Monte Carlo Tools for SM physics at the LHC

Mauro Moretti

Dept. of Phys. and INFN Ferrara

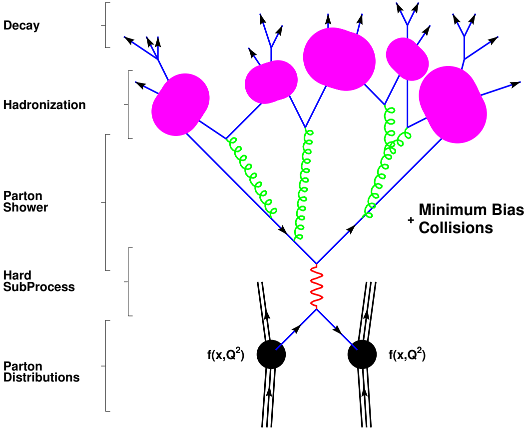
GGI 2009, Firenze, 26-30 October, 2009

Outline

- Brief introduction to PS MC
- LO matrix element generators and matching with QCD Parton Shower
- QCD NLO calculations and implementation in MC programs
- Matching NLO calculations with Parton Shower

Apologize for omitting several important contributions

evolution of a collision at hadron colliders

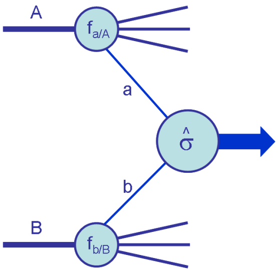


Factorization theorem (and assumption)

$$\sigma_{AB} = \sum_{ab} \int dx_a \quad dx_b \quad f_{a/A}(x_a, \mu_F^2)\, f_{b/B}(x_b, \mu_F^2)$$

$$\times \quad [\hat{\sigma}_{ab}(\mu_F^2, \mu_R^2, x_a p_a, x_b p_b)$$

$$+ \quad \mathcal{O}\left(\frac{\Lambda_{QCD}^n}{Q^n}\right)$$



Parton Shower MC event generators

- General-purpose tools
- Resum LL and some NLL QCD contributions ($\alpha_S^n L^{2n}$, $\alpha_S^n L^{2n-1}$)
- They describe the complete history of the hadron-hadron interaction. PS are the only tools where adronic final state is fully reconstructed and can be compared with data
- Only the hard subprocess is process dependent
- They provide an **exclusive** description of the events: complete information related to every particle is recorded
- Unweighted events are produced $\Rightarrow$ events are distributed in phase space as in the real experiment (provided the underlying theory is correct)
- Key theoretical ingredient: parton shower technique to generate higher order QCD corrections starting from a simple ($2 \to 1$ or $2 \to 2$) hard scattering

Parton Shower MC event generators

- PYTHIA
- HERWIG
- SHERPA
- ARIADNE

Continuos evolution of technical details (ordering variables, underlying events modelling, multiple interaction, C++ versions), but no major departure from original formulation.

At present no calculation can be compared with data without confronting with PS

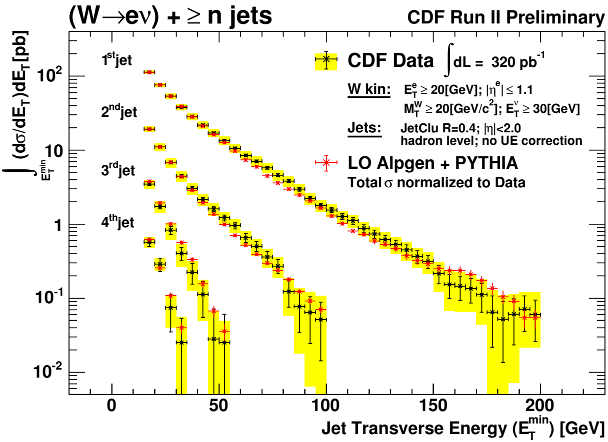
A numerical Monte Carlo solution of the DGLAP evolution equations, which allows the calculation of QCD (and also QED) higher order radiative corrections in the region of collinear parton branching and/or soft gluon emission. Leading logarithms automatically resummed

