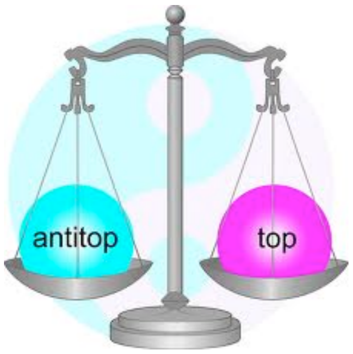
Measurement of the charge asymmetry in top quark pair production at 7 TeV

Ioana Anghel

University of Illinois at Chicago
(on behalf of the CMS Collaboration)



XXI International Workshop on Deep-Inelastic Scattering and Related Subjects, Newport News 2011

Outline

I. **Charge Asymmetry in top pair production in p-p collisions**

II. **Charge Asymmetry Measurement**

Event Selection

Top Quarks Reconstruction

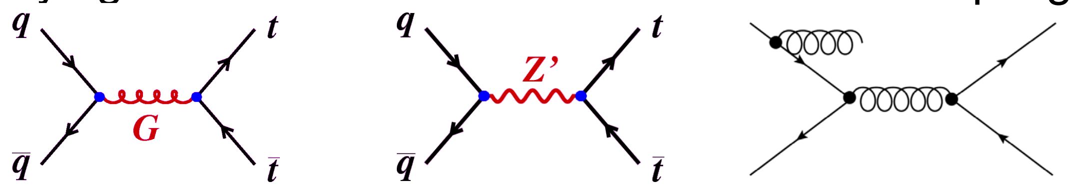
Correction of the Measurement Effects

III. **Systematics**

IV. **Summary**

Charge Asymmetry

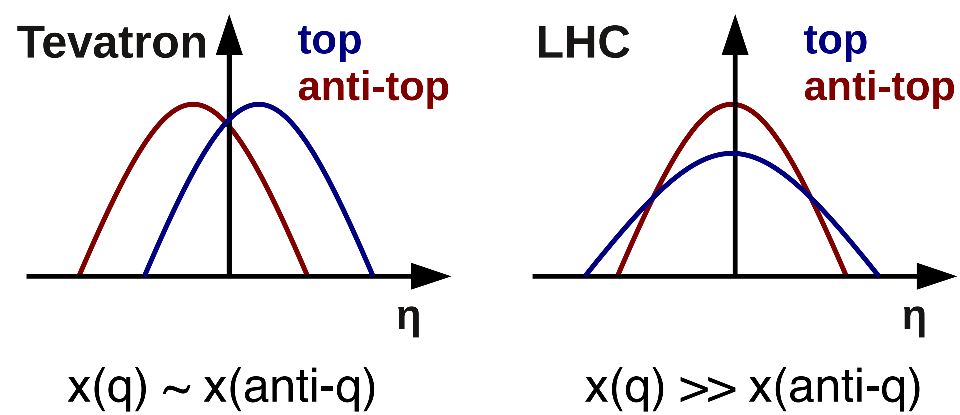
- The charge asymmetry provides the possibility to search for the unknown top quark production mechanisms
- Sources of asymmetry: new particles decaying to $t\bar{t}$ with different vector and axial couplings, interference with SM processes



  - top (anti-top) quarks preferably emitted in quark (anti-quark) direction

- The asymmetry appears only for asymmetric initial states - at LHC, the asymmetric initial states like $q\bar{q}$ are still possible, with the anti-quark coming from the sea quarks, but these events have a much smaller probability

- Proton-antiproton (Tevatron) - quark (anti-quark) are mainly from proton (antiproton).
  - the measured forward-backward asymmetry shows 2-3 σ deviation from the SM prediction



Sensitive Variables to Charge Asymmetry in p-p collisions

- Most theoretical predictions use the rapidity to calculate the charge asymmetry:

$$A_C(y_C) := \frac{N_t(|y| < y_C) - N_{\bar{t}}(|y| < y_C)}{N_t(|y| < y_C) + N_{\bar{t}}(|y| < y_C)}$$

In addition, cuts on rapidity and $m_{t\bar{t}}$ are used.

