THE SM AND NLO MULTILEG AND SM MC WORKING GROUPS: Summary Report

Convenors: J. Alcaraz Maestre[1], *G. Heinrich*[2], *J. Huston*[3], *F. Krauss*[4], *D. Maître,*[4,5], *E. Nurse,*[6], *R. Pittau*[7]

Contributing authors: S. Alioli[8], *J. R. Andersen*[9], *R. D. Ball*[10], *A. Buckley*[11], *M. Cacciari,*[12,13], *F. Campanario*[14], *N. Chanon*[15], *G. Chachamis*[16], *V. Ciulli*[17], *F. Cossutti*[18], *G. Cullen*[19], *A. Denner*[20], *S. Dittmaier*[21], *J. Fleischer*[22], *R. Frederix*[23], *S. Frixione*[5,24,25], *J. Gao*[26], *L. Garren*[27], *S. Gascon-Shotkin*[28], *N. Greiner*[2], *J. P. Guillet*[29], *T. Hapola*[9], *N. P. Hartland*[10], *G. Heinrich*[2], *G. Hesketh*[6], *V. Hirschi*[24], *H. Hoeth*[4], *J. Huston*[30], *T. Ježo*[31], *S. Kallweit*[23], *K. Kovařík*[14], *F. Krauss*[4], *A. Kusina*[26], *Z. Liang*[26], *P. Lenzi*[1], *L. Lönnblad*[32], *J. J. Lopez-Villarejo*[33,5], *G. Luisoni*[4], *D. Maître*[4,5], *F. Maltoni*[34], *P. Mastrolia*[2,35], *P. M. Nadolsky*[26], *E. Nurse,*[6], *C. Oleari*[36], *F. I. Olness*[26], *G. Ossola*[37,38], *E. Pilon*[29], *R. Pittau*[7], *S. Plätzer*[39], *S. Pozzorini*[23], *S. Prestel*[32], *E. Re*[4], *T. Reiter*[2], *T. Riemann*[19], *J. Rojo*[5], *G. P. Salam*[5,39,12], *S. Sapeta*[4,12], *I. Schienbein*[31], *M. Schönherr*[4], *H. Schulz*[41], *M. Schulze*[42], *M. Schwoerer*[43], *P. Skands*[5], *J. M. Smillie*[10], *G. Somogyi*[19], *G. Soyez*[44], *T. Stavreva*[31], *I. W. Stewart*[45], *M. Stockton*[46], *Z. Szőr*[47], *F. J. Tackmann*[39], *P. Torrielli*[24], *F. Tramontano*[5], *M. Tripiana*[48], *Z. Trócsányi*[47], *M. Ubiali*[49], *V. Yundin*[50], *S. Weinzierl*[51], *J. Winter*[5], *J. Y. Yu*[31,26], *K. Zapp*[4]

[1] CERN, CH-1211 Geneva 23, Switzerland
[2] Max-Planck Insitut für Physik, Föhringer Ring, 6, D-80805 München, Germany
[3] Physics and Astronomy Department, Michigan State University, East Lansing, MI, 48824 USA
[4] IPPP, University of Durham, Science Laboratories, South Rd, Durham DH1 3LE, UK
[5] PH-TH Department, CERN, CH-1211 Geneva 23, Switzerland
[6] Universtiy College London, Gower Street, WC1E 6BT, London, UK
[7] Departamento de Física Teórica y del Cosmos, Universidad de Granada
[8] Ernest Orlando LBNL, University of California, Berkeley, CA 94720, USA
[9] CP^3-Origins, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark
[10] Tait Institute, School of Physics and Astronomy, University of Edinburgh, EH9 3JZ, UK
[11] PPE Group, School of Physics, University of Edinburgh, EH9 3JZ, UK
[12] LPTHE, UPMC Univ. Paris 6 and CNRS UMR 7589, Paris, France
[13] Université Paris Diderot, Paris, France
[14] Institut für Theoretische Physik, Universität Karlsruhe, KIT, 76128 Karlsruhe, Germany
[15] Institute for particle physics, ETH Zürich, Switzerland
[16] Paul Scherrer Institut, CH-5232 Villigen PSI, Switzerland
[17] Università di Firenze and INFN, Sezione di Firenze, via G. Sansone 1, 50019 Sesto F. (FI), Italy
[18] INFN, Sezione di Trieste, Via Valerio 2, 34127 Trieste, Italia
[19] Deutsches Elektronen-Synchrotron, DESY, Platanenallee 6, 15738 Zeuthen, Germany
[20] Universität Würzburg, Institut für Theoretische Physik und Astrophysik, 97074 Würzburg, Germany
[21] Albert-Ludwigs-Universität Freiburg, Physikalisches Institut, 79104 Freiburg, Germany
[22] Fakultät für Physik, Universität Bielefeld, Universitätsstr. 25, 33615 Bielefeld, Germany
[23] Institut für Theoretische Physik, Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland
[24] ITPP, EPFL, CH-1015 Lausanne, Switzerland
[25] On leave of absence from INFN, Sezione di Genova, Italy
[26] Department of Physics, Southern Methodist University, Dallas, TX 75275-0175, USA
[27] Fermi National Accelerator Laboratory, Batavia, USA
[28] Université de Lyon, Université Claude Bernard Lyon 1/Institut de Physique Nucléaire de Lyon IN2P3-CNRS
[29] LAPTH, 9 Chemin de Bellevue, B.P. 110, Annecy-le-Vieux 74951, France
[30] Department of Physics and Astronomy, Michigan State University, United States of America
[31] Laboratoire de Physique Subatomique et de Cosmologie, Université Joseph Fourier/CNRS-IN2P3/INPG,53 Avenue des Martyrs, 38026 Grenoble, France
[32] Department of Astronomy and Theoretical Physics, Lund University, Sölvegatan 14A, SE-22362 Lund, Sweden
[33] Madrid, Autonoma U. & IFT, 28049 Cantoblanco, Spain
[34] Centre for Cosmology, Particle Physics and Phenomenology (CP3) Université catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium

Operated by Fermi Research Alliance, LLC under Contract No. De-AC02-07CH11359 with the United States Department of Energy.

[35] Dipartimento di Fisica, Universita di Padova, Italy
[36] Università di Milano-Bicocca and INFN, Sezione di Milano-Bicocca, 20126 Milan, Italy
[37] Physics Department, New York City College of Technology, City University of New York, USA
[38] The Graduate School and University Center, City University of New York, USA
[39] Theory Group, Deutsches Elektronen-Synchrotron (DESY), D-22607 Hamburg, Germany
[40] Department of Physics, Princeton University, Princeton, NJ 08544, USA
[41] Physics Dept., Berlin Humboldt University, Germany
[42] High Energy Physics Division, Argonne National Laboratory, Argonne, IL 60439, USA
[43] LAPP, 9 Chemin de Bellevue, B.P. 110, Annecy-le-Vieux 74951, France
[44] Institut de Physique Théorique, CEA Saclay, France
[45] Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[46] Department of Physics, McGill University, Montreal QC, Canada
[47] University of Debrecen and Institute of Nuclear Research of HAS, H-4001 P.O.Box 51, Hungary
[48] Instituto de Física de La Plata, UNLP-CONICET, La Plata, Argentina
[49] Institut für Theoretische Teilchenphysik und Kosmologie, RWTH Aachen University, D-52056 Aachen, Germany
[50] Niels Bohr International Academy and Discovery Center, Niels Bohr Institute, University of Copenhagen, Blegdamsvej 17, DK-2100, Copenhagen, Denmark
[51] Institut für Physik, Universität Mainz, D -55099 Mainz, Germany

Abstract

The 2011 Les Houches workshop was the first to confront LHC data. In the two years since the previous workshop there have been significant advances in both soft and hard QCD, particularly in the areas of multi-leg NLO calculations, the inclusion of those NLO calculations into parton shower Monte Carlos, and the tuning of the non-perturbative parameters of those Monte Carlos. These proceedings describe the theoretical advances that have taken place, the impact of the early LHC data, and the areas for future development.

Report of the SM and NLO Multileg and SM MC Working Groups for the Workshop "Physics at TeV Colliders", Les Houches, France, 31 May–8 June, 2011.

Contents

Part I
INTRODUCTION

The workshop in 2011 was the first for which the long-awaited LHC data (at 7 TeV) was available for analysis and comparison to theory. Even though of limited statistical power compared to the ultimate goals of the LHC, this data accesses a very wide kinematic range, and probes regions where multiple scales are important. The presence of large scales for some processes, on the TeV level, points to the importance of electroweak corrections, which have been calculated only for some of the important processes. The first hints of a Higgs boson have now been observed. In order to search for signs of New Physics, as well as to completely understand the Standard Model at the LHC, it is important to understand the *perturbative framework* at the LHC. The data taken so far provides many challenges for perturbative QCD predictions; and it is clear that New Physics, if it is present in current data, is hiding well.

On the theoretical side, there has been a great deal of productivity in the area of multi-particle calculations at next-to-leading order (NLO) and next-to-next-to-leading order (NNLO). NLO is the first order at which the normalization, and in some cases the shape, of perturbative cross sections can be considered reliable [1]. A full understanding for both Standard Model and beyond the Standard Model physics at the LHC requires the development of fast, reliable programs for the calculation of multi-parton final states at NLO. There have been many advances in the development of NLO techniques, especially in the area of automation [2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12].

Some of these approaches also allow for relatively easy [13, 14, 12] and/or automatic [15] inclusion of the NLO matrix elements into parton shower Monte Carlo programs. For more details we refer to the individual contributions in these proceedings.

A prioritized list of NLO (and some NNLO) cross sections was assembled at Les Houches in 2005 [16] and added to in 2007 [17] and 2009 [18]. This list includes cross sections which are experimentally important, and which are theoretically feasible (if difficult) to calculate. As we stand now, basically all NLO $2 \to 3$ and $2 \to 4$ cross sections of interest have been calculated, see Tables 1,2 below, and even some processes which were not on the 2009 wishlist are available at NLO, see e.g. [19, 20, 21, 22, 23, 24, 25]. The success of automation techniques means that future NLO calculations of similar complexity can be completed without the man-years of labor previously required. Thus, we do not add to the NLO wish list in 2011. Instead, we comment on calculations needed at NNLO, and processes at NLO for which it is important to calculate the impact of electroweak corrections, and/or the

influence of interference effects with other processes with the same final state.

For many of the processes calculated at the LHC (such as for Higgs production), it is important either to apply a veto for the production of extra jets, or to bin the analysis results according to the jet multiplicity. While such cuts are useful for dealing with the experimental backgrounds, the exclusivity of the cross sections results in increases to the theoretical uncertainties obtained for the corresponding inclusive results, see e.g. [26]. The impact of such cuts is explored in the contribution of Stewart and Tackmann in these proceedings.

Much of the complexity for multi-parton NLO processes consists of the calculation of the non-leading color contributions. Such contributions typically contribute only at the level of a few percent and approximations to the non-leading color contributions should be accurate within a percent or so [70, 71, 51]. So it may be more time-prudent for groups carrying out such calculations to estimate the non-leading color effects before carrying out the full calculation.

To reach full utility, the codes for any of these complex NLO calculations should be made public and/or the authors should generate ROOT ntuples providing the parton level event information from which experimentalists can assemble any cross sections of interest. Where possible, decays (with spin correlations) should be included. A ROOT output option is especially useful where the creation of a user-friendly NLO program may be very time-consuming. We now have some experience with the use of ROOT ntuples with both MCFM and Blackhat+Sherpa calculations. The latter, in particular, does not exist as a public program, while ROOT tuples have been made available for NLO W/Z + n jet multiplicities (with n up to 4) for W/Z + jets, and (also for n up to 4) for inclusive jet production. The estimation of the *correct* scale for use in multi-parton NLO calculations, and the proper evaluation of the uncertainty on this scale, is more complex than for simpler calculations. The use of ROOT ntuples can make these evaluations easier to carry out. A contribution describing their use has been included in these proceedings.

While NLO is sufficient for most predictions, it is also crucial to understand certain critical cross sections at NNLO. To date, NNLO calculations have been carried out primarily for processes in e^+e^- annihilation [72, 73, 74], and in hadronic collisions for $2 \to 1$ processes, with the exception of VH [75, 76, 77] and $\gamma\gamma$ production [78].

To calculate a $2 \to 2$ scattering process at NNLO, the divergent contributions arising from the tree-level $2 \to 4$, the one-loop $2 \to 3$ and the two-loop $2 \to 2$ subprocesses have to be properly subtracted and cancelled, such that the finite remainders can be combined into a parton-level event generator. To combine the three contributions, an infra-red subtraction scheme for unresolved real radiation is required. Several approaches have been used and are being further developed: antenna subtraction [79], which currently is extended to hadronic and semi-hadronic initial states [80, 81, 82, 83, 84], a method based on sector decomposition appplied to real radiation [85, 86, 87] where the decomposition is guided by the physical singularity structure [88, 89], q_T-subtraction [90], which is very elegant but appplicable only to colourless final states, and the one of [91] described in these proceedings.

Further, two-loop amplitudes are interesting in their own right from a field theory point of view, for example to study asymptotic behaviour, or to gain insights into the all-order infared structure of massless field theories.

Below we construct a table of calculations needed at the LHC, and which are feasible within the next few years. Certainly, results for inclusive cross sections at NNLO will be easier to achieve than differential distributions, but most groups are working towards a partonic Monte Carlo program capable of producing fully differential distributions for measured observables.

- $t\bar{t}$ production:
  needed for accurate background estimates, top mass measurement, top quark asymmetry (which is zero at tree level, so NLO is the leading non-vanishing order for this observable, and a discrepancy of theory predictions with Tevatron data needs to be understood). Several groups are already well

Process ($V \in \{Z, W, \gamma\}$)	Comments
Calculations completed since Les Houches 2005	
1. $pp \to VV$ jet	WW jet completed by Dittmaier/Kallweit/Uwer [27, 28]; Campbell/Ellis/Zanderighi [29]. ZZ jet completed by Binoth/Gleisberg/Karg/Kauer/Sanguinetti [30] WZ jet,$W\gamma$ jet completed by Campanario et al. [31, 32]
2. $pp \to$ Higgs+2 jets	NLO QCD to the gg channel completed by Campbell/Ellis/Zanderighi [33]; NLO QCD+EW to the VBF channel completed by Ciccolini/Denner/Dittmaier [34, 35] Interference QCD-EW in VBF channel [36, 37]
3. $pp \to VVV$	ZZZ completed by Lazopoulos/Melnikov/Petriello [38] and WWZ by Hankele/Zeppenfeld [39], see also Binoth/Ossola/Papadopoulos/Pittau [40] VBFNLO [41, 42] meanwhile also contains $WWW, ZZW, ZZZ, WW\gamma, ZZ\gamma, WZ\gamma, W\gamma\gamma, Z\gamma\gamma, \gamma\gamma\gamma, W\gamma\gamma j$ [43, 44, 45, 46, 47, 21]
4. $pp \to t\bar{t}\,b\bar{b}$	relevant for $t\bar{t}H$, computed by Bredenstein/Denner/Dittmaier/Pozzorini [48, 49] and Bevilacqua/Czakon/Papadopoulos/Pittau/Worek [50]
5. $pp \to V$+3 jets	W+3 jets calculated by the Blackhat/Sherpa [51] and Rocket [52] collaborations Z+3jets by Blackhat/Sherpa [53]
Calculations remaining from Les Houches 2005	
6. $pp \to t\bar{t}$+2jets	relevant for $t\bar{t}H$, computed by Bevilacqua/Czakon/Papadopoulos/Worek [54, 55]
7. $pp \to VV\,b\bar{b}$, 8. $pp \to VV$+2jets	Pozzorini et al.[25],Bevilacqua et al.[23] W^+W^++2jets [56], W^+W^-+2jets [57, 58], VBF contributions calculated by (Bozzi/)Jäger/Oleari/Zeppenfeld [59, 60, 61]
NLO calculations added to list in 2007	
9. $pp \to b\bar{b}b\bar{b}$	Binoth et al. [62, 63]
NLO calculations added to list in 2009	
10. $pp \to V + 4$ jets	top pair production, various new physics signatures Blackhat/Sherpa: W+4jets [22], Z+4jets [20] see also HEJ [64] for $W + n$jets
11. $pp \to Wb\bar{b}j$	top, new physics signatures, Reina/Schutzmeier [11]
12. $pp \to t\bar{t}t\bar{t}$	various new physics signatures
also completed: $pp \to W\gamma\gamma$ jet	Campanario/Englert/Rauch/Zeppenfeld [21]
$pp \to 4$ jets	Blackhat/Sherpa [19]

Table 1: The updated experimenter's wishlist for LHC processes

Calculations beyond NLO added in 2007	
13. $gg \to W^*W^*$ $\mathcal{O}(\alpha^2\alpha_s^3)$ 14. NNLO $pp \to t\bar{t}$ 15. NNLO to VBF and Z/γ+jet	backgrounds to Higgs normalization of a benchmark process Higgs couplings and SM benchmark
Calculations including electroweak effects	
16. NNLO QCD+NLO EW for W/Z NLO EW to W/Z NLO EW to W/Z+jet NLO EW to WH/ZH	precision calculation of a SM benchmark [65, 66] [67, 68] [69]

Table 2: The updated experimenter's wishlist for LHC processes continued

on the way to complete NNLO results for $t\bar{t}$ production [92, 93, 94, 95].

- W^+W^- production:
  importand background to Higgs search. At the LHC, $gg \to WW$ is the dominant subprocess, but $gg \to WW$ is a loop-induced process, such that two loops need to be calculated to get a reliable estimate of the cross section. Advances towards the full two-loop result are reported in [96, 97].
- inclusive jet/dijet production:
  NNLO parton distribution function (PDF) fits are starting to become the norm for predictions and comparisons at the LHC. Paramount in these global fits is the use of inclusive jet production to tie down the behavior of the gluon distribution, especially at high x. However, while the other essential processes used in the global fitting are known to NNLO, the inclusive jet production cross section is only known at NLO. Thus, it is crucial for precision predictions for the LHC for the NNLO corrections for this process to be calculated, and to be available for inclusion in the global PDF fits. First results for the real-virtual and double real corrections to gluon scattering can be found in [98, 99].
- V+1 jet production:
  $W/Z/\gamma$ + jet production form the signal channels (and backgrounds) for many key physics processes, for both SM and BSM. In addition, they also serve as calibration tools for the jet energy scale and for the crucial understanding of the missing transverse energy resolution. The two-loop amplitudes for this process are known [100, 101], therefore it can be calculated once the parts involving unresolved real radiation are available.
- V+γ production:
  important signal/background processes for Higgs and New Physics searches. The two-loop helicity amplitudes for $q\bar{q} \to W^{\pm}\gamma$ and $q\bar{q} \to Z^0\gamma$ recently have become available [102].
- Higgs+1 jet production:
  As mentioned previously, events in many of the experimental Higgs analyses are separated by the number of additional jets accompanying the Higgs boson. In many searches, the Higgs + 0 jet and Higgs + 1 jet bins contribute approximately equally to the sensitivity. It is thus necessary to have the same theoretical accuracy for the Higgs + 1 jet cross section as already exists for the inclusive Higgs cross section, i.e. NNLO. The two-Loop QCD Corrections to the Helicity Amplitudes for $H \to 3$ partons are already available [103].

The contributions in this document are arranged as follows. In section II, various developments concerning techniques for NLO and NNLO calculations are described, in particular in view of providing

automated tools for NLO corrections. In section III, issues related to parton distribution functions are discussed. Section IV contains phenomenological studies of observables and uncertainties, based on theory input where higher order corrections obtained by different approaches are available. Section V includes phenomenological studies on the definition of experimental observables and corrections applied to data. Finally Section VI discusses issues related to the tuning of Monte Carlos and standardised Monte Carlo output formats.

Part II

NLO AUTOMATION AND (N)NLO TECHNIQUES

1. PJFRY – A C++ PACKAGE FOR TENSOR REDUCTION OF ONE-LOOP FEYNMAN INTEGRALS [1]

Abstract

The C++ package PJFry 1.0.0 [104, 105] – a one loop tensor integral library – is introduced. We use an algebraic approach to tensor reduction. As a result, the tensor integrals are presented in terms of scalar one- to four-point functions, which have to be provided by an external library, e.g. QCDLoop/FF or OneLOop or LoopTools/FF. The reduction is implemented until five-point functions of rank five. A numerical example is shown, including a special treatment for small or vanishing inverse four-point Gram determinants. Future modules of PJFry might cover the treatment of n-point functions with $n \geq 6$; the corresponding formulae are worked out. Further, an extremely efficient approach to tensor reduction relies on evaluations of complete contractions of the tensor integrals with external momenta. For this, we worked out an algorithm for the analytical evaluation of sums over products of signed minors with scalar products of chords, i.e. differences of external momenta. As a result, the usual multiple sums over tensor coefficients are replaced for the numerical evaluation by compact combinations of the basic scalar functions.

1.1 PJFry

The goal of the C++ package PJFry is a stable and fast open-source implementation of one-loop tensor reduction of Feynman integrals

$$I_n^{\mu_1\cdots\mu_R} \quad = \quad C(\epsilon) \int \frac{d^dk}{i\pi^{d/2}} \; \frac{\prod_{r=1}^R k^{\mu_r}}{\prod_{j=1}^n (k-q_j)^2 - m_j^2 + i\epsilon}, \tag{1}$$

suitable for any physically relevant kinematics.[2] The algorithm was invented in [105]. PJFry performs the reduction of 5-point 1-loop tensor integrals up to rank 5. The 4- and 3-point tensor integrals are obtained as a by-product. Main features are:

- Any combination of internal or external masses
- Automatic selection of optimal formula for each coefficient
- Leading $()_5$ are eliminated in the reduction

[1] Contributed by: J. Fleischer, T. Riemann, V. Yundin

[2] An extended description of notations and of the formalism may be found in [105, 106, 107, 108]. The normalization of PJFry follows that chosen in the scalar library. For QCDLoop, $C(\epsilon) = \Gamma(1-2\epsilon)/[\Gamma(1+\epsilon)\Gamma^2(1-\epsilon)]$.

- Small $()_4$ are avoided using asymptotic expansions where appropriate
- Cache system for tensor coefficients and signed minors
- Interfaces for C, C++, FORTRAN and Mathematica
- Uses QCDLoop [109, 110] or OneLOop [111] for 4-dim scalar integrals
- Available from the project webpage `https://github.com/Vayu/PJFry/` [104, 105]

The installation of PJFry may be performed following the instructions given at the project webpage.
The project subdirectories are
./src - the library source code
./mlink - the MathLink interface
./examples - the FORTRAN examples of library use, built with make check

A build on Unix/Linux and similar systems is done in a standard way by sequential performing ./configure, make, make install. See the INSTALL file for a detailed description of the ./configure options.

The functions for tensor coefficients for up to rank $R = 5$ pentagon integrals are declared in the Mathematica interface:

```
In:= Names["PJFry`*"]

Out= {A0v0, B0v0, B0v1, B0v2, C0v0, C0v1, C0v2, C0v3, \
D0v0, D0v1, D0v2, D0v3, D0v4, E0v0, E0v1, E0v2, \
E0v3, E0v4, E0v5, GetMu2, SetMu2}
```

The C++ and Fortran interface syntax is very close to that of e.g. LoopTools/FF:

```
E0v3[i,j,k,p1s,p2s,p3s,p4s,p5s,s12,s23,s34,s45,s15,m1s,m2s,m3s,m4s,m5s,ep=0]
```

where:[3]
`i,j,k` are indices of the tensor coefficient $(0 < i \leq j \leq k < n)$,
`p1s,p2s,...` are squared external masses p_i^2,
`s12,s23,...` are Mandelstam invariants $(p_i + p_j)^2$,
`m1s,m2s,...` are squared internal masses m_i^2,
`ep=0,-1,-2` selects the coefficient of the ϵ-expansion.

The average evaluation time per phase-space point on a 2 GHz Core 2 laptop for the evaluation of all 81 rank 5 tensor form-factors amounts to 2 ms.

A numerical example is shown, for a configuration as in figure 1, in figures 2 and 3 for a five-point rank $R = 4$ tensor coefficient in a region, where one of the 4-point sub-Gram determinants vanishes [at $x = 0$]:

$$E_{3333}(0, 0, -6\times10^4(x+1), 0, 0, 10^4, -3.5\times10^4, 2\times10^4, -4\times10^4, 1.5\times10^4, 0, 6550, 0, 0, 8315)$$

The red curve is produced with standard PJFry, and the blue one with Passarino-Veltman [PV] reduction [112]; we mention that for the case treated here ($x \to 0$), the PV reduction is no standard option. Our expansion in terms of higher dimensional scalar 3-point functions in case of vanishing 4-point sub-Gram determinants uses only functions I_3^{d+2l} [105]. These are tensor coefficients of the pure $g^{\mu\nu}$ type [113], and so our method is different from others with a mixed numerical approach [114] or with use of additional tensor coefficients [115].

Tensor reduction by PJFry is used as one option of the GoSam package [12]. An older version of the algorithm, as described in [116], has been implemented independently in [11].

[3] One has to carefully control accuracies; e.g. the on-shell conditions for massless particles have to be fulfilled with a numerical precision expected by the scalar functions library in use; for QCDLoop this means on default at least 10 digits.

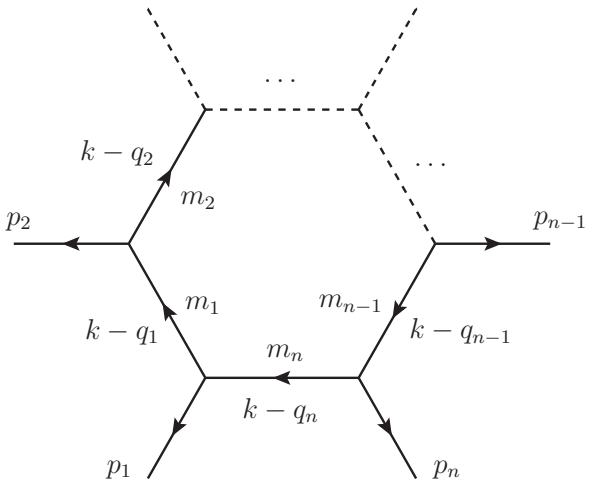


Fig. 1: Momenta definitions for PJFry.

1.2 POTENTIAL UPGRADES

1.21 Tensor reduction for higher-point functions

So far, PJFry is foreseen for 5-point functions and simpler ones. The extension to 6-point functions is known from e.g. [114, 115, 116]. In [107] we solve analytically generalized recursions for $n \geq 6$, derived in [114]:

$$I_n^{\mu_1\mu_2\cdots\mu_R} = -\sum_{r=1}^{n} C_r^{\mu_1}(n) I_{n-1}^{\mu_2\cdots\mu_R,r}, \tag{2}$$

where in $I_{n-1}^{\mu,\cdots,r}$ the line r is scratched. The coefficients for 6-point functions are:

$$C_r^{s,\mu}(6) = \sum_{i=1}^{5} \frac{1}{\binom{0}{s}_6} \begin{pmatrix} 0r \\ si \end{pmatrix}_6 q_i^{\mu_1}, \quad s = 0\ldots 6, \tag{3}$$

where the q_i are chords, and $\binom{0r}{si}_6$ etc. are signed minors with arbitrary s. For the 7-point and 8-point functions, we found several representations, among them

$$C_r^{st,\mu}(7) = \sum_{i=1}^{6} \frac{1}{\binom{st}{st}_7} \begin{pmatrix} sti \\ str \end{pmatrix}_7 q_i^{\mu} \tag{4}$$

and

$$C_r^{stu,\mu}(8) = \sum_{i=1}^{7} \frac{1}{\binom{stu}{stu}_8} \begin{pmatrix} stui \\ stur \end{pmatrix}_8 q_i^{\mu} \tag{5}$$

The upper indices s, t and u stand for the redundancy of the solutions and can be freely chosen.

1.22 Evaluation of contracted tensor integrals using sums over signed minors

The contraction of a tensor integral with chords may be written as a sum over basic scalar integrals (at a stage where they are free of tensor coefficient indices), multiplied by (multiple) sums over chords times

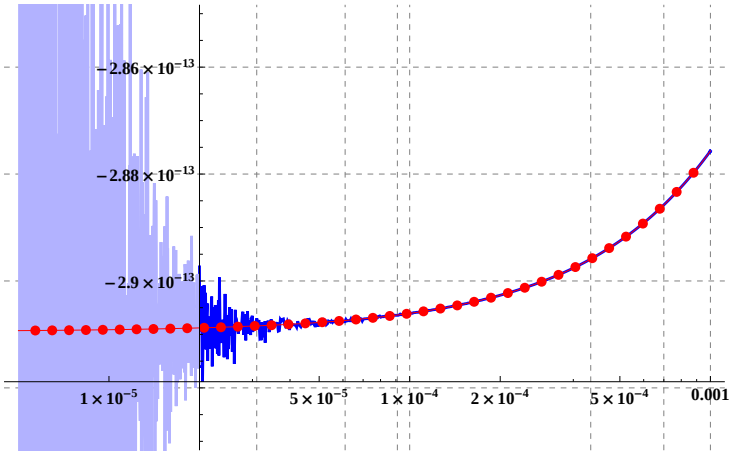


Fig. 2: Absolute accuracy of E_{3333} in the region of vanishing sub-Gram determinant. Blue curve: conventional Passarino-Veltman reduction, red curve: PJFry.

signed minors. If one may perform these sums algebraically, the method becomes very efficient. And this has been systematically worked out in [106], see also [108].

We reproduce here two 7-point examples.

The rank $R = 2, 3$ integrals become by contraction

$$q_{a,\mu} q_{b,\nu}\, I_7^{\mu\nu} = \sum_{r,t=1}^{7} K^{ab,rt} I_5^{rt}, \tag{6}$$

$$q_{a,\mu} q_{b,\nu} q_{c,\lambda}\, I_7^{\mu\nu\lambda} = \sum_{r,t,u=1}^{7} K^{abc,rtu} I_4^{rtu}, \tag{7}$$

where I_5^{rt} and I_4^{rtu} are scalar 5- and 4-point functions, arising from the 7-point function by scratching lines $r, t, \ldots$ In the general case, we have at this stage higher-dimensional integrals $I_n^{d+2l}, n = 2, \ldots, 5$, to be further reduced following the known scheme, if needed. Here, the I_5^{rt} have to be expressed by 4-point functions.

The expansion coefficients are factorizing here,

$$K^{ab,rt} = K^{a,r} K^{b,rt}, \tag{8}$$

$$K^{abc,rtu} = -K^{a,r} K^{b,rt} K^{c,rtu}, \tag{9}$$

and the sums over signed minors have been performed analytically:

$$K^{a,r} = \frac{1}{2} \left(\delta_{ar} - \delta_{7r} \right), \tag{10}$$

$$K^{b,rt} = \sum_{j=1}^{6} (q_b q_j) \frac{\binom{rst}{rsj}_7}{\binom{rs}{rs}_7} \equiv \frac{\Sigma_b^{1,stu}}{\binom{rs}{rs}_7} = \frac{1}{2} \left(\delta_{bt} - \delta_{7t} \right) - \frac{1}{2} \frac{\binom{rs}{ts}_7}{\binom{rs}{rs}_7} \left(\delta_{br} - \delta_{7r} \right), \tag{11}$$

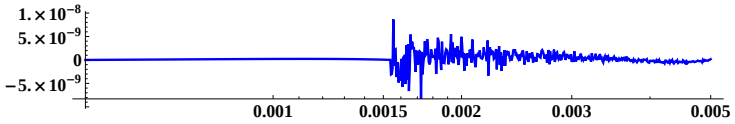


Fig. 3: Relative accuracy of E_{3333} in the region of vanishing sub-Gram determinant. At $x \sim 0.0015$, PJFry switched to the asymptotic expansion.

$$
\begin{aligned}
K^{a,stu} &= \sum_{i=1}^{6} (q_a q_i) \begin{pmatrix} 0stu \\ 0sti \end{pmatrix}_7 \equiv \Sigma_a^{2,stu} \qquad (12)\\
&= \frac{1}{2}\left\{ \begin{pmatrix} stu \\ st0 \end{pmatrix}_7 (Y_{a7} - Y_{77}) + \begin{pmatrix} 0st \\ 0st \end{pmatrix}_7 (\delta_{au} - \delta_{7u}) - \begin{pmatrix} 0st \\ 0su \end{pmatrix}_7 (\delta_{at} - \delta_{7t}) - \begin{pmatrix} 0ts \\ 0tu \end{pmatrix}_7 (\delta_{as} - \delta_{7s}) \right\},
\end{aligned}
$$

with

$$Y_{jk} = -(q_j - q_k)^2 + m_j^2 + m_k^2. \qquad (13)$$

Conventionally, $q_7 = 0$.

The sums may be found in eqns. (A.15) and (A.16) of [106]. The s is redundant and fulfils $s \neq r, b, 7$ in $K^{b,rt}$. In $K_0^{a,stu}$ it is $s, t, u = 1, \ldots 7$ with $s \neq u, t \neq u$.

ACKNOWLEDGEMENTS

J.F. thanks DESY for kind hospitality. Work is supported in part by Sonderforschungsbereich/Transregio SFB/TRR 9 of DFG "Computergestützte Theoretische Teilchenphysik" and European Initial Training Network LHCPHENOnet PITN-GA-2010-264564.

2. THE GOSAM APPROACH TO AUTOMATED ONE-LOOP CALCULATIONS [4]

Abstract

We describe the GoSam framework for the automated computation of multi-particle scattering amplitudes at the one-loop level. The amplitudes are generated explicitly in terms of Feynman diagrams, and can be evaluated using either d-dimensional reduction at the integrand level or tensor decomposition. GoSam can be used to compute one-loop QCD and EW corrections to Standard Model processes, and it is ready to link generic model files for theories Beyond the Standard Model.

[4] Contributed by: G. Cullen, N. Greiner, G. Heinrich, G. Luisoni, P. Mastrolia, G. Ossola, T. Reiter, F. Tramontano

2.1 Introduction and General Motivations

In the last few years we observed major advances in the direction of constructing packages for fully automated one-loop calculations, which profited from the new developments in the field of NLO QCD calculations [17, 18]. The continuous improvement of the techniques for one-loop computations led to important new results for processes with many particles [51, 70, 117, 71, 53, 48, 49, 50, 54, 62, 63, 23, 118, 25, 56, 57, 21, 22, 20, 19].

Very advanced calculations have been performed with improved algebraic reduction methods based on Feynman-diagrammatic algorithms, as well as with new numerical techniques based on the idea of reconstructing one-loop amplitudes from their unitarity cuts. These theoretical developments found an ideal counterpart in the integrand-level reduction algorithm, that allows for the reduction of any scattering amplitudes to scalar master integrals, simply by evaluating numerically the integrand at given fixed values of the integration momentum. In both scenarios, to tackle the increase in the complexity and in the number of diagrams that contribute to the amplitudes, automation becomes indispensable for processes with many external legs.

The purpose of the present document is to illustrate the main features of GoSam [12], a new framework which allows the automated calculation of one-loop scattering amplitudes for multi-particle processes. This approach combines the automated algebraic generation of d-dimensional unintegrated amplitudes obtained via Feynman diagrams, with the numerical integrand-level reduction provided by the d-dimensional extension [119, 120, 6] of the OPP integrand-level reduction method [121, 122, 123] and improved tensorial techniques [124, 125].

The integrands of the one-loop amplitudes are generated via Feynman diagrams, using `QGRAF` [126], `FORM` [127], `spinney` [128] and `haggies` [129]. The only task required from the user is the preparation of an "input card" to start the generation of the source code and its compilation, without having to worry about internal details of the code generation. The individual program tasks are efficiently managed by python scripts. Concerning the reduction, the program offers the possibility to use either the d-dimensional extension of the OPP method, as implemented in SAMURAI [6], or tensor reduction as implemented in `Golem95C` [130, 131] interfaced through tensorial reconstruction at the integrand level [124].

2.2 Algebraic approach to Automation

There are several approaches to the automated computation of multi-particle scattering amplitudes at the one-loop level, which provide different recipes for the construction of multi-purpose tools. The goal of such tools is the evaluation of one-loop scattering amplitudes for any choice of particles in the initial and final states, in a fully automated manner.

In the algebraic approach to multi-purpose automation, amplitudes can be generated from Feynman diagrams by employing tools for algebraic manipulation: already some time ago, the idea of automating NLO calculations has been pursued by public programs like FeynArts [132] and `QGRAF` [126] for diagram generation and FormCalc/LoopTools [133] and GRACE [3] for the automated calculation of NLO corrections, primarily in the electroweak sector.

When we combine the algebraic generation with the integrand-level reduction, the set of algebraic operations required are quite different with respect to a traditional tensorial reduction. Since the target is to provide the numerical value of the numerator function at given values of integration momentum, we should aim at expressions for the unintegrated numerator that are easily evaluated numerically. To achieve this task, for example, expressions in terms of spinor products are particularly convenient.

We briefly list here some of the advantages of the "algebraic approach": i) the algebraic generation is executed separately from the numerical reduction, therefore algebraic manipulations are possible before starting the numerical integration; CPU-time can be spent, once for all at the beginning of the calculation, to optimize and reduce the size of the integrands that will be evaluated numerically several times

later on during the reduction; ii) the algebraic method allows us to group sets of diagrams and cache all factors that do not depend on the integration momentum; iii) easy access to sub-parts of the computation; subsets of diagrams can be easily added or removed from the final results, simplifying comparisons and tests; iv) computer algebra can be performed in dimension d using alternative regularization schemes; v) the choice between different reduction algorithms can be performed at run-time, providing flexibility and internal cross-checks. In the next section we will briefly illustrate how these properties are used within GoSam.

Important progress in a similar direction has been also recently achieved by means of FeynArts, FormCalc and LoopTools [134, 2] to provide amplitudes that can be processed using the integrand-level reduction provided by CutTools [135] and/or SAMURAI [6] or with the traditional Passarino-Veltman reduction [112].

2.3 A brief introduction to GoSam

GoSam produces in a fully automated way all the code required to perform the calculation of virtual one-loop amplitudes. The only task left to the user is the preparation of an "input card" which contains all the information related to the particular process namely initial and final particles, model, helicities, selection rules to exclude particular sets of diagrams, regularization scheme. The card also contains flags to select the preferred reduction methods and some optimization flags to adapt the diagram generation to the needs of the user.

There are three main steps that GoSam follows in order to prepare the code for the calculation: the generation of diagrams that contribute to the process, the optimization and algebraic manipulation to simplify their expressions, and the writing of a FORTRAN code ready to be used within a phase-space integration. It is important to remember that these steps will only be performed once. After the code is generated, the reduction of unintegrated amplitudes to linear combinations of scalar (master) integrals is fully embedded in the process and can be performed with different options, all available at run-time. Only the last part, namely the reduction and evaluation of master integrals, will be repeated for all the different phase-space points that contribute to the cross-section.

2.31 Diagram Generation

For the diagram generation both at tree level and one-loop level we employ `QGRAF` [126] which we complemented by adding another filter over diagrams implemented in `Python`. This gives several advantages since it increases the ability of the code to distinguish certain classes of diagrams and group them according to the sets of their propagators, in order to fully optimize the reduction.

At this stage GoSam generates three sets of output files: an expression for each diagram for `FORM` [127], `Python` code for drawing each diagram, and `Python` code for computing the properties of the diagram. Information about the model is either read from the built-in Standard Model of `QGRAF` or can be defined by the user by means of LanHEP [136] or an Universal FeynRules Output (`UFO`) file [137]
.

The `Python` program automatically performs several operations: diagrams whose color factor turns out to be zero are dropped; the number of propagators containing the loop momentum, the tensor rank and the kinematic invariants of the associated loop integral are computed; diagrams with a vanishing loop integral associated are detected and flagged for the diagram selection; all propagators and vertices are classified for the diagram selection; diagrams containing massive quark self-energy insertions or closed massless quark loops are specially flagged.

During this phase, GoSam also generates a LaTeX file which contains, among other useful information of the generated process, the drawings of all contributing diagrams. To achieve this task, we use our own implementation of the algorithms described in Ref. [138] and Axodraw [139] to actually draw the diagrams.

2.32 Lorentz Algebra

Concerning the algebraic operations performed by GoSam to render the integral suitable for efficient numerical evaluation, one of the primary goals is to split the $(4-2\varepsilon)$ dimensional algebra into strictly four-dimensional objects and symbols representing the higher-dimensional remainder. All external vectors (momenta and polarisation vectors) are kept in four dimensions; internal vectors, however, are kept in the d-dimensional vector space.

We adopt the conventions used in [128], where $\hat{k}$ denotes the four dimensional projection of an in general d dimensional vector k. The $(d-4)$ dimensional orthogonal projection is denoted as $\tilde{k}$. For the integration momentum q we introduce in addition the symbol $\mu^2 = -\tilde{q}^2$, such that

$$q^2 = \hat{q}^2 + \tilde{q}^2 = \hat{q}^2 - \mu^2. \tag{14}$$

We also introduce suitable projectors by splitting the metric tensor

$$g^{\mu\nu} = \hat{g}^{\mu\nu} + \tilde{g}^{\mu\nu}, \quad \hat{g}^{\mu\nu}\tilde{g}_{\nu\rho} = 0, \quad \hat{g}^{\mu}_{\mu} = 4, \quad \tilde{g}^{\mu}_{\mu} = d-4. \tag{15}$$

Once all propagators and all vertices have been replaced by their corresponding expressions, according to the model file, all vector-like quantities and metric tensors are split into their four-dimensional and their orthogonal part. As we use the 't Hooft algebra, γ_5 is defined as a purely four-dimensional object, $\gamma_5 = i\epsilon_{\mu\nu\rho\sigma}\hat{\gamma}^{\mu}\hat{\gamma}^{\nu}\hat{\gamma}^{\rho}\hat{\gamma}^{\sigma}$. By applying the usual anti-commutation relation for Dirac matrices we can always separate the four-dimensional and $(d-4)$-dimensional parts of Dirac traces.

While the $(d-4)$-dimensional traces are reduced completely to products of $(d-4)$-dimensional metric tensors, the four-dimensional part, which will be reduced numerically, is treated such that the number of terms in the resulting expression is kept as small as possible.

2.33 Treatment of rational terms R_2

Instead of relying on the construction of R_2 from specialized Feynman rules [123, 140, 141, 142, 143], we can generate the R_2 part along with all other contribution using automated algebraic manipulations. The code offers the option between the *implicit* and *explicit* construction of the R_2 terms. The implicit construction treats the $4-$ and $(d-4)$ dimensional numerators on equal grounds: they are generated algebraically and reduced numerically. The explicit construction of R_2 is based on the fact that the $(d-4)$ dimensional part of the numerator function contains expressions for the corresponding integrals that are relatively simple and known explicitly. Therefore, after separating it using the algebraic manipulation described before, the $(d-4)$ dimensional part is computed analytically whereas the purely four-dimensional part is passed to the numerical reduction. This approach also allows for an efficient calculation of the part R_2 alone.

2.34 Reduction to scalar (master) integrals

GoSam allows to choose at run-time (i.e. without re-generating the code) the preferred method of reduction. Available options include the integral-level d-dimensional reduction, as implemented in SAMURAI, or traditional tensor reduction as implemented in `Golem95C` interfaced through tensorial reconstruction at the integrand level, or a combination of both. Concerning the scalar (tensorial) integrals, GoSam allows to choose among a variety of options, including QCDLoop [109], OneLoop [111], `Golem95C` [130], plus the recently added `PJFRY` [108]. Among these codes, OneLoop and `Golem95C` also fully support complex masses.

2.4 Installation and Usage

GoSam can be used within a standard Linux/Unix environment. In order to work, it requires some programs to be installed on the system: these include a recent version of `Python` (version $\geq$ 2.6), Java

($\geq$ 1.5), a Fortran95 compiler, `FORM` (version $\geq$ 3.3), and `QGRAF`. Further, at least one of the libraries SAMURAI or `Golem95C` needs to be present at the time the code is compiled.

To facilitate this task, we have prepared a package containing SAMURAI and `Golem95C` together with the libraries for the integrals: OneLOop, QCDLoop, and FF. The package, which is called `gosam-contrib-1.0.tar.gz` is available for download from

`http://projects.hepforge.org/gosam/`.

The installation procedure is facilitated by the use of `Autotools`.

The user can download the GoSam code either as a tar-ball or from the subversion repository at `http://projects.hepforge.org/gosam/`. The installation of GoSam is controlled by `Python Distutils` and can be performed by simply running the command

```
python setup.py install
```

In order to generate the code for a process, the user needs to prepare an input file (process card) which contains:

- *process* specific information, such as a list of initial and final state particles, their helicities (optional) and the order of the coupling constants;
- *scheme* specific information, such as the regularisation and renormalisation schemes, the underlying model, masses and widths which are set to zero;
- *system* specific information, such as paths to programs and libraries or compiler options;
- *optional* information for optimisations within the code generation.

Assuming that the process card is called `myprocess.in`, the generation of the code can be started by simply running the command `gosam.py myprocess.in`. All further steps are controlled by `makefile`s which are automatically generated by GoSam: the command `make compile` generates the source code and compiles all files relevant for the production of the matrix element. The code can be tested with the program `test.f90` (located in the subdirectory `matrix`) which provides, for a random phase-space point, the tree-level LO matrix element and the NLO result for the finite part, single and double poles. Examples of process cards for a selection of benchmark processes are provided with the main distubution.

For more details about the usage and installation of GoSam, we refer the user to a more technical presentation [144] or to the original publication [12] and the user manual which accompanies the code.

2.5 Examples of Applications

The BLHA interface [145] allows to link GoSam to a general Monte Carlo event generator, which is responsible for supplying the missing ingredients for a complete NLO calculation of a physical cross section. Among those, SHERPA [146] offers the possibility to compute the LO cross section and the real corrections with both the subtraction terms and the corresponding integrated counterparts [147]. Using the BLHA interface, we linked GoSam with SHERPA to compute the physical cross section for $W^{\pm}$ + 1-jet at NLO, which is described in Section 18..

The codes produced by GoSam have been tested on several processes of increasing complexity, some of which are shown in Table 1. The full list of processes produced by GoSam and compared to the literature where available is given in Ref. [12].

As an example of the usage of GoSam with a model file different from the Standard Model, we calculated the QCD corrections to neutralino pair production in the MSSM. The model file has been imported using the `UFO` interface. In this calculation, we combined the one-loop amplitude with the real radiation corrections to obtain results for differential cross sections. For the infrared subtraction terms

$e^+e^- \to u\bar{u}$	[148]
$e^+e^- \to t\bar{t}$	[149, 150], own analytic calculation
$u\bar{u} \to d\bar{d}$	[151, 8]
$gg \to gg$	[152]
$gg \to gZ$	[153]
$bg \to H\,b$	[154, 8]
$\gamma\gamma \to \gamma\gamma$ (W loop)	[155]
$\gamma\gamma \to \gamma\gamma\gamma\gamma$ (fermion loop)	[156]
$pp \to t\bar{t}$	[8], MCFM [157, 158]
$pp \to W^\pm\,j$ (QCD corr.)	[157, 158]
$pp \to W^\pm\,j$ (EW corr.)	for IR poles: [65, 159]
$pp \to W^\pm\,t$	[157, 158]
$pp \to W^\pm\,jj$	[157, 158]
$pp \to W^\pm b\bar{b}$ (massive b)	[157, 158]
$e^+e^- \to e^+e^-\gamma$ (QED)	[160]
$pp \to H\,t\bar{t}$	[8]
$pp \to Z\,t\bar{t}$	[10]
$pp \to W^+W^+jj$	[56, v3]
$pp \to b\bar{b}b\bar{b}$	[62, 63]
$pp \to W^+W^-b\bar{b}$	[8, 161]
$pp \to t\bar{t}b\bar{b}$	[8, 161]
$u\bar{d} \to W^+ggg$	[161]

Table 3: Some of the processes computed and checked with GoSam

we employed `MadDipole` [162], while the real emission part is calculated using MadGraph/MadEvent [163]. The virtual matrix element is renormalized in the $\overline{MS}$ scheme, while massive particles are treated in the on-shell scheme. The renormalization terms specific to the massive MSSM particles have been added manually. In Fig.4 we show the differential cross section for the $m_{\chi_1^0\chi_1^0}$ invariant mass, where we employed a jet veto to suppress large contributions from the channel $qg \to \chi_1^0\chi_1^0 q$ which opens up at order $\alpha^2\alpha_s$, but for large p_T^{jet} belongs to the distinct process of neutralino pair plus one hard jet production at leading order.

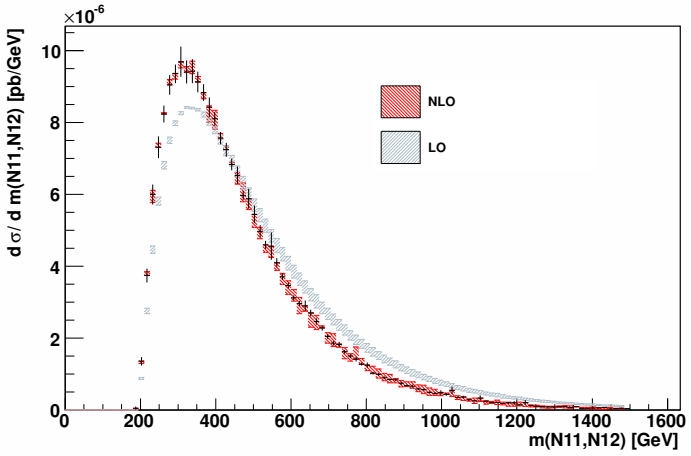


Fig. 4: Comparison of the NLO and LO $m_{\chi_1^0\chi_1^0}$ distributions for the process $pp \to \chi_1^0\chi_1^0$ with a jet veto on jets with $p_T^{jet} > 20$ GeV and $\eta < 4.5$. The band gives the dependence of the result on $\mu = \mu_F = \mu_R$ between $\mu_0/2$ and $2\mu_0$. We choose $\mu_0 = m_Z$.

Conclusions and Outlook

Several groups are currently working at the development of automated multi-purpose tools for one-loop calculations. For quite a long time, tree-level calculation have been fully automated and included in flexible multi-process tools [164, 165]. The level of automation achieved by one-loop calculations is suggesting the possibility of a similar success also at the next-lo-leading order. The target is to build efficient and flexible NLO programs which can be used to tackle the increasing need of precision required by the experimental collaborations.

GOSAM is a flexible and broadly applicable tool for the fully automated evaluation of one-loop scattering amplitudes. In this approach, scattering amplitudes are generated in terms of Feynman diagrams and their reduction to master integrals can be performed in several ways, which can be selected at run-time. GOSAM can be used to calculate one-loop corrections both in QCD and electro-weak theory and offers the flexibility to link general model files for theories Beyond the Standard Model. The code performed well in reproducing a wide range of examples and we are looking forward to tackle more challenging calculations and interfacing with other existing tools in the near future.

Acknowledgments

The work of G.C. was supported by DFG Sonderforschungsbereich Transregio 9, Computergestützte Theoretische Teilchenphysik. N.G. was supported in part by the U.S. Department of Energy under contract No. DE-FG02-91ER40677. P.M. and T.R. were supported by the Alexander von Humboldt Foundation, in the framework of the Sofja Kovaleskaja Award Project Advanced Mathematical Methods for Particle Physics, endowed by the German FederalMinistry of Education and Research. The work of G.O. was supported in part by the National Science Foundation under Grant PHY-0855489. G.O. also acknowledges the support of the Center for Theoretical Physics at City Tech. The research of F.T. is supported by Marie-Curie-IEF, project: SAMURAI-Apps. We also acknowledge the support of the Research Executive Agency (REA) of the European Union under the Grant Agreement number PITN-GA-2010-264564 (LHCPhenoNet).

3. AUTOMATION AND NUMERICAL LOOP INTEGRATION [5]

3.1 INTRODUCTION

Numerical methods are nowadays routinely used in fully differential fixed-order perturbative calculations for the integration over the phase-space of the final-state particles. The use of numerical methods for the phase-space integration allows the flexibility to compute any infrared-safe observable for a given process within a single numerical program. It is thus natural to investigate if numerical methods can also be applied for the loop integration in the virtual corrections [166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176]. A major breakthrough was achieved recently by showing that the numerical method is compatible in efficency with the commonly used approaches based on cut techniques and generalised unitarity or on Feynman graphs [70, 51, 53, 22, 20, 52, 117, 56, 54, 50, 48, 177, 111, 5, 178, 8, 12]. The implementation of the numerical method for the loop integration is process-independent and offers therefore the flexibility to compute several processes within one numerical program. We discuss the main principles of the numerical method for the loop integration at one-loop. In addition we give an outlook towards higher loops.

3.2 THE SUBTRACTION METHOD FOR THE LOOP INTEGRATION

The contributions to an infrared-safe n-jet observable observable O at next-to-leading order are given by

$$\langle O \rangle^{NLO} = \int_{n+1} O_{n+1} d\sigma^R + \int_n O_n d\sigma^V + \int_n O_n d\sigma^C. \tag{16}$$

[5] Contributed by: S. Weinzierl

Here a rather condensed notation is used. $d\sigma^R$ denotes the real emission contribution, whose matrix elements are given by the square of the Born amplitudes with $(n+3)$ partons $|A^{(0)}_{n+3}|^2$. $d\sigma^V$ denotes the virtual contribution, whose matrix elements are given by the interference term of the one-loop and Born amplitude Re $(A^{(0)*}_{n+2}A^{(1)}_{n+2})$ and $d\sigma^C$ denotes a collinear subtraction term, which subtracts the initial state collinear singularities. Each term is separately divergent and only their sum is finite.

The subtraction method is widely used to render the real emission part of a NLO calculation suitable for a numerical Monte Carlo integration. One adds and subtracts a suitably chosen piece to be able to perform the phase-space integrations by Monte Carlo methods:

$$\langle O\rangle^{NLO} \;=\; \int\limits_{n+1}\left(O_{n+1}d\sigma^R - O_n d\sigma^A\right) + \int\limits_{n}\left(O_n d\sigma^V + O_n d\sigma^C + O_n \int\limits_{1} d\sigma^A\right). \qquad (17)$$

The first term $\left(O_{n+1}d\sigma^R - O_n d\sigma^A\right)$ is by construction integrable over the $(n+1)$-particle phase-space and can be evaluated numerically. The result of the integration of the subtraction term over the unresolved one-parton phase-space is written in a compact notation as

$$d\sigma^C + \int\limits_{1} d\sigma^A \;=\; (\mathbf{I}+\mathbf{K}+\mathbf{P})\otimes d\sigma^B. \qquad (18)$$

The notation $\otimes$ indicates that colour correlations due to the colour charge operators $\mathbf{T}_i$ still remain. The terms with the insertion operators $\mathbf{K}$ and $\mathbf{P}$ pose no problem for a numerical evaluation. The term $\mathbf{I}\otimes d\sigma^B$ lives on the phase-space of the n-parton configuration and has the appropriate singularity structure to cancel the infrared divergences coming from the one-loop amplitude. Therefore $d\sigma^V + \mathbf{I}\otimes d\sigma^B$ is infrared finite.

We extend this subtraction method to the virtual part such that we can evaluate the one-loop integral of the one-loop amplitude numerically. The renormalised one-loop amplitude $\mathcal{A}^{(1)}$ is related to the bare amplitude $\mathcal{A}^{(1)}_{\rm bare}$ by $\mathcal{A}^{(1)} = \mathcal{A}^{(1)}_{\rm bare} + \mathcal{A}^{(1)}_{\rm CT}$, where $\mathcal{A}^{(1)}_{\rm CT}$ denotes the ultraviolet counterterm from renormalisation. The bare amplitude involves the loop integration

$$\mathcal{A}^{(1)}_{\rm bare} \;=\; \int \frac{d^Dk}{(2\pi)^D}\mathcal{G}^{(1)}_{\rm bare}. \qquad (19)$$

where $\mathcal{G}^{(1)}_{\rm bare}$ denotes the integrand of the bare one-loop amplitude. We introduce subtraction terms which match locally the singular behaviour of the bare integrand:

$$\mathcal{A}^{(1)}_{\rm bare} + \mathcal{A}^{(1)}_{\rm CT} = \int \frac{d^Dk}{(2\pi)^D}\left(\mathcal{G}^{(1)}_{\rm bare} - \mathcal{G}^{(1)}_{\rm soft} - \mathcal{G}^{(1)}_{\rm coll} - \mathcal{G}^{(1)}_{\rm UV}\right) + \left(\mathcal{A}^{(1)}_{\rm CT} + \mathcal{A}^{(1)}_{\rm soft} + \mathcal{A}^{(1)}_{\rm coll} + \mathcal{A}^{(1)}_{\rm UV}\right). \qquad (20)$$

Analogous to $\mathcal{G}^{(1)}_{\rm bare}$, the integrands of the subtraction terms $\mathcal{A}^{(1)}_x$ are denoted by $\mathcal{G}^{(1)}_x$, where x is equal to $\rm soft$, $\rm coll$ or $\rm UV$. The expression in the first bracket is finite and can therefore be integrated numerically in four dimensions. The integrated subtraction terms in the second bracket are easily calculated analytically in D dimensions. The result can be written as

$$2\ \mathrm{Re}\ \mathcal{A}^{(0)}\left(\mathcal{A}^{(1)}_{\rm CT} + \mathcal{A}^{(1)}_{\rm soft} + \mathcal{A}^{(1)}_{\rm coll} + \mathcal{A}^{(1)}_{\rm UV}\right)^* O_n d\phi_n \;=\; \mathbf{L}\otimes d\sigma^B. \qquad (21)$$

The insertion operator $\mathbf{L}$ contains the explicit poles in the dimensional regularisation parameter related to the infrared singularities of the one-loop amplitude. These poles cancel when combined with the insertion operator $\mathbf{I}$:

$$(\mathbf{I}+\mathbf{L})\otimes d\sigma^B \;=\; \text{finite}. \qquad (22)$$

The operator $\mathbf{L}$ contains, as does the operator $\mathbf{I}$, colour correlations due to soft gluons. In analogy to the one-loop amplitude we can write $d\sigma^V = d\sigma_{\rm CT} + \int \frac{d^Dk}{(2\pi)^D} d\sigma^V_{\rm bare}$ and then the NLO contributions reads

$$\langle O \rangle^{NLO} = \int\limits_{n+1} \left(O_{n+1} d\sigma^R - O_n d\sigma^A \right) + \int\limits_{n+\rm loop} O_n \left(d\sigma^V_{\rm bare} - d\sigma^{A'} \right) + \int\limits_{n} O_n \left(\mathbf{I} + \mathbf{L} + \mathbf{K} + \mathbf{P} \right) \otimes d\sigma^B. \tag{23}$$

In a condensed notation this reads

$$\langle O \rangle^{NLO} = \langle O \rangle^{NLO}_{\rm real} + \langle O \rangle^{NLO}_{\rm virtual} + \langle O \rangle^{NLO}_{\rm insertion}. \tag{24}$$

Every single term is finite and can be evaluated numerically.

3.3 THE SUBTRACTION TERMS

Amplitudes in QCD may be decomposed into group-theoretical factors (carrying the colour structures) multiplied by kinematic factors called partial amplitudes. At the loop level partial amplitudes may further be decomposed into primitive amplitudes. It is simpler to work with primitive one-loop amplitudes instead of a full one-loop amplitude. Our method exploits the fact that primitive one-loop amplitudes have a fixed cyclic ordering of the external legs and that they are gauge-invariant. The first point ensures that there are at maximum n different loop propagators in the problem, where n is the number of external legs, while the second property of gauge invariance is crucial for the proof of the method. We therefore consider in the following just a single primitive one-loop amplitude, which we denote by $A^{(1)}$, while keeping in mind that the full one-loop amplitude is just the sum of several primitive amplitudes multiplied by colour structures. We label the external momenta clockwise by p_1, p_2, ..., p_n and define $q_i = p_1 + p_2 + ... + p_i$, $k_i = k - q_i$. We can write the bare primitive one-loop amplitude in Feynman gauge as

$$A^{(1)}_{bare} = \int \frac{d^Dk}{(2\pi)^D} G^{(1)}_{bare}, \qquad G^{(1)}_{bare} = P(k) \prod_{i=1}^{n} \frac{1}{k_i^2 - m_i^2 + i\delta}. \tag{25}$$

$G^{(1)}_{bare}$ is the integrand of the bare one-loop amplitude. $P(k)$ is a polynomial in the loop momentum k. The $+i\delta$-prescription instructs us to deform – if possible – the integration contour into the complex plane to avoid the poles at $k_i^2 = m_i^2$. If a deformation close to a pole is not possible, we say that the contour is pinched. If we restrict ourselves to non-exceptional external momenta, then the divergences of the one-loop amplitude related to a pinched contour are either due to soft or collinear partons in the loop. These divergences are regulated within dimensional regularisation by setting the number of space-time dimensions equal to $D = 4 - 2\varepsilon$. A primitive amplitude which has soft or collinear divergences must have at least one loop propagator which corresponds to a gluon. An amplitude which just consists of a closed fermion loop does not have any infrared divergences. We denote by I_g the set of indices i, for which the propagator i in the loop corresponds to a gluon. The soft and collinear subtraction terms for massless QCD read [169]

$$\begin{aligned} G^{(1)}_{\rm soft} &= 16\pi\alpha_s i \sum_{j \in I_g} \frac{p_j \cdot p_{j+1}}{k_{j-1}^2 k_j^2 k_{j+1}^2} A_j^{(0)} \,, \\ G^{(1)}_{\rm coll} &= -8\pi\alpha_s i \sum_{j \in I_g} \left[\frac{S_j g_{\rm UV}(k_{j-1}^2, k_j^2)}{k_{j-1}^2 k_j^2} + \frac{S_{j+1} g_{\rm UV}(k_j^2, k_{j+1}^2)}{k_j^2 k_{j+1}^2} \right] A_j^{(0)} \,, \end{aligned} \tag{26}$$

where $S_j = 1$ if the external line j corresponds to a quark and $S_j = 1/2$ if it corresponds to a gluon. The function $g_{\rm UV}$ ensures that the integration over the loop momentum is ultraviolet finite. Integrating

the soft and the collinear part we obtain

$$
\begin{aligned}
S_\varepsilon^{-1}\mu_s^{2\varepsilon}\int\frac{d^Dk}{(2\pi)^D}\,G^{(1)}_{\rm soft} &= -\frac{\alpha_s}{4\pi}\frac{e^{\varepsilon\gamma_E}}{\Gamma(1-\varepsilon)}\sum_{j\in I_g}\frac{2}{\varepsilon^2}\Big(\frac{-2p_jp_{j+1}}{\mu_s^2}\Big)^{-\varepsilon}A_j^{(0)}\ +\mathcal{O}(\varepsilon),\\
S_\varepsilon^{-1}\mu_s^{2\varepsilon}\int\frac{d^Dk}{(2\pi)^D}\,G^{(1)}_{\rm coll} &= -\frac{\alpha_s}{4\pi}\frac{e^{\varepsilon\gamma_E}}{\Gamma(1-\varepsilon)}\sum_{j\in I_g}(S_j+S_{j+1})\frac{2}{\varepsilon}\Big(\frac{\mu_{uv}^2}{\mu_s^2}\Big)^{-\varepsilon}A_j^{(0)}\ +\mathcal{O}(\varepsilon),
\end{aligned}
\tag{27}
$$

$S_\varepsilon=(4\pi)^\varepsilon e^{-\varepsilon\gamma_E}$ is the typical volume factor of dimensional regularisation, γ_E is Euler's constant and μ is the renormalisation scale.

The ultraviolet subtraction terms correspond to propagator and vertex corrections. The subtraction terms are obtained by expanding the relevant loop propagators around a new ultraviolet propagator $(\bar{k}^2-\mu_{\rm UV}^2)^{-1}$, where $\bar{k}=k-Q$: For a single propagator we have

$$
\frac{1}{(k-p)^2} = \frac{1}{\bar{k}^2-\mu_{\rm UV}^2}+\frac{2\bar{k}\cdot(p-Q)}{\left(\bar{k}^2-\mu_{\rm UV}^2\right)^2}-\frac{(p-Q)^2+\mu_{\rm UV}^2}{\left(\bar{k}^2-\mu_{\rm UV}^2\right)^2}+\frac{\left[2\bar{k}\cdot(p-Q)\right]^2}{\left(\bar{k}^2-\mu_{\rm UV}^2\right)^3}+\mathcal{O}\left(\frac{1}{|\bar{k}|^5}\right).
$$

We can always add finite terms to the subtraction terms. For the ultraviolet subtraction terms we choose the finite terms such that the finite parts of the integrated ultraviolet subtraction terms are independent of Q and proportional to the pole part, with the same constant of proportionality for all ultraviolet subtraction terms. This ensures that the sum of all integrated UV subtraction terms is again proportional to a tree-level amplitude [167].

3.4 CONTOUR DEFORMATION

Having a complete list of ultraviolet and infrared subtraction terms at hand, we can ensure that the integration over the loop momentum gives a finite result and can therefore be performed in four dimensions. However, this does not yet imply that we can safely integrate each of the four components of the loop momentum k^μ from minus infinity to plus infinity along the real axis. There is still the possibility that some of the loop propagators go on-shell for real values of the loop momentum. If the contour is not pinched this is harmless, as we may escape into the complex plane in a direction indicated by Feynman's $+i\delta$-prescription. However, it implies that the integration should be done over a region of real dimension 4 in the complex space $\mathbb{C}^4$. Let us consider an integral corresponding to a primitive one-loop amplitude with n propagators minus the appropriate IR- and UV-subtraction terms:

$$
\int\frac{d^4\tilde{k}}{(2\pi)^4}\left(\mathcal{G}^{(1)}_{\rm bare}-\mathcal{G}^{(1)}_{\rm soft}-\mathcal{G}^{(1)}_{\rm coll}-\mathcal{G}^{(1)}_{\rm UV}\right) = \int\frac{d^4\tilde{k}}{(2\pi)^4}P(\tilde{k})\prod_{j=1}^{n}\frac{1}{\tilde{k}_j^2-m_j^2+i\delta}
\tag{28}
$$

where $P(\tilde{k})$ is a polynomial of the loop momentum $\tilde{k}^\mu$ and the integration is over a complex contour in order to avoid whenever possible the poles of the propagators. We set $\tilde{k}^\mu=k^\mu+i\kappa^\mu(k)$, where k^μ is real [170]. After this deformation our integral equals

$$
\int\frac{d^4k}{(2\pi)^4}\left|\frac{\partial\tilde{k}^\mu}{\partial k^\nu}\right|P(\tilde{k}(k))\prod_{j=1}^{n}\frac{1}{k_j^2-m_j^2-\kappa^2+2ik_j\cdot\kappa}.
\tag{29}
$$

To match Feynman's $+i\delta$-prescription we have to construct the deformation vector κ such that

$$
k_j^2-m_j^2 = 0 \quad\rightarrow\quad k_j\cdot\kappa\geq 0.
\tag{30}
$$

We remark that the numerical stability of the Monte Carlo integration depends strongly on the definition of the deformation vector κ.

3.5 NLO results for n-jets in electron-positron annihilation

We have calculated results for jet observables in electron-positron annihilation, where the jets are defined by the Durham jet algorithm [166]. The cross section for n jets normalised to the LO cross section for $e^+e^- \to$ hadrons reads

$$\frac{\sigma_{n-jet}(\mu)}{\sigma_0(\mu)} = \left(\frac{\alpha_s(\mu)}{2\pi}\right)^{n-2} A_n(\mu) + \left(\frac{\alpha_s(\mu)}{2\pi}\right)^{n-1} B_n(\mu) + \mathcal{O}(\alpha_s^n). \tag{31}$$

One can expand the perturbative coefficient A_n and B_n in $1/N_c$:

$$A_n = N_c \left(\frac{N_c}{2}\right)^{n-2} \left[A_{n,\mathrm{lc}} + \mathcal{O}\left(\frac{1}{N_c}\right)\right], \qquad B_n = N_c \left(\frac{N_c}{2}\right)^{n-1} \left[B_{n,\mathrm{lc}} + \mathcal{O}\left(\frac{1}{N_c}\right)\right].$$

We calculate the leading order coefficient $A_{n,\mathrm{lc}}$ and the next-to-leading order coefficient $B_{n,\mathrm{lc}}$ for $n \leq 7$ at the renormalisation scale μ equal to the centre-of-mass energy. The centre-of-mass energy is taken to be equal to the mass of the Z-boson. The scale variation can be restored from the renormalisation group equation. The calculation is done with five massless quark flavours. Fig. 5 shows the comparison of our

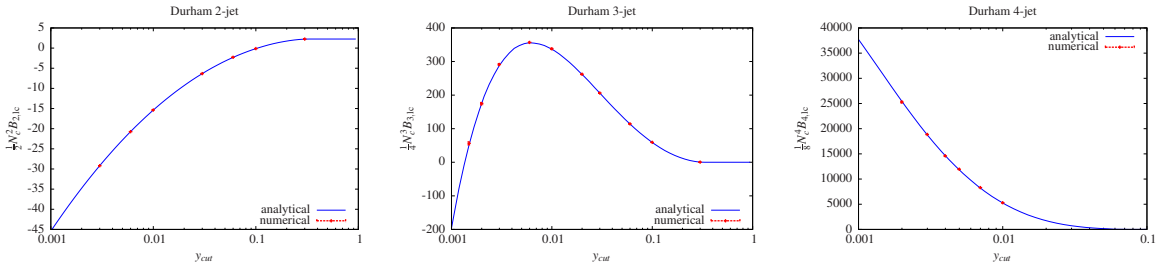


Fig. 5: Comparison of the NLO corrections to the two-, three- and four-jet rate between the numerical calculation and an analytic calculation. The error bars from the Monte Carlo integration are shown and are almost invisible.

numerical approach with the well-known results for two, three and four jets [179, 180, 181]. We observe an excellent agreement. The results for five, six and seven jets for the jet parameter $y_{cut} = 0.0006$ are

$$\begin{aligned}
\frac{N_c^4}{8} A_{5,\mathrm{lc}} &= (2.4764 \pm 0.0002) \cdot 10^4, & \frac{N_c^5}{16} B_{5,\mathrm{lc}} &= (1.84 \pm 0.15) \cdot 10^6, \\
\frac{N_c^5}{16} A_{6,\mathrm{lc}} &= (2.874 \pm 0.002) \cdot 10^5, & \frac{N_c^6}{32} B_{6,\mathrm{lc}} &= (3.88 \pm 0.18) \cdot 10^7, \\
\frac{N_c^6}{32} A_{7,\mathrm{lc}} &= (2.49 \pm 0.08) \cdot 10^6, & \frac{N_c^7}{64} B_{7,\mathrm{lc}} &= (5.4 \pm 0.3) \cdot 10^8.
\end{aligned} \tag{32}$$

3.6 FIRST STEPS TOWARDS NNLO

An NNLO calculation requires among other things also the calculation of the one-loop amplitude squared. The expansion in the dimensional regularisation parameter ε of the one-loop amplitude starts at order (-2) one would naively expect that up to order ε^0 the $\mathcal{O}(\uparrow)$- and $\mathcal{O}(\varepsilon^2)$-terms of the one-loop amplitude are needed for an NNLO calculation. However, it is by no means obvious how the approaches for one-loop amplitudes based on unitarity or the numerical method can be extended to include the higher-order terms in the ε-expansion. It turns out that the computation of these higher-order terms can be avoided, provided a method is known to compute the finite two-loop remainder function. The one- and two-loop amplitudes can be written as [182]

$$\begin{aligned}
\mathcal{A}^{(1)} &= \mathbf{Z}^{(1)}\mathcal{A}^{(0)} + \mathcal{F}^{(1)}_{\mathrm{minimal}}, \\
\mathcal{A}^{(2)} &= \left(\mathbf{Z}^{(2)} - \mathbf{Z}^{(1)}\mathbf{Z}^{(1)}\right)\mathcal{A}^{(0)} + \mathbf{Z}^{(1)}\mathcal{A}^{(1)} + \mathcal{F}^{(2)}_{\mathrm{minimal}},
\end{aligned} \tag{33}$$

where the operators $\mathbf{Z}^{(1)}$ and $\mathbf{Z}^{(2)}$ contain all the infrared poles and $\mathcal{F}^{(1)}_{\text{minimal}}$ and $\mathcal{F}^{(2)}_{\text{minimal}}$ are finite remainders. Here we used the convention that the operators $\mathbf{Z}^{(1)}$ and $\mathbf{Z}^{(2)}$ contain only pole terms, but no terms of order ε^k with $k \geq 0$. This corresponds to a minimal scheme. The operators $\mathbf{Z}^{(1)}$ and $\mathbf{Z}^{(2)}$ are well-known. At NNLO it is sufficient to know the ε^0-terms of $\mathcal{F}^{(1)}_{\text{minimal}}$ and $\mathcal{F}^{(2)}_{\text{minimal}}$, the ε^1- or ε^2-terms of $\mathcal{A}^{(1)}$ or $\mathcal{F}^{(1)}_{\text{minimal}}$ are not required [183].

ACKNOWLEDGEMENTS

I would like to thank Simon Plätzer for fruitful discussions during the workshop.

4. TOWARDS THE AUTOMATION OF ONE-LOOP AMPLITUDES [6]

Abstract

A program is presented that computes one-loop amplitudes automatically for processes with up to 6 external particles based on the Feynman-diagram approach. Additionally, universal one-loop building blocks, which can be used to compute several processes at NLO QCD are calculated.

4.1 INTRODUCTION

The calculation of processes with multi-particle final states beyond the leading order approximation has been an active field of research during the last years as a consequence of the demand of high accuracy for signal and background processes at the LHC. A next-to-leading (NLO) calculation consists of virtual and real radiation processes which are infrared divergent (IR) separately and can be computed numerically only after extracting the divergences of the real radiation contributions. The one-loop virtual calculation for multiple particles poseses a challenge of complexity not only due to the large number of contributing diagrams, but also concerning the stability of the numerical code to evaluate them. In the last years, an enormous progress has been achieved applying new techniques and using traditional Feynman-diagram approach, leading to new NLO predictions.

Due to the large number of processes of potential interest at the LHC, the scientific community has worked in the automation of the NLO calculations. The automation of the real contributions including their infrared subtraction terms has been successfully implemented in several packages and the automation of the virtual corrections, which is a harder problem, is currently being achieved in several programs (see [184] and references therein).

In Ref. [185], the early stage of a program, in the framework of Mathematica [186] and FeynCalc [187], to compute automatically one-loop amplitudes based on traditional Feynman-diagram techniques and involving up to $2 \to 4$ processes was presented. This program will become publicly available in the future. The method used is described in Section 4.2. In Section 4.3, we present a set of universal one-loop building blocks that has been used to compute recently several processes included in the VBFNLO package [42, 41].

4.2 TOWARDS AN AUTOMATIC ONE-LOOP AMPLITUDE GENERATOR

The program above mentioned automatically simplifies a set of amplitudes up to Hexagons of rank 5. The result is given in terms of scalar and tensor integrals following the Passarino-Veltman convention [112, 185], spinor chains, polarization vectors and model parameters. The simplified expression is written automatically to FORTRAN routines. For massless propagators, the amplitudes can be evaluated also in Mathematica with unlimited precision, which is used for testing purposes. To achieve that, the scalar integrals, the tensor reduction formalism to extract the tensor coefficient integrals, and also the helicity method described in Ref. [188, 189] to compute the spinor products have been implemented at

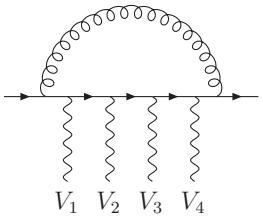
[6] Contributed by: F. Campanario

the FORTRAN and Mathematica level. For the determination of the tensor integrals up to the box level, the Passarino-Veltman tensor reduction formalism [112] is used applying the LU decomposition method to avoid the explicit calculation of inverse Gram matrices by solving a system of linear equations, which is a more stable procedure close to singular points. Finally, for singular Gram determinants, special tensor reduction routines following Ref. [115] have been implemented, however, the external momenta convention (Passarino-like) was used. The impact of these methods is discussed in detail in Ref. [185]. For pentagons, in addition to the Passarino-Veltman formalism, the method proposed by Denner and Dittmaier [115, 190], applied also to hexagons, has been implemented. For that, the recursion relations of Ref. [115] in terms of the Passarino-Veltman external momenta convention have been re-derived. This last method is used for the numerical implementation at the FORTRAN level.

The Mathematica function does several algebraic manipulations that are summarized as follows:

- Simultaneous extraction of rational terms based on Dirac algebra manipulations and cancelation of scalar products against propagators.
- Reduction to a minimal basis of tensor and scalar integrals.
- Reduction to a minimal basis of spinor chains.
- The use of Chisholm identities, which are only valid in 4 dimensions, for the contraction of Lorentz indices among different spinor chains is applied, if selected.
- Factorization of loop dependent and independent factors (Useful to perform gauge tests, Ward identities or the re-evaluation of the amplitudes for different helicity polarization of gluons and fermions at a lower CPU cost).

As an example of the notation used, the following Hexagon diagram is used. This is written as follows:



$$= \mathcal{M}_{V_1V_2V_3V_4,\tau} = g_\tau^{V_1 f} g_\tau^{V_2 f} g_\tau^{V_3 f} g_\tau^{V_4 f} \frac{g_0^2}{(4\pi)^2} \mathcal{C}_{ij}^{V_1V_2V_3V_4} \mathcal{M}_\tau^{ij}, \quad (34)$$

where g_0 is the strong unrenormalized coupling, $\mathcal{C}_{ij}^{V_1V_2V_3V_4}$ is a color diagram dependent factor, e.g, $\mathcal{C}_{ij}^{\gamma\gamma g\gamma} = (T_a)_{ij}(C_F - 1/2C_A)$. $g_\tau^{V_i f}$ are electroweak couplings and $\mathcal{M}_\tau^{ij}$ represents the amplitude considering generic off-shell vector bosons with color indices ij for a given helicity τ. The amplitude $\mathcal{M}_\tau^{ij}$, omitting color indices, is written in terms of

$$\mathcal{M}_\tau = \mathcal{M}_\tau^{D=4} + (D-4)\mathcal{M}_\tau^{DR}, \quad (35)$$

where $\mathcal{M}_\tau^{D=4}$ is the amplitude that one would obtain performing the Dirac algebra manipulation in four dimensions, $D = 4$, and $\mathcal{M}_\tau^{DR}$ contains the rational terms and vanishes in Dimensional Reduction (DR). These functions are decomposed in the form:

$$\mathcal{M}^{(D=4,DR)} = \sum_{i,j} \mathrm{SM}_{i,\tau}\ \mathrm{F1}_j, \quad (36)$$

where $\mathrm{SM}_{i,\tau}$ is a basis of Standard Matrix elements corresponding to spinor products describing the quark line of Eq. (34) which are computed following the helicity method [188, 189] with a defined helicity, τ. $\mathrm{F1}_j$ are complex functions which are further decomposed into dependent and independent loop integral parts,

$$\mathrm{F1}_j = \sum_{l,k} \mathrm{F}_l T_k \Big(\epsilon(p_n) \cdot p_m; \epsilon(p_i) \cdot \epsilon(p_r) \Big). \quad (37)$$

T_k is a monomial function at most for each polarization vector $\epsilon(p_x)$, i.e., $\epsilon(p_x)^0$ or $\epsilon(p_x)^1$. The first possibility, $\epsilon(p_x)^0$, implies that the polarization vector appears in the set of Standard Matrix elements

$\mathrm{SM}_{i,\tau}$. F_l contains kinematic variables $(p_i \cdot p_j)$, the scalar integrals (B_0, C_0, D_0), and the tensor integral coefficients $(B_{ij}, C_{ij}, D_{ij}, E_{ij}, F_{ij})$. Then, the full result is obtained from $\mathcal{M}_\tau^{D=4}$ and $\mathcal{M}_\tau^{DR}$ using the finite and the coefficients of the $1/\epsilon^n$ poles of the scalar and tensor coefficient integrals:

$$\mathcal{M}_v^{D=4} = \widetilde{\mathcal{M}}_v + \frac{\mathcal{M}_v^1}{\epsilon} + \frac{\mathcal{M}_v^2}{\epsilon^2}, \qquad (D-4)\mathcal{M}_v^{DR} = \widetilde{\mathcal{N}}_v + \frac{\mathcal{N}_v^1}{\epsilon}, \tag{38}$$

where, e.g., $\widetilde{\mathcal{M}}_v$ is the finite contribution obtained using the finite pieces of the scalar and tensor coefficient integrals including the finite contributions from rational terms arising in ultraviolet tensor coefficient integrals.

4.3 UNIVERSAL BUILDING BLOCKS

Based on the observation that the same one-loop virtual amplitudes appear in many processes (Fig. 6), we are aiming to collect a basis of universal building blocks, which can be used to compute all of the $2 \to 4$ processes at LHC at the QCD one-loop level (Similar to the philosophy of older versions of MADGRAPH [191] calling the HELAS [192] routines). This methodology of collecting topologies in groups has been proved very successful in the program VBFNLO, where for example a boxline routine, first line of Fig. 6, is computed and applied to $pp \to VV$, $pp \to VVV$, $pp \to VVj$ and EW production of $pp \to Vjj$ and $pp \to HVjj$.

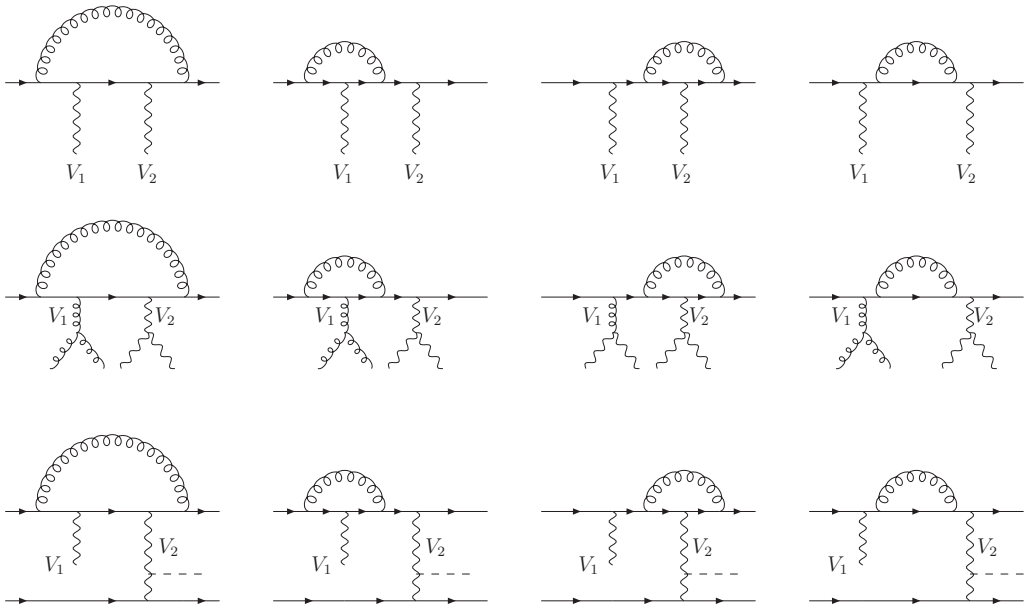


Fig. 6: Boxline contributions appearing in different processes.

To do that, we use the effective current approach described and applied in Refs. [39, 47, 31, 193, 21]. As illustration, the first diagram of the second raw of Fig. 6 is used. This can be written as,

$$A_{V_1V_2V_3V_4,\tau} = J^{\mu_1}_{V_1^*} J^{\mu_2}_{V_2^*} \mathcal{M}_{\mu_1\mu_2,\tau} \equiv \mathcal{M}_{V_1^*V_2^*,\tau}, \tag{39}$$

where the color indices have been omitted. Here, $J^{\mu_1}_{V_1^*}$ and $J^{\mu_2}_{V_2^*}$ represent effective polarization vectors in the unitarity gauge for the EW sector including finite width effects in the scheme of Refs. [194, 195] and propagator factors, e.g.,

$$J^{\mu_1}_{V_1^*}(q_1) = \frac{-i}{q_1^2 - M^2_{V_1^*} - i\, M_{V_1^*}\Gamma_{V_1^*}} \left(g^{\mu_1}_{\mu} - \frac{q_1^{\mu_1} q_{1\mu}}{q_1^2 - M^2_{V_1^*} - i\, M_{V_1^*}\Gamma_{V_1^*}} \right) \Gamma^{\mu}_{V_1^* V_1 V_3}, \tag{40}$$

with $\Gamma_{V_1^*}$, the width of the V_1^* vector boson, and $\Gamma^{\mu}_{V_1^* V_1 V_3}$, the triple vertex, which can also contain the leptonic decay of the EW vector bosons including all off-shell effects or BSM physics. In this manner, we can then concentrate in computing, instead of $A_{V_1 V_2 V_3 V_4, \tau}$, the virtual correction to two massive vector bosons attached to the quark line, $\mathcal{M}_{V_1^* V_2^*, \tau}$, or equivalently $\mathcal{M}_{\mu_1 \mu_2, \tau}$, where the polarization vectors or effective currents have been factored out. In our approach, this basic building block is the so-called Boxline, which is computed only once and re-used in different processes.

We plan to do a classification of all the topologies that appear at 1 loop level for up to $2 \to 4$ processes and install a library with all the basic one-loop building blocks already computed and simplified. This would be an advantage since, for example for $qq \to VVVV$ production, up to 24 hexagons for a single subprocess would appear, corresponding to the permutations of the vector bosons on the hexagon of Eq. 34. In this approach, the amplitude is obtained by calling the same one-loop amplitude 24 times with the corresponding ordering of momenta and polarization vectors. We aim towards an automation of this procedure, which will result into a faster and shorter final FORTRAN code generation. The specific building blocks are collected into groups with specific gauge and IR factorization properties, e.g, factorization of the IR divergences against the corresponding born, known behavior under Ward identity checks.

In Fig. 7, we present the topologies that have been computed and tested. In the first line, corrections to a quark line with the emission of V_n vector bosons in a fixed order are represented for 4 different topologies. (The first 2 were explained in detail in Ref. [185], including their stability behavior). We have only depicted the virtual amplitude with the higher complexity for a giving building block, e.g. the boxline of Fig.6 is obtained from the first diagram with two vector bosons attached, i.e., $n = 2$ in V_n. The first two topologies of the second line are collected by putting together all possible Feynman-diagrams with a fixed order of the vector bosons and attaching it to the quark lines in all possible ways. The crossing of the fermion lines are treated as independent building blocks and are not depicted. Finally, the fermion-loop corrections for a fixed order of vector bosons, V_n, are computed in the last diagram of the second line

The use of modular structure routines, as the above presented, has been proved to be an advantage in the program *VBFNLO* [42, 41] since once a structure is computed and checked it can be re-used for different processes. For example, using the building blocks of the first and second topology together with the fermion-loop diagrams, results at NLO QCD for all VVV [39, 47, 46, 45, 43, 44], several VVj [31, 32, 196, 197], $H\gamma jj$ [198] and $W\gamma\gamma j$ [21] production channels have been computed recently. The last one representing the first calculation at this accuracy falling in the category of $VVV + j$ production. Up to the pentagon level, these building blocks are publicly available as part of the *VBFNLO* [42, 41] package together with the tensor reduction routines, excluding the routines for small Gram determinants which will become available in the future, in addition to the other building blocks.

4.4 CONCLUSIONS

A program which automatically evaluates one-loop amplitudes for up to $2 \to 4$ processes has been presented based on the traditional Feynman-diagram approach. The program has been developed in the framework of Mathematica and FeynCalc and writes down automatically the simplified expression to FORTRAN. Up to the pentagon level and for massless propagators, the code can be evaluated numerically inside Mathematica with unlimited precision which can be used for testing purposes. For the reduction of tensor integrals, we have developed a library that includes expansion for small Gram determinants. Using the leptonic tensor formalism, we are building a library of universal one-loop building blocks, which can be used to compute several processes at NLO QCD. Recently, following this strategy, we have reported results for all VVV [39, 47, 46, 45, 43, 44], several VVj [31, 32, 196, 197], $H\gamma jj$ [198] and $W\gamma\gamma j$ [21] production channels inside the VBFNLO collaboration. The ultimate goal is to generalize the library to compute all of the $2 \to 4$ processes at LHC at the QCD one-loop level, similar to the philosophy of older versions of MADGRAPH [191] calling the HELAS [192] routines,

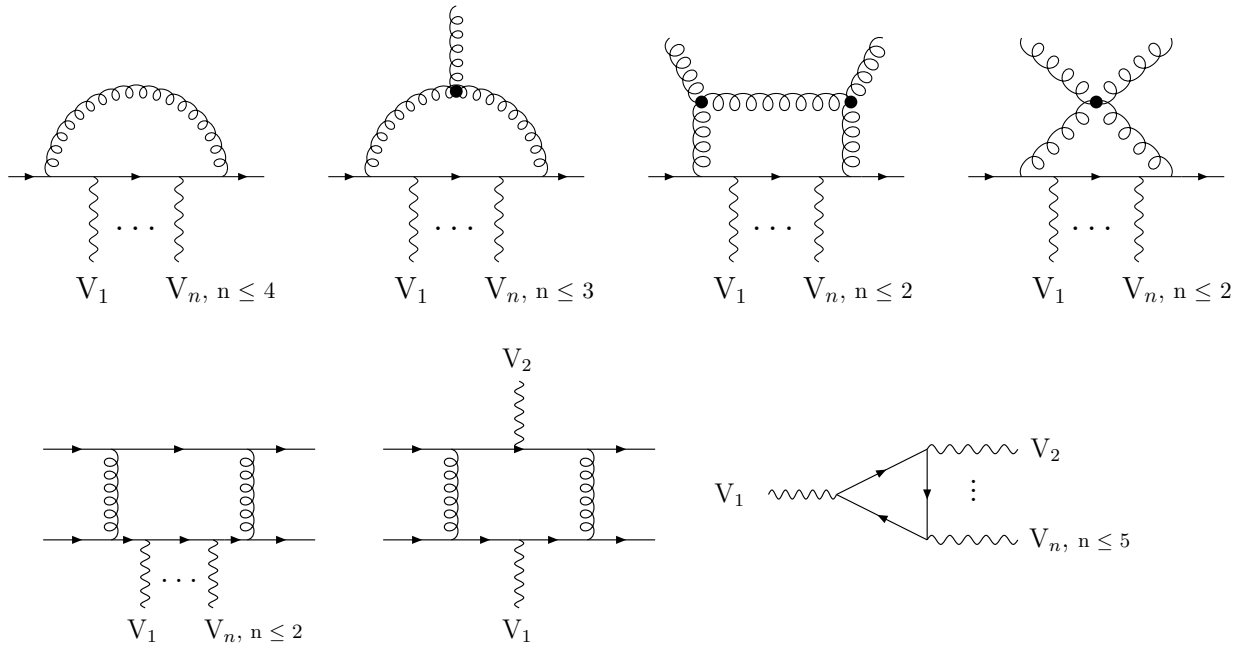


Fig. 7: Topologies of universal building one-loop blocks. Only the most complicated diagram of each topology is depicted, e.g, the boxline of Fig.6 is obtained from the first diagram with two vector bosons attached, i.e., $V_n, n = 2$.

and deliver a Mathematica package compatible with FeynArts [132], which can be used to compute full one-loop amplitudes automatically using the universal building blocks, resulting into a faster and shorter code generation.

ACKNOWLEDGEMENTS

F.C acknowledges fruitful discussions with the members of the VBFNLO collaboration and partial support of a postdoctoral fellowship of the Generalitat Valenciana, Spain (Beca Postdoctoral d'Excel·lència), by the European FEDER and Spanish MICINN under the grant FPA2008-02878 and by the Deutsche Forschungsgemeinschaft via the Sonderforschungsbereich/Transregio SFB/TR-9 "Computational Particle Physics". The Feynman diagrams in this paper were drawn using Axodraw [139].

5. THE TWO-LOOP QCD VIRTUAL AMPLITUDE FOR W PAIR PRODUCTION WITH FULL MASS DEPENDENCE [7]

5.1 INTRODUCTION

One of the main aims in the Large Hadron Collider (LHC) physics program is undoubtedly the discovery (or the exclusion) of the Higgs boson which is responsible for the fermion and gauge boson masses and also part of the mechanism of dynamical breaking of the Electroweak (EW) symmetry. Another important goal for the LHC is the precise measurement of the hadronic production of gauge boson pairs, WW, WZ, ZZ, $W\gamma$, $Z\gamma$, this in connection to the investigation of the non-Abelian gauge structure of the SM. W pair production,

$$q\bar{q} \to W^+ W^- , \tag{41}$$

[7] Contributed by: G. Chachamis

plays an essential role as it serves as a signal process in the search for New Physics and also is the dominant irreducible background to the Higgs discovery channel $pp \to H \to W^*W^* \to l\bar{\nu}\bar{l}'\nu'$, in the intermediate Higgs mass range [199]. Both ATLAS and CMS collaborations have released first values for the WW cross section [200, 201].

The process (41) is currently known at next-to-leading order (NLO) accuracy [202, 203, 204, 205, 206, 157, **?**]. The NLO corrections were proven to be large enhancing the tree-level result by almost 70% which falls to a (still) large 30% after imposing a jet veto. Therefore, if a theoretical estimate for the W pair production is to be compared against experimental measurements at the LHC, one is bound to go one order higher in the perturbative expansion, namely, to the next-to-next-to-leading order (NNLO). This would allow, in principle, an accuracy of around 10%.

High accuracy for the W pair production is also needed when the process is studied as background to Higgs production in order to match accuracies between signal and background. The signal process for the Higgs discovery via gluon fusion, $gg \to H$, as well as the process $H \to WW \to l\bar{\nu}\bar{l}'\nu'$ are known at NNLO [207, 208, 209, 210, 211, 212, 213, 214, 215, 216], whereas the EW corrections are known beyond NLO [217]. Another process that needs to be included in the background is the W pair production in the loop induced gluon fusion channel,

$$gg \to W^+W^- \,. \tag{42}$$

The latter contributes at $\mathcal{O}(\alpha_s^2)$ relative to the quark-anti-quark-annihilation channel but is nevertheless enhanced due to the large gluon flux at the LHC [218, 219].

The first main difficulty in studying the NNLO QCD corrections for W pair production is the calculation of the two-loop virtual amplitude since it is a $2 \to 2$ process with massive external particles. We have already computed the virtual corrections at the high energy limit [220, 97, 221]. However, this is not enough as it cannot cover the kinematical region close to threshold. Therefore, in order to cover all kinematical regions we proceed as follows. We perform a deep expansion in the W mass around the high energy limit which in combination with the method of numerical integration of differential equations [222, 223, 224] allows us the numerical computation of the two-loop amplitude with full mass dependence over the whole phase space.

5.2 THE HIGH ENERGY LIMIT

The methodology for obtaining the massive amplitude in the high energy limit, namely the limit where all the invariants are much larger than the W mass, is similar to the one followed in Refs. [225, 226]. The amplitude is reduced to an expression that only contains a small number of integrals (master integrals) with the help of the Laporta algorithm [227]. In the calculation for the two-loop amplitude there are 71 master integrals. Next step is the construction, in a fully automatised way, of the Mellin-Barnes (MB) representations [228, 229] of all the master integrals by using the **MBrepresentation** package [230]. The representations are then analytically continued in the number of space-time dimensions by means of the **MB** package [231], thus revealing the full singularity structure. An asymptotic expansion in the mass parameter (W mass) is performed by closing contours and the integrals are finally resummed, either with the help of **XSummer** [232] or the **PSLQ** algorithm [233]. The result is expressed in terms of harmonic polylogarithms.

5.3 POWER CORRECTIONS AND NUMERICAL EVALUATION

The high energy limit by itself is not enough, as was mentioned before. The next step, following the methods applied in Ref. [234], is to compute power corrections in the W mass. Power corrections are good enough to cover most of the phase space, apart from the region near threshold as well as the regions corresponding to small angle scattering.

We recapitulate here some of the notation of Ref. [221] for completeness. The charged vector-boson production in the leading partonic scattering process corresponds to

$$q(p_1) + \overline{q}(p_2) \;\; \rightarrow \;\; W^-(p_3, m) + W^+(p_4, m)\,, \tag{43}$$

where p_i denote the quark and W momenta and m is the mass of the W boson.

We have chosen to express the amplitude in terms of the kinematic variables x and m_s which are defined to be

$$x = -\frac{t}{s}, \;\; m_s = \frac{m^2}{s}, \tag{44}$$

where

$$s = (p_1 + p_2)^2 \;\; \text{and} \;\; t = (p_1 - p_3)^2 - m^2\,. \tag{45}$$

The variation then of x within the range $[1/2(1-\beta), 1/2(1+\beta)]$, where $\beta = \sqrt{1-4m^2/s}$ is the velocity, corresponds to angular variation between the forward and backward scattering.

It should be evident that any master integral M_i can be written then as

$$M_i = M_i\,(m_s, x, \epsilon) = \sum_{j=k}^{l} \epsilon^j I_{ij}(m_s, x), \tag{46}$$

where ϵ is the usual regulator in dimensional regularization ($d = 4 - 2\epsilon$) and the lowest power of ϵ in the sum can be -4.

The crucial point now is that the derivative of any Feynman integral with respect to any kinematical variable is again a Feynman integral with possibly higher powers of denominators or numerators which can also be reduced anew in terms of the initial set of master integrals. This means that one can construct a partially triangular system of differential equations in the mass, which can subsequently be solved in the form of a power series expansion, with the expansion parameter in our case being m_s following the conventions above.

Let us differentiate with respect to m_s and x, we will then have respectively

$$m_s \frac{d}{dm_s} M_i(m_s, x, \epsilon) = \sum_j C_{ij}(m_s, x, \epsilon)\; M_j(m_s, x, \epsilon) \tag{47}$$

and

$$x \frac{d}{dx} M_i(m_s, x, \epsilon) = \sum_j C'_{ij}(m_s, x, \epsilon)\; M_j(m_s, x, \epsilon)\,. \tag{48}$$

We use Eq. (47) to obtain the mass corrections for the master integrals calculating the power series expansion up to order m_s^{11} (see also Ref. [234] for more details). This deep expansion in m_s should be sufficient for most of the phase space but still not enough to cover the whole allowed kinematical region. The way to proceed from this point is to numerically integrate the system of differential equations.

In particular, we choose to work with the master integrals in the form of Eq. (46), where the ϵ dependence is explicit. We can then work with the coefficients of the ϵ terms and accordingly have

$$m_s \frac{d}{dm_s} I_i(m_s, x) = \sum_j J^M_{ij}\,(m_s, x)\; I_j(m_s, x) \tag{49}$$

and

$$x \frac{d}{dx} I_i(m_s, x) = \sum_j J^X_{ij}\,(m_s, x)\; I_j(m_s, x), \tag{50}$$

where the Jacobian matrices J^M and J^X have rational function elements.

By using this last system of differential equations, one can obtain a full numerical solution to the problem. What we are essentially dealing now with is an initial value problem and the main requirement is to have the initial conditions to proper accuracy. The initial conditions, namely the values of the master integrals at a proper kinematical point which we call initial point, are provided by the power series expansion. The initial point has to be chosen somewhere in the high energy limit region, where m_s is small and therefore, the values obtained by the power series are very accurate. Starting from there, one can evolve to any other point of the phase space by numerically integrating the system of differential equations Eq. (49) and Eq. (50).

We parametrise with a suitable grid of points the region close to threshold and then we calculate the master integrals for all points of the grid by evolving as described previously. Given that the master integrals have to be very smooth (we remain above all thresholds) one can use, after having the values for the grid points, interpolation to get the values at any point of the region. We use 1600 points for the grid and take as initial conditions the values of the master integrals at the point $m_s = 5 \times 10^{-3}$, $x = 1/4$. The relative errors at that point were estimated not to exceed 10^{-18}.

The numerical integration is performed by using one of the most advanced software packages implementing the variable coefficient multistep method (ODEPACK) [235]. We use quadruple precision to maximise accuracy. The values at any single grid point can be obtained in about 15 minutes in average (with a typical 2GHz Intel Core 2 Duo system) after compilation with the Intel Fortran compiler. The accuracy is around 10 digits for most of the points of the grid. It is also worth noting that in order to perform the numerical integration one needs to deform the contour in the complex plane away from the real axis. This is due to the fact that along the real axis there are spurious singularities. We use an elliptic contour and we achieve a better estimate of the final global error by calculating more than once for each point of the grid, using each time different eccentricities. Grids of solutions can actually be constructed, which will be subsequently interpolated when implemented as part of a Monte Carlo program.

One very stringent test we use to cross-check the correctness and also the accuracy of our calculation is to compare the infrared pole structure of our two-loop result against the one predicted by Catani [182] (see also Refs. [236, 237, 238]). According to Catani, the infrared poles of the interference of the tree and the two-loop amplitudes follow a generic formula which in our case, since we work with the rescaled variables m_s and x, can be cast into the following form:

$$\mathcal{C}^{(0\times 2)}_{atani}(m_s, x, \frac{s}{\mu}) = 2\mathrm{Re}\left\{\mathrm{I}^{(1)}(\text{ffl})\langle \mathrm{M}^{(0)}|\mathrm{M}^{(1)}\rangle + \mathrm{I}^{(2)}(\text{ffl})\langle \mathrm{M}^{(0)}|\mathrm{M}^{(0)}\rangle\right\}, \tag{51}$$

where $\mathrm{M}^{(0)}$ and $\mathrm{M}^{(1)}$ are the tree level and one-loop amplitudes respectively and μ is the renormalization scale. The operators $\mathrm{I}^{(1)}(\text{ffl})$ and $\mathrm{I}^{(2)}(\text{ffl})$ encode the information for the infrared pole structure and their exact expressions can be found in Ref. [97].

The way to perform the test is straightforward. For each point of the grid with coordinates $(m_{s(i)}, x_{(i)})$, we compute the numerical value of the two-loop amplitude ($\mathrm{M}^{(2)}$) interfered with the tree level amplitude

$$\mathcal{A}^{(0\times 2)}(m_{s(i)}, x_i, \frac{s}{\mu}) = \langle \mathrm{M}^{(0)}|\mathrm{M}^{(2)}\rangle + \langle \mathrm{M}^{(2)}|\mathrm{M}^{(0)}\rangle \tag{52}$$

by numerically integrating the differential equations as described previously and we also calculate the numerical value of the quantity $\mathcal{C}^{(0\times 2)}_{atani}(m_{s(i)}, x_{(i)}, \frac{s}{\mu})$ by using Eq. (51). Then, all we need to make sure is that the infrared singularities of the quantity $\left\{\mathcal{A}^{(0\times 2)}(m_{s(i)}, x_i, \frac{s}{\mu}) - \mathcal{C}^{(0\times 2)}_{atani}(m_{s(i)}, x_i, \frac{s}{\mu})\right\}$ cancel numerically for every point $(m_{s(i)}, x_{(i)})$ of the grid (ultraviolet divergencies have been removed by renormalization). We will not present here any numbers since the aim was to describe the general methods. The details and the results of the study will be presented in a future publication [239].

5.4 CONCLUSIONS

W pair production via quark-anti-quark-annihilation is an important signal process in the search for New Physics as well as the dominant irreducible background for one of the main Higgs discovery channels: $H \to WW \to 4$ leptons. Therefore, the accurate knowledge of this process is essential for the LHC. After having calculated the two-loop and the one-loop-squared virtual QCD corrections to the W boson pair production in the high energy limit we proceed to the next step. Namely, we use a combination of a deep expansion in the W mass around the high energy limit and of numerical integration of differential equations to compute the two-loop amplitude with full mass dependence over the whole phase space. A strigent cross-check of our calculation is to verify that the infrared structure of our result agrees with the prediction of the Catani formalism for the infrared structure of QCD amplitudes.

ACKNOWLEDGEMENTS

The author wishes to thank the Les Houches Workshop organizers for the friendly and stimulating atmosphere.

6. COMPUTATION OF INTEGRATED SUBTRACTION TERMS NUMERICALLY [8]

Abstract

We report on a numerical representation of the integrated subtraction terms of the NNLO subtraction scheme defined in Refs. [240, 241, 242, 243]. The integrated approximate cross sections themselves can be written as products of insertion operators (in colour space) times the Born, or the one-loop cross section. The insertion operator is constructed from the numerical representation of the integrated subtraction terms. We give selected results for the integrated doubly-collinear subtraction term.

