Recent top quark measurements at D0

- Introduction
- Differential cross sections
- Forward-backward asymmetries
- Mass
- Conclusions & Outlook

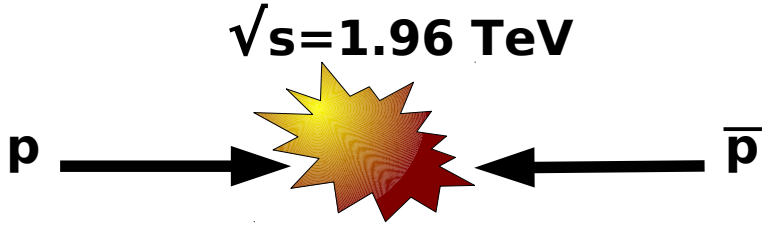
Fermilab

Andreas Jung

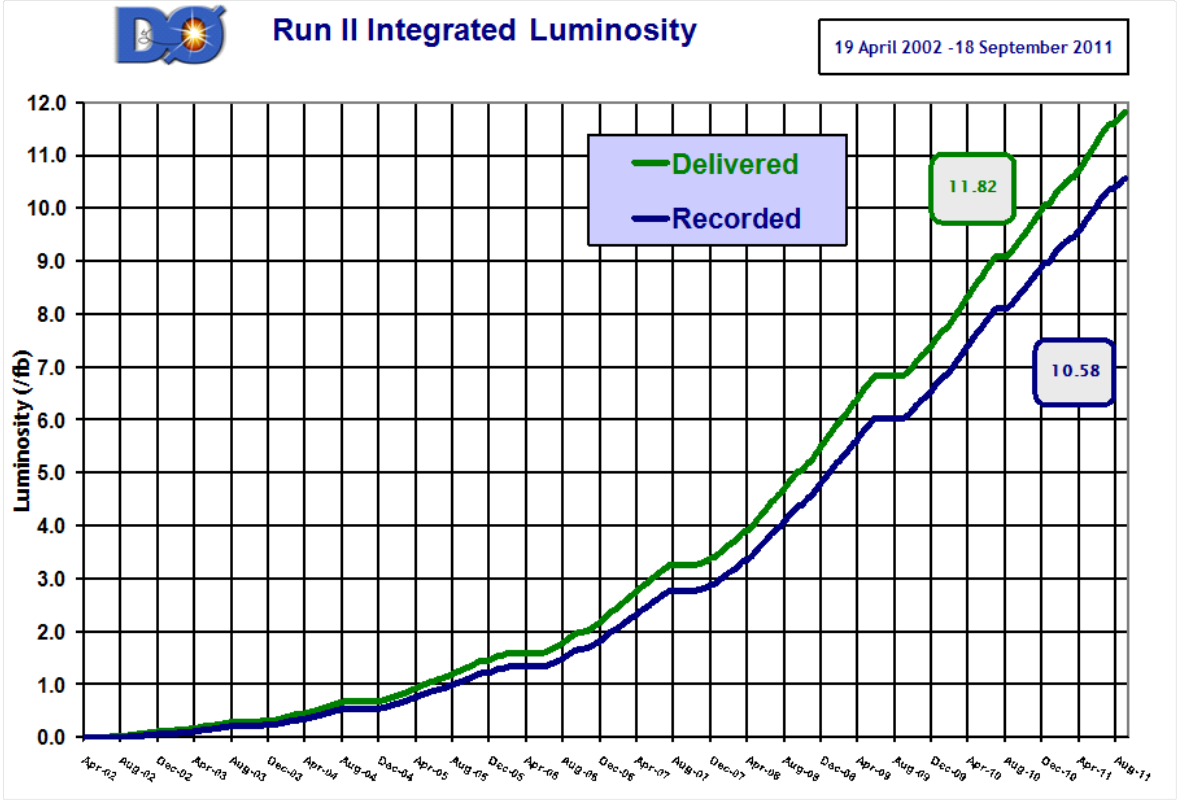
October 29th, 2014

Topic of the week seminar, LPC Fermilab

Tevatron – Introduction



- Peak luminosities: 3 – 4 x 10^{32} $cm^{-2}s^{-1}$
- ~10 fb^{-1}/experiment recorded

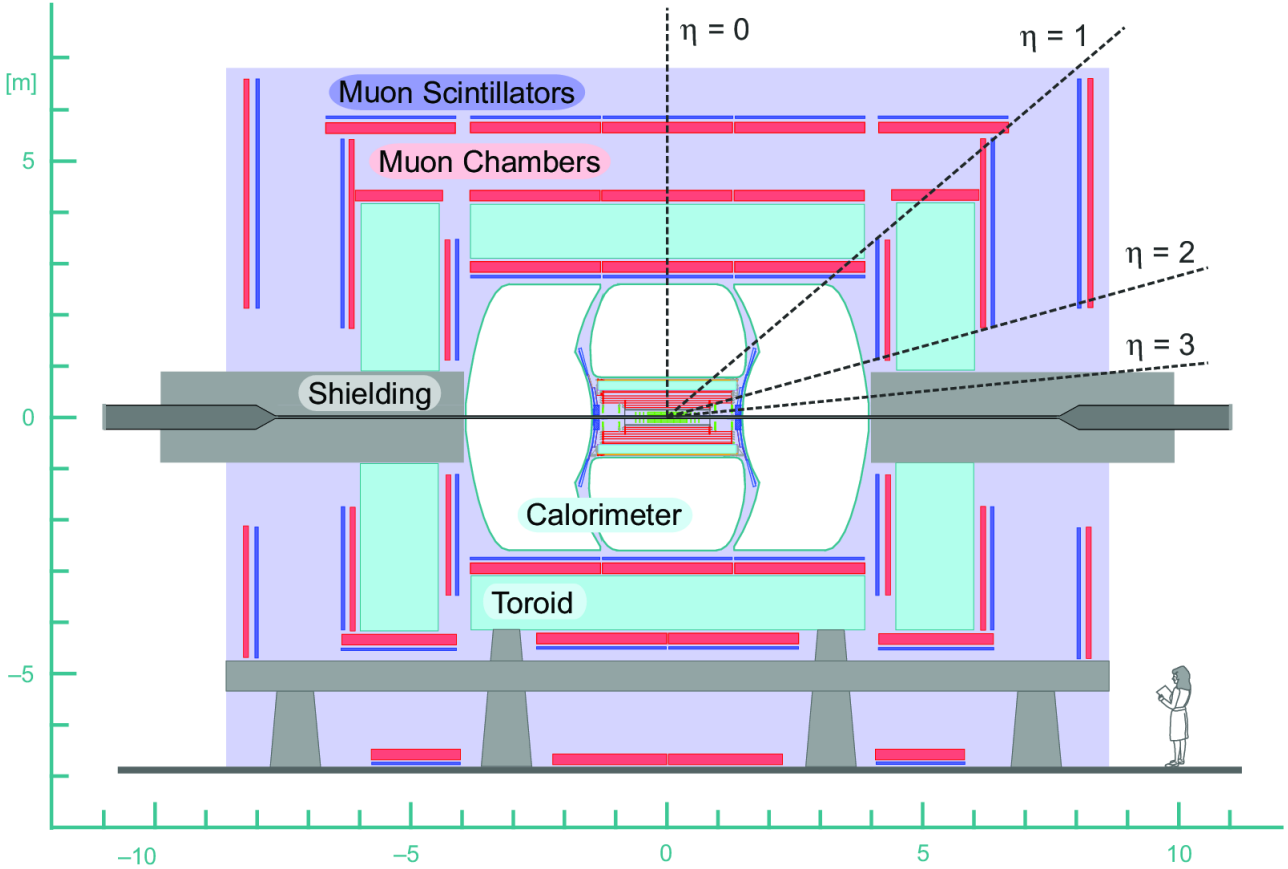


- Shutdown in September 2011

The D0 Detector

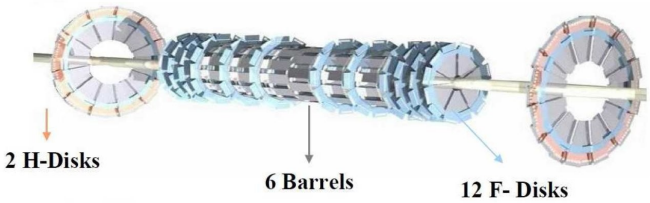
- General purpose 4π detectors:

Tracker: Detection and momentum measurement for charged particles
Calorimeter: Identification and energy measurement of jets and electrons
Muon system: Identification and momentum measurement of muons



Silicon Micro-strip Tracker:

- Identifies secondary vertices, typical resolutions for the impact parameter of typical tracks between 20-40 μm
- Vital for identification of *b* quarks

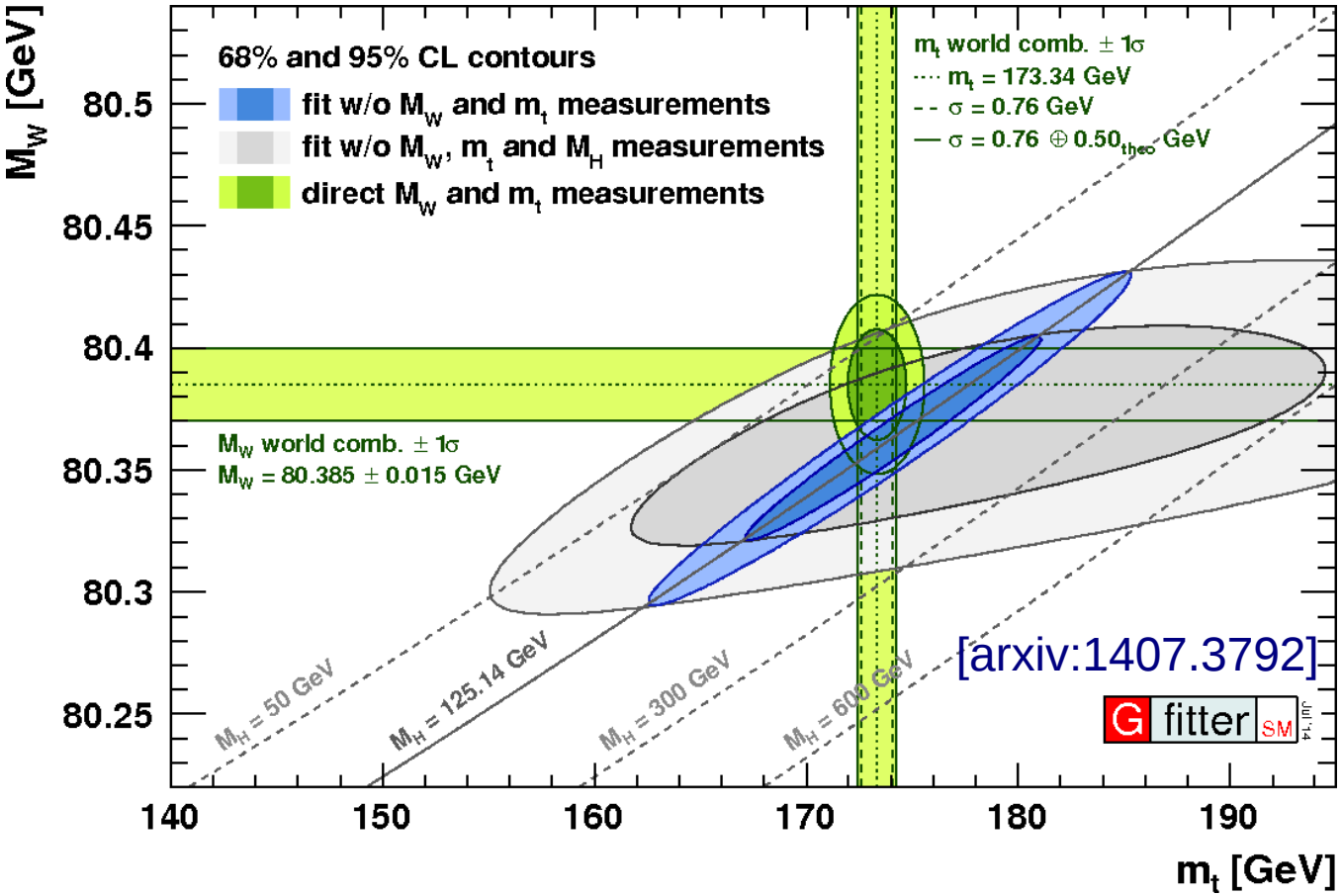


- More information: NIM A 634:8-46 (2011)
  Running Experience (arxiv:1202.5996)

- Similar detector coverage as CDF (& LHC detectors), D0 can reverse magnet polarity

Tevatron – Strengths

- Worlds largest $p\bar{p}$ data set for a long time
- Well understood detectors
- Less pile-up effects than LHC
- Initial state allows for unique measurements



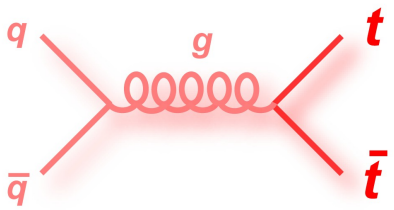
Why the top quark ?

- Top is the heaviest fundamental particle discovered so far:
  mass of gold atom [arxiv:1407.2682]
  → m_t = 174.3 ± 0.8 GeV
- Large Yukawa coupling to Higgs boson:
  → $\lambda_t \sim 1$
  - electron couples 10^5 less strong
  - special role in electroweak symmetry breaking ?
- Lifetime: $\tau \approx 5 \times 10^{-25}$ s << Γ_{QCD}
  → **Observe bare quark properties**

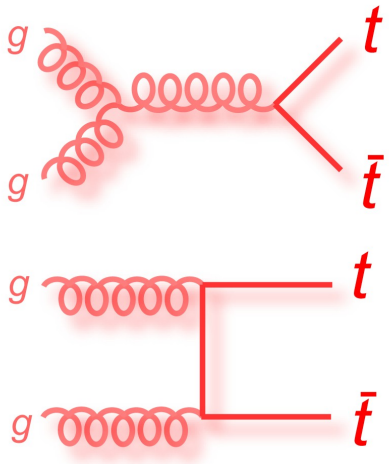
→ "Today's signal is tomorrows background"
Most dominant in many searches for new physics

Top quark production

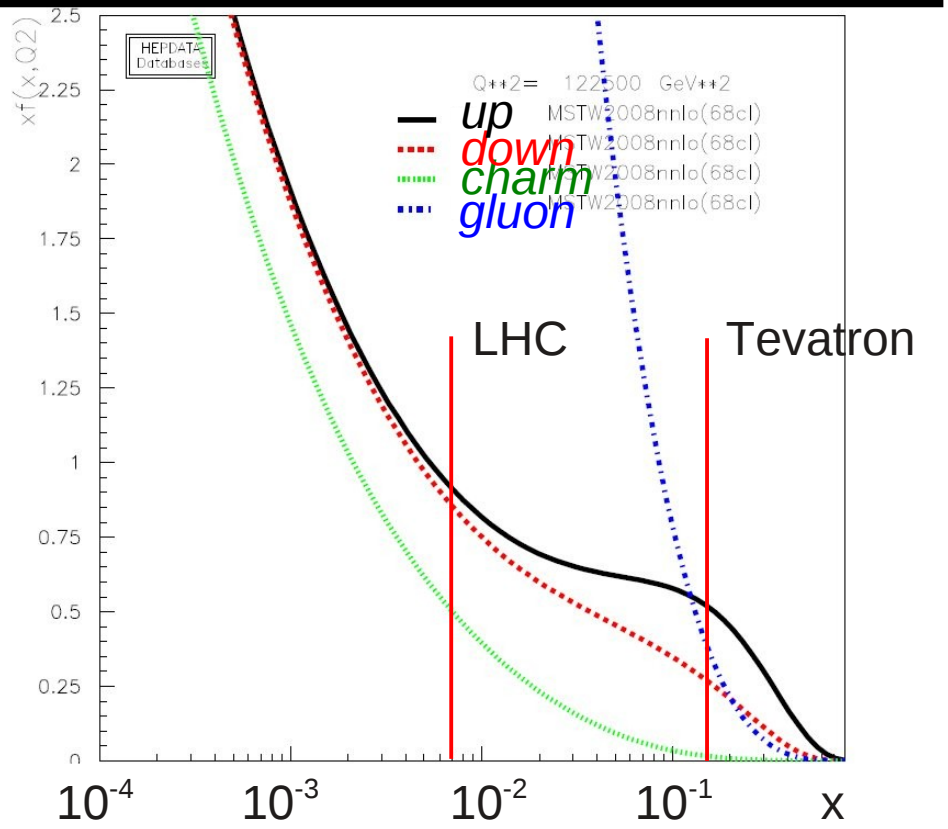
- Strong interaction: total cross section given by all amplitudes



$q\bar{q}$ annihilation



gg fusion



(Czakon, Fiedler, Mitov); [arxiv:1303.6254]
Phys. Rev. Lett. 110, 252004 (2013)

LHC vs. Tevatron:

$q\bar{q}$: ~85% ~15/13% (~10%, 14 TeV)

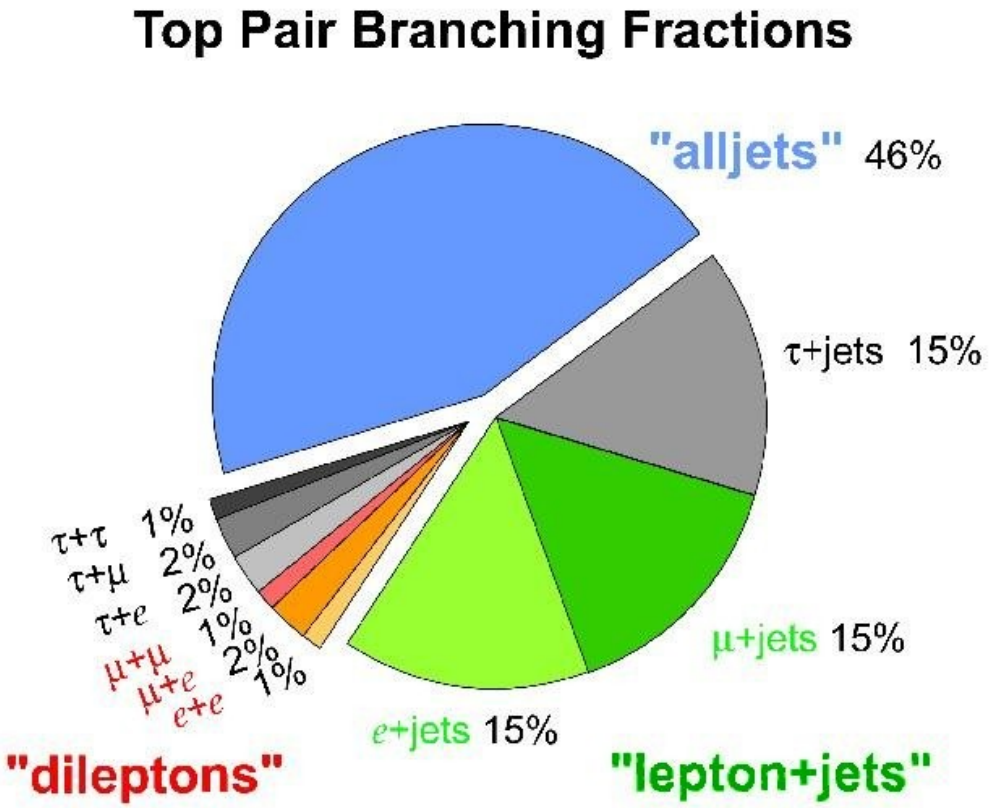
gg: ~15% ~85/87% (~90%, 14 TeV)
(1.96 TeV 7/8 TeV)

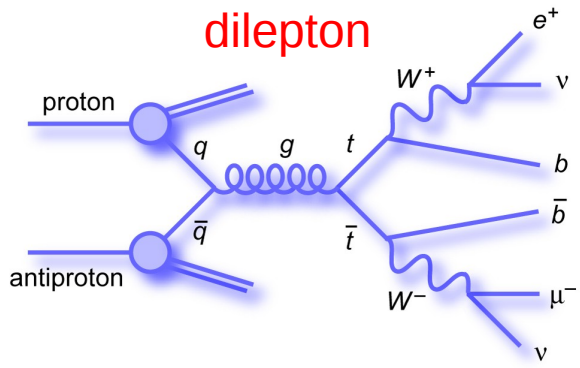
Collider	σ_{tot} [pb]	scales [pb]	pdf [pb]
Tevatron	7.164	$^{+0.110(1.5\%)}_{-0.200(2.8\%)}$	$^{+0.169(2.4\%)}_{-0.122(1.7\%)}$
LHC 7 TeV	172.0	$^{+4.4(2.6\%)}_{-5.8(3.4\%)}$	$^{+4.7(2.7\%)}_{-4.8(2.8\%)}$
LHC 8 TeV	245.8	$^{+6.2(2.5\%)}_{-8.4(3.4\%)}$	$^{+6.2(2.5\%)}_{-6.4(2.6\%)}$

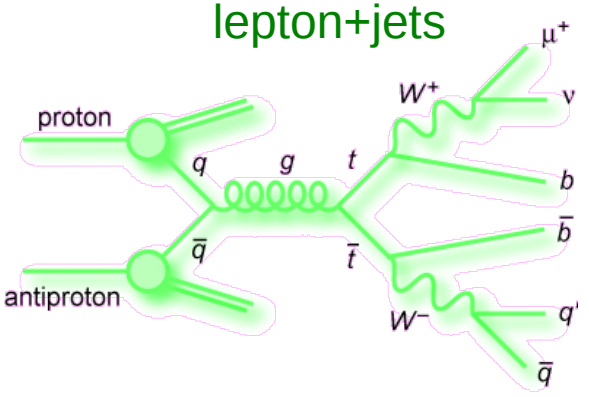
- Need high experimental precision to 'compete' with theory uncertainties

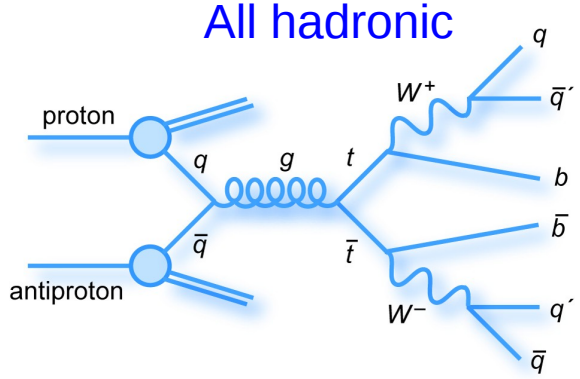
Top quark decay

- Different decay channels:





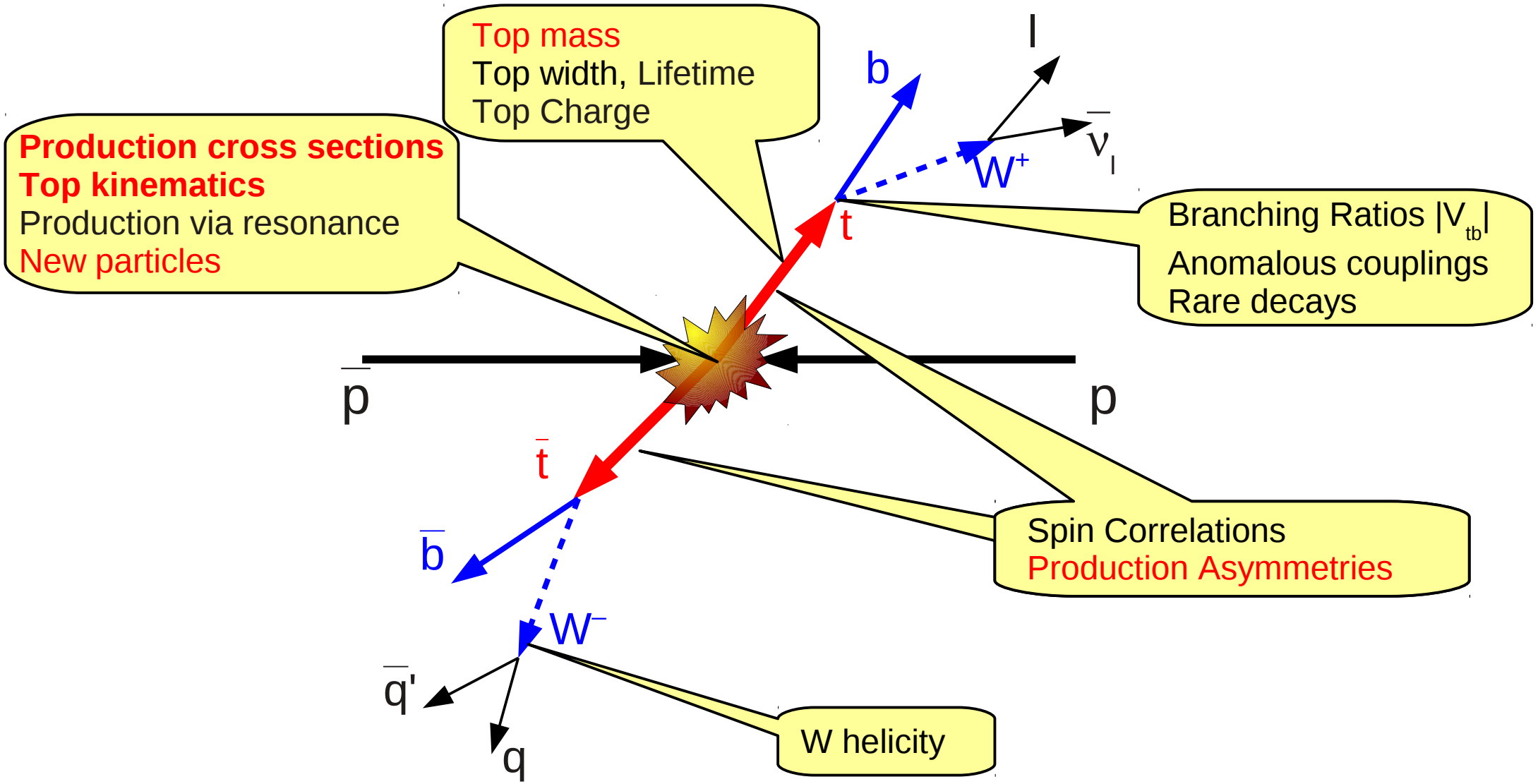




BR, bg decrease

BR, bg increase

A rich & wide field

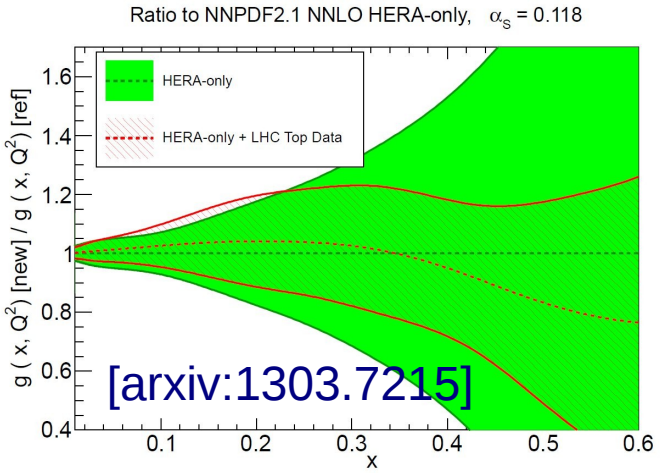


→ and can have all kinds of associated $t\bar{t}$ production: $t\bar{t}$ + H/Z/W/jets/$t\bar{t}$ or new particles

→ Recent review of top quark measurements at Tevatron: Publ. in MPLA (arxiv:1404.1842)

Why cross sections ?

- Inclusive and differential cross section distributions **test QCD calculations** and **enhance our understanding** of top quark production
  - Measurements are systematically limited

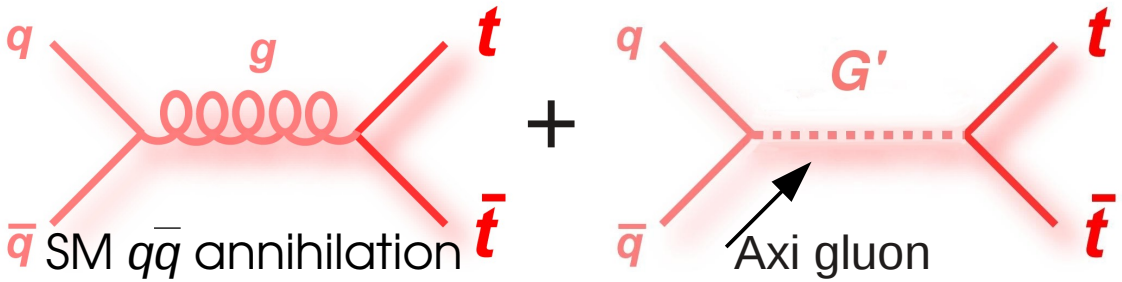


- **Constrain new physics models**, also new physics could 'contaminate' measurements

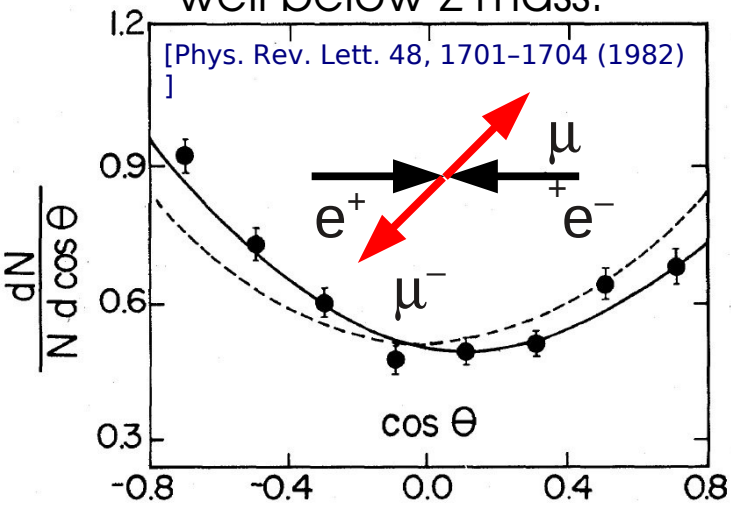
- Example: Angular distributions

- **Large Asymmetry due to:**

→ Axi gluons or *Z'*

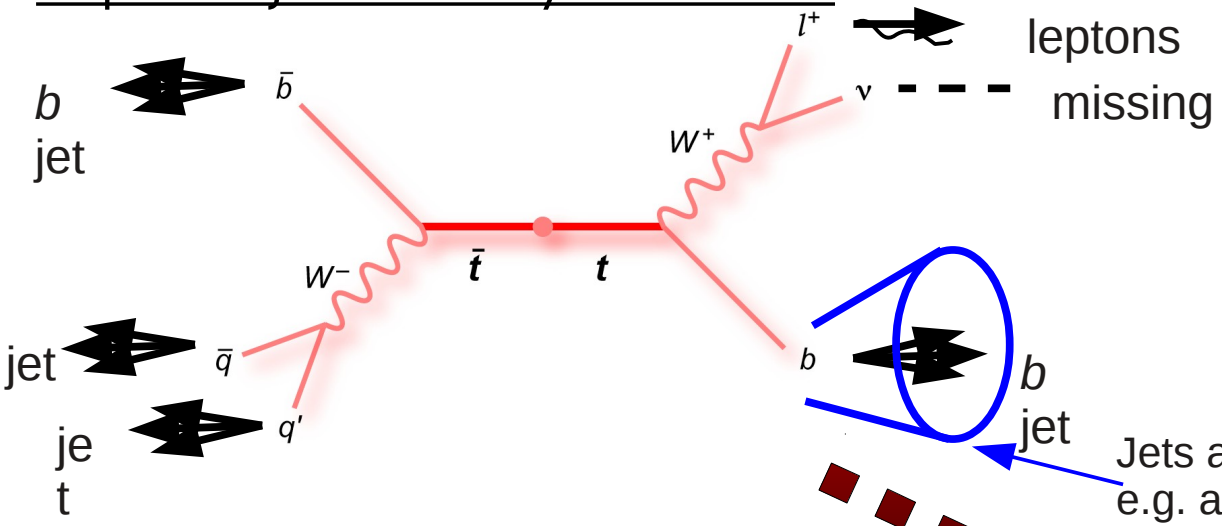


MARK-J at √s = 34.6 GeV, well below *Z* mass:



What to measure ?

- Lepton+jets decay channel:

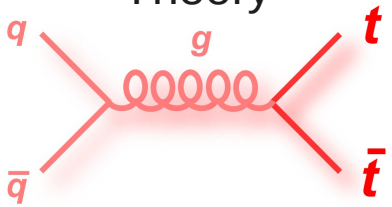


variable	kinematic range
lepton $\eta(\ell)$	$\lvert\eta(e)\rvert < 1.1$ and $\lvert\eta(\mu)\rvert < 2.0$
lepton $p_T(\ell)$	$p_T(\ell) > 20$ GeV
$\not\!\!E_T$	$\not\!\!E_T > 20$ GeV
jet $\eta(jet)$	$\lvert\eta(jet)\rvert < 2.5$
jet $p_T(jet)$	$p_T(jet) > 20$ GeV

Jets are defined by jet finding Algorithms, e.g. anti-k_T, cone, Cambridge-Aachen

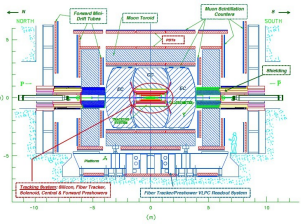
Measurement **"Data"**

Theory

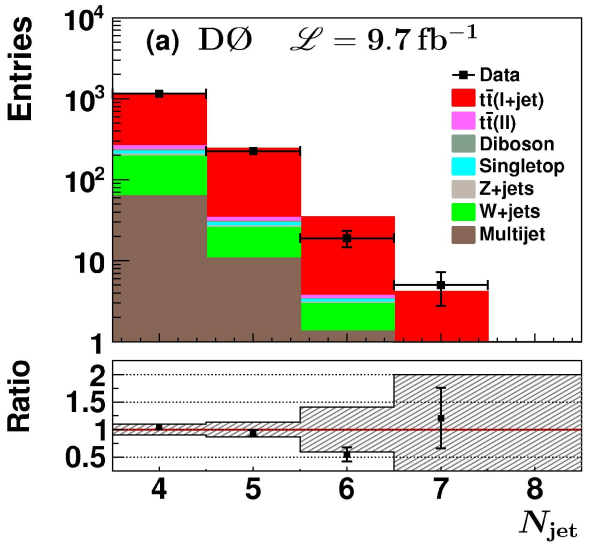


SM $q\bar{q}$ annihilation

Detector simulation



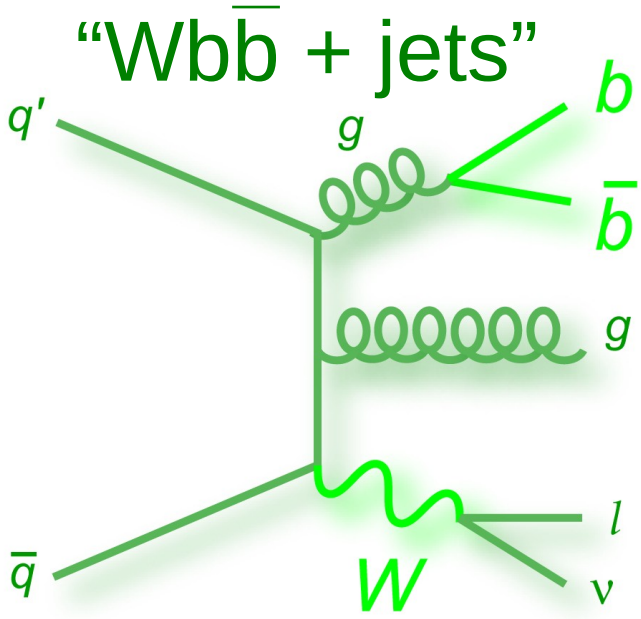
Prediction/ Calculation **"MC"**

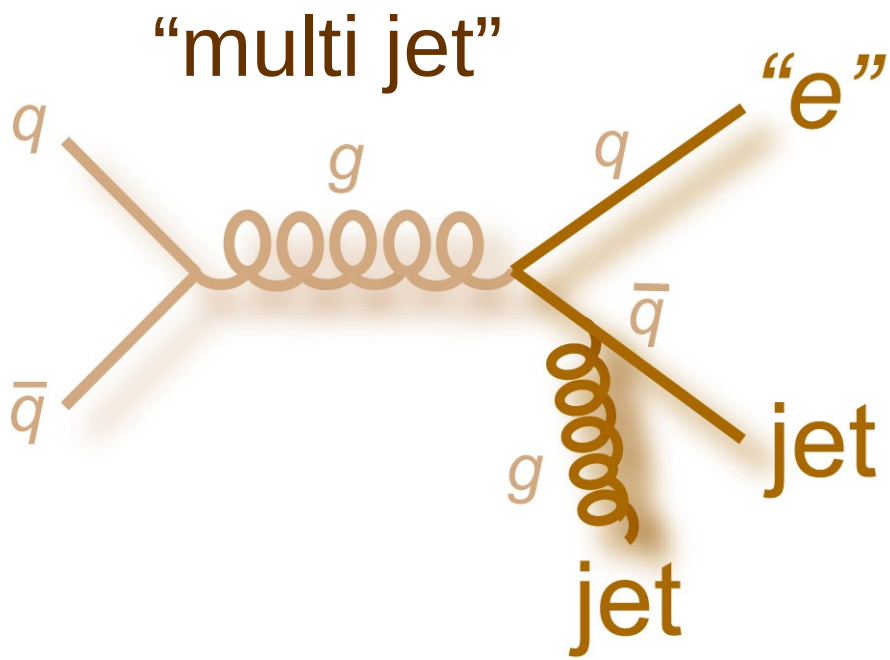


→ **Always two ways: Specific model compared to data or data corrected back to "theory-level" (aka "parton level"),** fiducial volume to full PS

Dominant backgrounds

- Background from other physics processes and instrumental sources, e.g.

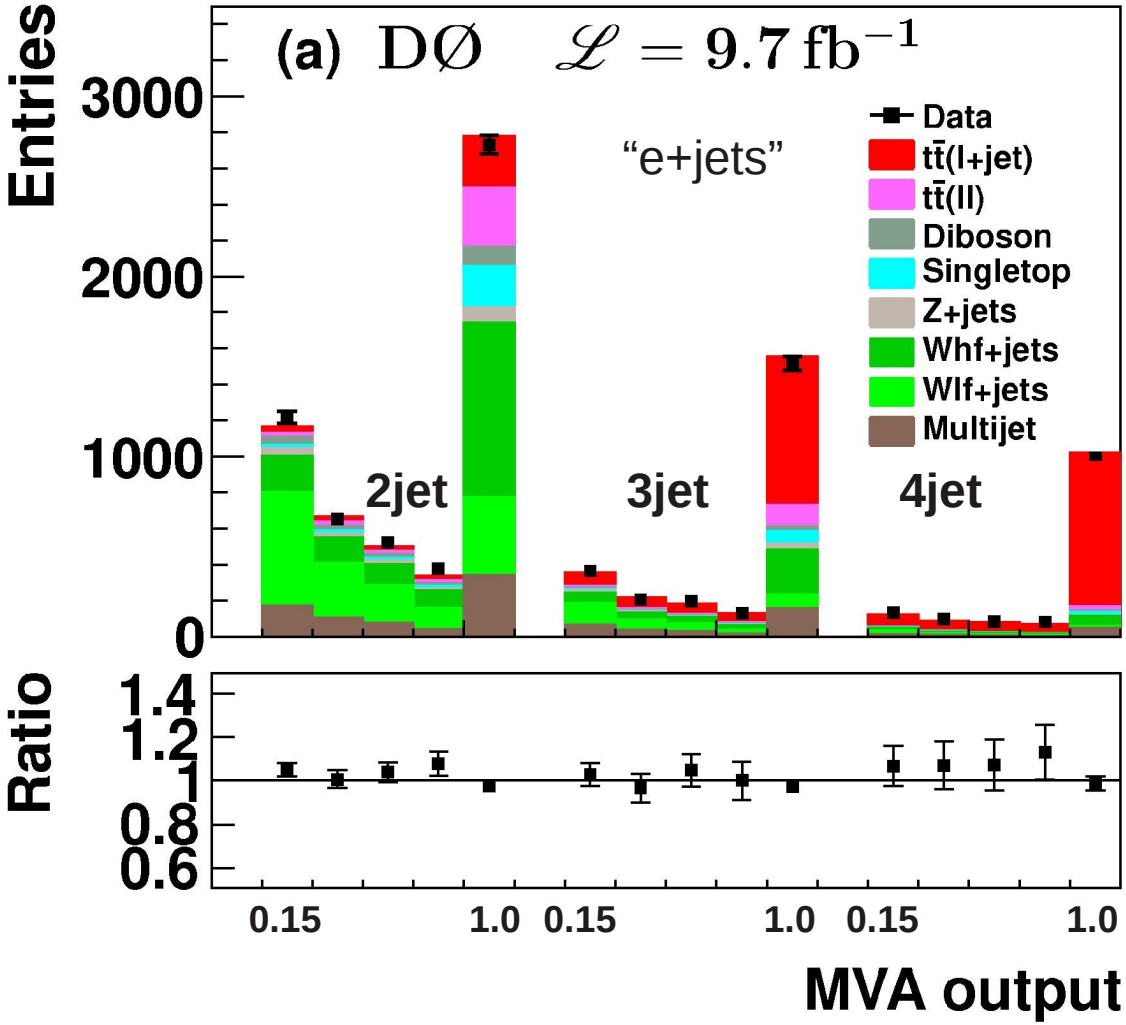


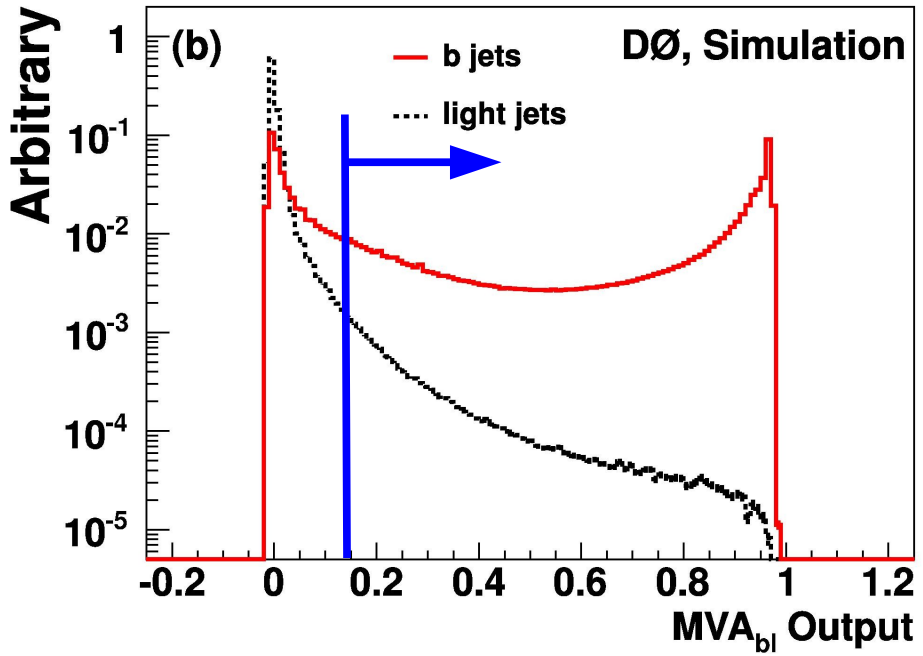


- Multi jet and W+heavy flavor (Whf) contribution derived from data
- Samples for signal and backgrounds:
  - $t\bar{t}$: Alpgen+Pythia, alternative also: Alpgen+Herwig, MC@NLO+Herwig
  - W/Z+jets & W/Zhf+jets: Alpgen+Pythia
  - Diboson: Pythia
  - Single top: CompHep

Enrich signal, reduce bg

- W + heavy flavor + jets (***W*+hf+jets**) contributions constrained by using the 2, 3, and ≥ 4 jet-bin output distributions of the multi-variate *b*-ID technique
- Apply 0.15 ("medium") cut on output values
- Simultaneous fit of ***W*+hf+jets** & $t\bar{t}$ contribution:

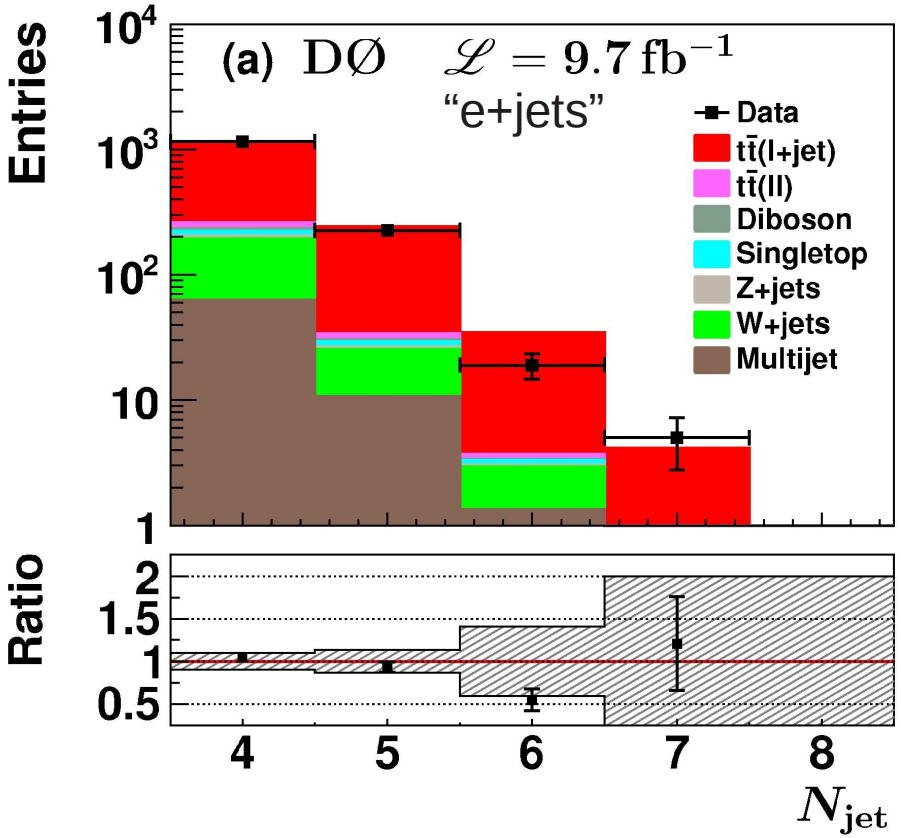


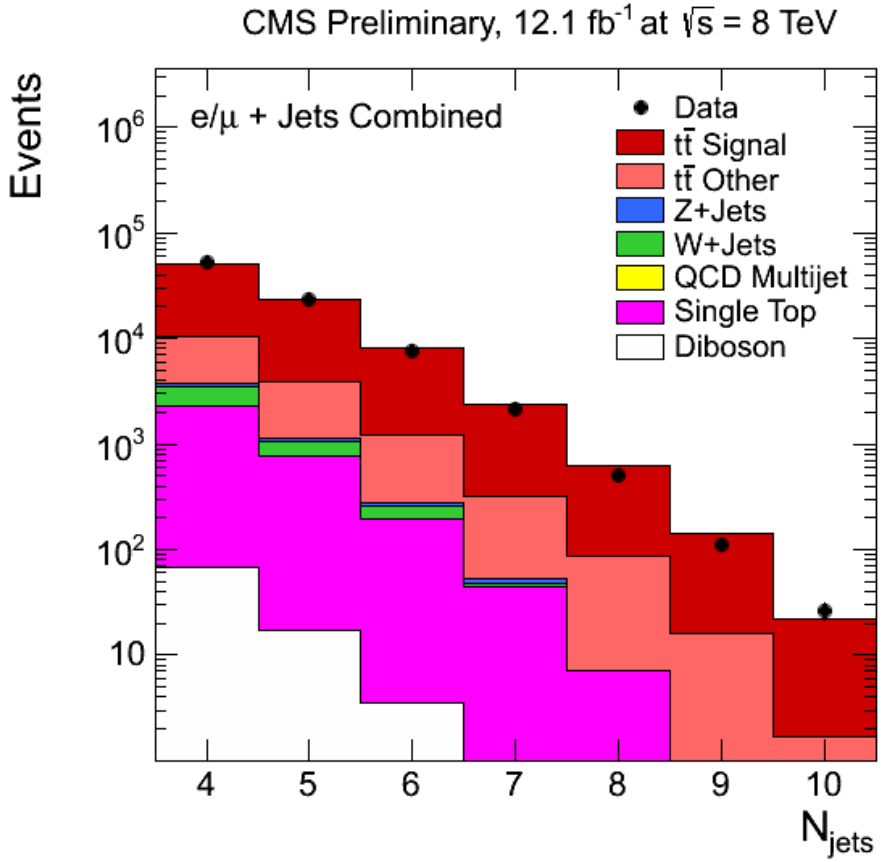


- At LHC: more complex techniques, like **top-tagging**, using **jet sub-structure** to identify tops

Control distributions

- Now check the signal region to measure differential cross sections: *e* or μ + ≥ 4 jets

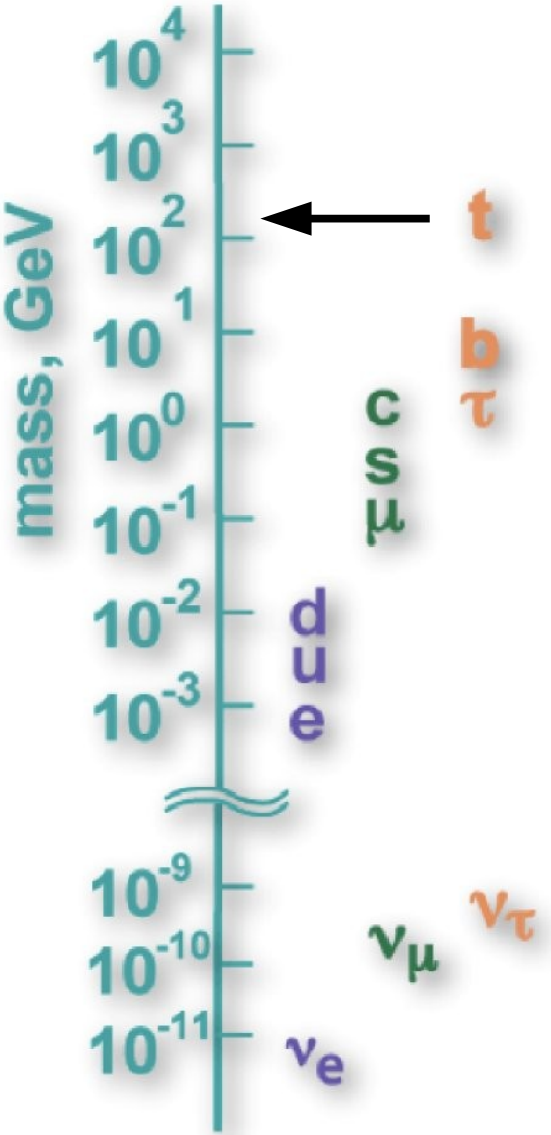




- Very good description/modeling of the data within systematic uncertainties
  (assuming the measured $t\bar{t}$ cross section)
- Higher energy at the LHC → results in higher number of jets and "boosted" top quarks (events at high scales, such as p_T, m($t\bar{t}$)) → new analysis techniques

Correct data to what ?

