

Natural Supersymmetry on trial at the CMS Experiment

**Dissertation
zur Erlangung des Doktorgrades
des Fachbereichs Physik
der Universität Hamburg**

vorgelegt von
Kristin Goebel geb. Heine
aus Gifhorn

Hamburg
2014

Gutachter der Dissertation: Prof. Dr. bla
Prof. Dr. bla

Gutachter der Disputation: Prof. Dr. bla
Prof. Dr. bla

Datum der Disputation: ?? . ?? 2014

Vorsitzender des Prüfungsausschusses: Dr. bla

Vorsitzender des Promotionsausschusses: Prof. Dr. bla

Leiterin des Fachbereichs Physik: Prof. Dr. bla

Dekan der Fakultät für Mathematik,
Informatik und Naturwissenschaften: Prof. Dr. bla

Abstract

The search for signals of new physics processes beyond the standard model of particle physics is one of the main goals of the CMS experiment at the CERN Large Hadron Collider. Since many theories like e.g. supersymmetry involve the production of new coloured particles which feature jets as their experimental signature, it is important to have a good understanding of jet properties in order to allow such searches.

In the first part of this thesis, a measurement of the jet transverse momentum resolution is presented based on the analysis of proton-proton collision data recorded at a center of mass energy of 8 TeV by the CMS experiment. The measurement utilizes the transverse momentum balance of dijet events at particle level. The main focus lies on the determination of the data-to-simulation ratio of the resolution which can be used to derive scale factors in order to correct the resolution in simulated events to match the observed resolution in data.

The second part of the thesis focuses on searches for supersymmetry in final states with jets and missing transverse momentum. First, a search performed on collision data recorded at $\sqrt{s} = 8$ TeV is presented which targets mainly signatures sensitive to the production of light squarks and gluinos as well as the production of gluino mediated third generation particles. Here the main focus lies on the estimation of background contributions arising from QCD multijet events. Second, a study based on simulated events at a center of mass energy of 13 TeV is shown investigating different search strategies towards the identification of directly produced top squarks. Utilizing algorithms for the identification of boosted hadronically decaying top quarks arising from the decay of heavy top squarks, a search sensitivity of top squark masses up to 1 TeV is expected for LSP masses less than approximately 200 GeV.

Kurzfassung

Die Suche nach Hinweisen auf Prozesse von neuer Physik jenseits des Standardmodells der Teilchenphysik ist eines der Hauptziele des CMS Experiments am CERN Large Hadron Collider. Da zahlreiche Theorien wie beispielsweise Supersymmetrie die Produktion von neuen farbgeladenen Teilchen, welche als experimentelle Signatur Jets aufweisen, beinhalten, ist es wichtig ein gutes Verständnis dieser Objekte zu erlangen um derartige Suchen zu ermöglichen.

Im ersten Teil dieser Arbeit wird eine Messung der Jet-Transversalimpuls-Auflösung vorgestellt, welche auf der Analyse von Proton-Proton-Kollisionsdaten basiert, die bei einer Schwerpunktsenergie von 8 TeV vom CMS Experiment aufgezeichnet wurden. Die Messung nutzt dabei die Transversalimpulsbalance von Zweijet-Ereignissen auf Teilchenniveau. Der Fokus liegt dabei auf der Bestimmung des Verhältnisses der Auflösung in Daten zu der Auflösung in simulierten Ereignissen, welches verwendet werden kann um Skalierungsfaktoren zu bestimmen, welche die in simulierten Ereignissen gemessene Auflösung an die in Daten beobachtete anpassen.

Der zweite Teil der Arbeit konzentriert sich auf Suchen nach Supersymmetrie in Endzuständen mit Jets und fehlendem Transversalimpuls. Zunächst wird eine Suche unter Verwendung von bei einer Schwerpunktsenergie von $\sqrt{s} = 8$ TeV aufgezeichneten Kollisionsdaten durchgeführt, die auf Signaturen abzielt, welche hauptsächlich sensitiv sind auf die Produktion von leichten Squarks und Gluinos sowie die Produktion von gluinoindizierten Teilchen der dritten Generation. Der Schwerpunkt liegt hier auf der Abschätzung des Untergrundbeitrags durch QCD Mulijet Ereignisse. Als zweites wird eine Studie vorgestellt, welche auf simulierten Ereignissen bei einer Schwerpunktsenergie von 13 TeV basiert und unterschiedliche Analysestrategien zur Identifikation von direkt produzierten Top Squarks untersucht. Unter Verwendung von Algorithmen zur Identifikation von geboosteten hadronisch zerfallenden Top Quarks aus den Zerfällen von Top Squarks, kann eine Sensitivität der Suche für Top Squarks Massen bis 1 TeV für LSP Massen bis circa 200 GeV erwartet werden.

Contents

1	Introduction	1
2	Theoretical Background	3
2.1	The Standard Model of Particle Physics	3
2.1.1	Limitations of the Standard Model	6
2.2	Supersymmetry	7
2.2.1	The MSSM	9
2.2.2	SUSY-Breaking	11
2.3	Searches for Supersymmetry and Current Constraints	12
2.3.1	Constraints from Indirect Measurements	12
2.3.2	Direct Searches at Collider Experiments	12
3	Experimental Setup	16
3.1	The Large Hadron Collider	16
3.2	The CMS Experiment	18
3.2.1	Coordinate Conventions and Kinematic Variables	19
3.2.2	Superconducting Magnet	20
3.2.3	Inner Tracking System	20
3.2.4	Electromagnetic Calorimeter	21
3.2.5	Hadron Calorimeter	23
3.2.6	Muon System	24
3.2.7	Trigger System	26
3.3	LHC Operation and Data Taking	27
3.4	Event Simulation	28
4	Object Reconstruction and Particle Identification	33
4.1	Global Event Description with the Particle-Flow Algorithm at CMS	33
4.2	Reconstruction of Jets	34
4.2.1	Jet Algorithms	34
4.2.2	Jet Types at CMS	37
4.2.3	Jet Transverse-Momentum Response	38
4.2.4	Jet Energy Calibration	39
4.3	Identification of b-Quark Jets	40
4.4	Identification of Boosted t-Quark Jets	42
5	Measurement of the Jet Transverse-Momentum Resolution	47
5.1	Components of the Jet Response	47
5.2	Basic Concept of the Dijet Asymmetry Method	48
5.3	Application to Realistic Collision Events	49
5.4	Samples and Event Selection	50

5.4.1	Datasets and Triggers	50
5.4.2	Selection Criteria	51
5.5	Definition of the Asymmetry Width	55
5.6	Corrections to the Dijet Asymmetry	55
5.6.1	Correction for Additional Jet Activity	55
5.6.2	Correction for Particle-Level Imbalance	58
5.6.3	Results of the Corrections to the Asymmetry	58
5.7	Determination of the Data-to-Simulation Ratio	60
5.8	Validation of the Method	63
5.8.1	Validation in Simulated Events	63
5.8.2	Validation of the Measured Data-to-Simulation Ratio	64
5.9	Systematic Uncertainties	64
5.10	Extension of the Method to the Forward Detector Region	71
5.11	Measurement for simulated events obtained with the HERWIG++ generator	73
5.12	Results	73
5.12.1	Determination of a combined result	73
5.12.2	Comparison to Other Measurements	74
5.13	Adjustment of the MC Resolution to Data	76
5.14	Impact of the Resolution Measurement on other Analyses	76
6	Search for New Physics with Jets and Missing Transverse Momentum	80
6.1	Event Selection	80
6.1.1	Data samples and trigger	80
6.1.2	Event Cleaning	80
6.1.3	Baseline Selection	80
6.1.4	Exclusive Search Regions	80
6.2	QCD Background Estimation with the Rebalance-And-Smear Method	80
6.2.1	Rebalance Procedure using Kinematic Fits	80
6.2.2	Response Smearing	80
6.2.3	Validation Tests	80
6.2.4	Systematic Uncertainties	80
6.2.5	QCD Background Prediction	80
6.3	Estimation of Non-QCD Backgrounds	80
6.3.1	Invisible Z Background	80
6.3.2	Hadronic τ Background	80
6.3.3	Lost-Lepton Background	80
6.4	Results and Interpretation	80
6.4.1	Comparison to Other Measurements	80
6.5	Status of natural supersymmetry after LHC Run-I	80
7	Prospect Studies for a Search for Top Squarks at $\sqrt{s} = 13$ TeV	84
7.1	Data samples	84
7.2	Top Tagging Efficiency Studies	84
7.2.1	Top Tag Efficiency	84
7.2.2	Misidentification Rate	84
7.3	Studies towards a Suitable Analysis Strategy	84
7.4	Results and Discussion	84

7.4.1	Comparison of the Performance of Various Selections	84
7.4.2	Discussion of Specific Simplified Assumptions in the Analysis	84
8	Conclusions	85

1 Introduction

The current knowledge and understanding of the fundamental constituents of matter and occurring interactions between them are summarized in the well-established standard model (SM) of particle physics. The SM which has been introduced in the early 1970's is to date a very successful theory as it was able to predict new particles in the past and is tested to very high precision. Beyond that there are several other fundamental questions unanswered like the origin of dark matter or the unification of coupling constants at a certain scale. One of such theories which goes beyond the standard model and is supposed to provide a solution to such open problems is *supersymmetry* (SUSY). A short introduction to the phenomenology of the standard model as well as supersymmetry can be found in Chapter 1 of this thesis.

In order to address all such questions about tests of the SM and new theories beyond, the CMS experiment located at the Large Hadron Collider (LHC) at CERN¹ has been build. It is designed to analyse particle collisions delivered by the LHC and is the experiment in whichs context this thesis has been performed. An overview of the technical setup of the CMS experiment as well as the current status of its' operation is given in Chapter 2. The general introduction to the CMS experiment is followed in Chapter 3 by an introduction to the reconstruction of certain objects recorded in the particle collisions as well as dedicated algorithms to identify specific decay processes, like e.g. decays of top-quarks.

As the name already indicates, the LHC is a hadron collider and thus the predominant objects to be measured in the detector are expected to be jets – the experimental signature of quarks and gluons. Thus it is very important to have a good understanding of these objects as e.g. a precise knowledge of the jet transverse momentum resolution. This can be measured utilizing events with a momentum balance in the transverse plane, like e.g. $\gamma + \text{jet}$ events, $Z + \text{jet}$ events or dijet events. A measurement of this quantity using dijet event topologies in data recorded by the CMS experiment at a center-of-mass energy of 8 TeV as well as simulated events is a main topic of this thesis and discussed in detail in Chapter 4.

The detailed knowledge about jets and their resolution can afterwards be exploited in a search for new physics mainly targeting signatures of decays involving supersymmetric particles in events with missing transverse energy and several jets. Such searches are especially sensitive to the production of gluinos and first and second generation squarks in supersymmetric models. The main focus regarding this new physics search lies on the estimation of background contributions arising from mismeasured jets and decays of heavy flavour quarks to the final states of interest. An overview of this analysis which is also based on the full dataset recorded by the CMS experiment at a center-of-mass energy of 8 TeV with special emphasis on the QCD multijet background determination is presented in Chapter 5.

Finally, Chapter 6 summarizes a prospect study for searches for supersymmetry at a

¹European Organization for Nuclear Research near Geneva, Suisse

center-of-mass energy of 13 TeV which will be most likely the operation energy of the LHC in 2015 when the next running period starts. In this study, it is investigated how the classical approach of the analysis presented in Chapter 5 can be adapted and improved in order to be sensitive not only to gluino and first and second generation squarks but to the direct production of top squarks. Special emphasis lies on the study of specific kinematic variables and the application of jet substructure tools.

The thesis ends with a short summary and outlook.

2 Theoretical Background

The standard model of particle physics (SM) describes the fundamental particles and interactions between them. It is a theory that successfully predicted the existence of several at that time undiscovered particles. In addition, it has been tested extensively in electroweak precision measurements at LEP.

Although the SM shows so far a remarkably successful performance there are also some open questions which can not be answered within the SM. Thus several theories have been developed to address problems which go beyond the SM. One of such well-motivated extensions is supersymmetry (SUSY) for which however no experimental evidence has been found so far.

After a short introduction to the phenomenology of the standard model including a discussion of specific shortcomings, the basic concepts of supersymmetry are introduced in this chapter. In addition, general concepts of searches for supersymmetry at collider experiments are discussed together with a summary of the current status of the results of such searches which have been performed in the past.

2.1 The Standard Model of Particle Physics

The description of the SM comprises the elementary particles and their interactions [1]. In general, one distinguishes between two types of particles: fermions and bosons. While matter particles are fermions with half-integer spin, the fundamental forces are mediated via bosons carrying integer spin. An overview of the contents of the SM is given in Fig. 2.1 where the particles are denoted together with their interactions¹.

Mathematically the standard model is a quantum field theory where interactions between particles are described via gauge symmetries. The underlying gauge group of the standard model is

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

where $SU(3)$ is the gauge group of the strong force and C indicates that this force acts on the colour charge, $SU(2)$ represents the weak force and L denotes that this force only acts on left-handed fermions, while $U(1)$ represents the electromagnetic force acting on the hypercharge Y .

A brief description of the properties of the particles contained in the SM and the corresponding interactions is given in the following:

Matter Constituents: In the SM, one distinguishes between twelve different fermions being the elementary constituents of matter. For each fermion there exists also an antiparticle which carries the opposite signed quantum numbers.

¹Gravity is not included in the current representation of the standard model and thus it is not discussed in this thesis.

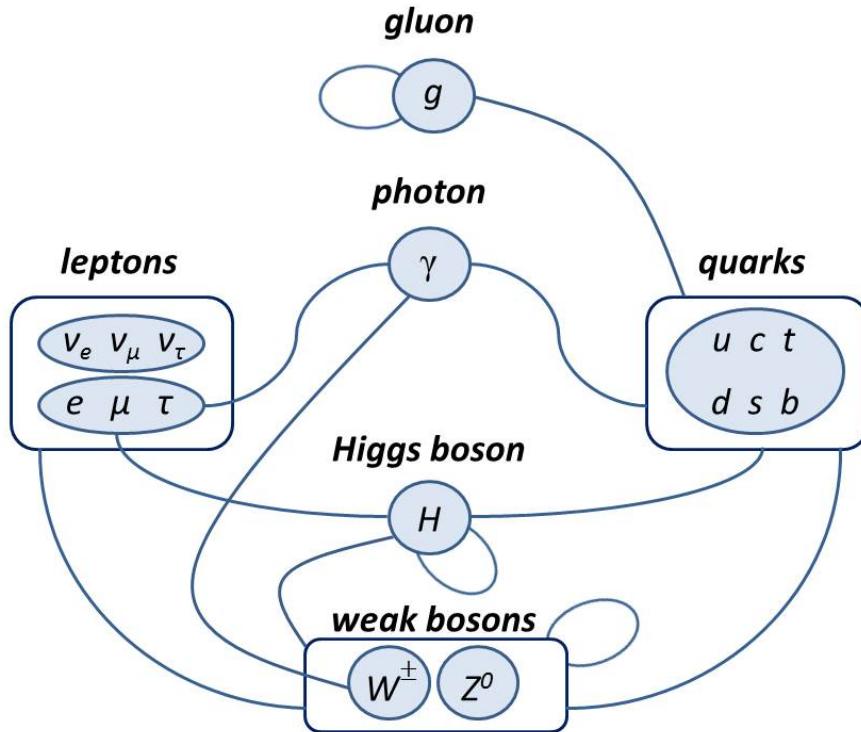


Figure 2.1: Overview of particles contained in the standard model. Blue lines indicate interactions between different particles.

Leptons: The SM contains in total six leptons which are three negatively charged leptons (e, μ, τ) and three neutral leptons (ν_e, ν_μ, ν_τ), the neutrinos. In addition to the charge, leptons are also distinguished according to the lepton numbers which are electron number $L_e = 1$ for electron and electron-neutrino, muon number $L_\mu = 1$ for muon and muon-neutrino and tauon number $L_\tau = 1$ for tauon and tauon-neutrino. Each pair of lepton and neutrino carrying the same lepton number is forming a so-called *generation* where e and ν_e belong to the first generation, μ and ν_μ to the second and τ and ν_τ to the third, respectively.

Quarks: The remaining six fermions in the SM are quarks and can be grouped into generations analogous to the leptons. The first generation is populated by up- and down-quark (u, d), the second by charm- and strange-quark (c, s) and the third by the top- and bottom-quark (t, b). All quarks carry electrical charge but different than for leptons it is not integer but $+2/3$ for the up-type quarks (u, c, t) and $-1/3$ for down-type quarks (d, s, b). Besides to the electrical charge, quarks also carry color charge which comes in three types.

In addition to the attributes described above, fermions are furthermore characterized by the weak isospin. Left-handed fermions in each generation form an isodoublet with a weak isospin of $\pm 1/2$ while right-handed components are isosinglets with a weak isospin of 0.

Fundamental Forces: Matter particles interact with each other through fundamental forces mediated via gauge bosons. These bosons arise from the principle of local gauge invariance under symmetry transformations.

Electromagnetic Force: The description of the electromagnetic force is based on the theory of *Quantum Electrodynamics* (QED). It is communicated between electrically charged particles, like the charged leptons and quarks, by the exchange of photons. These are massless and electrically neutral resulting in the property that the electromagnetic force is long ranged.

Weak Force: The weak force acts on the weak charge which is carried by all fermions in the SM. It is mediated by three vector bosons – namely the two charged W^\pm bosons and the neutral Z boson. These bosons are unlike the photon massive with masses of $W^\pm = 80.385 \pm 0.015 \text{ GeV}$ and $Z = 91.1876 \pm 0.0021 \text{ GeV}$ [1]. As a result the weak interaction is suppressed with respect to the electromagnetic force.

Weak interactions preferably take place within one generation, e.g. e or μ transform into the respective neutrinos when emitting a W^- . However, since the mass eigenstates in the weak interaction differ from the flavour eigenstates, W^\pm bosons can also couple to fermions between different generations. In the quark-sector typically a representation is chosen where the up-type flavour eigenstates correspond to the mass eigenstates and the down-type quarks mix. This mixing is described by the CKM-matrix [2,3]. This is an unitary matrix, described by three mixing angles and one CP-violating phase, which indicates the relative strength between individual transitions.

Strong Force: The theoretical framework describing the strong force is called Quantum Chromodynamics (QCD). It is mediated via eight massless gluons and acts on the colour charge which is carried for instance by quarks. In contrast to the photon which is electrically neutral and thus can not interact with itself, gluons carry a colour charge and hence are able to couple to themselves. The colour charge exists in three different states commonly denoted as 'red', 'green' and 'blue'. Quarks and gluons are collectively referred to as partons.

Regarding the dependence on the distance the strong force behaves differently than other fundamental forces: the coupling strength increases with rising distance. This is a consequence of the different colour states and the self-coupling property of gluons. It is usually referred to as *confinement* [4] and causes the effect that coloured objects can not exist freely. In fact when separated, coloured objects start to build new coloured particles until only a colour neutral formation is left. Such colourless objects linked by the strong force are named hadrons. On the other hand, particles taking part in the strong interaction start to behave quasi-free when they are brought close together. This feature is known as *asymptotic freedom* [5,6].

First proposed by Salam, Glashow and Weinberg [7,8], the electromagnetic and the weak force could be successfully unified into the electroweak force. As denoted earlier, the weak force acts on the weak isospin T_3 while the electromagnetic force acts on the hypercharge Y . These two quantities are connected via the following

relation to the electric charge Q

$$Q = T_3 + Y/2$$

In the electroweak theory, three gauge bosons $W_\mu^{1,2,3}$ are introduced for $SU(2)_L$ and one gauge boson B_μ for $U(1)_Y$. The physical states photon, W^\pm and Z are formed by mixing of these massless states. While the charged W_μ^\pm bosons are superpositions of W_μ^1 and W_μ^2 , the fields A_μ of the photon and Z_μ of the neutral vector boson are obtained by a mixing of the gauge fields W_μ^3 and B_μ parametrized by the weak mixing angle θ_W .

Formel?

Higgs Boson: The electroweak theory in the current representation requires that fermions and bosons are massless particles as mass terms violate the gauge invariance under $SU(2)_L \otimes U(1)_Y$ transformations. This is in contradiction to experimental observations which have shown that all particles, except for photon and gluon, in fact have mass.

An explanation for the generation of particle masses without violation of the principles of the electroweak theory is provided by the *Higgs-mechanism* [9–11] which is based on the concept of spontaneous symmetry breaking. The main idea behind this mechanism is that in general the principle of local gauge invariance is obeyed, but that it is explicitly broken by the ground state. In the context of the Higgs-mechanism this is realized by the introduction of the Higgs field described by a potential with a non-zero minimum value. This non-vanishing minimum represents the expectation value of the quantum field in the vacuum – the vacuum expectation value. The masses of particles are eventually generated by the couplings to the Higgs field.

Formel?

Furthermore, the Higgs-mechanism also predicts the existence of a new heavy boson, the Higgs boson, which is a quantum excitation of one of the components of the Higgs field. It is supposed to be a scalar and the Higgs mass one of the free parameters of the SM.

The discovery of a new boson at a mass of around 125 GeV has been announced by the ATLAS and CMS collaborations in 2012 [12, 13]. As all properties of this new boson are consistent with SM predictions so far (c.f. e.g. [14–17]), this indicates that the last remaining gap of the SM could finally be closed.

2.1.1 Limitations of the Standard Model

Although the SM has been incredibly successful so far and lead to several discoveries while withstanding numerous precision tests, it is known to be on the other hand also an incomplete theory. Some of known shortcomings of the SM are discussed in some detail in the following.

Gravity: As stated already earlier, the SM contains no description of gravity. In particular, it is currently not possible to unify general relativity and quantum theory in one common concept.

Matter antimatter asymmetry: According to the SM matter and antimatter exist to equal amounts in the universe which is in fact not the case. A theory which would be able to explain such an asymmetry needs some source of CP-violation.

The only source of CP-violation within the SM is arising from the CKM matrix as described in 2.1. However, this is not enough to be able to explain the degree of matter antimatter asymmetry in the universe [18].

Unification of couplings: The unification of the electromagnetic and the weak force leads to the question, if it is even possible to further unify the electroweak force with the strong force in order to build a combined theory usually referred to as Grand Unified Theory (GUT). This would imply that the coupling constants of the SM intersect when extrapolating them from the electroweak to the GUT scale. However, this feature is not observed within the SM.

Origin of dark matter and dark energy: There exist several cosmological observations that indicate that the matter described by the SM makes up only 4.9% of the universe [19]. A by far larger part of 26.8% is assigned to so-called *dark matter* which is presumably neutral and only weakly interacting. The only particles within the SM possessing such attributes are neutrinos. However, they are not able to account for the whole relic density present in the universe [20]. Furthermore, for the major part of the universe making up 68.3% there is no hint at all what its' nature could be. Thus it is typically denoted with *dark energy*.

Hierarchy problem: The observable mass of the Higgs boson is given by the bare mass of the Higgs boson plus contributions arising from higher order corrections caused by each massive SM particle. These higher order corrections are usually dependent on an ultraviolet cut-off scale. Conventionally this UV cut-off scale is chosen to be the Planck scale. This would result in a Higgs mass considerably larger than the observed mass of around 125 GeV and requires an enormous amount of fine tuning in order to address this problem.

Formel?

2.2 Supersymmetry

In order to overcome the weaknesses of the SM and to provide explanations for so far unsolved problems of which some have been discussed in 2.1.1, several theories have been developed which go beyond the SM. Among those, a favoured extension is *supersymmetry* (SUSY) as it is able to provide several benefits at once. In this section a brief introduction to the general concept of supersymmetry is given with focus on the *Minimal Supersymmetric Standard Model* (MSSM). For detailed reviews see e.g. [21, 22].

The basic idea of a supersymmetric theory is that a fermionic state is converted into a bosonic state and vice versa by the generator of a supersymmetry transformation Q according to

$$Q |\text{fermion}\rangle = Q |\text{boson}\rangle, \quad Q |\text{boson}\rangle = Q |\text{fermion}\rangle.$$

The supersymmetric fermionic and bosonic partner particles are called *superpartner* and form together the irreducible representations of the supersymmetry algebra named *supermultiplets* with the same number of fermionic and bosonic degrees of freedom. In case of unbroken supersymmetry, partner particles within one supermultiplet have the same mass as well as the same quantum numbers like electric charge, weak isospin and color

degrees of freedom, except for the spin. Commonly, supersymmetric particles are denoted *sparticles*.

Furthermore, in a general supersymmetric theory fulfilling the criteria of gauge invariance and renormalisability, processes are allowed violating either the lepton or baryon number conservation. However, such a process has not been observed experimentally so far and would imply for instance a rapid decay of protons. The current lower limit on the proton lifetime is ... and indicates that such processes must be suppressed. In order to achieve this, a new quantum number called *R-parity* is introduced according to

$$R = (-1)^{3(B-L)+2S}$$

with baryon number B , lepton number L and spin S . It is a multiplicative quantum number and amounts to $R = +1$ for SM particles while it is $R = -1$ for supersymmetric particles. Assuming R-parity conservation no baryon or lepton number violation processes occur ². In addition, the assumption of R-parity conservation leads to further phenomenological implications:

- SUSY particles can only be produced in pairs at collider experiments as only even numbers of supersymmetric particles can occur in an interaction vertex.
- The lightest supersymmetric particle (LSP) is stable and thus any decay chain of a supersymmetric particle finally ends in a state containing an odd number of LSPs.

Quelle

Quelle
RPV
SUSY in
Fussnote

A R-parity conserving supersymmetric theory as described above provides some elegant solutions to open questions as raised in section 2.1.1. As discussed there, the Higgs mass suffers from quadratically divergent contributions arising from higher order corrections caused by SM particles. However, since in SUSY each SM particle gets a supersymmetric partner these higher order corrections cancel as all supersymmetric particles contribute in an unbroken theory with equal size but opposite sign as their corresponding SM partner. Thus SUSY is able to provide a solution to the hierarchy problem. But the observation of such kind of supersymmetric particles with the exact same masses as their SM correspondents would have happened already. Since this is not the case, one knows from all experimental results that supersymmetry in fact has to be a broken symmetry. In order to still be able to provide a solution to the hierarchy problem, supersymmetric particles are expected to be not heavier than $\mathcal{O}(1 \text{ TeV})$. This is the main argument why one would expect supersymmetric particles to be in the TeV range, well within the reach of the LHC. Especially the superpartner of the top quark is expected to be not too heavy in order to be able to cancel the contributions from top loops. This is necessary as these give the largest contributions since the top quark is the heaviest particle in the SM.

In addition to this naturalness aspects, the coupling constants of the forces which do not intersect within the SM actually meet in one point when extrapolating the couplings from the electroweak to the GUT scale considering the existence of supersymmetric particles. This effect is illustrated in Fig. 2.2. It is visible that the evolution of the couplings is modified with respect to the SM at that energy scale where the supersymmetric particles enter. In general, this hints to the possibility of a grand unification.

²There exist also several R-parity violating SUSY models which are not in contradiction to the observed proton lifetime. However, these are not subject of this thesis and thus not discussed.

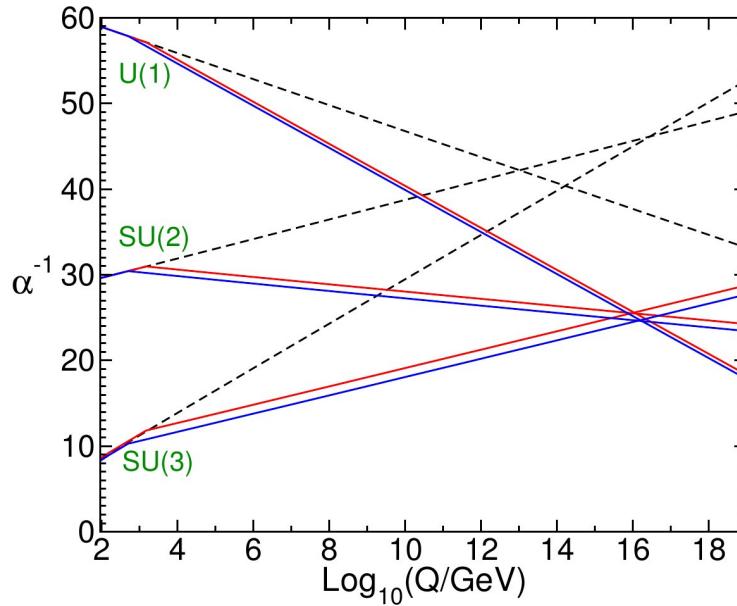


Figure 2.2: Comparison of the renormalization group evolution of the couplings α_a^{-1} in the SM (dashed lines) and the MSSM (solid lines) including two-loop effects. The masses of the supersymmetric particles in the MSSM are considered as a common threshold changing between 500 GeV and 1.5 TeV while $\alpha_3(m_Z)$ is varied between 0.117 and 0.121. Taken from [22].

Besides the two advantages discussed above, R-parity conserving SUSY models also provide a suitable dark matter candidate. As discussed previously, each decay of supersymmetric particles finally leads to the existence of an LSP which can not decay further. Thus, it is an adequate DM candidate when it is electrically uncharged and only weakly interacting.

2.2.1 The MSSM

The smallest possible supersymmetric extension of the SM including the full particle spectrum and interactions is comprised in the MSSM. An overview of the respective particle content is given in Tab. 2.2.1. This extension is minimal in that sense that it introduces the least feasible number of additional particles to the existing SM particles. Thus, each SM fermion gets an individual superpartner while the interactions and couplings of the supersymmetric particles are the same as for the SM counterparts. The different transformation of left- and right-handed fermions under the gauge groups implies the necessity to introduce separate superpartners for left- and right-handed states as well. These are arranged with their bosonic (spin 0) superpartner in a *chiral* supermultiplet. The labels indicating the left- and right-handed states refer to the helicity of the respective SM particle. These supersymmetric partners of fermions are named *sfermions* distinguishing between *sleptons* and *squarks* – the supersymmetric partners of leptons and quarks. In a similar manner the SM gauge bosons are arranged in *gauge* supermultiplets together with their fermionic (spin 1/2) supersymmetric correspondents. The SUSY partners of the gauge

Type	Spin	Gauge eigenstates	Mass eigenstates
Higgs bosons	0	$H_u^0 \ H_d^0 \ H_u^+ \ H_d^-$	$h^0 \ H^0 \ A^0 \ H^\pm$
Squarks	0	$\tilde{u}_L \ \tilde{u}_R \ \tilde{d}_L \ \tilde{d}_R$	see left
		$\tilde{s}_L \ \tilde{s}_R \ \tilde{c}_L \ \tilde{c}_R$	see left
		$\tilde{t}_L \ \tilde{t}_R \ \tilde{b}_L \ \tilde{b}_R$	$\tilde{t}_1 \ \tilde{t}_2 \ \tilde{b}_1 \ \tilde{b}_2$
Sleptons	0	$\tilde{e}_L \ \tilde{e}_R \ \tilde{\nu}_e$	see left
		$\tilde{\mu}_L \ \tilde{\mu}_R \ \tilde{\nu}_\mu$	see left
		$\tilde{\tau}_L \ \tilde{\tau}_R \ \tilde{\nu}_\tau$	$\tilde{\tau}_1 \ \tilde{\tau}_2 \ \tilde{\nu}_\tau$
Neutralinos	1/2	$\tilde{B}^0 \ \tilde{W}^0 \ \tilde{H}_u^0 \ \tilde{H}_d^0$	$\tilde{\chi}_1^0 \ \tilde{\chi}_2^0 \ \tilde{\chi}_3^0 \ \tilde{\chi}_4^0$
Charginos	1/2	$\tilde{W}^\pm \ \tilde{H}_u^\pm \ \tilde{H}_d^\pm$	$\tilde{\chi}_1^\pm \ \tilde{\chi}_2^\pm$
Gluino	1/2	\tilde{g}	see left
Gravitino	3/2	\tilde{G}	see left

Table 2.1: Supersymmetric particles contained in the MSSM neglecting mixing in the first two sfermion generations. Following [22].

bosons are named *gauginos* so that the superpartners in the gauge supermultiplets are the *gluino*, *wino* and *bino*. The corresponding gaugino mixtures of the neutral wino and the bino are the *photino* and the *zino*. Beyond that, the supersymmetric particle spectrum is extended by another supermultiplet containing the graviton (spin 2) and the respective supersymmetric partner – the *gravitino* (spin 3/2).

As described in Sec. 2.1, masses arise in the SM from the concept of spontaneous symmetry breaking implying the existence of the Higgs boson. The supersymmetric partner of the Higgs boson is named *higgsino*. While in the SM one Higgs doublet is sufficient to give mass to all particles, the Higgs sector needs to be extended in the MSSM. Here, two Higgs doublets are introduced where one doublet H_u gives mass to the up-type quarks and the other one H_d to the down-type quarks, respectively. These two doublets have together eight degrees of freedom of which three are needed to give mass to the gauge bosons of the weak interaction as in the SM. This results in five physical Higgs bosons which are the two scalar Higgs particles h^0, H^0 , the pseudoscalar A^0 as well as the charged Higgs bosons H^\pm . As further consequence there are two vacuum expectation values v_u and v_d present – each assigned to one Higgs doublet – whose ratio $\tan\beta = v_u/v_d$ is a free parameter of the model.

As in the SM, the gauge eigenstates of the theory do not necessarily have to be also the mass eigenstates. A mixing occurs especially in the gaugino sector. Here, the neutral components of the bino and wino mix with the neutral higgsinos and form four mass eigenstates – the *neutralinos* $\tilde{\chi}^0$. Similarly, also the charged gauginos and higgsinos mix to the four *charginos* $\tilde{\chi}^\pm$. Furthermore, mixing can also appear in the third squark and slepton generation. The mixing is supposed to be significant only for fermions of the third generation as the off-diagonal elements of the sfermion mass matrix are proportional to the mass of the respective SM partner.

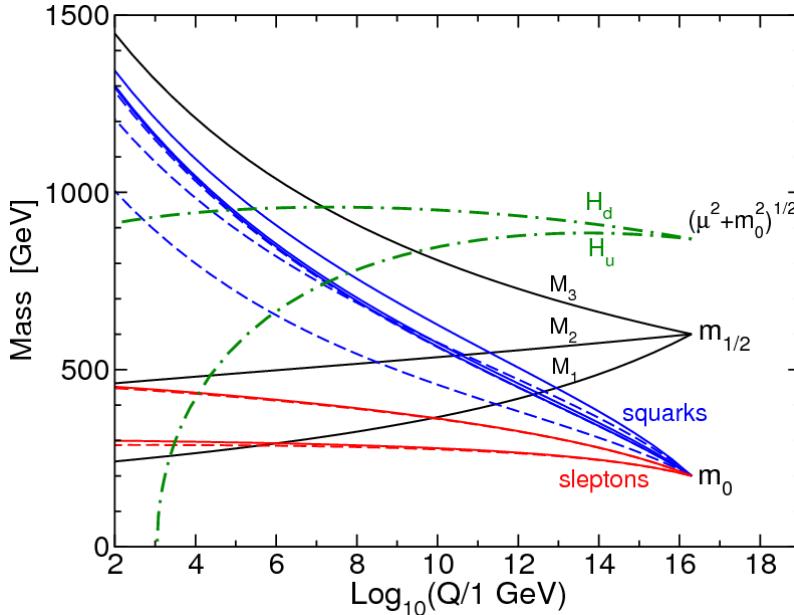


Figure 2.3: Evolution of scalar and gaugino mass parameters in the MSSM with mSUGRA boundary conditions imposed at $Q_0 = 2 \times 10^{16} \text{ GeV}$. The parameter $\mu^2 + m_{H_u}^2$ runs negative, provoking electroweak symmetry breaking. Taken from [22].