The subsequent parton emission is a stochastic Markov process in which successive values of the evolution variable Q, the momentum fraction z and the azimuthal angle ϕ are generated (allowing for kinematics reconstruction)

Motivations for LO and NLO matrix element event generators

- Top pair production, Drell-Yan, many gauge bosons production in association with hard, light jets $\Rightarrow$ crucial for SM studies
- Backgrounds to BSM searches $\Rightarrow$ large number of hard jets and leptons, missing E_T
  1. gluinos production and decay: Jets ($n \geq 4$) and missing E_T
  2. excited top states (composite models, kaluza klein excitations) production and decay: $t\bar{t}WWqq$ final states, four leptons four jets and missing E_T or lower number of leptons and larger number of jets
- Understanding as accurate as possible of associate production of a large number of jets, heavy quarks and EWK gauge bosons required

validation against Tevatron data



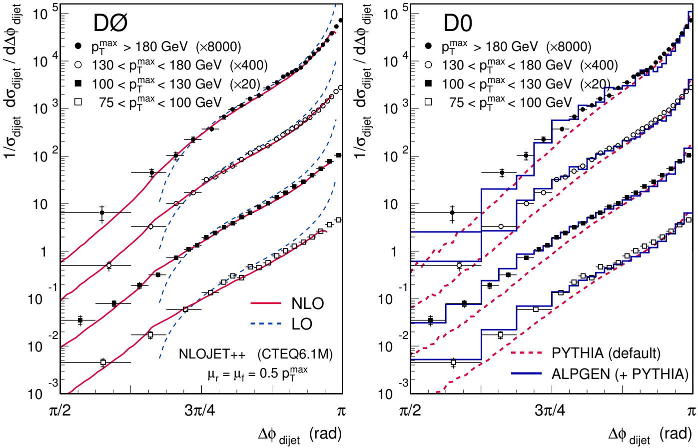


Figure: $\Delta\phi$ among the two leading jets (ordered in p_T), $\mathrm{jet} + \mathrm{jet} + X$ final state.

validation against Tevatron data

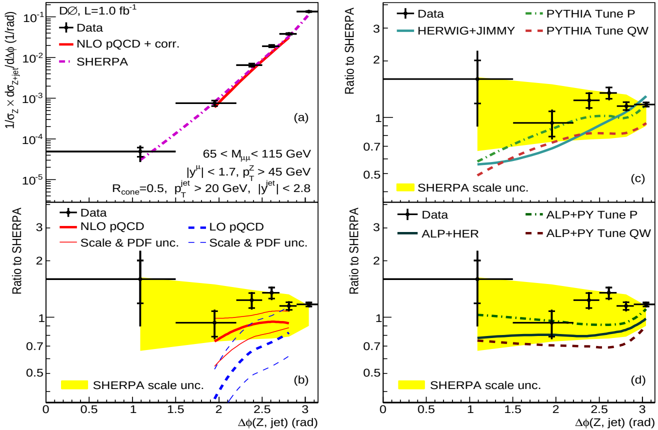


Figure: The measured cross section in bins of $\Delta\phi(Z, \mathrm{jet})$ for $Z/\gamma^* + \mathrm{jet} + X$ events for $p_T^Z > 25$ GeV. The distribution is shown in (a) and compared to various theoretical predictions (normalized to SHERPA one) in (b), (c) and (d).

validation against Tevatron data

This likely would fit better in the conclusions, but agreement with data provides now, perhaps, the best motivation

- LO matrix element generators miss the *total cross section*, however

  1. provides a relatively good description of distributions
  2. jet multiplicity is also reasonably described

  Notice that both the above statements hold true throught several order of magnitudes. Discrepancies still remain, however this comparisons has just began and the agreement would certainly benefit from a dedicated effort to tune these tools (provided man power is avaliable).

- NLO matrix element provides a good description of both normalization and shapes (on a more limited range of observables and jet multiplicity).