- Most precise predictions at approx. NNLO:
  → Require to correct to parton level
- Heaviest elementary particle, Top quark decays before it can hadronize:
  No fragmentation, unlike for *b* or *c* quarks
  → Effects from QCD and hadronization are reduced

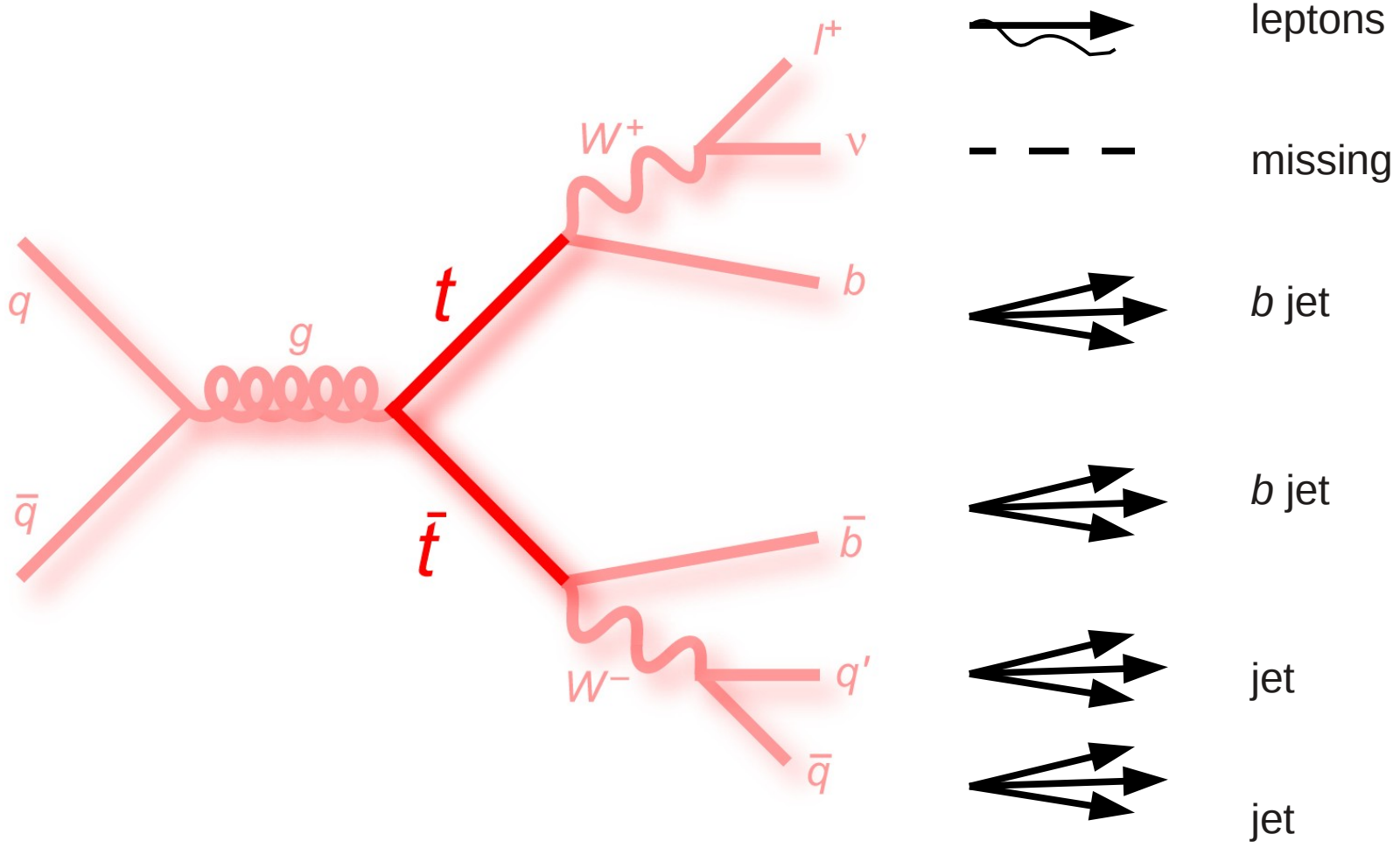


- At Tevatron $t\bar{t}$ pair produced almost at rest:
  → top quark decays almost always in detector acceptance
  → In other words: **Nature helps a lot...**
- At LHC fraction of tops can be boosted due to the higher available initial energy
  → special techniques: event structure / jet substructure

Top quark reconstruction

Associate leptons and jets with individual top quark decay particles:

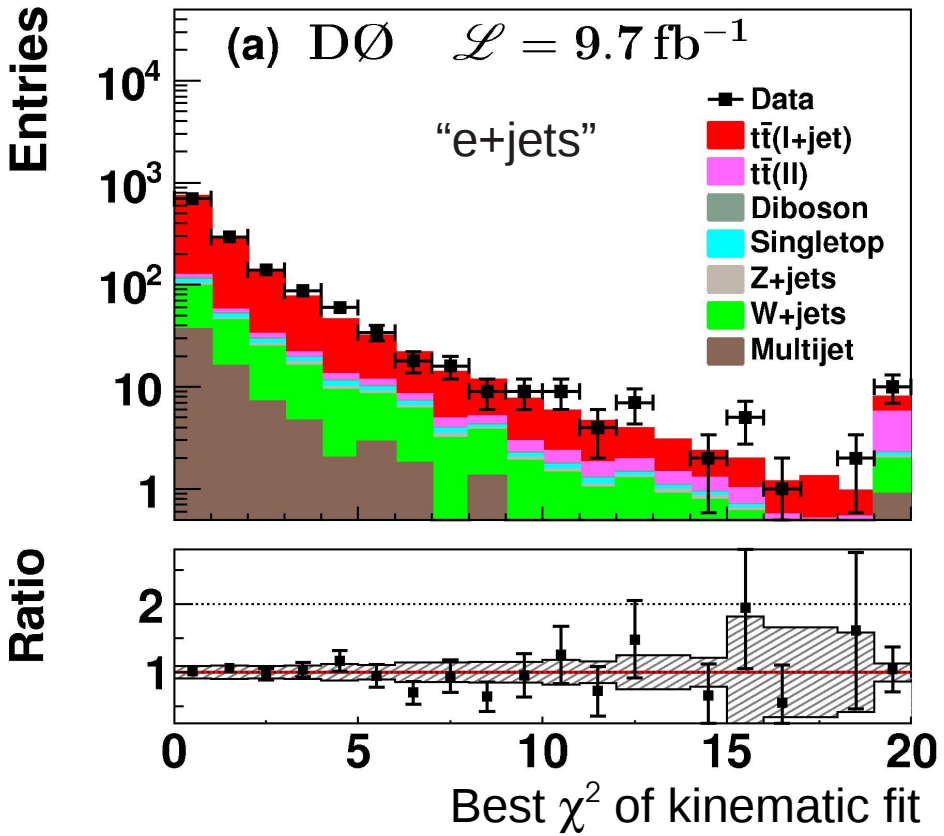
- 12 possible jet to quark assignments, reduced with *b* tag information
- Additional 2 solution for longitudinal momentum of neutrino
  → χ^2 based kinematic reconstruction

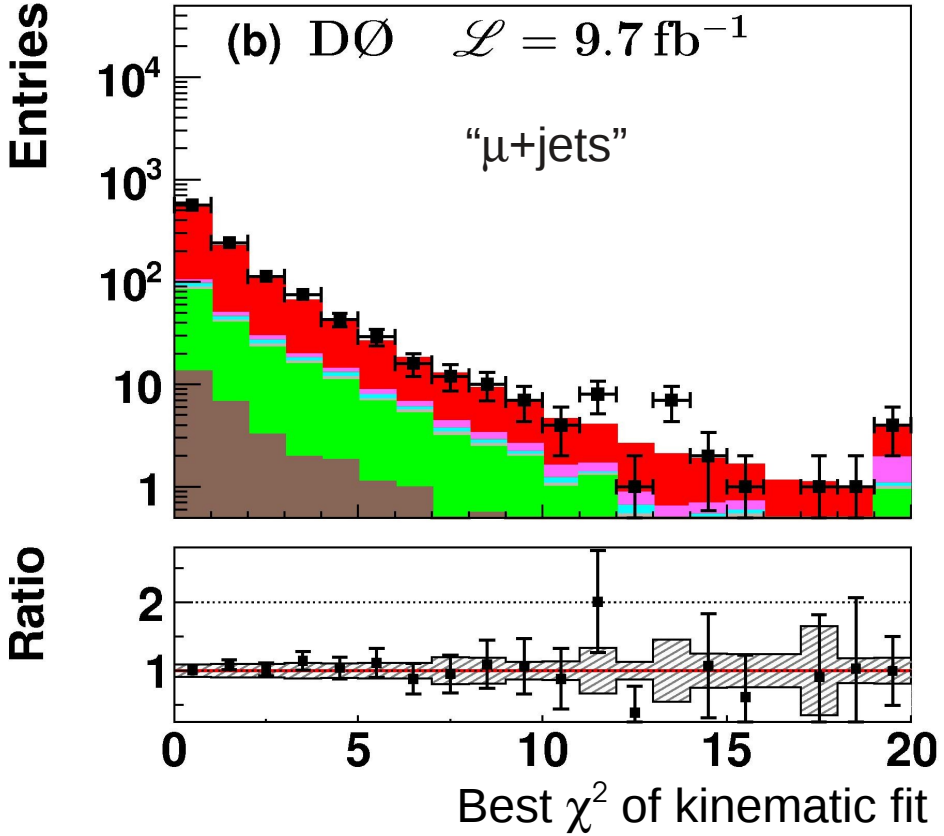


Top quark reconstruction

Associate leptons and jets with individual top quark decay particles:

- 12 possible jet to quark assignments, reduced with *b* tag information
- Additional 2 solution for longitudinal momentum of neutrino
  → χ^2 based kinematic reconstruction
    - Constrain mass of W and equal top masses
    - Best solution is input to the unfolding (background is subtracted)





Regularized Unfolding

- "Bin-by-bin" correction method:
  → Possibly strong model dependencies

$$\frac{d\sigma}{dX_i} = \frac{N_i^{\text{signal}}}{\epsilon \cdot \mathcal{L} \cdot \mathcal{BR} \cdot \Delta X_i}$$

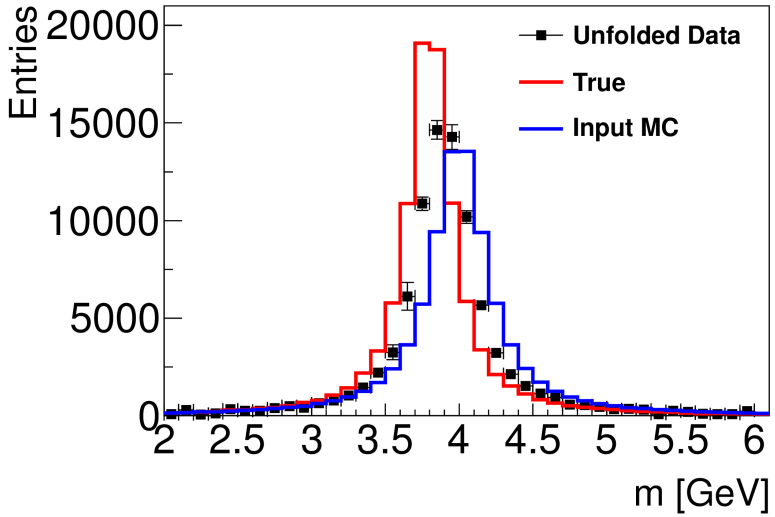
- Matrix Unfolding introduced:
  → Correct treatment of migrations
  → Reduced bias from choice of model
- For a discrete measurement with *n* bins:

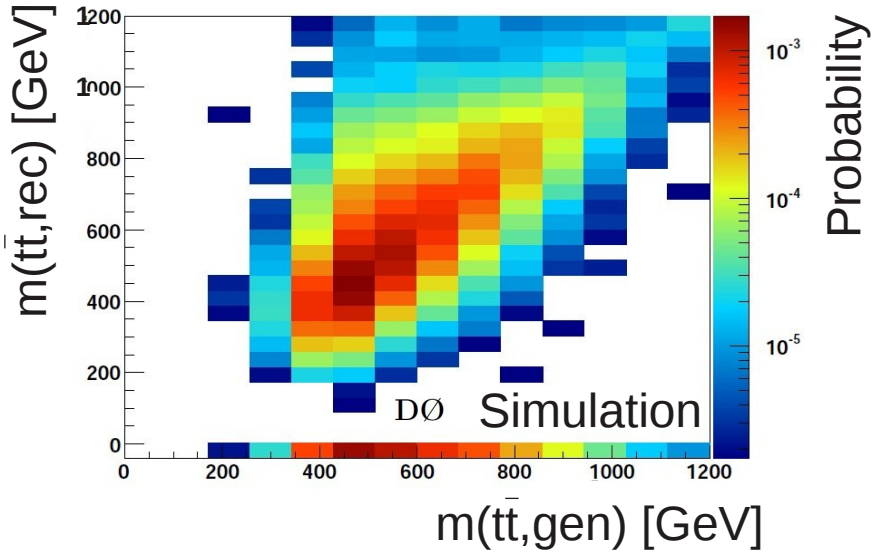
$$\boldsymbol{Ax = y \;\rightarrow\; x = A^{-1}y}$$

x = *n*-histogram of true variable *x*
y = *m*-histogram of measured variable *y*
A = *m* × *n* response matrix

- Need **regularization, can cause bias:**

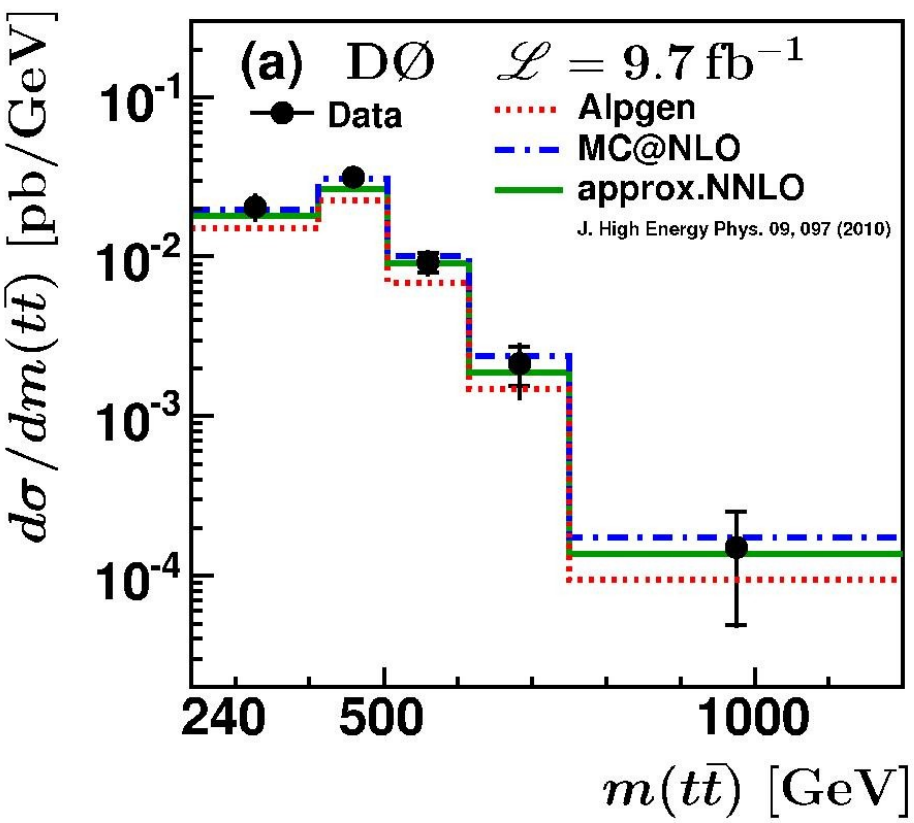
$$\chi^2 = \sum_i \frac{\left(y_i^{\text{data}} - \sum_j A_{ij} \cdot x_j^{\text{unfold}}\right)^2}{\left(\delta_i^{\text{data}}\right)^2} + \tau^2 \times L_{ij}(L_{ij})^T$$

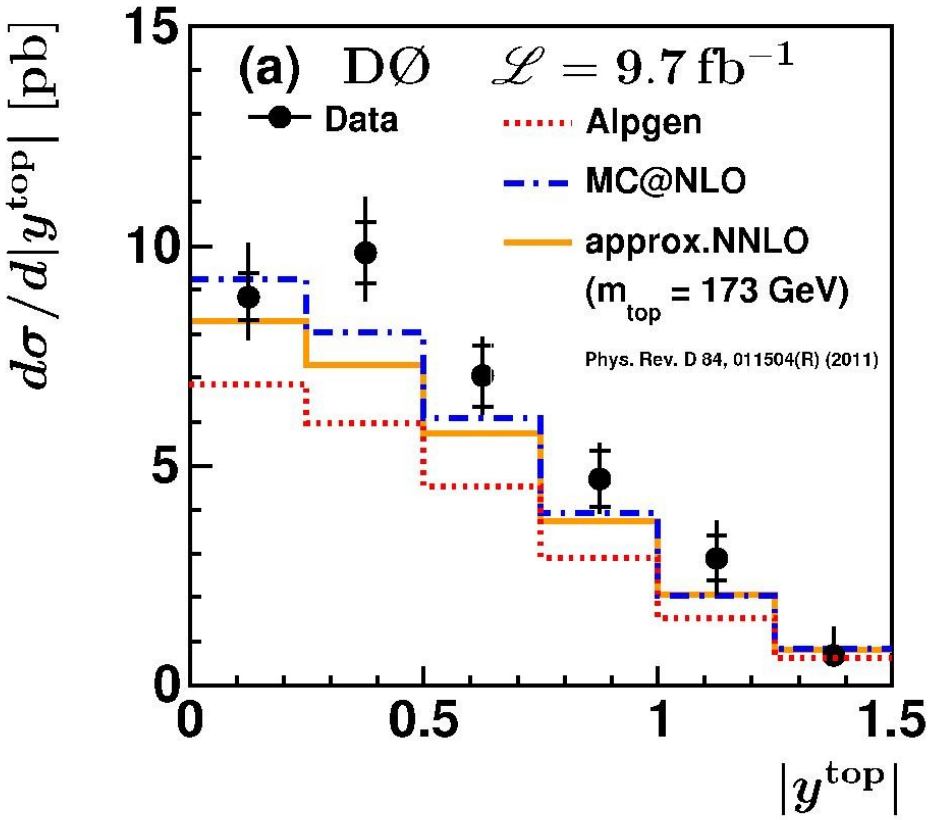




Toy model example: → Unfold smeared distribution to "true" using MC

Differential cross sections

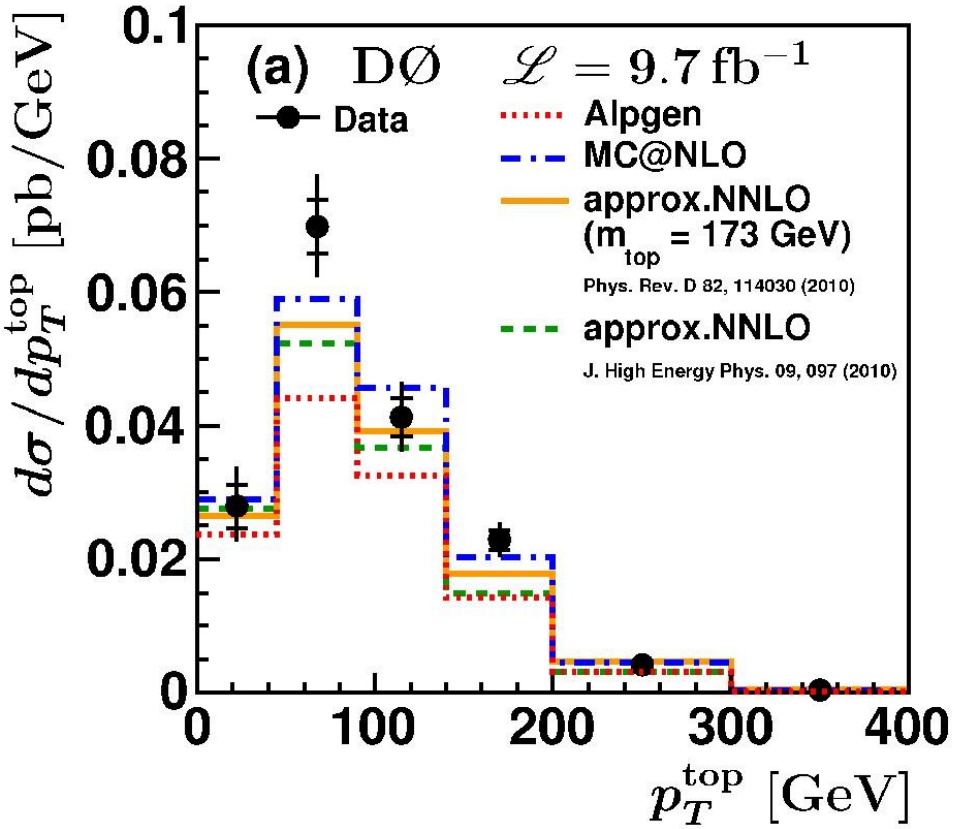


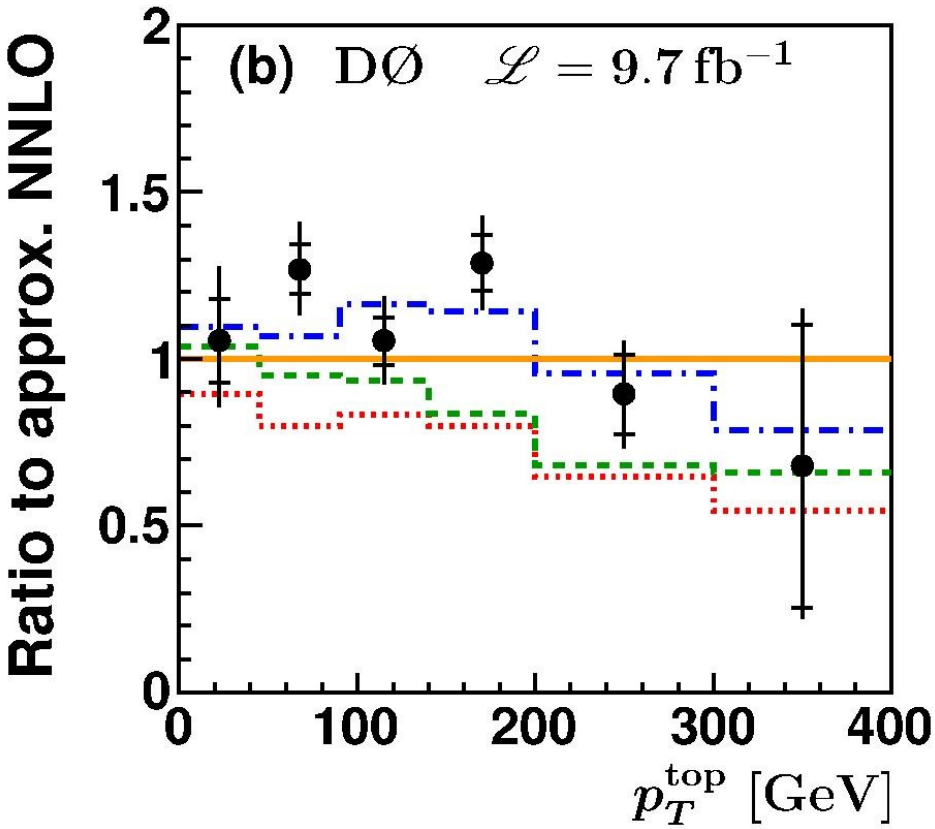


- Differential cross sections modeled by MC@NLO & approx. NNLO
- Systematically limited (dominant one is signal modeling)
- High tail sensitive to new physics (not the full information: Correlations)
  → no indication of new physics till 1.2 TeV

- LHC: no new physics till ~ 1.6 GeV
- Every measurements improves understanding of QCD and top quark kinematics

Differential cross sections

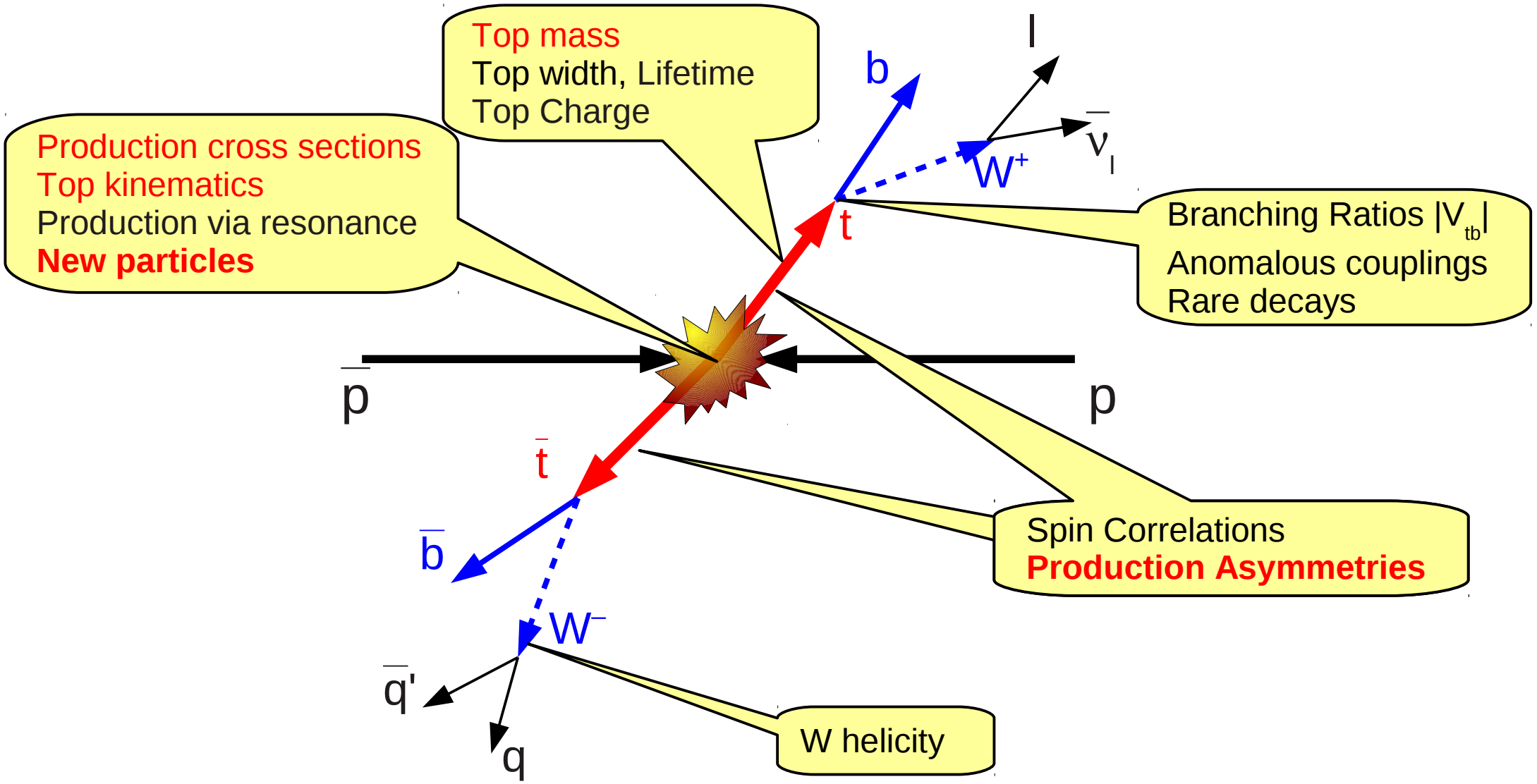




- Differential cross sections modeled by MC@NLO & approx. NNLO
- Well described distribution

- LHC (CMS): difficulties in describing the low p_T region
- Every measurements increases understanding of QCD and top quark kinematics

A rich & wide field

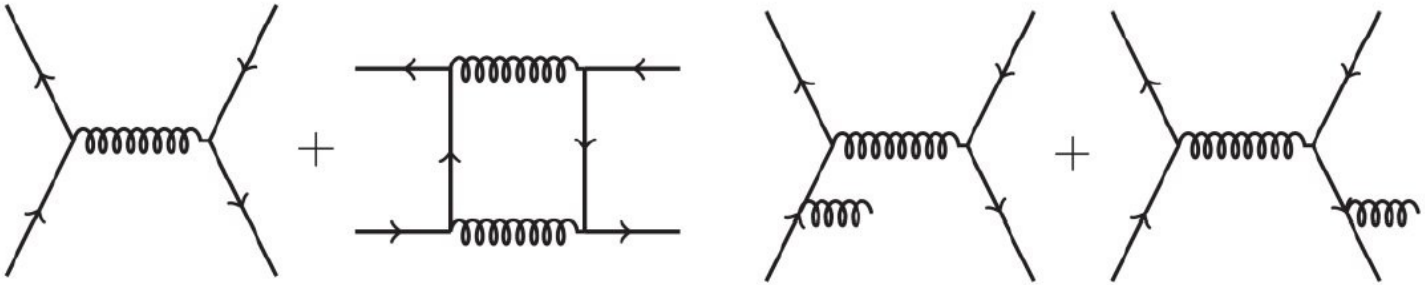


→ and can have all kinds of associated $t\bar{t}$ production: $t\bar{t}$ + H/Z/W/jets/$t\bar{t}$ or new particles

→ Recent review of top quark measurements at Tevatron: Publ. in MPLA (arxiv:1404.1842)

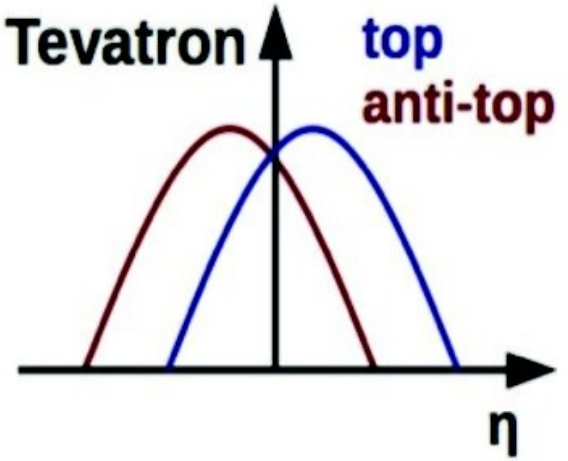
Top quark asymmetries

- Interference appears at NLO QCD:



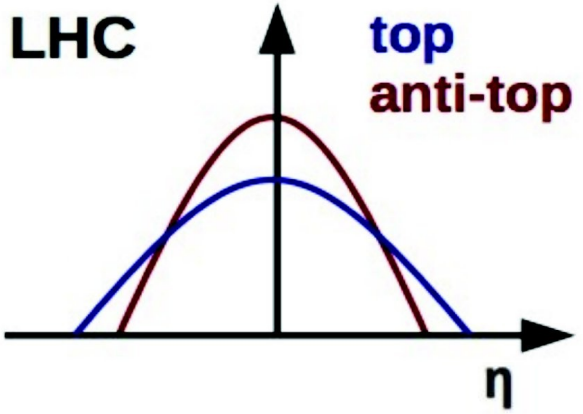
Positive asymmetry Negative asymmetry

→ Only occurs in $q\bar{q}$ initial state; gg is fwd-bwd symmetric



- This is a forward-backward asymmetry at Tevatron
- No valence anti-quarks at LHC → tops less central

$$A_{FB} = \frac{N(\Delta y > 0) - N(\Delta y < 0)}{N(\Delta y > 0) + N(\Delta y < 0)}$$



- SM predictions at NLO (QCD+EWK): **A_{FB} = 8.8 %**
  [Bernreuther & Si]

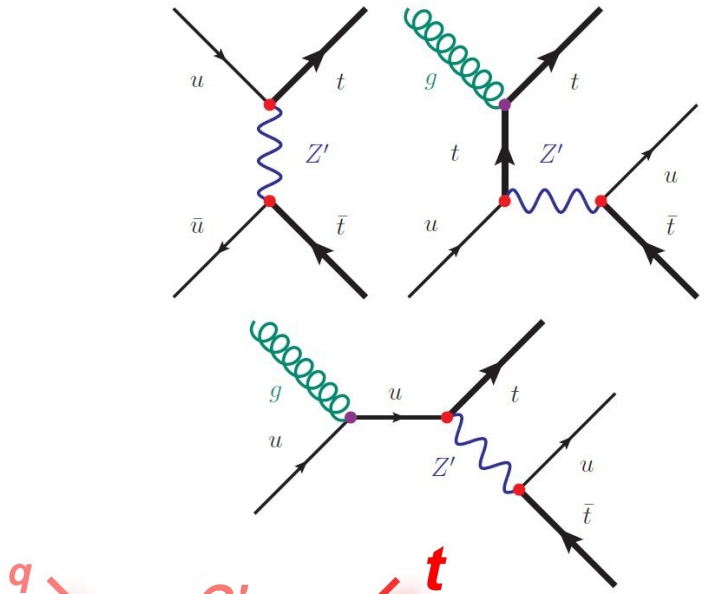
A_{FB} ~ 10 % (full NNLO pQCD predictions)
[Czakon TOP14]
and **A_{FB} = 12.5 %** (PMC scale setting)
[arxiv:1205.1232]

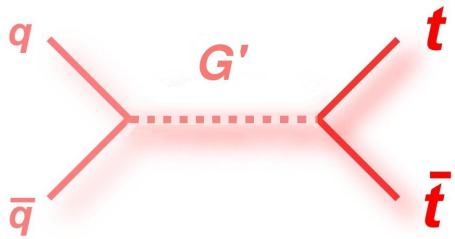
A_C ~ 1 % (waiting for full NNLO pQCD predictions)

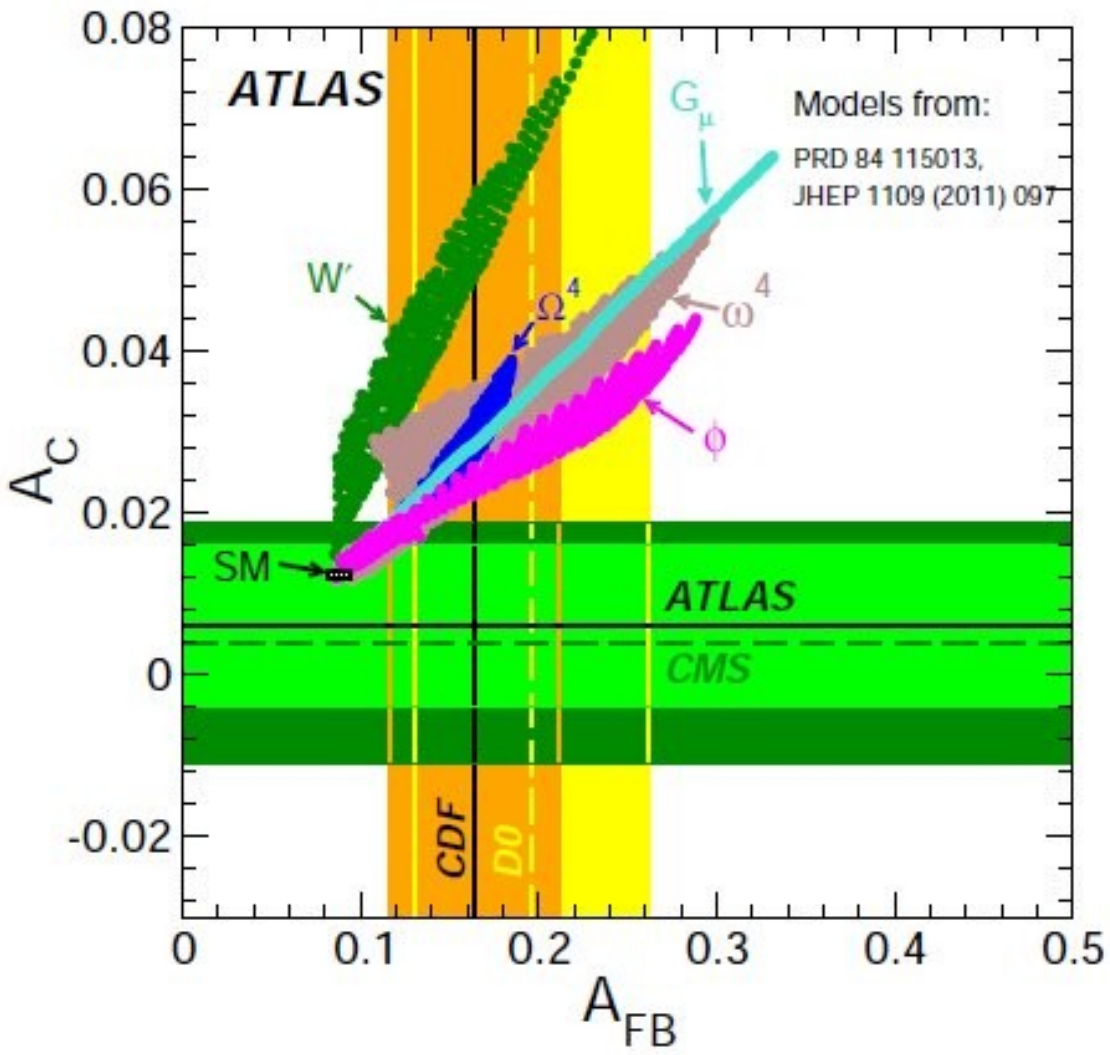
$$A_C = \frac{N(\Delta|y_t| > 0) - N(\Delta|y_t| < 0)}{N(\Delta|y_t| > 0) + N(\Delta|y_t| < 0)}$$

Top quark asymmetries

- Measurements at Tevatron & LHC are complimentary
  - Variety of models with wide parameter space still allowed
    → *W′, G, ω, φ, Ω*







Snowmass white paper on Kinematics of top quarks at upgraded LHC: [arxiv:1309.2889]

Top quark asymmetries

- Lots of effort went into maximizing expected significance:
  - → Understand & Improve systematic uncertainties
  - → Increase available sample size
    - Full Run II expected statistical uncertainty 4.6%, need 3% for 5 s.d. significance (given central value does not move) → Can we get to 3% ?