6.1 INTRODUCTION

We consider the NNLO correction to a generic m-jet observable,

$$\sigma^{\rm NNLO} = \int_{m+2} {\rm d}\sigma^{\rm RR}_{m+2} J_{m+2} + \int_{m+1} {\rm d}\sigma^{\rm RV}_{m+1} J_{m+1} + \int_{m} {\rm d}\sigma^{\rm VV}_{m} J_{m}\,. \tag{53}$$

The three contributions on the right hand side are separately divergent in $d = 4$ dimensions, but their sum is finite for IR safe observables. To obtain the finite NNLO correction, we first continue analytically all integrals to $d = 4 - 2\epsilon$ dimensions and then rewrite Eqn. (53) as

$$\sigma^{\rm NNLO} = \int_{m+2} {\rm d}\sigma^{\rm NNLO}_{m+2} + \int_{m+1} {\rm d}\sigma^{\rm NNLO}_{m+1} + \int_{m} {\rm d}\sigma^{\rm NNLO}_{m}\,, \tag{54}$$

that is a sum of three integrals where the integrands,

$${\rm d}\sigma^{\rm NNLO}_{m+2} = \Big\{ {\rm d}\sigma^{\rm RR}_{m+2} J_{m+2} - {\rm d}\sigma^{\rm RR,A_2}_{m+2} J_{m} - \Big[{\rm d}\sigma^{\rm RR,A_1}_{m+2} J_{m+1} - {\rm d}\sigma^{\rm RR,A_{12}}_{m+2} J_{m} \Big] \Big\}_{\epsilon=0}\,, \tag{55}$$

$${\rm d}\sigma^{\rm NNLO}_{m+1} = \Big\{ \Big[{\rm d}\sigma^{\rm RV}_{m+1} + \int_1 {\rm d}\sigma^{\rm RR,A_1}_{m+2} \Big] J_{m+1} - \Big[{\rm d}\sigma^{\rm RV,A_1}_{m+1} + \Big(\int_1 {\rm d}\sigma^{\rm RR,A_1}_{m+2} \Big)^{\rm A_1} \Big] J_{m} \Big\}_{\epsilon=0}\,, \tag{56}$$

and

$${\rm d}\sigma^{\rm NNLO}_{m} = \Big\{ {\rm d}\sigma^{\rm VV}_{m} + \int_2 \Big[{\rm d}\sigma^{\rm RR,A_2}_{m+2} - {\rm d}\sigma^{\rm RR,A_{12}}_{m+2} \Big] + \int_1 \Big[{\rm d}\sigma^{\rm RV,A_1}_{m+1} + \Big(\int_1 {\rm d}\sigma^{\rm RR,A_1}_{m+2} \Big)^{\rm A_1} \Big] \Big\}_{\epsilon=0} J_{m}\,, \tag{57}$$

[8] Contributed by: G. Somogyi, Z. Szőr, Z. Trócsányi

are integrable in four dimensions by construction. The approximate cross sections $\mathrm{d}\sigma^{\mathrm{RR,A_2}}_{m+2}$ and $\mathrm{d}\sigma^{\mathrm{RR,A_1}}_{m+2}$ regularise the doubly- and singly-unresolved limits of the real-emission contribution, $\mathrm{d}\sigma^{\mathrm{RR}}_{m+2}$ respectively. The double subtraction due to the overlap of these two terms is compensated by $\mathrm{d}\sigma^{\mathrm{RR,A_{12}}}_{m+2}$. These terms are given explicitly in Ref. [242]. Finally, $\mathrm{d}\sigma^{\mathrm{RV,A_1}}_{m+1}$ and $\left(\int_1 \mathrm{d}\sigma^{\mathrm{RR,A_1}}_{m+2}\right)^{\mathrm{A_1}}$ regularise the singly-unresolved limits of $\mathrm{d}\sigma^{\mathrm{RV}}_{m+1}$ and $\int_1 \mathrm{d}\sigma^{\mathrm{RR,A_1}}_{m+2}$ respectively. They are given explicitly in Ref. [243].

The construction of each approximate cross section in Eqns. (55–57) is based on the known and universal IR limits of tree level and one-loop squared matrix elements, and proceeds in two steps. First, the IR factorisation formulae are written in such a way that their complicated overlap structure can be disentangled ("matching of limits") [240, 244]. Second, we define "extensions" of the formulae, so that they are unambiguously defined away from the strict IR limits [241, 242, 243]. These extensions are defined by the use of various momentum mappings that map a set of $m+1$ or $m+2$ momenta into a set of m momenta,

$$\{p\}_{m+1} \longrightarrow \{\tilde{p}\}_m \quad \text{and} \quad \{p\}_{m+2} \longrightarrow \{\tilde{p}\}_m \,, \tag{58}$$

such that (i) the delicate structure of cancellations among the matched limit formulae in various limits is respected (ii) exact momentum conservation is implemented, and (iii) the original $m+1$ or $m+2$ particle phase space factorises exactly into the product of an m particle phase space and a one- or two-particle phase space measure,

$$\mathrm{d}\phi_{m+r}(\{p\}_{m+r}; Q) = \mathrm{d}\phi_m(\{\tilde{p}\}_m; Q)[\mathrm{d}p_{r,m}] \,, \qquad r = 1, 2 \,. \tag{59}$$

To finish the definition of the scheme, one must compute once and for all the one- and two-particle integrals, denoted formally as $\int_1$ and $\int_2$, appearing in Eqns. (56–57).

In general the integrated subtraction terms are integrals of extensions over the whole phase space of combinations of the QCD splitting functions and squared soft currents. In this proceedings we discuss two examples: (i) the singly-collinear subtractions $\mathcal{C}^{(\ell,0)}_{ir}$ and (ii) the doubly-collinear subtractions $\mathcal{C}^{(0,0)}_{ir,js}$, which are part of $\mathrm{d}\sigma^{\mathrm{RR,A_1}}_{m+2}$ and $\mathrm{d}\sigma^{\mathrm{RR,A_2}}_{m+2}$ in Eqn. (55), respectively. The precise definitions of these terms can be found in Ref. [242]. The meaning of the superscript is irrelevant for our present purpose (also explained in Ref. [242]).

Denoting a generic subtraction term by $\mathcal{X}^{(\ell,k)}$ (such as $\mathcal{C}^{(\ell,0)}_{ir}$) the integrated counterterms can be written in the following general form:

$$\int_r \mathcal{X}^{(\ell,k)} = \left[\frac{\alpha_{\mathrm{s}}}{2\pi} S_\epsilon \left(\frac{\mu^2}{Q^2}\right)^\epsilon\right]^{r+\ell} N_X(\epsilon) X^{(\ell)}(x,\ldots) \mathrm{Re}\langle \mathcal{M}^{(0)}_m(\{\tilde{p}\}) | \boldsymbol{T}_i \cdot \boldsymbol{T}_j \ldots | \mathcal{M}^{(k)}_m(\{\tilde{p}\})\rangle \,, \tag{60}$$

where $S_\epsilon = (4\pi)^\epsilon / \Gamma(1-\epsilon)$, and $X^{(\ell)}(x,\ldots)$ represents a function that depends on kinematical invariants of the factorized m-parton phase space. It results in the integration of the subtraction term $\mathcal{X}^{(\ell,k)}$ over the factorized phase spaces $[\mathrm{d}p_{r,m}]$ in Eqn. (59). In a NNLO computation the possible cases are $r+\ell+k=1$ with $\ell+k=0$ or 1, and $r=2$ with $\ell+k=0$. We use the colour- and spin-state notation of Ref. [236], when the amplitude for a scattering process involving m final-state momenta, $|\mathcal{M}^{(k)}_m\rangle$, is an abstract vector in colour and spin space; k denotes the number of loops. Colour interactions at QCD vertices are represented by associating colour charges $\boldsymbol{T}_i$ with the emission of a gluon from each parton i. There are $2r$ such colour charges. Then the functions $X^{(\ell)}$ are dimensionless in colour-space. For certain subtraction terms, universal, possibly ϵ-dependent numerical factors, $N_X(\epsilon)$ appear naturally, which can be factored out. Our purpose is to compute all functions $X^{(\ell)}$, which we discuss next.

6.2 INTEGRATING THE COUNTERTERMS

The actual computation of the integrated counterterms leads to a large number of multi-dimensional integrals. The ultimate goal is to find the analytical form of the coefficients of a Laurent expansion (in

ϵ) of these integrals, which turns out to be a rather tedious job. In order to compute these coefficients as efficiently as possible, we have explored several methods.

First, it is possible to extend the method of integration-by-parts identities and solving of differential equations, developed for computing multi-loop Feynman integrals [245, 246], to the relevant phase space integrations [247]. This method yields ϵ-expansions with fully analytical coefficients, with the final results being expressed in terms of two-dimensional harmonic polylogarithms (after a suitable basis extension, see Ref. [247] for details). This approach was used successfully to compute a class of singly-unresolved integrals [247].

Second, the phase space integrals that arise can be computed via the method of Mellin–Barnes (MB) representations [228, 229, 248]. Here we obtain the ϵ-expansion coefficients in terms of complex contour integrals over Γ-functions. Performing these integrals by the use of the residue theorem, a representation in terms of harmonic sums is obtained. In many cases, the sums can be evaluated in a closed form, yielding an analytical result. In some instances however, we find multi-dimensional MB integrals that are very difficult to compute fully analytically. Nevertheless, in these situations a direct numerical evaluation of the appropriate MB representations provides a fast and reliable way to obtain final results with small numerical uncertainties. We stress that for phenomenological applications, this is all that is required, since the numerical uncertainty of the complete computation is dominated by the phase space integrations. We have used the MB method to compute all singly-unresolved integrals [249], and all two-particle integrals appearing in $\int_2 \mathrm{d}\sigma_{m+2}^{RR,A_{12}}$ as well [91].

Finally, the method of iterated sector decomposition [250] can also be used to calculate the integrals we encounter [251]. Sector decomposition produces a representation of the ϵ-expansion where the coefficients are given in terms of (mostly quite cumbersome) finite integrals over the unit hypercube. The analytical evaluation of these integrals is not feasible except for the simplest cases. Nevertheless, this method is simple to implement and can be automated to a large extent. In fact there are several computer programs that use various implementations of sector decomposition to provide numerical values of coefficients of the powers of ϵ in the Laurent expansion of dimensionally regulated integrals [252, 253, 254]. We found the program `SecDec` powerful and flexible to generate sufficiently precise values of our integrated subtraction terms.

Choosing the Cuhre integrator implemented in `SecDec`, we can easily reach 10^{-7} relative precision for the integration. Such precision is sufficient for our purposes: (i) to demonstrate the cancellation of the ϵ poles numerically, and (ii) to compute the finite integrals in Eqns. (56) and (57). As the numerical uncertainty of the second item is limited more by the Monte Carlo integration over the $m+1$ and m particle phase spaces, for item (ii) much lower (not better than 10^{-3}) precision is sufficient. This looser requirement on the precision for the O(1) terms and the fact that the integrated subtraction terms are smooth functions of their parameters, with logarithmic behaviour for asymptotically small values of the parameters, makes possible that we find sufficient approximations to the integrated subtraction terms.

6.3 APPROXIMATE INTEGRATED SUBTRACTION TERMS

The computation of the integrated subtraction terms at any given values of the kinematical parameters, as required in the Monte Carlo integration over the phase space, is not feasible. In order to demonstrate the cancellation of the ϵ poles numerically we can choose several randomly selected phase space points and evaluate the necessary integrals with high precision. The cancellation cannot depend on the particular phase space point. In the case of the finite remainders, in order to compute the phase space integrals in Eqns. (56) and (57), we are able to find sufficiently precise approximations to the integrated subtraction terms using a procedure that can be automated to high degree. The latter point is also important as there are several hundred integrals to compute. In the following, we outline our procedure for two cases: (i) an example with integrals depending on one kinematical parameter, ($\int_1 \mathcal{X}^{(\ell,k)} = \int_1 \mathcal{C}_{ir}^{(\ell,0)}$) and (ii) another example with integrals depending on two kinematical parameters, ($\int_2 \mathcal{X}^{(\ell,k)} = \int_2 \mathcal{C}_{ir,js}^{(0,0)}$).

In order to compute $\int_1 \mathcal{C}_{ir}^{(\ell,0)}$, we have to integrate the azimuthally averaged Altarelli-Parisi splitting functions $P_{f_i f_r}^{(\ell)}(z_{i,r}, z_{r,i}; \epsilon)$ in $4-2\epsilon$ dimensions for the splitting process $f_{ir} \to f_i + f_r$, with z_i being the momentum fraction of parton f_i. It was discussed in Ref. [249] that the corresponding functions $C_{ir}^{(\ell)}$ can be expressed as combinations of the integrals (we changed the notation from $\mathcal{I}$ to $\mathcal{I}_C$)

$$\mathcal{I}_C(x; \epsilon, \alpha_0, d_0, \kappa, k, \delta, g_I^{(\pm)}) = \frac{16\pi^2}{S_\epsilon} Q^{2\epsilon} \int_1 [\mathrm{d}p_{1,m+1}^{(ir)}] \frac{z_r^{k+\delta\epsilon}}{s_{ir}^{1+\kappa\epsilon}} g_I^{(\pm)}(z_r) f(\alpha_0, \alpha_{ir}, d(m,\epsilon)) \,. \tag{61}$$

In terms of explicit integration variables these collinear integrals have the general form [249]

$$\begin{aligned} \mathcal{I}_C(x; \epsilon, \alpha_0, d_0; \kappa, k, \delta, g_I^{(\pm)}) &= x \int_0^{\alpha_0} \mathrm{d}\alpha \, \alpha^{-1-(1+\kappa)\epsilon} \, (1-\alpha)^{2d_0-1} \, [\alpha + (1-\alpha)x]^{-1-(1+\kappa)\epsilon} \\ &\times \int_0^1 \mathrm{d}v [v\,(1-v)]^{-\epsilon} \left(\frac{\alpha + (1-\alpha)xv}{2\alpha + (1-\alpha)x} \right)^{k+\delta\epsilon} \, g_I^{(\pm)} \left(\frac{\alpha + (1-\alpha)xv}{2\alpha + (1-\alpha)x} \right) . \end{aligned} \tag{62}$$

The necessary functions $g_I^{(\pm)}$ are listed in Ref. [249], where analytic results of these integrals for $\alpha_0 = 1$ and $d_0 = 3$ are also presented.

Our present goal is to provide sufficiently precise numerical approximations to the functions $\mathcal{I}_C(x)$ in a simple way. The motivation is that often it is difficult to perform the analytic computation with arbitrary values of the parameters. For instance, the derivation with $\alpha_0 = 1$ is rather different from a derivation with $\alpha_0 < 1$. Also, the choice for d_0 is to some extent arbitrary, and a new choice requires a completely new analytic computation. Thus, for the sake of flexibility we propose a fully numerical approach here.

First we used the program `SecDec`, modified such that it can compute the value of the integral at multiple values of the parameter x in a single run. For simplicity, we call the O(1) terms of the integral 'measurements'. Then, inspired by the analytic results in Ref. [249], we fitted these measurements by combinations of logarithms and polynomials in x of the form

$$\mathcal{F}_C(x; \kappa = 0, k, \delta = 0, g_I^{(\pm)} = 1) = \sum_{n=0}^{n_{\max}} P_n^{(m)}(x,k) \log^n(x) \,, \qquad P_n^{(m)}(x,k) = \sum_{n=0}^{m} a_n^{(k)} x^n \tag{63}$$

where the upper limit $n_{\max}$ is determined by the power $-n_{\max}$ of the leading pole in the Laurent-expansion (in ϵ) of the integral. As for the degree of the polynomials we tried several simple choices ($m = 1, 2, 3$). We found that splitting the region of the parameter space into an asymptotic ($0 < x \leq 10^{-4}$) and a non-asymptotic ($10^{-4} < x \leq 1$) region, we could provide a fit with $m = 2$ that approximates the analytic result within relative difference few times 10^{-4}. The loss of relative precision is associated with phase space points where the function changes sign, and its numerical value is close to zero (around $x = 0.2$).

In Fig. 8 we show the approximate function $\mathcal{F}_C(x; \alpha_0, d_0, 0, -1, 0, 1)$ together with the 'measurements', which coincide with the known exact analytic result to at least six digit accuracy. We find very good agreement, which is characterized by the ratio of the two values in the lower panels. In Fig. 8b we show the approximate function for $\alpha_0 = 0.1$ and $d_0 = 3 - 3\epsilon$ together with the corresponding 'measurements'. In this case the analytic results are not available.

Building on the experience gained in studying the one-parameter case, we worked out a similar strategy for the integrated subtraction term $\int_2 \mathcal{C}_{ir,js}^{(0,0)}$. The corresponding functions $C_{ir,js}^{(0,0)}$ can be ex-

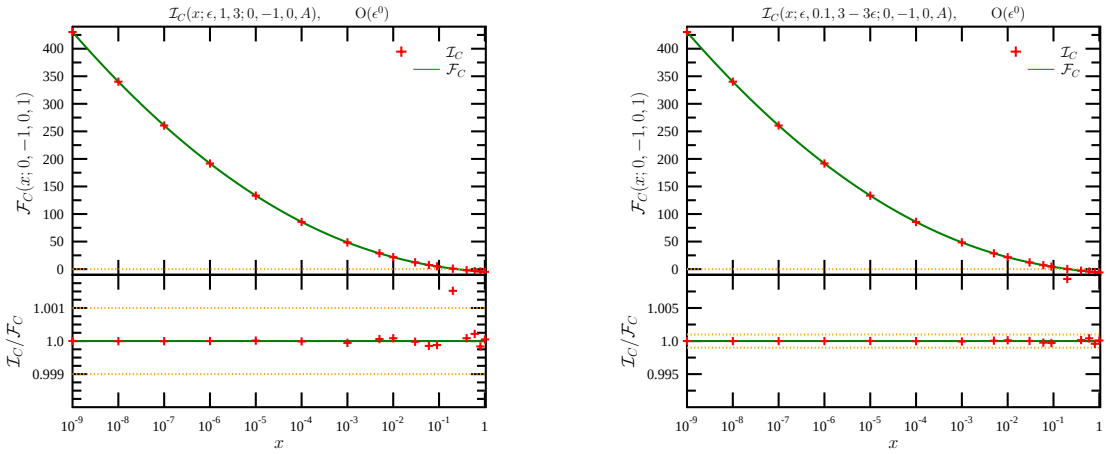


Fig. 8: The fitted function $\mathcal{F}_C(x;0,-1,0,1)$ compared to the integral $\mathcal{I}(x;0,-1,0,1)$ at a) $\alpha_0 = 1$ and $d_0 = 3$, b) $\alpha_0 = 0.1$ and $d_0 = 3 - 3\epsilon$.

pressed as combination of the integrals

$$\mathcal{I}_{2C}(x_i,x_j;\epsilon,\alpha_0,d_0;k,l) = x_i\,x_j \int_0^1 \mathrm{d}\alpha \int_0^1 \mathrm{d}\beta\,\Theta(\alpha_0-\alpha-\beta)$$
$$\times (1-\alpha-\beta)^{2d_0-2(1-\epsilon)}\alpha^{-1-\epsilon}\beta^{-1-\epsilon}\left(\alpha+(1-\alpha-\beta)x_i\right)^{-1-\epsilon}\left(\beta+(1-\alpha-\beta)x_j\right)^{-1-\epsilon}$$
$$\times \int_0^1 \mathrm{d}v\,v^{-\epsilon}(1-v)^{-\epsilon}\int_0^1 \mathrm{d}u\,u^{-\epsilon}(1-u)^{-\epsilon}\left(\frac{\alpha+(1-\alpha-\beta)x_i v}{2\alpha+(1-\alpha-\beta)x_i}\right)^k\left(\frac{\beta+(1-\alpha-\beta)x_j u}{2\beta+(1-\alpha-\beta)x_j}\right)^l . \tag{64}$$

We again run `SecDec` with $\alpha_0 = 0.1$ and $d_0 = 3-3\epsilon$ at several hundred different values of the kinematic parameters to obtain the 'measurements'. To reach 10^{-7} relative precision for all such 'measurements' takes several hours on a single CPU. Then we fitted these 'measurements' with the function

$$\mathcal{F}_{2C}(x_i,x_j;k,l) = \sum_{n_i=0}^{n_{\max}}\sum_{n_j=0}^{n_{\max}-n_i} P_{n_i}^{(m)}(x_i,k,l)P_{n_j}^{(m)}(x_j,k,l)\log^{n_i}(x_i)\log^{n_j}(x_j)\,. \tag{65}$$

We divide the parameter space $0 < x_i, x_j \leq 1$ into four regions: (i) $0 < x_i, x_j \leq 10^{-4}$, (ii) $0 < x_i \leq 10^{-2}$ and $10^{-4} < x_j \leq 1$, (iii) $0 < x_j \leq 10^{-2}$ and $10^{-4} < x_i \leq 1$, (iv) $10^{-2} < x_i, x_j \leq 1$. Using $m = 2$, we are able to fit the original function $\mathcal{I}_{2C}$ to per mille precision almost everywhere. The ratio of the fitted function $\mathcal{F}_{2C}$ to the numerical evaluation of $\mathcal{I}_{2C}$ is shown in Fig. 9 together whith the fitted function $\mathcal{F}_{2C}$ itself.

CONCLUSIONS

We have worked out a numerical procedure for providing simple approximations of the integrated subtraction terms of the NNLO subtraction scheme defined in Refs. [240, 241, 242, 243]. We use the publicly available program `SecDec` to compute the coefficients of the Laurent expansion of the necessary integrals to high numerical precision. We found that the integrals that depend on one or two kinematical invariants can be approximated with simple combinations of polynomials and logarithms. The precision of these approximations is usually at per mille or better.

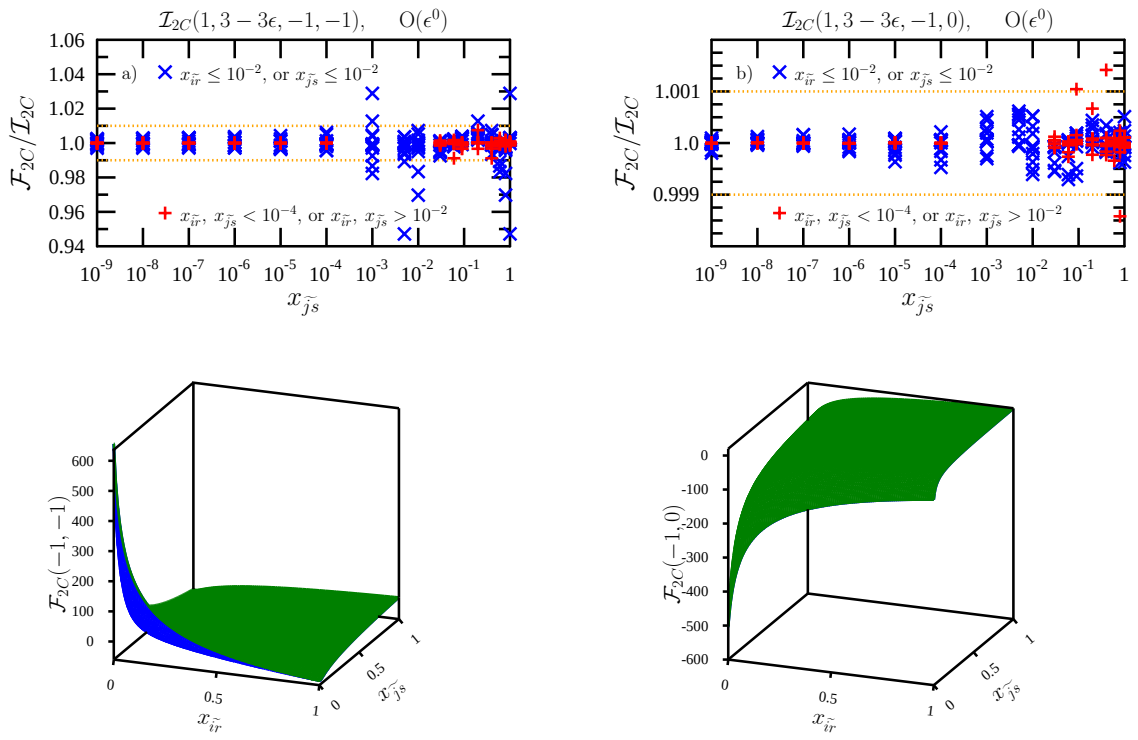


Fig. 9: The ratio of the fitted function $\mathcal{F}_{2C}(x_i, x_j; -1, l)$ to the integral $\mathcal{I}_{2C}(x_i, x_j; -1, l)$. a) $l = -1$ b) $l = 0$ Also shown the fitted function $\mathcal{F}_{2C}(x_i, x_j; -1, -l)$.

ACKNOWLEDGEMENTS

We are grateful to G. Heinrich for her help in modifying the `SecDec` package to our needs. This research was supported in part by the the LHCPhenoNet network PITN-GA-2010-264564.

Part III

PARTON DISTRIBUTION FUNCTIONS

7. WHICH EXPERIMENTS CONSTRAIN THE GLUON PDF IN A GLOBAL QCD FIT? [9]

Abstract

Based on computation of PDF-induced correlations, we identify the experiments in CTEQ and MSTW global QCD analyses that are sensitive to the gluon parton density in the proton. The Tevatron inclusive jet production at large momentum fractions x and DIS charm quark production at moderately small x show the strongest correlation with the gluon PDF. The strength of the PDF-induced correlation between the gluon PDF and inclusive (di)jet production data is different in the CTEQ and MSTW analyses.

7.1 Introduction

The parton distribution function (PDF) of gluons in a proton, $g(x, \mu)$, plays an important role in hadron collider phenomenology. It arises in cross sections for production of hadronic final states, massive scalar

[9] Contributed by: Z. Liang, P. M. Nadolsky

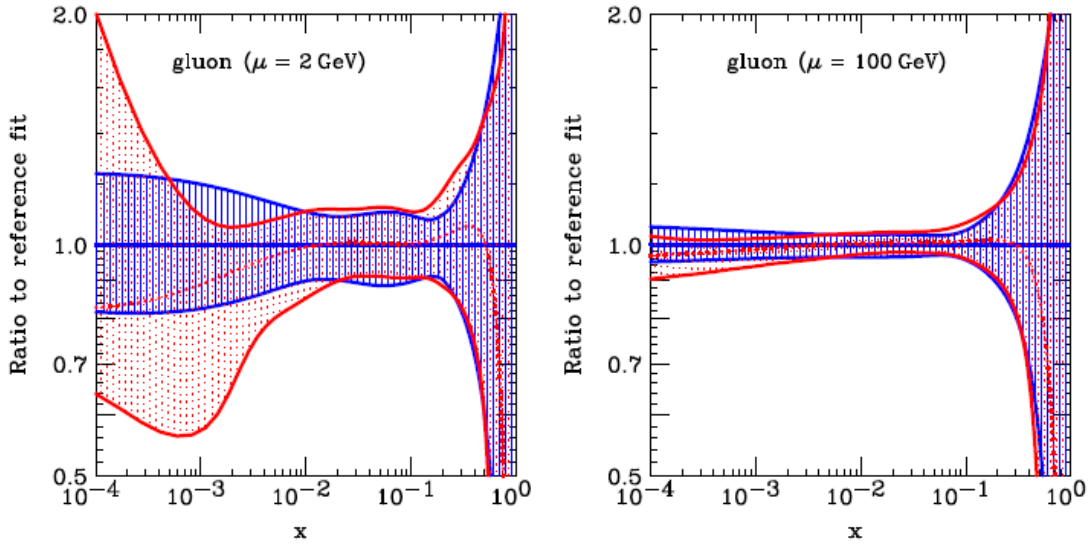


Fig. 10: CT10 and CTEQ 6.6 PDF uncertainty bands at $\mu = 2$ GeV (left) and 100 GeV (right), taken from Ref. [255]. The CTEQ 6.6 best-fit PDFs and uncertainties are indicated by solid curves and hatched bands, while those of CT10 are indicated by dashed curves and dotted bands.

bosons, and hypothetical elementary particles, often in a combination with an overall normalization prefactor proportional to α_s. The gluon distributions from CT10 [255] and CTEQ 6.6 [256] PDF sets are shown in Fig.10. The figure shows that the gluon PDF is constrained well by fitted experiments at the intermediate momentum fractions x, but the uncertainty grows in the region $x > 0.1$. We may ask which experiments in the global fit impose the most significant constraints on the the gluon PDF. It is often said that the precise neutral-current DIS data provides the tightest constraints on the gluon PDF at x of order 10^{-3}, while inclusive jet production at the Tevatron plays the key role in constraining the gluon at $x > 0.1$. However, the net PDF uncertainty reflects subtle interplay of numerous constraints imposed by QCD theory and multiple experiments, as well as various correlated uncertainties in experimental measurements. In this contribution, we identify the experiments with the strongest sensitivity to the gluon PDF by using a method of PDF-induced correlations that was developed in Refs. [256, 257, 258]. The analysis of correlations provides a systematic way to identify such experiments and also to establish specific ranges of x and Q where the correlations of the experimental data sets with the gluon PDF are the most pronounced.

7.2 Log-likelihood χ^2 and PDF-induced correlations

The quality of theory description of an experimental data set can be quantified by the log-likelihood function χ^2. Many high-energy physics experiments publish three kinds of measurement errors for each data point i: the statistical error σ_i, uncorrelated systematic error u_i, and correlated systematic errors $\{\beta_{1i}, \beta_{2i}, \beta_{3i}....\beta_{Ki}\}$ of K different types. To compare a theory prediction T_i to the data value D_i for a data point i, while accounting for all types of errors, the χ^2 function can be constructed as [259, 260]

$$\chi^2 = \sum_{expt.} \left[\sum_{i=1}^{N_e} \left(\frac{D_i - T_i(a) - \sum_{k=1}^{K} r_k \beta_{ki}}{\alpha_i^2} \right)^2 + \sum_{k=1}^{K} r_k^2 \right], \tag{66}$$

where $\alpha_i^2 = \sigma_i^2 + u_i^2$ is the combined uncorrelated error; r_k are random parameters describing each of K correlated errors (each distributed according to the standard normal distribution); N_e is the number of

the data points; and K is the number of the sources of the correlated systematic errors.

Analytic minimization of the function (66) with respect to the correlated systematic parameters r_k renders the following result [257, 259]:

$$r_k|_{\text{best fit}} = \sum_{k'=1}^{K} A^{-1}_{kk'} B_{k'}, \tag{67}$$

where $A_{kk'}$ and B_k are given by

$$A_{kk'} = \delta_{kk'} + \sum_{i=1}^{N_e} \frac{\beta_{ki}\beta_{k'i}}{\alpha_i^2}, \quad \text{and} \quad B_k = \sum_{i=1}^{N_e} \frac{\beta_{ki}(D_i - T_i)}{\alpha_i^2}. \tag{68}$$

Substituting Eq. (67) into Eq. (66), we obtain a reduced χ^2 function [257, 259],

$$\chi^2 = \sum_{expt.} \left[\sum_{i=1}^{N_e} \frac{(D_i - T_i)^2}{\alpha_i^2} - \sum_{k,k'=1}^{K} B_k A^{-1}_{kk'} B_{k'} \right]. \tag{69}$$

In this function, the information about the systematic shifts in r_k is included implicitly. Often, the influence of the correlated shifts on the PDFs is substantial.

Next, we wish to discuss correlations between PDF uncertainties of two variables, $X(\vec{a})$ and $Y(\vec{a})$, where $\vec{a} = \{a_1, a_2, ..., a_N\}$ is the vector of N PDF parameters. The correlations can be computed either in the Hessian [256, 257, 258] or Monte-Carlo [261] approaches. In this note we will adopt the Hessian approach.

A symmetric PDF uncertainty ΔX corresponds to the maximal variation of X for all combinations of PDF parameters that lie within the tolerance hypersphere $\Delta\chi^2 \leq T^2$. This uncertainty is given by

$$\Delta X = \frac{1}{2}\sqrt{\sum_{i=1}^{N} [X_i^+ - X_i^-]^2} \tag{70}$$

in terms of the value X_0 of X obtained with the central PDF set, and values X_i^+ and X_i^- of X obtained for maximal positive and negative displacements of each orthonormal PDF parameter a_i within the tolerance hypersphere. The same "master equation" defines ΔY, the PDF uncertainty of the variable Y.

In the linear approximation, the pairs of values of X and Y that are allowed within the PDF uncertainty correspond to the points inside an ellipse in the X-Y plane. The boundary of the ellipse is parametrically described by

$$X = X_0 + \Delta X \cos\theta, \tag{71}$$

$$Y = Y_0 + \Delta Y \cos(\theta + \varphi), \tag{72}$$

where the parameter θ varies between 0 and 2π, and the relative phase angle φ is a function of $X_i^\pm$ and $Y_i^\pm$. The PDF uncertainties ΔX and ΔY are calculated according to Eq. (70). The angle φ is included between the gradients $\vec{\nabla} X$ and $\vec{\nabla} Y$ of X and Y in the PDF parameter space. Its cosine,

$$\cos\varphi = \frac{\vec{\nabla} X \cdot \vec{\nabla} Y}{\Delta X \Delta Y} = \frac{1}{4\Delta X \Delta Y} \sum_{i=1}^{N} \left(X_i^{(+)} - X_i^{(-)} \right) \left(Y_i^{(+)} - Y_i^{(-)} \right), \tag{73}$$

quantifies the degree of similarity in the PDF dependence of X and Y. If X and Y are strongly correlated (corresponding to $\cos\varphi \to 1$) or anti-correlated ($\cos\varphi \to -1$), the PDF uncertainties of X and Y are driven by essentially the same combinations of PDF parameters. Conversely, the PDF dependence of X is independent from the PDF dependence of Y if $\cos\varphi \approx 0$.

7.3 Which experiments are sensitive to the gluon PDF?

If an experimental cross section σ strongly constrains a PDF $f_a(x,Q)$ for some combination of x and Q, we expect that Eq. (73) returns $|\cos\varphi|$ close to unity when using $X = f_{a/A}(x,Q)$ and $Y = \sigma$. If the experimental data set includes several data points, we can use $Y = \chi^2$. The strength of the constraint on the PDF from this experiment is determined by $|\cos\varphi|$ and the magnitude of χ^2. In the majority of the fitted experiments, χ^2/N_e is close to 1, so that $|\cos\varphi|$ tends to be more important for distinguishing between the sensitivities of the experiments than the magnitude of χ^2.

Following this approach, we compute $\cos\varphi$ between the NLO gluon PDF $g(x,Q)$ in various x ranges (for $Q^2 = 10$ GeV^2), and χ^2 for typical experimental data sets that are used in the PDF analysis. In this study, we compute $\cos\varphi$ for the experiments from the CT10 analysis that are listed in Table 4. In the figures, we refer to each experiment by its numerical ID that is shown in the left column of Table 4.

The $\cos\varphi$ values between the gluon PDF at a given x value and χ^2 for each experiment are plotted as two-dimensional contour plots for CT10 NLO PDFs [255] in the left panel of Fig. 11, and for MSTW'08 NLO PDFs [262] in the right panel. The horizontal axis indicates the range of x in $g(x,Q)$. The vertical axis indicates the ID of the experiment. At the bottom of the figure, we show the color legend adopted to draw the contour plots. The color legend is chosen so as to emphasize only cells with large correlation ($\cos\varphi > 0.5$, dark yellow-red colors) or large anticorrelation ($\cos\varphi < -0.5$, blue colors). The regions with $|\cos\varphi| < 0.5$ are filled with a light-yellow color. The χ^2 values for each data set are computed according to Eqs. (66) and (69) using the CTEQ fitting code for both CT10 and MSTW PDF sets.

Visual inspection of two panels of Fig. 11 reveals both similarities and differences in the pattern of correlations of the gluon PDF in the CT10 and MSTW PDF sets. In the case of the CT10 PDF (left panel), the gluon PDF has a pronounced anti-correlation (blue spots) with HERA charm and bottom SIDIS production data sets (experiments 140, 143, 145, 156, 157) at $x < 0.1$, as well as with Tevatron inclusive jet production data sets (experiments 504, 505, 514, and 515) at $x > 0.05$. Some correlations (brown and red spots) are also observed, but they are not as pronounced as the anti-correlations. Weaker (anti-)correlations can be noticed with the NMC F_2^p, CDHSW F_2^p, and E605 pp Drell-Yan process data, corresponding to experiments 103, 108, and 201.

While the gluon PDF of the MSTW'08 set (right panel of Fig. 11) also shows an (anti-)correlation with the heavy-quark DIS and jet production data, the overall pattern of the correlations is somewhat different from the CT10 case. Here, the gluon PDF is mostly correlated with high-x jet production (experiments 504, 505, 514, and 515), while it is either correlated or anti-correlated with heavy-quark DIS experiments (experiments 140, 143, 145, 156, 157). In addition, we observe significant (anti-)correlations with the combined HERA DIS data set (ID=159) and fixed-target DIS experiments (ID=101-124) that are not seen in the CT10 panel.

We now turn to the correlations of the gluon and u-quark PDFs with χ^2 values in individual bins of Tevatron inclusive jet and dijet production data. For this purpose, we represent χ^2 for one experimental data set in Eq. (69) as a sum of contributions χ_i^2 from individual data points i:

$$\chi^2 = \sum_{i=1}^{N_e} \chi_i^2, \tag{74}$$

where

$$\chi_i^2 = \frac{D_i - T_i}{\alpha_i} \sum_{j=1}^{N_e} \left\{ \delta_{ij} - \frac{D_j - T_j}{\alpha_j} \sum_{k,k'=1}^{K} \frac{\beta_{ki}}{\alpha_i} A_{kk'}^{-1} \frac{\beta_{k'j}}{\alpha_j} \right\}. \tag{75}$$

Each contribution χ_i^2 accounts for the effect of correlated systematic shifts through the term that includes $A_{kk'}^{-1}$ on the right-hand side of Eq. (75). Again, the constraining power of each point is determined both

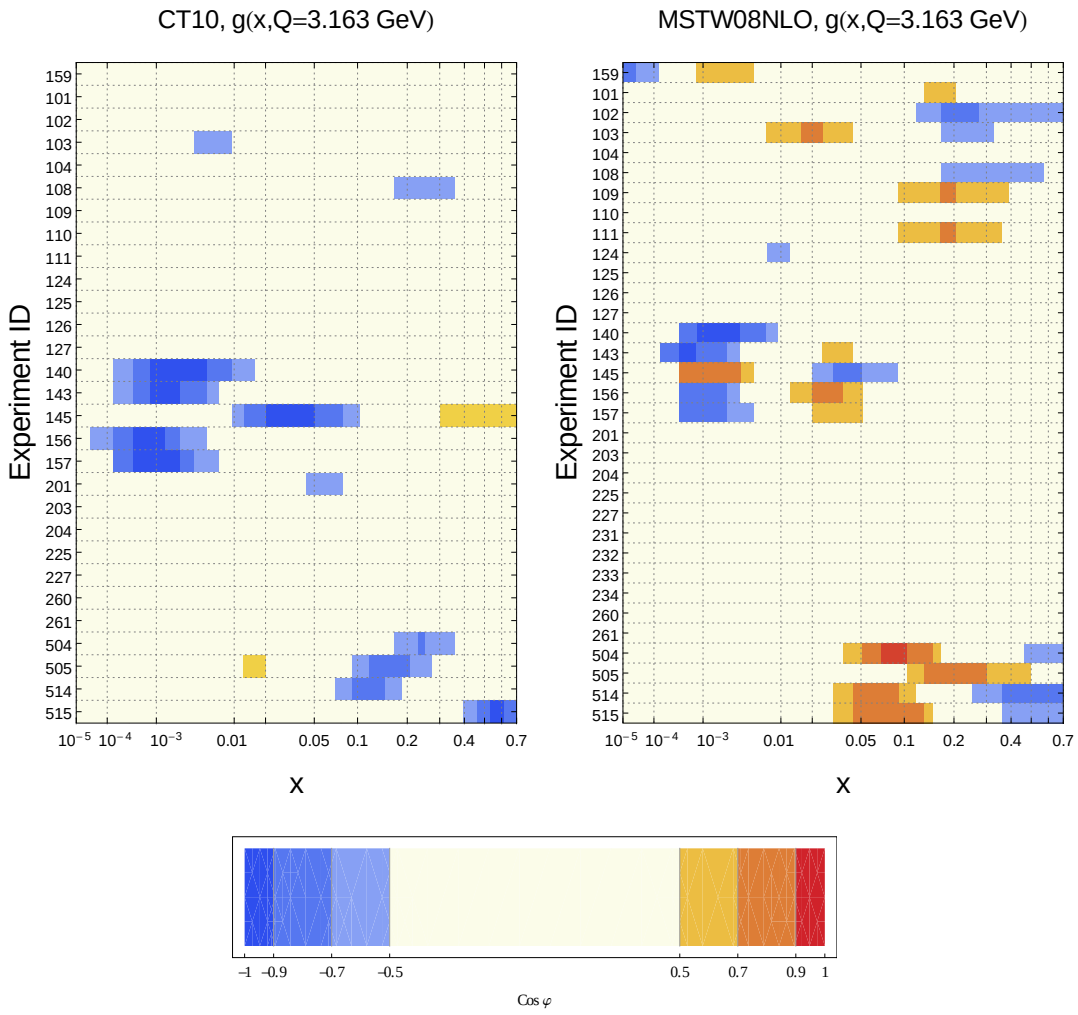


Fig. 11: Correlation between the gluon distribution from CT10 NLO (left) and MSTW2008 NLO (right) PDF sets and χ^2 for the experiments used in the CT10 global QCD analysis. The color of each cell indicates the value of $\cos\varphi$ according to the included legend. The ID's of individual experiments are listed in Table 4.

ID	Experimental data set
159	Combined HERA1 NC and CC DIS [263]
101	BCDMS F_2^p [264]
102	BCDMS F_2^d [265]
103	NMC F_2^p [266]
104	NMC F_2^d/F_2^p [266]
108	CDHSW F_2^p [267]
109	CDHSW F_3^p [267]
110	CCFR F_2^p [268]
111	CCFR xF_3^p [269]
124	NuTeV neutrino dimuon SIDIS [270]
125	NuTeV antineutrino dimuon SIDIS [270]
126	CCFR neutrino dimuon SIDIS [271]
127	CCFR antineutrino dimuon SIDIS [271]
140	H1 F_2^c [272]
143	H1 σ_r^c for $c\bar{c}$ [273, 274]
145	H1 σ_r^b for $b\bar{b}$ [273, 274]
156	ZEUS F_2^c [275]
157	ZEUS F_2^c [276]
201	E605 Drell-Yan process, $\sigma(pA)$ [277]
203	E866 Drell Yan process, $\sigma(pd)/(2\sigma(pp))$ [278]
204	E866 Drell-Yan process, $\sigma(pp)$ [279]
225	CDF Run-1 W charge asymmetry [280]
227	CDF Run-2 W charge asymmetry [281]
231-234	DØ Run-2 W charge asymmetry [282]
260	DØ Run-2 Z rapidity distribution [283]
261	CDF Run-2 Z rapidity distribution [284]
504	CDF Run-2 inclusive jet production [285]
505	CDF Run-1 inclusive central jet production [286]
514	DØ Run-2 inclusive jet production [287]
515	DØ Run-1 inclusive jet production [288]

Table 4: Experimental data sets examined in this analysis.

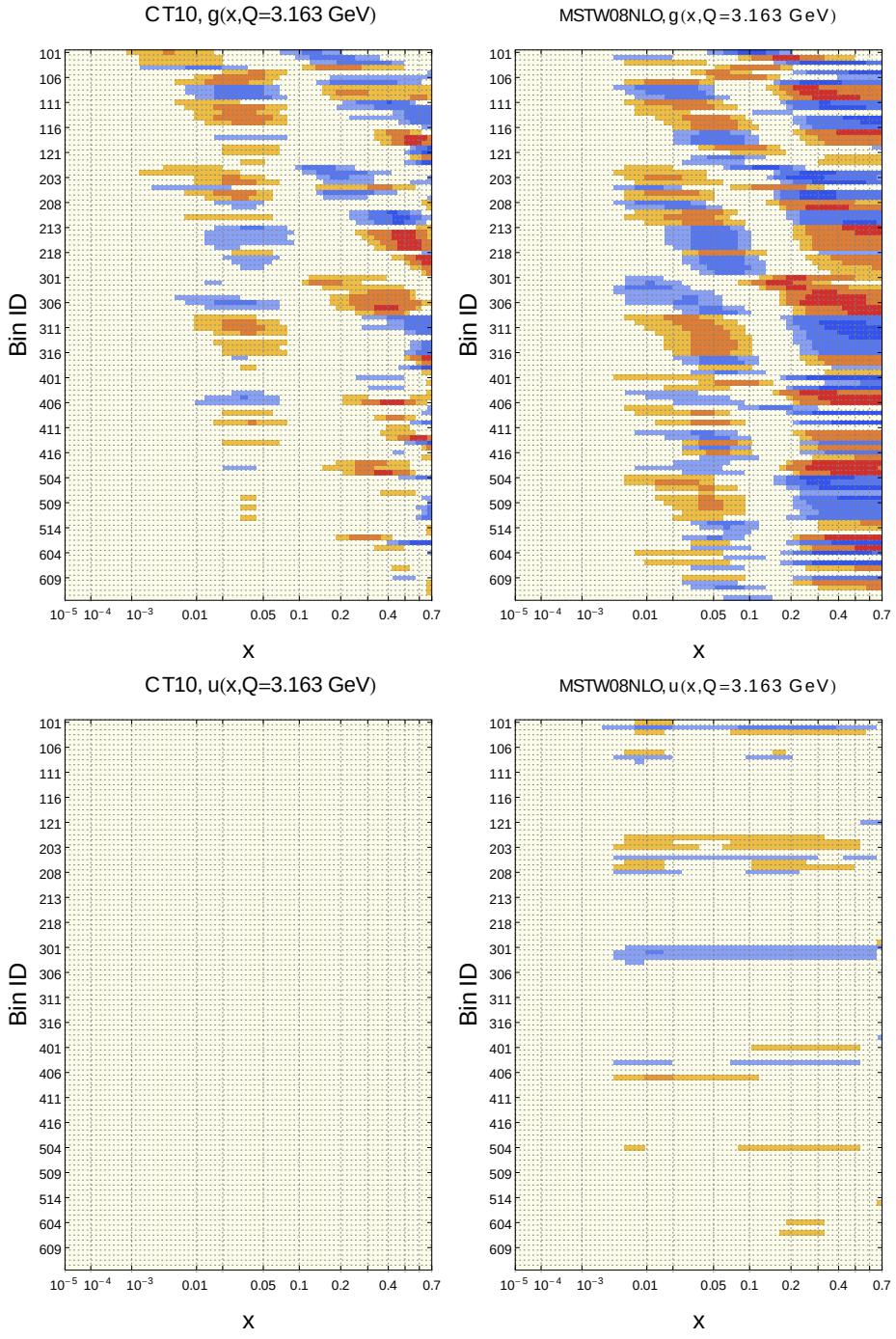


Fig. 12: Correlation cosine between χ^2_i in each p_T bin from DØ Run-2 inclusive jet production and gluon and u quark distributions from CT10 and MSTW 2008 NLO sets. The horizontal axis refers to the x value in the PDF. The vertical axis indicates the numerical ID of the experimental bin for which χ^2 is computed. The ID for each bin is indicated as $100\, i_y + i_{p_T}$, where $i_y = 1, ...6$ and i_{p_T} are the ID's of the corresponding rapidity interval and the p_T interval, respectively.

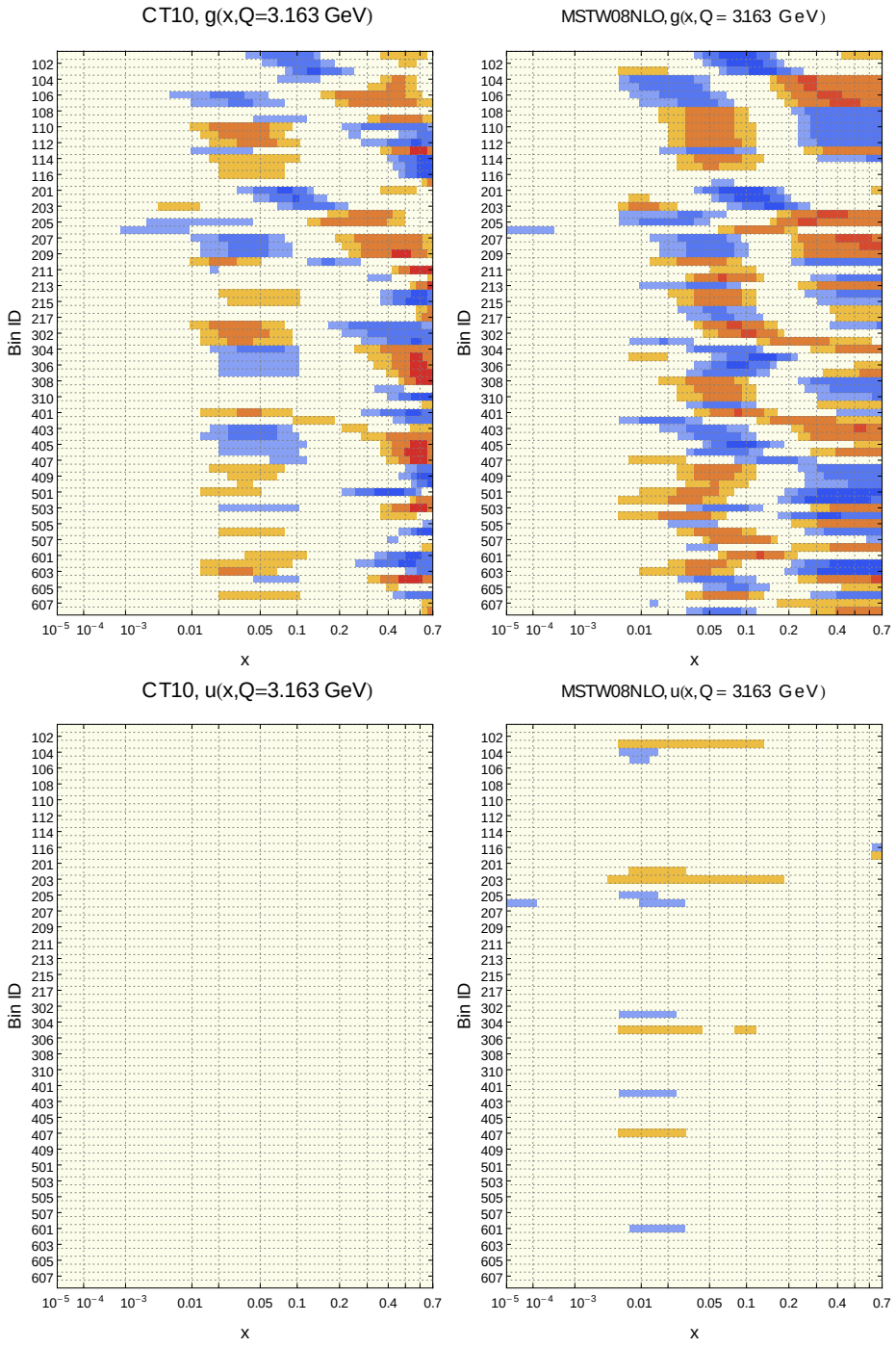


Fig. 13: Correlation cosine between χ^2_i in each m_{jj} bin from DØ Run-2 dijet production and gluon and u quark distributions from CT10 and MSTW 2008 NLO sets. The horizontal axis refers to the x value in the PDF. The vertical axis indicates the numerical ID of the experimental bin for which χ^2 is computed. The ID for each bin is indicated as $100\, i_{y_{max}} + i_{m_{jj}}$, where $i_{y_{max}} = 1, .., 6$ and $i_{m_{jj}}$ are the ID's of the corresponding intervals in y_{max} and m_{jj}, respectively.

by the value of $|\cos\varphi|$ and the magnitude of χ^2_i, with the latter being comparable to unity for the majority of the data points.

For DØ Run-2 single-inclusive jet cross sections [287], we plot the $\cos\varphi$ values for the gluon and u-quark PDFs, with χ^2_i computed for each bin of the jet's transverse momentum p_T and rapidity y. The resulting contour plots are shown in Fig. 12. Similarly, for DØ Run-2 dijet cross sections [289], Fig. 13 shows the contour plots of $\cos\varphi$ for χ^2_i in the bins of of dijet invariant mass m_{jj} and maximal absolute rapidity $|y| = \max(|y_1|, |y_2|)$ of the dijets. In both figures, theory cross sections are computed at NLO (without threshold resummation corrections) with the FASTNLO code [290, 291], using the settings described in Section 13.. The same color legend as in Fig. 11 is used. Similar patterns of correlations were found with the CDF Run-2 inclusive jet data (not shown).

The upper panels in both figures show $\cos\varphi$ for CT10 NLO and MSTW'08 NLO gluon PDFs. The correlated experimental errors modify the correlations by smearing the $\cos\varphi$ distribution. The pattern of $\cos\varphi$ indicates clearly that the (di)jet data are very sensitive to the gluon at x above 0.01. However, the correlation is weaker for the CT10 gluon PDF (left panel) then for MSTW'08 PDF (right panel), suggesting that the importance of the constraints on the gluon PDF from the jet data is not the same in two fits. In addition, the MSTW'08 u-quark PDF shows mild (anti-)correlation with both single-inclusive jet data and dijet data, as can be observed in the right lower panels in Figs. 12 and 13. No pronounced (anti-)correlations with the u-quark PDF or other quark PDFs of physical flavors are observed for the CT10 set, shown in the lower left panels.

The contour plots confirm the expectation that the inclusive jet data play an important role in constraining the gluon PDF. While the constraints are strongest at $x > 0.1$, they extend down to x as low as 0.05 for both CT10 and MSTW sets, as can be observed in Figs. 12 and 13. The gluon PDF is sensitive to constraints from heavy-quark semi-inclusive DIS production at even lower x values, cf. Fig. 11. As the HERA data on heavy-quark DIS production continue to improve, it will play an increasingly important role in constraining the low-x gluon density.

While the patterns of PDF-induced correlations are visually similar for the CT10 and MSTW'08 sets, they are not completely identical. Constraints on the gluon PDF from Tevatron jet production may not be as strong in the CT10 fit as in the MSTW'08 fit, according to Figs. 12 and 13. It remains to be investigated what causes the observed differences between CT10 and MSTW sets in the correlations involving the gluon PDF. Several features are different in these fits, including different heavy-quark DIS schemes, choice of experimental data sets, PDF parametrizations, and radiative contributions in theoretical cross sections. A combination of these effects may indirectly affect the strength of the constraints imposed on the gluon density by the collider jet data.

ACKNOWLEDGMENTS

This work was supported by the U.S. DOE Early Career Research Reward DE-SC0003870 and by Lightner-Sams Foundation. We thank M. Guzzi, J. Gao, J. Huston, J. Pumplin, D. Stump, and C.-P. Yuan for related discussions.

8. PDF CONSTRAINTS FROM ELECTROWEAK VECTOR BOSON PRODUCTION AT THE LHC [10]

Abstract

We present a study of the impact of the recent W and Z measurements from ATLAS, CMS and LHCb on parton distribution functions. We show that the NNPDF2.1 NNLO predictions are consistent with all the new data, but that these provide significant further constraints on the light quarks and antiquarks

[10] Contributed by: R. D. Ball, N. P. Hartland, J. Rojo and M. Ubiali

at medium and small-x. We conclude that these data already have the potential to play a useful role in future global PDF analyses.

8.1 LHC measurements sensitive to PDFs

The LHC has already provided an impressive set of measurements which are sensitive to parton distributions: inclusive jet and dijet data [292, 293, 294], electroweak vector boson production [295, 296, 297, 298, 299, 300] (both inclusive and in association with heavy quarks [301]) and direct photon production [302, 303]. The purpose of this contribution is to quantify the impact on PDFs of a subset of these data, the W and Z inclusive production measurements. In this first section we will review the status of LHC data relevant for PDF determination and then in the next section we will study how the W, Z data impact on the NNPDF analysis.

Let's begin this short review of LHC data with electroweak vector boson production. ATLAS has measured the W lepton and Z rapidity distributions using the 2010 data (36 pb^{-1}) and determined the full covariance matrix of correlated experimental uncertainties [295]. This measurement supersedes the original muon asymmetry measurement from W decays [296], for which the covariance matrix was not available. The CMS collaboration has presented a preliminary measurement of the muon asymmetry with 2011 data (234 pb^{-1}) [297] which supersedes the 2010 data [298]. In addition it has presented a measurement of the normalized Z rapidity distribution using 2010 data [299]. In neither of these two measurements has the full covariance matrix been made available. Finally, the LHCb Collaboration has presented preliminary results for the Z rapidity distribution, W lepton asymmetry and W lepton charge ratio using 2010 data [300].

Data Set	Ref.	$N_{\rm dat}$	$[\eta_{\rm min},\eta_{\rm max}]$	$\langle\sigma_{\rm stat}\rangle$ (%)	$\langle\sigma_{\rm sys}\rangle$ (%)	$\langle\sigma_{\rm norm}\rangle$ (%)
ATLAS W,Z 36 pb^{-1}	[295]	30	$[0,3.2]$	1.9	1.7	3.4
ATLAS W^+ 36 pb^{-1}	[295]	11	$[0,2.4]$	1.4	1.3	3.4
ATLAS W^- 36 pb^{-1}	[295]	11	$[0,2.4]$	1.6	1.4	3.4
ATLAS Z 36 pb^{-1}	[295]	8	$[0,3.2]$	2.8	2.4	3.4
CMS Z rapidity 36 pb^{-1}	[299]	35	$[0,3.6]$	12.3	-	0
CMS muon asymmetry 234 pb^{-1}	[297]	11	$[0,2.4]$	1.7	3.1	0
LHCb Z rapidity 36 pb^{-1}	[300]	5	$[2,4.5]$	20	5	3.4
LHCb W lepton asymmetry 36 pb^{-1}	[300]	5	$[2,4.5]$	16	21	0

Table 5: The number of data points, kinematical coverage and average statistical, systematic and normalization percentage uncertainties for each of the experimental LHC W and Z datasets considered in the present analysis. For the CMS Z rapidity data, the systematic uncertainty is included in the statistical uncertainty: there is no normalization uncertainty because these data are normalised to the total cross-section.

The kinematical coverage of each of the various LHC W and Z dataset with the corresponding average experimental uncertainties for each dataset are summarized in Table 5. As we can see the LHC electroweak data span a large range in rapidity up to η =4.5. Each of the three processes considered, W^+, W^- and Z is sensitive to different partonic subprocesses.

There are other LHC datasets potentially sensitive to PDFs. Jet production from the Tevatron has been a very important measurement not only to constrain the gluon at high x, but in determining the strong coupling from a global PDF analysis [304, 305]. Similar constraints are expected from the LHC jet data, extended into a wider kinematical range. From the 2010 (36 pb^{-1}) dataset inclusive jet and dijet production has been measured by both CMS [292, 293] and ATLAS [294], however only for ATLAS is the full experimental covariance matrix available. The LHC inclusive jet data can be treated within a global analysis framework using tools like FastNLO or APPLgrid [306]. Since the full NNLO corrections to the inclusive jet production are unknown, jet data in a NNLO analysis can be included only within some approximation: for example with NNLO PDF evolution and coupling running but with NLO matrix elements, or else with NLO matrix elements supplemented with Sudakov estimates of the

NNLO corrections. Another LHC measurement that has the potential to constrain the gluon PDFs is prompt photon production from ATLAS [302] and CMS [303]: its consistency with NLO QCD and their impact on the NNPDF2.1 PDFs will be discussed in detail in Ref. [307].

8.2 PDF constraints from LHC W and Z

Until recently all available NNPDF sets [308, 309, 261, 310, 311, 312, 313] were based on non-LHC data. NNPDF2.2 [314] was the first set to include LHC data, the W lepton asymmetry from ATLAS and CMS [298, 296]. However now these two datasets are outdated, the first because now the full correlation matrix of the W and Z lepton distributions is available, and the second because data from higher luminosities is also available. So we have chosen to continue to use as our baseline the NNPDF2.1 NNLO set.

We now study the impact of the latest LHC W and Z data on the NNPDF parton distributions. All our theoretical NNLO predictions will be computed with DYNNLO [315] with the same cuts and settings as in the respective measurements. The impact of the new data will be quantified using the reweighting method of Refs. [316, 314] applied to the $N_{\rm rep} = 1000$ replicas of the NNPDF2.1 NNLO set.

To begin with, we have computed the χ^2 for each of the datasets in Table 5 for the most recent NNLO PDF sets currently available on LHAPDF: NNPDF2.1, MSTW08 [262], ABKM09 [317], HERAPDF1.5 [318] and JR09 [319]. When available, we use the full experimental covariance matrix. Normalization uncertainties are included using the t_0 method [320]. This is important specially for the treatment of the ATLAS differential distributions where normalization uncertainties are comparable to the statistical and systematic uncertainties (See Table 5).

The results are summarized in Table 6. For the ATLAS W and Z lepton distributions we show the results both for the total dataset and the individual subsets, where in the latter case cross-correlations between subsets have been neglected. In all cases the theoretical NNLO predictions have been obtained with DYNNLO as discussed above. We can see that none of these PDF sets describes the ATLAS and CMS data perfectly, although NNPDF2.1 and HERAPDF1.5 give probably the best description, while ABKM09 and JR09 are significantly worse. All five sets give a reasonable description of the LHCb data within their large uncertainties.

Dataset	χ^2 NNPDF2.1	χ^2 MSTW08	χ^2 ABKM09	χ^2 JR09	χ^2 HERAPDF1.5
ATLAS	2.7	3.6	3.6	5.0	2.0
ATLAS W^+ 36 pb^{-1}	5.7	6.5	11.4	5.4	5.3
ATLAS W^- 36 pb^{-1}	2.5	4.1	5.4	8.0	6.4
ATLAS Z 36 pb^{-1}	1.8	3.7	4.2	6.5	2.9
CMS	2.0	3.0	2.8	3.6	2.8
CMS Z rapidity 36 pb^{-1}	1.9	2.9	2.7	2.0	3.0
CMS muon asymmetry 234 pb^{-1}	2.0	3.4	3.0	8.7	2.1
LHCb	0.8	0.7	1.2	0.4	0.6
LHCb Z rapidity 36 pb^{-1}	1.1	0.7	0.8	0.6	0.8
LHCb W lepton asymmetry 36 pb^{-1}	0.5	0.6	1.6	0.2	0.5

Table 6: Comparison between LHC W and Z data and the most recent NNLO PDFs. For each PDF set we provide the χ^2/dof between data and theory predictions, computed using the t_0-method.

For the ATLAS data, we would like to emphasize the importance of properly taking into account the correlations between datasets, specially the normalization: the description of the individual W^+, W^- and Z datasets is always worse than the overall description because of these cross-correlations. For the CMS Z rapidity distribution we find that the fixed order NNLO description seems rather worse than the NLO+LL prediction implemented in POWHEG [299]: the origin of this difference should be investigated in future studies.

We now discuss the impact of these LHC EW data into the NNPDF2.1 NNLO PDFs [313]. In Table 7 we summarize the initial χ^2 for each dataset, the χ^2 after reweighting, $\chi^2_{\rm rw}$. We find excellent agreement with all the LHC electroweak measurements after reweighting. Some comparisons between data and theory for a selected observables are shown in Fig. 14. From top to bottom we show the comparison with ATLAS, CMS and LHCb data. In each case we have included all the most updated electroweak datasets from each collaboration.

In Table 7 we also show the effective number of replicas left after the reweighting, defined as in Ref. [316] using the Shannon entropy,

$$N_{\rm eff} \equiv \exp\{\tfrac{1}{N_{\rm rep}} \sum_{k=1}^{N_{\rm rep}} w_k \ln(N_{\rm rep}/w_k)\} \, . \tag{76}$$

In each case we have performed the reweighting separately for each of the experimental datasets individually, for the combined datasets from each experiment, and finally with all three combined together.

Dataset	χ^2	$\chi^2_{\rm rw}$	$N_{\rm eff}$
ATLAS	2.7	1.2	16
ATLAS W^+ 36 pb^{-1}	5.7	1.5	17
ATLAS W^- 36 pb^{-1}	2.5	1.0	205
ATLAS Z 36 pb^{-1}	1.8	1.1	581
CMS	2.0	1.2	56
CMS Z rapidity 36 pb^{-1}	1.9	1.4	223
CMS muon asymmetry 234 pb^{-1}	2.0	0.4	200
LHCb	0.8	0.8	972
LHCb Z rapidity 36 pb^{-1}	1.1	1.0	962
LHCb W lepton asymmetry 36 pb^{-1}	0.8	0.5	961
All data combined	2.1	1.2	4

Table 7: The impact of LHC electroweak measurements on the NNPDF2.1 NNLO PDFs. For each dataset we show the initial χ^2, the χ^2 after reweighting these particular dataset and the effective number of replicas $N_{\rm eff}$ in this case. We show both the results for individual datasets as well as for the combined impact of all datasets within the same experiment. All the results have been computed starting with $N_{\rm rep} = 1000$ replicas.

When all the datasets are taken together, the initial $\chi^2 = 2.1$, already quite reasonable is reduced down to $\chi^2_{\rm rw} = 1.2$, thus obtaining a very good overall description of all the most recent LHC electroweak data. The effective number of replicas for all combined datasets is only $N_{\rm eff} = 4$ however: from this we conclude that to determine the combined impact of these data on PDFs would require many more replicas (around 25,000 in fact, to obtain reasonable statistical accuracy), or, more practically, a new fit. Note that the fact that the total effective number of replicas for the whole dataset is rather smaller than that of any individual subset confirms their mutual compatibility and the lack of any appreciable tension. Comparing the effective number of replicas for the individual datasets, the most constraining data are the ATLAS W and Z distributions, specially the very precise W^+ data. On the other hand the LHCb data have a rather small impact.

Let us now examine how various PDFs change when new experiments are added. In particular we show in Fig. 15 the NNPDF2.1 NNLO $d(x,Q^2)$ and $\bar{u}(x,Q^2)$ PDFs at $Q^2 = M_W^2$ as ratios to the central value before including the new data. As described above, we put together all the data from a given experiment. As can be seen, the ATLAS data give a moderate reduction in PDF uncertainties, and a somewhat softer small-x sea quarks, although the old and new PDFs agree at the 1–sigma level. For CMS the central values for the old and new PDFs are unchanged with a moderate error reduction at medium-x. Finally, for LHCb the PDF uncertainties are almost unaffected, due to the low constraining power of these datasets.

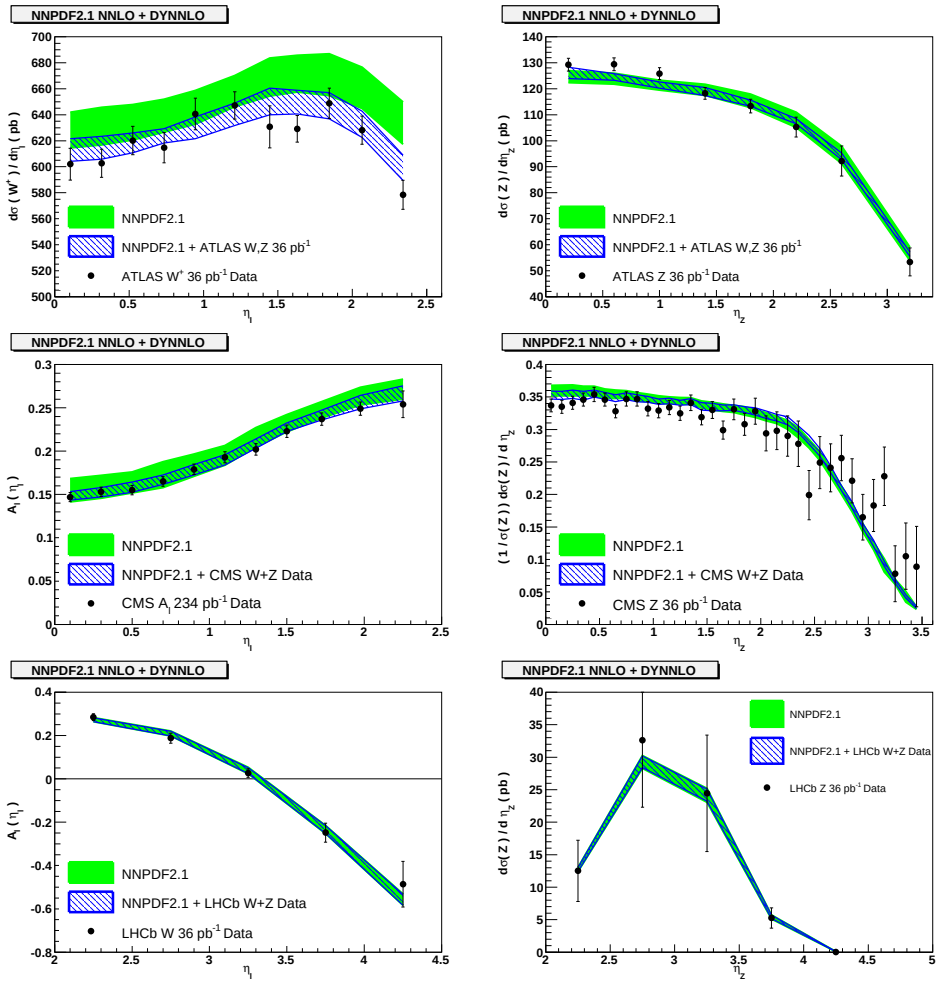


Fig. 14: Comparison between data and theory before and after reweighting for NNPDF2.1 NNLO compared to the various LHC EW datasets considered. From top to bottom we show comparisons with ATLAS (W^+ lepton Z rapidity distributions), CMS (W lepton asymmetry and Z rapidity distribution) and LHCb data (same as CMS). For the ATLAS data the error bars include statistical and systematic uncertaintes, but not the normalization uncertainties.

8.3 Conclusions

In this contribution we have quantified the impact of the most updated LHC electroweak data on the NNPDF2.1 NNLO parton distributions. NNPDF2.1 provides a reasonable description of all these datasets even before their impact on the PDFs is included. We find that all the datasets are mutually consistent, with no obvious tensions. The PDF uncertainties for the light quarks and antiquarks at medium and small-x are moderately reduced. The ATLAS W, Z data seem to prefer a softer small-x sea. It is clear from our results that the LHC W and Z data should play an important part in any future PDF global fit.

Acknowledgments

The research of J. R. has been supported by a Marie Curie Intra–European Fellowship of the European Community 7th Framework Programme under contract number PIEF-GA-2010-272515. M.U. is supported by the Bundesministerium für Bildung and Forschung (BmBF) of the Federal Republic of Germany (project code 05H09PAE).

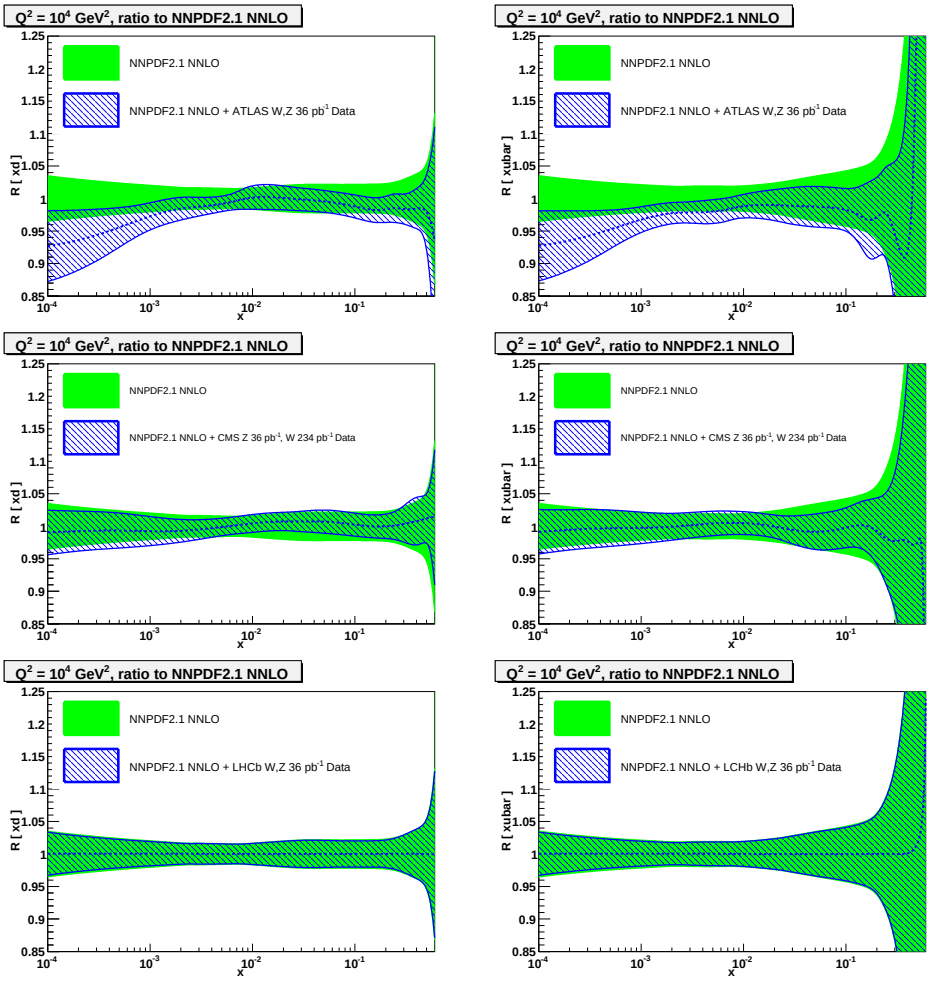


Fig. 15: Comparison between the NNPDF2.1 NNLO parton distributions before and after reweighting with the various LHC EW datasets considered. From top to bottom: ATLAS, CMS and LHCb data, where the left column we show the ratio of the d quark PDF and in the right column the ratio of the $\bar{u}$ quark PDF to their central values before reweighting.

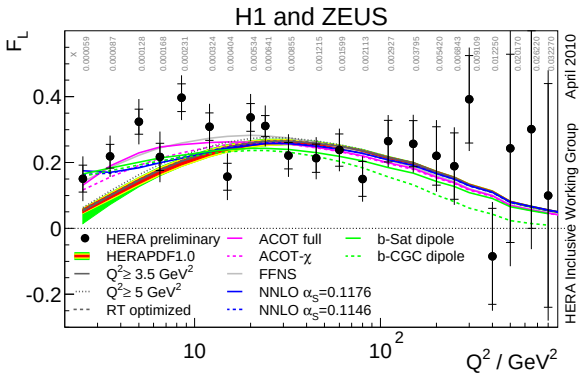


Fig. 16: F_L vs. Q^2 for the HERA combined measurements from H1 and ZEUS [321].

9. HEAVY QUARK PRODUCTION IN THE ACOT SCHEME AT NNLO AND N^{3}LO [11]

Abstract

We extend the ACOT scheme for heavy quark production to NNLO and N^{3}LO for the structure functions F_2 and F_L in deep-inelastic scattering (DIS). We use the fully massive ACOT scheme up to NLO, and estimate the dominant heavy quark mass effects at the higher orders using the massless Wilson coefficients together with a generalized slow-rescaling prescription. We present results for F_2 and F_L showing the effect of the higher orders and the contributions from the heavy flavors.

9.1 INTRODUCTION

The production of heavy quarks in high energy processes has become an increasingly important subject of study both theoretically and experimentally. The theory of heavy quark production in perturbative Quantum Chromodynamics (pQCD) is more challenging than that of light parton (jet) production because of the additional heavy quark mass scale. The correct theory must properly take into account the changing role of the heavy quark over the full kinematic range of the relevant process from the threshold region (where the quark behaves like a typical "heavy particle") to the asymptotic region (where the same quark behaves effectively like a parton, similar to the well known light quarks $\{u, d, s\}$).

With the ever-increasing precision of experimental data and the progression of theoretical calculations and parton distribution function (PDF) evolution to next-to-next-to-leading order (NNLO) of QCD, there is a clear need to implement the heavy quark schemes at this order and beyond. The most important case is arguably the heavy quark treatment in inclusive deep-inelastic scattering (DIS) since the very precise HERA data for DIS structure functions and cross sections form the backbone of any modern global analysis of PDFs. Here, the heavy quark structure functions contribute up to 30% or 40% to the inclusive structure functions at small momentum fractions x. Extending the heavy quark schemes to higher orders is relevant for extracting precision PDFs, and hence for accurate predictions of observables at the LHC.

An example where higher order corrections are particularly important is the longitudinal structure function F_L in DIS. The leading order $\mathcal{O}(\alpha_s^0)$ contributions to this structure function vanishes for massless quarks due to helicity conservation (Callan-Gross relation). Since the first unsuppressed contribution to F_L is at next-to-leading order, the NNLO and N^{3}LO corrections are more important than for F_2. In Fig. 16 we show the preliminary results for the F_L measurement from the H1 and ZEUS experiments [321]. In Fig. 17 displays sample Feynman diagrams at the various orders. Producing an accurate

[11] Contributed by: T. Stavreva, I. Schienbein, F. I. Olness, T. Ježo, K. Kovařík, A. Kusina, J. Y. Yu

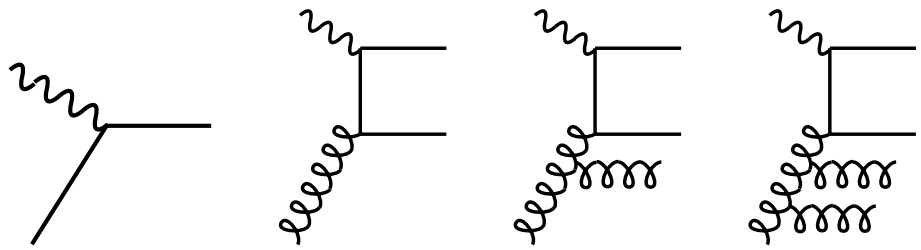


Fig. 17: Example Feynman diagrams contributing to DIS heavy quark production (from left): LO $\mathcal{O}(\alpha_S^0)$ quark-boson scattering $QV \to Q$, NLO $\mathcal{O}(\alpha_S^1)$ boson-gluon scattering $gV \to Q\bar{Q}$, NNLO $\mathcal{O}(\alpha_S^2)$ boson-gluon scattering $gV \to gQ\bar{Q}$ and N^{3}LO $\mathcal{O}(\alpha_S^3)$ boson-gluon scattering $gV \to ggQ\bar{Q}$.

prediction for F_L is a challenge, particularly in the region of low Q^2 and small x.

In this paper, we will briefly outline the method we used to incorporate the higher order terms, the key elements of the ACOT scheme, and the treatment of the heavy quark masses. We then present results for the F_2 and F_L neutral current DIS structure functions.

9.2 THE ACOT SCHEME AND ITS EXTENSION BEYOND NLO

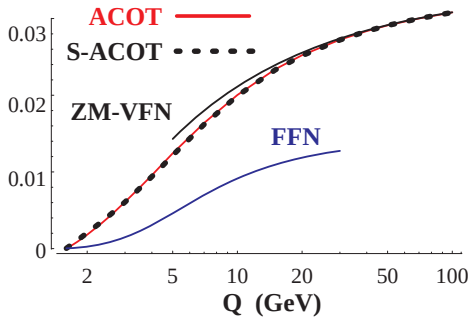


Fig. 18: Comparison of schemes for F_2^c at $x = 0.1$ for NLO DIS heavy quark production as a function of Q. We display calculations using the ACOT, S-ACOT, Fixed-Flavor Number Scheme (FFNS), and Zero-Mass Variable-Flavor-Number-Scheme (ZM-VFNS). The ACOT and S-ACOT results are virtually identical.

The ACOT scheme [322, 323] is based upon the factorization theorem for heavy quarks[324]; hence, it is valid at any order of perturbation theory. The factorization proof ensures that the ACOT scheme can be applied throughout the full kinematic regime, and that there is a smooth transition from a massless result ($m = 0$) to the heavy-mass decoupling limit ($m \to \infty$).

In the limit where the quark Q of mass m is relatively heavy compared to the characteristic energy scale ($\mu \lesssim m$), the ACOT result naturally reduces to the Fixed-Flavor-Number-Scheme (FFNS). In the FFNS, the heavy quark is treated as being extrinsic to the hadron, and there is no corresponding heavy quark PDF, $f_Q(x, \mu) = 0$. Conversely, in the limit where the quark mass is relatively light ($\mu \gtrsim m$), the ACOT result reduces to the $\overline{MS}$ Zero-Mass Variable-Flavor-Number-Scheme (ZM-VFNS) exactly–*without* any finite renormalizations. In this limit, the quark mass m no longer plays any dynamical role; it serves purely as a regulator. This feature is presented in Fig. 18 where we can see that the ACOT scheme precisely matches the results of the FFNS and ZM-VFNS schemes in their respective limits.

Additionally Fig. 18 shows the results obtained within the Simplified-ACOT scheme (S-ACOT) [325]. The S-ACOT scheme drops the heavy quark mass dependence for the hard-scattering

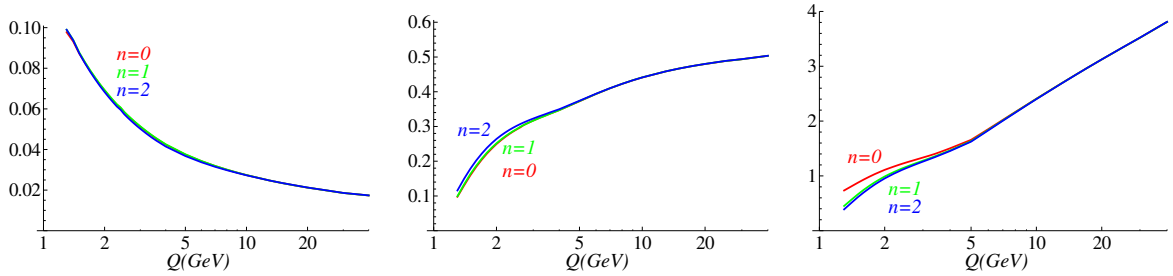


Fig. 19: F_L vs. Q at N^{3}LO for fixed $x = \{10^{-1}, 10^{-3}, 10^{-5}\}$, (left to right). The three lines show the mass effects via the scaling variable: $n = \{0, 1, 2\}$, (Red, Green, Blue). We observe the effect of the n-scaling is negligible except for small x and Q values.

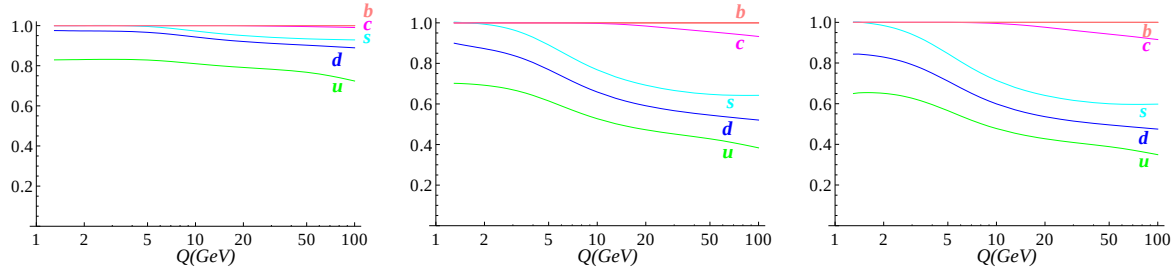
processes with incoming heavy quarks or with internal on-shell cuts on a heavy quark line. The S-ACOT scheme is *not* an approximation; it is an exact renormalization scheme, extensible to all orders. Note, the ACOT and S-ACOT results agree throughout the kinematic region.

9.21 Beyond NLO

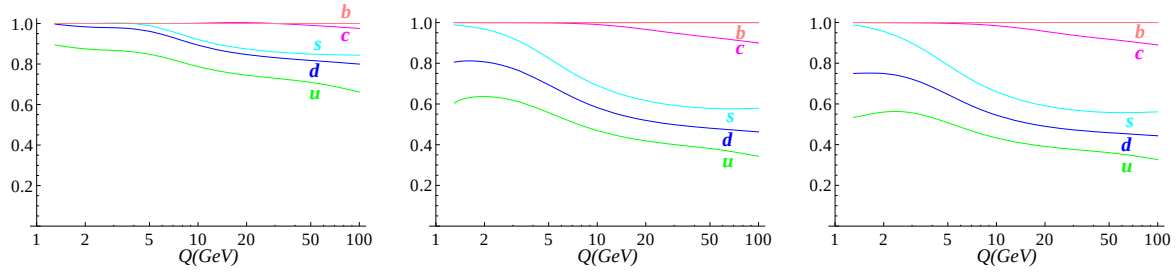
While there is no conceptual difficulty with extending the ACOT scheme beyond NLO, the fully massive Wilson coefficients have yet to be computed.[12] However massless calculations of NNLO and even N^{3}LO for F_2 and F_L structure functions are available.[13]

The question is: can we use these results, together with the knowledge that ACOT reduces to the massless $\overline{MS}$ (ZM-VFNS) for $m \to 0$, to estimate mass effects at NNLO and N^{3}LO? Obviously we cannot restore the fully massive ACOT result from the massless limit, but we can try to extract the dominant higher order contributions. There are two ways in which mass effects enter the calculation. The first is "dynamically" through the mass dependent Wilson coefficients. The second is "kinematically" via the restricted phase space. Comparisons using the fully massive results at NLO suggest that the kinematic mass effects are dominant, and that much of this dependence can be obtained with a rescaling of the Bjorken x variable. We introduce a generalized rescaling $x \to x[1 + (n\, m/Q)^2]$ where $n = 0$ is the massless result, $n = 1$ is the original Barnett[329] rescaling, and $n = 2$ is the χ-rescaling [330].

Thus, our strategy is as follows. We use the fully massive ACOT result to NLO [331], and add to this the massless NNLO and N^{3}LO contributions using the generalized rescaling prescription. By varying n, we can investigate the influence of the kinematic mass in our results. We argue that the massless Wilson coefficients at NNLO and N^{3}LO, together with the generalized rescaling prescription provide a good approximation of the exact result. At worst, the error is of order $\alpha\, \alpha_S^2 \times [m^2/Q^2]$, and comparative studies at NLO suggest the error is less.[14] For example, in Fig. 19 we display the results of F_L for $n = \{0, 1, 2\}$. The effects of the detailed mass dependence is most noticeable for low Q^2 and small x. While the massless scaling result ($n = 0$) does deviate from the other curves, comparing the $n = 1$ and $n = 2$ curves we observe the details of the mass rescaling are relatively small. While this is not a proof,[15] this result does give us confidence that the mass effects are under control.



(a) F_2 vs. Q.



(b) F_L vs. Q.

Fig. 20: Fractional contribution for each quark structure function $F_{2,L}^i$ for each flavor $i = \{u, d, s, c, b\}$ for (a) F_2^i and (b) F_L^i vs. Q at N^{3}LO for fixed $x = \{10^{-1}, 10^{-3}, 10^{-5}\}$ (left to right) for the ACOT-χ scheme.

9.3 RESULTS

In Fig. 20 we display the fractional contributions to the structure functions F_2 and F_L. At larger values of x and low Q, we observe that the heavy flavor contributions are minimal. For example, for $x = 10^{-1}$, we see that the u-quark structure function F^u comprises $\sim 80\%$ of the total structure function. In contrast, at $x = 10^{-5}$ and large Q we see that the contributions of the u and c quarks are comparable (as they couple with a factor 4/9), and the d and s quarks contributions are comparable (as they couple with a factor 1/9).

Figure 20 also shows how the χ-rescaling introduces a damping of the heavy quark contributions as we move from large Q^2 values to smaller values. The χ-rescaling ensures the heavy quarks (c, b) are appropriately suppressed for low Q^2 scales.

In Fig. 21a we display the results for F_2 vs. Q computed at various orders; the ratio to the N^{3}LO result is displayed in Fig. 21b. For large x (*c.f.* $x = 0.1$) we find the perturbative calculations are particularly stable. We see that the LO result is within 20% of the others at small Q, and within 5% at large Q. The NLO is within 2% at small Q, and indistinguishable from the NNLO and N^{3}LO for Q values above ~ 10 GeV. The NNLO and N^{3}LO results are essentially identical throughout the kinematic range. For smaller x values (10^{-3}, 10^{-5}), the contributions of the higher order terms are slightly larger. Here, the NNLO and N^{3}LO coincide for Q values above ~ 5 GeV, but the NLO result can differ by $\sim 5\%$ for low Q^2 scales.

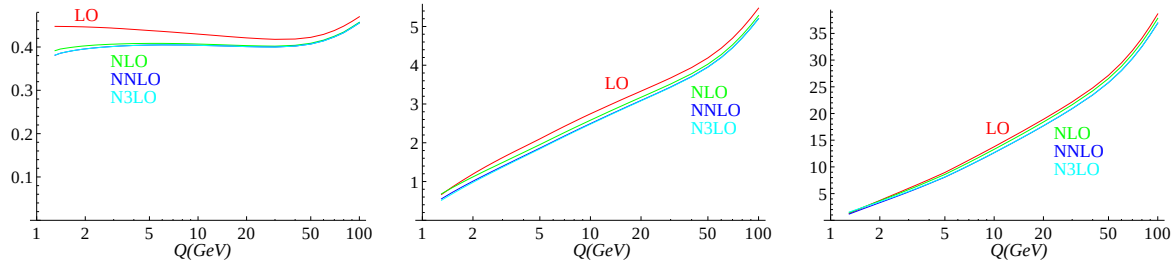
In Fig. 22 we display the results for F_L vs. Q computed at various orders. In contrast to F_2, we find that NLO corrections are large; this is expected because the LO corrections to F_L (which violate

[12]There has been a calculation of neutral current electroproduction of heavy quarks $\mathcal{O}(\alpha_s^2)$ in the FFNS [326]. however, extra contributions are still required for a VFNS calculation [327].

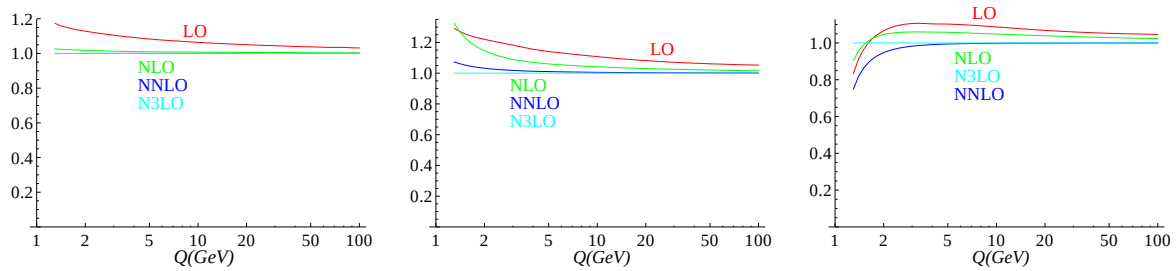
[13]See Ref. [328] and references therein.

[14]Details will be presented in a forthcoming publication.

[15]Of course, once the massive higher order Wilson coefficients have been computed, it is straightforward to incorporate these results into our calculations.



(a) F_2 vs. Q at {LO, NLO, NNLO, N^{3}LO} (red, green, blue, cyan) for fixed $x = \{10^{-1}, 10^{-3}, 10^{-5}\}$, (left to right) for ACOT-χ scheme.



(b) Ratio of F_2 vs. Q at {LO, NLO, NNLO, N^{3}LO} (red, green, blue, cyan) compared to F_2 at N^{3}LO for fixed $x = \{10^{-1}, 10^{-3}, 10^{-5}\}$, (left to right) for ACOT-χ scheme.

Fig. 21: F_2 vs. Q at {LO, NLO, NNLO, N^{3}LO}.

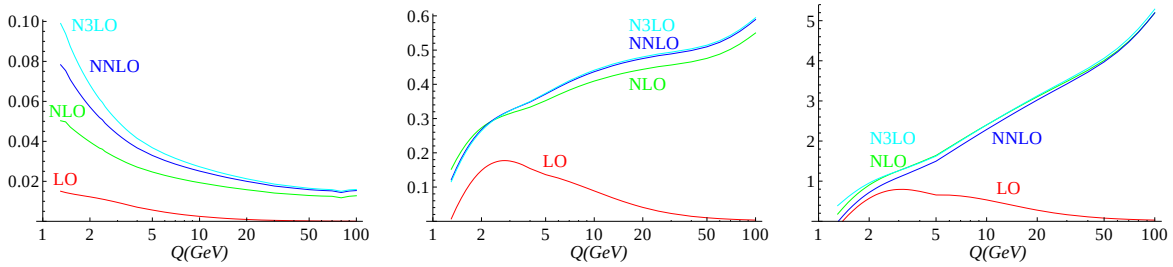
the Callan-Gross relation) are suppressed by (m^2/Q^2) compared to the dominant gluon contributions which enter at NLO. Consequently, we observe that the LO result for F_L receives large contributions from the higher order terms. Essentially, NLO is the first non-trivial order for F_L, and the subsequent contributions then converge. For example, at large x (*c.f.* $x = 0.1$) for $Q \sim 10$ GeV we find the NLO results yields $\sim 70\%$ of the total, the NNLO is a $\sim 20\%$ correction, and the N^{3}LO is a $\sim 10\%$ correction. For lower x values (10^{-3}, 10^{-5}) the convergence of the perturbative series improves, and the NLO results is within $\sim 10\%$ of the N^{3}LO result. Curiously, for $x = 10^{-5}$ the NNLO and N^{3}LO roughly compensate each other so that the NLO and the N^{3}LO match quite closely for $Q \gtrsim 2$ GeV.

9.4 CONCLUSIONS

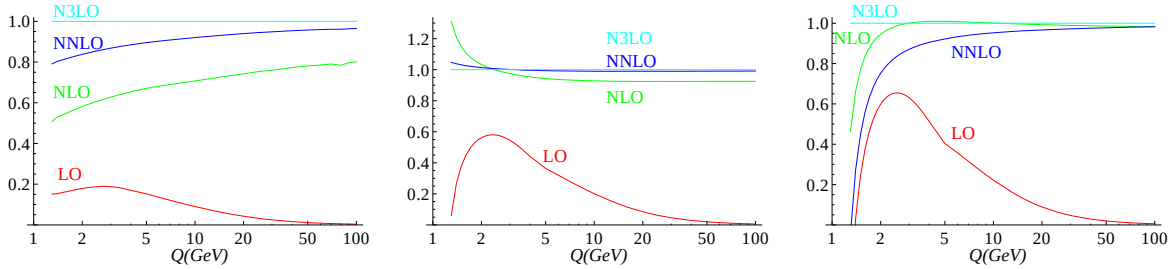
We have computed the F_2 and F_L structure functions in the ACOT scheme at NNLO and N^{3}LO. The full mass dependence is computed to NLO, and the dominant mass effects for the higher orders are approximated using a generalized rescaling; the details of this rescaling are demonstrated to be small. This allows us to make detailed predictions throughout the kinematic range investigated by HERA, and we obtain a reasonable estimate of the uncertainty due to the higher order mass effects. Together with the precise HERA data, these calculations facilitate accurate determination of the PDFs which are the foundation of the LHC calculations.

ACKNOWLEDGMENTS

We thank M. Botje, A. M. Cooper-Sakar, A. Glazov, C. Keppel, J. G. Morfín, P. Nadolsky, J. F. Owens, V. A. Radescu, for valuable discussions. F.I.O., I.S., and J.Y.Y. acknowledge the hospitality of Argonne, BNL, CERN, Fermilab, and Les Houches where a portion of this work was performed. F.I.O thanks the Galileo Galilei Institute for Theoretical Physics and the INFN for their hospitality and support. This work was supported by the U.S. Department of Energy under grant DE-FG02-04ER41299, and the Lightner-Sams Foundation. The work of J. Y. Yu was supported by the Deutsche Forschungsgemeinschaft (DFG) through grant No. YU 118/1-1. The work of K. Kovařík was supported by the ANR projects ANR-06-



(a) F_L vs. Q at {LO, NLO, NNLO, N^{3}LO} (red, green, blue, cyan) for fixed $x = \{10^{-1}, 10^{-3}, 10^{-5}\}$, (left to right) for ACOT-χ scheme.



(b) Ratio of F_L vs. Q at {LO, NLO, NNLO, N^{3}LO} (red, green, blue, cyan) compared to F_L at N^{3}LO for fixed $x = \{10^{-1}, 10^{-3}, 10^{-5}\}$, (left to right) for ACOT-χ scheme.

Fig. 22: F_L vs. Q at {LO, NLO, NNLO, N^{3}LO}.

JCJC-0038-01 and ToolsDMColl, BLAN07-2-194882.

Part IV

PHENOMENOLOGICAL STUDIES OF OBSERVABLES AND UNCERTAINTIES

10. FINITE-WIDTH EFFECTS IN TOP-QUARK PAIR PRODUCTION AND DECAY AT THE LHC [16]

Abstract

We investigate finite-top-width effects in top-quark pair production by comparing NLO QCD predictions for $\mathrm{pp} \to \mathrm{WWb\bar{b}}$ to corresponding $\mathrm{pp} \to \mathrm{t\bar{t}} \to \mathrm{WWb\bar{b}}$ results in the narrow-top-width limit. Finite-top-width effects, which result from non-resonant and off-shell contributions, are discussed in detail for the case of the inclusive cross section (with experimental cuts) and for selected differential observables in the di-lepton channel.

10.1 INTRODUCTION

Top-quark pair production at hadron colliders allows for key tests of the Standard Model and represents an omnipresent background to Higgs-boson and new-physics searches. The very large $\mathrm{t\bar{t}}$ samples from the Tevatron and the LHC, and the steadily increasing systematic precision call for a continu-

[16] Contributed by: A. Denner, S. Dittmaier, S. Kallweit, S. Pozzorini, M. Schulze

ous improvement of theory predictions.[17] In this context, a reliable theoretical description of experimental cuts and exclusive $\mathrm{t\bar{t}}$ observables, which depend on details of the $\mathrm{W^+W^-b\bar{b}}$ final state, requires higher-order calculations for top-pair *production and decay*. The first NLO QCD predictions for $\mathrm{pp} \to \mathrm{t\bar{t}} \to \mathrm{WWb\bar{b}} + X$ [335, 336, 337] have been obtained in the narrow-top-width limit, an approximation where the $2 \to 4$ particle process is factorised into on-shell $\mathrm{t\bar{t}}$ production and (anti)top decays, taking into account spin correlations. In this framework, it was shown that NLO QCD effects in top-quark decays have a significant impact on the kinematic properties of final-state leptons and b-jets [335, 336, 337], and play an important role for top-mass measurements at the LHC [338]. More recently, NLO QCD predictions for the complete $\mathrm{pp} \to \mathrm{W^+W^-b\bar{b}} + X$ process became available [25, 23], which include all effects related to the finite top-quark width, i.e. on- and off-shell intermediate top quarks, non-resonant contributions, and their interference with resonant $\mathrm{t\bar{t}}$ production. Besides new evidence for the importance of NLO corrections to $\mathrm{t\bar{t}}$ production and decay, these studies provided a first quantitative assessment of finite-width effects in the inclusive cross section. Applying a numerical $\Gamma_\mathrm{t} \to 0$ extrapolation to the NLO $\mathrm{pp} \to \mathrm{W^+W^-b\bar{b}}$ predictions, it was found that finite-top-width contributions to the $\mathrm{WWb\bar{b}}$ cross section at the Tevatron and the LHC (7 TeV) range from 0.2 to 1 percent [25, 23], which is perfectly consistent with the expected order of magnitude ($\Gamma_\mathrm{t}/m_\mathrm{t} \simeq 0.9\%$) of finite-top-width effects in inclusive observables.

In this study, we pursue the investigation of finite-top-width effects by means of a tuned comparison of the $\mathrm{pp} \to \mathrm{W^+W^-b\bar{b}}$ NLO calculation of Ref. [25] against the narrow-top-width approximation of Ref. [336]. This permits us, for the first time, to investigate Γ_t-effects in different phenomenologically interesting regions of the $\mathrm{WWb\bar{b}}$ phase space, where large off-shell and non-resonant contributions cannot be excluded a priori as in the case of inclusive observables.

10.2 NARROW-TOP-WIDTH APPROXIMATION AND FINITE-WIDTH EFFECTS

Let us start by recalling the main features of the NLO QCD calculations of $\mathrm{pp} \to \mathrm{t\bar{t}} \to \mathrm{W^+W^-b\bar{b}}$ in narrow-top-width approximation [336] and $\mathrm{pp} \to \mathrm{W^+W^-b\bar{b}}$ with finite-top-width effects [25]. For brevity, we denote them as $\mathrm{t\bar{t}}$ and $\mathrm{WWb\bar{b}}$ calculations, respectively. Both calculations implement leptonic W-boson decays in spin-correlated narrow-W-width approximation.

In the narrow-top-width limit of Ref. [336], top-quark resonances are approximated by

$$\lim_{\Gamma_\mathrm{t}/m_\mathrm{t}\to 0} \frac{1}{(p_\mathrm{t}^2 - m_\mathrm{t}^2)^2 + m_\mathrm{t}^2\Gamma_\mathrm{t}^2} = \frac{\pi}{m_\mathrm{t}\Gamma_\mathrm{t}}\delta\big(p_\mathrm{t}^2 - m_\mathrm{t}^2\big), \tag{77}$$

with delta functions that enforce the on-shell conditions, $p_\mathrm{t}^2 = m_\mathrm{t}^2$, and are accompanied by $1/\Gamma_\mathrm{t}$ factors. Contributions of $\mathcal{O}(\Gamma_\mathrm{t}/m_\mathrm{t})$, i.e. terms that do not involve two resonant top propagators, are systematically neglected. The differential $\mathrm{pp} \to \mathrm{t\bar{t}} \to \mathrm{WWb\bar{b}}$ cross section is factorised into the $\mathrm{pp} \to \mathrm{t\bar{t}}$ cross section times $\mathrm{t} \to \mathrm{Wb}$ partial decay widths, $\mathrm{d}\sigma = \left(\mathrm{d}\sigma_{\mathrm{t\bar{t}}}\,\mathrm{d}\Gamma_\mathrm{t}\mathrm{d}\Gamma_{\bar{\mathrm{t}}}\right)/\Gamma_\mathrm{t}^2$, taking into account top-quark spin correlations. The LO and NLO predictions can be schematically expressed as

$$\begin{aligned}
\mathrm{d}\sigma_\mathrm{LO} &= \Gamma_\mathrm{t,LO}^{-2}\left(\mathrm{d}\sigma_{\mathrm{t\bar{t}}}^0\,\mathrm{d}\Gamma_\mathrm{t}^0\mathrm{d}\Gamma_{\bar{\mathrm{t}}}^0\right), \\
\mathrm{d}\sigma_\mathrm{NLO} &= \Gamma_\mathrm{t,NLO}^{-2}\left[\left(\mathrm{d}\sigma_{\mathrm{t\bar{t}}}^0 + \mathrm{d}\sigma_{\mathrm{t\bar{t}}}^1\right)\,\mathrm{d}\Gamma_\mathrm{t}^0\mathrm{d}\Gamma_{\bar{\mathrm{t}}}^0 + \mathrm{d}\sigma_{\mathrm{t\bar{t}}}^0\,\left(\mathrm{d}\Gamma_\mathrm{t}^1\mathrm{d}\Gamma_{\bar{\mathrm{t}}}^0 + \mathrm{d}\Gamma_\mathrm{t}^0\mathrm{d}\Gamma_{\bar{\mathrm{t}}}^1\right)\right],
\end{aligned} \tag{78}$$

where the superscripts 0 and 1 indicate tree-level quantities and NLO corrections, respectively. The NLO prediction involves three terms, where the corrections are applied either to $\mathrm{d}\sigma_{\mathrm{t\bar{t}}}$ or to one of the decays. All ingredients of $\mathrm{d}\sigma_\mathrm{LO}$ and $\mathrm{d}\sigma_\mathrm{NLO}$ have to be evaluated with input parameters at the corresponding perturbative order. In particular, LO and NLO predictions must be computed using $\Gamma_\mathrm{t,LO}$ and $\Gamma_\mathrm{t,NLO}$ decay widths, as indicated in (78).[18] This guarantees that—up to higher-order corrections—the integration over

[17] Recent progress in the theoretical description of top-quark pair production at hadron colliders is reviewed in Refs. [332, 333, 334].

[18] We note that in the present study the factor $\Gamma_\mathrm{t,NLO}^{-2}$ in (78) is not expanded as $\left(\Gamma_\mathrm{t,LO} + \Gamma_\mathrm{t}^1\right)^{-2} = \Gamma_\mathrm{t,LO}^{-2}(1 - 2\,\Gamma_\mathrm{t}^1/\Gamma_\mathrm{t,LO})$, like in Eq. (6) of Ref. [336], since this procedure is not directly applicable to the full $\mathrm{WWb\bar{b}}$ calculation.

the phase space of each top decay in (78) is consistent with the branching fraction

$$\frac{\int \mathrm{d}\Gamma_{\mathrm{t}\to \mathrm{b}l\nu}}{\Gamma_{\mathrm{t}}} = \frac{\Gamma_{\mathrm{t}\to \mathrm{b}l\nu}}{\Gamma_{\mathrm{t}}} = \mathrm{BR}(\mathrm{t}\to \mathrm{b}l\nu). \tag{79}$$

In this context, let us point out that a consistent inclusion of finite-W-width corrections—both in the scattering amplitudes and the Γ_{t} input parameters—is expected to lead to doubly-suppressed effects. This is due to the fact that, in the $\Gamma_{\mathrm{t}} \to 0$ limit, $\mathcal{O}(\Gamma_{\mathrm{W}})$ corrections to the numerator and denominator of the branching fraction (79) cancel. Finite-W-width corrections are thus expected to produce very small effects of $\mathcal{O}(\frac{\Gamma_{\mathrm{W}}\Gamma_{\mathrm{t}}}{M_{\mathrm{W}} m_{\mathrm{t}}})$ in inclusive observables. This justifies the use of the narrow-W-width approximation in combination with finite-top-width contributions, which is the approach adopted in Ref. [25], although in kinematic regions where finite-Γ_{t} effects become large also finite-W-width corrections [23] might become non-negligible.

