Measurement of the Dijet Angular Distributions and Search for Quark Compositeness with the CMS Experiment

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH Aachen University zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigte Dissertation

vorgelegt von

Diplom-Physiker

Andreas Dominik Hinzmann

aus Düsseldorf

Berichter: Universitätprofessor Dr. Martin Erdmann

Universitätprofessor Dr. Thomas Hebbeker

Tag der mündlichen Prüfung: 7. Oktober 2011

Diese Dissertation ist auf den Internetseiten der Hochschulbibliothek online verfügbar.

Contents

1 Introduction **1**

2 The Standard Model of Elementary Particles and Quantum Chromodynamics **3**
2.1 The Standard Model of Elementary Particles 3
2.2 Quantum Chromodynamics in Proton-Proton Collisions 8
2.2.1 Perturbative QCD 8
2.2.2 The Factorization Theorem 10

3 The Dijet Angular Distributions and Quark Compositeness **13**
3.1 The Physics of Dijet Events 13
3.2 The Kinematics of Dijet Events 14
3.3 The Dijet Angular Distributions in QCD 15
3.4 Signatures of Quark Compositeness 17

4 The LHC Accelerator and the CMS Experiment **19**
4.1 The Large Hadron Collider 19
4.2 The CMS Detector 21
4.2.1 The Tracker 22
4.2.2 The Calorimeter 23
4.2.3 The Muon System 25
4.2.4 The Trigger System 25

5 Simulation of Collision Events **29**
5.1 Monte Carlo Event Generators 29
5.2 The CMS Detector Simulation 32

6 CMS Offline Computing and Analysis Software **33**
6.1 CMS Offline Computing 33
6.2 VISPA and PXL 36

7 Jet Reconstruction and Performance **41**
7.1 Jet Clustering Algorithms 41
7.2 Jet Reconstruction Methods in CMS 43
7.3 Noise Filtering 44
7.4 Jet Energy Calibration 46
7.5 Jet Energy Resolution 48
7.6 Jet Angular Resolution 49
7.7 Jet Reconstruction efficiency 51

8 The Data Sample and Event Selection **53**
8.1 The Dataset . . . 53
8.2 Dijet Reconstruction and Selection . . . 53
8.3 Triggers . . . 55
8.4 Event Cleaning . . . 57
8.5 Kinematic Distributions in Data and Simulation . . . 63

9 Measurement of the Dijet Angular Distributions **83**
9.1 Correction of Detector Effects . . . 83
9.2 Experimental Uncertainties . . . 87
9.2.1 Statistical Uncertainty . . . 87
9.2.2 Jet Energy Scale . . . 87
9.2.3 Jet Energy Resolution . . . 87
9.2.4 Detector Effects Correction . . . 90
9.2.5 Other Detector Effects . . . 90
9.2.6 Impact of Pile-up Interactions . . . 90
9.3 Data Distributions . . . 94

10 Predictions of the Dijet Angular Distributions **97**
10.1 QCD Predictions . . . 97
10.2 Theoretical Uncertainties of the QCD Predictions . . . 100
10.2.1 Renormalization and Factorization Scales . . . 100
10.2.2 Parton Distribution Functions . . . 103
10.2.3 Non-Perturbative Corrections . . . 103
10.3 Contact Interaction Predictions . . . 103

11 Search Techniques for Quark Compositeness **109**
11.1 Statistical Method . . . 109
11.2 Treatment of Uncertainties . . . 111
11.3 Limits Setting Procedure . . . 112
11.4 Cross checks . . . 113

12 Results and Discussion **115**
12.1 Test of Perturbative QCD . . . 115
12.2 Limits on New Physics . . . 118

13 Conclusions and Outlook **121**

Appendix **123**

Bibliography **132**

1 Introduction

The Large Hadron Collider (LHC) at the Conseil Européen pour la Recherche Nucléaire (CERN) allows to study the interactions of quarks and gluons in a yet unexplored energy regime. In 2010, the LHC delivered an integrated luminosity of more than 36 pb^{-1} of proton-proton collisions at a center-of-mass energy of $\sqrt{s} = 7$ TeV. In these proton-proton collisions, the interactions of the constituent quarks and gluons produced a considerable amount of jets of particles with transverse momenta above 1 TeV. Well suited for the study of these jet processes is the Compact Muon Solenoid (CMS) experiment situated at the LHC point 5 as it can measure jets with the necessary energy and angular resolutions over a large range of transverse momentum ($\sim$30 GeV$< p_T <\sim$7 TeV) and pseudo-rapidity ($|\eta| \leq 5$, where $\eta \equiv -\ln\tan(\theta/2)$ and θ is the polar angle relative to the proton beam). In this analysis, the data collected by the CMS experiment is used to test the theory of Quantum Chromodynamics which predicts the cross section of jet processes and at the same time to identify deviations as signals of new physics.

A natural idea of new physics beyond the Standard Model of elementary particle physics is the existence of a substructure of quarks (quark compositeness). Models describing quarks as bound states of constituent particles may be able to explain the number of quark generations, quark masses and charges. A common signature of these models are additional contact interactions between quarks in high-momentum-transfer interactions, observable in the cross section of jet processes.

Inspired by the Rutherford experiment, the scattering angle of two-jet processes (dijets) is measured to study the point-like quark and gluon scattering processes predicted by QCD. The dijet scattering angle is expressed in terms of $\chi_{\text{dijet}} = e^{|y_1 - y_2|}$, where y_1 and y_2 are the rapidities of the two jets, $y \equiv \frac{1}{2}\ln\left[(E + p_z)\,/\,(E - p_z)\right]$, and p_z is the projection of the jet momentum along the beam axis. The choice of the variable χ_{dijet} is motivated by the fact that the normalized differential cross section $\frac{1}{\sigma}\frac{d\sigma}{d\chi_{\text{dijet}}}$ (the dijet angular distribution) is flat in this variable for Rutherford scattering, characteristic for spin-1 particle exchange. In contrast to QCD which predicts a dijet angular distribution similar to Rutherford scattering, new physics, such as quark compositeness, that might have a more isotropic dijet angular distribution would produce an excess at low values of χ_{dijet}. Since the shapes of the dijet angular distributions for the $qg \to qg$, $qq' \to qq'$ and $gg \to gg$ scattering processes are similar, the QCD prediction does not strongly depend on the parton distribution functions (PDFs) which describe the momentum distribution of the partons inside the protons. Due to the normalization, the dijet angular distribution has a reduced sensitivity to several predominant experimental uncertainties (e.g. the jet energy scale and luminosity uncertainties). The dijet angular distribution is therefore well suited to test the predictions of QCD and to search for signals of new physics, in particular for signs of quark compositeness.

In the following a measurement of the dijet angular distributions and a search for quark compositeness with the CMS experiment is presented.

2 The Standard Model of Elementary Particles and Quantum Chromodynamics

2.1 The Standard Model of Elementary Particles

The Standard Model (SM) of elementary particle physics describes the structure and interactions of atoms and subatomic particles [1,2]. Its predictions have been tested with high precision and up to the present it endured all experimental checks.

The SM consists of a set of point-like particles and interactions between these particles. According to the SM all matter is made from these particles and in principle all interaction of matter can be understood from the interaction of its point-like constituents.

Force	Range [m]	Relative strength	Force carrier	Mass [GeV/c^2]	'Charge'
Strong force	10^{-15}	1	8 gluons (g)	massless	color
Electromagnetic force	inf	10^{-2}	photon (γ)	massless	charge
Weak force	10^{-18}	10^{-6}	$W^{\pm}$	80.399 ± 0.023	weak-isospin
			Z^0	91.1876 ± 0.0021	

Table 2.1: The three forces described by the Standard Model. The masses correspond to the force carrying bosons and are taken from Ref. [3].

Table 2.1 summarizes the three forces described by the SM. The strong force is responsible for binding the quarks in a nucleon. The electromagnetic force holds together the atomic nucleus and the orbiting electrons of an atom. The most famous effect of the weak force is the nuclear β decay.

The fourth force in nature, the gravitational force, is not described by the SM. However, its impact on small-scale particles is very small since its relative strength compared to the strong force is of order 10^{-39}. Anyway, the search for a theory including all four forces is one of the most interesting challenges of nowadays physics.

In the SM the interactions are described by quantum gauge field theories. In these theories the forces are considered to be due to the exchange of field quanta, which mediate the forces. These field quanta have spin 1 and are called gauge bosons. As presented in Tab. 2.1, the

electromagnetic force is mediated by the photon, the weak force by the $W^{\pm}$ and Z^0 bosons and the strong force by 8 gluons.

Table 2.1 also summarizes which 'charge' is responsible for interaction of particles with the three forces. The responsible quantum number for the electromagnetic force is the electric charge of a particle. The so called weak-isospin is introduced for the weak force. The counterpart of the charge for the strong force is called color charge.

In the following three sections the particles and interactions of the SM as well as the so called Higgs mechanism, which gives rise to particle masses in the SM, will be described.

The Elementary Particles

The SM consists of two classes of particles: the above mentioned gauge bosons with spin 1 which mediate the forces and the spin 1/2 particles, the fermions, of which matter is composed.

Particle	Generation	Mass [MeV/c^2]	Electric charge (Q)	3$^{\text{rd}}$ component of weak-isospin (T_3)	Color charge
Electron (e^-)	1	0.511	-1	+1/2	
Electron neutrino (ν_e)	1	$< 2 \times 10^{-6}$	0	-1/2	
Muon (μ^-)	2	106	-1	+1/2	
Muon neutrino (ν_μ)	2	< 0.19	0	-1/2	
Tau (τ^-)	3	1777	-1	+1/2	
Tau neutrino (ν_τ)	3	< 18.2	0	-1/2	
Up quark (u)	1	1.7-3.3	+2/3	-1/2	rgb
Down quark (d)	1	4.1-5.8	-1/3	+1/2	rgb
Charm quark (c)	2	$(1.27^{+0.07}_{-0.09}) \times 10^3$	+2/3	-1/2	rgb
Strange quark (s)	2	(101^{+29}_{-21})	-1/3	+1/2	rgb
Top quark (t)	3	$(172.0^{+0.9}_{-1.3}) \times 10^3$	+2/3	-1/2	rgb
Bottom quark (b)	3	$(4.19^{+0.18}_{-0.06}) \times 10^3$	-1/3	+1/2	rgb

Table 2.2: Leptons and quarks and their properties. Masses are taken from Ref. [3]. The third component of the weak-isospin is stated for left-handed particles only. 'rgb' in the last column indicates that the quarks exist in three different color charge states 'red', 'green' and 'blue'.

Table 2.2 gives a summary of the fermions in the SM. They can be grouped into leptons and quarks. There are 6 leptons (e^-, ν_e, μ^-, ν_μ, τ^-, ν_τ) and their corresponding antiparticles (e^+, $\bar{\nu}_e$, μ^+, $\bar{\nu}_\mu$, τ^+, $\bar{\nu}_\tau$), as well as 6 quarks (u, d, c, s, t, b) and antiquarks ($\bar{u}$, $\bar{d}$, $\bar{c}$, $\bar{s}$, $\bar{t}$, $\bar{b}$). While quarks participate in all three interactions, leptons do not participate in the strong interaction since they do not carry color charge. Within the leptons, the neutrinos play a special role. Since they carry neither electric charge nor color charge, they take part only in weak interactions which makes their detection very challenging.

Both the leptons and the quarks can be grouped into three generations as shown in Tab. 2.2. When looking at the particle masses a clear hierarchy of the three generation attracts attention.

This mass hierarchy is not explained by the SM and is one of the unanswered questions in particle physics. As a consequence of the mass hierarchy, most of the matter surrounding us consists of fermions from the first generation because they cannot decay any further. Atoms, for example, consist of nucleons formed by up and down quarks which are surrounded by electrons. Heavy particles such as the top quark can only be observed in physics experiments having sufficient energies for producing their particle masses.

From here on, natural units will be used, $\hbar = c = 1$, which means that $[\text{mass}] = [\text{energy}] = [\text{momentum}] = \frac{1}{[\text{length}]} = \frac{1}{[\text{time}]}$.

The Interactions

Electroweak Interactions

The electromagnetic and the weak force can be described in a joint quantum field theory. This theory is named Glashow-Weinberg-Salam model after its developers.

In the Glashow-Weinberg-Salam model fermions are represented by spinor fields ψ. The fermions are separated in left-handed fields (with $\psi_L = \frac{1}{2}(1-\gamma_5)\psi$) and right-handed fields (with $\psi_R = \frac{1}{2}(1+\gamma_5)\psi$), since the weak force was found to act on left-handed fermions only. The left-handed fermions are grouped into weak-isospin doublets while the right-handed fermions form singlets:

$\begin{pmatrix} u \\ d \end{pmatrix}_L$	$\begin{pmatrix} c \\ s \end{pmatrix}_L$	$\begin{pmatrix} t \\ b \end{pmatrix}_L$	$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L$	$\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L$	$\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L$
u_R	c_R	t_R			
d_R	s_R	b_R	e_R	μ_R	τ_R

Right-handed neutrinos are not listed because neutrinos were assumed to be massless in the original SM. Recent measurements have, however, shown that they indeed have a small mass [3].

All fermions are classified by two quantum numbers which define their characteristics in electroweak interactions: the third component of the weak-isospin T_3 and the weak hyper-charge Y. In each doublet the upper element has $T_3 = -\frac{1}{2}$ and the lower element $T_3 = +\frac{1}{2}$. The hyper-charge is given by the relation $Y = 2(Q - T_3)$, where Q is the electric charge.

The dynamics of electroweak forces can be derived from the free particle Lagrangian by demanding invariance under gauge transformations of $SU(2)_L \times U(1)_Y$. The free particle Lagrangian is given by

$$\mathcal{L}_0 = i\bar{\psi}\gamma^\mu\partial_\mu\psi \; . \tag{2.1}$$

The generators of the $SU(2)_L$ and $U(1)_Y$ are $\vec{T} = (T_1, T_2, T_3)^T$ and Y. The T_i follow the commuting relations $[T_i, T_j] = i\varepsilon_{ijk}T_k$, where ε_{ijk} is the Levi-Civita symbol. The Lagrangian is then demanded to be invariant under the transformations

$$\psi_L \to e^{ig\vec{\alpha}(x)\vec{T} + ig'\beta(x)Y}\psi_L \; , \tag{2.2}$$

$$\psi_R \to e^{ig'\beta(x)Y}\psi_R \; , \tag{2.3}$$

where g and g' determine the strengths of the electromagnetic and weak interactions.

To ensure the invariance under these transformations (the gauge invariance) additional terms have to be added to the free Lagrangian. These terms involve the spin 1 fields $\vec{W}_\mu = (W_{1\mu}, W_{2\mu}, W_{3\mu})^T$ and B_μ which are linear combinations of the above mentioned gauge bosons that mediate the forces ($W^\pm$, Z^0, γ). The resulting Lagrangian is then

$$\mathcal{L}_{EW} = \sum_{fermions} (i\bar{\psi}_L \gamma^\mu D_\mu \psi_L + i\bar{\psi}_R \gamma^\mu [\partial_\mu - ig' Y_R \frac{1}{2} B_\mu]\psi_R) - \frac{1}{4}\vec{W}_{\mu\nu}\vec{W}^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \ , \qquad (2.4)$$

where $D_\mu = \partial_\mu + ig\vec{W}_\mu \vec{T} + ig' Y \frac{1}{2} B_\mu$ is called the covariant derivative.

Strong Interactions

The quantum field theory that describes the strong interaction is called Quantum Chromodynamics (QCD). Similarly to the electroweak theory the Lagrangian can be derived from a gauge group which is $SU(3)_C$ for QCD.

In QCD the quarks are grouped in color charge triplets. Each quark carries one of the three different color charges which are called 'red', 'green' and 'blue'. Antiquarks carry the corresponding anti-color charges. Leptons do not carry color charge and thus do not participate in strong interactions.

The gauge bosons in QCD are eight gluons. In contrast to the neutral photons in electroweak theory, the gluons carry color charges themselves: one unit of color charge and one unit of anti-color charge. As a consequence gluons couple not only to quarks but also to themselves.

The dynamics of QCD are given by the Lagrangian:

$$\mathcal{L}_{QCD} = \sum_{flavors} i\bar{\psi}_a [\gamma^\mu D_\mu + im]\psi^a - \frac{1}{4} G^n_{\mu\nu} G^{\mu\nu}_n \ , \qquad (2.5)$$

where

$$D_\mu = \partial_\mu - ig_S \frac{\lambda^n}{2} A^n_\mu \qquad (2.6)$$

$$G^n_{\mu\nu} = \partial_\mu A^n_\nu - \partial_\nu A^n_\mu + g_S f^{nml} A^m_\mu A^l_\nu \qquad (2.7)$$

and ψ_a and A^n_μ are the quark and gluon fields, a = 1, 2, 3 and n, m, l = 1, 2, ..., 8 are color charge indices, λ^n and f^{mnl} are the Gell-Mann matrices and f-symbols, m are the quark masses and g_S determines the strength of the interaction.

In comparison to electroweak theory, the coupling strength of QCD is large. The coupling decreases with the momentum transfer of interactions, which has two major consequences. First, only in very high momentum transfer interactions, quarks can be regarded as free particles, the so called asymptotic freedom. Therefore, perturbative calculations in QCD are limited to interactions with high momentum transfer, where weak couplings guarantee a good description in perturbation theory with limited order. Second, free quarks are not observed in nature due to the strong coupling of QCD at large distances, known as confinement. It is an experimental fact that only combinations of quarks that form a color charge singlet state are observed. The possible states are baryons and mesons, $\varepsilon_{ijk} q_i q_j q_k$, $\varepsilon_{ijk} \bar{q}_i \bar{q}_j \bar{q}_k$ and $\bar{q}^i q_i$, where i, j, k are color charges.

The Higgs Mechanism

The Higgs mechanism, which is named after its inventor Peter Higgs, gives rise to particle masses in the SM. The particles described in electroweak theory, e.g. the gauge bosons $\vec{W}_\mu$ and B_μ, are massless. Mass-terms such as $\frac{1}{2}M^2 B_\mu B^\mu$ cannot be added to the Lagrangian since they would produce (unrenormalizable) divergencies in the theory. To solve this problem the Higgs mechanism is introduced.

A new scalar field of the form $\vec{\phi} = \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}$ is introduced and a new term $\mathcal{L}_H$ is added to the Lagrangian:

$$\mathcal{L}_H = |D_\mu \vec{\phi}|^2 - V(\vec{\phi}^\dagger \vec{\phi}) \tag{2.8}$$

$$V(\vec{\phi}^\dagger \vec{\phi}) = \mu^2 \vec{\phi}^\dagger \vec{\phi} + \lambda (\vec{\phi}^\dagger \vec{\phi})^2 \, . \tag{2.9}$$

The term $V(\vec{\phi}^\dagger \vec{\phi})$ is called the Higgs potential which is given by two the parameters μ^2 and λ.

By choosing the parameters $\mu^2 < 0$ and $\lambda > 0$ the vacuum expectation value (or ground state) of the Higgs potential $V(\vec{\phi}^\dagger \vec{\phi})$ is different from zero. If the fields ϕ_i ($i = 1,2,3,4$) are then redefined in such a way that their vacuum expectation value is zero, one gets three massless fields and one massive field, the so called Higgs boson field. This mechanism is called spontaneous symmetry breaking. In principle both parameterizations of the fields ϕ_i ($i = 1,2,3,4$) describe the same physics. However, since we calculate in perturbation theory with limited order, only the parametrization with the vacuum expectation value equal to zero will give correct results.

When spontaneous symmetry breaking is applied, the fields $\vec{W}_\mu$ and B_μ transform into four new massive fields $W^\pm_\mu$, Z^0_μ and the massless photon field A_μ. They are connected by the following relations:

$$\begin{pmatrix} A_\mu \\ Z^0_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W^3_\mu \end{pmatrix} \tag{2.10}$$

$$W^\pm_\mu = \frac{1}{\sqrt{2}} (W^1_\mu \mp i W^2_\mu) \, . \tag{2.11}$$

Here θ_W is the Weinberg angle, which is defined by the coupling constants $g'/g = \tan\theta_W$.

In addition to the four experimentally known $W^\pm$, Z^0, γ bosons, the Higgs mechanism predicts a massive Higgs boson. To date it is the only particle of the SM that has not been observed in experiments. Therefore its search is one of the most active fields in particle physics and it was one of the main motivations to construct the Large Hadron Collider at CERN [4,5].

The masses of fermions are also generated by spontaneous symmetry breaking by adding Yukawa interaction terms of the fermions with the Higgs field. This introduces new massive fields, the so called mass eigenstates of the original fermions (the electroweak eigenstates). For quarks these two states differ. When one sets the weak eigenstates of the up-type quarks equal to their mass eigenstates, the relation between the two states of the down-type quarks are given by

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \, . \tag{2.12}$$

This matrix V is called the Cabbibo-Kobayashi-Maskawa (CKM) matrix.

As a consequence the massive quarks observed in nature behave as mixtures of weak eigenstates. Weak interactions can thus lead to transitions between different quark generations. In principle the same applies to leptons. Since neutrinos are regarded as massless in the SM, mixing of flavors is not allowed for leptons. Extensions to the SM are needed in order to explain the results of neutrino oscillation experiments published in recent years [3].

2.2 Quantum Chromodynamics in Proton-Proton Collisions

The theory of Quantum Chromodynamics (QCD) describes the strong interactions of quarks and gluons. In proton-proton collider experiments this theory is probed by investigating inelastic scattering processes of protons. In this section it is explained how predictions from QCD are calculated in perturbation theory and how these can be related to proton-proton collision experiments.

2.2.1 Perturbative QCD

In the case of high momentum transfer interactions, the coupling constant of the strong interaction $\alpha_s = \frac{g_S^2}{4\pi}$ is small. For these interactions, QCD predictions can be calculated in perturbation theory using the formalism of the Feynman rules which are derived from the Lagrangian of QCD. Cross sections, or in other words rate predictions for interactions, are expressed as a product of a matrix element $|M|^2$, describing the interaction process, and a phase space factor, describing the possible initial states and final states of the interaction. For a $2 \to 2$ process the differential cross section in the center of mass frame is given by

$$\frac{d\hat{\sigma}}{d\Omega} = \frac{1}{64\pi^2 \hat{s}} \frac{p_f}{p_i} |M|^2 \,, \tag{2.13}$$

where Ω is the element of solid angle, $\hat{s}$ is the center of mass energy squared and p_i and p_f are the momenta of the incoming and outgoing particles respectively.

Figure 2.1 summarizes the Feynman diagrams and matrix elements of quark and gluon scattering in leading order perturbation theory. The variables $\hat{s}$, $\hat{t}$, $\hat{u}$ for a $AB \to CD$ processes are given by

$$\hat{s} = (p_A + p_B)^2 \tag{2.14}$$

$$\hat{t} = (p_A - p_C)^2 \tag{2.15}$$

$$\hat{u} = (p_A - p_D)^2 \,, \tag{2.16}$$

where p_A, p_B, p_C, p_D are the four-momenta of the incoming and outgoing particles.

At higher orders of perturbation theory also loop diagrams contribute which in principal lead to divergencies. In order to compute finite cross sections, a renormalization has to be performed.

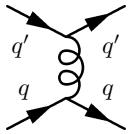
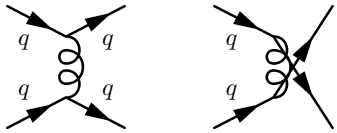
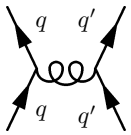
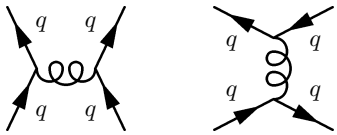
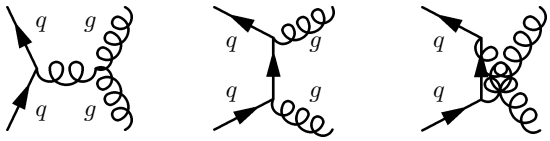
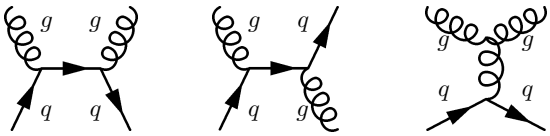
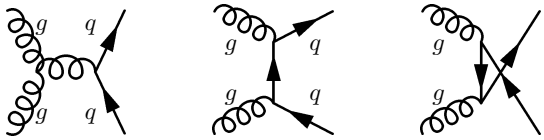
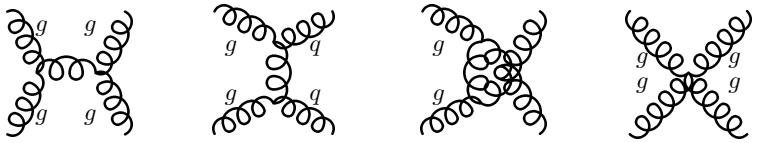
QCD process and Matrix element $\lvert M\rvert^2$	Feynman diagrams in leading order perturbation theory
$qq' \to qq' = q\bar{q}' \to q\bar{q}'$ $\frac{64}{9}\alpha_s^2 \left(\frac{\hat{s}^2+\hat{u}^2}{\hat{t}^2}\right)$	
$qq \to qq$ $\frac{64}{9}\alpha_s^2 \left(\frac{\hat{s}^2+\hat{u}^2}{\hat{t}^2} + \frac{\hat{s}^2+\hat{t}^2}{\hat{u}^2} - \frac{2}{3}\frac{\hat{s}}{\hat{u}\hat{t}}\right)$	
$q\bar{q} \to q'\bar{q}'$ $\frac{64}{9}\alpha_s^2 \left(\frac{\hat{t}^2+\hat{u}^2}{\hat{s}^2}\right)$	
$q\bar{q} \to q\bar{q}$ $\frac{64}{9}\alpha_s^2 \left(\frac{\hat{s}^2+\hat{u}^2}{\hat{t}^2} + \frac{\hat{t}^2+\hat{u}^2}{\hat{s}^2} - \frac{2}{3}\frac{\hat{u}^2}{\hat{s}\hat{t}}\right)$	
$q\bar{q} \to gg$ $\frac{128}{3}\alpha_s^2 \left(\frac{4}{9}\frac{\hat{t}^2+\hat{u}^2}{\hat{t}\hat{u}} - \frac{\hat{u}^2+\hat{t}^2}{\hat{s}^2}\right)$	
$qg \to qg$ $16\alpha_s^2 \left(\frac{\hat{s}^2+\hat{u}^2}{\hat{t}^2} - \frac{4}{9}\frac{\hat{s}^2+\hat{u}^2}{\hat{s}\hat{u}}\right)$	
$gg \to q\bar{q}$ $\frac{8}{3}\alpha_s^2 \left(\frac{1}{3}\frac{\hat{t}^2+\hat{u}^2}{\hat{t}\hat{u}} - \frac{3}{4}\frac{\hat{t}^2+\hat{u}^2}{\hat{s}^2}\right)$	
$gg \to gg$ $72\alpha_s^2 \left(3 + \frac{\hat{t}^2+\hat{u}^2}{\hat{s}^2} + \frac{\hat{s}^2+\hat{u}^2}{\hat{t}^2} + \frac{\hat{s}^2+\hat{t}^2}{\hat{u}^2}\right)$	

Figure 2.1: Feynman diagrams and matrix elements $\lvert M\rvert^2$ for the $2 \to 2$ QCD processes at leading order perturbation theory.

The divergencies are absorbed into a redefinition of the color charge. As a consequence of this, α_s is a running coupling constant as a function of a renormalization scale μ_R, and is given by

$$\alpha_s(\mu_R^2) = \frac{12\pi}{(33-2n_f)\log(\mu_R^2/\Lambda^2)} , \tag{2.17}$$

where n_f is the number of quark flavors. The parameter Λ is not predicted by the theory and was determined from experiment with a value in the range of 0.1 and 0.5 GeV depending on the number of quark flavors taken into account. The energy scale μ_R^2 characterizes the momentum scale of the process to be calculated and is typically set to the transverse momentum of the outgoing particles. From this formula one can see that α_s is small only for processes with a scale $\mu_R^2 \gg 1$ GeV^2 where perturbation theory is applicable.

2.2.2 The Factorization Theorem

In order to relate perturbative QCD predictions to the outcome of proton-proton collision experiments, the cross section for inelastic proton-proton scattering is factorized in the following way

$$d\sigma(a+b\to X) = \sum_{\text{partons } i,j} f_i^a(x_a,\mu_F^2) f_j^b(x_b,\mu_F^2) d\hat{\sigma}(i+j\to X') . \tag{2.18}$$

Here, $d\sigma(a+b\to X)$ is the cross section for two colliding protons a and b, while $d\hat{\sigma}(i+j\to X')$ is the cross section for two interacting partons i and j. The underlying idea comes from the parton model which describes a high energy proton beam as an unseparated beam of quarks, anti-quarks and gluons, called partons. Each parton carries a fraction x of the proton's momentum and energy. The parton momentum distribution function (PDF) $f_i(x,\mu_F^2)$ hereby describes the probability that a parton i carries a fraction x of the proton's momentum. By definition all partons in a proton need to add up to 1:

$$\sum_i \int dx f_i(x,\mu_F^2) = 1 \tag{2.19}$$

The PDFs are defined separately for each quark and anti-quark flavor and for gluons. They are constrained by a set of sum rules which come from the quantum numbers of the proton, e.g.

$$\int_0^1 [u(x,\mu_F^2) - \bar{u}(x,\mu_F^2)]dx = 2 \tag{2.20}$$

$$\int_0^1 [d(x,\mu_F^2) - \bar{d}(x,\mu_F^2)]dx = 1 \tag{2.21}$$

$$\int_0^1 [s(x,\mu_F^2) - \bar{s}(x,\mu_F^2)]dx = 0 , \tag{2.22}$$

where $u(x,\mu_F^2)$, $\bar{u}(x,\mu_F^2)$, $d(x,\mu_F^2)$, $\bar{d}(x,\mu_F^2)$, $s(x,\mu_F^2)$, $\bar{s}(x,\mu_F^2)$ are the PDFs for the up-, down-, strange-quarks and anti-quarks.

In QCD the PDFs depend on a factorization scale μ_F^2 at which the proton is probed. All interactions between quarks and gluons happening at scales below the scale μ_F^2 are absorbed into the PDFs. Therefore at small μ_F^2 the proton is observed basically as a combination of its

three valence quarks *uud*. At higher scales, however, it is dominated by sea quarks and gluons. The scale dependence of the PDFs in QCD is predicted by the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi evolution equation

$$\frac{d}{d\log\mu_F^2} f_i(x_i,\mu_F^2) = \frac{\alpha_s}{2\pi}\int_y^1 \frac{dy}{y} \sum_{j\in q,\bar{q},g} f_j(y,\mu_F^2) P_{ij}(x_i/y) \ , \tag{2.23}$$

where the $P_{ij}(x_i/y_j)$ are the so called splitting functions describing the probability for a parton i of momentum fraction x_i to originate from a parton j with a higher momentum fraction y_j. Using this equation the PDFs measured at a certain energy scale can be extrapolated to another scale, e.g. to the energies observed at the LHC.

The PDFs have been determined from structure function measurements in deep inelastic scattering electron-proton experiments as well as from cross section measurements in proton-anti-proton collisions. Several groups have performed global analyses of all available experimental data to parametrize the PDFs including the corresponding uncertainties [6,7].

3 The Dijet Angular Distributions and Quark Compositeness

3.1 The Physics of Dijet Events

At proton-proton colliders, the dominant inelastic scattering processes are strong interactions between partons. Since perturbative QCD makes predictions for the cross sections of these processes at scales where the strong coupling constant α_s is small, they offer a natural test environment for the theory of QCD.

Perturbative QCD in combination with the factorization theorem predicts the cross section of the production of partons in inelastic scattering of protons. At leading order perturbation theory there are two outgoing partons. Due to the confinement of quarks, the outgoing partons are not observed as free particles, but instead produce a shower of color-neutral particles, so called jets. Therefore one typically observes collision events with two outgoing particle showers, called dijet events.

As the transition from partons to particle jets cannot be fully calculated in perturbation theory, it is instead described by phenomenological models which will be discussed in detail in Chap. 5. To minimize the impact of these models in the comparison of measurements and predictions from perturbative QCD, the outgoing particles are clustered into two (or more) particle jets. The comparison of measurement and theory is done at the level of the four-momenta of jets, which are indeed closely related to those of the original partons.

In the history of particle physics the interest in dijets began with the understanding of the asymptotic freedom of QCD at high momentum transfers [8]. The observation of the first three jet events gave evidence in the existence of the gluon [9,10]. At the same time, these three jet events were interpreted as a demonstration of high-momentum-transfer QCD interactions in the limit of asymptotic freedom. Nowadays the physics of dijets is well established with many measurements performed at LEP [11] and the Tevatron [12].

For the CMS experiment, during the startup phase of the Large Hadron Collider (LHC) at CERN in 2009 and 2010, dijets provided the first well-known physics channel computable at perturbative level. The study of dijets had a large impact on the commissioning of the calorimeter and tracking detectors [13–15]. For the energy calibration of jets, dijets were used as a standard candle [16]. Dijets were the very first channel to search for new physics at the LHC due to its high cross section [17,18]. Furthermore, dijets provide a tool to tune showering algorithms of Monte Carlo event generators [19,20], which allow to predict the QCD background for searches, e.g. the search for the Higgs boson in channels with jets. The first tests of perturbative QCD were already performed using the collision data collected in early 2010 [21].

3.2 The Kinematics of Dijet Events

The scattering of two massless partons in their center-of-mass frame is demonstrated in Fig. 3.1. The kinematics of this process can be characterized by two variables: the center-of-mass energy $\sqrt{\hat{s}}$ and the scattering angle θ^*.

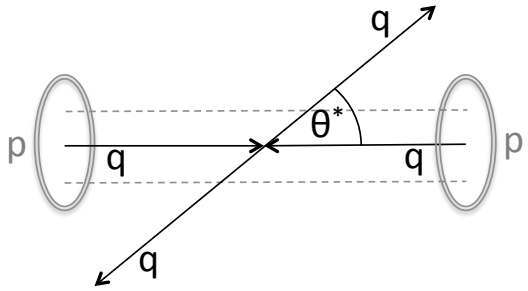


Figure 3.1: Scattering of two partons in their center-of-mass frame.

The center-of-mass energy $\sqrt{\hat{s}}$ is equivalent to the invariant mass M_{jj} of the two partons,

$$\hat{s} = M_{jj}^2, \tag{3.1}$$

which is given by

$$M_{jj}^2 = P_1^\mu P_{2,\mu} = (E_1 + E_2)^2 - (\vec{p}_1 + \vec{p}_2)^2, \tag{3.2}$$

where $P_i^\mu = (E_i, \vec{p}_i)$ are the four-momenta of the two partons. The invariant mass can be related to the transverse momenta p_{T_i} and rapidities y_i of the partons:

$$M_{jj}^2 = 2 p_{T_1} p_{T_2} cosh(y_1 - y_2), \tag{3.3}$$

where the rapidities are defined as

$$y = \frac{1}{2} \frac{E + p_z}{E - p_z}. \tag{3.4}$$

The scattering angle θ^* is related to the rapidities of the two partons:

$$|\cos\theta^*| = \tanh y^*, \tag{3.5}$$

where

$$y^* = \frac{1}{2}|y_1 - y_2|. \tag{3.6}$$

It is also related to the two kinematic variables $\hat{t}$, $\hat{u}$ and the squared center-of-mass energy $\hat{s}$:

$$\hat{t} = -\frac{\hat{s}}{2}(1 - \cos\theta^*), \tag{3.7}$$

$$\hat{u} = -\frac{\hat{s}}{2}(1 + \cos\theta^*). \tag{3.8}$$

In proton-proton collisions, the center-of-mass frame of the two interacting partons is boosted w.r.t. the detector frame as the two partons carry only the fractions x_1 and x_2 of the two proton momenta. Therefore one has to take into account the boost given by

$$y_{\text{boost}} = \frac{1}{2}(y_1 + y_2). \tag{3.9}$$

The relation between the center-of-mass energy s of the colliding protons and the center-of-mass energy $\hat{s}$ of the interacting partons can be expressed as

$$\sqrt{\hat{s}} = \sqrt{s\tau}, \tag{3.10}$$

$$\tau = \hat{s}/s = x_1 x_2. \tag{3.11}$$

Note that in proton-proton collisions some of the above relations are only approximate for two reasons. First, in higher orders of perturbation theory, scattering processes which produce more than two partons occur, and second, instead of massless partons, massive particle jets are measured. The relation between the center-of-mass property θ^* and the measured rapidity difference y^* (Equ. 3.5) is only approximate, as well as the relation between the invariant mass and the transverse momenta (Equ. 3.3). Also the center-of-mass energy is not equal to the invariant mass of the two outgoing jets anymore (Equ. 3.1).

3.3 The Dijet Angular Distributions in QCD

In parton-parton scattering, the angular distribution $d\hat{\sigma}/d\cos\theta^*$ of the outgoing partons is sensitive to the spin of the exchanged particle. For a spin 1 particle exchange the angular distribution is equal to the Rutherford scattering formula

$$\frac{d\hat{\sigma}}{d\cos\theta^*} \propto \frac{1}{(1-\cos\theta^*)^2}\ . \tag{3.12}$$

In QCD the exchanged particles are quarks (spin 1/2) and gluons (spin 1) and therefore a noticeable but small deviation from Rutherford scattering can be observed.

Instead of θ^*, the variable

$$\chi_{\text{dijet}} = e^{2y^\star} = e^{|y_1 - y_2|}\ , \tag{3.13}$$

related to θ^* in the case of scattering of two mass-less partons by

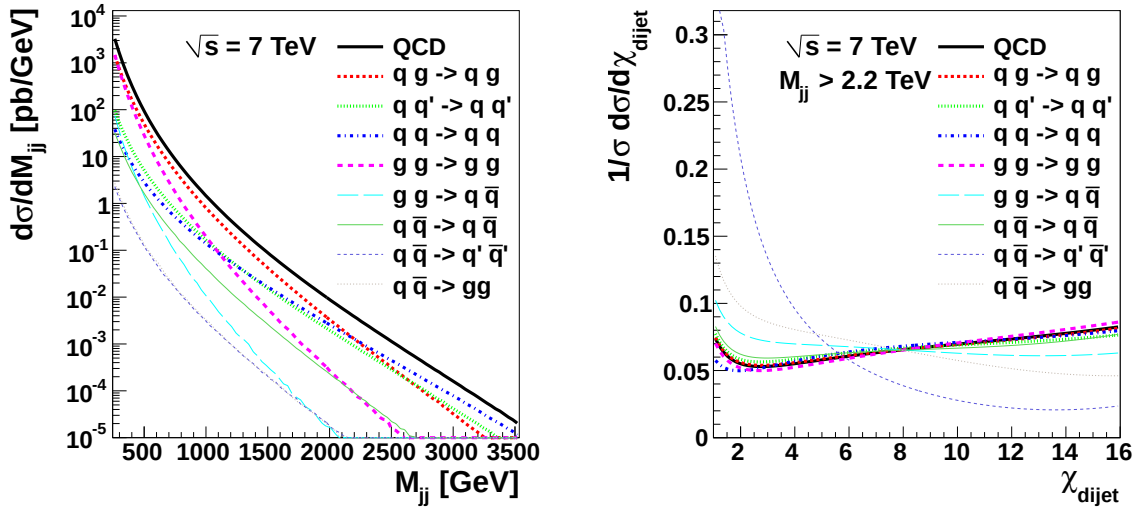
$$\chi_{\text{dijet}} = \frac{1+|\cos\theta^*|}{1-|\cos\theta^*|}\ , \tag{3.14}$$

is typically used to characterize the angular distribution of dijet events. This is motivated by the fact that

$$\frac{d\sigma}{d\chi_{\text{dijet}}} \propto (1-\cos\theta^*)^2 \frac{d\hat{\sigma}}{d\cos\theta^*} \tag{3.15}$$

is a flat distribution for Rutherford scattering.

In this analysis the normalized angular distributions $\frac{1}{\sigma}\frac{d\sigma}{d\chi_{\text{dijet}}}$ are measured in bins of M_{jj}. Due to the normalization, the most prominent experimental uncertainties, such as the integrated



(a) (b)

Figure 3.2: QCD prediction for (a) the dijet mass cross section and (b) the dijet angular distributions at the LHC, broken up into its subprocesses.

luminosity and the jet energy scale uncertainties, are strongly reduced and therefore allow a precise measurement. In the following $\frac{1}{\sigma}\frac{d\sigma}{d\chi_{\text{dijet}}}$ in bins of M_{jj} will be referred to as the dijet angular distributions.

Figure 3.2 (a) shows the differential cross section of various QCD scattering processes at the LHC as a function of M_{jj}. At low invariant masses the gluon-gluon and quark-gluon processes dominate. At high invariant masses quark-quark processes dominate. The relative contributions of the processes are explained by a convolution of the matrix elements and the parton distribution functions. The dominance of the gg processes at low M_{jj}, for example, comes from the large gluon distribution at small values of x combined with the large $gg \to gg$ matrix element (compare Fig. 2.1).

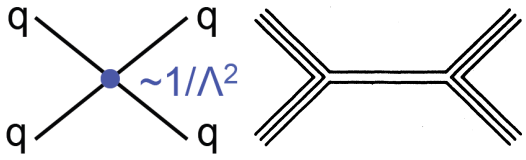
Figure 3.2 (b) shows predictions of the dijet angular distributions for different parton-scattering processes. All dominant processes are similar to a flat distribution as for Rutherford scattering. Only the negligible contribution of the $q\bar{q} \to q'\bar{q}'$ process of quark-antiquark annihilation which is driven by s-channel gluon exchange clearly deviates. The reason is that for s-channel scattering, the scattering angle θ^* is isotropically distributed and therefore the dijet angular distributions peak at low χ_{dijet}. From this figure it is clear that the shape of the dijet angular distributions only weakly depends on the relative contributions of the main processes ($qg \to qg$, $qq' \to qq'$, $qq \to qq$, $gg \to gg$) to the total cross section. Therefore the angular distributions are almost insensitive to the parton distribution functions, which predict the change of the relative contributions as a function of M_{jj}. The QCD prediction for the dijet angular distributions, thus, has a minimal uncertainty coming from the parton distribution functions and is well suited for a test of perturbative QCD.

The dijet angular distributions and related quantities have been measured by the UA1 [22], DØ [23,24], CDF [25] and ATLAS [26] collaborations, yielding good agreement with the QCD

predictions. Beyond searches for new physics phenomena, a measurement of the dijet angular distributions at the LHC is of particular interest as it extends previous measurements to higher $\sqrt{\hat{s}}$ and probes the dominant gluon contributions at low invariant masses.

3.4 Signatures of Quark Compositeness

The idea of substructure of quarks is motivated by the history of particle physics in which the frequent discovery of new sub-particles of molecules, atoms and nucleons have broken down the picture of particle physics to a small number of fundamental particles. A variety of models on sub-particles of quarks called preons bound together by a new strong interaction has emerged [27]. Compositeness of quarks may explain e.g. the number of quark generations as well as the quark masses and charges.



(a) (b)

Figure 3.3: Symbolic representation of (a) a contact interaction between quarks with a coupling strength Λ and (b) a possible underlying preon exchange process between quarks. The lines represent quarks in (a) and preons in (b).

These models have in common that at the scale of the constituents binding energy Λ, new interactions among quarks should appear. At energies below this scale the dominant new interaction between quarks would most commonly be a contact interaction between same flavor quarks as depicted in Fig. 3.3 (a). In the picture of preons, this interaction corresponds to an elastic interaction between composite quarks, where preon bound states with masses of order Λ are exchanged as depicted in Fig. 3.3 (b).

The Langrangian for a contact interaction between left-handed quarks is given by

$$L_{qq} = \pm\frac{g^2}{2\Lambda^2}(\bar{q}_L\gamma^\mu q_L)(\bar{q}_L\gamma_\mu q_L). \tag{3.16}$$

Note that the given Lagrangian is restricted to left-handed quarks while contact interactions may also be introduced for right-handed quarks. The sign of the term determines whether there is constructive (”-”) or destructive (”+”) interference with the corresponding QCD terms.

In principle the sub-structure of quarks can also be characterized by form factors [27]. This analysis, however, concentrates on the most common parametrization of quark compositeness signatures by the contact interaction model.

A contact interaction between quarks is observable in dijet production at high invariant masses where quark-quark interactions are least suppressed by the parton distribution functions. In principle the contact interaction signatures can be observed in any dijet observable,

such as the dijet mass, whose measurement is, however, subject to rather large experimental uncertainties. Contact interactions are best probed by measuring the dijet angular distributions which can be precisely measured with reduced experimental uncertainties. This method is inspired by Rutherford who discovered a point-like constituent, the nucleus, within the atom by measuring the scattering angle of α-particles on gold atoms.

Figure 3.4 shows the prediction for the dijet angular distributions at the LHC with dijet invariant masses above 2.2 TeV for QCD and for a model with a contact interaction between quarks at a scale of $\Lambda^{\pm} = 5$ TeV with constructive (Λ^{-}) and destructive(Λ^{+}) interference with the QCD terms. The contact interaction produces a peak at low χ_{dijet}, the same behavior as for the s-channel QCD diagrams. One can see that the contact interaction signal is clearly visible already at invariant masses well below the scale $\Lambda^{\pm}$. The Λ^{-} prediction is larger than the Λ^{+} prediction and distinguishable from QCD already in lower mass bins due to its constructive interference with the QCD terms.

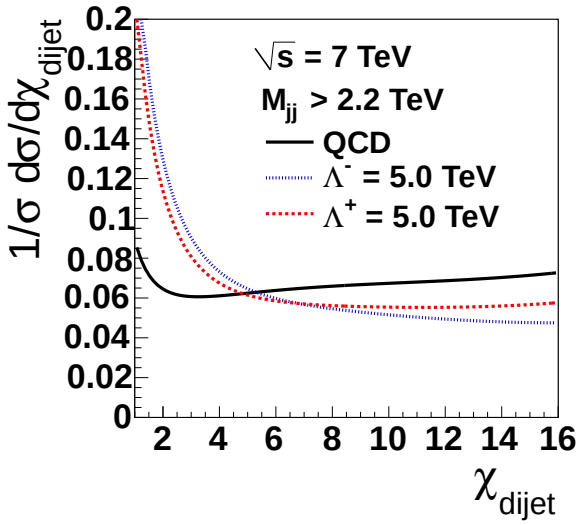


Figure 3.4: The dijet angular distributions for QCD and for a model with an additional contact interaction between quarks with constructive (Λ^{-}) and destructive(Λ^{+}) interference with the QCD terms.

Previous measurements of dijet events have set lower limits on the contact interaction scale [18, 22–26]. The most stringent limits before the start of the LHC were set by the DØ collaboration with $\Lambda^{\pm} > 3$ TeV [23].

4 The LHC Accelerator and the CMS Experiment

4.1 The Large Hadron Collider

The Large Hadron Collider (LHC) at the Conseil Européen pour la Recherche Nuclèaire (CERN) is as its name suggests the largest hadron collider ever constructed [5,28]. It is built in the tunnel of the former Large Electron-Positron Collider (LEP) with a circumference of 27 km.

The two main characteristics of a hadron collider are its beam energy and its luminosity. The beam energy defines the center-of-mass energy for the hadron collisions, which for two beams with equal energy is given by

$$\sqrt{s} = 2E_{\text{Beam}} \ . \tag{4.1}$$

The luminosity $\mathcal{L}$ characterizes the interaction rate. For a process with a cross section σ, the interaction rate is given by

$$\frac{dN_{ev}}{dt} = \sigma\mathcal{L} \ . \tag{4.2}$$

The number of interaction events in a period of running time of the collider is

$$N_{ev} = \sigma \int \mathcal{L}dt = \sigma L \ . \tag{4.3}$$

where L is called the integrated luminosity.

Designed to search for rare interactions at the TeV scale, the LHC supersedes previous high energy hadron colliders by a factor of 7 in its design center-of-mass energy of $\sqrt{s} = 14$ TeV for proton-proton collisions ($\sqrt{s} = 5.5$ TeV for Pb-Pb collisions) and by a factor of 100 in its design luminosity of $\mathcal{L} = 10^{34}$ cm^{-2}s^{-1} for proton-proton collisions ($\mathcal{L} = 10^{27}$ cm^{-2}s^{-1} for Pb-Pb collisions).

The colliding particles are accelerated in two separated beam pipes that have crossing points for collisions at four detector sites. The particle beams are kept on track by high-field superconducting NbTi dipole magnets which can produce magnetic fields of up to 8.3 T. They need to be cooled down to 1.9 K to reach their superconducting phase. For this purpose both beam pipes are included in a shared cryostatic system.

Figure 4.1 shows the accelerator and detector facilities at CERN. To reach energies of 7 TeV, the particles run through a chain of accelerators. Protons from a hydrogen source are first accelerated by the Linac to energies of 50 MeV. In the Booster, PS and SPS their energy is then increased sequentially to 1.4 GeV, 25 GeV and 450 GeV. Finally they are accelerated up

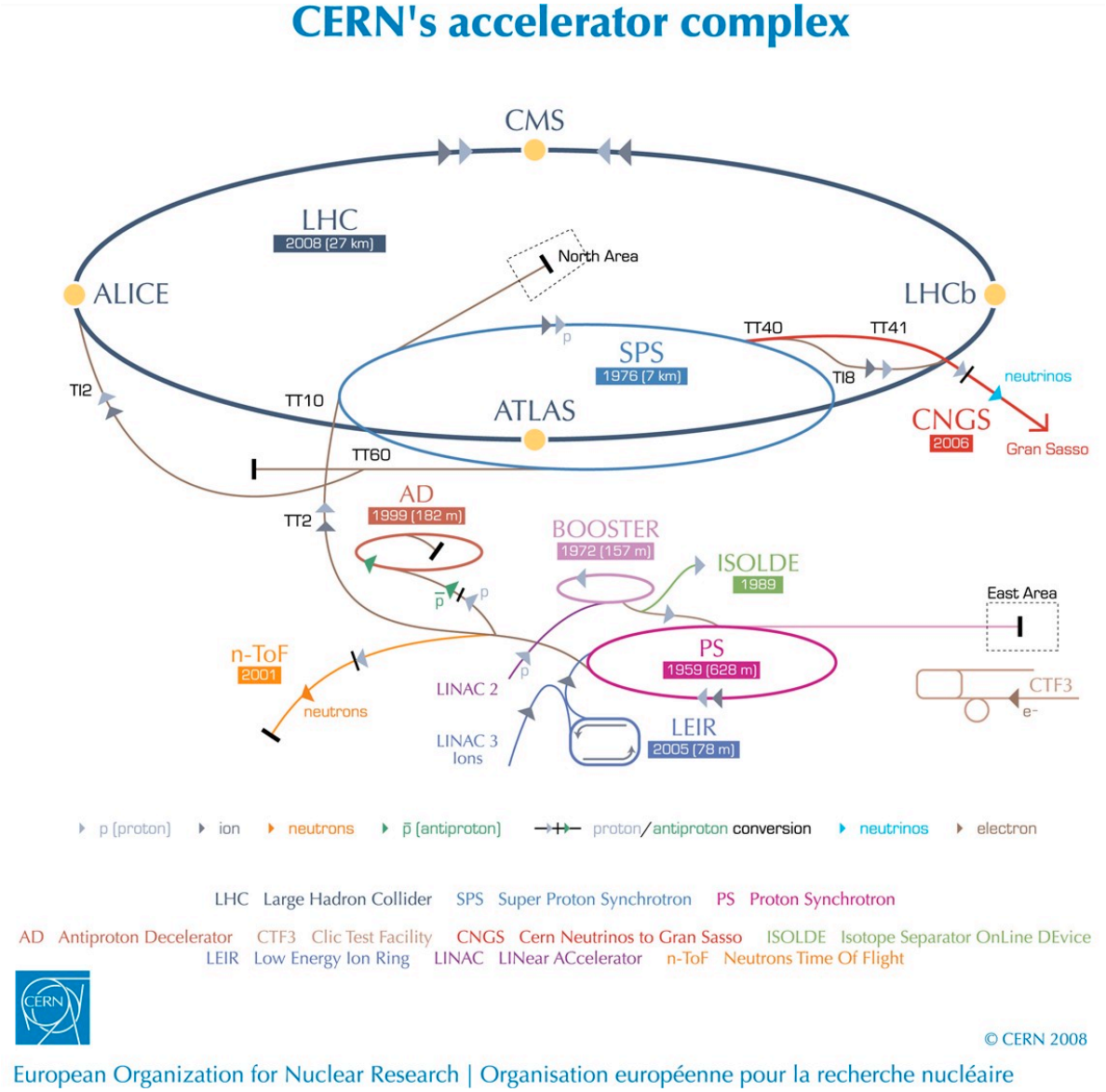


Figure 4.1: The accelerators and detectors at CERN [29].

to 7 TeV in the LHC. After acceleration the protons move through the LHC in 2808 bunches of 10^{11} protons with an average length of 7.5 cm and a bunch spacing of 25 ns. This corresponds to the design luminosity of 10^{34} cm^{-2}s^{-1}.

