, vol. 91, no. 7, p. 071 502, 2015. DOI: 10.1103/PhysRevD.91.071502. arXiv: 1501.01581 [hep-ph].

[231] CDF Collaboration, "Measurement of the top quark forward-backward production asymmetry and its dependence on event kinematic properties," *Phys. Rev. D*, vol. 87, no. 9, p. 092 002, 2013. DOI: 10.1103/PhysRevD.87.092002. arXiv: 1211.1003 [hep-ex].

[232] DØ Collaboration, "Measurement of the Forward-Backward Asymmetry in Top Quark-Antiquark Production in $p\bar{p}$ Collisions using the Lepton+Jets Channel," *Phys. Rev. D*, vol. 90, p. 072 011, 2014. DOI: 10.1103/PhysRevD.90.072011. arXiv: 1405.0421 [hep-ex].

[233] M. P. Rosello and M. Vos, "Constraints on four-fermion interactions from the $t\bar{t}$ charge asymmetry at hadron colliders," *Eur. Phys. J. C*, vol. 76, no. 4, p. 200, 2016. DOI: 10.1140/epjc/s10052-016-4040-x. arXiv: 1512.07542 [hep-ex].

[234] G. Choudalakis, *Fully bayesian unfolding*, 2012. arXiv: 1201.4612 [physics.data-an].

[235] S. Agostinelli *et al.*, "Geant4—a simulation toolkit," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers,*

Detectors and Associated Equipment, vol. 506, no. 3, pp. 250–303, 2003, ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(03)01368-8. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168900203013688.

[236] A. Hoecker *et al.*, *Tmva - toolkit for multivariate data analysis*, 2009. arXiv: physics/0703039 [physics.data-an].

[237] J. Erdmann *et al.*, "A likelihood-based reconstruction algorithm for top-quark pairs and the klfitter framework," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 748, 18–25, 2014, ISSN: 0168-9002. DOI: 10.1016/j.nima.2014.02.029. [Online]. Available: http://dx.doi.org/10.1016/j.nima.2014.02.029.

[238] DØ Collaboration, "Measurement of the top quark mass using dilepton events," *Physical Review Letters*, vol. 80, no. 10, 2063–2068, 1998, ISSN: 1079-7114. DOI: 10.1103/physrevlett.80.2063. [Online]. Available: http://dx.doi.org/10.1103/PhysRevLett.80.2063.

[239] M. Czakon, D. Heymes, A. Mitov, D. Pagani, I. Tsinikos, and M. Zaro, "Top-quark charge asymmetry at the LHC and Tevatron through NNLO QCD and NLO EW," *Phys. Rev. D*, vol. 98, no. 1, p. 014 003, 2018. DOI: 10.1103/PhysRevD.98.014003. arXiv: 1711.03945 [hep-ph].

[240] M. Czakon, D. Heymes, and A. Mitov, "Dynamical scales for multi-TeV top-pair production at the LHC," *JHEP*, vol. 04, p. 071, 2017. DOI: 10.1007/JHEP04(2017)071. arXiv: 1606.03350 [hep-ph].

[241] C. Degrande, G. Durieux, F. Maltoni, K. Mimasu, E. Vryonidou, and C. Zhang, "Automated one-loop computations in the standard model effective

field theory," *Phys. Rev. D*, vol. 103, no. 9, p. 096 024, 2021. DOI: 10.1103/PhysRevD.103.096024. arXiv: 2008.11743 [hep-ph].

[242] J. J. Ethier *et al.*, "Combined SMEFT interpretation of Higgs, diboson, and top quark data from the LHC," Apr. 2021. arXiv: 2105.00006 [hep-ph].

[243] R. Barlow, "Asymmetric errors," *eConf*, vol. C030908, L. Lyons, R. P. Mount, and R. Reitmeyer, Eds., WEMT002, 2003. arXiv: physics/0401042.

