QCD background modelling for semi-leptonic top pair production

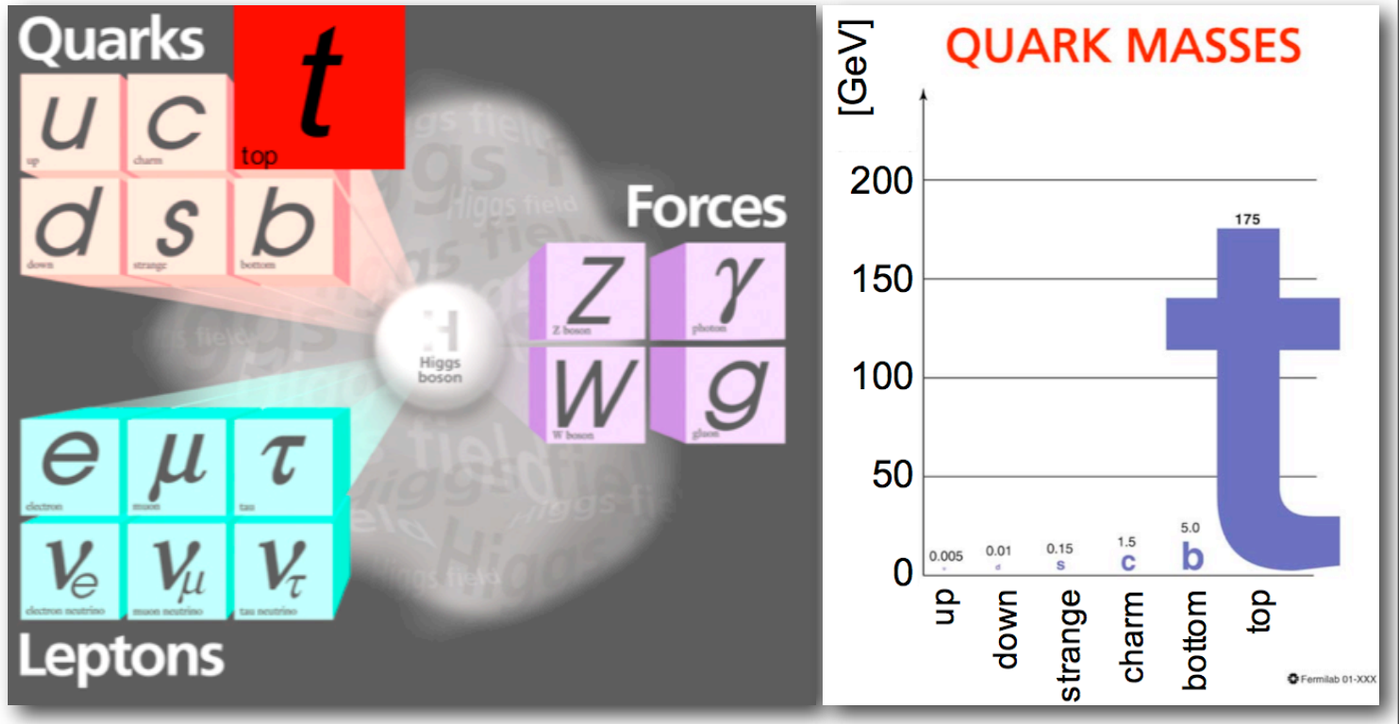
towards a top pair production cross-section measurement

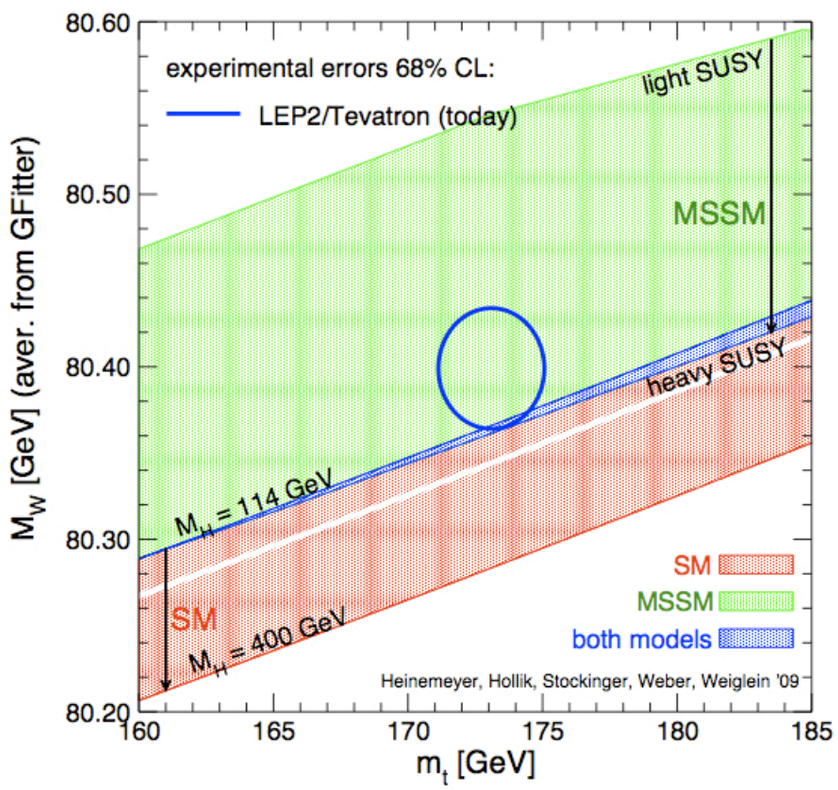
Clemens Lange (DESY)
Research Training Group
Mass, Spectrum, Symmetry (GK 1504)
Block Course
Humboldt-Universität zu Berlin
15th October, 2010

HELMHOLTZ ASSOCIATION

reminder: the top quark

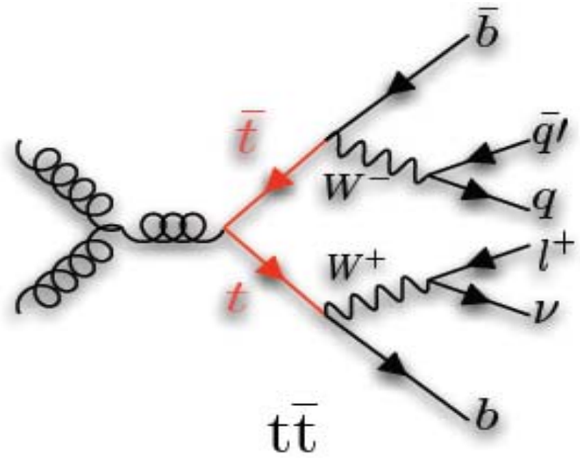
- discovered in 1995 at Tevatron
- expected properties:
  - spin 1/2, charge +2/3 e
- measured mass:
  (173.3 ± 1.1) GeV (arXiv:1007.3178 [hep-ex])
- average life time 4×10^{-25} s shorter than characteristic QCD hadronisation time of 28×10^{-25} s
  - no bound states (quarkonia)
  - spin is conserved
- perform SM precision measurements at LHC experiments
- constrain Higgs mass via loop corrections of W mass