2.2.2 SUSY-Breaking

As discussed above, SUSY particles are expected to have the same mass as their corresponding SM partner particle. However, none of such particles has been observed so far which implies that supersymmetry in fact must be broken and that sparticles are actually heavier than the SM counterpart. If SUSY is expected to provide a solution to the hierarchy problem, the mass difference should consequently not be too large in order to still allow a cancellation of loop contributions from SM particles to the Higgs mass. The characteristic mass scale should be around 1 TeV [22].

Typically, SUSY breaking is introduced such that so-called *soft breaking* terms are added to the theory in addition to the terms determining the gauge and Yukawa couplings. Since the MSSM does not contain such violating terms, it is extended assuming that the SUSY breaking occurs in a hidden sector which does not couple directly to the visible sector represented by the supermultiplets. Thus, a specific breaking mechanism must be assumed in order to describe the mediation of the supersymmetry breaking from the hidden to the visible sector. In total, the MSSM features several new phases, mixing angles and masses which add another 105 free parameters to the already existing parameters of the SM [23].

The two most popular SUSY breaking scenarios are either based on gravity-mediated or gauge-mediated interactions and known as *minimal supergravity* (mSUGRA) [24, 25] or *constrained MSSM* (CMSSM) [26, 27] and *gauge-mediated supersymmetry breaking* (GMSB) [28, 29]. Assuming a specific breaking scenario usually allows to drastically reduce the number of free parameters in the theory and determines the phenomenology of the respective model. In case of mSUGRA/CMSSM the whole model can be described by five

parameters which are the common scalar mass m_0 and the common mass of the gauginos and higgsinos $m_{1/2}$ at the GUT scale, the common trilinear coupling A_0 , $\tan\beta$ and the sign of the higgsino mass parameter μ . In Fig. 2.3 the evolution of the corresponding mass parameters to the electroweak scale are illustrated.

2.3 Searches for Supersymmetry and Current Constraints

The appealing attributes of supersymmetry which have been discussed above initiated a couple of indirect and direct searches looking for hints of supersymmetric particles. Although no sign for SUSY has been observed in nature so far, several results have been used in order to constrain the allowed parameter space. This section gives a short introduction to general search strategies for supersymmetry with main focus on collider experiments and provides a short overview of the status of experimental results. Results of direct searches at the LHC discussed here are restricted to results from searches performed previous to the analysis presented in this thesis. Further discussion of the status of supersymmetry after LHC Run-I follows in Sec. 6.5.

2.3.1 Constraints from Indirect Measurements

electroweak precision data, g-2, astrophysical observations, b-physics, CP-violation, FCNC

2.3.2 Direct Searches at Collider Experiments

Although the exploitation of indirect searches for supersymmetry is very useful in order to constrain the allowed SUSY parameter space, the most stringent exclusion limits are derived from direct searches at collider experiments. Typically, searches for SUSY at colliders make use of the specific production and decay properties which are caused by the assumption of R-parity conservation. These imply that sparticles are only produced in pairs and decay via cascades into the lightest supersymmetric particle which is often assumed to be the lightest neutralino. This leaves the experiment undetected and manifests in missing energy or missing momentum. Such missing energy signatures are thus a key-feature of searches for supersymmetry in collider experiments. Depending on the type of supersymmetric particle produced the missing energy can be accompanied by several leptons, photons or jets. Usually, searches are classified according to their targeted final state and are aimed at a specific kind of supersymmetric particle. If searches are designed quite generic in order to be sensitive to various types of particles and models, they are called *inclusive* searches.

Extensive searches resulting in the tightest exclusion limits in the pre-LHC era have been realised by the experiments performed at HERA, LEP and Tevatron:

HERA [30,31]: At HERA searches for R-parity conserving supersymmetric models have been mainly targeting processes involving the production of a selectron and a squark. The dominant MSSM process at HERA is the production of a selectron and a squark via the exchange of a neutralino where \tilde{e} and \tilde{q} subsequently decay into any lighter gaugino and their respective SM partner. Thus, a distinct experimental signature is given by an electron, hadrons and missing energy and momentum. Since no excess over the SM expectation has been observed, exclusion limits for the masses of the

selectron and squark in the context of the MSSM have been derived. Depending on the specific choices made to constrain the MSSM parameter space, the excluded region extends to $\approx 60 - 70$ GeV for squark masses and ≈ 40 GeV for the LSP mass.

LEP [32–36]: At LEP various searches for supersymmetry have been performed targeting different species of supersymmetric particles. Scalar leptons and quarks are mainly pair-produced in the s-channel via γ or Z exchange. In case of selectrons also the t-channel exchange of neutralinos yields important contributions. The typical energy scale at LEP results in decays into the respective SM partner particles – except for the scalar top – and the lightest neutralino while also cascade decays are possible. Typical final states contain missing energy and a pair of acoplanar leptons or jets where the direction of one of the leptons/jets is not in the plane defined by the direction of the second lepton/jet and the beam direction. Similarly, neutrinos and charginos are expected to be pair-produced via Z and Z/γ s-channel exchange or t-channel selectron or sneutrino exchange, respectively. Typically, the lightest neutralino and fermion pairs are the final decay products of such events. Hence, the final state in case of chargino production is characterized by missing energy accompanied by four jets, two jets and one lepton or only leptons depending on the specific decay mode of the chargino while the most important signature for neutralino production are acoplanar pairs of jets or leptons coming along with large missing momentum. However, the exact decay topologies strongly depend on the particular mass spectrum of the supersymmetric particles so that the above mentioned topologies could also be accompanied by photons or manifest in multiple jets or leptons in cascade decays. Interpretations of combined results from all four experiments in the mSUGRA model lead to exclusion limits showing that $m_{1/2}$ has to be greater than $\approx 100 - 200$ GeV over a wide range of m_0 up to the TeV-region for specific fixed other parameters. The lower limit on the LSP mass is found to be around 50 GeV nearly independent of other model constraints.

Tevatron [37–39]: The Tevatron accelerator was the best place to search for supersymmetry before the actual start of the LHC, as the center of mass energy exceeded that of HERA and LEP by several orders of magnitude. Since sparticles are expected to be heavy, thus a rich program of supersymmetric searches was enabled covering various final states of different lepton, photon or jet multiplicities. Of special interest is the search for colored sparticles like squarks and gluinos, as Tevatron is a hadron collider. Squarks are expected to decay preferably into a quark and the LSP, if they are heavier than the gluinos, resulting in an acoplanar-jet topology with missing transverse momentum. If gluinos are lighter than squarks, they will predominantly decay into a quark pair and the LSP producing a multijet final state accompanied by missing transverse momentum. Thus, searches for squarks and gluinos at Tevatron have been performed exploiting multijet and missing momentum final states. Results have been interpreted in the context of mSUGRA and extended the LEP results in the parameter region of $m_0 = 70 - 300$ GeV and $m_{1/2} = 125 - 165$ GeV. Thus, gluinos below $\approx 280 - 300$ GeV for all squark masses and squarks below 380 GeV independent of the gluino mass could be excluded.

The absence of SUSY-like signals at any collider experiment performed previous to the start of the LHC and exclusion limits at the order of a few hundred GeV on sparticle

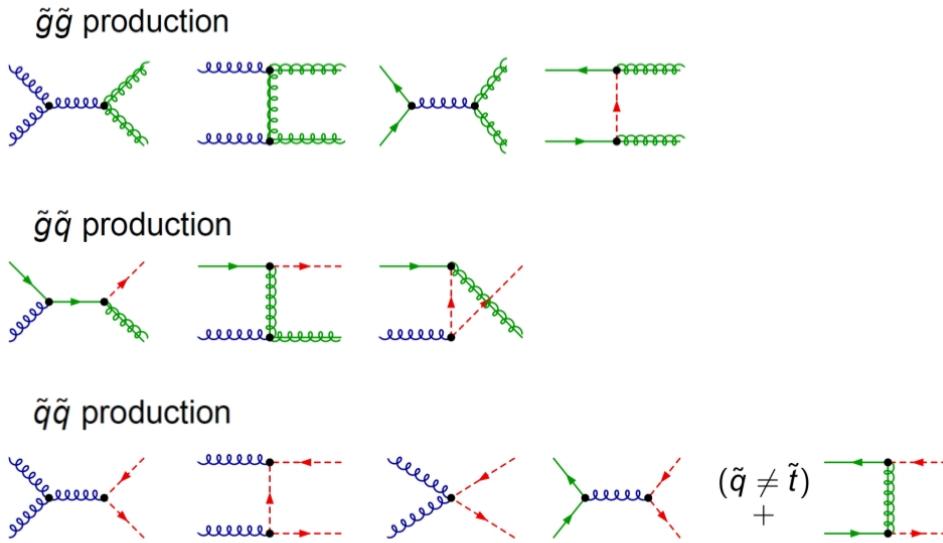


Figure 2.4: Example diagrams for the production of supersymmetric particles in hadron collisions at parton level.

masses activated an abundant amount of searches for supersymmetry at the LHC as it opened already with a center of mass energy of 7 TeV an even larger phase space than the Tevatron before. At the LHC a variety of searches targeting the various different production and decay modes of supersymmetric particles are performed. In general, they can be classified into searches for electroweakinos, third generation sfermions and searches for squarks and gluinos. Of particular interest are searches for colored particles as the LHC is as well as Tevatron a hadron collider.

At leading order sparticles in R-parity conserving models are predominantly produced in processes like [40–44]

$$pp \rightarrow \tilde{g}\tilde{g}, \tilde{g}\tilde{q}, \tilde{q}\tilde{q}. \quad (2.1)$$

Some example diagrams for the production modes of such processes at parton level are shown in Fig. 2.4. Typically, the squarks are assumed to be mass degenerate and refer to the partners of the light-flavour (u, d, s, c, b) quarks. Supersymmetric partners of the top quark are considered separately due to a potentially large mixing affecting the mass splitting. Most recent SUSY cross section calculations consider higher order corrections caused e.g. by quark radiation or gluon loops typically up to next-to-leading order (NLO) [45].

- decay
- backgrounds
- 7 TeV searches

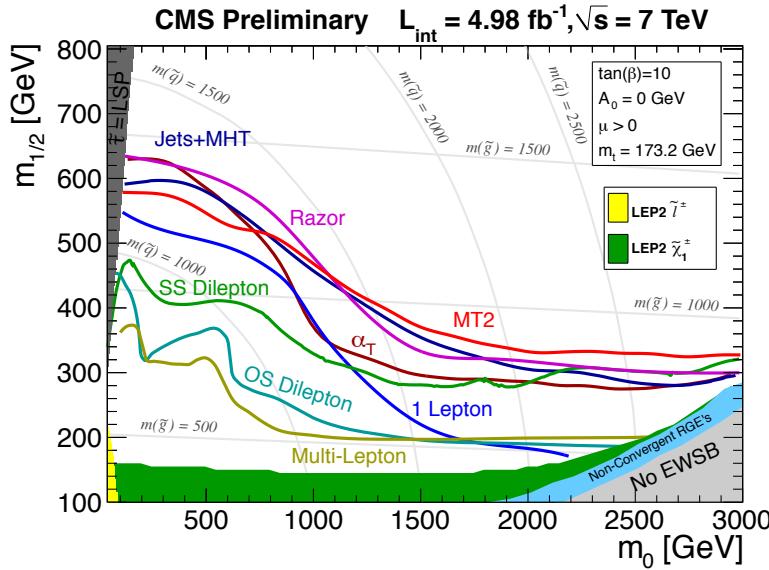


Figure 2.5: Interpretation of searches for supersymmetry at the CMS experiment within the CMSSM. Shown are the 95% CL exclusion limits in the m_0 and $m_{1/2}$ plane for various searches performed using different final state topologies [46].

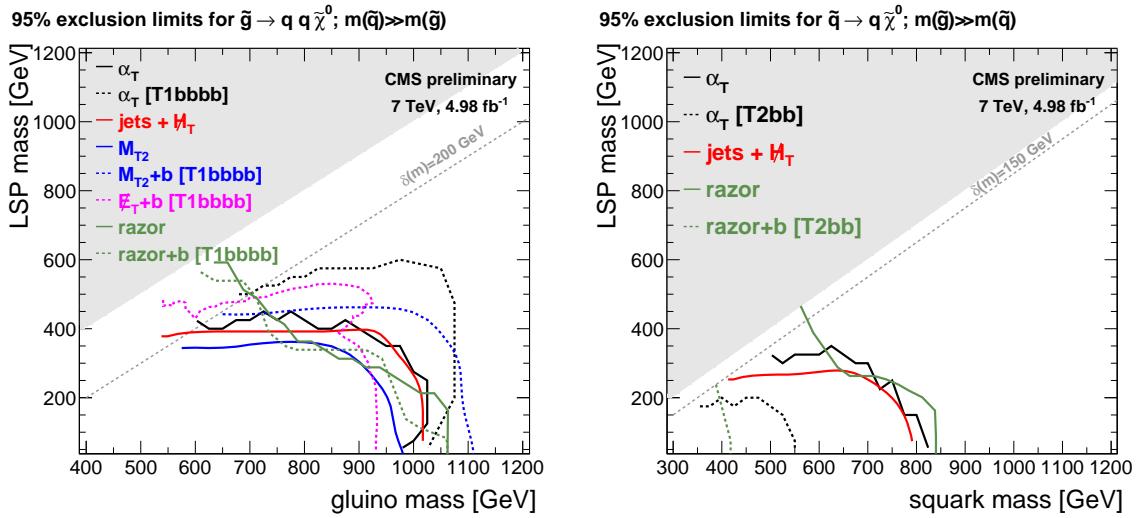


Figure 2.6: Interpretation of searches for supersymmetry at the CMS experiment with simplified models in $\tilde{g} \rightarrow q\bar{q}\tilde{\chi}_0^0$ (left) and $\tilde{q} \rightarrow q\tilde{\chi}_0^0$ (right) topologies. Shown are the 95% CL exclusion limits on the produced particle and LSP masses. The grey area represents the region where the respective decay mode is forbidden [47].

3 Experimental Setup

In order to probe the various aspects of the well-established standard model or search for hints of new physics beyond the SM, particle physics experiments preferentially make use of powerful particle accelerators where particles of a certain type are collided in order to probe the constituents of matter and interactions between them. The analyses presented in this thesis are all performed in the context of the CMS experiment located at the Large Hadron Collider (LHC) at CERN near Geneva.

The first part of this chapter provides an introduction to the LHC. This is followed by an overview of the detector system of the CMS experiment. Afterwards the hitherto periods of collision data taking at the LHC are discussed together with an introduction to the generation of simulated events which are used in the analysis of real data events.

3.1 The Large Hadron Collider

The Large Hadron Collider [48, 49] is a ring-accelerator designed to provide particle collisions of hadrons. It is built in the tunnel of the former LEP [50] collider 45 – 170 m below the ground and has a circumference of 26.7 km. The LHC is a particle-particle collider and thus composed of two rings with counter-rotating beams. The operation can be performed in different modes with either proton beams or heavy ions like e.g. lead ¹.

In each beam, protons are grouped together in bunches and accelerated in two evacuated beam pipes using superconducting radio-frequency cavities. With a nominal bunch spacing of 25 ns the bunch revolution frequency is 40 MHz. Each of the 2808 individual bunches per beam contains at design conditions 1.15×10^{11} protons. In order to bend the beams around the LHC ring superconducting dipole magnets are used with an operation temperature of 1.9 K. They provide a magnetic field of up to 8.33 T while additional quadrupole and sextupole magnets are utilized to squeeze and focus the beams.

Before the protons are injected into the LHC they are already pre-accelerated in various smaller accelerators up to a beam energy of 450 GeV while passing through the injector chain Linac2 – Proton Synchrotron Booster (PSB) – Proton Synchrotron (PS) – Super Proton Synchrotron (SPS). An overview of the accelerator complex at CERN is given in Fig. 3.1.

The main goal of the LHC is to provide proton-proton collisions to the experiments with center of mass energies up to 14 TeV in order to explore physics processes at novel energy regimes. The expected number of events N for a certain type of process is given by the product of the specific cross section σ of that process and the integral $L = \int \mathcal{L} dt$ of the instantaneous luminosity \mathcal{L} over time such that

$$N = \sigma \cdot L. \quad (3.1)$$

¹ All studies presented in this thesis are based on proton-proton collisions. Thus the operation with heavy ions is not discussed.

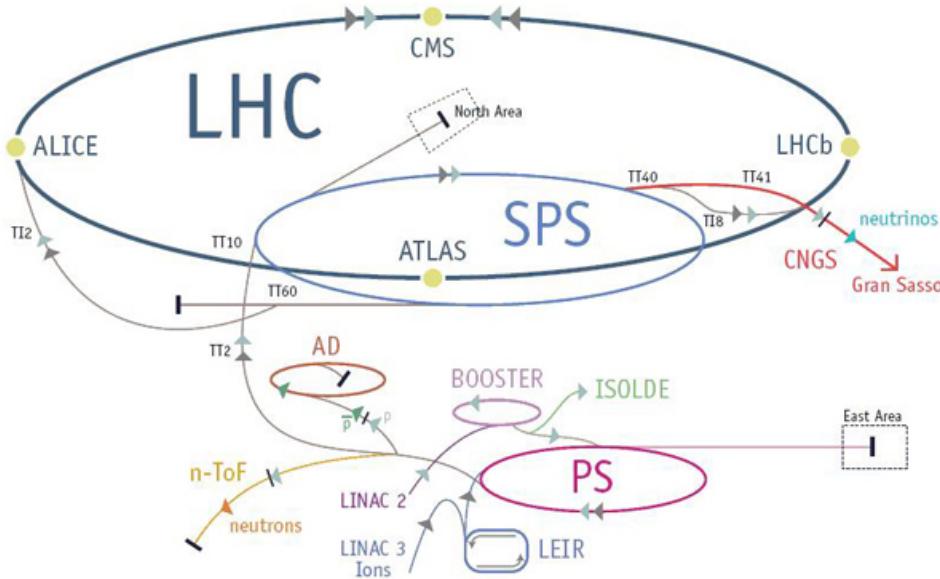


Figure 3.1: Illustration of the CERN accelerator complex including the injector chain of the LHC ring. Taken from [51].

The luminosity is a machine parameter and can be expressed for beams with Gaussian-shaped profiles as

$$\mathcal{L} = \frac{f n_1 n_2}{4\pi\sigma_x\sigma_y} \cdot F \quad (3.2)$$

with the revolution frequency f , the number of particles n_1 and n_2 contained in the two colliding bunches and the transverse beam sizes σ_x (σ_y) in the horizontal (vertical) directions. In order to take the inclination of the two beams into account, the geometrical correction factor F is introduced. The nominal peak luminosity of the LHC is $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. Since the total inelastic proton-proton cross-section at a center of mass energy of 14 TeV is close to 100 mb as indicated in Fig. 3.2, the expected event rate is approximately 10^9 events per second. This is resulting in high technical challenges for the experiments.

The four main experiments are located at the four locations along the LHC ring where the beams cross in order to measure the delivered particle collisions. The two high luminosity experiments ATLAS [53] and CMS [54, 55] are designed for multiple purposes like precision measurements of SM quantities, search for the standard model Higgs Boson or searches for signals indicating new physics processes. The LHCb detector [56] however is a specialised experiment focusing on the measurement of CP violation in the interactions of hadrons containing b-quarks. The only experiment designed especially for the analysis of heavy ion collisions is the ALICE [57] detector with the main emphasis on the physics of strongly interacting matter at extreme energy densities like for instance quark-gluon plasma.

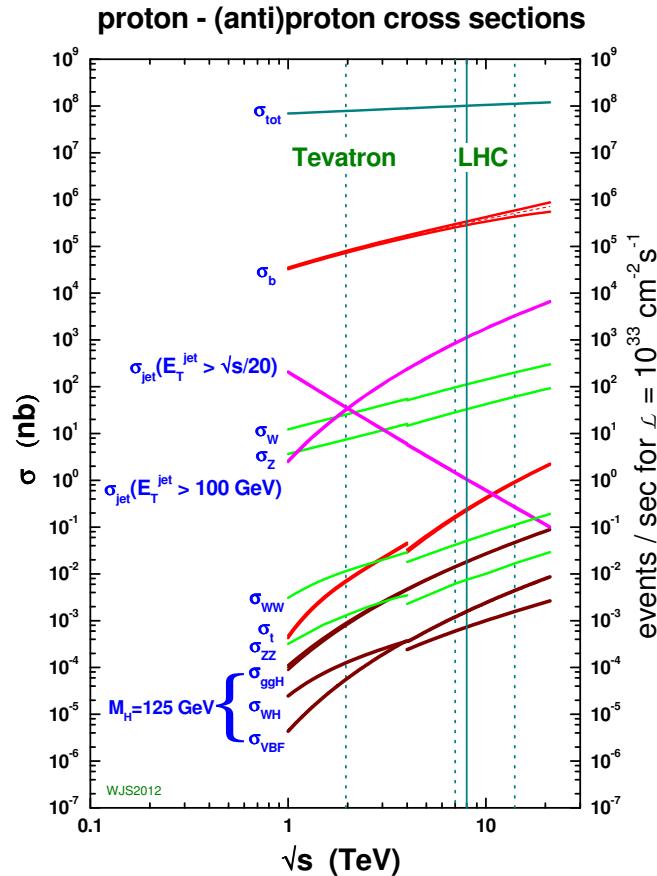


Figure 3.2: Summary of cross sections for various standard model processes in proton-proton and proton-antiproton collisions as function of the centre of mass energy. The right axis displays the corresponding event rate at a luminosity of $10^{33} \text{ cm}^{-2}\text{s}^{-1}$. Taken from [52].

3.2 The CMS Experiment

The CMS detector is one of the two experiments at the LHC designed to address a multitude of physics questions. In addition to tests of the SM at the TeV scale, studies of the nature of electroweak symmetry breaking which might show up in the presence of a Higgs boson and searches for so far unknown particles pointing to e.g. new symmetries in nature are the primary targets of these experiments. These ambitious physics goals can only be achieved by fully exploiting the by now unprecedented collision energy and luminosity.

The CMS detector with its typical cylindrical design of different sub-detector components around the beam line is designed to perfectly meet these particular conditions. A sketch of the CMS detector and the different sub-detectors is shown in Fig. 3.3. As a typical high-energy particle experiment the CMS detector makes mainly use of tracking

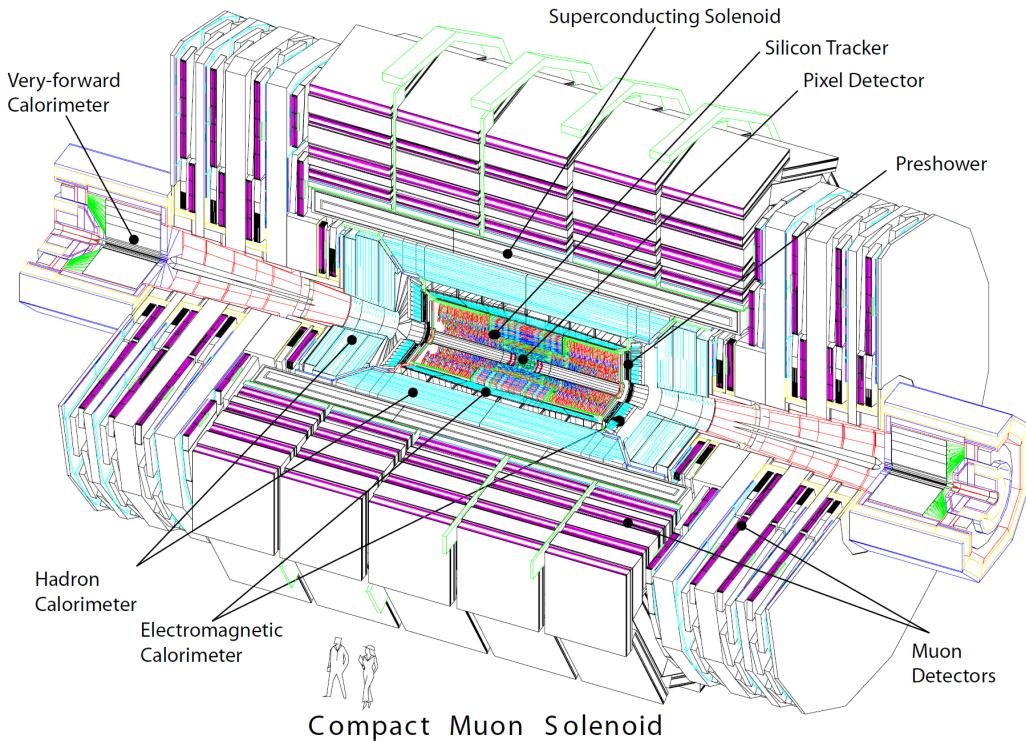


Figure 3.3: A perspective view of the CMS detector [54].

detectors and calorimeters to measure particles' momenta, energy depositions and flight directions in order to identify the objects emerging from the particle collisions. The entire CMS detector with a length of 21.6 m and a diameter of 14.6 m results in a total weight of 12500 t.

The following sections comprise a description of the CMS detector and individual sub-detector components focusing on the detector parts most relevant for the analyses presented in this thesis. A detailed discussion of the detector design can be found in [54, 55].

3.2.1 Coordinate Conventions and Kinematic Variables

In order to describe the particle collisions, the CMS experiment makes use of a right-handed coordinate system with its origin at the center of the detector at the nominal interaction point. While the z-axis is defined along the direction of the beam, the x-axis points to the center of the LHC ring and the y-axis vertically upwards. In this xy-plane the azimuthal angle ϕ is measured where $\phi = 0$ coincides with the x-axis. The polar angle θ however is defined with respect to the z-axis. A quantity closely related to the polar angle is the pseudorapidity η defined as

$$\eta = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right] \quad (3.3)$$

which is widely used in experimental particle physics as rapidity differences are Lorentz invariant. A pseudorapidity $\eta = 0$ corresponds to the direction perpendicular to the beam while $|\eta| \rightarrow \infty$ points along the beam. Based on the pseudorapidity the Lorentz invariant distance between two objects ΔR can be written as

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \quad (3.4)$$

At the LHC the initial conditions of the primary collisions are not known as the specific energy fraction of the proton which each parton carries can not be identified. Thus conservation of the total momentum can not be utilized directly to describe the momentum balance in the final state. However, it is known that the initial particles have no significant momentum orthogonal to the beam axis which is referred to as transverse momentum

$$p_T = p \cdot \sin\theta. \quad (3.5)$$

Thus, momentum conservation in the transverse plane is used to describe the final state conditions. Any difference between the total sum of all transverse momenta and zero is considered as missing energy E_T and often exploited to describe undetected particles.

3.2.2 Superconducting Magnet

The CMS experiment makes use of a large superconducting solenoid magnet which is a crucial component of the whole detector design and provides a magnetic field of up to 4 T. It allows to precisely determine the momenta and charge of charged particles from the bended tracks that they follow in the magnetic field.

With a length of 12.5 m and a diameter of the free bore of 6.3 m the total cold mass reaches 220 t. It is made up of a niobium-titanium coil which is winded in 4-layers. This configuration allows a storage of 2.6 GJ energy at full current. A 10000 t heavy-weight iron yoke is responsible for the return of the magnetic flux.

3.2.3 Inner Tracking System

The tracking system of the CMS experiment is the innermost part of the detector and installed directly around the interaction point completely contained in the bore of the magnet system. Its' purpose is to precisely measure the trajectories of charged particles arising from the collisions as well as to reconstruct secondary vertices. Due to the location close to the interaction point the tracking system has to cope with a high particle flux crossing the tracker associated with each bunch crossing. Hence high requirements on response time and granularity are set in order to properly identify the particles' tracks. In order to fulfill these tasks the CMS experiment makes use of a tracker design based on silicon detectors. It consists of mainly two components: the innermost part is made of silicon pixel detectors while these are surrounded by silicon strip modules. In total they add up to an active area of 200 m² with a length of 5.8 m and a diameter of 2.5 m covering the detector up to $|\eta| = 2.5$. A schematic overview of the whole tracking system is shown in Fig. 3.4.

Pixel Detector: The pixel detector consists of three barrel layers which extend from 4.4 cm to 10.2 cm and two endcap disks on each side. In total there are 1440 pixel modules installed. The size of one pixel cell is 100 x 150 μm^2 providing similar track

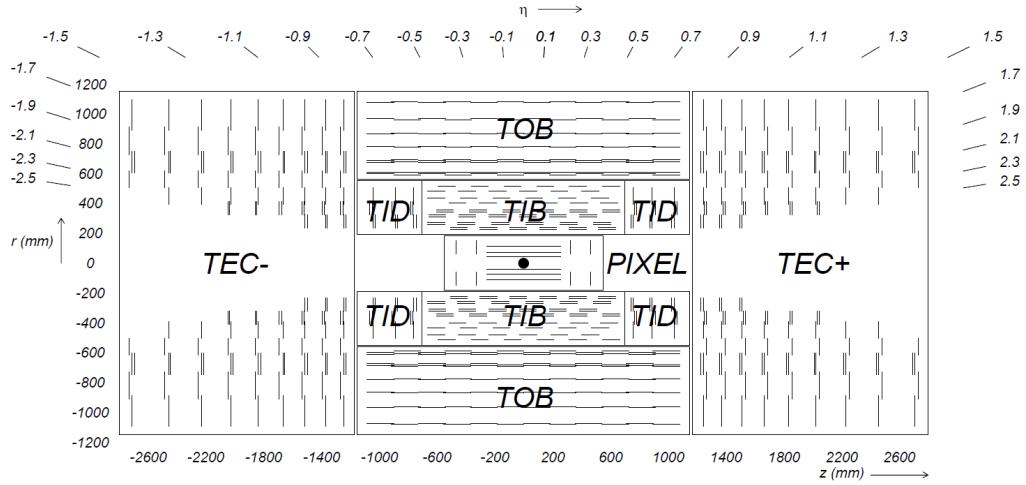


Figure 3.4: Sketch of the CMS tracking system in a r - z -view. Each tracker module is represented by one line. Taken from [54].

resolution quality in $r - \phi$ and z direction. This configuration provides for almost the whole range up to $|\eta| = 2.5$ three precise tracking hits. This is especially important for the reconstruction of secondary vertices.

Silicon Strip Tracker: The silicon strip detector which extends to a radius of 1.1 m comprises the pixel tracker. The more than 15000 individual strip detector modules are arranged in an inner and an outer detector part. The inner part of the strip tracker is build by the four Tracker Inner Barrel (TIB) layers which are accompanied by the three Tracker Inner Disks (TID) at the end sides. This inner part provides up to four track measurements in the $r - \phi$ plane. The TIB/TID system lies within the Tracker Outer Barrel (TOB) consisting of another six barrel layers while it is complemented by the Tracker EndCaps (TEC) which add another nine disks at each side of the tracking system. This layout provides at least around nine hits within the silicon strip system.

The tracking system with the design described above provides a very good impact parameter resolution and tracking efficiency [58]. The impact parameter resolution is of the order of $< 35 \mu\text{m}$ in the plane perpendicular to the beam (for particles with $p_T > 10 \text{ GeV}$) and reaches $75 \mu\text{m}$ in the longitudinal direction. Also the track reconstruction efficiency is expected to show a very good performance. The track reconstruction efficiency of high energetic electrons is above 90%, that of charged hadrons up to 95% (for $p_T > 10 \text{ GeV}$) and that for muons even better than 98% in the whole covered region up to $|\eta| = 2.5$ already for muons with very low transverse momenta around 1 GeV. Altogether, the relative transverse momentum resolution reaches a level of 1-2% for high momentum tracks ($\approx 100 \text{ GeV}$) in the central barrel for $|\eta| < 1.6$.

3.2.4 Electromagnetic Calorimeter

The CMS experiment makes use of a homogeneous electromagnetic calorimeter (ECAL) in order to precisely measure the energy deposits of electrons and photons. It is installed

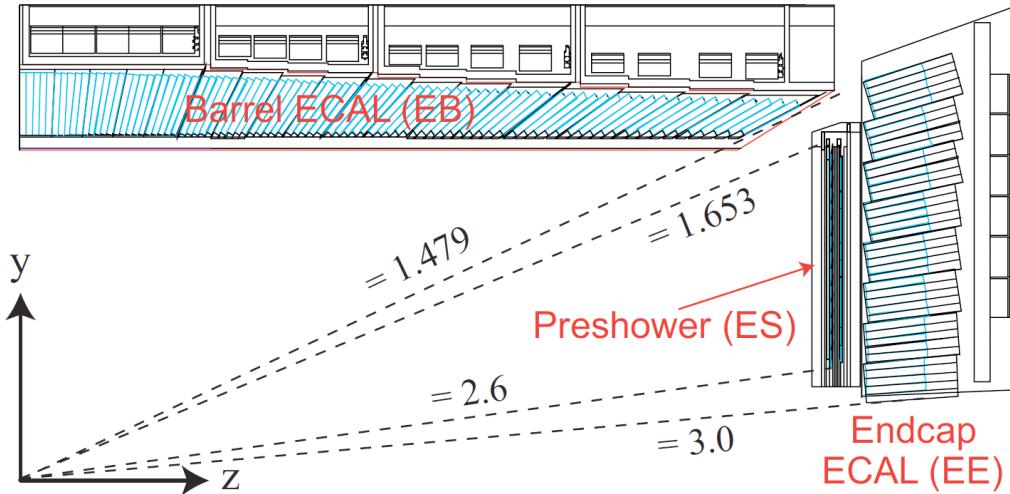


Figure 3.5: View of one quarter section of the CMS electromagnetic calorimeter in a yz -view. Taken from [55].

around the inner tracking system covering a range up to $|\eta| = 3.0$ and consists of lead tungstate (PbWO_4) crystals. These have been chosen as they provide a high density, short radiation length and a small Molière radius and hence allow to build a compact calorimeter with a fine granularity. As 80% of the scintillation light is emitted within 25 ns, this matches well the bunch crossing rate of the LHC machine. In order to collect the radiated light photodiodes are glued to the back of each crystal.

An overview of the ECAL layout is shown in Fig. 3.5. The individual sub-components are the following:

Barrel ECAL (EB): The barrel detector of the ECAL covers the pseudorapidity region up to $|\eta| = 1.479$. Within a radius of about 1.3 m a total number of 61200 crystals are installed. Each of them has a length of 230 mm resulting in a radiation length of 25.8 X_0 . The crystal cross-section in (η, ϕ) is $(0.0174, 0.0174)$. Avalanche photodiodes are used to detect the emitted scintillation light.

Endcap ECAL (EE): The EB is complemented on each side by an endcap which consists of two d-shaped halfs. The ECAL endcaps extend from $|\eta| = 1.479$ to $|\eta| = 3.0$. In total they contain another 14648 crystals with an individual length of 220 mm corresponding to 24.7 X_0 . For the collection of scintillation light vacuum phototriodes are used in the endcaps.