[244] "Combination of inclusive and differential $t\bar{t}$ charge asymmetry measurements using ATLAS and CMS data at $\sqrt{s} = 7$ and 8 TeV," *JHEP*, vol. 04, p. 033, 2018. DOI: 10.1007/JHEP04(2018)033. arXiv: 1709.05327 [hep-ex].

[245] CDF and DØ Collaborations, "Combined Forward-Backward Asymmetry Measurements in Top-Antitop Quark Production at the Tevatron," *Phys. Rev. Lett.*, vol. 120, no. 4, p. 042 001, 2018. DOI: 10.1103/PhysRevLett.120.042001. arXiv: 1709.04894 [hep-ex].

[246] I. Brivio *et al.*, "O new physics, where art thou? A global search in the top sector," *JHEP*, vol. 02, p. 131, 2020. DOI: 10.1007/JHEP02(2020)131. arXiv: 1910.03606 [hep-ph].

[247] D. Barducci *et al.*, "Interpreting top-quark LHC measurements in the standard-model effective field theory," J. A. Aguilar-Saavedra, C. Degrande, G. Durieux, F. Maltoni, E. Vryonidou, and C. Zhang, Eds., 2018. arXiv: 1802.07237 [hep-ph].

[248] N. P. Hartland *et al.*, "A Monte Carlo global analysis of the Standard Model Effective Field Theory: the top quark sector," *JHEP*, vol. 04, p. 100, 2019. DOI: 10.1007/JHEP04(2019)100. arXiv: 1901.05965 [hep-ph].

[249] J. Ellis, M. Madigan, K. Mimasu, V. Sanz, and T. You, "Top, Higgs, Diboson and Electroweak Fit to the Standard Model Effective Field Theory," *JHEP*, vol. 04, p. 279, 2021. DOI: 10.1007/JHEP04(2021)279. arXiv: 2012.02779 [hep-ph].

[250] J. Aparisi *et al.*, "m_b At m_H: The running bottom quark mass and the higgs boson," *Phys. Rev. Lett.*, vol. 128, p. 122 001, 12 2022. DOI: 10.1103/PhysRevLett.128.122001. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.128.122001.

[251] J. Vermaseren, S. Larin, and T van Ritbergen, "The 4-loop quark mass anomalous dimension and the invariant quark mass," *Physics Letters B*, vol. 405, no. 3–4, 327–333, 1997, ISSN: 0370-2693. DOI: 10.1016/s0370-2693(97)00660-6. [Online]. Available: http://dx.doi.org/10.1016/S0370-2693(97)00660-6.

[252] K. Chetyrkin, "Quark mass anomalous dimension to $\mathcal{O}(\alpha_s^4)$," *Physics Letters B*, vol. 404, no. 1–2, 161–165, 1997, ISSN: 0370-2693. DOI: 10.1016/s0370-2693(97)00535-2. [Online]. Available: http://dx.doi.org/10.1016/S0370-2693(97)00535-2.

[253] P. A. Baikov, K. G. Chetyrkin, and J. H. Kühn, "Quark mass and field anomalous dimensions to $\mathcal{O}(\alpha_s^5)$," *Journal of High Energy Physics*, vol. 2014, no. 10, 2014, ISSN: 1029-8479. DOI: 10.1007/jhep10(2014)076. [Online]. Available: http://dx.doi.org/10.1007/JHEP10(2014)076.

[254] S. Narison, "$\overline{m}_c$ and $\overline{m}_b$ from M_{Bc} and improved estimates of f_{Bc} and $f_{Bc(2S)}$," *Phys. Lett. B*, vol. 802, p. 135 221, 2020. DOI: 10.1016/j.physletb.2020.135221. arXiv: 1906.03614 [hep-ph].

[255] C. Peset, A. Pineda, and J. Segovia, "The charm/bottom quark mass from heavy quarkonium at N^{3}LO," *JHEP*, vol. 09, p. 167, 2018. DOI: 10.1007/JHEP09(2018)167. arXiv: 1806.05197 [hep-ph].

