ttH(H->bb) search with the ATLAS detector at the LHC experiment.

Thomas Calvet
Supervisor: Arnaud Duperrin
Co-supervisor: Georges Aad

CPPM *PhD day*
2016-10-07

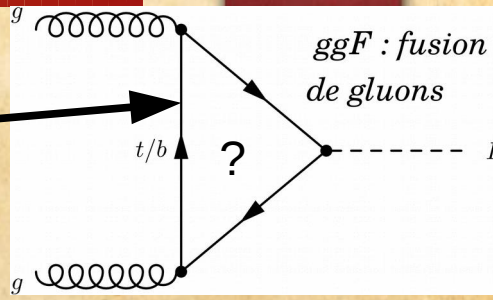
ttH(H->bb) search at ATLAS in Run 2, why ?

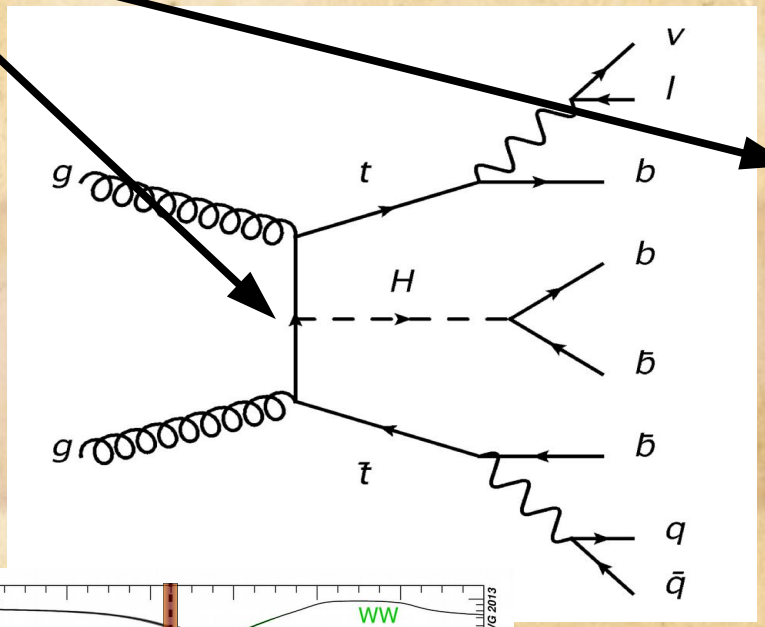
→ **ttH gives direct measurement of H<->t coupling:**

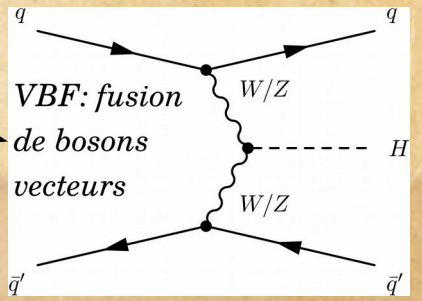
→ ggF: anything can happen in the loop.

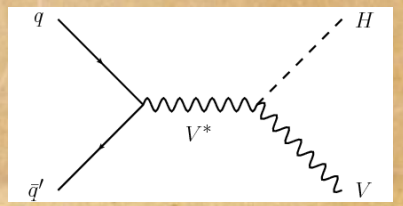
→ VBF, VH: no vertex H-t.

→ ttH: out of loops vertex !