- We used variable $|\eta_t| - |\eta_{\bar{t}}|$ to determine its asymmetry:

$$A_C = \frac{N^+ - N^-}{N^+ + N^-}$$ where N^+ / N^- are the number of events with positive/negative values of $|\eta_t| - |\eta_{\bar{t}}|$

- we used the pseudo-rapidity because it is based on angular information only, providing a good resolution

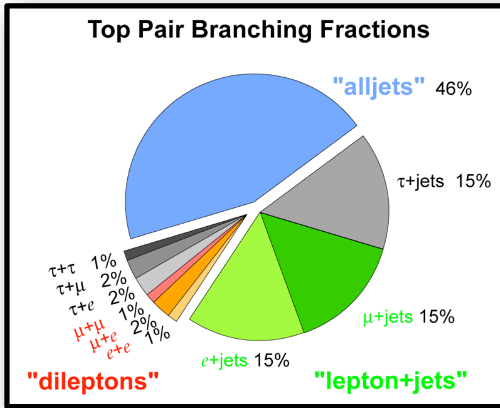
- no additional cuts were used, due to limited statistics

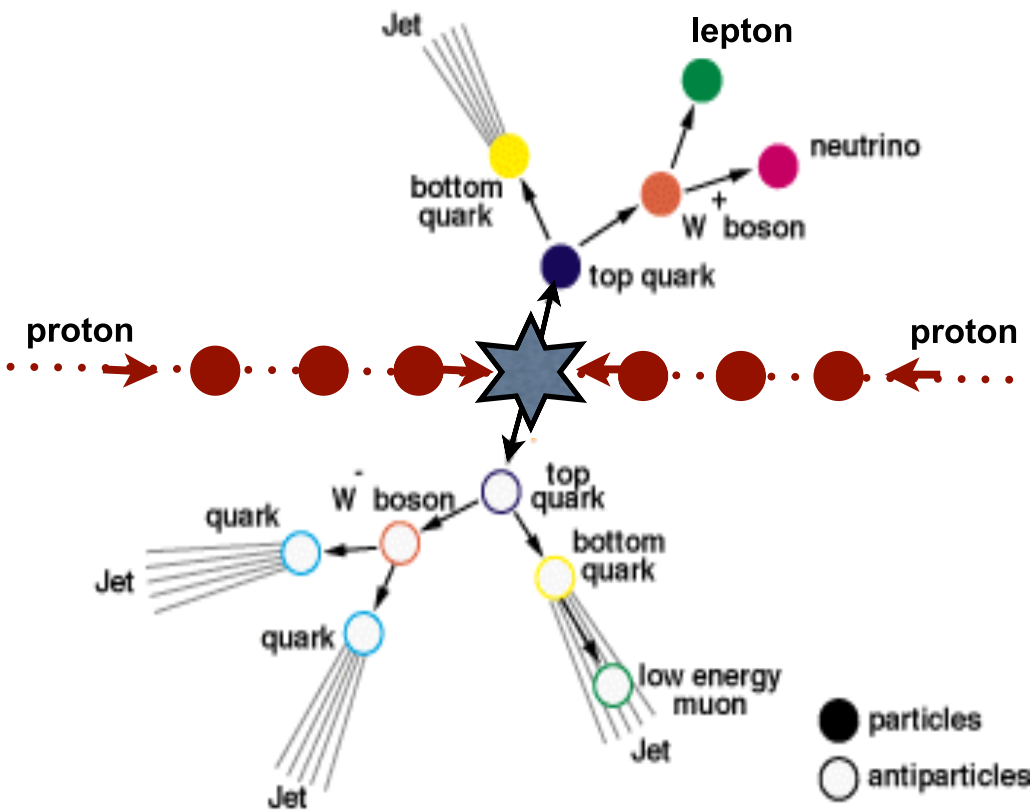
- A_C predictions:

- in SM: $A_C^{SM} = 0.0121$ G. Rodrigo et al., doi:10.1007/JHEP02(2010)051

Event Selection

- Data ~ 36 pb^{-1}
- We selected events in the semi-leptonic channel (muon and electron)





- **Event selection**:

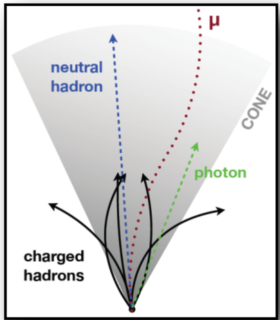
  - exactly one primary vertex (PV)

  - exactly one isolated muon or electron originating from the PV

  - missing transverse energy (neutrino)

  - *particle flow jets* reconstructed using anti-k_T algorithm (4 jets)

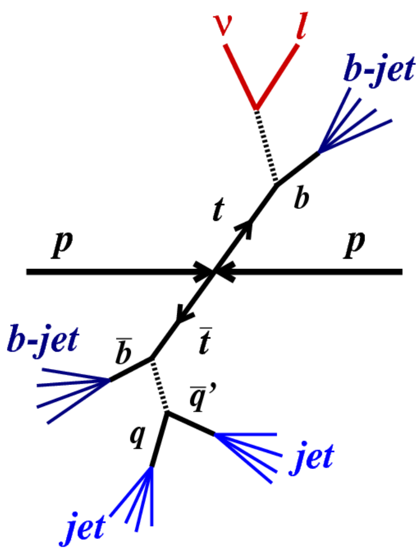
Particle flow jets - from identified particles using all detector components



