$Z \rightarrow \tau\tau \rightarrow \mu\ \tau_h\ 3\nu$

Data driven background estimation method for QCD and W+jets

BMBF-Forschungsschwerpunkt
ATLAS Experiment
Physics on the TeV-scale at the Large Hadron Collider
FSP 101
ATLAS

GRADUIERTEN KOLLEG
Masse-Spektrum-Symmetrie

PHYSICS AT THE TERA SCALE
Helmholtz Alliance

Frank Seifert

Supervised by Michael Kobel, Wolfgang Mader and Arno Straessner

TECHNISCHE UNIVERSITÄT DRESDEN

Introduction and cut flow

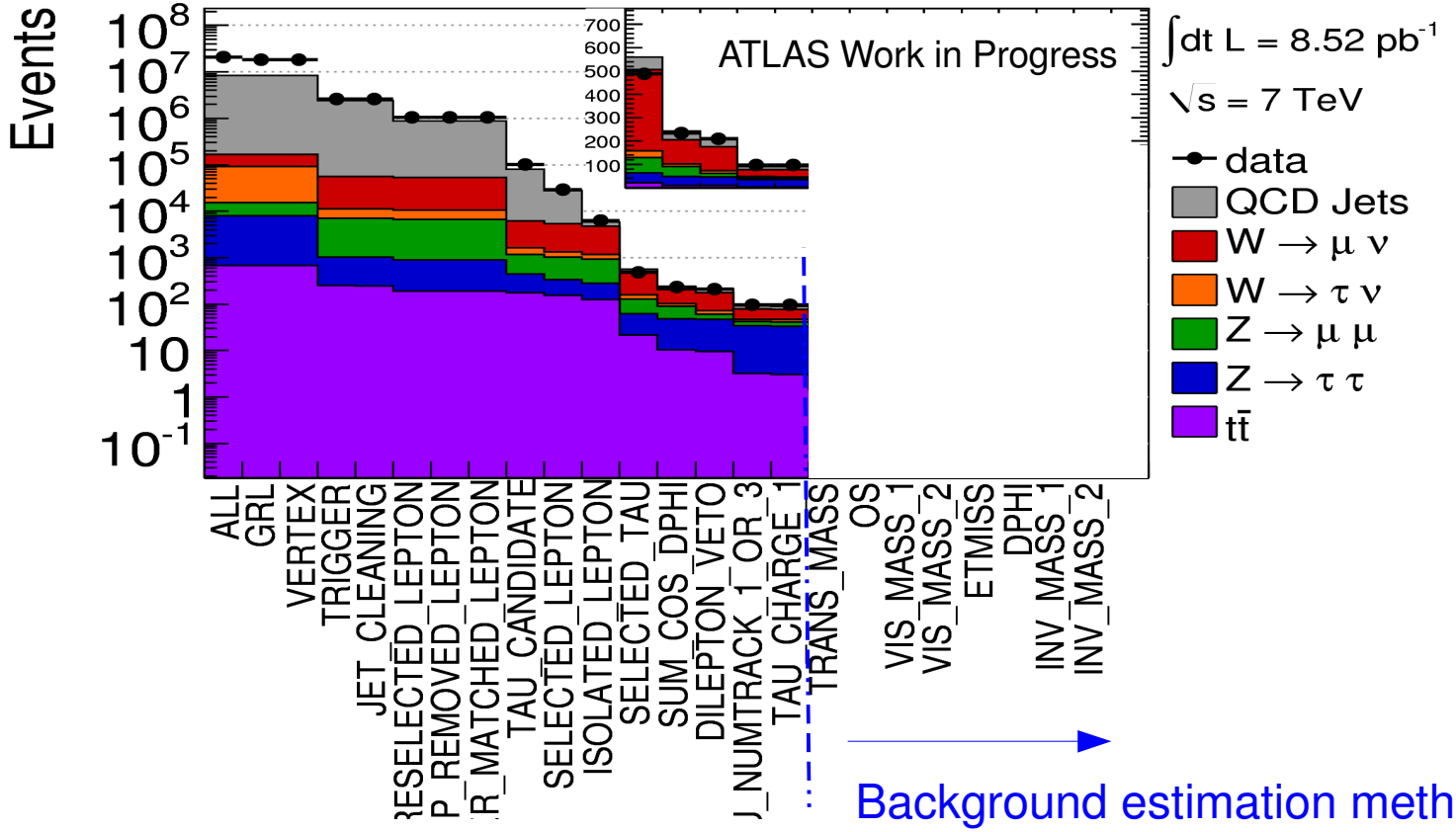
Background study is based on the $Z \rightarrow \tau\tau \rightarrow \mu\, \tau_h\, 3\nu$ analysis package of the ATLAS $Z \rightarrow \tau\tau$ working group.

Cut flow:

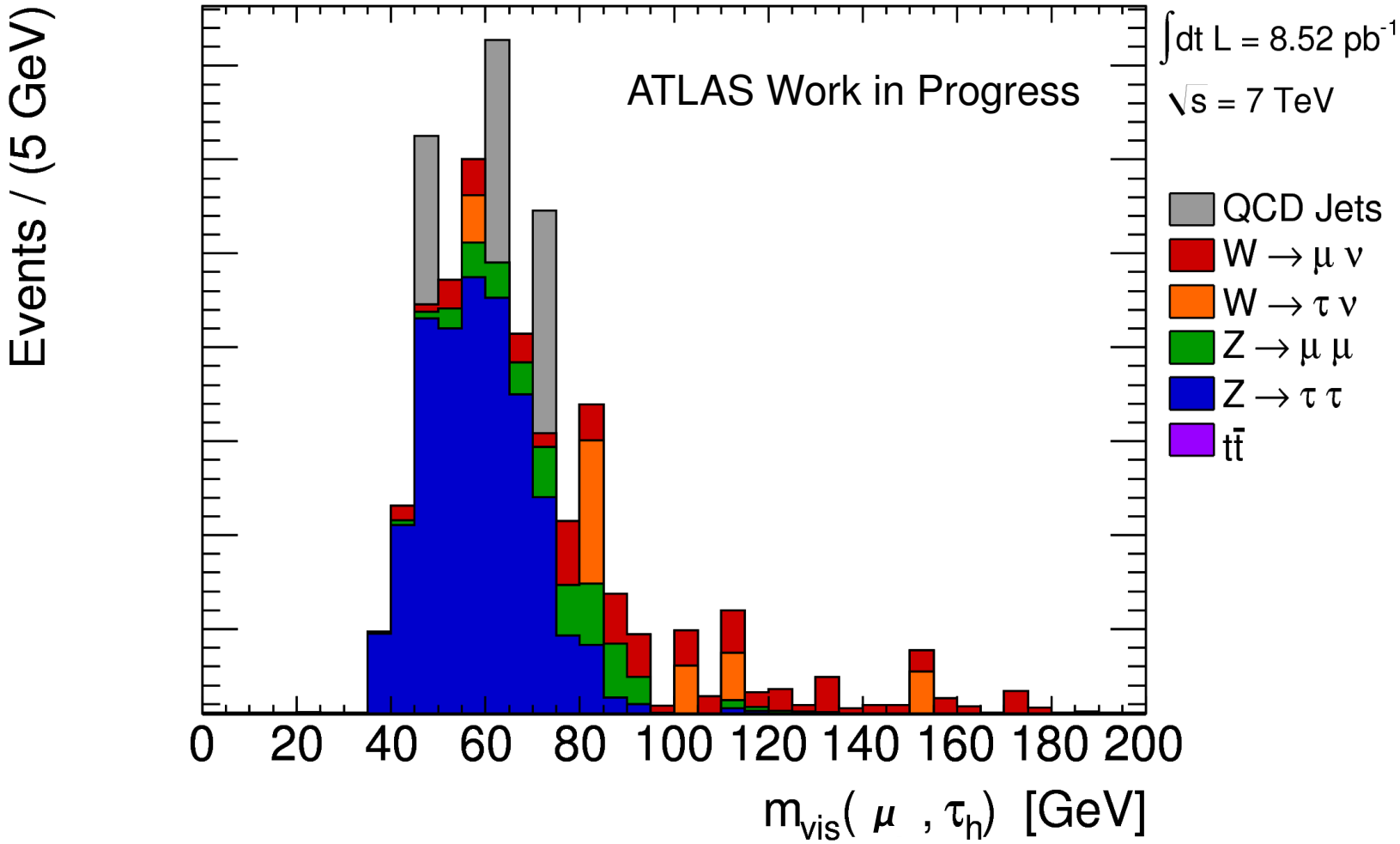
- ≥1 reconstructed tau candidate
- ≥1 muon (reconstructed in muon spectrometer matched to ID track), $p_T > 15$GeV
- Require muon isolation
- Identified tau passing strong quality criteria
- $\cos[\varphi(\mu)-\varphi(E_T^{miss})] + \cos[\varphi(\tau)-\varphi(E_T^{miss})] > -0.15$
- Dilepton veto
- τ_h : $N_{track} = 1$ or 3, $|charge| = 1$
- Transverse mass $M_T(\mu, E_T^{miss}) < 50$GeV
- Opposite charge of μ and tau
- $35\text{GeV} < M_{vis}(\mu, \tau_h) < 75\text{GeV}$
- $E_T^{miss} > 20$ GeV

Overview of event yields

- Data period D-G (8.52pb^{-1}) is used.
- Monte carlo (MC) with <$N_{\text{primary vertex}}$>=2 is used and a vertex reweighting method from data to MC is applied.



Visible mass distribution after M_T > 50GeV and OS cut



→ **Need to estimate QCD and W+jets background contribution from data**

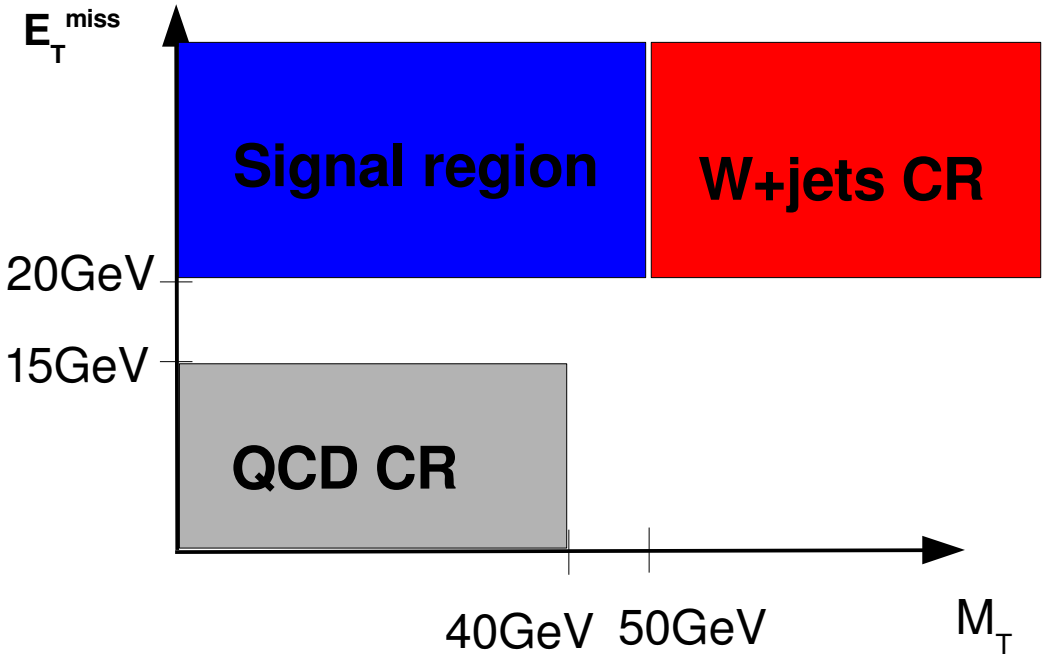
Background estimation method

➔ **Method should rely as much as possible on data**

• **Use control regions (CR) in Z → ττ → μ τ_h 3ν selection, where background is enriched → use OS/SS ratio of background contributions**