In 2010 the LHC operated at $\sqrt{s} = 7$ TeV for proton-proton collisions with a continuous ramp up in luminosity over the year. The peak luminosity was above $\mathcal{L} = 10^{32}$ cm^{-2}s^{-1} with protons in up to 368 bunches and a bunch spacing of 150 ns. The development of the integrated luminosity over the year 2010 is shown in Fig. 4.2, with a total of 47 pb^{-1}. The integrated luminosity has been measured with an uncertainty of 4% [31,32].

Four large experiments are installed at the points of the accelerator ring where the protons are brought to collision as shown in Fig. 4.1. The ATLAS (A Toroidal LHC ApparatuS) and CMS (Compact Muon Solenoid) experiments are general-purpose detectors which cover a wide range of physics analyses. LHC-b (Large Hadron Collider beauty Experiment) aims to analyze phenomena in decays of B-mesons with high precision and high statistics. ALICE (A Large Ion

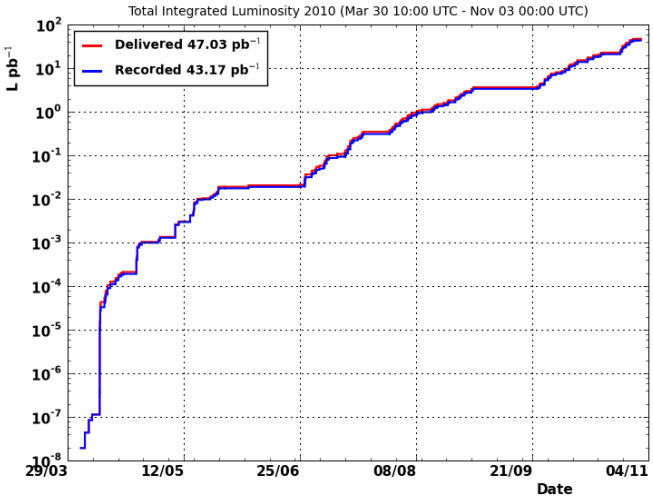


Figure 4.2: Integrated luminosity delivered by the LHC in 2010 [30].

Collider Experiment) focuses on the investigation of heavy ion collisions.

4.2 The CMS Detector

The CMS detector is designed for the search of the Higgs boson and new physics at the high energy frontier at the LHC [5,33,34]. To achieve these goals, it is required to detect and identify a variety of particle types produced in proton-proton collisions, including muons, electrons, photons as well as heavy and light quark jets, with good precision. These requriements were achieved by the design presented in the following.

Figure 4.3 shows a cutaway view of the CMS detector and its compounds. The CMS detector has a diameter of 15 m, is 21.6 m long and weighs about 12500 t. It surrounds the collision point in its center to detect all particles produced in high momentum transfer collisions, including an estimate of the transverse momentum carried away by invisible particles through the sum of all visible particles, called missing transverse energy. Several sub-detectors are organized in a barrel part and two forward parts. From the center outwards the CMS detector consists of the following compounds.

- The pixel detector and the silicon strip tracker together form the tracker which is the key component of CMS to measure the momenta of charged particles and identify vertices.
- The electromagnetic calorimeter, the pre-shower detector, the hadronic calorimeter and the very forward calorimeter provide information on the energies and directions of all charged and neutral particles.
- The NbTi superconducting solenoid generates a magnetic field of up to 4 T in the central part of the detector to enable determination of particle momenta in the tracker. The magnet is surrounded by several return yokes made of iron that return the magnetic field of the solenoid. They shape the magnetic field outside of the solenoid to enable determination of particle momenta also in the muon system.

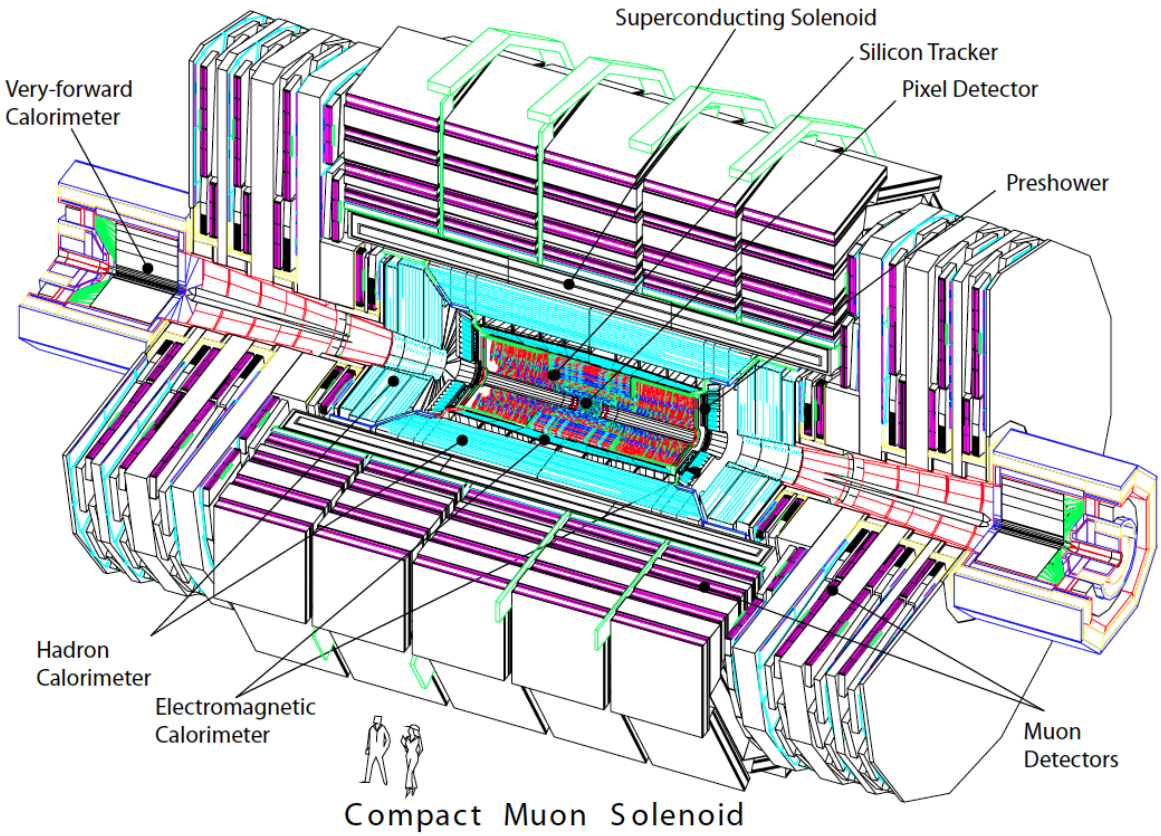


Figure 4.3: Cutaway view of the CMS detector and its compounds [5].

- The detectors of the muon system are installed in the return yokes of the magnet and provide identification of muons and measurement of their momenta.

For the description of the CMS detector the following coordinate system is used. The origin is set to the center of the detector. The z-axis is chosen parallel to the beam axis. The x-axis lies in the horizontal plane. Coordinates are then typically given by the z-component, the azimuthal angle ϕ measured in the x-y-plane from the x-axis and the pseudo-rapidity $\eta = -\ln\tan(\theta/2)$ where θ is the polar angle measured from the z-axis.

In the following sections the three main components of the CMS detector will be described. This is followed by a section on the triggering system of the CMS experiment.

4.2.1 The Tracker

The tracker measures tracks of charged particles with high precision and provides vertex information for both the identification of the primary vertex of a collision and the identification of jets originating from heavy quarks which contain B-hadrons decaying at secondary vertices [5,33]. For this purpose it is built in the region closest to the beam pipe, from a radius of 4 cm to 115 cm. In this region it has to deal with high particle fluxes. Therefore, the components of the tracker need to be radiation hard and provide a measurement with high granularity in order to distinguish neighboring particle tracks. The technology chosen for this purpose at CMS are

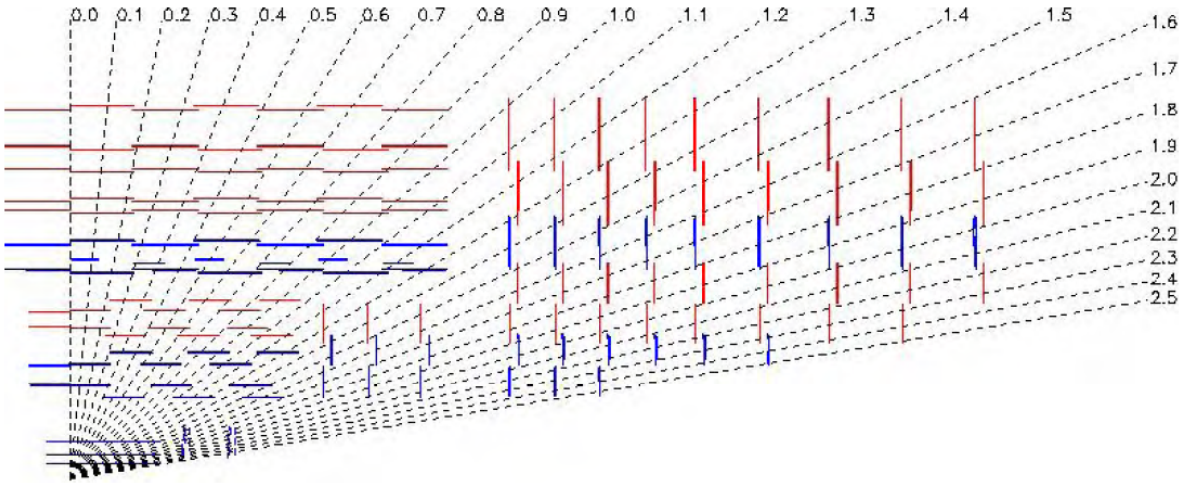


Figure 4.4: Layout of the tracker of the CMS experiment (1/4 of the z view) [5].

silicon detectors, cooled down to -10 °C to ensure radiation hardness. With more than 200 m^2 of silicon it is the largest detector of its kind.

Figure 4.4 shows the layout of the tracker. It consists of two main compounds that use different detection technology, the pixel detector and the silicon strip tracker.

The inner part of the tracker up to a radius of 11 cm, the pixel detector, is built from three layers of silicon pixel detectors in the barrel and two layers in the end-caps. The pixels have a size of 100×150 μm^2 and give three-dimensional information of particle tracks. The pixel detector is the essential part of the detector for the identification of vertices.

The outer part of the tracker, the silicon strip tracker, consists of several layers of silicon strip detectors with strip widths of 80 μm (20 cm$< r <$ 55 cm) to 180 μm ($r >$ 55 cm). The layers are arranged into two barrel layers and end-caps with different radii to provide a maximum number of detector layers over the whole η-range of the tracker. The strips supply two-dimensional information on the particle tracks. Three-dimensional information is derived by combining barrel and end-cap information. Additionally, stereo layers of silicon strips which are twisted by an angle of 100 mrad also supply three-dimensional information.

Overall, the tracker covers the region of $|\eta| < 2.5$ and has 5×10^7 readout channels with an occupancy of less than 1% at highest luminosity. An analysis of the minimum bias data at $\sqrt{s} = 7$ TeV collected in 2010 yields resolutions in x(y) and z close to 25 μm and 20 μm for primary vertexes reconstructed from more than 30 tracks [35]. Using muon pairs from J/Ψ decays, the momentum resolutions for muon tracks at the J/Ψ invariant mass of 3.096 GeV was found to be within 0.6% and 3% increasing with pseudo-rapidity and transverse momentum, in perfect agreement with the simulation [36].

4.2.2 The Calorimeter

The calorimeter measures the energies and directions of all neutral and charged particles traversing the detector, with the exception of muons and neutrinos [5,33]. It consists of two parts, the

electromagnetic calorimeter (ECAL) and the hadronic calorimeter (HCAL). The ECAL aims at the measurement of photons and charged particles such as electrons. Both calorimeters are used for measuring hadronic particle jets and determine the missing transverse energy. Figure 4.5 shows the layout of the two calorimeters. In contrast to other multi-purpose detectors, the calorimeter of the CMS experiment is located almost fully inside the magnet coil. This design minimizes energy loss of particles before entering the calorimeter.

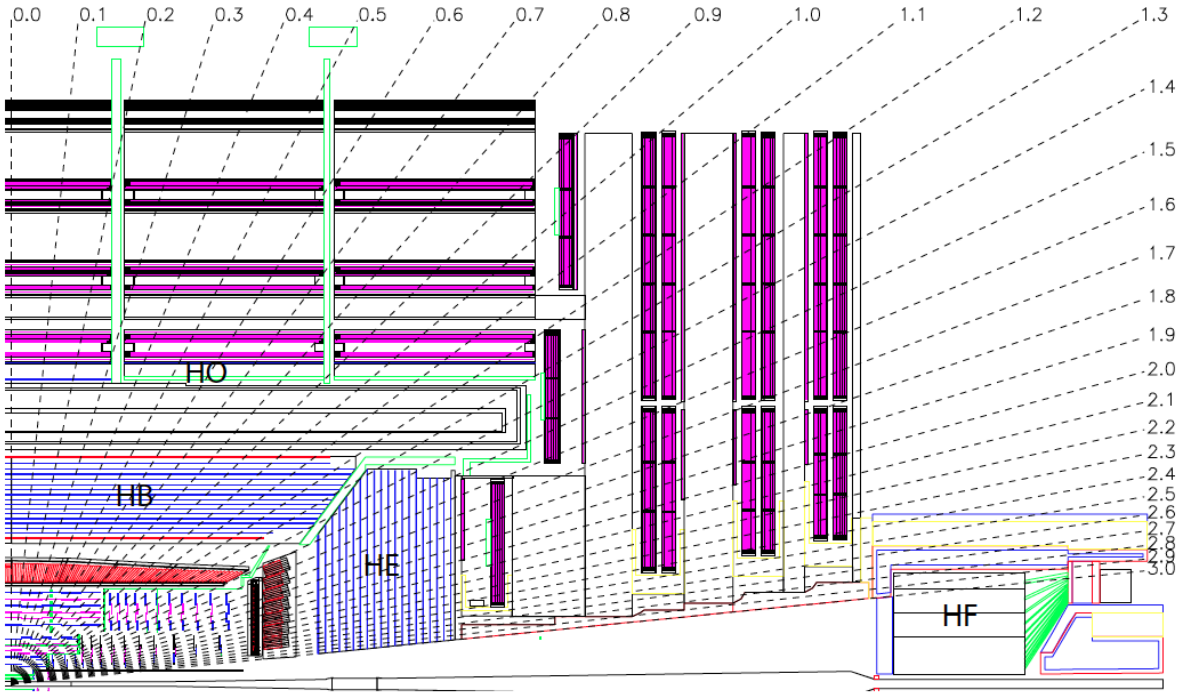


Figure 4.5: Layout of the calorimeters of the CMS experiment (1/4 of the z view) [33].

The ECAL consists of 61200 scintillating $PbWO_4$ crystals that provide excellent energy resolution, radiation hardness and scintillation speed. In the barrel region they are read out by silicon avalanche photodiodes (APDs) which can operate in high magnetic fields. In the end-caps vacuum photo-triodes (VPTs) are used which are more radiation hard. The length of the crystals is 23 cm (22 cm) in the barrel (end-caps) which corresponds to 26 (25) radiation lengths. In the forward region particle identification is enhanced by an additional pre-shower detector in front of the calorimeter crystals. Overall, the ECAL covers an η-range of $|\eta| < 3.0$. The energy resolution is designed to be

$$\frac{\sigma_E}{E} = \frac{2.7\%}{\sqrt{E/\mathrm{GeV}}} \oplus \frac{20\%}{E/\mathrm{GeV}} \oplus 0.55\% \tag{4.4}$$

in the barrel and

$$\frac{\sigma_E}{E} = \frac{5.7\%}{\sqrt{E/\mathrm{GeV}}} \oplus \frac{25\%}{E/\mathrm{GeV}} \oplus 0.55\% \tag{4.5}$$

in the end-caps. The energy response of the ECAL was measured using the first 2010 collision data and found to agree with simulation within 3% [37].

The HCAL consists of copper absorbers interleaved with plastic scintillators which are read out with wavelength-shifting fibers (WLS). In the barrel region the HCAL is extended by a

tail catcher calorimeter outside the magnet coil to expand the depth of the HCAL to about 10 interaction lengths for the measurement of highly energetic jets. In the forward region the HCAL is extended by an iron/quartz-fiber calorimeter, the very forward calorimeter, to allow detection up to $|\eta| < 5.0$, thus ensuring a comprehensive measurement of the missing transverse energy. The energy resolution is designed to be

$$\frac{\sigma_E}{E} = \frac{70\%}{\sqrt{E/\mathrm{GeV}}} \oplus 9.5\% \tag{4.6}$$

in the barrel and

$$\frac{\sigma_E}{E} = \frac{127\%}{\sqrt{E/\mathrm{GeV}}} \oplus 9\% \tag{4.7}$$

in the end-caps. The energy response of the HCAL was measured using the first 2010 collision data and found to agree with simulation within 8% [38].

The calorimeter is the key detector component for measuring particle jets up to transverse momenta of several TeV, which are the key ingredient to this analysis. Therefore a detailed discussion of the commissioning and performance of jets will be given in Chap. 7

4.2.3 The Muon System

The muon system identifies muons and measures their momenta and charges [5,33]. It is located outside the magnet. Since muons are the only charged particles that penetrate the calorimeter and the magnet, tracks in the muon system can be identified as muons. Figure 4.6 shows the layout of the muon system.

The muon system is integrated within the return yoke of the magnet which provides a magnetic field for momentum measurements. In the barrel region drift tubes (DTs) are installed. In the forward region cathode strip chambers (CSCs) are used instead because of their good performance in strong magnetic fields and at high rates. Additionally, both the barrel and the forward region are equipped with resistive plate chambers (RPCs) which provide fast timing and are used for trigger purposes.

Overall, the muon system covers the region of $|\eta| < 2.4$ and has about one million readout channels. Best momentum resolution is achieved by combining tracker and muon system information. The global momentum resolution is strongly η-dependent. For $p_T < 100$ GeV (1 TeV) it is designed to be 6%-20% (15%-35%). At low p_T, good momentum resolution is achieved by the tracker, whereas at high p_T the muon system sets in and improves the global resolution. Details on the commissioning of muon identification and momentum measurement can be found in Ref. [36,39].

4.2.4 The Trigger System

At the design luminosity of 10^{34} cm^{-2}s^{-1}, on average about 20 collisions per bunch crossing every 25 ns are expected [5,33]. With $\sim 10^8$ readout channels which corresponds to about 1 MB

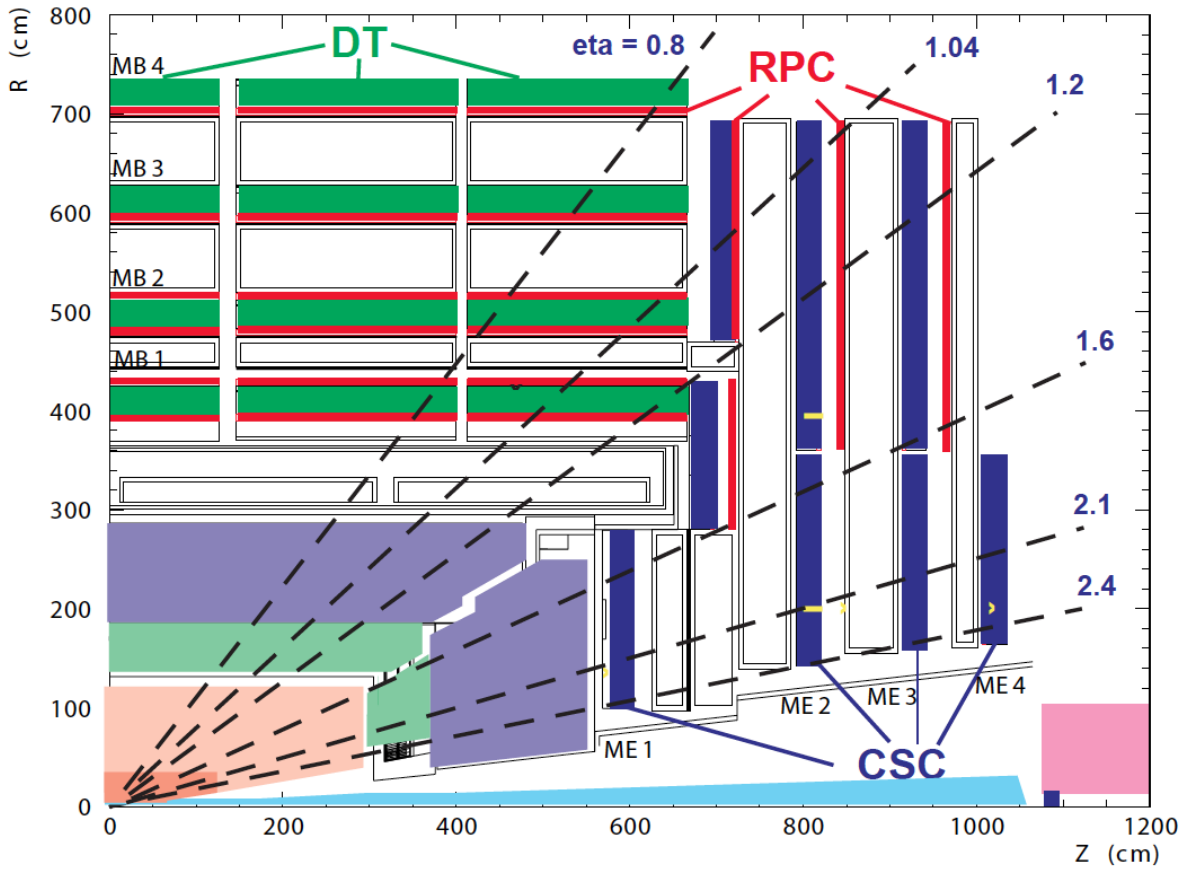


Figure 4.6: Layout of the muon system of the CMS experiment (1/4 of the z view) [5].

of data per event, the CMS detector is expected to produce about 100 TB of data per second. The data storage system, however, can only handle of the order of $\sim$ 100 MB/s. To reduce the data rate to this level, a two-level trigger system is used at the CMS experiment to filter out the events of interest for physics.

The Level-1 trigger (L1) consists of custom hardware processors. It uses information from the muon system and the calorimeter to construct simple 'trigger primitive' objects such as muons, electrons/photons, jets and the sums of E_T and E_T^{miss}. Simple cuts are applied to these objects to limit the event rate to the order of $\sim$ 100 kHz.

The High-Level Trigger (HLT) is software-based and runs on a farm of commercial processors. The HLT uses objects with finer granularity and higher precision than the Level-1 trigger. Also track information is exploited using simple algorithms. The final decision of the HLT can even be based on a full event reconstruction. Thereby, the event rate is reduced to the order of $\sim$ 100 Hz.

The decisions of the HLT trigger are made according to the so-called trigger menu which consists of several paths which have different requirements to an event. Events that pass at least one trigger path are passed on to the data storage system.

During the 2010 run, the trigger menu was continuously adapted to the increasing luminosity delivered by the LHC. With increasing luminosity the requirement conditions, e.g. on particle transverse momentum thresholds or particle isolation, were tightened in order to be

able to record all the events of physics processes of interest for CMS. Lower threshold triggers remained active but their event rate was reduced by rejecting a fraction of events, called pre-scaling. A trigger pre-scaled with a factor of 10, for example, evaluates every 10^{th} event while rejecting the rest.

5 Simulation of Collision Events

Scattering processes of partons in proton-proton collisions can be calculated in perturbative QCD making use of the factorization theorem as described in Sec. 2.2. In order to compare these predictions with data from proton-proton collisions, one has to simulate the full proton-proton interaction process including the formation of particle jets from the outgoing partons, which cannot be fully described in perturbative QCD. These processes are modeled in Monte Carlo (MC) event generators. Further, one has to understand the exact response of the detector to the outgoing particles produced in proton-proton collisions. This is done by passing the outgoing particles predicted by MC event generators through a full simulation of the detector. In this chapter, MC event generators are described in detail, followed by a brief description of the CMS detector simulation.

5.1 Monte Carlo Event Generators

Monte Carlo event generators model the outcome of proton-proton collisions by combining perturbative predictions from QCD with phenomenological models [40–42]. Collision events are randomly generated according to the probability for each possible interaction process. Summing over these MC events, cross sections for processes with all kinds of requirements on the events, e.g. transverse momentum requirements on jets, can be calculated.

The proton-proton collision process is factorized into different steps decoupled by the timescale on which they happen as depicted in Fig. 5.1. First, the parton-parton interaction is calculated in perturbation theory as described in Sec. 2.2, typically at leading order, drawing partons according to the parton distribution functions from the protons. Second, emissions of additional partons from the in- and outgoing partons are simulated making use of perturbative QCD. Third, the so called beam remnants, the remaining parts of the colliding protons, produce additional interactions which are modeled phenomenologically. Forth, the outgoing partons form hadrons in the so called hadronization step which is also modeled phenomenologically. Finally, unstable outgoing hadrons decay.

Among the most popular MC generators for the LHC are PYTHIA [41] and HERWIG++ [42]. The following description follows closely the implementations in PYTHIA, but is in general similar for other MC generators.

Parton showering

Perturbative calculations in QCD are limited to processes in which the coupling constant $\alpha_s(\mu)$ is small, and by the complexity of higher order calculations preventing their evaluation. Parton

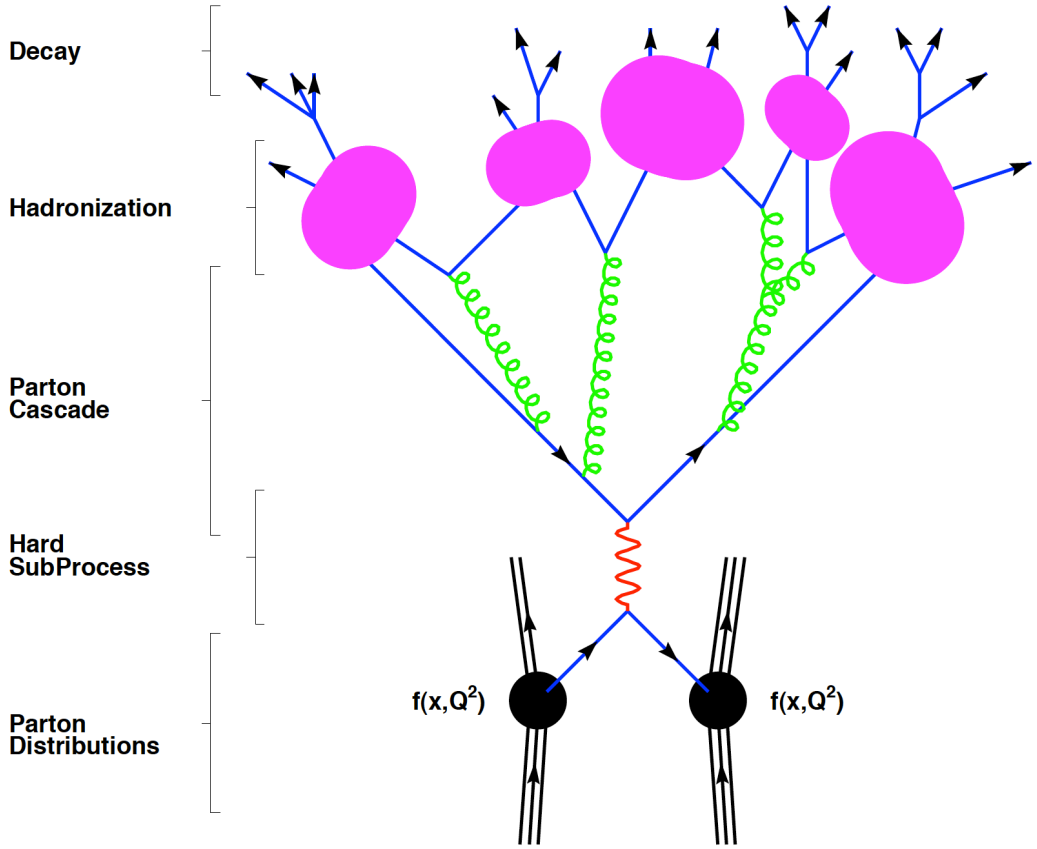


Figure 5.1: Hadron collision process as implemented in MC event generators [43].

showering algorithms extend the fixed order calculations beyond these limiting factors by calculating emissions of additional partons from the in- and outgoing partons of the main interaction. This approach in principle takes into account emissions of an unlimited number of partons, but, as opposed to full higher order calculations, does not take into account loop diagrams.

Parton showering algorithms make use of the splitting functions described in Sec. 2.2.2, which give the probability of a parton j to split into two partons. Using these probabilities, a cascade of splittings is constructed, where at each splittings the four-momentum of the ingoing parton is shared among the outgoing partons. Initial state and final state radiation are treated separately. While initial state radiation is typically propagated backward in time from the main interaction, final state radiation is propagated forward in time. The evolution is characterized by an ordering parameter describing the decreasing energy scale of the splittings. Constrained by color coherence effects in QCD, parton showering algorithms take into account, an angular ordering of the shower which means a decrease of the emission angle throughout the evolution. This feature leads to the fact that jets from high-momentum partons are restricted within an angular region. The cascade of splittings is stopped once the energy scale reaches below a typical value of 1 GeV, when the coupling constant $\alpha_s(\mu)$ becomes large.

Multiple Parton Interactions

In a proton-proton collision, apart from the main interaction of two partons, a lot of additional activity takes place, which is called the underlying event. This additional activity in the event comes mainly from interactions of the remaining partons in the two proton-remnants. These interactions are typically much weaker than the main interaction. As a consequence, it is usually not possible to describe them in the framework of perturbative QCD. Instead a number of phenomenological models for these processes of multiple parton interaction have been developed.

A multiple parton interaction model is built on assumptions on the initial state for these interactions, the structure of the proton-remnants, and also the interaction processes. A typical model as implemented in PYTHIA assumes that the additional interactions are initiated independently from the main interaction according to Poisson statistics and behave the same as the $2 \rightarrow 2$ perturbative QCD interactions, but extrapolated to low transverse momenta, where in principle the predictions break down. At a tunable transverse momentum scale smaller than 2-3 GeV a regularization for the divergent cross section is introduced. In more advanced models, the multiple parton interactions are then showered together with the partons of the main interaction and also the connection of color charge between the interactions is taken into account.

Hadronization

The formation of color-neutral hadrons from the colored partons, happening at an energy scale around 1 GeV, is treated in phenomenological models. A popular model is the Lund string fragmentation as implemented in the PYTHIA event generator. It starts from the assumption that the process of building quark pairs and the formation of actual mesons from those is basically decoupled. In a first step, quark pairs are formed in an iterative procedure. The partons are assumed to be connected via strings. When two partons are separated and their string energy increases, quark anti-quark pairs are pulled out of the vacuum assuming a tunneling probability. The tunneling probability depends on the quark mass and therefore defines the distribution of flavors in the final state. Two of the quarks form a pair, and the other two, still connected by a string, continue the iterative procedure. The momentum is distributed between the quarks according to a fragmentation function, which thereby defines the momentum distribution between different hadrons in the final state. In a second step, the quark pairs are transformed into mesons and baryons according to spin counting arguments and the measured contributions of meson and baryons. Finally, unstable particles are decayed using their measured branching fractions.

Generators and Tunes

The field of application for MC event generators in particle physics is twofold. On the one hand the implemented models make predictions and their physics content can be probed with collision data. On the other hand MC event generators can be used to study the detector performance. For this purpose it is important to select a suited model and to tune its parameters to describe the data in an optimal way. Therefore, a variety of models and parameter tunes for each

step of the MC event generation are probed with collision data. In the following the differences between the models and parameter tunes used in this study are briefly described.

In PYTHIA 6.422 there are two different showering algorithms available, which differ primarily in their choice of the ordering parameter. While the old model implements virtuality-ordered showers [41], the new model uses transverse-momentum-ordered showers [44]. Another showering algorithm is available in HERWIG++ 2.4.2, which is based on the coherent branching algorithm [45].

For the multiple parton interaction modeling two different parameter tunes for the PYTHIA event generator are used in this study. While the D6T [46] tune was created to primarily describe the underlying event data of the CDF experiment at Tevatron, the Z2 [46,47] tune was specifically designed to describe the available underlying event CMS data. HERWIG++ implements a different multiple parton interaction model based on the model from JIMMY [48].

For the hadronization process, two different models are used: the Lund string model [49] implemented in PYTHIA as described in Sec. 5.1, and the clustering model [50] implemented in HERWIG++.

5.2 The CMS Detector Simulation

For a detailed understanding on how interactions in proton-proton collisions at the LHC are observed by the CMS detector, a dedicated simulation of the whole detector and the subatomic interactions is needed [5]. Both the propagation of particles through the detector material as well as the response of the active detector components and their digital output need to be simulated. The input to the detector simulation are collections of particles produced by MC event generators. The output is the digital signal from all detector components in the same format that is used for real data (RAW).

The CMS simulation is based on the GEANT4 [51] toolkit. It allows modeling of the full detector geometry and the simulation of particles' propagation through magnetic fields as well as electromagnetic and hadronic interactions with the crossed material. The simulation is performed in two steps: the simulation step and the digitization step. In the simulation step, particles are propagated through the detector allowing for interaction with detector material. Hits are produced on interaction with sensitive detector components. In the digitization step, the electronic readout of these hits is simulated, taking into account resolution and detector response effects.

6 CMS Offline Computing and Analysis Software

This chapter gives a summary of the computing and software involved in this analysis, starting from the data recorded by the CMS experiment, describing the complete path to the high-level data analysis. First, the LHC computing grid is introduced, followed by a description of the CMS event data model. Then, the analysis specific software layer of CMS, the Physics Analysis Toolkit (PAT), is described. Finally the high-level data analysis development environment VISPA is described, which was used to design the dijet angular distribution analysis. The ROOT framework version 5.22 is also used extensively in the analysis and is described elsewhere [52].

6.1 CMS Offline Computing

The LHC computing Grid

CMS as well as the other LHC experiments are facing new computing challenges w.r.t. previous experiments [5,53]. The users of the CMS experiment are distributed among a large number of institutes worldwide. Datasets with large statistics will be collected over a running period of approximately 15 years, and high flexibility is required since the analysis needs are unforeseeable. The computing is therefore organized in a worldwide computing grid maintained by the Worldwide LHC Computing Grid (WLCG) [54] project, a collaboration between the LHC experiments, computing centers worldwide and grid middleware providers. The computing infrastructure and the data are organized in a tiered structure as depicted in Fig. 6.1, consisting of a central computing center located at CERN, the Tier-0, 8 large computing centers distributed around the world, the Tier-1s, and about 50 smaller scale Tier-2 centers worldwide.

The typical flow of data collected by the experiment to the end user's analysis is as follows. The Tier-0 receives the data collected by the detector online system in the RAW data-tier, containing the full event information read from the detector, with event sizes of $\sim$1-2MB. An express reconstruction of part of the data for alignment and calibration purposes as well as high priority analyses is performed. At the same time, the data are split up into primary datasets according to trigger definitions (e.g. jet dataset, muon dataset, etc.) in order to have datasets of appropriate sizes for analyses using specific objects. Then the Tier-0 performs a prompt reconstruction of physics objects stored in the RECO data-tier of $\sim$0.5MB/events which is then exported together with the RAW to at least two Tier-1 centers for permanent storage. The Tier-0

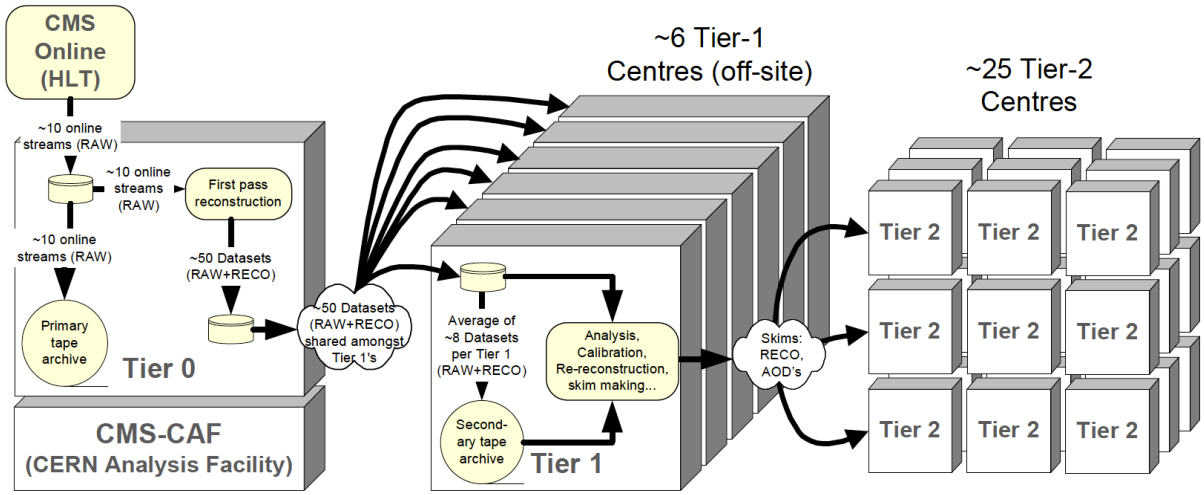


Figure 6.1: Organization of the CMS Computing infrastructure and data into a tiered structure [33].

is capable of buffering the data until safe transfer to the Tier-1s is completed. The Tier-1 centers perform second-pass-reconstructions and provide the Tier-2s with copies and subsets of the data, usually in form of the AOD data-tier which is designed to serve 90% of the analysis needs with a typical size of ∼100kB/event. The Tier-2 centers are dedicated for analysis activities and Monte Carlo dataset production.

The CMS Event Data Model

The CMS application framework, called CMS Event Data Model (EDM), is a modular framework centered around the concept of an event [55–57]. Independent C++ plugin modules can be executed according to Python configuration files. These plugins can read and write from/to a flexible event content with objects organized by their type as shown in Fig. 6.2. The data format is based on ROOT [52] and objects are stored in independent collections (TTrees and TBranches).

In the process of event reconstruction, typically lower level detector objects such as hits in the tracker and energy deposits in the calorimeters, are combined to higher-level analysis objects, such as electrons or jets in several independent steps, each represented by a module. The combined analysis objects are linked to their input objects via references, minimizing redundancy in persistent information. Due to the modularity, each reconstruction step is easily replaceable by an alternative algorithm, and the results of several algorithms can be stored in parallel. A data provenance instance keeps track of all utilized modules and their relevant configuration parameters that lead to the reconstruction of each object [58]. In this way, even after several replication steps, a clear identification of events and object definitions can be guaranteed.

Physics Analysis Tools in CMS

While the EDM is optimized for performance and flexibility of the reconstruction, the information in RECO/AOD is difficult to access for high-level data analyses. Therefore common

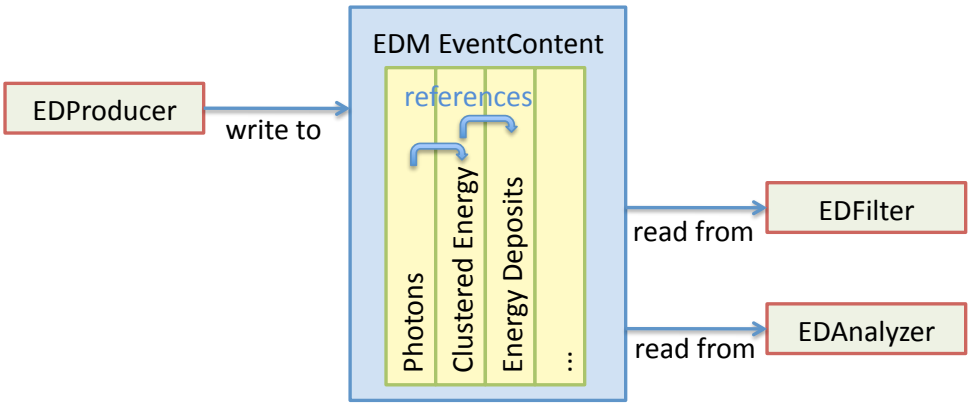


Figure 6.2: Demonstration of the CMS Event Data Model (EDM). C++ plugins modules of types EDProducer, EDAnalyzer, EDFilter read and write from/to the event content. Reconstructed objects, produced from lower levels objects, carry references to their input objects [57].

analysis data formats and algorithms are provided in form of the Physics Analysis Toolkit (PAT) [57]. Common data formats are provided to simplify controlling the event size for high-level data analysis purposes and avoid the adoption of different data formats by the various working groups. It facilitates transfer and comparison of analyses by common skims within physics analysis groups. In addition, PAT provides physics analysis groups with a common interface to the latest reconstruction algorithms developed within the collaboration.

Easy access to all data relevant for high-level analyses is achieved in PAT by introducing new data formats for physics objects, such as electrons or jets, which serve as entry point to all information associated to them. After the reconstruction step, objects of these types are produced consistently using C++ plug-in modules that can be fully configured using Python configuration files. This additional analysis layer allows to select, gather and combine all information relevant to a specific analysis.

PAT is completely configurable to any analysis specific need. For clearly arranged object selection, a user-friendly concept is available, by specification of cuts as strings in the configuration files [59]. All member functions, including those of referenced objects, are applicable. As an example, the module `selectedPatJets` is configured using a string parameter `cut` to specify the selection criteria: `selectedPatJets.cut="pt>30"`.

When dealing with huge configuration systems as in the CMS experiment, visualization is essential. Therefore, an editor with a graphical user interface, the ConfigEditor, was developed to facilitate the understanding and creation of configuration files [60–62]. With the ConfigEditor, CMS configuration files can be inspected, debugged and verified interactively.

Figure 6.3 shows a screenshot of the ConfigEditor. From top-left to right, the 4 main panels provide functionality for: interactive creation of user-defined configuration files, a tree-like representation of the workflow defined in the configuration file, a graphical representation of all modules and their dependencies, and a list of all configurable parameters for a selected module. The graphical user interface of the ConfigEditor is based on the VISPA [63] development environment for physics analyses which is described in the following section.

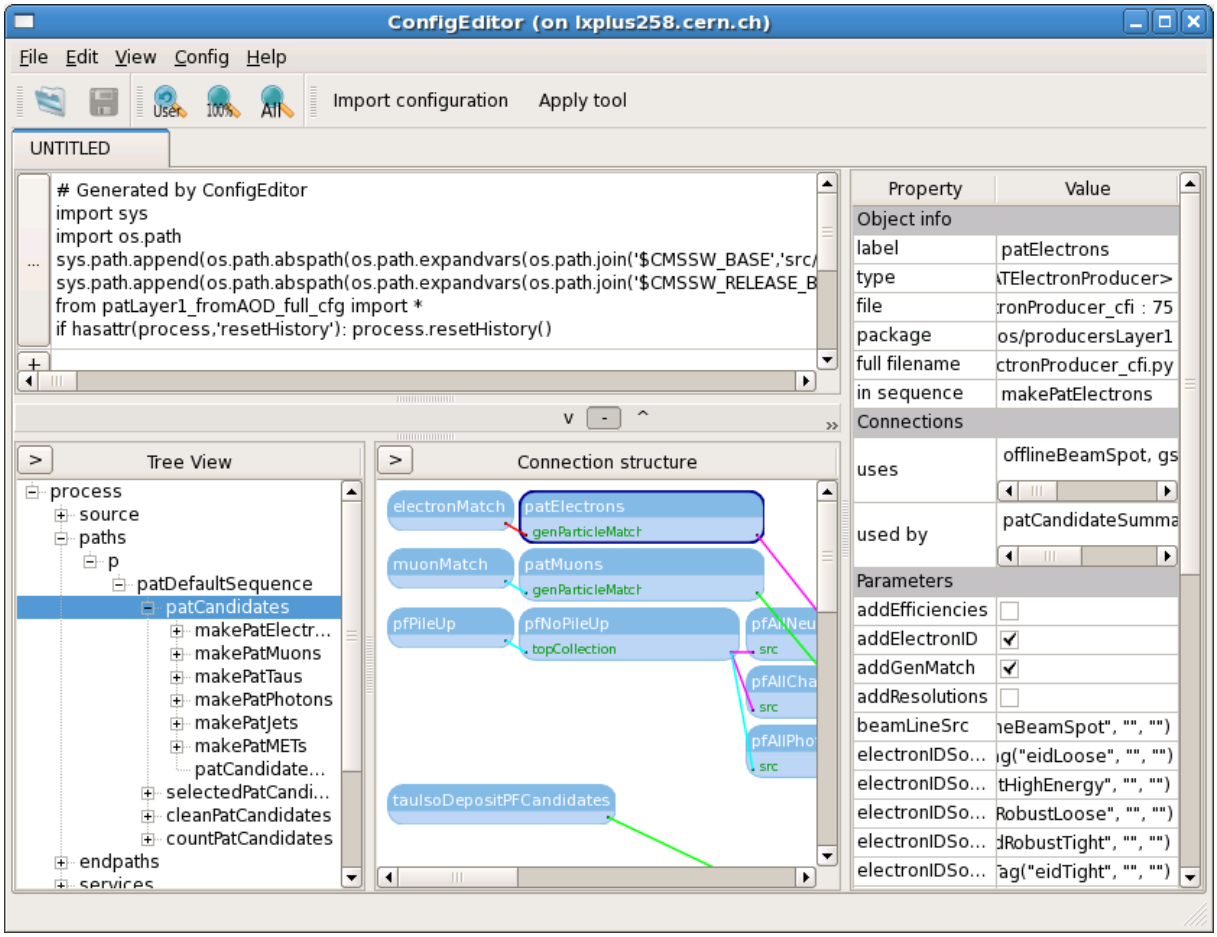


Figure 6.3: Screenshot of the ConfigEditor [62].

6.2 VISPA and PXL

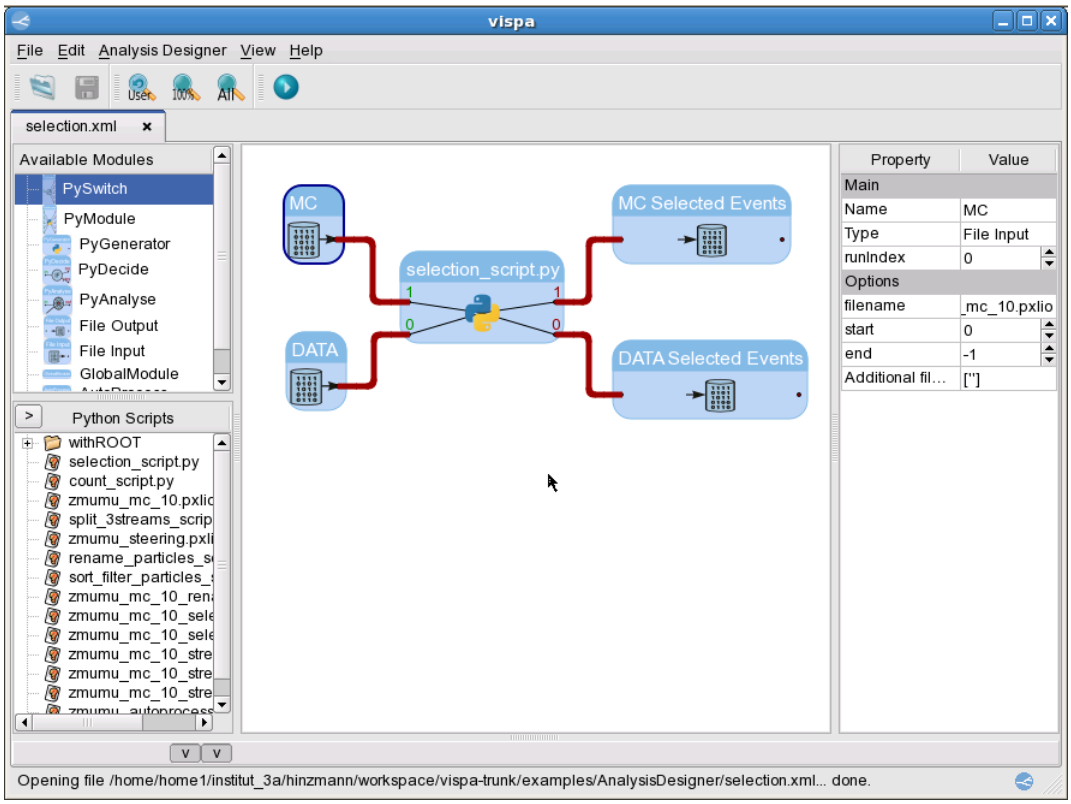
VISPA (Visual Physics Analysis) is a visual development environment for physics data analyses [63,64]. Its key idea is the development of physics analyses combining graphical and textual programming. Visual tools to design and execute modular analyses, to create analysis templates, and to browse physics event data are provided in a multipurpose window. Analysis modules can be programmed in both C++ and Python depending on the convenience and performance needs. VISPA is available for Linux, Windows and Mac OS X and supplies its own data format including physics objects and containers to allow convenient transport of analyses between different computers. In the following, first, the underlying C++ toolkit for physics analyses, PXL, is described, followed by a description of the Python-based graphical user interface of VISPA.

The Physics eXtension Library (PXL)

The Physics eXtension Library (PXL) is a C++ toolkit offering a collection of tools to support experiment independent physics analyses [63,65]. Each component is accessible within VISPA, but can be used separately as well. The main components of PXL are discussed in the following.

PXL provides interfaces for physics analyses in high energy and astroparticle physics, including physics objects (e.g. pxl::Particle), containers (e.g. pxl::Event), object relations (e.g. the pxl::Relations of pxl::Particles) and an interface allowing for the addition of user specific data to any of these objects (e.g. pxl::UserRecords). All classes are accessible in C++ and Python [66], enabling the implementation of analysis modules in both languages. The Python classes are wrapped around the C++ classes using the SWIG [67] tool.

The fast, compressed and flexible I/O format of PXL allows for simple splitting and merging of data at the file level and exchange of data between all common platforms (Linux, Windows and MacOS). The common mother class of all classes in PXL, the pxl::Serializable, defines the interface for their serialization (deserialization) to (from) disk.



(a)

Figure 6.4: Screen shot of the analysis designer in VISPA [63].