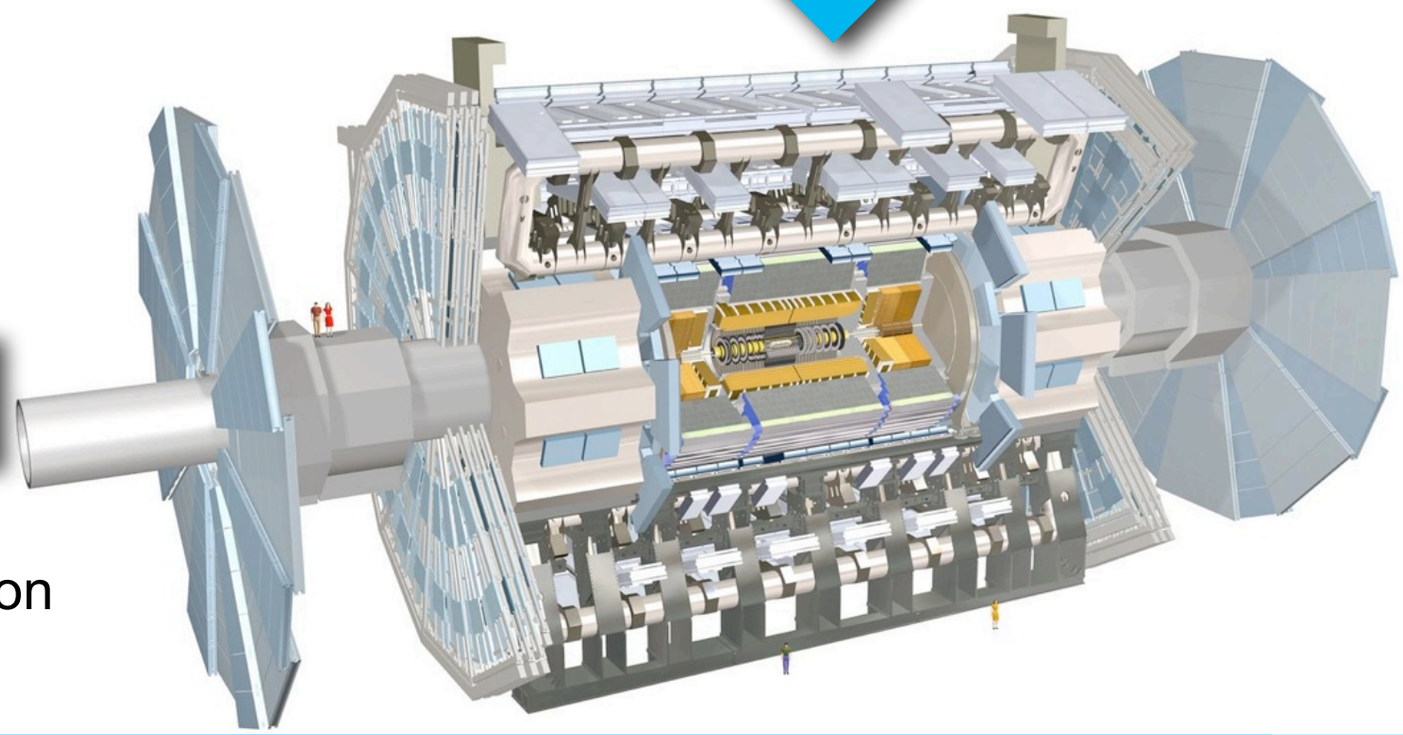
top quark measurement at ATLAS

pp collisions
at 7 TeV

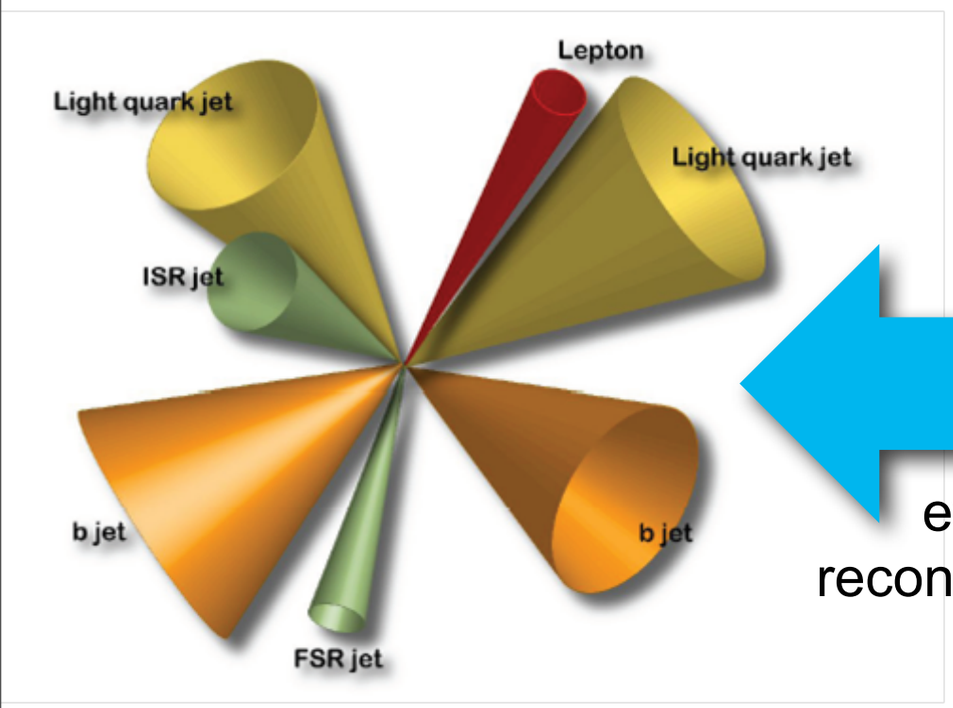


$t\bar{t}$

decay of top quark

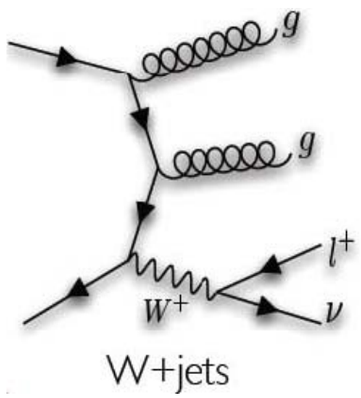


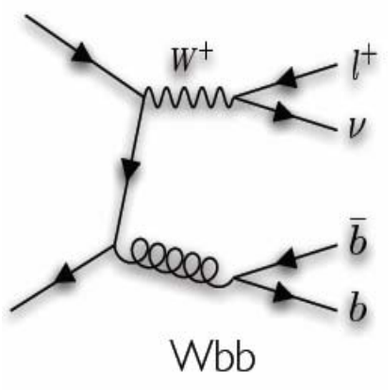
event reconstruction

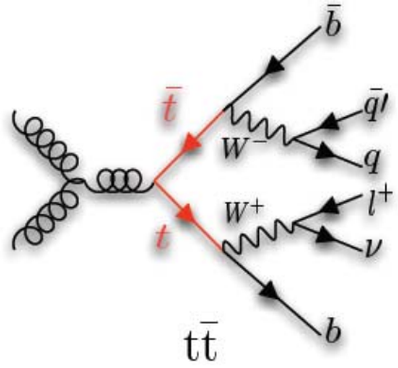


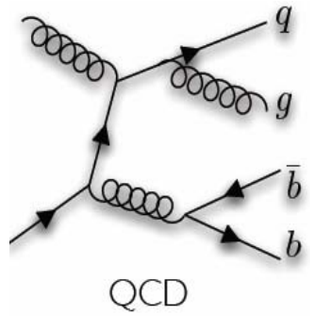
top quark measurement at ATLAS (2)

