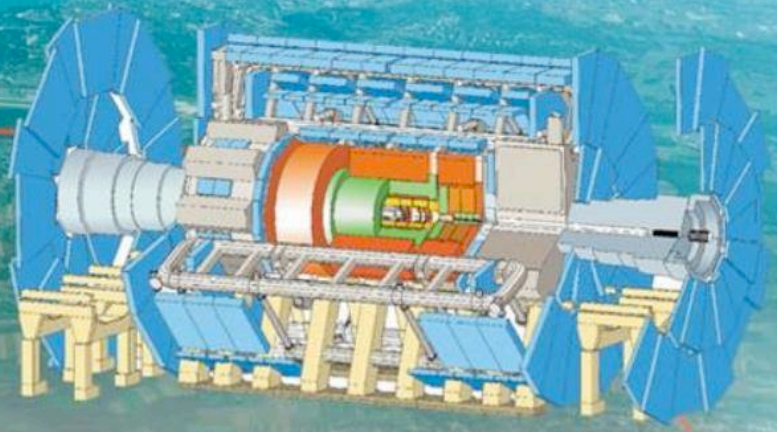
Background Shape Study for the $t\bar{t}H^0$, $H^0 \rightarrow b\bar{b}$ Channel



Catrin Bernius

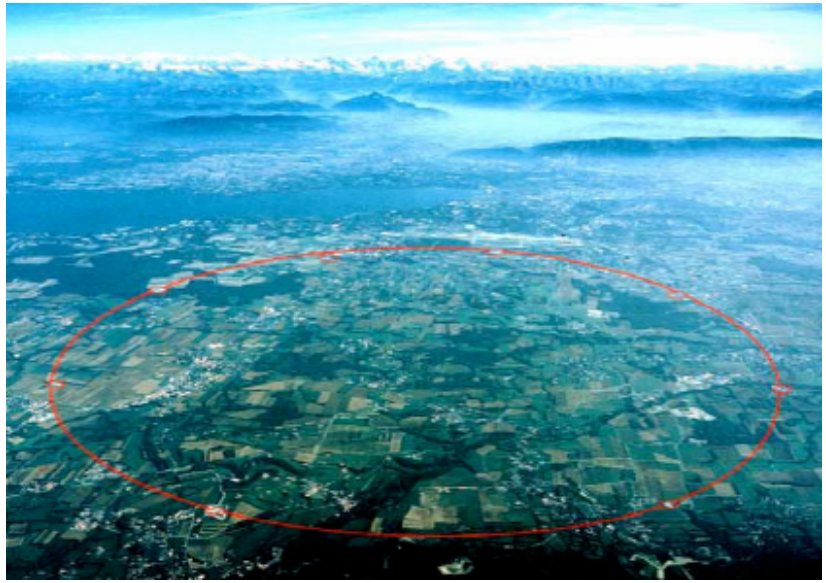
First year talk

15th June 2007

Outline

- LHC & ATLAS Detector
- $t\bar{t}H^0$ Channel and Backgrounds
- Motivation of Background Shape Study
- Background Shape Study
  - Tight Selection
  - Background Fit
  - Estimation of Signal Contribution
    - Sideband Fit
    - Combined Fit
- Summary & Outlook

LHC & ATLAS Detector

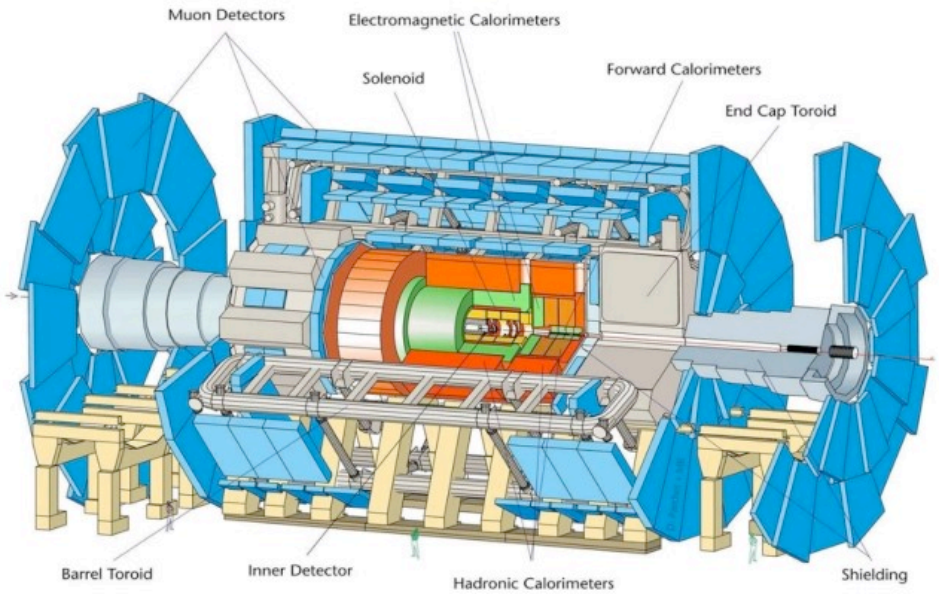


Large Hadron Collider (LHC):

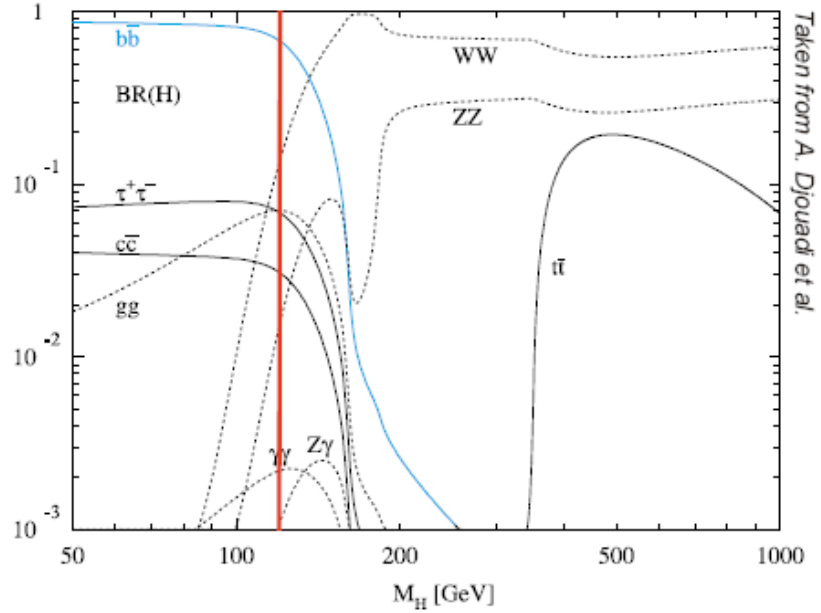
- 27 km circumference ring
- collide two counter-rotating 7TeV proton beams at four collision points
- design luminostiy: $10^{34}cm^{-2}s^{-1}$

