Summary of Top Activities

Kerim Suruliz

INFN Trieste and Udine/ICTP Group

Workshop ATLAS Italy, Bologna 16-17 June 2009

Various groups in Italy are involved in top physics at ATLAS.

- Milano: Ilaria Besana, Tommaso Lari, Clara Troncon, Simone Montesano.
- Bologna: Lorenzo Bellagamba, Graziano Bruni, Riccardo Di Sipio.
- Genova: Guido Gagliardi, Bianca Osculati, Stefano Passaggio (just joined; will increase in size later on).
- Udine/ICTP: Bobby Acharya, Marina Cobal, Michele Pinamonti, Umberto de Sanctis, Kerim Suruliz.

Activities

- Milano: data-driven estimates of the W+jets background for commissioning analysis top x-sec measurement in the single lepton channel.
- Bologna: top MC comparison between MC@NLO and POWHEG (with Herwig++).
- Genova: concentrate on b-tagging related studies intially, probably in the commissioning analysis context (using their expertise in pixel), moving to study the all-hadronic channel later on.
- Udine/ICTP: cut and count commissioning analysis for top x-sec measurement in single lepton channel, $t\bar{t}$ production mechanism studies at lowered centre of mass energies.

Important contributions at coordination level: co-convener of x-sec group (Marina) and editor of INT/PUB x-sec notes (Bobby).

Strategy for $t\bar{t}$ cross section measurement with early data (commissioning analysis):

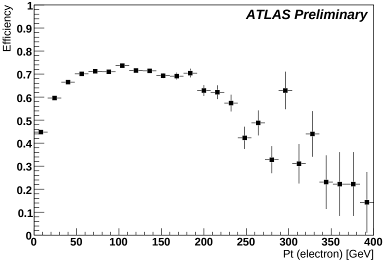
- Assuming 200 pb^{-1} of data at 10 TeV
- Analysis designed to identify top signal without relying on b-tagging (but b-tagging greatly improves S/B).
- Unprescaled single lepton (e, μ) triggers considered
- Definition and strategy for the main systematics.

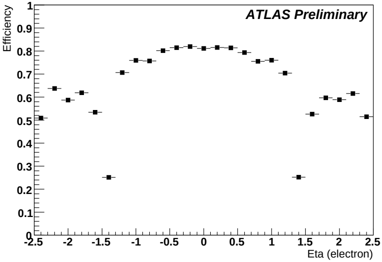
This analysis is the baseline analysis for the single lepton channel x-sec PUB note, currently under review for approval.

- Cut and Count Method – use MC to estimate everything
  - Advantage: simple, doesnt rely on distribution shapes eg Top Mass
  - Disadvantage: relies on a non-optimal MC and is sensitive to the $W+$jets background uncertainty
  - Can be supplemented with a data-driven estimate of the background which reduces the systematics

Definition similar to CSC:

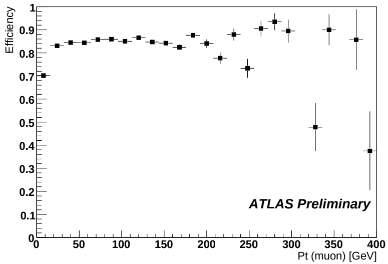
1. egamma isEM ElectronMedium
2. author $= 1$ or author $= 3$
3. $P_T > 20$GeV
4. $|\eta| < 1.37$ or $1.52 < |\eta| < 2.47$
5. etcone20 < 6 GeV

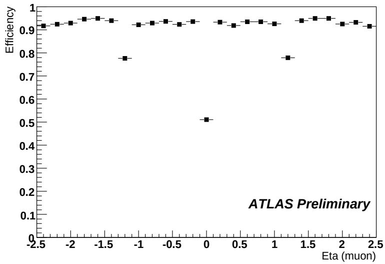




Use STACO muons. Require isCombined.

