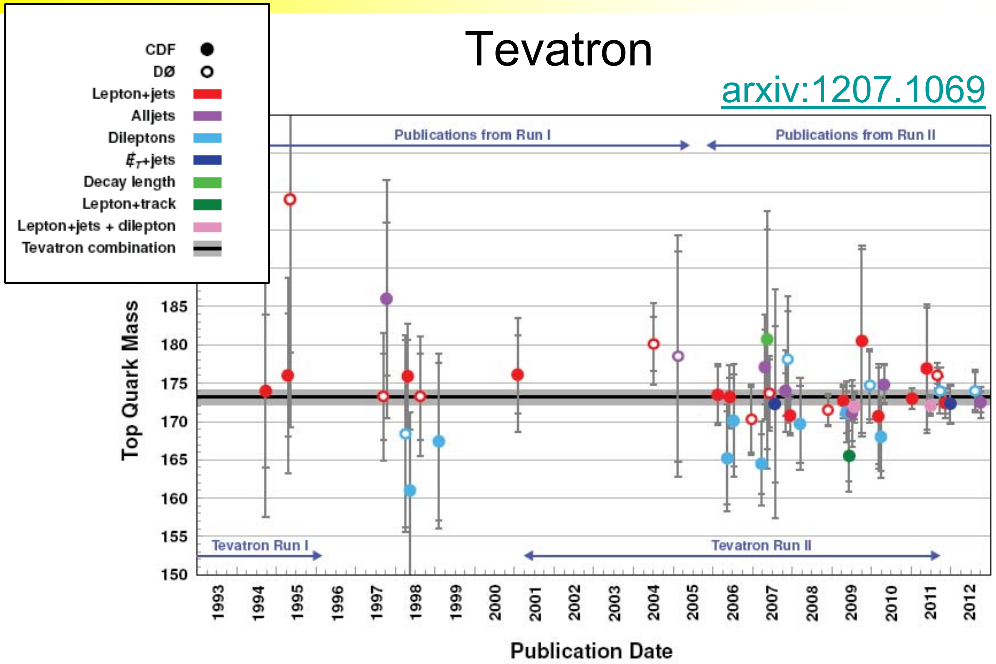
Fifth International Workshop on Top Quark Physics
Winchester U.K. September 16-21, 2012

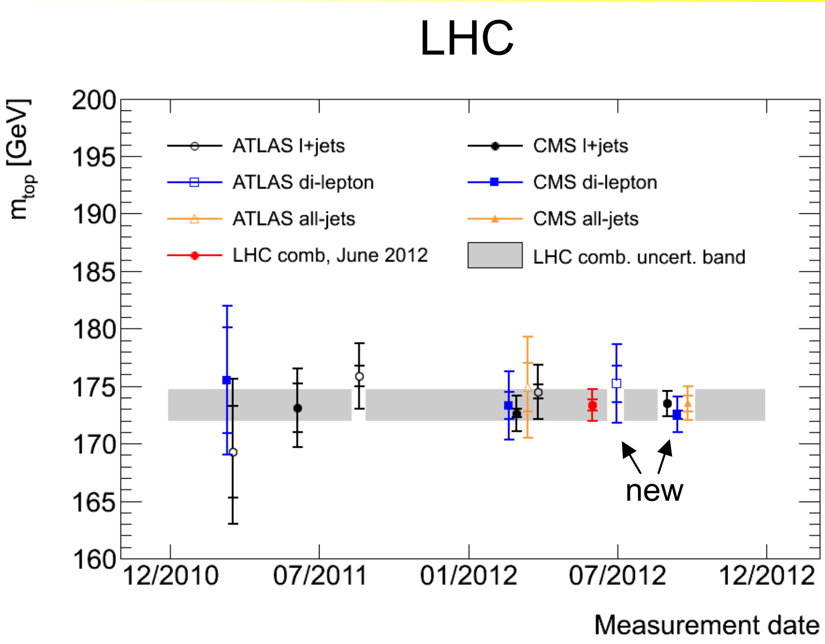
Experimental and theoretical uncertainties in top-quark mass measurements at Tevatron and LHC

Giorgio Cortiana MPP Munich

Max-Planck-Institut für Physik
(Werner-Heisenberg-Institut)

Measurements overview





- Time evolution of the Tevatron and LHC m_{top} measurements
  - Several measurements from 1995 to 2012
  - Almost since the beginning (except for the very early Tevatron measurements) systematic uncertainties are dominating

	Tevatron		LHC	
[GeV]	CDF	D0	ATLAS	CMS
Stat.	***0.5*** – ***10.3***	***0.8 – 12.3***	***0.6 – 4.0***	***0.3*** – ***4.6***
Syst.	***1.0*** – ***5.7***	***1.2 – 3.9***	***2.3 – 4.6***	***1.0*** – ***4.6***
Tot	***1.1*** – ***11.5***	***1.5 – 12.8***	***2.4 – 6.3***	***1.1*** – ***6.5***

I will mainly discuss the uncertainties on the current Tevatron and LHC, m_{top} combinations input measurements

► but also some updates (in red) as from talks by:
Pavol Bartos (Tevatron)
Markus Seidel (LHC)

Talk roadmap

Systematic uncertainties on top-quark mass measurements

- Theory uncertainties:
  - Signal simulation
  - Event modelling and environment
- Experimental uncertainties
  - Physics objects and detector modelling
    - identification, reconstruction, and calibration
  - Energy scales
    - (in particular for jets)

- How the systematics uncertainties are evaluated by different experiments
- What are the techniques adopted and/or prospects to reduce the uncertainties

I will discuss the uncertainties on direct m_{top} determination. m_{top} extraction from x-section will have in addition theoretical uncertainties related to the x-sec calculation (PDF, scales)

Talk roadmap

Systematic uncertainties on top-quark mass measurements

- Theory uncertainties:
  - Signal simulation
  - Event modelling and environment

⇧

- Background contamination

⇩

- Experimental uncertainties
  - Physics objects and detector modelling
    - identification, reconstruction, and calibration
  - Energy scales
    - (in particular for jets)

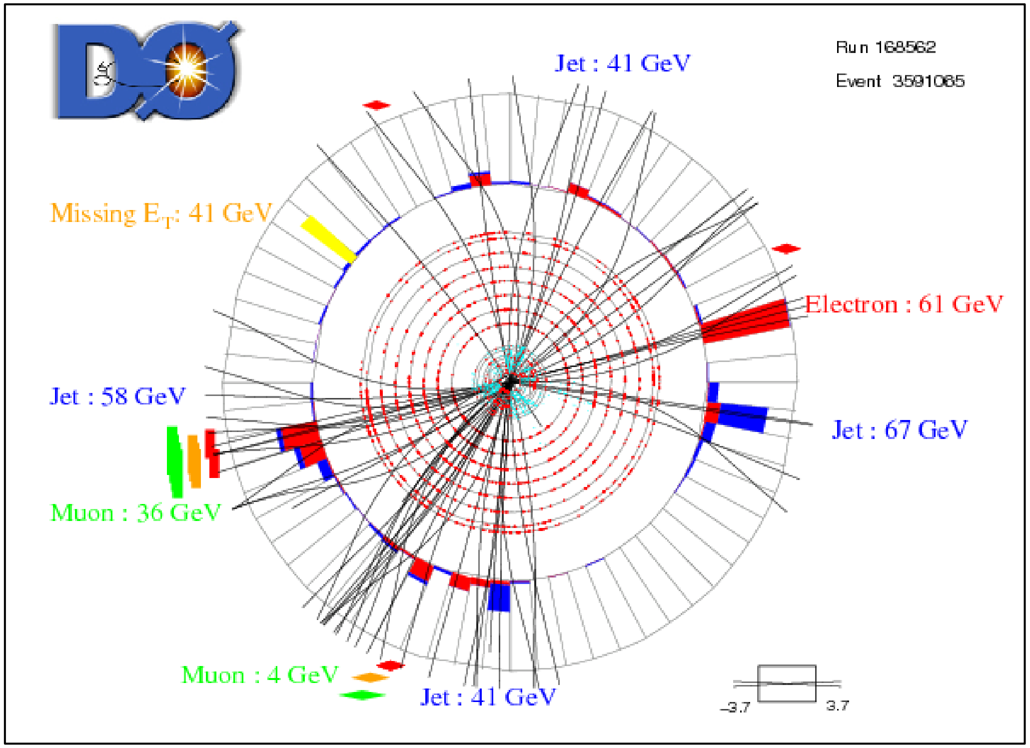
- How the systematics uncertainties are evaluated by different experiments

- What are the techniques adopted and/or prospects to reduce the uncertainties

I will discuss the uncertainties on direct m_{top} determination. m_{top} extraction from x-section will have in addition theoretical uncertainties related to the x-sec calculation (PDF, scales)

Event displays

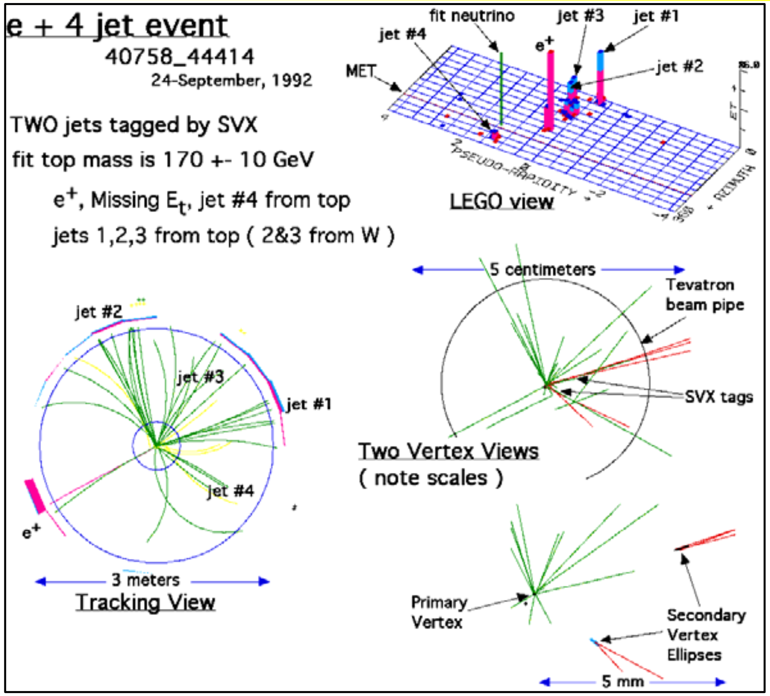
http://www-d0.fnal.gov/Run2Physics/displays/



Run 168562

Event 3591085

http://www-cdf.fnal.gov/events/tej.html



e + 4 jet event

40758_44414

24-September, 1992

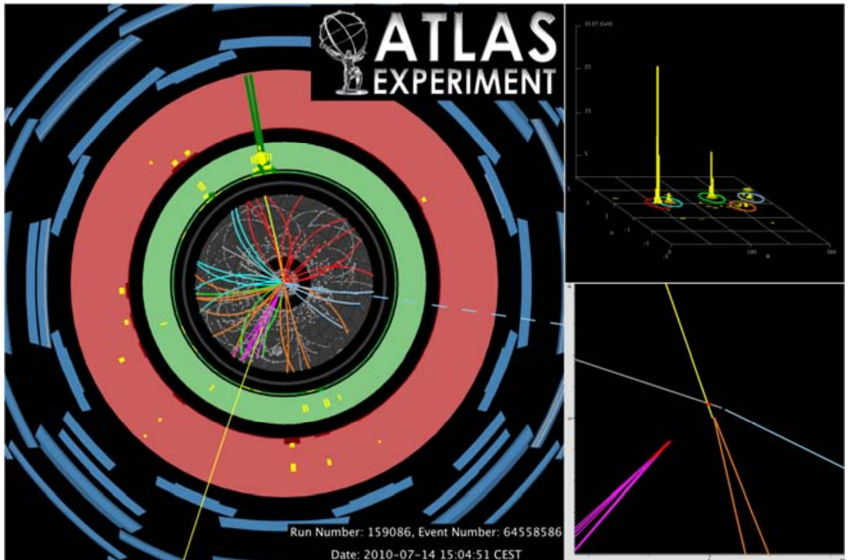
TWO jets tagged by SVX

fit top mass is 170 +- 10 GeV

e^+, Missing E_t, jet #4 from top

jets 1,2,3 from top (2&3 from W)