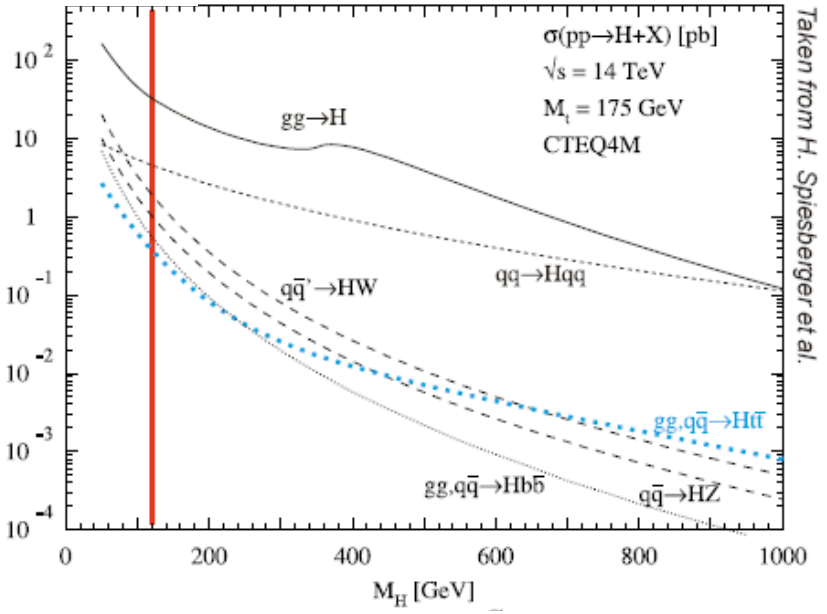
A Toroidial Lhc ApparatuS (ATLAS):

- width: 44m, diameter: 22m, 7000t
- cylindrical, multilayer structure
  - Inner Detector
  - Calorimeter (EM & HAD)
  - Muon Detectors

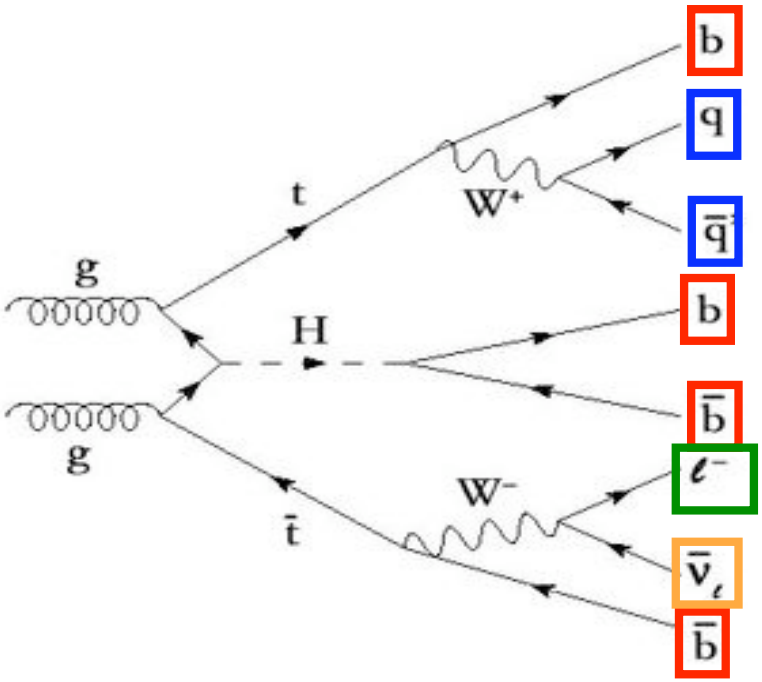


$t\bar{t}H$, $H^0 \rightarrow b\bar{b}$ Channel





- $H \rightarrow b\bar{b}$ main decay mode for m_H<135 GeV
- $H \rightarrow b\bar{b}$ almost impossible to detect in direct production $gg \rightarrow H$
- use lepton as trigger: WH, ZH, $\underline{t\bar{t}H}$
- final state: 4 b-jets, 2 light jets, 1 isolated lepton, Missing E_T

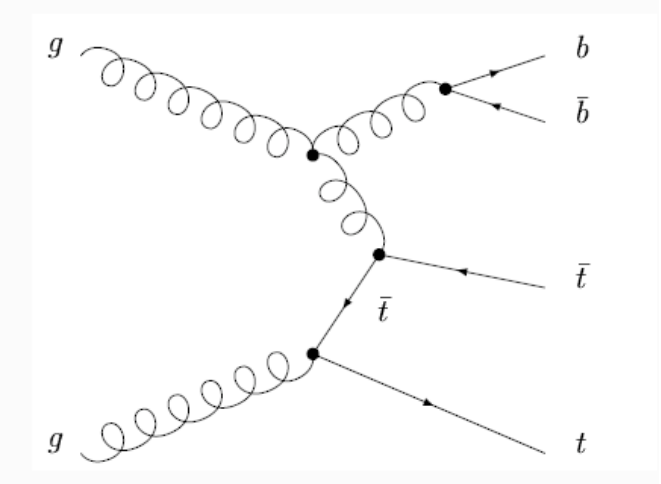


Backgrounds to $t\bar{t}H$, $H^0 \rightarrow b\bar{b}$

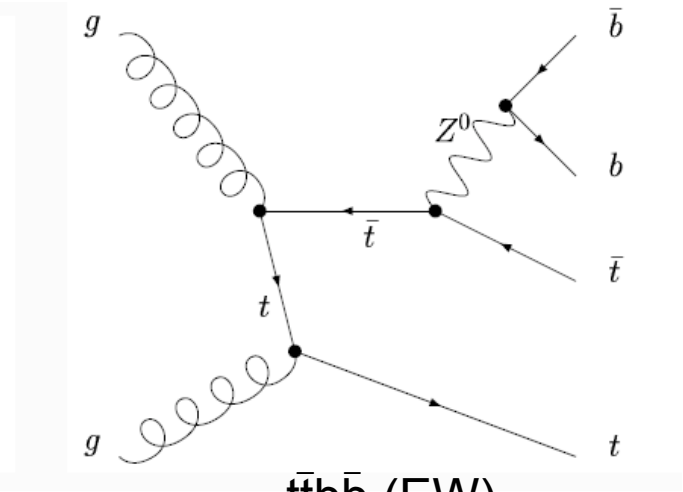
- reducible backgrounds: $t\bar{t}jj$
  - light jets misidentified as b-jets
  - suppression when requiring 4 b-jets

- irreducible background: $t\bar{t}b\bar{b}$ (QCD & EW)
  - same final states as signal

Process	σ (pb)
$t\bar{t}H$ (m_H=120GeV)	0.52
$t\bar{t}jj$	474
$gg \rightarrow t\bar{t}b\bar{b}$	8.1
$qq \rightarrow t\bar{t}b\bar{b}$	0.5
$gg \rightarrow Z/g/W \rightarrow t\bar{t}b\bar{b}$	0.9



$t\bar{t}b\bar{b}$ (QCD)

