ATLAS+CMS top quark pair modelling uncertainties and correlations with single top

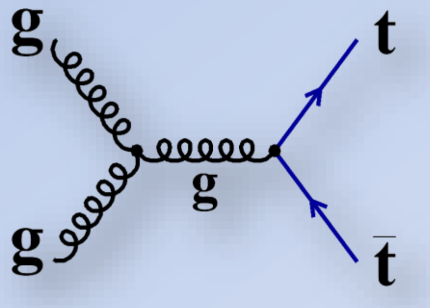
Dominic Hirschbühl
Benedikt Maier

BERGISCHE
UNIVERSITÄT
WUPPERTAL

TOPLHCWG meeting
12.01.2015

Introduction

- Modelling and uncertainties for the $t\bar{t}$ production @ NLO
  - Three main generators are available
    - Powheg → focus today
    - Madgraph5_aMC@NLO
    - Sherpa
  - Focus on radiation systematic for NLO generators



- Scale variations / radiation systematic for Powheg + Pythia6
  - Variations of the hard scatter scale
  - Variations of the damping parameter h_{damp}
    - Setting h_{damp} to a finite value effectively corresponds to damping of high-p_T radiation in Powheg
    - Needs to be switched on, otherwise scale dependence underestimated at high-p_T
    - Powheg-specific, no need to do it in (Madgraph5_a)MC@NLO!

- Correlations of uncertainties between $t\bar{t}$ and single top

Systematic uncertainties to be discussed

- Up to now:
  - https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TheorySystematics
  - Focus today mostly on NLO+PS generators

- Baseline of list of systematic uncertainties is list from Rikkert for NLO + PS generators
  - Scale Variations/Radiation → main focus
  - Parton shower
  - PDF uncertainty
  - NLO-subtraction method
  - DR vs. DS (only for Wt) → discussed in talk by Reinhard

- Processes for which uncertainties should be considered
  - $t\bar{t}$, t-channel single top, s-channel single top, Wt channel

- Discuss possible correlation between processes
  - For Wt analysis most important, but also for t-channel

Scale variations / Radiation

Systematic connected with the scales of the event. Sometimes also called as *ISR/FSR* or *radiation systematic*. Typically scale(s) and/or α_s are changed.
Variations should be checked with data, i.e. jet-gap / N-jet Rivet analysis

Summary of already used variations in top publications:

- Liza's talk in last open TOPLHCWG meeting: https://indico.cern.ch/event/301787/session/10/contribution/22/material/slides/0.pdf
- ATLAS PUB note: ATL-PHYS-PUB-2014-005

- CMS: MadGraph +Pythia6 (multi-leg), vary ren. scale (Q [GeV]) in ME and PS simultaneously by 1/2 and 2. In addition: vary fac. scale simultaneously as well.
- ATLAS (since summer 2013): Alpgen +Pythia6 (multi-leg), vary ren. scale (Q [GeV]) in ME and PS simultaneously by 1/2 and 2. In addition: retune UE. No fac. scale variation.
- ATLAS: AcerMC +Pythia6 (Born-level), vary ren. scale (Q [GeV]) in PS (but not in ME) by $\sim$ 1/2 and 2 (exact range from data limits). In addition: vary hardest emission scale PARP(67)(data limits)

Rivet analyses used for validation of variations

- Gap fraction - Eur. Phys. J. C72 (2012) 2043
- Jet multiplicities - arXiv:1407.0891
- Top parton distributions - Phys. Rev. D 90, 072004 (2014)

CMS analyses are being implemented but not ready yet.

Gap fraction analysis

Eur. Phys. J. C72 (2012) 2043

Study fraction of $t\bar{t}$ events, that do not contain an additional jet(s):

- Sensitive to the amount of extra radiation
- Use dilepton events with two reconstructed b-quark jets
  → additional (radiated) jets easily to identify

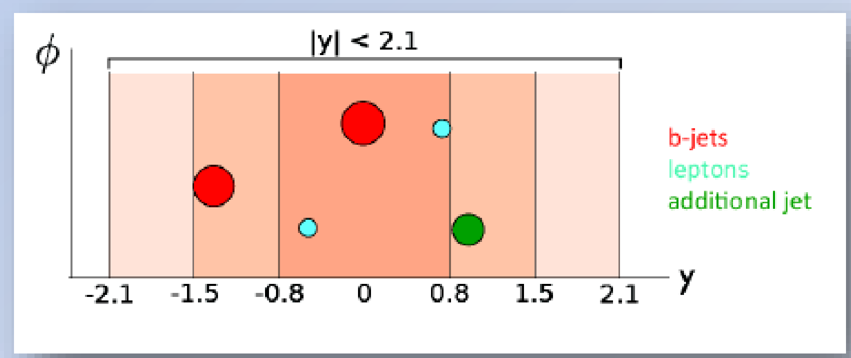
Provided unfolded distributions

- Fraction of events that do not contain an additional jet in a central rapidity region with $p_T > Q_0$:

$$f_{gap}(Q_0) = \frac{n_{gap}(Q_0)}{N_{t\bar{t}}}$$

- Sum of the p_T of the jets falling into each rapidity region

$$f_{gap}(Q_{sum}) = \frac{n_{gap}(Q_{sum})}{N_{t\bar{t}}}$$



Official Rivet routine since Rivet 1.8.1

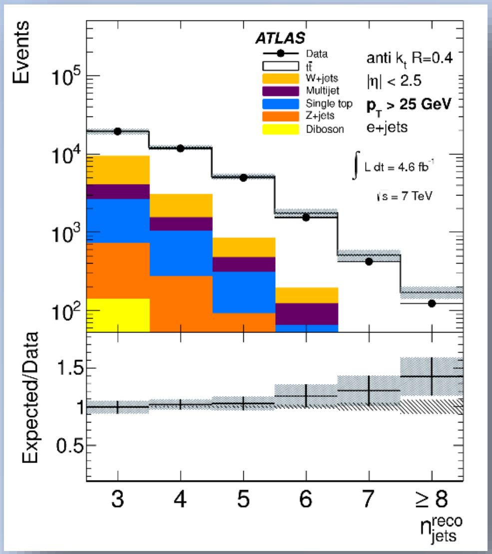
Similar Analysis from CMS:

- 7 TeV: arXiv:1404.3171
- 8 TeV: CMS-PAS-TOP-12-041

Jet multiplicity and jet transverse momentum

Particle level definition of objects:

$E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV & $m_{\mathrm{T}}(W) > 35$ GeV One or more b-jets Three or more jets with $p_{\mathrm{T}} >25$ GeV & $\lvert\eta\rvert < 2.5$
e (μ) with $p_{\mathrm{T}} > 25$ GeV & $\lvert\eta\rvert < 2.5$ No additional e (μ) with $p_{\mathrm{T}} > 15$ GeV & $\lvert\eta\rvert < 2.5$ No μ (e) with $p_{\mathrm{T}} > 15$ GeV & $\lvert\eta\rvert < 2.5$
No jet-jet pair with $\Delta R < 0.5$ No jet-electron or jet-muon pair with $\Delta R < 0.4$



- Selection of semileptonic $t\bar{t}$ events, i.e. 4 jets “belong” to the $t\bar{t}$ process, the 5$^{\mathrm{th}}$ jet, ordered in p_T, corresponds to the first additional emission → should be correlated to $f_{gap}(Q_0)$.
- B-tagging is done using ghost tagging, i.e. adding B-hadrons to the jet clustering

Provided unfolded distributions

- Jet multiplicities for jets with $p_T > 25$ GeV, $p_T > 40$ GeV, $p_T > 60$ GeV, $p_T > 80$ GeV
- Jet p_T for the first five jets

arXiv:1407.0891

Official Rivet routine since Rivet 2.2.0

Normalized differential distributions

Definition of top quarks:

Selection of semileptonic $t\bar{t}$ events
Reconstruction of top quarks from jets, leptons and missing transverse momentum
Unfolded to parton level where the top quark is defined directly before the decay and after QCD radiation.

Provided unfolded distributions

- Transverse momentum of the top quark
- Transverse momentum and rapidity of the $t\bar{t}$ system
- Invariant mass of the $t\bar{t}$ system

"Private" Rivet routine

- No chance to be included in Rivet because of parton level quantities
- So far, only compared with Pythia6 and Herwig
  (possible ambiguity in event record with C++ generators)

Phys. Rev. D 90, 072004 (2014)

Contents of studies

Variations for Powheg

Can vary scales of the ME and the shower and h_{damp}

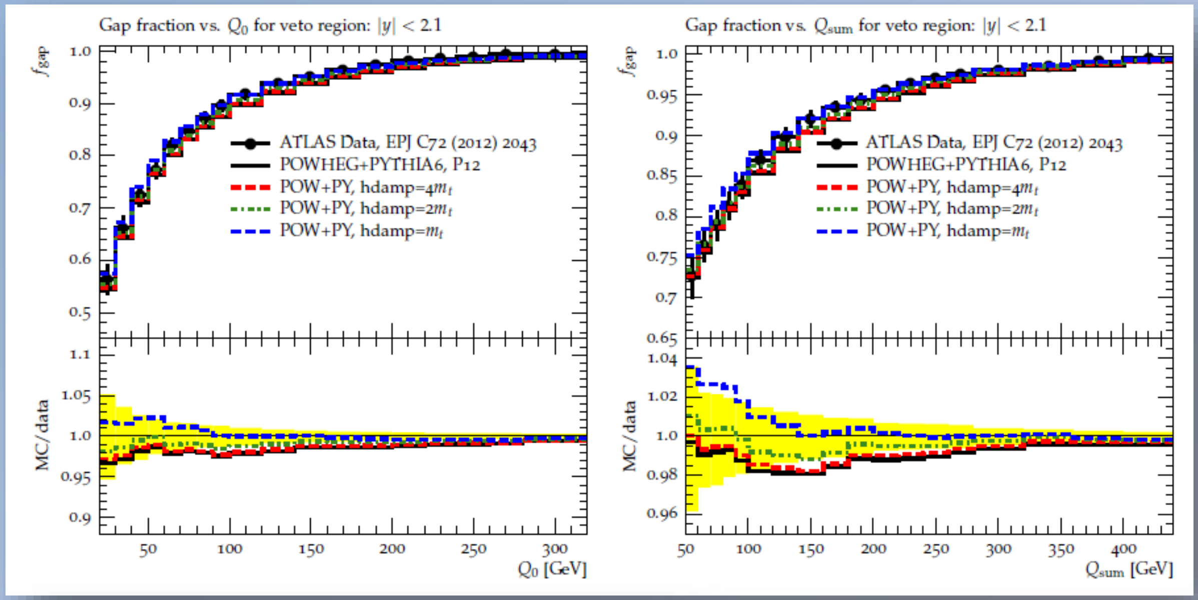
- Scan of h_{damp} for $h_{damp} = 0.5 \cdot m_t, m_t, 2 \cdot m_t, 4 \cdot m_t, \infty$
- Variations of renormalization/factorization scale μ by a factor of 0.5 and 2
- Change Pythia6 tune: PerugiaRadHi / PerugiaRadLo

