NLO high multiplicity processes

Journal of Physics: Conference Series **608** (2015) 012075 doi:10.1088/1742-6596/608/1/012075

NLO high multiplicity processes

D. Maître

Institute for Particle Physics Phenomenology, University of Durham, Durham DH1 3LE, UK

E-mail: daniel.maitre@durham.ac.uk

Abstract.

In this presentation some aspects of Next-to-Leading Order (NLO) calculations in QCD are presented. The focus is brought to aspects of such calculations for processes involving a high final-state particle multiplicity.

1. Introduction

Events with a large number of jets are common place at the LHC, both in Standard Model (SM) measurements and as signature or background to beyond the standard model (BSM) models. Benchmark SM processes such as vector boson and jets have been measured with up to 6 jets[1, 2, 3, 4].

Precise theoretical predictions are needed to compare with these precise measurement and also provide a good understanding of these processes when they are a background to new physics searches. A good theoretical understanding also helps designing so-called data-driven background estimation.

NLO calculations improve upon Born-level calculations by reducing the dependence of the latter on the unphysical renormalisation scale. The differential shape of distributions are described in a more precise way and the structure of jets start to be resolved at that order. The unphysical renormalisation scale dependence enters the cross section calculation through the strong coupling constant. The problem is exacerbated when a large number of jets – each yielding one power of the coupling constant – is considered. Figure 1 illustrates the increase in the renormalisation scale dependence as the power of the renormalisation constant increases. At NLO the situation is greatly improved since the virtual matrix elements they include have an additional dependence on the renormalisation scale which cancel the tree-level scale dependence to a large degree.

The uncertainty resulting from the scale dependence is usually assessed by varying the scales by factors between 1/2 and 2 around a central scale. Figure 1 shows the relative uncertainties obtained using this procedure for the total cross section for W+jets production as a function of hte jet multiplicity from Ref. [5].

Until recently the bottleneck for NLO calculations was the calculation of the virtual matrix element. This bottleneck has been largely cleared in the last year, in what has been called the "NLO revolution". A lot of this progress has been achieved through the development of so-called unitarity techniques (for a review, see [6]). A sketch of how this method works is presented in the following section.

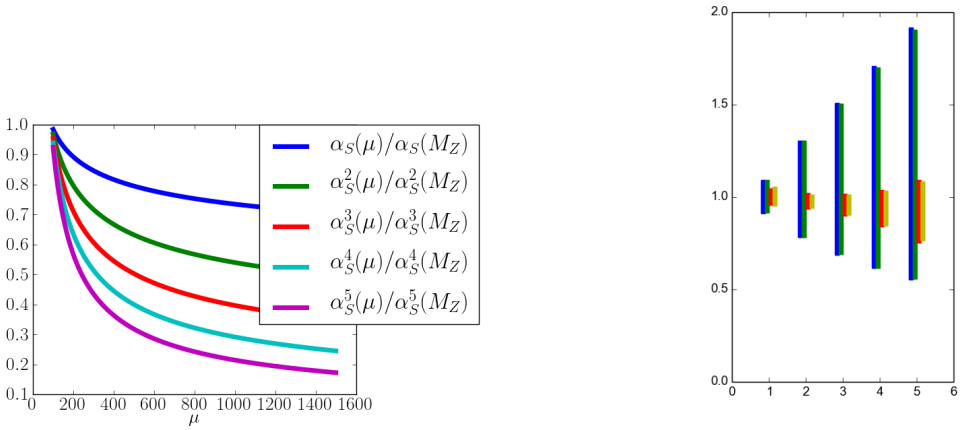


Figure 1. Left pane: Renormalisation scale dependence of the strong coupling constant. Each curve represents the dependence of the strong coupling constant on the renormalisation scale. The curves are normalised to their values at the Z boson mass. Right pane: Uncertainty due to renormalisation and factorisation scales for total cross section for the production of a W-boson as a function of the jet multiplicity, normalised by the central value. The blue and green bars represent the normalised uncertainty at LO for W^- and W^+ respectively. The red and yellow bars show the NLO scale variation for W^- and W^+ respectively.

2. Unitarity technique

To calculate NLO cross sections one has to include real emission matrix element and the virtual one-loop matrix elements for the process to account for one additional radiation and the influence of virtual particles respectively. In this section we concentrate on describing the unitarity technique approach to computing the virtual matrix element.