Preshower (ES): In front of the endcap crystals a preshower detector is placed. It covers the pseudorapidity range of $|\eta| = 1.653 - 2.6$. The main purpose is to identify photons emerging from the decay of neutral pions. It is a two-layer sampling calorimeter with lead as absorber material and silicon strip sensors measuring the deposited energy. The total thickness of the preshower is 20 cm (3 X_0).

In order to achieve a stable and uniform energy measurement across the whole ECAL an accurate calibration is of crucial importance. In addition to the calibration of the absolute energy scale especially channel-to-channel effects – referred to as intercalibration –

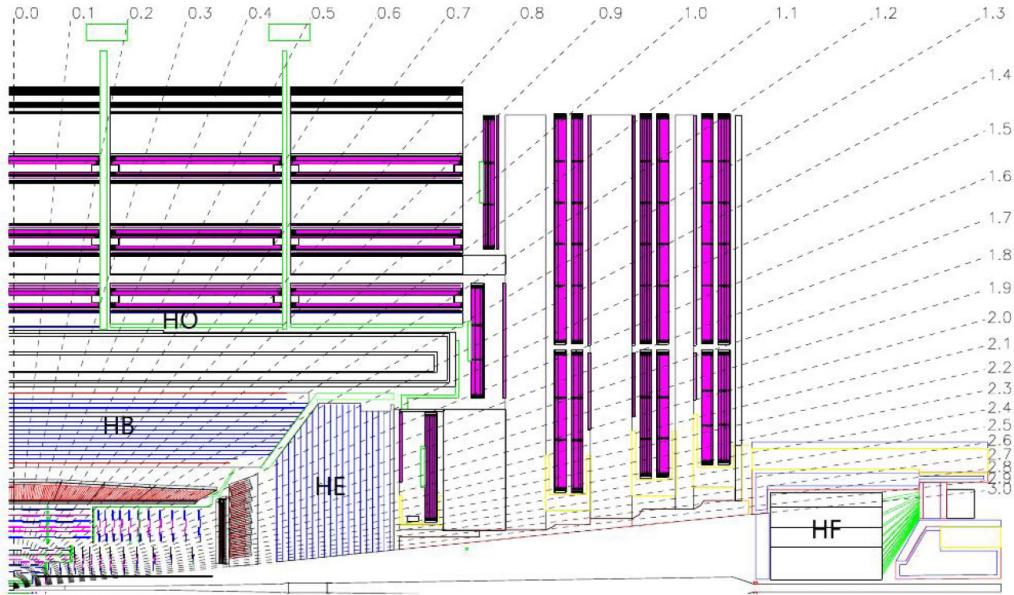


Figure 3.6: Longitudinal view of one quarter of the CMS detector showing the location of the individual HCAL sub-detector parts. Taken from [54].

have to be accounted for which is mainly done based on physics events. Changes in the transparency of the ECAL crystals during operation caused by irradiation are monitored by a dedicated laser system based on the injection of reference laser pulses into the crystals. The typical achievable relative energy resolution for 120 GeV electrons with this calorimeter configuration is of the order of 0.5%.

3.2.5 Hadron Calorimeter

In addition to the previously described electromagnetic calorimeter, the calorimetry of the CMS experiment is completed by the hadron calorimeter (HCAL). It is designed to provide an accurate energy measurement of hadron jets and indirectly also of invisible particles like e.g. neutrinos by the determination of missing transverse energy. In order to obtain a measure of the missing transverse energy it is important that the calorimeter is hermetic in the sense that it provides a large geometric coverage to potentially measure all particles emerging from an interaction. Thus the HCAL is build such that a pseudorapidity range up to $|\eta| = 5.2$ is comprised.

The hadron calorimeter completely surrounds the inner tracking system and the electromagnetic barrel calorimeter while it is mainly contained within the magnet system. Hence its radial dimensions are limited on the one hand by the outer circumference of the barrel ECAL and on the other hand by the inner border of the magnet coil. To ensure that the complete hadronic showers are contained in the HCAL, an additional calorimeter component is installed outside the solenoid in the barrel part.

An overview of the layout of the CMS hadron calorimeter is shown in Fig. 3.6. It is a typical sampling calorimeter with alternating layers of absorber material and active scintillator layers. The individual sub-components are the following:

Hadron barrel (HB): The barrel part of the CMS hadron calorimeter covers the pseudorapidity range up to $|\eta| = 1.3$ and is composed of two half barrels each containing 36 identical azimuthal wedges. These wedges hold the absorber plates which are flat brass plates arranged parallel to the axis of the beam. For reasons of stability the first and last layers are made of stainless steel. The total thickness of the absorber material ranges from 5.82 interactions lengths (λ_I) at $|\eta| = 0.0$ to 10.6 λ_I at $|\eta| = 1.3$. The 17 active plastic scintillator layers alternate with the absorber plates and have a segmentation in $(\Delta\eta, \Delta\phi)$ of $(0.087, 0.087)$.

Each half barrel is divided into 16 η -regions for which the individual tiles are optically linked together using wavelength shiftig fibres and thus form so-called *HCAL towers*. The read-out of each longitudinal tower is carried out using pixelated hybrid photodiodes.

Hadron outer (HO): The calorimeters in the central pseudorapidity region do not provide a sufficient depth in order to fully contain all hadronic showers. Therefore the HB is complemented by the outer hadron barrel part which is placed outside the solenoid covering $|\eta| \leq 1.26$. The HO makes use of the solenoid as additional absorber material and adds another one or even two layers in the most central part of scintillators to the barrel region. Thus the total depth is extended to $11.8 \lambda_I$.

Hadron endcap (HE): The hadron barrel calorimeter is supplemented by the hadron endcap. It is mounted on the endcap iron yoke and covers the pseudorapidity region of $1.3 \leq |\eta| \leq 3.0$ using 18 scintillator layers inserted into brass absorber plates. The granularity of the endcap calorimeter is the same as for the barrel up to $|\eta| = 1.6$ and gets coarser for larger pseudorapidities with $(\Delta\eta, \Delta\phi) \approx (0.17, 0.17)$.

Hadron forward (HF): The forward hadron calorimeter extends the pseudorapidity coverage from $|\eta| = 2.9$ (slightly overlapping with the HE) up to $|\eta| = 5.2$. It is located 11.2 m from the nominal interaction point and has to be in particular radiation hard to cope with the vast particle flux. Thus the HF is made of steel absorber plates with radiation hard quartz fibres integrated as active material. These fibres are arranged parallel to the beam line and form towers with a size in $(\Delta\eta, \Delta\phi)$ of $\approx (0.175, 0.175)$. The signal is detected as Cerenkov light originating from the quartz fibres.

Similar to the ECAL also the performance of the HCAL has to be well calibrated and further monitored during operation. Thus an initial calibration using a radioactive source is combined with test beam data to derive the absolute energy scale. A continous update of this calibrations is performed using isolated energetic particles e.g. from decays of W or Z bosons.

3.2.6 Muon System

The outermost part of the CMS detector – as seen from the interaction point – is made up of the muon system. This important component of the detector is assembled in the return yoke of the CMS magnet and consists of a central barrel cylinder which is complemented by endcap disks in the forward region. This results in a coverage of the pseudorapidity range up to $|\eta| = 2.4$. In the barrel part four layers of detectors are installed alternating

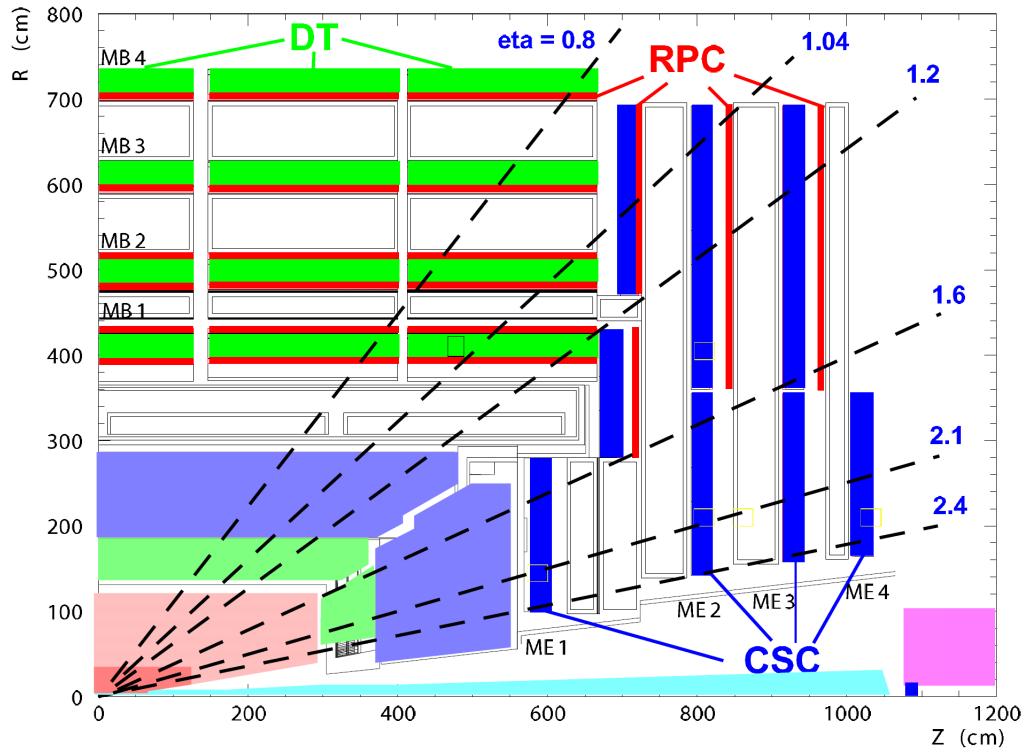


Figure 3.7: Taken from [55].

with the iron yoke while the detectors in the endcap are mounted on four discs perpendicular to the beam. In total about 25000 m^2 detection planes are employed. The layout of the muon system is illustrated in Fig. 3.7.

In order to feature a good muon momentum resolution given the different radiation conditions and variations in the homogeneity of the magnetic field depending on the pseudorapidity region it is made use of three different types of gaseous detectors. These different types of tracking chambers are described in the following:

Drift tube (DT) chambers: In the barrel region for $|\eta| < 1.2$ where the background due to neutrons is low and residual effects from the magnetic field likewise the muon system is equipped with drift tube chambers. While in all four stations in the barrel the muon coordinates in the $r\phi$ -plane are measured, only the first three layers provide also a measurement of the z -direction. The maximum drift length was chosen to be 21 mm resulting in a negligible occupancy while keeping the number of active channels at an acceptable level. Furthermore, a technology based on tubes was chosen avoiding the issue of possibly broken wires. The resolution in $r\phi$ is designed to reach a precision of 100 μm .

Cathode strip chambers (CSC): The endcap regions are equipped with cathode strip chambers and cover the pseudorapidity range $0.9 < |\eta| < 2.4$. These are chosen as they provide a fast response time and fine segmentation while they are resistant against radiation. Thus they are well suited for the forward region where the muon and background rates are largely increased and the magnetic field is high and non-

uniform. The CSCs which are multiwire proportional chambers where anode wires are interlaced with cathode panels perform a precise position measurement in the $r\phi$ bending plane with a spatial resolution of 75–150 μm .

Resistive plate chambers (RPC): Resistive plate chambers are used to complement the drift tube and cathode strip chambers in the range $|\eta| < 1.6$. These are gaseous parallel-plate detectors with a spacial resolution coarser than the DTs and CSCs but usable at high particle rates while providing a very fast response and good time resolution. Thus they are able to very efficiently detect the correct bunch-crossing a muon track is associated to.

The global muon reconstruction efficiency is in general about 95 – 99% and only drops for some $|\eta|$ regions like e.g. in the transition region between the barrel and endcap part around $|\eta| = 1.2$. Since muons reaching the muon system are affected by multiple scattering the resolution for muons with low transverse momenta below $\approx 200 \text{ GeV}$ is in general better based on the inner tracking system than for the muon system alone. For highly energetic muons it is comparable. In general, the muon momentum resolution can be improved by combining the information from the inner tracker and the muon system due to an improved fault finding. This combined approach results in a relative muon momentum resolution for muons with high momenta around 1 TeV of about 5%.

3.2.7 Trigger System

The LHC operating at design conditions provides particle collisions with an interaction rate of 40 MHz. This results in an enormous amount of data events which have to be processed and stored for later offline analyses. With an approximate event size of 1 MB it is technically impossible to record all such events. However, as illustrated in Fig. 3.2 the event rate of interesting potentially new physics events is orders of magnitudes smaller than the proton-proton cross section. Thus, this allows to already perform online a suitable event preselection in order to reduce the amount of data to a storable size still containing the information of interest. Hence, the trigger system also makes up the first step in the physics analysis process.

In order to achieve the necessary rate reduction, the CMS experiment makes use of a two-stage trigger system. This is described in the following.

Level-1 (L1) Trigger: The L1 Trigger consists of custom-made fast programmable hardware. It makes use of data received from fast detector components – namely the calorimeters and the muon system – at reduced granularity. For instance at L1 the calorimeter is divided into so-called *trigger towers* which cover an area in (η, ϕ) of (0.087, 0.087) up to $|\eta| = 1.74$ and get even coarser for higher $|\eta|$. At L1 the trigger decision is based on energy deposits in those trigger towers or certain hit patterns in the muon chambers forming trigger primitive objects as electrons/photons, muons or jets and global quantities like sums of E_T or \cancel{E}_T . Events are accepted, if these trigger objects pass some predefined criteria like certain p_T thresholds. The L1 trigger latency between the actual bunch crossing and the delivery of the L1 trigger decision to the front-end electronics is $3.2 \mu\text{s}$. During this period the high resolution data is pipelined in readout buffers for further processing. Following this procedure the L1 trigger reduces the event rate to 100 kHz.

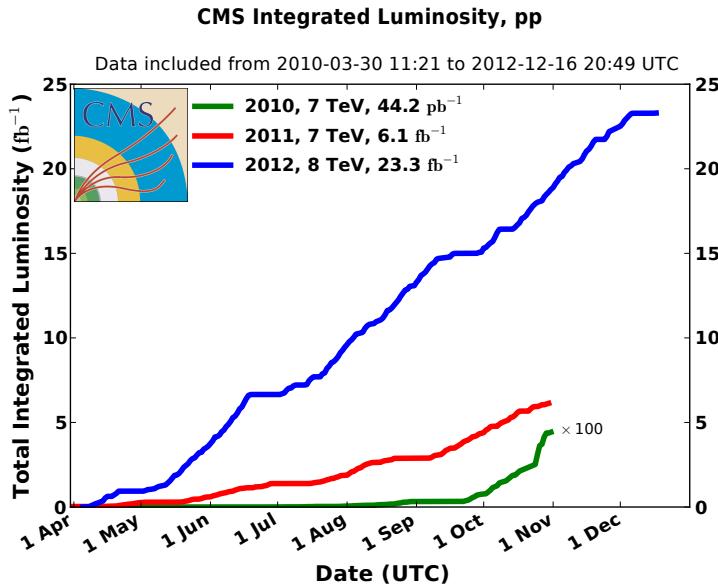


Figure 3.8: Cumulative integrated luminosity versus day delivered to CMS during stable beams for pp collisions. Different data-taking periods are indicated as follows: green for 2010, red for 2011 and blue for 2012. Taken from [59].

High-Level Trigger (HLT): Events that are accepted by the L1 are transferred to the High-Level trigger for further processing. The HLT is a software system running on several thousand commercial processors. It has access to the full information from all sub-detectors and performs an event reconstruction similar to the later offline reconstruction. Thus it allows a further rate reduction to the final output rate of a few hundred Hz of events that are finally stored for analyses. Since the HLT is software based, it allows to continuously adjust the used algorithms in order to perfectly meet changing conditions during operation.

3.3 LHC Operation and Data Taking

The LHC was put into operation the first time in September 2008. After a major cooling incident only a few days later requiring a longer technical stop, beams were circulated again in November 2009. The first collisions at a center of mass energy of 7 TeV finally took place end of March 2010 [60]. In the following running period in 2010, data corresponding to an integrated luminosity of 44.2 pb^{-1} were delivered to the experiments with a maximum peak instantaneous luminosity of $2.05 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$. These data allowed intense studies of the detector performance and made first searches for new physics possible. The next data taking period performed during 2011 at the same center of mass energy even lead to much more pp collision data provided to the experiments with a total amount of 6.1 fb^{-1} reaching a peak instantaneous luminosity of $3.5 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$. Following another technical shutdown during winter, the center of mass energy was finally increased to 8 TeV for the running period during 2012 and the peak luminosity reached with values of up to $7.7 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ already almost design conditions. In total, 23.3 fb^{-1} of integrated

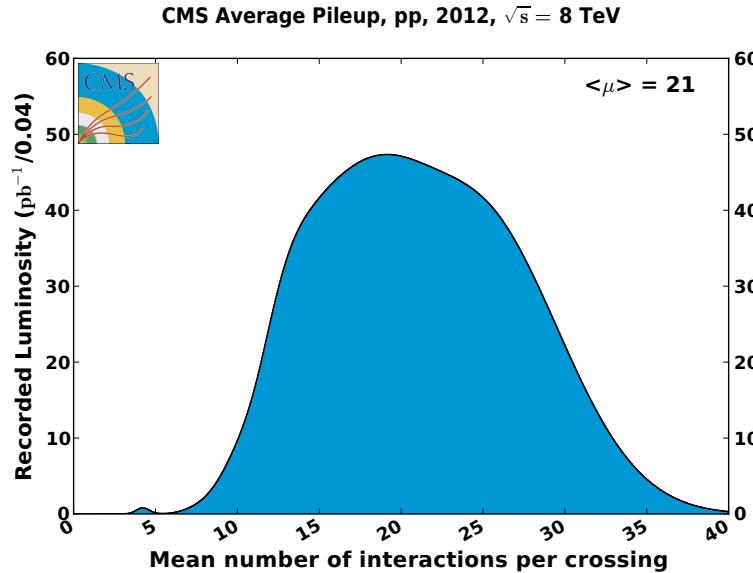


Figure 3.9: Mean number of interactions per bunch crossing in pp collisions at $\sqrt{s} = 8 \text{ TeV}$.
Taken from [59].

luminosity pp collisions were produced by the LHC operating at stable beam conditions. The evolution of the integrated luminosity versus days is also illustrated in Fig. 3.8.

During most of the operation in 2012, the LHC was circulating 1380 bunches per beam with a distance of 50 ns. The average bunch intensity, i.e. the number of protons per bunch, was varying from 1.6 to 1.7×10^{11} extending already the design value [59, 61]. This was on average resulting in 21 pp collisions per bunch crossing. This process of multiple interactions per bunch crossing is known as *pileup*. An overview of the pileup profile observed in collision data taken in 2012 is given in Fig. 3.9. The maximum number of pileup events even reached a value of ≈ 40 . This effect caused by the required high luminosity generates challenging conditions for the experiments as many objects arising from such multiple interactions have to be assigned to the correct process and particular collision objects have to be filtered. Dedicated techniques in order to cope with this problem have been developed of which some are discussed in Chapter 4.

3.4 Event Simulation

An important tool in high energy physics is the use of simulation in order to acquire a good understanding of the behaviour of particle collisions and related collision products observed in the detector. Thus, simulated events often serve as benchmark for the development of new detector concepts. Moreover, they are heavily exploited in the validation and interpretation of the results of actual collision experiments, likewise for the LHC. For instance simulated events are used to derive expectations for certain physics properties or the detector performance. In particular they allow to estimate how signals of new physics events would look like in the experiment.

This section provides a brief introduction to the principles of event simulation in hadron

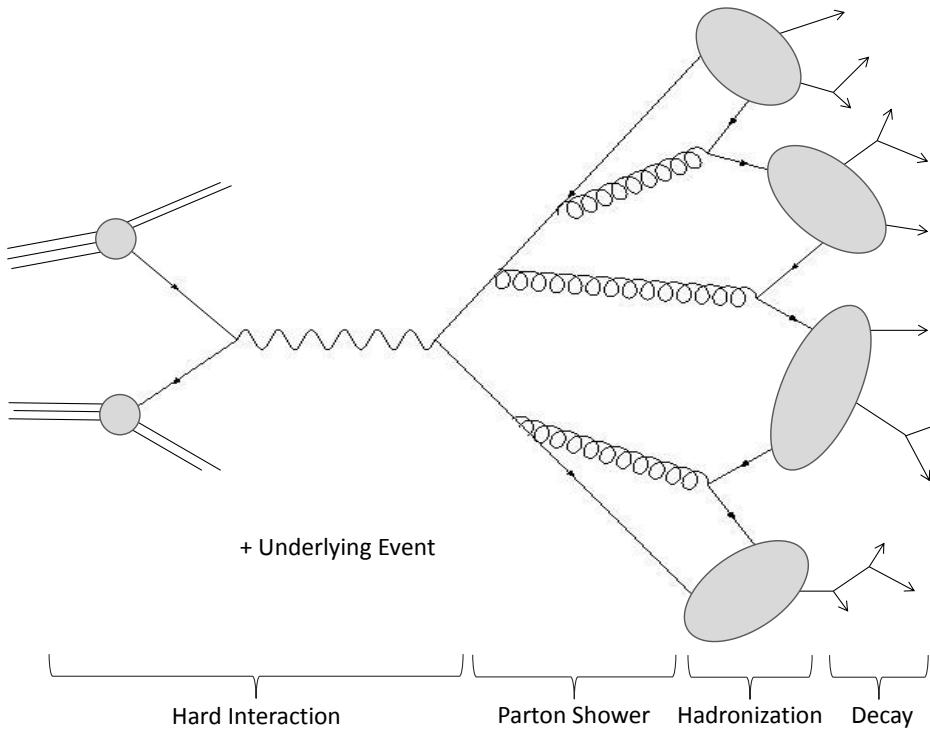


Figure 3.10: Sketch of individual steps in the generation process of simulated events.

collisions and introduces some event generators including different approaches for the simulation of the detector. A broader overview of event simulation and respective generators for LHC physics can be e.g. found in [62, 63].

The simulation of high energy collisions is a quite challenging task as each collision involves typically several hundreds of particles with momenta ranging over some orders of magnitude. Furthermore, the collisions being subject to quantum chromodynamics are only calculable within approximation schemes where also numerical methods do not provide solutions within a reasonable amount of time. Thus, event simulation is primarily utilizing *Monte Carlo* (MC) techniques which rely on the repeated sampling of random numbers, cf. e.g. [64]. For convenience events obtained from simulation are often denoted by the label 'MC' in this thesis.

Typically, the generation of an event follows several subsequent steps which are illustrated in Fig. 3.10 and described in the following:

Hard process: The first step in the simulation of collision events is the description of the nominal proton-proton interaction which is typically referred to as *hard process*. The proton itself is not a fundamental particle but exhibits an internal structure . The constituents of which a proton is made off are known as *partons*. Thus two protons interact when there is a momentum transfer Q taking place between two individual partons. The probability for individual partons to take part in the hard interaction is parametrized by the *parton-distribution functions* (PDFs) which have been determined experimentally . The value of x denotes the fraction of the

Quelle:
Proton
Struktur

Quelle:PDFs

longitudinal momentum carried by an individual parton. Consequently, the initial state of a proton-proton collision is not exactly known. Nevertheless, the cross section of a specific process can be calculated following the *factorization theorem* [65, 66]. Here, the hard interaction is described via perturbation theory and low-energy processes are considered in the phenomenological models of the respective PDFs. Thus the cross section for a process $ab \rightarrow n$ is given according to

$$\sigma = \sum_{a,b} \int_0^1 dx_a dx_b \int f_a(x_a, \mu_F) f_b(x_b, \mu_F) d\hat{\sigma}_{ab \rightarrow n}(\mu_F, \mu_R). \quad (3.6)$$

Here, $f(x, \mu_F)$ are the PDFs of the interacting protons which depend on the factorization scale μ_F whereas $\sigma_{ab \rightarrow n}$ indicates the parton-level cross section for the production of a final state n from partons a and b . This parton-level cross section depends on the final-state phase space, the factorization scale and the renormalisation scale μ_R as well as the corresponding matrix element. The factorization and renormalization scale are unphysical and have to be chosen for the generation process. Often the process possesses a typical hard scale Q^2 so that the choice $\mu_F = \mu_R = Q^2$ is made. However, this choice is not fixed by first principles.

Parton shower: After the production in the hard interaction the outgoing partons start to form a shower, i.e. cascades of further partons emerge. Typically, this happens with descending amounts of momentum transfer from the high scales down to low scales around 1 GeV. This evolution is typically described by a probabilistic shower algorithm. In addition to a parton shower related to the outgoing partons which is referred to as *final-state radiation* (FSR), partons can already radiate off other partons before the actual hard process takes place. This effect is known as *initial-state radiation* and can be described by similar principles as FSR. However, in the case of ISR it is inevitable to consider that not all partons from the ISR shower necessarily take part in any hard process and thus have to be related to the proton remnant.

Hadronization: The evolution of the parton shower continues until low scales in momentum transfer are reached and so the final state partons eventually start to form colour neutral hadrons. In the context of simulation the term *hadronization* describes the particular model which is used in order to specify the transition from the partonic state to the complete hadronic final state. Since the hadronization involves low-energetic processes it can not be treated within perturbation theory. In principle one distinguishes between *string models* (cf. e.g. [67]) and *cluster models* (cf. e.g. [68, 69]) in order to characterize the hadronization. While the first model describes directly the transition from the parton to the hadron based on the assumption of linear confinement, the second model introduces an intermediate step of cluster objects with mass scales around a few GeV. Particles arising from the hadronization are denoted with *generator-level* particles and labelled 'gen' within this thesis.

Decay: After the hadronization process a couple of unstable hadrons are present in the event whose decay into stable particles must be modelled. Stable in this context means that they do not decay further within the typical collider timescales. Thus,

the observable final-state hadrons result from a convolution of the hadronization process with the decay modelling. Concerning the decay process a choice has to be made regarding the hadrons which shall be included in the simulation and the respective decay modes which have to be considered. Mainly, these choices are based on experimental results in combination with theoretical assumptions. Typical differences occur e.g. in the consideration of excited mesons or heavy baryon multiplets as well as an appropriate treatment of matrix elements and spin correlations during the decay.

Underlying event: In addition to hadrons emerging from the process connected directly to the hard interaction, further contributions to the event can be present which is referred to as *underlying event* (UE). The UE activity can be e.g. caused by additional interactions occurring for incoming partons, which is known as *multiple parton interaction* (MPI), or by interactions arising from the proton remnants. In general, such effects contribute to the total amount of scattered energy and increase the number of particles appearing in the hadronization process.

The various sub-processes described above are the basis of various event generators which differ in some aspects concerning the treatment of individual sub-processes in the event simulation. In this thesis, different generators are used:

PYTHIA [70]: PYTHIA is a general purpose event generator which has been used already extensively at previous collider experiments. It is designed to simulate collisions of either hadrons or same-generation leptons. This means that it is well suited to model the required pp collisions, but is currently not usable for e.g. ep or γp processes. The highest multiplicity of particles involved in the hard interaction that can be simulated with PYTHIA are $2 \rightarrow 3$ processes. Nonetheless, more than three final state particles can arise for instance from the parton shower. Besides the simulation of whole physics events including all relevant steps, PYTHIA can also be interfaced to other generator programs in order to carry out the parton shower and hadronization step only.

HERWIG++ [71]: HERWIG++ is as well as PYTHIA a general purpose MC generator. It provides the possibility to simulate high-energy collisions for lepton-lepton, lepton-hadron and hadron-hadron processes with special emphasis on the modelling of QCD radiation. Thus, the distinct feature of HERWIG++ is that colour coherence effects are taken into account by employing angular ordered parton showers meaning that the coherence of soft radiation is treated correctly.

MADGRAPH [72]: MADGRAPH is a general-purpose matrix-element program and designed to provide accurate descriptions of multiparton processes. It can commonly be interfaced to PYTHIA for the realisation of the showering and hadronization.

POWHEG [73]: The POWHEG programme provides accurate QCD computations based on matrix elements up to next-to-leading order following the concept proposed in [74]. It can be interfaced to shower programs like PYTHIA or HERWIG++ in order to perform the subsequent showering process.

In order to evaluate the interaction of the generated particles with the detector material the passage of the particles through the detector is simulated. This is done based on a

detailed model of the CMS apparatus utilizing the GEANT4 programme [75] and referred to as *full simulation*. Full simulation is quite extensive in computing time and thus for several purposes also a simplified scheme of the detector model is used which is known as *fast simulation* [76]. Especially in analyses where many signal samples are needed to scan a wide parameter range, like e.g. in SUSY analyses, the usage of the fast simulation is very beneficial. A comparison of several relevant kinematic distributions in real collision events to simulated events obtained from fast simulation shows a quite good agreement [77]. In particular, also jet related quantities like the sum of calorimetric jet momenta or the missing transverse momentum are modelled with a similar accuracy in full and fast simulation. Finally, recorded collision data as well as simulated events including the modelling of the passage through the detector are present in the same data format which allows to apply equal reconstruction techniques in order to derive physics objects from the obtained events.

4 Object Reconstruction and Particle Identification

Particles produced in the pp collisions traverse through the detector and interact with the detector sub-components in a characteristic manner. Thus, it is possible to reconstruct the corresponding event and identify the types of particles which actually emerged from the collision.

The approach for the event reconstruction and identification of specific particles used in CMS is discussed in this Chapter. First, the *Particle-Flow (PF) algorithm* used for a global description of the collision event is introduced. Beyond that, in particular the reconstruction of jets is discussed in Section 4.2. Furthermore, the identification of decays from b-hadrons and boosted top quarks is reviewed in Sections 4.3 and 4.4, respectively.

4.1 Global Event Description with the Particle-Flow Algorithm at CMS

The CMS experiment introduced the Particle-Flow algorithm for the reconstruction of collision events. This algorithm is designed to identify all stable particles in an event and can be applied to data events as well as to simulated events in an identical manner. Electrons and photons, charged and neutral hadrons as well as muons are thereby distinguishable and all sub-components of the detector are used by the PF algorithm to reconstruct the particles four-momenta. The CMS detector is very well suited for this task. The silicon tracker enclosed by the uniform magnetic field enables a very efficient track reconstruction yielding only a small track fake rate down to small transverse momenta of 150 MeV/c. Furthermore, the strength of the magnetic field together with a high ECAL granularity allows photons to be separated from charged-particle energy deposits. A detailed introduction to the PF algorithm can be found in [78].

The event reconstruction starts with the identification of fundamental objects in the sub-detectors which are charged-particle tracks, calorimeter clusters and muon tracks. Tracks emerging from charged particles are formed following an iterative tracking algorithm [79]. Starting from an initial seed trajectory, tracks are extrapolated to further tracker layers by taking into account multiple scattering and energy loss in the material following the equations of motion of a charged particle in a constant magnetic field. Each step proceeds with a removal of unambiguously allocated hits from the previous iteration. With this approach a high tracking efficiency as well as a low fake track rate can be achieved which is a crucial requirement for a successful particle-flow event reconstruction. Furthermore, calorimeter clusters are formed in each sub-detector separately based on adjacent calorimeter cells. Neighbouring cells are combined to form clusters when their energy exceeds a pre-defined threshold.

A particle traversing through the detector gives typically rise to several of such elementary

components so that a dedicated link algorithm is applied in order to connect these elements and form blocks while removing a potential double-counting of the same object in different detector parts. First, charged-tracks are associated to calorimeter clusters, if the extrapolated trajectory matches the cluster within the cluster boundaries. This is done considering effects like gaps and cracks between detector components, uncertainty on the shower position or multiple-scattering. Photons from Bremsstrahlung are considered by extrapolating also tangents of the tracks to the respective energy clusters. In a similar manner ECAL and HCAL clusters can be connected to each other as well by linking clusters in the more granular calorimeter to clusters in the less granular one. At last global muons can be defined by associating charged-tracks from the tracker with muon tracks reconstructed in the muon system.

After the identification of such blocks of elements the PF algorithm proceeds to finally create a list of all particles contained in the event applying dedicated quality criteria interpreting the blocks in terms of particles. The identification of muons and a removal of their tracks from the blocks is followed by an assignment of electrons and associated Bremsstrahlung from tracks and linked ECAL clusters. After these have been removed from the list of blocks as well, remaining good quality tracks are considered to be charged hadrons. Their momenta are determined from combining the track momentum and the respective energy in the calorimeter cluster. If the cluster energy largely exceeds the measured momentum from the track beyond the detector resolution, it constitutes a photon and if the excess is larger than the total ECAL energy also a neutral hadron. Finally, remaining ECAL and HCAL clusters not linked to any track give rise to photons and neutral hadrons.

The complete set of particles can then be used to derive further objects and quantities like e.g. jets as discussed in Section 4.2, missing transverse energy E_T^{miss} which is the magnitude of the vector momentum imbalance perpendicular to the beam direction or decay products of tau leptons. More detailed information on the specific quality criteria required for the identification of certain particles is given in the Chapters 5, 6 and 7 for each analysis presented in this thesis, individually.

4.2 Reconstruction of Jets

As stated already earlier jets are the experimental signatures of quarks and gluons and constitute of a collimated spray of hadrons. However, in order to identify a particular jet and relate its properties to the original parton a proper jet definition is needed. Typically, a *jet algorithm* determines how to cluster particles into a jet. Furthermore, it has to be defined how to assign a momentum to the jet. At CMS, the standard procedure is to assign the four-momentum sum of all jet constituents to the jet.

Different jet algorithms are introduced in Section 4.2.1 followed by a discussion of different jet types used at CMS in Section 4.2.2. Furthermore, the jet transverse-momentum response is defined in Section 4.2.3 and the calibration of jet energies is reviewed in Section 4.2.4.