[256] Y. Kiyo, G. Mishima, and Y. Sumino, "Determination of m_c and m_b from quarkonium 1S energy levels in perturbative QCD," *Phys. Lett.*, vol. B752, pp. 122–127, 2016, [Erratum: Phys. Lett.B772,878(2017)]. DOI: 10.1016/j.physletb.2015.11.040,10.1016/j.physletb.2017.09.024. arXiv: 1510.07072 [hep-ph].

[257] A. A. Penin and N. Zerf, "Bottom Quark Mass from Υ Sum Rules to $\mathcal{O}(\alpha_s^3)$," *JHEP*, vol. 04, p. 120, 2014. DOI: 10.1007/JHEP04(2014)120. arXiv: 1401.7035 [hep-ph].

[258] A. Alberti, P. Gambino, K. J. Healey, and S. Nandi, "Precision Determination of the Cabibbo-Kobayashi-Maskawa Element V_{cb}," *Phys. Rev. Lett.*, vol. 114, no. 6, p. 061 802, 2015. DOI: 10.1103/PhysRevLett.114.061802. arXiv: 1411.6560 [hep-ph].

[259] M. Beneke, A. Maier, J. Piclum, and T. Rauh, "The bottom-quark mass from non-relativistic sum rules at NNNLO," *Nucl. Phys. B*, vol. 891, pp. 42–72, 2015. DOI: 10.1016/j.nuclphysb.2014.12.001. arXiv: 1411.3132 [hep-ph].

[260] B. Dehnadi, A. H. Hoang, and V. Mateu, "Bottom and Charm Mass Determinations with a Convergence Test," *JHEP*, vol. 08, A. A. Petrov, Ed., p. 155, 2015. DOI: 10.1007/JHEP08(2015)155. arXiv: 1504.07638 [hep-ph].

[261] W. Lucha, D. Melikhov, and S. Simula, "Accurate bottom-quark mass from Borel QCD sum rules for f_B and f_{B_s}," *Phys. Rev. D*, vol. 88,

p. 056 011, 2013. DOI: 10.1103/PhysRevD.88.056011. arXiv: 1305.7099 [hep-ph].

[262] S. Bodenstein, J. Bordes, C. Dominguez, J. Penarrocha, and K. Schilcher, "Bottom-quark mass from finite energy QCD sum rules," *Phys. Rev. D*, vol. 85, p. 034 003, 2012. DOI: 10.1103/PhysRevD.85.034003. arXiv: 1111.5742 [hep-ph].

[263] A. Laschka, N. Kaiser, and W. Weise, "Quark-antiquark potential to order 1/m and heavy quark masses," *Phys. Rev. D*, vol. 83, p. 094 002, 2011. DOI: 10.1103/PhysRevD.83.094002. arXiv: 1102.0945 [hep-ph].

[264] K. Chetyrkin *et al.*, "Charm and Bottom Quark Masses: An Update," *Phys. Rev. D*, vol. 80, p. 074 010, 2009. DOI: 10.1103/PhysRevD.80.074010. arXiv: 0907.2110 [hep-ph].

[265] H1 and Z. Collaborations, "Combination and QCD analysis of charm and beauty production cross-section measurements in deep inelastic *ep* scattering at HERA," *Eur. Phys. J. C*, vol. 78, no. 6, p. 473, 2018. DOI: 10.1140/epjc/s10052-018-5848-3. arXiv: 1804.01019 [hep-ex].

[266] C. Schwanda *et al.*, "Measurement of the Moments of the Photon Energy Spectrum in B —> X(s) gamma Decays and Determination of —V(cb)— and m(b) at Belle," *Phys. Rev. D*, vol. 78, p. 032 016, 2008. DOI: 10.1103/PhysRevD.78.032016. arXiv: 0803.2158 [hep-ex].

