

## Jet physics in hadronic collisions: theoretical remarks

Michelangelo L. Mangano<sup>1,a</sup>

<sup>1</sup>CERN, PH-TH, 1211 Geneva 23, Switzerland

**Abstract.** I review some of the lessons that are emerging from the analysis of jet production at the LHC.

### 1 Introduction

Jets are the most direct and frequent signal of hard interactions in hadronic collisions (for a recent review, see [1]). They reflect the dynamics of the quarks and gluons produced in large-momentum-transfer ( $Q^2$ ) processes. Jets may arise from the decay of heavier particles, and from higher-order radiative processes. In the former case, they can be used as discovery probes, or as means to measure the properties of their parent particles (e.g. the top quark mass). In the latter case, they directly probe the QCD dynamics (e.g. higher-order corrections or parton distribution functions (PDFs)), and they determine final states that can act as important backgrounds to searches or measurements.

We should therefore look at jet physics in hadronic collisions from three perspectives:

1. **Opportunities.** Fully exploit jet probes to:
  - do precision measurements (PDFs,  $m_{top}$ ,  $\alpha_s$ , hard diffraction, ...);
  - probe the existence of BSM phenomena.
  - Question: how much can we push, and trust, the precision of these predictions?
2. **Challenges.** Verify and improve the suitability of our calculations under different and difficult dynamical regimes:
  - high orders in perturbation theory (PT): e.g. loop corrections, or processes proportional to large powers of  $\alpha_s$ , as in multijet production;
  - resummation: e.g. in processes characterized by very different hard scales, say  $\mu_1 \gg \mu_2$ , where the expansion parameter  $\alpha_s \log(\mu_1/\mu_2)$  is large. Typical examples here include production of multiple soft jets accompanying the creation of very heavy objects ( $t\bar{t}$ +multijets), production of  $b$  quarks in high- $E_T$  jets ( $\mu_1 \sim E_T$  and  $\mu_2 \sim m_b$ ), etc.
  - Question: how reliable are our approximations? What is the best theoretical approach to a given observable (e.g. shower or fixed-order)?

3. **Needs:** which calculations/measurements are needed to improve on (1), or validate (2)?

As shown in the various contributions to the plenary and parallel sessions of this Conference, the amount of measurements emerging from the analysis of data at the Tevatron, HERA and LHC is immense. Rather than reviewing all of this material, I shall focus here on a few specific examples, to give an overall assessment of the status of theoretical predictions, and highlight outstanding challenges and opportunities that emerge from the data.

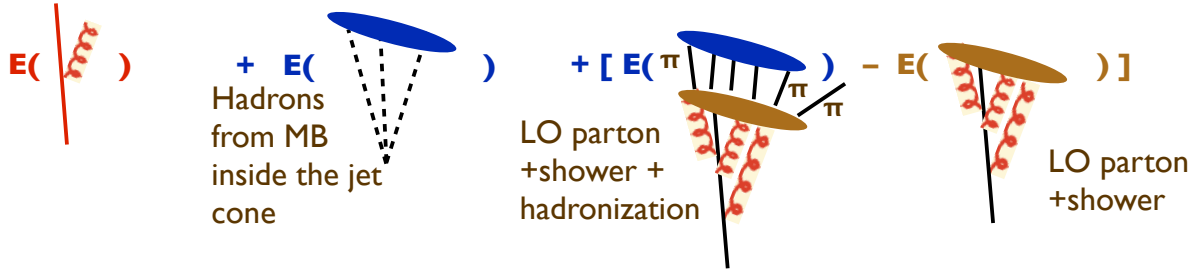
I would like to dedicate this contribution to the memory of Professor Kunitaka Kondo (1935-2011). A gentle and generous man, who greatly contributed to the Tevatron program, to the understanding of jets in  $p\bar{p}$  collisions, and to their use as analysis tools and probes of BSM phenomena, and who organized the last Hadron Collider conference to take place in Japan [2].

