Jets and Heavy Flavour at LHC with ATLAS and CMS

Anne-Marie Magnan on behalf of the ATLAS and CMS collaborations

Imperial College HEP Dept., Prince Consort Road, London SW7 2BW, UK

DOI: http://dx.doi.org/10.3204/DESY-PROC-2009-03/Magnan

The LHC experiments ATLAS and CMS plan to take advantage of large multi-jet samples with and without heavy flavour tagging and vector boson production to test QCD at the TeV scale. Initial multi-jet cross section measurements at LHC will demonstrate the understanding of the calibration of the detectors, the jet energy scale systematics and the trigger. Further in the LHC run, measurements of inclusive di-jet cross sections with heavy flavour tag, which provides the process hard scale, will probe QCD at scales never tested before. Jet production measurements with associated W and Z bosons provide a separate test of QCD in different and complementary channels. Measurements of these processes are essential to demonstrate the understanding of major backgrounds to Higgs and SUSY channels, such as those of top-quark production or W+jet/Z+jet.

1 Overview of ATLAS and CMS

A detailed description of the ATLAS (A Toroidal LHC ApparatuS) and CMS (Compact Muon Solenoid) experiments can be found respectively in [1] and [2]. ATLAS (CMS) has an overall length of 44 m (22 m), a diameter of 25 m (15 m), and weighs 7 000 tons (12 500 tons).

ATLAS is composed of a thin 2 T superconducting solenoid surrounding the inner-detector cavity, a high granularity liquid-argon (LAr) electromagnetic sampling calorimeter, followed by scintillator-tile/LAr hadronic calorimeters, three large superconducting toroids arranged with an eight-fold azimuthal symmetry around the calorimeters, and a muon spectrometer. The inner detector is made of semiconductor pixel and strip detectors, surrounded by straw-tube tracking detectors with the capability to generate and detect transition radiation. LAr forward calorimeters extend the pseudo-rapidity coverage from $|\eta| > 3$ to $|\eta| < 4.9$.

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter. Within the field volume are the silicon pixel and strip tracker, the lead-tungstate crystal electromagnetic calorimeter (ECAL) and the brass-scintillator hadronic calorimeter (HCAL). Muons are measured in gas chambers embedded in the iron return yoke. CMS also has extensive forward calorimetry, extending the pseudo-rapidity coverage of the calorimeters from $|\eta| > 3$ to $|\eta| < 5$.

In ATLAS (CMS), the ECAL has an energy resolution of about 1 % (0.5 %) at 100 GeV, and represents 22 to 26 X_0 (24.7 to 25.8 X_0). The HCAL, when combined with the ECAL, measures jets with a resolution $\Delta E/E \approx 50\,\%/\sqrt{E} \oplus 3\,\%$ ($\approx 100\,\%/\sqrt{E} \oplus 5\,\%$). The calorimeter cells are grouped in projective towers, of granularity $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ (0.087×0.087) at central rapidity and 0.2×0.1 (0.175×0.175) at forward rapidity. The resolution in the ATLAS (CMS)

tracker is expected to be $\sigma/p_T \approx 5 \times 10^{-5} \times p_T \oplus 0.01$ ($\approx 1.5 \times 10^{-5} \times p_T \oplus 0.005$). Both apparatuses provide the vertex position with $\approx 100\,\mu$m accuracy for 1 GeV tracks, and below $20\,\mu$m accuracy for tracks above 20 GeV.

2 QCD and the LHC

QCD processes constitute the dominant source of interactions at the LHC due to their large cross sections relative to other processes, as detailed in Tab. 1 [3]. This makes QCD an attractive topic for early physics at LHC. By measuring jets, several objectives can be attained, both from theoretical and experimental points of view: commissioning of the detectors, confrontation of perturbative QCD (pQCD) at the TeV scale, tests of PDF evolution schemes, probes of α_S, understanding of multi-jet production (background to other searches), sensitivity to new physics.