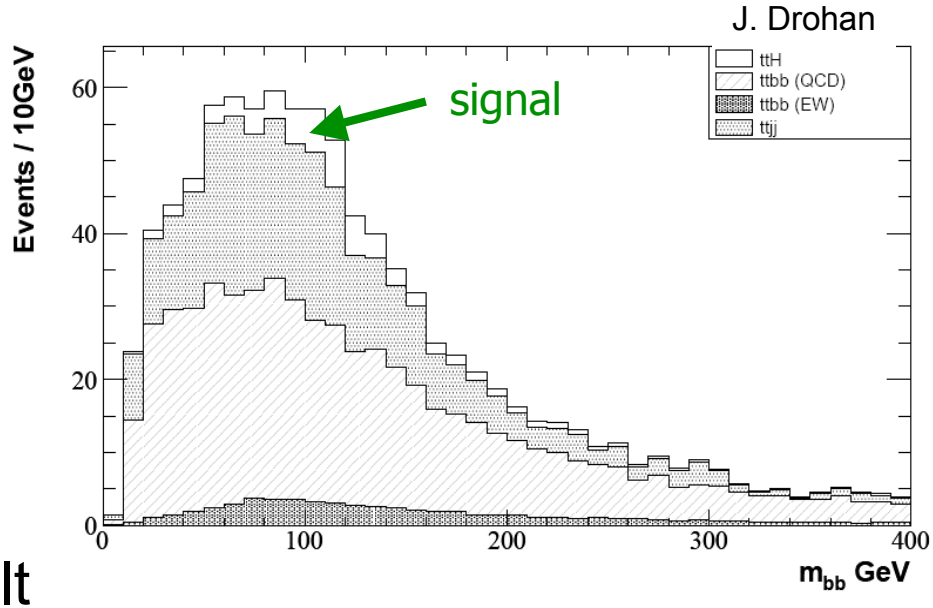


$t\bar{t}b\bar{b}$ (EW)

- combinatorial background: mis-pairing of b-jets in signal events

Motivation for Background Shape Study

- broad mass peak of signal events
- signal shape mimics $t\bar{t}b\bar{b}$ shape



- signal extraction very difficult
- need to know background shape well to constrain background contribution and estimate signal contribution

<u>study $t\bar{t}b\bar{b}$ background shape in invariant Higgs mass system</u>

Strategy

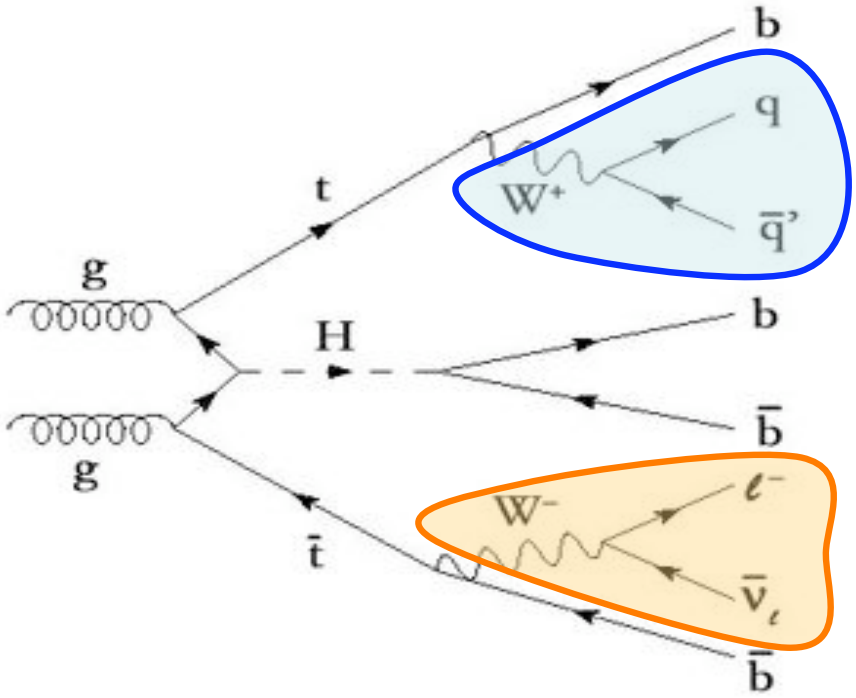
- more significant difference between shapes of signal and background in reconstructed m_{bb} spectrum
  - → "tight selection"
- describe background shape
  - → Fit
- possible approaches for signal estimation and extraction
  - → sideband fit
  - → combined fit

Idea: apply the developed methods in data, when LHC starts taking data

Signal Reconstruction (Standard Selection)

Signal reconstruction is done as in TDR & J. Drohan:

- select events with > 1 isolated lepton (electron, muon)
- require >2 jets and >4 b-tagged jets
  - p_t>15GeV, $|\eta|$<5
- Reconstruction from final state particles

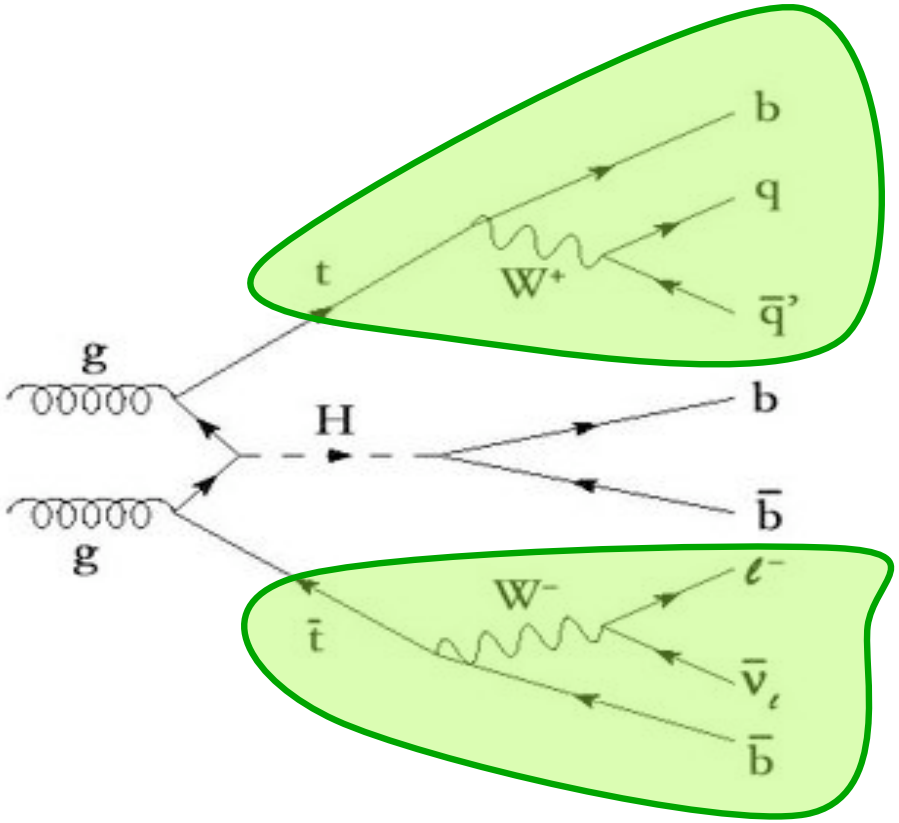


- $W \rightarrow l\nu$ and $W \rightarrow jj$ candidates

Signal Reconstruction (Standard Selection)

