Monte Carlo generators for the LHC

1. Generators and experimental setup
2. Tuning and validation
3. Next steps towards data

Roberto Chierici
CNRS/IPN Lyon

on behalf of the ATLAS and CMS collaborations
(but slightly biased on CMS)

LISHEP2009, Rio de Janeiro, Brazil, January 2009

Generators and experimental setup

- Physics and software requirements
- Modern event generators and tools
- Main choices by CMS and ATLAS
- Input settings

Introduction

Monte Carlo event generators for HEP are used to simulate and predict, given certain initial conditions, what happens in a particle collision.

Mandatory tools for experimentalists. Experimentalists must be able to:

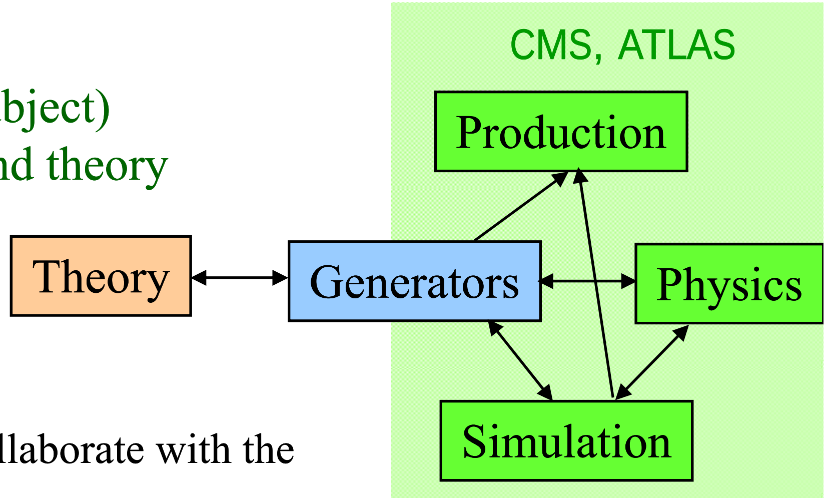
0. Plan experiments
   - how should I design my detector and the global acceptance to be able to see, at 95% CL, a signal of new physics in the channel X and expected cross section Y, with a total integrated luminosity L?
1. Understand what they see in data. Not as straightforward as one might believe:
   - is what I see predicted by simulation?
   - am I seeing a physics effect, or a detector effect?
   - is it something expected? Should I claim something?
   - is it a bug? A bad tuning? …
2. Making a measurement
   - motivate, validate and tune analyses strategies and cuts
   - model the expected backgrounds (and subtract them to data, eventually)
   - know the expected signal efficiencies

In all this, we must take care of using those tools which are the best suited (modern, accurate and flexible) for the particular physics (signal and backgrounds) under study.

Introduction

- This period is crucial for refining the choices related to Monte Carlo productions at the LHC.
- It is the moment for the first Monte Carlo generations for data taking: a sound choice of generators and massive event generations are of extreme importance to prepare the analyses and shape the strategies for data.

- Monte Carlo generators (and people working on the subject) represent the natural interface between experimental and theory communities.
- Setting up generators in an experimental community implies both software and physics aspects.



PH

SW

- collect the physics requirements, communicate/collaborate with the theory community
- build a coherent generation for doing physics
- steer the generator tuning with data
- interface, validate, maintain the generator in the experiment's software framework
- plan the event generation from the different physics groups' requests

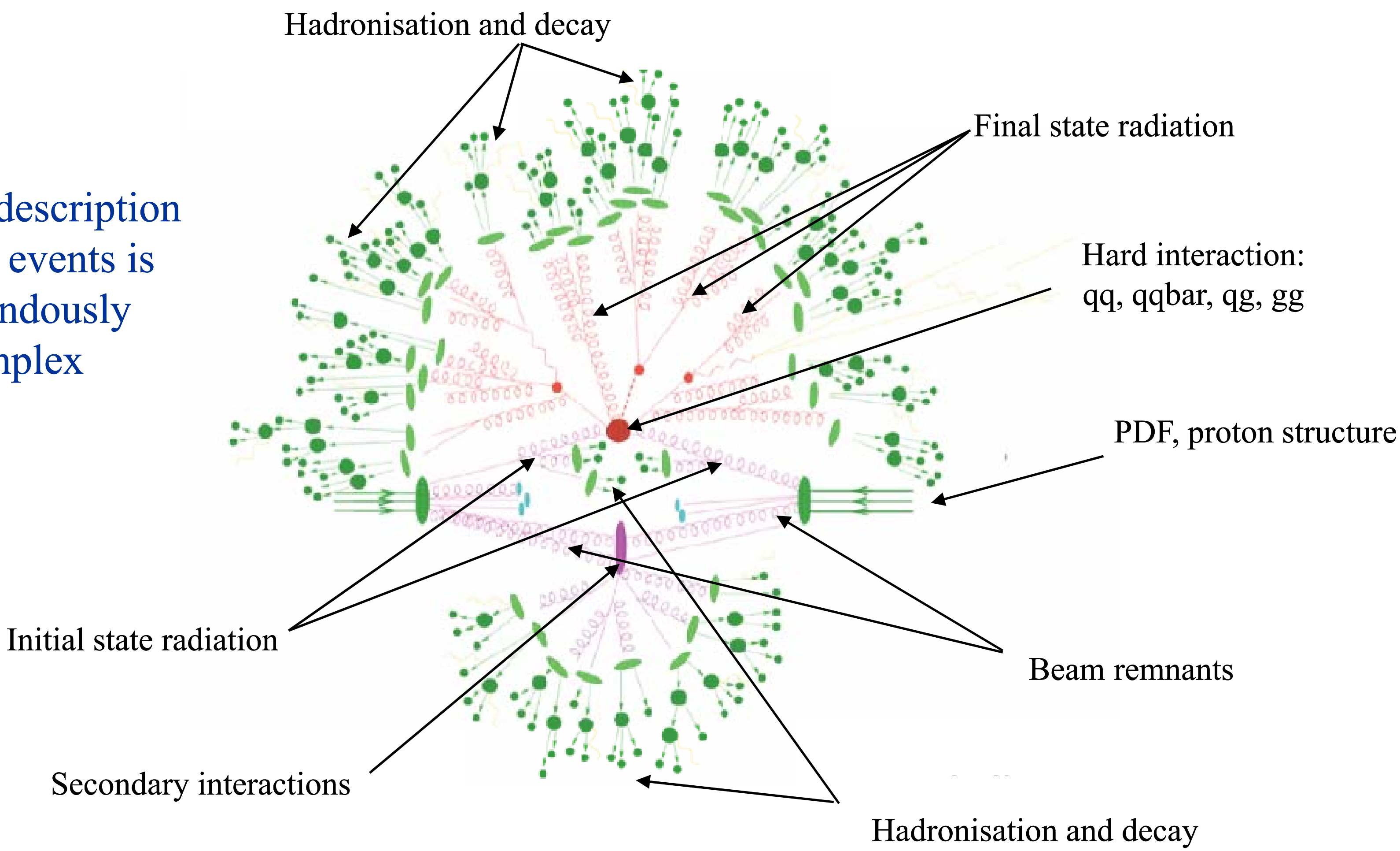
This talk reports about the present experience of CMS and ATLAS in generators for SM and beyond.

→ This talk is not a generator review.

→ This talk is biased on generators for description of high p_T physics

Physics requirements

The MC description of LHC events is tremendously complex



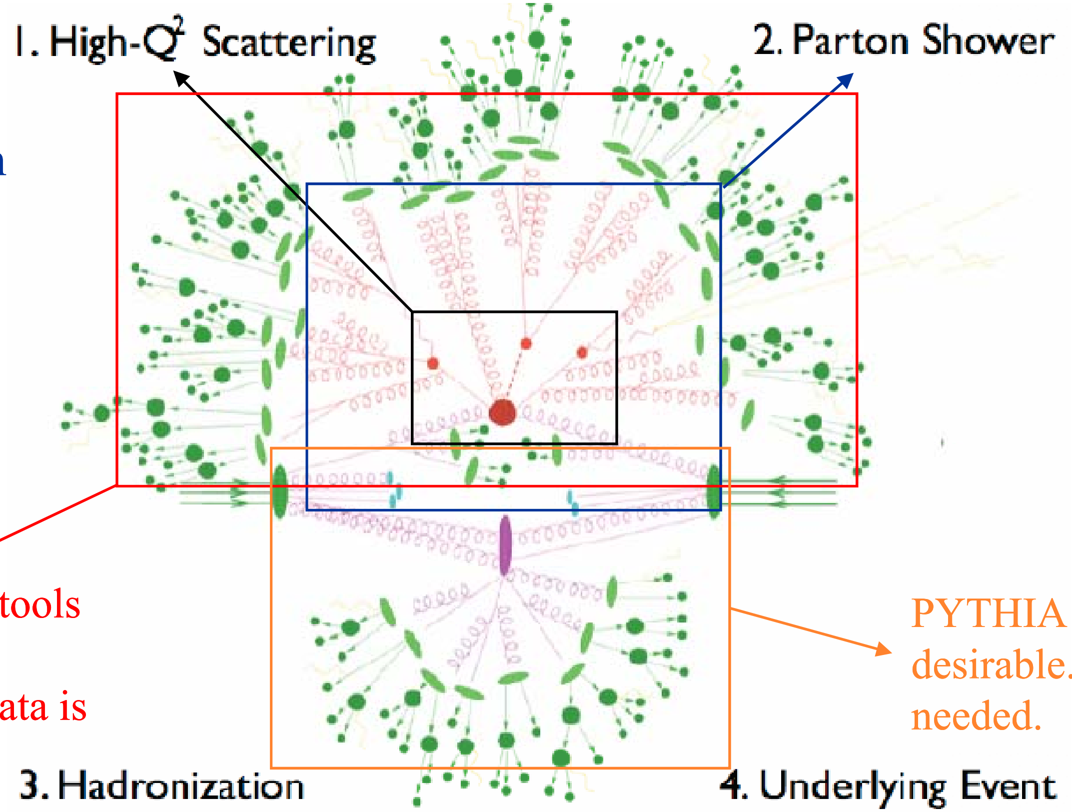
This is a schematization to be able to cut down the problem in pieces and model them in a different way. The "pieces" are correlated !

Physics requirements

Extra gluon emission described with ME at the highest possible order (+matching). Spin correlations needed.

Interface to dedicated tools (PYTHIA/HERWIG) is needed. Tuning with data is also needed.

The MC description of LHC events is tremendously complex



Interface to dedicated tools (PYTHIA/HERWIG) needed. Tuning with data is also needed.

PYTHIA MPI. HERWIG/JIMMY desirable. Tuning with data is needed.

Other desirable features, from the experimentalist's viewpoint:

- output in the Les Houches standard format
- as much complete as possible coverage of SM phase space
- user friendly inclusion of new physics signals
- support ☺

Event generators and decay tools at the LHC

Both ATLAS and CMS try to use as many event generators as reasonable:

- Latest FORTRAN event generator version, moving to the newest C++ versions

  PYTHIA6, HERWIG

  ✓ extensively used as baseline event generators (SM, BSM) and for PS/fragmentation from external MEs

  PYTHIA8, HERWIG++

  ✓ being used in parallel. First central productions performed in CMS

  SHERPA

  ✓ interfaced in both experiments. Extensively used in ATLAS already

- Several decay packages and add-on are used

  TAUOLA (τ decays)

  ✓ work with all event generators, used where the description of τ decays is relevant

  PHOTOS (QED corrections)

  ✓ work with all event generators, used where real QED description in simple processes important

  EvtGen (for B hadron decays)

  ✓ work with all event generators, used only in signal description

- Several ME parton level generators are used for LO, NLO, HLO QCD description of most of the SM physics processes:

  Enormous progress in the recent years → next slides

QCD emission: ME vs PS

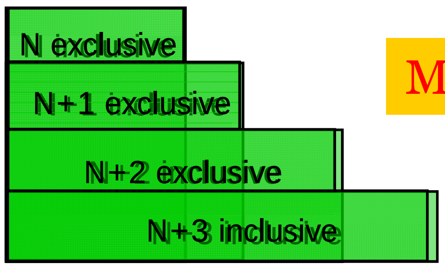
No generator adequately reproduces the physics processes for the whole CMS program
Essential to understand which techniques are applicable to which kinematic regime.

- Parton Shower: infinite serie in α_S keeping only singular terms (collinear approx.):
  - ☞ Excellent at low p_T, with emission at any order, simple interface with hadronization
  - ☞ Large uncertainties away from singular regions
  - ☞ To be used for soft (compared to signal scale) jets.
- Fixed order matrix elements: truncated expansion in α_S
  - ☞ Full helicity structure to the given order
  - ☞ To be used for hard (compared to signal scale) jets.

High jet multiplicity events are bound to be better described by ME.

- NLO generators with PS now exist: predict correct normalization and shape at NLO
- Techniques for correcting the first emission in PS to get the shapes of the first additional hard jet in the event correct also exist
- What for higher multiplicity (HLO)? How to match PS and ME?
  - Cutoff? Where? How to avoid double counting? (ME_N+PS has parts of ME_{N+1}+PS)

1. ME_N+PS
2. ME_{N+1}+PS
3. ME_{N+2}+PS
4. ME_{N+3}+PS



MLM

- (cone-)cluster showered event → njets
- match partons from the ME to the clustered jets
- if not all partons are matched, discard the event

Works independent of the generation procedure

Beware: PS(today) ≠ PS(yesterday). Tunings need to adapt to the choice of the matching

Parton level generators at the LHC

- PYTHIA, HERWIG: 2→2, 3 LO ME
  - ✓ reference for QCD production, (PYTHIA also for BSM) in both ATLAS and CMS
- Alpgen, MadEvent/MadGraph, SHERPA, HELAC: 2→2(+3, 4) LO ME + ME-PS matching
  - ME-PS matching (MLM, CKKW) to higher LO diagrams (up to four additional q,g jets)
  - ✓ ALPGEN, MG/ME reference generators for high p_T physics in both ATLAS and CMS
  - ✓ SHERPA used for boson production in ATLAS, being validated in CMS
  - ✓ Interest in HELAC in CMS
- POWHEG, MC@NLO: NLO ME, many process implemented, in continuous evolution
  - With ME-PS matching to NLO
  - ✓ MC@NLO reference generator for top and electroweak physics in CMS and ATLAS
  - ✓ Extreme interest in POWHEG in CMS, under validation
- AcerMC, CompHEP, Phantom: 2→n LO ME
  - ✓ AcerMC reference generator for top physics in ATLAS
  - ✓ CompHEP used for multi-jet and some BSM in ATLAS
  - ✓ Phantom used for 6-fermion final states in CMS (top physics, boson scattering)
- TopRex, SingleTop: dedicated 2→2, 3 LO ME
  - ✓ used for top physics in the recent past in both ATLAS and CMS

Example of generator integration: CMS

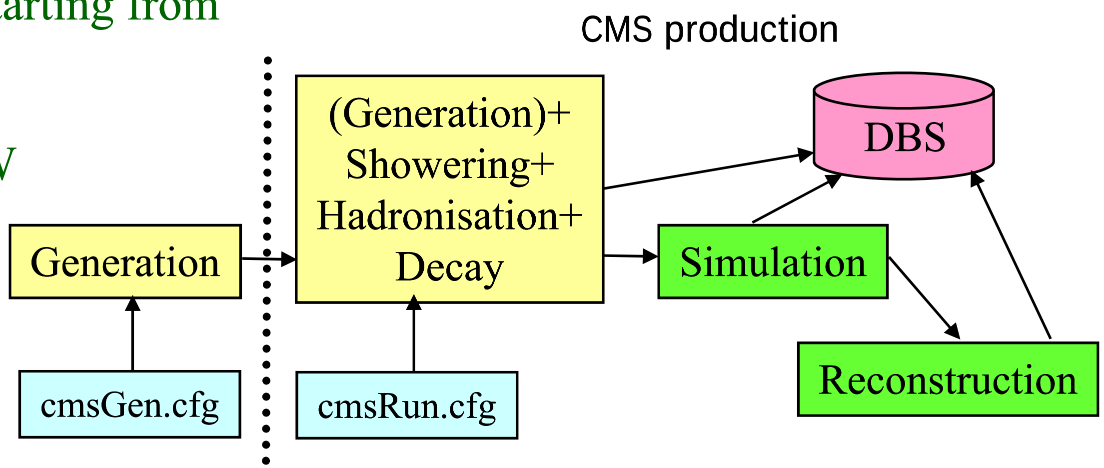
- The physics program at the LHC is very rich!

