Strategies for top pair cross section measurements in ATLAS

Duc Bao Ta
University of Bonn
on behalf of the ATLAS top groups

Why precise cross section measurements?

General Strategies and Analyses

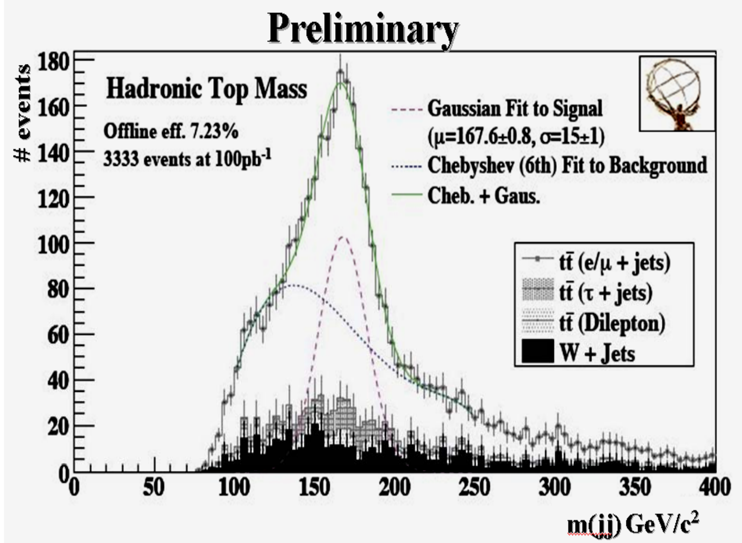
General Issues

Remarks

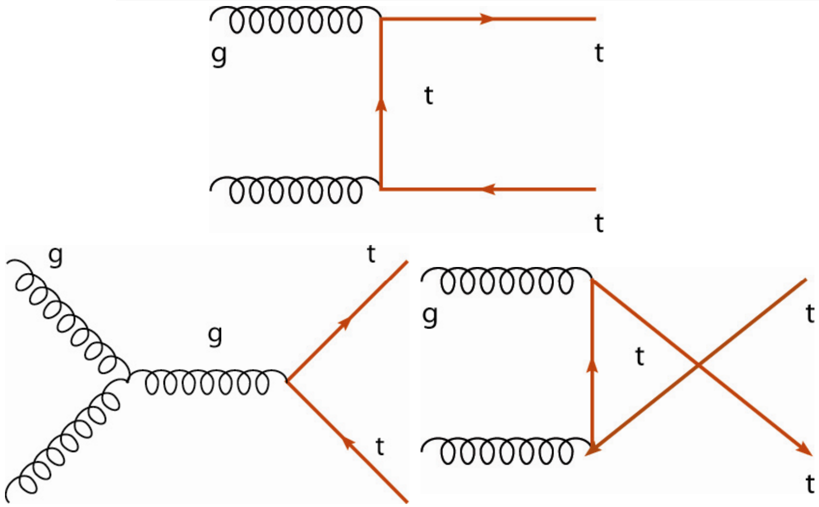
- All plots and results shown in this talk are preliminary
- The consolidation and finalizing of the results is ongoing right now
- Will see the definitive plots and results in the forthcoming CSC notes

Why precise cross section measurement?

- 1 $t\bar{t}$ pair per second, top peak visible after a few weeks
  - Theoretical prediction: NLO+NLL 833+52-39pb [Boniciani, Catani, Mangano, Nucl Phys B 529 (1998) 424-250]

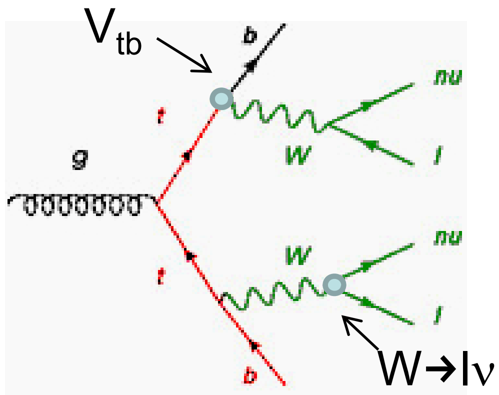


- ~90% top pair production at LHC via gluons
  - Measurement would test QCD calculation, gluon pdfs

