Experiences on QCD Monte Carlo simulations: a user point of view on the inclusive jet cross-section simulations

Paolo Francavilla
INFN Sez. di Pisa, University of Pisa, School of Graduate Studies G. Galilei
CERN
E-mail: paolo.francavilla@cern.ch

Abstract. In the last years, important progresses in the theoretical description of the QCD high $p_T$ processes have been carried out. In this proceeding, a review of the tools and techniques used to simulate QCD cross sections will be presented from a user point of view. The benchmark process selected for the discussion is the inclusive jet cross section. The proceeding will focus on the uncertainties of the Next to Leading Order (NLO) cross sections, on the strategies adopted to correct for the non-perturbative effects such as the hadronization and the underlying event, and the new techniques derived during the last years to incorporate in a coherent way the NLO matrix elements in the Monte Carlo generators.

1. Introduction
The imminent advent of the measurements at the Large Hadron Collider (LHC) guided important progresses to automatize the theoretical prediction of the Standard Model processes in hadron colliders in the last years. Some of them are oriented at the development of efficient methods to estimate the Next to Leading Order (NLO) - or in some cases even the higher orders - contributions for a big number of different processes. Other important progresses are aimed at merging the exact perturbative expansion approach to approximate models, trying to incorporate in the final prediction part of the contributions of all the higher orders. This proceeding is not a complete review of these techniques, but a report of the hands-on experiences in estimating the theoretical prediction for a specific process.

The process under investigation is the jet production at Large Hadron Collider (LHC), in proton-proton collisions. The jets are a spray of collimated particles produced in high energy collisions. Their production is the dominant high transverse-momentum ($p_T$) process at the hadron colliders, and it gives us the first glimpse of physics at the TeV scale.

Furthermore, the LHC collaborations have recently reported the measurement of the inclusive jet cross section with the 2010 data [1, 2, 3]. This gives the possibility to directly test the predictions against the measured spectrum.

2. Fixed order predictions
Fixed-order predictions, which involve the first terms in the Quantum Chromo-Dynamics (QCD) perturbative expansion for a given observable, are conceptually quite simple: it is easy to
state which contributions are included, and as one includes further orders in the expansion one can reasonably hope to see systematic improvement in the accuracy of ones predictions. The expansion is in powers of the strong coupling constant $\alpha_S$. Usually, only the first 2 or 3 terms in the expansion are known. The first term is known as the Leading Order (LO).

2.1. Leading Order
The LO theoretical prediction for the inclusive jet cross section is done calculating the QCD $2 \rightarrow 2$ matrix elements.

In this calculation, the initial protons are replaced by the initial partons (quarks or gluons). The two initial partons carry a fraction of the initial protons momentum, which is statistically described by the proton Parton Density Functions (PDF). The spatial momentum of these incoming partons have only a longitudinal component (parallel to the beam axis), and no transverse momentum. At the LO the proton remnants are not taken into account, with the assumption that their contribution to the hard scattering is minimal.

Once the initial partons are extracted from the proton, the QCD amplitude for the $2 \rightarrow 2$ parton processes can be calculated. The number of sub-processes which must be taken into account is high, and they can be divided according to the initial partonic configuration (i.e. gluon-gluon, quark-gluon, quark-quark).

The final state of the LO calculation consists of two outgoing partons. They are back-to-back in the azimuthal angle, with identical $p_T$ (to obey the momentum conservation). For most of the common jet definitions, this is a sufficient condition to avoid any merging of the two partons in one final jet. Each final state parton is identified as a final jets, with no differences due to the jet definition in use.

A final state with only two partons is clearly a rough approximation, because the multiplicity of particles in the measured events can be really high (much bigger than 2 particles). The discrepancy in the multiplicity of the final state objects is reduced by the use of a jet algorithm. Nevertheless, the measured jets are complex QCD observables, and different jet definitions produce jets with different properties in the measurement, reinforcing the need of a more accurate description of the final state system in the theoretical prediction.