### 2 Inclusive jet cross section

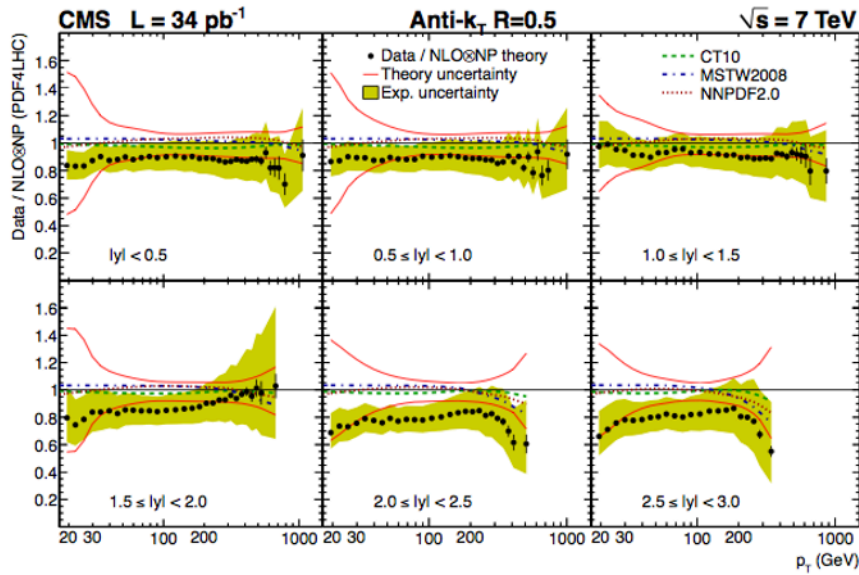
The  $E_T$  distribution of inclusive jets is a direct probe of PDFs and of  $\alpha_s$ . Its behaviour at the largest values of  $E_T$ , furthermore, allows to test the point-like nature of quarks. The robustness of such test, on the other hand, relies on the precision with which PDFs are known [3], stressing once more the relevance of the direct PDF determination, particularly in the large- $x$  region. In this respect, measurements at large rapidity ( $y$ ) are particularly important, since at large  $y$  one can probe the large- $x$  behaviour of PDFs in a region of  $Q^2$  where the absence of possible new physics effects has already been verified, and these do not influence the interpretation of the results in terms of PDFs.

Next-to-leading order (NLO) parton-level QCD predictions for inclusive jet cross sections at hadron colliders have been available for a long time [4–6]. To compare these predictions with data, the jet is defined as the sum of the one or two partons that can be clustered into an individual jet. This relies on the fact that, for inclusive quantities such as a jet cross section, the shower evolution, and thus the full-order resummation of collinear and soft emissions, modifies the predictions only by terms that are formally beyond NLO. The parton-level jet energy is then

<sup>a</sup>e-mail: michelangelo.mangano@cern.ch



**Figure 1.** Schematic representation of the corrections needed to compare a fixed-order NLO calculation of jet rates with data:  $E = E_{NLO} + E_{MB} + \Delta_{had}E$ , where  $E_{NLO}$  is the energy of the 1 or 2 partons in the NLO final state,  $E_{MB}$  is the contribution from the fragments of the protons, and  $\Delta_{had}E$  is the energy shift before and after the hadronization of the partons emitted during the shower evolution.



**Figure 2.** CMS measurement [8] of inclusive jet  $E_T$  spectra, in different rapidity ranges, at  $\sqrt{s} = 7$  TeV. Jets are defined by the anti- $k_T$  algorithm [7], with  $R = 0.5$ . The data points are normalized to a parton-level NLO calculation, corrected for the effects of hadronization and underlying event. The theoretical systematics includes both scale and PDF uncertainty.

corrected for two non-perturbative (non-PT) effects, which cannot be accounted for by the pure parton-level result: the presence of the underlying event (namely energy deposited inside the jet by the fragments of the colliding protons), and the hadronization corrections (partonic energy can be dragged in or out of the jet cone when partons are brought together to form hadrons). This is shown schematically in fig. 1.