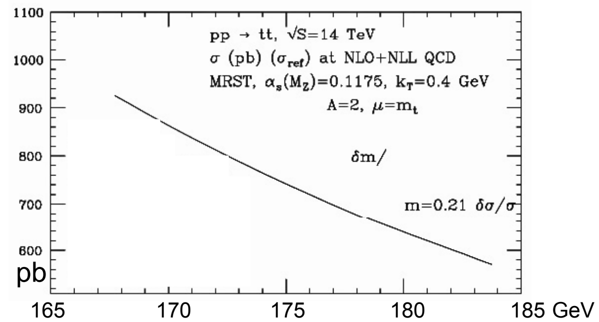


universitätbonn

- Test of EW
  - Top quark decay (V_{tb}) of EW nature
  - W properties (BR $W \rightarrow l\nu$)
  - sensitivity on top mass



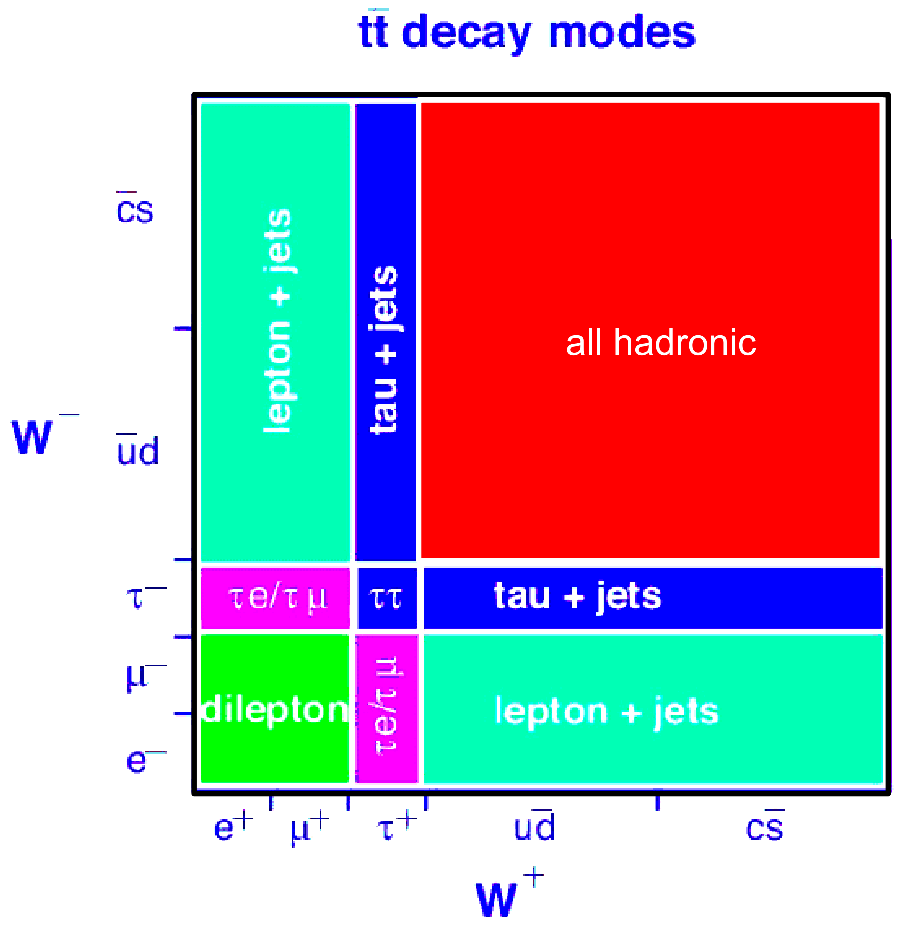
- New physics searches
  - $t \rightarrow H^{+}+b$, rare top decays
  - resonances $Z' \rightarrow t\bar{t}$
- Background to other analyses
  - Single Top, Higgs, SUSY



- Measurement of cross section can provide consistency test of SM and physics beyond SM

General strategies

- Measurement in different W decay modes:
  - all hadronic
    - 6-jet environment, most difficult
  - lepton+jets
    - golden channel
  - di-leptonic
    - cleanest selection possible
    - include leptonically decaying τ
    - exploit τ-tagging
  - different backgrounds, different systematic issues, use as consistency check
- with or without b-tagging
  - w/o b-tagging suitable for first measurements
  - with b-tagging purer selection at later stage