Reconstruction of the Top 4-vector

- The measurement of pseudo-rapidity for the top (anti-top) quarks requires the reconstruction of the 4-vector momentum:

  - not trivial in data: we need to rely on the measured quantities for the final states(lepton, missing transverse energy and jets)



- The reconstruction involves several hypotheses tested in MC - constructed based on associating jets to the 2 top quarks. The hypothesis selection is improved by using the b-tagging
- Ψ **variable** - constructed for each hypothesis, the hypothesis with the **smallest** Ψ value is chosen

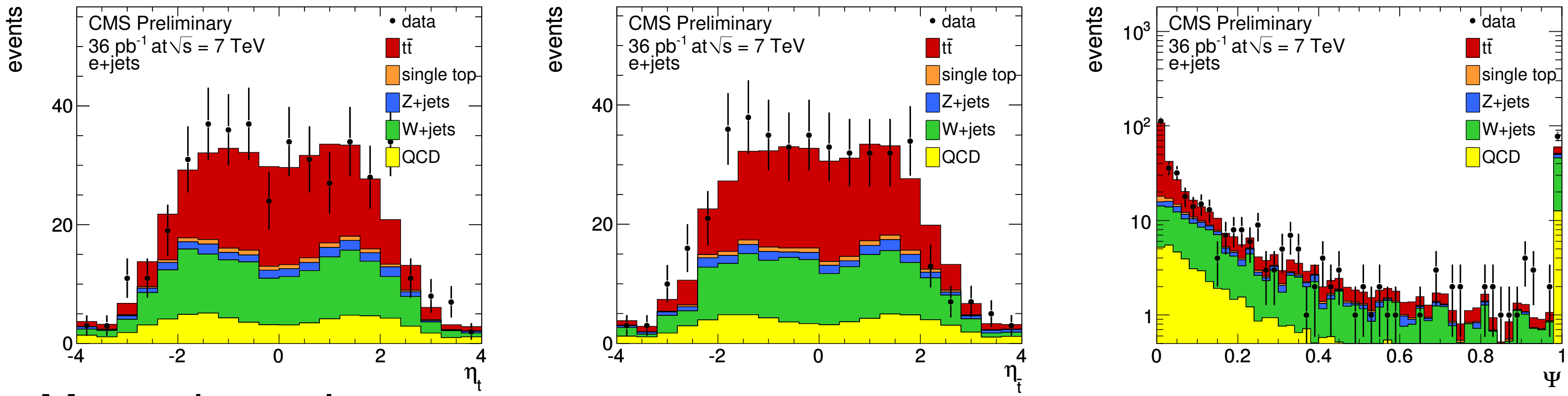
$$\Psi = \left[\frac{(M^{rec}_{W,had} - M^{b.p}_{W,had})^2}{\sigma^2_{W,had}} + \frac{(M^{rec}_{t,had} - M^{b.p}_{t,had})^2}{\sigma^2_{t,had}} + \frac{(M^{rec}_{t,lep} - M^{b.p}_{t,lep})^2}{\sigma^2_{t,lep}} \right] * P_b(q_1) * P_b(q_2) * (1 - P_b(q_1)) * (1 - P_b(q_2))$$

where $M_i^{b.p}$ and σ_i are determined from the best possible hypothesis on MC, while $P_b(q_i)$ is the probability that a certain jet q is assigned to a b quark and not to one of the light quarks in the best possible hypothesis