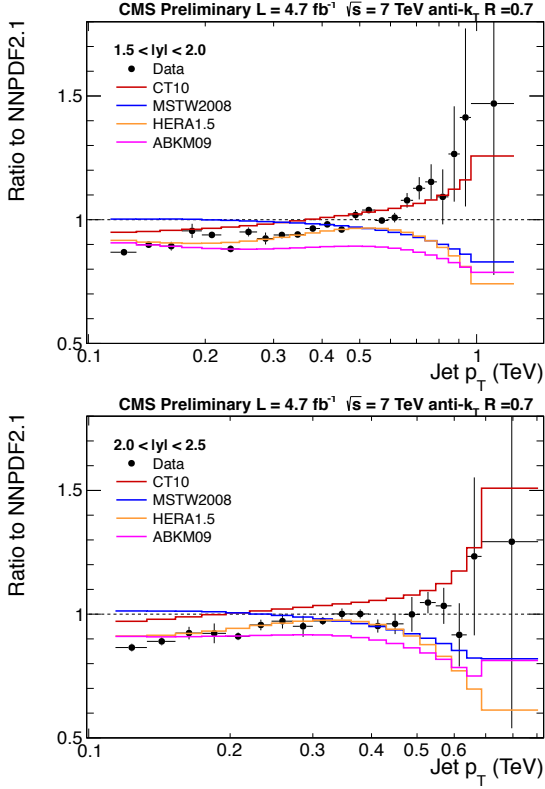
As an example, fig. 2 shows the results published by CMS, from the first 7 TeV run [8]. Overall the agreement, within the quoted uncertainties, is excellent. Few remarks are nevertheless in order.

The large theoretical systematics at small  $E_T$  are dominated by non-PT effects; for central rapidities, the agreement of the data with the central value of the prediction is however excellent, and this therefore suggests that this systematics is likely overestimated. The data could therefore be used to constrain more tightly the impact of non-PT physics. The agreement however deteriorates at larger  $y$ , where it is only by fully exploiting the non-PT systematics that data agree with theory within  $1\sigma$ . Thus the question

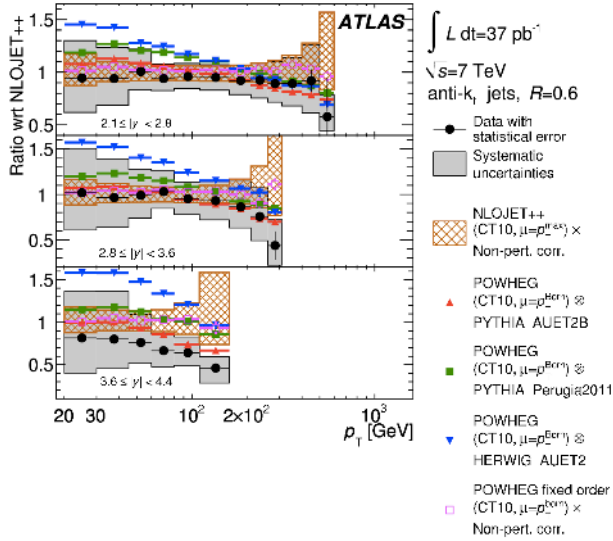
is whether these systematics are correlated or not across the  $y$  range. If they are, then there are indications of a discrepancy that cannot be attributed to non-PT physics.

At large  $E_T$ , on the other hand, the data tend to be on the low side of the predictions. Different PDFs are quite compatible with each other, and what dominates the theory systematics is the spread of any individual PDF set. There are therefore indications that one could use these data to improve the PDF knowledge. The discrepancy becomes bigger and bigger at large  $y$ , where however also the experimental systematics, driven by the jet energy scale, grows. The question is, once again, how much of this uncertainty is correlated between the central and forward  $y$  regions.

The trend is different with a larger jet cone radius, as shown in fig. 3, where higher-statistics CMS data [9], analyzed with a radius of 0.7, are on the high side of the theory curves. Since  $E_T$  slopes at large  $y$  are much steeper than at central  $y$ , they are much more sensitive to small changes in the jet energy. It is therefore suggestive to conclude that higher-order perturbative corrections due to multiple gluon emission in the shower have a non-negligible nu-



**Figure 3.** Same as fig. 2, but with a cone radius  $R = 0.7$ , and for a larger statistics analysis of CMS data [9].



**Figure 4.** ATLAS results [10] for jet production in the forward  $y$  region, normalized to parton-level QCD predictions (scaled by non-PT effects). The data are compared to predictions from various NLO+shower results, using different shower models and underlying event tunes.

merical impact on the jet energy, in spite of their being formally of higher order.

