Measurement of the charge asymmetry in highly boosted $t\bar{t}$ events in the single-lepton channel at 13TeV with the CMS detector

Hugo Becerril[1], Xuan Chen[2],Cecilia Gerber[1], Andreas Jung[3], Titas Roy[1]

University of Illinois at Chicago[1], Cornell University[2], Purdue University[3]

Lake Louise Winter Institute 2022

CMS

UIC University of Illinois at Chicago

Outline

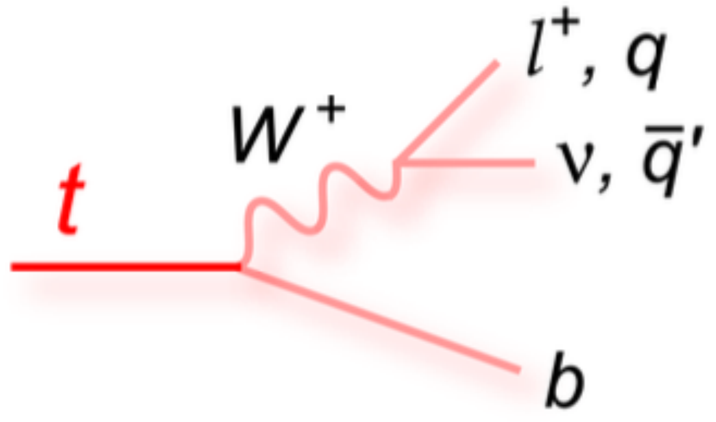
- Introduction and Motivation
- Ac measurement
- Conclusions

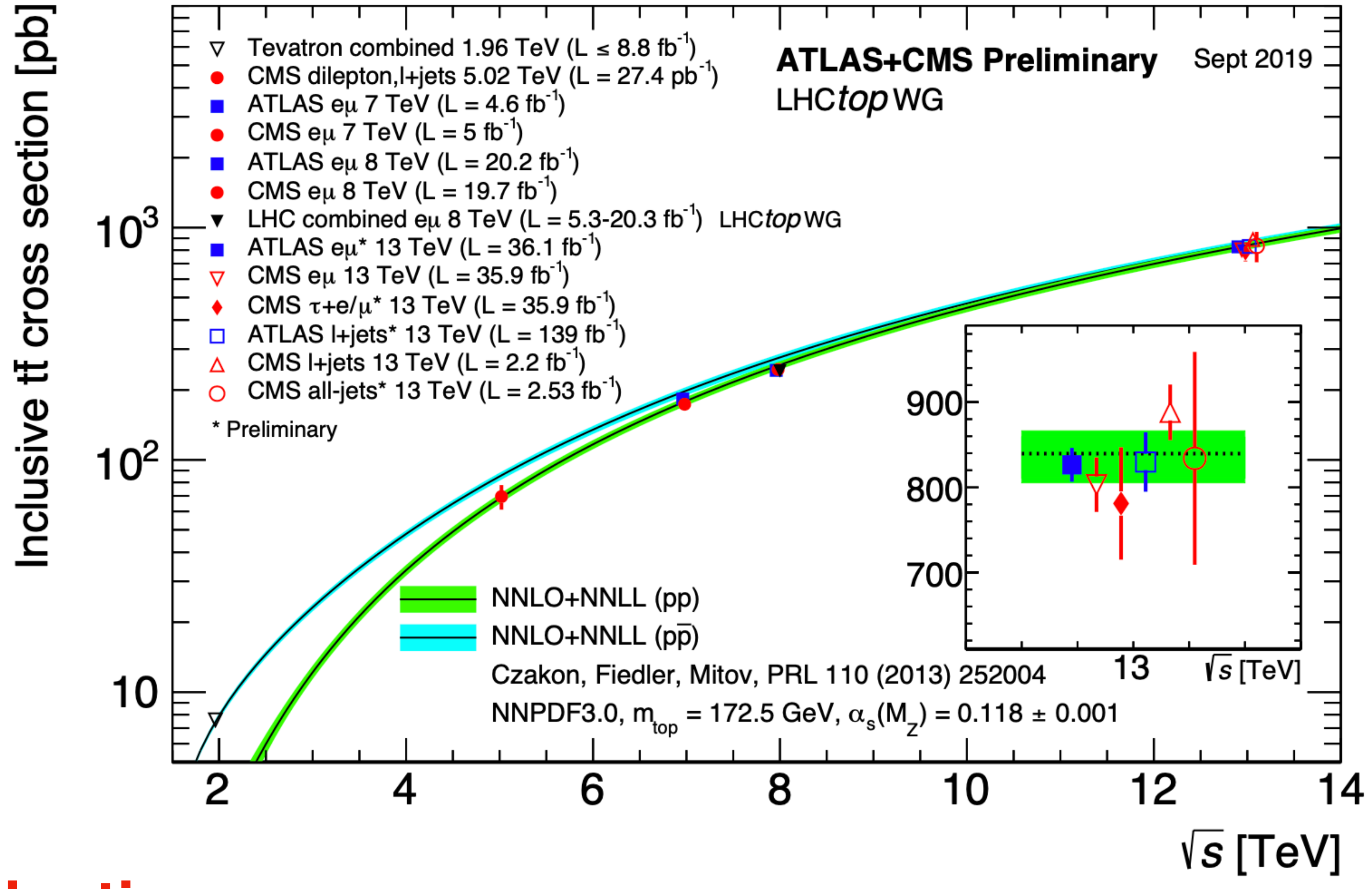
Introduction and Motivation

Top quark pair production at hadron colliders

The top quark, a unique particle

- **Most massive elementary particle $m_{top} \approx 175$ GeV**
- **Large coupling to Higgs boson & special role in EWSB**
- **Decays before hadronising, allowing study of bare quarks**