For this illustration we restrict ourselves to massless internal particles. In this case a general one-loop amplitude can be written in terms of known scalar two-, three- and four-point integrals.

$$\mathcal{A} = R + \sum d_i I_i^{(4)} + \sum c_i I_i^{(3)} + \sum b_i I_i^{(2)}, \tag{1}$$

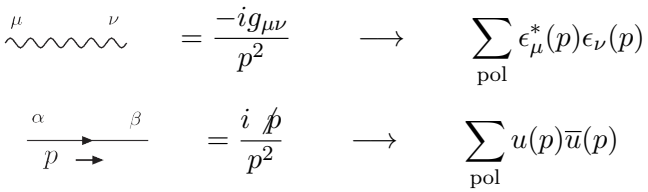
where $I^{(n)}$ are the n-point scalar integral and the sums run over all possible propagator combinations. The scalar integrals are process independent. The first term R is a rational part free of logarithm or polylogarithms. The coefficients d_i, c_i and b_i are rational functions of the spinor products. The calculation of the one-loop integral is reduced to the calculation of these coefficients.

The way unitarity techniques work is by applying so-called unitary cuts on both sides of Equation 1. A unitarity cut is the replacement under the integral:

$$\frac{1}{P^2} \to 2\pi i \delta(P^2)\,. \tag{2}$$

This replacement is done in the full loop integral on the left-hand side and in each individual scalar integral in the right-hand side. If the propagator being cut is not present in the expression the result of the cut is simply 0, unitarity cuts can thus be seen as projectors onto structures that exhibit a given set of propagators. To determine the coefficients in Equation 1 one needs to apply two, three or four cuts simultaneously. The delta-function constraints imposed by the cuts can often only be satisfied by complex momenta.

While the cuts on the right-hand side of Equation 1 will result in a set of integrals being replaced by a Jackobian factor and the other with 0, the cuts on the left-hand side will affect the propagators, transforming them into a sum of external wave-function factors. This is illustrated in Figure 2. As a result of the unitarity cut the integrand on the left-hand side of Equation 1



$$= \frac{-ig_{\mu\nu}}{p^2} \longrightarrow \sum_{\text{pol}} \epsilon^*_\mu(p)\epsilon_\nu(p)$$

$$= \frac{i \not{p}}{p^2} \longrightarrow \sum_{\text{pol}} u(p)\overline{u}(p)$$

Figure 2. Effect on the gluon and quark propagator after application of a unitarity cut.

collapses into a product of trees. Depending on the number of cuts an integration over all momenta satisfying the cut conditions might remain. The next remaining step is to perform this integration. The nature of the one-loop integrand imposes constraints on the parametric dependence of the integrand as a function of the remaining parameter. Typically the functional form will be a rational function of the integration parameters. Through the knowledge of this form one can extract the value of the integral by evaluating the integrand at a very small number of points. To illustrate how this can be achieved we consider the example where the general form of the integrand is given by

$$i(t) = \frac{1}{t} \underbrace{\left(\frac{c_{-3}}{t^3} + \frac{c_{-2}}{t^2} + \frac{c_{-1}}{t} + c_0 + c_1 t + c_2 t^2 + c_3 t^3\right)}_{f(t)}$$

and the integration range of t is a circle of unit radius in the complex plane. The result of the integral is clearly $2\pi c_0$ and the value of c_0 can be directly extracted from the combination of evaluations of the integrand at specific values:

$$I = \frac{2\pi i}{7} \sum_{j=0}^{6} i(e^{\frac{2\pi i j}{7}}) \, .$$

All integral coefficients can be extracted in this way starting from the box coefficients and using know coefficient to help determine lower-point coefficients. The procedure is illustrated in Figure 3. Once all coefficients of the scalar integrals are determined the last ingredient is the

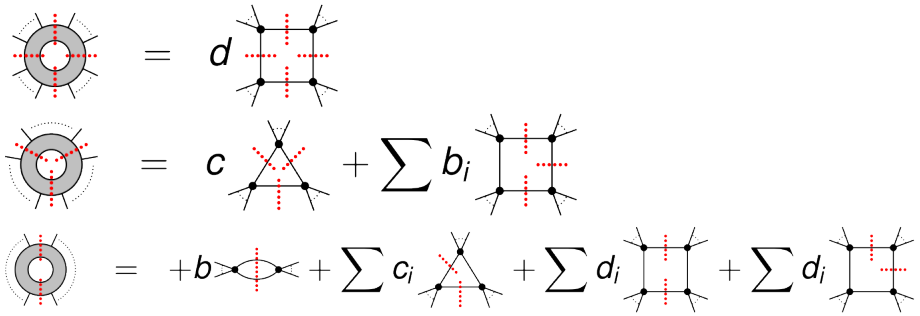


Figure 3. Illustration of the cut procedure to determine the scalar integral coefficients.

rational term. There are different ways of computing this term. A detailed description is beyond

the scope of this illustration, details can be found in Ref. [6]. The unitarity method works very well at NLO and attempts are being made to apply it at NNLO (see for example ACAT 2013 proceedings [7] for a more detailed description). In this case the situation is more complicated, owing to the fact that there is no known analogous to Equation 1 at two-loop level.