This conclusion is consistent with the results of an ATLAS study [10], in which the data are compared to fixed-order NLO predictions from [11] and to the results of a complete NLO+shower calculation [12], done with POWHEG and either Pythia [13] or Herwig [14] for the shower evolution. In the case of Pythia, POWHEG's agreement with data is significantly better than the fixed-order NLO result, particularly at large  $E_T$  and large  $y$ . Furthermore, noticeable differences appear in general between Pythia and Herwig showers, and different tunes of the underlying event.

So, while the overall agreement of data and theory is certainly satisfactory, I believe that few things need to be better understood in order to reliably use high- $E_T$  and high- $y$  jet data to improve PDFs.

### 3 Jet cross section ratios

The availability of LHC data at different energies opens a new opportunity for precise measurements with jets. In fact, a large fraction of the experimental and theoretical systematics in the jet cross section measurements have some degree of correlation at different energies, and may therefore cancel in the measurement of cross section ratios. Historically, jet production at different energies have been compared as a function of the  $x_T = E_T/E_{beam}$  variable. This reduces the theoretical systematics related to PDFs, since these are probed at the same  $x$  values. A direct comparison of  $E_T$  spectra, on the other hand, enjoys a correlated theoretical scale uncertainty (the underlying hard process is the same at the two energies), as well as more closely correlated experimental systematics (e.g. jet energy scale). The larger PDF systematics of this ratio, compared to  $x_T$  ratios, is also desirable, since it gives potential sensitivity to improve our knowledge of the PDFs themselves!

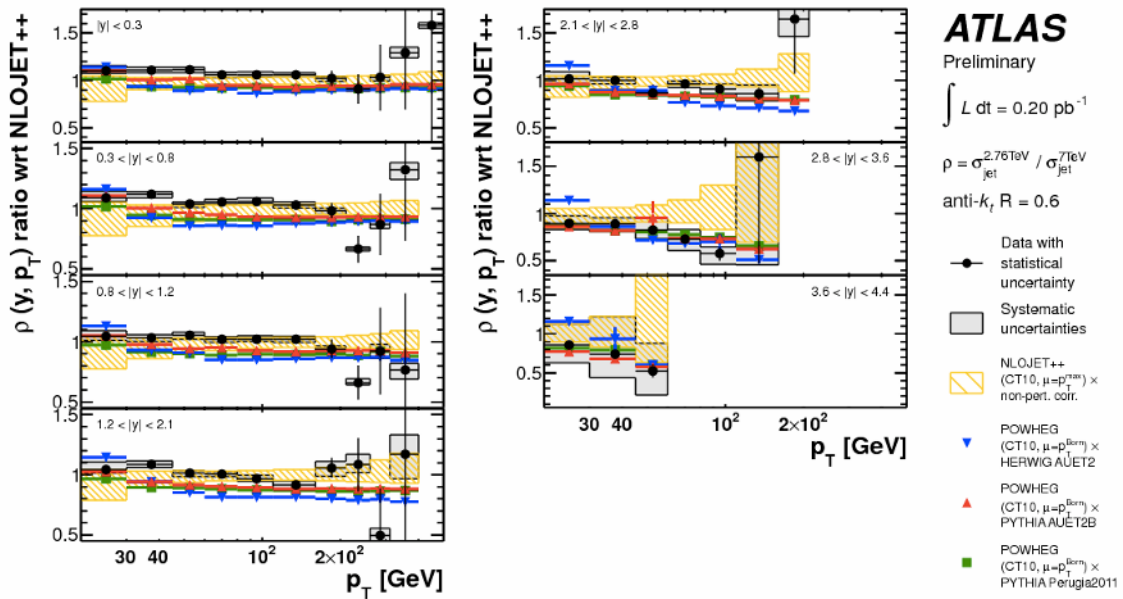
Several examples of the remarkable theoretical precision that can be achieved through cross section ratios at different energies, are given in [15]. For example, that study finds that:

$$\frac{\sigma(E_T > 2\text{TeV}; \sqrt{s} = 8\text{TeV})}{\sigma(E_T > 2\text{TeV}; \sqrt{s} = 7\text{TeV})} = 7.65 \pm 1.5\%(\text{scale}) \pm 7\%(\text{PDF}) \quad (1)$$

$$\frac{\sigma(E_T > 2\text{TeV}; \sqrt{s} = 14\text{TeV})}{\sigma(E_T > 2\text{TeV}; \sqrt{s} = 8\text{TeV})} = 182 \pm 4\%(\text{scale}) \pm 10\%(\text{PDF}) \quad (2)$$

Notice that the scale uncertainty is much smaller than the PDF uncertainty, giving the opportunity to improve the PDF knowledge with a direct comparison with data.

A first experimental exploration of  $E_T$  cross section ratios, using data from the short 2011 run at  $\sqrt{s} = 2.76$  TeV and from  $\sqrt{s} = 7$  TeV, has been documented recently by ATLAS [16]. Figure 5 shows the cross section ratios, as a function of jet  $E_T$  and for different  $y$  windows, compared



**Figure 5.** ATLAS results [16] for the ratio of jet  $E_T$  distributions at  $\sqrt{S} = 2.76$  and at 7 TeV. The ratio is normalized to NLO QCD plus non-PT corrections, and compared to the predictions of various NLO+shower calculations.

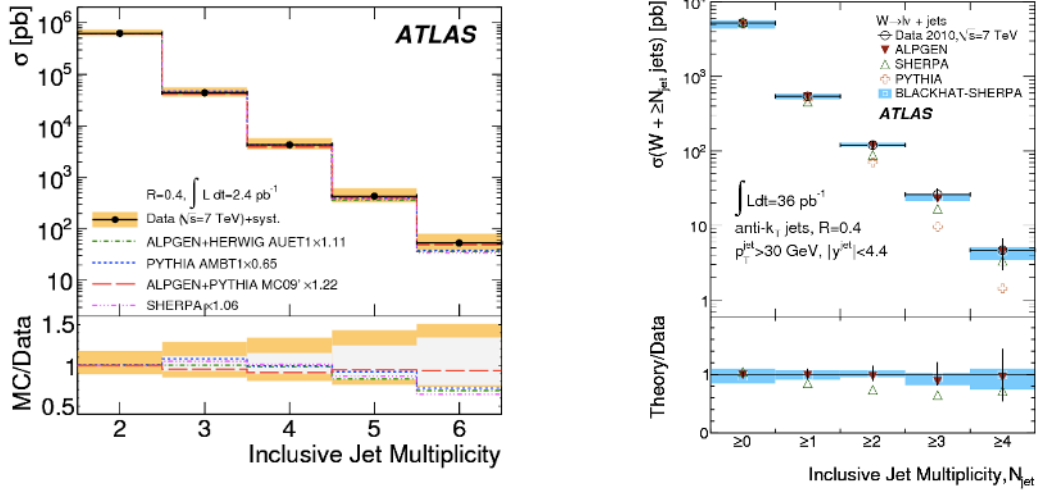
against parton level NLO and against the NLO+shower result of POWHEG, using different choices of parton showers and underlying event tunes. The comparison with the results of fig. 4 is quite telling: on one side, for a fixed PDF the spread of theoretical predictions is much smaller in the ratio than in the 7 TeV cross section, while the PDF systematics remains quite large. Notice also the very small experimental uncertainty, which blows up only at the largest  $E_T$  values, dominated by the poor statistics at 2.76 TeV. Reference [16] also includes a first study of the impact of the inclusion of these data in PDF fits, showing that they can indeed shift the central values, and reduce the overall uncertainty, of the gluon density over a wide range of  $x$ . The much larger statistics available today at both 7 and 8 TeV, and the more accurate luminosity determination, will certainly offer a very powerful tool to carry these studies even further.

