Jet pair production with POWHEG

Emanuele Re*

IPPP, Durham University

LoopFest X

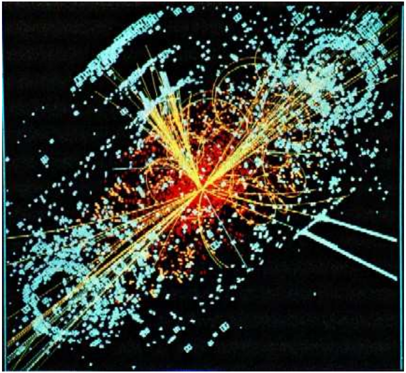
Northwestern University, 13 May 2011

*in collaboration with S. Alioli, K. Hamilton, P. Nason and C. Oleari

[JHEP 1104:081,2011]

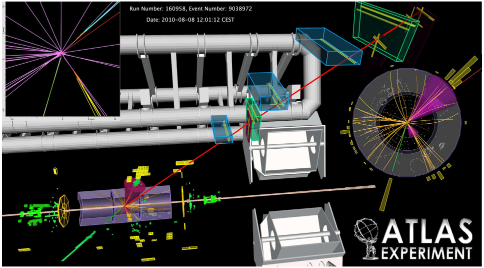
Outline

- Quick description of the method
  - theoretical motivation
  - ingredients and accuracy
- Jet pair production in `POWHEG`
- Results and comments
- Comparison with data
- Outlook

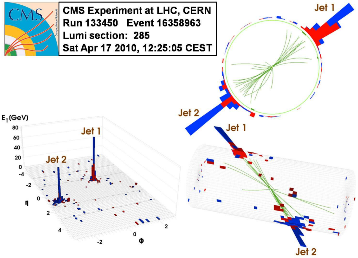


Theoretical predictions for hadron collider physics

- LHC is running, data are collected, many publications already present, a lot of experimental effort...
- and Tevatron is still running too!
- Main goals: understand EWSB mechanism (Higgs boson) and search for new Physics.
- Many steps to achieve this goal:
    - Understand the detectors.
    - Rediscover what we already know.



[ATLAS $t\bar{t}$-pair candidate]



[CMS dijet event]

- Disentangle signal and backgrounds (analysis strategies).
- Compare signals with best available predictions.

traditionally used Th. inputs: parton-level calculations / Monte Carlo event generators.

NLO vs. SMC's (LO + Parton Shower)

- NLO

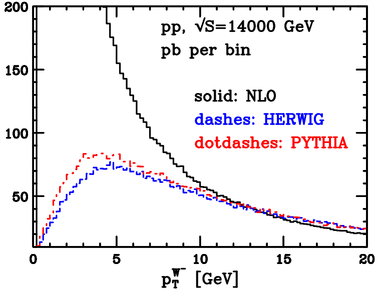
✓ NLO accuracy for inclusive observables (not only rates).

✓ reduced theoretical uncertainty (less sensitive to μ_R and μ_F choices).

✓ accurate shapes at high-p_T (for the 1st emission).

✗ wrong shapes in small-p_T region (or generically where you want to resum logs).

✗ description only at the parton level.



- SMC's

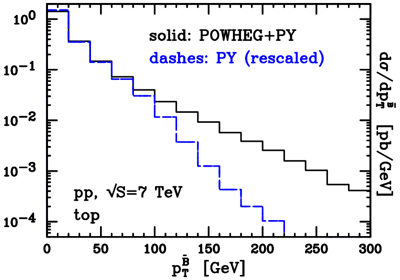
✗ total normalization accurate only at LO.

✗ poor description of high-p_T emissions.

✓ Sudakov suppression of small p_T emissions (LL resummation, via parton showers).

✓ simulate high-multiplicity events at the hadron level, modelling also NP effects.

✓ largely used by experimental collaborations at various stages.



NLO vs. SMC's (LO + Parton Shower)

- NLO

✓ NLO accuracy for inclusive observables (not only rates).

✓ reduced theoretical uncertainty (less sensitive to μ_{R} and μ_{F} choices).

✓ accurate shapes at high-p_{T} (for the 1st emission).

✗ wrong shapes in small-p_{T} region (or generically where you want to resum logs).

✗ description only at the parton level.

- SMC's

✗ total normalization accurate only at LO.

✗ poor description of high-p_{T} emissions.

✓ Sudakov suppression of small p_{T} emissions (LL resummation, via parton showers).