✓ pp General purpose: Pythia6, Herwig6, (Pythia8, Herwig++), Sherpa
✓ pp HLO: Alpgen, MadGraph, Helac, Sherpa
✓ pp NLO: MCatNLO, POWHEG
✓ pp Others: CompHEP, TopRex, Phantom
✓ Diffractive physics: Pomwig, Exhume, EDDE
✓ Decayers: EvtGen, Tauola, Photos
✓ Heavy Ions: Hydjet, Pyquen
✓ Detector specific: Cosmics, particle guns, beam halo, beam-gas
✓ New physics specific: Charybdis

- A lot of software integration work is needed to properly integrate many external generators to the experiment SW (and distributed production) scheme.
  - → Standards like intermediate Les Houches Event (LHE) parton level files help the task.
  - → Using externals from the LCG project GENSER guarantees code blessed by the authors, ported/tested on the necessary platforms and a first systematic basic sanity check

- CMS prefer generators integrated in the experiment software, that can directly be used in production, but now the possibility of starting from externally produced LHE files exist.

- The generator interfaced to the CMSSW must undergo:
  ✓ a technical validation to show its ability to run standalone and in production
  ✓ a physics validation wrt a generator with similar physics content



Generators setup and validation

soft
↓
hard

- MB-UE tuning
- MPI and PYTHIA
- ME matching vs PS for W+jets
- ME-PS vs NLO in the top sector
- A comment on theory uncertainties

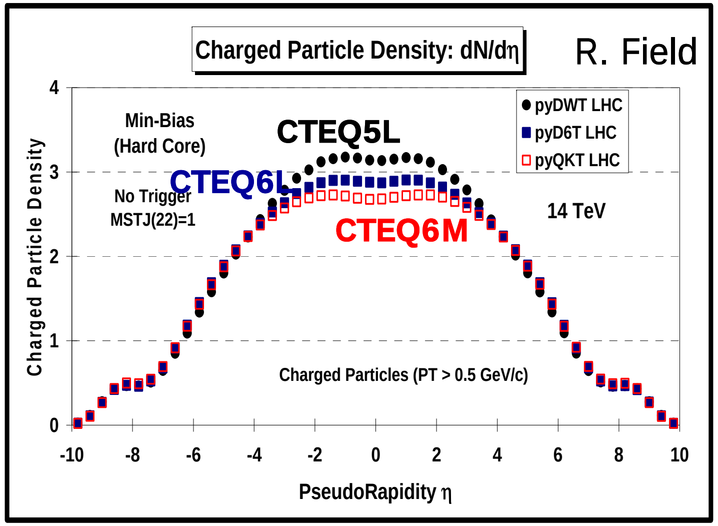
General reference MC settings

Tools for shower+hadronisation: PYTHIA primary choice (6.409), HERWIG (6.510) useful crosscheck (and used for MC@NLO). Growing interest for Sherpa.
PYTHIA: CMS uses old Q^2 ordered shower, ATLAS the new one, p_T ordered.

PDF settings: LO PDF CTEQ6L1 (LHAPDF=10042).
NLO PDF used only for NLO generators and for determining errors.
Need iterations to tune PDF with the use of LHC data.
Consensus in the theory community on using Modified Leading Order PDFs (LO*) for all LO calculations, or calculations including leading order matrix element corrections

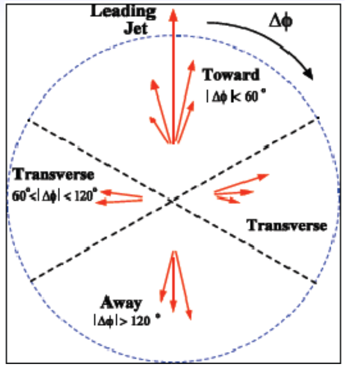
PS radiation and fragmentation (b, light quarks)
settings: used from LEP tunings (see for instance CMS note 2005/013).
Need to re-tune to LHC data. Maybe a different approach necessary when using external ME to Parton Showers?

MB/UE tuning: either by the use of Jimmy (for HERWIG) or, for PYTHIA:
ATLAS: via the new MPI interleaved model, tuned to published Tevatron data
CMS: via the new D6T tuning (by R. Field), using CDF data and ATLAS extrapolation at LHC energies.
MB/UE ↔ PS ↔ Fragmentation !



Roberto Chierici

Underlying event tunings

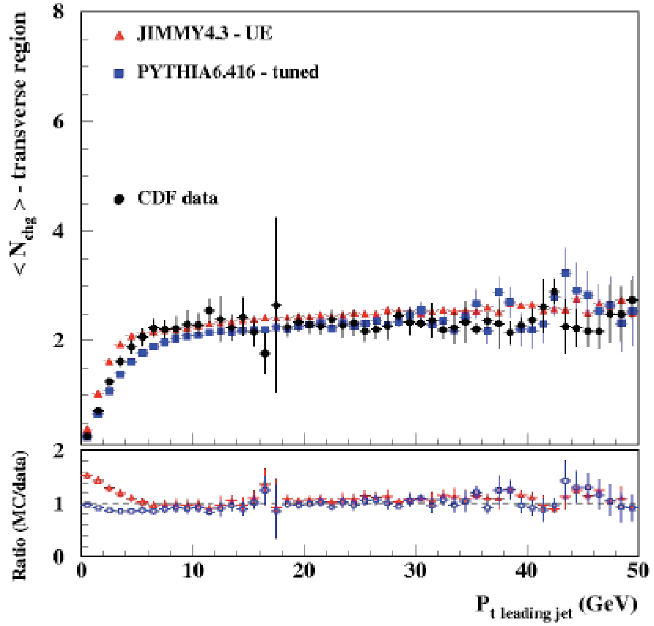


Tune on data distributions (differential in η, φ) of $\langle N_{chg} \rangle$ and $\langle \Sigma P_T \rangle$ in the region transverse to the leading jet

Use CDF data at 630 GeV and 1800 GeV (important for energy extrapolation)

$N_{jets} > 1$,
$|\eta_{jet}| < 2.5$,
$E_T^{jet} > 10$ GeV,

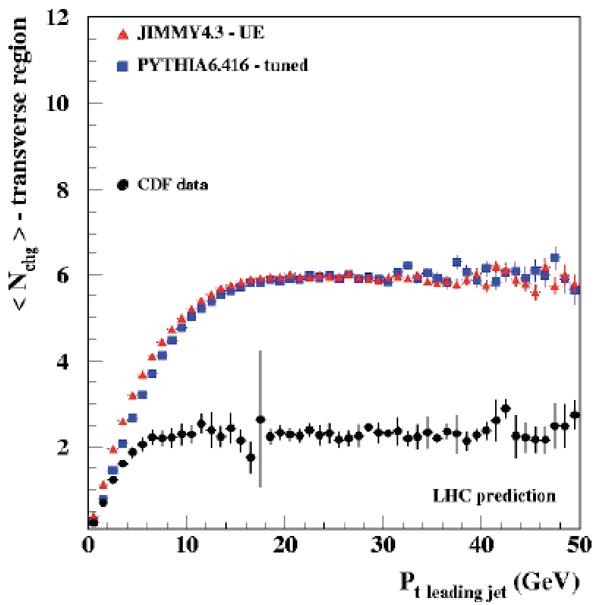
$|\eta_{track}| < 2.5$,
$p_T^{track} > 1$ GeV/c

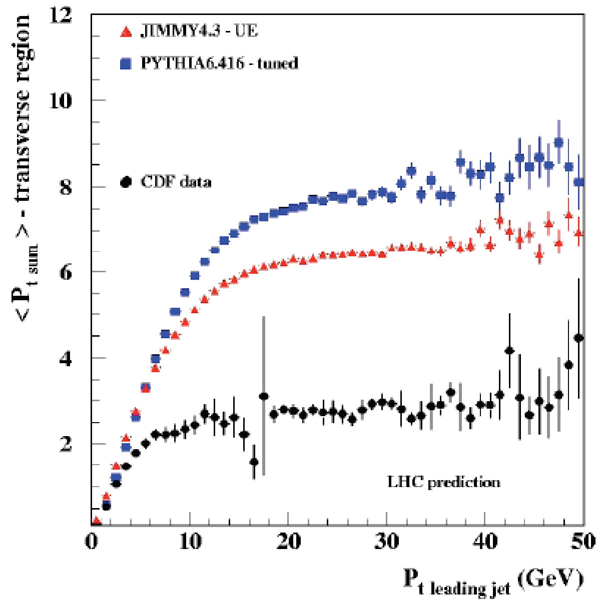


- Leave radiation (IS/FS) as default.
- Tune colour reconnection parameters, matter distribution parameters.
- Main parameter is the p_T cutoff

$$\sigma(\widehat{P_T}) \rightarrow \sigma(\widehat{P_T}) \cdot \frac{(\widehat{P_T})^4}{((\widehat{P_{T0}})^2 + (\widehat{P_T})^2)^2}$$

Extrapolation at LHC energies





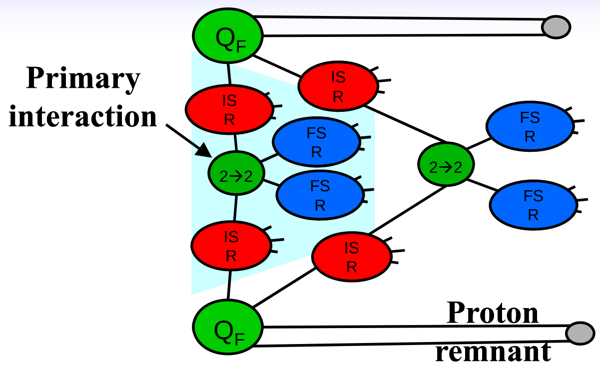
Roberto Chierici

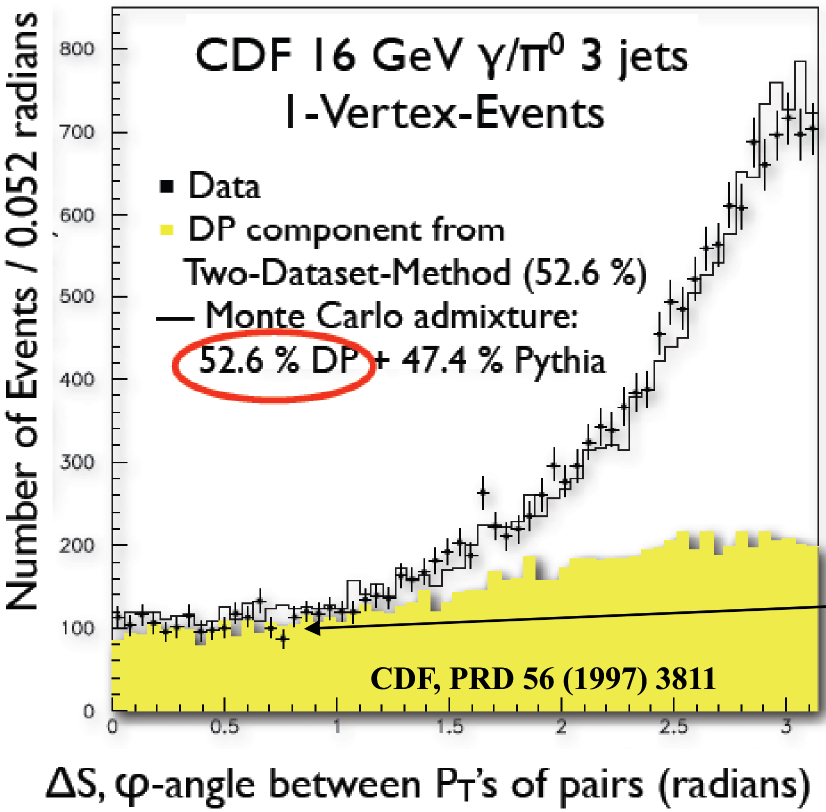
MPI

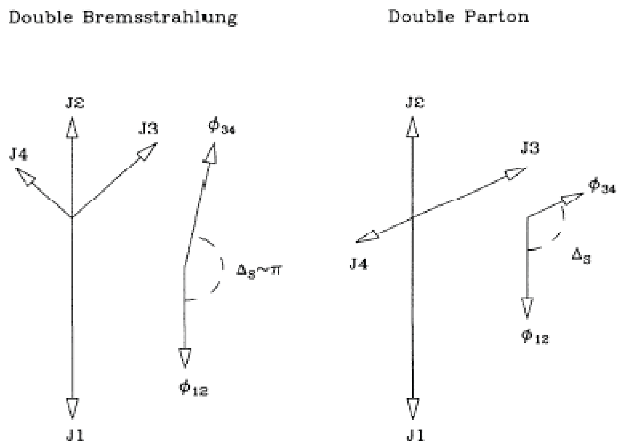
• QCD diverges at low p_T, multiple parton-parton collisions should occur in the same event: 4→4, 3→3, 3→2

• Additional non-perturbative phenomena? (BE correlations, Colour Reconnections,…)

• Evidence at the Tevatron from γ+jets







azimuthal angle between the p_T vectors of γ+j and j+j

Need the inclusion of double parton interactions to correctly describe data !