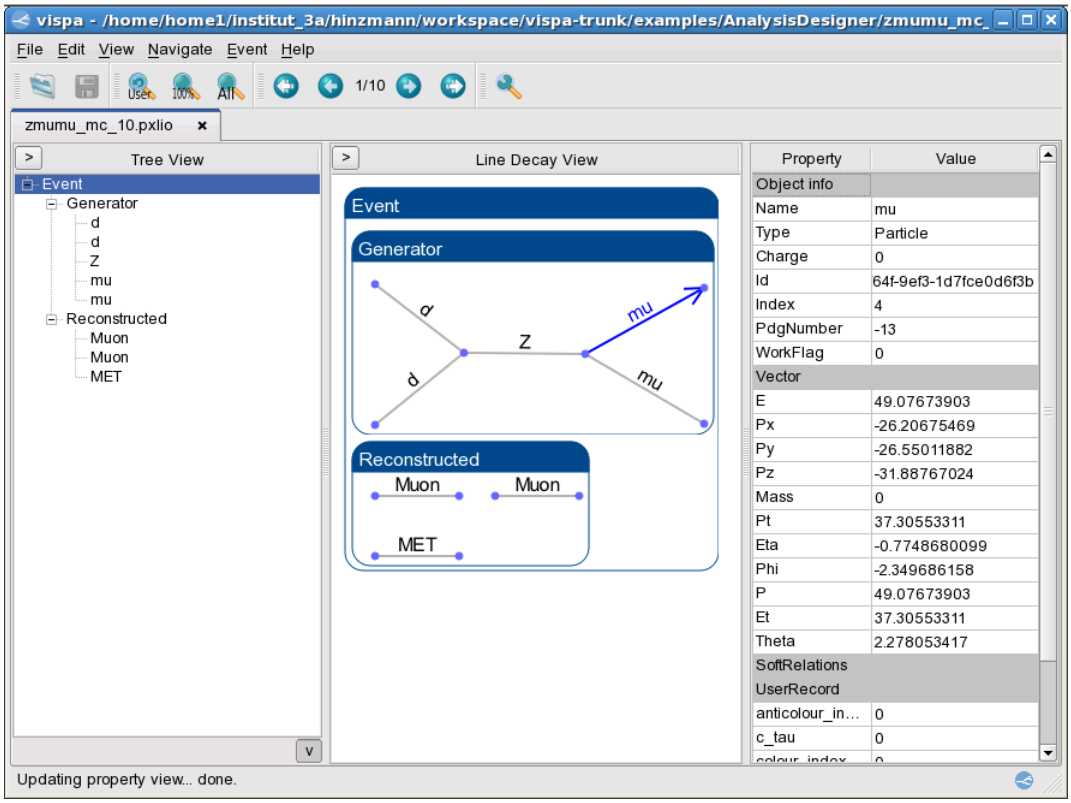
The framework for modular physics analyses in PXL defines the analysis logic for VISPA. A simple example analysis is shown in Fig. 6.4. The data flow between modules, visualized by connection lines from left to right, is controlled via multiple sinks and sources of the modules which can be connected to each other. The modules and connections are accessible via a central interface, the pxl::Analysis, which can be saved and loaded in XML format. The graphical user interface of VISPA provides the functionality to automatically create a tar-ball of a pxl::Analysis in order to exchange complete analyses within working groups. Analyses can be run on a batch system or interactively using the graphical user interface of VISPA.

The modules exchange data via a common interface, which for high energy physics is the pxl::Event, but in principle any pxl::Serializable can be exchanged. This interface for modules

guaranties easy sharing and reuse of common modules within a working group. A variety of module interfaces and examples are available in PXL, from which users derive their analysis specific modules. Both C++ and Python modules can be used within the same analysis. While C++ modules are typically used for performance-sensitive analysis tasks, Python modules are used for fast prototyping and analysis logic.

Visual development of physics analyses

Figure 6.4 shows the graphical user interface for the visual development of analyses in VISPA. New analyses are designed by connecting and configuring a chosen set of modules. On the left side of the window, a list of the available C++ and Python modules in shown. These can be inserted into the analysis and connection lines can be drawn using drag and drop. The properties grid on the right side of the window allows to modify the parameters of the modules. The source code of each module can be edited by a double click on the module. A double click on input and output modules opens a data browser for inspection of the corresponding file, explained in the following.



(a)

Figure 6.5: Screen shot of the data browser in VISPA [63].

Figure 6.5 shows the data browser of VISPA, which provides a visual representation of all data contained in a PXL data file. Decay trees are drawn according to the relations of the displayed objects. The properties grid on the right side of the window allows to inspect all properties including user defined data of an object. In this way every single object and

parameter in a file can be inspected on an event by event basis. The data browser further serves as an editor for event templates. These can be used to create steering files for automated event reconstruction, described in detail in Ref. [68].

The graphical user interface is based on the platform independent application framework PyQt4 [69,70]. To allow straightforward implementation and integration with experiment specific software, the graphical platform provides a plugin mechanism for new applications and a large collection of graphical components. An example for a successful integration with experiment specific software is the ConfigEditor [60] in CMS, described in the previous section.

7 Jet Reconstruction and Performance

Outgoing partons of high momentum transfer QCD interactions form jets of particles. Hence, jets are an important tool to probe the predictions from perturbative QCD. In order to compare theory predictions and experimental data, a common algorithm to cluster particles into jets and reconstruct the energy and direction of the original parton needs to be defined. While in theory jets are clustered from the parton or particle final state, in experiments the energy deposited throughout the detector is clustered to reconstruct the energy of the jets of particles coming from the parton final state.

In this chapter the reconstruction of jets and their performance is described. First, jet clustering algorithms are introduced which are necessary to compare theory calculations with experimental measurements. Then, the methods to reconstruct jets from detector signals in CMS are explained, followed by a description on how detector noise is filtered. This is followed by an explanation of the jet energy calibration procedure in CMS and its precision. Finally, the jet performance in energy resolution, angular resolution and reconstruction efficiency is shown.

7.1 Jet Clustering Algorithms

Jet clustering algorithms are used to compare theoretical predictions from perturbative QCD with experimental results. They reduce the complex structure of particle jets from a scattered parton to a simple four-momentum, the main property of particle jets. Jet clustering algorithms allow comparison of theory and experiment, without the need of a complete understanding of the hadronization process and also without the need to detect and identify each single particle inside the jet.

Two properties of jet clustering algorithms are desired to be able to calculate meaningful theory predictions. Clustering algorithms need to be infrared-safe, which means that the emission of infinitesimally-low-energy partons from partons inside a jet does not affect the jet properties. Further they need to be collinear-safe, which means that the splitting of a parton inside a jet into two collinear partons does not affect the jet properties.

Jet algorithms for hadron colliders can be divided into two classes: cone algorithms [71] and sequential clustering algorithms [72–76]. The LHC experiments have converged on a particular sequential clustering algorithm which fulfills the theory requirements and has good properties for experimental measurements: the anti-k_T algorithm [76].

The anti-k_T algorithm is an iterative algorithm defined as follows. For all combinations of particles in an event, the distance variable d_{ij} between two particles i and j, defined as

$$d_{ij} = \frac{1}{\max(k_{Ti}^2, k_{Tj}^2)} \frac{\Delta R_{ij}^2}{R^2}, \tag{7.1}$$

as well as the distance d_{iB} of each particle i w.r.t. the beam pipe, defined as

$$d_{iB} = \frac{1}{k_{Ti}^2}, \tag{7.2}$$

are calculated. The $\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ describe the angular distances between two particles i and j. The k_{Ti} and k_{Tj} are the transverse momenta of the two particles. The distance parameter R is responsible for defining the angular size of the jets. If the d_{iB} of a particle i is smaller than all d_{ij}, it is considered to be a jet and is removed from the list of particles in the event. Otherwise, the two particles with the smallest d_{ij} are combined into a single object by adding their 4-vectors and another iteration is run. The iteration ends once all particles in the event have been clustered into jets.

This algorithm is collinear-safe as the clustering is driven by the angular distance between two particles. Gluons emitted at small angles are picked up by the algorithm in early steps of the iteration and therefore don't affect the jet properties. The algorithm is also infrared-safe, but in a more indirect way. Due to the weighting with the inverse maximum squared transverse momentum, the emission of additional low energy gluons are picked up rather late in the clustering process. They are picked up after all hard emissions at small angles, but, importantly, before two soft particles can cluster with each other. Soft emissions will therefore not cluster into separate jets, preserving infrared-safety.

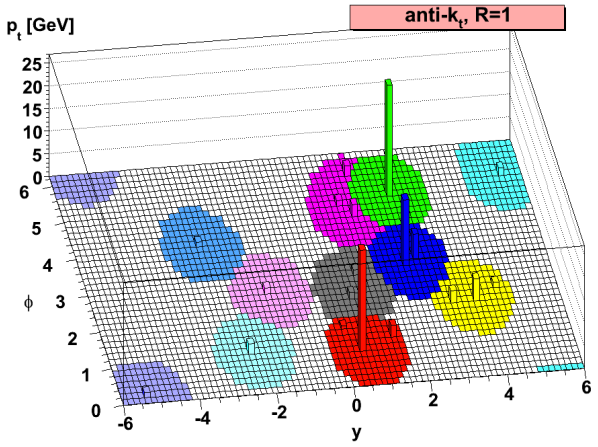


Figure 7.1: Example event clustered using the anti-k_T algorithm [76].

The anti-k_T algorithm has very good experimental properties. By construction it tends to form almost circular jets with a radius of R. This is demonstrated for an example event in Fig. 7.1. Only if two jets are closer than $2R$, the shape of the two jets varies from being circular. Experimentally this feature is very important as it allows straight-forward calibration and understanding of the detector acceptance. Due to its limited cone size, the measurement of jets can be restricted to active detector regions. Due to the fixed cone size, the number of detector cells to be clustered into a jet is constant over certain detector regions and therefore straight-forward to calibrate.

The choice of the distance parameters R, depends on the problem to be addressed by the measurement. While large cone size jets collect all energy from the scattered parton, they

also pick up a large contribution of background energy from the underlying event or pile-up interactions. Small cone size jets pick up little contamination, but may not collect all energy from the scattered parton. The choice also depends on the p_T of the jets to be measured as the average jet radius depends on the jet p_T. The two default choices in CMS are $R = 0.5$ and $R = 0.7$. This study uses $R = 0.5$ to be rather safe against background contributions, but we will see that the results are independent of the choice of the distance parameter in Sec. 11.4.

7.2 Jet Reconstruction Methods in CMS

In CMS four standard methods for jet reconstruction are available as depicted in Fig. 7.2. These methods make use of different detector components, e.g. the tracker and the calorimeters, and give different reconstructed objects as input to the above explained jet clustering algorithms. This variety of methods was very valuable during startup phase of the LHC to commission the jet reconstruction methods in comparison to each other [13–15]. The four methods are briefly described in the following, with primary focus on calorimeter jets which are used in this analysis.

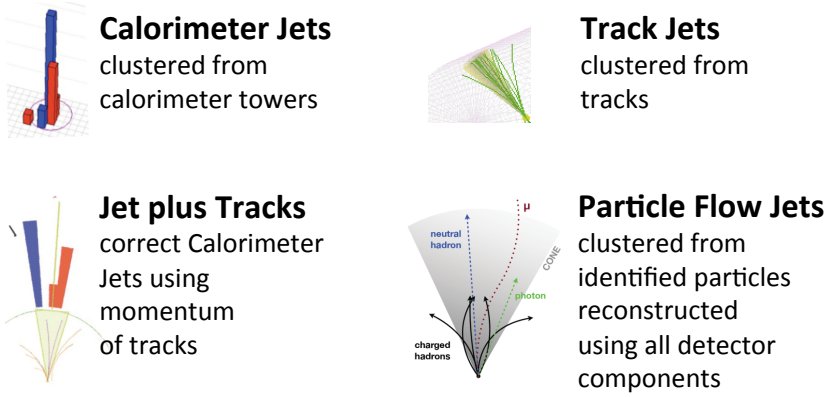


Figure 7.2: Graphical representation of the four jet reconstruction methods in CMS [77].

Calorimeter jets (CaloJets) are clustered from the energy measured in the electromagnetic and hadronic calorimeters (ECAL and HCAL) combined into towers [14]. Due to the finer granularity of the ECAL, typically several ECAL cells and a single HCAL cell form one tower. Together, the calorimeters collect electromagnetic energy (electrons and photons) and hadronic energy (charged and neutral hadrons).

A part of the jet energy is not measured since low momentum charged particles do not reach the calorimeters as they are deflected by the magnetic field. This affects the energy calibration of the jets and also the resolution in the ϕ-direction of the jet measurement. Also, some electromagnetic energy is lost in the 1-3 radiation lengths of the material in the tracker. Further, highly energetic jets may punch through the hadronic calorimeter which is limited to 10 interaction lengths, but this energy is partly recovered by a tail catcher calorimeter located outside the solenoid magnet.

Due to the fact that the collisions may happen along the beam axis within a distance of up to 10 cm from the detector center, the angular measurement of jets cannot simply be done w.r.t. to

the detector center, but must include the information of the primary vertex of a collision. The reconstructed jet angles are therefore reconstructed w.r.t. to the primary vertex of the collision.

As the reconstruction of vertices is crucial for the measurement of jet angles, it is briefly described in this paragraph. A detailed description can be found in Ref. [35]. CMS uses an iterative algorithm to reconstruct tracks from hits in the silicon pixel and strip detectors [78]. In several iterations, tracks are reconstructed using a combinatorial track finder using e.g. pixel hit triplets as seeds. Prompt tracks originating from the primary interaction region are selected based on a set of quality requirements, including the transverse impact parameter significance with respect to the beam line, the number of strip and pixel hits, and the normalized χ^2 of the track fit. The selected tracks are then clustered into vertices based on their position along the beam axis. The exact positions of the vertices are obtained in a fit to all assigned tracks.

Another jet reconstruction method in CMS are jets clustered from reconstructed charged particle tracks (TrackJets) [15]. They contain on average about 60% of the energy of the original parton, as typically 30% of the particles in a jet are neutral particles (mostly π^0). In contrast to CaloJets, they have high efficiency at low transverse momenta where particles are deflected by the magnetic field and do not reach the calorimeters. This approach is completely orthogonal to CaloJets as independent detector components are used.

The JetsPlusTracks (JPT) method [79] combines reconstructed CaloJets with the information of the reconstructed tracks. The energy of the reconstructed CaloJets is corrected for the momenta carried by the tracks deflected by the magnetic field outside and inside the jet cone. This procedure significantly improves the jet energy resolution at low p_T without the need to match single tracks to calorimeter cells, but simply by assignment of tracks to the jet cone.

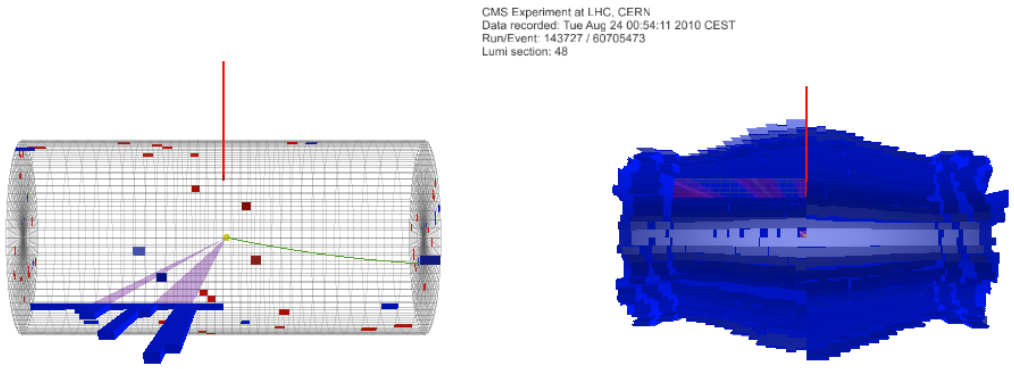
The ParticleFlow (PF) method [80,81] makes use of all detector components to reconstruct each single particle passing the detector. Electrons, photons, muons, charged hadrons and neutral hadrons are identified using the tracker, the calorimeters and the muon system. Particle trajectories measured in the tracker are extrapolated through the detector volume and are matched to the energy deposits in the calorimeters. While the momentum of neutral particles is measured in the calorimeters, the momentum of charged particles is measured by the tracker with a better resolution. In this way both the angular and energy measurements are improved as this algorithm makes use of the full granularity of the tracker and the ECAL which is much higher than that of the HCAL.

7.3 Noise Filtering

Noise filtering is an important part of jet reconstruction, as noise in calorimeter cells can lead to fake jets and mis-reconstruction of jet energies and directions. The description in the following corresponds to the algorithms used for the reconstruction of CaloJets. The noise filtering is performed at two steps of the reconstruction: first, on the basis of single calorimeters cells, and second, on the basis of reconstructed jets.

The filter algorithm based on single calorimeter cells is run before the reconstruction of towers and jets. Channels which are known to be dead or noisy are excluded from the following

clustering. So called ECAL spikes which arise when particles traverse the readout Avalanche Photo Diodes (APDs) of the ECAL are removed based on timing cuts and by a cut on the ratio of the cell energy divided by the energy collected in the surrounding 4 cells [82].



(a) (b)

Figure 7.3: Events with electronic noise (a) in a single HPD and (b) in the full HCAL.

The filter based on reconstructed jets, called jet ID, identifies jets which do not come from noise based on a set of quality requirements [83]. The requirements are motivated by observations in noise-only samples. Three typical kinds of noise are observed in the Hybrid Photo Detectors (HPDs), the Readout BoXes (RBXs) and the full HCAL. The HPDs measure the scintillation light in the HCAL barrel. Each HPD measures a set of cells arranged at same ϕ. Therefore electronic noise in an HPD is observed in several cells arranged at same ϕ. An example HPD noise event is shown in Fig. 7.3 (a). The RBXs collect the signals from three neighboring HPDs and produce very similar noise patterns. Figure 7.3 (b) is another extreme example of a noise event, where the full HCAL shows noise.

Table 7.1: "Loose" jet ID criteria for CaloJets [83].

variable	loose JetID
EMF ($\|\eta\| < 2.6$)	> 0.01
n^{90}_{hits}	> 1
f_{HPD}	< 0.98

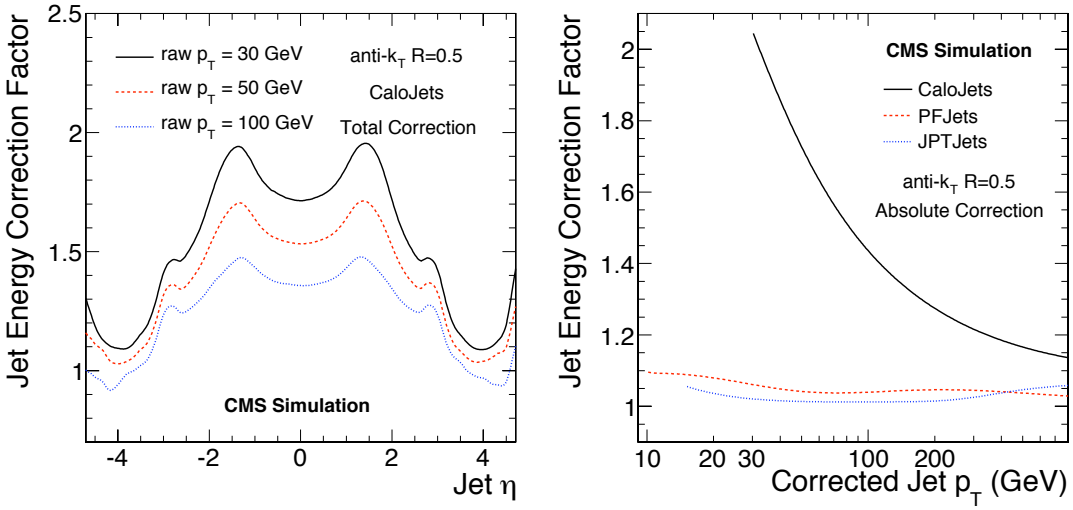
The jet ID requirements for CaloJets are designed to filter this noise, at the same time being $> 99\%$ efficient for non-noise jets. The criteria of the "loose" jet ID are given in Tab. 7.1. Jets are required to have a non-zero electromagnetic energy fraction (EMF). This criterion rejects jets which do not deposit energy in the ECAL, meaning unphysical noise in the HCAL. A second cut is required on the number of hits which contain 90% of the jet energy (n^{90}_{hits}), rejecting jets from noise in single calorimeter towers. The third cut on the fraction of energy contributed by the hottest (highest energy) HPD (f_{HPD}) is designed to reject HPD noise.

7.4 Jet Energy Calibration

The jet energy calibration is an important ingredient for any jet cross section measurement, in this study particularly for the measurement of the dijet mass. Goal of the jet energy calibration is to relate the jet energies measured in the detector to the energies of the jets of particles which traverse the detector.

Each calorimeter cell in CMS has been energy-calibrated using test beam data [38]. However, various detector effects, such as the deflection of particles by the magnetic field before they reach the calorimeters, change the measured jet energies w.r.t. the energy of the particle jet. CMS uses a factorized approach to correct for these effects [14,16]. The three correction steps are explained in the following.

The first step in the CMS approach is a correction to the jet energies to eliminate pile-up effects [14]. Pile-up collisions generate additional energy in the event which is picked up in the jet reconstruction of the main collision. In the 2010 data, with a median of 2-3 pile-up collisions, these effects are found to be rather small. Therefore the correction is not applied and instead treated as an uncertainty on the jet energy scale.



(a) (b)

Figure 7.4: Simulation-driven jet energy corrections as a function of jet (a) pseudo-rapidity and (b) transverse momentum [14].

Second, a simulation driven jet energy response correction is applied [14]. The detector simulation takes into account effects of particles deflected by the magnetic field, energy lost when traversing the detector material, particle conversions, and a detailed detector geometry. A correction function to flatten the jet energy response as function of jet pseudo-rapidity and transverse momentum is derived from simulated dijet events, by comparing the transverse momenta of the particle jets and the simulated CaloJets. This correction function is show in Fig. 7.4 (a,b). It differs strongly depending on the jet reconstruction method. While CaloJets need a large correction factor at low transverse momentum due to particles which do not reach the calorimeters within the jet cone as they are deflected by the magnetic field, PF jets and JPT jets need almost no correction, as they make additional use of the reconstructed particle tracks.

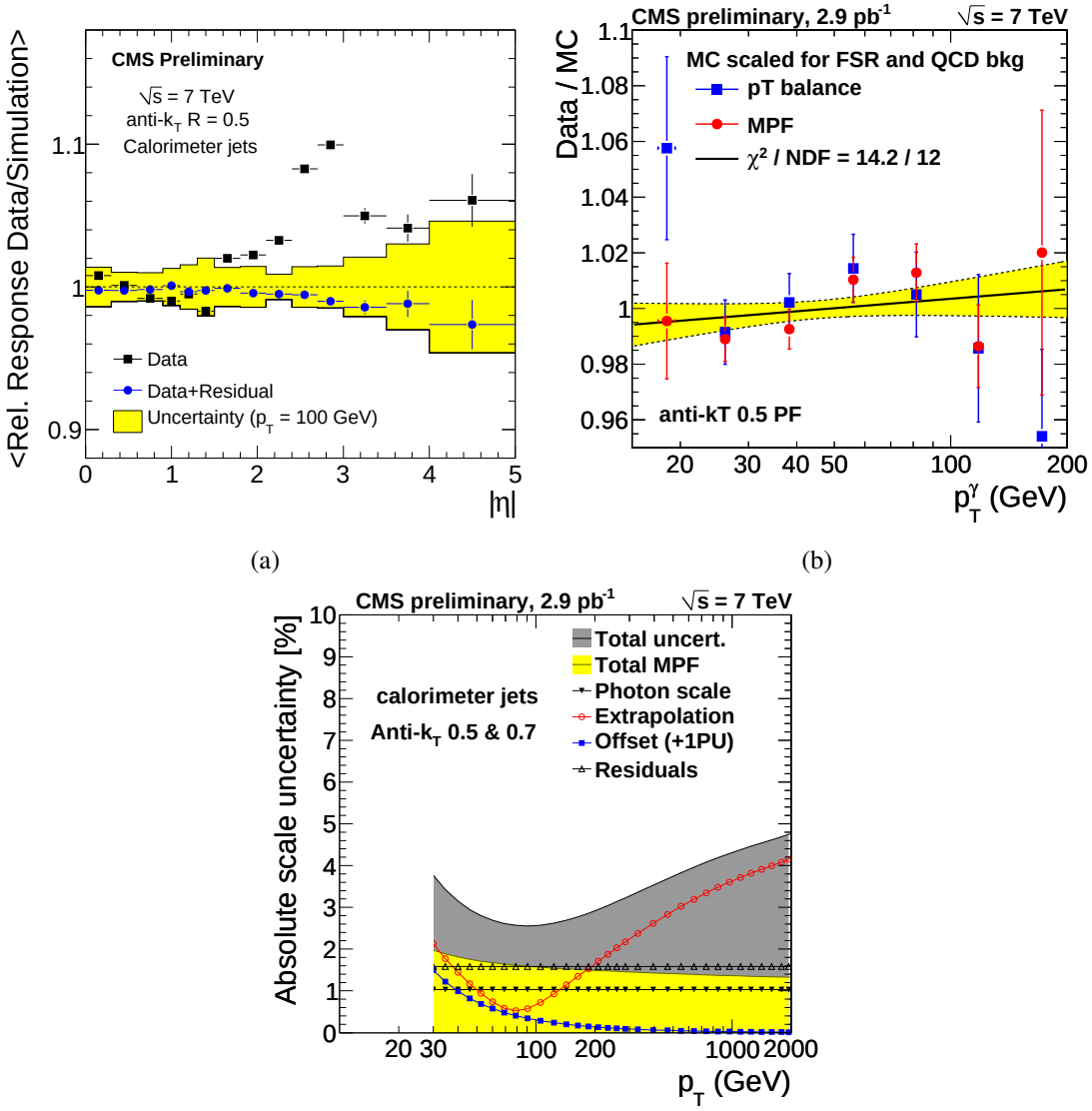


Figure 7.5: Data-driven jet energy corrections in CMS [16]: (a) result of the residual jet energy correction using di-jet events, (b) result of the absolute jet energy correction using photon+jet events, (c) uncertainty of the absolute jet energy scale.

Finally data-driven residual corrections are applied which correct for any measurable difference between the detector simulation and the jets measured in data [16]. This correction is done in two steps: first, the calibration as a function of jet pseudo-rapidity, and second, as a function of transverse momentum.

The fact that dijet events are balanced in transverse momentum is used to calibrate the jet energies as a function of pseudo-rapidity w.r.t. jets reconstructed in the barrel part of the detector. Only events with back-to-back dijets and little additional activity in the event are used, to avoid any impact from unbalanced events. In Fig. 7.5 (a) the resulting residual correction is demonstrated, by showing the ratio of data and simulation before ("Data") and after ("Data+Residual") correction. The dominant uncertainties on this correction are the limited statistics of dijet events and the uncertainty on the modeling of the jet energy resolution in simulation, used to correct for resolution effects influencing the dijet balance method. The correction is derived as a func-

tion of the absolute of the pseudo-rapidity $|\eta|$ due to the limited data statistics. Asymmetries in $+\eta$ and $-\eta$ in the detector response are covered by the assigned uncertainty. The uncertainty is shown in Fig. 7.5 (a) as a yellow band. The relative jet energy calibration is typically known to the level of 1% per unit of pseudo-rapidity.

The jet energy is calibrated as a function of transverse momentum using photon+jet events in two independent approaches. The simpler approach is the photon+jet balancing method. It makes use of the fact that the jet transverse momentum should be equal to the photon transverse momentum which is measured using the precisely calibrated ECAL. Hereby, additional event activity has to be carefully taken into account. The more advanced technique, the so called Missing energy Projection Fraction (MPF) method, takes into account the missing transverse energy measured in the calorimeters to balance the photon and jet momenta:

$$R_\gamma \vec{p_T}^{\gamma} + R_{recoil} \vec{p_T}^{recoil} = \vec{E_T}^{miss} \tag{7.3}$$

Here, R_γ and R_{recoil} are the detector responses to the photon and the jet. In this method the additional event activity is taken into account by the missing transverse energy E_T^{miss}. Therefore, additional jets in the event have only a small impact on the measurement which is nevertheless taken into account. The results of both methods are summarized in Fig. 7.5 (b). The ratio of data and simulation, meaning the residual correction as a function of transverse momentum, is very close to unity. The uncertainties on the absolute jet energy scale are given in Fig. 7.5 (c). The dominant sources are the Photon (ECAL) calibration, the uncertainties of the MPF method, and the extrapolation to ranges of transverse momentum where no photon+jet data are available for calibration (e.g. $p_T > 200$ GeV). The total uncertainties on the jet energy scale have been parametrized as a function of η and p_T and range from 3%-6% for CaloJets with $p_T >$30 GeV.

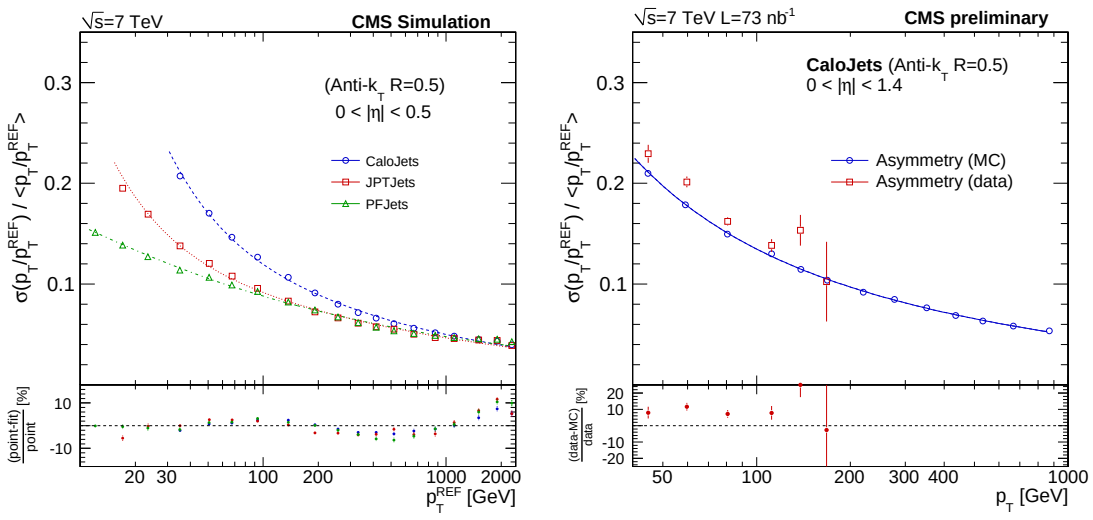
7.5 Jet Energy Resolution

The energy resolution of jets is limited by energy fluctuations from electronic noise, energy sampling, and from non-uniformities and non-linearities in the detector response. In the following, a brief description of the derivation of the jet energy resolutions from simulation and their validation with data is given [14].

The jet energy resolution is derived from simulation by comparing simulated CaloJet momenta with those of their matching particle jets. The resolution is parametrized using a double-sided Crystal-Ball function describing the resolutions as a Gaussian core with tails. Figure 7.6 shows the Gaussian core resolution for CaloJets as a function of jet transverse momentum, which is around 10% at $p_T = 100$ GeV. The PF jets and JPT jets have a better resolution compared to CaloJets jets as they make use of the tracker.

Dijet events allow a direct measurement of the jet energy resolution from data which is used to validate the resolution derived from simulation. The so called dijet asymmetry method makes use of the fact that for perfectly balanced dijet events, the connection between the Gaussian resolution and the asymmetry in transverse momentum is given by

$$\frac{\sigma(p_T)}{p_T} = \sqrt{2}\sigma\left(\frac{p_T^{jet1} - p_T^{jet2}}{p_T^{jet1} + p_T^{jet2}}\right) \tag{7.4}$$



(a) (b)

Figure 7.6: (a) Jet energy resolution derived from simulation. (b) Jet energy resolution measured using the dijet asymmetry [14].

This relation is used to extract the jet energy resolution from dijet events. Additional activity in the event has to be properly taken into account, e.g. by extrapolating the result for events with different cuts on the p_T of an additional third jet in the event to $p_T = 0$ GeV. The resulting resolution in shown in Fig. 7.6 (b) for data and simulation. Overall the resolution observed in data is about 10% higher than expected from the simulation. During the startup phase of the LHC, this difference is accounted for as an uncertainty on the resolution of 10%. In future studies this difference will be corrected [84].

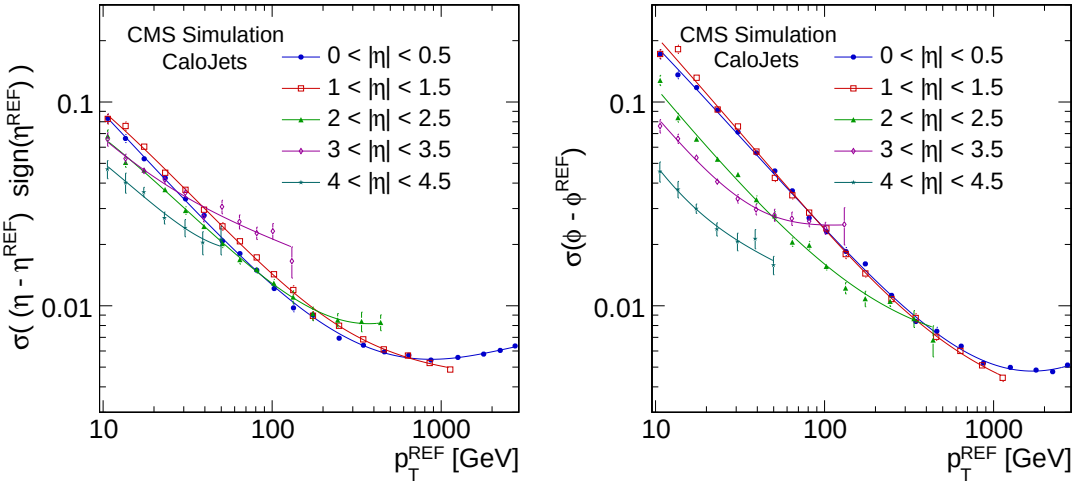
7.6 Jet Angular Resolution

The angular resolutions of jets in η and ϕ are an essential ingredient for the measurement of the dijet angular distributions. They are mainly limited by the size of the calorimeter towers and the deflection of particles out of the jet cone by the magnetic field. In the following, a brief description of the derivation of the jet angular resolutions from simulation and their validation with data is given [14].

The angular resolutions are derived from simulation by comparing simulated CaloJet η and ϕ with those of their matching particle jets [14]. They are parametrized using a piecewise-defined function describing the resolutions as a Gaussian core with tails. Figure 7.7 shows the Gaussian core resolution which is around 0.01 in η and 0.02 in ϕ at $p_T = 100$ GeV, and is therefore sufficient for what will be needed for the measurement of the dijet angular distributions.

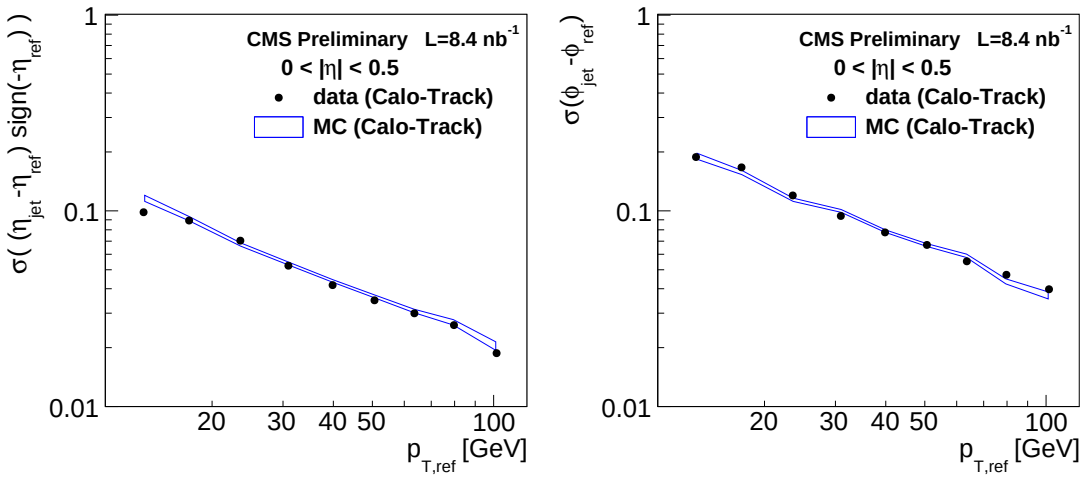
The parametrized resolutions are validated in data by measuring the difference in jet direction of the same jet measured using CaloJets and TrackJets:

$$\sigma(\eta_{\text{CaloJet}} - \eta_{\text{TrackJet}}) sign(\eta_{\text{TrackJet}}) \tag{7.5}$$



(a) (b)

Figure 7.7: Jet angular resolutions derived from simulation [14].

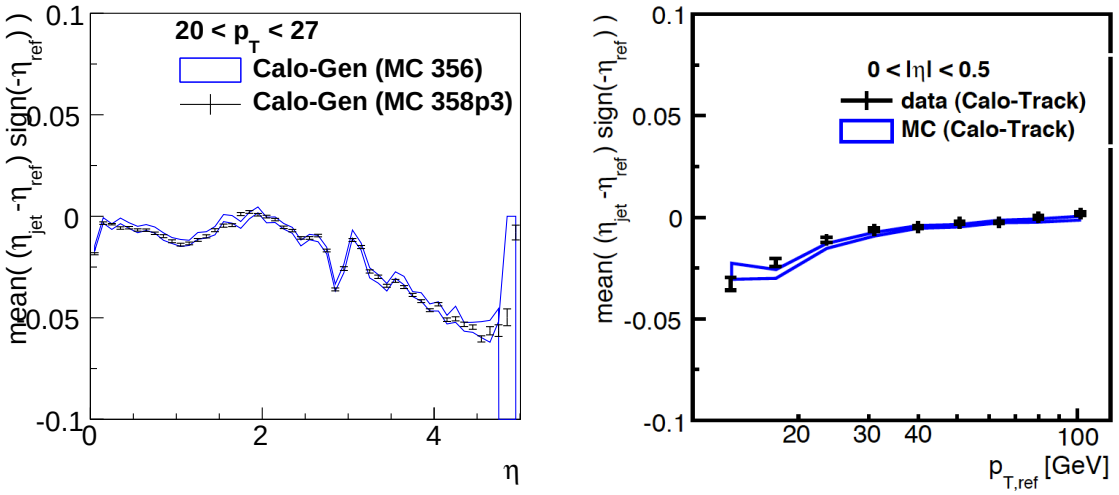


(a) (b)

Figure 7.8: Validation of the jet angular resolutions in data [14].

This variable contains both CaloJet and TrackJet resolution information but is dominated by the much larger CaloJet angular resolution and therefore a good quantity to validate the CaloJet resolution in data. The results are shown in Fig. 7.8 for data and simulation. A good agreement between data and simulation is observed, giving confidence in the modeling of the angular resolutions in simulation.

Another important aspect of the jet angle measurement is the average mis-measurement of the jet pseudo-rapidity, which can be derived from simulation from the mean offset in η of the reconstructed CaloJets w.r.t. the particle jets. It is shown in Fig. 7.9 (a) as a function of pseudo-rapidity. In the region of $|\eta| < 2.5$ the mis-measurement is smaller than 0.02. Above $|\eta| > 2.5$, where part of the jet is measured by forward calorimeters at $|\eta| > 3$, the mis-measurement is larger. This is due to the fact that the forward calorimeter towers are not projective, meaning that their surface is not pointing towards the detector center, but parallel to the beam axis instead.



(a) (b)

Figure 7.9: (a) Pseudo-rapidity offset of CaloJets in simulation. (b) Pseudo-rapidity offset of CaloJets w.r.t. TrackJets in data and simulation.

Therefore a correction needs to be applied for the fact that the jet direction is not equal to the pseudo-rapidity of the tower center, but rather to the pseudo-rapidity inside the tower where the shower maximum of a jet appears.

In order to validate that this offset is also visible in data, the offset between CaloJets and TrackJets, which are known to measure the pseudo-rapidity accurately, are compared in data and simulation as shown in Fig. 7.9 (b). It is found that the offset in data and simulation agrees, and above $p_T > 30$ GeV, the jet pseudo-rapidity is measured more accurately than 0.01.

7.7 Jet Reconstruction efficiency

For measuring jet cross sections, it is essential to have a stable jet reconstruction efficiency in the kinematic region of the measurement. From simulation it is known that the CMS detector has a reconstruction efficiency very close to 100% for CaloJets of transverse momenta above $p_T > 30$ GeV [85]. This is validated in first data using TrackJets which have practically 100% reconstruction efficiency even at lower transverse momenta [15].

Figure 7.10 shows the efficiency of finding a CaloJet within a cone of $\Delta R < 0.5$ near a reconstructed TrackJet, i.e. their matching efficiency. The transverse momentum of the TrackJet, shown on the x-axis, is corrected for the missing neutral fraction by an η- and p_T-dependent correction of about 1.5. The matching efficiency is a measure of the reconstruction efficiency for CaloJets. Is is, however, distorted by the fact that the energy resolution of the TrackJets is worse than of CaloJets. Nevertheless, this measurement clearly shows that around 30 GeV, the jet reconstruction efficiency approaches 99%. The turn-on energy in simulation and data agrees within 5%, validating the modeling of the reconstruction efficiency in the simulation.

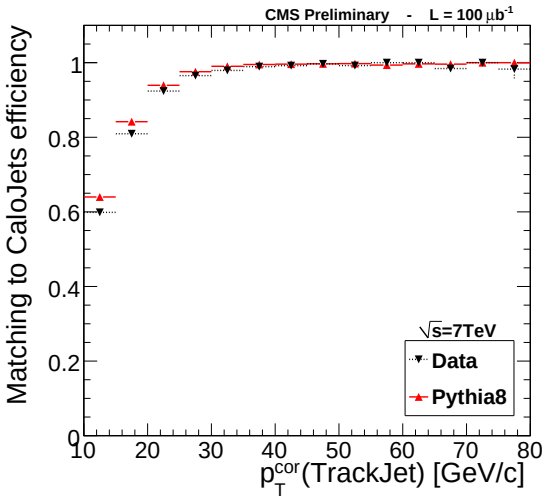


Figure 7.10: Matching efficiency for CaloJets w.r.t. Trackjets in data and simulation [15].

8 The Data Sample and Event Selection

In this chapter the data sample used for this analysis is described. First, the details about the data recording of CMS in 2010 are given. Then the kinematic reconstruction and selection of dijet events is explained, followed by a discussion of the trigger efficiencies and event cleaning cuts. Finally kinematic distributions in data and simulation are compared and conclusions on the accuracy of the simulation are drawn.

8.1 The Dataset

The dataset used in this analysis was recorded in the 2010 run of the LHC at $\sqrt{s} = 7$ TeV. Of the 47 pb^{-1} of integrated luminosity delivered by the LHC, CMS recorded 43 pb^{-1}, corresponding to an overall data taking efficiency of 90%. In 36 pb^{-1} of these data, validated for physics analyses, the magnetic field was at the nominal value of 3.8 T and all sub-detectors were in perfect conditions, each with at least 98% of all channels operational.

Input to this analysis were the jet primary datasets of the 2010A(first 3 pb^{-1}) and 2010B(remaining 33 pb^{-1}) data taking periods in AOD format. All data validated for physics analyses from the runs 136033-149442 were used.

8.2 Dijet Reconstruction and Selection

Dijet events are reconstructed and selected as follows. First, calorimeter jets are clustered using the anti-kT algorithm with a distance parameter of $R = 0.5$ in the pseudo-rapidity range $|\eta| < 5$. The two jets with highest transverse momenta $p_{T,1}$ and $p_{T,2}$, called the two leading jets, are selected. The variables of the dijet system are then given by:

$$M_{jj}^2 = (E_1 + E_2)^2 - (\vec{p}_1 + \vec{p}_2)^2 \tag{8.1}$$

$$y_{\text{boost}} = \frac{1}{2}(y_1 + y_2) \tag{8.2}$$

$$\chi_{\text{dijet}} = e^{2y^\star} = e^{|y_1 - y_2|} \tag{8.3}$$

In this analysis, dijet events are counted in bins of M_{jj} and χ_{dijet}. The binning in χ_{dijet} is 1-2, 2-3, 3-4, 4-5, 5-6, 6-7, 7-8, 8-9, 9-10, 10-12, 12-14, 14-16 and some of these bins are merged in regions of M_{jj} where few data statistics are available. The binning is chosen to be fine enough to resolve new physics signals from the predictions of QCD and coarser than the

detector resolution. There are 9 bins in M_{jj}: 250-350, 350-500, 500-650, 650-850, 850-1100, 1100-1400, 1400-1800, 1800-2200, 2200-∞ GeV. The minimum mass thresholds of the 5 lowest M_{jj} bins are chosen at the points where the jet triggers are fully efficient in order to optimize the number of events collected in these bins (see Sec. 8.3 for details). The remaining dijet mass bins are chosen with a width larger than the dijet mass resolution, requiring at least 20 events in the highest mass bin to resolve the shape of the angular distributions.

Two kinematic cuts are applied in order to restrict the measurement within a well understood region of the detector, avoiding the transition region between the HCAL and the forward calorimeters at $2.8 < |\eta| < 3.2$. By requiring

$$\chi_{\text{dijet}} < 16 \tag{8.4}$$

$$|y_{\text{boost}}| < 1.11 \tag{8.5}$$

the rapidities of the two jets are effectively restricted to

$$|y_1| < 2.5 \wedge |y_2| < 2.5 \tag{8.6}$$

through the inequality:

$$\max(|y_1|, |y_2|) < |y_{\text{boost}}| + y^\star < 1.11 + \frac{1}{2}\ln(16) < 2.5 \tag{8.7}$$

The reason for this combination of cuts instead of a direct restriction of y_1 and y_2 is that these cuts conserve the shape of the χ_{dijet} distribution. While the χ_{dijet} cut restricts the distribution to the range 1 to 16, the y_{boost} cut is an orthogonal cut which does not change the distribution. The cut values of 16 and 1.11 were chosen in such a way that one has a large enough χ_{dijet} range to discriminate the angular distributions of QCD and new physics signals and also that one has a weak enough cut on y_{boost} in order to not cut away any potential new physics events appearing at high dijet masses.

The above described selection is visualized in Fig. 8.1 using simulated dijet events. The correlation of the two jet rapidities after the selection in the lowest and highest mass bins are plotted in (a) and (c). The diagonal borders with positive slope correspond to the restriction to $\chi_{\text{dijet}} < 16$. The diagonal borders with negative slope correspond to the $|y_{\text{boost}}| < 1.11$ cut. While the y_{boost} cut rejects events in the lowest mass bin, no events are rejected in the highest mass bin, as the dijet events at high masses have a smaller y_{boost}. This is due to the fact that the parton distribution functions favor events of balanced $x_1 \approx x_2$ of the two partons at high masses.

Figures 8.1 (b) and (d) show $p_{T,2}$ as a function of χ_{dijet} for the lowest and highest mass bins. It demonstrates the distribution of jet transverse momenta in bins of M_{jj}. The lowest transverse momentum jets selected in this analysis have $p_{T,2} > 30$ GeV. The highest transverse momentum jets selected in this analysis have $p_{T,2} > 1$ TeV.

As an example of a reconstructed dijet event, the one with the highest reconstructed dijet mass selected in this analysis is shown in Fig. 8.2. The energy collected in the ECAL (HCAL) is represented by red (blue) bars, reconstructed tracks with $p_T > 2$ GeV by green lines and reconstructed muons by red lines. Three jets with $p_T > 30$ GeV are reconstructed and their corrected transverse momentum is represented by purple bars. In Fig. 8.2 (a) one can see that the event is perfectly back-to-back in ϕ which is expected from momentum conservation in dijet production.

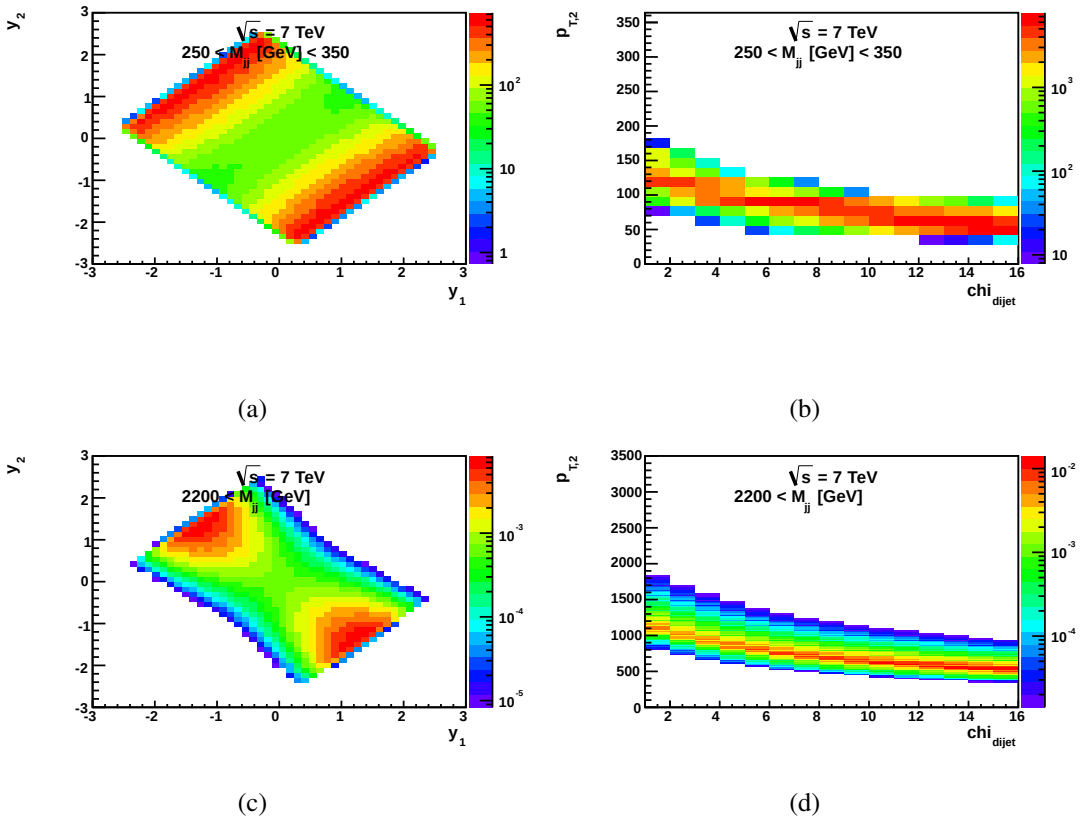


Figure 8.1: Correlation of y_1, y_2 and $p_{T,2}$, χ_{dijet} in simulated dijet events.

8.3 Triggers

The trigger paths used in this analysis are the inclusive single jet triggers named HLT_Jet*U. They require a Level-1 trigger jet and an HLT jet with certain thresholds of transverse momentum (e.g. HLT_Jet30U requires p_T >20 GeV at L1 and p_T >30 GeV at HLT). The jet energy corrections described in Sec. 7.4 are not applied at trigger level, and therefore the given trigger thresholds translate into higher thresholds for analysis level jets.

Table 8.1 summarizes the triggers used in this analysis. The indicated effective integrated luminosities for each trigger differ because the lower threshold triggers were pre-scaled during the later running periods in 2010 when the LHC delivered higher instantaneous luminosities. In later runs, slightly modified versions of the triggers with optimized matching of Level-1 and HLT requirements were introduced (e.g. *_v3). It was made sure that the later versions have the same efficiencies as the old ones.

