Hard QCD — Theoretical Aspects

Giulia Zanderighi
Rudolf Peierls Centre for Theoretical Physics
1 Keble Road, OX1 3PN, Oxford, UK

1 Introduction

ATLAS and CMS had collected around 45 pb^{-1} in 2010, 700 pb^{-1} by June 2011, the time of this conference, and more than 3 fb^{-1} by the write-up of these proceedings. With the 2010 and early 2011 data, all major Standard Model (SM) processes have already been re-established, including single-top and di-boson production, challenging measurements (because of the small/large cross sections/backgrounds) that have been performed at the Tevatron only in recent years. By now, we have entered a new territory in the search of physics beyond the SM (BSM) with sensitivities already well exceeding those of LEP and the Tevatron.

One important question concerns the role of QCD for LHC measurements and new-physics searches. Understanding how QCD works is essential in order to make accurate predictions for both the signal and background processes. This typically requires complex calculations to higher order in the perturbative expansion of the coupling constant. Understanding QCD dynamics can however also help reducing backgrounds and sharpen the structure of the signal. This can for instance be achieved by designing better observables, by employing appropriate jet algorithms, by using jet-substructure, or by exploiting properties of boosted kinematics. Finally, once discovery is be made, QCD will be crucial to extract the properties (masses, spins, and couplings) of the new states found. Therefore, at the LHC, no matter what physics you do, QCD will be part of your life.

It is interesting to first recall a recent measurement that was the origin of considerable excitement. In April 2011, CDF reported the observation of a peak in the m_{jj} distribution in W + dijet events [1]. The first measurement had a 3.2σ significance, and was based on 4.3 fb^{-1}. Subsequently, more data (7.3 fb^{-1}) has been analyzed, leading to a significance of more than 4σ [2]. Since then, a large number of tentative BSM explanations appeared on the arXiv, along with few SM analysis that address the question of whether this effect can be attributed to a mismodelling of one of the SM backgrounds (in particular single top) [3, 4, 5]. The excitement was curbed right at the time of this conference, when D0 announced that it did not confirm the excess seen by CDF [6]. It is yet unclear what the reasons for the discrepancy between CDF and D0 findings are, if any. However, this example demonstrates that even in the case where one identifies a mass peak in the tail of a distribution (a scenario that was considered "an easy discovery") a robust control of SM backgrounds remains mandatory, in particular when the shape of the backgrounds is one of the issues.

The important question becomes then what are the tools at our disposal to make precise predictions, and whether we have the solid control of backgrounds that is needed in order to claim discoveries. In the following, I will review the current status of our tools, and will discuss a few recent ideas to further improve on the way we perform technically challenging calculations.

2 Perturbative tools

2.1 Monte Carlo and leading order matrix elements

Every analysis at the LHC uses a Monte Carlo (MC) program for the simulation of the signal process of interest, for the simulation of the backgrounds, for subtracting the underlying event and the non-perturbative contributions, and/or for efficiency studies and modeling of the detector response. The current level of sophistication is such that (essentially) not a single study relies on `Pythia/Herwig` alone. It is well understood that in multi-parton processes it is important to describe the multiple hard QCD radiation at least using exact matrix elements, employing for instance `Alpgen` [7], `Madgraph` [8], or `Sherpa` [9].