SM signals

Avaliable for all the most relevant standard model signals:

1. LO + multi-jets interfaced to PS
2. NLO MC

- N-jets
- W^*, Z^* NNLO; NLO$_{QED}$ (+ QED PS); soft gluon resummation NLL, NNLL;
- $\bar{t}t$ (NLO large); soft gluon resummation
- H (gluon fusion) NLO; NNLO soft gluon resummation NLL; N.B. *everything (in MC)* in $m_t \to \infty$ limit, situation much less satisfactory for $m_H \geq 2m_t$
- H (Weak Bosons Fusion) (no NLO matched with PS)

Many, but not all, relevant backgrounds known to NLO, strong progresses in this field however

LO matrix element event generators

- Many (hard) jets final states are not well described by traditional Parton Shower event generators (e.g. `HERWIG` and `PYTHIA`) and matrix element generators are needed
- example: the statistical significance of the $t\bar{t}H(\to b\bar{b})$ channel at LHC has been lowered after complete simulations with matrix element event generators instead of plain Pyhtia
- Now we have the technology for the calculation of exact Leading Order (LO) multiparton processes of essentially arbitrary complexity

available LO generators

- ACERMC, ALPGEN, CompHEP, GRACE, MADEVENT, HELAC/PHEGAS, NJETS, PHANTOM, SHERPA, VECBOS, OMEGA, WIZARD
- some of them can deal with every SM final state, while others are designed for particular channels
- recent activity by few teams (essentially at present ALPGEN, ARIADNE HELAC, MADEVENT and SHERPA) in combining consistently matrix element predictions with Parton Shower, in order to exploit at the same time the positive features of matrix element and parton shower description

- Usually matrix element event generators are used to produce parton-level unweighted events which are then used as input for the shower evolution and hadronization given by a parton shower programme
- Generating final states with QCD partons, we need parton level cuts, and the final results will depend on these unphysical cuts and will DIVERGE in the soft/collinear limit IRRESPECTIVELY of the cuts applied at analysis level.
- Moreover if we put together, after showering, samples obtained with different parton-level multiplicities we meet the "double counting" problem: the same jet multiplicity can be obtained from different parton-level multiplicities. We need a resolution cutoff to separate the region covered by the matrix element and the one covered by the parton shower
- the problem has been studied for e^+e^- collisions and proved that the dependence on the cutoff can be shifted at NLL level by using matrix elements reweighted with Sudakov form factors (giving the probability of no further emissions) and vetoed parton shower (algorithm known as CKKW procedure)

CKKW for hadronic collisions

- A recipe has been proposed few years ago by F. Krauss, but no formal proof of NLL accuracy exists up to now.
- Indipendent proposal by M.L. Mangano, based on the idea of parton-jet matching (so called MLM prescription). Events are rejected *after the shower* if the PS produces jets "harder" than the ME ones. Sudakov reweighting is not analytic but induced by the veto procedure. The "interplay" of the ME with PS is kept minimal (no need to modify the shower).
- Extension of the CKKW algorithm to the dipole cascade model of parton evolution implemented in `ARIADNE` (Lavesson and Lonnblad)
- Recent activity has been devoted to the comparison between different implementations of CKKW matching in order to quantify the differences and understand their origin

S. Höche et al.

comparison at LHC for W+jets

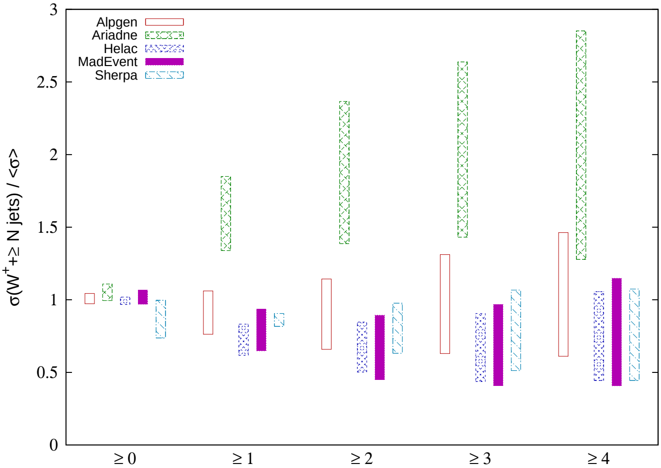


Figure: Range of variation for the LHC cross-section rates of the five codes, normalized to the average value of the default settings for all codes in each multiplicity bin.