The efficiencies of the triggers have been evaluated using lower p_T threshold triggers as a reference (e.g. Jet15U as reference for Jet30U). The trigger efficiency is given by the number of events triggered by the trigger of interest divided by the number of events triggered by the reference trigger, scaled by their effective integrated luminosities:

$$\text{trigger efficiency} = \frac{N^{\text{trigger of interest}}_{\text{selected events}}}{N^{\text{reference trigger}}_{\text{selected events}}} \frac{L^{\text{reference trigger}}}{L^{\text{trigger of interest}}} \tag{8.8}$$

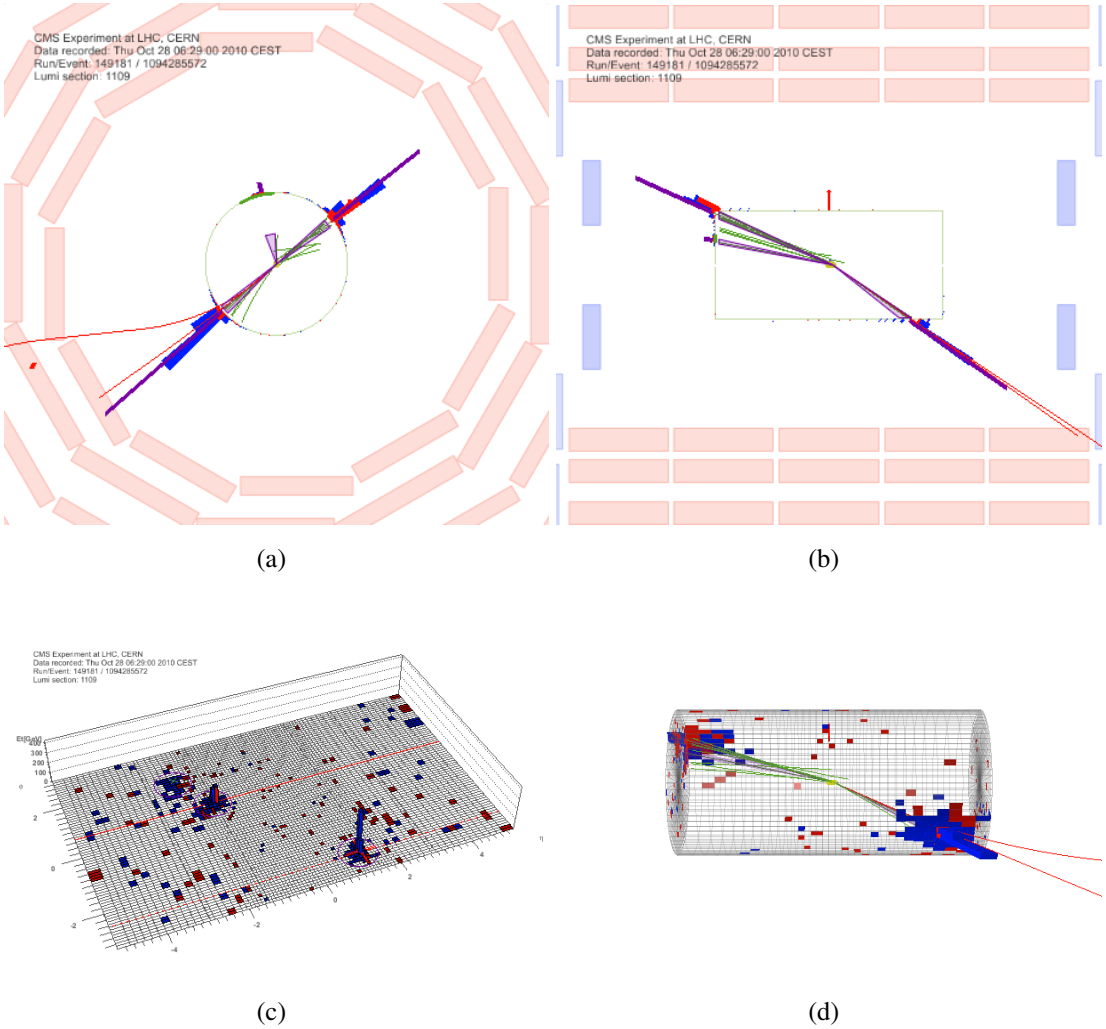


Figure 8.2: Event display of the highest dijet mass event selected in this analysis. (a) shows the ρ-ϕ-projection where the event is viewed from along the beam-axis z. (b) shows the ρ-z-projection of the event where the beam-axis crosses the center of the display horizontally. (c) demonstrates the transverse energy distribution of the event in bins of ϕ and η. (d) shows a 3D-view of the event where the cylinder corresponds to the inner calorimeter surface.

Table 8.1: Triggers used for the measurement of different mass bins and their corresponding effective integrated luminosity.

Mass bin [GeV]	HLT trigger	L [pb^{-1}]
250-350	Jet30U, Jet30U_v3	0.352
350-500	Jet50U, Jet50U_v3	3.50
500-650	Jet50U, Jet70U, Jet70U_v2, Jet70U_v3	9.17
650-850	Jet50U, Jet100U, Jet100U_v2, Jet100U_v3	19.8
850-1100	Jet50U, Jet100U, Jet140U_v1, Jet140U_v3	36.1
1100-1400	Jet50U, Jet100U, Jet140U_v1, Jet140U_v3	36.1
1400-1800	Jet50U, Jet100U, Jet140U_v1, Jet140U_v3	36.1
1800-2200	Jet50U, Jet100U, Jet140U_v1, Jet140U_v3	36.1
2200-∞	Jet50U, Jet100U, Jet140U_v1, Jet140U_v3	36.1

This method is approximate in the sense that the efficiency can statistically fluctuate above one when the reference trigger is pre-scaled. To verify the results, the efficiency was crosschecked in running periods where both the trigger of interest and the reference trigger were not pre-scaled and found to agree.

Figure 8.3 shows the efficiencies of the triggers used in this analysis as a function of dijet mass. They show a very similar behavior and therefore the highest p_T threshold trigger is discussed representatively in the following. Above a certain threshold in invariant mass, 830 GeV for Jet140U, the trigger is fully efficient. This is due to the fact that the single jet triggers are designed to be fully efficient above a certain transverse momentum of the leading jet in the event, and, as we have seen in Figure 8.1, a threshold in dijet mass can be directly translated into a threshold on the transverse momentum of the leading jet. Below this threshold the trigger becomes inefficient, showing an intermediate step before reaching zero efficiency at low masses. The step is a feature which comes from the fact that the trigger efficiency varies as a function of pseudo-rapidity which is connected to the dijet mass via $M_{jj}^2 = 2p_{T_1}p_{T_2}cosh(\eta_1 - \eta_2)$.

In order to use as much data statistics as possible in the analysis, the first 5 mass bin borders in Tab. 8.1 are chosen according to the trigger efficiencies. The mass bin borders are chosen in such a way that the trigger efficiency is greater than 99%. This is required in order to avoid any impact of the trigger efficiency and its description in simulation on the analysis. The trigger turn-on points where the triggers are 99% efficient are determined by fitting an error function to the trigger efficiency as a function of dijet mass.

8.4 Event Cleaning

In addition to the kinematic selection cuts for dijet events, described in Sec. 8.2, events are required to fulfill certain quality requirements in order to reduce background from noise and non-collision events.

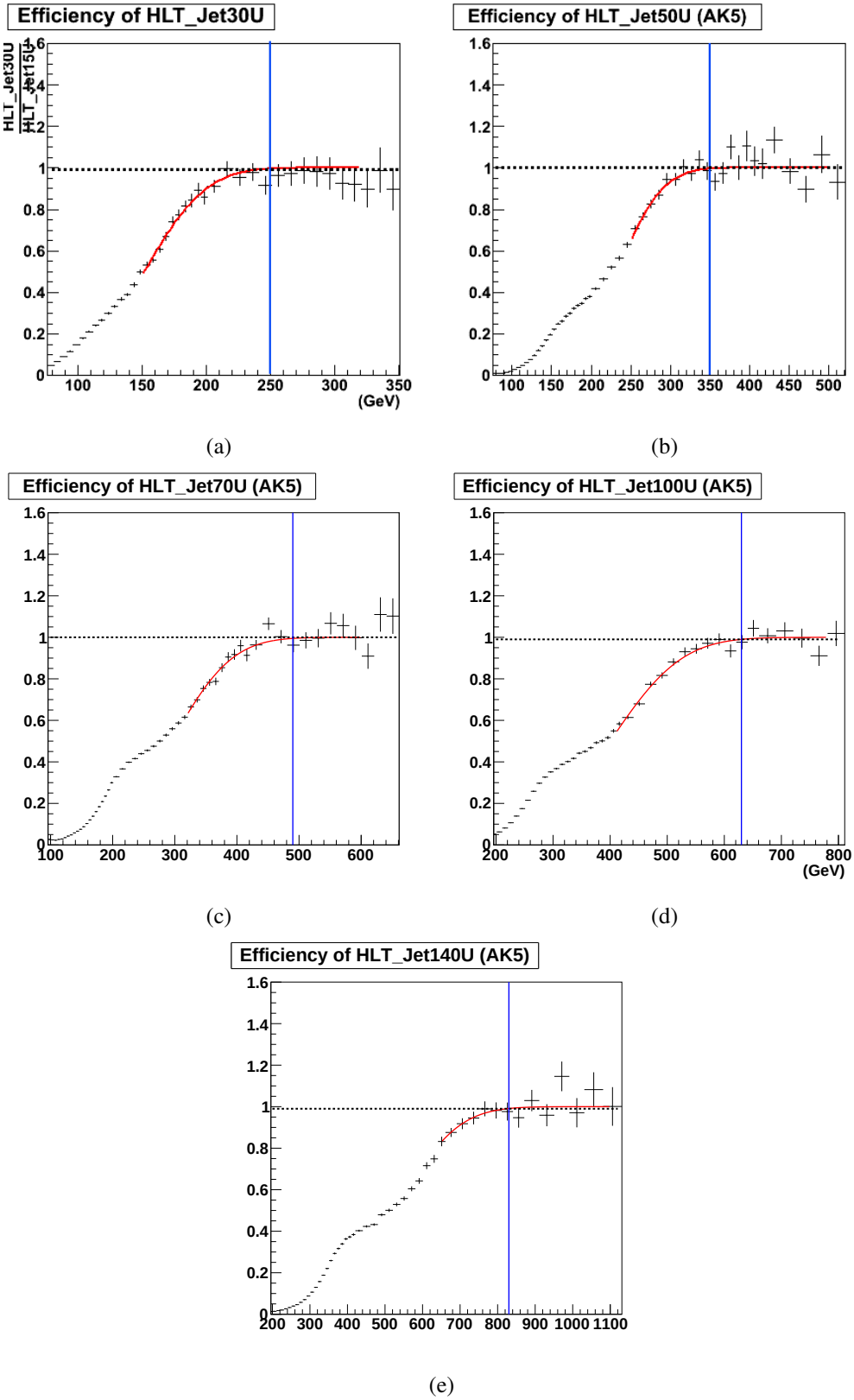


Figure 8.3: Trigger efficiencies as a function of dijet mass.

A non-collision background observed in the early running of the LHC in 2009 are the so-called beam scraping events [35]. These are events with an interaction of the beam with material

in the beam pipe, resulting in particles passing the pixel detector in parallel to the beam pipe which penetrate a large number of pixels, reconstructed as fake tracks. These events are filtered by cutting on the fraction of high-purity tracks [78], which fulfill certain quality requirements designed to reject fake tracks, with respect to the total number of tracks. In this analysis, the fraction is required to be at least 25% when the event has at least 10 tracks.

To protect the analysis against background from collisions happening outside the volume of the magnet, a primary vertex [35] is required to be reconstructed from at least 5 tracks ($n_{\mathrm{dof}} > 4$) with a maximum distance $|z_{\mathrm{PV}}| < 24$ cm along the beam-axis from the detector center and a maximum distance $\rho_{\mathrm{PV}} = \sqrt{x_{\mathrm{PV}}^2 + y_{\mathrm{PV}}^2} < 2$ cm orthogonal to the beam-axis.

Finally events containing fake jets reconstructed from detector noise are rejected by requiring the two leading jets in the event to pass the "loose" jet ID as described in Sec. 7.3.

Table 8.2: Number of data events in each mass bins and the fraction of events that pass the scraping event filter, the vertex requirements and the jet ID requirements with all other cuts applied. The corresponding efficiencies for dijet events evaluated from simulation are given in brackets.

Mass bin [GeV]	# events	scraping filter (MC) [%]	vertex (MC) [%]	jet ID (MC) [%]
250-350	65735	100.00 (100.00)	99.89 (99.98)	99.89 (99.96)
350-500	157549	100.00 (100.00)	99.92 (99.98)	99.91 (99.97)
500-650	75379	100.00 (100.00)	99.83 (99.97)	99.90 (99.97)
650-850	47876	100.00 (100.00)	99.86 (99.97)	99.90 (99.97)
850-1100	22147	100.00 (100.00)	99.82 (99.96)	99.79 (99.97)
1100-1400	5216	100.00 (100.00)	99.94 (99.95)	99.48 (99.98)
1400-1800	1174	100.00 (100.00)	99.8 (99.95)	95.8 (99.98)
1800-2200	180	100.0 (100.00)	100.0 (99.93)	94.7 (99.98)
2200-∞	48	100 (100.00)	100 (99.91)	79 (99.97)

The number of data events in each mass bin after applying all cuts is summarized in Tab. 8.2. The fraction of events that pass the scraping event filter, the vertex requirements and the jet ID requirements with all other cuts applied is given, together with the corresponding efficiencies for dijet events evaluated from simulation.

The scraping event filter rejects no selected dijet event as expected from simulation since the beam scraping events are already rejected by the tight kinematic selection of dijet events.

The vertex requirements rejects 0.1-0.2% of the events in all mass bins. The efficiency for dijet events evaluated from simulation, which is known to describe the efficiency in data very well [35], is higher than 99.9%. Therefore, the vertex requirements are found to reject a small contamination by non-central collisions of 0-0.1%.

The jet ID requirements reject less than 1% of the events for $M_{jj} < 1400$ GeV. Above $M_{jj} > 1.4$ TeV a high fraction of noise events is rejected. These events have been identified to be due to single instances of RBX/HPD noise which are being reconstructed as two jets at

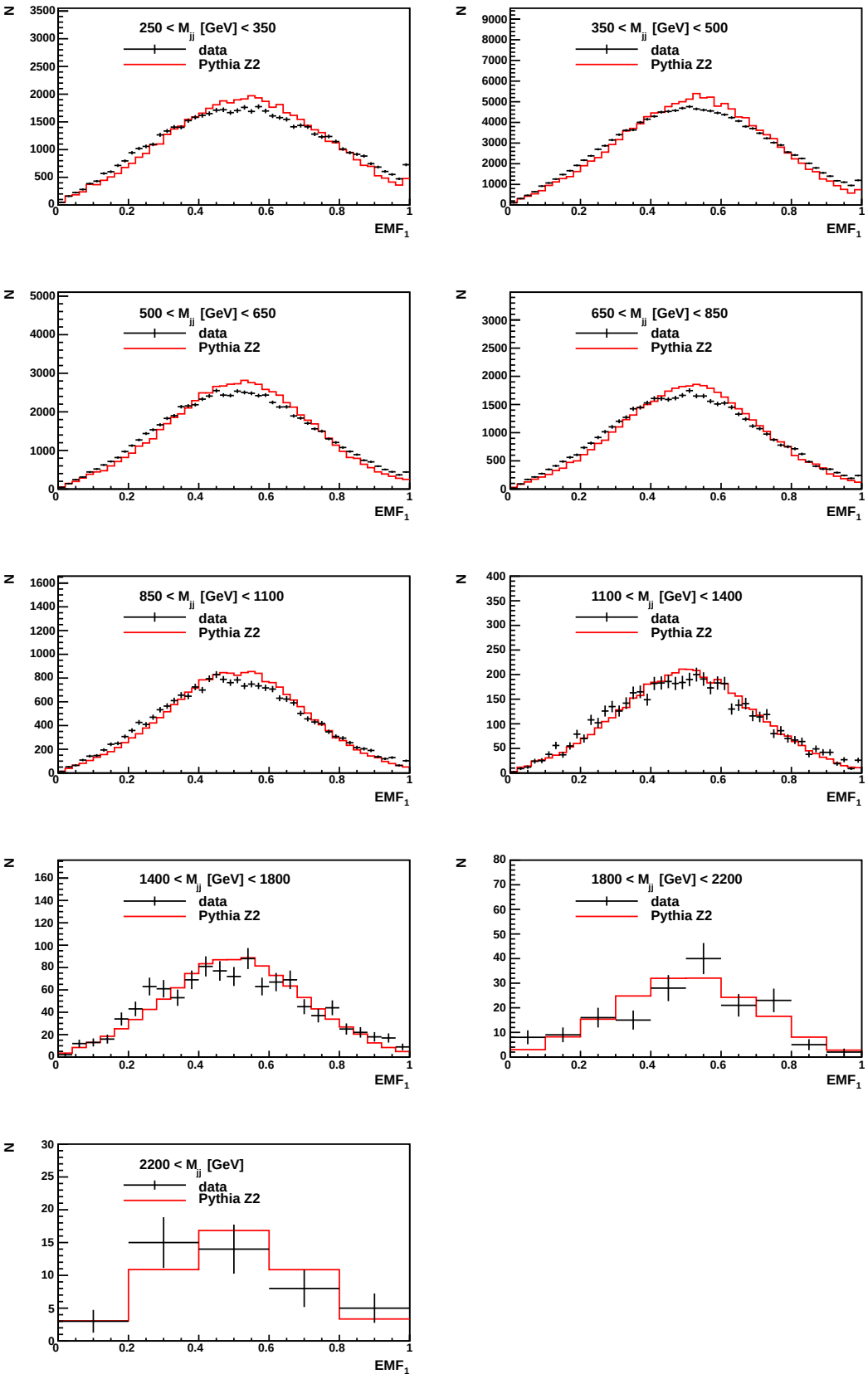


Figure 8.4: Electromagnetic energy fraction (EMF) of the leading jet in data and simulation.

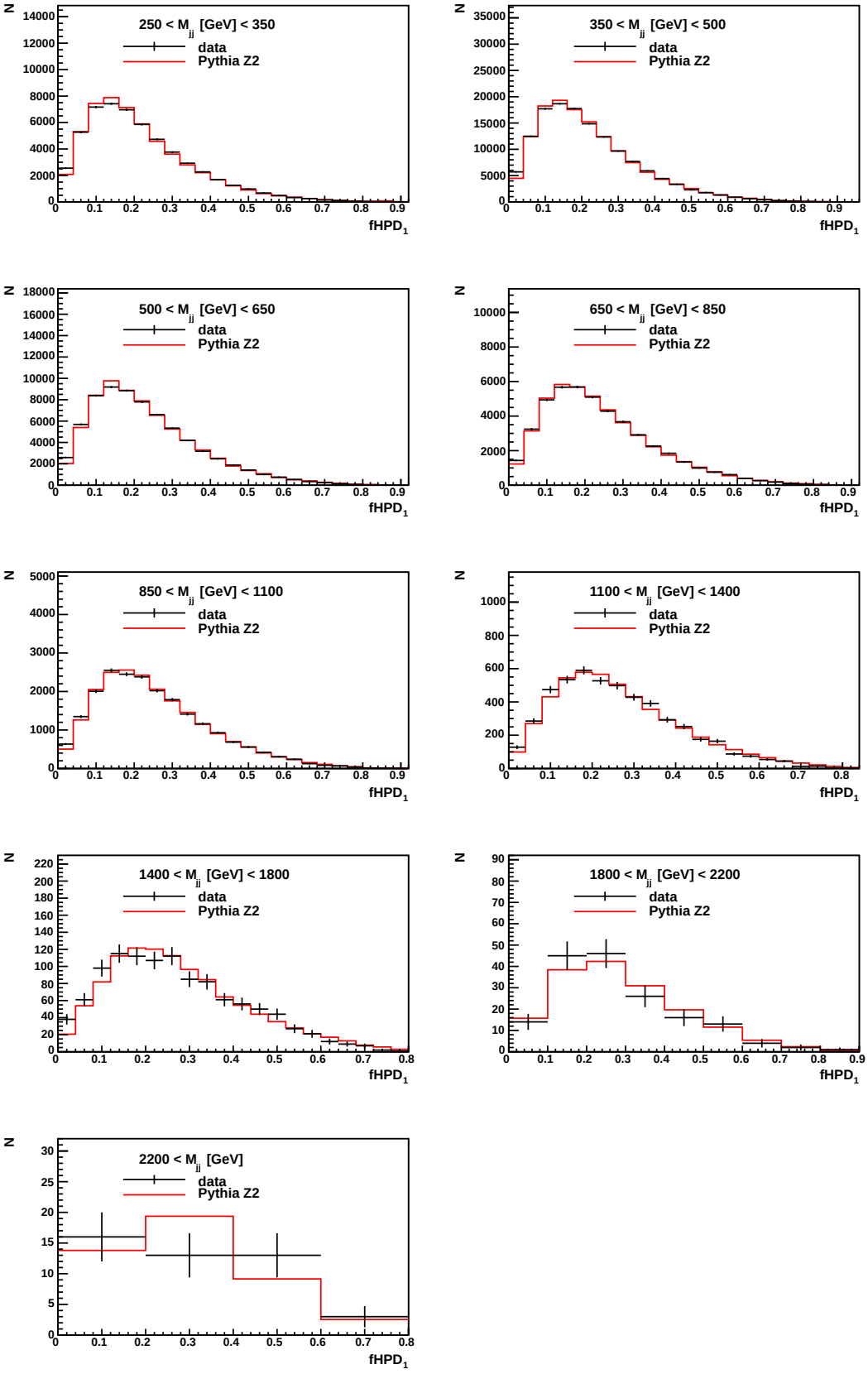


Figure 8.5: Fraction of energy contributed by the hottest (highest energy) HPD (f_{HPD}) of the leading jet in data and simulation.

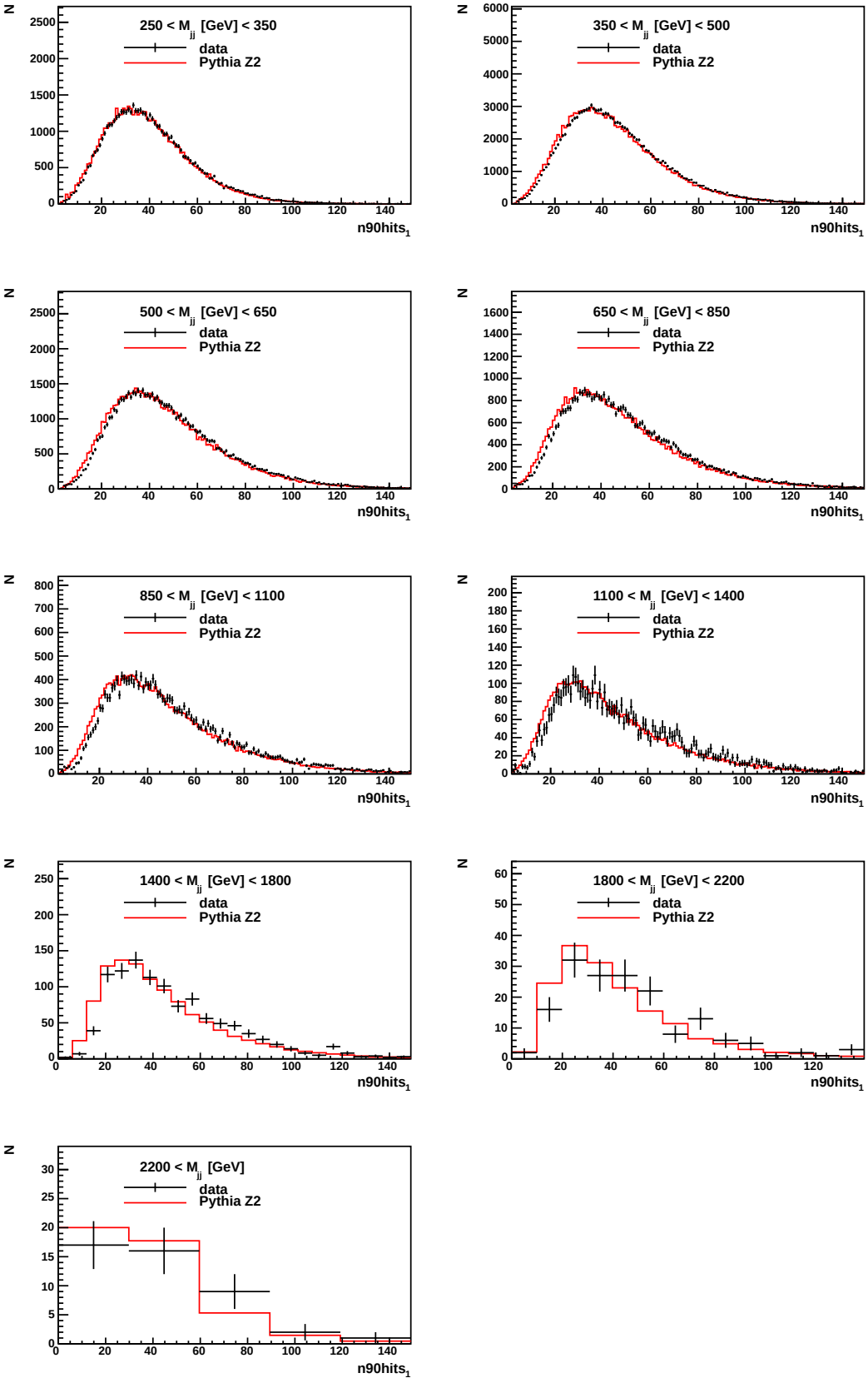


Figure 8.6: Number of hits which contain 90% of the jet energy (n^{90}_{hits}) of the leading jet in data and simulation.

the same ϕ. A typical event display of such an event is shown in Fig. 7.3 (a). They are typically rejected by both the f_{HPD} cut and the EMF cut.

From simulation the efficiency for real dijet events to pass the jet ID requirements is found to be higher than 99.9%. In order to make sure that this efficiency also holds on data, the jet ID variables are compared in data and events, simulated using the full detector simulation. Figures 8.4, 8.5 and 8.6 show the EMF, f_{HPD} and n^{90}_{hits} of the leading jet. The data distributions with all selection cuts applied are compared to simulated events. The description is reasonable in all mass bins and no unexpected artifacts are observed. The largest discrepancy is observed in the n^{90}_{hits} distribution at high dijet masses, which tends to be larger in data than in simulation. For this analysis the description in the region of the jet ID cuts is important. As no significant data excess is observed in the region around EMF ≈ 0.01, $f_{\mathrm{HPD}} \approx 0.98$, $n^{90}_{\mathrm{hits}} \approx 1$, the high jet ID efficiency in simulation is expected to hold for the data.

8.5 Kinematic Distributions in Data and Simulation

As this is the first dijet analysis in CMS using the full 2010 dataset, it is particularly important to validate both the data and the simulation. This is done by investigating the most important dijet variables both in data and simulation and comparing them to each other. The full event selection including the trigger requirements, cleaning cuts and kinematic selection is applied to the data, while only the kinematic selection is applied to the simulation. Trigger requirements are not applied to the simulation as the trigger efficiency has been measured from data to be >99% and it is difficult to model in simulation. The event cleaning cuts have negligible impact on the simulated events as seen in the previous section and are therefore not applied on the simulation as well.

Three different simulation samples are probed, which will be used to study systematic effects in the simulation, explained in the next chapter. The samples use a simplified simulation of detector effects and contain $\sim 10^8$ events. Particle jets are clustered from the event generator particles and are then smeared in energy and angle according to a parametrization of the jet resolution, derived from MC truth [14]. Since jet resolutions are the leading detector effect impacting the kinematics of dijet events, this simplified version of the detector simulation is sufficient for this study. The generators and tunes used to produce the three dijet samples are: PYTHIA with tune Z2, PYTHIA with tune D6T and HERWIG++ (for details see Sec. 5.1).

As the invariant mass spectrum of dijets is steeply falling and computing resources are limited, it is important to optimize the distribution of the event statistics over the spectrum. Therefore, the simulated samples are generated in 20 bins of $\hat{\mathrm{p}}_{\mathrm{T}}$, the transverse momentum of the leading parton. The complete spectrum is reconstructed by weighing each $\hat{\mathrm{p}}_{\mathrm{T}}$-bin according to its cross section and number of events.

In Fig. 8.7, 8.9, 8.11 and 8.13 the p_T and y distributions are shown for both the leading and 2^{nd} leading jet in the event. The predictions from simulation are normalized to the number of data events in each mass bin and the shapes are compared. The shape of the p_T distribution is explained by the binning in dijet invariant mass which defines the steep rise at the low edge of

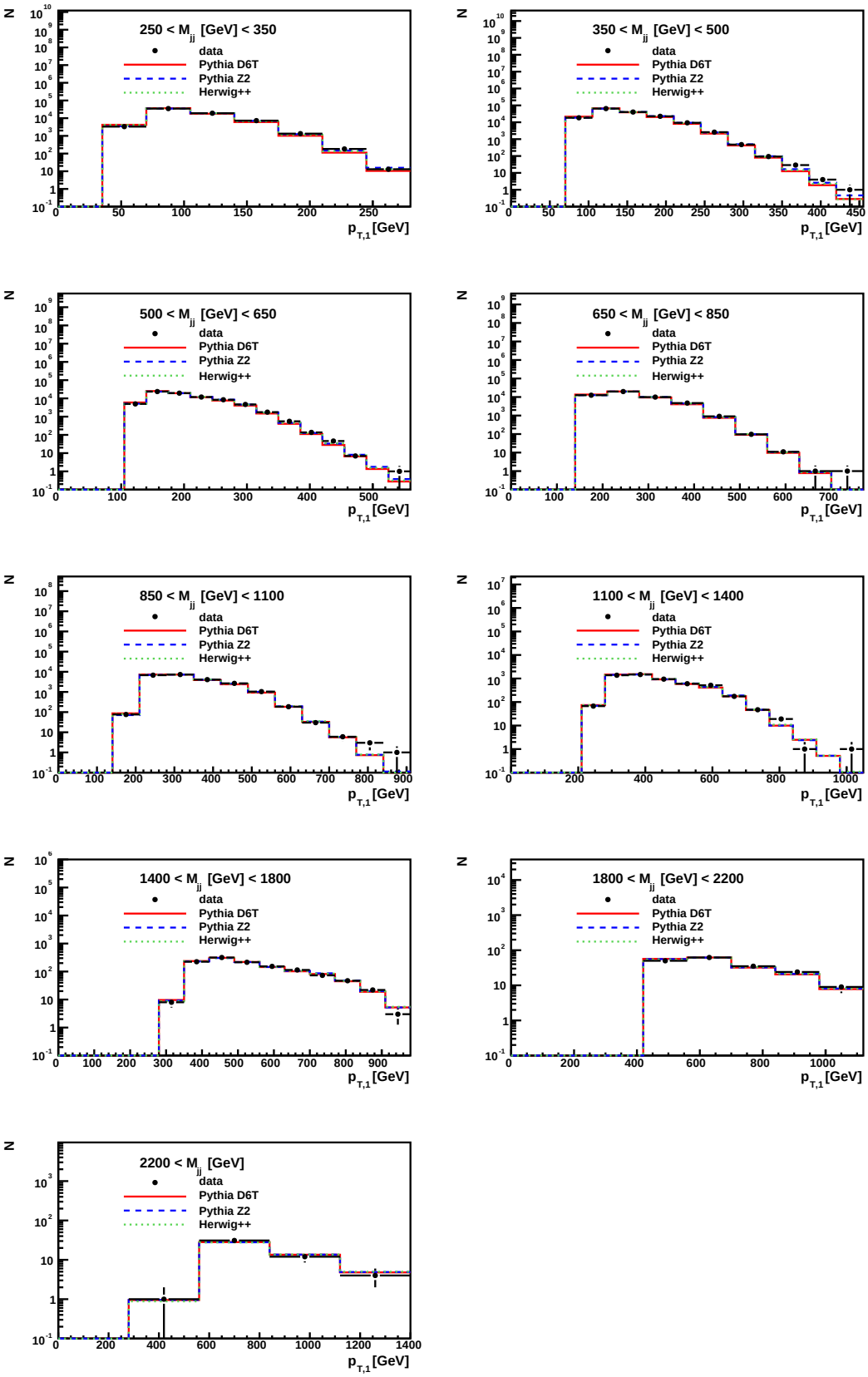


Figure 8.7: p_T of the leading jet in data and simulation.

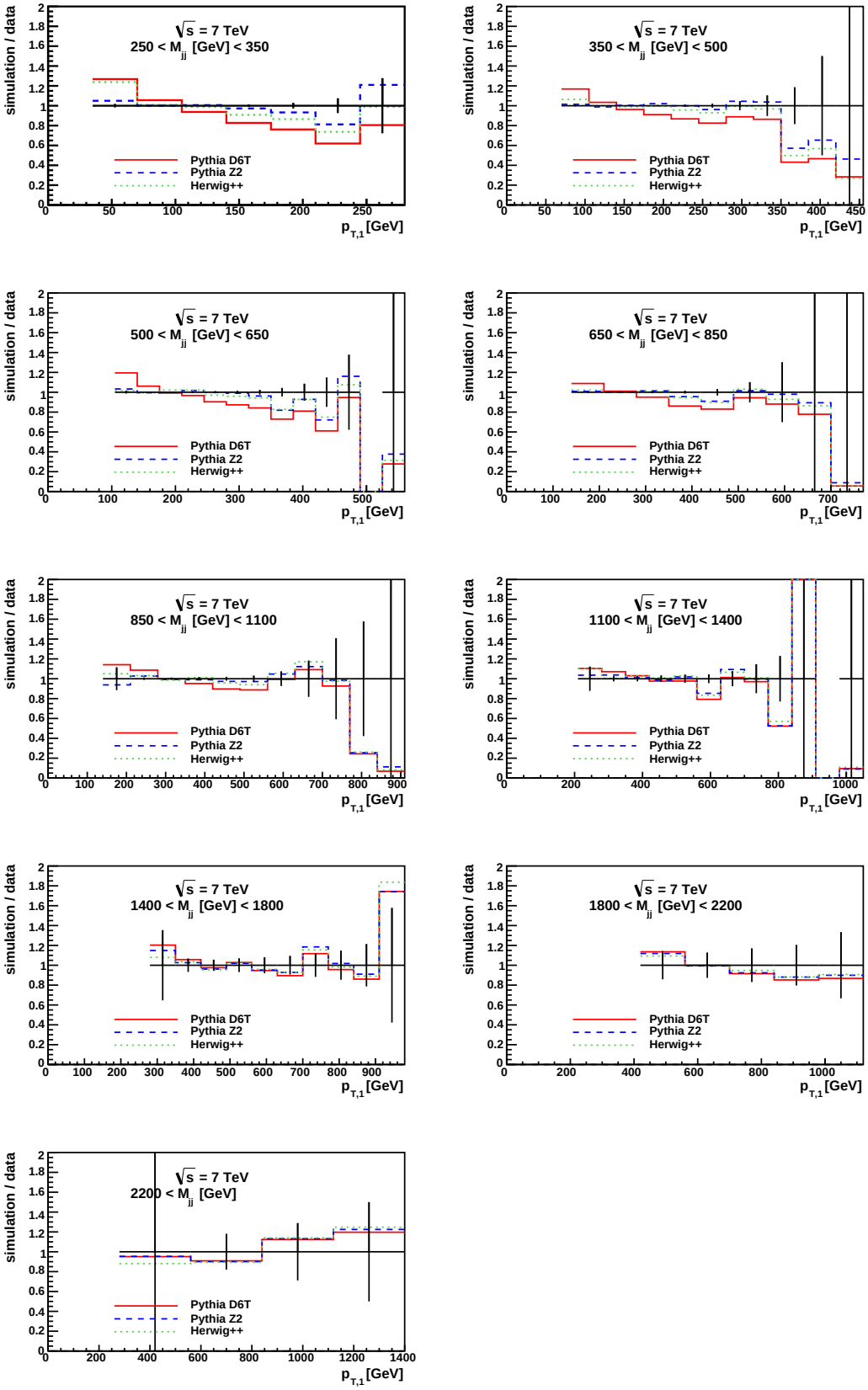


Figure 8.8: Ratio of simulation and data for the p_T of the leading jet.

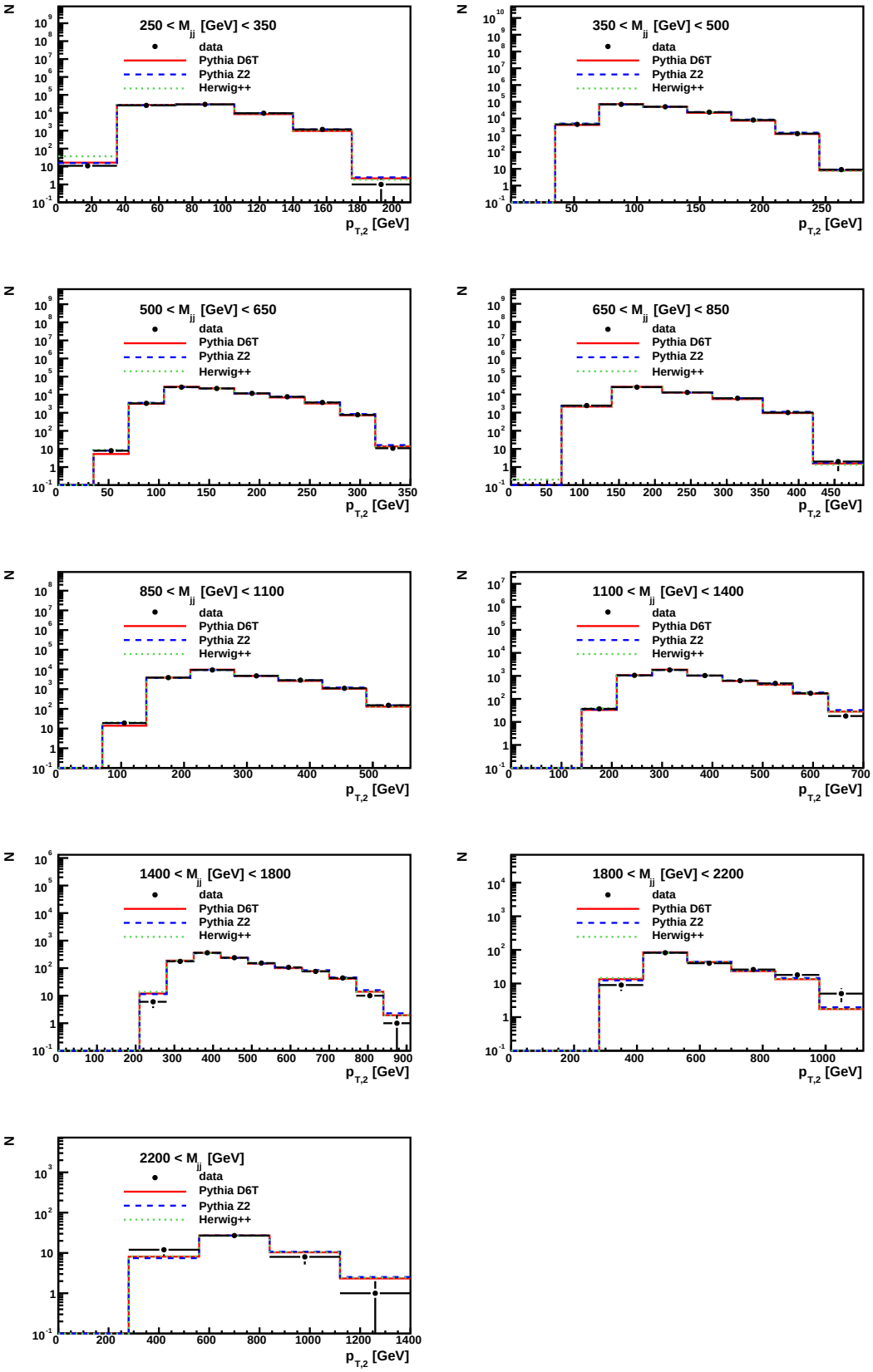


Figure 8.9: p_T of the 2^{nd} leading jet in data and simulation.

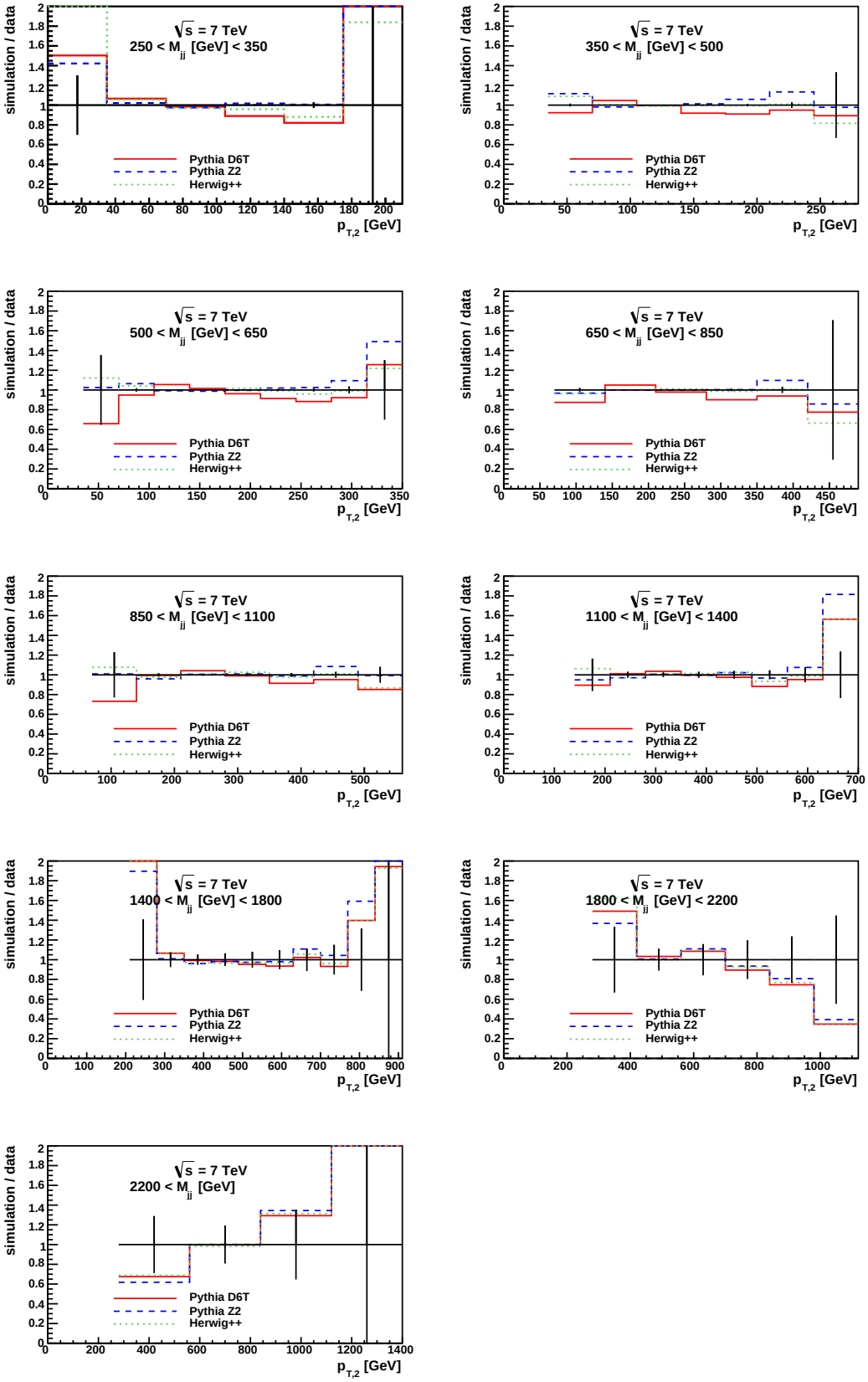


Figure 8.10: Ratio of simulation and data for the p_T of the 2^{nd} leading jet.

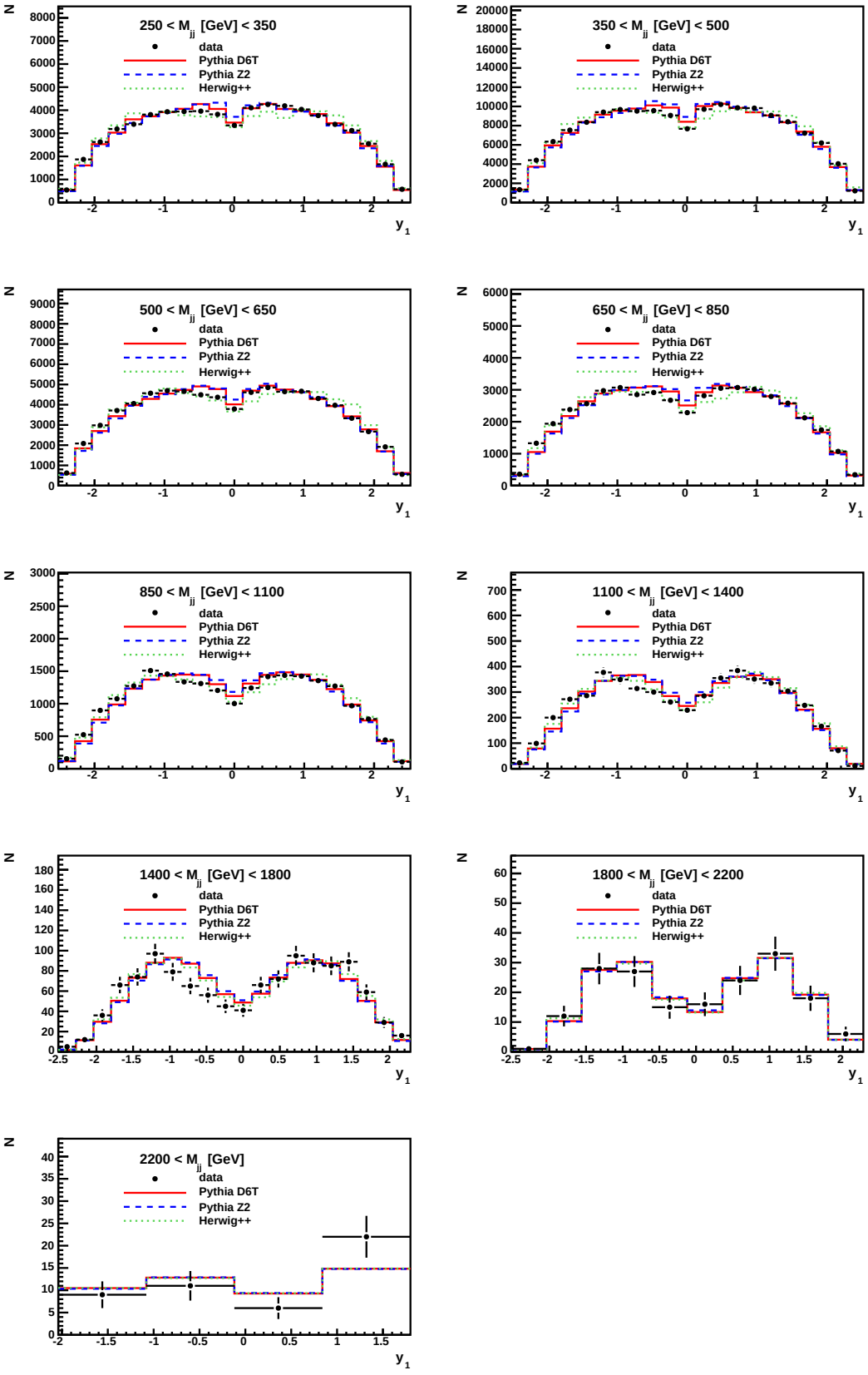


Figure 8.11: Rapidity of the leading jet in data and simulation.

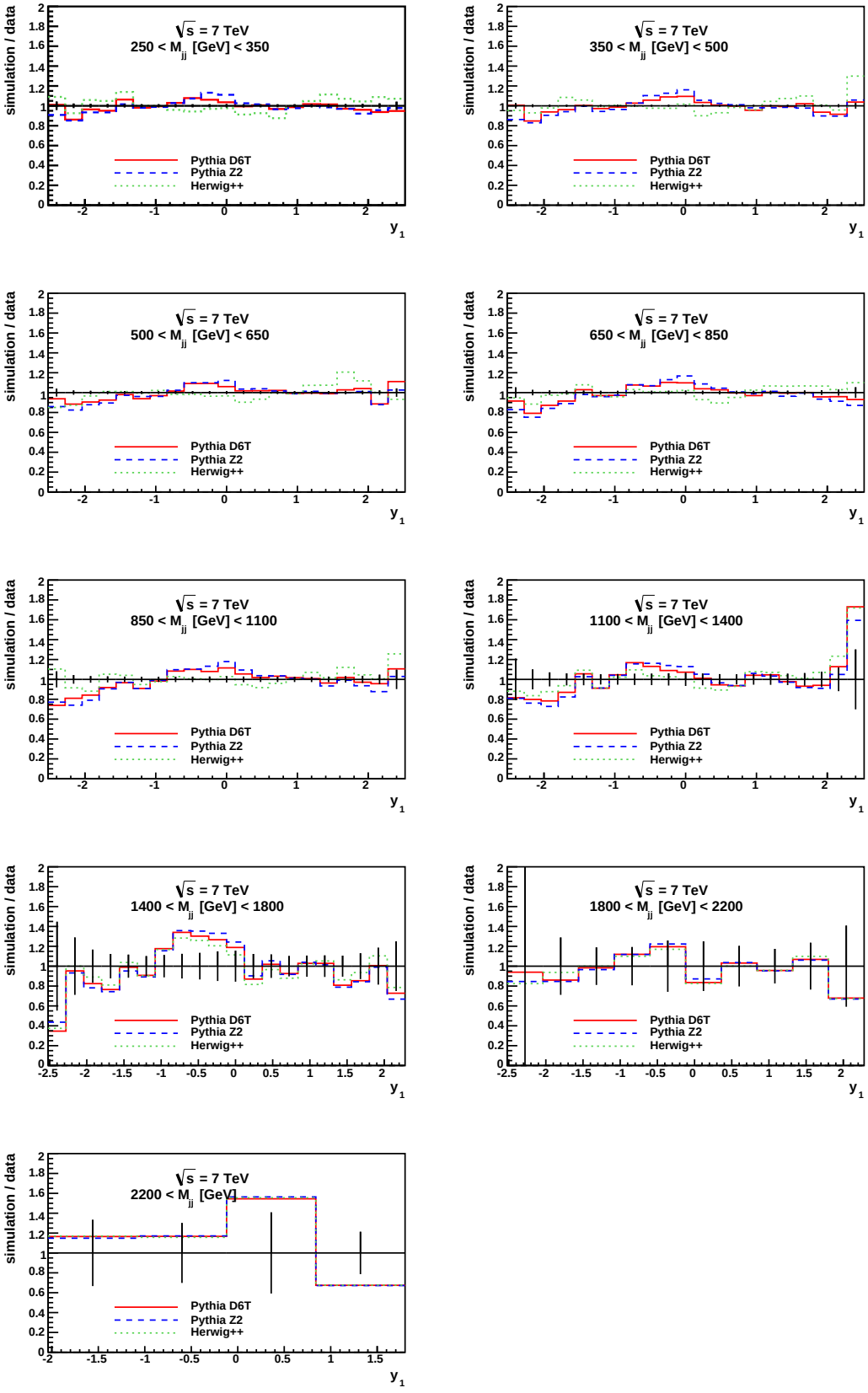


Figure 8.12: Ratio of simulation and data for the rapidity of the leading jet.