Since experimental studies rely heavily on all these leading-order (LO) tools, there is continuous progress in their development, and the `Herwig/Pythia` codes that we have today bear little resemblance to their original version of the '80s. In particular, in `Pythia` 8.1 [10] (a C++ code) there is a new fully interleaved p_t-ordered MPI+ISR+FSR evolution (the original mass-ordered evolution is not supported any longer), a richer mix of underlying event processes (γ, J/Ψ, DY), the possibility to select two hard interactions in the same event, an x-dependent proton size in the MPI framework, the full hadron–hadron machinery for diffractive systems, several new processes in and beyond the SM, and various other new features. `Herwig++` [11] (the current version is 2.5.1) has new next-to-leading order (NLO) matrix elements, including weak boson pair production, a colour reconnection model, diffractive processes, additional models of BSM physics, and new LO elements for hadron-hadron, lepton-lepton collisions and photon-initiated processes. `Sherpa` [9] (1.3) has improved integration routines in Comix, a simplified kinematics reconstruction algorithm of the parton shower (PS), leading to numerically more stable simulations, HepMC output for NLO events and various other improvements/bug-fixes. `Madgraph` [8] (v5) has a completely new diagram generation algorithm, which makes optimal use of model-independent information, has an efficient decay-chain package, and a new library for the colour calculations. Altogether, there is a continuous, fast progress in various directions. So far, it is amazing how well these tools work, once the normalization is fixed using data. A very recent comparison of data with `Alpgen` up to seven jets (a control region for BSM searches) can be found in [12]. These LO programs will undergo a stress test in the coming years.

2.2 The NLO revolution

Theorists like to advertise NLO by using the reduction of scale (theory) uncertainties as an argument. However, the strongest argument in support of NLO calculations is their past success in accurately describing LEP and Tevatron data. *Recent* revolutionary ideas in the way NLO computations are performed include sewing together tree-level amplitudes to compute loop amplitudes (using on-shell intermediate states, cuts, unitarity ideas, ...) [13], the OPP algorithm, i.e. an algebraic way to extract coefficients of master integrals by evaluating the amplitudes at specific values of the loop momentum [14], and D-dimensional unitarity, i.e. a practical numerical tool to evaluate full amplitudes, including the rational part, with unitarity ideas [15]. For a pedagogical review on unitarity methods see [16].

These methods led in the last 2 to 3 years to a number of $2 \to 4$ calculations at hadron colliders. These include $W + 3$ jets [17, 18], $Z + 3$ jets [19], $t\bar{t}b\bar{b}$ [20], $t\bar{t} \to W^+W^-b\bar{b}$ [21], $W^+W^+ + 2$ jets [22], $W^+W^+ + 2$ jets [23], $t\bar{t} + 2$ jets [24], and, via crossing of the process $Z + 3$ jets, $e^+e^- \to 5$ jets [25].

Feynman diagram methods have also been applied successfully to $2 \to 4$ calculations, this is for instance the case for quark-induced $b\bar{b}b\bar{b}$ [26], $t\bar{t}b\bar{b}$ [27], and $W^+W^-b\bar{b}$ [28] production. Note that only few years ago performing these type of calculations with Feynman diagrams was considered an impossible task.

Given that both Feynman diagram and unitarity based methods allowed us to compute $2 \to 4$ processes it might be unclear where the revolution advocated in the heading of the subsection lies in. The revolution, I believe, is not yet in the applications that we see today, rather in the prospect for low-cost fully computer-automated NLO calculations even beyond $2 \to 4$ in the near future. Indeed, two $2 \to 5$ processes have already been computed at NLO, namely $W + 4$ jets [29] and $Z + 4$ jets [30].[1] As far as the full automation is concerned, let me highlight only one interesting recent general approach [31]. It is based on Feynman diagrams, uses the OPP procedure for virtual calculation, and FKS subtraction of divergences, together with clever and efficient procedures to deal with instabilities. More improvements and refinements are to be expected soon. At present there is no public code, instead there are plans to provide N-tuples.

[1] In both cases the leading colour approximation has been used, and six-quark processes have been neglected. Both approximations are expected to give rise to small (percent) corrections only.

2.3 Merging NLO and Parton showers

While NLO predictions provide relatively accurate results for inclusive cross sections, they do not furnish an exclusive description of the final state that can be compared with actual particles in the detectors, as Monte Carlo programs do. It is therefore useful to combine the best features of both approaches. Two public frameworks exist for this purpose, namely `MC@NLO` [32] and `POWHEG` [33]. These tools are almost 10 years old now, and since their conception a long list of processes has been implemented in both frameworks.

