Search for the Higgs boson produced in association with a vector boson and decaying into a pair of *b*-quarks using large-R jets with the ATLAS detector

Giulia Di Gregorio

16th September 2019

Outline

- Introduction
- Motivation and status of $H \to b\bar{b}$ search
- Analysis strategy
- Preliminary results
- Conclusions
  - Personal contributions
  - Next steps

Introduction: Higgs sector

Run 1

2009

2012

2013

- Higgs discovery
- Precise determination of the Higgs boson mass
- Observation of decays in vector bosons

Run 2

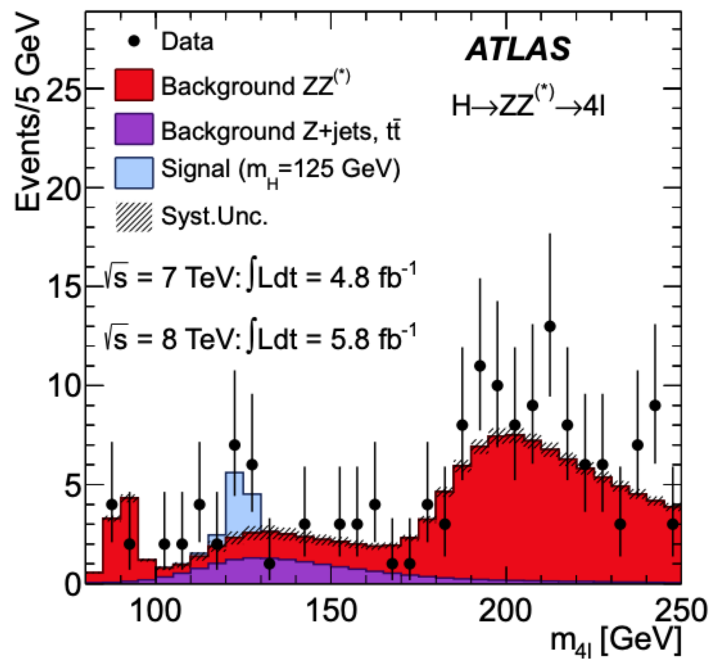
2015

2018

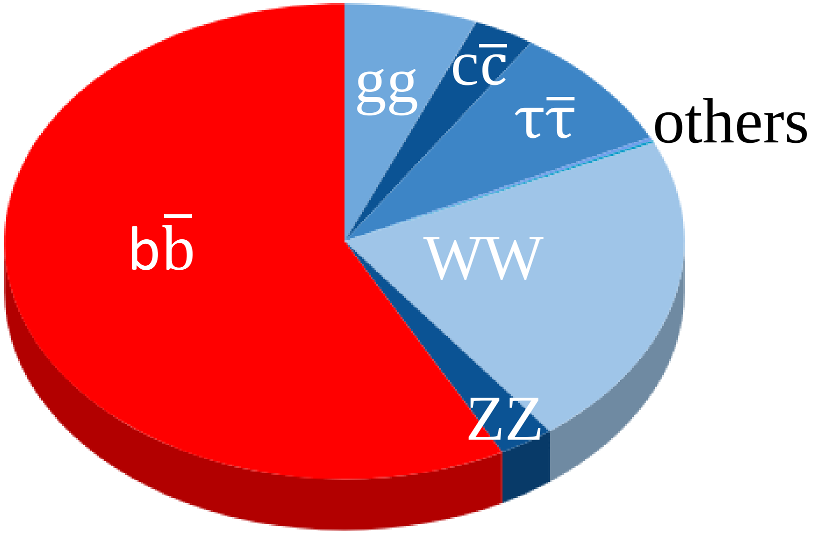
- Precise measurement of Higgs properties
- Observation of Yukawa couplings to fermions: tau, top and ***b*** quarks