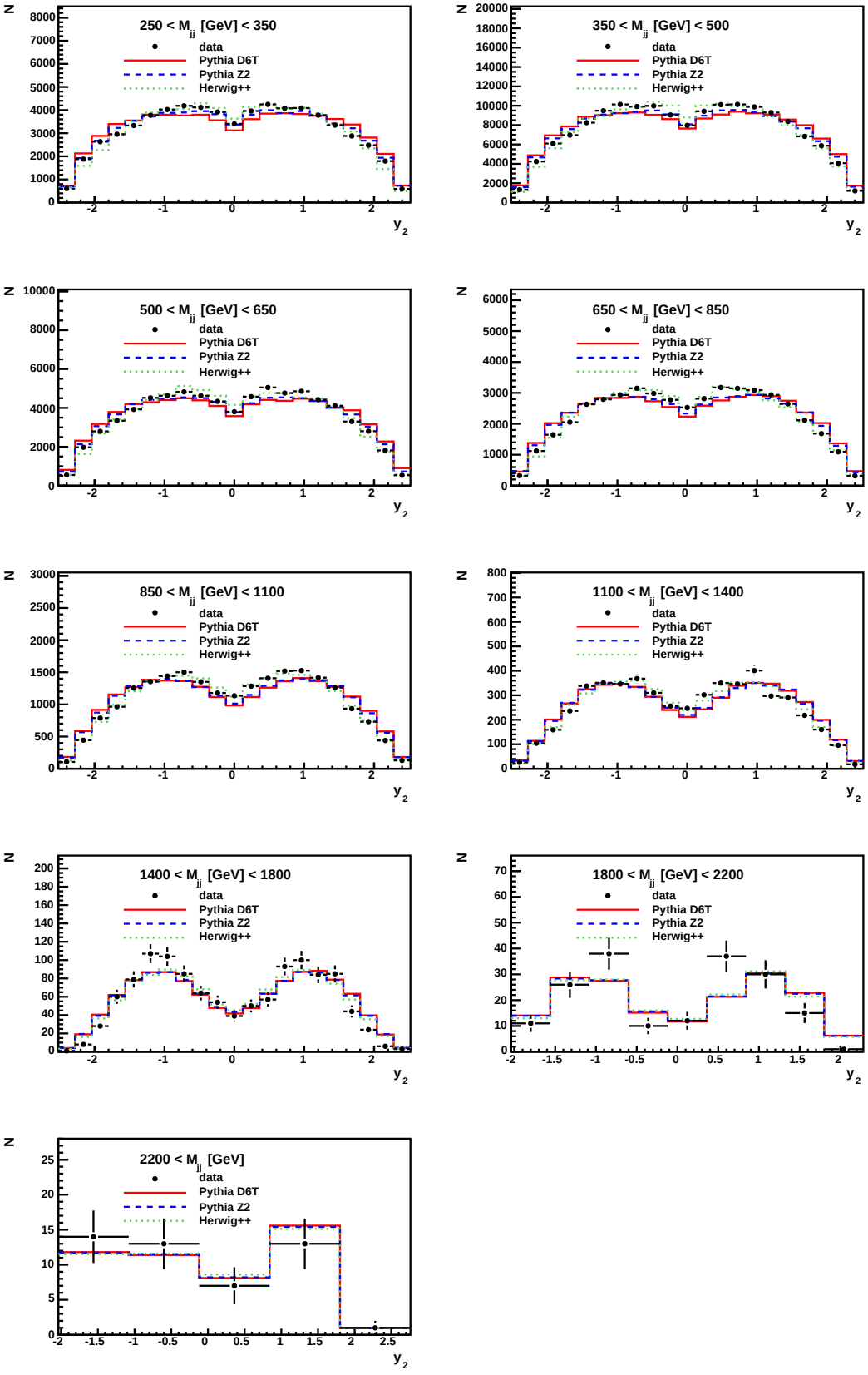


Figure 8.13: Rapidity of the 2^{nd} leading jet in data and simulation.

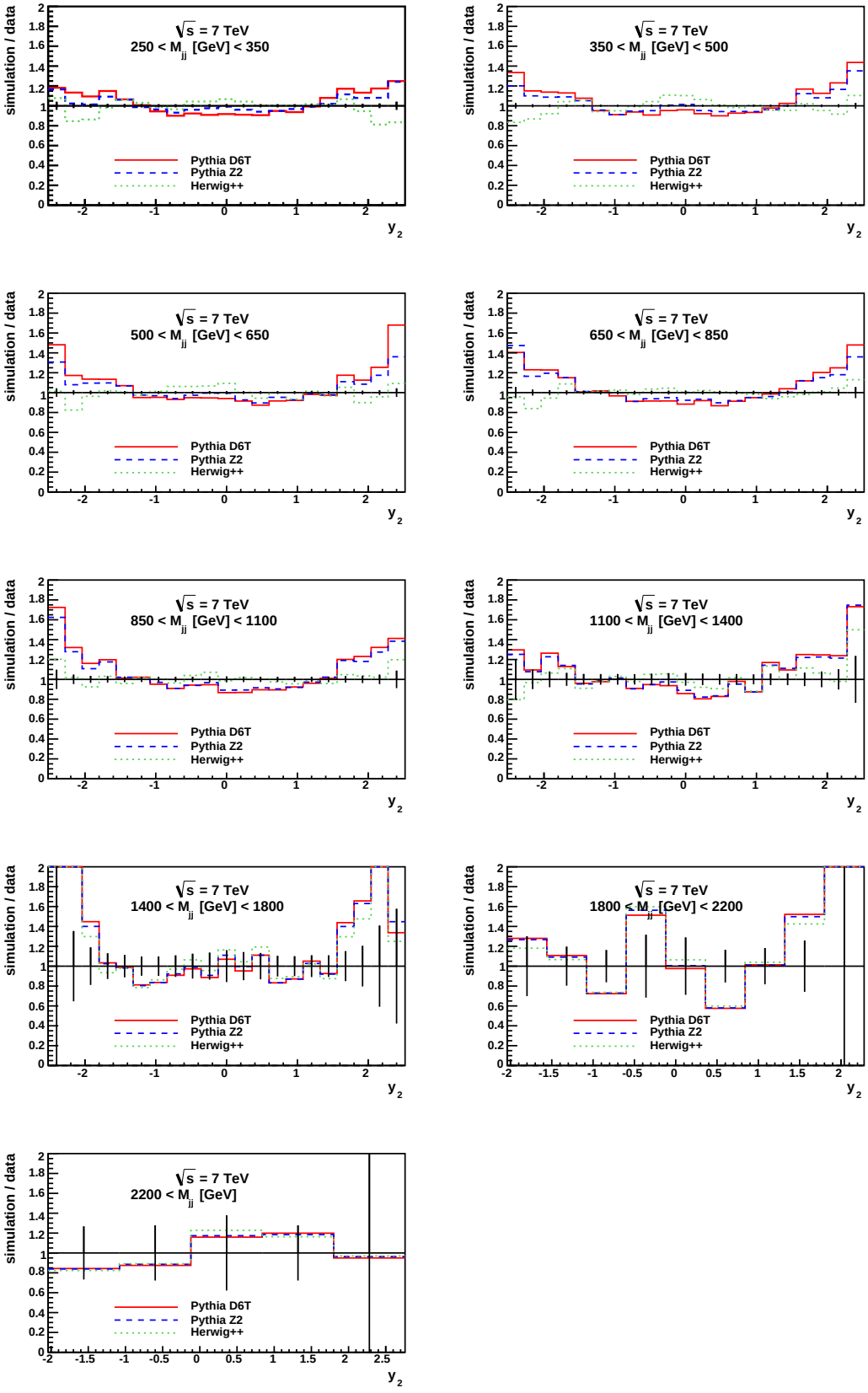


Figure 8.14: Ratio of simulation and data for the rapidity of the 2^{nd} leading jet.

the p_T spectrum. The y distributions are limited within $|y| < 2.5$ due to the kinematic selection applied. The dip at 0 in the y distribution also comes from the event selection as can be understood from Fig. 8.1 (a). With increasing dijet mass, the y distribution becomes narrower due to the fact that at high invariant masses the x values of the two partons need to be of the order of 0.1-1, and are therefore close to each other $x_1 \approx x_2$.

The ratio of simulation and data are shown in Fig. 8.8, 8.10, 8.12 and 8.14. All three simulation samples give a reasonable description of the p_T and y distributions in data. The differences between the generators and tunes are of the same order as the differences between simulation and data. Typically the difference between simulation and data are covered by the variation of the generators and tunes. This is important to note, as the variation of the generators will be used as a measure for the degree to which simulation is able to describe the data.

Figure 8.15 shows the dijet mass distributions for the data from all five jet triggers used in this analysis with a lower cut on the dijet mass which corresponds to the lowest mass bin border where data of a trigger was used. The agreement of data and simulation is remarkable as this distribution ranges over 5 orders of magnitudes. The ratio in Fig. 8.16 shows slight fluctuations of the order of 10% of simulation and data which is however within the expected statistical fluctuations.

The boost of the dijet system is shown in Fig. 8.17 and 8.18. It is restricted to $|y_{\text{boost}}| < 1.11$ by the event selection. From the distribution it becomes clear that this selection does remove a considerable amount of dijet events in the lower mass bins. However in the highest mass bins where new physics signals may appear, all events are kept. This variable is well described by all generators and tunes.

Figures 8.19 and 8.20 show another important property of dijet events. The variable $\sum p_T = \sqrt{(p_{x,1}+p_{x,2})^2+(p_{y,1}+p_{y,2})^2}$ is a measure for the momentum balance of the dijet system. For a pure dijet event without any additional radiation it is expected to be 0. In the case of a multi-parton process with more than two jets in the final state, the transverse momentum of the first two jets is not balanced and this variable therefore measures the transverse momentum sum of all additional particles. In addition to testing multi-jet production, this variable is sensitive to the jet energy resolution, as a mis-measurement of one of the energies of the two jets results in an effective transverse momentum imbalance of the event. From the lowest to the highest mass bin, the median of $\sum p_T$ increases in agreement with an expected raise of the absolute jet energy resolution. In the highest mass bin the median is slightly higher than expected, but within $\sim$1-2σ of the statistical uncertainties. As these distributions would point out problems in the jet energy reconstruction, the agreement gives confidence that the reconstruction algorithms are understood even at highest transverse momenta.

The χ_{dijet} distribution itself is compared in Figs. 8.21 and 8.22. Equally to the previous comparisons, the data are not corrected for detector effects, and the generator predictions include the simulation of the detector resolution. The simulated distributions are normalized to the number of data events. Interestingly the predictions from the generators and tunes vary significantly in this variable. PYTHIA with tune Z2 gives a very precise description of the data. HERWIG++ predicts less events at low χ_{dijet} and more at high χ_{dijet}. PYTHIA with tune D6T predicts the least events at low χ_{dijet} and the most at high χ_{dijet}. The differences are due to the different

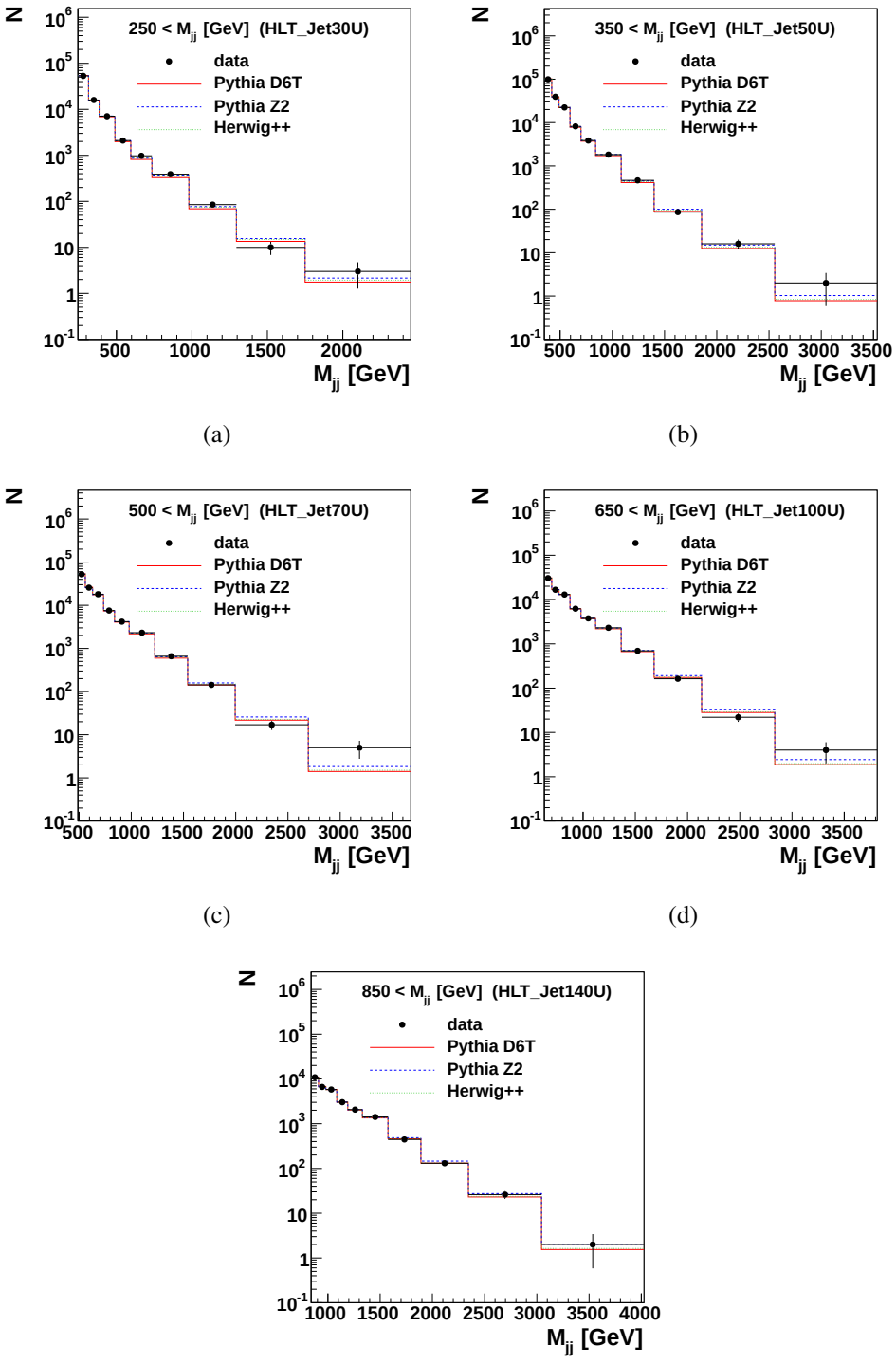
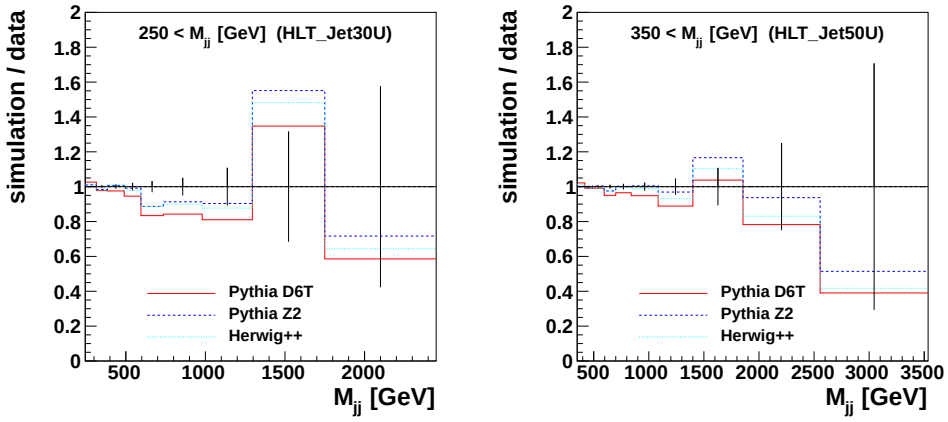
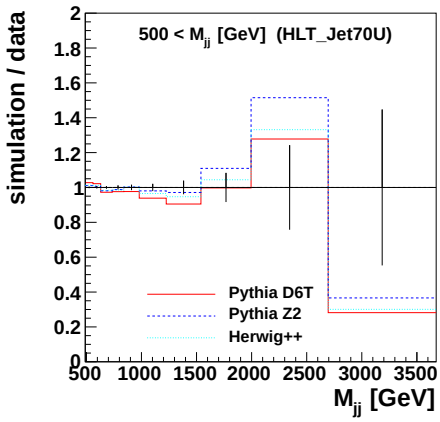


Figure 8.15: M_{jj} in data and simulation for M_{jj} greater than the first 5 mass bins borders. Note that each of these first five mass bins uses a different trigger with different effective integrated luminosity.

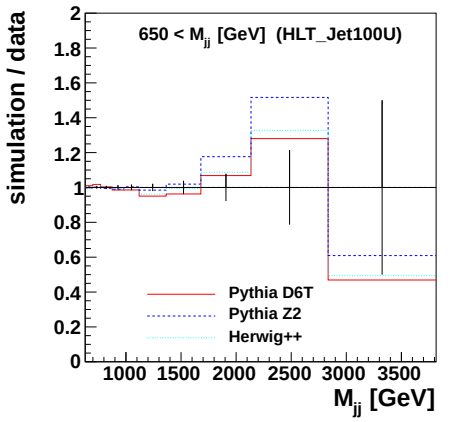


(a)

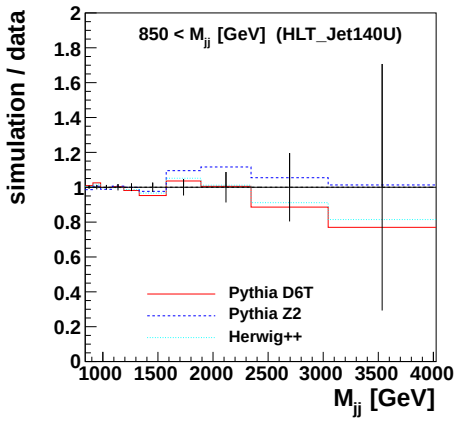
(b)



(c)



(d)



(e)

Figure 8.16: Ratio of simulation and data for M_{jj}. Note that each of these first five mass bins uses a different trigger with different effective integrated luminosity.

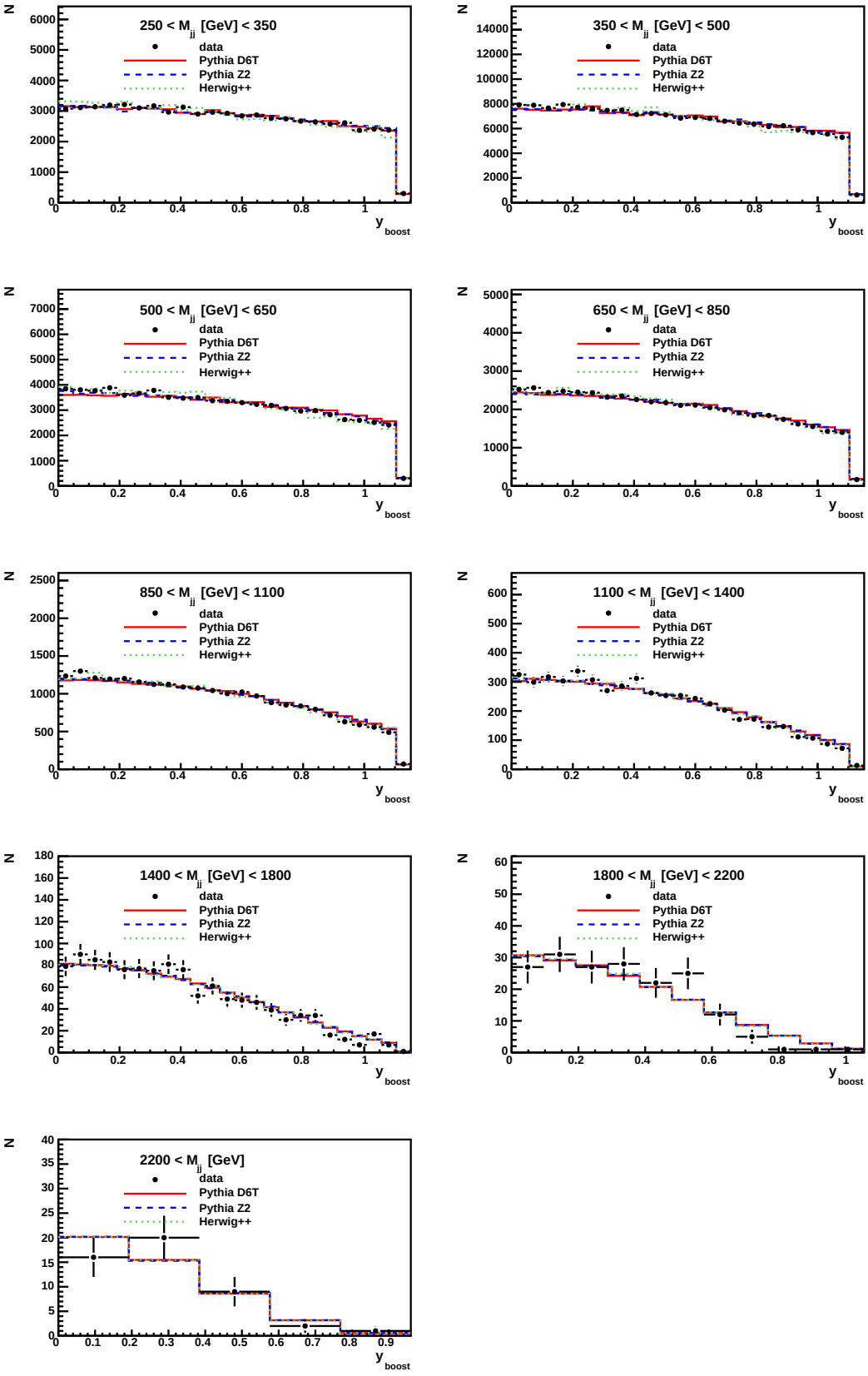


Figure 8.17: y_{boost} in data and simulation.

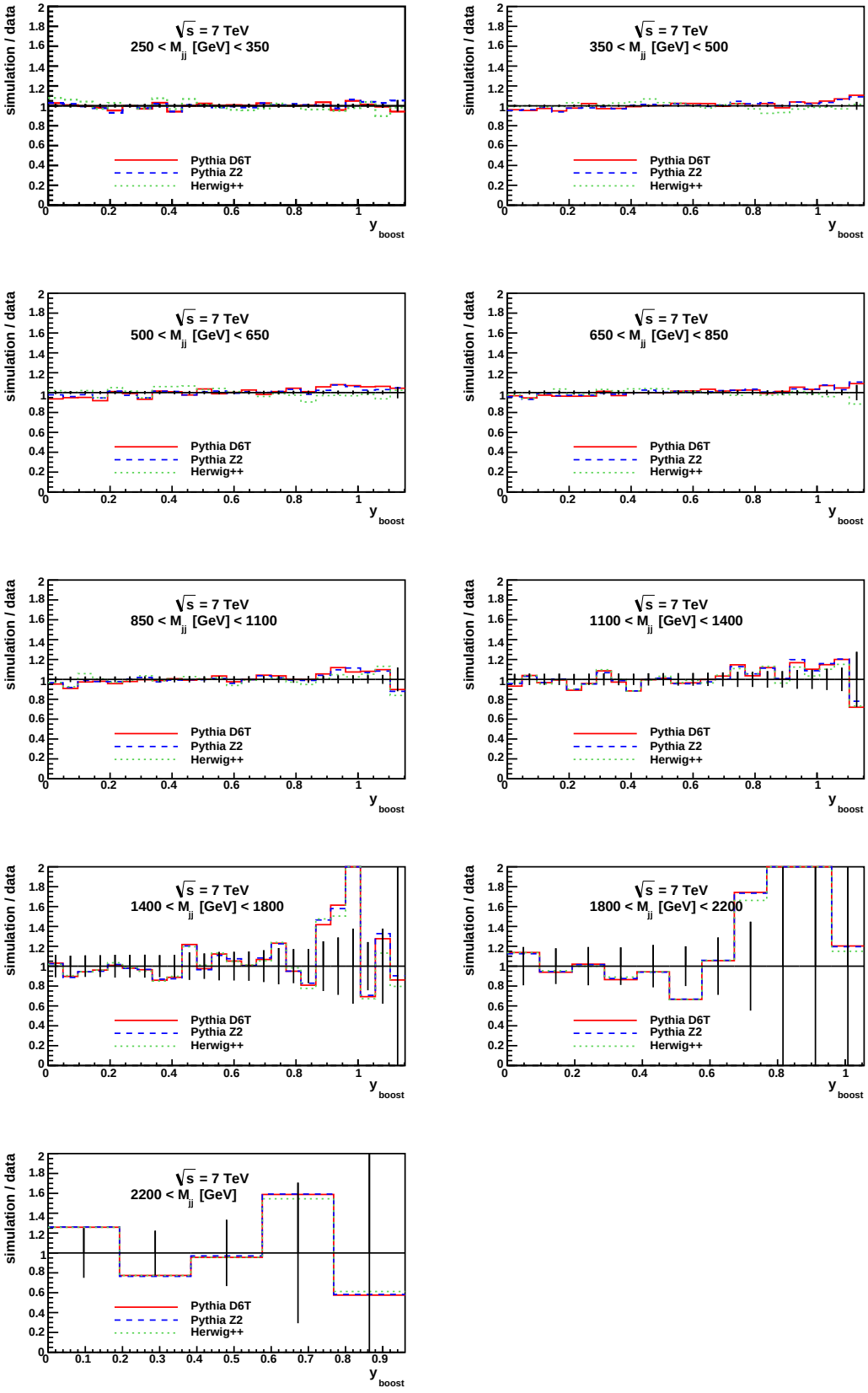


Figure 8.18: Ratio of simulation and data for y_{boost}.

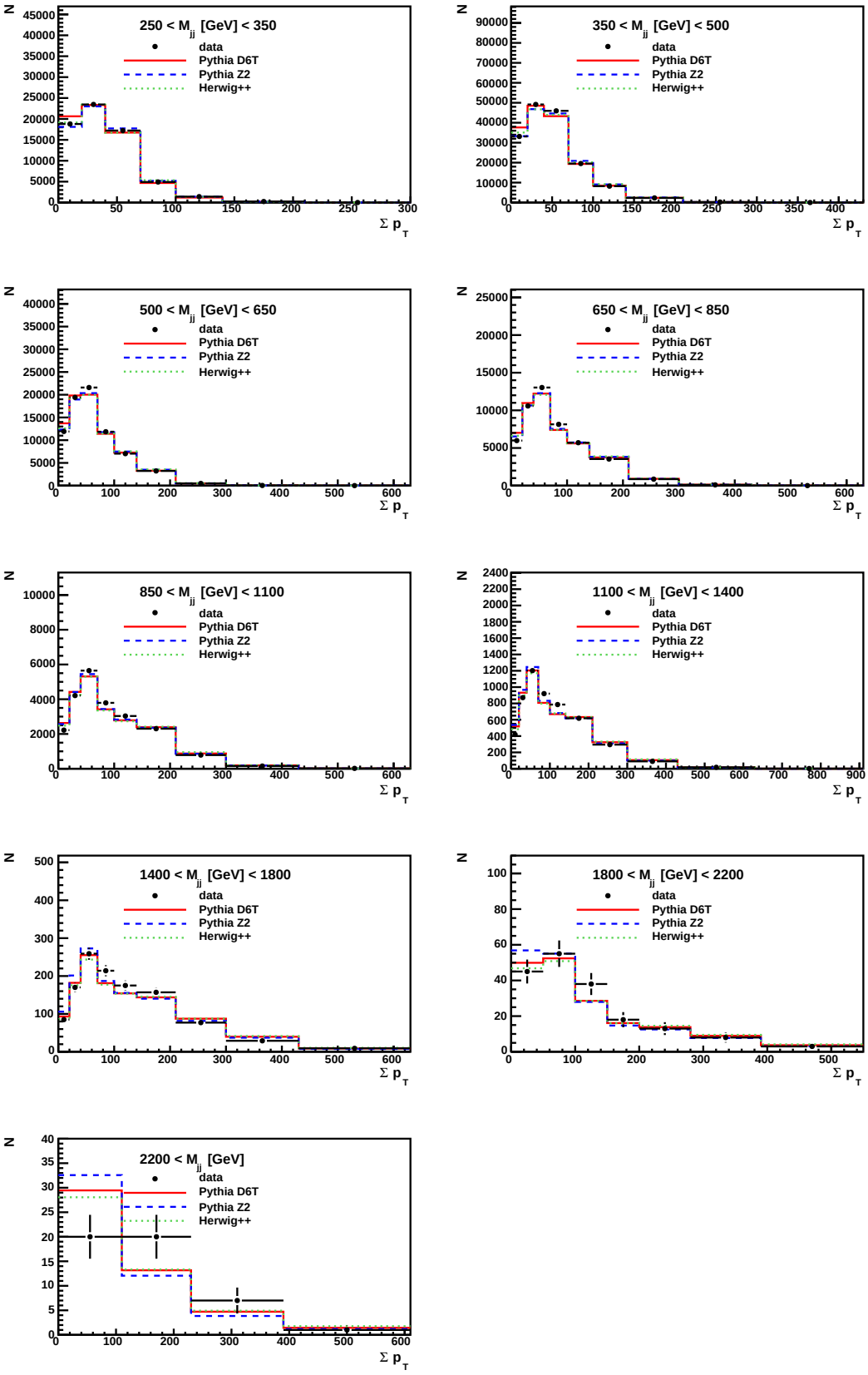


Figure 8.19: $\sum p_T = \sqrt{(p_{x,1}+p_{x,2})^2+(p_{y,1}+p_{y,2})^2}$ in data and simulation.

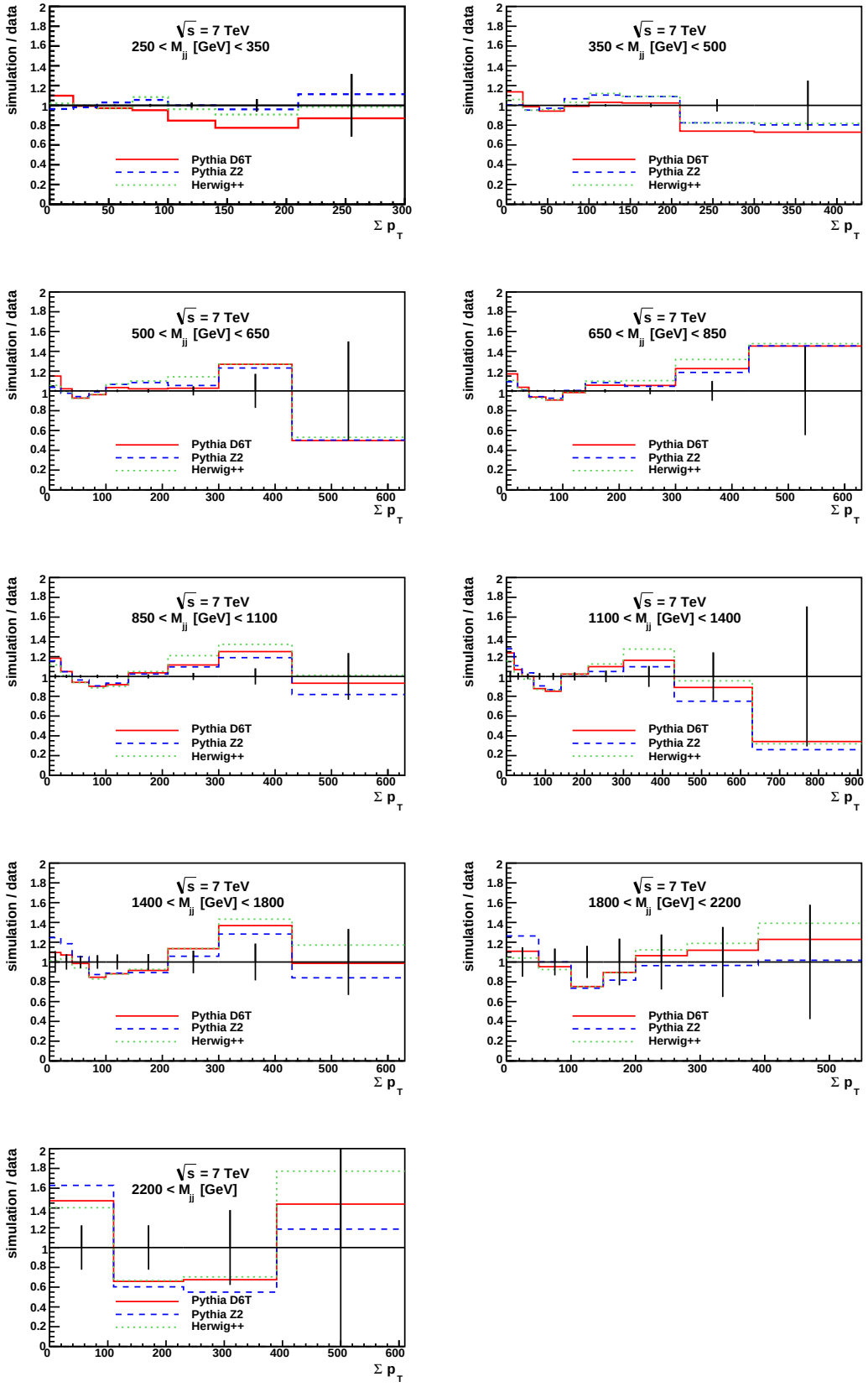


Figure 8.20: Ratio of simulation and data for $\sum p_T = \sqrt{(p_{x,1}+p_{x,2})^2+(p_{y,1}+p_{y,2})^2}$.

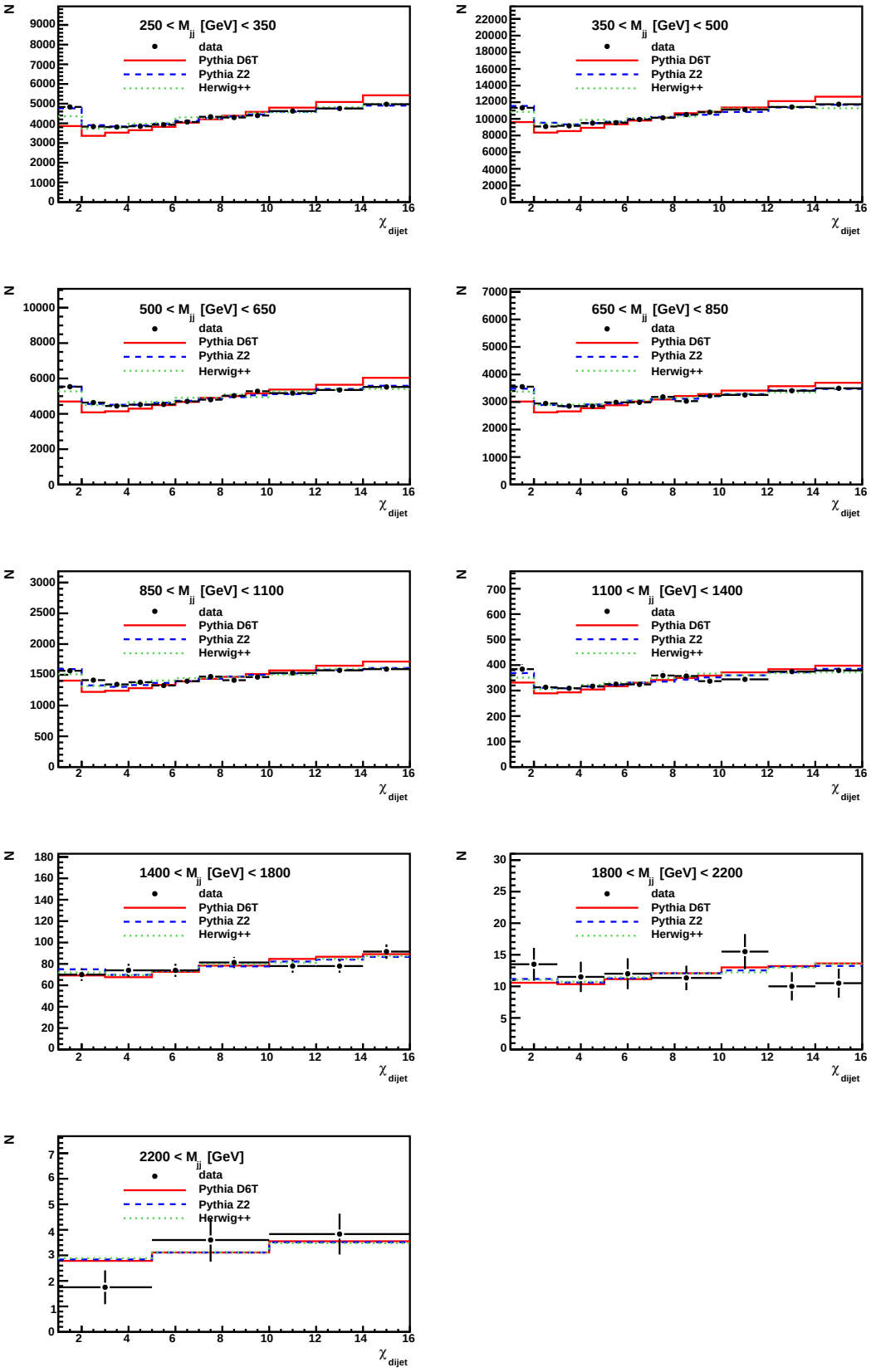


Figure 8.21: χ distributions in data and simulation.

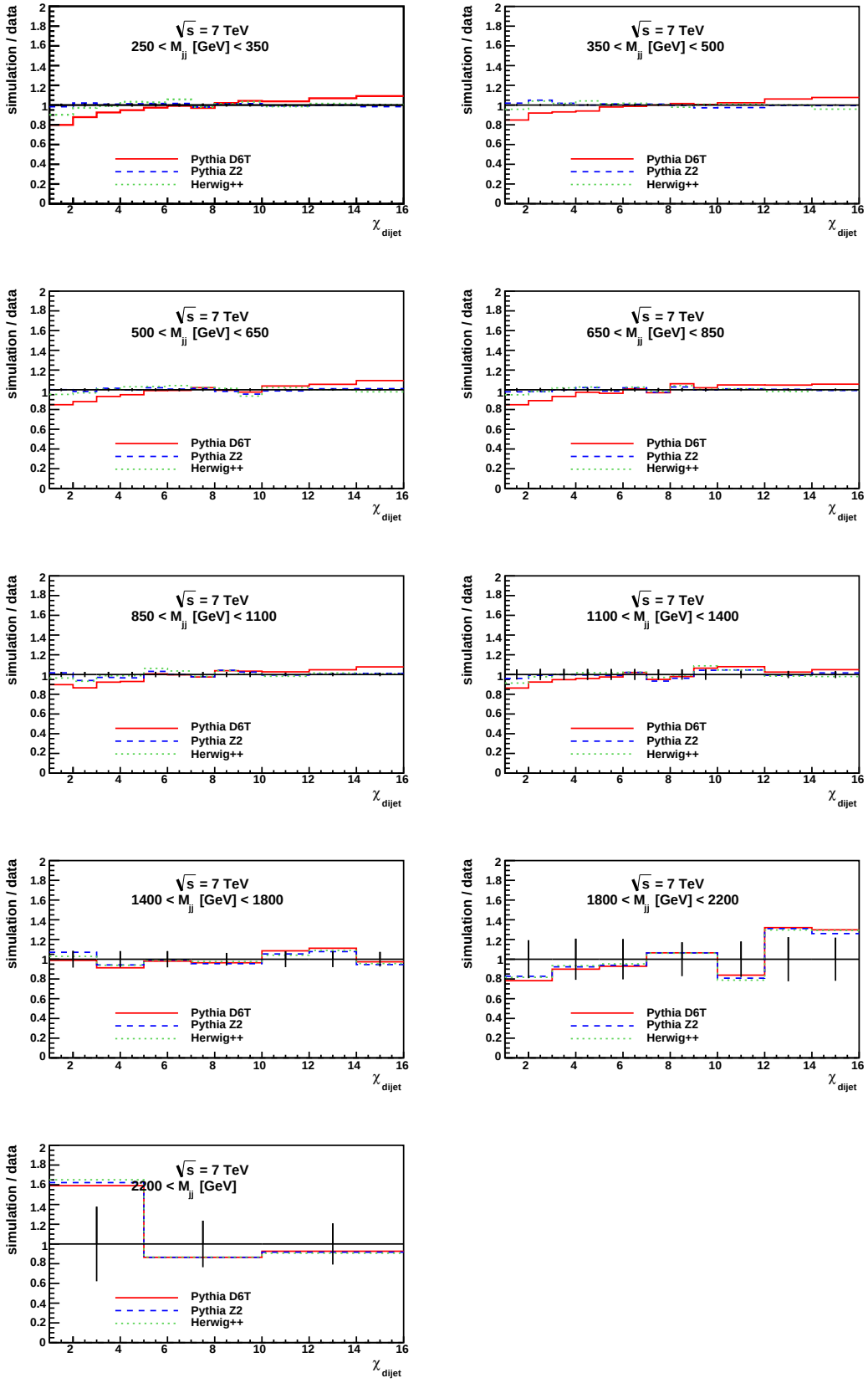


Figure 8.22: Ratio of simulation and data for the χ distributions.

showering algorithms in the three generators and tunes (for details see Sec. 5.1). This can be understood from Fig. 8.23 which shows the PYTHIA D6T prediction with different settings for the amount of initial state ratiation (PARP(67)) and final state radiation (PARP(71)). One can see that the effect of the initial state radiation is of the same order as the difference between the three generators and tunes which use different showering algorithms. The effect of the final state radiation is comparably small.

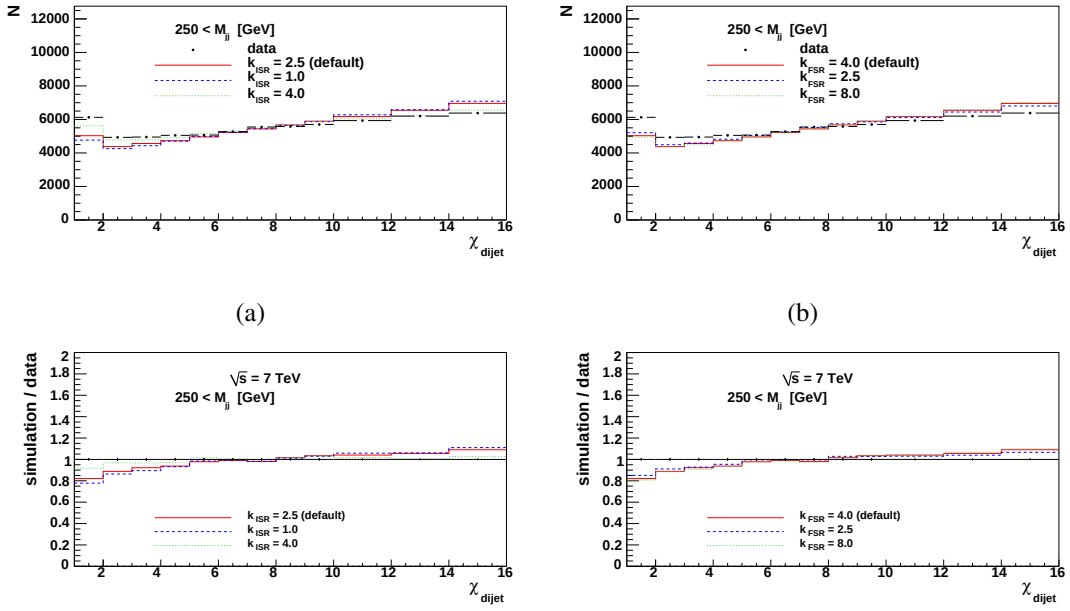


Figure 8.23: Impact of initial and final state radiation on the dijet angular distributions.

9 Measurement of the Dijet Angular Distributions

The measurement of the dijet angular distributions includes the understanding and correction of all detector effects as well as a determination of the experimental uncertainties of the measured distributions. In this chapter first the correction for detector effects will be described. Then the experimental uncertainties are discussed in detail. Finally the measured distributions are presented.

9.1 Correction of Detector Effects

The dijet angular distributions are by construction a quantity that can be measured with low systematic uncertainties with the CMS detector. The normalization strongly reduces the impact of the most prominent experimental uncertainties, such as the integrated luminosity and the jet energy scale uncertainties, and the measurement of angles can be performed with high precision as we have seen in Chap. 7. Nevertheless, there remain two major categories of detector effects which have impact on the dijet angular distributions. First, the loss of events due to missing detector acceptance and trigger inefficiencies, and second, mis-measurement of the jet angles and energies.

As the M_{jj} bins have been chosen to ensure $>99\%$ trigger efficiency in all mass bins, losses due to trigger inefficiencies are negligible. The kinematic event selection ensures $>99\%$ jet reconstruction efficiency in all M_{jj} and χ_{dijet} bins. Also, the event cleaning cuts described in Sec. 8.4 are $>99.9\%$ efficient and therefore negligible. Therefore no correction is applied for these effects and the corresponding uncertainties are neglected.

Mis-measurement of the jet angles and energies are the leading detector effects for which the data have to be corrected. The largest effects come from the finite energy resolution of the reconstructed jets, which arises from energy fluctuations due to electronic noise, energy sampling, non-uniformities and non-linearities of the detector, as well as the jet angular resolution which arises from the finite size of the calorimeter towers. These effects lead to the migration of events between the M_{jj} and χ_{dijet} bins. As the jet energy and angular resolution are by far the dominant effects, a correction for these is applied while all other detector effects are treated as systematic uncertainties.

To study the impact of the jet resolutions on the dijet angular distributions, simulated dijet events are used. The two samples generated using PYTHIA with tune D6T and HERWIG++

smeared with parametrized jet resolutions described in Sec. 8.5 are used. The impact of the detector resolution on the dijet angular distributions is investigated by calculating the correlation of M_{jj} and χ_{dijet} between particle-level dijets and detector-resolution-smeared dijets. The resulting correlation matrix is shown in Fig. 9.1. A strong correlation between the measured and the true χ_{dijet} is observed. In 99.9% of the cases, events are found within the first neighboring χ_{dijet} bin. This is due to the coarse binning in χ_{dijet} compared to the jet angular resolution. Most migrations happen due to the jet energy resolution, which moves events between different M_{jj} bins. This effect is however moderate with a purity of $>75\%$ in each bin.

Using the correlation matrix, it is possible to evaluate the dijet angular distributions at particle-level with all detector effects removed. This procedure is called unfolding. A variety of procedures exist for this purpose [86]. As the correlation matrix is well behaved, being diagonal with small correlations, this analysis uses an approximate approach to unfold the data, the "bin-by-bin unfolding" or "correction factor" method. Instead of the full matrix information, the ratio of the dijet angular distributions of the particle-level dijet and the corresponding smeared dijet is used to correct for detector effects. This approximation neglects bin-to-bin correlations which are small in this case. As the statistical uncertainties on the corrected distributions may be underestimated using this method, a cross check of these is performed in Sec. 11.4.

Figure 9.2 shows the ratio of dijet angular distributions of particle-level dijets and the detector-resolution-smeared dijets for PYTHIA (D6T) and HERWIG++. This correction factor is used to correct the data for detector effects. It lies between 90% and 98% in all mass bins. The fact that this value is smaller than 1 is explained as follows. On average the smearing of the jet energy due to the limited detector resolution moves the same amount of events upwards and downwards in dijet mass. As, however, the dijet mass spectrum is steeply falling, the upward moving events have a much larger impact on the measured number of events as the downward moving ones, and therefore the measured number of smeared dijets is higher than the particle-level dijets. The correction on the total number of events in each mass bins has no impact on the result of the analysis as normalized distributions are measured. It is however important to note that this factor is close to unity, in order to derive consistent results using the "correction factor" method. While the normalization of the correction factors in Fig. 9.2 has no impact on the dijet angular distributions, the shape of the correction factor is important. It varies over the range of $1 < \chi_{\text{dijet}} < 16$ by less than $\pm 3\%$ in all mass bins. The largest effects are in the lowest mass bin, where the jet energy resolution is worst.

The correction factors from PYTHIA (D6T) and HERWIG++ are fit to remove statistical fluctuations in the simulated events using a first order polynomial giving a good description. Both generators yield similar results varying by less than $\pm 1.5\%$. In Sec. 8.5, we have seen that both PYTHIA and HERWIG++ give a reasonable description of the data. As it is unclear which one is more accurate to describe the measured dijet events, the mean of the two is used for the corrections factor and their spread is used as a measure for the corresponding uncertainty which will be discussed in Sec. 9.2.4.

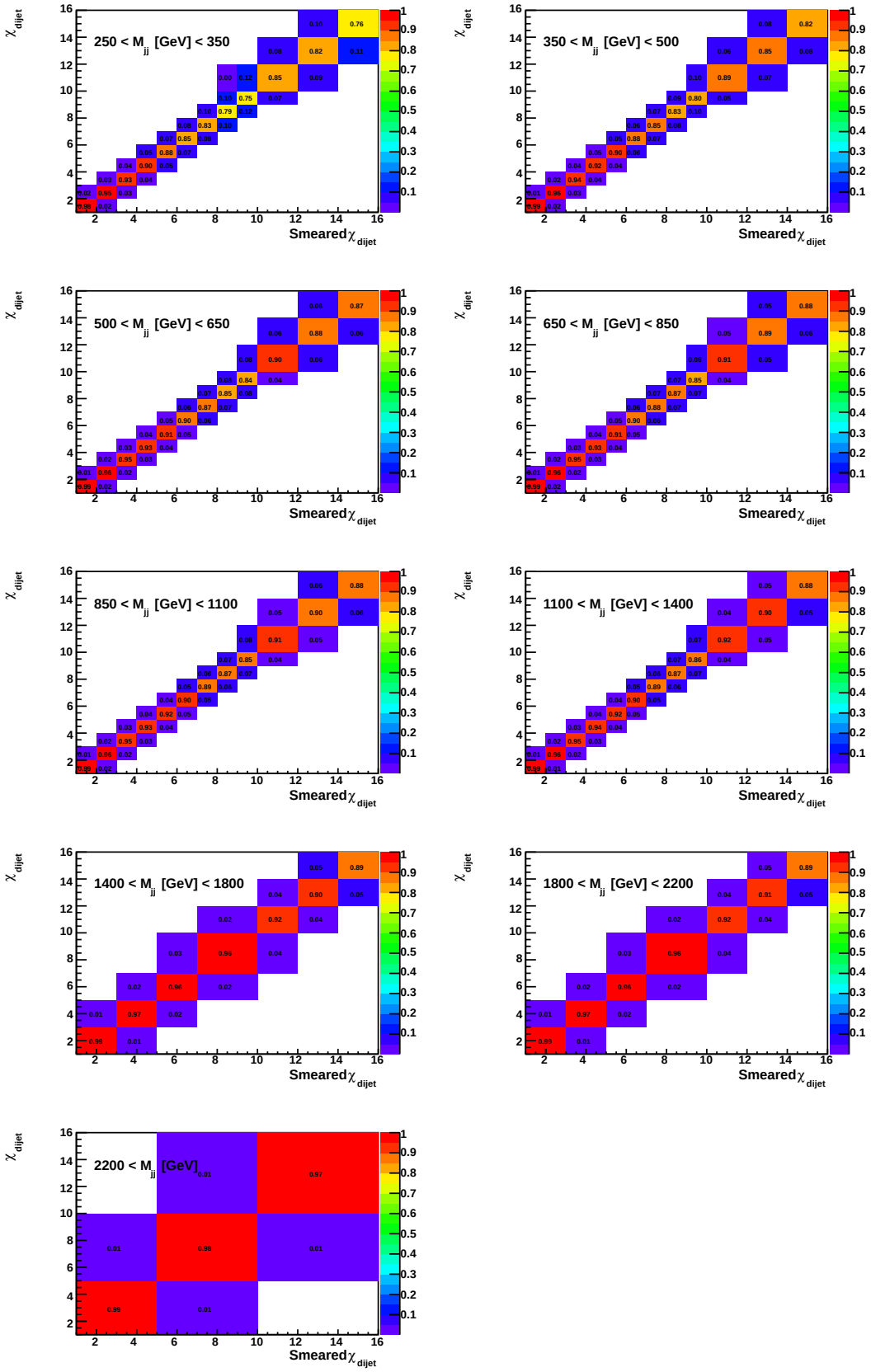


Figure 9.1: Correlation matrix between particle-level dijets and detector-resolution-smeared dijets. For each smeared dijet on the x-axis, the corresponding particle-level dijet is filled on the y-axis. The numbers indicate the fraction of smeared dijets in a certain mass bin which originate from a particle-level dijet from any other bin. All fractions above 1‰ are shown.