New kinematic fit:

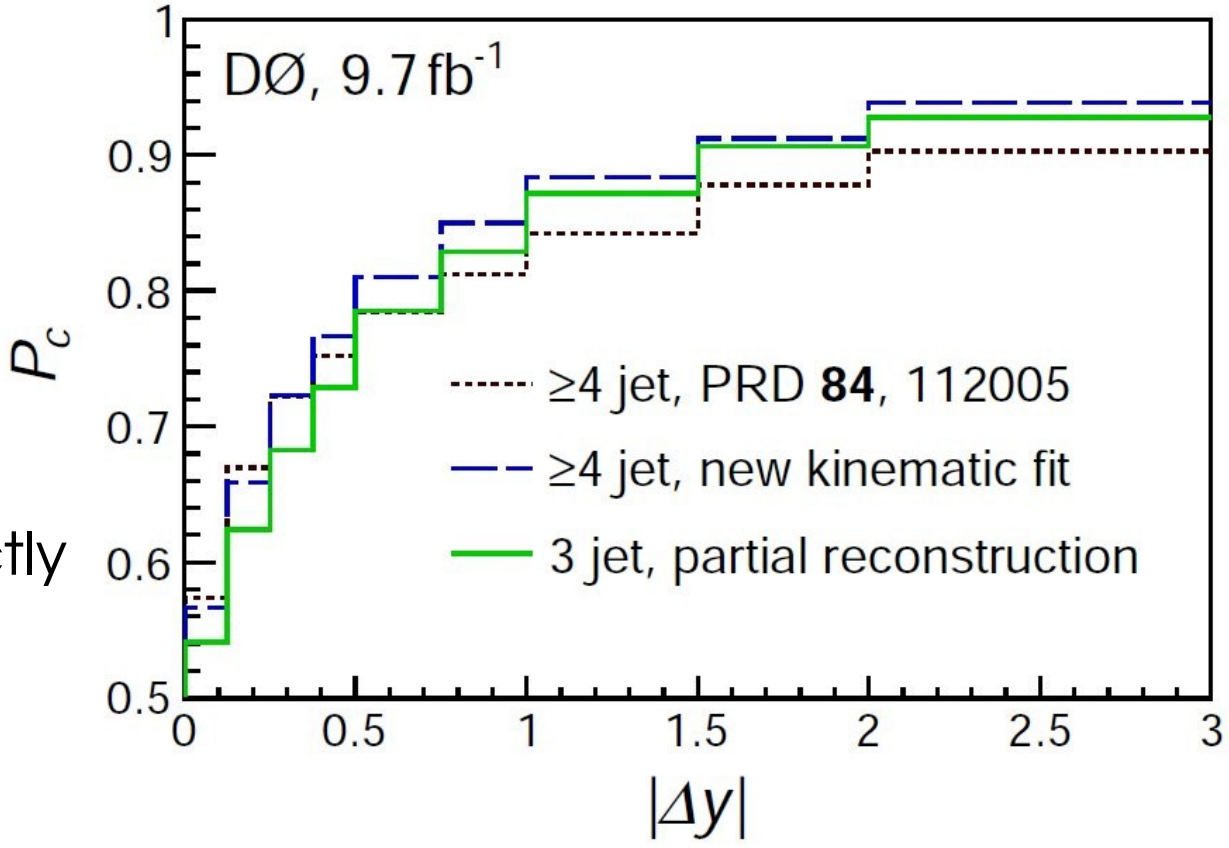
- Flavor dependent transfer fcts
- Analytical solution for p(ν)

  (Betchart, Demina, Harel NIM A 736 (2014) 169)

Partial $t\bar{t}$ reconstruction:

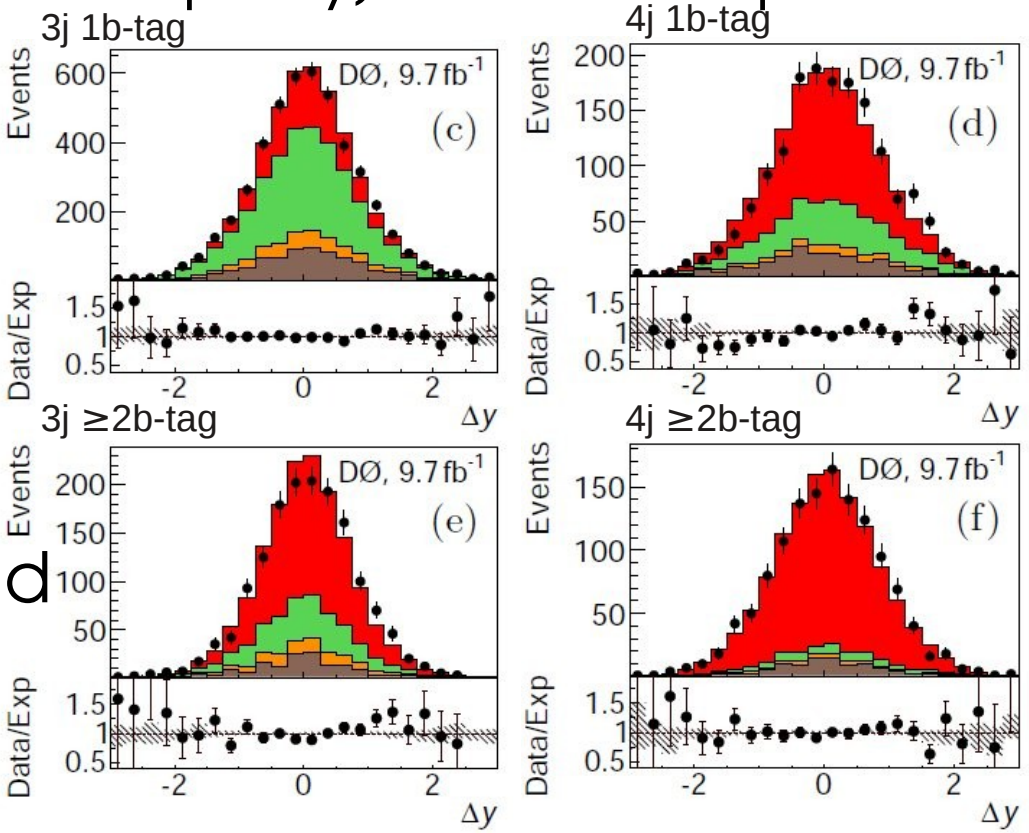
- Lost jet is in 80% from hadronic top
- Construct a likelihood & correctly reconstruct sign of Δy:

→ P_C (≥ 4j) = 77.6 %
P_C (= 3j) = 74.5 %



Top quark asymmetries

- Shares same selection as the one for the leptonic asymmetry measurement, also employs Discriminant D_c
- Combine several channels (w different purity) & unfold to parton level [arxiv:1405.0421]
  → used modified TUnfold



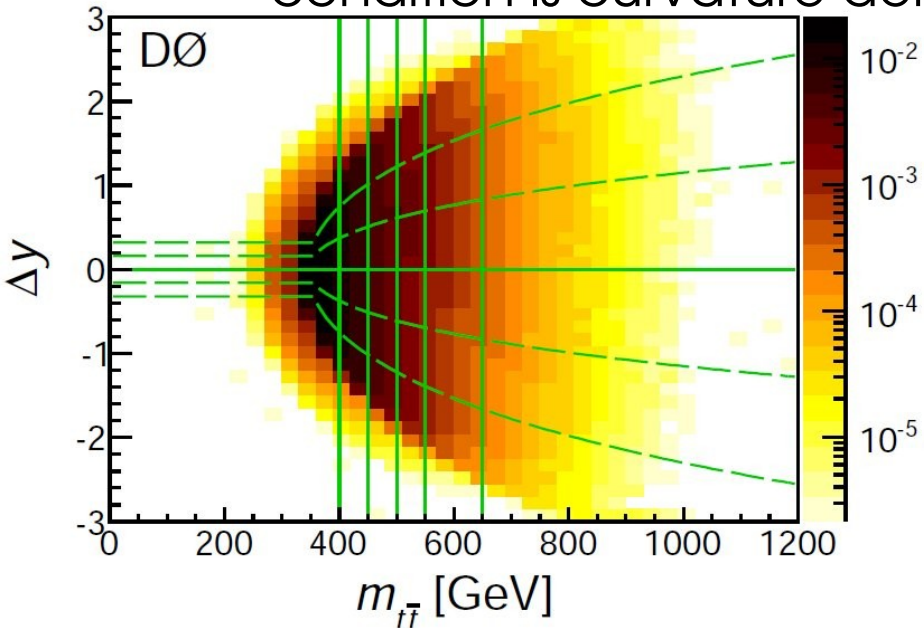
- Achieved improvement beyond luminosity scaling with 3% total uncertainty
- SM @NNLO: ~10% (Czakon, Mitov, et al.)

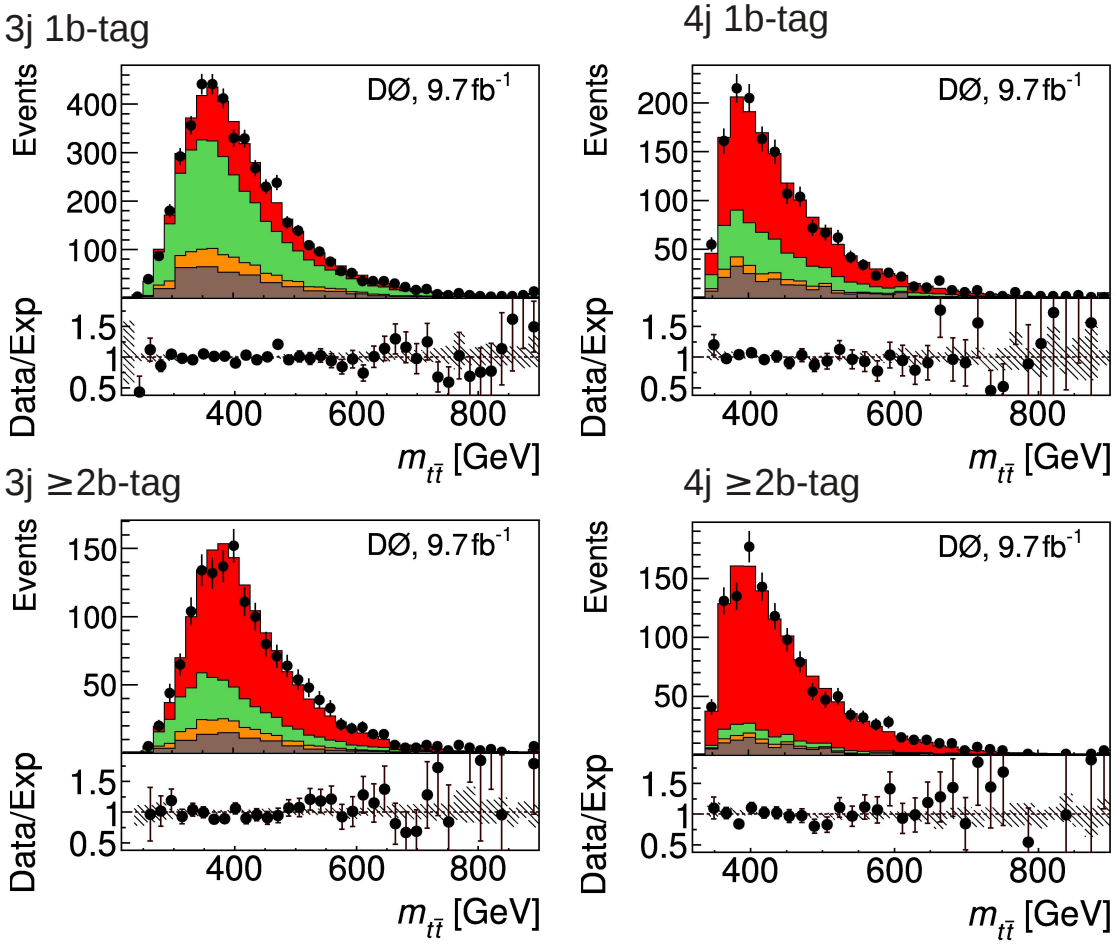
→ A_{FB} = 10.6 ± 3.0 (tot.) % → D0 agrees with SM within uncertainties

CDF: A_{FB} = 16.4 ± 4.5 (tot.) % → CDF higher than SM predictions

Top quark asymmetries

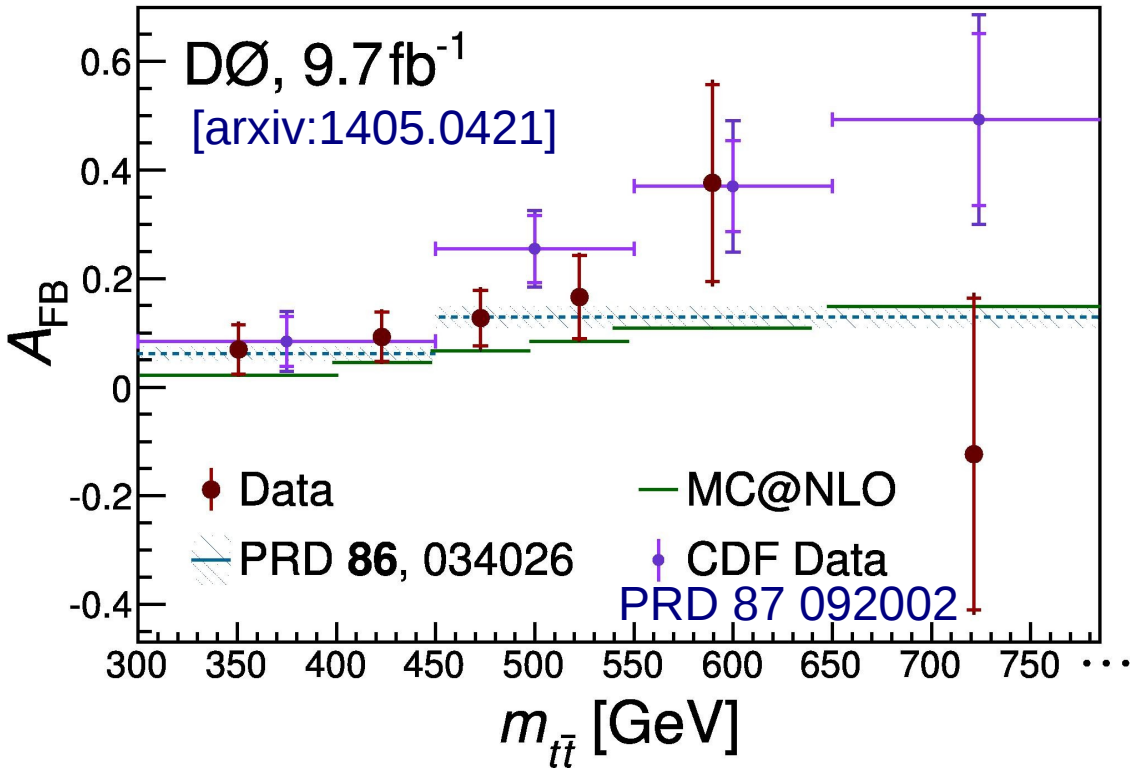
- Shares same selection as the one for the leptonic asymmetry measurement, also employs Discriminant D_C [arxiv:1405.0421]
- Combine several channels (purity!) & unfold to parton level
  → used modified TUnfold
- $m(t\bar{t})$ dependence of A_{FB}:
  - 2D regularized unfolding
    → account for variable bin-size
    → condition is curvature density

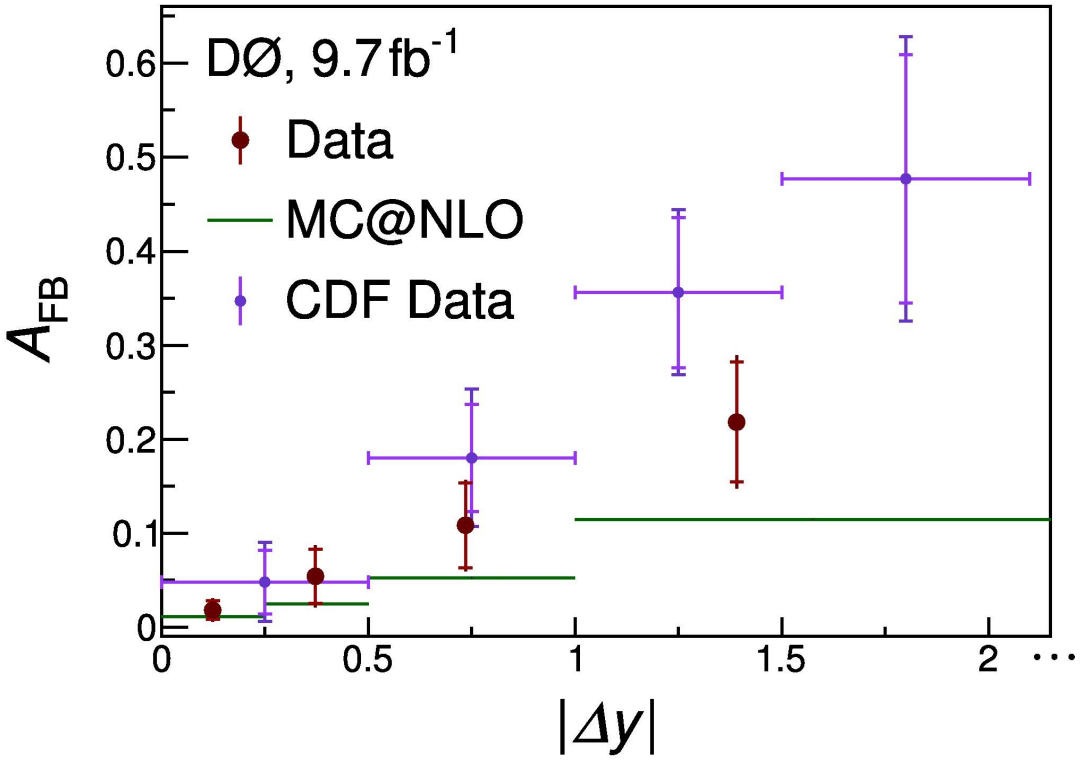




Top quark asymmetries

- Existing CDF measurement in l+jets decay channel also measured kinematic dependence of A_{FB}



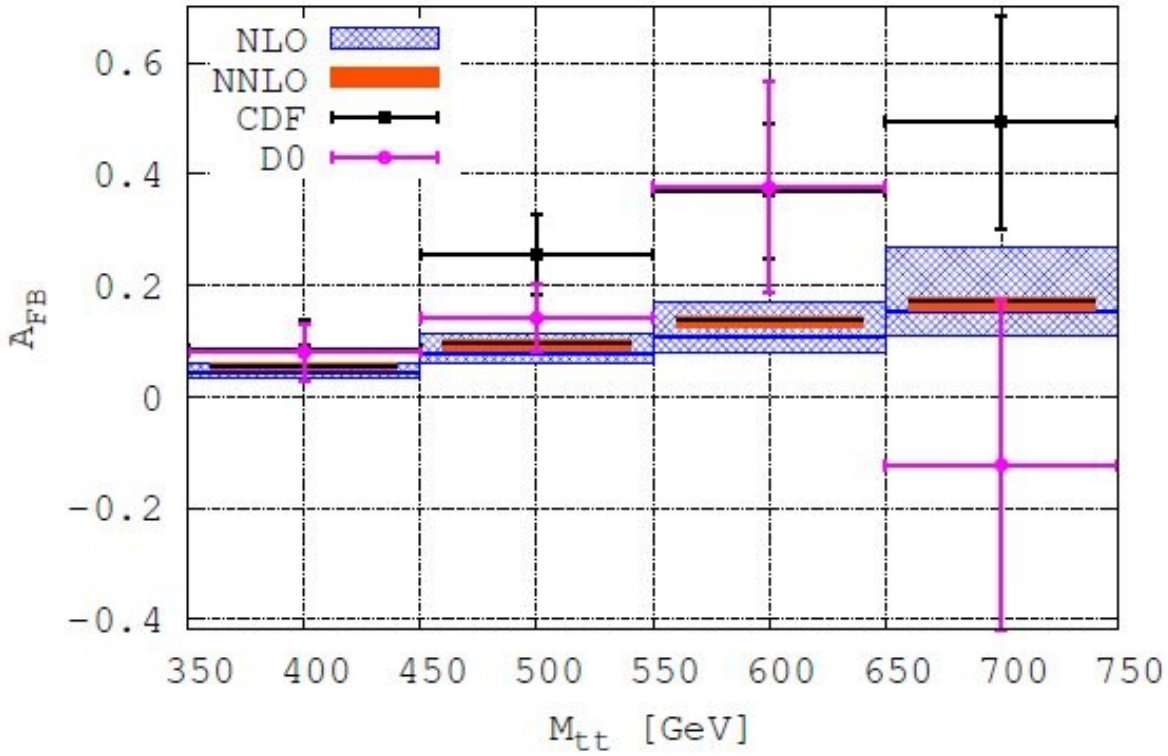


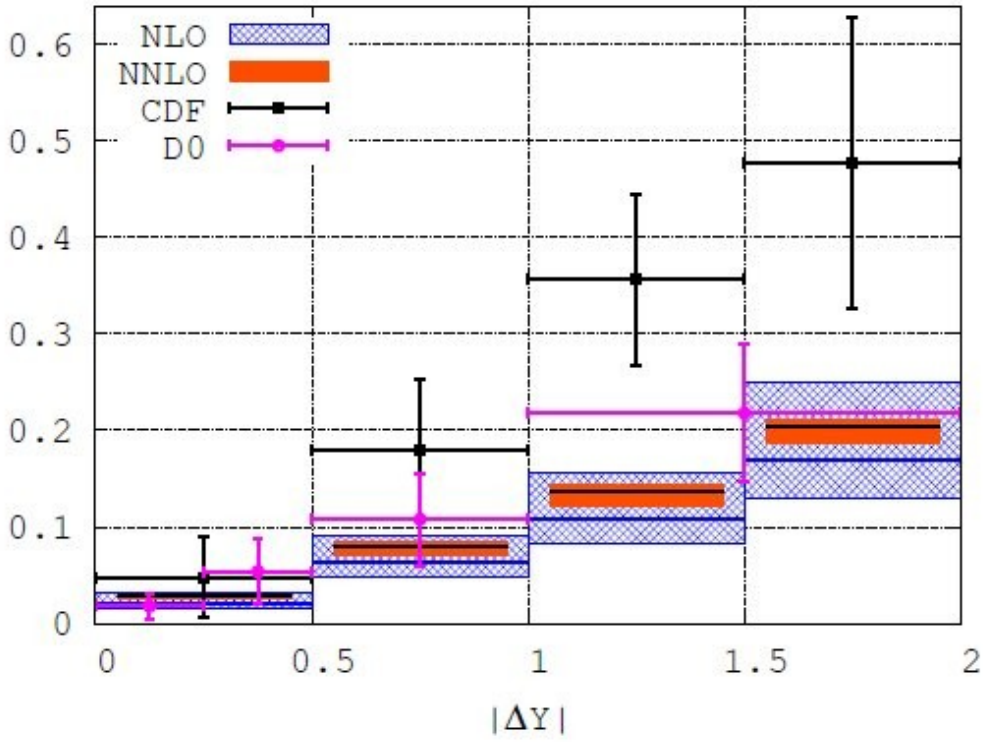
- Kinematic dependencies larger than "currently" predicted by SM
- Good agreement of D0 data with most recent pQCD @NNLO
  (see TOP14 talk by Michael Czakon)

Asymmetries

- Eagerly awaiting the NNLO pQCD predictions: **Arrived !**

Czakon





- Basically data (though on the high side) agrees with new predictions in case of D0
- CDF data slightly higher

Top quark asymmetries

NEW

- Reconstructing ttbar kinematics: challenging in dilepton channel
- Both *W* boson decay to lepton and neutrino
- For a kinematic quantity "z":

$$L_z = \frac{1}{\mathcal{A}\cdot\sigma_{tot}} \sum_{flavors} \int_{x,q_1,q_2,p_T^{t\bar{t}},\phi^{t\bar{t}}} W(x,z)\,W(p_T^{t\bar{t}})\,f_{PDF}(q_1)\,f_{PDF}(q_2)\,d^6\sigma_{hs}(x)\,dp_T^{t\bar{t}}\,d\phi^{t\bar{t}}\,dq_1\,dq_2$$

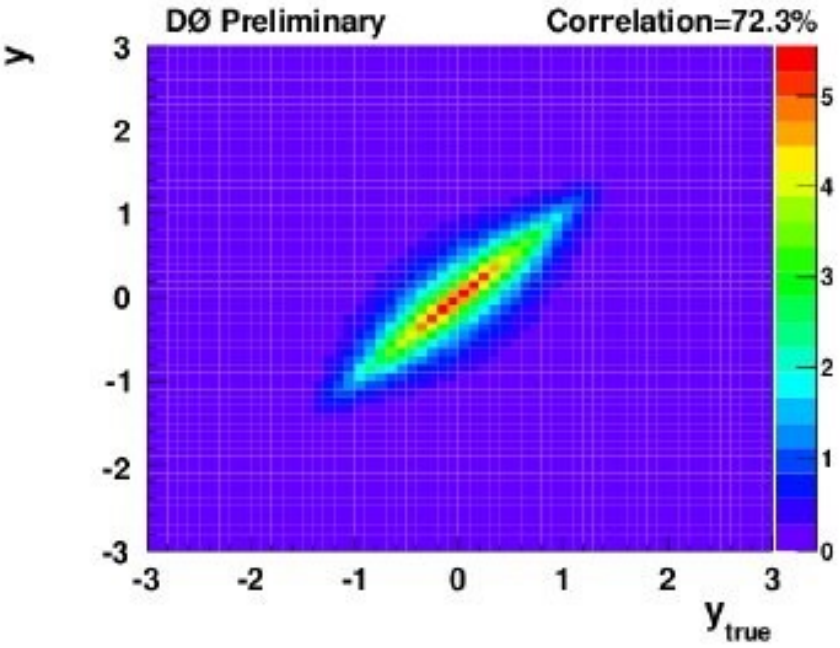
$d^6\sigma(x)$ differential cross-section

$f_{PDF}(q_i)$: parton density function

W(x,z) transfer functions to relate reconstructed quantities z to true 4-vectors x

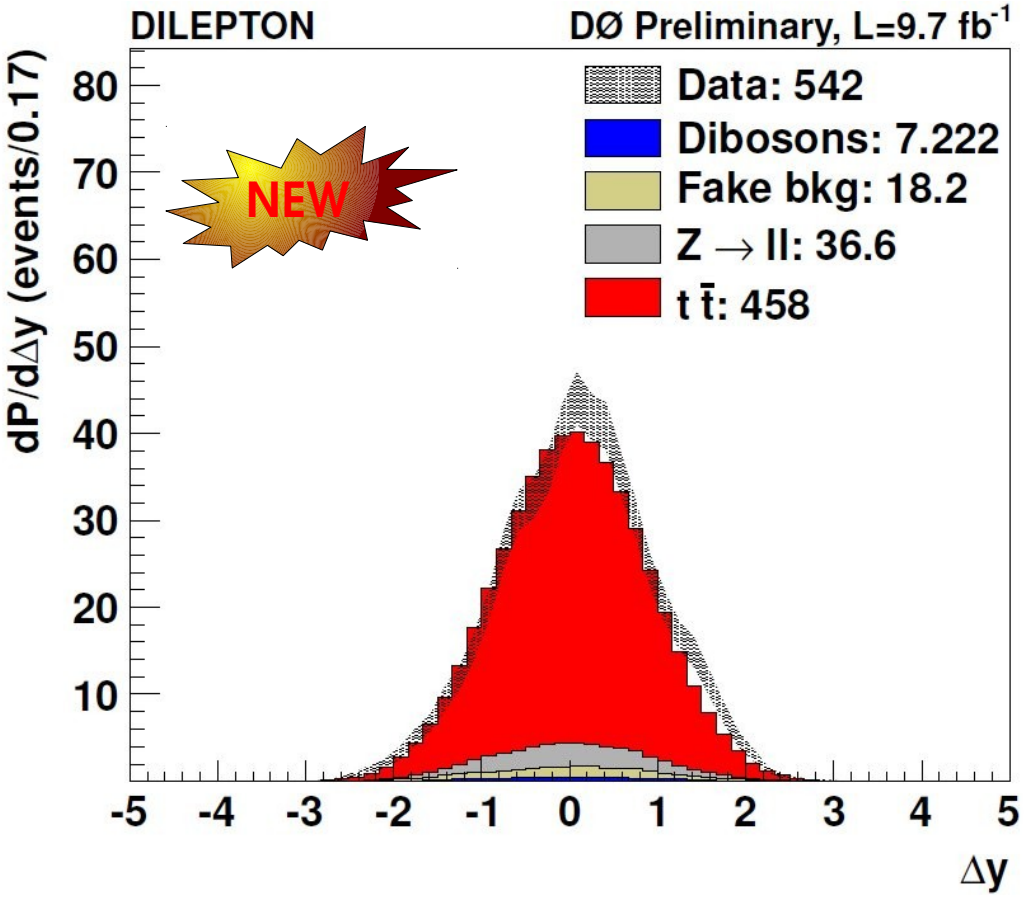
$W(p_T)$: pdf for ttbar transverse momentum

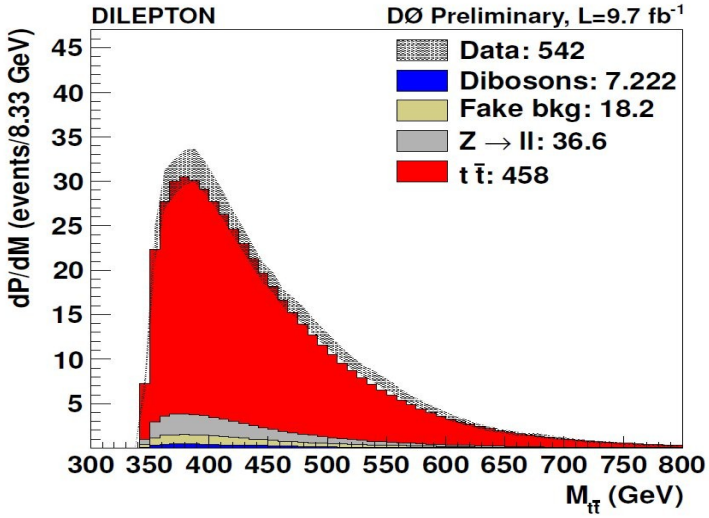
- Good correlation between true and reconstructed quantity

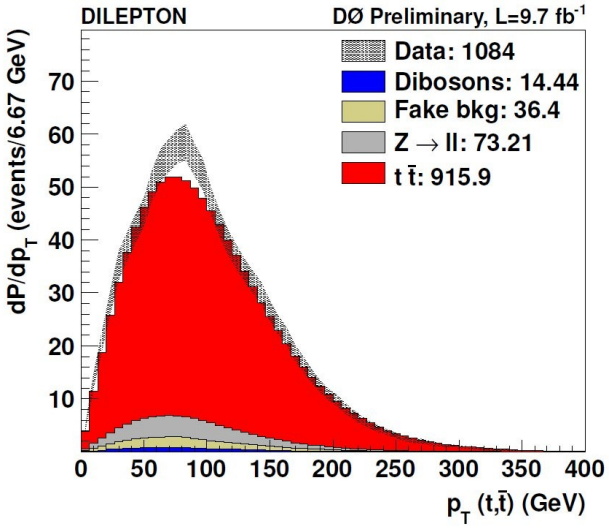


Top quark asymmetries

- New measurement by D0 in the dilepton channel employing the matrix element method:
  → assign a likelihood per event for most probably Δy (x) value





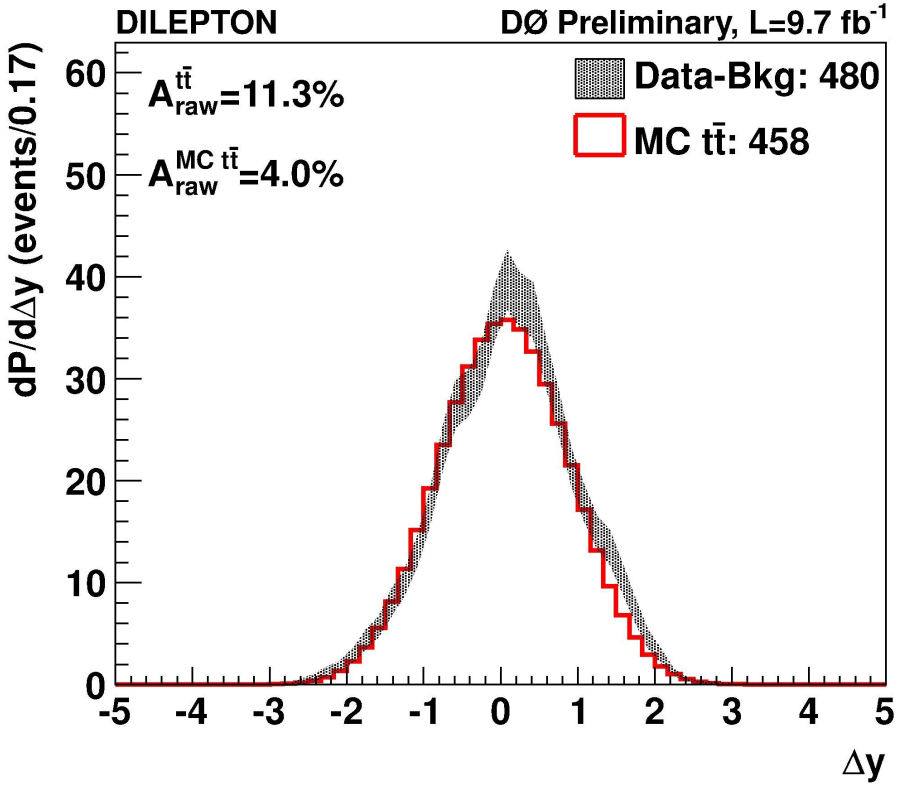


- Control distributions show reasonable data modeling by MC → extract asymmetry

Top quark asymmetries

- Systematic uncertainties dominated by signal modeling, in particular hadronization and showering and calibration

NEW



Source of uncertainty	Uncertainty on $A_{FB}^{t\bar{t}}$ (%)
Detector modeling	
jet energy scale	0.14
jet energy resolution	0.17
flavor-dependent jet response	0.03
b-tagging	0.11
Signal modeling	
ISR/FSR	0.32
forward/backward ISR	0.36
hadronisation and showering	→ 1.08
higher order correction	0.80
PDF	0.60
Background model	
fake background normalization	0.35
fake background shape	0.35
background normalization	0.53
Calibration	
$\Delta y_{t\bar{t}}$ model	→ 2.7
calibration statistics	0.4
Total	3.3

- Test SM or test for BSM, but need to consider unknown polarization in calibration

→ A_{FB} = 18.0 ± 6.9 (tot.) %

A_{FB} = 18.0 ± 6.9 (tot.) ± 5.1(model) %

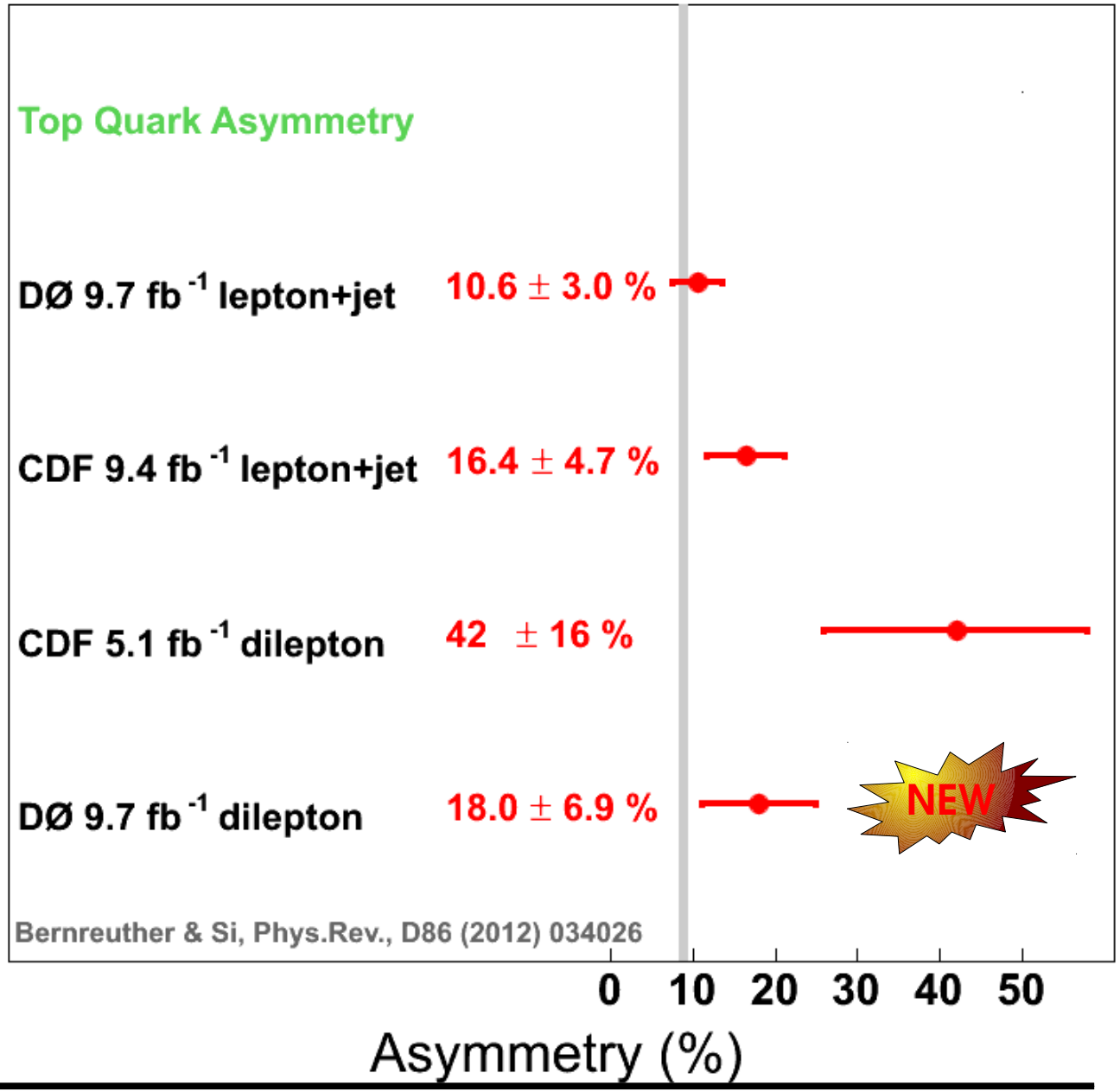
Top quark asymmetries

Summary of the current situation:

- Most Tevatron results use full data sets
- Combinations to come

Theory vs Data:

- Latest @NNLO is ≈ 10%
- lepton+jets:
  → agreement with D0, while CDF little higher
- Dilepton:
  → latest D0 result on the high side

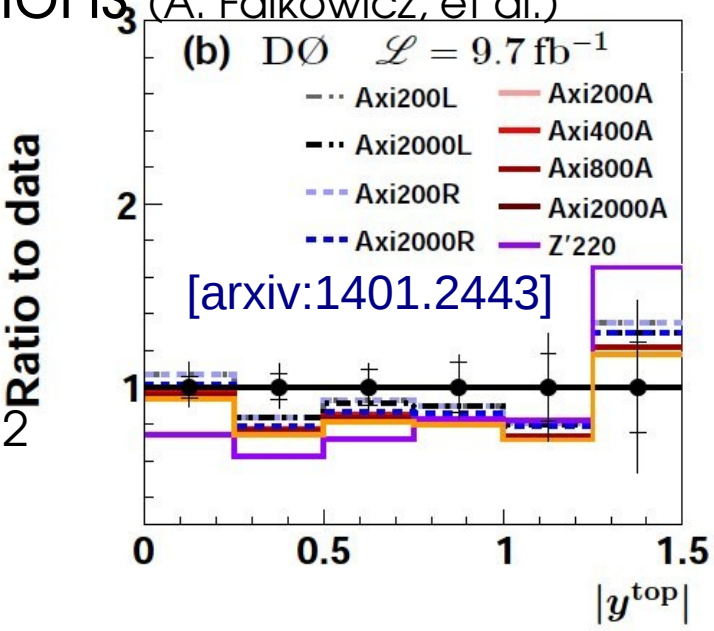


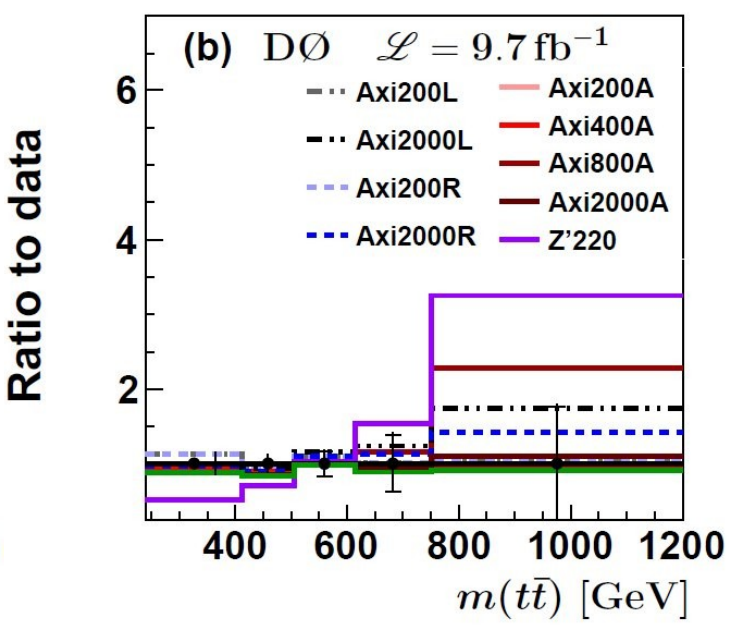
Axi gluons & Z' models

- Various axi gluon models with different couplings, differential cross section predictions (A. Falkowicz, et al.)
  - Compare various models to unfolded cross section data using full covariance matrix:

$$\chi^2 = \sum_{i,j} (y-\mu)_i \cdot \mathrm{cov}^{-1}_{i,j} \cdot (y-\mu)_j$$

  - Models with masses of 0.2 to 2 TeV for L (left), R (right), A (axial)
  - Models are constrained by σ(tt), A_{FB} and high tail of dσ/dm(tt) at Tevatron and LHC
  - Some models are in tension with the presented data !
    → Z', some axi gluons

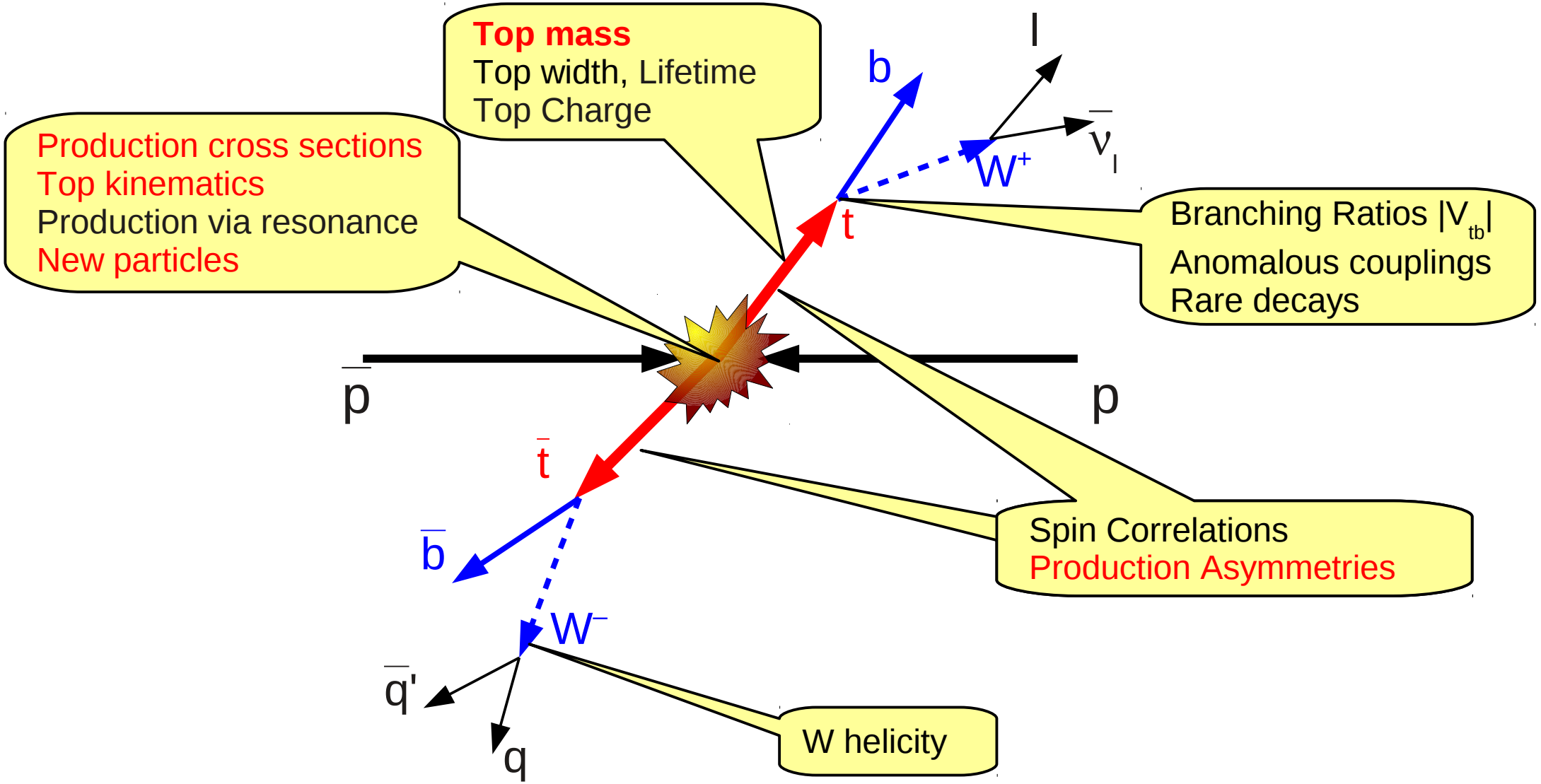




	$M(t\bar{t})$ $[\chi^2/ndf]$	p_T^{top} $[\chi^2/ndf]$	$\lvert y^{\mathrm{top}}\rvert$ $[\chi^2/ndf]$
axi200L	0.96	1.07	1.20
axi200R	0.96	1.07	1.20
axi200A	0.85	3.55	3.88
axi400A	0.44	2.65	3.26
axi800A	0.97	2.86	3.23
axi2000L	0.58	1.27	3.78
axi2000R	0.43	1.94	2.75
axi2000A	0.88	3.56	4.11
Z'220	4.95	8.27	7.48