Another weak point in the LO prediction is the intrinsic uncertainty in the perturbative expansion. This is a general aspect for the perturbative calculations, and will be discussed in the next section.

2.2. Higher orders
The hope to improve the accuracy of the perturbative predictions pushes to provide the calculation of the higher order terms in the expansion. The next term, after the leading, is the Next to Leading Order (NLO). To estimate the NLO cross section, two different and divergent terms must be calculated: the real emission of an extra parton, and the virtual contribution. These two contributions diverge in some configurations (infrared and collinear limits), but if they are summed together in the proper way, they counterbalance. This requires an infra-red and collinear safe observable. For the inclusive jet cross section, the jet definition plays this role. This proceeding does not report a description of the problems caused by the selection of an infra-red and/or collinear unsafe jet algorithm (it can be found in several references, i.e. [7]).

By using an infra-red and collinear safe jet definition, the NLO prediction for the inclusive jet cross section can be calculated with NLOJET++ [4] (Figure 1). In this figure, proton-proton collisions at $\sqrt{s} = 7$ TeV are simulated by using the CTEQ6.5 PDF [5], and the anti-$k_T$ with $R=0.6$ to reconstruct the jets [6]. The renormalization and the factorization scales are equal to the leading jet $p_T$. 



2.3. Comparing Leading Order and Next to leading Order

The use of the second term in the expansion could change the prediction in some special point of the phase space. To check the difference between the leading order and the next to leading order prediction, the ratio of the two cross sections in bins of rapidity and $p_T$ is shown in Figure 2.

In this plot, two different radii for the anti-$k_T$ jet algorithm are used: $R=0.6$ on the left, and $R=0.4$ on the right. To reduce the effect of the PDF, the CTEQ 6.5 is used for both the LO and for the NLO predictions. In most of the phase space, the ratio is close to 1, confirming that the contribution of the second term is small, for the selected scales. In some regions, close to the end of the phase space, the difference is bigger. The radius $R=0.4$ seems to be slightly more
affected by the contribution of the second order.

2.4. Stability of the fixed order predictions

The LO (NLO) prediction is the perturbative expansion truncated to the first (second) non trivial order. The accuracy of a truncated expansion (in this case at the first or second order) is good if the contributions from the higher orders are small.

A rough estimate of the goodness of the truncated series can be obtained from the stability of the expansion, when shifting the renormalization ($\mu_r$) and factorization ($\mu_f$) scale. These scales are usually representative of the energy scale of the process. The final prediction should not depend on the selection of these scales (since there is no correspondence in the measured observable), but a spurious dependence remains in the truncated perturbative expansion. One can reasonably hope to see a smaller dependence and a better stability in including further orders in the expansion.

No strong and fundamental argument fixes these scales to a precise and defined energy for a certain observable, and different definitions are possible. For the inclusive jet cross section, the scales can be of the order of the $p_T$ of the leading jet and a possible variation can be $0.5 \times p_{T,\text{max}} < \mu_r, \mu_f < 2 \times p_{T,\text{max}}$.

To check the stability of the inclusive jet cross section, its dependence as a function of the renormalization and factorization scale variations is shown in Figure 3 on the left. This figure shows the dependence of the inclusive jet cross section $\int d\eta d\sigma/dp_T$ integrated in the $p_T$ interval between 170 GeV and 210 GeV and in the pseudo-rapidity range $|\eta| < 5.0$. The calculation, performed with NLOJET++, is done for proton-proton collisions at $\sqrt{s} = 14$ TeV$^1$. The PDFs for the NLO calculation is the CTEQ6.1 and for the LO the CTEQ6.LL.

The scales for the calculation is proportional to the $p_T$ of the leading jet, and three different schemes of variation are used for the LO and the NLO predictions:

$$\mu_f = \mu_r = \alpha \times p_{T,\text{max}}$$
$$\mu_r = \alpha \times p_{T,\text{max}}, \mu_f = p_{T,\text{max}}$$
$$\mu_f = \alpha \times p_{T,\text{max}}, \mu_r = p_{T,\text{max}}$$

in which both the scales are simultaneously or independently shifted.

