Search for the SM Higgs boson in full-hadronic di-tau events in the ATLAS experiment

Daniele Zanzi, Sandra Kortner

MPP ATLAS-MDT

ATLAS

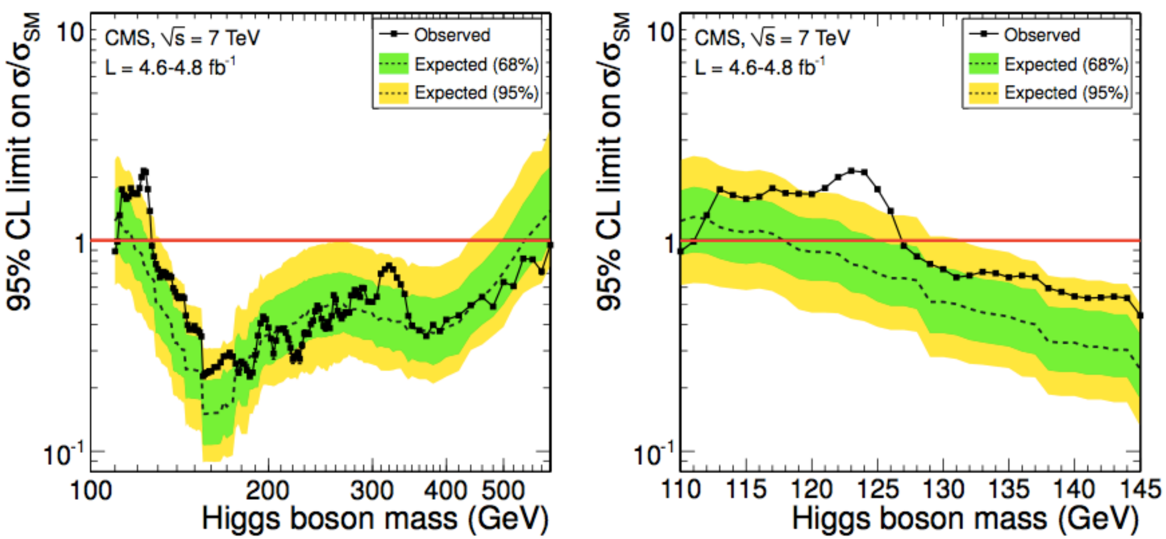
Max-Planck-Institut für Physik
(Werner-Heisenberg-Institut)

Outline

1. Status of the SM Higgs searches
2. Motivation for the search in the di-tau channel
3. The full-hadronic final state
4. Hadronic taus
5. Strategy of the search
6. Event selection
7. Background Estimation
8. Expected result
9. Summary

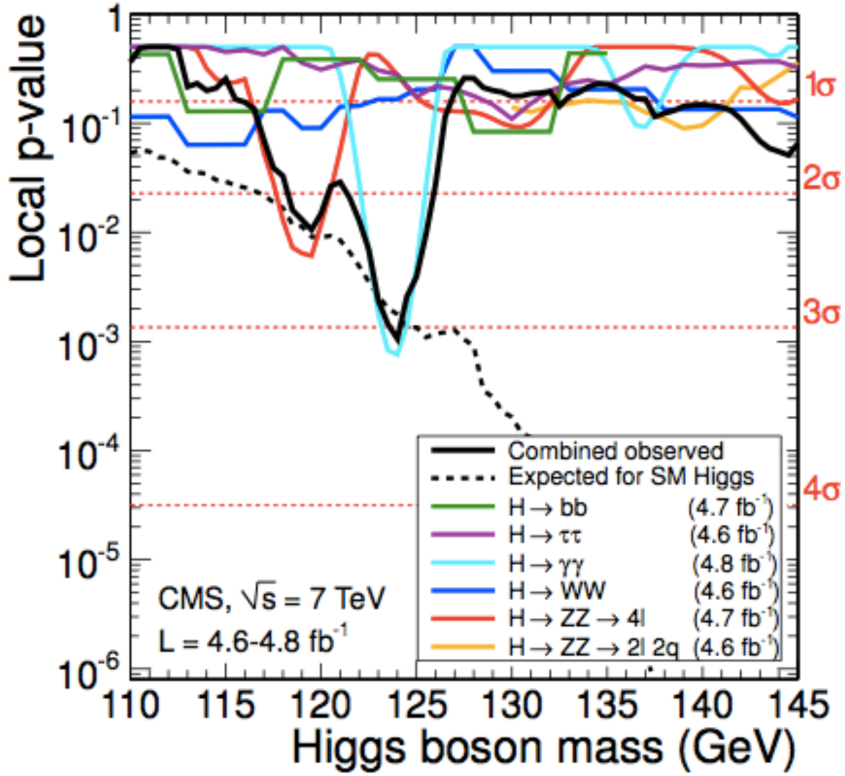
SM Higgs Searches

1. On Wednesday ATLAS and CMS submitted the results of the searches for the SM Higgs based on the data collected in 2011 (~5 fb^{-1})
2. The mass range around 125 GeV is getting harder and harder to exclude and hints of excess wrt SM prediction are coming out