## 4 Multijets

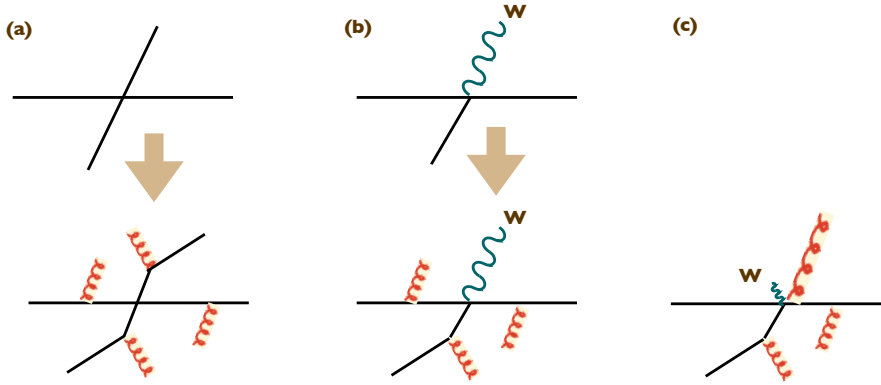
Final states with many hard jets are very interesting: on one side they pose a challenge to the theoretical calculations, which require the evaluation of huge numbers of complex Feynman diagrams, and possibly the resummation of large logarithms emerging from the existence of very disparate scales (e.g. the individual jet  $E_T$  versus the transverse energy of the overall multijet system). On the other, multijet final states (possibly accompanied by other objects like gauge bosons or top quarks) are the dominant background to many searches of phenomena beyond the Standard Model. Their study is therefore a crucial component of the LHC physics programme.

Available analyses from ATLAS and CMS are still limited to the low-statistics 2010 run, but provide already very valuable information. Figure 6 shows ATLAS's measurements of the inclusive multijet [17] and  $W$ +multijet [18] rates, compared to the predictions of several theoretical approaches. In the multijet case, the comparison is done with leading-order matrix element calculations, merged with shower evolution (Sherpa [19], and Alpgen [20] plus Herwig and Pythia), and with a pure parton shower approach (Pythia). In the  $W$ +jets case, there is also the comparison with parton-level NLO results [21].

The leading-order matrix element plus shower results and the parton-level NLO results agree well with data, and the occasional minor discrepancies are consistent with the theoretical systematics, and can in principle be “tuned away”, for example, by adapting the scale choice or the details of the merging algorithms. The pure shower-evolution results from Pythia agree well with the multijet rates, but fall short of predicting the rates for  $W$ +multijets. This apparently contradictory conclusion has a possible explanation, graphically outlined in fig. 7. In the case of multijets, Pythia starts the shower evolution from a “seed” parton level configuration with two hard, recoiling, partons. Additional jets are emitted from both initial and final states and, to the extent that they are softer than the leading jets, the shower approximation appears to work well. In the case of  $W$ +jets, however, the partonic configuration that seeds the shower has a single jet recoiling against the  $W$ . The emission of softer jets from the initial states and from the final-state parton, will dominantly lead to configurations in which the  $W$  maintains a large transverse momentum. This evolution entirely misses, however, the important region of phase space in which the  $W$  boson is soft



**Figure 6.** ATLAS measurements of inclusive jet rates, in multijet [17] (left) and  $W$ +multijet [18] (right) final states.



**Figure 7.** Schematic picture of the shower generation of multijet (a) and  $W$ +multijet (b) final states, starting from the leading-order processes included in shower Monte Carlo programs. In the case of  $W$ +jets, only configurations with a large value of  $p_T(W)$  can be generated (lower graph of (b)), while the important phase space region where  $p_T(W)$  is small (c) is neglected.