Tevatron data adds sensitivity at low mass
→ specific models constrained

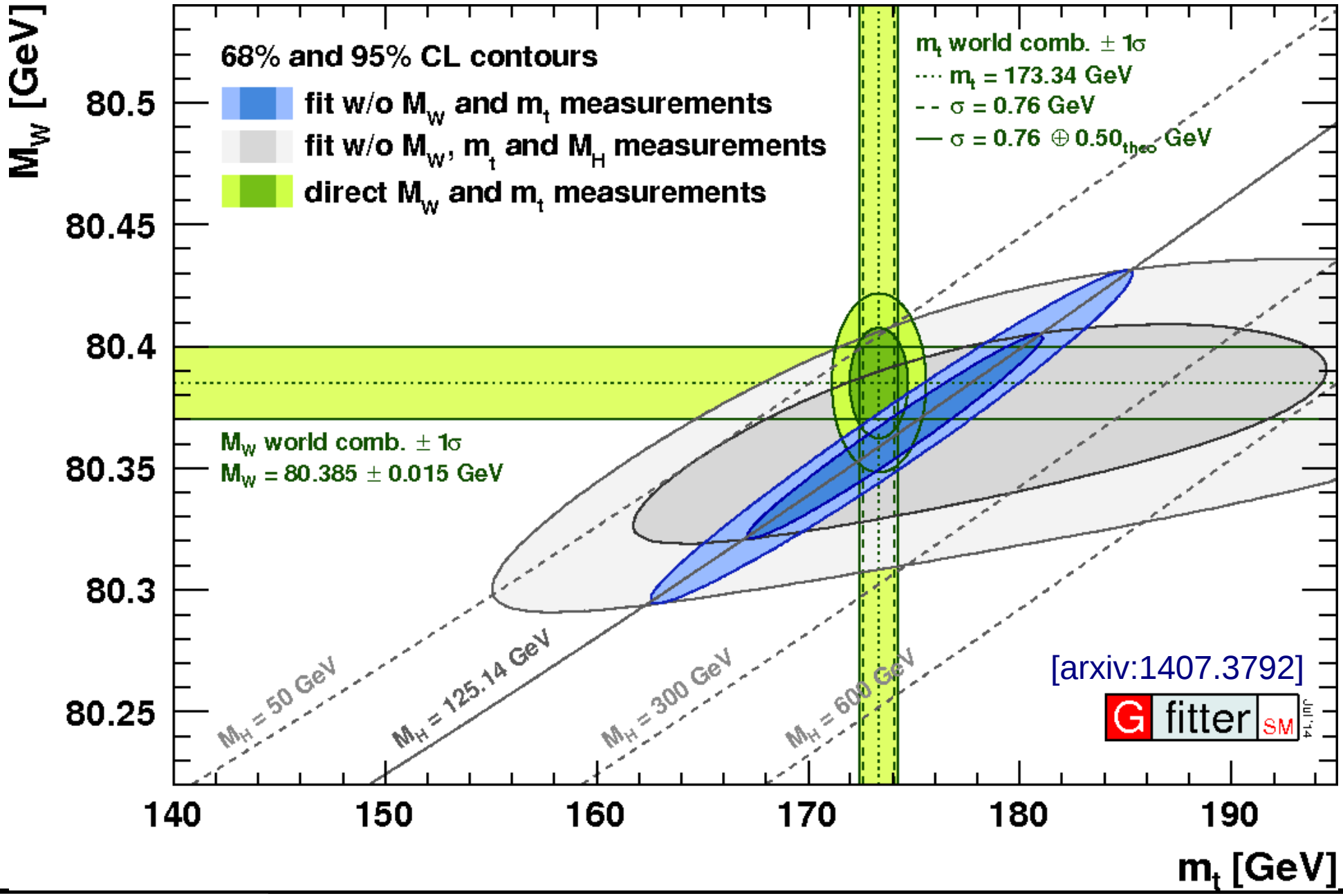
A rich & wide field



→ and can have all kinds of associated $t\bar{t}$ production: $t\bar{t}$ + H/Z/W/jets/$t\bar{t}$ or new particles

→ Recent review of top quark measurements at Tevatron: Publ. in MPLA (arxiv:1404.1842)

SM self-consistency



Top mass: methods

- **Matrix Element method** (ME) calculates event probability densities (PD) from **differential cross sections** and detector resolutions
  - Matrix element maximizes statistical information, current implementations are leading order

- **Ideogram method** event likelihood based on Breit-Wigner (signal) convoluted with detector resolutions

$$\mathcal{L}\left(\text{sample}|m_{\text{t}},\text{JSF}\right) = \prod_{\text{events}}\left(\sum_{i=1}^{n} P_{\text{gof}}\left(i\right)\left(\sum_{j} f_j P_j(m_{\text{t},i}^{\text{fit}}|m_{\text{t}},\text{JSF}) \times P_j(m_{\text{W},i}^{\text{reco}}|m_{\text{t}},\text{JSF})\right)\right)^{w_{\text{event}}}$$

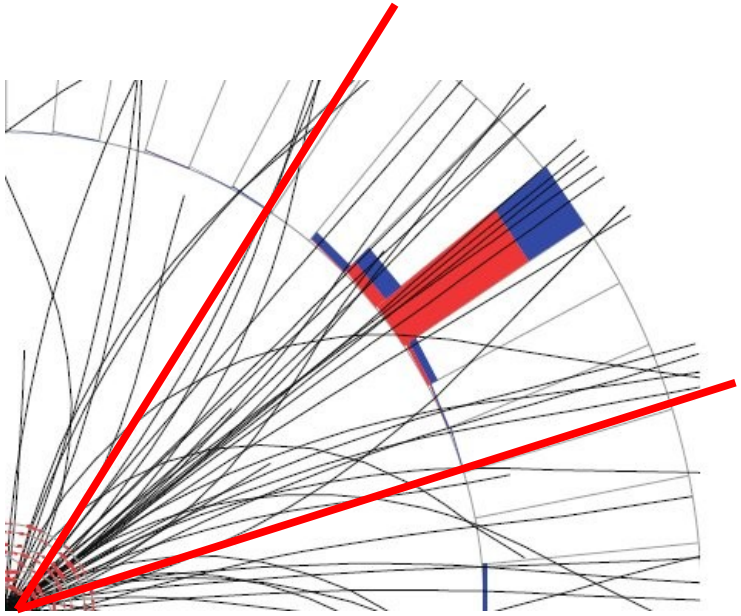
- **Template method** compares histograms in data to the ones derived from simulations (including detector resolutions) for different m_t

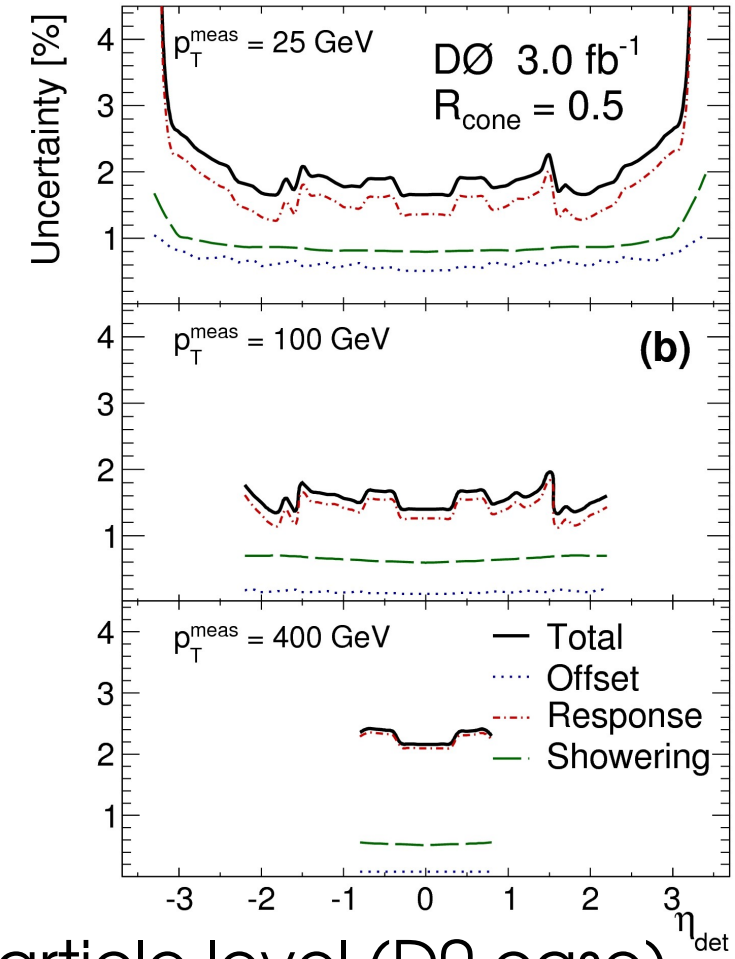
$$\chi^2_t = \frac{\left(m^{(1)}_{jj} - M_W\right)^2}{\Gamma^2_W} + \frac{\left(m^{(2)}_{jj} - M_W\right)^2}{\Gamma^2_W} + \frac{\left(m^{(1)}_{jjb} - m^{rec}_t\right)^2}{\Gamma^2_t} + \frac{\left(m^{(2)}_{jjb} - m^{rec}_t\right)^2}{\Gamma^2_t} + \sum_{i=1}^{6} \frac{\left(p^{fit}_{T,i} - p^{meas}_{T,i}\right)^2}{\sigma^2_i}$$

- These "standard" methods depend on MC → We measure MC mass

Jet Energy Scale (JES)

- Its not just a "jet" from a quark...complex:
  - Initial state & final state radiation, underlying events, multiple parton interaction, the parton shower





- JES corrects reconstructed energy back to particle level (D0 case)

$$E^{\mathrm{ptcl}} = \frac{E^{\mathrm{meas}} - E_{\mathrm{O}}}{R \cdot S}$$

E_O – offset correction
R – response correction
S – Showering corrections, In/Out of cone

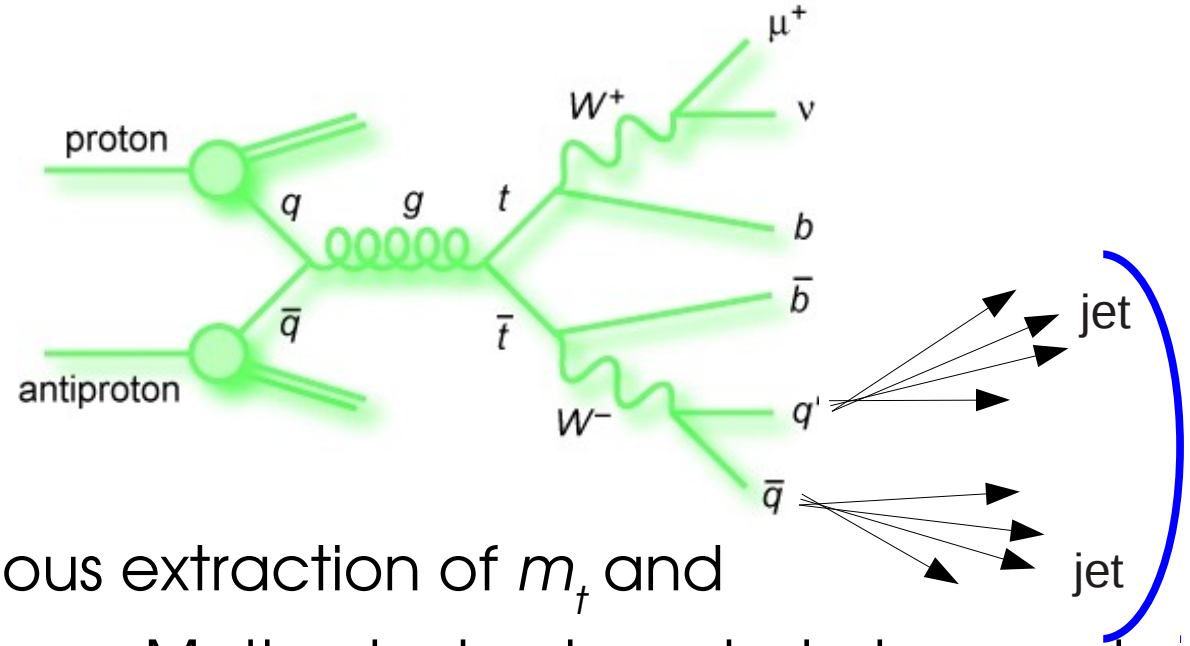
  - To be exact: D0 JES only valid for Pythia Tune A

Matrix element method

- Matrix Element method (ME) calculates event probability densities (PD) from differential cross sections and detector resolutions:

$$P(x, m_t) = \frac{1}{\sigma(m_t)} \int \sum \underbrace{d\sigma(y, m_t)}_{\text{LO ME}}\, dq_1 dq_2\, \underbrace{f(q_1) f(q_2)}_{\text{PDFs}}\, \underbrace{W(y, x, k_{\text{JES}})}_{\text{Transfer function}}$$

- k_{JES} is a global factor for the Jet Energy Scale, used for **in-situ calibration** using the hadronic W decay (does not account for b-JES)



Constrain jets of hadronic W decay to W mass

- Simultaneous extraction of m_t and global JES scale factor k_{JES} → Method robust against changes in event kinematics

Matrix element method

Updates compared to last publication/measurement:

- More data 3.6/fb → 9.7/fb (full Run II)
- Improved object IDs (e, μ, b)
- Faster method:
  - Random number generation
  - Modify treatment of kJES
    → Verified that method gets same result as with "old" method, but factor of **~ 100 faster**
  - Allowed **dramatic increase** in MC statistics

	Phys. Rev. D 84, 032004 (2011)
Higher-order effects	±0.25
ISR/FSR	±0.26
Hadronization and UE	±0.58
Color reconnection	±0.28
Multiple $p\bar{p}$ interactions	±0.07
Modeling of background	±0.16
W+jets heavy-flavor scale factor	±0.07
Modeling of b jets	±0.09
Choice of PDF	±0.24
Residual jet energy scale	±0.21
Data-MC jet response difference	±0.28
b-tagging efficiency	±0.08
Trigger efficiency	±0.01
Lepton momentum scale	±0.17
Jet energy resolution	±0.32
Jet ID efficiency	±0.26

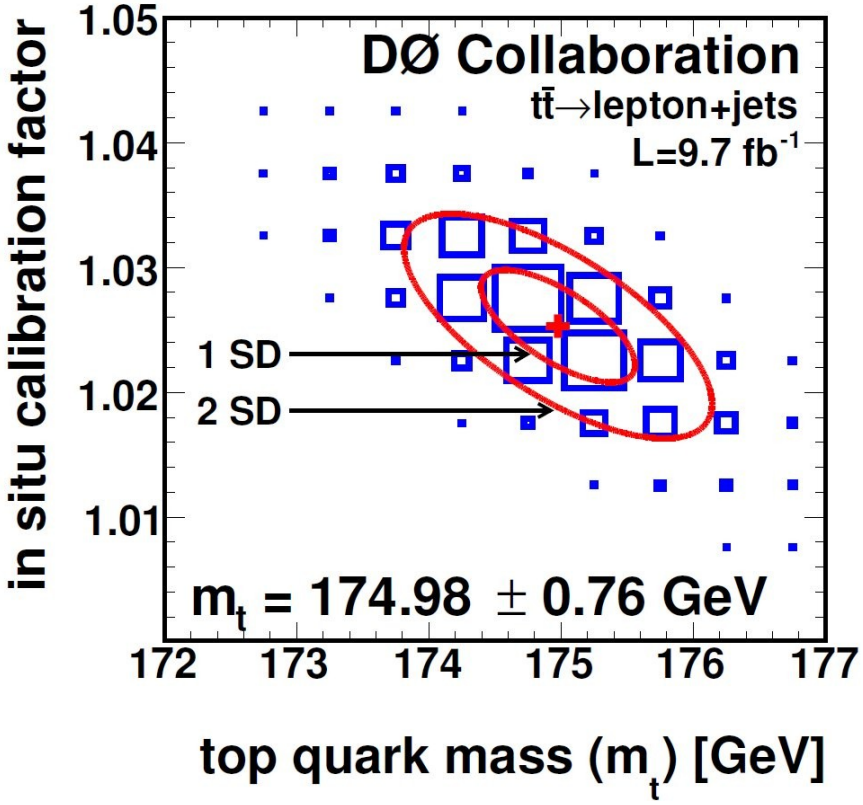
→ Typical statistical uncertainty:
~0.25 GeV → ~ 0.01 – 0.05 GeV

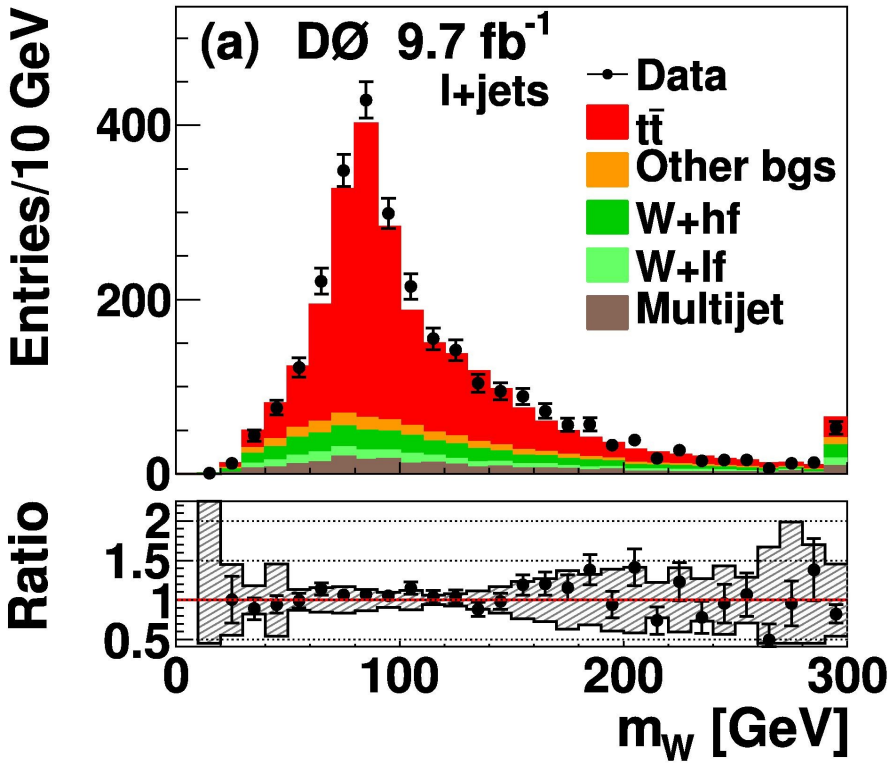
Systematic Uncertainties

- Complete list for 9.7/fb, see right
- Can discuss in more detail about the ← ones
- Couple of systematic uncertainties benefit from new calibrations and improved methods →

Source of uncertainty	Effect on m_t (GeV)
Signal and background modeling:	
Higher order corrections	+0.15
Initial/final state radiation	±0.09 ←
Hadronization and UE	+0.26 ←
Color reconnection	+0.10
Multiple $p\bar{p}$ interactions	−0.06
Heavy flavor scale factor	±0.06
b-jet modeling	+0.09
PDF uncertainty	±0.11
Detector modeling:	
Residual jet energy scale	±0.21
→ Flavor-dependent response to jets	±0.16 ←
→ b tagging	±0.10
Trigger	±0.01
→ Lepton momentum scale	±0.01
→ Jet energy resolution	±0.07
→ Jet ID efficiency	−0.01
Method:	
Modeling of multijet events	+0.04
Signal fraction	±0.08
MC calibration	±0.07

Top mass: matrix element





- l+jets selection with exactly 4 jets & ≥ 1 *b*-tag
- "Post-fit" distributions applying ME results are modeling the data – significant improvement comparing to "Pre-fit" case

PRL [arxiv:1405:1756]

→ m_t = 174.98 ± 0.58 (stat.+JES) ± 0.49 (syst) GeV

m_t = 174.98 ± 0.76 (total) GeV

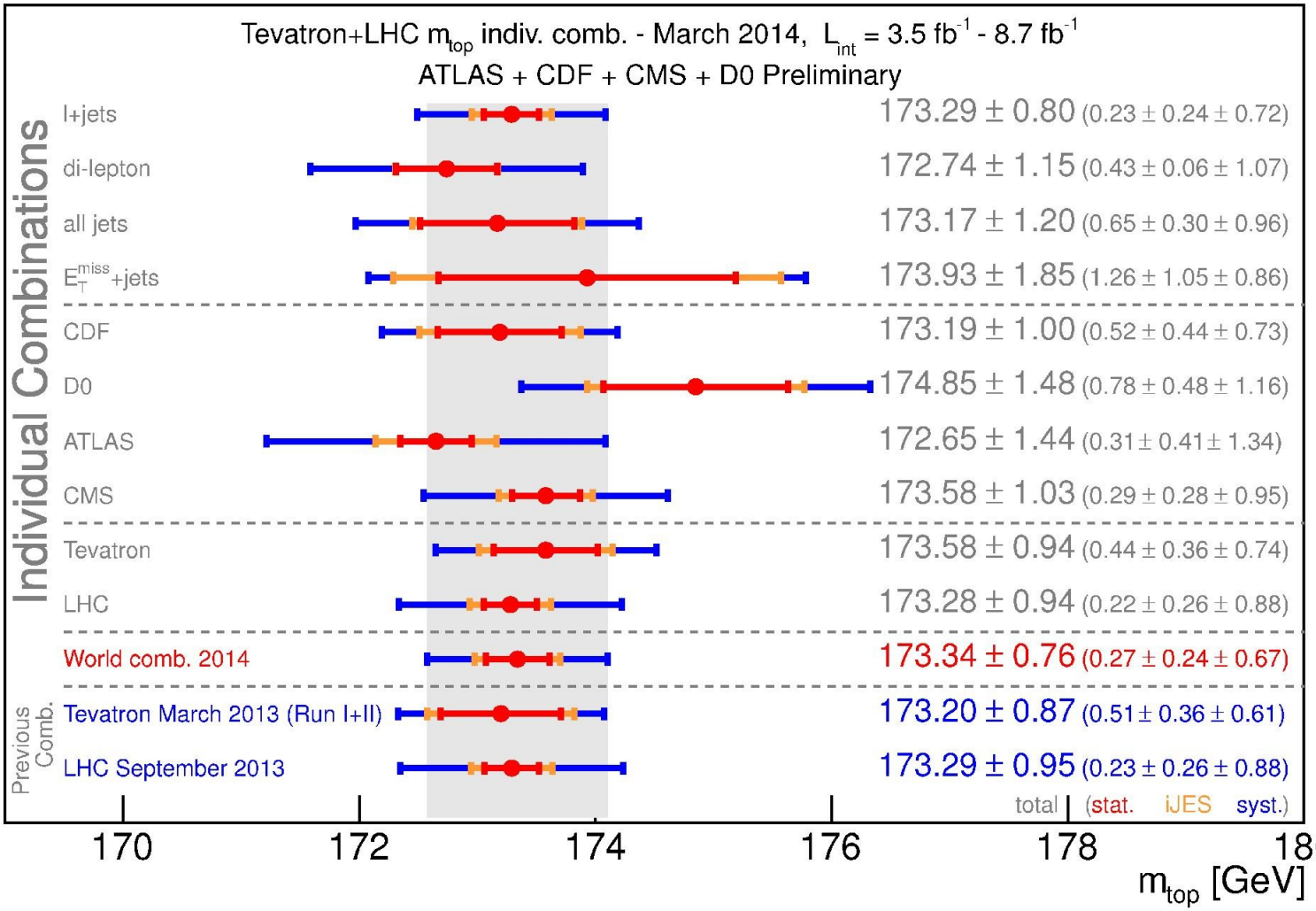
0.43% relative uncertainty

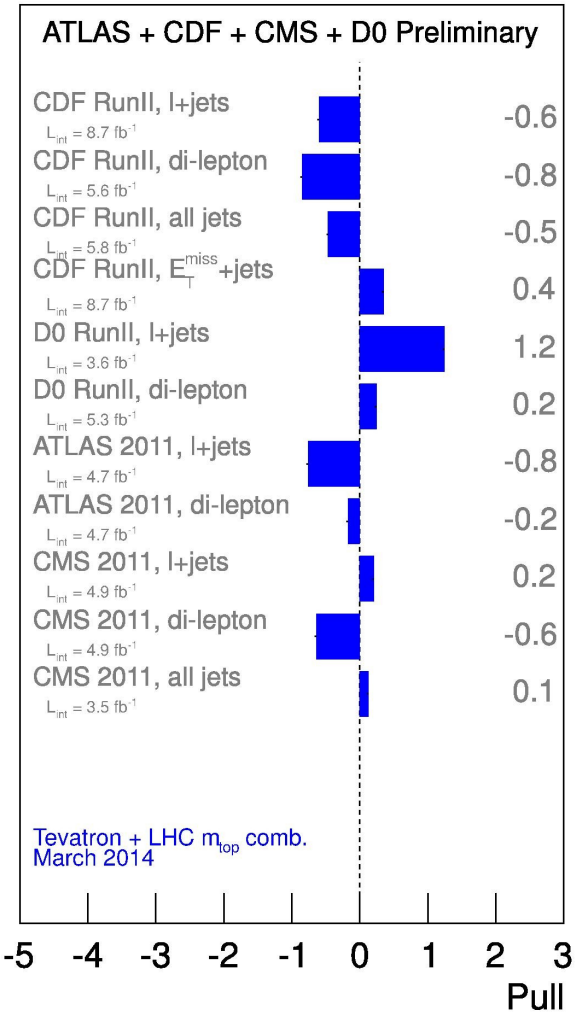
For comparison CMS: rel. precision is 0.45%

m_t = 172.02 ± 0.77 GeV

World combination

- First World Combination of top mass measurements
- Detailed study of inter-experiment & inter-collider correlations

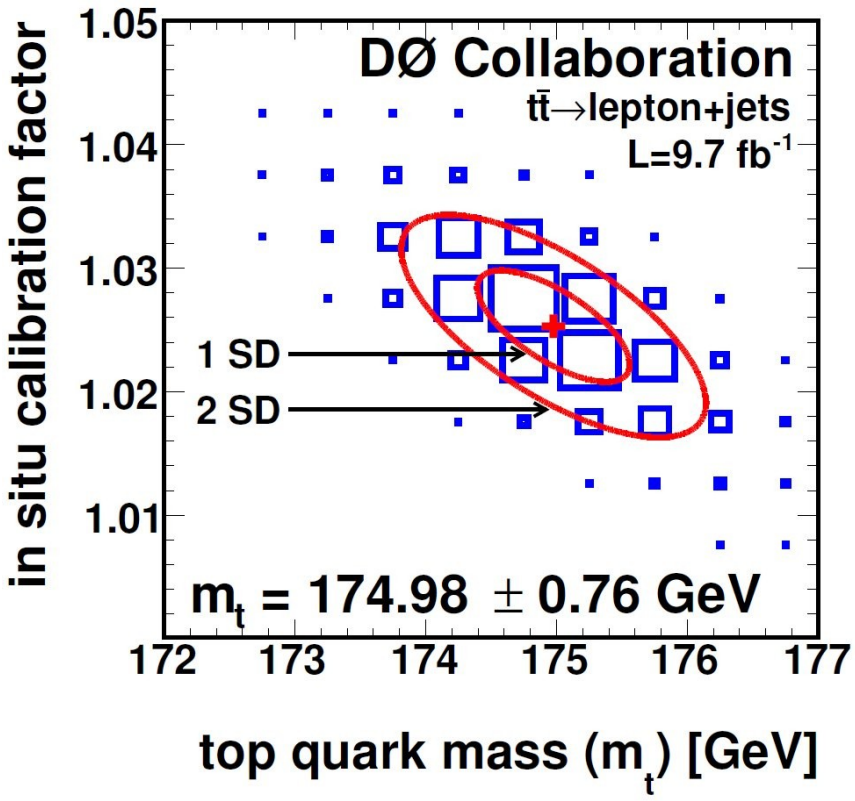


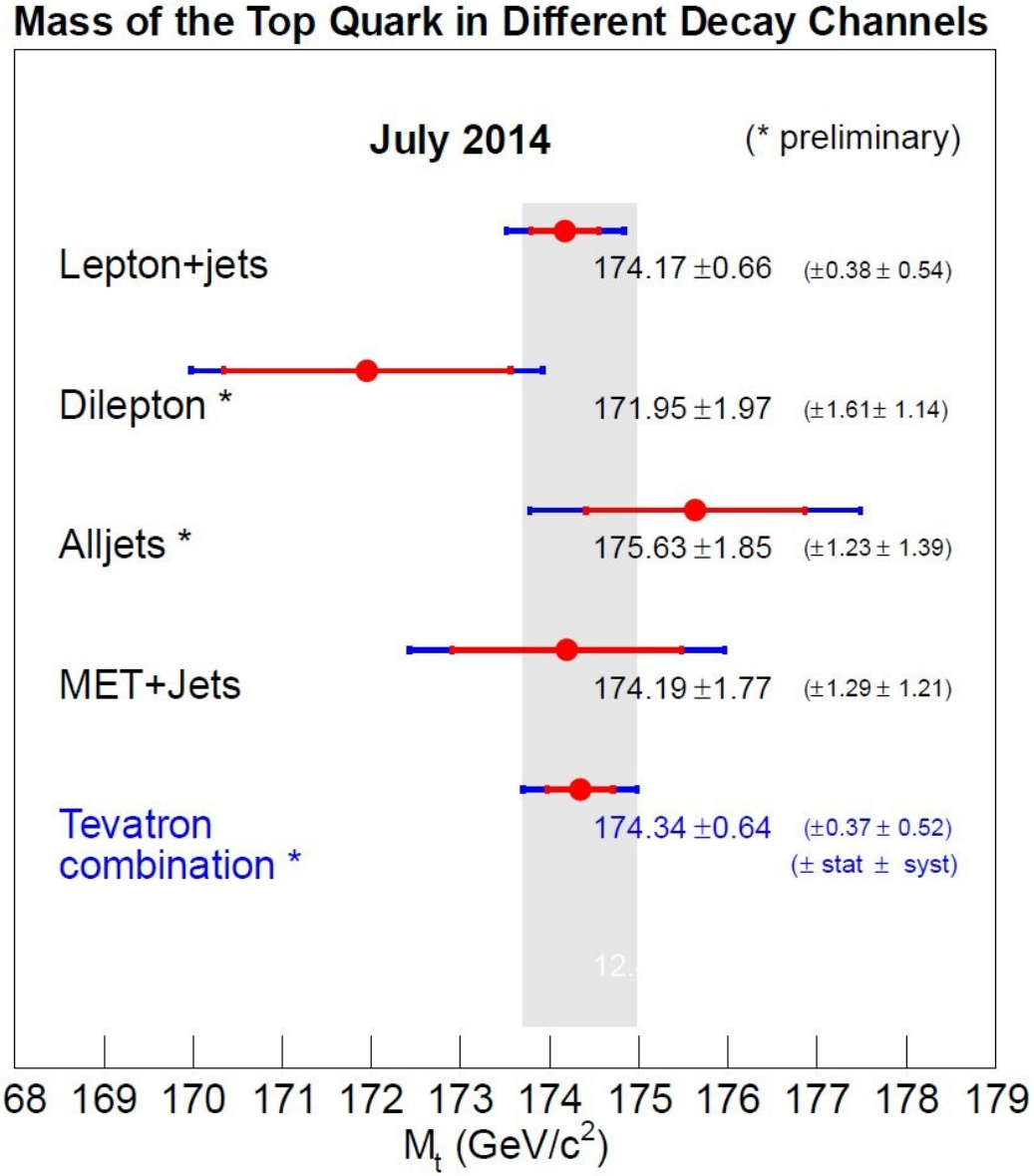


→ $m_t = 173.34 \pm 0.76$ (stat. + JES + syst.) GeV 0.44% rel. unc.

Mass – Tevatron

- Combination using BLUE
- Dominated by D0 l+jets result with precision of 0.43%





- New combination: precision is 0.37%

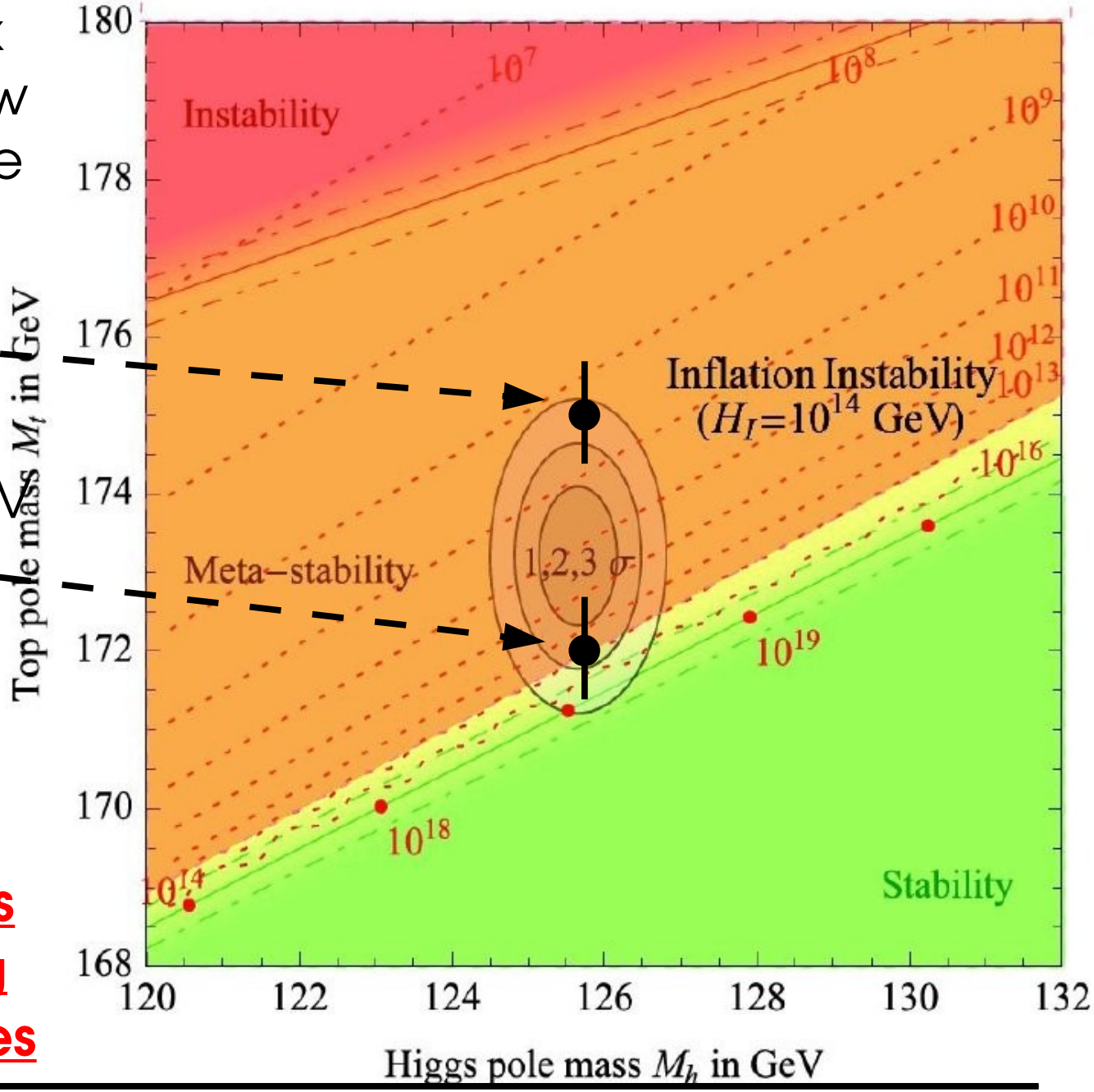
→ m_t = 174.34 ± 0.64 (stat. + JES + syst.) GeV 0.37% rel. unc.

Fate of THE UNIVERSE

- Stability of the electroweak vacuum assuming no new physics up to Planck scale
- Add latest measurements:

  D0 m_t = 174.98 ± 0.76 GeV

  CMS m_t = 172.02 ± 0.77 GeV



- New physics ?
- MC/Detector issue ?
- **Needs more measurements and better understanding of systematic uncertainties**

Summary & Conclusions

- Top quark physics is an exciting field
  - Able to constrain and test the standard model
  - Ideal probe to search for contributions of new physics

- Forward-backward asymmetry: Theory moved up, Experimental moved down (D0 ljets) but CDF higher, new D0 result higher
- and differential distributions are described by approx. NNLO pQCD

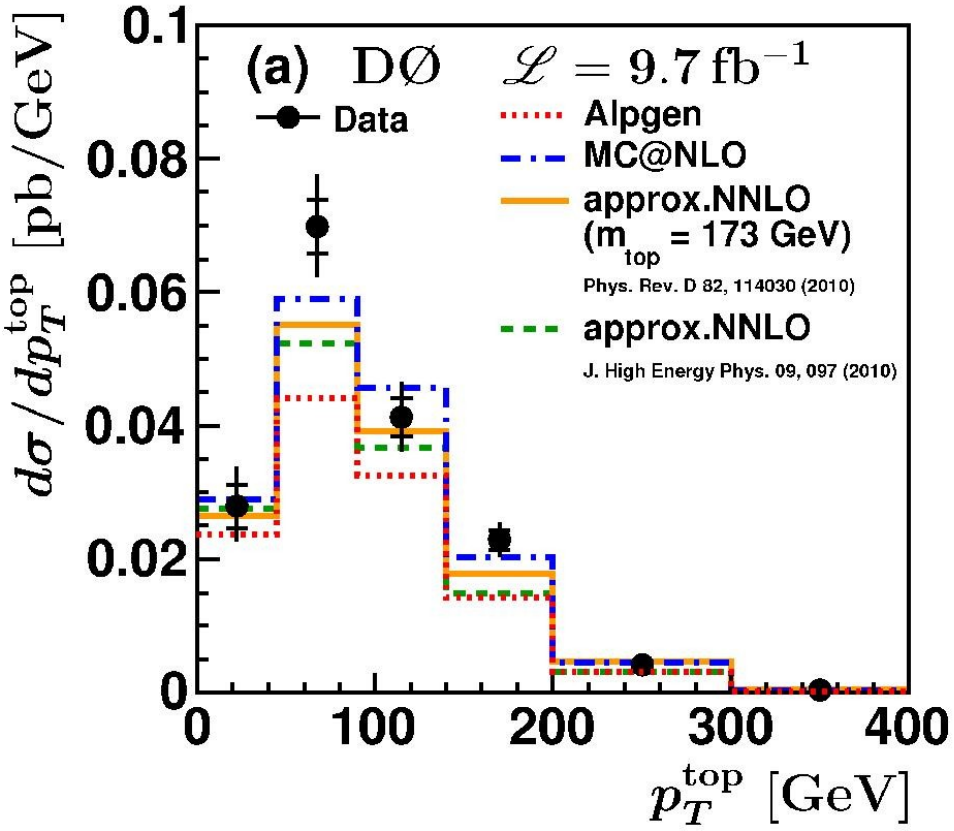
- High precision top mass measurements:
  - Systematic uncertainties are small
  - Need more understanding of what we measure exactly
  - Four collaboration wide effort

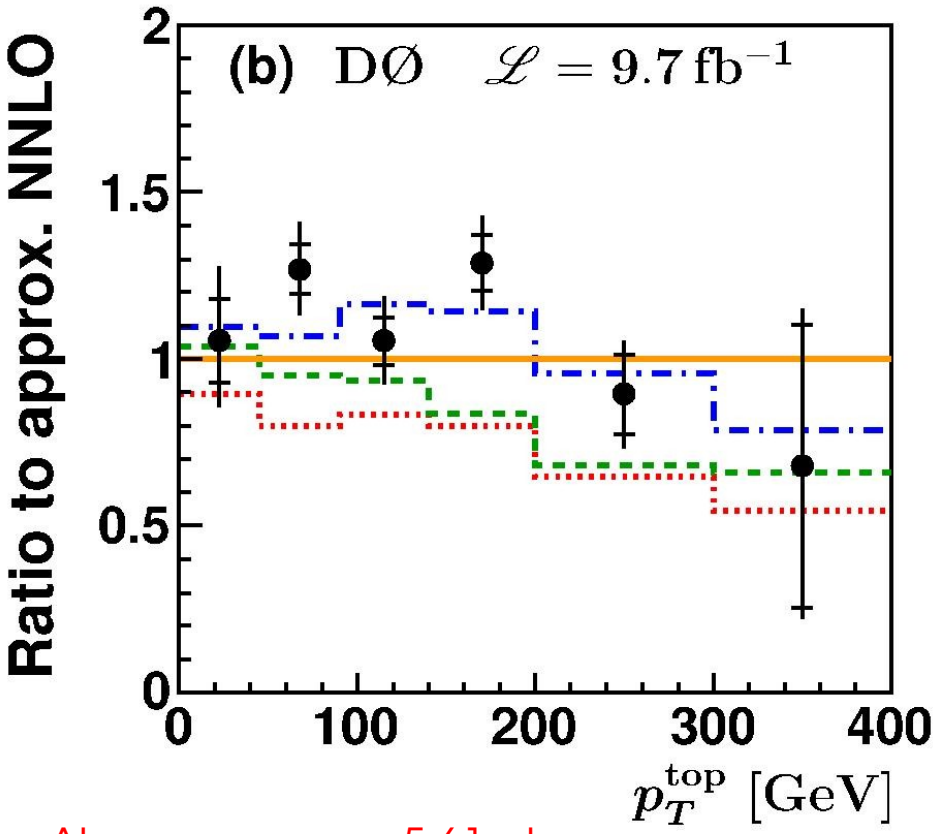
Outlook:

- More D0 measurements in the pipeline and expected to come out very soon – all part of the Legacy of the Tevatron

Backup

Differential cross sections





- Average t and $\bar{t}$ cross section
- Differential shape of data modeled by MC@NLO & approx. NNLO
- Not the full information: Correlations, all provided in the paper

- Alpgen: σ_{tot} = 5.61 pb
  MC@NLO: σ_{tot} = 7.56 pb
- app. NNLO by N.Kidonakis: p_T , arXiv:1009.4935
  σ_{tot} = 7.05 + 0.10 – 0.17 pb
- app. NNLO by L.Yang et al.: JHEP 09 (2010) 097
  σ_{tot} (NNLO + NNLL): σ_{tot} = 7.24 $^{+\ 0.24}_{-\ 0.26}$ pb
  [arXiv:hep-ph/1204.5201]

Systematic uncertainties

- Systematic uncertainties (D0) and their effect on differential cross sections:
  - Migration matrix derived from simulation with variation of parameters or modified background subtracted data
  - Systematic uncertainty is difference in unfolded results to nominal

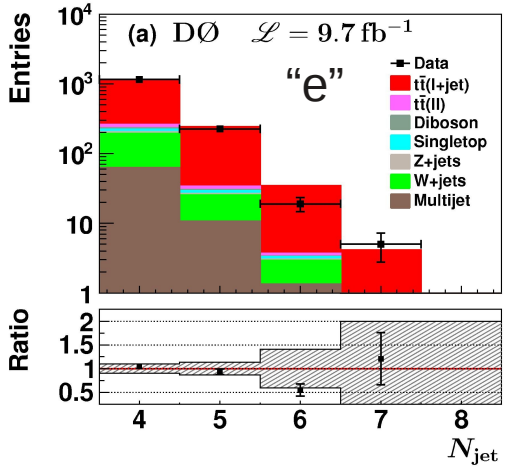
Source of uncertainty	Uncertainties, %	
	δ_{incl}	$\lvert\delta_{\text{diff}}\rvert$
Signal modeling	$^{+5.2}_{-4.4}$	4.0 – 14.2
PDF	$^{+3.0}_{-3.4}$	0.9 – 4.4
Detector Modeling	$^{+4.0}_{-4.1}$	3.1 – 13.7
Sample composition	± 1.8	2.8 – 9.2
Regularization strength	± 0.2	0.8 – 2.1
Integrated luminosity	± 6.1	6.1 – 6.1
Total systematic uncertainty	$^{+9.6}_{-9.3}$	8.5 – 23.1

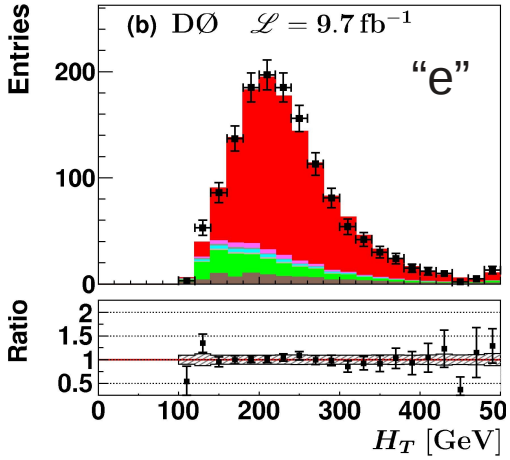
(for covariance matrix: systematic uncertainties assumed 100% correlated between bins)

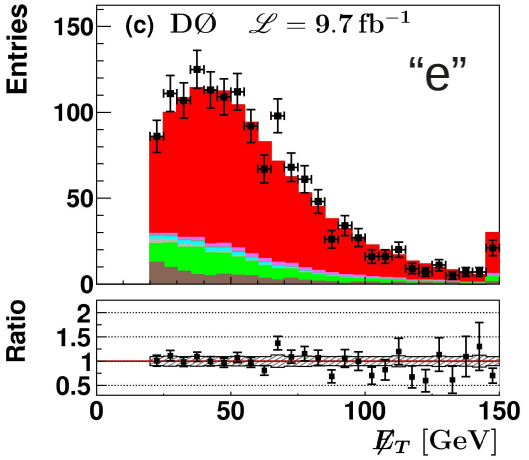
- Total systematic uncertainty (includes 6.1% luminosity): **9.4 % inclusive**