Data/MC Comparison for Reconstructed Objects

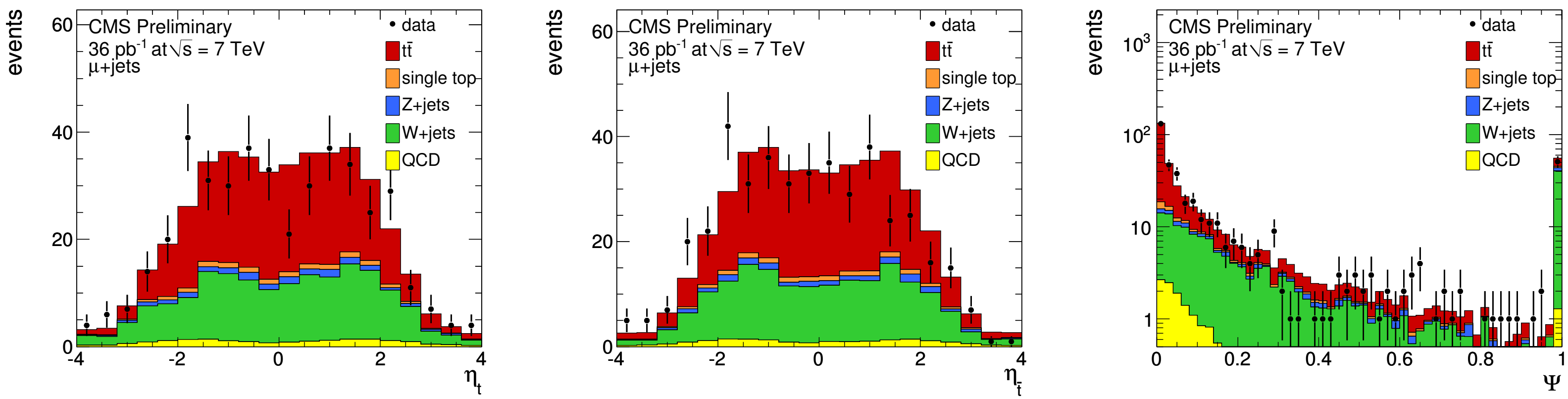
- Electron channel

chosen hypothesis



- Muon channel

chosen hypothesis



Good agreement between data and MC

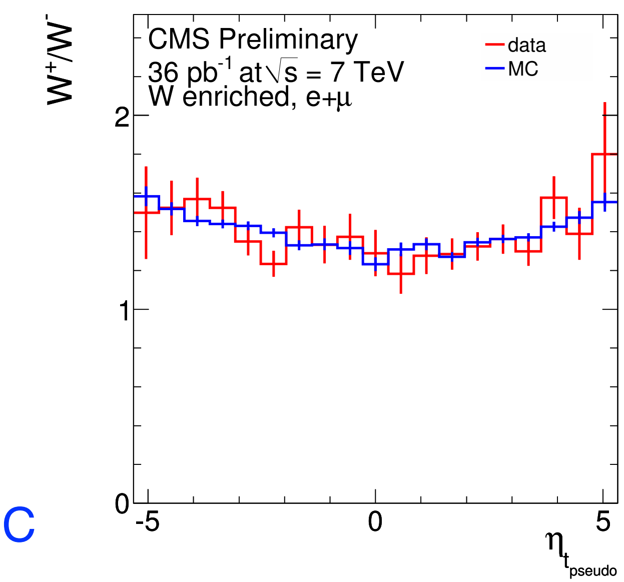
- QCD is modeled using a data-driven method

- all other templates are taken from MC simulations

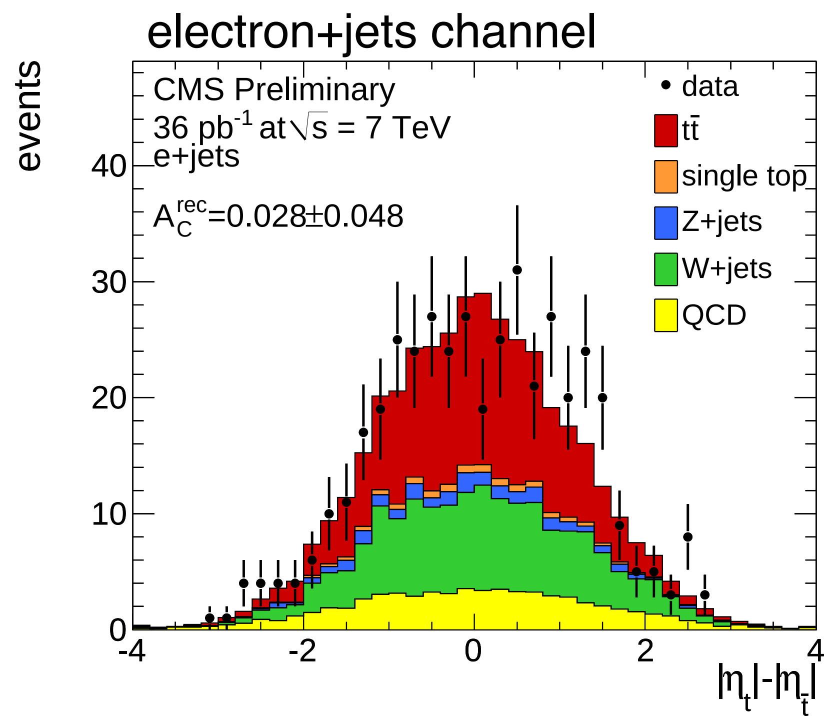
- signal and background processes are normalized to the fit result of the cross section measurement