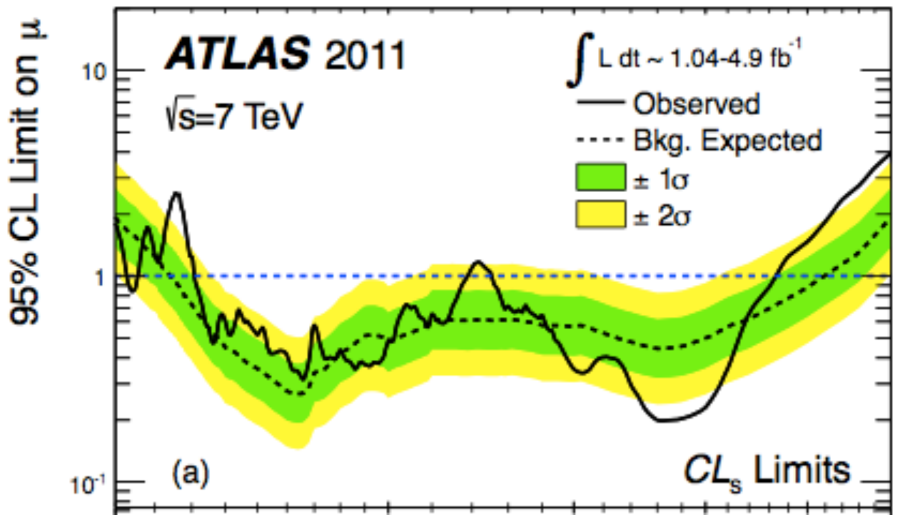
CMS Results

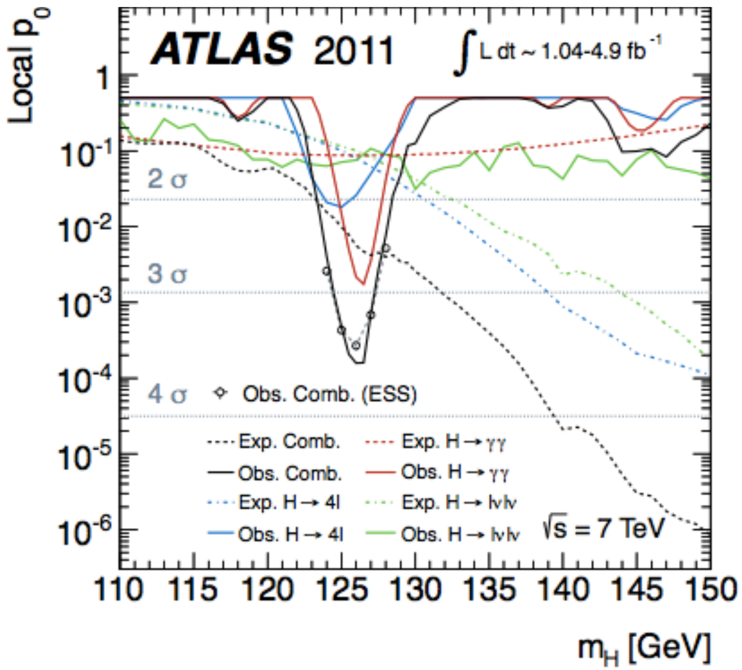
Combined results are reported from searches for the standard model Higgs boson in proton-proton collisions at $\sqrt{s} = 7$ TeV in five Higgs boson decay modes: $\gamma\gamma$, bb, $\tau\tau$, WW, and ZZ. The explored Higgs boson mass range is 110–600 GeV. The analysed data correspond to an integrated luminosity of 4.6–4.8 fb^{-1}. The expected excluded mass range in the absence of the standard model Higgs boson is 118–543 GeV at 95% CL. The observed results exclude the standard model Higgs boson in the mass range 127–600 GeV at 95% CL, and in the mass range 129–525 GeV at 99% CL. An excess of events above the expected standard model background is observed at the low end of the explored mass range making the observed limits weaker than expected in the absence of a signal. The largest excess, with a local significance of 3.1σ, is observed for a Higgs boson mass hypothesis of 124 GeV. The global significance of observing an excess with a local significance $\geq$3.1σ anywhere in the search range 110–600 (110–145) GeV is estimated to be 1.5σ (2.1σ). More data are required to ascertain the origin of this excess.



ATLAS Results

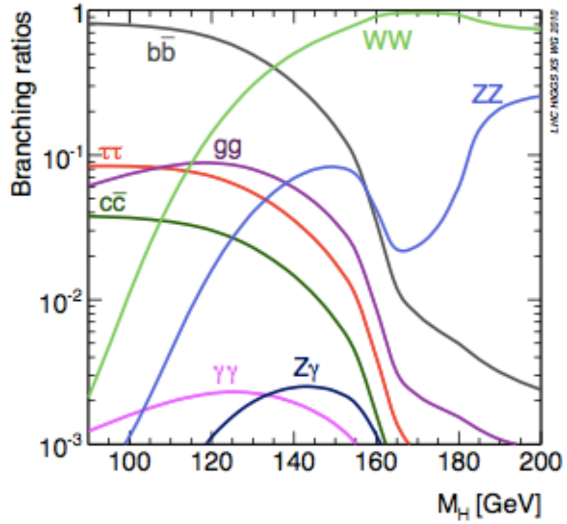
A combined search for the Standard Model Higgs boson with the ATLAS experiment at the LHC using datasets corresponding to integrated luminosities from 1.04 fb^{-1} to 4.9 fb^{-1} of pp collisions collected at $\sqrt{s}$ = 7 TeV is presented. The Higgs boson mass ranges 112.9–115.5 GeV, 131–238 GeV and 251–466 GeV are excluded at the 95% confidence level (CL), while the range 124–519 GeV is expected to be excluded in the absence of a signal. An excess of events is observed around m_H~126 GeV with a local significance of 3.5 standard deviations (σ). The local significance of $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ^{(*)} \rightarrow \ell^+\ell^-\ell'^+\ell'^-$ and $H \rightarrow WW^{(*)} \rightarrow \ell^+\nu\ell'^-\bar{\nu}$, the three most sensitive channels in this mass range, are 2.8σ, 2.1σ and 1.4σ, respectively. The global probability for the background to produce such a fluctuation anywhere in the explored Higgs boson mass range 110–600 GeV is estimated to be ~1.4% or, equivalently, 2.2σ.

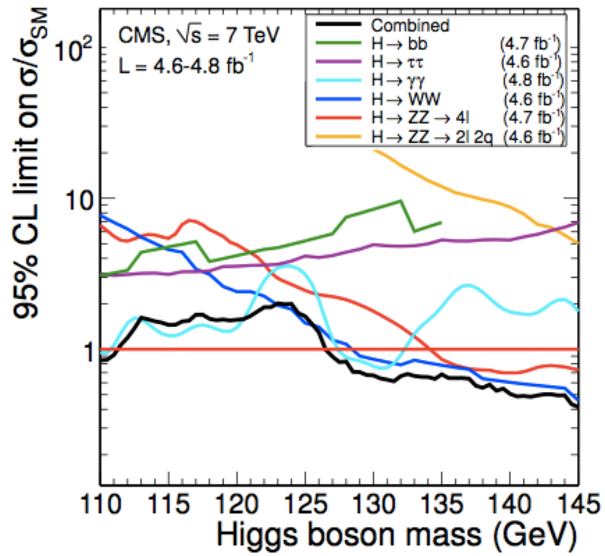




The H→$\tau\tau$ channel

1. At LHC the mass range around 120-130 GeV is the most difficult to explore because bb decay mode is almost blinded by the huge multijets background
2. The decay mode with highest sensitivity in the low mass range is H→$\gamma\gamma$, even though the BR is very little
3. In order to increase the sensitivity all channels have to be exploited and the $\tau\tau$ channel can provide a significant help





