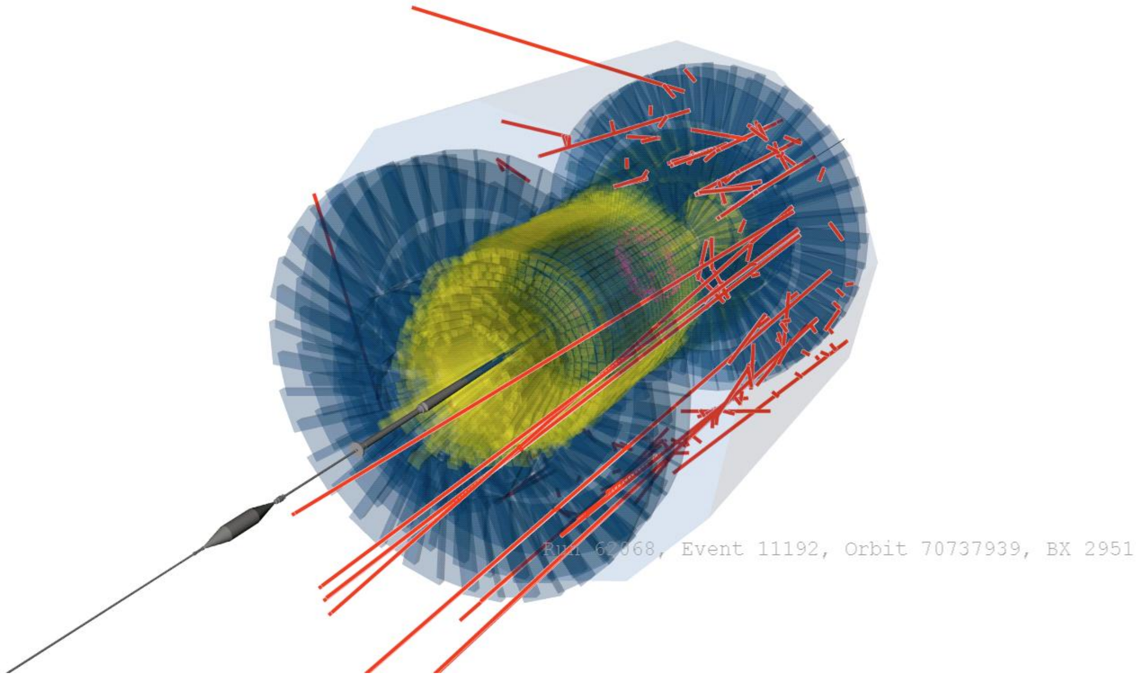


INSTRUMENTAL EFFECTS AND DATA-DRIVEN BACKGROUND ESTIMATES

Joe Boudreau (ATLAS/University of Pittsburgh)

on behalf of ATLAS, CMS, CDF, D0 experiments

Top 2012

This talk covers these main topics

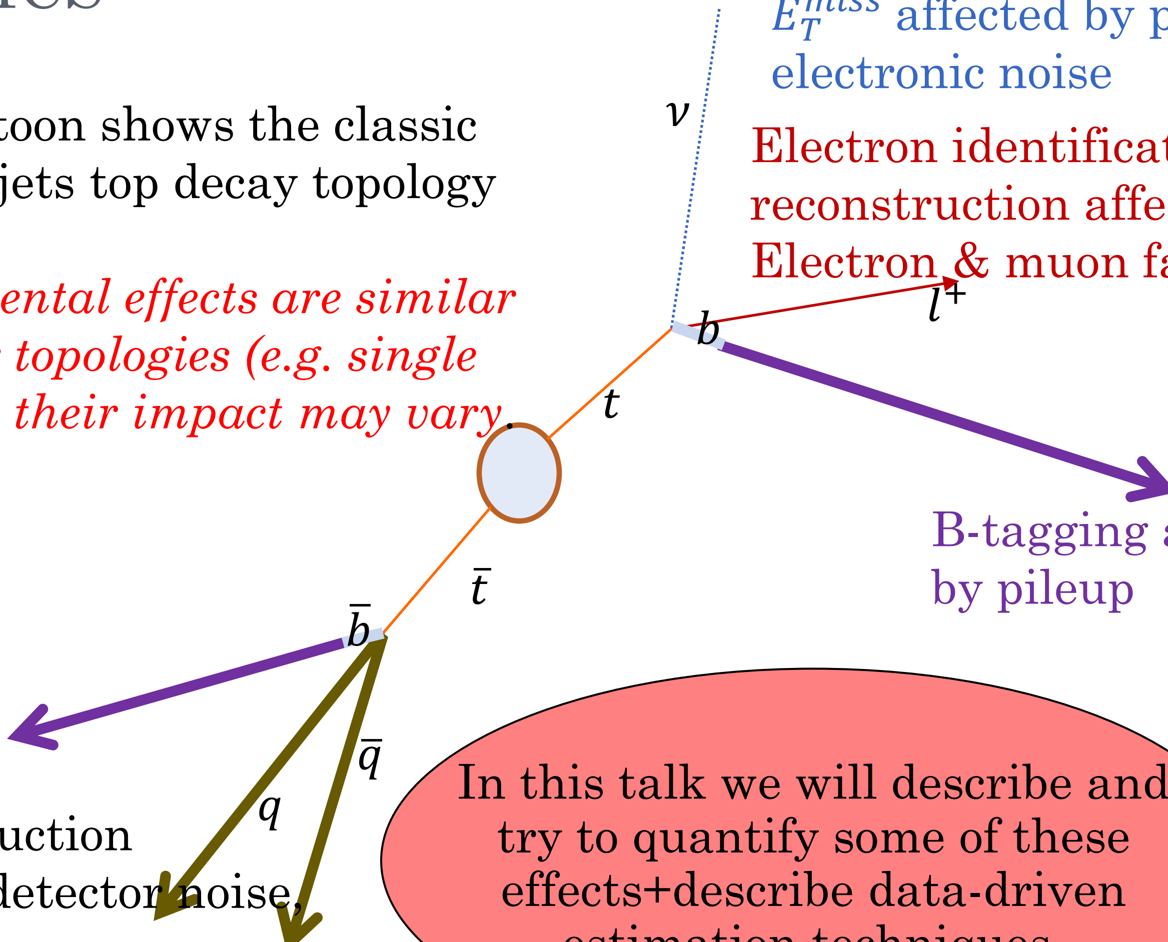
- Instrumental effects & instrumental backgrounds
- Evolution of pileup & its impact on analysis
- Estimation of fake lepton contamination

Impact of these effects on top physics

- this cartoon shows the classic lepton+jets top decay topology
- *instrumental effects are similar in other topologies (e.g. single top) but their impact may vary.*

E_T^{miss} affected by pileup and electronic noise

Electron identification + reconstruction affected
Electron & muon fakes arise

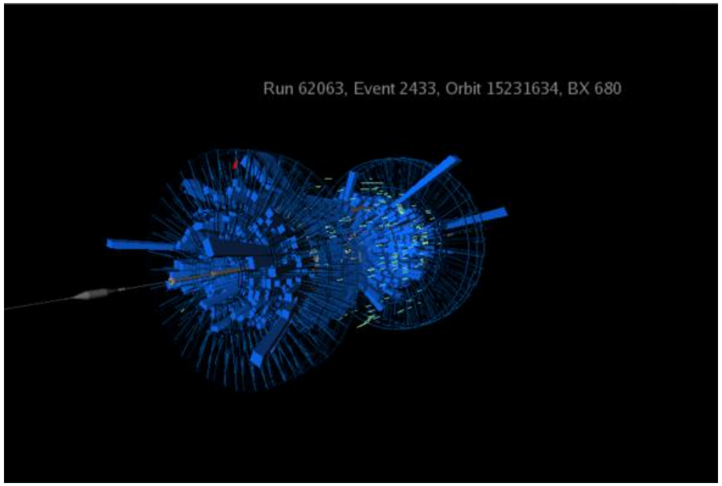


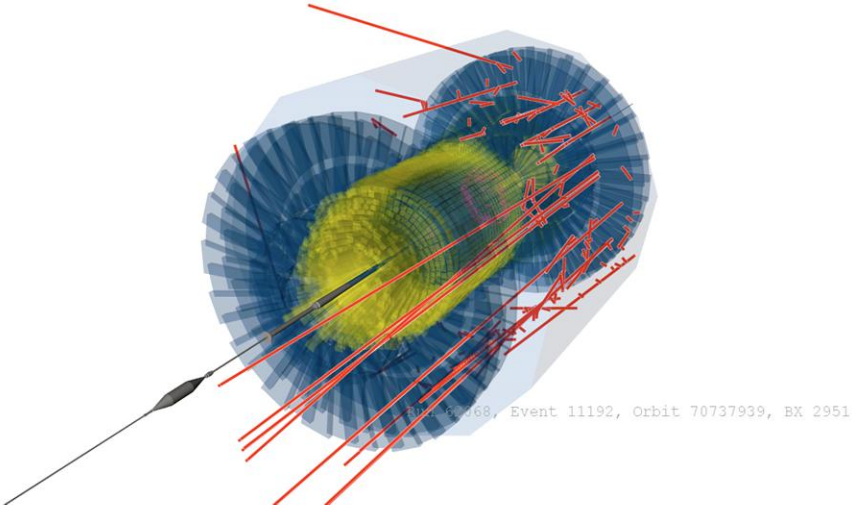
B-tagging affected by pileup

Jet reconstruction affected by detector noise, & pileup

In this talk we will describe and try to quantify some of these effects+describe data-driven estimation techniques