4.2.1 Jet Algorithms

A jet algorithm usually provides a prescription how to combine individual particles into a single jet based on some distance criterium. Good jet algorithms though should be able

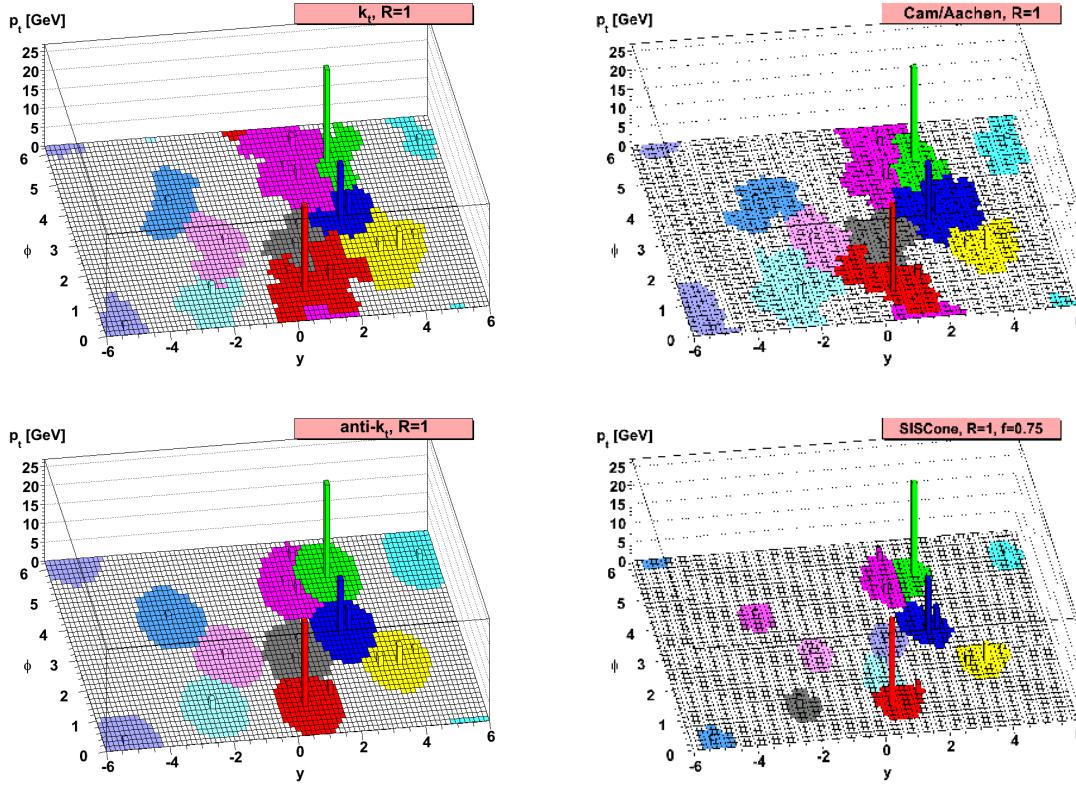


Figure 4.1: Sample parton-level event generated with HERWIG++, adding many random soft particles, that is clustered with the k_T -algorithm (*top left*), the C/A-algorithm (*top right*), the SIScone-algorithm (*bottom left*) and the $\text{anti}-k_T$ -algorithm (*bottom right*). Colours illustrate the active areas of the resulting hard jets. Taken from [80].

to identify jets that are neither sensitive to the emission of soft particles (*infrared safety*) nor to the collinear splitting of particles (*collinear safety*). These features named *IRC safety* are desirable as otherwise e.g. cross sections would be dependent on particular effects of hadronisation involving random collinear splittings and soft emissions making it difficult to compare the results from experiments to theoretical predictions. This effect is even enforced because detectors intrinsically provide regularisation of IRC effects to some extent due to a finite resolution. Furthermore, calculations in QCD perturbation theory rely on the cancellation of divergences related to IRC processes. Thus, if jets are sensitive to such effects cancellations are not ensured and can lead to infinite cross sections. An introduction to the most commonly used jet algorithms known as *cone algorithms* and *sequential recombination algorithms* is given in the following. A comprehensive overview of jet algorithms and properties can be found in [80].

Cone Algorithms: Cone algorithms were the first developed jet algorithms. They are based on the general idea that the main kinematics in an event are not changed by specific effects from hadronisation and thus a jet is defined by a set of particles within a stable cone around their centre of mass. Typically, separate angular or

energy parameters are used to perform the jet finding. A very common approach is implemented in *iterative cone* (IC) algorithms. Here, a seed constituent i which is for instance the constituent with the highest transverse momentum defines the initial direction and momenta of all constituents j within a cone defined by

$$\Delta R_{i,j}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2 < R^2 \quad (4.1)$$

are added to the momentum of the seed. The resulting direction is used as new seed direction and the whole procedure is repeated until a stable cone is achieved. The dimensionless parameter R hence defines the jet radius. After the finding of such a jet, all constituents are removed from the input list and further jets are clustered from the remaining objects. This progressive removal approach avoids to form jets with overlapping cones. However, such a procedure is not IRC safe as collinear splittings can lead to different final ensembles of jets.

The issue with IRC safety in cone algorithms can be avoided by instead of iteratively forming stable cones one after the other identifying all stable cone solutions at once. This procedure is denoted *seedless cone* (SC) algorithm. The usage of such algorithms though is typically impractical as the computation time increases exponentially with the number of particles to be considered so that even for 100 particles it is not solveable at any reasonable timescale. A feasible implementation of a seedless cone algorithm featuring only a logarithmic time-dependence is given by the SISCone algorithm [81]. As it is usually nonetheless still more time consuming than sequential algorithms as described in the next part, the SISCone algorithm is not used by CMS.

Sequential Recombination Algorithms: The basic concept of sequential clustering algorithms is to group in an iterative procedure pairs of particles together based on some distance measure and thus reconstructs to some extent the evolution of a parton shower. A suitable metric to be used at hadron colliders where the total energy of a collision is unknown based on variables invariant under longitudinal boosts is

$$d_{ij} = \min(k_{T,i}^{2p}, k_{T,j}^{2p}) \frac{\Delta R_{ij}^2}{R^2}$$

$$d_{iB} = k_{T,i}^{2p}$$

with the distance d_{ij} between final state objects i and j carrying transverse momentum k_T and the distance of the object to the beam d_{iB} . While ΔR_{ij} denotes the spatial separation in the (η, ϕ) -plane, R and p are free parameters of the algorithm. The recombination is done by first calculating d_{ij} and d_{iB} for all objects in the final state by identifying the minimum value. If the minimum is d_{ij} , the two objects i and j are combined and all distances are computed again. However, if the minimum is d_{iB} , object i is declared a jet and removed from the input list. This procedure is repeated until all objects are assigned. In this context the parameter R acts as an angular cut-off and thus has a similar role as the jet radius in cone algorithms. Depending on the choice of the parameter p different types of algorithms are distinguished which are all IRC safe. An illustration of an arbitrary example event where jets are clustered with the different jet algorithms is shown in Fig 4.1. The

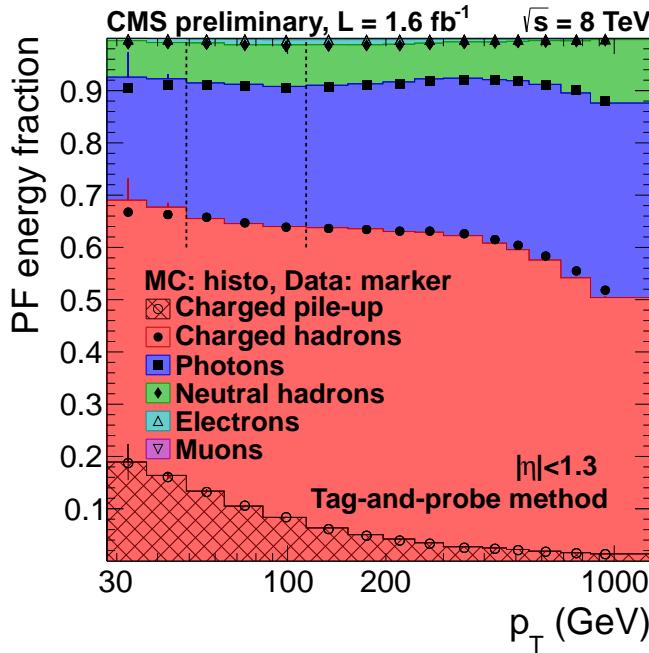


Figure 4.2: Composition of the PF jet energy versus jet p_T in the barrel detector region $|\eta| < 1.3$ in simulated events (coloured histograms) and data (solid markers). Taken from [86].

k_T -algorithm uses $p = 1$ [82] and thus clusters soft particles first. This results mainly in irregular shaped jets (cf. Fig 4.1) and makes them sensitive to radiation in the event. Consequently, they are difficult to calibrate and therefore impractical to use at hadron colliders. The *Cambridge-Aachen-* or short *CA*-algorithm [83, 84] utilizes $p = 0$. Thus, it does not rely on momentum information but clusters jets based on the angular separation of input objects. This allows a direct geometric interpretation of jets and makes this type of jet algorithm especially suited for analyses of jet substructure as will be discussed later in Section 4.4. These jets do still reflect the structure of the parton shower, but are less affected by soft radiation than the k_T -algorithm. Finally, the *anti- k_T* -algorithm uses $p = -1$ [85] and starts the jet clustering beginning with the hardest objects in the event. Hence, the evolution of the parton shower is not reproduced. The anti- k_T algorithm tends to form very regular shaped jets, as the four-momentum of the core of the jet is not much affected by the softer components which are clustered later in the process. Typically, the regular shape allows an easy calibration of anti- k_T jets.

4.2.2 Jet Types at CMS

The default jets to be used at the CMS experiment are jets clustered by the anti- k_T algorithm using a distance parameter of $R = 0.5$. This can be applied to reconstructed detector signals resulting in *detector-level* jets or to final-state particles after hadronisation, as this is possible in simulated events, giving rise to *particle-level* jets. Implementations of various jet algorithms are provided by the *FastJet* package [87, 88].

Based on the information used from the various CMS subdetectors for the jet clustering different types of detector-level jets are distinguished. Some of them are the following [89]:

Calorimeter (Calo) jets: Calo jets are clustered from energy deposits in the calorimeters. For this purpose, calorimeter towers are defined which consist of at least one HCAL cell and the geometrically associated ECAL cells. In case of the barrel detector region a calorimeter tower consists for instance of one HCAL cell and 5×5 ECAL cells. The four-momentum of each tower is defined by the tower position as seen from the interaction point and the energy deposit in the tower above a certain threshold assuming a mass of zero.

Jet-Plus-Track (JPT) jets: JPT jets are reconstructed from calorimeter jets complemented by tracking information [90]. Tracks of charged particles can be associated to calo jets based on the separation of the jet axis and the momentum vector of the track in the (η, ϕ) -plane. Associated tracks are projected to the jet cone and are exploited to correct the jet energy and direction resulting in an improved jet response.

Particle-Flow (PF) jets: PF jets are clustered from the four-momentum vectors of particle-flow candidates as identified by the PF algorithm described in Section 4.1. Typically, these types of jets show the best performance as the excellent resolution of the tracking system and the ECAL are utilized. Only neutral hadrons which constitute around 15% of the jet energy (cf. Fig. 4.2) rely on the energy measurement of the HCAL with its relatively poor resolution. Thus, PF jets are the default jets to be used at the CMS experiment, as it is done within this thesis. In order to mitigate influences from pile-up, charged hadrons unambiguously associated to vertices other than the primary vertex can be subtracted from the event before performing the actual jet clustering. This technique is referred to as *charged hadron subtraction* (CHS) and used as default throughout this thesis.

4.2.3 Jet Transverse-Momentum Response

In general, the jet transverse momentum as measured at detector level is not necessarily equal to the energy of the original particle due to for instance the limited detector resolution. This effect is quantified by the *jet transverse-momentum response* R which is defined as

$$R = \frac{p_T}{p_T^{\text{particle}}} \quad (4.2)$$

where p_T denotes the transverse momentum of the jet measured at detector level and p_T^{particle} is the transverse momentum of the original particle-level jet. Thus, the jet response provides a measure of the fraction of the jet momentum visible in the detector compared to the actual momentum of the particle after hadronisation.

The average response $\langle R \rangle$ is referred to as *jet energy scale* (JES) and has to be calibrated such that $\langle R \rangle = 1$ for fixed p_T^{particle} . The response usually depends on the jet momentum as well as on the pseudorapidity. This is expected since the quality of the jet measurement is directly related to the energy of the particles and the resolution of the detector sub-components. This is for instance caused by the specific track-reconstruction efficiency or the individual amount of detector material.

Furthermore, the width of the response distribution corresponds to the relative *jet transverse-momentum resolution* (JER) and hence is a function of p_T and η as well.

4.2.4 Jet Energy Calibration

In order to relate the measured jet momentum on average to the momentum of the corresponding particle-level jet a jet calibration procedure is applied. Within the CMS experiment a factorized approach is utilized which is described in detail in [89].

The calibrated jet four-momentum vector p^{cor} is obtained from the raw jet momentum four-vector p^{raw} by scaling the raw momentum with a correction factor C according to

$$p^{cor} = C \cdot p^{raw} = C_{\text{offset}}(p_T^{raw}, \eta) \cdot C_{\text{rel}}(\eta) \cdot C_{\text{abs}}(p'_T) \cdot C_{\text{res}}(p''_T, \eta) \cdot p^{raw} \quad (4.3)$$

where C is composed of the offset correction C_{offset} , the calibration factors C_{rel} and C_{abs} as well as residual correction factors C_{res} applied to data. Each correction factor is applied sequentially after the other in a fixed order such that p'_T denotes the transverse momentum after the application of the offset correction and p''_T is the transverse momentum after all previous corrections. Some details for each individual correction are given in the following:

L1 Offset: The L1 offset correction is designed to compensate for additional energy contributions arising from instrumental noise or pile-up events. The p_T offset is estimated in dependence of η , the effective jet area A_j and the p_T -density ρ . The jet area is determined by adding a large number of infinitely small four-momentum vectors to the event. The active jet area is hence defined as the region in the (η, ϕ) -plane occupied by the soft particles after clustering them together with the true hard jet components. The p_T -density ρ is defined on an event-by-event basis as the median of the distribution p_T^j/A_j where j denotes all reconstructed jets in the event. The offset in p_T is estimated from simulated QCD multijet events by directly comparing the jet momentum of a jet with and without pile-up overlaid.

L2 Relative + L3 Absolute Correction: The L2 relative correction is designed to make the jet energy scale uniform with respect to η while the L3 absolute correction ensures a uniform response versus p_T . Both corrections are entirely estimated from simulated QCD multijet events. The correction is defined as the inverse of the average response $1/\langle R \rangle$.

L2L3 Residual Correction: In order to compensate for remaining response differences between simulated events and data, residual correction factors are derived. These are applied to data only and correct for remaining differences in the data-to-simulation ratio of the relative jet energy scale. Such corrections can be derived from events featuring a momentum balance in the transverse plane like dijet events (used for the L2 residual determination) or $Z/\gamma +$ jet events (used for the measurement of the L3 residual correction).

The calibration factors are obtained with respect to the average flavour composition of a QCD multijet sample. Thus, further steps of correction factors can be applied for specific analysis purposes correcting e.g. the different response for various jet flavours. However, such higher order corrections are not used in this thesis.

The actual set of correction factors used in this thesis, if not stated otherwise, is documented in [91] and illustrated in Fig. In total, the size of the JES corrections are at the order of

Fig. JEC
size of
JEC

4.3 Identification of b-Quark Jets

Jets arising from the hadronization of bottom quarks are usually referred to as *b-jets*. As these are existent in many physics processes as e.g. the decay of top quarks, it is of essential importance to be able to identify b jets and distinguish them from jets initiated by gluons or light-flavour quarks. Typically, the identification of b-jets is denoted *b-tagging* and the distinct properties of b quarks are exploited for the identification of the respective jets. In general, hadronic jets arising from the fragmentation of b quarks possess quite large masses, long lifetimes and daughter particles featuring hard momentum spectra. The CMS experiment with the ability to perform a precise charged-particle tracking and lepton identification exploits the specific b jet properties in dedicated b-tagging algorithms for an efficient b-jet identification [92]:

Track Counting (TC) algorithm: A powerful discriminator for the decay products of a b hadron from prompt tracks is the *impact parameter* (IP) of a track with respect to the primary vertex. Its significance can be computed by taking the ratio of the IP to its respective uncertainty. Tracks in a jet are sorted by decreasing values of the IP significance by the TC algorithm. Depending on whether the IP significance of the second or the third ranked track is chosen as discriminator value the algorithm is denoted *Track Counting High Efficiency* (TCHE) or *Track Counting High Purity* (TCHP) algorithm.

Jet Probability (JP) algorithm: The JP algorithm extends the simple TC algorithm by connecting the information about the IP from a couple of tracks in the jet. A likelihood is calculated that all tracks of the jet stem actually from the primary vertex. This approach can be varied by giving more weight to tracks with the highest IP significance. The maximum of such tracks is four and matches the average number of reconstructed charged particles from the decay of b hadrons. This version is called *Jet B Probability* (JBP) algorithm.

Simple Secondary Vertex (SSV) algorithm: A further useful discriminating feature for b tagging is the presence of a secondary vertex and related kinematic variables like the flight distance and direction which can be determined from the vector between the primary and secondary vertex. The SSV exploits the significance of the flight distance which is given by the flight distance divided by the associated uncertainty. Two different versions of this algorithm exist targeting on the one hand a *High Efficiency* (SSVHE) and on the other hand a *High Purity* (SSVHP). While the SSVHE is based on vertices with at least two associated tracks, the SSVHP uses vertices with more than three tracks. Typically, the efficiency of the algorithm is limited by the reconstruction efficiency of secondary vertices which amounts to about 65%.

Combined Secondary Vertex (CSV) algorithm: The CSV algorithm provides an efficient identification of b jets also in cases when no secondary vertex could be

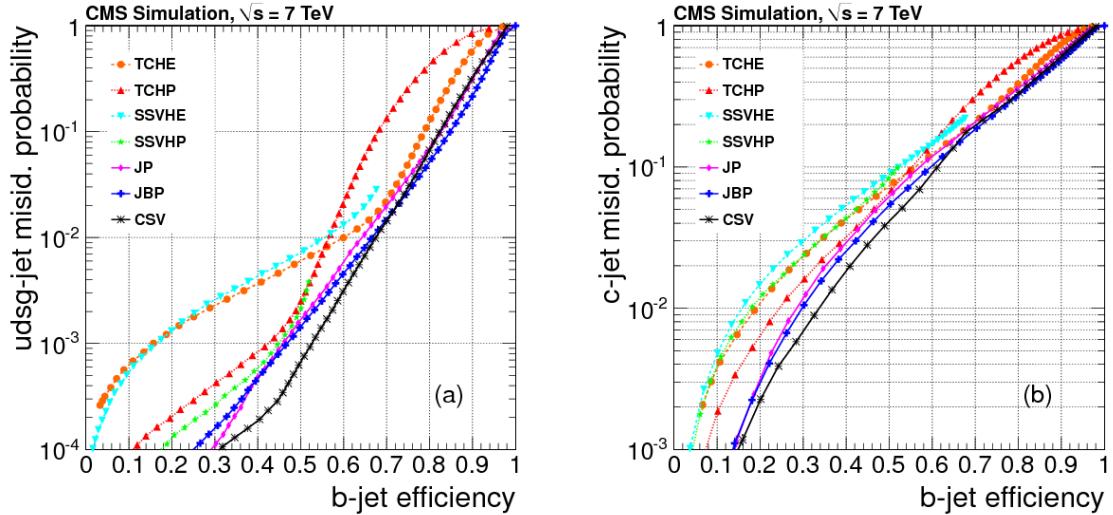


Figure 4.3: Performance curves obtained from simulation for the algorithms described in the text. (a) light-parton- and (b) c-jet misidentification probabilities as a function of the b-jet efficiency. Taken from [92].

reconstructed. It utilizes an approach combining information from secondary vertices as well as track-based lifetime information and thus is able to exceed the efficiency of SSV algorithms. Often, pseudo-vertices can be formed from tracks even when failing the reconstruction of an actual secondary vertex which allows to derive some secondary vertex related quantities. Variables used in the CSV algorithm are e.g. flight distance significance, vertex mass, number of tracks at the vertex, number of tracks in the jet or the IP significances for the tracks in the jet. These variables are used to compute two likelihood ratios which can be used to distinguish either c and b jets or light-parton and b jets.

Each algorithm determines a discriminator value per jet indicating how b-jet-like a jet behaves. Based on that, working points are defined corresponding to a specific minimum threshold of the discriminator value. These working points are named *loose*, *medium* and *tight* and correspond to a misidentification probability, i.e. the probability to identify a non-b jet as b jet, of 10%, 1% and 0.1% for an average jet momentum of 80 GeV, respectively.

In order to determine the quality of a particular b-tagging algorithm, typically the misidentification probability as function of the b-tag efficiency is compared for various taggers. Such a performance comparison is illustrated in Fig 4.3 for the tagging algorithms described above. The misidentification probability is derived separately for light-flavour and gluon initiated jets as well as c-jets. The curves are derived from simulated multijet events using jets with $p_T > 60$ GeV. For loose working points the b-tag efficiency is around $\approx 80 - 85\%$ and the JBP algorithm shows the best performance. In case of medium and tight selections, the b-tag efficiency drops to $\approx 45 - 55\%$ and the CSV algorithm performs best.

B-tagging algorithms used for analyses of data obtained at $\sqrt{s} = 8$ TeV in 2012 where the

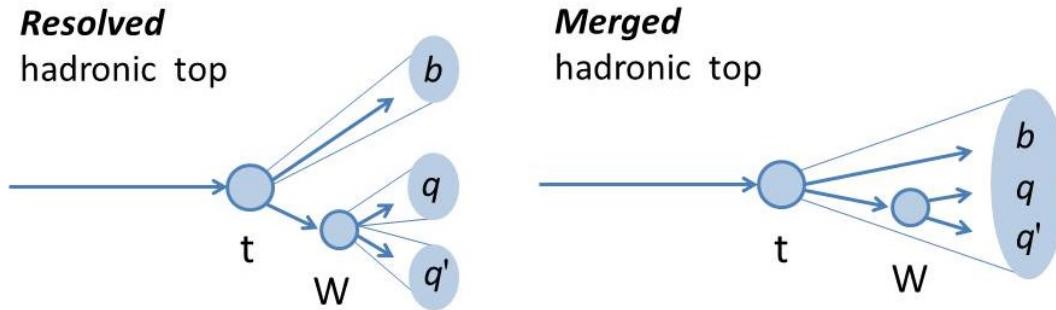


Figure 4.4: Schematic diagrams of the decay of a top-quark in the resolved case (*left*) and the boosted scenario (*right*).

TCHP, JP and CSV algorithm [93]. In order to compare the b-tagging performance in data and simulation two different event samples have been selected. On the one hand, an inclusive multijet sample is selected requiring at least one jet with $p_T = 60 - 500$ GeV which is dominated by jets from light-flavour quarks and gluons. On the other hand, a sample dominated by top-pair production is selected requiring an electron, a muon and at least two jets with $p_T > 30$ GeV such that this sample is enriched in b-jets. These samples allow to compare quantities relevant for b-tagging in data and simulation like e.g. the number of secondary vertices or b-tag discriminator values. In general, quantities relevant for b-tagging show a good agreement between data and simulation with deviations within 20%. Typically, efficiencies and misidentification probabilities are measured as function of jet transverse momentum and pseudorapidity. A potential disagreement between data and simulation can be expressed in terms of *scale factors* and corrected for in analyses such that b-tagging efficiency and misidentification probability in data and simulation match.

4.4 Identification of Boosted t-Quark Jets

Supersymmetric models or other scenarios describing physics beyond the Standard Model predict the existence of new massive particles. Often, the coupling of these particles especially to quarks belonging to the third generation is sizeable like e.g. in decays of top squarks which predominantly are expected to decay into top quarks. Consequently, such processes lead to highly-energetic top-quarks in the final state which can be identified exploiting the specific properties of the top-quark.

The top-quark is the heaviest quark with a mass of 173.34 ± 0.76 GeV [94] resulting in a rapid decay. This decay happens even faster than the hadronisation timescale is such that the top-quark behaves in the detector as a bare quark without forming color-neutral hadrons. The decay takes place via the weak interaction and as denoted in Sec. 2.1, the quark mixing is parametrized by the CKM-matrix. Since the corresponding matrix element $V_{tb} \approx 1$, the top quark decays thus almost exclusively into a W boson and a b-quark. Therefore, the decay of the top-quark is strongly characterized by the decay of the W boson. In general, the W boson can decay into a charged lepton and its corresponding neutrino or a pair of light quarks which are up, down, charm and strange. Taking into

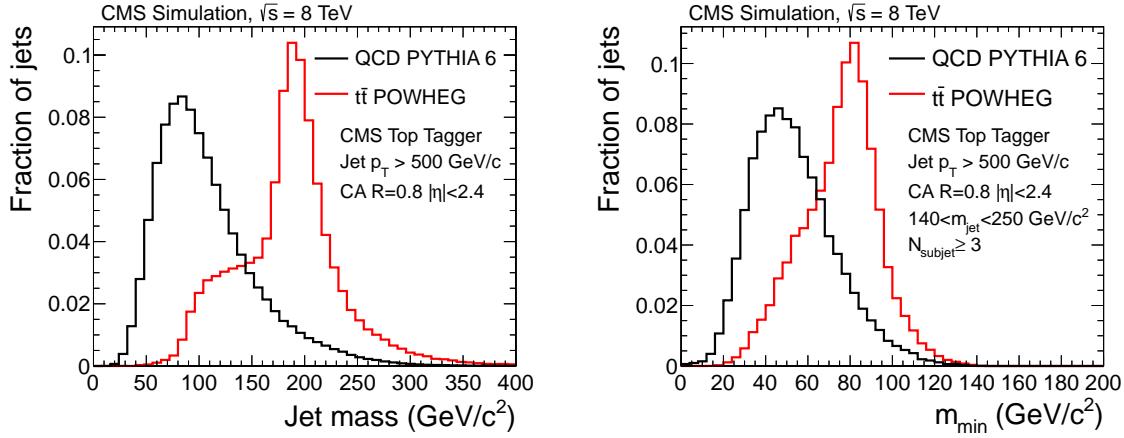


Figure 4.5: Jet mass (*left*) and minimum pairwise subjet mass (*right*) for CA8 jets with $p_T > 500 \text{ GeV}$ from a simulated $t\bar{t}$ POWHEG sample (red) and from a simulated QCD PYTHIA sample (black). Taken from [95].

account the three possible colour states for each quark pair this results in nine different W decay modes. Consequently, two thirds of top quark decays result exclusively in hadrons which is typically referred to as *hadronic top*. If top quarks are produced with low energy ($< 2m_t$), the decay will predominantly occur at rest and the top decay products show up as three distinct signatures in the detector. In the case of hadronic tops these are three well separated jets. However, if the top transverse momentum is high, the decay products are *boosted* and thus collimated in the forward direction. Consequently, they might overlap and merge into a single large jet (*fat jet*). The opening angle of the decay products ΔR is expected to scale as

$$\Delta R \approx 2m/p_T \quad (4.4)$$

where m and p_T are the mass and the transverse momentum of the decaying particle, respectively. Schematic diagrams of resolved and boosted top quark decays are illustrated in Fig. 4.4. The identification of boosted top-quark decays – here restricted to hadronic tops – is typically known as *top tagging* and aims at the determination of the decay

$$t \rightarrow W + b \rightarrow qq' + b \quad (4.5)$$

by analysing the substructure of fat jets and applying kinematic selections.

Within the CMS experiment several top tagging algorithms (or short *taggers*) are commissioned [95]. Two of them are described in some detail in the following:

CMS Top Tagger: The CMS Top Tagger [96] is based on the Top Tagger developed by Kaplan et al. [97] and acts on jets clustered by the Cambridge-Aachen algorithm with distance parameter $R = 0.8$ (CA8 jet). These jets used as input for the algorithm are denoted as *hard jets*. Since the decay products of the hadronic top are not expected to be all collimated in one jet with $R = 0.8$ for low transverse momenta, only jets with $p_T^{jet} > 350 \text{ GeV}$ are considered. In order to identify subjets within the hard jets, a two stage decomposition procedure is applied. First, the algorithm aims at

jet
grooming

Ref DP
note

splitting the hard jet into two subclusters (*primary decomposition*) and second, it is attempted to further split the clusters emerging from the first step (*secondary decomposition*). Hence, the pairwise clustering sequence used to form the hard jet is performed in reverse order to identify subclusters. Typically, subjets are found when they are spatially well separated and carry a significant fraction of the momentum of the hard jet. Details on the actual splitting criteria can be found in [95]. With this approach up to four individual subjets are identified within the hard jet. After a successful decomposition procedure kinematic criteria can be applied to the identified subjets in order to tag top jets. In the CMS top-tagging algorithm – as employed in this thesis – the following selections are used:

- Number of subjets ≥ 3
- The jet mass m_{jet} , i.e. the mass of the four-vector sum of the constituents of the hard jet, has to be close to the top-mass: $m_{\text{jet}} = 140 - 250 \text{ GeV}$.
- The invariant mass of each pair of the three subjets highest in p_T is calculated. The minimum of the pairwise masses m_{\min} has to be close to the W mass: $m_{\min} > 50 \text{ GeV}$.

In Fig. 4.5, the distributions for jet mass and minimum pairwise mass are illustrated for a simulated $t\bar{t}$ (red) and QCD multijet (black) sample at $\sqrt{s} = 8 \text{ TeV}$.

HEP Top Tagger: The HEP Top Tagger [98] uses jets with an even larger distance parameter than the CMS Top Tagger of $R = 1.5$ (CA15 jets). This makes the HEP top-tagging algorithm especially suitable for tops with moderate boost. Fat jets with a transverse momentum greater than 200 GeV are used as input for the algorithm. Since a larger jet size is in principle more prone to disturbing effects from the underlying event or pile-up, a sophisticated decomposition procedure is applied to distinguish hard subjets from soft components. Similar to the decomposition done for the CMS tagger, the identification of subjets is based on going through the cluster history of a jet in reversed order. First, the jet is decomposed into subclusters applying a *mass drop condition* discarding too soft components. Afterwards, all combinations of three clusters resulting from the mass drop decomposition are reclustered into subjets. For each combination the subjets are filtered by keeping only the five subjets highest in p_T . Finally, the combination with the mass determined from the filtered subjets closest to the top mass is kept and reclustered to force three subjets. Details on the mass drop decomposition and reclustering criteria can be found in [95]. Kinematic selections are applied to these three final subjets in order to identify the top jets. The following quantities are used based on the invariant mass of combinations of subjets:

- The invariant mass of the sum of the four-vectors of the three subjets is required to be in the top mass window: $m_{123} = 140 - 250 \text{ GeV}$.
- In order to select the W mass the jet has to satisfy at least one of the following conditions based on the subjet pairwise masses

$$0.2 < \tan \frac{m_{13}}{m_{12}} < 1.3 \text{ and } R_{\min} < \frac{m_{23}}{m_{123}} < R_{\max}$$

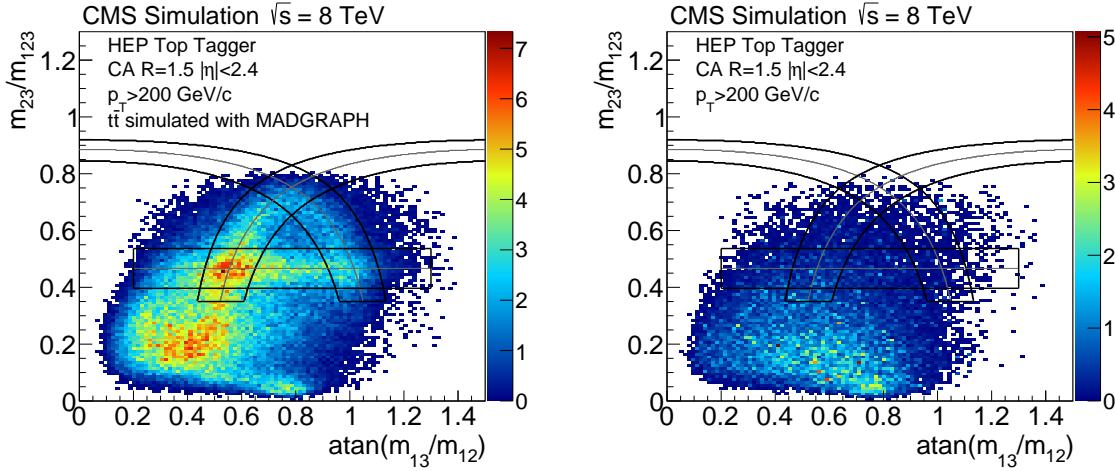


Figure 4.6: Two-dimensional distributions of m_{23}/m_{123} versus $\text{atan}(m_{13}/m_{12})$ for HEP Top Tagger subjets from CA15 jets with $p_T > 200 \text{ GeV}$ for a simulated $t\bar{t}$ MADGRAPH sample (*left*) and for a simulated background sample composed of cross-section weighted boson + jets, diboson, single-top, $t\bar{t}$ all-hadronic and $t\bar{t}$ leptonic events (*right*). The A-shaped region indicates the selected region by the HEP Top Tagger. Taken from [95].

$$R_{\min}^2(1 + (\frac{m_{13}}{m_{12}})^2) < 1 - (\frac{m_{23}}{m_{123}})^2 < R_{\max}^2(1 + (\frac{m_{13}}{m_{12}})^2) \text{ and } \frac{m_{23}}{m_{123}} > 0.35$$

$$R_{\min}^2(1 + (\frac{m_{12}}{m_{13}})^2) < 1 - (\frac{m_{23}}{m_{123}})^2 < R_{\max}^2(1 + (\frac{m_{12}}{m_{13}})^2) \text{ and } \frac{m_{23}}{m_{123}} > 0.35$$

where the indices of m indicate the rank of the considered subjets with respect to the transverse momentum, $R_{\min} = (1 - f_W) \times m_W/m_t$ and $R_{\max} = (1 + f_W) \times m_W/m_t$ for the W mass width chosen to be $f_W = 0.495$.

The criteria related to the selection of the W mass are illustrated in Fig. 4.6 for signal events as well as background events and indicated by the black solid lines resulting in an A-shaped region.

In analogy to the b-tagging algorithms, different working points for each top tagger can be defined characterized by a specific top-tag efficiency and misidentification rate. Typically, the working points are chosen such that they have a minimum mistag rate for a given signal efficiency. Commonly, the top-tag efficiency for simulated events is defined as the number of jets passing the top-tagging selection divided by the number of jets spatially matched to a simulated hadronic top or anti-top passing a certain p_T selection. Similarly, the mistag rate is defined as the number of jets passing the top-tagging selection divided by the number of jets spatially matched to a simulated quark or gluon from the hard process passing the p_T selection. The top-tag efficiency determined from simulated $t\bar{t}$ POWHEG events for the selection criteria described above amounts to 35.3% for a matched parton- p_T

of > 200 GeV for the HEP Top Tagger and 38.3% for a matched parton- p_T of > 400 GeV for the CMS Top Tagger while the mistag rates as determined from simulated PYTHIA QCD multijet events are 2.6% and 2.5%, respectively. However, these efficiencies and mistag rates show a moderate dependence on pile-up such that in high pile-up environments a performance degradation is expected. For instance, in the case of the CMS Top Tagger the efficiency decreases as function of the number of primary vertices with a slope of $0.031\% \pm 0.034\%$ while the mistag rate increases with a slope of $0.095\% \pm 0.006\%$.

5 Measurement of the Jet Transverse-Momentum Resolution

Many measurements of standard model properties or searches for new physics beyond the standard model performed within the CMS experiment rely on events with jets in the final state. Hence a good understanding of jet properties, like for instance the resolution of the jet transverse momentum is of major importance and a crucial ingredient for such kind of analyses. For example, many new physics searches are carried out based on jet final states. Here, QCD multijet events can fake the signature of possible new physics events and constitute a background process, since a mismeasurement of the jet momenta due to the limited detector resolution or the decay of heavy flavour quarks leads to a momentum imbalance in the event and consequently to measurable missing energy. The knowledge of the jet resolution is hence keypoint in the prediction of such background contributions, as discussed in Chapter 6.

In this chapter, an analysis is presented where the jet- p_T resolution in data and in simulated events is derived. The method is based on momentum conservation in the transverse plane of dijet events and offers the possibility to cover a large phase space in p_T and η . A similar approach was already used in previous studies at $\sqrt{s} = 7\text{ TeV}$ [89, 99] while a complementary approach utilizes $\gamma + \text{jet}$ events [100–102]. The measurement shown here is based on collision data corresponding to an integrated luminosity of 19.7 fb^{-1} recorded at $\sqrt{s} = 8\text{ TeV}$ in 2012.