The $H \rightarrow \tau\tau$ channel (II)

There are 3 different final states:

1. full leptonic (12%)
    - clean events with two leptons
    - main backgrounds: two real leptons from $Z\tau\tau$, Zll, top, di-boson
    - little branching ratio
2. semi leptonic (46%)
    - high branching ratio
    - main backgrounds: one real lepton and 1 fake tau from Z+jets, W+jets, ttbar, QCD, …
3. full hadronic (42%)
    - high branching ratio
    - main backgrounds: events with 2 real taus like $Z\tau\tau$ or with two fake taus like multijets
    - identification efficiency of two hadronic taus

The full hadronic final state

1. The full hadronic final state is typically considered as ‘not-worth-to-try’ because of the huge multijets background
2. Therefore this final state has never been used for the SM Higgs search
3. We want to prove that this channel is not blinded by multijets events and that it can increase the combined sensitivity of the $H \rightarrow \tau\tau$ channel

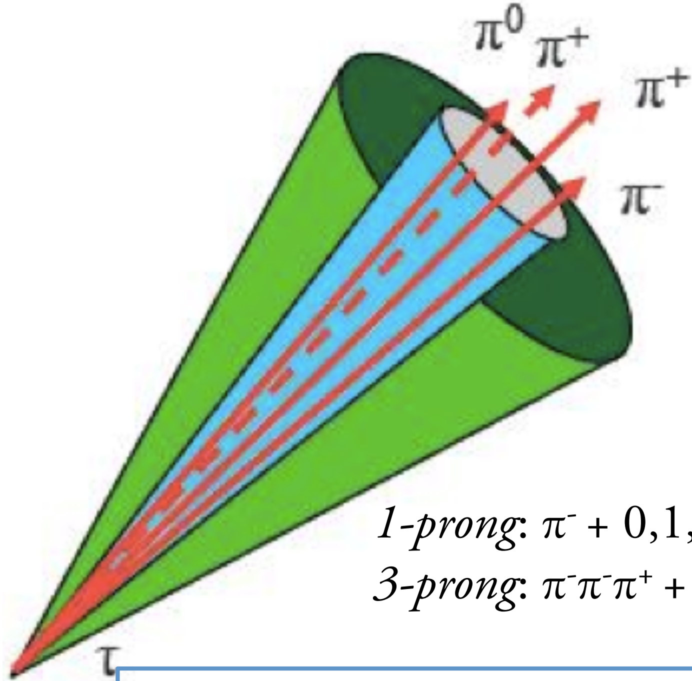
Why hadronic τ are difficult to deal with?

1. <u>Hadronic τ are not leptons, but jets!</u>

2. It's very difficult to distinguish between taus and jets

3. It's necessary to make a trade-off between identification efficiency and rejection power against jet

4. If we need <u>2 taus</u> this trade-off is doubled



1-prong: $\pi^- + 0,1,2\ \pi^0$
3-prong: $\pi^-\pi^-\pi^+ + 0,1\ \pi^0$

- Collimated and isolated jet with low track multiplicity
- Possible secondary vertex
- Energy deposit both in EM and Hadronic calorimeter

Strategy of the search

Our analysis is based on:

1. Collinear Mass Approximation for event selection and background rejection

2. Track Multiplicity for background estimation

Collinear Mass Approx

1. H→$\tau\tau$→hνhν
2. It's not possible to reconstruct the invariant mass of the $\tau\tau$ system, i.e. the mass of the Higgs candidate, due to the presence of 2 neutrinos
3. We can assume that the neutrinos are *collinear* to the visible product of the tau decay
4. We also assume that all the missing energy in the event is carried by the 2 neutrinos

$$p_{\tau,i} = p_{h,i} + p_{\nu,i}$$

$$p_{h,i} = (E_{h,i}, \vec{p_{h,i}})$$

$$p_{\nu,i} = \left(E_{\nu,i}, E_{\nu,i}\frac{\vec{p_{h,i}}}{|\vec{p_{h,i}}|}\right)$$

$$E_{\mathrm{x}}^{\mathrm{miss}} = E_{\nu,1,x} + E_{\nu,2,x} = E_{\nu,1}\frac{p_{h,1,x}}{|\vec{p_{h,1}}|} + E_{\nu,2}\frac{p_{h,2,x}}{|\vec{p_{h,2}}|},$$

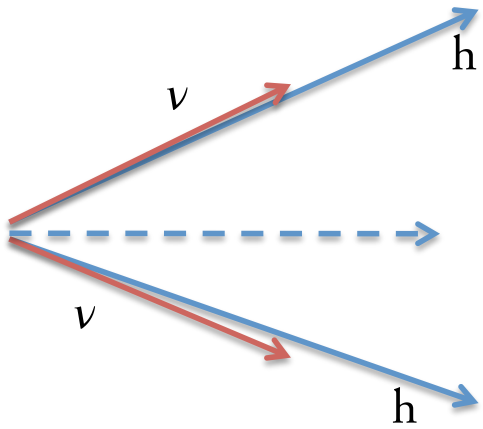
$$E_{\mathrm{y}}^{\mathrm{miss}} = E_{\nu,1,y} + E_{\nu,2,y} = E_{\nu,1}\frac{p_{h,1,y}}{|\vec{p_{h,1}}|} + E_{\nu,2}\frac{p_{h,2,y}}{|\vec{p_{h,2}}|}.$$

Collinear Mass Approx (II)

1. In order to get unique and physical solutions we need to require that:

$$\cos \Delta\phi(\tau_{vis,1}, \tau_{vis,2}) > -1$$

$$x_i = \frac{E_{h,i}}{E_{\tau,i}} = \frac{E_{h,i}}{E_{h,i} + E_{\nu,i}} \qquad 0 < x_1 < 1 \quad , \quad 0 < x_2 < 1$$