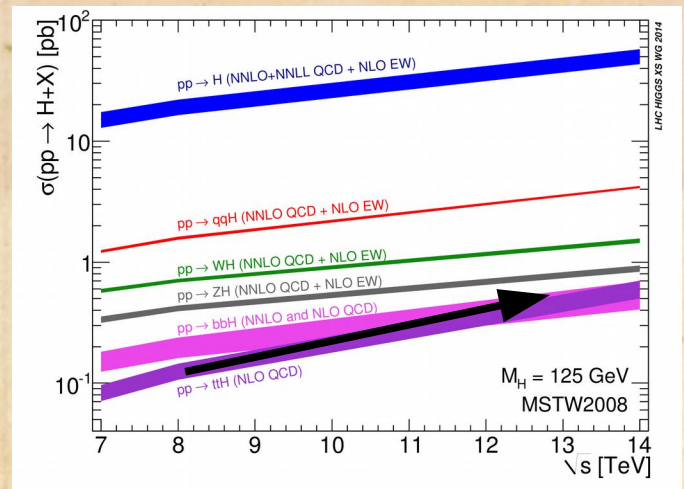


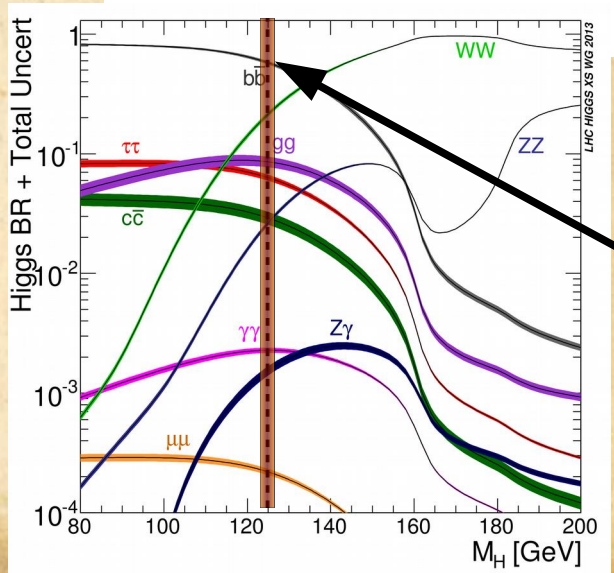


→ **Not discovered in Run 1:**

→ Run 2: (7,8)TeV → 13TeV

=> ttH cross section x3.7





→ **(H→bb):**

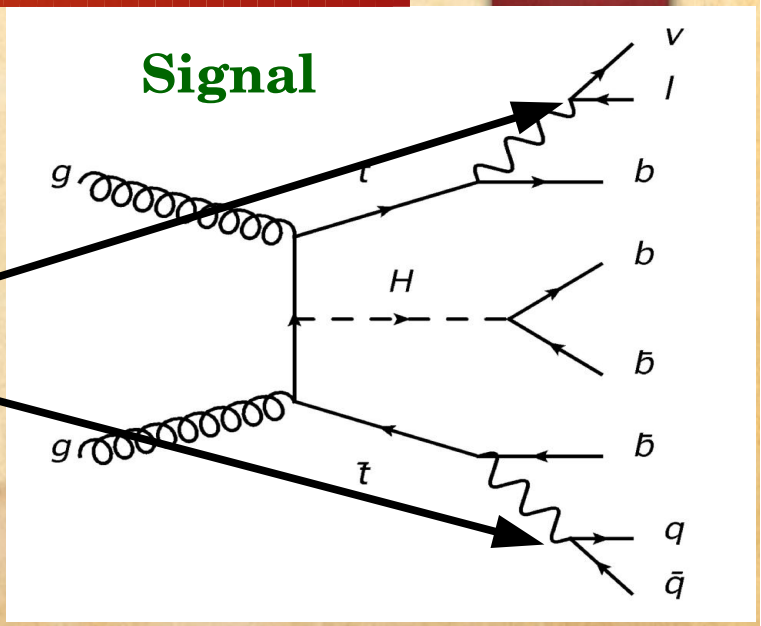
→ Not yet observed !

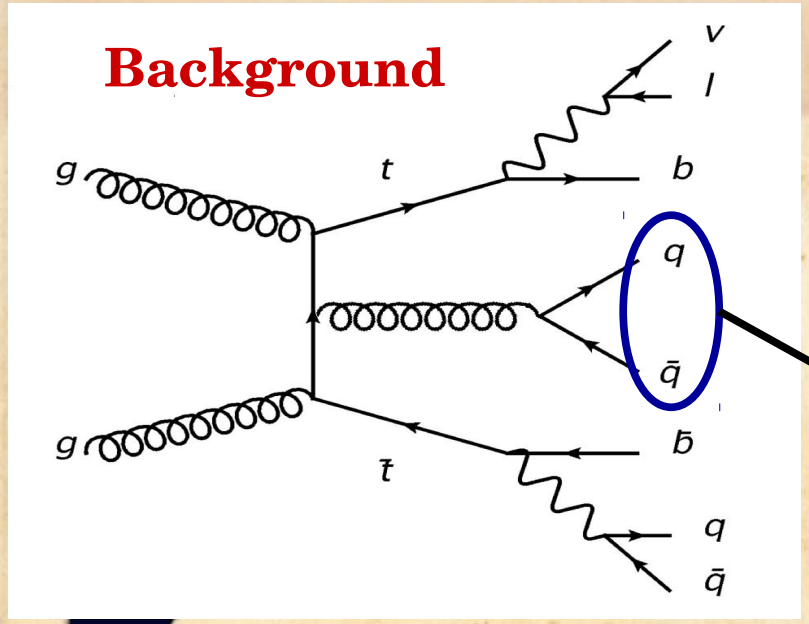
→ Highest BR for M_H=125Gev.

→ Expect more stat :)

ttH(H->bb) lepton+jets background

→ ttH(H→bb) signal final state:
- → 4b-quarks and 2W.
- → In this talk, single lepton channel.