Signal region	QCD CR	W+jets CR
E_T^{miss} > 20GeV	E_T^{miss} < 15GeV	E_T^{miss} > 20GeV
M_T < 50GeV	M_T < 40GeV	M_T > 50GeV

+ Drop lepton isolation to increase statistics in the QCD control region!



Calculation of the background components

$$N_{OS,Bkg.}^{SigReg} = R_{OS/SS,QCD} * N_{SS,QCD}^{SigReg} + R_{OS/SS,W} * N_{SS,W}^{SigReg}$$

From Data QCD CR

$$= R_{SR/CR,SS,QCD} * N_{SS,QCD}^{CR}$$

From MC

From Data QCD CR

→ Trust only on MC E_T^{miss} shape of QCD

Assumptions to be proven:

- $R_{OS/SS}$ is the same in signal region (SR) and CR for each background component
- E_T^{miss} and M_T distributions are the same for OS and SS events

OS and SS events in the control regions with isolation

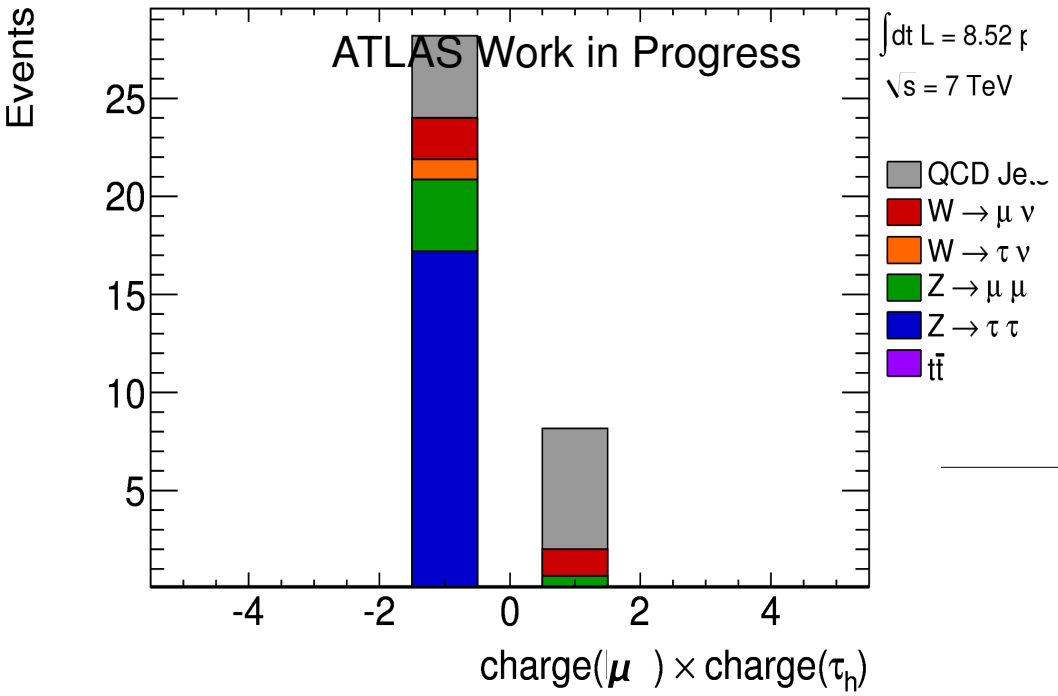
Number of events left in signal- and control regions.