Process	σ (nb)
Total	10^8
$W^{\pm} \to e\nu$	20
$Z \to e^+e^-$	2
$t\bar{t}$	0.8
$b\bar{b}$	5×10^5
$c\bar{c}$	10^7
central jets	
$p_T > 10$ GeV	2.5×10^6
$p_T > 100$ GeV	10^3
$p_T > 1000$ GeV	1.5×10^{-3}

Table 1: Cross-sections expected at the LHC for a few processes, at $\sqrt{s} = 10$ TeV.

The number of jets per bin in transverse momentum p_T, for a centre of mass energy of 10 TeV expected at start-up of the LHC, is shown in Fig. 1, for different ranges in rapidity y. With only 10 pb^{-1} of integrated luminosity, several tens of events are still expected with jets above 1 TeV, and so early measurements are possible for a large range in energy.

3 Jets at the LHC

3.1 Definition of a jet

From a theoretical point of view, a so-called *parton jet* originates from the proton-proton collision, and should contain the partons produced and the particles from initial- and final-state radiation (ISR/FSR). From an experimental point of view, a parton jet then undergoes hadronisation (decays, or interactions in the beam pipe/tracker material), after which point it can be reconstructed as a *particle jet* if individual particles are identified (so called *particle flow* algorithms). Electromagnetic and hadronic components will finally shower in the calorimeters, so that pure *calorimeter jets* can be reconstructed. Two types of algorithms exist: cone-based and sequential recombination. Cone-based can be seeded (at LHC, iterative, with sizes $\Delta R = 0.4$ (CMS 0.5) and 0.7), in which case they are not infrared- or collinear-safe, but are fast and reliable for triggering, or seedless (the Seedless Infrared Safe - SIScone - algorithm in CMS). The sequential recombination algorithm k_T is used in both ATLAS and CMS with sizes 0.4 and 0.6. To compare jets at each step, the same jet reconstruction algorithm should be employed. Inputs to the algorithms are hence either calorimetric energy depositions (towers or clusters), tracks, particle or energy flow reconstructed objects, simulated or generated particles.

The particle content of a jet is shown in Fig. 2 [4], and is independent of the jet transverse momentum, as expected since jet fragmentation functions are independent of the energy. Charged particles will carry 65% of the energy, hence use should be made of the good tracker resolution of both detectors. Photons will carry 25% of the energy, and the excellent EM

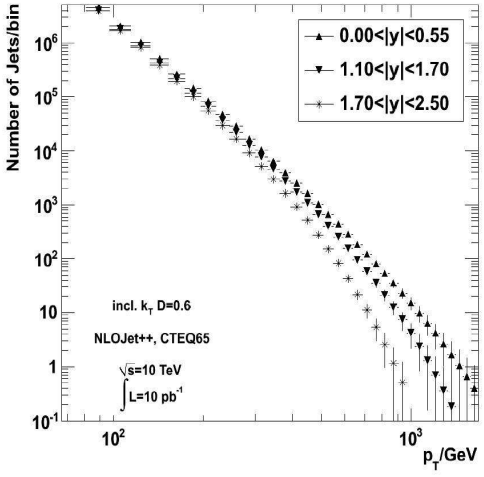


Figure 1: Expected number of jets per bin in p_T for $10\,\mathrm{pb}^{-1}$ of integrated luminosity at $\sqrt{s} = 10\,\mathrm{TeV}$.

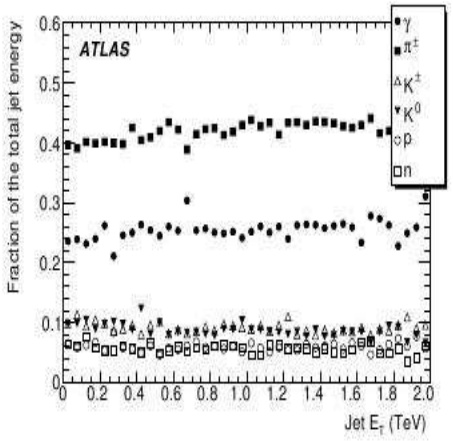


Figure 2: Particle content of a jet as a function of its transverse energy.

calorimeter resolution should help significantly in the overall jet energy resolution. Neutral particles will carry the remaining 10% of the energy, and represent the limiting factor to jet energy resolution.

