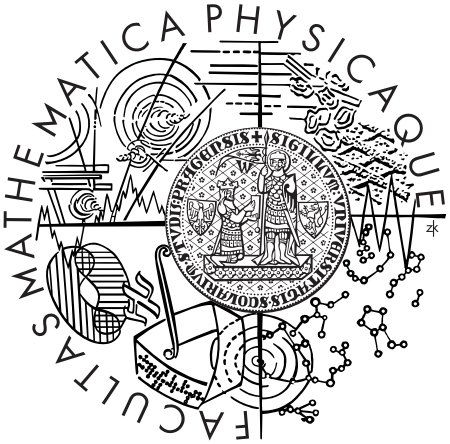
Charles University in Prague
Faculty of Mathematics and Physics

Doctoral Thesis



Lukáš Přibyl

pp interactions at the ATLAS experiment at the LHC accelerator

Institute of Particle and Nuclear Physics, Charles University
Thesis Supervisor: **RNDr. Stanislav Němeček, CSc.**
Thesis Consultant: **RNDr. Tomáš Davídek, PhD.**

Study Programme: Physics

Subject Field: F9 – Subnuclear Physics
February 2008

pp interactions at the ATLAS experiment at the LHC accelerator

Lukáš Přibyl

Institute of Physics
Academy of Sciences of the Czech Republic

Thesis Supervisor: **RNDr. Stanislav Němeček, CSc.**
Thesis Consultant: **RNDr. Tomáš Davídek, PhD.**

Abstract

In this thesis the potential of the ATLAS detector to exclude or discover quark compositeness is studied in detail. There were four different sensitive analysis methods developed for this purpose. These methods use a detailed information about inclusive dijet p_{T}-spectrum or dijet angular distribution and two of them use Bayesian technique (as is usual in similar analyses at Tevatron). For all four methods a larger number of systematic errors was taken into account and their influence was studied in detail too.

It is also shown, that the early data exclusion limit on quark compositeness reachable by the ATLAS experiment is $\Lambda^{+}_{LL} = 10.3$ TeV at 95% C.L with 100 pb^{-1} of collision data of a sufficient quality. This limit is considerably higher than the current limit known from Tevatron: $\Lambda^{+}_{LL} = 2.7$ TeV at 95% C.L. The highest exclusion limit on quark compositeness with expected total amount of data to be collected by the ATLAS experiment is above 22 TeV, assuming 3% jet energy scale uncertainty.

Prohlašuji, že jsem dizertační práci napsal samostatně a výhradně s použitím citovaných pramenů. Souhlasím se zapůjčováním práce.

.....
Lukáš Přibyl
V Praze, dne 22. února 2008

to my parents Jitka and Jiří
and to my wife Adéla

Acknowledgements

This thesis would hardly exist in this shape without many fruitful discussions with my colleagues. I would like to express my big thanks to Stanislav Němeček and Tomáš Davídek for their unremitting willingness to discuss all the aspects of my analysis and sometimes for pushing me in writing of this thesis. I would like to thank to Tomáš Davídek for his invaluable help with computational aspects of the analysis. I also owe thanks to Iacopo Vivarelli and Chiara Roda for jet energy scale and jet analysis discussions, to Samir Ferrag and Gustaaf Broojimans for comments on quark compositeness, to prof. Jiří Chýla and Saša Kupčo for QCD discussions, to Jiří Mašík and Saša Solodkov for their invaluable help with the always so quickly evolving framework Athena, to Tancredi Carli and Olda Kepka for their support with the almost undocumented Monte Carlo generator NLOJET++, to Marek Taševský for Monte Carlo discussions, to Jan Smolík for discussions and always friendly atmosphere in our office.

I must also thank for the opportunity to participate on the ATLAS detector installation and commissioning. The interactions with ATLAS colleagues and the experience with the hardware of the ATLAS detector, for which I would like to thank to Stanislav Němeček, were very enriching.

Finally, I would like to thank to my family for their support and understanding and patience and strength they give me.

Contents

Chapter 1

Introduction

The Large Hadron Collider (LHC) at CERN will move the frontiers of contemporary particle physics much further, with its extraordinary high luminosity and collision energy. With its two multipurpose detectors ATLAS and CMS it will also bring an unprecedented opportunity to study a substructure of quarks – quark compositeness – a subject of interest in this paper. In this chapter we will first outline the physics expectations laid on the ATLAS experiment and physics motivation of our study and then briefly describe the structure of this paper.

1.1 Physics motivation

Let us recall some of the main questions we would like to answer in particle physics:

Do the forces in nature unify? Recalling the history of unification (electricity and magnetism, electromagnetic and weak forces) it would be desirable to unify strong, electroweak and gravitational force. Within Standard Model (SM) this can not happen at Planck scale (10^{19} GeV), but already if Supersymmetry (SUSY) is added it is feasible. So is the Supersymmetry the right way to go? If yes, the ATLAS detector has a great potential to discover some of the predicted supersymmetric particles.

What gives particles mass? According to SM as the particles are moving through the Higgs field, it increases their resistance to movement and they acquire mass. The Higgs boson, creating this field, is the last missing cornerstone of SM. Today we know its mass must lie in the range 114 Gev $< M_H <$ 1.2 TeV. Several processes of Higgs boson decay were chosen as a benchmark for the design of the ATLAS detector.

What is dark mater? Planets, stars, baryonic dark matter,... all these make up to $\approx 4\%$ of the universe. 22 % is non-baryonic Dark matter and 74 % Dark Energy [71]. An evidence for Dark matter comes from observation of galaxy rotation curves and collisions of galaxy clusters, microwave background anisotropies, type Ia supernovae, weak lensing, etc. Some of the cold dark matter particle candidates, such as SUSY neutralino χ might be measured also under ATLAS experimental conditions.

Apart from discovering new kinds of particles, the ATLAS experiment should confirm characteristics of particles already known. Today quarks are known as fundamental point-like particles with no observed substructure. This fact can be expressed in a lower cut on quark compositeness scale Λ. The model of quark compositeness used here is based on left-left isoscalar contact interaction effective lagrangian with destructive interference [11], Λ is then denoted as Λ^+_{LL}. The latest results from Tevatron (see Sec. 3.4) show that $\Lambda^+_{LL} > 2.7$ TeV at 95% C.L. As will be shown in this study, the ATLAS detector can move this frontier quickly much further – to $\Lambda^+_{LL} > 10.3$ TeV at 95% C.L with already 100 pb^{-1} of collision data of a sufficient quality, which might be available already during the first year of its operation. Moreover with all expected collected data at 300 fb^{-1}, the compositeness frontier will be moved above 22 TeV, unless quark compositeness is really discovered. The ATLAS detector potential of quark compositeness discovery is also analysed in this study.

Quark compositeness would be the most pronounced in high-p_{T} region of inclusive dijet p_{T} spectra, where it would cause an excess of events, or in inclusive dijet angular distribution spectra, where it would cause and excess of events in the central region. Four analysis methods for these two kind of distributions were developed and their sensitivity to quark compositeness and to systematic errors is discussed in detail.

1.2 Structure of this paper

The content of this paper is organised as follows:

In chapter 2 aspects of perturbative QCD necessary for our study are discussed. Following chapter 3 brings a brief overview of quark compositeness models and the description of the model used for this study and its implementation in used Monte Carlo generator. Chapter 4 presents the ATLAS detector at LHC with its sub-detectors and read-out systems. In the next chapter 5 the strategy of jet reconstruction and calibration used for full simulation data, and in the near future also for the physics data, is discussed together with expected jet energy scale uncertainties that must be taken into account in every analysis using jets. Chapter 6 brings a description of simulated datasets used in this study and an overview of jet reconstruction in fast simulation. All four analysis methods and sources of systematic errors are also discussed therein. Results of analysis of inclusive dijet p_{T} spectrum are contained in chapter 7, and results of analysis of inclusive dijet angular distribution are presented in chapter 8. Finally, results from all four analysis methods are briefly summarised in chapter 9, conclusions are presented in the last chapter 10.

Chapter 2

Aspects of QCD

An underlying theory for this study is perturbative Quantum Chromodynamics (pQCD), a very successful theory of strong interaction. In our search for quark compositeness we will focus on spotting of discrepancies between QCD predictions and measured data. In this Chapter a brief review of QCD aspects, the most important ones for this study, is presented. Unless stated otherwise, the facts in this Chapter were taken from [1], [2] and a review article [3], tailored to the needs of physics at LHC.

2.1 QCD Langrangian

An experimental evidence that quarks come in three colours was obtained already several decades ago. Another quark quantum number, colour, was required in order to explain an existence of baryon Δ^{++} with spin 3/2. Moreover, cross-sections and decay rates of various processes depend on number of colours, N_c, e.g $\sigma(e^+e^- \to \text{hadrons}) \propto N_c$ and $\Gamma(\pi_0 \to \gamma\gamma) \propto N_c^2$.

QCD is a Yang-Mills gauge theory with SU(3) symmetry. Quarks are put in triplets:

$$\Psi^q(x) \equiv \begin{pmatrix} \Psi_1^q(x) \\ \Psi_2^q(x) \\ \Psi_3^q(x) \end{pmatrix}, \tag{2.1}$$

where q is a quark flavour. Invariance under local SU(3) transformations is required. QCD Lagrangian reads:

$$\begin{aligned} \mathcal{L}_{\text{QCD}} &= -\frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} + \sum_q \overline{\Psi}^q_i(i\gamma^\mu D_{\mu ij} - m_q\delta_{ij})\Psi^q_j + \mathcal{L}_{\text{gauge fixing}} \\ &\quad + \mathcal{L}_{\text{ghost}}, \\ F^a_{\mu\nu} &= \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g_s f^{abc} A^b_\mu A^c_\nu, \\ D_{\mu ij} &= \delta_{ij}\partial_\mu - ig_s T^a_{ij} A^a_\mu, \end{aligned} \tag{2.2}$$

where g_s is the QCD coupling constant, f^{abc} are the structure constants of SU(3): $[T^a, T^b] = if^{abc}T^c$ (a, b, c = 1,...,8), A^a_μ are the 8 gluon fields. T^a_{ij} are

8 'colour matrices', i.e. generators of the SU(3) transformation acting on the fundamental (triplet) representation:

$$T^a_{ij} = \frac{1}{2}\lambda^a_{ij}, \tag{2.3}$$

where λ^a_{ij} are Gell-Mann matrices. This corresponds to the normalisation

$$\mathrm{Tr}(T^a T^b) = T^a_{ij} T^b_{ji} = \frac{1}{2}\delta_{ab}. \tag{2.4}$$

The QCD Lagrangian is invariant under local SU(3) transformations:

$$\Psi \rightarrow \exp\left(i\sum_{a=1}^{8} T^a \omega^a(x)\right)\Psi \tag{2.5a}$$

$$A^a_\mu \rightarrow A^a_\mu - \frac{1}{g_s}\partial_\mu \omega^\mu(x) - \sum_{b,c=1}^{8} f^{abc}\omega^b(x) A^c_\mu. \tag{2.5b}$$

In order to quantise the theory and reduce the number of degrees of freedom of the gauge fields, one needs to introduce a gauge fixing term. In case of *covariant gauges*, the $\mathcal{L}_{\text{gauge fixing}}$ term in Eq. (2.2) takes form

$$\mathcal{L}_{\text{gauge fixing}} = \frac{1}{2\alpha}\sum_a (\partial^\mu A^a_\mu)^2, \tag{2.6}$$

and additional non-physical ghost fields are required to guarantee the unitarity of the theory. Putting $\alpha = 1$ defines so-called Feynman gauge.

In case of *non-covariant (axial)* gauges, the gauge fixing term reads

$$\mathcal{L}_{\text{gauge fixing}} = \frac{1}{2\alpha}\sum_a (n^\mu A^a_\mu)^2, \tag{2.7}$$

and no ghost fields are required. Nevertheless, it was proven, that physical quantities are independent of the choice of gauge fixing term.

2.1.1 QCD coupling constant

When calculating invariant amplitudes corresponding to diagrams with loops, one must face the ultraviolet (UV) divergences, associated with infinite loop momenta. In order to manage such divergences and also to give the theory a better physical meaning the renormalisation procedure is used, which replaces divergent contributions with finite expressions in a systematic way. The price one has to pay for that, is an introduction of renormalisation scale μ. Physical quantities are independent of μ, only if all orders of perturbation expansion are involved. The higher the number of orders involved, the smaller the dependence on μ.

Renormalised strong coupling constant satisfies

$$g(\mu) = g_0 + g_0^3\left[b\ln\frac{M}{\mu} + c\right] + \ldots, \tag{2.8}$$

where g_0 is the bare coupling and M the UV cut-off for divergent loop integrals. Eq. (2.8) represents already a particular convention. We could have defined another coupling, $g\prime$, by say replacing $\mu \to 2\mu$. But in general, the coupling constants in two different schemes will be related by:

$$g\prime(\mu) = g(\mu)\left[1 + \kappa g^2(\mu) + \ldots\right]. \tag{2.9}$$

The dependence of $g(\mu)$ on μ can be written as

$$\mu\frac{\partial}{\partial\mu}g(\mu) = -b_{\mathrm{QCD}}\, g^3(\mu) + \mathcal{O}(g^5(\mu)) = -\beta(g(\mu))g(\mu), \tag{2.10}$$

where n_f is the number of quark flavours, $b_{\mathrm{QCD}} = \frac{1}{16\pi^2}\left(11 - \frac{2}{3}n_f\right)$ and the β-function reads

$$\beta(g) = \beta_0\frac{g^2}{16\pi^2} + \beta_1\left(\frac{g^2}{16\pi^2}\right)^2 + \ldots. \tag{2.11}$$

Note, that in Eq. (2.10), the factor b_{QCD} is positive - that is an opposite situation with respect to QED.

In principle one can solve the Eq. (2.10) in terms of $g(\mu_0)$, determined from experiment. Explicit leading order solution (so-called running coupling constant) reads

$$g^2(\mu) = \frac{g^2(\mu_0)}{1 + g^2(\mu_0)b\ln(\mu^2/\mu_0^2)}. \tag{2.12}$$

Introducing $\alpha_s = g^2/(4\pi)$, the experimental value for strong coupling constant is $\alpha_s(M_{\mathrm{Z}}) = 0.118$ at a scale $\mu_0 = M_{\mathrm{Z}}$. The interpretation of this result is that the closer the quarks are to each other (the higher the scale μ is), the weaker is the 'colour charge' and the quarks behave almost like free particles. This phenomenon is called *asymptotic freedom* and it is one of the corner-stones of QCD.

2.1.2 Relevant invariant amplitudes

QCD processes will form most of the background, when searching for quark compositeness. Leading order invariant amplitudes for such relevant processes are listed in Tab. 2.1, as taken from [2]. The normalisation is such that

$$\frac{\mathrm{d}\sigma}{\mathrm{d}t} = \frac{1}{16\pi s^2}|\overline{M}|^2, \tag{2.13}$$

s, t and u are Mandelstam variables.

Relative contribution of various initial parton state channels to jet inclusive p_{T} spectrum are depicted in Fig. 2.1.

Process	$\frac{1}{g^4}\lvert\overline{M}\rvert^2$
$q_\alpha q_\beta \to q_\alpha q_\beta$	$\frac{2}{9}\left[\frac{2(s^2+u^2)}{t^2} + \left(\frac{2(t^2+s^2)}{u^2} - \frac{1}{3}\frac{4s^2}{ut}\right)\delta_{\alpha\beta}\right]$
$q_\alpha \bar{q}_\beta \to q_\alpha \bar{q}_\beta$	$\frac{2}{9}\left[\frac{2(s^2+u^2)}{t^2} + \left(\frac{2(t^2+u^2)}{s^2} - \frac{1}{3}\frac{4u^2}{st}\right)\delta_{\alpha\beta}\right]$
$qg \to qg$	$\left[\left(1-\frac{us}{t^2}\right) - \frac{4}{9}\left(\frac{s}{u}+\frac{u}{s}\right) - 1\right]$
$gg \to q\bar{q}$	$\frac{1}{6}\left[\frac{u}{t}+\frac{t}{u}\right] - \frac{3}{4}\left[1-\frac{ut}{s^2}\right] + \frac{3}{8}$
$q\bar{q} \to gg$	$\frac{64}{9}M(gg \to q\bar{q})$
$gg \to qq$	$\frac{8}{9}\left[-\frac{33}{4} - 4\left(\frac{us}{t^2}+\frac{ut}{s^2}+\frac{st}{u^2}\right)\right] + \frac{9}{16}\left[45 - \left(\frac{s^2}{ut}+\frac{t^2}{us}+\frac{u^2}{ts}\right)\right]$

Table 2.1: Spin and colour averaged absolute squares of invariant amplitudes for the simplest two-body QCD processes.

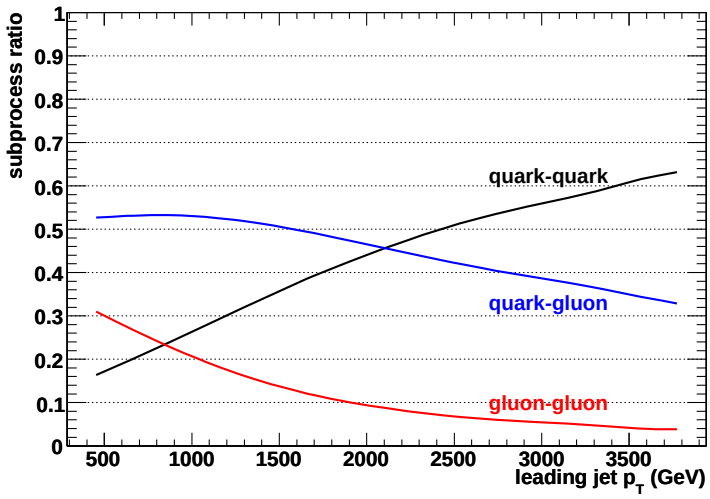


Figure 2.1: Relative contribution of various initial parton state channels to jet inclusive p_{T} spectrum. Calculated in PYTHIA with CTEQ6L1 pdf.

2.2 Parton showers

Using a leading-order matrix elements leads to an inherent dependence on unphysical renormalisation and factorisation scale, which is often large. One way to reduce this source of theoretical uncertainty is to use higher orders in the perturbative expansion of a given observable. Another way, widely used in Monte Carlo generators, is a sort of "all orders" approach - *parton showers*.[1]

By use of parton showers a few partons produced in a hard interaction at a high energy scale can be related to partons at an energy scale close to Λ_{QCD}[2].

At this lower energy scale, a universal non-perturbative model can then be used to provide the transition from partons to the hadrons that are observed experimentally. This is possible because the parton showering allows for the evolution, using the DGLAP formalism, of the parton fragmentation function. The solution of this DGLAP evolution equation can be rewritten with the help of Sudakov form factor, which indicates the probability of evolving from a higher scale to a lower scale without the emission of a gluon greater than a given value. For the case of parton showers from the initial state the evolution proceeds backwards from the hard scale of the hard process to the cutoff scale with the Sudakov form factors being weighted by the pdfs at the relevant scales.

We can write an expression for the Sudakov form factor of an initial state parton as

$$\Delta(t) \equiv \exp\left[-\int_{t_0}^{t} \frac{\mathrm{d}t'}{t'} \int \frac{\mathrm{d}z}{z} \frac{\alpha_S}{2\pi} P(z) \frac{f(x/z,t)}{f(x,t)}\right], \tag{2.14}$$

where t is the hard scale, t_0 is the cutoff scale and $P(z)$ is the branching function (see also Sec. 2.3.1) for the branching under consideration. The Sudakov form factor has a similar form for the final state parton but without the pdf weighting. The introduction of the Sudakov form factor re-sums all the effects of soft and collinear gluon emission, which leads to well-defined predictions even in these kinematic regions.

To sum up, parton showers provide an excellent description in regions which are dominated by soft and collinear gluon emission. On the other hand matrix element calculations provide a good description of processes, where the partons are energetic and widely separated and, in addition, include the effects of interference between amplitudes with the same external partons. One attempt to merge these approaches, CKKW formalism, is implemented in Monte Carlo generator SHERPA and an approximate version also in ALPGEN 2.0.

2.3 Parton distribution functions

Protons are not fundamental particles, they have an inner structure. What we observe during hard collisions of protons are interactions between proton

[1] Numerical implementation done in Monte Carlo generators PYTHIA or HERWIG and SHERPA.

[2] The scale parameter Λ_{QCD} can be formulated in terms of μ_0 from Eq. (2.12) as $\Lambda^2_{\mathrm{QCD}} = \mu_0^2 \exp(-1/(b\, g^2(\mu_0))$. Eq. (2.12) can be then rewritten as $g^2(\mu) = 1/(b \ln(\mu^2/\Lambda^2_{\mathrm{QCD}}))$.

constituents – partons (quarks and gluons). The proton structure was probed in detail in deep inelastic scattering (DIS) experiments – collisions of leptons and protons. A scheme of DIS process is depicted in Fig. 2.2.

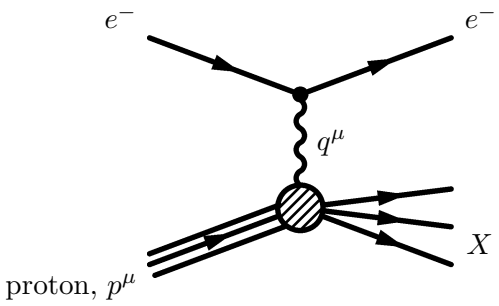


Figure 2.2: A scheme of deep inelastic scattering process $e + p \to e + X$.

Standard variables used to describe this process are

$$Q^2 = -q^2, \tag{2.15a}$$

$$x = \frac{Q^2}{2pq}, \tag{2.15b}$$

$$y = \frac{Q^2}{xs}, \tag{2.15c}$$

where s is the the Mandelstam variable. In general, for a corresponding cross-section, we can write

$$\frac{\mathrm{d}\sigma}{\mathrm{d}x\mathrm{d}Q^2} = \frac{4\pi\alpha^2}{Q^4}\left[y^2 F_1(x,Q^2) + 2(1-y)\frac{1}{x}F_2(x,Q^2)\right], \tag{2.16}$$

where the $F_i(x,Q^2)$ are called structure functions. At the beginning of these measurements, the experimental data showed, that for $Q^2 > 1$ GeV^2 the structure functions almost do not depend on Q^2 and are constant to a good approximation. Thus $F_i(x,Q^2) \to F_i(x)$. This is so-called Bjorken scaling. Later we will comment on the fact, how this independence is violated.

There is an important relation between the two structure functions, called Callan-Gross relation:

$$F_2(x) = 2xF_1(x). \tag{2.17}$$

This relation was confirmed experimentally and provides a convincing evidence that charged partons carry spin 1/2.

The structure function F_2 is directly related to the parton distribution functions (pdfs), $q_i(x)$, describing the probability density of finding a parton type i, carrying a fractional nucleon momentum x inside the nucleon:

$$F_2(x) = \sum_i x\, e_i^2 q_i(x), \tag{2.18}$$

where e_i is the electric charge of parton type i. The independence on Q reflects our assumption of partons behaving as asymptotically free particles.

2.3.1 pdf evolution

Later experiments showed, that the structure function data do exhibit systematic violations of Bjorken scaling. This is due to quarks emitting gluons during the collision. So far we have considered quarks and gluons as free particles before and after the collision, even though they are confined within the hadrons and behave like free particles only when probed at short distances. Renormalised, or dressed, pdfs take this behaviour into account and they also explain the violation of Bjorken scaling.

Let us first consider interactions of partons with their own chromodynamic field. These interactions lead to a non-zero probability of finding parton of one type within another parton, which is expressed in QCD branching functions. In leading order[3]:

$$\begin{aligned}
P_{qq}^{(0)}(x) &= P_{\bar{q}\bar{q}}^{(0)}(x) = \frac{4}{3}\left[\frac{1+x^2}{1-x}\right]_+, \\
P_{Gq}^{(0)}(x) &= P_{G\bar{q}}^{(0)}(x) = \frac{4}{3}\left[\frac{1+(1-x)^2}{x}\right], \\
P_{qG}^{(0)}(x) &= P_{\bar{q}G}^{(0)}(x) = \left[\frac{x^2+(1-x)^2}{2}\right], \\
P_{GG}^{(0)}(x) &= 6\left(\left[\frac{x}{1-x}\right]_+ + \frac{1-x}{x} + x(x-1) + \left(\frac{33-2n_f}{36} - 1\right)\delta(1-x)\right),
\end{aligned} \tag{2.19}$$

where e.g. $P_{qG}^{(0)}$ is a probability to find a quark inside gluon that would carry a fractional momentum x of the original gluon momentum. All these branching functions are the same for all quark flavours.

In a situation described in previous section, corresponding to Eq. (2.18), the quark was assumed to have no chromodynamic field around itself – so-called *bare* quark. Now we can consider a situation, when the bare quark, before interacting with the probing photon, emits gluons as depicted in Fig. 2.3. When we sum up all the contributions in Fig. 2.3, we define the *renormalised*, or *dressed* quark distribution functions, $q(x,\mu)$:

[3]The "+" distribution in first function is defined as:
$[f(x)]_+ \equiv \lim_{\beta\to 0^+}\Big(f(x)\theta(1-x-\beta) - \delta(1-x-\beta)\int_0^{1-\beta} f(y)\mathrm{d}y\Big)$.

$$
\begin{aligned}
q(x,\mu) \;\equiv\; & q_0(x) + \int_x^1 \frac{\mathrm{d}y}{y}\left[P_{qq}^{(0)}\left(\frac{x}{y}\right)\int_{m^2}^{\mu^2}\frac{\mathrm{d}\tau_1}{\tau_1}\frac{\alpha_s(\tau_1)}{2\pi}\right] q_0(y) + \\
& + \int_x^1 \frac{\mathrm{d}y}{y}\int_y^1 \frac{\mathrm{d}w}{w}\int_{m^2}^{\mu^2}\frac{\mathrm{d}\tau_1}{\tau_1}\int_{m^2}^{\tau_1}\frac{\mathrm{d}\tau_2}{\tau_2}\frac{\alpha_s(\tau_1)}{2\pi}\frac{\alpha_s(\tau_2)}{2\pi}\times \\
& \times\left[P_{qq}^{(0)}\left(\frac{x}{y}\right)P_{qq}^{(0)}\left(\frac{y}{w}\right)\right] q_{\mathrm{NS},0}(x) + \ldots, \qquad (2.20)
\end{aligned}
$$

where m is an infrared cut-off, the scale μ has the meaning of maximal virtuality of the quark which interacts with the photon in the upper, QED, vertex of Fig. 2.3 and it corresponds to the factorisation scale. At the leading order, the dressed distribution functions correspond to the valence ones used in Eq. (2.18).

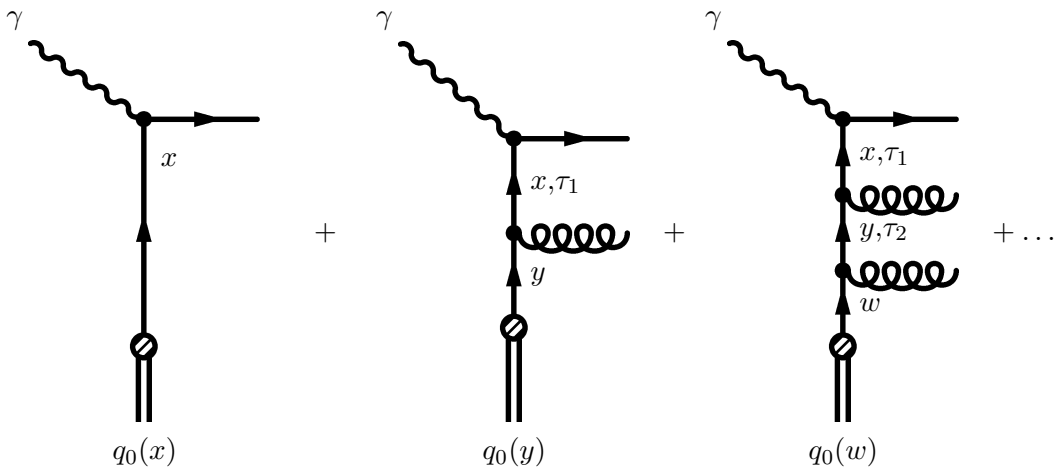


Figure 2.3: Diagrams contributing to the dressed parton distribution function of the non-singlet quark inside a given hadron. τ_i denote absolute value of the virtuality of a given intermediate state.

$q(x,\mu)$ are not calculable in perturbation theory. Nevertheless its scale (μ) dependence is. In the leading order this dependence is given by DGLAP evolution equations, [7]:

$$
\begin{aligned}
\frac{\mathrm{d}q_i(x,\mu)}{\mathrm{d}\ln\mu} &= \frac{\alpha_S(\mu)}{\pi}\left[\int_x^1\frac{\mathrm{d}y}{y}P_{qq}^{(0)}\left(\frac{x}{y}\right)q_i(y,\mu) + \int_x^1\frac{\mathrm{d}y}{y}P_{qG}^{(0)}\left(\frac{x}{y}\right)G(y,\mu)\right], \\
\frac{\mathrm{d}\bar{q}_i(x,\mu)}{\mathrm{d}\ln\mu} &= \frac{\alpha_S(\mu)}{\pi}\left[\int_x^1\frac{\mathrm{d}y}{y}P_{qq}^{(0)}\left(\frac{x}{y}\right)\bar{q}_i(y,\mu) + \right. \\
&\quad \left. + \int_x^1\frac{\mathrm{d}y}{y}P_{qG}^{(0)}\left(\frac{x}{y}\right)G(y,\mu)\right], \qquad (2.21)
\end{aligned}
$$

where the gluon distribution function $G(x,\mu)$ satisfies similar evolution equation

$$\frac{\mathrm{d}G(x,\mu)}{\mathrm{d}\ln\mu} = \frac{\alpha_S(\mu)}{\pi}\left[\int_x^1 \frac{\mathrm{d}y}{y} P_{Gq}^{(0)}\left(\frac{x}{y}\right)\sum_{i=1}^{n_f}[q_i(y,\mu)+\bar{q}_i(y,\mu)] + \int_x^1 \frac{\mathrm{d}y}{y} P_{GG}^{(0)}\left(\frac{x}{y}\right) G(y,\mu)\right]. \quad (2.22)$$

While $P^{(0)}$ are uniquely defined, $P^{(1)}$'s, used in higher order, are arbitrary. Choice of $P^{(1)}$'s defines the factorisation scheme. Similarly higher order terms in Eq. (2.11) are arbitrary and their choice defines the renormalisation scheme.

A comparison of HERA and LHC parton kinematics is shown in Fig. 2.4. In case of pdfs, a transition in Q^2 from HERA to LHC parton kinematics region can be done using the DGLAP equations.

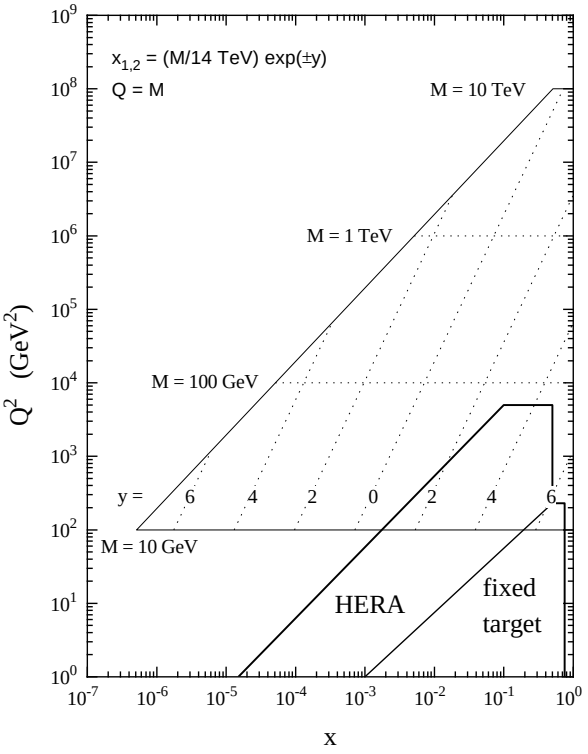


Figure 2.4: Accessible LHC Parton Kinematics. Dependence on rapidity y is also shown.

2.3.2 pdf uncertainties

The calculation of the production cross section at hadron colliders relies upon a knowledge of the distribution of the momentum fraction x of the partons in a proton in the relevant kinematic range. These pdfs are determined by global fits to deep inelastic scattering (DIS), Drell-Yan (DY) and jet production at current energy ranges. In this study we use two major pdf sets, CTEQ [5], and MRST [6], enabling also uncertainty studies.

A global pdf analysis carried out next-to-leading order needs to be performed in a specific renormalisation and factorisation scheme. Among others, the $\overline{\rm MS}$ scheme is the most often used one. All global analyses use a generic form for the parameterisation of both the quark and gluon distribution at some reference value Q_0:

$$F(x, Q_0) = A_0 x^{A_1} (1-x)^{A_2} P(x; A_3, \ldots) \tag{2.23}$$

The reference value Q_0 is usually chosen in the range of 1-2 GeV. The parameter A_1 is associated with small-x behaviour, while A_2 is associated with large-x valence counting rules. The term $P(x; A_3, \ldots)$ is a suitably chosen smooth function.

A conventional method of estimating parton distribution uncertainties has been to compare different published parton distributions. This is unreliable, since most published sets of pdfs adopt similar assumptions. The sum of the quark distributions is, in general well-determined over a wide range of x and Q^2. The individual quark pdfs are known to less accuracy. And the most uncertain are the gluon distribution functions.