SOURCES OF TRUE INSTRUMENTAL BACKGROUNDS





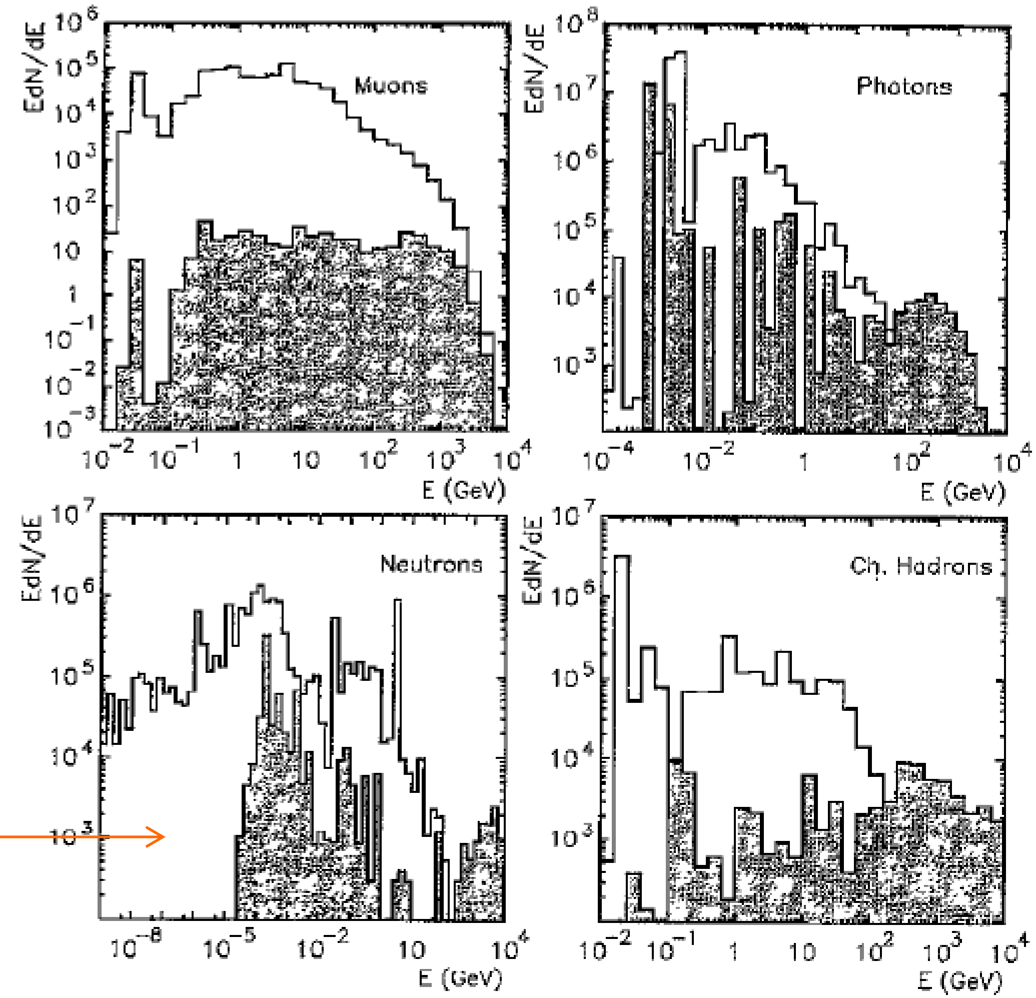
- protons in the LHC strike limiting apertures or beam gas and produce n, γ, $K^{\pm}, \pi^{\pm} \rightarrow \mu^{\pm}$
- Cosmic rays which contaminate events with a probability of $O(10^{-4})$
- Electronic noise in the calorimeter.

BEAM HALO CHARACTERISTICS

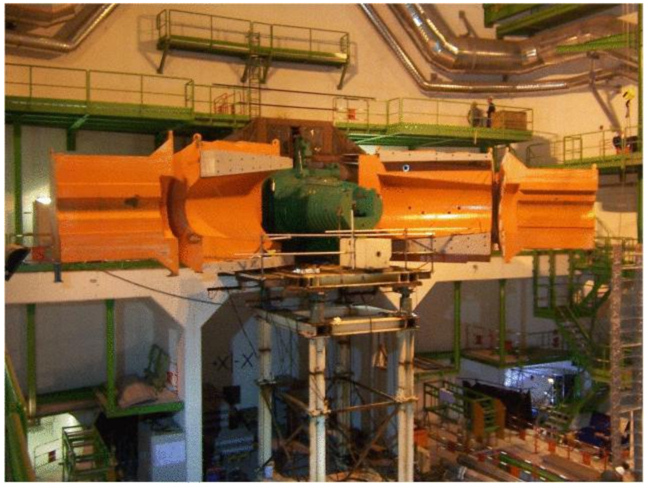
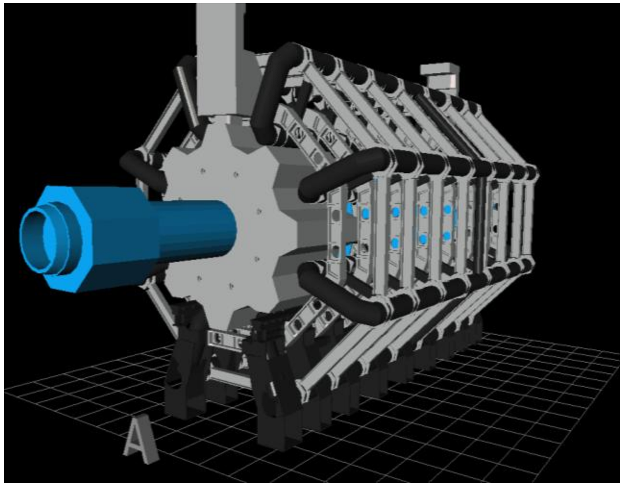
Simulated beam backgrounds in CMS

- 25 m away from the interaction point.
- Dark shaded: radial distance from beam line less than 20 cm.

Nucl. Instrum. Meth. A381 (1996) 531-544



Suppression of non-collision events

- Non-collision background in the detector was known to be a potential problem when the detector was designed and the beam division + the experiments worried about it a lot during construction.
- First line of defense: beam cleaning system (scrapers) in the LHC
- Second line of defense: shielding to plug the tunnel entrance in the experimental halls
-  

- Third line of defense: cuts at the analysis level

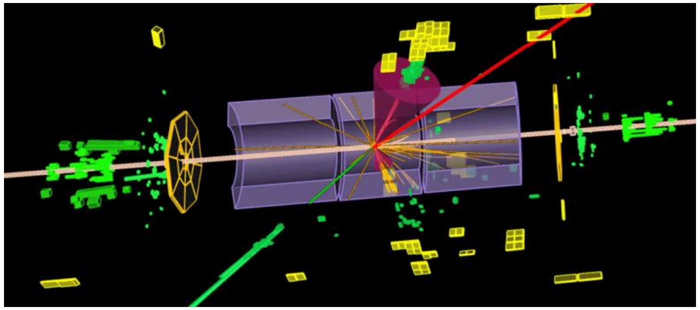
  Require in-time collisions and a primary vertex in the event.

BEAM HALO REJECTION:

*and cosmic rays, too.

- CMS (Eur. Phys. J. C (2011) 71:1721)
  - require primary vertex to be found in a cylinder around nominal IP with length 24 cm and radius 2 cm.
- ATLAS ATLAS-CONF-2010-038
  - require a primary vertex with at least five tracks.
  - Cosmic ray rejection: discard events muon pairs with $\Delta\phi > 3.1$, $|d_0|>0.5$ mm, for both muons, opposite-sign d_0 in dilepton analyses.
  - Rate of beam background events overlapping with pp collisions is ~3×10^{-8}.