New PYTHIA has interleaved evolution of PS and secondary interactions (with PDFs derived from sum rules)

MPI and PYTHIA8 validation

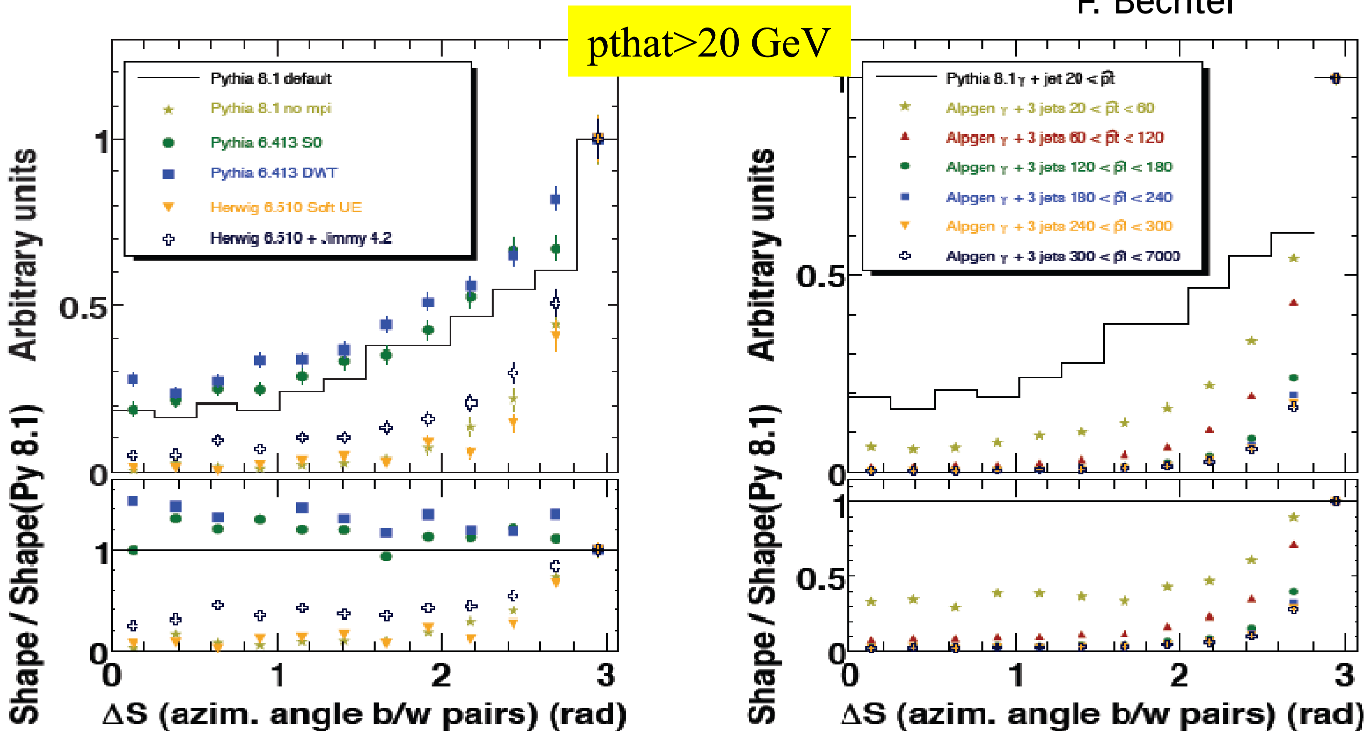
- Compare PYTHIA8 with/without MPI with PYTHIA6, HERWIG, ALPGEN for γjjj
- Simple selection: γ from MC, jets with midPointCone, $p_T(\gamma, jets) > 30$ GeV, $\Delta R_{ij} > 0.8$, choose pairing on the basis of the minimum sum of relative p_Ts

- PYTHIA 6.413, DWT or S0 tuning (more colour reconnection)
- PYTHIA 8.1 default (tuning~S0), with and without MPI

- HERWIG 6.510, JIMMY 4.2 (and no MPI)
- ALPGEN γ+3j inclusive, PS with PYTHIA 6.409, DWT tuning

F. Bechtel

- Large difference in HERWIG and PYTHIA
- MPI-no MPI give a sizable difference
- A better ME description of the process has no effect
- The S0, DWT tunings mimic the presence of MPI



PYTHIA8 now for the first time used for official MC production in CMS

Roberto Chierici

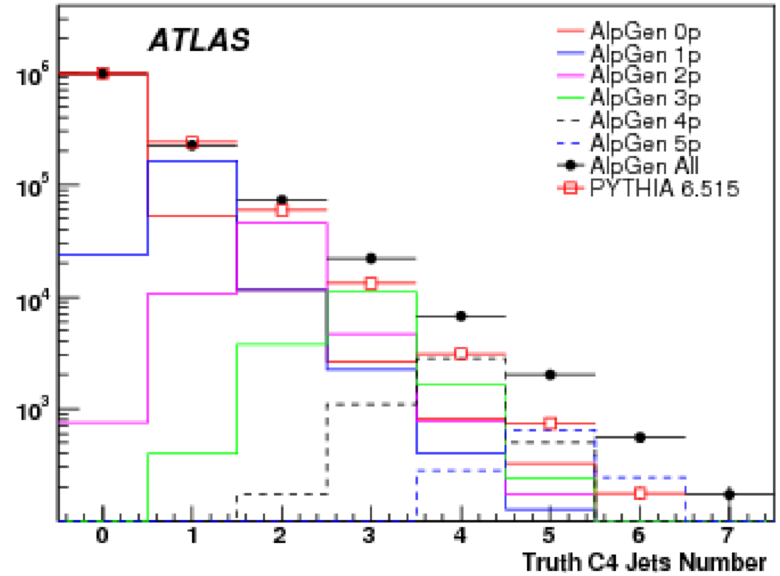
Matched ME-PS vs standard PS: W+jets

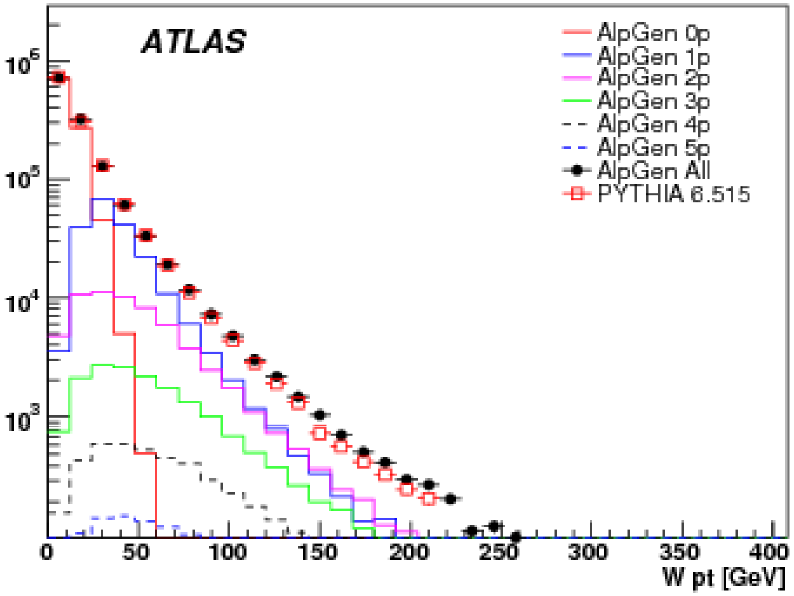
- Compare ALPGEN W+jets and PYTHIA 6.5 W
(normalization to 100/pb and ALPGEN cross-section)

- Good agreement with PYTHIA for all W/lepton observables.
- Deviations in the description of the hard hadronic environment.
- Largest discrepancies where W+jets is background to something else (example: top).

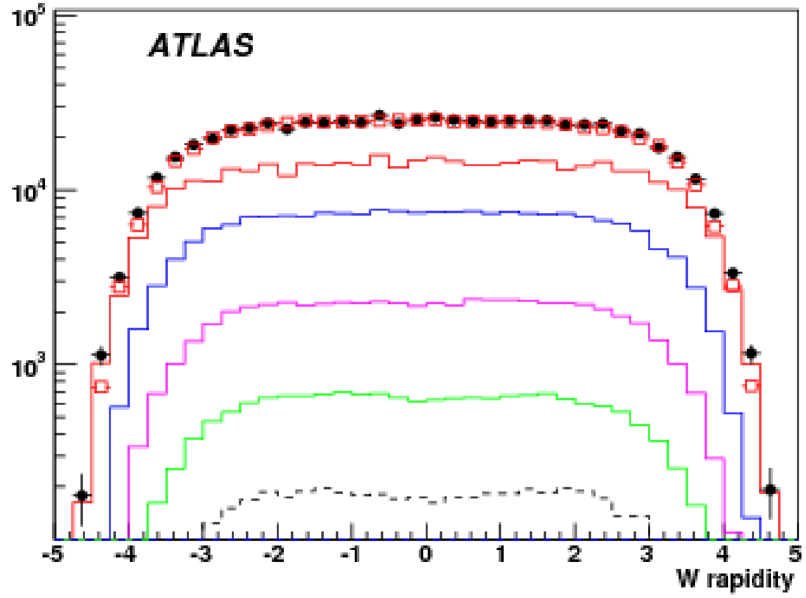
MC N C4Jet (pT>20GeV)

O. Jinnouchi



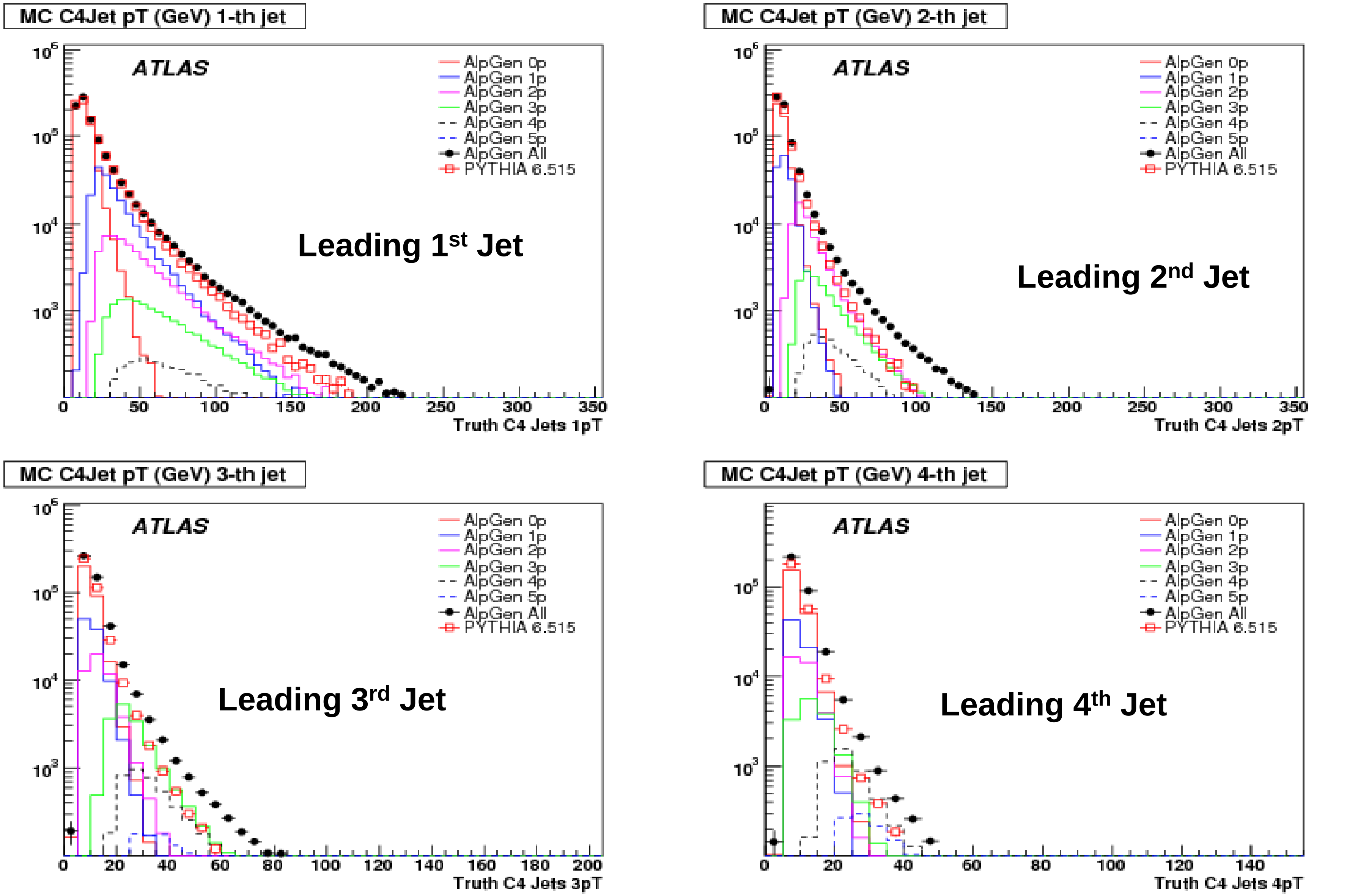


MC W rapidity (pT>10GeV)



Matched ME-PS vs standard PS: W+jets

Large differences in transverse variables related to high transverse energy radiation.



Validation after analysis: W+jets

- It is important to know the modeling uncertainties on predictions: how well is W+jets described? The question is particularly interesting in the many-jet regions
- Use ALPGEN and vary matching scale and minimum ΔR parton separation to study the sensitivity in the analyses.
- Simple selection: 1 lepton pt>20 GeV $|\eta|<2.5$
  Missing E_T >20 GeV
  Jet algorithm is cone 0.4; 4 jets >20 GeV $|\eta|<2.5$; 3 jets >40 GeV

Matching parameter	σ vis(%) W+0p	σ vis(%) W+1p	σ vis(%) W+2p	σ vis(%) W+3p	σ vis(%) W+4p	σ vis(%) W+5p inc	Relative difference
pt10,r07	0.4%	0.1%	0.3%	2.9%	15.7%	80.5%	1.48
pt10,r03	0.4%	0.1%	0.4%	0.5%	7.2%	91.3%	1.86
pt20,r07	0.8%	0.1%	1.6%	12.2%	42.0%	43.0%	0.87
pt20,r03	***0.4%***	***0.1%***	***0.7%***	***2.8%***	***34.4%***	***61.3%***	**1**
pt40,r07	1.0%	0.2%	7.6%	45.7%	35.4%	9.4%	0.60
pt40,r03	1.2%	0.3%	6.1%	36.7%	38.1%	16.5%	0.53

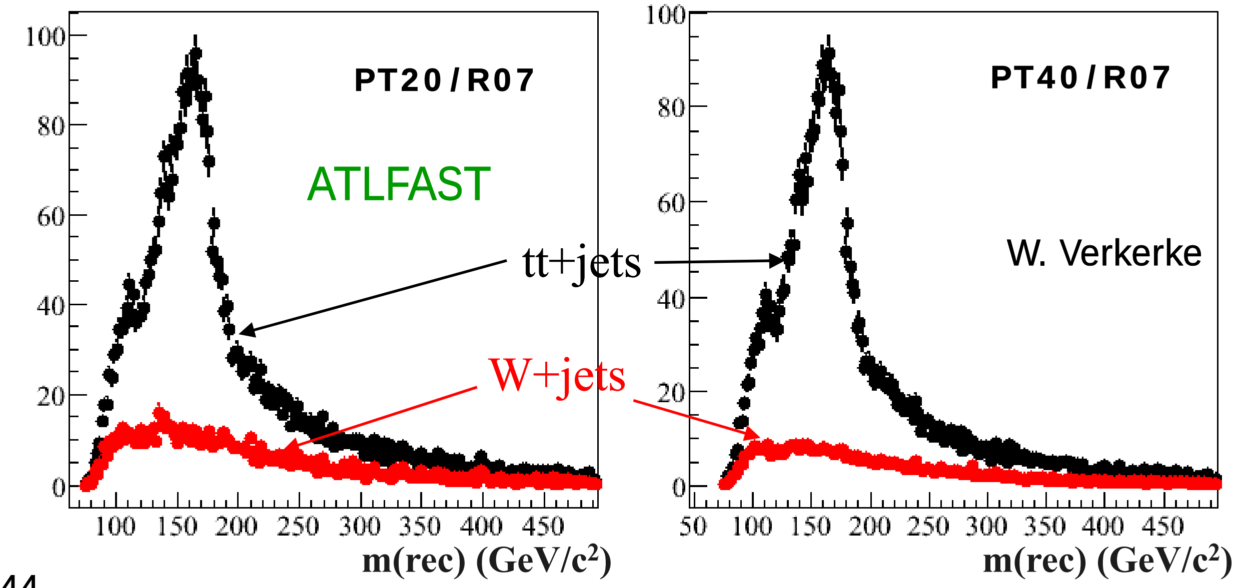
Accepted σ after cuts, normalised to the pt20,r03 point
(the error on the relative differences is a few %)

- Relative importance of different parton multiplicities as expected. Total accepted cross-section after selection varies up to a factor 2!