The calculation of Ref. [25] provides a full description of $\mathrm{pp} \to \mathrm{W^+W^-b\bar{b}}$ at order $\mathcal{O}(\alpha_{\mathrm{S}}^3\alpha^2)$. The top-quark width is incorporated into the complex top mass, $\mu_{\mathrm{t}}^2 = m_{\mathrm{t}}^2 - \mathrm{i}m_{\mathrm{t}}\Gamma_{\mathrm{t}}$, in the complex-mass scheme [339]. In this way, off-shell-top contributions are consistently described by Breit–Wigner distributions. Besides contributions with two intermediate top resonances, also singly- and non-resonant diagrams are taken into account, including interferences. A few representative tree diagrams are shown in Fig.23. The NLO $\mathrm{WWb\bar{b}}$ predictions involve factorisable corrections to doubly-resonant diagrams, which provide the off-shell extension of NLO corrections in $\mathrm{t\bar{t}}$ approximation (78). In addition, there are non-factorisable corrections, where $\mathrm{t\bar{t}}$ production and decay parts of the process are connected via exchange of QCD partons, and NLO corrections to singly- and non-resonant topologies. Further technical aspects are discussed in the original publications [336, 25].

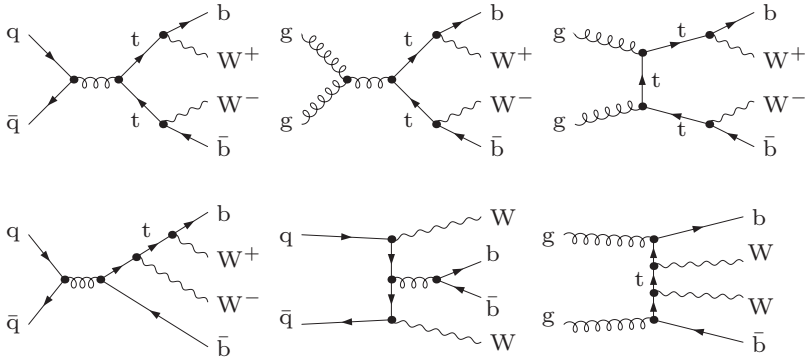


Fig. 23: Representative LO diagrams of doubly-resonant (upper line), singly-resonant (first diagram in lower line), and non-resonant type (last two diagrams in lower line).

10.3 NUMERICAL RESULTS

10.31 Input parameters and setup

In the following we compare $\mathrm{t\bar{t}}$ and $\mathrm{WWb\bar{b}}$ predictions for $\mathrm{W^+}(\to \nu_{\mathrm{e}}\mathrm{e^+})\mathrm{W^-}(\to \mu^-\bar{\nu}_\mu)\mathrm{b\bar{b}}$ production at the Tevatron ($\mathrm{p\bar{p}}$ collisions at 1.96 TeV) and the LHC (pp collisions at 7 and 14 TeV). These results are based on the same input parameters and cuts as in Ref. [25]. In NLO (LO) QCD we employ MSTW2008NLO (LO) parton distributions [262] and describe the running of the strong coupling constant α_{S} with two-loop (one-loop) accuracy, including five active flavours. Contributions induced by the strongly suppressed bottom-quark density are neglected. For the gauge-boson and top-quark masses we use $m_{\mathrm{t}} = 172\,\mathrm{GeV}$, $M_{\mathrm{W}} = 80.399\,\mathrm{GeV}$, and $M_{\mathrm{Z}} = 91.1876\,\mathrm{GeV}$. The masses of all other quarks, including b-quarks, are neglected. In view of the negligibly small Higgs-mass dependence we adopt the $M_{\mathrm{H}} \to \infty$ limit, i.e. we omit diagrams involving Higgs bosons. The electroweak couplings are derived

Collider	$\sqrt{s}$ [TeV]	approx.	$\sigma_{\mathrm{t\bar{t}}}$ [fb]	$\sigma_{\mathrm{WWb\bar{b}}}$ [fb]	$\sigma_{\mathrm{t\bar{t}}}/\sigma_{\mathrm{WWb\bar{b}}} - 1$	Ref. [25]
Tevatron	1.96	LO	$44.691(8)^{+19.81}_{-12.58}$	$44.310(3)^{+19.68}_{-12.49}$	$+0.861(19)\%$	$+0.8\%$
		NLO	$42.16(3)^{+0.00}_{-2.91}$	$41.75(5)^{+0.00}_{-2.63}$	$+0.98(14)\%$	$+0.9\%$
LHC	7	LO	$659.5(1)^{+261.8}_{-173.1}$	$662.35(4)^{+263.4}_{-174.1}$	$-0.431(16)\%$	-0.4%
		NLO	$837(2)^{+42}_{-87}$	$840(2)^{+41}_{-87}$	$-0.41(31)\%$	-0.2%
LHC	14	LO	$3306.3(1)^{+1086.8}_{-763.6}$	$3334.6(2)^{+1098.5}_{-771.2}$	$-0.849(7)\%$	$---$
		NLO	$4253(3)^{+282}_{-404}$	$4286(7)^{+283}_{-407}$	$-0.77(19)\%$	$---$

Table 8: Integrated $\nu_{\mathrm{e}}\mathrm{e}^+\mu^-\bar{\nu}_\mu\mathrm{b\bar{b}}$ cross section in narrow-with approximation ($\sigma_{\mathrm{t\bar{t}}}$) and including finite-top-width effects ($\sigma_{\mathrm{WWb\bar{b}}}$). The relative error of the narrow-width approximation (sixth column) is compared to the prediction of Ref. [25] (seventh column). Factor-two scale variations in $\sigma_{\mathrm{t\bar{t}}}$ and $\sigma_{\mathrm{WWb\bar{b}}}$ are shown as sub- and super-scripts, while statistical errors are given in parenthesis.

from the Fermi constant $G_\mu = 1.16637 \times 10^{-5}\,\mathrm{GeV}^{-2}$ in the G_μ-scheme, where the sine of the mixing angle and the electromagnetic coupling read $s_{\mathrm{w}}^2 = 1 - M_{\mathrm{W}}^2/M_{\mathrm{Z}}^2$ and $\alpha = \sqrt{2}G_\mu M_{\mathrm{W}}^2 s_{\mathrm{w}}^2/\pi$. For consistency, we perform the LO and NLO calculations using the top-quark widths $\Gamma_{\mathrm{t,LO}} = 1.4655\,\mathrm{GeV}$ and $\Gamma_{\mathrm{t,NLO}} = 1.3376\,\mathrm{GeV}$ [340], respectively. Since the leptonic W-boson decay does not receive NLO QCD corrections we employ the NLO W-boson width $\Gamma_{\mathrm{W}} = 2.0997\,\mathrm{GeV}$ everywhere.

Final-state quarks and gluons with pseudo-rapidity $|\eta| < 5$ are converted into infrared-safe jets using the anti-k_{T} algorithm [341]. For the Tevatron (LHC) we set the jet-algorithm parameter $R = 0.4\,(0.5)$ and apply the transverse-momentum and pseudo-rapidity cuts $p_{\mathrm{T,b-jet}} > 20\,(30)\,\mathrm{GeV}$, $|\eta_{\mathrm{b-jet}}| < 2.5$. Moreover, we require a missing transverse momentum of $p_{\mathrm{T,miss}} > 25\,(20)\,\mathrm{GeV}$ and charged leptons with $p_{\mathrm{T},l} > 20\,\mathrm{GeV}$ and $|\eta_l| < 2.5$.

For the renormalisation and factorisation scales we adopt the central value $\mu = m_{\mathrm{t}}$ and study factor-two variations of $\mu = \mu_{\mathrm{ren}} = \mu_{\mathrm{fact}}$, i.e. we compare predictions at $\mu/m_{\mathrm{t}} = 0.5, 1, 2$. The scale variations are applied also to $\Gamma_{\mathrm{t,NLO}}$, but not to Γ_{W}.

10.32 Integrated cross section

Results for the integrated $\nu_{\mathrm{e}}\mathrm{e}^+\mu^-\bar{\nu}_\mu\mathrm{b\bar{b}}$ cross sections and scale uncertainties at the Tevatron and the LHC are reported in Table 8. While the $\sigma_{\mathrm{WWb\bar{b}}}$ results for Tevatron and LHC at 7 TeV correspond to those of Ref. [25][19], the ones for LHC at 14 TeV as well as all $\sigma_{\mathrm{t\bar{t}}}$ predictions are new. Comparing all $\mathrm{WWb\bar{b}}$ and $\mathrm{t\bar{t}}$ predictions we find that finite-top-width effects never exceed one percent, both in LO and NLO. The statistical precision of the calculations permits us to assess the error of the NWA, $\sigma_{\mathrm{t\bar{t}}}/\sigma_{\mathrm{WWb\bar{b}}} - 1$, with an accuracy of 1–3 permille. At the Tevatron, the NWA overestimates the $\mathrm{WWb\bar{b}}$ cross section by an amount very close to $\Gamma_{\mathrm{t}}/m_{\mathrm{t}} \simeq 0.9\%$, both in LO and NLO. The error of the NWA at the 7(14) TeV LHC ranges between 4 and 8 permille. As shown in the last column of Table 8, these finite-width effects are in very good agreement with the results of the $\Gamma_{\mathrm{t}} \to 0$ extrapolation in Ref. [25]. Similar results can be found also in Ref. [23].

[19] To be more precise, in Ref. [25] the scale dependence was assessed using a fixed Γ_{t} input, while here we take into account the μ-dependence of $\Gamma_{\mathrm{t,NLO}}$, which results into slightly different $\sigma_{\mathrm{WWb\bar{b}}}$ variations at NLO.

10.33 Differential distributions

The small finite-width corrections to the integrated cross section demonstrate that—in presence of standard LHC and Tevatron cuts—the NWA provides a fairly accurate description of inclusive $\mathrm{WWb\bar{b}}$ production. It is thus interesting to investigate to which extent this conclusion applies to the various phenomenologically important regions of the $\mathrm{WWb\bar{b}}$ phase space. To this end we have compared $\mathrm{t\bar{t}}$ and $\mathrm{WWb\bar{b}}$ predictions for a few differential observables that are relevant for top-pair production, either as signal or as background to Higgs production or new physics. Note that we refrain from selecting kinematic variables like the top-quark invariant mass or imposing cuts of type $M_{\mathrm{Wb}} > 200\,\mathrm{GeV}$, which would lead to obvious enhancements of non-resonant contributions.

In Figs. 24–27 we present predictions for some invariant-mass and transverse-momentum distributions, restricting ourselves to the case of the 7 TeV LHC. For each observable we display $\mathrm{t\bar{t}}$ (dashed curves) and $\mathrm{WWb\bar{b}}$ (solid curves) results in LO (blue) and NLO (red) approximation. Absolute predictions (left plots) are complemented by the ratios $(\mathrm{d}\sigma_{\mathrm{LO}} - \mathrm{d}\sigma_{\mathrm{NLO}})/\mathrm{d}\sigma_{\mathrm{NLO}}$ (upper right plots) and $(\mathrm{d}\sigma_{\mathrm{t\bar{t}}} - \mathrm{d}\sigma_{\mathrm{WWb\bar{b}}})/\mathrm{d}\sigma_{\mathrm{WWb\bar{b}}}$ (lower right plots), which indicate the relative error of LO and narrow-width approximations w.r.t. the best predictions, i.e. NLO and $\mathrm{WWb\bar{b}}$.

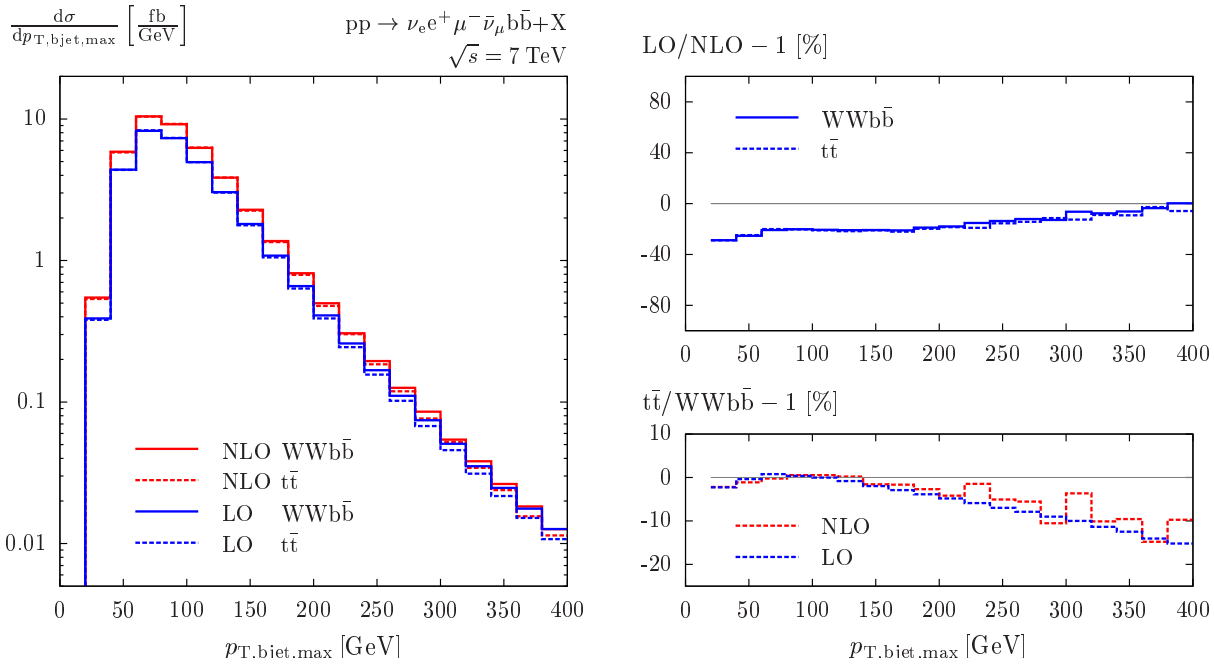


Fig. 24: Distribution in the transverse momentum of the harder b-jet at the 7 TeV LHC: LO (blue) and NLO (red) predictions in narrow-width approximation ($\mathrm{t\bar{t}}$, dashed) and including finite-top-width effects ($\mathrm{WWb\bar{b}}$, solid). Plotted are absolute predictions (left) and relative deviations of LO (upper-right) and narrow-width (lower-right) approximations w.r.t. NLO and $\mathrm{WWb\bar{b}}$ predictions, respectively.

The transverse-momentum distribution of the harder b-jet is shown in Fig. 24. In the range below 200 GeV, which contains the bulk of the cross section, the NLO and finite-width corrections behave similarly as for the integrated cross section: LO predictions deviate from NLO ones by about -20%, and the error of the NWA ranges between $+1$ and -4%. Finite-width effects tend to increase with p_{T} and reach the 10% level around 300 GeV. Within the entire p_{T} range the LO/NLO ratios resulting from the $\mathrm{t\bar{t}}$ and $\mathrm{WWb\bar{b}}$ calculations are almost equal. Equivalently, we find the same $\mathrm{d}\sigma_{\mathrm{t\bar{t}}}/\mathrm{d}\sigma_{\mathrm{WWb\bar{b}}}$ ratios in LO and NLO.

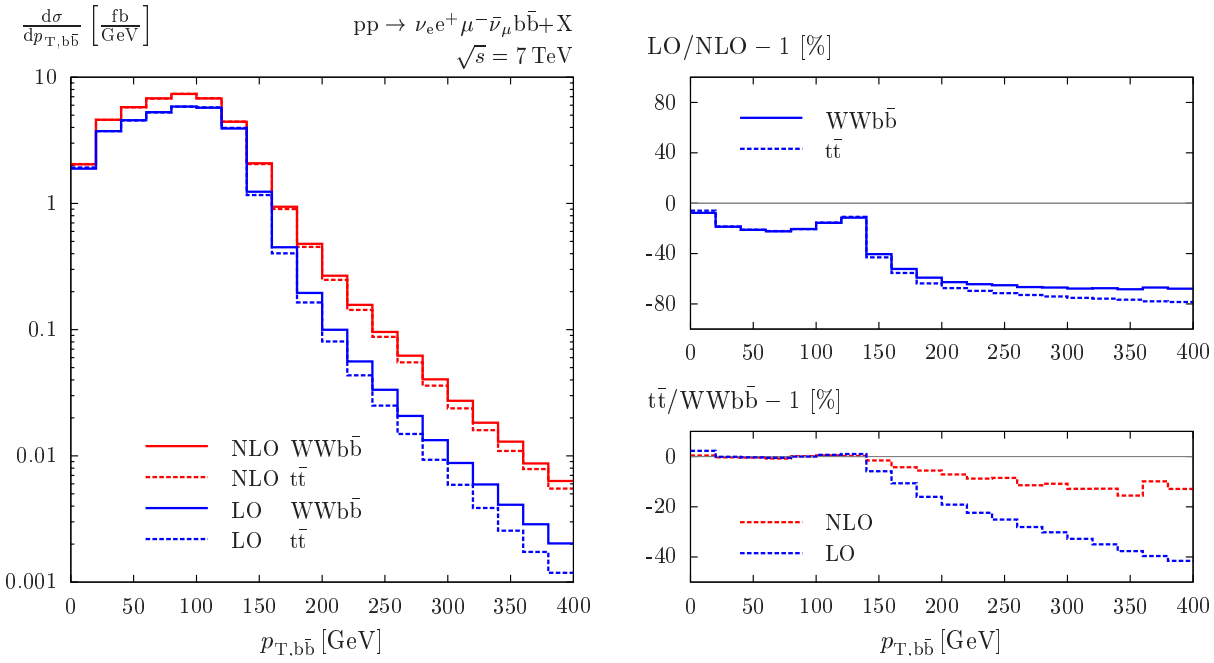


Fig. 25: Distribution in the transverse momentum of the $b\bar{b}$ di-jet system at the 7 TeV LHC: LO (blue) and NLO (red) predictions in narrow-width approximation ($t\bar{t}$, dashed) and including finite-top-width effects ($WWb\bar{b}$, solid). Plotted are absolute predictions (left) and relative deviations of LO (upper-right) and narrow-width (lower-right) approximations w.r.t. NLO and $WWb\bar{b}$ predictions, respectively.

In Fig. 25 we show the transverse-momentum distribution of the $b\bar{b}$ di-jet system. This kinematic variable plays an important role in boosted-Higgs searches with a large $t\bar{t}$ background. In particular, the strategy proposed in Ref. [342] to extract a $pp \to H(\to b\bar{b})W$ signal at the LHC is based on the selection of boosted $H \to b\bar{b}$ candidates with $p_{T,b\bar{b}} > 200\,\text{GeV}$, which permits to reduce $t\bar{t}$ contamination (and other backgrounds) in a very efficient way. As can be seen from Fig. 25, the suppression of $t\bar{t}$ production is indeed particularly strong at $p_{T,b\bar{b}} \gtrsim 150\,\text{GeV}$. This is due to kinematic constraints that characterise the LO and narrow-width approximations: in order to acquire $p_{T,b} > (m_t^2 - M_W^2)/(2m_t) \simeq 65\,\text{GeV}$ b-quarks need to be boosted via the p_T of their parent (anti)top quarks, and the fact that top and antitop quarks have opposite transverse momenta (at LO) makes it difficult to generate a $b\bar{b}$ system with high p_T. The NLO and finite-width corrections undergo less stringent kinematic restrictions, resulting into a significant enhancement of $WWb\bar{b}$ events at large $p_{T,b\bar{b}}$. This is clearly reflected in the differences between the various curves in the left plot of Fig. 25. The most pronounced effect comes from the NLO corrections, where the $t\bar{t}$ system can acquire large transverse momentum by recoiling against extra jet radiation. As indicated by the right-upper plot, the NLO correction represents 50–80% of the cross section at high p_T, corresponding to a huge K-factor of 2–5. Finite-width effects (lower-right plot) lead to a further significant, although less dramatic, enhancement; for example, non-resonant topologies can lead to direct $b\bar{b}$ production via high-p_T gluons that recoil against W^+W^- pairs. For $p_{T,b\bar{b}} > 200\,\text{GeV}$, we find that 20–40% of the LO $WWb\bar{b}$ cross section is due to finite-width contributions, while this fraction decreases to 7–15% at NLO. This reduction is related to the dominance of the jet-emission contribution, which we expect to be rather well described by the NWA. On the other hand, an optimal suppression of the $t\bar{t}$ background will require a very tight jet-veto [342], and in this case we expect finite-width corrections to the NLO $t\bar{t}$ predictions to be as large as in LO.

The distribution in the missing transverse momentum, i.e. the vector sum of the ν_e and $\bar{\nu}_\mu$ trans-

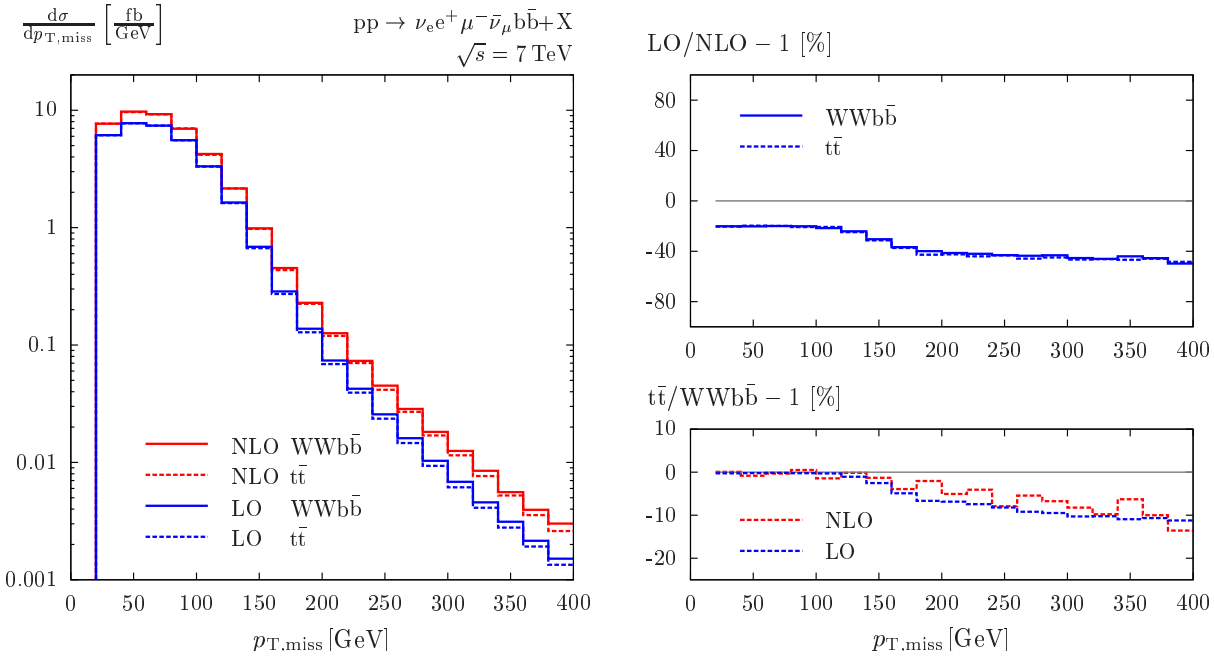


Fig. 26: Distribution in the missing transverse momentum at the 7 TeV LHC: LO (blue) and NLO (red) predictions in narrow-width approximation ($t\bar{t}$, dashed) and including finite-top-width effects ($WWb\bar{b}$, solid). Plotted are absolute predictions (left) and relative deviations of LO (upper-right) and narrow-width (lower-right) approximations w.r.t. NLO and $WWb\bar{b}$ predictions, respectively.

verse momenta, is displayed in Fig. 26. This distribution is relevant for new-physics searches based on missing transverse energy plus jets and leptons. Its tail features a qualitatively similar behaviour as in the case of $p_{T,b\bar{b}}$, due to analogous kinematic constraints. However, in the case of $p_{T,miss}$ the corrections are less pronounced: the NLO correction does not exceed 40–50% of the full prediction, and finite-width contributions stay below roughly 10%.

Figure 27 displays the distribution in the invariant mass of the positron and a b-jet, i.e. the visible products of a top-quark decay. More precisely, assuming that the charge of the b-jet is not known, the e^+b pair is built by selecting the b-jet that yields the smallest invariant mass.[20] In narrow-width and LO approximation this kinematic quantity is characterised by a sharp upper bound, $M^2_{e^+b} < m_t^2 - M_W^2 \simeq (152\,\text{GeV})^2$, which renders it very sensitive to the top-quark mass. The value of m_t can be extracted with high precision using, for instance, the invariant-mass distribution of a positron and a J/ψ from a B-meson decay [343, 338], an observable that is closely related to M_{e^+b}. In the region below the kinematic bound, the NLO corrections to M_{e^+b} vary between 5–30%, and the impact of the NLO shape distortion on a precision m_t-measurement is certainly significant. For $M_{e^+b} < 150\,\text{GeV}$, the NWA agrees with the $WWb\bar{b}$ predictions at the 1% level or better. In contrast, in the vicinity of the kinematic bound the impact of finite-width (and NLO) corrections becomes clearly more important, giving rise to a tail that extends above $M^2_{e^+b} = m_t^2 - M_W^2$. The resulting contribution to the total cross section is fairly small, but the impact of such finite-width effects on the top-mass measurement might be non-negligible, given the high m_t-sensitivity of the $M^2_{e^+b} \simeq m_t^2 - M_W^2$ region.

[20] Note that the M_{e^+b} distribution in Ref. [25] was defined by selecting (based on the Monte Carlo truth) jets that involve negatively charged b-quarks, such that the e^+b pairs are consistent with top-quark decays.

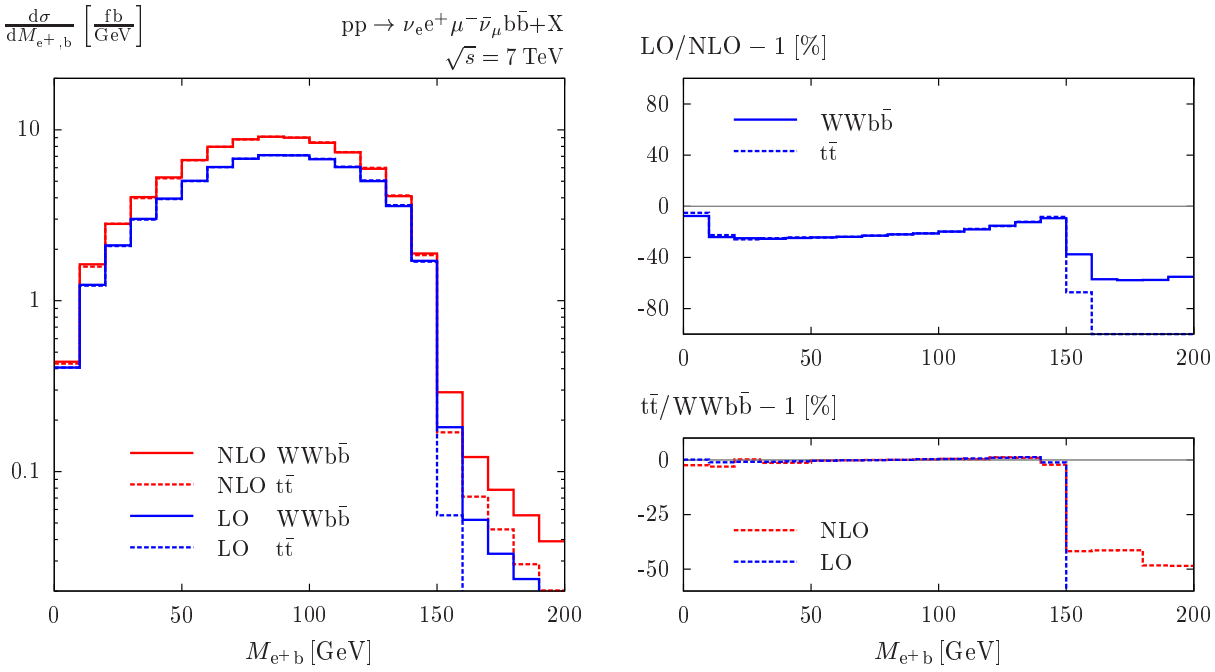


Fig. 27: Distribution in the invariant mass of the positron–b-jet system (as defined in the text) at the 7 TeV LHC: LO (blue) and NLO (red) predictions in narrow-width approximation ($t\bar{t}$, dashed) and including finite-top-width effects ($WWb\bar{b}$, solid). Plotted are absolute predictions (left) and relative deviations of LO (upper-right) and narrow-width (lower-right) approximations w.r.t. NLO and $WWb\bar{b}$ predictions, respectively.

10.4 CONCLUSIONS

Based on recent NLO QCD calculations, we have presented a systematic comparison of top-pair production and decay in narrow-top-width approximation, $pp \to t\bar{t} \to WWb\bar{b}$, against the complete $pp \to WWb\bar{b}$ process, which involves finite-top-width effects of non-resonant and off-shell type.

At the Tevatron and the LHC (7 and 14 TeV), finite-top-width contributions to the integrated cross section (in the di-lepton channel) turn out not to exceed one percent. This confirms previous estimates based on the $\Gamma_t \to 0$ extrapolation of $pp \to WWb\bar{b}$ predictions. At the 7 TeV LHC, we also investigated differential observables that are relevant either for top-pair production as a signal or as a background in Higgs or new-physics searches. In the case of the b-jet transverse momentum and $p_{T,miss}$ distributions, finite-width effects remain very small over a large kinematic range and reach the 10% level only around 300 GeV. In contrast, the p_T-distribution of the $b\bar{b}$ di-jet system receives Γ_t-corrections beyond 20–30% for $p_{T,b\bar{b}} \gtrsim 200\,\text{GeV}$, a kinematic region that plays an important role in $pp \to H(\to b\bar{b})W$ searches based on boosted $H \to b\bar{b}$ candidates. For the lepton–b-jet invariant-mass distribution—an observable that provides high sensitivity to the top-quark mass—finite-width corrections do not exceed one percent in the range that contains the bulk of the cross section, but become more sizable in the region of highest m_t-sensitivity. This motivates more detailed studies of finite-width effects in the context of high-precision m_t-measurements at the LHC. The results of this investigation of finite-width effects in $t\bar{t}$ production give also useful insights into possible limitations of treating *associated* top-pair production processes in the narrow-width approximation, since NLO calculations for $pp \to WWb\bar{b}j$ and similar reactions will not be available too soon.

ACKNOWLEDGMENTS

We thank K. Melnikov for discussions and comments on the manuscript. S.P. would like to acknowledge financial support from the SNSF. M.S. is grateful for support from the Director's Fellowship of Argonne National Laboratory under DOE grant DE-AC02-06CD11357. The presented results of the calculation described in Ref. [336] were performed on the Homewood High Performance Cluster of Johns Hopkins University supported by grant NSF-OCI-108849.

11. Strong and Smooth Ordering in Antenna Showers [21]

Abstract

We comment on strong and smooth ordering in antenna showers, and extend the definition of smooth ordering to include the case of $g \to q\bar{q}$ splittings. We define three observables in hadronic Z decays that can be used to probe the subleading properties of shower models.

11.1 INTRODUCTION

Traditional parton showers are based on collinear factorization, and the shower evolution proceeds via $1 \to 2$ branchings, on which additional constraints have to be imposed to ensure momentum conservation and QCD coherence (see [344]). Antenna showers are instead based on momentum-conserving and intrinsically coherent $2 \to 3$ branchings, as pioneered by Ariadne [345, 346]. This note concerns the antenna shower implementation in the Vincia code [347], a plug-in to Pythia 8 [348], though we emphasize that the notion of smooth ordering could be applied to other shower types as well.

In leading-logarithmic (LL) antenna showers, the fundamental step is a Lorentz-invariant $2 \to 3$ branching process by which two on-shell "parent" partons are replaced by three on-shell "daughter" partons. This $2 \to 3$ process makes use of three ingredients [349]:

1. An *antenna function* that captures the leading tree-level singularities of QCD matrix elements.
2. An antenna *phase space* — an exact, momentum-conserving and Lorentz-invariant factorization of the pre- and post-branching phase spaces.
3. A *kinematics map*, specifying how the global orientation of the post-branching momenta are related to the pre-branching ones.

Antenna showers come in two varieties: global and sector. The two kinds differ in how the collinear singularities of gluons are partitioned among neighboring antennae, see [350, 351]. Here, we shall only be concerned with the global type [345, 352, 347, 349], in which the gluon-collinear singularity is partitioned such that two neigbouring antennae each contain "half" of it; their sum reproduces the full singularity.

If each antenna in a global shower is allowed to emit in its full phase space, the resulting shower evolution amounts to an incoherent addition of independently radiating dipoles. This tends to overcount regions in which several dipole terms contribute at the same level, i.e., in regions where dipole-dipole interference effects (or, equivalently, multipole effects) are important [353, 351]. The situation is analogous to, though less severe than, the case of traditional parton showers with virtuality-ordering [354], which represent an incoherent addition of independent monopoles. In parton/monopole showers, multiparton interference effects for soft radiation can be taken into account by the requirement of angular ordering [355], while in dipole/antenna showers, typically a measure of transverse momentum is used, such as

$$p_{\perp A}^2 = \frac{s_{ij}s_{jk}}{s_{IK}} \,, \tag{80}$$

for a branching $IK \to ijk$, with $s_{ab} \equiv 2p_a \cdot p_b = (p_a + p_b)^2$ for massless partons. Some alternative possibilities are compared in [349].

[21] Contributed by: J. J. Lopez-Villarejo, P. Skands

11.2 STRONG AND SMOOTH ORDERING

In a strongly-ordered shower, each consecutive branching is required to occur at a lower scale in the evolution variable than that of the previous one: $Q_{n+1} < Q_n$. This can be represented as a step function in the evolution variable, multiplying the branching kernels. In a smoothly-ordered shower [349], the step function is replaced by a smooth dampening factor designed to leave the soft and collinear limits unchanged while suppressing radiation at scales above $\sim Q_n$. Specifically, for evolution in $p_\perp$, we replace the strong-ordering condition as follows,

$$\Theta(\hat{p}_\perp - p_\perp)\, P_{LL} \;\;\rightarrow\;\; P_{imp}\, P_{LL} \;\equiv\; \frac{\hat{p}_\perp^2}{\hat{p}_\perp^2 + p_\perp^2}\, P_{LL}\,, \tag{81}$$

where $\hat{p}_\perp$ characterizes the scale of the previous branching[22], $p_\perp$ is the scale of the emission under consideration, and P_{LL} is an ordinary LL shower kernel, which in our case is represented by a gluon-emission antenna function. (We return to the case of $g \to q\bar{q}$ below.)

Thus, for $p_\perp \ll \hat{p}_\perp$ (the strongly-ordered limit) the smooth-ordering factor P_{imp} tends to unity, while for $p_\perp \sim \hat{p}_\perp$ (the ordering threshold) it tends to $1/2$, and finally for $p_\perp \gg \hat{p}_\perp$ (highly unordered), it tends to zero $\propto \hat{p}_\perp^2/p_\perp^2$. Note that, since P_{LL} is likewise $\propto 1/p_\perp^2$, the net effect of the suppression factor is to modify the behavior of the splitting kernel from $1/p_\perp^2$ in the strongly-ordered limits to $1/p_\perp^4$ for highly unordered branchings, similar to what has been studied for initial-state parton showers in [356]; above the strong-ordering threshold, the branching probability is explicitly suppressed beyond LL.

For a rigourous interpretation of the P_{imp} factor one would have to analyze the $2 \to 4$ antennae [79] and check that the combination of two $2 \to 3$ antennae times this factor does indeed reproduce subleading aspects of the full $2 \to 4$ function. In the absence of such a study, one may still physically interpret its purpose in the following way: the LL antenna functions are derived assuming the outgoing partons/jets to be massless. This is a good approximation if the virtuality that they can acquire (through further showering) is restricted by the strong-ordering threshold. When allowing unordered branchings, however, the corresponding Feynman diagrams contain highly off-shell propagators, which the P_{imp} factor attempts to mimic by introducing an "effective mass" in the denominator of eq. (81).

For gluon emissions, it was shown in [349] that the smooth-ordering condition does lead to a systematic improvement in the shower. Since it simultaneously guarantees a complete phase-space coverage (contrary to the case for strong ordering [357, 349]), it is the default option in Vincia.

Antenna showers including $g \to q\bar{q}$ splittings were studied in [358], in which evolution in $m^2_{q\bar{q}}$ was introduced for such branchings. This is based on the observation [359] that the scale controlling the divergences of $g \to q\bar{q}$ splittings is the invariant mass of the pair, not its $p_\perp$. By analogy with the physical interpretation given to the P_{imp} factor for gluon emissions above, it therefore seems well-motivated to study a "generalized" P_{imp} factor where each scale depends on whether we are dealing with a gluon or a quark:

$$P_{imp} = \frac{\hat{Q}_E^2}{\hat{Q}_E^2 + Q_E^2}, \tag{82}$$

where Q_E is the evolution variable: $p_\perp$ for gluons and invariant mass for quark-antiquark pairs.

We can assess the improvement that this produces in the shower by plotting the ratio of the shower approximation vs. the LO matrix element for $Z \to q\bar{q}'q'\bar{q}$ and $Z \to q\bar{q}'q'g\bar{q}$. This is shown in fig. 28, where the histograms represent the distribution of $\log_{10}(\mathrm{PS}/\mathrm{ME})$ in a flat phase-space scan, normalized to unity (i.e., the same type of distributions that were shown in [349, 358, 351]). Points to the left of zero are undercounted by the shower approximation, while points to the right are overcounted. Although the agreement is by no means perfect, we do observe a slight improvement in the shower approximation

[22] We take $\hat{p}_\perp$ to be the smallest $p_\perp$ scale among all the color-connected partons in the parent configuration, i.e., a global measure of the "current" $p_\perp$ scale of that topology. This makes the shower a true Markov chain (i.e., history-independent) which has beneficial consequences for matching to matrix elements [349].

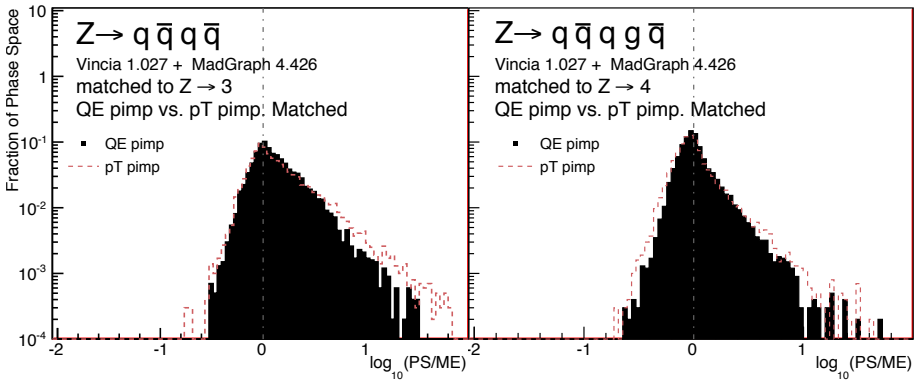


Fig. 28: Comparison between generalized $\tilde{P}_{imp}$ and "old" P_{imp} factor in the global shower approximations to LO matrix elements, for processes involving a $g \to q\bar{q}$ splitting. *Left:* $Z \to q\bar{q}'q'\bar{q}$. *Right:* $Z \to q\bar{q}'q'g\bar{q}$. In both cases, GKS matching to the LO matrix element for the preceding multiplicity ($Z \to 3$ and $Z \to 4$, respectively) has been included, and the Ariadne factor was applied to $g \to q\bar{q}$ splittings.

when the P_{imp} factor is defined in terms of Q_E (solid black histogram), as compared to the definition used previously (dashed histogram). Note that we used the so-called Ariadne factor in the shower approximation for all cases, see [358], and that the distributions were made including GKS matching to the preceding multiplicities [349].

11.3 SENSITIVE OBSERVABLES IN HADRONIC Z DECAYS

The properties of shower and matrix-element matching algorithms are coming under increasing scrutiny, not least due to the desire of achieving reliable descriptions of jet production and jet properties, such as jet substructure, for signal and background estimates at the LHC.

For final-state radiation, i.e., jet broadening and jet splitting, hadronic Z decays are the main reference, with a large set of events shapes and jet resolutions/rates being used to constrain and tune shower algorithms (see, e.g., [360, 344]). However, in the logarithmically dominated regions, these observables are typically dominated by leading logs, and are well described by all coherent and reasonably well-tuned shower algorithms on the market. In order to probe the subleading properties in a more dedicated way, we have found the following three simple observables useful, each designed to isolate a specific aspect.

We consider hadronic Z events (photon ISR is switched off, and matching beyond 3 jets is switched off for the strongly-ordered showers) and use the k_T clustering algorithm [361] to cluster all events back to two jets. The $3 \to 2$ clustering scale is denoted $y_{23} = k_{T3}^2/m_Z^2$, and so on for higher jet numbers. We require all y_{ij} entering in the observables below to be greater than 0.005, to remove contamination from B decays and lower scales. Since the original topology contains two jets, we also keep track of which "side" each clustering happens on. Strong ordering corresponds to $y_{23} \gg y_{34} \gg \ldots$, while events with, e.g., $y_{34} \sim y_{23}$ should be more sensitive to the ordering condition and to the effective $1 \to 3$ spliting kernels.

The first observable is thus simply the ratio y_{34}/y_{23}, in events where the $4 \to 3$ and $3 \to 2$ clusterings happen in the same jet. This distribution is illustrated in the left-hand pane of fig. 29, with logarithmic axes. Vertical error bars indicate the expected 1σ statistical error with 400k hadronic Z decays. Since the k_T algorithm allows for unordered clustering scales, the distribution extends beyond $\xi_{24} = \ln(y_{34}/y_{24}) = 0$. Default Pythia (thick solid line) is compared to three different Vincia settings: smooth (thin solid) and strong (dashed) ordering in $p_\perp$ and strong ordering in dipole virtuality, m_D (dotted). Note here that ordering in the variables $p_\perp$ or m_D does not directly imply ordering in k_T.

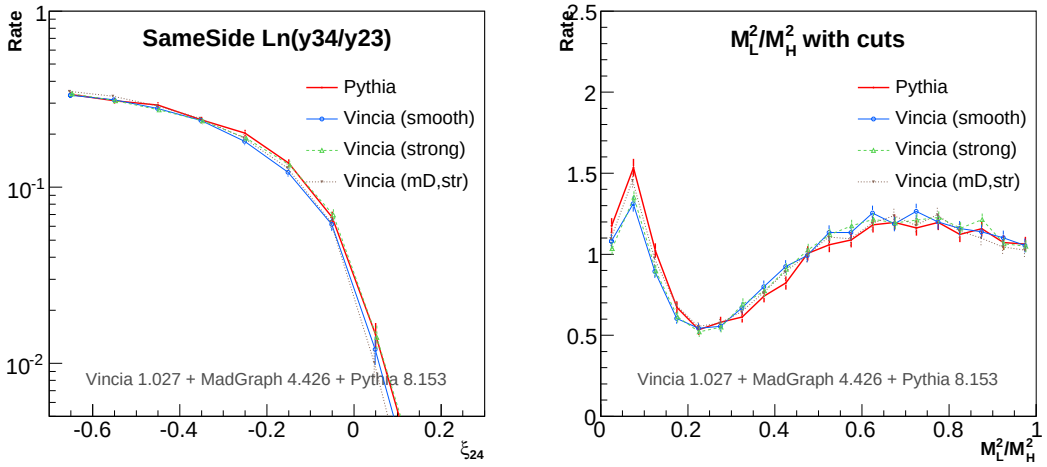


Fig. 29: *Left:* $\xi_{24} = \ln(y_{34}/y_{23})$ in "same-side" 4-jet events. *Right:* Ratio of jet masses, m_L^2/m_H^2, in "compressed" 4-jet events. Error bars indicate expected 1σ statistical errors with 400k hadronic Z decays.

The fact that the P_{imp} factor also suppresses branchings slightly below the strong-ordering threshold is manifest in the thin solid line lying below the other ones in the region just below zero, which should be statistically significant with a sample size of ~ 0.5M events. Note as well that these distributions become indistinguishable if one does not make the requirement of sameside clustering (not shown), presumably since opposite-side collinear splittings then dominate.

A related observable is shown in the right-hand pane of fig. 29. To force a "compressed" scale hierarchy, we impose the cut $y_{34} > 0.5\, y_{23}$, and plot the ratio M_L^2/M_H^2 of the masses of the jets at the end of the clustering. With four partons at LO, the light jet mass is zero if both the $4 \to 3$ and $3 \to 2$ clusterings happen in the same jet, while it is non-zero otherwise. Thus, the region close to zero isolates events with a $1 \to 3$ splitting occurring in one of the jets, while the region above ~ 0.25 is dominated by opposite-side $1 \to 2$ splittings. In Pythia and in mass-ordered Vincia, the peak at zero is stronger than in the $p_\perp$-ordered Vincia cases, while there is no difference between strong and smooth ordering in this variable. It thus serves as a useful complement to ξ_{24}.

Finally, in fig. 30, we consider 4-jet events in which the second and third jets (ordered in energy) are nearly collinear and back-to-back to the hardest jet. Specifically, we impose the cuts $\theta_{12} > 120^\circ$, $\theta_{13} > 120^\circ$, and $\theta_{23} < 30^\circ$. We then plot the angle of the fourth (softest) jet with respect to the hardest one. Again the strong and smooth ordering options are indistinguishable, but interesting differences with respect to both Pythia and mass-ordered Vincia are visible. Mass-ordering tends to produce a broader distribution, with more radiation at right angles to the hardest jet (consistent with mass-ordering prioritizing wide-angle emissions over collinear ones), and the $p_\perp$-ordered Vincia showers exhibit a stronger collinear peak than the Pythia one. A similar observable was proposed in [362].

We conclude that, if all three observables could be measured with an accuracy of $\sim 5 - 10\%$ or better, a useful and multi-dimensional constraint on the subleading shower aspects would be obtained, including sensitivity both to the type and shape of the ordering condition, and to the form of the effective $1 \to 3$ probabilities produced by the shower. We emphasize that we have here restricted our attention to shower models that are virtually indistinguishable on all other observables we have considered.

ACKNOWLEDGEMENTS

We thank D. Kosower, S. Plätzer, and G. Salam for discussions on definitions of sensitive observables. This work was supported by the European Commission (HPRN-CT- 200-00148), FPA2009-09017 (DGI

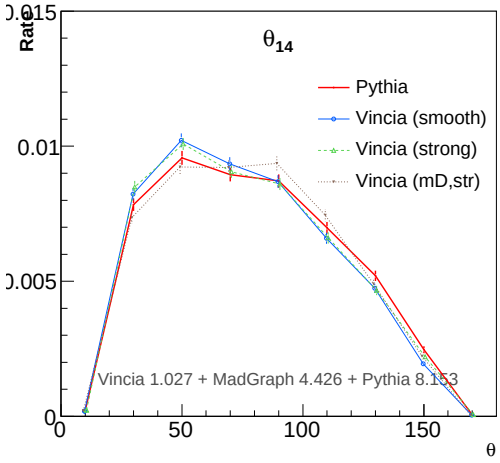


Fig. 30: Angle between the hardest ($1^{\rm st}$) and softest ($4^{\rm th}$) jets in "collinear" 4-jet events. Error bars indicate expected 1σ statistical errors with 400k hadronic Z decays.

del MCyT, Spain) and S2009ESP-1473 (CA Madrid). J.J. L-V is supported by a MEC grant, AP2007-00385, and wants to thank the CERN Theory Division for its hospitality.

12. PERTURBATIVE UNCERTAINTIES AND RESUMMATION FOR EXCLUSIVE JET CROSS SECTIONS [23]

12.1 Introduction

In this writeup we discuss predictions for exclusive jet cross sections, which have a particular number of jets in the final state. There are several motivations for analyzing events by dividing the data into exclusive jet bins, in particular when the relevant backgrounds strongly depend on the number of jets, or when the sensitivity can be increased by optimizing the analysis for the individual jet bins. As our primary example we will consider the Higgs analysis in the $H \to WW$ channel, which is performed separately in exclusive 0-jet, 1-jet, and 2-jet bins [363, 364, 365]. Other examples are vector-boson fusion analyses, which are typically performed in the exclusive 2-jet channel, boosted $H \to b\bar{b}$ analyses that include a veto on additional jets, as well as $H \to \tau\tau$ and $H \to \gamma\gamma$ which benefit from improved sensitivity when the Higgs recoils against a jet. The importance of the Higgs + 1 jet channel in $H \to \tau\tau$ and $H \to WW^*$ was demonstrated explicitly in Refs. [366, 367]. Another motivation for studying exclusive jet bins are the $W+$ jets channels, which are important backgrounds for new physics searches. We will use the notation σ_N for an *exclusive* N-jet cross section (with exactly N jets), and the notation $\sigma_{\geq N}$ for an *inclusive* N-jet cross section (with N or more jets).

To explore the implications of the jet bin restrictions, consider a simple example where we divide the total cross section, $\sigma_{\rm total}$, into an exclusive 0-jet bin, $\sigma_0(p^{\rm cut})$, and the remaining inclusive (≥ 1)-jet bin, $\sigma_{\geq 1}(p^{\rm cut})$,

$$\sigma_{\rm total} = \int_0^{p^{\rm cut}} dp\, \frac{d\sigma}{dp} + \int_{p^{\rm cut}} dp\, \frac{d\sigma}{dp} \equiv \sigma_0(p^{\rm cut}) + \sigma_{\geq 1}(p^{\rm cut})\,. \tag{83}$$

Here p denotes the kinematic variable which is used to divide up the cross section into jet bins. A typical choice is $p \equiv p_T^{\rm jet}$, defined by the largest p_T of any jet in the event, such that $\sigma_0(p_T^{\rm cut})$ only contains events with jets having $p_T \leq p_T^{\rm cut}$, and $\sigma_{\geq 1}(p_T^{\rm cut})$ contains events with at least one jet with $p_T \geq p_T^{\rm cut}$. By defining $\sigma_0(p_T^{\rm cut})$ and $\sigma_{\geq 1}(p_T^{\rm cut})$ one has divided up initial-state radiation from the

[23] Contributed by: Iain W. Stewart, Frank J. Tackmann

colliding hard partons and soft radiation in the event. This restriction on additional emissions changes the coefficients appearing in the α_s expansion and leads to the appearance of double and single logarithms of the form $\alpha_s \ln^2(p^{\rm cut}/Q)$ and $\alpha_s \ln(p^{\rm cut}/Q)$ (with higher powers $\alpha_s^n \ln^{m\leq 2n}(p^{\rm cut}/Q)$ appearing at higher orders in perturbation theory). Here Q is the hard scale of the process, such as $Q = m_H$ for Higgs production, and most often we have $p^{\rm cut} \ll Q$. These changes to the perturbation series can modify the convergence of fixed-order results and make it prudent to consider resummed cross section predictions that include an all-orders resummation of the large logarithms. For N jets the analog of Eq. (83) is $\sigma_{\geq N} = \sigma_N(p_{N+1}^{\rm cut}) + \sigma_{\geq N+1}(p_{N+1}^{\rm cut})$ and the same discussion applies regarding the large logarithms of $p_N^{\rm cut}$ that are not present in $\sigma_{\geq N}$, but are present in each of σ_N and $\sigma_{\geq N+1}$.

The definition of $\sigma_0(p^{\rm cut})$ may include dependence on rapidity and on the grouping of particles. For a jet-based variable like $p_T^{\rm jet}$ the former is induced by only considering jets within the rapidity range $|\eta^{\rm jet}| \leq \eta^{\rm cut}$, and the latter enters through the choice of jet algorithm. These dependencies make theoretical predictions more difficult. In Higgs production via gluon fusion the cross section is known to next-to-next-to-leading order (NNLO) [208, 368, 207, 209, 210, 211, 369, 370], and NNLO results including full kinematic information are available through FeHiP [214, 371] and HNNLO [90, 216] (as well as by combining the total NNLO cross section with MCFM [157, 372] for some distributions). When the measurements are performed in exclusive jet bins, the perturbative uncertainties in the theoretical predictions must also be evaluated separately for each individual jet bin [373]. When combining channels with different jet multiplicities, the correlations between the theoretical uncertainties can be significant and must be taken into account [26]. The perturbative predictions can be made more precise by including a resummation of large $p^{\rm cut}$ dependent logarithms on top of the fixed-order predictions. At the leading logarithmic level this can be achieved with standard parton shower Monte Carlo programs, regardless of the precise definition of $p^{\rm cut}$. So far a next-to-next-to-leading logarithmic (NNLL) resummed result for a jet-veto variable only exists for beam thrust [374], $\mathcal{T}_{\rm cm}$, which is a rapidity weighted E_T-like inclusive variable. The definitions of the jet-veto variables we will use are

$$p_T^{\rm jet} = \Big| \sum_{k\in {\rm jet}} \vec{p}_{Tk} \Big| , \qquad\qquad \mathcal{T}_{\rm cm} = \sum_k |\vec{p}_{Tk}| e^{-|\eta_k|} = \sum_k (E_k - |p_k^z|) \,. \tag{84}$$

For $p_T^{\rm jet}$ our jets are defined using anti-k_T [341] with $R = 0.5$, and we consider jets that satisfy a rapidity cut $|\eta| \leq \eta^{\rm cut}$. For $\mathcal{T}_{\rm cm}$ the sum is over all objects in the final state except the Higgs decay products, and can in principle be considered over particles, topo-clusters, or jets with a small R parameter. In all our results we consistently use MSTW2008 NNLO PDFs [262].

In this writeup we will explore fixed NNLO and resummed NNLL+NNLO predictions for $H+$ 0-jet cross sections and compare various methods for evaluating the uncertainty as a function of cuts on $p_T^{\rm jet}$ and $\mathcal{T}_{\rm cm}$. The three methods we will discuss for evaluating the uncertainties in exclusive jet cross sections are

A) "Direct Exclusive Scale Variation". Here the uncertainties are evaluated by directly varying the renormalization and factorization scales in the fixed-order predictions for each exclusive jet cross section σ_N. This implies that the uncertainties are 100% correlated for different Ns.

B) "Combined Inclusive Scale Variation", as proposed in Ref. [26] and utilized in Refs. [363, 364, 365]. Here, the perturbative uncertainties in the inclusive N-jet cross sections, $\sigma_{\geq N}$, are treated as the primary uncertainties that can be evaluated by scale variations in fixed-order perturbation theory. These uncertainties are treated as uncorrelated for different N. The exclusive N-jet cross sections are obtained using $\sigma_N = \sigma_{\geq N} - \sigma_{\geq N+1}$. The uncertainties and correlations follow from standard error propagation, including the appropriate anticorrelations between σ_N and $\sigma_{N\pm 1}$ related to the division into jet bins.

C) "Uncertainties from Resummation." Resummed calculations for exclusive jet cross sections can provide uncertainty estimates that allow one to simultaneously include both types of correlated

and anticorrelated uncertainties as in methods A and B. The magnitude of the uncertainties may also be reduced from the resummation of large logarithms.

In all three methods, adding the exclusive jet cross sections yields the expected scale variation in the total cross section. Method B avoids a potential underestimate of the uncertainties in individual jet bins due to strong cancellations that can potentially take place in method A. Method B produces realistic perturbative uncertainties for exclusive jet cross sections when using fixed-order predictions for various processes, since it accounts for the presence of large logarithms at higher orders caused by the jet binning. In Method C one utilizes higher-order resummed predictions for the exclusive jet cross sections, which allow one to obtain improved central values and further refined uncertainty estimates.

The basic structure of the large logarithms in the perturbative series is discussed in Sec. 12.2. In Sec. 12.3 we discuss and compare the above three methods to determine the perturbative uncertainties. The work discussed here regarding methods A, B, and C builds on work done in Refs. [375, 26], was initiated at Les Houches, and has also been incorporated in the second Higgs Yellow Book report [376] (Secs. 5.2 and 5.5.) We also review recent work by others that can be found in [376](Sec. 5.3).

Note that here we are only discussing the theoretical uncertainties due to unknown higher-order perturbative corrections, which are commonly estimated using scale variation. Parametric uncertainties, such as PDF choices and $\alpha_s(m_Z)$ uncertainties, must be treated appropriately as common sources for all investigated channels.

12.2 Theoretical Motivation

12.21 Structure of the Perturbative Series

We begin by discussing the structure of the large logarithms in exclusive jet cross sections. For Higgs production from gluon fusion with $p_T^{\rm jet} \le p_T^{\rm cut}$ the leading double logarithms appearing at $\mathcal{O}(\alpha_s)$ are

$$\sigma_0(p_T^{\rm cut}) = \sigma_B\Big(1 - \frac{3\alpha_s}{\pi}\, 2\ln^2\frac{p_T^{\rm cut}}{m_H} + \cdots\Big)\,, \tag{85}$$

where σ_B is the Born (tree-level) cross section.

The total cross section only depends on the hard scale $Q = m_H$, which means by choosing the factorization and renormalization scales $\mu_f \simeq \mu_r \simeq m_H$, the fixed-order expansion does not contain large logarithms and has the structure

$$\sigma_{\rm total} \simeq \sigma_B\big[1 + \alpha_s + \alpha_s^2 + \mathcal{O}(\alpha_s^3)\big]\,. \tag{86}$$

Our expressions for perturbative series such as this one are schematic, showing the scaling of the terms without the coefficient functions. The convolution with the parton distribution functions (PDFs) are also not displayed. For $gg \to H$, the coefficients of this series can be large, corresponding to the well-known large K factors. As usual, varying the scale in $\alpha_s(\mu)$ (and the PDFs) one obtains an estimate of the size of the missing higher-order terms in this series, which we denote by $\Delta_{\rm total}$.

The inclusive 1-jet cross section has the perturbative structure

$$\sigma_{\ge 1}(p^{\rm cut}) \simeq \sigma_B\big[\alpha_s(L^2 + L + 1) + \alpha_s^2(L^4 + L^3 + L^2 + L + 1) + \mathcal{O}(\alpha_s^3 L^6)\big]\,, \tag{87}$$

where the logarithms $L = \ln(p^{\rm cut}/m_H)$. For $p^{\rm cut} \ll m_H$ these logarithms can get large enough to overcome the α_s suppression. In the limit $\alpha_s L^2 \simeq 1$, the fixed-order perturbative expansion breaks down and the logarithmic terms must be resummed to all orders in α_s to obtain a meaningful result. For typical experimental values of $p^{\rm cut}$ fixed-order perturbation theory can still be considered, but the logarithms cause large corrections at each order and dominate the series.

The exclusive 0-jet cross section is equal to the difference between Eqs. (86) and (87), and so has the schematic structure

$$\sigma_0(p^{\rm cut}) = \sigma_{\rm total} - \sigma_{\ge 1}(p^{\rm cut})$$

$$\simeq \sigma_B \Big\{ \big[1 + \alpha_s + \alpha_s^2 + \mathcal{O}(\alpha_s^3)\big] - \big[\alpha_s(L^2 + L + 1) + \alpha_s^2(L^4 + L^3 + L^2 + L + 1) + \mathcal{O}(\alpha_s^3 L^6)\big] \Big\} . \tag{88}$$

In this difference, the large positive corrections in $\sigma_{\rm total}$ partly cancel against the large negative logarithmic corrections in $\sigma_{\geq 1}$. For example, at $\mathcal{O}(\alpha_s)$ there is a value of L for which the α_s terms in Eq. (88) cancel exactly. At this $p^{\rm cut}$ the NLO 0-jet cross section has vanishing scale dependence and is equal to the LO cross section, $\sigma_0(p^{\rm cut}) = \sigma_B$. Due to this cancellation, a standard use of scale variation in $\sigma_0(p^{\rm cut})$ does not actually probe the size of the large logarithms, and does not provide an estimate of $\Delta_{\rm cut}$. This issue impacts the uncertainties in the experimentally relevant region for $p^{\rm cut}$.

For example, for $gg \to H$ (with $\sqrt{s} = 7\,\text{TeV}$, $m_H = 165\,\text{GeV}$, $\mu_f = \mu_r = m_H/2$), one finds [214, 371, 90, 216]

$$\begin{aligned} \sigma_{\rm total} &= (3.32\,\text{pb})\big[1 + 9.5\,\alpha_s + 35\,\alpha_s^2 + \mathcal{O}(\alpha_s^3)\big] , \\ \sigma_{\geq 1}\big(p_T^{\rm jet} \geq 30\,\text{GeV}, |\eta^{\rm jet}| \leq 3.0\big) &= (3.32\,\text{pb})\big[4.7\,\alpha_s + 26\,\alpha_s^2 + \mathcal{O}(\alpha_s^3)\big] . \end{aligned} \tag{89}$$

In $\sigma_{\rm total}$ one can see the impact of the well-known large K factors. (Using instead $\mu_f = \mu_r = m_H$ the $9.5\alpha_s$ and $35\alpha_s^2$ coefficients in $\sigma_{\rm total}$ increase to $11\alpha_s$ and $65\alpha_s^2$.) In $\sigma_{\geq 1}$, one can see the impact of the large logarithms on the perturbative series. Taking their difference to get σ_0, one observes a sizeable numerical cancellation between the two series at each order in α_s.

12.22 Perturbative Series for the Event Fraction

Experimentally the desired quantity which incorporates the jet-veto cut is the exclusive 0-jet event fraction

$$f_0(p^{\rm cut}) = \frac{\sigma_0(p^{\rm cut})}{\sigma_{\rm total}} = 1 - \frac{\sigma_{\geq 1}(p^{\rm cut})}{\sigma_{\rm total}} . \tag{90}$$

One option for treating $f_0(p^{\rm cut})$ is to consider it as a derived quantity, given the basic observables $\{\sigma_0, \sigma_{\rm total}\}$ or $\{\sigma_{\geq 1}, \sigma_{\rm total}\}$. In this approach, which was utilized in Ref. [26] and Ref. [376](Secs. 5.2 and 5.5), one propagates the uncertainties from the σ_is to derive those for $f_0(p^{\rm cut})$. This approach is natural from the perspective of utilizing log-resummed computations for $\sigma_0(p^{\rm cut})$. In particular, it maintains the constraint that for large $p^{\rm cut}$ we have monotonic convergence of $\sigma_0 \to \sigma_{\rm total}$ and $f_0 \to 1$, a property that relies on a phase space cut reducing the cross section, but does not depend on perturbation theory.

When using fixed-order predictions for the various cross sections, an alternative to Eq. (90) considered in Ref. [376](Sec. 5.3) is to analyze the perturbation theory for $f_0(p^{\rm cut})$ directly. In this case different schemes of organizing the perturbation series, by keeping or dropping various $\mathcal{O}(\alpha_s^3)$ terms, give a method to estimate the size of the higher-order perturbative corrections. Three such schemes were considered in Ref. [376](Sec. 5.3) (which we label here by schemes 1,2,3). It is convenient to define the perturbative corrections to the cross section by dividing each of them by the Born cross section σ_B, such that we can write

$$\begin{aligned} \sigma_{\rm total} &= \sigma_B\big[1 + \hat{\sigma}^{(1)}_{\rm total} + \hat{\sigma}^{(2)}_{\rm total} + \mathcal{O}(\alpha_s^3)\big] , \\ \sigma_{\geq 1}(p^{\rm cut}) &= \sigma_B\big[\hat{\sigma}^{(1)}_{\geq 1}(p^{\rm cut}) + \hat{\sigma}^{(2)}_{\geq 1}(p^{\rm cut}) + \mathcal{O}(\alpha_s^3)\big] . \end{aligned} \tag{91}$$

With this notation the result of treating f_0 as a derived quantity is

$$\big[f_0(p^{\rm cut})\big]^{(\text{scheme 1})} = 1 - \frac{\hat{\sigma}^{(1)}_{\geq 1}(p^{\rm cut}) + \hat{\sigma}^{(2)}_{\geq 1}(p^{\rm cut})}{1 + \hat{\sigma}^{(1)}_{\rm total} + \hat{\sigma}^{(2)}_{\rm total}} + \mathcal{O}(\alpha_s^3) , \tag{92}$$

while at the same order in perturbation theory we can also consider the following expressions for f_0:

$$\begin{aligned}
\left[f_0(p^{\rm cut})\right]^{(\text{scheme 2})} &= 1 - \frac{\hat\sigma^{(1)}_{\geq 1}(p^{\rm cut}) + \hat\sigma^{(2)}_{\geq 1}(p^{\rm cut})}{1+\hat\sigma^{(1)}_{\rm total}} + \mathcal{O}(\alpha_s^3)\,, \\
\left[f_0(p^{\rm cut})\right]^{(\text{scheme 3})} &= 1 - \left[\hat\sigma^{(1)}_{\geq 1}(p^{\rm cut}) + \hat\sigma^{(2)}_{\geq 1}(p^{\rm cut})\right] + \hat\sigma^{(1)}_{\geq 1}\,\hat\sigma^{(1)}_{\rm total} + \mathcal{O}(\alpha_s^3)\,.
\end{aligned} \tag{93}$$

We will contrast using the expressions in Eq. (92) and Eq. (93) with various methods for analyzing the uncertainty in our discussion below.

12.3 Uncertainty Analysis for Exclusive Jet Bins

As described in Sec. 12.21, the phase space restriction defining σ_0 changes its perturbative structure compared to that of $\sigma_{\rm total}$. In general this gives rise to an additional perturbative uncertainty due to missing higher-order terms depending on $p^{\rm cut}$. We will call the associated jet-binning uncertainty $\Delta_{\rm cut}$. This can be thought of as an uncertainty related to the presence of large logarithms of $p^{\rm cut}$ at higher orders in perturbation theory. In Eq. (83) both σ_0 and $\sigma_{\geq 1}$ depend on the phase space cut, $p^{\rm cut}$, and by construction this dependence cancels in $\sigma_0 + \sigma_{\geq 1}$. Hence, the additional uncertainty $\Delta_{\rm cut}$ induced by $p^{\rm cut}$ must be 100% anticorrelated between $\sigma_0(p^{\rm cut})$ and $\sigma_{\geq 1}(p^{\rm cut})$, such that it cancels in their sum. For example, using a covariance matrix to model the uncertainties and correlations, the contribution of $\Delta_{\rm cut}$ to the covariance matrix for $\{\sigma_0, \sigma_{\geq 1}\}$ must be of the form

$$C_{\rm cut} = \begin{pmatrix} \Delta^2_{\rm cut} & -\Delta^2_{\rm cut} \\ -\Delta^2_{\rm cut} & \Delta^2_{\rm cut} \end{pmatrix}\,. \tag{94}$$

The questions then are: (1) How can we estimate $\Delta_{\rm cut}$ in a simple way, and (2) how is the perturbative uncertainty $\Delta_{\rm total}$ of $\sigma_{\rm total}$ related to the uncertainties of σ_0 and $\sigma_{\geq 1}$?

12.31 Perturbative Uncertainties for Method A

When using method A to estimate the perturbative uncertainties one simply uses a common scale variation to estimate the uncertainty Δ_0 in σ_0 and the uncertainty $\Delta_{\geq 1}$ in $\sigma_{\geq 1}$. By doing so the uncertainties are 100% correlated, corresponding to a covariance matrix in method A for $\{\sigma_0, \sigma_{\geq 1}\}$ given by

$$C_A = \begin{pmatrix} \Delta^2_0 & \Delta_0\,\Delta_{\geq 1} \\ \Delta_0\,\Delta_{\geq 1} & \Delta^2_{\geq 1} \end{pmatrix}\,. \tag{95}$$

Here $\Delta_{\rm total} = \Delta_0 + \Delta_{\geq 1}$ is the scale uncertainty in $\sigma_{\rm total}$. When instead of σ_0 we directly calculate the 0-jet event fraction f_0 using Eq. (92) or one of the expressions in Eq. (93), we can again determine the method A uncertainty estimate by scale variation in f_0 (we will refer to these results as methods A_1, A_2, and A_3 respectively).

In this method $\Delta_{\rm cut}$ is not included because, as explained below Eq. (88), varying the perturbative scale in Δ_0 does not probe the presence of the higher order large logarithms depending on $p^{\rm cut}$. This method can lead to an underestimate of the perturbative uncertainty in σ_0 (and hence f_0), since there is a region of $p^{\rm cut}$ values where scale variation is no longer a reasonable estimate of higher order corrections because of the vanishing of the μ dependence.

12.32 Perturbative Uncertainties for Method B

Since the perturbative series for $\sigma_{\geq 1}$ in Eq. (87) is dominated by the large logarithms of $p^{\rm cut}$, we can use its scale variation $\Delta_{\geq 1}$ to get an estimate for their size by taking $\Delta_{\rm cut} = \Delta_{\geq 1}$ [26]. Since $\Delta_{\rm cut}$ and $\Delta_{\rm total}$ are by definition uncorrelated, by setting $\Delta_{\rm cut} = \Delta_{\geq 1}$ we are effectively treating the perturbative

series for $\sigma_{\rm total}$ and $\sigma_{\geq 1}$ as independent with uncorrelated perturbative uncertainties. That is, considering $\{\sigma_{\rm total}, \sigma_{\geq 1}\}$, the covariance matrix is diagonal,

$$\begin{pmatrix} \Delta^2_{\rm total} & 0 \\ 0 & \Delta^2_{\geq 1} \end{pmatrix}, \tag{96}$$

where $\Delta_{\rm total}$ and $\Delta_{\geq 1}$ are evaluated by separate scale variations in the fixed-order predictions for $\sigma_{\rm total}$ and $\sigma_{\geq 1}$. This is consistent, since for small $p^{\rm cut}$ the two series have very different structures. In particular, there is no reason to believe that the same cancellations in σ_0 will persist at every order in perturbation theory at a given $p^{\rm cut}$. It follows that the perturbative uncertainty in $\sigma_0 = \sigma_{\rm total} - \sigma_{\geq 1}$ is given by $\Delta^2_{\rm total} + \Delta^2_{\geq 1}$, and the resulting covariance matrix for $\{\sigma_0, \sigma_{\geq 1}\}$ in method B is

$$C_B = \begin{pmatrix} \Delta^2_{\geq 1} + \Delta^2_{\rm total} & -\Delta^2_{\geq 1} \\ -\Delta^2_{\geq 1} & \Delta^2_{\geq 1} \end{pmatrix}. \tag{97}$$

Note that all of $\Delta_{\rm total}$ occurs in the uncertainty for σ_0. This is reasonable from the point of view that σ_0 starts at the same order in α_s as $\sigma_{\rm total}$ and contains the same leading virtual corrections. The method B uncertainty for the event fraction f_0 follows most naturally by error propagation from the cross sections, treating it as a derived quantity.

The limit $\Delta_{\rm cut} = \Delta_{\geq 1}$ that Eq. (97) is based on is of course an approximation. However, the preceding arguments show that it is a more reasonable starting point than method A, since the latter does not account for the additional $p^{\rm cut}$ induced uncertainties.

The generalization of the above discussion to more jets and several jet bins is straightforward. For the N-jet bin we replace $\sigma_{\rm total} \to \sigma_{\geq N}$, $\sigma_0 \to \sigma_N$, and $\sigma_{\geq 1} \to \sigma_{\geq N+1}$. If the perturbative series for $\sigma_{\geq N}$ exhibits large α_s corrections due to its logarithmic series or otherwise, then the presence of a different series of large logarithms in $\sigma_{\geq N+1}$ will again lead to cancellations when we consider the difference $\sigma_N = \sigma_{\geq N} - \sigma_{\geq N+1}$. These two cross sections will have different series for their double logarithms since the number of active partons and their color structure differ. In this situation $\Delta_{\geq N+1}$ will again give a better estimate for the extra $\Delta_{\rm cut}$ type uncertainty that arises from separating $\sigma_{\geq N}$ into σ_N and $\sigma_{\geq N+1}$.

12.33 Perturbative Uncertainties for Method C

In method C we assess the perturbative uncertainties using resummed predictions for variables $p^{\rm cut}$ that implement a jet veto, following Refs. [375, 26]. An advantage of using resummed predictions is that they contain perturbation theory scale parameters which allow for an evaluation of two components of the theory error, one which is 100% correlated with the total cross section (as in method A), and one related to the presence of the jet-bin cut which is anti-correlated between neighboring jet bins (as in method B).

The resummed H + 0-jet cross section predictions of Ref. [375] follow from a factorization theorem for the 0-jet cross section [374], $\sigma_0(\mathcal{T}^{\rm cut}_{\rm cm}) = H\,\mathcal{I}_{gi}\,\mathcal{I}_{gj} \otimes S f_i f_j$, where H contains hard virtual effects, the $\mathcal{I}$s and S describe the veto-restricted collinear and soft radiation, and the fs are standard parton distributions. Fixed-order perturbation theory is carried out at three scales, a hard scale $\mu_H^2 \sim m_H^2$ in H, and beam and soft scales $\mu_B^2 \sim m_H \mathcal{T}^{\rm cut}_{\rm cm}$ and $\mu_S^2 \sim (\mathcal{T}^{\rm cut}_{\rm cm})^2$ for $\mathcal{I}$ and S, and are then connected by NNLL renormalization group evolution that sums the jet-veto logarithms, which are encoded in ratios of these scales. The perturbative uncertainties can be assessed by considering two sources: i) an overall scale variation that simultaneously varies $\{\mu_H, \mu_B, \mu_S\}$ up and down by a factor of two which we denote by Δ_{H0}, and ii) individual variations of μ_B or μ_S that each hold the other two scales fixed [375], whose envelope we denote by the uncertainty Δ_{SB}. Here Δ_{H0} is dominated by the same sources of uncertainty as the total cross section $\sigma_{\rm total}$, and hence should be considered 100% correlated with its uncertainty

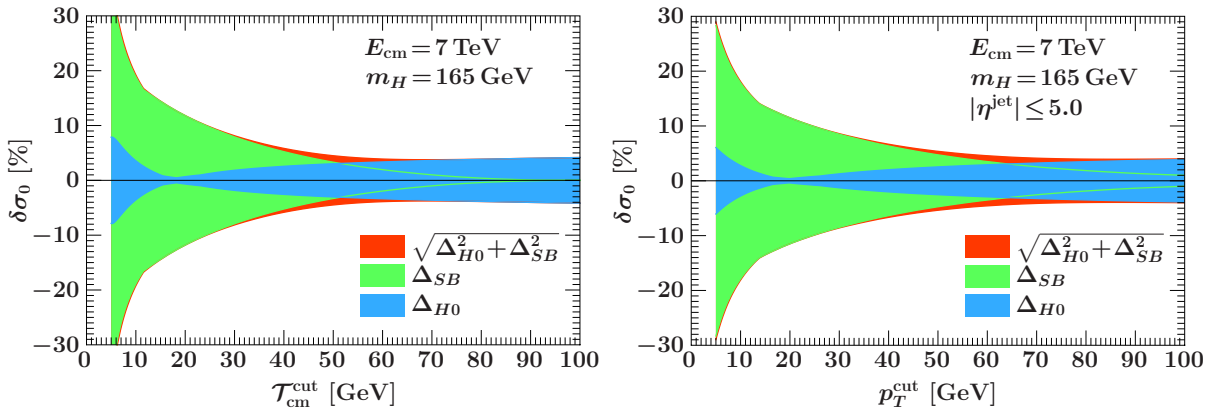


Fig. 31: Relative uncertainties for the 0-jet bin cross section from resummation at NNLL+NNLO for beam thrust $\mathcal{T}_{\rm cm}$ on the left and $p_T^{\rm jet}$ on the right.

$\Delta_{\rm total}$. The uncertainty Δ_{SB} is only present due to the jet-bin cut, and hence gives the $\Delta_{\rm cut}$ uncertainty that is anti-correlated between neighboring jet bins.

If we simultaneously consider the cross sections $\{\sigma_0, \sigma_{\geq 1}\}$ then the full correlation matrix in method C is

$$C_C = \begin{pmatrix} \Delta_{SB}^2 & -\Delta_{SB}^2 \\ -\Delta_{SB}^2 & \Delta_{SB}^2 \end{pmatrix} + \begin{pmatrix} \Delta_{H0}^2 & \Delta_{H0}\,\Delta_{H\geq 1} \\ \Delta_{H0}\,\Delta_{H\geq 1} & \Delta_{H\geq 1}^2 \end{pmatrix}, \tag{98}$$

where $\Delta_{H\geq 1} = \Delta_{\rm total} - \Delta_{H0}$ encodes the 100% correlated component of the uncertainty for the (≥ 1)-jet inclusive cross section. Computing the uncertainty in $\sigma_{\rm total}$ gives back $\Delta_{\rm total}$.

Eq. (98) can be compared to C_A for method A in Eq. (95), which corresponds to taking $\Delta_{SB} \to 0$ and obtaining the analog of Δ_{H0} by up/down scale variation without resummation ($\mu_H = \mu_B = \mu_S$). It can also be compared to C_B for method B in Eq. (97), which corresponds to taking $\Delta_{SB} \to \Delta_{\geq 1}$ and $\Delta_{H\geq 1} \to 0$, such that $\Delta_{H0} \to \Delta_{\rm total}$. The numerical dominance of Δ_{SB}^2 over $\Delta_{H0}\Delta_{H\geq 1}$ in the 0-jet region is another way to justify the preference for using method B when only given a choice between methods A and B. For example, for $p_T^{\rm cut} = 30\,{\rm GeV}$ and $|\eta^{\rm jet}| \leq 5.0$ we have $\Delta_{SB}^2 = 0.17$ and $\Delta_{H0}\Delta_{H\geq 1} = 0.02$.

In Fig. 31 we show the uncertainties Δ_{SB} (light green) and Δ_{H0} (medium blue) as a function of the jet-veto variable, as well as the combined uncertainty adding these components in quadrature (dark orange). From the figure we see that the Δ_{H0} dominates at large values where the veto is turned off and we approach the total cross section, and that the jet-cut uncertainty Δ_{SB} dominates for the small cut values that are typical of experimental analyses with Higgs jet bins. The same pattern is observed in the left panel which directly uses the NNLL+NNLO predictions for $\mathcal{T}_{\rm cm}^{\rm cut}$, and the right panel which shows the result from reweighting these predictions to $p_T^{\rm cut}$ as explained in Sec. 12.34 below.

12.34 Comparison of Uncertainty Methods

In Fig. 32 we compare the uncertainties for the 0-jet bin cross section from methods A (medium green), B (light green), and C (dark orange). In the upper panels we use $\mathcal{T}_{\rm cm}^{\rm cut}$ as the jet-veto variable and full results for the NNLO and NNLL+NNLO cross sections, while in the lower panels we use $p_T^{\rm cut}$ as the jet-veto variable with the full NNLO and the reweighted NNLL+NNLO results (as explained below). The upper panels use a cut on beam thrust, $\mathcal{T}_{\rm cm}^{\rm cut}$ while the lower panels use $p_T^{\rm cut}$. The right panels show the same results as those on the left, but are normalized to the highest-order result to better show the relative differences and uncertainties. The uncertainties in methods A, B, and C are computed from the upper left entry of the matrices C_A, C_B, and C_C, respectively.

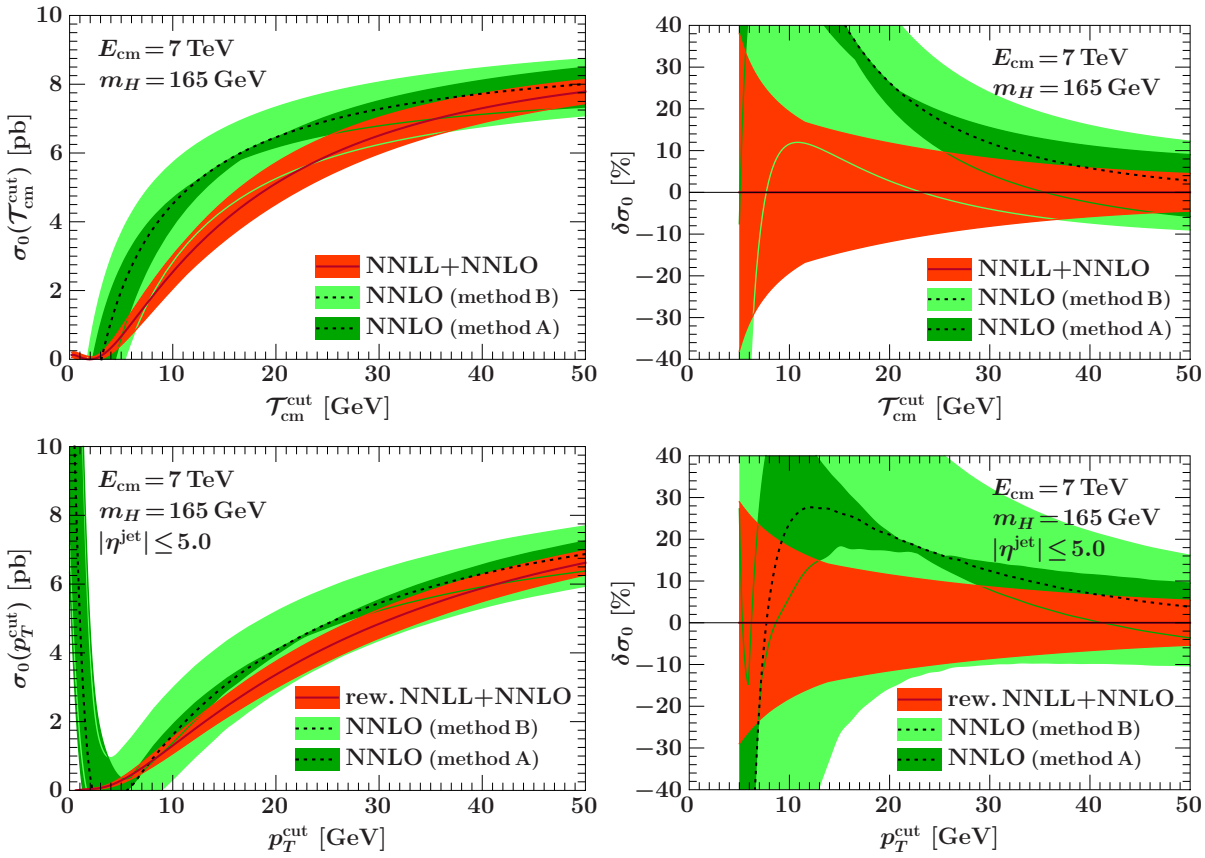


Fig. 32: Comparison of uncertainties for methods A, B, C for the 0-jet bin cross section for beam thrust $\mathcal{T}_{\rm cm}$ (top) and $p_T^{\rm jet}$ (bottom). Results are shown at NNLO with uncertainties from methods A and B and for the NNLL+NNLO resummed result using method C (reweighted for $p_T^{\rm cut}$). On the right all curves are normalized relative to the NNLL+NNLO central value.

From Fig. 32 we see that in method A (medium green bands) for small values of $p_T^{\rm cut}$ the cancellations that take place in $\sigma_0(p^{\rm cut})$ cause the error bands to shrink and eventually almost vanish at $p_T^{\rm cut} \simeq 25\,\text{GeV}$, where there is an almost exact cancellation between the two series in Eq. (88). This is avoided by using method B (light green bands). For large values of $p_T^{\rm cut}$ method B reproduces the method A scale variation, since $\sigma_{\geq 1}(p^{\rm cut})$ becomes small. On the other hand, for small values of $p_T^{\rm cut}$ the uncertainties estimated using method B are more realistic, because they explicitly estimate the uncertainties due to the presence of higher order large logarithmic corrections.

The features of this plot are quite generic. In particular, the same pattern of uncertainties is observed for the Tevatron, when using $\mu = m_H$ as our central scale (with $\mu = 2m_H$ and $\mu = m_H/2$ for the range of scale variation), whether or not we only look at jets at central rapidities, or when considering the exclusive 1-jet cross section. We also note that using independent variations for μ_f and μ_r does not change this picture, in particular the μ_f variation for fixed μ_r is quite small.

For method C with $\mathcal{T}_{\rm cm}$ we make use of resummed predictions for $H + 0$ jets from gluon fusion at next-to-next-to-leading logarithmic order (NNLL+NNLO) from Ref. [375]. This includes the correct NNLO fixed-order corrections for $\sigma_0(\mathcal{T}_{\rm cm}^{\rm cut})$ for any cut. The resulting cross section $\sigma_0(\mathcal{T}_{\rm cm}^{\rm cut})$ has the jet veto implemented by a cut $\mathcal{T}_{\rm cm} \leq \mathcal{T}_{\rm cm}^{\rm cut}$. This cross section contains a resummation of large logarithms at two orders beyond standard LL parton shower programs. A similar resummation for the case of $p_T^{\rm jet}$ is not available. Instead, we use MC@NLO and reweight it to the resummed predictions in $\mathcal{T}_{\rm cm}$, doing so for both the central curve as well as each of the six scale variation curves needed for the uncertainty

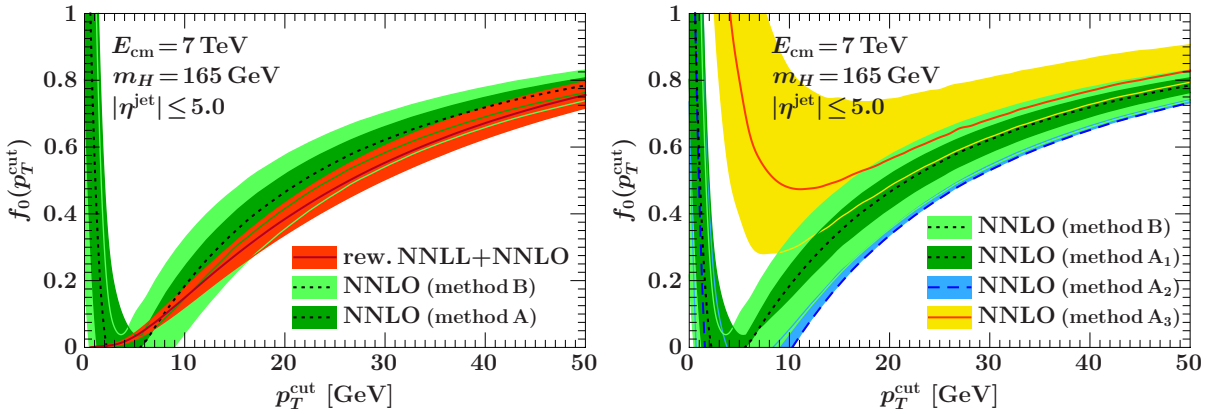


Fig. 33: In the left panel we show the same three curves as in the bottom-left panel of Fig. 32, but for the event fraction $f_0(p_T^{\rm cut})$ treated as a derived quantity from the jet-bin cross sections. In the right panel we contrast the uncertainties obtained using Eqs. (92) and (93) together with method A, with the uncertainty obtained using method B.

determination in method C.[24] We then use the reweighted Monte Carlo sample to obtain cross section predictions for the standard jet veto, $\sigma_0(p_T^{\rm cut})$. We will refer to this as the reweighted NNLL+NNLO result. Since the Monte Carlo here is only used to provide a transfer matrix between $\mathcal{T}_{\rm cm}$ and $p_T^{\rm jet}$, and both variables implement a jet veto, one expects that most of the improvements from the higher-order resummation are preserved by the reweighting. However, we caution that this is not equivalent to a complete NNLL+NNLO result for the $p_T^{\rm cut}$ spectrum, since the reweighting may not fully capture effects associated with the choice of jet algorithm and other effects that enter at this order for $p_T^{\rm cut}$. The dependence on the Monte Carlo transfer matrix also introduces an additional uncertainty, which should be studied and is not included in our numerical results. The transfer matrix is obtained at the parton level, without hadronization or underlying event, since we are reweighting a partonic NNLL+NNLO calculation.

From Fig. 32 one observes that the resummation of the large jet-veto logarithms (dark red central curve) lowers the cross section for both $\mathcal{T}_{\rm cm}^{\rm cut}$ and $p_T^{\rm cut}$. Comparing to NNLO for cut values $\gtrsim 25\,{\rm GeV}$ the relative uncertainties in the resummed result of method C (dark orange bands) and the reduction in the resummed central value are similar for both jet-veto variables. Since one expects resummation to decrease the uncertainties, one can also see that the NNLO uncertainties from method B are more consistent with the higher order NNLL+NNLO resummed method C results than those in method A. We observe that the uncertainties in method C are reduced by about a factor of two compared to those in method B. Since the zero-jet bin plays a crucial role in the $H \to WW$ channel for Higgs searches, and these improvements will also be reflected in uncertainties for the one-jet bin, the improved theoretical precision obtained with method C has the potential to be quite important.

In Fig. 33 we show results for the 0-jet event fraction f_0, with $p_T^{\rm cut}$ as the jet-veto variable. In the left panel we compare the uncertainties in $f_0(p_T^{\rm cut})$ that result from propagating the uncertainties from the jet-bin cross sections obtained from methods A (medium green), B (light green), and C (dark orange). The conclusions are analogous to the corresponding cross-section results in the bottom-left panel of Fig. 32, namely that method B provides a better estimate for the perturbative fixed-order uncertainties than method A, and that the higher-order logarithmic summation present in method C leads to a slightly smaller central value together with the decrease to the uncertainty one expects from incorporating the resummation. In the right panel of Fig. 33 we show the results of the different perturbative schemes for f_0 defined in Eq. (92) (middle dark green band) and Eq. (93) (lower narrow blue band and upper wide

[24] We thank Fabian Stöckli for collaboration on this NNLL+NNLO reweighting analysis for $p_T^{\rm cut}$.

	method A	method B	method C
$\delta\sigma_0(p_T^{\text{cut}})$	3%	19%	9%
$\delta\sigma_{\geq 1}(p_T^{\text{cut}})$	19%	19%	14%
$\rho(\sigma_{\text{total}}, \sigma_0)$	1	0.78	0.15
$\rho(\sigma_{\text{total}}, \sigma_{\geq 1})$	1	0	0.65
$\rho(\sigma_0, \sigma_{\geq 1})$	1	-0.63	-0.65
$\delta f_0(p_T^{\text{cut}})$	6%	13%	9%
$\delta f_{\geq 1}(p_T^{\text{cut}})$	10%	21%	11%
$\rho(\sigma_{\text{total}}, f_0)$	-1	0.43	-0.38
$\rho(\sigma_{\text{total}}, f_{\geq 1})$	1	-0.43	0.38

Table 9: Example of relative uncertainties δ and correlations ρ obtained for the LHC at 7 TeV for $p_T^{\text{cut}} = 30\,\text{GeV}$ and $|\eta^{\text{jet}}| \leq 5.0$.

yellow band) each at NNLO and in each case obtaining the uncertainties using method A (direct scale variation) [376](Sec. 5.3). For comparison, the middle light green band shows the uncertainties obtained from method B. The different method A schemes have a wide spread, which demonstrates the large size of the higher-order perturbative corrections in the total and inclusive 1-jet cross sections. The central values of the alternative methods A_2 and A_3 are not covered by the method A_1 uncertainty band, but all three central curves are covered by the larger uncertainty band from method B (except at small p_T^{cut} where scheme 3 starts to diverge earlier than the other schemes). This can be taken as a confirmation that method A tends to underestimate the perturbative uncertainties in the fixed-order results [376](Sec. 5.3), while method B produces more realistic fixed-order uncertainties.

To appreciate the effects of the different methods on the correlation matrix we consider as an example the results for $p_T^{\text{cut}} = 30\,\text{GeV}$ and $|\eta^{\text{jet}}| \leq 5.0$. The inclusive cross sections are $\sigma_{\text{total}} = (8.76 \pm 0.80)\,\text{pb}$ at NNLO, and $\sigma_{\geq 1} = (3.31 \pm 0.64)\,\text{pb}$ at NLO. The relative uncertainties and correlations at these cuts for the three methods are shown in Table 9. The numbers for the cross sections are also translated into the equivalent results for the event fractions, $f_0(p_T^{\text{cut}}) = \sigma_0(p_T^{\text{cut}})/\sigma_{\text{total}}$ and $f_{\geq 1}(p_T^{\text{cut}}) = \sigma_{\geq 1}(p_T^{\text{cut}})/\sigma_{\text{total}}$. Note that method A should not be used due to the lack of a contribution corresponding to Δ_{cut} in this method, and the resulting underestimated $\delta\sigma_0$. In methods B and C we see, as expected, that σ_0 and $\sigma_{\geq 1}$ have a substantial anti-correlation due to the jet-bin boundary they share.

12.4 Conclusion

To summarize, we have discussed the implications of separating LHC cross sections into jet bins, using Higgs production from gluon fusion as a concrete example. The jet binning induces logarithmic dependences on the jet-bin boundary which is important to properly take into account when making predictions and estimating perturbative uncertainties. When using fixed-order predictions only, the additional logarithms at higher orders in perturbation theory caused by the jet binning can be taken into account in the perturbative uncertainty estimate using method B. By resumming the jet-binning logarithms one can obtain improved predictions with reduced (and more sophisticated) uncertainties using method C.

Here we have focused our discussion on σ_0 and $\sigma_{\geq 1}$ and how to take into account the resulting jet-bin boundary. To further separate $\sigma_{\geq 1}$ into a one-jet bin σ_1 and a $\sigma_{\geq 2}$ one can use method B for this boundary by treating $\Delta_{\geq 2}$ as uncorrelated with the total uncertainty for $\sigma_{\geq 1}$ from either methods B or C. Examples of utilizing method B for this jet bin boundary can be found in Ref. [26]. Once it becomes available one can also use a resummed prediction with uncertainties for this boundary with method C.

Acknowledgments

This work was supported in part by the Office of Nuclear Physics of the U.S. Department of Energy under the grant DE-FG02-94ER40818. Preprint: MIT–CTP 4340

13. A NLO BENCHMARK COMPARISON FOR INCLUSIVE JET PRODUCTION AT HADRON COLLIDERS [25]

Abstract

We present a benchmark comparison of two next-to-leading order (NLO) calculations for inclusive jet and jet pair production at hadron colliders. A new version of the NLO code EKS is adapted for computation of differential cross sections and compared to an independent calculation based on the FastNLO code. A percent-level agreement between the two codes is observed for specified settings of computations at typical transverse momenta and rapidities of Tevatron and LHC measurements. We identify theoretical prerequisites for achieving such level of agreement and comment on the stability of NLO calculations with respect to the factorization scale choice.

13.1 INTRODUCTION

Inclusive jet production at hadron colliders provides an excellent opportunity to test perturbative QCD (PQCD) and look for possible new physics beyond the Standard Model (SM) over a wide range of energy scales. Single inclusive jet production in the Tevatron Run-2 has been recently used to determine the QCD coupling constant [377] and constrain parton distribution functions (PDF) in the proton in global QCD analyses by several groups [255, 378, 262, 311]. Jet production data provide constraints on the gluon PDF at large x values, possibly in a combination with small-x quark PDFs, as discussed in Section 7.. Invariant mass distributions of dijets [379], angular distributions [380, 381], and other jet observables at the LHC [382, 293, 292] provide a unprecedented opportunity to extend searches for quark compositeness and new particle resonances toward the highest energies attainable.

In this contribution, we examine agreement between the computer programs that are available for NLO calculations of jet production cross sections. NLO QCD predictions for jet production work remarkably well in a wide kinematical range and across many orders of magnitude of the cross sections. Nonetheless, the latest PDF analysis evaluates many scattering processes up to NNLO in perturbative QCD. Jet production observables are pivotal for constraining the large-x gluon PDF, but remain known to NLO only. We identify and document main factors affecting NLO jet cross sections at a few-percent level of accuracy and compare the numerical results for typical collider kinematics. Differences between the programs used, and choices for the theoretical inputs made, may be responsible for some differences observed between CT10 and other PDFs, as explained below. Such NLO benchmark comparison will be useful for quantifying or reducing the uncertainties on the resulting PDFs and for the future implementation of NNLO and higher-order resummed contributions to the jet cross sections.

From the experimental point of view, jet production has an advantage of very high statistics and a drawback of sizeable systematical errors associated with complexities of jet reconstruction. NLO theoretical uncertainties due to the QCD scale dependence and the fixed-order model for the jet algorithm are comparable to the experimental errors. Control of numerical accuracy involves, in particular, careful tuning of Monte-Carlo integration to handle steeply falling jet cross sections.

An early numerical code (EKS) for the NLO calculation of single-inclusive jet and dijet distributions was developed by S. D. Ellis, Z. Kunszt and D. E. Soper in 1990's [383] based on the subtraction method. Two other widely used numerical programs are NLOJET++ [384, 385] and FastNLO [290, 291].

[25] Contributed by: J. Gao, Z. Liang, P. M. Nadolsky

The latter provides a fast interface to obtain NLO predictions in kinematical bins of already published experimental jet cross sections by interpolating table files produced by NLOJET++. Besides these *fixed-order* calculations, POWHEG combines the NLO jet production cross sections with *leading-logarithm* QCD showering effects [386]. Some phenomenological studies also include partial NNLO contributions to jet cross sections obtained by threshold resummation [387].

The agreement between the above NLO numerical programs is not automatically met, which motivates the present benchmark comparison. The past CTEQ PDF analyses computed NLO jet cross sections using NLO K-factor tables produced by the EKS code, while other PDF analysis groups use FastNLO. Since the CT10 NLO gluon PDF behaves somewhat differently at large x than the gluon PDF from MSTW'08 or other groups [262], one must compare the EKS and FastNLO computations for the same input values to confirm that these programs do not cause the observed disagreement.

Here we show that the results for the Tevatron ($\sqrt{s} = 1.96$ TeV) and LHC ($\sqrt{s} = 7$ TeV) from EKS and FastNLO agree well when the computation parameters are chosen as described in the next section. These settings must be consciously controlled in order to reach acceptable agreement. As a result of this work, the EKS code has been revised to improve its stability and efficiency and to implement output into new differential cross sections [388].

13.2 Theoretical setup and inputs

Several theoretical inputs must be matched exactly between the EKS and FastNLO programs in order to reach the level of agreement shown in the figures below.

- **Jet algorithm.** When calculating the distribution of jet observables, we need to use the same jet algorithms as the ones in the experimental measurements. In this comparison, we utilized the cone-based Midpoint algorithm [389] for the Tevatron observables and cluster-based anti-k_T algorithm [341] for the LHC. The only difference between the Midpoint algorithm and modified Snowmass algorithm [389] used in the original EKS program is that the Midpoint algorithm always starts with the middle point between the two partons' directions as a seed for a new protojet, no matter how large their separation is. In the NLO theoretical calculations for single-jet or dijet production that include at most three final-state partons, the cluster-based k_T [390], anti-k_T, and Cambridge-Aachen (CA) [391] algorithms are equivalent.
- **The recombination scheme** is a procedure for merging two nearby partons into one jet. For example, the energy scheme (4D, based on adding the 4-momentum) or E_T scheme (based on adding the scalar E_T, then averaging over the partons' directions using E_T as the weights) can be employed to find the momentum of the merged jet [392]. Our comparison uses the energy scheme for both the Tevatron and LHC measurements, as it is often used by the recent experiments. Different choices of the recombination scheme can cause differences of up to ten percent in the NLO distributions, as will be shown later. Note that, with the energy scheme, the jet could be massive, which means that the jet's pseudorapidity will not be equal to its rapidity.
- **The jet trigger** imposes acceptance conditions on each jet's p_T or rapidity when deciding if this jet's contribution is included into the jet observable. In NLO calculations of single-inclusive jet distributions, the jet trigger conditions have no influence. In dijet production, they may change the cross sections by small amounts by affecting the selection of two leading jets in some cases. In our dijet calculations we choose $p_T > 40\,\mathrm{GeV}$, $|y| < 3$ for each jet at the Tevatron and $p_T > 30\,\mathrm{GeV}$, $|y| < 3$ at the LHC.
- **Renormalization and factorization scales.** The scale choice is only related to theory and has no correspondence in experiment. It is conventional to choose the renormalization and factorization scales to be of order of the typical transverse momentum p_T of the jet(s): $\mu_R \sim \mu_F \sim p_T$. In contributions with two resolved jets, p_T naturally corresponds to the transverse momentum of either of the final-state jets (which are equal by momentum conservation). More ambiguity

is present in contributions with three resolved jets, when p_T can correspond to the transverse momentum of either of the jets in each event or to a combination of three transverse momenta. *A meaningful comparison must use equivalent definitions of "jet p_T" in the renormalization and factorization scales of both NLO calculations.*

When FastNLO interpolates tables of NLOJET++ cross sections for single inclusive-jet production, it sets μ_R and μ_F proportional to the p_T value at a fixed point in each p_T bin of the experimental data. Given the high precision of the latest PDF analyses, the FastNLO scale convention produces a numerically different result than the scale proportional to the p_T of the leading jet or the average p_T of two leading jets in each event. It depends on the binning of the experimental data and is numerically close to the average p_T in each bin for small enough bins.

In the EKS calculations for single-jet production, we set the scale proportional to p_T of each individual jet in any p_T bin, which means that we repeat the evaluation of the matrix elements with three resolved jets (contributing to three p_T bins) by successively setting $\mu_{R,F}$ to be proportional to the p_T of each jet in the event. Such matrix elements are thus evaluated three times. This event-level scale setting of EKS turns out to be numerically close to the bin-level scale setting of FastNLO if the bin sizes are small. However, a few-percent differences are still observed at the largest rapidities and p_T. For dijet production, FastNLO and EKS choose the μ_R and μ_F scales that are proportional to the average $|p_T| = (|p_{T1}| + |p_{T2}|)/2$ of the two leading jets.

- **Monte-Carlo integration.** Precision calculations for jet production are numerically challenging because of the rapid falloff of the cross sections with the jet's p_T and rapidity, and also because of large numerical cancellations occurring between some $2 \to 2$ and $2 \to 3$ contributions. Both EKS and NLOJET++ evaluate differential cross sections by Monte-Carlo integration, which requires to generate of order 10^9 of sample points to achieve percent-level accuracy for the whole kinematical region. The upgraded EKS code performs the Monte-Carlo integration using the VEGAS method from the CUBA2.1 library [393]. The EKS output is produced in the form of two-dimensional cross sections ($d^2\sigma/(dp_T\,dy)$, $d^2\sigma/(dM_{jj}\,dy)$, ...) and stored in finely binned two-dimensional histograms. Such output is "almost fully differential" in the sense that the finely grained histograms can be rebinned into any set of coarse bins of the given experiment at the stage of the user's final analysis. This format is different from the FastNLO format, which provides the cross sections in coarse bins taken from pre-existing experimental publications.

  The fine binning in EKS is introduced at the stage of Monte-Carlo integration in order to improve convergence and to better handle the NLO cancellations. The Monte Carlo sampling pattern is tuned automatically to ensure that all fine bins are filled with comparable numbers of sample points, regardless of the momentum and scattering angle values associated with each bin. Then we get uniform relative errors on the cross sections in all bins without consuming too much CPU time, and despite the dramatic variation of cross sections across the bins. Finally, EKS includes a module to allow for flexible choices of scales μ_R and μ_F, and another module for calculating differential cross sections of user-provided jet observables.

13.3 RESULTS

Figs. 34-39 compare our representative numerical results with the ones provided by FastNLO for p_T distributions of single jets, invariant mass distributions of dijets, and (in the case of D0 Run-2) angular distributions (χ) of dijets. Kinematical bins of the Tevatron ($\sqrt{s} = 1.96$ TeV) [287, 285, 289, 394] and LHC ($\sqrt{s} = 7$ TeV) [293, 292] measurements, and CTEQ6.6 PDFs [256] were used. The cone sizes R of the jets are indicated in the figures.

Left panels in the figures show ratios of EKS to FastNLO cross sections, $\sigma_{\rm EKS}/\sigma_{\rm FastNLO}$, at the LO (red points) and NLO=LO+NLO correction (blue points), in kinematical bins provided by the experiments. The horizontal axis indicates the ID of each bin, which are arranged in the order of increasing jet rapidity y and then jet's p_T for inclusive jet production, y and then M_{jj} for dijet production, and M_{jj}

then χ for dijet angular dependence. Vertical lines indicate the boundaries of each rapidity interval for single-jet and dijet distributions, and of each dimass interval for the χ distribution. For example, Fig. 34 shows $\sigma_{\rm EKS}/\sigma_{\rm FastNLO}$ in 6 bins of jet rapidity, with bins 1...23 corresponding to the first rapidity bin ($|y| < 0.4$), bins 24...45 corresponding to the second rapidity bin ($0.4 < |y| < 0.8$), and so on.