✓ simulate high-multiplicity events at the hadron level, modelling also NP effects.

✓ largely used by experimental collaborations at various stages.

natural to try to merge the 2 approaches, keeping the good features of both.

real emissions included in both approaches

- NLO: exact $n+1$-body matrix element.
- PS's: multiple emissions in the collinear approximation.

main problem: avoid to double-count them !

many proposals, currently two fully tested solutions: `MC@NLO` [Frixione, Webber 2001] and `POWHEG` [Nason 2004].

We start by looking to the formula for a NLO calculation and for the first branching of a LO Parton Shower.

- NLO cross section:

$$d\sigma_{\mathrm{NLO}} = d\Phi_n \left\{ B(\Phi_n) + V(\Phi_n) + \underbrace{[R(\Phi_{n+1}) - C(\Phi_{n+1})]}_{\text{finite}} d\Phi_r \right\}$$

where

$$d\Phi_{n+1} = d\Phi_n d\Phi_r \ , \ \Phi_r = \{t, z, \varphi\} \ , \ \ V(\Phi_n) = \underbrace{V_{div}(\Phi_n) + \int d\Phi_r C(\Phi_n, \Phi_r)}_{\text{finite}}$$

and

$$\frac{R(\Phi_{n+1})}{B(\Phi_n)} d\Phi_r \rightarrow \left(\frac{\alpha_s}{2\pi} \frac{1}{t} P(z) \right) dt\, dz\, \frac{d\phi}{2\pi} \quad \text{when} \ \ t \rightarrow 0$$

coll. factorization

- SMC first emission:

$$d\sigma_{\mathrm{SMC}} = B(\Phi_n)\, d\Phi_n \left[\Delta(t_{\max}, t_0) + \Delta(t_{\max}, t)\, \frac{\alpha_s}{2\pi} \frac{1}{t} P(z)\, d\Phi_r \right]$$

$$\Delta(t_{\max}, t) = \exp\left\{ - \int_t^{t_{\max}} d\Phi_r' \frac{\alpha_s}{2\pi} \frac{1}{t'} P(z') \right\}$$

SMC Sudakov form factor

Idea: *Modify $d\sigma_{\rm SMC}$ in such a way that*, expanding in $\alpha_{\rm S}$, *one recovers the NLO cross section*

- With the substitutions

$$B(\Phi_n) \quad \Rightarrow \quad \bar{B}(\Phi_n) = B(\Phi_n) + V(\Phi_n) + \int \left[R(\Phi_{n+1}) - C(\Phi_{n+1})\right] d\Phi_r$$

$$\Delta(t_{\rm max}, t) \quad \Rightarrow \quad \Delta(\Phi_n; k_{\rm T}) = \exp\left\{ -\int \frac{R(\Phi_n, \Phi'_r)}{B(\Phi_n)} \theta(k'_{\rm T} - k_{\rm T})\, d\Phi'_r \right\}$$ POWHEG Sudakov

we get the POWHEG “master formula” for the hardest emission:

$$d\sigma_{\rm POW} = \bar{B}(\Phi_n)\, d\Phi_n \left\{ \Delta(\Phi_n; k_{\rm T}^{\rm min}) + \Delta(\Phi_n; k_{\rm T}) \frac{R(\Phi_n, \Phi_r)}{B(\Phi_n)} d\Phi_r \right\}$$

[Nason, JHEP 0411:040,2004]

- generated events have positive weight: $\bar{B}(\Phi_n)$ is usually positive: POsitive Weight Hardest Emission Generator
- to avoid double-counting, *subsequent emissions must be $p_{\rm T}$ vetoed !*
- large $k_{\rm T}$ accuracy preserved: since $\Delta(k_{\rm T}) \to 1$,

$$d\sigma_{\rm POW} \approx \bar{B}(\Phi_n) \times \frac{R(\Phi_{n+1})}{B(\Phi_n)}\, d\Phi_{n+1} \approx R(\Phi_{n+1})\, d\Phi_{n+1} \times \left(1 + \mathcal{O}(\alpha_{\rm S})\right)$$

- small $k_{\rm T}$ LL accuracy of SMC's preserved:

$$\frac{R(\Phi_n, \Phi_r)}{B(\Phi_n)} d\Phi_r \approx \frac{\alpha_{\rm S}}{2\pi} \frac{1}{t} P(z)\, dt\, dz\, \frac{d\phi}{2\pi}$$