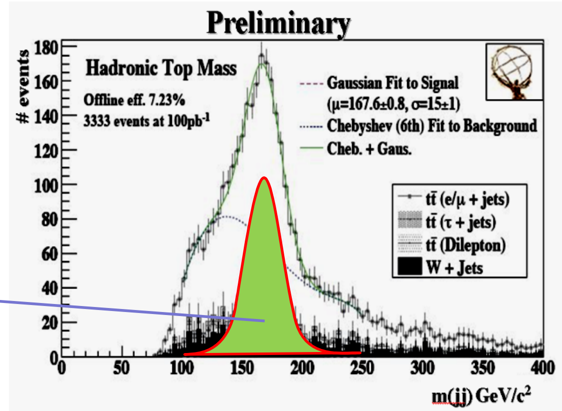


- different data amount aims
  - commissioning analysis, early data
    - also used for detector commissioning and calibration
  - first data analysis
    - without b-tagging, not perfect alignment
  - later stage analysis
    - make use of b-tagging
    - more statistics for rare decay channels (e.g. τ to leptons)
    - systematic error will dominate soon

universitätbonn

General strategies (II)

- direct measurement
  - cut and count
    - simple and powerful for first measurements



$$\sigma = \frac{N_{measured} - N_{background}}{A \cdot \varepsilon \cdot L}$$

    - selection from top mass measurement or use selection for fruther top property measurement
  - Likelihood methods
    - more sensitive to actual top events shapes (from MC)
  - Phase space resolved measurements
    - expect high statistics available for fine phase space binning, sensitive to new physics mimicking top decay events

Cut and count (I)

- lepton+jets channel: *B. Acharya et al.* INFN Udine, CERN
  - commissioning study, W+jets background only
  - MC divided into 1/3 "MC" and 2/3 "data"
  - efficiencies and acceptance determined by "MC"

Cuts
p_T(lepton) > 20 GeV
3 jets with p_T > 40 GeV
4^{th} jet with p_T > 20 GeV
η(lepton) < 2.5, η(jet) < 2.5
Missing E_T > 20 GeV
Top is reconstructed as the 3-jet comb.
Require m(2 of 3 jets) within 10 GeV of M_W

ε from MC	**N_{bkg} from MC**	**S+B from data**
2.7 %	1243.1	1418.0

Total BR	**Result for $100 pb^{-1}$**
5/9	820.4pb± 3.2%(stat) ± 11% (JES)

- lepton+jets channel: *A. Bangert et al.* MPI Munich
  - looking into the effect of different jet algorithms

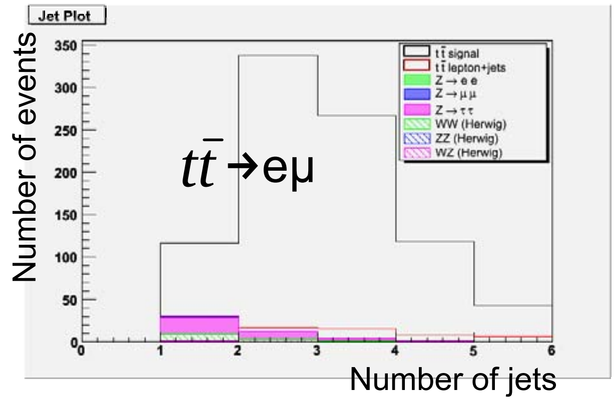
Cut and count (II)

- di-lepton channel: *D.B Ta and J. Sjölin et al.* University of Bonn, University of Stockholm
  - $t\bar{t}$ lepton+jets, Z→ll and di-boson backgrounds
  - both group obtain comparable numbers

Cuts for best S/sqrt(S+B) ratio	
2 opposite charged, isolated leptons	
at least 2 jets	
p_T (leptons, jets) > 20 GeV, η(lepton, jets) < 2.5	
remove overlap of muons with jets	
ee / μμ veto on M(ll) 85GeV - 95GeV	
eμ:$E_{T,Miss}$>20 GeV	ee / μμ: $E_{T,Miss}$>35 GeV

<table>
<tr><th></th><th>efficiency</th><th>S/B</th><th>S/sqrt(S+B)</th></tr>
<tr><td>sig eμ</td><td>(5.76±0.8)%</td><td rowspan="2">13.01 ±0.60</td><td rowspan="2">22.13±0.03</td></tr>
<tr><td>bkg</td><td>(0.01±0.001)%</td></tr>
<tr><td>ee</td><td colspan="2">6.71 ±0.61</td><td>12.43±0.05</td></tr>
<tr><td>μμ</td><td colspan="2">5.16 ±0.57</td><td>14.60±0.04</td></tr>
</table>