The left panel includes, from top to bottom, three plots obtained with the renormalization and factorization scales equal to 1/2, 1, and 2 times the center scale. We can see a good overall agreement between EKS and FastNLO both at LO and NLO. The only significant discrepancies are found in the highest p_T bins for both the Tevatron and LHC single inclusive jet production, which may be due to the difference in the scale choices used in EKS and FastNLO. [These differences reduce when going to NLO]. In the EKS single-jet calculation, we use the actual p_T of the partonic jet filled into the bin as the scale input. FastNLO sets the scale according to a fixed p_T value in each experimental bin, which tends to be different from the EKS scale in the highest p_T bins, which have large widths. The same reason causes a small normalization shift in the other p_T bins.

For dijet production, we only observe random fluctuations at highest M_{jj} that are mainly due to numerical integration errors.

In the right panels of Figs. 34-39, we present plots of the NLO K factor from EKS for each distribution, defined as the ratio of the NLO differential cross section to the LO one. The value of the K factor and its stability with respect to the scale choice may provide an indication of the magnitude of yet higher-order corrections.

To minimize the potential effect of higher-order terms, one might opt to choose the renormalization and factorization scales that bring the K factor close to unity in most of the kinematical region. An alternative approach for setting the scale is based on the minimal sensitivity method, which suggests to choose the μ_R and μ_F values (taken to be equal and designated as μ in the following) at the point where the scale dependence of the NLO cross section is the smallest.

In (di)jet production at central rapidities at the Tevatron, both requirements ($K \approx 1$ and $d\sigma_{NLO}(\mu)/d\mu \approx 0$) could be satisfied by choosing $\mu \approx 0.5 p_T$; see, *e.g.*, the appendix in Ref. [395]. For this reason, the scale $p_T/2$ was used in the CT10 study. However, the point of the minimal sensitivity shifts to higher values (close to p_T or even higher) at forward rapidities at the Tevatron or at all rapidities at the LHC. For such higher scales, however, it is hard to satisfy the requirement that K remains close to unity at the same time.

This point is illustrated by our plots of the K factors. At the central rapidities and $\mu_R = \mu_F = 0.5 p_T$ at the Tevatron (the lowest 3 rapidity bins in Figs. 34-37), $K \approx 1$ and is relatively independent of p_T, as seen in the top subpanels. However, with this scale choice the K factor deviates significantly from unity and has strong kinematic dependence if the rapidity and p_T are large. If one chooses the scale that is equal to p_T or even $2p_T$ (the middle and bottom figures), in accord with the minimal sensitivity method for the forward bins, the kinematical dependence of the K factor reduces, but its value increases to 1.3-1.6 in most of the bins.

For CMS kinematics (Figs. 38-39), the K factor has significant kinematical dependence for all central scale choices, however, the choice $\mu_R = \mu_F = p_T$ (the middle subpanels) results in a comparatively flatter K factor that is also closer to unity. We can see that it is hard to find a fixed scale (or a scale of the type $p_T \times$ (a function of y) [383]) that would simultaneously reduce the magnitude of the NLO correction and stabilize its scale dependence and kinematical dependence. The scale $0.5 p_T$ may be slightly more optimal at the Tevatron, and the scale p_T may be slightly better at the LHC. In the absence of a clearly superior scale choice, it may be necessary to vary the scale of jet cross sections in the global fit in order to estimate its effect on the PDF errors.

In Figs. 40 and 41, we plot the ratios of the NLO distributions calculated using different recombination schemes, where σ_{4D} is obtained with the energy scheme, and σ_{E_T} is with the E_T scheme. For single inclusive jet production at both the Tevatron and LHC, σ_{E_T} is larger then σ_{4D}. An opposite trend

is observed in dijet production. Differences of the predictions based on the two schemes are larger with the Midpoint algorithm (used at the Tevatron) than with the anti-k_T algorithm (used at the LHC). In an NLO calculation, the Midpoint algorithm allows a larger maximal angular separation ($2R$) between the two partons forming a jet, compared to the anti-k_T algorithm that only allows the angular separation up to R. This produces the shown kinematical differences between the two schemes.

CONCLUSIONS

Jet production plays an important role at hadron colliders and is a main background process in the bulk of new physics searches. A benchmark comparison of NLO QCD predictions for jet production from different numerical codes can be useful for both the ongoing phenomenological studies and upcoming higher-order calculations. In this work we modify the original EKS program and compare the single-jet and dijet cross sections that it produces with the ones from the FastNLO program. We find a good agreement between two programs, apart from differences of up to 5-10% occuring at the highest jet p_T's and rapidities. We document the exact combination of theoretical settings in EKS that are needed to reproduce the FastNLO results. Based on the EKS calculation, we attempted to identify the choice of the renormalization and factorization scales that could simultaneously reduce the magnitude of NLO K factors and/or scale dependence of the NLO cross section. Since we could not easily find such a scale combination, we propose to vary the factorization and renormalization scales in future (N)NLO PDF fits to better estimate theoretical uncertainties in the resulting PDFs. There is a plan to publish the updated EKS program in the near future [388].

ACKNOWLEDGMENTS

The work on this preprint was supported by the U.S. DOE Early Career Research Reward DE-SC0003870 and by Lightner-Sams Foundation. We thank our collaborators who have contributed to the development of the EKS code, Hung-Liang Lai, Zhao Li, Dave Soper, and C.-P. Yuan. PMN appreciates helpful discussions with J. Huston and CTEQ members. He also benefited from discussing related work with participants of the Workshop "High Energy QCD at the start of the LHC" at the Galileo Galilei Institute of Theoretical Physics (INFN Florence, September 2011). PMN thanks the organizers of this workshop for the financial support and hospitality.

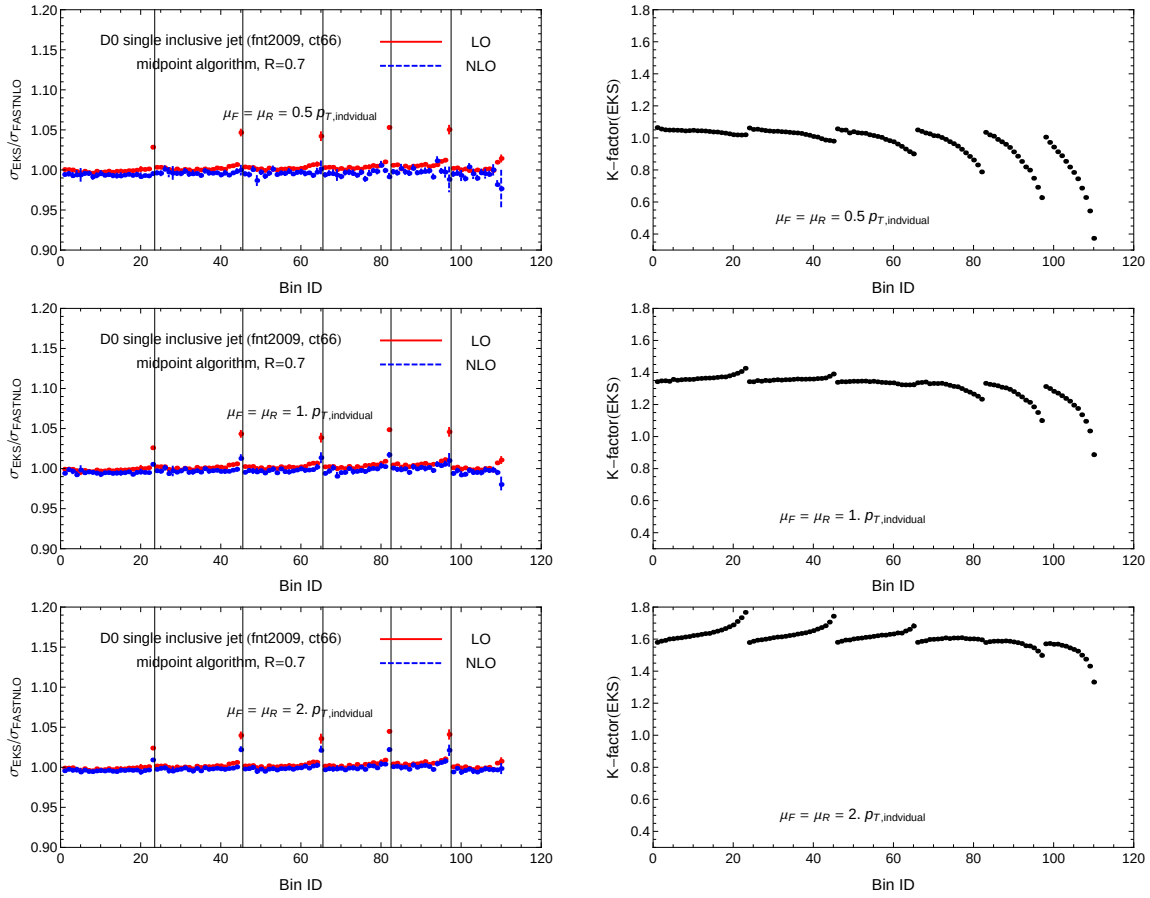


Fig. 34: Comparison of p_T distributions for single inclusive jet production from EKS and FastNLO for D0 Tevatron Run II measurement.[287]

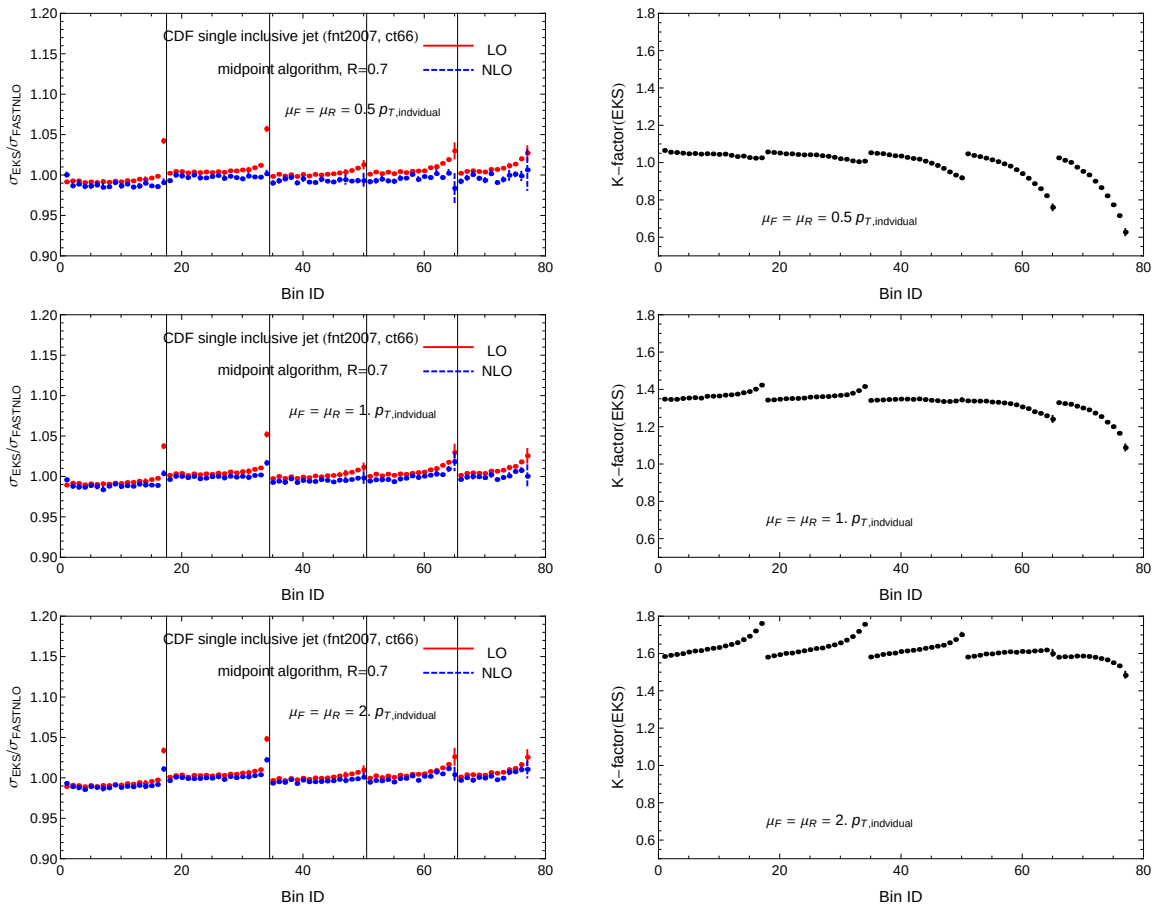


Fig. 35: Comparison of p_T distributions for single inclusive jet production from EKS and FastNLO for CDF Tevatron Run II measurement.[285]

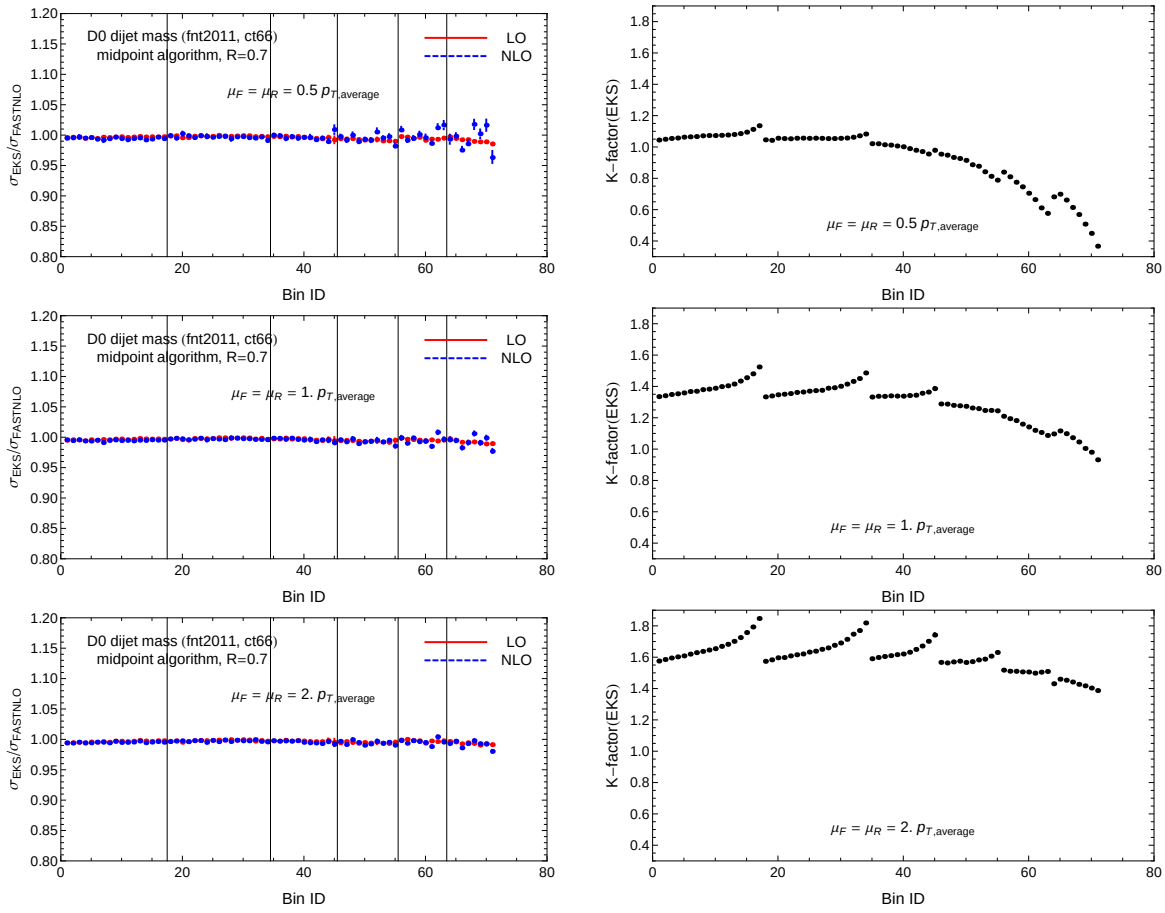


Fig. 36: Comparison of invariant mass distributions for dijet production from EKS and FastNLO for D0 Tevatron Run II measurement.[289]

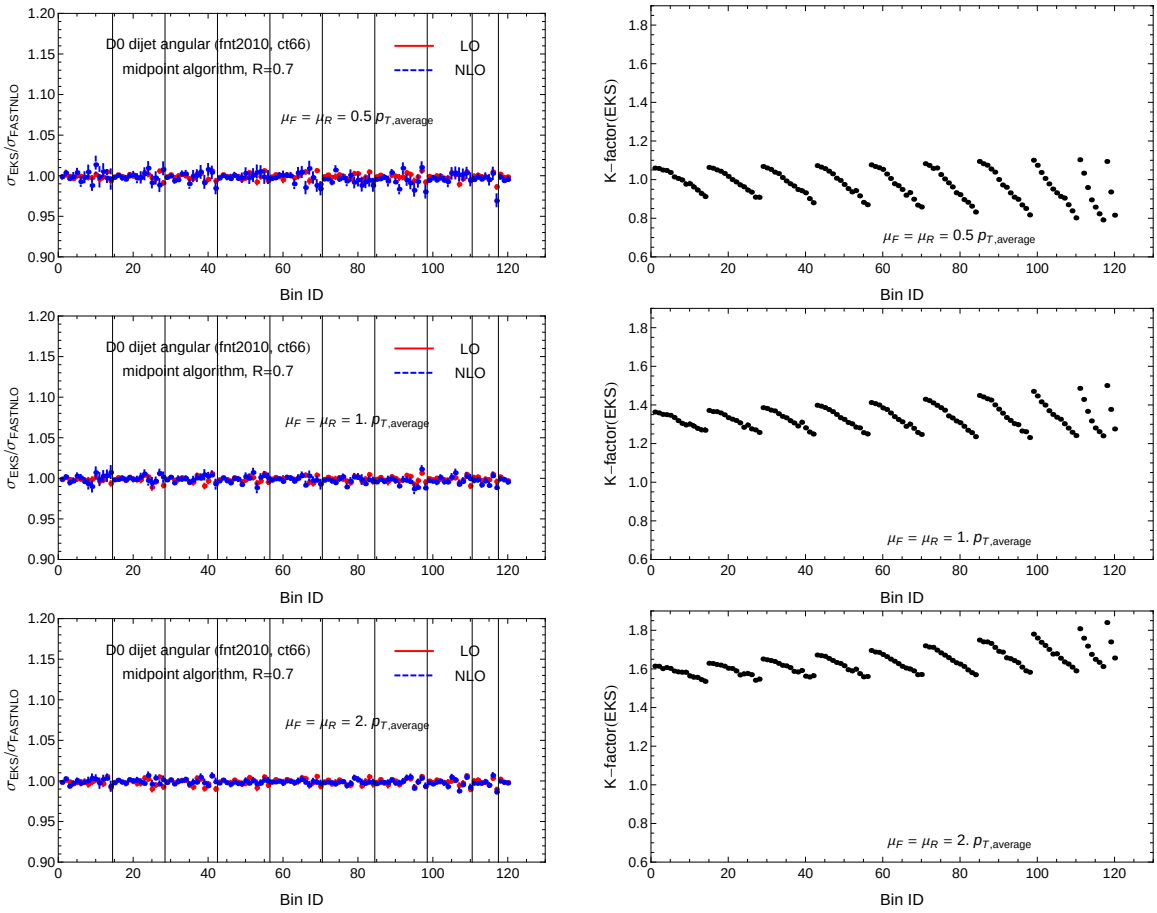


Fig. 37: Comparison of angular (χ) distributions for dijet production from EKS and FastNLO for D0 Tevatron Run II measurement.[394]

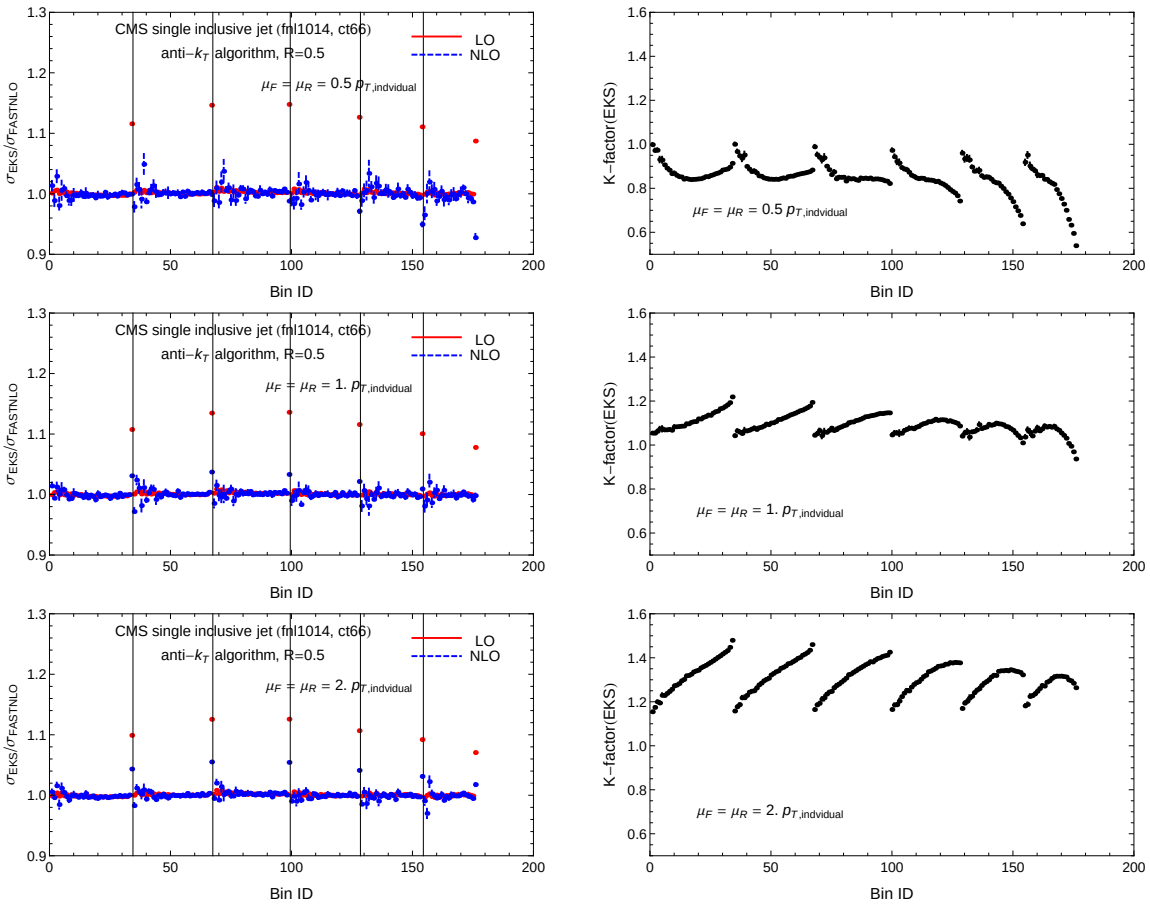


Fig. 38: Comparison of p_T distributions for single inclusive jet production from EKS and FastNLO for CMS LHC (7 TeV) measurement.[292]

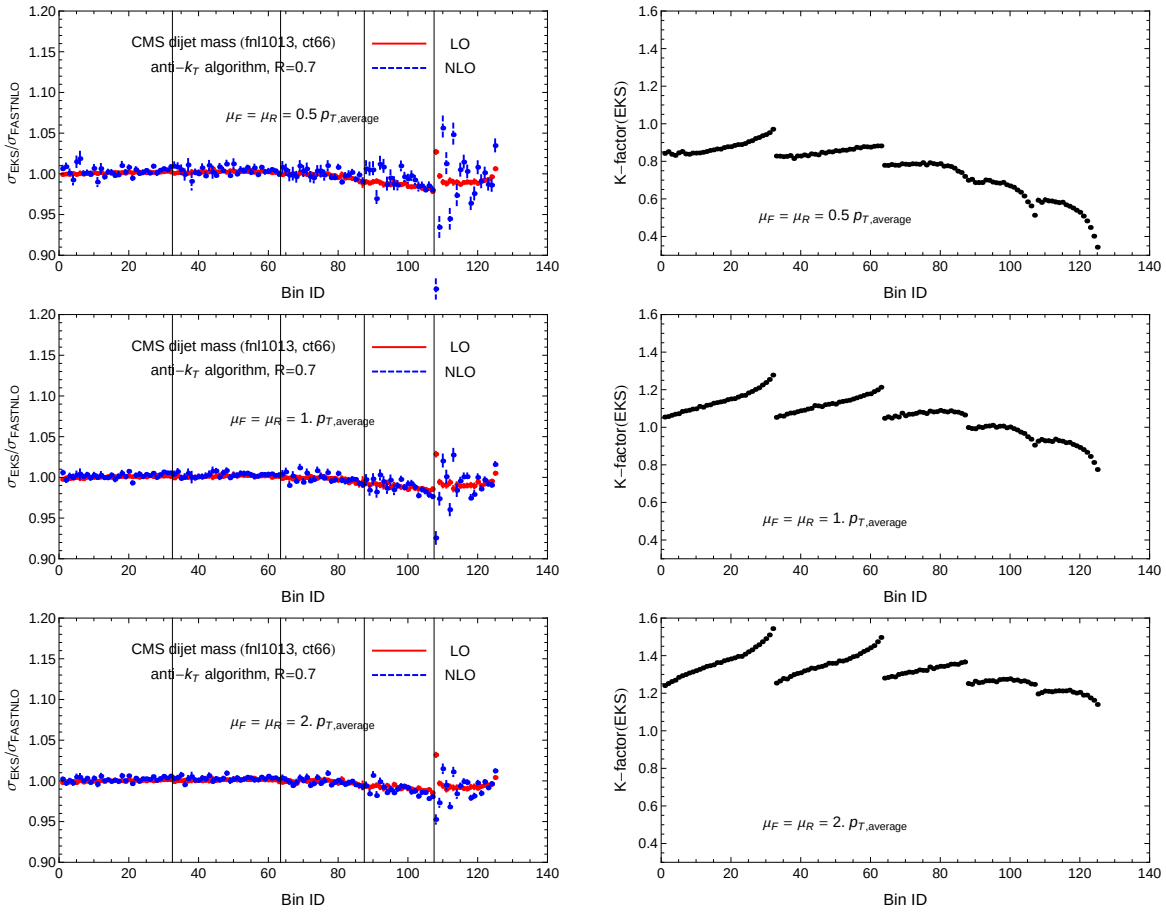


Fig. 39: Comparison of invariant mass distributions for dijet production from EKS and FastNLO for CMS LHC (7 TeV) measurement.[293]

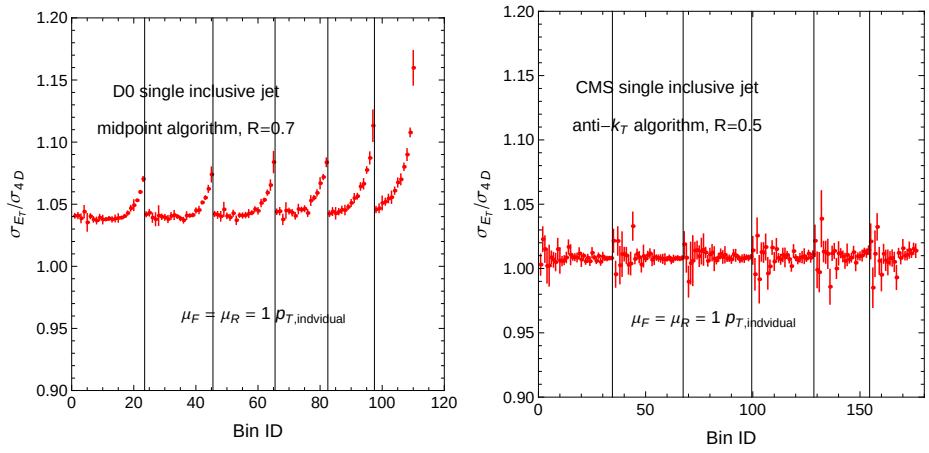


Fig. 40: Comparison of p_T distributions for single inclusive jet production using different recombination schemes.

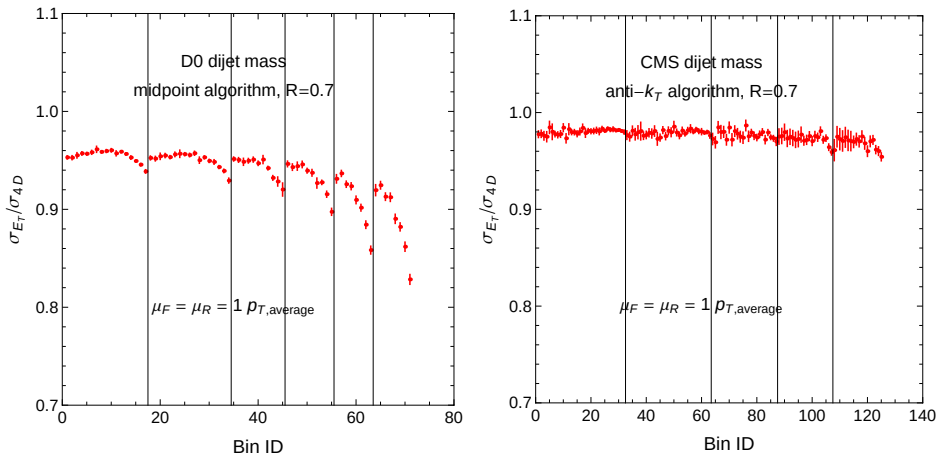


Fig. 41: Comparison of invariant mass distributions for the dijet production using different recombination schemes.

14. PHENOMENOLOGICAL STUDIES WITH AMC@NLO [26]

Abstract

We present four phenomenological studies of hadron collider processes performed within the aMC@NLO framework

14.1 Introduction

aMC@NLO (http://amcatnlo.cern.ch) is a fully automated approach to complete event generation and subsequent parton shower at the NLO accuracy in QCD, which allows accurate and flexible simulations for both signals and backgrounds at hadron colliders. All calculational aspects in aMC@NLO are automated. One-loop contributions are evaluated with MadLoop [8, 396], that uses the OPP integrand reduction method [121] as implemented in CutTools [135]. The other matrix-element contributions to the cross sections, their phase-space subtractions according to the FKS formalism [397], their combinations with the one-loop results, and their integration are performed by MadFKS [398] [27]. The matching of the NLO results with HERWIG [399] or PYTHIA [400] parton showers is performed with the MC@NLO method [401], and it is also completely automatic. Finally, aMC@NLO can compute scale and PDF uncertainties at no extra CPU-time cost with the help of the process-independent reweighting technique described in [402].

For all technical details we refer to the original publications. We report here on the physics results obtained with aMC@NLO for observables of interest at hadron colliders [15, 403, 402, 404]. We stress that they are simulated at the hadron level, namely including parton shower and hadronization effects. In Sects. 14.2, 14.3, and 14.4 we present results for the production of ttH, Vbb, and four-lepton final states at the LHC, respectively. Section 14.5 reports on a study of the Wjj process at Tevatron. Finally, in sect. 14.6 we draw our conclusions. The list of the processes considered here should convince the reader that one can perform realistic analyses of experimental data, for signals and backgrounds, entirely within the aMC@NLO framework.

14.2 The ttH process at the LHC

The production process of a H boson in association with a top pair [15] is a classic mechanism for Higgs production at the LHC [24, 405], where the large ttH Yukawa coupling and the presence of top quarks can be exploited to extract the signal from its QCD multi-jet background. As an example of the use of aMC@NLO for this process we present, in Fig. 42, the Higgs transverse momentum distribution and the transverse momentum of the ttH or ttA system at the $\sqrt{s}$= 7 TeV LHC for a Standard Model (scalar) Higgs with $M_H = 120$ GeV and for a pseudoscalar one with $M_A = 120/40$ GeV. The total NLO cross sections in the three cases are $\sigma_{\rm NLO}(M_H = 120) = 103.4$ fb, $\sigma_{\rm NLO}(M_A = 120) = 31.9$ fb, and $\sigma_{\rm NLO}(M_A = 40) = 77.3$ fb, respectively. At moderate values of the Higgs transverse momentum, the scalar and pseudoscalar cases are clearly distinguishable, while at larger values the three distributions tend to coincide. Parton shower effects give in general small corrections with respect to the a pure NLO calculation, except for variables involving all produced particles, such as the transverse momentum of the ttH or ttA system shown in the right panel of Fig. 42.

14.3 The Vbb process at the LHC

With Vbb we understand $\ell\nu bb$ and $\ell^+\ell^- bb$ final states [403], which are the main backgrounds to searches for SM Higgs production in association with vector bosons (WH/ZH), with the subsequent Higgs decay into a bb pair. The aMC@NLO framework allows a realistic study including

- NLO corrections;

[26] Contributed by: R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, R. Pittau, P. Torrielli

[27] The validation of MadLoop and MadFKS in the context of hadronic collisions has been presented in Ref. [8].

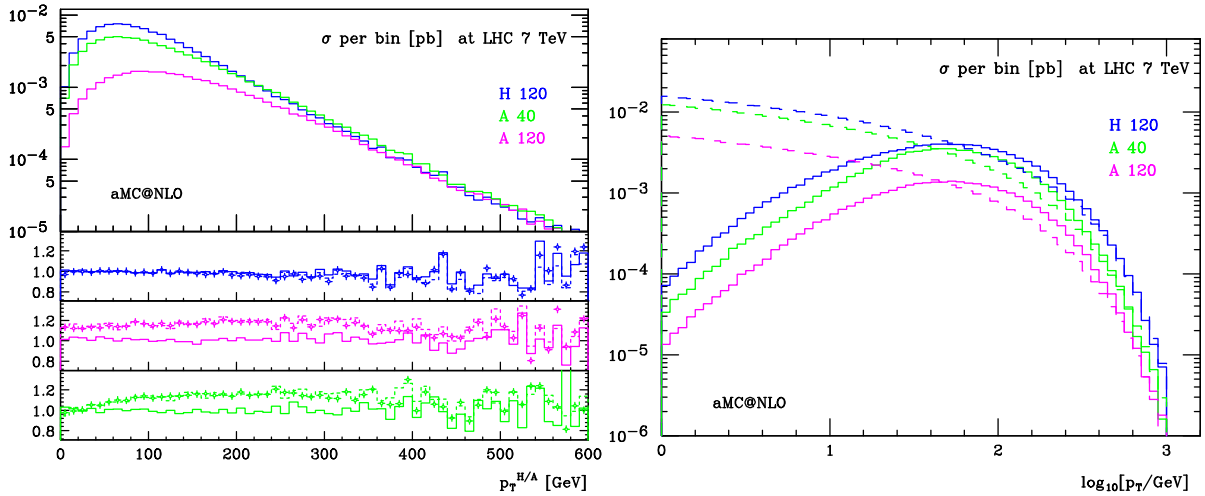


Fig. 42: Higgs transverse momentum distributions (left) and transverse momentum of the ttH or ttA system (right) in ttH/ttA events at the LHC ($\sqrt{s}$=7 TeV), with aMC@NLO in the three cases: Scalar (blue) and pseudoscalar (magenta) Higgs with $m_{H/A} = 120$ GeV and pseudoscalar (green) with $m_A = 40$ GeV. In the lower panels of the left part, the ratios of aMC@NLO over LO (dashed), NLO (solid), and aMC@LO (crosses) are shown. Solid histograms in the right panel are relevant to aMC@NLO, dashed ones to a pure NLO calculation.

- bottom quark mass effects;
- spin-correlation and off-shell effects;
- showering and hadronization.

As an example we show, in Fig. 43, the invariant mass of the pair of the two leading b-jets, compared with the signal distributions for a standard Higgs with $m_H = 120$ GeV. Fig. 43 is interesting because both signal and background are studied at the NLO accuracy. It should be noted that, since completely hadronized events are simulated, sophisticated studies of the jet sub-structure are possible within the aMC@NLO framework, as presented in Fig. 44, where the fractions of events containing zero b-jets, exactly one b-jet, and exactly two b-jets are plotted. The b-jet fractions are fairly similar for Wbb and Zbb production, and the effects of the NLO corrections are consistent with the fully-inclusive K factors. On the other hand, the bb-jet contribution to the b-jet rate is seen to be more than three times larger for $\ell^{\pm}\nu bb$ than for $\ell^+\ell^- bb$ final states. This fact is related to the different mechanisms for the production of a bb pair in the two processes. At variance with the case of $\ell^{\pm}\nu bb$ production, in a $\ell^+\ell^- bb$ final state the two b's may come from the separate branchings of two initial-state gluons, and thus the probability of them ending in the same jet is much smaller than in the case of a $g \to bb$ final-state branching, which gives the only possible contribution to a $\ell^{\pm}\nu bb$ final state.

14.4 Four-lepton production at the LHC

Vector boson pair production is interesting in at least two respects. Firstly, it is an irreducible background to Higgs signals, in particular through the W^+W and ZZ channels which are relevant to searches for a standard model Higgs of mass larger than about 140 GeV. Secondly, di-boson cross sections are quite sensitive to violations of the gauge structure of the Standard Model, and hence are good probes of scenarios where new physics is heavy and not directly accessible at the LHC, yet the couplings in the vector boson sector are affected. We consider here the neutral process [402]

$$pp \to (Z/\gamma^*)(Z/\gamma^*) \to \ell^+\ell^-\ell^{(\prime)+}\ell^{(\prime)-} ,$$

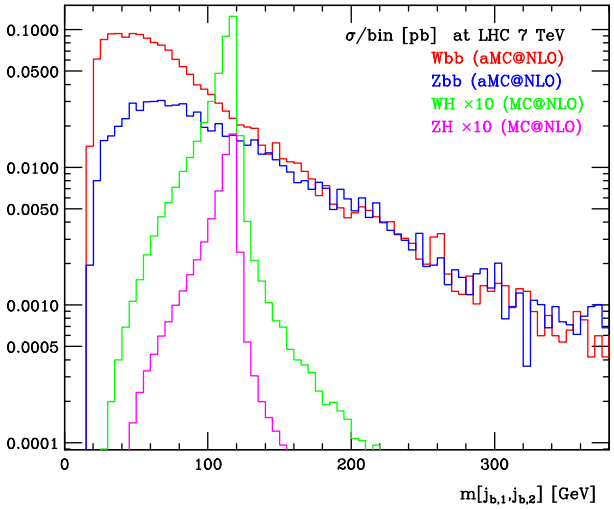


Fig. 43: Invariant mass of the pair of the two leading b-jets. $WH(\to \ell\nu bb)$, $ZH(\to \ell^+\ell^- bb)$, $\ell\nu bb$, and $\ell^+\ell^- bb$ results are shown, with the former two rescaled by a factor of ten.

Process	Cross section (fb)		
	$q\bar{q}/qg$ channels		gg channel
	$\mathcal{O}(\alpha_s^0)$	$\mathcal{O}(\alpha_s^0)+\mathcal{O}(\alpha_s)$	$\mathcal{O}(\alpha_s^2)$
$pp \to e^+e^-\mu^+\mu^-$	9.19	$12.90^{+0.27(2.1\%)+0.26(2.0\%)}_{-0.23(1.8\%)-0.22(1.7\%)}$	$0.566^{+0.162(28.6\%)+0.012(2.1\%)}_{-0.118(20.8\%)-0.014(2.5\%)}$
$pp \to e^+e^-e^+e^-$	4.58	$6.43^{+0.13(2.1\%)+0.11(1.7\%)}_{-0.13(2.0\%)-0.10(1.6\%)}$	

Table 10: Total cross sections for $e^+e^-\mu^+\mu^-$ and $e^+e^-e^+e^-$ production at the LHC ($\sqrt{S} = 7$ TeV) within the cuts $M(\ell^\pm\ell^{(\prime)\mp}) \geq 30$ GeV. The first and second errors affecting the results are the scale and PDF uncertainties (also given as fractions of the central values).

which, although smaller than the W^+W^- channel, may provide a cleaner signal due to the possibility of fully reconstructing the decay products of the two vector bosons. aMC@NLO predictions for the cross sections are given in Tab. 10, which also includes aMC@NLO estimates for scale and PDF uncertainties. The four-lepton invariant mass and the transverse momentum distribution are presented in Fig. 45, where comparisons between the results obtained with aMC@NLO matched to HERWIG and to PYTHIA are also given. We stress that these results include the contributions due to gg-initiated processes, which have also been computed automatically. These are formally of NNLO, but may play a non-negligible phenomenological role owing to their parton-luminosity dominance at a large-energy collider such as the LHC.

14.5 Wjj at Tevatron

In [406] CDF reported an excess of events in two-jet production in association with a W boson, in the form of a broad peak centered at M_{jj} = 144 GeV in the dijet invariant mass [28]. Motivated by this fact, we present in Fig. 46 the aMC@NLO prediction [404] for the dijet invariant mass in Wjj events, using the same cuts as CDF and D0 in the signal region, also comparing with a pure NLO computation and with the ALPGEN [408] findings (one-, two-, and three-parton multiplicities have been consistently matched to

[28] Such an excess has so far failed to be confirmed by a very similar D0 analysis [407].

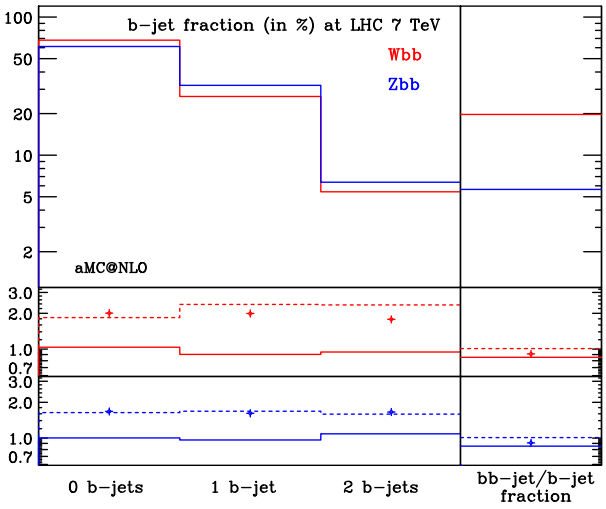


Fig. 44: Fractions of events (in percent) that contain: zero b-jets, exactly one b-jet, and exactly two b-jets. The rightmost bin displays the fraction of b-jets which are bb-jets. The two insets show the ratio of the aMC@NLO results over the corresponding NLO (solid), aMC@LO (dashed), and LO (symbols) ones, separately for Wbb (upper inset) and Zbb (lower inset) production.

obtain the latter). Perturbative, parton-level results agree well with those obtained after shower, and PDF and scale uncertainties (also reported in Fig. 46) are well under control. In summary, we do not observe any significant effects in the shape of distributions due to NLO corrections, which therefore cannot be responsible for the excess of events observed by the CDF collaboration.

14.6 Conclusions

The results we have presented in this contribution are based on the strategic assumption that, for the word *automation* to have its proper meaning, the only operation required from a user is that of typing-in the process to be computed, and other analysis-related information (such as final-state cuts). In particular, the codes that achieve the *automation* may only differentiate between processes depending on their general characteristics, but must never work on a case-by-case basis. The aMC@NLO framework is based on such an assumption, providing a very powerful tool to compare, at the NLO accuracy including showering and hadronization, theory and experiment in high energy collisions. As an example of the flexibility of aMC@NLO we have presented results for the processes $pp \to ttH$, $pp \to Vbb$, $pp \to \ell^+\ell^-\ell^{(\prime)+}\ell^{(\prime)-}$ at the LHC, and a study of $p\bar{p} \to Wjj$ at Tevatron.

Acknowledgements

This research has been supported by the Swiss National Science Foundation (SNF) under contracts 200020-138206 and 200020-129513, by the Belgian IAP Program, BELSPO P6/11-P and the IISN convention 4.4511.10. S.F., F.M. and R.P. acknowledge the financial support of the MICINN project FPA2011-22398 (LHC@NLO).

15. PROBING CORRECTIONS TO DIJET PRODUCTION AT THE LHC [29]

Abstract

We compare and discuss a few kinematic distributions for dijet production at the LHC, computed with a fixed next-to-leading order code, with the

[29] Contributed by: S. Alioli, J. R. Andersen, C. Oleari, E. Re, J. M. Smillie

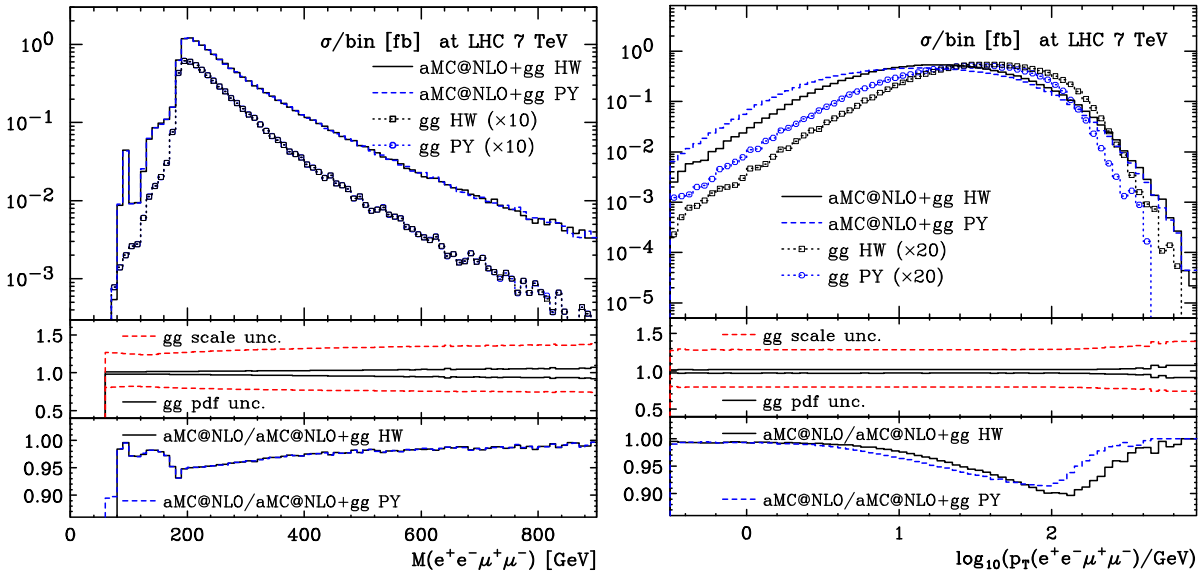


Fig. 45: Four-lepton invariant mass and the transverse momentum distributions for aMC@NLO +gg HERWIG (solid black) and PYTHIA (dashed blue) results. The rescaled gg contributions with HERWIG (open black boxes) and PYTHIA (open blue circles) are shown separately. Middle insets: scale (dashed red) and PDF (solid black) fractional uncertainties. Lower insets: aMC@NLO/(aMC@NLO+gg) with HERWIG (solid black) and PYTHIA (dashed blue).

POWHEG BOX and with HEJ. Previous experimental studies have dealt with kinematic distributions where the predictions of the three approaches were very similar. In this proceeding, we investigate kinematic distributions where the resummed effects in POWHEG and HEJ are clearly shown and enhanced with respect to the fixed NLO result, since different QCD-radiation regimes are probed.

15.1 Introduction

Dijet production is one of the cornerstone processes at the LHC. The cross section for jet production is very large, making it an important testing ground for our understanding of QCD at high-energy scales. In addition, jet production is an important background for many searches for new physics. It is therefore essential to probe and test our theoretical predictions. Dijet-production studies can bring insights in jet production in association with other particles too: for example, Higgs boson production plus two jets in gluon fusion, a key process for assessing the CP properties of the Higgs boson, can benefit from these studies.

There have been a number of very interesting experimental studies in dijet production by both the ATLAS [410, 411, 294] and CMS [412, 413, 414, 415] Collaborations. It is already clear that higher order QCD contributions beyond a fixed order, low multiplicity calculation can be important because the large available phase space for jet emission at the LHC compensates for the suppression of extra powers of the strong coupling constant.

In this contribution, we compare two theoretical approaches to dijet production that include higher order effects: POWHEG [416, 417, 386, 13] and HEJ [418, 419, 64]. The POWHEG method successfully merges a fixed next-to-leading order (NLO) calculation with a parton shower program, that resums leading logarithmic contributions from collinear emissions. In this study, the POWHEG results obtained with the POWHEG BOX [13] are interfaced with the transverse-momentum-ordered shower

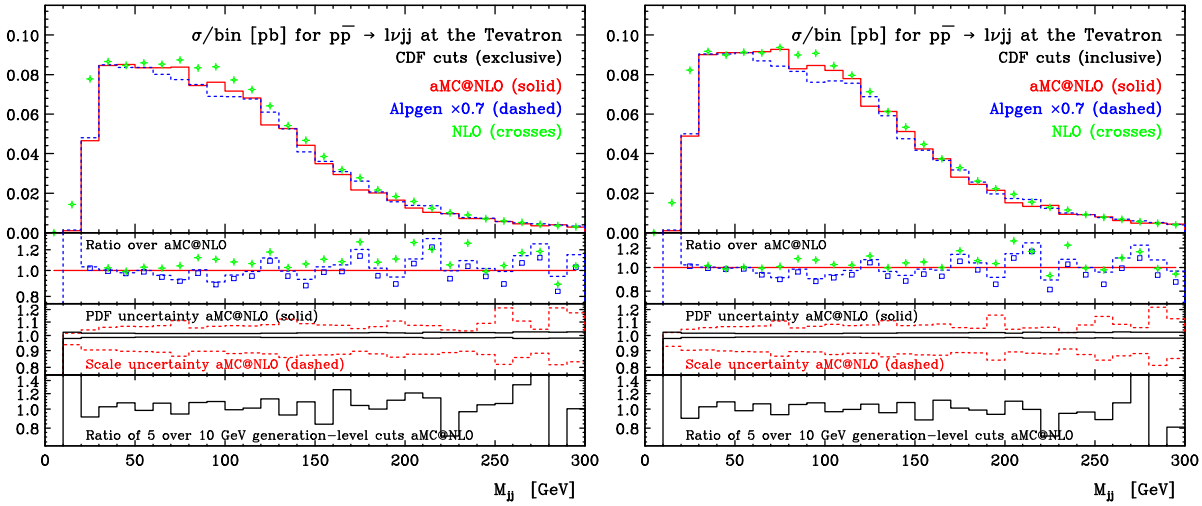


Fig. 46: Invariant mass of the pair of the two hardest jets, with CDF/D0 cuts of [406] (left) and of [409] (right).

provided by PYTHIA 6.4.21 [400]. In contrast, the starting point for HEJ is an all-order approximation to the hard scattering matrix element in the regime of wide-angle QCD emissions. HEJ is accurate at leading logarithmic precision in the invariant mass of any two jets. This is then supplemented with the missing contributions (through a merging and reweighting-procedure) necessary to also ensure tree-level accuracy for final states with up to four jets. The tree-level matrix elements are taken from Standalone Madgraph [163].

The POWHEG and HEJ approaches are clearly very different in their description of QCD radiation. Nevertheless, for several kinematic distributions (see for example ref. [410]) the predictions from POWHEG and HEJ are very similar. In this study, we investigate various observables which can expose the differences in the two approaches and we compare them with the fixed NLO results.

15.2 A comparison between NLO, POWHEG and HEJ in dijet production

In order to avoid biasing our event sample, we impose a minimal set of cuts, avoiding symmetric cuts on the jet transverse momenta that would give an unphysical cross section at fixed NLO level [420, 421], due to the presence of unresummed logarithms. Neither the POWHEG or HEJ descriptions suffer from this instability. However, in order to have a sensible fixed NLO cross section to compare with, we impose asymmetric cuts

$$p_{\mathrm{T}}^{j} > 35\ \mathrm{GeV}, \quad p_{\mathrm{T}}^{j_1} > 45\ \mathrm{GeV}, \quad |y_j| < 4.7\,, \tag{99}$$

i.e. all jets are required to have a minimum transverse momentum of 35 GeV, and the hardest-jet transverse momentum, $p_{\mathrm{T}}^{j_1}$, is required to be greater than 45 GeV. In order to comply with the experimental acceptance, all jets are further required to have an absolute rapidity $|y_j|$ less than 4.7. Jets are defined according to the anti-kt jet algorithm, with radius $R = 0.5$. Only events with at least two jets fulfilling Eq. (99) are kept.

In the following, we compare the fixed NLO cross section with the POWHEG first emission results, with the POWHEG results showered by PYTHIA and with the HEJ predictions. The renormalization and factorization scales have been chosen equal to the transverse momentum of the hardest jet in each event, for the HEJ predictions. For the NLO computation (and for computing the POWHEG $\bar{B}$ function), scales are set to the transverse momentum of the so called underlying Born configuration. Scale-uncertainty bands obtained by varying these scales by a factor of two in each direction are shown for the NLO and HEJ results. The scales entering in the evaluation of parton distribution functions and

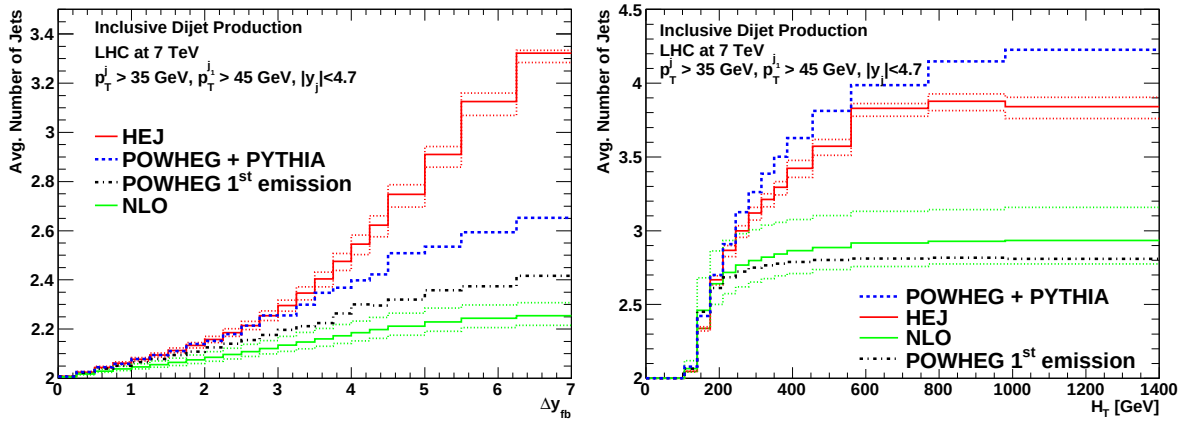


Fig. 47: The average number of jets as a function of $\Delta y_{\rm fb}$ (left plot) and of $H_{\rm T}$ (right plot), as predicted by a fixed NLO calculation, by POWHEG first emission, by POWHEG+PYTHIA and by HEJ. The dotted red lines around the HEJ prediction and the green ones around the NLO result are obtained by varying the renormalization and factorization scales by a factor of two around their central value.

of the strong coupling in the POWHEG Sudakov form factor are instead evaluated with a scale equal to the transverse momentum of the POWHEG hardest emission [417, 386].

In Fig. 47 we plot the average number of jets as a function of the rapidity difference between the most forward and most backward of the jets fulfilling Eq. (99), Δy_{fb}, on the left-hand side, and as a function of $H_{\rm T} = \sum_j p_{\rm T}^j$ on the right-hand side. The wide-angle resummation implemented in HEJ produces more hard jets than POWHEG and the fixed NLO calculation, as the rapidity separation between the most forward and the most backward jet in the event increases. Both the NLO and the first-emission POWHEG results have at most 3 jets, so that the average number of jets cannot exceed 3, and give similar results. Additional jets are instead produced by the PYTHIA shower, so that the average number of jets is increased by roughly 20% with respect to the NLO one, for $\Delta y_{\rm fb} \approx 7$. For the same separation in rapidity, the HEJ prediction is 45% larger than the NLO result, with a chance to distinguish among the three approaches.

The dependence of the average number of jets from $H_{\rm T}$ (right plot) displays a different behaviour: here the showered events have on average more jets than HEJ and the NLO results, as the sum of the transverse momentum of all the final-state jets increases. It is interesting here to comment on the NLO result obtained with the factorization and renormalization scales set to $p_{\rm T}^{\rm UB}/2$, i.e. half of the transverse momentum of the underlying Born configuration. In fact, from the plot, an unphysical behaviour of this quantity emerges: the average number of jets is greater than 3 above $H_{\rm T} \approx 270$ GeV. This is due to the fact that the high $H_{\rm T}$ region is populated mostly by events with 3 jets, two of which have approximately the same high transverse momentum, and the third one is softer with respect to the other two (the cuts in Eq. (99) are always in place). In this configuration, the exclusive two-jet cross section becomes negative, due to incomplete cancellation of the virtual (negative) contribution, now enhanced by a higher value of the strong coupling constant, evaluated at a lower renormalization scale. A more detailed discussion can be found in ref. [422].

As a last example of a kinematic distribution that displays different behaviour if evaluated at NLO or using POWHEG or HEJ, we plot in Fig. 48 the average value of $\cos(\pi - \phi_{\rm fb})$, where $\phi_{\rm fb}$ is the azimuthal angle between the most forward and backward jets, as a function of their rapidity separation $\Delta y_{\rm fb}$. For dijet events at tree-level, $\phi_{\rm fb} = \pi$ since the two jets must be back-to-back, and the average value of the cosine is 1. Deviation from 1 then indicates the presence of additional emissions, so that

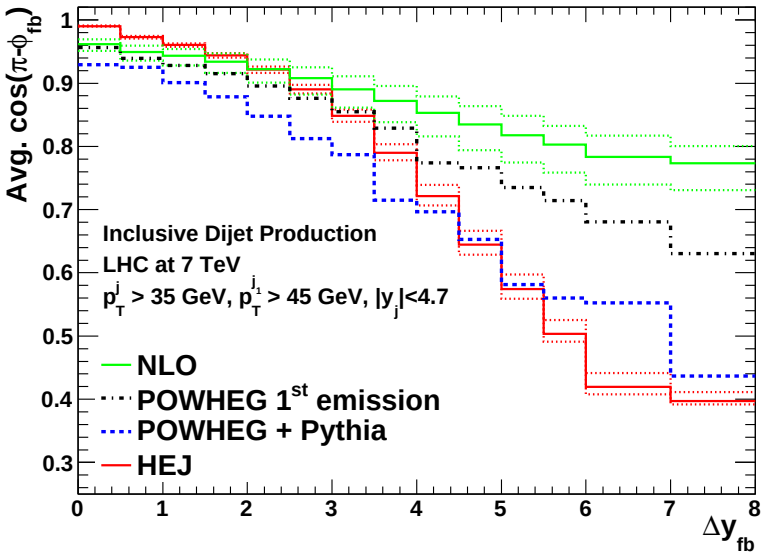


Fig. 48: The average value of $\cos(\pi - \phi_{\rm fb})$ as a function of $\Delta y_{\rm fb}$, where $\phi_{\rm fb}$ is the azimuthal angle separation between the most forward and most backward jet. The dotted red and green lines are obtained by varying the renormalization and factorization scales by a factor of 2 in both directions.

this kinematic distribution carries information on the decorrelation between the two jets. This quantity is more inclusive than the average number of jets as it is sensitive also to emissions below the jet $p_{\rm T}$ cut. The higher radiation activity in POWHEG+PYTHIA and in HEJ, with respect to the fixed NLO and the POWHEG first-emission results, is clearly visible in the figure: the stronger jet activity produced by HEJ at higher rapidity separation (see the left plot of Fig. 47) lowers the average value of the cosine below the POWHEG+PYTHIA result. As expected, the average value predicted by the POWHEG first-emission and the NLO calculation is closer to 1, since they contain at most one radiated parton.

Conclusions

In this proceeding, we have discussed the results obtained using a fixed NLO calculation, HEJ and POWHEG+PYTHIA, in the description of three kinematic distributions, selected in order to display more clearly the differences among the three approaches: the average number of jets and azimuthal decorrelation between the most forward and the most backward jet, plotted as a function of the rapidity separation of the most forward and the most backward jet, and the average number of jets plotted as a function of the sum of the transverse momenta of all the jets in the event.

While the limitations of the NLO calculation are clearly visible when we probe regions of the phase space where multi-jet emissions becomes important, the predictions of POWHEG+PYTHIA and HEJ are distinguishable when dealing with the average number of jets as a function rapidity span. Less marked differences are found as a function of Ht, and in the study of the azimuthal decorrelation of the most forward and backward jet.

An experimental analysis of the dijet data, collected at the LHC, should then follow to investigate to which extent our theoretical knowledge for these kinematic distributions is under control.

Acknowledgments

All of the authors would like to thank the organizers for an extremely successful and inspiring workshop, and the staff of the Ecole de Physique des Houches for their hospitality. The work of ER and JMS is supported by the UK Science and Technology Facilities Council (STFC). ER and SA acknowledge

financial support from the LHCPhenoNet network under the Grant Agreement PITN-GA-2010-264564 for travel expenses.

16. W+JETS PRODUCTION AT THE LHC: A COMPARISON OF PERTURBATIVE TOOLS [30]

Abstract

In this contribution, we discuss several theoretical predictions for W plus jets production at the LHC, compare the predictions to recent data from the ATLAS collaboration, and examine possible improvements to the theoretical framework.

16.1 Motivation

Experimentalists are reliant on a number of tools, at LO and NLO, at parton level and at hadron level, in order to understand both simple and complex final states at the LHC. One of the benchmark processes, for both signals to new physics and for their backgrounds, is the production of W plus jets. In this contribution, we discuss several different predictions for the W plus jets final state, concentrating on the H_T distribution. We examine where the predictions agree, and where they disagree and compare the predictions to LHC data. We introduce the idea of NLO 'Exclusive Sums', and discuss the performance of this technique and consider also how LoopSim may be able to improve the predictions. We document the use of ROOT ntuples for W plus jets predictions produced by the BlackHat+Sherpa collaboration, indicating how they can be used to examine the variation of the cross sections with jet size/algorithm, PDFs, and scale choices. We also study the possibility of using the LoopSim method together with BlackHat+Sherpa type ntuples, since this may offer the opportunity to improve on the results from NLO Exclusive Sums.

16.2 Theory tools: strengths and weaknesses

NLO is the first order at which the normalization (and sometimes the shape) of LHC cross sections can be realistically calculated. The state of the art is in parton-level programs such as BlackHat+Sherpa, where $W + n$-jet cross sections are available, with n up to 4 at NLO [70, 51, 22] (and soon up to 5 [423]). Of course, such parton-level final states do not allow for the full comparisons to the data allowed by the full parton shower Monte Carlo programs such as Sherpa. NLO matrix elements have been included into parton shower Monte Carlos, but only for relatively simple final states (although we note that the NLO matrix elements for $W + 2$ jets [404] and $W + 3$ jets [424] have recently been implemented in parton shower Monte Carlo programs).

The Sherpa Monte Carlo program [146, 425] includes the exact LO $W + n$-parton ($W + n$-jet) matrix elements, with n up to 4 (in this study), using the newer ME&TS scheme as introduced in Refs. [426, 427, 428] for the addition of states with different jet multiplicities with the correct normalizations. The newer matrix-element plus parton-shower merging scheme improves over the CKKW [429, 430] formalism by allowing for a better interplay between the matrix-element and parton-shower descriptions. This in particular required the implementation of truncated showers ('TS'). As before, additional jets are, of course, then produced by the parton shower. Both BlackHat+Sherpa and Sherpa rely on DGLAP-based evolution of gluon emission, on the assumption that the gluon emissions are strongly ordered in transverse momentum. For an alternative prediction, we use the program HEJ [418, 419, 431]. The High Energy Jets (HEJ) framework provides a leading-log resummation of the dominant terms in the limit of large invariant mass between jets. In addition, HEJ contains a merging procedure to ensure tree-level accuracy for final states with two, three or four jets.

[30] Contributed by: J. R.'Andersen, J. Huston, D. Maître, S. Sapeta, G. P. Salam, J. M. Smillie, J. Winter

A NLO n-jet prediction produces events with with either n or $n+1$ partons. For observables for which higher multiplicities have a significant impact, this limitation can be detrimental. If one has predictions for different multiplicities, one can try to combine them by avoiding double counting by requiring that the n-jet prediction is used only to describe n-jet events (except for the highest multiplicity where $(n+1)$-jets configurations are allowed). This procedure is crude and does not increase the formal accuracy of the prediction which is that of NLO of the smallest multiplicity. The idea is that, in observables where higher multiplicities events dominate, a better prediction might be obtained. This has been denoted as the 'Exclusive Sums' technique. The impact of the Exclusive Sums approach depends on the kinematic variable under consideration. For this contribution, we consider only the H_T variable, defined as the sum of the transverse momenta of all of the leptons (including neutrinos) and jets in the event. The impact of the approach is expected to depend on the observable under consideration and it may be more beneficial for variables sensitive to multi-jet radiation, such as H_T, than for more inclusive variables such as $p_{t,W}$. Comparisons for the latter are left to a study now in progress.

16.3 Use of BlackHat+Sherpa ntuples

As has been partially detailed in these proceedings, there have been many advances in the computation of the NLO corrections for multi-parton final states. Often such calculations do not exist in a compact user-friendly form, and other means must be taken to allow experimentalists to have access to the results. The BlackHat+Sherpa collaboration has chosen to make available ROOT tuples that contain all of the parton-level information needed to form flexible predictions. The ROOT ntuple framework is a very efficient way to store such information and the use of ROOT tuples is very familiar to experimentalists.

The ROOT ntuples store the four-vectors for the final state partons, as well as their flavor information. The calculation is originally performed using a specific choice of PDF, $\alpha_s(m_Z)$, renormalization scale μ_R and factorization scale μ_F, but weight information is also stored in the ntuples that allows each event to be easily re-weighted to any other (reasonable) values for the above parameters. (PDFs are varied through calls to LHAPDF [432].) No jet clustering has been performed on the final state partons; jet reconstruction is left to the user, for any jet algorithm/size for which the correct counter-events are present in the ntuple. For the results presented here, the SISCone [433], k_T [361] and anti-k_T [341] algorithms, with jet radii R of 0.4, 0.5, 0.6 and 0.7 can be used. Each of the above jet algorithms were run and the results stored in SpartyJet ntuples.[31] The SpartyJet tuples were 'friended' with the BlackHat+Sherpa ntuples, allowing the analysis script access to all jet information. Such a flexibility allows for an investigation of the dependence of the physics on the details of the manner in which the partons are combined into jets, in a manner difficult to achieve prior to this.

The four-vector information stored in the BlackHat+Sherpa ntuples is shown in Table 11. Note the variety of entries needed for the re-weighting of the cross section results, especially for the case of the variation of the two scales μ_R and μ_F. Information is stored in separate ntuples for the different categories of events, which are typically Born, loop (leading color and sub-leading color), real and subtraction terms. For large n, in $W+n$-parton final states, there are many divergences present when two partons become collinear or one parton becomes soft. These divergences are controlled using the traditional Catani–Seymour approach [236], which involves the generation of many counter-events. Many of the events have negative weights; only the sum is guaranteed to be positive-definite. Predictions with reasonable statistical precision may require the sum of billions of events. The resultant tuples may amount to several Terabytes. However, the output can be subdivided into ROOT files of order 5–10 GB, allowing for simultaneous parallel processing of the events over multiple nodes, such as in the Tier3 facility at Michigan State University used for these comparisons.

[31] SpartyJet [434] is a set of software tools for jet finding and analysis, built around the FastJet library of jet algorithms [435]. SpartyJet provides four key extensions to FastJet: a simple Python interface to most FastJet features, a powerful framework for building up modular analyses, extensive input file handling capabilities, and a graphical browser for viewing analysis output and creating new on-the-fly analyses.

16.4 BlackHat+Sherpa predictions

We have generated NLO predictions with the BlackHat+Sherpa predictions implementing the cuts used in the 2010 ATLAS W plus jets paper [436]. For completeness, the cuts are reproduced below:

- $p_T^{\rm lepton} > 20$ GeV,
- $|\eta^{\rm lepton}| < 2.4$,
- $E_T^{\rm miss} > 25$ GeV,
- $m_{T,W} > 40$ GeV,
- $p_T^{\rm jet} > 30$ GeV,
- $|y^{\rm jet}| < 4.4$,
- $\Delta R^{\rm lepton-jet} > 0.5$.

In Figure 49, we show the NLO BlackHat+Sherpa prediction for the H_T distribution for $W+ \geq 1$ jets (left) using the anti-k_T jet algorithm with $R = 0.4$. As the prediction is an inclusive NLO calculation for $W+ \geq 1$ jets, there are contributions from both the one-jet and the two-jet final states. Note that as H_T increases, the contributions from the $W + 2$-jet subprocess also increases. On the right, we again show the H_T distribution, but now compute the prediction using the 'Exclusive Sums' technique, adding in the NLO $W + 2$-jet information. Now there is a significant contribution at high H_T from the $W + 3$-

branch name	type	notes
id	I	id of the event. Real events and their associated counter-terms share the same id. This allows for the correct treatment of statistical errors.
nparticle	I	number of particles in the final state
px	F[nparticle]	array of the x components of the final state particles
py	F[nparticle]	array of the y components of the final state particles
pz	F[nparticle]	array of the z components of the final state particles
E	F[nparticle]	array of the energy components of the final state particles
alphas	D	α_s value used for this event
kf	I	PDG codes of the final state particles
weight	D	weight of the event
weight2	D	weight of the event to be used to treat the statistical errors correctly in the real part
me_wgt	D	matrix element weight, the same as weight but without pdf factors
me_wgt2	D	matrix element weight, the same as weight2 but without pdf factors
x1	D	fraction of the hadron momentum carried by the first incoming parton
x2	D	fraction of the hadron momentum carried by the second incoming parton
x1p	D	second momentum fraction used in the integrated real part
x2p	D	second momentum fraction used in the integrated real part
id1	I	PDG code of the first incoming parton
id2	I	PDG code of the second incoming parton
fac_scale	D	factorization scale used
ren_scale	D	renormalization scale used
nuwgt	I	number of additional weights
usr_wgts	D[nuwgt]	additional weights needed to change the scale

Table 11: Branches in a BlackHat+Sherpa ROOT file.

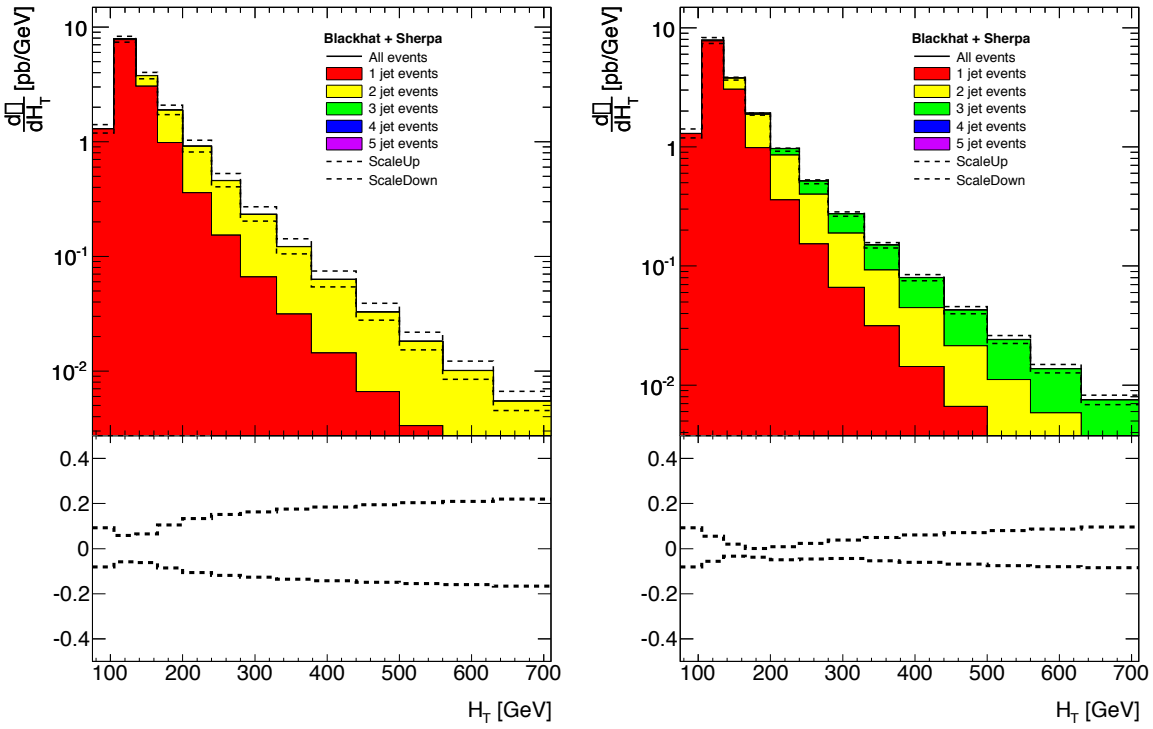


Fig. 49: The W plus jets cross section, as a function of H_T, for the NLO inclusive $W+ \geq 1$-jet prediction (left) and for the Exclusive Sums approach, adding in $W + 2$-jet production at NLO (right). The cross sections have been evaluated at a central scale of $H_T/2$ and the uncertainty is given by varying the renormalization and factorization scales independently up and down by a factor of 2, while ensuring that the ratio of the two scales is never larger than a factor of 2.

jet final state as well. In Figure 50, the H_T prediction is shown using the Exclusive Sums approach, adding 1+2+3 jets at NLO (left) and 1+2+3+4 jets at NLO (right). It is evident that as H_T increases, contributions from higher jet multiplicities that are only present implicitly in a traditional inclusive NLO $W+ \geq 1$-jet calculation, become important. The Exclusive Sums H_T predictions agree with that for the inclusive NLO $W+ \geq 1$-jet calculation at low H_T, but are larger at higher H_T, and in better agreement with the ATLAS data (as discussed below).

However, it can also be noticed that the scale dependences for the Exclusive Sums predictions apparently get better when the 2-jet NLO information is added, but significantly worse when the 3-jet and 4-jet information is added. As discussed in the Appendix, the reduction in scale dependence with the addition of the 2-jet NLO terms may be due to the stabilization of the predictions for the $qq \to Wq'q$ topologies. Adding the 3-jet and 4-jet NLO terms seems to destabilize the predictions. There are missing Sudakov terms needed to properly 'stitch' the different multiplicity samples together; it is hoped that the LoopSim technique may offer one way in supplying those missing terms.

Below in Figure 51, we show the NLO BlackHat+Sherpa predictions for the H_T distribution for $W+ \geq 2$ jets: the inclusive calculation to the left, the Exclusive Sums result adding 2+3-jet NLO information in the middle and the Exclusive Sums result adding 2+3+4-jet NLO information to the right. Over the kinematic range covered in these plots, the Exclusive Sums technique adds less to the cross section at high H_T, although there is still a degradation of the scale dependence.

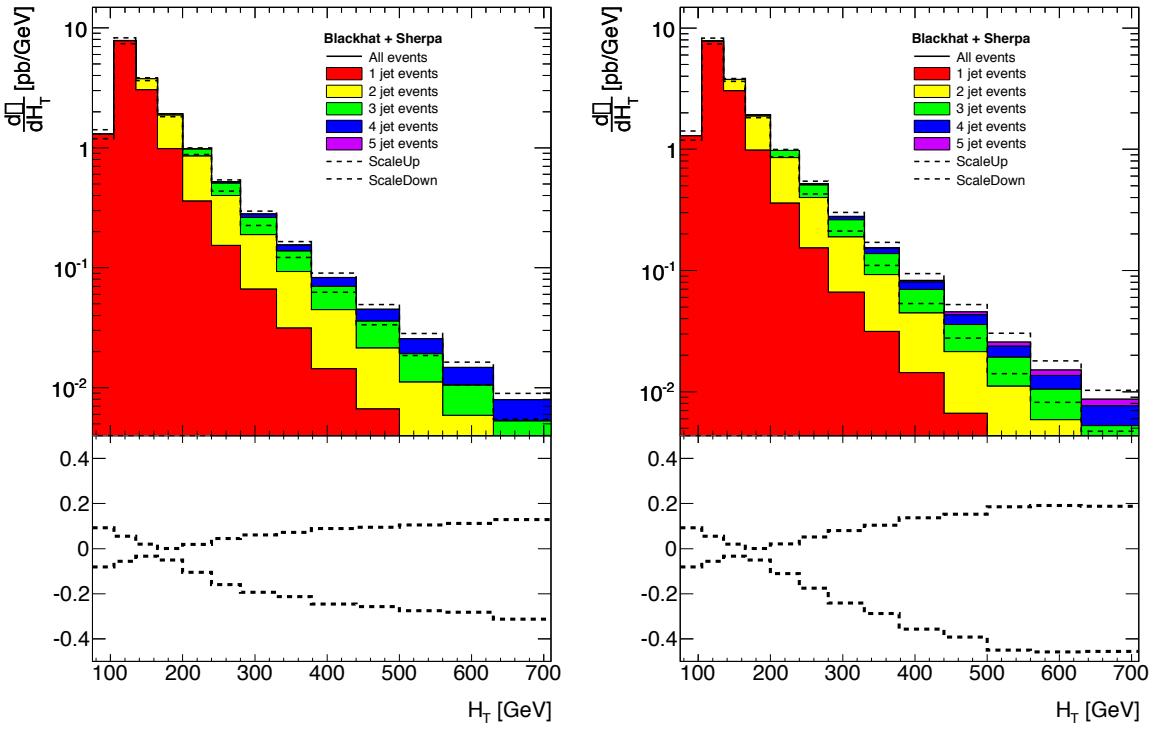


Fig. 50: The W plus jets cross section, as a function of H_T, for $W+ \geq$ 1-jet production using the Exclusive Sums approach, and adding up to 3 jets at NLO (left) and 4 jets at NLO (right). The cross sections have been evaluated at a central scale of $H_T/2$ and the uncertainty is given by varying the renormalization and factorization scales independently up and down by a factor of 2, while ensuring that the ratio of the two scales is never larger than a factor of 2.

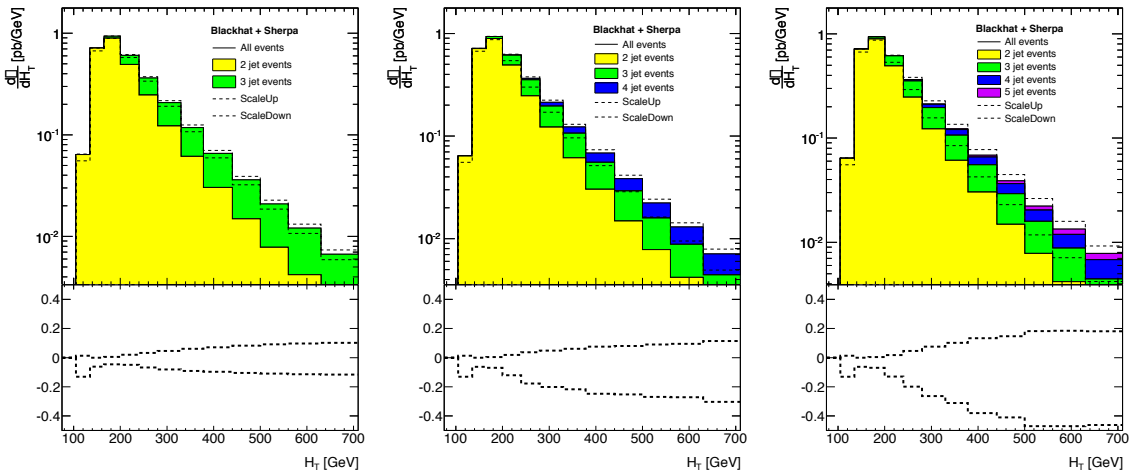


Fig. 51: The W plus jets cross section, as a function of H_T, for $W+ \geq$ 2-jet production using the inclusive NLO production (left) and the Exclusive Sums approach, adding up to 3 jets at NLO (center) and 4 jets at NLO (right). The cross sections have been evaluated at a central scale of $H_T/2$ and the uncertainty is given by varying the renormalization and factorization scales independently up and down by a factor of 2, while ensuring that the ratio of the two scales is never larger than a factor of 2.

16.5 Towards interfacing BlackHat+Sherpa ntuples with LoopSim

LoopSim is a method to simulate higher order QCD corrections, in particular those beyond NLO. It is expected to work best for processes with large NLO-to-LO K-factor, however it was found to be advantageous even in some cases where the K-factor is moderate [437]. The method is based on unitarity and its main ingredient is a procedure that takes events from a process with $n+m$ partons in the final state and produces counter-term events with $n+m-1, n+m-2, \ldots, n$ particles, which approximate 1-loop, 2-loop, etc. contributions. In contrast to the Exclusive Sums method, it enables one to introduce (approximate) virtual corrections beyond 1-loop, thus ensuring that the $\alpha_s L^2$ type terms cancel for all the orders that are included. While we will not show LoopSim results that are directly comparable to the ATLAS data (the samples were generated before those cuts were made public), we will examine below the dependence on the $p_{t,\min}$ choice (which sets the size of $L = \ln O_t/p_{t,\min}$ where O_t is a transverse observable) and see that it vanishes as $p_{t,\min} \to 0$.

To distinguish between the exact result at the order N^{p}LO and the result with simulated loops we use a notation in which we replace N by $\bar{n}$ for the orders simulated by LoopSim. So for example, $W+1$ jet at $\bar{n}$LO has approximate 1-loop diagrams and is obtained by combining $W+1$ jet at LO with $W+2$ jet at LO where the latter is passed through LoopSim. Similarly $W+1$ jet at $\bar{n}$NLO has exact 1-loop diagrams but simulated 2-loop contributions (by using $W+2$ jet at NLO as an input to LoopSim).

As argued in the previous section, the BlackHat+Sherpa ntuples allow one to efficiently perform a broad range of analyses. They have however a limitation. In order to reduce the size of stored files, the only partonic events that are recorded for the $W+n$-jet sample are those in which there are at least n jets above a 20 GeV threshold. Since this threshold is below the jet cuts used by ATLAS and CMS, it is adequate for any NLO study of LHC jet cross sections. The situation is slightly more complex if we want to use the BlackHat+Sherpa ntuples to compute predictions beyond NLO using LoopSim. This is because the cut that is present in the $W+2$-jet BlackHat+Sherpa NLO sample eliminates part of the real contribution to the $W+1$-jet phase space at NNLO, for example $W+3$-parton events in which the 3 partons all form part of a single jet, or in which 2 partons form part of one jet, while the third is well separated in angle but below the 20 GeV jet threshold.

Since we plan to use LoopSim interfaced to BlackHat+Sherpa ntuples in our future study of multi-jet processes, it is important to directly check the effect of the finite generation p_t cut, $p_{t,\mathrm{gen}}^{\min}$, on the predictions of the p_t and H_T distributions. We have performed such a study for $W^- + 1$ jet generated with MCFM, where we varied a 'parton'-p_t generation cut from 1 to 20 GeV.[32] This is not entirely equivalent to the cut in the BlackHat+Sherpa samples (which is applied to the standard jets, not to the partons), but should be adequate from the point of view of estimating the potential order of magnitude of finite generation cuts. The output from MCFM was interfaced to LoopSim which produced the additional loop diagrams. Then, the events were analyzed with the following set of cuts: $|y^{\mathrm{lepton}}| < 2.5$, $p_T^{\mathrm{lepton}} > 20$ GeV , $|y^{\mathrm{jet}}| < 4.5$, $p_T^{\mathrm{jet}} > 25$ GeV , $m_{T,W} < 20$ GeV, where the anti-k_T algorithm with $R = 0.4$ was used for clustering.

The results are presented in Figure 52 where the ratios of cross sections obtained with a range of generation cuts are shown as functions of the p_t of the leading jet and $H_{T,\mathrm{jets}}$. At NLO, the only artefact we see is for the $p_{t,\mathrm{gen}}^{\min}$ of 20 GeV in the bin below 40 GeV. This is as expected, since a 20 GeV cut on each of two partons can at most affect jets up to 40 GeV (such an artefact would not be present in the BlackHat+Sherpa samples). At $\bar{n}$LO and $\bar{n}$NLO, however, the dependence on $p_{t,\mathrm{gen}}^{\min}$ is extended to a larger range of $p_{t,\mathrm{lead.jet}}/H_{T,\mathrm{jets}}$ and it is visible also for values of $p_{t,\mathrm{gen}}^{\min} < 20$ GeV. However, even if the $p_{t,\mathrm{gen}}^{\min}$ dependence of the $\bar{n}$LO and $\bar{n}$NLO results is stronger than at NLO, it dies out quickly with increasing $p_{t,\mathrm{lead.jet}}/H_{T,\mathrm{jets}}$ and becomes irrelevant at ~ 100 GeV, depending on the observable and the

[32] Strictly speaking, the events destined for the $W+n$-jet sample were clustered with the k_T algorithm with a very small separation, $R = 0.03$, and accepted when at least n small-R jets were above the generation cut. These events were then passed through the standard analysis cuts (with or without LoopSim).

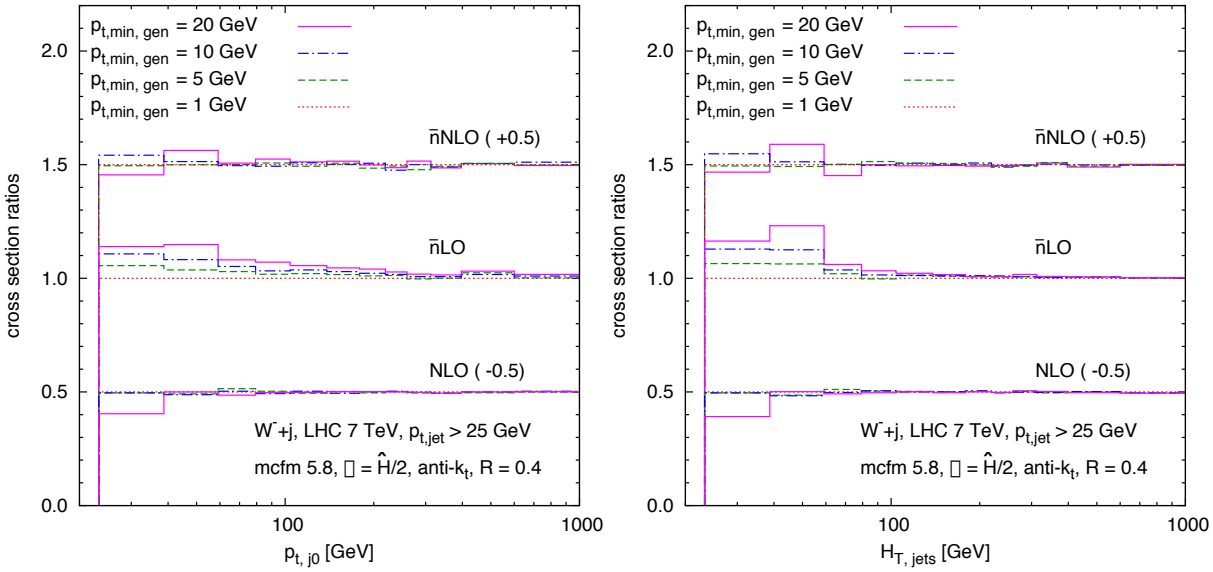


Fig. 52: Ratios of cross sections from runs with a certain range of $p_{t,\mathrm{gen}}^{\mathrm{min}}$ values taken wrt. the cross section generated with $p_{t,\mathrm{gen}}^{\mathrm{min}} = 1$ GeV for the distributions of p_t of the leading jet (left) and the scalar sum of jets' transverse momenta, $H_{T,\mathrm{jets}}$ (right).

order. This appears to be consistent with the expectation that the effect of the cut should vanish as a power of $p_{t,\mathrm{gen}}^{\mathrm{min}}/p_{t,\mathrm{lead.jet}}$ or $p_{t,\mathrm{gen}}^{\mathrm{min}}/H_{T,\mathrm{jets}}$.

Therefore we conclude, that in spite of the finite generation cut one should be able to trust the results obtained using BlackHat+Sherpa ntuples, above a moderate p_t limit, even for more complex analyses such as those involving LoopSim.

16.6 Comparisons to data, Sherpa and HEJ predictions

In Figure 53 (left), we compare the ratio of the 2010 ATLAS W plus jets data for the H_T distribution for $W+ \geq 1$ jets to predictions using the generic NLO calculation for $W+ \geq 1$ jet, the Exclusive Sums approach adding up to 4 jets at NLO and the Monte Carlo event generator Sherpa. The agreement between the data and the pure NLO result is rather poor; it improves substantially with the inclusion of the Exclusive sums up to two jets at NLO, with further small improvements coming from higher multiplicities. As a reminder, we previously noted that the scale dependence improved when adding the 2-jet NLO information, but degraded when adding higher jet multiplicities. The Sherpa prediction slightly overshoots the data for H_T in the inclusive $W + 1$-jet bin. We however note that the data versus Sherpa H_T ratio has been formed based on the absolute normalization as given by the Monte Carlo simulation. Comparing the inclusive 1-jet cross sections, we find a factor of 0.97 between the data and the Sherpa result.

In Figure 53 (right), we compare the ratio of the 2010 ATLAS W plus jets data for the H_T distribution for $W+ \geq 2$ jets to predictions using the generic NLO calculation for $W+ \geq 2$ jet, the Exclusive Sums approach adding up to 4 jets at NLO, and to predictions from HEJ and from Sherpa. As noted previously, there is some increase in the predictions from the Exclusive Sums approach at the highest H_T values, but not nearly as much as in the $W+ \geq 1$-jet case. These increases go in the direction of closer agreement with the data, but the statistical error does not allow a clear judgement to be made. The Sherpa and HEJ predictions for this ratio are in reasonable agreement with the data but appear to fall off somewhat more rapidly at large H_T than either the data or the various BlackHat+Sherpa predictions. Again this partly is the result of relying on the absolutely normalized Monte Carlo predictions, which yield $W+ \geq 2$-jet normalization factors of 0.95 or 0.93 between data and Sherpa or HEJ, respectively.

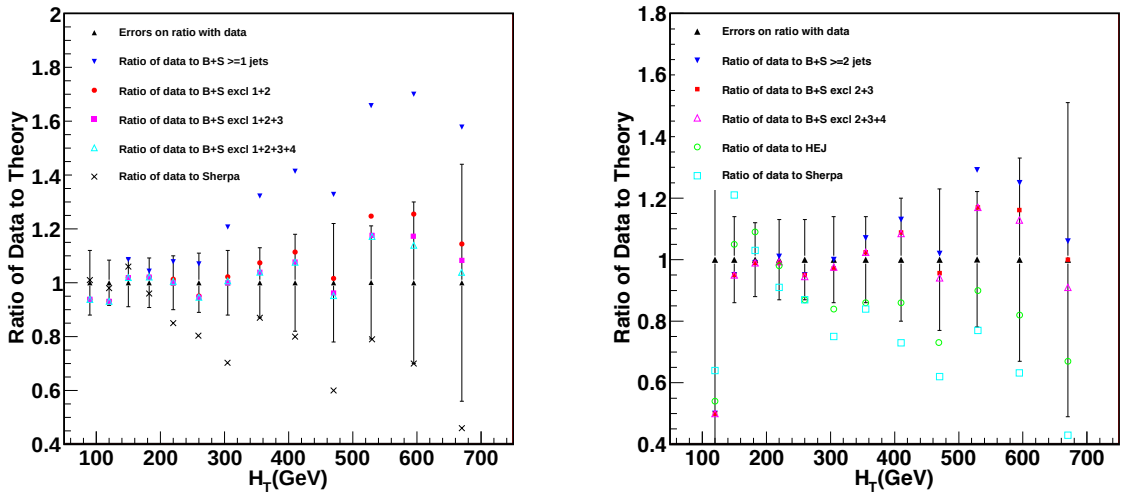


Fig. 53: The ratios of the ATLAS $W+ \geq 1$-jet (left) and $W+ \geq 2$-jet (right) cross sections, as a function of H_T, taken wrt. various theory predictions. The absolute normalization has been kept as given by the calculations. The error bars represent the total fractional error (statistical plus systematic added in quadrature) at each point.

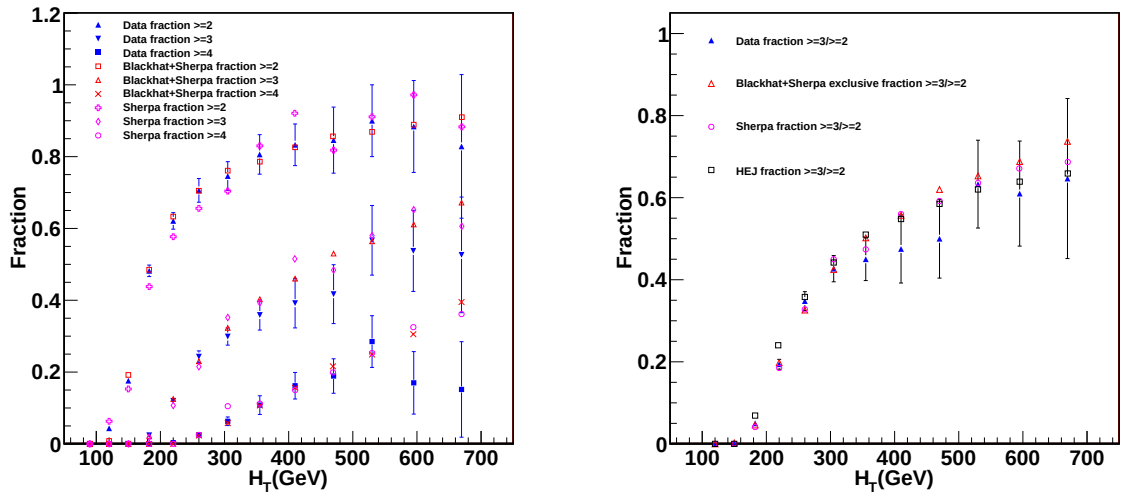


Fig. 54: (left) The fraction of the H_T cross section for $W+ \geq 1$-jet events arising from the $W+ \geq 2$-jet, $W+ \geq 3$-jet and $W+ \geq 4$-jet final states derived from the Exclusive Sums approach, from Sherpa and from HEJ, compared to the 2010 ATLAS data. (right) The ratio of the cross sections for $W+ \geq 3$ jets to $W+ \geq 2$ jets, as a function of H_T, using predictions from the Exclusive Sums approach, from Sherpa and from HEJ, compared to the ratio from the 2010 ATLAS data.

In Figure 54 (left), we show the predictions for the fractions of the H_T cross section in the inclusive $W + 1$-jet bin arising from the inclusive $W + 2$-jet, $W + 3$-jet and $W + 4$-jet final states as obtained from the Exclusive Sums approach and from Sherpa, compared to the 2010 ATLAS data. In Figure 54

(right), we show the ratio of the cross sections for $W+ \geq 3$ jets to $W+ \geq 2$ jets, as a function of H_T, again using predictions from the Exclusive Sums approach and from Sherpa but also from HEJ. We again compare to the ratio given by the 2010 ATLAS data. All three predictions agree with each other and with the data over the range considered, despite the big differences in the approaches. There may be an indication of some separation between the predictions at the very highest H_T values.

16.7 Conclusions, outlook and future studies

The advances achieved over the last few years in calculating NLO corrections for multi-jet final states allow a more serious consideration of the possibility to combine various n-jet NLO predictions into an inclusive jet sample. The Exclusive Sums approach discussed in this contribution is a first promising step into this direction. More studies are required to understand the uncertainties related to this procedure. One way of doing so would be to test the stability of the predictions against variation of the jet algorithm and/or parameters of the jet algorithm used to obtain and separate the different NLO predictions for the fixed-multiplicity sets that eventually make up the sum of exclusive n-jet contributions.[33]

For the Exclusive Sums approach, outlined here for the case of $W+ \geq 1$ jets, contributions are added proportional to α_s^2 ($W + 1$ jet at NLO), α_s^3 ($W + 2$ jets at NLO), α_s^4 ($W + 3$ jets at NLO) and α_s^5 ($W + 4$ jets at NLO), i.e. this procedure mixes powers of α_s and thus is missing essential Sudakov form factors that effectively bring each term to the same power of α_s. One could imagine accomplishing this by embedding the NLO matrix elements in a parton shower Monte Carlo framework, however the technology for merging different multiplicities of NLO calculations with a parton shower is still under development. Note that at LO the tree-level matrix-element plus parton-shower merging methods (e.g. as implemented in Sherpa) are designed to satisfy this same-$\mathcal{O}(\alpha_s)$ requirement by including the (all-orders) leading-log effects to the 'LO Exclusive Sums' exhibiting the LO analog of the Exclusive Sums discussed here. Compared to the matrix-element plus parton-shower merging, we see that the 'NLO Exclusive Sums' technique only accounts for Sudakov effects up to $\mathcal{O}(\alpha_s)$ while it describes each jet bin at full NLO instead of LO accuracy.

Relying on the parton shower Monte Carlo framework is not the only way to go in refining the Exclusive Sums strategy. Alternatively, the LoopSim method can be used to provide approximations to the higher-loop terms missing in the Exclusive Sums approach. As we have seen here, prospects for using it together with BlackHat+Sherpa ntuples seem promising. A detailed comparison of the LoopSim results to LHC data is however beyond the scope of this Les Houches contribution, though we look forward to it being carried out in the near future.

The ATLAS data taken in 2011 is about a factor of 130 times as large as the data taken in 2010 (the only published data for W plus jets so far). This will allow a much further reach in all kinematic variables. To get an idea, we show in Figure 55 the ratio of the predictions from the Exclusive Sums to the respective inclusive NLO predictions for $W+ \geq 1, 2, 3$ jets. At an H_T value of 2 TeV, the ratio for $W \geq 1$ jet is of the order of 2; the ratio for $W \geq 2$ jets rises to about 1.4. The NLO-to-LO K-factor for $W+ \geq 1$ jet rises rapidly with increasing H_T, while the K-factor for $W+ \geq 2$ jets increases only moderately (because no new subprocesses are being introduced). It will be interesting to see if (a) the additional factor of 2 (for the $W+ \geq 1$-jet case) and (b) the additional factor of approximately 1.4 (for the $W+ \geq 2$ jet case) lead to better agreement with the data. The LHC data from 2011 (and the higher statistics expected in 2012) will reach these kinematic values and should shed further light on the necessity and the efficacy of this theoretical technique, not only for $W+ \geq 1$ jet, but for higher jet multiplicities as well.

[33] In this study, the same jet definition (anti-k_T with $R = 0.4$) was used for both establishing the separation of the fixed jet-multiplicity contributions and evaluating the cuts and observables.

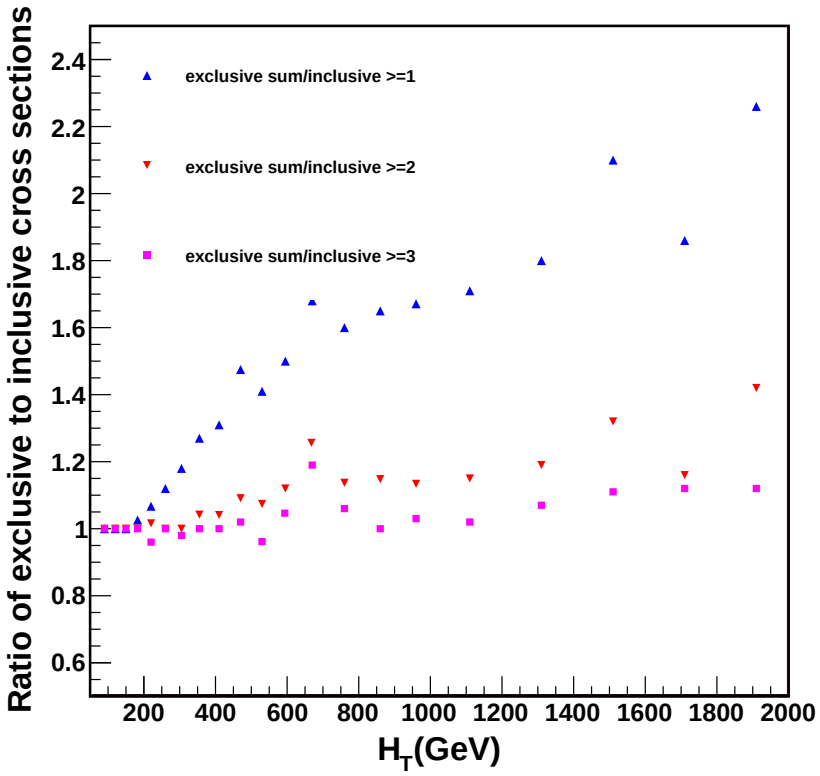


Fig. 55: The ratio of the predictions obtained from the NLO Exclusive Sums approach to the inclusive NLO predictions for $W \geq 1$ jets, $W \geq 2$ jets and $W \geq 3$ jets. The 'jitter' is due to the limited BlackHat+Sherpa statistics for these predictions.

Acknowledgments

We would like to thank the Les Houches organizers for a very stimulating workshop, and Zvi Bern, Lance Dixon and Kemal Ozeren for useful discussions. GPS and SS gratefully acknowledge support from the Agence Nationale de la Recherche under grant ANR-09-BLAN-0060. JMS is supported by the UK Science and Technology Council (STFC). JH acknowledges support from the National Science Foundation.

Appendix: a double logarithmic analysis of the Exclusive Sums method

To help understand the structure of the Exclusive Sums method, it can be useful to consider how it works in a simple double logarithmic approximation. We use $p_{t,\min}$ to represent the minimum p_t for the jets in the Exclusive Sums sample, and first study the cross section for W production as a function of $p_{t,W}$ at high $p_{t,W}$ ($\gg m_W$), considering in particular the terms that go as $\alpha_s^n L^{2n}$ where $L = \ln p_{t,W}/p_{t,\min}$. The 0-jet sample does not contribute at all to non-zero $p_{t,W}$, so the first term comes from the exclusive 1-jet contribution. If calculated to all orders in the double logarithmic approximation (DLA), it would have the form

$$\sigma^{\rm DLA}_{1,\rm excl}(p_{t,W}) \;=\; \sigma^{\rm LO}_1(p_{t,W})\,\exp\left(-\frac{2\,C\alpha_s}{\pi}L^2\right)\,, \tag{100}$$

where $C = 2\,C_F + C_A$ for the (dominant) $qg \to W^{\pm}q'$ scattering process. The n exclusive jet rate would be given by

$$\sigma^{\text{DLA}}_{n,\text{excl}}(p_{t,W}) \;=\; \sigma^{\text{LO}}_{1}(p_{t,W})\,\frac{1}{(n-1)!}\left(\frac{2\,C\alpha_s}{\pi}L^2\right)^{n-1}\exp\left(-\frac{2\,C\alpha_s}{\pi}L^2\right), \tag{101}$$

and one sees that the sum over all multiplicities is given by

$$\sigma(p_{t,W})^{\text{DLA}} \;=\; \sum_{n=1}^{\infty}\sigma^{\text{DLA}}_{n,\text{excl}}(p_{t,W}) \;=\; \sigma^{\text{LO}}_{1}(p_{t,W})\,, \tag{102}$$

i.e. in the double logarithmic approximation, there are no corrections to the $p_{t,W}$ distribution at high $p_{t,W}$. Now let us consider what happens if we expand each of the exclusive sums to NLO. For the n-jet cross section, we have

$$\sigma^{\text{NLO(DLA)}}_{n,\text{excl}}(p_{t,W}) \;\simeq\; \sigma^{\text{LO}}_{1}(p_{t,W})\,\frac{1}{(n-1)!}\left(\frac{2\,C\alpha_s}{\pi}L^2\right)^{n-1}\left(1-\frac{2\,C\alpha_s}{\pi}L^2\right). \tag{103}$$

Performing the sum over n, which corresponds to summing an infinite tower of NLO exclusive jet calculations, leads to

$$\begin{aligned}
\sigma(p_{t,W})^{\text{DLA}} \;&=\; \sum_{n=1}^{\infty}\sigma^{\text{NLO(DLA)}}_{n,\text{excl}}(p_{t,W}) && \text{(104a)}\\
&=\; \sigma^{\text{LO}}_{1}(p_{t,W})\exp\left(\frac{2\,C\alpha_s}{\pi}L^2\right)\left(1-\frac{2\,C\alpha_s}{\pi}L^2\right) && \text{(104b)}\\
&=\; \sigma^{\text{LO}}_{1}(p_{t,W})\left(1-\frac{1}{2}\left(\frac{2\,C\alpha_s}{\pi}L^2\right)^2+\mathcal{O}(\alpha_s^3L^6)\right). && \text{(104c)}
\end{aligned}$$

As long as L^2 is not large, the difference between this and the correct answer of Eq. (102) is a straightforward NNLO correction, i.e. small. However when $p_{t,W} \gg p_{t,\min}$ the logarithms become large, the $\alpha_s^2L^4$ term can be of order 1 and the Exclusive Sums method may then no longer be a good approximation. A similar analysis can be performed for an exclusive sum truncated at some finite order, as used in our study.