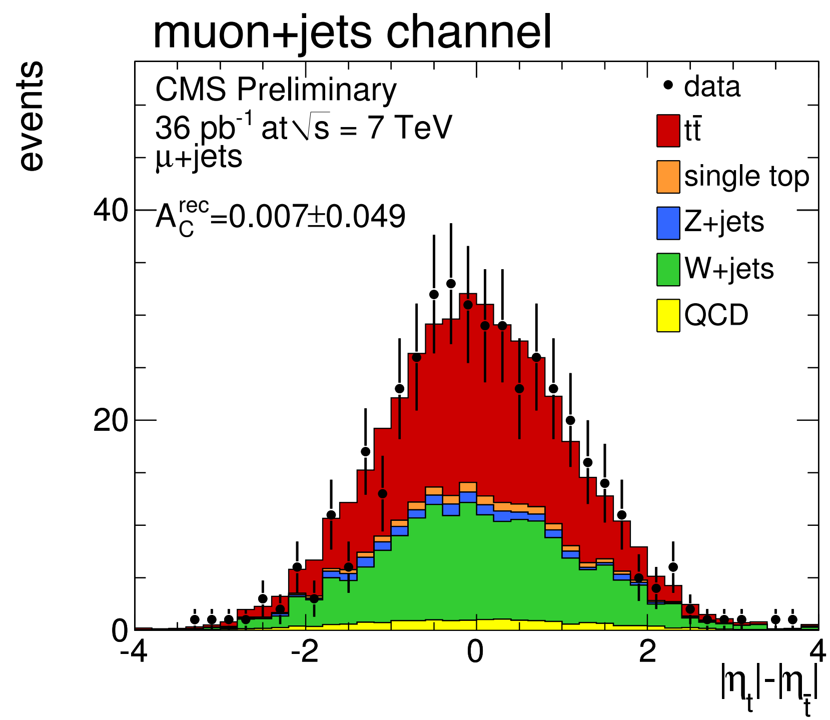
W+jets background asymmetry

- After selection, 50% of the events are top quark pair events, the background being dominated by W+jets
- W^+ and W^- events show different shape for $|\eta_t| - |\eta_{\bar{t}}|$ in theory (expected)
- Need to understand if there is an asymmetry in W+jets events which is not described by MC
- We apply a slightly different selection to obtain an enriched sample in W+jets:
  - one lepton
  - only 1 or 2 jets in the event
  - a cut threshold for the transverse mass of the reconstructed W (to reduce the QCD background)
- We reconstruct a "pseudo"-top (leptonic branch) and compare the ratio of the $\eta_{t_{pseudo}}$ distribution for W^+ and W^- between data and MC
- The data and MC ratio distributions looks very similar - the differences between W+ and W- events are correctly described by MC



Uncorrected A_C





- By combining the electron+jets and muon+jets channels: **A_C^{rec}=0.018±0.034(stat)**

- The measured value may be distorted by several effects:

  - the **amount of background .** When subtracted, **$A_C^{bkg\text{-}subtr}$=0.035±0.070(stat)**

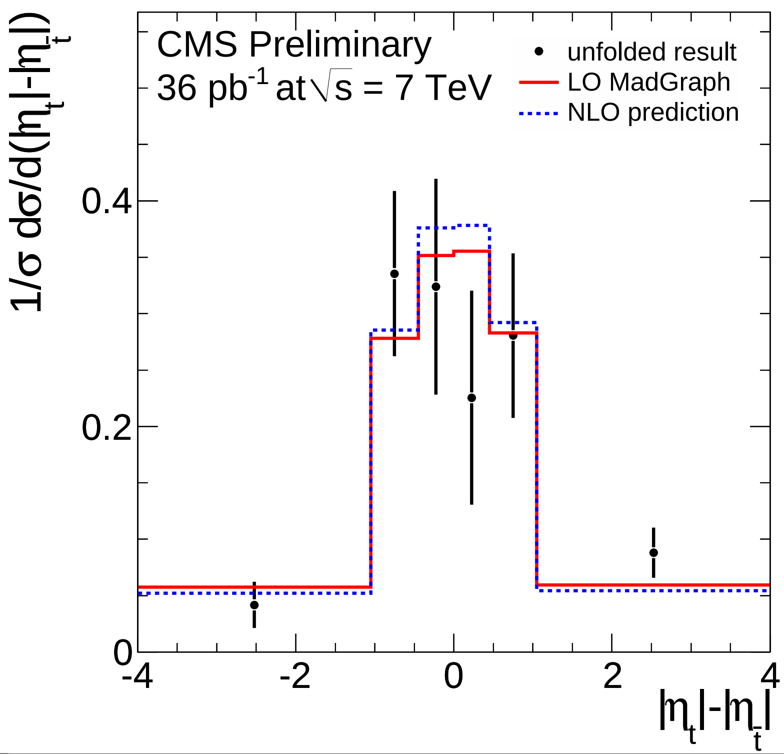
  - **smearing**: reconstructed $|\eta_t| - |\eta_{\bar{t}}|$ does not directly correspond to true value

  - **lepton efficiency**

Smearing Effects Correction

- This correction is done with a **regularized unfolding** technique:

  - from the smearing performed for the MADGRAPH generated $t\bar{t}$ sample, the smearing matrix A between true and reconstructed $|\eta_t| - |\eta_{\bar{t}}|$ is extracted
  - the smearing matrix A also account for efficiency losses
  - the matrix A connects the measured spectrum $\overrightarrow{y}$ to the true spectrum $\overrightarrow{x}$: $\overrightarrow{y} = A\overrightarrow{x}$
  - equation is solved for $\overrightarrow{x}$ through a generalized process to invert the matrix A



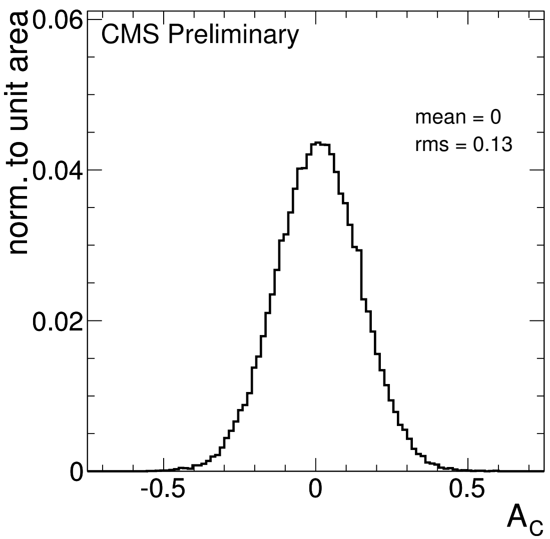
The unfolded $|\eta_t| - |\eta_{\bar{t}}|$ spectrum and theory curves agree within errors

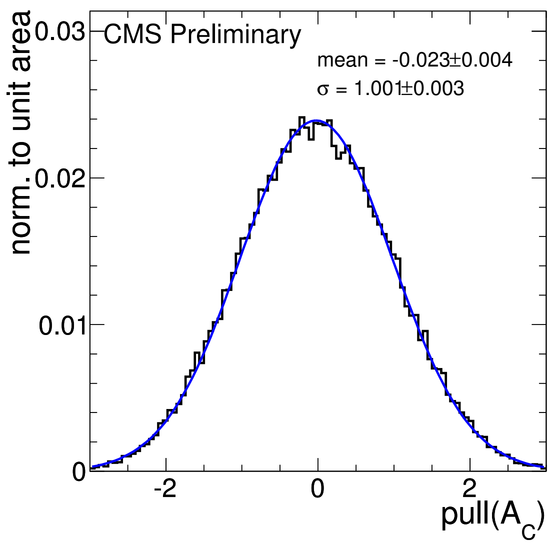
A_C=0.060±0.134(stat)

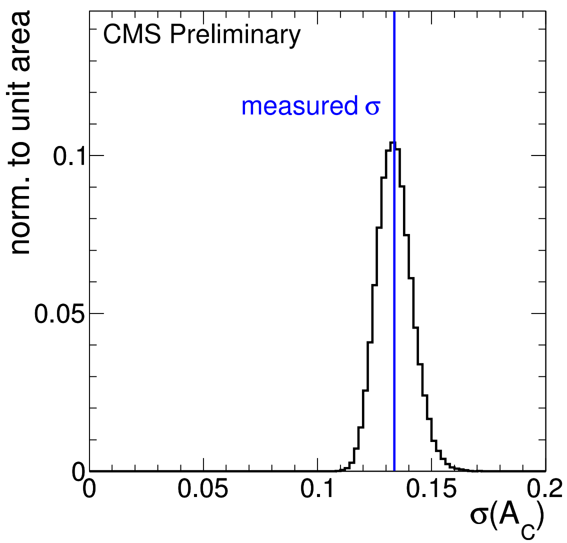
- from SM: $A_C^{SM} = 0.0121$

Unfolding Validation

- the method is tested with pseudo-experiments.
- for each experiment, pseudo data is constructed from random distributions of the MC samples (Poisson distribution, involving also the statistical uncertainties from the covariance matrix)







$$pull(A_C) = \frac{A_C^{true} - A_C^{rec}}{\sigma_{unfolding}}$$

The pull width is ~ 1, meaning that the error propagation is done correctly by the unfolding method.