3.2 Jet energy scale and jet energy resolution

In ATLAS, the jet energy scale is obtained by a calibration procedure described in detail in [4]. Several methods are used in order to improve the jet energy resolution. One of them involves using the track content of a jet. The method is illustrated in Fig. 3, left. The overall jet response is centred on the expected energy, but different bins in the fraction $f_{\mathrm{trk}} = \frac{p_T^{\mathrm{tracks}}}{p_T^{\mathrm{calo}}}$ show different central values for the response, leading to an artificially larger spread in the overall response. By correcting the energy as a function of f_{trk}, the jet energy resolution can be improved by $\approx 10\%$ at 40 GeV, as illustrated in Fig. 3, right, leaving the overall jet response unchanged.

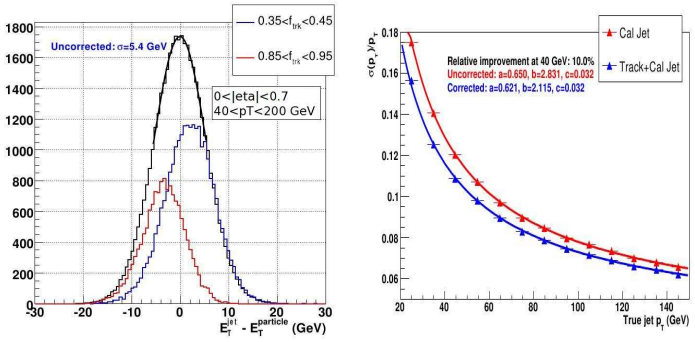


Figure 3: Track-based correction procedure in ATLAS: jet energy scale for different fractions of energy carried by the tracks associated to the jet (left) and jet energy resolution before and after corrections (right).

In CMS, the jet energy calibration uses a factorised approach, after which the jet response for calorimeter jets is flat in transverse momentum and pseudo-rapidity [5]. By using a more complete reconstruction of the events with a particle flow (PF) algorithm, making use of both iterative tracking and calorimeter clustering using calibrated clusters, it is possible to improve greatly the jet energy resolution [6]. The jet response before any correction is shown in Fig. 4, left, for both calorimeter and PF jets. The jet response is already nearly flat and close to the expected value for PF jets. The jet energy resolution after corrections of the calorimeter jets is shown in Fig. 4, right. PF reconstruction of the event leads to an improvement of $\approx 40\%$ on the jet energy resolution at 40 GeV, allowing to recover a value compatible with the one obtained in ATLAS.

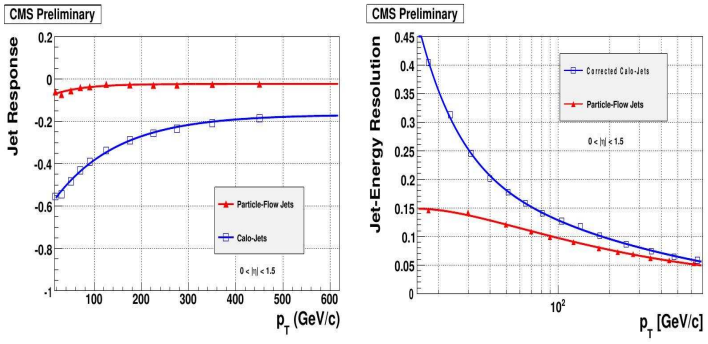


Figure 4: Jet response (left) and jet energy resolution (right) for calorimeter and particle flow jets in CMS.