2. The selected events have two taus not back-to-back with the missing energy pointing in the middle of the two

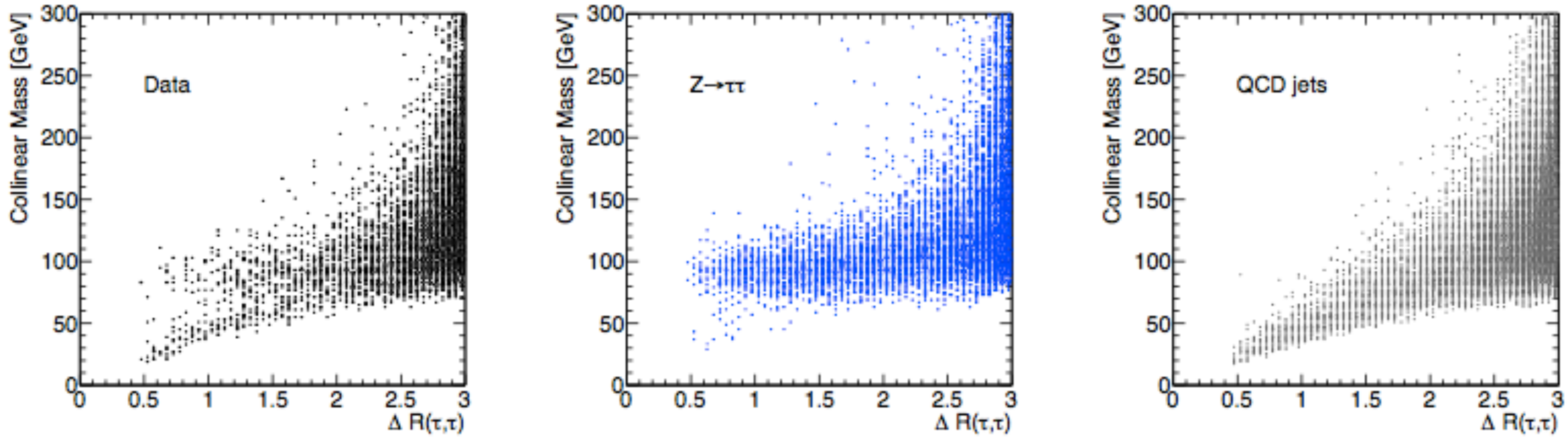
3. This approximation has a strong rejection power against not-di-tau events, like multijets, but the signal efficiency is little because the Higgs needs to be boosted or recoiling from a hard jet

4. This means that the analysis will be not inclusive, but a H+1jet search

$\Delta R(\tau,\tau)$

1. Instead of cutting in $\Delta\phi$, we selects events cutting on $\Delta R = \sqrt{(\Delta\phi^2 + \Delta\eta^2)}$

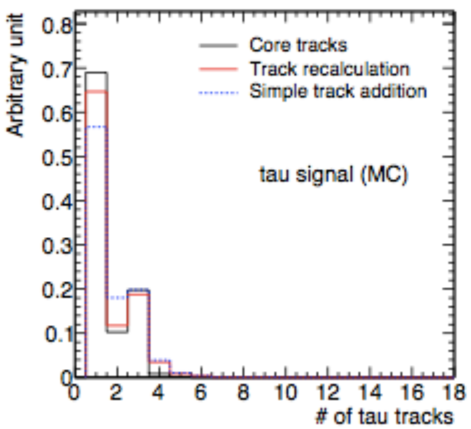
2. This increases the rejection against multijets and also reduces the tail of the Z mass shape

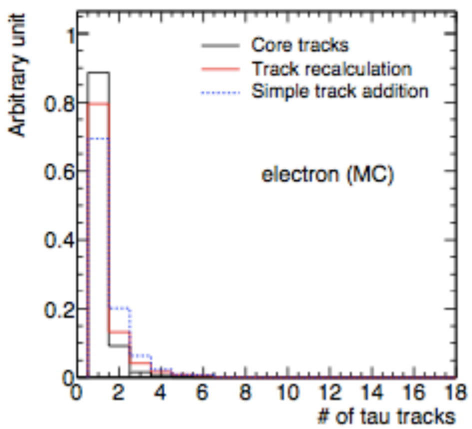


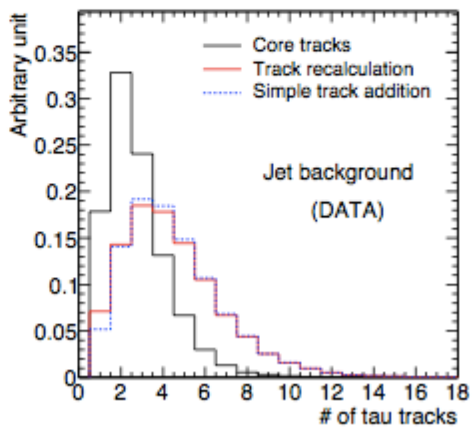
Track Multiplicity

1. Usually a hadronic taus has 1 or 3 tracks produced by the charged pions
2. These tracks are emitted in a narrow cone and in the tau reconstruction the track association is performed in a cone of radius $\Delta R=\sqrt{(\Delta\phi^2 + \Delta\eta^2)}<0.2$
3. In order to enhance the discrimination against quark or gluon initiated jets, we count the tracks in a wider cone of radius $\Delta R<0.6$
4. In order to not include tracks from the underlying event or pileup events, we take into account the pt correlation of the tracks inside the *core* ($\Delta R<0.2$) and outside with a k_T-like algorithm

$$D = \max\left\{\frac{p_T^{\text{core}}\,\Delta R(\text{core, outer})}{p_T^{\text{outer}}}\right\} \leq 4$$