CTEQ and MRST use Lagrange Multiplier and Hessian techniques. In the Hessian method a large matrix (20×20 for CTEQ, 15×15 for MRST), with dimension equal to the number of free parameters used for the fit, has to be diagonalized. The result is 20 (15) orthonormal eigenvector directions for CTEQ (MRST) which provide the basis for determination of the pdf error for any cross section. The eigenvectors are now admixtures of the 20 pdf parameters left free in the global fit. The larger eigenvalues correspond to directions which are well-determined; for example, eigenvectors 1 and 2 are sensitive primarily to the valence quark distribution at moderate x, a region where they are well-constrained. Eigenvector 5 (pdfs 9 and 10) is most sensitive to the low x behaviour of the gluon distribution. Eigenvector 15 in the CTEQ6M1 error pdf set is sensitive to the high x gluon behaviour and thus influences high-$p_{\rm T}$ jet cross section behaviour.

Each error pdf results from an excursion along the "+" and "−" directions for each eigenvector. In CTEQ6M1 the excursions in parametric space result in symmetric "+" and "−" error pdfs for the first 10 eigenvectors but can result in asymmetric[4] error pdfs for the last 10. CTEQ6M1 has 40 error pdfs.

Perhaps the most controversial aspect of pdf uncertainties is the determination of the $\Delta\chi^2$ excursion from the central value to determine a reasonable error. CTEQ uses $\Delta\chi^2$ value of 100 (corresponding to 90 % CL limit) while MRST uses a value of 50 (corresponding also roughly to 90 % CL limit). Thus, in general, the pdf uncertainties will be larger for CTEQ than for MRST. The uncertainties for all predictions should be linearly dependent on the tolerance parameter used; thus it should be reasonable to scale the uncertainty for an observable from 90 %CL limit provided by the CTEQ/MRST error pdf to a one-sigma error by dividing by a factor of 1.64.

[4]In other words, the resulting "+" and "−" pdfs for a given eigenvector both lie below or both lie above the central "best fit" pdf.

2.3.3 Master equation

For an observable X let us denote its value using the central pdf X_0. X_i^+ (X_i^-) is the value of that variable using the pdf corresponding to the "+" ("−") direction for the eigenvector i. In order to calculate the pdf error for an observable, the *Master Equation* is used:

$$\Delta X_{max}^+ = \sqrt{\sum_{i=1}^{N} \Big[max(X_i^+ - X_0, X_i^- - X_0, 0)\Big]^2}$$
$$\Delta X_{max}^- = \sqrt{\sum_{i=1}^{N} \Big[max(X_0 - X_i^+, X_0 - X_i^-, 0)\Big]^2} \tag{2.24}$$

The addition in quadrature is justified by the eigenvectors forming an orthonormal basis. Either X_0 and $X_i^\pm$ can be calculated separately in matrix element/Monte Carlo program (requiring the program to be run $2N+1$ times) or X_0 can be calculated with the program and at the same time the ratio of the pdf luminosities (the product of the two pdfs and the x values used in the event generation) for eigenvector i ($\pm$) to that of the central fit can be calculated and stored (*pdf re-weighting*). This results in an effective sample with $2N+1$ weights, but identical kinematics, requiring a substantially reduced amount of time to generate.

An example of pdfs and pdf uncertainties for various partons is shown in Fig. 2.5.

2.3.4 pdf re-weighting

For each event n, generated with the central pdf, and each pdf member i from given pdf set (containing 1 central + N_{PDF} error members), a weight W_n^i can be calculated.

$$W_n^i = \frac{f(x_1, Q; S_i) f(x_2, Q; S_i)}{f(x_1, Q; S_0) f(x_2, Q; S_0)}, \tag{2.25}$$

where $n = 1, \ldots, N_{events}$, $i = 0, \ldots, N_{PDF}$ and S_i denotes individual pdf set member. Weight W_n^0 for the central pdf is equal to 1 by definition. Compared to standard approach, each event is simulated only once, so the kinematics do not change and there is no residual statistical variation in uncertainty.

One concern involved in this method is that re-weighting events does not correctly modify the Sudakov form factors, used for parton showers calculations (see Sec. 2.2), it only affects the hard process. However, the impact of this was shown negligible [8]. The weighting method is theoretically correct only in the limit, when all possible initial states are populated. For this reason, it is important that reasonably high statistics is generated when using this technique. Any analysis sensitive to the extreme tails of distribution should use this method with caution. More detailed information can be found in [9].

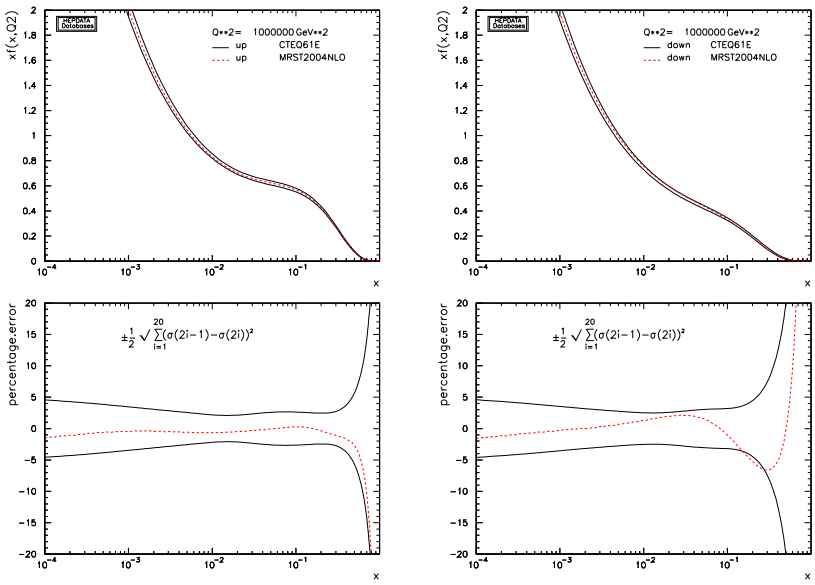


Figure 2.5a: CTEQ6M1 pdf error bands for up and down quark (left, right). MRST2004 NLO central pdf values are also shown (red line), and usually lie within the CTEQ error band. Differences between central pdfs from various pdf sets are usually smaller, than pdf uncertainties within a given pdf set due to the fit procedure.

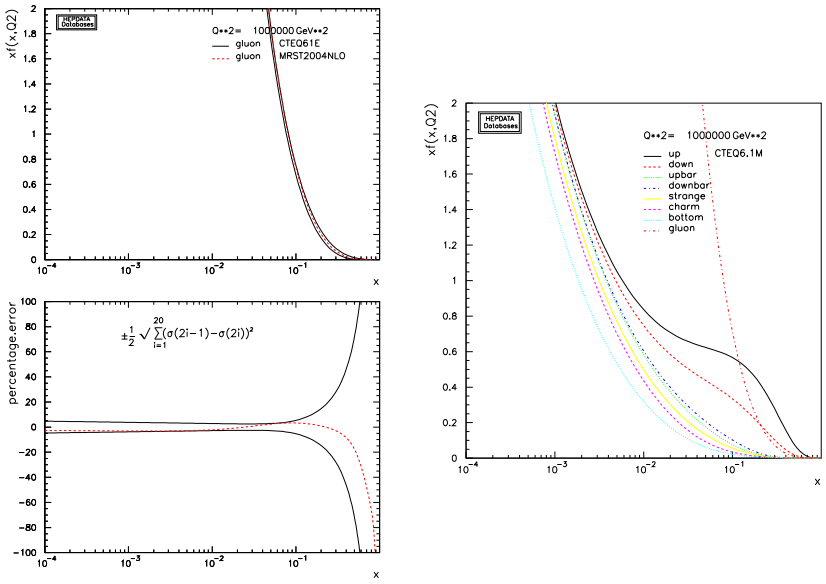


Figure 2.5b: The largest pdf uncertainties at high x are caused by our ignorance of gluon distribution (left). Comparison of CTEQ6M1 central pdf values for various partons (right). All figures obtained using a macro at http://durpdg.dur.ac.uk/hepdata/pdf3.html.

2.4 Hard scattering formalism and the QCD factorisation theorem

The following was adopted from [3]. We begin 30 years ago with the production of massive lepton pair by quark-antiquark annihilation - the Drell-Yan process. It was postulated that the hadronic cross-section $\sigma(AB \to \mu^+\mu^- + X)$ could be obtained by weighting the subprocess cross section $\hat{\sigma}$ for $q\bar{q} \to \mu^+\mu^-$ with the parton distribution functions $f_{q/A}(x)$ extracted from deep inelastic scattering:

$$\sigma_{AB} = \int \mathrm{d}x_a \mathrm{d}x_b f_{a/A}(x_a) f_{b/B}(x_b) \hat{\sigma}_{ab\to X}, \tag{2.26}$$

where for the Drell-Yan process, $X = l^+l^-$ and $ab = q\bar{q}$, $\bar{q}q$. The good agreement between theoretical predictions and the measured cross sections provided confirmation of the parton model formalism. Studies were successfully extended to other hard scattering processes, for example the production of hadrons and photons with large transverse momentum. Problems, however, appeared with perturbative corrections from real and virtual gluon emissions. Large logarithms from gluons emitted collinear with the incoming quarks spoil the convergence of the perturbative expansion. But it turned out that they can be factored into renormalised parton distributions, (2.26) then becomes:

$$\sigma_{AB} = \int \mathrm{d}x_a \mathrm{d}x_b f_{a/A}(x_a, Q^2) f_{b/B}(x_b, Q^2) \hat{\sigma}_{ab\to X}. \tag{2.27}$$

The Q^2 that appears in the pdfs is a large momentum scale that characterises the hard scattering, e.g. $M^2_{l^+l^-}$, p^2_{T}, ...

The last step was the recognition that the finite corrections left behind after the logarithms had been factored were not universal and had to be calculated separately for each process. Schematically

$$\sigma_{AB} = \int \mathrm{d}x_a \mathrm{d}x_b f_{a/A}(x_a, \mu_F^2) f_{b/B}(x_b, \mu_F^2) \times [\hat{\sigma}_0 + \alpha_S(\mu_R^2)\hat{\sigma}_1 + \ldots]_{ab\to X}. \tag{2.28}$$

Here μ_F is the *factorisation scale*, which can be thought of as the scale that separates the long- and short-distance physics, and μ_R is the *renormalisation scale* for the QCD running coupling.

Chapter 3

Quark compositeness

The quark and lepton compositeness models were the most popular among theorists in 1980th. Until present, no signs of compositeness have been observed by any experiment. Nevertheless, it still remains an interesting question, how precisely the inclusive jets cross-sections follow the Standard model predictions, whether possible discrepancies between theory and data are large enough to point on compositeness or other physics beyond the Standard model. Such a question should also be answered by the data from ATLAS detector.

In this section, let us start with a brief review of more advanced compositeness models, then describe a model used for our analysis, and eventually review the present-day limits for quark compositeness.

3.1 Rishon and preon models

When speaking about compositeness, let us mention at least some of the theoretical models describing the hypothetical constituents of quarks and leptons.

3.1.1 Rishon model

The rishon model [19] is based upon $\mathrm{SU(3)_H} \times \mathrm{SU(3)_C} \times \mathrm{U(1)_{EM}}$ gauge invariant lagrangian. The fundamental particles are two types of rishons (one fermion with charge 1/3 e denoted by T and one neutral fermion, denoted by V), hypergluons, gluons and the photon. No fundamental scalar exists. Below the hypercolor scale Λ_H only $\mathrm{SU(3)_H}$ singlets exist. The simplest composite fermions are made of three rishons or three antirishons and reproduce the observed properties of one generation of quarks and leptons. The weak interaction appears only at the composite level as residual short-range interaction among hypercolor singlets.

The ingredients of first generation of composite fermions are presented in Tab. 3.1.

The rishon model thus solves the vanishing sum of all quark and lepton charges in one generation. The renormalisability of the standard model requires the absence of ABJ anomalies, leading to the constraint

TTT (e^+)	$\bar{T}\bar{T}\bar{T}$ (e^-)
TTV (u)	$\bar{T}\bar{T}\bar{V}$ ($\bar{u}$)
TVV (d)	$\bar{T}\bar{V}\bar{V}$ ($\bar{d}$)
VVV (ν_e)	$\bar{V}\bar{V}\bar{V}$ ($\bar{\nu_e}$)

Table 3.1: The ingredients of first generation of composite fermions in the rishon model.

$$Q(e^-) + Q(\nu_e) + 3Q(u) + 3Q(d) = 0. \tag{3.1}$$

The total rishon content of e^-, ν, u and d is $6T + 6\bar{T} + 6V + 6\bar{V}$. Hence the sum of all electric charges vanishes.

One might raise the question, what is the value of scale Λ_H. The answer comes from an analysis of proton decay in rishon model. The authors claim, that to the lowest order in Λ_H (i.e. $\tau_p \sim \mathcal{O}(\Lambda_H^4)$), proton decay is forbidden. It is allowed in higher orders and one then obtains

$$\tau_p \sim \Lambda_H^8 / M^9, \tag{3.2}$$

where the mass factor M may involve quark, lepton, proton or W masses. In case of proton decay the experimental limit on τ_p yields $\Lambda_H \sim 10^8$ GeV, well beyond capabilities of nowadays colliders.

3.1.2 Preons, tweedles, helons

The first model based on compositeness of quarks and leptons was developed by Pati and Salam [20] in 1974. They called their hypothetical constituent particles *preons* and this name was later adopted to refer to sub-quark/sub-lepton particles of any model. Unfortunately, this model lacked any real explanatory power.

The Rishon model, mentioned above, was further used as an inspiration for the Helon model, pedagogically described in [21]. The Helon model uses even lower level fundamental objects, they could be represented as twists through $\pm\pi$ in a ribbon. A twist through π is called U and twist through $-\pi$ is called E. They both are called *tweedles* and can be left- or right-handed. Tweedles combine in pairs called *helons*, forming three types of helons H_+, H_- and H_0. Helons are then bound into triplets.

In terms of the number of particles, the helon model is more economical, than even the rishon model, despite allowing helons to be composite. The helon model also improves on the original rishon model by explaining why the ordering of helons (which are analogous to rishons) should matter. As an example, the helon triplet $H_+H_+H_+$ forms e^+, $H_+H_+H_0$ forms u_B, but $H_+H_0H_+$ forms u_G, where the subscripts of up quark denote colour charges of QCD.

In [21] there is also shown, how the Helon model describes antiparticles, bosons, annihilation, electroweak interaction and other aspects of the Standard Model. It has also ambitions to explain gravity and the origin of mass.

Many other models of quark and lepton compositeness have been developed in a hope to explain the existence of three generations of quarks and leptons, the repetitive pattern of these generations, the connection between the electric charges of quarks and leptons and the existence of more than 20 free parameters in the standard model. All of these models are now lacking experimental data. When the ATLAS data will be available, the first search for compositeness will lead us to comparison of data and standard model predictions.

3.2 Contact interaction

If quarks and leptons are made of constituents, then at the scale of constituent binding energy, there should appear new interactions among quarks and leptons. If the compositeness scale (Λ) is much larger than $\sqrt{\hat{s}}$, the centre of mass energy of the colliding partons, these interactions are suppressed by inverse powers of Λ and the quarks and leptons would appear to be point-like. The dominant effect should come from the lowest dimensional interactions with four fermions (contact terms), whose most general chirally invariant form reads [10]:

$$\begin{aligned}\mathcal{L} = \; & \frac{g^2}{2\Lambda^2}\Big[\eta_{LL}\overline{\Psi}_L\gamma_\mu\Psi_L\overline{\Psi}_L\gamma^\mu\Psi_L + \eta_{RR}\overline{\Psi}_R\gamma_\mu\Psi_R\overline{\Psi}_R\gamma^\mu\Psi_R \\ & + 2\,\eta_{LR}\overline{\Psi}_L\gamma_\mu\Psi_L\overline{\Psi}_R\gamma^\mu\Psi_R\Big].\end{aligned} \tag{3.3}$$

Usually less general form of Eq. (3.3) is used for practical calculations. Typical combinations of values of the coefficients $\eta_{\alpha\beta}$ are

$$\begin{aligned}(\eta_{LL},\eta_{RR},\eta_{LR}) &= (\pm1,0,0) \text{ for } \Lambda=\Lambda^{\pm}_{LL}, \\ (\eta_{LL},\eta_{RR},\eta_{LR}) &= (0,\pm1,0) \text{ for } \Lambda=\Lambda^{\pm}_{RR}, \\ (\eta_{LL},\eta_{RR},\eta_{LR}) &= (\pm1,\pm1,\pm1) \text{ for } \Lambda=\Lambda^{\pm}_{VV}, \\ (\eta_{LL},\eta_{RR},\eta_{LR}) &= (\pm1,\pm1,\mp1) \text{ for } \Lambda=\Lambda^{\pm}_{AA}.\end{aligned} \tag{3.4}$$

Such interactions can arise by constituent interchange (when the fermions have common constituents, e.g. for $qq{\rightarrow}qq$) and/or by exchange of the binding quanta, whenever binding quanta couple to constituents of both particles.

Another typical consequence of compositeness is the appearance of excited leptons and quarks (l^* and q^*). Phenomenologically, an excited lepton (quark) is defined to be a heavy lepton (quark) sharing its leptonic number (flavour) with one of the existing leptons (quarks).

The contact interaction compositeness model used in this study is depicted in Fig. 3.1 and described by the following effective Lagrangian:

$$\mathcal{L}_{qqqq}(\Lambda) = \frac{\eta g^2}{2(\Lambda^{+}_{LL})^2_q}\bar{\Psi}^L_q\gamma^\mu\Psi^L_q\bar{\Psi}^L_q\gamma^\mu\Psi^L_q, \tag{3.5}$$

where $\eta = 1$ (destructive interference) and $g^2/4\pi = 1$. Only quarks are considered composite, the case of composite leptons is not studied. For the sake of

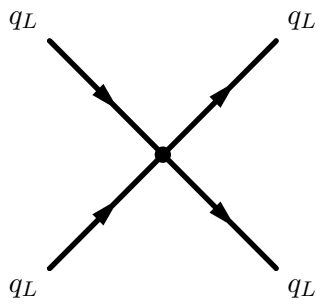


Figure 3.1: Feynman diagram of contact four-fermion interaction.

brevity let us further omit the sub- and superscripts of the compositeness scale: $\Lambda \equiv \Lambda^{+}_{LL}$. The lagrangian used for our simulations is the sum of the lagrangian in Eq. (3.5) and QCD langrangian

$$\mathcal{L}(\Lambda) = \mathcal{L}_{QCD} + \eta \mathcal{L}_{qqqq}(\Lambda). \tag{3.6}$$

When calculating relevant cross-sections, this yields for instance

$$\frac{\mathrm{d}\sigma^{\Lambda}}{\mathrm{d}p_{\mathrm{T}}} = \frac{\mathrm{d}\sigma}{\mathrm{d}p_{\mathrm{T}}}(QCD) + \frac{1}{\Lambda^2} I + \frac{1}{\Lambda^4} C, \tag{3.7}$$

where the first term on the r.h.s. is the Standard Model QCD contribution and p_{T} is the transverse momentum of an inclusive jet. I is caused by an interference of QCD and the contact term (CT), and C is the pure contact term contribution to the cross-section.

Employing Eq. (3.6) yields absolute squares of invariant amplitudes[1] [11]

$$|\mathcal{A}(ud{\to}ud)|^2 = \frac{4}{9}\alpha_s^2(Q^2)\frac{s^2+u^2}{t^2} + \left(\frac{\eta s}{\Lambda^2}\right)^2, \tag{3.8a}$$

$$|\mathcal{A}(u\bar{d}{\to}u\bar{d})|^2 = \frac{4}{9}\alpha_s^2(Q^2)\frac{s^2+u^2}{t^2} + \left(\frac{\eta u}{\Lambda^2}\right)^2, \tag{3.8b}$$

$$|\mathcal{A}(u\bar{u}{\to}d\bar{d})|^2 = \frac{4}{9}\alpha_s^2(Q^2)\frac{t^2+u^2}{s^2} + \left(\frac{\eta u}{\Lambda^2}\right)^2, \tag{3.8c}$$

$$\begin{aligned} |\mathcal{A}(uu{\to}uu)|^2 &= \frac{4}{9}\alpha_s^2(Q^2)\left[\frac{s^2+t^2}{u^2} + \frac{s^2+u^2}{t^2} - \frac{2}{3}\frac{s^2}{ut}\right] \\ &\quad + \frac{8}{9}\alpha_s^2(Q^2)\frac{\eta}{\Lambda^2}s^2\left(\frac{1}{t}+\frac{1}{u}\right) + \frac{8}{3}\left(\frac{\eta s}{\Lambda^2}\right)^2, \end{aligned} \tag{3.8d}$$

$$\begin{aligned} |\mathcal{A}(u\bar{u}{\to}u\bar{u})|^2 &= \frac{4}{9}\alpha_s^2(Q^2)\left[\frac{u^2+t^2}{s^2} + \frac{s^2+u^2}{t^2} - \frac{2}{3}\frac{u^2}{st}\right] \\ &\quad + \frac{8}{9}\alpha_s^2(Q^2)\frac{\eta}{\Lambda^2}u^2\left(\frac{1}{t}+\frac{1}{s}\right) + \frac{8}{3}\left(\frac{\eta u}{\Lambda^2}\right)^2, \end{aligned} \tag{3.8e}$$

such that

[1] Substitution $q \leftrightarrow \bar{q}$ on l.h.s. yields substitution $s{\leftrightarrow}u$ on r.h.s.

$$\frac{\mathrm{d}\sigma}{\mathrm{d}t} = \frac{\pi}{s^2}|\mathcal{A}|^2. \tag{3.9}$$

The strong coupling constant α_s in Eqs. (3.8) is to be evaluated at a scale Q^2 typical of the process in question. The choice of Q^2 is somewhat ambiguous and thus brings in a systematic error, evaluated in Sec. 7.2.2 and Sec. 8.2.2.

3.2.1 Quark compositeness in Pythia

The Pythia program [14] was used to generate a large part of the Monte Carlo data for this study. For this purpose, several questions or obstacles must have been solved. At some point, these solutions have eventually become a part of Pythia. Analysis that used older versions of Pythia are biased. Therefore it might be useful to outline, what was found.

Firstly, there are two references in Pythia manual [13] concerning compositeness. These are [12] and [11]. The formulas for invariant amplitudes $|\mathcal{A}(ud{\to}ud)|^2$ and $|\mathcal{A}(uu{\to}uu)|^2$ referenced therein are in contradiction. Eqs. (3.8) are in agreement with [11] and we consider them as correct, [15]. We have verified, that these formulas were also implemented in Pythia [17].

Secondly, there was a problem of implementation of the interference term in Eq. (3.7), which was after our intervention solved in Pythia 6.223 [17].

Finally, there was an obstacle with double counting of initial-state radiation, when compositeness was switched on[2]. It was solved in version 6.404, [18, 16]. See also Fig. 3.2 for an illustration.

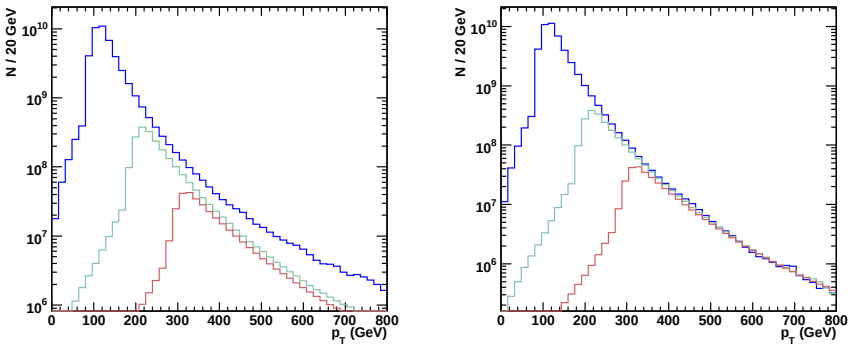


Figure 3.2: Inclusive dijet p_T spectra with various lower p_T cuts (100, 200 and 300 GeV), generated in Pythia with quark compositeness switched on (Λ = 3 TeV, MSEL=51, ITCM(5)=2, RTCM(42)=1), corresponding integral luminosity 30 fb^{-1}. Situation before version 6.404 (left) and from 6.404 on (right), where the spectra at higher p_T agree correctly with each other.

[2]To be precise, when MSEL=51 was turned on.

3.2.2 NLO k-factors

The inherent uncertainty in the lowest order calculation derives from its dependence on the unphysical renormalisation and factorisation scales, which is often large. In order to compare with predictions that have smaller theoretical uncertainties Next-to-leading order (NLO) calculations should be used. When the NLO calculations of matrix elements are not available for a given physics channel, that is also the case for quark compositeness, they can be approximated by using k-factor - a simple ratio of NLO and LO cross-sections. The ratio can depend quite strongly on the pdfs that were used in both the NLO and LO evaluations. k-factor depends strongly on the renormalisation and factorisation scales at which it is evaluated.

The k-factor for quark compositeness is applied as follows:

$$\sigma(sm+ct)_{NLO} = \frac{\sigma(sm)_{NLO}}{\sigma(sm)_{LO}}\sigma(sm+ct)_{LO}, \tag{3.10}$$

where $\sigma(sm)$ is a cross-section for a given observable with QCD only and $\sigma(sm+ct)$ is a cross-section including the contact term (3.5) switched on.

3.3 Bayesian Technique

Before proceeding with Sec. 3.4, where present-day quark compositeness limits are presented, it is necessary to illustrate Bayesian technique, [22], a widely used statistical approach, when determining a lower limit on compositeness scale Λ. Bayesian technique is also used in this study.

3.3.1 Bayes' theorem

Let us have probability functions $P(A)$, $P(B)$ and conditional probability functions $P(A|B)$ and $P(B|A)$. The Bayes' theorem reads

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}. \tag{3.11}$$

When combined with law of total probability

$$P(B) = \sum_i P(B|A_i)P(A_i) \tag{3.12}$$

it gives

$$P(A|B) = \frac{P(B|A)P(A)}{\sum_i P(B|A_i)P(A_i)}, \tag{3.13}$$

where the subset A could, for example, be one of the A_i. The most commonly, the probability $P(A)$ is assigned a value equal to the limiting frequency of occurrence of A. This interpretation forms the basis of frequentists statistics.

Using subjective probability, however, $P(A)$ is interpreted as a degree of belief that the hypothesis A is true. Subjective probability is used in Bayesian, as opposed to frequentists, statistics. Bayes' theorem can be written as

$$P(theory|data) \propto P(data|theory)P(theory), \tag{3.14}$$

where $P(theory)$ is the *prior* probability, that the theory (any hypothesis) is true. *Prior* must be guessed and is subjective. $P(data|theory)$ is a probability to have gotten the data actually obtained, given the theory, which is also called the *likelyhood*. We are interested in resulting $P(theory|data)$ - *posterior* probability. Using subjective *prior*, being a matter of choice, brings a certain systematic error into the analysis. Nevertheless, the *posterior* should depend on *prior* only marginally, if the analysis is to be taken seriously.

Finally, the lower limit on compositeness scale Λ_{lim} corresponding to chosen confidence level, C.L., is defined as:

$$\int_0^{\Lambda_{lim}^{-2}} P(\Lambda^{-2}|d_0)\mathrm{d}\Lambda^{-2} = \text{C.L.} \tag{3.15}$$

3.4 Present-day quark compositeness limits

The latest lower limits on quark compositeness come from analysis of the Tevatron data [25] and were obtained using Bayesian technique with prior flat in $\xi = 1/\Lambda$, CTEQ3M pdfs and a scale $\mu = 0.5E_{\mathrm{T}}^{max}$. The values of the limits are:

$\Lambda_{LL}^{+} > 2.7$ TeV at 95 % C.L.,
$\Lambda_{LL}^{-} > 2.4$ TeV at 95 % C.L.

A comparison of dijet invariant mass data and theory is depicted in Fig. 3.3.

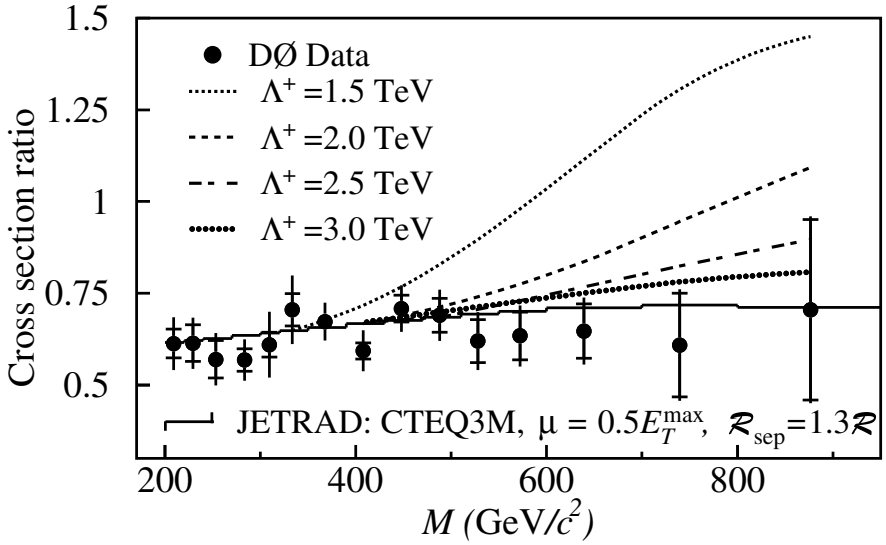


Figure 3.3: The ratio of cross sections for $|\eta_{jet}| < 0.5$ and $0.5 < |\eta| < 1.0$ for data (solid circles) and theory (various lines). The error bars show the statistical and systematic uncertainties added in quadrature, and the crossbar shows the size of the statistical error. Taken from [25].

Chapter 4

The ATLAS detector at LHC

The new physics discoveries, outlined in Sec. 1.1, as well as precision measurements in SM sector will require new experiments with outstanding performance, both physical and technical. Nowadays there are four such experiments being built along the LHC ring at CERN: two general-purpose experiments ATLAS [29, 30, 31], and CMS [32], experiment ALICE [33] dedicated to the study of heavy-ion collisions, and LHCb [34], dedicated to the study of B mesons rare decays and CP violation.

In this Chapter let us start with a short overview of the LHC machine, providing pp collisions at the centre-of-mass energy $\sqrt{s} = 14$ TeV at a design luminosity of 10^{34} cm^{-2}s^{-1} and then focus on the description of the ATLAS experiment, whose potential to reveal quark compositeness we study.

4.1 The Large Hadron Collider

The Large Hadron Collider (LHC), [35], is a superconducting proton synchrotron, being constructed at CERN inside the 27 km long tunnel used in the past years by LEP. LHC will accelerate two beams of protons (or heavy ions), travelling in opposite directions, from 450 GeV to 7 TeV. The existing machines at CERN will do the first stages of acceleration (Fig. 4.1): first the protons are accelerated up to 50 MeV in the proton linac, then the Proton Synchrotron Booster (PSB) will let them reach 1.8 GeV. The Proton Synchrotron (PS) will accelerate them up to 26 GeV. Finally, the Super Proton Synchrotron (SPS) will be used to inject the 450 GeV protons into the LHC.

The LHC dipole magnets will house in one single twin bore magnet (with the same yoke and cryostat) two different magnetic channels. The magnets will provide a magnetic field of 8.36 T, which allows the colliding beam particles to reach the design energy of 7 TeV.

At the LHC energies, a total inelastic non-diffractive pp cross-section is about 70 mb. Since the interesting processes have cross-sections several orders of magnitude lower (the cross-section for $t\bar{t}$ pairs production is $\sigma_{t\bar{t}} \simeq 800$ pb, the inclusive Higgs production cross-section is well below 1 nb for any Higgs mass), a very high luminosity and interaction rate are needed.

Before reaching the design luminosity, it is foreseen to have one year at

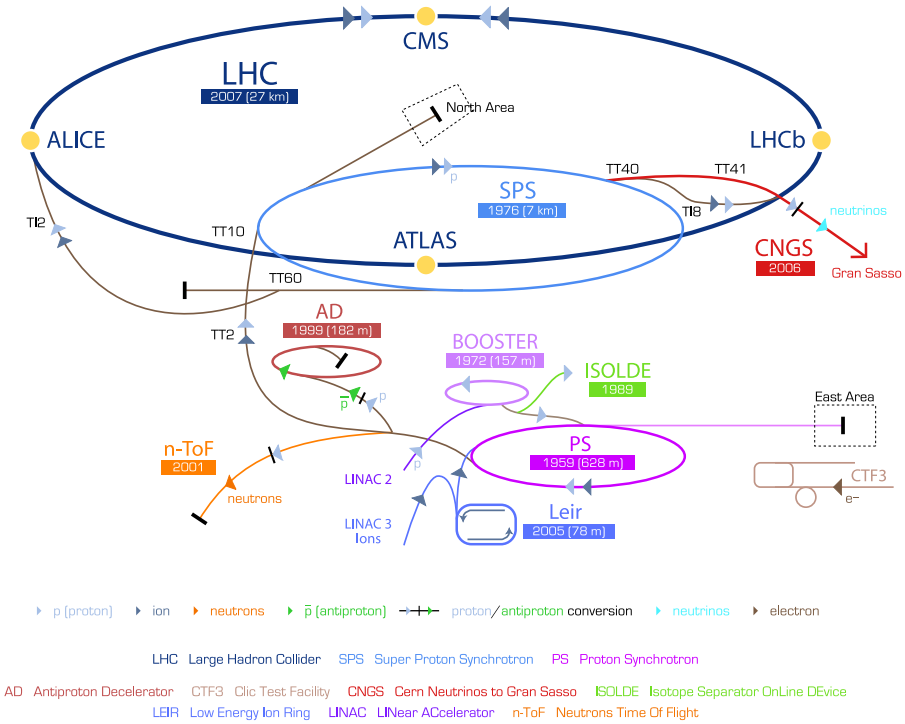


Figure 4.1: Layout of CERN accelerator complex. Before the particles are finally injected into the LHC, they are pre-accelerated in a chain of other accelerators. It is also worth noticing that the oldest part of this complex, the PS, has been in operation since 1959. Image not to scale. Courtesy of CERN Press office.

the so called low luminosity (2.10^{33} cm^{-2}s^{-1}), which should provide enough statistics to perform the first physics measurements and machine and detector tuning.