3.3 First measurements with jets

In order to determine the jet energy scale with real data, different ranges in energy are treated differently. Jets with $10 < p_T < 200\,\text{GeV}$ ($200 < p_T < 500\,\text{GeV}$) are corrected using Z+jets (γ+jets) events. In ATLAS, the jet energy scale is expected to be measured with a statistical uncertainty of 1% (1-2%) with $300\,\text{pb}^{-1}$ ($100\,\text{pb}^{-1}$) of integrated luminosity [4] [7]. The systematic uncertainties, at the level of 5-10% at low p_T, reducing to 1-2% for $p_T > 100$ GeV, are due mainly to theoretical uncertainties on ISR/FSR and on the underlying event (UE). Above 500GeV, a multi-jet p_T-balance method is used: low-p_T jets with known jet energy scale (JES) are balanced against high-p_T jet with unknown JES. A statistical (systematic) uncertainty of 2% (7%) is expected for $1\,\text{fb}^{-1}$ of integrated luminosity.

Another important step in the commissioning of the detector is the measurement of di-jet cross-sections. A small amount of data is shown to be enough to exceed the Tevatron p_T reach (700 GeV). With $10\,\text{pb}^{-1}$ at 14 TeV, the sensitivity to contact interactions goes beyond the Tevatron limit of 2.7 TeV [8]. With $100\,\text{pb}^{-1}$, the sensitivity to objects decaying into 2 jets (di-jet resonances: q*, Z', etc.) goes also beyond the Tevatron limit of 0.87 TeV. Uncertainties for such measurements are shown in Fig. 5 as a function of the jet p_T,

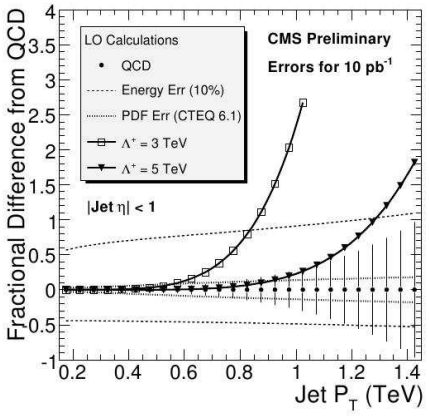


Figure 5: Uncertainties expected for $10\,\text{pb}^{-1}$ in di-jet events in CMS.

with the dominant one being due to the jet energy scale. Constraining the PDFs with such measurements will require a profound knowledge of the systematic uncertainties.

When the first data arrive, the first measurements involving jets will however be to characterise the underlying event, and to put constraints on the current Monte Carlo (MC) simulations. A method employed in CMS is described in detail in [9] and emphasises the search for variables allowing to discriminate between different MC models.

4 Heavy flavour at the LHC

4.1 B-tagging algorithms

Heavy-flavoured particles are characterised by a large lifetime: $c\tau$ =125-300 (500) μm for D (B) mesons. It is hence crucial for their identification to have a good reconstruction of tracks/vertices displaced from the primary vertex. In addition, semi-leptonic decays are important, with branching ratios BR(b→l+X) = 20% and BR(b→c→l+X) = 20%. Soft-lepton tagging methods will improve the identification. Furthermore, B hadrons take away about 70% of the b quark energy, so *high mass* states are looked for.

These criteria are used and combined differently in eleven algorithms in CMS, from the simple track-counting based algorithms to more evolved secondary-vertex finder algorithms. The current expected performance of a secondary-vertex tagging algorithm is of 1% mis-tagging rate for 50% (15%) b-tagging efficiency if no misalignment (start-up) scenario is applied.

4.2 Calibration on real data

In order to fully understand the detectors, the b-tagging efficiency and the mis-tagging rate must be extracted from real data. Two categories of methods are being developed, depending on the energy of the jets.