Systematics Results

- Only systematics influencing the direction of the reconstructed top quark momenta can change the charge asymmetry value.
- Systematics are evaluated by drawing pseudo-experiments from systematically shifted templates

source of systematic	positive shift in A_C	negative shift in A_C
jet energy scale	0.017	-
jet energy resolution	0.007	-0.006
Q^2 scale	0.003	-0.007
ISR/FSR	0.005	-0.0006
matching threshold	0.004	-0.006
PDF	0.004	-0.011
b tagging	0.007	-
lepton efficiency	0.017	-0.018
QCD model	0.005	-0.005
overall	$\pm$0.026	

$$A_C = 0.060 \pm 0.134(\text{stat})^{+0.028}_{-0.025}(\text{syst})$$

Summary

- We developed a method to measure the top pair charge asymmetry A_C in proton-proton collisions with the CMS detector in the lepton+jets decay channel.
- For the A_C measurement, the spectrum of $|\eta_t| - |\eta_{\bar{t}}|$ is corrected for selection and reconstruction inefficiencies using an unfolding technique.
- The measured value A_C obtained for 36 pb^{-1} of data is dominated by the statistical uncertainty and is consistent with the Standard Model prediction.

$A_C = 0.060 \pm 0.134(stat)^{+0.028}_{-0.025}(syst)$

- With a larger data set (> 1fb^{-1}), the measurement of the top quark charge asymmetry can reach the same sensitivity as the Tevatron results.

Reference: CMS Physics Analysis Summary TOP-10-010

Back-up slides

QCD templates

- MC QCD templates:

  - hard to model isolated fake leptons in MC

  - modeled with Pythia 2-to-2 - insufficient statistics is available, especially for high jet multiplicity bins

- QCD templates are obtained from data, using sidebands regions:

  e+jets - use events passing at least two of the following criteria:

  i) 0.1< relative isolation < 0.5

  ii) impact parameter (measured wrt beam spot) > 0.02 cm

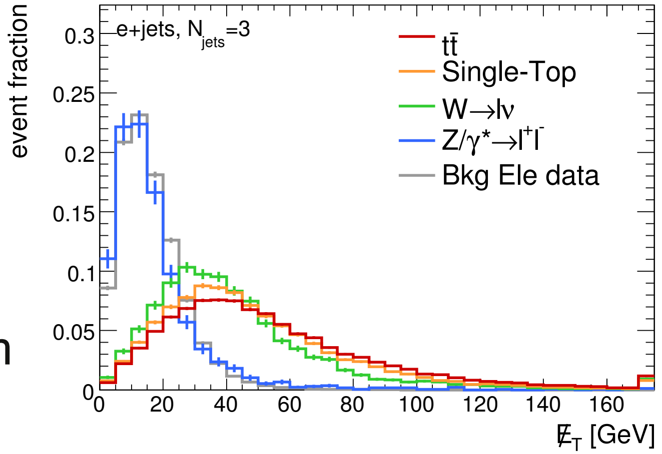
  iii) fail standard tight electron identification

  muon+jets - use events satisfying:

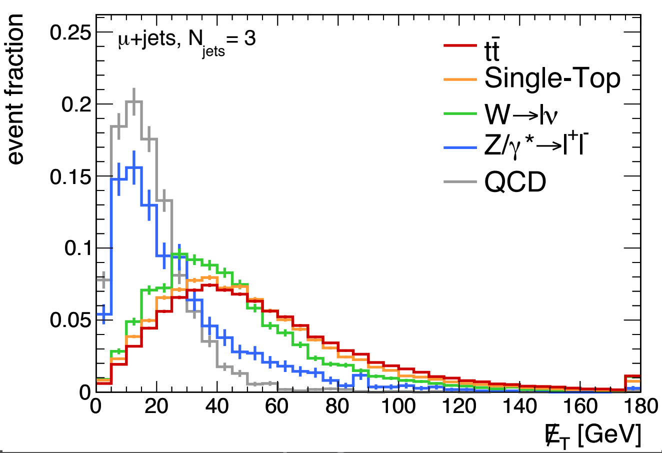
  i) 0.2 < relative isolation < 0.5

- Discriminating variables between signal and background: missing transverse energy and M3