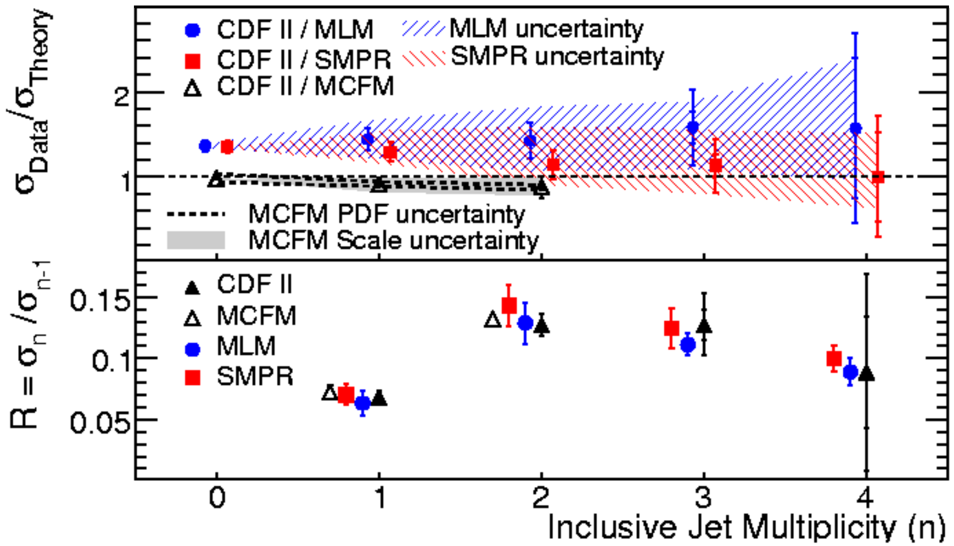
Validation after analysis: W+jets (cont)

Differences in angular shape and normalization for extra jet production

1 lepton pt>20 GeV |η|<2.5
Missing E_T >20 GeV
Jet algorithm is cone 0.4
4 jets >20 GeV |η|<2.5
3 jets >40 GeV



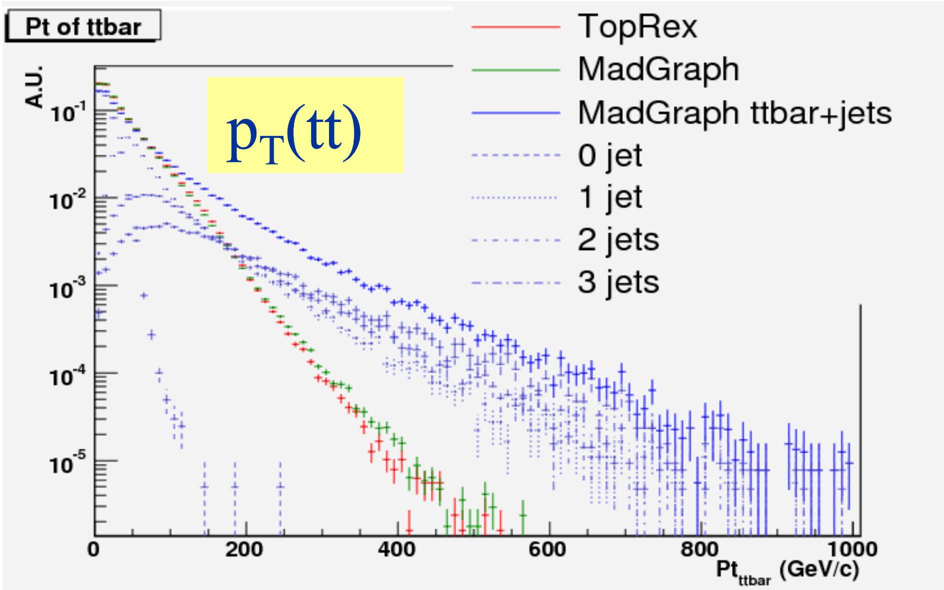
CDF Collaboration, arXiv:0711.4044



- Large theory uncertainties, and not enough data at the Tevatron to further constrain the Monte Carlo in that region
- Tunings are important, need to use LHC data

Validating ME-PS vs PS in top physics

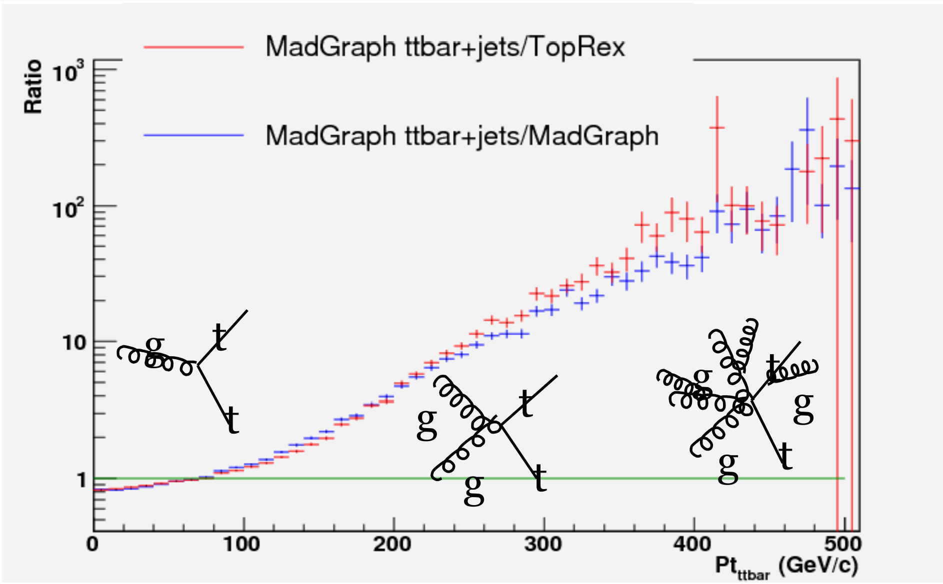
Large differences in transverse variables related to radiation

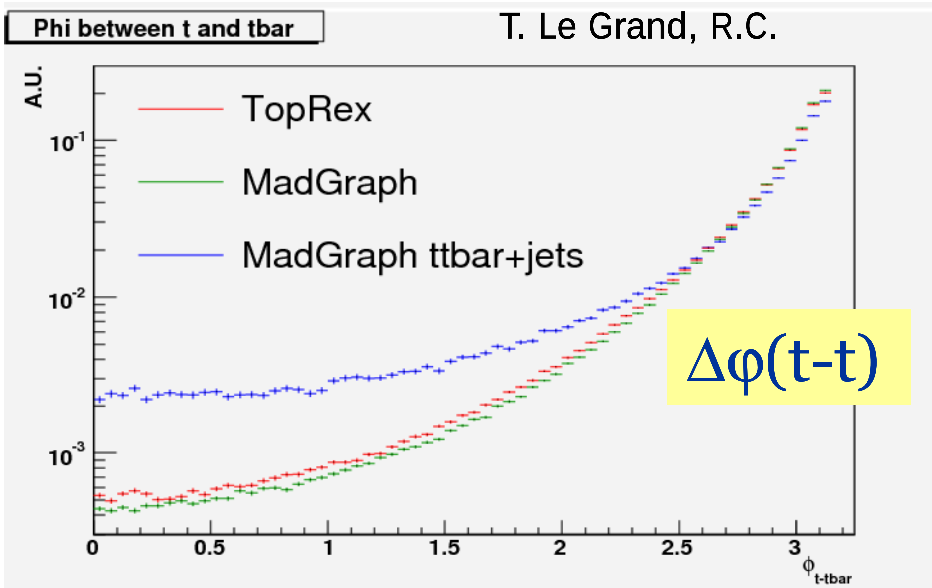


- Large effects at high p_T(tt)=p_T(radiation)
- Average p_T(tt)~60-70 GeV !
- 40% probability that a tt system recoils against a radiation larger than 50 GeV

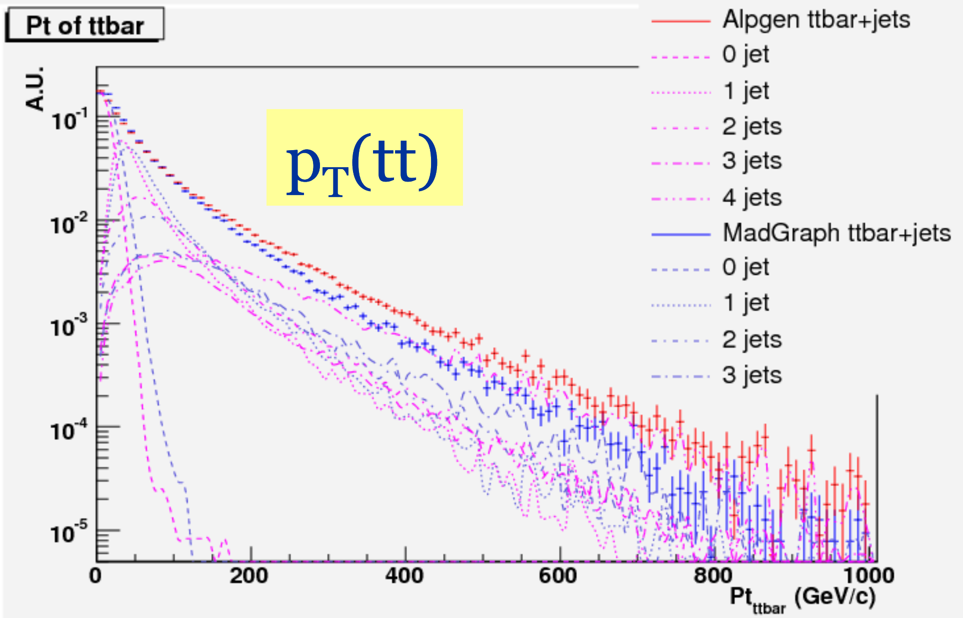
→ effect on reconstruction

→ Mandatory to use the same strategies for physics backgrounds like W/Z+Njets





ALPGEN vs MadGraph with matching

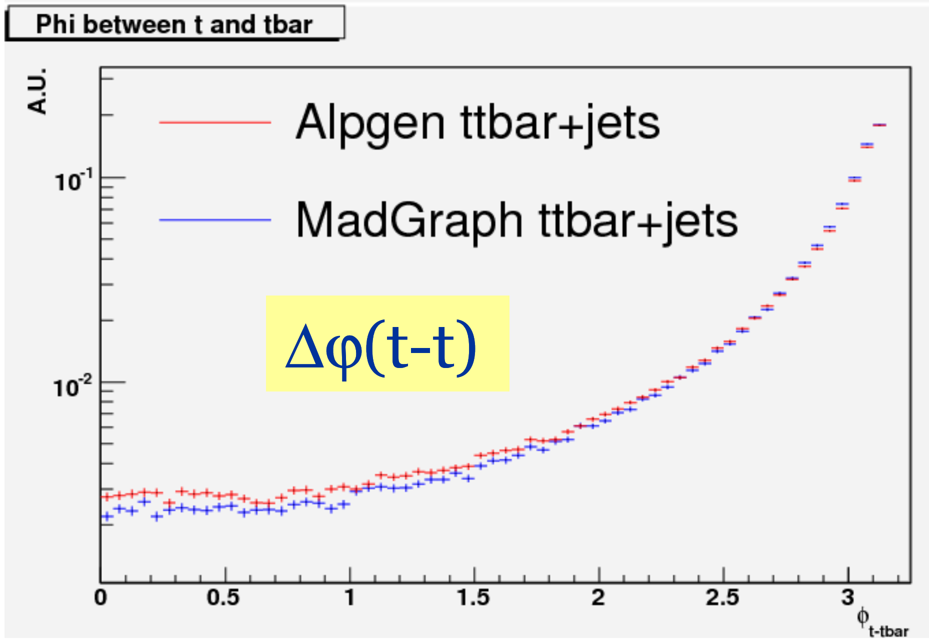


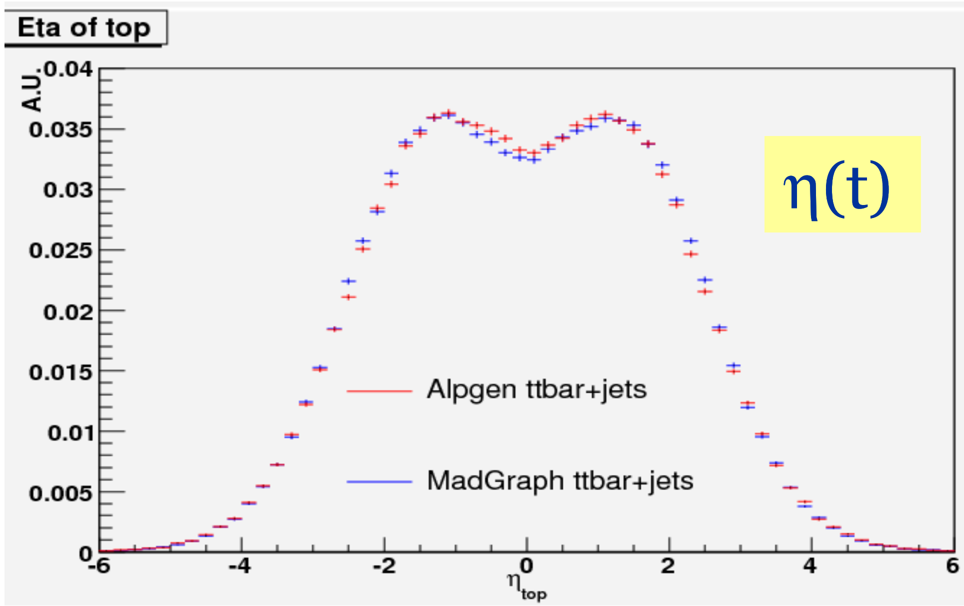
ALPGEN and MadGraph differ by at most 50% on the p_T prediction