Given the above discussion, one may wonder then if there are any circumstances in which the Exclusive Sums method will bring benefits. For the observable studied in this contribution, H_T, the key difference with respect to $p_{t,W}$ is that it is subject to a 'giant' K-factor at NLO. This phenomenon is associated with 'dijet' topologies in which a soft or collinear W is radiated off the dijet system, leading to a double logarithmic (electroweak) enhancement. In addition these topologies can be created by qq type scattering (whereas the LO process involves only gq or $q\bar{q}'$ scattering), leading to further enhancement in pp collisions at large H_T. Dijet type topologies contribute significantly to the H_T distribution, even when the W is soft, because the variable sums all particles' transverse momenta (whereas the softness of the W limits these topologies' contribution to the $p_{t,W}$ distribution).

Because of the giant K-factor, for the H_T variable the behaviour of the Exclusive Sums method is more subtle than for $p_{t,W}$: while the $\sigma^{\text{NLO}}_{W+2}$ contributions destabilize the prediction for the $qg \to Wq'$ type topologies, they instead stabilize the prediction for the much larger $qq \to Wq'q$ topologies (present only at LO in a NLO $W+1$-jet calculation). Going further in the exclusive sum, however, i.e. including $\sigma^{\text{NLO}}_{W+3}$ and $\sigma^{\text{NLO}}_{W+4}$ contributions can however destabilize the predictions for both kinds of topologies. Traces of this behaviour were visible in the numerical studies shown above.

17. W PRODUCTION IN ASSOCIATION WITH MULTIPLE JETS AT THE LHC [34]

Abstract

We compare the results from four different theoretical predictions for the production of a W boson in association with at least two jets at the Large Hadron Collider. We discuss a possible method for combining next-to-leading order samples with different jet multiplicity from BLACKHAT+SHERPA. We then compare these results with the next-to-leading order W plus two jet calculation, the leading order ME&TS merged approach of SHERPA and the high-energy resummation approach of HIGH ENERGY JETS in an attempt to determine if these approaches can be distinguished at the LHC.

17.1 INTRODUCTION

The production of a W boson in association with jets at the Large Hadron Collider (LHC) is an extremely important process. It contributes to three distinct areas of the rich physics program at the LHC. Firstly, it is a key Standard Model signal and therefore important to test our understanding of the Standard Model in the TeV-scale energy range. Secondly, it is an important background in many searches for new physics where, for example, new heavy coloured particles have cascade decay chains. Thirdly, it provides an ideal testing ground for experimental techniques such as a jet veto: what is learned in the relatively well-understood treatment of W plus jets can be directly applied to Higgs searches for example.

It has been observed that the ratio of $W+(n+1)$-jet events to $W+n$-jet events can be substantially larger than one might naïvely expect by considering the α_s suppression only. This is especially true in phase-space regions of large four-momenta, such as the high-H_T tail, because the available phase space for extra jet emission at the LHC is extremely large. It can therefore compensate for the effect of an additional factor of the strong coupling. This effect is more visible in distributions where additional radiation leads to a significant change in the value of the observable, as is the case for the H_T distribution, the scalar sum of the transverse momenta of identified leptons, jets and missing energy. The change will be more moderate in an observable like $H_{T,2}$, whose definition differs from that of H_T by truncating the jet sum to include only the two hardest jets in the event. To make an impact here requires the radiation to lead to an additional jet with transverse momentum as large as that of the second hardest jet, not only larger than the jet p_T threshold. The effects will also be smaller in more inclusive variables like the transverse momentum of the W boson, $p_{T,W}$, or the leading jet, p_{T,j_1}.

There are a number of different theoretical approaches to describing the emission of large numbers of jets. In order to probe to what extent the differences in these will be accessible at the LHC, we will compare, in this study, the predictions for the jet activity in inclusive $W(\to e\nu) + 2$-jet production from (a) BLACKHAT+SHERPA (*BHS*) [70, 51, 22], (b) combined *BHS* samples (to be described below), (c) SHERPA [146, 425] run in ME&TS mode (*S-MEPS*) [426, 427, 428] and (d) HIGH ENERGY JETS (*HEJ*), an all-order resummation of wide-angle radiation [418, 419, 438].

The current state-of-the-art next-to-leading order (NLO) predictions for W production in association with jets are those of *BHS*, which have been calculated up to W plus four jets with a leading-colour approximation for the virtual part [22], and up to W plus three jets with a full color treatment [70, 51]. In this study, we consider the inclusive $W + 2$-jet prediction at NLO accuracy, and further, discuss and show predictions from an inclusive sample where $W + 2, 3, 4$-jet events generated by *BHS* are combined in a simple manner, nevertheless without introducing any double counting of phase-space regions.

The *S-MEPS* predictions are obtained from merging at leading order (LO) tree-level Matrix Elements for $W + 0, \ldots, n$-parton final states with (Truncated) parton Showers (hence the name ME&TS) preserving the leading logarithmic accuracy to which soft and collinear multiple emissions are described by the parton shower. The newer ME&TS merging scheme was introduced in Ref. [426] and optimised

[34] Contributed by: J. R. Andersen, D. Maître, J. M. Smillie, J. Winter

as documented in Refs. [427, 428] to improve over the original SHERPA implementation based on the CKKW approach [429, 430]. ME&TS guarantees a better matching regarding the usage of scales as occurring in the evaluation of the matrix elements and those scales driving parton showering. The *S-MEPS* sample used in our study was generated by including $W(\to e\nu)$ production matrix elements with up to five extra partons (massless quarks, u, d, s, c, b, and gluons).

The *HEJ* framework is a resummation of the leading logarithmic terms occurrung in pure, or W, Z or H plus, multi-jet production in the limit of large invariant mass between each pair of jets, to all orders in α_s. This is then matched to tree-level accuracy for final states with two, three or four jets. In principle, the *HEJ* framework can be merged with a parton shower to add the collinear pieces which are not included in the *HEJ* description (*HEJ* does include soft emissions down to around 2 GeV). First steps in this direction for pure jet production were taken in [431]. Here, the *HEJ* predictions are calculated at the parton level.

In the 17.2 section, we will elaborate on the method (b) for combining NLO samples of different jet multiplicities. Then, in the 17.3 section, we will first show explicit results of the sizable impact of large multiplicity events by comparing predictions from the combined *BHS* sample and the *S-MEPS* merged sample. Secondly, we will study variables chosen to probe the differences in the treatment of the QCD radiation. We will show and compare the predictions for all four descriptions mentioned above focusing on the following observables:

- the average number of jets as a function of $H_T = \sum_i p_{T,j_i} + p_{T,e} + p_{T,\nu}$ and Δy, the rapidity difference between the most forward and most backward jets, and also
- the ratio of the inclusive 3-jet rate to the inclusive 2-jet rate as a function of H_T and Δy.[35]

We will then discuss the areas of agreement and difference that we find, before we finally conclude in the 17.4 section.

17.2 NLO EXCLUSIVE SUMS

An NLO n-jet prediction contains events with n or $n+1$ partons. For observables for which higher multiplicites have a significant impact, this limitation can be detrimental. If one has predictions for different multiplicities $(m, m+1, \ldots, M)$, one can try to combine them by avoiding double counting by requiring that the n-jet prediction is used only to describe n-jet events (except for the highest multiplicity where $(n+1)$-jets configurations are allowed). The total cross section can be rewritten as a decomposition based on exclusive (exc) and inclusive (inc) jet bins:

$$\sigma^{\text{tot}} \equiv \sigma_m^{\text{inc}} = \sum_{n=m}^{M-1} \sigma_n^{\text{exc}} + \sigma_M^{\text{inc}} . \tag{105}$$

The exclusive-sums procedure describes each jet bin at NLO accuracy, i.e. at $\mathcal{O}(\alpha_s^{n+1})$, or, alternatively, only the $(M+1)$-th (inclusive) jet bin is predicted with LO precision. We hence note that the combination of the terms shown in Eq. (105) occurs at different orders of the strong coupling. Furthermore, the definition of an exclusive n-jet sample requires a detailed treatment of jet vetoing. For these reasons, the simple combination procedure is crude and does not increase the formal accuracy of the prediction, which is that of NLO of the smallest multiplicity. However, one can hope that the procedure will lead to a better prediction in observables where higher multiplicity events dominate.

More studies are required to understand the uncertainties related to this procedure. One way of doing so would be to vary the jet algorithm and/or parameters of the jet algorithm used to separate the different NLO predictions into fixed multiplicities sets and test the stability of the prediction.[36] This is left to a future study.

[35] Given the definition of H_T, we note $H_{T,2} = p_{T,j_1} + p_{T,j_2} + p_{T,e} + p_{T,\nu}$.

[36] In this study we used the same jet algorithm for both defining the partitioning in jet multiplicities as well as applying cuts and determining observables.

$\lvert\eta_e\rvert < 2.5$	$p_{T,e} > 20$ GeV
$M_{\perp,W} > 20$ GeV	$p_{T,\nu} > 20$ GeV
$\lvert\eta_j\rvert < 4.5$	$p_{T,j} > 25$ GeV

Table 12: Summary of the cuts applied in the analysis.

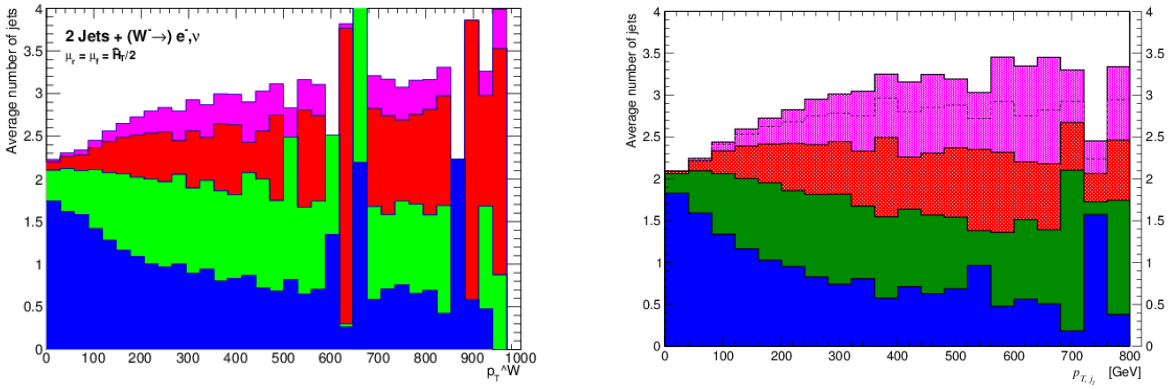


Fig. 56: The average number of jets as a function of $p_{T,W}$ (left) and p_{T,j_1} (right). The $p_{T,W}$ plot shows the *BHS* exclusive sums prediction, while the p_{T,j_1} plot is obtained from *S-MEPS*.

17.3 RESULTS OF THE COMPARISON

In this section, we compare the results of different theoretical descriptions for $W + n$-jets production at the LHC. The number n can take values from 2 and above, as we will mostly consider inclusive samples. The four descriptions, which we will compare here in more detail, are

- the *BHS* calculation of $W + 2$-jets at NLO,
- the combined sample of $W + 2, 3, 4$-jet events at NLO from *BHS*, as described in the 17.2 section,
- the *S-MEPS* merged $W + n$-jets sample using LO tree-level matrix elements up to $n = 5$, and
- the approach of *HEJ*.

Throughout this study, we will consider inclusive samples of W^- boson production in association with at least two hard jets identified by the anti-k_T jet algorithm using $R = 0.4$. The jets are required to have $p_{T,j} > 25$ GeV. We look only in the $W^- \to e^-\bar{\nu}_e$ decay channel and use the cuts given in Tab. 12 where $M_{\perp,W}$ is defined as $M_{\perp,W} = \sqrt{(|\vec{p}_{T,e}| + |\vec{p}_{T,\nu}|)^2 - (\vec{p}_{T,e} + \vec{p}_{T,\nu})^2}$.

The *HEJ* predictions use the geometric mean of the jet transverse momenta to determine the renormalisation and factorisation scale, i.e. $(\prod p_{T,j})^{1/n}$. This central choice will be varied by a factor of two in either direction to provide an envelope (marked by dotted lines in the corresponding figures) around the *HEJ* default prediction. The *BHS* predictions instead use $\hat{H}'_T/2$ as the NLO calculation becomes unstable for a scale which is too low. In the *S-MEPS* calculation, scales are chosen according to the default prescription given by ME&TS [426].

The variables $H_{T,2}$, $p_{T,W}$ and p_{T,j_1} are less sensitive to the presence of additional radiation than H_T, as discussed in the introduction. The plots, which we present in Figs. 56 and 57 address the alternative question: given a particular value of H_T, $H_{T,2}$ etc. how many jets are typically found in the event?

Figs. 56 and 57 show the stacked results for the average number of jets as a function of $p_{T,W}$, p_{T,j_1}, H_T and $H_{T,2}$ visualising the contributions from each exclusive $2, 3, 4$-jet sample and the inclusive 5-jet

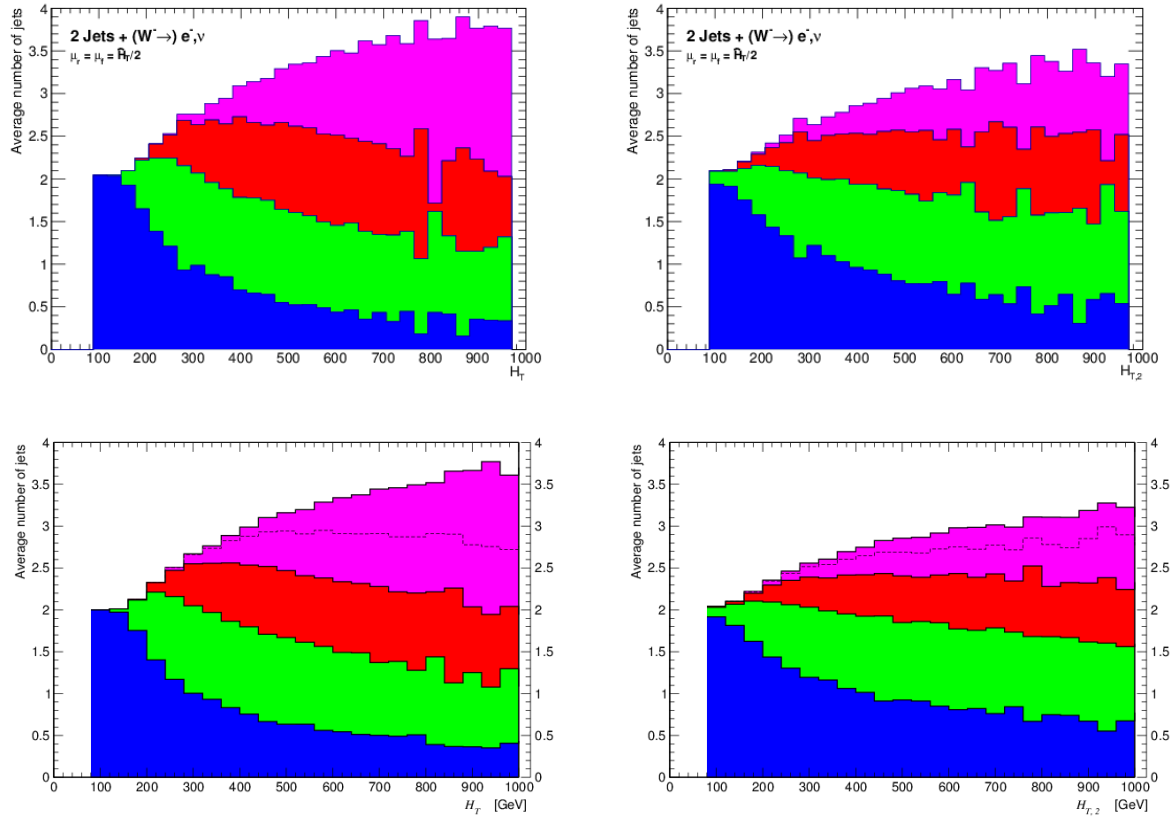


Fig. 57: The contribution from different multiplicities to the average number of jets as a function of H_T and $H_{T,2}$. The upper plots show the *BHS* exclusive sums prediction, while the lower ones are extracted from *S-MEPS*.

sample. The left (right) plot in Fig. 56 and the upper (lower) rows of plots in Fig. 57 depict the results as obtained from the combined *BHS* sample (the *S-MEPS* sample). In all cases the different colours correspond to the terms in the numerator of the formula for the average number of jets,

$$\langle N\rangle_5 \;=\; \frac{\sum\limits_{i=2,3,4} i\, n_i^{\rm exc} + 5\, n_5^{\rm inc}}{\sum\limits_{i=2,3,4} n_i^{\rm exc} + n_5^{\rm inc}} \;=\; \frac{\sum\limits_{i=2,3,4} i\, n_i^{\rm exc} + 5\, n_5^{\rm inc}}{n_2^{\rm inc}}\,, \tag{106}$$

where blue, green, red and magenta stand for $i = 2, 3, 4$ and $i = 5$, respectively. The subscript to $\langle N\rangle$ clarifies that we truncate the determination of the average after the fifth jet bin, noting that $\langle N\rangle_k \to \langle N\rangle$ for a sufficiently large number of jet bins. This makes no difference for the *BHS* predictions employed here since the jet multiplicity de facto is limited to five, but it does for the *S-MEPS* and *HEJ* computations where events with $i > 5$ jets do occur. We have defined $n_k^{\rm exc/inc} = d\sigma_k^{\rm exc/inc}/dO$ where O denotes an observable like H_T, or Δy presented later on. Note that in Fig. 57 the 5-jet part contributes to the average number of jets with a factor of 5, while the 2-jet part, for example, contributes with a factor of 2 only.

The layout of Fig. 58 (including the colour coding) is the same as before: here, we however display, wrt. $n_2^{\rm inc}$, the relative fractions of the different multiplicities corresponding to the terms in the denominator of Eq. (106). In other words, in Fig. 58 we consider the partitioning of

$$1 \;=\; \frac{\sum\limits_{i=2,3,4} n_i^{\rm exc} + n_5^{\rm inc}}{n_2^{\rm inc}}\,. \tag{107}$$

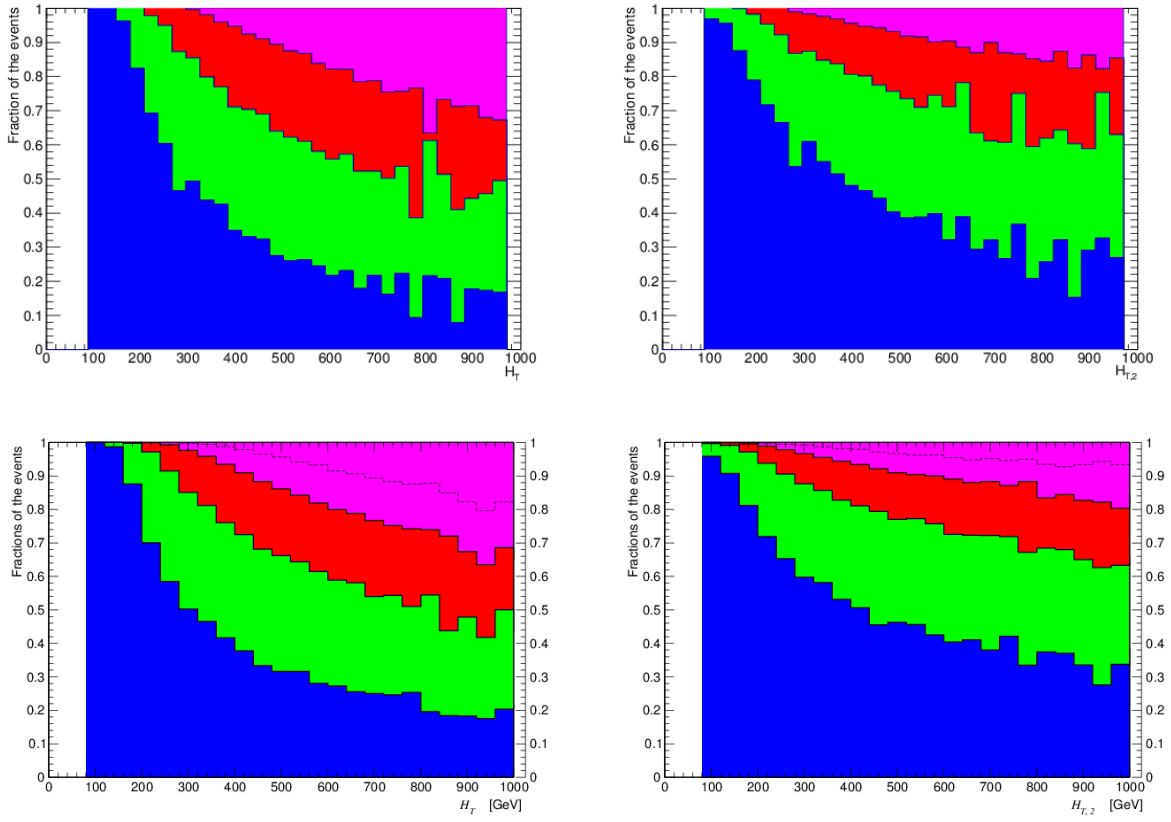


Fig. 58: The fraction of the total rate from different multiplicities as a function of H_T and $H_{T,2}$. The upper plots show the *BHS* exclusive sums prediction, while the lower ones are extracted from *S-MEPS*.

Although there is just a 30% fraction of inclusive 5-jet events to the total cross section, we observe that their contribution to the build-up of $\langle N \rangle(H_T)$ for very large H_T gets close to 50%. Also, for an $H_T \sim 500$ GeV, the average number of jets is composed evenly between the $2, 3$-jet and $4, 5$-jet contributions, while the relative fraction of the $2, 3$-jet events is nearly 70%. This emphasizes the dominance of multi-jet events in forming large H_T values. It also can be seen that for medium H_T values, $400 < H_T < 700$ GeV, all the multiplicities give roughly the same contribution to the variable $\langle N \rangle(H_T)$, while for low H_T, the average is primarily described by 2-jet events.

Going clockwise through Figs. 56 and 57 we see that the average number of jets is indeed sensitive to higher multiplicities when considered as a function of $p_{T,W}$, p_{T,j_1} and $H_{T,2}$, but in all these cases this happens to a lesser extent as if considered as a function of H_T. As expected, the dependence is mildest for $p_{T,W}$, the most inclusive observable studied here. We also observe that the jet-bin decomposition of p_{T,j_1} and $H_{T,2}$ turns out very similar. Most strikingly we note the increase in the contribution from the highest multiplicity events, the ones containing more or at least five jets. For $H_{T,2}$, we furthermore display to the right of Fig. 58 the relative fractions as done in the H_T case. Even for largest $H_{T,2}$ values, the fraction arising from $2, 3$-jet events remains close to 65% stressing once more the lower sensitivity of $H_{T,2}$ versus H_T regarding multiple jet production.

Finally, we compare the plots from the combined *BHS* samples in all figures to the corresponding ones generated with the *S-MEPS* sample. Interestingly, the outcome looks very similar although ME&TS handles the single terms in Eq. (105) rather differently. They are calculated at least at leading (soft/collinear) logarithmic accuracy improved by LO n-jet effects. Presumably, for the exclusive jet bins, this description (which allows a better treatment of jet vetoes) is not too far off the exclusive sums

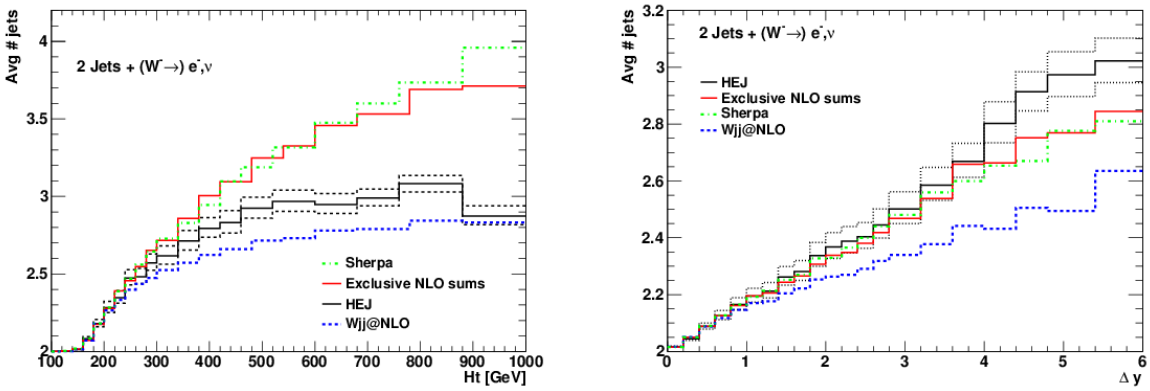


Fig. 59: Average number of jets as a function of H_T (left) and Δy (right) in two *BHS* descriptions, from *HEJ* and from *S-MEPS*, the latter using the $\langle N \rangle_7$ definition. The bands shown with dotted lines for the *HEJ* prediction are a result of varying the scale by a factor of 2 in each direction.

approach, since the unresolved $\mathcal{O}(\alpha_s)$ corrections are also present in the Sudakov form factors applied in the ME&TS approach. Also, the combined *BHS* samples as well as the *S-MEPS* sample use the same tree-level matrix elements, namely up to $W + 5$-parton matrix elements. Clearly, it has to be studied further whether this similarity in the results is a coincidence or not.

It is clear that the impact of the higher multiplicity samples is significant throughout, especially in the high H_T tail. This is precisely the region, which would be probed for signs of new physics, and therefore it is essential that we fully understand our theoretical descriptions in this region. This is the subject of the remainder of this contribution, where we compare all four different methods of modelling hard QCD radiation in inclusive $W + 2$-jet events.

The left plot of Fig. 59 shows the final comparison plot between the exclusive sums and inclusive 2-jet *BHS* results as well as the *HEJ* and *S-MEPS* predictions for the average number of jets as a function of H_T. The differences in the descriptions are significantly larger than the scale uncertainty band on the *HEJ* prediction. For the $W + 2$-jet NLO result, the number of jets rises to 2.6 already at $H_T = 500$ GeV but that levels off significantly below the *S-MEPS*, exclusive *BHS* sum and *HEJ* results. The *HEJ* results level off at a higher value of about 3.0, starting to clearly disagree with the exclusive sums and *S-MEPS* predictions above 500 GeV, from where those two curves keep rising to a final level of around 3.7 to 4.0. The *S-MEPS* comes in highest at largest H_T, where $\langle N \rangle_7$ is shown, cf. Eq. (106), in order to determine the average number of jets for this *S-MEPS* result. The reason for giving slightly higher $\langle N \rangle$ than the exclusive sums lies in the contribution of additional parton-shower jets present in the *S-MEPS* calculation and more accurately accounted for by the use of the $\langle N \rangle_7$ definition as compared to the earlier result based on $\langle N \rangle_5$ presented in Fig. 57 to the lower left.

In the right panel of Fig. 59, we have plotted the average number of jets as a function of the rapidity span, Δy, instead of H_T as before. Again the differences are larger than the scale variation shown on the *HEJ* result, but the ordering is different to that of the left plot of Fig. 59. All four descriptions increase linearly with Δy but the gradient is steepest for the *HEJ* predictions where the average rises above 3.0 for Δy values as large as 6.0. The *BHS* exclusive sum result is consistently below this, reaching about 2.8 at $\Delta y = 6.0$, and agrees pretty well with the *S-MEPS* result based on $\langle N \rangle_7$. The NLO $W + 2$-jet prediction given by *BHS* is lower still, between 2.4 and 2.5 for $\Delta y \sim 5.0$.

It may seem surprising that on the plot on the left-hand side the exclusive sums and *S-MEPS* lie higher for most of the distribution whereas on the right-hand side these approaches as well as *HEJ* give predicitions that are commensurate. The region of high H_T and that of high Δy however are largely distinct as it is very expensive to have both a large rapidity and large p_T for the jets. Also while radiating

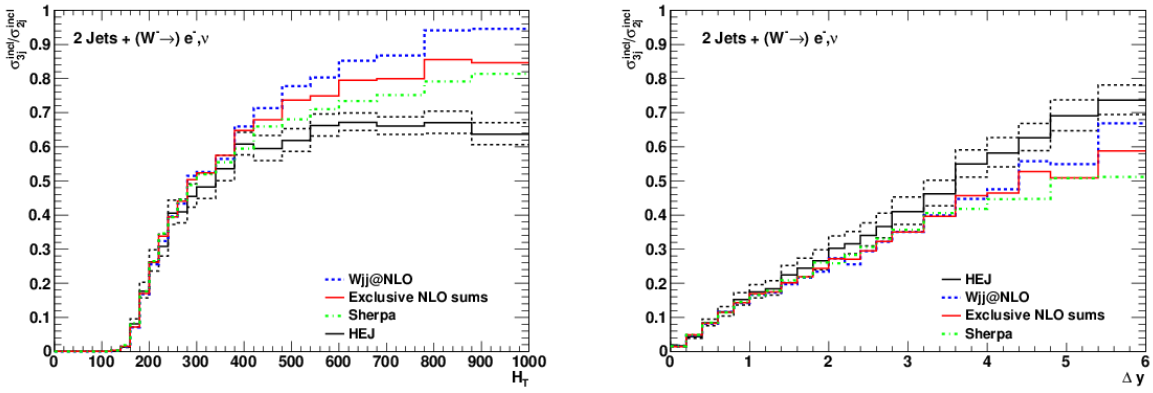


Fig. 60: The ratio of the inclusive 3-jet and 2-jet rates in the inclusive W + 2-jet NLO and exclusive sum description of *BHS* as well as in the *S-MEPS* and *HEJ* approaches as a function of H_T (left) and Δy (right). Again, the dotted lines indicate the uncertainty band from varying the scale in *HEJ* by a factor of 2 in each direction.

an additional jet automatically moves an event towards the higher H_T direction, radiating an additional jet tends to not change the rapidity difference. So, we expect the higher multiplicities to have a smaller effect on the average number of jets as a function of Δy compared to as a function of H_T. This is indeed the case in Fig. 59.

Lastly, in Fig. 60 we plot the ratio of the inclusive 3-jet to the inclusive 2-jet rate as a function of H_T (left) and Δy (right), again for all four descriptions used here. The predicted $(d\sigma_3^{\text{inc}}/dH_T)/(d\sigma_2^{\text{inc}}/dH_T)$ all agree very well below 400 GeV. The fixed order *BHS* result for $W + 2$ jets is highest for large H_T, however is known to become unreliable here, since the probability that an inclusive 2-jet event is at least a 3-jet event turns too large, being in conflict with the expected behaviour of an $\mathcal{O}(\alpha_s)$ correction. The *BHS* exclusive sums, the *S-MEPS* and the *HEJ* results, in this order, level off considerably lower with the *HEJ* fraction staying below 60% to 70%, which leaves the other predictions again above the *HEJ* uncertainty envelope. In contrast, when the same ratio of jet rates is plotted against Δy, the *HEJ* prediction is consistently higher throughout. This again emphasises that differences in the descriptions come to light in different kinematic regions. However, in both cases here the magnitude of the differences is relatively small and would be rather difficult to distinguish in present experimental data.

17.4 CONCLUSIONS

We have compared a number of theoretical descriptions of W^- production in association with at least two jets. After outlining one possible method of combining NLO calculations of different multiplicities, we compared this with a pure NLO calculation of $W+2$-jets production obtained by BLACKHAT+SHERPA, a sample of leading-order events merged using the ME&TS method of SHERPA, and the high-energy resummation of the HEJ framework.

We studied the average number of jets and the ratio of the 3-jet and 2-jet inclusive cross sections as a function of Δy and of H_T. We find, with these simple cuts, some clear differences in the predictions when we study the average number of jets as a function of both Δy and H_T. Smaller differences, which would be more difficult to disentangle experimentally, are found when we study the ratio of inclusive rates.

It would be very valuable to have an experimental study, which probed the average number of jets in W production in association with at least two jets, to test our different descriptions of these important Standard Model processes.

ACKNOWLEDGEMENTS

The authors would all like to thank the organisers and participants for the extremely stimulating workshop. DMs work was supported by the Research Executive Agency (REA) of the European Union under the Grant Agreement number PITN-GA-2010-264564 (LHCPhenoNet). JMS is supported by the UK Science and Technology Facilities Council (STFC).

18. UNCERTAINTIES IN THE SIMULATION OF W+ JETS – A CASE STUDY [37]

Abstract

In this contribution, uncertainties in the simulation of a large variety of observables related to the production of W in association with jets at the LHC and the Tevatron are discussed. This work aims to

- serve as a compendium of currently publicly accessible tools in addition to the ones presented in a previous publication [439] with a similar topic, and to compare their results;
- discuss the origin and generic size of various uncertainties in the simulation of perturbative and non-perturbative aspects of this process;
- trace the interplay of these uncertainties in various stages of the full event simulation;
- hint at those uncertainties in each of the various tools considered here which the respective authors find relevant;
- guide their users in how to assess the related uncertainties in a way the authors recommend.

18.1 Introduction

The production of W-bosons in association with jets constitutes an important process at the Tevatron and the LHC, for a variety of reasons. First of all, it represents a major background to Standard Model signatures such as top-pair and single-top production,and it also plays a role in searches for the Higgs boson in the Standard Model. Furthermore, this reaction, together with the fairly similar channel of Z-production in association with jets, provides one of the most important backgrounds in those searches for new physics where large missing transverse energy and high jet multiplicities characterise the respective signal. Thirdly, this process has become a standard reaction for QCD studies at hadron colliders, ranging from the validation of simulation tools for multijet signatures to measurements related to multiple parton scattering. Finally, this process also provides one of the main testbeds for novel techniques in the automation of higher-order QCD corrections and their matching or merging with subsequent parton showers in the framework of event generators.

In the spirit of this last point, providing a testbed for the combination of fixed order calculations with the parton shower, this process has been analysed in quite some depth in [439] about five years ago. A number of reasons provide motivation to update and extend this previous study, namely

- the LHC being up and running and starting to provide highly precise data such that a proper treatment of uncertainties becomes an important issue;
- major improvements in the ability to calculate higher-order corrections including up to four jets in the final state accompanying the W bosons [51, 117, 22];
- the advent of such next-to leading order calculations – albeit for lower final state multiplicities – fully matched to the parton shower [440, 404, 424];
- an improved understanding of the leading order merging prescription for towers of multijet multiplicities with the parton shower [426, 441];

[37] Contributed by: S. Alioli, J. R. Andersen, V. Ciulli, F. Cossutti, T. Hapola, H. Hoeth, F. Krauss, P. Lenzi, L. Lönnblad, G. Luisoni, D. Maître, C. Oleari, S. Prestel, E. Re, T. Reiter, M. Schönherr, J. Smillie, F. Tramontano, J. Winter, K. Zapp

- the combination of matching and merging methods [442, 443];
- and new methods to simulate multijet topologies based on the high-energy limit [64].

Therefore this study aims at being a first step towards a more complete update of [439], with a shift in focus towards a discussion of theoretical uncertainties in different approximations, including perturbative and non-perturbative effects. Apart from tracing the origin and determining the generic size of various uncertainties in the theoretical description of various observables related to this process, also the interplay of them at various stages of the simulation, from the matrix element to the hadron level will be discussed. Consequently, the most important causes for theory uncertainties in various tools are highlighted. Therefore, one of the more practically relevant goals is to also provide methods to reliably and robustly estimate such uncertainties for the various tools used in this study, as recommended by authors or users.

The outline is as follows: After briefly presenting the various tools included in this study and discussing the way they have been used here in Sec. 18.2, example results for them will be presented individually, tool by tool in Sec. 18.3. In Sec. 18.4 these results are compared in order to see and quantify relative differences. In this endeavour, experimental results have not yet been included. We reserve this comparison with relevant data for a later, full-fledged analysis, which will hopefully include even more tools.

18.2 Codes

In this work a variety of different codes has been employed, which allow to study the process at various different stages:

1. Fixed order matrix elements:
   By now, the description of W boson production in association with jets is possible for up to 4 additional jets at NLO. Here, results from two NLO codes, GoSam+Sherpa [12, 425, 146] and BlackHat+Sherpa [70, 51, 22], which are either publicly available or provide publicly available event files, are presented. The corresponding results therefore are on the matrix element level.
2. All-order resummed matrix elements:
   Approximations to the partonic matrix elements for the processes of n-jet production, and $W, Z, H + n$-jets, $n \geq 2$, was recently calculated to any multiplicity, and including all-order resummations for the leading virtual corrections. The all-order scheme [418, 419], implemented in the HEJ [64] code, becomes exact in the limit of large invariant mass between each parton (the MRK limit of BFKL). The resummation scheme is merged with LO matrix elements (much like in MEPS, see later). The resummation of HEJ can also be interfaced to a parton shower [431]; the results presented here, however, are on the matrix element level. It should also be stressed that due to the nature of the approximation of HEJ, the simulation here are relevant for the production of *at least* two jets in addition to the W boson.
3. Parton showers:
   The pure parton shower code relies on the collinear approximation to produce additional jets. By using a matrix element reweighting, however, in the process of W production, typically one additional jet can correctly be described. For this simulation, Pythia8 [348] has been used here, with results available on the parton shower level, hadron level and hadron level including UE.
4. LO matrix elements merged with the parton shower (MEPS):
   By now, the use of towers of multijet matrix elements with increasing multiplicity merged to the parton shower following ideas presented in [444, 429, 445, 430] is common practise in the experimental collaborations. In fact, a first comparison of different codes and implementation has been presented a while ago [439]. Here, three implementations of these ideas are included, namely the ones in MadGraph+Pythia [165, 446, 163, 191, 400], Pythia8+ME [441] and Sherpa [146]. Here results are available on all levels matrix element level, parton shower level, hadron level,

hadron level including UE, and hadron level including UE and QED final state radiation in different combinations of codes.

5. NLO matrix elements matched to the parton shower (NLO⊗PS and MENLOPS):
   In principle two methods by now have been proposed and fully implemented which consistently match full NLO calculations to the parton shower, namely MC@NLO [401] and POWHEG [416, 417]. Here the latter is being used, with its implementation in the POWHEG BOX [13], and interfaced to the PYTHIA [400] parton shower in its k_T-ordered version [447]. In addition, a combination of such matching with the merging methods described in the previous point is available [442, 443], ranging under the name MENLOPS. In this paper we use an implementation of such methods provided in the SHERPA framework. In both cases, results are available on all levels matrix element level, parton shower level, hadron level, and hadron level including UE.

18.21 BLACKHAT + SHERPA

The NLO predictions are obtained by combining BLACKHAT [4] for the virtual part and SHERPA [147, 448] for the real part. It is currently possible to obtain predictions at NLO for a W-boson in combination with up to four jets [70, 51, 22].

The plots have been produced by re-analysing large event files produced by the combination of BLACKHAT and SHERPA. These files contain particle four-momenta as well as the coefficients of all scale dependent functions, including the PDFs so that it makes it possible to easily change factorisation and renormalisation scales as well as the PDF set.

We used a common factorisation and renormalisation scale $\mu_F = \mu_R = \hat{H}'_T/2$ with $\hat{H}'_T = \sum_j p_T^j + E_T^W$ where the sum runs over all jets and $E_T^W = \sqrt{M_W^2 + (p_T^W)^2}$.

Estimation of uncertainties The estimation of the uncertainties for the NLO calculation obtained with BLACKHAT+SHERPA is obtained by combining in quadrature the pdf uncertainties obtained using the pdf error set and the uncertainties obtained by varying the factorisation and renormalisation scales simultaneously by factors of 1/2 and 2. To this error we also add in quadrature the integration error estimate. Another way of estimating the uncertainties due to the choice of scales is to compare predictions obtained using different choice of basis scales, but this has not been done for this study.

We used the CTEQ6.6 PDF set. The value of α_s used for this calculation has also been taken as that provided with this PDF set. The PDF uncertainties are estimated using the hessian method and PDF 'error' set provided with the CTEQ6.6 PDF set.

18.22 GOSAM + SHERPA

GOSAM [12] is a new framework which allows the automated computation of one-loop scattering amplitudes for multi-particle processes. The one-loop scattering amplitudes are generated in terms of algebraic d-dimensional unintegrated amplitudes, which are obtained via Feynman diagrams. This allows to perform symbolic manipulations of the expressions prior any numerical step. For the reduction, the program offers the possibility to use either a d-dimensional extension of the OPP method [121, 122, 119], as implemented in SAMURAI [6], or tensor reduction as implemented in golem95 [130, 131] interfaced through tensorial reconstruction at the integrand level [124].

The GOSAM framework can be used to calculate one-loop corrections within both QCD and electroweak theory. Beyond the Standard Model theories can be interfaced using FEYNRULES [137] or LANHEP [136].

To produce results for a certain process specified by the user, the program must be fed with an "input card" with the details of the process. Alternatively, when interfacing the program with a Monte

Carlo (MC) event generator which supports the Binoth-Les-Houches-Accord (BLHA) interface [145], the specific order file produced by the MC event generator can be passed to GoSam.

The analysis presented here was performed using this latter generation mode and Sherpa [425, 146] was chosen as MC event generator. Sherpa provides therefore the matrix elements for the production of W and exactly one jet at the Born-level and the NLO real corrections to it, together with the needed subtraction terms and their integrated counter-parts. GoSam provides the NLO virtual-part. The generation of the code follows the standards of the BLHA-interface [145]. During the first call of Sherpa an "order file" is written by the MC program. This file is read-in by GoSam to produce the code for the one-loop evaluation of needed process. If this happens successfully, a contract file with information on the different possible subprocesses is produced by GoSam and can be later read by the MC generator to recognize the numbering of the different partonic subprocesses. At running time all information between GoSam and Sherpa is also passed using the BLHA-interface standards.

The steering of the event generation and the analysis interface with Rivet [360] is done using Sherpa cards. Each curve in the analysis consists of 100 combined runs of 50 million events. The renormalisation and factorisation scales are set according to the choice made for this analysis in Les Houches to

$$\mu_F = \mu_R = \hat{H}'_T/2\,,$$

where $\hat{H}'_T$ is defined in the previous section.

Estimate of uncertainties The estimation of the uncertainties for the NLO calculation obtained with GoSam+Sherpa is done combining in quadrature the PDF uncertainties with the uncertainty coming from the separate variation of factorisation and renormalisation scale by factors of 1/2 and 2. Ideally also the integration error should be added in quadrature to the previous estimate, however the MC integration error obtained with Rivet at NLO is not reliable because of the incapacity of Rivet to take into account properly the correlation between real and subtraction events. For this reason and because of the very high statistics of the MC sample, the MC integration error is neglected. To assess the PDF uncertainty we compute the envelope of the results obtained using the three different PDF sets CT10 [255], used as nominal set, MSTW08 [262] and NNPDF2.1 [312]. The total scale uncertainty is determined by adding in quadrature the factorisation and renormalisation scale uncertainties. Each of them is found by computing the maximum between the nominal value and the up and down variations.

18.23 HEJ

The *High Energy Jets* (Hej) framework [418, 419] provides an alternative description of collider events to the standard fixed order calculations (possibly interfaced to a parton shower). Instead, Hej uses approximations to the hard scattering matrix element to all orders in α_s which become exact in the High Energy limit. The approximation results in sufficiently simple matrix elements, that these can be explicitly regulated, integrated and summed over any (relevant) multiplicity. This results in an explicit *all-order* resummation of the dominant contributions from wide-angle QCD radiation.

The building blocks of the Hej framework ensure the correct leading logarithmic behaviour in the Multi-Regge Kinematic limit (aka. the High Energy Limit) of large invariant mass between all partons, for both the real and virtual corrections. The resummed n-jet rate is then further matched to tree-level accuracy for events with up to and including four jets, using a merging procedure for the soft radiation.

This procedure has so far been applied to the production of jets [64], W plus jets [449], Z plus jets and Higgs boson plus jets and has currently been implemented in a fully flexible Monte Carlo for the first two of these processes. The implementation integrates explicitly over any number of QCD emissions from a (W, Z, H+) dijet system, and hence produces event samples for processes with two jets or more. Note that one has access to the momenta of all final state particles for every event and it is therefore extremely simple to restrict to a subset of the events if required, e.g. 3-jet exclusive events.

The HEJ resummation includes emissions at large transverse momentum which are increasingly important as the centre-of-mass energy of particle collisions increases. HEJ is currently the only available flexible Monte Carlo generator to obtain leading logarithmic accuracy in the limit of large invariant mass between emissions. However, the HEJ framework does not include any systematic resummation in the collinear limit. This is included in a parton shower, but a careful merging procedure is required to link one with HEJ, as there is significant overlap between the soft emissions included in each approach; the first steps in this direction have been taken for jet production [431] and are ongoing. In the current study though, only parton level predictions are given.

Estimate of Uncertainty The HEJ framework does not contain any tunable parameters other than the choice of renormalisation and factorisation scale (just like any fixed order calculation). In this study, in common with other approaches, we choose both of these to be given by the geometric mean of the transverse momentum of the jets:

$$\mu_R \;=\; \mu_F \;=\; \left(\prod_{j=1}^{n} p_T^j\right)^{1/n} \quad , \tag{18.2.1}$$

where the jets are defined according to the relevant cuts in each analysis. This is however only an arbitrary choice, as the framework admits any choice for the scale, including H_T, p_T of the hardest jet and a fixed scale. For a given scale, α_s is evaluated according to the relevant PDF.

In common with standard convention, we calculate the scale variation by changing this scale by a factor of two in both directions. In principle, one could also include the PDF uncertainty, but this is not done in this study (as the scale uncertainty dominates). As described above, HEJ contains matching to tree-level accuracy for up to four jets. However, unlike the merging procedure in a showered sample, the merging scale here is not a free parameter. There is only one rational choice for the merging scale: the minimum p_T of a jet in the relevant analysis. However, in an inclusive sample with at least two jets, one could use as a further estimator of uncertainty the variation obtained when matching to three and four jet LO matrix elements. This procedure will be studied in detail in Ref. [449], but in the present study, we quote the uncertainty only from the scale variation.

18.24 MADGRAPH + PYTHIA

MADGRAPH [165, 446, 163, 191] is a general purpose leading order matrix element (ME) generator, with a broad variety of models available and easily extensible thanks to its modular structure. The event generation is performed by the MADEVENT component, a tool implementing the Single Diagram Enhanced algorithm for multi-channel phase space integration. When a user provided process is specified, MADGRAPH automatically generates the amplitudes for all the relevant subprocesses and produces the mappings for the integration over the phase space. This process-dependent information is then used by MADEVENT, where the process specific code generated allows the user to calculate cross sections and to produce unweighted events. Once the parton level events have been generated, a traditional parton shower (PS) Monte Carlo library can be run on top of the MADGRAPH output to describe additional QCD radiation, and possibly allow to produce hadron level generated events if a suitable hadronisation model is then applied.

In order to avoid double counting of QCD radiation from the matrix element and the parton shower, the MLM matching approach is used in its ktMLM implementation provided by the MADGRAPH team [191].

For the present study MADGRAPH-5.1.1 [191] has been used for the matrix element generation, while the parton shower and hadronisation has been provided by PYTHIA 6.4.2.4 [400]. The $W + n$ jet process has been simulated up to 4 additional partons. The PDF used in both calculations has been

CTEQ6L1, and in the matrix element calculation the strong coupling costant has been setup to be equal to the one from the PDF used. The factorisation scale and the hadronisation scale are set to the W transverse mass, $m_{\perp,W}$. The parton level clusterisation scale *xqcut* has been set to 10 GeV, while the ME - PS matching scale *qcut* has been set to the optimal value of 20 GeV, determined ensuring the smoothness of the differential jet rate.

The PYTHIA settings have been defined according the so called Tune Z2, an adjustment of Tune Z1 described in [450] for CTEQ6L1, where the $p_\perp$ cutoff for the multiple parton interactions is set to `PARP(82)=1.832` obtained on top of LHC data as far as the underlying event and multiple parton interactions are concerned, while the fragmentation parameters are those optimized on LEP data by the PROFESSOR [451] team.

Estimate of uncertainties To estimate the uncertainties due to the factorisation scale and the renormalisation scale, which are set to $m_{\perp,W}$, we varied them simultaneously by a factor two. In addition we have independently varied by a factor two the ME - PS matching scale.

While applying these modifications, the total cross-section is kept fixed to the value obtained with the default parameters, 27.77 nb. That is because we are only interested in shape variations of the distributions, rather than in the total cross-section of the process calculated by MADGRAPH, which is accurate only at the leading order.

18.25 POWHEG BOX + PYTHIA8

The POWHEG BOX [13] is a computer framework to ease the POWHEG [416] implementation of new processes. It only requires as input the individual components of the NLO calculation under consideration, *i.e.*, the Born process, its virtual radiative corrections and the real emission contributions. Then it automatically combines them, canceling the emerging soft and collinear singularities in the Frixione-Kunszt-Signer (FKS) subtraction scheme, and produces the required events. The POWHEG BOX is also a library, where previously implemented processes are available in a common framework.

For the present study we make use of the W+jet implementation presented in [440].

The produced events are passed to PYTHIA8 [348] through the Les Houches interface [452] and showered with the default transverse-momentum ordered shower, vetoing further emissions harder than the one already present in the input events. This is achieved by setting the starting scale of the shower as the transverse momentum of the hardest emission.[38]

When multiple partonic interactions (MPI) are turned on, these are allowed to be harder than the first POWHEG emission. Indeed, since the $W + 1$ jet process is not accounted for in MPI, there is no over-counting.

Eventually, the relevant distributions are evaluated by interfacing the MonteCarlo output to the RIVET [360] analysis, for the two given sets of ATLAS and CMS cuts.

Since we have simulated events starting from a hard process where a W is produced in association with one jet, only observables built from events where at least 1 jet is present will be shown.

Generation of predictions and estimate of uncertainties Predictions presented here are based on a merged sample of 4M $W^+ + j$ and 4M $W^- + j$ weighted events, produced with the default POWHEG BOX choice of parameters. In particular, we have required a minimum cut $p_{\mathrm{T}} = 5$ GeV on the associated jet at the generation level and, in order to enhance the statistical sampling of the high-p_{T} tail, we have further suppressed the rapidly rising contribution at low jet p_{T} by the factor $p_{\mathrm{T}}^2/\left(p_{\mathrm{T,supp}}^2 + p_{\mathrm{T}}^2\right)$, with $p_{\mathrm{T,supp}}^2 = 100$ GeV. The inverse of this factor enters the event weight.

[38] An extra veto may be required at this stage, due to the different definitions of the transverse momentum used in the POWHEG BOX – either for initial or final state radiation – and in PYTHIA8.

We have adopted the POWHEG BOX default values for EW parameters, namely

$$M_W = 80.398 \text{ GeV}\,, \quad \Gamma_{\rm W} = 2.141 \text{ GeV}\,, \quad (\text{ff}_{\rm em})^{-1} = 128.89\,, \quad \sin^2 \text{`}_{\rm W} = 0.222645 \tag{18.2.2}$$

and we have assumed a CKM matrix with a mixing between the first two generations only

$$|V_{ud}| = |V_{cs}| = 0.975\,, \quad |V_{us}| = |V_{cd}| = 0.222\,, \text{ and } |V_{tb}| = 1\,. \tag{18.2.3}$$

Finally, we have resctricted the integration region to the interval $0 < M_W < 2221$ GeV.

For the computation of the POWHEG $\bar{B}$ function, the renormalisation and factorisation scale was chosen equal to

$$\mu_R = \mu_F = p_{\perp,j}\,, \tag{18.2.4}$$

where $p_{\perp,j}$ corresponds to the transverse-momentum of the (single) parton recoiling against the W boson in the so-called underlying Born kinematics [417]. We have also run the code using

$$\mu_R = \mu_F = 1/2 \left(\sqrt{M_W^2 + p_{\perp,W}^2} + p_{\perp,j}\right)\,, \tag{18.2.5}$$

but no relevant differences were observed with respect to the aforementioned choice, being the two scales similar for the $W + 1$ jet processes at hand.

The scales entering in the evaluation of parton distribution functions and of the strong coupling in the POWHEG Sudakov form factor are chosen to be equal to the transverse momentum of the POWHEG hardest emission [417, 386].

Scale-uncertainty bands obtained by varying the factorisation and renormalisation scales entering the $\bar{B}$ function by a factor of two in either directions are used as an estimate of the theoretical error associated to higher order missing effects.

The uncertainty due to the PDF choice was estimated generating events using three different sets (CT10 [255], MSTW2008 [262], and NNPDF2.1 [312]). The value of the strong coupling constant at M_Z is consistently read from the PDF table used. The further showering performed by PYTHIA8 is instead performed with default PDF and α_s definitions, the difference being beyond the claimed accuracy of the calculation. In this study, we have used PYTHIA8, version 8.153.

18.26 PYTHIA8

PYTHIA8 [348] is the latest incarnation of event generators of the PYTHIA family. At the heart of the generator are parton showers that evolve high-scale processes to the scale of hadronisation, by generating splittings with DGLAP splitting kernels. The splitting scales are ordered in relative transverse momentum [348, 447], and the phase space is constructed in a dipole-like manner in order to capture soft gluon coherence effects [453]. A key point of the evolution of partonic states in PYTHIA8 is that all perturbative components are interleaved [348, 447, 454], i.e. multiple partonic interactions, space-like and time-like showers are all generated in one transverse-momentum ordered evolution sequence. This means that due to the competition for phase space, all steps in the event generation are correlated. For a detailed discussion how parameters of the interleaved shower evolution are tuned to collider data, see [455]. PYTHIA8 with additional matrix element corrections has so far not been tuned to data. Since in [441], only very small differences were seen for LEP between PYTHIA8 with and without matrix element merging, we expect only small re-tuning effects in the parameters of the Lund string model [456]. Similarly, since we keep the low-scale modelling of PYTHIA8 largely intact, only small changes in the underlying event tuning are expected. We however expect that some re-tuning will be needed for jet shape data.

It should be noted that PYTHIA8 includes a selection $2 \to 1$ and $2 \to 2$ processes, as well as a limited variety of $2 \to 3$ processes, but does not contain a general ME generator. New processes, particularly for higher jet multiplicities, have to be made available in form of Les Houches Event (LHE)

[452] files. By virtue of matrix element corrections, PYTHIA8 describes the first emission in $\mathrm{W}+\mathrm{jets}$ with the full matrix element probability. When introducing matrix elements with one additional jet within matrix element merging, this allows to fully cancel the merging scale dependence for the first emission, while small merging scale dependencies enter when including further jets. Current versions of PYTHIA8 include a general implementation of the CKKW-L matrix element merging prescription [445]. Please consult [441] for a detailed discussion of the implementation in PYTHIA8.

Generation of the predictions To generate predictions with stand-alone PYTHIA8 (i.e. without inclusion of matrix elements for W production in association with two or more jets), the built-in $\mathrm{q\bar{q}} \to \mathrm{W}$ matrix element in PYTHIA8 was used to generate the initial configuration. This was then evolved with to the hadronisation scale and the ensemble of partons hadronised using the Lund string model. For this study, we use the publicly available PYTHIA 8.157, with CTEQ6L1 parton distribution functions, and the associated Tune 4C. Since [441] showed a large dependence of the quality of the matrix element merging on whether rapidity-ordered emissions are explicitly forbidden in space-like showers, results are presented with and without enforced rapidity ordering.

The inclusion of matrix elements for additional jets into PYTHIA8 is achieved with CKKW-L merging. All merging tasks are handled internally in PYTHIA 8.157, allowing for a high degree of automation. This means that the user only needs to supply

- Matrix element configurations in form of LHE files.
- An identifier giving the hard process of interest.
- A value of the merging scale. Facilities to allow the user to implement a her/his own merging scale definition are available.

For this report, matrix element configurations with additional jets were generated with MADGRAPH/ MADEVENT [163], and read into PYTHIA8 in form of Les Houches Events. PYTHIA8 then derives all possible parton shower histories for an event, probabilistically chooses a history, and uses the reconstructed states and splitting scales to perform a re-weighting with Sudakov factors and α_s values. This means each event will have a weight

$$
\begin{aligned}
w_{\text{CKKWL}} \quad = \quad & \frac{x_n^+ f_n^+(x_n^+,\rho_n)}{x_n^+ f_n^+(x_n^+,\mu_F^2)} \frac{x_n^- f_n^-(x_n^-,\rho_n)}{x_n^- f_n^-(x_n^-,\mu_F^2)} \\
& \times \prod_{i=1}^{n} \left[\frac{\alpha_s(\rho_i)}{\alpha_{\text{sME}}} \frac{x_{i-1}^+ f_{i-1}^+(x_{i-1}^+,\rho_{i-1})}{x_{i-1}^+ f_{i-1}^+(x_{i-1}^+,\rho_i)} \frac{x_{i-1}^- f_{i-1}^-(x_{i-1}^-,\rho_{i-1})}{x_{i-1}^- f_{i-1}^-(x_{i-1}^-,\rho_i)} \Pi_{S_{+i-1}}(\rho_{i-1},\rho_i) \right] \Pi_{S_n}(\rho_n, t_{\text{MS}})
\end{aligned}
$$

where ρ_i and $x_i^{\pm}$ are the the reconstructed shower splitting scales and momentum fractions of the incoming partons in $\pm$z-direction, and $\Pi_{S_{+i}}(\rho_i,\rho_{i+1})$ the parton shower no-emission probability when evolving the state S_{+i} from scale ρ_i to ρ_{i+1}. α_{sME} gives the strong coupling used in the matrix element calculation. All reweighting factors are generated dynamically with help of the shower. The interleaved evolution of PYTHIA8 is accommodated by consistently including effects of multiple interactions into the no-emission probabilities. A detailed description of the formalism is given in [441].

As input for the current analysis, we have produced LHE files for $\mathrm{W}^+ + \mathrm{jets}$ with up to four (three) additional jets at Tevatron (LHC) energies. The renormalisation scale in MADGRAPH was fixed to $\mu_R = \mathrm{M_Z}$. For hadronic cross sections, CTEQ6L1 parton distributions (as implemented in LHAPDF [457]) have been chosen at a factorisation scale $\mu_F = \mathrm{M_W}$, and the strong coupling in the ME was correspondingly fixed to $\alpha_s(\mathrm{M_Z}) = 0.129783$. To regularise QCD divergences and act as a merging scale, a cut in

$$
k_\perp^2 = \min\left\{ \min(p_{T,i}^2, p_{T,j}^2), \min(p_{T,i}^2, p_{T,j}^2) \frac{(\Delta\eta_{ij})^2 + (\Delta\phi_{ij})^2}{D^2} \right\} \quad \text{with} \quad D = 0.4
$$

and a cut value of $k_{\perp,min} = t_{\text{MS}} = 15$ GeV has been applied to the matrix element.

Merged PYTHIA8 predictions are given for the default settings, i.e. using the parameters of Tune 4C, for Tune A2 [458], and for Tune 4C without enforced rapidity ordering (dubbed Tune X). Again, it should be noted that so far, no tuning including additional jets has so far been conducted.

Estimate of uncertainties To estimate uncertainties of a merged prediction of $\mathrm{W}+\mathrm{jets}$, it is interesting to study the dependence on the merging scale value. For this, we have generated LHE files with three different $k_{\perp,min} = t_{\text{MS}}$ cuts ($t_{\text{MS}} = 15, 30, 45$) GeV, and performed CKKW-L merging on these samples. Furthermore, to show the effect of tuning, the $t_{\text{MS}} = 15$-GeV-sample was processed for two adequate tunes, Tune 4C and A2.

Uncertainties related to shower ordering In [441], it was shown that restricting shower emissions in PYTHIA8 to regions of phase space ordered both in transverse momentum and rapidity leads to non-negligible effects in merged predictions. This can be seen as an effect of limiting the shower accuracy by reducing the phase space over which splitting kernels are integrated, meaning the accuracy of Sudakov form factors is impaired. Loosely speaking, if above the merging scale, the matrix element, integrated over the full phase space[39], differs substantially from the splitting probabilities integrated over the allowed parton shower phase space, merged results will exhibit substantial merging scale dependencies. Such problems are obviously introduced if the parton shower phase space is heavily constrained.

Changing the phase space regions in which the shower is allowed to radiate thus allows us to estimate the uncertainties of the merging procedure in conjunction with the underlying shower. Particularly, this procedure can test the quality of the matrix element merging beyond the first few emissions, and give hints on how the shower resummation may be improved.

To emphasise the impact of the shower transition probabilities, we choose a fairly small merging scale ($t_{\text{MS}} = 15$ GeV) to regularise the tree-level matrix elements for this investigation. Then, for each matrix element state, we generate *all* possible parton shower histories for a matrix element state, by clustering emissions. This is achieved by inverting the shower momentum- and flavour-mappings.

When merging matrix elements with rapidity-ordered showers, we investigate two ways of biasing the selection of a particular history, from which to generate the necessary Sudakov form factors:

1. In a "y-blind" sample, we do not include an additional discriminant based on rapidity. This means that – just like in the standard case – ρ-ordered will be preferred over ρ-unordered ones.
2. In a "y-conscious" sample, we pick histories with rapidity-unordered splittings only if no rapidity-ordered histories were found. Adopting this strict ordering criterion, histories ordered in ρ *and* rapidity will be chosen predominantly, and only if no such history exists, histories un-ordered in either ρ and/or rapidity are picked.

It should be noted that to the accuracy of the parton shower, both these prescriptions are equivalent, and switching the choice of histories gives a real estimate of the quality of the merging in conjunction the underlying shower. We believe that including this uncertainty gives a pessimistic view on how wide the range of predictions of one merged calculation can be, indicating that although standard by now, matrix element merging in PYTHIA8 should be applied with care. However, with reasonable settings, including additional jets can improve the description of multiple hard jets substantially.

18.27 SHERPA

SHERPA [425, 146] is a full-fledged event generator capable of simulating all aspects of particle collisions as they occur at particle accelerators such as the Tevatron or the LHC. It includes two independent matrix element generators, AMEGIC++ [147] and COMIX [459], to generate cross sections and distributions

[39] Particularly for high jet multiplicities, it could be imagined that phase space integrators have difficulties to fully sample the phase space, especially close to kinematic limits. By full phase space, we mean the region that the phase space generator actually filled.

for final state multiplicities of up to six to ten particles. In the former one, methods to automatically generate dipole subtraction terms in the widely used Catani–Seymour scheme [236, 150] have been incorporated [448]; the SHERPA package also supports the BLHA [145] for the interface to one-loop programs such as BLACKHAT or GOSAM. For parton showering, SHERPA employs an algorithm based on Catani-Seymour subtraction kernels, proposed in [460] and implemented in the SHERPA framework in [461]. For the hadronisation, SHERPA uses either its native hadronisation scheme, based on the cluster fragmentation model [462] and its implementation described in [463] or an interface to PYTHIA [400] providing access to the routines of the Lund string model [456]. Both have been successfully tuned to LEP data within the SHERPA framework, with a similar quality in describing the data. The hadron decays are also fully provided in the SHERPA framework, as well as QED final state radiation to both the W-boson and the hadron decays, simulated using the YFS approach [464, 465].

In this work, the most recent, publically available SHERPA version, SHERPA-1.3.1, has been used in two ways of running the simulation, namely

1. in the MEPS mode:
   In this method, towers of LO matrix elements with increasing jet multiplicity, in the case at hand $W, W+1, W+2, \ldots, W+n_J$ jets, are merged in the spirit of [429, 430] to yield an inclusive sample. In fact, codes relying on such algorithms have been compared in a previous publication [439], which helped to establish and validate the methods and their various implementations. In contrast to the original implementation in SHERPA [466], which used analytical forms of Sudakov form factors etc., the current version of the method [426] directly uses the parton shower for Sudakov rejections etc. and is thus closer in spirit to the variant presented in [445, 467] for multijet merging.
2. in the MENLOPS mode:
   This method can be understood as the combination of a matching of the parton shower to a NLO matrix element and a merging of additional towers of LO matrix elements with even higher jet multiplicities. Thus, in the case at hand, inclusive W production calculated at NLO accuracy is merged, as above, with LO matrix elements for $W + 1$, $W + 2$, $\ldots$, $W + n_J$ jets. This method has been pioneered in [442, 443] where the implementation employed within SHERPA has been detailed in the second reference.

The respective settings and relevant details for both simulation modes are described below.

SHERPA in MEPS mode In the MEPS mode SHERPA was run with up to $n_J = 6$ jets in the matrix element evaluation including all possible massless (anti-)quark and gluon initial and final states. All matrix elements were generated using COMIX. The MEPS-separation parameter was set to $Q_{\text{cut}} = 20$ GeV, for its precise definition see [426]. The scales are chosen as

$$\alpha_s^{k+n}(\mu_{\text{eff}}) \;=\; \alpha_s^k(\mu) \cdot \alpha_s(p_{\perp,1}) \cdot \ldots \cdot \alpha_s(p_{\perp,n}) \;, \tag{18.2.6}$$

wherein the relative transverse momenta $p_{\perp,i}$ are the nodal values of the final state partons of the $W + n$ parton matrix element as obtained from recombining it using the inverted splitting probablities given by the parton shower. The core scale μ is then chosen as the partonic centre-of-mass energy of the reconstructed core process, i.e. $\mu^2 = \hat{s}_{2\to 2}$ where $k = 0$ in the process at hand. In all stricly perturbative setups a parton shower cutoff of $t_0 = (0.7\text{ GeV})^2$ has been used.

The parton shower cutoff and all fragmentation parameters of both the internal cluster hadronisation and the interfaced Lund string fragmentation models have been tuned to LEP data and give a similarly good description. Similarly, the parameters of SHERPA's MPI model have been tuned to Tevatron and LHC data using the CT10 [255] parton density parametrisation. These parameters are given in App. 18.71.

SHERPA in MENLOPS mode In the MENLOPS mode SHERPA is run with essentially the same parameters as in the MEPS mode, described in the previous subsection. Hence, $n_J = 6$ and $Q_{\text{cut}} = 20\,\text{GeV}$. To be able to describe the inclusive W production process at NLO accuracy, AMEGIC++ was used for all parts of the NLO W production matrix elements (supplemented with a hardcoded one-loop matrix element from the internal library) and the LO $W + 1$ parton matrix element. Consecutively, the scales were chosen as above with $k = 0$ for all tree-level parts and $k = 1$ for the real and virtual corrections entering the next-to-leading order correction of the core process. All non-perturbative parameters remain unchanged wrt. the MEPS mode.

Estimate of uncertainties In order to estimate the uncertainites of the SHERPA predictions, the following procedures have been applied:

(A) PDF uncertainties:
Unlike in the PDF4LHC presciption [468], here only the central predictions of the three NLO PDFs, CT10 [255], MSTW2008 [262] and NNPDF2.1 [312] are compared to estimate the PDF uncertainties. The different parametrisations of PDFs as well as their corresponding value of α_s, both its value at M_Z and its running, enter in the calculation of the matrix elements, the parton shower and the underlying event.

(B) Scale uncertainties:
In a global manner, all scales, renormalisation and factorisation scales are simultaneously modified by the canonical multiplication with 2 and 1/2. This, however, is not only applied to the evaluation of the matrix elements but also to that of the parton shower, the hadronisation, the underlying event simulation and the hadron decays. Regarding the matrix-element evaluation, the MEPS default scale choice forms the starting point for the scale variations to be executed.

(C) Hadronisation uncertainty:
Here the intrinsic modeling uncertainties are evaluated by changing the hadronisation model operating on SHERPA's parton shower final states, namely switching from SHERPA's default cluster hadronisation to PYTHIA's string fragmentation. For both schemes, an independently tuned set of parameters has been employed to perform the parton-to-hadron transition.

(D) Underlying event uncertainty:
To this end the tune of the underlying event based on using the CT10 PDF has been modified such that the plateau of the number of charged particles and sum of transverse momenta in the transverse region are increased or decreased by 10%. This change in the amount of MPI activity is accomplished by varying the σ_{ND} correction factor (SIGMA_ND_FACTOR) by -0.04 or $+0.05$, respectively.

18.3 Results

In this section we compile results for the individual codes for a number of representative observables at the different levels of the simulation. It should be noted, though, that in all results presented in this section PDF uncertainties have been estimated by typically varying only over a few different sets rather than employing the full procedure as suggested by the PDF4LHC accord [468].

18.31 BLACKHAT + SHERPA

The following results have been obtained with BLACKHAT+SHERPA. Uncertainties due to the factorisation/renormalisation scale variation and the that due to the PDF uncertainties are shown. The yellow band corresponds to the addition in quardature of these two uncertainties and the statistical estimation on the integration error. All observables are defined using the ATLAS cuts, cf. App. 18.6.

Fig. 64 displays the inclusive cross section for a W boson in association with n jets, where n=1,2,3,4. A NLO computation of W+4 jets also provides a leading order calculation of the W+5

jets rate, but since it is not at NLO accuracy we refrain here from including it.

In all the plots presented in this section the uncertainties are dominated by the uncertainty arising from the scale variation (it is not the case when the central scale of the process is chosen close to a local maximum, in which case the upper boundary of the scale variation is very close or identical with the central value, as can be seen from the plots corresponding to W+3,4 jets). This is partially due to the fact that for the assessment of the PDF uncertainty only error sets have been employed that are closely related to the central set. In addition, the functional form of the scale definition as given by the kineamtics of the final state has not been changed, but rather the emerging scales μ_F and μ_R have been multiplied in parallel by factors of 2 and 1/2.

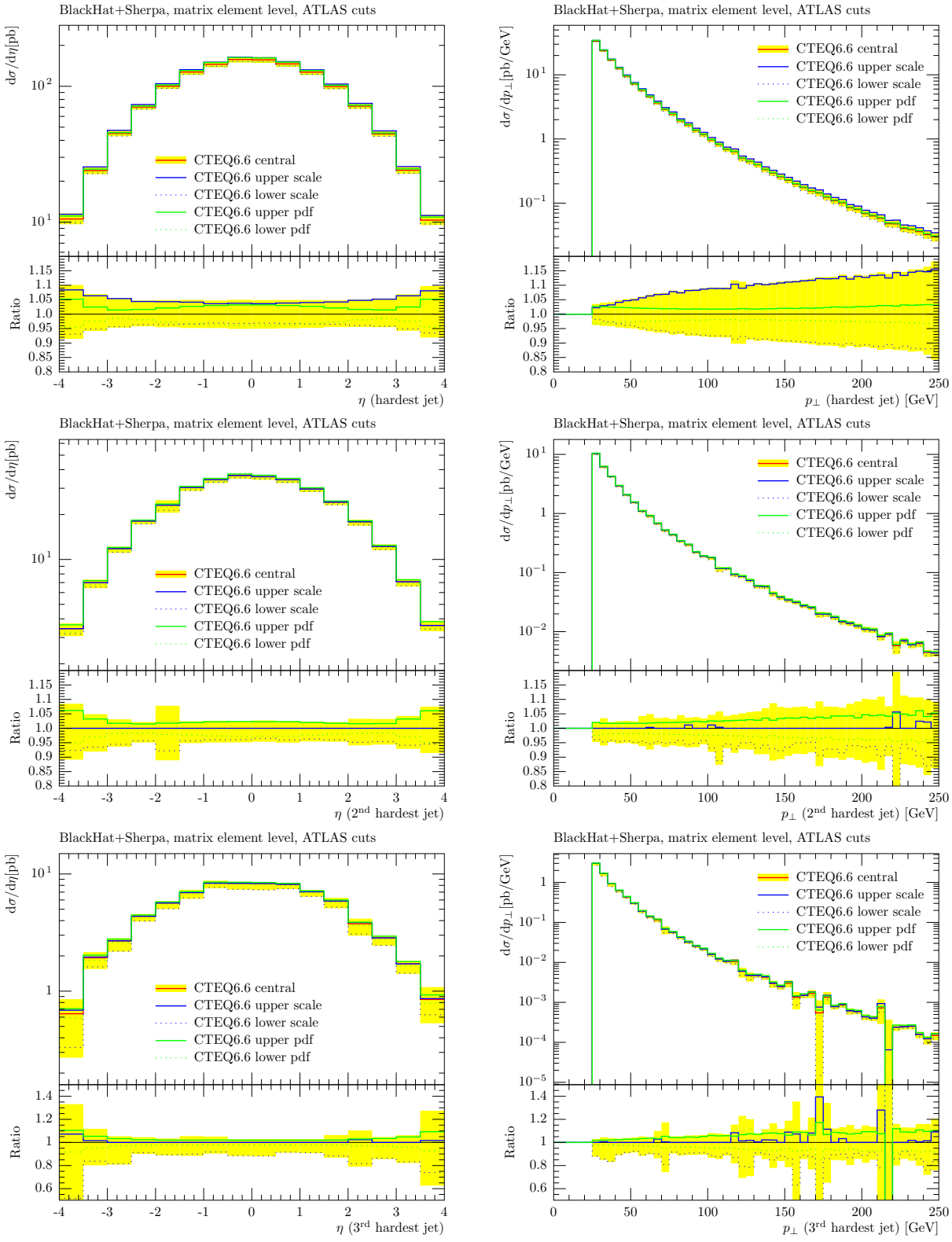


Fig. 61: Pseudo-rapidity and transverse momentum distributions for the first jet in inclusive $W + 1$ jet production (upper panel), for the second jet in inclusive $W + 2$ jet production (central panel), anf for the third jet in inclusive $W + 3$ jet production (lower panel).

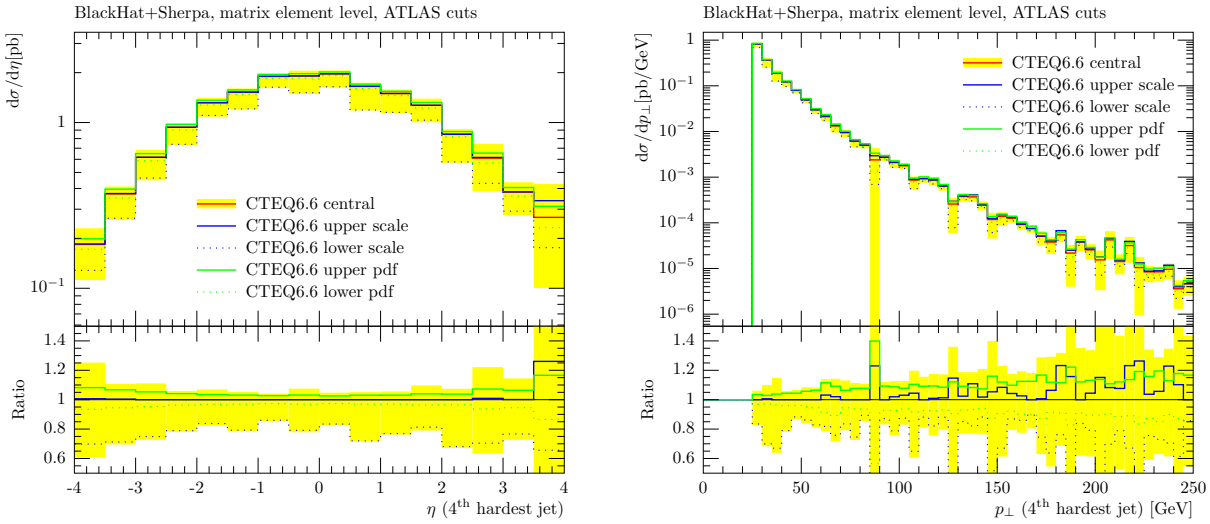


Fig. 62: Pseudo-rapidity and transverse momentum distributions for the fourth jet in $W+4$ jet production.

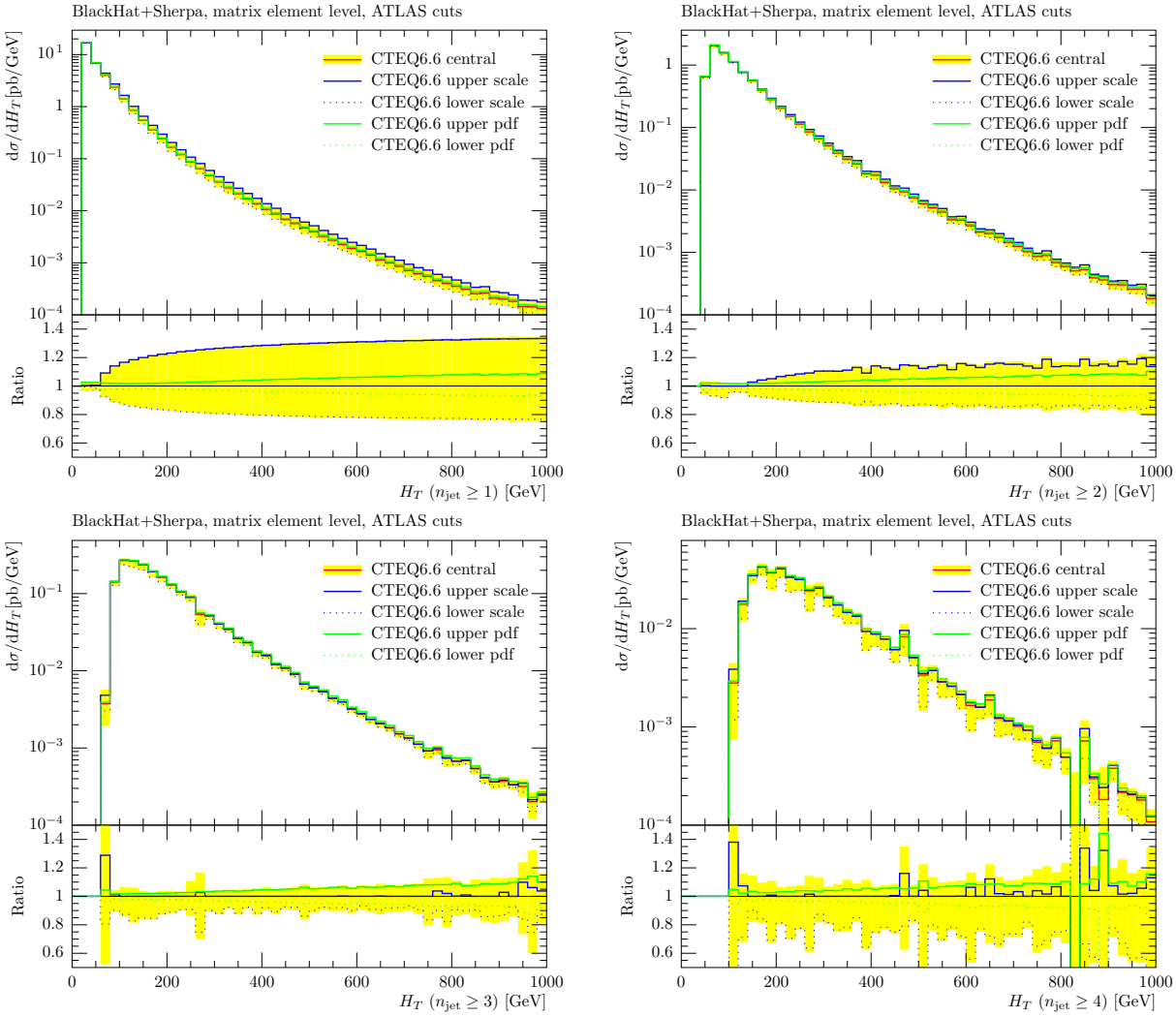


Fig. 63: HT distributions for event with at least one (top left), two (top right), three (bottom left) or four (bottom right) jets.

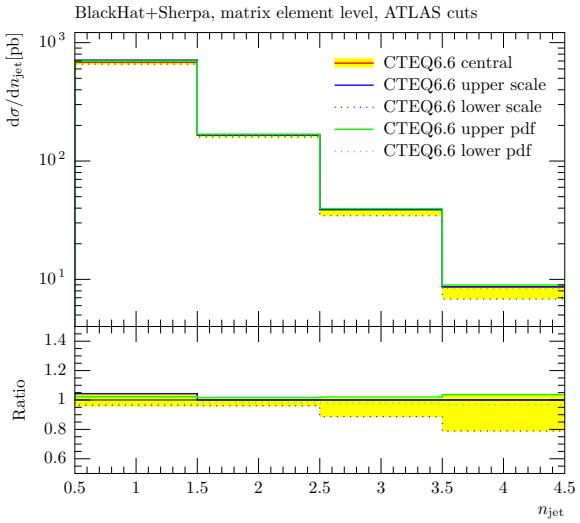


Fig. 64: Inclusive cross section for $W + n$ jet production.

18.32 GoSam + Sherpa

The setup described in the previous section for the analysis using GoSam+Sherpa gives the following theoretical uncertainties. The plots show that in general the scale uncertainties are bigger then the PDF uncertainties and that the renormalisation scale dependence is usually bigger then the dependence on the factorisation scale. To illustrate the decrease in the scale uncertainty given by the NLO calculation we also include the distributions for the pseudo-rapidity and transverse momentum of the second hardest jet, which have only tree-level accuracy. All observables shown are defined using the ATLAS cuts, cf. App. 18.6. Note that errors in the 2-jet configuration are increased w.r.t. those provided by BlackHat+Sherpa, since here only $W + 1$ jet configurations are dealt with at NLO, and the 2-jet configurations therefore are descibed at LO only.

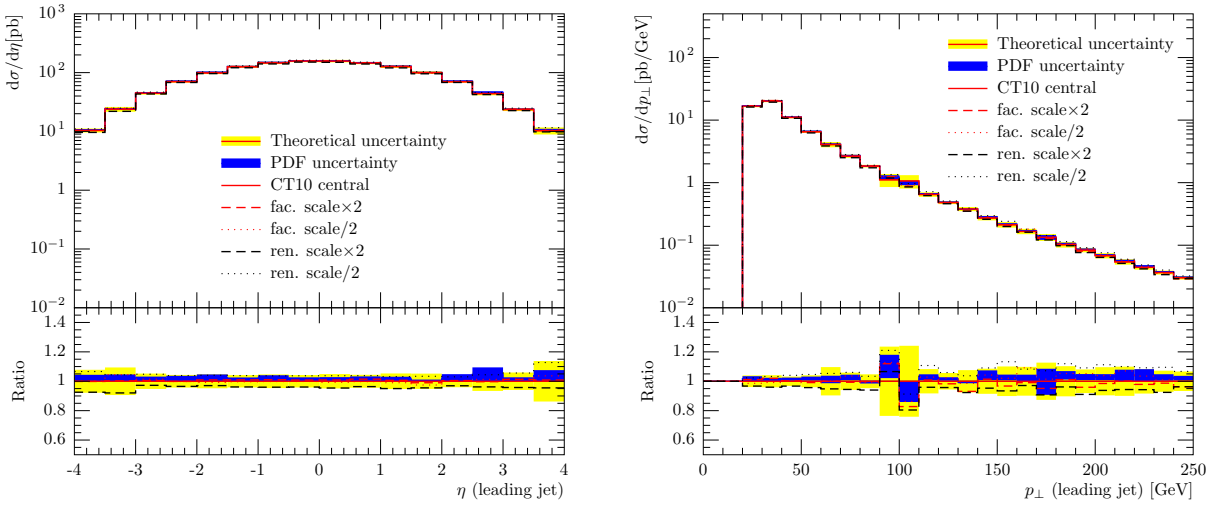


Fig. 65: Pseudo-rapidity and transverse momentum distributions for the hardest jet.

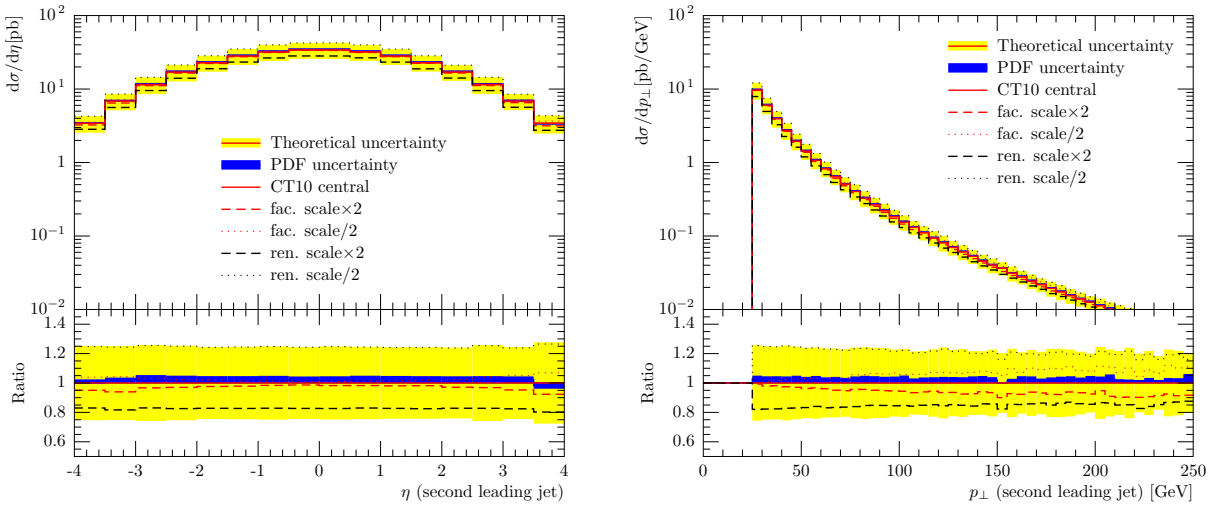


Fig. 66: Pseudo-rapidity and transverse momentum distributions for the second hardest jet. This distribution have formally leading order accuracy and have therefore a much larger scale dependence than the same distribution for the hardest jet, for which a genuine NLO prediction is available.

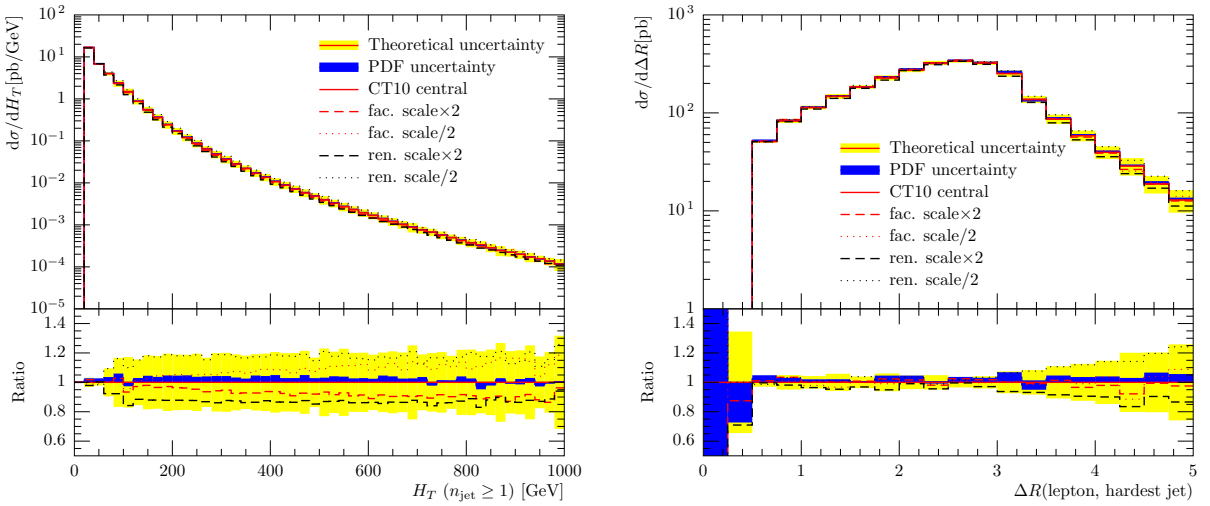


Fig. 67: HT distributions (left) and ΔR between lepton and hardest jet (right) for events with at least one jet.

18.33 HEJ

This section contains the predictions from the High Energy Jets (HEJ) event generator. This gives predictions for the production of a W boson in association with *at least two* jets. Throughout, we show results for CTEQ, MSTW and NNPDF parton distributions. We show a scale uncertainty band only for the first of these for clarity. The results for the other two are very similar. The yellow band in the ratio panel shows the statistical uncertainty in each case. The scale variation is seen to be dominant over the statistical uncertainty and the differences in choice of pdf. All observables are defined using the CMS cut definitions, cf. App. 18.6.

As discussed in Sec. 18.23, the resummation contained in the HEJ framework is supplemented with a merging procedure to ensure tree-level accuracy for events with up to and including four jets. This leads to the larger drop from the four jet to the five jet cross section, compared to the drop either from

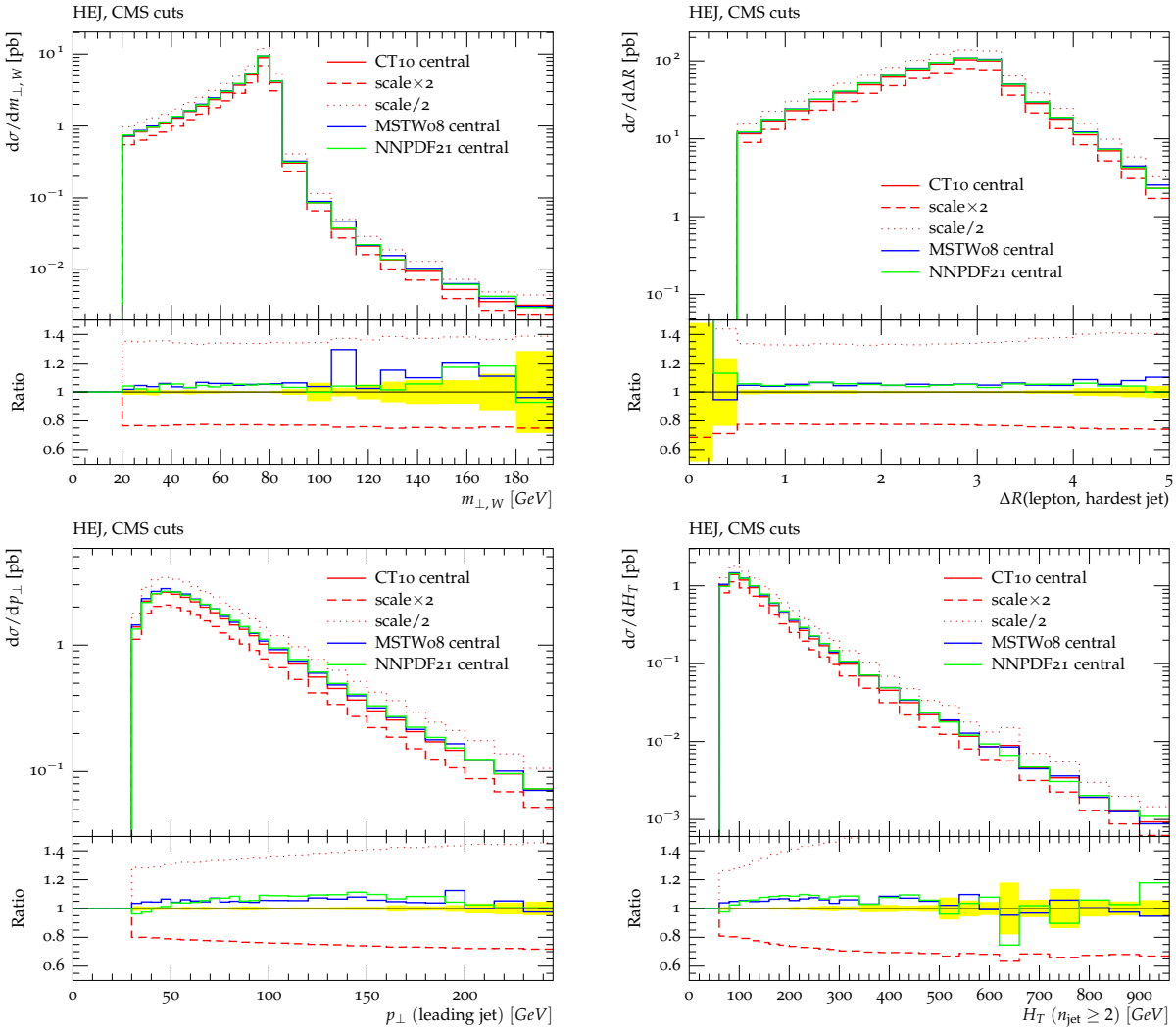


Fig. 68: The HEJ prediction for the distribution of the transverse mass of the W boson (top left) and for the angle between the hardest jet and the charged lepton from the decay of the W boson (top right), the transverse momentum of the hardest jet (bottom left) and for the H_T distribution (bottom right) in events where a W boson was produced in association with at least two jets.

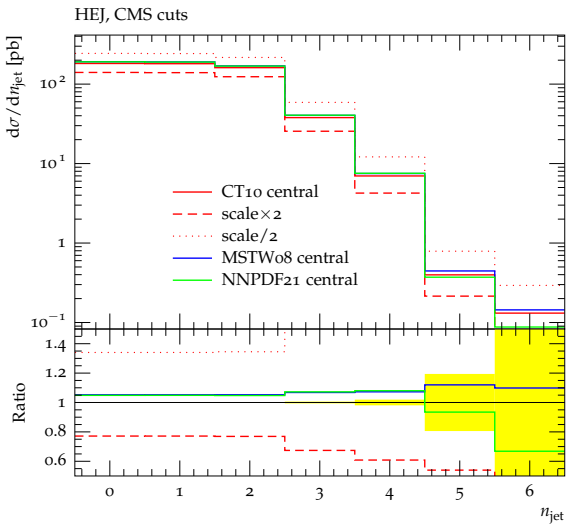


Fig. 69: The HEJ prediction for the cross sections of W plus n jets.

three-jet to four-jet, or from five-jet to six-jet. This can be clearly seen in Fig. 69.

18.34 MADGRAPH + PYTHIA

The following results have been obtained with MADGRAPH+PYTHIA. Uncertainties due to the factorisation and renormalisation scale and MEPS matching scale are shown for results on hadron level including UE and QED final state radiation. A comparison of results on parton shower level, hadron level, hadron level including UE, and hadron level including UE and QED final state radiation, is also presented. All observables shown are defined using the CMS cuts, cf. App. 18.6.

From these results we can conclude that the largest uncertainty on all observables is due to the factorisation and renormalisation scale. In addition to that, a large effect is found by switching off the final state QED radiation, while only a small difference is oberved between results at the shower level and all other results prior to the QED radiation.

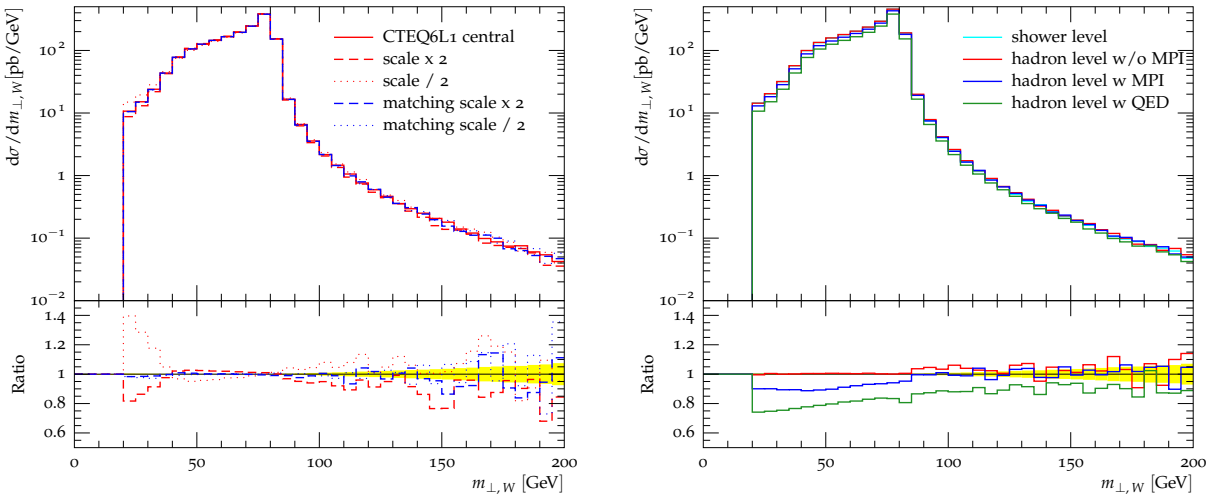


Fig. 70: MADGRAPH+PYTHIA results for W transverse mass.

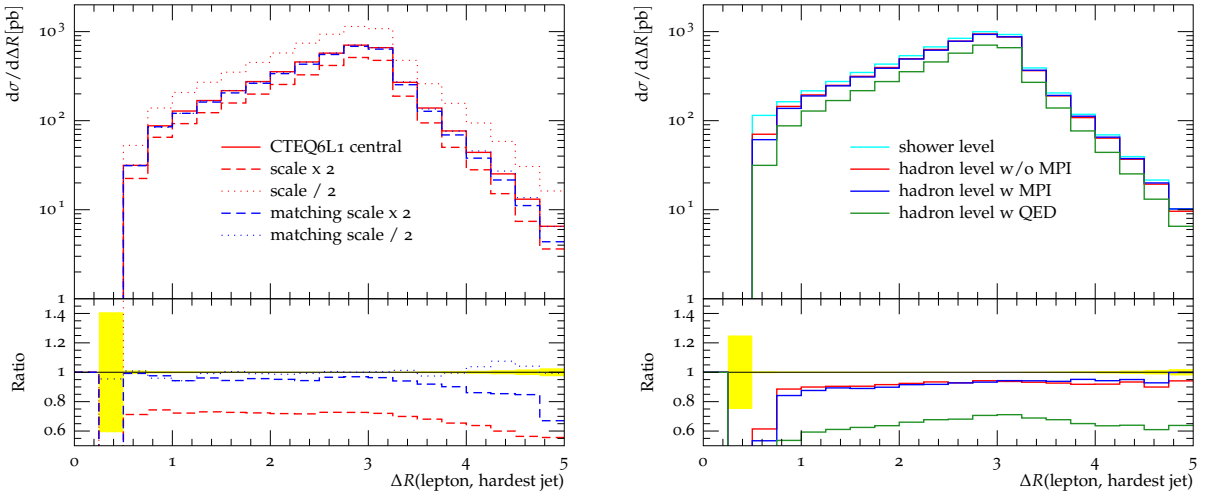


Fig. 71: MADGRAPH+PYTHIA results for ΔR between lepton and hardest jet.

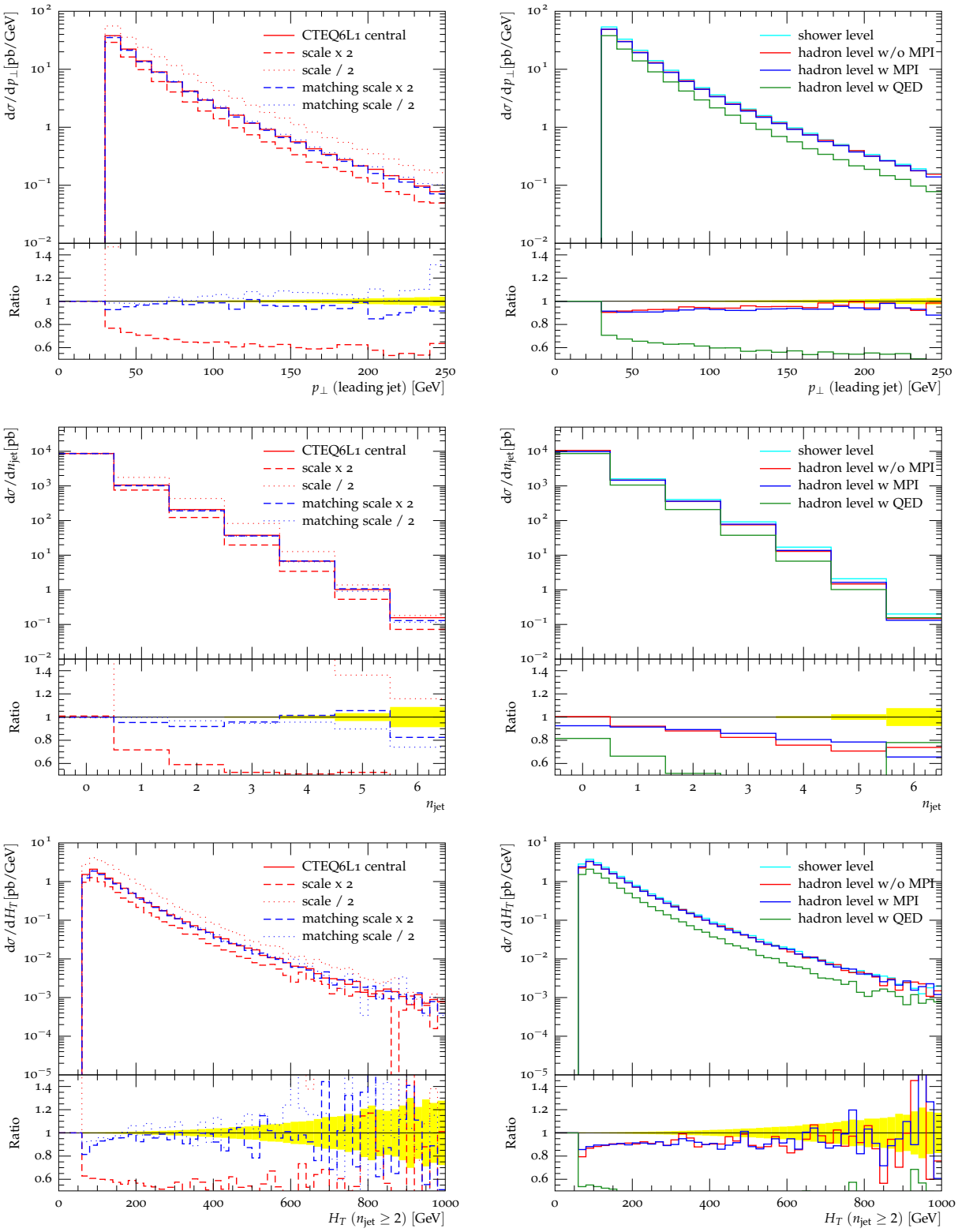


Fig. 72: MADGRAPH+PYTHIA results for for $p_\perp$ of hardest jet (top), number of jets (middle) and H_T of events with at least 2 jets (bottom).

18.35 POWHEG BOX + PYTHIA8

In this section we show results obtained by running the POWHEG BOX implementation of $W + 1$ jet together with PYTHIA8. In all the following plots of this section, CMS analysis cuts have been enforced, see App. 18.6. For this study, in the left panels of Figs. 73-75, we show uncertainties obtained from variations of renormalisation and factorisation scales by a factor of two in either directions and by choosing different PDF sets in the computation of the hard scattering. Results are shown at the final level, after the shower, the hadronisation and the inclusion of MPI, all performed by PYTHIA8. In general, we notice that the uncertainty due to scale variations is greater than the changes in the results due to different PDF choices.

In the right panels of Figs. 73-75 we show our results at different stages of the simulation, for a fixed PDF set (chosen to be CT10). The stages considered include from the first emission level up to the full showered events in PYTHIA8, including MPI and also effects due to QED radiation off leptons and quarks. Various stages of the simulation have been obtained setting the PYTHIA8 switches as reported in Sec. 18.73.

We recall here that results should be considered to be physical only after the the hadron level is reached (possibly including MPI and QED effects). In particular, we stress that the results at the parton level are obtained considering only the POWHEG first emission, and they are therefore only intermediate: indeed at this stage only the hardest radiation has been generated and effects due to further showering are not yet taken into account.

For most of the observables results do not show large variations going from a simulation level to another. In particular, for truly NLO predictions such the plots in Fig. 74 or the bin $n_{\text{jet}} = 1$ of Fig. 73, the major effects that arise at each successive stage of the simulation are a change in the normalisation, due to a slightly different number of events passing the analysis cuts when multiple emissions are allowed, and a moderate shape distortion in the low end of the spectrum. Both these effects may be attributed to multiple QCD radiation due to Sudakov effects introduced by the parton shower. As expected, these effects are of the same size, or smaller, than the theoretical uncertainty due to scale and PDF's variations, when propagated to the hadronic level. Similar effects are also observed when the QED radiation is turned on. In this case, results are lowered as a consequence of the cuts on the lepton transverse momentum and rapidity.

Due to the requirement of having at least two jets, the remaining observables are predicted only at leading order or with leading log accuracy by the POWHEG simulation of $W + 1$ jet. This is also reflected

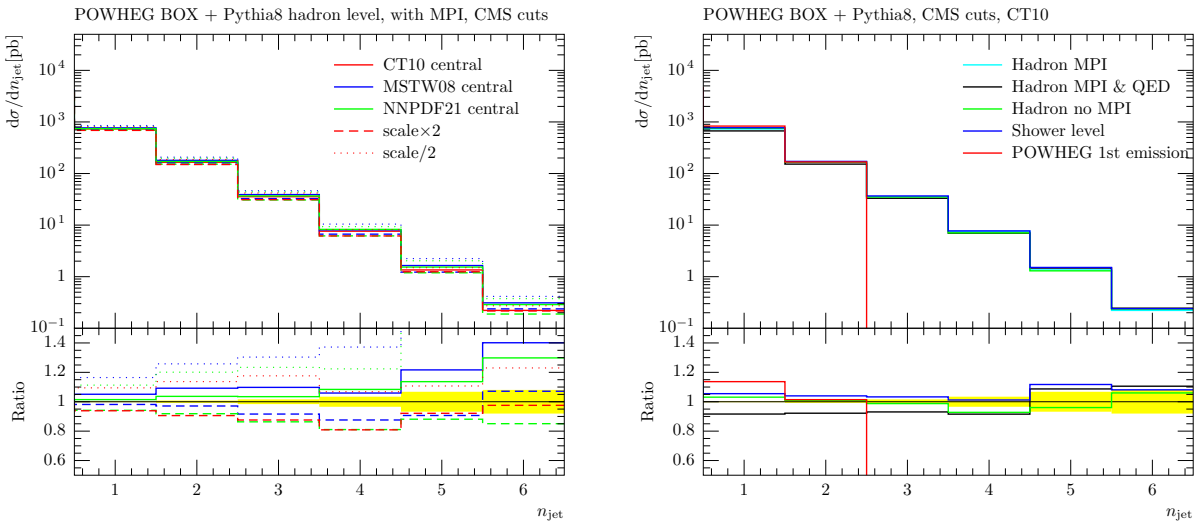


Fig. 73: The number of jets, as predicted by POWHEG BOX + PYTHIA8.