Signal reconstruction is done as in TDR & J. Drohan:

- select events with > 1 isolated lepton (electron, muon)
- require >2 jets and >4 b-tagged jets
  - p_t>15GeV, $|\eta|$<5
- Reconstruction from final state particles



- $W \rightarrow l\nu$ and $W \rightarrow jj$ candidates
- simultaneous reconstruction of two tops by minimizing

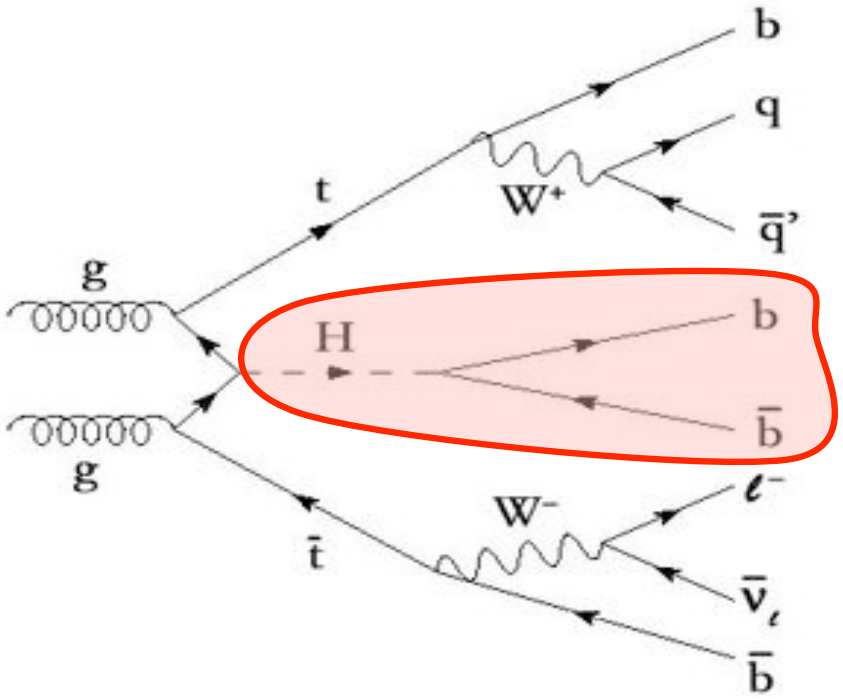
$$\Delta^2 = (m_{l\nu b} - m_t)^2 + (m_{jjb} - m_t)^2$$

- accept only solutions where $|m_t^{reco} - m_t| < 20\text{GeV}$

Signal Reconstruction (Standard Selection)

Signal reconstruction is done as in TDR & J. Drohan:

- select events with > 1 isolated lepton (electron, muon)
- require >2 jets and >4 b-tagged jets
  - p_t>15GeV, $|\eta|$<5
- Reconstruction from final state particles



- $W \rightarrow l\nu$ and $W \rightarrow jj$ candidates
- simultaneous reconstruction of two tops by minimizing

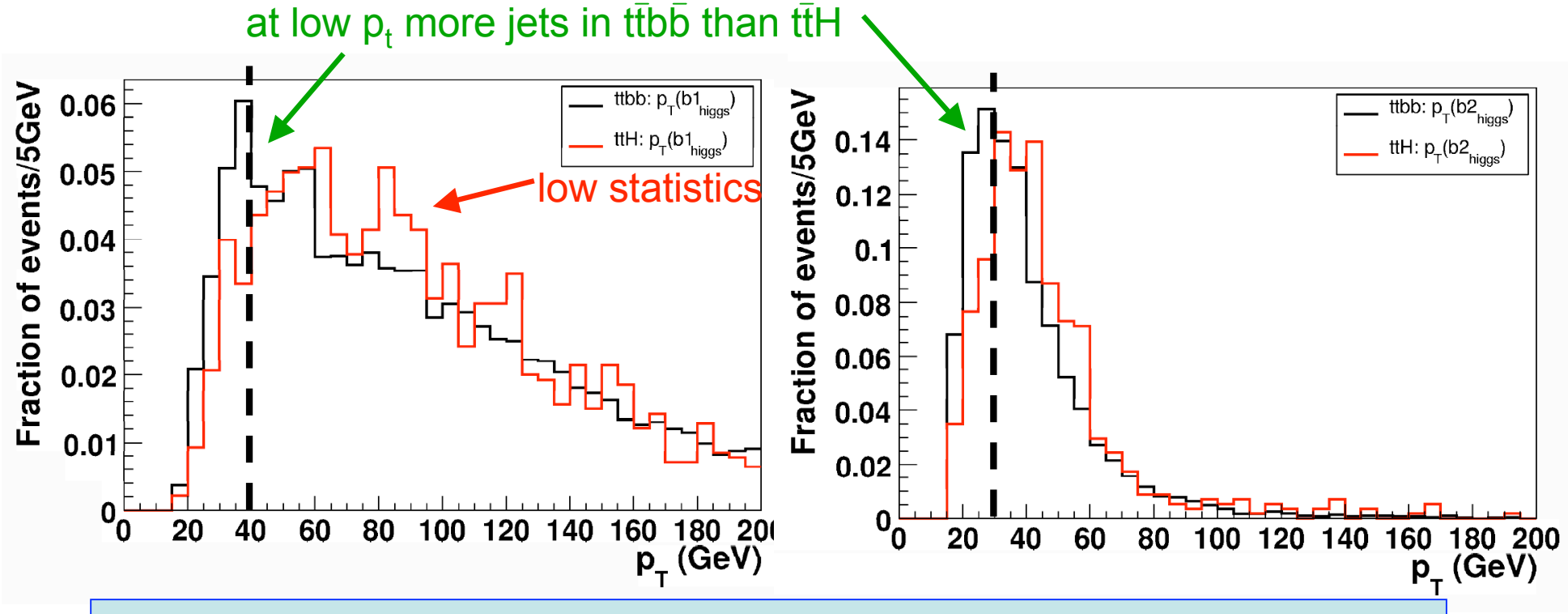
$$\Delta^2 = (m_{l\nu b} - m_t)^2 + (m_{jjb} - m_t)^2$$

- accept only solutions where $|m_t^{reco} - m_t| < 20\text{GeV}$
- with remaining two b jets reconstruct Higgs decay
- only accept events where $|m_H^{reco} - m_H| < 30\text{GeV}$ (m_H=120GeV)

Tight Selection

- Intention: more distinctive difference between background and signal shapes
- several kinematic cuts of analysis were varied
- p_t distribution of the two b-jets from reconstructed Higgs candidates