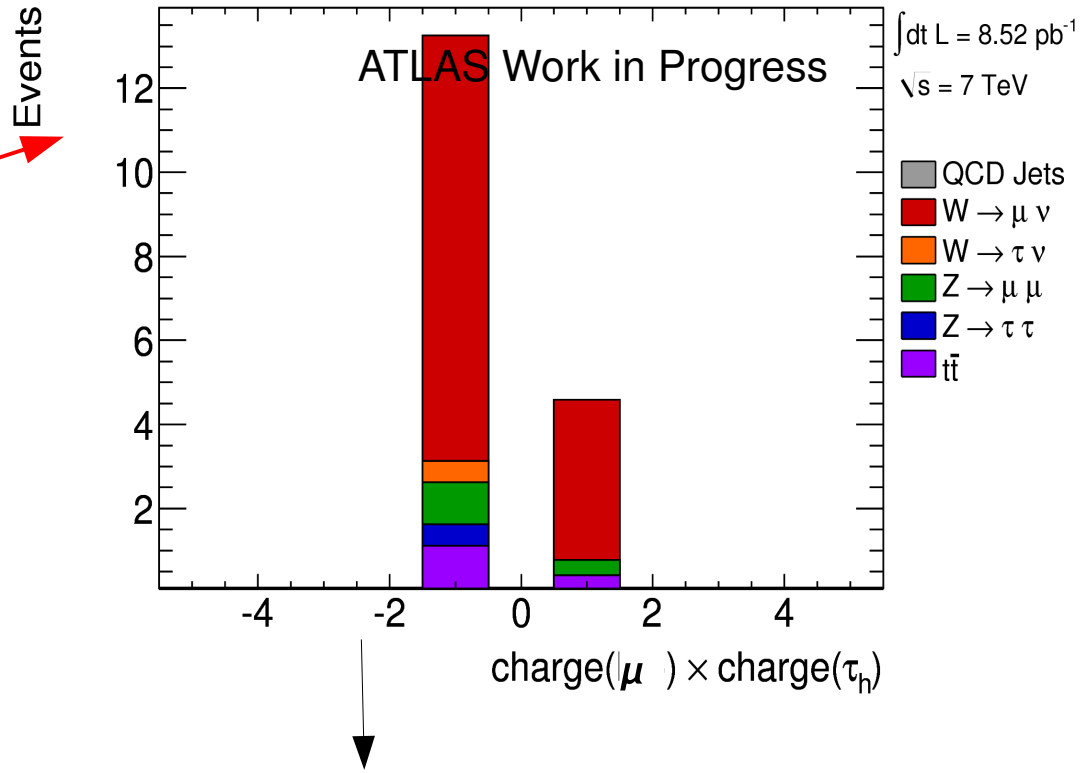
OS and SS

With muon isolation

QCD CR

W+jets CR





- W+jets CR dominated by W+jets events

- QCD CR dominated by signal events

➔ Drop isolation criteria on muons to increase QCD contribution

OS and SS events in the control regions without isolation

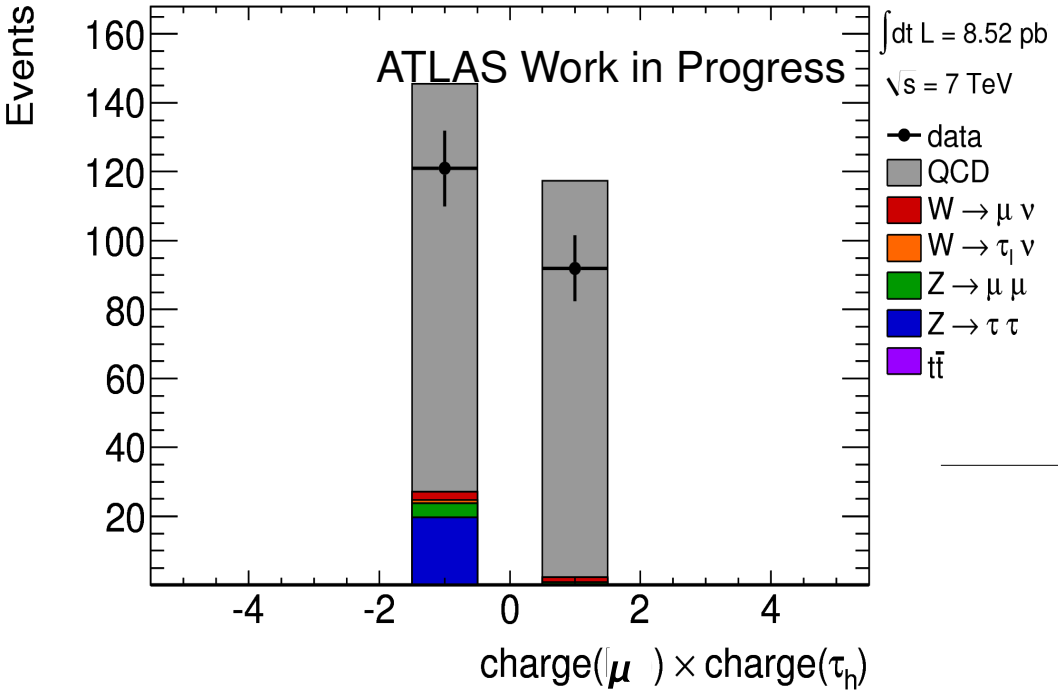
Number of events left in signal- and control regions.