Important to understand the residual theory error on the distributions:

- Effect of renormalisation and factorisation scales on the predictions
- Effect of the chosen ME-PS matching scale

Excellent agreement on other variables



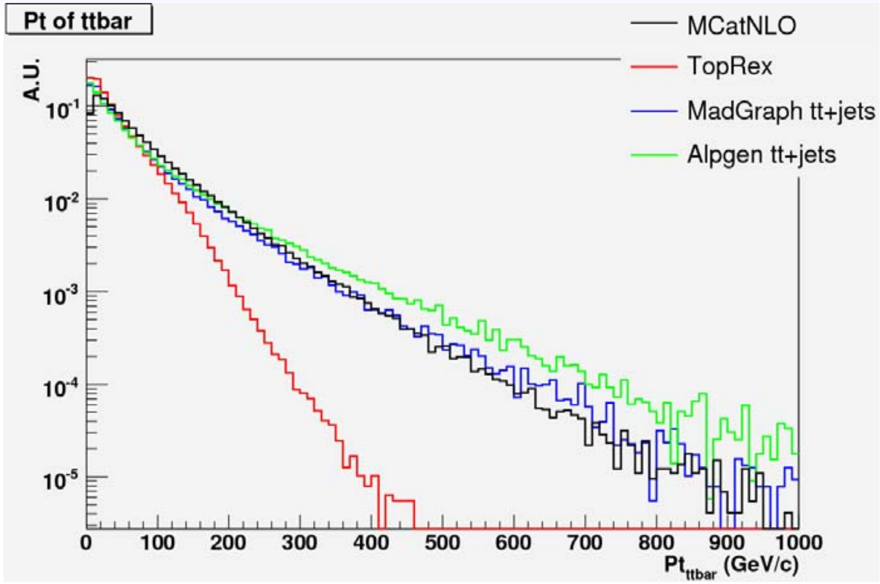


Roberto Chierici

Matched vs MC@NLO

Comparisons to MC@NLO ongoing in CMS.
Conceptual difficulties in interpreting the results:

- Non perturbative part treated by HERWIG/JIMMY.
- Should compare to a matched tt0j(exc)+tt1j(inc) production

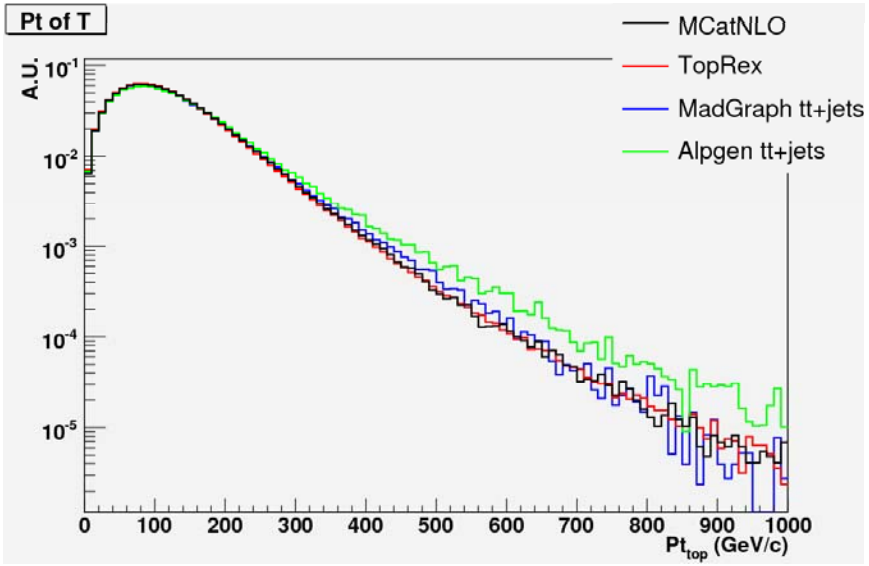


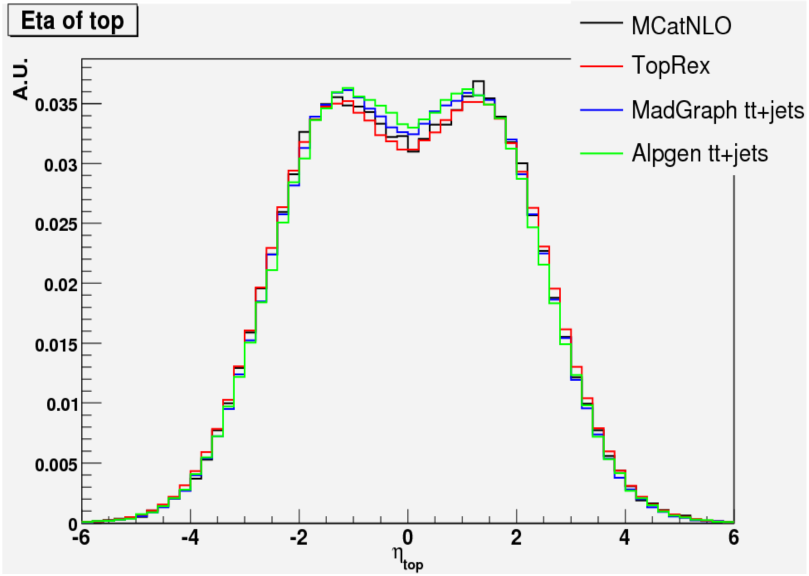
Still a very important step in understanding high p_T radiation and increase our confidence in the process description. Also gives indications on:

- Relative importance of first emission
- Normalization
- Indication of systematic errors associated to the description of radiation.

Preliminary generator level study

Essential agreement in the p_T(tt) tail. Good agreement in other distributions.





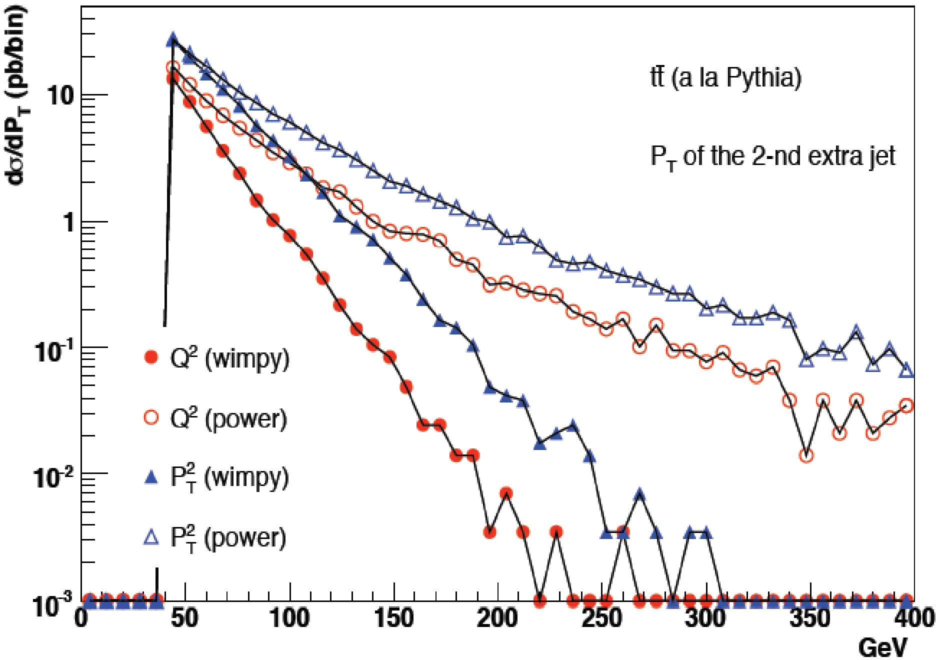
The importance of ME tools

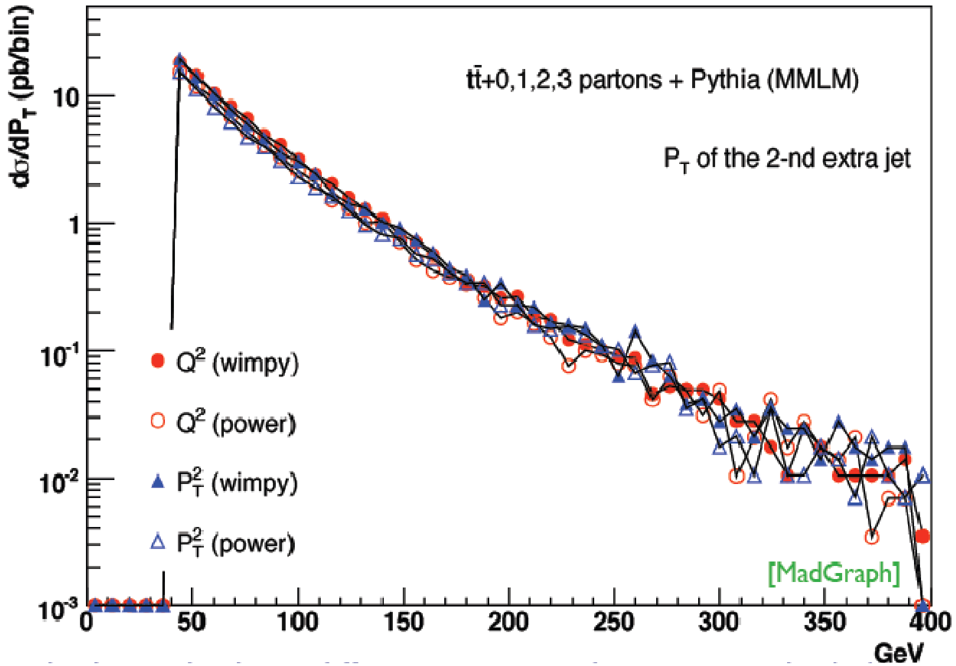
A parton shower is by construction an highly tunable tool.

For a matched calculation the effect of tunings in the hard regions are less relevant because this is described by the Matrix Element

- more predictive power
- less sensitivity to the MC tunings
- systematic errors due to theory/modelling are smaller (should include theory uncertainties of the matching itself)

F. Maltoni, top 2008



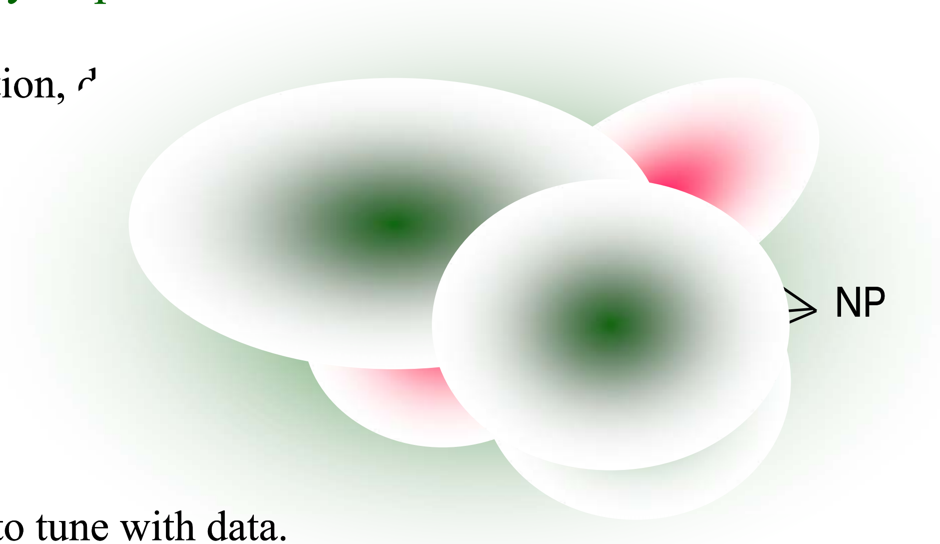


Roadmap towards physics

- Present and future Monte Carlo generations
- Strategies and current productions at CMS
- Conclusions

Generation strategy

• Take care of the SM as correctly and uniformly as possible
  → Satisfactory ME description.
  → Coherent interface to showering, fragmentation, d
  → Tuned underlying event. Use LHC data.
  → Uniform choice for input parameter settings and PDFs



• Add generator redundancy in crucial portions of phase space
  → One prediction is always not enough: HLO vs NLO vs different approaches.
  → Different interface to showerings – prepare to tune with data.
  → Different settings to study systematics (tunings, PDFs,…).
  → Sensitivity of analyses and reconstruction methods to "theory/modeling" effects.

• Add new physics samples
  → Main SUSY and BSM points to train analyses

• Determine tails
  → Study what tails are most interesting and refine studies there.

⇒ The use of an as much coherent (IPS, PDF, cuts) as possible set-up will ease enormously the tasks of the analyses
  • compare SM and BSM on a similar (equal) footing. Include SM-SM and even SM-BSM interference !
  • disentangle detector/simulation effects from the physics input to the generation
  • speed up things

The CMS way

- A large Monte Carlo production effort is needed in preparation for the data taking period
  - → trigger studies, definition and overlaps between primary datasets
  - → full SM coverage for training the analyses, especially for QCD studies and rejection
  - → systematic studies where relevant, redundancy of generators for validation
- We want to use the right tools in the right portions of the phase space
  - → calculations can sometime become very CPU intensive (e.g. W+4jets with ME)

⇒ Better handle large Monte Carlo production centrally

→ reduces the possibility of errors and effort duplication

- How to make a coherent generation?
  - → partition the phase space, avoiding double counting of processes or duplication of MC samples
  - → use as much as possible a reference generation setup
- CMS way: $\text{MadGraph}_{\text{+PYTHIA}}$ as a reference ME generator for SM and BSM.
  - + can treat all phase space coherently, including SM+BSM interferences
  - + do not give up higher leading order matched QCD contribution
  - + flexibility of including any new physics
- Use $\text{ALPGEN}_{\text{+PYTHIA}}$ and $\text{MCatNLO}_{\text{+HERWIG}}$ as primary comparisons for the analyses

- Definition of different portions of phase space in collaboration with the MG/ME team, with theory-validated LHE files and corresponding binaries for Monte Carlo productions.
  - → Agree on the file contents (processes, cuts, settings)
  - → So far all SM is covered, BSM to follow

From: http://cp3wks05.fynu.ucl.ac.be/twiki/bin/view/Library/MadGraphSamples

Recent production efforts in CMS

- Very large fast simulation production (~600M events), corresponding to 3-6 months of data taking ay 20% efficiency and 300 Hz rate to storage.
  - More than 400M are QCD. H_T binning for jets and γ+jets needed to enhance tails.
  - Matched productions organized in "cocktails", with all parton multiplicities together in the same dataset and with the right proportion.