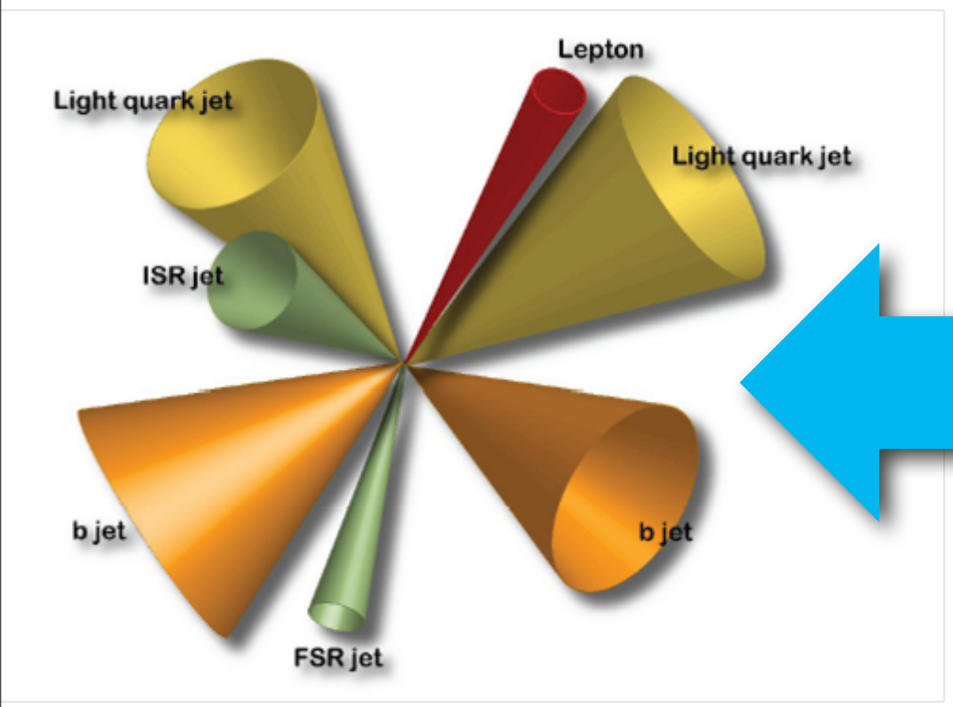
?