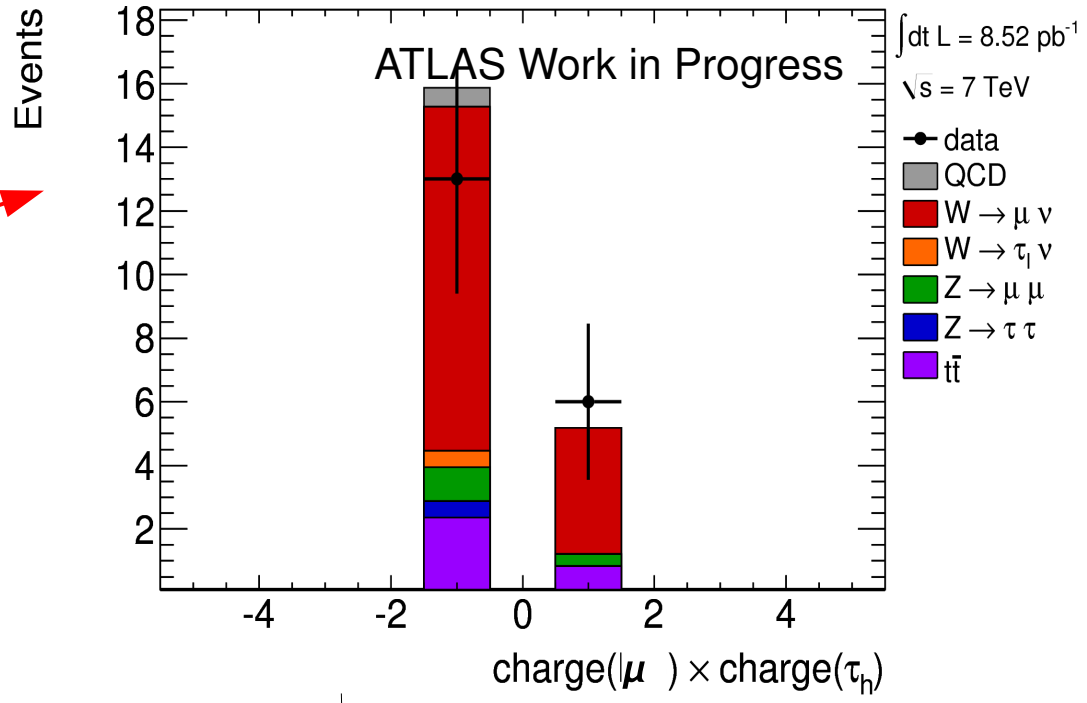
OS and SS

Without muon isolation

QCD CR

W+jets CR

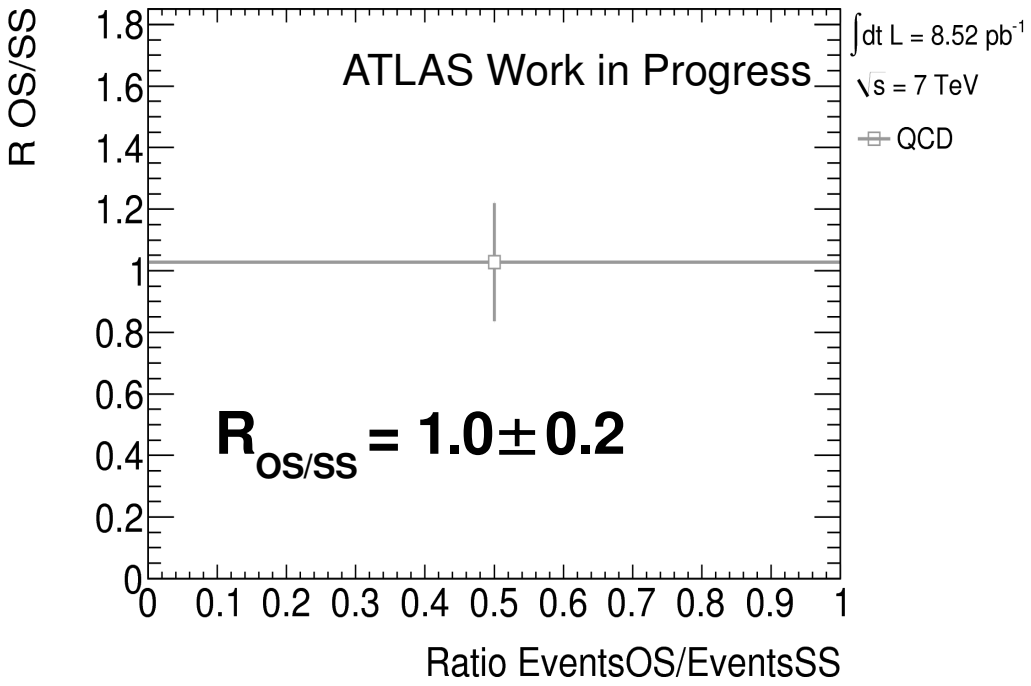




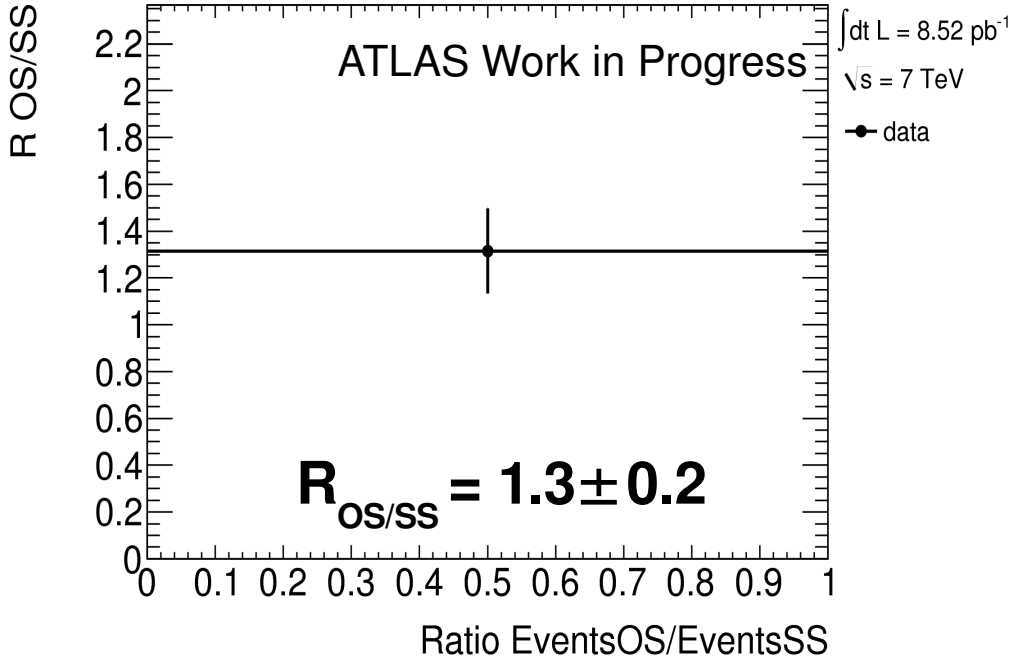
- W+jets CR dominated by W+jets events
- QCD CR now dominated by QCD events
- Normalization of QCD CR to be investigated

OS/SS ratio of QCD events from QCD CR

Events QCD MC (QCD CR)



Events Data (QCD CR)