1. $P_T > 20\text{GeV}$
2. $|\eta| < 2.5$
3. etcone20 $<$ 6 GeV (NB etcone30 in Athena)



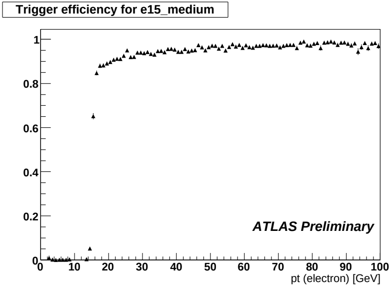


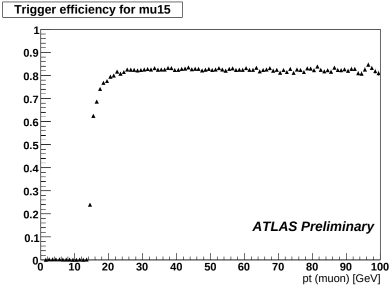
Object definitions - jets

Jets are Cone4TowerJets, with $P_T > 20\text{GeV}$, $|\eta| < 2.5$.

Overlap removal with good electrons: remove the jet if such an electron is found within $\Delta R = 0.2$
Also muon/jet overlap removal: remove muon if there is a good jet within $\Delta R = 0.3$.

Use e15_medium for electron channel, mu_15 for muon channel.





Trigger efficiency for e15_medium and mu15 for reprocessed data.

Serious study on which triggers are optimal for the early data x-sec measurement not yet done - space for contributions.

Decided on a common treatment of the systematics. See https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TopPubSystematics

1. Lepton trigger efficiency - measured from Z events, error expected to be around 1% for 200pb^{-1}.
2. Lepton id efficiency - around 1% with early data, while fake rate 50% (20%) for electrons (muons).
3. Jet Energy Scale (JES). Here consider two scenarios: Default: scale jets with $|\eta| < 3.2$ by 5% and by 10% with $|\eta|$ above. Pessimistic: scale jets with $|\eta| < 3.2$ by 10% and 20% with $|\eta|$ above. The $\not{E}_T$ is rescaled accordingly.
4. Initial/Final State Radiation (ISR/FSR) - evaluated using AcerMC+Pythia with different sets of ISR and FSR Pythia parameters.
5. PDFs: baseline CTEQ6m vs. CTEQ6.6

Systematics - continued

1. W+jets normalisation. Large uncertainty on normalisation of exlusive W+k jet MC samples. Use data-driven methods to estimate the normalisation, via the relation

$$(\frac{W^{SR}}{W^{CR}})_{\texttt{data}} = (\frac{Z^{SR}}{Z^{CR}})_{\texttt{data}} C_{MC} \tag{1}$$

   Error on measurement expected to be roughly 20%. Used Pythia vs Alpgen for a generator error.

2. Generator uncertainty: compare differences between different Monte Carlo programs - MC@NLO, AcerMC, ALPGEN.
3. Luminosity - expect a 20% uncertainty with early data.

Commissioning analysis selection cuts

1. cut 1: trigger - e15_medium or mu15.
2. cut 2: 1 good e, 0 good μ (el channel), or 0 good e, 1 good μ (muon channel).
3. cut 3: $\not{E}_T > 20$ GeV
4. cut 4: 3 jets with $P_T > 40$GeV.
5. cut 5: 4 jets with $P_T > 20$GeV.

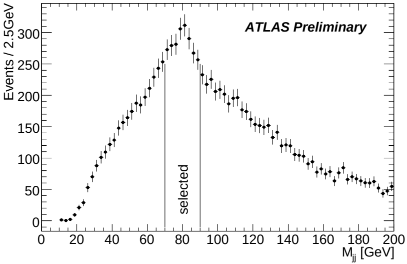
Top candidates can be reconstructed in events which pass all the cuts as the 3-jet combinations with highest combined P_T.

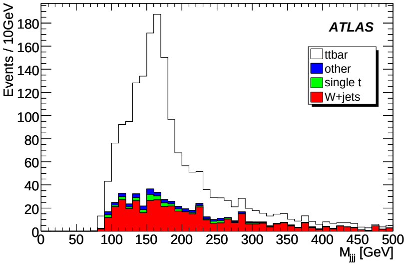
Two additional cuts are then possible.