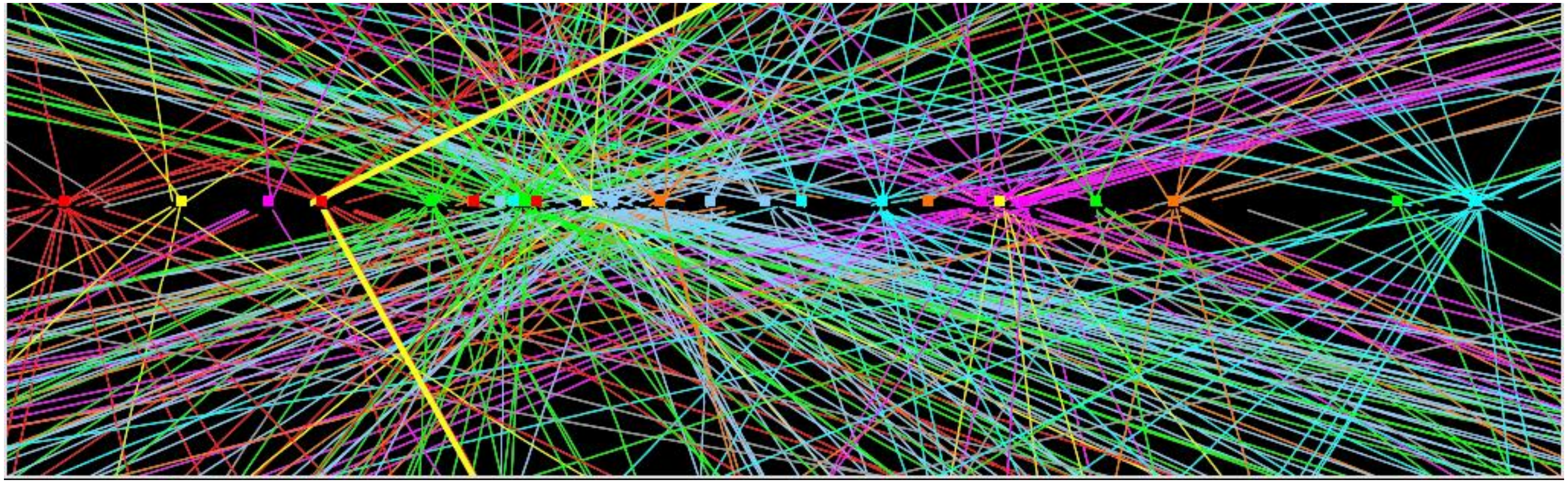
  Use timing information from calorimeter and scintillators.



PILEUP AT THE LHC

50 ns bunch spacing was predominant in 2011/12

In-time pileup: superimposed on every triggering hard collision are multiple interaction vertices from soft QCD:

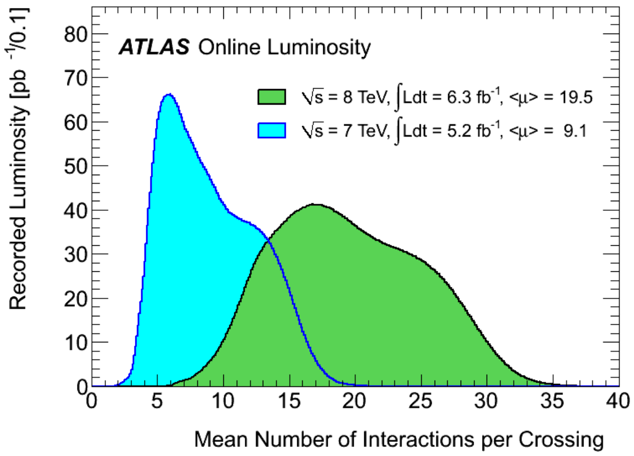


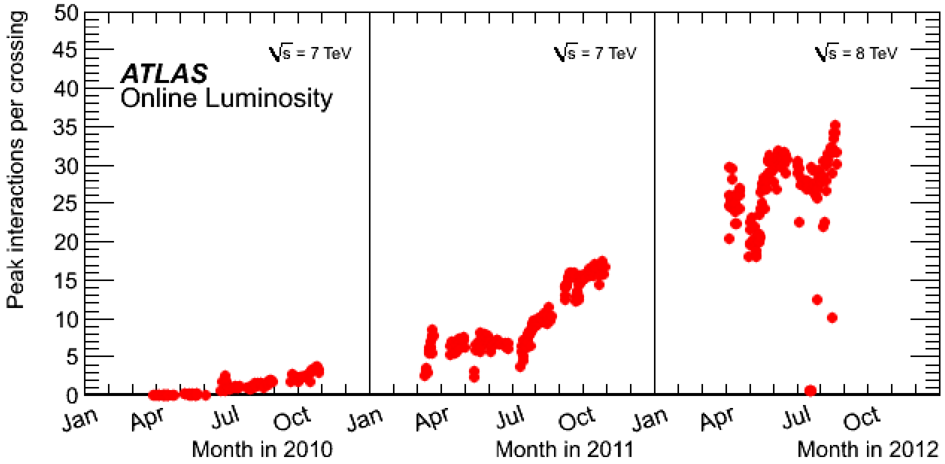
Out-of time pileup: Collisions occur -150 ns, -100 ns, -50 ns, 50 ns, 100 ns... after the triggering hard collision.

* ATLAS: 250-600 ns LAr drift time;

* CMS 100 ns integration time in hadron calorimeters (reduced from 200 ns in 2012 to mitigate pileup effects)

Evolution of pileup with time





This has an impact on both reconstructed physics objects and on the trigger:

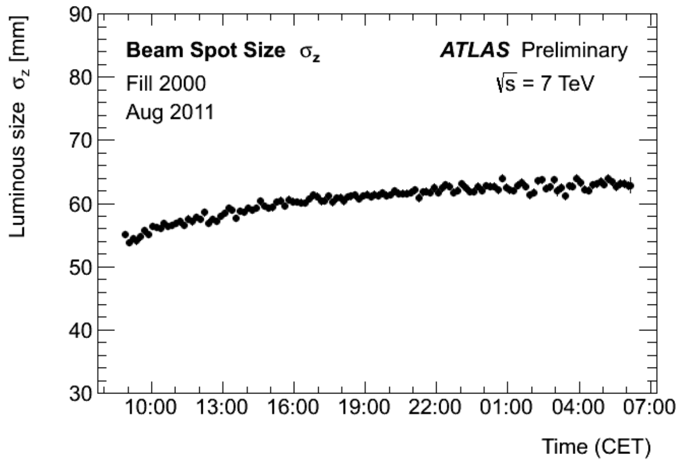
- thresholds vary in both ATLAS and CMS.
- efficiency of single muon trigger decreases because isolation cuts.

Remarks: in-time pileup rate depends only on the luminosity

Out of time pileup depends also on bunch structure

TEVATRON PILEUP EFFECTS:

- CDF/D0: Approximately 2.4 interaction vertices per bunch crossing
- ATLAS & CMS: designed for up to ~25 interaction vertices per bunch crossing
- The density of these vertices is dramatically higher.
- Beamspot length in ATLAS/CMS ~ 6 cm, (c.f. CDF/D0 ~ 30 cm).



- Density ~ 50x larger in LHC than Tevatron when both contributions included!
- Three layers of pixels in ATLAS/CMS provide excellent resolution and ability to resolve multiple interaction vertices.

PILEUP AT THE TEVATRON EXPERIMENTS

TABLE II: Summary of systematic uncertainties on ΔM_{top}. CDF 10777 ($t\bar{t}$ mass difference)

	CDF II Preliminary 8.7 fb^{-1}
Source	Uncertainty (GeV/c^2)
Signal modeling	0.14
Parton showering	0.17
Next Leading Order	0.16
b and $\bar{b}$ jets asymmetry	0.38
Jet energy scale	0.07
Parton distribution functions	0.12
b-jet energy scale	0.05
Background shape	0.20
Gluon fusion fraction	0.05
Initial and final state radiation	0.10
Monte Carlo statistics	0.07
Lepton energy scale	0.06
Multiple hadron interaction	0.05
Color reconnection	0.23
Total systematic uncertainty	0.59

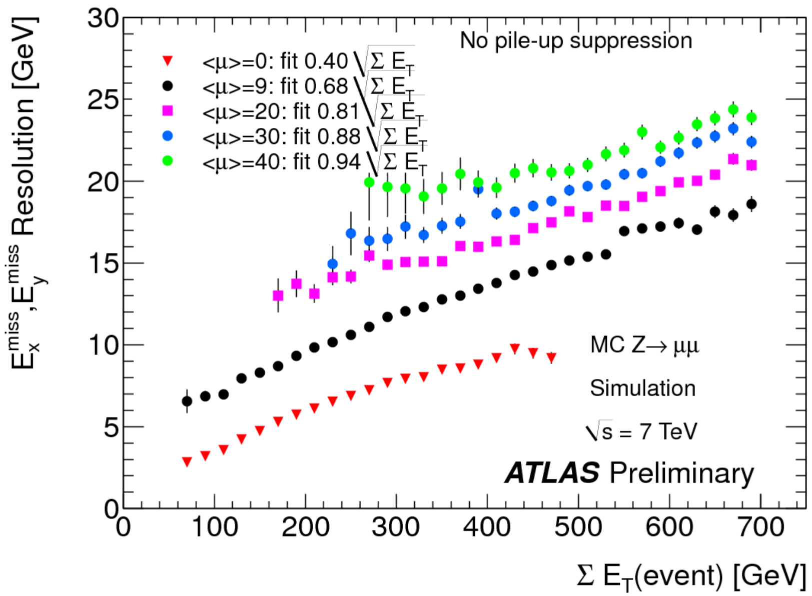
CDF: simulates extra vertices. D0 overlays real data. Small effects and small systematic errors.