The shift of the scales for the LO prediction produces a drastic variation of the prediction, and the introduction of the next order in the expansion improves the stability of the result.

The envelopes of the three schemes of variation reported above, limiting $\alpha$ in the range $[0.5, 2]$ are shown as a function of $p_T$ in Figure 3.right. In this plot, the variations are normalized to the NLO prediction with $\mu_f = \mu_r = p_{T,\text{max}}$. The width of the bands is an estimate of the stability of the prediction, and it shows the reduction of uncertainty from approximately 30-40% at the LO to a 5-10% at the NLO.

As expected, the corrections due to the second leading order in the perturbative expansion improve the stability of the prediction.

In addition, the real emission of an extra final parton in the NLO is a first step to overcome the simplicity of the final state discussed above for the LO. The jets in the NLO calculation of the inclusive jet cross section have an internal structure, formed by one or two partons.

3. Some approximations: the parton showers

The NLO calculation provides an accurate prediction as long as the observables are infra-red and collinear safe (i.e. jets), but it is not describing of the production of particles (i.e. pions or kaons) in the jets.

$^1$ The major features of the LO and of the NLO stability of the cross sections due to renormalization and factorization scale variation marginally depend on the center of mass energy.
Figure 3. Dependence of the LO and NLO inclusive jet cross section on the renormalization and factorization scale. Left: variation of the cross section in the range in $p_T$ [170 GeV, 210 GeV] for different schemes of scale variation. Right: stability bands for the scale variation as a function of $p_T$ for the inclusive jet cross section, normalized to the NLO prediction with $\mu_f = \mu_r = p_T, max$.

The divergences in the real emission of an extra parton can be interpreted as the footprint of the high probability to emit gluons in the collinear and soft regions. With this idea, one can perform an approximated sum over all the perturbative orders to calculate the probability to emit collinear or soft gluons. Whereas the real emission of one extra gluon have a bare infinity if one took the collinear and soft limit, in this approximated calculation, the emission probability is simply bounded to be between 0 and 1.

Several code were proposed and used to simulate this approximated emission of partons [8, 9, 10, 11, 12, 13]. These code recursively simulate the emission of multiple partons, producing a multi parton final state. This gives a parton shower event. To get a prediction for the inclusive jet cross section, the parton shower can be merged/matched with the LO or the NLO matrix elements. Different Monte Carlo generators can produce parton shower events for the jet production [8, 9, 10, 11, 13] starting from the LO matrix element.

The results for the inclusive jet cross section simulated with Pythia 6.421 [8] at the level of the parton shower are shown in Figure 4,left. In this plot the cross section is calculated in different regions of rapidity for the anti-$k_T$ jet algorithm with $R=0.6$. The impact of the parton shower for these two radii is shown in Figure 4,right. In this plot, the predictions of the inclusive jet cross section for the anti-$k_T$ jet with $R=0.6$ and $R=0.4$, are divided by the pure LO prediction. Since the pure LO prediction does not depend on the parameter $R$, the difference between $R=0.4$ and $R=0.6$ is due to the parton shower. The anti-$k_T$ jet with $R=0.6$ is producing an higher cross section, simply for the effect of the radiation of gluons.
Figure 4. Left: Inclusive jet cross section estimated by Pythia at the end of the parton shower for the anti-$k_T$ jet algorithm with $R=0.6$. Right: Comparison of the inclusive jet cross section at the end of the parton shower simulation for anti-$k_T$ with $R=0.4$ and $R=0.6$. The simulation is performed in the central rapidity region $|y|<0.3$ with the Pythia Monte Carlo simulation. The two estimates are normalized to the LO parton level prediction obtained by Pythia.

in the parton shower.