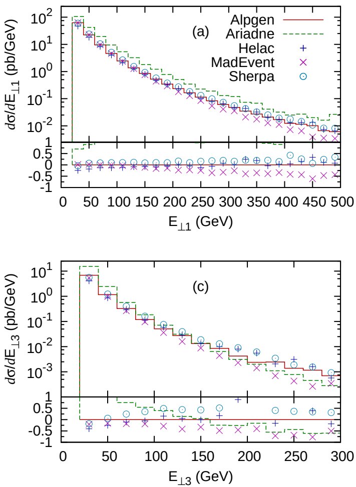


Figure: Inclusive $E_{\perp}$ spectra of the leading 4 jets at the LHC (pb/GeV). In all cases the full line gives the ALPGEN results, the dashed line gives the ARIADNE results and the "+", "x" and "o" points give the HELAC , MADEVENT and SHERPA results respectively.

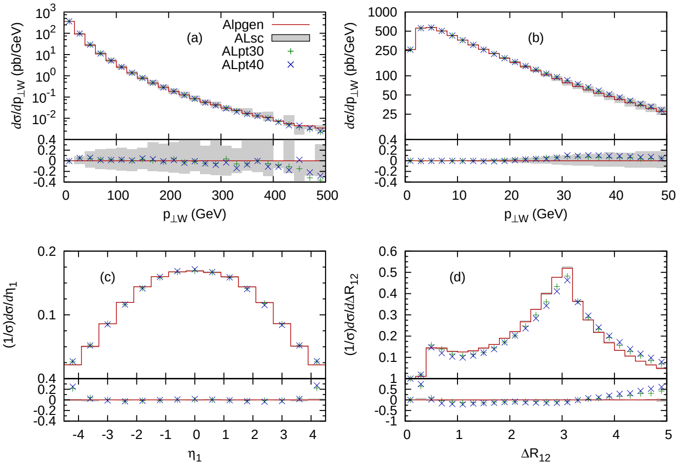


Figure: ALPGEN systematics at the LHC. (a) and (b) show the $p_\perp$ spectrum of the W, (c) shows the pseudo-rapidity distribution of the leading jet, (d) shows the ΔR separation between the two leading jets. The full line is the default settings of ALPGEN , the shaded area is the range between two different scale choiches for α_S, while the points represent two different choiches of matching parameters.

QCD NLO calculations and tools

- Very intense activity by several groups in working out missing NLO calculations for processes relevant at LHC
- NLO calculations allows to make more stable predictions with respect to LO ones
- By now all $2 \to 2$ and $3 \to 3$ processes are are known at NLO accuracy
- Most of the calculations are implemented in Monte Carlo programs allowing to study distributions with any kind of kinematical cuts
- main limitations: no unweighted events, final state particles are partons and not hadrons

QCD NLO calculations and tools

The situation can be summarized by the processes implemented in latest version of `MCFM`

J.M. Campbell and R.K. Ellis

- $pp \to W^\pm/Z$
- $pp \to W^\pm + Z/\gamma/g^* \to b\bar{b}$
- $pp \to W^+W^-$
- $pp \to ZZ$
- $pp \to W/Z + b$
- $pp \to W/Z + b\bar{b}$
- $pp \to W/Z + b + j$
- $pp \to W/Z + j$
- $pp \to W/Z + 2j$

- $pp \to H$
- $pp \to W^\pm/Z + H$
- $pp \to H + j$
- $pp \to H + 2j$
- $pp \to t\bar{t}$
- $pp \to t + X$
- $pp \to t + W$

NLO calculations

- VVj

  Dittmaier et. al.; Campbell et. al; Binoth et al.