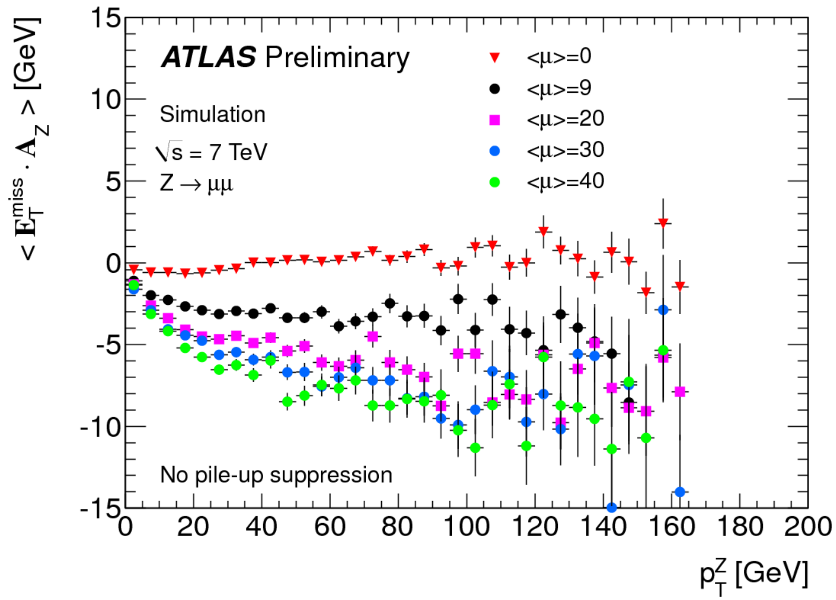
CDF: this highly subdominant systematic uncertainty due to pileup comes from the modeling of pileup (~1.9 in simulation vs. ~2.4 in the data).

By contrast, ATLAS/CMS approach requires active suppression of pileup effects.

Pileup and E_T^{miss}

ATLAS-CONF-2012-101





Pileup events result in additional soft energy deposited into the calorimeter, which degrades the E_T^{miss} resolution, bias, and scale.

Pileup is simulated in Monte Carlo.
Monte Carlo-data mismatch is handled through reweighting.
Suppression techniques were developed to reduce the effect of pileup (2011).

Pileup suppression

E_T^{miss} is the vector sum of contributions from various physics objects including electrons, photons, τ's, μ's, jets, soft jets and unassociated cells.

$$E_{\mathrm{x(y)}}^{\mathrm{miss}} = E_{\mathrm{x(y)}}^{\mathrm{miss},e} + E_{\mathrm{x(y)}}^{\mathrm{miss},\gamma} + E_{\mathrm{x(y)}}^{\mathrm{miss},\tau} + E_{\mathrm{x(y)}}^{\mathrm{miss,jets}}$$
$$+ E_{\mathrm{x(y)}}^{\mathrm{miss,softjets}} + (E_{\mathrm{x(y)}}^{\mathrm{miss,calo},\mu}) + E_{\mathrm{x(y)}}^{\mathrm{miss,CellOut}} + E_{\mathrm{x(y)}}^{\mathrm{miss},\mu},$$

Jets with 7 GeV p_T<20 GeV (EM scale)

energy from cells not associated to e, γ, μ, τ, or jets.

"soft terms" : soft QCD from pileup contributes to both of these.

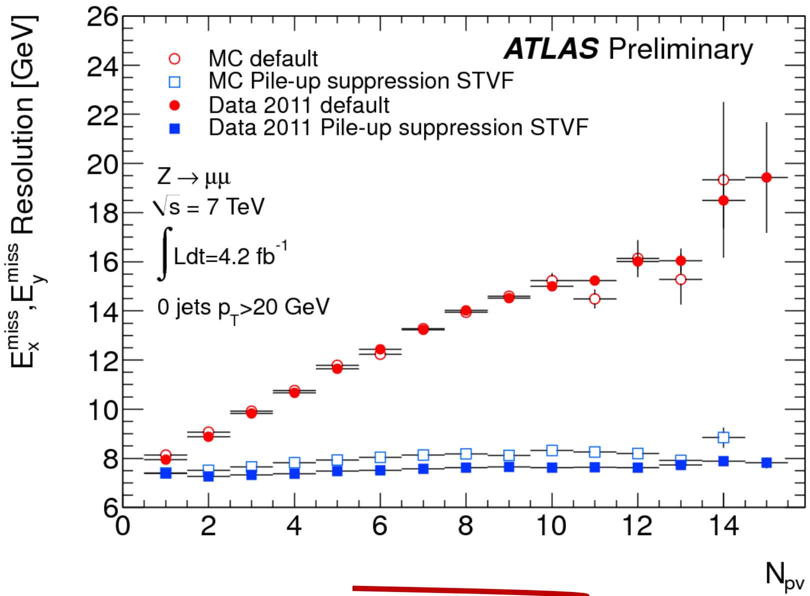
This portion is rescaled to the fraction of tracks p_T coming from the hard scattering primary vertex:

Define the "soft term vertex fraction":

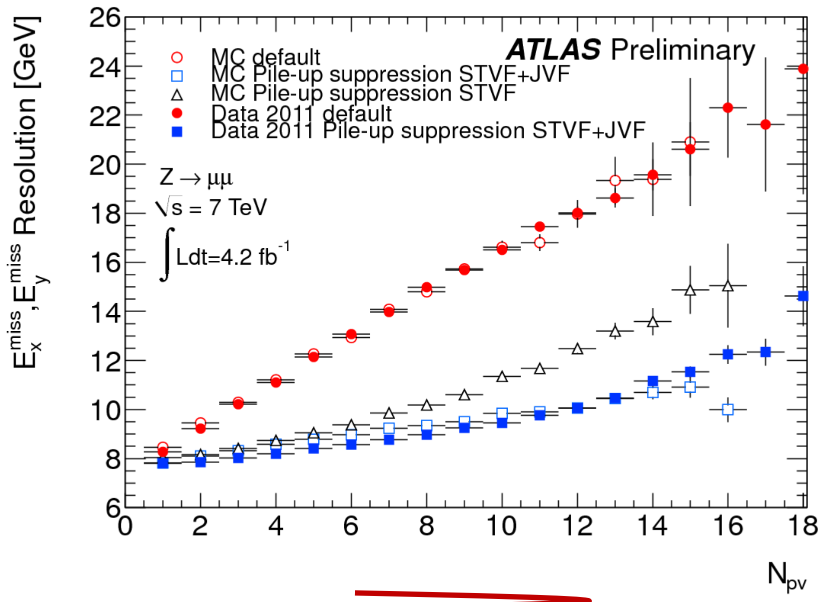
STVF = $\sum_{track,PV} p_T \,/\, \sum_{track} p_T$

To correct for this energy left by the pileup tracks, rescale the soft terms by this fraction.

PILEUP SUPPRESSION



$Z \to \mu^+\mu^-$, no jets $p_T > 20$ GeV



$Z \to \mu^+\mu^-$ inclusive

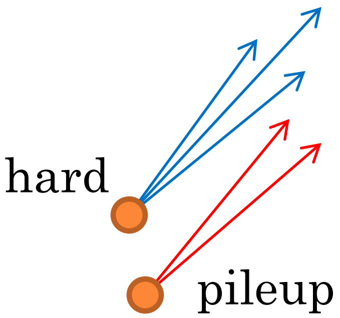
- other pileup suppression techniques have been investigated but this one presently has the best performance.
- STVF pileup up suppression not yet applied to *top* analyses in this conference, but expected for future updates based on 2012 data.