Today

2012. arXiv:1207.7214



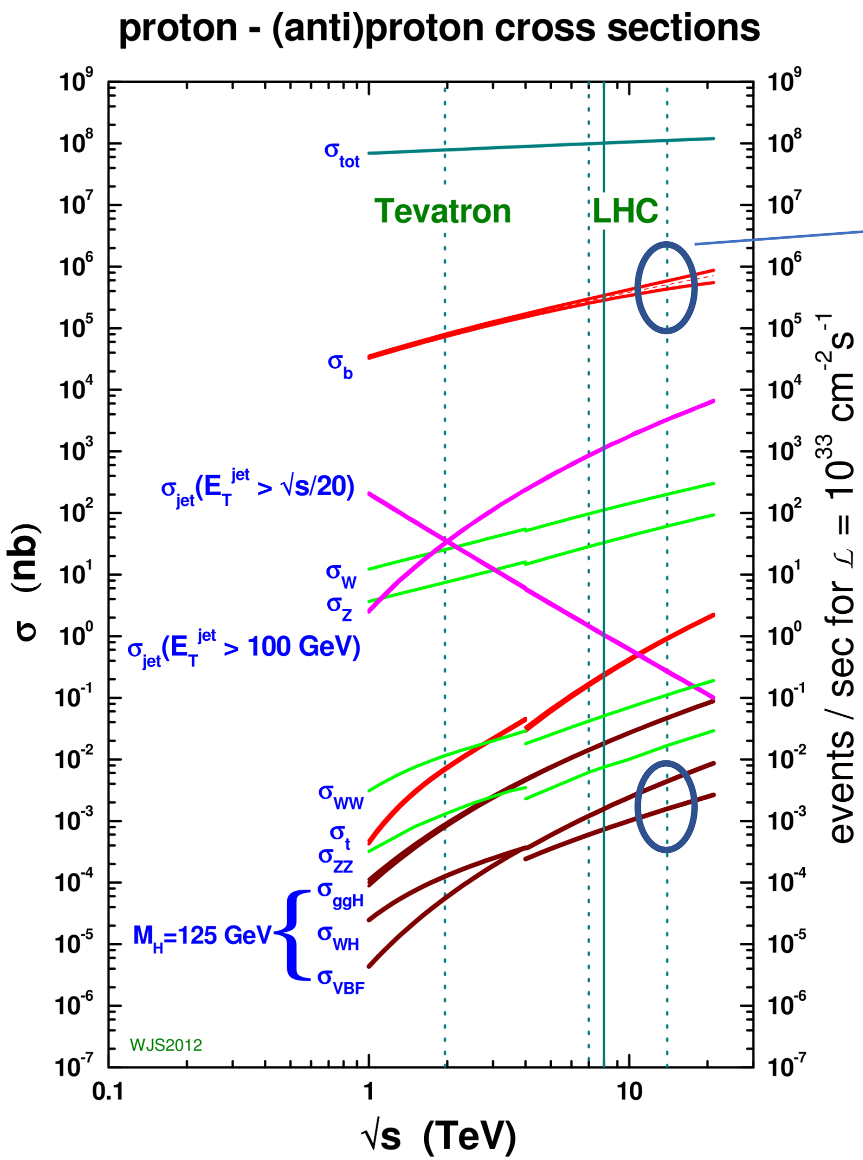
Why is the $H \rightarrow b\bar{b}$ channel interesting?

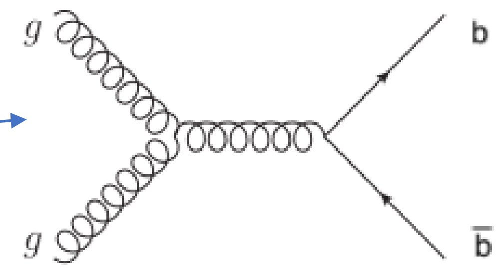


- Large BR (58%)
- Direct coupling between Higgs and fermions

- Why did it take so long to find the largest Higgs boson decay mode?