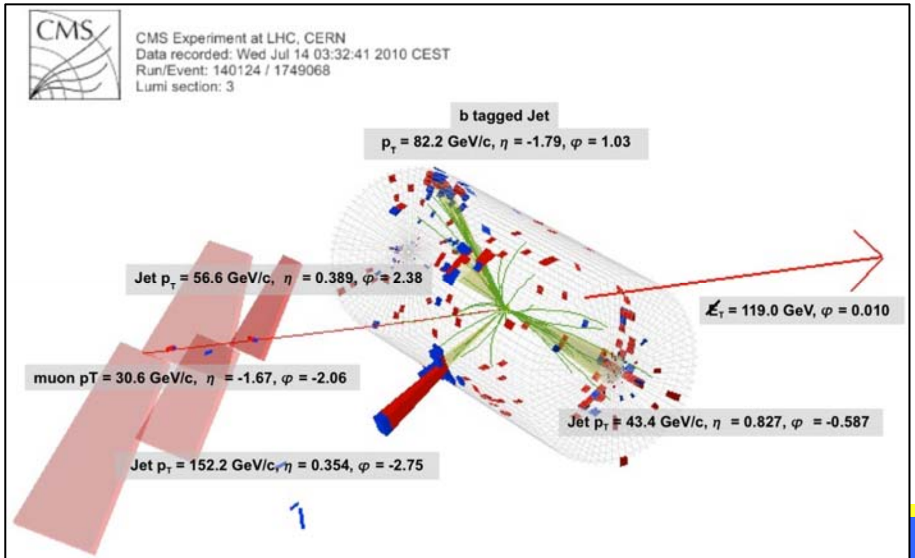
ATLAS EventDisplay



Run Number: 159086, Event Number: 64558586

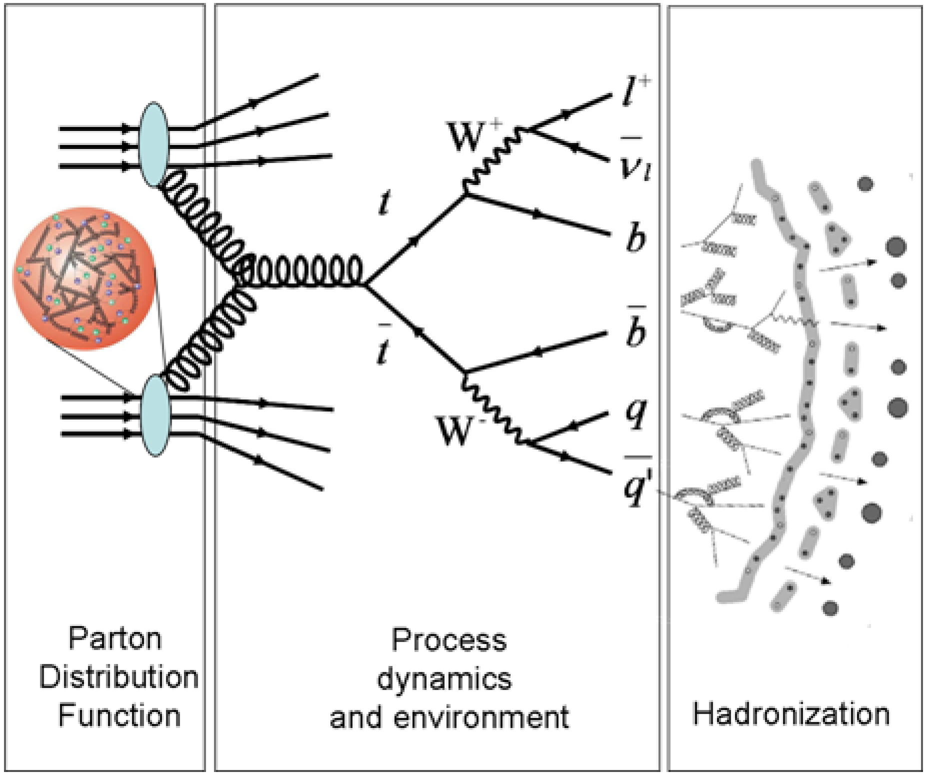
Date: 2010-07-14 15:04:51 CEST

CMS TOP Candidate 2010



CMS Experiment at LHC, CERN
Data recorded: Wed Jul 14 03:32:41 2010 CEST
Run/Event: 140124 / 1749068
Lumi section: 3

Anatomy of a top-quark pair event



- Signal simulation
  - Proton PDFs
  - MC generator
  - Hadronization models

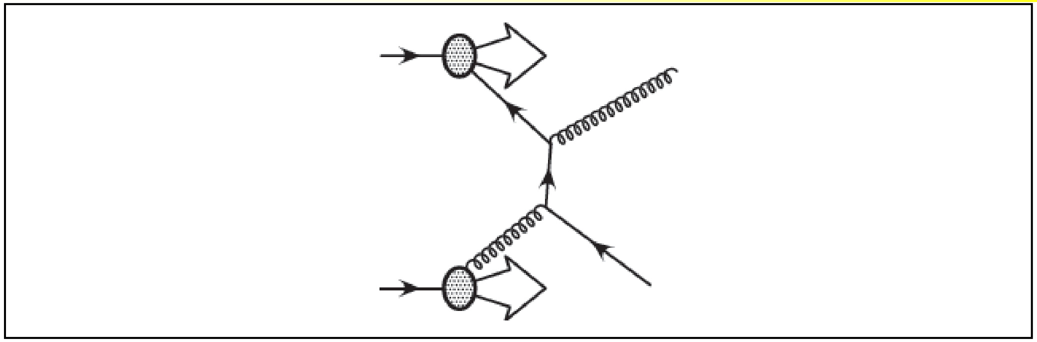
►see also talks by: Ivor Fleck (ATLAS), Martijn Gosselin (CMS), Viatcheslav Sharyy (Tevatron)

- Event modelling and environment
  - Underlying event and MC tunes
  - Color reconnection
  - QCD radiation
  - Pileup (multiple interactions)

- Theory uncertainty evaluated by:
  - using different generators / tools / inputs (ie PDF) for the simulation of the hard process and/or the parton shower (varied steering parameters in the simulation programs).
  - May have stat components due to the limited MC sample sizes

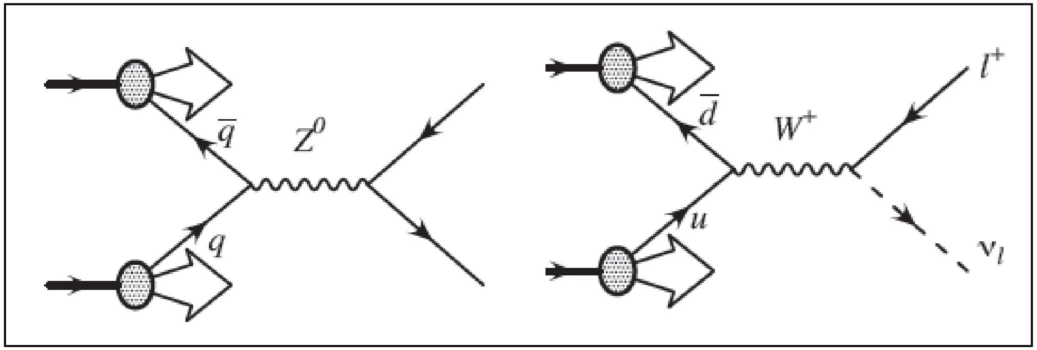
Proton PDF

- How well do we understand the (anti) proton composition impacts the top-pair production x-sec and the event kinematics

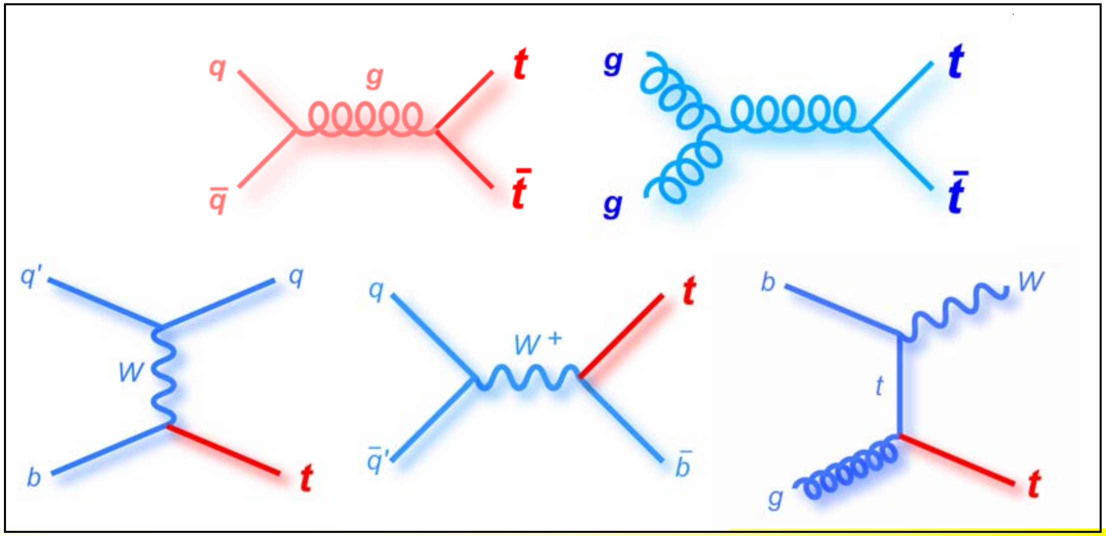


- The PDF are obtained global fits from different groups.

  Use mainly measurements from e-p DIS and in addition:



- Jets physics
- Weak bosons production:
  - Inclusive W/Z production
  - W/Z asymmetries
- (Top-quark production)
  - Top pair
  - Single top

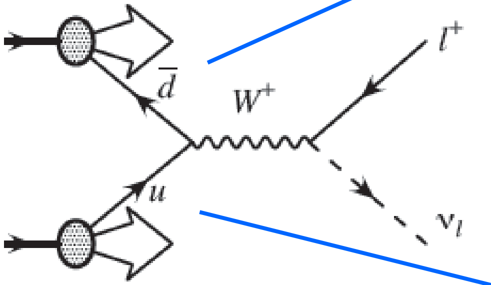


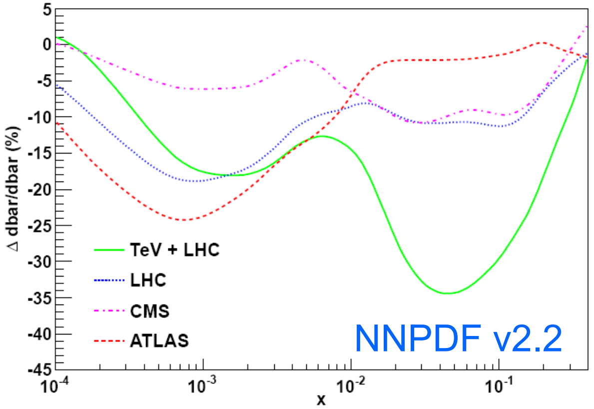
PDF

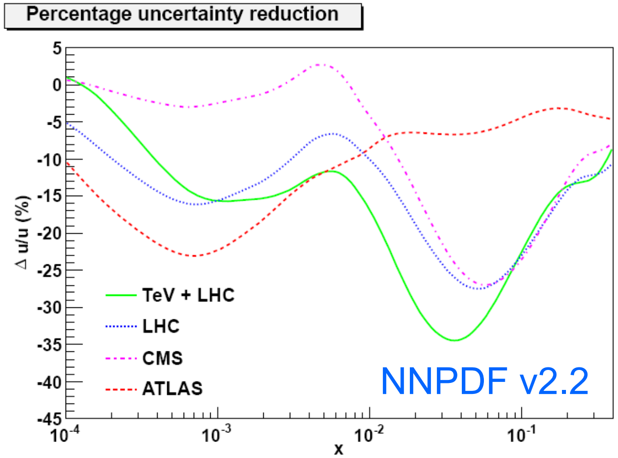
	Tevatron		LHC	
	CDF	D0	ATLAS	CMS
PDF reweighting	*CTEQ6M / MRST98L*	*CTEQ6M*	*CTEQ6.6*	*MSTW08/ CTEQ6.6 / NNPDF2.0*
Unc. on m_{top} [GeV]	*~0.1*	*~0.2*	*0.1 – 0.6*	*0.1 – 0.5*

- Uncertainties are in general small, but new data keeps helping reducing them
  (LHC probes new x-Q^2 regions with respect to previous experiments)

- Weak bosons production
  - W/Z asymmetries
    - to reduce PDF uncertainty on valence and sea quarks







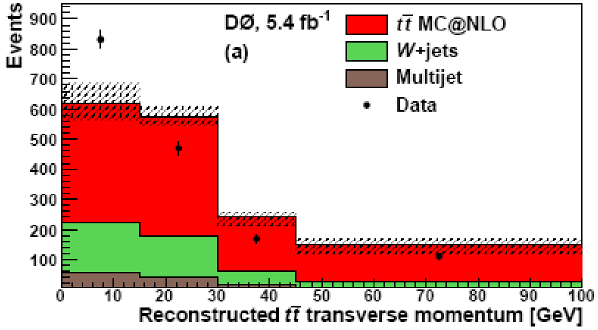
http://arxiv.org/abs/1108.1758

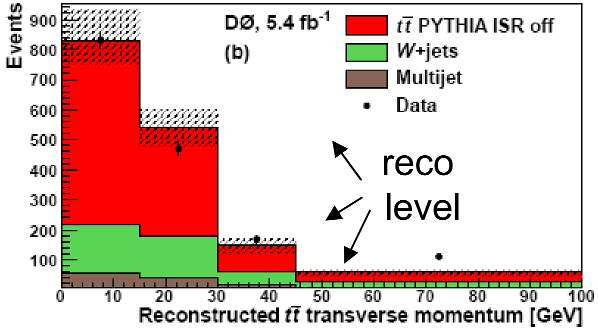
Similarly W+1jet can be used to constrain the gluon PDF