The design luminosity of 10^{34} cm^{-2}s^{-1} with 2808 bunches crossing at 25 ns intervals is expected to be reached a few years after the first collisions. With the inelastic proton proton cross-section of 70 mb, this gives approximately 23 events per bunch-crossing, or a total rate of 10^9 inelastic events/sec. This means around 1000 particles will emerge from the collision points every 25 ns within pseudorapidity $|\eta|$ less than 2.5. A serious experimental difficulty at LHC is therefore that every candidate event for new physics will on average be accompanied by 23 inelastic events occurring simultaneously in the detector. The main parameters of the LHC for pp and heavy-ion collisions are shown in Tab. 4.1.

Parameters	p–p	Pb^{82+}–Pb^{82+}
Beam energy (TeV)	7.0	7.0
Centre of mass energy (TeV)	14	1262
Injection energy (GeV)	450	190.6
Bunch spacing (ns)	25	124.75
Particles per bunch	1×10^{11}	6.2×10^7
R.M.S. bunch length (m)	0.075	0.075
Number of bunches	2808	608
Initial luminosity (cm^{-2}s^{-1})	10^{33}	1.95×10^{27}
Luminosity (cm^{-2}s^{-1})	10^{34}	1.8×10^{27}
Luminosity lifetime (h)	10	10
Dipole field (T)	8.3	8.3

Table 4.1: The LHC settings for pp and heavy-ion collisions.

4.2 The ATLAS detector design criteria

In this section let us first briefly describe the coordinate system used for the ATLAS detector and then proceed with its design criteria.

The beam direction defines the z-axis and the x-y plane is transverse to the beam direction. The positive x-axis is defined as pointing from the interaction point to the centre of the LHC ring and the positive y-axis is defined as pointing upwards. The side-A of the detector is defined as that with positive z and side-C is that with negative z. The azimuthal angle ϕ is measured around the beam axis, and the polar angle θ is the angle from the beam axis. The pseudorapidity η is defined as $\eta = -\ln tan(\theta/2)$. The transverse momentum p_{T}, the transverse energy E_{T}, and the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ are defined in the x-y plane unless stated otherwise.

The design criteria for the ATLAS (A Toroidal LHC ApparatuS) detector are as follows:

- Large acceptance in pseudorapidity (η) with almost full azimuthal (ϕ) angle coverage.
- Good charged particle momentum resolution and reconstruction efficiency in the inner tracker.
- Very good electromagnetic (EM) calorimetry for electron and photon identification and measurements completed by full-coverage hadronic calorimetry for accurate jet and missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) measurements.
- Good stand-alone muon identification and momentum resolution over a wide range of momenta and the ability to determine unambiguously the charge of high p_{T} muons.
- Triggering on low transverse momenta objects is important to maintain high kinematic efficiency with sufficient background rejection to obtain an acceptable trigger rate for most physics processes of interest at the LHC.
- Due to the experimental conditions at the LHC, the detector needs fast, radiation hard electronics and sensor elements. In addition, a very high granularity is needed to be able to handle the particle fluxes and to reduce the influence of overlapping events.

The criteria for the ATLAS calorimeters design follow from the physics processes to be measured [59]. For EM calorimeters following stringent requirements are defined:

- High jet energy resolution at $E_{\mathrm{T}} \approx 100$ GeV and coverage to low E_{T} required by rare physics events $\mathrm{H} \rightarrow \gamma\gamma$ and $\mathrm{H} \rightarrow \mathrm{ZZ} \rightarrow 4e$,
- e reconstruction down to a few GeV required for b-physics measurements,
- dynamic range of the EM calorimeters should span from mip to $\mathrm{Z}' \rightarrow ee$ at a few TeV.

In order to fulfil these requirements, the EM calorimeters should have the following performance at $|\eta| < 2.5$:

- energy resolution (E in GeV):

$$\frac{\sigma(E)}{E} = \frac{8-11\%}{\sqrt{E}} \oplus \frac{0.2-0.4}{E} \oplus 0.7\%, \tag{4.1}$$

- linearity better than 0.1%.

In case of hadron and forward calorimeters the physics channels $\mathrm{H} \rightarrow \mathrm{WW} \rightarrow jjX$ and Z, W and top quark production require good dijet mass resolution. Moreover Higgs fusion requires good forward jet tagging and $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement demands good jet energy resolution and small jet energy scale uncertainties. It should be stressed, that jet energy scale uncertainty is one of the most limiting factors for measurement of quark compositeness, as will be discussed in Chapters 7 and 8. The hadronic and forward calorimeters should thus meet following criteria of jet energy resolution:

- for $|\eta| < 3$:

$$\frac{\sigma(e)}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\% \tag{4.2}$$

- for $3 < |\eta| < 4.9$:

$$\frac{\sigma(e)}{E} = \frac{100\%}{\sqrt{E}} \oplus 5\%. \tag{4.3}$$

4.3 The detector description

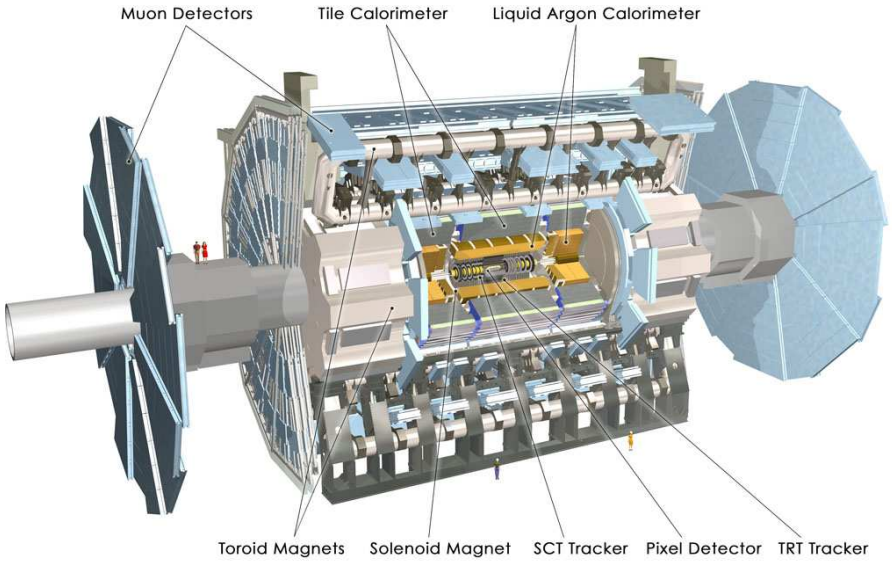


Figure 4.2: The ATLAS detector model. The diameter is 22 m and overall length 46 m. The overall weight of the ATLAS detector is approximately 7000 tons, most of it contained in calorimeters. Courtesy of CERN, Geneva.

The overall detector layout is shown in Fig. 4.2. The two independent magnetic systems (the solenoidal one in the Inner Detector and the toroidal one in the muon spectrometer) consist of a thin super-conducting solenoid surrounding the Inner Detector and three sets of eight independent coils arranged with an eight-fold symmetry outside the calorimeters. The central solenoid envelopes the Inner Detector and provides a 2 Tesla field oriented along the beam axis. The Inner Detector makes use of three different technologies, at different distances from the interaction point (IP). Three inner layers of pixels allow good secondary vertex identification and, together with the four layers of silicon micro-strips, good momentum measurements. The tracking is then completed by continuous straw-tube detectors with transition radiation detection capability in the outer part.

The calorimetry uses radiation-hard liquid argon (LAr) technology for the EM barrel and end-cap, for the Hadronic End-Cap (HEC) and for the Forward (FCAL) calorimeters. In the barrel region the cryostat is shared with the superconducting solenoid, while the EM end-cap, the HEC and the FCAL share

the same cryostat in the forward region. In the barrel region the longitudinal hermeticity is provided by the Tile calorimeter (Tilecal). Scintillating tiles are used as an active material, while the passive material is iron. The Tilecal is subdivided into three regions: a barrel region ($|\eta| < 1$) and two extended barrel ($1 < |\eta| < 1.7$) regions. The two gaps between these three regions are covered by the Inter Tilecal Calorimeters (ITC) and the Intermediate Gap Scintillators, which allow the recover part of the energy lost in the gap.

Outside the calorimeters there is the Muon Spectrometer (MS). The magnetic field is provided by the 25 m long coils in the central region. The coverage at small angles is completed by two end-cap toroids. The magnetic field bends the particles inside the open structure that constitutes the support for the muon chambers. The multiple scattering is therefore minimised. This allows a very good measurement of the muon momentum with three stations of high precision tracking chambers. The muon detector includes fast response trigger chambers, which operate in coincidence to provide a fast trigger decision on the muon p_{T}.

The total radius of the ATLAS experiment, from the interaction point to the last muon chamber, is about 11 m. The total length is about 46 m, the overall weight about 7000 tons. Almost 90% of the total ATLAS volume is occupied by the toroids and by the Muon Spectrometer.

The ATLAS detector is being installed and commissioned at Point 1 of LHC ring. The status of the installation as of February 2007 is shown in Fig. 4.3.

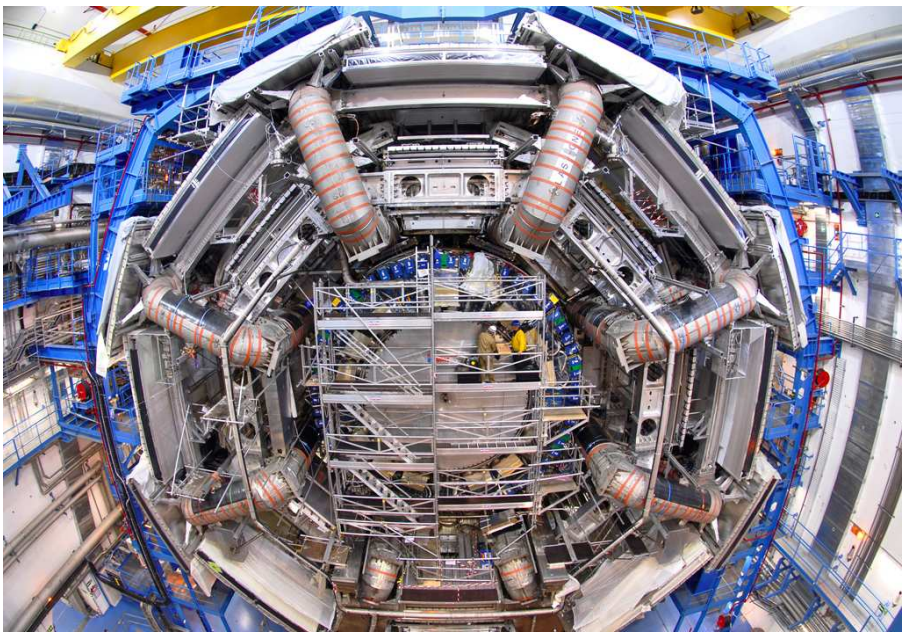


Figure 4.3: The ATLAS detector during assembly in the underground cavern as of February 2007. Parts of cryostats of eight large barrel toroid coils are visible, surrounding Tile Calorimeter with inserted LAr cryostat. Muon stations among the BT coils are also visible. Courtesy of CERN, Geneva.

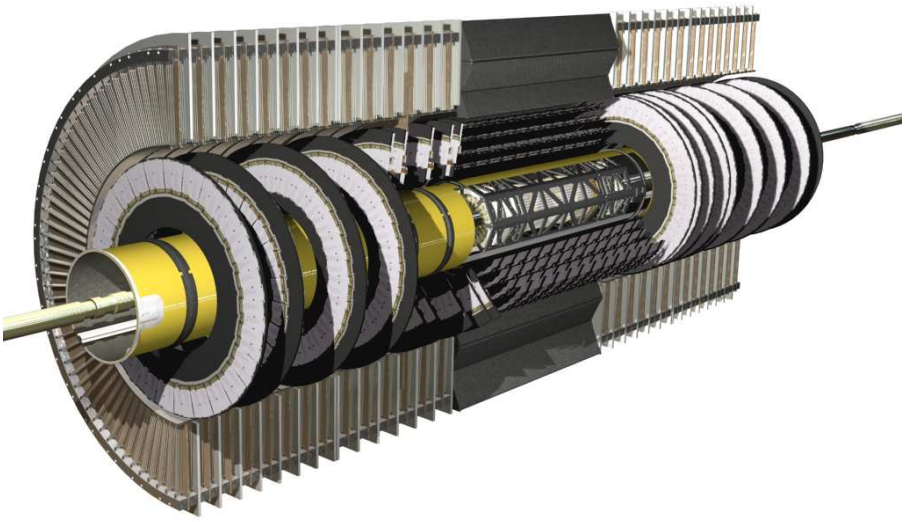


Figure 4.4: Overall view of the inner detector.

4.3.1 Tracking

The strategy used for the ATLAS Inner Detector (ID) [36] is to combine few high precision measurements close to the interaction point with a large number of lower precision measurements in the outer radius. The inner detector is embedded in the 2 T magnetic field provided by the Central Solenoid (CS). The CS extends over a length of 5.3 m and has a bore of 2.5 m. The position of CS in front of the EM calorimeter provides a careful minimisation of the dead material in order to achieve the desired calorimeter performance. As a consequence, the CS and the LAr calorimeter share one common vacuum vessel.

The structure of the inner detector is shown in Fig. 4.4. Within a radius of 56 cm from the interaction point, pixel and silicon micro-strip technologies offer a fine-granularity, thus a high precision of the track measurement. Charged particle proceeding from interaction point typically hits three layers of pixels (which measure both $R-\phi$ and z coordinate) and 4 layers of SCT modules, for a total of 7 tracking points. In each SCT module, two strip sensors are mounted together, with 40 mrad angle between them, allowing the measurement of the three coordinates. In the barrel region (which covers up to $|\eta| = 1$ for a total length of 160 cm), the pixels and SCT are arranged in concentric cylinders around the beam axis, while in the end-cap (up to $|\eta| = 2.5$) they are arranged in disks perpendicular to the beam axis.

A large number of tracking points (36) is provided by the Transition Radiation Tracker (TRT) that also can give e/π separation identifying the transition radiation emitted by electrons travelling at a high speed. TRT consists of straw tubes arranged parallel to the beam axis in the barrel region and in wheels around the beam axis in the end-cap. The reduced resolution with respect to the silicon detectors is compensated by the higher radius and by the number

of points measured. Therefore, the relative precisions of the measurements of the TRT and pixels/SCT are comparable. The TRT detector is intrinsically radiation hard.

The outer radius of the inner detector cavity is 115 cm, while the total length is 7 m. The layout provides full tracking coverage within $|\eta| < 2.5$, including impact parameter measurement and vertexing for heavy flavours and τ tagging. The expected precision for the detector is

$$\sigma_{R-\phi}(\mu\text{m}) = 13 \oplus \frac{62}{p_{\text{T}}\sqrt{\sin\theta}}, \tag{4.4}$$

$$\sigma_{z}(\mu\text{m}) = 39 \oplus \frac{90}{p_{\text{T}}\sqrt{\sin\theta}}. \tag{4.5}$$

While the radiation impact is lower on the TRT detector, it is not in particular for the pixels, which are more exposed to the radiation since they are closer to the interaction point. The intrinsic radiation weakness of the silicon would probably impose their substitution after a few years of operation, depending on the luminosity profile.

4.3.2 Calorimeters

A view of the ATLAS calorimeters is presented in Fig. 4.5. The calorimetry consists of an electromagnetic calorimeter covering the pseudorapidity region $|\eta| < 3.2$, a hadronic barrel calorimeter covering $|\eta| < 1.7$, hadronic end-cap calorimeters covering $1.5 < |\eta| < 3.2$, and forward calorimeters covering $3.1 < |\eta| < 4.9$. Over the pseudorapidity range $|\eta| < 1.8$, the EM calorimeter is preceded by a presampler detector.

The EM calorimeter is a lead/liquid-argon detector (LAr) with accordion geometry. The barrel EM calorimeter is contained in a barrel cryostat, which surrounds the inner detector cavity. The hadronic barrel iron/scintillating tiles calorimeter (Tilecal) is a cylinder divided into three sections: the central barrel and two identical extended barrels. Two end-cap cryostats house the LAr end-cap EM and hadronic calorimeters, as well as the integrated forward calorimeter. The hadronic end-cap calorimeter is a copper LAr detector with parallel-plate geometry, and the forward calorimeter, a dense LAr calorimeter with rod-shaped electrodes in a copper and tungsten matrix. The barrel and extended barrel tile calorimeters support the LAr cryostats and also act as the main solenoid flux return.

The pseudorapidity coverage, granularity, and longitudinal segmentation of the calorimeters are summarised in Tab. 4.2.

The LAr electromagnetic calorimeter

The EM calorimeter, [45], is divided into a barrel part ($|\eta| < 1.5$) and two end-caps ($1.4 < |\eta| < 3.2$). The barrel calorimeter consists of two identical half-barrels, separated by a small gap (6 mm) at $z = 0$. Each end-cap calorimeter is mechanically divided into two coaxial wheels: an outer wheel covering the region

EM Calorimeter	Barrel	End-Cap	
Coverage	$\|\eta\| < 1.475$	$1.375 < \|\eta\| < 3.2$	
Long. segmentation	3 samplings	3 samplings	$1.5 < \|\eta\| < 2.5$
		2 samplings	$1.375 < \|\eta\| < 1.5$
			$2.5 < \|\eta\| < 3.2$
Granularity ($\Delta\eta \times \Delta\phi$)			
Sampling 1	0.003×0.1	0.25×0.1	$1.375 < \|\eta\| < 1.5$
		0.003×0.1	$1.5 < \|\eta\| < 1.8$
		0.004×0.1	$1.8 < \|\eta\| < 2.0$
		0.006×0.1	$2.0 < \|\eta\| < 2.5$
		0.1×0.1	$2.5 < \|\eta\| < 3.2$
Sampling 2	0.025×0.025	0.025×0.025	$1.375 < \|\eta\| < 2.5$
		0.1×0.1	$2.5 < \|\eta\| < 3.2$
Sampling 3	0.05×0.025	0.05×0.025	$1.5 < \|\eta\| < 2.5$
Presampler	Barrel	End-cap	
Coverage	$\|\eta\| < 1.52$	$1.5 < \|\eta\| < 1.8$	
Long. segmentation	1 samplings	1 samplings	
Granularity ($\Delta\eta \times \Delta\phi$)	0.025×0.1	0.025×0.1	
Tilecal	Barrel	Extended Barrel	
Coverage	$\|\eta\| < 1.0$	$0.8 < \|\eta\| < 1.7$	
Long. segmentation	3 samplings	3 samplings	
Granularity ($\Delta\eta \times \Delta\phi$)			
Sampling 1 and 2	0.1×0.1	0.1×0.1	
Sampling 3	0.2×0.1	0.2×0.1	
Hadronic LAr		End-cap	
Coverage		$1.5 < \|\eta\|3.2$	
Long. segmentation		4 samplings	
Granularity ($\Delta\eta \times \Delta\phi$)	0.1×0.1	$1.5 < \|\eta\| < 2.5$	
	0.2×0.2	$2.5 < \|\eta\| < 3.2$	
FCAL Calorimeter		Forward	
Coverage		$3.1 < \|\eta\| < 4.9$	
Long. segmentation		3 samplings	
Granularity ($\Delta\eta \times \Delta\phi$)		0.2×0.2	

Table 4.2: Design parameters of the ATLAS calorimeters.

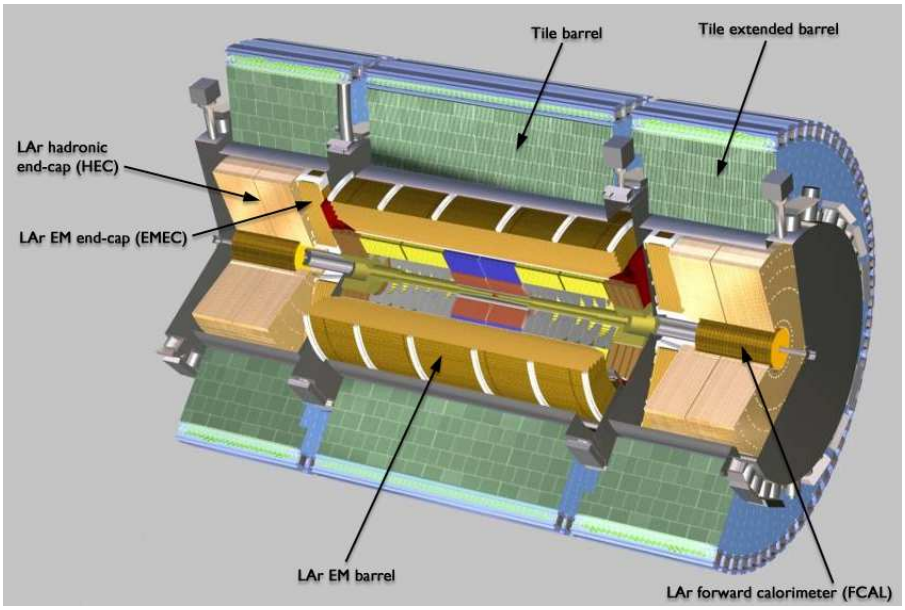


Figure 4.5: Overall view of the ATLAS calorimeters.

$1.375 < |\eta| < 2.5$, and an inner wheel covering the region $2.5 < |\eta| < 3.2$. The EM calorimeter is a lead LAr detector with accordion-shaped kapton electrodes and lead absorber plates over its full coverage. The accordion geometry provides complete symmetry without azimuthal cracks. The lead thickness in the absorber plates has been optimised as a function of η in terms of EM calorimeter performance in energy resolution. The total thickness of the EM calorimeter is larger than 24 radiation lengths (X_0) in the barrel and larger than 26 X_0 in the end-caps. Over the region devoted to precision physics ($|\eta| < 2.5$), the EM calorimeter is segmented into three longitudinal sections. For $|\eta| > 2.5$, i.e. for the end-cap inner wheel, the calorimeter is segmented in two longitudinal sections and has a coarser lateral granularity than for the rest of the acceptance. This is sufficient to satisfy the physics requirements (reconstruction of jets and measurement of $E_{\mathrm{T}}^{\mathrm{miss}}$).

The total material seen by an incident particle before the calorimeter front face is approximately 2.3 X_0 at $\eta = 0$, and increases with pseudorapidity in the barrel because of the particle trajectory angle. In region $|\eta| < 1.8$, a presampler is used to correct for the energy lost by electrons and photons upstream of the calorimeter. The presampler consists of an active LAr layer of thickness 1.1 cm (0.5 cm) in the barrel (end-cap) region.

At the transition between the barrel and the end-cap calorimeters, i.e. at the boundary between the two cryostats, the amount of material in front of the calorimeter reaches a localised maximum of about 7 X_0 . In this region, the presampler is complemented by a scintillator slab inserted in the crack between the barrel and end-cap cryostats and covering the region $1.0 < |\eta| < 1.6$. The region $1.37 < |\eta| < 1.52$ is not used for precision physics measurements involving photons because of the large amount of material situated in front of the EM

calorimeter.

The linearity of the EM calorimeters has been verified (both for the barrel and for the extended barrel) with electron test beams up to 350 GeV. Obtained linearity is better than 1%. The resolution found for the barrel at $\eta = 0.9$ is

$$\frac{\sigma(E)}{E} = \frac{10\%}{\sqrt{E(GeV}} \oplus \frac{0.38 \text{ GeV}}{E} \oplus 0.3\%. \tag{4.6}$$

The hadronic calorimeters

The ATLAS hadronic calorimeters cover the range $|\eta| < 4.9$ using different techniques suited for the widely varying requirements and radiation environment over the large η-range.

An important parameter in the design of the hadronic calorimeter is its thickness: it has to provide good containment for hadronic showers and reduce to a minimum punch-through into the muon system. Close to 10 λ of active calorimeter are adequate to provide good resolution for high energy jets. The total thickness, including 1.5 λ from the outer support, is 11 λ interaction lengths (λ) at $\eta = 0$ and has been shown both by measurements and simulation to be sufficient to reduce the punch-through well below the irreducible level of prompt or decay muons. Together with the large η-coverage, this will also guarantee a good $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement, which is important for many physics signatures and in particular for SUSY particle searches.

The Tile calorimeter. The large hadronic barrel calorimeter, [46], is a non-compensating sampling calorimeter using iron as the absorber and scintillating tiles as the active material. The tiles are placed radially and staggered in depth. The structure is periodic along z. The scintillating tiles are 3 mm thick and the total thickness of the iron plates in one period is 14 mm. The two opposite sides of the scintillating tiles are read out by wavelength shifting (WLS) fibres into two separate photo-multipliers (PMTs), that is read out channels. A read out cell of the calorimeter is formed by a group of scintillating tiles read out by such two channels. The read out cells then form pseudoprojective towers, providing an essential information for the detector trigger. The tile calorimeter is composed of one barrel and two extended barrels. Radially, the tile calorimeter extends from an inner radius of 2.28 m to an outer radius of 4.25 m. It is longitudinally segmented in three layers, approximately 1.4, 4.0 and 1.8 interaction lengths thick at $\eta = 0$. Azimuthally, the barrel and extended barrels are divided into 64 modules. In η, the readout cells, built by grouping fibres into PMTs, are pseudo-projective towards the interaction region. The total number of channels is approximately 10,000. The calorimeter is placed behind the EM calorimeter. The total thickness at the outer edge of the tile-instrumented region is 9.2 λ at $\eta = 0$.

The barrel cylinder covers the region of $|\eta| < 1.0$. A vertical gap of 68 cm provides space for cables from the ID, feed-throughs, and 5 service pipes for the EM calorimeter and the CS; it also houses front-end electronics for the EM calorimeter. The extended barrel covers the region $0.8 < |\eta| < 1.7$. The energy lost in the inactive materials in the gaps between the tile barrel and

extended barrel calorimeters is sampled by the scintillators of ITC, which has the same segmentation as the rest of the tile calorimeter. It is composed of two radial sections attached on the face of the extended barrel. The ITC is extended further inwards by a scintillator sheet, covering the inner part of the extended barrel and extending to the region between the LAr barrel and end-cap cryostats over $1.0 < |\eta| < 1.6$. This scintillator samples the energy lost in the cryostat walls and other dead materials.

The Tilecal standalone resolution for single pions was measured at the test beam, [47]. Resolution was obtained from analysis of 20 to 350 GeV pion beams impinging at pseudorapidities 0.2 to 1.2:

$$\frac{\sigma(E)}{E} = \frac{(52 \pm 2)\%}{\sqrt{E(GeV)}} + (5.0 \pm 0.2\%). \tag{4.7}$$

LAr hadronic end-cap calorimeters (HEC). Each HEC consists of two independent wheels of outer radius 2.03 m. The upstream wheel is built out of 25 mm copper plates, while the second one further from the interaction point uses 50 mm plates. In both wheels, the 8.5 mm gap between consecutive copper plates is equipped with three parallel electrodes, splitting the gap into four drift spaces of about 1.8 mm. The readout electrode is the central one.

Each wheel is built out of 32 identical modules, assembled with fixtures at the periphery and at the central bore. Each wheel is divided into two longitudinal segments.

To minimise the dip in the material density at the transition between the end-cap and the forward calorimeter (around $|\eta| = 3.1$), the HEC calorimeter reaches $|\eta| = 3.2$, thereby overlapping the forward calorimeter.

The HEC standalone resolution for single pions (6–200 GeV) was measured at the test beam, [44]. The result is:

$$\frac{\sigma(E)}{E} = \frac{(70.6 \pm 1.5)\%}{\sqrt{E(GeV)}} \oplus (5.8 \pm 0.2). \tag{4.8}$$

LAr forward calorimeter. The forward calorimeter (FCal) has a front face at about 4.5 m from the interaction point and is integrated into the end-cap cryostat as this provides clear benefits in terms of uniformity of the calorimetric coverage as well as reduced radiation background levels in the muon spectrometer. In order to minimise the amount of neutron albedo in the ID cavity, the front face of the FCal is recessed by about 1.2 m with respect to the EM calorimeter front face. This severely limits longitudinal space for installing about 9.5 active interaction lengths, and therefore calls for a high-density design, which also avoids energy leakage from the FCal to its neighbours.

The FCal consists of three sections: the first one is made of copper, while the other two are made of tungsten. In each section the calorimeter consists of a metal matrix with regularly spaced longitudinal channels filled with concentric rods and tubes. The rods are at positive high voltage while the tubes and matrix are grounded. The LAr in the gap between is the sensitive medium. This geometry allows for an excellent control of the gaps which are as small as 0.25 mm in the first section.

4.3.3 Muon spectrometer

The conceptual layout of the muon spectrometer is shown in Fig. 4.6. One of the most important features of the muon spectrometer, [37], is the possibility of a precise standalone measurement of the muon momentum. The magnetic field provided by the superconducting air-core toroid magnets deflects the muon trajectories that are measured by high precision tracking chambers. The magnetic field in the $|\eta| < 1.4$ range is provided by the large barrel toroid (BT), [38], while the region $1.6 < |\eta| < 2.7$ is covered by two smaller end-cap toroid magnets (ECT) inserted at the both ends of the barrel toroid. In the so called transition region ($1.4 < |\eta| < 1.6$) the combined contributions of both the barrel and end-cap toroid magnets provide the magnetic field coverage. The magnetic field is mostly orthogonal to the muon trajectory in the covered pseudorapidity range, while the effect of multiple scattering is minimised. The ECT coil system is rotated by 22.5^o with respect to the BT coil system in order to provide optimised radial overlap. The BT provides a bending power of 2 to 6 Tm and the ECT contributes with approximately 1 to 8 Tm.

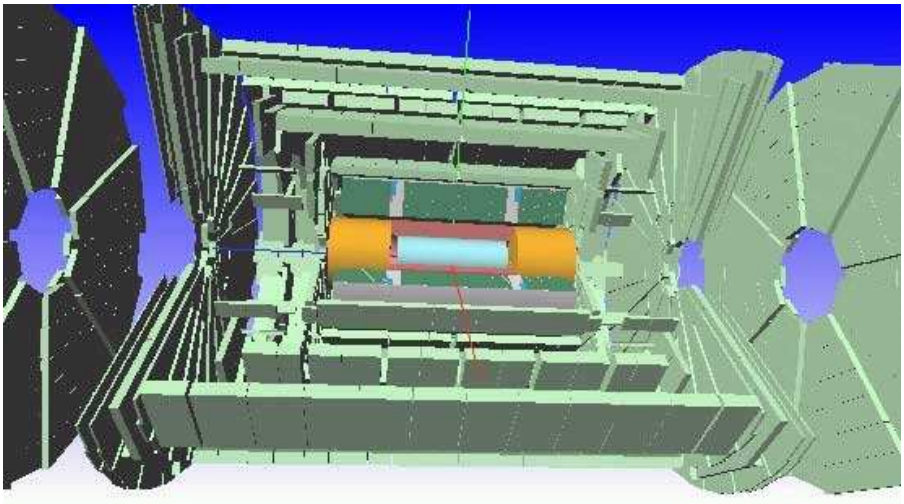


Figure 4.6: Overall view of the ATLAS muon system.

In the barrel region, the muon chambers are arranged in three cylindrical layers (stations), while in the end-cap regions they form three vertical walls. The transition region is instrumented with one extra station.

The azimuthal layout follows the magnet structure: there are 16 sectors. The so called Large sectors lie between the coils, and they overlap with the Small sectors, placed in correspondence with the coils themselves. The $R - z$ layout of the chambers is shown in Fig. 4.7.

The choice of the different chamber technologies has been driven by the particle fluxes foreseen in the different regions of the detector. Criteria of rate capability, granularity, ageing properties and radiation hardness have been considered. Tab. 4.3 summarises the chamber technologies used in the various pseudorapidity regions.

The measurement of the track bending coordinate (η) is provided (in most

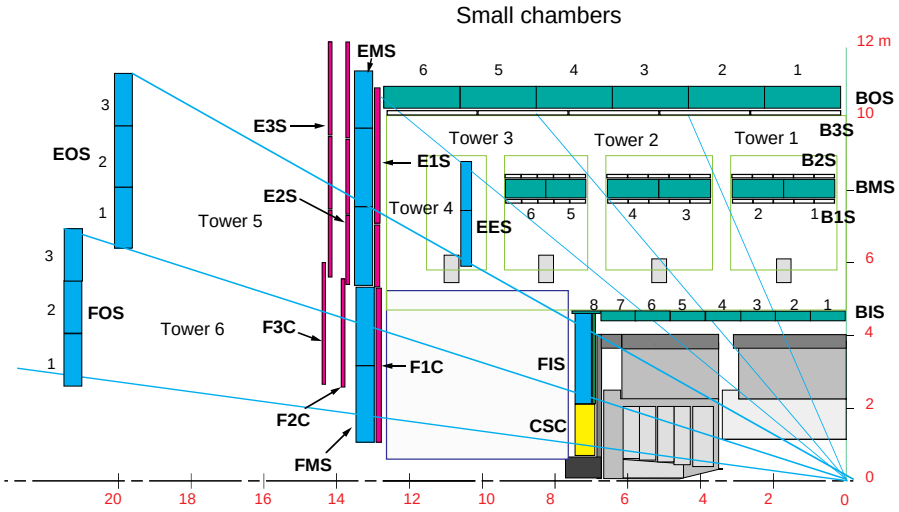


Figure 4.7: A view of one ϕ plane ($\phi = 22^o$) of the muon chamber layout.

of the η region) by the Monitored Drift Tubes (MDT), while at large pseudo-rapidity, the higher granularity Cathode Strip Chambers (CSC) are used.