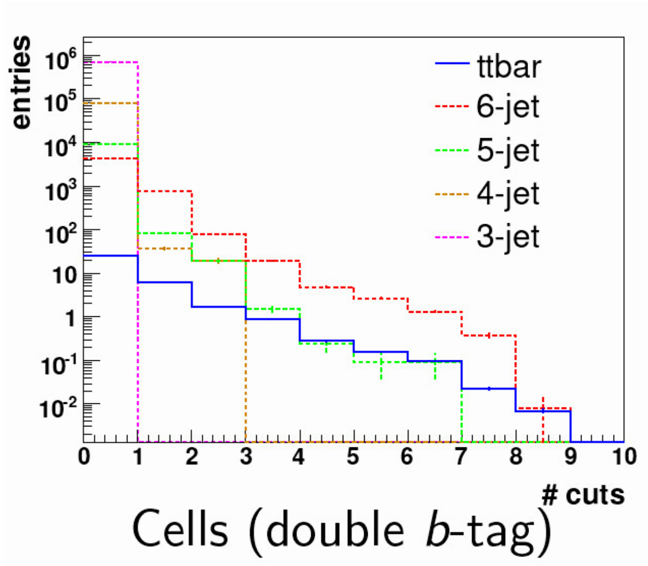
- expected relative error: $(\Delta\sigma/\sigma)^2$(eμ,100pb^{-1}) ~(4.4%)² (stat)



Cut and count (III)

- all hadronic channel: *M. Lambacher et al.* LMU Munich:
  - $t\bar{t}$ events generated with PYTHIA
  - main background QCD from ALPGEN (3,4,5,6+ fs partons)/PYTHIA
  - ATLFAST simulation study
  - S/B ~1/16, with b-tagging S/B ~1/1
  - Going towards likelihood measurements or neural network?

	Cuts
0	$\lvert\eta_{jets}\rvert < 3$
1	$N_{jets} == 6$
2	$p_T >$ 115,90,70,55,40,30 GeV
3	Σp_T (jets) > 140 GeV
4	75 GeV < m(jj) < 150 GeV
5	165 GeV < m (jjj) < 400 GeV
6	Σp_T (jjj) > 250 GeV
7	Aplanarity > 0.1
8	(double b-tag)



universitätbonn

Likelihood methods

$$Likelihood = -\sum_{i=1}^{N_{tot}} \ln(G(x_i \mid N_s, N_b)) + N_{tot}$$

- Multidimensional likelihood PDF $G(x_i \mid N_{signal}$ and $N_{background})$ with constraint on a fixed total number of events N_{tot}

- PDFs mostly obtained from MC, need trusted MC and high statistics
  - sensitive to certain signal and background shapes unlike cut and count

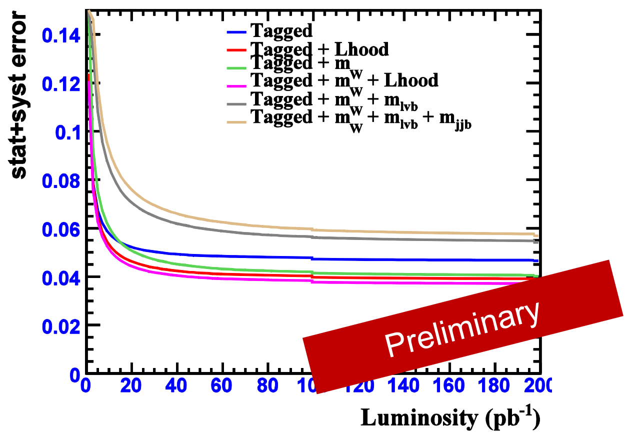
- Can fit more quantities (e.g. b-jet efficiencies) at the same time with sensible PDFs

Likelihood methods (II)

- lepton+jets channel: *F. Chevallier* LPSC Grenoble
  - Cross section extraction in 16 excusive channels
    - 1,2 b-tag, e or μ channel with 3,4,5,6+ jets

$$Likelihood = \prod_{16\,\text{channels}} P[N_{obs} \mid N_{pred}(\sigma_{t\bar{t}})]$$