5.1 Components of the Jet Response

As introduced in Section 4.2.3 the transverse momentum of a jet at detector level does not equal the momentum of the corresponding particle-level jet which is typically expressed in the jet response. In simulated events the transverse momentum of the particle-level jet is given by the p_T of the generated jet which is clustered from all stable particles after hadronisation including neutrinos. Thus, the *MC-truth response* can be determined as

$$R = \frac{p_T}{p_T^{\text{gen}}} . \quad (5.1)$$

In Fig. 5.1 an example for a jet response distribution derived from simulation is shown. It is obtained from a QCD multijet sample generated with PYTHIA 6 tune Z2 [103] using the CTEQ6L1 PDF set [104] and processed with the full detector simulation. Since the cross-section of the process has been scaled by $\hat{p}_T^{4.5}$ with the scale parameter \hat{p}_T describing the momentum transfer in the hard process, the sample is reweighted with the inverse in order to regain the physical spectrum. For the calculation of the truth response, the two generated jets in the event with highest transverse momentum are selected. The event is rejected, if one of these two generator jets does not have a corresponding reconstructed jet within a distance of $\Delta R < 0.25$. Otherwise the response is calculated in intervals of

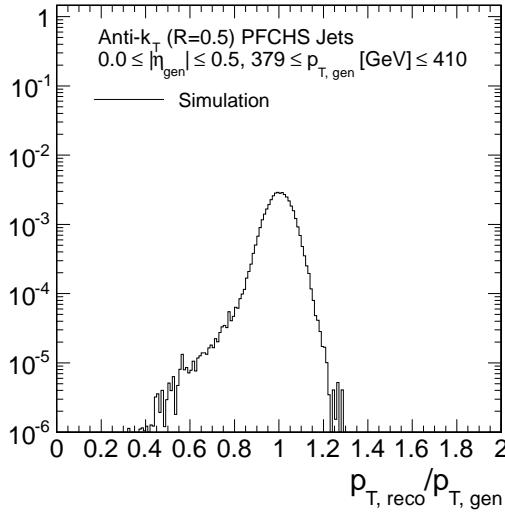


Figure 5.1: Jet response function for one example $|\eta_{\text{gen}}|$ and $p_{T,\text{gen}}$ interval derived from simulation.

p_{T}^{gen} and $|\eta^{\text{gen}}|$ since a dependence on the momentum and the respective detector region is expected, as discussed in Section 4.2.3. This is done after applying the jet energy corrections discussed in Section 4.2.4 to the detector-level jet momenta.

Apparently the jet response consists generally of two components – the Gaussian-shaped core around the mean referred to as jet resolution and non-Gaussian components referred to as *tails*. The core of the response is mainly caused by the intrinsic resolution of the various sub-detector components and the precision of the PF and jet clustering algorithms. The response tails however are predominantly caused by severe jet-momentum mismeasurements. These can be due to detector effects like shower leakage and detector noise or also to real physics processes. For instance the lower response tails get populated by semi-leptonic decays of heavy-flavour quarks. These contain neutrinos that carry a certain amount of the momentum and leave the detector unnoticed .

more info
on in-
trinsic
resolution

5.2 Basic Concept of the Dijet Asymmetry Method

As introduced in Section 4.2.3 and 5.1 the jet transverse momentum resolution corresponds to the Gaussian-shaped core part of the jet response. In simulated events the particle-level jet momenta are given by the generator-level jet momenta. In data events however, no such equivalent is present so that the jet resolution is not accessible directly.

One possibility to measure the resolution of the jet transverse momenta in data as well as in simulated events is to utilize the dijet asymmetry A . For events with at least two jets it is defined as

$$A = \frac{p_{T,1} - p_{T,2}}{p_{T,1} + p_{T,2}}. \quad (5.2)$$

In this equation $p_{T,1}$ and $p_{T,2}$ correspond to the randomly ordered transverse momenta of the two leading jets.

For a sufficient number of events the asymmetry is approximately normally distributed

and the standard deviation is given as

$$\sigma_A = \left| \frac{\partial A}{\partial p_{T,1}} \right| \cdot \sigma(p_{T,1}) \oplus \left| \frac{\partial A}{\partial p_{T,2}} \right| \cdot \sigma(p_{T,2}) \quad (5.3)$$

In an ideal dijet topology the two jets are exactly balanced at particle level. If they belong in addition to the same η region, then $\langle p_{T,1} \rangle = \langle p_{T,2} \rangle = \langle p_T \rangle$ and $\sigma(p_{T,1}) = \sigma(p_{T,2}) = \sigma(p_T)$. This allows the simplification of Eq. 5.3 and provides the following important relation between the width of the asymmetry σ_A and the jet- p_T resolution $\sigma(p_T)$

$$\frac{\sigma(p_T)}{\langle p_T \rangle} = \sqrt{2} \cdot \sigma_A \quad (5.4)$$

Already at the Tevatron experiments [105, 106], the ATLAS experiment [107] or in previous CMS analyses as stated above this relationship was utilized to measure the jet resolution from dijet events.

5.3 Application to Realistic Collision Events

The measurement of the jet transverse momentum resolution in collision events is based on Eq. 5.4. As discussed in Section 4.2.3 the resolution is a function of p_T and η . Consequently, the asymmetry distribution is a function of p_T and η as well. Typically, the jet- p_T spectrum is affected by migration effects. Due to the limited resolution a particular interval of reconstructed jet momenta is populated not only by jets whose true momentum belongs to that bin, but also from jets outside. In case of a steeply falling spectrum, as this is the case for the jet momenta, more jets are supposed to migrate into a specific interval than out. Consequently, the measured response is systematically higher and the measured relative response gets biased in favor of the object with the worse resolution. In order to reduce this resolution bias in the analysis, the measurement is performed in intervals of the average momentum of the two leading jets in the event

$$p_T^{\text{ave}} = \frac{1}{2}(p_{T,1} + p_{T,2}) . \quad (5.5)$$

Beyond that the ideal dijet topology with exactly two jets being perfectly balanced is interfered with additional effects in realistic collision events. Very often further jet activity is occurring as momentum of the hard scattering process is transferred to soft-particles or jets arising from initial or final state radiation leading to momentum imbalance in the event. This additional jet activity can in good approximation be described by the variable α which is defined as the ratio of the transverse momentum of the third jet to the average momentum according to

$$\alpha = \frac{p_{T,3}}{p_T^{\text{ave}}} . \quad (5.6)$$

The presence of further jets is neglected in the parametrization of the additional activity in the event, as these have consecutively less and less high momentum due to the strongly decreasing jet production cross section versus jet- p_T [108]. The presence of additional jets and the thereby introduced imbalance leads to a broadening of the observed asymmetry distribution. This effect is also illustrated later in Section 5.4.2. In order to determine the

Table 5.1: Trigger paths with offline p_T^{ave} thresholds at which the trigger efficiency reaches the 99% efficiency plateau. Thresholds are given for PFCHS jets.

Trigger	p_T^{ave} threshold [GeV]
HltDiPFJetAve40	62
HltDiPFJetAve80	107
HltDiPFJetAve140	175
HltDiPFJetAve200	242
HltDiPFJetAve260	310
HltDiPFJetAve320	379
HltDiPFJetAve400	467

intrinsic resolution from such events the measured resolution has to be corrected for this bias.

A further source of momentum differences between the particle-level and the detector-level jet leading to an overall momentum imbalance in an event is arising from out-of-cone showering effects in the jet-clustering procedure. Typically, some particles might be too soft to be included in the clustered jet. Furthermore, additional contributions from pile-up or the underlying event might be wrongly associated to a jet. Such effects require a correction to the resolution measurement as well. The actual procedure how to determine and apply the required corrections is discussed in Section 5.6.

5.4 Samples and Event Selection

5.4.1 Datasets and Triggers

In this analysis, multijet events from pp collisions are considered which have been recorded in 2012 with the CMS detector at $\sqrt{s} = 8$ TeV. From these data only those are considered where all subdetectors have been reliably operating. The collected data sample used in this analysis corresponds to an integrated luminosity of 19.7 fb^{-1} with an uncertainty of 2.5% (syst.) + 0.5% (stat.) [109].

Multijet events are pre-selected by a set of triggers based on the average transverse momentum of the two leading jets in the event. In order to obtain a good coverage of the p_T^{ave} spectrum, different trigger paths are combined. Since they have different minimum p_T^{ave} thresholds, a broad range in p_T^{ave} is considered. In Tab. 5.1 the different trigger paths used for this analysis are listed and the particular offline p_T^{ave} value for which the respective trigger reaches 99% of the efficiency plateau is given (taken from [110]).

The simulated event sample is generated with PYTHIA, as discussed in Section 5.1 .

details
trigger
selection
+ eff

Table 5.2: Overview of the $|\eta|$ and p_T^{ave} interval boundaries used for the resolution measurement.

$ \eta $	p_T^{ave} [GeV]
0.0, 0.5, 1.1, 1.7, 2.3, 2.8, 3.2, 5.0	62, 107, 175, 205, 242, 270, 310, 335, 379, 410, 467, 600, 1000, ∞

5.4.2 Selection Criteria

The physics objects used in the analysis are reconstructed with the particle-flow algorithm [78]. In order to reduce a possible impact from pileup, charged hadrons associated with a pileup vertex are removed from the event. This is referred to as charged-hadron subtraction (CHS). All jets – in data as well as in simulation – are clustered with the anti- k_T algorithm using a distance parameter of $R = 0.5$. In addition, jets are calibrated by applying jet energy corrections containing residual corrections for data. The event reconstruction is done using the CMSSW framework series CMSSW_5_3_X with the global tag FT_53_V21_AN6 for data and START53_V27 for simulation.

The two leading jets in each event have to fulfill loose jet identification criteria . If there is a third jet in the event, it has to fulfill this requirement as well. In the analysis a minimum p_T threshold of 10 GeV is introduced for the first three leading jets. If an event has no third jet with p_T above 10 GeV, the event is rejected in order to mitigate effects from pile-up.

In order to enrich the sample with events close to the ideal dijet topology with two jets pointing into opposite directions in the transverse plane and to reject events with very large asymmetries the two leading jets have to fulfill

$$|\Delta\phi| > 2.7 \text{ with } \Delta\phi = \Delta\phi(p_{T,1}, p_{T,2}). \quad (5.7)$$

As the resolution is a function of the transverse momentum and $|\eta|$ the asymmetry is calculated in various intervals of $|\eta|$ and p_T^{ave} which are summarized in Tab. 5.2. The p_T^{ave} intervals are chosen such that the interval boundaries correspond to the offline p_T^{ave} -thresholds of the various different triggers. This ensures that the intervals belong to only one trigger. If a trigger provides enough statistical precision, the interval is separated into two.

The $|\eta|$ intervals are chosen such that they correspond to different sub-detector components, if possible. In the forward region of the detector, mainly jets with low transverse momentum are expected corresponding to triggers with low p_T^{ave} thresholds. Since these are highly prescaled only few events populate the high $|\eta|$ region and prohibit a finer η -binning there.

As discussed in Section 5.3 events with a topology of exactly two high- p_T jets and no further jet activity are almost not existent in realistic collisions. To select a sample which offers many events that maintain a dijet-like structure, a maximum α threshold of $\alpha_{\max} = 0.25$ is introduced.

In order to account for the α dependence of the measured asymmetry distributions the p_T^{ave} and $|\eta|$ intervals for the resolution measurement are further subdivided in various α intervals. The upper boundaries of the α_{\max} intervals are 0.1, 0.125, 0.15, 0.175, 0.2, 0.225, 0.25.

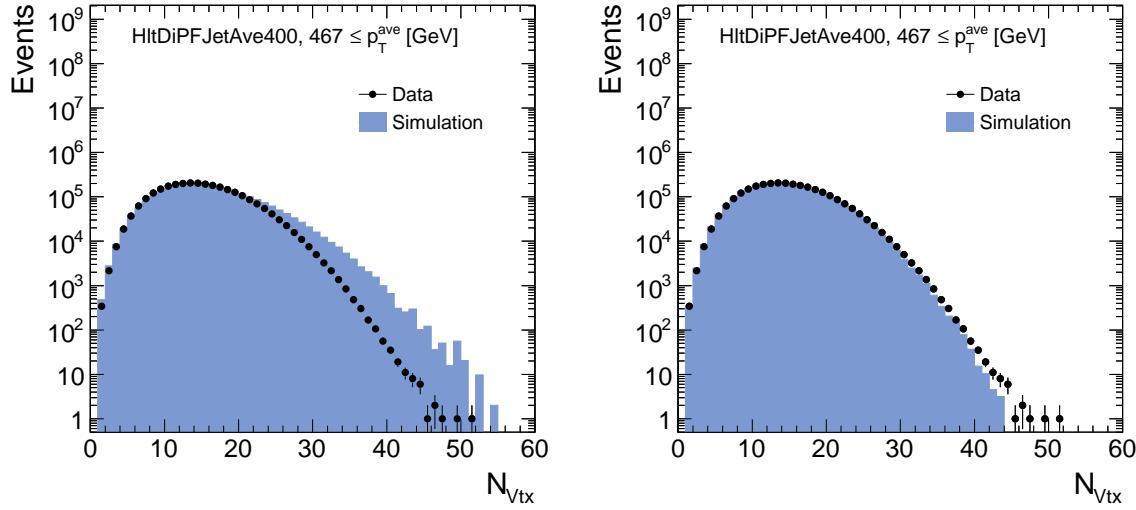


Figure 5.2: Distribution of number of primary vertices in data and simulation before (left) and after (right) reweighting of the pileup scenario in simulation for trigger path HltDiPFJetAve400.

Each interval ranges from $\alpha = 0.0$ to the respective upper boundary.

For the simulated sample two further things have to be considered related to the modelling of pileup. The first problem is that sometimes the \hat{p}_T of an additional pileup event is larger than the \hat{p}_T of the hard interaction resulting in events with extremely large weights. In order to protect the sample against such cases simulated events are rejected, if the p_T of one of the two leading jets is larger than $2 \cdot \hat{p}_T$.

The second issue is that the QCD sample is generated with a certain estimated pileup scenario. This is not necessarily the same as the observed pileup distribution in data. Therefore the simulated events are reweighted to match the pileup conditions in data. The pileup distribution in data is different for each individual trigger. This is due to the fact that the pileup conditions changed within the running periods in 2012 and due to the different prescales of the individual triggers which also changed over time each trigger path has an individual observed pileup distribution. Therefore each simulated event is reweighted to match the pileup scenario in the data distribution of that particular trigger path to which the simulated event would belong depending on the p_T^{ave} of the event. This procedure can be checked by comparing the distribution of the number of primary vertices in data and simulation before and after reweighting since the number of primary vertices is a measure for the pileup activity in an event. The result is shown in Fig. 5.2 for the trigger path with the highest p_T^{ave} threshold used in the analysis and shows a good agreement especially in the bulk of the distribution. The distributions of primary vertices before and after reweighting the pileup scenario of the other trigger paths relevant for the analysis are shown in Appendix . The pileup weight is considered as a multiplicative factor for each simulated event.

The resulting inclusive p_T^{ave} spectrum of events after the described selection in data and in simulation as well as the p_T spectra of the first three leading jets are shown in Fig. 5.3. The number of simulated events in each trigger p_T^{ave} interval is normalized to the respective

app:pu
reweight

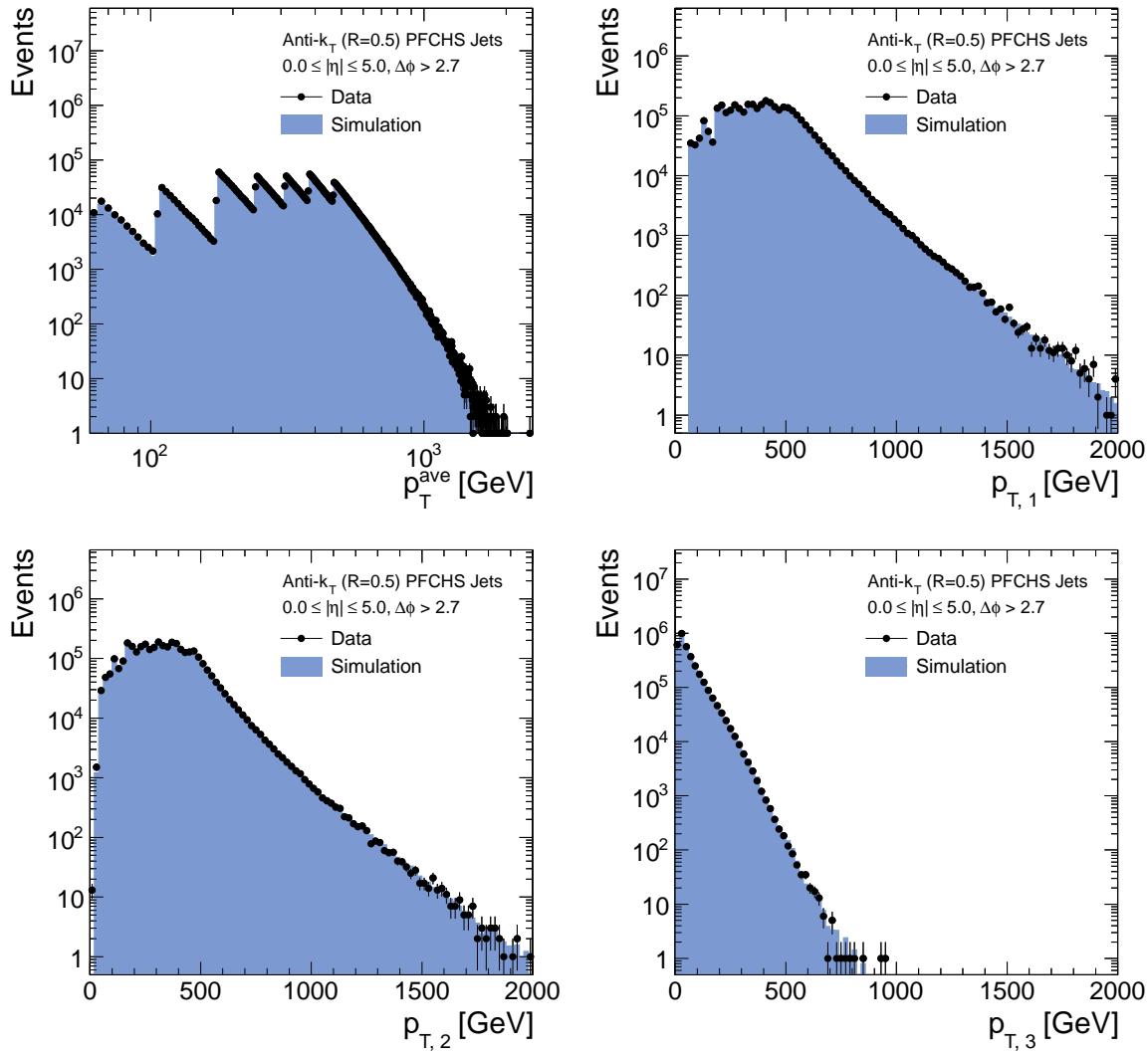


Figure 5.3: Inclusive p_T^{ave} spectrum of events after the total selection in data and in simulation (top left), leading jet- p_T (top right), subleading jet- p_T (bottom left) and third jet- p_T (bottom right).

integral in data.

Finally, some exemplary asymmetry distributions obtained after the dedicated selection described here are depicted in Fig. 5.4 for data and simulated events. Since the relation between resolution and width of the asymmetry distribution as expressed in Eq. 5.4 is based on the assumption that the asymmetry is normally distributed, only asymmetry distributions with at least 100 events in the specified p_T^{ave}, η and α intervals are considered for the analysis.

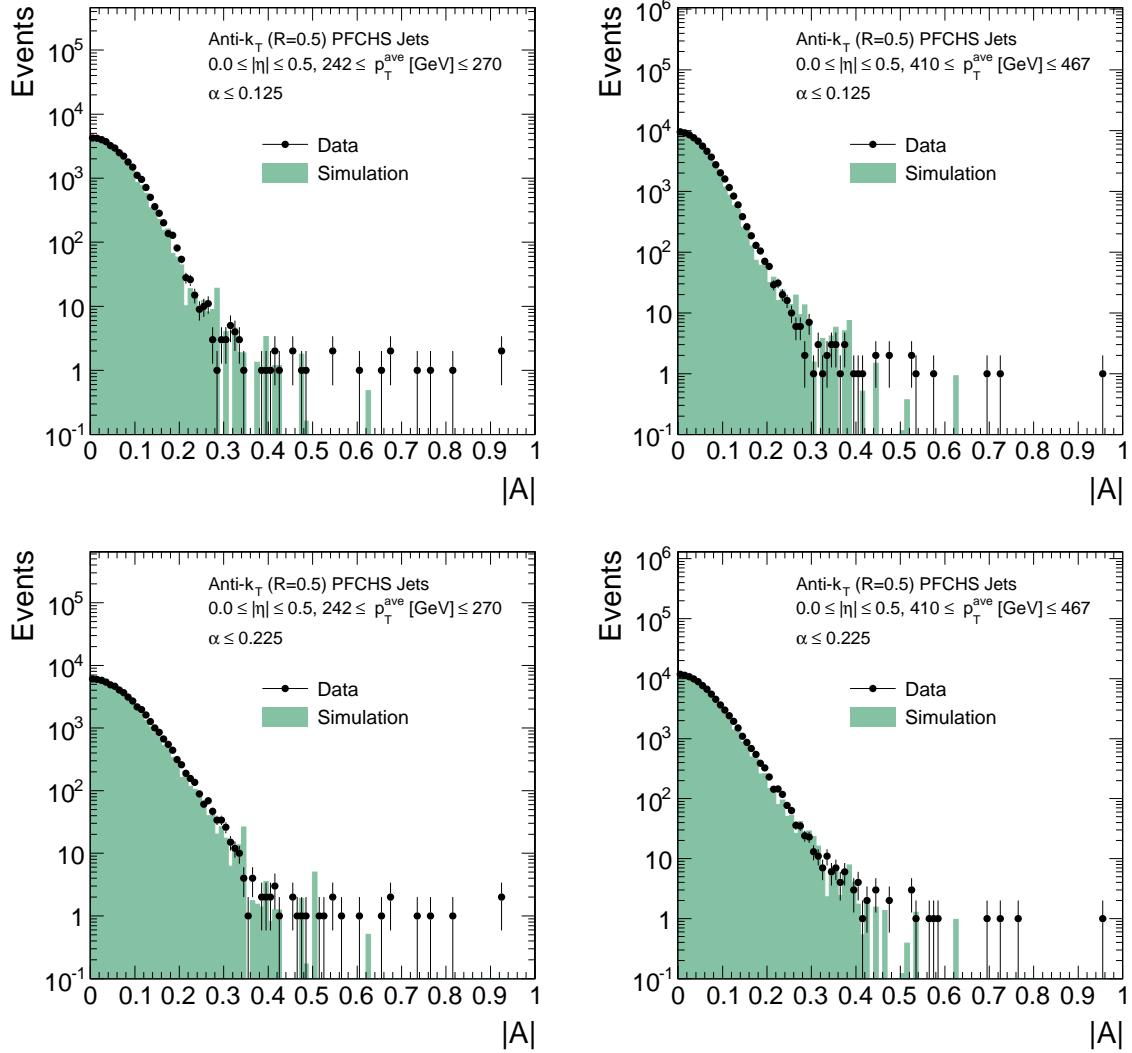


Figure 5.4: Some example asymmetry distributions for the lowest $|\eta|$ region, two medium p_T^{ave} intervals and for a low (top row) and a high (bottom row) α interval.

5.5 Definition of the Asymmetry Width

As the measurement of the resolution in collision data is based on the width of the dijet asymmetry distribution, a proper definition of the asymmetry width is of major importance.

From the tails of the detector response events with large asymmetries can emerge leading to non-gaussian tails in the asymmetry distributions. These shall not be considered in the calculation of the asymmetry width in order to not bias the determination of the asymmetry width one is interested in for the measurement of the resolution. Hence the asymmetry width has to be defined such that the central part of the distribution is approximated reasonably well by a gaussian distribution. In Fig. 5.5 the asymmetry distribution in data and in simulation for a certain $|\eta|$, p_T^{ave} and α interval is compared to gaussian distributions where the standard deviation of the gaussian function is chosen such that it corresponds to different asymmetry widths. The width of the asymmetry is determined by taking the whole distribution into account or by truncating a certain percentage in the tail regions. It is visible that the core part of the asymmetry distribution can in good approximation be described by a gaussian function if the tails are neglected in the calculation of the asymmetry width as long as the truncation is not rejecting parts of the gaussian core. Thus the asymmetry width is in this analysis calculated as 98.5%-truncated root-mean-square

$$\sigma_A = \text{RMS}_{98.5\%} = \sqrt{\frac{1}{\sum_i y_i} \sum_i y_i \cdot A_i^2} \quad (5.8)$$

where y_i denotes the frequency of the asymmetry value A_i and the sum over i includes all values such that 98.5% of the total asymmetry distribution are covered. The statistical uncertainty is given by

$$\Delta\sigma_A = \Delta\text{RMS}_{98.5\%} = \frac{\text{RMS}_{98.5\%}}{\sqrt{2 \cdot n_{\text{eff}}}} \quad (5.9)$$

with the number of effective entries n_{eff} in the specified interval for which the asymmetry distribution has been derived.

5.6 Corrections to the Dijet Asymmetry

The fundamental relation between the width of the asymmetry distribution and the jet energy resolution as expressed in Eq. 5.4 holds in this form only for the case of an ideal dijet topology. In real collision events various effects can occur that disturb the exact balance of the two jets. Such effects can be soft radiation or additional jets originating from the hard scattering. They lead to imbalance in the event and hence broaden the observed asymmetry distribution resulting in an increased measured resolution. In order to determine the intrinsic resolution, the measured one has to be corrected for such effects. These corrections are explained in the following Sections 5.6.1 and 5.6.2.

5.6.1 Correction for Additional Jet Activity

In addition to the two jets originating from the hard interaction further jets can occur in an event from the emission of additional partons or pileup. The additional jet activity in the event is quantified by α introduced in Eq. 5.6 which is the fractional p_T of the third

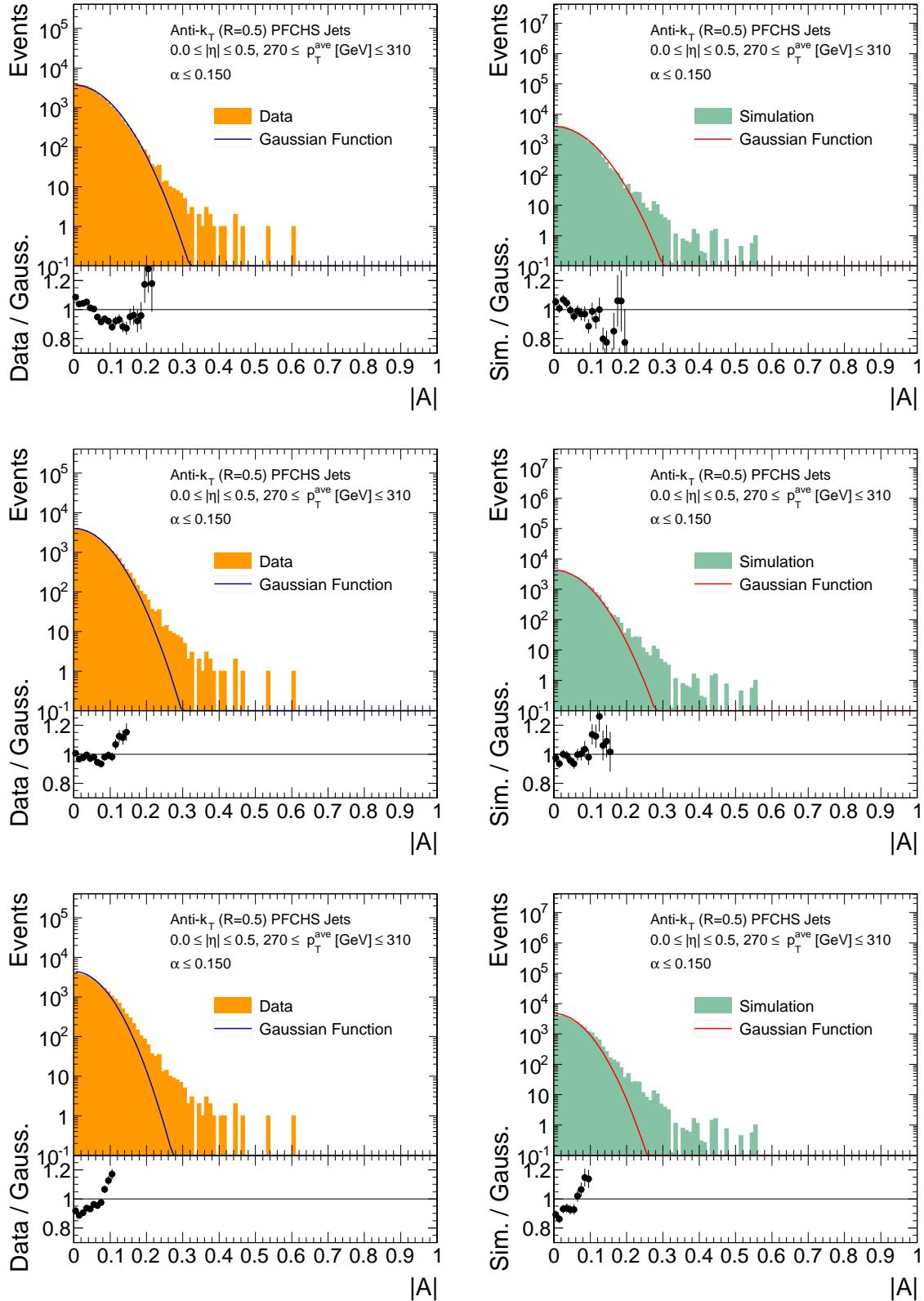


Figure 5.5: Asymmetry distributions for data (left) and simulation (right) compared to gaussian distributions obtained with standard deviations corresponding to 0% truncation (top), 1.5% truncation (middle) and 5% truncation (bottom) of the tail regions.

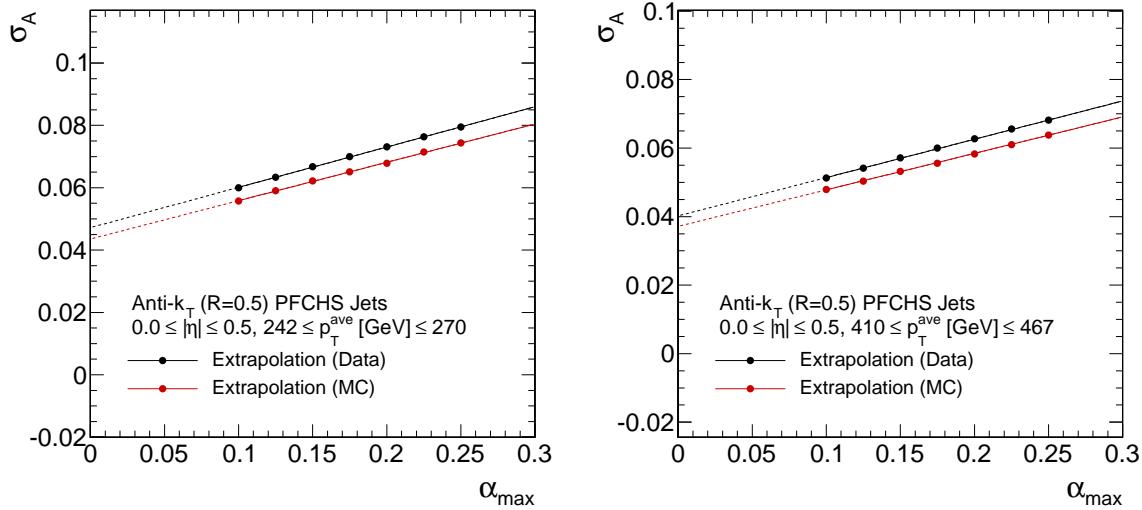


Figure 5.6: Two exemplary extrapolations of measured values for σ_A in data and simulation to obtain the result at zero additional jet activity.

jet. The impact from jets beyond the third one is neglected since there impact is expected to be small due to the strongly decreasing jet-production cross-section.

The imbalance contribution arising from additional jet activity is taken care of by an extrapolation procedure. As described in Section 5.4.2 the asymmetry distribution is calculated for several intervals in $|\eta|$ and p_T^{ave} with different selections on the maximum value of α ($= \alpha_{max}$). For each of these individual selections the width of the asymmetry is determined as stated in Section 5.3. The measured values of $\sigma_A(\alpha_{max})$ are extrapolated to $\alpha_{max} \rightarrow 0$ assuming a linear behaviour. The y-intercept of the fitted linear function represents the resolution without further jet activity in the event while the statistical uncertainty is given by the respective fit error. The extrapolation procedure for two exemplary $|\eta|$ and p_T^{ave} intervals is illustrated in Fig. 5.6 (further extrapolations are shown in Appendix).

As stated in Section 5.4.2 the selection is performed in inclusive intervals of $\alpha < \alpha_{max}$. This results in a correlation of the measured values of σ_A for a particular p_T^{ave} interval. In order to obtain a realistic estimate of the statistical uncertainty the correlation is propagated to the extrapolation fit.

In some more detail this means that the measured data points shall be described by a linear function $y = a \cdot x + b$. Thus the parameters a and b have to be determined for known x , y and covariance matrix C by minimizing the following χ^2

$$dy^T \cdot C^{-1} \cdot dy \quad (5.10)$$

where $dy = y_{\text{measured}} - y_{\text{predicted}}$. In this analysis the following relations apply $y = \sigma_A(\alpha_{max})$, $x = \alpha_{max}$ and $b = \sigma_A(\alpha_{max} \rightarrow 0)$.

Under the assumption that all events that belong to α interval i are also completely

appendix:
extrapol

included in the next higher α interval j ¹ the covariance matrix is given by

$$C_{i,j}(\sigma_{A_i}, \sigma_{A_j}) = (\Delta\sigma_{A_i})^2 \frac{\sigma_{A_i}}{\sigma_{A_j}} \left(\frac{n_i}{n_j} \right)^2 \quad (5.11)$$

where n is the number of events in that particular interval. The function minimization itself is performed with the Minuit package [111].

In Fig. 5.7 an overview of a goodness-of-fit test is shown for the extrapolation fits in data and simulation. The resulting χ^2 distribution over the number of degrees of freedom is shown as function of p_T^{ave} for each individual $|\eta|$ interval. In general the fit quality is quite good but worsens for larger p_T^{ave} values. Especially for the higher p_T^{ave} intervals in the central $|\eta|$ regions the statistical uncertainty is rather low and so the fit is very sensitive to even small deviations from a linear behaviour resulting in a minor fit quality. To consider this for the final result a possible non-linearity is covered with a dedicated systematic uncertainty.

5.6.2 Correction for Particle-Level Imbalance

In addition to an imbalance in dijet events caused by the presence of additional jets differences between the momentum of the parton and the corresponding particle-level jet can also arise e.g. from the jet clustering algorithm. This fact is referred to as particle-level imbalance (PLI) and estimated from simulation.