- inclusive observables have NLO accuracy

Accuracy of the POWHEG Sudakov

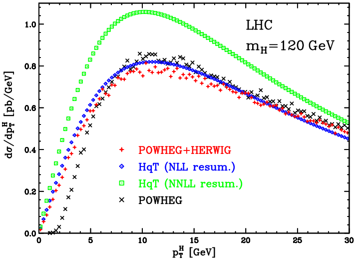
logs that exponentiate ($\sim B$) are resummed, since they are contained in R/B:

- LL OK: double soft and collinear logs are included [✓]
- single collinear logs (NLL) are also included [✓]

to go to full NLL:

- bremsstrahlung scheme: $\alpha_s \to \alpha_s \left(1 + \frac{\alpha_s}{2\pi} K\right)$ [✓]
- include soft non-collinear logs ($\sim B_{ij}$), that in general don't exponentiate. [✗]
- included in POWHEG if no more than 3 colored particles at the Born level. [✓]
- recover these logs in the large N_C limit shown to be possible but not explicitely implemented until now.

⇒ for simple processes, should have NLL accuracy:



Role of the subsequent shower

- it is vetoed: therefore it is responsible for the accuracy of radiation softer than the 1st one.
- in an angular ordered shower, the hardest emission is not the first: a truncated shower is needed to restore soft wide-angle radiation effects.

the POWHEG BOX framework

- Although it may look easy, the actual implementation of the algorithm is not straightforward. [Frixione,Nason,Oleari, JHEP 0711:070,2007]
- Our automation of the algorithm led to the POWHEG BOX package, which has been available for more than 1 year now.
- General features:
  - automation of the POWHEG algorithm using the FKS subtraction scheme.
  - all previous implementations and new ones included in a single and public framework:

    $V, H(gg$ fusion and VBF$), Q\bar{Q},$ single-top $(s, t, Wt), ZZ, V+j, jj, WWjj, Wb\bar{b}, Q\bar{Q}j$

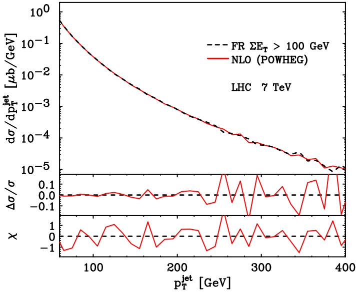
  - it produces LHE files, ready to be showered through HERWIG or PYTHIA.
  - once needed ingredients are provided, it can be used as a "black-box", although all the details were carefully described.

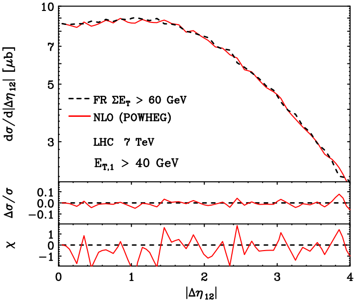
  [Alioli,Nason,Oleari,ER, JHEP 1006:043,2010]
- Other features:
  - we want to keep as much as possible the original goal of independence from the parton-shower. If needed, will try to refine the interface.
  - until now effects of neglecting truncated-shower (when HERWIG is used) were found to be negligible. If needed, this is a point where there is space for improvements.
  - we will continue keeping our code completely available for interested theorists, and if you implement your process, we would be happy to include it in the repository.

http://powhegbox.mib.infn.it

Jet pair production with POWHEG

- Dijet production is by far the most frequent hard scattering in hadronic collisions.
- from the technical point of view, it is up to now the more complicated process implemented in POWHEG.
  This means also a serious test for the POWHEG BOX program.
- All ingredients have been known since the late 80's: [Ellis, Sexton], [Kunszt, Soper]
  - $2 \to 2$ and $2 \to 3$ tree-level amplitudes
  - virtual corrections
  - color-linked amplitudes
  - $2 \to 2$ amplitudes in the planar limit, to assign color structure before showering.
- Check with independent NLO computation by Frixione-Ridolfi:

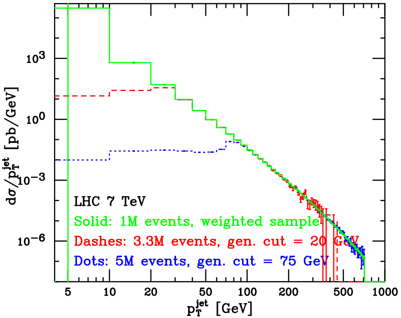




$\Rightarrow$ D0 midpoint, $R = 0.7$, $f = 0.5$

$$\frac{\Delta\sigma}{\sigma} = \frac{\sigma_1 - \sigma_2}{\sigma_2} \qquad \chi = \frac{\sigma_1 - \sigma_2}{\sqrt{\delta\sigma_1^2 + \delta\sigma_2^2}}$$