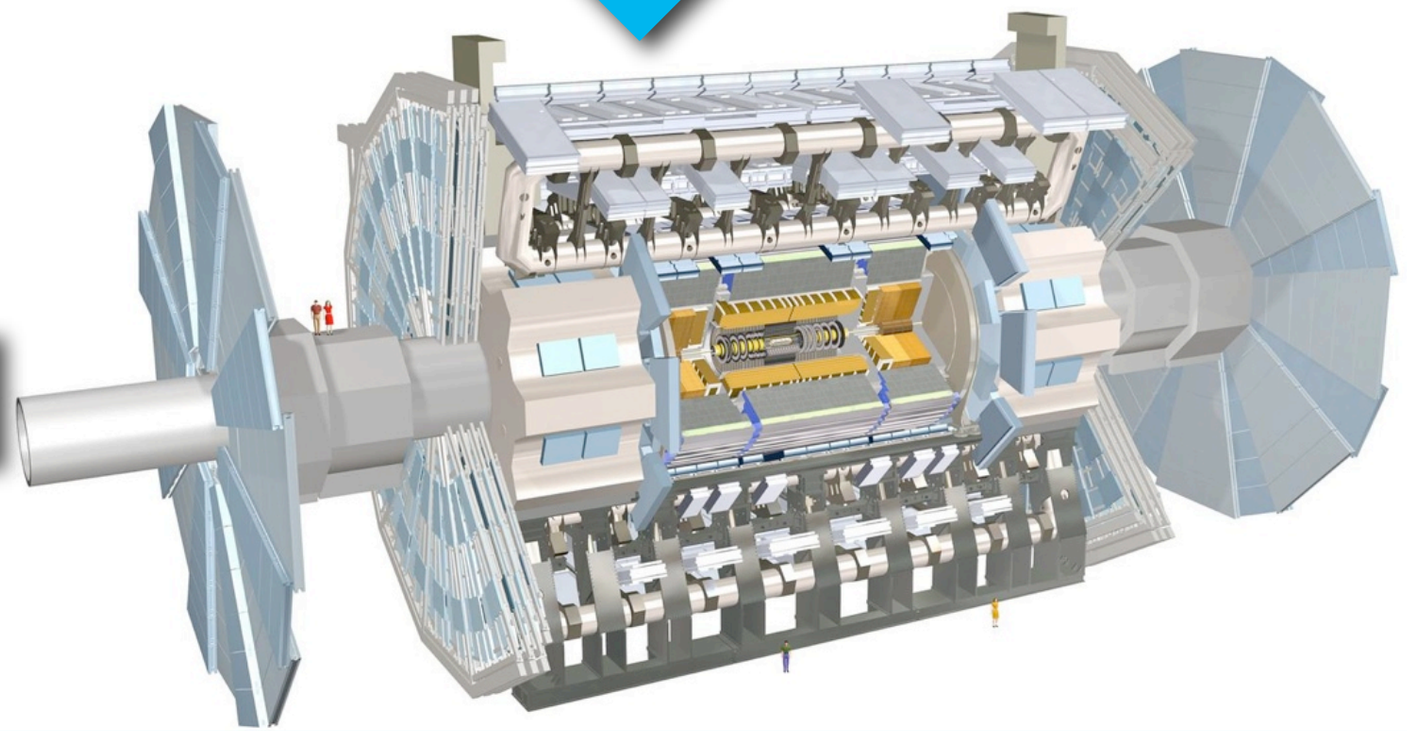




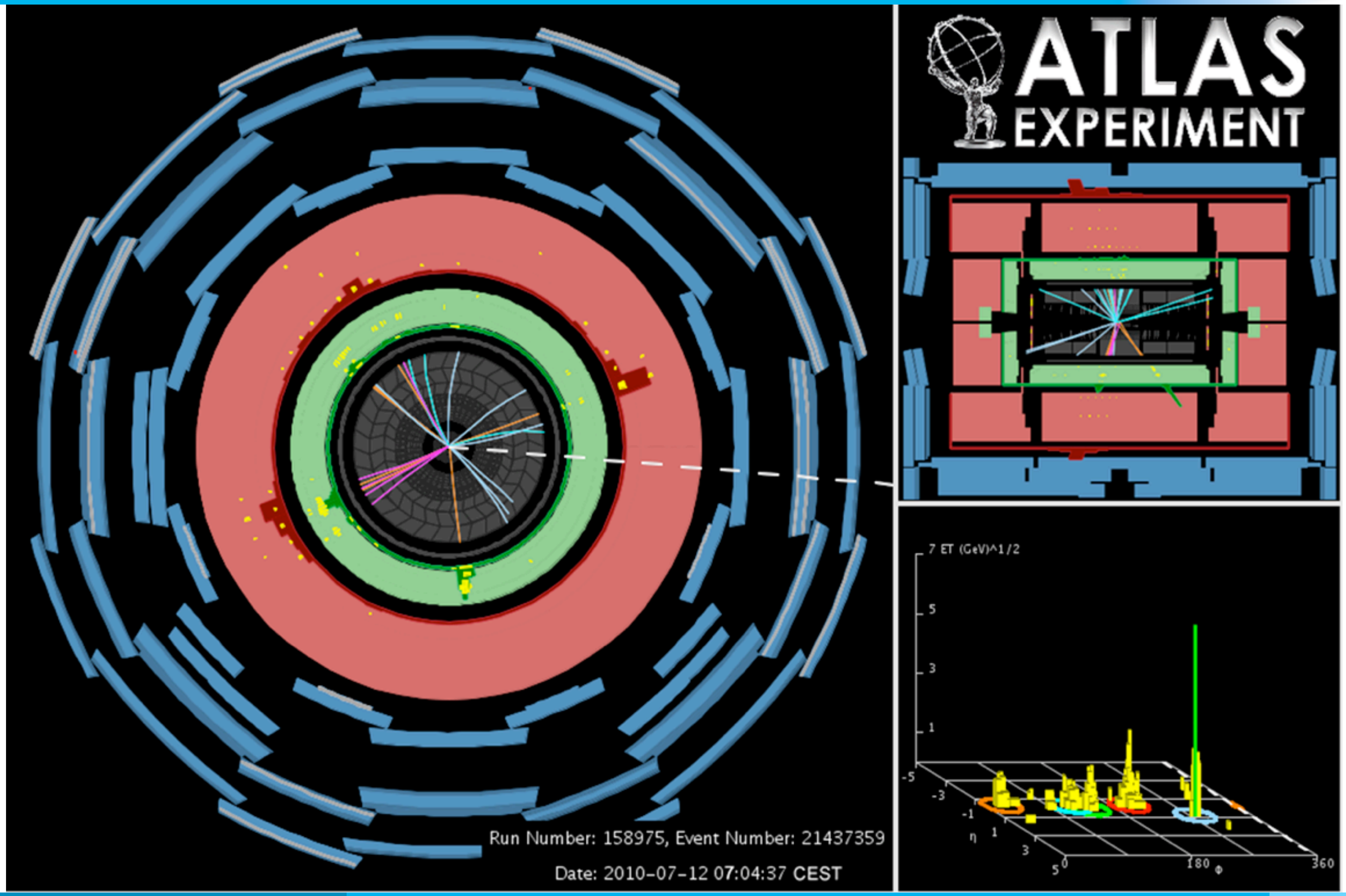








top quark candidate at ATLAS



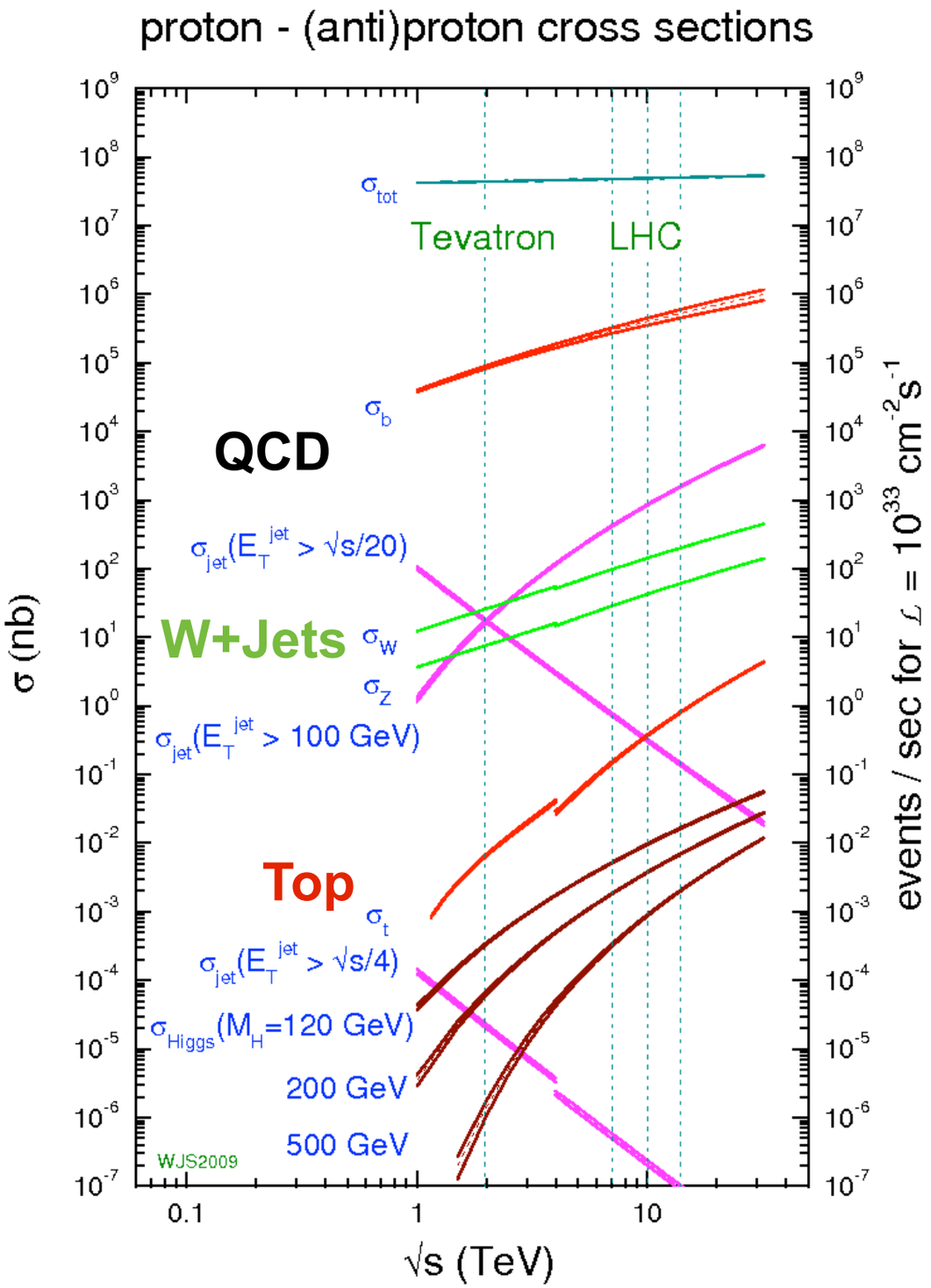
top cross-section measurement and backgrounds