CMS Physics Analysis Summary TOP-10-002

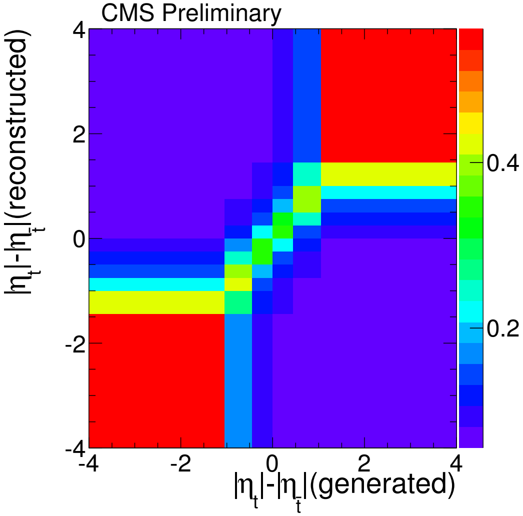


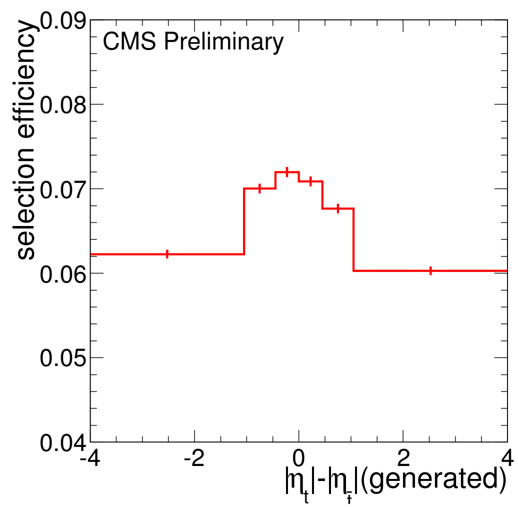
CMS Physics Analysis Summary TOP-10-002



Smearing and Selection Effects

- Selection efficiency - not flat. This can affect the charge asymmetry measurement
- Matrix A - from MADGRAPH generated top pair events. The resolution for $|\eta_t| - |\eta_{\bar{t}}|$ is ~ 0.5.





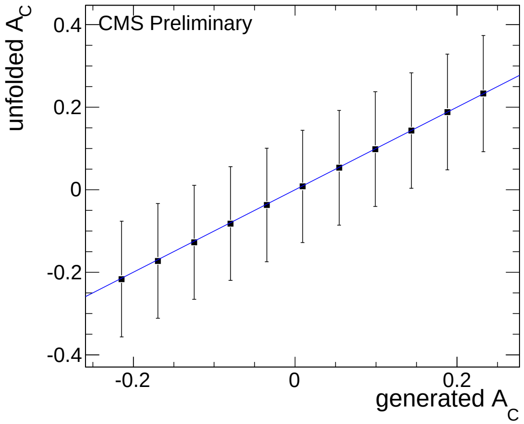
The result of the unfolding is a full covariance matrix for the measured spectrum. The elements of the covariance matrix are relative bin uncertainties multiplied with the corresponding bin by bin correlation $\sigma_i\sigma_j * Cor(i,j)$

CMS Preliminary

	bin 1	bin 2	bin 3	bin 4	bin 5	bin 6
bin 1	0.241	-0.011	-0.070	-0.031	0.007	0.000
bin 2	-0.010	0.048	0.027	-0.040	-0.022	0.003
bin 3	-0.070	0.027	0.087	0.027	-0.034	-0.013
bin 4	-0.032	-0.040	0.026	0.177	0.043	-0.053
bin 5	0.006	-0.023	-0.035	0.043	0.068	-0.003
bin 6	0.001	0.003	-0.013	-0.052	-0.002	0.064

Unfolding Checks

- **linearity check**: construct pseudo data with MC samples with different $|\eta_t| - |\eta_{\bar{t}}|$ asymmetries.

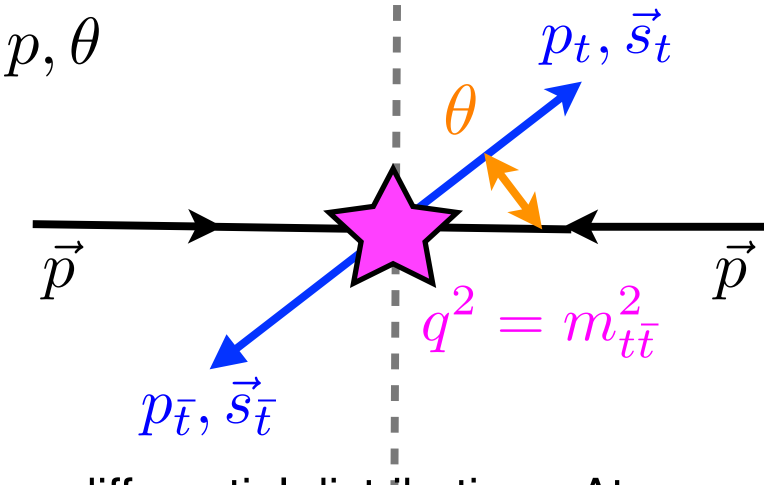


- the error bars correspond to the mean of the expected statistical uncertainty in each bin

Generated asymmetry is found correctly with the unfolding method.

Top Quark Pair Production at LHC

★ Parameters characterizing a top quark pair event: $q^2, \sigma_{t\bar{t}}, p, \theta$



$q^2, \sigma_{t\bar{t}}$ - measured to be within the SM

$m^2_{t\bar{t}}$ - reconstructed, might show if BSM particles are present

- wide distribution, which hides such resonances

★ At low order, there are no differences between top and anti-top differential distributions. At higher order (α^3), the distributions are not longer equal.