MC choice

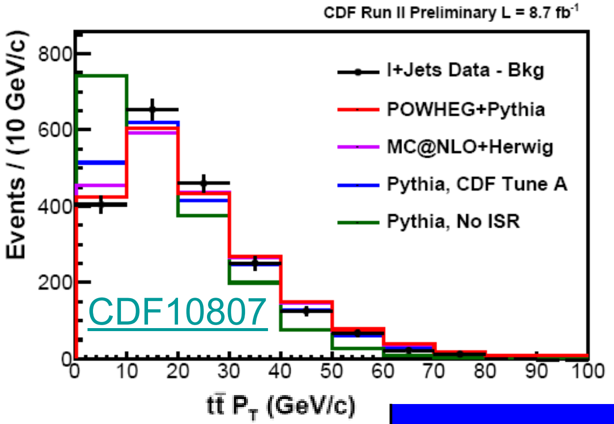
- Various simulation engines are available to simulate signal (and background) events.
  - LO (Pythia, Herwig)
  - NLO (MC@NLO, PowHeg)
  - multi-leg generators (Alpgen, Madgraph)

http://arxiv.org/abs/1107.4995

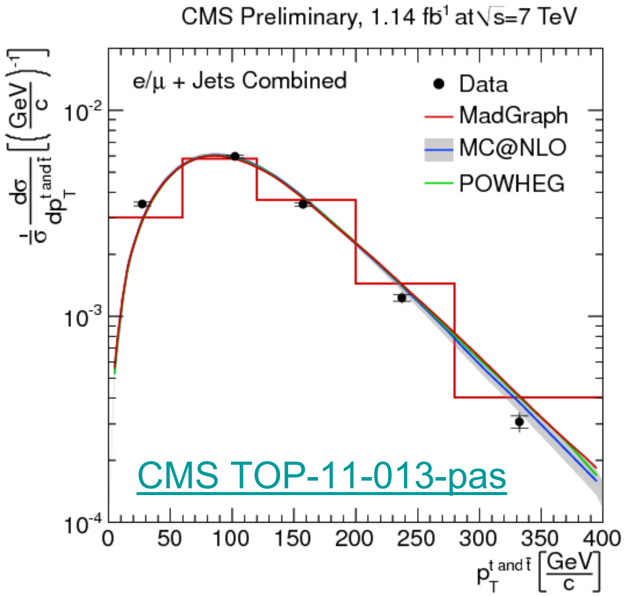




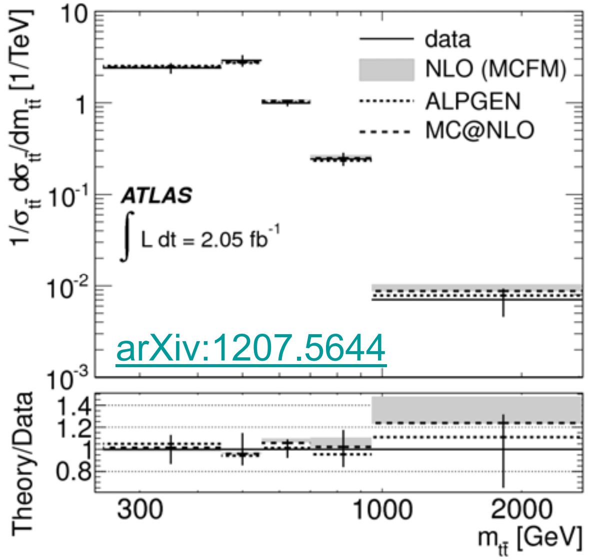
reco level



CDF10807



CMS TOP-11-013-pas



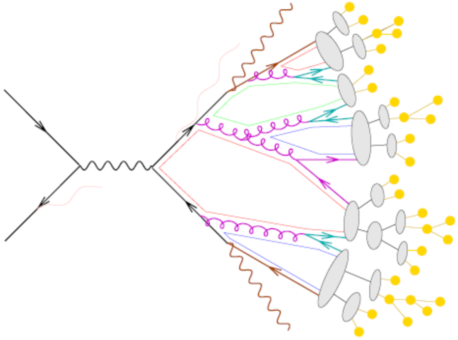
arXiv:1207.5644

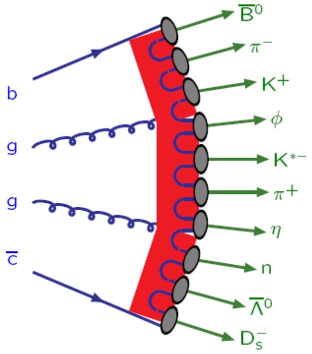
	Tevatron		LHC	
	CDF	**D0**	**ATLAS**	**CMS (dilep only)**
MC generator Systematics	***Pythia / MC@NLO***	***Alpgen / MC@NLO***	***MC@NLO / PowHeg***	***Madgraph/PowHeg/Alpgen***
Unc. on m_{top} [GeV]	***0.1 – 0.7***	***0.25 – 0.6***	***0.3 – 1.3***	***<0.1 – 0.4***

Hadronization models

- Describe the transition from final state partons to colorless hadrons

- Cluster and string hadronization models are implemented in Herwig and Pythia, respectively





- Hadronization systematics are considered in the determination of both the JES (Jet Energy Scale) and the MC modelling uncertainties:
  - the component in the JES refers mainly to single isolated jets
  - the hadronization unc. from top-pair MC accounts also for the multi-jet environment

there could be a sizeable double counting

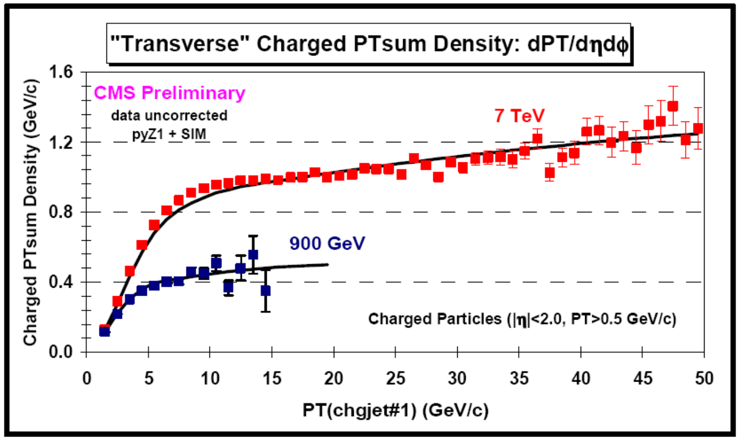
	Tevatron		LHC	
	CDF	**D0**	**ATLAS**	**CMS**
Hadronization	***Pythia/Herwig*** ***incl. UE syst***	***Alpgen*** ***Pythia/Herwig*** ***incl. UE syst***	***Powheg*** ***Pythia/Herwig***	***(considered in the JES syst determination)***
Unc. on m_{top} [GeV]	***0.2 – 0.3***	***0.6***	***0.2 – 0.9***	***n/e***

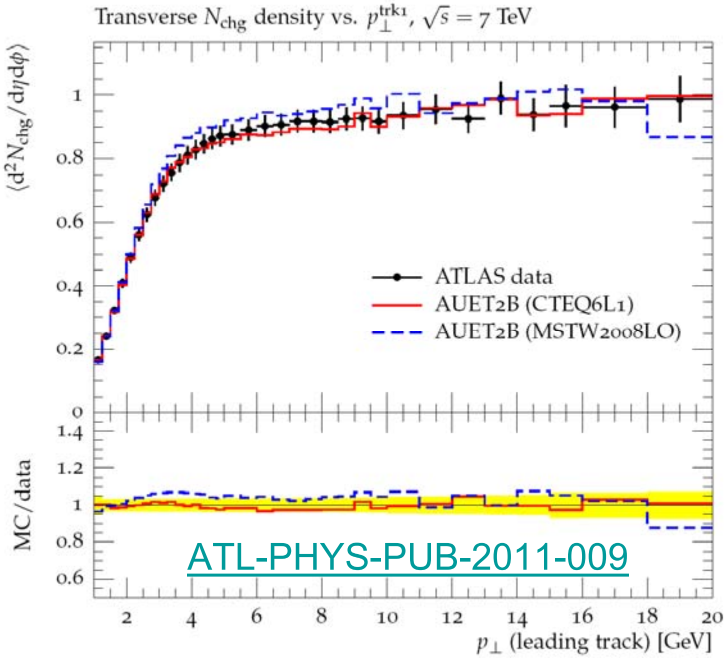
Event modelling and environment - 1

UE and MC tunes

- Additional physics (soft QCD) processes can contribute to the total observed activity in hard hadron collisions
  - Interactions between proton remnants

http://arxiv.org/pdf/1010.3558v1





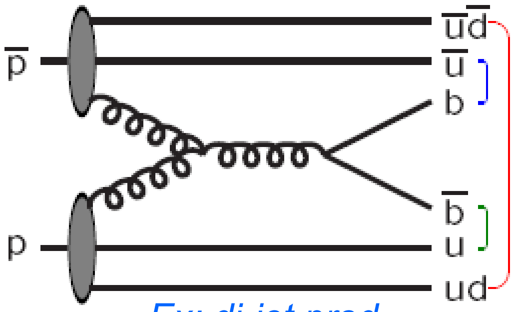
ATL-PHYS-PUB-2011-009

	Tevatron		LHC	
	CDF	D0	ATLAS	CMS
UE tune	*Pythia/Herwig*		*Varying Pythia tunes*	
Unc. on m_{top} [GeV]	*Incl. in the hadroniz. syst*	*Incl. in the hadroniz. syst*	***0.2** – 0.6*	***0.2 – 1.4***

Event modelling and environment - 2

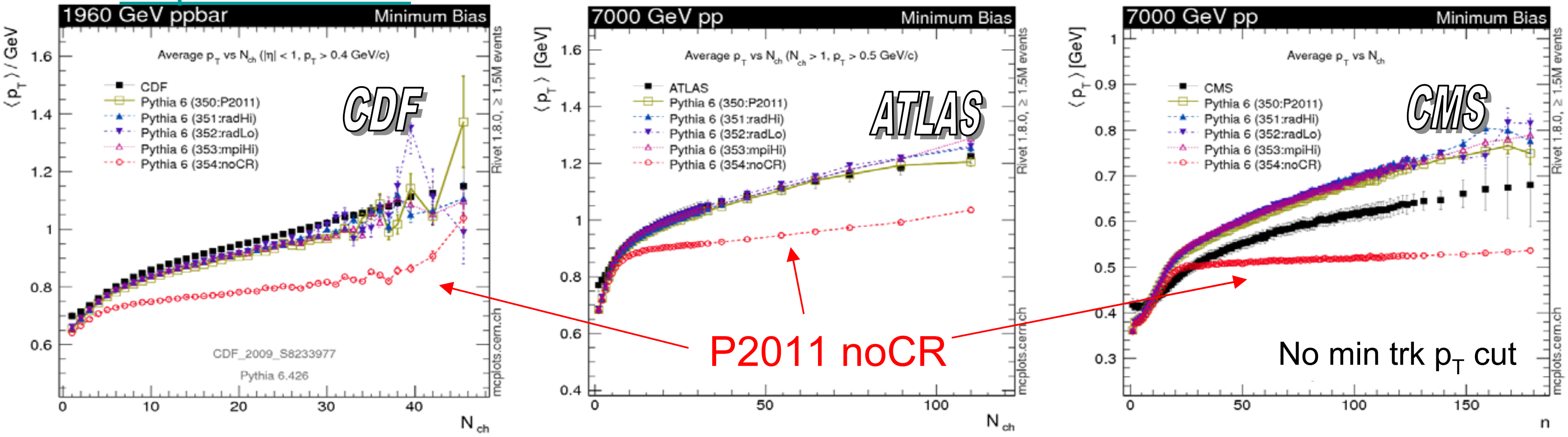
Color reconnection

- cross-talks between the hard interaction and the beam remnants in different color systems (see for example hep-ph:0703081)



Ex: di-jet prod.

mcplots.cern.ch



	Tevatron		LHC	
	CDF	D0	ATLAS	CMS
Colour reconnection (CR or noCR)	***Pythia TuneA variation***		***Pythia Tune A and Perugia variations***	***Pythia Perugia variations***
Unc. on m_{top} [GeV]	***0.2 – 0.5***	***0.3 – 0.5***	***0.6 – 1.2***	***0.1 – 0.5***

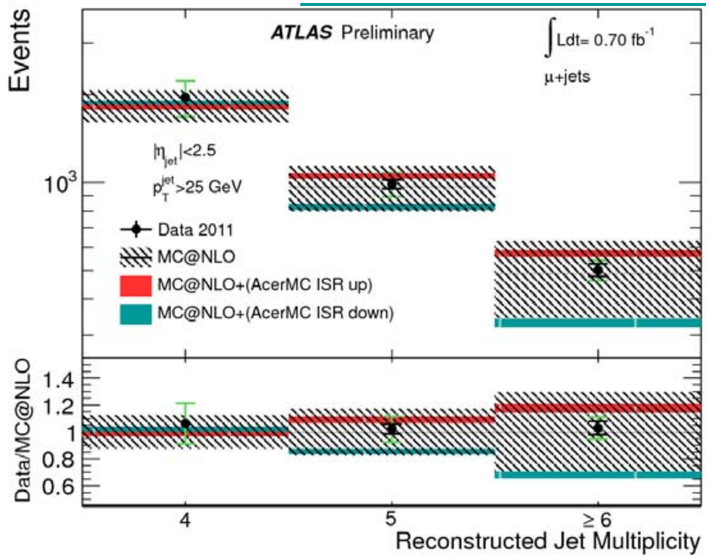
Event modelling and environment - 3

Radiation systematics (ISR/FSR)

- Increased/decreased levels of extra QCD radiation modify the event kinematics and topology (ie jet multiplicity) and alter the top-pair decay reconstruction

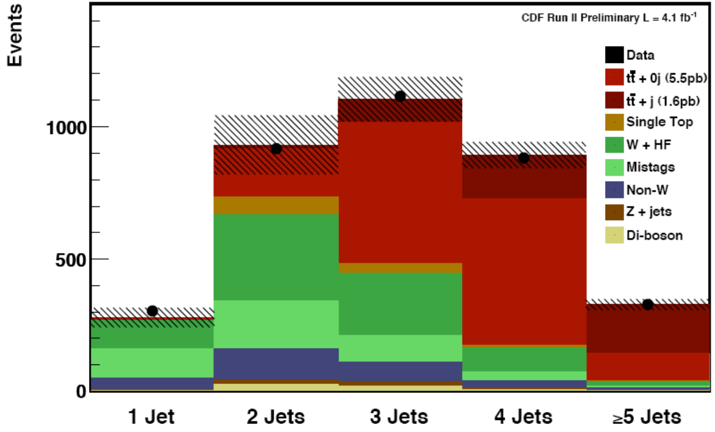
- Can be constrained using top-pair+jets: but in general large systematics due to detector effects (JES, b-tagging) and luminosity.