3. NLO event files

NLO calculations for high multiplicity processes are computationally intensive and it is therefore advisable to try to reuse these calculations for as many analysis and observables as possible. The most expensive part of the calculation is the computation of the matrix element, while other steps in a typical analysis – such as the computation of the observables, parton distribution functions, jet clustering, – are relatively inexpensive. One solution to the problem is to store the computed matrix elements and the coefficient of scale dependent quantities in a file along with the phase space point. These files can then be reread for another analysis, which is less computationally expensive than the original computation. Details of a practical implementation of this strategy using ROOT files [**?**] as a storage back-end can be found in Ref. [9].

The advantages of this strategy is that analysis observables and cuts can be adapted to other analysis (provided of course that the event files have been generated for an inclusive enough phase-space for the analysis.) By storing the coefficients of factorisation and renormalisation scales logarithms allows for the inexpensive calculation of scale and pdf variation – an operation that is otherwise extremely costly. Event files may be easier to handle for experimental groups than the program that generated them. In addition, the validation of an event file is easier than the validation of an entire NLO program, which can facilitate the dissemination of NLO results. These advantages come at the price of large event files, and a balance between inclusivity and file size has to be found.

There are complications due to nature of the NLO calculation that one must be aware of when using NLO event files. The first is that NLO event files cannot be valid for arbitrary jet algoithms or arbitrary values of the parameters of a given jet algorithm such as the jet radius. The files are typically generated in such a way that they are compatible with a wide range of jet algorithms. The second issue concerns the estimation of the statistical error and has implications for the extraction of cumulative distributions and re-binning of histograms.

The issue is cause by the infrared divergences of the real matrix element. At NLO these infrared divergences are caused by the additional radiated particle going either collinear to another particle or soft. This divergence is cancelled by the infrared divergences in the virtual part, but this cancellation is happening in different phase space. To organise this cancellation across phase spaces the most fequently used strategy is to construct a subtraction term to the real matrix element that contains all the divergent behaviour of the real matrix element, while being simple enough to be integrated over the unresolved emission phase space. This integrated subtraction part is added to the virtual part to cancel its explicit infrared divergences.

The implication for the NLO event files is that the events for the real emission will contain events in the infrared regions that have associated subtraction terms. The weights of these events will be large (because of the infrared divergence) and anti-correlated (because the subtraction has been constructed for exactly that purpose.)

This artificial large correlation will lead to a severe over-estimation of the statistical integration error when using the usual formula for hte variance:

$$\begin{aligned}\sigma^2 &= \langle(\omega - \langle\omega\rangle)^2\rangle = \langle\omega^2\rangle - (\langle\omega\rangle)^2 \\ &\simeq \frac{\sum\omega_i^2}{N} - \left(\frac{\sum\omega_i}{N}\right)^2 = \frac{1}{N}\left(\sum\omega_i^2 - \frac{(\sum\omega_i)^2}{N}\right) .\end{aligned}$$

To illustrate the problem, let us assume that all real weights and subtractions come in pairs,

both of approximate size $\omega_+ \simeq -\omega_- \simeq L$ and that their difference (the actual physically relevant quantity) is of order s. In this case the error will be estimated by

$$\sigma^2 \simeq \frac{1}{N}\left(NL^2 + \frac{(Ns)^2}{N}\right) = L^2 - s^2 \simeq L^2 \tag{3}$$

Which overestimate the error. We get a better estimation by summing the contributions of the real part and the associated subtraction *before* squaring

$$\sigma^2 \simeq \frac{1}{N}\left(\sum(\omega_{i,+} + \omega_{i,-})^2 - \frac{\left(\sum \omega_i\right)^2}{N}\right). \tag{4}$$

In some cases even using NLO event files may be too inefficient. This can happen for example when one is interested in a particular distribution to perform a pdf fit or strong coupling determination. In this case one can use dedicated programs such as FastNLO [10] or ApplGrid [11] to generate an interpolation table for a specific distribution that allows to predict the distribution with a different pdf set or one that uses a different value of the strong coupling constant. FastNLO works by approximating the pdf

$$f_a(x) \simeq \sum_i f_a(x_i)E(x). \tag{5}$$

using a set of interpolating functions $E(x)$ subject to the constraint

$$\sum_i E_i(x) = 1\,, \qquad E_i(x_j) = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