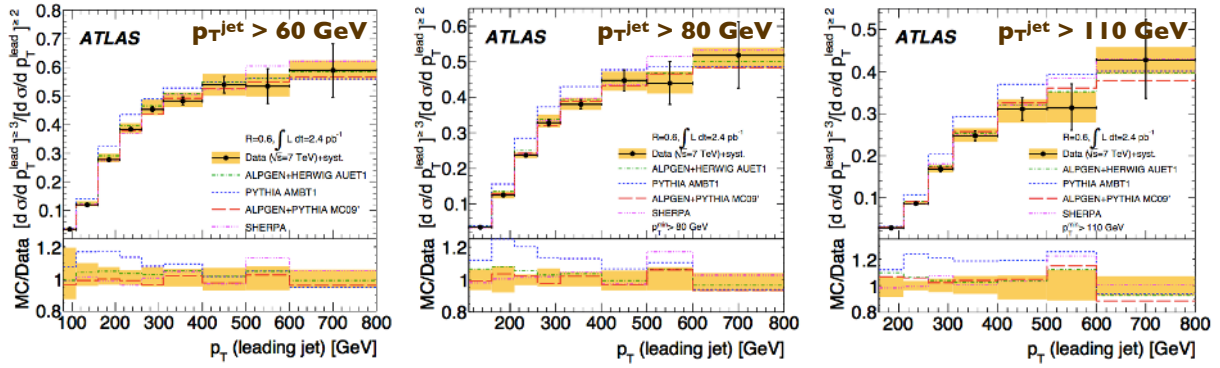
with respect to the rest of the event, as shown in inset (c) of the figure. The more jets we require in the final state, the larger must be the initial  $p_T(W)$ , to allow the recoiling jet and the initial state to have enough energy to radiate the required jets. This reduces even more the accessible phase-space for the  $W$ , giving an overall rate that drops more and more at larger jet multiplicity.

One more observation is worthwhile. Even though the shower-only approach gives a good description of inclusive multijet rates, there is no guarantee that kinematical correlations among the jets are equally well predicted. Consider for example fig. 8, taken again from [17]. It describes the fraction of events with three or more jets, as a function of  $E_T^{lead}$ , the  $E_T$  of the leading jet. The three plots correspond to different  $E_T$  thresholds for the definition of the jets,  $E_T > 60, 80$  and  $110$  GeV. Pythia seems less accurate in the region of smaller  $E_T^{lead}$ . There appears to be a change in the agreement pattern for  $E_T^{lead} > 300, 400$  and  $500$  GeV, respectively. Given that the total transverse energy of the jets,  $H_T$ , must be larger than  $2 \times E_T^{lead}$ , we see that the best agreement is obtained when  $H_T \gtrsim 0.1 \times E_T^{min} \sim$

$\alpha_S \times E_T^{min}$ , in other words when the  $E_T$  of the radiated jet is indeed *soft* with respect to the  $Q$  scale of the event. Notice also that Alpgen and Sherpa reproduce these distributions quite well over the full range of  $E_T^{lead}$ .

## 5 Conclusions and outlook

Jet physics in hadron collisions is as healthy as ever. The multitude of tests and probes explored by the Tevatron and LHC experiments gives us great confidence that the theory tools we have available are good, and much work is in place to improve them even further. The leading-order matrix element generators, merged with shower evolution, are very successful in describing complex higher-order topologies and rates. NLO calculations are making great progress: they are now available for high multiplicity final states, beyond the most optimistic hopes of only few years ago. New techniques are emerging, which I could not review here, to fully automatize the calculation of the NLO cross sections and their parton showers evolution [22], extending also the possibility of merging



**Figure 8.** ATLAS measurement [17] of  $\geq 3$ -jet fractions, as a function of the leading jet  $E_T$ , and for different jet thresholds.

together samples of different multiplicity [23–25], maintaining the NLO accuracy. With NLO results close to becoming a straightforward enterprise, limited only by CPU power, NNLO results start appearing. For example, after the end of the Workshop, a major new achievement was reported, namely the calculation of the NNLO inclusive  $E_T$  jet cross section, limited to the  $gg$  initial state [26]. The calculation shows a significant reduction of the scale uncertainty, with a mild increase in rate, at the 20% level relative to NLO.