ATLAS-CONF-2011-142



see also poster by Sarah Allwood-Spiers

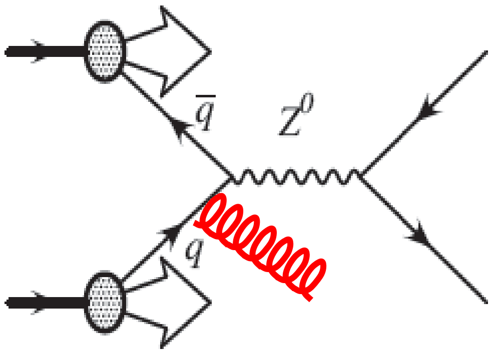
cdf9850

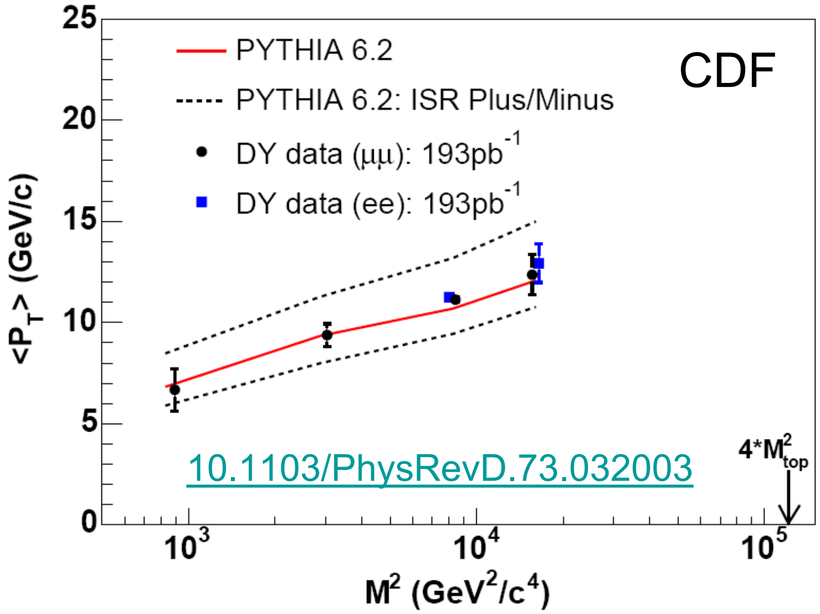


Constraining ISR with data

At Tevatron:

- Drell-Yan events
  - Same initial state as most top-pair events, but no final-state radiation
  - The mean p_T of the produced di-lepton pairs is sensitive to the ISR settings
  - The values of Λ_{QCD} and Q^2 in the MC that bracket the data are used to tune the ISR variations (extrapolated to the top-pair region)
- Since the initial and final-state shower algorithms are controlled by the same QCD evolution equation, the same parameter variations are used to estimate the effect of final-state radiation

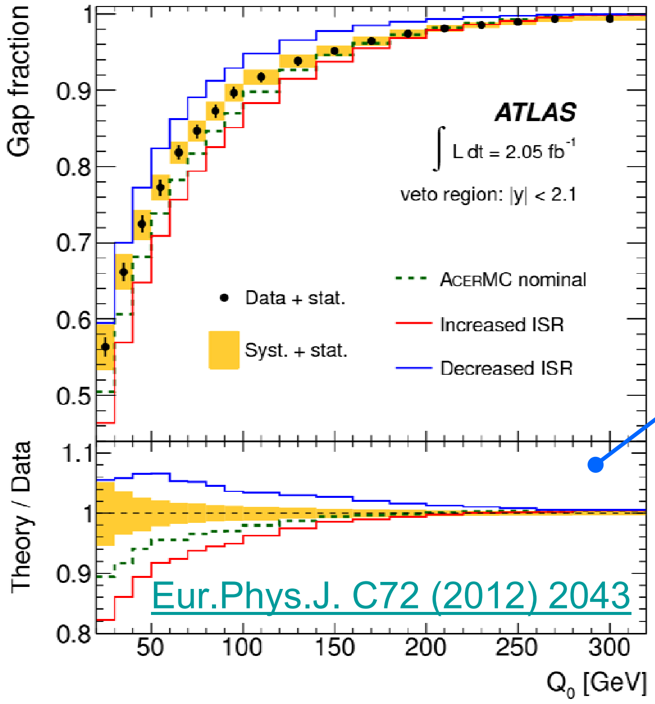




10.1103/PhysRevD.73.032003

Jet shapes can be used to complement FSR tunings

ISR



Eur.Phys.J. C72 (2012) 2043

Unfolded data/MC comparison using AcerMC (LO) + Pythia ISR variations

At LHC (for example),

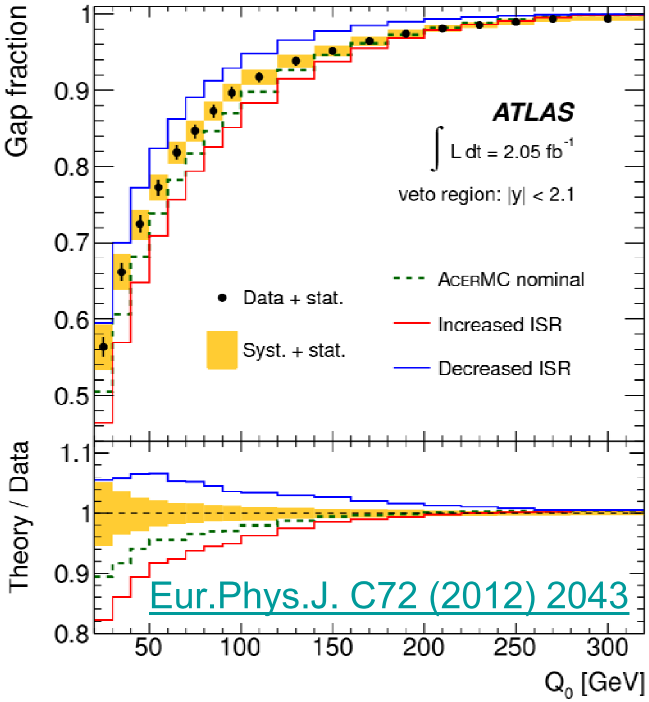
- As an alternative to top-pair+jets, use the Gap Fraction in top-pair di-lepton events
  - measure the fraction of top-pair dilepton events which do not contain an extra jet with $p_T > Q_0$

$$\boxed{f(Q_0) = \frac{n(Q_0)}{N}}$$

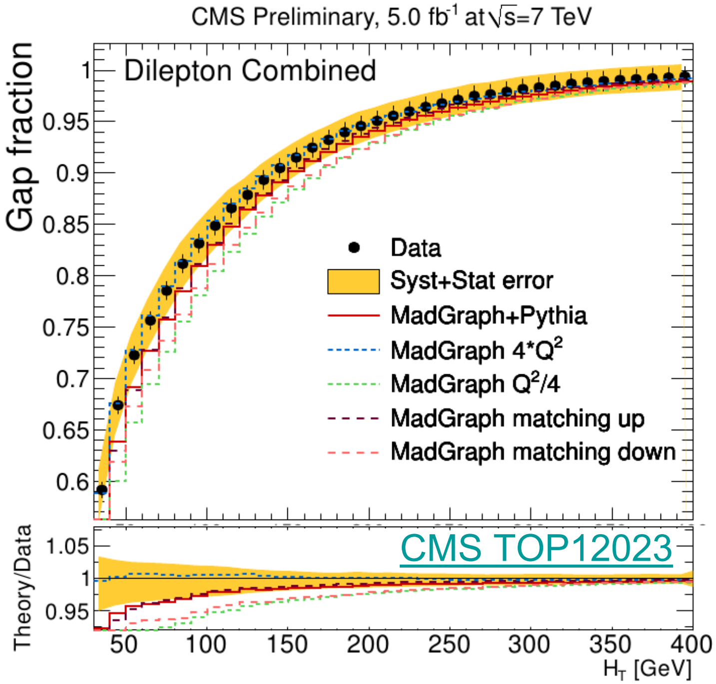
sensitive to the leading p_T emission

- Many of the experimental systematic uncertainties cancel in the ratio
  - The data are use to constrain the modelling of quark and gluon radiation in top-quark pair events

ISR



Eur.Phys.J. C72 (2012) 2043



CMS TOP12023

Unfolded data/MC comparison using AcerMC (LO) + Pythia ISR variations

At LHC (for example),

- As an alternative to top-pair+jets, use the Gap Fraction in top-pair di-lepton events
  - measure the fraction of top-pair dilepton events which do not contain an extra jet with p_T>Q_0
  - the veto criterion can be extended to probe jet activity beyond the leading additional jet.

$$\boxed{f(Q_0) = \frac{n(Q_0)}{N}}$$

sensitive to the leading p_T emission

$$\boxed{f(Q_{\text{sum}}) = \frac{n(Q_{\text{sum}})}{N}}$$

sensitive to all hard emissions accompanying the system (Q_{sum} = H_T)

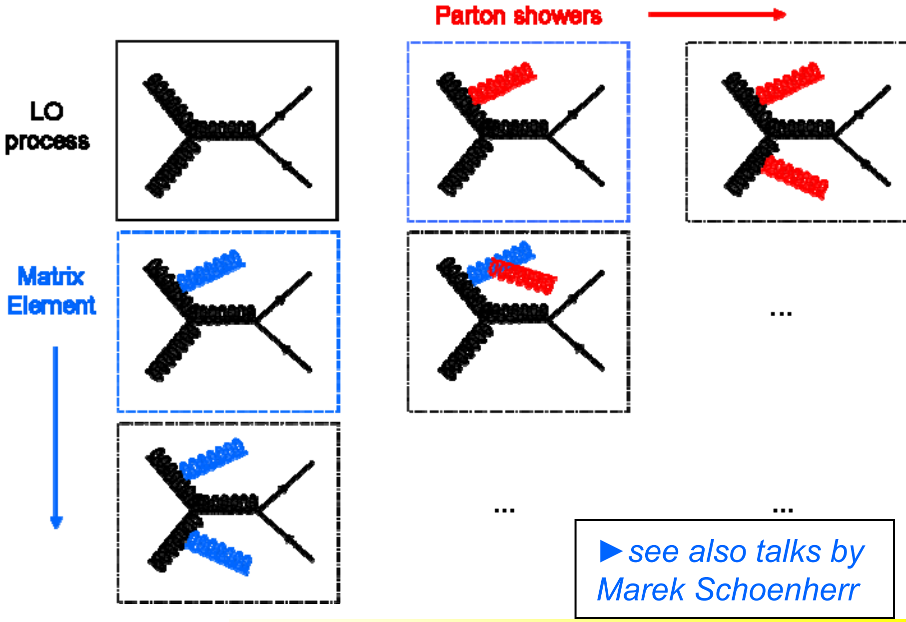
- Many of the experimental systematic uncertainties cancel in the ratio
  - The data are use to constrain the modelling of quark and gluon radiation in top-quark pair events

Radiation systematics summary

	Tevatron		LHC	
	CDF	**D0**	**ATLAS**	**CMS**
QCD radiation (ISR/FSR)	***Pythia More/Less PS***		***ACEMC + Pythia More/Less PS***	***Q^2 variation in Madgraph+Pythia + variation in the ME/PS jet-to-parton matching scale (for initial measurements also Pythia More / Less PS)***
Unc. on m_{top} [GeV]	***0.1 – 2.7***	***0.3 – 0.4***	***0.5 / 1.0 – 2.5 (Before gap fraction analysis)***	***0.3 – 1.2***

- In multi-leg MC, the same multi parton final state can be produced by the matrix element or by parton shower
  - To avoid double counting need a matching procedure

- matching scale + Q^2 scale variations describe to a large extent the same physics effect as ISR/FSR variations (preliminary investigation within the TopLHC WG)
  - This approach is used with MadGraph + Pythia by the CMS collaboration

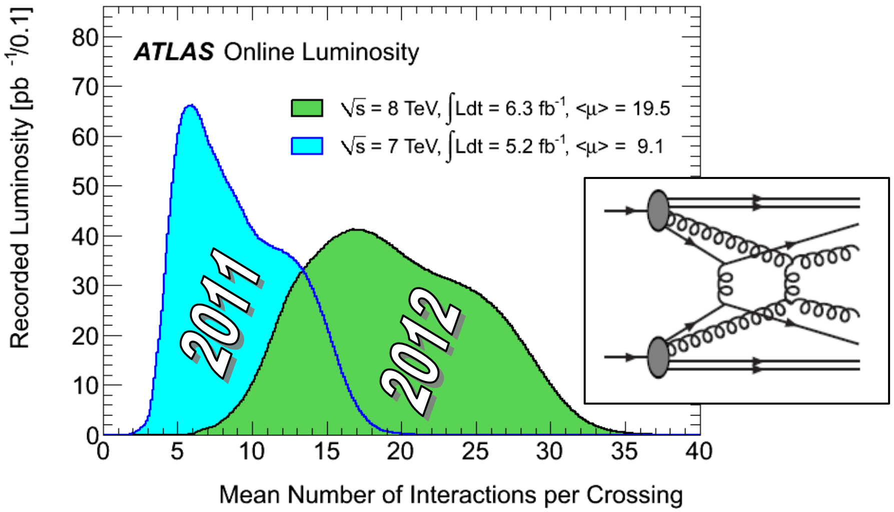


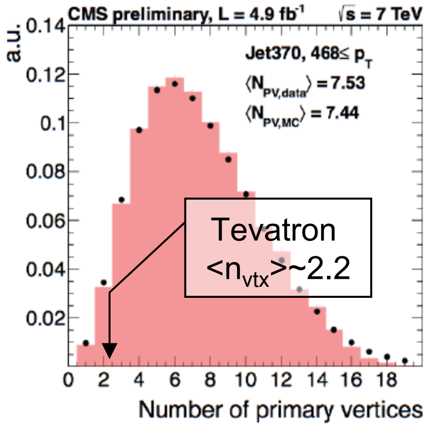
► *see also talks by Marek Schoenherr*

Event modelling and environment - 4

Pileup

- Overlay of additional low-p_T events (from MC or data for D0):
    - contribute to building up the total amount of energy and
    - cause color exchanges between the remnants, thereby increasing the number of particles produced in the hadronization stage.
- Affects both the JES determination (sep. component), the m_{top} resolution and the event kinematics
- Reweighting of MC event to match the data luminosity profile is required.