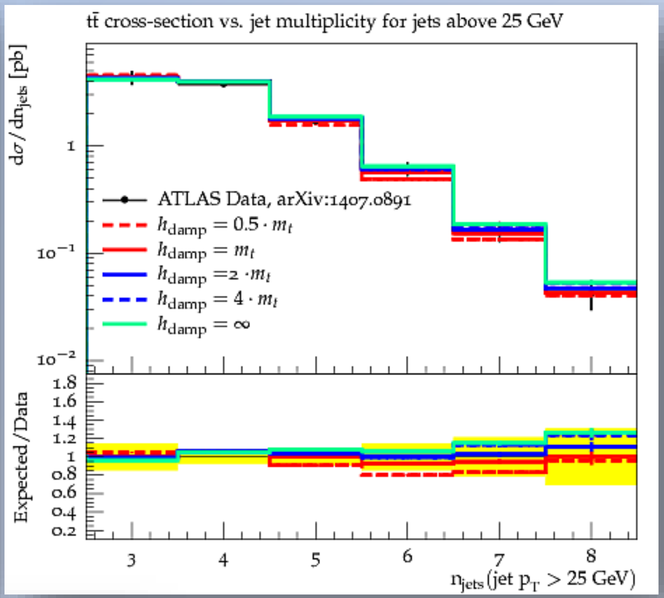
Question to theorists on Powheg:

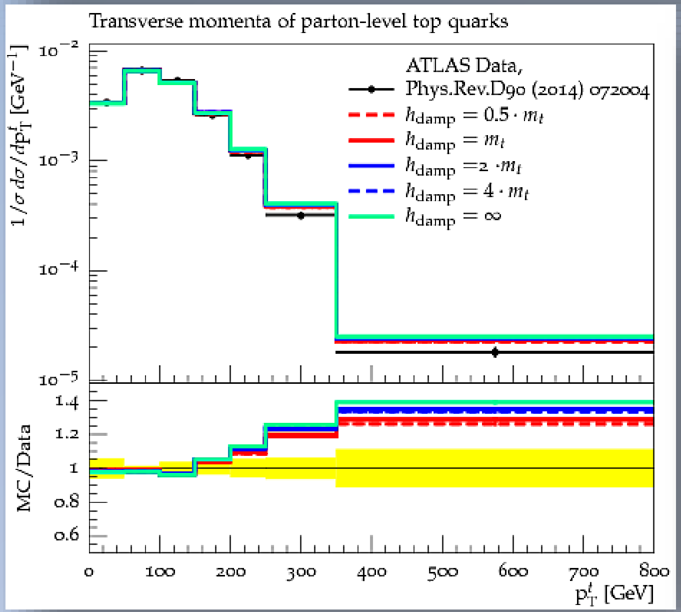
Which strategy is appropriate for choosing scales and h_{damp} ?

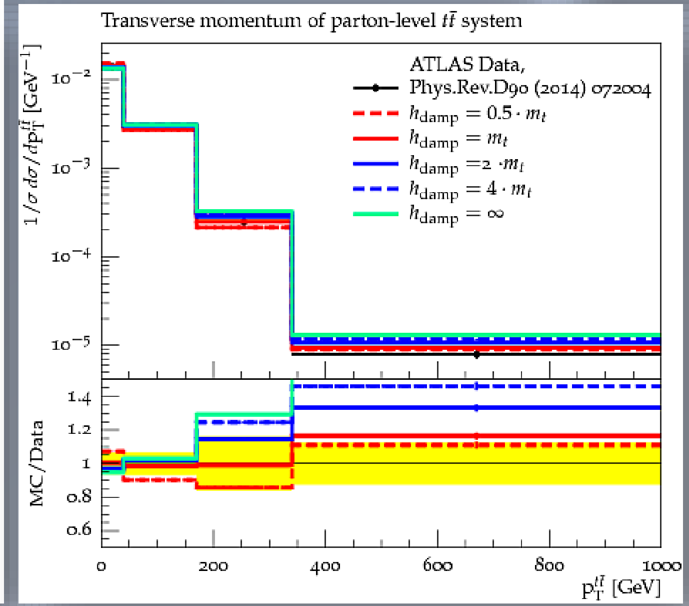
- Use only suggestion from "theory"
  - $h_{damp} = m_t$, $\mu = 1$ as central values and variations of μ and h_{damp}
- Tune central values of h_{damp} and μ to data and use variations (suggest from theory) as systematic
- Tune central values and variations to data
  → which results are appropriate/sufficient to cover with uncertainty?