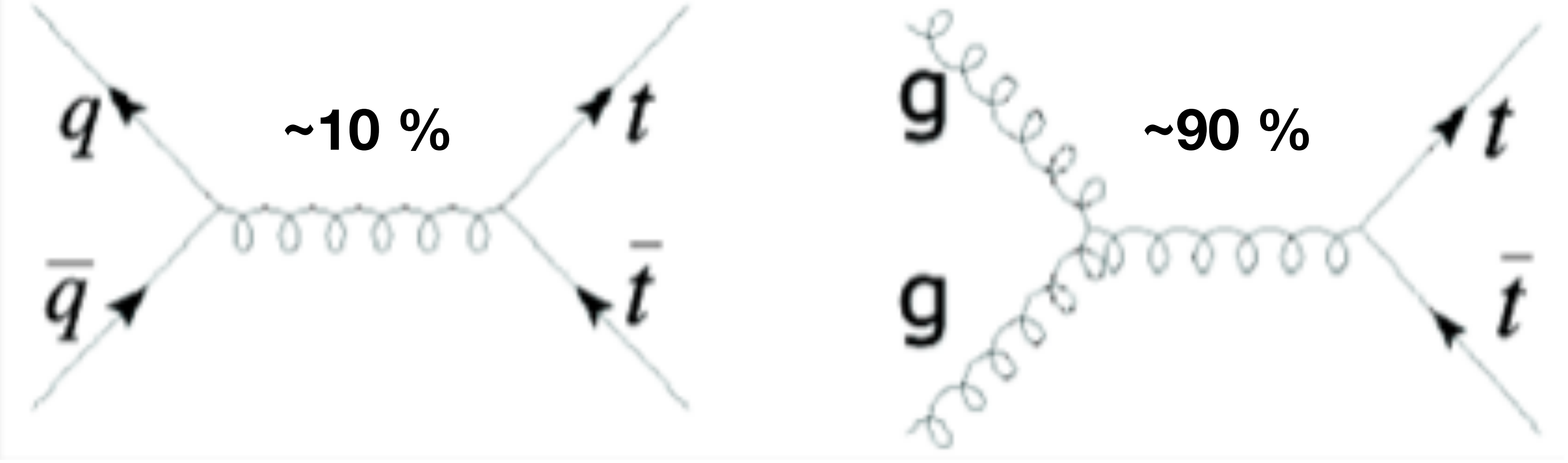


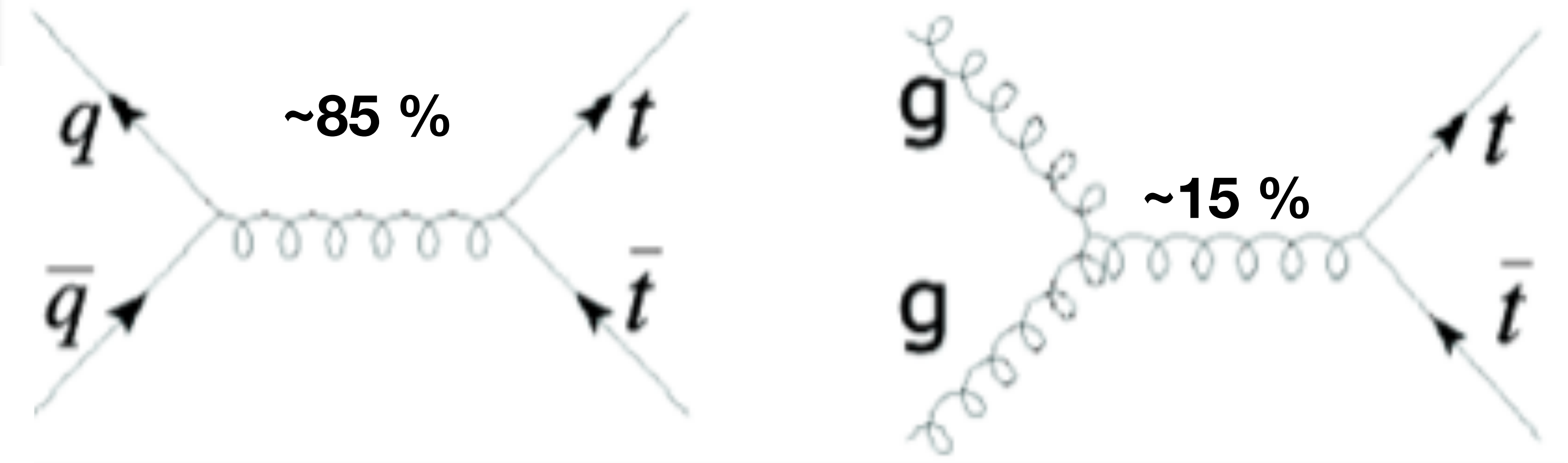


CERN $(pp)\ \sqrt{s} = 13 TeV$

FERMILAB
$\sqrt{s} = 1.96 TeV\ (p\bar{p})$

Top quark pair production





CMS
UIC University of Illinois at Chicago

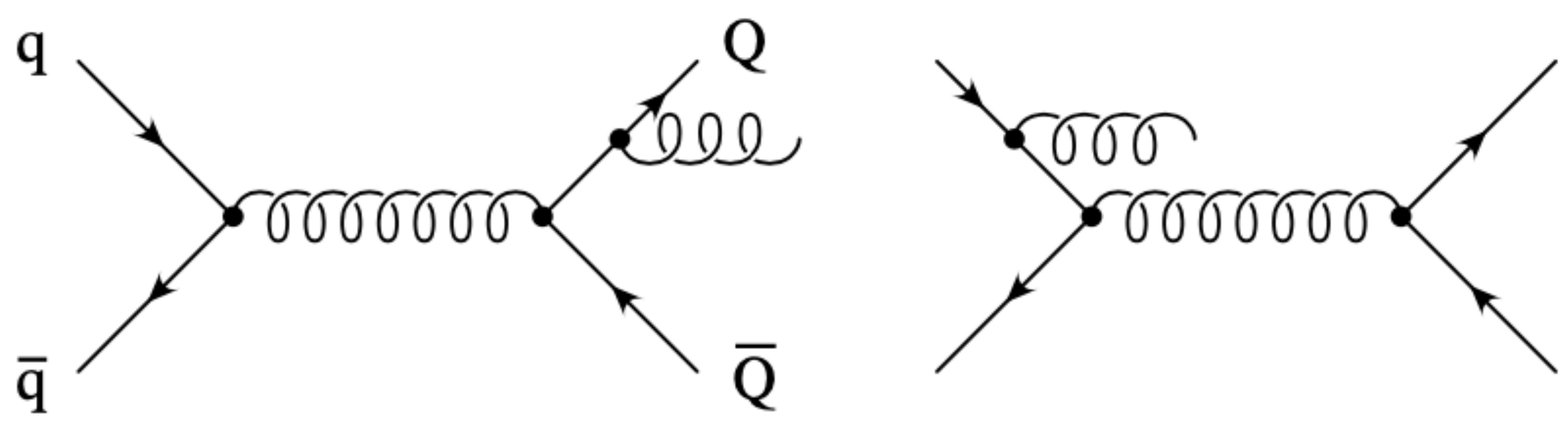
Charge asymmetry in hadron colliders

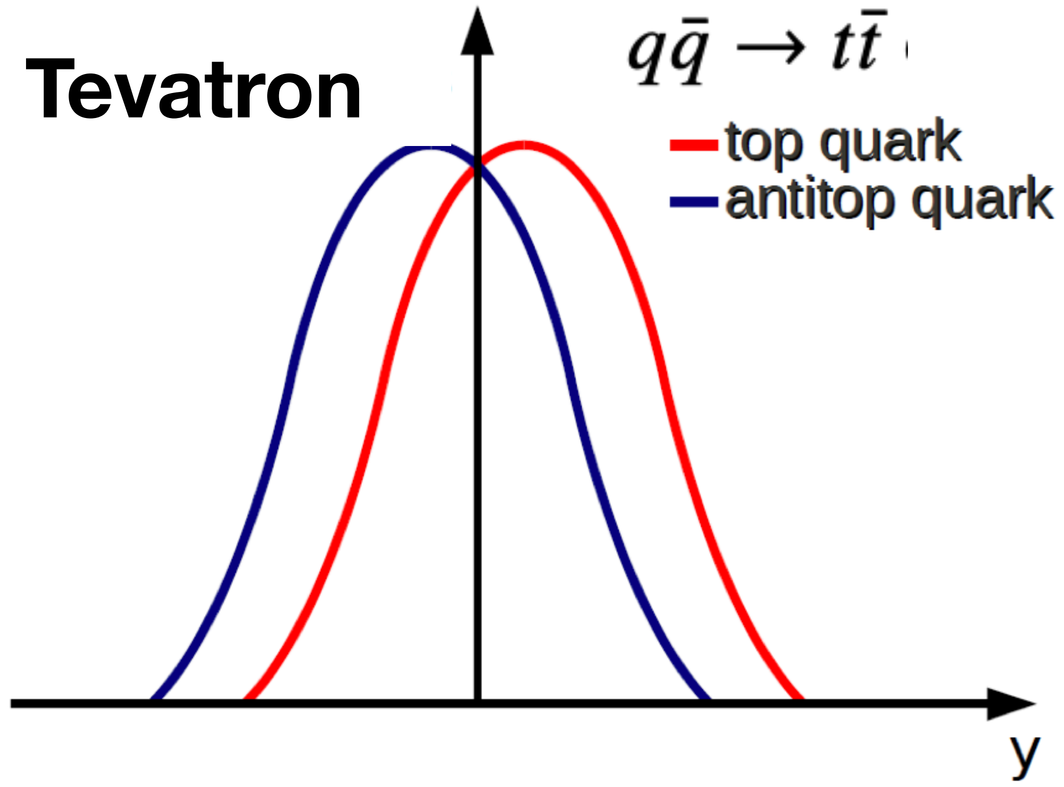
The relevant processes at LO for heavy flavor are symmetric under charge conjugation.

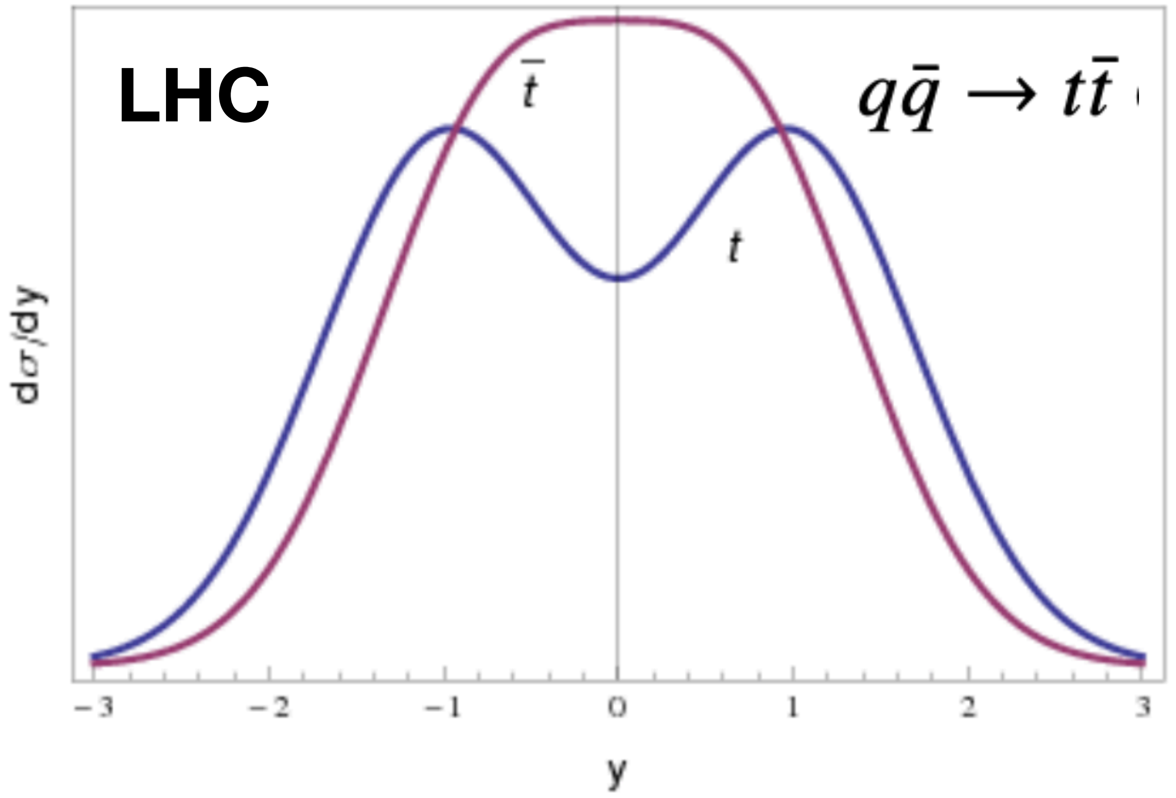
$$g + \bar{g} -> Q + \bar{Q}$$

$$q + \bar{q} -> Q + \bar{Q}$$

- Radiative corrections lead to a charge asymmetry.



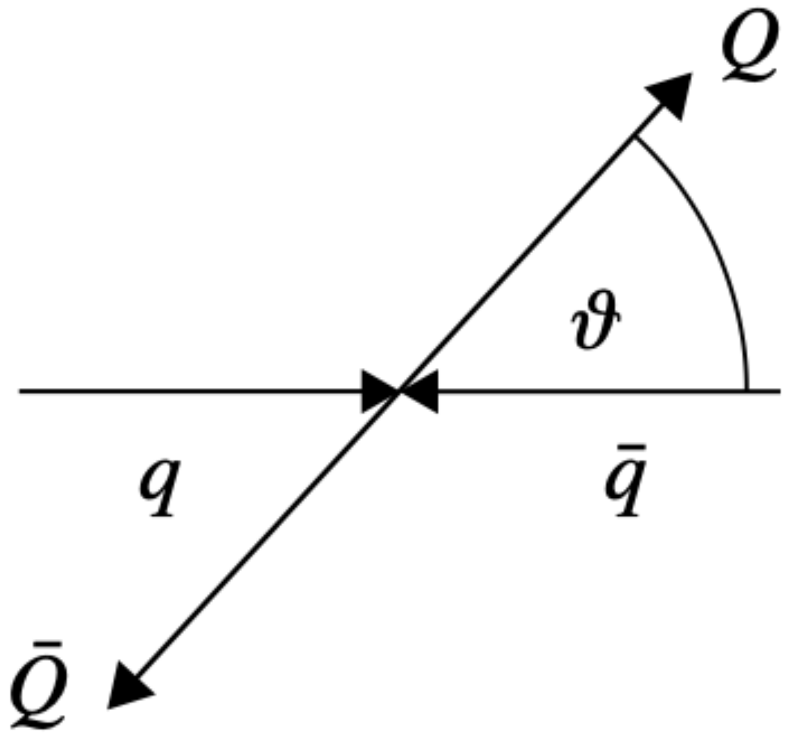




$$A^{FB} = \frac{N(\Delta y > 0) - N(\Delta y < 0)}{N(\Delta y > 0) + N(\Delta y < 0)}$$

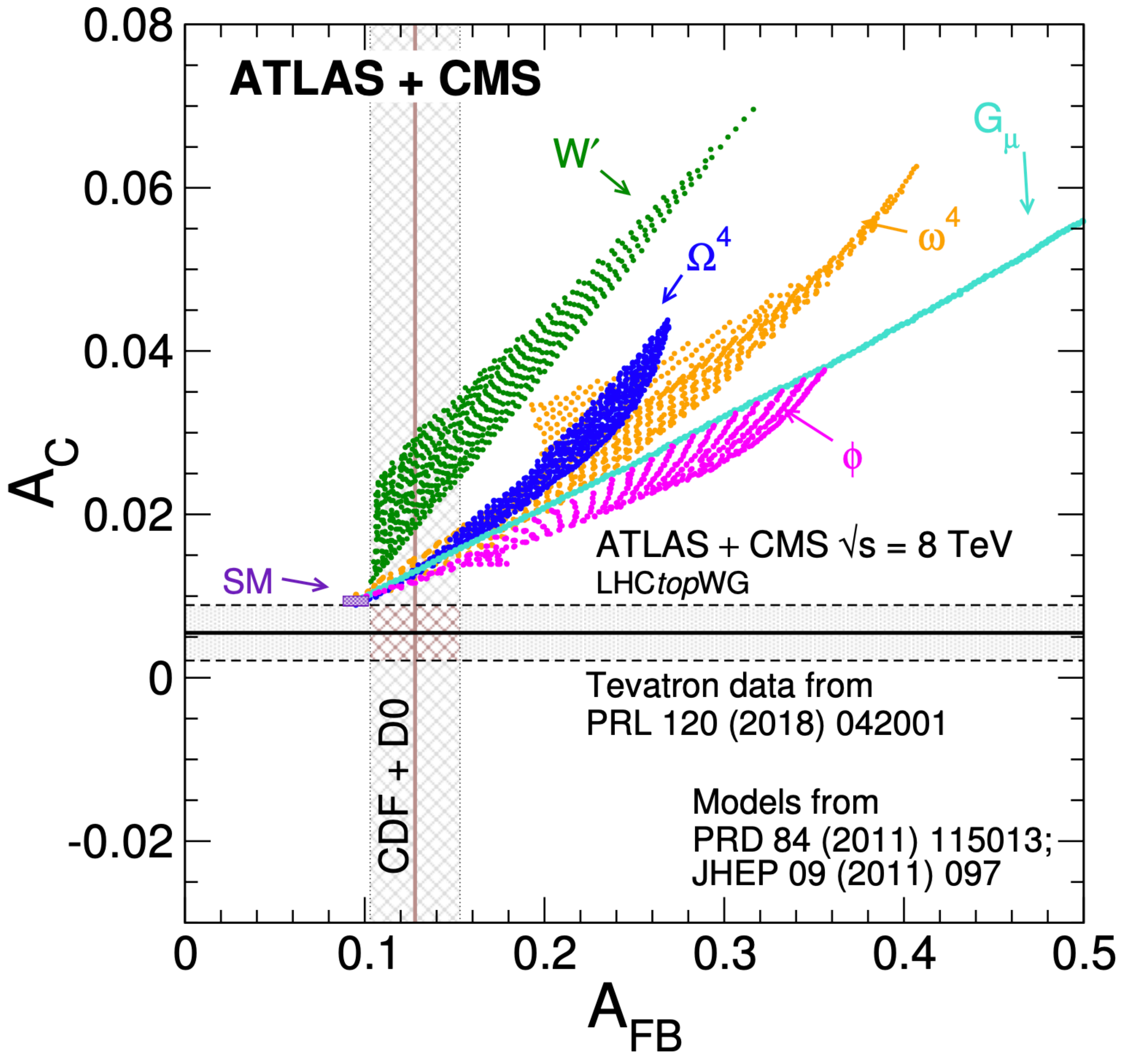
$$A_C = \frac{N(\Delta|y| > 0 - N(\Delta|y| < 0)}{N(\Delta|y| > 0 + N(\Delta|y| < 0)}$$

As a consequence of these asymmetries, the top quark is preferentially produced in the direction of the incoming quark.