Region		station I	station E	station M	station O
Barrel	$\|\eta\| < 1$	MDT		MDT RPC	MDT RPC
End-Caps	$1 < \|\eta\| < 1.4$	MDT TGC	MDT		
	$1.4 < \|\eta\| < 2$	MDT TGC		MDT TGC	
	$2 < \|\eta\| < 2.4$	CSC			MDT
	$2.4 < \|\eta\| < 2.7$	CSC		MDT TGC	

	TRIGGER CHAMBERS		PRECISION CHAMBERS	
Technologies used	RPC	TGC	MDT	CSC
Number of channels	354K	440K	372K	67K
Area (m^2)	3650	2900	5500	27
Time resolution	< 5 ns	< 7 ns	500 ns	<7 ns
Spatial resolution	5-10 mm		80 μm	60 μm

Table 4.3: Design parameters of the Muon spectrometer.

The requirements on the momentum resolution ($\Delta p_{\mathrm{T}}/p_{\mathrm{T}} \simeq 10\%$ at 1 TeV/c) call for an accuracy of the relative positioning of chambers traversed by a muon track that matches the intrinsic resolution and the mechanical tolerances of the precision chambers.

The knowledge of the chamber positioning with an accuracy of 30 μm is required within a projective tower. The accuracy required for the relative positioning of different towers to obtain adequate mass resolutions for multi-muon final states is in the millimetre range. This accuracy can be achieved by the initial positioning and survey of chambers at the installation time. The relative alignment of muon spectrometer, calorimeters and ID will rely on the measurement of the high-momentum muon trajectories.

The MDT chambers are equipped with a in-plane alignment system aiming at a measurement of the tube position displacements, with respect to their nominal positions at the assembly phase, with a precision better than 10 μm. In

order to achieve such a precision, the spectrometer is equipped with a RASNIK system: a laser, mounted at one side of a chamber, projects a pattern to a CCD camera positioned at the other end of the chamber. From the displacement of the pattern-figure with respect to what is expected, corrections for chambers deformation can be computed.

The chambers for the LVL1 muon trigger system covers the region $|\eta| < 2.4$. Resistive Plate Chambers (RPC) are used in the barrel region, while the Thin Gap Chambers (TGC) are used in the end-cap region. Their first task is to identify without any ambiguity the bunch crossing of the triggered event. This requires a time resolution better than 25 ns. Next, they have to provide a well defined p_{T}cut-off for the LVL1 choice. This is obtained considering a window of a size defined by the LVL1 p_{T}threshold considered on the second RPC (or TGC) station once a hit has been obtained in the first station. Finally, the trigger chambers measure the non-bending coordinate (ϕ), in a plane orthogonal to the one measured by the precision chambers, with a typical precision of 5–10 mm.

4.3.4 Forward detectors

Three smaller detector systems are built to cover the forward region of ATLAS. These are closely connected to the luminosity determination in ATLAS, but are in addition foreseen to study forward physics. In order of their distance from the ATLAS interaction point, the first system is a Cerenkov detector called LUCID (LUminosity measurement using Cerenkov Integrating Detector, [40]). LUCID is the main luminosity monitor in ATLAS, detecting inelastic pp scattering in the forward direction, and is located 17 m away from the interaction point.

The second system is the so-called Zero Degree Calorimeter (ZDC, [41]) which is located at a distance of 140 m from the IP. This corresponds to the location where the LHC beam-pipe is divided into two and the ZDC is located between the beam pipes, thus it will measure neutral particle at 0^o polar angle.

The most remote system is the Absolute Luminosity for ATLAS (ALFA) system. It consists of scintillating fibre trackers located inside roman pots at a distance of 240 m from the ATLAS IP, and determines the absolute luminosity of CERN LHC at ATLAS IP by measuring elastic scattering in the Coulomb interface region to a precision $\Delta L/L \sim 3\%$.

4.3.5 Data Acquisition, Trigger and Control systems

Data Acquisition

The ATLAS trigger and data acquisition systems (TDAQ) and the detector control system (DCS) are responsible for the data-flow of data from detector front-end electronics modules to data storage at CERN computing centre, for the selection of events and for the control and monitoring of the detector, [39].

A schematic diagram is presented in Fig. 4.8 and can be divided into four principal systems, namely:

- The Data Flow System - responsible for receiving of the detector data, serving data to the trigger system, and transporting the data for selected

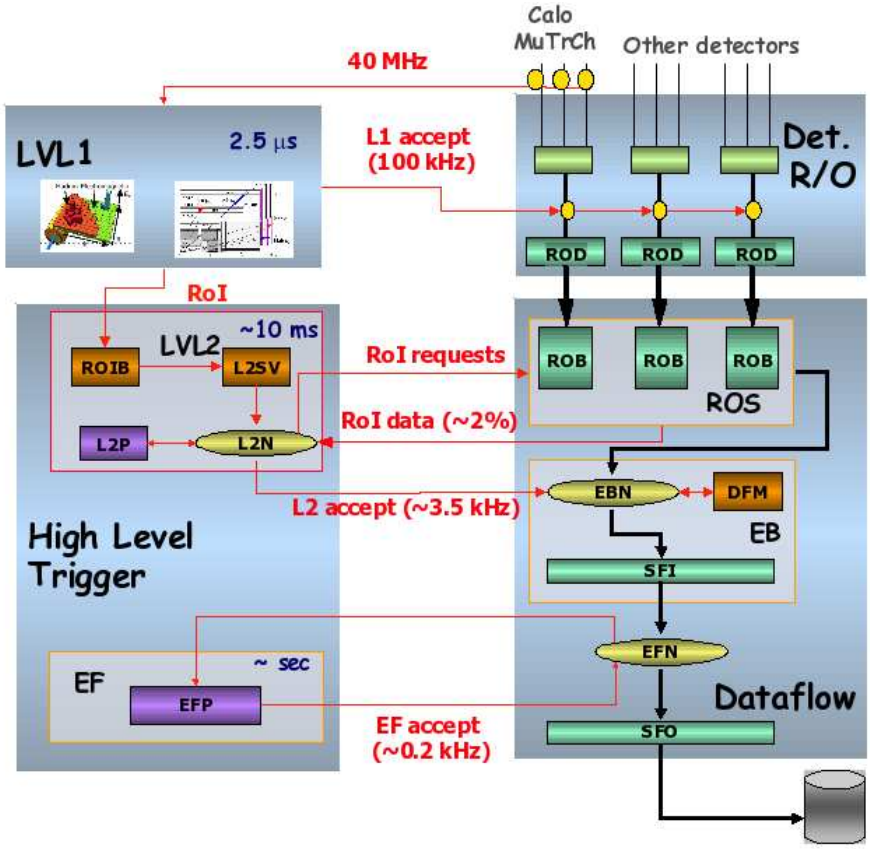


Figure 4.8: Main components of the Data Flow (right) and trigger systems (left). Black lines represent a path of a physics data, red lines movements of trigger data.

events to mass storage,

- The Trigger System - subdivided into three layers. The first level (LVL1) purely hardware based, other two levels software based (collectively called High Level Trigger, HLT),
- The Online System - responsible for all aspects of experiment and TDAQ operation and control during data-taking, and during testing and calibration runs, and
- The DCS - responsible for a coherent and safe operation of the ATLAS detector, as well as for the interface with external systems and services, including the LHC itself.

The Online system is implicitly understood to be connected to all elements in Fig. 4.8, and the DCS to all hardware elements which need to be monitored and controlled.

In the right-half of Fig. 4.8 the data flow system is represented in the different components. The signals coming from the detectors (upper part), read by the input-output systems are managed by the back-end electronics boards: the so called Read Out Drivers (ROD). They are detector-specific and they have the responsibility of the reading of the data from the front-end electronics when a Level 1 trigger request is propagated to the system. The RODs, equipped with Digital Signal Processors (DSP), have also the responsibility to make the first pre-formatting of the data to the ATLAS TDAQ final format and, in some cases, like for the calorimeters, apply some simple algorithms and performing calculations. From this point on, the components of the Data Flow are no more detector specific.

The data are then moved to the ROSes (Read Out System), normal PCs equipped with one or more buffer cards, the Read Out Buffers (ROBs). Here they will be further formatted and moved, after request, to the Event Building Network (EBN), passed to the last stage of trigger, and in case of acceptance, stored to a set of local disk pools (SFO: Sub Farm Output).

Trigger

Starting from an initial bunch-crossing rate of 40 MHz, the rate of selected events must be reduced to about 200 Hz for permanent storage. The event rate is determined by the total pp cross section, and is about 10^9 evt/sec at the nominal luminosity. Hence an overall rejection factor of about 5×10^6 against minimum bias events is required. This strong requirement must match an excellent trigger efficiency for rare physics processes of interest.

The LVL1 trigger makes an initial selection based on a reduced granularity information from a subset of detectors. Objects searched by the calorimeter trigger are high p_{T}electrons and photons, jets, and τs decaying into hadrons, as well as large missing and total transverse energies. High and low transverse momentum muons are identified using only the muon trigger chambers.

The calorimeter selections are based on a reduced-granularity information from all the calorimeters. In case of the electron/photon and hadron/τ triggers,

energy isolation cuts can be applied. The missing and total scalar transverse energies used in the LVL1 trigger are calculated by summing over trigger towers. In addition, a trigger on the scalar sum of jet transverse energies is also available. No tracking information is used at LVL1 due to timing restrictions and the inherent complex nature of the information from the inner detector.

The maximum rate at which the ATLAS front-end systems can accept LVL1 triggers is limited to 100 kHz. An essential requirement on the LVL1 trigger is that it should uniquely identify physics events of interest. Given the short (25 ns) bunch-crossing interval, this is a non-trivial task. In case of the muon trigger, the physical size of the muon spectrometer is comparable to the bunch spacing. For the calorimeter trigger, a serious challenge is, that the pulse shape of calorimeter signals extends over many bunch crossings. During this time, information for all detector channels is stored in pipeline memories. The LVL1 latency, measured from the time of the *pp* collision until the trigger decision is available to the front-end electronics, is required to be less than 2.5 μs. In order to achieve this, the LVL1 trigger is implemented as a system of purpose-built hardware processors.

Results from the L1 muon and calorimeter triggers are processed by the Central Trigger Processor, which implements a trigger 'menu' made up of combinations of trigger elements. Prescaling of the trigger menu items is also available, allowing optimal use of the rate bandwidth as luminosity and background conditions change.

Another important functionality of the LVL1 system is the identification of the Regions Of Interest (ROIs) representing the position of the triggering objects in the (η,ϕ) space. This is one of the main peculiarities of ATLAS triggering system. This information is used to greatly reduce the needed computation time at the LVL2 triggering system and the size of data to be transferred in the system.

The LVL2 runs offline-like algorithms, optimised for the on-line use, using the full granularity information from the inner detector as well as from the muon detectors and calorimetry. However, it is structured to process the data belonging only to a spatial window around the ROIs identified by the LVL1 trigger. Simulations showed that this corresponds to roughly 2-5% of the overall ATLAS data size. LVL2 has a maximum latency time of 10 ms, after this time the event is selected (and hence moved to the Event Filter system for further processing) or discarded and removed from the Data Flow chain (up to this moment the event fragments have been buffered in the ROBs. The final LVL2 rate is expected to be about 1-2 KHz.

After LVL2, the last stage of the on-line selection is performed by the Event Filter. It employs offline algorithms and methods, slightly adapted to the on-line environment, and uses the most up to date available calibration and alignment information and the magnetic field map, thus making a complete event reconstruction. Events, that have passed the Event Filter, are written to mass storage for a subsequent off-line analysis.

The time available for a decision at the event filter is 1 s. The output rate from LVL2 should be reduced by an order of magnitude, giving about 200 Hz. The final event size is expected to be 1.5 MBytes corresponding to an output

data rate of about 300 MByte/s resulting in 10^{15} bytes of data per year.

Detector Control System

In order to enable coherent and safe operation of the ATLAS detector, a Detector Control System (DCS) has been defined and implemented. It has the task to set up the detector hardware in a selected state and to continuously monitor its operation. The DCS consists of two parts: the Front-end systems (FE) and the Back-end (BE) control station.

The FE connects to the detector hardware and the equipment to supervise ranges from simple sensors to complex devices like software controllable power supplies. A small set of commercial devices has been selected as a standard, such as crates and power supplies. A general purpose I/O concentrator has been developed, called Embedded Local Monitor Board (ELMB). It comprises a multiplexed ADC (64 channels with 16 bit resolution), 24 digital I/O lines and a serial bus to drive external devices. The ADC part of the ELMB can be configured for various types of sensors.

The BE is organised in three layers: the Global Control Station (GCS) with human interfaces in the ATLAS control room for overall operations, the Subdetector Control Stations (SCS) for stand-alone operation of a subdetector, and the Local Control Stations (LCS) for process control of subsystems.

4.3.6 Computing

The complexity and size of the ATLAS experiment imposes the use of new paradigms also in what the processing of the data is concerned once they are made available on mass storage. The events rate of 200 Hz, the size of the events (approximately 1.5 MB per event), the number of physicists involved in the analysis requires that the data distribution, processing and analysis is carried out according to a multi-tier schema, that is well suited to distribute the computing and storage loads among the different participating institutes. Even if similar strategies have been used in the past it is the first time that this kind of distributed analysis are performed on a ATLAS-size scale requiring the development of completely new tools extremely performing, [42].

At the output of the Event Filter the raw data are transferred to the CERN's computing centre, known as Tier-0, that is the first layer of the ATLAS analysis system. Here a complete copy of the raw data is stored and a first-pass reconstruction is applied producing ESD (Event Summary Data) and AOD (Analysis Object Data). The ESD data-format contains the reconstructed quantities measured by the detector (energy in the calorimetric cells, clusters information, tracks) as well as the reconstructed physics objects (electrons and gammas, jets, taus, muons). The small-sized data in AOD format are well suited for distribution to the physicists groups, to reduce their size only the physics objects are recorded. Each event can be characterised by few quantities, like, for example, the number of jets in the event, p_{T} of the leading jet, lepton multiplicity and so on, this information, produced by Tier-0 and stored in the TAGs and allows for a very fast filtering of the datasets. Tier-0 has also the responsibility to run

calibration and alignment algorithms that will be refined in future steps.

Distribution of data to the community is done copying raw data, ESDs, AODs and TAGs to the Tier-1s. Tier-1s are big regional computer centres spread around the world (at the moment ATLAS foresees 10 of these centres). A copy of the raw data is divided among all the Tier-1s (each one having on average 10% of the entire raw data) while a complete copy of the ESDs, AODs and TAGs is distributed to each Tier-1. Tier-1s have also the responsibility to reprocess raw data performing more accurate reconstructions and produce updated version of ESDs, AODs and TAGs spread them among the different computer centres.

Most of the physics analyses are performed at the Tier-2s centres, in average 5 Tier-2s are directly connected to one Tier-1 and typically receive a copy of one third of the most updated ESD and AOD data and a complete copy of the TAGs. Tier-2s have the entire responsibility for the official Monte Carlo production (the simulated data is stored in the more reliable Tier-1s). The physics groups analysis and the development and refinement of calibration and reconstruction algorithms are also performed at the Tier-2 centres. The physics analysis will be performed mainly on the AOD data set (with the help of TAGs for pre-selections) or on even more compact derived formats like ntuples.

The multi-tier paradigm is deployed using *grid* technology and middle-ware that completely hides to the physicists the complex multi-tier structure, [43].

Chapter 5

Jet reconstruction

Asking for a certain number of partons being present in the final state is not meaningful in pQCD. Particles are therefore combined into groups, called jets. Jets can be defined at parton level, particle level (from final states), or detector level (from detector towers). Jets are the observable final states also for our study of quark compositeness.

In this chapter we will start with a short description of jet clustering algorithms, then briefly summarise current ATLAS procedure of reconstructing and calibrating jets, describe basic methods of in-situ validation of jet calibration, and finally mention expected values of jet energy scale uncertainty in early data.

5.1 Jet clustering algorithms

The two default jet-clustering algorithms in ATLAS are a seeded fixed cone algorithm and a successive recombination ($k_\perp$) algorithm. Both algorithms are used to produce narrow jets (for events containing large multiplicities of jets etc.) and wider jets (for studies of dijet final states or multi-jet final states at low luminosities).

The seeded cone algorithm uses two parameters, the transverse energy threshold for a seed, $E_{\mathrm{T}} = 1$ GeV for all cone jets, and the cone size, $R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$, with $R = 0.4$ for narrow jets and $R = 0.7$ for wide jets. In all cases, a split-and-merge step follows the actual cone building, with an overlap fraction threshold of 50%. The cone algorithm is sensitive to soft gluon emission and embodies collinear sensitivity [60]. However, this sensitivity is greatly reduced for larger p_{T} jets. When jet reconstruction becomes 100% efficient, the collinear sensitivity is reduced to a second order effect. As will be described in Sec. 6.1, in this study we use only jets with $p_{\mathrm{T}} > 350$ GeV, for which the reconstruction is almost 100% efficient.

The $k_\perp$ algorithm in ATLAS is implemented following the suggestions in [61], which makes it efficient even for a rather large number of input objects and avoids the usual pre-clustering step. The distance parameter $D = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ is adjusted for narrow jets to $D = 0.4$ and for wide jets to $D = 0.6$. The physics performance is very similar to the one of the corresponding cone configurations. The $k_\perp$ algorithm is infrared and collinear safe to all orders of calculation.

In all cases the full four-momentum recombination is used to calculate the jet kinematics after each clustering step.

5.2 Input to jet reconstruction

The inputs for calorimeter or reconstructed jet-finding are calorimeter signals. Fig. 5.1 presents an overview of the reconstruction flow for calorimeter jets. Calorimeter jets are reconstructed by applying a jet-clustering algorithm to calorimeter signals, typically followed by a calibration step. Two different signals from the calorimeter are used for jet-finding, towers and topological cell clusters (topoclusters). Towers are formed by collecting cells into bins of a rectangular $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ grid. All calorimeter cells are used in the towers. Towers with negative signals dominated by noise are recombined with nearby positive signal towers until the net signal is positive.

Topoclusters represent an attempt to reconstruct three-dimensional energy deposition in the calorimeter [62, 63]. First, nearest neighbours are collected around seed cells with a significant absolute signal above the major seed threshold, such as $|E_{cell}| > 4\sigma_{cell}$ of the total noise (electronics plus pile-up, that is incoherent and coherent noise). If the absolute value of signal of these neighbouring cells is above a secondary seed threshold, typically $|E_{cell}| > 2\sigma_{cell}$, they are considered secondary seeds, and their direct neighbours are also collected. Finally, all surrounding cells above a very low threshold (typically set to $0\sigma_{cell}$ are added if no more secondary seeds are among the direct neighbours.

Contrary to the signal tower formation, topological cell clustering includes noise suppression, meaning that cells with no signal at all are most likely not included in the cluster. This results in substantially less noise. However, the topological cell clustering will require careful validation with real data, in particular in terms of the possible impact of long-range noise correlations and of detailed studies of pile-up effects as the luminosity increases.

5.3 Jet calibration

The following was adopted from [64]. The ATLAS calorimeter jet calibration is based on the application of cell signal weighting similar to the original approach developed for the H1 calorimeter [65]: all calorimeter cells with four-momenta $(E_i, \overrightarrow{p}_i)$, where $E_i = |\overrightarrow{p}_i|$, in tower or cluster jets are considered and re-summed with weighting functions, w, such that the resulting new jet four-momentum is:

$$(E_{rec}, \overrightarrow{p}_{rec}) = \left(\sum_i^{N_{cells}} w(\rho_i, \overrightarrow{X}_i) E_i, \sum_i^{N_{cells}} w(\rho_i, \overrightarrow{X}_i) \overrightarrow{p}_i, \right). \tag{5.1}$$

The weighting functions w depend on the cell signal density, $\rho_i = E_i/V_i$, and on the cell location in the calorimeter, $\overrightarrow{X}_i$, consisting basically of module and compartment identifiers. They are fitted using simulated QCD di-jet events, covering the whole kinematic range expected at the LHC, and matching calorimeter cone-tower jets, with $R = 0.7$, with nearby truth-particle cone jets

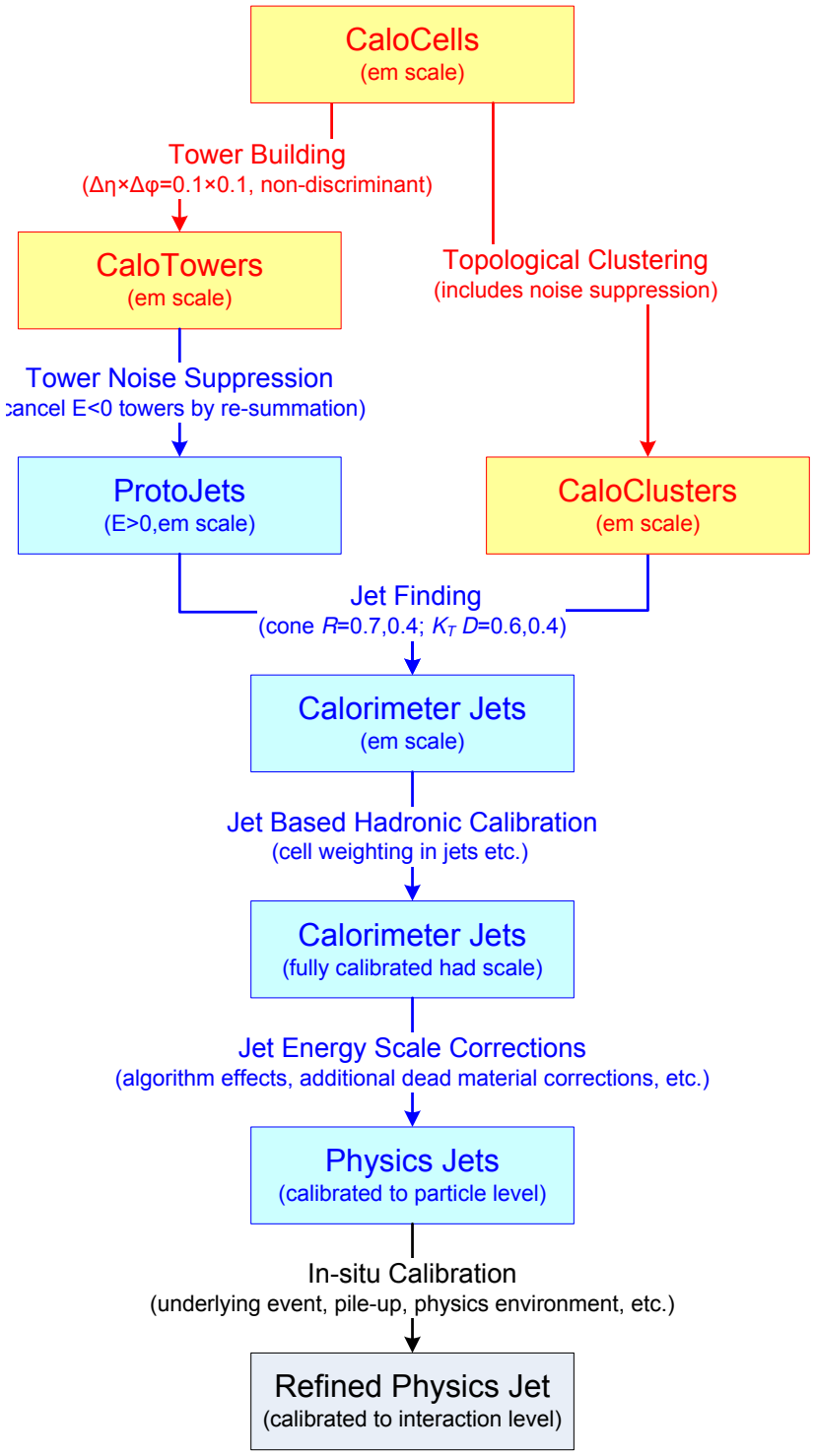


Figure 5.1: Jet reconstruction flow for calorimeter jets from towers or clusters.

(jets builded from all stable particles before interacting with the detector or its magnetic field) of the same size and with energy E_{truth}, and then constraining E_{rec} in Eq. (5.1) to E_{truth} by:

$$\frac{\partial \chi^2}{\partial w(\rho_i, \overrightarrow{X}_i)} = \frac{\partial}{\partial w(\rho_i, \overrightarrow{X}_i)} \left[\sum_{\text{matched jets}} \frac{((E_{rec} + E_{DM}) - E_{truth})^2}{E_{truth}} \right] = 0. \quad (5.2)$$

The weighting functions determined in this way absorb all detector effects. Implicitly included also are corrections for energy loss in inactive materials, except for losses between the electromagnetic barrel and tile barrel calorimeters, which are parametrised in Eq. (5.2) as

$$E_{DM} = \alpha \sqrt{E_{EMB3} E_{TILE0}}, \quad (5.3)$$

where E_{EMB3} is the sum of the energies of the cells in the last compartment of the barrel EM calorimeter belonging to the jet and E_{TILE0} is the corresponding sum in the first compartment of the hadronic tile calorimeter. Both quantities are reconstructed at the electromagnetic energy scale. The parameter α was assumed to be independent of energy and η and was determined together with $w(\rho_i, \overrightarrow{X}_i)$ in a combined fit according to Eq. (5.2).

The calibration applied in this way only corrects to the level of the truth-particle jet. The extracted weighting functions were obtained for cone-tower jets with $R = 0.7$ and are not universal, since they depend on the choice of calorimeter signals used, on the jet algorithm chosen and on its specific configuration, and on the choice of (simulated) physics calibration samples used to extract them. Residual mis-calibrations of all other jet-clustering algorithms and their configurations are corrected for by functions depending on $|\eta|$ and p_{T} of each measured jet. Whole this jet calibration procedure is called global calibration.

5.3.1 Jet energy linearity and resolution

The following signal features are extracted from simulations including a model for the electronic noise in each calorimeter cell, tuned with parameters extracted from various beam test measurements. Pile-up fluctuations are not included.

The signal linearity for calorimeter jets shown in Fig. 5.2 is expressed by the ratio of the reconstructed jet energy and the matched truth-jet energy, E_{rec}/E_{truth}, in simulated QCD dijet events. The signal linearity of global calibrated jets is reasonable over the whole energy range. It is also shown, that the energy of jets at the electromagnetic energy scale (that is without any hadronic calibration applied) corresponds to only $\sim 70\%$ to $\sim 80\%$ of the true jet energy.

The fractional energy resolution for the same jets, again after global calibration is shown for two different η regions in Fig. 5.3. The energy resolution was fitted by a formula

$$\frac{\sigma}{E} = \sqrt{\frac{a^2}{E} + \frac{b^2}{E^2} + c^2}. \quad (5.4)$$

In the central region $0.2 < |\eta| < 0.4$, the stochastic term is $\approx 60\%\sqrt{\text{GeV}}$, while the high-energy limit of the resolution $c \approx 3\%$. The noise term b increases from 0.5 GeV to 1.5 GeV when going from the barrel to the end-cap η ranges.

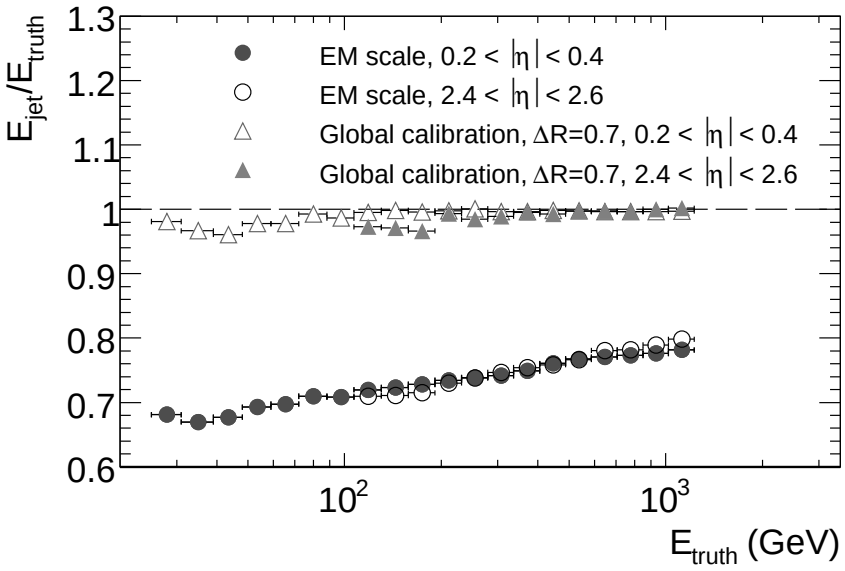


Figure 5.2: Signal linearity for cone-tower jets with $R = 0.7$, as expressed by the ratio of reconstructed tower jet energy to the matching truth-jet energy E_{rec}/E_{truth}.

5.3.2 Jet signal uniformity

The variation of the jet energy response with jet direction depicted in Fig. 5.4 is a measure of the jet signal uniformity across the full rapidity coverage of the calorimeters. The dips in response for $1.2 < |\eta| < 2.0$ and $2.8 < |\eta| < 3.4$ correspond to the two transition regions with higher abundance of dead material, while the decrease of response in the last bin $|\eta| > 4.4$ is caused by the limited fiducial coverage of the jet reconstruction. The dips are much more apparent at lower transverse energies.

5.4 In-situ validation of jet calibration

Several final states at the LHC provide signals for validation of the jet energy calibration, and even the extraction of further corrections. The final states with a well measured electromagnetic object balancing one or more jets in transverse momentum are suitable for this task. The $\gamma+\text{jet(s)}$ process provides high statistics in the p_{T} range from 40 to 400 GeV, but lower purity than the $\text{Z} + \text{jet(s)}$ process covering precisely lower p_{T} region up to 100 GeV.

One example of measuring the jet response using $\gamma + \text{jet(s)}$ is the missing transverse momentum projection fraction. The basic idea of this method is to

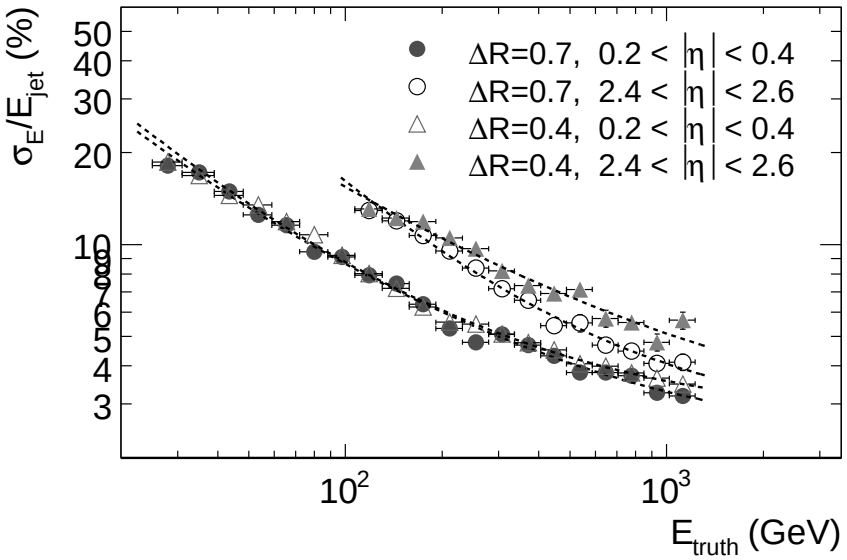


Figure 5.3: Fractional energy resolution for calibrated cone-tower jets reconstructed with $R = 0.7$ and $R = 0.4$.

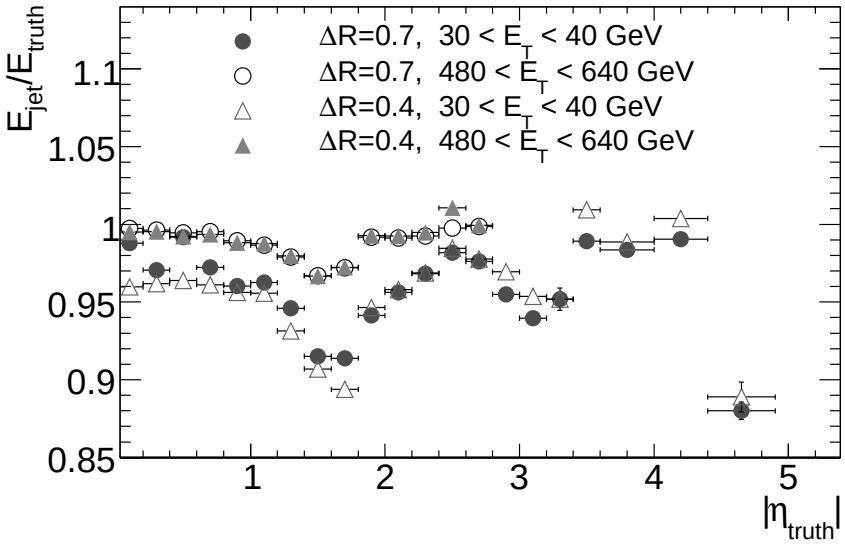


Figure 5.4: Signal uniformity for QCD dijets in two E_{T} ranges, as a function of η of the matched truth-particle jet.

project the hadronic p_{T} vectors onto the p_{T} vector of the photon and to measure the apparent $E_{\mathrm{T}}^{\mathrm{miss}}$ fraction. In events where the photon is back-to-back with the jet, the jet response R_{jet} can then be determined by

$$R_{jet} = -\frac{\sum_{\mathrm{signals}} \overrightarrow{p}_{\mathrm{T,had}} \cdot \hat{\eta}_{\gamma}}{p_{\mathrm{T},\gamma}}, \tag{5.5}$$

where $\hat{\eta}_{\gamma} = \overrightarrow{p}_{\mathrm{T}\gamma}/p_{\mathrm{T},\gamma}$ is the direction of the photon in the transverse plane. The hadronic transverse momentum can be calculated using reconstructed jets, or just using the sum of cluster signals. R_{jet} can be directly measured and used for a global jet energy scale calibration derived from collision data.

The other important final state for jet calibration are hadronically decaying W bosons in $t\bar{t}$ production. Here, m_W constraints the energy scale of the two quark jets. With further in-situ corrections aimed at re-scaling je energies as a function of $|\eta|$ to obtain a uniform response, a linearity of better than 2% can be achieved up to values of the true transverse momentum of the W-boson as high as 200 GeV.