Top mass cut: the 3-jet invariant mass is 149GeV$< M_{jjj} < 189$GeV.

W mass cut: among the 3 jets there are 2 whose combined invariant mass is within 10GeV of M_W.

TOP AND W PLOTS





Left: Invariant masses of all pairs of jets consituting the top candidate, after baseline selection.

Right: Invariant mass for triplets of jets forming the top candidate, after imposing the M_W cut.

Cut and Count Method: electrons

Electron analysis						
	10TeV (200 pb^{-1})			14TeV (100 pb^{-1})		
Sample	default	W con.	m_t win	default	W con.	m_t win
ttbar	2600	1286	581	2555	1262	561
W+jets	1305	448	108	761	241	60
single top	210	81	27	183	67	23
$Z \rightarrow ll$ +jets	148	43	11	115	35	8
hadronic $t\bar{t}$	16	10	2	11	4	0.0
$W\ b\bar{b}$	21	7	2	44	15	3
$W\ c\bar{c}$				19	6	1
WW	11	6	2	7	4	0.4
WZ	3	1	0	4	1	0.4
ZZ	0.4	0.2	0.1	0.5	0.2	0.1
Signal	2600	**1286**	581	2555	1262	561
Background	1715	**598**	154	1144	374	96
S/B	1.5	**2.1**	3.8	2.2	3.4	5.8

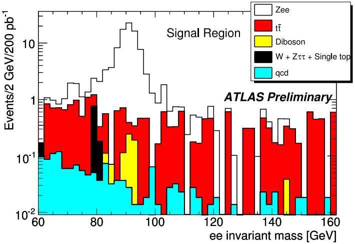
Summary of systematics: Cut and Count

Source	e	e M_W-cut	μ	μ M_W-cut
Stat.	2.5	3.4	2.3	3.1
Lep ID eff	1.0	1.0	1.0	1.0
Lep trig eff	1.0	1.0	1.0	1.0
50% W+jets	25.1	17.4	28.1	19.8
20% W+jets	10.0	7.0	11.2	7.9
JES (10%)	+43.0-38.0	+22.0-32.0	+35.0-38.7	+12.4-34.7
JES (5%)	+21.0-19.0	+11.0-15.0	+16.5-20.2	+7.6-17.6
PDFs	1.9	1.9	1.2	1.4
ISR/FSR	+9.1-9.1	+7.6-8.2	+8.2-8.2	+5.2-8.3
Signal MC	5.0	8.4	2.2	8.0
Theory	1.3	0.9	1.1	0.9
10% Lumi	16.7	14.7	17.0	15.0
20% Lumi	33.2	29.3	34.0	30.1

$$\frac{\Delta\sigma}{\sigma}(\mathrm{e}) = 3.4\mathtt{stat}\;(+17.4-20.4)\mathtt{sys.}\;\pm 1.9\mathrm{PDF} \pm\; 29.3\mathtt{lumi}$$

$$\frac{\Delta\sigma}{\sigma}(\mu) = 3.1\mathtt{stat}\;(+14.6-22.5)\mathtt{sys.} \pm\; 1.4\mathrm{PDF} \pm\; 30.1\mathtt{lumi}$$

Data driven estimation of the W+jets background @ Milano



- Select a control region of Z+jets events - low jet multiplicity
- Perform the baseline selection on it to obtain a Z+jets 'signal sample'
- Select a control sample of W events
- Use the ratio of W to Z to extrapolate W into signal region

TABLE: Number of events in the Z$\rightarrow ll$ + 1 jet control samples and the Z$\rightarrow ll$ signal-like control samples for an integrated luminosity of 200 pb^{-1}.