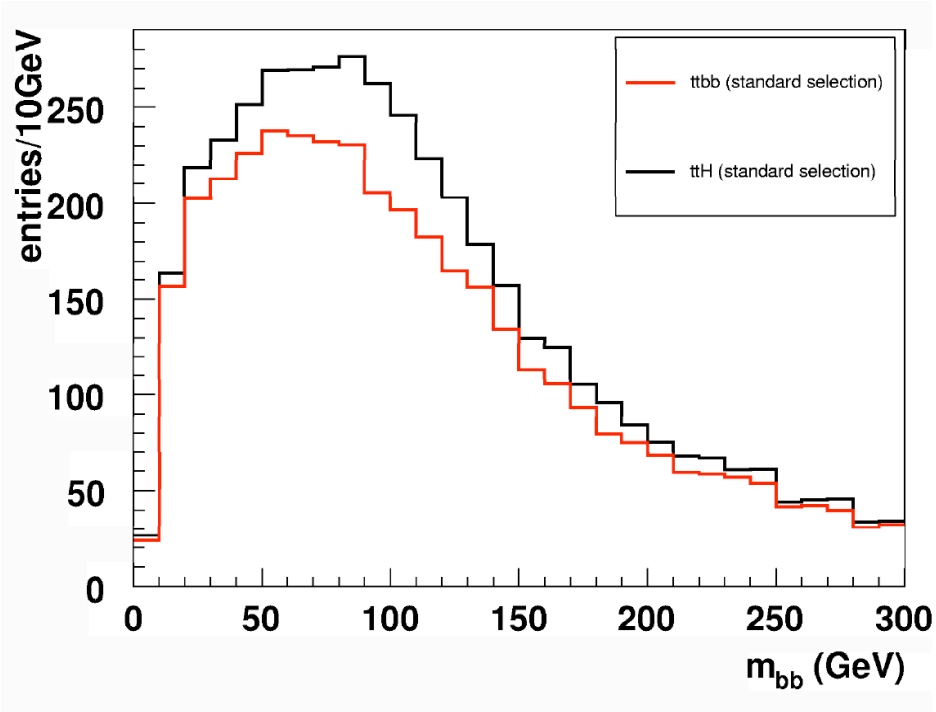
1M $t\bar{t}b\bar{b}$ (AcerMC), 38950 $t\bar{t}H$ (Pythia)



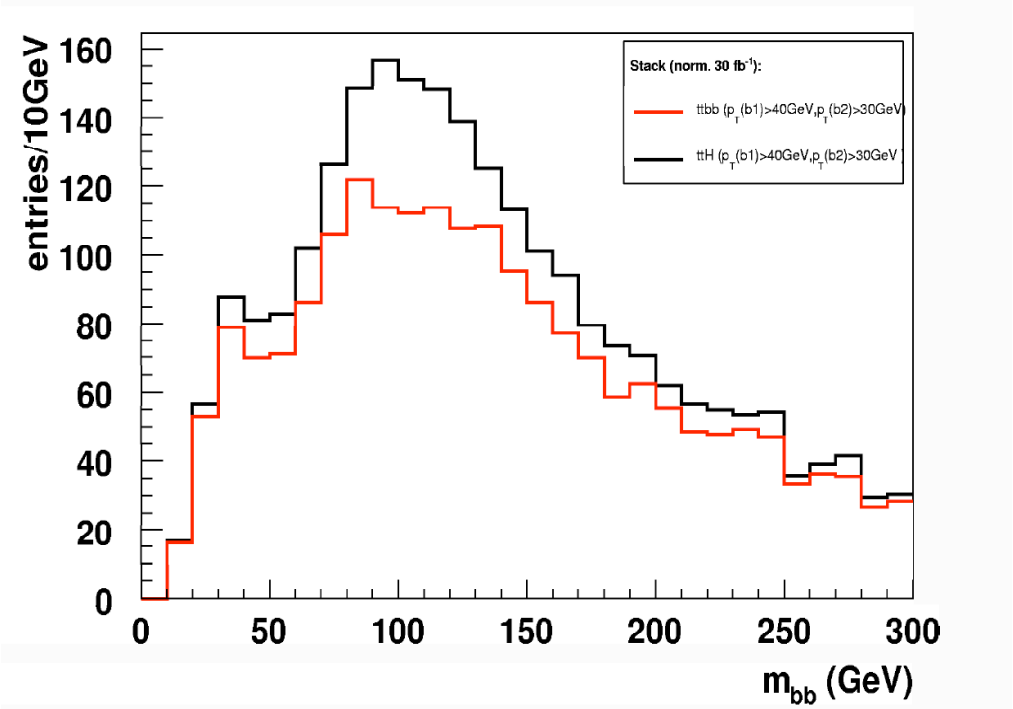
Tight selection: $p_t(b^1_{Higgs})$> 40 GeV, $p_t(b^2_{Higgs})$> 30 GeV

Shape Comparison

Standard Selection



Tight Selection



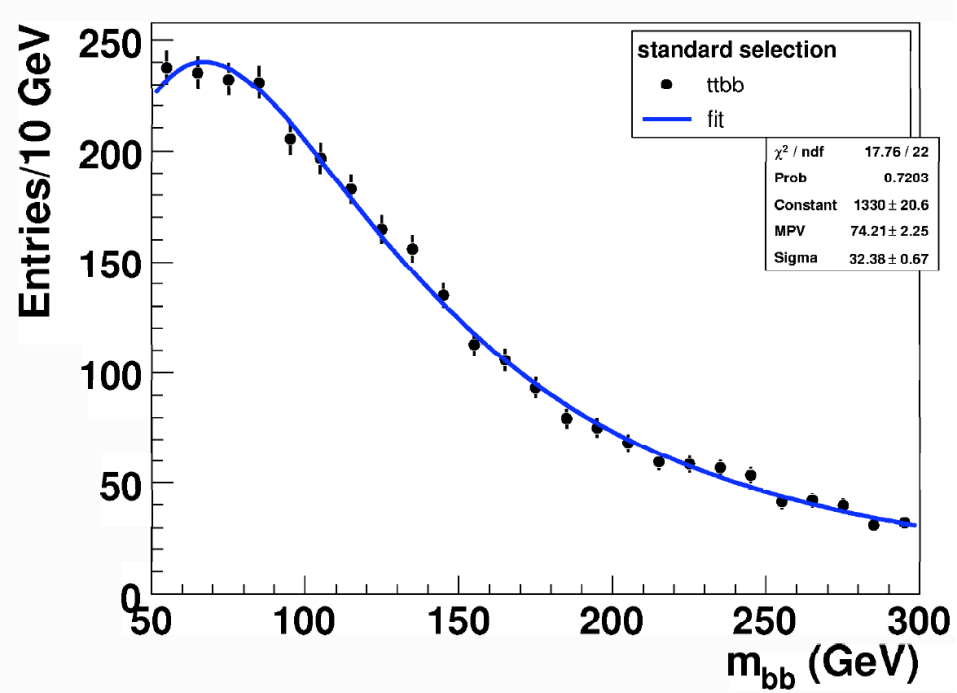
- number of selected signal & background events normalized to 30 fb^{-1}(“combined sample“)
- there is a significant change in shape:
  - more distinct peak in tight selection
  - greater difference in shape between ttbb & ttH events in tight selection