5.5 Validation of high-p_{T} jet calibration

The upper p_{T} limit of the above discussed in-situ methods to validate the global calibration is about 400 GeV, where these methods are expected to fail due to insufficient statistics. In order to validate even higher p_{T} regions, the multi-jet bootstrap method [66] was developed. This method utilises already validated p_{T} region as its base and iteratively validates the calibration in higher p_{T} regions, using QCD jets only, where large statistics is available up to several TeV region.

In the bootstrap method, events with at least three jets are selected, with one jet having significantly higher p_{T} than all others. This jet is balanced in p_{T} with the vector sum of the remaining jets and used to set the jet energy scale in higher p_{T} region. This process can be iterated. The bootstrap method also includes corrections for effects spoiling the vector momentum sum balance, such as very soft radiation favouring one side of the event and migration effects caused by limited jet energy resolution. With the bootstrap method, the jet energy scale set in the base p_{T} region (up to 400 GeV) can be bootstrapped up to 1 TeV with precision of $\pm 1\%$ with 1 fb^{-1}of collected data and the limits of this method with 300 fb^{-1}are expected to be around 3 TeV [67].

Jet energy scale in the p_{T} region above 3 TeV cannot be validated directly with experimental data and we must rely on our extrapolation. The systematic error of this extrapolation above 4 TeV due to uncertainty in our knowledge of e/h of the ATLAS non-compensating calorimeters might be as small as 1 – 2% [68]. Nevertheless at such very high p_{T} region other sources of systematic errors will arise, for example due to a punch-through of hadronic showers and leak of the jet energy into the Muon Spectrometer. These uncertainties still need to be studied and will be specified also using the real data.

5.6 Early data JES systematic errors

Due to the composite character of jets and a corresponding lack of test beam data it is impossible to specify precisely the jet energy scale uncertainty prior to having a collision data of sufficient quality. The values of JES systematic errors recommended by the ATLAS community for the early data analysis (0.1 – 1 fb^{-1}) are listed in Tab. 5.1, [69, 70]. It is very likely that with increasing amount of collected collision data these uncertainties will decrease.

	JES systematic error
$\lvert\eta\rvert < 3.2$	7%
$3.2 < \lvert\eta\rvert < 4.9$	15%

Table 5.1: Recommended values of JES uncertainty to be taken into account in early data analysis.

Chapter 6

Simulated Monte Carlo data and analysis strategy

In this chapter let us describe the data samples analysed in following chapters. The analysis strategy used for these samples is also briefly introduced.

6.1 Simulated data sets

As the startup of the LHC machine is expected in 2008, no physics data from the ATLAS detector to study quark compositeness is available yet. Nevertheless, the aim of this work is to be prepared for the real data and to find an optimal analysis approach to search for quark compositeness. For this purpose, several kinds of large Monte Carlo data samples were simulated.

The private compilation of Pythia 6.323 program was used to generate the particle collisions. The only difference with respect to the official Pythia 6.323 code is, that in our version the ISR error affecting compositeness generation was corrected, as described in Sec. 3.2.1. Our private code is, in terms of compositeness, equivalent to that of official Pythia 6.404, which was not available at the moment of generation of the data samples.

The full simulation of the ATLAS detector response was done in ATLAS dedicated framework Athena 11.0.42, [48]. The full GEANT4 [49, 50] simulation of the ATLAS response is extremely CPU time consuming, therefore it is available only for limited number of events and only for QCD dijets. The most of the statistics was obtained in fast simulation of the ATLAS detector response ATLFAST [51, 52], in Athena release 11.0.41 (which is, in case of ATLFAST, equivalent to 11.0.42). A detailed overview of settings of Pythia parameters used for ATLFAST simulation is presented in Appendix A.

Only jets with $p_{\mathrm{T}} > 350$ GeV were used. The value of this threshold is based on the ATLAS inclusive physics prescaled trigger *2j350*, passing events with at least 2 jets with $p_{\mathrm{T}} > 350$ GeV. The reconstruction efficiency for such rather high-p_{T} jets is taken to be 100%.

6.1.1 Full simulation data

The full simulation of the ATLAS detector response is considered as more realistic than the fast simulation (see Sec. 6.1.2), unfortunately it is also extremely CPU time consuming. In order to use its computing resources effectively, the ATLAS community decided to manage the production of full simulation data centrally. One of the latest results of this effort are so-called CSC (Computing System Commissioning) data samples.

There are two sets of CSC dijet samples used in this study: PYTHIA generated dijet samples used for a comparison with fast simulation and HERWIG [72] generated (with JIMMY plug-in for multiparton interactions [73]) dijet samples used for an estimate of a systematic error due to the choice of Monte Carlo generator. A detailed overview of used settings of Monte Carlo generator parameters is presented in Appendix A.

The PYTHIA generated CSC data consists of several QCD inclusive dijet samples with various values of lower and upper cuts on p_{T} of the jets. Physical spectra from these samples are weighted according to their total cross-sections and then summed together (or sewed, in one word) in order to reconstruct the full inclusive dijet QCD spectrum. An overview of these samples is available in Tab. 6.1.

sample	p_{T} cuts (GeV)	σ (pb)	Number of events
J4	140 to 280	3.08×10^5	167947
J5	280 to 560	1.25×10^4	277491
J6	560 to 1120	360	246195
J7	1120 to 2240	5.71	246345
J8	$>$ 2240	0.24	259588

Table 6.1: Selected CSC Pythia dijet samples. Numbers of events correspond to all AOD events fully simulated and reconstructed in Athena 11.0.42.

The JIMMY/HERWIG generated CSC data consists of several QCD inclusive dijet samples covering a kinematic region of J4 – J7 samples and denoted similarly as J4 Jimmy – J7 Jimmy. The amount of fully reconstructed J4 Jimmy – J7 Jimmy events is only about 20% of PYTHIA generated J4 – J7 events.

The CSC dijet samples were simulated using the CTEQ6L1 pdfs (LO fit with LO α_s), in agreement with ATLAS recommendation to use LO pdfs with LO generator (such as PYTHIA). The jet reconstruction procedure used for these jets is described in chapter 5.

To be complete, the jets were retrieved from Analysis Object Data (AOD, [54]) containers. The Cone jet algorithm with a cone size of $\Delta R = 0.7$ was used.

Full simulation data for compositeness switched on is not available.

6.1.2 Fast simulation

The following was adopted from [52]. Fast simulation of the ATLAS detector response is performed by ATLFAST. No detailed simulations of any interactions

of the particles in the detector media are performed. These interactions are implicitly taken into account by using a parametrised detector response. The following steps are carried out:

- Stable particles from the event generators are tracked through the magnetic field by using a helix model as an explicit track model. In order to calculate the particle tracks, no multiple scattering, energy loss or nuclear interactions are taken into account. These effects are, however, implicitly taken into account by the application of appropriate resolution functions, when calculating the response of the calorimeters.

- The energies of the electrons, photons and hadrons are deposited in a calorimeter cell map. The response of the calorimeter is assumed to be the same for EM and hadronic showers (that means $e/h = 1$) and uniform over the full calorimeter region. The energy of the particle is entirely deposited in the hit calorimeter cells (there is no punch-through), assuming a granularity of the calorimeter cell map of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ up to $|\eta| < 3.2$ and $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ for $3.2 < |\eta| < 4.9$. Neither lateral nor longitudinal shower development is simulated.

- The reconstruction of the physics objects in ATLFAST relies to a large extent on the Monte Carlo truth information. No reconstruction layer based on the simulated detector information is implemented.

The steps carried out to classify physics objects are the following:

1. **Clusters.** Cluster reconstruction is carried out on the map of energies deposited in cells. A cone algorithm with a cone size of $\Delta R = 0.7$ is used. The cone is initiated by the seed cells with energies above 1.5 GeV. The cluster energy must pass a threshold of 5 GeV. The clusters may get reclassified as electrons, photons, taus or jets in one of the following steps. If they are associated with one of these objects, they get removed from the list of clusters.

2. **Electrons.** For each true electron a calorimeter cluster is searched that could be matched to the electron in (η, ϕ) space. The separation is chosen to be less than $\Delta R = 0.15$. Other isolation criteria are then used and the reconstructed electron energy is smeared.

3. **Photons** In terms of isolation criteria, photons are treated in the same way as electrons. In addition a smearing of the photon direction in η is applied.

4. **Muons.** For each true muon with $p_{\mathrm{T}} > 0.5$ GeV[1] the reconstructed momentum is calculated from the true muon momentum, using Gaussian smearing.

[1] In this study $c = 1$ and the unit of particle momentum is (eV).

5. **Jets.** All clusters that have not been assigned to a true electron, photon or a muon are considered as jets, if their transverse energy exceeds a threshold of 10 GeV. The jet energy is smeared according to the following resolution functions:

   for $|\eta| \leq 3.2$

   $$\frac{\sigma}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\%, \tag{6.1}$$

   for $3.2 < |\eta| < 4.9$

   $$\frac{\sigma}{E} = \frac{100\%}{\sqrt{E}} \oplus 7\%. \tag{6.2}$$

   This corresponds to the calorimeter design criteria as described in Sec. 4.2. The jet direction is taken to be the cluster direction. Since the response function of the calorimeter is set to one ($e/h = 1$), no jet calibration is needed to correct for the lower and non-uniform calorimeter response to hadrons. However, an out-of-cone energy correction is needed.

6. Next steps include identification of taus, b and c jets, missing transverse momentum and tracks.

The labelling of the jets is done using the MC truth, possible labels are b, c, τ and u for jets originated by b quark, c quark, τ lepton and other particles, respectively. The ATLFAST jets may be also calibrated in ATLFASTB algorithm:

- Based on efficiency functions of tagging that have been evaluated from full simulation, it gives a reconstruction label to light jets and to b, c and τ jets.
- It applies calibration constants to jets depending on the jet p_{T} and η and reconstruction label. The calibration constants take into account mainly the out of cone energy. The applied calibration factor is $K_{jet} = p_{\mathrm{T}}^{\mathrm{parton}}/p_{\mathrm{T}}^{\mathrm{jet}}$, where $p_{\mathrm{T}}^{\mathrm{parton}}$ denotes the transverse momentum of the parton that initiated the jet.

Concerning our data, no ATLFASTB corrections were applied to ATLFAST jets. The treatment of ATLFAST jets in release 11.0.41 is the one obtained in [53].

The ATLFAST data was also generated with CTEQ6L1 pdfs. It consists of samples with compositeness switched off (SM only) and switched on (with values of Λ = 3, 5, 10, 15 and 20 TeV). There were 450 millions of ATLFAST events simulated. To be complete, the jets were retrieved from Analysis Object Data (AOD, [54]) containers (which are created by copying information from internal ATLFAST objects).

6.1.2.1 Comparison with CSC samples

In order to benefit from the full simulation/reconstruction CSC samples and to transfer the results of full simulation of ATLAS detector response also to the ATLFAST data, a comparison of CSC samples with corresponding ATLFAST data was done. In this case, the same generated events are used both for full simulation/reconstruction and for ATLFAST. In both cases with jet cone algorithm parameter $\Delta R = 0.7$.

We should point out here some intrinsic differences between these two kinds of data. While ATLFAST builds clusters (initiated from cluster seeds ordered in p_{T}) without applying a split-and-merge procedure, the full reconstruction uses a standard seed cone algorithm [55]. This can lead to differences in the reconstructed final state. Moreover, the largest effects to be corrected for in the full reconstruction are detector effects, mainly the non-linearities coming from the non-compensating ATLAS calorimeter system. Such effects are not present in ATLFAST.

6.1.3 pdf uncertainty studies

The pdf experimental uncertainty studies were done at a parton level using two pdf error sets of CTEQ6M1 and MRST2001E (both with NLO fit and NLO α_s, as there are no LO pdf error sets.) and Pythia. The simulated data consists of samples with compositeness switched off (SM only) and switched on (with values of Λ = 3, 5, 10, 15 and 20 TeV). There were 350 millions of events simulated.

In case of dijet p_{T} spectra only, the pdf uncertainties were also studied with NLOJET++ generator [56, 57] and CTEQ6M1, in order to verify, that the order of the matrix elements used in the generator (LO or NLO) does not affect significantly the value of obtained relative pdf errors. This was verified to an accuracy of one per mille.

6.1.4 Next-to-leading order approach

As already said, the full simulation/reconstruction and the fast simulation dijet samples were generated with CTEQ6L1 pdfs (LO fit, LO α_s) and Pythia (LO generator). In order to be consistent with the pdf uncertainty results (NLO) and to benefit from NLO dijet cross-section calculations the following two factors were applied to this data:

- **k-factors:** a ratio of NLO and LO cross-sections for a given observable (SM spectra only), as calculated in NLOJET++ at parton level (hard process only), and
- a ratio of cross-sections as calculated in Pythia with NLO and LO pdfs, here referred to as **ll2nn** factors ("LO fit + LO α_s pdf to NLO fit + NLO α_s pdf"). There are two sets of these factors: one for the step from CTEQ6L1 to CTEQ6M1 and the other for the step from CTEQ6L1 to MRST2001E. The ll2nn factors were calculated with compositeness

switched off (SM only) and switched on (with values of Λ = 3, 5, 10, 15 and 20 TeV).

Both of these kinds of factors are functions of the variable they are used for, e.g. k-factor used for inclusive dijet p_{T} spectrum is a function of p_{T}.

6.2 Analysis strategy

In order to find an analysis method the least sensitive to systematic errors and the most sensitive to quark compositeness, four different analysis approaches were studied:

- p_{T}**-Bayes** method. In Sec. 7.3 the Bayesian technique was applied to inclusive dijet p_{T}-spectrum. This method uses a detailed information about the p_{T}-spectrum shape.
- **Rdist3** "ratio" method is based on a comparison of a numbers of events in p_{T} spectra below and above some defined p_{T} cut. Hopes laid on this method – a smaller sensitivity to systematic errors – were not fully met, as shown in Sec. 7.4.
- **R1** "ratio" method is based on a comparison of a numbers of events in dijet angular distribution spectra (or χ spectra) below and above some defined χ cut. This is studied in Sec. 8.3.
- χ**-Bayes** method applies Bayesian technique to dijet angular distribution spectra in Sec. 8.4.

The results from all four methods are compared in Sec. 9.

The following systematic errors were taken into account for each of these methods:

- **pdf uncertainties**. The uncertainties of parton distribution functions caused by the experimental uncertainties of data used for the pdf global fits. CTEQ6M1 and MRST2001E pdf error sets were used.
- **renormalisation and factorisation scale uncertainties**. Used k-factors depend on non-physical scales μ_R and μ_f. The uncertainty caused by this dependence must be somehow quantified, if we neglected it, we would obtain higher exclusion limits for compositeness, stating, that there is not uncertainty at fixed-order perturbation calculation, which is not true. The $k = k(\mu_r, \mu_F)$ dependence could be a subject to a further study, nevertheless for the purpose of this paper a standard approach was used. The k-factors for each analysis method were evaluated at following scales: $\mu_R = \mu_F$ = 0.5, 1.0 and 2.0 p_{T}^{max}, where p_{T}^{max} is the highest jet p_{T} in a given event.

- **Jet energy scale (JES) uncertainty.** At the point of real physics data analysis, the JES uncertainty will depend on both p_{T} and η, it might be as good as a few percent up to $p_{\mathrm{T}} = 3$ TeV and might be much worse above $p_{\mathrm{T}} = 3$ TeV, depending on accumulated integral luminosity. For the purpose of this paper an "all-inclusive" JES uncertainty of 3% for all values of p_{T} and η was used, in order to make a comparison of all analysis methods more simple. The JES uncertainty of 3% might be reached with larger amount of sufficient quality collision data. In case of early data (0.1 – 1.0 fb^{-1}) larger JES uncertainty is foreseen (see Tab. 5.1), a prediction of exclusion limits on quark compositeness involving this higher JES uncertainty is presented in Sec. 9.2.

- **Absolute luminosity measurement uncertainty** of 3% (see Sec. 4.3.4 for details).

- **Uncertainty of the comparison with full simulation.** This uncertainty rises from limited statistics of full simulation data samples and from systematic differences between the full and fast reconstruction. It is different for each analysis method.

In addition, two other systematic errors were studied in special cases. Firstly, in case of the most sensitive Bayesian method p_{T}-Bayes, a dependence on prior probabilities was studied. Secondly, in case of dijet inclusive p_{T} spectrum, a dependence on used Monte Carlo generator was studied with Pythia and Jimmy/Herwig generated CSC dijet samples.

Yet another study concerning systematic errors would be interesting: the dependence of systematic errors on the type and parameters of used jet algorithm. For a certain jet algorithm the dependence of cross-sections on renormalisation and factorisation scales might be minimal or the jet energy scale uncertainty might be minimal. Nevertheless, such a study was not performed in this paper.

Chapter 7

Inclusive dijet p_{T}-spectrum

In this Chapter the effect of quark compositeness on inclusive dijet p_{T}-spectrum is studied with two methods, p_{T}-Bayes and Rdist3 (in Sections 7.3 and 7.4). Let us start with the description of the p_{T}-spectra and related systematic errors.

In order to cover almost ten orders of magnitude range of differential cross-section $\mathrm{d}\sigma/\mathrm{d}p_{\mathrm{T}}$, the data was sewed from 5 p_{T}-slices with following cuts[1] on p_{T}: 200, 350, 600, 1200, and 2000 GeV as shown in Fig. 7.1. This procedure was done for SM and for all values of Λ too.

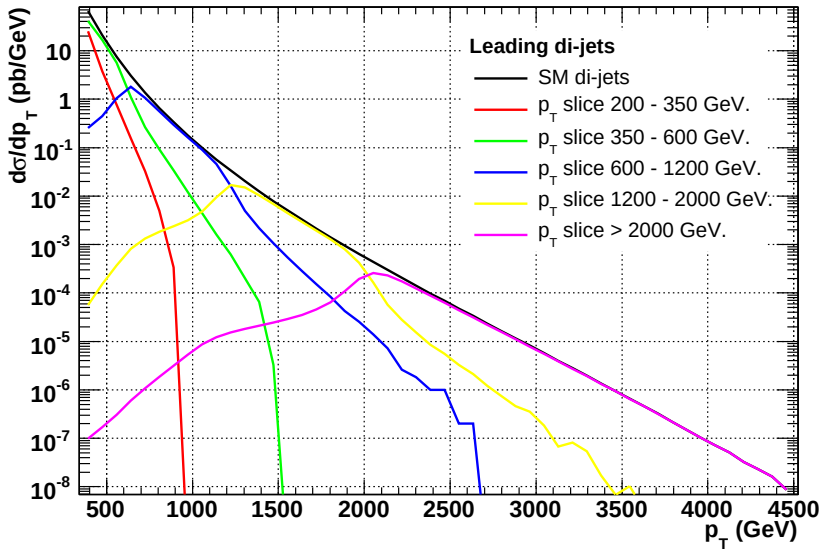


Figure 7.1: Summing QCD inclusive leading dijet p_{T}-spectra from p_{T}-slices. ATLFAST.

Inclusive leading dijet p_{T}-spectra for various values of compositeness scale Λ are shown in Fig. 7.2, the spectra consists of 50 p_{T}-bins in a range from 350 to 4500 GeV. Note, that for every value of Λ there always exist a p_{T}-range or bin where the expected number of events is lower than that of pure QCD spectrum.

[1] PYTHIA CKIN(3) and CKIN(4) cuts.

This is due to the interference terms $\mathcal{O}(\Lambda^{-2})$ in Eqs. (3.8) and it is visible in Fig. 7.8. The higher the Λ value, the higher in p_{T} these undershoot bins are.

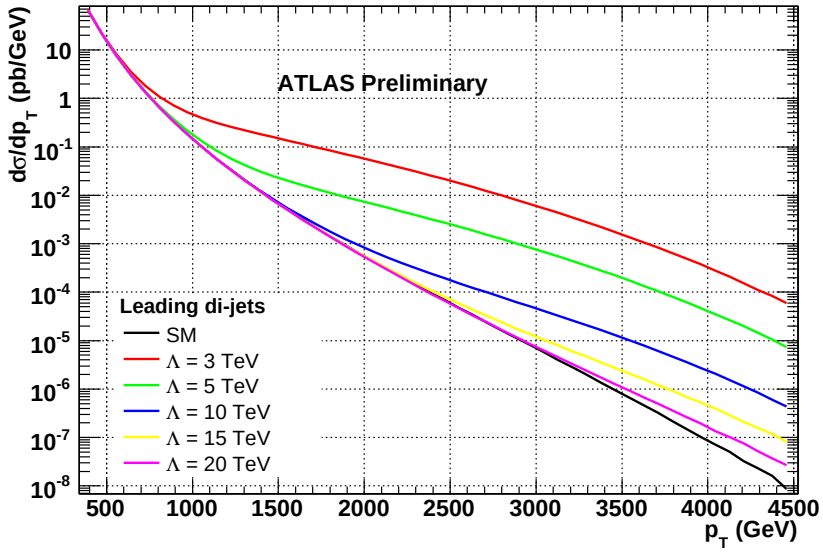


Figure 7.2: Inclusive leading dijet p_{T}-spectra for Λ = 3, 5, 10, 15, 20 TeV and QCD only as obtained with ATLFAST, PYTHIA and CTEQ6L1.

7.1 Λ-fit

The simulated data samples include five different values of Λ, but for our studies we need to know the p_{T}-spectra for a whole continuum of Λ values and it is possible to obtain this p_{T}-spectra Λ dependence with the following procedure.

The differential cross-section $\mathrm{d}\sigma/\mathrm{d}p_{\mathrm{T}}$ in k^{th} p_{T}-bin is proportional to the convolution of pdf contribution and the matrix element:

$$\sigma_k \propto f(x_1, Q^2, flavour_1) f(x_2, Q^2, flavour_2) \otimes \hat{\sigma}, \tag{7.1}$$

where the factorisation scale is set to Q^2. Within the given p_{T}-bin the value of Q is well limited and the dependence of pdfs on x_1, x_2 and flavours of incident partons is very small. So the major dependence of $\mathrm{d}\sigma/\mathrm{d}p_{\mathrm{T}}$ in k^{th} on Λ follows from Eqs. (3.8) and the cross-sections in p_{T}-bins for simulated values of Λ can be fitted by the following formula:

$$\sigma_k(\Lambda) = p_1^k + \frac{p_2^k}{\Lambda^2} + \frac{p_3^k}{\Lambda^4}, \tag{7.2}$$

where p_i^k are the fit parameters for the given bin. This method is called Λ-fit and it is one of the corner stones of this study.

Of course, the most rigorous method would be to make Λ-fit on contents of (p_{T}, x, flavour) bins at parton level (even without detector simulation). So,

what are the consequences of our simple approximation (including detector simulation)?

Firstly, the fit is very efficient: the difference between the 6 fitted data points and the fit function is not more than 0.3 %, which can be neglected with respect to other sources of errors (statistical error, pdf uncertainties, JES uncertainties). Secondly, (p_{T}, x, flavour) bins would require a substantial increase of statistics, in order to keep their statistical errors reasonably small. That is hardly feasible.

We consider the Λ-fit as an enough precise approximation of the ideal Λ-fit. An example of this fit for two different p_{T} bins is shown in Fig. 7.3. The resulting Λ-fit functions are used for calculation of number of events in a given p_{T}-bin for any given value of Λ.

7.2 Systematic errors

In this Section systematic errors of inclusive dijet p_{T}-spectra are shown. The effect of 3% uncertainty in absolute luminosity measurement is not shown, nevertheless it is included in further calculations.

7.2.1 PDF uncertainties

Inclusive dijet p_{T}-spectra pdf errors were obtained with the pdf re-weighting method (see Sec. 2.3.4). The SM p_{T}-spectra obtained with 40 error pdfs of CTEQ6M1 (30 error pdfs of MRST2001) normalised to the p_{T}-spectrum obtained with the central CTEQ6M1 pdf (central MRST2001 pdf) are shown in Fig. 7.4. The yellow master error bands correspond to 90% C.L. The comparison of MRST and CTEQ relative p_{T} master errors is presented in Fig. 7.5, where the p_{T} spectra obtained with error pdfs of both groups were always normalised to the p_{T}-spectrum obtained with CTEQ6M1 central pdf. The MRST master error band lies mostly within the CTEQ master error band, but the p_{T}-spectrum obtained with central MRST pdf is decreasing at high p_{T} with respect to the p_{T}-spectrum obtained with central CTEQ pdf. At $p_{\mathrm{T}} = 4$ TeV the difference is almost 40%.

In order to calculate the pdf error for any value of Λ, Λ-fit was performed on simulated p_{T}-spectra (corresponding to SM and $\Lambda = 3$, 5, 10, 15 and 20 TeV) obtained with every pdf error set member (from CTEQ and MRST) and resulting master error, depending on Λ and p_{T} was calculated and divided by a factor of 1.64 to correspond to 1-sigma uncertainty.

7.2.2 NLO k-factors

NLOJET++ generator was used to calculate dijet p_{T}-spectra at parton level (hard process only) at LO and NLO order and their renormalisation and factorisation scale dependence, see Fig. 7.6. The red hatched band in the right-hand side plot is obtained with setting $\mu_R = \mu_F = 0.5$, 1.0, 2.0 p_{Tmax} and is taken as a contribution to systematic error.

Complementary ll2nn factors, that are ratios of p_{T}-spectra obtained with NLO and LO pdfs (see Sec. 6.1.3) are shown in Fig. 7.7. It is apparent, that the

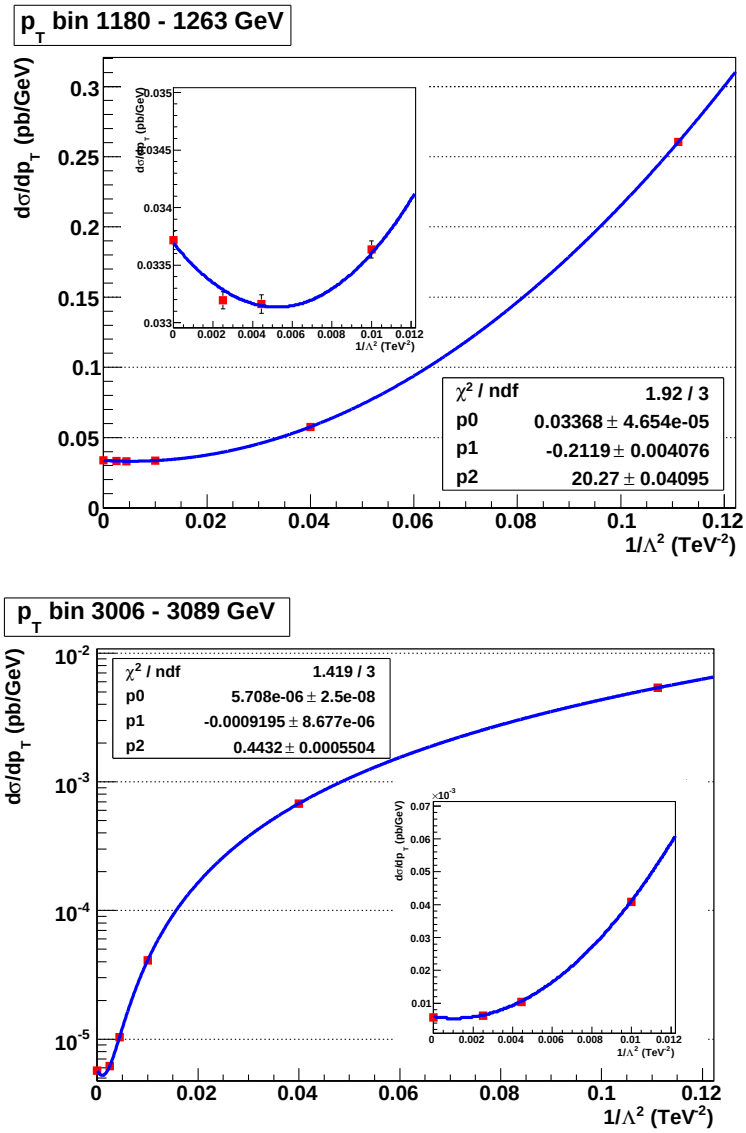


Figure 7.3: An example of various p_{T}-bin Λ-fit. The red squares correspond to the values obtained directly from simulated data samples with compositeness switched off and on (Λ = 3, 5, 10, 15 and 20 TeV). Statistical errors of the simulated data are usually smaller than the red squares. In both plots the low $1/\Lambda^2$ areas are zoomed in floating windows. The resulting Λ-fit functions are used for calculation of number of events in a given p_{T}-bin for any given value of Λ.

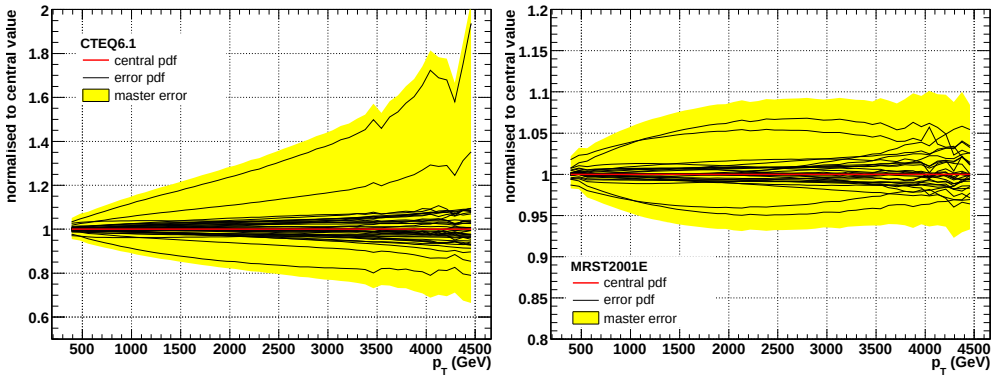


Figure 7.4: Relative pdf errors of SM inclusive dijet p_{T}-spectrum as obtained with CTEQ6M1 error pdfs (left) and MRST2001E error pdfs (right). The yellow master error band corresponds to 90% C.L. In case of CTEQ6M1 the relative master error grows rapidly with p_{T} up to about $+70\%$ and -30%. In case of MRST2001E the relative master error in p_{T}-region from 2000 to 4500 GeV is about $+10\%$ and -7% and is roughly constant with p_{T}. This behaviour is caused mainly by MRST2001E error pdf members number 17, 18, 19 and 20. At high-p_{T} region (above 4 TeV) statistical error influences this uncertainty too.

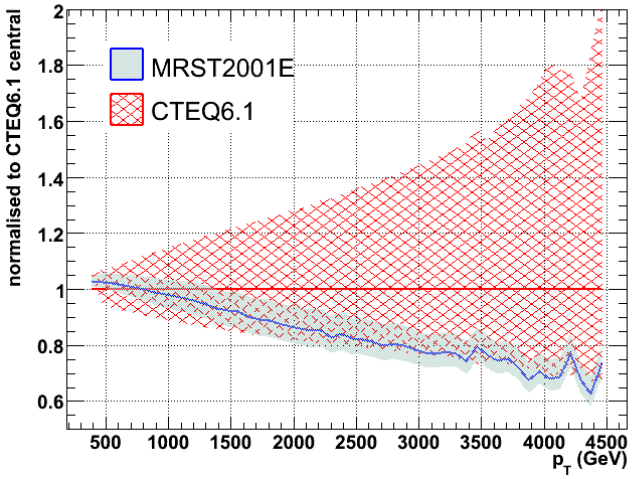


Figure 7.5: Inclusive dijet p_{T}-spectra master error bands and central values (solid lines) as obtained with CTEQ6M1 and MRST2001E, normalised to p_{T}-spectrum obtained with CTEQ central pdf. The master error bands correspond to 90% C.L. The MRST master error band lies mostly within the CTEQ master error band, but the p_{T}-spectrum obtained with central MRST pdf is decreasing at high p_{T} with respect to the p_{T}-spectrum obtained with central CTEQ pdf. At high-p_{T} region (above 4 TeV) statistical error influences this uncertainty too.

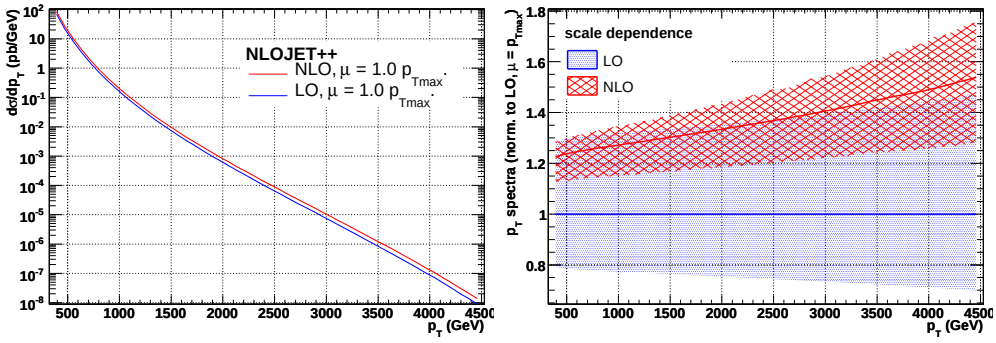


Figure 7.6: Inclusive dijet p_{T}-spectra at parton level (hard process only) as generated with NLOJET++ at LO and NLO order (left) and their renormalisation and factorisation scale dependence (right), obtained with setting $\mu_R = \mu_F =$ 0.5, 1.0, 2.0 p_{Tmax}.

ll2nn factors strongly depend on Λ. This is caused by the fact, that the relative abundance of quarks and gluons depends on the value of Λ and different pdf sets give different results for each kind of partons.

7.2.3 Jet energy scale uncertainty

The effect of 3% jet energy scale uncertainty on SM p_{T}-spectrum is shown in Fig. 7.8. In order to calculate the systematic error due to this uncertainty for any value of Λ, the JES for simulated samples (that is for SM, and Λ = 3, 5, 10, 15 and 20 TeV) was shifted $\pm$3% and Λ-fit was performed on resulting shifted JES p_{T}-spectra.