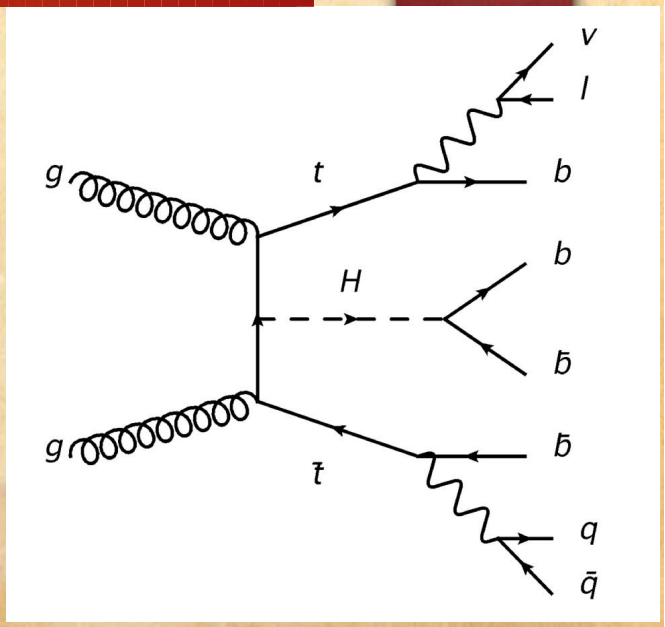
→ tt+jets background:
- → Main background.
- → σ (tt+jets) >> σ (ttH).
- → tt+jets cross section increased by a factor 3.2 from Run 1 to Run 2.

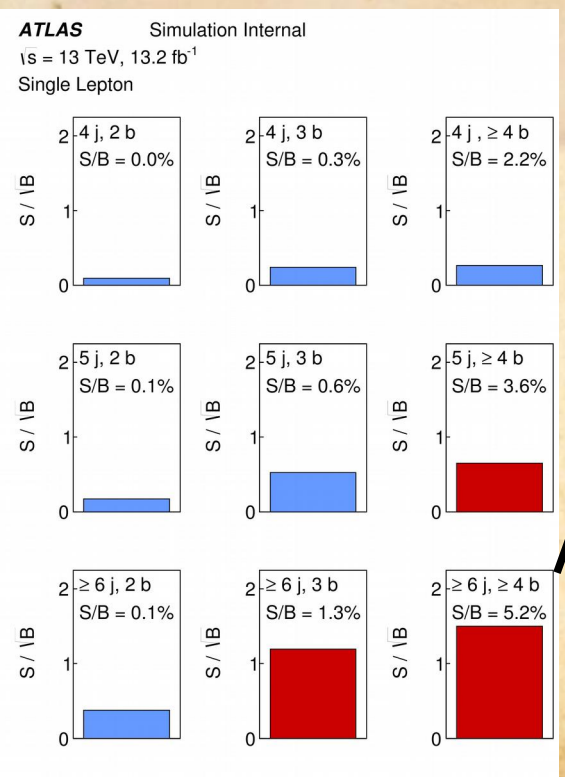
→ Splitted according to additionnal jet flavour:
- → tt+bb : irreducible
- → tt+light & tt+cc : reducible

ttH(H->bb) analysis strategy – Region definition.

→ **Select ttbar like events:**
- → **at least 4jets** out of which **at least 2b-tags**.
- → Exactly **one isolated lepton**.
- → At this point much more background than signal :s

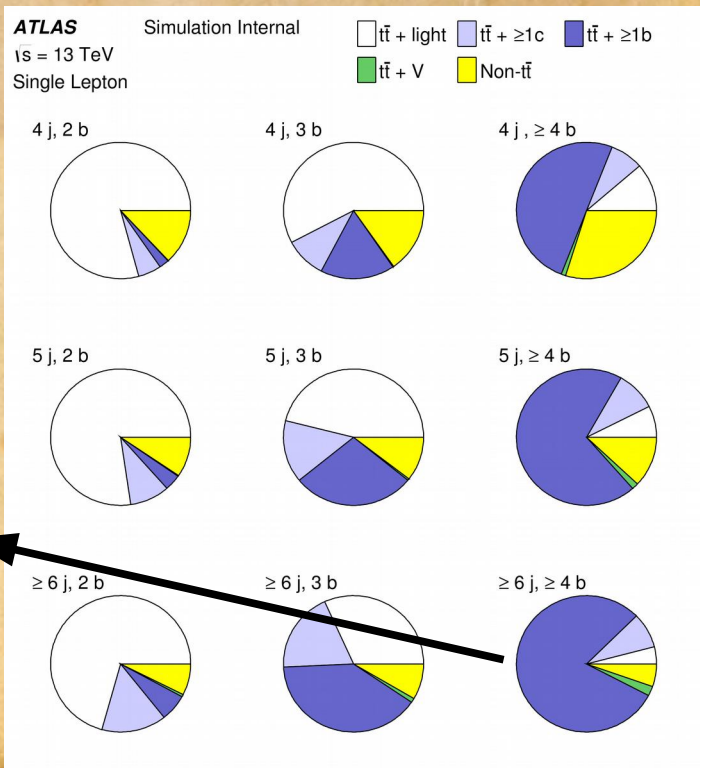
→ **Further split in regions:**
- → Increase signal purity.
- → Define **Signal Regions** at high N(jets)/N(b-rags) and **Control Regions.**





Most signal like region only 5.2% ttH.