Tevatron
$<n_{vtx}>$~2.2

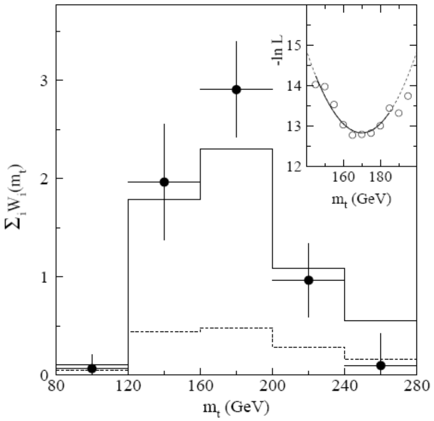
CMS DP2012_006

	Tevatron		LHC	
	CDF	**D0**	**ATLAS**	**CMS**
Pileup	*Data/MC diff in n_{vtx}*	*Unc in the reweighting*	*Data/MC after reweighting*	
Unc. on m_{top} [GeV]	*0.1 – 0.2*	*0.1 – 0.2 (1.2 for RunI di-lep)*	*0.05 – 0.7*	*0.1 – 1.0*

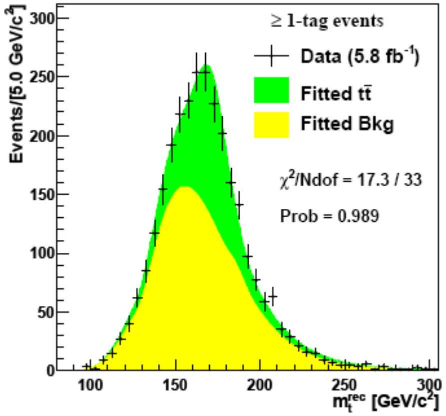
Background contamination

- Uncertainties on the background normalization and shape induce m_{top} unc (via the fitting procedure)
  - analysis, kin. sel. and decay channel dependent
- Uncertainties can be classified as
  - from MC: W/Z+jets shape, and sub-dominant contributions from single top and di-boson production
  - from data: mainly W+jets norm and QCD contributions and/or fake leptons (statistical unc. components can be reduced increasing the samples size)

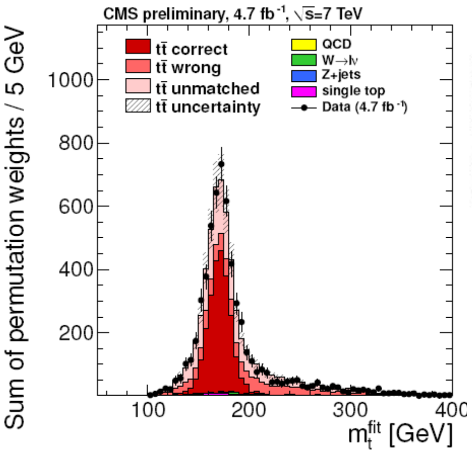
D0 di-lep Run I



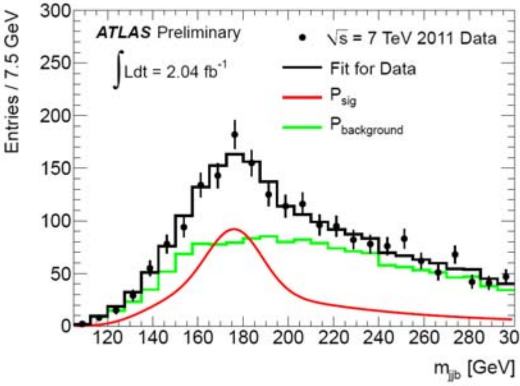
CDF all-jet Run II



CMS μ+jets



ATLAS all-jet

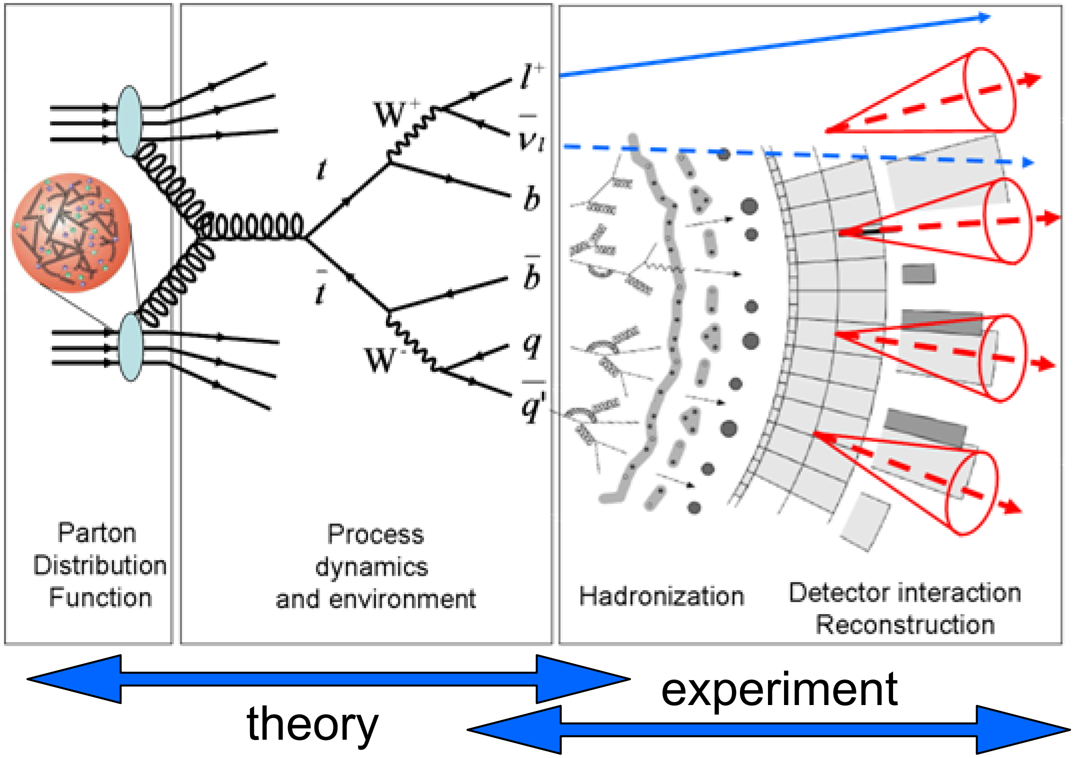


	Tevatron		LHC	
Syst type / Values in GeV	**CDF**	**D0**	**ATLAS**	**CMS**
Background from MC (theory)	***<0.01 – 1.7***	***<0.01 – 1.0***	***0.1 – 1.8***	***0.1 – 0.2***
Background from data	***0.1 – 0.6***	***0.2***	***0.6 – 1.9***	***0.2 – 0.4***

▲ *see also talks by: Frederic Deliot and Joseph Boudreau*

Max-Planck-Institut für Physik

Physics objects and detector modelling

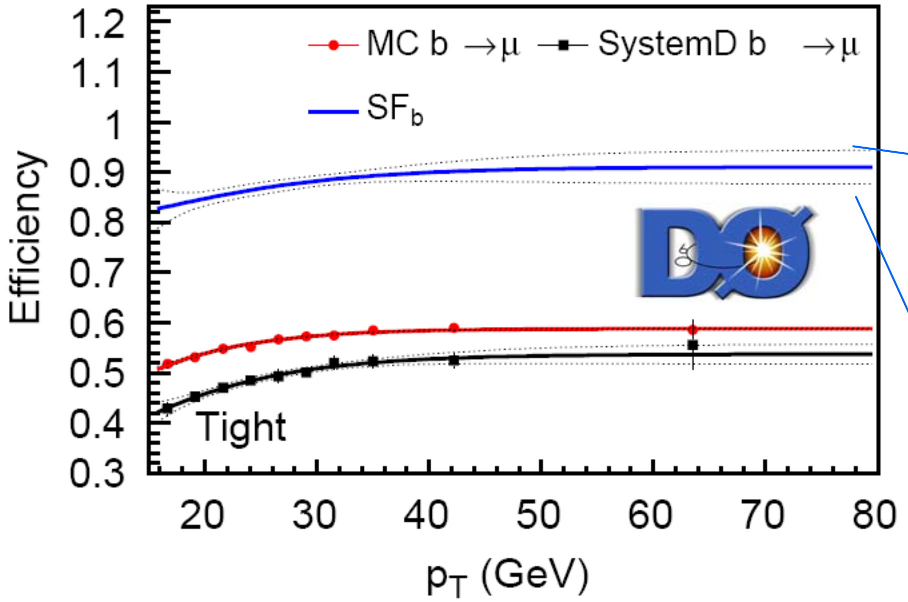


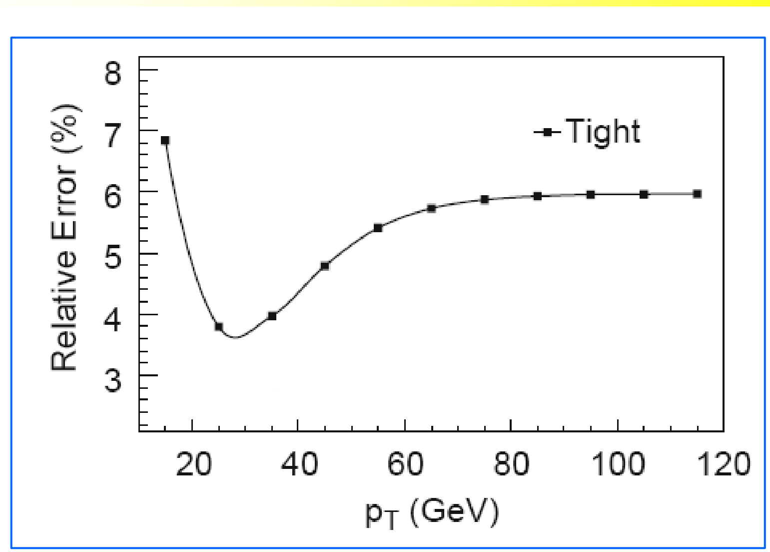
- b-tagging
- Jet energy resolution
- Jet reconstruction efficiency
- E_T^{miss} modelling
- Leptons momentum and energy scale

- Jet Energy Scale
  [deserves separate treatment]

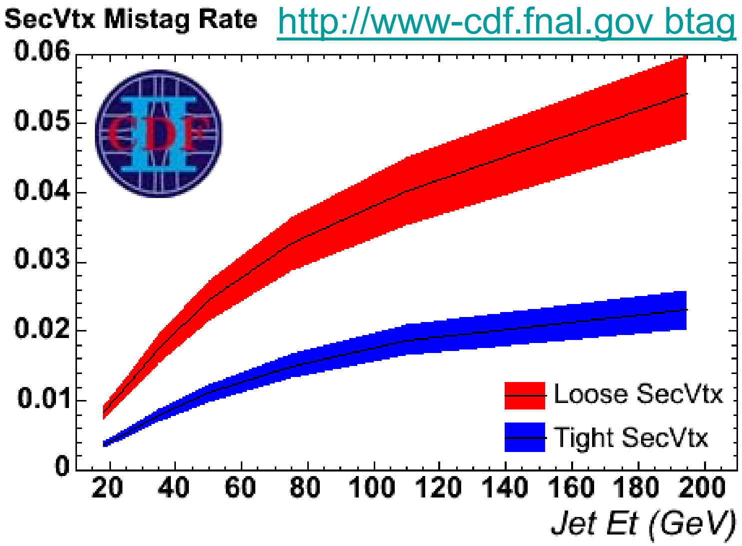
► *see also talks by:*
Carmen Diez Pardos (CMS)
Veronique Boisvert (ATLAS)