In particular, recently the `POWHEG BOX` was released [34], which is a general framework for implementing NLO calculations in shower MC programs according to the `POWHEG` method. The user essentially only needs to provide a simple set of routines (Born, colour-correlated Born, virtual, real, and phase space) that are part of any NLO calculation. The first $2 \to 4$ process that has been implemented in the `POWHEG BOX` is $pp \to W^+W^+ +$ 2 jets [35]. This is a relatively simple $2 \to 4$ process since the cross section is finite without any cut on the jets. As expected, for inclusive observables there are only minor differences between pure NLO and `POWHEG`+PS, but for exclusive observables, depending on the details of the observable definition, there can be important differences.

`aMCNLO` is a novel approach to a complete event generation at NLO. It has been used for the calculation of scalar and pseudo-scalar Higgs production in association with a $t\bar{t}$ pair [36], and $W/Zb\bar{b}$ [37]. As yet, no public code is available.

2.4 `MENLOPS` and `LoopSim`

`MENLOPS` [38, 39] is a method to further improve on NLO + PS predictions with matrix elements involving more partons in the final state. For example, for W production it includes, as in `MC@NLO` or `POWHEG`, W production at NLO, the PS, but also $W +$ $1, 2, 3 \ldots$ jets using exact matrix elements. Roughly speaking, it uses a jet-algorithm to define two different regimes, and then corrects the 1-jet fraction using exact matrix elements and the 2-jet fraction using the NLO K-factor. This achieves NLO quality accuracy for inclusive quantities but an improved sensitivity to hard radiation and multi-parton kinematic features.

A further recent theoretical development is `LoopSim`. If one considers the process $W + 1$ jet, the three observables $p_{t,Z}$, $p_{t,j}$, and $H_{T,jets} = \sum_j p_{t,j}$ are identical at LO. However, at NLO $p_{t,Z}$ has a moderate K-factor ($\lessapprox 2$), $p_{t,j}$ has a large K-factor (~ 5) and $H_{T,jets}$ has a giant K-factor (~ 50). The reason for the very large K factors in the last two observables is that the NLO result is dominated by configurations where there are two hard jets and a soft W (these are enhanced by electroweak logarithms), additionally there is an important enhancement coming from incoming qq channels. `LoopSim` [40] is a procedure that uses a sequential algorithm, close to the Cambridge/Aachen one, to determine the branching history, "loops" over soft particles (i.e. they are removed from the event and the residual event is adjusted), and it uses a unitary operator to cancel divergences. In essence, this is a way to extend a calculation that is exact at a given order in perturbation theory, in an approximate way to higher orders. The procedure is expected to be more accurate the larger the K-factor is. On the same line as `MENLOPS` and `LoopSim`, one might expect other extensions of the MLM/CKKW matching procedure in the near future.

2.5 (Approximate) NNLO

Drell-Yan is the most accurately predicted process at the LHC, and state-of the art codes are described in [41, 42]. NNLO calculations have been available since many years, and now that precise LHC data have been compared to those predictions, one can not but praise the impressive agreement between NNLO theory and experiment (see e. g. [43]). In particular, not only cross sections have been measured, but also W/Z properties, e.g. the weak-mixing angle [44]. More details can be found in the proceedings of S. Forte and F. Petriello. I will not mention any recent developments in Higgs physics, since we had various dedicated contributions at this conference. For an update on recent theoretical results for SM Higgs productions, I refer the reader to the proceedings of F. Petriello and F. Piccinini.

The top is the most interesting SM quark. Its large mass implies a large Yukawa coupling, which causes the top to be a prominent decay product in many BSM models. LHC data have already been successfully compared to approximate NNLO predictions [45, 46], however various approximate NNLO predictions, based on a threshold resummation, do not fully agree within quoted uncertainties [47, 48, 49, 50], so a full NNLO calculations is highly desirable. A better perturbative control of the top production cross section is also important to further constrain gluon parton distribution functions, to have an accurate extraction of the top mass from the cross section, and to improve our perturbative control of the top forward-backward asymmetry. In fact, an almost 3σ deviation from the SM is observed by CDF, which becomes a 4.2σ effect in the high-mass region, $M_{t\bar{t}} > 450$ GeV. This effect, seen both by CDF and D0 (though their results are not fully compatible), seems to be a persistent over many measurements, and the enhancement in the high mass region is particularly tantalizing. However, one has to bear in mind that this a difficult measurement given the presence of neutrinos in the final state, the combinatorics in the reconstruction of the tops, and the limited statistics at the Tevatron. Nevertheless, various suggestions have been made recently to explain the asymmetry in terms of BSM physics, but all proposals face the problem that they have to preserve good agreement with symmetric observables, respect dijet bounds and/or they must evade the strong bounds on like-sign top decays. Fervid activity is therefore currently devoted towards a complete NNLO calculation of $t\bar{t}$ production (see [51] and references therein).