- Divergent at tree-level !
- In a NLO computation: observable $\mathcal{O}$ is IR-safe, and vanish fast enough when 2 singular regions are approached (i.e. we ask for 2 or more jets)
  $\Rightarrow$ just integrate and fill histograms
- In POWHEG, we start by generating $2 \to 2$ kinematics:
  $\Rightarrow$ a *generation* cut is needed



2 options:

- choose generation cut $\ll$ analysis cut, and check that results don't depend from small variations
- weighted generation:

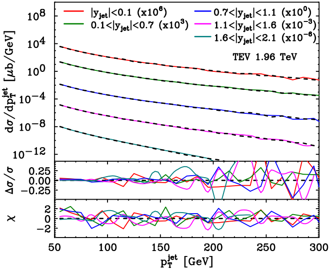
$$\bar{B}(\Phi_2) \to \bar{B}(\Phi_2)\, F(k_T)$$

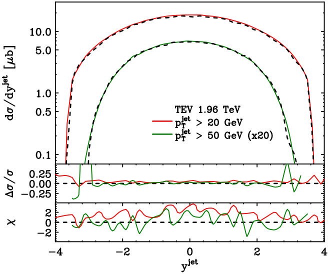
$$F(k_T) = \left(\frac{k_T^2}{k_T^2 + k_{T,s}^2} \right)^3$$

$\Rightarrow$ small k_T suppression
$\Rightarrow$ event weight: $F(k_T)^{-1}$

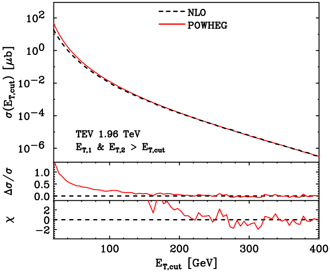
Jet pair production with POWHEG

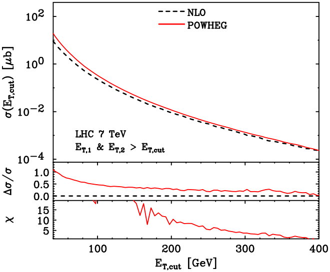
- for inclusive observables, we obtain the expected agreement between NLO and POWHEG: POWHEG = first emission (colored line)





- however, in presence of symmetric cuts:



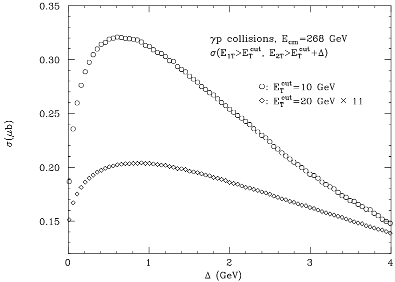


Inclusive dijet processes and the role of cuts

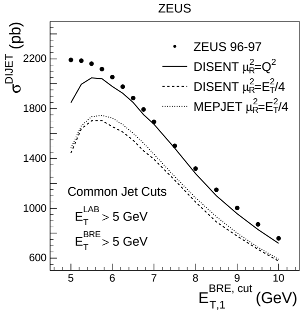
- The most inclusive measurement in jet production is the total cross section. It depends on the cuts used to define jets.
- Despite its simplicity, nontrivial QCD effects take place also when considering the simple observable $\sigma(\Delta)$, where

$$E_{T,2} > E_{T,cut} \quad E_{T,1} > E_{T,cut} + \Delta$$

- From simple considerations on phase space, we expect $\sigma'(\Delta) = d\sigma/d\Delta < 0$, instead NLO prediction has a peak.



γp predictions (from Frixione-Ridolfi)



ZEUS data (from hep/ex:0109029)

- Of course, experimentally there is nothing “special” in using symmetric cuts, as data above show.

Why this problem?