7.2.4 Comparison with Full Simulation

Standard model inclusive dijet p_{T}-spectrum sewed from full simulation CSC samples J4 – J8 is depicted in Fig. 7.9. For the purpose of this comparison, the generated events, used for this full simulation were also used as an input for ATLFAST.

The ratio of full and fast inclusive dijet p_{T}-spectra (FA factor – "full over ATLFAST") is depicted in Fig. 7.9. The p_{T}-spectrum ranging from 350 to 4500 GeV is divided into 50 bins. It is apparent, that full simulation predicts smaller number of high-p_{T} jets with respect to the p_{T} spectrum obtained by ATLFAST.

7.2.5 Dependence on Monte Carlo Generator

There are many Monte Carlo generators capable of computing hard process matrix elements, but only a few of them (such as Pythia, Herwig or Sherpa) are able to generate the complete event. The description of parton showers and hadronisation in Pythia and Herwig is based on different models. In order to estimate the systematic error caused by the selection of Pythia as a main

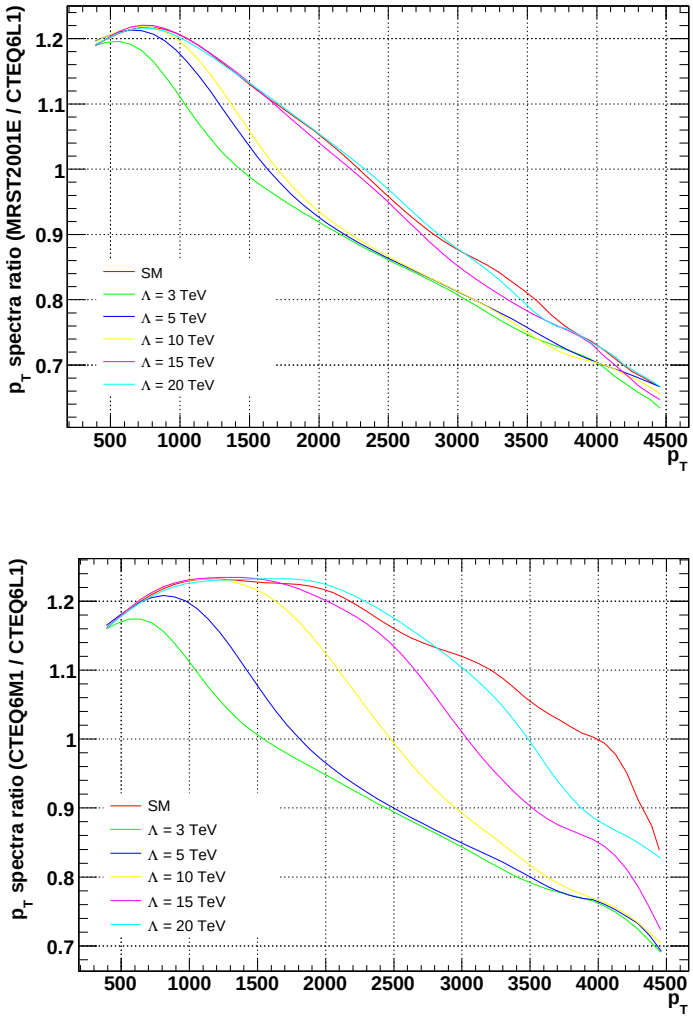


Figure 7.7: Ratios of inclusive dijet p_{T}-spectra obtained in Pythia with MRST2001E and CTEQ6L1 pdfs (top) and CTEQ6M1 and CTEQ6L1 pdfs (bottom).

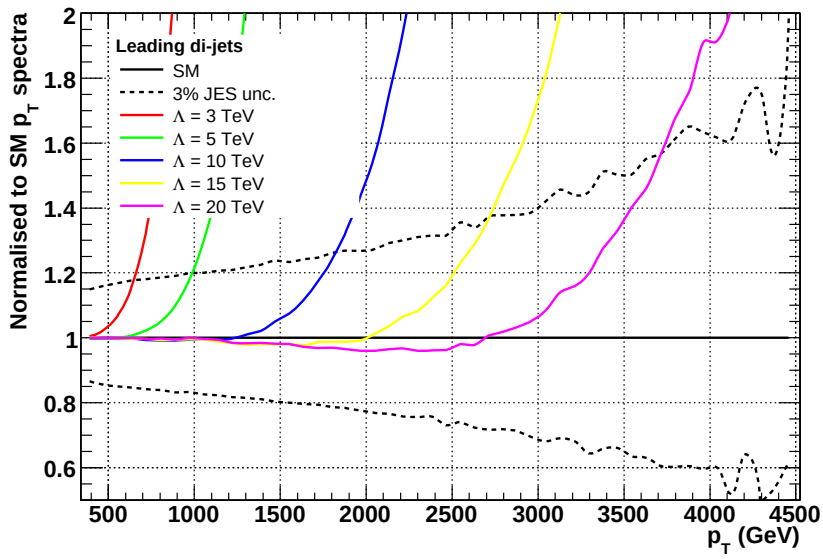


Figure 7.8: Inclusive dijet p_{T}-spectra for various values of Λ normalised to SM spectrum. The two dashed lines show the systematic error caused by 3% jet energy scale uncertainty. ATLFAST, PYTHIA, CTEQ6L1.

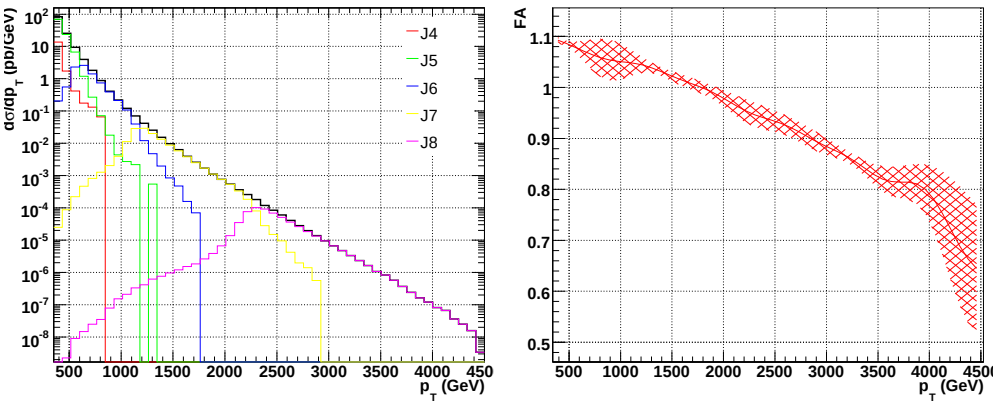


Figure 7.9: Full simulation of SM p_{T}-spectrum (left). Ratio of inclusive dijet p_{T} spectra from full and fast simulation (right). Showed error band corresponds to a statistical error due to limited amount of events available for full simulation. It is apparent, that full simulation predicts smaller number of high-p_{T} jets.

Monte Carlo generator for this study, a comparison of inclusive dijet p_{T}-spectra generated with PYTHIA and JIMMY/HERWIG was performed. The result is presented in Fig. 7.10. It is apparent, that JIMMY/HERWIG predicts about 15% less events than PYTHIA in whole studied p_{T}-range. This systematic error was not taken into account in further calculations, it would cause a decrease of compositeness exclusion limits in an order of per cent.

The detailed settings of both Monte Carlo generators used for this comparison are available in Appendix A.

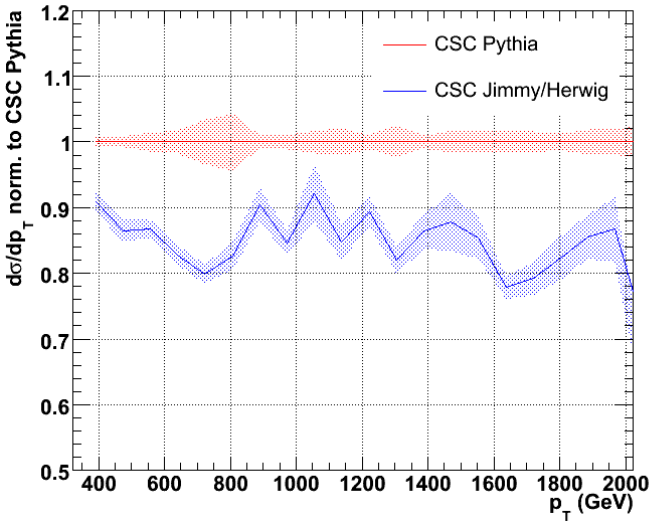


Figure 7.10: Inclusive dijet p_{T}-spectra at the detector level generated with PYTHIA and JIMMY/HERWIG, normalised to the PYTHIA case. Full simulation CSC dijet samples were used. The shape of JIMMY/HERWIG normalised p_{T}-spectrum is affected by a smaller statistics available in full simulation. It is apparent, that JIMMY/HERWIG predicts about 15% less events than PYTHIA.

7.3 Bayesian approach

Bayesian technique is widely used in order to determine the lower limit on compositeness scale [25, 23]. The output of this technique is the posterior probability density $P(\Lambda|d)$, telling us how probably is given value of Λ incorporated in nature, given the data d we measured in our experiment. Employing the Bayes' theorem (3.14) for inclusive dijet p_{T}-spectrum is done as follows:

$$P(\Lambda|d) = \frac{1}{Z} P(\Lambda) \prod_{k=1}^{k_{max}(L)} P^k(N_d^k|N_{exp}^{\Lambda,k}) \otimes G_k(JES(\Lambda), pdf(\Lambda), \delta L, \mu, FA), \tag{7.3}$$

where $1/Z$ is a normalisation factor and the product runs over p_{T}-bins starting at 350 GeV up to the $k_{max}(L)$ bin, which is the highest p_{T}-bin of Standard

Model QCD spectrum containing at least one event at a given total integral luminosity L. The likelyhood function takes a form of Poisson distribution[2]:

$$P^k(N_d^k|N_{exp}^{\Lambda,k}) = \frac{e^{-N_{exp}^{\Lambda,k}}(N_{exp}^{\Lambda,k})^{N_d^k}}{N_d^k!}, \tag{7.4}$$

where N_d^k is a number of events in k^{th} p_{T}-bin in experimental data and $N_{exp}^{\Lambda,k}$ is an *expected* number of events in k^{th} p_{T}-bin as predicted by theory for a given value of Λ. The values $N_{exp}^{\Lambda,k}$ were obtained as a product of FA factor (at given p_{T}) and a result a Λ-fit of ATLFAST data, to which k-factors and ll2nn factors were applied. Similarly for N_d^k the values of $N_{exp}^{\Lambda,k}$ were used, either with $\Lambda = \infty$ corresponding to pure SM, when we are interested in exclusion limits, or with finite value of Λ, when we are interested in discovery of compositeness.

In order to include smearing due to systematic errors, the likelyhood is convoluted with a Gaussian distribution $G_k(JES(\Lambda), pdf(\Lambda), \delta L, \mu, FA)$ of mean equal to zero and σ involving contributions from jet energy scale uncertainty and parton distribution function uncertainty (both Λ dependent), absolute luminosity measurement uncertainty, renormalisation and factorisation scale uncertainty and FA factor uncertainty. All systematic errors are treated as independent.

Concerning the prior probability density $P(\Lambda)$, a common choice is $P(\Lambda) \propto 1/\Lambda^2$ or $P(\Lambda) \propto 1/\Lambda^4$, as inspired by the form of langrangian in Eq. (3.5). Another choice, $P(\Lambda) \propto \exp(-\Lambda/\mu)$, where μ is a chosen scale, was used in [24]. We tried this last prior with $\mu = 10$ TeV, this value is higher, than the current limit on quark compositeness set by Tevatron, but such a compositeness scale still could be measured by ATLAS.

There is also a good statistical reason, why to choose one of the priors above. It is so-called maximum information entropy choice of prior probability, [58]. As our current knowledge about Λ is very limited (we know only its lower limit), the prior pdfs should represent our ignorance of Λ and their information entropy should be the largest. This is really the case: $P(\Lambda) \propto 1/\Lambda^2$ gives a maximum entropy for $1/\Lambda^2$ in interval $[0,1/\Lambda_{min}^2]$, $P(\Lambda) \propto 1/\Lambda^4$ gives a maximum entropy for $1/\Lambda^4$ in interval $[0,1/\Lambda_{min}^4]$ and $P(\Lambda) \propto \exp(-\Lambda/10\ \mathrm{TeV})$ gives a maximum entropy for Λ in interval $[\Lambda_{min},\infty)$, when a mean value $\mu = 10$ TeV of this pdf is given. As the best choice of prior we would favour the first choice: $P(\Lambda) \propto 1/\Lambda^2$, as $1/\Lambda^2$ is the parameter present in the effective Lagrangian in Eq. (3.5). The least favourable choice is $P(\Lambda) \propto \exp(-\Lambda/10\ \mathrm{TeV})$; firstly Λ^1 is not present in the Lagrangian, and secondly this prior requires an arbitrary parameter μ.

The value Λ_{min} can be set to 2.7 TeV, according to the limit already measured (see Sec. 3.4), as long as this limit is far enough from the limits achievable by ATLAS, it does not affect the results anyway. Finally, the results of Bayesian technique can only be relevant, if the bias introduced by the choice of prior, is reasonably small. This will be discussed in Sec. 7.3.1.

The exclusion limits Λ_{lim} obtained with prior $P(\Lambda) \propto 1/\Lambda^2$ and corresponding to a given confidence level C.L. are calculated according to Eq. (3.15) and their values are listed in Tab. 7.1 and shown in Fig. 7.11.

[2]in case of $N_d^k > 30$ this is approximated by a Gaussian distribution for technical reasons.

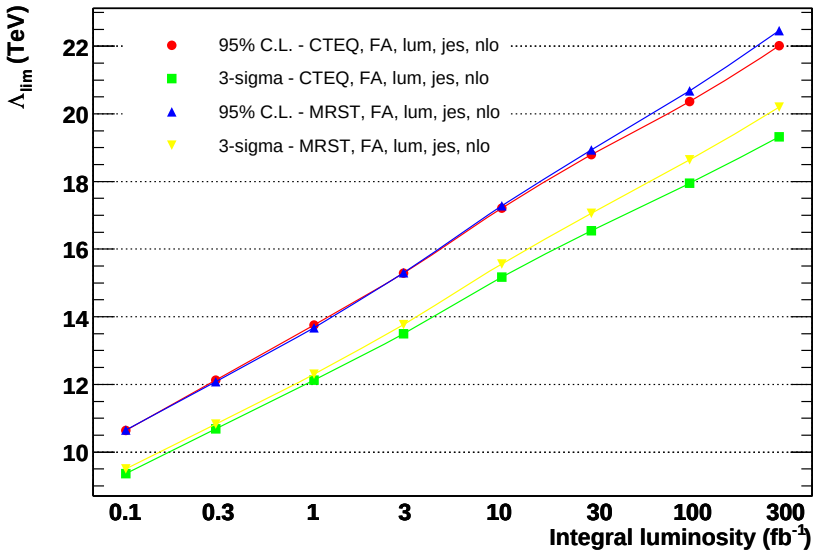


Figure 7.11: Dependence of p_{T}-Bayes exclusion limits Λ_{lim} (TeV) on total integral luminosity.

Int. luminosity	95% C.L.		3-sigma	
(fb^{-1})	MRST	CTEQ	MRST	CTEQ
0.1	10.65	10.64	9.51	9.36
0.3	12.08	12.13	10.82	10.69
1.0	13.68	13.75	12.30	12.13
3.0	15.30	15.29	13.77	13.50
10.	17.28	17.21	15.56	15.17
30.	18.94	18.80	17.06	16.55
100.	20.68	20.36	18.64	17.95
300.	22.46	22.01	20.20	19.32

Table 7.1: Values of exclusion limits Λ_{lim} (TeV) at 95% C.L. and 3-sigma level obtained by p_{T}-Bayes with CTEQ6M1 or MRST2001E pdfs for various values of total integral luminosity.

7.3.1 Dependence of posterior probabilities on priors

In this section we study the dependence of posterior probabilities on priors. Posterior pdf computed with various priors are shown in Fig. 7.12. Corresponding values of Λ_{lim} are listed in Tab. 7.2. Although the shapes of posteriors apparently differ, the difference in exclusion limits is less pronounced: 5% between the smallest and the largest value. This difference is still not small, but it is not surprising: after the measurement our ignorance of Λ is still large - we still do not know its value, we only know its lower limit.

On the contrary to the case, when measured data corresponds to SM p_{T}-spectrum, the dependence of posterior on prior is much more modest, when the likelyhood function is more narrow, that is when measured data corresponds to p_{T}-spectrum with some finite value of Λ switched on. This is studied in next section.

Yet another issue must be clarified. If the measured data corresponds to SM p_{T}-spectrum, why the maximum of the posterior pdf is not at $1/\Lambda^2 = 0$, as we see in Fig. 7.12? Firstly, it depends on the prior, the priors $1/\Lambda^4$ and $\exp(-\Lambda/10\ \mathrm{TeV})$ obviously move the maximum position to the right. But even in case of prior flat in $1/\Lambda^2$ the maximum is not at zero. That is caused by a combination of two effects: the existence of undershoot p_{T}-bins for spectra with compositeness switched on (see introduction of Sec. 7 and Fig. 7.8) and the dependence of systematic error on Λ which follows these undershoot bins. But as we are aware of this effect we can take it into account. The posterior pdf maximum position corresponds much better to the measured value of Λ, when compositeness would be discovered, as illustrated in the next section.

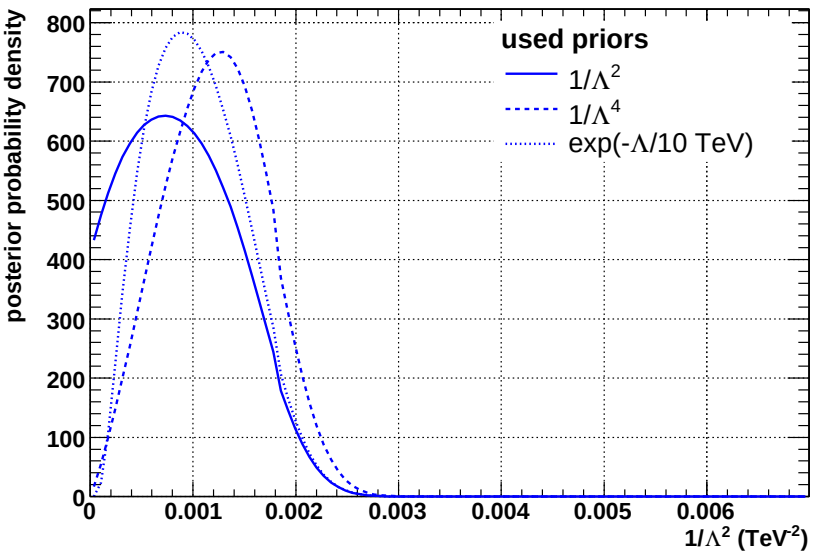


Figure 7.12: Posterior pdf computed with various priors. Measured data = SM only, compositeness switched off. These pdfs are used for the calculation of exclusion limits. CTEQ6M1 pdf used, total integrated luminosity is 300 fb^{-1}.

prior	Λ_{lim} (TeV)
$1/\Lambda^2$	22.01
$1/\Lambda^4$	20.86
$\exp(-\Lambda/10$ TeV)	21.86

Table 7.2: Values of exclusion limits Λ_{lim} (TeV) at 95% C.L. for p_{T}-Bayes method at 300 fb^{-1} computed with various prior probabilities. CTEQ6M1 pdfs used.

7.3.2 Discovery potential

The p_{T}-Bayes method can be used to test any kind of measured p_{T} spectra for compositeness. In this section two other cases are studied with compositeness switched on: $\Lambda = 15$ TeV and $\Lambda = 10$ TeV. Compare Figs. 7.13 and 7.14 for the shapes of posterior pdf obtained with various priors. In these cases, when compositeness would be discovered a 5-sigma discovery interval can also be calculated. As the posterior pdfs are non-Gaussian asymmetric functions, the shortest 5-sigma (and for illustration 3-sigma) confidence levels were found, as stated in Tabs. 7.3 and Tab. 7.4.

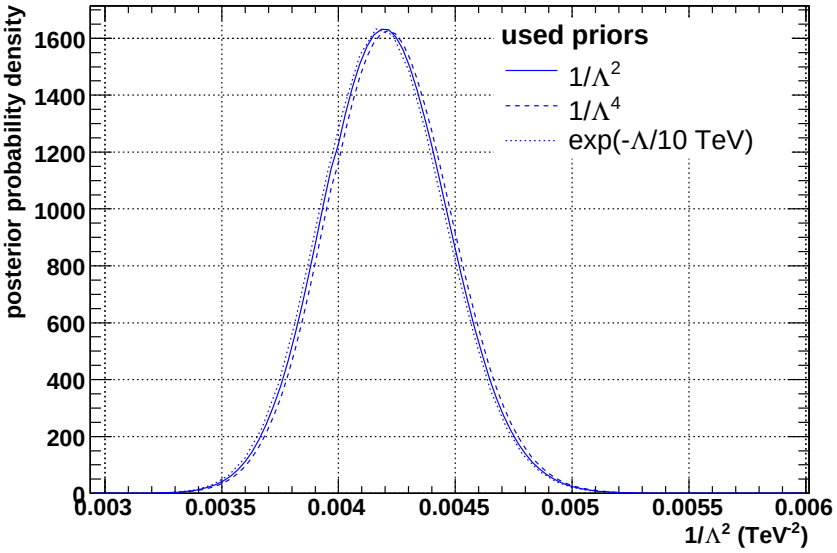


Figure 7.13: Posterior probability density for $\Lambda = 15$ TeV and various priors at 300 fb^{-1}. CTEQ6M1 pdfs used.

We also observe, that the lower the value of Λ, the better the positions of posterior maxima correspond to it and the smaller dependence on priors is present – in case of $\Lambda = 10$ TeV it almost disappears. In this case our ignorance of Λ would be very small, so our choice of prior naturally plays only a negligible role.

prior	Λ_{max} (TeV)	3-sigma	5-sigma
$1/\Lambda^2$	15.44	14.09 to 17.08	13.30 to 18.51
$1/\Lambda^4$	15.41	14.07 to 17.04	13.27 to 18.45
$\exp(-\Lambda/10$ TeV)	15.46	14.11 to 17.11	13.31 to 18.55

Table 7.3: Values of the most probable value Λ_{max} and the shortest 5-sigma interval for various priors when compositeness with $\Lambda = 15$ TeV is switched on. CTEQ6M1 pdfs, 300 fb^{-1}.

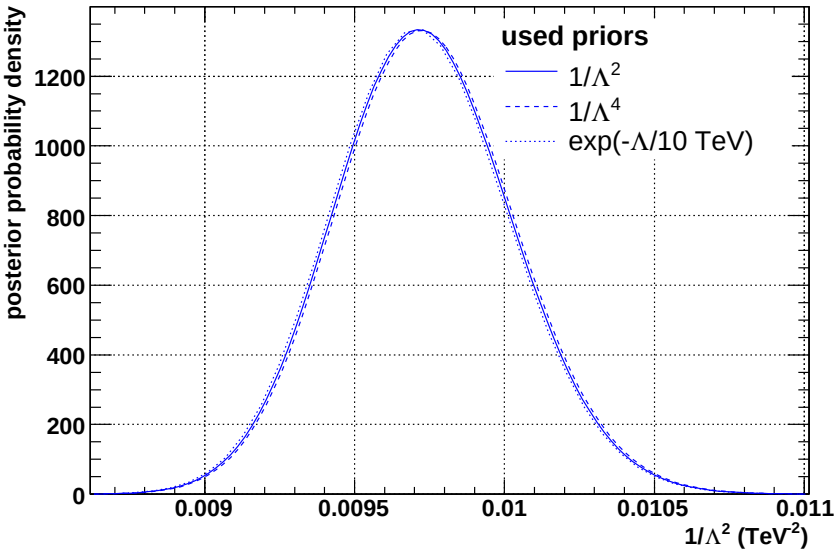


Figure 7.14: Posterior probability density for $\Lambda = 10$ TeV and various priors at 300 fb^{-1}. CTEQ6M1 pdfs used.

prior	Λ_{max} (TeV)	3-sigma	5-sigma
$1/\Lambda^2$	10.146	9.684 to 10.612	9.379 to 10.933
$1/\Lambda^4$	10.141	9.680 to 10.608	9.374 to 10.928
$\exp(-\Lambda/10$ TeV)	10.150	9.689 to 10.617	9.384 to 10.938

Table 7.4: Values of the most probable value Λ_{max} and the shortest 5-sigma interval for various priors when compositeness with $\Lambda = 10$ TeV is switched on. CTEQ6M1 pdfs, 300 fb^{-1}.

7.4 Rdist3

In this method, in order to quantify the difference between p_{T} spectra with compositeness switched on and off, the following ratio of number of event below and above chosen p_{T}^0 cut is used:

$$R(\Lambda) = \left(\frac{N(p_{\mathrm{T}} > p_{\mathrm{T}}^0)}{N(p_{\mathrm{T}} < p_{\mathrm{T}}^0)} \right)_{\Lambda} \tag{7.5}$$

The difference between SM and SM+CT spectra, measured in standard deviation of the difference between R values of these spectra is simply:

$$R_{dist3} = \frac{R(\Lambda) - R(SM)}{\sqrt{\sigma^2_{R(\Lambda)} + \sigma^2_{R(SM)}}} \tag{7.6}$$

The value of p_{T}^0 cut is chosen to get maximum value of $Rdist3$ for every simulated Λ, as shown in Fig. 7.15. In order to be able to calculate $Rdist3$ for any value of Λ, we must fit the p_{T}^0 cut dependence on Λ and the resulting fit function use as a new definition of the p_{T}^0 cut (see Fig. 7.15). The only requirement for the fit function is, that it should be a smooth function, that fits p_{T}^0 cut values for simulated Λ well. p_{T}^0 cut is now *defined* according to the formula:

$$p_{\mathrm{T}}^0 \text{ (GeV)} = -520.7 + 109.7 \times \Lambda(\text{TeV}) + 1099 \times (\Lambda(\text{TeV}) - 2 \text{ TeV})^{0.1} \tag{7.7}$$

One might wonder, which value of Λ shall we use for the calculation of p_{T}^0 cut when the value of Λ is the one we want to measure and we do not know that value prior to the experimental measurement. Therefore, we must calculate $Rdist3$ for all possible cases of Λ and corresponding cuts. The highest value of $Rdist3$ will be reached with p_{T}^0 cut optimal for the Λ which is experimentally observed. Practically, $R(SM)$ in Eq. 7.6 will be based on our Monte Carlo prediction and $R(\Lambda)$ in Eq. 7.6 will be calculated from real measured data.

The values of $N(p_{\mathrm{T}} > p_{\mathrm{T}}^0)$ and $N(p_{\mathrm{T}} < p_{\mathrm{T}}^0)$ for a given Λ were calculated using the Λ-fit functions from Sec. 7.1. There are 50 bins of dijet p_{T}-spectra from 350 to 4500 GeV: edges of its 35 p_{T}-bins ranging from 848 to 3753 GeV correspond to p_{T}^0 cuts optimised for values of Λ spanning from 2.75 TeV to 25.24 TeV, as expressed in Eq. 7.7. The contents of p_{T}-bins below and above p_{T}^0 cuts were summed in order to calculate $Rdist3 = Rdist3(\Lambda)$ – see Fig. 7.16.

7.4.1 Exclusion limits

When calculating $Rdist3$ exclusion limits, the following factors and systematic errors are included: k-factors and ll2nn factors, pdf uncertainties, JES uncertainties, FA factor and its uncertainties. The uncertainty on integral luminosity measurement cancels in $Rdist3$ by definition. A comparison of relative $Rdist3$ systematic errors is in Fig. 7.17. It is apparent, that the largest contribution to the systematic error comes from the JES uncertainty.

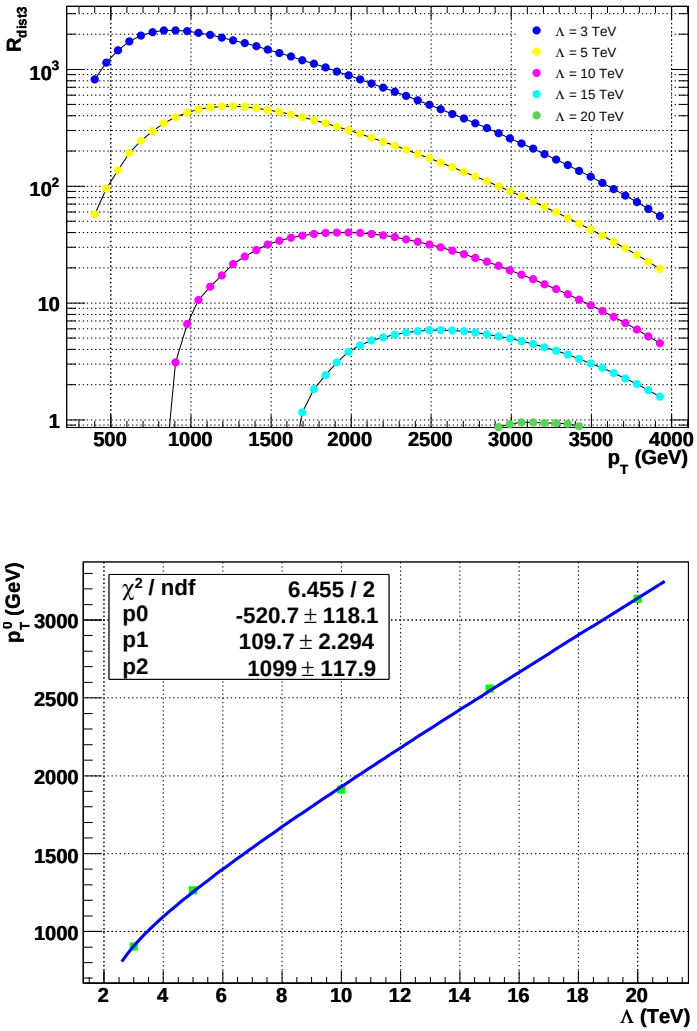


Figure 7.15: *Rdist*3 dependence on p_{T}^0 cut at 30 fb^{-1}. ATLFAST, Pythia, CTEQ6L1 (top). Note, that the maximum of the curves is rather flat. Fitted dependence of p_{T}^0 cuts on Λ (bottom). The resulting fitting function is used as a definition of the optimal p_{T}^0 cut for a given value of Λ.

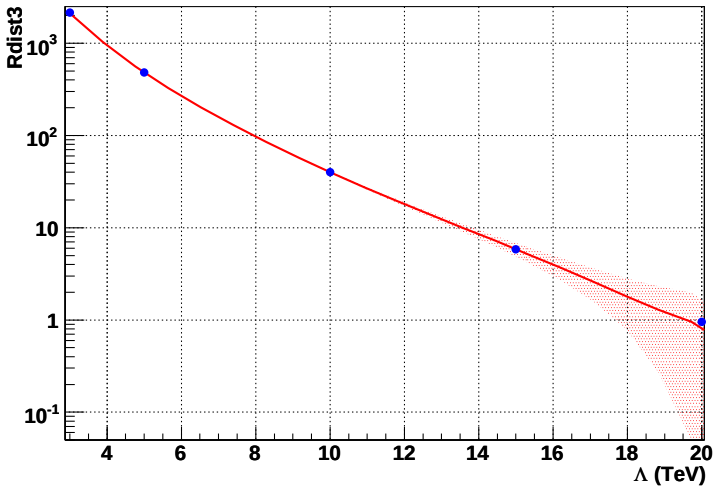


Figure 7.16: *Rdist*3 dependence on Λ at 30 fb^{-1} as obtained with the aid of Λ-fit and p_{T}^0 cut definition in Eq. (7.7). Statistical error shown by the red band. The blue points represent the values of *Rdist*3 as calculated directly from simulated data samples.

Finally, *Rdist*3 exclusion limits are depicted in Fig. 7.18 and listed in Tab. 7.5. At higher integral luminosities *Rdist*3 exclusion limits grow more slowly due to larger systematic errors – an effect of collecting higher statistics is not that large anymore.

To be complete, the discovery limits for *Rdist*3 method are not calculated, as it is not as sensitive as p_{T}-Bayes method.

Int. luminosity	95% C.L.		3-sigma	
(fb^{-1})	MRST	CTEQ	MRST	CTEQ
0.1	10.49	10.39	9.47	9.36
0.3	11.72	11.61	10.62	10.49
1.0	13.09	12.96	11.84	11.68
3.0	14.25	14.10	12.87	12.69
10.	15.25	15.09	13.80	13.60
30.	15.95	15.77	14.25	14.04
100.	16.31	16.12	14.44	14.22
300.	16.40	16.21	14.47	14.25

Table 7.5: Values of exclusion limits Λ_{lim} (TeV) at 95% C.L. and 3-sigma level obtained by *Rdist*3 method with CTEQ6M1 or MRST2001E.