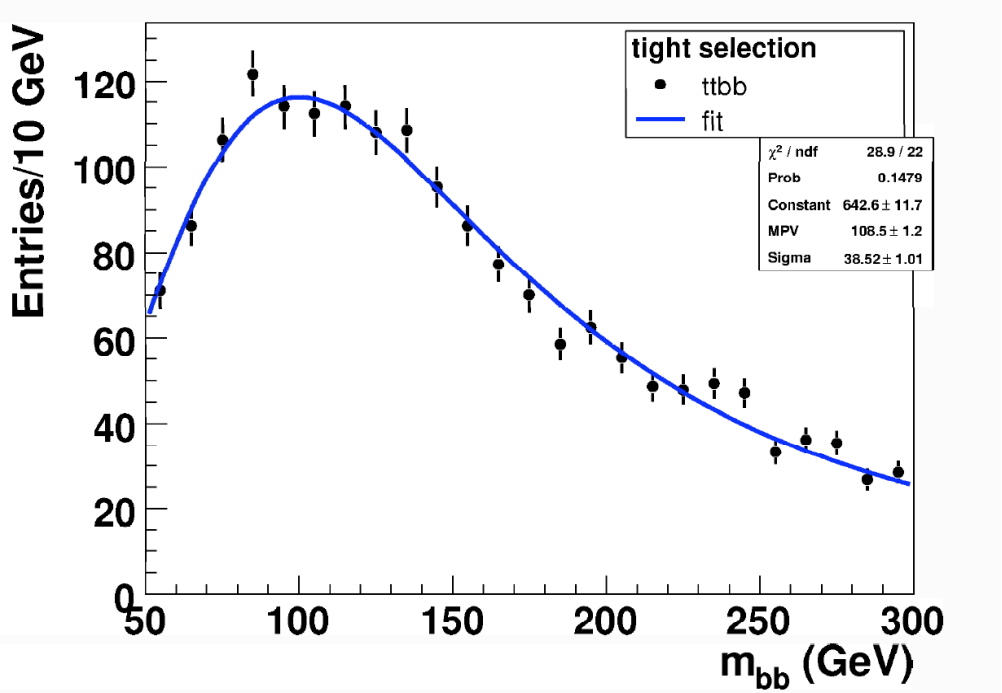
Fitting the Background Shape

- Description of background shape
- Fit with Landau distribution:

$$\text{Landau} = c\frac{1}{x\sqrt{2\pi s^2}}e^{-(\ln x - m)^2/(2s^2)}$$

m = most probable value
s = sigma
c = constant

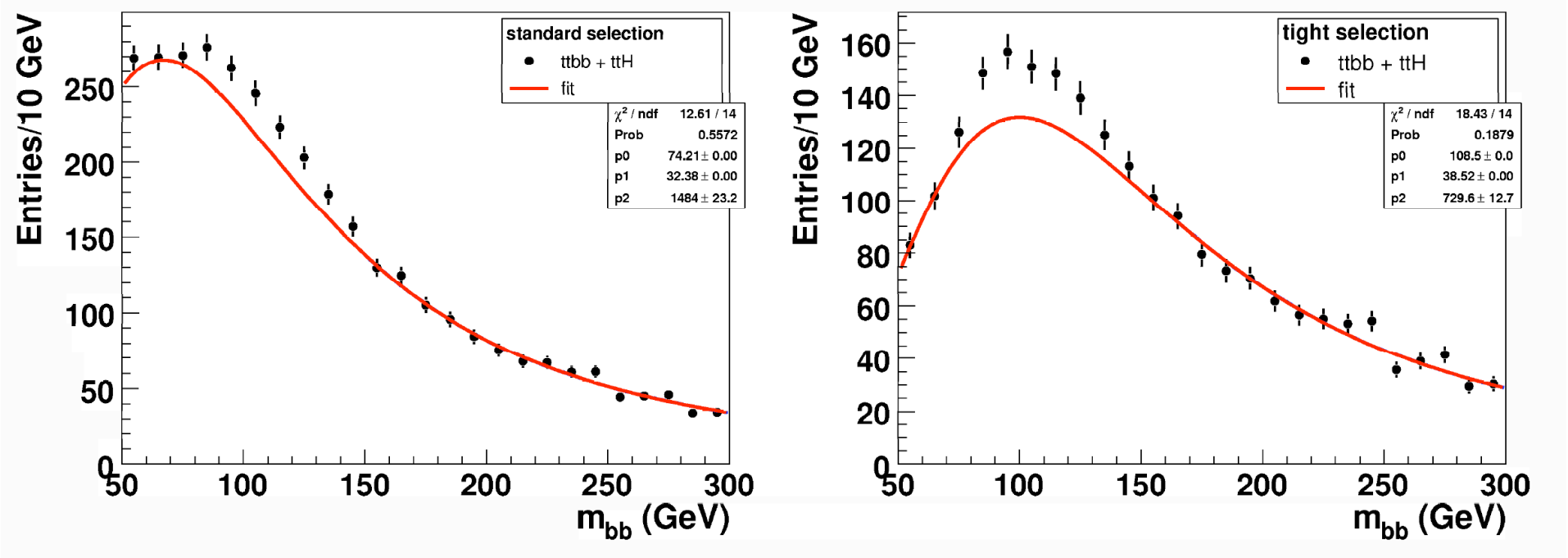




selection	m	s	c	χ^2/n.d.f.
standard	74.21±2.25	32.38±0.67	1330±20.6	17.76/22
tight	108.5±1.2	38.52±1.01	642.6±11.7	28.9/22