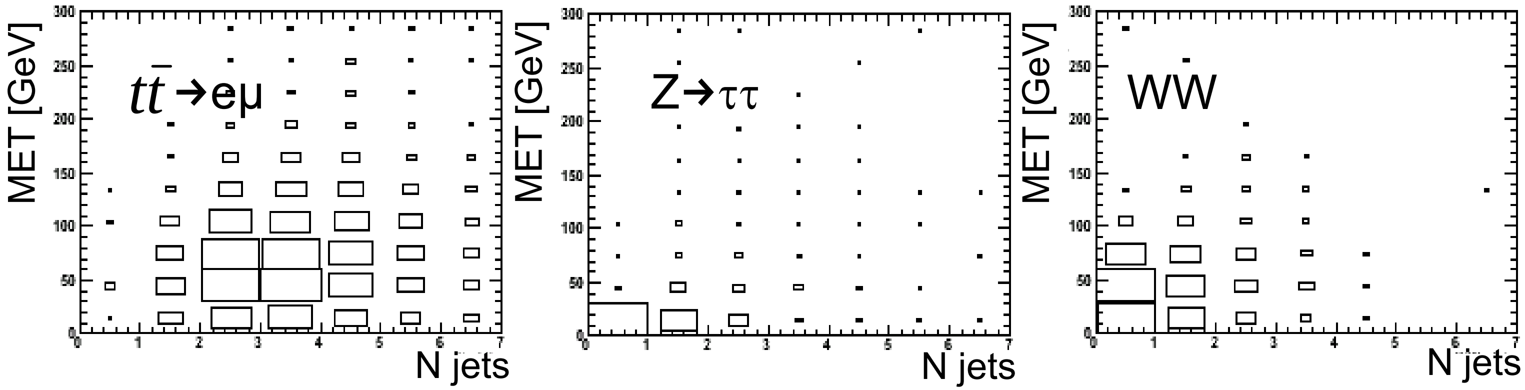
  - select events with lepton+jets cuts
    - one isolated lepton p_T>20 GeV $|\eta|$<2.5, $E_{T,Miss}$ > 20 GeV
    - at least 2 jets p_T >30 GeV, $|\eta|$<2.5, at least1 b-tag p_T >40 GeV, $|\eta|$<2.5
    - Hadronic W: m(jj) = m_W ± 30 GeV, associate b-jet giving best m_{top}
    - Likelihood selection on topologcal variables
  - sensitivity of method:
    - m_{top} and luminosity uncertainties excluded



universitätbonn

Likelihood methods (III)

- di-lepton channel: *J. Sjölin*, University of Stockholm
    - adopted from CDF method
    - backgrounds WW, $Z \rightarrow \tau\tau$
    - select di-lepton events:
        - $E_T(e) > 20$ GeV, $E_T(\mu) > 20$ GeV, Jets $E_T > 20$GeV (γ and τ are jets)
        - ee and $\mu\mu$ with MET/sqrt(ΣE_T)>2.5
    - binned fit in MET and number of jets space



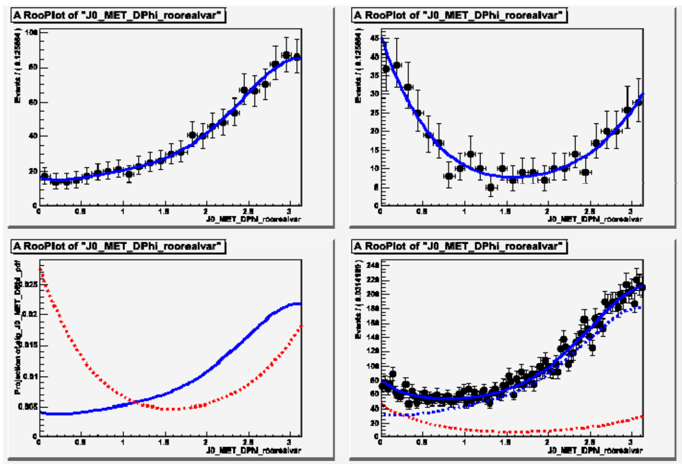
- uncertainties from pseudo experiments

Likelihood methods (IV)

- dilepton channel: *D.B. Ta et al.* University of Bonn
  - Selection of events with cut from the cut and count method
    - but have to include all channels due to limited MC statistics, MET>30 GeV
  - pdfs fitted to 5 good variables:
    - $|\Delta\phi|$ Lepton0 MET, $|\Delta\phi|$ Lepton1 MET, $|\Delta\phi|$ Jet0 MET, $|\Sigma\eta|$ Leptons, $|\Delta\phi|$ Leptons
  - Sensitivity with pseudo experiments and different S/B fractions:

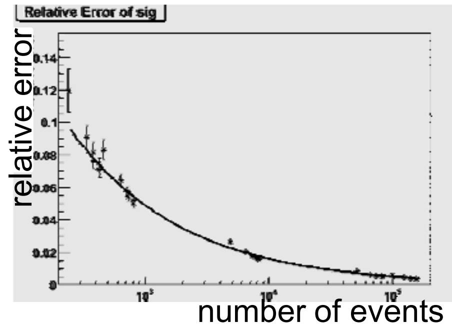
Example:
$|\Delta\phi|$ Jet0 MET

Normalized pdf-fit
→ fit and sum gives G



S/B 5:1 mix
and random data
from pdfs

Sensitivity (statistical error only):