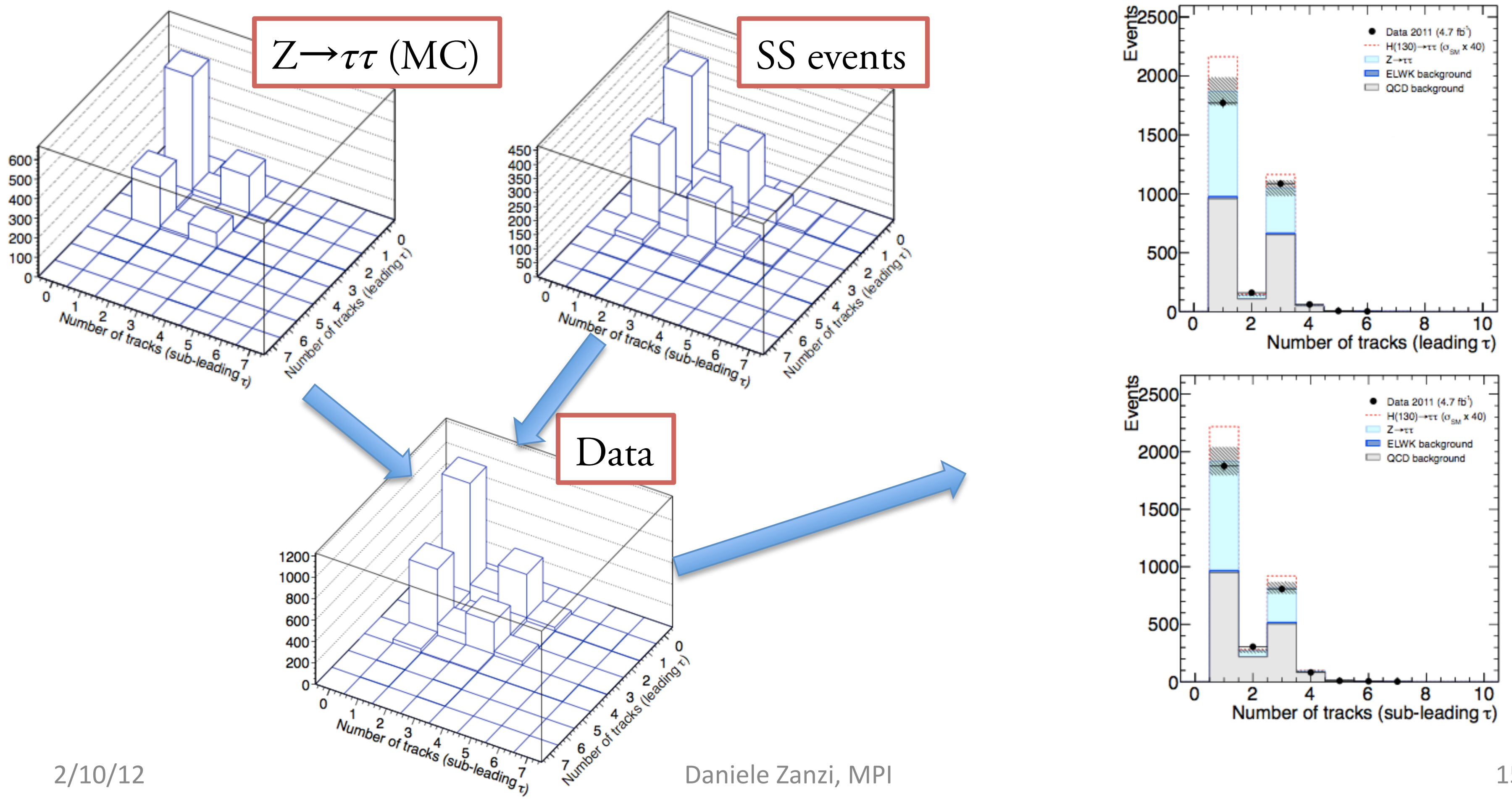






Track Fit

The track multiplicity is used to in a 2D fit to estimate the fraction of real and fake taus in the selected events:



Z shape

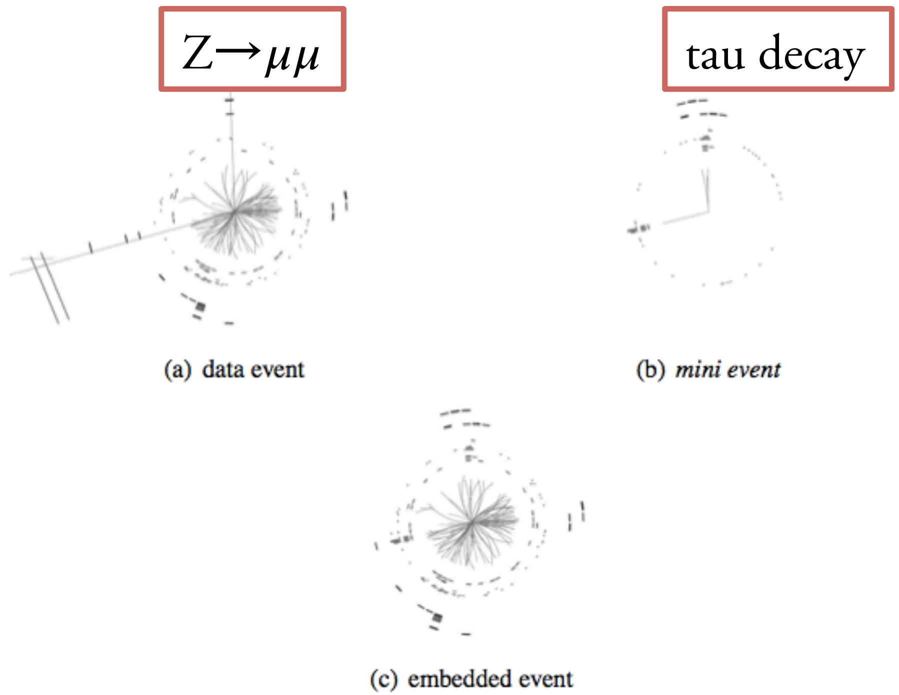
1. The final result will be based on the collinear mass fit so the modelling Z shape is a very sensitive topic

2. The mass shape is heavily dependent on features of the events which are difficult to simulate, such as hard jets multiplicity, underlying events, pileup, …

3. We want to use a shape taken directly from data, but it's not possible to perform a signal-free and pure $Z\rightarrow\tau\tau$ event selection, so…

Embedded Z→$\mu\mu$

- We select Z→$\mu\mu$,
- and we replace the muons with taus!
- These are hybrid events where only the hadronic tau decay is simulated and the rest of the events comes from real data
- The kinematics of the Z decay is kept considering also the effect of the tau mass