PILEUP AND JET RECONSTRUCTION (ATLAS)

- Suppress single-cell "jets" in the HEC
- Coherent noise "bursts" in the LAr
- Out of time jets are rejected $|\Delta t| = |t_{jet} - t_{event}| > 50$ ns
- Suppress jets seeded from masked calorimeter cells.
- Suppress jets whose main contribution comes from the poorly-calibrated "tile gap 3" scintillators. TG3 is the gap between tilecal barrel & endcap.

- JVF > 0.75
- JVF > 0.50 (2012)

$$\text{JVF} = \sum_{track,PV} p_T \, / \sum_{track} p_T$$



JVF represents the fraction of jet p_T which is associated with the hard-scatter vertex (PV).

The sum runs over all tracks in a jet.

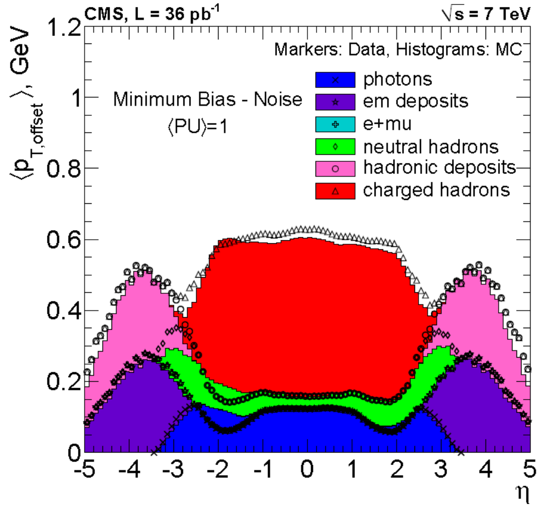
This cut is used to reduce pileup.

- Finally, offset corrections are added to compensate for pileup.

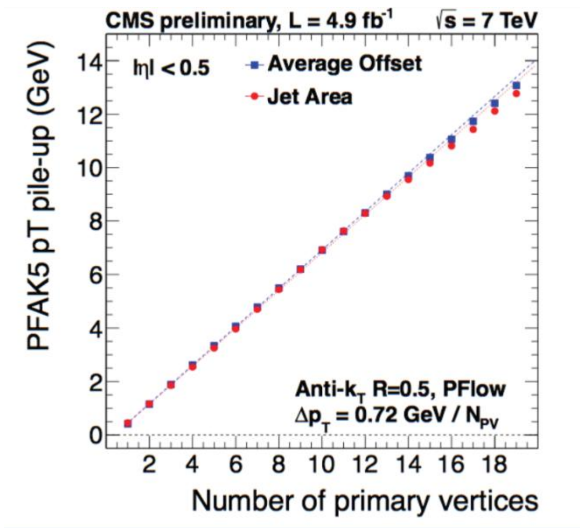
PILEUP AND JET RECONSTRUCTION (CMS)

- CMS has compared two ways of correcting Jet Energy: average offset and area-based.
- Jet area: a measure of a jet's susceptibility to underlying event and pileup.

- Average offset corrects jet energy for the average deposits of a zero-bias event. This depends on η:



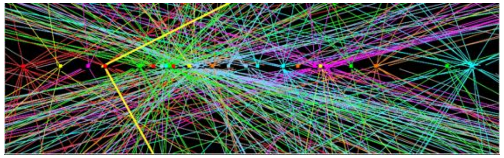
CMS, arXiv:1107.4277



CMS DP/2012-006

- The corrections are in good agreement with area-based corrections.
- Notice that the most of the additional p_T comes from charged hadrons

E_T^{miss} AND ISOLATION (CMS).



In CMS charged hadron subtraction removes particle flow objects not from the primary vertex from jets, E_T^{miss} and isolation discriminants

- E_T^{miss} calculation
  (CMS) is computed from particle flow (PF) objects.
  Charged PF objects not associated with the hard scattering vertex are not included in the calculation.

  Maintains performance in the presence of pileup.

- Isolation discriminant:
  (CMS) Energy in a cone ΔR<0.3 around the lepton candidate is summed. $I_{rel} \equiv \sum E_T / p_T^{lepton}$. The sum does not include charged particle flow (PF) objects not associated to the hard scattering vertex.

ref: CMS-PAS-TOP-11-005

Sources of lepton fakes

Fake electrons

1. Track/photon overlap
2. Photon conversion
3. *b*-quarks and *c*-quarks decaying semileptonically
4. Jets with fluctuations to
   -- few charged tracks.
   -- little energy in hadronic compartments.

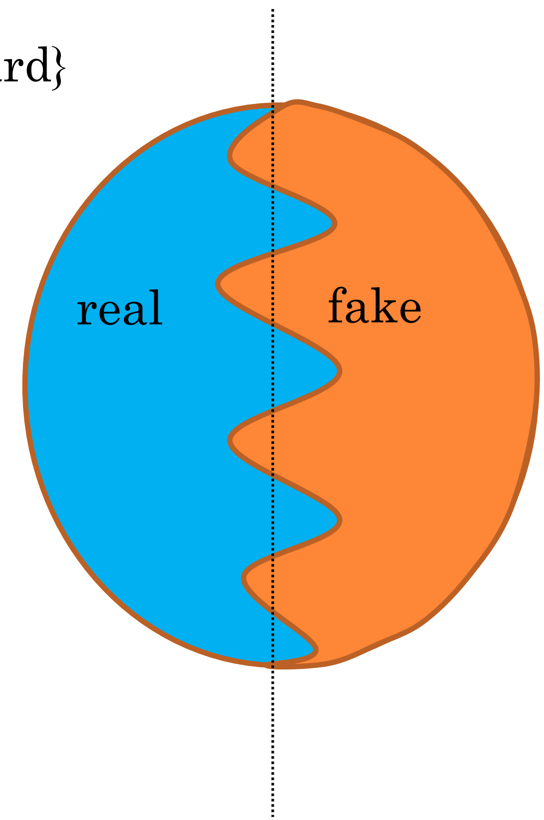
Fake muons

1. Pions or kaons which decay in flight within tracking region.
2. *b*-quarks and *c*-quarks decaying semileptonically
3. hadrons which do not shower in the calorimeter. Interaction length of steel: 17 cm.
4. punch-through hadrons

Simulation of fake electron and fake muon samples from QCD events is *both* unreliable and impractical.

THE MATRIX METHOD: LEPTON + JETS

DØ Collaboration, B. Abbott, et al., Phys. Rev. D 61 (2000) 072001.



$$N^{\text{loose}} = N^{\text{loose}}_{\text{real}} + N^{\text{loose}}_{\text{fake}},$$
$$N^{\text{std}} = rN^{\text{loose}}_{\text{real}} + fN^{\text{loose}}_{\text{fake}},$$

$$\begin{pmatrix} N^{loose} \\ N^{std} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ r & f \end{pmatrix} \circ \begin{pmatrix} N^{loose}_{real} \\ N^{loose}_{fake} \end{pmatrix}$$

r is the marginal efficiency of standard cuts.
f is the same, for background sources

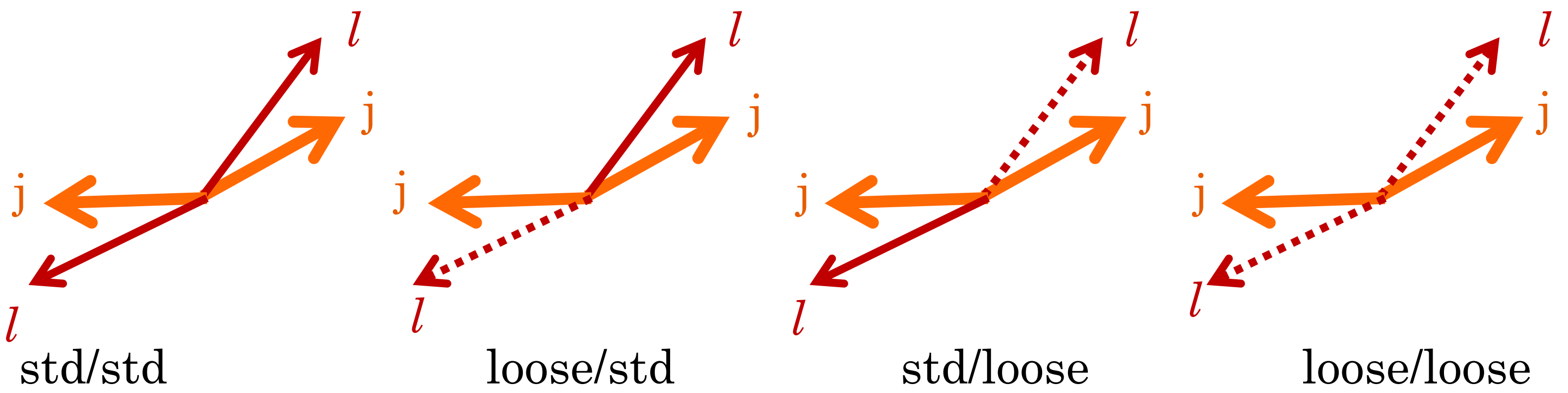
Both can be measured in pure or background event subtracted samples