Using this approximation one can then write any cross section or distribution as

$$\sigma = \sum_{a,b,i,j} f_a(x_i) f_b(x_j) \underbrace{\int dx_1\, dx_2\, E_i(x_1) E_j(x_2) \int d\phi\, \hat{\sigma}(x_1, x_2; \phi)}_{c_{a,b;i,j}}\,. \tag{7}$$

The coefficients $c_{a,b;i,j}$ can be computed using NLO event files by replacing the pdfs with the functions $E(x)$, which is possible since the appropriate coefficients are stored to allow for a change of pdf or factorisation scale. The use of such tables also reduces the size of the files to be handled from the hundreds of Gb to a few Mb.

We now have a situation where one can gain evaluation speed for a NLO calculation at the cost of loss of generality. First on can use NLO event files, where one is restricted to a fixed set of jet algorithms and a set of generation cut. Within these constraint, it is possible to perform any analysis and compute any observable. One can improve the speed further by choosing a fixed analysis and a fixed observable and using fastNLO.

4. High multiplicity

To illustrate the importance of high multiplicity processes let us consider the distribution of the total transverse energy H_T in Z/W production with at least one jet as measured by Atlas [12, 4] and shown in Figure 4.

We can see that the NLO description is not good for large values of H_T. A closer investigation shows that this is due to the fact that higher multiplicity processes are important for large values of H_T. This can be seen in Figure 5 from Ref. [14]. This is one example illustrating that high

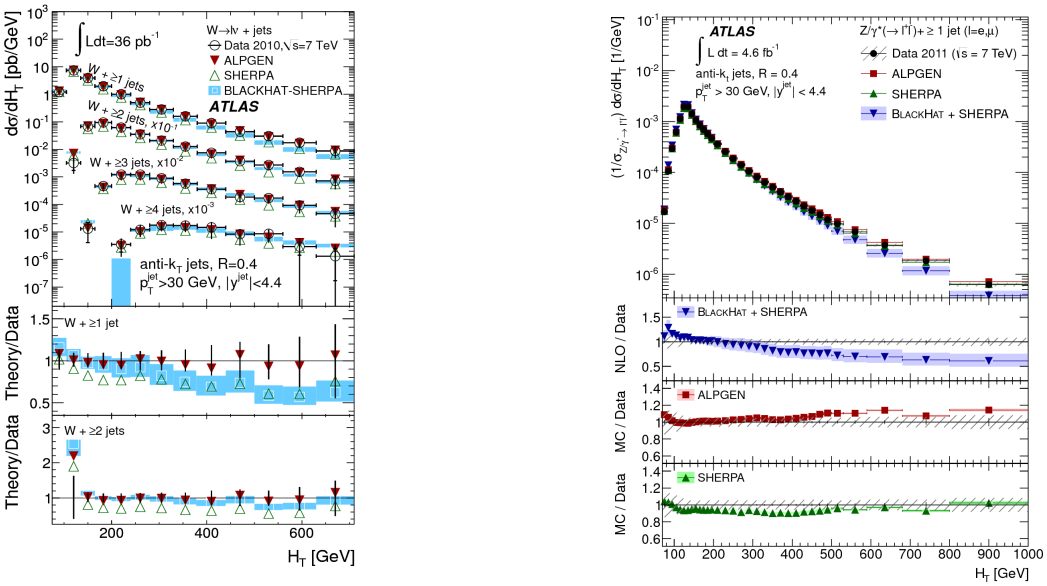


Figure 4. H_T distribution in $W + 1$ jet (left) and $Z + 1$ jet (right) as measured by Atlas.

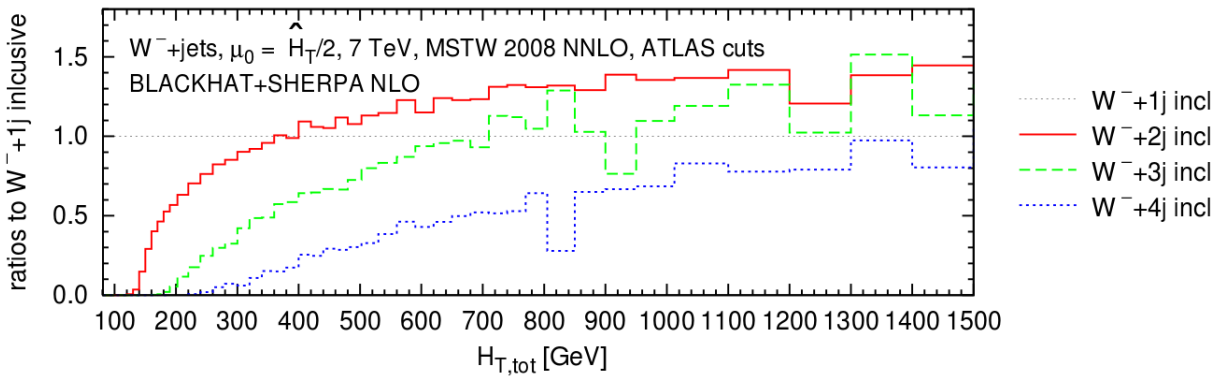


Figure 5.

multiplicity processes can be important in the tails of some distributions, where many searches for new physics focus their efforts.