h_{damp} variations



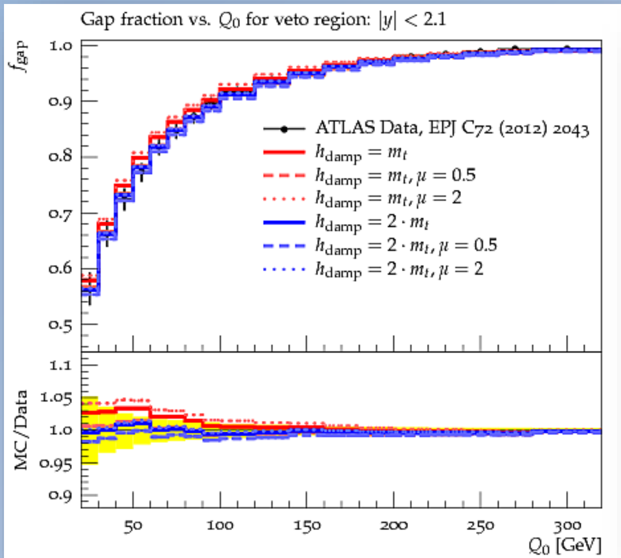


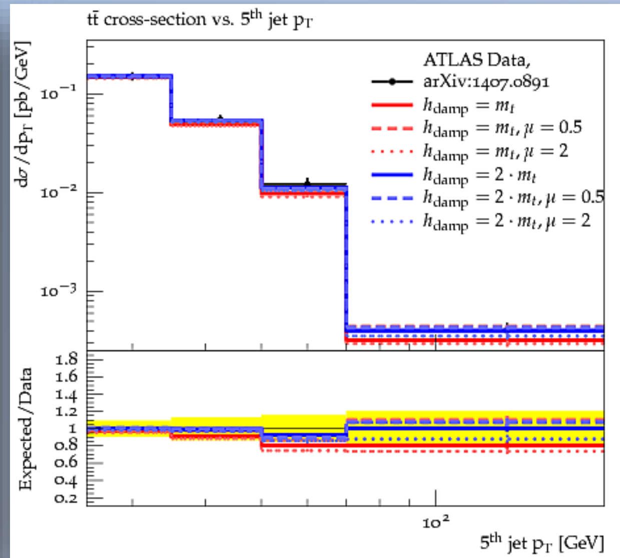


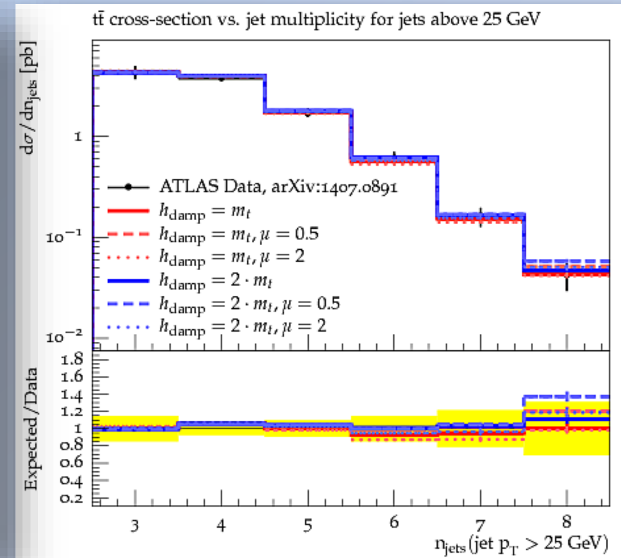


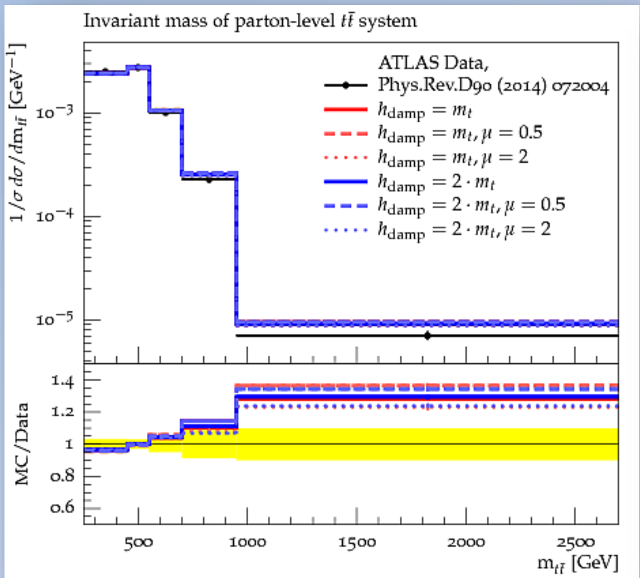
- h_{damp} effects mainly on $p_T(t\bar{t})$ an gap fraction
- mild influence on high p_T tail of top quark
- no effect on rapidities.