- signal and background cross-sections calculated in theory
- top pair production at 7 TeV and m_t = 172.5 GeV (Moch, Uwer):

$$\sigma_{t\bar{t}} = 164.57^{+8.34}_{-11.33}\ \mathrm{pb}$$

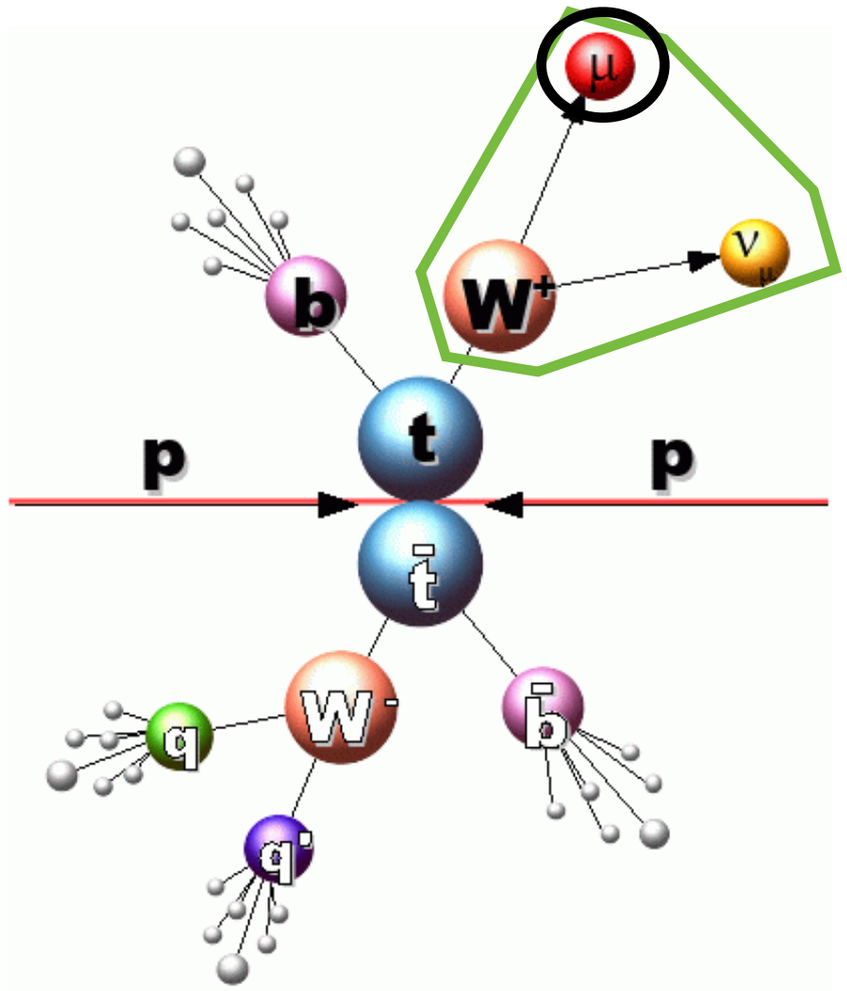
- W+jets production: σ ~ 30 nb
- QCD production: σ ~ μb/mb
- to calculate cross-section in cut and count experiment, need to evaluate background:

$$\sigma = \frac{N_{\mathrm{tot}} - N_{\mathrm{bg}}}{\epsilon \times \mathcal{L}}$$



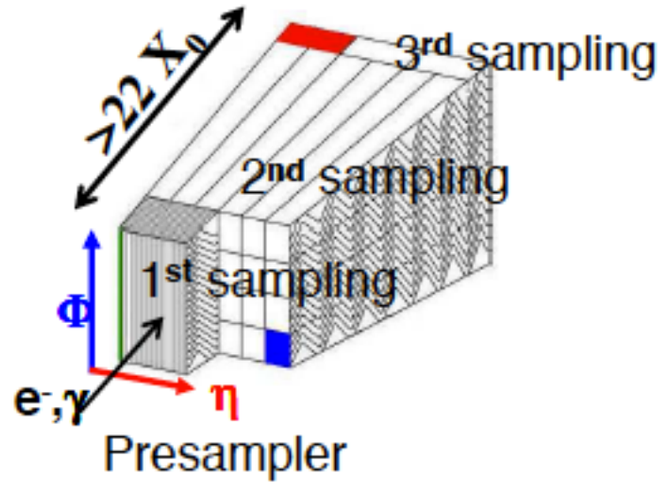
estimating background contribution to signal selection

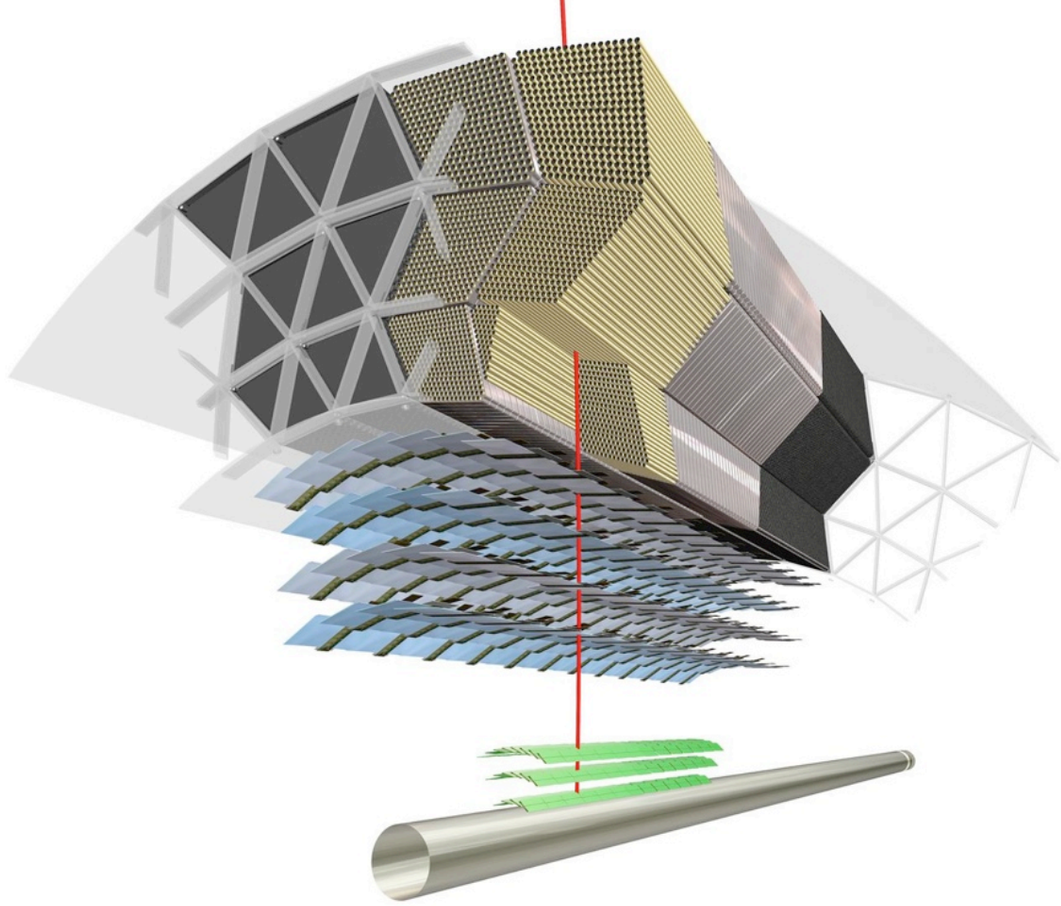
- model all processes in MC
- apply tight selection cuts to reject backgrounds:
  - select high quality lepton
  - select ≥ 4 high-p_T jets
  - require b-tag
  - cut on missing transverse energy (neutrino)
- cannot reject W+jets background completely due to similar topology
- QCD contributes only due to mis-identified leptons (electron fake rate of order 10^{-3}-10^{-4})
- QCD MC production limited by computing power
- mostly instrumental background, difficult to model → **take from data**