→ **OS/SS ratio of Data and MC in QCD CR compatible, but big uncertainties**

→ **can be improved**

Preliminary Numbers for QCD

$N_{QCD,SS}^{SR} = 8.6 \pm 4.0$ → statistical uncertainty limits precision

$N_{QCD,SS}^{QCD\ CR} = 115 \pm 15$

$N_{Data,SS}^{QCD\ CR} = 92 \pm 10$

$$N_{OS,QCD}^{SigReg} = R_{OS/SS,QCD} * N_{SS,QCD}^{SigReg}$$

From Data QCD CR

$$= R_{sig/CR,SS,QCD} * N_{Data,SS}^{QCD\ CR}$$

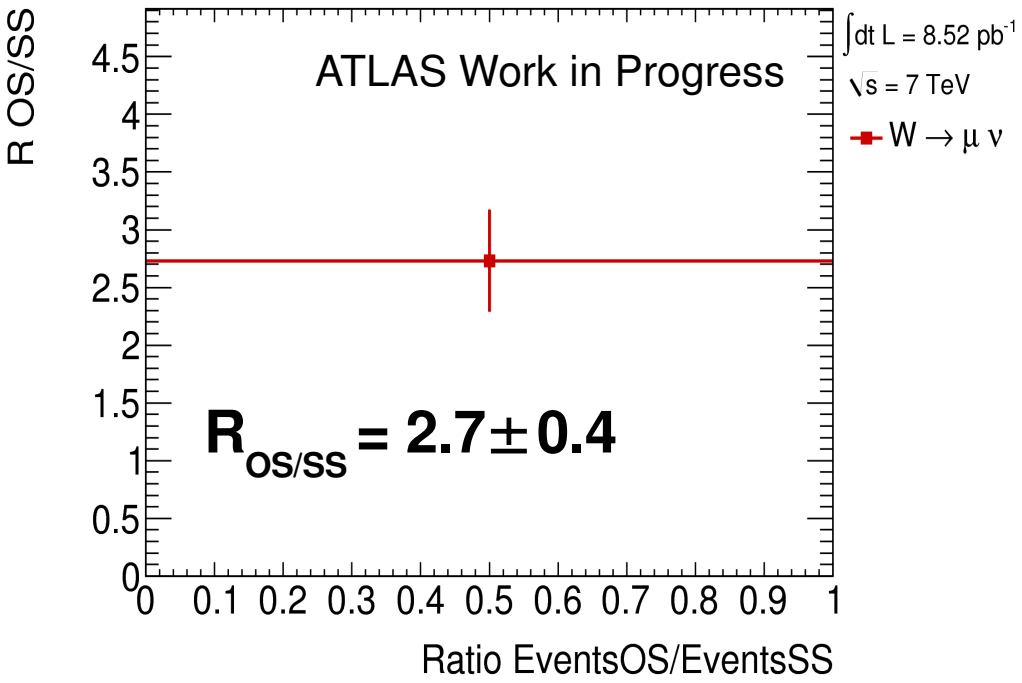
$$\text{QCD} = (1.3 \pm 0.2) * (0.075 \pm 0.036) * (92 \pm 10)$$

$$\text{Background:} = \underline{9.0 \pm 4.6}$$

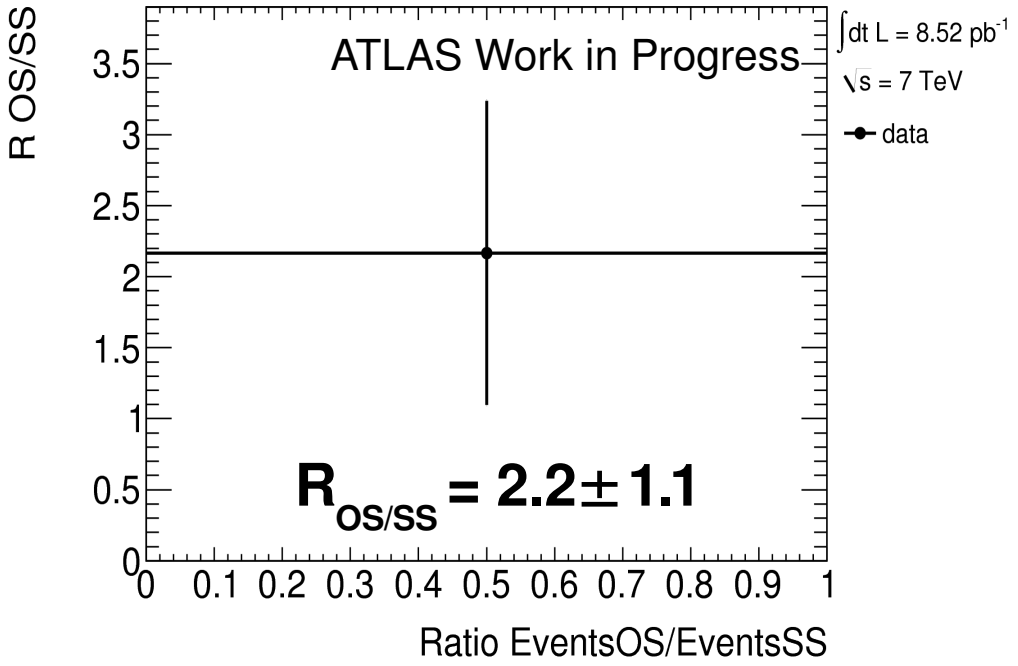
Numbers without lepton isolation criteria!