(a) data event

(b) *mini event*

(c) embedded event

Summary of the strategy

1. The collinear approximation provides the bkg rejection needed to suppress multijets events

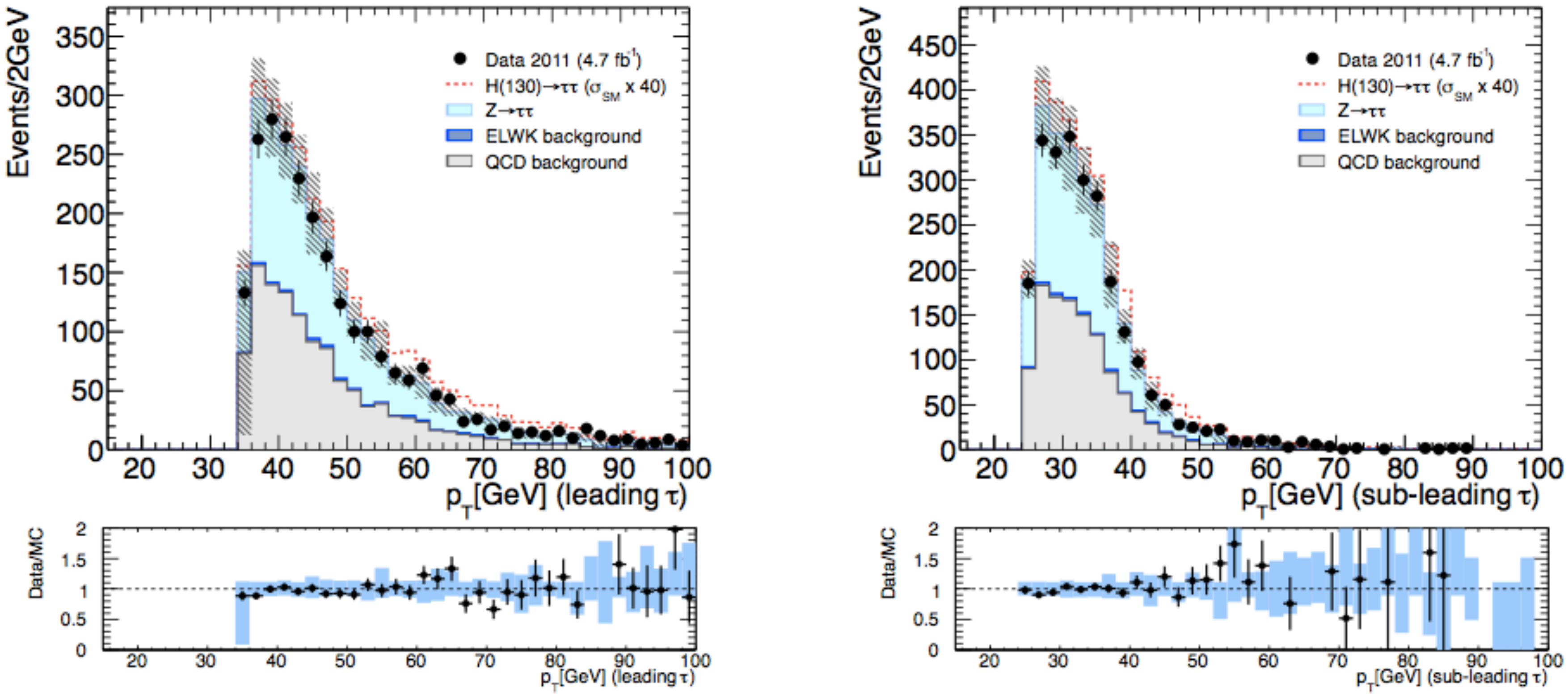
2. The two sources of background are estimated from data:

   – <u>$Z \rightarrow \tau\tau$</u>: shape from embedded $Z \rightarrow \mu\mu$ events and normalization from track fit

   – <u>Multijets</u>: shape from not-OS events and normalization from track fit

3. This strategy reduces the impact of the uncertainties on the MC simulation of background events

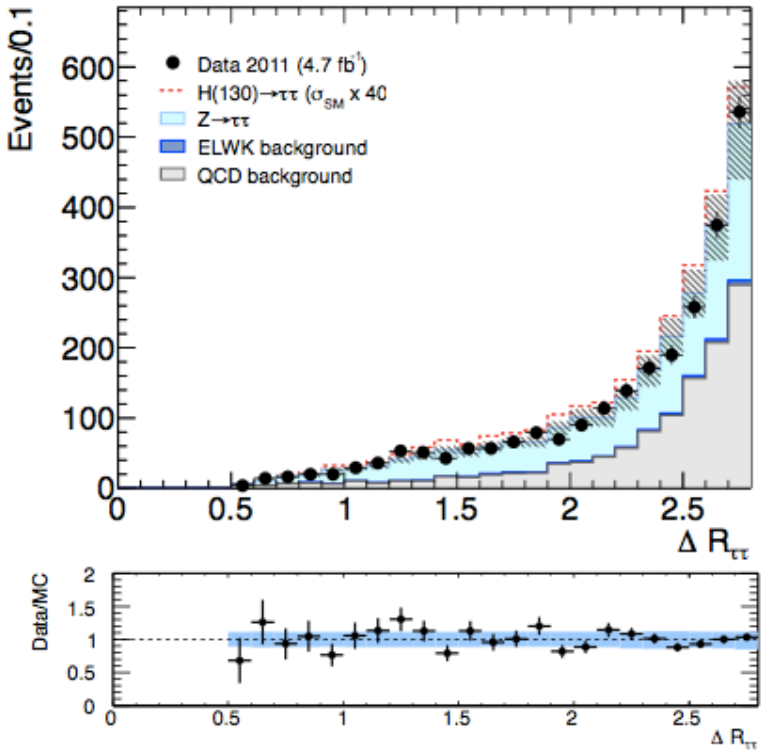
Tau Pt

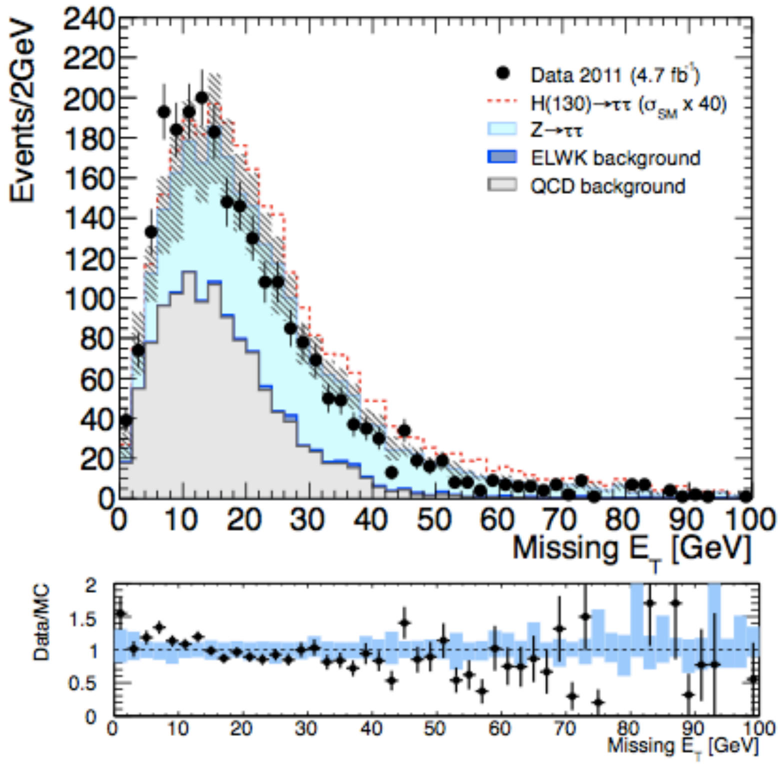


Data 2011 (4.7 fb-1)
H(130)→ττ (σSM x 40)
Z→ττ
ELWK background
QCD background
Events/2GeV
pT[GeV] (leading τ)
pT[GeV] (sub-leading τ)
Data/MC