[267] B. Aubert *et al.*, "Measurement and interpretation of moments in inclusive semileptonic decays anti-B —> X(c) l- anti-nu," *Phys. Rev. D*, vol. 81, p. 032 003, 2010. DOI: 10.1103/PhysRevD.81.032003. arXiv: 0908.0415 [hep-ex].

[268] M. S. Bilenky, G. Rodrigo, and A. Santamaria, "Three jet production at LEP and the bottom quark mass," *Nucl. Phys. B*, vol. 439, pp. 505–535, 1995. DOI: 10.1016/0550-3213(94)00586-4. arXiv: hep-ph/9410258.

[269] G. Rodrigo, A. Santamaria, and M. S. Bilenky, "Do the quark masses run? Extracting m-bar(b) (m(z)) from LEP data," *Phys. Rev. Lett.*, vol. 79, pp. 193–196, 1997. DOI: 10.1103/PhysRevLett.79.193. arXiv: hep-ph/9703358 [hep-ph].

[270] W. Bernreuther, A. Brandenburg, and P. Uwer, "Next-to-leading order QCD corrections to three jet cross-sections with massive quarks," *Phys. Rev. Lett.*, vol. 79, pp. 189–192, 1997. DOI: 10.1103/PhysRevLett.79.189. arXiv: hep-ph/9703305 [hep-ph].

[271] A. Brandenburg and P. Uwer, "Next-to-leading order QCD corrections and massive quarks in e+ e- —> three jets," *Nucl. Phys. B*, vol. 515, pp. 279–320, 1998. DOI: 10.1016/S0550-3213(97)00790-6. arXiv: hep-ph/9708350.

[272] P. Nason and C. Oleari, "Next-to-leading order corrections to momentum correlations in Z0 —> b anti-b," *Phys. Lett. B*, vol. 407, pp. 57–60, 1997. DOI: 10.1016/S0370-2693(97)00721-1. arXiv: hep-ph/9705295.

[273] P. Nason and C. Oleari, "Next-to-leading order corrections to the production of heavy flavor jets in e+ e- collisions," *Nucl. Phys. B*, vol. 521, pp. 237–273, 1998. DOI: 10.1016/S0550-3213(98)00125-4. arXiv: hep-ph/9709360.

[274] DELPHI Collaboration, "Determination of the b quark mass at the mz scale with the delphi detector at lep," *The European Physical Journal C*, vol. 46, no. 3, 569–583, 2006, ISSN: 1434-6052. DOI: 10.1140/epjc/s2006-

02497-6. [Online]. Available: http://dx.doi.org/10.1140/epjc/s2006-02497-6.

[275] DELPHI Collaboration, "Study of b-quark mass effects in multijet topologies with the delphi detector at lep," *The European Physical Journal C*, vol. 55, no. 4, 2008, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-008-0631-5. [Online]. Available: http://dx.doi.org/10.1140/epjc/s10052-008-0631-5.

[276] ALEPH Collaboration, "A measurement of the b-quark mass from hadronic z decays," *The European Physical Journal C*, vol. 18, no. 1, 1–13, 2000, ISSN: 1434-6052. DOI: 10.1007/s100520000533. [Online]. Available: http://dx.doi.org/10.1007/s100520000533.

[277] OPAL Collaboration, "Determination of the b quark mass at the z mass scale," *The European Physical Journal C*, vol. 21, no. 3, 411–422, 2001, ISSN: 1434-6052. DOI: 10.1007/s100520100746. [Online]. Available: http://dx.doi.org/10.1007/s100520100746.

[278] A. Brandenburg, P. Burrows, D. Muller, N. Oishi, and P. Uwer, "Measurement of the running b-quark mass using e^+e^- events," *Physics Letters B*, vol. 468, no. 1–2, 168–177, 1999, ISSN: 0370-2693. DOI: 10.1016/s0370-2693(99)01194-6. [Online]. Available: http://dx.doi.org/10.1016/S0370-2693(99)01194-6.