A new era of quantitative and precise applications of hard QCD at the LHC is opening, much like the quantitative tests that came from LEP and HERA. The full benefits of precision jet physics at the LHC are yet to be explored. But since new physics, if any is there, is hiding well, our safest bet today to pull it out of nasty SM backgrounds is to invest in a full-scale campaign of improvement in the precision with which we can measure/predict hard processes at the LHC.

## References

- [1] J. M. Butterworth, G. Dissertori and G. P. Salam, *Ann. Rev. Nucl. Part. Sci.* **62** (2012) 387 [arXiv:1202.0583 [hep-ex]].
- [2] K. Kondo, (ed.) and S. Kim, (ed.), “Proton anti-proton collider physics. Proceedings, 9th Topical Workshop, Tsukuba, Japan, October 18-22, 1993,” Tokyo, Japan: Universal Acad. Pr. (1994) 469 p
- [3] S. Forte and G. Watt, arXiv:1301.6754 [hep-ph].
- [4] R. K. Ellis and J. C. Sexton, *Nucl. Phys. B* **269** (1986) 445.
- [5] F. Aversa, P. Chiappetta, M. Greco and J. P. Guillet, *Nucl. Phys. B* **327** (1989) 105.
- [6] S. D. Ellis, Z. Kunszt and D. E. Soper, *Phys. Rev. Lett.* **64** (1990) 2121.
- [7] M. Cacciari, G. P. Salam and G. Soyez, *JHEP* **0804** (2008) 063 [arXiv:0802.1189 [hep-ph]].
- [8] S. Chatrchyan *et al.* [CMS Collaboration], *JHEP* **1206** (2012) 036 [arXiv:1202.0704 [hep-ex]].
- [9] CMS Collaboration, CMS-PAS-QCD-11-004.
- [10] G. Aad *et al.* [ATLAS Collaboration], *Phys. Rev. D* **86** (2012) 014022 [arXiv:1112.6297 [hep-ex]].
- [11] Z. Nagy, *Phys. Rev. D* **68** (2003) 094002 [hep-ph/0307268].
- [12] S. Alioli, K. Hamilton, P. Nason, C. Oleari and E. Re, *JHEP* **1104** (2011) 081 [arXiv:1012.3380 [hep-ph]].
- [13] T. Sjostrand, S. Mrenna and P. Z. Skands, *JHEP* **0605** (2006) 026 [hep-ph/0603175].
- [14] G. Corcella *et al.* hep-ph/0210213.
- [15] M. L. Mangano and J. Rojo, *JHEP* **1208** (2012) 010 [arXiv:1206.3557 [hep-ph]].
- [16] ATLAS Collaboration, ATLAS-CONF-2012-128.
- [17] G. Aad *et al.* [ATLAS Collaboration], *Eur. Phys. J. C* **71** (2011) 1763 [arXiv:1107.2092 [hep-ex]].
- [18] G. Aad *et al.* [ATLAS Collaboration], *Phys. Rev. D* **85** (2012) 092002 [arXiv:1201.1276 [hep-ex]].
- [19] F. Krauss, A. Schalicke, S. Schumann and G. Soff, *Phys. Rev. D* **72** (2005) 054017 [hep-ph/0503280].
- [20] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau and A. D. Polosa, *JHEP* **0307** (2003) 001 [hep-ph/0206293].
- [21] C. F. Berger *et al.*, *Phys. Rev. Lett.* **106** (2011) 092001 [arXiv:1009.2338 [hep-ph]].
- [22] V. Hirschi *et al.*, *JHEP* **1105** (2011) 044 [arXiv:1103.0621 [hep-ph]].
- [23] K. Hamilton, P. Nason and G. Zanderighi, *JHEP* **1210** (2012) 155 [arXiv:1206.3572 [hep-ph]].
- [24] S. Hoeche, F. Krauss, M. Schonherr and F. Siegert, arXiv:1207.5030 [hep-ph].
- [25] R. Frederix and S. Frixione, *JHEP* **1212** (2012) 061 [arXiv:1209.6215 [hep-ph]].
- [26] A. G. D. Ridder, T. Gehrmann, E. W. N. Glover and J. Pires, arXiv:1301.7310 [hep-ph].