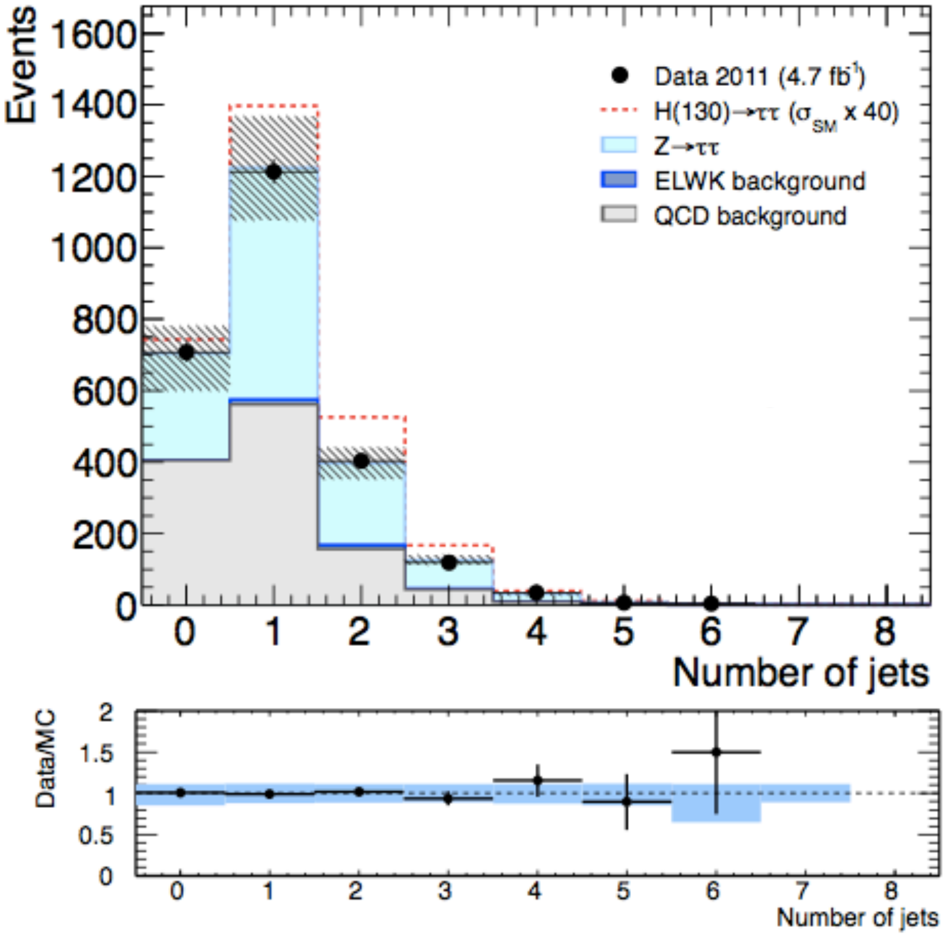
(a) (b)