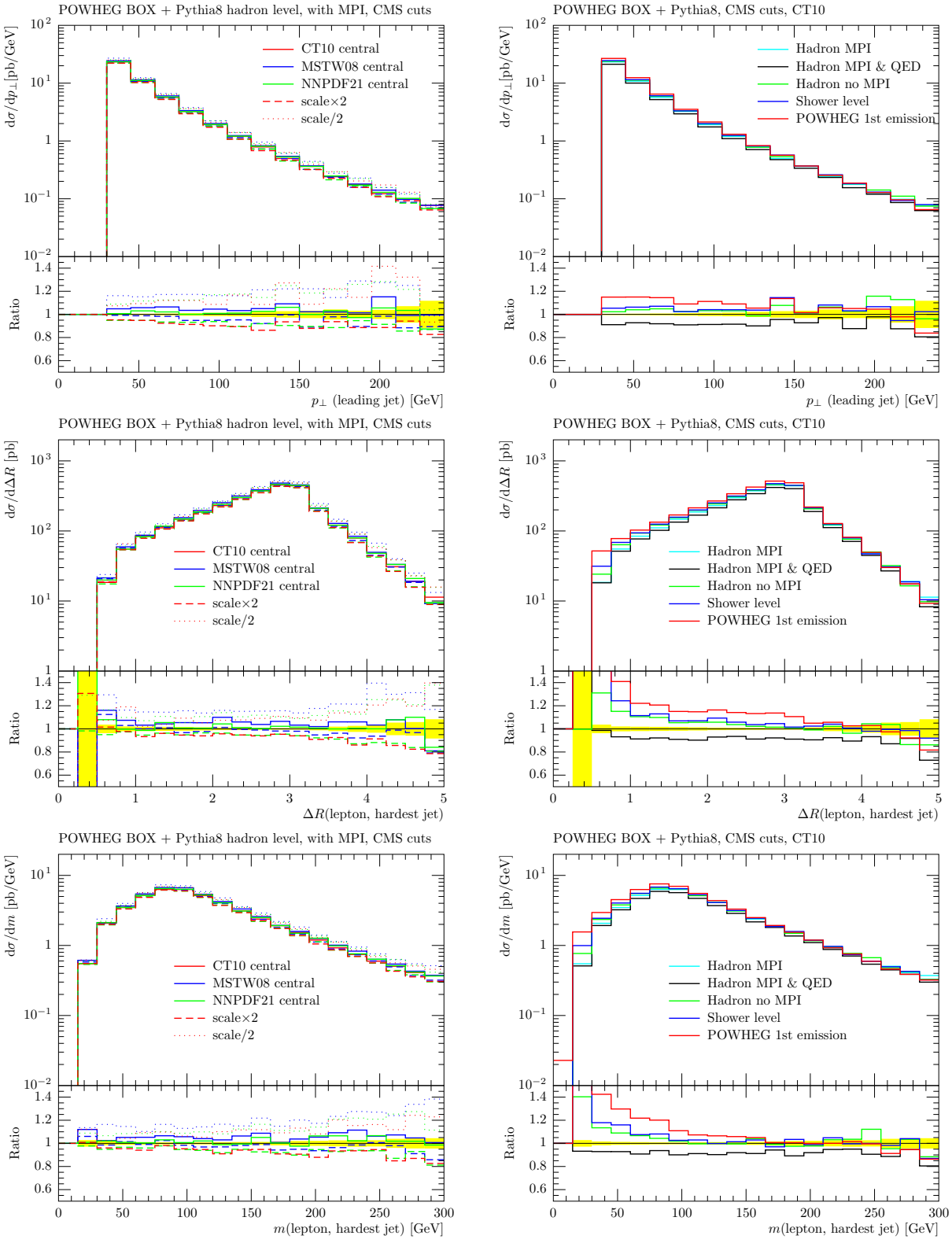


Fig. 74: The hardest jet transverse momentum distribution (upper plots), the ΔR separation (middle plots) and the invariant mass m (lower plots) of the hardest jets and the hardest lepton, as predicted by POWHEG BOX + PYTHIA8.

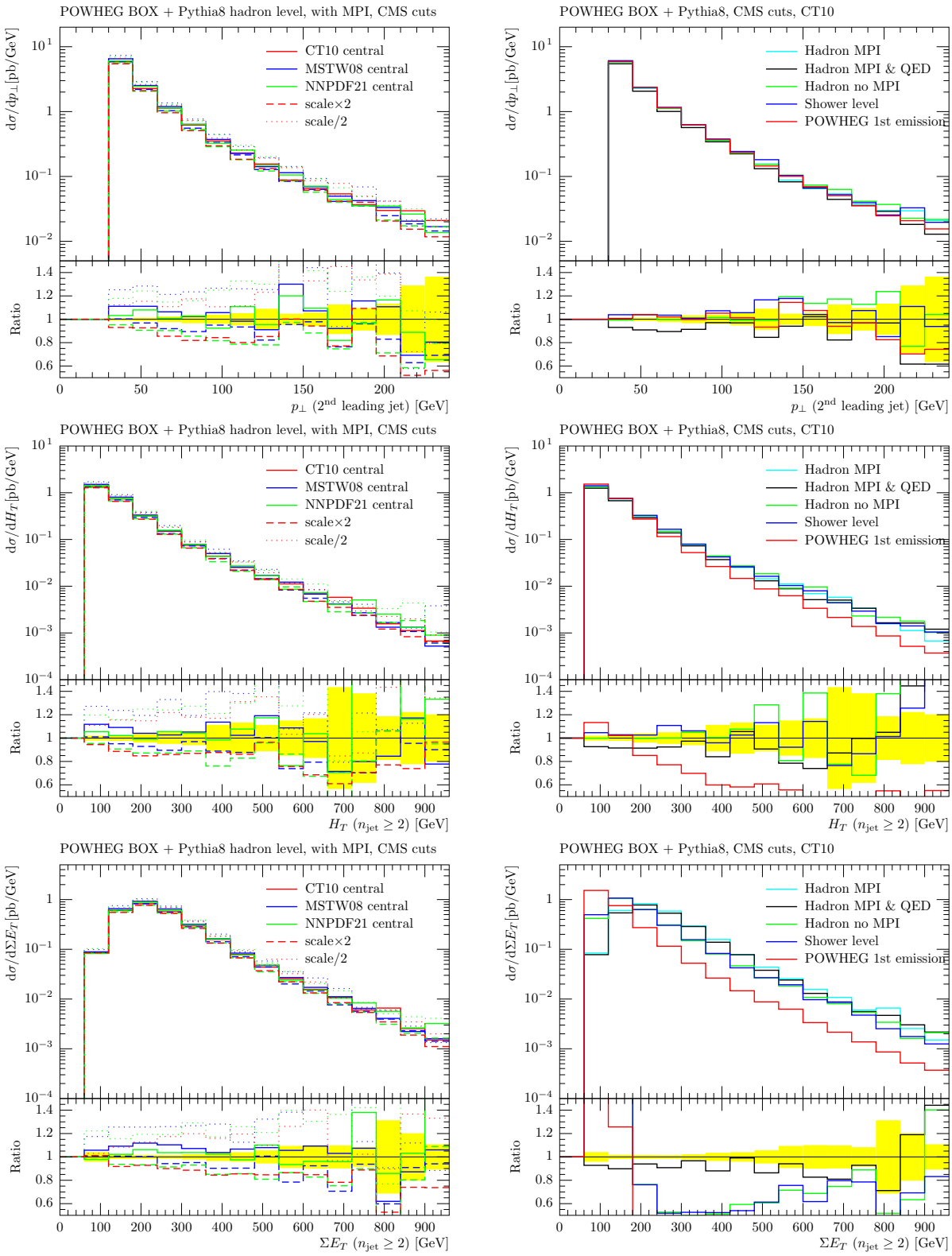


Fig. 75: The transverse momentum $p_\perp$ of the next-to-hardest jet, the scalar sum of the jet transverse energy H_T of events with at least 2 jets and the sum of the transverse energies of all the particles in events with 2 or more jets, as predicted by POWHEG BOX + PYTHIA8.

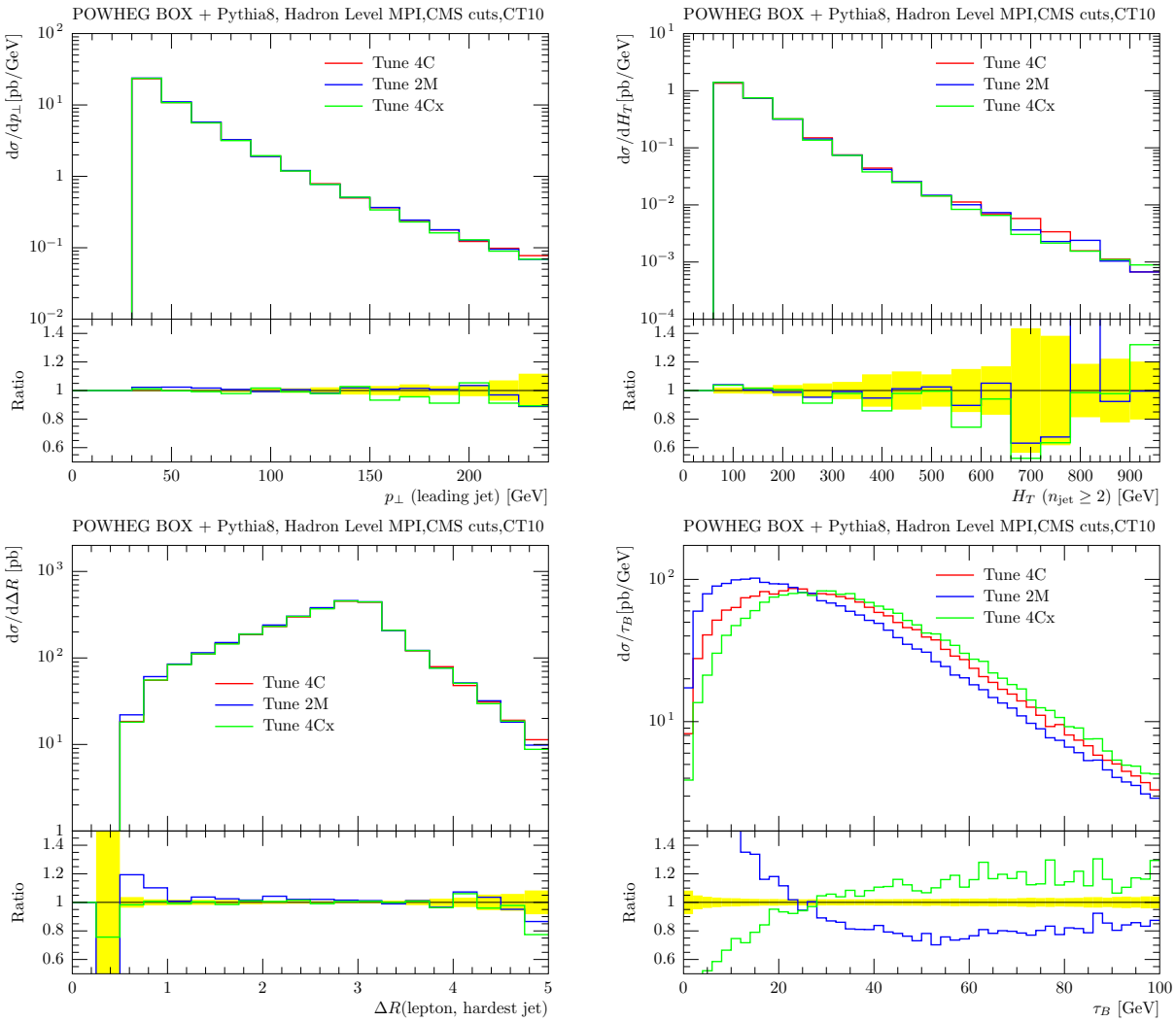


Fig. 76: Comparison between predictions using different PYTHIA8 tunes, at hadron level with MPI, as predicted by POWHEG BOX + PYTHIA8.

in the larger band associated with the scale variations.

Observables such as H_T, the scalar sum of the transverse energy of the jets for events with two or more jets, show an enhancement in the high-H_T tail. This effect mostly arise as a consequence of the showering, since the successive stages do not change the predictions any longer. The same behaviour, even more enhanced, is also observed in the scalar sum of the transverse energy of all particles, always in events with two or more jets.

In Fig. 76 we instead compare the effect of using different PYTHIA8 tunes on our predictions, obtained in this case at the hadron level, including MPI. Essentially all the observables turned out to be extremely stable under the variations of the PYTHIA8 tune, as shown in Fig. 76. Major differences only appears for the beam thrust, when it is defined at the particle level (see App. 18.62).

18.36 PYTHIA8

For this study, PYTHIA8 has been run stand-alone and including matrix elements with additional jets. Note that in PYTHIA8, multiple interactions are interleaved with space- and time-like showers, meaning that in general, MPI and parton showers cannot be disentangled by just switching off secondary scatterings. When referring to "Hadron Level", we mean after the interleaved evolution (including QED

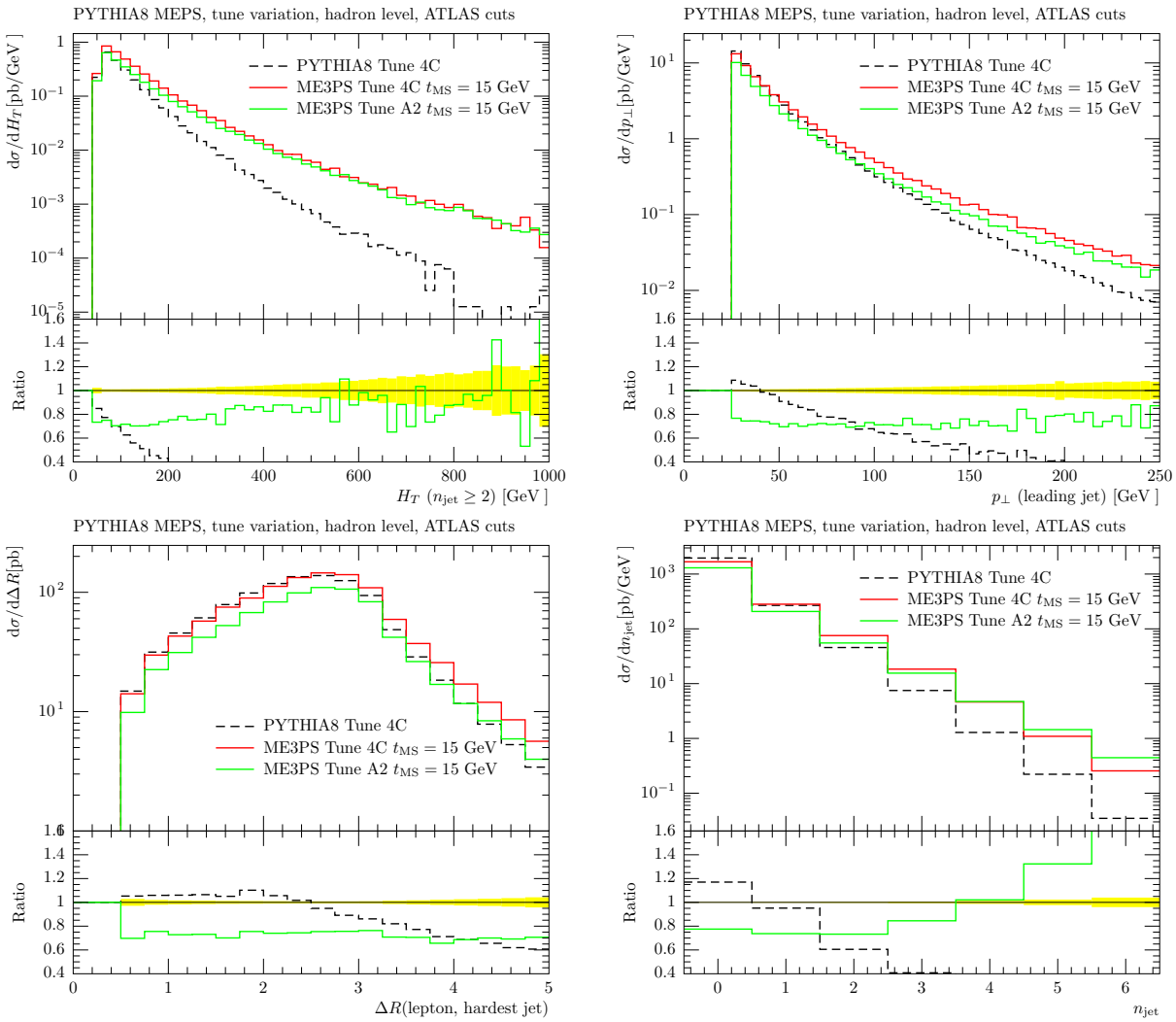


Fig. 77: Tuning variations for PYTHIA8 at hadron level. The plots show the H_T-distribution when requiring at least two jets (upper left), the $p_\perp$ of the hardest jet (upper right), the ΔR-separation of lepton and the hardest jet (lower left), and the number of jets (lower right). The lower insets show the ratio of the samples in the upper half to ME3PS (Tune 4C, y-blind treatment). All merged plots are produced with a merging scale of $t_{\mathrm{MS}} = 15$ GeV.

splittings), and after hadronisation. For the sake of comparison, "Shower Level" indicates results after (interleaved) final- and initial-state radiation, switching multiparton interactions off. All results presented in this section are generated with CTEQ6L1 parton distributions for protons colliding at $E_{\mathrm{CM}} = 7000$ GeV. CKKW-L-merged samples include up to three additional jets, taken from MADGRAPH/MADEVENT.

Fig. 77 exemplifies how changes in the tuning of the event generator can affect the outcome of merged calculations in PYTHIA8. For this, we produce predictions for Tune 4C [455] and Tune A2 [458]. In general we observe only modest shape changes of up to about 20% in observables, when comparing the two merged predictions, lending confidence to the statement that the tuning did not artificially produce hard scale physics. Normalisation changes between 4C and A2 can be explained by a difference in Sudakov suppression: Since Tune 4C integrates the splitting kernels over a smaller region of phase space, the suppression generated by trial showers is less pronounced. The increase in the number of jets in Tune A2 with respect to Tune 4C, after the third jet, is expected, because the generation of the fourth jet is handled solely by the parton shower. Since 4C allows less phase space for these emissions by

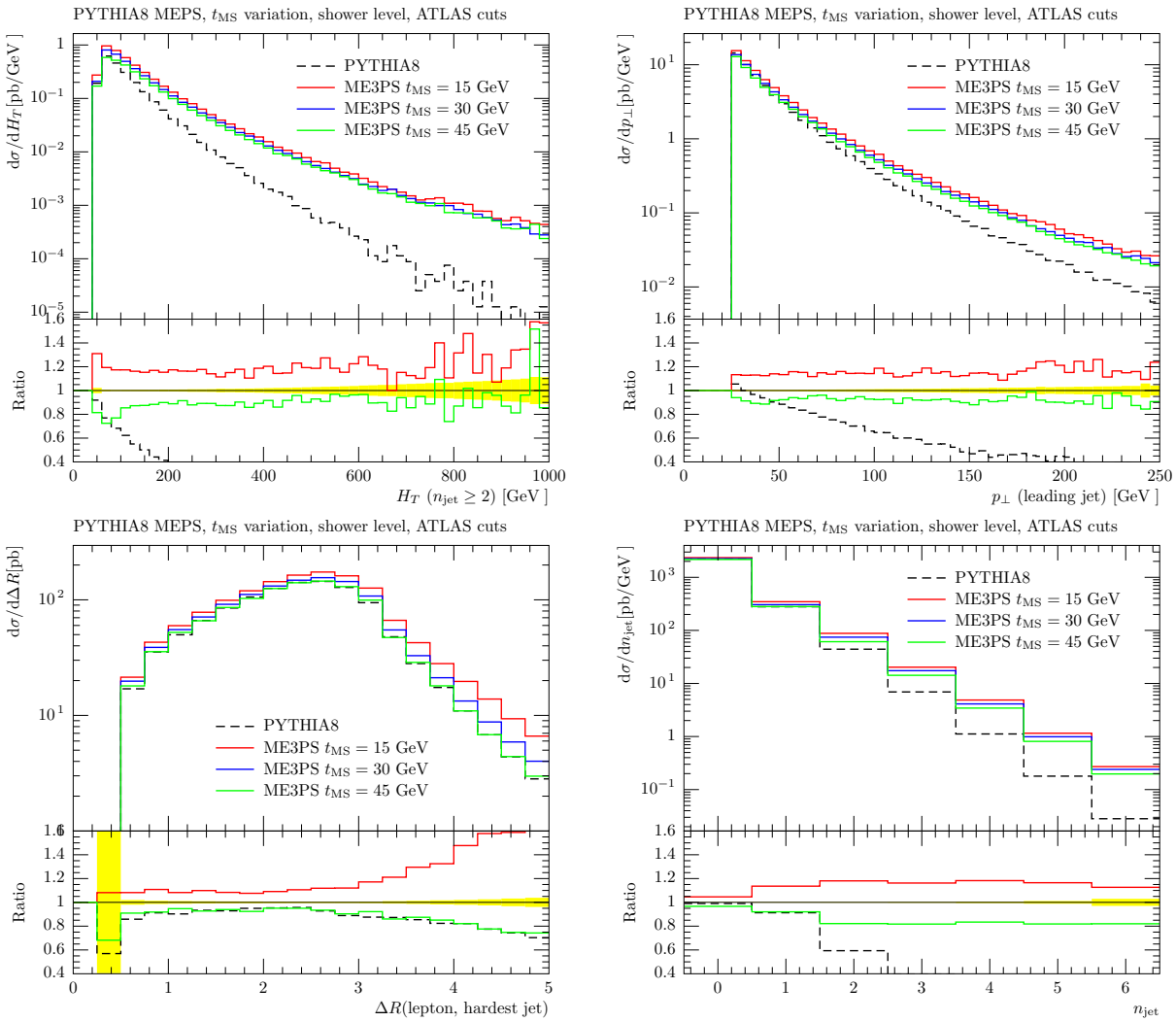


Fig. 78: Variation of the merging scale value for PYTHIA8 at shower level. The plots show the H_T-distribution when requiring at least two jets (upper left), the $p_\perp$ of the hardest jet (upper right), the ΔR-separation of lepton and the hardest jet (lower left), and the number of jets (lower right). The lower insets show the ratio of the samples in the upper half to ME3PS for $t_{\rm MS} = 30$ GeV. All plots are generated using Tune 4C (y-blind treatment).

enforcing rapidity ordering, A2 will look harder. It is debatable whether including rapidity ordering into the tuning makes the tune mimic hard scale effects. The scales at which the fourth jet is produced are certainly close to the scale of (hard) multiple interactions, which is in turn closely connect to soft physics. Although the enforced rapidity ordering in Tune 4C might be considered questionable, we here take the pragmatic approach of considering the evolution both with and without enforced rapidity ordering. From the fact that up to three jets, the merged predictions of Tune 4C and Tune A2 only differ in normalisation, we anticipate that the effect of rapidity ordering will be reduced by merging more jets, since then, the number of jets above a cut-off will be dictated by the matrix element.

In Fig. 78, we investigate the impact of changes in the merging scale value. Again, we mainly see normalisation changes and only small changes in shape, which in most cases are smaller than changes due to different tunes. The R-separation between lepton and hardest jet ΔR(lepton,hardest jet) shows significant shape changes above π. This again is an effect of Tune 4C, and is greatly reduced in Tune A2[40], as

[40]For the sake of compactness, merging scale variation plots for Tune A2 could not be included here. We hope the reader is

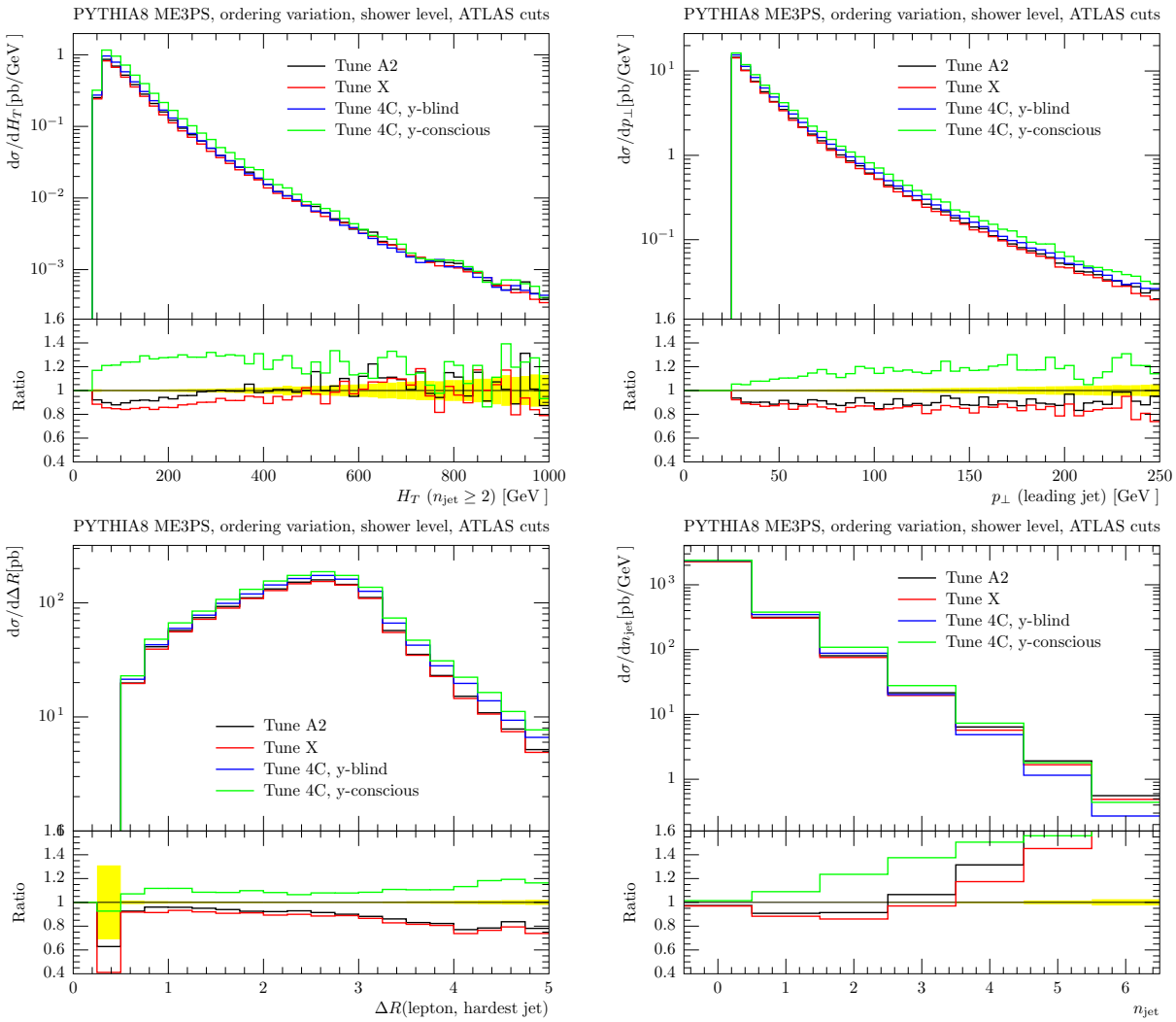


Fig. 79: Variation of the criterion employed to favour "ordered histories" in PYTHIA8 at shower level. The plots show the H_T-distribution when requiring at least two jets (upper left), the $p_\perp$ of the hardest jet (upper right), the ΔR-separation of lepton and the hardest jet (lower left), and the number of jets (lower right). The lower insets show the ratio of the samples in the upper half to ME3PS (Tune 4C, y-blind treatment). All merged plots are produced with a merging scale of $t_{\text{MS}} = 15$ GeV.

can be inferred from the tune variation. However, even in Tune A2, small shape changes remain, with the change becoming less pronounced when comparing two large merging scales. We take this as an indication that the shower splitting probability – giving radiative contributions to $\Delta R(\text{lepton,hardest jet}) > \pi$ for high t_{MS} – and the the matrix element, which fills the same region in for the low merging scale case, are indeed different from the second jet on. This also explains the difference between Tune 4C and Tune A2, which differ by the phase space regions over which the splitting kernels are integrated.

Finally, in Fig. 79, we address the interplay of matrix element merging and ordering in the underlying shower more carefully. The effect of different choices manifests itself again mainly in changes of the normalisation of the plots, and is comparable in magnitude to the impact of merging scale variations. At first, the changes may seem counter-intuitive, and need clarification. For this, it is important to remember the definition of "y-blind" and "y-conscious" in section 18.26. The y-blind treatment will – irrespectively of rapidity configurations – mainly choose histories ordered in the shower evolution variable ρ, and only

nevertheless willing to consider the following argument – which is only supported by the omitted results.

pick ρ-unordered histories if no other ones have been constructed. However, in the y-conscious approach, once no history ordered both in rapidity *and* ρ is found, one amongst all un-ordered histories is chosen probabilistically, irrespectively of the history being y-/ρ-/or y- and ρ-unordered. Since the ordering criterion is stricter, un-ordered histories will be chosen more frequently, meaning that ρ-unordered ones will also contribute more, compared to the y-blind case. Matrix element states with no ordered histories will have a number of jets at very similar scales, so that the Sudakov suppression generated by trial showers will be smaller. Moreover, for matrix element states in which the last reconstructed splitting is unordered, the parton shower will be started at the larger of the unordered scales[41], which can result in a slightly harder spectrum of resolved parton shower jets. Because ρ-unordered states are picked more often when requiring a tighter ordering criterion, this leads in visible differences. The y-conscious method might seem somewhat artificial, considering that it introduces a larger dependence on states outside the range of even the y-unordered shower variant. Nevertheless, the y-blind and y-conscious prescriptions are equivalent to the accuracy of the (y-ordered) shower, so that both should be investigated when assessing the quality of the merging. From the visible changes, we can infer that different treatments of formally sub-leading effects do matter. For the y-ordered evolution, these are more visible since the accuracy of the shower itself is worse, so that the effects of including matrix element states cancel to a lesser degree. It is interesting to note that the deviations between the different prescriptions are considerably smaller if the merging scale is increased, again hinting at a reduced shower accuracy if the evolution is ordered in multiple variables.

Fig. 79 further shows distributions labelled Tune X, which have been generated by using Tune 4C, removing the rapidity constraint on space-like emissions, and treating histories y-blind. Results of these runs, as expected, closely follow Tune A2. The outcome of both Tune A2 and Tune X differs only slightly from the Tune 4C (y-blind) curves, consolidating the conclusion that shifting fractions of ρ-un-ordered histories are responsible for the deviations between the y-blind and y-conscious methods. As in the discussion of tuning variation, the similarity in the results of the merged calculation for Tune 4C and Tune A2 breaks down once we examine jets that are solely produced by the shower, i.e. starting from the fourth jet.

18.37 SHERPA

As described in Sec. 18.27, SHERPA has been run in two modes for this comparison of LHC predictions. The results for the conventional merging of towers of tree-level matrix elements, SHERPA MEPS, are presented in Sec. 18.37 while the results of its enhancement to NLO accuracy in the core W production process, SHERPA MENLOPS, are displayed in Sec. 18.37. As detailed earlier, all parameters have been chosen identically otherwise. The precise requirements regarding the event selection and the definitions of the observables used in this comparsion follow the CMS cut specifications and can be found in App. 18.6.

SHERPA MEPS Figs. 80-83 show the results as obtained by running SHERPA in the MEPS mode for a variety of inclusive and multi-jet observables at different levels of the event generation. All central results are displayed together with their respective uncertainties related to the different sources listed in Sec. 18.27. The layout in all figures is the same: the upper left and right panels respectively show the matrix element level and parton shower level predictions for a given observable. The matrix element level is defined as the event generation phase right before the parton showering. For the MEPS approach this means that modifications necessary for the procedure to work like α_s reweighting and Sudakov rejection have been already included at this level. The predictions presented in all centre panels were generated after enhancing the event generation to include corrections induced by the parton-to-hadron transition and decays of the therein produced primordial hadrons. On top of these soft physics effects, one has to also

[41] This is the default choice in PYTHIA8. Other choices are available to the user.

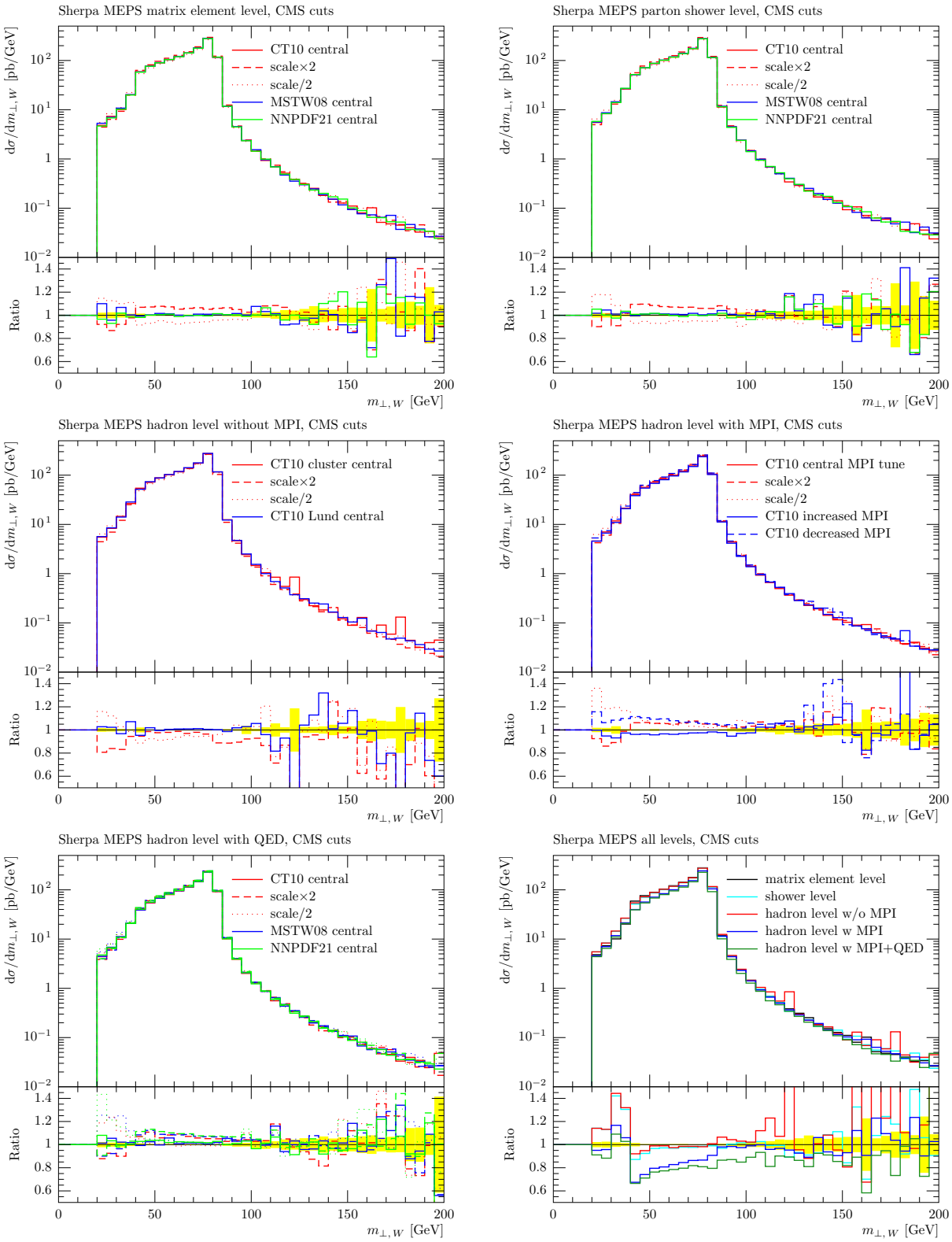


Fig. 80: SHERPA MEPS. Uncertainty of the transverse mass of the reconstructed W on the matrix element level (upper left), after parton showering (upper right), including hadronisation correction (centre left), multiple parton interactions (centre right), and QED corrections (lower left). The lower right panel shows the evolution of the central value.

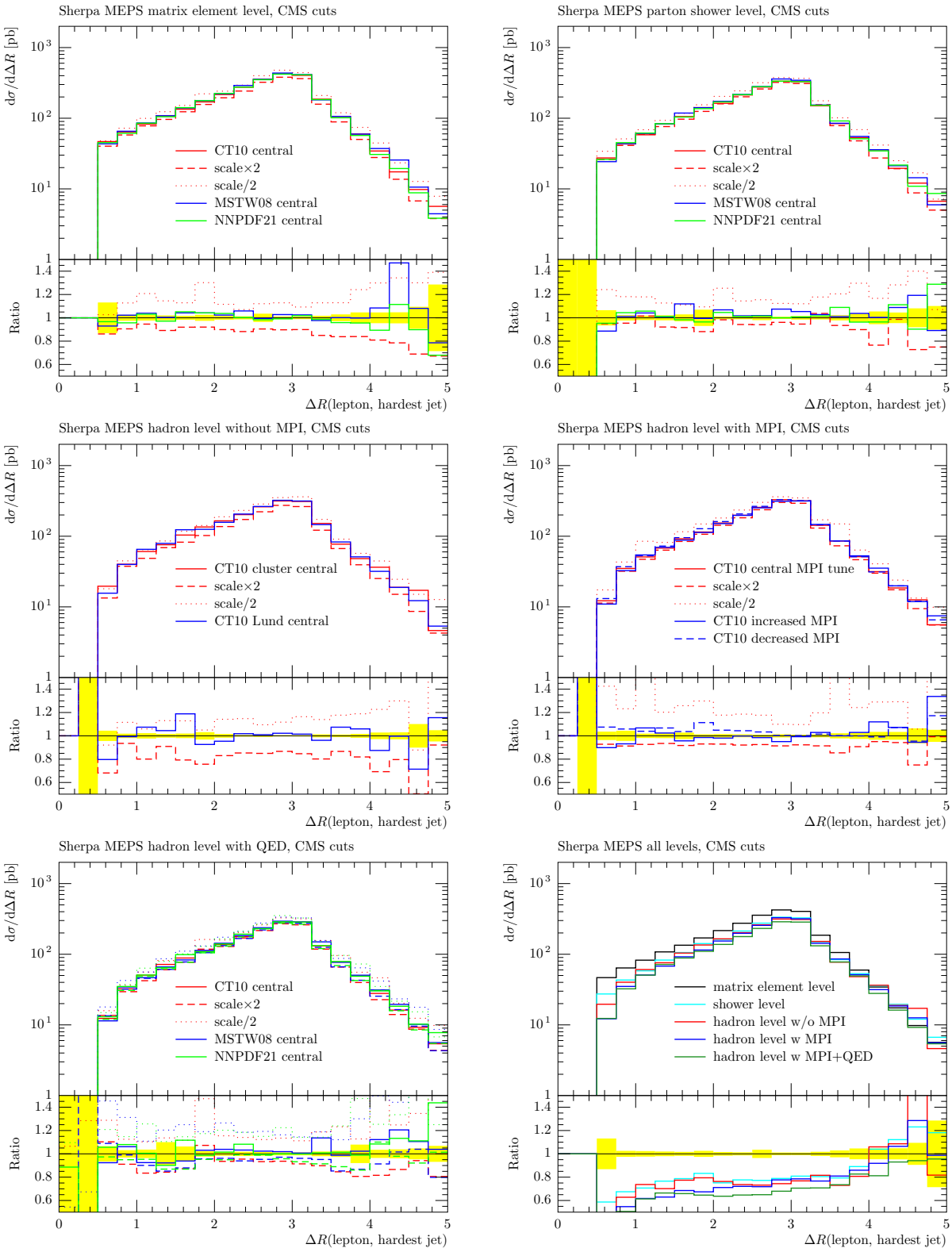


Fig. 81: SHERPA MEPS. Uncertainty of the angular separation of the charged lepton and the hardest jet on the matrix element level (upper left), after parton showering (upper right), including hadronisation correction (centre left), multiple parton interactions (centre right), and QED corrections (lower left). The lower right panel shows the evolution of the central value.

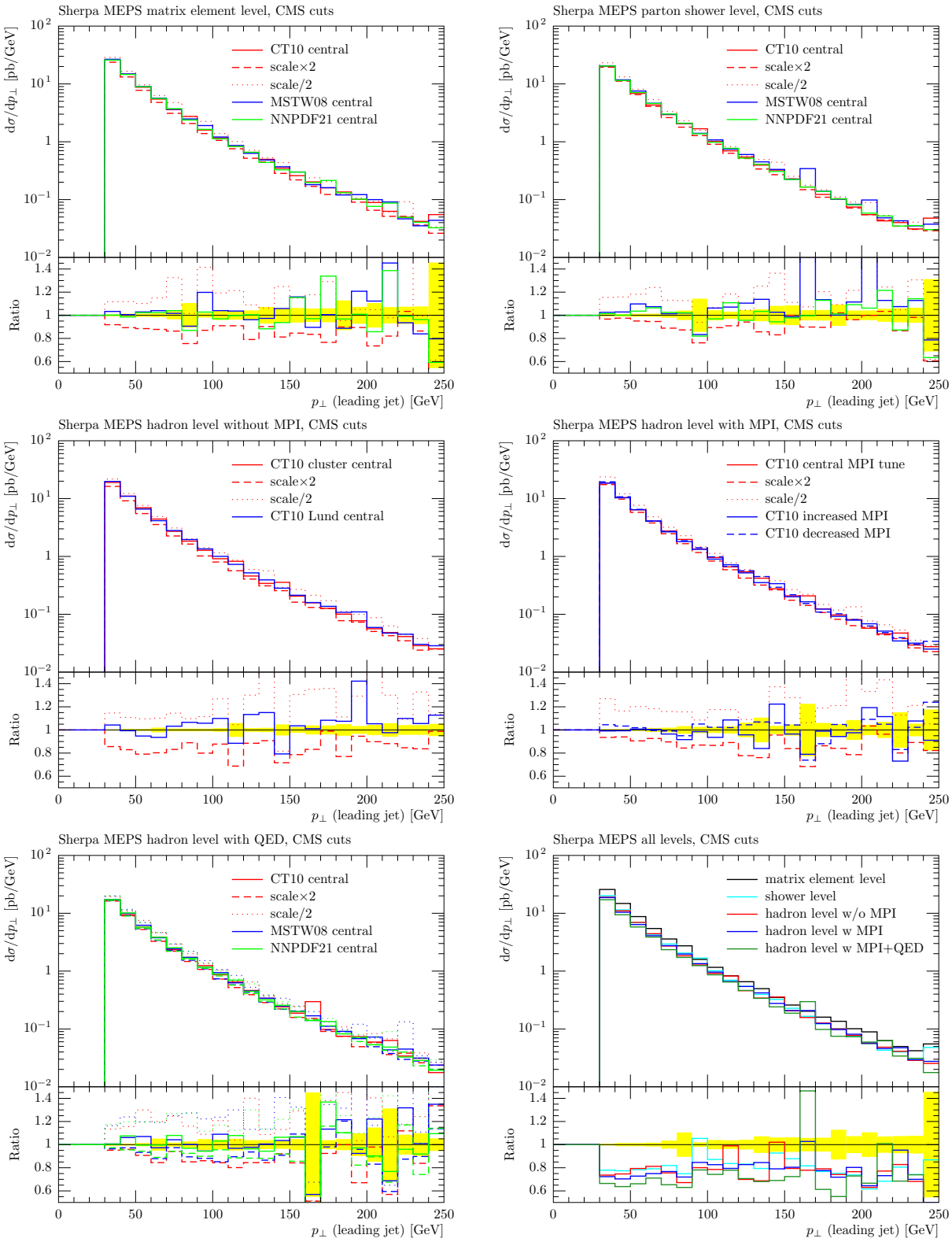


Fig. 82: SHERPA MEPS. Uncertainty of the transverse momentum of the hardest jet on the matrix element level (upper left), after parton showering (upper right), including hadronisation correction (centre left), multiple parton interactions (centre right), and QED corrections (lower left). The lower right panel shows the evolution of the central value.

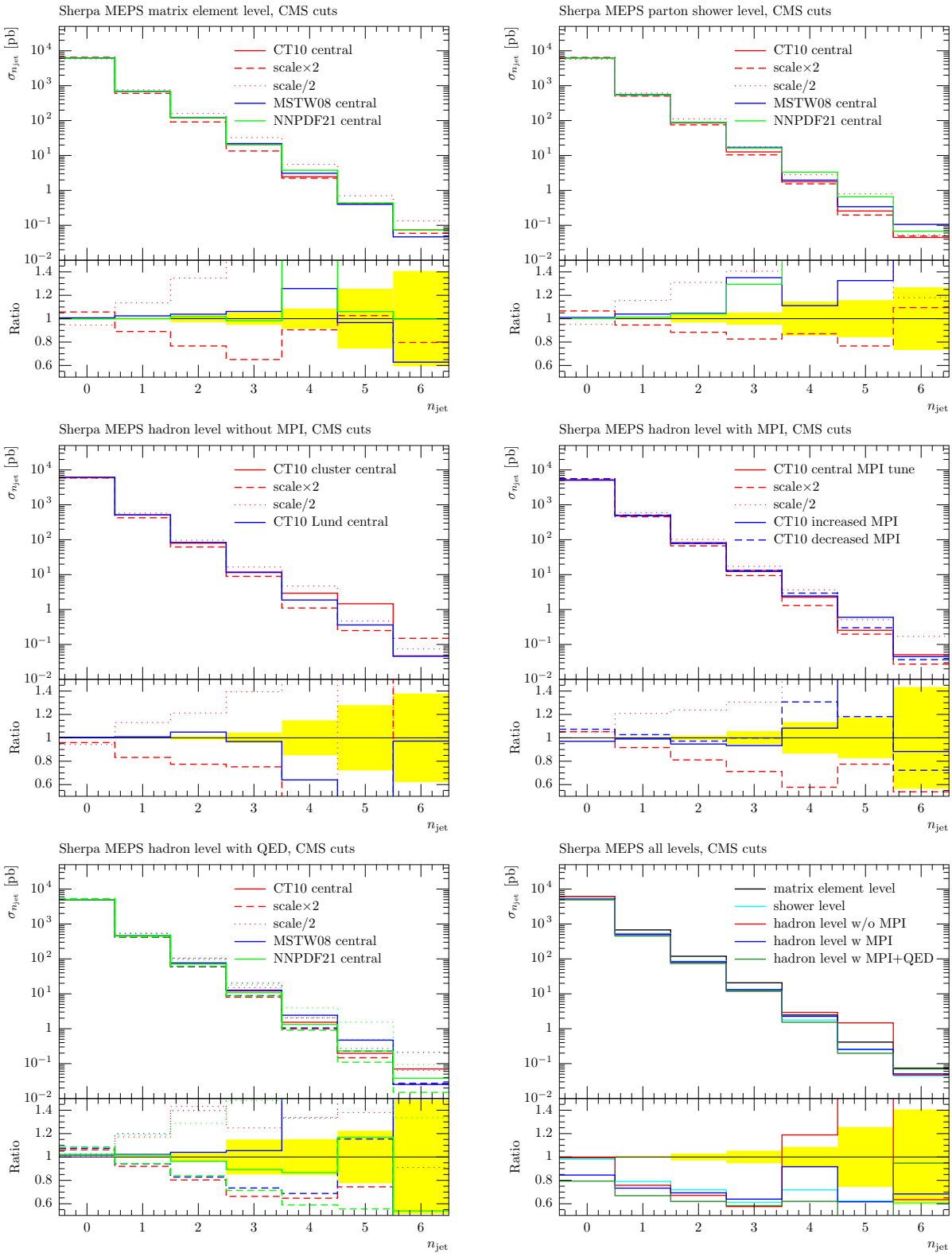


Fig. 83: SHERPA MEPS. Uncertainty of the inclusive jet multiplicity on the matrix element level (upper left), after parton showering (upper right), including hadronisation correction (centre left), multiple parton interactions (centre right), and QED corrections (lower left). The lower right panel shows the evolution of the central value.

account for multiple parton interactions. The results in the centre right panels of all figures incorporate these additional corrections. Finally, all plots to the lower left show the most complete hadron level predictions, which were obtained by adding to the event generation QED radiation effects as occurring in the decays of the vector boson and the hadrons. To allow a direct comparison of the impact of the consecutive event stages, the way the central results change is summarized in all plots to the lower right of Figs. 80-83. In these, as in all other panels, the main plots are supplemented by ratio plots stressing the magnitude of the differences and uncertainties. Note that the yellow band throughout illustrates the statistical uncertainty on the central event sample.

Apart from the summary plots at the lower right, all other cases depict predictions documenting the uncertainty of the central predictions at the different levels of event generation. These uncertainty estimates are gained following the procedures outlined in Sec. 18.27. At all event simulation phases, the scales are varied as described under this section's point (B). Note that the variation is applied to all phases used to make up the respective central (or default) sample, which is taken as the reference under all circumstances. For the matrix element level, parton shower level and full hadron level results, PDF variations according to point (A) are shown in addition, whereas for the centre panel plots, the focus is on the outcomes of the model and tune variations instead, as specified in point (C) and point (D) of Sec. 18.27. Notice that the lower left panels also contain the outcomes of scale variations utilizing the alternative PDFs mentioned under point (A); they are much alike the ones stemming from the default set.

As an example for an inclusive observable the transverse mass of the reconstructed W boson is shown in Fig. 80. The scale uncertainties amount to $\sim$15% at all generation levels, whereas the uncertainties due to the choice of PDF are much smaller. Similarly, the hadronisation uncertainty is negligible. The $m_{\perp,W}$ observable however is more sensitive to the tuning of the MPI model as can be seen from the $\pm$10% envelope in the centre right plot of Fig. 80. The uncertainty is of the same order as for the scale variations, which generally are more pronounced in the soft region. When considering the impact of each perturbative and non-perturbative event stage (see the plot to the lower right), it is the MPI corrections that are largest in the region of $m_{\perp,W} < m_W$, ranging up to $\sim$30% wrt. the matrix element level prediction. They are small above m_W. In this region the dominant effect comes from the QED corrections, which themselves are rather small, but they lead to a contamination of the electron isolation. The application of the isolation cuts then yields a reduction of the overall normalisation of the event sample. Finally there is a small shift towards lower transverse masses, pronouncing the deviation in the tail of the distribution somewhat further.

Fig. 81 and Fig. 82 depict observables that require the presence of at least one jet. In the former the geometric separation, ΔR, between the hardest jet and the electron is shown, while in the latter, focus is on the transverse momentum, $p_\perp$ of the hardest jet only. As before the dependence of the predictions on PDF and hadronisation model changes remains negligible. While the scale dependence of the ΔR and $p_\perp$ variables increases to $\sim$30%, the uncertainty due to the tuning of the MPI model decreases to $\sim$5% when compared to the findings concerning the more inclusive observable considered above. In both cases the reason for the sensitivity change obviously lies in demanding at least one (hard) jet. The scale uncertainties primarily result from changes in the overall cross section. Again, comparing the results of the different event stages, one clearly observes the large impact parton showering has on modifying the matrix element level predictions. The non-perturbative effects go in the same direction amplifying the parton shower effects, but as expected this amplification turns out to be rather mild in the well separated and/or hard phase space regions. QED corrections only play a minor role, and are far less important than for the $m_{\perp,W}$ variable.

In Fig. 83 one of the simplest examples of a multi-jet observable is presented, namely the distribution of the inclusive $W + n$ jet cross sections as a function of n_{jet}. Qualitatively, the parameter and model dependencies of the predictions are found to behave as for the inclusive one-jet variables. As one would expect, the scale uncertainties subsequently increase with the order of the jet bin. The same can be noticed for the variation of the PDFs used in the calculation – even though here the effect is considerably

smaller.

SHERPA MENLOPS Following the outline of the previous subsection, Figs. 84-87 compile the results, which were obtained by executing SHERPA in the MENLOPS mode, cf. Sec. 18.27. The presentation is based on the same set of figures where the selection of the observables has been taken as in the MEPS case. Again, all (central) predictions are examined towards their scale, PDF, non-perturbative modeling and QED simulation dependence. One small difference has to be pointed out: the plots to the lower left now depict exclusively to what extent the additional QED corrections modify the outcomes including multiple parton interactions and hadronisation effects.

Fig. 84 shows the transverse mass of the reconstructed W boson. In the MENLOPS approach, this observable is described at NLO accuracy, which leads to a reduction of the associated scale uncertainties. The scale variation results for $m_{\perp,W}$ nicely confirm this expectation as can be seen in the upper four plots of Fig. 84. The deviations from the central prediction are much smaller than those found for the MEPS scenario exhibited in Fig. 80; they now are of similar magnitude as the PDF uncertainties. While the scale dependence is reduced, PDF and MPI tune variations as well as QED corrections manifest themselves as in the MEPS case. In particular, the discussion around Fig. 80 explaining the effects of extra QED emissions (as being most relevant in the W decay) can be used to understand the findings illustrated in the bottom left panel of Fig. 84.

The MENLOPS method primarily improves the precision of the description of the core process, here the description of the W production process. One also benefits from improving the overall normalisation. However, processes with additional partons in the final state are described in the MENLOPS approach at the same level of accuracy as in the MEPS approach – in both cases by tree-level matrix elements. Thus, the one-jet observables, ΔR between the lepton and leading jet and the $p_\perp$ of the leading jet, and their related uncertainties turn out to be predicted in a very similar manner. This can be clearly observed by comparing Figs. 85-86 with Figs. 81-82. Unlike the findings for $m_{\perp,W}$, it particularly can be noticed that the scale dependence associated with the one-jet observables shown here remains unchanged when compared to the respective MEPS results.

Fig. 87 depicts the distribution of the inclusive $W + n$ jet cross sections as obtained for the MENLOPS case. Using the above reasoning, one can understand these results as for the one-jet variables. Note that the scale dependence of the zeroth jet bin shows the expected decrease owing to the NLO accuracy underlying the description of the core process.

Fig. 88 finally, highlights the evolution and uncertainties of two definition of the beamthrust, cf. App. 18.6: a physical observable summing over all final state particles excluding the W-constituent lepton and a pseudo-observable including the W itself. For both observables small perturbative uncertainties are completely burried underneath much larger non-perturbative effects and modelling uncertainties, an effect also seen in results from the POWHEG BOX+PYTHIA simulation, cf. Fig. 76. This can only be interpreted as this observable being dominated by non-perturbative effects and in particular the underlying event, which somewhat invalidates statements about the merit of this observable in a clean determination of initial state radiation effects made in [469, 26].

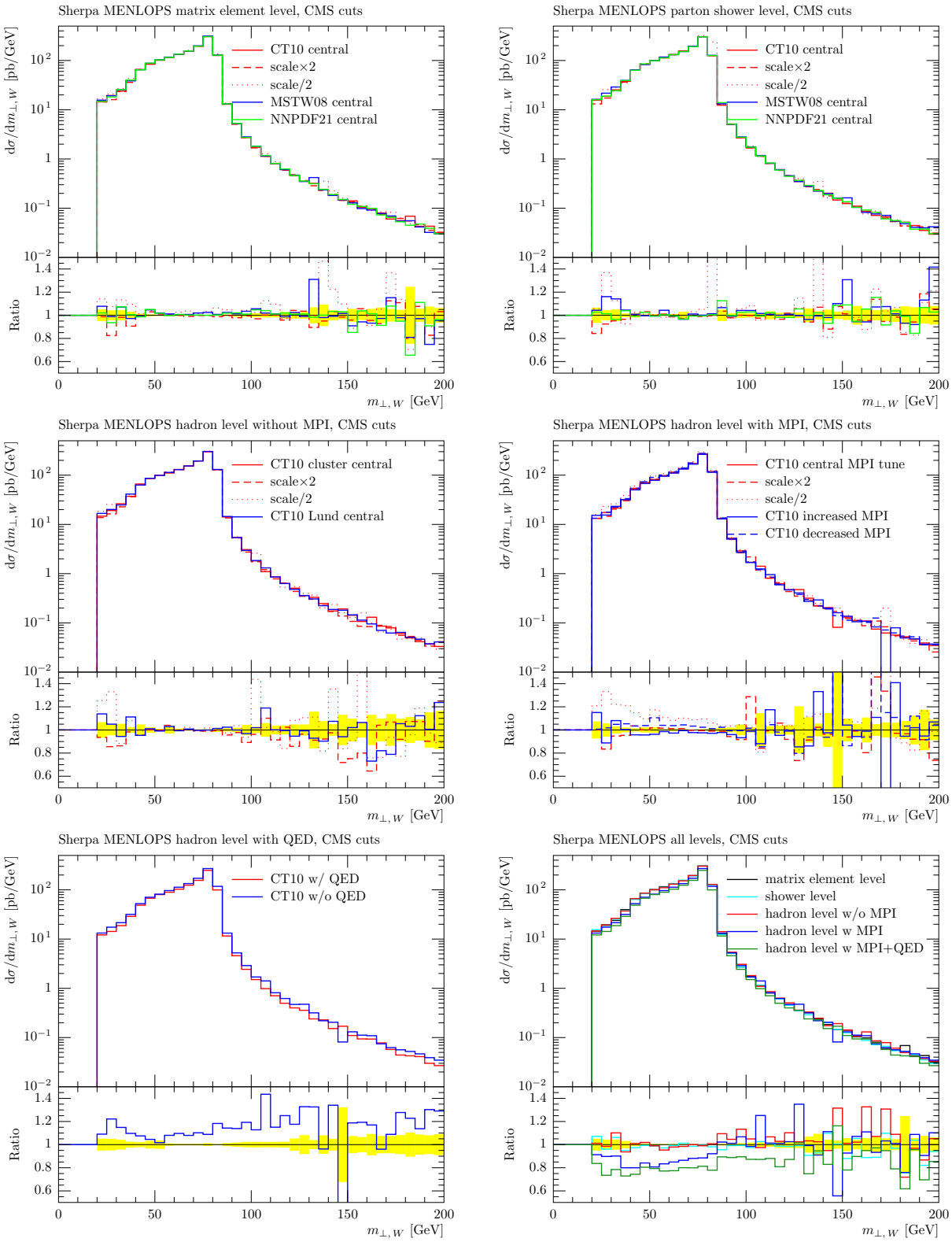


Fig. 84: SHERPA MENLOPS. Uncertainty of the transverse mass of the reconstructed W on the matrix element level (upper left), after parton showering (upper right), including hadronisation correction (centre left), multiple parton interactions (centre right), and QED corrections (lower left). The lower right panel shows the evolution of the central value.

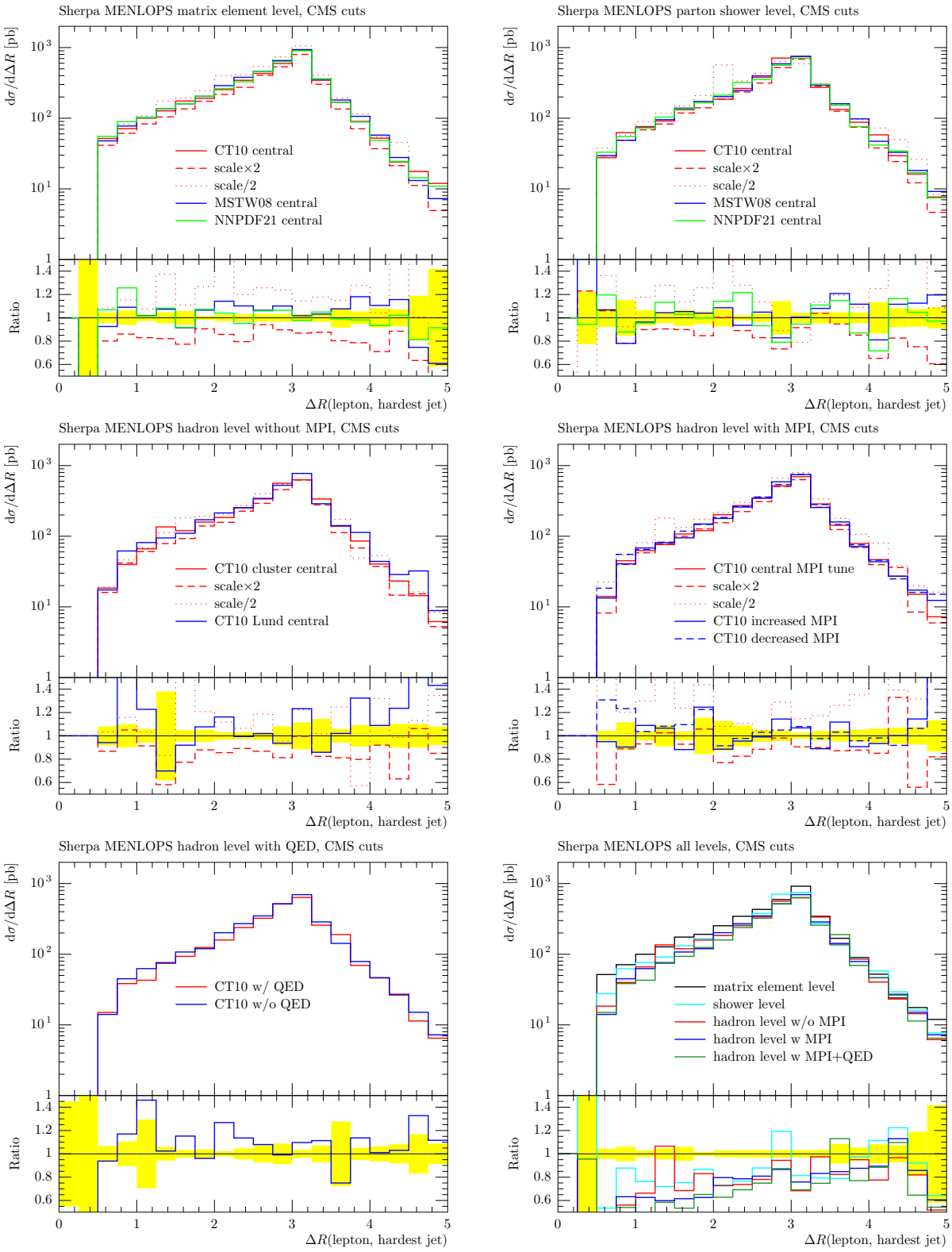


Fig. 85: SHERPA MENLOPS. Uncertainty of the angular separation of the charged lepton and the hardest jet on the matrix element level (upper left), after parton showering (upper right), including hadronisation correction (centre left), multiple parton interactions (centre right), and QED corrections (lower left). The lower right panel shows the evolution of the central value.

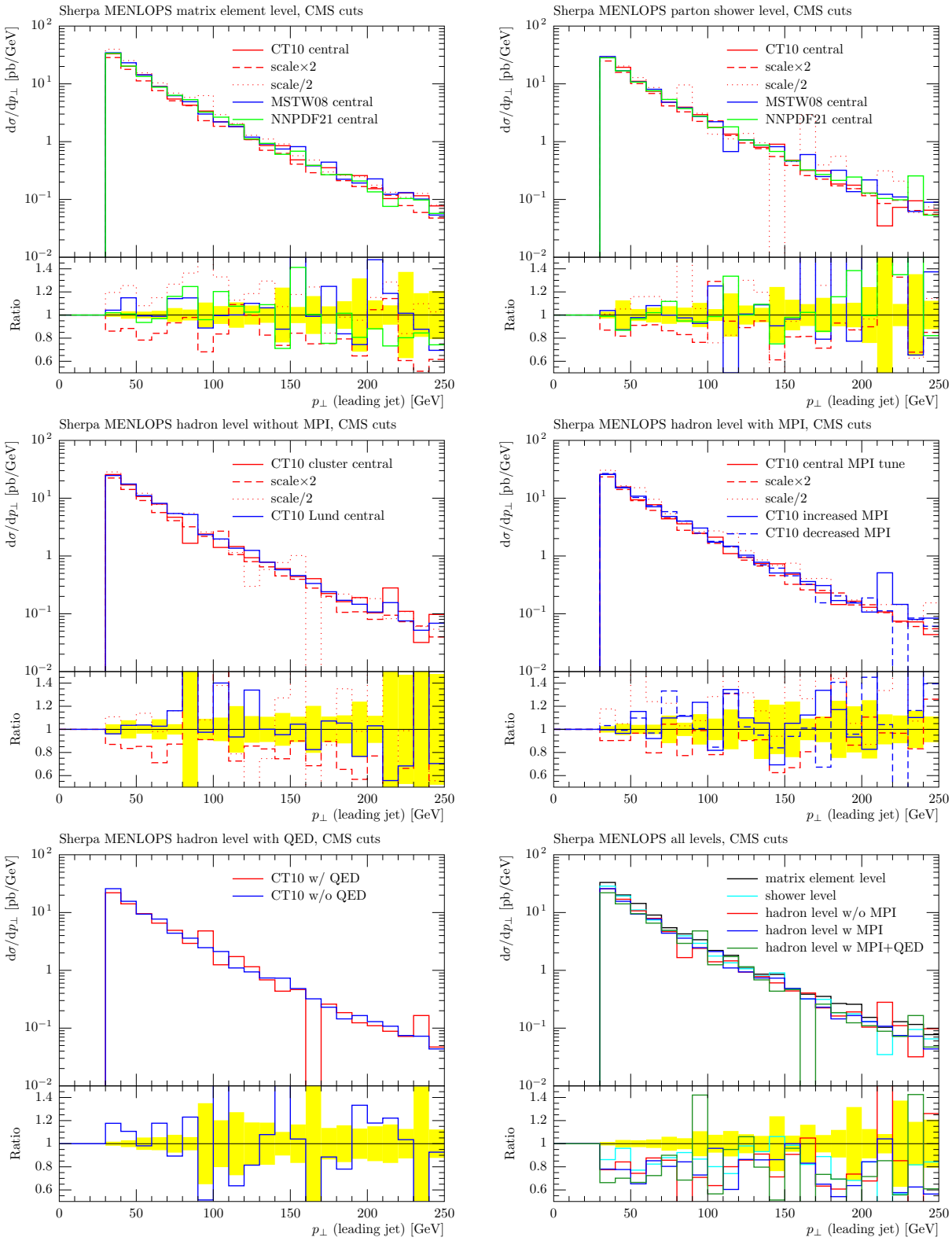


Fig. 86: SHERPA MENLOPS. Uncertainty of the transverse momentum of the hardest jet on the matrix element level (upper left), after parton showering (upper right), including hadronisation correction (centre left), multiple parton interactions (centre right), and QED corrections (lower left). The lower right panel shows the evolution of the central value.

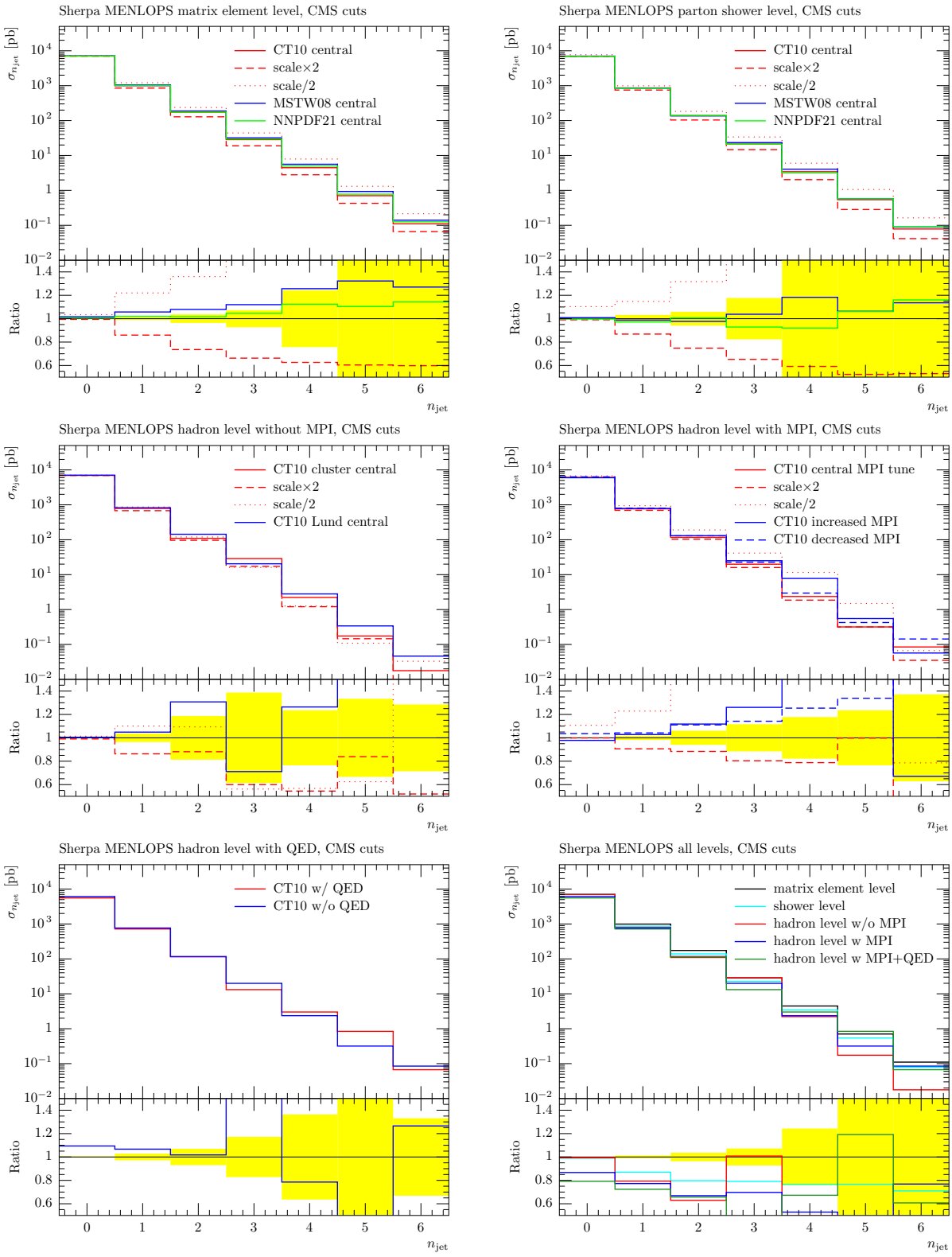


Fig. 87: SHERPA MENLOPS. Uncertainty of the inclusive jet multiplicity on the matrix element level (upper left), after parton showering (upper right), including hadronisation correction (centre left), multiple parton interactions (centre right), and QED corrections (lower left). The lower right panel shows the evolution of the central value.

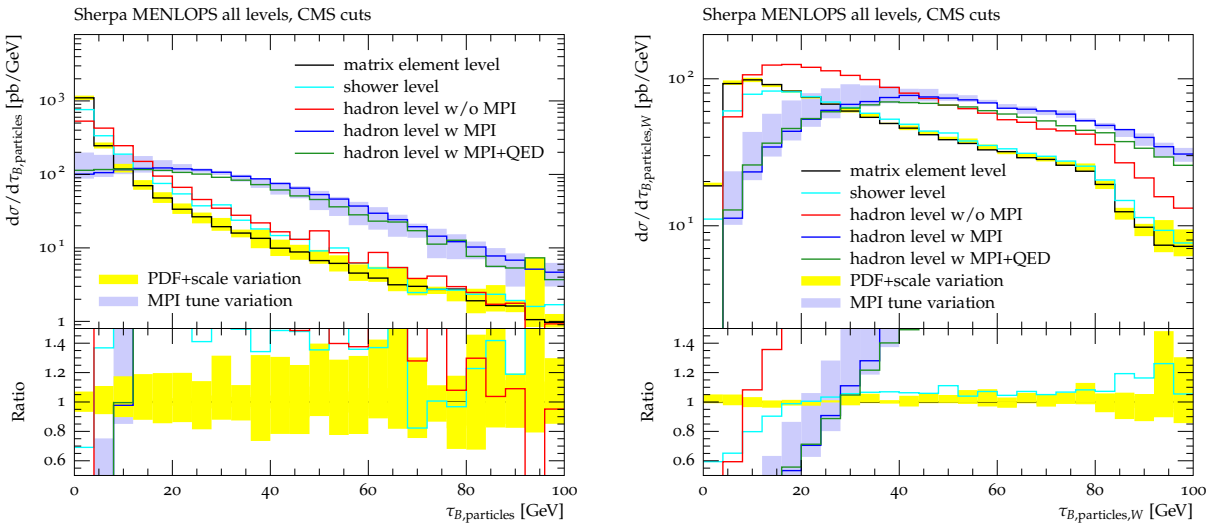


Fig. 88: SHERPA MENLOPS. Evolution and uncertainty of two definitions of the beamthrust, calculated using all particles not constituting the W (left) and including the W (right). Exemplary, the combined PDF and scale uncertainty on the matrix element level prediction (yellow) and the modeling uncertainty of the hadron level prediction (blue) are shown.

18.4 Comparisons

In this section we compare the results of different tools with each other. While the aim of this study was to have a fairly tuned comparison with as many aspects of the calculations as possible being centrally defined, there are still important residual differences in the various results. Obviously, the different codes produced results at different stages of the simulation, which are not always directly comparable; in addition, some of these stages are not very straightforward to obtain: for instance, running PYTHIA8 without multiple parton interactions included in the interleaved showering obviously changes the overall logic of the parton shower model of this code. In addition, other, more obvious differences occur, ranging from inconsistent choices of PDFs to different strategies in scale setting procedures. For the case of the PDFs, by directly comparing results obtained with BLACKHAT+SHERPA using CTEQ6.6 and with GOSAM+SHERPA using CT10, it appears as if at NLO these differences are minor. However, it is not clear how much of the differences between MADGRAPH+PYTHIA and PYTHIA8, which both employ CTEQ6L1, and the other codes, which employ NLO PDFs, can be attributed to differences in PDFs.

In addition, results obtained with the NLO codes typically include at least one jet - POWHEG BOX+PYTHIA8 and GOSAM+SHERPA take $W + 1$ jet at NLO as their core process - while HEJ starts at $W + 2$ jets, and BLACKHAT+SHERPA presents results for up to 4 jets accompanying the W boson in different jet bins. Obviously, on the other hand, the multijet merged samples of MADGRAPH+PYTHIA, PYTHIA8 MEPS and SHERPA include LO matrix elements for up to 3 to 6 jets.

In the plots in this section each code is shown with a yellow error band, which is the envelope of the variations presented in Sec. 18.3. The only exception is BLACKHAT+SHERPA, which is shown with a blue error band. In the ratio plots the codes are plotted relative to BLACKHAT+SHERPA, also at the parton shower level.

18.41 Inclusive observables

In this section we present some inclusive observables, which are typically all obtained from codes employing multijet merging. By and large, all codes agree in the shapes of the $m_{\perp,W}$ distribution at different stages, although there are sizable differences in the respective normalisation of the samples.

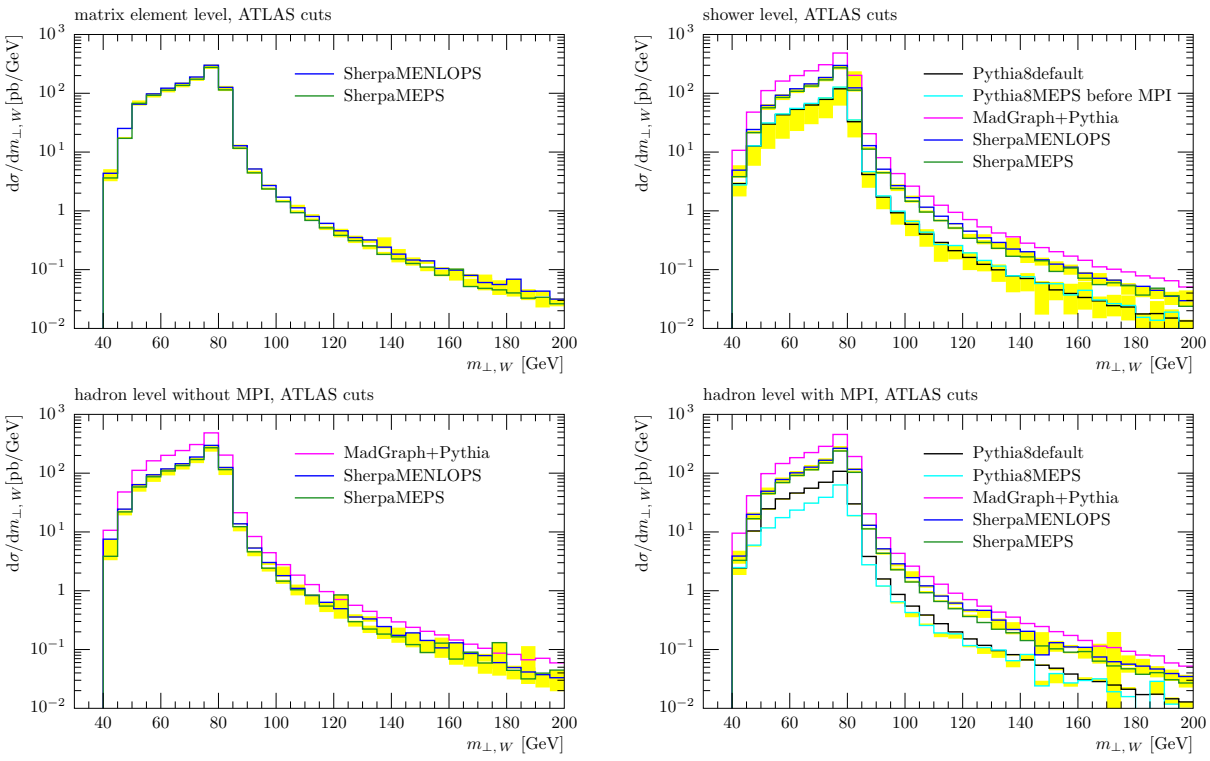


Fig. 89: Transverse mass of the reconstructed W on all levels of the simulation, for the exact definition see App. 18.62 and for the cuts employed in the analysis App. 18.61. Note that PYTHIA8 and MADGRAPH+PYTHIA use the CTEQ6L1 pdf, while SHERPA uses CT10.

18.42 Observables with at least one jet

As a first and fairly telling observable the $p_\perp$-spectrum of the hardest jet is compared, cf. Fig. 90. At the parton level, the results of the NLO calculations – BLACKHAT+SHERPA and GOSAM+SHERPA– agree nearly perfectly with each other and within about 20% with the multijet merged samples of SHERPA, both at LO (SHERPA MEPS) and in the MENLOPS (SHERPA MENLOPS) sample. The increase of the latter with respect to the former at relatively low transverse momenta of about 50 GeV or below can probably be related to the different scale definition in the argument of the strong coupling, where the NLO calculations choose $\mu_R^2 = (H_T'/2)^2 \approx M_W^2/4 + p_{\perp,j}^2$ while in the SHERPA simulation the transverse momentum of the jet has been chosen. Clearly, for small transverse momenta this will lead to visible differences. Going from the matrix element to the parton shower level typically leads to the jets becoming softer and to losing some of them, due to partons emitted outside the jet and a corresponding energy loss. This explains why the SHERPA distribution at the shower level is softer than the NLO result, and thus the SHERPA result at the matrix element level, although the size of the difference seems to be larger than one would naïvely expect. This finding is, however, somewhat at odds with the results obtained from MADGRAPH+PYTHIA, which seem to be slightly harder in shape and significantly larger in normalisation. The PYTHIA8 MEPS sample, on the other hand, has a smaller one-jet inclusive cross section than SHERPA, but the jet spectrum exhibits a somewhat harder tail, corresponding to a shape difference of about 30-40% with respect to both the SHERPA results. The same finding, a somewhat harder tail, is also true for the POWHEG BOX+PYTHIA8 results. The same trends can be also found at the hadron and hadron + MPI level. For the POWHEG BOX result the difference can be attributed to the usage of a scale defined at the "underlying-Born" level (cf. Sec. 18.25 for more details). Indeed it has been checked explicitly that a NLO computation performed with the same scale choice used in POWHEG BOX gives a result in complete agreement with the POWHEG BOX result shown here. Clearly, the differences between different

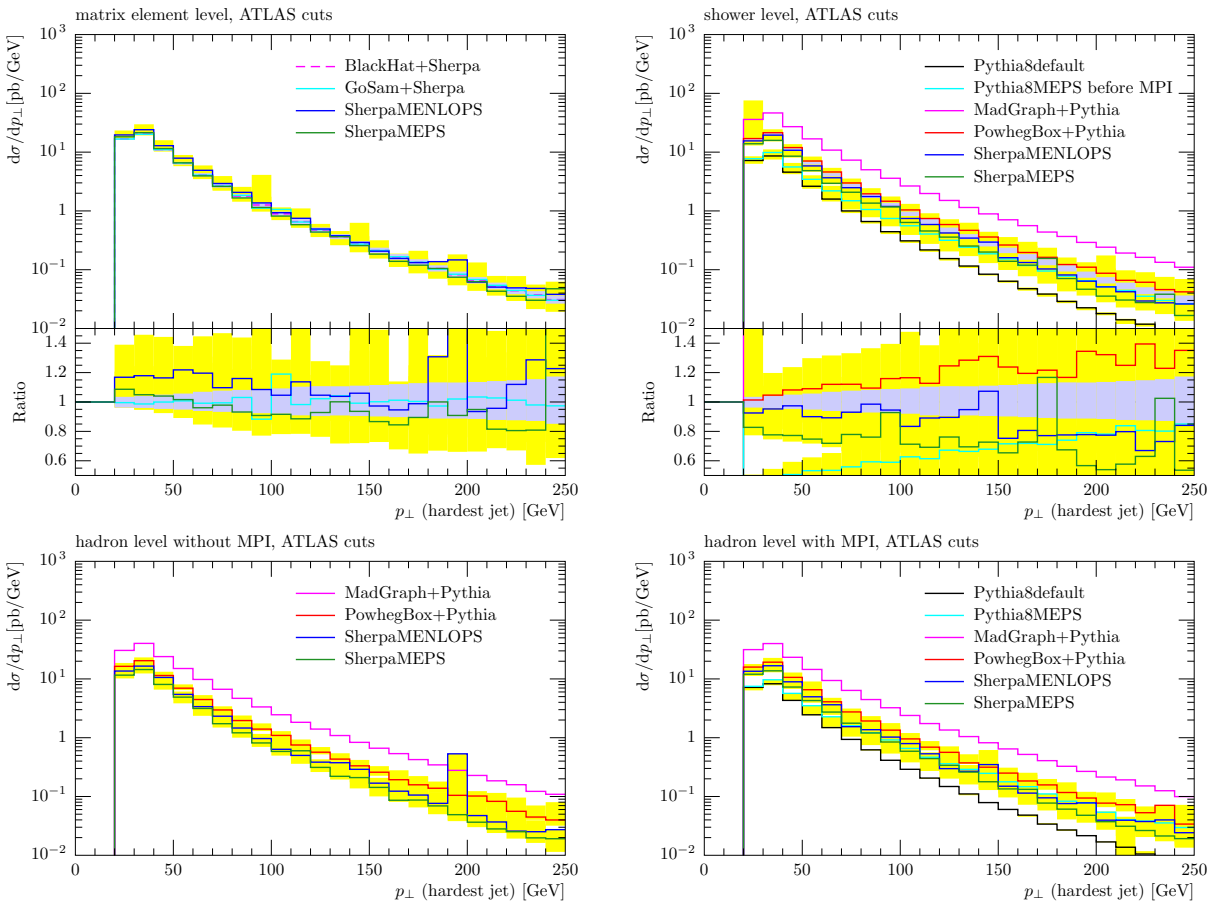


Fig. 90: Transverse momentum of hardest jet on all levels of the simulation, where jets are reconstructed using the anti-$k_\perp$ with $R = 0.4$ within $|\eta| < 4.4$ (for exact definitions and cuts see App. 18.61 and App. 18.62). Note that BLACKHAT uses the CTEQ6.6 pdf, PYTHIA8 and MADGRAPH+PYTHIA CTEQ6L1 and all the others use CT10. In both ratio plots the ratio is taken with respect to BLACKHAT+SHERPA (on matrix element level).

calculations and codes exhibited here deserve a more in-depth study, which, unfortunately, is beyond the scope of this comparison.

Similar findings are also true for the next observable, the ΔR distribution between the lepton stemming from the W decay and the hardest jet displayed in Fig. 91. Again, the two SHERPA samples are compared with the two NLO samples, this time exhibiting a sizable shape difference towards an increase at smaller and a decrease at larger distances of about 40% relative cross section. While higher jet configurations typically tend to be a bit more central, it seems far-fetched to attribute this difference only to them. At the same time, large differences in R are most likely due to jets which are pretty much forward[42]. This region of phase space for jet production, however, is known to be quite susceptible to mismatches in scale and/or PDF definitions. However, it is worth noting that this difference vanishes almost completely at the parton shower level. The PYTHIA8 MEPS sample, despite a sizable difference in cross section, appears to follow the shape of the NLO and SHERPA results. Further comparing these results to those of the other codes at the shower level suggests that the MADGRAPH+PYTHIA merged sample, apart from a drastically enhanced cross section, also shows an enhancement in shape at smaller $\Delta R \leq 2$ w.r.t. the NLO result. Interestingly enough, the POWHEG BOX+PYTHIA8 sample exhibits the

[42] Assuming the lepton and the jet to be back-to-back, $\Delta\phi = \pi$, one still needs $\Delta\eta \approx 4$ to obtain $(\Delta R)^2 = (\Delta\phi)^2 + (\Delta\eta)^2 \approx 5^2$.

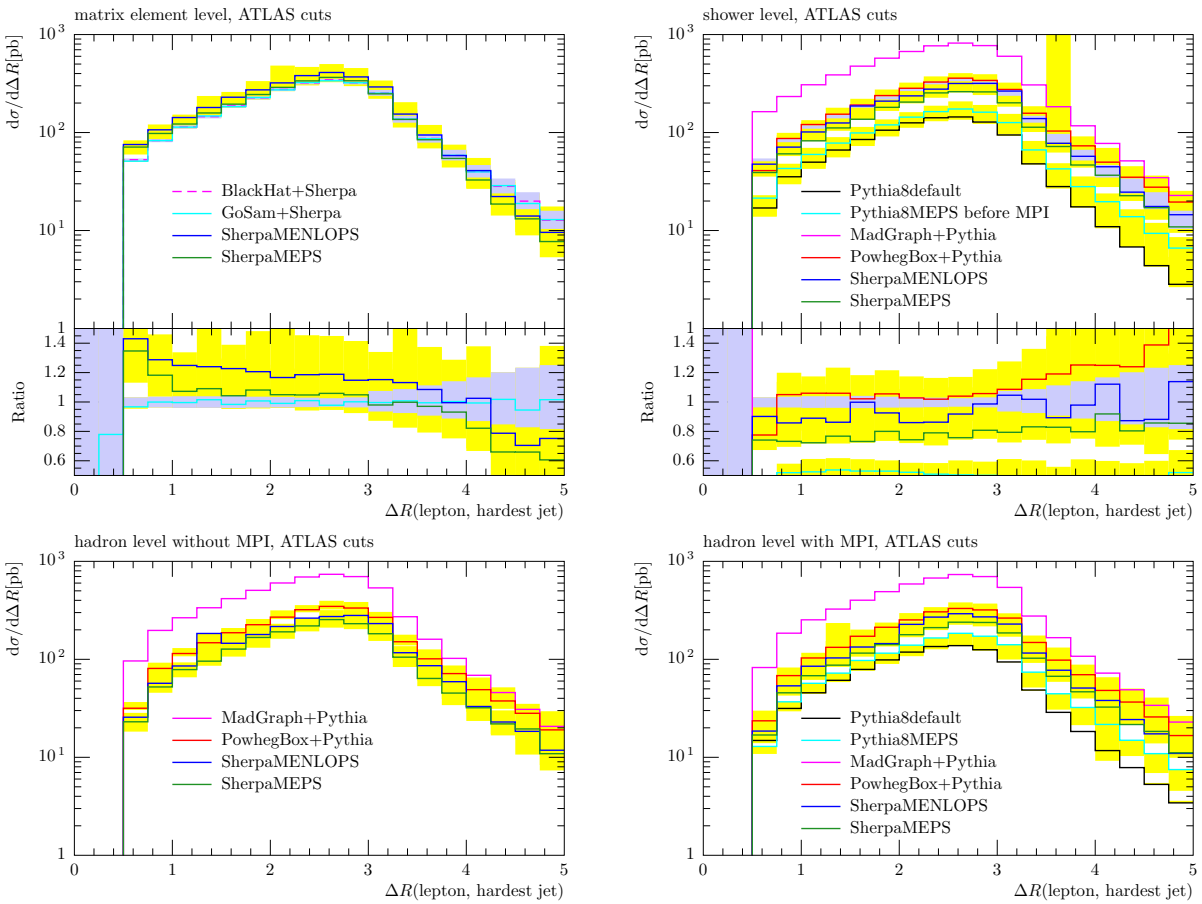


Fig. 91: ΔR between hardest lepton and hardest jet on all levels of the simulation (for exact definitions and cuts see App. 18.61 and App. 18.62). Note that BLACKHAT uses the CTEQ6.6 pdf, PYTHIA8 and MADGRAPH+PYTHIA CTEQ6L1 and all the others use CT10. In both ratio plots the ratio is taken with respect to BLACKHAT+SHERPA (on matrix element level).

opposite behaviour: while the cross section seems fairly consistent with the SHERPA and the NLO ones, the shape shows some enhancement of up to 40% at large distance ΔR, which following the reasoning for the jet-$p_\perp$ spectrum may also hint at being due to a difference in the definition of scales. As before, the same trends visible at the parton shower level can also be found at the hadron and hadron + MPI level.

18.43 Multi-jet observables

In observables including at least two jets, consider first the case of the H_T distribution depicted in Fig. 92. Over the full range and obscured by large statistical fluctuations both SHERPA samples seem to follow the NLO prediction from BLACKHAT+SHERPA. The LO result from GOSAM+SHERPA, on the other hand, appears to fall off at the hard end of the distribution. The prediction from HEJ is a bit more subtle to judge: at low H_T (around 100 GeV), we see that it is in good agreement with the predictions from the other approaches. However, as higher values of H_T are probed, the HEJ prediction becomes noticeably larger than the fixed-order descriptions, including those from SHERPA where different multiplicities are merged. This is the region in H_T where we would expect high multiplicities to have a noticeable effect, and therefore where we would expect to see the impact of the resummation in HEJ. This is, however, slighlty at odds with the fact that the SHERPA prediction included up to 6 jets and that the multijet rates and the $p_\perp$ distributions of the fifth and sixth jet from HEJ undershoot those from SHERPA, cf. Fig. 93

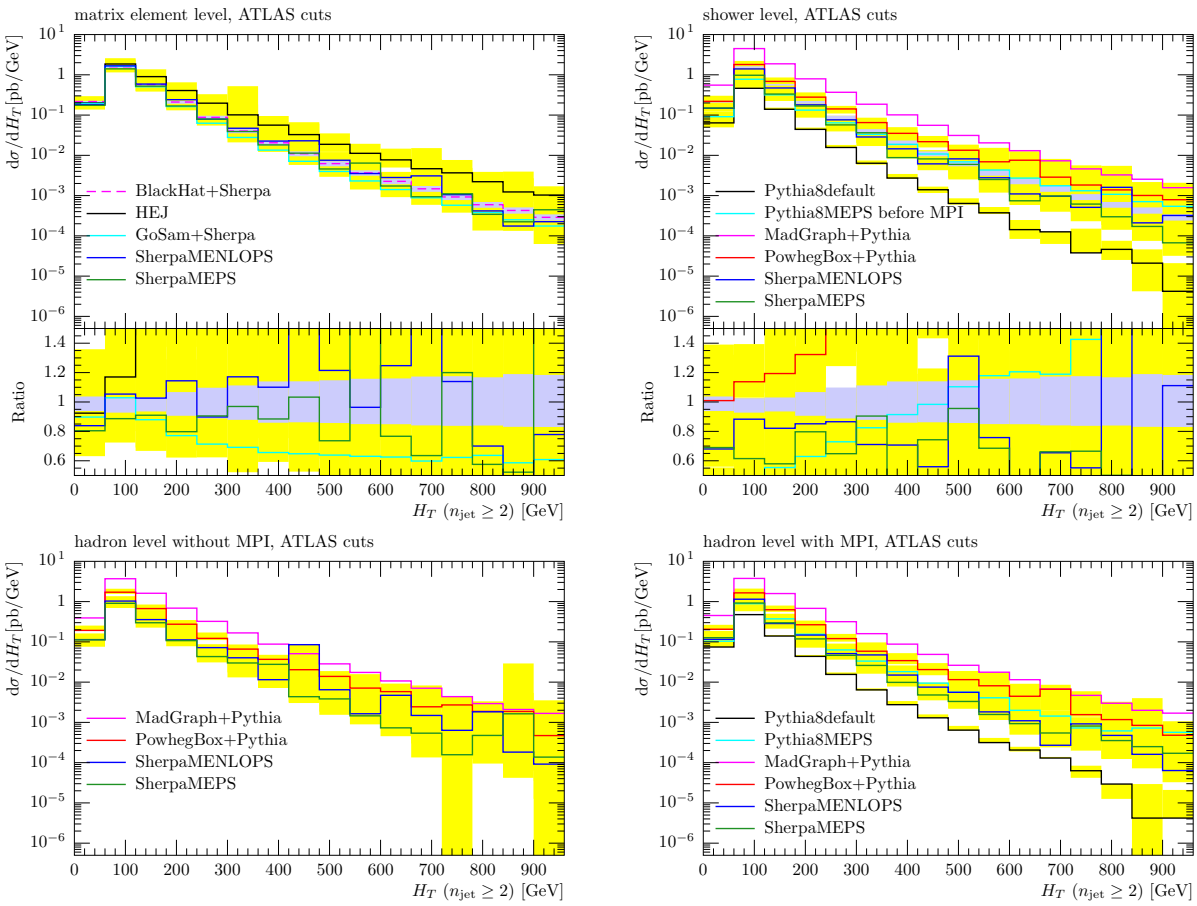


Fig. 92: $H_T = \sum_{i \in \{\text{jets}\}} E_{\perp i}$ of events with at least 2 jets on all levels of the simulation (for exact definitions and cuts see App. 18.61 and App. 18.62). Note that BLACKHAT uses the CTEQ6.6 pdf, PYTHIA8 and MADGRAPH+PYTHIA CTEQ6L1 and all the others use CT10. In both ratio plots the ratio is taken with respect to BLACKHAT+SHERPA (on matrix element level).

and Fig. 95. However, a similar trend concerning the hard tail of this distribution appears also on the shower level in the MADGRAPH+PYTHIA sample, which includes up to 4 extra jets, and in the POWHEG BOX+PYTHIA8 sample, which includes 2 jets at LO and 1 jet at NLO. The trend is even more pronounced with an even harder tail for the PYTHIA8 MEPS sample, which includes 3 extra jets. At this level, SHERPA more or less follows the NLO result. It should be noted, though, that all approaches remain within the scale variation band indicated on the BlackHat prediction. This findings are consistently carried over to the hadron and hadron+MPI level.

Turning to the n-jet rates, at the matrix element level, SHERPA follows fairly closely the NLO results in different jet multiplicity bins, while HEJ seem to overshoot the central value in the 3- and 4-jet bin, but staying inside the NLO scale uncertainty band. going back to the tree-level result of SHERPA in the 5- and 6-jet bins. As discussed in Sec. 18.33, the HEJ framework includes tree-level matching for final states with up to and including four jets in the final state. Therefore it is fair to assume that the absence of matching for five jets and above leads to the larger drop in cross section observed in Fig. 93 from four-jet to five-jet as compared to that from either three-jet to four-jet or from five-jet to six-jet and lends support to the suspicion that in HEJ a matched sample would also provide larger 5- and 6-jet multiplicities. At the shower level, the trend already visible at the H_T distribution repeats itself. The smaller cross section in the PYTHIA8 MEPS sample is mainly due to the low multiplicity bins, such that the shape of the n-jet distribution also has a relatively harder tail than the SHERPA sample. In contrast,

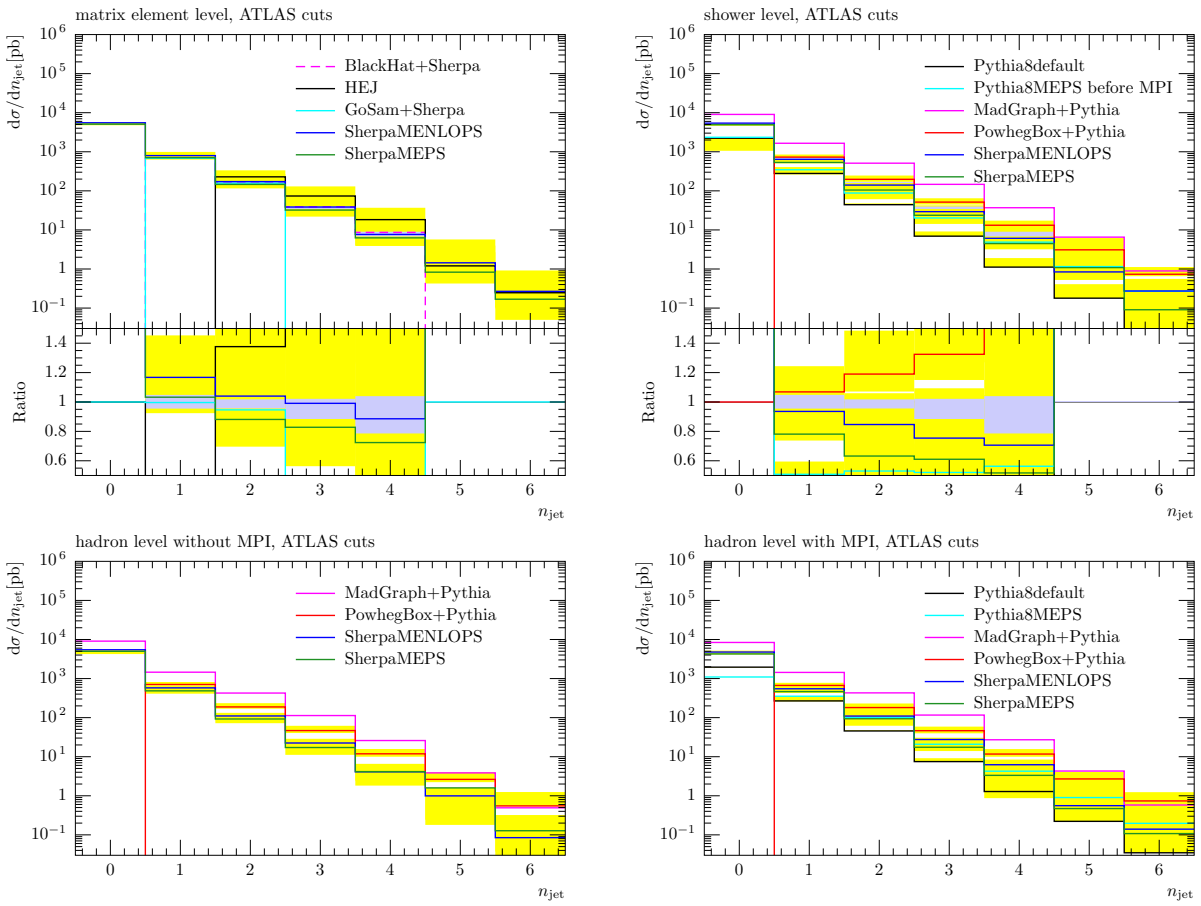


Fig. 93: Number of jets on all levels of the simulation (for exact definitions and cuts see App. 18.61 and App. 18.62). Note that POWHEG BOX+PYTHIA8 and GOSAM calculate $W + 1$ jet on matrix element level, while HEJ starts with $W + 2$ jets and that BLACKHAT uses the CTEQ6.6 pdf, PYTHIA8 and MADGRAPH+PYTHIA CTEQ6L1 and all the others use CT10. In both ratio plots the ratio is taken with respect to BLACKHAT+SHERPA (on matrix element level).

the POWHEG BOX +PYTHIA8 result, starting consistently at 1 jet, appear to be at the upper end of the NLO uncertainties throughout.

Looking at the correlation of the two leading jets in Fig. 94 at the matrix element only, both the ΔR and the m_{12} distribution provided by SHERPA have a slight tilt against the NLO prediction from BLACKHAT+ SHERPA, undershooting the latter result by up to about 40% for large ΔR and by up to about 20% for large m_{12}. While HEJ seems to roughly follow the shape of SHERPA for ΔR, it is significantly harder than SHERPA and the NLO result for large values of m_{12}. In addition, in both cases, HEJ also predicts a larger cross section that the other tools.

Fig. 95 shows the transverse momentum distributions for the third to sixth jets ordered in $p_\perp$, and at the matrix element level. For the third hardest jet, the prediction from HEJ is similar in shape but higher in cross section than the results obtained at NLO from BLACKHAT+SHERPA or the two SHERPA samples. For the fourth jet, the HEJ cross section still seems higher than the other ones, but this discrepancy seems to be mainly around comparably low jet $p_\perp$. For larger values of $p_\perp$ all tree-level type or resummed predictions are below the NLO result. Surprisingly, for the fifth and sixth hardest jets, the HEJ predictions follow the SHERPA ones for low values of $p_\perp$ below about 60 GeV, before they fall off nearly instantly. This again may be an artefact of tree-level matching not being included in HEJ for the production of five- and six-jets or of missing statistical support in this region of phase space.

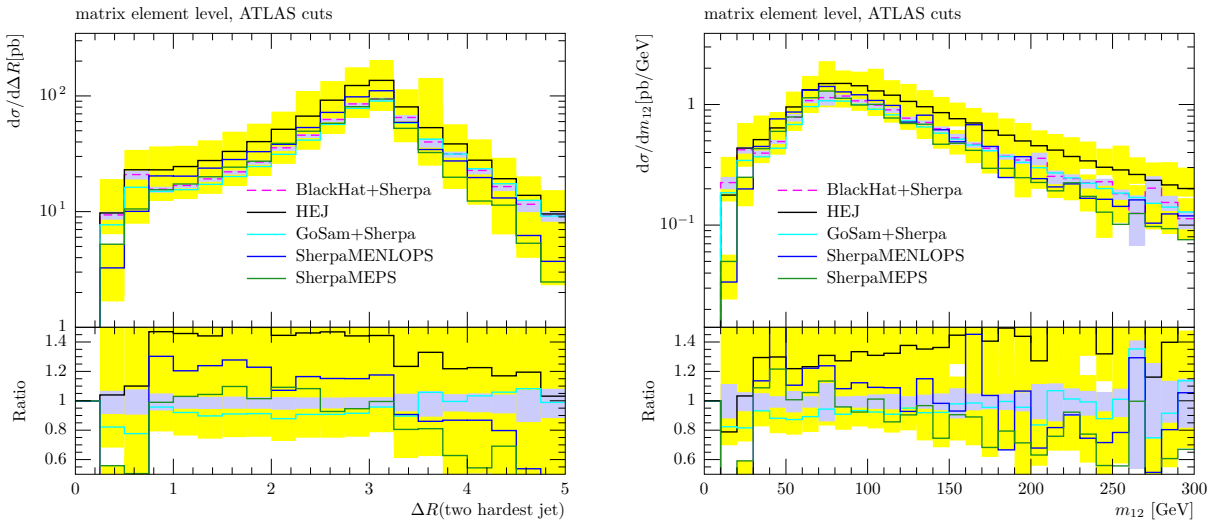


Fig. 94: ΔR of two leading jets (left) and invariant mass of two hardest jets (right) matrix element level (for exact definitions and cuts see App. 18.61 and App. 18.62). Note that BLACKHAT uses the CTEQ6.6 pdf, PYTHIA8 and MADGRAPH+PYTHIA CTEQ6L1 and all the others use CT10. In both ratio plots the ratio is taken with respect to BLACKHAT+SHERPA (on matrix element level).