Then $$\begin{pmatrix} N^{loose}_{real} \\ N^{loose}_{fake} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ r & f \end{pmatrix}^{-1} \circ \begin{pmatrix} N^{loose} \\ N^{std} \end{pmatrix}$$

& you can completely determine the composition of standard selection

MATRIX METHOD IN DILEPTON CHANNELS, SINGLE TOP WT CHANNEL

Same idea, more categories (j=jet, *l*=lepton)



Model is

$$\begin{pmatrix} N^{l,l} \\ N^{l,s} \\ N^{s,l} \\ N^{s,s} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ r_2 & f_2 & r_2 & f_2 \\ r_1 & f_1 & r_1 & f_1 \\ r_1 r_2 & r_1 f_2 & f_1 r_2 & f_1 f_2 \end{pmatrix} \circ \begin{pmatrix} N^{l,l}_{r,r} \\ N^{l,l}_{r,f} \\ N^{l,l}_{f,r} \\ N^{l,l}_{f,f} \end{pmatrix}$$

& other inputs and the method of solution is similar

DETAILS, MATRIX METHOD

Domain of application:

- $t\bar{t}$ and boosted $t\bar{t}$ in lepton+jets (with a 2x2 matrix).
- $t\bar{t}$ in dilepton modes, single top Wt channel (4x4 matrix).
- b-tagging calibration, to measure the b-jet fake rate

Source of real (leptons) for the measurement of r

- usually is taken to be $Z \to e^+e^-$ or $Z \to \mu^+\mu^-$ w/ additional jets.
- MC corrected for Data/MC differences is some cases

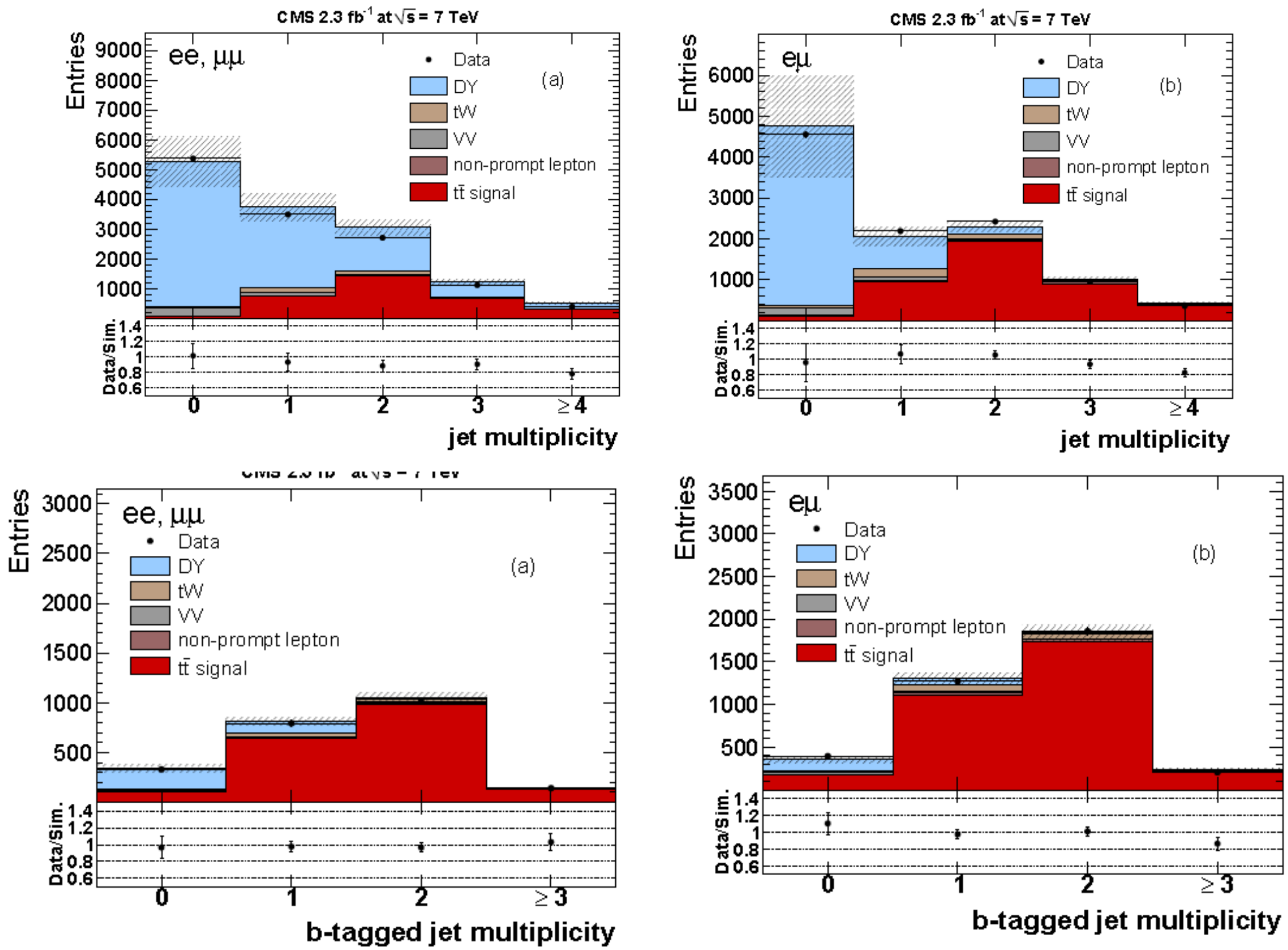
Source of fake leptons for measurement of f

- events enriched in multijet events and/or heavy flavor
- Systematic errors from fake sample choice on f.

+ the algorithm is applied within bins after correcting the real and/or fake samples from any additional contamination.

EXAMPLES:

arXiv:1208.2671

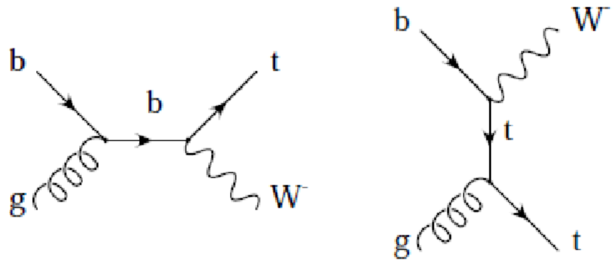


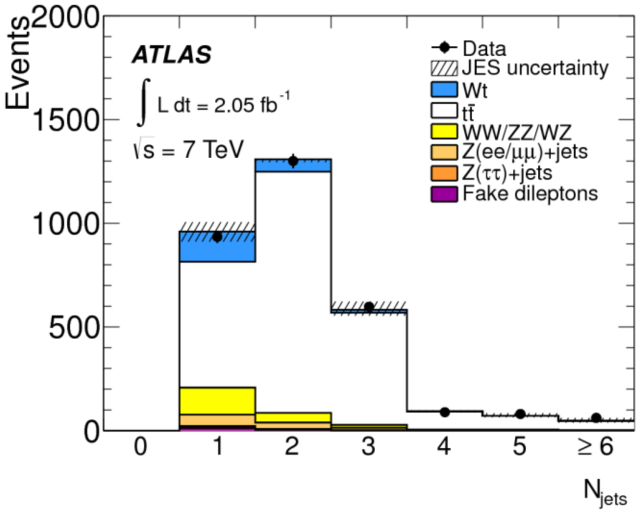
	Before b tagging			Requiring ≥1 b-tagged jet		
	e^+e^-	$\mu^+\mu^-$	$e^\pm\mu^\mp$	e^+e^-	$\mu^+\mu^-$	$e^\pm\mu^\mp$
N_W	7.8 ± 5.9	14.9 ± 7.1	63.8 ± 16.8	1.8 ± 4.8	9.8 ± 5.6	42.4 ± 14.6
N_{MJ}	0.7 ± 0.6	0.4 ± 0.3	21.1 ± 10.0	0.6 ± 0.5	0.2 ± 0.1	7.5 ± 3.9

N_W: one real lepton, "W-like"
N_{MJ}: zero real lepton, "MJ-like"

EXAMPLES:

Single-top *Wt* channel
ATLAS, Phys. Lett. B 716 (2012) 142-159





small contribution, 100% syst. uncertainty assigned.

The jet-electron method: basic idea

- QCD fakes arise from jets with few tracks which shower early in the calorimeter.