Δy is a suitable variable to study the asymmetry because it is Lorentz invariant with respect to boosts along the z-axis.
Δy is defined as:

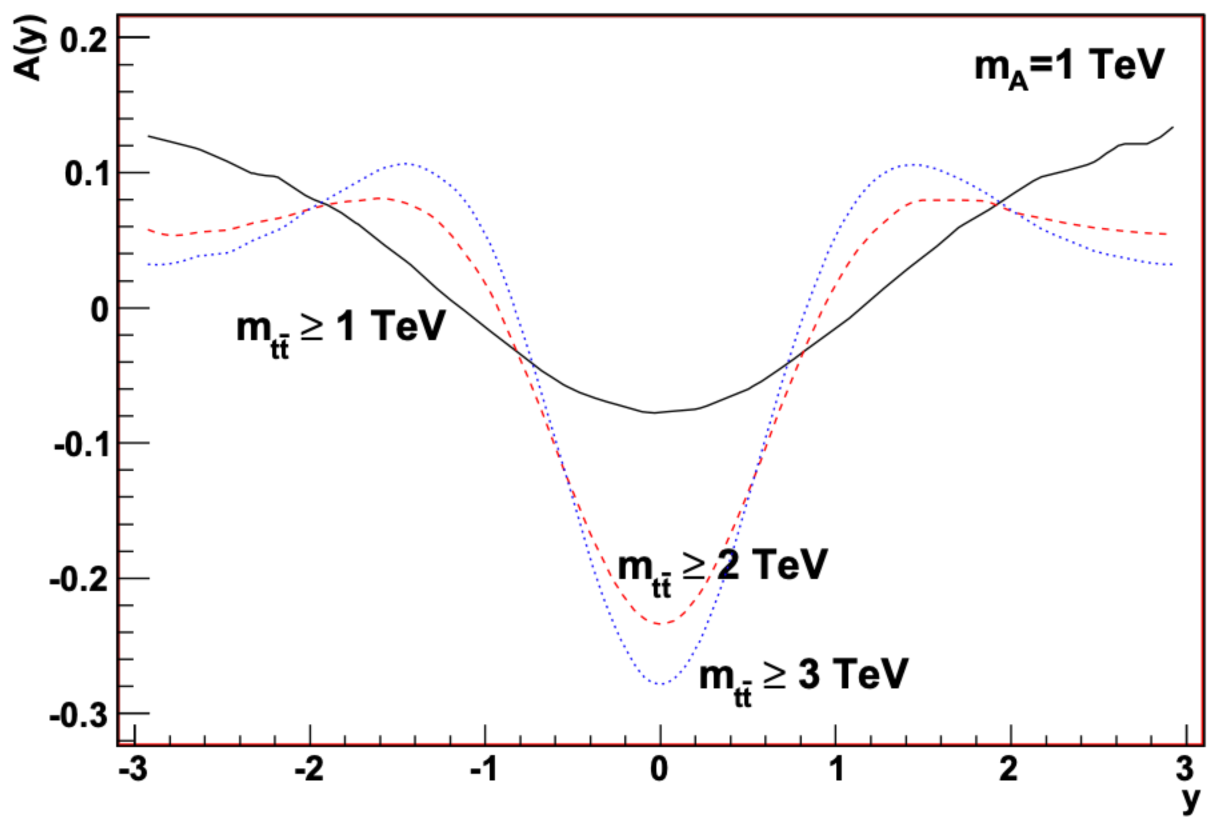
$$\Delta y = y_t - y_{\bar{t}}$$



https://doi.org/10.1007/JHEP04(2018)033

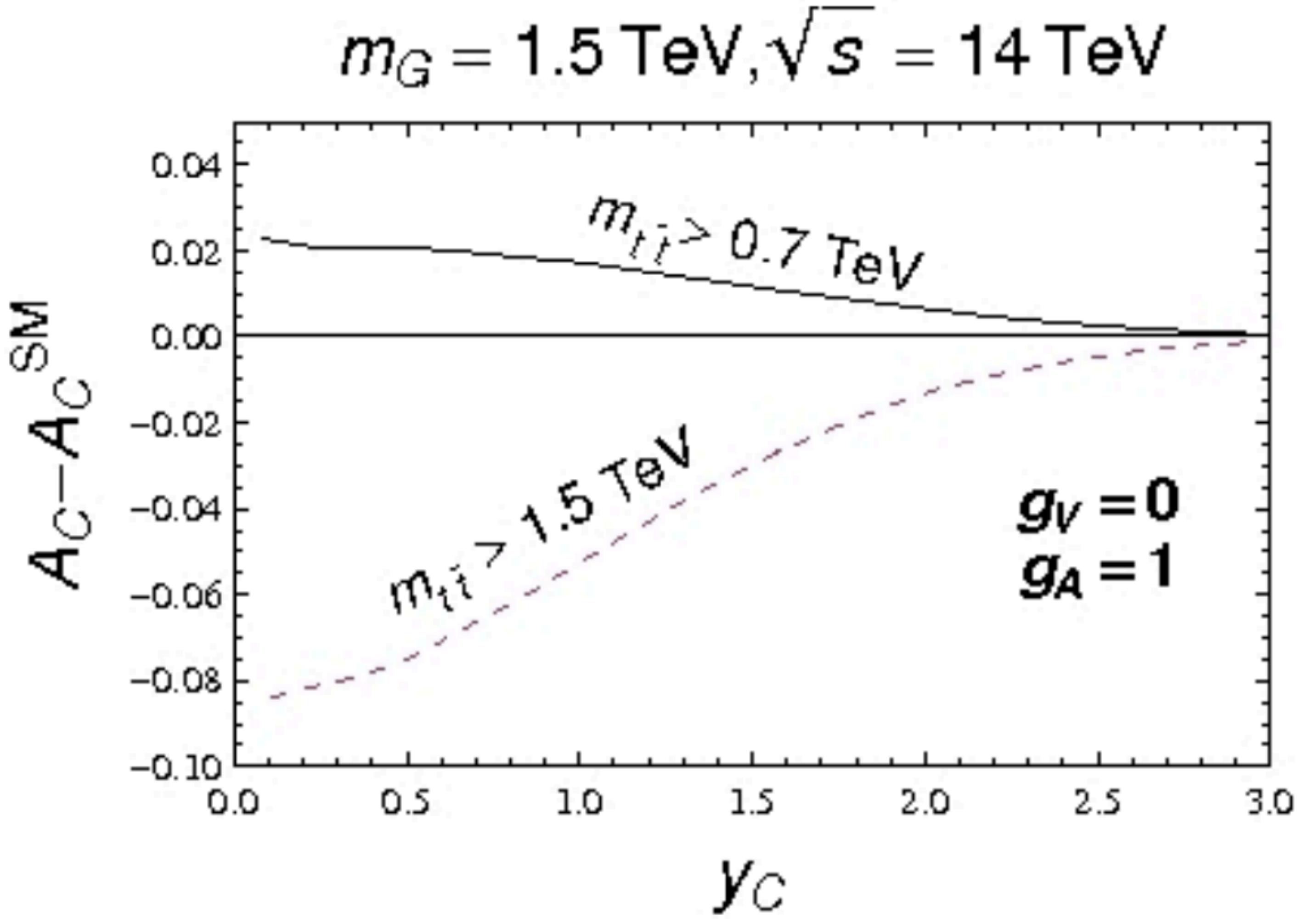
Charge Asymmetry in BSM processes

- There are still several BSM models that predict charge asymmetry which vary as a function of the invariant mass *m*tt. BSM effects are expected to be enhanced for large invariant mass because at high transfer momentum fractions the relative contribution of valance quarks increase.



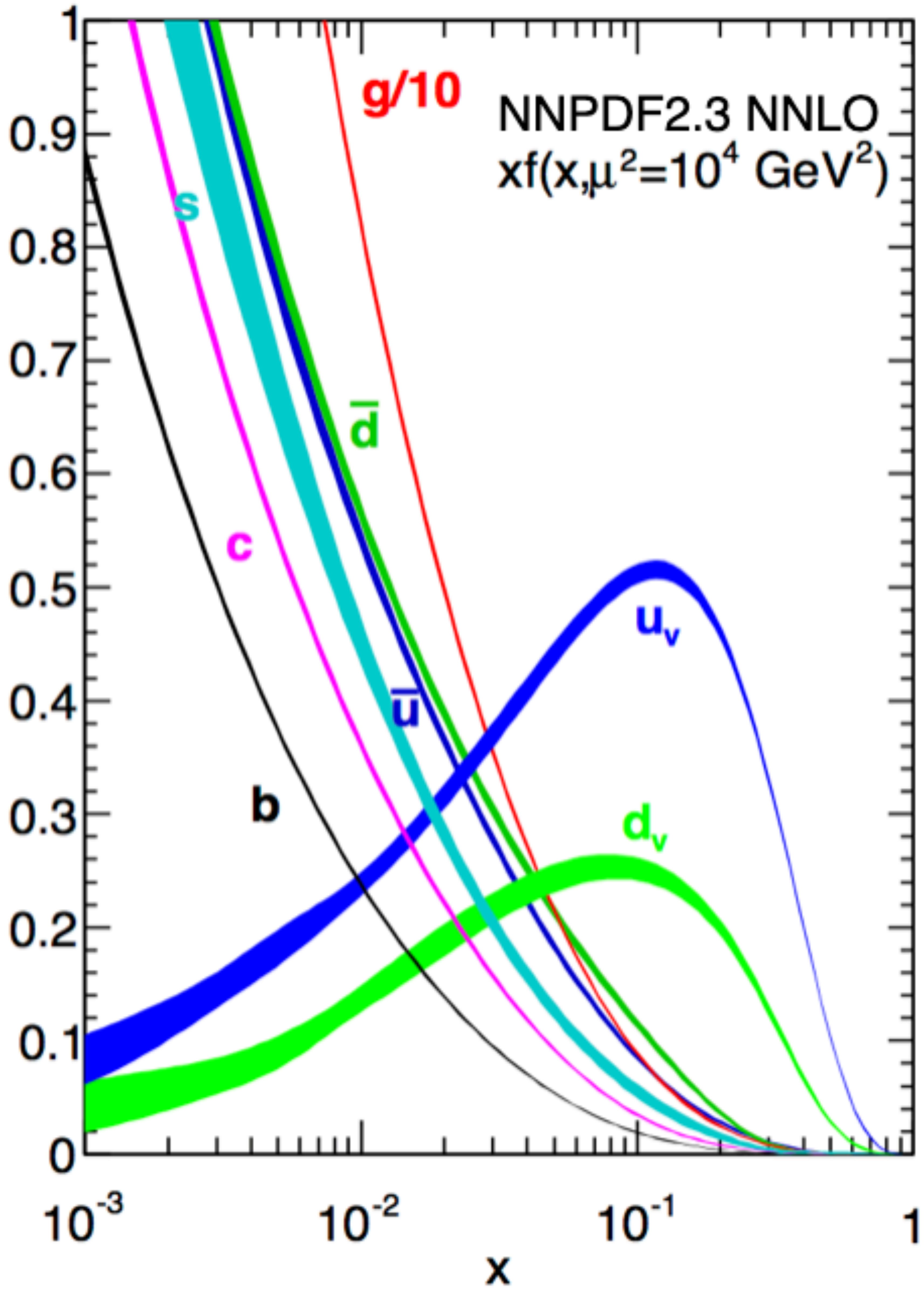
Phys. Rev. D 77, 014003

Top quarks, axigluons, and charge asymmetries at hadron colliders



Phys. Rev. D 78, 094018

Massive color-octet bosons and the charge asymmetries of top quarks at hadron colliders



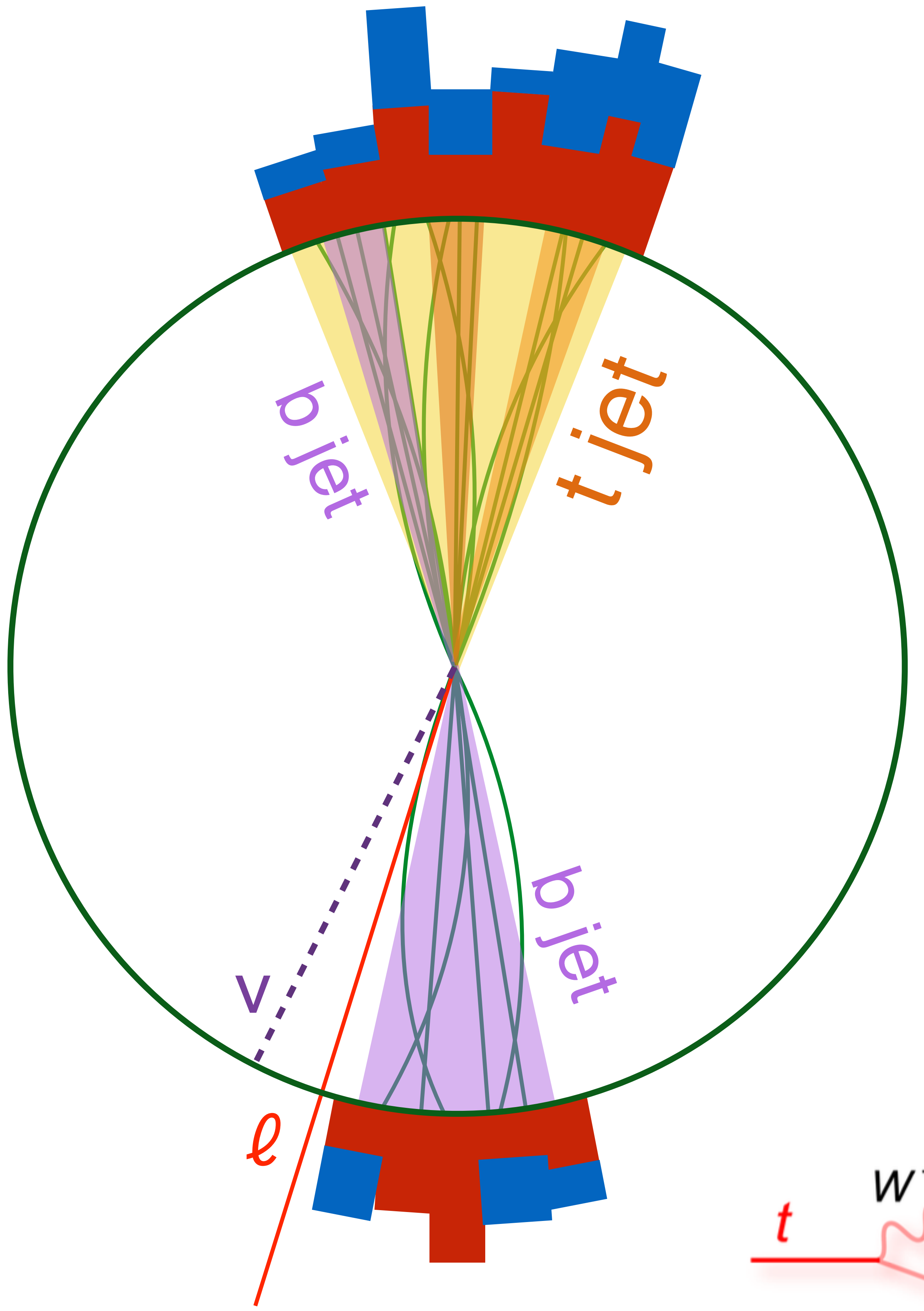
Ac measurement

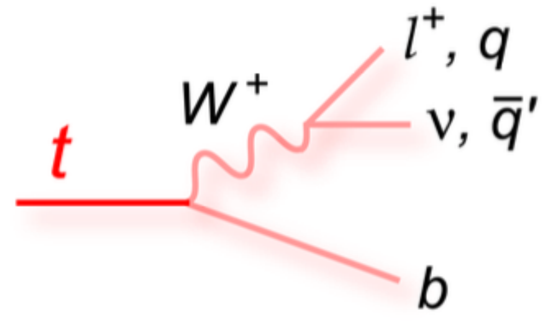
Baseline selection in the l+jets channel for Full Run II data