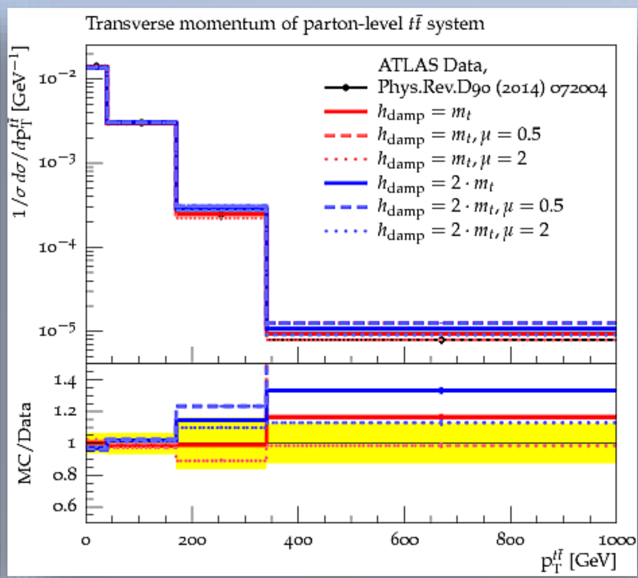
ME scale variations for different h_{damp}









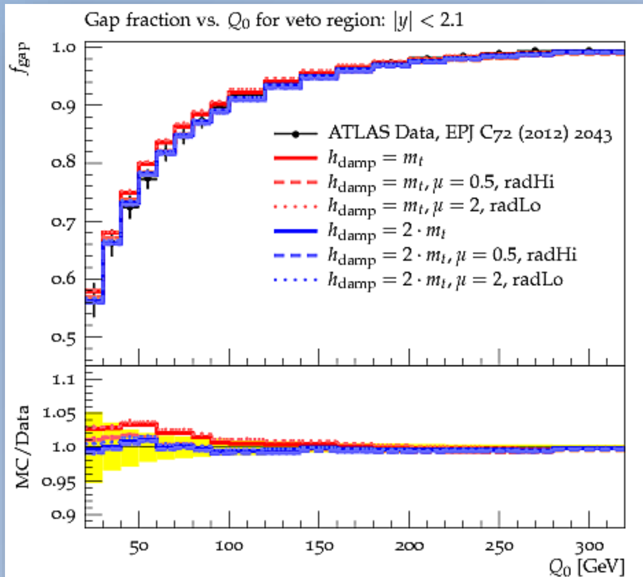


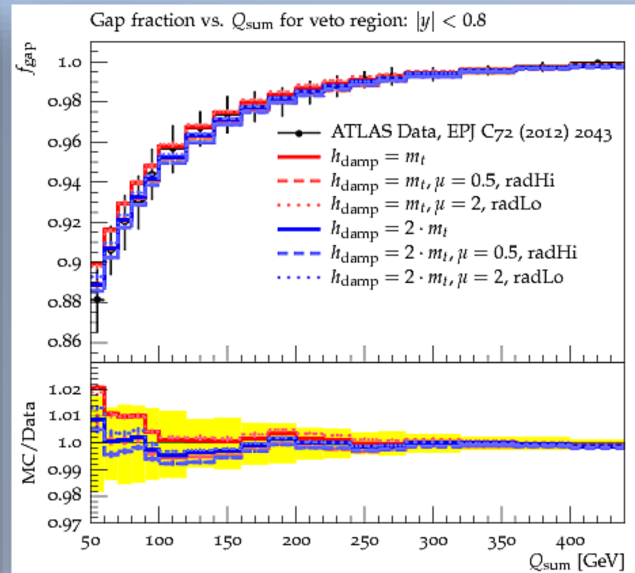
For rapidity gap scale variations ≈ h_{damp} variations

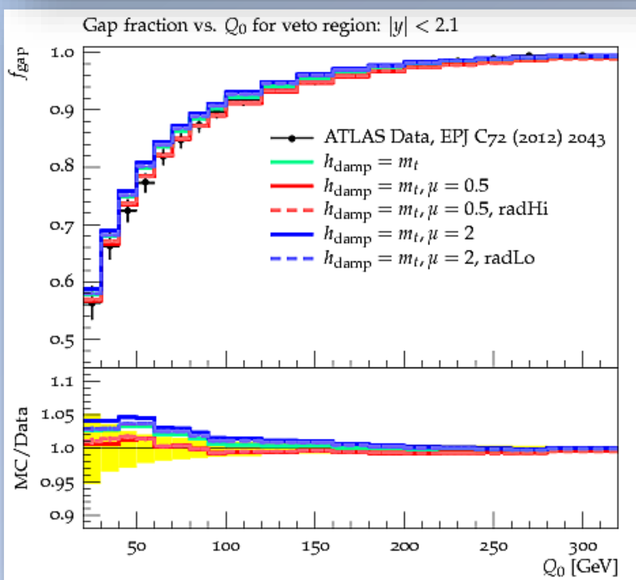
Jet multiplicity doesn't show big changes

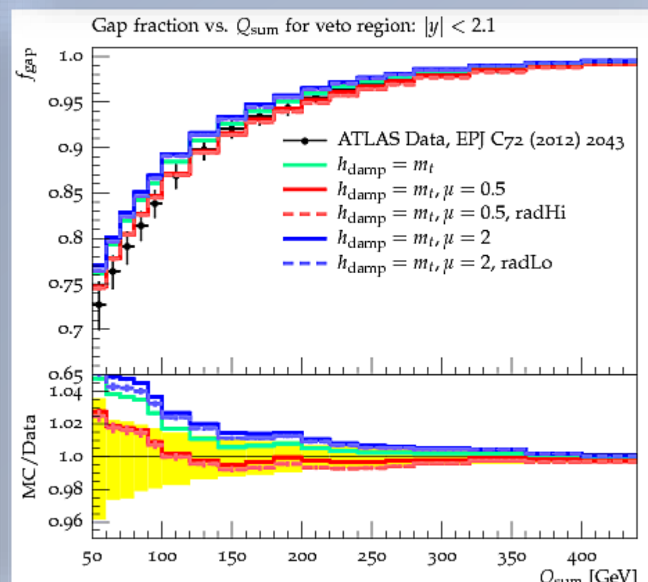
Huge change in $p_T(t\bar{t})$ and small changes in $m(t\bar{t})$