3.1. Non perturbative effects

Even if the transition between the parton level and the hadron level (hadronization) is not completely understood, different models have been developed. The Monte Carlo generators can produce final states which are formed by hadrons. In addition to the hadronization, the Monte Carlo generators have models which take into account the effects due to the proton remnants, which are not directly taking part to the hard scattering. The effects of these "spectators" (which are color connected with the rest of the event, and could eventually produce a second hard scattering) are named underlying event effects. To reproduce as well as possible the properties of the particle production at hadron colliders, these models are tuned to some measured observables.

The non perturbative effects have an impact on the final cross section. To separate these effects from the parton shower effects, the cross section at hadron level $\sigma_{Had}$ is compared to the one obtained at the end of the parton shower simulation $\sigma_{PS}$. The ratio:

$$C_{NP} = \frac{\sigma_{Had}}{\sigma_{PS}}$$

is called non perturbative correction.

Since different models of parton showers, of hadronization and of underlying event exists, different corrections can be obtained. Figure 5 shows the non perturbative corrections for the inclusive jet cross section at $\sqrt{s} = 7$ TeV for the anti-$k_T$ jet algorithm with $R=0.4$ and $R=0.6$ in the region in rapidity $|y|<0.3$. Different tunes and Monte Carlo generators (Pythia [8] and Herwig [10]) are used to derive these results\(^2\).

\(^2\) The results are obtained by running the Monte Carlo generators with the Agile interface in the Rivet framework.
The corrections are really close to 1 for $p_T$ above 200 GeV, which means that in this regime the effect of the hadronization and underlying event is small. The effect of the hadronization and underlying event depends on the radius $R$ of the anti-$k_T$ jets.

4. **A solid prediction?**

Even if the NLO prediction is not producing a complete list of particles in the final state, and the jet substructure is produced by the emission of only one extra parton, the calculation is usually compared with the measurement. The non perturbative correction, shown in Figure 5 can be used to overcome the lack in the description of the hadronization and the underlying event. These corrections are used as a bin-by-by factors to be multiply to the NLO cross section. These factor are useful for the comparison of the measured cross section with the theoretical predictions, and to perform QCD fits (i.e. the PDF fits). To avoid any source of ambiguity, an agreement on how to define the factors, on how to estimate their uncertainties and on how to use them would be really useful. For example, in the region where the correction are big, (i.e. 1.5) the perturbative calculation is not accurate, and the contribution from the non perturbative physics is important. Can the NLO prediction, corrected with the non perturbative factor, be trusted as an accurate estimate of the cross section in these regions? Which is the proper way to estimate the theoretical uncertainties?

4.1. **A coherent prediction**

To overcome the weak points of using a non perturbative correction, one needs to perform a NLO calculation, merged (or matched) with a parton shower which can be interfaced with the hadronization and the underlaying event models.

In the last years several important progresses were done in this direction. The two main formalisms to perform this coherent predictions are MC@NLO [14] and Powheg [15]. Only recently the dijet production was introduced in the Powheg formalism, and it can be used as a natural benchmark to test the method and to check the predictions.

Other processes are calculated with these formalisms, but only few of them have two real final partons in the leading order contribution. The checks on the jet production is important.

**Figure 5.** Non perturbative corrections for the inclusive jet cross section for different Monte Carlo generators and tunes. Left: anti-$k_T$ jet algorithm with $R=0.6$. Right: anti-$k_T$ jet algorithm with $R=0.4$. 

Jet Reconstruction and Spectroscopy at Hadron Colliders IOP Publishing
Journal of Physics: Conference Series 323 (2011) 012016 doi:10.1088/1742-6596/323/1/012016
in the optic of moving to other processes with two or more jets in association with other objects in the final state.

5. Conclusion
In this proceeding, several tools to estimate the inclusive jet cross section at hadron colliders have been discussed. The goal of the discussion is not a precise report of all the different aspect of the available calculations, but the review of some aspects of the different approaches.