Event selection:

- HLT_Mu50_v*
- Exactly one muon (CutedBasedGlobalHightPt), pT > 55GeV and $|\eta| < 2.4$
- MET > 50 GeV
- At least two high pT jets: AK4 CHS pT (jet1) > 150GeV, pT(jet2) > 50 GeV
- HT_lep > 150 GeV
- AK8 PUPPI jets with pT > 400 GeV and $|\eta| < 2.4$
- Veto on events with second lepton pT > 20 GeV, $|\eta| < 2.4$

- HLT Ele50 CaloIdVT GsfTrkIdT PFJet165 v* OR HLT Ele115 CaloIdVT GsfTrkIdT v*
- Exactly one electron (cutes-based electron Tight), pT > 85GeV and $|\eta| < 2.4$.
- MET > 120 GeV
- At least two high pT jets: AK4 CHS pT (jet1) > 185GeV, pT(jet2) > 50 GeV and $|\eta| < 2.4$
- AK8 PUPPI jets with pT > 400 GeV and $|\eta| < 2.4$
- Veto on events with second lepton pT > 20 GeV, $|\eta| < 2.4$



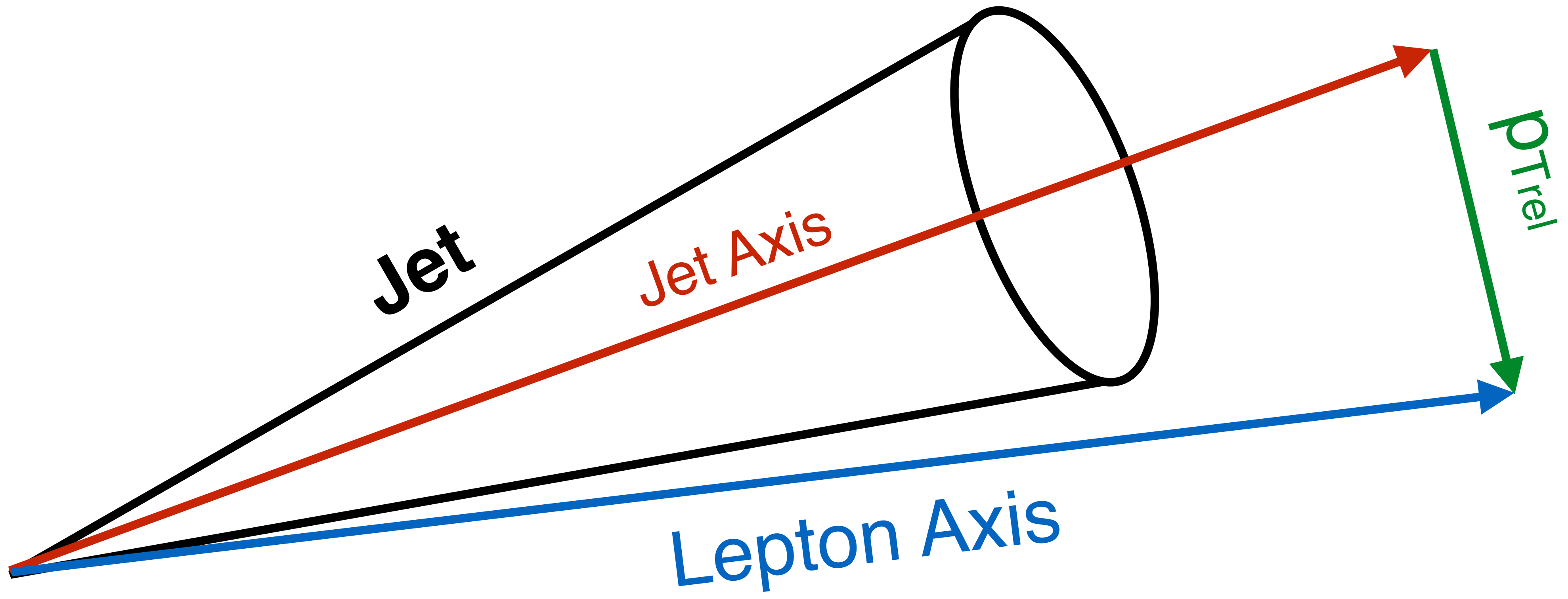


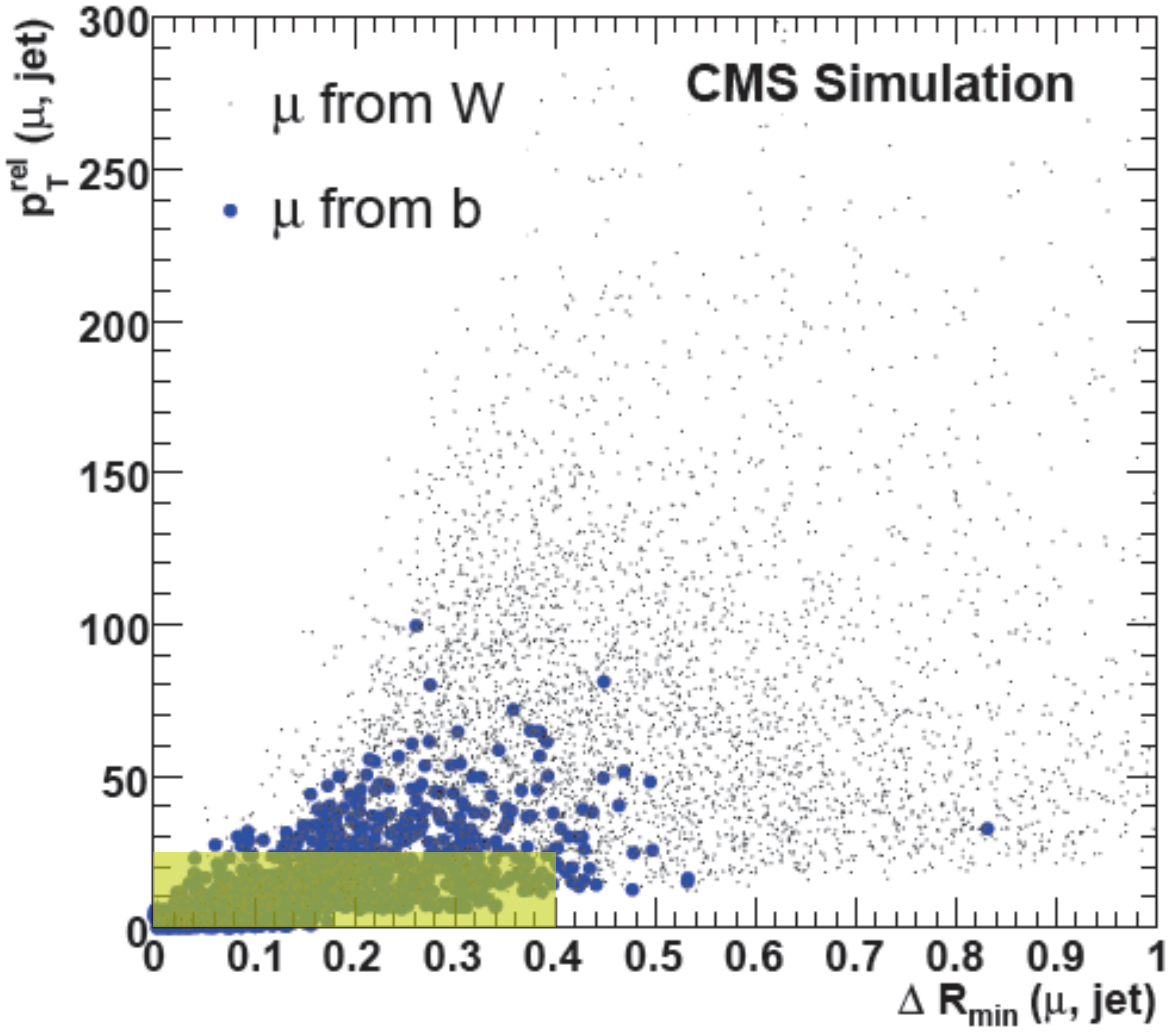
Single lepton channel

2D-Momentum Isolation

Published in the Journal of High Energy Physics as doi:10.1007/JHEP04(2019)031.

- Identification of leptons inside jets
  - if $\Delta R_{ij} > 0.4$: No p_{Trel} Selection
  - if $\Delta Rij < 0.4$: $p_{Trel} > 25$ GeV

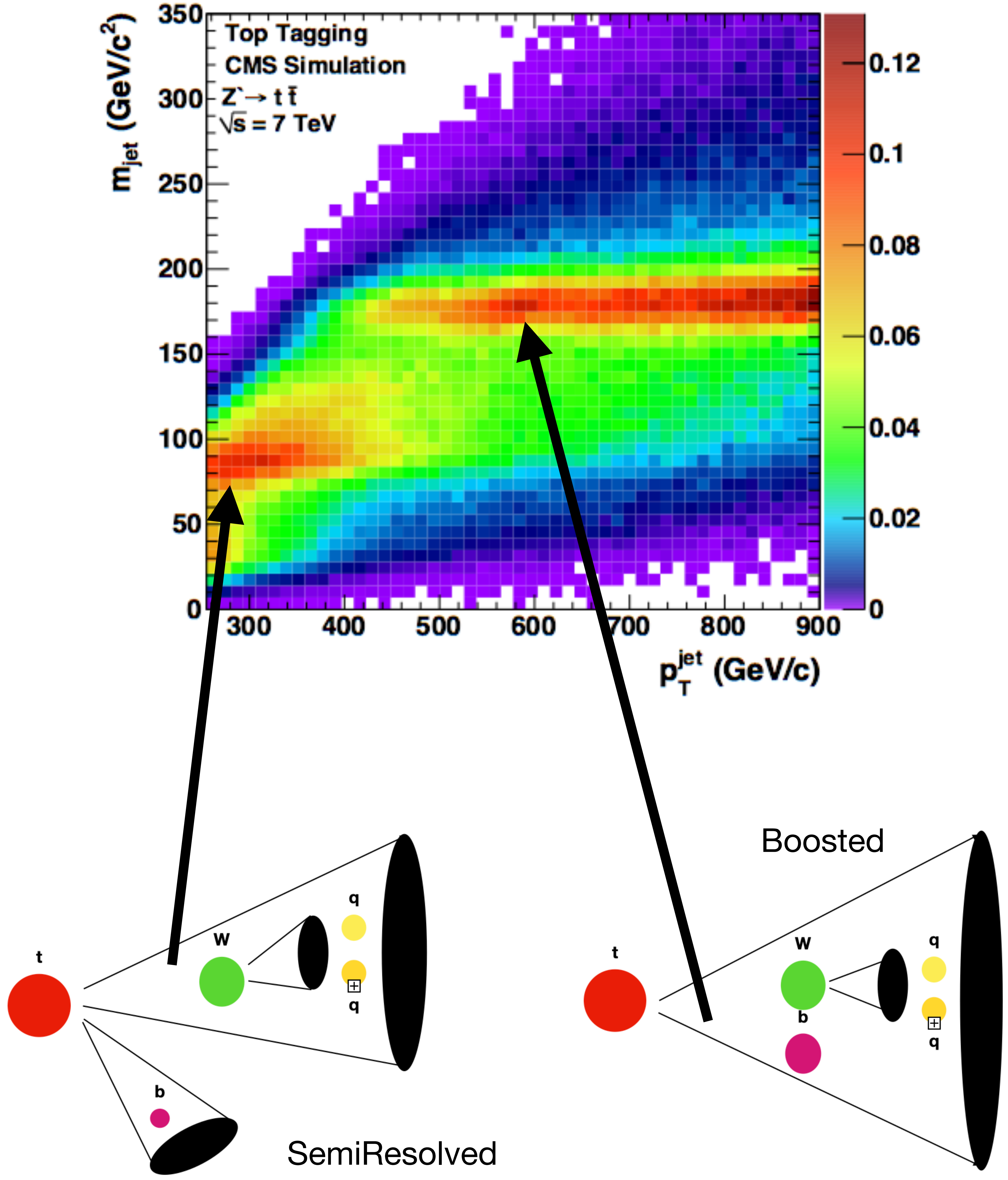




AK8 Top/W tagging

- Jets from a hadronic boosted top quark (W boson), will have three (two) groups of clusters where the jet energy is concentrated. These structures correspond to the quarks from the decay.