2.6 Jet algorithms

For a long time, infrared (IR) unsafe algorithms were used at hadron colliders, with several "patches" to minimize the effect of the IR-unsafety. At the LHC, both ATLAS and CMS have adopted as default the anti-k_t algorithm [52]. Given that this algorithm was proposed only three years ago, it shows how flexible experimentalists are today in adopting new, successful ideas.[2] Using this algorithm both collaborations have already explored up to the 4 TeV region and could place constraints on various BSM models, in particular those models that would give rise to a resonance in the M_{jj} distribution (such as axigluons, massive colored bosons, black-holes, ...).

Other IR safe algorithms like the Cambridge-Aachen or SISCone are in use as well. This is particularly the case for studies that exploit the fact that when a massive boosted object decays, which gives rise to a "fat jet" with a none-trivial jet-substructure. Looking at the internal structure of these jets using jet-grooming techniques like filtering, pruning or trimming has a huge potential to make discoveries "easier" [53]. The potential of these studies has been demonstrated in several examples, however sophisticated jet studies are still a young field, and as to now there are no precise rules on how to make discoveries easier.

3 Conclusions

QCD is a dynamic field, there has been a spectacular progress in recent years, which includes amazing technical achievements (higher multiplicities and/or loops), clever merging procedures to catch best features of different calculations, ingenuity in refining observables, sophisticated techniques for looking inside jets, and spectacular formal developments (IR/UV structures, N=4 or N=8 SYM, twistors, Wilson loops amplitudes, symbols ...) that I did not have time to mention.

The SM has been already re-established at the LHC, we are now waiting eagerly for signs of new physics. We have the right tools to make the most out of observations at the LHC, but is it more important than ever to choose the right observables and tools for a given physics analysis.

References

[1] T. Aaltonen *et al.* [CDF Collaboration], Phys. Rev. Lett. **106** (2011) 171801.

[2] http://www-cdf.fnal.gov/physics/ewk/2011/wjj/

[3] T. Plehn, M. Takeuchi, J. Phys. G **G38** (2011) 095006.

[2] A minor downside to this is that ATLAS and CMS use a different radius, the choices for ATLAS are 0.5 and 0.7, while for CMS they are 0.4 and 0.6.

[4] J. M. Campbell, A. Martin, C. Williams, Phys. Rev. **D84** (2011) 036005.

[5] Z. Sullivan, A. Menon, [arXiv:1108.4676 [hep-ph]].

[6] V. M. Abazov *et al.* [D0 Collaboration], Phys. Rev. Lett. **107** (2011) 011804.

[7] M. L. Mangano, *et al.*, JHEP **0307** (2003) 001.

[8] J. Alwall, *et al.*, JHEP **1106** (2011) 128.

[9] T. Gleisberg, *et al.*, JHEP **0902** (2009) 007.

[10] T. Sjostrand, S. Mrenna, P. Skands, Comput. Phys. Commun. **178** (2008) 852-867.

[11] S. Gieseke, *et al.*, [arXiv:1102.1672 [hep-ph]].

[12] G. Aad *et al.* [Atlas Collaboration], [arXiv:1110.2299 [hep-ex]].

[13] R. Britto, F. Cachazo, B. Feng, Nucl. Phys. **B725** (2005) 275-305.