Process	Z$\rightarrow ee$		Z$\rightarrow \mu\mu$	
	+ 1 jet	signal-like	+ 1 jet	signal-like
Z(ee)	**10214.0**	**81.5**	0.0	0.0
Z($\mu\mu$)	0.0	0.0	**15747**	**150.1**
Z($\tau\tau$)	0.1	0.1	0.9	0.0
W	8.4	0.0	0.0	0.0
$t\bar{t}$	6.0	2.8	10.5	5.0
single top	2.9	0.0	2.3	0.0
Wbb	0.0	0.0	0.0	0.0
Diboson	24.9	0.5	40.0	1.0
QCD	107.7	0.35	0.0	0.0

Systematics

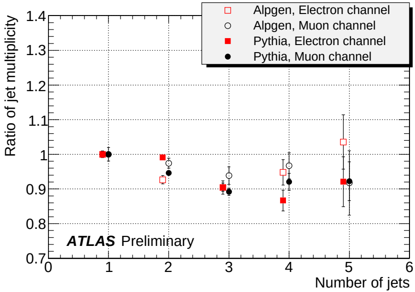


Figure: Ratio of W and Z events as a function of the number of jets with $p_{\mathrm{T}} > 20$ GeV. The ratio has been normalized to unity in the one-jet bin. The results obtained at Monte Carlo truth and reconstruction level are compared.

A source of systematic error is the uncertainty in C_{MC}, the MC correction factor. Comparing ALPGEN with PYTHIA, a roughly 10% error on this is obtained.

The dominant sources of systematic errors are

- Purity of control samples. Z samples have very high purity , and one can use a sideband subtraction method for background - therefore uncertainty here is negligible. Dominant uncertainty from W sample. Assume 50% uncertainty of background (QCD predominantly).
- Uncertainty of MC correction factor C_{MC}

Table: The expected relative uncertainties on the background estimation.

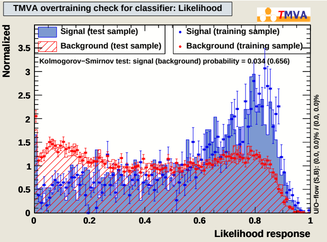
	electron channel	muon channel
Stats. 200 pb^{-1}	11.3%	8.3%
Sample purity	18.8%	11.9%
MC uncert.	10.1%	10.1%
JES 10%	3.6%	2.3%
JES 5%	3.0%	0.7%
Lep. scale	0.4%	0.7%
total error	24.1%	17.8%

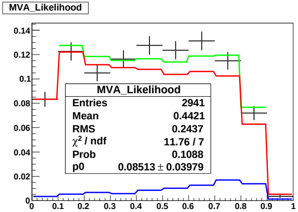
Work by Michele Pinamonti in collaboration with Pamela Ferrari (NIKHEF). The aim is to study the characteristics of top production at different com energies, comparing results at 6, 8, 10 and 14TeV. PGS (Pretty Good Simulation) was used for producing the MCs.

It is also hoped that one can perform a measurement of the relative contributions of $q\bar{q} \to t\bar{t}$ and $gg \to t\bar{t}$ production channels.

To do this, a number of discriminant distributions is considered which take into account the expected differences in kinematics (e.g. gluons are more likely to radiate a soft gluon than quarks; also, PDFs for quarks and antiquarks are different giving differences in total $t\bar{t}$ system boost and angular distributions)

A paper is in preparation.





Left: likelihood functions for $q\bar{q}$ (blue) and gg (red).

Right: fit to extract the relative proportion of the two contributions.

Use TMVA (Toolkit for MultiVariate Analysis) to build likelihood function from MC which can be applied to the data. Work in progress.