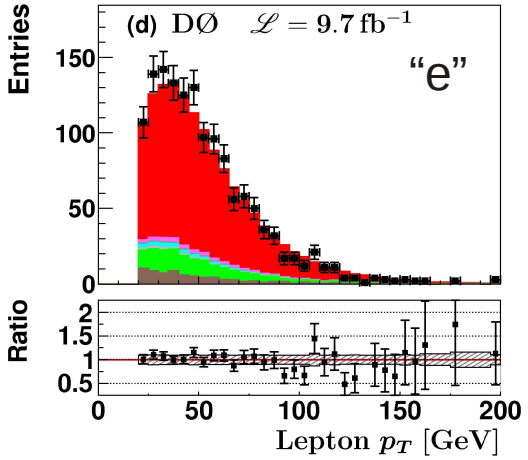
Control distributions

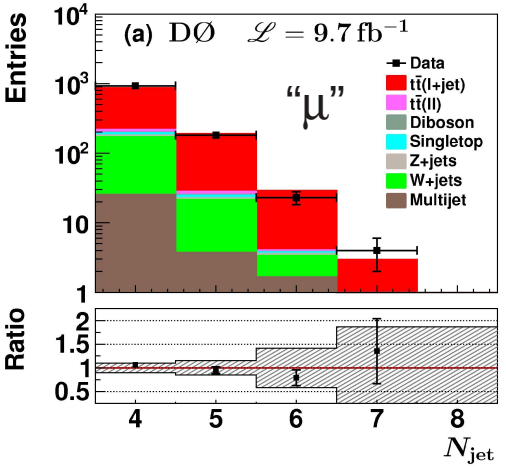
- Now check the signal region to measure differential cross sections: e or μ + ≥ 4 jets

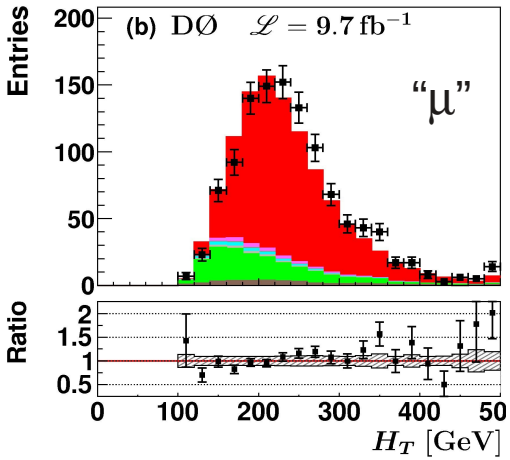


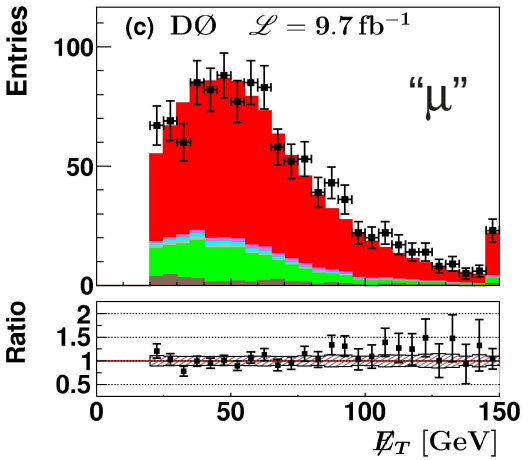


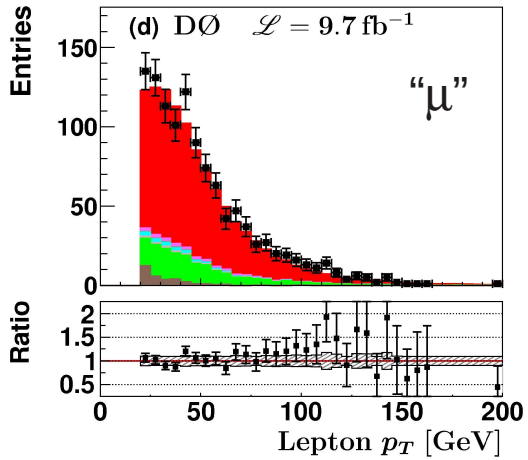






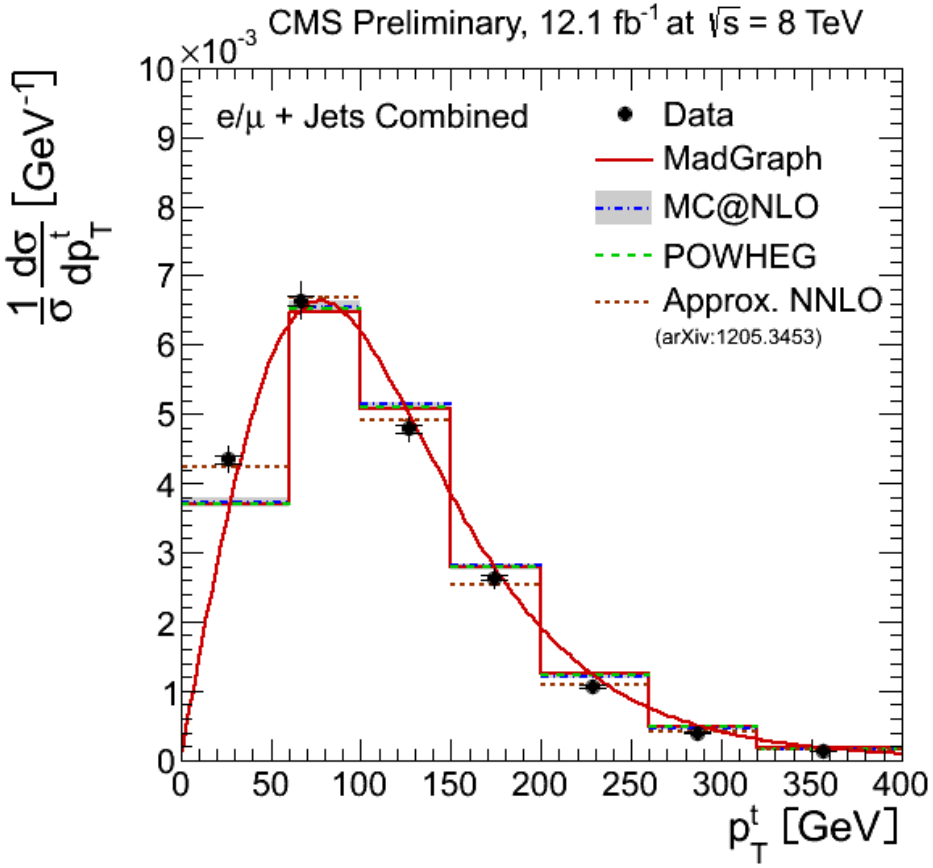




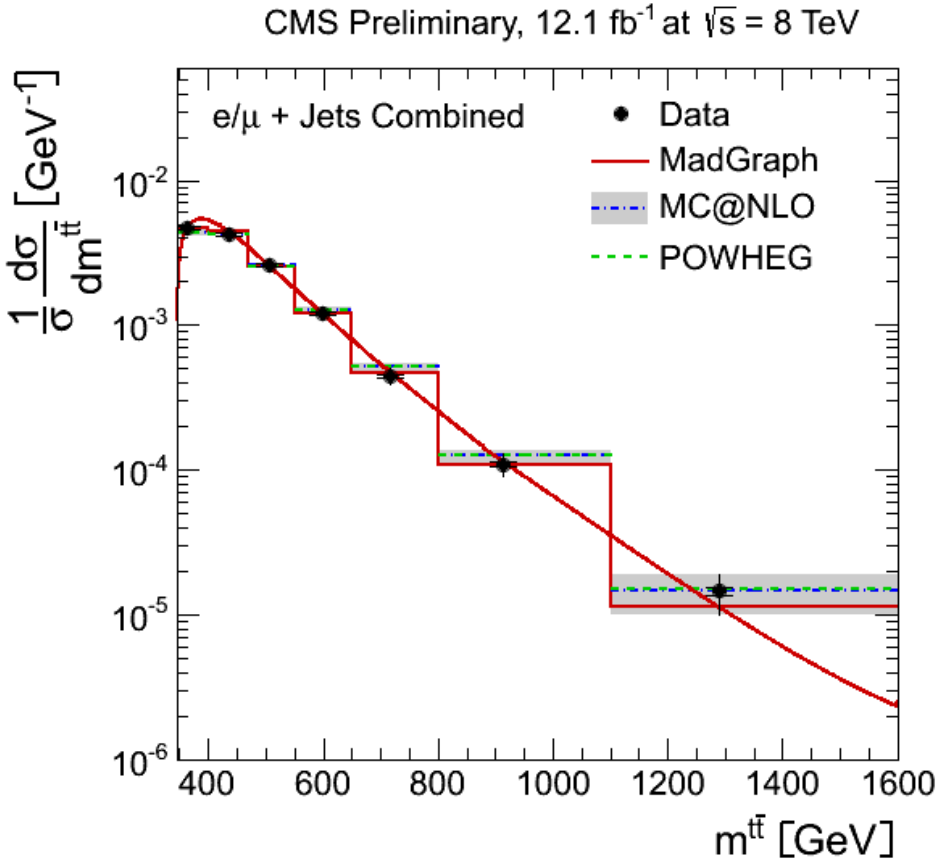


(assuming the measured $t\bar{t}$ cross section)

Differential cross sections



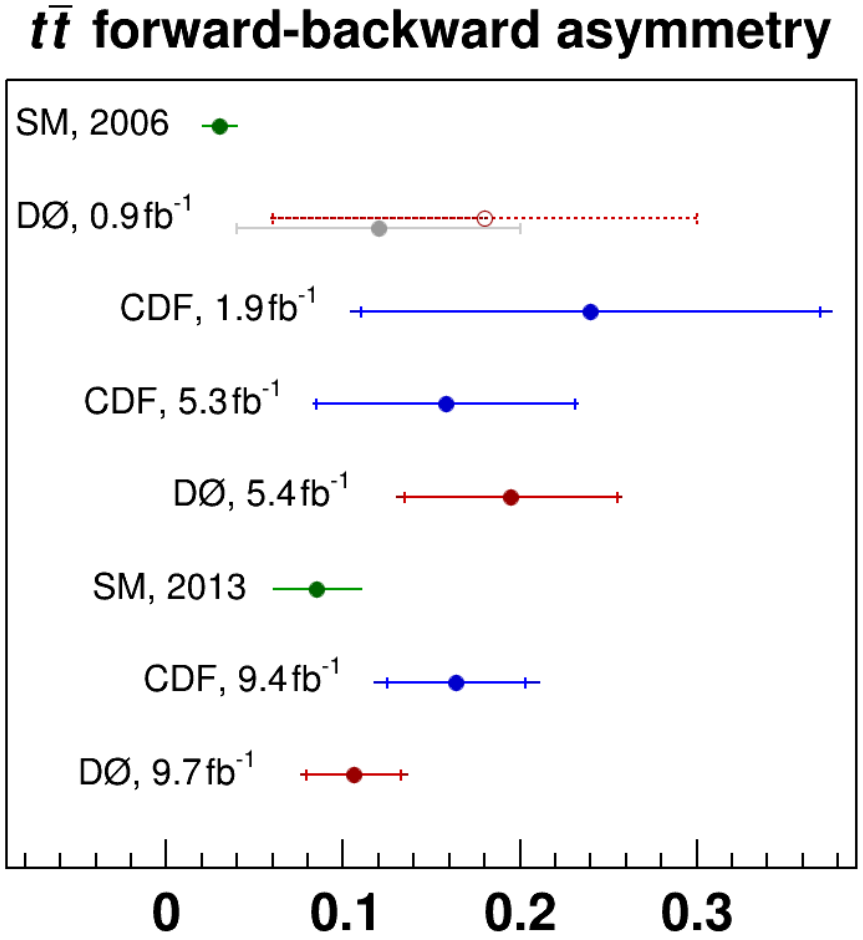
- Softer p_T in Data compared to MC@NLO, Powheg, Madgraph
  - Better described by approx. NNLO
- Similarly seen in 7 TeV results by CMS
- Some tension with ATLAS in first p_T bin

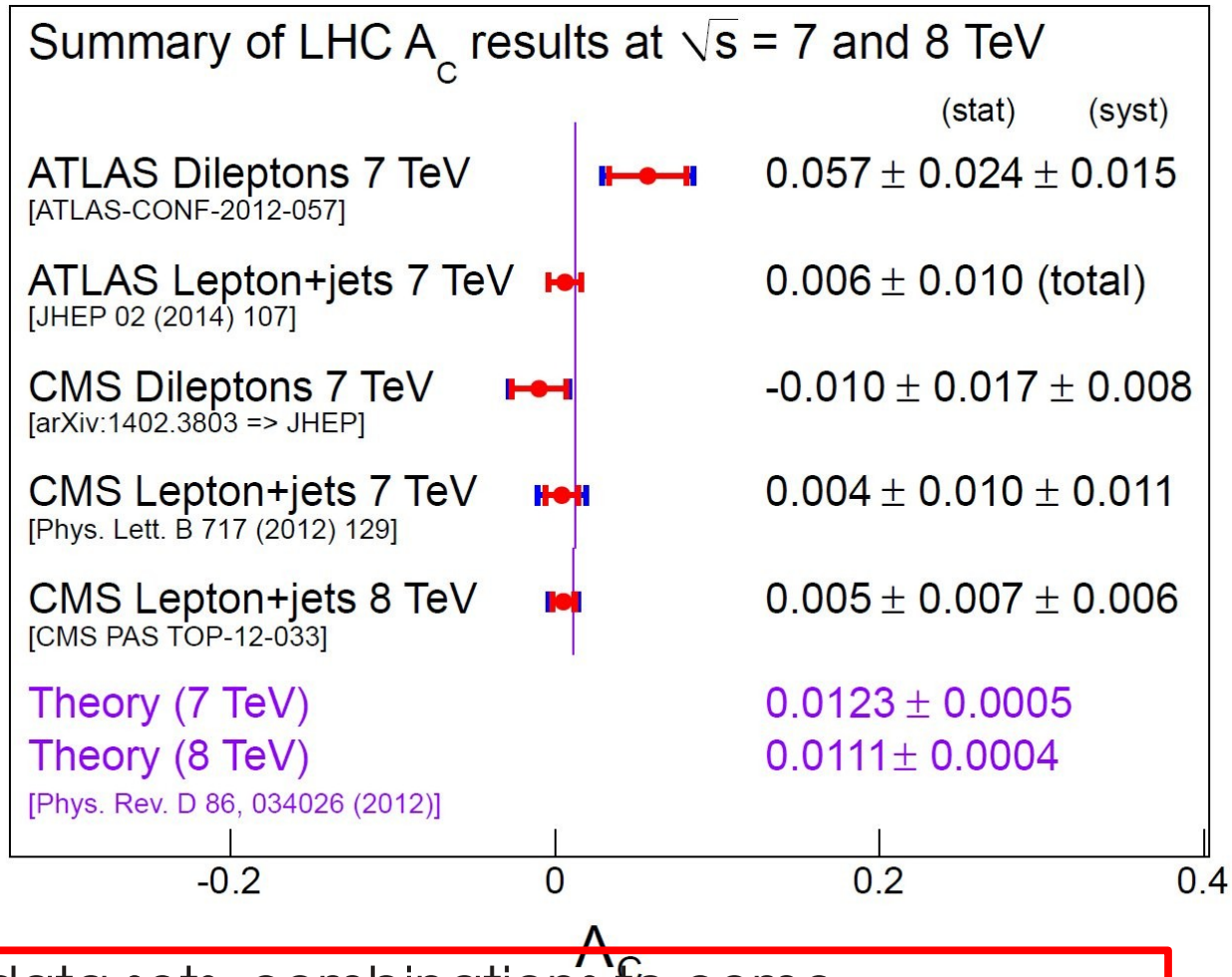


- Nicely described by basically all predictions
- No indications of new physics up to 1.6 TeV

Top quark asymmetries

- A summary of the current situation:





→
- Most Tevatron results use full data sets, combinations to come
- LHC measurements at increased √s expected to observe asymmetries
- Theory: Need QCD predictions at NNLO

Lepton asymmetries

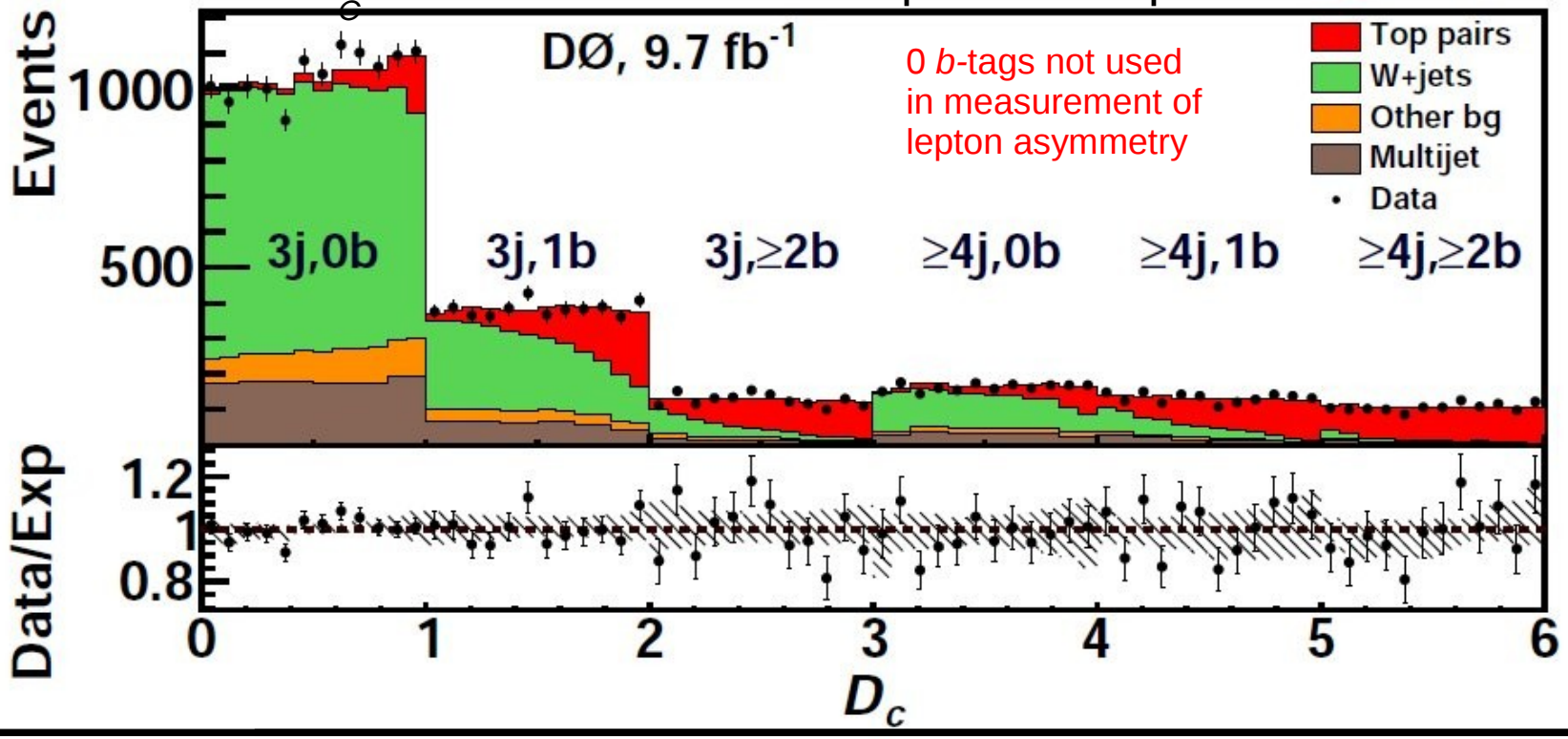
Acc. by PRD: [arxiv:1403:1294]

Leptonic asymmetries less affected by reconstruction effects:

- lepton+jets, full Run II data set, **Final measurements**

→ Maximize acceptance & precision: Include the 3 jet bin

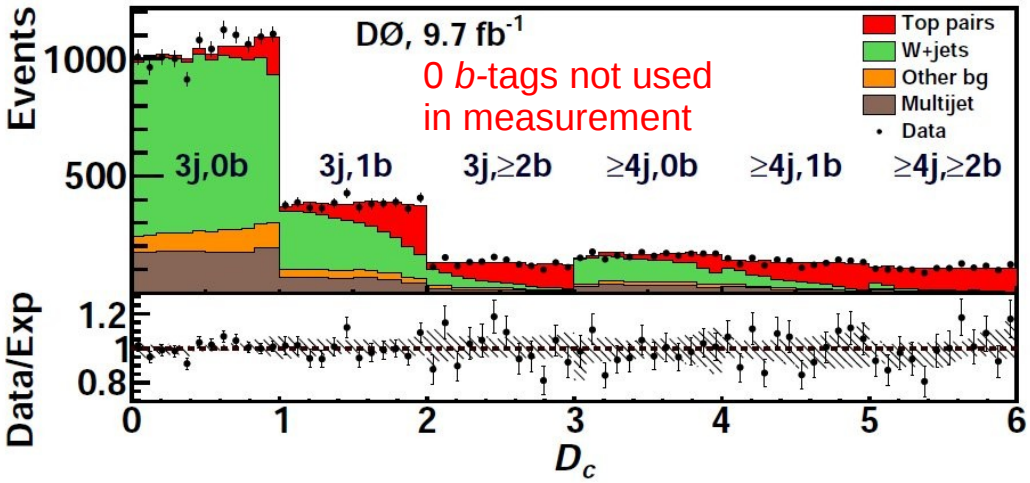
- Larger contribution of backgrounds in 3 jet bin
  → need to calibrate W+jets, use 0 *b*-tag
  → Improved background rejection and modeling
- Discriminant D_c to determine sample composition

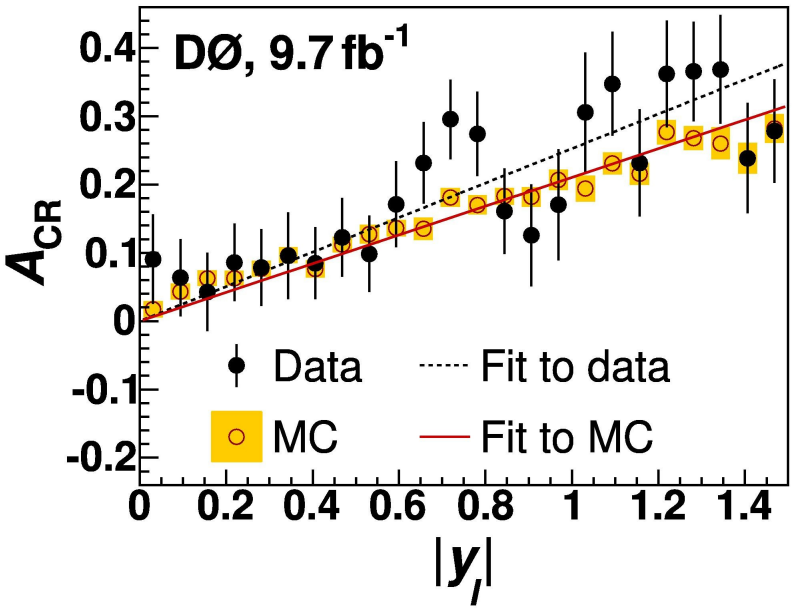


Lepton asymmetries

Acc. by PRD: [arxiv:1403:1294]

- lepton+jets, updated measurement of leptonic asymmetries
- Discriminant D_c to determine sample composition
- Need to calibrate W+jets, use 0 *b*-tag

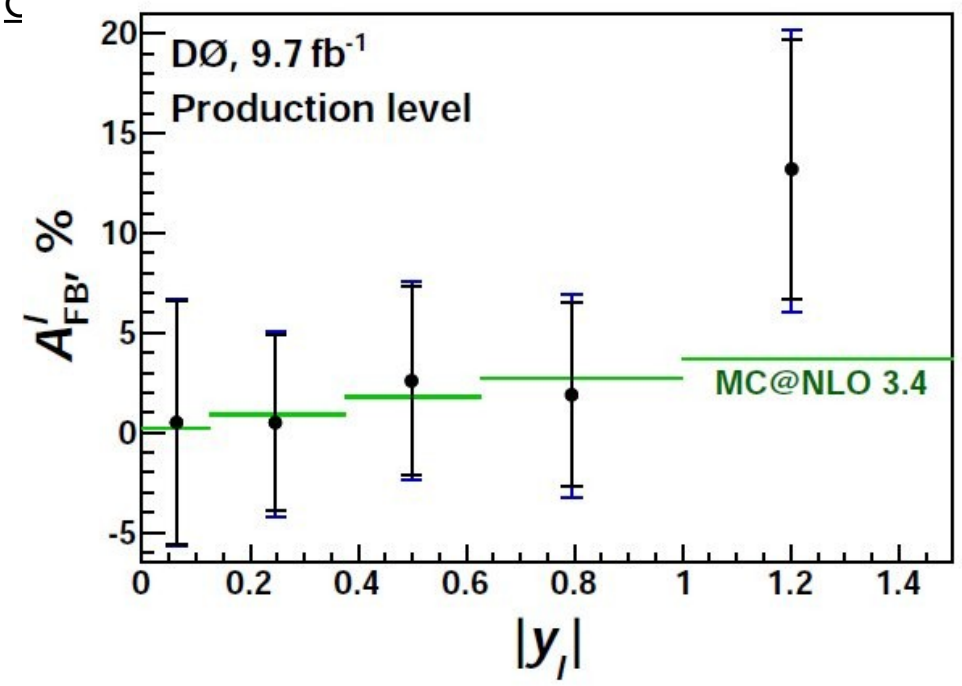




- Asymmetry in W+jets control region (CR) different from MC
- PDF uncertainty shown by yellow bars
- Full difference between data and MC slope taken as systematic uncertainty

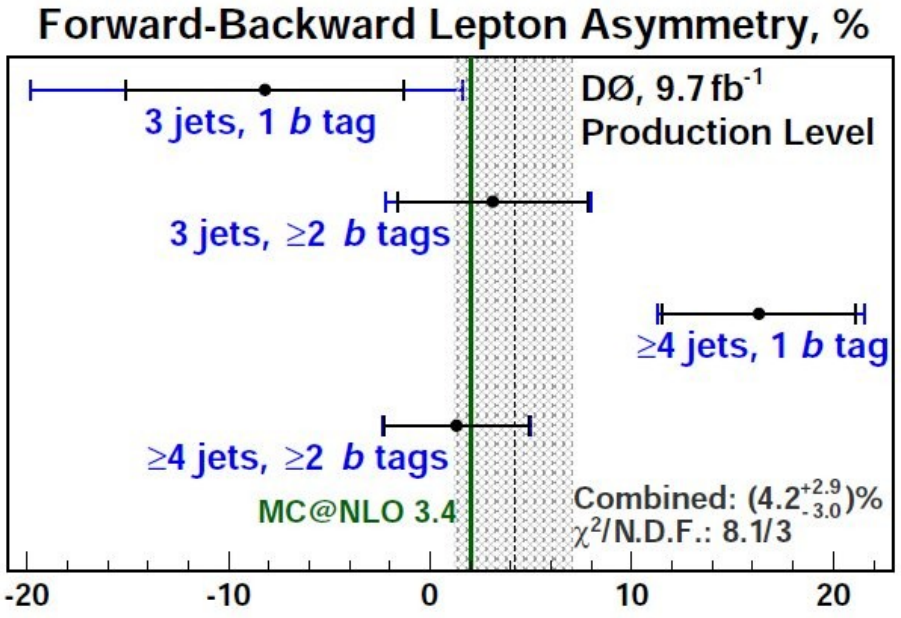
Lepton asymmetries

- C

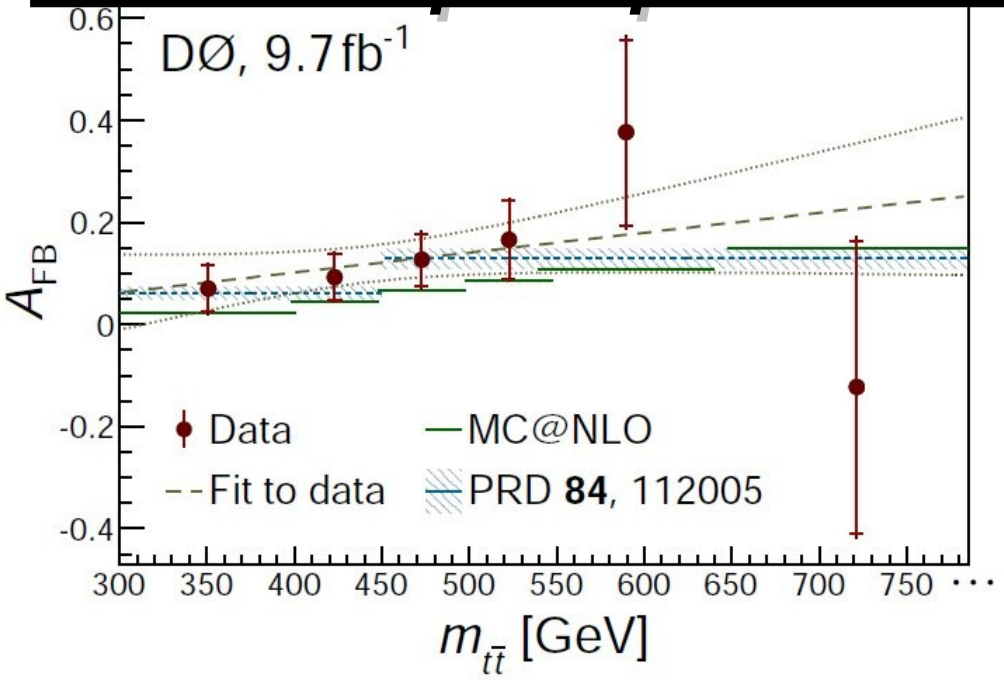


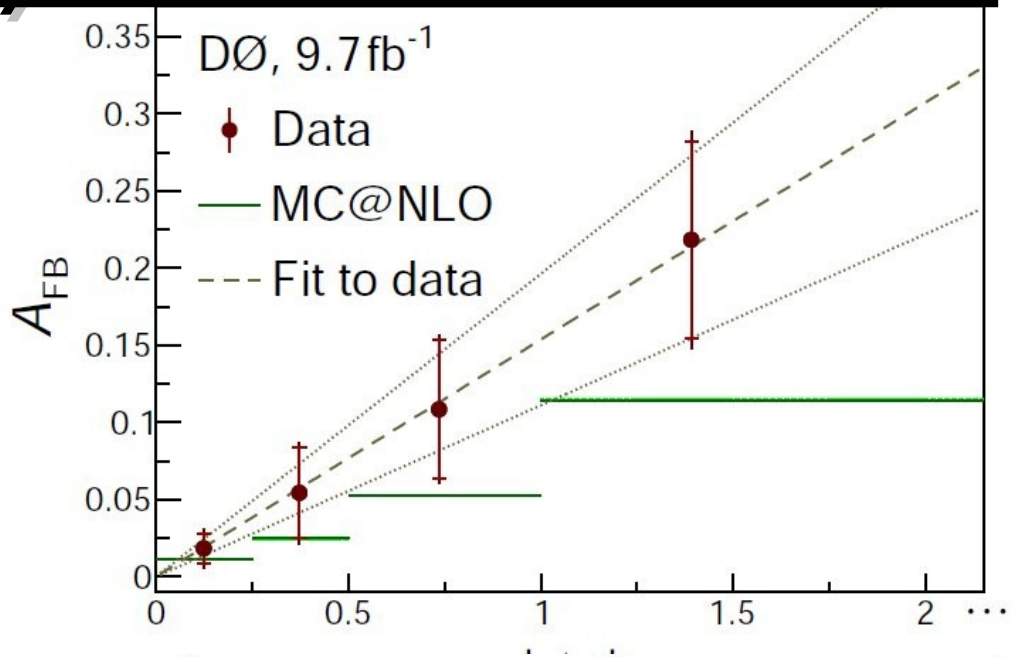
Source	Absolute uncertainty, % Reconstruction level Prediction	Reconstruction level Measurement	Prod. level Measurement
Jet reco	−0.1	–	–
JES/JER	+0.1	+0.1/−0.3	+0.2/−0.3
Signal modeling	–	−0.2	+0.6/−0.4
b tagging	±0.1	+0.5/−0.8	+0.8/−1.1
Bg subtraction	n/a	+0.1/−0.3	+0.1/−0.3
Bg modeling	n/a	+1.4/−1.5	+1.3/−1.5
PDFs	–	+0.3/−0.2	+0.1/−0.2
Total	±0.1	+1.5/−1.7	+1.7/−2.0

Channel	A^l_{FB}, % Data	MC@NLO
l+3 jets, 1 b tag	$-6.8 \pm 6.0\,(\text{stat.})^{+6.1}_{-5.6}\,(\text{syst.})$	2.7 ± 0.4
l+3 jets, ≥2 b tags	$3.7 \pm 4.3\,(\text{stat.})^{+1.1}_{-1.2}\,(\text{syst.})$	2.8 ± 0.3
l+≥4 jets, 1 b tag	$14.8 \pm 4.2\,(\text{stat.})^{+1.1}_{-1.2}\,(\text{syst.})$	0.5 ± 0.3
l+≥4 jets, ≥2 b tags	$-0.9 \pm 3.2\,(\text{stat.})^{+0.3}_{-0.9}\,(\text{syst.})$	1.1 ± 0.2
Total	$2.9 \pm 2.1\,(\text{stat.})^{+1.5}_{-1.7}\,(\text{syst.})$	1.6 ± 0.2

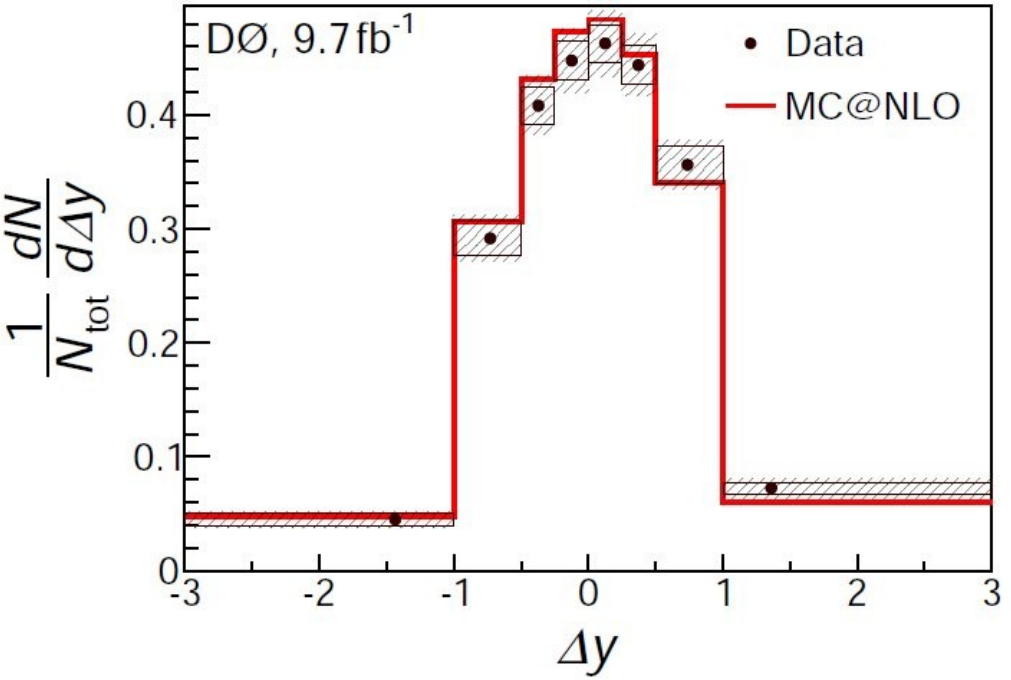


Top quark asymmetries





Source	Reco. level inclusive	Production level inclusive	Production level 2D
Background model	+0.7/−0.8	1.0	1.1–2.8
Signal model	< 0.1	0.5	0.8–5.2
Unfolding	N/A	0.5	0.9–1.9
PDFs and pileup	0.3	0.4	0.5–2.9
Detector model	+0.1/−0.3	0.3	0.4–3.3
Sample composition	< 0.1	< 0.1	< 0.1
Total	+0.8/−0.9	1.3	2.1–7.5



Hadronization/UE

- Compare **Alpgen+Pythia versus Alpgen+Herwig**, drawbacks:
  - D0 JES not applicable to Herwig, by "default" results in difference
  - A residual JES uncertainty in p_T, η is already applied

    → double-counting

- <u>To avoid double-counting refined strategy:</u>
  - Factorize effects in this comparison by using particle level jets

    → no effects from JES
  - Match particle level with detector level jets: $\Delta R \leq 0.25$
  - $p_T(t\bar{t})$ differs in Alpgen+Pythia versus Alpgen+Herwig, a larger effect already included via ISR/FSR

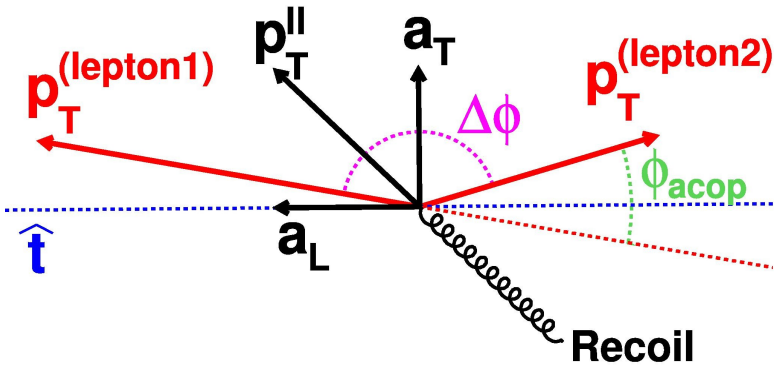
    → reweight $p_T(t\bar{t})$ in Herwig to default (Pythia)

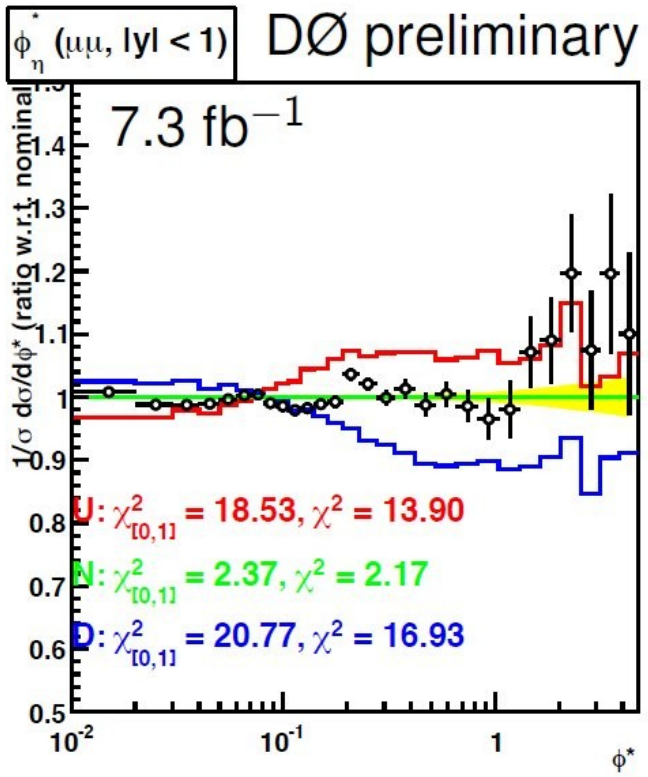
→ <u>Hadronization/UE uncertainty:</u>

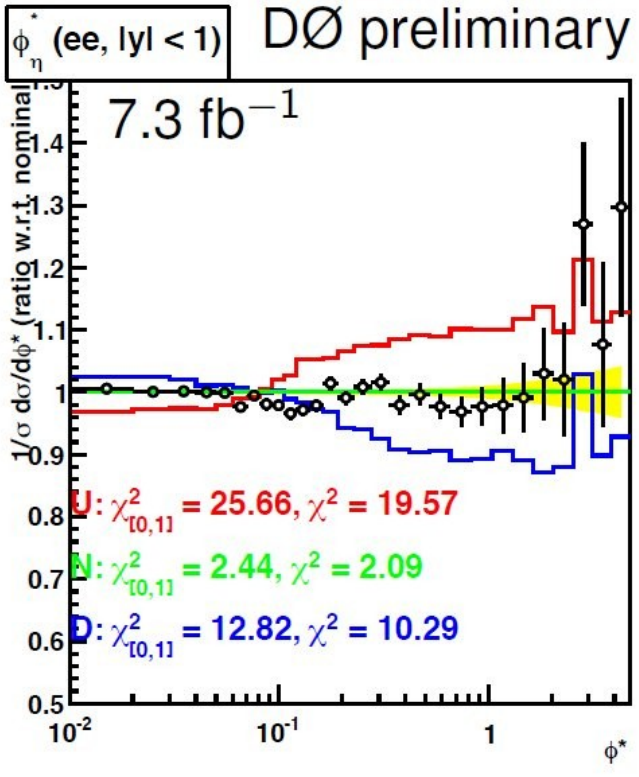
0.26 GeV (previously 0.58 GeV)

Initial/Final state radiation

- Vary renormalization scale in Alpgen
  - Constrain variation by employing D0 φ^* measurement, ($|y| < 1$; $1 < |y| < 2$; $|y| < 2$)
  - Variation by factor 1.5 covers spread of differences between data and MC





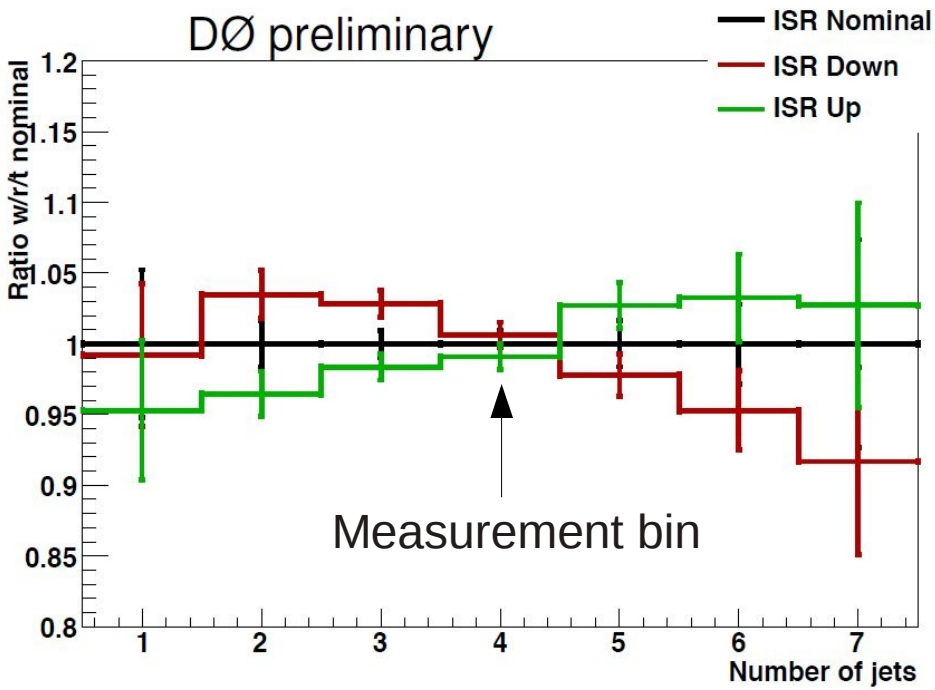


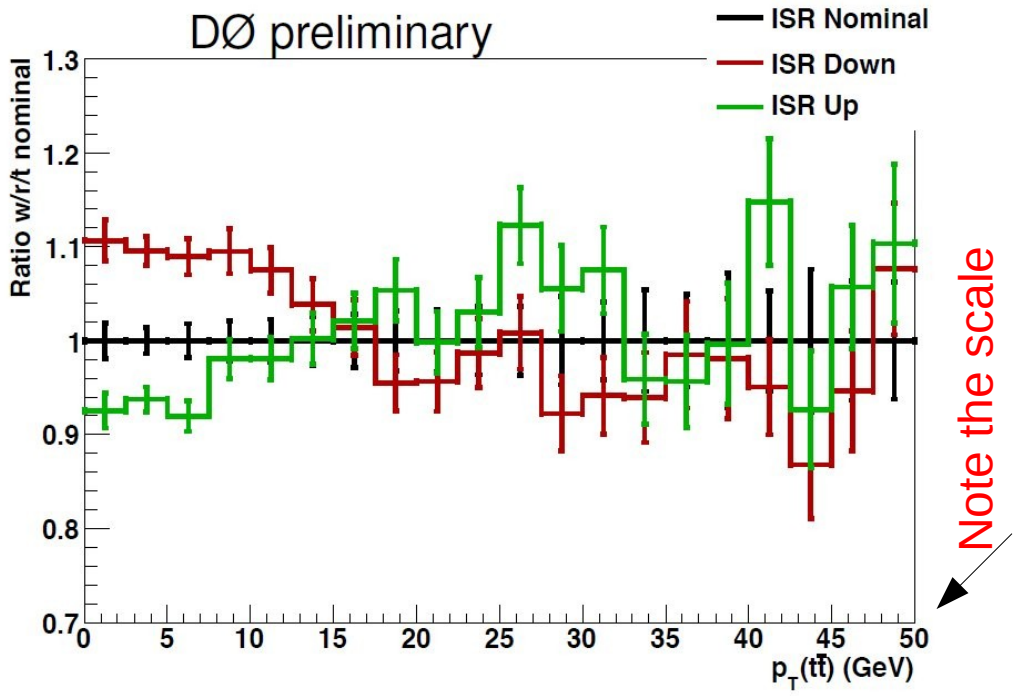
- Measurement only depends on directions of two leptons
  → measured precisely

Phys. Rev. Lett. 106, 122001 (2011)

Initial/Final state radiation

- Vary renormalization scale in Alpgen
  - Constrain variation by employing D0 φ^* measurement, ($|y| < 1$; $1 < |y| < 2$; $|y| < 2$)
  - Variation by factor 1.5 covers spread of differences between data and MC