- $t\bar{t}H$, $t\bar{t}j$, Wj NLO α_S and α_{EW}

  Dittmaier et. al.

- Hjj

  Campbell et al.; Ciccolini et al.

- VVV

  Lazopoulos et al.; Hankele et al.

dedicated NLO programs

- `GG2WW`: $gg \to W^*W^* \to$ leptons

  Binoth, Ciccolini, Kauer and Krämer

- vector boson pairs in VBF (*i.e.* with two fwd jets)

  Oleari, Zeppenfeld et al.

  - $pp \to WZ \to$ leptons via VBF
  - $pp \to W^+W^- \to$ leptons
  - $pp \to ZZ \to$ leptons

- `FEWZ`: Drell-Yan with NLO and NNLO accuracy and completely exclusive control on lepton momenta

  Melnikov and Petriello

- $pp \to t\bar{t}H$

  Beenakker et al, Dawson et al.

- `DIPHOX`: $\gamma\gamma$ final states

  Aurenche et al.

- `NLOJET++`: $pp \to$ jets

  Nagy

dedicated NLO programs

- $\texttt{FEHIP}$: $pp \rightarrow H$ at NLO and NNLO accuracy

  Anastasiou, Melnikov and Petriello

- HNNLO (Higgs production at NNLO & NNLL), Drell-Yan avaliable soon Catani, Grazzini

multileg NLO calculations

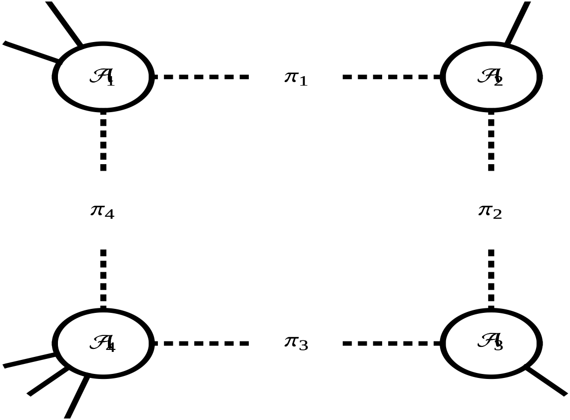
The bottleneck for the extension of NLO calculations to processes with many legs (which are of great interest for background evalutation at LHC): the virtual corrections. Standard analytical Feynman diagrams methods become untractable for many external legs

Recently important steps forward:

- reconstruct the one loop integrand by properly cutting the *scattering amplitudes*
- the one loop integrand is given as a convolution of *tree level* amplitudes with *on shell, complex* external momenta.

multileg NLO calculations

1. A generic NLO terms, before integration over loop momenta, looks like
$$\phi = \frac{A}{\pi_1 \pi_2 \ldots \pi_n}$$
where $\pi_j = (l - q_j)^2 - m_j^2$, l loop momenta and A a polinomial in l
2. In four dimensions
$$\phi \sim \frac{A_{j_1 j_2 j_3 j_4}}{\pi_{j_1} \pi_{j_2} \pi_{j_3} \pi_{j_4}} + \frac{A_{j_1 j_2 j_3}}{\pi_{j_1} \pi_{j_2} \pi_{j_3}} + \ldots$$
3. The form of the $A_{\ldots}$ coefficient is highly constrained and they can be derived computing the convolution of 4 (3 or 2) scattering amplitudes with complex momenta determined in such a way that 4 (3 or 2) propagators diverge
4. the NLO amplitude is therefore
$$\mathcal{A}_{NLO} \sim \tilde{A}_{j_1,j_2,j_3,j_4} \mathcal{I}_{j_1,j_2,j_3,j_4}$$
5. In D dimension some additional complications: rational terms



Cutting four propagators and letting the corresponding momenta to go on shell (complex momenta) the coefficient $\mathcal{A}_{1234}$ is uniquely singled out. It is obtined convoluting the $\mathcal{A}_j$ scattering amplitudes.