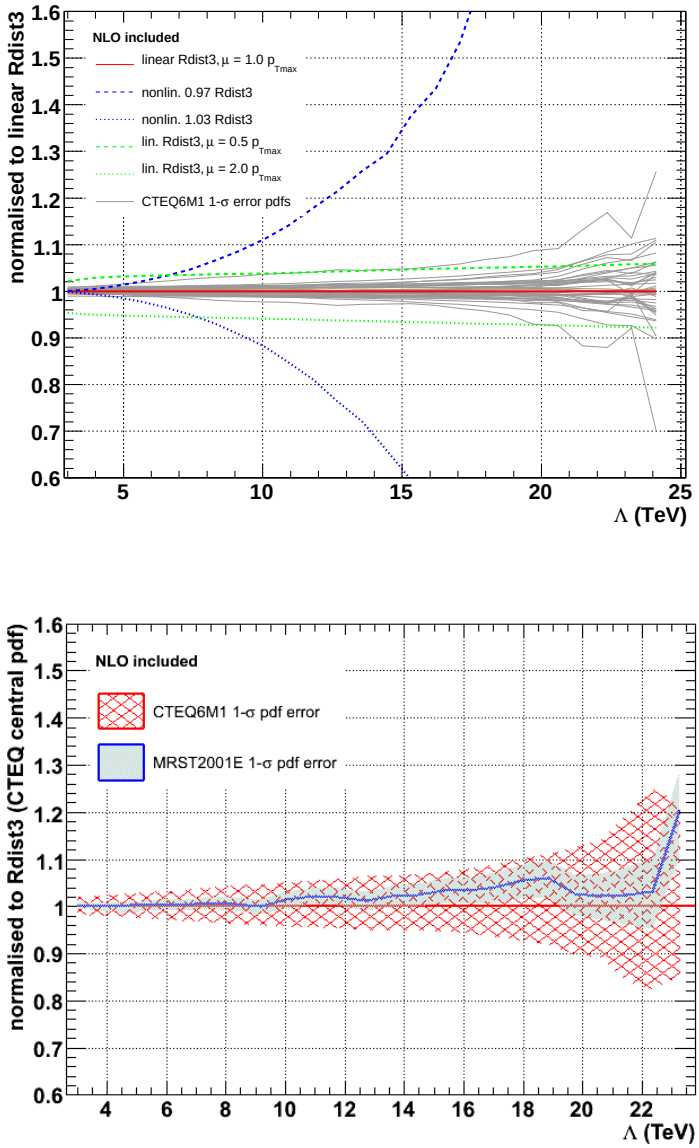


Figure 7.17: *Rdist*3 relative systematic error comparison (top). It is apparent, that the largest contribution to the systematic error of *Rdist*3 comes from JES uncertainty. *Rdist*3 pdf master error comparison as obtained with CTEQ6M1 and MRST2001E pdf sets (bottom). The pdf errors were divided by a factor of 1.64 in order to scale down from original 90% C.L. to a 1-sigma level.

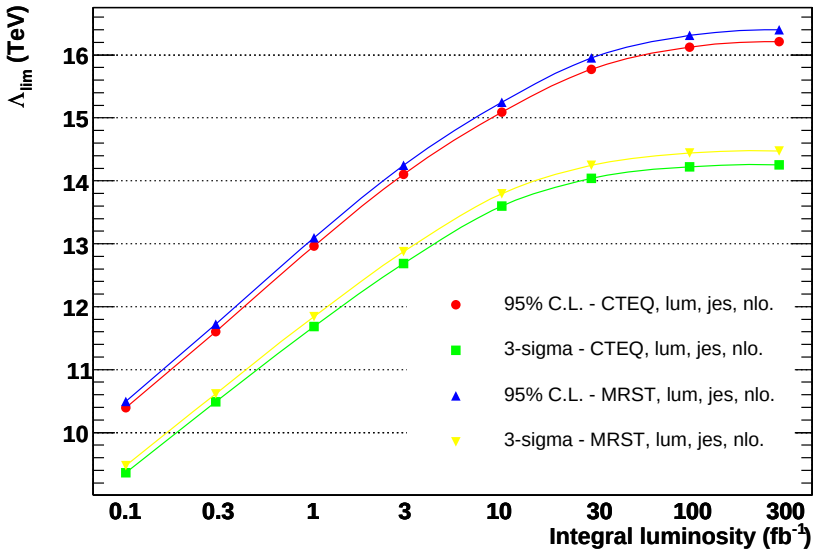


Figure 7.18: *Rdist*3 exclusion limits, as obtained with CTEQ6M1 and MRST2001E.

Chapter 8

Dijet angular distribution

In this Chapter the effect of quark compositeness on inclusive dijet angular distribution is studied with two methods, R1 and χ-Bayes (in Sections 8.3 and 8.4). Let us start with the description of the inclusive dijet angular distribution and related systematic errors.

To study a leading dijet angular distribution it is useful to employ the variable χ, which is Lorentz-invariant for massless objects:

$$\chi = \exp|\eta_1 - \eta_2|, \tag{8.1}$$

where η_1 and η_2 are pseudorapidities of the two leading jets. A typical example of χ-distribution for various values of Λ, normalised to the total number of events, passing the cuts, is shown in Fig. 8.1. Quark compositeness causes an excess of events at small values of χ, as can be derived from Eqs. 3.8.

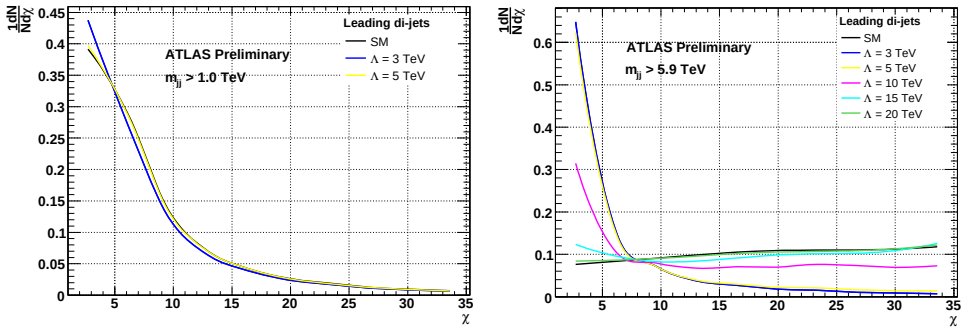


Figure 8.1: Normalised χ-distribution for various values of Λ with cut on dijet invariant mass $m_{jj} > 1$ TeV (left) and $m_{jj} > 5.9$ TeV (right). χ is in the range of 1 to 35, and the spectra are binned into 10 bins.

One would like to benefit from measuring angular distributions and get a rid of JES uncertainties. But when we are interested in measuring quark compositeness, this is unfortunately not possible – without a lower cut on leading dijet invariant mass m_{jj}, the effect of compositeness is very week, as presented in Fig. 8.1: with $m_{jj} > 1$ TeV the angular distribution for $\Lambda = 5$ TeV almost

coincides with SM distribution, while with $m_{jj} > 5.9$ TeV even $\Lambda = 15$ TeV can be spot, despite very reduced statistics at such a high m_{jj}.

8.1 Λ-fit

Analogically to p_{T}-spectra Λ-fit performed in Sec. 7.1, the Λ-fit method was performed on the contents of dijet angular distribution bins and Eq. (7.1) was employed. The χ-distribution is binned into 10 bins in the range of χ from 1 to 35 in order to keep the statistical error of the simulation reasonably small. The highest cut on dijet invariant mass was chosen to be $m_{jj} = 5.9$ TeV. This value is somewhat arbitrary (in fact, it was inspired by R1 method, as will be shown later), but it would not make much sense to set even higher m_{jj} cut, as the uncertainty in JES of corresponding jets might be rather large and statistics fairly low. An example of χ-bin Λ-fit is shown in Fig. 8.2. The largest difference of fitted data and the fit function in case of χ-bin Λ-fit is 0.5%. Corresponding χ-Λ plane is depicted in Fig. 8.3. The results of Λ-fit performed on χ-distribution, where k-factors and ll2nn factors were applied, were used for χ-Bayes method in Sec. 8.4. For the needs of R1 method, Λ-fit was performed on different angular distributions, as described in Sec. 8.3.

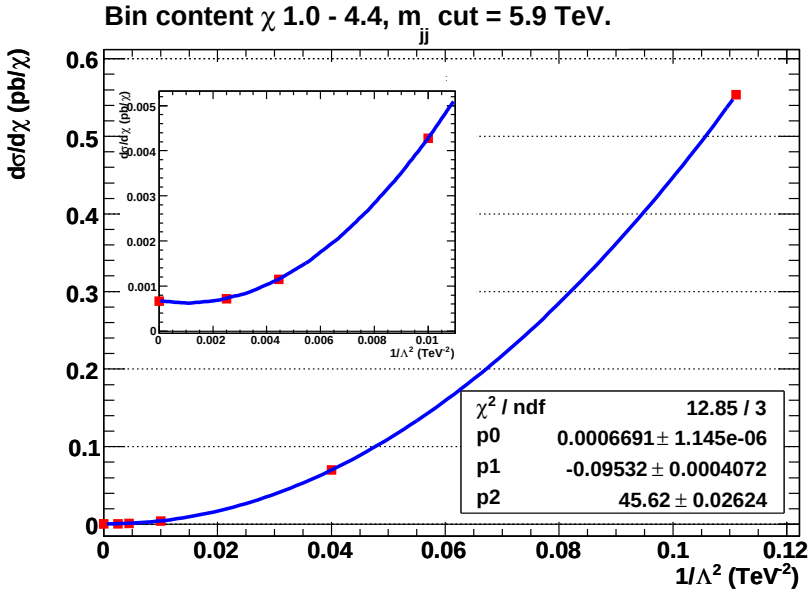


Figure 8.2: The Λ-fit of χ-bin (1.0; 4.4) with cut $m_{jj} > 5.9$ TeV applied. ATLFAST, Pythia, CTEQ6L1. The red squares correspond to the values obtained directly from simulated data samples with compositeness switched off and on ($\Lambda = 3$, 5, 10, 15 and 20 TeV). The low $1/\Lambda^2$ area is zoomed in a floating window. The resulting fitting function is used for calculation of number of events in the given bin for any given value of compositeness scale Λ.

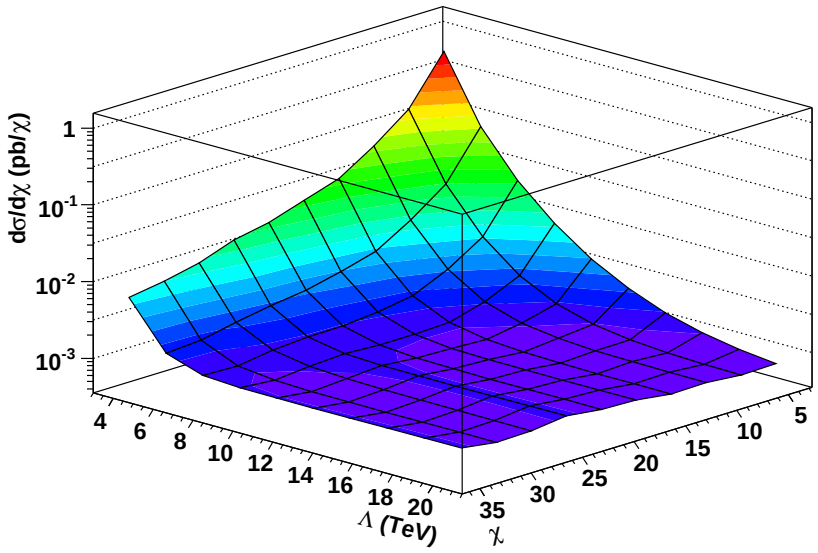


Figure 8.3: Dependence of differential cross-section $\mathrm{d}\sigma/\mathrm{d}\chi$ on χ and Λ when cut $m_{jj} > 5.9$ TeV is applied. One can see, that with Λ not only the shape of χ-distribution changes, but the absolute values of $\mathrm{d}\sigma/\mathrm{d}\chi$ change rapidly, too. ATLFAST, PYTHIA, CTEQ6L1.

8.2 Systematic errors

In this Section systematic errors of χ-spectra are shown. The effect of 3% uncertainty in absolute luminosity measurement is not shown, nevertheless it is included in further calculations.

8.2.1 PDF uncertainties

Analogically to p_{T}-spectra pdf uncertainties in Sec. 7.2.1, the pdf uncertainties were calculated for χ-distribution obtained from dijets passing the cut $m_{jj} > 5.9$ TeV as shown in Fig. 8.4. A comparison of master error bands from CTEQ6M1 and MRST2001 is illustrated in Fig. 8.5. Again MRST master error band lies mostly within CTEQ master error band and is significantly more narrow. MRST pdfs also predict smaller number of events in χ-distribution at high m_{jj}.

8.2.2 NLO k-factors

NLO k-factors and their scale dependence were calculated with NLOJET++ generator analogically to procedure in Sec. 7.2.2. The k-factor systematic error due to scale dependence was again obtained by varying $\mu_R = \mu_F = 0.5$, 1.0, 2.0 p_{Tmax} (see Fig. 8.6). Complementary ll2nn factors were applied to χ-spectra too.

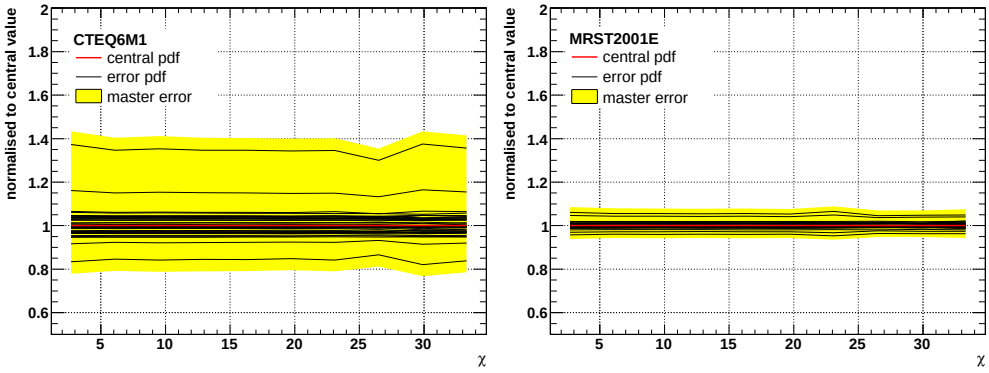


Figure 8.4: Relative pdf uncertainties of SM χ-distribution with cut $m_{jj} > 5.9$ TeV. CTEQ6M1 (left) and MRST2001E (right) sets used. Uncertainties correspond to 90% C.L.

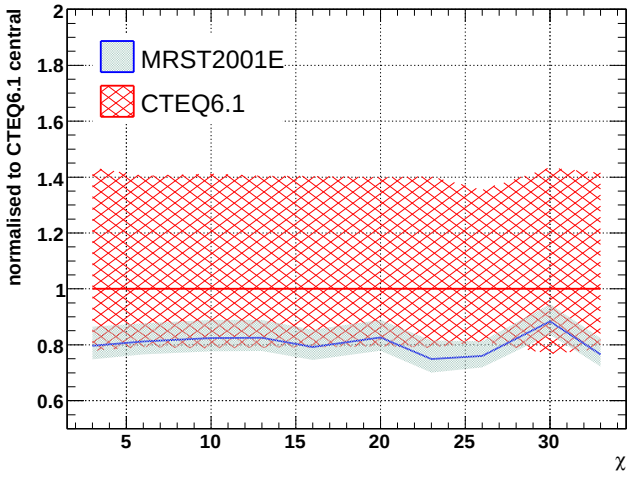


Figure 8.5: A comparison of pdf uncertainty of SM χ-distribution with cut $m_{jj} > 5.9$ TeV applied as obtained with CTEQ6M1 and MRST2001E. The MRST master error band lies almost within the CTEQ master error band. The MRST predicts smaller mean value of the χ-distribution than CTEQ.

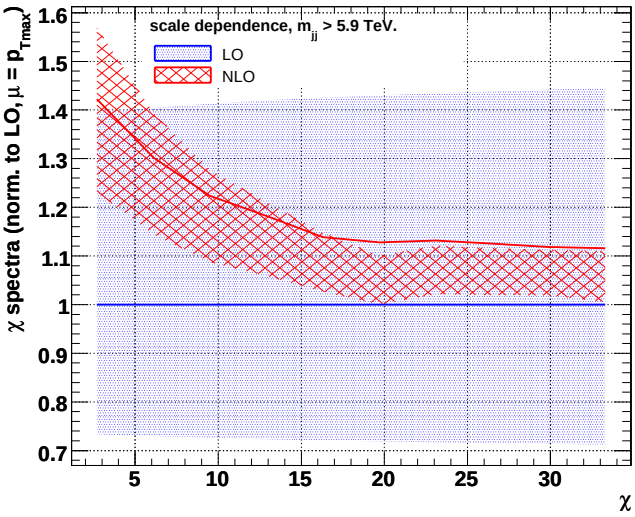


Figure 8.6: χ-distribution k-factor (solid red line). The blue and red hatched band were obtained by setting $\mu_R = \mu_F = 0.5,\ 1.0,\ 2.0\ p_{\mathrm{Tmax}}$ at LO, NLO, respectively.

8.2.3 Jet energy scale uncertainty

The jet energy scale uncertainty is introduced into angular distribution spectra due to the necessity to use lower cuts on dijet invariant mass (see introduction of Sec. 8). An effect of 3% JES on χ-distribution calculated with different m_{jj} cuts is illustrated in Fig. 8.7.

In order to specify χ JES uncertainty for any value of Λ, the JES for simulated samples (that is for SM, and $\Lambda = 3$, 5, 10, 15 and 20 TeV) was shifted $\pm 3\%$ and Λ-fit was performed on resulting shifted JES χ-spectra.

8.2.4 Comparison with Full Simulation

Full simulation CSC samples J4 – J8 were used to study the difference between full and fast simulation. For the purpose of this comparison, the generated events, used for full simulation were also used as an input for ATLFAST.

The ratio of full and fast χ-spectra (FA factor – "full over ATLFAST") is depicted in Fig. 8.8. Only events with dijet invariant mass m_{jj} passing a threshold of 5.9 TeV were used in this case. The region of dijets with such high m_{jj} is very important for compositeness study, unfortunately the statistical error due to a limited amount of events available in full simulation is so large, that in case of χ-spectra the FA factor was not applied at all.

8.3 R1

Analogically to *Rdist*3 method for p_{T}-spectrum (Sec. 7.4) the $R1$ method uses a similar ratio of number of event below and above chosen χ_{cut} to quantify the

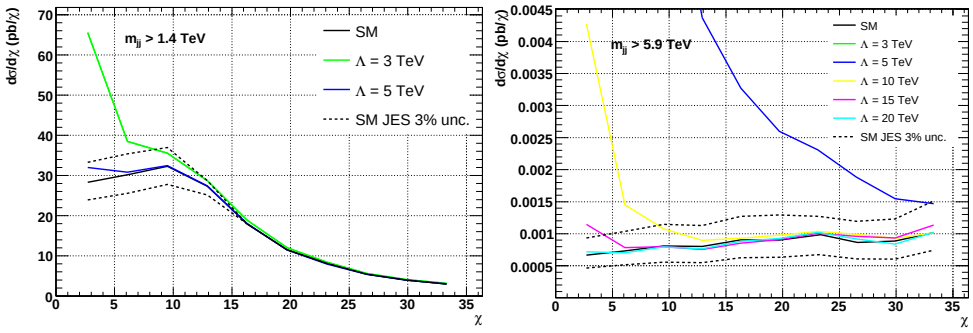


Figure 8.7: χ-distribution for SM and various values of Λ obtained with $m_{jj} > 1.4$ TeV (left) and $m_{jj} > 5.9$ TeV (right). Black dashed lines show 3% JES uncertainty band for SM. With cut $m_{jj} > 1.4$ TeV applied the χ-distributions for $\Lambda \geq 10$ TeV would practically sit on the top of SM χ-distribution and therefore they are not shown. Even $\Lambda = 5$ TeV is almost "lost" within the JES uncertainty. With cut $m_{jj} > 5.9$ TeV on the other hand, even $\Lambda = 15$ TeV is distinguishable from the SM case.

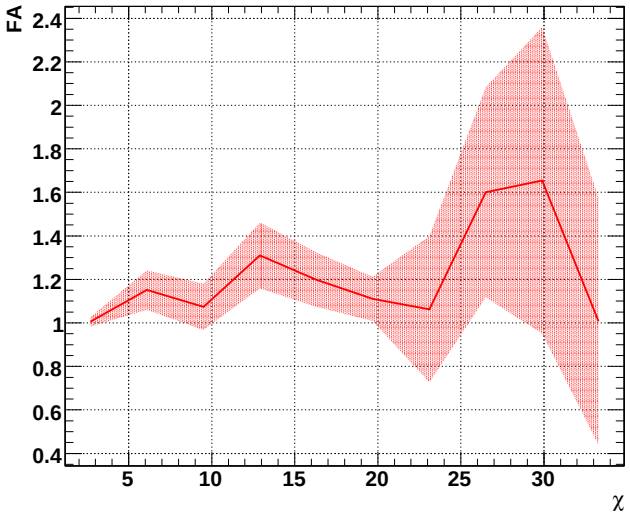


Figure 8.8: Ratio of χ spectra from full and fast simulation, cut $m_{jj} > 5.9$ TeV used. Showed error band corresponds to a statistical error due to limited amount of events available in full simulation. In this case the statistical error is so large, that the obtained FA factor for χ spectrum was not used.

difference between χ spectra with compositeness switched on and off:

$$R_\chi(\Lambda) = \left(\frac{N(\chi < \chi_{\text{cut}})}{N(\chi > \chi_{\text{cut}})} \right)_\Lambda \tag{8.2}$$

The difference between SM and SM+CT spectra, measured in standard deviation of the difference between R_χ values of these spectra is then:

$$R_1 = \frac{R_\chi(\Lambda) - R_\chi(SM)}{\sqrt{\sigma_\Lambda^2 + \sigma_{SM}^2}} \tag{8.3}$$

As already mentioned, it is necessary to introduce a lower cut on m_{jj} in order to obtain χ-spectra, where the effect of compositeness can be spot. Thus, for the $R1$ method, there are now two cuts to be optimised to get the highest value of $R1$ for a given Λ: χ_{cut} and a lower cut on dijet invariant mass m_{jj}. An example of $R1$ dependence on both these cuts is depicted in Fig. 8.9 for $\Lambda = 10$ TeV. Optimal values of χ_{cut} and m_{jj} cut were found for all five simulated values of Λ (3, 5, 10, 15 and 20 TeV).

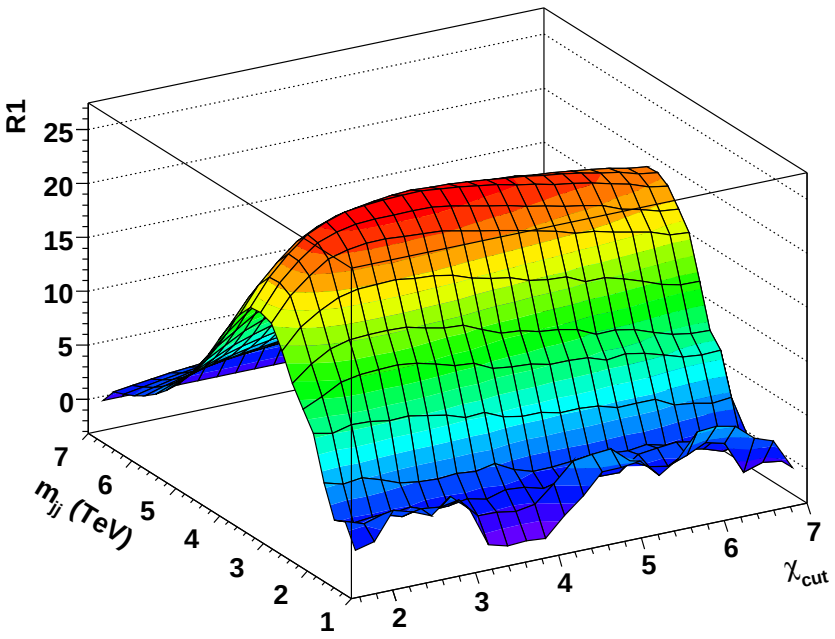


Figure 8.9: χ-cut and m_{jj}-cut $R1$ dependence for $\Lambda = 10$ TeV. It is apparent, that the value of $R1$ near its maximum depends only a little on both cuts.

In order to be able to calculate $R1$ for any value of Λ, the optimal values of χ_{cut} and m_{jj} cut for simulated Λ were fitted by suitable smooth functions, as presented in Fig. 8.10. The resulting fit functions are used as a new definition of the χ-cut and m_{jj}-cut.

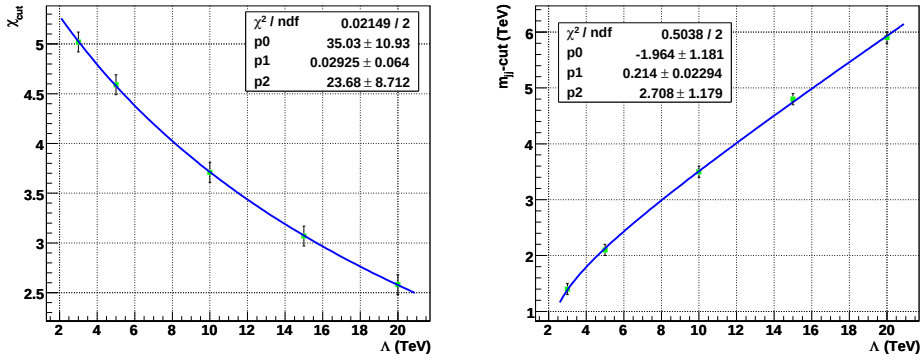


Figure 8.10: Dependence of optimal χ-cut on Λ (left). Dependence of optimal m_{jj}-cut on Λ (right). The resulting fitting functions are used as new definitions of χ-cuts and m_{jj}-cuts.

χ-cut is now *defined* according to the formula:

$$\chi_{\text{cut}} = 35.03 + 0.02924 \times \Lambda(\text{TeV}) - 23.67 \times (\Lambda(\text{TeV}) + 8\ \text{TeV})^{0.1} \qquad (8.4)$$

m_{jj}-cut is now *defined* according to the formula:

$$m_{jj\text{cut}}(\text{GeV}) = -1.9640 + 0.2140 \times \Lambda(\text{TeV}) + 2.7076 \times (\Lambda(\text{TeV}) - 2\ \text{TeV})^{0.1} \qquad (8.5)$$

One might wonder, which value of Λ shall we use for the calculation of χ-cut and m_{jj}-cut when the value of Λ is the one we want to measure and we do not know that value prior to the experimental measurement. Therefore, we must calculate $R1$ for all possible cases of Λ and corresponding cuts. The highest value of $R1$ will be reached with χ-cut and m_{jj}-cut optimal for the Λ which is experimentally observed. This procedure is also illustrated in Sec. 8.3.2.

Due to the range of χ and m_{jj} cuts it was necessary to make a dedicated Λ-fit for the needs of R1. The following binning was chosen. Let us have 44 bins in Λ from 3 to 25 TeV (bin width of 0.5 TeV). According to Eqs. (8.4) – (8.5) we obtain corresponding values of χ and m_{jj} cut for each Λ bin edge. For each m_{jj} bin, the χ spectrum was divided into two bins: $(1; \chi_{cut})$ and $(\chi_{cut}; 35)$. These bins were then fitted. The largest relative difference between the fitted data and the fitting function was 0.3%. Thanks to this $R1$ dedicated Λ-fit and Eqs. (8.4) – (8.5) the $R1 = R1(\Lambda)$ dependence is calculated as shown in Fig. 8.11.

8.3.1 Exclusion limits

When calculating $R1$ exclusion limits, the following factors and systematic errors are included: k-factors (see Fig. 8.12) and ll2nn factors, pdf uncertainties, JES uncertainties, FA factor and its uncertainties. The uncertainty of absolute luminosity measurement cancels in $R1$ by definition. A comparison of relative $R1$ systematic errors is in Fig. 8.13. In this case all sources of systematic errors are similarly important.

$R1$ exclusion limits are depicted in Fig. 8.14 and listed in Tab. 8.1.

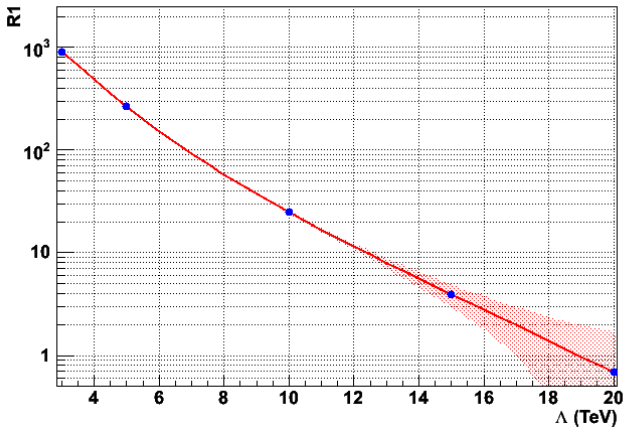


Figure 8.11: $R1$ dependence on Λ at 30 fb^{-1}. CTEQ6L1. The red band shows statistical error.

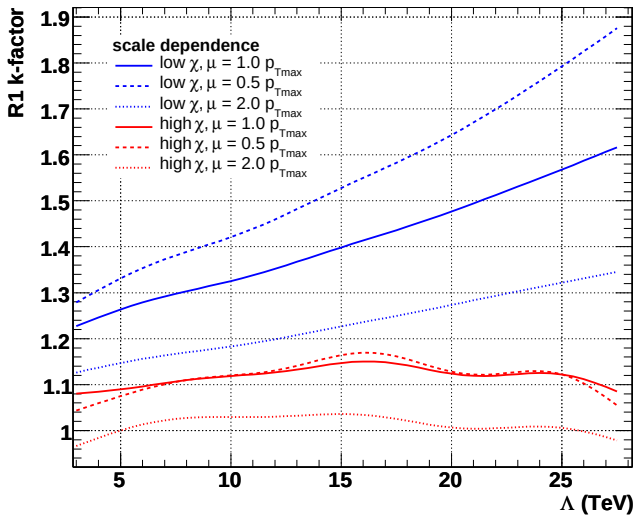


Figure 8.12: $R1$ k-factors and their scale dependence calculated in NLOJET++. Low χ-bin contains χ in interval $(1;\chi_{cut})$, high χ-bin contains χ in interval $(\chi_{cut};35)$.

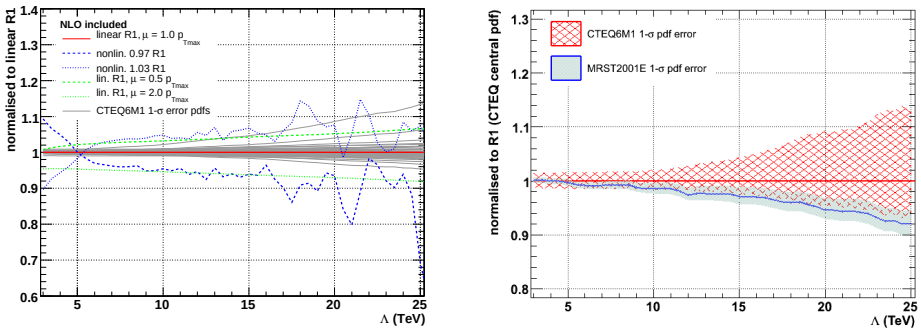


Figure 8.13: $R1$ relative systematic error comparison (left). It is apparent, that the CTEQ6M1 pdf uncertainties is one of the largest contributions to the systematic error. $R1$ pdf master error comparison as obtained with CTEQ6M1 and MRST2001E pdf sets (right). The pdf errors were divided by a factor of 1.64 in order to scale down from original 90% C.L. to a 1-sigma level.

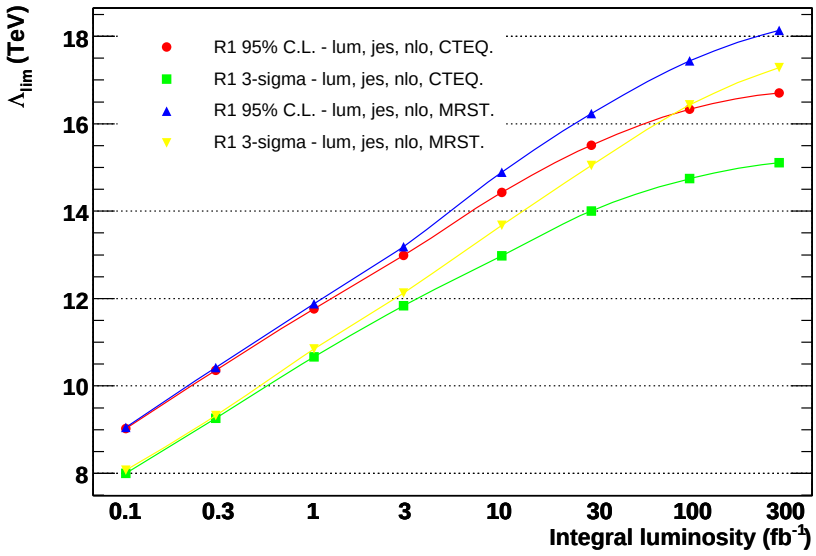


Figure 8.14: $R1$ exclusion limits with CTEQ6M1 and MRST2001E pdf sets.

Int. luminosity	95% C.L.		3-sigma	
(fb^{-1})	MRST	CTEQ	MRST	CTEQ
0.1	9.05	9.03	8.07	8.00
0.3	10.43	10.36	9.32	9.26
1.0	11.88	11.77	10.85	10.67
3.0	13.19	12.98	12.13	11.84
10.	14.90	14.43	13.68	12.97
30.	16.24	15.51	15.05	14.01
100.	17.43	16.33	16.43	14.74
300.	18.14	16.70	17.29	15.11

Table 8.1: Values of exclusion limits Λ_{lim} (TeV) at 95% C.L. and 3-sigma level obtained by $R1$ method with CTEQ6M1 or MRST2001E pdf uncertainties.

8.3.2 Discovery potential

$R1$ discovery potential at 300 fb^{-1} is illustrated in Fig. 8.15. When we calculate $R1$, then for a given value of Λ we know the best value of these cuts according to Eqs. (8.4) – (8.5). The solid blue line in the Fig. 8.15 corresponds to our prediction of the highest $R1$ values for Λ given by the horizontal axis. When we measure χ spectrum by the experiment, corresponding to some fixed value $\Lambda = \Lambda_{exp}$, we do not know the value of Λ_{exp} in advance and so we do not know, what values of χ_{cut} and m_{jj} to choose. Therefore we calculate the $R1$ curve with values of these cuts for any meaningful value of Λ. An example of such a curve for $\Lambda_{exp} = 10$ TeV and $\Lambda_{exp} = 15$ TeV is also shown in Fig. 8.15. In these cases, Λ given by the horizontal axis influences only the values of cuts, but the angular spectra are fixed. The intersections of these "experimental" curves with our $R1$ prediction indicate the value Λ_{exp} that was eventually measured. The highest value of $R1$ is obtained with m_{jj}-cut and χ-cut optimal for the measured Λ_{exp}.