- "Electron-like" jets are those with few tracks which have showered early in the calorimeter. EG in ATLAS:

  - 80-95% of the energy must be deposited in EM section of the calo.
  - $p_T > 25$ GeV , same η requirement as electron candidates.
  - at least 4 tracks
  - no other electron or muon candidate in the event.

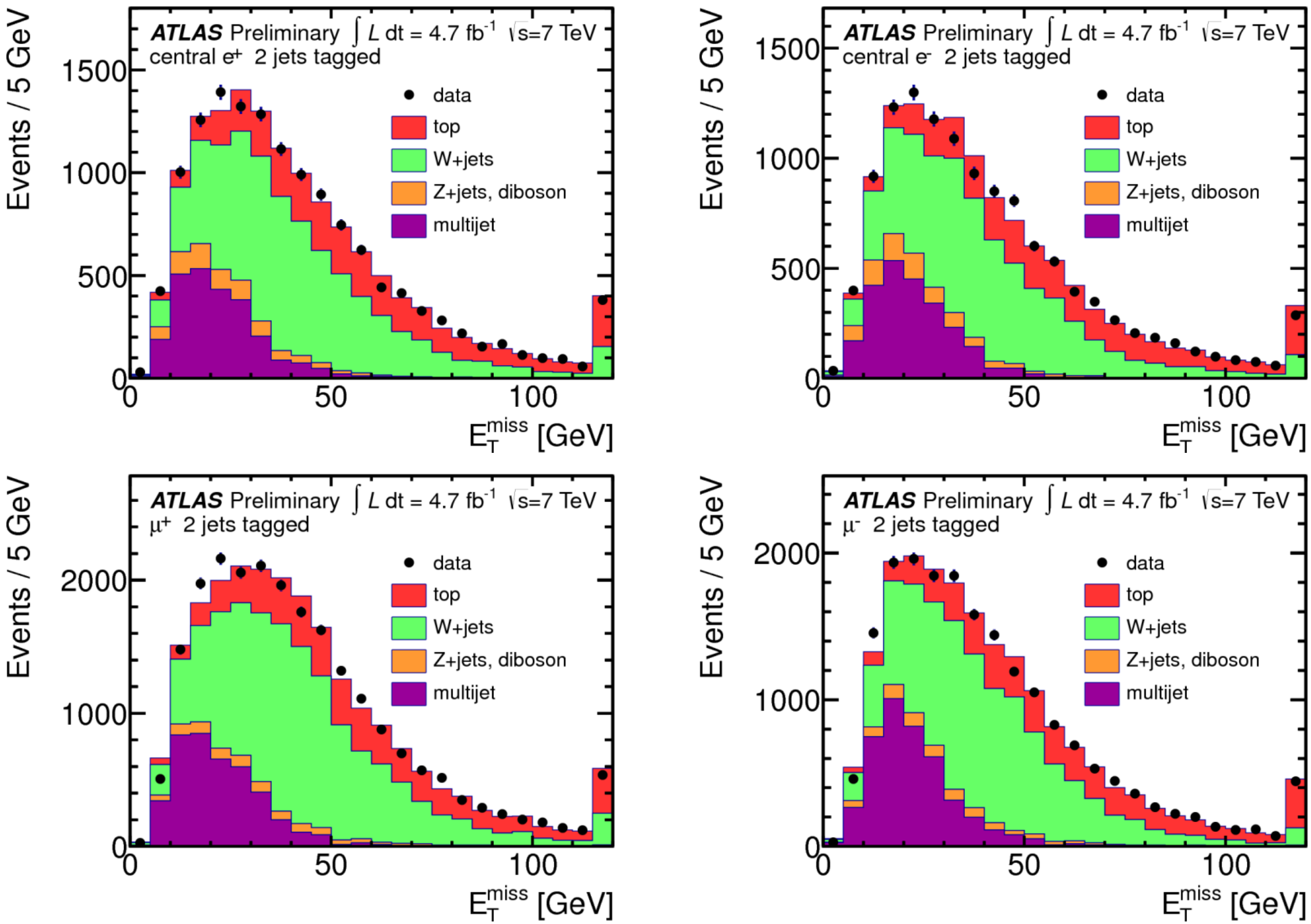
- The shape of the multijet background kinematic distributions are taken from the jet electron sample.

- The normalization is determined from a fit to the E_T^{miss} distribution (before the E_T^{miss} cut is applied).

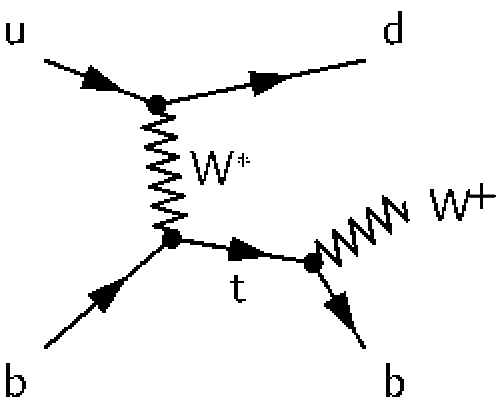
  * E_T^{miss} shape for electroweak processes taken from MC
  * E_T^{miss} shape for QCD multijets taken from data

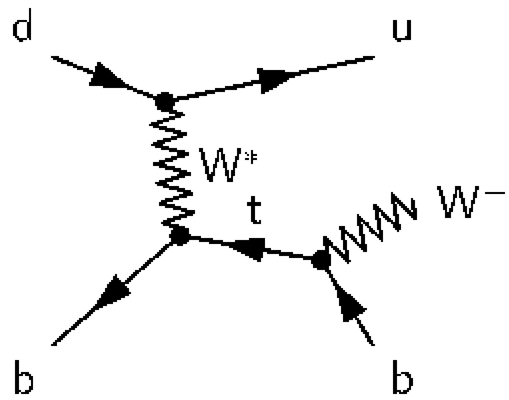
- The jet-electron sample is also used to estimate muon fakes from QCD multijet events.

Jet electron in ATLAS single top (ATLAS-CONF-2012-056)



SINGLE TOP T-CHANNEL CHARGE ASYMMETRY

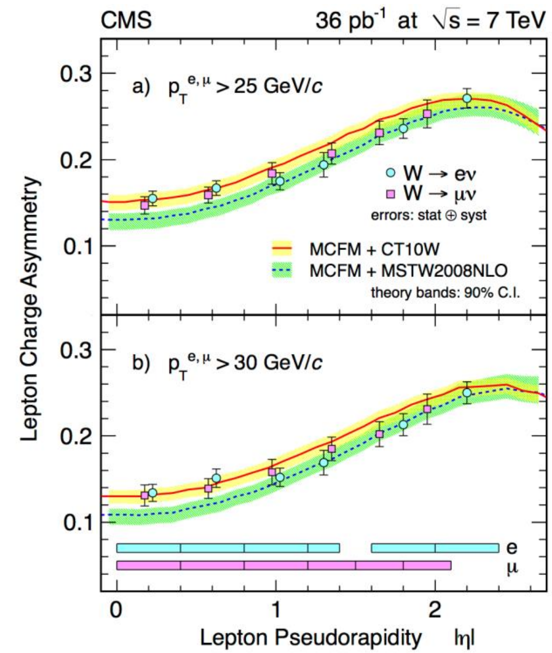




Because of the predominance of *u*-flavored valance quarks over *d*-flavored valence quarks in the proton, the positive leptons prevail over negative leptons in the final state.

The W + jets background is also asymmetric

CMS, JHEP 1104:050,2011;
ATLAS, Phys. Lett. B 701 (2011) 31-49

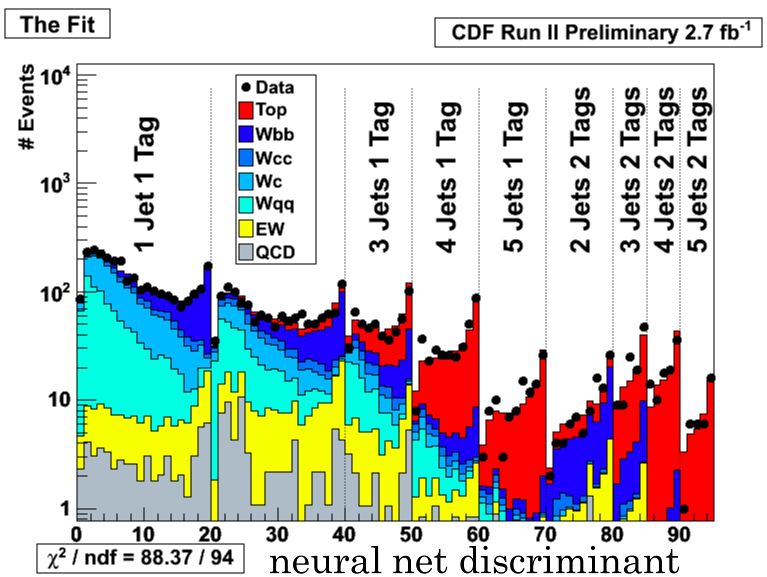


$$\mathcal{A}(\eta) = \frac{d\sigma/d\eta(W^+ \to \ell^+\nu) - d\sigma/d\eta(W^- \to \ell^-\bar{\nu})}{d\sigma/d\eta(W^+ \to \ell^+\nu) + d\sigma/d\eta(W^- \to \ell^-\bar{\nu})},$$

A JetElectron analysis on previous slide is carried out for positive and negative muons and electrons (4 bins).