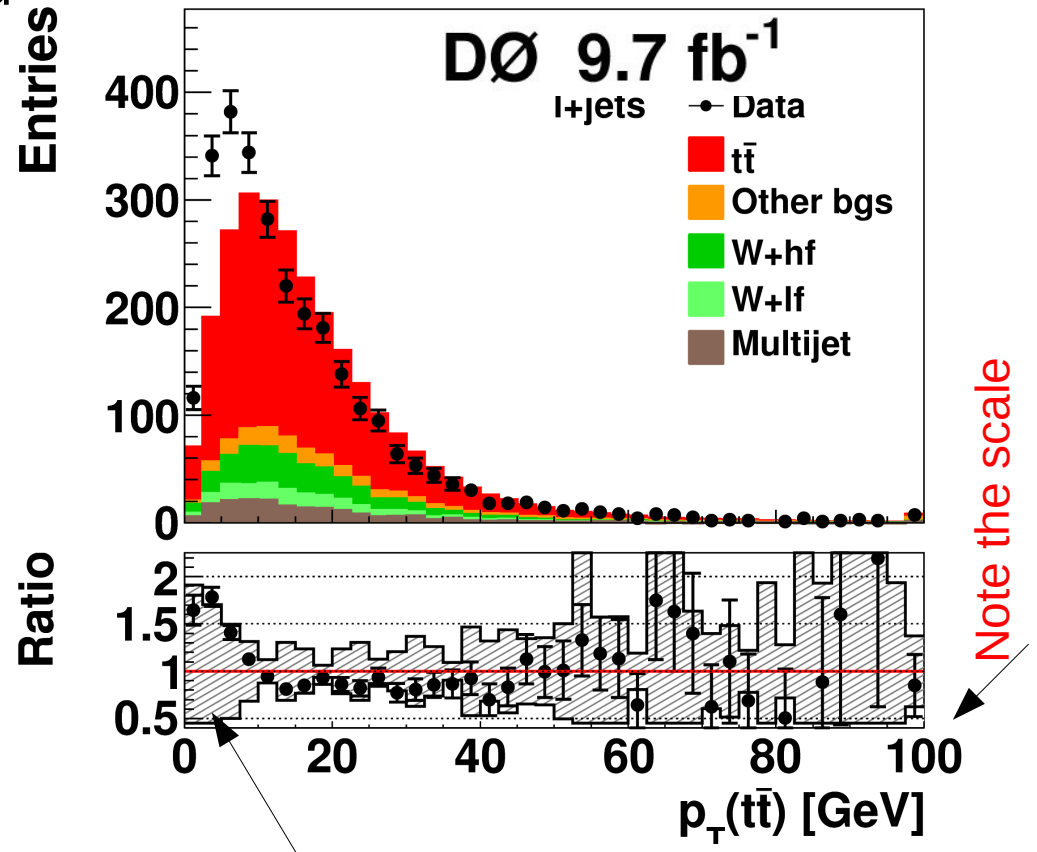




- Results in an uncertainty of: 0.06 GeV

Initial/Final state radiation

- The $p_T(t\bar{t})$ distribution is not modeled by MC:
  - Known issue...$p_T(t\bar{t})$ is closely tied to Initial/Final state radiation (see previous slide)
  - Assign a systematic uncertainty based on data-MC difference in $p_T(t\bar{t})$ → 0.07 GeV

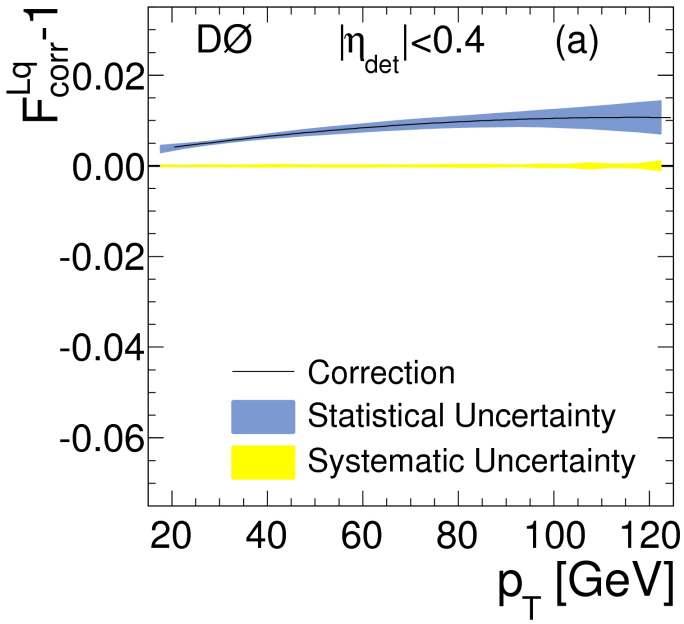


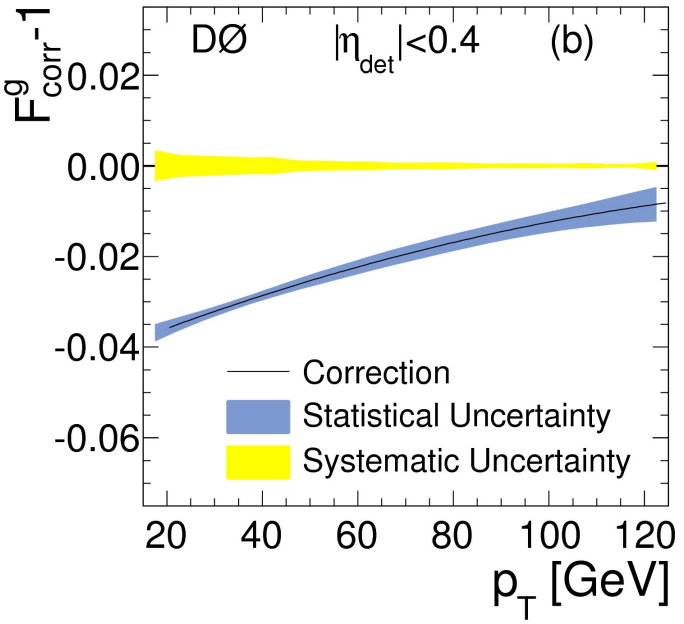
- Add in quadrature the effects from scale variations constrained by φ^* measurement and $p_T(t\bar{t})$ reweighting:

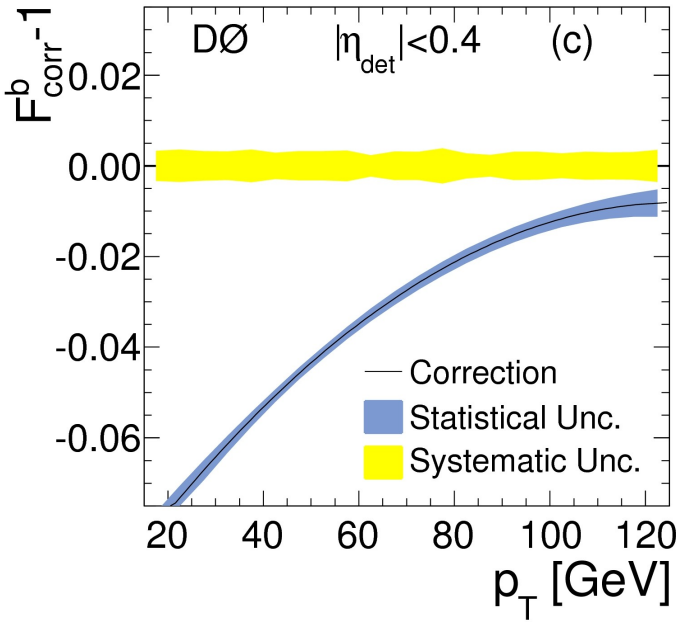
→ Initial/Final state radiation uncertainty:
0.09 GeV (previously 0.26 GeV)

Single Particle Response

- Improves MC description of jets and reduces sample-dependency
- Resulting F_{corr} for different jet flavors and their uncertainties







- Small correction to light quarks (*u,d,s,c*), several % for *g* and *b* quarks
- Without that correction, the measurement (see 1/fb result) suffers an uncertainty for the b/light response ratio of 0.83 GeV, by far dominant source Phys. Rev. Lett. 101, 182001 (2008)

Calibration

- Employ pseudo-experiments to calibrate the method in m_t and k_{JES}
  - Calibrate m_t by using m_t = 165, 170, 172.5, 175, 180 (k_{JES} =1) and in k_{JES} = 0.95, 1.00, 1.05
  - Employ measured signal fractions per run period
  - Calibrated quantities defined as:

$$m_t' = \frac{(m_t - 172.5\ \text{GeV}) - p_0^{m_t}}{p_1^{m_t}} + 172.5\ \text{GeV}$$

$$k_{\text{JES}}' = \frac{(k_{\text{JES}} - 1) - p_0^{k_{\text{JES}}}}{p_1^{k_{\text{JES}}}} + 1$$

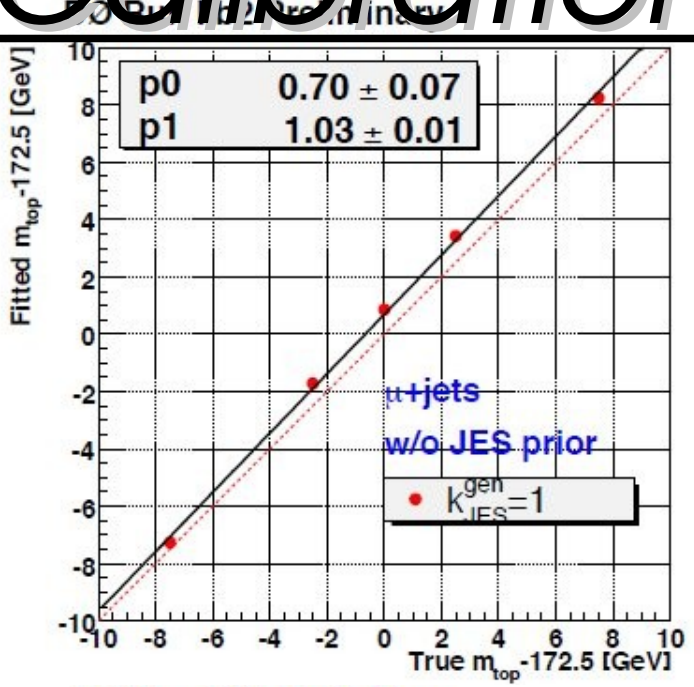
- Statistical uncertainties corrected by using the pull: $pull = \frac{\langle x \rangle - \bar{x}}{\sigma}$

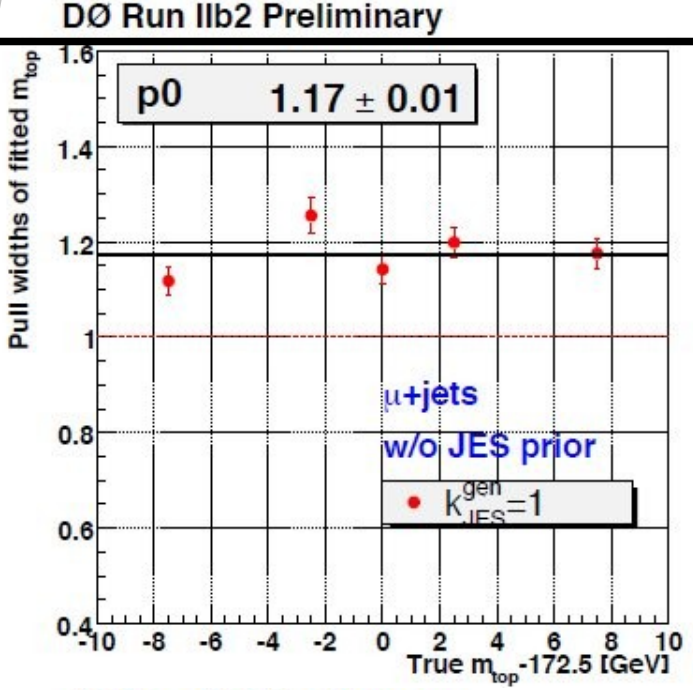
$$\sigma'(m_t') = \sigma(m_t') \times w_{\text{pull}}(m_t)$$

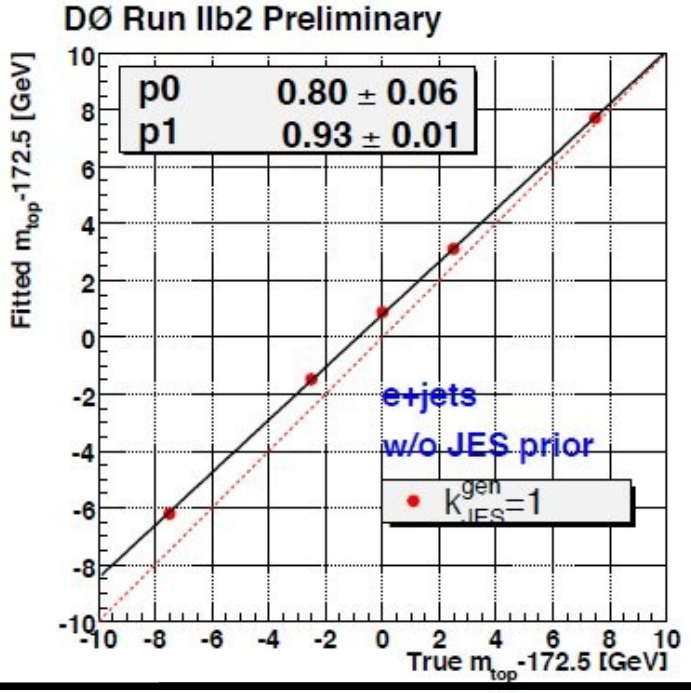
$$\sigma'(k_{\text{JES}}') = \sigma(k_{\text{JES}}') \times w_{\text{pull}}(k_{\text{JES}})$$

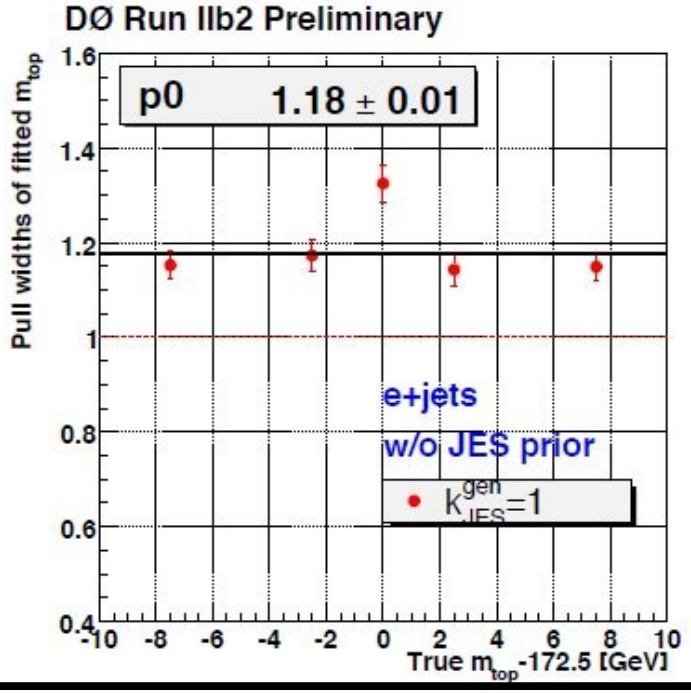
Width of the pull distribution

Calibration



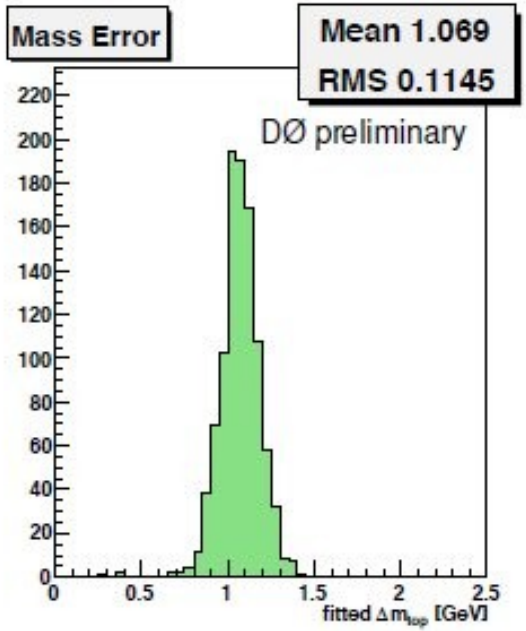


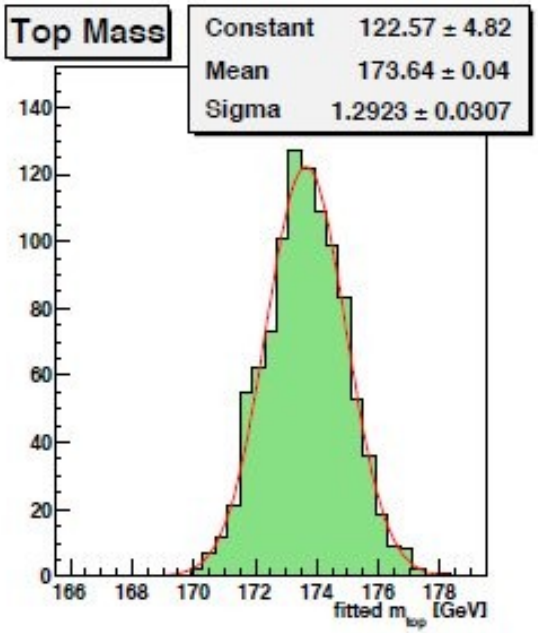


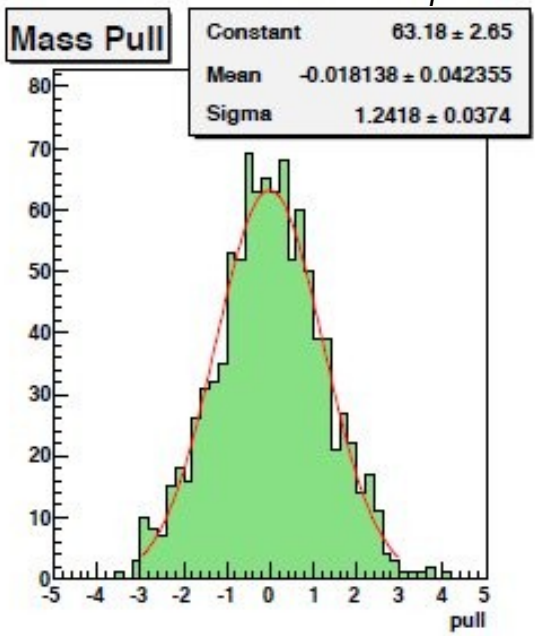


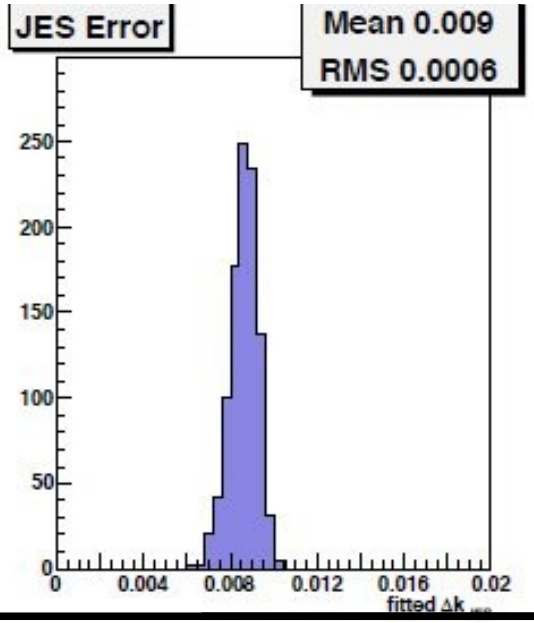
Calibration

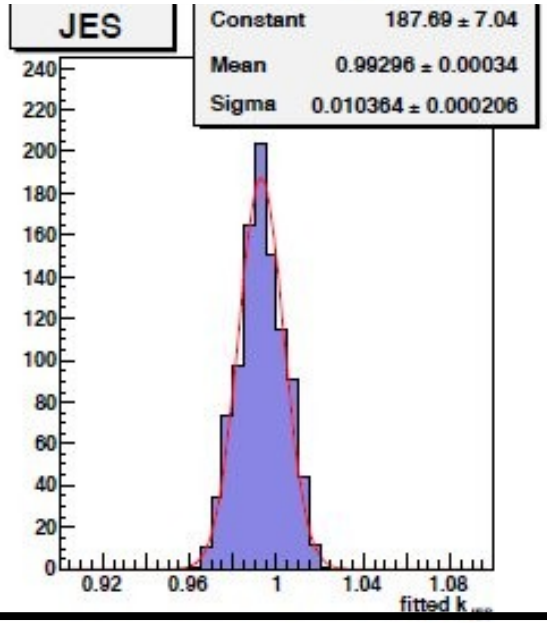
- Employ pseudo-experiments to calibrate the method in m_t and k_{JES}

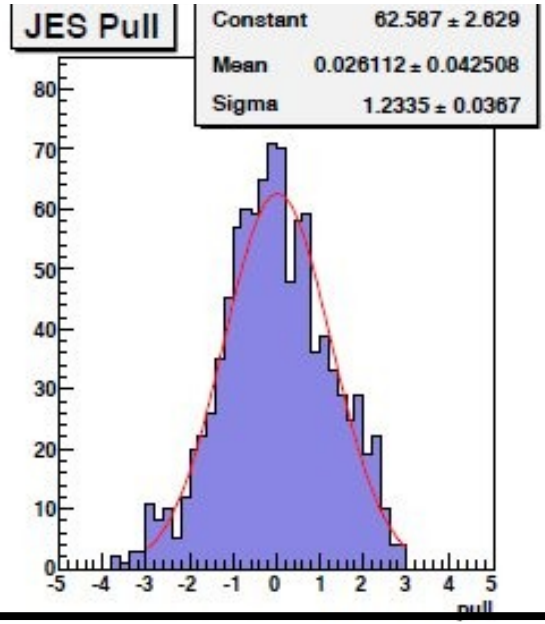








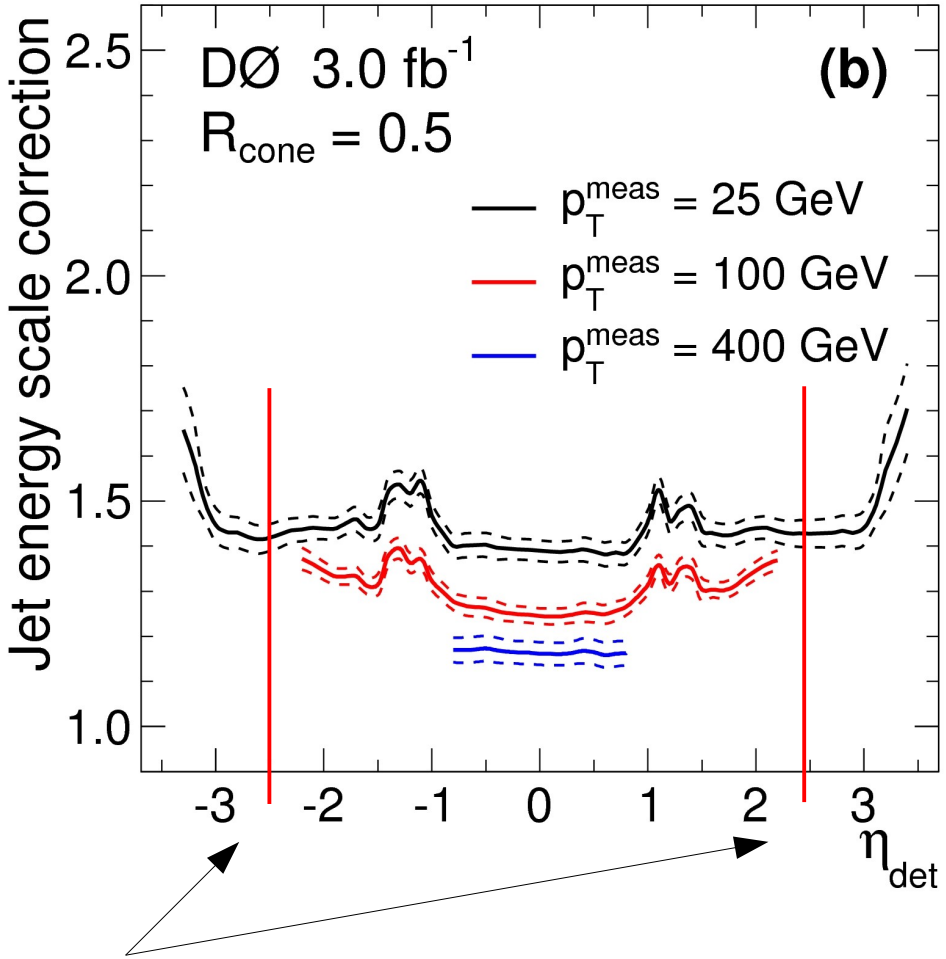




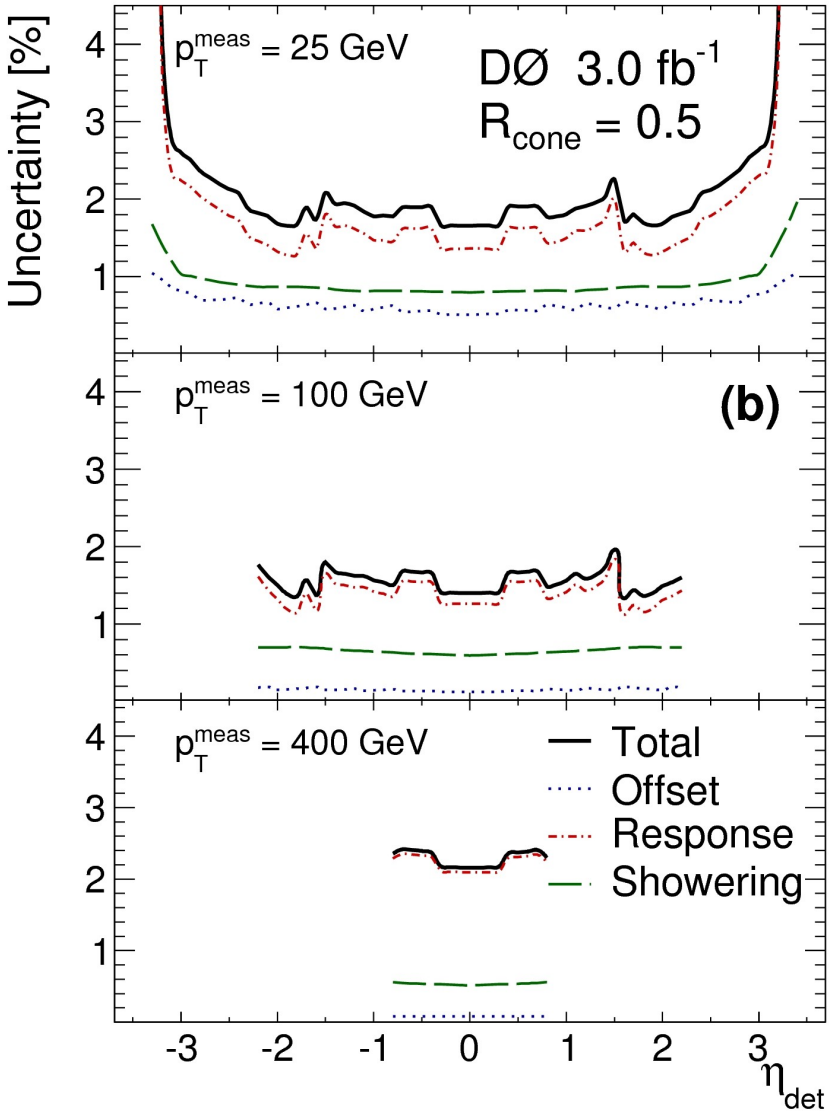
Jet Energy Scale (JES)

- Latest JES calibration applied
- Correction depends on jet energy, uncertainties of 1-2%

Accepted by NIM



Jet cuts required in top quark measurements



Single Particle Response

- JES does not distinguish quark and gluon jets
- Employ γ+jets events to calibrate response in MC to data
- Derive a correction factor F_{corr} for MC:

$$F_{\mathrm{corr}} = \frac{1}{\langle F \rangle_{\gamma+\mathrm{jet}}} \cdot \frac{\sum_i E_i \cdot R_i^{\mathrm{data}}}{\sum_i E_i \cdot R_i^{\mathrm{MC}}}$$

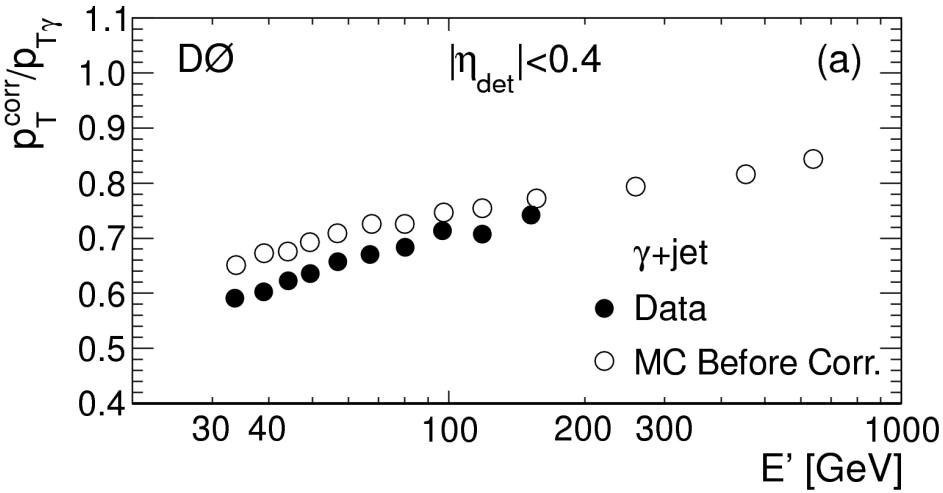
- Response in calorimeter for data or MC for particle *i* inside particle jet
- Matching with $dR < 0.25$

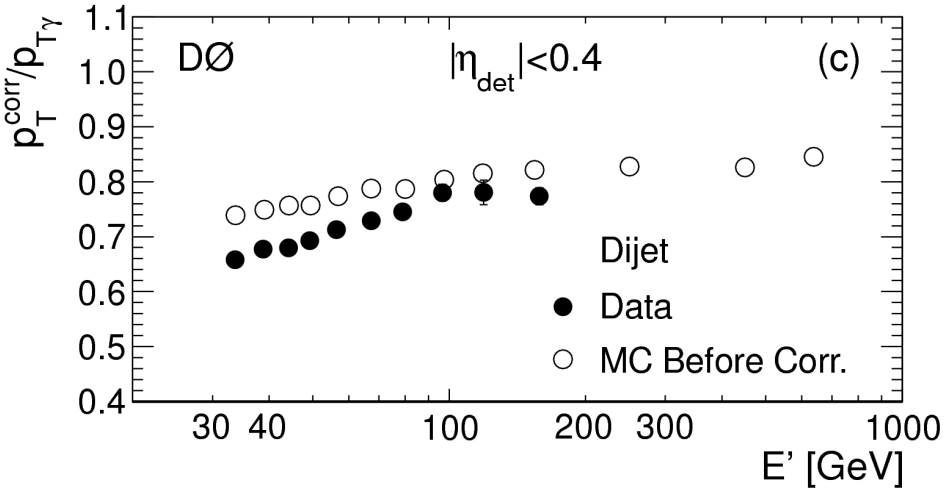
- Preserves standard JES calibration
- <u>Calibration by using:</u>
  - Reconstructed jet p_T with offset correction: p_T^{corr}
  - p_T of EM cluster with tight photon ID: p_T^{γ}
  - Calibrate ratio $p_T^{corr} / p_T^{\gamma}$

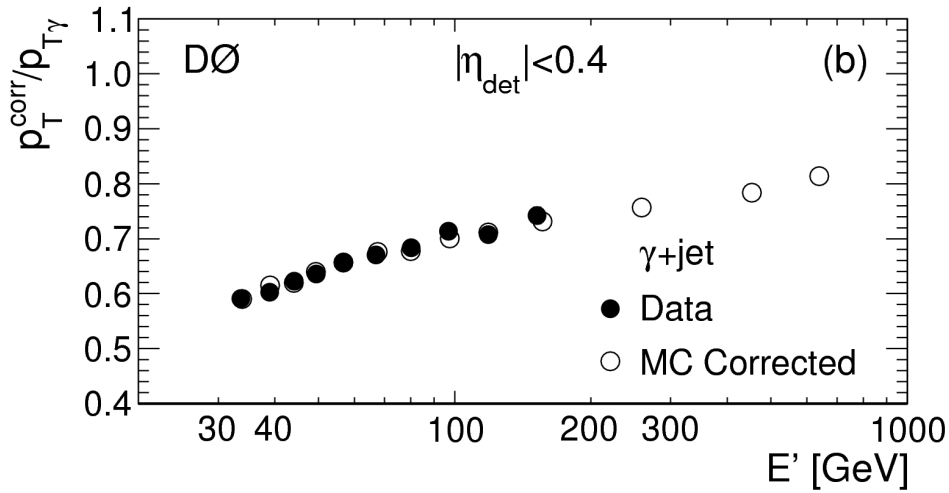
Assuming the single particle composition as in MC

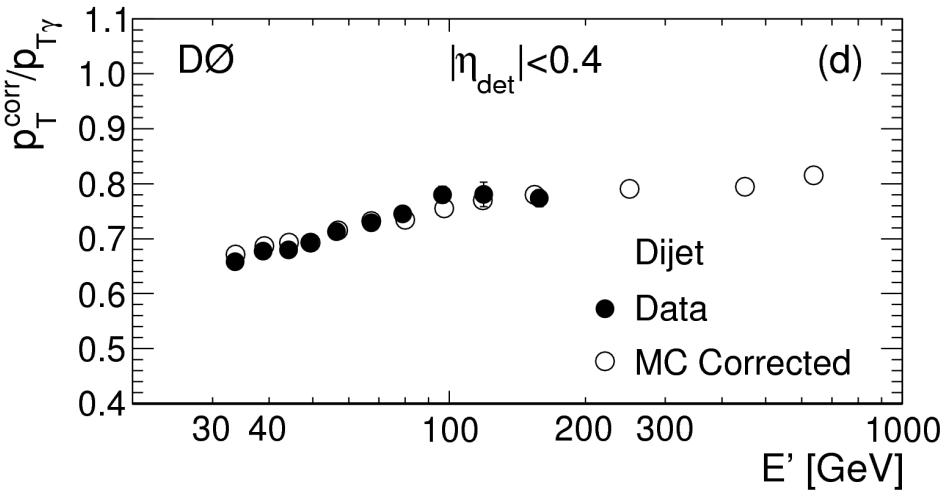
Single Particle Response

- Does not correct the response of gluon and quark jets to be the same
  → the response of jets in MC and data to be the same









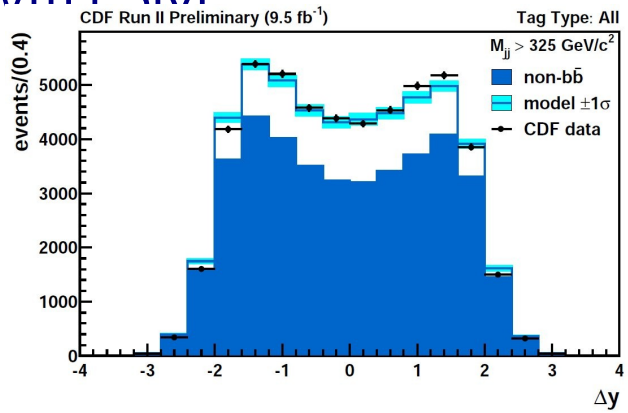
Other asymmetries...

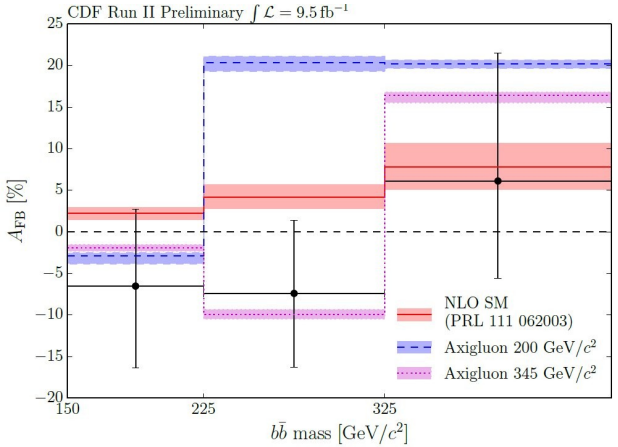
What about *b*-quark asymmetries ?

- CDF: Data side-band technique to estimate light quark asymmetry
  - Identify *b* and $\bar{b}$ employing jet charge
  - Bayesian technique to correct for detector effects
  - Results agree with SM

NEW

Conf. Note 11092

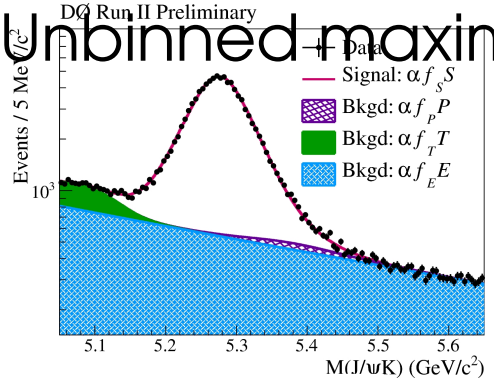




- D0 uses $B^{\pm} \rightarrow J/\psi\, K^{\pm}$ to probe asymmetry of *b*-quarks

D0 Note 6441-CONF

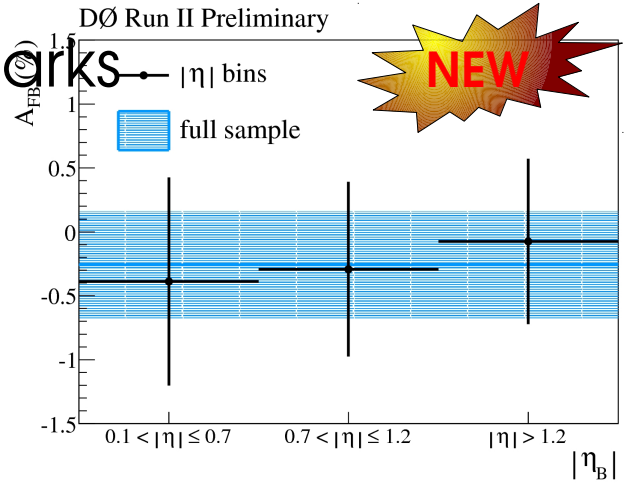
- Unbinned maximum likelihood fit



A_{FB} = -0.26 ± 0.41 ± 017 %

MC@NLO: 1.63 ± 0.43 ± x.xx %

Stat. To come: scales,pdf

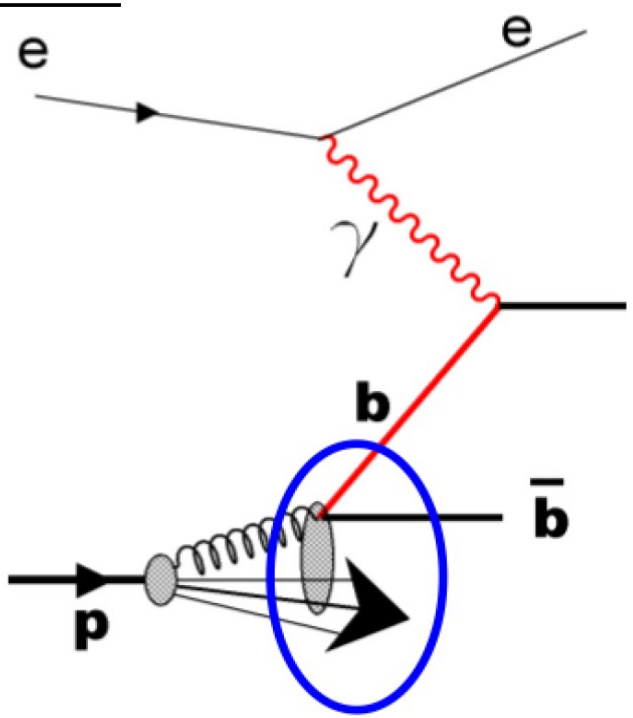


NEW

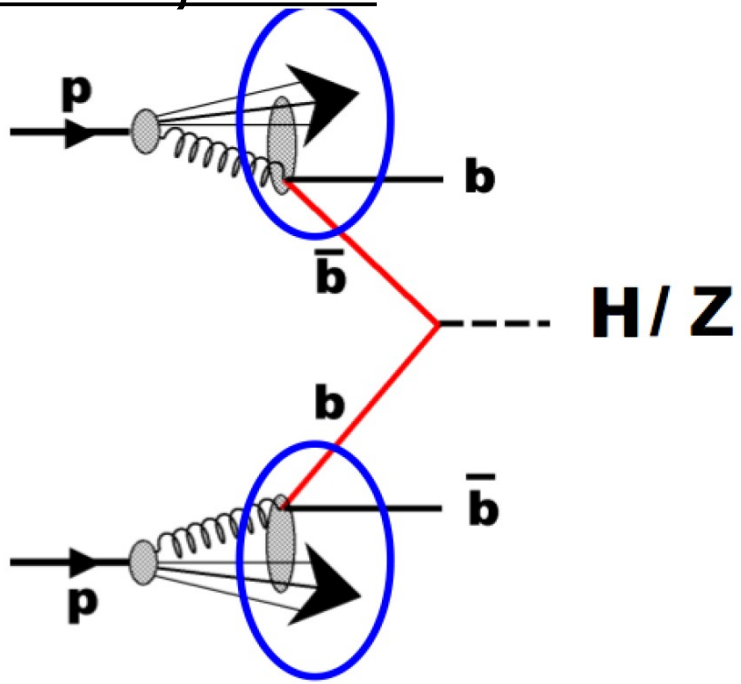
Factorization theorem

$$d\sigma = \sum_{i,j,k} f_i^A(x_1, \mu_f) \otimes f_j^B(x_2, \mu_f) \otimes d\hat{\sigma}_{ij \to kX}(\mu_f)$$

HERA:



Tevatron, LHC:



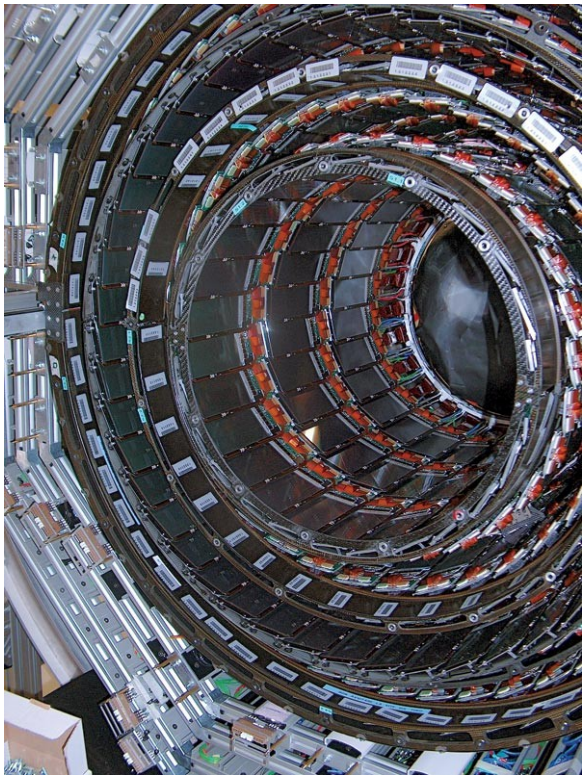
Three ingredients: **Parton distribution functions (PDFs)** and **matrix element**

Detectors – comparison

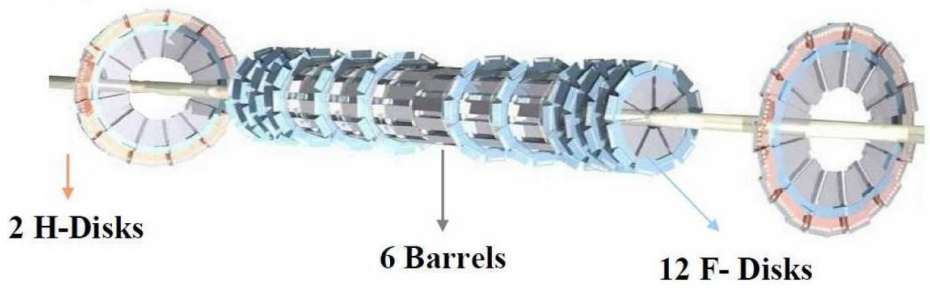
component	DØ	CMS
Inner Magnetic field [T]	1.2 (surrounds Tracker)	4 (surrounds Calo)
Silicon Tracker:		
Readout channels	800k	76M
Area [m^2]	3	205
Weight [t]	5000	14000
Dimension [m]	12×20	15×28.7

- Huge difference in all aspects of the silicon detector, e.g. N° channels and area
- Consequences for object IDs using silicon detector, e.g. *b*-tagging

CMS silicon tracker:



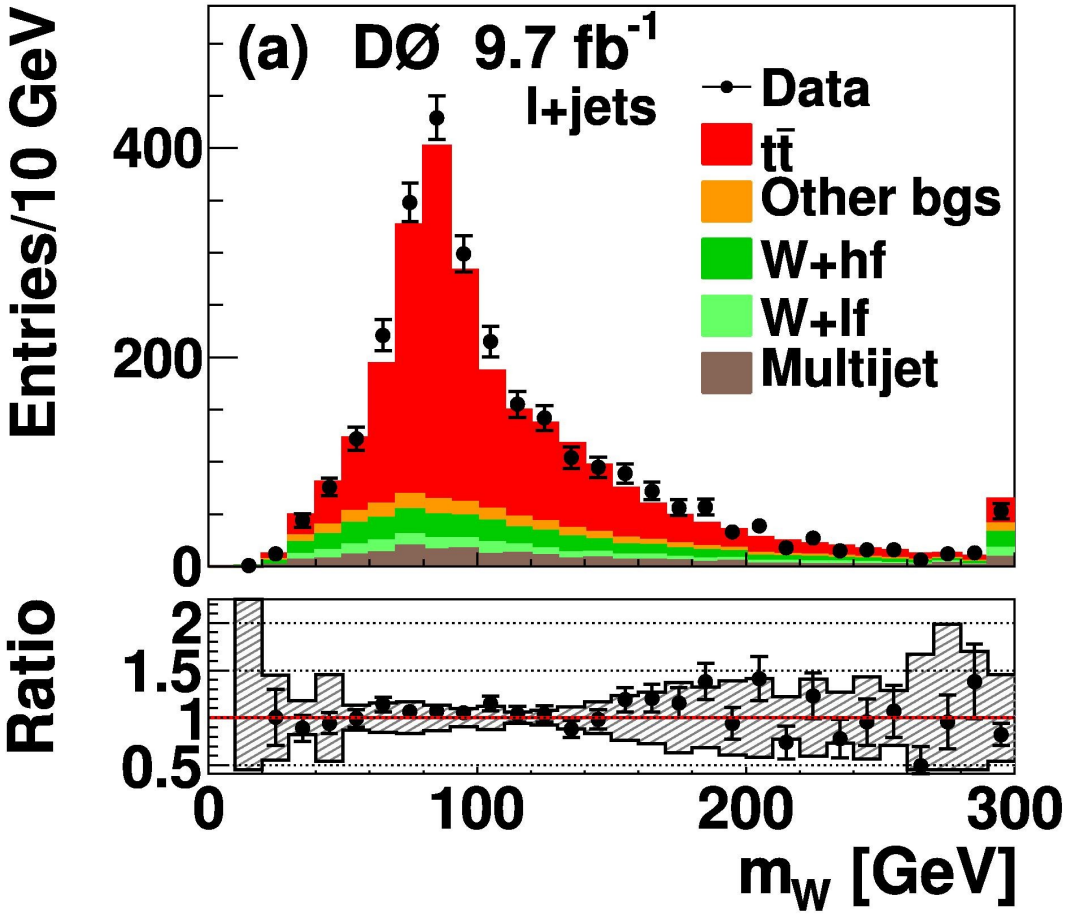
D0 silicon tracker:

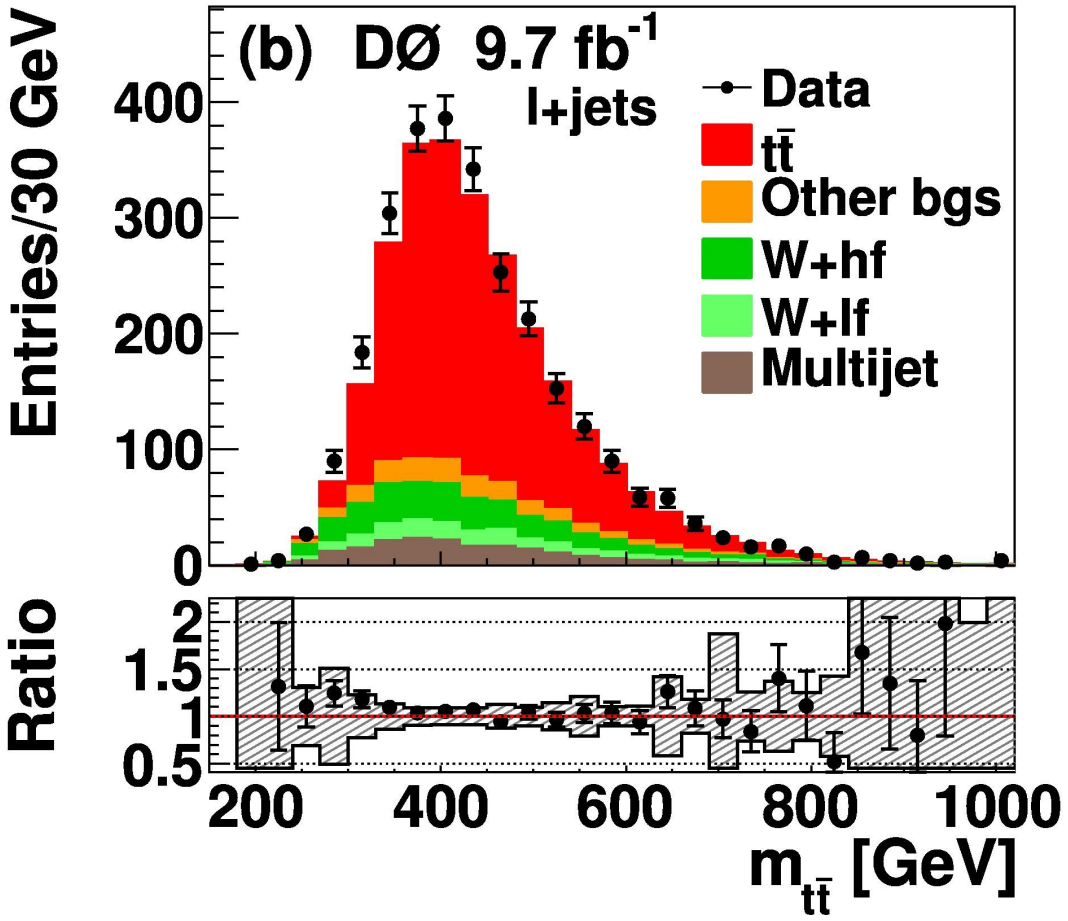


NIM A 634:8-46 (2011)

Control distributions

- Apply ME results to control distributions in m_W and $m_{t\bar{t}}$
  - MC (Alpgen+Pythia) with m_t=175 GeV
  - $\sigma_{t\bar{t}}$ as on previous slide



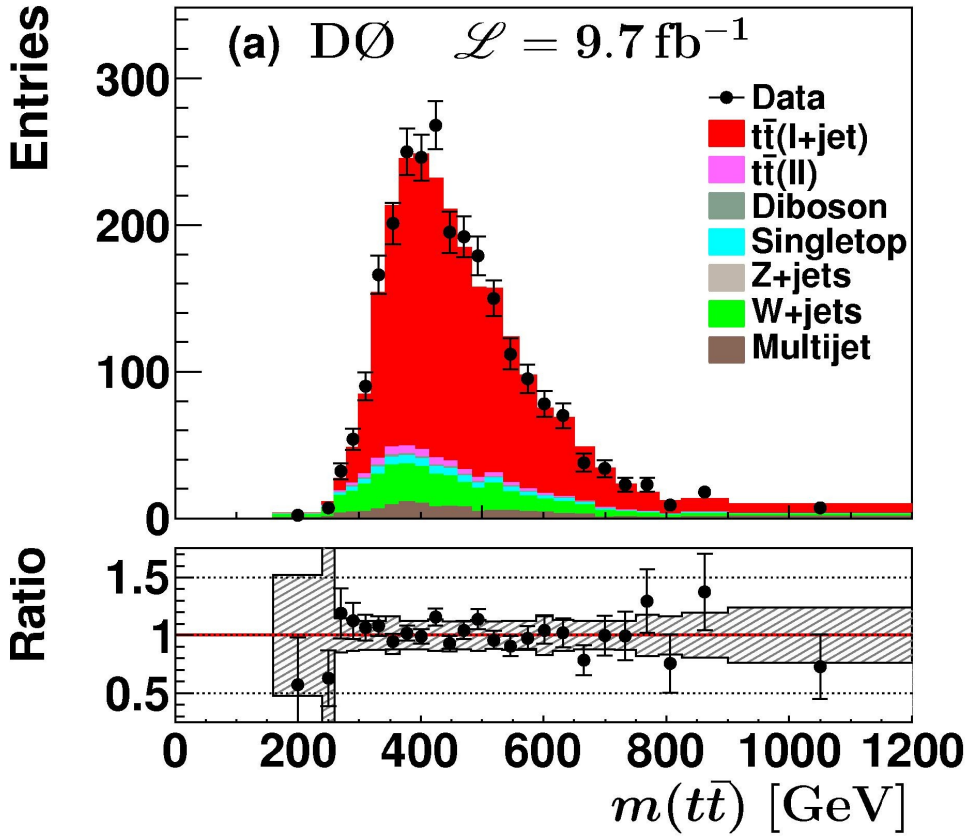


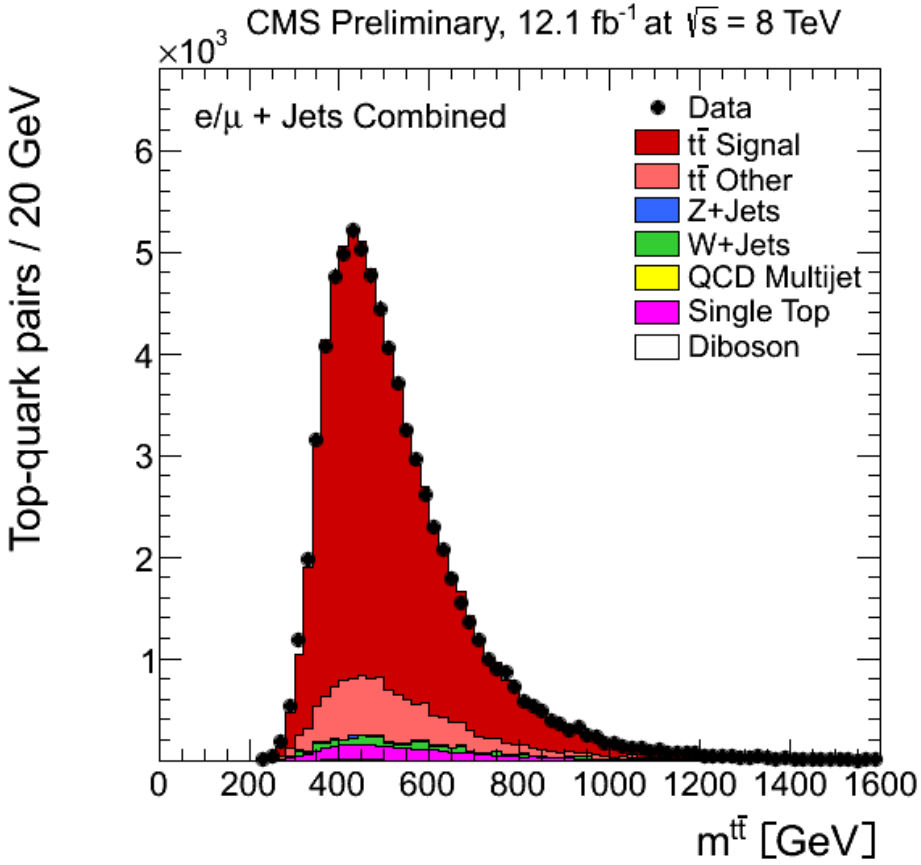
- Nicely modeled by MC, improved description compared to pre-ME

Top quark reconstruction

- Kinematic top quark reconstruction: Assign observed particles to true decay particles of the top quark
- As an example invariant mass of the $t\bar{t}$ pair

e or μ + ≥ 4 jets

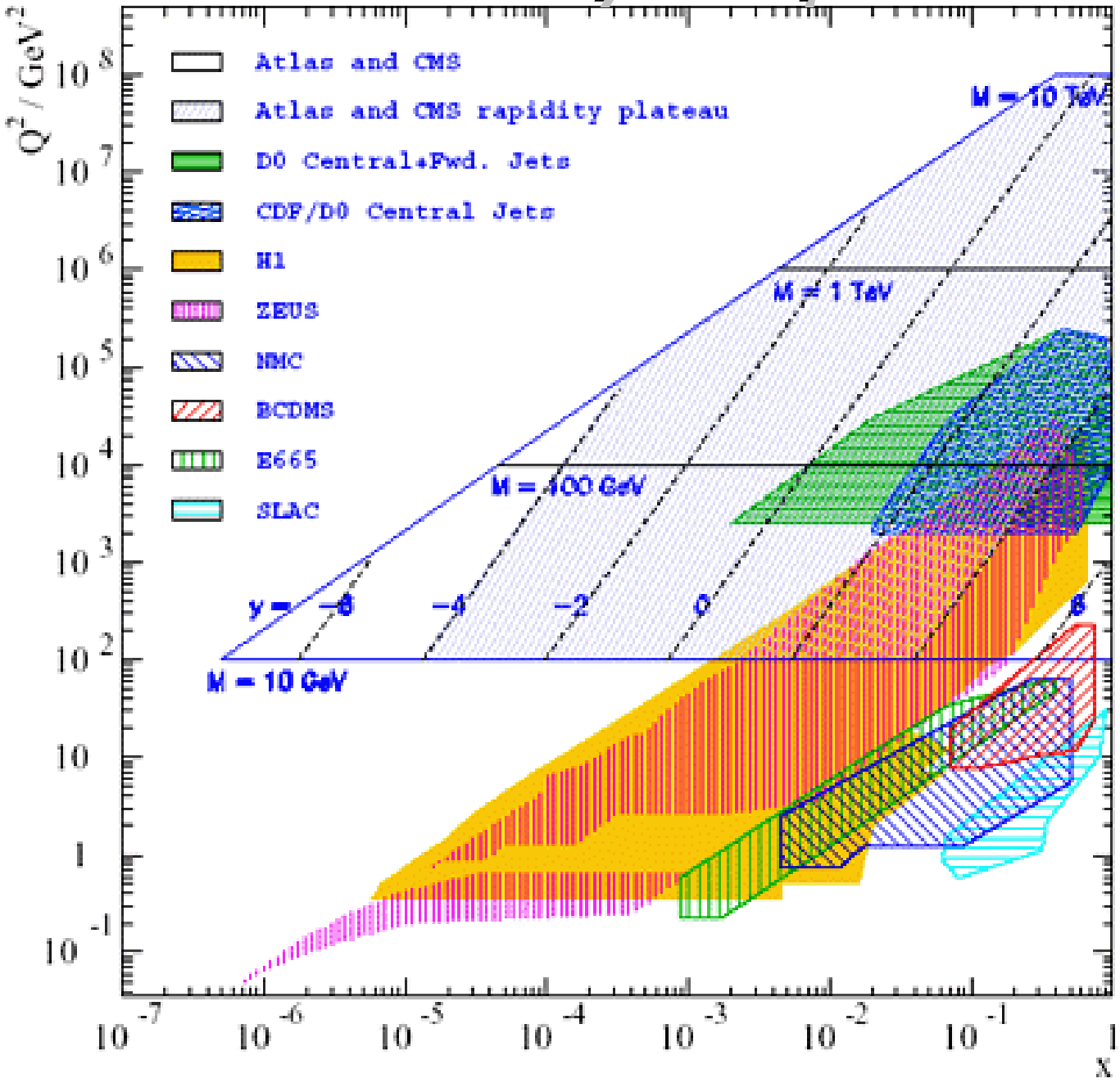




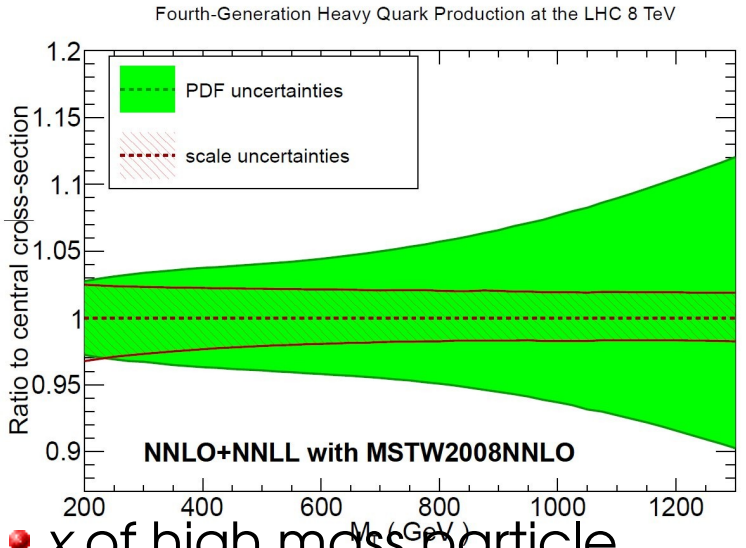
- Reconstructed top quark distributions well modeled by MC

(assuming the measured $t\bar{t}$ cross section)

Why top data



- Uncertainties at high masses dominated by PDF:



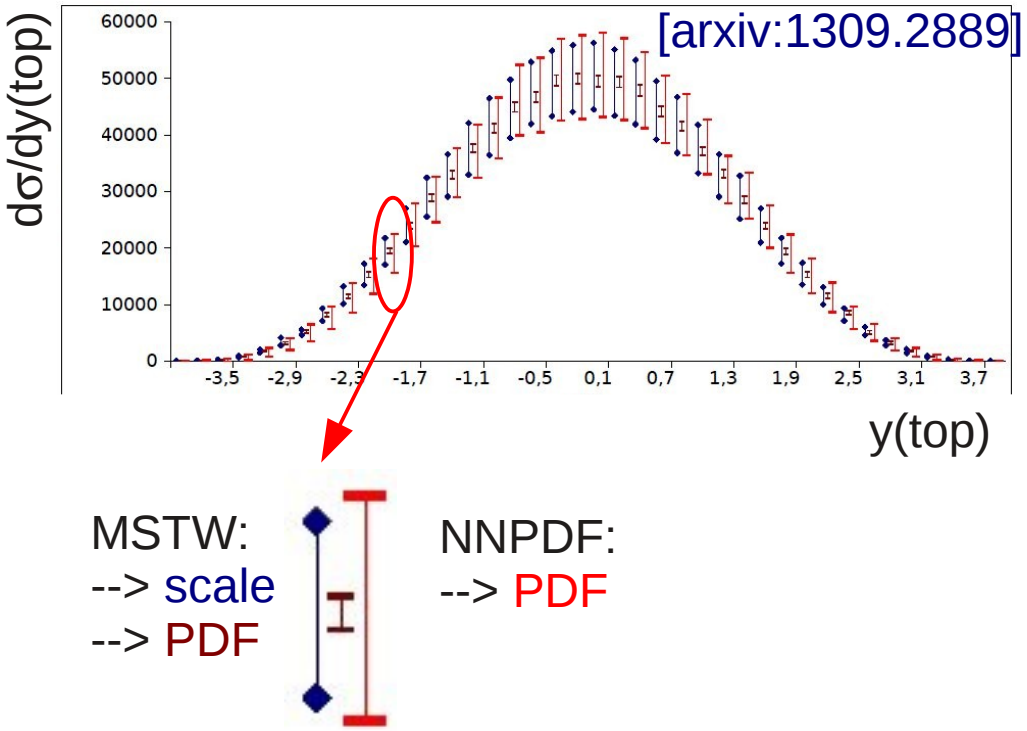
- *x* of high mass particle production ?

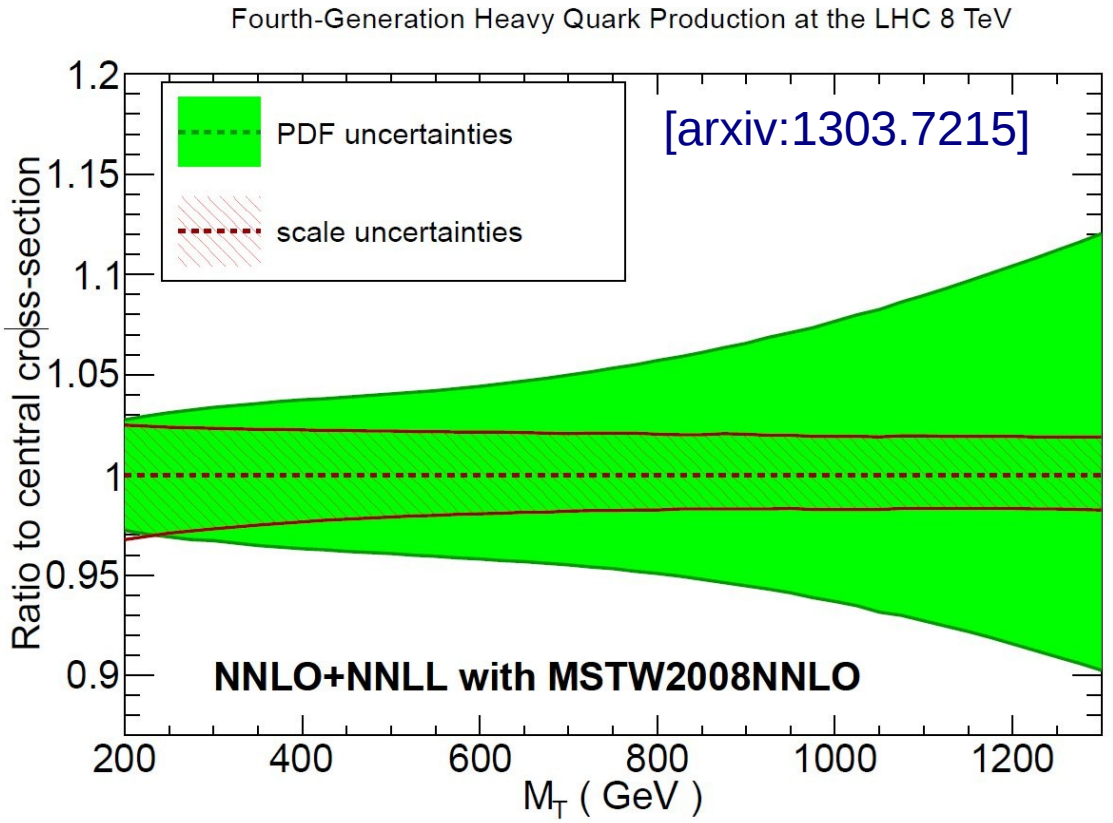
$$\hat{s} = x_1 x_2 \cdot s$$

Motivation

- Differential cross section distributions **test QCD calculations** and **enhance our understanding** of top quark production:
  - Inputs for estimating **Parton Distribution Functions (PDF)**, or the **Proton structure**

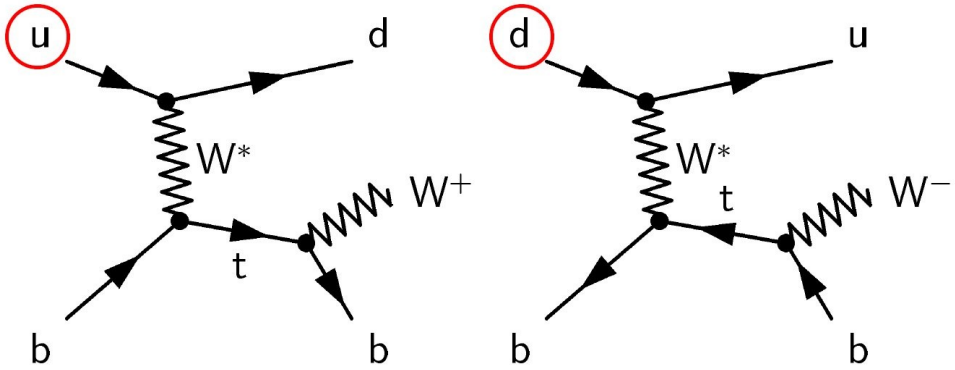
NLO calculation at 14 TeV

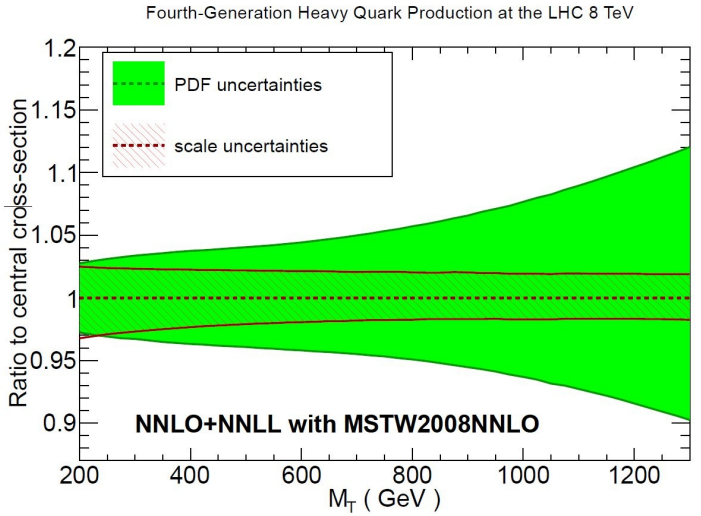


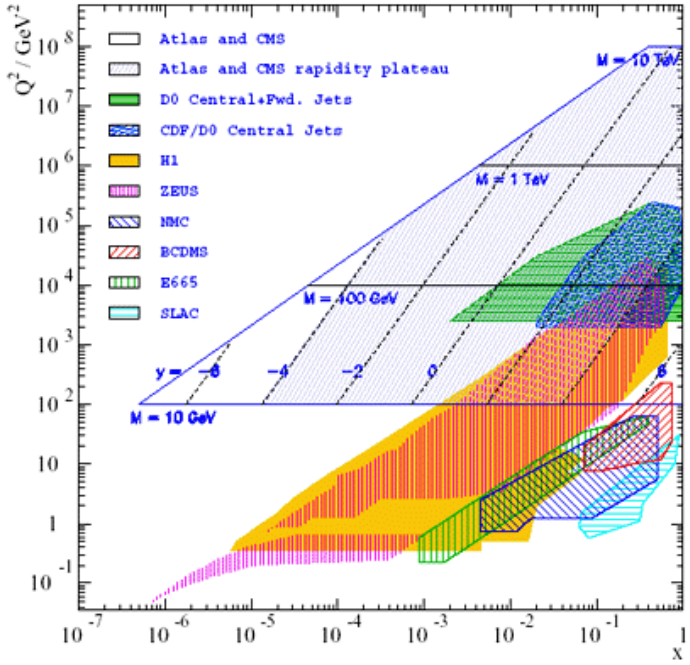


Why top data

- Need data of production of (fairly) heavy particles with sufficient precision: → **Top quarks**
- In future: single top t-channel sensitive to *u*/*d* ratio



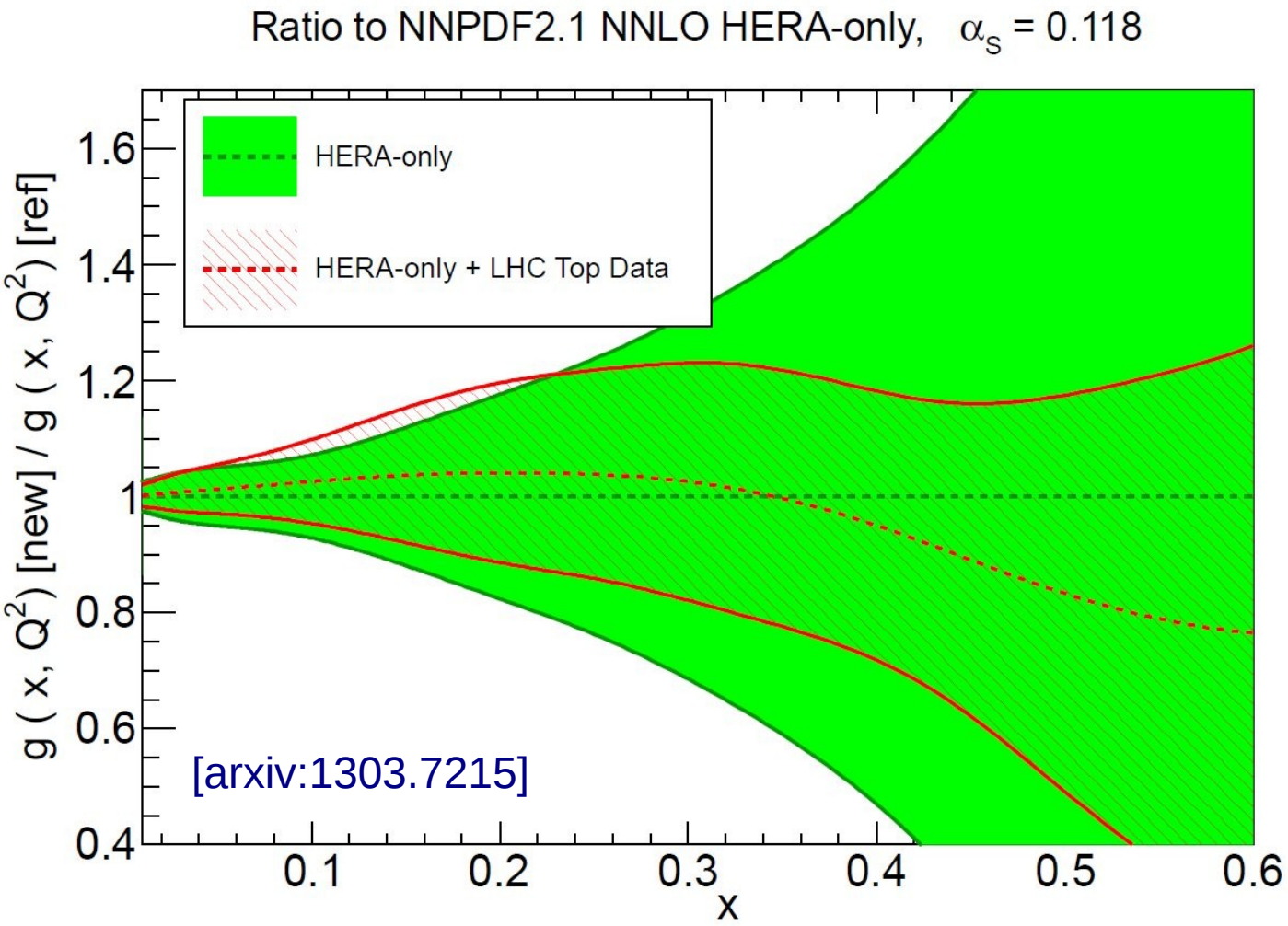




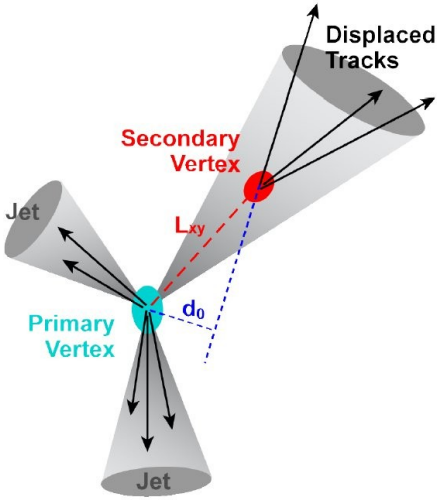
Top cross section for PDF

Global analysis by the NNPDF group:

- Compares HERA only versus HERA + Top data, Large impact at large x
- **Significant reduction** of uncertainties for the gluon content of the proton

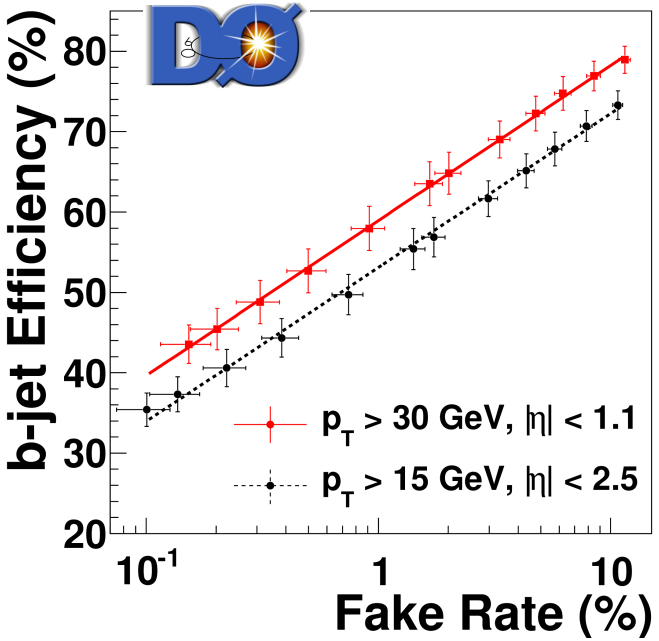


"b-tagging" technique

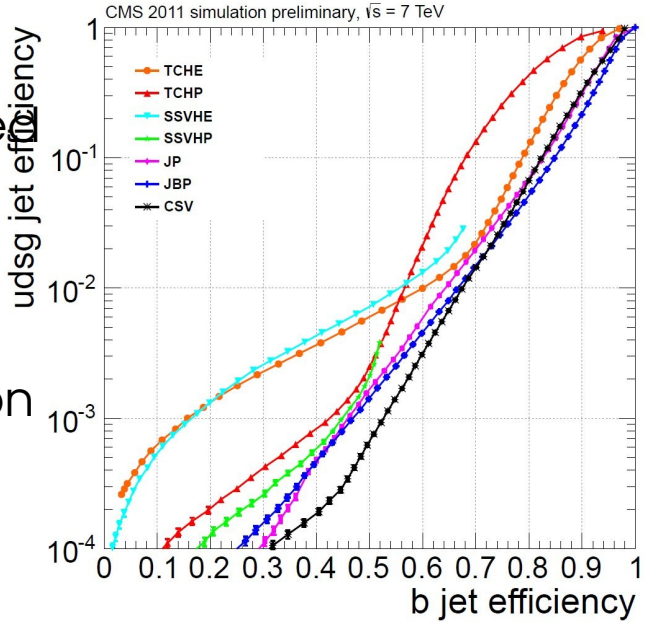


- *b* quarks hadronize before decaying into a *c* quark:
  - → Long-lived *b* hadrons decay some mm away, < 10 mm
  - → Identified using Multivariate Analysis technique

Subm. to NIM [arxiv:1312.7623]



- Similar strategy in both experiments:
  - → Impact parameter (secondary decay vertices) of charged particle tracks one of the most important quantity
  - → Many different algorithms used at CMS
  - → D0 merged all algorithms into one MVA
- Overall CMS *b*-tagging efficiencies and light jet misidentification rates improved a lot compared to D0

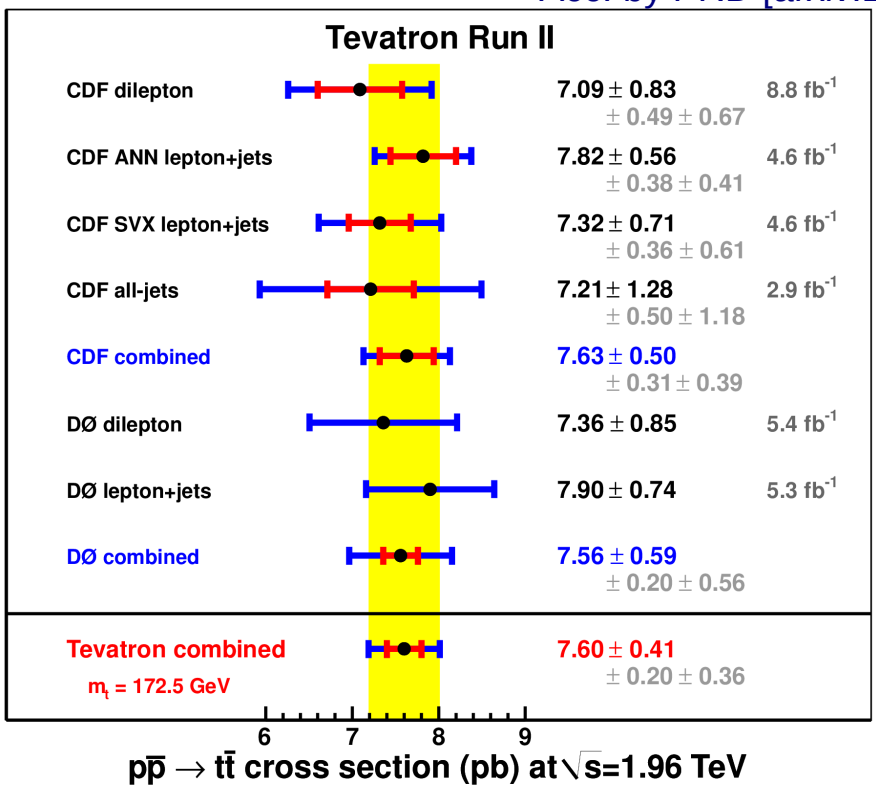


Top quark production

Tevatron: Combination

- CDF inputs: 4 measurements, < 8.8 fb^{-1}
- D0 inputs: 2 measurements, < 5.4 fb^{-1}
- Correlations taken into account

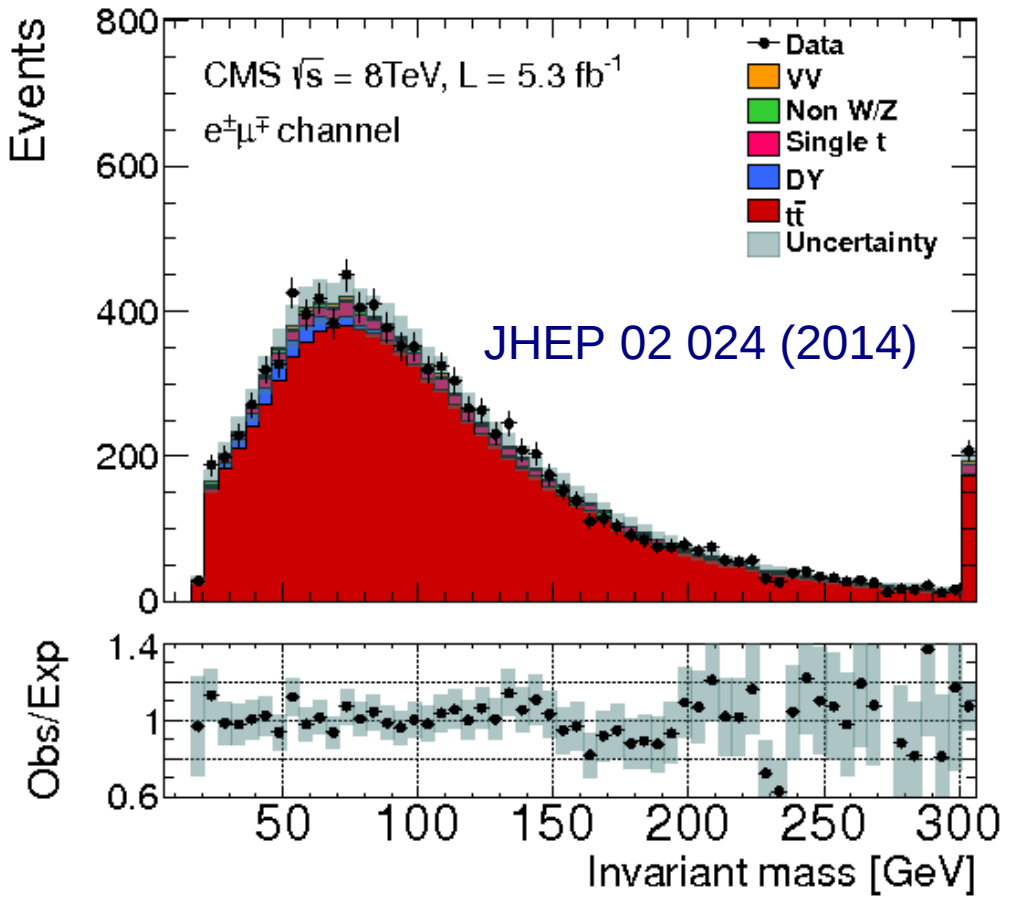
Acc. by PRD [arxiv:1309.7570]



5.4% precision in combination

LHC: single experiment, dilepton

- extremely clean, very high purity



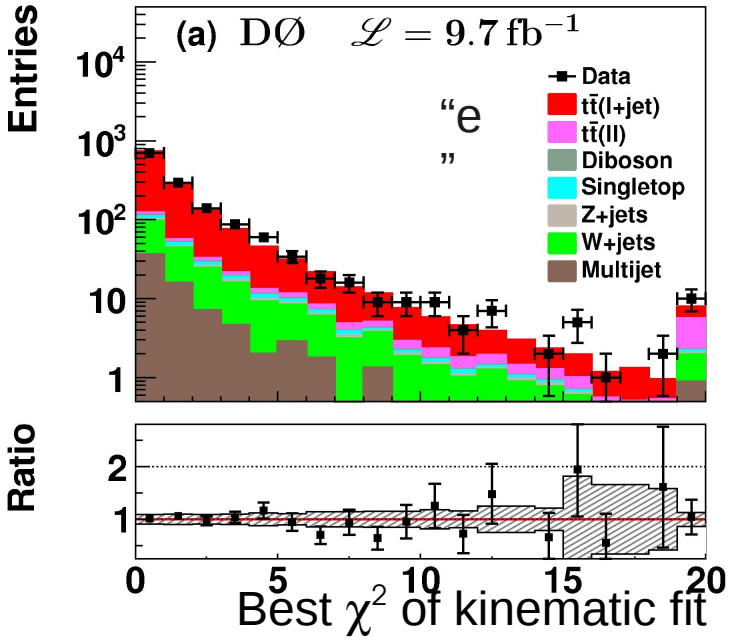
$\sigma(t\bar{t})$ = (239 ± 2 (stat) ± 11 (syst) ± 6 (lumi))pb ~5.4%
(one measurement)

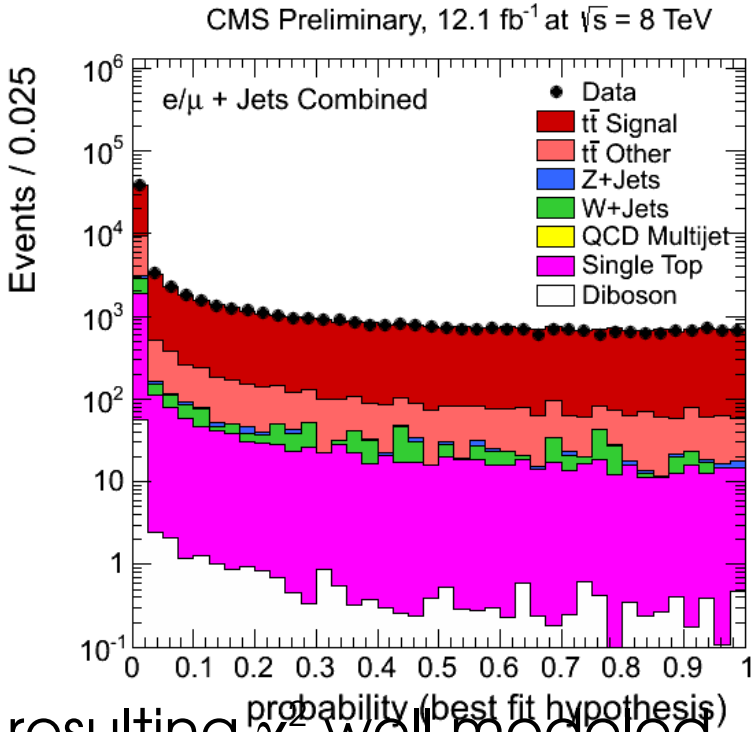
Top quark reconstruction

Associate leptons and jets with individual top quarks by a constrained kinematic fit:

- D0 constrains m_t and m_W, whereas CMS only constrains m_W
- Experimental resolution taken into account & assign *b*-tagged jets to b quarks

Values:
m_t=172.5 GeV
m_W=80.4 GeV





- Choose best solutions, distribution of resulting χ^2 well modeled

Regularized Unfolding

- Bin-by-bin correction method (common) in HEP:
  - Possibly strong model dependencies
  - Errors in general too optimistic

$$\frac{d\sigma}{dX_i} = \frac{N_i^{\text{signal}}}{\epsilon \cdot \mathcal{L} \cdot \mathcal{BR} \cdot \Delta X_i}$$

- Matrix Unfolding introduces (ideal case):
  - No bias with respect to a particular model of the physical process and MC
  - No (small) bias, with respect to general requirements of solution (smoothness)

  In Reality: → Migrations are treated more properly compared to Bin-by-bin, but:

  → Result tends to depend on the MC used, model uncertainty is reduced compared to Bin-by-bin but not zero

  → Regularized unfolding needs to be checked very carefully

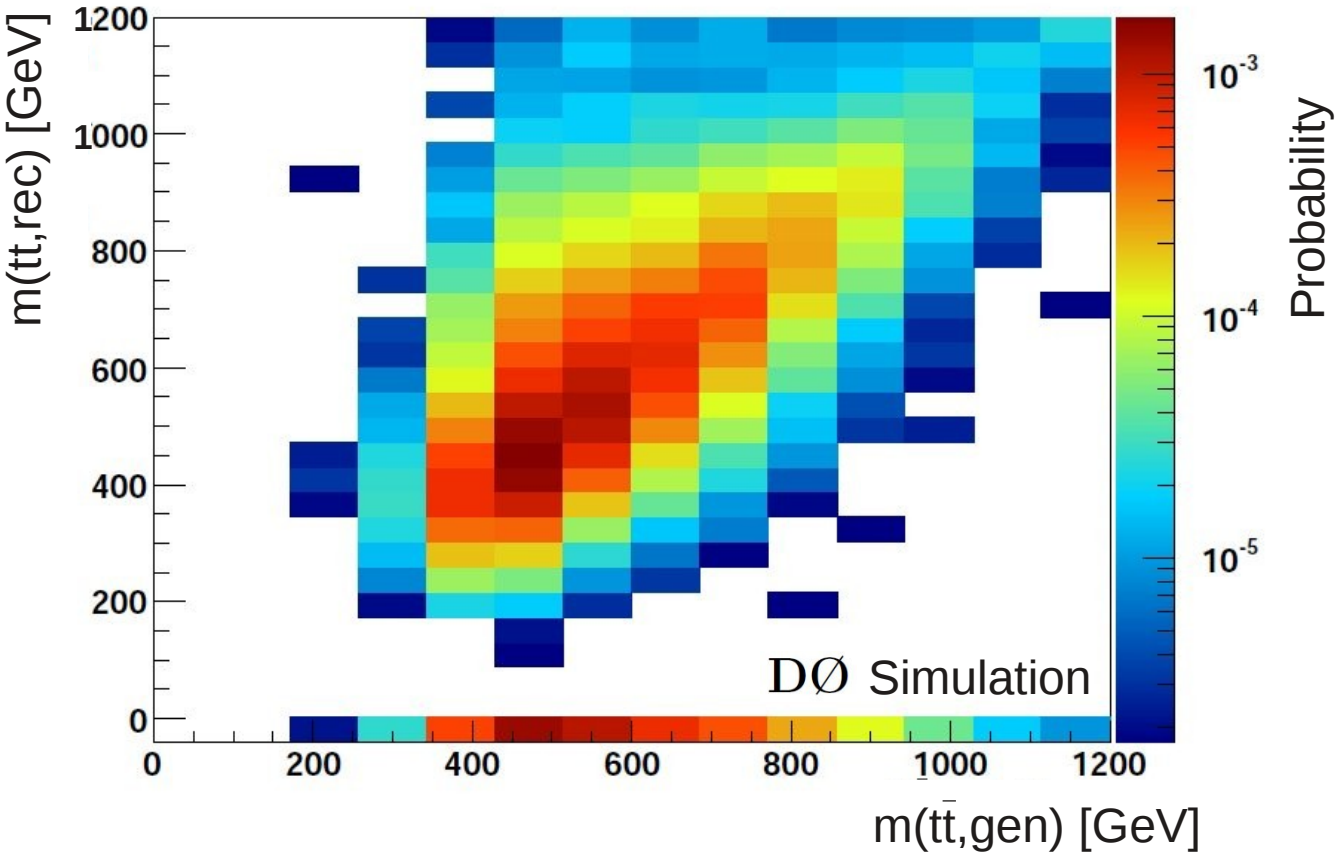
- For a discrete measurement with *n* bins:

$$\boldsymbol{Ax = y \;\rightarrow\; x = A^{-1}y} \qquad \boldsymbol{V_x = A^{-1}V_y(A^{-1})^T}$$

$\boldsymbol{x}$	= *n*-histogram of true variable *x*	$\boldsymbol{V_x}$	= error propagation for true variable *x*
$\boldsymbol{y}$	= *m*-histogram of measured variable *y*	$\boldsymbol{V_y}$	= error of measured variable *y*
$\boldsymbol{A}$	= $m \times n$ response matrix		

Correcting the data...