[279] SLD Collaboration, "Improved test of the flavor independence of strong interactions," *Physical Review D*, vol. 59, no. 1, 1998, ISSN: 1089-4918. DOI: 10.1103/physrevd.59.012002. [Online]. Available: http://dx.doi.org/10.1103/PhysRevD.59.012002.

[280] J. Kieseler, "A method and tool for combining differential or inclusive measurements obtained with simultaneously constrained uncertainties,"

The European Physical Journal C, vol. 77, no. 11, 2017, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-017-5345-0. [Online]. Available: http://dx.doi.org/10.1140/epjc/s10052-017-5345-0.

[281] ATLAS Collaboration, "Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector," *Phys. Lett. B*, vol. 786, pp. 59–86, 2018. DOI: 10.1016/j.physletb.2018.09.013. arXiv: 1808.08238 [hep-ex].

[282] CMS Collaboration, "Observation of Higgs boson decay to bottom quarks," *Phys. Rev. Lett.*, vol. 121, no. 12, p. 121 801, 2018. DOI: 10.1103/PhysRevLett.121.121801. arXiv: 1808.08242 [hep-ex].

[283] ATLAS and CMS Collaborations, "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s}$ = 7 and 8 TeV," 2015. [Online]. Available: http://cds.cern.ch/record/2052552.

[284] CMS Collaboration, "Combined measurements of Higgs boson couplings in proton–proton collisions at $\sqrt{s}$ = 13 TeV," *Eur. Phys. J.*, vol. C79, no. 5, p. 421, 2019. DOI: 10.1140/epjc/s10052-019-6909-y. arXiv: 1809.10733 [hep-ex].

[285] ATLAS Collaboration, "Combined measurements of Higgs boson production and decay using up to 80 fb^{-1} of proton-proton collision data at $\sqrt{s}$ = 13 TeV collected with the ATLAS experiment," *Phys. Rev. D*, vol. 101, no. 1, p. 012 002, 2020. DOI: 10.1103/PhysRevD.101.012002. arXiv: 1909.02845 [hep-ex].

[286] ATLAS Collaboration, "A combination of measurements of Higgs boson production and decay using up to 139 fb^{-1} of proton–proton collision data at $\sqrt{s}$ = 13 TeV collected with the ATLAS experiment," *ATLAS-CONF-*

2020-027, 2020. [Online]. Available: https://cds.cern.ch/record/2725733.

[287] E. Braaten and J. P. Leveille, "Higgs Boson Decay and the Running Mass," *Phys. Rev. D*, vol. 22, p. 715, 1980. DOI: 10.1103/PhysRevD.22.715.

[288] N. Sakai, "Perturbative QCD Corrections to the Hadronic Decay Width of the Higgs Boson," *Phys. Rev. D*, vol. 22, p. 2220, 1980. DOI: 10.1103/PhysRevD.22.2220.

[289] T. Inami and T. Kubota, "Renormalization Group Estimate of the Hadronic Decay Width of the Higgs Boson," *Nucl. Phys.*, vol. B179, pp. 171–188, 1981. DOI: 10.1016/0550-3213(81)90253-4.

[290] M. Drees and K.-I. Hikasa, "Note on QCD Corrections to Hadronic Higgs Decay," *Phys. Lett.*, vol. B240, p. 455, 1990, [Erratum: Phys. Lett.B262,497(1991)]. DOI: 10.1016/0370-2693(90)91130-4.

[291] M. Drees and K.-I. Hikasa, "Heavy Quark Thresholds in Higgs Physics," *Phys. Rev.*, vol. D41, p. 1547, 1990. DOI: 10.1103/PhysRevD.41.1547.

[292] S. G. Gorishnii, A. L. Kataev, and S. A. Larin, "The Width of Higgs Boson Decay Into Hadrons: Three Loop Corrections of Strong Interactions," *Sov. J. Nucl. Phys.*, vol. 40, pp. 329–334, 1984, [Yad. Fiz.40,517(1984)].