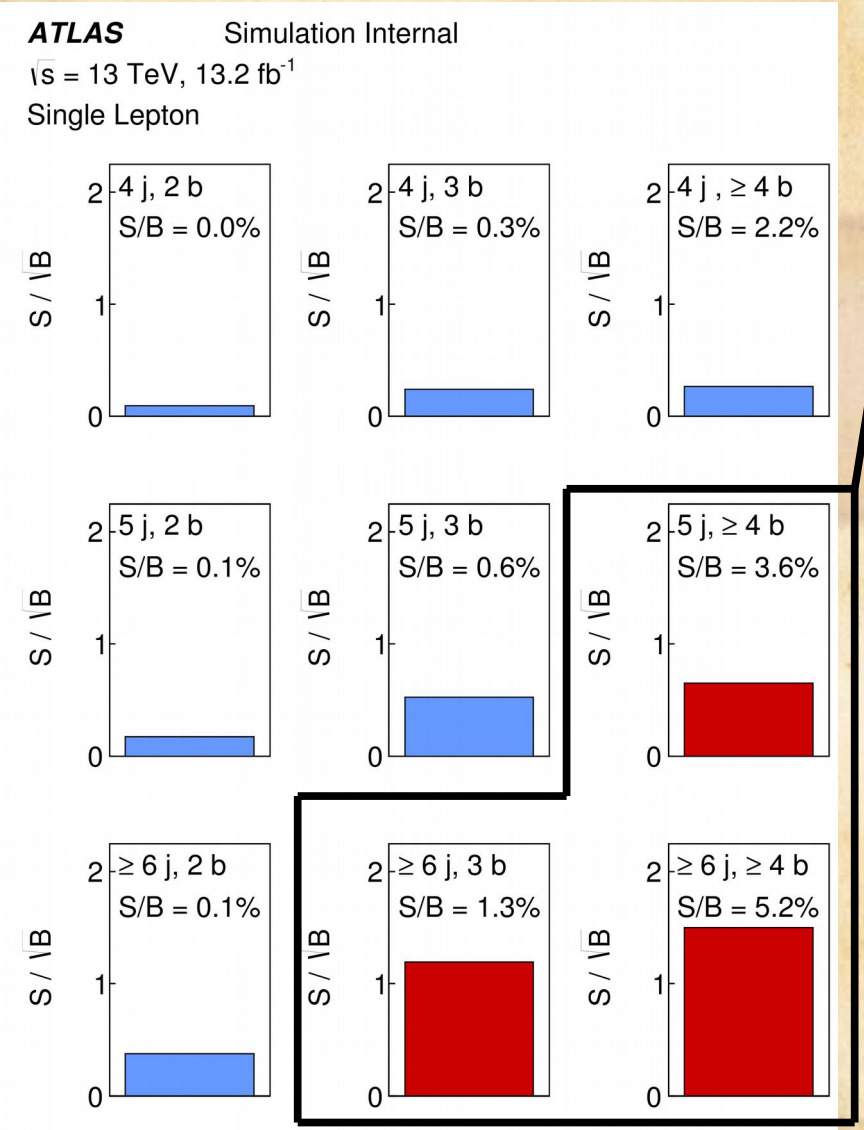
Completely dominated by irreducible tt+bb.



ttH(H->bb) analysis strategy – Signal Regions.

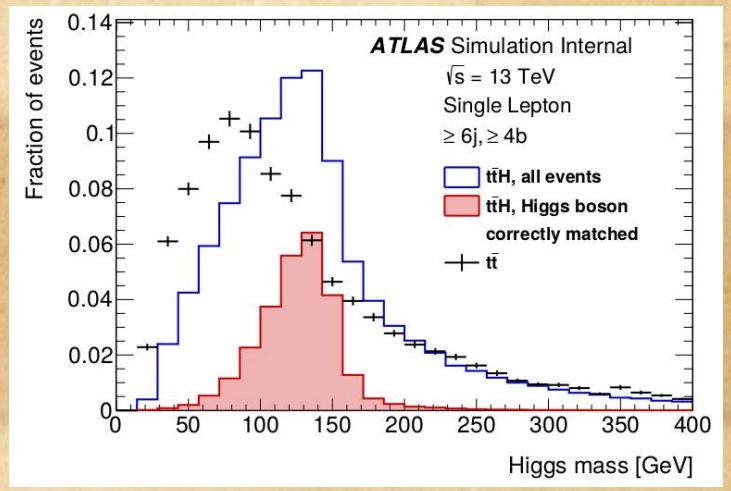
→ Select signal like events.

→ Further split in regions.