Inclusive dijet processes with symmetric cuts

- as first noticed by Frixione-Ridolfi, NLO curve alone is “wrong” when symmetric cuts are applied $\Rightarrow$ unbalanced cancellation of soft-collinear emissions close to the cut.
- argument by Banfi-Dasgupta (for DIS): $\sigma(E_{T,c},\Delta) = f \otimes C_0(E_{T,c},\Delta)$

  leading-order

$$C_0(\Delta) = \int d\Phi_2 |M_2|^2 \Theta(E_{T,1} - (E_{T,c}+\Delta))\Theta(E_{T,2} - E_{T,c})$$

$$= \int d^2\vec{k}_{T,1}\ J\ |M_2|^2 \Theta(k_{T,1} - (E_{T,c}+\Delta))$$

$$C_0'(\Delta) = -\int d^2\vec{k}_{T,1}\ J\ |M_2|^2 \delta(k_{T,1} - (E_{T,c}+\Delta)) \quad \Rightarrow \sigma' < 0$$

  real + virtual emission, in the soft+coll limit:

$$C_1'(\Delta) \sim -\int d^2\vec{k}_{T,1}\ J\ |M_2|^2 \delta(k_{T,1} - (E_{T,c}+\Delta)) \times \int d\Phi_r S(k_r)[\Theta(\Delta - |k_{r,x}|) - 1]$$

  where

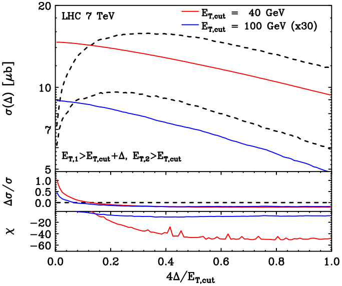
  - $|k_{r,x}| = |E_{T,1} - E_{T,2}|$
  - $|k_{r,x}| < \Delta$ needed to have $E_{T,2} > E_{T,c}$
  - assume k_r not recombined with $k_{T,1}$ or $k_{T,2}$

  NLO, in the soft limit: $C'_{NLO}(\Delta) = C_0'(\Delta)\ W_{NLO}(\Delta)$

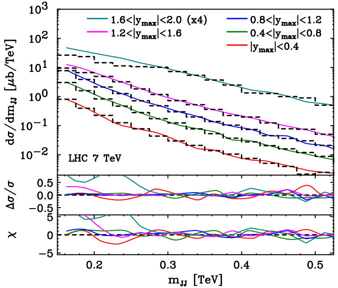
$$W_{NLO} = 1 + \int d\Phi_r S(k_r)[\Theta(\Delta - |k_{r,x}|) - 1] = 1 - c\frac{\alpha}{\pi}\log^2\left(\frac{Q}{\Delta}\right)$$

need of resummation to restore the correct behaviour

Inclusive dijet processes with symmetric cuts



- Observed the same pattern of FR in dijet hadroproduction with POWHEG
- Resummation performed by the shower works well (here shown POWHEG first emission). Notice that in this case it's a LL resummation.



$\Leftarrow \quad |y| = \max(|y_1|, |y_2|)$

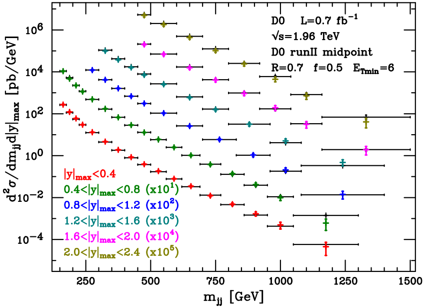
- Although in $\sigma(\Delta)$ the effect is huge, symmetric cuts may affect also other distribution...

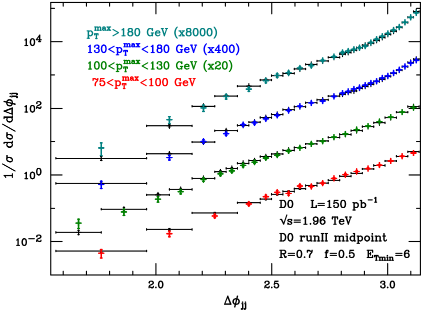
$$E_T \sim \frac{m_{jj}}{2\cosh|y|}$$

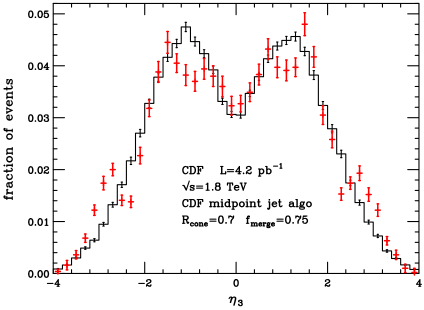
Here we used $E_{T,cut} = 40$ GeV:

$$y \sim 1.8 \quad \Rightarrow \quad m_{jj} \sim 250 \text{ GeV}$$
$$y \sim 1.4 \quad \Rightarrow \quad m_{jj} \sim 170 \text{ GeV}$$

comparison with Tevatron data

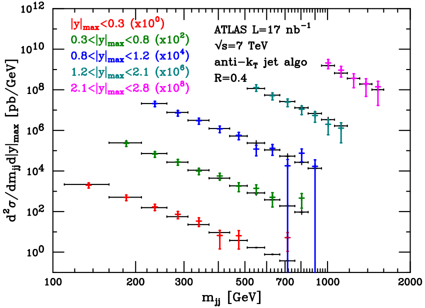


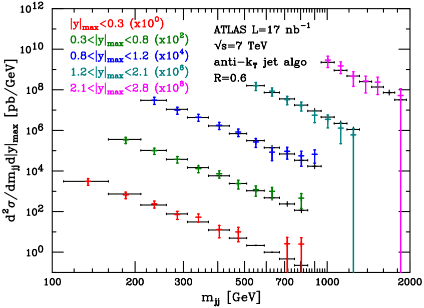


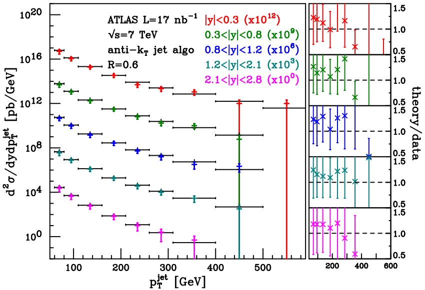


- black: POWHEG+ PYTHIA, Perugia tune
- direct comparison with D0 data: no K factors, no parton-to-hadron corrections
- 5M weighted events, $k_{T,cut} = 1$ GeV, $F(p_T) = \left(\frac{p_T^2}{p_T^2+(600)^2}\right)^3$, folded integration.

comparison with ATLAS data



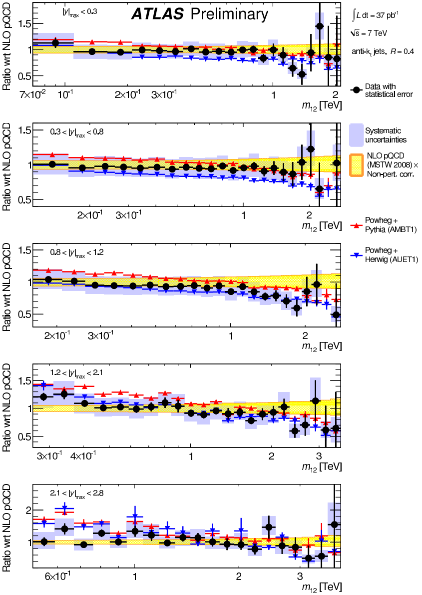




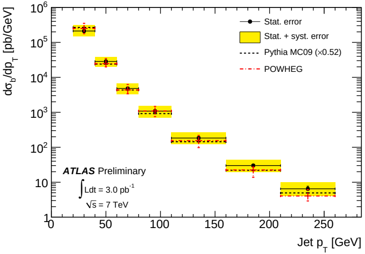
- 5M weighted events, $k_{T,cut} = 1$ GeV, $F(p_T) = \left(\frac{p_T^2}{p_T^2 + (200)^2}\right)^3$, folded integration.
- when comparing with first ATLAS data [Eur.Phys.J.C71:1512(2011)], we found good agreement.
- with more recent data, an ATLAS note showed a sizeable disagreement, especially in m_{jj} with $R = 0.6$.
- problem is currently under study.

ATLAS studies: m_{jj} and p_T

Program already used in ATLAS-CONF-2011-038,-047,-056,-057 CMS-PAS-FWD-10-003,-006

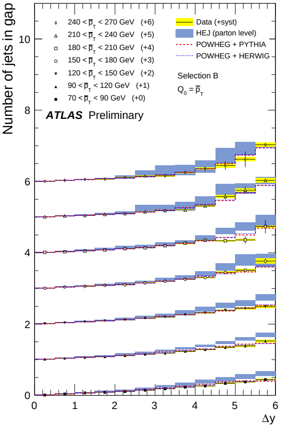


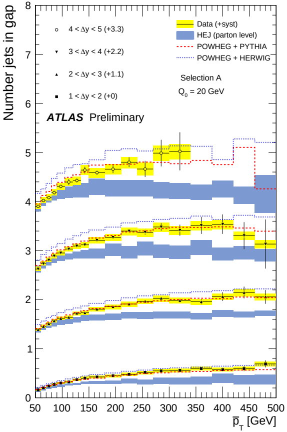
- dijet invariant mass, $R = 0.4$
- cuts: $p_T^{j1} > 30$ GeV, $p_T^{j2} > 20$ GeV, $|y^j| < 4.4$
- observed disagreement, especially when $R = 0.6$