Sideband Fit

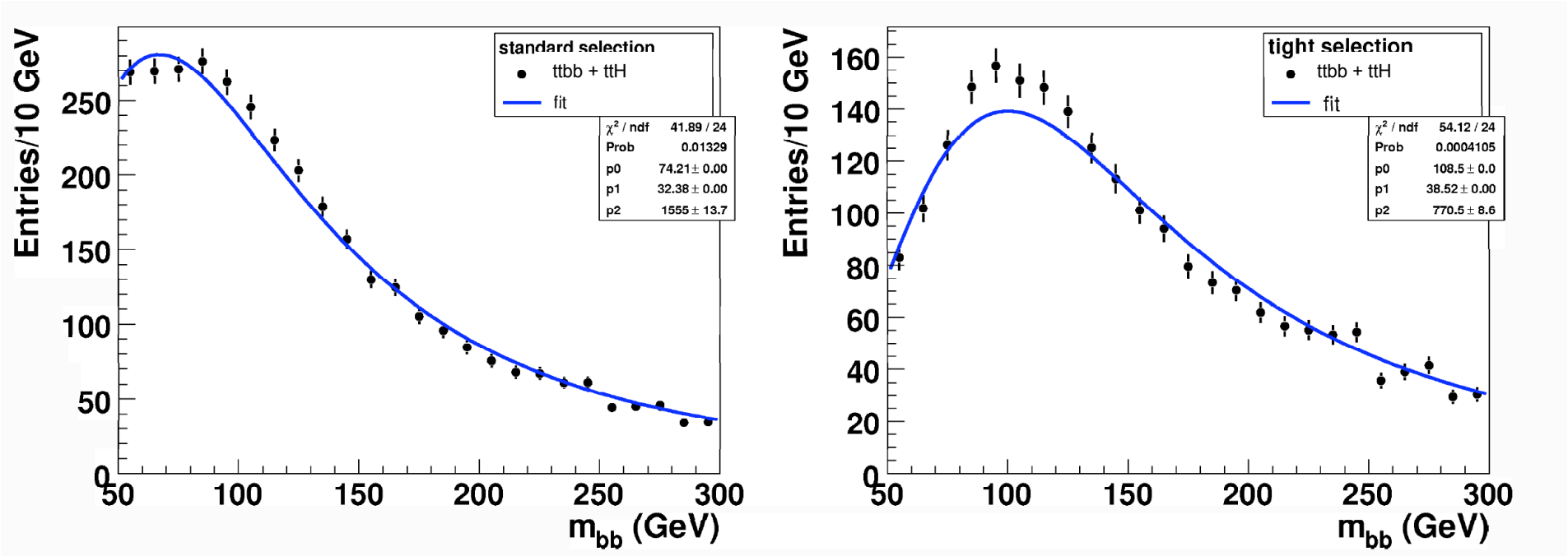
- Assume: no signal contribution in combined sample in sideband region 150GeV<m_{bb}<300GeV
- Plan: fit landau distribution in sideband region, use parameters from background fit to fix the mpv and sigma, leave only normalization free, draw function over whole range
- more distinct contribution of non background events in tight selection



- But careful: in sideband also small signal contribution
  → underestimating signal events

Combined Fit (1)

- Fit Landau now over whole range with fixed mpv and sigma from background fit



selection	χ^2/n.d.f.
standard	41.89/24
tight	54.12/24

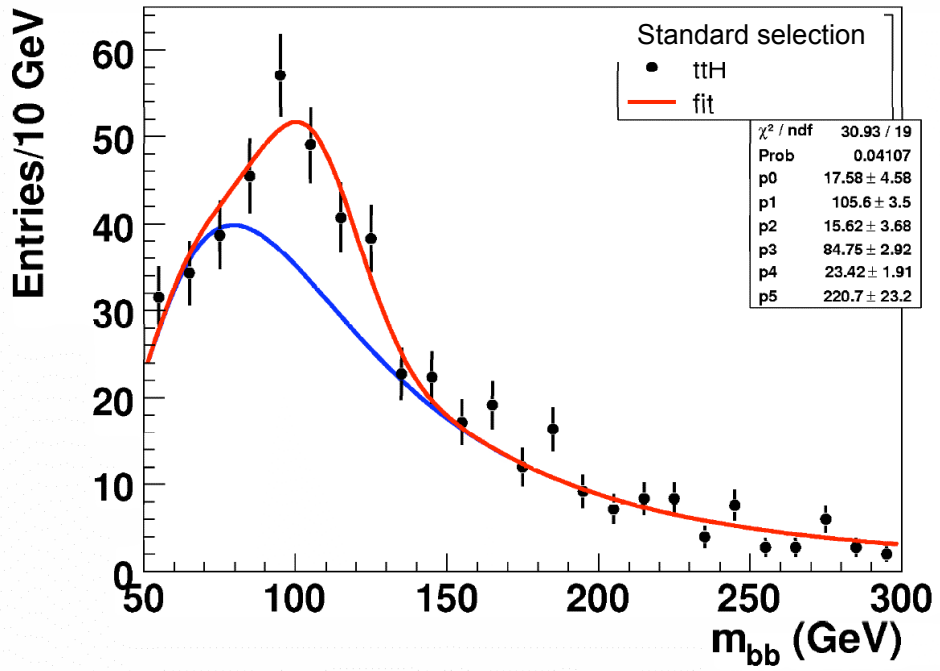
- with tight selection: worse fit result
- greater disagreement of shapes of background only and combined sample compared to standard selection

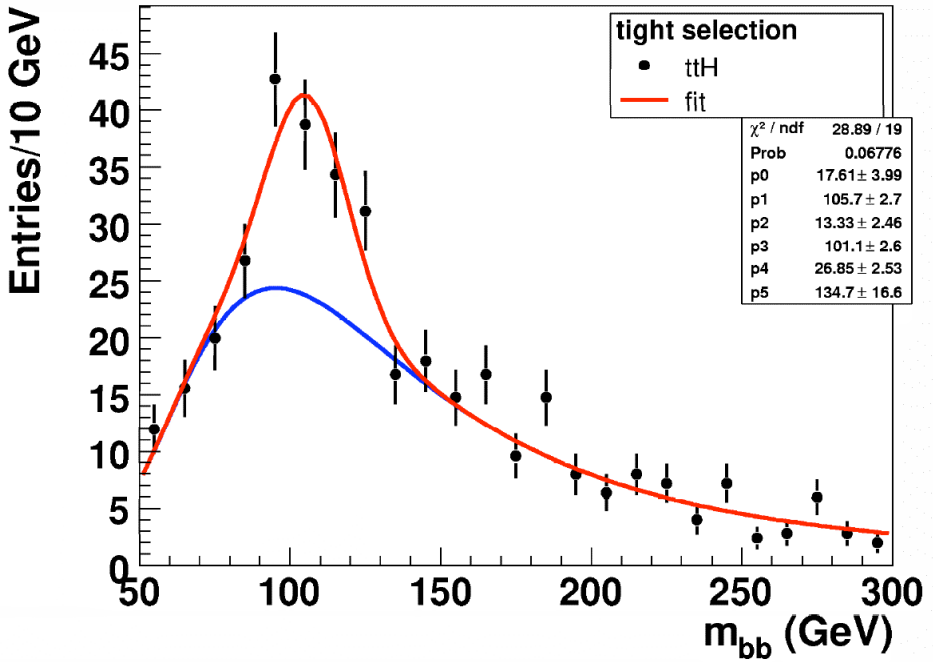
Combined Fit (2)