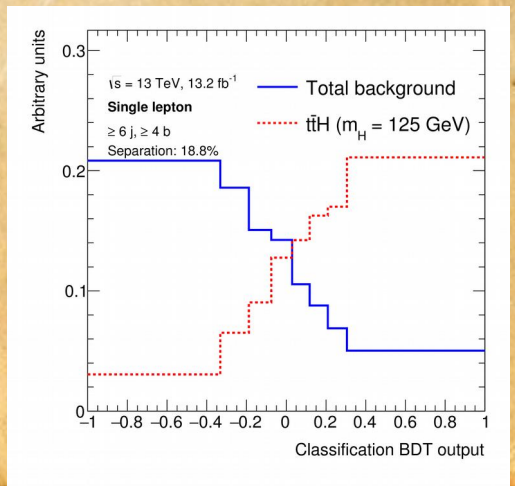


Need to increase sensitivity in SR.

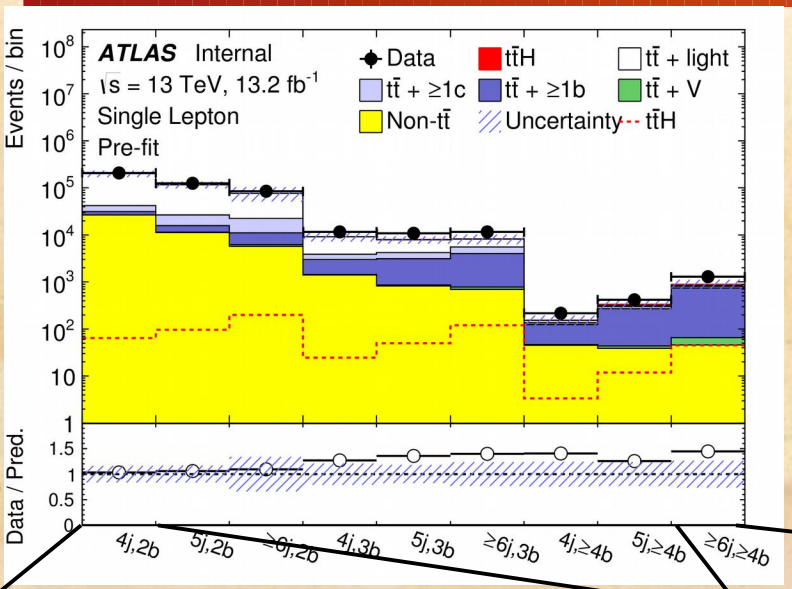
→ Step 1: reconstruct the ttH system with a Boosted Decision Tree.



→ Step 2: differentiate ttbar event from ttH events with a BDT.



ttH(H->bb) analysis strategy – Fit philosophy.

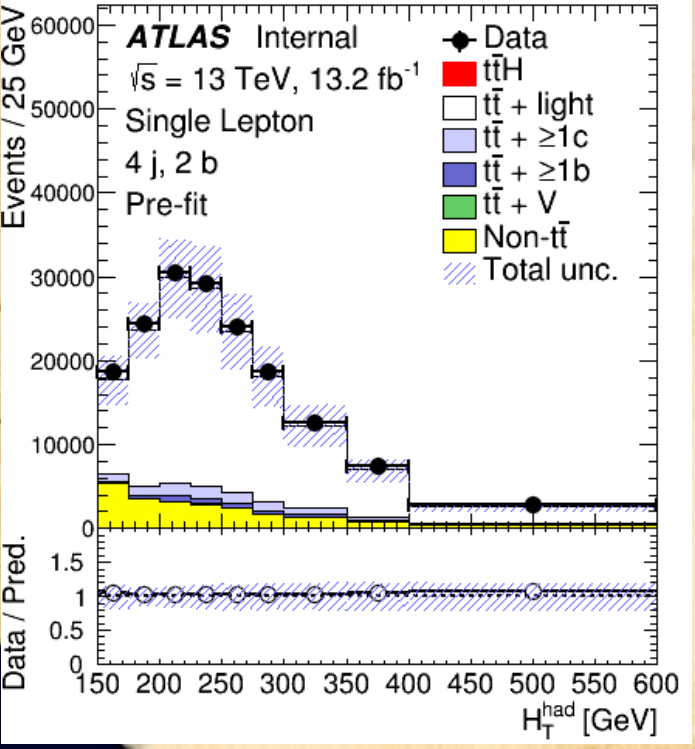


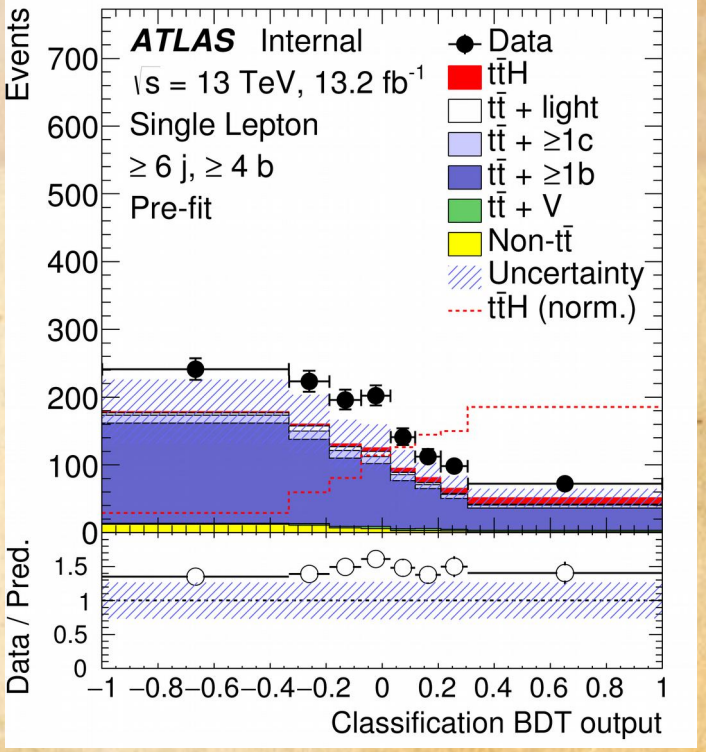
ttbar miss-modelling:

→ Add **large systematics uncertainties** to cover miss-modelling.

→ **Fit also control regions** to constrain ttbar modelling.

→ Use simple variable in control region: $H_t^{had}=\sum p_T(jets)$.





Fit BDT in SR:

→ High signal sensitivity

Use several bins in each region:

→ **Fit also the shape** of the distributions.

→ Binning need also optimization.

ttH(H->bb) systematics.

→ **Systematics dominated analysis.**

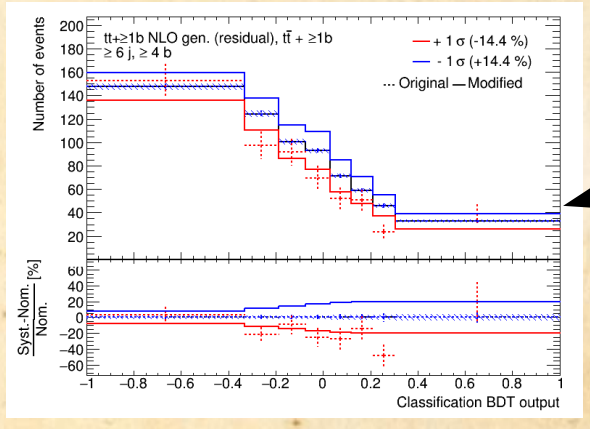
→ **132 systematics !**

→ Reco objets (leptons and **jets**), **flav-tagging**, **ttbar** and other bkg modelling.

→ **Apply tt+HF normalizations as free floating parameters in the fit.**

→ **Large Data/MC differences.**

→ **Important slope in N(jets) and N(b-jets).**



Systematic uncertainty	Type	Components
Luminosity	N	1
Reconstructed Objects		
Electron trigger+reco+ID+isolation	SN	5
Electron energy scale+resolution	SN	2
Muon trigger+reco+ID+isolation	SN	6
Muon momentum scale+resolution	SN	3
Pileup modelling	SN	1
Jet vertex tagger	SN	1
Jet energy scale	SN	19
Jet energy resolution	SN	1
Missing transverse momentum	SN	3
b-tagging efficiency	SN	5
c-tagging efficiency	SN	4
Light-jet tagging efficiency	SN	14
High-p_{T} tagging	SN	2
Background and Signal Model		
$t\bar{t}$ cross section	N	1
$t\bar{t}$+HF: normalisation	N	2
$t\bar{t}$+≥ 1b: NLO Shape	SN	10
$t\bar{t}$+≥ 1c: NLO Shape	SN	1
$t\bar{t}$ modelling: residual Radiation	SN	3
$t\bar{t}$ modelling: residual NLO generator	SN	3
$t\bar{t}$ modelling: residual parton shower+hadronisation	SN	3
$t\bar{t}$ NNLO reweighting	SN	4
W+jets normalisation	N	6
Z+jets normalisation	N	6
Single top cross section	N	2
Single top model	SN	2
Diboson normalisation	N	1
Fakes normalization	SN	7
$t\bar{t}V$ cross section	N	4
$t\bar{t}V$ modelling	SN	2
$t\bar{t}H$ cross section	N	2
$t\bar{t}H$ branching ratios	N	4
$t\bar{t}H$ modelling	SN	2

ttH(H->bb) the fit.

→ Use binned Maximum Likelihood function:

9 regions

from 3 to 11 bins (total 68)

$$L_{Pois}(\mu,\kappa,\theta) \approx \prod_{reg}^{r} \prod_{bin}^{r} \frac{\left(N_{r,i}^{\exp}(\mu,\kappa,\theta)\right)^{N_{r,i}^{data}}}{N_{r,i}^{data}} e^{-N_{\exp}^{r,i}(\mu,\kappa,\theta)}$$