- As an example: D0 case
- The migration matrix is <u>the</u> crucial ingredient for unfolding:

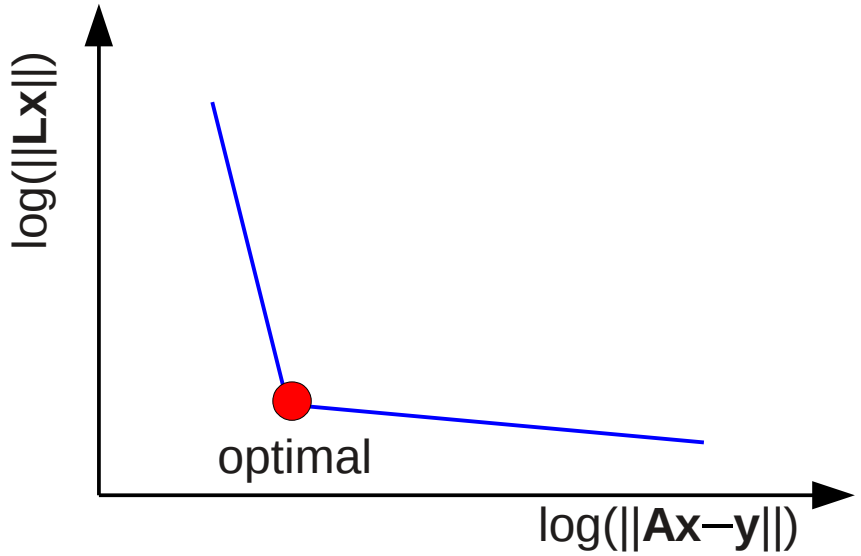


$$\frac{d\sigma_i}{dX} = \frac{N_i^{\text{unfold}}}{\mathscr{L} \cdot B \cdot \Delta X_i}$$

- Use twice as many bins on 'reconstructed' level as on 'generated' level
  → provides more information to the unfolding process
- Use a generator row (1st row) to correct for detector efficiency

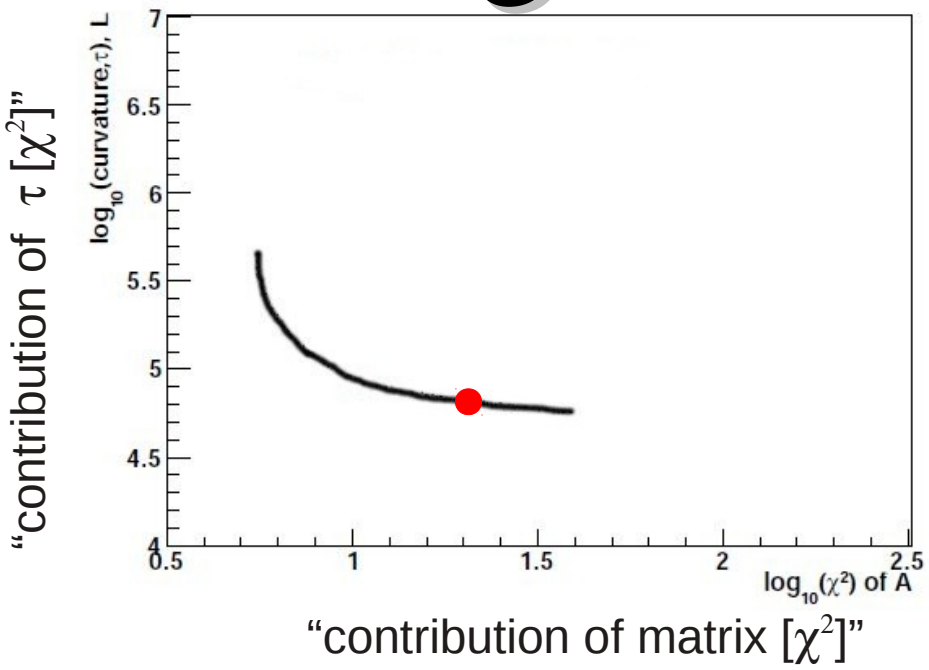
Regularized Unfolding

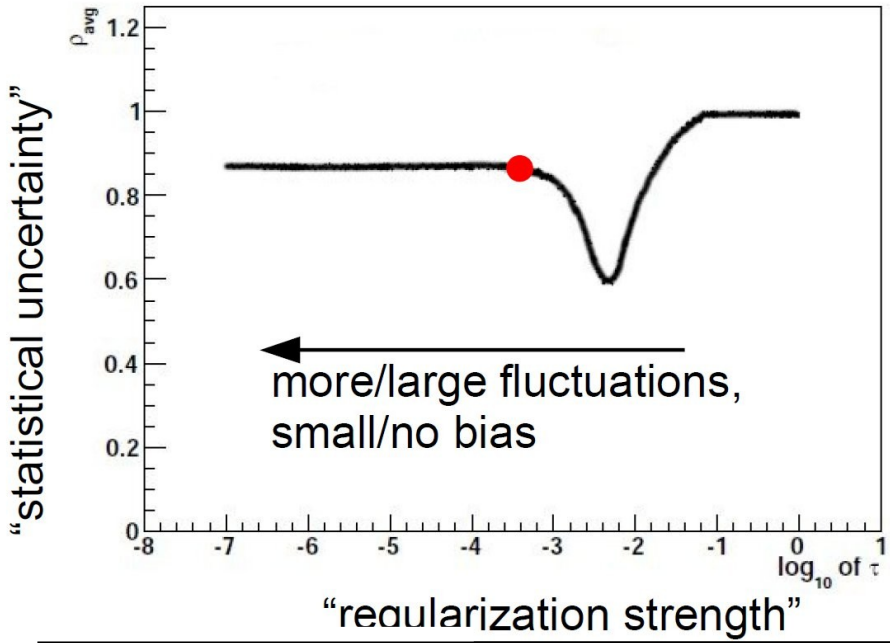
- Bin size ≈ resolution → large fluctuations from contributions of non-significant bins
- Use regularization (or if possible: larger bins):
  - $F(x) = ||\boldsymbol{Ax} - \boldsymbol{y}||^2 + \tau\,||\boldsymbol{Lx}||^2 = min \qquad \tau > 0$
  - '*L*' Condition based on: size, curvature, etc.
    they act on (y-x_0) where x_0 is taken from the migration matrix
- Optimal value of τ:
  - L–curve method



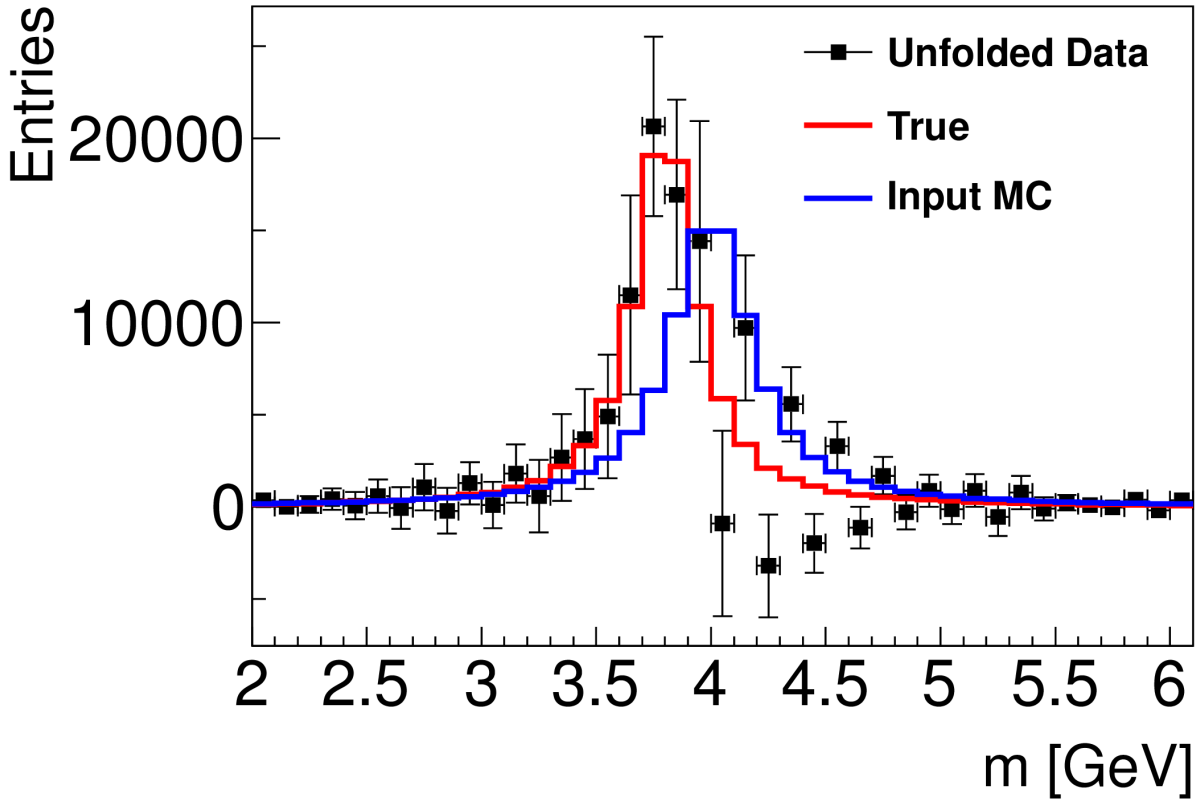
→ Several conditions for the regularization possible – choice is non-trivial & depends on the problem/analysis

Regularized Unfolding



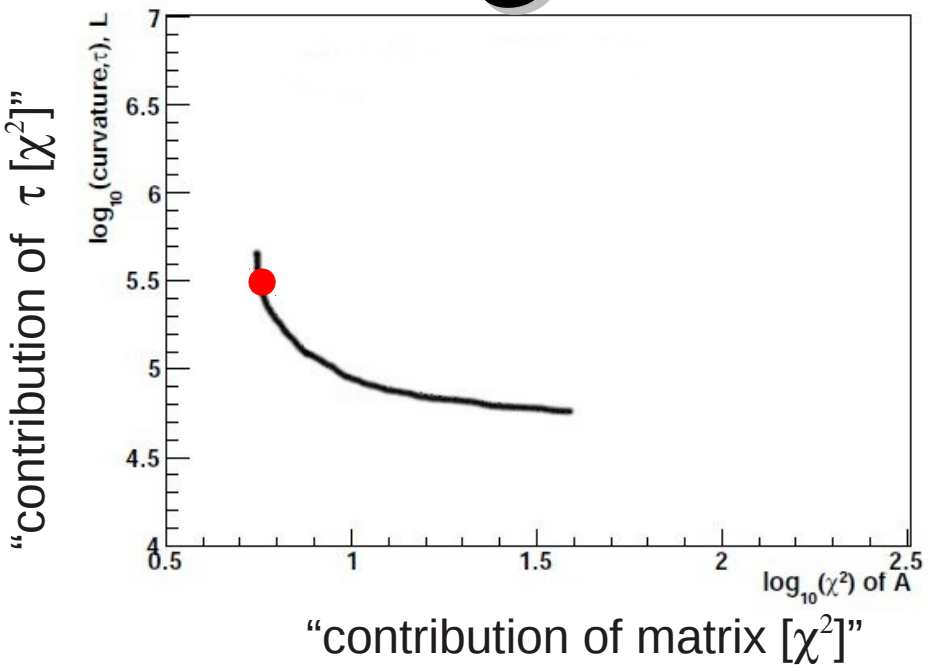


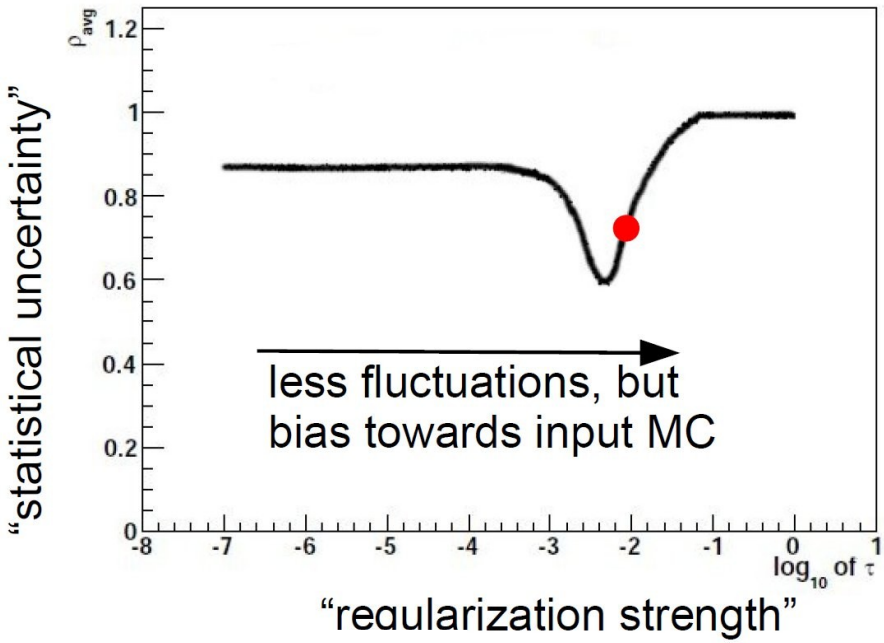
Toy model example: → Unfold smeared distribution to "true"

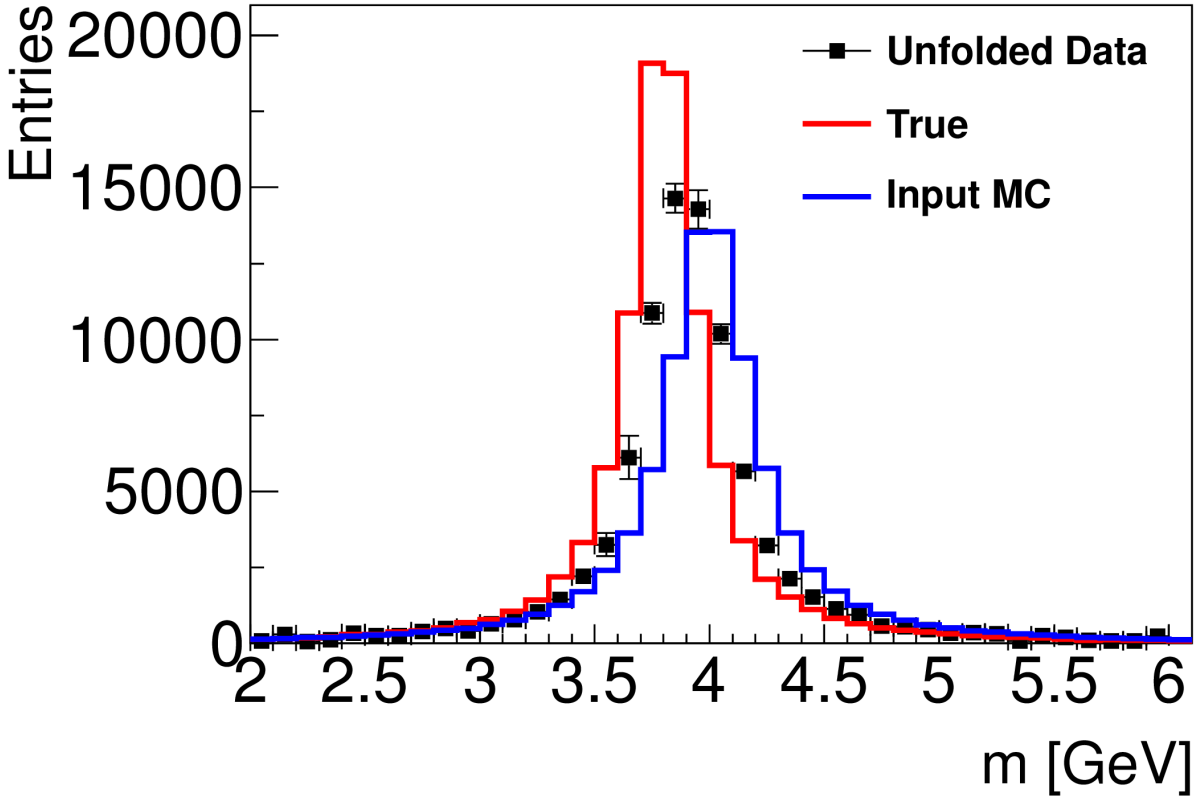


Regularized Unfolding

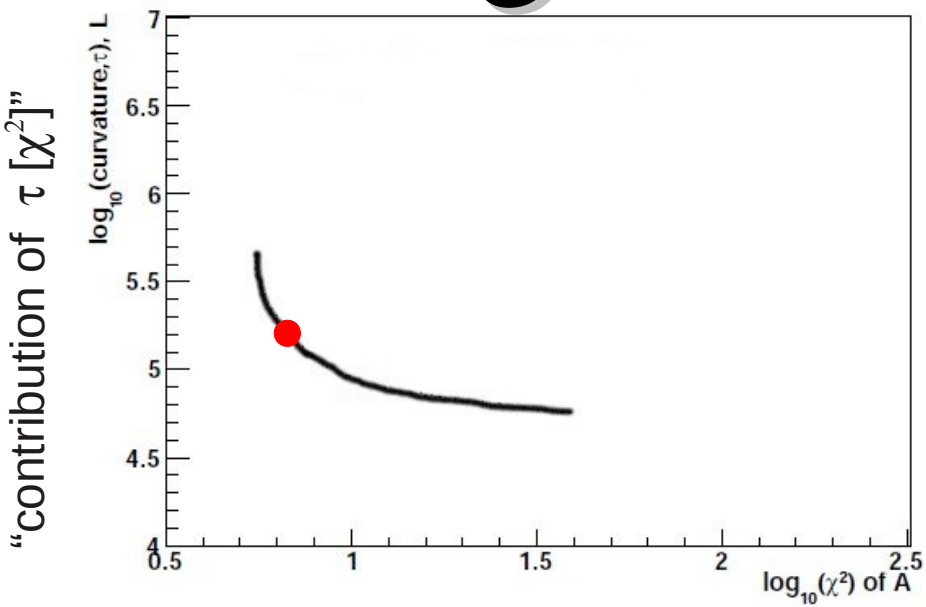
Toy model example: → Unfold smeared distribution to "true"

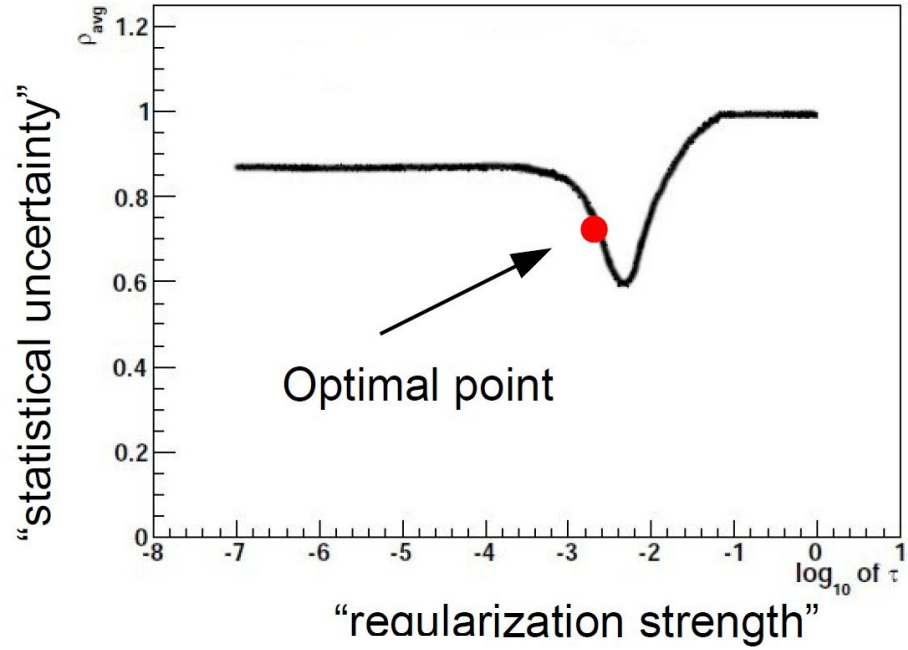




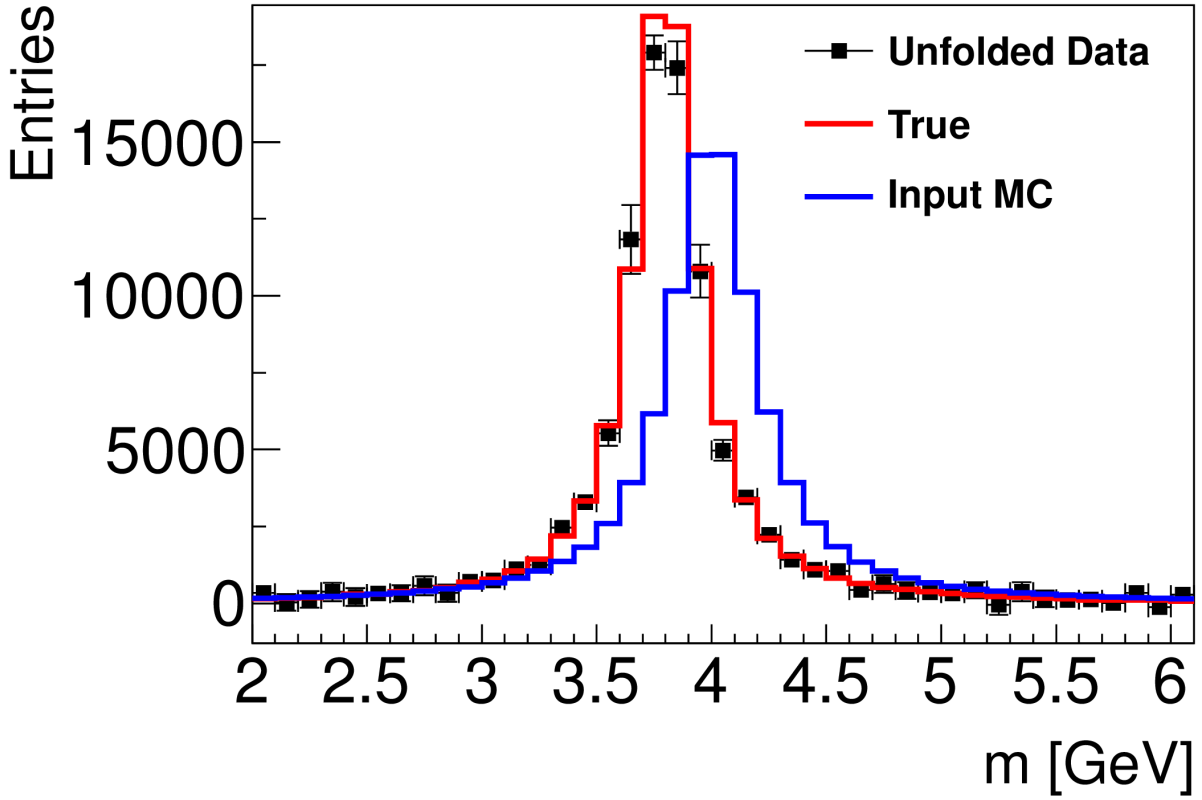


Regularized Unfolding

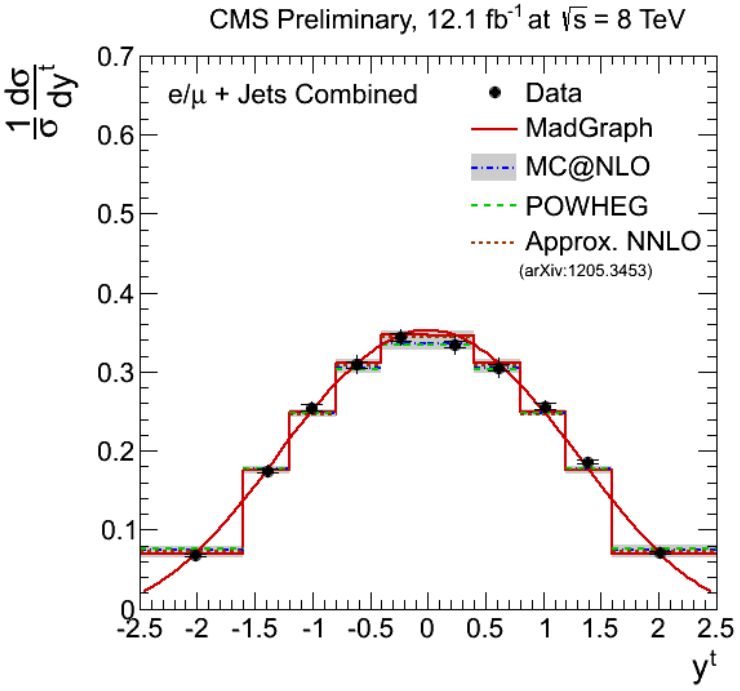


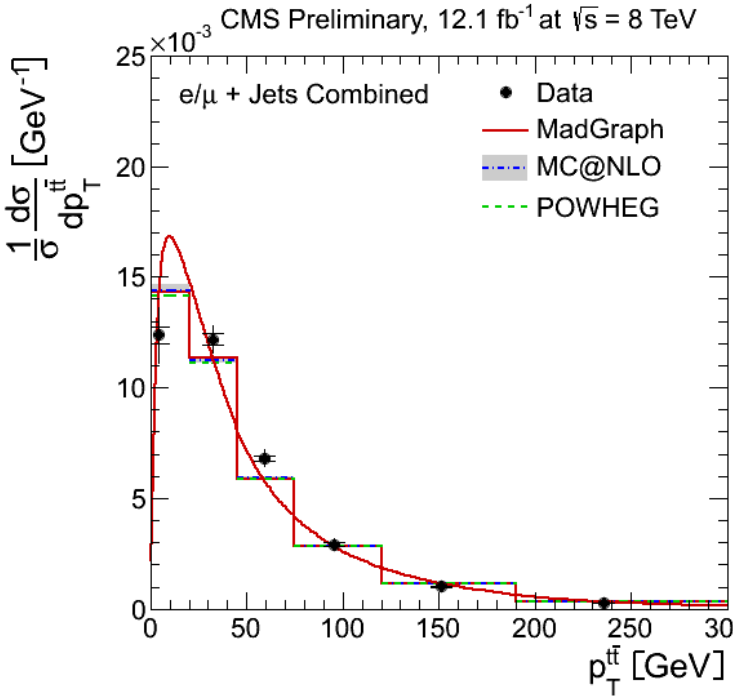


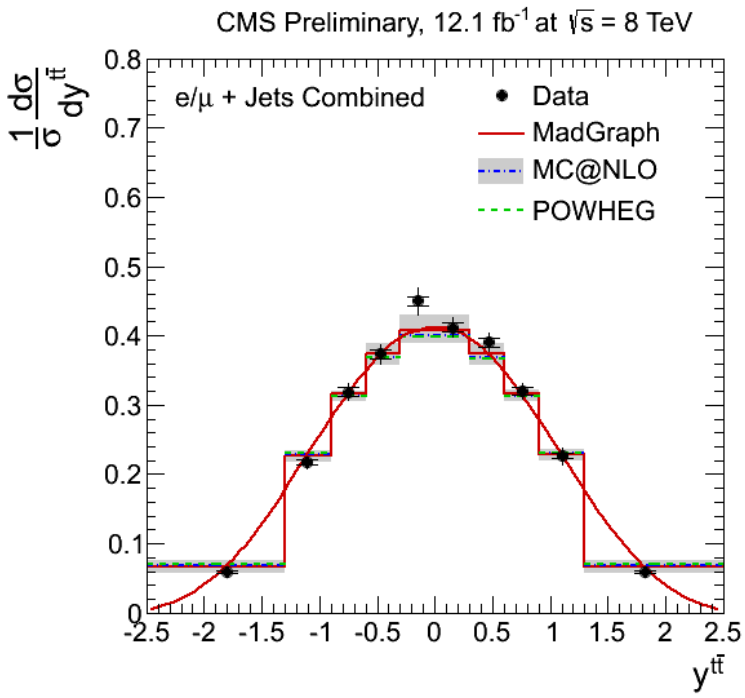
Toy model example: → Unfold smeared distribution to "true"



Differential cross sections





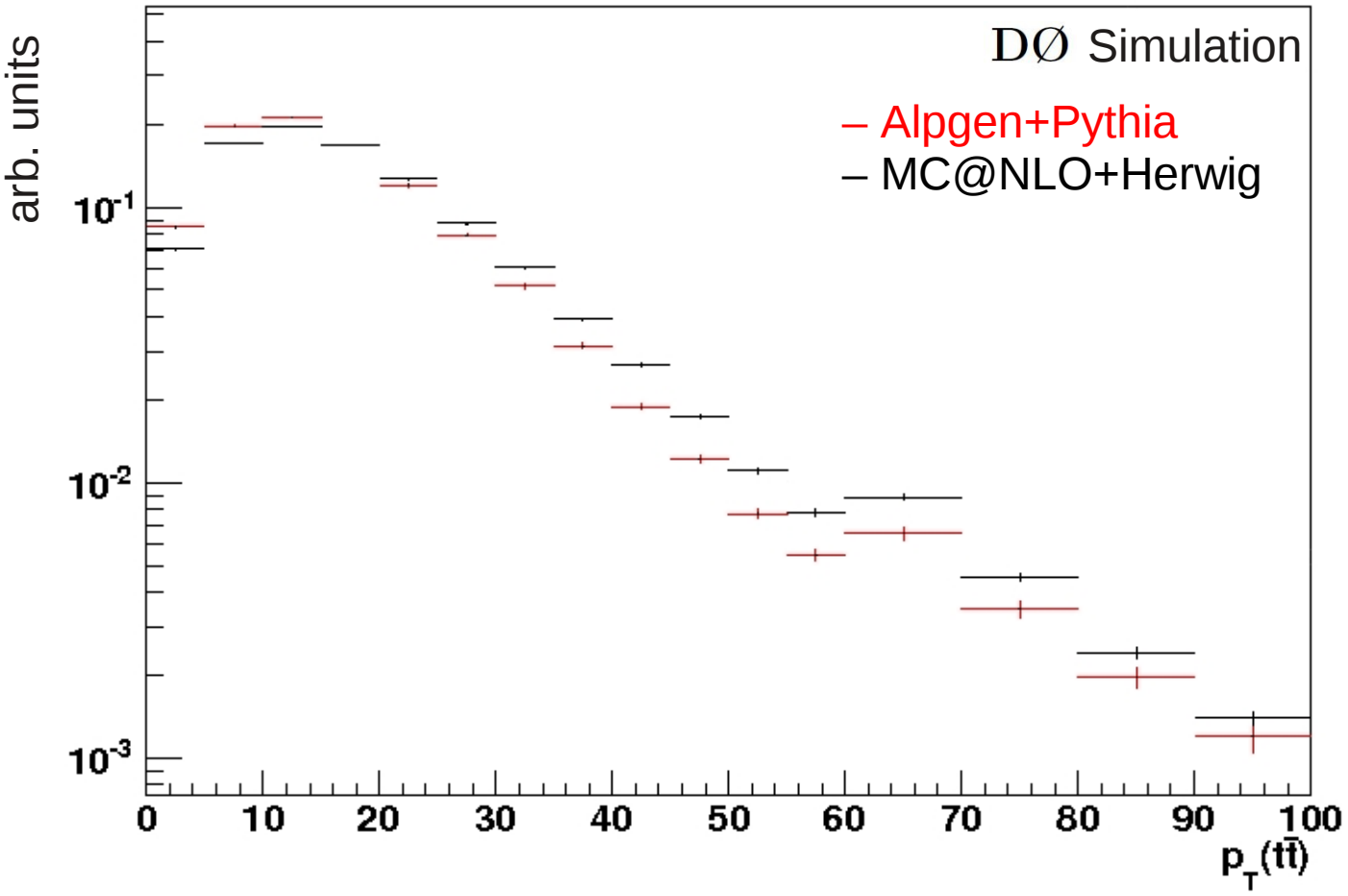


- Good description by predictions

Systematic uncertainties

Study systematic uncertainties and their effect on differential cross sections by:

- Migration matrix derived from simulation with variation of parameters or modified background subtracted data
- Systematic uncertainty is difference in unfolded results to nominal



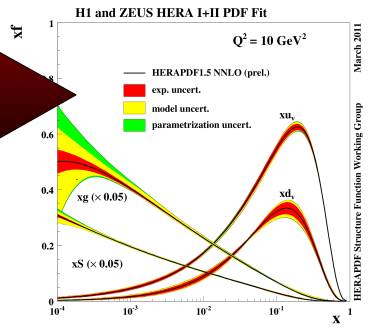
Global analyses / fits

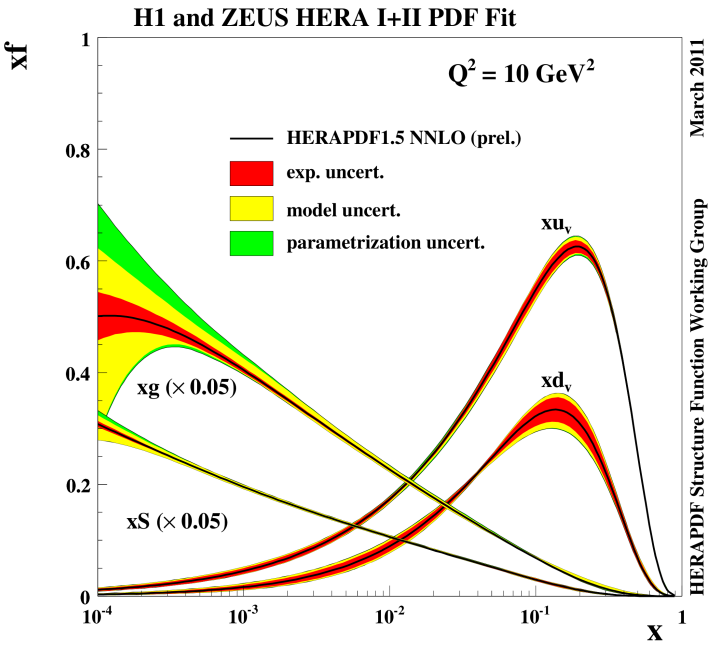
Data: HERA, Tevatron, LHC fixed target data
Processes: DIS, Jets, W & Z production, Drell-Yan, **Top**

Theory: Heavy Flavor scheme, Calculations for process

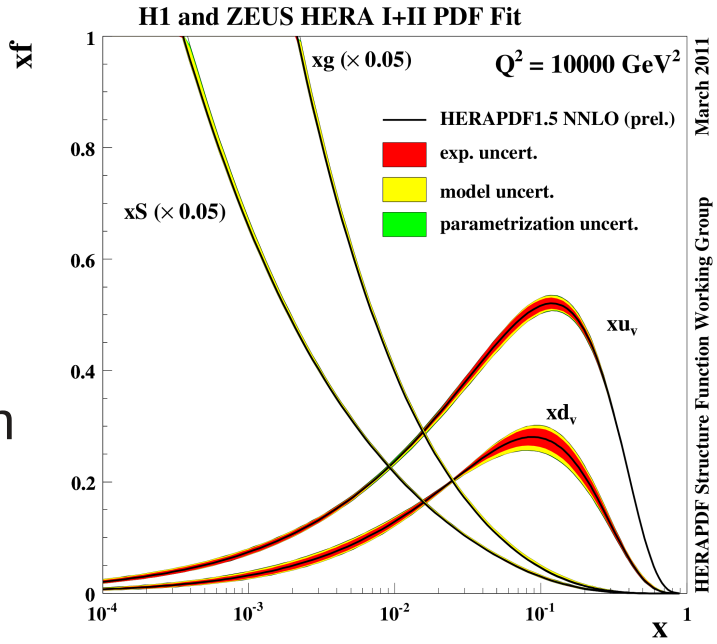
MSTW, CTEQ, NNPDF
HERAPDF, ABM, GJR

Parton distribution functions at starting scale

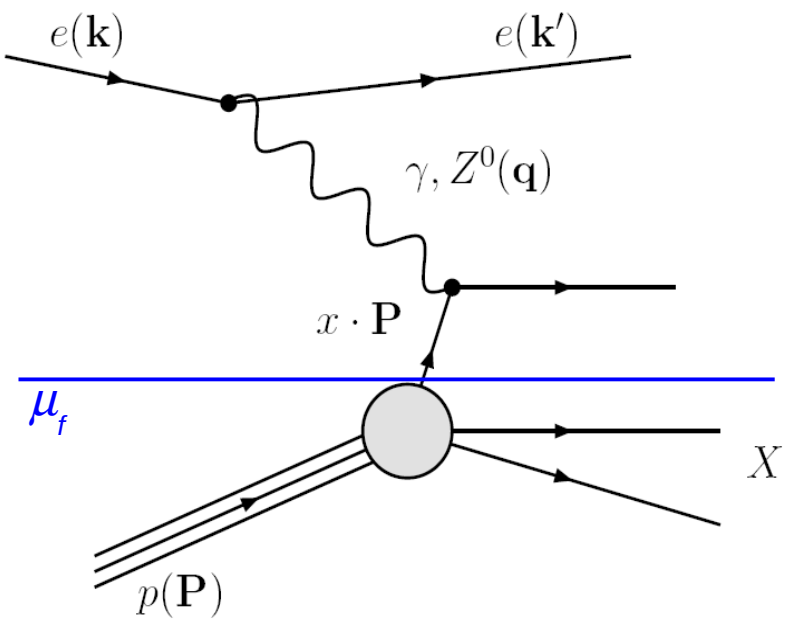




Extrapolate in Q^2 by evolution equations (DGLAP), here: $10 \rightarrow 10000$ GeV^2



Proton structure



Kinematics of DIS given by Lorentz invariant quantities:

Photon virtuality / Resolving power:

$$Q^2 := -q^2 = -(\mathbf{k} - \mathbf{k}')^2$$

Inelasticity:

$$y := \frac{(\mathbf{p} \cdot \mathbf{q})}{(\mathbf{p} \cdot \mathbf{k})}$$

Bjørken x:

$$x := \frac{Q^2}{2(\mathbf{p} \cdot \mathbf{q})}$$

(momentum fraction of scattered parton)

Related by:

$$Q^2 = sxy$$

→ DIS is a unique tool to:

- Test perturbative QCD & measure substructure of proton:
  parton **d**istribution **f**unctions (PDFs)

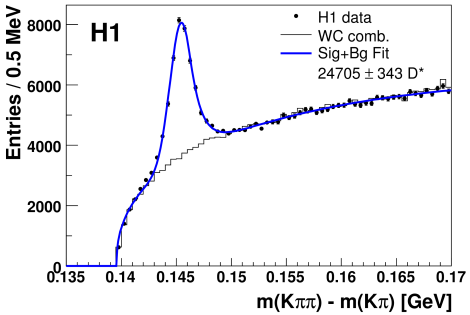
Factorization theorem:

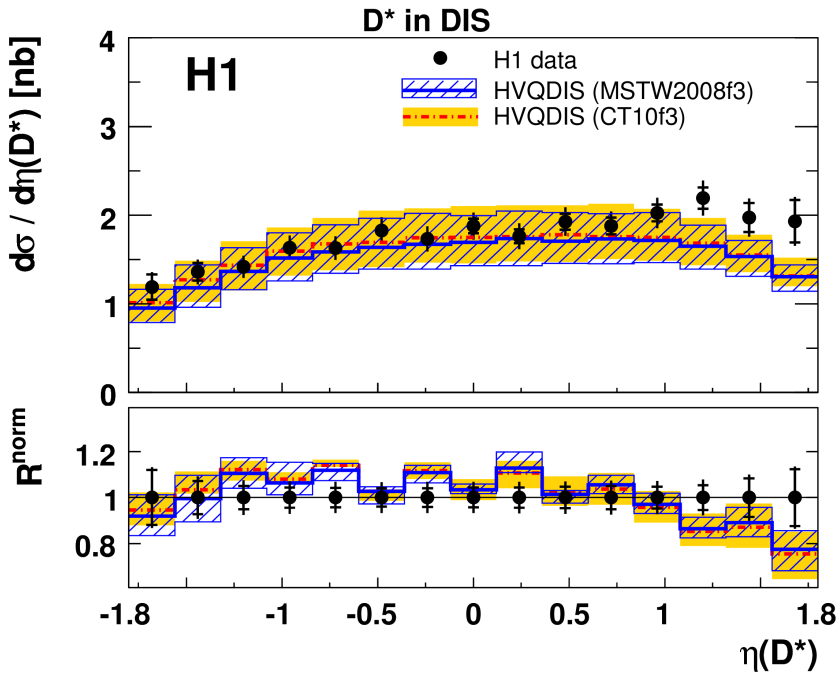
$$d\sigma = \sum_{i,j,k} f_i^A(x_1, \mu_f) \otimes f_j^B(x_2, \mu_f) \otimes d\hat{\sigma}_{ij \to kX}(\mu_f) \otimes D_k^H(z, \mu_f)$$

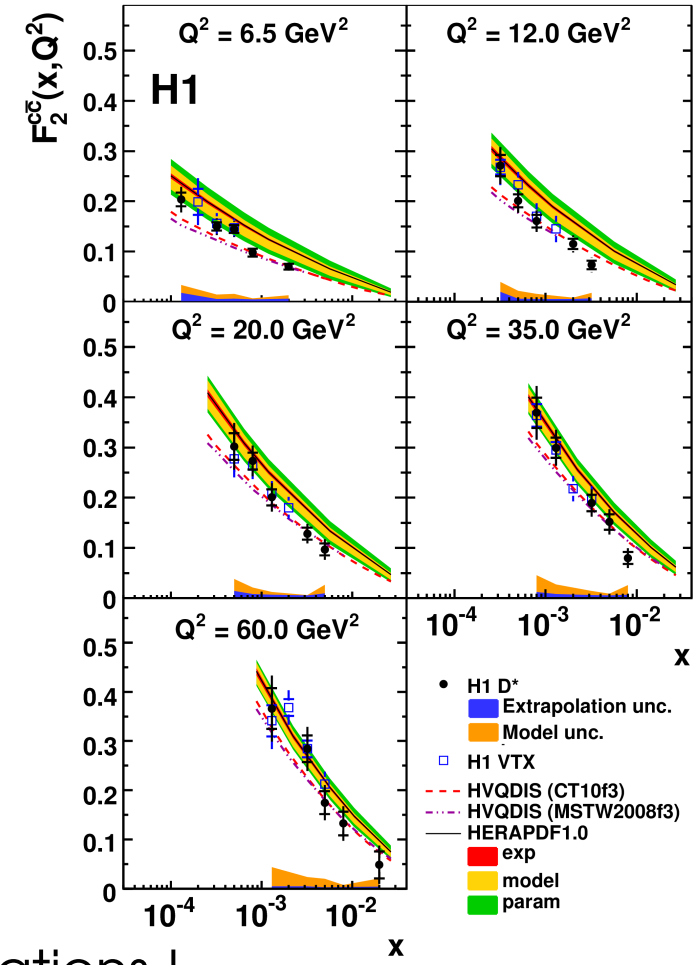
Proton structure

- Full HERA II data sample is used, largest D* phase space coverage at HERA
- Employs regularized Matrix unfolding, reduces model uncertainties
- Significantly reduced systematic uncertainties

Eur.Phys.J. C71 (2011) 1769







- Data are reasonably described by pQCD NLO calculations !
- Double Differential cross sections → measure $F^2_c(x,Q^2)$, used for PDF fits