Correlated variations of ME and PS scales



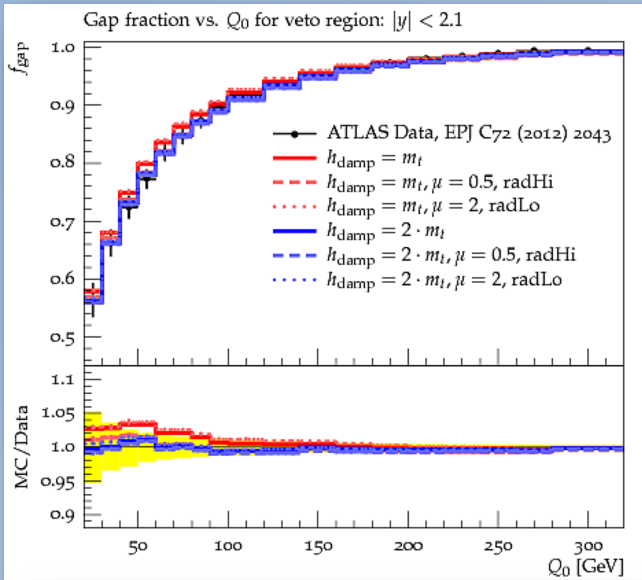


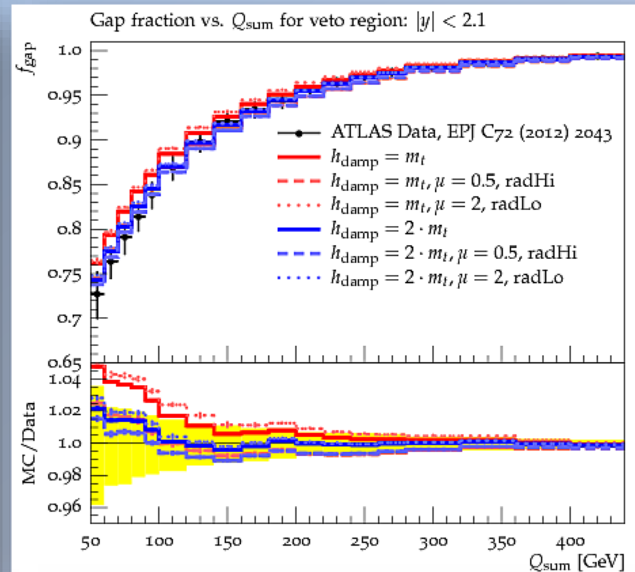


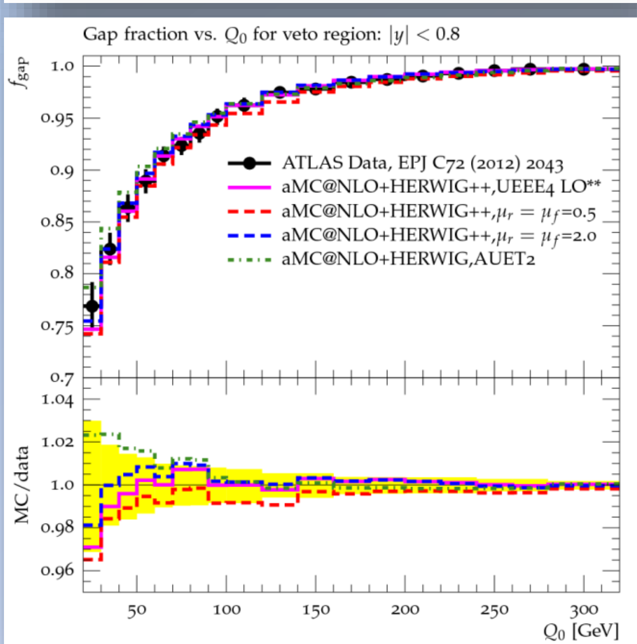


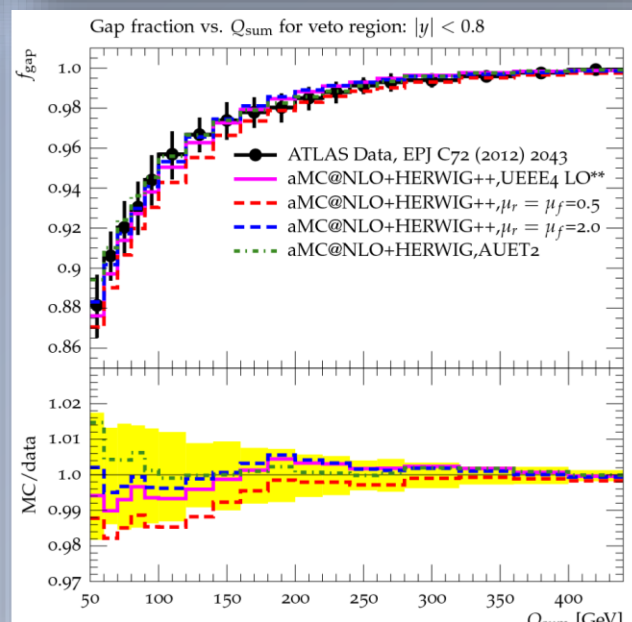
No big effect on gap fraction

Comparison with Madgraph5_aMC@NLO









Scale/h_{damp} variations in Powheg have approx. the same size as scale variations in Madgraph5_aMC@NLO

ATL-PHYS-PUB-2014-005

Summary scale variations

Summary:

h_{damp} and ME scale variation have similar impact on observables.
$h_{damp} + 2\cdot$scale and $h_{damp} + 0.5\cdot$scale covers the envelope of independent variations

Correlation with PS scale doesn't have a big influence in the studied observables.