5-sigma discovery intervals are listed in Tab. 8.2. It is also interesting to compare these intervals with corresponding values (obtained with prior flat in $1/\Lambda^2$) in Tabs. 7.3 – 7.4 obtained with p_{T}-Bayes method. As the p_{T}-Bayes method is more sensitive to quark compositeness, its discovery intervals are also narrower.

Λ (TeV)	3-sigma	5-sigma
10	9.18 to 11.02	8.64 to 11.75
15	12.95 to 18.46	11.85 to 21.24

Table 8.2: 3 and 5-sigma discovery intervals of "measured" Λ_{exp} (TeV) at 300 fb^{-1} obtained by R1 method with CTEQ6M1 pdf uncertainties for various values of Λ implemented in the "measured data".

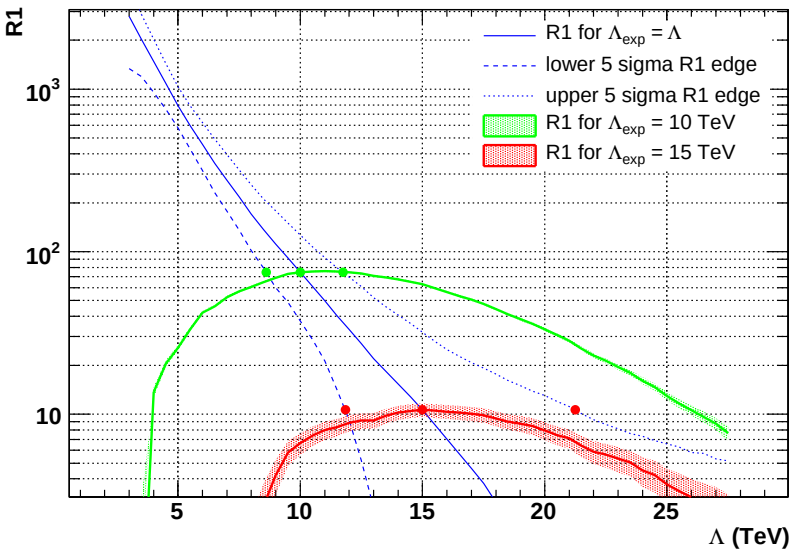


Figure 8.15: $R1$ 5-sigma discovery potential at 300 fb^{-1}, CTEQ6M1. Three green (red) markers show 5-sigma interval for $\Lambda = 10$ TeV ($\Lambda = 15$ TeV). Green and Red band correspond to statistical errors of the "experimental" data. See text for details.

8.4 Bayesian approach

The χ-Bayes method is based on a comparison of measured and predicted χ-spectra obtained with a given lower cut on dijet invariant mass m_{jj}. Analogically to p_{T}-Bayes method, the χ-Bayes posterior probability density function is calculated as:

$$P(\Lambda|d) = \frac{1}{Z} P(\Lambda) \prod_{k=1}^{10} P^k(N_d^k | N_{exp}^{\Lambda,k}) \otimes G_k(JES(\Lambda), pdf(\Lambda), \delta L, \mu, FA). \quad (8.6)$$

The only difference is, that the product on the right-hand side of Eq. (8.6) runs over a fixed number of 10 χ-bins ranging from 1 to 35. The χ-spectra were obtained with a fixed lower cut on dijet invariant mass $m_{jj} > 5.9$ TeV. The meaning of the symbols in Eq. (8.6) is the same as o those in (7.3), except for the fact, that now we deal with a given χ-spectra.

Although the m_{jj} cut strongly reduces the statistics, the higher it is, the more quark compositeness is pronounced and observable, as shown in Fig. 8.16. In this figure, the NLO k-factors and scale dependence errors are not included, qualitatively they would not change the results. In case of lower m_{jj} cuts (especially 1.4, 2.1 and 3.5 TeV) the dependence of exclusion limits on integral luminosity is almost flat. This is caused by systematic errors, overwhelming the indistinctive effect of quark compositeness in this kinematic region.

One might wonder, why the values of m_{jj} cuts were chosen as stated in the legend of Fig. 8.16. These are the optimal values of m_{jj} cuts in $R1$ method for

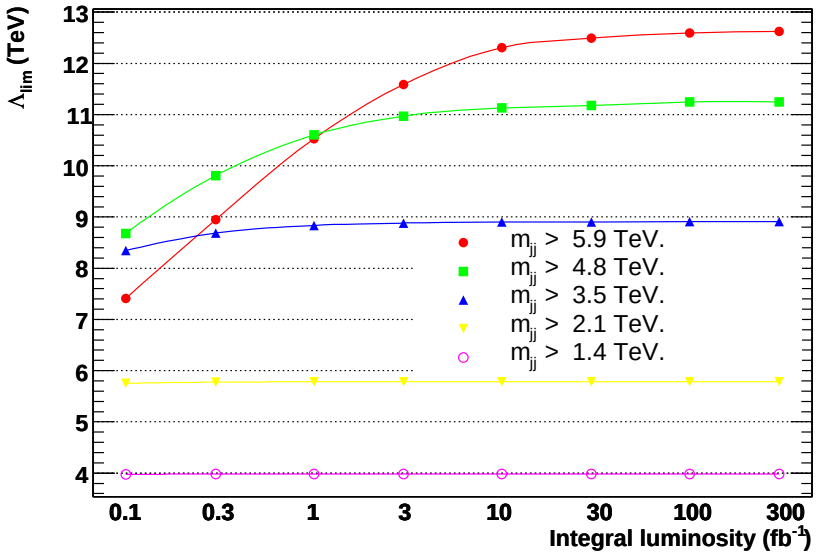


Figure 8.16: Dependence of χ-Bayes 3-sigma exclusion limits on m_{jj}-cuts and total integral luminosity. At higher integral luminosities, the highest exclusion limits are obtained with the highest m_{jj}-cut. In case of lower integral luminosities, this is not completely the case, as the statistical error plays a larger role and the higher m_{jj}-cut reduces the statistics significantly. CTEQ6M1 pdfs used. NLO k-factors are not included.

individual simulated data sets (with $\Lambda = 3$, 5, 10, 15 and 20 TeV), so they were chosen as a good reference to start with. Any other values of m_{jj} cuts would do the similar job.

8.4.1 Exclusion limits

The χ-Bayes method exclusion limits with $m_{jj} > 5.9$ TeV are depicted in Fig. 8.17 and listed in Tab. 8.3.

To be complete, the discovery limits were not calculated with this method, as it is not as sensitive as the $R1$ method.

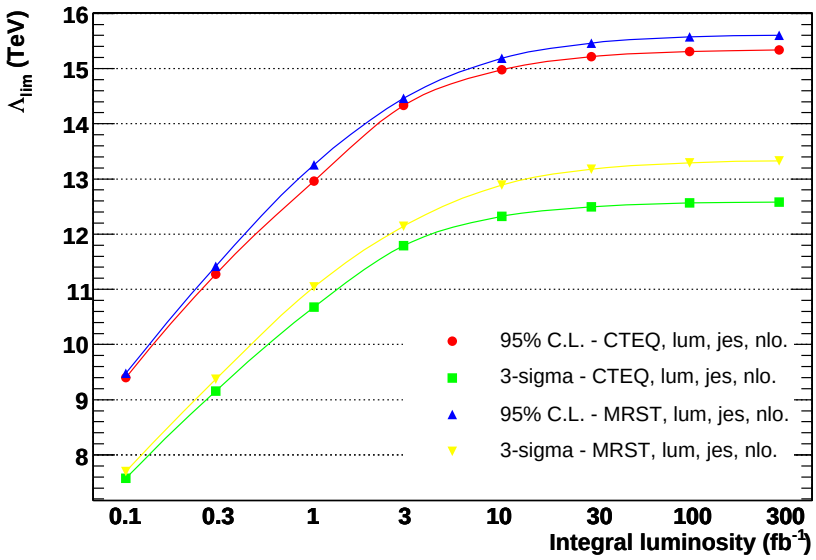


Figure 8.17: Dependence of χ-Bayes exclusion limits Λ_{lim} (TeV) on total integral luminosity.

Int. luminosity	95% C.L.		3-sigma	
(fb^{-1})	MRST	CTEQ	MRST	CTEQ
0.1	9.48	9.40	7.70	7.57
0.3	11.42	11.27	9.39	9.16
1.0	13.26	12.96	11.04	10.68
3.0	14.46	14.33	12.14	11.79
10.	15.18	14.98	12.89	12.32
30.	15.46	15.21	13.18	12.49
100.	15.57	15.31	13.29	12.56
300.	15.60	15.33	13.33	12.58

Table 8.3: Values of exclusion Λ_{lim} (TeV) at 95% C.L. and 3-sigma level obtained by χ-Bayes method with CTEQ6M1 or MRST2001E pdfs for various values of total integral luminosity.

Chapter 9

Summary of results

In this chapter the results obtained with all four analysis methods are briefly summarised and compared. The value of jet energy scale uncertainty involved is 3% unless stated otherwise.

9.1 Exclusion limits

Exclusion limits for quark compositeness computed with CTEQ6M1 pdfs are shown in Fig. 9.1 and listed in Tab. 9.1. Exclusion limits for quark compositeness computed with MRST2001 pdfs are shown in Fig. 9.2 and listed in Tab. 9.2.

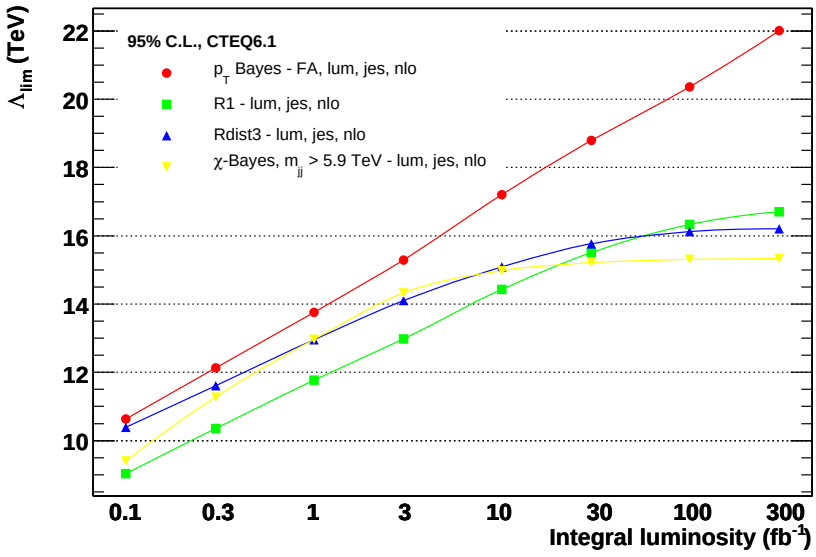


Figure 9.1: Dependence of exclusion limits of Λ_{lim} at 95% C.L. on total integral luminosity as obtained with all four analysis methods. CTEQ6M1 pdfs used.

The largest values of Λ_{lim} are obtained with p_{T}-Bayes method with prior flat in $1/\Lambda^2$, and 5% systematic error due to choice of prior. The largest values

Int. luminosity (fb^{-1})	p_T-Bayes	R1	Rdist3	χ-Bayes
0.1	10.64	9.03	10.39	9.40
0.3	12.13	10.36	11.61	11.27
1.0	13.75	11.77	12.96	12.76
3.0	15.29	12.98	14.10	14.33
10.	17.21	14.43	15.09	14.98
30.	18.80	15.51	15.77	15.21
100.	20.36	16.33	16.12	15.31
300.	22.01	16.70	16.21	15.33

Table 9.1: Values of exclusion limits Λ_{lim} (TeV) at 95% C.L. for various methods and integral luminosities with CTEQ6M1.

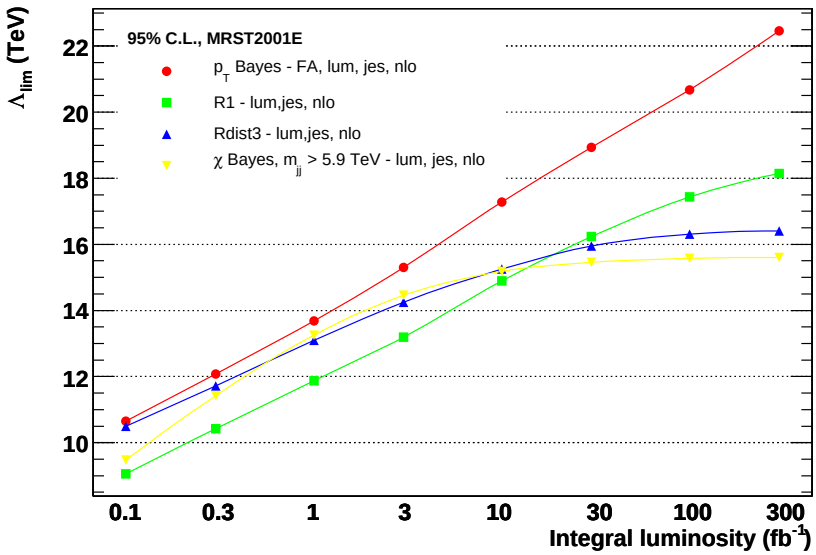


Figure 9.2: Dependence of exclusion limits Λ_{lim} at 95% C.L. on total integral luminosity as obtained with all four analysis methods. MRST2001E pdfs used.

Int. luminosity (fb^{-1})	p_T-Bayes	R1	Rdist3	χ-Bayes
0.1	10.65	9.05	10.49	9.48
0.3	12.08	10.43	11.72	11.42
1.0	13.68	11.88	13.09	13.26
3.0	15.30	13.19	14.25	14.46
10.	17.28	14.90	15.25	15.18
30.	18.94	16.24	15.95	15.46
100.	20.68	17.43	16.31	15.57
300.	22.46	18.14	16.40	15.60

Table 9.2: Values of exclusion limits Λ_{lim} (TeV) at 95% C.L. for various methods and integral luminosities with MRST2001E.

of Λ_{lim} based on angular distribution analysis are obtained with $R1$ method. Exclusion limits computed with MRST2001E pdfs are slightly higher than those computed with CTEQ6M1. This is caused by a prediction of smaller systematic pdf error in case of MRST, nevertheless that does not mean, that MRST limits are better. We have calculated only predictions based on nowadays knowledge of pdfs, this knowledge will hopefully evolve to a higher precision during running the ATLAS experiment and so our compositeness exclusion limits might still increase a little. The crucial influence on the exclusion limits follows from our knowledge of jet energy scale.

9.2 Early data exclusion limits

In this section the exclusion limits on quark compositeness using our most effective method p_{T}-Bayes and involving foreseen early data JES uncertainty (7% for $|\eta| < 3.2$, see Tab. 5.1) are listed. See Tab. 9.3 for details. Compared to results presented in previous Sec. 9.1, the early data exclusion limits involving a higher JES uncertainty are naturally lower. However, this difference is not very large due to the fact, that the statistical error for early data plays a larger role, compared to the systematic error. Concerning the systematic error itself, the largest contribution is now caused by JES uncertainty, therefore the difference between exclusion limits calculated using CTEQ or MRST pdfs is almost negligible.

Int. luminosity	p_{T}-Bayes	
(fb^{-1})	MRST2001E	CTEQ6M1
0.1	10.340	10.334
0.3	11.809	11.800
1.0	13.403	13.401

Table 9.3: Values of exclusion limits Λ_{lim} (TeV) at 95% C.L. calculated using p_{T}-Bayes method. Expected early data JES uncertainty (see Tab. 5.1) involved.

9.3 Discovery potential

The discovery potential – the ability of our methods to measure quark compositeness in a situation when it is discovered – was studied with p_{T}-Bayes and $R1$ method for compositeness scales $\Lambda = 10$ TeV and $\Lambda = 15$ TeV. Both these methods are the most promising ones for analysis of inclusive dijet p_{T}-spectra and dijet angular distribution, respectively. It should be stressed, that if quark compositeness is measured, it must be measured in both p_{T} and χ-spectra or even in its two dimensional (p_{T}, χ) combination. There is a whole spectrum of new physics predicted to be measured at ATLAS (and there might always be something unpredicted) that might fake quark compositeness either in p_{T} or in χ-spectra. The discovery intervals are listed in Tab. 9.4. As p_{T}-Bayes method is more sensitive to compositeness, its discovery intervals are also narrower, but

both methods are independent and both should be used and checked, whether they give compatible results.

Λ (TeV)	10	15
p_{T}-Bayes ($1/\Lambda^2$)	9.38 to 10.93	13.30 to 18.51
$R1$	8.64 to 11.75	11.85 to 21.24

Table 9.4: The shortest 5-sigma intervals for p_{T}-Bayes method (with prior flat in $1/\Lambda^2$) and $R1$ method with compositeness scales $\Lambda = 10$ TeV and $\Lambda = 15$ TeV switched on. CTEQ6M1 pdfs, 300 fb^{-1}.

Chapter 10

Conclusions

In this paper the potential of the ATLAS detector to exclude quark compositeness or to discover it was studied in great detail. There were four different analysis methods developed for this purpose. Two of them use Bayesian technique (as is usual in similar analyses at Tevatron) and a detailed information about inclusive dijet p_{T} spectrum or dijet angular distribution. The other two "ratio" methods use the ratio of events below and above a given cut in the investigated spectra. It was newly shown, that the value of this cut should be optimised in order to obtain the best sensitivity to quark compositeness. Moreover, larger number of systematic errors was taken into account in mentioned analysis methods and their influence was studied too. It turned out, that even the "ratio" methods are affected by the systematic errors. The most sensitive method for quark compositeness measurement is the p_{T}-Bayes method, the most sensitive method using angular distribution is the $R1$ method. Both methods should be used in order to ensure, that the results correspond indeed to quark compositeness and not to some other kind of new physics.

As a very important spin-off of this study, two bugs in PYTHIA Monte Carlo generator were identified and later corrected for in the official PYTHIA code.

It was also shown, that the early data limits (including expected early data jet energy scale uncertainty) on quark compositeness reachable by ATLAS are much higher than those currently known from Tevatron. The highest limit on quark compositeness with expected total amount of data to be collected by the ATLAS experiment is above 22 TeV, assuming 3% jet energy scale uncertainty.

Finally, this paper brings new, more detailed insight into quark compositeness studies at the ATLAS detector. Very sensitive analysis methods were developed and tested, including systematic errors, and expected limits on quark compositeness and discovery potentials were calculated.

Appendix A

Monte Carlo generator settings

In this Appendix the complete settings of Monte Carlo generators PYTHIA and JIMMY/HERWIG used for simulation of analysed samples are listed. In case of production of CSC dijet samples (that is full simulation), these settings were completely defined by the ATLAS collaboration, including the underlying event settings. The settings of PYTHIA for ATLFAST mass production (that is almost all the data analysed in this study) were defined by the author.

The format of the settings is copied from the original jobOption files used for steering of jobs in Athena and written in PYTHON. However, its connection to FORTRAN format is straightforward.

A.1 CSC dijet samples J4 – J8 Pythia settings

```
Pythia.PythiaCommand = [
  "pysubs msel 0",
  "pysubs ckin 3 140.",
  "pysubs ckin 4 280.",
  "pysubs msub 11 1",
  "pysubs msub 12 1",
  "pysubs msub 13 1",
  "pysubs msub 68 1",
  "pysubs msub 28 1",
  "pysubs msub 53 1",
  "pypars mstp 82 4",
  "pyinit pylisti 12",
  "pyinit pylistf 1",
  "pystat 1 3 4 5",
  "pyinit dumpr 1 5"
]
```

The values of upper and lower p_{T}-cuts CKIN(3)=140 and CKIN(4)=280 correspond to J4 sample. Values of these cuts for other samples were set according

to Tab. 6.1. Additional underlying event settings:

```
Pythia.PythiaCommand += [
  "pypars mstp 70 2",
  "pypars mstp 72 0",
  "pypars mstp 81 21",
  "pypars mstp 82 4",
  "pypars mstp 84 1",
  "pypars mstp 85 1",
  "pypars mstp 86 2",
  "pypars mstp 87 4",
  "pypars mstp 88 0",
  "pypars mstp 89 1",
  "pypars mstp 90 1",
  "pypars mstp 95 1",
  "pypars parp 78 0.2",
  "pypars parp 80 0.01",
  "pypars parp 82 1.9",
  "pypars parp 83 0.3",
  "pypars parp 84 0.5",
  "pypars parp 89 1800",
  "pypars parp 90 0.22",
  "pydat1 parj 81 0.14"
]
```

A.2 CSC dijet samples J4 – J7 Jimmy/Herwig settings

```
Jimmy.JimmyCommand = [
  "iproc 11500",
  "modpdf 10042",
  "autpdf HWLHAPDF",
  "msflag 1",
  "jmbug 0",
  "jmueo 1",
  "ptjim 4.91",
  "jmrad 73 1.8",
  "pltcut 0.0000000000333",
  "prsof 0",
  "rmass 198 80.425",
  "rmass 199 80.425",
  "rmass 200 91.19",
  "gamw 2.124",
  "gamz 2.495",
  "clpow 1.20",
  "ptmin 140.",
```

```
  "ptmax 280."
]
```

The values of upper and lower p_{T}-cuts `ptmin=140` and `ptmax=280` correspond to J4 sample. Values of these cuts for other samples were set according to Tab. 6.1.

A.3 Pythia settings for ATLFAST mass production

```
Pythia.PythiaCommand = [
  "pysubs msel 51",        # compositeness processes
  "pysubs ckin 3 350.",    # lower pT cut
  "pysubs ckin 4 600.",    # upper pT cut
  "pydat1 mstu 54 1",
  "pydat1 paru 53 10.",
  "pytcsm rtcm 41 15000.", # Lambda (GeV)
  "pytcsm rtcm 42 1",      # interference sign
  "pytcsm itcm 5 2",       # 2-on,1-ud,0-off
  "pypars mstp 82 4",      # Rome settings
  "pydat1 mstj 11 3",      # Rome settings
  "pydat1 mstj 22 2",      # Rome settings
  "pydat1 parj 54 -0.07",  # Rome settings
  "pydat1 parj 55 -0.006", # Rome settings
  "pypars parp 82 1.8",    # Rome settings
  "pypars parp 84 0.5"     # Rome settings
]
```

An example of Pythia settings in jobOptions file with a lower p_{T}-cut of 350 GeV and an upper p_{T}-cut of 600 GeV. In this case the value of compositeness scale $\Lambda = 15000$ GeV, compositeness is switched on for all quarks (ITCM(5)=2) with positive interference sign (RTCM(42)=1). The so-called Rome settings were recommended by the ATLAS collaboration as a tune of Pythia during Rome production, which was a predecessor of CSC production. So the CSC settings are more up-to-date than Rome settings, but a correction for that is taken into account in FA factors, where possible – that is in case of inclusive dijet p_{T} spectra.

Bibliography

[1] R.K. Ellis, W.J.Stirling, B.R. Webber: QCD and Collider Physics, Cambridge University Press, 1996.

[2] J. Chýla: Quarks, Partons and Quantum Chromodynamics, a study text, available at: http://www-hep2.fzu.cz/Centrum/uc_texty.html, 2003.

[3] J.W. Huston, J.M. Campbell, W.J. Stirling: A Primer for LHC Physics, ATL-COM-PHYS-2006-077, hep-ph/0610012, 2006.

[4] J. Pumplin, D.R. Stump, J. Huston, H.L. Lai, W. Tung, S. Kuhlmann, J.F. Owens: Inclusive Jet Production, Parton Distributions, and the Search for New Physics, hep-ph/0303013, 2003.

[5] J. Pumplin, D.R. Stump, J. Huston, H.L. Lai, P. Nadolsky, W.K. Tung: New Generation of Parton Distributions with Uncertainties from Global QCD Analysis, hep-ph/0201195, 2002.

[6] A.D. Martin, R.G. Roberts, W.J. Stirling, R.S. Thorne: Uncertainties of predictions from parton distributions. I: experimental errors, hep-ph/0211080, 2002.

[7] L. N. Lipatov, Sov. J. Nucl. Phys. **20** (1975) 95;
V. N. Gribov and L. N. Lipatov, Sov. J.Nucl. Phys. **15** (1972) 438;
G. Altarelli and G. Parisi, Nucl. Phys. **B126** (1977) 298;
Yu. L. Dokshitzer, Sov. Phys. JETP **46** (1977) 641.

[8] S. Gieseke, JHEP **01** 058 (2005), hep-ph/0412342.

[9] D. Bourilkov, R.C. Group, and M.R. Whalley: LHAPDF: PDF Use from the Tevatron to the LHC, hep-ph/0605240, 2006.

[10] E.J. Eichten, K.D. Lane, and M.E. Peskin, Phys. Rev. Lett. **50**, 811 (1983).

[11] P. Chiappetta and M. Perrottet, in 'Large Hadron Collider Workshop', eds. G. Jarlskog and D. Rein, CERN 90-10 (Geneva 1990), Vol. II, p806.

[12] E. Eichten, I. Hincliffe, K. Lane and C. Quigg, Rev. Mod. Phys 56 (1984), 579; Rev. Mod. Phys. 58 (1985), 1065.

[13] T. Sjöstrand, L. Lönnblad, S. Mrenna, P. Skands: Pythia 6.3 Manual, hep-ph/0308153, August 2003.

[14] T. Sjöstrand, P. Edén, C. Friberg, L. Lönblad, G. Miu, S. Mrenna and E. Norrbin, Computer Physics Commun. **135** (2001) 238.

[15] T. Davídek - private communication, Prague.

[16] T. Sjöstrand - private communication.

[17] L. Přibyl: Studies on quark compositeness in ATLAS, WDS, June 2004, Prague.

[18] L. Přibyl: Quark compositeness in ATLAS, LHC Days in Split, October 2006, Split.

[19] H. Harari, N. Seidberg: The rishon model, Nuclear Physics B204 (1982) 141-167.

[20] J.C. Pati, A. Salam, Phys. Rev. **D10**, 275 (1974).

[21] S.O. Bilson-Thompson: A topological model of composite preons, hep-ph/0503213, March 2005.

[22] H. Jeffreys, *Theory of probability* [Claredon Press, Oxford, 1939 (revised 1988)], p. 94.

[23] S. Jain, A.K. Gupta, N.K. Mondal: Search for quark-lepton Compositeness at Tevatron and LHC, hep-ex/0005025v1, May 2000.

[24] A. Kupčo: Compositeness limits from dijet angular distribution in RunIIa, DØ Collaboration Meeting, September 2001.

[25] B. Abbott *et al.*: Dijet Mass Spectrum and a Search for Quark Compositeness in $\bar{p}p$ Collisions at $\sqrt{s} = 1.8$ TeV, Phys. Rev. Lett., Vol. 82, 12, March 1999.

[26] ATLAS Collaboration: *ATLAS Detector and Physics Performance*, CERN/LHCC/99-15, 1999.

[27] Z.U. Usubov: A search for quark compositeness at the LHC, ATLAS Internal Note ATL-COM-PHYS-99-052 (1999).

[28] L. Přibyl: Quark compositeness in ATLAS, February 2006, RTN Workshop, MFF UK, Praha.

[29] ATLAS Collaboration: ATLAS Technical proposal for a General-Purpose pp Experiment at the Large Hadron Collider, CERN/LHCC/94-43, 1994.

[30] ATLAS Collaboration: ATLAS Detector and Physics Performance, CERN/LHCC/99-14, 1999.

[31] ATLAS Collaboration: ATLAS Technical paper, under preparation, 2007.

[32] CMS Collaboration: The Compact Muon Solenoid Techncal Proposal, CERN/LHCC/94-38, 1994.

[33] ALICE Collaboration: A Large Ion Collider Experiment Technical Proposal, CERN/LHCC/95-71, 1995.

[34] LHCb Collaboration: LHCb Technical Proposal, CERN LHCC 98-4, LHCC/P4, 1998.

[35] LHC Collaboration: LHC Design Report, available at http://ab-div.web.cern.ch/ab-div/Publications/LHC-DesignReport.html, CERN, 2004.

[36] ATLAS Collaboration: ATLAS Inner Detector Technical Design Report, CERN/LHCC/97-16, 1997.

[37] ATLAS Collaboration: ATLAS Muon Spectrometer Technical Design Report, CERN/LHCC/97-22, 1997.

[38] ATLAS Collaboration: ATLAS Magnet System Technical Design Report, CERN/LHCC/97-18, 1997.

[39] ATLAS Collaboration: ATLAS High-Level Triggers, DAQ and DCS Technical Proposal, CERN/LHCC/2000-17, 2000.

[40] ATLAS Collaboration: Luminosity Letter of Intent, CERN/LHCC/2004-10, 2004.

[41] G. Atoian, A. Denisov, V. Isaakov, A. Poblaguev, S. White, M. Zeller: A Zero Degree Calorimeter for ATLAS, CERN-ATL-COM-PHYS-2006-014, 2006.

[42] ATLAS Collaboration: ATLAS Computing Technical Design Report, CERN/LHCC/2005-022, 2005.

[43] M. Aderholz *et al.*: Models of networked analysis at regional centres for LHC experiments (MONARC). Phase 2 report (2000), CERN/LCB/2001-001, 2000.

[44] B. Dowler *et al.*: Performance of the ATLAS hadronic end-cap calorimeter in beam tests, Nucl. Instr. and Meth., A482:94, 2002.

[45] ATLAS Collaboration: Liquid Argon Calorimeter, CERN/LHCC 96-41, CERN, 1996.

[46] ATLAS Collaboration: *ATLAS Tile Calorimeter TDR*, CERN/LHCC/96-42, 1996.

[47] V. Giangiobbe, P. Johansson, K. Jon-And, C. Santoni: Hadronic calorimeter performance in the ATLAS combined testbeam 2004, ATLAS note, ATLAS-TILECAL-PUB-2005-008, 2005.

[48] The ATLAS collaboration: Athena framework, https://twiki.cern.ch/twiki/bin/view/Atlas/AthenaFramework.

[49] GEANT4 Home Page: http://geant4.web.cern.ch/geant4/.

[50] GEANT4 Physics Reference Manual: http://geant4.web.cern.ch/geant4/G4UsersDocuments/UsersGuides/PhysicsReferenceManual/html/PhysicsReferenceManual.html

[51] E. Richter-Was, D. Froidevaux and L. Poggioli: ATLFAST, A fast simulation package for ATLAS, ATLAS note, ATL-PHYS-98-131, 1998.

[52] The Atlfast Validation Taskforce: Performance of the ATLAS fast simulation ATLFAST, ATLAS note, ATL-PHYS-INT-2007-005, 2007.

[53] P. Brückman and E. Richter-Was: How well can we reconstruct continuum background → b-jet calibration for the ATLFAST Fortran framework, ATLAS note, ATL-PHYS-2002-025.

[54] Asai, Set al.: Physics Analysis Tools 2006 Workshop Summary Report, ATL-COM-SOFT-2006-009, CERN. The ATLAS Collaboration: AOD Class Summary, https://twiki.cern.ch/twiki/bin/view/Atlas/AODClassSummary.

[55] G.C. Blazey *et al.*: Run 2 Jet Physics, arXiv:hep-ex/0005012 v2, 2002.

[56] Z. Nagy: NLOJET++ Home page, http://nagyz.web.cern.ch/nagyz/Site/NLOJET++/NLOJET++.html.

[57] Z. Nagy: Next-to-leading order calculation of three-jet observables in hadron-hadron collision, arXiv:hep-ph/0307268.

[58] D.S. Sivia: A Bayesian tutorial, Claredon press, Oxford, 1996.

[59] M. Lefebvre, P. Loch: Introduction to Hadronic calibration in ATLAS, 3$^{\text{rd}}$ Hadronic Calibration Workshop, Milan, Italy, 26-27 April, 2007.

[60] G.C. Blazey *et al.*: Run II Jet Physics, arXiv:hep-ex/005012v2, May 2000.

[61] M. Cacciari and G. Salam, Dispelling the N**3 myth for the k(t) jet-finder, Phys. Lett. **B641** (2006) 5761.

[62] B. Andrieu *et al.*: Results from pion calibration runs for the H1 liquid argon calorimeter and comparisons with simulations, 1993.

[63] C. Cojocaru *et al.*: Hadronic calibration of the ATLAS liquid argon endcap calorimeter in the pseudorapidity region $1.6 < |\eta| < 1.8$ in beam tests, Nucl. Inst. Meth. **A531** (2004) 481514.

[64] ATLAS Collaboration: The ATLAS Experiment at the CERN Large Hadron Collider, Draft 4, 2008.

[65] I. Abt *et al.* (H1 Collaboration): The tracking, calorimeter and muon detectors of the H1 experiment at HERA, Nucl. Inst. Meth. **A386** (1997) 348.

[66] F. Ruehr: Multi-jet Bootstrapping – JES calibration in the TeV regime, Exotic Monojets/Dijets group meeting, CERN, January 2008.

[67] F. Ruehr – private communication, January 2008.

[68] ATLAS Collaboration: ATLAS Detector and Physics Performance, ATLAS TDR14, CERN/LHCC 99-14, 1999, page 273.

[69] ATLAS Collaboration: https://twiki.cern.ch/twiki/bin/view/Atlas/CSCNotesList#Systematic_Uncertainties.

[70] ATLAS Collaboration: https://twiki.cern.ch/twiki/pub/Atlas/NewHiggsCSCPage/systematics_4.ppt.

[71] N. Palanque-Delabrouille,J. Rich: Overview of astroparticle physics and cosmology, LHC days in Split, October 2006.

[72] HERWIG 6.5, G. Corcella, I.G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M.H. Seymour and B.R. Webber, JHEP 0101 (2001) 010 [hep-ph/0011363]; hep-ph/0210213.

[73] J.M. Butterworth ,J.R. Forshaw, M.H. Seymour: Multiparton interactions in photoproduction at HERA, Z.Phys.C72:637-646, hep-ph/9601371, 1996.