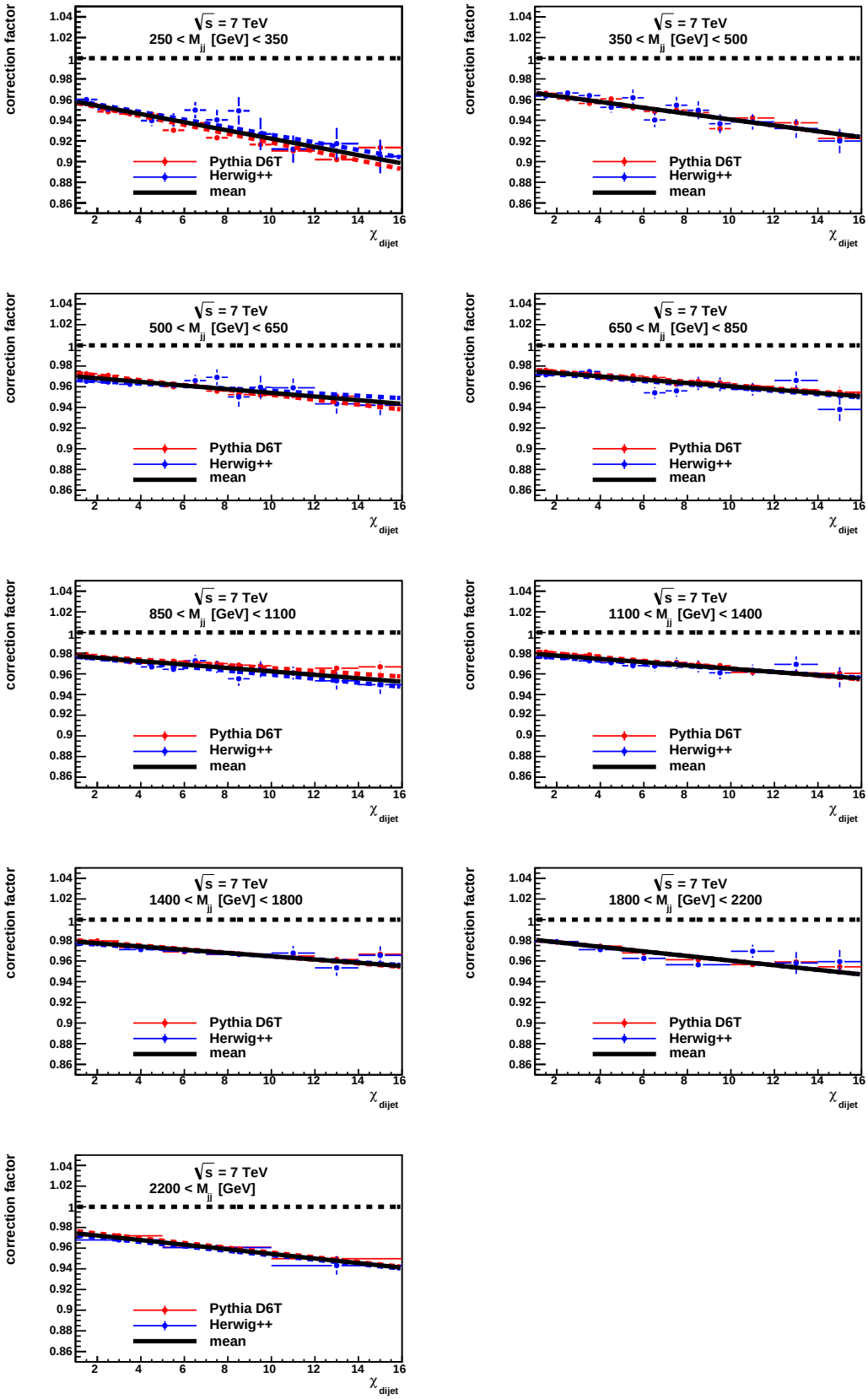


Figure 9.2: The ratio of dijet angular distributions of particle-level dijets and the detector-resolution-smeared dijets for PYTHIA (D6T) (red) and HERWIG++ (blue), the fits to both (red and blue dashed lines) and their mean (black line).

9.2 Experimental Uncertainties

The dominant sources of experimental uncertainty on the dijet angular distributions arise from the data statistics, the jet energy scale (JES), the jet energy resolution (JER) and the detector resolution correction. The contributions to the total uncertainty from each of these sources are described below.

9.2.1 Statistical Uncertainty

The statistical uncertainty of the dijet angular distributions varies from $\sim$1.5% in the lowest mass bin to $\sim$30% in the highest mass bin. In principle the underlying statistics for normalized distributions is a multinomial distribution. However for the high number of bins used, it is well approximated by the Poisson distribution. In the most extreme case of the highest mass bin with only 3 bins in χ_{dijet}, the use of a multinomial statistics would reduce the standard deviation by 10%.

9.2.2 Jet Energy Scale

As described in Sec. 7.4, the uncertainty of the jet energy scale has been parametrized in bins of jet p_T and η [16]. The resulting uncertainty on the shape of the dijet angular distributions is evaluated by varying the four-momenta of all jets within their uncertainties and evaluating the effect on the χ_{dijet} distributions. Two variations are applied, varying the jet four-momenta one standard deviation upwards and one standard deviation downwards according to the p_T and η dependent uncertainties. This method is approximate in a sense that correlations of the uncertainty in different regions of p_T and η are not fully taken into account, but is sufficient at the given level of jet energy scale precision and therefore the recommended procedure in CMS [16]. The resulting effect on the dijet angular distributions in shown in Fig. 9.3. It is found to be less than 3% in all mass bins, symmetric and fairly linear. The uncertainty is maximal in the highest mass bin, due to the uncertainty in the extrapolation of the jet energy scale from the transverse momentum range where it was measured in photon+jet events.

9.2.3 Jet Energy Resolution

The jet energy resolution is known with a precision of 10% (independent of p_T) [14]. To evaluate the effect of the jet energy resolution uncertainty on the dijet angular distributions, the particle-level jets are smeared with the jet energy resolution varied by 10% from the nominal values. The ratio between the dijet angular distributions with varied resolutions and the nominal resolution is shown in Fig. 9.4. The resulting uncertainty is less than 1.5% in all mass bins. It is largest in the lowest mass bin where the jet resolution is worse than in the higher mass bins.

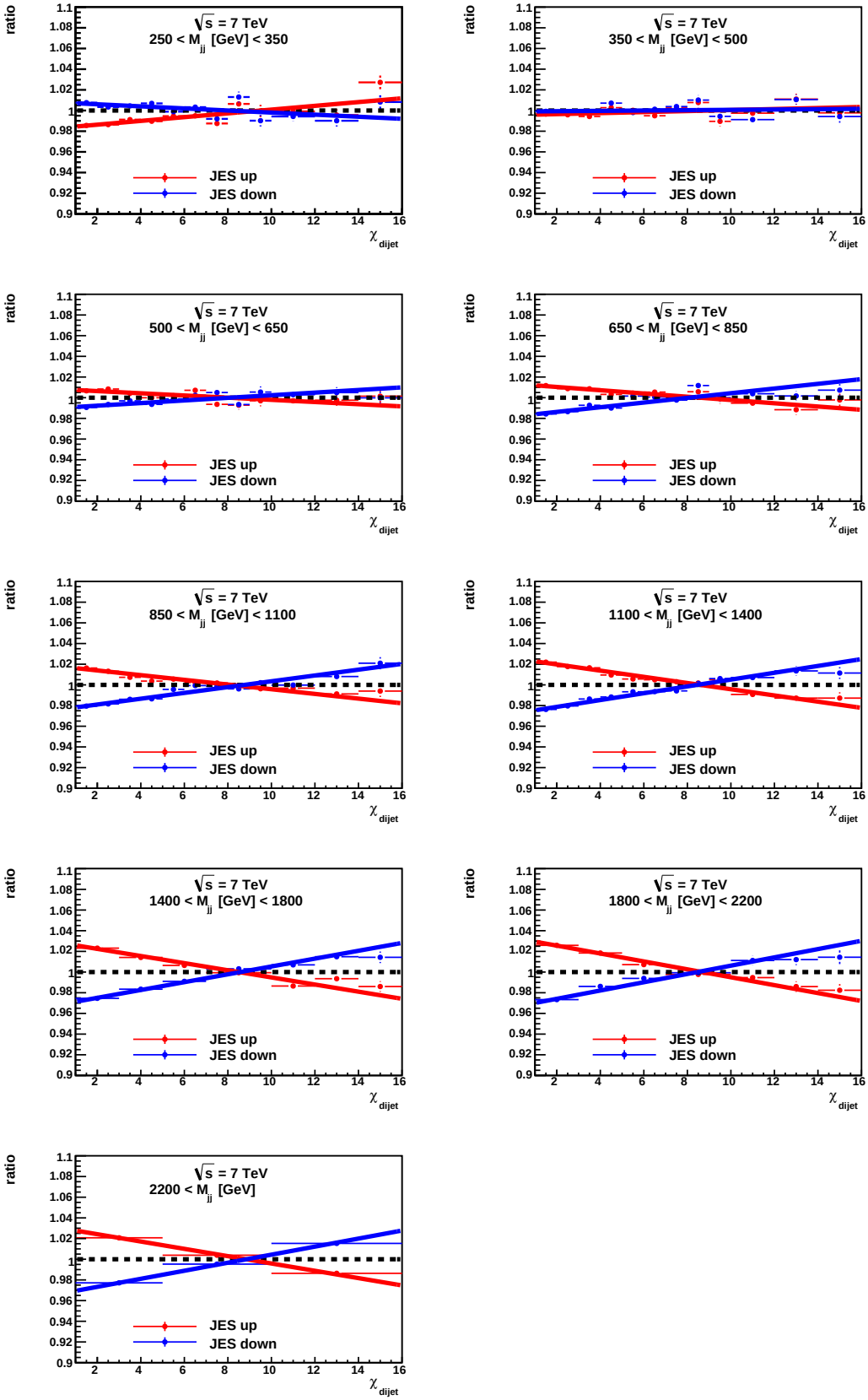


Figure 9.3: The effect of the jet energy scale uncertainty on the dijet angular distributions.

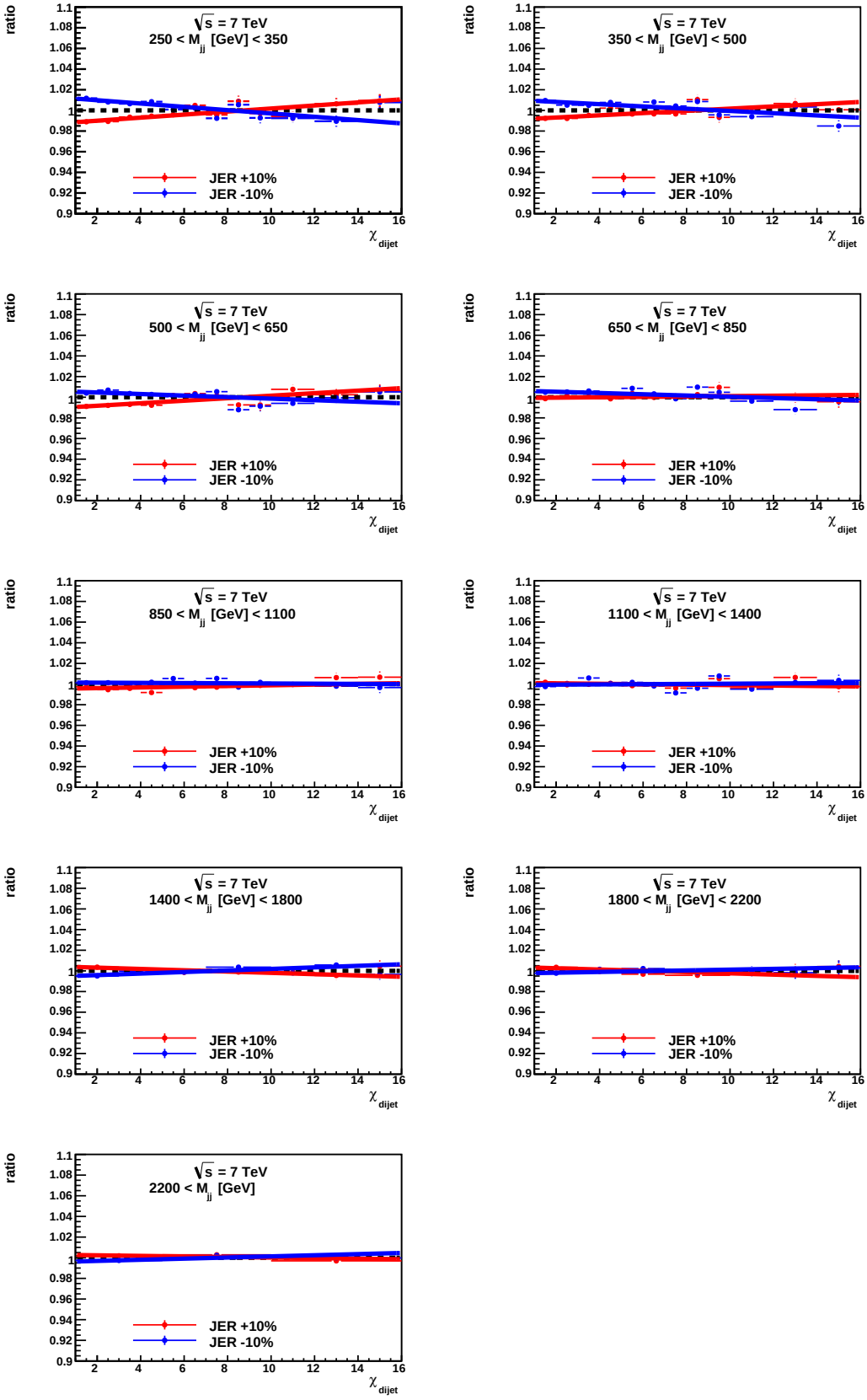


Figure 9.4: The effect of 10% jet energy resolution uncertainty on the dijet angular distributions.

9.2.4 Detector Effects Correction

The correction for detector effects, described in the previous section, relies on the description of dijet events by the simulation. Event generators are used to evaluate the effect of the jet resolution on the dijet angular distribution. Therefore an uncertainty quantifying the knowledge of the simulation of dijet events has to be assigned.

Half of the value of the correction factor on the shape of the dijet angular distributions is used as an estimate of this systematic uncertainty, as shown in Fig. 9.5. This is motivated as follows. As we have seen in Sec. 8.5, the variation of the generators and tunes is a good estimate for the agreement of dijet data and simulation. Figure 9.5 shows the detector effect correction derived from each generator and tune. Their variation is well described by the assigned uncertainty, which should therefore be a good estimate. The uncertainty is less than 1.5% in all mass bins, with its highest value in the lowest mass bin, where the different underlying event models of the generators have their largest impact.

9.2.5 Other Detector Effects

The correction for detector effects takes into account the jet energy resolutions, while other detector effects are neglected. In order to quantify the remaining effects, the jet-resolution-smeared simulation is compared to a full CMS detector simulation. In this sample particles from the PYTHIA event generator with tune Z2 are propagated through a full GEANT4-based CMS detector simulation with a total of 10^7 events. The ratio of the jet-resolution-smeared simulation and the full CMS detector simulation as shown in Fig. 9.6 is treated as an uncertainty. Due to the limited statistics of the full simulation sample, the uncertainty is approximated by the weighted root mean square for each mass bin, defined as

$$\sigma_{\text{Detector smearing}} = \sqrt{\frac{1}{\sum_{i\in\chi bins}\frac{1}{\sigma_i^2}}\sum_{i\in\chi bins}\frac{\Delta^2_{\text{Reco-GenSmear}}}{\sigma_i^2}}\,. \tag{9.1}$$

The resulting uncertainty, indicated as a red band in Fig. 9.6, is less than 1% for all mass bins.

9.2.6 Impact of Pile-up Interactions

Pile-up interactions, i.e. collisions overlapping in the same bunch-crossing, may impact the measurement of the dijet angular distributions in the following way. As it is unlikely to have two high-p_T dijet interactions in a single event, the pile-up interactions typically have a much smaller p_T than the main interaction in an event. Therefore, the pile-up interactions impact the dijet angular distributions mainly by a change of the jet energy scale due to the additional activity in the event. As this effect is already taken into account in the jet energy scale uncertainty, it is not treated as an additional uncertainty.

In order to exclude any unexpected impact of the pile-up interactions to the dijet angular distributions, the ratio of data events with exactly one reconstructed primary vertex and all events with at least one reconstructed primary vertex is studied in Fig. 9.7. On average about 50% of the events are reconstructed with more than one primary vertex. The effect of pile-up is found to lie within the statistical uncertainties.

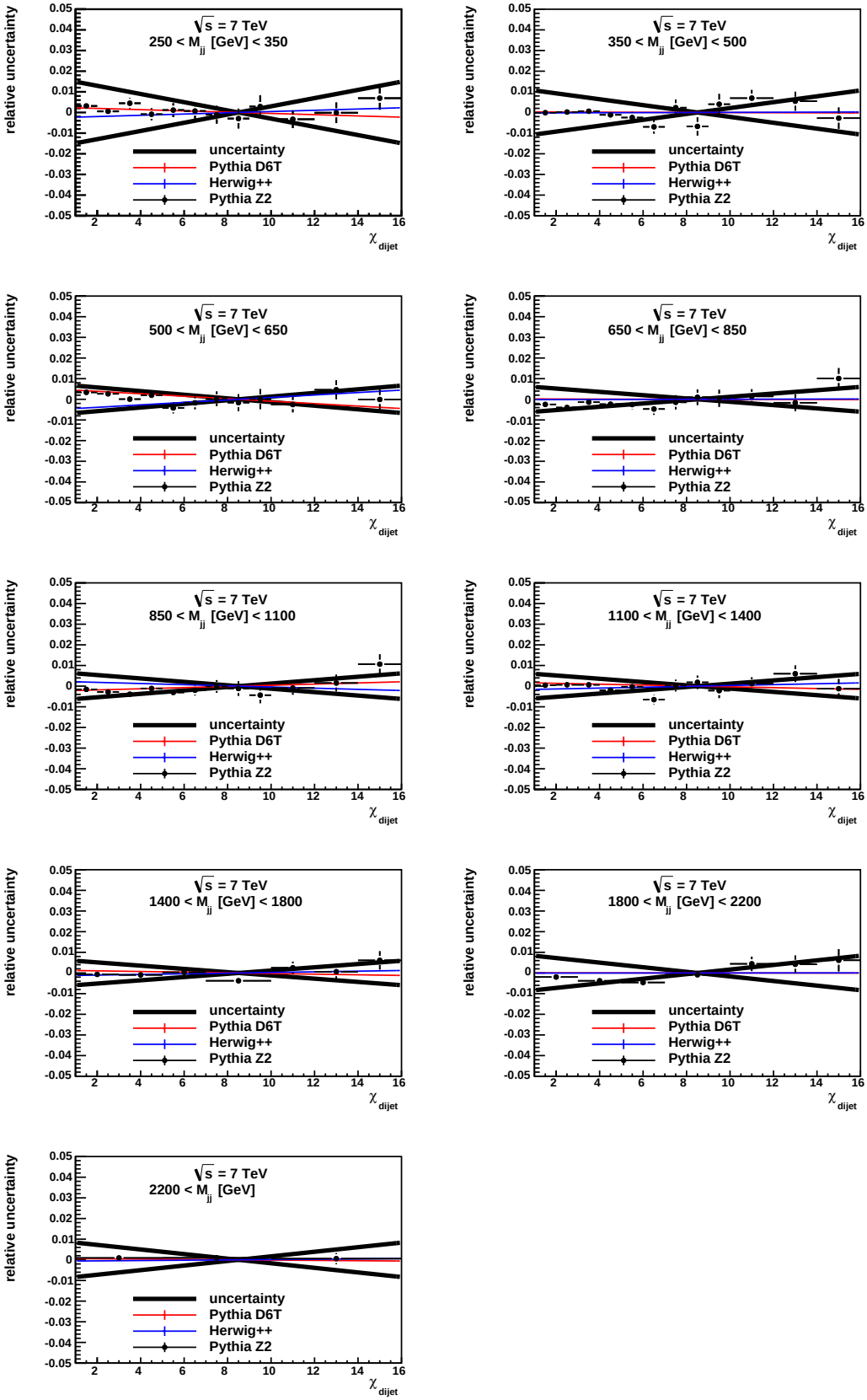


Figure 9.5: Systematic uncertainty on the detector effects (black lines) and comparison with HERWIG++ (blue line) and PYTHIA tune D6T (red line) and Z2 (black markers).

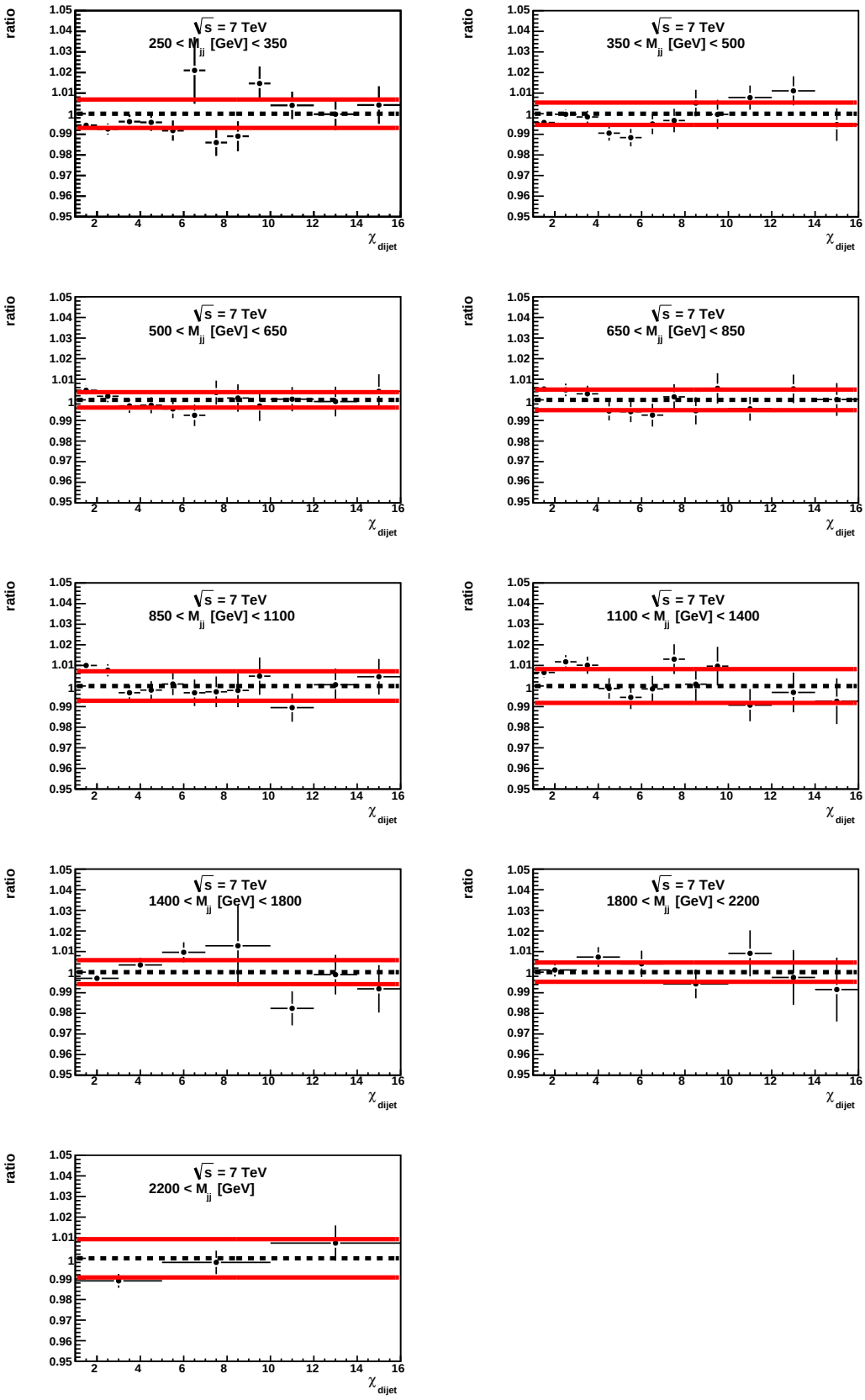


Figure 9.6: The ratio of dijet angular distributions of the smeared jets and the corrected calorimeter jets (black markers). The corresponding assigned uncertainty is indicated as red lines.

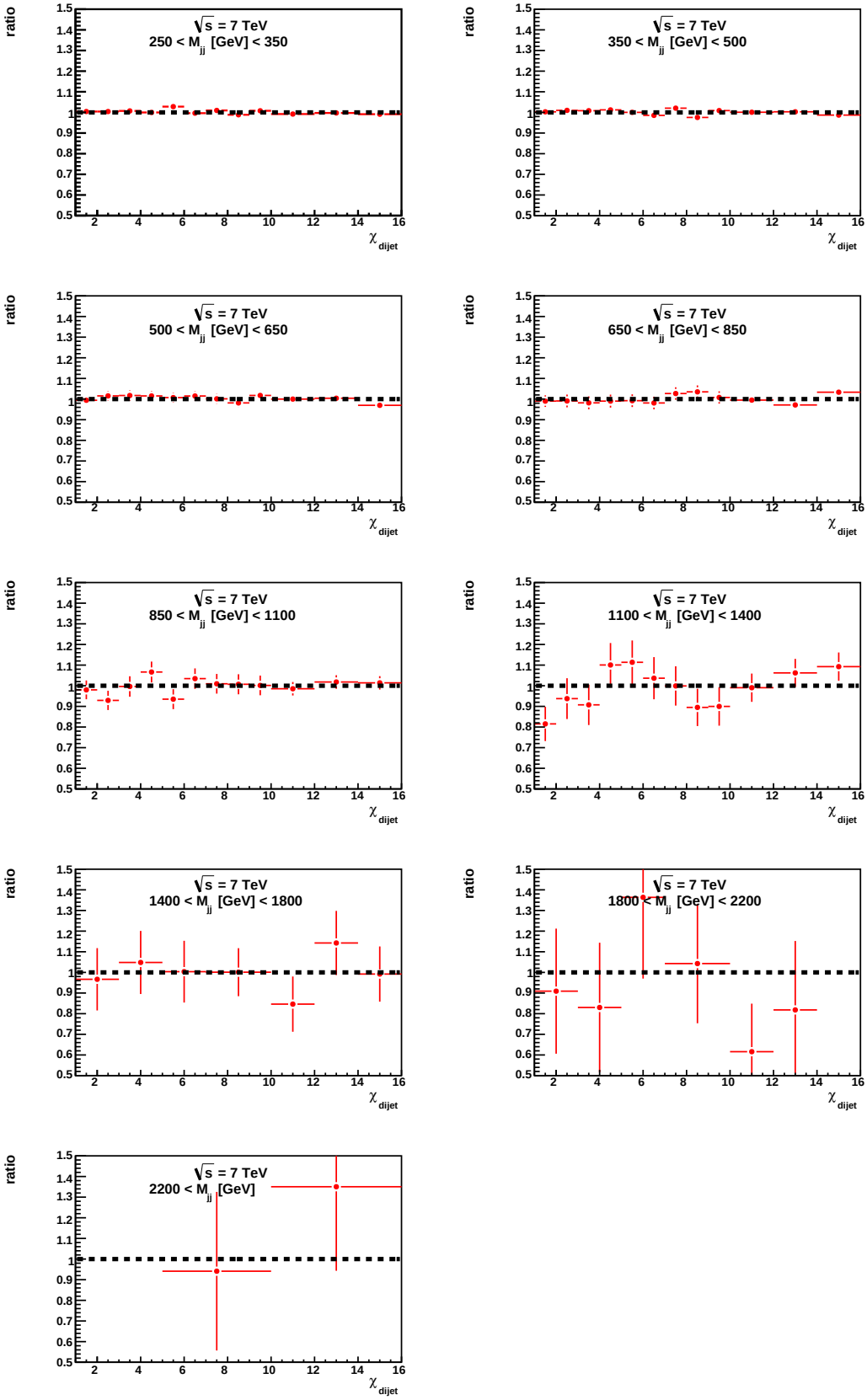


Figure 9.7: The ratio of the angular distributions restricted to events with exactly one reconstructed primary vertex and all events with at least one reconstructed primary vertex.

9.3 Data Distributions

Fig. 9.8 shows the measured dijet angular distributions. They are corrected for detector effects and all previously described uncertainties are included. The systematic uncertainties are displayed as an error bar with ticks. Note that they are correlated between the χ_{dijet} bins. The quadratic sum of systematic and statistical uncertainties is displayed as an error bar with open ends. By comparing the two errors bars, it becomes clear that the measurement of the dijet angular distributions is dominated by systematic uncertainties in the lower mass bins, and completely dominated by statistical uncertainties in the highest mass bins. A complete listing of the data points and uncertainties can be found in the Appendix. A summary table of all uncertainties will be presented in Chap. 12.

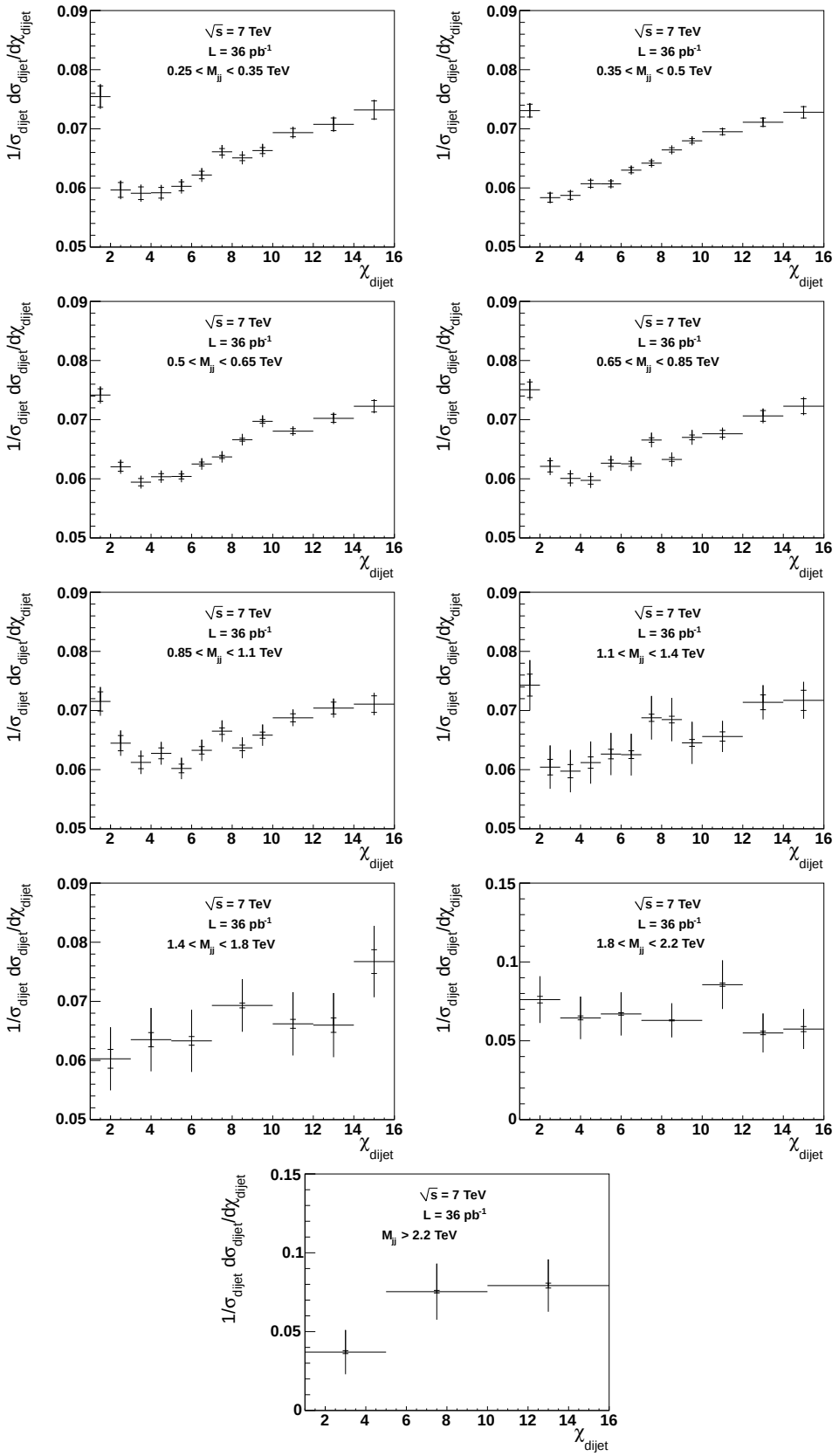


Figure 9.8: Dijet angular distributions in data, corrected for detector effects. The error bars represent statistical and systematic uncertainties combined in quadrature; the innermost interval with ticks indicates the systematic uncertainties only.

10 Predictions of the Dijet Angular Distributions

In this chapter the theory predictions for the dijet angular distributions are described. The nowadays most precise available predictions from perturbative QCD at next-to-leading order (NLO) are used. These are combined with non-perturbative effects from MC event generators. In the following, first, the QCD predictions and the associated uncertainties are discussed, followed by the predictions for contact interaction signals.

10.1 QCD Predictions

The predictions for the dijet angular distributions are calculated to next-to-leading order using the NLOJET++ program [87]. This program is also integrated into the FASTNLO package which is used to calculate PDF and scale uncertainties with fast performance [88]. The calculation is performed using the CTEQ6.6 NLO parton distribution functions (PDFs) [6] with the same value of α_S used in the matrix element calculation and in the PDFs. The factorization scale μ_F and the renormalization scale μ_R are set to the average p_T of the two leading jets ($\mu_F = \mu_R = \langle p_T \rangle$). The resulting dijet angular distribution before normalization are shown in Fig. 10.1 together with a leading order calculation using the CTEQ6L1 PDFs.

In order to compare these predictions to data, non-perturbative effects have to be taken into account. These effects include multiple parton interactions additional to the hard scattering process and hadronization, which transforms the parton final state into a particle final state. The non-perturbative corrections are calculated using MC event generators and then applied to the NLO predictions:

$$\sigma^{QCD}_{NLO+\text{Non-Pert.}} = \sigma^{QCD}_{NLO} \cdot (\text{non-perturbative corrections}) \tag{10.1}$$

PYTHIA 6.422 with tune D6T and HERWIG++ 2.4.2 are used to derive independent estimates of the non-perturbative effects. They have different modeling of multiple parton interactions and hadronization (for details see Sec. 5.1). The corrections are derived by dividing the default prediction from each generator by the prediction where hadronization (`MSTJ(1)` in PYTHIA, `LHCHandler:MultipleInteractionHandler` in HERWIG++) and multiple parton interactions (`MSTP(81)` in PYTHIA, `LHCHandler:HadronizationHandler` in HERWIG++) are switched off.

The resulting corrections are shown in Fig. 10.2. They range from 20% in the lowest mass bin to 2% in the highest mass bin. Note that only the shape of the corrections, which varies by

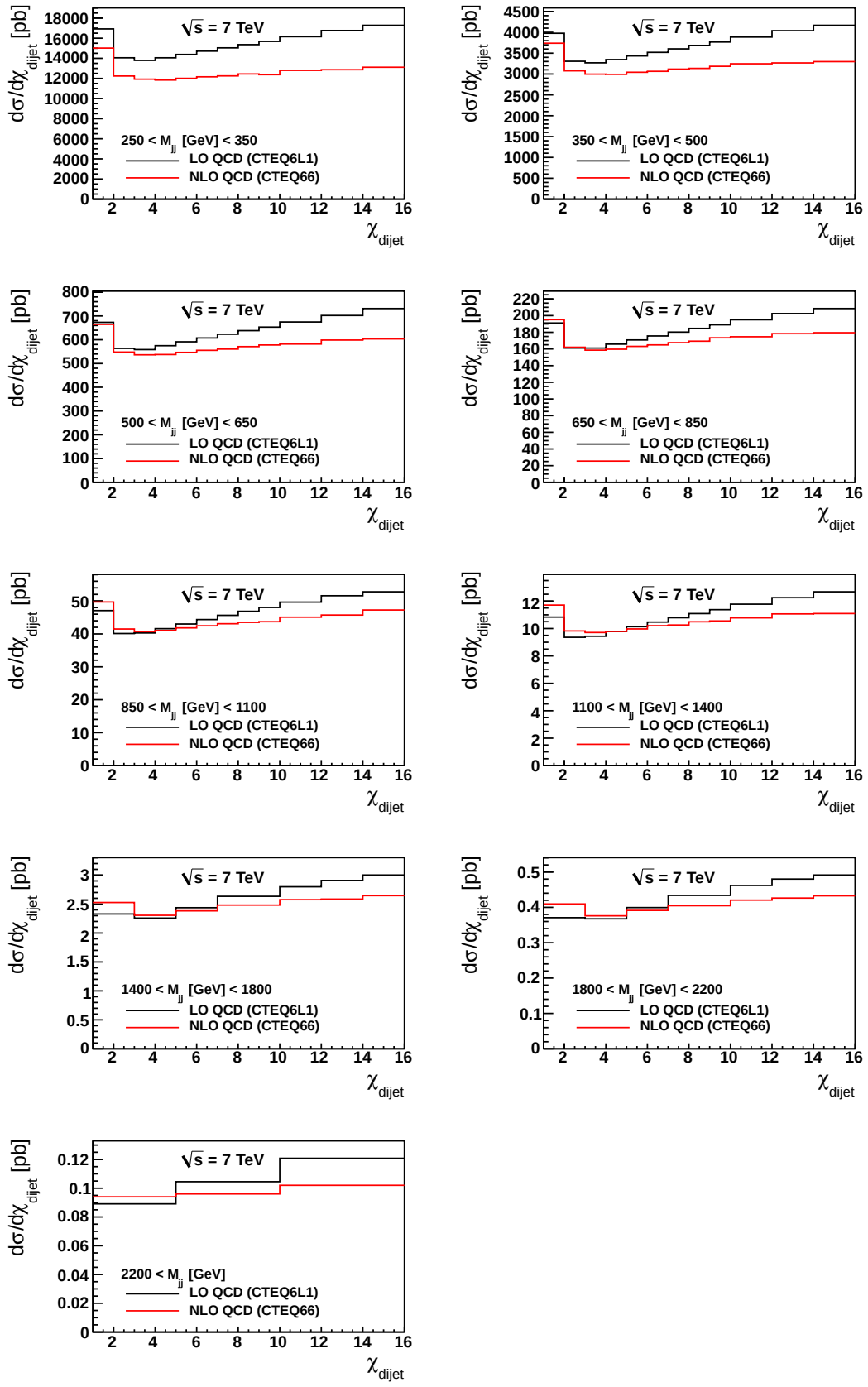


Figure 10.1: Comparison of LO and NLO QCD predictions for the dijet angular distributions before normalization.

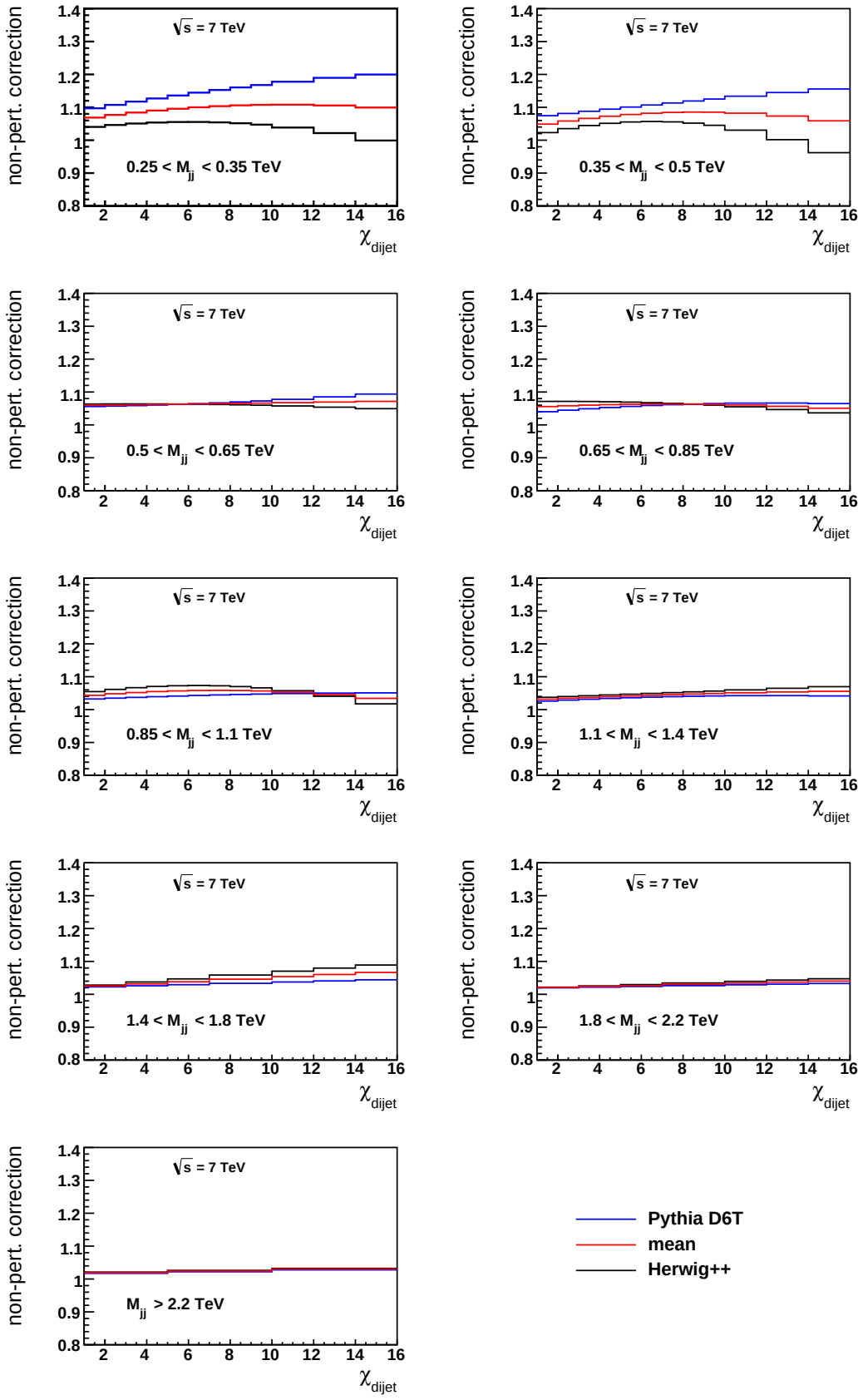


Figure 10.2: Non-perturbative corrections derived from the mean (red) of PYTHIA (blue) and HERWIG++ (black).

4% to 0.1%, is important for the dijet angular distributions. As it is not clear which of the two non-perturbative models gives a more accurate description of the effects, the mean of the two is used for the correction and their spread is used as a measure for the corresponding uncertainty which will be discussed in Sec. 10.2.3.

Figure 10.3 shows the full prediction for the normalized dijet angular distributions from NLO QCD with non-perturbative corrections applied.

10.2 Theoretical Uncertainties of the QCD Predictions

The dominant sources of uncertainty on the QCD prediction come from the renormalization and factorization scales, the parton distribution functions and the modeling of the non-perturbative effects. Each of these uncertainties is discussed in the following.

10.2.1 Renormalization and Factorization Scales

As the QCD predictions are derived in perturbation theory, their higher order corrections are uncertain. A commonly used procedure to estimate the size of the expected corrections is to vary the choice of the renormalization scale μ_R and evaluate its effect on the prediction. This is motivated by the fact that in higher order calculations the effect of the choice of μ_R should vanish. Also the exact choice of the factorization scale μ_F is ambiguous and is therefore varied to estimate the uncertainty related to its choice.

It is common practice to vary them at the same time by a factor of 2 upwards and downwards. However a more sophisticated approach is used here, as there is no reason why these scales should be correlated. The two scales are varied independently in the ranges

$$\langle p_T \rangle /2 \leq \mu_R \leq 2 \langle p_T \rangle \ , \tag{10.2}$$

$$\mu_R/2 \leq \mu_F \leq 2\mu_R \ , \tag{10.3}$$

resulting in the following 6 combinations:

$$(\mu_F, \mu_R) = (\langle p_T \rangle /2, \langle p_T \rangle /2), (\langle p_T \rangle /2, \langle p_T \rangle), (\langle p_T \rangle, \langle p_T \rangle /2), (2\langle p_T \rangle, 2\langle p_T \rangle), \tag{10.4}$$

$$(2\langle p_T \rangle, \langle p_T \rangle), (\langle p_T \rangle, 2\langle p_T \rangle) \ . \tag{10.5}$$

The maximum effect on the angular distributions of all variations is assigned as systematic uncertainty. Note that the resulting uncertainty can be asymmetric. It is presented in Fig. 10.4 and ranges from 5% in the lowest mass bin to 9% in the highest mass bin. This is the largest contribution to the theoretical uncertainty for the dijet angular distributions.

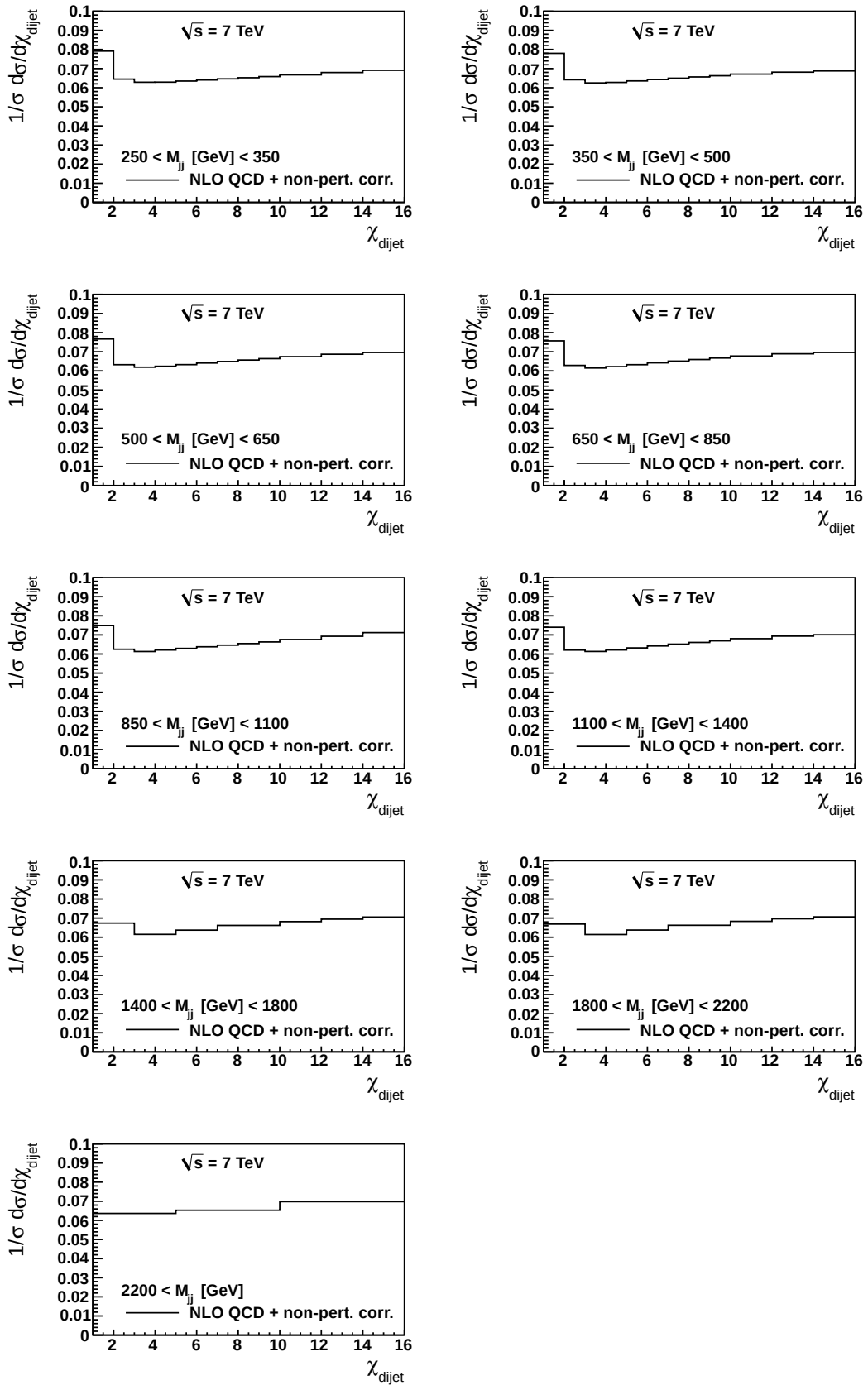


Figure 10.3: NLO QCD predictions for the dijet angular distributions with non-perturbative corrections applied.

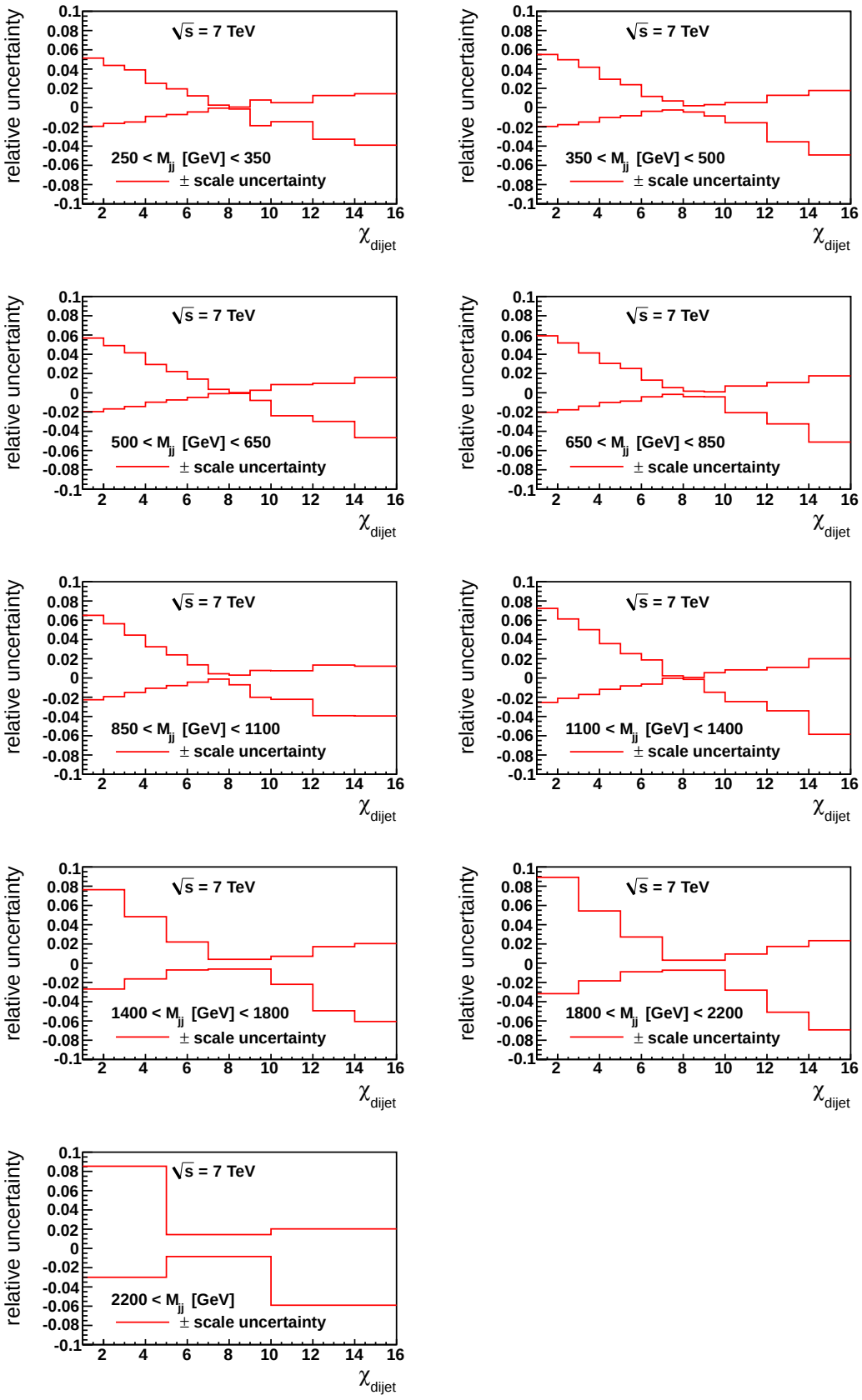


Figure 10.4: Scale uncertainty of the QCD NLO prediction.