None of the variations give good agreement in $p_T(t)$

Input from theory needed!

- How much should we tune these parameters and how much can we constrain the uncertainties from data?
- Which observables are suitable?
- Is there any parameter changing $p_T(t)$?

For the (near) future:

How to deal with Multileg NLO matched and merged samples?

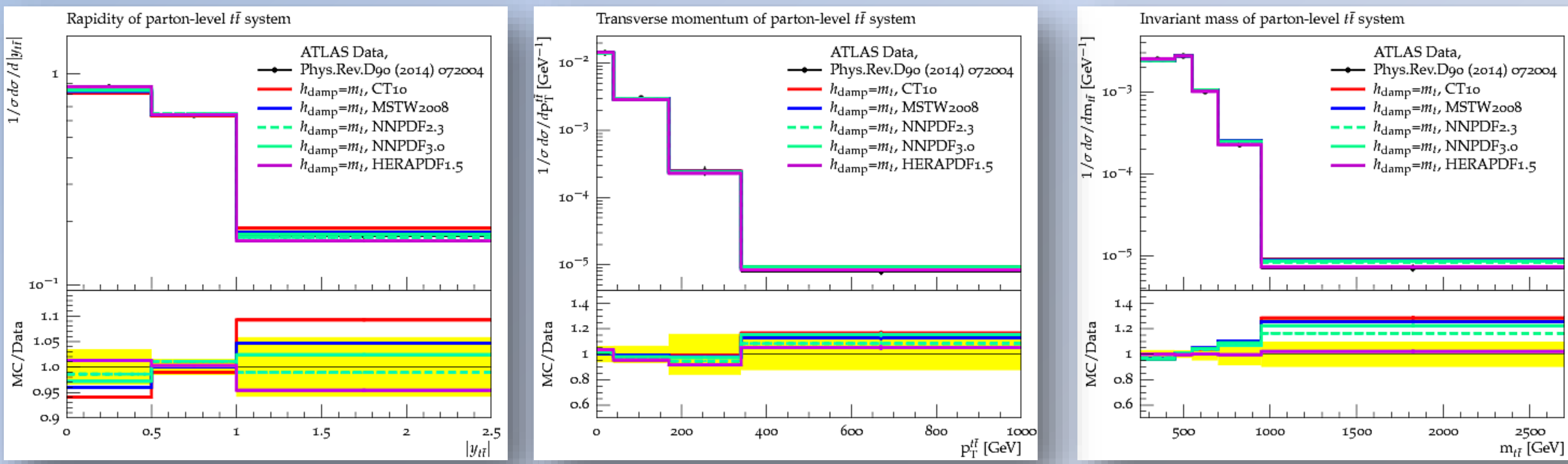
Other uncertainties

PDF: Follow PDF4LHC recommendation (current one and then the future one)
→ a clear recipe should be provided by PDF4LHC
(Building the envelope of 200 EVs is horrible → see also correlations)

1) Simplicity. The "midpoint prescription" is not very suitable.
   → Would like to have one PDF as default and then determine uncertainties, like it is done for all other model uncertainties.
2) Reasonable uncertainties.
   → Envelope approach may not be best estimate of uncertainty.

Choice of central PDF:

- NNPDF3.0 looks similar to the last generation of PDFs but has better modelling of $y(t\bar{t})$.



Other uncertainties

NLO subtraction:

Comparison between two different methods, i.e. MC@NLO vs. Powheg

→ Is it possible to provide a single parameter variation within one NLO setup?

Parton shower / Hadronisation

Comparison between two different shower MC, e.g. Pythia vs. Herwig

→ We should ideally find out a way to estimate the double counting with other uncertainties

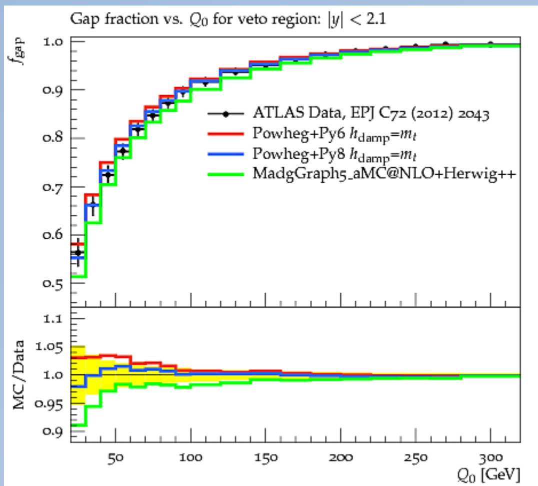
See also the two top mass talks from Fabrice and Markus tomorrow