θ = Nuisance Parameters
= Systematics + gammas
gammas= MC stat bin errors

$$N_{\exp}(\mu,\kappa,\theta) = \mu \cdot N_{sig}^{MC}(\theta_i) + \sum_k \kappa_k \cdot N_{bkg\,k}^{MC}(\theta_i) + N_{bkg\,n,\ldots}(\theta_i)$$

ttH

tt+≥1b and tt+≥1c

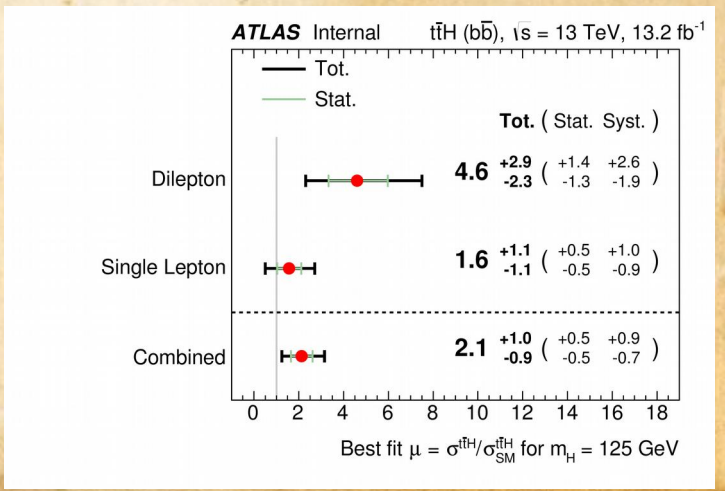
12 others

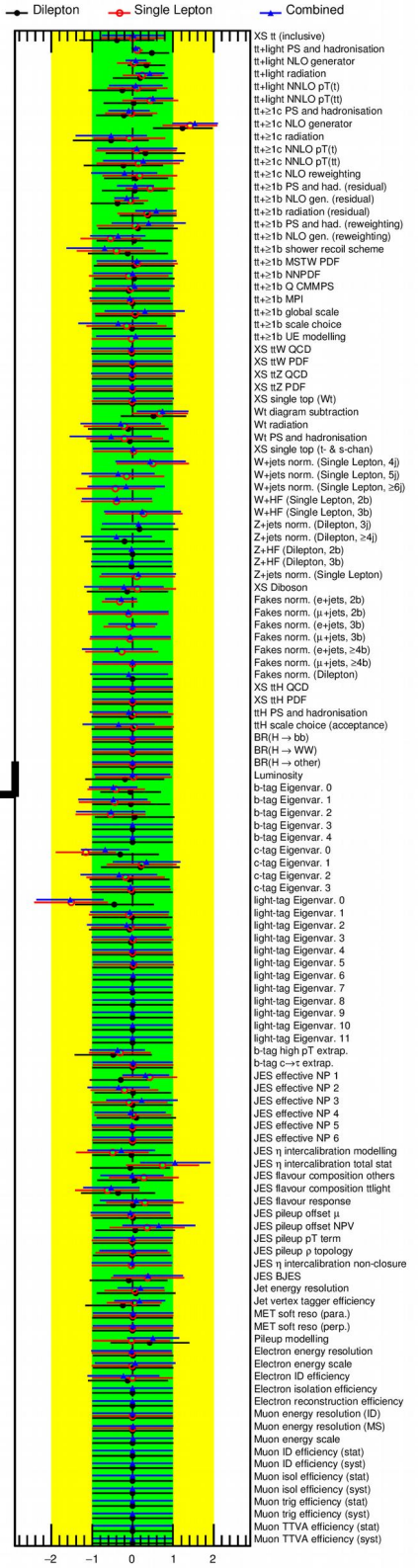
$$L(\mu,\kappa,\theta) = L_{Pois}(\mu,\kappa,\theta) \prod_p \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\theta_p^2}{2}\right)$$

prior or penalty term

TtHFitter

Uncertainty source	$\Delta\mu$	
$t\bar{t}+\geq 1b$ modelling	+0.53	−0.53
Jet flavour tagging	+0.26	−0.26
$t\bar{t}H$ modelling	+0.32	−0.20
Background model statistics	+0.25	−0.25
$t\bar{t}+\geq 1c$ modelling	+0.24	−0.23
Jet energy scale and resolution	+0.19	−0.19
$t\bar{t}$+light modelling	+0.19	−0.18
Other background modelling	+0.18	−0.18
Jet-vertex association, pileup modelling	+0.12	−0.12
Luminosity	+0.12	−0.12
$t\bar{t}Z$ modelling	+0.06	−0.06
Light lepton (e, μ) ID, isolation, trigger	+0.05	−0.05
Total systematic uncertainty	+0.90	−0.75
$t\bar{t}+\geq 1b$ normalisation	+0.34	−0.34
$t\bar{t}+\geq 1c$ normalisation	+0.14	−0.14
Statistical uncertainty	+0.49	−0.49
Total uncertainty	+1.02	−0.89





ttH(H->bb) Run 2 results.