intermezzo: electron identification

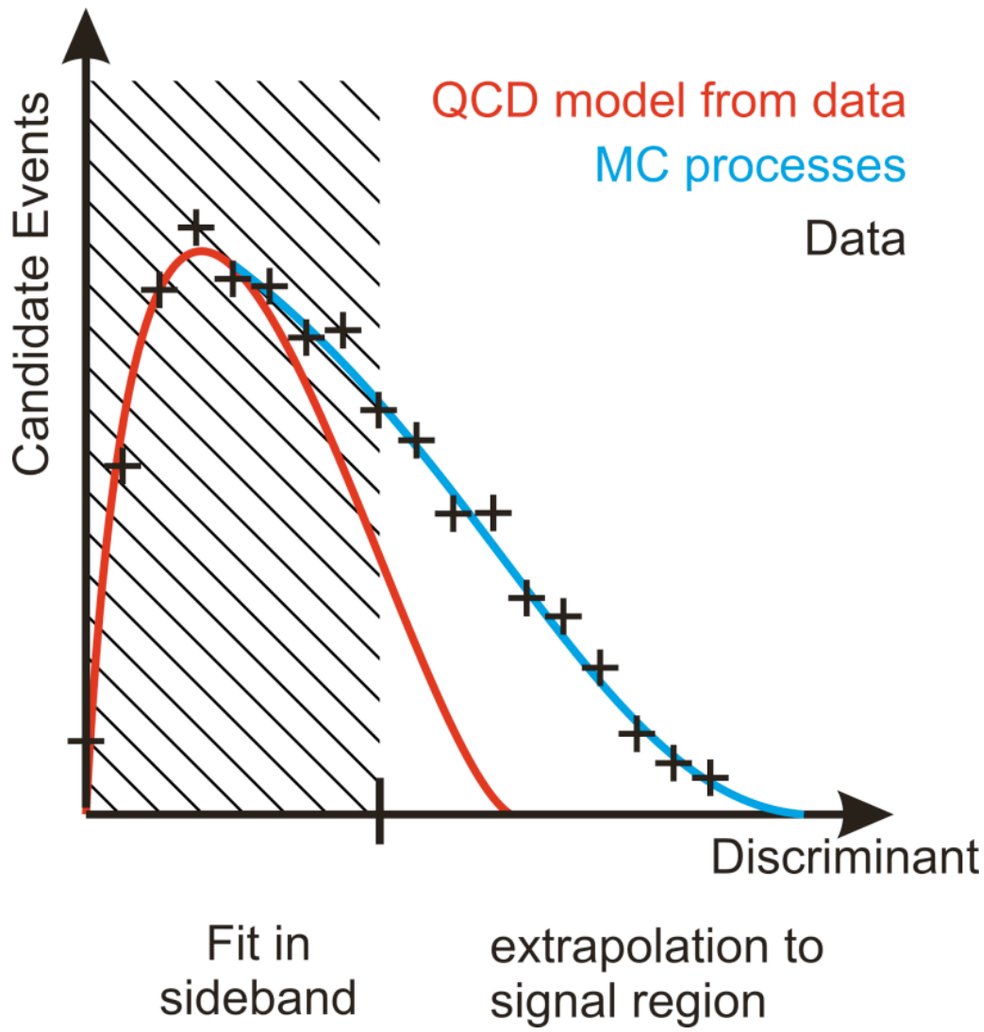
- electron produced at interaction point travel through inner detectors
- track can be reconstructed from hits
- electron is then stopped in electro-magnetic calorimeter
- to obtain high-quality electron candidates we can cut on (particle identification (PID) variables):
  - inner detector track quality
  - track match with calorimeter cluster
  - shower shape
  - leakage into hadronic calorimeter
- also require isolation





modelling QCD from data

- need to find region where QCD dominates while staying as close as possible to signal selection
- select sample orthogonal to standard top selection by inverting cut on electron PID selection → **anti-electron sample**
- find distribution that is sensitive to lepton fakes → **missing transverse energy** (QCD here mostly instrumental background)
- shape of QCD background taken from data, but model provides no cross-section → **determine amount of QCD background from fit**



details on the fitting procedure

- binned Likelihood fit using TMinuit
- allow background contributions to vary from Standard Model expectation within expected uncertainties

Likelihood function

$$-\log \mathcal{L} = \sum_{i}^{\text{bins}} (\mu_i - N_i^{data} \log \mu_i) + \mathcal{G}$$