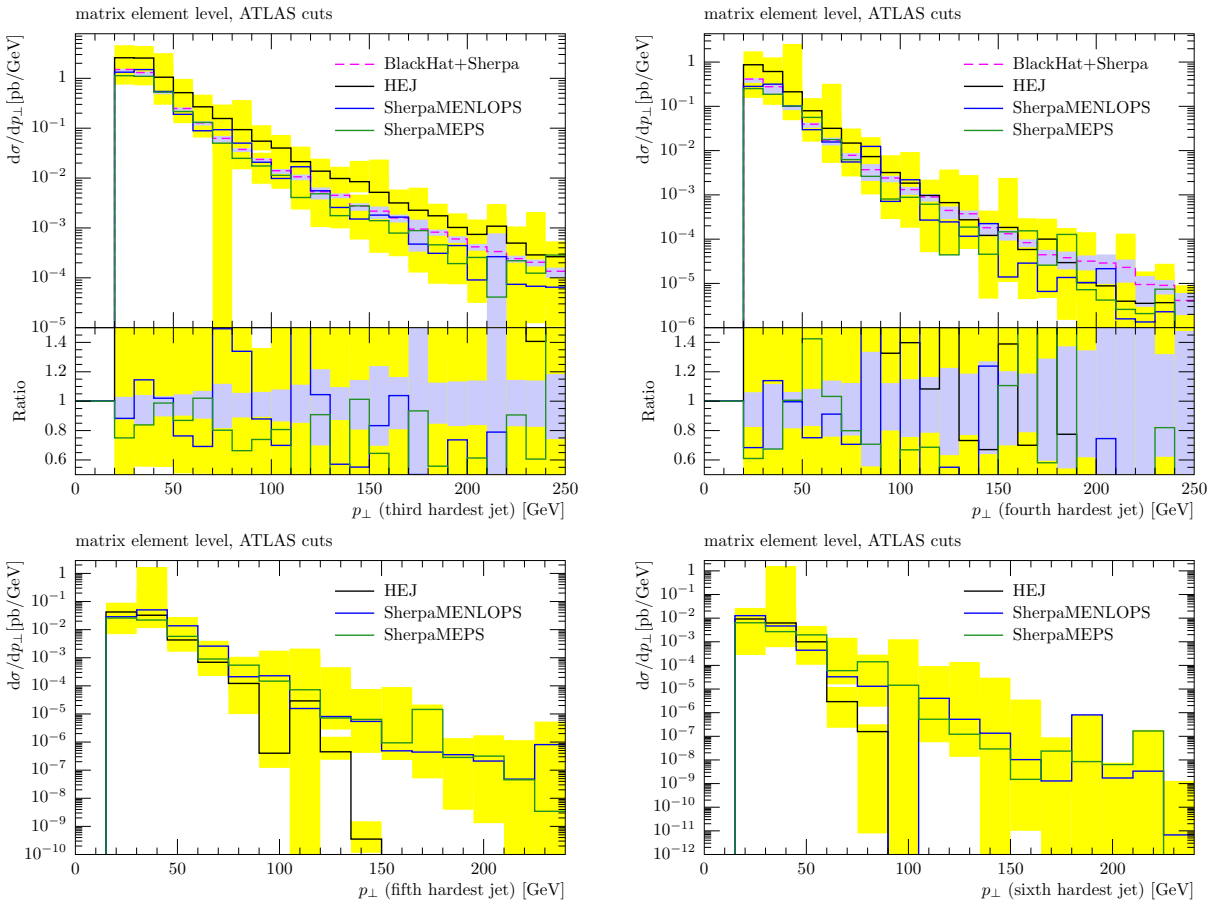


Fig. 95: Transverse momentum of third to sixth hardest jet on matrix element level (for exact definitions and cuts see App. 18.61 and App. 18.62). Note that BLACKHAT uses the CTEQ6.6 pdf, PYTHIA8 and MADGRAPH+PYTHIA CTEQ6L1 and all the others use CT10. In both ratio plots the ratio is taken with respect to BLACKHAT+SHERPA (on matrix element level).

18.5 Conclusions

In this study first steps towards an update and extension of the comparison in [439] have been made. In contrast to the older study, a larger variety of tools including fixed-order and resummation tools as well as NLO matched and tree-level merged simulations have been included. Not surprisingly, some observables appear to be described fairly consistently between different tools, while others exhibit large deviations, sometimes clearly beyond the formal accuracy claimed by the different methods, and also beyond the best estimates of intrinsic modelling or calculational uncertainties provided by the authors. In some instances the relative differences are way beyond naïve expectations by most of the authors of this study. This clearly hints at the need to carefully cross-validate different tools before deploying them for large scale simulations, and it also necessitates an increased collaboration of the authors of such tools in order to arrive at a more consistent picture.

We hope that this study triggered some future work towards the latter goal.

Acknowledgements

We would like to thank the organisers of Les Houches 2012 for repeatedly hosting such a fantastic and fruitful workshop. We would also like to thank all participants for a lively and stimulating atmosphere and lots of inspiring discussions.

MS's and 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). SA and ER also acknowledge funding from LHCPhenoNet for travel expenses. LL's and SP's work was supported by the Swedish research council (contract 621-2009-4076). JMS and GL are supported by the UK Science and Technology Facilities Council (STFC). T.R. is supported by the Alexander von Humboldt Foundation, in the framework of the Sofja Kovaleskaja Award Project "Advanced Mathematical Methods for Particle Physics", endowed by the German Federal Ministry of Education and Research. F.T. is supported by Marie-Curie-IEF, project: "SAMURAI-Apps".

18.6 Cuts and observables

18.61 Cuts

	ATLAS	CMS
lepton $p_\perp$	$> 20\,\text{GeV}$	$> 20\,\text{GeV}$
lepton $\lvert\eta\rvert$	< 2.5 $(e,\ \mu)$	< 2.5 $(e), 2.1$ (μ)
$\not{E}_\perp$	$> 25\,\text{GeV}$	no cut
$m_{\perp,W}$	$> 40\,\text{GeV}$	$> 20\,\text{GeV}$
jet $p_\perp$	$> 25\,\text{GeV}$	$> 30\,\text{GeV}$
jet $\lvert\eta\rvert$	< 4.4	< 2.4
jet radius	0.4 (anti-$k_\perp$)	0.5 (anti-$k_\perp$)
lepton isolation	$< 10\,\%$ of lepton energy in cone with R=0.5	$< 10\,\%$ of lepton energy in cone with R=0.5

Table 13: Cuts used in this study inspired by common ATLAS and CMS cuts.

Tab. 13 presents the cuts applied to define the event selection in both the ATLAS and CMS specifications.

18.62 Analysis procedure and definition of observables

A common analysis was implemented within the RIVET framework and used by all codes providing individual events. This analyses is carried out as defined in the following:

1. remove all neutrinos from all final states (i.e. 'all particles' from now on means 'all particles without neutrinos')
2. find hardest isolated lepton (electron or muon) ('lepton' from now on means 'hardest isolated lepton')
3. cut on lepton $p_\perp$ and $|\eta|$
4. compute missing transverse energy $\not{E}_\perp$:
   (a) sum the three-momenta of all particles within $|\eta| < 10$, this yields $-\not{p}$
   (b) compute missing energy as $\not{E} = |\not{p}|$
   (c) assume resulting four-vector $\not{p}$ corresponds to neutrino
5. for ATLAS cut on $\not{E}_\perp$
6. resonstruct W four-momentum as $p^W = p^{\text{lepton}} + \not{p}$
7. compute W transverse mass squared as $m^2_{\perp,W} = (p^{\text{lepton}}_\perp + \not{p}_\perp)^2 - (p^W_\perp)^2$
8. cut on W transverse mass
9. remove lepton from final state
10. cluster into jets keeping only those passing the $p_\perp$ and $|\eta|$ cuts
11. compute $H_T = \sum_{i\in\{\text{jets}\}} E_{\perp\, i}$
12. compute beam thrust $\tau_B = \sum_{i\in\{\text{particles}\}} (E_i - |p^z_i|)$ using all visible particles

It should be noted that this defintion of the W is infra-red safe only for transverse observables.

18.7 Detailed settings

18.71 SHERPA

For this study SHERPA-1.3.1 was used. Except for the underlying event, which was tuned for the CT10 [255] parton distribution functions and whose parameters are given below, all other non-perturbative parameters were kept at their default values. The underlying model was tuned for the cluster hadronisation.

K_PERP_MEAN_1	1.17
K_PERP_MEAN_2	1.17
K_PERP_SIGMA_1	0.760
K_PERP_SIGMA_2	0.760
PROFILE_PARAMETERS	0.576, 0.353
RESCALE_EXPONENT	0.238
SCALE_MIN	2.52
SIGMA_ND_FACTOR	0.465

18.72 PYTHIA8

To produce the results, we have used two tunes of PYTHIA8, Tune 4C and Tune A2, both of which use CTEQ6L1 parton distributions. Tune 4C is the default tune in PYTHIA8– no additional input settings are necessary. For completeness, below we list all parameters that are implicitly set by choosing the default Tune 4C.

```
PDF:pSet = 8
SigmaProcess:alphaSvalue = 0.135
SigmaDiffractive:dampen = on
SigmaDiffractive:maxXB = 65.0
SigmaDiffractive:maxAX = 65.0
SigmaDiffractive:maxXX = 65.0
```

```
TimeShower:dampenBeamRecoil = on
TimeShower:phiPolAsym = on
SpaceShower:alphaSvalue = 0.137
SpaceShower:samePTasMPI = false
SpaceShower:pT0Ref = 2.0
SpaceShower:ecmRef = 1800.0
SpaceShower:ecmPow = 0.0
SpaceShower:rapidityOrder = on
SpaceShower:phiPolAsym = on
SpaceShower:phiIntAsym = on
MultipartonInteractions:alphaSvalue = 0.135
MultipartonInteractions:pT0Ref = 2.085
MultipartonInteractions:ecmRef = 1800.
MultipartonInteractions:ecmPow =  0.19
MultipartonInteractions:bProfile = 3
MultipartonInteractions:expPow = 2.0
BeamRemnants:primordialKTsoft = 0.5
BeamRemnants:primordialKThard = 2.0
BeamRemnants:halfScaleForKT = 1.0
BeamRemnants:halfMassForKT =  1.0
BeamRemnants:reconnectRange = 1.5
```

A detailed discussion of these choices can be found in [455]. All other parameters remain with their default values. For our purposes, it might be interesting to remark that the starting value for α_s-evolution in time-like splittings is given by

```
SpaceShower:alphaSvalue = 0.1383
```

To investigate the impact of rapidity ordering in space-like showers, we chose to remove enforced rapidity ordering by setting

```
SpaceShower:rapidityOrder = off
```

If rapidity ordering is enforced in ISR, the question arises how it should be treated when picking histories. For this purpose, PYTHIA8 supplies the switch

```
Merging:enforceStrongOrdering
```

When switched "on", this parameter will result in picking non-rapidity-ordered histories only if no rapidity-ordered paths where found, thus disfavouring non-rapidity-ordered parton shower histories for matrix element states. To have a more complete understanding of the impact of tuning, we also changed to the recently proposed Tune A2 [458]. For this, we have to set

```
Tune:pp = 7
```

PYTHIA8 will then reset the following parameters:

```
PDF:pSet = 8
SigmaProcess:alphaSvalue = 0.135
SigmaDiffractive:dampen = on
SigmaDiffractive:maxXB = 65.0
SigmaDiffractive:maxAX = 65.0
```

```
SigmaDiffractive:maxXX = 65.0
TimeShower:dampenBeamRecoil = on
TimeShower:phiPolAsym = on
SpaceShower:alphaSvalue = 0.137
SpaceShower:samePTasMPI = false
SpaceShower:pT0Ref = 2.0
SpaceShower:ecmRef = 1800.0
SpaceShower:ecmPow = 0.0
SpaceShower:rapidityOrder = false
SpaceShower:phiPolAsym = on
SpaceShower:phiIntAsym = on
MultipartonInteractions:alphaSvalue = 0.135
MultipartonInteractions:pT0Ref = 2.18
MultipartonInteractions:ecmRef = 1800.
MultipartonInteractions:ecmPow = 0.22
MultipartonInteractions:bProfile = 4
MultipartonInteractions:a1 = 0.06
BeamRemnants:primordialKTsoft = 0.5
BeamRemnants:primordialKThard = 2.0
BeamRemnants:halfScaleForKT = 1.0
BeamRemnants:halfMassForKT = 1.0
BeamRemnants:reconnectRange = 1.55
```

Apart from not enforcing rapidity ordering in space-like splittings, this tune differs from Tune 4C in that the proton size is considered $x-$dependent. This is in the spirit of Tune 4CX, which was introduced in [470]. In general, since we include matrix element states for two and three jets, we do not apply additional matrix element corrections in PYTHIA8 after the first emission, by setting

```
SpaceShower:MEafterFirst = off
TimeShower:MEafterFirst = off
```

18.73 POWHEG BOX + PYTHIA8

For this study we used POWHEG BOX rev1282 and PYTHIA 8.153. Except for the specific subprocess requested, the parton distribution functions set and the renormalisation/factorisation scale factors chosen, all the other parameters were kept fixed below during all the runs. Here is a sample POWHEG BOX input file:

```
! Wˆ+ + jet production parameter
idvecbos 24         ! PDG id of vector boson (24: W+, -24: W-)
vdecaymode 1        ! decay channel (1: electron, 2: muon, 3: tau)
numevts 4000000     ! number of events to be generated
ih1 1               ! hadron 1 (1 for protons, -1 for antiprotons)
ih2 1               ! hadron 2 (1 for protons, -1 for antiprotons)
ebeam1 3500d0       ! energy of beam 1 in GeV
ebeam2 3500d0       ! energy of beam 2 in GeV
lhans1  192800      ! pdf set for hadron 1 (LHA numbering)
lhans2  192800      ! pdf set for hadron 2 (LHA numbering)
ncall1 100000       ! number of calls for initializing the ...
itmx1    5          ! number of iterations for initializing the ...
ncall2 250000       ! number of calls for computing the integral ...
```

```
itmx2    4          ! number of iterations for computing the ...
foldcsi   1         ! number of folds on csi integration
foldy     1         ! number of folds on  y  integration
foldphi   1         ! number of folds on phi integration
nubound 100000      ! number of bbarra calls to setup norm of ...
icsimax  1          ! <= 100, number of csi subdivision when ...
iymax    1          ! <= 100, number of y subdivision when ...
xupbound 2d0        ! increase upper bound for radiation generation
renscfact  1d0      ! (default 1d0) ren scale factor: muren  = ...
facscfact  1d0      ! (default 1d0) fac scale factor: mufact = ...
withdamp    1       ! (default 0, do not use) use Born-zero ...
iseed    12345679   ! initialize random number sequence
bornktmin  5d0      ! (default 0d0) kt min at Born level for ...
bornsuppfact 100d0  !  (default 0d0) mass param for Born ...
withnegweights 1    ! (default 0) allows negative weighted ...
runningscale 1      ! (default 0) ren. and fact. scales set to ...
masswindow_low 1000 ! restricts phase space to ...
masswindow_high 1000! restricts phase space to ...
```

When interfacing to PYTHIA8 we have changed the following settings with respect to PYTHIA8 defaults, for the various stages under investigations:

```
///Hadron Level w MPI and QED
BeamRemnants:reconnectRange = 1.50000
MultipleInteractions:alphaSvalue = 0.13500
MultipleInteractions:bProfile = 3
MultipleInteractions:ecmPow = 0.1900
MultipleInteractions:expPow = 2.0000
MultipleInteractions:pT0Ref = 2.0850
PDF:pSet = 8
SigmaDiffractive:dampen = on
SigmaDiffractive:maxAX = 65.0000
SigmaDiffractive:maxXB = 65.0000
SigmaDiffractive:maxXX = 65.0000
SigmaProcess:alphaSvalue = 0.13500
SpaceShower:MEafterFirst = off
SpaceShower:MEcorrections = off
SpaceShower:pTmaxMatch = 0
SpaceShower:rapidityOrder = on
TimeShower:MEcorrections = off
TimeShower:MEafterFirst = off
TimeShower:pTmaxMatch  = 0

//Hadron Level w MPI (added)
SpaceShower:QEDshowerByQ = off
SpaceShower:QEDshowerByL = off
TimeShower:QEDshowerByQ = off
TimeShower:QEDshowerByL = off

//Hadron Level w/o MPI (added)
PartonLevel:MI = off
```

```
//Shower Level (added)
HadronLevel:All = off

//Parton Level (added)
PartonLevel:ISR = off
PartonLevel:FSR = off
PartonLevel:Remnants = on
```

and, most important, we have vetoed shower emissions with a transverse momentum greater than the value of `SCALUP` read from the Les Houches event file for the corresponding event.

18.74 MADGRAPH + PYTHIA

For this study MADGRAPH/MADEVENT 5.1.1.0 and PYTHIA 6.4.2.4 is used. The LHE files are generated for events with a W and up to four additional partons, *i.e.* for the process:

```
pp>w- -> l-vl~ ; l-vl~~j ; l-vl~~jj ;l-vl~~jjj ; l-vl~~jjjj ;
l-vl~ ; l-vl~j ; l-vl~jj ;
l-vl~jjj ; l-vl~ jjjj ()
```

The mass of the b quark is set to zero. The strong constant $\alpha_s(M_Z^2)$ is set to 0.1300 both in the matrix element calculation and in the proton PDF, that is the CTEQ6L1.

PYTHIA is used for the parton shower and the hadronisation with the following parameters modified according to tune Z2.

```
MSTU(21)=1      ! Check on possible errors during program execution
MSTJ(22)=2      ! Decay those unstable particles
PARJ(71)=10 .   ! for which ctau  10 mm
MSTP(33)=0      ! no K factors in hard cross sections
MSTP(2)=1       ! which order running alphaS
MSTP(51)=10042  ! structure function chosen (external PDF CTEQ6L1)
MSTP(52)=2      ! work with LHAPDF
PARP(82)=1.832  ! pt cutoff for multiparton interactions
PARP(89)=1800.  ! sqrts for which PARP82 is set
PARP(90)=0.275  ! Multiple interactions: rescaling power
MSTP(95)=6      ! CR (color reconnection parameters)
PARP(77)=1.016  ! CR
PARP(78)=0.538  ! CR
PARP(80)=0.1    ! Prob. colored parton from BBR
PARP(83)=0.356  ! Multiple interactions: matter distribution para...
PARP(84)=0.651  ! Multiple interactions: matter distribution para...
PARP(62)=1.025  ! ISR cutoff
MSTP(91)=1      ! Gaussian primordial kT
PARP(93)=10.0   ! primordial kT-max
MSTP(81)=21     ! multiple parton interactions 1 is Pythia default
MSTP(82)=4      ! Defines the multi-parton model
PMAS(5,1)=4.8   ! b quark mass
PMAS(6,1)=172.5 ! t quark mass
MSTJ(1)=1       ! Fragmentation/hadronization on
MSTP(61)=1      ! Parton showering on
```

For additional studies, we set

```
MSTJ(41)=3      ! switch off lepton FSR
MSTP(81)=20     ! switch off MPI
MSTJ(1)=0       ! Fragmentation/hadronization off
```

to switch off, respectively, final state QED radiation, multi-particle interactions, and hadronisation.

Part V
EXPERIMENTAL DEFINITIONS AND CORRECTIONS

19. PHOTON ISOLATION AND FRAGMENTATION CONTRIBUTION [43]

Abstract

Photon isolation and its link with the fragmentation contribution is explored via NLO matrix-element generator and parton-shower Monte-Carlo.

Firstly the dependence of the inclusive photon and di-photon NLO cross sections to the choice of isolation criteria are investigated. The isolation criteria used is the discretized version of the Frixione isolation, with parameters chosen for those most practical at an experimental level. As an extention, a more generalized version of the standard Frixione isolation is also studied. The selection of scale is also investigated in search of the 'saddle point', which would give the optimal scale choice. In addition the choice of jet algorithm is investigated for the photon with associated jet cross section.

Secondly, properties of the fragmentation contribution in parton-shower Monte-Carlos are investigated. The distance profile of the photon to the other generator level particles in the event is explored in the case of neutral mesons, fragmentation photons and direct photons. Next the impact of a "hollow" or "crown" isolation criterion, expected to enhance the fragmentation contribution, is explored. Then, to complement the NLO inclusive studies, the impact of typical Frixione isolation criteria on the fragmentation component are investigated in the parton-shower Monte-Carlos.

Finally conclusions are made comparing the properties of the fragmentation contribution in NLO generators and parton-shower generators.

19.1 INTRODUCTION

Experimental measurements of single photons and di-photons require the application of isolation cuts to reduce the copious backgrounds arising from jet fragmentation. Such cuts also have the impact of reducing the fragmentation contributions of photon production. On the theoretical side, including fragmentation contributions of photon production can greatly increase the complexity of the calculations, while the application of appropriate isolation cuts can effectively remove those fragmentation contributions.

[43] Contributed by:N. Chanon, S. Gascon-Shotkin, J. P. Guillet, J. Huston, E. Pilon, M. Schwoerer, M. Stockton, M. Tripiana

19.11 Frixione isolation

In the following we will study the Frixione isolation criterion [471], which was designed to suppress the fragmentation contribution. It has been shown to reduce the fragmentation contribution in NLO generators [18]. The question is to know whether or not the behavior is still applicable using parton shower Monte-Carlo and if it can be used experimentally.

We consider the following function for the isolation criterion :

$$E_T^{iso}(R) < f(R) = \epsilon \cdot p_{T,\gamma} \cdot \left(\frac{1 - cos(R)}{1 - cos(R_0)} \right)^n \tag{19.1.1}$$

where $E_T^{iso}(R)$ is the isolation sum of all particles inside a cone of $R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ around the photon, ϵ is the strength or the scale of the isolation criterion, $p_{T,\gamma}$ is the transverse energy of the photon, R_0 is the first considered cone and n the power of the isolation criterion. This formula can be altered by replacing $p_{T,\gamma}$ with a fixed threshold. Functional forms $f(R)$ for different ϵ and n are shown Fig. 96.

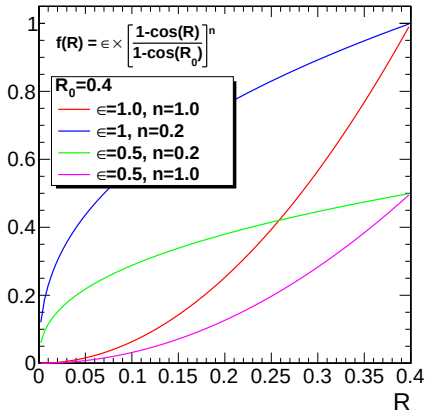


Fig. 96: Examples of Frixione functional form for different parameters.

The Frixione isolation is tighter and tighter when decreasing the R cone size. With matrix element NLO generators (as in Jetphox [472, 473] and Diphox [474]), the only contribution possible at a given $R \neq 0$ is the one coming from the suplementary hard jets in the event, while the fragmentation debris are emitted colinearly to the photon at $R = 0$ (angle information is lost due to the fragmentation function which is integrated over the angle). With parton-shower generator, fragmentation photons are emitted off quarks at a non-zero angle during the showering process. In the following sections we will study the link between isolation and fragmentation in NLO generators, then with parton-shower Monte-Carlo.

19.12 Experimental complications

There are various mismatches between isolation cuts applied to theoretical calculations and isolation cuts applied to data (or to Monte Carlo). First of all, we wish to apply the isolation cut only to energy related to the hard scatter. Experimentally, most of the energy inside an isolation cone is due to the underlying event associated with the hard scatter, or the remnants of additional interactions in the same crossing. Techniques such as jet area subtraction [475, 476] can be used to remove an amount of energy from the isolation cone roughly equal to the expected contamination from underlying event/pileup, leaving only energy related to the hard scatter and specifically to fragmentation processes. Since there is no underlying event/pileup in partonic level theory calculations, only the isolation cut needs to be applied to the theory.

Full details of the use of this correction within the ATLAS collaboration can be found in the inclusive cross section measurement [477]. The ATLAS isolation definition uses a cone around the cluster of cells that are identified as a photon. These photon cluster cells are not included in the sum, so there is first a correction for any leakage of the photon shower into the surrounding cells (typically a few percent of the photon p_T). The pile-up/underlying event correction is then applied by calculating per event the ambient energy from the jet activity in that specific event. This follows the jet area corrections method mentioned above, where all jets are reconstructed without any minimum momentum threshold. The energy density of each jet is calculated and the median density is used for the correction. In 2010 this typically resulted in a correction of around $900\mathrm{MeV}$.

The original Frixione isolation scheme assumed that an isolation cut could be applied continuously as a function of R (distance from the photon). Actual detectors have a finite granularity. A solution to this was the adoption of a discretized version of Frixione isolation, allowing this granularity to be taken into account [18]. However, it is not possible to place an isolation cut on the inner-most cone (typically $R \sim 0.1$), because of the presence of the photon itself. While the separation between the fragmentation photon and the jet remnants is finite in data (and in Monte Carlo) , fragmentation is treated as a collinear process in partonic cross sections. The inability to apply the isolation cut down to $R = 0$, results in a greatly reduced ability to discriminate against fragmentation processes in the partonic level theory.

To rectify this the Frixione calculation could be modified into a 'crown' isolation, whereby the last cone is missed from the calculation. Unfortunately as most of the radiation is collinear in the fragmentation events, it is likely to reduce its effect of removing these events. Other studies [478] have shown that the photon quality cuts applied by the experiments will reduce the fraction of fragmentation photons accepted, where substantial fragmentation energy is collinear with the photon. However, the rejection is not 100%, so we are still left with a smaller reduction of fragmentation contributions in the partonic level theory than are actually (presumably) present in the data. In these proceedings, we will discuss how to more properly incorporate the correct level of rejection in the theory.

19.13 Choice of fragmentation scale

In addition to the experimental difficulties with applying the isolation criteria there are also difficulties in choosing appropriate scales for the theoretical calculation. This is discussed in the following text, along with other considerations for applying Frixione isolation at a theoretical level.

Fragmentation is treated as a collinear process in partonic calculations. In this framework, the original "continuous" Frixione criterion [471] was designed to inhibit the appearance of final state photon-parton collinear singularities which otherwise require absorption in a fragmentation function $D(z, M_F)$. Thus, cross sections for the production of prompt photons isolated with this criterion involve no fragmentation contribution. Discretized variants of this criterion have been proposed which aim at matching better what can be actually implemented experimentally [18]. They consist in a limited number of nested cones $\mathcal{C}_{j=1,\cdots,n}$ with respective radii $R_1 = R_{min} < R_2 < \cdots < R_n = R_{max}$ defined in the azimuthal and rapidity differences with respect to the photon direction, and requiring recursively that the accompanying hadronic transverse energies inside every successive cone $\mathcal{C}_j$ be less than an ordered sequence of maximum values[44] $E_{T\,j}^{iso}$ such that $0 < E_{T\,1}^{iso} < \cdots < E_{T\,n}^{iso}$. However, in contrast with the continuous criterion, such discretized variants still involve a fragmentation contribution, though the latter is expected to be small, since the situation in the innermost cone shares some similarity with the standard cone criterion. When quantifying the magnitude of the fragmentation contribution with such discretized criteria, a potentially tricky issue concerns the fragmentation scale dependence and the "best choice" of scale.

This issue matters for isolation with the standard cone criterion when the radius R of the cone is $\ll 1$ while E_T^{iso} is kept fixed. Whereas the natural fragmentation scale M_F in the non-isolated case

[44] Whether the $E_{T\,j}^{iso}$ take fixed values or are functions of the photon's transverse momentum does not matter here.

is $\sim p_T^\gamma$, this choice can lead to very poor theoretical estimates at Next to Leading Order (NLO) in perturbative QCD when $R \ll 1$ [472]. The scale dependence near the choice $M_F \sim p_T^\gamma$ is then large and, worse, the theoretical prediction may eventually exhibit an unphysical violation of unitarity whereby the predicted NLO cross section for photons becomes *larger* than the inclusive one, so that even for only moderately small R the reliability of the prediction is questionable. On the other hand, as[45] $D(z, M_F) \sim \log(M_F/\Lambda_{QCD})$, with the choice $M_F \sim Rp_T^\gamma$ the fragmentation contribution is suppressed compared with $M_F \sim p_T^\gamma$. The situation is improved regarding both scale dependence and unitarity, although it does not solve the problem completely. One actually faces a multiscale problem: $\Lambda_{QCD} \ll Rp_T^\gamma \ll p_T^\gamma$, and a one-scale compromise is possibly insufficient depending on the kinematical regime explored. The atypical choice $M_F \sim Rp_T^\gamma$ has in principle to be supplemented by a resummation of the logarithmic R dependence coming form outside the cone, if at all possible. At leading-$\log R$ (LLR) accuracy at least, such a resummation is actually feasible, which furthermore allows to solve the apparent puzzle why scale choices should be very different in the cases with isolation in a narrow cone vs. broad cone or without isolation.

The concern about the discretized Frixione criteria is that the innermost cone size is quite small. The choice for the fragmentation scale M_F shall then arguably be $M_F \sim \mathcal{O}(R_{min}p_T^\gamma)$. On the other hand, as the allowed transverse energy deposit $E_T^{iso}(R_{min})$ inside this cone is correspondingly small, the width of the interval in the fragmentation variable on which the fragmentation function is convoluted with the partonic cross section is restricted to a rather narrow range $0 < 1 - z < E_T^{iso}(R_{min})/p_T^\gamma \sim \epsilon(R_{min}/R_{max})^n$. This leads to a quite suppressed fragmentation contribution. The combination of the two effects: a low fragmentation scale and a narrow z-range, is the discrete counterpart of the inhibition of fragmentation by the continuous criterion. We may thus expect that the issue of the narrow cone is less worrying for the reliability of the NLO calculation in this case than if only R were taken small while keeping E_T^{iso} fixed. In order to assess the uncertainty on the fragmentation contribution we may perform the calculation for the "arguably better" scale $M_F \sim R_{min}\, p_T^\gamma$ and compare it to the expectedly larger result for the standard choice $M_F \sim p_T^\gamma/2$.

19.2 ISOLATION FOR INCLUSIVE PHOTONS AND DIPHOTONS AT NLO

The study at NLO uses the Jetphox generator to calculate the inclusive photon cross section and Diphox for the di-photon cross section. Details of how to use the software and to obtain predictions with errors can be found in [480] and the selection criteria used are listed in the appendix. Previous results from Les Houches [18] showed that the discretized Frixione isolation criteria did manage to reduce the fragmentation contribution, here we extend that study in several ways. Firstly the cross section returned has been compared to that calculated from using the standard cone isolation, as used in current measurements. A generalized form of the Frixione isolation is discussed, aimed to satisfy both the experimental and the theoretical requirements on the isolation cut for different p_T regimes. In addition the effects of changing the number of cones used in the calculation and of choosing an E_T cut, rather than relating it to the photon p_T, are investigated. In addition, further complications to comparing theoretical and experimental isolation calculations are discussed. Finally there are further brief studies using Jetphox to look at scale and jet algorithm choices.

19.21 Discretized prescription

The parameters used to define different selections, according to Eq. 19.1.1, were:

a: ϵ=0.05 n=0.2	b: ϵ=1 n=0.2	c: ϵ=1 n=1	d: ϵ=0.5 n=1
e: ϵ=0.05 n=1	f: ϵ=1 n=0.1	g: ϵ=1 n=0.5	

[45] The logarithmic behaviour holds in when $M_F \gg \Lambda_{QCD}$. In the non-perturbative regime instead no logarithm is expected to develop and one expects rather a power-suppressed behaviour; see ref. [479] for more details.

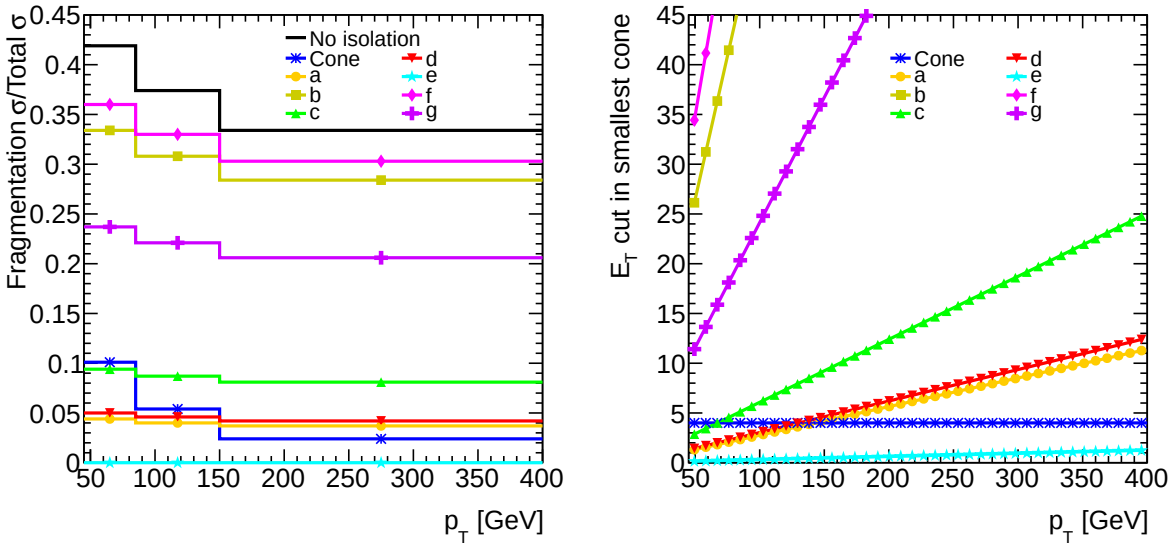


Fig. 97: Left: Fragmentation fraction for the cone and Frixione isolation criteria. Right: The applied E_T cut on the isolation sum as a function of p_T for the 0.4 cone or 0.1 cone in the Frixione criteria.

where all but the last two were based on the previous study. In all cases R_0 was chosen to be 0.4, with R being set to either: 0.4, 0.3, 0.2, 0.1 or 0.4, 0.35, 0.3, 0.25, 0.2, 0.15, 0.1.

The comparison to the cone isolation in Fig. 97 shows that out of the chosen parameters only 1 set removes the fragmentation contribution more than what is removed by the cone algorithm, although two are lower until high p_T. It also shows that criteria b and f are not much better than applying no isolation criteria at all. When altering Eq. 19.1.1 to use a fixed $E_T = 4\text{GeV}$ instead of $p_{T,\gamma}$ the results are more promising but this is because it applies a cut in the 0.1 cone that is below the experimental accuracy (of the order 100 MeV due to detector resolution/noise). Unfortunately Fig. 97 also shows that this is also the case for the p_T requirements, as case e (the only criteria to perform better than the standard cone) also applies a cut that is not viable experimentally in the 0.1 cone.

There are some positive outcomes from these studies, firstly the Frixione criteria b and d maybe useful criteria to use experimentally as they keep the fragmentation contribution similar and low in all bins, which could help with understanding of the systematic errors/correlation between bins. Secondly the comparison of the number of cones used in the Frixione criteria resulted in a difference of around 1% on the total cross section and almost no effect on the fragmentation fraction. This means that it is fine to use the lower number of cones case, and that the discrete Frixione criteria is most likely very similar to that of the continuous version.

19.22 Generalized prescription

As seen previously, to remove the fragmentation contribution in the theory, a small value of ε is needed. However, given the effects of finite resolution and granularity on the experimental description of the isolation energy, a minimum threshold has to be allowed in the isolation cone, especially at low p_T. A typical value of 2-4 GeV is used as experimental cut, to optimize the rejection of hadronic background coming from the decay of light mesons. Now, at high p_T this cut might result too tight, particularly on the theoretical side given that an isolation cut much smaller than the photon p_T can cause large logs in the calculations, this effect was not observed in the previous Les Houches study.

As a good compromise of these two requirements, it has been proposed [481] to extend the original

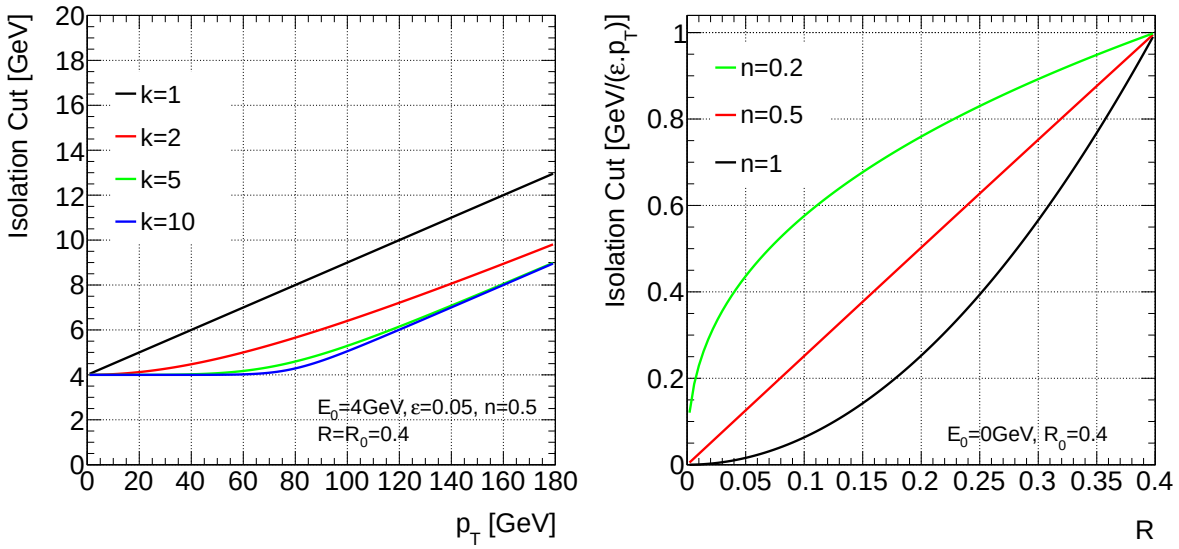


Fig. 98: Effect of the two terms in the modified Frixione isolation prescription effects of the different pieces of the generalized Frixione isolation prescription on the isolation cut as a function of p_T (left) and the cone radius R (right).

Frixione prescription (Eq. 19.1.1) to a more general form:

$$E_T^{iso} < \Big((E_0)^k + (\varepsilon.p_T)^k\Big)^{1/k} \left(\frac{1-\cos R}{1-\cos R_0}\right)^n \tag{19.2.1}$$

where:

- R_0 is the maximum cone size
- E_0 is the minimum energy pedestal allowed in a cone of size R_0
- ε is the fraction of the photon p_T allowed in the cone of size R_0
- k determines the shape of the isolation profile in p_T
- n determines the shape of the isolation profile in R (see Fig. 98[right])

The extra parameters give enough flexibility to ensure a (finite) tight cut at low p_T ($\sim E_0$) and, at the same time, a loose cut at high end of the spectrum driven by the photon p_T. The k parameter controls how quickly/smoothly is the transition from one regime to the other (see Fig. 98[left]).

This generalized prescription[46] has been implemented in Jetphox recently and some possible configurations are explored here. The studied configurations vary ε (=0.05,1) and k (=2,5,10), and have a fixed value for $E_0 = 4\text{GeV}$ (the typical cut applied in ATLAS) and $n = 0.5$ (given the linear behaviour of isolation distribution width observed for direct photons in ATLAS [482]).

The high-ε configurations (ε=1), show a worse performance at removing the fragmentation contribution with respect to the fixed cone approach and are practically insensitive to the value of k in the formula. The remaining fragmentation fraction is $\sim 25\%$ at 45 GeV decreasing to 20% in the highest p_T bin. On the other hand, as seen in Fig. 99, all the configurations for a low value of ε (=0.05) show an improvement in fragmentation rejection compared to both the no isolation and fixed cone cases, in the whole p_T region ($45\text{GeV} < \text{p}_\text{T} < 600\text{GeV}$). The p_T profile for the smaller cone ($R = 0.1$) in this case (Fig. 99[right]) looks also more promising in terms of its applicability at the experimental level.

[46] Indeed for $E_0 = 0$ the original Frixione prescription is restored.

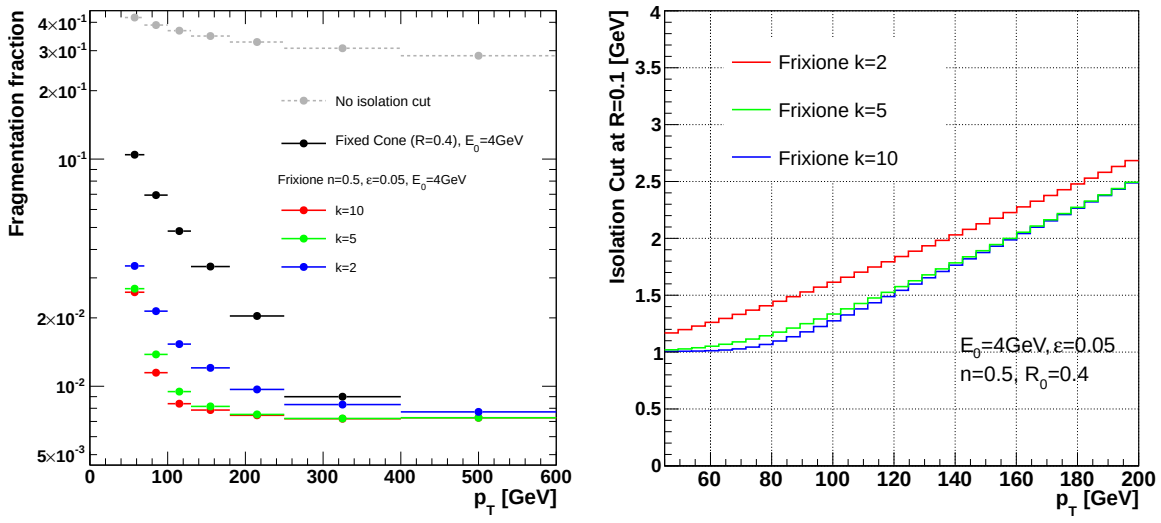


Fig. 99: Left: Fragmentation fraction for the cone and the (generalized) Frixione isolation criteria. Right: The applied isolation cut as a function of p_T for 0.1 cone in the (generalized) Frixione criteria.

19.23 Continuous and discretized Frixione criteria in di-photon events

Following the previous studies with inclusive photons, we now consider the production of photon pairs. The aims of this study are i) to assess the effect on the magnitude of the fragmentation contribution by comparing results from using the continuous Frixione criterion with those using several variants of the discretized version, implemented in the NLO programme Diphox, thereby providing a NLO assessment of how much fragmentation may be missing in the NNLO calculation of Catani et al. [78] which includes no fragmentation and therefore uses the continuous criterion; ii) to probe the dependence of the prediction with respect to the fragmentation scale choice. It supplements a similar comparison which had been performed for inclusive photon production in [18].

Fig. 100 provides a comparison of the original continuous criterion to the discretized version of the criterion based on four nested cones with respective radii $R_{min} = 0.1, R_2 = 0.2, R_3 = 0.3$ and $R_{max} = 0.4$. Four variants of the energy profile $E_T^{iso}(R)$ as defined in Eq. 19.1.1 have been considered: $(\epsilon, n) = (0.05, 0.2), (0.05, 1), (0.5, 1)$ and $(1, 1)$. Fig. 100 (left) presents the distribution in invariant mass of photon pairs in the range 40 GeV $\leq m_{\gamma\gamma} \leq$ 300 GeV. The discretized criterion $(0.05, 1)$ suppresses fragmentation so much that there is practically no difference between the discretized and continuous versions. With the criterion $(1, 1)$, the discretized version leads to a distribution $O(10-12\%)$ larger than the continuous one. The choice $(0.5, 1)$ displays a similar feature, though quantitatively less important. The energy profile of the fourth choice is not suited for an efficient isolation unless ϵ is chosen very small. A similar comparison is shown on Fig. 100 (right) for the distribution in the difference in azimuthal angle $\Delta\phi$ between the two photons. Whereas the distribution in invariant mass is dominated by the direct contribution, the tail of the distribution in $\Delta\phi$ tail at low $\Delta\phi$ is more sensitive to the fragmentation contribution. Therefore, the conclusions are qualitatively similar to the one drawn for the distribution in invariant mass, yet the effects are quantitatively larger.

Fig. 101 assesses the dependence on the fragmentation scale M_F, for the distributions in invariant mass (left) and in $\Delta\phi$ (right) respectively. Two choices were considered: $M_F = R_{min} \min\{p_T^{\gamma 1}, p_T^{\gamma 2}\} = 0.1 \min\{p_T^{\gamma 1}, p_T^{\gamma 2}\}$ vs. $M_F = R_{max} \min\{p_T^{\gamma 1}, p_T^{\gamma 2}\} = 0.4 \min\{p_T^{\gamma 1}, p_T^{\gamma 2}\}$ closer to a standard choice[47]. As expected, the distribution in invariant mass, which is not very sensitive to the fragmen-

[47] Fragmentation scale options depending only on the p_T of the photon from fragmentation are not available because Diphox

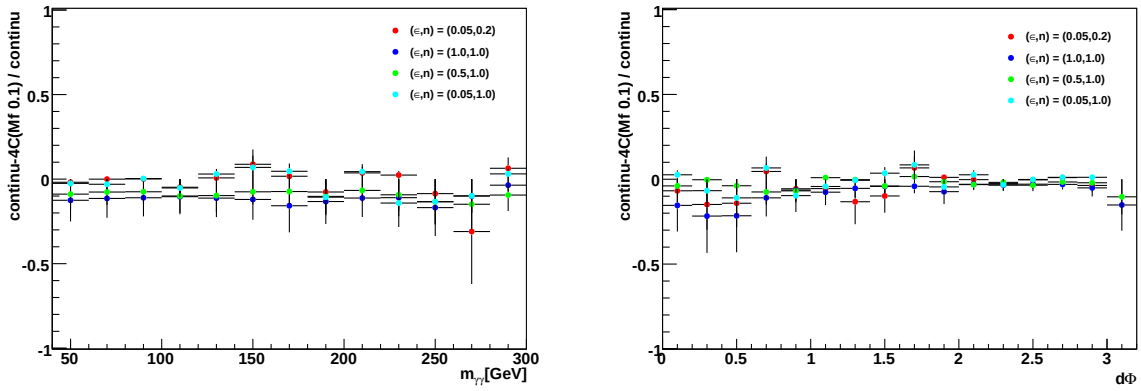


Fig. 100: Comparing continuous and discretized Frixione criterions for the distributions in $m_{\gamma\gamma}$ (left) and $\Delta\phi$ (right) of photon pairs, for four variants of the criterion.

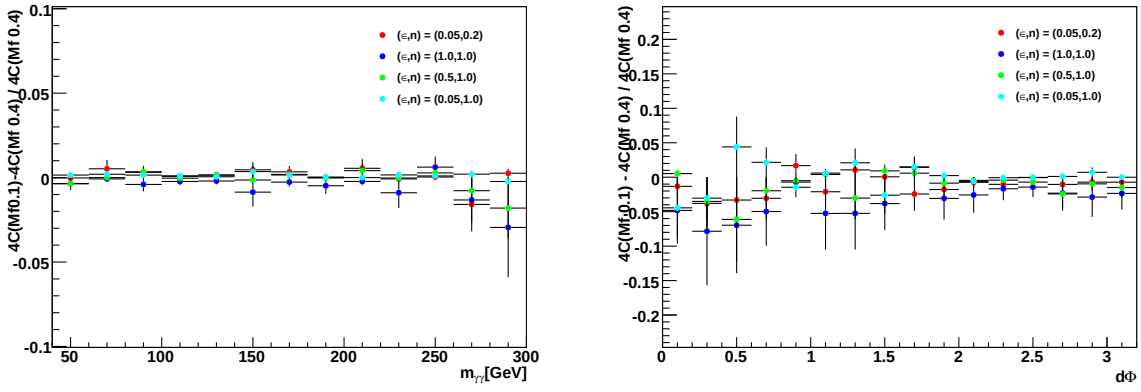


Fig. 101: Fragmentation scale dependence of the distribution in $m_{\gamma\gamma}$ (left) and $\Delta\phi$ (right) of photon pairs, for four variants of the discretized Frixione criterion.

tation contribution, is practically not impacted by the choice. The distribution in $\Delta\phi$ is more sensitive to the fragmentation component and the sensitivity to the fragmentation scale choice is larger than for the distribution in invariant mass. The sensitivity to the fragmentation scale choice is the largest in the case of the criterion $(\epsilon, n) = (1, 1)$, for which the predictions are 5-7% smaller, rather uniformly, with the lower scale choice than with the more standard one.

In conclusion this preliminary study shows that the impact of the fragmentation contribution passing the discretized criterion seems to be almost negligible on the distribution in invariant mass, and remains small even on the tail of the distribution in azimuthal angle. Notwithstanding, the conclusions shall have limited use depending on how isolation is actually implemented experimentally in the innermost cone. We here stick to a discretized version of the Frixione criterion which respects the original idea of a transverse energy deposit decreasing towards zero with the cone radius. If instead any experimental constraint would allow a more permissive condition in the innermost cone, a dedicated study would be mandatory.

19.24 Additional studies at NLO

In addition to the isolation studies with Jetphox, we present here two brief studies as an attempt to reduce the theoretical errors from the NLO calculation. These study the choice of renormalizaion and factorization scale parameter and secondly the jet algorithm parameters.

As studied in [480], the scale choice is set to a fraction of the photon p_T. By altering this fraction around the central value of 1.0, it is hoped to gain an uncertainty on the terms missed in the NLO calculation. The best selection for this central value would be to be at a 'saddle point', where moving in any direction from this point gives similar changes in the cross section. However, it is found that as the scale is reduced (in steps: 2.0, 1.0, 0.5, 0.25, 0.1, 0.05 and 0.01) the cross section increases, when moving the two scales coherently or independantly. One difference in this result to the previous study was that it was carried out in three p_T bins, but the result remained the same for all (only the highest bin was able to be calculated with a scale of 0.01). Similarly the addition of using Frixione isolation instead of the standard cone isolation also resulted in the same cross section behaviour. The summary of this is that there must be large contributions needed from NNLO. However, on the positive side, in all 3 p_T bins, the variation between 0.5-1.0-2.0 resulted in differences of similar magnitude around 1.0, so this is likely a safe estimate of the uncertainty.

After the inclusive photon measurements, the next step experimentally is to require the addition of at least one jet. Using a jet of $10\mathrm{GeV}$ the cross section was calculated for two algorithms each for multiple sizes:

- Kt algorithm with $\Delta R = 0.3$, 0.4, 0.5 or 0.6
- Cone with $\Delta R = 0.4$, 0.5 or 0.6

These choices had an affect of $< 1\%$ on the cross section computed in 3 photon p_T bins, suggesting that this will not increase the error for the NLO calculation when moving from the inclusive cross section to that with an additional jet.

19.3 FRAGMENTATION PHOTONS IN PARTON-SHOWER MONTE-CARLO

The second part of this study continues to investigate photon isolation, but now in di-photon events using parton-shower Monte-Carlo generators; again the selection used is listed in the appendix. The study begins by investigating the distance between the photon and other particles. It then moves into studying several different styles of isolation criteria, including Frixione criteria as done in the inclusive NLO studies.

encodes the two photons in a symmetrized way.

19.31 Topology of fragmentation photons

We consider three sets of parton-shower Monte-Carlo samples for the $\gamma\gamma$+X process:

- Pythia [400] $\gamma\gamma$ Born and Box direct processes, plus the Pythia γ+jet process with the jet fragmenting into a photon (20 million events were generated for the γ+jet sample and 1000 times more would have been needed for the dijet fragmenting to two photons due to the low $q \to \gamma$ branching ratio for isolated photons).
- Pythia $\gamma\gamma$ Born and Box direct processes, plus the Pythia γ+jet process with the jet fragmenting into a photon and the Pythia dijet process with the two jets fragmenting into photons. Both Pythia γ+jet and dijet samples were generated with a filter which enhances the presence of events with isolated electromagnetic particles.
- Madgraph [165] $\gamma\gamma$ + up to two supplementary hard jets, with fragmentation/hadronization done with Pythia.

The fragmentation contribution is included as a bremsstrahlung contribution in Madgraph at matrix element level, while it is included as a showering contribution in Pythia γ+jet and dijet (in the PYTHIA samples we identify fragmentation photons as those having a quark or gluon[48] as parent). The fragmentation fraction found is compatible with ref [483]. We consider additionally the case where the two jets fragment into boosted neutral mesons (π^0, η, ρ and ω) that can experimentally mimic direct or fragmentation photons at reconstructed level because of the finite granularity of the detector. These samples include an underlying event but were generated without pile-up.

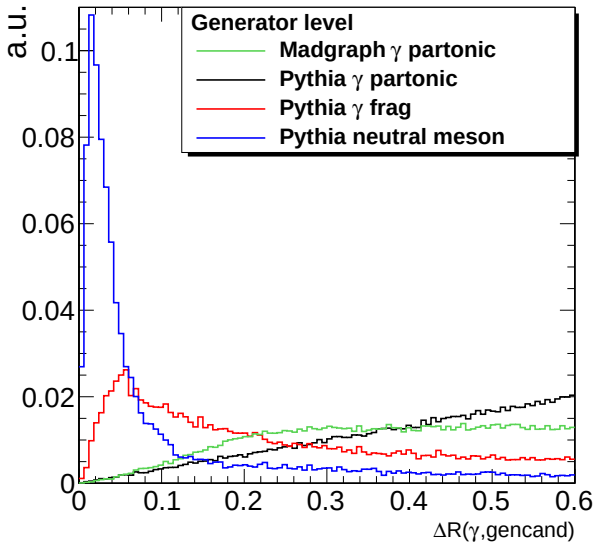


Fig. 102: ΔR distance distribution between the photon and the other generator-level particle candidates in the event, for neutral mesons, fragmentation photons and partonic photons.

Fig. 102 shows the ΔR distance between the photon or neutral mesons and the other particle candidates in the event. Partonic photons, fragmentation photons and neutral mesons have different properties as a function of ΔR. Partonic photons in Pythia have a linear behavior, which is expected because the only contribution that can enter in the isolation sum is the underlying event and pile-up (with also a small contribution from QCD radiation at the shower level) which is expected to be uniform in space. As each bin consists of an annulus with radius growing linearly as a function of R, the quantity

[48] Photon radiation directly by a gluon is of course not physically possible; the representation as such in PYTHIA is a technical shortcut for the actual physical process

of particles grows linearly with R in the area of the annulus. Neutral mesons have a radically different profile, with a peak of the ΔR distribution close to 0. The peak is caused by the decay of particles resulting from jet fragmentation close to the neutral meson direction. Pythia fragmentation photons have a behavior somehow in between that of neutral mesons and partonic photons. The peak at low ΔR is still present but much reduced with respect to that of neutral mesons. Madgraph partonic photons exhibit a modulation of the Pythia partonic photon ΔR distribution, probably because Madgraph includes fragmentation as a bremsstrahlung contribution.

From this we can expect that the smaller the ΔR cone used in Frixione isolation (until $\Delta R \simeq 0.1$), the higher the discrimination against the neutral mesons and fragmentation photons. The discrimination against neutral mesons is higher than that against fragmentation photons (as is well-known experimentally). This can be seen in Fig. 103, which shows the isolation sum profile divided by the transverse energy of the photon for different cone sizes.

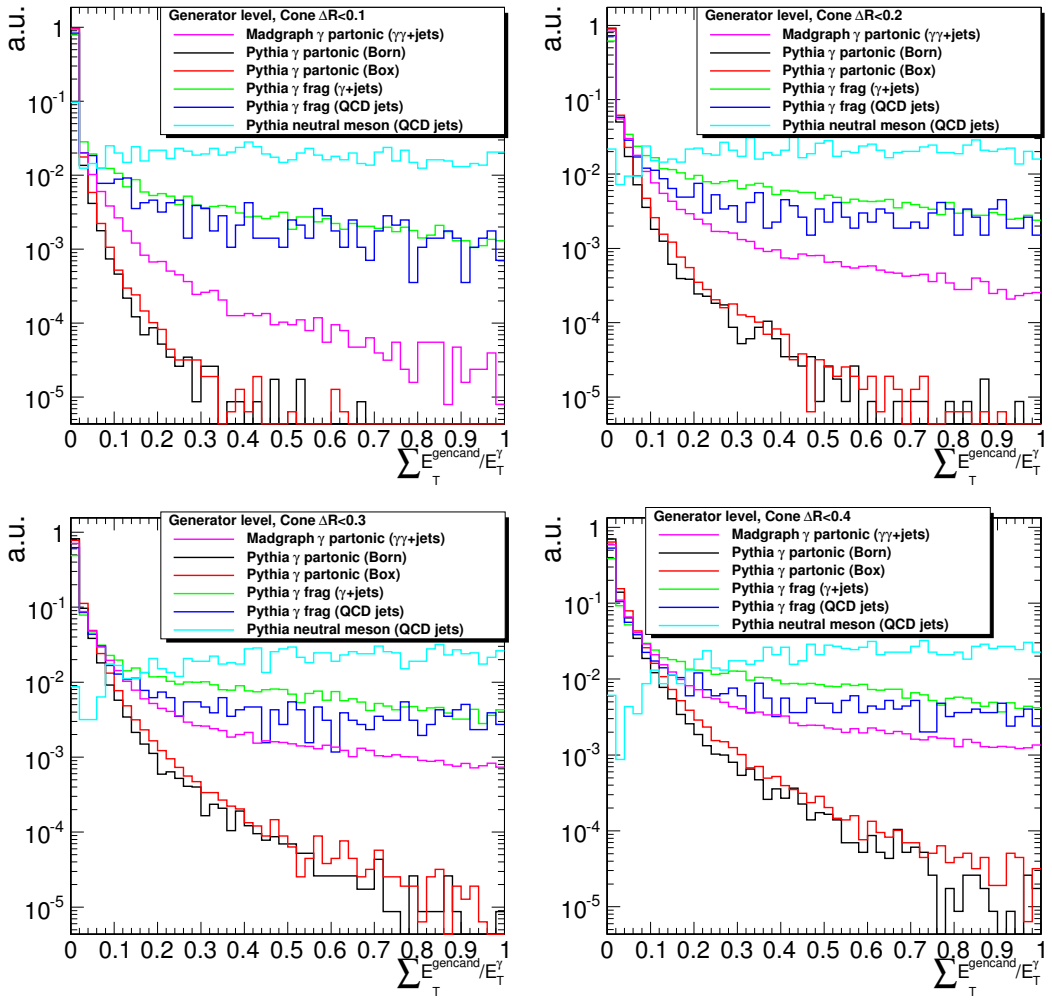


Fig. 103: Isolation sum normalized to the photon energy computed in cones of size $\Delta R < 0.1$ (top left), $\Delta R < 0.2$ (top right), $\Delta R < 0.3$ (bottom left), $\Delta R < 0.4$ (bottom right), for neutral mesons, fragmentation photons and partonic photons.

Table 14: Fraction represented by the 1-fragmentation and 2-fragmentation contributions for various Frixione isolation criteria in Pythia two-prompt photon samples (with electromagnetic enrichment filter).

Criteria	1-frag fraction	2-frag fraction	1,2-frag fraction
Solid $\Delta R < 0.4$, $E_T^{iso} < 5$ GeV	0.335	0.157	0.492
Hollow $0.1 < \Delta R < 0.4$, $E_T^{iso} < 4$ GeV	0.337	0.168	0.505
Frixione 8 cones, $E_T = 20$ GeV, $\epsilon = 0.05$, $n = 1.0$	0.322	0.145	0.467
Frixione 8 cones, $E_T = 20$ GeV, $\epsilon = 0.05$, $n = 0.2$	0.318	0.147	0.466
Frixione 8 cones, $E_T = 20$ GeV, $\epsilon = 1.0$, $n = 0.2$	0.372	0.228	0.599
Frixione 8 cones, $E_T = 20$ GeV, $\epsilon = 1.0$, $n = 0.1$	0.374	0.232	0.601
Frixione 8 cones, $E_T = 20$ GeV, $\epsilon = 1.0$, $n = 1.0$	0.353	0.192	0.545
Frixione 8 cones, $E_T = 20$ GeV, $\epsilon = 1.0$, $n = 0.5$	0.365	0.212	0.577
Frixione 8 cones, $E_T = 20$ GeV, $\epsilon = 0.5$, $n = 1.0$	0.343	0.176	0.518

19.32 Impact of hollow cones on the fragmentation contribution

In NLO generators, the products of the quark fragmentation are along the fragmentation photon direction. In parton-shower generators we have seen that this is not necessarily true. In NLO generators, the "hollow" or "crown" isolation, where the energy sum has to be below a fixed threshold in a region $R_1 < \Delta R < R_2$ while in the region $R < R_1$ any arbitrary amount of energy is admitted, has been shown to enhance the fragmentation contribution with respect to the usual "solid" isolation. This "hollow" isolation is interesting also because this criterion is closer than the "solid" cone to what is used experimentally (it allows to exclude from the isolation sum the energy deposited by the photon itself). The first two lines of Tables 14 and 15 show that in general the fragmentation fraction does not increase significantly when moving from solid to hollow cone isolation, for the PYTHIA samples.

19.33 Impact of Frixione isolation on the fragmentation contribution

Tables 14 and 15 report the fragmentation fraction inside the Pythia two-prompt sample for different Frixione isolation criteria. Eight isolation cones were used : $\Delta R < 0.05$, $\Delta R < 0.1$, $\Delta R < 0.15$, $\Delta R < 0.2$, $\Delta R < 0.25$, $\Delta R < 0.3$, $\Delta R < 0.35$, $\Delta R < 0.4$. The results are almost identical if instead four cones are used (0.1, 0.2, 0.3, 0.4), as found at NLO. The tables show that in both the electromagnetically-enriched samples and non-enriched samples, the discrete Frixione isolation with the usual functional form $f(R)$ does not reduce the fragmentation contribution with respect to the standard isolation criterion (with a cone $\Delta R < 0.4$) except when the parameter ϵ is at its smallest value, $\epsilon = 0.05$, for which modest reductions of between 5 and 8% can be achieved. The cause of this apparent non-optimal behavior can be explained by the non-collinearity of the fragmentation debris around the fragmentation photon in PYTHIA. Frixione isolation is designed to apply tighter and tighter isolation criteria $E_T^{iso} < f(R) \to 0$ as $\Delta R \to 0$, assuming that most of the fragmentation debris are around $\Delta R \simeq 0$. As it is seemingly not the case in the parton-shower Monte-Carlo studied here, the criterion loses most of its discrimination power.

The previous study suggests that the previous working points studied with the Frixione functional form $f(R)$ might not be optimal for the rejection of fragmentation debris. In figure 104 we compare the performance of three different sets of criteria: 1) non-Frixione isolation in a single cone $\Delta R < 0.4$, 2) optimized working points for the parameters in the Frixione functional form (four cones 0.1, 0.2, 0.3, 0.4 were used to make the algorithm converge faster), 3) re-optimized 'Frixione' isolation criteria on cones $\Delta R < 0.1, 0.2, 0.3, 0.4$ without using the explicit functional form (we no longer constrain the events to satisfy $E_T^{iso} < f(R)$ and let $f(R)$ free). In the second case, an optimization procedure is performed scanning over the parameters ϵ and n to find the best working points (corresponding to a maximum

Table 15: Fraction represented by the 1-fragmentation contribution for various Frixione isolation criteria in Pythia two-prompt photon samples (without enrichment filter).

Criteria	1-frag fraction
Solid $\Delta R < 0.4$, $E_T^{iso} < 5$ GeV	0.455
Hollow $0.1 < \Delta R < 0.4$, $E_T^{iso} < 4$ GeV	0.458
Frixione 8 cones, $E_T = 20$ GeV, $\epsilon = 0.05$, $n = 1.0$	0.420
Frixione 8 cones, $E_T = 20$ GeV, $\epsilon = 0.05$, $n = 0.2$	0.419
Frixione 8 cones, $E_T = 20$ GeV, $\epsilon = 1.0$, $n = 0.2$	0.514
Frixione 8 cones, $E_T = 20$ GeV, $\epsilon = 1.0$, $n = 0.1$	0.519
Frixione 8 cones, $E_T = 20$ GeV, $\epsilon = 1.0$, $n = 1.0$	0.489
Frixione 8 cones, $E_T = 20$ GeV, $\epsilon = 1.0$, $n = 0.5$	0.503
Frixione 8 cones, $E_T = 20$ GeV, $\epsilon = 0.5$, $n = 1.0$	0.465

efficiency for a given s/b). In the last case, an optimization code is used to find the best selection criteria to be applied on E_T^{iso} for each ΔR cone. The optimization takes as input the target value of s/b (partonic signal over fragmentation background ratio), then relaxes and tightens each cut separately with an iterative procedure to find the best signal efficiency for this s/b target. The procedure was performed to find the working points corresponding to the s/b obtained with the first Frixione criterion.

Figure 104 shows that the optimized working points for the Frixione functional form perform slightly better than the standard isolation for a given photon efficiency, and that optimization using no functional form in turn performs slightly better than the Frixione functional form; for the same value of single-photon efficiency, lower values of fragmentation fraction are attainable. It should be noted that this optimisation leads to a looser cut on the first cone, $\Delta R < 0.1$, than the usual functional form. Nevertheless, to obtain reductions in the fragmentation fraction of more than 10%, increasingly significant reductions in single photon efficiency are required, since the fragmentation reduction becomes nearly flat.

All in all, with the definition of the isolation in a cone of ΔR used here, which is a usual way of defining isolation at the experimental level (where one has however to remove the footprint of the photon from the isolation sum and to cope with pile-up), rejecting fragmentation photons can be done only at a cost of a lowered signal efficiency. With this optimization procedure it was found that to decrease the fragmentation fraction by 10%, a signal loss of about 60% has to be achieved, leading to extremely tight cuts probably not applicable in experimental analysis.

CONCLUSIONS

Firstly for the NLO cross sections, it was found that only one of the Frixione isolation criteria suggested in [18] actually performs better at removing the fragmentation contribution in the inclusive case than that of the standard cone, although potentially too tight to use experimentally. However, it is useful to see that the results are independent of the number of cones used. This is also the case for the di-photon cross section where it compares well to the continuous criteria. A more promising result in the inclusive case is that the generalized version of Frixione isolation, with small values of ϵ, do significantly reduce the fragmentation fraction without applying too tight a cut in the smallest cone. In addition it is confirmed that the 'saddle point' can not be found when altering the scale choice for the inclusive cross-section, suggesting more corrections needed at NNLO. Finally when moving to the photon with associated jet cross section, from the inclusive cross section, it is reported that the jet algorithm (or size) used at NLO makes little difference to the cross section.

In the study at the parton-shower level it was shown that the isolation profile of fragmentation

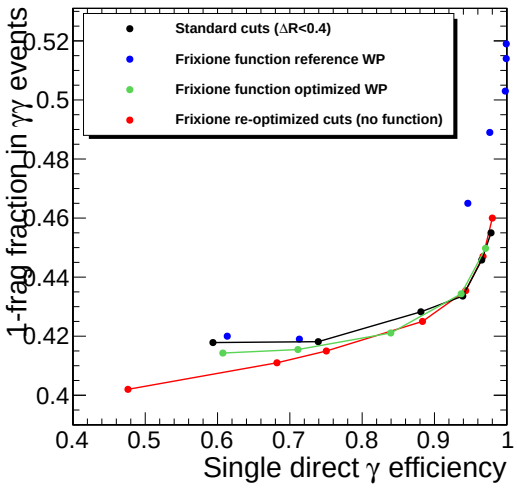


Fig. 104: The fraction of the 1-fragmentation contribution vs single γ efficiency for various sets of criteria. Blue : tested working points reported on table 15. Black : selection criteria on isolation in $\Delta R < 0.4$. Green : Optimized Frixione isolation using the usual functional form. Red : Re-optimized isolation criteria on cones $\Delta R < 0.1, 0.2, 0.3, 0.4$ without using the functional form.

photons is no longer collinear anymore, whereas the modelization of the fragmentation function in NLO generators leads the quark/gluon debris to be collinear to the photon. Furthermore, the hadronization process of the quark/gluon that emitted the photon leads to a ΔR profile which is no longer peaked at zero (but close to zero). In parton-shower programs isolation still has increasing discriminating power when going lower in ΔR. However, it was found that a 10% decrease in fragmentation fraction in di-photon events with respect to standard isolation leads to a drop in single photon signal efficiency to approximately 60% of the initial value. The usual functional form for Frixione isolation was shown to be not completely optimal for suppressing the fragmentation contribution while preserving high signal efficiency. This can be mitigated by re-optimizing the cuts for each ΔR of the discrete Frixione prescription, which allows a looser cut in the innermost cone. Further studies using other fragmentation modelizations in parton-shower programs like SHERPA [427] (LO matrix-element where photons and jets in the shower are matched to matrix element level) or POWHEG [484] (NLO matrix-element with consistent fragmentation photon matching) would need to be investigated.

In conclusion the results from the two studies show there are differences and similarities at the two levels. Regarding the fragmentation fraction, this is far more reduced at the NLO level than at the parton-shower level. However, the two levels show agreement that the results from Frixione isolation are independant of the number of cones used and that similar shape cuts can be obtained by retuning cuts at the parton-shower level and by using the generalized prescription at NLO.

ACKNOWLEDGEMENTS

Thanks to the whole of the Les Houches photon subgroup, especially to J.P. Guillet and E. Pilon for their expert input.

APPENDIX: Selection details

All of the studies are carried out for pp collisions at $\sqrt{S} = 7$ TeV. For simplicity the the inclusive studies are carried out for the region where the photon lies in $|\eta| < 0.6$. When calculating the cone isolation around the photon a cone of 0.4 is used with the requirement that the energy in the cone is less than $4\mathrm{GeV}$. The renormalization scale μ and initial state factorization scale M are set to the photon p_T,

unless stated otherwise, and the CTEQ6.6 PDF[256] is used and the photon fragmentation functions are BFG set II [479]. For the generalized Frixione isolation case, the cones used are: R = 0.4, 0.35, 0.3, 0.25, 0.2, 0.15 and 0.1.

In the NLO di-photon studies, the photons have: $p_T^{\gamma\,1} \leq 25$ GeV, $p_T^{\gamma\,2} \geq 22$ GeV, in the rapidity range $|\eta^\gamma| \leq 2.5$ for both photons, and a separation $\Delta R_{\gamma\gamma} \geq 0.4$ is required between the two photons. The mass range considered is 40 GeV $\leq m_{\gamma\gamma} \leq 300$ GeV. In this case the scales μ and M are chosen equal to min $\{p_T^{\gamma\,1}, p_T^{\gamma\,2}\}$.

In the parton shower di-photon studies, the photons are selected with: $M_{\gamma\gamma} > 80$ GeV, p_T >21,20 GeV, $|\eta|$ <2.5 and $E_T^{iso} < 5$ GeV.

20. EVENT-BY-EVENT PILEUP SUBTRACTION USING JET AREAS [49]

Abstract

In these proceedings, we compare the efficiency of several jet-area-based subtraction methods to correct for pile-up contamination at hadronic colliders. We study the dependence on various variables like the p_t and rapidity of the jets, the number of pile-up vertices or the Monte-Carlo generator variations. We conclude that estimations of the pile-up density using a median computed over grid-cell patches, including a rescaling to correct for the rapidity dependence, perform particularly well, though alternative methods are possible.

20.1 Introduction

With the LHC running at larger and larger luminosities, hard pp interactions are accompanied by an increasing number of pile-up (PU) collisions: from a few PU events per bunch crossing in spring 2011, operation with ~ 20 PU events is now routine. Considering only in-time PU, this would lead to an extra transverse momentum of ~ 750 GeV deposited in the event, and a jet of a typical radius $R = 0.5$ would see its transverse momentum shifted by ~ 10 GeV. In order to obtain a good energy resolution for the jets it is therefore mandatory to correct for this contamination.

In these proceedings, we review several methods — both existing methods and new refinements — to subtract the contamination due to PU and provide a systematic study of their efficiency.

It is important to note already now that PU has not only the effect to shift the momentum of the jets: it also smears their momentum. Indeed, the number of PU vertices varies from one collision to the next (following a Poisson distribution varying with the beam conditions), all PU interactions, *i.e.* minimum bias collisions, do not lead to the same energy deposit, and finally, the energy produced in a minimum bias collision is not deposited uniformly across the detector. Altogether, on top of an average shift, PU will add two sources of resolution smearing: *event-to-event* and *in-event* fluctuations corresponding respectively to variations of the PU activity from one event to another and from one point to another in a single event.

Here we shall primarily study *in-time PU*, that is the effects coming from multiple pp interactions that occur in the same bunch crossing as the hard interaction one triggers on. Because of the response time inherent to each detector this would come with a second effect, *out-of-time PU*, corresponding to the PU activity in the few bunch crossings preceding the one with the hard interaction. Since these heavily depend on the details of each individual detector — and even varies from one sub-detector to another — it goes beyond the scope of this theoretical study. However, as we shall discuss in further detail later on, the PU subtraction methods proposed here do not make any assumption about a distinction between in-time and out-of-time PU and thus should be robust enough in more complex cases.

[49] Contributed by: M. Cacciari, G. P. Salam and G. Soyez

20.2 Subtraction method(s)

We are interested in the situation where a *hard event* is contaminated by a *background* coming from additional pileup interactions. A reconstructed jet in that *full event* (hard event + background), which we shall call a *full jet*, differs from the *hard jet* in the original hard event because of the presence of the background. By *background subtraction*, we mean correcting the full jet in such as to recover the momentum of the original hard jet, *i.e.* subtract the pileup contamination from the jet's momentum.

20.21 Background effects

Our starting point is to realise [485] that a uniform background affects the momentum of a jet in two ways: it shifts its momentum because of the background particles clustered with the jet, and it modifies the way the hard particles themselves are clustered because the background particles are not infinitely soft.

This means that the reconstructed momentum has the form[50]

$$p_{t,full} = p_{t,\mathrm{hard}} + \rho A \pm \sigma\sqrt{A} + \Delta p_t^{BR} \tag{20.2.1}$$

where p_t denotes the transverse momentum of the reconstructed jet, $p_{t,\mathrm{hard}}$ the momentum of the original hard jet (in the absence of PU), A the jet area, ρ the background density per unit area within a given event, σ the fluctuations of that background (per unit area) from place to place within the event, and Δp_t^{BR} the back-reaction describing the effect of the background particles on the clustering of the hard ones.

If the background has a positional dependence (*e.g.* depends on rapidity) then ρ and σ will depend on the position of the jet one tries to subtract.

Eq. (20.2.1) characterises the fact that the background has the effects of shifting the transverse momentum of the jet and to degrade its resolution. The shift comes from the "ρ" term in (20.2.1) and from potential back-reaction systematic effects. Using the anti-k_t jet algorithm the shift due to back-reaction is negligible[51]. Resolution smearing effects come from various sources: the fluctuations of the background from within an event, *i.e.* the "σ" term in (20.2.1), fluctuations of the background from one event to another, that is the fact that ρ is not the same in every event, and the fluctuations in the back-reaction.

20.22 Central subtraction formula

From (20.2.1), the natural way to subtract the background contamination is to define the *subtracted jet* as [485]

$$p_{t,\mathrm{sub}} = p_t - \rho_{\mathrm{est}} A \tag{20.2.2}$$

where ρ_{est} is the estimated value for the background density per unit area.

To apply this subtraction we need to compute the jet area and find an estimation ρ_{est} for the background density per unit area. The jet areas are readily available using FastJet, so we just need to focus on ρ_{est}. The main goal of these proceedings is to investigate various methods of obtaining ρ_{est} which are listed below. In all cases, it is primordial to realise that the determination of ρ_{est} is performed event-by-event, and even jet-by-jet when the positional dependence of the background is taken into account.

As we shall see later on, the fact that ρ is estimated for each individual event is crucial: it corrects for the fluctuations of the background from one event to another. If instead one uses an averaged value for ρ_{est} (over many events), one would get an extra resolution smearing due to the fluctuations of ρ across different events. Similarly, the jet area A in (20.2.2) has to be computed for each individual jet. Using an average area would lead to an additional source of fluctuations of the form $\rho\sqrt{\langle A^2\rangle - \langle A\rangle^2}$.

[50] This can be defined for the 4-momentum of the jet but we shall only discuss its transverse momentum for simplicity.

[51] For the k_t or Cambridge/Aachen algorithms, it is usually negative and can be of the order of a GeV.

Using *seen vertices* Since experimentally it might be possible — within some level of accuracy that goes beyond the scope of this discussion — to count the number of pileup vertices using charged track reconstruction, one appealing way to estimate the background density in a given event would be to count these vertices and subtract a pre-determined number for each of them:

$$\rho_{\text{est}}^{(n\text{PU})}(y) = f(y)\, n_{\text{PU,seen}}, \tag{20.2.3}$$

where we have made explicit the fact that the proportionality constant $f(y)$ can carry a rapidity dependence. $f(y)$ can be studied from minimum bias collisions (see Section 20.32 below) and can take into account the fact that only a fraction of the PU vertices will be reconstructed.

Median subtraction This technique divides the rapidity-azimuthal angle plane in patches and estimates ρ for each event using

$$\rho_{\text{est}}^{(\text{global})} = \underset{i\in \text{patches}}{\text{median}} \left\{ \frac{p_{t,i}}{A_i} \right\} \tag{20.2.4}$$

This is motivated by the observation that many regions in the event are populated just by the background. In these regions, p_t/A is an estimate of ρ and the use of the median, rather than the average, which ensures reduced bias from the hard jets.

This method was originally proposed in [485] using jets (from a k_t or Cambridge/Aachen clustering) as patches. Here, we shall also test a new option where the $y-\phi$ plane is simply subdivided into grid cells that we use as patches.

Using a local range Eq. (20.2.4) provides a unique, global, estimate of ρ for the event but does not take into account the positional-dependence of the background. One option, assuming one wants to estimate ρ at the location of a jet j, is to limit the computation of the median to the jets in the vicinity of j, that is[52]

$$\rho_{\text{est}}^{(\text{local})}(j) = \underset{\text{jets } i\in\mathcal{R}(j)}{\text{median}} \left\{ \frac{p_{t,i}}{A_i} \right\} \tag{20.2.5}$$

where $\mathcal{R}(j)$ is a *local range* around j. A typical example, that we shall study later on, is the case of a *strip range* where only the jets with $|y-y_j| < \Delta$ are included. This option was already proven to be powerful in [486].

Using rescaling Another option to correct for the rapidity dependence of the background[53] is to introduce a pre-computed rapidity-reshaping function $f(y)$ (see Section 20.32) and use

$$\rho_{\text{est}}^{(\text{resc.})}(y) = f(y) \underset{i\in \text{patches}}{\text{median}} \left\{ \frac{p_{t,i}}{A_i f(y_i)} \right\} \tag{20.2.6}$$

where now all patches (jets or grid cells) are included in the computation of the median.

20.3 Performance tests

20.31 Testing framework

The remainder of these proceedings will be devoted to an in-depth comparison of the subtraction methods proposed in Section 20.2. Our testing framework will be very similar to the one used in [486]: we embed a hard event into a pileup background (see again Section 20.2, we reconstruct and subtract the jets in

[52] This option will only be considered in the case where jets are used as patches.

[53] The same technique should also work for the azimuthal-angle dependence of the underlying-event in heavy-ion collisions.

both the hard and full events[54], for each jet in the hard event, we find the matching jet in the full event and compute the shift

$$\Delta p_t = p_t^{\text{full,sub}} - p_t^{\text{hard,sub}}, \tag{20.3.1}$$

i.e. the difference between the reconstructed-and-subtracted jet with and without pileup. A positive (resp. negative) Δp_t would mean that the PU contamination has been underestimated (resp. overestimated).

Though in principle there is some genuine information in the complete Δp_t distribution — *e.g.* it could be useful to deconvolute the extra smearing brought by the pileup, see *e.g.* [486] and [487] — we shall focus on two simpler quantities: the average shift $\langle \Delta p_t \rangle$ and the dispersion $\sigma_{\Delta p_t}$. While the first one is a direct measure of how well one succeeds at subtracting the pileup contamination on average, the second quantifies the remaining effects on the resolution. One thus wishes to have $\langle \Delta p_t \rangle$ close to 0 and $\sigma_{\Delta p_t}$ as small as possible. Note that these two quantities can be studied as a function of variables like the rapidity and transverse momentum of the jets or the number of pileup interactions. In all cases, a flat behaviour would indicate a robust subtraction method.

The robustness of our conclusions can be checked by varying many ingredients:

- one can study various hard processes with the hope that the PU subtraction is not biased by the hard event. In what follows we shall study dijets with p_t ranging from 50 GeV to 1 TeV, as well as fully hadronic $t\bar{t}$ events as a representative of busier final states.
- The Monte-Carlo used to generate the hard event and PU can be varied. For the hard event, we have used Pythia 6.4.24 [400] with the Perugia 2011 tune, Pythia 8.150 with tune 4C [348] and Herwig 6.5.10 [488] with the ATLAS tune and we have switched multiple interactions on (our default) or off. For the minimum bias sample used to generate PU, we have used Pythia 8, tune 4C, and checked that our conclusions remain unchanged when using Herwig++ [489] (tune LHC-UE7-2).

Additional details of the analysis For the sake of completeness, we list here the many other details of how the Δp_t analysis has been conducted: we have considered particles with $|y| \leq 5$ with no p_t cut or detector effect; jets have been reconstructed with the anti-k_t algorithm with $R = 0.5$ keeping jets with $|y| \leq 4$; for area computations, we have used active areas with explicit ghosts with ghosts placed[55] up to $|y| = 5$; for jet-based background estimations, we have used the k_t algorithm with $R = 0.4$ though other options will be discussed (and the 2 hardest jets in the set have been excluded from the median computation to reduce the bias from the hard event); for grid-based estimations the grid extends up to $|y| = 5$ with cells of edge-size 0.56 (other sizes will be investigated); for estimations using a local range, a strip range of half-width 1.5 has been used and we refer to the Section 20.32 below for more information about the rapidity rescaling. Jet reconstruction, area computation and background estimation have all been carried out using FastJet (v3) [361, 435]. Pile-up is generated as a superposition of a Poisson-distributed number of minimum bias events and we will vary the average number of pileup interactions. We shall always assume pp collisions with $\sqrt{s} = 7$ TeV. Finally, the matching of a full jet to a hard jet is made by requiring that their common constituents contribute for at least 50% of the transverse momentum of the hard jet. We shall not discuss matching efficiencies here but they are extremely good: for a reconstructed (full) jet of 50 GeV and 20 PU events, the matching efficiency is 99.9% and this increases to 99.98% for $p_t \geq 50$ GeV and 5 PU events and 99.995% for $p_t \geq 100$ GeV and 20 PU events.

20.32 Minimum bias and rapidity shape

Before discussing the performances of the subtraction methods described in Section 20.2, there is still a building block that has to be discussed, namely the rapidity dependence of the background $f(y)$ that

[54] One may argue whether or not one should subtract the jets in the hard event. We decided to do so to cover the case where the hard event contains Underlying Event which, as a relatively uniform background, will also be subtracted together with the pileup.

[55] Note that we have used the ghost placement of FastJet 3 which differs slightly from the one in v2.4.

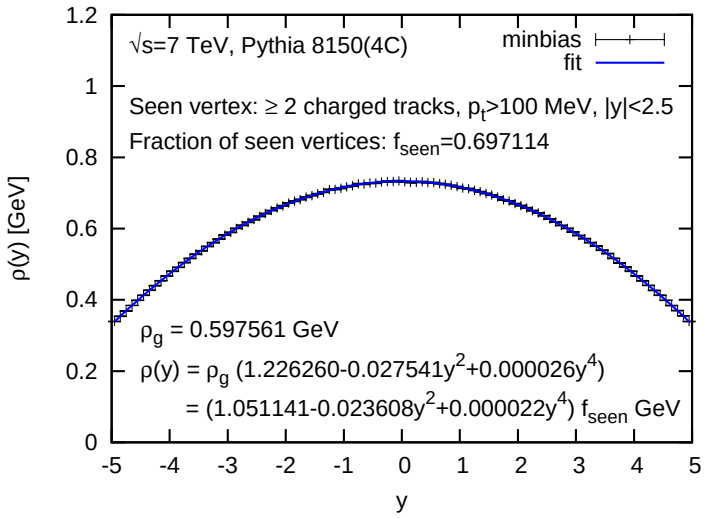


Fig. 105: Rapidity dependence of the transverse energy per unit area deposited in minimum bias events (obtained from Pythia 8, tune 4C). The normalisation of the fit is such that $f_{\rm seen}$ is the fraction of *seen* minimum bias events *i.e.* the fraction of events which have at least 2 charged tracks with $|y| \leq 2.5$ and $p_t \geq 100$ MeV.

enters in Eqs. (20.2.3) and (20.2.6). Letting aside the question of in-time vs. out-of-time PU and non-linear effects in the detectors, the shape $f(y)$ can be obtained directly from minimum bias events.

In our case, we have generated minimum bias events with Pythia 8 (tune 4C) and studied the rapidity dependence of the transverse momentum deposited per unit area. The result is shown on Fig. 105 together with a quartic fit. If $f(y)$ is used to rescale median-based estimates of ρ, Eq. (20.2.6), any global normalisation factor would cancel, but in the case of Eq. (20.2.3) *i.e.* for the "seen vertices" method, the normalisation has to match what we mean by a *seen* PU vertex. In what follows, we shall define that as a minimum bias interaction that has at least 2 charged tracks with $|y| \leq 2.5$ and $p_t \geq 100$ MeV, which corresponds to 69.7% of the events[56]. In these conditions, we have found that the rapidity dependence is well reproduced by

$$f(y) = 1.051141 - 0.023608\, y^2 + 0.000026\, y^4. \tag{20.3.2}$$

20.33 Generic performance and rapidity dependence

Let us begin our performance benchmarks by the study of the rapidity dependence of PU subtraction. First of all, Fig. 106 shows the residual average shift ($\langle \Delta p_t \rangle$) as a function of the rapidity of the hard jet. These results are presented for different hard processes, generated with Pythia 8 and assuming an average of 10 PU events per hard interaction. Robustness w.r.t. that choice will be discussed in the next Section but does not play any significant role for the moment.

The first observation is that the subtraction based on the number of seen PU vertices does a very good job in all 3 cases. Then, global median-based (using jets or grid cells) estimations of ρ, *i.e.* the (red) square symbols, do a fair job on average but, as expected, fail to correct for the rapidity dependence of the PU contamination. If one now restricts the median to a rapidity strip around the jet, the (blue) triangles, or if one uses rapidity rescaling, the (black) circles, the residual shift is very close to 0, typically a few hundreds of MeV, and flat in rapidity.

Note that the strip-range approach seems to have a small residual rapidity dependence and overall offset for high-p_t processes or multi-jet situations. That last point, more clearly observed with some

[56] This is a bit optimistic but does not affect in any way our discussion.

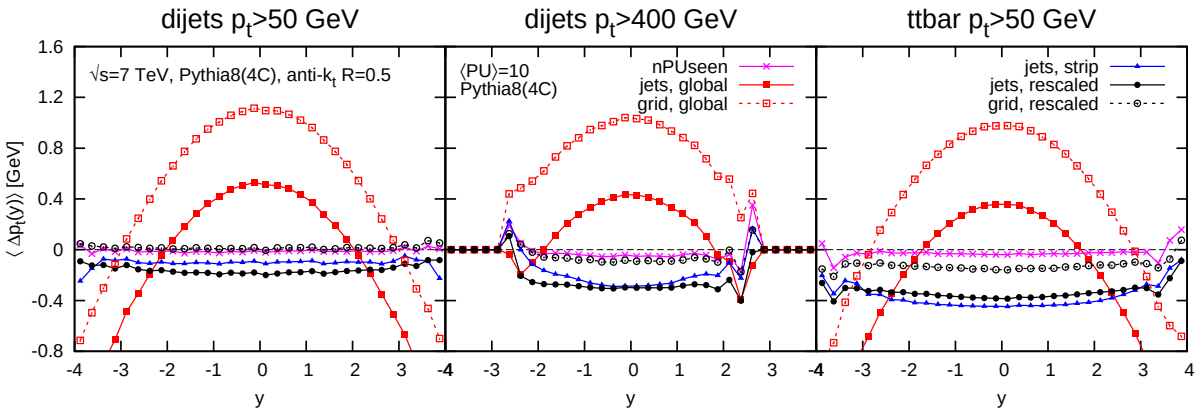


Fig. 106: Residual average shift as a function of the jet rapidity for all the considered subtraction methods. For the left (resp. centre, right) plot, the hard event sample consists of dijets with $p_t \geq 50$ GeV (resp. dijets with $p_t \geq 400$ GeV, and jets above $p_t \geq 50$ GeV in $t\bar{t}$ events), generated with Pythia 8 (tune 4C) in all cases. The typical PU contamination (for unsubtracted jets) is around 5 GeV.

Monte-Carlo generators like Pythia 6 than with others, may be due to the fact that smaller ranges tend to be more affected by the presence of the hard jets (see *e.g.* Appendix A.2 of [486]), an effect which is reinforced for multi-jet events. The fact that the residual shift seems a bit smaller for grid-based estimates will be discussed more extensively in the next Section.

Next, we turn to the dispersion of Δp_t, a direct measure of the impact of PU fluctuations on the p_t resolution of the jets. Our results are plotted in Fig. 107 as a function of the rapidity of the hard jet (left panel), the number of PU vertices (central panel) and the transverse momentum of the hard jet (right panel). All subtraction methods have been included as well as the dispersion one would observe if no subtraction were performed.

The results show a clear trend: first, a subtraction based on the number of seen PU vertices bring an improvement compared to not doing any subtraction; second, median-based estimations of ρ give a more significant improvement; and third, all median-based approaches perform similarly well.

The reason why median-based estimations of ρ outperform the estimation based on the number of seen PU vertices is simply because minimum bias events do not all yield the same energy deposit and this leads to an additional source of fluctuations in the "seen vertices" estimation compared to all median-based ones. This is the main motivation for using an event-by-event determination of ρ based on the energy deposited in the event. This motivation is further strengthened by the fact that additional issues like vertex resolution or out-of-time PU would affect both $\langle \Delta p_t \rangle$ and $\sigma_{\Delta p_t}$ if estimated simply from the number of seen vertices while median-based approaches are more robust.

Note finally that even though local ranges and rapidity rescaling do correct for the rapidity dependence of the PU on average, the dispersion still depends on rapidity. The increase with the number of PU vertices is in agreement with the expected $\sqrt{n_{\mathrm{PU}}}$ behaviour and the increase with the p_t of the hard process can be associated with *back-reaction*, see [486]. These numbers can also be compared to the typical detector resolutions which would be $\sim$10 GeV for 100 GeV jets and $\sim$20 GeV at $p_t = 400$ GeV [490, 491].

20.34 Robustness and Monte-Carlo dependence

The last series of results we want to present addresses the stability and robustness of the median-based estimation of the PU density per unit area.

To do that, the first thing we shall discuss is the Monte-Carlo dependence of our results. In Fig.

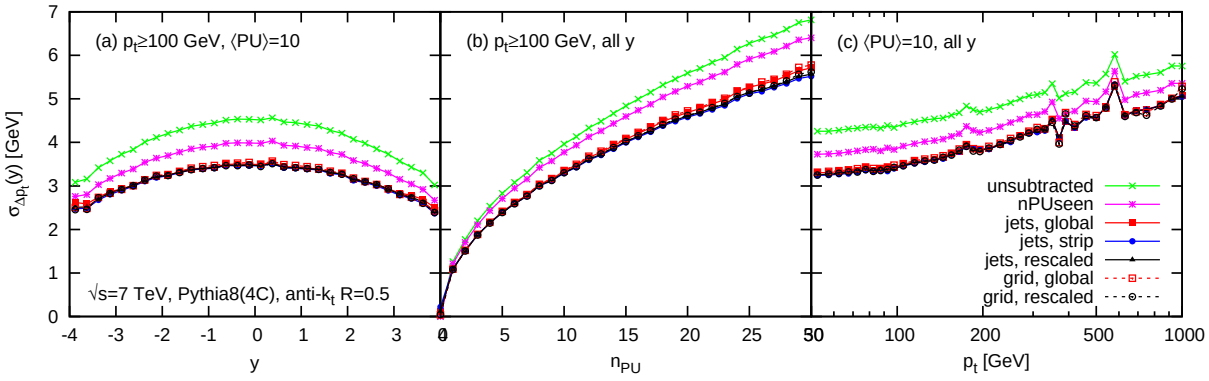


Fig. 107: Dispersion $\sigma_{\Delta p_t}$. Each curve corresponds to a different subtraction method and the results are presented as a function of different kinematic variables: left, as a function of the rapidity of the hard jet for a sample of jets with $p_t \geq 100$ GeV and assuming an average of 10 PU events; centre: as a function of the number of PU events for a sample of jets with $p_t \geq 100$; right: as a function of the p_t of the hard jet, assuming an average of 10 PU events

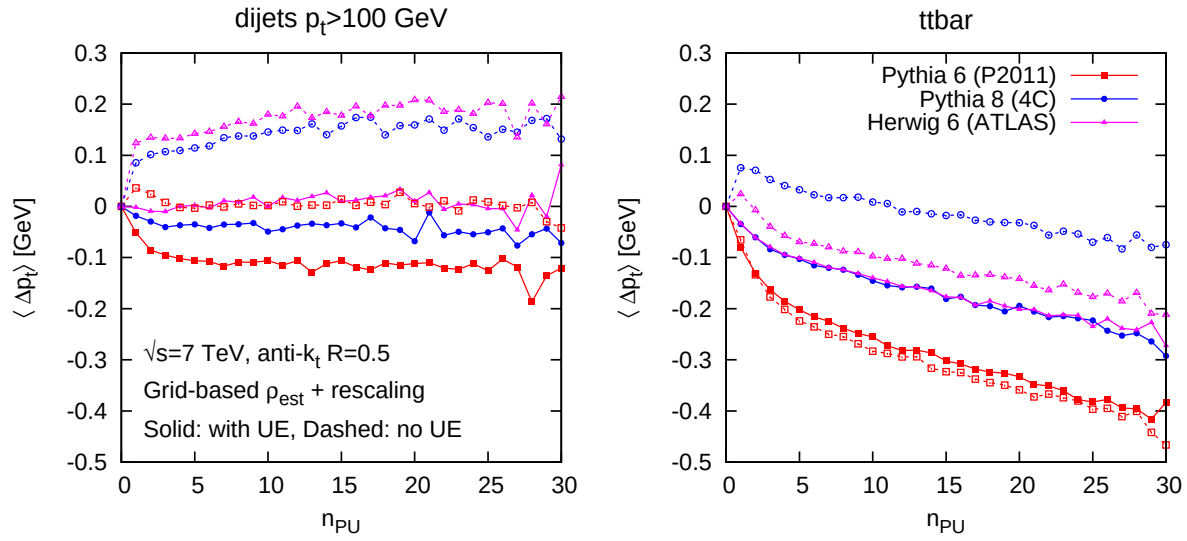


Fig. 108: Dependence of the average p_t shift as a function of the number of PU vertices for various Monte-Carlo generators. For the left plot, the hard sample is made of dijets with $p_t \geq 100$ GeV while for the right plot, we have used a hadronic $t\bar{t}$ sample. For each generator, we have considered both the case with the Underlying Event switched on (filled symbols) and off (open symbols). All results have been obtained using a grid-based median estimation of ρ using rapidity rescaling.

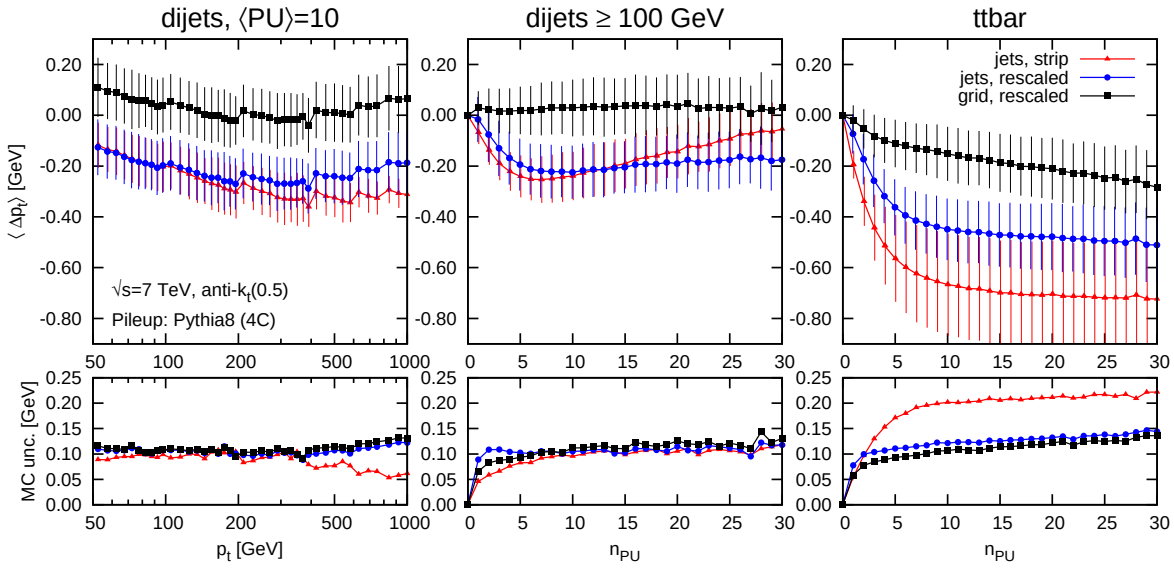


Fig. 109: Average residual shift after PU subtraction. $\langle \Delta p_t \rangle$ is plotted as a function of the p_t of the jet for an average of 10 PU events (left panel), or as a function of the number of PU vertices for dijets with $p_t \geq 100$ GeV (central panel) and for $t\bar{t}$ events (right panel). In all cases, we compare 3 methods: the rapidity-strip range, (red) triangles, the jet-based approach with y-rescaling, (blue) circles, and the grid-based approach with y rescaling, (black) squares. Each curve is the result of averaging over the various Monte-Carlo generator options and the dispersion between them is represented both as error bars on the top row and directly on the bottom row.

108 we compare the different Monte-Carlo predictions for the $\langle \Delta p_t \rangle$ dependence on the number of PU vertices in the case of a grid-based median estimate of ρ with rapidity rescaling. For each of the three considered Monte-Carlos, we have repeated the analysis with and without Underlying Event (UE) in the hard event. The first observation is that all the results span a range of 300-400 MeV in Δp_t and have a similar dependence on the number of PU vertices. The dependence on n_{PU} is flat for dijet events but shows a small decrease for the busier $t\bar{t}$ events. The 300-400 MeV shift splits into a 100-200 MeV effect when changing the generator, which is likely due to the small but non-zero effect of the hard event on the median computation, and a 100-200 MeV effect coming from the switching on/off of the UE.

This question of subtracting the UE deserves a discussion: since the UE is also a soft background which is relatively uniform, it contributes to the median estimate and, therefore, one expects the UE, or at least a part of it, to be subtracted together with the PU. Precisely for that reason, when we compute Δp_t, our subtraction procedure is not applied only on the "full jet" (hard jet+PU) but also on the hard jet, see Eq. (20.3.1). The 100-200 MeV negative shift observed in Fig. 108 thus means that, when switching on the UE, one subtracts a bit more of the UE in the full event (with PU) than in the hard event alone (without PU). This could be due to the fact (see [492] for details) that for sparse events, as is typically the case with UE but no PU, the median tends to slightly underestimate the "real" ρ, *e.g.* if half of the event is empty, the median estimate would be 0. This is in agreement with the fact that for $t\bar{t}$ events, where the hard event is busier, switching on the UE tends to have a smaller effect. Note finally that as far as the size of the effect is concerned, this 100-200 GeV shift has to be compared with the $\sim$1 GeV contamination of the UE in the hard jets.

Finally, we wish to compare the robustness of our various subtraction methods for various processes *i.e.* hard events and PU conditions. In order to avoid multiplying the number of plots, we shall treat the Monte-Carlo (including the switching on/off of the UE) as an error estimate. That is, an average measure and an uncertainty will be extracted by taking the average and dispersion of the 6 Monte-Carlo setups. The results of this combination are presented on Fig. 109 for various situations and subtraction

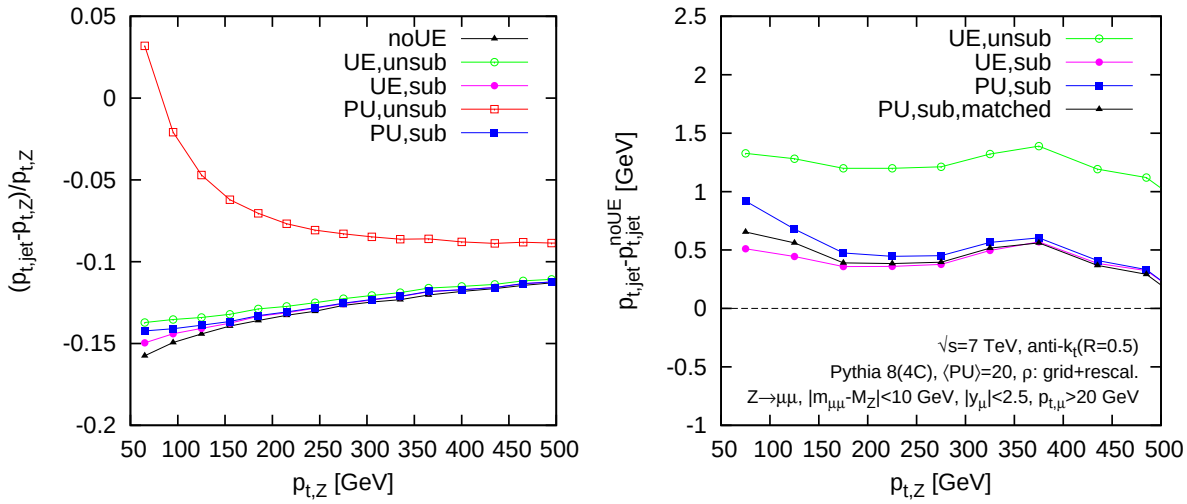


Fig. 110: Left: relative difference between the reconstructed jet and the reconstructed Z boson transverse momenta. Right: at a given p_t of the reconstructed Z boson, difference between the reconstructed p_t of the jet and the ideal p_t with no UE or PU, i.e. p_t shift w.r.t. the "noUE" curve, the (black) triangles, on the left panel. See the text for the details of the analysis.

methods. For example, the 6 curves from the left plot of Fig. 108 have been combined into the (black) squares of the central panel in Fig. 109.

Two pieces of information can be extracted from these results. First of all, for dijets, the quality of PU subtraction is, to a large extent, flat as a function of the p_t of the jets and the number of PU vertices. When moving to multi-jet situations, we observe an additional residual shift in the 100-300 MeV range, extending to ∼500 MeV for the rapidity-strip-range method. This slightly increased sensitivity of the rapidity-strip-range method also depends on the Monte-Carlo. While in all other cases, our estimates vary by ∼ 100 MeV when changing the details of the generator, for multi-jet events and the rapidity-strip-range approach this is increased to ∼200 MeV.

Overall, the quality of the subtraction is globally very good. Methods involving rapidity rescaling tends to perform a bit better than the estimate using a rapidity strip range, mainly a consequence of the latter's greater sensitivity to multi-jet events. In comparing grid-based to jet-based estimations of ρ, one sees that the former gives slightly better results, though the differences remain small.

Since the grid-based approach is considerably faster than the jet-based one, as it does not require an additional clustering of the event[57], the estimation of ρ using a grid-based median with rapidity rescaling comes out as a very good default for PU subtraction. One should however keep in mind local-range approaches for the case where the rapidity rescaling function cannot easily be obtained.

20.4 PU v. UE subtraction: an analysis on Z+jet events

To give further insight on the question of what fraction of the Underlying Event gets subtracted together with the pileup, we have performed an additional study of Z+jet events. We look at events where the Z boson decays into a pair of muons. We have considered 5 different situations: events without PU or UE, events with UE but no PU subtracted or not, and events with both UE and PU again subtracted or not. Except for the study of events without UE, this analysis could also be carried out directly on data.

Practically, we impose that both muons have a transverse momentum of at least 20 GeV and have $|y| \leq 2.5$, and we require that their reconstructed invariant mass is within 10 GeV of the nominal Z

[57] Note that the clustering of the main event still needs to include the computation of jet areas since they are needed in Eq. (20.2.2).

mass. As previously, jets are reconstructed using the anti-k_t jet algorithm and the pileup subtraction is performed using the grid-based-median approach with rapidity rescaling and a grid size of 0.55. All events have been generated with Pythia 8 (tune 4C) and we have assumed an average PU multiplicity of 20 events.

In Fig. 110, we have plotted the ratio $p_{t,\mathrm{jet}}/p_{t,Z} - 1$, with $p_{t,\mathrm{jet}}$ the transverse momentum of the leading jet, for the various situations under considerations. Compared to the ideal situation with no PU and no UE, the (black) triangles, one clearly sees the expected effect of switching on the UE, the empty (green) circles, or adding PU, the empty (red) squares: the UE and PU add to the jet $\sim$1.2 and 13 GeV respectively.