POWHEG vs MC@NLO comparison @ Bologna

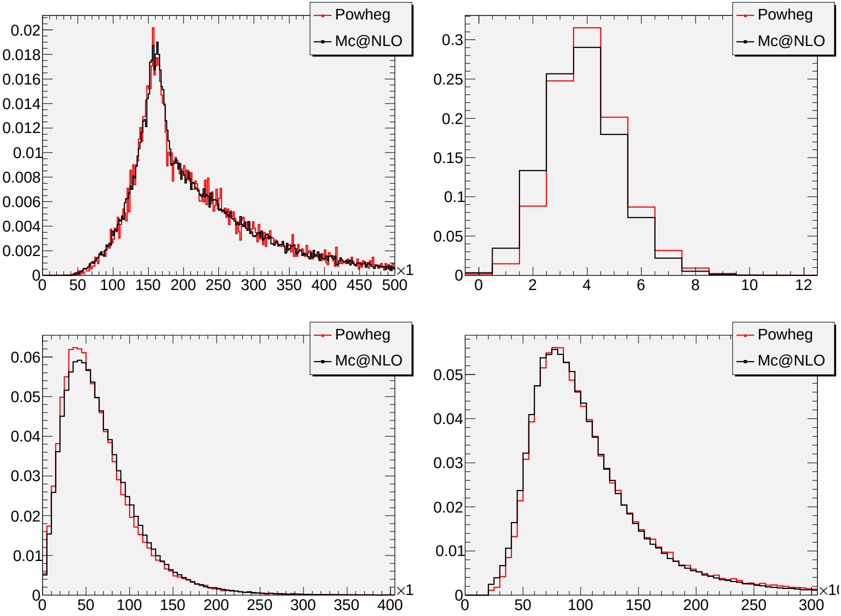
Work done in Bologna by Riccardo di Sipio.

POWHEG (POsitive Weight Hard Event Generator) is a NLO Monte Carlo generator which does not have weights, unlike MC@NLO (in which events have weight ± 1).

POWHEG can be interfaced with any shower MC, as opposed to MC@NLO which has been interfaced only with HERWIG.

The aim is to commission POWHEG as this will be important for $t\bar{t}$ systematics.

Studies at generator level completed, moving to AtlFast-II simulated comparisons.


Powheg
Mc@NLO
0.02
0.018
0.016
0.014
0.012
0.01
0.008
0.006
0.004
0.002
0
0 50 100 150 200 250 300 350 400 450 500
×1
Powheg
Mc@NLO
0.3
0.25
0.2
0.15
0.1
0.05
0
0 2 4 6 8 10 12
Powheg
Mc@NLO
0.06
0.05
0.04
0.03
0.02
0.01
0
0 50 100 150 200 250 300 350 400
×1
Powheg
Mc@NLO
0.05
0.04
0.03
0.02
0.01
0
0 50 100 150 200 250 300

Plans for the future

Need to coordinate efforts in anticipation of early data.

Good level of interaction between Milano and ICTP/Udine during preparation of INT/PUB note. We plan to coordinate software (e.g. format of ntuples) so once data arrives we can easily perform cross-checks etc.

Need to define a common strategy for the benchmark points:

- first few pb^{-1}
- $50\mathrm{pb}^{-1}$
- $200\mathrm{pb}^{-1}$

Need to learn about running on ESDs in case of delays in preparation and distribution of AODs?

With very first data, look at simpler objects/measurements before moving to top physics. E.g. distributions (transverse momentum, pseudorapidity) of charged particles, W, Z peaks, inclusive jet cross section, trying various jet algorithms.

50PB^{-1}

With 50pb^{-1} can hope to observe a clear top signal.

Start studying fake rates (matrix method?).

Once we have significant Z+jets samples, we can start the data-driven background estimation.

Start checking the JES via the hadronic W peak.

Start tuning the signal and background MC to better describe data wherever possible.

200PB^{-1}

With 100-200pb^{-1} the commissioning cross-section analysis can be performed.

Aim for a publication as soon as possible, in which the Italian community hopes to play a key role!

To make this a reality, we have to work together as much as possible, coordinate software infrastructure and share the tasks.

10TeV analysis was a good proof of principle that this is achievable!

	Trig	Lep	$\not{E}_T$	Jet req. 1	Jet req. 2
	eff (%)	eff (%)	eff (%)	eff (%)	eff (%)
$t\bar{t}$ (e) (10TeV)	25.0	70.2	90.6	47.5	79.8
$t\bar{t}$ (μ) (10TeV)	31.0	67.0	91.2	47.7	80.0

Efficiency given wrt the whole non-fully hadronic ttbar sample.