10.2.2 Parton Distribution Functions

As the parton distribution functions have been measured in experiments with limited precision, an uncertainty is assigned to them. In the CTEQ6.6 PDF sets, the uncertainty is parametrized as variations along a set of 22 eigenvectors. The uncertainty on the dijet angular distributions is evaluated using the procedure described in Ref. [6]. For each variation along an eigenvector i, the dijet angular distributions are calculated. The uncertainty for a bin value X is then given by

$$\Delta X = \frac{1}{2}\sqrt{\sum_{i=1}^{22}(X_i^{(+)} - X_i^{(-)})^2}\,. \tag{10.6}$$

In Fig. 10.5 the resulting uncertainties are shown. They are smaller than $\pm 0.5\%$ for all mass bins. Therefore this uncertainty is almost negligible. As explained in Sec. 3.3, this is due to the fact that the shape of the angular distributions is almost insensitive to the mixture of quark flavors and gluons in the initial state.

10.2.3 Non-Perturbative Corrections

The non-perturbative corrections have been evaluated using both PYTHIA with tune D6T and HERWIG++ which have different underlying event and hadronization models. In Sec. 8.5 we have seen that the data description by both generators is reasonable and their spread is of the same order as the differences to data. Therefore half the difference between PYTHIA and HERWIG++ is assigned as a systematic uncertainty for the non-perturbative corrections, shown in Fig. 10.6. Note that the difference in shape and not the absolute value of the angular distribution is evaluated. In order to cross-check the assigned systematic uncertainty, the result for PYTHIA with tune Z2 is overlaid and found to be of the same order as the assigned uncertainty.

10.3 Contact Interaction Predictions

Four different models of contact interactions are considered in this study. The focus is on color- and isospin-singlet contact interactions between left-handed quarks as described in Sec. 3.4. In the following the NLO and LO predictions for a contact interaction with constructive (Λ^-) and destructive (Λ^+) interference with the QCD interactions are described. Both NLO and LO predictions are of interest. While the NLO predictions are at the state-of-the-art, the LO predictions are important for comparisons with older results, since the NLO predictions only became available recently [89].

The NLO predictions are calculated using the formalism described in Ref. [89] using the CTEQ66 PDFs. These calculations include the contact interaction term as well as the interference terms with QCD. The full prediction is obtained by adding the NLO QCD prediction. Non-perturbative corrections due to hadronization and multiple parton interactions were also applied to the prediction.

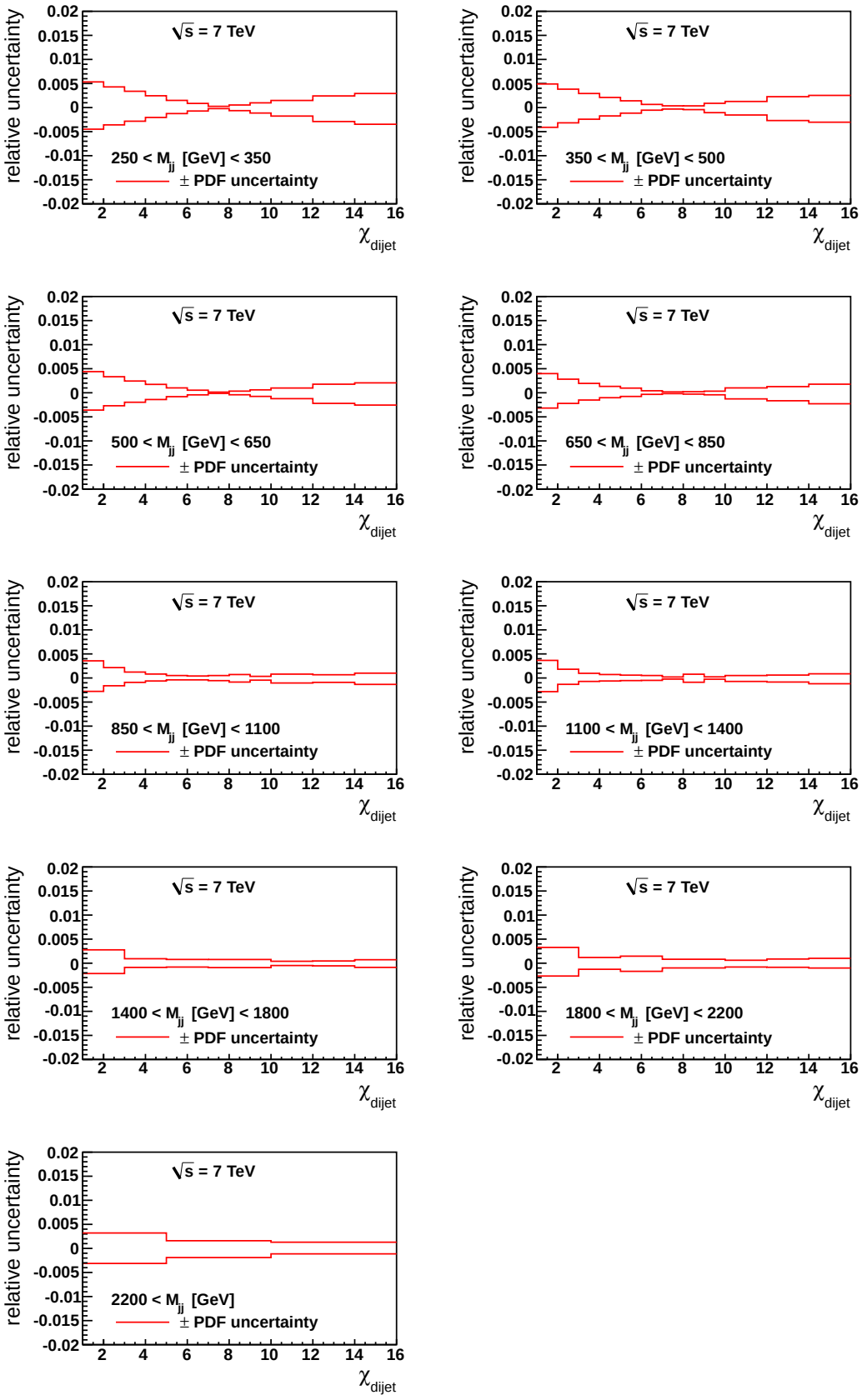


Figure 10.5: PDF uncertainty of the QCD NLO prediction.

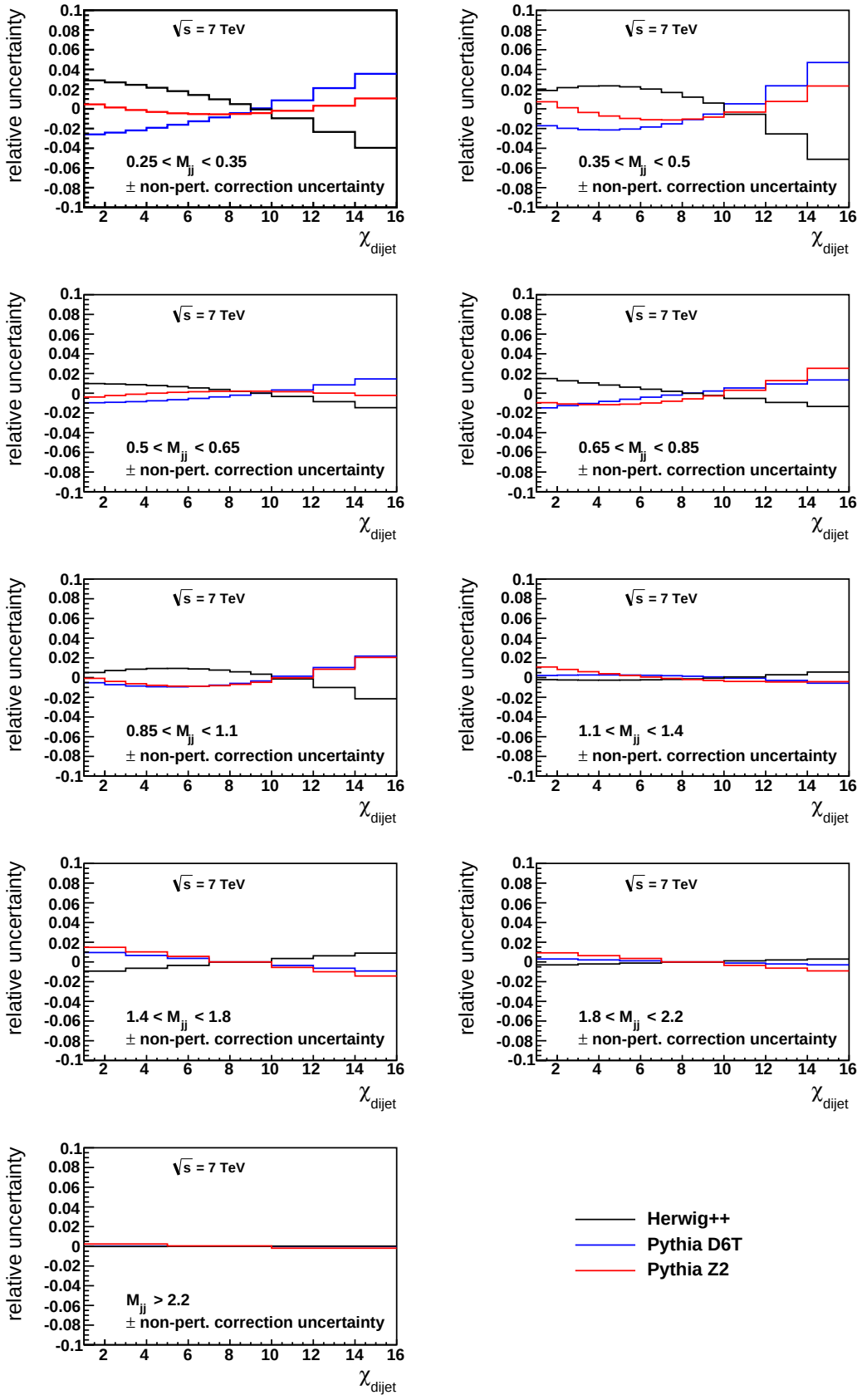


Figure 10.6: Uncertainty on the non-perturbative corrections (blue, black) and comparison with PYTHIA tune Z2 (red).

Table 10.1: Pythia6 parameters for the generation of color- and isospin-singlet contact interactions between left-handed quarks.

parameter	value	description
ITCM(5)	2	Contact interactions between all quarks flavors
RTCM(41)	Λ	Contact interaction scale in GeV
RTCM(42)	± 1	Sign of the contact interaction term
MSUB(381), MSUB(382)	1	qq and $q\bar{q}$ via contact interaction
MSUB(13), MSUB(28), MSUB(53), MSUB(68)	1	all gluon processes as in QCD

The LO predictions are calculated with PYTHIA using the CTEQ6L1 PDFs and the parameters given in Tab. 10.1. A QCD plus Contact Interaction (QCD+CI) prediction, where the QCD terms are corrected to NLO while the CI terms are calculated at LO, is obtained by adding the difference $\sigma_{NLO}^{QCD} - \sigma_{LO}^{QCD}$ (from Fig. 10.1) bin-by-bin to the CI prediction:

$$\sigma^{QCD+CI} = (\sigma_{NLO}^{QCD} - \sigma_{LO}^{QCD} + \sigma_{LO}^{QCD+CI}) \cdot (\text{non-perturbative corrections}) \tag{10.7}$$

In the limit of $\Lambda = \infty$ this formula yields the NLO QCD prediction.

Figure 10.7 shows the predictions for the dijet angular distributions for various values of Λ^+. With increasing contact interaction scale, the contribution of the contact interaction term to the measured mass bins decreases. Figure 10.8 shows the predictions for the angular distributions for all four models. The deviation from the QCD predictions are larger for the Λ^- than for Λ^+ due to the constructive interference terms with QCD. The NLO predictions for both Λ^- and Λ^+ are smaller than the LO prediction as explained in Ref. [89].

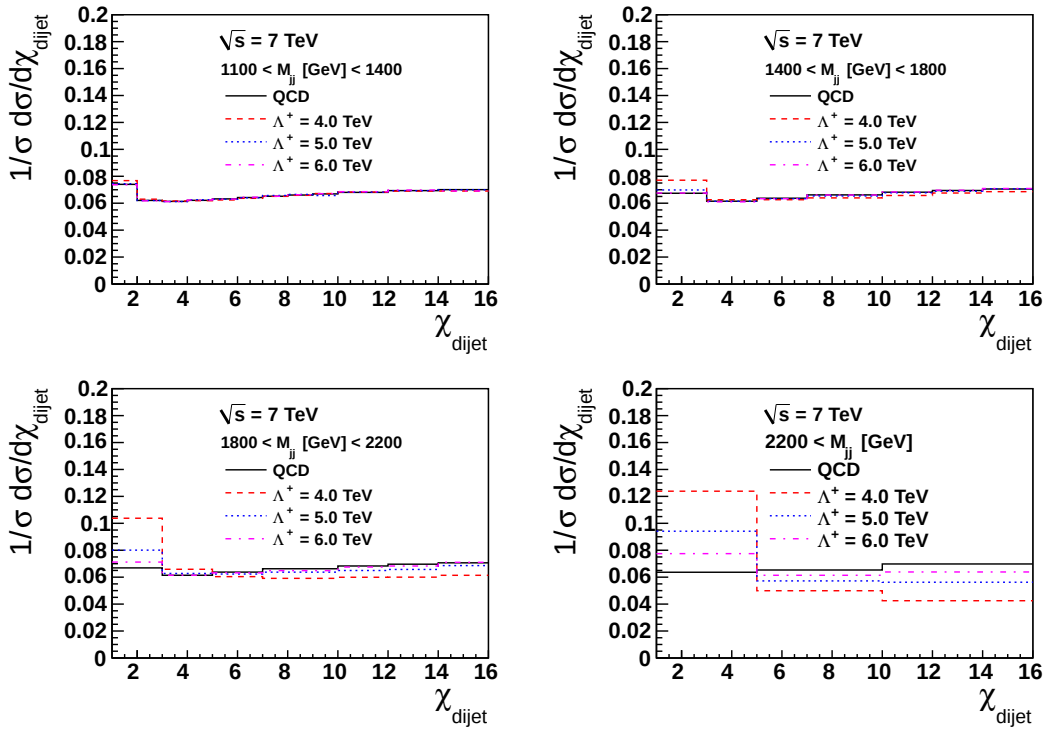


Figure 10.7: Theory prediction of the angular distributions for QCD and the contact interaction model for various scales of Λ^+.

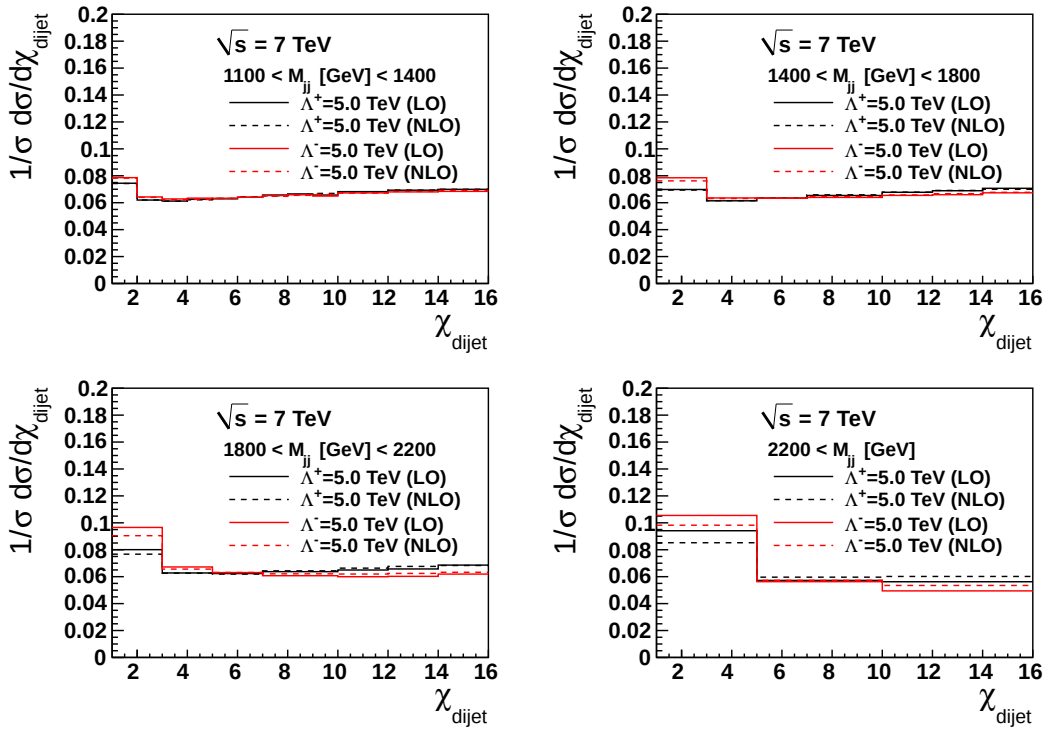


Figure 10.8: Theory prediction of the angular distributions for the LO and NLO contact interaction models with constructive (Λ^-) and destructive (Λ^+) interference with the QCD interactions

11 Search Techniques for Quark Compositeness

The quark compositeness model makes clear predictions for deviations of the dijet angular distributions from the QCD predictions. Therefore, a hypothesis test to decide between the QCD and quark compositeness model can be performed. In principle one can either test the QCD hypothesis and exclude it if there is a large variation of the data from the QCD prediction, or test the quark compositeness hypothesis and exclude it if there is a large variation of the data from the quark compositeness prediction. Only the latter is discussed in the following.

The statistical test is performed using a modified frequentist approach based on the quantity CL_s which makes use of the log-likelihood-ratio. Both statistical and systematic uncertainties of the measurement and the theory predictions are taken into account. The test is used to set limits on the contact interaction (CI) scale Λ of the quark compositeness model. In the following sections, first, the statistical method is explained, then the limit setting procedure and finally several cross checks to support the analysis are presented. A summary of the results will be given in the next chapter.

11.1 Statistical Method

The statistical test discriminating between the QCD only hypothesis and the quark compositeness (QCD+CI) hypothesis is based on the log-likelihood-ratio

$$Q = \text{test statistics} = -2\ln(\frac{L_{QCD+CI}}{L_{QCD}}) \; . \tag{11.1}$$

The likelihood functions L_{QCD+CI} and L_{QCD} are modeled as a product of Poisson likelihood functions for each bin in χ_{dijet} and M_{jj} in the four highest M_{jj} ranges:

$$L(x_i \text{ for } i \in \text{ bins}|\lambda_i \text{ for } i \in \text{ bins}) = \prod_{i \in \text{ bins}} \frac{\lambda_i^{x_i}}{x_i!} \cdot e^{-\lambda_i} \tag{11.2}$$

The λ_i correspond to the entries in each bin of the QCD (QCD+CI) model, the x_i to the entries of the data. The theory predictions for each M_{jj} range are normalized to the number of data events in that range.

For each hypothesis the probability distribution functions $P_{\text{QCD+CI}}(\text{Q})$ and $P_{\text{QCD}}(Q)$ are obtained using ensembles of pseudo-experiments. Figure 11.1 demonstrates the $P_{\text{QCD+CI}}(\text{Q})$ and $P_{\text{QCD}}(Q)$ for various contact interaction scales Λ^+ as blue and red lines, respectively.

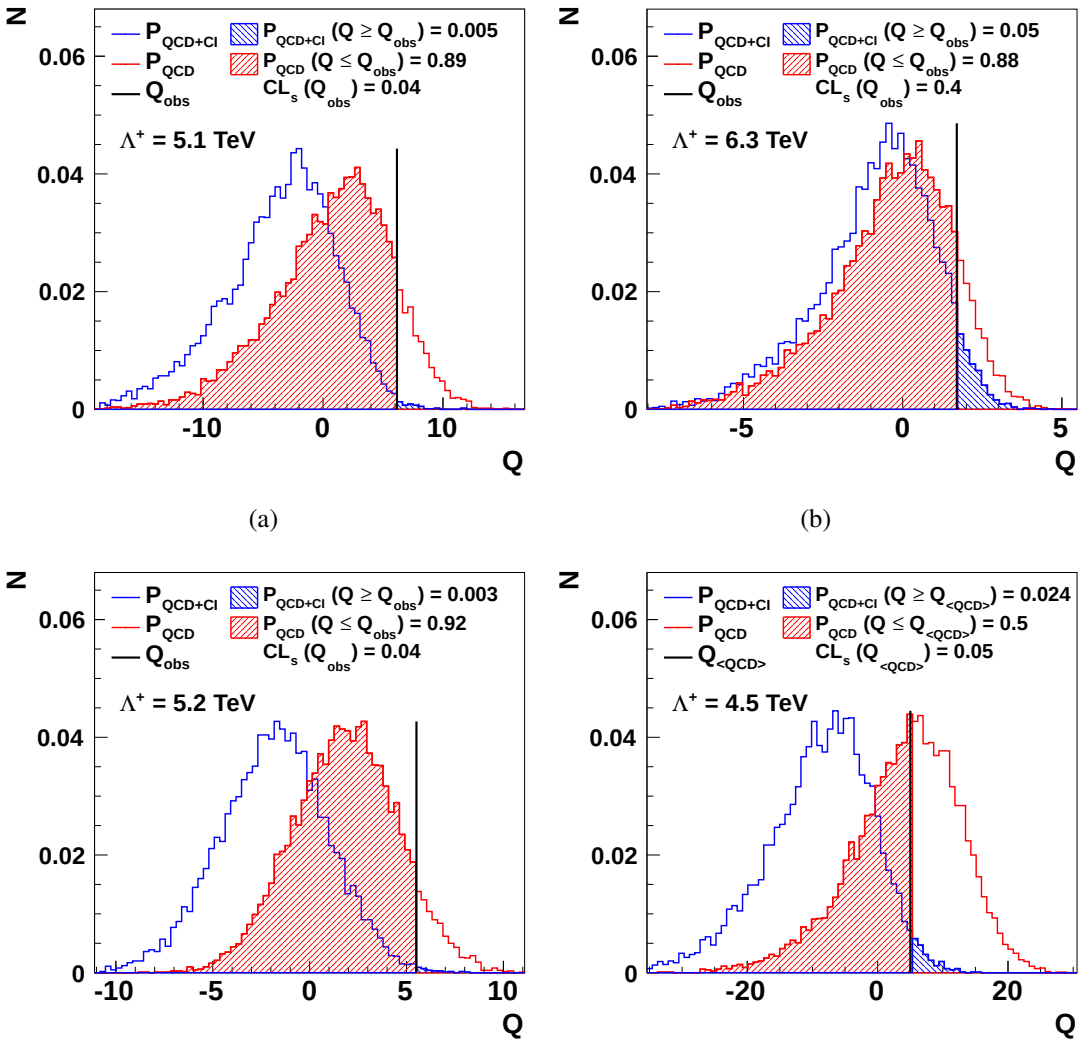


Figure 11.1: Test statistics Q for data (black), for 10000 QCD pseudo-experiments (red) and for 10000 QCD+CI pseudo-experiments (blue) at four different scales of Λ^+. (c) demonstrates the result without systematic uncertainties, while (a,b,d) include systematic uncertainties. (d) demonstrates the expected CL_s where the data test statistics is set to the median of the QCD distribution.

With increasing Λ^+ the two distributions become more and more similar and indistinguishable eventually, as the contact interaction signal vanishes from the analysis mass bins.

The p-values for the two hypotheses $P_{\mathrm{QCD+CI}}(Q \geq Q_{obs})$ and $P_{\mathrm{QCD}}(Q \leq Q_{obs})$ are obtained by integrating over $P_{\mathrm{QCD+CI}}(\mathrm{Q})$ and $P_{\mathrm{QCD}}(Q)$. Here, Q_{obs} is the log-likelihood-ratio observed in data and is indicated as a black lines in Fig. 11.1. $P_{\mathrm{QCD+CI}}(Q \geq Q_{obs})$ is the probability to observe a log-likelihood-ratio Q in the QCD+CI model that is higher than the log-likelihood-ratio Q_{obs} observed in data and $P_{\mathrm{QCD}}(Q \leq Q_{obs})$ is the probability to observe a log-likelihood-ratio Q in the QCD model that is lower than the log-likelihood-ratio Q_{obs} observed in data.

A modified frequentist approach [90–92] based on the quantity

$$CL_s = \frac{P_{QCD+CI}(Q \geq Q_{obs})}{1 - P_{QCD}(Q \leq Q_{obs})} \tag{11.3}$$

is used to make the decision of the hypothesis test. The QCD+CI model is excluded at 95% confidence level if $CL_s < 0.05$. As an example, the model with $\Lambda^+ = 5.1$ TeV in Fig. 11.1 (a) with $CL_s = 0.04$, is very close to the exclusion criterion. Therefore the QCD+CI with higher values of Λ^+ will be excluded, as we will see in the Sec. 11.3. For the evaluation of the ensembles of pseudo-experiments and CL_s, the ROOSTATS [93,94] package and in particular the `HybridCalculator` were used.

The reason for using the $CL_s < 0.05$ instead of the simple p-value $P_{QCD+CI}(Q \geq Q_{obs}) < 0.05$ for the test decision becomes clear when looking at Fig. 11.1 (b). At $\Lambda^+ = 6.3$ TeV, the $P_{QCD+CI}(Q \geq Q_{obs}) = 0.05$ reaches the value for an exclusion. However, in this case, the probability distributions for QCD and QCD+CI are almost indistinguishable, in other words, the data have little sensitivity to discriminate between QCD and QCD+CI. Therefore, the QCD+CI model should not be excluded by the statistical method in this case. The CL_s compensates this problem by dividing by 1 minus the p-value for the QCD prediction. This, however, results in an over-coverage, or in other words the limits at a given confidence level are less stringent than those obtained directly from the p-value of QCD+CI. In summary the use of CL_s prevents an exclusion that could result from statistical fluctuations in situations where one has little sensitivity to new physics but results in over-coverage.

11.2 Treatment of Uncertainties

The uncertainties of the measurement are taken into account as follows. The statistical uncertainties are treated as Poisson distributed around the number of observed data events in each bin in the ensembles of pseudo-experiments. Systematic uncertainties are introduced via Bayesian integration [95] by varying them as nuisance parameters in the ensembles of pseudo-experiments. Note that only shape changing systematic uncertainties which do not change the yield are taken into account since the normalized angular distributions are measured. The systematic uncertainties are modeled using 3 histograms with varied scales of -1σ, 0σ, +1σ of the expected uncertainty, interpolated using cubic splines with linear extrapolation beyond $\pm 1\sigma$. The scale is varied according to a Gaussian prior probability density function, resulting in combination in a Gaussian distribution convoluted with the shape variation induced to the χ_{dijet} distributions. The variation is truncated at 0 events in less than 1% of the cases.

The two leading systematic uncertainties, namely the jet energy scale and the μ_F, μ_R scales, are taken into account. For the scale uncertainty the maximum variations of the shape of the χ_{dijet} distribution are used to model the uncertainty. The +1σ (-1σ) histogram is chosen to have the maximal positive (negative) slope. Note that the choice of the Gaussian prior probability density function is ambiguous in this case, as the scale uncertainty is just a rough estimator for higher order perturbative effects and therefore its distribution is unknown. The jet energy scale uncertainty has been derived independently for each scale of Λ, as it differs from the jet energy

scale effect on the QCD distribution. In this way, the effect of event migration between mass bins it takes into account. As the quark compositeness signal is only prominent in the higher mass bins, event migrations change the prediction in each mass bin, and has to be taken into account.

The effect of the systematic uncertainties on the $P_{\mathrm{QCD+CI}}(\mathrm{Q})$ can be seen in Fig. 11.1 (a) and (c). The distribution in (a) with systematic uncertainties included is slightly broadened compared to the distribution without systematic uncertainties included in (c). Therefore the CL_s value for exclusion is reached at a 0.1 TeV smaller value of Λ^+ when including systematic uncertainties. The effect is very small as the hypothesis test is mainly driven by the highest mass bins which are dominated by statistical uncertainties.

11.3 Limits Setting Procedure

Having defined the exclusion criterion $CL_s < 0.05$ at 95% confidence level, one can exclude parameter ranges for the quark compositeness model. The quark compositeness model is described by a single parameter Λ. As the cross section of the contact interaction in the mass bins of the analysis decreases continuously with increasing Λ, one can set lower limits on the contact interaction scale Λ from the measured dijet angular distributions.

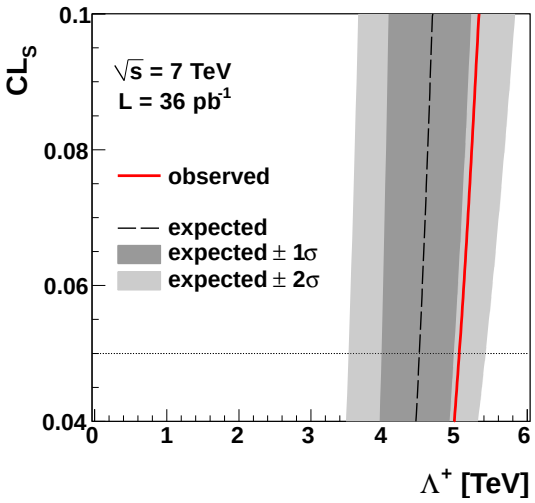


Figure 11.2: Observed CL_s (red line) and expected CL_s (dashed black line) with 1(2) standard deviation bands indicated in dark(light) grey band as a function of the contact interaction scale Λ^+.

The limits are obtained by calculating the CL_s as a function of Λ and finding the crossing point with $CL_s = 0.05$. This is demonstrated in Fig. 11.2. The CL_s as a function of Λ^+ obtained from data is shown as a red line which crosses the dotted $CL_s = 0.05$ line at $\Lambda^+ = 5.1$ TeV. The resulting lower limit on the contact interaction scale Λ^+ is then $\Lambda^+ > 5.1$ TeV.

For this calculation, a continuous model of the dijet angular distributions as a function of Λ is needed. It is constructed by interpolating the angular distributions in each mass and χ_{dijet}

bin between the contact interaction predictions at the scales 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 8 and 9 TeV using cubic spline interpolation. The precision of the interpolation has been evaluated to be better than 0.01 TeV in terms of the limit on Λ by comparing to an interpolation with the 5 TeV prediction removed.

In addition to the limit observed in data, it is important to study also the expected limit given the number of observed data events in each mass bin. The expected limit quantifies the sensitivity of the experiment independent from statistical fluctuations in the data and is therefore important for comparisons with other measurements. The evaluation of the expected CL_s is demonstrated in Fig. 11.1 (d). It is extracted by using the log-likelihood-ratio $Q_{\langle QCD \rangle}$ at the median of the $P_{\mathrm{QCD}}(\mathrm{Q})$ instead of the Q_{obs} of the data. In this way a limit is extracted under the assumption that the data lies at the center of the expectation for the QCD hypothesis. The resulting expected CL_s as a function of Λ^+ is shown in Fig. 11.2 as a black bashed line.

Additional to the expected limit one can also extract the variation of the expected limit within the uncertainties. By using the log-likelihood ratio Q at 16% (2.2%) and 84% (97.7%) values of $P_{\mathrm{QCD}}(\mathrm{Q})$, instead of the median (=50%), 1σ (2σ) bands reflecting the variation of the expected limit, are obtained. They are shown as grey bands in Fig. 11.2.

11.4 Cross checks

This section will summarize various crosschecks which were performed to check the consistency of the analysis. With the limit setting procedure in place, it is possible to check the impact of several effects on the limit to understand their impact on the analysis.

The first crosscheck was performed to verify the validity of the use of the ”correction factor” method to unfold the data. The potential danger of the ”correction factor” method is an underestimation of the statistical uncertainties. For example in an experiment with bad dijet mass resolution, a large number of events may propagate to higher mass bins and reduce the relative statistical uncertainty in those. This effect is not taken into account in the ”correction factor” method. The crosscheck for the validity of the unfolding procedure is to compare the following to two limits:

1. from the unfolded data w.r.t. the theory prediction and
2. from the raw data w.r.t. the detector resolution smeared theory prediction.

In other words, if the statistical evaluation at particle-level and detector level yield the same result, the unfolding procedure does not introduce any bias on the statistical uncertainties. The test was performed using a PYTHIA D6T prediction for QCD and QCD+CI and statistical uncertainties only were taken into the account for the limits. The result was an agreement of the two limits within 0.02 TeV. Therefore, the bias of the unfolding procedure on the limit is negligible.

The analysis was performed using an anti-k_T clustering algorithm with a distance parameter of 0.5. In principle, the choice of the distance parameter has the potential to improve the sensitivity of the analysis. For example by choosing a larger distance parameter, one could collect a larger part of final state radiation and therefore measure higher invariant masses. However as

this is true for both QCD and QCD+CI, it is not clear whether a gain is possible. Therefore the limits on the contact interaction scale for a distance parameter of 0.7 and 0.5 were compared. The resulting limits agree within 0.02 TeV, which shows that the sensitivity of the analysis is practically independent of the choice of the distance parameter.

In order to study the contribution of different mass bins to the limit, the expected limit was derived with different subsets of the mass bins, summarized in Tab. 11.1.

Table 11.1: Expected limits on Λ^+ and Λ^- using different sets of mass bins.

mass bins	expected Λ^+ limit [TeV]	expected Λ^- limit [TeV]
highest	4.4	5.0
two highest	4.5	5.4
four highest	4.5	5.4

From this table one can see that the highest mass bin is responsible for 98% of the Λ^+ limit and 93% of the Λ^- limit. The second highest mass bin contributes 2% to the Λ^+ limit and 7% to the Λ^- limit. The contributions of the third and fourth mass bins are less than 1%. Note that due to the interference term with QCD, the contact interaction with constructive interference (Λ^-) has a higher contribution from the second highest mass bin.

12 Results and Discussion

The results of this analysis are on the one hand a test of perturbative QCD, and on the other hand a search for quark substructure. In the following two sections the two results will be presented and discussed in the context of other measurements and theoretical models.

12.1 Test of Perturbative QCD

In Fig. 12.1 the measured dijet angular distributions are compared to the prediction from perturbative QCD. The corresponding ratio of the data divided by the QCD prediction is shown in Fig. 12.2. The data are corrected for detector effects and are shown with statistical and systematic uncertainties added in quadrature. The QCD prediction is calculated at NLO with non-perturbative corrections applied and shown with the theoretical uncertainties indicated as a grey band. The dijet angular distributions are normalized to the total number of events within each M_{jj} region. Each distribution is shifted by an offset for presentation in a single plot.

Table 12.1: Uncertainties on the shape of the dijet angular distributions in the lowest and highest mass bins.

Uncertainty	$0.25 < M_{jj}$ [TeV] < 0.35	$2.2 < M_{jj}$ [TeV]
Statistical	1.6%	38%
Jet Energy Scale (3-5%)	1.5%	2.3%
Jet Energy Resolution (10%)	1.1%	0.3%
Detector Effects Correction (50% of correction)	1.4%	0.6%
Other Detector Effects (FullSim-GenSmear)	0.7%	0.9%
NLO scale (6 variations of μ_R and μ_F)	5.1%	8.5%
NLO PDF (CTEQ6.6 90CL)	0.5%	0.3%
Non-pert. correction (PYTHIA-HERWIG++)	4.0%	0.02%

A summary of the sources of uncertainty is given in Tab. 12.1. For demonstration, the uncertainties are given for the lowest and highest mass bins. The numbers represent the maximum variation in all χ_{dijet} bins in the mass bin of interest. Note that the largest experimental uncertainties are the jet energy scale and the statistical uncertainty. The largest theoretical uncertainty results from the scale variations. A complete listing of the uncertainties for all χ_{dijet} and mass bins can be found in the Appendix.

This measurement extends the kinematic region of dijet studies by previous experiments [22–26]. The data are found to be in good agreement with the theoretical predictions. The largest

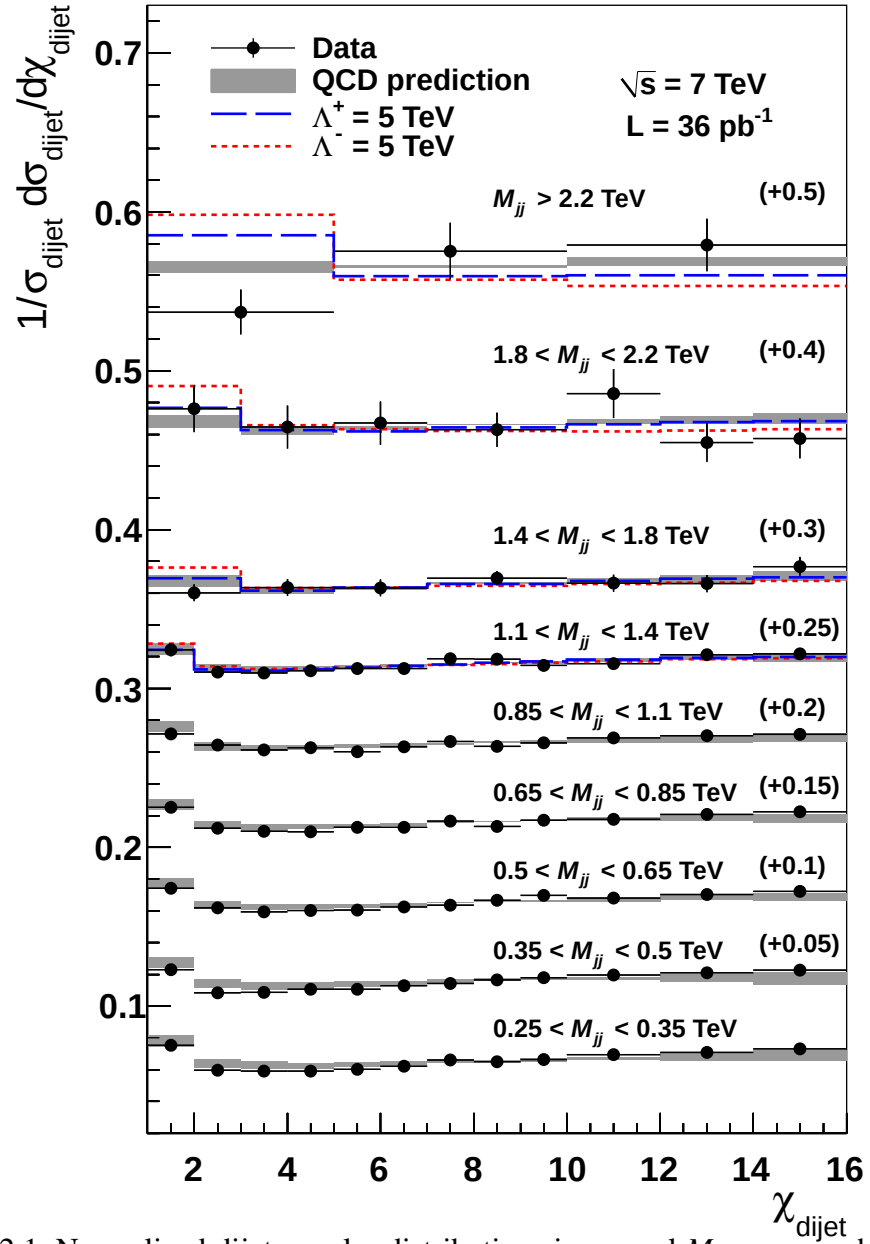


Figure 12.1: Normalized dijet angular distributions in several M_{jj} ranges, shifted vertically by the offsets given in parentheses in the figure for clarity. The data points include statistical and systematic uncertainties. The results are compared with the predictions from QCD at NLO (shaded band) and with the predictions including a contact interaction term at NLO of compositeness scale $\Lambda^+ = 5$ TeV (dashed histogram) and $\Lambda^- = 5$ TeV (dotted histogram). The shaded band shows the uncertainty of the NLO QCD prediction due to μ_R and μ_F scale variations and PDF uncertainties as well as the uncertainties from the non-perturbative corrections added in quadrature.

deviation is found in the highest mass bin and the lowest χ_{dijet} bin. These fluctuations are well within 2 standard deviations of the statistical uncertainty and therefore not significant. In general the data points are at the edge of the theoretical uncertainty band which is correlated between the χ_{dijet} bins. Therefore, this can be understood as an overall deviation of 1 standard

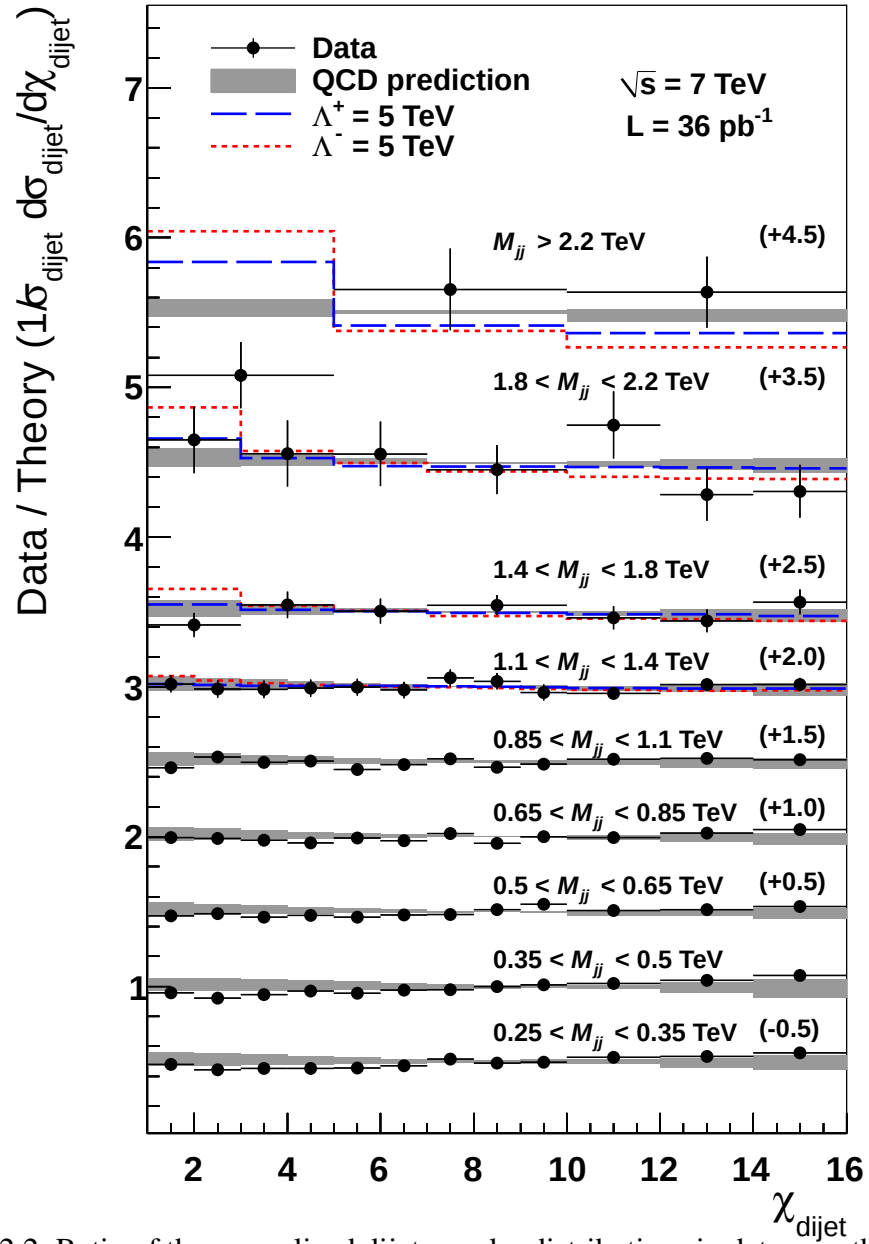


Figure 12.2: Ratio of the normalized dijet angular distributions in data over the QCD prediction in several M_{jj} ranges, shifted vertically by the offsets given in parentheses in the figure for clarity. The details about the content of the plot are given in Fig. 12.1.

deviation, though the theoretical uncertainty cannot be interpreted as Gaussian distributed. In summary the data are well described by the theory.

Due to the low experimental uncertainties in the lower mass bins, this data are already able to constrain the theory. As the analysis is insensitive to the PDFs and the non-perturbative corrections are small with the exception of the lowest mass bin, the data can be used to test or constrain the higher order perturbative terms in QCD. The data could be used to test NNLO QCD calculations as well as to tune showering algorithms of MC event generators which simulate higher order perturbative terms. In Sec. 8.5 it was already shown that the data prefer the p_T ordered showering model implemented in PYTHIA tune Z2 to the old model in tune D6T. It is,

however, important to use a combination of several variables sensitive to the showering algorithm (e.g. M_{jj}) in order to tune it, as it is not clear from first principles if showering algorithms are able to describe higher order terms properly. To make this possible, the data of this analysis have been made public under Ref. [96,97].

12.2 Limits on New Physics

As the data are found to be in good agreement with the QCD predictions, limits are set on the contact interaction scale Λ restricting quark compositeness models. Models of a contact interaction between left-handed quarks with destructive (Λ^+) and constructive (Λ^-) interference with the QCD terms are studied. The lower limits on the contact interaction scale in these models at 95% confidence level ($CL_s < 0.05$) are given in Tab. 12.2. The results are obtained using NLO predictions for the contact interaction terms. Since these predictions became available only recently [89], the LO results which have been published in Ref. [97] are also provided for comparison with previous results.

Table 12.2: Limits on quark compositeness models with a contact interaction between left-handed quarks with destructive (Λ^+) and constructive (Λ^-) interference with the QCD terms.

Model	observed limit [TeV]	expected limit [TeV]
Λ^+ (NLO)	5.1	$4.5^{+0.5}_{-0.5}$
Λ^- (NLO)	6.2	$5.4^{+1.0}_{-0.8}$
Λ^+ (LO)	5.6	$5.0^{+0.4}_{-0.5}$
Λ^- (LO)	6.7	$5.8^{+1.0}_{-1.0}$

The observed limits agree within 1-1.4 standard deviations with the expected limits. They are higher than the expected limits because at $M_{jj} > 2.2$ TeV the measured dijet angular distribution at low χ_{dijet} is lower than, although statistically compatible with, the QCD prediction. Note that in this case, the denominator in $CL_s = (P_{QCD+CI}(Q \geq Q_{obs}))/(1 - P_{QCD}(Q \leq Q_{obs}))$ actually reduces the limit to protect against an exclusion due to the 1-1.4 standard deviation fluctuation where the data have no sensitivity. It is therefore crucial to quote the CL_s limit instead of a simple $P_{QCD+CI}(Q \geq Q_{obs})$.

Figure 12.3 compares the obtained limits on $\Lambda^{\pm}$ calculated at LO to the previously published results [18,22,23,25,26]. The most stringent limit of $\Lambda^+ > 4.0$ TeV was obtained in the analysis of the dijet centrality ratio by CMS with 2.9 pb^{-1} [18]. The limits set by this analysis significantly improve all previously published results.

Note that in contrast to the Tevatron results [23,25], the limits on Λ^+ and Λ^- differ significantly in this analysis. This is due to the fact that the interference term differs for pp-collisions, while it is almost the same for $p\bar{p}$ collisions [27]. This is the first analysis using pp-collisions to set limits on Λ^+ and Λ^- separately.

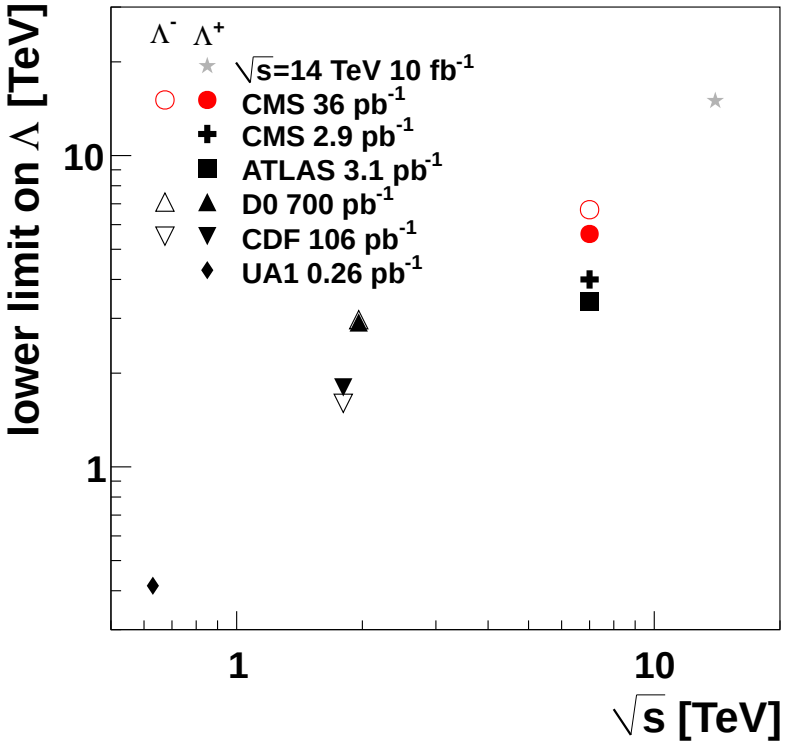


Figure 12.3: Limits on a contact interaction between quarks calculated at LO from various experiments [18,22,23,25,26,97]. The limits obtained in this analysis are marked in red. The expected limit from CMS at $\sqrt{s} = 14$ TeV and 10 fb^{-1} [98] is marked in grey.

Recently, results from the ATLAS collaboration using 36 pb^{-1} including limits on quark compositeness have been made public [99]. The observed and expected limits on Λ^+ at 95% confidence level, requiring $P_{QCD+CI}(Q \geq Q_{obs}) < 0.05$, are $\Lambda >$9.5 TeV and $\Lambda >$5.7 TeV, respectively. As the data fluctuates more than 1 standard deviation from the expected value, the $P_{QCD+CI}(Q \geq Q_{obs}) < 0.05$ limit has to be regarded with care as it is not protected against an exclusion where the data have no sensitivity. A direct comparison of the experiments using only the dijet angular distributions in the highest mass bin and the expected $P_{QCD+CI}(Q \geq Q_{obs}) <$ 0.05 limit, yields an agreement within 0.2 TeV. In other words, the two experiments have almost the same sensitivity when using the same statistical evaluation of the data. As CMS and ATLAS both have a good angular resolution for dijets, no major difference is expected. The remaining difference of 3.7 TeV between the observed limits of the ATLAS result ($\Lambda >$9.5 TeV) and this analysis ($\Lambda >$5.6 TeV) comes from the choice of the statistical method.

Right after the data of this analysis were presented publicly, preliminary results extracting bounds on large extra dimensions were made public [100]. A limit on the coefficient of the effective operator generated by tree-level graviton exchange in large extra dimensions of $M_T >$ 3.4 TeV is extracted from the dijet angular distributions data of this analysis. This shows the importance of publishing the data corrected for all detector effects, so any new theory can be tested with the documented data.

The future reach of contact interaction searches in CMS is indicated in Fig. 12.3 as a grey

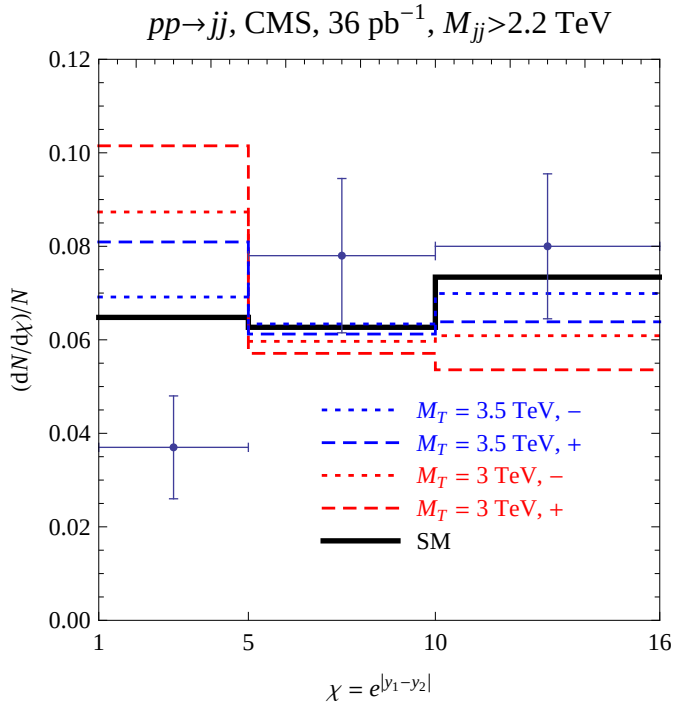


Figure 12.4: Signatures of extra dimensions in the dijet angular distributions.

star. With an increased integrated luminosity of 10 fb^{-1} expected in the year 2012 and the nominal center-of-mass energy of the LHC of $\sqrt{s} = 14$ TeV expected in the year 2014, the expected limit on the contact interaction scale reaches $\Lambda^+ > 15$ TeV. Higher values may also be possible with more integrated luminosity and improved analysis techniques. The sensitivity of the analysis may e.g. be improved by optimizing the granularity of the binning in χ_{dijet} and M_{jj}. So far, CMS has independently explored the two extreme cases of, on the one hand, fine binning in χ_{dijet} and coarse binning in M_{jj} (this analysis), and, on the other hand, fine binning in M_{jj} and only two bins in rapidity which approximately corresponds to exploring two bins in χ_{dijet} (analysis in Ref. [18]). The combination of both approaches may improve sensitivity by the order of 5-10% in terms of the limit on the contact interaction scale.