$\Delta sig_{rel}(N)=\exp(a)\, N^{-b}$
$a=0.32\pm0.02$
$b=0.483\pm0.002$

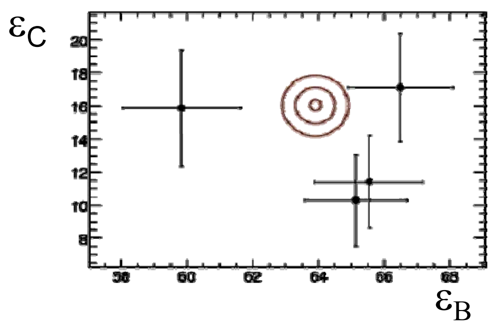
Likelihood methods (V)

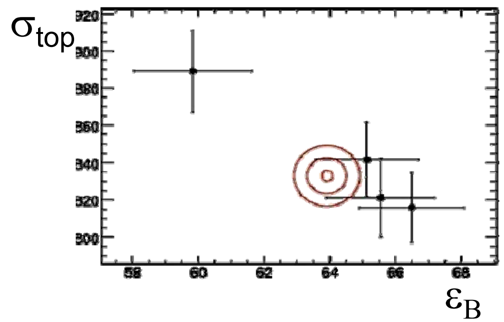
- lepton+jets and di-lepton channel: *H. Bachacou et al.* Saclay
  - simultaneous fit of ε(b-jet), ε(c-jet), N_{sig}
  - count number of events with 0,1,2,3 tagged jets
  - assume BR(t→Wb)=1 and ε(light jets)=0.01

$$< N_n >=$$
$$(L\cdot\sigma_{t\bar{t}}\cdot A_{pre-tag})\cdot\Sigma_{i,j,k}F_{i,j,k}\Sigma_{combi.}A_i^{i'}\cdot\epsilon_b^{i'}\cdot(1-\epsilon_b)^{i-i'}\cdot A_j^{j'}\cdot\epsilon_c^{j'}\cdot(1-\epsilon_c)^{j-j'}\cdot A_k^{k'}\cdot\epsilon_l^{k'}\cdot(1-\epsilon_l)^{k-k'}$$

  - select events:
    - one isol. lepton |η|<2.5, p_T>20 GeV, $E_{T,Miss}$>20 GeV
    - at least 4 jets E_T >40 GeV, |η|<2.5, b-tag weight > 6.5

mean uncertainty
for 4 subsamples
of 100pb^{-1}
ε_b 2%
ε_c 3%
σ_{top} 21pb



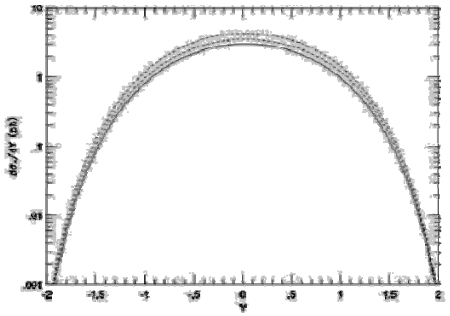


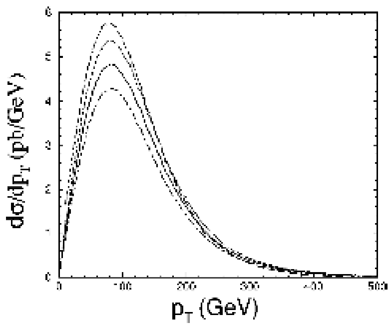
universitätbonn

Phase space resolved cross section

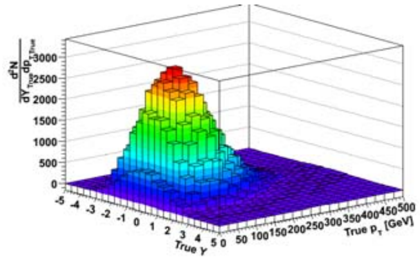
- lepton+jets channel: *M. Bunse et al.* University of Dortmund
  - assumed bin-size of 0.2x20GeV in $|\eta|<3$ and $p_T < 250$ GeV
    one has 375 phase space cells
  - for Tevatron only ~130 events/cell, LHC: ~21.000 events/cell