OS/SS ratio of W+jets events from W CR

Events MC W+jets (W CR)



Events Data (W CR)



→ **Low statistics in W control region**

→ **Could modify the W CR to get higher statistics**

Preliminary Numbers for W+jets

$N_{W,SS}^{SR} = 1.8 \pm 0.4$

$N_{W,SS}^{W\ CR} = 4.0 \pm 0.5$

$N_{Data,SS}^{W\ CR} = 6.0 \pm 2.5$ → statistical uncertainty limits precision

$$N_{OS,W}^{SigReg} = R_{OS/SS,W} * N_{SS,W.}^{SigReg}$$

From Data W CR

$$= R_{sig/CR,SS,W} * N_{Data,SS}^{W\ CR}$$

For W+jets:

$$= (2.2 \pm 1.1) * (0.45 \pm 0.11) * (6.0 \pm 2.5)$$

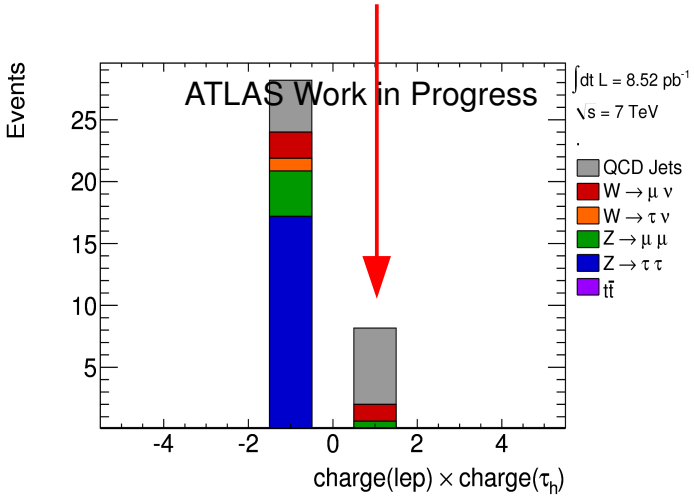
$$= \underline{5.9 \pm 3.4}$$

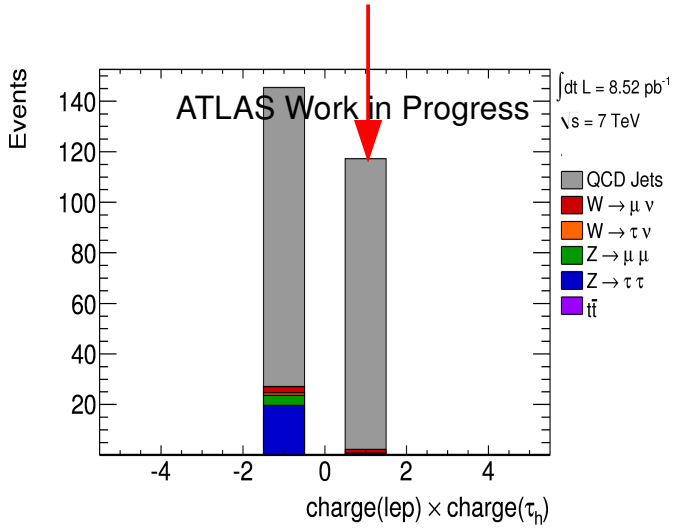
Application to Signal region <u>including</u> lepton isolation

- **Take the Number of SS events in the QCD CR with and without lepton isolation to get the difference of QCD events in the SR with and without isolation**

→ Take $R_{noIso/Iso}(N_{SS}^{QCD\ CR})=(9.0\pm2.9)$ from data to scale down to SR with isolation

QCD CR →





→ Divide the (9.0±4.6) events by (9.0±2.9) to get QCD estimate in SR <u>including</u> lepton isolation

Preliminary numbers including lepton isolation

Taking into account the efficiency of lepton isolation leads to the following number of QCD- and W+jets events in the signal region:

Estimated number of events:

- **1.0±0.6 QCD events in signal region**
- **5.9±3.4 W+jet events in signal region**

Crosscheck with Monte Carlo numbers:

- **0.0 QCD events in signal region**
  (uncertainty to be investigated)
- **4.5±1.1 W+jet events in signal segion**

Conclusion

- The estimated numbers for QCD and W+jets backgrounds match with the MC prediction
- → The method looks in principle promising

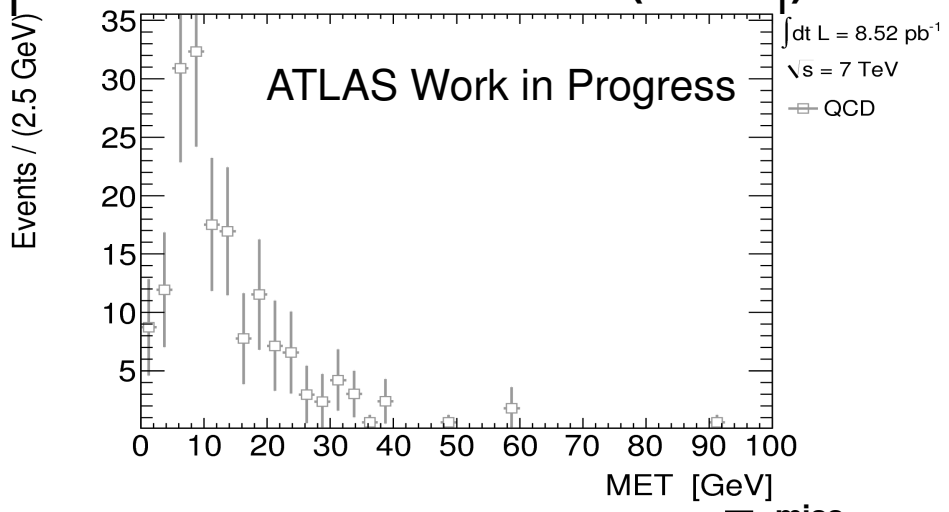
- Statistical uncertainties will become smaller with more data available
- → Could evaluate method to estimate cross-feed between signal and control regions ("tiles method", "ABCD method")

- QCD CR contains too much signal → drop cut on E_T^{miss} for signal selection → this method cannot be applied in the near future
- → Can be used at a later step, when selecting the signal for calculating the inv. $Z \rightarrow \tau\tau$ mass