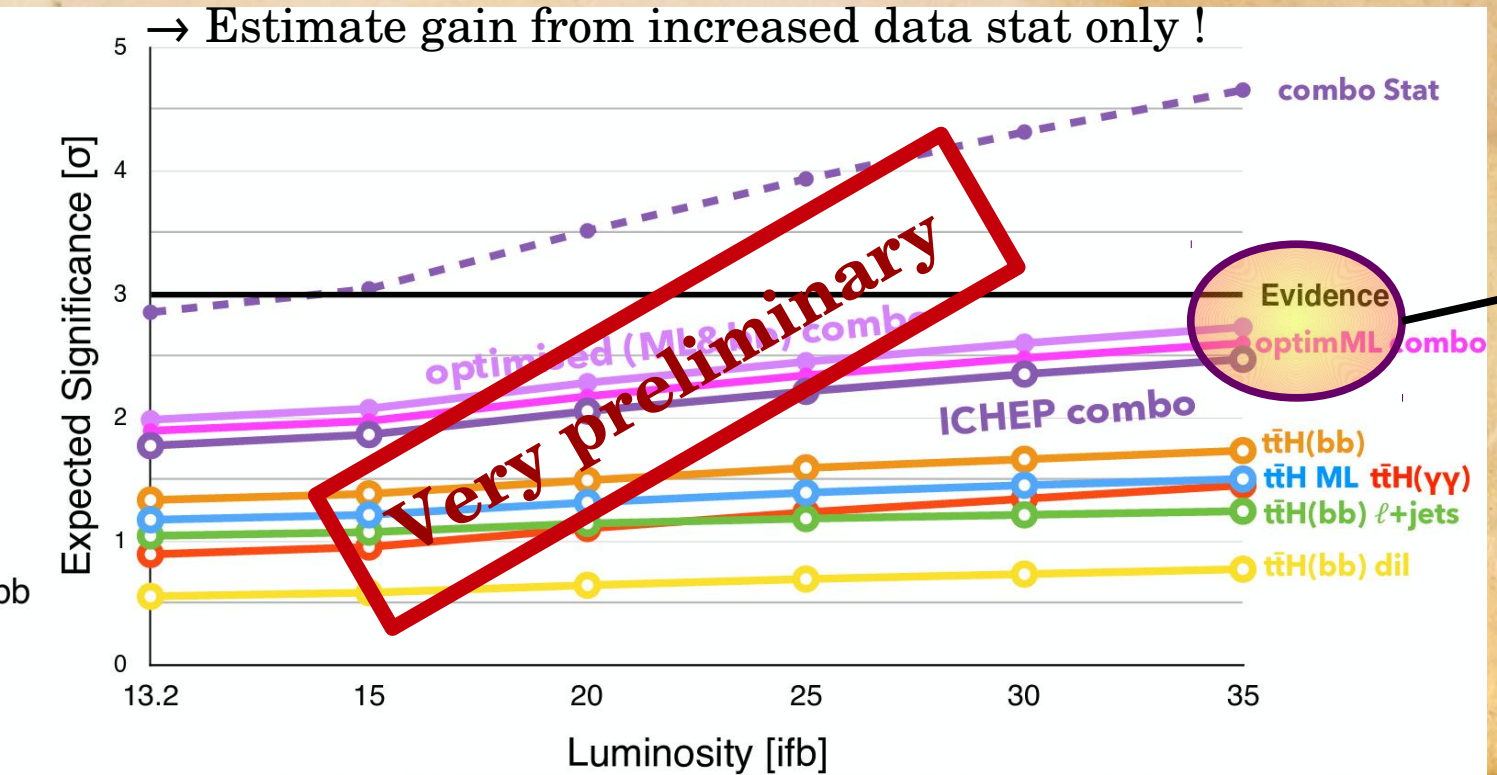
→ **Significance improved compared to Run 1:**

	√s	Luminosity	Observed Significance	Expected Significance
ATLAS Run 1	8 TeV	20.3 fb^{-1}	1.4 σ	1.1 σ
ATLAS Run 2	13 TeV	13.2 fb^{-1}	2.4 σ	1.2 σ

→ **Expected projections towards paper:**

→ Around 35 fb^{-1} should be available.

→ Estimate gain from increased data stat only !



Evidence is not so far for ttH combination :)

Summary

→ **ttH(H→bb) is an important analysis for Higgs boson characterization:**

- → Very **challenging** process:
  - → **Lots of jets/b-jets** in final state.
  - → **Low signal purity** (5.1%) even in the most sensitive region !
- → Apply complicated techniques to **increase sensitivity**:
  - → Final state object **reconstruction with BDT**.
  - → **ttH/tt+jets separation** with a second layer **BDT**.

→ **Statistical analysis of data with binned Maximum Likelihood:**

- → **Many fitted parameters**: 9 regions with their binning, 15 samples, 132 systematics, ...
- → Powerful fitting tool **TtHFitter.**
- → Large **differences** between **expectations and data**.
  - → **Fit uses a lot of parameters to correct miss-modelling**.

→ **Not yet reached evidence:**

- → Analysis published in a **conf note for ICHEP 2016**.
- → **Better results than Run 1**
- → May have **ttH evidence with 35 fb^{-1}** :)

Backup.