Theory

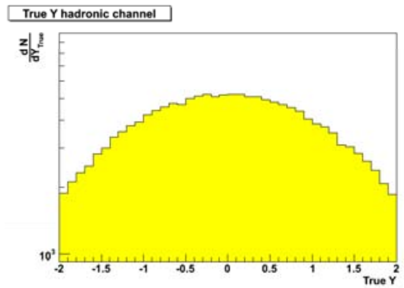


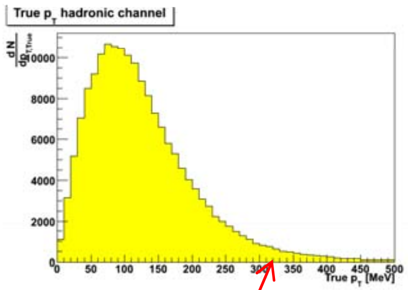


Top from truth

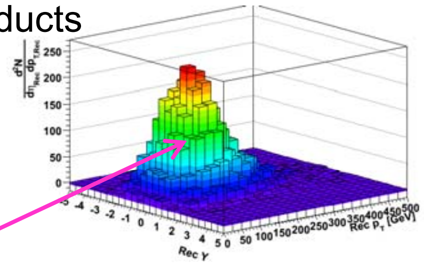


True events





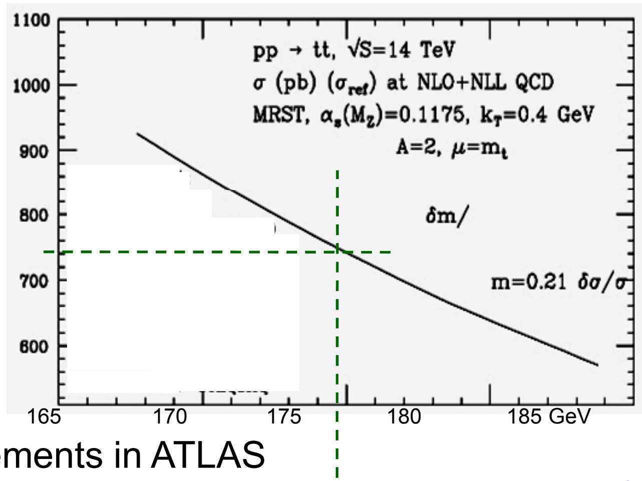
Top from reconstructed decay products



- regions with small statistics
- understand top from reconstructed events

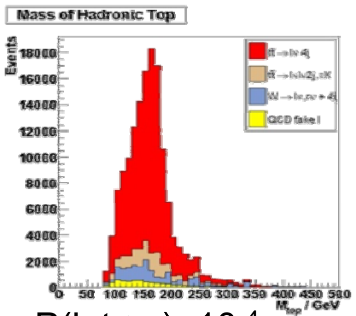
General strategies (III)

- cross section ratios
  - common uncertainties on *L*, higher order effects, PDFs etc. cancel out
- in streaming test data
  - testing different data stream modes (incl./exclu.)
  - 18pb^{-1} (~30 min @10^{33} ,3.6 Mevents) in 10 runs with different L
  - streaming test data as if real data (SM mix, no truth)
  - ideal for cross section exercise, test of “from data methods”: cross section ratio Z->ll/ttbar, lepton efficiencies from Z->ll
- Top mass vs. cross section
  - no one looked into that $\Delta m/m \sim 0.21 \Delta\sigma/\sigma$

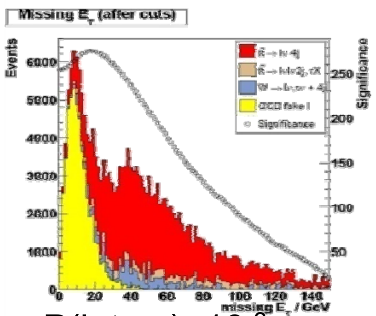


Cross section ratios

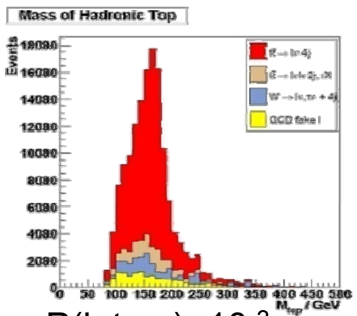
- cross section ratio lepton+jets and di-lepton channel: *R.Mameghani et al.* LMU Munich
    - focus on effect of fake rate
    - lepton+jets with QCD background and assumed fake rate jet→e