13 Conclusions and Outlook

In the first year of the LHC at a center-of-mass energy of $\sqrt{s} = 7$ TeV, a large sample of dijet events was recorded with the CMS experiment. Right from the start of the LHC, these events enabled tests of Quantum Chromodynamics (QCD) in a new kinematic regime, and searches for signals of new physics phenomena in a so far unexplored territory. Of particular interest was the search for contact interactions between quarks which would indicate substructure within quarks. Discovery would allow for new physics theories where the values of quark masses, charges and generations may be explained.

In this analysis, the dijet angular distributions, expressed in the variable $\chi_{\text{dijet}} = e^{|y_1 - y_2|}$, were measured in 9 dijet invariant mass bins ranging from 0.25 TeV to >2.2 TeV. The uncertainties of less than 3% in the range $1 < \chi_{\text{dijet}} < 16$ achieved in the lower mass bins allowed for a test of the predictions of perturbative QCD. As the kinematic regime of multi-TeV dijet masses has never been explored before, the results are an important test to clarify whether QCD predictions are valid at these energies. Overall a good agreement with the predictions from perturbative QCD calculated at NLO and corrected for non-perturbative effects is observed, showing the validity of the theory of QCD in the multi-TeV energy regime.

The dijet angular distributions have been used to exclude potential signals from quark compositeness in the data delivered by the LHC in 2010. Limits were set on the energy scale of contact interactions between left-handed quarks. This analysis was the first to use the most accurate NLO predictions for the contact interactions terms and to distinguish constructive and destructive interference terms with QCD in pp-collisions where they have different signatures (as opposed to $p\bar{p}$-collisions). The limits at the 95% confidence level on the contact interaction scale with destructive (Λ^+) and constructive (Λ^-) interference terms with QCD were obtained using a modified frequentist approach and resulted in $\Lambda^+ > 5.1$ TeV and $\Lambda^- > 6.2$ TeV. These values can be compared with the limits of $\Lambda^+ > 4.5$ TeV and $\Lambda^- > 5.4$ TeV expected for the number of events recorded. The limits obtained in this analysis significantly improve all previous measurements at the Tevatron and the LHC.

In the coming years of the LHC, contact interaction searches in dijet events will remain a top priority analysis as the sensitivity to contact interaction scales will increase to values above $\Lambda^+ = 15$ TeV with the LHC operating at its design luminosity and center-of-mass energy.

Appendix

Table 1: Number of data events, detector effects correction factor and dijet angular distributions for data, QCD prediction and contact interaction model prediction.

M_{jj} / χ_{dijet}	N_{raw}	$c_{\text{unsm.}}$	$\frac{1}{N}\frac{dN}{d\chi}$	QCD	Λ^+=5 TeV
250-350					
1-2	4809	0.956	0.0754	0.0771	0.0771
2-3	3819	0.952	0.0597	0.0633	0.0626
3-4	3799	0.948	0.0591	0.0621	0.0622
4-5	3820	0.944	0.0592	0.0622	0.0623
5-6	3907	0.940	0.0603	0.0630	0.0634
6-7	4048	0.936	0.0622	0.0641	0.0646
7-8	4322	0.932	0.0661	0.0651	0.0654
8-9	4273	0.928	0.0651	0.0659	0.0660
9-10	4373	0.924	0.0663	0.0666	0.0663
10-12	9205	0.918	0.0694	0.0675	0.0670
12-14	9474	0.910	0.0708	0.0686	0.0688
14-16	9886	0.902	0.0732	0.0692	0.0694
350-500					
1-2	11266	0.965	0.0731	0.0762	0.0761
2-3	9023	0.962	0.0584	0.0633	0.0631
3-4	9109	0.959	0.0587	0.0621	0.0621
4-5	9443	0.956	0.0607	0.0627	0.0626
5-6	9468	0.953	0.0607	0.0635	0.0636
6-7	9859	0.951	0.0630	0.0645	0.0647
7-8	10074	0.948	0.0642	0.0655	0.0659
8-9	10457	0.945	0.0664	0.0665	0.0669
9-10	10732	0.942	0.0680	0.0673	0.0676
10-12	22048	0.938	0.0695	0.0681	0.0682
12-14	22696	0.932	0.0711	0.0683	0.0679
14-16	23374	0.926	0.0728	0.0679	0.0676
500-650					
1-2	5517	0.969	0.0742	0.0762	0.0764
2-3	4623	0.967	0.0620	0.0629	0.0632
3-4	4437	0.966	0.0594	0.0617	0.0620
4-5	4514	0.964	0.0604	0.0619	0.0620
5-6	4526	0.962	0.0604	0.0628	0.0628
6-7	4692	0.960	0.0625	0.0639	0.0638
7-8	4791	0.958	0.0637	0.0649	0.0647
8-9	5019	0.957	0.0666	0.0657	0.0655
9-10	5263	0.955	0.0697	0.0664	0.0663
10-12	10303	0.952	0.0681	0.0675	0.0677
12-14	10671	0.949	0.0702	0.0692	0.0697
14-16	11023	0.945	0.0723	0.0700	0.0692
650-850					
1-2	3551	0.974	0.0751	0.0754	0.0747
2-3	2943	0.972	0.0621	0.0628	0.0625
3-4	2851	0.971	0.0601	0.0615	0.0613
4-5	2840	0.969	0.0597	0.0623	0.0622
5-6	2983	0.968	0.0626	0.0632	0.0631
6-7	2983	0.966	0.0625	0.0642	0.0641
7-8	3180	0.964	0.0666	0.0652	0.0651
8-9	3028	0.963	0.0633	0.0662	0.0661
9-10	3212	0.961	0.0670	0.0670	0.0671
10-12	6499	0.959	0.0676	0.0681	0.0683
12-14	6812	0.956	0.0706	0.0689	0.0692
14-16	6994	0.952	0.0723	0.0691	0.0693

Table 2: Number of data events, detector effects correction factor and dijet angular distributions for data, QCD prediction and contact interaction model prediction.

M_{jj} / χ_{dijet}	N_{raw}	$c_{\text{unsm.}}$	$\frac{1}{N}\frac{dN}{d\chi}$	QCD	Λ^{+}=5 TeV
850-1100					
1-2	1565	0.976	0.0715	0.0745	0.0742
2-3	1413	0.975	0.0645	0.0624	0.0623
3-4	1344	0.973	0.0612	0.0615	0.0613
4-5	1380	0.971	0.0628	0.0624	0.0629
5-6	1326	0.970	0.0602	0.0634	0.0638
6-7	1396	0.968	0.0633	0.0644	0.0645
7-8	1470	0.967	0.0665	0.0653	0.0652
8-9	1410	0.965	0.0637	0.0661	0.0659
9-10	1460	0.963	0.0658	0.0668	0.0666
10-12	3058	0.961	0.0688	0.0677	0.0675
12-14	3142	0.957	0.0704	0.0688	0.0685
14-16	3183	0.954	0.0711	0.0702	0.0707
1100-1400					
1-2	383	0.978	0.0743	0.0729	0.0734
2-3	312	0.977	0.0604	0.0613	0.0612
3-4	309	0.975	0.0597	0.0607	0.0606
4-5	317	0.974	0.0612	0.0617	0.0619
5-6	325	0.972	0.0626	0.0628	0.0629
6-7	325	0.971	0.0625	0.0639	0.0638
7-8	358	0.969	0.0688	0.0649	0.0648
8-9	357	0.967	0.0685	0.0660	0.0659
9-10	337	0.966	0.0645	0.0670	0.0670
10-12	687	0.963	0.0656	0.0685	0.0686
12-14	750	0.960	0.0714	0.0702	0.0701
14-16	756	0.957	0.0717	0.0707	0.0705
1400-1800					
1-3	140	0.977	0.0603	0.0661	0.0684
3-5	148	0.974	0.0635	0.0607	0.0606
5-7	148	0.971	0.0633	0.0630	0.0626
7-10	244	0.967	0.0693	0.0663	0.0655
10-12	156	0.963	0.0662	0.0688	0.0684
12-14	156	0.960	0.0660	0.0701	0.0696
14-16	182	0.957	0.0767	0.0719	0.0721
1800-2200					
1-3	27	0.978	0.0761	0.0662	0.0791
3-5	23	0.974	0.0645	0.0610	0.0623
5-7	24	0.969	0.0670	0.0635	0.0621
7-10	34	0.964	0.0629	0.0662	0.0638
10-12	31	0.958	0.0856	0.0686	0.0653
12-14	20	0.954	0.0550	0.0701	0.0663
14-16	21	0.949	0.0574	0.0713	0.0693
2200-∞					
1-5	7	0.970	0.0370	0.0636	0.0941
5-10	18	0.960	0.0753	0.0653	0.0572
10-16	23	0.948	0.0792	0.0698	0.0562

Table 3: Statstical, experimental, theoretical and the total uncertainties.

M_{jj} / χ_{dijet}	Stat.[%]	Total experiment[%]	Total theory[%]	Total[%]
250-350				
1-2	±1.44	±2.39	+5.92−3.27	+6.54−4.30
2-3	±1.62	±2.10	+5.15−2.94	+5.79−3.96
3-4	±1.62	±1.81	+4.63−2.67	+5.22−3.61
4-5	±1.62	±1.53	+3.31−2.14	+3.99−3.09
5-6	±1.60	±1.26	+2.65−1.77	+3.34−2.70
6-7	±1.57	±1.02	+1.86−1.34	+2.64−2.31
7-8	±1.52	±0.83	+1.00−0.87	+2.00−1.94
8-9	±1.53	±0.72	+0.49−0.46	+1.76−1.75
9-10	±1.51	±0.77	+0.79−1.89	+1.87−2.54
10-12	±1.04	±1.01	+1.01−1.76	+1.77−2.28
12-14	±1.03	±1.51	+2.46−4.05	+3.06−4.44
14-16	±1.01	±2.09	+3.84−5.57	+4.49−6.03
350-500				
1-2	±0.94	±1.48	+5.84−2.64	+6.10−3.17
2-3	±1.05	±1.31	+5.43−2.68	+5.68−3.16
3-4	±1.05	±1.14	+4.79−2.61	+5.03−3.03
4-5	±1.03	±0.98	+3.76−2.38	+4.02−2.77
5-6	±1.03	±0.84	+3.26−2.22	+3.52−2.59
6-7	±1.01	±0.71	+2.32−1.89	+2.62−2.25
7-8	±1.00	±0.61	+1.80−1.55	+2.14−1.94
8-9	±0.98	±0.55	+1.20−1.18	+1.65−1.63
9-10	±0.97	±0.57	+0.67−1.03	+1.31−1.53
10-12	±0.67	±0.71	+0.74−1.68	+1.23−1.94
12-14	±0.66	±0.99	+2.67−4.38	+2.93−4.54
14-16	±0.65	±1.32	+5.03−7.11	+5.24−7.26
500-650				
1-2	±1.35	±1.42	+5.78−2.22	+6.10−2.95
2-3	±1.47	±1.23	+5.00−1.93	+5.36−2.72
3-4	±1.50	±1.04	+4.25−1.68	+4.62−2.48
4-5	±1.49	±0.86	+3.05−1.25	+3.50−2.13
5-6	±1.49	±0.69	+2.31−1.00	+2.83−1.92
6-7	±1.46	±0.54	+1.52−0.72	+2.17−1.71
7-8	±1.44	±0.42	+0.52−0.39	+1.59−1.55
8-9	±1.41	±0.38	+0.21−0.22	+1.48−1.48
9-10	±1.38	±0.43	+0.27−0.80	+1.47−1.65
10-12	±0.99	±0.63	+0.92−2.41	+1.49−2.68
12-14	±0.97	±0.97	+1.31−3.11	+1.89−3.40
14-16	±0.95	±1.34	+2.16−4.89	+2.72−5.16
650-850				
1-2	±1.68	±1.76	+6.11−2.54	+6.58−3.51
2-3	±1.84	±1.52	+5.34−2.18	+5.85−3.24
3-4	±1.87	±1.29	+4.27−1.75	+4.84−2.87
4-5	±1.88	±1.07	+3.16−1.31	+3.83−2.52
5-6	±1.83	±0.86	+2.60−1.06	+3.29−2.29
6-7	±1.83	±0.68	+1.37−0.58	+2.39−2.04
7-8	±1.77	±0.57	+0.57−0.26	+1.95−1.88
8-9	±1.82	±0.53	+0.17−0.40	+1.90−1.93
9-10	±1.76	±0.61	+0.24−0.48	+1.88−1.93
10-12	±1.24	±0.86	+0.88−2.14	+1.75−2.62
12-14	±1.21	±1.28	+1.43−3.37	+2.27−3.80
14-16	±1.20	±1.74	+2.22−5.30	+3.07−5.70

Table 4: Statstical, experimental, theoretical and the total uncertainties.

M_{jj} / χ_{dijet}	Stat.[%]	Total experiment[%]	Total theory[%]	Total[%]
850-1100				
1-2	±2.53	±2.28	+6.54−2.34	+7.37−4.13
2-3	±2.66	±2.00	+5.69−2.07	+6.59−3.92
3-4	±2.73	±1.73	+4.54−1.73	+5.57−3.66
4-5	±2.69	±1.47	+3.37−1.42	+4.56−3.38
5-6	±2.75	±1.22	+2.58−1.23	+3.96−3.25
6-7	±2.68	±1.00	+1.62−0.99	+3.29−3.02
7-8	±2.61	±0.83	+0.88−0.78	+2.87−2.85
8-9	±2.66	±0.74	+0.66−0.92	+2.84−2.91
9-10	±2.62	±0.79	+0.86−2.04	+2.87−3.41
10-12	±1.81	±0.99	+0.77−2.22	+2.20−3.03
12-14	±1.78	±1.44	+1.69−4.04	+2.85−4.65
14-16	±1.77	±1.97	+2.50−4.50	+3.64−5.22
1100-1400				
1-2	±5.11	±2.50	+7.24−2.56	+9.21−6.24
2-3	±5.66	±2.18	+6.14−2.13	+8.63−6.43
3-4	±5.69	±1.87	+5.03−1.73	+7.82−6.23
4-5	±5.62	±1.57	+3.58−1.21	+6.84−5.96
5-6	±5.55	±1.29	+2.54−0.86	+6.24−5.76
6-7	±5.55	±1.05	+1.89−0.68	+5.96−5.69
7-8	±5.29	±0.89	+0.29−0.19	+5.37−5.36
8-9	±5.29	±0.82	+0.17−0.21	+5.36−5.36
9-10	±5.45	±0.90	+0.56−1.49	+5.55−5.72
10-12	±3.82	±1.20	+0.85−2.45	+4.09−4.69
12-14	±3.65	±1.76	+1.14−3.42	+4.21−5.30
14-16	±3.64	±2.38	+2.08−5.88	+4.82−7.31
1400-1800				
1-3	±8.45	±2.64	+7.70−2.85	+11.74−9.30
3-5	±8.22	±1.88	+4.88−1.76	+9.74−8.61
5-7	±8.22	±1.16	+2.24−0.79	+8.60−8.34
7-10	±6.40	±0.59	+0.41−0.62	+6.44−6.46
10-12	±8.01	±1.14	+0.80−2.22	+8.13−8.39
12-14	±8.01	±1.86	+1.83−4.98	+8.42−9.61
14-16	±7.41	±2.62	+2.23−6.13	+8.17−9.97
1800-2200				
1-3	±19.25	±2.74	+8.93−3.19	+21.39−19.70
3-5	±20.85	±1.92	+5.43−1.85	+21.63−21.02
5-7	±20.41	±1.16	+2.73−0.92	+20.63−20.47
7-10	±17.15	±0.49	+0.32−0.73	+17.16−17.17
10-12	±17.96	±1.19	+0.96−2.80	+18.03−18.22
12-14	±22.36	±1.99	+1.75−5.10	+22.52−23.02
14-16	±21.82	±2.81	+2.36−6.93	+22.13−23.07
2200-∞				
1-5	±37.80	±2.55	+8.54−3.02	+38.83−38.00
5-10	±23.57	±1.08	+1.44−0.86	+23.64−23.61
10-16	±20.85	±1.95	+2.03−5.90	+21.04−21.76

Table 5: Experimental uncertainties from jet energy scale, jet energy resolution, detector resolution correction and other detector effects and the total experimental uncertainties.

M_{jj} / χ_{dijet}	JES[%]	JER[%]	Correction [%]	Other [%]	Total experiment[%]
250-350					
1-2	±1.46	±1.07	±1.39	±0.69	±2.39
2-3	±1.28	±0.93	±1.19	±0.69	±2.10
3-4	±1.10	±0.78	±0.99	±0.69	±1.81
4-5	±0.91	±0.63	±0.80	±0.69	±1.53
5-6	±0.73	±0.48	±0.60	±0.69	±1.26
6-7	±0.55	±0.33	±0.40	±0.69	±1.02
7-8	±0.37	±0.19	±0.20	±0.69	±0.83
8-9	±0.18	±0.06	±0.00	±0.69	±0.72
9-10	±0.15	±0.22	±0.20	±0.69	±0.77
10-12	±0.30	±0.46	±0.50	±0.69	±1.01
12-14	±0.63	±0.78	±0.89	±0.69	±1.51
14-16	±1.00	±1.10	±1.29	±0.69	±2.09
350-500					
1-2	±0.37	±0.87	±1.00	±0.54	±1.48
2-3	±0.32	±0.77	±0.85	±0.54	±1.31
3-4	±0.28	±0.66	±0.71	±0.54	±1.14
4-5	±0.23	±0.55	±0.57	±0.54	±0.98
5-6	±0.18	±0.44	±0.43	±0.54	±0.84
6-7	±0.13	±0.33	±0.28	±0.54	±0.71
7-8	±0.08	±0.22	±0.14	±0.54	±0.61
8-9	±0.04	±0.11	±0.00	±0.54	±0.55
9-10	±0.06	±0.12	±0.14	±0.54	±0.57
10-12	±0.10	±0.28	±0.36	±0.54	±0.71
12-14	±0.19	±0.49	±0.64	±0.54	±0.99
14-16	±0.29	±0.71	±0.93	±0.54	±1.32
500-650					
1-2	±0.84	±0.89	±0.62	±0.37	±1.42
2-3	±0.71	±0.76	±0.53	±0.37	±1.23
3-4	±0.58	±0.64	±0.44	±0.37	±1.04
4-5	±0.46	±0.52	±0.35	±0.37	±0.86
5-6	±0.33	±0.40	±0.26	±0.37	±0.69
6-7	±0.20	±0.28	±0.18	±0.37	±0.54
7-8	±0.08	±0.16	±0.09	±0.37	±0.42
8-9	±0.05	±0.03	±0.00	±0.37	±0.38
9-10	±0.18	±0.10	±0.09	±0.37	±0.43
10-12	±0.37	±0.27	±0.22	±0.37	±0.63
12-14	±0.62	±0.51	±0.40	±0.37	±0.97
14-16	±0.87	±0.76	±0.57	±0.37	±1.34
650-850					
1-2	±1.49	±0.56	±0.56	±0.50	±1.76
2-3	±1.26	±0.50	±0.48	±0.50	±1.52
3-4	±1.04	±0.44	±0.40	±0.50	±1.29
4-5	±0.81	±0.38	±0.32	±0.50	±1.07
5-6	±0.58	±0.32	±0.24	±0.50	±0.86
6-7	±0.36	±0.26	±0.16	±0.50	±0.68
7-8	±0.17	±0.20	±0.08	±0.50	±0.57
8-9	±0.10	±0.13	±0.00	±0.50	±0.53
9-10	±0.32	±0.10	±0.08	±0.50	±0.61
10-12	±0.66	±0.13	±0.20	±0.50	±0.86
12-14	±1.11	±0.17	±0.36	±0.50	±1.28
14-16	±1.57	±0.26	±0.52	±0.50	±1.74

Table 6: Experimental uncertainties from jet energy scale, jet energy resolution, detector resolution correction and other detector effects, and the total experimental uncertainties.

M_{jj} / χ_{dijet}	JES[%]	JER[%]	Correction [%]	Other [%]	Total experiment[%]
850-1100					
1-2	±2.04	±0.41	±0.58	±0.71	±2.28
2-3	±1.76	±0.38	±0.50	±0.71	±2.00
3-4	±1.48	±0.34	±0.41	±0.71	±1.73
4-5	±1.20	±0.31	±0.33	±0.71	±1.47
5-6	±0.92	±0.27	±0.25	±0.71	±1.22
6-7	±0.64	±0.24	±0.17	±0.71	±1.00
7-8	±0.35	±0.21	±0.08	±0.71	±0.83
8-9	±0.08	±0.17	±0.00	±0.71	±0.74
9-10	±0.31	±0.14	±0.08	±0.71	±0.79
10-12	±0.65	±0.09	±0.21	±0.71	±0.99
12-14	±1.19	±0.02	±0.37	±0.71	±1.44
14-16	±1.75	±0.05	±0.54	±0.71	±1.97
1100-1400					
1-2	±2.29	±0.08	±0.56	±0.82	±2.50
2-3	±1.96	±0.06	±0.48	±0.82	±2.18
3-4	±1.63	±0.04	±0.40	±0.82	±1.87
4-5	±1.30	±0.03	±0.32	±0.82	±1.57
5-6	±0.97	±0.02	±0.24	±0.82	±1.29
6-7	±0.64	±0.03	±0.16	±0.82	±1.05
7-8	±0.33	±0.05	±0.08	±0.82	±0.89
8-9	±0.03	±0.07	±0.00	±0.82	±0.82
9-10	±0.35	±0.09	±0.08	±0.82	±0.90
10-12	±0.84	±0.12	±0.20	±0.82	±1.20
12-14	±1.50	±0.17	±0.36	±0.82	±1.76
14-16	±2.16	±0.21	±0.52	±0.82	±2.38
1400-1800					
1-3	±2.49	±0.41	±0.51	±0.58	±2.64
3-5	±1.73	±0.26	±0.35	±0.58	±1.88
5-7	±0.98	±0.11	±0.20	±0.58	±1.16
7-10	±0.03	±0.10	±0.00	±0.58	±0.59
10-12	±0.92	±0.27	±0.20	±0.58	±1.14
12-14	±1.68	±0.42	±0.35	±0.58	±1.86
14-16	±2.44	±0.56	±0.51	±0.58	±2.62
1800-2200					
1-3	±2.59	±0.26	±0.72	±0.47	±2.74
3-5	±1.79	±0.13	±0.50	±0.47	±1.92
5-7	±1.03	±0.03	±0.28	±0.47	±1.16
7-10	±0.07	±0.15	±0.00	±0.47	±0.49
10-12	±1.02	±0.30	±0.28	±0.47	±1.19
12-14	±1.82	±0.42	±0.50	±0.47	±1.99
14-16	±2.62	±0.55	±0.72	±0.47	±2.81
2200-∞					
1-5	±2.28	±0.25	±0.61	±0.93	±2.55
5-10	±0.53	±0.08	±0.11	±0.93	±1.08
10-16	±1.61	±0.30	±0.50	±0.93	±1.95

Table 7: Theoretical uncertainties from scale variations, pdf uncertainties and non-perturbative corrections, and the total theoretical uncertainties.

M_{jj} / χ_{dijet}	Scale[%]	PDF[%]	NPC[%]	Total theory[%]
250-350				
1-2	+5.14−1.95	+0.53−0.45	+2.89−2.59	+5.92−3.27
2-3	+4.38−1.65	+0.43−0.36	+2.69−2.41	+5.15−2.94
3-4	+3.92−1.50	+0.34−0.28	+2.44−2.19	+4.63−2.67
4-5	+2.52−0.92	+0.25−0.21	+2.14−1.92	+3.31−2.14
5-6	+1.94−0.73	+0.15−0.12	+1.80−1.61	+2.65−1.77
6-7	+1.22−0.46	+0.09−0.07	+1.41−1.26	+1.86−1.34
7-8	+0.25−0.06	+0.03−0.02	+0.97−0.87	+1.00−0.87
8-9	+0.05−0.15	+0.05−0.06	+0.48−0.43	+0.49−0.46
9-10	+0.78−1.88	+0.10−0.11	+0.05−0.06	+0.79−1.89
10-12	+0.51−1.47	+0.15−0.18	+0.86−0.96	+1.01−1.76
12-14	+1.25−3.29	+0.24−0.29	+2.10−2.34	+2.46−4.05
14-16	+1.44−3.91	+0.29−0.35	+3.55−3.96	+3.84−5.57
350-500				
1-2	+5.51−1.96	+0.49−0.41	+1.87−1.71	+5.84−2.64
2-3	+4.97−1.78	+0.38−0.32	+2.15−1.98	+5.43−2.68
3-4	+4.19−1.50	+0.29−0.24	+2.31−2.12	+4.79−2.61
4-5	+2.94−1.02	+0.21−0.17	+2.33−2.14	+3.76−2.38
5-6	+2.38−0.85	+0.14−0.12	+2.23−2.05	+3.26−2.22
6-7	+1.16−0.39	+0.07−0.05	+2.01−1.84	+2.32−1.89
7-8	+0.69−0.25	+0.04−0.03	+1.66−1.52	+1.80−1.55
8-9	+0.19−0.46	+0.04−0.04	+1.19−1.09	+1.20−1.18
9-10	+0.30−0.87	+0.09−0.11	+0.59−0.54	+0.67−1.03
10-12	+0.52−1.57	+0.13−0.15	+0.51−0.55	+0.74−1.68
12-14	+1.27−3.56	+0.23−0.27	+2.34−2.55	+2.67−4.38
14-16	+1.77−4.92	+0.25−0.30	+4.70−5.12	+5.03−7.11
500-650				
1-2	+5.68−1.96	+0.44−0.36	+0.98−0.97	+5.78−2.22
2-3	+4.90−1.68	+0.33−0.27	+0.93−0.92	+5.00−1.93
3-4	+4.15−1.43	+0.24−0.20	+0.87−0.86	+4.25−1.68
4-5	+2.94−0.98	+0.17−0.14	+0.78−0.77	+3.05−1.25
5-6	+2.21−0.74	+0.10−0.08	+0.67−0.66	+2.31−1.00
6-7	+1.42−0.48	+0.05−0.04	+0.54−0.53	+1.52−0.72
7-8	+0.35−0.09	+0.01−0.01	+0.38−0.38	+0.52−0.39
8-9	+0.02−0.07	+0.03−0.04	+0.20−0.20	+0.21−0.22
9-10	+0.26−0.79	+0.06−0.08	+0.01−0.01	+0.27−0.80
10-12	+0.85−2.39	+0.10−0.12	+0.33−0.33	+0.92−2.41
12-14	+0.98−2.98	+0.18−0.22	+0.85−0.86	+1.31−3.11
14-16	+1.59−4.66	+0.21−0.26	+1.45−1.47	+2.16−4.89
650-850				
1-2	+5.92−2.03	+0.40−0.32	+1.48−1.48	+6.11−2.54
2-3	+5.18−1.77	+0.28−0.22	+1.26−1.26	+5.34−2.18
3-4	+4.14−1.39	+0.19−0.15	+1.04−1.04	+4.27−1.75
4-5	+3.05−1.00	+0.13−0.10	+0.83−0.83	+3.16−1.31
5-6	+2.52−0.87	+0.10−0.08	+0.61−0.61	+2.60−1.06
6-7	+1.31−0.42	+0.04−0.03	+0.40−0.40	+1.37−0.58
7-8	+0.54−0.17	+0.02−0.02	+0.19−0.19	+0.57−0.26
8-9	+0.16−0.40	+0.02−0.03	+0.01−0.01	+0.17−0.40
9-10	+0.10−0.42	+0.03−0.04	+0.22−0.22	+0.24−0.48
10-12	+0.70−2.07	+0.10−0.13	+0.53−0.53	+0.88−2.14
12-14	+1.07−3.23	+0.13−0.17	+0.94−0.94	+1.43−3.37
14-16	+1.76−5.12	+0.18−0.23	+1.34−1.34	+2.22−5.30

Table 8: Theoretical uncertainties from scale variations, pdf uncertainties and non-perturbative corrections, and the total theoretical uncertainties.

M_{jj} / χ_{dijet}	Scale[%]	PDF[%]	NPC[%]	Total theory[%]
850-1100				
1-2	+6.51−2.26	+0.36−0.28	+0.52−0.52	+6.54−2.34
2-3	+5.64−1.94	+0.22−0.16	+0.72−0.72	+5.69−2.07
3-4	+4.45−1.50	+0.12−0.09	+0.85−0.86	+4.54−1.73
4-5	+3.24−1.06	+0.08−0.06	+0.92−0.93	+3.37−1.42
5-6	+2.40−0.79	+0.05−0.04	+0.93−0.94	+2.58−1.23
6-7	+1.37−0.43	+0.04−0.04	+0.88−0.89	+1.62−0.99
7-8	+0.44−0.12	+0.05−0.05	+0.76−0.77	+0.88−0.78
8-9	+0.30−0.71	+0.07−0.08	+0.58−0.59	+0.66−0.92
9-10	+0.78−2.01	+0.03−0.04	+0.34−0.35	+0.86−2.04
10-12	+0.75−2.22	+0.08−0.10	+0.14−0.14	+0.77−2.22
12-14	+1.34−3.91	+0.07−0.09	+1.02−1.01	+1.69−4.04
14-16	+1.23−3.94	+0.10−0.13	+2.18−2.15	+2.50−4.50
1100-1400				
1-2	+7.23−2.54	+0.37−0.28	+0.22−0.21	+7.24−2.56
2-3	+6.13−2.11	+0.18−0.13	+0.25−0.24	+6.14−2.13
3-4	+5.02−1.71	+0.10−0.07	+0.26−0.26	+5.03−1.73
4-5	+3.57−1.18	+0.07−0.06	+0.27−0.26	+3.58−1.21
5-6	+2.53−0.82	+0.06−0.05	+0.26−0.25	+2.54−0.86
6-7	+1.88−0.64	+0.05−0.05	+0.23−0.23	+1.89−0.68
7-8	+0.22−0.03	+0.02−0.02	+0.19−0.19	+0.29−0.19
8-9	+0.06−0.14	+0.08−0.09	+0.13−0.13	+0.17−0.21
9-10	+0.55−1.48	+0.02−0.03	+0.07−0.06	+0.56−1.49
10-12	+0.84−2.45	+0.05−0.07	+0.06−0.06	+0.85−2.45
12-14	+1.10−3.41	+0.06−0.08	+0.28−0.29	+1.14−3.42
14-16	+2.00−5.85	+0.09−0.12	+0.56−0.57	+2.08−5.88
1400-1800				
1-3	+7.64−2.68	+0.28−0.21	+0.95−0.93	+7.70−2.85
3-5	+4.83−1.63	+0.09−0.09	+0.66−0.64	+4.88−1.76
5-7	+2.20−0.70	+0.08−0.08	+0.36−0.35	+2.24−0.79
7-10	+0.40−0.61	+0.08−0.09	+0.00−0.00	+0.41−0.62
10-12	+0.72−2.19	+0.04−0.05	+0.35−0.36	+0.80−2.22
12-14	+1.72−4.94	+0.05−0.06	+0.62−0.64	+1.83−4.98
14-16	+2.04−6.06	+0.07−0.09	+0.90−0.92	+2.23−6.13
1800-2200				
1-3	+8.92−3.16	+0.33−0.27	+0.30−0.30	+8.93−3.19
3-5	+5.43−1.83	+0.12−0.13	+0.21−0.21	+5.43−1.85
5-7	+2.72−0.89	+0.15−0.17	+0.11−0.11	+2.73−0.92
7-10	+0.31−0.73	+0.08−0.10	+0.00−0.00	+0.32−0.73
10-12	+0.95−2.80	+0.06−0.08	+0.11−0.11	+0.96−2.80
12-14	+1.74−5.09	+0.09−0.08	+0.20−0.20	+1.75−5.10
14-16	+2.34−6.93	+0.10−0.10	+0.29−0.29	+2.36−6.93
2200-∞				
1-5	+8.54−3.00	+0.32−0.31	+0.02−0.02	+8.54−3.02
5-10	+1.43−0.84	+0.16−0.19	+0.00−0.00	+1.44−0.86
10-16	+2.03−5.90	+0.13−0.11	+0.02−0.02	+2.03−5.90

Bibliography

[1] F. Halzen and A. D. Martin, "Quarks and Leptons". John Wiley and Sons, 1984.

[2] I. Dremin and A. Kaidalov, "Quantum chromodynamics and phenomenology of strong interactions", *Physics Uspekhi* **49 (3)** (2006) 263–273.

[3] K. Nakamura et al., "The Review of Particle Physics", *J. Phys. G* **37** (2010) 075021.

[4] The TEVNPH Working Group, "Combined CDF and D0 Upper Limits on Standard Model Higgs Boson Production with up to 8.2/fb of Data", *FERMILAB-CONF-11-044-E* (2011).

[5] CMS Collaboration, "CMS Physics Technical Design Report Volume I: Detector Performance and Software", *CERN-LHCC-2006-001* (2006).

[6] P. Nadolsky et al., "Implications of CTEQ global analysis for collider observables", *Phys. Rev.* **D78** (2008) 013004, `arXiv:0802.0007`.

[7] A. Martin et al., "Parton distributions for the LHC", *Eur.Phys.J.C* **63** (2009) 189–285.

[8] D. Gross, "Twenty Five Years of Asymptotic Freedom", `arXiv:hep-th/9809060`.

[9] W. B. et al. (JADE Collaboration), "Observation of planar three-jet events in e+e annihilation and evidence for gluon bremsstrahlung", *Physics Letters B* **91** (1980) 142.

[10] D. B. et al. (MARK J Collaboration), "Discovery of Three-Jet Events and a Test of Quantum Chromodynamics at PETRA", *Physical Review Letters* **43** (1979) 830.

[11] S. Bethke, "Eleven Years of QCD at LEP", *Eur.Phys.J.direct C4* **1** (2002).

[12] A. Warburton, "QCD Results from the Fermilab Tevatron proton-antiproton Collider", *PoS(HCP2009)* **013** (2009).

[13] CMS Collaboration, "Jets in 0.9 and 2.36 TeV pp Collisions", *CMS Physics Analysis Summary* **JME-10-001** (2010).

[14] CMS Collaboration, "Jet Performance in pp Collisions at $\sqrt{s} = 7\ TeV$", *CMS Physics Analysis Summary* **JME-10-003** (2010).

[15] CMS Collaboration, "Commissioning of TrackJets in pp Collisions at $\sqrt{s} = 7\ TeV$", *CMS Physics Analysis Summary* **JME-10-006** (2010).

[16] CMS Collaboration, "Jet Energy Corrections determination at $\sqrt{s} = 7\ TeV$", *CMS Physics Analysis Summary* **JME-10-010** (2010).

[17] CMS Collaboration, "Search for Dijet Resonances in $\sqrt{s} = 7\ TeV$ pp Collisions at CMS", *Phys. Rev. Lett.* **105** (2010) 211801.

[18] CMS Collaboration, "Search for Quark Compositeness with the Dijet Centrality Ratio in pp Collisions at $\sqrt{s} = 7\ TeV$", *Phys. Rev. Lett.* **105** (2010) 262001.

[19] CMS Collaboration, "Dijet Azimuthal Decorrelations in pp Collisions at $\sqrt{s} = 7\ TeV$", *Phys. Rev. Lett.* **106** (2011) 122003.

[20] CMS Collaboration, "First measurement of hadronic event shapes in pp collisions at $\sqrt{s} = 7\ TeV$", *Physics Letters B* **699** (2011), no. 1-2, 48–67.

[21] CMS Collaboration, "Measurement of the Inclusive Jet Cross Section in pp Collisions at $\sqrt{s} = 7\ TeV$", *CMS Physics Analysis Summary* **QCD-10-011** (2010).

[22] UA1 Collaboration, "Angular distributions for High Mass Jet Pairs and a Limit on the Energy Scale of Compositeness for Quarks from the CERN $p\bar{p}$ Collider", *Phys. Lett. B* **177** (1986).

[23] DØ Collaboration, "Measurement of Dijet Angular Distributions at $\sqrt{s} = 1.96\ TeV$ and Searches for Quark Compositeness and Extra Spatial Dimensions", *Phys. Rev. Lett.* **103** (2009) 191803.

[24] DØ Collaboration, "High-p_T jets in $p\overline{p}$ collisions at $\sqrt{s} = 630$ and 1800 GeV", *Phys. Rev. D* **64** (2001) 032003.

[25] CDF Collaboration, "Measurement of Dijet Angular Distributions at CDF", *[Erratum-ibid.* **78***, 4307 (1997)] Phys. Rev. Lett.* **77** (1996) 5336.

[26] ATLAS Collaboration, "Search for Quark Contact Interactions in Dijet Angular Distributions in pp Collisions at $\sqrt{s} = 7\ TeV$ Measured with the ATLAS Detector", *Phys. Lett. B* **694** (2010) 327–345, arXiv:1009.5069.

[27] E. Eichten et al., "Supercollider physics", *Rev. Mod. Phys.* **56** (1984) 4.

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

[29] "The CERN accelerator complex", http://public.web.cern.ch/public/en/research/AccelComplex--en.html.

[30] "CMS Public Luminosity Results", https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults2010.

[31] CMS Collaboration, "Measurement of CMS Luminosity", *CMS Physics Analysis Summary* **EWK-10-004** (2010).

[32] CMS Collaboration, "Absolute luminosity normalization", *CMS Detector Performance Summaries* **DP-2011-002** (2011).

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

[34] G. Jarlskog and D. Rein, eds., "Large Hadron Collider Workshop: Proceedings Vol. I-III". 1990.

[35] CMS Collaboration, "Tracking and Primary Vertex Results in First 7 TeV Collisions", *CMS Physics Analysis Summary* **TRK-10-005** (2010).

[36] CMS Collaboration, "Measurement of Momentum Scale and Resolution using Low-mass Resonances and Cosmic Ray Muons", *CMS Physics Analysis Summary* **TRK-10-004** (2010).

[37] CMS Collaboration, "Electromagnetic calorimeter calibration with 7 TeV data", *CMS Physics Analysis Summary* **EGM-10-003** (2010).

[38] CMS Collaboration, "Single-Particle Response in the CMS Calorimeters", *CMS Physics Analysis Summary* **JME-10-008** (2010).

[39] CMS Collaboration, "Performance of muon identification in pp collisions at $\sqrt{s} = 7$ *TeV*", *CMS Physics Analysis Summary* **MUO-10-002** (2010).

[40] M. Dobbs et al., "Les Houches Guidebook to Monte Carlo Generators for Hadron Collider Physics", *proc. "Physics at TeV Colliders", Les Houches, France, May 2003* (2004).

[41] T. Sjostrand, S. Mrenna, and P. Skands, "PYTHIA 6.4 Physics and Manual", *JHEP* **05** (2006) 026, arXiv:hep-ph/0603175.

[42] M. Bahr et al., "Herwig++ Physics and Manual", *Eur. Phys. J. C* **58** (2008) 639.

[43] M. Dobbs and J. B. Hansen, "The HepMC C++ Monte Carlo Event Record for High Energy Physics", *Comput. Phys. Commun.* **134** (2001) 41.

[44] T. Sjostrand and P. Skands, "Transverse-momentum-ordered showers and interleaved multiple interactions", *Eur. Phys. J. C* **39** (2005) 129.

[45] S. Gieseke, P. Stephens, and B. Webber, "New Formalism for QCD Parton Showers", *JHEP* **0312** (2003) 045.

[46] R. Field, "Early LHC Underlying Event Data-Findings and Surprises", arXiv:1010.3558.

[47] "The PYTHIA6 Z2 tune is identical to the Z1 tune described in [46] except that Z2 uses the CTEQ6L PDF while Z1 uses CTEQ5L.",.

[48] J. Butterworth, J. Forshaw, and M. Seymour, "Multi-Parton Interactions in Photoproduction at HERA", *Z. Phys. C* **72** (1996) 637–646.

[49] B. Andersson, G. Gustafson, G. Ingelman et al., "Parton fragmentation and string dynamics", *Phys. Rept.* **97** (1983) 31.

[50] B. R. Webber, "A QCD Model for Jet Fragmentation including Soft Gluon Interference", *Nucl. Phys. B* **238** (1984) 492.

[51] GEANT4 Collaboration (S. Agostinelli et al.), "GEANT4: A simulation toolkit", *Instrum. and Methods* **A506** (2003) 250–303.

[52] R. Brun and F. Rademakers, "ROOT - An Object Oriented Data Analysis Framework", *Nucl. Inst. & Meth. in Phys. Res. A* **398** (1996) 81–86 http://root.cern.ch.

[53] CMS Collaboration, "The CMS computing model", *CERN-LHCC-2004-035/G-083* (2004).

[54] The LCG TDR Editorial Board, "LCG Computing TDR", *CERN-LHCC-2005-024* (2005).

[55] CMS Collaboration, "CMS Computing TDR", *CERN-LHCC-2005-023* (2005).

[56] R. Wolf, "A Tour Through the CMS Data Model", *Proc. Ninth International Conference of Computational Methods in Sciences and Engineering (ICCMSE 2009), Rhodes, Greece* (2009).

[57] A. Hinzmann, "Tools for Physics Analysis in CMS", *Proc. Computing in High Energy Physics (CHEP2010), Taipei, Taiwan* (2010).

[58] C. Jones et al., "File Level Provenance Tracking in CMS", *Proc. 17th International Conference on Computing in High Energy and Nuclear Physics (CHEP 2009), Prague, Czech Republic, J. Phys.: Conf. Ser.* **219** (2009) 032011.

[59] C. Jones, L. Lista, and G. Petrucciani, "Expression and cut parser for CMS event", *Proc. 17th International Conference on Computing in High Energy and Nuclear Physics (CHEP 2009), Prague, Czech Republic, J. Phys.: Conf. Ser.* **219** (2009) 032050.

[60] A. Hinzmann, "Visualization of the CMS Python Configuration System", *Proc. 17th International Conference on Computing in High Energy and Nuclear Physics (CHEP 2009), Prague, Czech Republic, J. Phys.: Conf. Ser.* **219** (2009) 042008.

[61] A. de Cosa, "CMS Configuration Editor: GUI based application for user analysis job", *Proc. Computing in High Energy Physics (CHEP2010), Taipei, Taiwan* (2010).

[62] "CMS Software WorkBook: ConfigEditor", `http://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookConfigEditor`.

[63] M.Brodski et al., "Visual Physics Analysis - Applications in High Energy and Astroparticle Physics", *Proceedings ACAT2010, Jaipur, India, PoS(ACAT2010)* **064** (2010).

[64] "Visual Physics Analysis (VISPA)", `http:/vispa.sourceforge.net`.

[65] "Physics eXtension Library (PXL)", `http://pxl.sourceforge.net`.

[66] "Python Programming Language", `http://www.python.org`.

[67] "Simplified Wrapper and Interface Generator (SWIG)", `http://www.swig.org`.

[68] O. Actis et al., "Automated Reconstruction of Particle Cascades in High Energy Physics Experiments", `arXiv:0801.1302`.

[69] "Qt - A cross-platform application and UI framework", `http://www.qtsoftware.com`.

[70] "PyQt", `http://www.riverbankcomputing.co.uk`.

[71] G. P. Salam and G. Soyez, "A practical seedless infrared-safe cone jet algorithm", *JHEP* **0705** (2007) 086.

[72] S. Catani, Y. L. Dokshitzer, M. H. Seymour et al., "Longitudinally invariant Kt clustering algorithms for hadron hadron collisions", *Nucl.Phys.B* **406** (1993) 187–224.

[73] S. D. Ellis and D. E. Soper, "Successive Combination Jet Algorithm For Hadron Collisions", *Phys.Rev.D* **48** (1993) 3160.

[74] Y. Dokshitzer, G. Leder, S. Moretti et al., "Better jet clustering algorithms", *JHEP* **08** (1997) 001.

[75] M. Wobisch and T. Wengler, "Hadronization corrections to jet cross sections in deep-inelastic scattering", arXiv:hep-ph/9907280.

[76] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-kt Jet Clustering Algorithm", *JHEP* **0804** (2008) 063.

[77] A. Hinzmann, "Jets and high-pT Hadron Production in p-p collisions", https://indico.cern.ch/contributionDisplay.py?contribId=39&confId=91353.

[78] CMS Collaboration, "Tracking and Vertexing Results from First Collisions", *CMS Physics Analysis Summary* **TRK-10-001** (2010).

[79] CMS Collaboration, "Jet Plus Tracks Algorithm for Calorimeter Jet Energy Corrections in CMS", *CMS Physics Analysis Summary* **JME-09-002** (2009).

[80] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and E_T^{miss}", *CMS Physics Analysis Summary* **PFT-09-001** (2009).

[81] CMS Collaboration, "Commissioning of the Particle-Flow Reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV", *CMS Physics Analysis Summary* **PFT-10-002** (2010).

[82] "ECAL Spikes in CMS", http://cms--project--ecal--p5.web.cern.ch/cms--project--ECAL--P5/approved/Spike2.php.

[83] CMS Collaboration, "Calorimeter Jet Quality Criteria for the First CMS Collision Data", *CMS Physics Analysis Summary* **JME-09-008** (2009).

[84] CMS Collaboration, "Jet Energy Resolution in CMS at $\sqrt{s} = 7\ TeV$", *CMS Physics Analysis Summary* **JME-10-014** (2011).

[85] CMS Collaboration, "Measurement of the Jet Energy Resolutions and Jet Reconstruction Efficiency at CMS", *CMS Physics Analysis Summary* **JME-09-007** (2009).

[86] G. Cowan, "A Survey of Unfolding Methods for Particle Physics", *Proc. Advanced Statistical Techniques in Particle Physics, Durham* (2002).

[87] Z. Nagy, "Three-jet cross sections in hadron hadron collisions at next-to-leading order", *Phys. Rev. Lett.* **88** (2002) 122003, arXiv:hep-ph/0110315. doi:10.1103/PhysRevLett.88.122003.

[88] T. Kluge, K. Rabbertz, and M. Wobisch, "Fast pQCD calculations for PDF fits", in *14th International Workshop on Deep Inelastic Scattering (DIS 2006), 20-24 Apr 2006*, pp. 483–486. 2006. arXiv:hep-ph/0609285.

[89] J. Gao et al., "Next-to-leading QCD effect to the quark compositeness search at the LHC", *Phys.Rev.Lett.* **106** (2011) 142001.

[90] T. Junk, "Modified frequentist analysis of search results", *Nucl. Instrum. Methods* (1999) A434, 435.

[91] A. Read, "Modified frequentist analysis of search results (The CL(s) method)", *First Workshop on Confidence Limits, CERN, Geneva, Switzerland* (2000) 81–101.

[92] A. L. Read, "Presentation of search results: the CLs technique", *J. Phys. G: Nucl. Part. Phys.* **28** (2002) 2693.

[93] "Roostats Manual", http://root.cern.ch/root/html/ROOFIT_ROOSTATS_Index.html.

[94] "Roofit Manual", http://root.cern.ch/drupal/content/roofit.

[95] R. Cousins and V. Highland, "Incorporating systematic uncertainties into an upper limit", *Nucl. Instrum. Meth.* **A320** (1992) 331. doi:10.1016/0168-9002(92)90794-5.

[96] CMS Collaboration, "Measurement of Dijet Angular Distributions and Search for Quark Compositeness in pp Collisions at $\sqrt{s} = 7$ TeV", *The Durham HepData Project* (2011) http://hepdata.cedar.ac.uk/view/p7951.

[97] CMS Collaboration, "Measurement of Dijet Angular Distributions and Search for Quark Compositeness in pp Collisions at 7 TeV", *Phys.Rev.Lett.* **106** (2011) 201804, arXiv:1102.2020.

[98] CMS Collaboration, "CMS Physics Technical Design Report Volume II: Physics Performance", *CERN-LHCC-2006-021* (2006).

[99] ATLAS Collaboration, "Search for New Physics in Dijet Mass and Angular Distributions in pp Collisions at sqrt(s) = 7 TeV Measured with the ATLAS Detector", *New J.Phys.* **13** (2011) 053044, arXiv:1103.3864.

[100] R. Franceschini et al., "LHC bounds on large extra dimensions", arXiv:1101.4919.

Danksagung (Acknowledgements)

Als erstes geht mein Dank an Prof. Dr. Martin Erdmann für die ausgezeichnete Betreuung dieser Doktorarbeit. Die wöchentlichen Arbeitsgruppentreffen unter Leitung von Prof. Dr. Martin Erdmann, sowie die persönliche Betreuung waren hervorragende Vorraussetzungen für das Gelingen dieser Arbeit.

Prof. Dr. Thomas Hebbeker gilt ein großes Dankeschön für die Übernahme der Zweitkorrektur, sowie die Leitung der wöchentlichen Arbeitsgruppentreffen der 3A-CMS Gruppe.

A special thank goes to Prof. Nikos Varelas who acted as the contact person for this analysis in CMS and lead the dijet angular distributions analysis team through the complete process of publication from the first day of data taking.

I would like to thank the dijet angular distributions analysis team – Leonard Apanasevich, Suvadeep Bose, Dr. Klaus Rabbertz and Nikos Varelas – for fruitful discussions in our regular meetings and very good teamwork.

Many thanks also to the analysis review committee – Robert Klanner and Marek Zielinsk – who guided this analysis through the process of publication.

Desweiteren möchte ich meinem Mit-Doktoranden Jan Steggemann für den täglichen Chat über Physik- und Statistikprobleme, sowie sonstigen CMS-Tratsch, danken.

Unserer CMS-Arbeitgruppe – Martin Erdmann, Robert Fischer, Rebekka Höing, Tatsiana Klimkovich, Dennis Klingebiel, Joschka Lingemann und Jan Steggemann – möchte ich für den interessanten Austausch in den wöchentlichen Meetings und auch während der gemeinsamen Exkursionen zur DPG bzw. zum CERN danken.

Mein Dank gilt außerdem Dr. Benedikt Hegner (auch bekannt als Benedetto Cola Calzone) für die Betreuung und Untersützung meiner Arbeit in der Physics Analysis Tools Gruppe in CMS.

Außerdem möchte ich dem gesamten VISPA-Team – Martin Erdmann, Robert Fischer, Dennis Klingebiel, Matthias Komm, Joschka Lingemann, Gero Müller, Jan Steggemann und Tobias Winchen – für die gute Zusammenarbeit und die interessanten Diskussionen über Softwarekonzepte und Implementierung danken.

Für die Unterstützung bei technischen Problemen mit dem Aachener GRID cluster gilt mein besonderer Dank – Dr. Thomas Kress, Dr. Andreas Nowak und Dr. Manuel Giffels.

Jan Steggemann und Martin Erdmann möchte ich außerdem für das Korrekturlesen dieser Arbeit danken.

Regina möchte ich für die großartige Unterstützung während unserer gesamten Zeit in Genf danken.

Abschließend möchte ich meinen Eltern für die Unterstützung während des gesamten Studiums danken.

Lebenslauf

Persönliche Daten

Name:	Hinzmann
Vorname:	Andreas
Geburstag:	30.01.1982
Geburtsort:	Düsseldorf
Staatsangehörogkeit:	Deutsch

Qualifikationen

2001:	Abitur am Gymnasium Hochdahl
2002-2008:	Studium der Physik an der RWTH Aachen University, Abschluss Diplomphysiker
seit 2008:	Beginn der Promotion an der RWTH Aachen University