$\Delta R(\tau,\tau)$ and E_T^{miss}

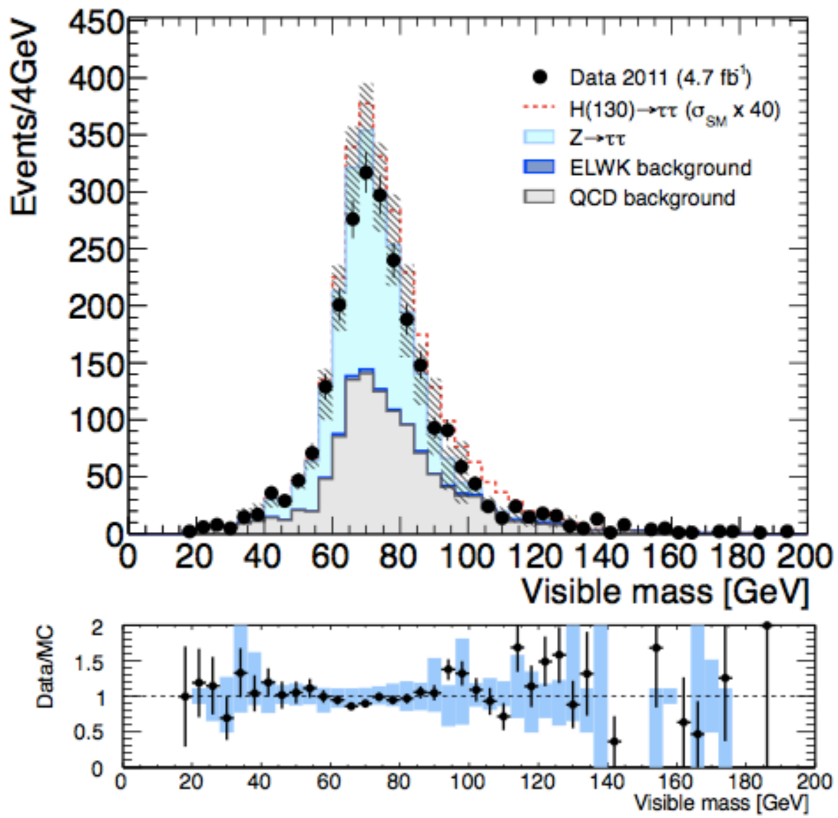


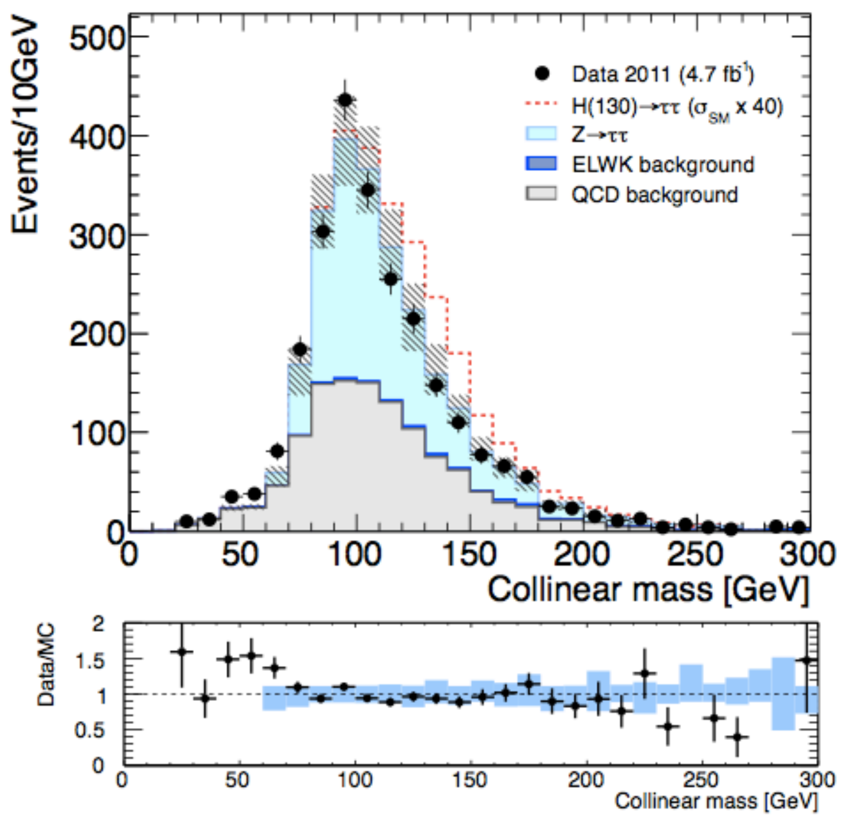


#jets with p_T>25 GeV



Mass Distribution

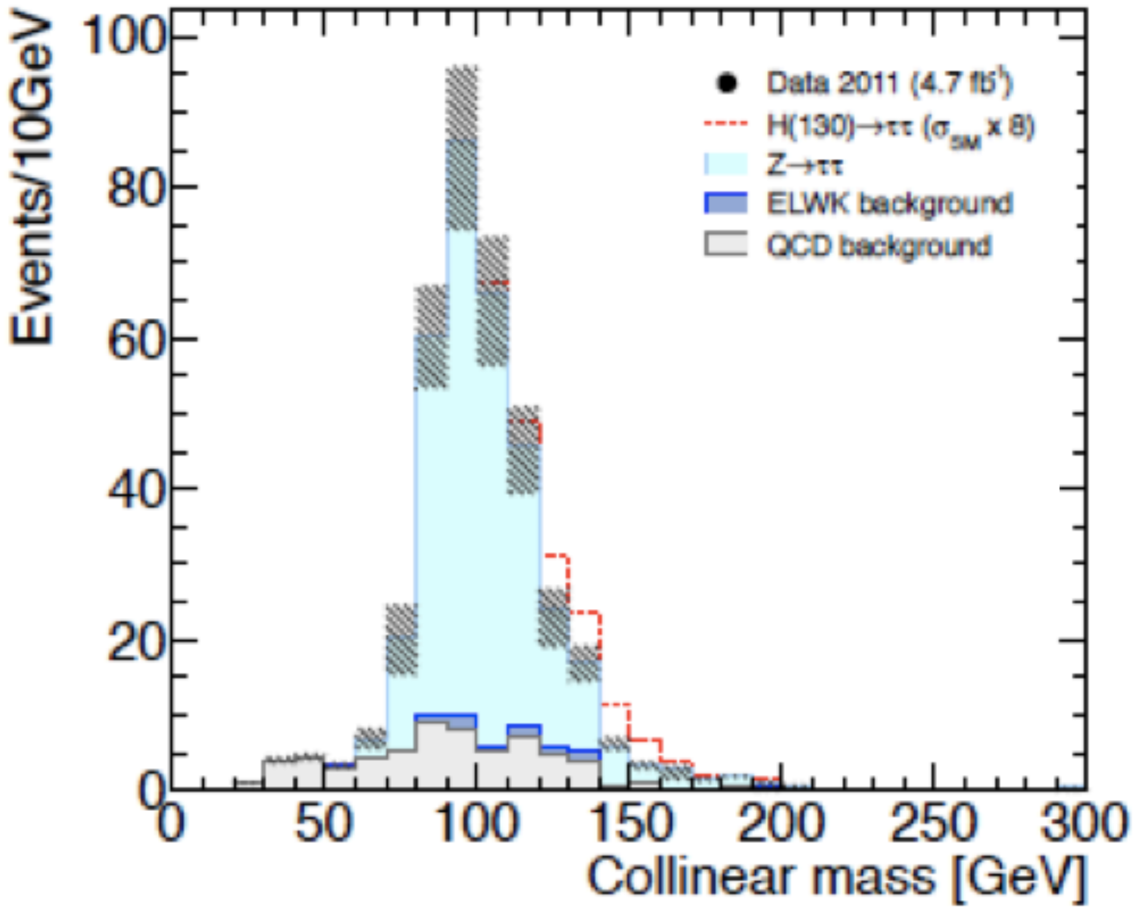




Systematic Uncertainties

1. The systematic errors on the signal acceptance come from tau related uncertainties:
   - Tau ID efficiency (4%+4%)
   - Tau Trigger 4%
   - Jet and Tau Energy Scale (~10%)
   - Theoretical prediction (up to 15% for gluon fusion)
2. Systematic errors on backgrounds estimation:
   - uncertainty on the track fit
   - uncertainty on the Z shape due to tau energy scale

EXPECTED RESULT



SUMMARY

1. We show that the search for the SM Higgs in the full hadronic di-tau final state is *not* hopeless

2. We define an event selection which is able to have a very little event yield from multijets and the rest is irreducible background, i.e. $Z\rightarrow\tau\tau$

3. All background estimations is performed with data-driven methods

4. Now that the evaluation of the systematic errors is done, we can understand better the weak points of the analysis and improve the strategy for the analysis on the new coming data