The contribution arising from the particle-level imbalance is estimated from the dijet asymmetry at generator level which is defined along with the asymmetry introduced for the resolution measurement itself as

$$A_{gen} = \frac{p_{T,1}^{gen} - p_{T,2}^{gen}}{p_{T,1}^{gen} + p_{T,2}^{gen}} \quad (5.12)$$

where $p_{T,1}^{gen}$ and $p_{T,2}^{gen}$ refer to the momenta of the two leading generated jets. This distribution is affected the same way as the asymmetry defined in Eq. 5.2 by additional jet activity in the event. Therefore the measured values of $\sigma_{A_{gen}}$ are extrapolated to zero additional jet activity $\alpha_{max} \rightarrow 0$ with the same extrapolation procedure as described in Section 5.6.1. Some exemplary extrapolations are shown in Fig. 5.8.

The contribution due to PLI is then given by

$$\sigma_{PLI} = \sqrt{2} \cdot \sigma_{A_{gen}}(\alpha_{max} \rightarrow 0). \quad (5.13)$$

5.6.3 Results of the Corrections to the Asymmetry

The results of the various extrapolation fits to determine $\sigma_A(\alpha_{max} \rightarrow 0)$ in data and simulation as well as for the particle-level imbalance are shown together in Fig. 5.9 as function of p_T^{ave} for the different η regions.

In order to obtain the jet energy resolution the fit result of the extrapolated σ_{PLI} is subtracted in quadrature from the measured σ_A after the correction for additional jet

¹This assumption is only almost true since the asymmetry distributions are truncated to reject non-gaussian components.

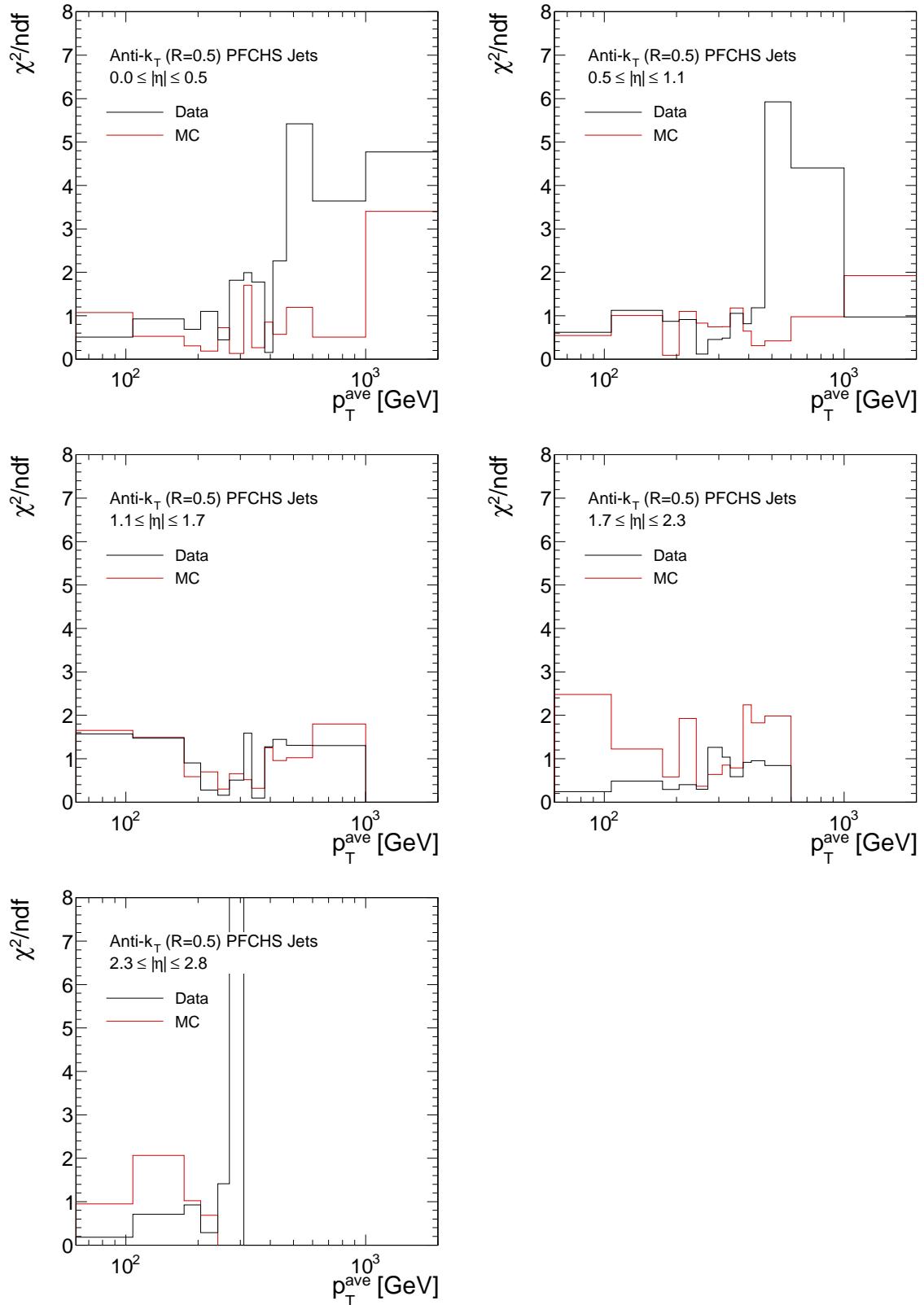


Figure 5.7: Goodness-of-fit test for the extrapolation fits in data and simulation in each $|\eta|$ region as function of p_T^{ave} .

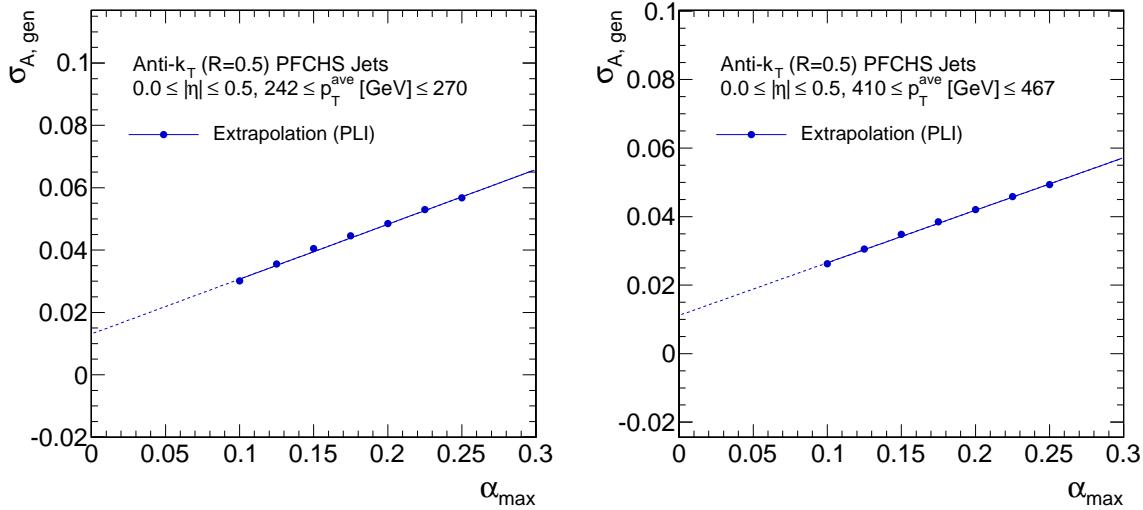


Figure 5.8: Two exemplary extrapolations of measured values for $\sigma_{A,gen}$ in simulation to determine the contribution from PLI.

activity

$$\sigma_{JER} = \sqrt{2} \cdot \sigma_A(\alpha_{max} \rightarrow 0) \ominus \sigma_{PLI}. \quad (5.14)$$

As can be seen in Fig. 5.9 the correction due to particle-level imbalance is small compared to the measured asymmetry widths.

5.7 Determination of the Data-to-Simulation Ratio of the Jet Transverse Momentum Resolution

The measured resolution in data and simulation can be compared after applying all corrections described in Section 5.6 by calculating the data-to-simulation ratio

$$c(\text{Data/MC}) = \frac{\sigma_{JER, \text{data}}}{\sigma_{JER, \text{simulation}}}. \quad (5.15)$$

The resulting distributions as function of p_T^{ave} for various $|\eta|$ regions are shown in Fig. 5.10. As no significant p_T dependence is observed, the ratio is parametrized by a constant fit. The fit result shows in each $|\eta|$ region a value larger than one meaning that the resolution in data is in general worse than in simulation. The constant fit is also visualized in Fig. 5.10 (red line) together with the statistical uncertainty resulting from the fit (grey shaded area). Due to statistical limitations no data-to-simulation ratio could be determined for the two highest $|\eta|$ intervals. Results are summarized later in Tab 5.3.

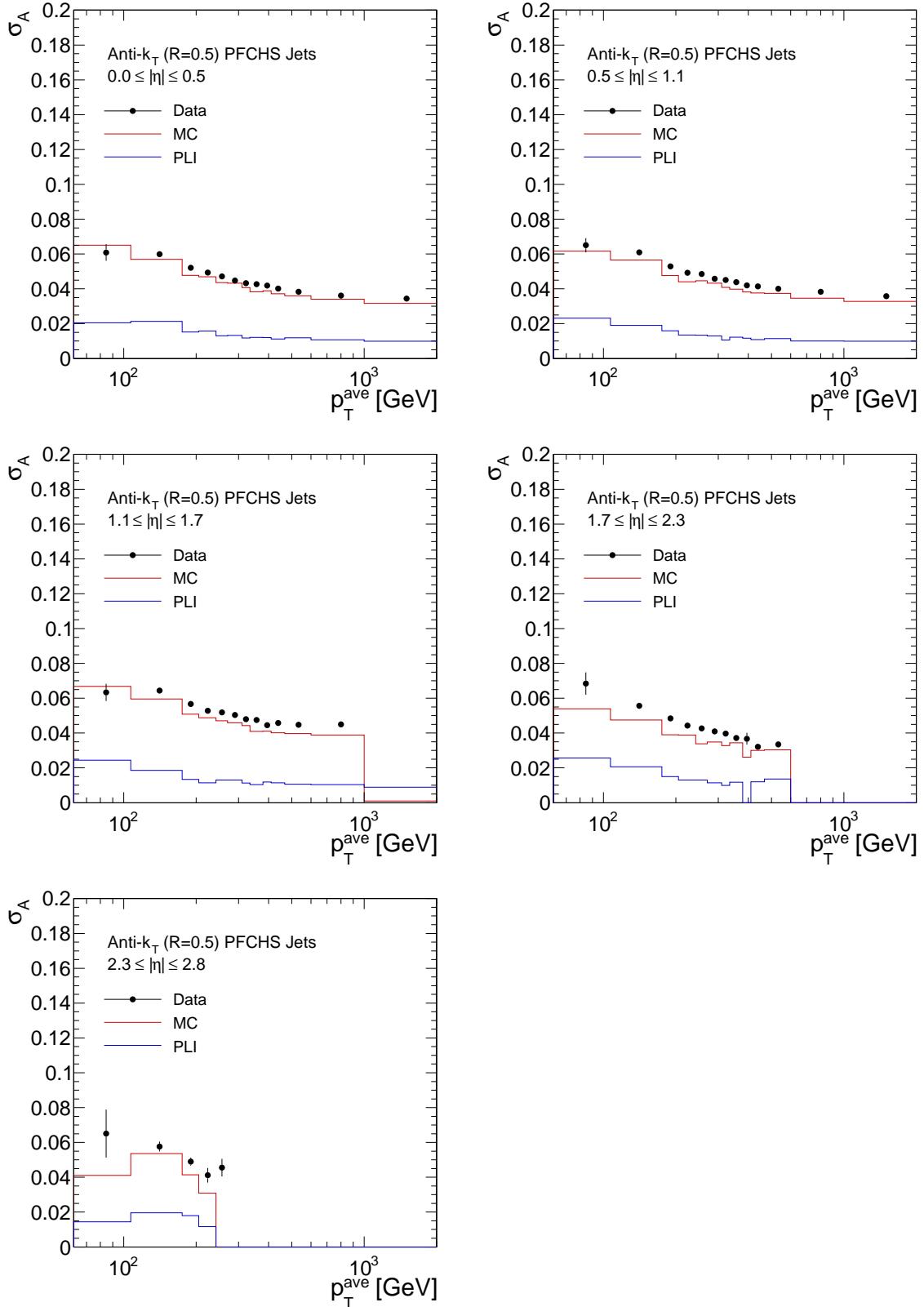


Figure 5.9: Results of extrapolation fits in different $|\eta|$ regions as function of p_T^{ave} for data, simulation and particle-level imbalance.

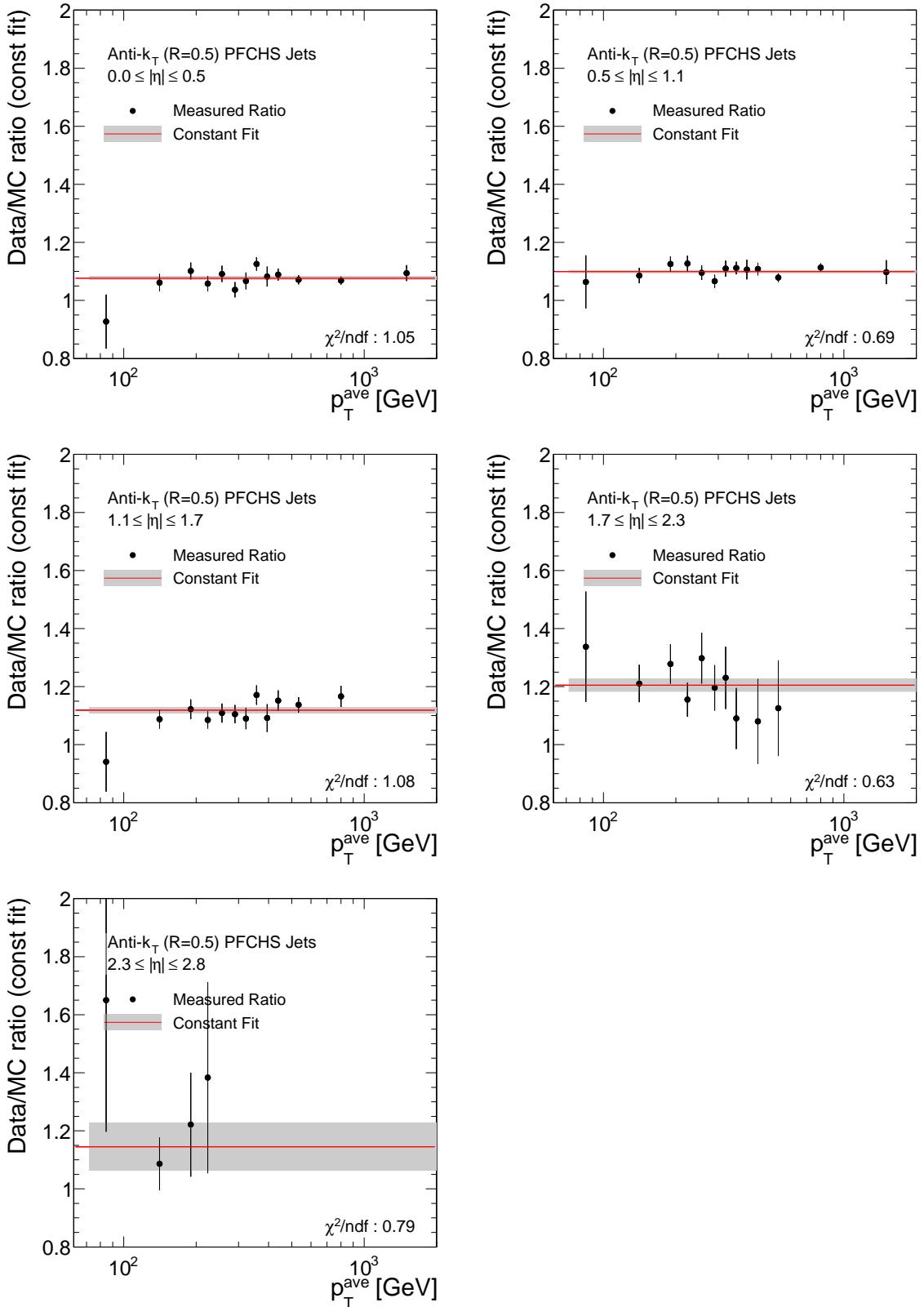


Figure 5.10: Ratio of the resolution in data to the resolution in the QCD sample in different $|\eta|$ regions as function of p_T^{ave} shown together with a constant fit.

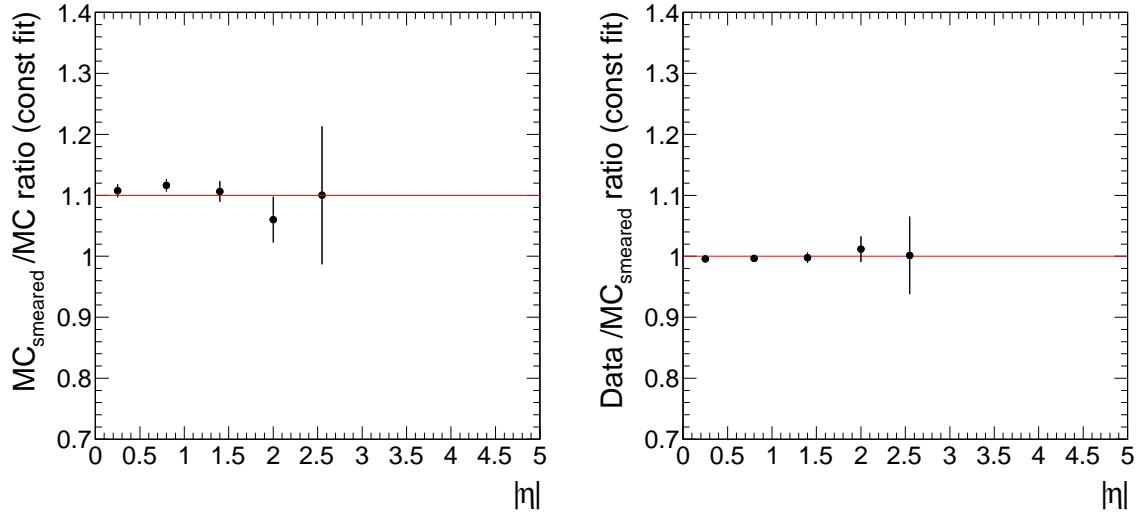


Figure 5.11: Ratio $c(\text{MC}_{\text{smeared}}/\text{MC})$ of the resolution in the smeared QCD sample to the resolution of the unsmeared QCD sample as function of $|\eta|$ with statistical uncertainty (*left*) and ratio $c(\text{Data}/\text{MC}_{\text{smeared}})$ of the resolution in data to the resolution of the QCD sample smeared with the measured data-to-simulation ratio as function of $|\eta|$ with statistical uncertainty (*right*). The red line indicates the target value of each validation test.

5.8 Validation of the Method

5.8.1 Validation in Simulated Events

In order to test the quality of the method to predict the correct data-to-simulation ratio a so-called closure test on simulated events is performed.

Therefore the simulated QCD sample is divided into two sub-samples with equal number of events. In one of the sub-samples the leading corrected jets² in p_T in each event are smeared with a smearing factor c to increase the p_T resolution, if they have a minimum p_T threshold of 10 GeV. This means that for each reconstructed jet which has an associated generated jet within a ΔR cone size of 0.25 the transverse momentum is scaled as follows

$$p'_T = p_T^{\text{gen}} + c \cdot (p_T - p_T^{\text{gen}}). \quad (5.16)$$

After the smearing procedure the jets are re-ordered again descendent in p_T .

For the closure test the smearing factor has been chosen to be $c = 1.1$ for each $|\eta|$ interval. The ratio of the measured resolution in the smeared QCD sub-sample to the resolution in the unsmeared sub-sample is derived with the asymmetry method including all corrections as described in Section 5.6. This should result in a measured ratio of the input value $c(\text{MC}_{\text{smeared}}/\text{MC}) = 1.1$. The resulting ratios from a constant fit as function of p_T^{ave} in each $|\eta|$ region with statistical uncertainties are shown in Fig. 5.11 (left).

It can be seen that the input scaling factor of 1.1 is nicely recovered with this procedure

²It is necessary to correct all jets which might get relevant for the analysis i.e. become one of the leading three jets. In this analysis the five leading jets in p_T are corrected.

within the statistical uncertainties. This allows the conclusion that a potential resolution difference in the ratio when comparing data and simulation can be determined with this approach as well.

5.8.2 Validation of the Measured Data-to-Simulation Ratio

In addition to the closure test on simulated events the measured data-to-simulation ratio can also be directly validated with a similar approach as described in Section 5.8.1.

For that the individual jet momenta in the simulated QCD sample are scaled with the measured resolution ratio values $c(\text{Data}/\text{MC})$ depending on the respective $|\eta|$ region. Applying the measurement procedure of the data-to-simulation ratio again after the smearing procedure one expects to obtain ratios of $c(\text{Data}/\text{MC}_{\text{smeared}}) = 1$ since the resolution differences between data and simulation have been adjusted before.

The result of this test is illustrated in Fig. 5.11 (right) with statistical errors. A nice agreement with the expected value of 1 is visible even without the consideration of systematic uncertainties in this consistency test.

5.9 Systematic Uncertainties

In addition to the uncertainty of the resolution measurement due to statictical effects, there are also systematic components which influence the outcome of the data-to-simulation ratio. The different contributions of systematic uncertainties to the measurement are discussed in this section.

All systematic uncertainties are determined by evaluating the shift of the data-to-simulation ratio when varying a certain aspect in the measurement procedure. The uncertainty δc is calculated by the determination of the ratio for a certain shift Δ and comparing it to the nominal ratio

$$\delta c(\text{Data}/\text{MC}) = c(\text{Data}/\text{MC})_{\Delta} - c(\text{Data}/\text{MC}) \quad (5.17)$$

It is symmetrized by taking the average shift, if the variation leads to an upward and a downward shift of the ratio. If the variation results only in an upward or downward shift, the maximum of both is taken and quoted as symmetric uncertainty.

PU reweighting: The trigger dependent pileup distributions in data which are taken in order to reweight the simulated sample to match the pileup distribution in data are calculated with a nominal minimum bias cross section of 69.4 mb. In order to evaluate the uncertainty due to this choice the minimum bias cross section is varied to 73.5 mb and the pileup scenario of the simulated QCD sample is reweighted to the data distributions obtained with this changed minimum bias cross section. Apart from that the measurement of the data-to-simulation ratio stays the same.

Since the minimum bias cross section is only shifted upwards, the resulting shift in the data-to-simulation ratio is considered as symmetric uncertainty.

Particle-level imbalance: The measured resolution in data and simulation is corrected for an imbalance at particle level due to out-of-cone showering effects using simulation only. In order to account for the uncertainty on σ_{PLI} the PLI-correction factor for each measured resolution value is shifted up and down by 25%. This means that the

changed ratio is calculated as

$$c(\text{Data}/\text{MC})_{\text{PLI}} = \frac{\sigma_{\text{Data}(\alpha \rightarrow 0)} \ominus (f \cdot \sigma_{\text{PLI}})}{\sigma_{\text{MC}(\alpha \rightarrow 0)} \ominus (f \cdot \sigma_{\text{PLI}})} \quad (5.18)$$

with $f = 0.75, 1.25$ respectively.

Jet energy scale: The jet energy scale has been corrected to particle level by the application of certain calibration factors which themselves have an uncertainty assigned. To evaluate the impact of this uncertainty on the data-to-simulation ratio of the resolution all jet momenta in the simulated sample are shifted up and down by the JEC uncertainty. The jet momenta in data stay unchanged.

α -spectrum: To account for additional jet activity in the event the measured width of the asymmetry distributions is extrapolated to zero additional jet activity by a linear function. This linear behaviour is more an empirically found relation than a theoretically fundamental connection. One influencing factor of the linear behaviour is the functional form of the α -spectrum. It determines how many events are added to the asymmetry distribution in the next higher α interval and consequently how much the asymmetry distribution is broadened. This can result e.g. in a smaller or bigger slope of the linear function or disturb the linear behaviour.

In order to evaluate the influence of the α -spectrum on the data-to-simulation ratio the observed inclusive α -spectrum in the simulated sample is reweighted to roughly match that in data. In Fig. 5.12 a comparison of the α -spectrum in data and simulation is shown. The bottom part displays the ratio Data/MC. The red curve overlaid in the ratio of the left distribution is used to reweight the events in the simulation. For each event a weight $w(\alpha)$ is calculated according to

$$w(\alpha) = 0.545 \cdot (\text{erf}(13.5 \cdot \alpha - 0.02) + 1) \quad (5.19)$$

where erf indicates the error-function³. This weight is considered as multiplicative factor onto the usual event weight in simulation.

α -range: The functional form of the α -spectrum is only one part of the question concerning the validity of the linear assumption for the extrapolation procedure. There is also the question for which α -range this assumption is valid. The choice of the linear function implies that the linear behaviour holds also for small α -values where no additional points are measured and up to $\alpha = 20 - 25\%$.

In order to study the linear behaviour of the extrapolation towards smaller α -values the minimum p_T cut of 10 GeV for the third jet is dropped and an additional α -interval 0.0 – 0.05 is introduced. In Fig. 5.13 some exemplary extrapolations with the additional α -interval are shown. In general the fit quality decreases slightly but is still reasonable (shown in App.). The resulting difference in the data-to-simulation ratio is considered as systematic uncertainty.

app:fit
quality

Non-gaussian tails: As discussed in Sec. 5.3 the width of the asymmetry distribution is calculated as a truncated root-mean-square in order to reject contributions from non-

³ $\text{erf} = \frac{1}{\sqrt{\pi}} \int_{-x}^x e^{-t^2} dt$

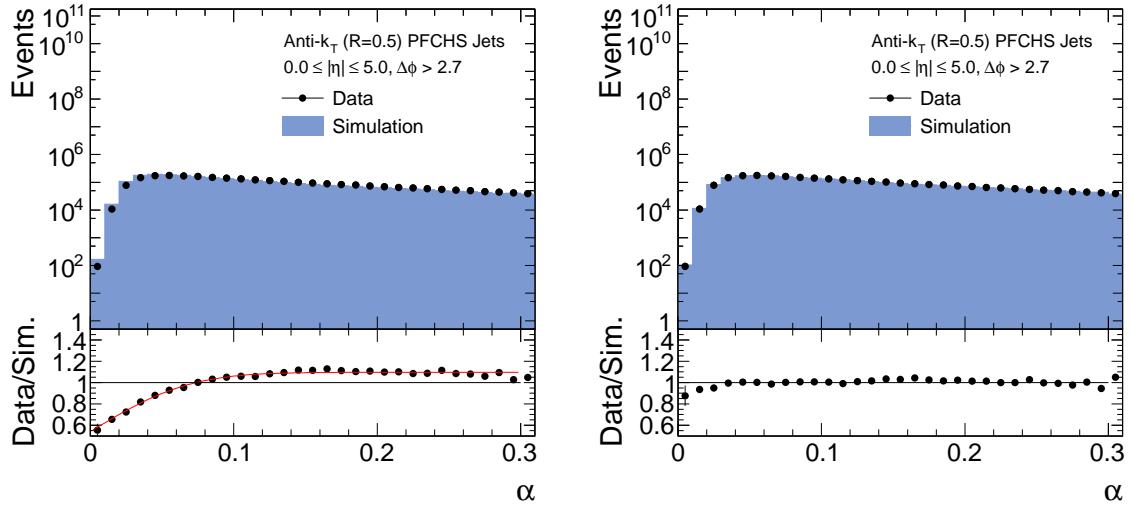


Figure 5.12: Inclusive α -spectrum before (left) and after (right) reweighting the α -spectrum of the simulated sample. The red curve in the bottom of the left plot illustrates the function used for the α -spectrum reweighting.

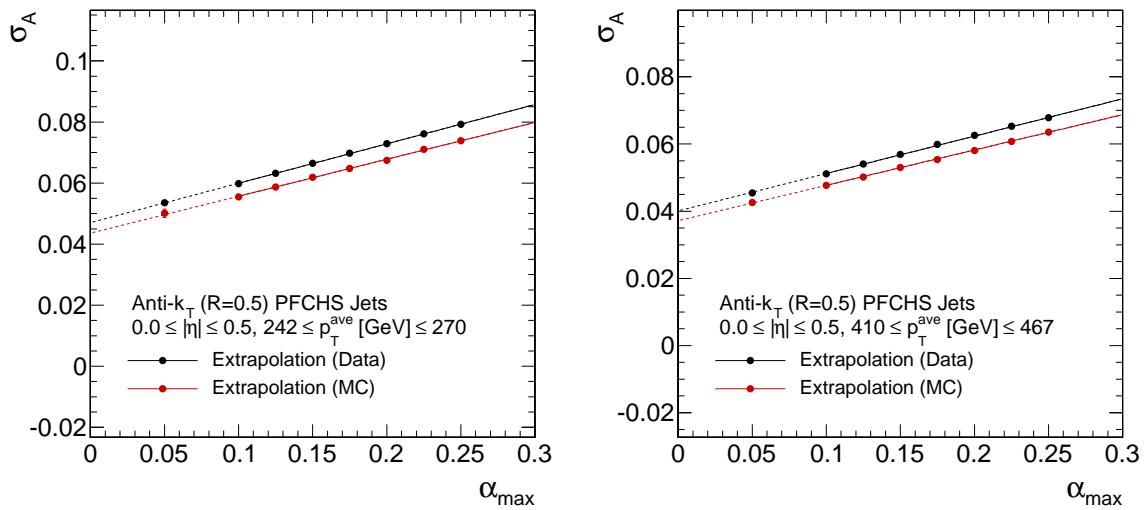


Figure 5.13: Two exemplary extrapolations of measured values for σ_A in data and simulation to obtain the result at zero additional jet activity when adding one additional α interval 0.0 – 0.05.

gaussian tails. The truncation was chosen to be 1.5% for both data and simulation. In general it is possible that the tail contributions in data and simulation differ and consequently do not cancel out in the ratio. In order to evaluate this effect the data-to-simulation ratio is calculated when truncating 5% of the original distribution instead of 1.5%.

Flavour uncertainty: In general the jet energy resolution for light (u,d,s) and heavy quarks (c,b) can be quite different due to neutrinos occurring in the decay of heavy quarks. This resolution difference related to the quark flavour cancels in principle out in the data-to-simulation ratio, if the flavour composition is the same in data and simulation. Since this is not necessarily the case the impact of heavy quarks produced in gluon splitting processes on the data-to-simulation ratio is evaluated by varying the event weights for events with gluon splitting into heavy quarks.

An event is considered as gluon splitting event, if one of the two leading jets has a gluon jet flavour with respect to the physics definition for jet flavour. It is distinguished between gluon splitting into light and heavy jets by identifying if the jet considered as gluon in the physics definition has a b/c-flavour or a light flavour using the algorithmic definition. All events where at least one of the two leading jets is identified as a gluon splitting into heavy quarks get an event weight of 1.5. The data-to-simulation ratio is then derived for the resolution from data to the resolution in the reweighted simulated sample.

Shape of the data-to-simulation ratio: The data-to-simulation ratio in this measurement is determined by fitting the ratio of $\sigma_{\text{JER, data}}$ to $\sigma_{\text{JER, simulation}}$ with a constant. This assumes that the ratio is flat as function of p_T^{ave} . In order to test this assumption one can fit the $\sigma_{\text{JER, data}}$ and $\sigma_{\text{JER, simulation}}$ distributions separately and scale the parameters for data with a certain factor.

The following function (NSC – function) is used to fit the resolution in simulation

$$f(p_T) = \sqrt{\frac{N^2}{p_T^2} + \frac{S^2}{p_T} + C^2} \quad (5.20)$$

with $p_T = p_T^{ave}$ and parameters N, S and C. If the data-to-simulation ratio is flat as function of p_T^{ave} one expects to find one common scale factor k_{NSC} for the N, S and C parameters to describe the data. A strong influence from e.g. pile-up which might be modelled differently than it actually occurs in data could result in individual scale factors for these parameters.

In order to test the assumption of a flat ratio the data points are fitted with the following function

$$f(p_T) = \sqrt{\frac{(k_{NS} \cdot N)^2}{p_T^2} + \frac{(k_{NS} \cdot S)^2}{p_T} + (k_C \cdot C)^2} \quad (5.21)$$

which uses one scale factor for parameters N and S as well as one scale factor for C while the parameters N, S and C are fixed to the results of the fit to the simulation. The fits are summarized in Fig. 5.14.

It turns out that the scale factors k_{NS} and k_C are not necessarily equal. In order

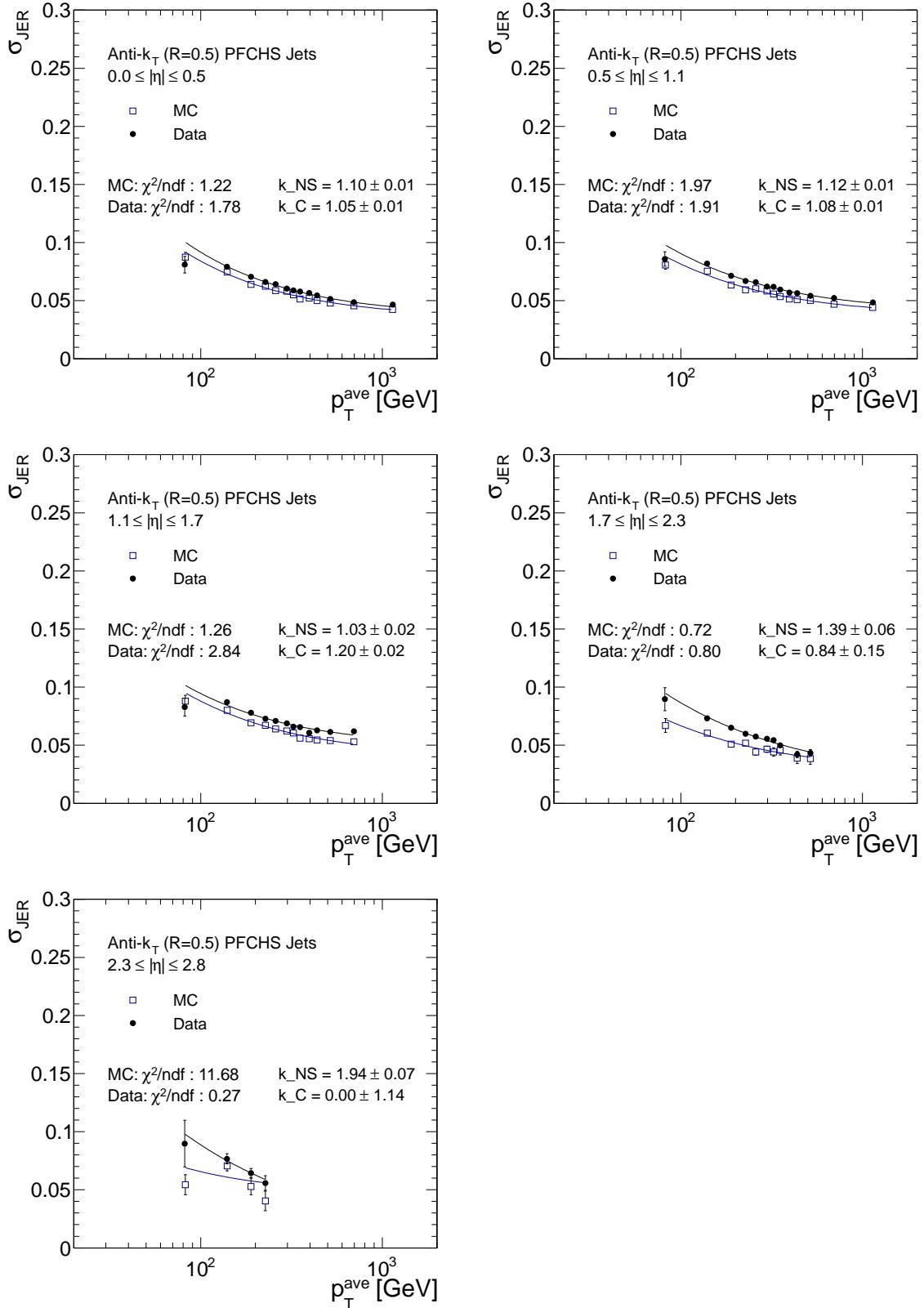


Figure 5.14: Results of fitting the resolution in data and simulation with the respective NSC-functions in each $|\eta|$ interval.

to see, if such a behaviour is compatible with the measured points for the data-to-simulation ratio, the ratio of the two fitted functions for data and simulation is shown together with the measured values and the constant fit in Fig. 5.15.