We now turn to the cases where the soft background is subtracted, *i.e.* the filled (blue) squares and (magenta) circles, for the cases with and without PU respectively. There are two main observations:

- with or without PU, the UE is never fully subtracted: from the original 1-1.5 GeV shift, we do subtract about 800 MeV to be left with a 0-500 MeV effect from the UE. That effect becomes smaller and smaller when going to large p_t.
- in the presence of PU, the subtraction produces results very close to the corresponding results without PU and where only the UE is subtracted. This nearly perfect agreement at large $p_{t,\mathrm{jet}}$ slightly degrades into an additional offset of a few hundreds of MeV when going to smaller scales. This comes about for the following reason: the non-zero p_t resolution induced by pileup (even after subtraction) means that in events in which the two hardest jets have similar p_t, the one that is hardest in the event with pileup may not correspond to the one that is hardest in the event without pileup. This introduces a positive bias on the hardest jet p_t (a similar bias would be present in real data even without pileup, simply due to detector resolution). The "matched" curve in Fig. 110 (right) shows that if, in a given hard event supplemented with pileup, we explicitly use the jet that is closest to the hardest jet in that same event without pileup, then the offset disappears, confirming its origin as due to resolution-related jet mismatching.

20.5 Conclusions and discussion

In these proceedings, we have investigated several methods to correct for the pile-up contamination to jets. They are all based on the observation that the average PU contribution to a jet is on average proportional to its area, which directly leads to eq. (20.2.2). The various methods then differ by the method used to estimate the PU activity per unit area, ρ. The subtraction efficiency has been studied by embedding hard events into PU backgrounds and investigating how jet reconstruction was affected by measuring the remaining p_t shift after subtraction ($\langle \Delta p_t \rangle$) as well as the impact on resolution ($\sigma_{\Delta p_t}$).

There are 3 broad approaches to the estimation of ρ: (a) using an average contamination per PU vertex, the *seen vertices* approach, (b) using an event-by-event estimation and, the *median* approach with jets or grid cells as patches, and (c) using an event-by-event and jet-by-jet method, the *local range* or *rescaling* approaches.

The first important message is that, though all methods give a very good overall subtraction ($\langle \Delta p_t \rangle \approx 0$), event-by-event methods should be preferred because their smaller PU impact on the p_t resolution (see Fig. 107). This is mostly because the "seen vertices" method has an additional smearing coming from the fluctuations between different minimum bias collisions. This does not happen in event-by event methods that are only affected by point-to-point fluctuations in an event. Note also that event-by-event methods are very likely more robust than methods based on identifying secondary vertices when effects like vertex identification and out-of-time PU are taken into account.

The next observation is that event-by-event and jet-by-jet methods have the additional advantage that they correct for positional-dependence of the background like its rapidity dependence (see Fig. 106). The median approach using a local range (with jets as patches) or rapidity rescaling (using jets or grid cells as patches) all give an average offset in the 0-300 MeV range, independently of the rapidity of

the jet, its p_t or the number of PU vertices, see Fig. 109 and are thus very suitable methods for PU subtraction at the LHC. Pushing the analysis a bit further one may argue that the local-range method has a slightly larger offset when applied to situations with large jet multiplicity like $t\bar{t}$ events (the right panel of Fig. 109) though this argument seems to depend on the Monte-Carlo used to generate the hard-event sample. Also, since it avoids clustering the event a second time, the grid-based method has the advantage of being faster than the jet-based approach.

At the end of the day, *we can recommend the median-based subtraction method with rapidity rescaling and using grid cells as patches* as a powerful default PU subtraction method at the LHC. But one should keep in mind that the use of jets instead of grid cells also does a very good job and that local-ranges can be a good alternative to rapidity rescaling if the rescaling function cannot be computed. Also, though we have not discussed that in detail, a grid cell size of 0.55 is a good default as is the use of k_t jets with $R = 0.4$.

To conclude, let us make a few general remarks. First, our suggested method involves relatively few assumptions, which helps ensure its robustness. Effects like in-time v. out-of-time PU or detector response should not have a big impact. Many of the studies performed here can be repeated with "real data" rather than Monte-Carlo simulations. The best example is certainly the Z+jet study of Section 20.4 which could be done using data samples with different PU activity from 2010 and 2011. Also, the rapidity rescaling function can likely be obtained from minimum bias collision data and the embedding of a hard event into pure PU events could help quantifying the remaining $\mathcal{O}$(100 MeV) bias. Experimentally, it would also be interesting to investigate hybrid techniques where one would discard the charged tracks that do not point to the primary vertex and apply the subtraction technique described here to the rest of the event. This would have the advantage to further reduce fluctuation effects (roughly by a factor $\sim \sqrt{1/(1-f_{\rm chg})} \approx 1.6$, where $f_{\rm chg} \approx 0.61$ is the fraction of charged particles in an event). Finally, all the facilities to compute jet areas and background estimation — including jets or grid-cells as patches, local ranges and rescaling functions — are readily available from FastJet (v3.0.0 onward) using *e.g.* the `GridMedianBackgroundEstimator` or `Subtractor` tools.

Acknowledgements

We thank Andy Buckley, Suzanne Gascon-Shotkin, Paolo Francavilla, Peter Loch, Emily Nurse, and Mark Stockton for very stimulating discussion during and after the workshop. This work was supported in part by grants ANR-09-BLAN-0060, ANR-10-CEXC-009-01 and PITN-GA-2010-264564.

Part VI
MC TUNING AND OUTPUT FORMATS

21. TUNE KILLING: QUANTITATIVE COMPARISONS OF MC GENERATORS AND TUNES [58]

Abstract

We summarise the implementation, status, and scope of the "tune killing" project, which classifies MC generator codes and tunes according to their quality of data description across a range of LHC-relevant observables. The primary aim of the project is to provide sufficiently clear information about generator performance that the current large collection of available tunes may be objectively reduced to a more manageable standard set for common use by LHC experiments and phenomenologists. We make final recommendations as

[58] Contributed by: A. Buckley, G. Hesketh, H. Hoeth, F. Krauss, E. Nurse, S. Plätzer, H. Schulz

to which generators and tunes are in rude health, and those which are obvious candidates for retirement from active service.

21.1 INTRODUCTION

Popular MC generators are nowadays associated with a bewildering array of standard parameter configurations, called "tunes". This proliferation of tunes is due to the ongoing project to provide optimised descriptions of LEP, Tevatron and LHC data: as new data and techniques have become available, new tunes have been created, usually but not always with increasing quality of data description. This process looks set to continue, and hence there is a need for agreement on which tunes are of most common interest at a given time.

The PYTHIA6 [400] event generator in particular has been the *de facto* testbed for tuning due to the wealth of community expertise and its ubiquity of tuning parameters for physical processes. At the time of writing there are 77 tunes available via the built-in `PYTUNE` routine, and a further 10 or more presented by the ATLAS experiment alone (this counting of ATLAS tunes includes equivalently weighted tunes for multiple PDFs, but not systematic variation tunes, of which there are many more). With such a profligacy of configuration options, it is difficult to objectively decide which are to be preferred for LHC simulation without manually cross-referencing hundreds of plots. It is hence not uncommon for different experimental or phenomenological studies to use entirely disjoint MC generator setups, making comparison difficult. Ideally we would have a much smaller set of agreed-upon generator setups, but choosing such a privileged subset requires clear information on which to base our preferences.

As a first step to addressing this issue, we present here a comparative study of event generator codes and tunes across a range of observables, particularly those of relevance for LHC physics. The study is based on analyses from the Rivet [360] toolkit, and the resulting data descriptions are quantitatively scored based on measures of deviation from the data values, including χ^2 and median/maximum bin-wise deviations (in units of combined experimental, statistical, and theoretical uncertainties). The results are presented as a series of Web pages, using colour coded tables which are hyperlinked to provide the necessary information in a compact, hierarchical form.

21.2 Analysis system

The data analysed for this project was produced by individual runs of various generator/tune configurations into the Rivet analysis system. A choice of Rivet analyses was made, intended to cover a number of core QCD modelling aspects for LHC physics: these are documented in Table 16. In some cases only the most relevant range in the distribution is included, as indicated in the Table. The generators and tunes used are documented in Table 17.

Note that not all observables are suitable for all generators. For example, AlpGen has not been used for LEP fragmentation, although in a future iteration we will extend the AlpGen coverage to include underlying event observables, where the hard jets could interfere with those from the multiple parton interaction (MPI) mechanism. Several observables, notably hard photon physics, minimum bias observables, and fragmentation/strangeness from RHIC and LHC have not yet been included: this is envisaged as a future extension of the project.

A Python program was written to load the histogram files for each generator/tune combination from a hierarchical directory structure, and to perform some basic statistical characterisation on each bin, histogram, and semantic group of histograms. At the histogram level, specifications are used to determine which bins are to be considered in the statistical comparisons, and to add a nominal "theoretical uncertainty". In this study a 10% theoretical uncertainty was added to the underlying event and fragmentation observables and a 5% theoretical uncertainty on the rest. The combined uncertainty for each bin b is then computed from the sum in quadrature of the reference data error, the MC statistical error and the theoretical uncertainty, and is used to compute a MC–data deviation for that bin, expressed in units of

Observable	Rivet analysis	Ref.	Range
Underlying event			
Transverse region N_{ch} vs. $p_{\perp}^{\text{lead}}, p_{\perp}^{\text{ch}} > 500\,\text{MeV}$	ATLAS_2010_S8894728	[493]	$p_{\perp}^{\text{lead}} > 5\,\text{GeV}$
Transverse region $\sum p_{\perp}$ vs. $p_{\perp}^{\text{lead}}, p_{\perp}^{\text{ch}} > 500\,\text{MeV}$	ATLAS_2010_S8894728	[493]	$p_{\perp}^{\text{lead}} > 5\,\text{GeV}$
Transverse region $\langle p_{\perp} \rangle$ vs $N_{\text{ch}}, p_{\perp}^{\text{ch}} > 500\,\text{MeV}$	ATLAS_2010_S8894728	[493]	
Jets			
Toward region N_{ch} vs $p_{\perp}^{\text{lead}}, p_{\perp}^{\text{ch}} > 500\,\text{MeV}$	ATLAS_2010_S8894728	[493]	$p_{\perp}^{\text{lead}} > 5\,\text{GeV}$
Toward region $\sum p_{\perp}$ vs $p_{\perp}^{\text{lead}}, p_{\perp}^{\text{ch}} > 500\,\text{MeV}$	ATLAS_2010_S8894728	[493]	$p_{\perp}^{\text{lead}} > 5\,\text{GeV}$
Jet shapes, $30 < p_{\perp} < 40\,\text{GeV}$, $\lvert y \rvert < 2.8$	ATLAS_2011_S8924791	[494]	
Jet shapes, $310 < p_{\perp} < 400\,\text{GeV}$, $\lvert y \rvert < 2.8$	ATLAS_2011_S8924791	[494]	
Dijet $\Delta\phi$, $110 < p_{\perp} < 160\,\text{GeV}$	ATLAS_2011_S8971293	[495]	$3\pi/4 \to \pi$
Dijet $\Delta\phi$, $310 < p_{\perp} < 400\,\text{GeV}$	ATLAS_2011_S8971293	[495]	$3\pi/4 \to \pi$
Dijet mass, $0.3 < \lvert y \rvert < 0.8$, anti-$k_{\perp}(0.4)$	ATLAS_2010_S8817804	[382]	
Transverse thrust, $90\,\text{GeV} < p_{\perp}^{\text{jet1}} < 125\,\text{GeV}$	CMS_2011_S8957746	[496]	
ISR/intrinsic-$k_{\perp}$			
DØ ϕ^*, $\lvert y \rvert < 1.0$	D0_2010_S8821313	[497]	$\phi^* < 0.4$
DØ ϕ^*, $1.0 < \lvert y \rvert < 2.0$	D0_2010_S8821313	[497]	$\phi^* < 0.4$
Fragmentation			
N_{ch}, π^+/π^-, K^+/K^- at LEP	DELPHI_1996_S3430090	[498]	
ρ/π, K/π, $\Sigma^{\pm,+,-,0}/\pi$, p/π, Λ/π	PDG_HADRON_MULTIPLICITIES	[499]	
Inclusive x_p, thrust (+ major & minor)	DELPHI_1996_S3430090	[498]	
B fragmentation	DELPHI_2002_069_CONF_603	[500]	

Table 16: Observables used in the tune killing exercise.

Generator and version	Tunes
Sherpa 1.3.1[146]	Default (CTEQ6.6)
Herwig++ 2.5.2[489]	LHC-UE-EE-3 series (LO$*$$*$ and CTEQ6L1)
Pythia 8.150[348]	4C
PYTHIA 6.425[400]	D6T, DW[501], Z2, AMBT1[502], AUET2B (LO$*$$*$ and CTEQ6L1)[503], Perugia 2010[504], Perugia 2011[504], prof-Q^2[451]
AlpGen[505] + PYTHIA 6.425 (*)	Same tunes as PYTHIA6. Perugia 2011 using matched ME/PS Λ_{QCD}.[506]
HERWIG 6.510[488] + JIMMY 4.31[507]	AUET2 LO$*$$*$[508]
AlpGen + HERWIG 6.5 + JIMMY 4.31 (*)	Same as for HERWIG+JIMMY.

Table 17: Generators and tunes used in the tune killing exercise. (*) Jet and Z boson ϕ^* observables only.

the total bin error, $\text{dev}_b = (\text{MC}_b - \text{data}_b)/\text{err}_b$.

For each active histogram, the system then reports the χ^2/N_{bin}, and the median, mean, and maximum bin-wise deviation. A total "metric" value for each histogram is reported as the maximum bin-wise deviation if that is greater than 10σ, otherwise the greater of the median and mean deviations. This hybrid treatment of the metric allows the system to flag up histograms in which there are either widespread moderate deviations or a small number of very discrepant bins which might be missed with a pure median or mean deviation treatment. An HTML table and set of histograms are rendered by the system for each observable, with a continuous colour coding scheme used to highlight the relative quality of data description from ideal (green) to very poor (red).

The histograms are grouped to collect together observables from different sources which reflect related aspects of QCD modelling. The current groups are "Underlying Event (UE)", "Dijets", "Multijets", "Jet shapes", "W and Z", "Fragmentation", and "B fragmentation". In these groups, the same χ^2/N_{bin}, and mean/median/maximum deviation statistics are calculated as before. For visual compactness of classification we again use a hybrid performance metric for each histogram group: again this is the maximum bin-wise deviation found in the contained histograms if that is greater than 10σ, otherwise the maximum histogram-wise deviation metric in the group if that is greater than 5σ, otherwise the maximum of the median/mean bin-wise deviation.

The Web pages generated to present this data in a compact way consist of a single top level page containing a colour-coded table of tune performance metrics for each histogram group. Each cell in the table is hyperlinked to a more detailed table for that tune/group where the various χ^2/N, max/mean/median deviation and hybrid metric are presented, again colour-coded, for each histogram in the group. The table rows are then hyperlinked to a plot page showing explicitly the tune/generator behaviour for each histogram and indicating the active range of the histograms where appropriate. These pages are shown in Figures 111 to 113. This form of presentation allows a rapid assessment of generator/tune performance, while still permitting detailed investigation of any flagged-up issues with a few mouse clicks. The system is easily extensible to more observables, groups, and different theory uncertainty / visual classification thresholds.

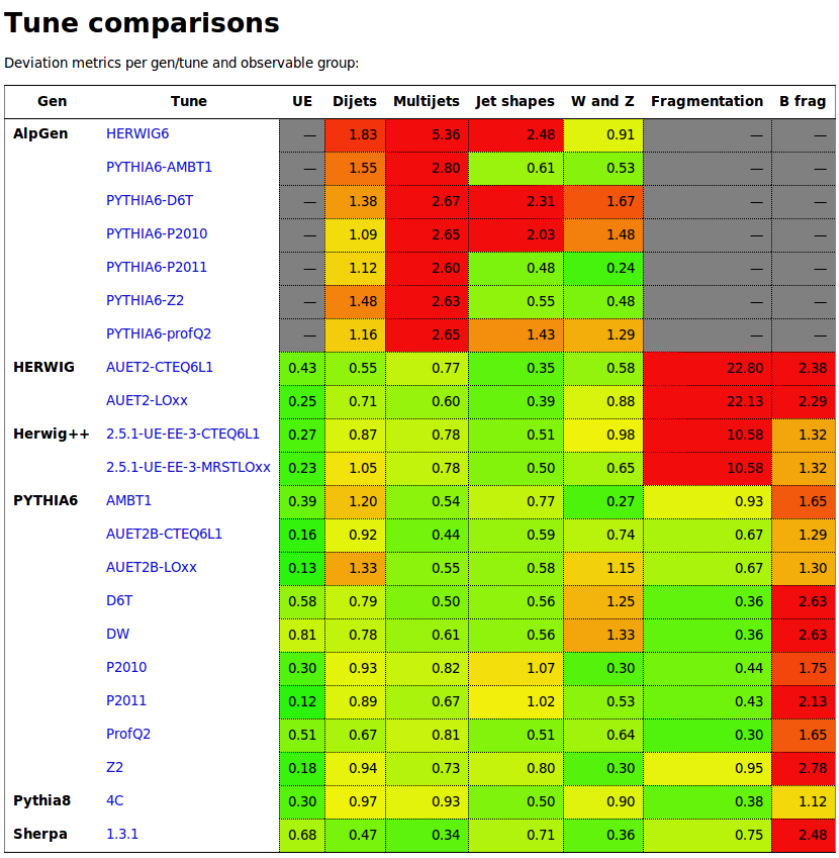
The classification colours for each performance figure are generated in HSB colour space as a linear variation in deviation x between green (120) and red (0) in the Hue parameter, i.e. $H = 120(1 - \min(x/x_{\text{bad}}, 1.0))$, with fixed Saturation and Brightness parameters. The visual threshold x_{bad} was chosen to be different for each metric type: 5σ for maximum deviations, 4σ for χ^2/N, and 2σ for mean and median deviations, and for the hybrid performance metrics. These thresholds were iterated from initial suggestions to the point where distinctions could be made between the models: similar iteration of the discriminating criteria are envisaged while significant model/tune variations exist as the motivation of this study is model discrimination rather than passing or failing a natural performance figure.

21.3 Results

As the central theme of this project has been to provide a comprehensible visualisation of the relative performance of generators and tunes, and hierarchical presentation via Web pages was key to achieving this, it would be self-defeating to attempt to present the same information in this summary. Additionally, the nature of tune comparison is that it evolves as new data, tunes, and generator versions become available. Hence, for up-to-date status information we refer the reader to the persistent "tune killing" web page at `http://projects.hepforge.org/rivet/tunecmp/`.

However, it *is* worth mentioning some of the most striking features of generators which have been made more evident by this collating of data–MC comparisons:

- The general quality of jet and W/Z data description is in fact better than expected: among PYTHIA tunes in particular there is sufficient variation in parton shower parameters that significant deviations in jet observables would reasonably be expected, but in fact the majority of tunes describe



Tune comparisons

Deviation metrics per gen/tune and observable group:

Gen	Tune	UE	Dijets	Multijets	Jet shapes	W and Z	Fragmentation	B frag
AlpGen	HERWIG6	—	1.83	5.36	2.48	0.91	—	—
	PYTHIA6-AMBT1	—	1.55	2.80	0.61	0.53	—	—
	PYTHIA6-D6T	—	1.38	2.67	2.31	1.67	—	—
	PYTHIA6-P2010	—	1.09	2.65	2.03	1.48	—	—
	PYTHIA6-P2011	—	1.12	2.60	0.48	0.24	—	—
	PYTHIA6-Z2	—	1.48	2.63	0.55	0.48	—	—
	PYTHIA6-profQ2	—	1.16	2.65	1.43	1.29	—	—
HERWIG	AUET2-CTEQ6L1	0.43	0.55	0.77	0.35	0.58	22.80	2.38
	AUET2-LOxx	0.25	0.71	0.60	0.39	0.88	22.13	2.29
Herwig++	2.5.1-UE-EE-3-CTEQ6L1	0.27	0.87	0.78	0.51	0.98	10.58	1.32
	2.5.1-UE-EE-3-MRSTLOxx	0.23	1.05	0.78	0.50	0.65	10.58	1.32
PYTHIA6	AMBT1	0.39	1.20	0.54	0.77	0.27	0.93	1.65
	AUET2B-CTEQ6L1	0.16	0.92	0.44	0.59	0.74	0.67	1.29
	AUET2B-LOxx	0.13	1.33	0.55	0.58	1.15	0.67	1.30
	D6T	0.58	0.79	0.50	0.56	1.25	0.36	2.63
	DW	0.81	0.78	0.61	0.56	1.33	0.36	2.63
	P2010	0.30	0.93	0.82	1.07	0.30	0.44	1.75
	P2011	0.12	0.89	0.67	1.02	0.53	0.43	2.13
	ProfQ2	0.51	0.67	0.81	0.51	0.64	0.30	1.65
	Z2	0.18	0.94	0.73	0.80	0.30	0.95	2.78
Pythia8	4C	0.30	0.97	0.93	0.50	0.90	0.38	1.12
Sherpa	1.3.1	0.68	0.47	0.34	0.71	0.36	0.75	2.48

Fig. 111: Screenshot of the top-level summary page produced by the tune comparison system.

Jet shapes

Histo	chi2/Ndf	Median deviation	Mean deviation	Max deviation	Metric
Jet shape ρ for $p_\p... (ATLAS_2011_S8924791/d01-x06-y01)	0.59	0.77	0.69	1.17	0.77
Jet shape ρ for $p_\p... (ATLAS_2011_S8924791/d09-x06-y01)	0.14	0.36	0.30	0.61	0.36
Central Transv. Thrust, $9... (CMS_2011_S8957746/d01-x01-y01)	0.37	0.43	0.53	1.08	0.53
Central Transv. Minor, $90... (CMS_2011_S8957746/d02-x01-y01)	0.34	0.38	0.48	1.14	0.48

W and Z

Histo	chi2/Ndf	Median deviation	Mean deviation	Max deviation	Metric
Muon channel ($|y_Z|<1$) (D0_2010_S8821313/d02-x01-y01)	0.77	0.70	0.79	1.52	0.79
Muon channel ($1<|y_Z|<2$) (D0_2010_S8821313/d02-x01-y02)	0.27	0.38	0.45	1.08	0.45

Fragmentation

Histo	chi2/Ndf	Median deviation	Mean deviation	Max deviation	Metric
Scaled momentum, $x_p = |p... (DELPHI_1996_S3430090/d07-x01-y01)	0.74	0.28	0.47	3.29	0.47
$1-\text{Thrust}$ (DELPHI_1996_S3430090/d11-x01-y01)	4.12	0.24	1.13	6.26	1.13
Thrust major, M (DELPHI_1996_S3430090/d12-x01-y01)	7.24	0.50	1.61	9.37	1.61
Thrust minor, m (DELPHI_1996_S3430090/d13-x01-y01)	9.69	0.40	1.57	10.58	10.58
Mean charged multiplicity (DELPHI_1996_S3430090/d35-x01-y01)	0.08	0.28	0.28	0.28	0.28

Fig. 112: Screenshot of the mid-level performance metric page produced by the tune comparison system. This specific example is part of the performance metrics for the Herwig++ LHC-UE-EE-3 LO$*$ tune.

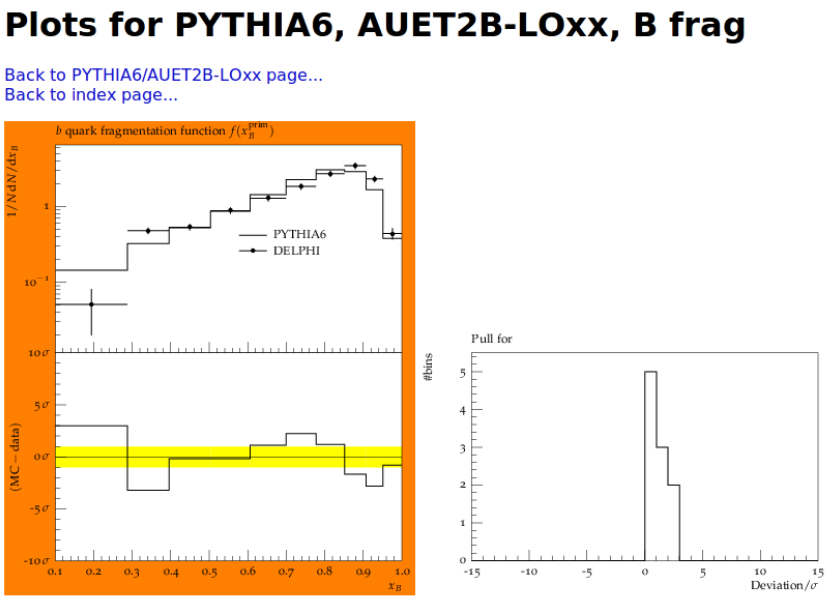


Fig. 113: Screenshot of the observable plot page produced by the tune comparison system. This specific example shows the B fragmentation performance of the PYTHIA6 AUET2B LO$*$$*$ tune.

data fairly well.

- The UE in particular has been a focus of tuning activity and this is evident in the consistency of UE data description. The worst performance in this group is from the DW tune of PYTHIA, but even this pre-LHC tune with the "old" PYTHIA MPI model achieves a deviation metric of less than 1σ on LHC UE observables.
- PYTHIA D6T outperforms PYTHIA DW – an unexpected result since the MPI energy evolution of D6T is fixed to the default and disfavoured form $p_\perp^0(s) \sim (\sqrt{s}/1800\,\text{GeV})^{0.16}$, whereas the exponent in DW is closer to the tuned consensus of ~ 0.25. This may be a lucky behaviour at 7 TeV, and hence care is needed with extrapolation of D6T to 8, 10, or 14 TeV, but it is clear that the PYTHIA Q^2-ordered parton shower is not yet dead on purely physics grounds. The best tune of this PYTHIA configuration, however, is Prof-Q^2, which in addition to general small improvements, is significantly better than DW or D6T at describing the vector boson $p_\perp$ distribution.
- Pythia8 is generally seen to perform very well, and provides significant improvements over PYTHIA6 for jet shapes and B fragmentation. Tuning focus is accordingly beginning to shift towards Pythia8, also for minimum bias observables not yet considered here.
- AlpGen interacts strongly with tunes on jet shape and vector boson data descriptions. In particular there appears to be little motivation to use AlpGen with the D6T or Perugia 2010 tunes of PYTHIA6. AlpGen+HERWIG also has significant problems with jet shapes in particular, and the indication of this study is that AlpGen+PYTHIA Perugia 2011 is the most performant configuration, closely followed by AlpGen+PYTHIA Z2. Notably, the Perugia 2011 tune of PYTHIA was specifically developed to minimise ME/PS merging artefacts when used with AlpGen.
- Both HERWIG and Herwig++ have problems describing LEP fragmentation data, but Herwig++ is a very significant improvement over its Fortran cousin. The identified hadron rates are in particular much improved, although $K^\pm$ and Σ^0 remain anomalous. However, a known problem with Herwig++ is the poor description of the LEP thrust distribution, which overshoots significantly in the multi-parton region.
- AlpGen seems to have difficulty describing dijet azimuthal decorrelations, even when restricted to the 2/3 parton region of the plot. This is particularly surprising as AlpGen is intended to provide

the multi-parton configurations needed to describe this observable.

- B fragmentation is in general quite poorly described. The best descriptions are by Pythia8, the AUET2B tunes of PYTHIA6, and Herwig++. Other generators and tunes are in decidedly dodgy shape for B-specific predictions at the LHC.

Insofar as it is within the scope of this project to make recommendations for canonical generator and tune choices, we note that the Perugia 2011, AUET2B, and Z2 tunes of PYTHIA6 provide the best data descriptions currently available with that generator and that the Prof-Q^2 tune is the best available configuration using the Q^2-ordered PYTHIA parton shower. We hence recommend these 4 PYTHIA tunes as the current minimal set of PYTHIA tunes for general use at the LHC, particularly once an update of the ATLAS AUET2B tune has fixed the tuning issue with the Z $p_\perp$.

Among the other generators, where there is not such a proliferation of tunes, we note again the apparent performance issues with AlpGen – this is clearly in need of further pursuit. However, to reduce the amount of comparison needed, we note that Perugia 2011 is the only PYTHIA6 tune now optimised for use with AlpGen with avoidance of the worst effects of ME/PS coupling mismatches: hence future studies can quite happily restrict themselves to this AlpGen+PYTHIA configuration. As AlpGen+HERWIG has several problems with jet description, HERWIG itself has serious problems with both light and B fragmentation, and no further tuning of the JIMMY MPI model is envisaged, the HERWIG generator cannot be recommended for future use in any capacity where an alternative exists.

The "new" C++ generators Herwig++, Pythia8, and Sherpa all perform well, with the exception of Sherpa's B fragmentation and the Herwig++ light fragmentation. Pythia8 generally behaves well but some tuning or development may be needed to improve inter-jet observables and the Z $p_\perp$ spectrum. In general, the C++ generators are in good health, and we anticipate further improvements as the focus of tuning studies shifts to them.

21.4 Outlook

This project has put in place a system and a set of classification criteria which have proven useful for summarising and investigating MC generator model and tune predictivity for a variety of QCD phenomena. While we claim no mandate to truly "kill" certain tunes or generators, and wish to emphasise that a poor performance in a single observable type (in particular B fragmentation) certainly does not render that generator useless, the results from these comparisons do provide strong arguments for deprecation of at least several PYTHIA6 tunes and of the Fortran HERWIG generator in general.

It is the nature of a project like this that results are continually being updated, and there are many natural avenues for extension which we wish to pursue, in particular:

- Extra observables, e.g. minimum bias and $E_\perp$ flow, LHC and Tevatron photon physics, LHC W/Z $p_\perp$ data, strangeness data from LHC and RHIC, explicit multijet observables, etc..
- Extra generators and tunes, in particular POWHEG+PYTHIA/HERWIG/Pythia8/Herwig++, MadGraph+PYTHIA/Pythia8, MC@NLO+HERWIG/Herwig++. Comparison between Sherpa with the CTEQ6.6 and CTEQ6L1 PDFs. New Pythia8 and PYTHIA6 tunes from ATLAS.

Greater automation of the data generation will be important, as finding resources (human rather than CPU!) to produce and run combinatoric numbers of generator/tune/PDF/observable combinations has been troublesome. We suggest that this project can make use of the output of the CERN LPCC MCplots system (also Rivet-based) for future extension. We also look forward to a forthcoming major upgrade of the Rivet histogramming system which will greatly simplify the treatment of multi-leg generators for which the n-parton samples must be explicitly merged, e.g. AlpGen, MadGraph, etc.

22. COMPACT ASCII OUTPUT FORMAT FOR HEPMC [59]

Abstract

[59] Contributed by: A. Buckley, L. Garren, G. Hesketh, H. Hoeth, L. Lönnblad, E. Nurse, S. Plätzer, G. Salam, G. Soyez

We discuss the possibility of reducing the footprint of HepMC event files. Different compression options are discussed, and a suggestion for an update of the HepMC ASCII file format is presented.

22.1 Introduction

The HepMC [509] event record has become the de-facto standard for communicating events between event generators and different kinds of analysis programs. HepMC also provides an ASCII-based file format for storing and retrieving events to and from disk, which has also become the standard. This file format is not at all optimized for size, and although disk space today is fairly cheap, there are still problems associated with handling very large files.

A typical minimum-bias 7 TeV LHC HepMC event occupies around 50 kB when written on disk. More interesting events are usually bigger than this and one would typically want to store many events to get anywhere near the statistics collected by any of the LHC experiments; it is clear that such event files will become very large and difficult to handle. Even with standard compression algorithms such as `gzip` and `bzip2`, where these file sizes can be reduced by a factor 3 or more, the problem is still substantial.

One could imagine using a binary output format to reduce the event size. Writing a 4 byte floating point number in an ASCII file typically takes 10-12 characters, so here one could expect to reduce file sizes up to a factor 3. However, standard compression algorithms are rather good at identifying strings of numbers and compressing them, so there is normally not much to be gained by using a compressed binary format compared to a compressed ASCII one. In addition one would lose the advantage of ASCII files that they are (somewhat) readable to the human eye.

Instead the key to reducing file sizes is to remove redundant and unnecessary information stored in the files. This could involve completely reversible operations such as removing the information about the momentum of an intermediate particle, as this can be reconstructed from its decay products. It could also involve irreversible operations such as reducing the precision on the momenta. In the following we describe a number of such operations, which allows us to reduce the file sizes by almost a factor 30.

22.2 The Benchmarking procedure

We started out by generating 1000 non-diffractive QCD events with Pythia 6.425[400] using the AGILe [360] interface. The resulting file size was 48 MB, which can be reduced to 16 MB or 13 MB using `gzip` or `bzip2` respectively. We then investigated several ways of reducing this size.

Removing irrelevant particles The HepMC format contains quite a lot of information about how the event was generated, such as intermediate particles in the hard sub process, which may be generator-dependent (and often unphysical) and is not relevant when comparing to experimental data. In principle one could argue that the only thing that should be written out is final-state stable particles (with HepMC status code 1). However, there are circumstances where information about intermediate unstable hadrons (status code 2) is relevant. The AGILe event generator interface already includes facilities for keeping only particle entries with status code 1 or 2.

Reconstructible information Some information in the HepMC file is redundant in the sense that it can be reconstructed from other information in the file. Here are some examples.

- Both energy, momentum and invariant mass of each particle is written out. Clearly, we can eg. reconstruct the energy given the three-momentum and mass[60].

[60] One could also reconstruct the mass given energy and momentum, but this typically gives large precision problems for small masses.

Format	Status codes	no comp. (MB)	`gzip` (MB)	`bzip2` (MB)
Standard	All	48	16	13
	1 & 2	43	15	13
	1	17	6.0	4.8
Compact	All	18	3.3	2.1
	1 & 2	13	2.9	1.9
	1	4.0	1.9	1.6
Compact binary	1	1.8	1.7	1.7

Table 18: Size of the benchmark file after applying different compression methods.

- The three-momenta of decayed hadrons (status code 2) can be reconstructed from the sum of the momenta of the decay products.
- The mass of a stable particle can be deduced from the particle ID.
- The position of a vertex can be deduced from the previous vertex position and the life-time and momentum direction of the connecting particle.
- Each particle in a HepMC event has a unique *bar code*, which is an otherwise arbitrary integer. No loss of information would result from renumbering the particles, simply inferring their bar code from the order in which they appear in the event.

Precision Clearly, having 8 byte floating point numbers is not very relevant for many of the numbers in an event file. When comparing with experimental data, there is no point in having much larger precision than what is achievable in the experiment, and it makes sense to match the the information in the HepMC file to the precision of the actually measured variables in the experiments.

A possible example is to store masses and transverse momenta as integers in units of 0.1 MeV, azimuthal angles as integers in units of $0.00002\times\pi$, pseudorapidities as integers in units of 0.00001 and vertex positions as integers in units of 0.001 mm.

22.3 Benchmark Results

We have investigated several of the options listed in the previous section, and the resulting file sizes when applied to the benchmark file is presented in table 18. Firstly we see the size reduction using the standard format and simply reducing the number of particles, keeping only those with status code 2 and 1 or only 1. Next we present the same results, but using a compact format which keeps the structure of the HepMC ASCII file but applies all optimizations discussed above. Finally, for reference, we present an aggressively compacted *Binary* format which uses the following optimizations for each particle: stores 1 float for transverse and 1 float for longitudinal momentum, a 3-byte integer for phi, and 1 byte for PDG IDs (rare PDG IDs are written out with 4 full bytes). This format loses the HepMC structure of the event and in some sense this represents the target size, below which it is difficult to go.

It is clear that one does not gain much by using a binary format provided one uses the optimizations presented above together with `bzip2` compression algorithm.

22.4 Outlook

Given the results above, the work to include a more efficient file format for the HepMC has begun. The suggestion is to keep the current structure of the file format, but to add options to exclude all particles except those with status code 1 (or 2). Furthermore options for the representation and precision of momenta and vertex positions will be included as well as options for excluding (simply replacing with a single exclamation mark for easy parsing) information which can be reconstructed. The new format will be included in a forthcoming HepMC version during 2012.

In this report we have not looked carefully at the time it takes to read the different formats. With the default HepMC format this can be many times larger than the time taken for typical particle-level analyses, while for minimal binary formats it is of the same order. We defer detailed study of this question to future work.

References

[1] J. M. Campbell, J. Huston, and W. Stirling Rept.Prog.Phys. **70** (2007) 89, arXiv:hep-ph/0611148 [hep-ph].

[2] S. Agrawal, T. Hahn, and E. Mirabella arXiv:1112.0124 [hep-ph].

[3] G. Belanger, F. Boudjema, J. Fujimoto, T. Ishikawa, T. Kaneko, et al. Phys.Rept. **430** (2006) 117–209, arXiv:hep-ph/0308080 [hep-ph].

[4] C. Berger, Z. Bern, L. Dixon, F. Febres Cordero, D. Forde, et al. Phys.Rev. **D78** (2008) 036003, arXiv:0803.4180 [hep-ph].

[5] S. Badger, B. Biedermann, and P. Uwer Comput.Phys.Commun. **182** (2011) 1674–1692, arXiv:1011.2900 [hep-ph].

[6] P. Mastrolia, G. Ossola, T. Reiter, and F. Tramontano JHEP **1008** (2010) 080, arXiv:1006.0710 [hep-ph].

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

[8] V. Hirschi, R. Frederix, S. Frixione, M. V. Garzelli, F. Maltoni, et al. JHEP **1105** (2011) 044, arXiv:1103.0621 [hep-ph].

[9] J. M. Campbell, R. Ellis, and C. Williams JHEP **1110** (2011) 005, arXiv:1107.5569 [hep-ph].

[10] G. Bevilacqua, M. Czakon, M. Garzelli, A. van Hameren, A. Kardos, et al. arXiv:1110.1499 [hep-ph].

[11] L. Reina and T. Schutzmeier arXiv:1110.4438 [hep-ph].

[12] G. Cullen et al. arXiv:1111.2034 [hep-ph].

[13] S. Alioli, P. Nason, C. Oleari, and E. Re JHEP **1006** (2010) 043, arXiv:1002.2581 [hep-ph].

[14] M. Garzelli, A. Kardos, C. Papadopoulos, and Z. Trocsanyi Europhys.Lett. **96** (2011) 11001, arXiv:1108.0387 [hep-ph].

[15] R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, R. Pittau, et al. Phys.Lett. **B701** (2011) 427–433, arXiv:1104.5613 [hep-ph].

[16] C. Buttar, S. Dittmaier, V. Drollinger, S. Frixione, A. Nikitenko, et al. arXiv:hep-ph/0604120 [hep-ph].

[17] NLO Multileg Working Group Collaboration, Z. Bern et al. arXiv:0803.0494 [hep-ph].

[18] SM and NLO Multileg Working Group Collaboration, J. R. Andersen et al. arXiv:1003.1241 [hep-ph].

[19] Z. Bern et al. arXiv:1112.3940 [hep-ph].

[20] H. Ita, Z. Bern, L. Dixon, F. Cordero, D. Kosower, et al. arXiv:1108.2229 [hep-ph].

[21] F. Campanario, C. Englert, M. Rauch, and D. Zeppenfeld Phys.Lett. **B704** (2011) 515–519, arXiv:1106.4009 [hep-ph].

[22] C. Berger, Z. Bern, L. J. Dixon, F. Cordero, D. Forde, et al. Phys.Rev.Lett. **106** (2011) 092001, arXiv:1009.2338 [hep-ph].

[23] G. Bevilacqua, M. Czakon, A. van Hameren, C. G. Papadopoulos, and M. Worek JHEP **1102** (2011) 083, arXiv:1012.4230 [hep-ph].

[24] LHC Higgs Cross Section Working Group Collaboration, S. Dittmaier et al. arXiv:1101.0593 [hep-ph].

[25] A. Denner, S. Dittmaier, S. Kallweit, and S. Pozzorini Phys.Rev.Lett. **106** (2011) 052001, arXiv:1012.3975 [hep-ph].

[26] I. W. Stewart and F. J. Tackmann arXiv:1107.2117 [hep-ph].

[27] S. Dittmaier, S. Kallweit, and P. Uwer Phys.Rev.Lett. **100** (2008) 062003, arXiv:0710.1577 [hep-ph].

[28] S. Dittmaier, S. Kallweit, and P. Uwer Nucl.Phys. **B826** (2010) 18–70, arXiv:0908.4124 [hep-ph].

[29] J. M. Campbell, R. Ellis, and G. Zanderighi JHEP **0712** (2007) 056, arXiv:0710.1832 [hep-ph].

[30] T. Binoth, T. Gleisberg, S. Karg, N. Kauer, and G. Sanguinetti Phys.Lett. **B683** (2010) 154–159, arXiv:0911.3181 [hep-ph].

[31] F. Campanario, C. Englert, M. Spannowsky, and D. Zeppenfeld Europhys.Lett. **88** (2009) 11001, arXiv:0908.1638 [hep-ph].

[32] F. Campanario, C. Englert, S. Kallweit, M. Spannowsky, and D. Zeppenfeld JHEP **1007** (2010) 076, arXiv:1006.0390 [hep-ph].

[33] J. M. Campbell, R. Ellis, and G. Zanderighi JHEP **0610** (2006) 028, arXiv:hep-ph/0608194 [hep-ph].

[34] M. Ciccolini, A. Denner, and S. Dittmaier Phys.Rev.Lett. **99** (2007) 161803, arXiv:0707.0381 [hep-ph].

[35] M. Ciccolini, A. Denner, and S. Dittmaier Phys.Rev. **D77** (2008) 013002, arXiv:0710.4749 [hep-ph].

[36] J. Andersen, T. Binoth, G. Heinrich, and J. Smillie JHEP **0802** (2008) 057, arXiv:0709.3513 [hep-ph].

[37] A. Bredenstein, K. Hagiwara, and B. Jager Phys.Rev. **D77** (2008) 073004, arXiv:0801.4231 [hep-ph].

[38] A. Lazopoulos, K. Melnikov, and F. Petriello Phys.Rev. **D76** (2007) 014001, arXiv:hep-ph/0703273 [hep-ph].

[39] V. Hankele and D. Zeppenfeld Phys.Lett. **B661** (2008) 103–108, arXiv:0712.3544 [hep-ph].

[40] T. Binoth, G. Ossola, C. Papadopoulos, and R. Pittau JHEP **0806** (2008) 082, arXiv:0804.0350 [hep-ph].

[41] K. Arnold, J. Bellm, G. Bozzi, M. Brieg, F. Campanario, et al. arXiv:1107.4038 [hep-ph].

[42] K. Arnold, M. Bahr, G. Bozzi, F. Campanario, C. Englert, et al. Comput.Phys.Commun. **180** (2009) 1661–1670, arXiv:0811.4559 [hep-ph].

[43] G. Bozzi, F. Campanario, M. Rauch, and D. Zeppenfeld Phys.Rev. **D83** (2011) 114035, arXiv:1103.4613 [hep-ph].

[44] G. Bozzi, F. Campanario, M. Rauch, and D. Zeppenfeld Phys.Rev. **D84** (2011) 074028, arXiv:1107.3149 [hep-ph].

[45] G. Bozzi, F. Campanario, M. Rauch, H. Rzehak, and D. Zeppenfeld Phys.Lett. **B696** (2011) 380–385, arXiv:1011.2206 [hep-ph].

[46] G. Bozzi, F. Campanario, V. Hankele, and D. Zeppenfeld Phys.Rev. **D81** (2010) 094030, arXiv:0911.0438 [hep-ph].

[47] F. Campanario, V. Hankele, C. Oleari, S. Prestel, and D. Zeppenfeld Phys.Rev. **D78** (2008) 094012, arXiv:0809.0790 [hep-ph].

[48] A. Bredenstein, A. Denner, S. Dittmaier, and S. Pozzorini Phys.Rev.Lett. **103** (2009) 012002, arXiv:0905.0110 [hep-ph].

[49] A. Bredenstein, A. Denner, S. Dittmaier, and S. Pozzorini JHEP **1003** (2010) 021, arXiv:1001.4006 [hep-ph].

[50] G. Bevilacqua, M. Czakon, C. Papadopoulos, R. Pittau, and M. Worek JHEP **0909** (2009) 109, arXiv:0907.4723 [hep-ph].

[51] C. Berger, Z. Bern, L. J. Dixon, F. Febres Cordero, D. Forde, et al. Phys.Rev. **D80** (2009) 074036, arXiv:0907.1984 [hep-ph].

[52] R. Ellis, K. Melnikov, and G. Zanderighi JHEP **0904** (2009) 077, arXiv:0901.4101 [hep-ph].

[53] C. Berger, Z. Bern, L. J. Dixon, F. Cordero, D. Forde, et al. Phys.Rev. **D82** (2010) 074002, arXiv:1004.1659 [hep-ph].

[54] G. Bevilacqua, M. Czakon, C. Papadopoulos, and M. Worek Phys.Rev.Lett. **104** (2010) 162002, arXiv:1002.4009 [hep-ph].

[55] G. Bevilacqua, M. Czakon, C. Papadopoulos, and M. Worek Phys.Rev. **D84** (2011) 114017, arXiv:1108.2851 [hep-ph].

[56] T. Melia, K. Melnikov, R. Rontsch, and G. Zanderighi JHEP **1012** (2010) 053, arXiv:1007.5313 [hep-ph].

[57] T. Melia, K. Melnikov, R. Rontsch, and G. Zanderighi arXiv:1104.2327 [hep-ph].

[58] N. Greiner et al. arXiv:1202.6004 [hep-ph].

[59] B. Jager, C. Oleari, and D. Zeppenfeld JHEP **0607** (2006) 015, arXiv:hep-ph/0603177 [hep-ph].

[60] B. Jager, C. Oleari, and D. Zeppenfeld Phys.Rev. **D73** (2006) 113006, arXiv:hep-ph/0604200 [hep-ph].

[61] G. Bozzi, B. Jager, C. Oleari, and D. Zeppenfeld Phys.Rev. **D75** (2007) 073004, arXiv:hep-ph/0701105 [hep-ph].

[62] T. Binoth, N. Greiner, A. Guffanti, J. Reuter, J.-P. Guillet, et al. Phys.Lett. **B685** (2010) 293–296, arXiv:0910.4379 [hep-ph].

[63] N. Greiner, A. Guffanti, T. Reiter, and J. Reuter arXiv:1105.3624 [hep-ph].

[64] J. R. Andersen and J. M. Smillie JHEP **1106** (2011) 010, arXiv:1101.5394 [hep-ph].

[65] J. H. Kuhn, A. Kulesza, S. Pozzorini, and M. Schulze Nucl.Phys. **B797** (2008) 27–77, arXiv:0708.0476 [hep-ph].

[66] W. Hollik, T. Kasprzik, and B. Kniehl Nucl.Phys. **B790** (2008) 138–159, arXiv:0707.2553 [hep-ph].

[67] A. Denner, S. Dittmaier, T. Kasprzik, and A. Muck JHEP **0908** (2009) 075, arXiv:0906.1656 [hep-ph].

[68] A. Denner, S. Dittmaier, T. Kasprzik, and A. Muck JHEP **1106** (2011) 069, arXiv:1103.0914 [hep-ph].

[69] A. Denner, S. Dittmaier, S. Kallweit, and A. Muck arXiv:1112.5142 [hep-ph].

[70] C. Berger, Z. Bern, L. J. Dixon, F. Febres Cordero, D. Forde, et al. Phys.Rev.Lett. **102** (2009) 222001, arXiv:0902.2760 [hep-ph].

[71] K. Melnikov and G. Zanderighi Phys.Rev. **D81** (2010) 074025, arXiv:0910.3671 [hep-ph].

[72] A. Gehrmann-De Ridder, T. Gehrmann, E. Glover, and G. Heinrich JHEP **0712** (2007) 094, arXiv:0711.4711 [hep-ph].

[73] A. Gehrmann-De Ridder, T. Gehrmann, E. Glover, and G. Heinrich Phys.Rev.Lett. **100** (2008) 172001, arXiv:0802.0813 [hep-ph].

[74] S. Weinzierl JHEP **0906** (2009) 041, arXiv:0904.1077 [hep-ph].

[75] O. Brein, A. Djouadi, and R. Harlander Phys.Lett. **B579** (2004) 149–156, arXiv:hep-ph/0307206 [hep-ph].

[76] O. Brein, R. Harlander, M. Wiesemann, and T. Zirke arXiv:1111.0761 [hep-ph].

[77] G. Ferrera, M. Grazzini, and F. Tramontano Phys.Rev.Lett. **107** (2011) 152003, arXiv:1107.1164 [hep-ph].

[78] S. Catani, L. Cieri, D. de Florian, G. Ferrera, and M. Grazzini arXiv:1110.2375 [hep-ph].

[79] A. Gehrmann-De Ridder, T. Gehrmann, and E. Glover JHEP **0509** (2005) 056, arXiv:hep-ph/0505111 [hep-ph]. Erratum added online, 8/18/06.

[80] A. Daleo, A. Gehrmann-De Ridder, T. Gehrmann, and G. Luisoni JHEP **1001** (2010) 118, arXiv:0912.0374 [hep-ph].

[81] T. Gehrmann and E. Glover Phys.Lett. **B676** (2009) 146–151, arXiv:0904.2665 [hep-ph].

[82] T. Gehrmann and P. F. Monni arXiv:1107.4037 [hep-ph].

[83] G. Abelof and A. Gehrmann-De Ridder JHEP **1104** (2011) 063, arXiv:1102.2443 [hep-ph].

[84] R. Boughezal, A. Gehrmann-De Ridder, and M. Ritzmann JHEP **1102** (2011) 098, arXiv:1011.6631 [hep-ph].

[85] G. Heinrich Nucl. Phys. Proc. Suppl. **116** (2003) 368–372, hep-ph/0211144.

[86] C. Anastasiou, K. Melnikov, and F. Petriello Phys. Rev. **D69** (2004) 076010, hep-ph/0311311.

[87] T. Binoth and G. Heinrich Nucl. Phys. **B693** (2004) 134–148, hep-ph/0402265.

[88] M. Czakon Phys.Lett. **B693** (2010) 259–268, arXiv:1005.0274 [hep-ph].

[89] R. Boughezal, K. Melnikov, and F. Petriello arXiv:1111.7041 [hep-ph].

[90] S. Catani and M. Grazzini Phys.Rev.Lett. **98** (2007) 222002, arXiv:hep-ph/0703012 [hep-ph].

[91] P. Bolzoni, G. Somogyi, and Z. Trocsanyi JHEP **1101** (2011) 059, arXiv:1011.1909 [hep-ph].

[92] R. Bonciani, A. Ferroglia, T. Gehrmann, A. Manteuffel, and C. Studerus JHEP **1101** (2011) 102, arXiv:1011.6661 [hep-ph].

[93] G. Abelof and A. G.-D. Ridder arXiv:1112.4736 [hep-ph].

[94] M. Czakon Nucl.Phys. **B849** (2011) 250–295, arXiv:1101.0642 [hep-ph].

[95] M. Czakon and A. Mitov arXiv:1112.5675 [hep-ph].

[96] G. Chachamis PoS **RADCOR2009** (2010) 059.

[97] G. Chachamis, M. Czakon, and D. Eiras JHEP **12** (2008) 003, arXiv:0802.4028 [hep-ph].

[98] A. Gehrmann-De Ridder, E. Glover, and J. Pires arXiv:1112.3613 [hep-ph].

[99] J. Pires and E. Glover Nucl.Phys.Proc.Suppl. **205-206** (2010) 176–181, arXiv:1006.1849 [hep-ph].

[100] L. Garland, T. Gehrmann, E. Glover, A. Koukoutsakis, and E. Remiddi Nucl.Phys. **B627** (2002) 107–188, arXiv:hep-ph/0112081 [hep-ph].

[101] L. Garland, T. Gehrmann, E. Glover, A. Koukoutsakis, and E. Remiddi Nucl.Phys. **B642** (2002) 227–262, arXiv:hep-ph/0206067 [hep-ph].

[102] T. Gehrmann and L. Tancredi arXiv:1112.1531 [hep-ph].

[103] T. Gehrmann, M. Jaquier, E. Glover, and A. Koukoutsakis arXiv:1112.3554 [hep-ph].

[104] V. Yundin. C++ package PJFry, available at the webpage https://github.com/Vayu/PJFry/.

[105] J. Fleischer and T. Riemann Phys.Rev. **D83** (2011) 073004, arXiv:1009.4436 [hep-ph].

[106] J. Fleischer and T. Riemann Phys.Lett. **B701** (2011) 646–653, arXiv:1104.4067 [hep-ph].

[107] J. Fleischer and T. Riemann arXiv:1111.5821 [hep-ph].

[108] J. Fleischer, T. Riemann, and V. Yundin arXiv:1112.0500 [hep-ph].

[109] R. K. Ellis and G. Zanderighi JHEP **02** (2008) 002, arXiv:0712.1851 [hep-ph].

[110] G. J. van Oldenborgh Comput. Phys. Commun. **66** (1991) 1–15.

[111] A. van Hameren Comput.Phys.Commun. **182** (2011) 2427–2438, arXiv:1007.4716 [hep-ph].

[112] G. Passarino and M. Veltman Nucl.Phys. **B160** (1979) 151.

[113] A. I. Davydychev Phys. Lett. **B263** (1991) 107–111.

[114] T. Binoth, J. Guillet, G. Heinrich, E. Pilon, and C. Schubert JHEP **10** (2005) 015, hep-ph/0504267.

[115] A. Denner and S. Dittmaier Nucl.Phys. **B734** (2006) 62–115, arXiv:hep-ph/0509141 [hep-ph].

[116] T. Diakonidis, J. Fleischer, J. Gluza, K. Kajda, T. Riemann, and J. Tausk Phys. Rev. **D80** (2009) 036003, arXiv:0812.2134 [hep-ph].

[117] R. Ellis, K. Melnikov, and G. Zanderighi Phys.Rev. **D80** (2009) 094002, arXiv:0906.1445 [hep-ph].

[118] G. Bevilacqua, M. Czakon, C. Papadopoulos, and M. Worek arXiv:1108.2851 [hep-ph].

[119] R. Ellis, W. T. Giele, Z. Kunszt, and K. Melnikov Nucl.Phys. **B822** (2009) 270–282, arXiv:0806.3467 [hep-ph].

[120] P. Mastrolia, G. Ossola, C. G. Papadopoulos, and R. Pittau JHEP **06** (2008) 030, arXiv:0803.3964 [hep-ph].

[121] G. Ossola, C. G. Papadopoulos, and R. Pittau Nucl.Phys. **B763** (2007) 147–169, arXiv:hep-ph/0609007 [hep-ph].

[122] G. Ossola, C. G. Papadopoulos, and R. Pittau JHEP **07** (2007) 085, arXiv:0704.1271 [hep-ph].

[123] G. Ossola, C. G. Papadopoulos, and R. Pittau JHEP **0805** (2008) 004, arXiv:0802.1876 [hep-ph].

[124] G. Heinrich, G. Ossola, T. Reiter, and F. Tramontano JHEP **1010** (2010) 105, arXiv:1008.2441 [hep-ph].

[125] T. Reiter et al. arXiv:1011.6632 [hep-ph].

[126] P. Nogueira J.Comput.Phys. **105** (1993) 279–289.

[127] J. Vermaseren arXiv:math-ph/0010025 [math-ph].

[128] G. Cullen, M. Koch-Janusz, and T. Reiter Comput.Phys.Commun. **182** (2011) 2368–2387, arXiv:1008.0803 [hep-ph].

[129] T. Reiter Comput.Phys.Commun. **181** (2010) 1301–1331, arXiv:0907.3714 [hep-ph].

[130] T. Binoth, J.-P. Guillet, G. Heinrich, E. Pilon, and T. Reiter Comput.Phys.Commun. **180** (2009) 2317–2330, arXiv:0810.0992 [hep-ph].

[131] G. Cullen, J. Guillet, G. Heinrich, T. Kleinschmidt, E. Pilon, et al. Comput.Phys.Commun. **182** (2011) 2276–2284, arXiv:1101.5595 [hep-ph].

[132] T. Hahn Comput.Phys.Commun. **140** (2001) 418–431, arXiv:hep-ph/0012260 [hep-ph].

[133] T. Hahn and M. Perez-Victoria Comput.Phys.Commun. **118** (1999) 153–165, arXiv:hep-ph/9807565 [hep-ph].

[134] T. Hahn PoS **ACAT2010** (2010) 078, arXiv:1006.2231 [hep-ph].

[135] G. Ossola, C. G. Papadopoulos, and R. Pittau JHEP **0803** (2008) 042, arXiv:0711.3596 [hep-ph].

[136] A. Semenov arXiv:1005.1909 [hep-ph].

[137] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer, et al. arXiv:1108.2040 [hep-ph].

[138] T. Ohl Comput.Phys.Commun. **90** (1995) 340–354, arXiv:hep-ph/9505351 [hep-ph].

[139] J. Vermaseren Comput.Phys.Commun. **83** (1994) 45–58. Axodraw can be obtained from anonymous ftp from ftp.nikhef.nl. It is located in directory pub/form/axodraw. The author's email address is: t68@nikhef.nl.

[140] P. Draggiotis, M. Garzelli, C. Papadopoulos, and R. Pittau JHEP **0904** (2009) 072, arXiv:0903.0356 [hep-ph].

[141] M. Garzelli, I. Malamos, and R. Pittau JHEP **1001** (2010) 040, arXiv:0910.3130 [hep-ph].

[142] M. Garzelli, I. Malamos, and R. Pittau JHEP **1101** (2011) 029, arXiv:1009.4302 [hep-ph].

[143] M. Garzelli and I. Malamos Eur.Phys.J. **C71** (2011) 1605, arXiv:1010.1248 [hep-ph].

[144] G. Cullen et al. arXiv:1111.6534 [hep-ph].

[145] T. Binoth, F. Boudjema, G. Dissertori, A. Lazopoulos, A. Denner, et al. Comput.Phys.Commun. **181** (2010) 1612–1622, arXiv:1001.1307 [hep-ph]. Dedicated to the memory of, and in tribute to, Thomas Binoth, who led the effort to develop this proposal for Les Houches 2009.

[146] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, et al. JHEP **0902** (2009) 007, arXiv:0811.4622 [hep-ph].

[147] F. Krauss, R. Kuhn, and G. Soff JHEP **0202** (2002) 044, arXiv:hep-ph/0109036 [hep-ph].

[148] R. Ellis, W. Stirling, and B. Webber Camb.Monogr.Part.Phys.Nucl.Phys.Cosmol. **8** (1996) 1–435.

[149] J. Jersak, E. Laermann, and P. Zerwas Phys.Rev. **D25** (1982) 1218.

[150] S. Catani, S. Dittmaier, M. H. Seymour, and Z. Trocsanyi Nucl.Phys. **B627** (2002) 189–265, arXiv:hep-ph/0201036 [hep-ph].

[151] R. Ellis and J. Sexton Nucl.Phys. **B269** (1986) 445.

[152] T. Binoth, J. Guillet, and G. Heinrich JHEP **0702** (2007) 013, arXiv:hep-ph/0609054 [hep-ph].

[153] J. van der Bij and E. Glover Nucl.Phys. **B313** (1989) 237.

[154] J. M. Campbell, R. Ellis, F. Maltoni, and S. Willenbrock Phys.Rev. **D67** (2003) 095002, arXiv:hep-ph/0204093 [hep-ph].

[155] G. Gounaris, P. Porfyriadis, and F. Renard Eur.Phys.J. **C9** (1999) 673–686, arXiv:hep-ph/9902230 [hep-ph].

[156] C. Bernicot arXiv:0804.1315 [hep-ph].

[157] J. M. Campbell and R. Ellis Phys.Rev. **D60** (1999) 113006, arXiv:hep-ph/9905386 [hep-ph].

[158] J. M. Campbell, R. Ellis, and C. Williams JHEP **1107** (2011) 018, arXiv:1105.0020 [hep-ph].

[159] T. Gehrmann and N. Greiner JHEP **12** (2010) 050, arXiv:1011.0321 [hep-ph].

[160] S. Actis, P. Mastrolia, and G. Ossola Phys.Lett. **B682** (2010) 419–427, arXiv:0909.1750 [hep-ph].

[161] A. van Hameren, C. Papadopoulos, and R. Pittau JHEP **0909** (2009) 106, arXiv:0903.4665 [hep-ph].

[162] R. Frederix, T. Gehrmann, and N. Greiner JHEP **0809** (2008) 122, arXiv:0808.2128 [hep-ph].

[163] J. Alwall et al. JHEP **09** (2007) 028, arXiv:0706.2334 [hep-ph].

[164] A. Kanaki and C. G. Papadopoulos Comput.Phys.Commun. **132** (2000) 306–315, arXiv:hep-ph/0002082 [hep-ph].

[165] T. Stelzer and W. Long Comput.Phys.Commun. **81** (1994) 357–371, arXiv:hep-ph/9401258 [hep-ph].

[166] S. Becker, D. Goetz, C. Reuschle, C. Schwan, and S. Weinzierl arXiv:1111.1733 [hep-ph].

[167] S. Becker, C. Reuschle, and S. Weinzierl JHEP **12** (2010) 013, arXiv:1010.4187 [hep-ph].

[168] M. Assadsolimani, S. Becker, C. Reuschle, and S. Weinzierl Nucl. Phys. Proc. Suppl. **205-206** (2010) 224–229, arXiv:1006.4609 [hep-ph].

[169] M. Assadsolimani, S. Becker, and S. Weinzierl Phys. Rev. **D81** (2010) 094002, arXiv:0912.1680 [hep-ph].

[170] W. Gong, Z. Nagy, and D. E. Soper Phys. Rev. **D79** (2009) 033005, arXiv:0812.3686 [hep-ph].

[171] C. Anastasiou, S. Beerli, and A. Daleo JHEP **05** (2007) 071, hep-ph/0703282.

[172] Z. Nagy and D. E. Soper Phys. Rev. **D74** (2006) 093006, hep-ph/0610028.

[173] Z. Nagy and D. E. Soper JHEP **09** (2003) 055, hep-ph/0308127.

[174] D. E. Soper Phys. Rev. **D64** (2001) 034018, hep-ph/0103262.

[175] D. E. Soper Phys. Rev. **D62** (2000) 014009, hep-ph/9910292.

[176] D. E. Soper Phys. Rev. Lett. **81** (1998) 2638–2641, hep-ph/9804454.

[177] R. Frederix, S. Frixione, K. Melnikov, and G. Zanderighi JHEP **1011** (2010) 050, arXiv:1008.5313 [hep-ph].

[178] F. Cascioli, P. Maierhofer, and S. Pozzorini arXiv:1111.5206 [hep-ph].

[179] S. Weinzierl and D. A. Kosower Phys. Rev. **D60** (1999) 054028, hep-ph/9901277.

[180] S. Weinzierl Eur. Phys. J. **C71** (2011) 1565, arXiv:1011.6247 [hep-ph].

[181] S. Weinzierl Eur. Phys. J. **C45** (2006) 745–757, hep-ph/0510157.

[182] S. Catani Phys.Lett. **B427** (1998) 161–171, arXiv:hep-ph/9802439 [hep-ph].

[183] S. Weinzierl Phys. Rev. **D84** (2011) 074007, arXiv:1107.5131 [hep-ph].

[184] T. Gehrmann PoS **DIS2010** (2010) 004, arXiv:1007.2107 [hep-ph].

[185] F. Campanario JHEP **1110** (2011) 070, arXiv:1105.0920 [hep-ph].

[186] I. Wolfram Research, *Mathematica Edition: Version 5.2*. Wolfram Research, Inc., 2005.

[187] R. Mertig, M. Bohm, and A. Denner Comput.Phys.Commun. **64** (1991) 345–359.

[188] K. Hagiwara and D. Zeppenfeld Nucl.Phys. **B274** (1986) 1.

[189] K. Hagiwara and D. Zeppenfeld Nucl.Phys. **B313** (1989) 560.

[190] A. Denner and S. Dittmaier Nucl.Phys. **B658** (2003) 175–202, arXiv:hep-ph/0212259 [hep-ph].

[191] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer JHEP **1106** (2011) 128, arXiv:1106.0522 [hep-ph].

[192] H. Murayama, I. Watanabe, and K. Hagiwara, *HELAS: HELicity amplitude subroutines for Feynman diagram evaluations*, 1992. KEK-91-11.

[193] F. Campanario, M. Kubocz, and D. Zeppenfeld Phys.Rev. **D84** (2011) 095025, arXiv:1011.3819 [hep-ph].

[194] A. Denner, S. Dittmaier, M. Roth, and D. Wackeroth Nucl.Phys. **B560** (1999) 33–65, arXiv:hep-ph/9904472 [hep-ph].

[195] C. Oleari and D. Zeppenfeld Phys.Rev. **D69** (2004) 093004, arXiv:hep-ph/0310156 [hep-ph].

[196] F. Campanario, C. Englert, and M. Spannowsky Phys.Rev. **D82** (2010) 054015, arXiv:1006.3090 [hep-ph].

[197] F. Campanario, C. Englert, and M. Spannowsky Phys.Rev. **D83** (2011) 074009, arXiv:1010.1291 [hep-ph].

[198] K. Arnold, T. Figy, B. Jager, and D. Zeppenfeld JHEP **1008** (2010) 088, arXiv:1006.4237 [hep-ph].

[199] M. Dittmar and H. K. Dreiner Phys. Rev. **D55** (1997) 167–172, arXiv:hep-ph/9608317.

[200] ATLAS Collaboration, H. Yang arXiv:1109.2576 [hep-ex].

[201] CMS Collaboration, *Measurement of the WW, WZ and ZZ cross sections at CMS*, 2011. CMS-PAS-EWK-11-010.

[202] R. W. Brown and K. O. Mikaelian Phys. Rev. **D19** (1979) 922.

[203] J. Ohnemus Phys.Rev. **D44** (1991) 1403–1414.

[204] S. Frixione Nucl. Phys. **B410** (1993) 280–324.

[205] L. J. Dixon, Z. Kunszt, and A. Signer Nucl. Phys. **B531** (1998) 3–23, arXiv:hep-ph/9803250.

[206] L. J. Dixon, Z. Kunszt, and A. Signer Phys.Rev. **D60** (1999) 114037, arXiv:hep-ph/9907305 [hep-ph].

[207] M. Spira, A. Djouadi, D. Graudenz, and P. M. Zerwas Nucl. Phys. **B453** (1995) 17–82, arXiv:hep-ph/9504378.

[208] S. Dawson Nucl. Phys. **B359** (1991) 283–300.

[209] R. V. Harlander and W. B. Kilgore Phys. Rev. Lett. **88** (2002) 201801, arXiv:hep-ph/0201206.

[210] C. Anastasiou and K. Melnikov Nucl. Phys. **B646** (2002) 220–256, arXiv:hep-ph/0207004.

[211] V. Ravindran, J. Smith, and W. L. van Neerven Nucl. Phys. **B665** (2003) 325–366, arXiv:hep-ph/0302135.

[212] S. Catani, D. de Florian, and M. Grazzini JHEP **01** (2002) 015, arXiv:hep-ph/0111164.

[213] G. Davatz, G. Dissertori, M. Dittmar, M. Grazzini, and F. Pauss JHEP **05** (2004) 009, arXiv:hep-ph/0402218.

[214] C. Anastasiou, K. Melnikov, and F. Petriello Phys. Rev. Lett. **93** (2004) 262002, arXiv:hep-ph/0409088.

[215] C. Anastasiou, G. Dissertori, and F. Stockli JHEP **09** (2007) 018, arXiv:0707.2373 [hep-ph].

[216] M. Grazzini JHEP **02** (2008) 043, arXiv:0801.3232 [hep-ph].

[217] A. Bredenstein, A. Denner, S. Dittmaier, and M. M. Weber Phys. Rev. **D74** (2006) 013004, arXiv:hep-ph/0604011.

[218] T. Binoth, M. Ciccolini, N. Kauer, and M. Kramer JHEP **03** (2005) 065, arXiv:hep-ph/0503094.

[219] T. Binoth, M. Ciccolini, N. Kauer, and M. Kramer JHEP **0612** (2006) 046, arXiv:hep-ph/0611170 [hep-ph].

[220] G. Chachamis Acta Phys. Polon. **B38** (2007) 3563–3568, arXiv:0710.3035 [hep-ph].

[221] G. Chachamis, M. Czakon, and D. Eiras arXiv:0806.3043 [hep-ph].

[222] M. Caffo, H. Czyz, S. Laporta, and E. Remiddi Nuovo Cim. **A111** (1998) 365–389, arXiv:hep-th/9805118.

[223] R. Boughezal, M. Czakon, and T. Schutzmeier JHEP **09** (2007) 072, arXiv:0707.3090 [hep-ph].

[224] M. Czakon and T. Schutzmeier JHEP **07** (2008) 001, arXiv:0712.2762 [hep-ph].

[225] M. Czakon, A. Mitov, and S. Moch Phys. Lett. **B651** (2007) 147–159, arXiv:0705.1975 [hep-ph].

[226] M. Czakon, A. Mitov, and S. Moch Nucl. Phys. **B798** (2008) 210–250, arXiv:0707.4139 [hep-ph].

[227] S. Laporta Int. J. Mod. Phys. **A15** (2000) 5087–5159, arXiv:hep-ph/0102033.

[228] V. A. Smirnov Phys. Lett. **B460** (1999) 397–404, arXiv:hep-ph/9905323.

[229] J. B. Tausk Phys. Lett. **B469** (1999) 225–234, arXiv:hep-ph/9909506.

[230] G. Chachamis and M. Czakon Publication in preparation .

[231] M. Czakon Comput. Phys. Commun. **175** (2006) 559–571, arXiv:hep-ph/0511200.

[232] S. Moch and P. Uwer Comput. Phys. Commun. **174** (2006) 759–770, arXiv:math-ph/0508008.

[233] H. R. P. Ferguson and D. H. Bailey Math. of Comput. **53** (1989) 649.

[234] M. Czakon Phys. Lett. **B664** (2008) 307–314, arXiv:0803.1400 [hep-ph].

[235] P. N. Brown, G. D. Byrne, and A. C. Hindmarsh SIAM J. Sci. Stat. Comput. (1989) 1038.

[236] S. Catani and M. Seymour Nucl.Phys. **B485** (1997) 291–419, arXiv:hep-ph/9605323 [hep-ph].

[237] W. Giele and E. Glover Phys.Rev. **D46** (1992) 1980–2010.

[238] Z. Kunszt, A. Signer, and Z. Trocsanyi Nucl.Phys. **B420** (1994) 550–564, arXiv:hep-ph/9401294 [hep-ph].

[239] G. Chachamis and M. Czakon Publication in preparation .

[240] G. Somogyi, Z. Trócsányi, and V. Del Duca JHEP **0506** (2005) 024, arXiv:hep-ph/0502226 [hep-ph].

[241] G. Somogyi and Z. Trócsányi arXiv:hep-ph/0609041 [hep-ph].

[242] G. Somogyi, Z. Trócsányi, and V. Del Duca JHEP **0701** (2007) 070, arXiv:hep-ph/0609042 [hep-ph].

[243] G. Somogyi and Z. Trócsányi JHEP **0701** (2007) 052, arXiv:hep-ph/0609043 [hep-ph].

[244] Z. Nagy, G. Somogyi, and Z. Trócsányi arXiv:hep-ph/0702273 [HEP-PH].

[245] A. Kotikov Phys.Lett. **B267** (1991) 123–127.

[246] E. Remiddi Nuovo Cim. **A110** (1997) 1435–1452, arXiv:hep-th/9711188 [hep-th].

[247] U. Aglietti, V. Del Duca, C. Duhr, G. Somogyi, and Z. Trócsányi JHEP **0809** (2008) 107, arXiv:0807.0514 [hep-ph].

[248] V. A. Smirnov Springer Tracts Mod.Phys. **211** (2004) 1–244.

[249] P. Bolzoni, S.-O. Moch, G. Somogyi, and Z. Trócsányi JHEP **0908** (2009) 079, arXiv:0905.4390 [hep-ph].

[250] G. Heinrich Int.J.Mod.Phys. **A23** (2008) 1457–1486, arXiv:0803.4177 [hep-ph].

[251] G. Somogyi and Z. Trócsányi JHEP **0808** (2008) 042, arXiv:0807.0509 [hep-ph].

[252] C. Bogner and S. Weinzierl Comput.Phys.Commun. **178** (2008) 596–610, arXiv:0709.4092 [hep-ph].

[253] A. Smirnov and M. Tentyukov Comput.Phys.Commun. **180** (2009) 735–746, arXiv:0807.4129 [hep-ph].

[254] J. Carter and G. Heinrich Comput.Phys.Commun. **182** (2011) 1566–1581, arXiv:1011.5493 [hep-ph].

[255] H.-L. Lai et al. Phys. Rev. **D82** (2010) 074024.

[256] P. M. Nadolsky et al. Phys. Rev. **D78** (2008) 013004, arXiv:0802.0007 [hep-ph].

[257] J. Pumplin, D. Stump, R. Brock, D. Casey, J. Huston, et al. Phys.Rev. **D65** (2001) 014013.

[258] P. M. Nadolsky and Z. Sullivan eConf **C010630** (2001) P510, arXiv:hep-ph/0110378 [hep-ph].

[259] D. Stump, J. Pumplin, R. Brock, D. Casey, J. Huston, et al. Phys.Rev. **D65** (2001) 014012.

[260] J. Pumplin, D. Stump, J. Huston, H. Lai, P. M. Nadolsky, et al. JHEP **0207** (2002) 012.

[261] NNPDF Collaboration, R. D. Ball et al. Nucl.Phys. **B809** (2009) 1. Erratum: *ibid.*, B816 293 (2009).

[262] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt Eur. Phys. J. **C63** (2009) 189–285, arXiv:0901.0002 [hep-ph].

[263] H1 and ZEUS Collaboration, F. Aaron et al. JHEP **1001** (2010) 109.

[264] BCDMS Collaboration, A. C. Benvenuti et al. Phys. Lett. **B223** (1989) 485.

[265] BCDMS Collaboration, A. C. Benvenuti et al. Phys. Lett. **B237** (1990) 592.

[266] New Muon Collaboration, M. Arneodo et al. Nucl.Phys. **B483** (1997) 3.

[267] J. P. Berge et al. Z. Phys. **C49** (1991) 187.

[268] CCFR/NuTeV Collaboration, U.-K. Yang et al. Phys.Rev.Lett. **86** (2001) 2742.

[269] W. G. Seligman et al. Phys. Rev. Lett. **79** (1997) 1213.

[270] D. A. Mason, *Measurement of the strange - antistrange asymmetry at NLO in QCD from NuTeV dimuon data*. PhD thesis, 2006.

[271] NuTeV Collaboration, M. Goncharov et al. Phys.Rev. **D64** (2001) 112006.

[272] H1 Collaboration, C. Adloff et al. Phys.Lett. **B528** (2002) 199.

[273] H1 Collaboration, A. Aktas et al. Eur.Phys.J. **C40** (2005) 349.

[274] H1 Collaboration, A. Aktas et al. Eur.Phys.J. **C45** (2006) 23.

[275] ZEUS Collaboration, J. Breitweg et al. Eur.Phys.J. **C12** (2000) 35.

[276] ZEUS Collaboration, S. Chekanov et al. Phys.Rev. **D69** (2004) 012004.

[277] G. Moreno, C. Brown, W. Cooper, D. Finley, Y. Hsiung, et al. Phys.Rev. **D43** (1991) 2815.

[278] FNAL E866/NuSea Collaboration, R. Towell et al. Phys.Rev. **D64** (2001) 052002.

[279] NuSea Collaboration, J. Webb et al. Phys.Rev.Lett. (2003) , arXiv:hep-ex/0302019 [hep-ex].

[280] CDF Collaboration, F. Abe et al. Phys.Rev.Lett. **77** (1996) 2616.

[281] CDF Collaboration, D. Acosta et al. Phys.Rev. **D71** (2005) 051104.

[282] D0 Collaboration, V. Abazov et al. Phys.Rev.Lett. **101** (2008) 211801.

[283] D0 Collaboration, V. Abazov et al. Phys.Lett. **B658** (2008) 112.

[284] CDF Collaboration, T. A. Aaltonen et al. Phys.Lett. **B692** (2010) 232.

[285] CDF Collaboration, T. Aaltonen et al. Phys. Rev. **D78** (2008) 052006.

[286] CDF Collaboration, T. Affolder et al. Phys.Rev. **D64** (2001) 032001.

[287] D0 Collaboration, V. M. Abazov et al. Phys. Rev. Lett. **101** (2008) 062001.

[288] D0 Collaboration, B. Abbott et al. Phys.Rev.Lett. **86** (2001) 1707.

[289] D0 Collaboration, V. M. Abazov et al. Phys. Lett. **B693** (2010) 531.

[290] T. Kluge, K. Rabbertz, and M. Wobisch arXiv:hep-ph/0609285.

[291] See http://projects.hepforge.org/fastnlo/form/index.html.

[292] CMS Collaboration, S. Chatrchyan et al. Phys. Rev. Lett. **107** (2011) 132001.

[293] CMS Collaboration, S. Chatrchyan et al. Phys. Lett. **B700** (2011) 187.

[294] ATLAS Collaboration, G. Aad et al. arXiv:1112.6297 [hep-ex].

[295] ATLAS Collaboration, G. Aad et al. arXiv:1109.5141 [hep-ex].

[296] ATLAS Collaboration, G. Aad et al. Phys.Lett. **B701** (2011) 31–49, arXiv:1103.2929 [hep-ex].

[297] CMS Collaboration, *Inclusive W and Z production cross sections*, 2011. CMS-PAS-EWK-11-005.

[298] CMS Collaboration, S. Chatrchyan et al. JHEP **1104** (2011) 050, arXiv:1103.3470 [hep-ex].

[299] CMS Collaboration, S. Chatrchyan et al. Phys. Rev. D 85, **032002** (2012) , arXiv:1110.4973 [hep-ex].

[300] LHCb Collaboration, *Updated measurements of W and Z production at $\sqrt{s}$ = 7 TeV with the LHCb experiment*, 2011. LHCb-CONF-2011-039.

[301] CMS Collaboration, *Study of associated charm production in W final states at sqrt s = 7 TeV*, 2011. CMS-PAS-EWK-11-013.

[302] ATLAS Collaboration, G. Aad et al. Phys.Lett. **B706** (2011) 150–167.

[303] CMS Collaboration, S. Chatrchyan et al. Phys.Rev. **D84** (2011) 052011, arXiv:1108.2044 [hep-ex].

[304] R. D. Ball et al. Phys.Lett. **B707** (2012) 66–71.

[305] S. Lionetti et al. Phys.Lett. **B701** (2011) 346–352, arXiv:1103.2369 [hep-ph].

[306] T. Carli et al. Eur.Phys.J. **C66** (2010) 503–524, arXiv:0911.2985 [hep-ph].

[307] D. d'Enterria and J. Rojo arXiv:1202.1762 [hep-ph].

[308] NNPDF Collaboration, L. Del Debbio et al. JHEP **0503** (2005) 080, arXiv:hep-ph/0501067 [hep-ph].

[309] NNPDF Collaboration, L. Del Debbio et al. JHEP **0703** (2007) 039, arXiv:hep-ph/0701127 [hep-ph].

[310] The NNPDF Collaboration, R. D. Ball et al. Nucl.Phys. **B823** (2009) 195–233, arXiv:0906.1958 [hep-ph].

[311] R. D. Ball et al. Nucl. Phys. **B838** (2010) 136–206.

[312] R. D. Ball, V. Bertone, F. Cerutti, L. Del Debbio, S. Forte, et al. Nucl.Phys. **B849** (2011) 296–363, arXiv:1101.1300 [hep-ph].

[313] The NNPDF Collaboration, R. D. Ball et al. Nucl.Phys. **B855** (2012) 153–221.

[314] R. D. Ball et al. Nucl.Phys. **B855** (2012) 608–638, arXiv:1108.1758 [hep-ph].

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

[316] The NNPDF Collaboration, R. D. Ball et al. Nucl.Phys. **B849** (2011) 112–143, arXiv:1012.0836 [hep-ph].

[317] S. Alekhin et al. Phys.Rev. **D81** (2010) 014032.

[318] H1 and ZEUS Collaboration, *HERAPDF1.5 NNLO (Preliminary)*, 2011. H1prelim-11-042, ZEUS-prel-11-002.

[319] P. Jimenez-Delgado and E. Reya Phys.Rev. **D80** (2009) 114011.

[320] NNPDF Collaboration, R. D. Ball et al. JHEP **1005** (2010) 075, arXiv:0912.2276 [hep-ph].

[321] H1 and ZEUS Collaborations, *Combined Measurement of the Inclusive e^+p Scattering Cross Sections at HERA for Reduced Proton Beam Energy Runs and Determination of Structure Function F_L*, . H1prelim-10-044, ZEUS-prel-10-008.

[322] M. A. G. Aivazis, F. I. Olness, and W. K. Tung Phys. Rev. **D50** (1994) 3085–3101, hep-ph/9312318.

[323] M. A. G. Aivazis, J. C. Collins, F. I. Olness, and W. K. Tung Phys. Rev. **D50** (1994) 3102–3118, hep-ph/9312319.

[324] J. C. Collins Phys. Rev. **D58** (1998) 094002, hep-ph/9806259.

[325] M. Krämer, F. I. Olness, and D. E. Soper Phys. Rev. **D62** (2000) 096007, hep-ph/0003035.

[326] E. Laenen, S. Riemersma, J. Smith, and W. L. van Neerven Nucl. Phys. **B392** (1993) 162–228.

[327] M. Guzzi, P. M. Nadolsky, H.-L. Lai, and C. P. Yuan arXiv:1108.5112 [hep-ph].

[328] S. Moch, J. A. M. Vermaseren, and A. Vogt Phys. Lett. **B606** (2005) 123–129, arXiv:hep-ph/0411112.

[329] R. M. Barnett Phys. Rev. Lett. **36** (1976) 1163–1166.

[330] J. Amundson, F. I. Olness, C. Schmidt, W. K. Tung, and X. Wang. To be published in the proceedings of 6th International Workshop on Deep Inelastic Scattering and QCD (DIS 98), Brussels, Belgium, 4-8 Apr 1998.

[331] S. Kretzer and I. Schienbein Phys. Rev. **D58** (1998) 094035, hep-ph/9805233.

[332] N. Kidonakis and B. D. Pecjak arXiv:1108.6063 [hep-ph].

[333] S. Weinzierl arXiv:1201.4025 [hep-ph].

[334] R. Bonciani and A. Ferroglia arXiv:1201.4382 [hep-ph].

[335] W. Bernreuther et al. Nucl. Phys. **B690** (2004) 81–137, arXiv:hep-ph/0403035 [hep-ph].

[336] K. Melnikov and M. Schulze JHEP **08** (2009) 049, arXiv:0907.3090 [hep-ph].

[337] W. Bernreuther and Z.-G. Si Nucl. Phys. **B837** (2010) 90–121, arXiv:1003.3926 [hep-ph].

[338] S. Biswas, K. Melnikov, and M. Schulze JHEP **1008** (2010) 048, arXiv:1006.0910 [hep-ph].

[339] A. Denner, S. Dittmaier, M. Roth, and L. Wieders Nucl.Phys. **B724** (2005) 247–294, arXiv:hep-ph/0505042 [hep-ph].

[340] M. Jezabek and J. H. Kühn Nucl. Phys. **B314** (1989) 1.

[341] M. Cacciari, G. P. Salam, and G. Soyez JHEP **04** (2008) 063.

[342] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam Phys.Rev.Lett. **100** (2008) 242001, arXiv:0802.2470 [hep-ph].

[343] A. Kharchilava Phys. Lett. **B476** (2000) 73–78, arXiv:hep-ph/9912320 [hep-ph].

[344] A. Buckley et al. Phys.Rept. **504** (2011) 145–233.

[345] G. Gustafson and U. Pettersson Nucl. Phys. **B306** (1988) 746.

[346] L. Lönnblad Comput. Phys. Commun. **71** (1992) 15.

[347] W. T. Giele, D. A. Kosower, and P. Z. Skands Phys.Rev. **D78** (2008) 014026.

[348] T. Sjöstrand, S. Mrenna, and P. Skands Comput. Phys. Commun. **178** (2008) 852–867, arXiv:arXiv:0710.3820 [hep-ph].

[349] W. Giele, D. Kosower, and P. Skands Phys.Rev. **D84** (2011) 054003.

[350] P. Skands and D. Soper. in Les Houches NLO Multileg WG Summary report, arXiv:0803.0494 [hep-ph].

[351] J. J. Lopez-Villarejo and P. Skands JHEP **1111** (2011) 1, arXiv:1109.3608 [hep-ph].

[352] J.-C. Winter and F. Krauss JHEP **0807** (2008) 040.

[353] P. Z. Skands and S. Weinzierl Phys.Rev. **D79** (2009) 074021.

[354] M. Bengtsson and T. Sjöstrand Nucl.Phys. **B289** (1987) 810.

[355] G. Marchesini and B. R. Webber Nucl. Phys. **B238** (1984) 1.

[356] R. Corke and T. Sjöstrand Eur.Phys.J. **C69** (2010) 1.

[357] B. Andersson, G. Gustafson, and C. Sjögren Nucl.Phys. **B380** (1992) 391.

[358] A. G.-D. Ridder, M. Ritzmann, and P. Skands Phys.Rev. **D85** (2012) 014013, arXiv:1108.6172 [hep-ph].

[359] M. Seymour Z.Phys. **C63** (1994) 99.

[360] A. Buckley, J. Butterworth, L. Lönnblad, H. Hoeth, J. Monk, et al. arXiv:1003.0694 [hep-ph].

[361] M. Cacciari and G. P. Salam Phys.Lett. **B641** (2006) 57–61, arXiv:hep-ph/0512210 [hep-ph].

[362] S. Plätzer and S. Gieseke JHEP **1101** (2011) 024.

[363] The Tevatron New Phenomena and Higgs Working Group Collaboration, D. Benjamin et al. arXiv:arxiv:1108.3331 [hep-ex].

[364] CMS Collaboration, K. Sung et al. arXiv:arXiv:1109.2457 [hep-ex].

[365] ATLAS Collaboration, G. Aad et al. arXiv:arXiv:1112.2577 [hep-ex].

[366] B. Mellado, W. Quayle, and S. L. Wu Phys. Lett. B **611** (2005) 60–65, arXiv:hep-ph/0406095 [hep-ph].

[367] B. Mellado, W. Quayle, and S. L. Wu Phys. Rev. D **76** (2007) 093007, arXiv:arXiv:0708.2507 [hep-ph].

[368] A. Djouadi, M. Spira, and P. M. Zerwas Phys. Lett. B **264** (1991) 440–446.

[369] A. Pak, M. Rogal, and M. Steinhauser JHEP **02** (2010) 025, arXiv:arXiv:0911.4662 [hep-ph].

[370] R. V. Harlander, H. Mantler, S. Marzani, and K. J. Ozeren Eur. Phys. J. C **66** (2010) 359–372, arXiv:arXiv:0912.2104 [hep-ph].

[371] C. Anastasiou, K. Melnikov, and F. Petriello Nucl.Phys. **B724** (2005) 197–246, arXiv:hep-ph/0501130 [hep-ph].

[372] J. M. Campbell, R. Ellis, and C. Williams Phys.Rev. **D81** (2010) 074023, arXiv:1001.4495 [hep-ph].

[373] C. Anastasiou, G. Dissertori, M. Grazzini, F. Stöckli, and B. R. Webber JHEP **08** (2009) 099, arXiv:arXiv:0905.3529 [hep-ph].

[374] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn Phys. Rev. D **81** (2010) 094035, arXiv:arXiv:0910.0467 [hep-ph].

[375] C. F. Berger, C. Marcantonini, I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn JHEP **04** (2011) 092, arXiv:arXiv:1012.4480 [hep-ph].

[376] LHC Higgs Cross Section Working Group Collaboration, S. Dittmaier et al. arXiv:arXiv:1201.3084 [hep-ph]. 275 pages, 136 figures, to be submitted to CERN Report. Working Group web page: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections.

[377] D0 Collaboration, V. M. Abazov et al. Phys. Rev. **D80** (2009) 111107.

[378] M. Guzzi et al. arXiv:arXiv:1101.0561 [hep-ph] [hep-ph].

[379] CMS Collaboration, S. Chatrchyan et al. Phys. Lett. **B704** (2011) 123.

[380] CMS Collaboration, V. Khachatryan et al. Phys. Rev. Lett. **106** (2011) 201804.

[381] ATLAS Collaboration, G. Aad et al. New J. Phys. **13** (2011) 053044.

[382] ATLAS Collaboration, G. Aad et al. Eur. Phys. J. **C71** (2011) 1512.

[383] S. D. Ellis, Z. Kunszt, and D. E. Soper Phys. Rev. Lett. **69** (1992) 1496.

[384] Z. Nagy Phys. Rev. Lett. **88** (2002) 122003.

[385] Z. Nagy Phys. Rev. **D68** (2003) 094002, arXiv:hep-ph/0307268.

[386] S. Alioli, K. Hamilton, P. Nason, C. Oleari, and E. Re JHEP **1104** (2011) 081, arXiv:1012.3380 [hep-ph].

[387] N. Kidonakis and J. F. Owens Phys. Rev. **D63** (2001) 054019.

[388] J. Gao, Z. Liang, P. Nadolsky, D. Soper, C.-P. Yuan, et al. in preparation.

[389] G. C. Blazey et al. arXiv:hep-ex/0005012.

[390] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber Nucl. Phys. **B406** (1993) 187.

[391] M. Wobisch and T. Wengler arXiv:hep-ph/9907280.

[392] G. P. Salam Eur. Phys. J. **C67** (2010) 637.

[393] T. Hahn Comput. Phys. Commun. **168** (2005) 78.

[394] D0 Collaboration, V. Abazov et al. Phys.Rev.Lett. **103** (2009) 191803.

[395] D. Stump, J. Huston, J. Pumplin, W.-K. Tung, H.-L. Lai, S. Kuhlmann, and J. F. Owens JHEP **0310** (2003) 046.

[396] V. Hirschi arXiv:1111.2708 [hep-ph].

[397] S. Frixione, Z. Kunszt, and A. Signer Nucl.Phys. **B467** (1996) 399–442, arXiv:hep-ph/9512328 [hep-ph].

[398] R. Frederix, S. Frixione, F. Maltoni, and T. Stelzer JHEP **0910** (2009) 003, arXiv:0908.4272 [hep-ph].

[399] G. Corcella, I. Knowles, G. Marchesini, S. Moretti, K. Odagiri, et al. JHEP **0101** (2001) 010, arXiv:hep-ph/0011363 [hep-ph].

[400] T. Sjöstrand, S. Mrenna, and P. Skands JHEP **05** (2006) 026, arXiv:hep-ph/0603175.

[401] S. Frixione and B. R. Webber JHEP **0206** (2002) 029, arXiv:hep-ph/0204244 [hep-ph].

[402] R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, R. Pittau, and P. Torrielli arXiv:1110.4738 [hep-ph].

[403] R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, R. Pittau, et al. JHEP **1109** (2011) 061, arXiv:1106.6019 [hep-ph].

[404] R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, R. Pittau, et al. arXiv:1110.5502 [hep-ph].

[405] L. H. C. S. W. Group arXiv:1201.3084 [hep-ph].

[406] CDF Collaboration, T. Aaltonen et al. Phys.Rev.Lett. **106** (2011) 171801, arXiv:1104.0699 [hep-ex].

[407] D0 Collaboration, V. Abazov et al. Phys.Rev.Lett. **107** (2011) 011804, arXiv:1106.1921 [hep-ex].

[408] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa JHEP **0307** (2003) 001, arXiv:hep-ph/0206293 [hep-ph].

[409] CDF Collaboration http://www-cdf.fnal.gov/physics/ewk/2011/wjj/7_3.html.

[410] ATLAS Collaboration, G. Aad et al. JHEP **1109** (2011) 053, arXiv:1107.1641 [hep-ex].

[411] ATLAS Collaboration arXiv:1107.2092 [hep-ex].

[412] CMS Collaboration, S. Chatrchyan et al. Phys.Lett. **B702** (2011) 336–354, arXiv:1106.0647 [hep-ex].

[413] CMS Collaboration, *Cross section measurement for simultaneous production of a central and a forward jet in proton-proton collisions at sqrt(s)=7 TeV*, 2011. CMS-FWD-10-006.

[414] CMS Collaboration, V. Khachatryan et al. Phys.Rev.Lett. **106** (2011) 122003, arXiv:1101.5029 [hep-ex].

[415] CMS Collaboration, *Measurement of inclusive and exclusive dijet production ratio at large rapidity intervals at* $\sqrt{s} = 7$ *TeV*, 2011. CMS-FWD-10-014.

[416] P. Nason JHEP **0411** (2004) 040, arXiv:hep-ph/0409146 [hep-ph].

[417] S. Frixione, P. Nason, and C. Oleari JHEP **0711** (2007) 070, arXiv:0709.2092 [hep-ph].

[418] J. R. Andersen and J. M. Smillie JHEP **1001** (2010) 039, arXiv:0908.2786 [hep-ph].

[419] J. R. Andersen and J. M. Smillie Phys.Rev. **D81** (2010) 114021, arXiv:0910.5113 [hep-ph].

[420] S. Frixione and G. Ridolfi Nucl. Phys. **B507** (1997) 315, hep-ph/9707345.

[421] A. Banfi and M. Dasgupta JHEP **01** (2004) 027, arXiv:hep-ph/0312108.

[422] S. Alioli, J. R. Andersen, C. Oleari, E. Re, and J. M. Smillie. In preparation.

[423] Z. Bern , private communication, February, 2011.

[424] S. Hoeche, F. Krauss, M. Schonherr, and F. Siegert arXiv:1201.5882 [hep-ph].

[425] T. Gleisberg, S. Hoeche, F. Krauss, A. Schalicke, S. Schumann, et al. JHEP **0402** (2004) 056, arXiv:hep-ph/0311263 [hep-ph].

[426] S. Hoeche, F. Krauss, S. Schumann, and F. Siegert JHEP **0905** (2009) 053, arXiv:0903.1219 [hep-ph].

[427] S. Hoeche, S. Schumann, and F. Siegert Phys. Rev. **D81** (2010) 034026, arXiv hep-ph/0912.3501.

[428] T. Carli, T. Gehrmann, and S. Hoeche Eur.Phys.J. **C67** (2010) 73–97, arXiv:0912.3715 [hep-ph].

[429] S. Catani, F. Krauss, R. Kuhn, and B. Webber JHEP **0111** (2001) 063, arXiv:hep-ph/0109231 [hep-ph].

[430] F. Krauss JHEP **0208** (2002) 015, arXiv:hep-ph/0205283 [hep-ph].

[431] J. R. Andersen, L. Lönnblad, and J. M. Smillie JHEP **1107** (2011) 110, arXiv:1104.1316 [hep-ph].

[432] http://hepforge.cedar.ac.ul/lhapdf/.

[433] G. P. Salam and G. Soyez JHEP **0705** (2007) 086, arXiv:0704.0292 [hep-ph].

[434] P.-A. Delsart, K. L. Geerlings, J. Huston, B. T. Martin, and C. K. Vermilion arXiv:1201.3617 [hep-ex]. 24 pages, 3 figures.

[435] M. Cacciari, G. P. Salam, and G. Soyez arXiv:1111.6097 [hep-ph]. To appear in *Eur.Phys.J.C*. See also http://www.fastjet.fr.

[436] ATLAS Collaboration, G. Aad et al. arXiv:1201.1276 [hep-ex].

[437] M. Rubin, G. P. Salam, and S. Sapeta JHEP **1009** (2010) 084, arXiv:1006.2144 [hep-ph].

[438] J. R. Andersen, T. Hapola, and J. M. Smillie (In preparation) .

[439] J. Alwall, S. Hoche, F. Krauss, N. Lavesson, L. Lonnblad, et al. Eur.Phys.J. **C53** (2008) 473–500, arXiv:0706.2569 [hep-ph].

[440] S. Alioli, P. Nason, C. Oleari, and E. Re JHEP **1101** (2011) 095, arXiv:1009.5594 [hep-ph].

[441] L. Lonnblad and S. Prestel arXiv:1109.4829 [hep-ph].

[442] K. Hamilton and P. Nason JHEP **1006** (2010) 039, arXiv:1004.1764 [hep-ph].

[443] S. Hoche, F. Krauss, M. Schonherr, and F. Siegert JHEP **1108** (2011) 123, arXiv:1009.1127 [hep-ph].

[444] M. L. Mangano, M. Moretti, and R. Pittau Nucl.Phys. **B632** (2002) 343–362, arXiv:hep-ph/0108069 [hep-ph].

[445] L. Lonnblad JHEP **0205** (2002) 046, arXiv:hep-ph/0112284 [hep-ph].

[446] F. Maltoni and T. Stelzer JHEP **0302** (2003) 027, arXiv:hep-ph/0208156 [hep-ph].

[447] T. Sjöstrand and P. Z. Skands Eur. Phys. J. C **39** (2005) 129–154, arXiv:hep-ph/0408302.

[448] T. Gleisberg and F. Krauss Eur.Phys.J. **C53** (2008) 501–523, arXiv:0709.2881 [hep-ph].

[449] J. R. Andersen, T. Hapola, and J. M. Smillie (In preparation) .

[450] R. Field arXiv:1010.3558 [hep-ph].

[451] A. Buckley, H. Hoeth, H. Lacker, H. Schulz, and J. E. von Seggern Eur. Phys. J. C **65** (2010) 331–357, arXiv:arXiv:0907.2973 [hep-ph].

[452] J. Alwall, A. Ballestrero, P. Bartalini, S. Belov, E. Boos, et al. Comput.Phys.Commun. **176** (2007) 300–304, arXiv:hep-ph/0609017 [hep-ph].

[453] G. Gustafson Phys.Lett. **B175** (1986) 453.

[454] T. Sjöstrand and P. Z. Skands JHEP **03** (2004) 053, arXiv:hep-ph/0402078.

[455] R. Corke and T. Sjostrand JHEP **1103** (2011) 032, arXiv:1011.1759 [hep-ph].

[456] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand Phys.Rept. **97** (1983) 31–145.

[457] M. R. Whalley, D. Bourilkov, and R. C. Group arXiv:hep-ph/0508110.

[458] ATLAS Collaboration, G. Aad et al. (2011) , https://cdsweb.cern.ch/record/1400677.

[459] T. Gleisberg and S. Hoeche JHEP **0812** (2008) 039, arXiv:0808.3674 [hep-ph].

[460] Z. Nagy and D. E. Soper arXiv:hep-ph/0601021 [hep-ph].

[461] S. Schumann and F. Krauss JHEP **0803** (2008) 038, arXiv:0709.1027 [hep-ph].

[462] B. Webber Nucl.Phys. **B238** (1984) 492.

[463] J.-C. Winter, F. Krauss, and G. Soff Eur.Phys.J. **C36** (2004) 381–395, arXiv:hep-ph/0311085 [hep-ph].

[464] D. Yennie, S. C. Frautschi, and H. Suura Annals Phys. **13** (1961) 379–452.

[465] M. Schonherr and F. Krauss JHEP **0812** (2008) 018, arXiv:0810.5071 [hep-ph].

[466] A. Schalicke and F. Krauss JHEP **0507** (2005) 018, arXiv:hep-ph/0503281 [hep-ph].

[467] N. Lavesson and L. Lonnblad JHEP **0507** (2005) 054, arXiv:hep-ph/0503293 [hep-ph].

[468] M. Botje, J. Butterworth, A. Cooper-Sarkar, A. de Roeck, J. Feltesse, et al. arXiv:1101.0538 [hep-ph].

[469] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn Phys. Rev. Lett. **106** (2011) 032001, arXiv:arXiv:1005.4060 [hep-ph].

[470] R. Corke and T. Sjostrand JHEP **1105** (2011) 009, arXiv:1101.5953 [hep-ph].

[471] S. Frixione Phys.Lett. **B429** (1998) 369–374, arXiv hep-ph/9801442.

[472] S. Catani et al. JHEP **05** (2002) 028.

[473] P. Aurenche et al. Phys. Rev. **D73** (2006) 094007.

[474] T. Binoth, J. P. Guillet, E. Pilon, and M. Werlen Eur. Phys. J. **C16** (2000) 311–330, arXiv hep-ph/9911340.

[475] M. Cacciari, G. P. Salam, and G. Soyez JHEP **04** (2008) 005, arXiv hep-ph/0802.1188.

[476] M. Cacciari, G. P. Salam, and S. Sapeta JHEP **04** (2010) 065.

[477] ATLAS Phys. Rev. **D83** (2011) 052005.

[478] M. Tripiana. Private communication, 2011.

[479] L. Bourhis, M. Fontannaz, and J. Guillet Eur.Phys.J. **C2** (1998) 529–537, arXiv hep-ph/9704447.

[480] R. Blair, B. Brelier, F. Bucci, S. Chekanov, M. Stockton, and M. Tripiana CERN-OPEN-2011-041, CERN, Geneva, Sep, 2011.

[481] E. Pilon. Private communication, 2010.

[482] M. Hance. Private communication, 2011.

[483] CDF Collaboration CollaborationPhys. Rev. Lett. **107** (Sep, 2011) 102003. http://link.aps.org/doi/10.1103/PhysRevLett.107.102003.

[484] L. D'Errico and P. Richardson arXiv hep-ph/1106.3939.

[485] M. Cacciari and G. P. Salam Phys.Lett. **B659** (2008) 119–126, arXiv:0707.1378 [hep-ph].

[486] M. Cacciari, J. Rojo, G. P. Salam, and G. Soyez Eur.Phys.J. **C71** (2011) 1539, arXiv:1010.1759 [hep-ph].

[487] P. Jacobs (STAR collaboration) arXiv:1012.2406 [nucl-ex].

[488] G. Corcella et al. arXiv:hep-ph/0210213.

[489] S. Gieseke, D. Grellscheid, K. Hamilton, A. Papaefstathiou, S. Platzer, et al. arXiv:1102.1672 [hep-ph].

[490] For the ATLAS collaboration, the jet resolution can be extracted from the first plot from http://twiki.cern.ch/twiki/pub/AtlasPublic/JetResolutionPreliminaryResults, as extracted on 13 April 2011.

[491] For the CMS collaboration, see *e.g.* CMS PAS JME-10-014.

[492] M. Cacciari, G. P. Salam, and S. Sapeta JHEP **1004** (2010) 065, arXiv:0912.4926 [hep-ph].

[493] ATLAS Collaboration, G. Aad et al. Phys. Rev. D **83** (2011) 112001, arXiv:1012.0791 [hep-ex].

[494] ATLAS Collaboration, G. Aad et al. Phys.Rev. **D83** (2011) 052003.

[495] ATLAS Collaboration, G. Aad et al. Phys.Rev.Lett. **106** (2011) 172002, arXiv:1102.2696 [hep-ex].

[496] CMS Collaboration, V. Khachatryan et al. Phys.Lett. **B699** (2011) 48–67, arXiv:1102.0068 [hep-ex].

[497] D0 Collaboration, V. M. Abazov et al. Phys.Rev.Lett. **106** (2011) 122001, arXiv:1010.0262 [hep-ex].

[498] DELPHI Collaboration, P. Abreu et al. Z. Phys. C **73** (1996) 11–60.

[499] Particle Data Group Collaboration, C. Amsler et al. Phys.Lett. **B667** (2008) 1–1340.

[500] DELPHI Collaboration, *A Study of the b-Quark Fragmentation Function with the DELPHI Detector at LEP I*, 2002. DELPHI 2002-069-CONF-603 (ICHEP 2002).

[501] R. Field, *Min-bias and the underlying event at the Tevatron and the LHC*, A talk presented at the Fermilab MC Tuning Workshop, Fermilab, Oct, 2002.

[502] The ATLAS Collaboration, *Charged particle multiplicities in inelastic pp interactions with ATLAS and the ATLAS Minimum Bias Tune 1*, 2010. ATLAS-CONF-2010-031.

[503] The ATLAS Collaboration, *ATLAS tunes of PYTHIA 6 and PYTHIA 8 for MC11*, 2010. ATL-PHYS-PUB-2011-009.

[504] P. Skands arXiv:0905.3418v1 [hep-ph].

[505] M. L. Mangano and M. Moretti and F. Piccinini and R. Pittau and A. D. Polosa JHEP **0307** (2003) 001. doi:10.1088/1126-6708/2003/07/001.

[506] B. Cooper, J. Katzy, M. Mangano, A. Messina, L. Mijovic, et al. arXiv:1109.5295 [hep-ph].

[507] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour Z. Phys. C **72** (1996) 637–646, arXiv:hep-ph/9601371.

[508] The ATLAS Collaboration, *New ATLAS event generator tunes to 2010 data*, 2011. ATL-PHYS-PUB-2011-008.

[509] M. Dobbs and J. B. Hansen Comput.Phys.Commun. **134** (2001) 41–46.