- $R = 0.4$, $|y^j| < 2.1$, cuts: $p_T^j > 20$ GeV,
- MC b-jets = jet with a b-flavoured hadron within $\Delta R = 0.4$.
- POWHEG as it is, PYTHIA corrected with a K-factor to match total measured cross section.

ATLAS studies: activity between jets



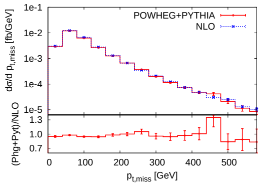


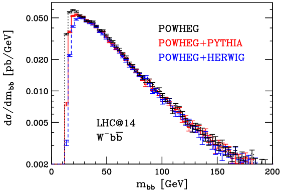
- cuts: $p_T^j > 20$ GeV, $|y^j| < 4.5$
- gap region = 2 highest-y jets, with $\bar{p}_T > 50$ GeV
- gap events $=$ no jets harder than Q_0 within the gap (here $Q_0 = \bar{p}_T$)

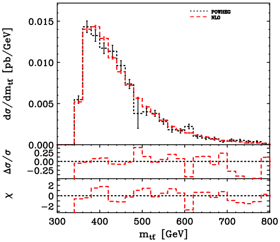
- cuts: $p_T^j > 20$ GeV, $|y^j| < 4.5$
- gap region = 2 highest-p_T jets, with $\bar{p}_T > 50$ GeV
- gap events $=$ no jets harder than Q_0 within the gap (here $Q_0 = 20$ GeV)

Conclusions and outlook

- Many $2 \to 2$ SM processes are available within the `POWHEG BOX` package.
- Implementing jet-pair was a serious test for our automation of the algorithm.
- Together with other `POWHEG` implementations (in `HERWIG++` and `SHERPA`) and with `MC@NLO` it is already possible to simulate almost all $2 \to 2$ SM processes with NLO+PS accuracy.
- $2 \to 3$ implementations are work in progress, and a $2 \to 4$ implementation was already possible.

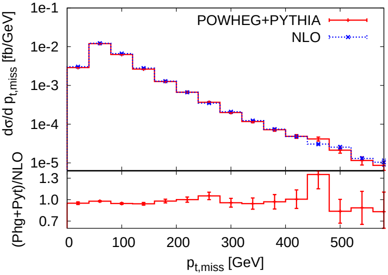






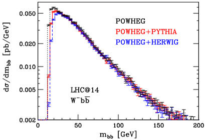
Conclusions and outlook

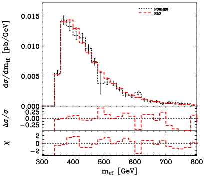
- Many $2 \to 2$ SM processes are available within the `POWHEG BOX` package.
- Implementing jet-pair was a serious test for our automation of the algorithm.
- Together with other `POWHEG` implementations (in `HERWIG++` and `SHERPA`) and with `MC@NLO` it is already possible to simulate almost all $2 \to 2$ SM processes with NLO+PS accuracy.
- $2 \to 3$ implementations are work in progress, and a $2 \to 4$ implementation was already possible.



W^+W^+jj

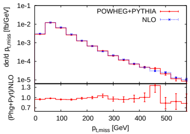
[Melia,Nason,Rontsch,Zanderighi, arXiv:1102.4846]

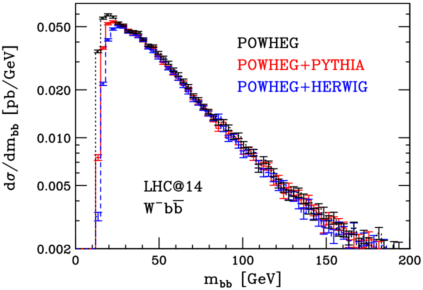


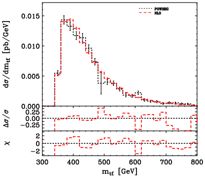


Conclusions and outlook

- Many $2 \to 2$ SM processes are available within the `POWHEG BOX` package.
- Implementing jet-pair was a serious test for our automation of the algorithm.
- Together with other `POWHEG` implementations (in `HERWIG++` and `SHERPA`) and with `MC@NLO` it is already possible to simulate almost all $2 \to 2$ SM processes with NLO+PS accuracy.
- $2 \to 3$ implementations are work in progress, and a $2 \to 4$ implementation was already possible.