One can see that the ratio derived from the NSC – fits is also compatible with the measured values for the ratio which means that in principle a trend versus p_T^{ave} could be hidden in the measured points. The difference of the measured scale factors k_{NS} and k_C to the central value of the measurement from the constant fit is considered as systematic uncertainty. This also covers possible differences of the scaling factors for jets with very low or very high transverse momenta which can not be directly measured in this analysis. As the NSC-Fits in the higher $|\eta|$ intervals suffer from the low number of events there, a 2% uncertainty on the shape of the data-to-simulation ratio is considered for all $|\eta|$ intervals.

A summary of the data-to-simulation ratios measured with this method together with all systematic uncertainties considered for the ratio is given in Tab 5.3.

Table 5.3: Summary of the measured data-to-simulation ratios $c(\text{Data}/\text{MC})$ with statistical uncertainty and systematic uncertainty for each uncertainty source in different $|\eta|$ regions.

	$ \eta $						
	0.0 – 0.5	0.5 – 1.1	1.1 – 1.7	1.7 – 2.3	2.3 – 2.8	2.8 – 3.2	3.2 – 5.0
$c(\text{Data}/\text{MC})$	1.077	1.100	1.119	1.205	1.145	-	-
Stat. uncertainty	± 0.007	± 0.006	± 0.010	± 0.027	± 0.078	-	-
PU reweighting	± 0.004	± 0.003	± 0.004	± 0.004	± 0.013	-	-
Particle-level imbalance	± 0.004	± 0.005	± 0.005	± 0.015	± 0.010	-	-
Jet energy scale	± 0.005	± 0.007	± 0.008	± 0.015	± 0.020	-	-
α -spectrum	± 0.004	± 0.007	± 0.004	± 0.006	± 0.003	-	-
α -range	± 0.005	± 0.009	± 0.004	± 0.011	± 0.012	-	-
Non-gaussian tails	± 0.004	± 0.003	± 0.003	± 0.013	± 0.007	-	-
Jet Flavour	± 0.007	± 0.004	± 0.005	± 0.006	± 0.003	-	-
Ratio shape	± 0.022	± 0.022	± 0.022	± 0.024	± 0.023	-	-
Total syst. uncertainty	± 0.025	± 0.027	± 0.026	± 0.038	± 0.038	-	-

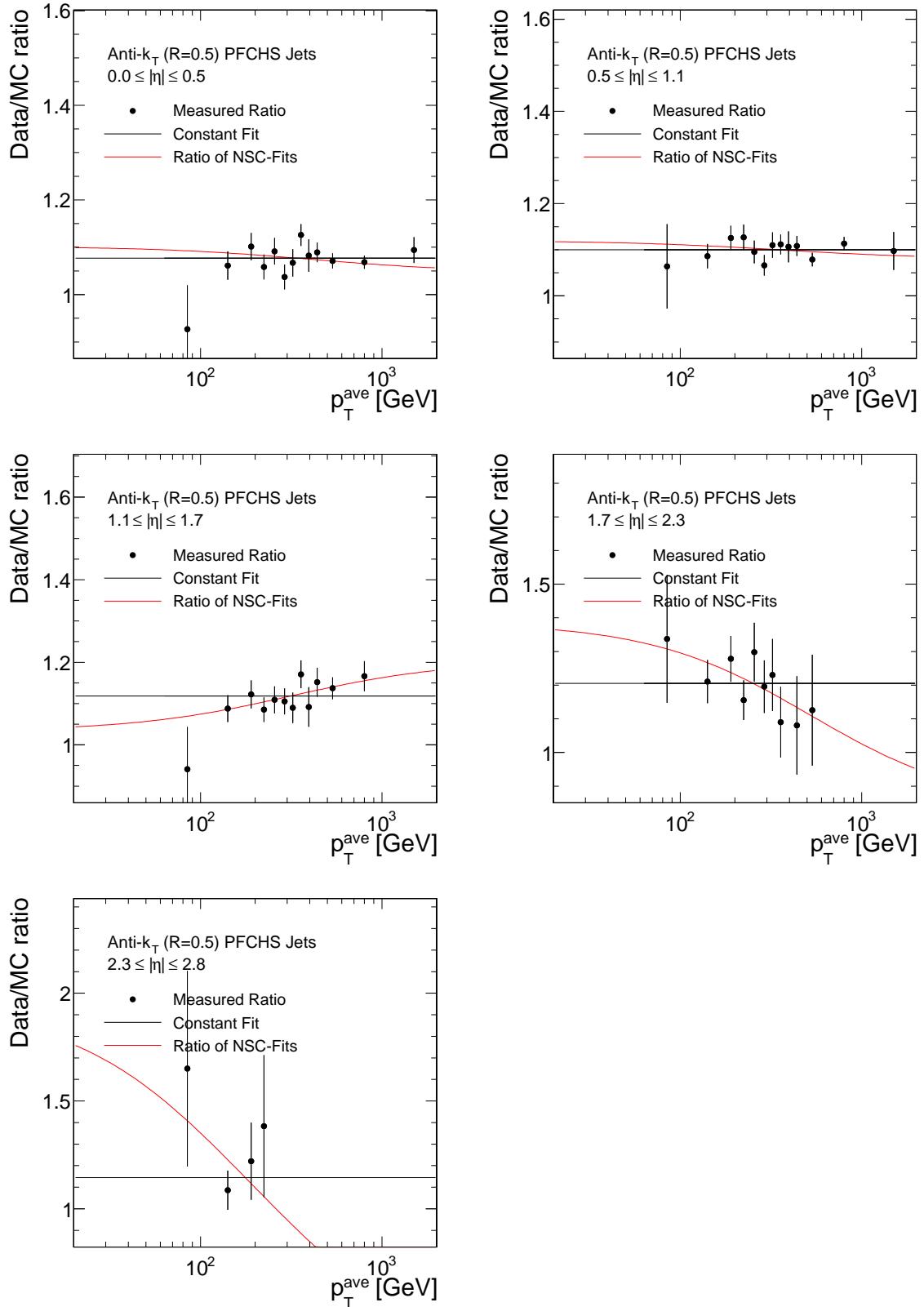


Figure 5.15: Ratio of the results fitting the resolution in data and simulation with the NSC-functions in each $|\eta|$ interval together with the measured ratio and a constant fit.

5.10 Extension of the Method to the Forward Detector Region

The measurement presented until here is based on the requirement that the two leading jets in an event belong both to the same $|\eta|$ region. This means that the resolution of these two jets is the same and results in the relation that $\frac{\sigma(p_T)}{\langle p_T \rangle} = \sqrt{2} \cdot \sigma_A$. This requirement of $|\eta_{\text{jet},1}| = |\eta_{\text{jet},2}|$ reduces significantly the available number of events especially in the forward region of the detector as the jets in these events have low transverse momenta and consequently have to be triggered by the lowest p_T^{ave} triggers which are highly prescaled. Thus for the intervals $|\eta| > 2.8$ no data-to-simulation ratio could be determined.

In order to extend the analysis such that a measurement in the forward region of the detector is also possible one can drop the requirement that $|\eta_{\text{jet},1}| = |\eta_{\text{jet},2}|$ and select events where the two leading jets belong to different pseudorapidity regions. The resolution $\sigma(p_T^{\text{jet,probe}})$ in a probe interval $|\eta_{\text{probe}}|$ can then be measured, if the resolution $\sigma(p_T^{\text{jet,ref}})$ in a reference interval $|\eta_{\text{ref}}|$ is already known.

Starting again with the relation in Eq. 5.3, it can be shown that for $\langle p_{T,\text{probe}} \rangle = \langle p_{T,\text{ref}} \rangle = \langle p_T \rangle$ and $\sigma(p_{T,\text{probe}}) \neq \sigma(p_{T,\text{ref}})$ one obtains the relation

$$\frac{\sigma(p_{T,\text{probe}})}{\langle p_T \rangle} = \sqrt{4 \cdot \sigma_A(|\eta_{\text{probe}}| \neq |\eta_{\text{ref}}|) - \left(\frac{\sigma(p_{T,\text{ref}})}{\langle p_T \rangle} \right)^2} \quad (5.22)$$

where $\sigma_A(|\eta_{\text{probe}}| \neq |\eta_{\text{ref}}|)$ is the width of the asymmetry calculated from the two leading jets in an event of which one jet is in a reference $|\eta|$ interval for which the resolution has been already measured with the same- $|\eta|$ requirement and the probe jet belonging to another $|\eta|$ region. The asymmetry width is taken after the correction for additional jet activity and the particle-level imbalance determined from $\sigma_{A,\text{gen}}(|\eta_{\text{probe}}| \neq |\eta_{\text{ref}}|)$. The reference interval is preferred to be in the central detector region where the statistical precision from the same- $|\eta|$ measurement has already been sufficient. The statistical uncertainty of $\frac{\sigma(p_{T,\text{ref}})}{\langle p_T \rangle}$ is propagated to the statistical uncertainty of $\frac{\sigma(p_{T,\text{probe}})}{\langle p_T \rangle}$.

As long as both jets belong to the same $|\eta|$ region residual effects from jet energy scale differences in data and simulation affect both jets the same way and should not have an impact on the resolution measurement. Since residual effects from the jet energy scale where the mean of the asymmetry is shifted and not exactly zero become more important, if both jets belong to different probe and reference intervals, the asymmetry is defined as

$$A = \frac{p_{T,\text{probe}} - p_{T,\text{ref}}}{p_{T,\text{probe}} + p_{T,\text{ref}}} \quad (5.23)$$

for the reference-and-probe interval measurement for the transverse momenta $p_{T,\text{ref}}$ and $p_{T,\text{probe}}$ of the reference and probe jet, respectively. Accordingly, in the determination of the asymmetry width σ_A the mean of the distribution A_{mean} is also estimated and the width calculated as

$$\sigma_A = \text{RMS}_{98.5\%} = \sqrt{\frac{1}{n} \cdot \sum_i (A_i - A_{\text{mean}})^2} \quad (5.24)$$

with the number of possible values n for the individual asymmetry values A_i where the sum over i includes all values such that 98.5% of the total asymmetry distribution are

covered symmetric around the mean.

The measurement with the reference-and-probe interval selection is performed taking the three innermost $|\eta|$ regions $0.0 - 0.5$, $0.5 - 1.1$ and $1.1 - 1.7$ as reference intervals each. This allows a measurement of the data-to-simulation ratio also in the intervals $2.8 - 3.2$ and $3.2 - 5.0$. In addition, it provides for the innermost $|\eta|$ intervals an enhancement of the available number of events and thus a further reduction of the statistical uncertainties. The systematic uncertainties of the two new measurements are evaluated the same way as for the same- $|\eta|$ measurement by varying a certain aspect in the determination of the data-to-simulation ratio and taking the deviation from the nominal ratio as uncertainty. The variation is done for the $\sigma_{A(|\eta_{\text{probe}}| \neq |\eta_{\text{ref}}|)}$ value as well as for $\frac{\sigma(p_T, \text{ref})}{\langle p_T \rangle}$ for each uncertainty source.

A summary of the resulting data-to-simulation ratios determined in the measurements with the two different reference intervals and the corresponding uncertainties is shown in Table 5.4 for $|\eta_{\text{ref}}| = 0.0 - 0.5$, Table 5.5 for $|\eta_{\text{ref}}| = 0.5 - 1.1$ and Table 5.6 for $|\eta_{\text{ref}}| = 1.1 - 1.7$, respectively.

Table 5.4: Summary of the measurement with reference bin $|\eta| = 0.0 - 0.5$ showing the nominal data-to-simulation ratio $c(\text{Data}/\text{MC})$ with statistical uncertainty and systematic uncertainty for each uncertainty source in different $|\eta|$ regions.

		$ \eta_{\text{ref}} = 0.0 - 0.5$						
		$ \eta_{\text{probe}} $						
		0.0 – 0.5	0.5 – 1.1	1.1 – 1.7	1.7 – 2.3	2.3 – 2.8	2.8 – 3.2	3.2 – 5.0
$c(\text{Data}/\text{MC})$	-	1.106	1.133	1.227	1.253	1.410	1.171	
Stat. uncertainty	-	± 0.008	± 0.009	± 0.025	± 0.047	± 0.068	± 0.116	
PU reweighting	-	± 0.002	± 0.001	± 0.001	± 0.021	± 0.025	± 0.007	
Particle-level imbalance	-	± 0.005	± 0.005	± 0.008	± 0.009	± 0.006	± 0.004	
Jet energy scale	-	± 0.006	± 0.010	± 0.021	± 0.034	± 0.022	± 0.066	
α -spectrum	-	± 0.010	± 0.006	± 0.011	± 0.006	± 0.009	± 0.012	
α -range	-	± 0.010	± 0.014	± 0.046	± 0.089	± 0.015	± 0.018	
Non-gaussian tails	-	± 0.001	± 0.007	± 0.005	± 0.044	± 0.046	± 0.029	
Jet Flavour	-	± 0.004	± 0.004	± 0.001	± 0.021	± 0.013	± 0.008	
Ratio shape	-	± 0.022	± 0.023	± 0.025	± 0.025	± 0.028	± 0.023	
Total syst. uncertainty	-	± 0.028	± 0.030	± 0.058	± 0.112	± 0.067	± 0.079	

5.11 Measurement for simulated events obtained with the Herwig++ generator

In order to test if the data-to-simulation ratio is generator independent the ratio is also derived for simulated events generated with HERWIG++⁴.

Fig. 5.16 shows the inclusive p_T^{ave} spectrum compared in data and simulation for simulated events taken from HERWIG++.

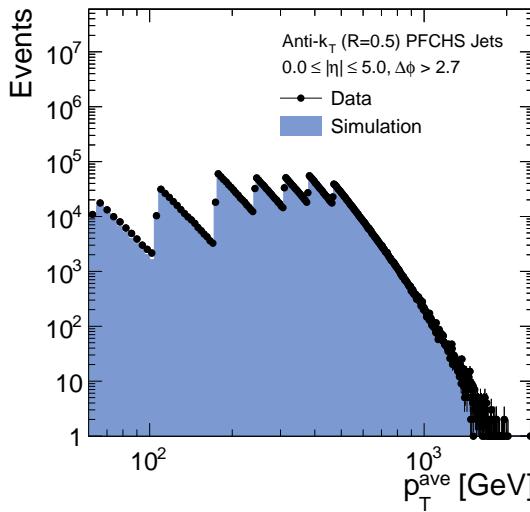


Figure 5.16: Inclusive p_T^{ave} spectrum of events after the event selection in data and in simulated events generated with HERWIG++.

The agreement of the p_T^{ave} spectrum between data and simulation looks reasonable and is similar to the spectrum in simulated events from PYTHIA.

The resulting data-to-simulation ratios for HERWIG++ from the different measurements together with statistical uncertainties are summarized in Tab. 5.7 In general, the obtained data-to-simulation ratios for HERWIG++ are quite similar to that obtained from PYTHIA while some further discussion can be found in Sec. 5.12.1.

5.12 Results

5.12.1 Determination of a combined result

In order to determine one data-to-simulation ratio for each $|\eta|$ region the results obtained from the individual measurements are combined into one result. This combination is for all $|\eta|$ regions – except for the very last $|\eta|$ region ranging from 3.2 – 5.0 which is discussed separately – derived from the measurements done with simulated events from PYTHIA while the HERWIG++ results are only used as cross check. The combined result for all $|\eta|$ regions up to 3.2 using events from HERWIG++ are shown in Appendix .

⁴The HERWIG++ sample is /QCD_Pt_15to3000_TuneEE3C_Flat_8TeV_herwigpp created in the series /Summer12_DR53X – PU_S10_START53_V7A – v1/AODSIM

app:combination
herwig

The combined data-to-simulation ratios are calculated as weighted mean from the individual measurements with the same- $|\eta|$ requirement and the three central-forward combinations. Each measurement is weighted by its' statistical uncertainty. For the evaluation of the systematic uncertainties considered for the combined result the systematic shift of the weighted mean caused by the systematic shifts of the single measurements is determined. This means that for each uncertainty source the individual up and down shifts to the nominal ratio of each measurement are taken and combined as the weighted mean using the statistical uncertainty just as for the nominal mean. Consequentially, the difference of each systematically shifted weighted mean to the nominal weighted mean is the respective systematic uncertainty for the combination. The total systematic uncertainty quoted for the weighted mean of the data-to-simulation ratio is the average of the individual systematic uncertainties in each $|\eta|$ interval.

The $|\eta|$ region from $3.2 - 5.0$ covering the most forward region of the detector – namely the hadronic forward – behaves a bit special compared to the others. In contrast to the other $|\eta|$ regions, the individual measurements from PYTHIA and HERWIG++ have a very large spread of 0.83 up to 1.32 in this particular interval. If one wants to have one combined ratio factor for this interval, it is in this case reasonable to use all six individual measurements from PYTHIA as well as from HERWIG++ to find a result which covers all of these partly very different ratios. If one combines the measurements following the same procedure as described for the other intervals and taking the systematic uncertainty from the combined PYTHIA results the combined total result would be 1.056 ± 0.048 (stat. unc.) ± 0.079 (syst. unc.) ($= 0.092$ total unc.). In order to test the consistency of this combined value with the single measurements the χ^2/ndf is calculated which results in $\chi^2/ndf = 4.3$. This shows that the obtained total uncertainty is too small to reasonably cover all individual results. Hence the systematic uncertainty is increased to a value of ± 0.185 which gives a total uncertainty of ± 0.191 and results in a $\chi^2/ndf = 1.0$ such that the single measurements are well represented by this combined result.

The obtained data-to-simulation ratios for different $|\eta|$ regions are summarized in Tab. 5.8 together with the statistical and systematic uncertainties. In addition, Fig. 5.17 shows the measured ratios with the total uncertainty which is the quadratic sum of the statistical and systematic uncertainties.

The determined ratios vary from 1.06 to 1.40 and are larger than one in all detector parts. Possible sources of this difference can be e.g. mismodelled noise effects or inhomogenities of the detector.

5.12.2 Comparison to Other Measurements

Earlier analyses using dijet events for data collected at $\sqrt{s} = 7$ TeV have obtained similar results showing data-to-simulation ratios greater than one. A complementary approach is followed using $\gamma +$ jet events which offer a very precise opportunity to measure the jet resolution due to the excellent resolution of the photon energy. A comparison of the numbers presented here to the latest $\sqrt{s} = 7$ TeV results from dijet events [99] and $\sqrt{s} = 8$ TeV results for $\gamma +$ jet measurements [102] are shown together in Fig. 5.18.

The results obtained from different methods and for the different center-of-mass energies are well compatible with each other. The main advantage of the dijet measurement presented in this note compared to the $\gamma +$ jet analysis is that one obtains also a number for the outermost $|\eta|$ region. In addition, the total uncertainties in the other $|\eta|$ regions

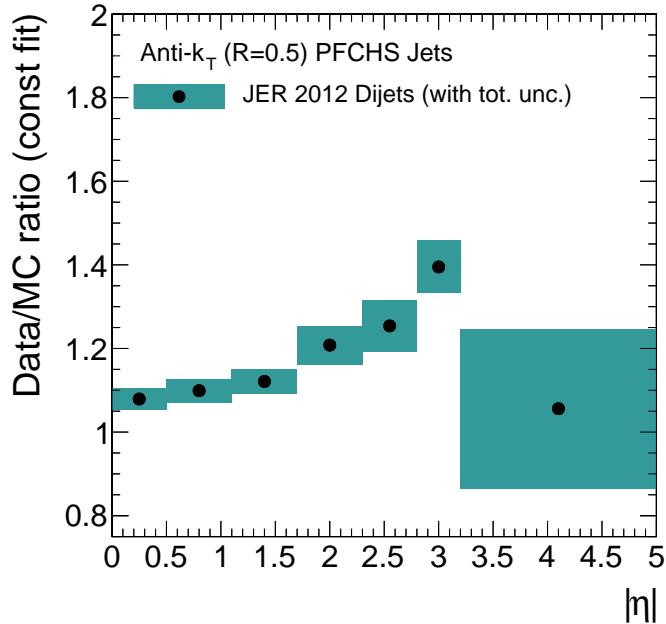


Figure 5.17: Measured data-to-simulation ratio in various $|\eta|$ regions displayed with total uncertainty.

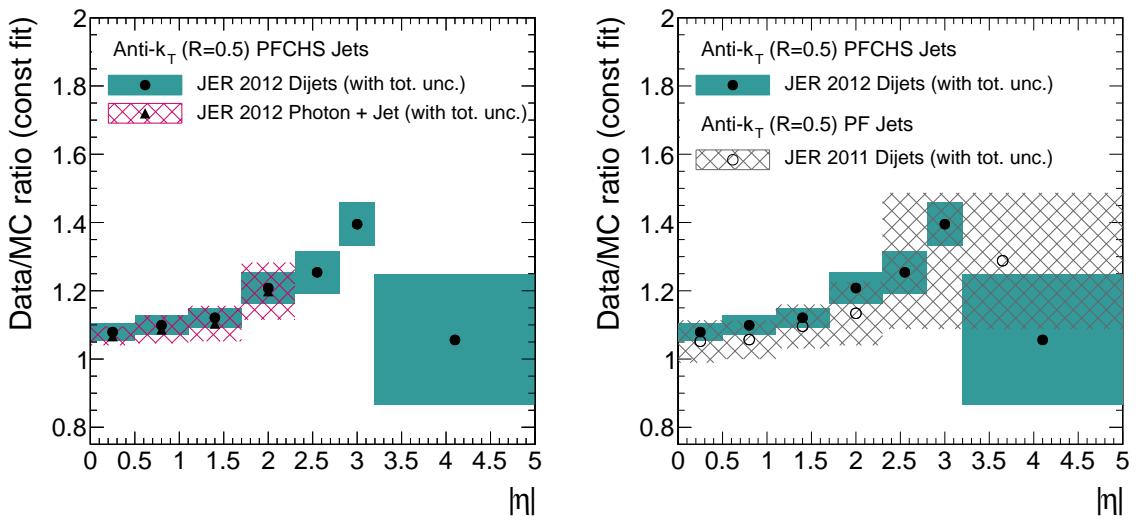


Figure 5.18: Measured data-to-simulation ratio in various $|\eta|$ regions displayed with total uncertainty. Comparison to results obtained for $\sqrt{s} = 8$ TeV data from $\gamma + \text{jet}$ events (left) and to $\sqrt{s} = 7$ TeV data from dijet events (right).

are at a same accuracy or even slightly lower.

In comparison to the $\sqrt{s} = 7$ TeV results from dijet events the total uncertainty could be significantly reduced by incorporating the correlation among the different inclusive α regions in the extrapolation procedure. In the previous analysis the statistical uncertainties were underestimated by not considering the above mentioned correlation. This effect was compensated by introducing a rather conservative systematic uncertainty on the extrapolation to zero additional jet activity. Since the treatment of the statistical uncertainty has been changed in this analysis, the estimation of systematic effects could be adjusted accordingly and lead to the overall reduced uncertainty values summarized in Table 5.8.

5.13 Adjustment of the MC Resolution to Data

5.14 Impact of the Resolution Measurement on other Analyses

Table 5.5: Summary of the measurement with reference bin $|\eta| = 0.5 - 1.1$ showing the nominal data-to-simulation ratio $c(\text{Data}/\text{MC})$ with statistical uncertainty and systematic uncertainty for each uncertainty source in different $|\eta|$ regions.

	$ \eta_{\text{ref}} = 0.5 - 1.1$ $ \eta_{\text{probe}} $						
	0.0 – 0.5	0.5 – 1.1	1.1 – 1.7	1.7 – 2.3	2.3 – 2.8	2.8 – 3.2	3.2 – 5.0
$c(\text{Data}/\text{MC})$	1.081	-	1.111	1.206	1.300	1.356	0.829
Stat. uncertainty	± 0.008	-	± 0.009	± 0.023	± 0.047	± 0.058	± 0.082
PU reweighting	± 0.001	-	± 0.003	± 0.005	± 0.002	± 0.024	± 0.039
Particle-level imbalance	± 0.005	-	± 0.005	± 0.010	± 0.017	± 0.009	± 0.021
Jet energy scale	± 0.008	-	± 0.009	± 0.021	± 0.009	± 0.061	± 0.134
α -spectrum	± 0.007	-	± 0.004	± 0.003	± 0.008	± 0.036	± 0.002
α -range	± 0.004	-	± 0.012	± 0.016	± 0.043	± 0.048	± 0.041
Non-gaussian tails	± 0.004	-	± 0.008	± 0.010	± 0.061	± 0.023	± 0.012
Jet Flavour	± 0.005	-	± 0.007	± 0.009	± 0.013	± 0.042	± 0.007
Ratio shape	± 0.022	-	± 0.022	± 0.024	± 0.026	± 0.027	± 0.017
Total syst. uncertainty	± 0.026	-	± 0.030	± 0.039	± 0.082	± 0.105	± 0.149

Table 5.6: Summary of the measurement with reference bin $|\eta| = 1.1 - 1.7$ showing the nominal data-to-simulation ratio $c(\text{Data}/\text{MC})$ with statistical uncertainty and systematic uncertainty for each uncertainty source in different $|\eta|$ regions.

	$ \eta_{\text{ref}} = 1.1 - 1.7$ $ \eta_{\text{probe}} $						
	0.0 – 0.5	0.5 – 1.1	1.1 – 1.7	1.7 – 2.3	2.3 – 2.8	2.8 – 3.2	3.2 – 5.0
$c(\text{Data}/\text{MC})$	1.084	1.082	-	1.189	1.250	1.432	1.137
Stat. uncertainty	± 0.012	± 0.012	-	± 0.031	± 0.051	± 0.066	± 0.105
PU reweighting	± 0.005	± 0.004	-	± 0.024	± 0.001	± 0.002	± 0.029
Particle-level imbalance	± 0.004	± 0.004	-	± 0.016	± 0.017	± 0.019	± 0.006
Jet energy scale	± 0.012	± 0.010	-	± 0.029	± 0.041	± 0.055	± 0.018
α -spectrum	± 0.006	± 0.008	-	± 0.019	± 0.006	± 0.019	± 0.009
α -range	± 0.016	± 0.022	-	± 0.036	± 0.032	± 0.016	± 0.048
Non-gaussian tails	± 0.010	± 0.010	-	± 0.005	± 0.018	± 0.012	± 0.047
Jet Flavour	± 0.005	± 0.007	-	± 0.012	± 0.015	± 0.031	± 0.054
Ratio shape	± 0.022	± 0.022	-	± 0.024	± 0.025	± 0.029	± 0.023
Total syst. uncertainty	± 0.033	± 0.036	-	± 0.063	± 0.065	± 0.077	± 0.096

Table 5.7: Measured data-to-simulation ratio in various $|\eta|$ regions with statistical uncertainty for simulated events obtained with the HERWIG++ generator.

$ \eta $	$ \eta_1 = \eta_2 $ $c(\text{Data}/\text{MC})$	$ \eta_{\text{ref}} = 0.0 - 0.5$ $c(\text{Data}/\text{MC})$	$ \eta_{\text{ref}} = 0.5 - 1.1$ $c(\text{Data}/\text{MC})$	$ \eta_{\text{ref}} = 1.1 - 1.7$ $c(\text{Data}/\text{MC})$
0.0 – 0.5	1.090 ± 0.008	-	1.091 ± 0.011	1.129 ± 0.017
0.5 – 1.1	1.107 ± 0.008	1.111 ± 0.010	-	1.121 ± 0.016
1.1 – 1.7	1.117 ± 0.013	1.177 ± 0.013	1.149 ± 0.012	-
1.7 – 2.3	1.297 ± 0.038	1.212 ± 0.035	1.229 ± 0.033	1.187 ± 0.041
2.3 – 2.8	1.085 ± 0.080	1.231 ± 0.072	1.178 ± 0.061	1.363 ± 0.087
2.8 – 3.2	-	1.368 ± 0.094	1.259 ± 0.061	1.488 ± 0.112
3.2 – 5.0	-	1.245 ± 0.158	1.124 ± 0.128	1.324 ± 0.215

Table 5.8: Measured data-to-simulation ratio in various $|\eta|$ regions with statistical and systematic uncertainty as well as the total uncertainty.

$ \eta $	$c(\text{Data}/\text{MC})$	stat.	syst.	tot.
0.0 – 0.5	1.079	± 0.005	± 0.026	± 0.026
0.5 – 1.1	1.099	± 0.005	± 0.028	± 0.028
1.1 – 1.7	1.121	± 0.005	± 0.029	± 0.029
1.7 – 2.3	1.208	± 0.013	± 0.045	± 0.046
2.3 – 2.8	1.254	± 0.026	± 0.056	± 0.062
2.8 – 3.2	1.395	± 0.036	± 0.051	± 0.063
3.2 – 5.0	1.056	± 0.048	± 0.185	± 0.191

6 Search for New Physics in the Multijet and Missing Transverse Momentum Final State at $\sqrt{s} = 8$ TeV

Introduction and motivation with 8 TeV cross section and rate

6.1 Event Selection

6.1.1 Data samples and trigger

6.1.2 Event Cleaning

6.1.3 Baseline Selection

6.1.4 Exclusive Search Regions

6.2 QCD Background Estimation with the Rebalance-And-Smear Method

6.2.1 Rebalance Procedure using Kinematic Fits

6.2.2 Response Smearing

6.2.3 Validation Tests

6.2.4 Systematic Uncertainties

6.2.5 QCD Background Prediction

6.3 Estimation of Non-QCD Backgrounds

6.3.1 Invisible Z Background

6.3.2 Hadronic τ Background

6.3.3 Lost-Lepton Background

6.4 Results and Interpretation

6.4.1 Comparison to Other Measurements

6.5 Status of natural supersymmetry after LHC Run-I

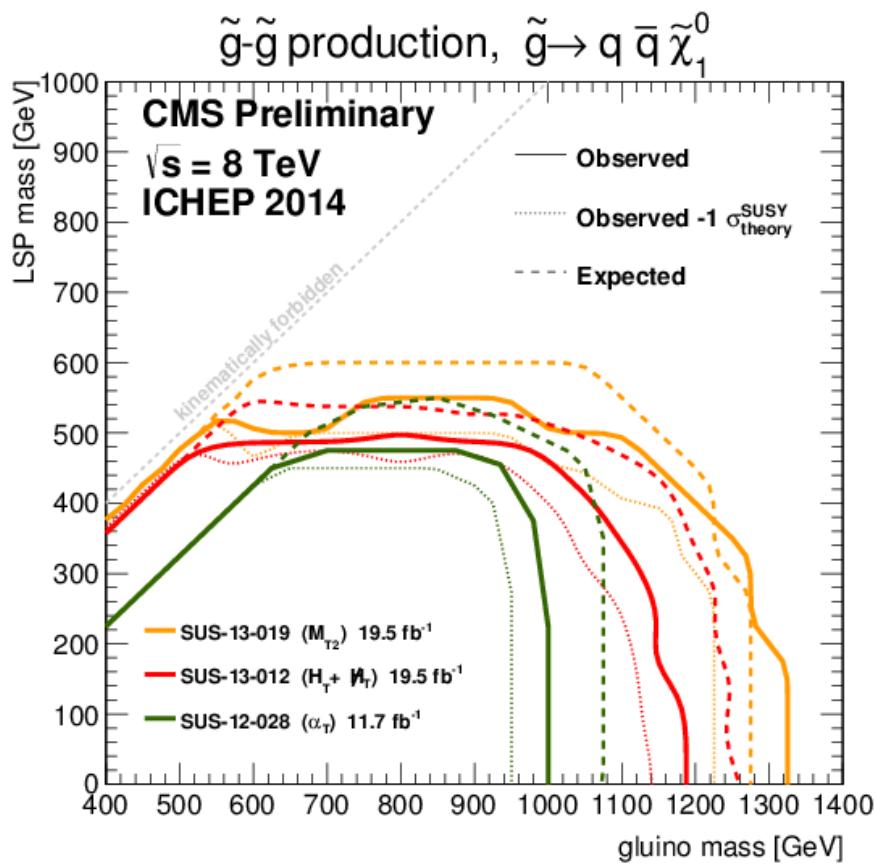


Figure 6.1: Comparison of various exclusion limits derived by different CMS analyses for the SMS T1qqqq. Taken from ... (cms susy public results).

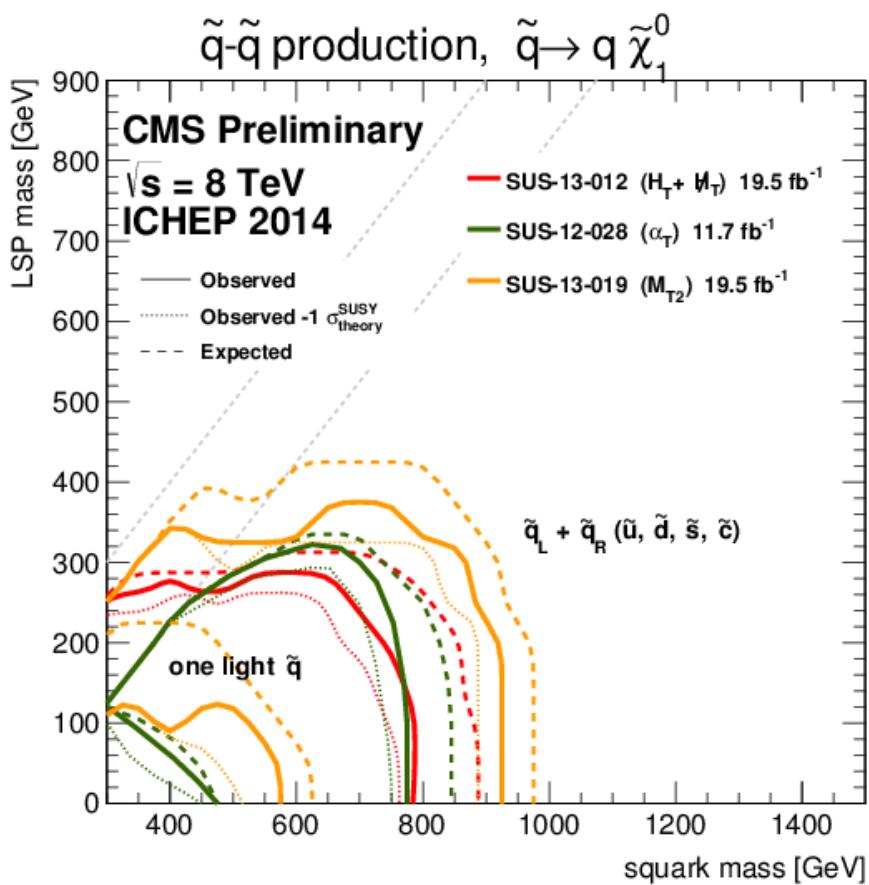


Figure 6.2: Comparison of various exclusion limits derived by different CMS analyses for the SMS T2qq. Taken from ... (cms susy public results).