- Consider now also signal events
- Fit signal distribution with Gaussian and Landau distribution

$$\text{Gauss} = \frac{1}{\sigma\sqrt{2\pi}} e^{\cdot (x-\mu)^2/(2\sigma)}$$

$$\text{Landau} = \frac{1}{x\sqrt{2\pi s^2}} e^{-(\ln x - m)^2/(2s^2)}$$



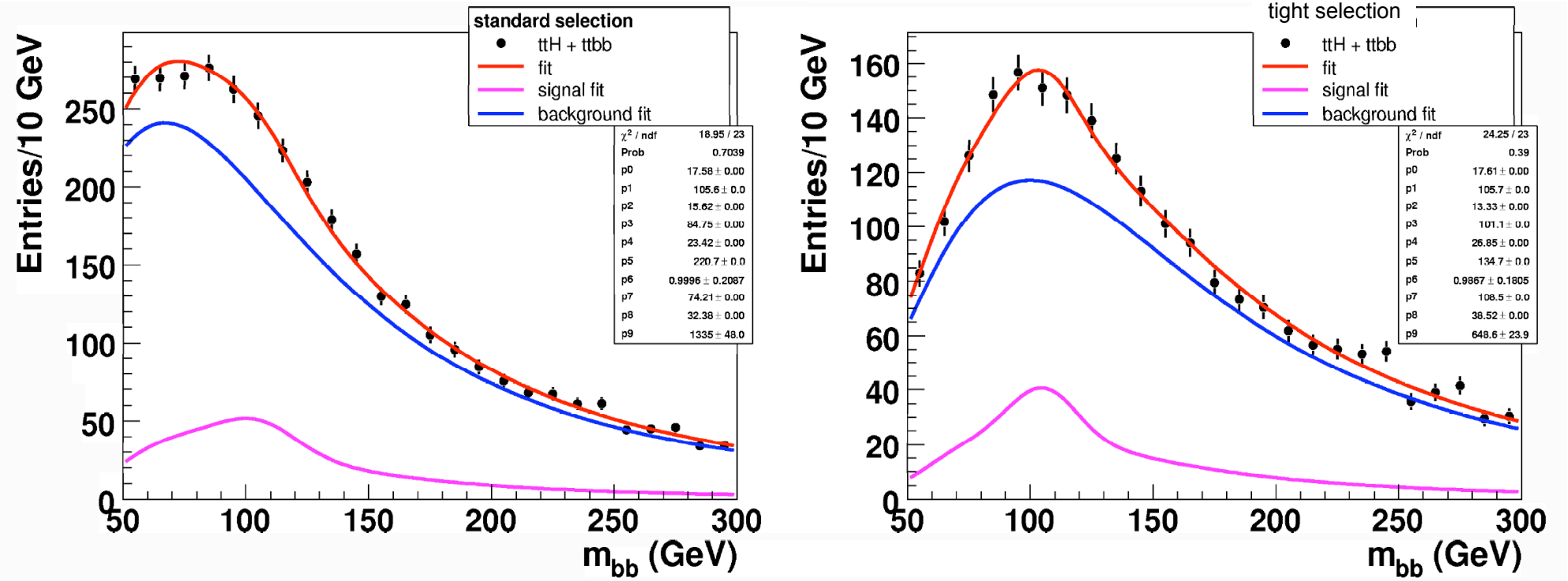


selection	μ	σ	n_{gauss}	m	s	n_{Landau}	χ^2/n.d.f.
standard	105.6±3.5	15.6±3.7	17.6±4.6	84.8±2.9	23.4±1.9	220.7±23.2	30.9/19
tight	105.7⊥2.7	13.3⊥2.5	17.6⊥3.9	101.1⊥2.6	26.9⊥2.5	134.7⊥16.6	28.9/19

Combined Fit (3)

- Combine background and signal fit:

$$\text{combined fit} = a \times \text{signal fit} + b \times \text{background fit}$$



- Evaluate goodness of fit with Kolmogorov-Smirnov test:
  - sensitive to differences in shapes
  - takes max. distance between cumulative distributions of sample and fit function

selection	fit applied	P(d)	P(χ^2,n.d.f.)
standard	background	0.3972	0.0133
standard	combined	1	0.7038
tight	background	0.1487	0.0004
tight	combined	0.9998	0.3899

Summary & Outlook

- Good starting point in order to apply developed techniques in data

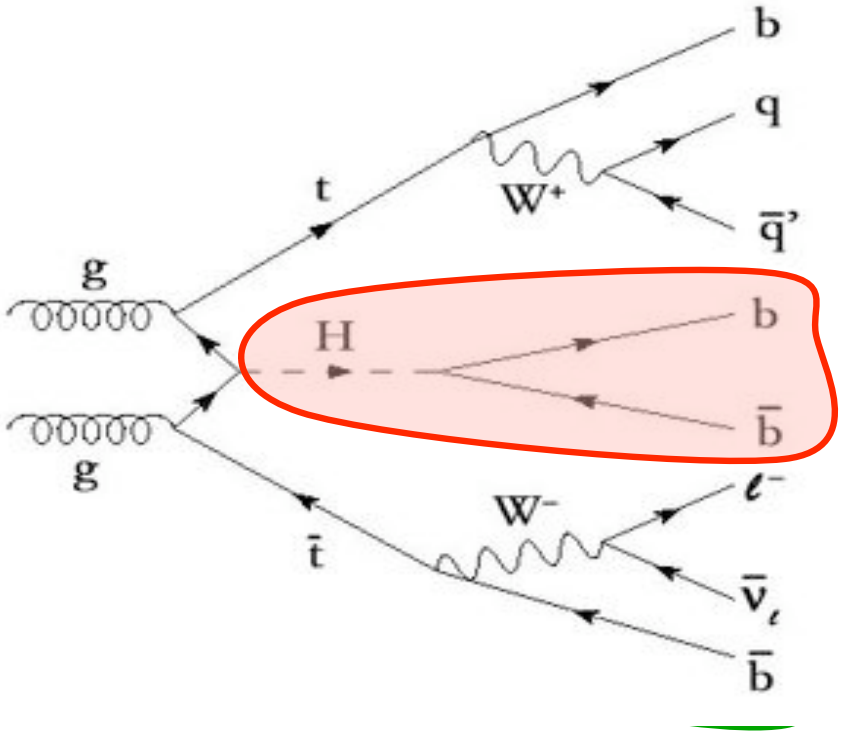
- More to do:
  - Include other backgrounds (esp. ttjj)
  - Further development of sideband/combined fit method and of tight selection

- Level 2 trigger work
- Continue work with ATLAS distributed analysis
  - Basically try to make jobs work on Grid
  - TWiki page: https://www.hep.ucl.ac.uk/twiki/bin/view/Main/AtlasGanga

Signal Reconstruction (Standard Selection)

Signal reconstruction is done as in TDR & J. Drohan:

- select events with > 1 isolated lepton (electron, muon)
- require >2 jets and >4 b-tagged jets
  - p_t>15GeV, $|\eta|$<5
- Reconstruction from final state particles



- $W \rightarrow l\nu$ and $W \rightarrow jj$ candidates
- simultaneous reconstruction of two tops by minimizing

$$\Delta^2 = (m_{l\nu b} - m_t)^2 + (m_{jjb} - m_t)^2$$

- accept only solutions where $|m_t^{reco} - m_t| < 20GeV$
- with remaining two b jets reconstruct Higgs decay
- only accept events where $|m_H^{reco} - m_H| < 30GeV$ (m_H=120GeV)