The *anti-electron method* is a variant of the JetElectron method: use a sample of leptons failing some of the lepton cuts as the template, fit to the E_T^{miss} distribution.

--CDF's favorite method e.g. PRD 84, 031101 (R) (2011)



CDF public note 10137

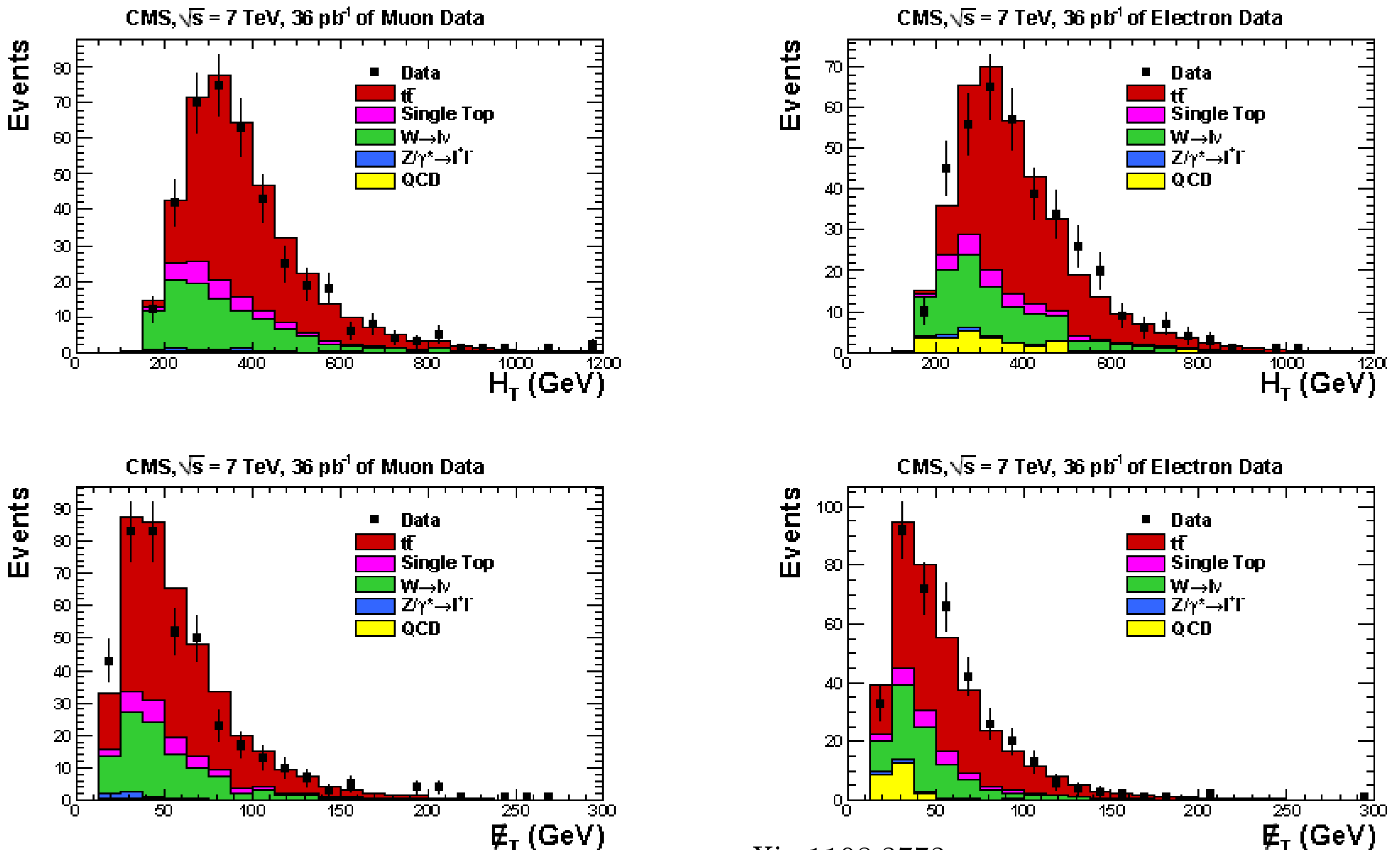
$\sigma(t\bar{t})$ in lepton+jets

Uncertainty	This Analysis	Previous Analysis
Statistical	0.33	0.36
Heavy Flavor Correction	—	0.27
Jet Energy Scale*	0.29	0.29
b-Tagging*	0.23	0.39
Mistags*	0.08	0.17
Q^2*	0.21	—
ISR/FSR*	0.01	0.06
Luminosity	0.45	0.43
QCD Shape	0.01	0.06
Parton Showering	0.11	0.21
KIT Correction	0.10	—
Color Reconnection	0.03	—
Lepton ID	0.05	0.04
Z_0	0.02	0.02
PDF	0.05	0.04
Total	0.73	0.84

- evaluate systematic errors by varying template shape.
- vary template shape by using two-track lepton candidates.

ANTIELECTRON METHOD IN CMS LEPTON+JET

- nonisolated leptons ($I_{rel} > 0.2$) used for template shape (muons) w/ additional inputs from MC and "marginal failures" for electrons—candidates failing two out of six electron id cuts.
- normalization determined from the sideband region $E_T^{miss} < 20$ GeV and by fit to Vertex Mass.



arXiv:1108.3773

METHODOLOGIES IN CMS AND ATLAS

Exper-iment	Channel	Goal	lumi fb^{-1}	Ref.	QCD Bkg Method
ATLAS	lepton+jets	$\sigma(t\bar{t})$	0.035	Phys. Lett. B711 (2012) 244-263	MM(μ), AE(e)
ATLAS	dilepton	$\sigma(t\bar{t})$	0.7	JHEP 1205 (2012) 059	MM
ATLAS	single top lepton+jets	σ(t), t-chan	1.04	arXiv:1205.313	JE
CMS	lepton+jets	$\sigma(t\bar{t})$	0.036	arXiv:1108.3773	AE
CMS	dilepton	$\sigma(t\bar{t})$	2.3	arXiv:1208.2671	MM
CMS	single top lepton+jets	σ(t), t-chan	0.036	Phys.Rev.Lett.107:091802,2011	AE

MM=matrix method,
JE =jet electron
AE =antielectron

METHODOLOGIES IN CDF AND D0

Exper-iment	Channel	Goal	lumi fb-1	Ref.	QCD Bkg Method
CDF	lepton+jets	$\sigma(t\bar{t})$	2.7	Phys.Rev.D84:031101,2011	AE
CDF	dilepton	$\sigma(t\bar{t})$	2.8	Phys.Rev.D82:02002, 2010	*other
CDF	single top lepton+jets	σ(t), s,t,Wt chan	3.2	Phys.Rev.D82:112005,2010	JE/AE hybrid
D0	lepton+jets	$\sigma(t\bar{t})$	0.9	PRL 100, 192004 (2008)	MM
D0	dilepton	$\sigma(t\bar{t}), m_t$	1.0	PLB 679 , 177 (2009)	AE(e)
D0	single top	σ(t), t-chan	2.3	Fermilab-Pub-09/372-E	MM

*other=

- Consider only W+fake lepton.
- Parameterize the fake rate in of fakeable objects in a sample of generic jets with trigger requirement $E_T^{trig} > 20, 50, 70,$ and 100 GeV.
- Apply this to fake rate to fakeable events to obtain W+fake estimate.
- Systematic errors from variation of the selection criteria

CONCLUSIONS

- Basic suppression of beam halo, calorimeter noise, cosmic rays are well established and reduce backgrounds to acceptable levels.
- Pileup is a continuing problem as the number of primary vertices has nearly reached its design value. ATLAS, CMS simulates this and has put in place algorithms to suppress the effect on physics analysis.
- QCD events containing fake leptons are a small background but irreducible. Matrix method, Jet Electron method, and anti-electron method obtain accurate estimates at the LHC and Tevatron for the precision top measurements now in progress.
- *Instrumental effects are largely under control thanks to the hard work of the performance groups.*