- Sub-structure techniques are applied to large footprint jets to identify top quarks:
  - Jet mSD near to the top (W) mass (Soft drop mass: is a jet grooming algorithm to remove the soft and wide angle radiation from a jet) JHEP03(2011)015
  - Compatibility of a large radius jet having 3 (2 or 1 subjets) N-subjetiness: JHEP05(2014)146,



Currently using these working points

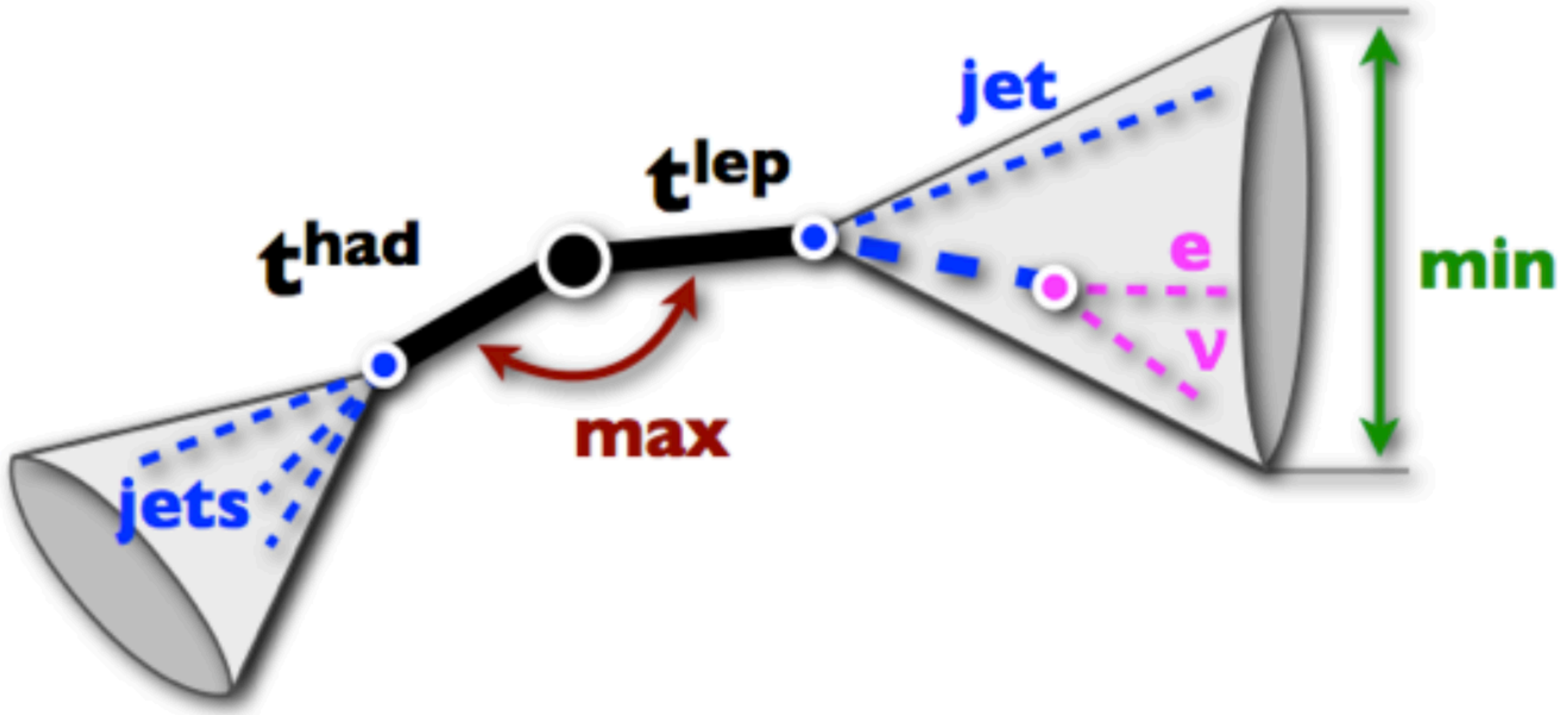
	mSD	$\tau_{32} = \tau_3/\tau_2$
Toptagging	105 < m_jet < 220	< 0.65
	mSD	$\tau_{21} = \tau_2/\tau_1$
W tagging	65 < m_jet < 105	< 0.45

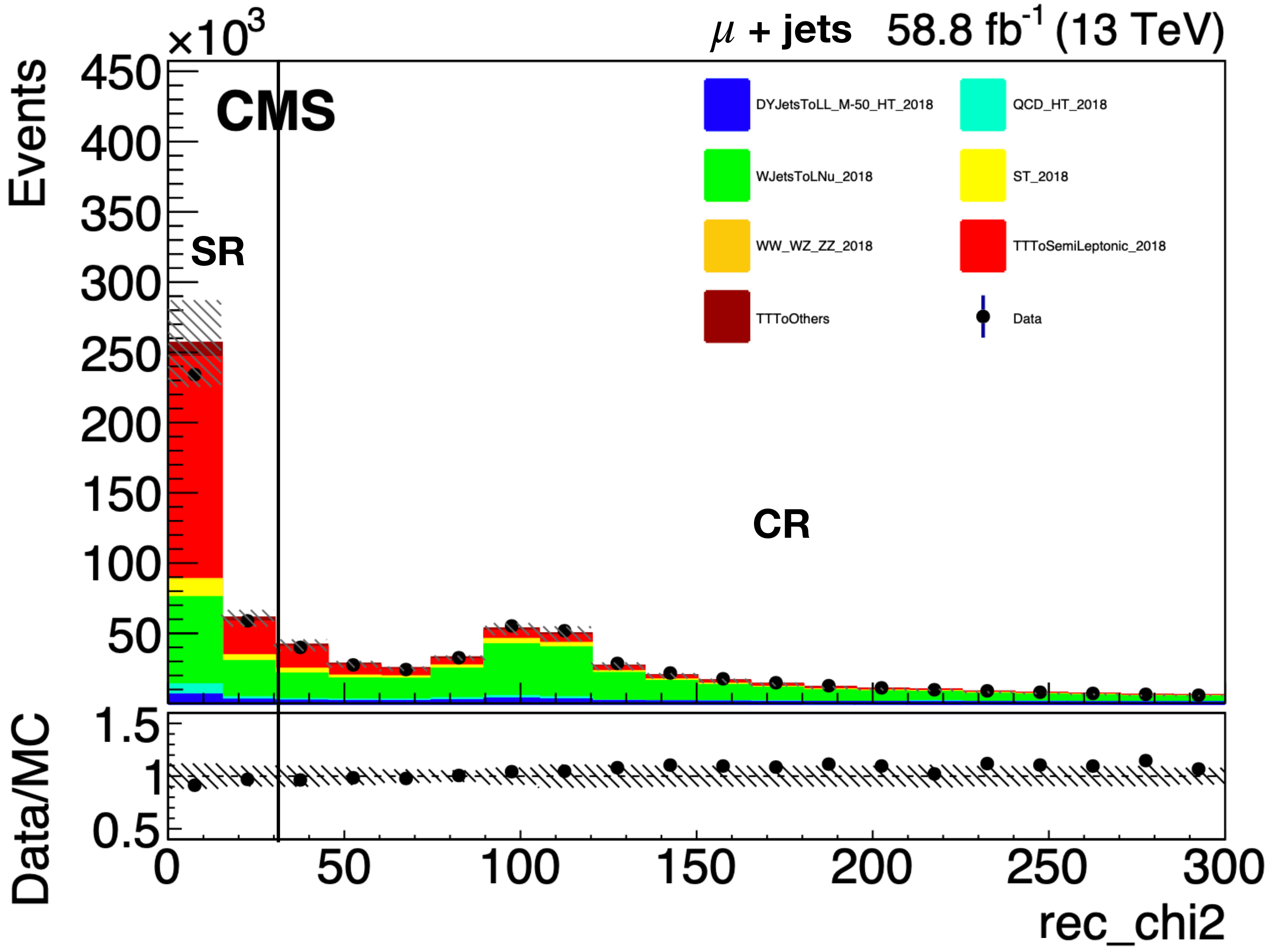
University of Illinois at Chicago

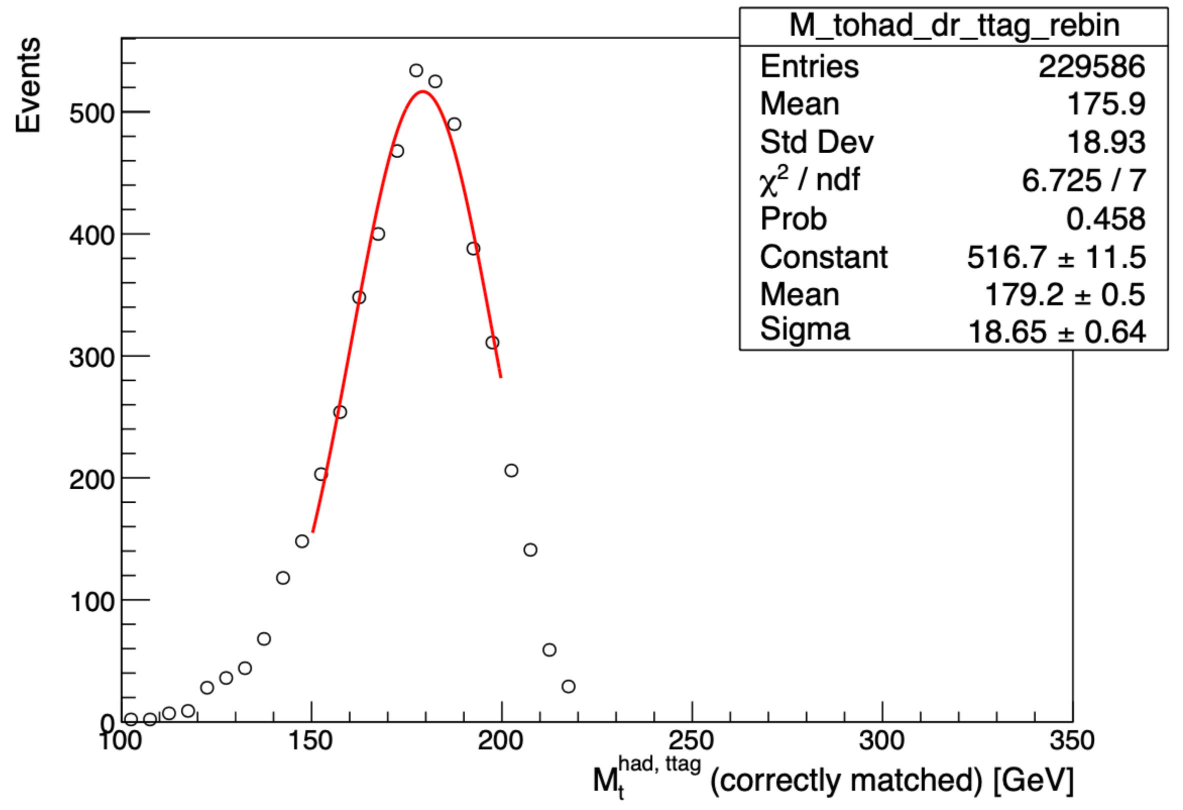
Event reconstruction

- χ^2 techniques solves $t\bar{t}$ system
  - All combinations tested
    - Keep only the minimum
  - $\chi^2 < 30$

$$\chi^2 = \chi^2_{lep} + \chi^2_{had} = [\frac{M_{lep} - \bar{M}_{lep}}{\sigma_{M_{lep}}}]^2 + [\frac{M_{had} - \bar{M}_{had}}{\sigma_{M_{had}}}]^2$$







$\bar{M}_{lep}$ = 169 GeV, $\bar{\sigma}_{M_{lep}}$ = 21 GeV, $\bar{M}_{had}$ = 175.9, $\bar{\sigma}_{M_{had}}$ = 18.93

CMS UIC University of Illinois at Chicago

Final events selection and categorization

The measurements is performed in different channels corresponding to boosted, semi-resolved and resolved topologies defined as:

- Boosted: #Top tagged jets =1 **AND** #W tagged jets = 0 **AND** #b-tagged jets >= 1
- SemiResolved: #Top tagged jets = 0 **AND** #W tagged jets = 1 **AND** #b-tagged jets >= 1
- Resolved: #Top tagged jets = 0 **AND** #W tagged jets = 0 **AND** #b-tagged jets >= 1

We use events in the highly boosted region i.e. $M_{t\bar{t}} > 750$ GeV.