Backup: Efficiencies

Event type	Electron analysis Trigger+Selection (%)			Muon analysis Trigger+Selection (%)		
		W const.	$m_t$win		$+\ W$ const.	$m_t$win
$t\bar{t}$ (e)	18.9	9.9	4.7	0.0	0.0	0.0
$t\bar{t}$ (μ)	0.0	0.0	0.0	22.9	12.1	5.7
$t\bar{t}$ (τ)	1.3	0.7	0.3	1.6	0.9	0.4
$t\bar{t}$ (dilep)	3.0	1.1	0.3	3.4	1.2	0.3
$t\bar{t}$ (hadron)	0.0	0.0	0.0	0.0	0.0	0.0

Breakdown of efficiencies according to event type at truth level.

Backup: Cut and Count Jet scale uncertainty

Recalculate the cross section σ and find $\Delta\sigma/\sigma$ after rescaling jets and $\not{E}_T$. This reduces to computing

$$\epsilon = (S' + B' - B - S)/S$$

Table: Relative uncertainty on the cross section for the default selection and the default selection + M_W, for the electron and muon analyses.

Error	**Variation**			
(S'+B'-B-S)/S	+5%	+10%	-5%	-10%
electron				
default	21%	43%	-19%	-38%
default + M_W	**11.2%**	22.2%	**-15%**	-32.3%
muon				
default	16.5%	35.0%	-20.2%	-38.7%
default + M_W	**7.6%**	12.4%	**-17.6%**	-34.7%

W+jets will be determined from Z+jets using data driven methods.

We expect the error on the measured W+jets background to be of order 20%.

TABLE: Events which pass the default selection and the default selection + M_W for different normalization of the W+ jets samples.

Variation of W+jets	**electron default**	**electron default + M_W**	**muon default**	**muon default + M_W**
50% more	25.1	17.4	28.1	19.8
20% more	10.0	7.0	11.2	7.9
10% more	5.0	3.5	5.6	4.0

Backup: Cut and Count results: muons

Muon analysis						
	10TeV (200 pb^{-1})			14TeV (100 pb^{-1})		
Sample	default	W con.	m_t win	default	W con.	m_t win
ttbar	3144	1584	712	3274	1606	755
W+jets	1766	628	148	1052	319	98
single top	227	98	33	227	99	25
$Z \rightarrow ll$ +jets	144	49	13	84	23	3
hadronic $t\bar{t}$	11	5	2	35	17	7
$W\ b\bar{b}$	32	10	3	64	19	4
$W\ c\bar{c}$				26	9	3
$W\ W$	14	7	2	7	3	0.7
$W\ Z$	5	2	0.2	7	3	0.8
$Z\ Z$	0.5	0.2	0.1	0.7	0.3	0.1
Signal	3144	**1584**	712	3274	1606	755
Background	2199	**799**	201	1497	495	143
S/B	1.4	**2.0**	3.5	2.2	3.2	5.3

BACKUP: W+JETS CONTROL SAMPLE FOR DATA-DRIVEN BACKGROUND ESTIMATION

TABLE: Composition of the $W \to l\nu$ + 1 jet control sample, for an integrated luminosity of 200 pb^{-1}.

Process	$W \to e\nu$	$W \to \mu\nu$
$W(e\nu)$	**148682**	0
$W(\mu\nu)$	43	**190274**
$W(\tau\nu)$	5571	6822
$Z(ee)$	1157	0
$Z(\mu\mu)$	1	8066
$Z(\tau\tau)$	879	1130
$t\bar{t}$	203	241
single top	272	308
Wbb	97	119
Diboson	84	557
QCD	47600	28000

Note: difference between Zee and Zmm believed to be genuine in the sense that a 'missing muon' has a different contribution to missing energy than a 'missing electron'. Under further study.