MB	PYTHIA	100M	tt+jets	MadGraph$_{+PYTHIA}$	10M
QCD jets	MadGraph$_{+PYTHIA}$	217M	W+jets	MadGraph$_{+PYTHIA}$	63M
MB bb	PYTHIA	21M	Z+jets	MadGraph$_{+PYTHIA}$	5M
QCD bb	Madgraph $_{+PYTHIA}$	23M	γ+jets	MadGraph$_{+PYTHIA}$	35M
QCD jets	PYTHIA	45M	γ+jets	PYTHIA	4M
QCD e.m.	PYTHIA	33M			

- A full detector simulation, corresponding to a total of 200M events.
  - ✓ Min bias (40 Mevt), QCD (light/b) (30 Mevt) and γ+jets (10 Mevt)
  - ✓ Electrons/muons from b or in-flight decays (25/40 Mevt)
  - ✓ Drell-Yan and Onia (10 Mevt)

  } PYTHIA $p_T$hat bins for QCD

  - ✓ QCD (light/b) (plus jets) (30 Mevt) and γ+jets (5 Mevt)
  - ✓W/Z (+j) (10 Mevt), others EWK (VVj, Wc, VQQ, γ*+j, Z→νν) (10 Mevt)
  - ✓Top (2 Mevt)

  } MadGraph, ME-PS matching

  - ✓ +500K γ+jets with PYTHIA8
  - ✓ 5M QCD with HERWIG++
  - ✓ +700K single diffractive with POMWIG
  - ✓ 1M EWK+Top with MC@NLO

  } Generation for validation

- "Signals" (SM small cross-sections, BSM, Higgs) are handled outside central production.
  - ✓ Made with PYTHIA, ALPGEN, MC@NLO

Summary

- This is a crucial year for Monte Carlo production in CMS and ATLAS.
- It is essential to favour a consistent (same generation strategy for signal and background) and coherent (same settings, full phase space coverage) configuration for the reference MC samples.
- We consider crucial to have generator redundancy for cross-checks and validation:
  - we do not want to rely on just one prediction
  - learn about sensitivity to theory modeling when data alone is not sufficient

- Use the best tools on the market (for signals and for backgrounds):
  - new PDFs, new MPI models, modern C++ event generators
  - ME-PS matching is necessary for describing multi-jet associated productions
  - NLO Monte Carlo invaluable for cross-checks and normalizations
  - Flexible implementation of new physics signals is very desirable

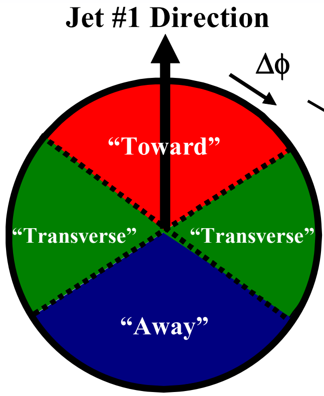
- A good MC setup is not enough: get ready for tuning with data (MB/UE, radiation, fragmentation, PDFs,…).

- ATLAS and CMS need+want to keep very much alive the level of communication with the theory groups. Examples of very profitable collaboration already in place.

Monte Carlos will shape the way how we will do physics. At the startup and not only.

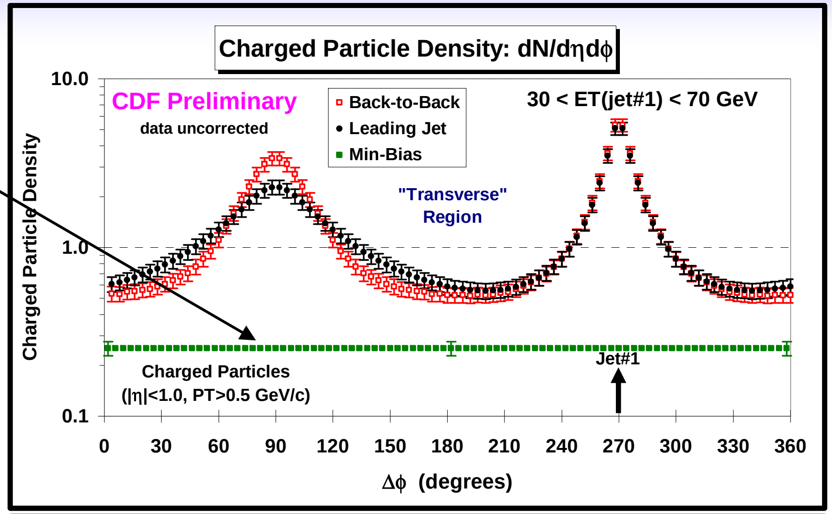
Backup

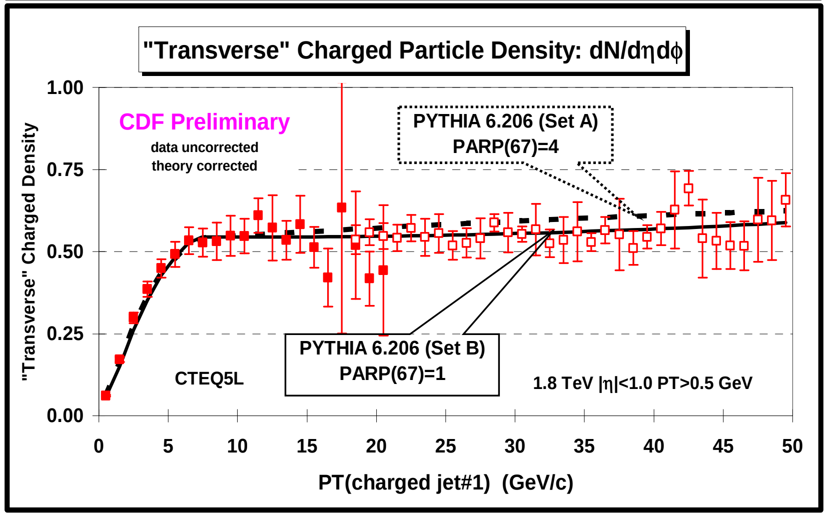
Pedestal effect



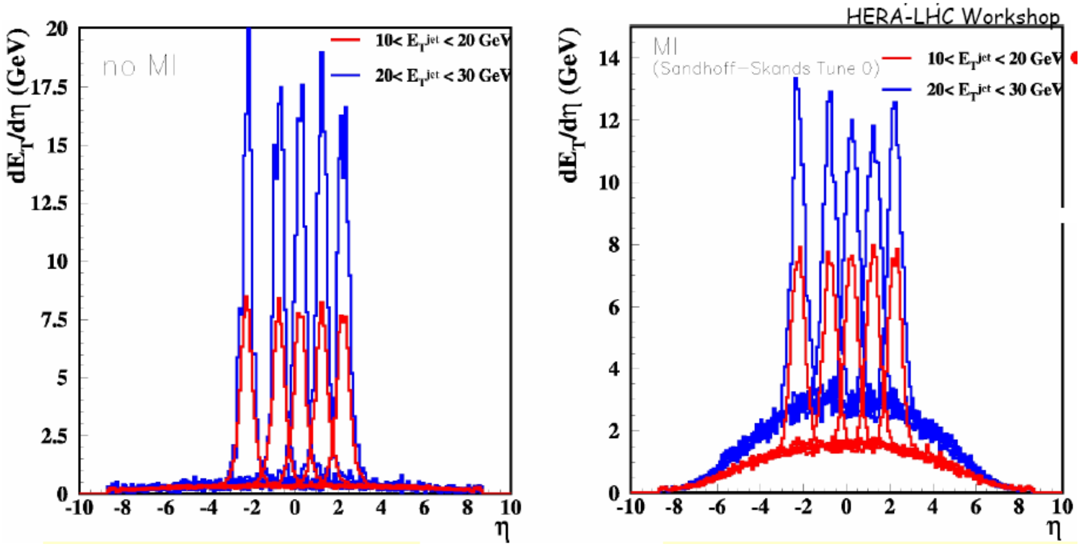
- The "transverse" region is very sensitive to the underlying event activity

- Study the event properties, in particular the particle density, as a function of the azimuthal angle, defined with respect to the leading jet in the event





- Tune the PS MC to the relevant differential distributions



- The underlying event gives the so-called pedestal effect
- Particles from the underlying event contribute to the jet energies

Underlying event tunings

PYTHIA6.416 - PYEVNT

CTEQ6ll (LO fit with LO α_s) → LHAPDF (set number 10042)

	Default:		Tuned:	
	MSTP(81)=21	→	new multiple interaction model & new parton shower selected!	
MPI	MSTP(82)=4			
ISR	MSTP(84)=1	} →	allowed ISR and FSR after first (hardest) interaction	
FSR	MSTP(85)=1			
MPI	MSTP(86)=2			
colour reconnection	MSTP(87)=4			
	MSTP(88)=1			
	MSTP(89)=1			
	MSTP(90)=0			
	MSTP(95)=1	→	MSTP(95)=2	minimize the total string length
	PARP(78)=0.025	→	PARP(78)=0.3	regulates the number of attempted colour reconnections (choose shorter strings)
p_{T0} scale	PARP(82)=2.0	→	PARP(82)=2.1	
matter distribution	PARP(83)=0.5	→	PARP(83)=0.8	
	PARP(84)=0.4	→	PARP(84)=0.7	
p_{T0} scale	PARP(89)=1800			
	PARP(90)=0.16			

Phase space definitions

QCD Jets

Process	Stars	Couplings	Phase space region	Matching	Banner	Event files	Remarks
jets (2)	3	QCD only	Ht bins (100-250 GeV, 250-500 GeV, 500-1000 GeV, 1000-inf GeV)	0,1,2,3,4+			light jets are u,d,c,s,g; Need to veto the first gluon splitting into bb in the PS
bb~ + jets	3	QCD only	Ht bins (100-250 GeV, 250-500 GeV, 500-1000 GeV, 1000-inf GeV)	0,1,2,3+			massive b; Need to veto the first gluon splitting into bb in the PS
bb~ bb~+ jets	1	QCD only		0,1+			massive b

Vector Boson

Process	Stars	Couplings	Phase space region	Matching	Banner	Event files	Remarks
W (-> l v)+ jets	3	EW=2 + QCD	all	0,1,2,3,4+			W=W+,W- ; l=(e,mu,tau)
Z /a* (-> l+l-)+ jets	3	EW=2 + QCD	m(l+l-)>50 GeV	0,1,2,3,4+			photon is included ; l=(e,mu,tau)
Z /a* (-> l+l-)+ jets	3	EW=2 + QCD	50>m(l+l-)>10 GeV	0,1,2,3,4+			photon is included ; l=(e,mu,tau)
Z (-> vv)+ jets	3	EW=2 + QCD	pt(Z)>50 GeV	0,1,2,3,4+			
V (-> l l')+ QQ~ + jets	1	EW=2+QCD	all	0,1,2+			V=W+,W-,Z ; l=(e,mu,tau,v), (Z->vv included) Q=b
a + jets	3	EW=1 + QCD	pt(a)>20 GeV, abs(eta(a))<2.5, DeltaR(a,jet)>0.3, in Ht bins (40-100 GeV, 100-200 GeV, 200-inf GeV)	0,1,2,3,4+			photon
a + QQ~ + jets	1	EW=1 + QCD	pt(a)>20 GeV, abs(eta(a))<2.5, DeltaR(a,jet)>0.3	0,1,2+			photon; Q=b

Vector Bosons

Process	Stars	Couplings	Phase space region	Matching	Banner	Event files	Remarks
VV(-> 4l)+ jets	3	EW=2+QCD	all	0,1+			V=W+,W-,Z l=(e,mu,tau,v)
VV (-> 4l) + QQ~	1	EW=1 + QCD	all	no			V=W+,W-,Z l=(e,mu,tau,v), Q=b
aV(-> 2l)+ jets	1	EW=2+QCD	all	0,1+			V=W+,W-,Z l=(e,mu,tau,v)
a a + jets	1	EW=2+QCD	pt(a)>20 GeV, abs(eta(a))<2.5, DeltaR(a,jet)>0.3	0,1,2+			photon
a a + QQ~ + jets	1	EW=1 + QCD	pt(a)>20 GeV, abs(eta(a))<2.5, DeltaR(a,jet)>0.3	no			photon; Q=b
V V V	3	EW=3	all	no			V=W+,W-,Z
a a a	3	EW=3	pt(a)>20 GeV, abs(eta(a))<2.5, DeltaR(a,jet)>0.3	no			

in collaboration with
F. Maltoni and the MG team

Top

Process	Stars	Couplings	Phase space region	Matching	Banner	Event files	Remarks
tt + jets	3	QCD only	all	0,1,2,3+			top decays into everything. Done with DECAY
tt + bb~	3	QCD only	all	no			top decays into everything. Done with DECAY
tjb	3	EW=2, QCD=1	all	no			t-channel, b massive, no top decay
tj	3	EW=2, QCD=0	all	no			t-channel, no top decay
tb	3	EW=2, QCD=0	all	no			s-channel, b massive, no top decay
tW	3	EW=2, QCD=1	all	no			tW-channel, no top decay
tWb	3	EW=2, QCD=2	all	no			tW-channel, b-massive, doub-res diagram subtraction, no top decay

Setup and Input Parameter Settings

Scales: set dynamically in MadGraph to m_T. We should use them as our default.

PDFs: proposal to use CTEQ6L1, for which an UE tune exists.

Input Parameter Settings: all listed here:
https://twiki.cern.ch/twiki/bin/view/CMS/MadgraphProduction2008Proposal
Change in the pole masses to match the PDG everywhere except for W and top, where the most recent world averages are used.

Note: we must take care and port these settings into other general purpose generators used in CMS

Tau decays: MadGraph can handle simple tau decays correctly, do we want to use this option?
Follow Tauola closely instead…

Block SMINPUTS # Standard Model inputs

- 1 1.32350785E+02 # alpha_em(MZ)(-1) SM MSbar
- 2 1.16639000E-05 # G_Fermi
- 3 1.18000000E-01 # alpha_s(MZ) SM MSbar
- 4 9.11880000E+01 # Z mass (as input parameter)

Block MGYUKAWA # Yukawa masses m/v=y/sqrt(2)

- PDG YMASS
- 5 4.20000000E+00 # mbottom for the Yukawa y_b
- 4 1.42000000E+00 # mcharm for the Yukawa y_c
- 6 1.64500000E+02 # mtop for the Yukawa y_t
- 15 1.77700000E+00 # mtau for the Yukawa y_ta

Block MGCKM # CKM elements for MadGraph ?