[293] S. G. Gorishnii, A. L. Kataev, S. A. Larin, and L. R. Surguladze, "Corrected Three Loop QCD Correction to the Correlator of the Quark Scalar Currents and $\Gamma_{tot}(H^0 \to \text{Hadrons})$," *Mod. Phys. Lett.*, vol. A5, pp. 2703–2712, 1990. DOI: 10.1142/S0217732390003152.

[294] S. G. Gorishnii, A. L. Kataev, S. A. Larin, and L. R. Surguladze, "Scheme dependence of the next to next-to-leading QCD corrections to $\Gamma_{tot}(H^0 \to \text{hadrons})$ and the spurious QCD infrared fixed point," *Phys. Rev.*, vol. D43, pp. 1633–1640, 1991. DOI: 10.1103/PhysRevD.43.1633.

[295] A. L. Kataev and V. T. Kim, "The Effects of the QCD corrections to $\Gamma(H^0 \to b\bar{b})$," *Mod. Phys. Lett.*, vol. A9, pp. 1309–1326, 1994. DOI: 10.1142/S0217732394001131.

[296] L. R. Surguladze, "Quark mass effects in fermionic decays of the Higgs boson in $\mathcal{O}(\alpha_s^2)$ perturbative QCD," *Phys. Lett.*, vol. B341, pp. 60–72, 1994. DOI: 10.1016/0370-2693(94)01253-9. arXiv: hep-ph/9405325 [hep-ph].

[297] K. G. Chetyrkin, "Correlator of the quark scalar currents and $\Gamma_{tot}(H \to$ hadrons) at $\mathcal{O}(\alpha_s^3)$ in pQCD," *Phys. Lett.*, vol. B390, pp. 309–317, 1997. DOI: 10.1016/S0370-2693(96)01368-8. arXiv: hep-ph/9608318 [hep-ph].

[298] K. Melnikov, "Two loop $\mathcal{O}(N_f\alpha_s^2)$ correction to the decay width of the Higgs boson to two massive fermions," *Phys. Rev.*, vol. D53, pp. 5020–5027, 1996. DOI: 10.1103/PhysRevD.53.5020. arXiv: hep-ph/9511310 [hep-ph].

[299] P. A. Baikov, K. G. Chetyrkin, and J. H. Kuhn, "Scalar correlator at O(alpha(s)**4), Higgs decay into b-quarks and bounds on the light quark masses," *Phys. Rev. Lett.*, vol. 96, p. 012 003, 2006. DOI: 10.1103/PhysRevLett.96.012003. arXiv: hep-ph/0511063.

[300] F. Herzog, B. Ruijl, T. Ueda, J. A. M. Vermaseren, and A. Vogt, "On Higgs decays to hadrons and the R-ratio at N^{4}LO," *JHEP*, vol. 08, p. 113, 2017. DOI: 10.1007/JHEP08(2017)113. arXiv: 1707.01044 [hep-ph].

[301] K. G. Chetyrkin and A. Kwiatkowski, "Second order QCD corrections to scalar and pseudoscalar Higgs decays into massive bottom quarks," *Nucl. Phys. B*, vol. 461, pp. 3–18, 1996. DOI: 10.1016/0550-3213(95)00616-8. arXiv: hep-ph/9505358.

[302] S. A. Larin, T. van Ritbergen, and J. A. M. Vermaseren, "The Large top quark mass expansion for Higgs boson decays into bottom quarks and into gluons," *Phys. Lett.*, vol. B362, pp. 134–140, 1995. DOI: 10.1016/0370-2693(95)01192-S. arXiv: hep-ph/9506465 [hep-ph].