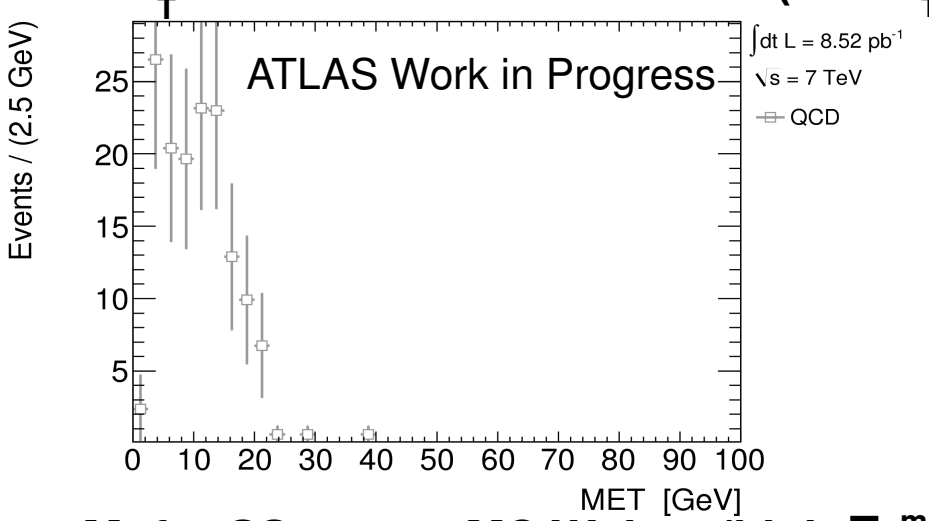
Backup

Monte carlo control histograms

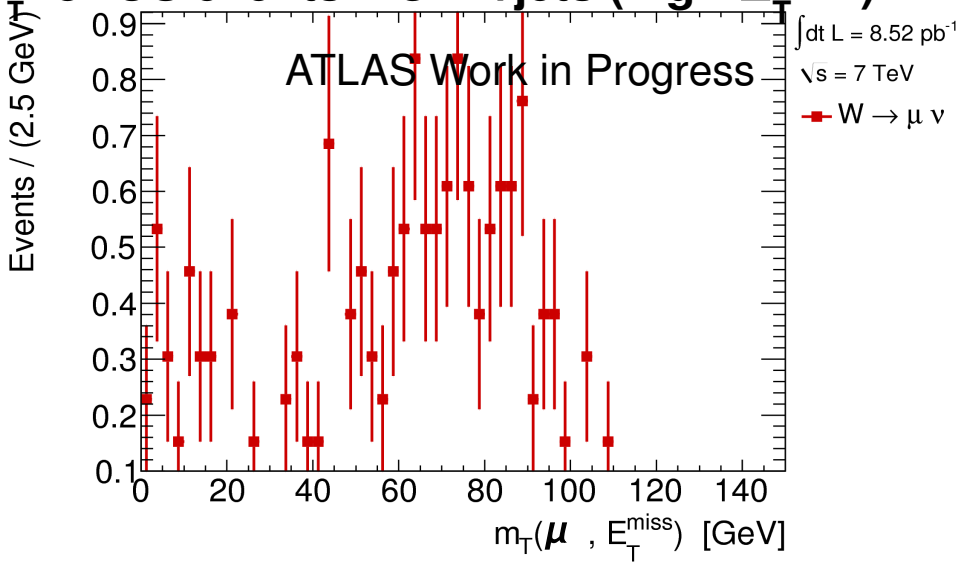
E_T^{miss} for OS events MC QCD (low M_T)



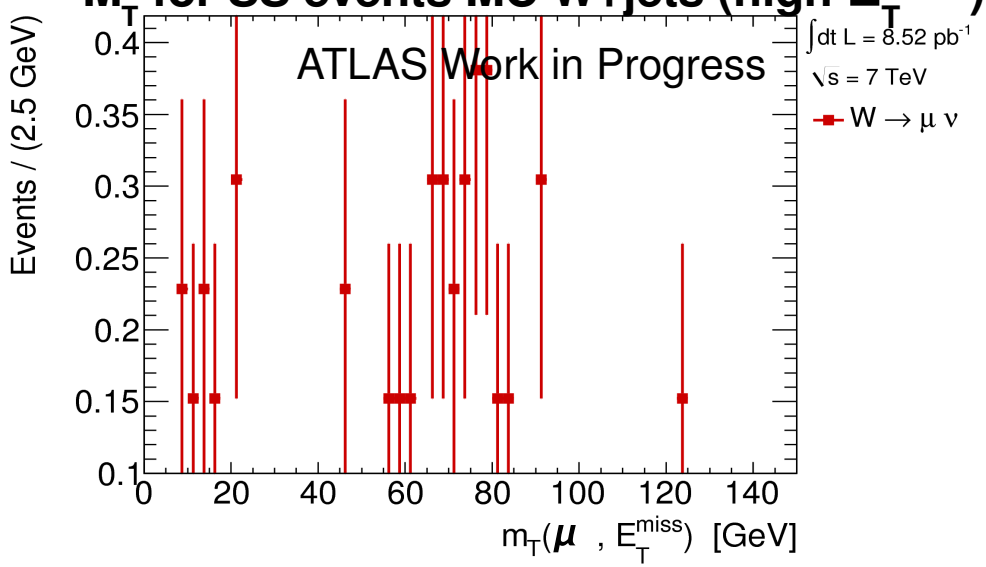
E_T^{miss} for SS events MC QCD (low M_T)



M_T for OS events MC W+jets (high E_T^{miss})



M_T for SS events MC W+jets (high E_T^{miss})



→ Too low MC statistic for good proof of assumptions