- 1 1 9.75000000E-01 # Vud for Cabibbo matrix

Block MASS # Mass spectrum (kinematic masses)

- PDG Mass
- 5 4.70000000E+00 # bottom pole mass
- 6 1.70900000E+02 # top pole mass
- 15 1.77700000E+00 # tau mass
- 23 9.11880000E+01 # Z mass
- 24 8.04000000E+01 # W mass
- 25 1.20000000E+02 # H mass

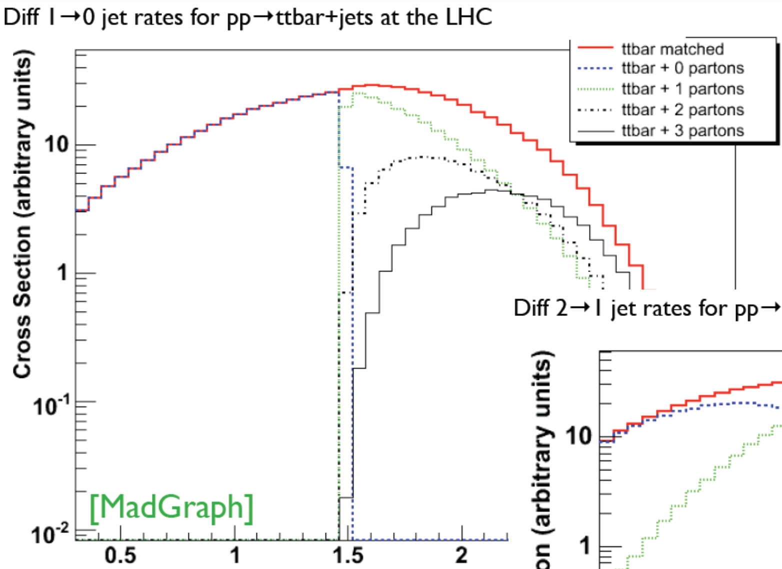
DECAYS

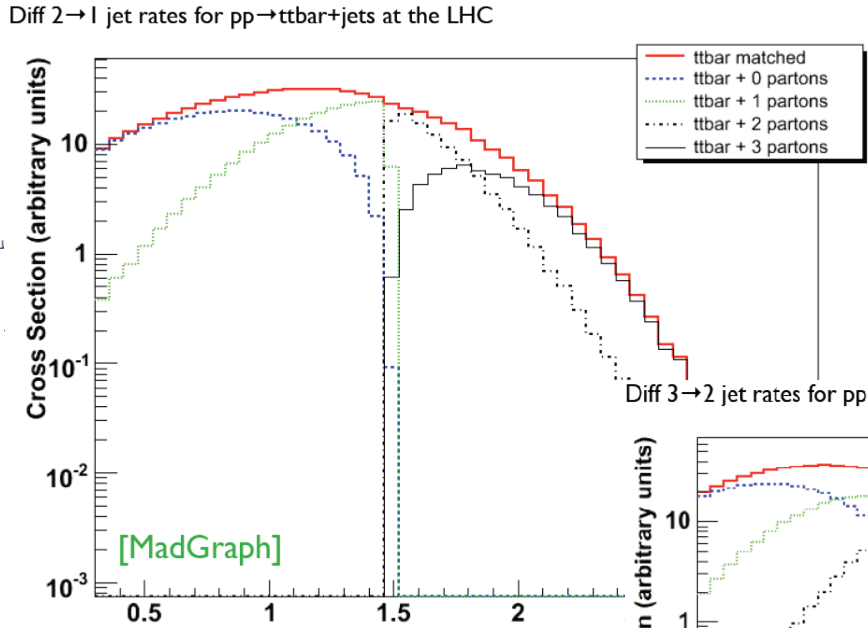
- PDG Width
- 6 1.40736322E+00 # top width
- 23 2.44041149E+00 # Z width
- 24 2.04614832E+00 # W width
- 25 5.75450381E-03 # H width

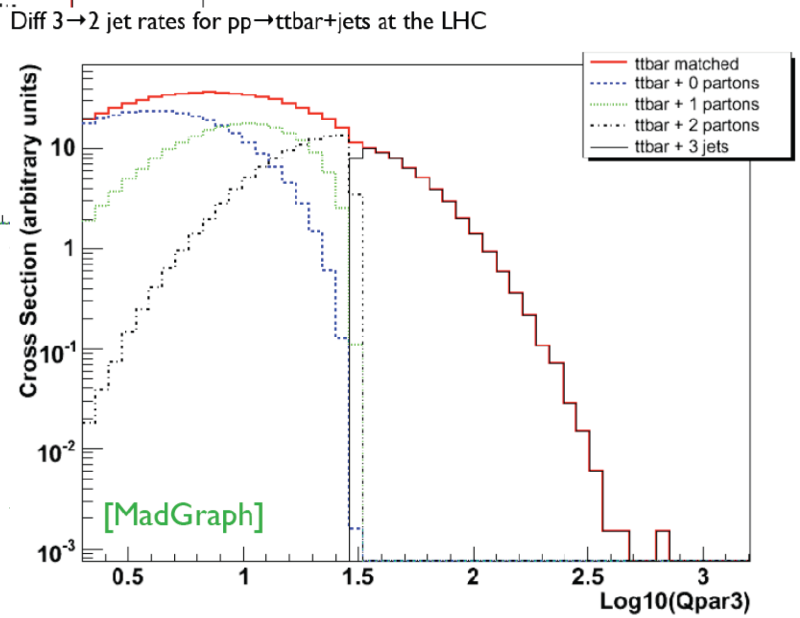
BRANCHING RATIOS

- BR NDA ID1 ID2
- 8.27247499E-02 2 4 -4 # BR(H -> c cbar)
- 7.17633150E-01 2 5 -5 # BR(H -> b bbar)
- 0.00000000E+00 2 6 -6 # BR(H -> t tbar)
- 4.31613962E-02 2 15 -15 # BR(H -> tau- tau+)
- 6.90147194E-03 2 23 23 # BR(H -> Z Z^(*))
- 7.47460022E-02 2 24 -24 # BR(H -> W W^(*))
- 3.02414236E-02 2 21 21 # BR(H -> g g)
- 1.43040946E-03 2 22 22 # BR(H -> A A)

Matching validation





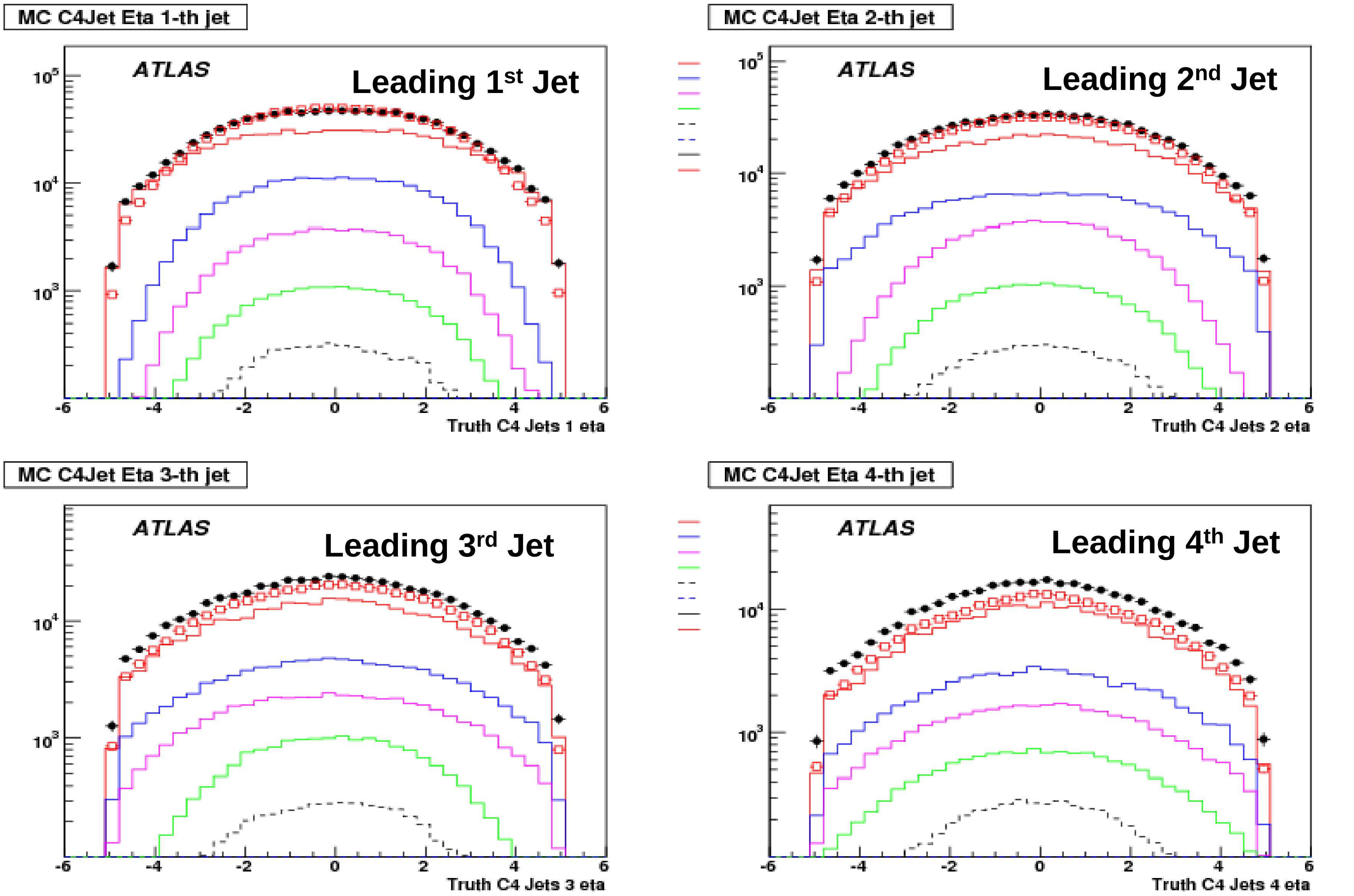


- Jet rates are
  - smooth at the cutoff scale
  - independent upon the cutoff scale (under reasonable

Roberto Chierici

Matched ME-PS vs standard PS: W+jets

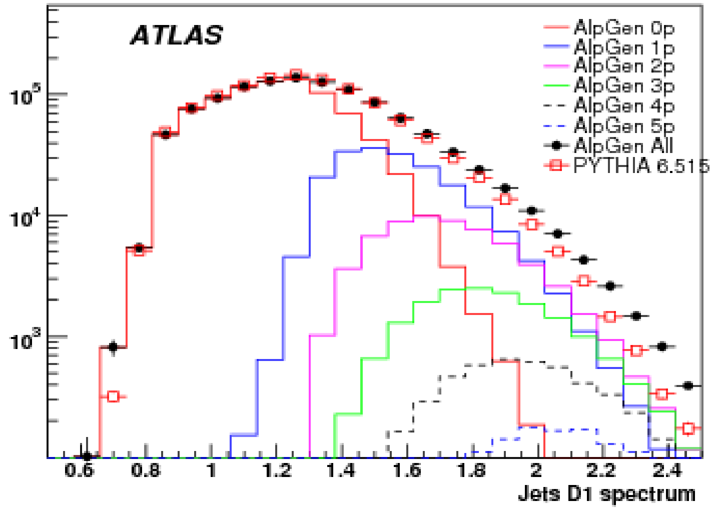
Differences in angular shape and normalization for extra jet production