$$\mu_i = \sum_{k}^{\text{processes}} N_i^{\text{MC}} \cdot \beta_k$$

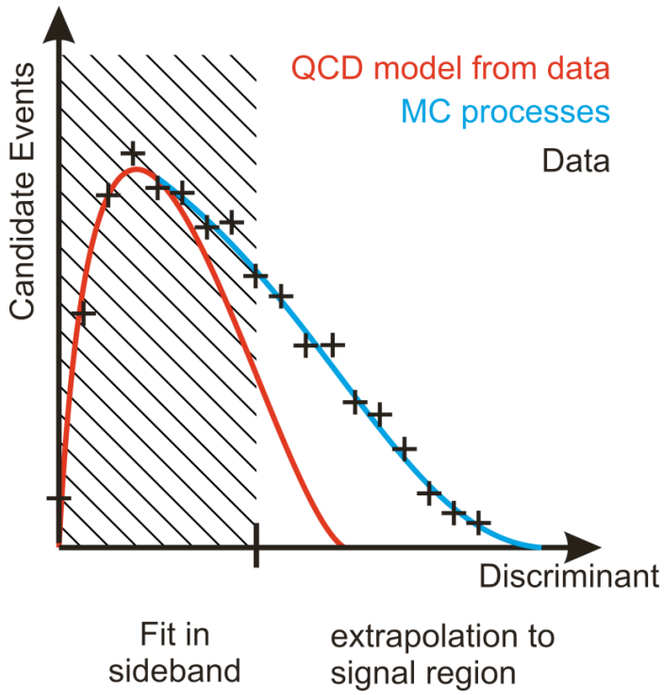
Gaussian constraint

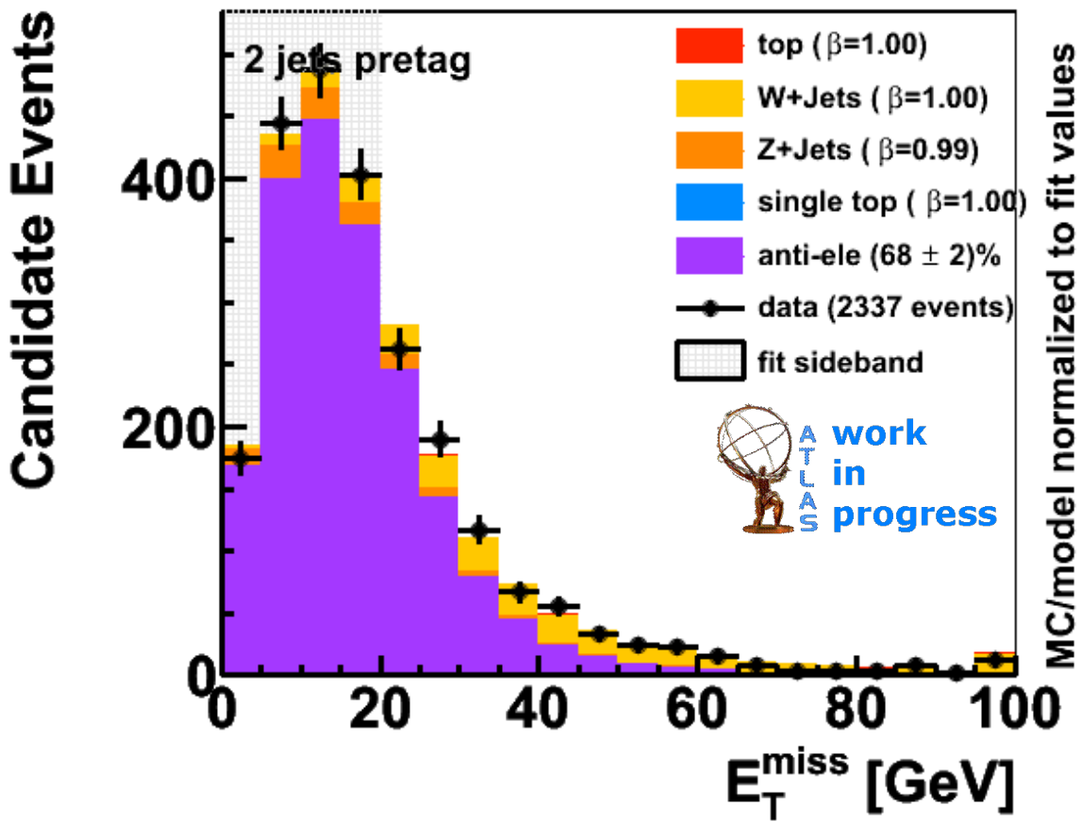
$$\mathcal{G} = -\frac{1}{2} \sum_{k}^{\text{processes}} \left(\frac{\beta_k - 1}{\sigma_k} \right)^2$$