Unfolding with combine

We are using a likelihood-based unfolding, using the combine Higgs tool (command line interface to apply many statistical techniques available inside RooFit/RooStats)

$$\mathcal{L} = \prod_{i \in \mathrm{reco}} \mathcal{P}\Big(x_{\mathrm{reco},i} \Big| \sum_{j \in \mathrm{gen}} \mu_j \mathbf{R}_{ij} + \vec{b}_i\Big)$$

Where I = 2,j = 2

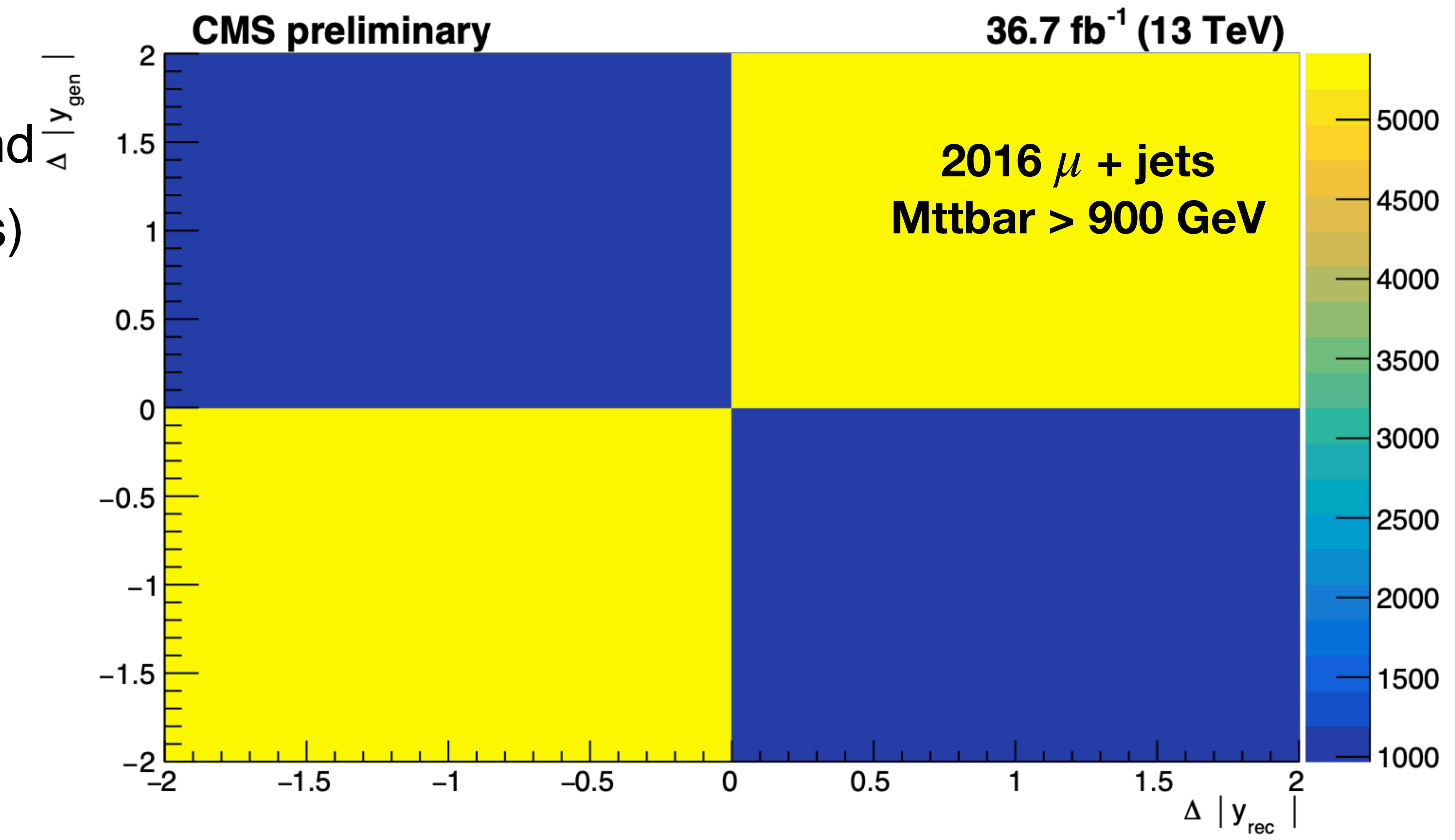
The way we construct the likelihood is via the datacard

- Each datacard represents a reconstructed bin
- Each datacard describes the contribution from each of the common truth bins $\Delta|y_{gen}| > 0, \Delta|y_{gen}| < 0.$

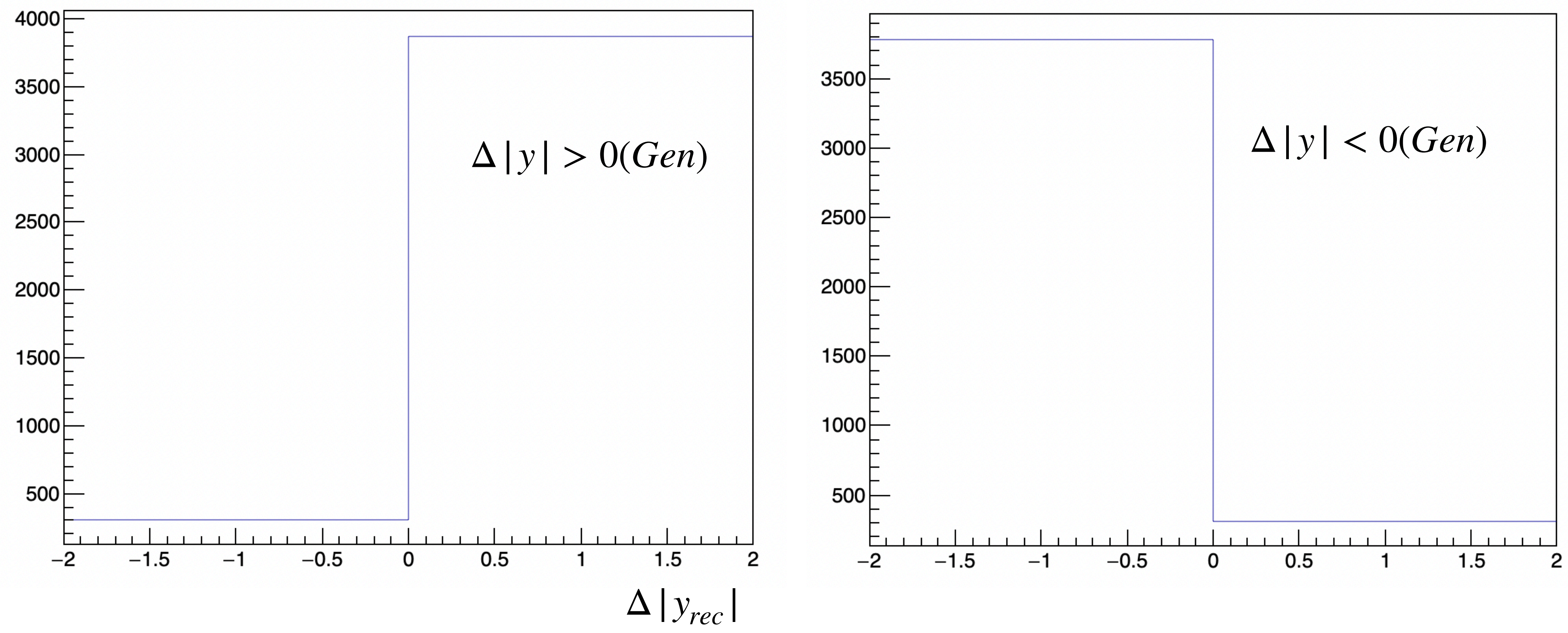
The number of datacards is given by:

2 lepton flavors (electrons and muons)
2 Mass bins (750 Gev < Mttbar < 900 GeV , Mttbar > 900 GeV)
3 years (2016,2017,2018)

In total, we have **12** datacards.
$\Delta|y|$ migrations arte taking into account with the datacard.
Mttbar migrations are included by counting Under Flow and Over Flow bins



No need of regularized unfolding based on the conditional number of each matrix cond(A) < 10



$$A_C^{unf} = \frac{N_{unf}(\Delta|y|>0) - N_{unf}(\Delta|y|<0)}{N_{unf}(\Delta|y|>0) + N_{unf}(\Delta|y|<0)}.$$

CMS UIC University of Illinois at Chicago

Charge asymmetry (Fiducial/Full Phase-space)

Ac is defined as:

$$A_C^{full} = \frac{N_{unf}(\Delta|y| > 0) - N_{unf}(\Delta|y| > 0)}{N_{unf}(\Delta|y| > 0) + N_{unf}(\Delta|y| > 0)}$$

Where N_{unf} is defined as:

$$N_{unf}(\Delta|y| > 0) = r_{pos} \times \frac{N_{truth}(\Delta|y| > 0)}{\alpha\varepsilon^{pos}},$$
$$N_{unf}(\Delta|y| < 0) = r_{neg} \times \frac{N_{truth}(\Delta|y| < 0)}{\alpha\varepsilon^{neg}},$$

Where $\alpha\epsilon$ is the acceptance measured at generator level that corrects back from the fiducial phase space of a given channel to the full phase space. In this situation:

$$A_C^{full} = \frac{\alpha\varepsilon^{neg} \times r_{pos} \times N_{truth}(\Delta|y| > 0) - \alpha\varepsilon^{pos} \times r_{neg} \times N_{unf}(\Delta|y| < 0)}{\alpha\varepsilon^{neg} \times r_{pos} \times N_{unf}(\Delta|y| > 0) + \alpha\varepsilon^{neg} \times r_{neg} \times N_{unf}(\Delta|y| < 0)}.$$

It is more convenient to directly obtain the unfolded charge asymmetry with its corresponding uncertainties with no need for further error propagation.

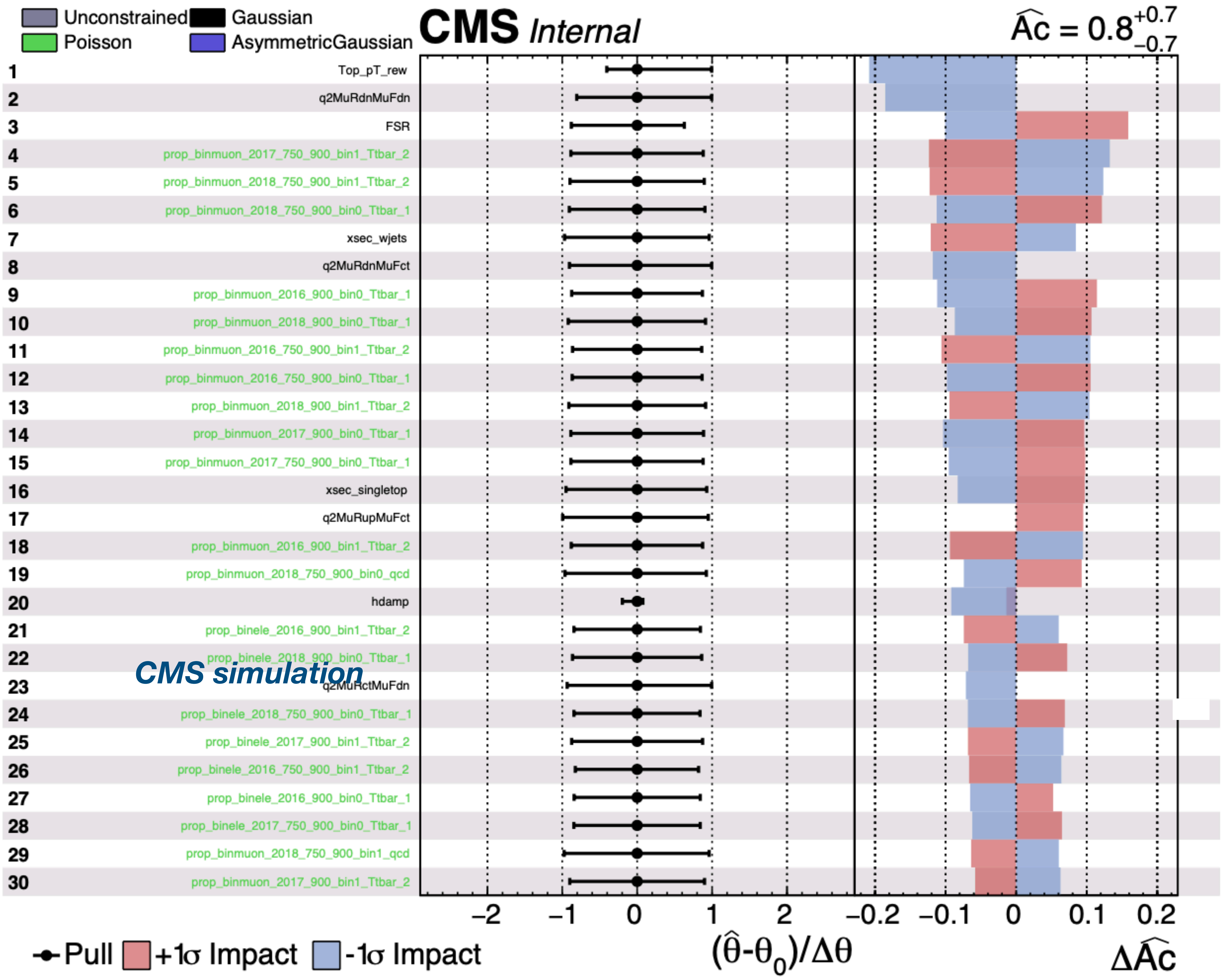
$$r_{pos} = \frac{\alpha\varepsilon^{pos}}{\alpha\varepsilon^{neg}} r_{neg} \times \frac{N_{Truth}(\Delta|y| < 0)}{N_{Truth}(\Delta|y| > 0)} \times \frac{1 + A_C^{full}}{1 - A_C^{full}}$$