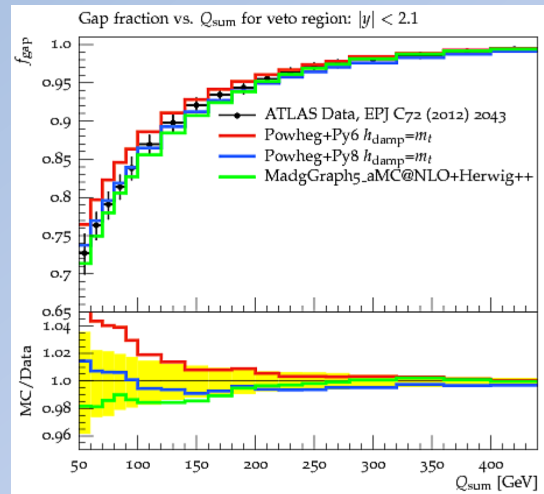
New generators

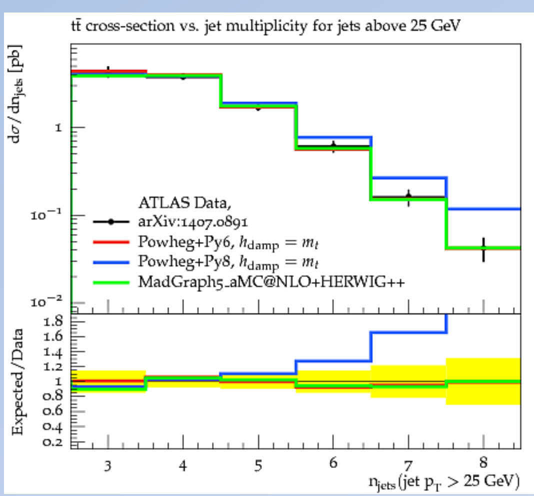
Comparison of new generators with Powheg+Pythia6

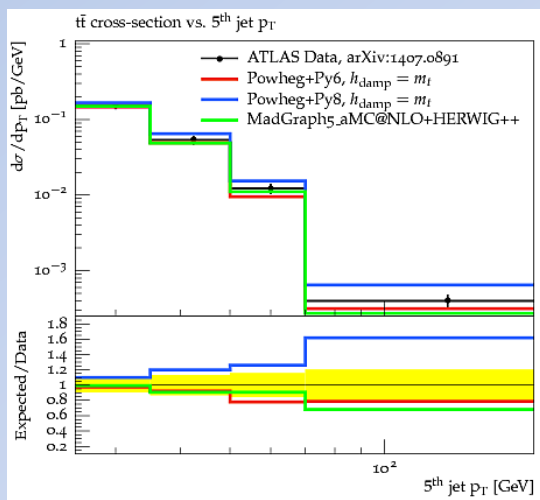
Powheg+Pythia8 → uses vetoed showers and AU2 tune

aMC@NLO+Herwig++ → UEEE5 tune









aMC@NLO + Herwig++ looks promising

Powheg+Pythia8 predicts much harder additional radiation than Powheg+Pythia6 for the jet multiplicity, but only for higher jet bins.

How does the h_{damp} interplay with the matching to Pythia8?

Correlations between processes

In general we always vary the same parameters.

- PDF uncertainty
  → correlate (technically very difficult)
  What about correlation between theory prediction and acceptance (e.g., NNLO vs NLO)? Currently we assume them to be uncorrelated

- Scale Variations / Radiation
  → don't correlate: scale variation can have different impact, e.g. different initial states.

  Question to theorists:
  Does the systematic on the radiation depend on the initial state of the process, e.g. gluons vs. quarks?
  Should h_{damp} also be set to a finite value for single top processes?

- Parton shower
  → correlate

- NLO-subtraction method
  → correlate

If different generators for different processes are used, the story can be different.

Summary

- Extend studies of scale / h_{damp} variations to new unfolded data

- Still very limited phase space for data comparisons
  - CMS is working on implementing their unfolded results in Rivet
  - Final 8 TeV analyses still to come

- Developed approach to access radiation systematic for NLO generators (particularly for the POWHEGBOX)
  - Variations of ren/fac scale by factor of 2 up and down combined with variation of h_{damp} by a factor of 2 seems reasonable.
  - Variations within parton shower doesn't show a big effect on the observables considered.

- Correlations of systematics between single top and $t\bar{t}$ are being investigated $\rightarrow$ Open questions about radiation systematic

- Started to explore new C++ generators, aMC@NLO + Herwig++ looks promising, Powheg+Pythia8 is outside the uncertainty of Powheg+Pythia6 for the jet multiplicity, but only for higher jet bins.

Backup

Suggestion from theory

Uncertainty estimates

- For observables that have NLO precision, the theory/generator uncertainties can be estimated by
  - Independent renormalisation and factorisation scale variations
  - PDF error sets (preferably following the PDF4LHC agreement)
  - Matching an NLO computation to at least 2 different parton showers
  - These PDF and scale variations can be obtained via reweighting in aMC@NLO and POWHEG, not yet possible in Sherpa.
- For observables that do not have NLO precision, further uncertainties are coming from the shower starting scale ("Power" or "Wimpy" shower). Currently these cannot be approximated with the (a)MC@NLO program, but not really relevant because why use an NLO+PS computation for these observables in the first place?
  They can be estimated more correctly in the NLO Sherpa program.

Rikkert Frederix

7

Scale variations for NLO generators

μ_f	μ_r	Shower
0.5	0.5	radHi
1	0.5	radHi
0.5	1	default
1	1	default
2	1	default
1	2	radLo
2	2	radLo