P(jet→e)=10^{-4}
MET>0 GeV



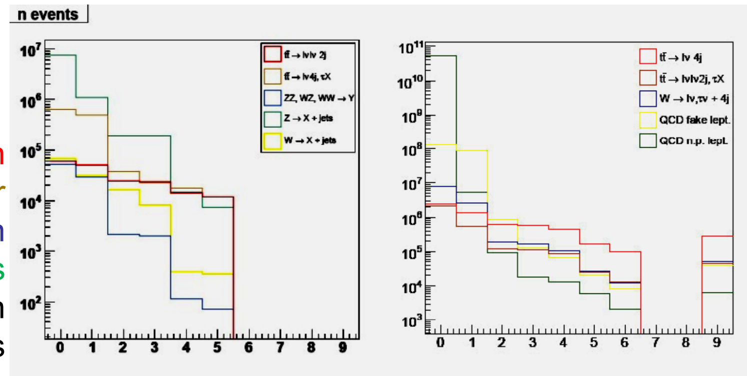
P(jet→e)=10^{-3}



P(jet→e)=10^{-3}
MET>30 GeV

- cut flow for both analysis with same fake electron rates 10^{-4}

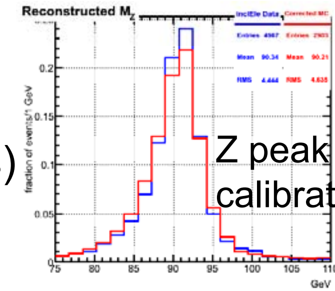
ttbar dilepton
other ttbar
diboson
Z+jets
yellow: W+jets with fake electrons



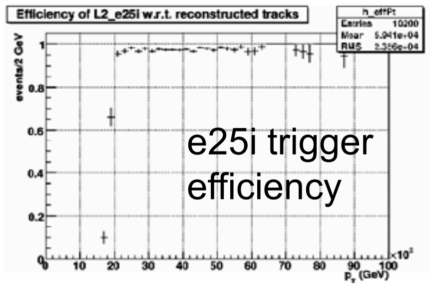
ttbar l+jets
other ttbar
QCD non prompt
yellow: QCD, fake

Streaming test analysis

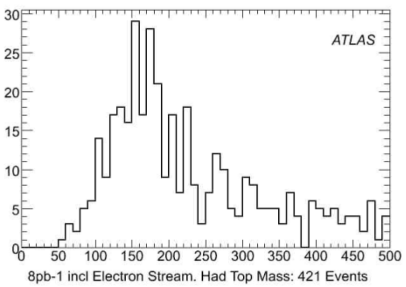
- streaming test data analysis: *A. Holloway, M. Shapiro et al.* LBL, *B. Acharya et al.* INFN Udine
  - Original CDF top discovery analysis in electron+jets
    - use tag and probe Z→ee (efficiencies/energy corrections)
    - measure W+0,1jet→e and predict W+4jet numbers, compare with data, interpret excess as top
    - $\sigma = (N(W+\geq 4)_{stream} - N_{W+4jets\ predicted})/(L_{stream}\ \varepsilon_{Trig}\ (A^{*}\varepsilon)_{reco})$
    - $\sigma = 941 \pm 96$ (stat) pb
    - in principle works on “data” !
  - similar method with Z+4,5jets events
    - also using three jet mass



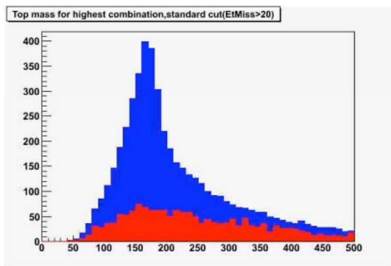
Z peak after calibration



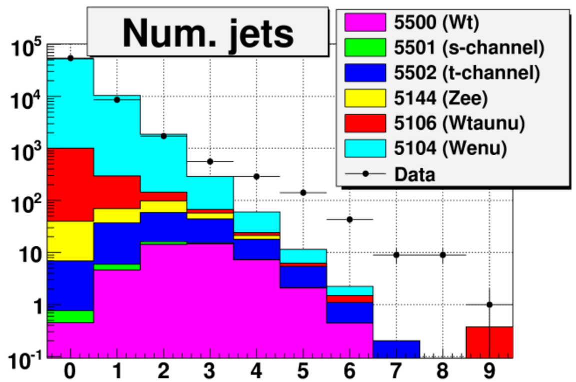
e25i trigger efficiency



in streaming test data



in MC data



General issues

- Obvious systematics
  - trigger evaluation
  - PDF uncertainties
  - ISR/FSR variation

- Main issue: additional leptons
  - hard to estimate from MC
    - from QCD background, from W+jets background
    - detector effects leading to fake leptons
    - non-prompt leptons from b-decay
  - a lot of work is put into developing determination methods
  - also need real data

Summary

- ATLAS analyses make use of many methods
  - cut and count
  - likelihood
  - phase space resolved cross section
  - cross section ratios
  - relative cross section measurement to standard candles / (efficiencies from) in streaming test data
- Measurement of the cross section gives a good understanding on selecting top events
  - need knowledge to reject the main background for new physics
- Analyses will give final results soon in CSC note: **“T6:Top Cross Section”**