One possible way of including higher multiplicity contributions to the distribution is add the contribution of exclusive NLO predictions for different multiplicities. By exclusive is meant here to put a cut on the transverse momentum of the additional $(n + 1)$-th jet in each $V + n$ jets sample to avoid double-counting. This method is very crude and has the formal accuracy of the first sample, so from that point of view no improvement is to be expected, but for observables for which high multiplicity processes are relevant this can yield an improvement. Figure 6 taken from ref. [13] shows how higher multiplicity processes gain a larger importance as one moves towards the tail of the distribution.

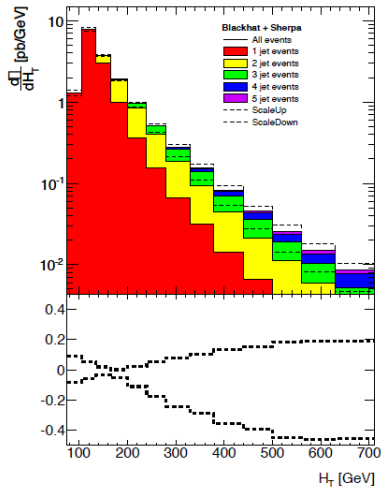


Figure 6. Various multiplicity contributions to the H_T distribution.

Acknowledgements

DM's work was supported by the Research Executive Agency (REA) of the European Union under the Grant Agreement number PITN–GA–2010–264564 (LHCPhenoNet).

References

[1] V. Khachatryan *et al.* [CMS Collaboration], arXiv:1408.3104 [hep-ex].
[2] V. Khachatryan *et al.* [CMS Collaboration], arXiv:1406.7533 [hep-ex].
[3] G. Aad *et al.* [ATLAS Collaboration], JHEP **1307** (2013) 032 [arXiv:1304.7098 [hep-ex]].
[4] G. Aad *et al.* [ATLAS Collaboration], Phys. Rev. D **85** (2012) 092002 [arXiv:1201.1276 [hep-ex]].
[5] Z. Bern, L. J. Dixon, F. Febres Cordero, S. Höche, H. Ita, D. A. Kosower, D. Maître and K. J. Ozeren, Phys. Rev. D **88**, 014025 (2013) [arXiv:1304.1253 [hep-ph]].
[6] H. Ita, J. Phys. A **44** (2011) 454005 [arXiv:1109.6527 [hep-th]].
[7] S. Badger, H. Frellesvig and Y. Zhang, J. Phys. Conf. Ser. **523** (2014) 012061 [arXiv:1310.4445 [hep-ph]].
[8] R. Brun and F. Rademakers, Nucl. Instrum. Meth. A **389**, 81 (1997); I. Antcheva *et al.*, Comput. Phys. Commun. **182**, 1384 (2011).
[9] Z. Bern, L. J. Dixon, F. Febres Cordero, S. Hche, H. Ita, D. A. Kosower and D. Maitre, Comput. Phys. Commun. **185** (2014) 1443 [arXiv:1310.7439 [hep-ph]].
[10] D. Britzger *et al.* [fastNLO Collaboration], arXiv:1208.3641 [hep-ph].
[11] M. Sutton, T. Carli, D. Clements, A. Cooper-Sarkar, C. Gwenlan, G. P. Salam, F. Siegert and P. Starovoitov, PoS DIS **2010** (2010) 051.
[12] G. Aad *et al.* [ATLAS Collaboration], JHEP **1307** (2013) 032 [arXiv:1304.7098 [hep-ex]].
[13] J. Alcaraz Maestre *et al.* [SM AND NLO MULTILEG and SM MC Working Groups Collaboration], arXiv:1203.6803 [hep-ph].
[14] D. Matre and S. Sapeta, Eur. Phys. J. C **73** (2013) 2663 [arXiv:1307.2252].
[15] J. R. Andersen *et al.* [SM and NLO Multileg Working Group Collaboration], arXiv:1003.1241 [hep-ph].