CMS UIC University of Illinois at Chicago

systematic uncertainty	uncertainty	type	Signal	Background
luminosity	2.5%, 2.3%, 2.5%	Normalization	Yes	All
pileup reweighing	$\sigma_{minbias} \pm 4.6\%$	Normalization & Shape	Yes	All
muon ID & Reco	$\pm 1\sigma(p_{\mathrm{T}}, \eta)$	Normalization & Shape	Yes	All
muon trigger	$\pm 1\sigma(p_{\mathrm{T}}, \eta)$	Normalization & Shape	Yes	All
electron ID & Reco	$\pm 1\sigma(p_{\mathrm{T}}, \eta)$	Normalization & Shape	Yes	All
electron trigger	$\pm 1\sigma$	Normalization & Shape	Yes	All
Jet Energy Scale	$\pm 1\sigma(p_{\mathrm{T}}, \eta)$	Normalization & Shape	Yes	All
Jet Energy Resolution	$\pm 1\sigma(\eta)$	Normalization & Shape	Yes	All
b-tagging	$\pm 1\sigma(p_{\mathrm{T}}, \eta)$	Normalization & Shape	Yes	All
t-tagging	$\pm 1\sigma(p_{\mathrm{T}})$	Normalization & Shape	Yes	All
t-tagging mistag rate	$\pm 20\%$	Normalization & Shape	Yes	All
W-tagging	$\pm 1\sigma(p_{\mathrm{T}})$	Normalization & Shape	Yes	All
W-tagging mistag rate	$\pm 20\%$	Normalization & Shape	Yes	All
PDFs	NNPDF 3.0, 3.1	Normalization &Shape	Yes	All
$\mathrm{t\bar{t}}$ cross section	5%	Normalization	Yes	No
Single Top cross section	30%	Normalization	No	Single Top
$W+$jets cross section	30%	Normalization	No	$W+$jets
Drell-Yan cross section	30%	Normalization	No	Drell-Yan
Diboson cross section	30%	Normalization	No	Dibosons
QCD cross section	50%	Normalization	No	QCD
μ_R and μ_F at ME level	envelope($\times 2$, $\times 0.5$)	Normalization &Shape	Yes	$W+$jets
top-pT	$\pm 1\sigma$	Normalization &Shape	Yes	No
ISR	$\pm 1\sigma$	Normalization &Shape	Yes	No
FSR	$\pm 1\sigma$	Normalization &Shape	Yes	No
h_{damp}	$\pm 1\sigma$	Normalization &Shape	Yes	No

Table 11: List of systematic uncertainties considered in the analysis. Uncertainties can affect the rate and shape of the signal ($\mathrm{t\bar{t}}$ semileptonic) and/or background processes ($t\bar{t}_{2lep}$, $t\bar{t}_{had}$, Single Top, $W+$jets, Dibosons and Multijet QCD) as indicated in the table.

Impacts l+jets (Mttbar > 750 GeV)



Charge asymmetry in the fiducial phase space (Asimov data)

Summary of the charge asymmetry measurements corrected back to the fiducial phase space.

Channel	$M_{t\bar{t}}$ (GeV)	$N_{unfold}(\Delta\|y\| > 0)$	$N_{unfold}(\Delta\|y\| < 0)$	A_C^{fid} (%)
ℓ + jets Full Run II	≥ 750	69410	68351	$0.76^{+0.70}_{-0.70}$
ℓ + jets Full Run II	[750 − 900]	32646	32252	$0.6^{+[illegible]}_{-1.06}$
ℓ + jets Full Run II	> 900	36764	36099	$0.91^{+1.01}_{-1.01}$
μ + jets Full Run II	≥ 750	54239	53446	$0.73^{+0.7}_{-0.8}$
e + jets Full Run II	≥ 750	15171	14905	$0.88^{+1.3}_{-1.45}$

Ac_13TeV_nnlo+ew theory calculations

https://arxiv.org/abs/1711.03945

Mttbar.	Nominal	Up	Down
750-900	8.7361e-03	6.2000e-04	-6.7930e-04
>900	1.0133e-02	5.7245e-04	-7.3356e-04
>750	8.9668e-03	2.5252e-03	-1.4078e-03

ATLAS-CONF-2019-026

	Data 139 fb^{-1}					
	A_C	Stat.	Syst.	MC stat.	Bias	Total unc.
Inclusive	0.0060	0.0011	0.0009	0.0005	0.0001	0.0015
$m_{t\bar{t}}$ < 500 GeV	0.0045	0.0028	0.0034	0.0013	0.0001	0.0045
500-750 GeV	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
750-1000 GeV	0.0100	0.0049	0.0046	0.0021	0.0001	0.0070
1000-1500 GeV	0.0169	0.0072	0.0027	0.0029	0.0004	0.0083
> 1500 GeV	0.0121	0.0277	0.0150	0.0092	0.0005	0.0329

University of Illinois at Chicago

Conclusions

- First CMS measurement of the top quark charge asymmetry at 13TeV
- Novel compared with the previous results because it focuses exclusively on the very high Lorentz boost regime, using dedicated reconstruction techniques not just for the hadronically-decaying top quark (AK8 top and W tag) but also for the leptonically-decaying top quark
- Competitive result with the (unpublished ATLAS-CONF-2019-026) ATLAS result at 13TeV

Backup

Forward-Backward Asymmetry in top pair production at the Tevatron

- Investigation of the charge asymmetry in heavy quark production was performed at the Tevatron accelerator by CDF and D0 experiments.

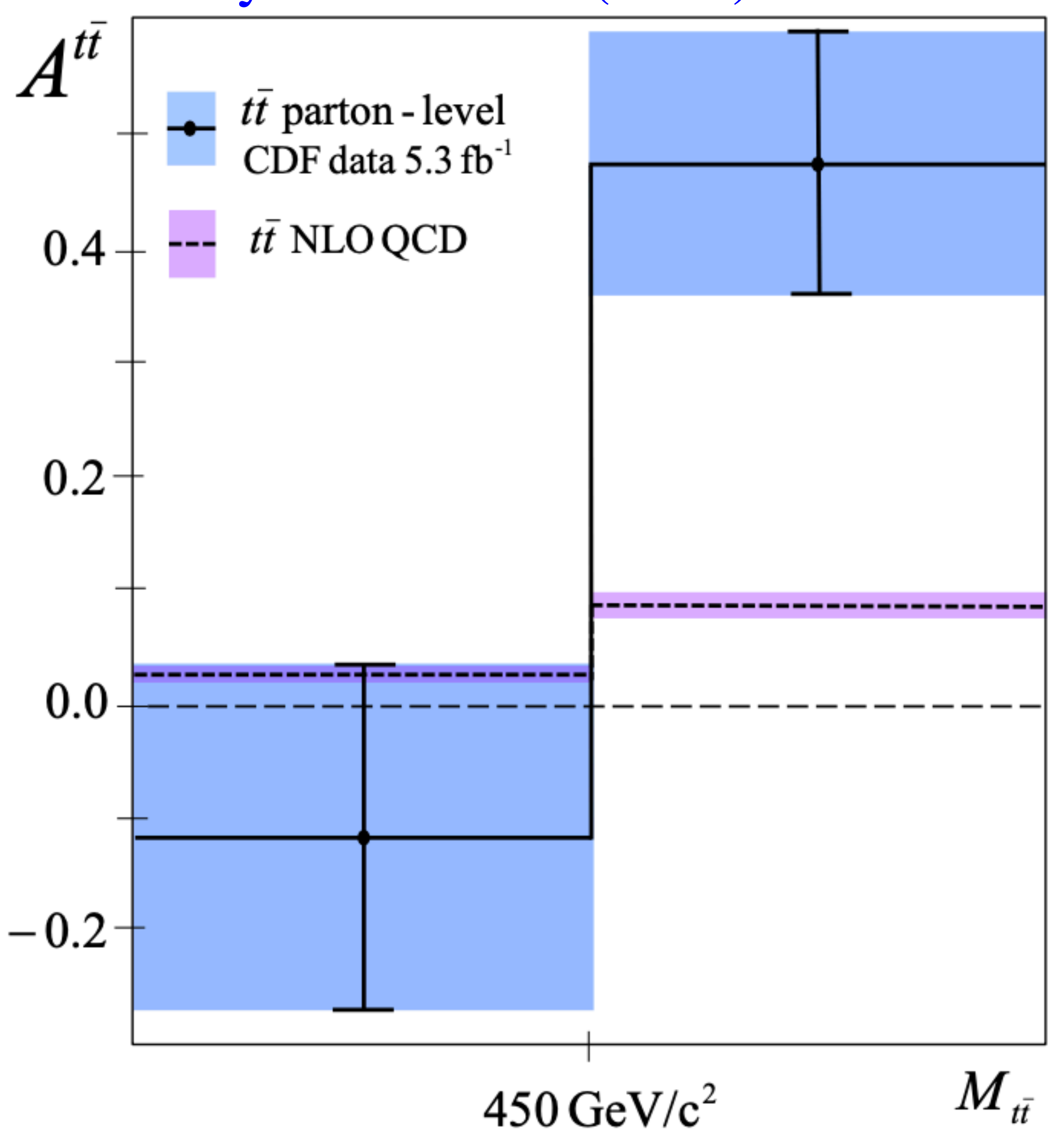
- Tevatron was a very suitable collider for studying $t\bar{t}$ charge asymmetry due to the dominant $q\bar{q} \rightarrow t\bar{t}$ production channel.

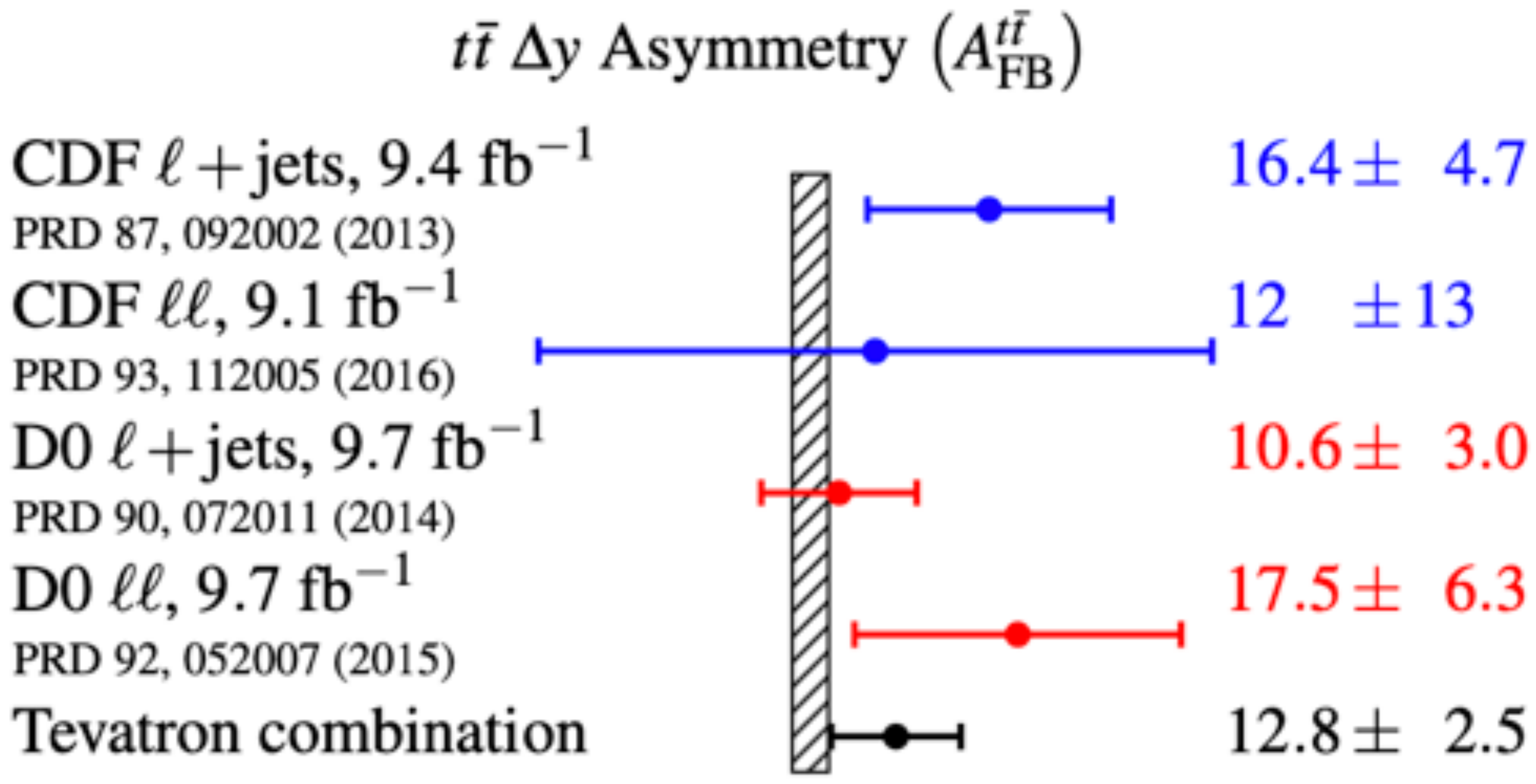
- Asymmetry, defined as:

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

- First results measured in D0 and CDF disagree with the MC@NLO based predictions, with most significant discrepancy above 3σ.