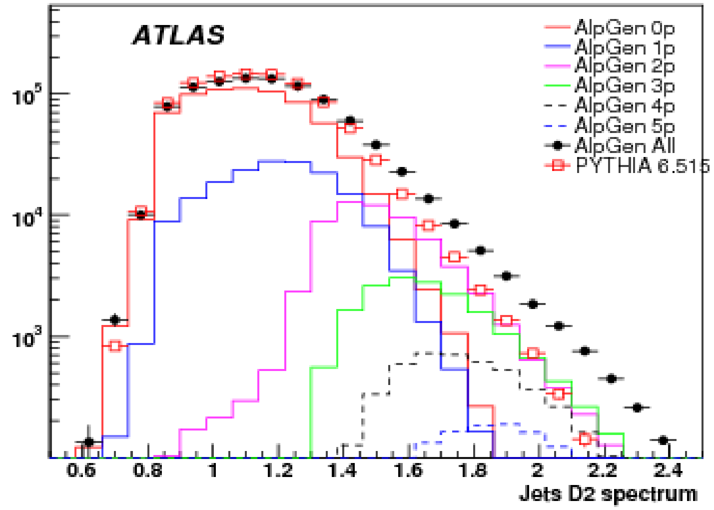
Matched ME-PS vs standard PS

Large differences in transverse variables related to radiation

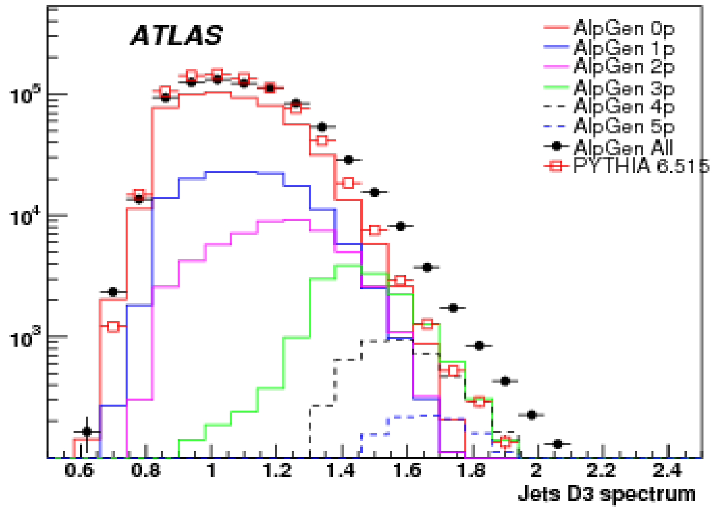
d_1 spectrum **0→1 jet rate**



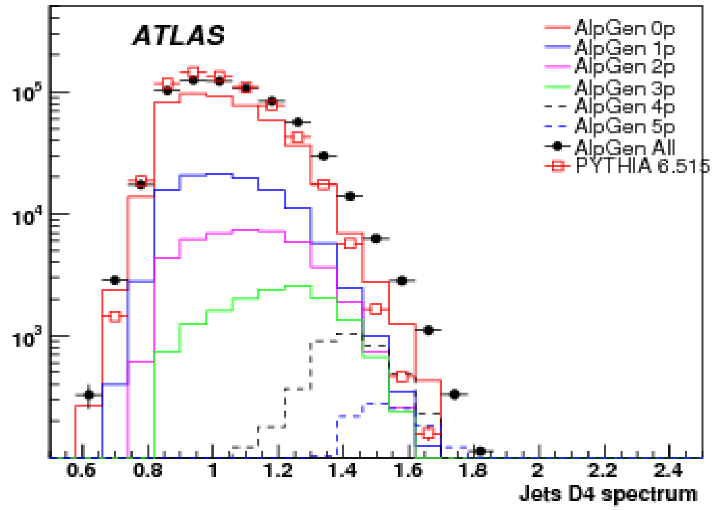
d_2 spectrum **1→2 jet rate**



d_3 spectrum **2→3 jet rate**



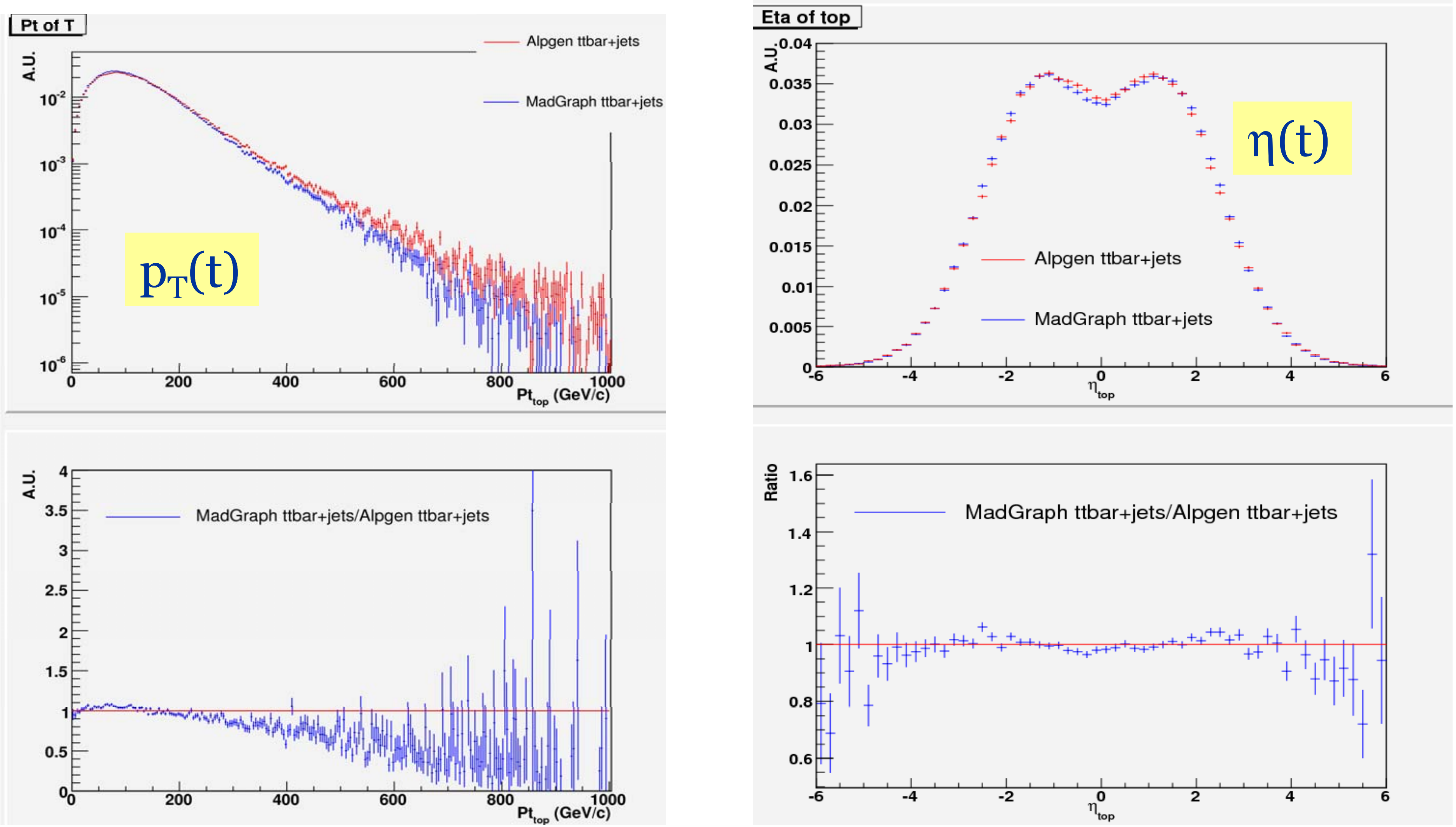
d_4 spectrum **3→4 jet rate**



ALPGEN vs MadGraph with matching

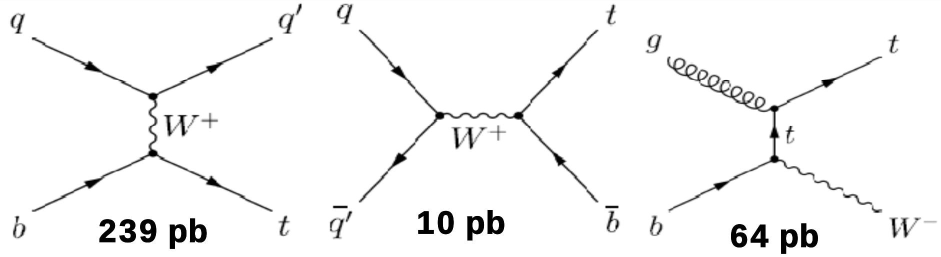
Excellent description of other variables for top physics

Residual differences due to different generator input settings and scales chosen?



Single top in ATLAS

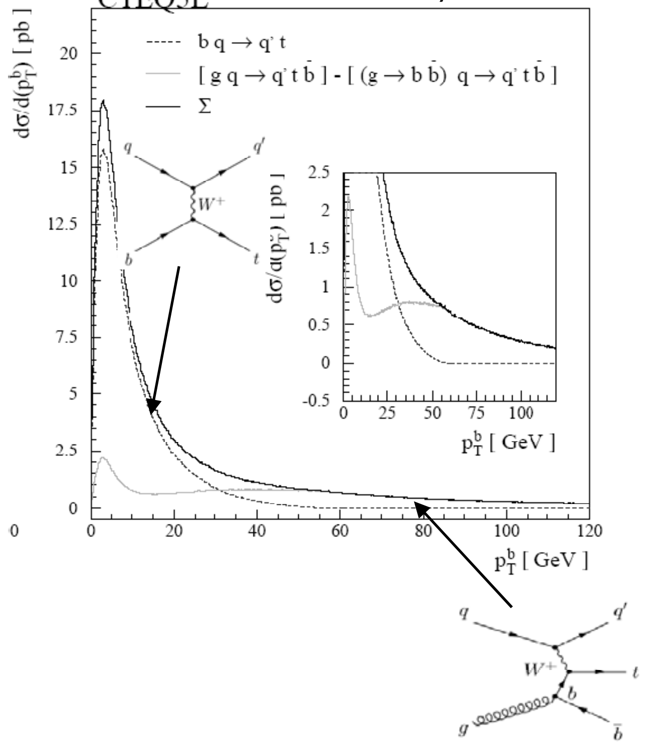
- Handled with AcerMC
  - ✓ s-channel, t-channel Wt, t-channel tq matched to NLO tqb with the ACOT prescription



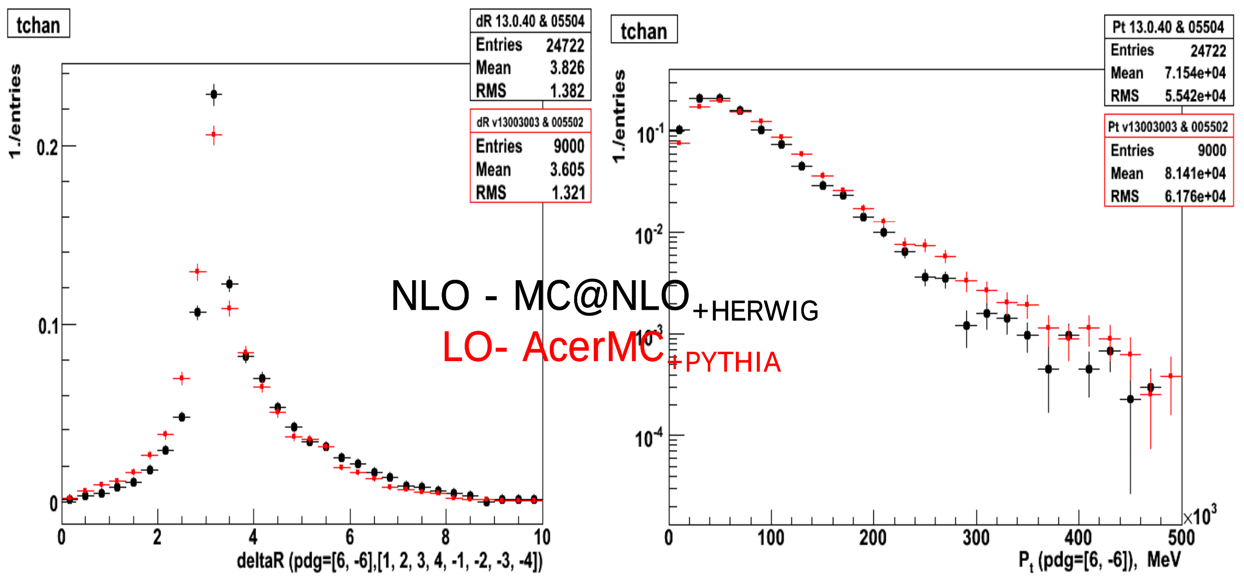
- The matching recovers part of the NLO contributions
  - ⇒ hardest second b spectrum

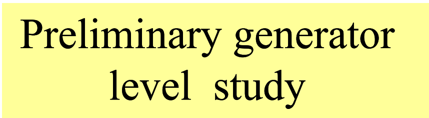
- As a comparison also MCatNLO envisaged. Validation of the generator has started, encouraging first results for the s-channel and the t-channel qt.

B. Kersevan, I. Hinchliffe

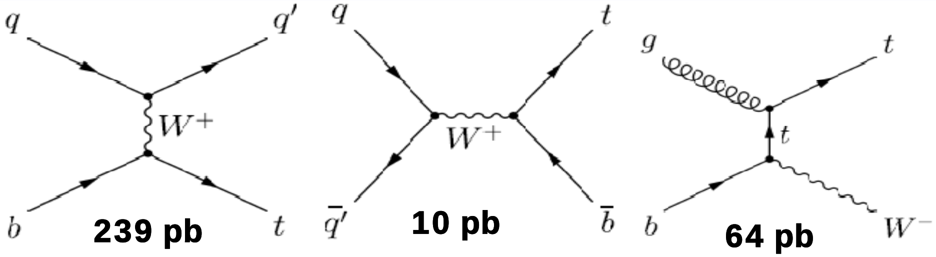


G.Khoriauli, M.Cristinziani and ATLAS MC group - t-channel tq(b) only



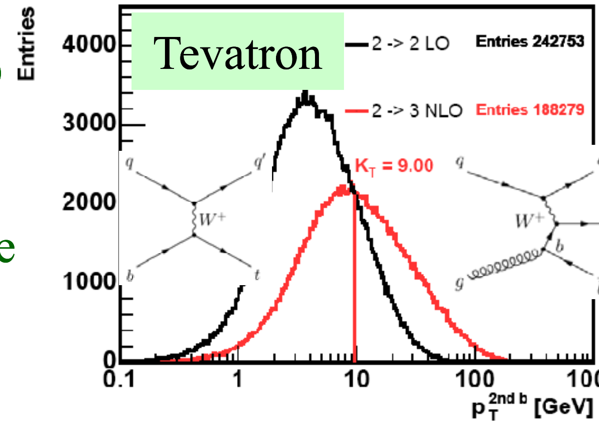


Single top in CMS

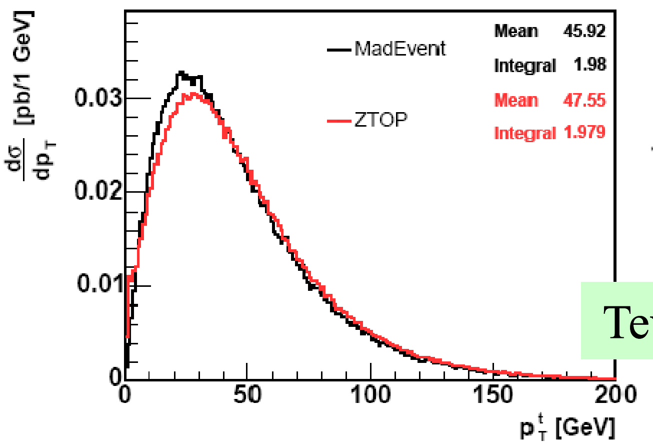


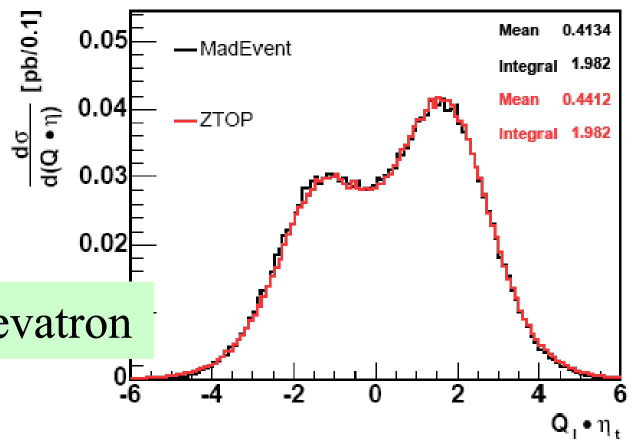
- Handled with MadGraph and MCatNLO
- For LO generation pp→qt, mimic NLO effects by matching 2→2 with 2→3 LO processes a la Tevatron according to a cut on the p_T of the second b. The sum of the contributions is normalized to the NLO prediction and the cut is chosen to have a continuous curve.
- pp→tW NLO mixes with tt LO, tWb needs to subtract the double resonant top contribution to avoid double counting.

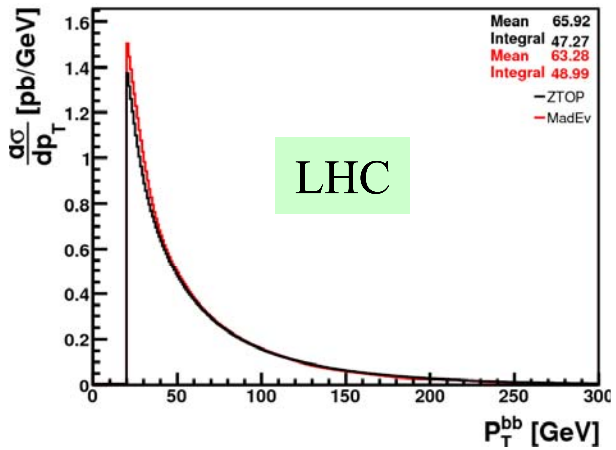
⇒ Best to use an NLO MC approach when possible for single top analyses



P. Sturm, W. Wagner, J. Weinelt







Top pair generation comparison

Aim: test generators and validate their physics content (for top pair physics) in the frame of the CMS software. This will test them in the way we use them (debug !) and help understanding their features and make generation choices.

- Particular focus on the validation of MadGraph.
- Work ongoing on MC@NLO.

Input parameter settings as uniform as possible:

Parameter	TopRex	MadGraph	ALPGEN	Phantom
PDFs	CTEQ5L	CTEQ5L	CTEQ5L	CTEQ5L
Renormalization scale	m_T	m_t	m_T	m_T → $\sum_{tops}(m^2+p_T^2)$
Factorization scale	m_T	m_t	m_T	m_T
Top mass (Gev/c^2)	175	174.3	175	175
Λ_{QCD} (PARP(61), PARP(62)) (GeV)	0.25	0.25	0.25	0.25
Q^2_{max} switch (PARP(67))	2.5	2.5	2.5	2.5

Interface to CMS software via Les Houches file standard where possible.
Amount of statistics compatible with the CMS production allocated "bandwidth":

Generator	Production	Type	Interface to CMSSW	Events
TopRex	Offical CMS	inclusive $t\bar{t}$	Direct	$1,5 \times 10^6$
Alpgen	Official CMS	$t\bar{t}$ +Njets matched, N=0,...,4	Converted ROOT files	2×10^5
Phantom	Private	pp→$b\ell\nu$bqq'	LH files	5×10^5
MadGraph	Private	inclusive $t\bar{t}$	LH files	3×10^6
		$t\bar{t}$ +Njets matched, N=0,...,3	LH files	2×10^5