At low p_T, the efficiency is extracted from *muon-in-jet* QCD samples, using two methods in CMS. The p_T^{rel} method is based on estimating the muon content of jets, and the particularity of the projection of p_T^{μ} on the jet axis (p_T^{rel}), which is different for b- and u,d,s,g,c-jets. The System8 method [10] uses two di-jet samples of differing b-quark content, and two uncorrelated tagging algorithms (typically the soft muon one and the algorithm to be calibrated), to form a system of eight equations with eight unknown, from which the b-tag efficiency can be extracted. The result is shown in Fig. 6 for a track-counting based algorithm, for 10 pb^{-1} of integrated luminosity in CMS, and a 1% mis-tagging rate, as a function of the jet p_T. For 100 pb^{-1} of integrated luminosity in CMS, and a mis-tagging rate of 1%, an uncertainty of ≈8.6% is expected on the measurement of the b-tagging efficiency. Whereas the System8 method is expected to give reliable results at low p_T, it is however not suited for p_T larger than 80 GeV.

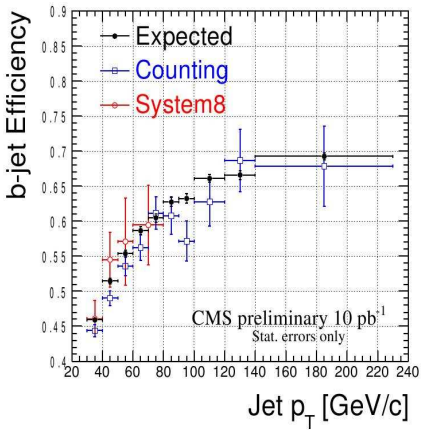


Figure 6: Expected b-tagging efficiency as a function of jet p_T, for 10 pb^{-1} of integrated luminosity in CMS.

At high p_T, a method developed in ATLAS relies on using jets from $t\bar{t}$ events to isolate a highly enriched b-jet sample [4]. Assuming that both top-quark decay into W+b, the events will indeed contain at least two b-jets. In addition, depending on the W decay mode, the topology studied will contain two leptons, or one lepton and two jets.

These events can be identified using a counting method: the number of events expected as a function of the number of tagged jets is shown in Fig. 7 for different MC samples, for $100\,\text{pb}^{-1}$ of integrated luminosity in ATLAS. When requiring more than one jet, $t\bar{t}$ events dominate by more than one order of magnitude compared to other samples. The b-tagging efficiency ϵ_b can be obtained with an uncertainty of $\Delta\epsilon_b/\epsilon_b \approx 2.7(4.2)(\text{stat.}) \oplus 3.4(3.5)(\text{syst.})\%$ for lepton+jets (di-lepton) final states with $100\,\text{pb}^{-1}$ of integrated luminosity.

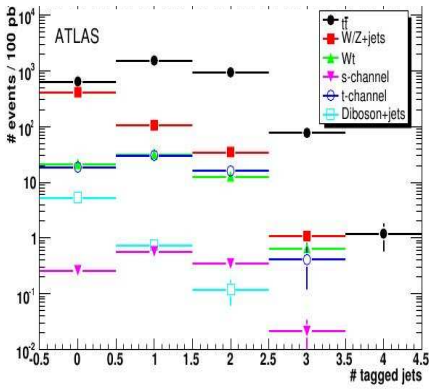


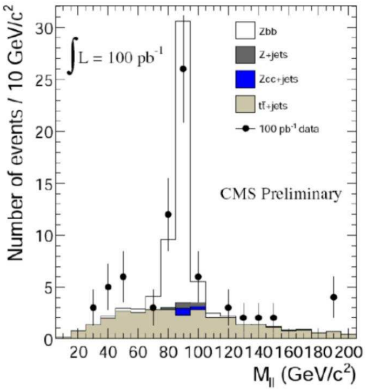
Figure 7: Number of b-tagged jets expected for different samples, with $100\,\text{pb}^{-1}$ of integrated luminosity in ATLAS.

$t\bar{t}$ events can also be identified using kinematic, topological or likelihood requirements. These methods require background subtraction, but allow to measure ϵ_b as a function of E_T and η of the jet. The resulting uncertainty is expected to be $\pm 10\%$ with 100 pb^{-1} in ATLAS (6-10% in CMS with $1\,\text{fb}^{-1}$).