[14] G. Ossola, C. G. Papadopoulos, R. Pittau, Nucl. Phys. **B763** (2007) 147-169.

[15] W. T. Giele, Z. Kunszt, K. Melnikov, JHEP **0804** (2008) 049.

[16] R. K. Ellis, Z. Kunszt, K. Melnikov, G. Zanderighi, [arXiv:1105.4319 [hep-ph]].

[17] R. K. Ellis, K. Melnikov, G. Zanderighi, Phys. Rev. **D80** (2009) 094002.

[18] C. F. Berger, *et al.*, Phys. Rev. Lett. **102** (2009) 222001.

[19] C. F. Berger, *et al.*, Phys. Rev. **D82** (2010) 074002.

[20] G. Bevilacqua, *et al.*, JHEP **0909** (2009) 109.

[21] G. Bevilacqua, *et al.*, JHEP **1102** (2011) 083.

[22] T. Melia, *et al.*, JHEP **1012** (2010) 053.

[23] T. Melia, *et al.*, Phys. Rev. **D83** (2011) 114043.

[24] G. Bevilacqua, *et al.*, [arXiv:1108.2851 [hep-ph]].

[25] R. Frederix, *et al.*, JHEP **1011** (2010) 050.

[26] T. Binoth, *et al.*, Phys. Lett. **B685** (2010) 293-296.

[27] A. Bredenstein, *et al.*, JHEP **1003** (2010) 021.

[28] A. Denner, *et al.*, Phys. Rev. Lett. **106** (2011) 052001.

[29] C. F. Berger, *et al.*, Phys. Rev. Lett. **106** (2011) 092001.

[30] H. Ita, *et al.*, [arXiv:1108.2229 [hep-ph]].

[31] V. Hirschi, *et al.*, JHEP **1105** (2011) 044.

[32] S. Frixione, *et al.*, [arXiv:1010.0819 [hep-ph]].

[33] P. Nason, JHEP **0411** (2004) 040.

[34] S. Alioli, P. Nason, C. Oleari, E. Re, JHEP **1006** (2010) 043.

[35] T. Melia, *et al.*, Eur. Phys. J. **C71** (2011) 1670.

[36] R. Frederix, *et al.*, Phys. Lett. **B701** (2011) 427-433.

[37] R. Frederix, *et al.*, JHEP **1109** (2011) 061.

[38] K. Hamilton, P. Nason, JHEP **1006** (2010) 039.

[39] S. Alioli, K. Hamilton, E. Re, [arXiv:1108.0909 [hep-ph]].

[40] M. Rubin, G. P. Salam, S. Sapeta, JHEP **1009** (2010) 084.

[41] S. Catani, *et al.*, Phys. Rev. Lett. **103** (2009) 082001.

[42] R. Gavin, *et al.*, Comput. Phys. Commun. **182** (2011) 2388-2403.

[43] S. Chatrchyan *et al.* [CMS Collaboration], [arXiv:1108.0566 [hep-ex]].

[44] S. Chatrchyan *et al.* [the CMS Collaboration], [arXiv:1110.2682 [hep-ex]].

[45] G. Aad *et al.* [ATLAS Collaboration], arXiv:1108.3699 [hep-ex].

[46] S. Chatrchyan *et al.* [CMS Collaboration], arXiv:1108.3773 [hep-ex].

[47] Y. Kiyo, *et al.* Eur. Phys. J. **C60** (2009) 375-386. [arXiv:0812.0919 [hep-ph]].

[48] V. Ahrens, *et al.* [arXiv:1103.0550 [hep-ph]].

[49] M. Beneke, *et al.* [arXiv:1109.1536 [hep-ph]].

[50] N. Kidonakis, [arXiv:1109.3231 [hep-ph]].

[51] M. Czakon, Nucl. Phys. B **849** (2011) 250 [arXiv:1101.0642 [hep-ph]].

[52] M. Cacciari, G. P. Salam, G. Soyez, JHEP **0804** (2008) 063.

[53] A. Abdesselam, *et al.*, Eur. Phys. J. **C71** (2011) 1661.