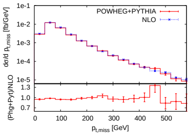


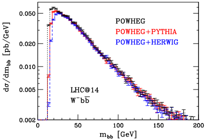
$Wb\bar{b}$

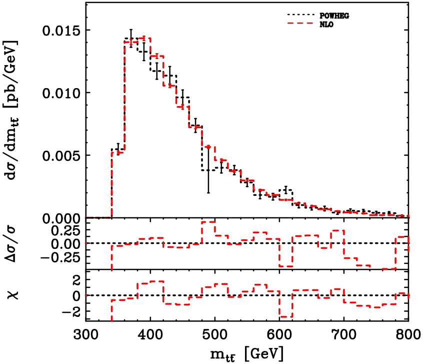
[Oleari,Reina, PRELIMINARY]

Conclusions and outlook

- Many $2 \to 2$ SM processes are available within the POWHEG BOX package.
- Implementing jet-pair was a serious test for our automation of the algorithm.
- Together with other POWHEG implementations (in HERWIG++ and SHERPA) and with MC@NLO it is already possible to simulate almost all $2 \to 2$ SM processes with NLO+PS accuracy.
- $2 \to 3$ implementations are work in progress, and a $2 \to 4$ implementation was already possible.







$Q\bar{Q}j$

[Kardos, Papadopoulos, Trocsanyi, arXiv:1101.2672]

[Alioli,Moch,Uwer, PRELIMINARY]

Conclusions and outlook

- Many $2 \to 2$ SM processes are available within the `POWHEG BOX` package.
- Implementing jet-pair was a serious test for our automation of the algorithm.
- Together with other `POWHEG` implementations (in `HERWIG++` and `SHERPA`) and with `MC@NLO` it is already possible to simulate almost all $2 \to 2$ SM processes with NLO+PS accuracy.
- $2 \to 3$ implementations are work in progress, and a $2 \to 4$ implementation was already possible.
- Understand the origin of the disagreement with ATLAS dijets data is work in progress.
- In general, the validation of the code will be demanding for more complicated processes:

  $\Rightarrow$ code running properly $\neq$ implementation fully understood
  $\Rightarrow$ this could be especially relevant for processes with multijets

Outlooks:

- Many interesting processes yet to be implemented (DY with EW corrections, V+multijets, heavy flavours with jets, exact mass effects in Higgs gluon fusion, BSM).

  $\Rightarrow$ use them to do some phenomenology
  $\Rightarrow$ allow experimentalists to have accurate tools
- Interfacing to modern codes for virtual corrections.
- Further studies and improvements are possible, for example `MENLOPS`

  [Hamilton,Nason], [SHERPA]

  $\Rightarrow$ include multileg accuracy to a NLO+PS simulation.

Conclusions and outlook

- Many $2 \to 2$ SM processes are available within the `POWHEG BOX` package.
- Implementing jet-pair was a serious test for our automation of the algorithm.
- Together with other `POWHEG` implementations (in `HERWIG++` and `SHERPA`) and with `MC@NLO` it is already possible to simulate almost all $2 \to 2$ SM processes with NLO+PS accuracy.
- $2 \to 3$ implementations are work in progress, and a $2 \to 4$ implementation was already possible.
- Understand the origin of the disagreement with ATLAS dijets data is work in progress.
- In general, the validation of the code will be demanding for more complicated processes:

  $\Rightarrow$ code running properly $\neq$ implementation fully understood
  $\Rightarrow$ this could be especially relevant for processes with multijets

Outlooks:

- Many interesting processes yet to be implemented (DY with EW corrections, V+multijets, heavy flavours with jets, exact mass effects in Higgs gluon fusion, BSM).

  $\Rightarrow$ use them to do some phenomenology
  $\Rightarrow$ allow experimentalists to have accurate tools
- Interfacing to modern codes for virtual corrections.
- Further studies and improvements are possible, for example `MENLOPS`

  [Hamilton,Nason], [SHERPA]

  $\Rightarrow$ include multileg accuracy to a NLO+PS simulation.

Thanks for your attention!