b-tagging calibration



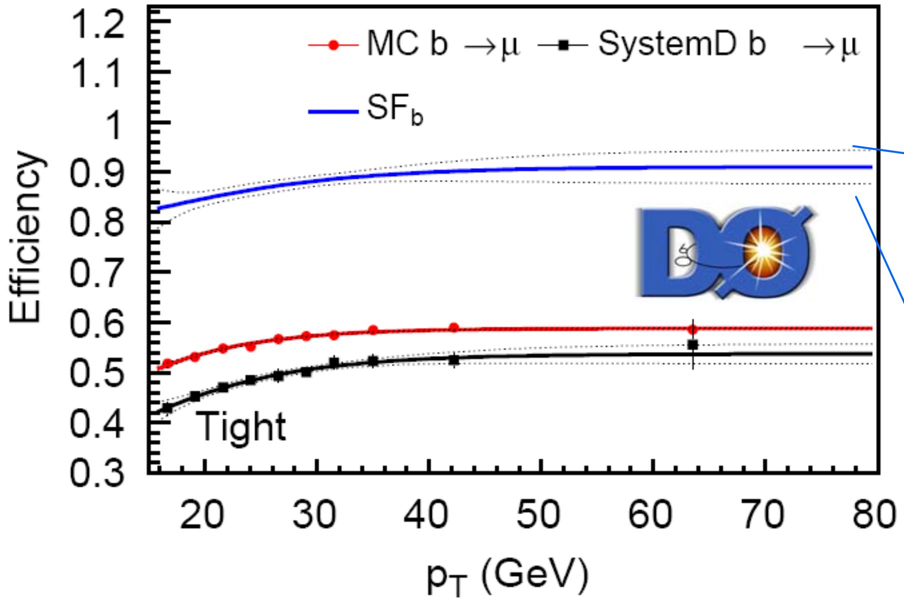


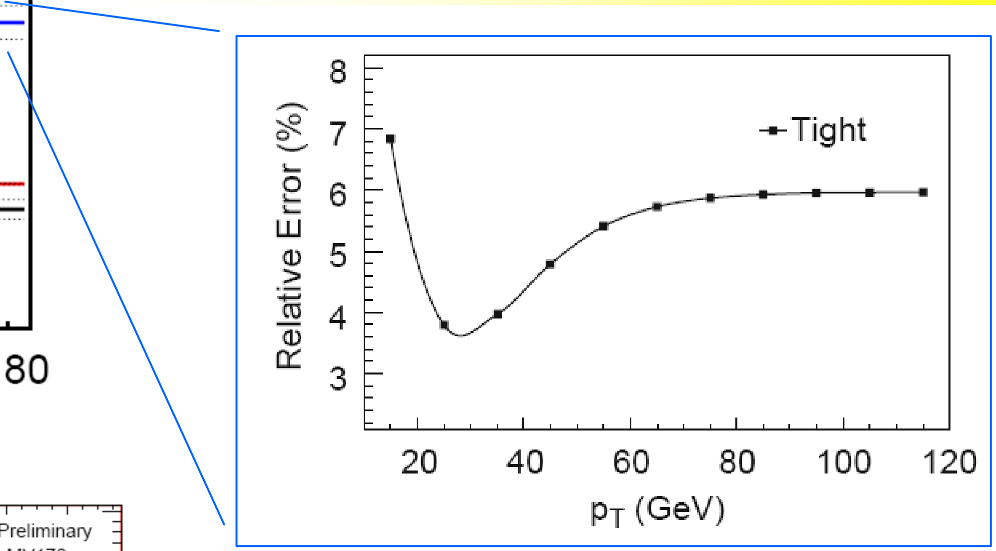
- *b*-tagging is an important handle for background reduction
  - Algorithms need to be calibrated, to account for detector effects and possible mis-modeling of the jet/trk/event variables in the MC with respect to the data.
- Data samples are used to determined:
  - *b-tagging efficiency*
  - *mis-tag rate*
  - and the corresponding uncertainties.
    - ***Include b-jet modelling uncertainties such semilep BR, b-jet fragmentation***



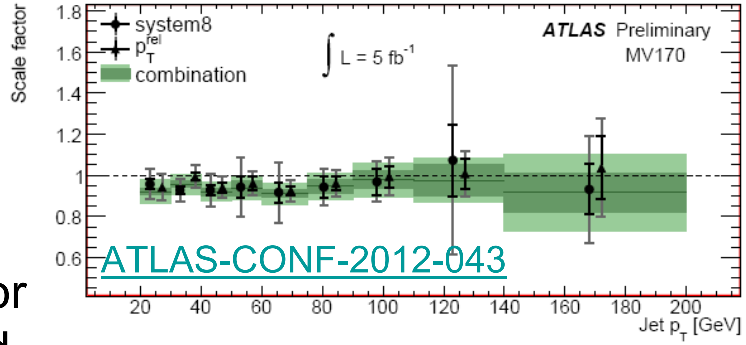
http://www-cdf.fnal.gov btag

b-tagging calibration

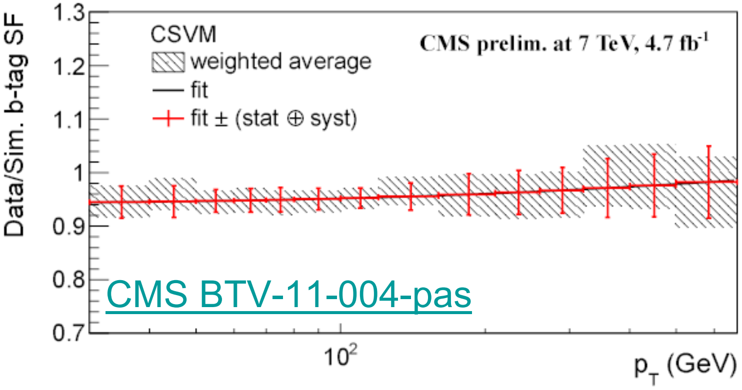




Similar concepts for ATLAS and CMS

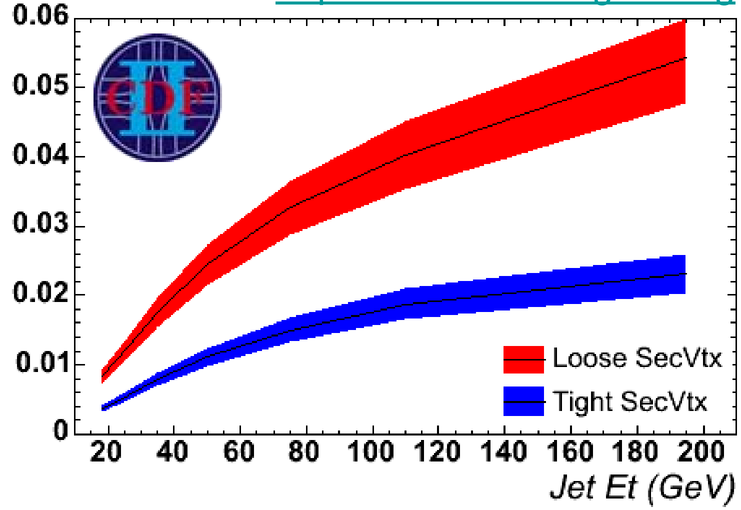


ATLAS-CONF-2012-043



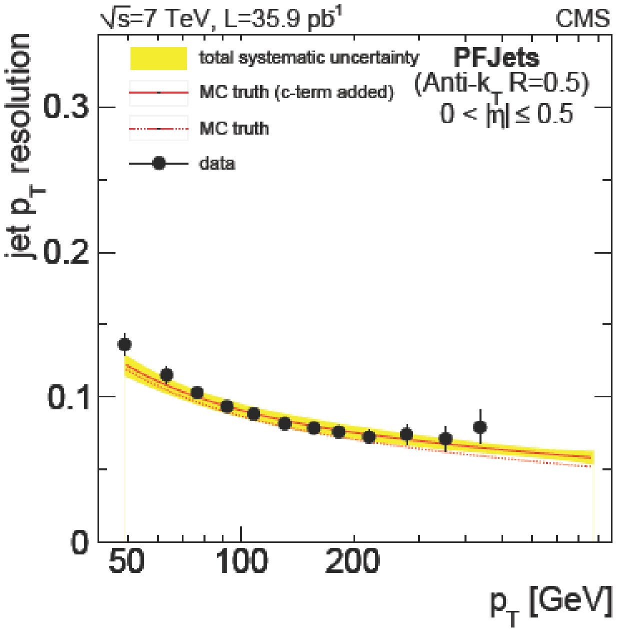
CMS BTV-11-004-pas

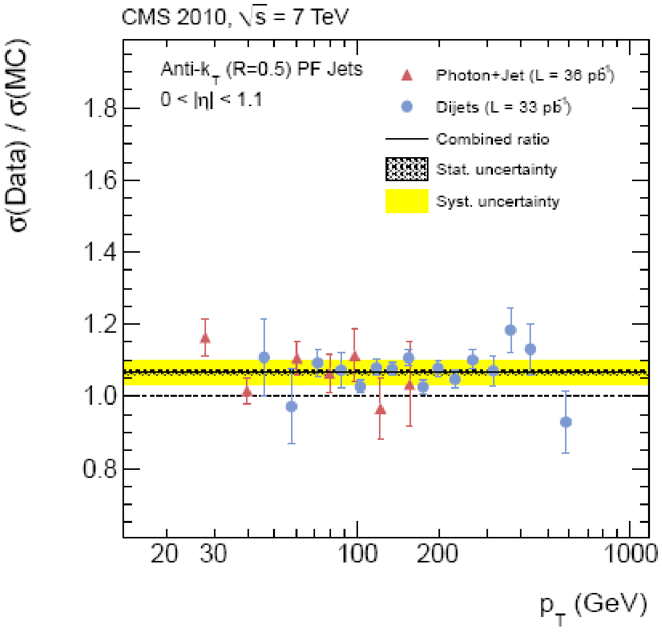
SecVtx Mistag Rate http://www-cdf.fnal.gov btag



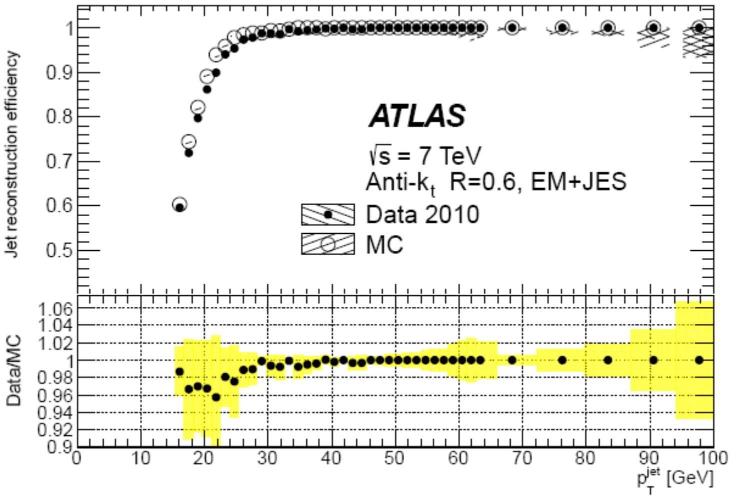
Jet energy resolution and reco efficiency

- **Jet Resolution**
  - Difference in the jet energy resolution between data and MC have to be accounted for
  - ►can result in broader/narrower invariant mass distrib.





- **Jet Reconstruction efficiency**
  - The jet reco efficiency is measured relative to track jets in data and MC.
  - The Data/MC difference, is propagated to the analyses



important only for low p_T jets

Summary of detector uncertainties

	Tevatron		LHC	
	CDF	**D0**	**ATLAS**	**CMS**
B-tagging / b-jet modelling	***Data/MC scale factors varied within uncertainties***			
Unc. on m_{top} [GeV]	***0.2 – 2.5 (for L_{xy})***	***0.1 – 0.2***	***0.3 – 0.5***	***0.1 – 0.5***
Quite some statistical components in the b-tagging calibration, can be reduced by increased sample sizes				
Jet energy resolution / reco eff	***ne***	***MC Jet energy smearing and random jet drop***		***MC Jet energy smearing***
Unc. on m_{top} [GeV]	***negligible***	***0.2 – 0.3***	***0.1 – 0.9***	***0.1 – 0.5***

Summary of detector uncertainties

Missing E_T modelling

- the E_T^{miss} absorbs the scale and resolution uncertainty of all physics objects like e, γ, μ, τ and jets.
- There are also “soft terms” including un-clustered or energy not associated to physics objects

	Tevatron		**LHC**	
	CDF	**D0**	**ATLAS**	**CMS**
B-tagging / b-jet modelling	***Data/MC scale factors varied within uncertainties***			
Unc. on m_{top} [GeV]	***0.2 – 2.5 (for L_{xy})***	***0.1 – 0.2***	***0.3 – 0.5***	***0.1 – 0.5***
	Quite some statistical components in the b-tagging calibration, can be reduced by increased sample sizes			
Jet energy resolution / reco eff	***ne***	***MC Jet energy smearing and random jet drop***		***MC Jet energy smearing***
Unc. on m_{top} [GeV]	***negligible***	***0.2 – 0.3***	***0.1 – 0.9***	***0.1 – 0.5***
E_T^{miss} modelling (soft term)	***ne***	***ne***	***Variation of the soft term of E_T^{miss} within uncertainties***	
Unc. on m_{top} [GeV]	***–***	***–***	***0.1***	***0.1 – 0.4***