multileg NLO calculations

- two approaches (which shares several aspects)
  1. Ossola, Papadopoulos, Pittau
  2. Dixon, Bern, Kosower. Ellis, Giele, Kunszt Zanderighi

multileg NLO calculations

First results:

- $t\bar{t}b\bar{b}$ at NLO; Bredenstein et al., Bevilacqua et al.
- $W + 3\mathrm{j}ets$ at NLO ; Berger et al., Ellis et al.
- $\bar{b}b\bar{b}b$ Binoth et al. (partial: $q\bar{q}$ only)

Parton shower MC at NLO

NLO calculations do a fairly good job in reproducing inclusive cross section as well as a (somewhat limited) number of (IR/Collinear safe) observables. The comparison of data demand a fairly important work to extrapolate experimental data to the parton level (adding to the systematics).

$\Rightarrow$ Quest for PS at NLO to interface NLO calculations with PS and hadronization models.

Problems:

- Retain both NLO and (N)LL accuracy avoiding double counting and preserving a smooth interplay between the two approximations
- Negative weights potentially unsuitable for a MC approach

Parton shower MC at NLO

Insofar two working approaches

- MC@NLO (Frixione, Webber, Nason, Laenen, Motylinski) (Fucks et al. Z')
  1. $$\sigma = \sigma_{NLO} + \sigma_{NLL} - \sigma_{NLL,\, Truncated}$$
  2. implemented for a large number of processes
  3. $10 \sim 20\%$ of negative weights
  4. it requires a detailed knowledge of MC kinematics
  5. the hardest emission is provided either from NLO or by the PS (basically depending on p_T)

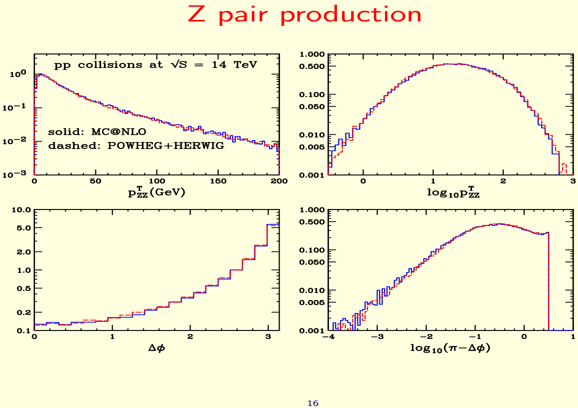
Parton shower MC at NLO

- POWHEG (Nason, Ridolfi, Latunde-Dada, Gieseke, Webber, Frixione, Re, Alioli, Oleari), (Hamilton, Richardson, Tully)

1. $$\sigma \sim \frac{\sigma_{NLO}}{\sigma_{LO}}(\sigma_{NLL} - \sigma_{NLL,\ Truncated} + \sigma_{LO})$$
2. implemented for a fairly large number of processes
3. hardest emission from POWHEG $\Rightarrow$ interaction with PS kept to minimal, just needs modified (or truncated for NLL) showers to avoid radiation from PS harder than NLO one.
4. positive weights almost guaranteed
5. retains NLO accuracy for total σ and the same IR/Collinear accuracy of underling PS
6. in progress POWHEG BOX: input user NLO calculation $\Rightarrow$ output PS at NLO

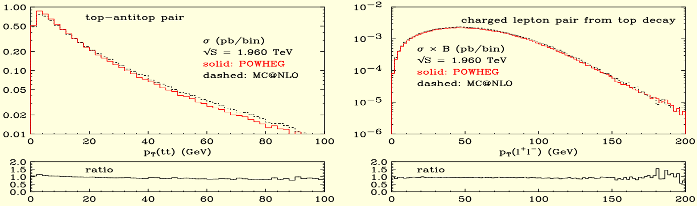
MC@NLO versus POWHEG

- Difference at NNLO level
- Compare well on most observables



MC@NLO versus POWHEG

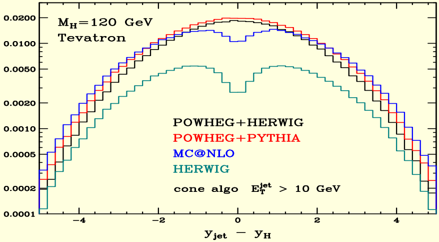
POWHEG and MC@NLO comparison:
Top pair production