The advent of the NLO predictions, which give a more stable prediction with respect to the LO, merged with the parton shower approach, which approximately describes the emission of extra partons, is one of the main improvement of the last years. These simulations are easily interfaced with the hadronization models and underlying event models, which describe the non perturbative part of the proton-proton scattering.

This new approach to obtain a theoretical prediction for QCD processes is really promising, overcoming the weak points of the pure NLO prediction, such as the lack of description of the final state, and the necessity to derive the non perturbative corrections.

Acknowledgments
The author would like to thank the organizer of the conference for the precious opportunity to discuss these aspects. The author would like to thank C. Roda, J. Butterworth, P. Loch, P. Starovoitov, T. Carli, G. Salam, K. Perez, Z. Nagy, M. Seymour, P. Nason, C. Oleari, E. Re, P. Skands T. Sjöstrand and R. Torre for their help and for the useful discussions.

References
[1] ATLAS Collaboration, Measurement of inclusive jet and dijet cross sections in proton-proton collisions at 7 TeV centre-of-mass energy with the ATLAS detector, Eur. Phys. J. C71 (2011) 1512, arXiv:1009.5908 [hep-ex].
[2] ATLAS Collaboration, Measurement of inclusive jet and dijet cross sections in proton-proton collision data at 7 TeV centre-of-mass energy using the ATLAS detector, ATLAS-CONF-2011-047.
[3] CMS Collaboration, Measurement of the Inclusive Jet Cross Section in pp Collisions at $\sqrt{s} = 7$ TeV, arXiv:1105.2087v1 [hep-ex].
[4] Z. Nagy, Next-to-leading order calculation of three jet observables in hadron hadron collision, Phys. Rev. D68 (2003) 094002, arXiv:hep-ph/0307268.
[5] W. K. Tung, H. L. Lai, A. Belyaev, J. Pumpkin, D. Stump and C. P. Yuan, JHEP 0702 (2007) 053 [arXiv:hep-ph/061254].
[6] M. Cacciari, G. Salam, and G. Soyez, The anti-kt jet clustering algorithm, JHEP 0804 (2008) 063, arXiv:0802.1189.
[7] C. Buttar et al., Standard Model Handles and Candles Working Group: Tools and Jets Summary Report, arXiv:arXiv:0803.0678 [hep-ph].
[8] T. Sjostrand, S. Mrenna, and P. Skands, PYTHIA 6.4 physics and manual, JHEP 05 (2006) 026, hep-ph/0603175.
[9] T. Sjostrand, S. Mrenna and P. Skands, Comput. Phys. Commun. 178 (2008) 852 [arXiv:0710.3820 [hep-ph]]; http://home.thep.lu.se/~torbjorn/Pythia.html.
[10] G. Corcella et al., JHEP 0101 (2001) 010 [arXiv:hep-ph/0011363]; http://hepwww.rl.ac.uk/theory/seymour/herwig/.
[11] M. Bahr et al., Eur. Phys. J. C58 (2008) 639 [arXiv:0803.0883 [hep-ph]]; http://projects.hepforge.org/herwig/.
[12] L. Lonnblad, Comput. Phys. Commun. 71 (1992) 15.
[13] T. Gleisberg, S. Hoche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert and J.Winter, JHEP 0902 (2009) 007 [arXiv:0811.4622 [hep-ph]]; http://projects.hepforge.org/sherpa/.
[14] S. Friction, F. Stoeckli, P. Torrielli, B. R. Webber, and C. D. White, The MC@NLO 4.0 Event Generator, arXiv:1010.0819 [hep-ph].
[15] S. Alioli, P. Nason, C. Oleari, and E. Re, A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX, arXiv:1002.2581 [hep-ph].
[16] S. Alioli, K. Hamilton, P. Nason, C. Oleari, and E. Re, Jet pair production in POWHEG, arXiv:1012.3380 [hep-ph].