4.3 Measurement of the bbZ cross-section

An example for a measurement involving b-tagging is the measurement of the b(b)Z cross-section. The gains of such an analysis are both theoretical and experimental: the same techniques are employed to calculate the b(b)H cross-section, and large theoretical uncertainties ($\approx 20\%$ uncertainty due to renormalisation and factorisation scales, and an additional 5-10% due to PDFs) exist, which could be constrained by a measurement [11].

A preliminary study has been done in CMS [12], using a selection of at least two leptons (e ($\eta < 2.5$) or μ ($\eta < 2$)) with $p_T > 20\,\text{GeV}$, with in addition at least two jets with $\eta < 2.4$ and $E_T > 30\,\text{GeV}$. A track-counting b-tagging discriminant is used, with a working point leading to a mis-tagging rate smaller than 1% (0.1%) for c (light-flavour) jets. The $t\bar{t}$ background is further reduced by a cut on the transverse missing energy $ME_T < 50\,\text{GeV}$. The main systematic uncertainties are due to: jet energy scale ($\pm 7.6\%$), ME_T ($\pm 7.4\%$), difference between NLO and LO for generator level cuts (-10%), luminosity



$$\delta\sigma = \pm^{21\%}_{25\%}\,(syst) \pm 15\%\,(stat.)$$

Figure 8: Invariant mass of the two leptons in a selection of bbZ events, with $100\,\text{pb}^{-1}$ of integrated luminosity in CMS.

(±10.%), b-tagging (±16%), mis-tagging (±1.%), $t\bar{t}$ background subtraction (±4.6%), with the numbers calculated for an integrated luminosity of $100\,\mathrm{pb}^{-1}$. With early data, it will hence already be possible to measure the bbZ cross-section with a total uncertainty of the order of the theoretical uncertainties.

5 Conclusion

LHC can probe (p)QCD, but the dominant experimental uncertainty, due to the jet energy scale, must be controlled. A large integrated luminosity will be needed to obtain a 1% error on the jet energy scale. Vice versa, QCD is essential to LHC discoveries: a better understanding of hard-scattering processes will lead to a better understanding of the backgrounds to new physics. Contact interactions and resonances decaying into di-jets can be discovered early on, even with 10% JES uncertainty at start-up. Theoretical uncertainties also need to be reduced to the size of the experimental uncertainties to increase the sensitivity to new physics. Feedback loops between measurements and theory are important in this respect.

References

[1] The ATLAS Collaboration, JINST 3 (2008) S08003.

[2] The CMS Collaboration, JINST 3 (2008) S08004.

[3] J.M. Campbell, J.W. Huston and W.J. Stirling, *Hard interactions of quarks and gluons: a primer for LHC physics*, Rep. Prog. Phys. **70** (2007).

[4] The ATLAS Collaboration, CERN-OPEN-2008-020 (2008) (arXiv:0901.0512).

[5] http://cms-physics.web.cern.ch/cms-physics/public/JME-07-002-pas.pdf.

[6] http://cms-physics.web.cern.ch/cms-physics/public/PFT-09-001-pas.pdf.

[7] C. Buttar, *Early QCD measurements with ATLAS*, XVII International Workshop on Deep-Inelastic Scattering and Related Subjects, Madrid, Spain (2009), http://indico.cern.ch/contributionDisplay.py?contribId=241&sessionId=3&confId=53294.

[8] http://cms-physics.web.cern.ch/cms-physics/public/SBM-07-001-pas.pdf.

[9] http://cms-physics.web.cern.ch/cms-physics/public/QCD-07-003-pas.pdf.

[10] http://cms-physics.web.cern.ch/cms-physics/public/BTV-07-001-pas.pdf.

[11] J. Campbell, R.K. Ellis, F. Maltoni and S. Willenbrock, arXiv:hep-ph/0510362 (2008).

[12] http://cms-physics.web.cern.ch/cms-physics/public/EWK-08-001-pas.pdf.