[303] A. Primo, G. Sasso, G. Somogyi, and F. Tramontano, "Exact Top Yukawa corrections to Higgs boson decay into bottom quarks," *Phys. Rev.*, vol. D99, no. 5, p. 054013, 2019. DOI: 10.1103/PhysRevD.99.054013. arXiv: 1812.07811 [hep-ph].

[304] J. Fleischer and F. Jegerlehner, "Radiative Corrections to Higgs Decays in the Extended Weinberg-Salam Model," *Phys. Rev.*, vol. D23, pp. 2001–2026, 1981. DOI: 10.1103/PhysRevD.23.2001.

[305] D. Yu. Bardin, B. M. Vilensky, and P. K. Khristova, "Calculation of the Higgs boson decay width into fermion pairs," *Sov. J. Nucl. Phys.*, vol. 53, pp. 152–158, 1991, [Yad. Fiz.53,240(1991)].

[306] A. Dabelstein and W. Hollik, "Electroweak corrections to the fermionic decay width of the standard Higgs boson," *Z. Phys.*, vol. C53, pp. 507–516, 1992. DOI: 10.1007/BF01625912.

[307] B. A. Kniehl, "Radiative corrections for $H \to$ f anti-f (γ) in the standard model," *Nucl. Phys.*, vol. B376, pp. 3–28, 1992. DOI: 10.1016/0550-3213(92)90065-J.

[308] A. L. Kataev, "The Order $\mathcal{O}(\alpha\alpha_s)$ and $\mathcal{O}O(\alpha^2)$ corrections to the decay width of the neutral Higgs boson to the anti-b b pair," *JETP Lett.*, vol. 66, pp. 327–330, 1997. DOI: 10.1134/1.567516. arXiv: hep-ph/9708292 [hep-ph].

[309] B. A. Kniehl and M. Spira, "Two loop $\mathcal{O}(\alpha_s G_F m_t^2)$ correction to the $H \to b\bar{b}$ decay rate," *Nucl. Phys.*, vol. B432, pp. 39–48, 1994. DOI: 10.1016/0550-3213(94)90592-4. arXiv: hep-ph/9410319 [hep-ph].

[310] A. Kwiatkowski and M. Steinhauser, "Corrections of order $\mathcal{O}(G_F \alpha_s m_t^2)$ to the Higgs decay rate $\Gamma(H \to b\bar{b})$," *Phys. Lett.*, vol. B338, pp. 66–70, 1994, [Erratum: Phys. Lett.B342,455(1995)]. DOI: 10.1016/0370-2693(94)01527-J,10.1016/0370-2693(94)91345-5. arXiv: hep-ph/9405308 [hep-ph].

[311] K. G. Chetyrkin, B. A. Kniehl, and M. Steinhauser, "Virtual top quark effects on the $H \to b\bar{b}$ decay at next-to-leading order in QCD," *Phys. Rev. Lett.*, vol. 78, pp. 594–597, 1997. DOI: 10.1103/PhysRevLett.78.594. arXiv: hep-ph/9610456 [hep-ph].

[312] L. Mihaila, B. Schmidt, and M. Steinhauser, "$\Gamma(H \to b\bar{b})$ to order $\alpha\alpha_s$," *Phys. Lett.*, vol. B751, pp. 442–447, 2015. DOI: 10.1016/j.physletb.2015.10.078. arXiv: 1509.02294 [hep-ph].

[313] E. Chaubey and S. Weinzierl, "Two-loop master integrals for the mixed QCD-electroweak corrections for $H \to b\bar{b}$ through a $Ht\bar{t}$-coupling,," 2019. arXiv: 1904.00382 [hep-ph].

[314] A. Djouadi, J. Kalinowski, M. Mühlleitner, and M. Spira, "HDECAY: Twenty$_{++}$ years after," *Comput. Phys. Commun.*, vol. 238, pp. 214–231, 2019. DOI: 10.1016/j.cpc.2018.12.010. arXiv: 1801.09506 [hep-ph].