Summary of detector uncertainties

Missing E_T modelling

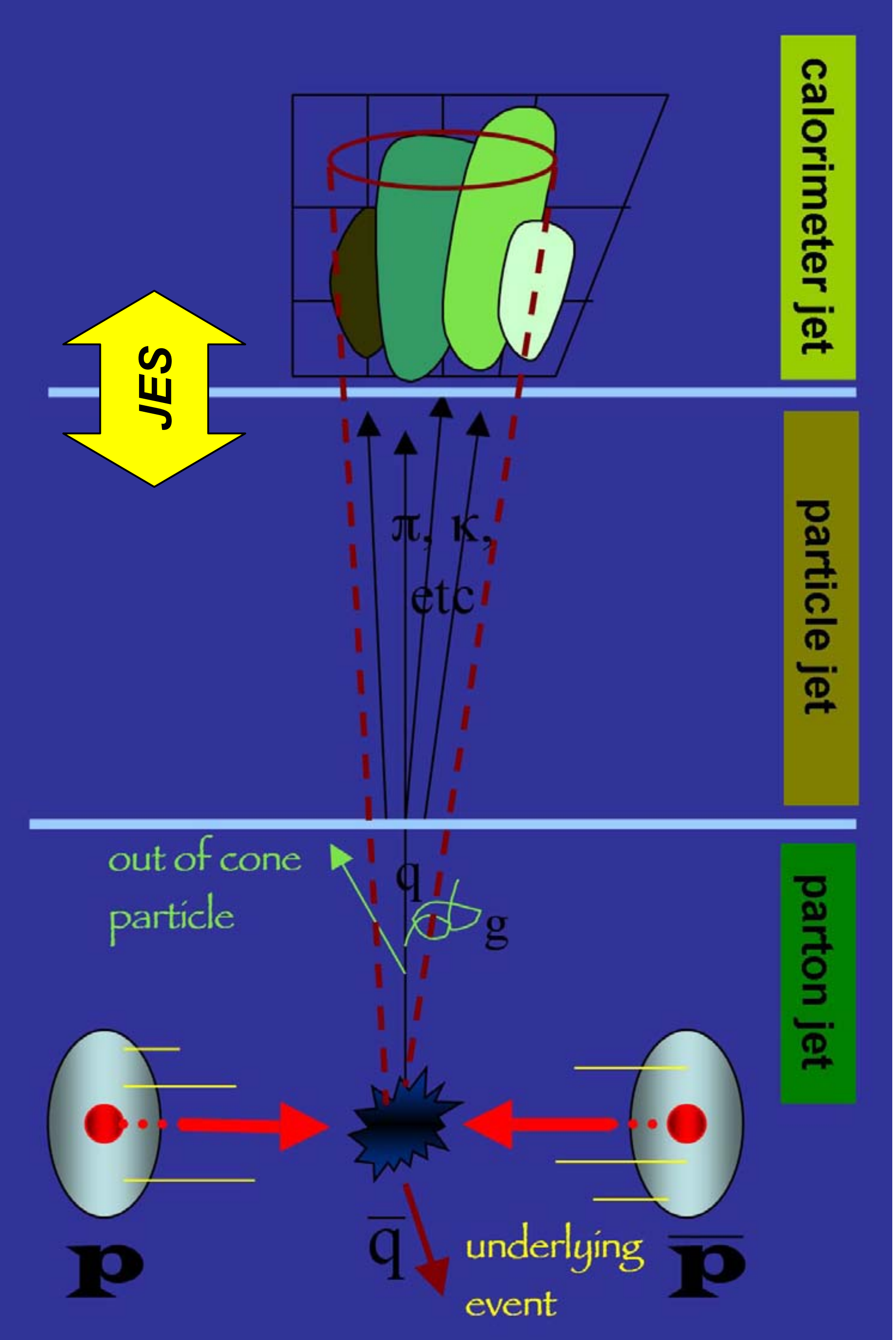
- the E_T^{miss} absorbs the scale and resolution uncertainty of all physics objects like e, γ, μ, τ and jets.
- There are also “soft terms” including un-clustered or energy not associated to physics objects

	Tevatron		**LHC**	
	CDF	**D0**	**ATLAS**	**CMS**
B-tagging / b-jet modelling	***Data/MC scale factors varied within uncertainties***			
Unc. on m_{top} [GeV]	***0.2 – 2.5 (for L_{xy})***	***0.1 – 0.2***	***0.3 – 0.5***	***0.1 – 0.5***
	Quite some statistical components in the b-tagging calibration, can be reduced by increased sample sizes			
Jet energy resolution / reco eff	***ne***	***MC Jet energy smearing and random jet drop***		***MC Jet energy smearing***
Unc. on m_{top} [GeV]	***negligible***	***0.2 – 0.3***	***0.1 – 0.9***	***0.1 – 0.5***
E_T^{miss} modelling (soft term)	***ne***	***ne***	***Variation of the soft term of E_T^{miss} within uncertainties***	
Unc. on m_{top} [GeV]	–	–	***0.1***	***0.1 – 0.4***

Lepton momentum / energy scale

- In general a sub-dominant uncertainty contribution that varies between <0.1 GeV (l+jets) to 0.2 – 0.3 GeV (di-lep)

JES

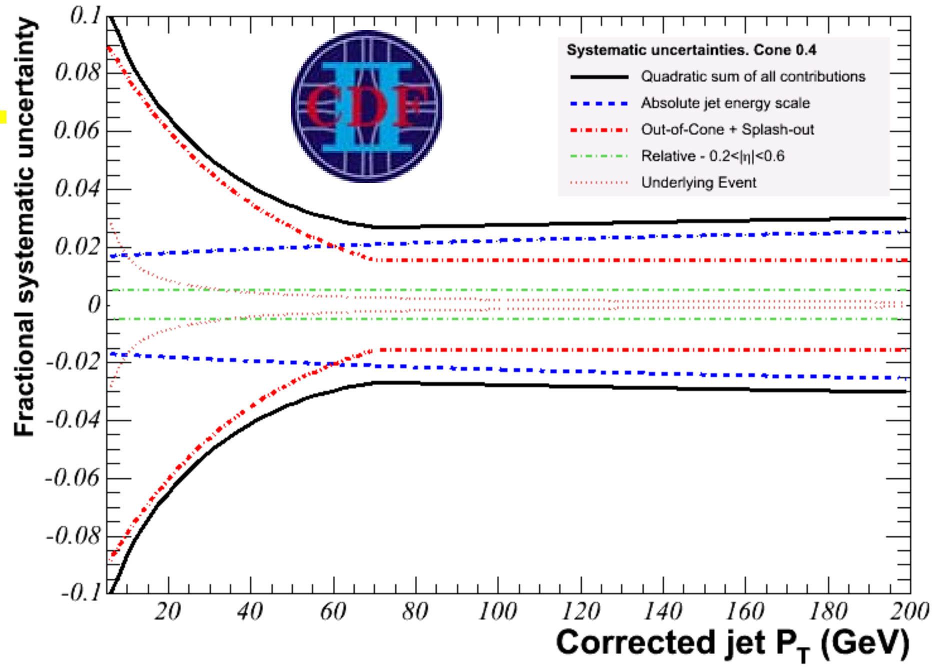


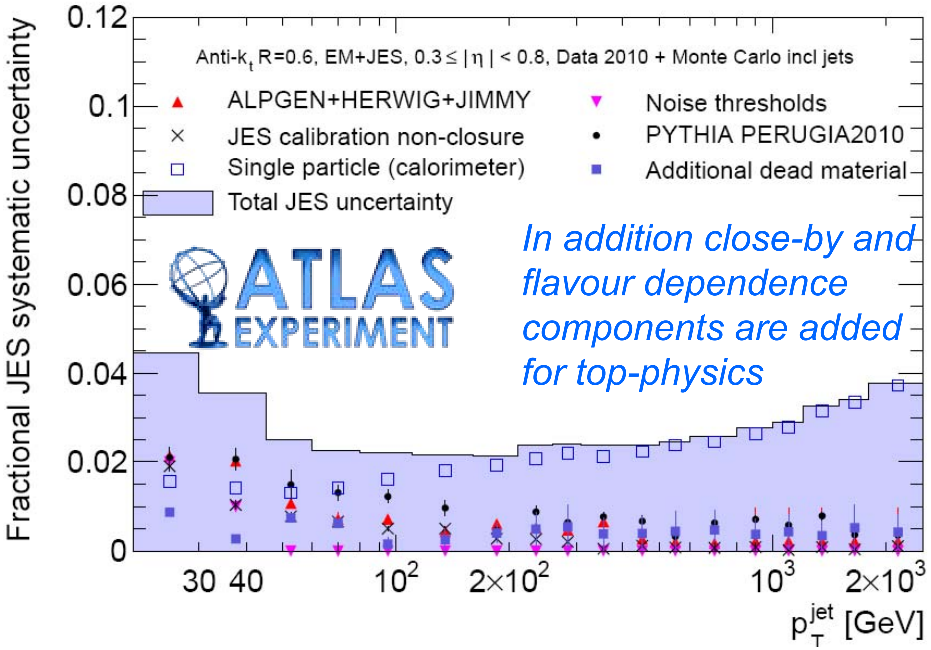
- Describes difference between reconstructed jet and particle jet (or partons in the CDF case)

- Non-trivial determination.
  - involve both MC and in-situ techniques (ie γ/Z+jet, di-jet balance)

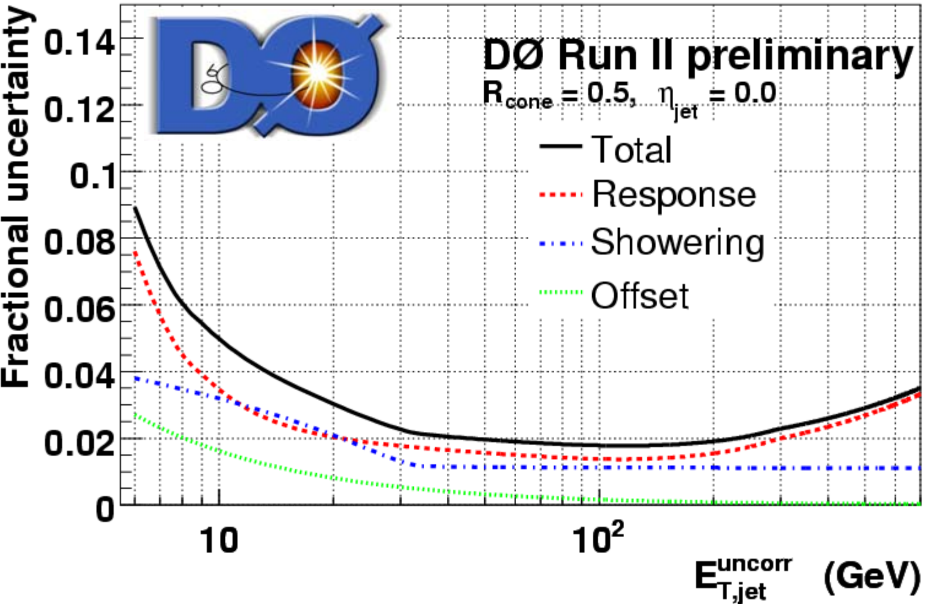
- The jet energy scale uncertainty is derived from in-situ measurements along with systematic variations of the Monte Carlo simulation

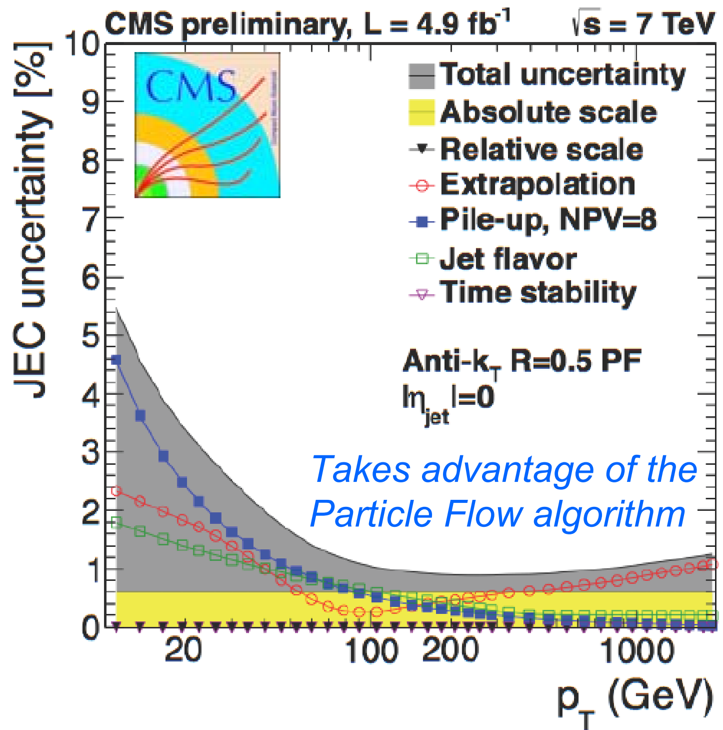
- Typically uncertainties are O(few %) and vary with jet properties (p_T, η)
  - ~100% correlated with m_{top} (when using jet info for its reconstruction)
  - ***σ(JES)= 3% → $\sigma(m_{top})$~5 GeV***





In addition close-by and flavour dependence components are added for top-physics

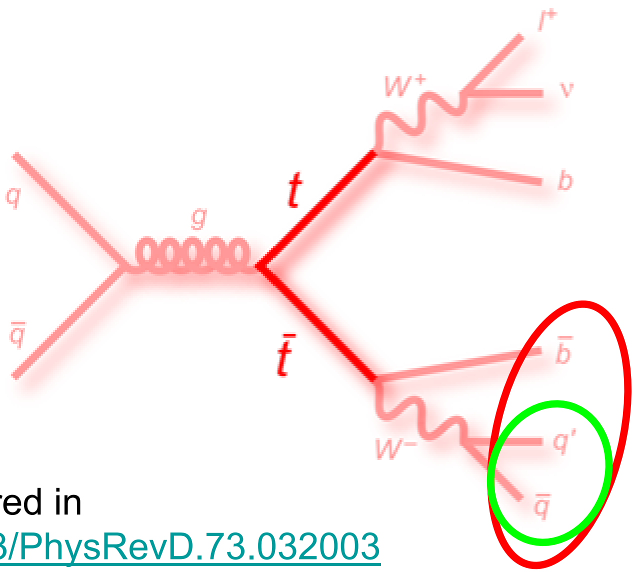




Takes advantage of the Particle Flow algorithm

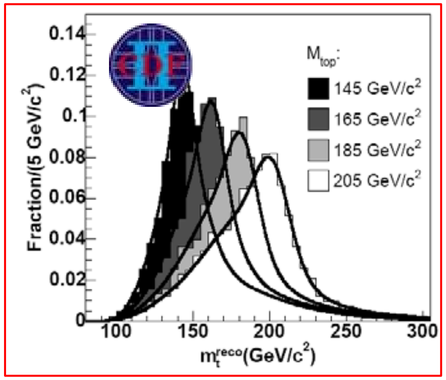
In-situ constraint of the light-quark JES

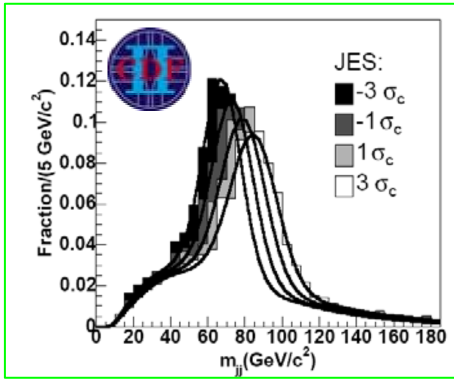
- Most of the recent m_{top} measurements in the l+jets (E_T^{miss}+jets) and all-jets channels have developed techniques to constrain in-situ the light-quark jet energy scale using in top events information from the W→jj decay.