Good agreement for most observables considered
(differences can be ascribed to different treatment of higher order terms)

19

Jet rapidity in h production

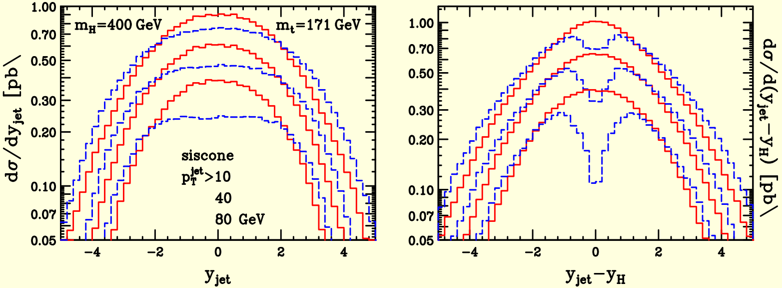


Dip in MC@NLO inerithed from even deeper dip in HERWIG
(MC@NLO tries to fill dead regions in HERWIG, a mismatch remains).

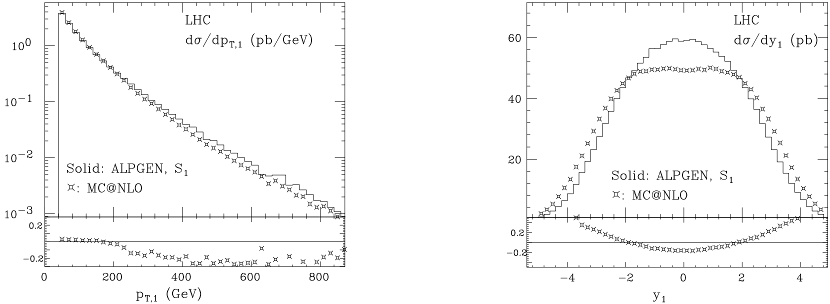
25

- Dip at $\eta = 0$ likely from a dead zone inherited from HERWIG. Nothing like this in POWHEG: hardest radiation always from NLO.
- Differencies starts at NNLO

Gets worse for larger E_T cuts:



looking at more exclusive distributions

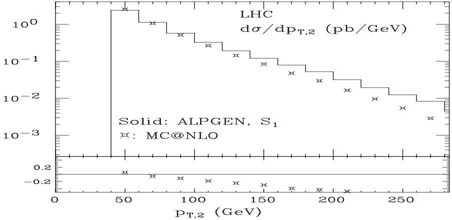


Behaviour first observed comparing `ALPGEN` and MC@NLO.

Always compare output of different codes.

NLO PS guarantee improved accuracy on *inclusive quantities* (non zero at LO), but in phase space corner LO codes can behave better, always check!

looking at higher jet multiplicities



NLO PS montecarlos get the first emissions (with the limitations shown previously) at LO accuracy subsequent one at LL
LO multileg event generators get several emissions at LO accuracy so likely more accurate for higher orders high p_T emissions.

Conclusions

- A variety of tools is avaliable for LO calculations with essentially arbitrary multiplicity.
  1. Once properly interfaced with PS tools they provide a good description of data.
  2. Dedicated effort on experimental side to test and tune this tools
  3. Essential for new physics searces (usually performed in negligible small corners of phase space)
- An impressive compilation of NLO calculatin covering essentially all $2 \to 3$ is avaliable and compare well with data over a broad range of variables
- Getting NLO from LO essentially completed $\Rightarrow$ authomatic NLO computation, very large multiplicities (a few $2 \to 4$ to start with but no limitation in principle)
- Successful implementation of NLO PS matching, black box with NLO input NLO+PS output in progress renedering realistic a full merging of these approaches.