[315] A. Djouadi, J. Kalinowski, and M. Spira, "HDECAY: A Program for Higgs boson decays in the standard model and its supersymmetric extension," *Comput. Phys. Commun.*, vol. 108, pp. 56–74, 1998. DOI: 10.1016/S0010-4655(97)00123-9. arXiv: hep-ph/9704448 [hep-ph].

[316] A. Bredenstein, A. Denner, S. Dittmaier, and M. M. Weber, "Radiative corrections to the semileptonic and hadronic Higgs-boson decays H —> W W / Z Z —> 4 fermions," *JHEP*, vol. 02, p. 080, 2007. DOI: 10.1088/1126-6708/2007/02/080. arXiv: hep-ph/0611234.

[317] A. Bredenstein, A. Denner, S. Dittmaier, and M. M. Weber, "Precise predictions for the Higgs-boson decay H —> WW/ZZ —> 4 leptons," *Phys. Rev. D*, vol. 74, p. 013 004, 2006. DOI: 10.1103/PhysRevD.74.013004. arXiv: hep-ph/0604011.

[318] A. Denner, S. Dittmaier, and A. Mück, "PROPHECY4F 3.0: A Monte Carlo program for Higgs-boson decays into four-fermion final states in and beyond the Standard Model," *Comput. Phys. Commun.*, vol. 254, p. 107 336, 2020. DOI: 10.1016/j.cpc.2020.107336. arXiv: 1912.02010 [hep-ph].

[319] D. de Florian *et al.*, "Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector," vol. 2/2017, Oct. 2016. DOI: 10.23731/CYRM-2017-002. arXiv: 1610.07922 [hep-ph].

[320] A. H. Hoang, C. Lepenik, and V. Mateu, "REvolver: Automated running and matching of couplings and masses in QCD," Feb. 2021. arXiv: 2102.01085 [hep-ph].

[321] P. Zyla *et al.*, "Review of Particle Physics," *PTEP*, vol. 2020, no. 8, p. 083C01, 2020. DOI: 10.1093/ptep/ptaa104.

[322] CMS Collaboration, "Running of the top quark mass from proton-proton collisions at $\sqrt{s}$ = 13TeV," *Phys. Lett. B*, vol. 803, p. 135 263, 2020. DOI: 10.1016/j.physletb.2020.135263. arXiv: 1909.09193 [hep-ex].

[323] M. Cepeda *et al.*, "Report from Working Group 2," *CERN Yellow Rep. Monogr.*, vol. 7, pp. 221–584, 2019. DOI: 10.23731/CYRM-2019-007.221. arXiv: 1902.00134 [hep-ph].

[324] J. de Blas *et al.*, "Higgs Boson Studies at Future Particle Colliders," *JHEP*, vol. 01, p. 139, 2020. DOI: 10.1007/JHEP01(2020)139. arXiv: 1905.03764 [hep-ph].

[325] T. Barklow *et al.*, "Improved Formalism for Precision Higgs Coupling Fits," *Phys. Rev. D*, vol. 97, no. 5, p. 053 003, 2018. DOI: 10.1103/PhysRevD.97.053003. arXiv: 1708.08912 [hep-ph].

[326] H. Abramowicz *et al.*, "Higgs physics at the CLIC electron–positron linear collider," *Eur. Phys. J. C*, vol. 77, no. 7, p. 475, 2017. DOI: 10.1140/epjc/s10052-017-4968-5. arXiv: 1608.07538 [hep-ex].

[327] K. Fujii *et al.*, "Tests of the Standard Model at the International Linear Collider," Aug. 2019. arXiv: 1908.11299 [hep-ex].

[328] P. Bambade *et al.*, "The International Linear Collider: A Global Project," Mar. 2019. arXiv: 1903.01629 [hep-ex].

[329] J. Fuster *et al.*, "Prospects for the measurement of the bottom quark mass at the ILC," *ILD note*, eprint: ILD-PHYS-PUB--2021-001.