An example also exists for the transfer of the JES constraint in the l+jets channel to the di-lepton one (taking into account the environmental difference http://arxiv.org/abs/1201.5172v1)

Pioneered in 10.1103/PhysRevD.73.032003





- This allow a large reduction of the impact of the JES uncertainty on the m_{top} determination, at the price of an additional contribution to the stat uncertainty (2D fit)

Statistical component

	Tevatron		LHC	
Syst type / Values in GeV	**CDF**	**D0**	**ATLAS**	**CMS**
light JES	*0.2 (Lxy) 0.5– 4.4*	*0.6 – 3.4*	*0.7 – 2.1*	*0.2 – 2.1*
Jet Scale Factor (from in-situ W→jj)	*0.5 – 1.5*	*0.5 – 0.6*	*0.4*	*0.4*

In-situ constraint of the light-quark JES - 2

- The method generally works well,
  however at the precision level of the latest measurements (and JES determinations) can become less powerful

  - In the recent CMS m_{top} result in the all-had channel (CMS TOP-11-017-pas), the gain of 2D method for the jet energy scale uncertainty reduction was not compensated by the increase of other systematics effects altering the m_W (and m_{top}) distribution, namely the CR and UE event syst.
    ►The 1D result at the end was found to be more precise

    - This may be attributed to a sensitivity of the 2D method to additional relative differences in the light- and b-jet responses that can be introduced by the event modelling.

b- (or flavor-) specific JES

- The standard JES is set using light-quark/gluon jets (similar arguments for the 2D in-situ method)
- The jet response to heavy-flavor jets can be different:
  - Semileptonic decays
  - Differences in the shower and particle content of b-jet wrt the light-jets
- CDF measures the b-to-light-quark energy scale difference applying single particle responses measured in data or predicted by MC to samples of jets from top pair events.

$$\frac{\left(p_T^{data} / p_T^{MC}\right)_{b-jets}}{\left(p_T^{data} / p_T^{MC}\right)_{light-jets}} = 1.01$$

	Tevatron	
Syst type / Values in GeV	**CDF**	**D0**
b-jet JES (response and model)	*0.1 – 0.8*	*0.3 – 0.7*

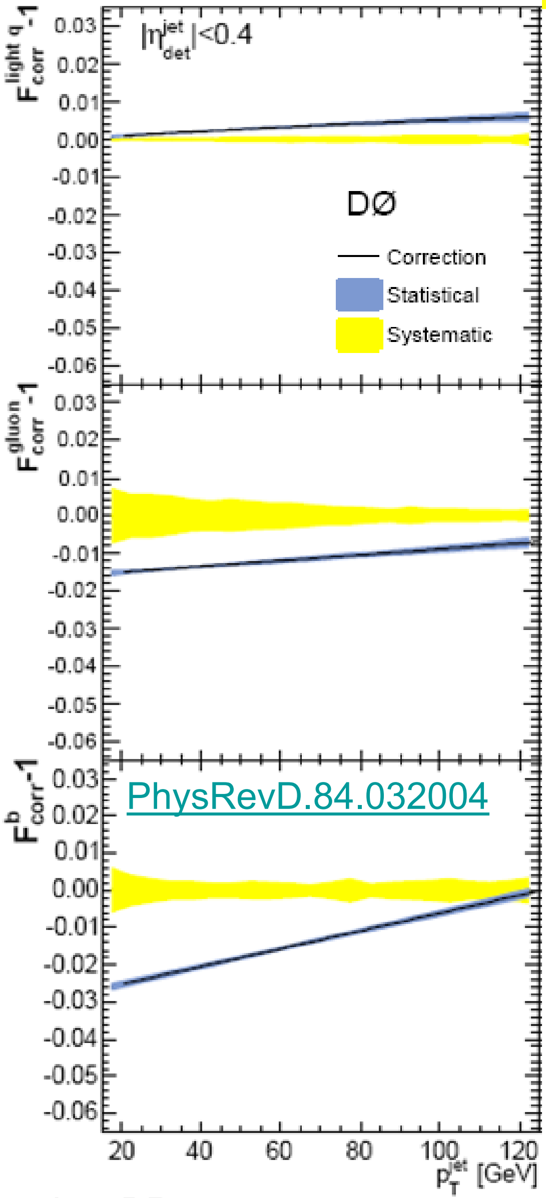
- D0 applies flavor dependent jet energy corrections to the MC to bring the response closer to that of the data and thus minimize the uncertainty.

single particle response

$$D = \frac{\sum E_i \cdot R_i^{data}}{\sum E_i \cdot R_i^{MC}}$$

$$F_{corr}^{\beta} = D_\beta / \left\langle D_{\gamma+jet} \right\rangle$$

JES for light jet mixture

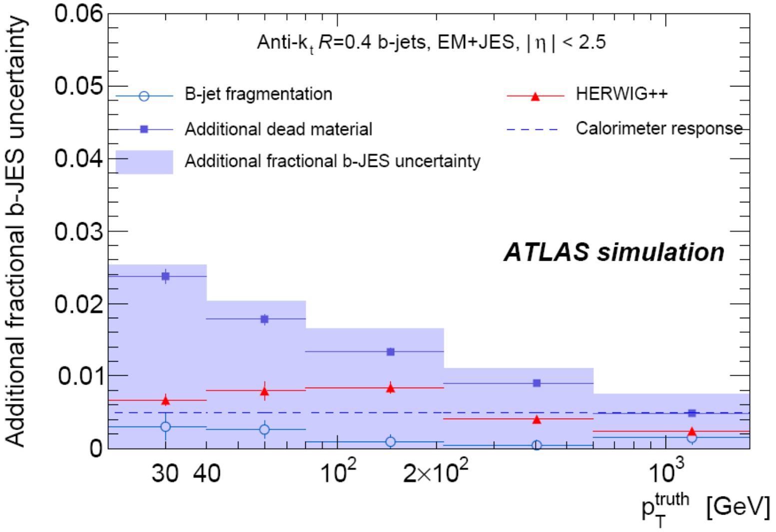


PhysRevD.84.032004

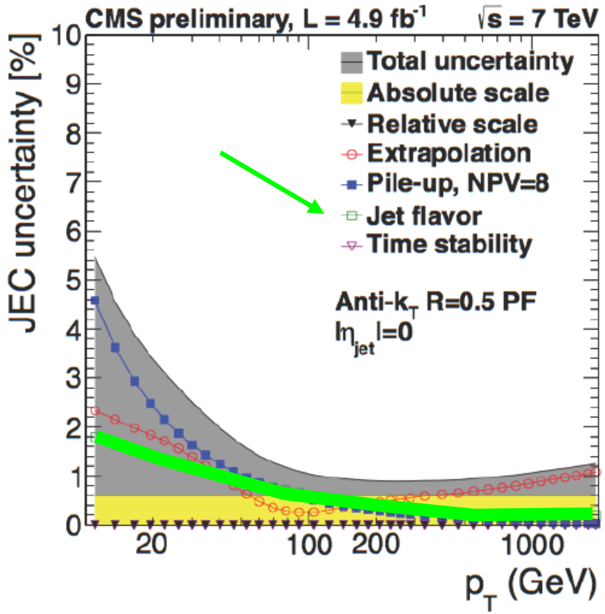
In addition, fragmentation and semilep BR uncertainties are considered (b-jet modelling)

b- (or flavor-) specific JES at LHC

- ATLAS:
  - Additional b-JES uncertainty evaluated from:
    - calorimeter response (single particle)
    - b-jet fragmentation
    - Pythia/Herwig++
    - limited knowledge of the detector material



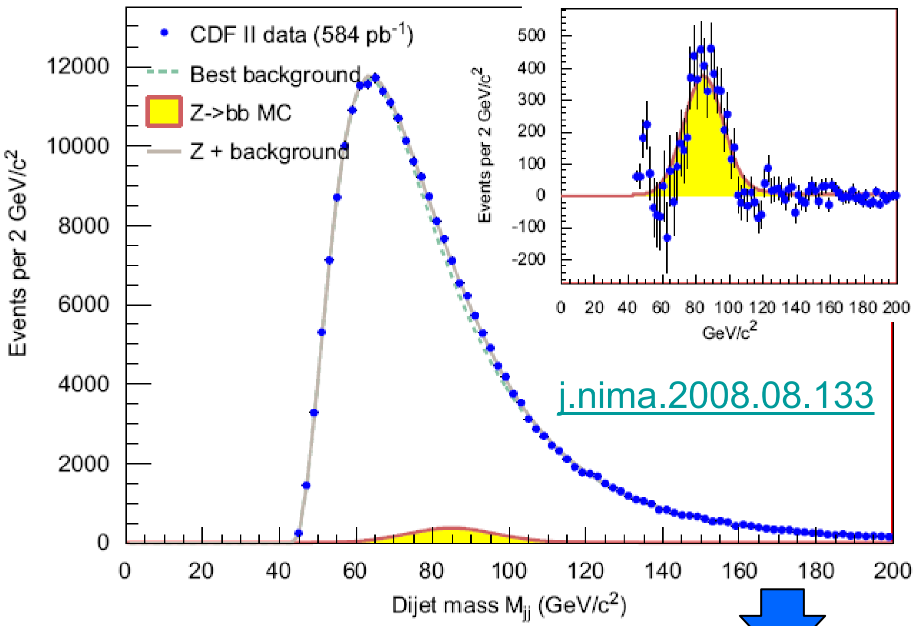
- CMS:
  - Use the max Pythia/Herwig response differences, for pure quark flavor at low p_T and for pure gluon flavor at high p_T, as a flavor uncertainty applicable to any jet flavor mixture.
  - Detector effects are considered to be negligible due to the extensive use of tracking to reconstruct individual particles in the jet (Particle Flow)



	Tevatron		**LHC**	
Syst type / Values in GeV	**CDF**	**D0**	**ATLAS**	**CMS**
b-jet JES (response and model)	***0.1 – 0.8***	***0.3 – 0.7***	***1.4 – 2.5***	***0.5 – 1.1***

bJES checks with data?

- Z→bb @ CDF used to check the b-jet energy scale. Very large background from QCD di-jet production.
  - impractical @ LHC due to larger background cross section.



j.nima.2008.08.133

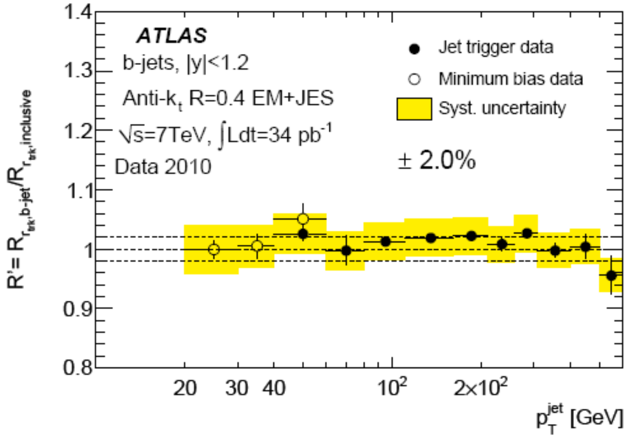
$$\text{bJSF} = 0.974 \pm 0.011(\text{stat})^{+0.017}_{-0.014}(\text{syst}) = 0.974^{+0.020}_{-0.018}(\text{total}).$$

- ATLAS used tracks in jets to validate the residual b-to-light-quark JES:

$$r_{trk} = \frac{\left|\sum \vec{p}_T^{\,track}\right|}{\left|\vec{p}_T^{\,jet}\right|}$$

$$R = \frac{\langle r_{trk} \rangle_{Data}}{\langle r_{trk} \rangle_{MC}}$$

$$R' = R^{b-jets} \,/\, R^{light-jets}$$



It opens the possibility for an in-situ b-jet specific calibration!

Conclusions

- Experiments at Tevatron and LHC have been and are doing a great job in reducing the m_{top} uncertainty related to the physics objects and the detector description
  - This will continue in the future taking advantage of the large statistics available at the LHC.
  - In-situ constraint of the light jets energy scale helped in the mitigation of the JES (but also of the radiation) systematics on $m_{top.}$
  - Next steps will include beating down the flavor dependent JES uncertainty (b-JES), either by using data to constrain it, or (as done by D0) by employing flavour specific jet energy corrections.

- To reach and consolidate the sub-GeV m_{top} precision this must go hand in hand with the reduction of the theory uncertainties
  - Close collaboration between experiment and theory will be needed.
  - Reduce possible sources of double counting (in each systematic category try to separate out experimental and theoretical components)
  - Use available datasets to further tune MC generators

- Last but not least: make sure we understanding what we are measuring
  - the m_{top} definition and interpretations ►see talk by Sven-Olaf Moch

Thanks for your attention!

- backup -

More info on the CMS ParticleFlow JES

CMS DP2012_006