Phys. Rev. D **83** (2011) 112003.

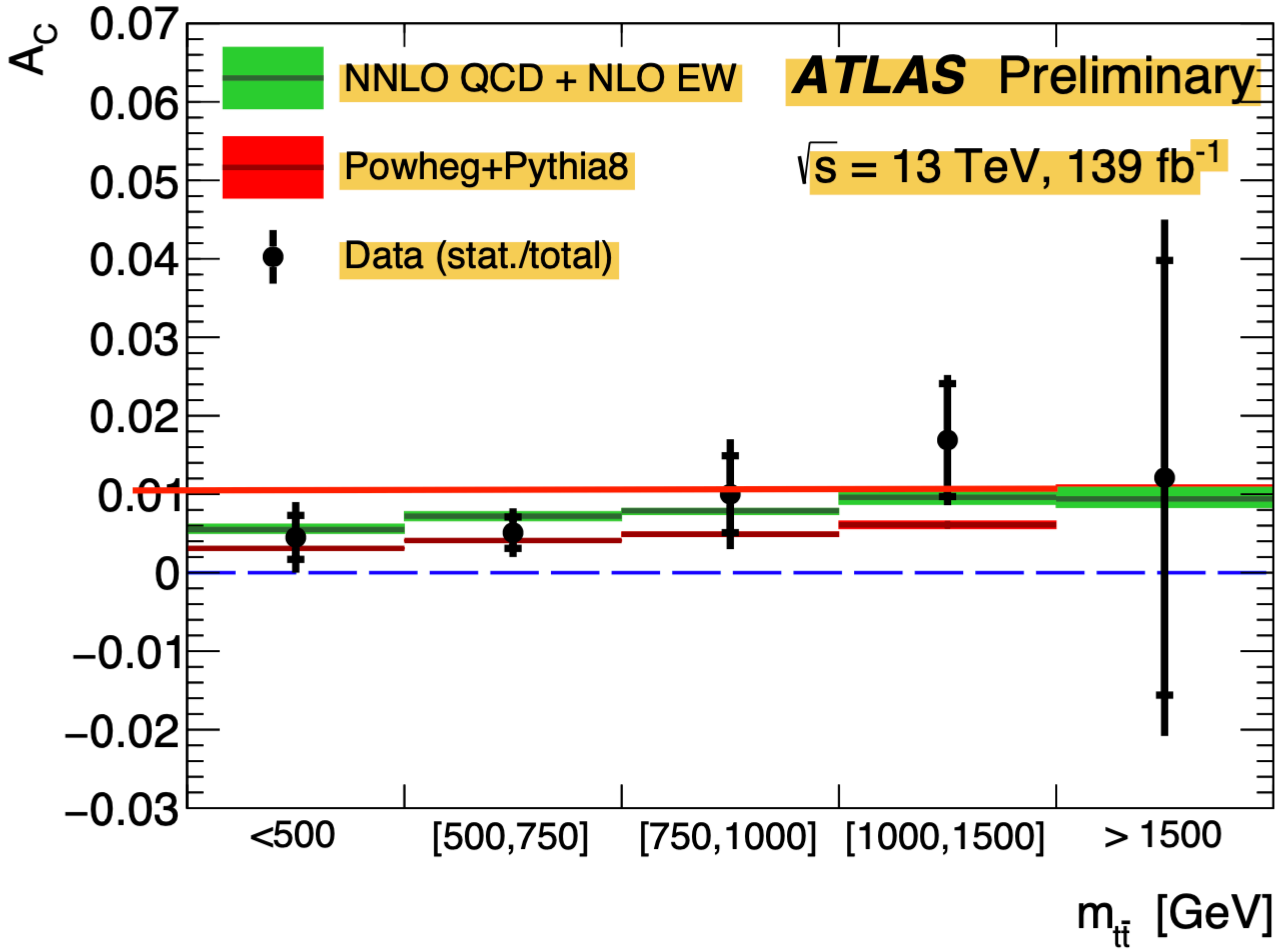




Phys. Rev. Lett. 120 (2018) 042001.

- ATLAS-CONF-2019-026

		Data 139 fb^{-1}					
		A_C	Stat.	Syst.	MC stat.	Bias	Total unc.
Inclusive		0.0060	0.0011	0.0009	0.0005	0.0001	0.0015
$m_{t\bar{t}}$	< 500 GeV	0.0045	0.0028	0.0034	0.0013	0.0001	0.0045
	500-750 GeV	0.0051	0.0020	0.0021	0.0009	<0.0001	0.0031
	750-1000 GeV	0.0100	0.0049	0.0046	0.0021	0.0001	0.0070
	1000-1500 GeV	0.0169	0.0072	0.0027	0.0029	0.0004	0.0083
	> 1500 GeV	0.0121	0.0277	0.0150	0.0092	0.0005	0.0329

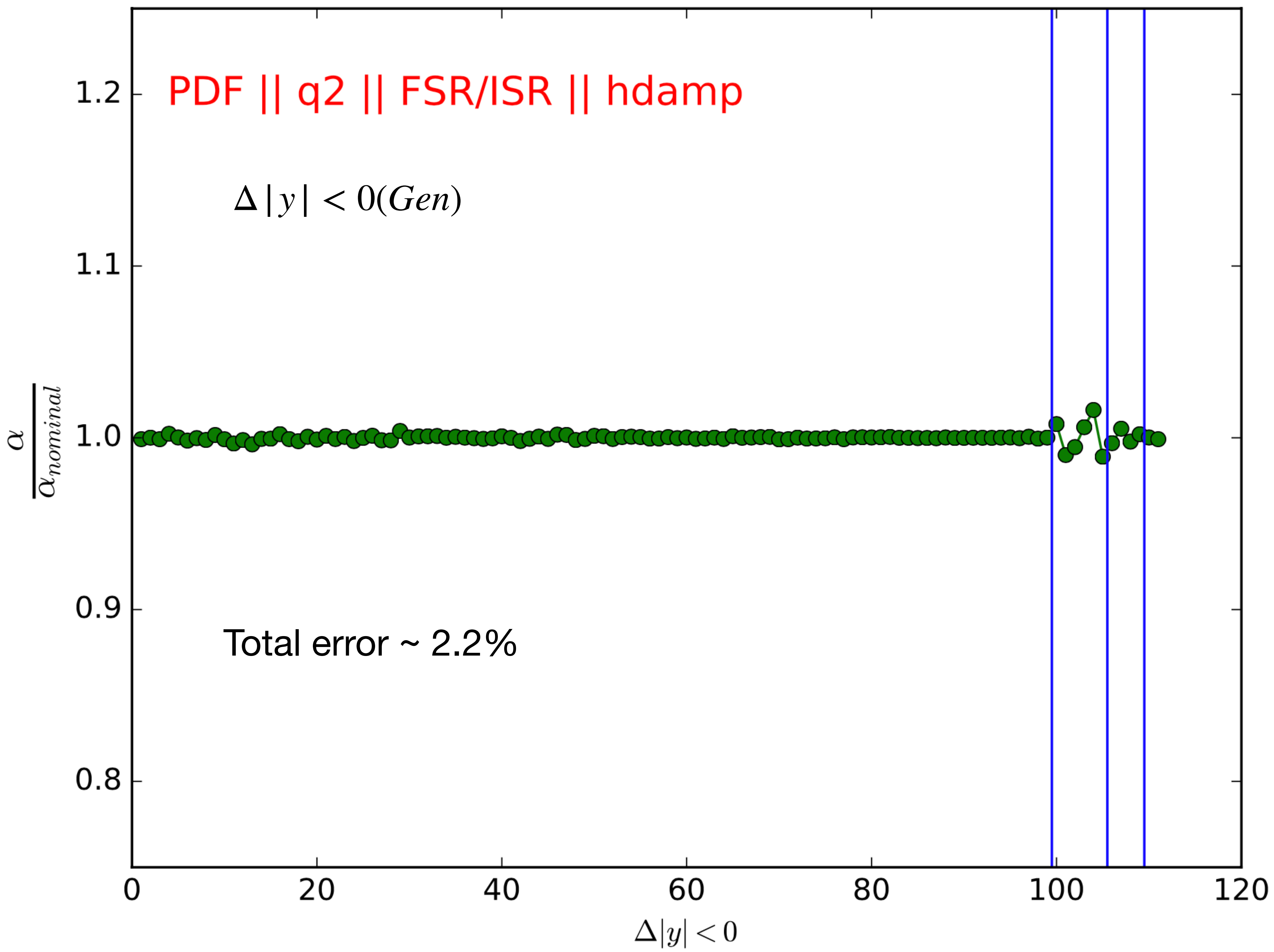


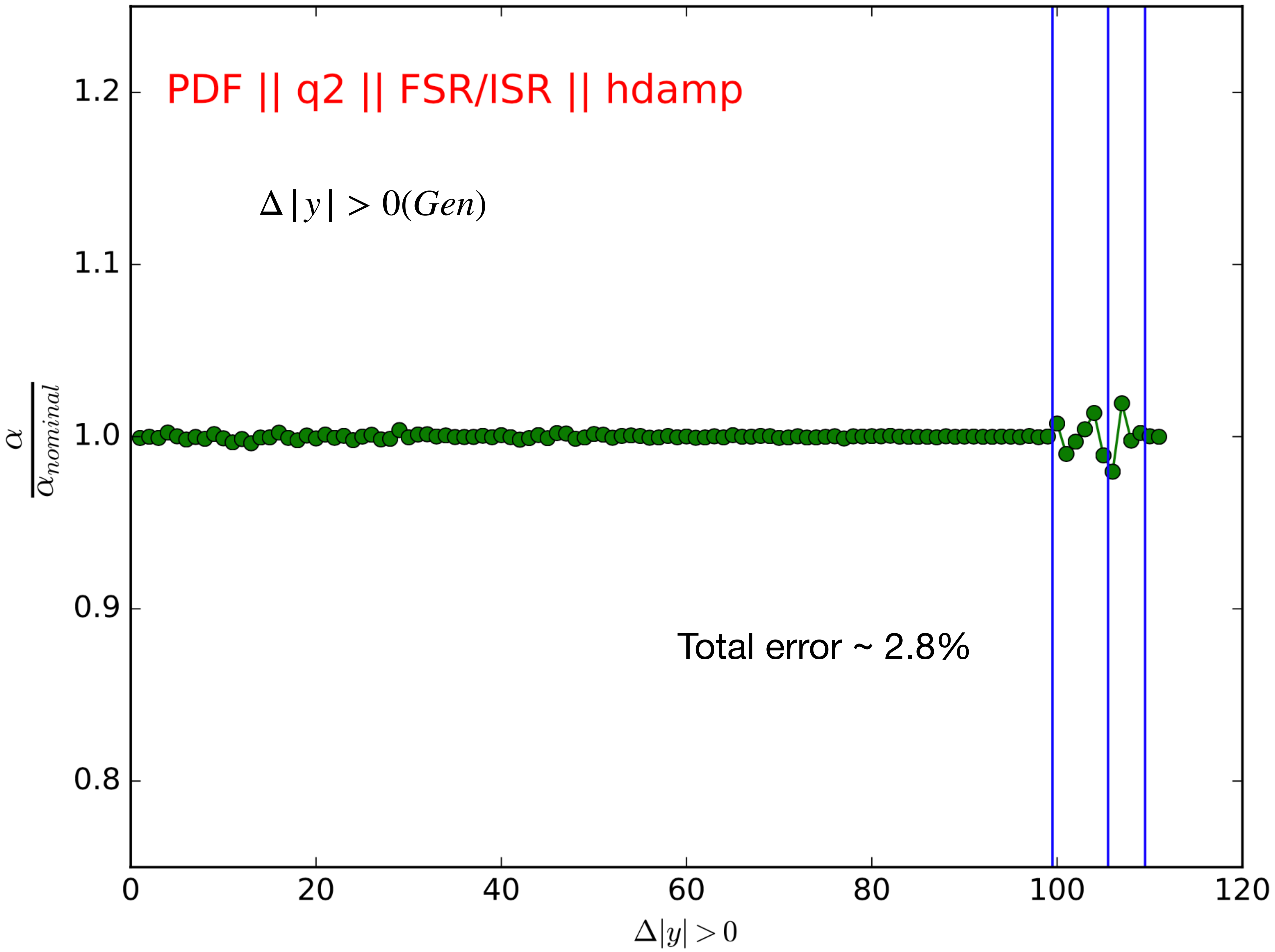
Acceptance x efficiency error

- ~ 10,000,000 events per lepton flavor (2016,2017,2018)

TTToSemiLeptonic_TuneCP5_13TeV_pythia8

- We calculate the ratio between the nominal acceptance and the acceptance affected by theoretical weights (FSR/ISR, hdamp, PDF, q2).





μ+jets channel 750 GeV < Mttbar < 900 GeV (other channels have similar results)