$$\beta_k = \frac{N^{\text{fit}}}{N^{\text{expected}}}$$

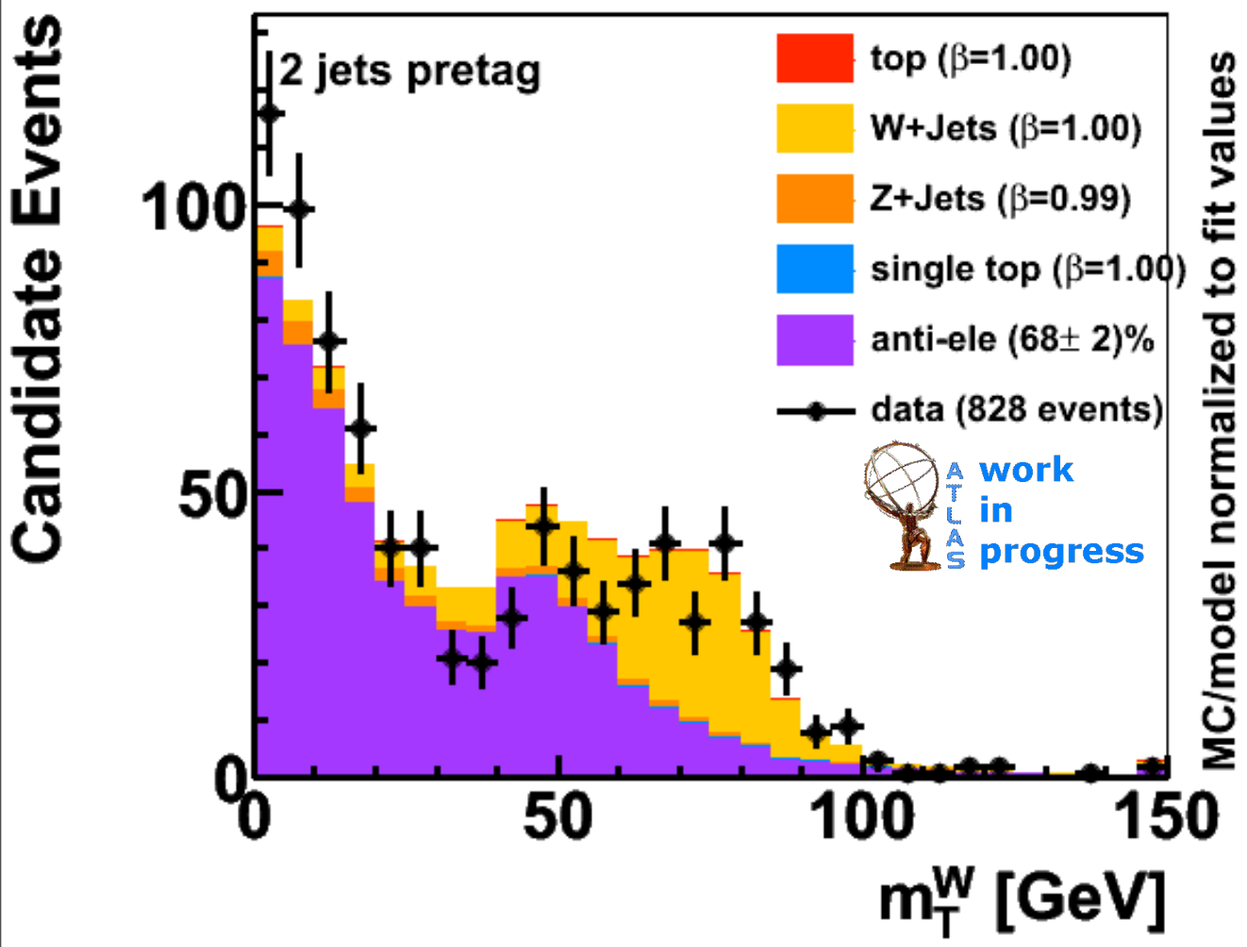
anti-electron model

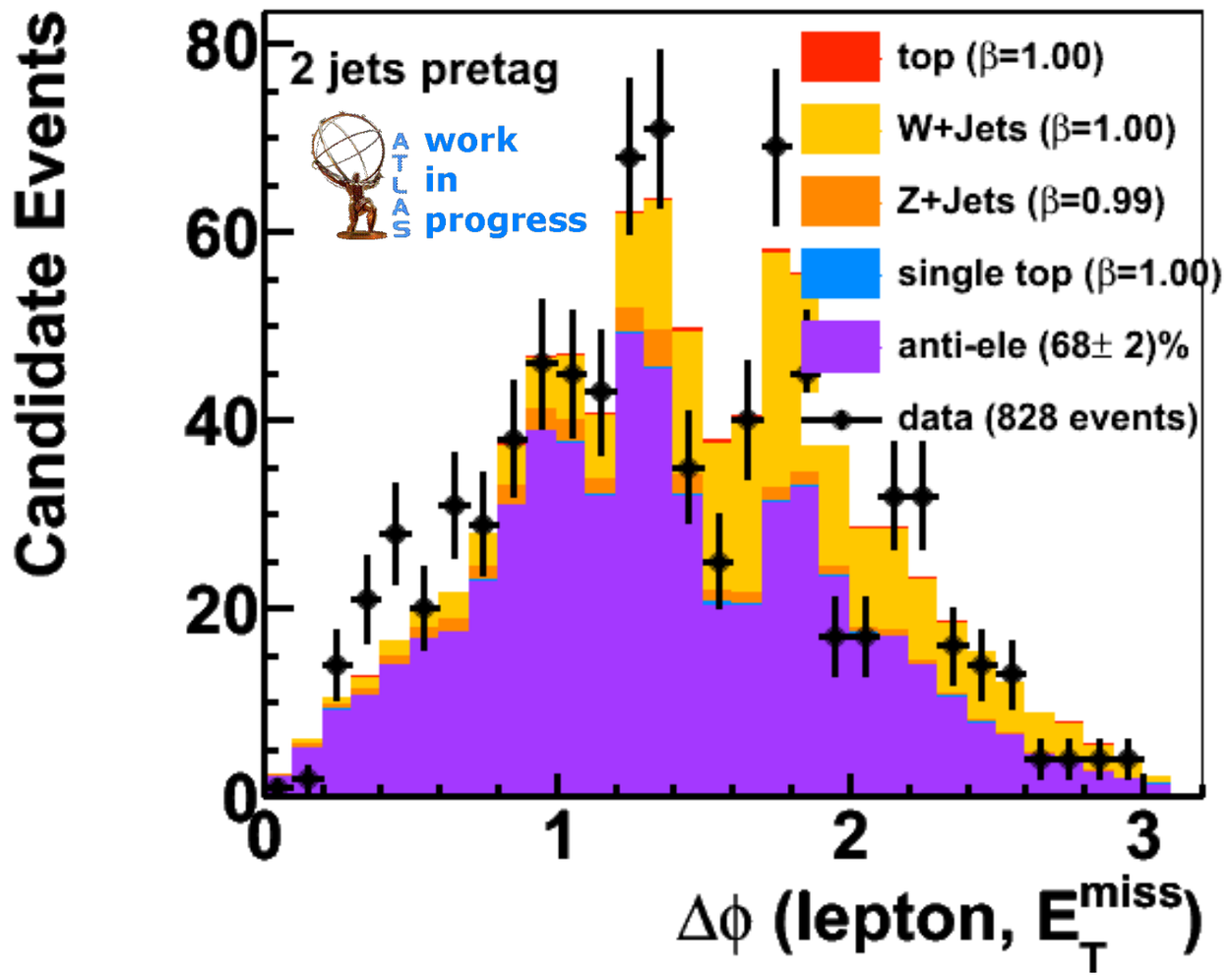
- if we can fit in sideband we can perform another fit (e.g. W+jets background) in the signal region later
- top selection applies MET cut of 20 GeV → **use 0-20 GeV sideband for fit and extrapolate**
- all processes taken from MC and scaled to SM expectation
- **challenge**: tune background model to agree with signal-selected data
- having a full background model should also allow us to look into other kinematic distributions





control plots

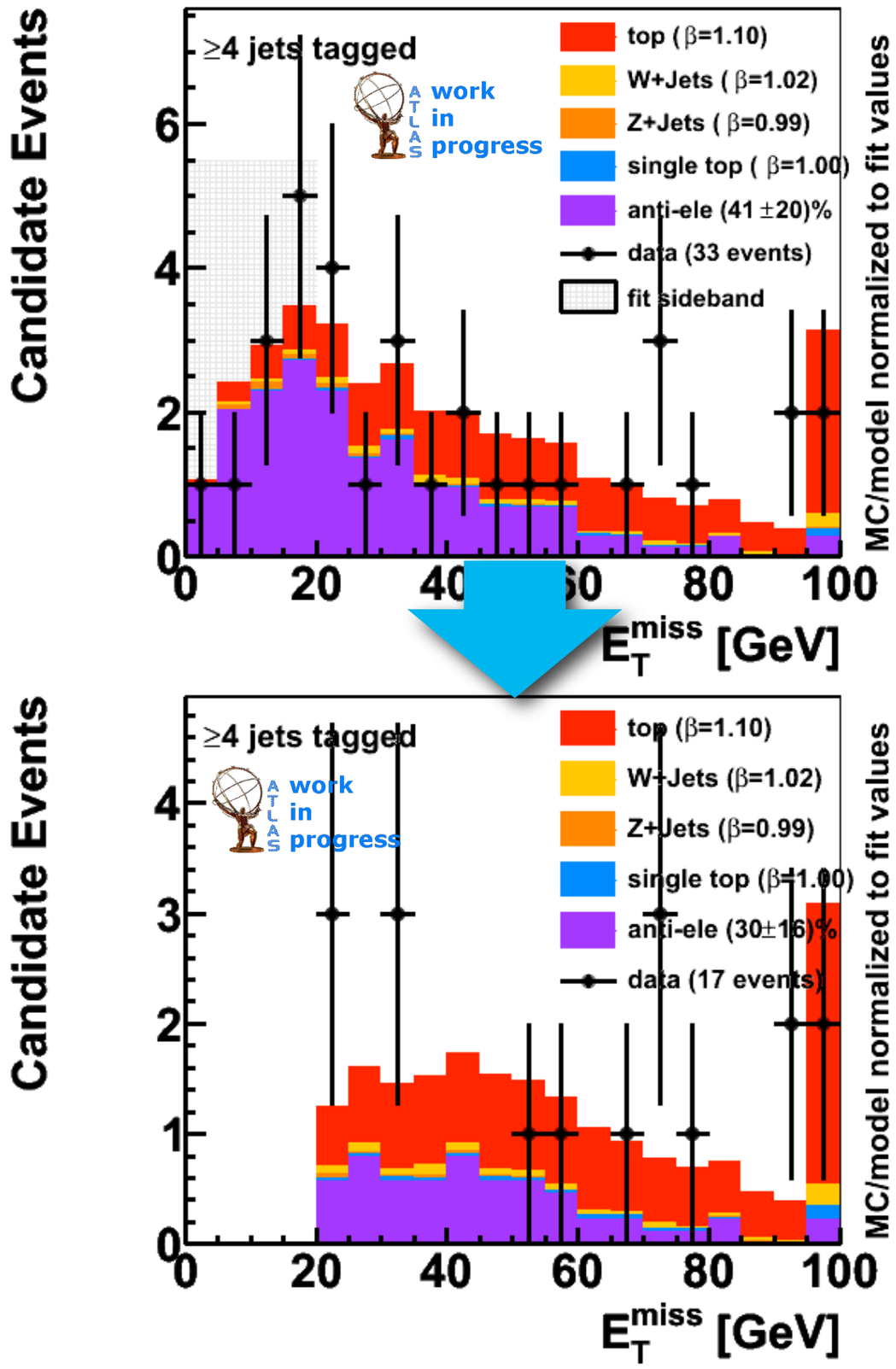




good agreement with data!

top signal region

- estimated QCD fraction very high in previous plots
- need to reduce QCD contribution to „see“ signal
- introduced additional QCD-killer cut: MET + W transv. mass > 60 GeV
- problem: kills sideband region
- solution: fit before cut is applied, extrapolate and then apply cut (for each process)
- QCD fraction reduced substantially (though very limited statistics here)



summary & outlook

- top pair cross-section measurement challenging
- anti-electron model allows to model QCD background from data
- need more statistics to fully validate model

- further backgrounds need to be evaluated
- cross-section measurement to be performed