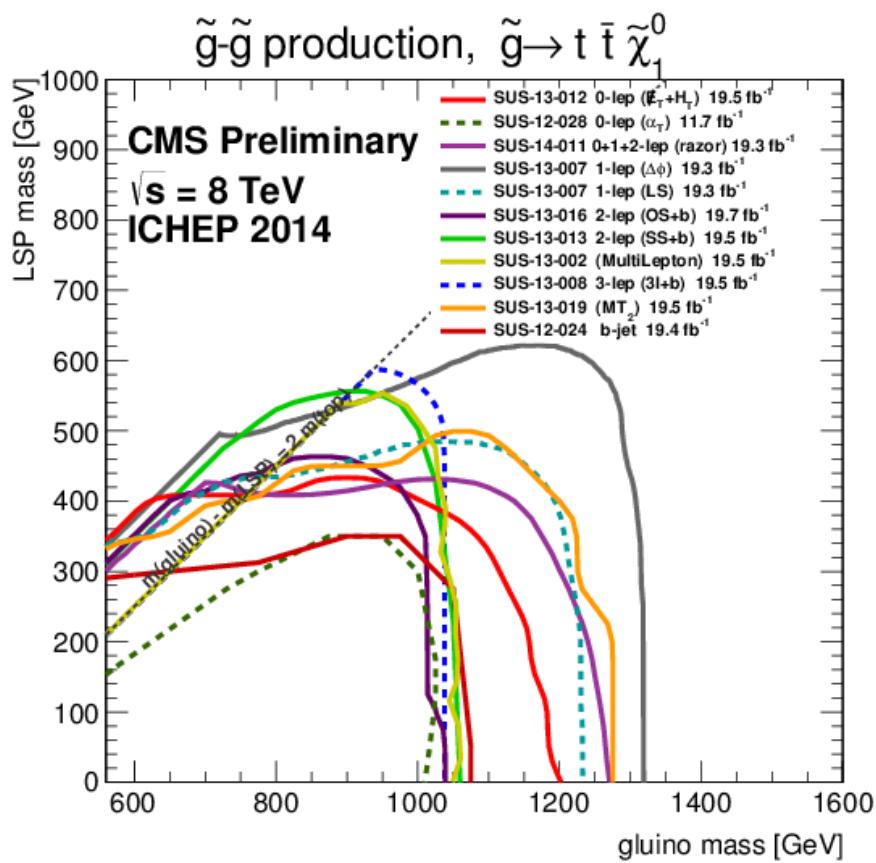


Figure 6.3: Comparison of various exclusion limits derived by different CMS analyses for the SMS T1tttt. Taken from ... (cms susy public results).

7 Prospect Studies for a Search for Top Squarks in Events with Jets and Missing Transverse Momentum at $\sqrt{s} = 13$ TeV

7.1 Data samples

7.2 Top Tagging Efficiency Studies

7.2.1 Top Tag Efficiency

7.2.2 Misidentification Rate

7.3 Studies towards a Suitable Analysis Strategy

7.4 Results and Discussion

7.4.1 Comparison of the Performance of Various Selections

7.4.2 Discussion of Specific Simplified Assumptions in the Analysis

8 Conclusions

Supersymmetry is among the favoured extensions of the standard model of particle physics and one of the main target of searches for new physics within the CMS experiment. Since especially coloured SUSY particles are expected at a high rate at the LHC which predominantly manifest in final states containing jets and missing transverse momentum, it is promising to perform searches based on such signatures which require in addition a good understanding of jet properties.

In this thesis, a measurement of the jet transverse momentum resolution in dijet events corresponding to data with an integrated luminosity of 19.8 fb^{-1} recorded at $\sqrt{s} = 8 \text{ TeV}$ in 2012 by the CMS experiment has been presented. A similar approach to previous analyses was used, but the precision of the measurement could be significantly improved. This was achieved by a correct treatment of statistical uncertainties by considering correlations among inclusive distributions which subsequently also allowed a revised treatment of systematic uncertainties which have been conservatively overestimated in the past. Furthermore, the method has been extended to be able to measure the resolution in the forward part of the detector with higher precision. For $|\eta| = 0.0 - 5.0$ the data-to-simulation ratio of the resolution has been measured. Since no significant trend as function of p_T^{ave} was apparent, the ratio is determined as function of $|\eta|$ only. The ratios obtained for the various $|\eta|$ regions increase from 1.079 ± 0.026 in the central region up to 1.395 ± 0.063 at $|\eta| = 2.8 - 3.2$ and drop again for the outermost region $|\eta| = 3.2 - 5.0$ down to 1.056 ± 0.191 . These values can be utilized to adjust the resolution in simulation to match the one observed in data.

In the second part of this thesis, a search for supersymmetry in events with $N_{\text{jets}} = 3 - 5, 6 - 7, \geq 8$ and large missing momentum in proton-proton collisions at $\sqrt{s} = 8 \text{ TeV}$ in data corresponding to an integrated luminosity of 19.5 fb^{-1} has been presented. Special emphasis is put on the determination of the QCD multijet background which relies strongly on the precise knowledge of the jet response. These QCD background contributions are estimated directly from multijet events in data by modelling momentum mismeasurements based on the jet response. A similar approach has been used already in earlier versions of this analysis where the search has been performed inclusive in the jet multiplicity requiring at least three jets. With the extension of the analysis to multijet events, an adjustment of the method to predict QCD multijet background contributions became inevitable. A dedicated correction to the existing method has been introduced in order to predict the jet multiplicity correctly. Moreover the assignment of systematic uncertainties has been revised in order to consider e.g. the challenging conditions due to pile-up appropriately. In total, the QCD multijet background could be estimated with a precision of approximately 70 – 80% in search regions with non-negligible QCD background contributions. The observed number of events in data are consistent with the expected number of events from standard model processes such that exclusion limits are derived for various simplified supersymmetric models. In the context of this simplified models the production of squarks below 780 GeV and that of gluinos up to 1.1 – 1.2 TeV can be excluded at 95% CL for

LSP masses not exceeding 100 GeV.

In addition to inclusive searches targeting gluinos and squarks, CMS has also performed searches for direct production of top squarks with data obtained at $\sqrt{s} = 8$ TeV. With this data direct production of top squarks decaying into top and LSP could be excluded up to top squarks masses of approximately 750 GeV for LSP masses below 100 GeV. In order to extend the mass reach of such searches to even higher top squark masses in the next run period of the LHC which will start in 2015 with a center of mass energy of 13 TeV suitable selection criteria for such analyses have been investigated in the third part of this thesis. Studies presented here are based on events with several jets, large momentum imbalance and no leptons. One key aspect is the application of dedicated algorithms for the identification of decay products emerging from boosted top decays in order to separate possible SUSY signal events from standard model background. Various different discriminating variables have been investigated. Moreover the sensitivity of several selections – also that one of the all-hadronic stop search performed at $\sqrt{s} = 8$ TeV – has been compared by deriving the expected exclusion reach for data corresponding to 19.5 fb^{-1} at $\sqrt{s} = 13$ TeV. Selection criteria could be identified extending the mass reach of direct stop searches up to roughly 1 TeV for LSP masses below 100 GeV.

Summary
in den
Kontext
"Natural
SUSY"
einbetten

Bibliography

- [1] J. B. et al. (Particle Data Group)), “Review of Particle Physics”, *Phys. Rev. D* **86** (Jul, 2012) 010001. doi:10.1103/PhysRevD.86.010001.
- [2] N. Cabibbo, “Unitary Symmetry and Leptonic Decays”, *Phys. Rev. Lett.* **10** (Jun, 1963) 531 – 533. doi:10.1103/PhysRevLett.10.531.
- [3] M. Kobayashi and T. Maskawa, “*CP*-Violation in the Renormalizable Theory of Weak Interaction”, *Progress of Theoretical Physics* **49** (1973), no. 2, 652–657. doi:10.1143/PTP.49.652.
- [4] R. Alkofer and J. Greensite, “Quark Confinement: The Hard Problem of Hadron Physics”, *J.Phys.* **G34** (2007) S3, arXiv:hep-ph/0610365. doi:10.1088/0954-3899/34/7/S02.
- [5] H. D. Politzer, “Reliable Perturbative Results for Strong Interactions?”, *Phys. Rev. Lett.* **30** (Jun, 1973) 1346 – 1349. doi:10.1103/PhysRevLett.30.1346.
- [6] D. J. Gross and F. Wilczek, “Ultraviolet Behavior of Non-Abelian Gauge Theories”, *Phys. Rev. Lett.* **30** (Jun, 1973) 1343 – 1346. doi:10.1103/PhysRevLett.30.1343.
- [7] S. Glashow, “Partial Symmetries of Weak Interactions”, *Nucl.Phys.* **22** (1961) 579–588. doi:10.1016/0029-5582(61)90469-2.
- [8] S. Weinberg, “A Model of Leptons”, *Phys.Rev.Lett.* **19** (1967) 1264–1266. doi:10.1103/PhysRevLett.19.1264.
- [9] P. W. Higgs, “Broken Symmetries and the Masses of Gauge Bosons”, *Phys. Rev. Lett.* **13** (1964) 508 – 509. doi:10.1103/PhysRevLett.13.508.
- [10] F. Englert and R. Brout, “Broken Symmetry and the Mass of Gauge Vector Mesons”, *Phys. Rev. Lett.* **13** (1964) 321 – 323. doi:10.1103/PhysRevLett.13.321.
- [11] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, “Global Conservation Laws and Massless Particles”, *Phys. Rev. Lett.* **13** (1964) 585 – 587. doi:10.1103/PhysRevLett.13.585.
- [12] ATLAS Collaboration Collaboration, “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”, *Phys.Lett.* **B716** (2012) 1–29, arXiv:1207.7214. doi:10.1016/j.physletb.2012.08.020.

- [13] CMS Collaboration Collaboration, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”, *Phys.Lett.* **B716** (2012) 30–61, [arXiv:1207.7235](https://arxiv.org/abs/1207.7235). doi:[10.1016/j.physletb.2012.08.021](https://doi.org/10.1016/j.physletb.2012.08.021).
- [14] ATLAS Collaboration Collaboration, “Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC”, *Phys.Lett.* **B726** (2013) 88–119, [arXiv:1307.1427](https://arxiv.org/abs/1307.1427). doi:[10.1016/j.physletb.2013.08.010](https://doi.org/10.1016/j.physletb.2013.08.010).
- [15] ATLAS Collaboration Collaboration, “Measurements of Higgs boson production and couplings in the four-lepton channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector”, [arXiv:1408.5191](https://arxiv.org/abs/1408.5191).
- [16] ATLAS Collaboration Collaboration, “Measurement of Higgs boson production in the diphoton decay channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector”, [arXiv:1408.7084](https://arxiv.org/abs/1408.7084).
- [17] CMS Collaboration Collaboration, “Precise determination of the mass of the Higgs boson and studies of the compatibility of its couplings with the standard model”, Technical Report CMS-PAS-HIG-14-009, CERN, Geneva, 2014.
- [18] M. Beyer, ed., “CP Violation in Particle, Nuclear, and Astrophysics”. Lecture Notes in Physics. Springer, 2010. ISBN 3-642-07830-3.
- [19] Planck Collaboration Collaboration, “Planck 2013 results. XVI. Cosmological parameters”, *Astron.Astrophys.* (2014) [arXiv:1303.5076](https://arxiv.org/abs/1303.5076). doi:[10.1051/0004-6361/201321591](https://doi.org/10.1051/0004-6361/201321591).
- [20] G. Bertone, D. Hooper, and J. Silk, “Particle dark matter: Evidence, candidates and constraints”, *Phys.Rept.* **405** (2005) 279–390, [arXiv:hep-ph/0404175](https://arxiv.org/abs/hep-ph/0404175). doi:[10.1016/j.physrep.2004.08.031](https://doi.org/10.1016/j.physrep.2004.08.031).
- [21] I. J. Aitchison, “Supersymmetry and the MSSM: An Elementary introduction”, [arXiv:hep-ph/0505105](https://arxiv.org/abs/hep-ph/0505105).
- [22] S. P. Martin, “A Supersymmetry primer”, *Adv.Ser.Direct.High Energy Phys.* **21** (2010) 1–153, [arXiv:hep-ph/9709356](https://arxiv.org/abs/hep-ph/9709356).
- [23] S. Dimopoulos and D. W. Sutter, “The Supersymmetric flavor problem”, *Nucl.Phys.* **B452** (1995) 496–512, [arXiv:hep-ph/9504415](https://arxiv.org/abs/hep-ph/9504415). doi:[10.1016/0550-3213\(95\)00421-N](https://doi.org/10.1016/0550-3213(95)00421-N).
- [24] A. H. Chamseddine, R. L. Arnowitt, and P. Nath, “Locally Supersymmetric Grand Unification”, *Phys.Rev.Lett.* **49** (1982) 970. doi:[10.1103/PhysRevLett.49.970](https://doi.org/10.1103/PhysRevLett.49.970).
- [25] L. Alvarez-Gaume, J. Polchinski, and M. B. Wise, “Minimal Low-Energy Supergravity”, *Nucl.Phys.* **B221** (1983) 495. doi:[10.1016/0550-3213\(83\)90591-6](https://doi.org/10.1016/0550-3213(83)90591-6).
- [26] G. L. Kane, C. F. Kolda, L. Roszkowski et al., “Study of constrained minimal supersymmetry”, *Phys.Rev.* **D49** (1994) 6173–6210, [arXiv:hep-ph/9312272](https://arxiv.org/abs/hep-ph/9312272). doi:[10.1103/PhysRevD.49.6173](https://doi.org/10.1103/PhysRevD.49.6173).

- [27] H. Baer, C. Balazs, A. Belyaev et al., “Updated constraints on the minimal supergravity model”, *JHEP* **0207** (2002) 050, [arXiv:hep-ph/0205325](https://arxiv.org/abs/hep-ph/0205325).
- [28] M. Dine and W. Fischler, “A Phenomenological Model of Particle Physics Based on Supersymmetry”, *Phys.Lett.* **B110** (1982) 227.
`doi:10.1016/0370-2693(82)91241-2`.
- [29] L. Alvarez-Gaume, M. Claudson, and M. B. Wise, “Low-Energy Supersymmetry”, *Nucl.Phys.* **B207** (1982) 96. `doi:10.1016/0550-3213(82)90138-9`.
- [30] H1 Collaboration, “A Search for Selectrons and Squarks at HERA”, *Phys. Lett.* **B380** (1996) 461, [arXiv:hep-ex/9605002](https://arxiv.org/abs/hep-ex/9605002). `doi:10.1016/0370-2693(96)00640-5`.
- [31] ZEUS Collaboration Collaboration, “Search for selectron and squark production in e^+p collisions at HERA”, *Phys.Lett.* **B434** (1998) 214–230,
[arXiv:hep-ex/9806019](https://arxiv.org/abs/hep-ex/9806019). `doi:10.1016/S0370-2693(98)00817-X`.
- [32] ALEPH Collaboration, “Absolute Mass Lower Limit for the Lightest Neutralino of the MSSM from e^+e^- data at \sqrt{s} up to 209 GeV”, *Phys. Lett.* **B 583** (2004) 247. See also references therein. `doi:10.1016/j.physletb.2003.12.066`.
- [33] DELPHI Collaboration, “Searches for Supersymmetric Particles in e^+e^- Collisions up to 208 GeV and Interpretation of the Results within the MSSM”, *Eur. Phys. J. C* **31** (2003) 421, [arXiv:hep-ex/0311019](https://arxiv.org/abs/hep-ex/0311019). See also references therein.
`doi:10.1140/epjc/s2003-01355-5`.
- [34] L3 Collaboration, “Search for Scalar Leptons and Scalar Quarks at LEP”, *Phys. Lett.* **B 580** (2004) 37, [arXiv:hep-ex/0310007](https://arxiv.org/abs/hep-ex/0310007). See also references therein.
`doi:10.1016/j.physletb.2003.10.010`.
- [35] OPAL Collaboration, “Search for Chargino and Neutralino Production at $\sqrt{s} = 192$ GeV to 209 GeV at LEP”, *Eur. Phys. J. C* **35** (2004) 1, [arXiv:hep-ex/0401026](https://arxiv.org/abs/hep-ex/0401026). See also references therein. `doi:10.1140/epjc/s2004-01780-x`.
- [36] ALEPH, DELPHI, L3 and OPAL Collaborations, “Joint SUSY Working Group”, *LEPSUSYWG/02-06-2* (2004). <http://lepsusy.web.cern.ch/lepsusy>.
- [37] CDF Collaboration Collaboration, “Inclusive Search for Squark and Gluino Production in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ -TeV”, *Phys.Rev.Lett.* **102** (2009) 121801, [arXiv:0811.2512](https://arxiv.org/abs/0811.2512). `doi:10.1103/PhysRevLett.102.121801`.
- [38] DØ Collaboration, Abazov et al., “Search for squarks and gluinos in events with jets and missing transverse energy using 2.1 fb^{-1} of $p\bar{p}$ collision data at $\sqrt{s} = 1.96 \text{ TeV}$ ”, *Phys. Lett.* **B 660** (2008) 449, [arXiv:0712.3805](https://arxiv.org/abs/0712.3805).
- [39] DØ Collaboration, “Search for associated production of charginos and neutralinos in the trilepton final state using 2.3 fb^{-1} of data”, *Phys. Lett.* **B 680** (2009), no. 1, 34 – 43. `doi:DOI: 10.1016/j.physletb.2009.08.011`.
- [40] G. L. Kane and J. Leveille, “Experimental Constraints on Gluino Masses and Supersymmetric Theories”, *Phys.Lett.* **B112** (1982) 227.
`doi:10.1016/0370-2693(82)90968-6`.

- [41] P. Harrison and C. Llewellyn Smith, “Hadroproduction of Supersymmetric Particles”, *Nucl.Phys.* **B213** (1983) 223. doi:10.1016/0550-3213(83)90510-2.
- [42] E. Reya and D. Roy, “SUPERSYMMETRIC PARTICLE PRODUCTION AT p anti-p COLLIDER ENERGIES”, *Phys.Rev.* **D32** (1985) 645. doi:10.1103/PhysRevD.32.645.
- [43] S. Dawson, E. Eichten, and C. Quigg, “Search for Supersymmetric Particles in Hadron - Hadron Collisions”, *Phys.Rev.* **D31** (1985) 1581. doi:10.1103/PhysRevD.31.1581.
- [44] H. Baer and X. Tata, “Component Formulae for Hadroproduction of Left-handed and Right-handed Squarks”, *Phys.Lett.* **B160** (1985) 159. doi:10.1016/0370-2693(85)91484-4.
- [45] M. Kramer, A. Kulesza, R. van der Leeuw et al., “Supersymmetry production cross sections in pp collisions at $\sqrt{s} = 7$ TeV”, arXiv:1206.2892.
- [46] CMS Collaboration, “CMS Supersymmetry Physics Results”, *Public CMS Wiki* (2014). <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSUS>, Topic revision: r15.
- [47] CMS Collaboration Collaboration, “Interpretation of Searches for Supersymmetry with simplified Models”, *Phys.Rev.* **D88** (2013), no. 5, 052017, arXiv:1301.2175. doi:10.1103/PhysRevD.88.052017.
- [48] O. S. Brüning, P. Collier, P. Lebrun et al., “LHC Design Report”, volume 1: The LHC Main Ring. CERN, Geneva, 2004. CERN-2004-003-V-1.
- [49] L. Evans and P. Bryant, “LHC Machine”, *Journal of Instrumentation* **3** (2008), no. 08, S08001. doi:10.1088/1748-0221/3/08/S08001.
- [50] CERN, “LEP Design Report”, volume 2: The LEP Main Ring. CERN, Geneva, 1984. CERN-LEP-84-01.
- [51] CERN, “LHC: the guide (English version)”. CERN-Brochure-2009-003-Eng, 2009.
- [52] 2012. W.J. Stirling, private communication.
- [53] ATLAS Collaboration, “The ATLAS Experiment at the CERN Large Hadron Collider”, *JINST* **3** (2008) S08003. doi:10.1088/1748-0221/3/08/S08003.
- [54] CMS Collaboration, “The CMS experiment at the CERN LHC”, *JINST* **3** (2008) S08004. doi:10.1088/1748-0221/3/08/S08004.
- [55] CMS Collaboration, “CMS Detector Performance and Software”, volume 1 of *CMS Physics Technical Design Report*. CERN, Geneva, 2006. CMS-TDR-008-1. CERN-LHCC-2006-001.
- [56] LHCb Collaboration, “The LHCb Detector at the LHC”, *JINST* **3** (2008) S08005. doi:10.1088/1748-0221/3/08/S08005.

- [57] ALICE Collaboration, “The ALICE experiment at the CERN LHC”, *JINST* **3** (2008) S08002. doi:[10.1088/1748-0221/3/08/S08002](https://doi.org/10.1088/1748-0221/3/08/S08002).
- [58] CMS Collaboration, “The CMS tracker system project”, volume 5 of *Technical Design Report CMS*. CERN, Geneva, 1997. CMS-TDR-005. CERN-LHCC-98-006.
- [59] CMS Collaboration, “Luminosity Information for the 2012 Proton Run”, 2012. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>.
- [60] “LHC machine outreach”, 2013. <http://lhc-machine-outreach.web.cern.ch/lhc-machine-outreach/>.
- [61] M. Lamont, “Status of the LHC”, *J.Phys.Conf.Ser.* **455** (2013) 012001. doi:[10.1088/1742-6596/455/1/012001](https://doi.org/10.1088/1742-6596/455/1/012001).
- [62] M. H. Seymour and M. Marx, “Monte Carlo Event Generators”, [arXiv:1304.6677](https://arxiv.org/abs/1304.6677).
- [63] A. Buckley, J. Butterworth, S. Gieseke et al., “General-purpose event generators for LHC physics”, *Phys.Rept.* **504** (2011) 145–233, [arXiv:1101.2599](https://arxiv.org/abs/1101.2599). doi:[10.1016/j.physrep.2011.03.005](https://doi.org/10.1016/j.physrep.2011.03.005).
- [64] N. Metropolis and S. Ulam, “The Monte Carlo Method”, *Journal of the American Statistical Association* **44** (1949), no. 247, 335–341. doi:[10.2307/2280232](https://doi.org/10.2307/2280232).
- [65] J. C. Collins and D. E. Soper, “The Theorems of Perturbative QCD”, *Ann.Rev.Nucl.Part.Sci.* **37** (1987) 383–409. doi:[10.1146/annurev.ns.37.120187.002123](https://doi.org/10.1146/annurev.ns.37.120187.002123).
- [66] J. C. Collins, D. E. Soper, and G. F. Sterman, “Factorization of Hard Processes in QCD”, *Adv.Ser.Direct.High Energy Phys.* **5** (1988) 1–91, [arXiv:hep-ph/0409313](https://arxiv.org/abs/hep-ph/0409313).
- [67] B. Andersson, G. Gustafson, G. Ingelman et al., “Parton fragmentation and string dynamics”, *Physics Reports* **97** (1983), no. 2 - 3, 31 – 145. doi:[http://dx.doi.org/10.1016/0370-1573\(83\)90080-7](http://dx.doi.org/10.1016/0370-1573(83)90080-7).
- [68] D. Amati and G. Veneziano, “Preconfinement as a property of perturbative QCD”, *Physics Letters B* **83** (1979), no. 1, 87 – 92. doi:[http://dx.doi.org/10.1016/0370-2693\(79\)90896-7](http://dx.doi.org/10.1016/0370-2693(79)90896-7).
- [69] A. Bassetto, M. Ciafaloni, and G. Marchesini, “Color singlet distributions and mass damping in perturbative QCD”, *Physics Letters B* **83** (1979), no. 2, 207 – 212. doi:[http://dx.doi.org/10.1016/0370-2693\(79\)90687-7](http://dx.doi.org/10.1016/0370-2693(79)90687-7).
- [70] T. Sjostrand, S. Mrenna, and P. Z. Skands, “PYTHIA 6.4 Physics and Manual”, *JHEP* **0605** (2006) 026, [arXiv:hep-ph/0603175](https://arxiv.org/abs/hep-ph/0603175). doi:[10.1088/1126-6708/2006/05/026](https://doi.org/10.1088/1126-6708/2006/05/026).
- [71] M. Bahr, S. Gieseke, M. Gigg et al., “Herwig++ Physics and Manual”, *Eur.Phys.J.* **C58** (2008) 639–707, [arXiv:0803.0883](https://arxiv.org/abs/0803.0883). doi:[10.1140/epjc/s10052-008-0798-9](https://doi.org/10.1140/epjc/s10052-008-0798-9).

- [72] J. Alwall, R. Frederix, S. Frixione et al., “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”, *JHEP* **1407** (2014) 079, [arXiv:1405.0301](https://arxiv.org/abs/1405.0301).
[doi:10.1007/JHEP07\(2014\)079](https://doi.org/10.1007/JHEP07(2014)079).
- [73] C. Oleari, “The POWHEG-BOX”, *Nucl.Phys.Proc.Suppl.* **205-206** (2010) 36–41, [arXiv:1007.3893](https://arxiv.org/abs/1007.3893). [doi:10.1016/j.nuclphysbps.2010.08.016](https://doi.org/10.1016/j.nuclphysbps.2010.08.016).
- [74] P. Nason, “A New method for combining NLO QCD with shower Monte Carlo algorithms”, *JHEP* **0411** (2004) 040, [arXiv:hep-ph/0409146](https://arxiv.org/abs/hep-ph/0409146).
[doi:10.1088/1126-6708/2004/11/040](https://doi.org/10.1088/1126-6708/2004/11/040).
- [75] S. Agostinelli, J. Allison, K. Amako et al., “G4 – a simulation toolkit”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** (2003), no. 3, 250 – 303.
[doi:10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8).
- [76] CMS Collaboration Collaboration, “The fast simulation of the CMS detector at LHC”, *J.Phys.Conf.Ser.* **331** (2011) 032049.
[doi:10.1088/1742-6596/331/3/032049](https://doi.org/10.1088/1742-6596/331/3/032049).
- [77] CMS Collaboration Collaboration, “Comparison of the Fast Simulation of CMS with the first LHC data”, CMS-DP-2010-039.
- [78] CMS Collaboration, “Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and E_T^{miss} ”, *CMS Physics Analysis Summary CMS-PAS-PFT-09-001* (2009).
- [79] W. Adam, B. Mangano, T. Speer et al., “Track Reconstruction in the CMS tracker”, Technical Report CMS-NOTE-2006-041, CERN, Geneva, Dec, 2006.
- [80] G. P. Salam, “Towards Jetography”, *Eur.Phys.J.* **C67** (2010) 637–686, [arXiv:0906.1833](https://arxiv.org/abs/0906.1833). [doi:10.1140/epjc/s10052-010-1314-6](https://doi.org/10.1140/epjc/s10052-010-1314-6).
- [81] G. P. Salam and G. Soyez, “A Practical Seedless Infrared-Safe Cone jet algorithm”, *JHEP* **0705** (2007) 086, [arXiv:0704.0292](https://arxiv.org/abs/0704.0292).
[doi:10.1088/1126-6708/2007/05/086](https://doi.org/10.1088/1126-6708/2007/05/086).
- [82] S. D. Ellis and D. E. Soper, “Successive combination jet algorithm for hadron collisions”, *Phys. Rev. D* **48** (Oct, 1993) 3160 – 3166.
[doi:10.1103/PhysRevD.48.3160](https://doi.org/10.1103/PhysRevD.48.3160).
- [83] Y. L. Dokshitzer, G. Leder, S. Moretti et al., “Better jet clustering algorithms”, *JHEP* **9708** (1997) 001, [arXiv:hep-ph/9707323](https://arxiv.org/abs/hep-ph/9707323).
[doi:10.1088/1126-6708/1997/08/001](https://doi.org/10.1088/1126-6708/1997/08/001).
- [84] M. Wobisch and T. Wengler, “Hadronization corrections to jet cross-sections in deep inelastic scattering”, [arXiv:hep-ph/9907280](https://arxiv.org/abs/hep-ph/9907280). DESY-THESIS-2000-049.
- [85] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- k_T jet clustering algorithm”, *Journal of High Energy Physics* **2008** (2008), no. 04, 063, [arXiv:0802.1189](https://arxiv.org/abs/0802.1189).
[doi:10.1088/1126-6708/2008/04/063](https://doi.org/10.1088/1126-6708/2008/04/063).

- [86] CMS Collaboration Collaboration, “Jet Energy Corrections and Uncertainties. Detector Performance Plots for 2012.”,.
- [87] M. Cacciari, G. P. Salam, and G. Soyez, “FastJet user manual”, [arXiv:1111.6097](https://arxiv.org/abs/1111.6097).
- [88] M. Cacciari and G. P. Salam, “Dispelling the N^3 myth for the k_t jet-finder”, *Phys.Lett.* **B641** (2006) 57–61, [arXiv:hep-ph/0512210](https://arxiv.org/abs/hep-ph/0512210).
[doi:10.1016/j.physletb.2006.08.037](https://doi.org/10.1016/j.physletb.2006.08.037).
- [89] CMS Collaboration, “Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS”, *Journal of Instrumentation* **6** (2011) P11002, [arXiv:1107.4277](https://arxiv.org/abs/1107.4277). [doi:10.1088/1748-0221/6/11/P11002](https://doi.org/10.1088/1748-0221/6/11/P11002).
- [90] CMS Collaboration, “The Jet Plus Tracks Algorithm for Calorimeter Jet Energy Corrections in CMS”, *CMS Physics Analysis Summary CMS-PAS-JME-09-002* (2009).
- [91] CMS Collaboration Collaboration, “8 TeV Jet Energy Corrections and Uncertainties based on 19.8 fb^{-1} of data in CMS”,.
- [92] CMS Collaboration Collaboration, “Identification of b-quark jets with the CMS experiment”, *JINST* **8** (2013) P04013, [arXiv:1211.4462](https://arxiv.org/abs/1211.4462).
[doi:10.1088/1748-0221/8/04/P04013](https://doi.org/10.1088/1748-0221/8/04/P04013).
- [93] CMS Collaboration Collaboration, “Performance of b tagging at $\sqrt{s}=8 \text{ TeV}$ in multijet, ttbar and boosted topology events”, Technical Report CMS-PAS-BTV-13-001, CERN, Geneva, 2013.
- [94] ATLAS Collaboration, CDF Collaboration, CMS Collaboration, D0 Collaboration Collaboration, “First combination of Tevatron and LHC measurements of the top-quark mass”, [arXiv:1403.4427](https://arxiv.org/abs/1403.4427).
- [95] CMS Collaboration Collaboration, “Boosted Top Jet Tagging at CMS”,.
- [96] CMS Collaboration Collaboration, “A Cambridge-Aachen (C-A) based Jet Algorithm for boosted top-jet tagging”, Technical Report CMS-PAS-JME-09-001, CERN, 2009. Geneva, Jul, 2009.
- [97] D. E. Kaplan, K. Rehermann, M. D. Schwartz et al., “Top Tagging: A Method for Identifying Boosted Hadronically Decaying Top Quarks”, *Phys.Rev.Lett.* **101** (2008) 142001, [arXiv:0806.0848](https://arxiv.org/abs/0806.0848). [doi:10.1103/PhysRevLett.101.142001](https://doi.org/10.1103/PhysRevLett.101.142001).
- [98] T. Plehn, M. Spannowsky, M. Takeuchi et al., “Stop Reconstruction with Tagged Tops”, *JHEP* **1010** (2010) 078, [arXiv:1006.2833](https://arxiv.org/abs/1006.2833).
[doi:10.1007/JHEP10\(2010\)078](https://doi.org/10.1007/JHEP10(2010)078).
- [99] M. Schröder, “Quality of Jet Measurements and Impact on a Search for New Physics at CMS”. PhD thesis, Universität Hamburg, Germany, 2012.
DESY-THESIS-2012-042.
- [100] F. Pandolfi, D. del Re, M. Voutilainen, “Jet Response and Resolution Measurement with Photon+Jet Events at $\sqrt{s} = 7 \text{ TeV}$ ”, *CMS AN-2010/141* (2010).

- [101] F. Pandolfi, D. del Re, M. Voutilainen, “Update on Jet Response and Resolution Measurements with Photon+Jet Events at $\sqrt{s} = 7$ TeV”, *CMS AN-2011/004* (2011).
- [102] T. Lenz, C. Sander, P. Schleper et al., “Jet Energy Resolution Measurement using Photon+Jet Events at $\sqrt{s} = 8$ TeV”, *CMS AN-2013/179* (2013).
- [103] CMS Collaboration, “Measurement of the Underlying Event Activity at the LHC with $\sqrt{s} = 7$ TeV and Comparison with $\sqrt{s} = 0.9$ TeV”, *JHEP* **1109** (2011) 109, [arXiv:1107.0330](https://arxiv.org/abs/1107.0330). doi:10.1007/JHEP09(2011)109.
- [104] J. Pumplin, D. Stump, J. Huston et al., “New generation of parton distributions with uncertainties from global QCD analysis”, *JHEP* **0207** (2002) 012, [arXiv:hep-ph/0201195](https://arxiv.org/abs/hep-ph/0201195). doi:10.1088/1126-6708/2002/07/012.
- [105] D0 Collaboration, “High- p_T jets in $p\bar{p}$ collisions at $\sqrt{s} = 630$ GeV and 1800 GeV”, *Phys.Rev.* **D64** (2001) 032003, [arXiv:hep-ex/0012046](https://arxiv.org/abs/hep-ex/0012046). doi:10.1103/PhysRevD.64.032003.
- [106] D0 Collaboration, “Measurement of the inclusive jet cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV”, *Phys.Rev.* **D85** (2012) 052006, [arXiv:1110.3771](https://arxiv.org/abs/1110.3771). doi:10.1103/PhysRevD.85.052006.
- [107] ATLAS Collaboration, “Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector”, *Eur.Phys.J.* **C73** (2013) 2306, [arXiv:1210.6210](https://arxiv.org/abs/1210.6210). doi:10.1140/epjc/s10052-013-2306-0.
- [108] CMS Collaboration, “Measurement of Differential Jet Cross Sections at $\sqrt{s} = 7$ TeV with the CMS Detector”, *CMS Physics Analysis Summary CMS-PAS-QCD-11-004* (2012).
- [109] CMS Collaboration Collaboration, “CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update”, Technical Report CMS-PAS-LUM-13-001, CERN, Geneva, 2013.
- [110] “Hamburg Wiki Calibration”. <https://twiki.cern.ch/twiki/bin/viewauth/CMS/HamburgWikiAnalysisCalibration>, Jan., 2014.
- [111] F. James and M. Roos, “Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations”, *Comput.Phys.Commun.* **10** (1975) 343–367. doi:10.1016/0010-4655(75)90039-9.