LHC cross sections

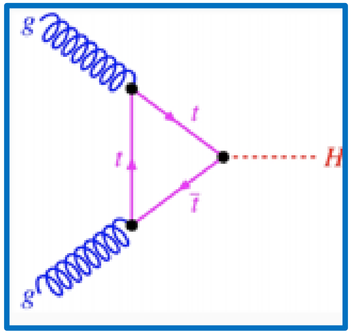


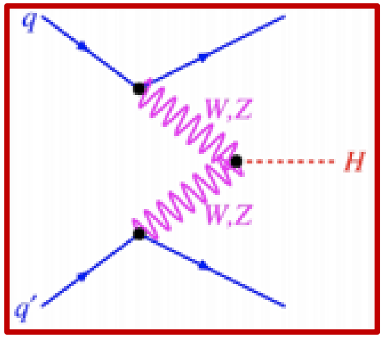


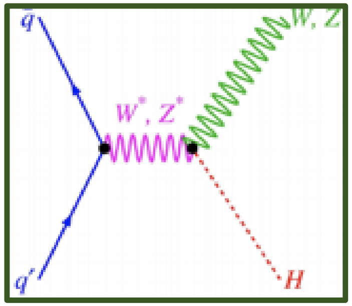
Background of multi-jet production is 8 order of magnitude higher!

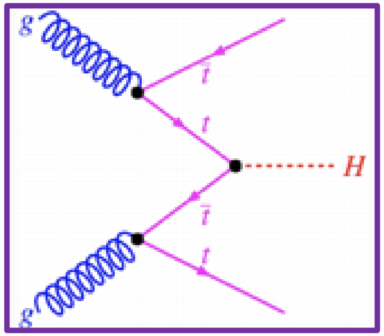
Higgs production modes at LHC

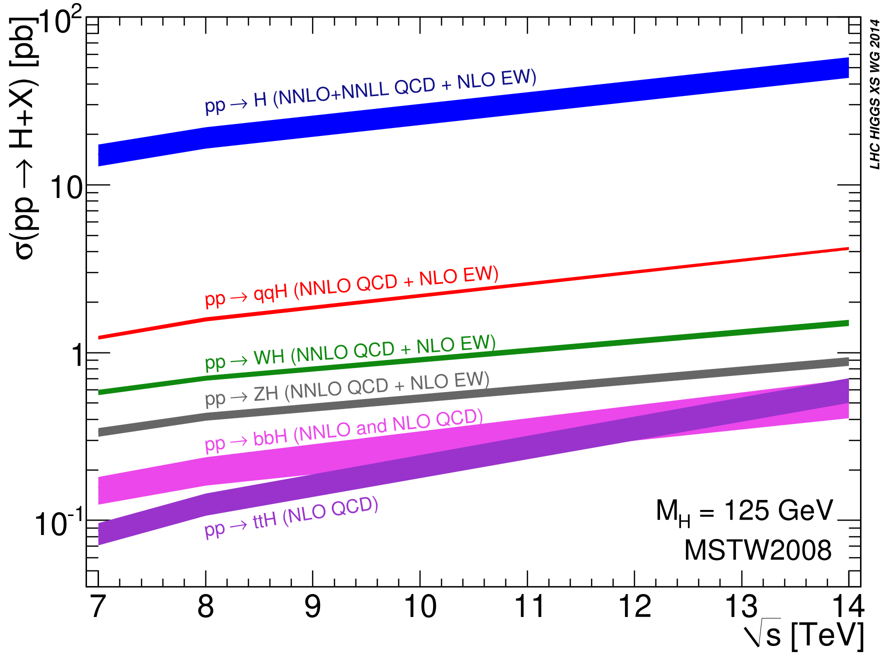
- **Gluon-gluon fusion → σ = 49 pb**
- **Vector boson fusion → σ = 4 pb**
- **Associated with a vector boson V (V=Z or V=W) → σ = 2 pb**
- **Associated with top quarks → σ = 1 pb**







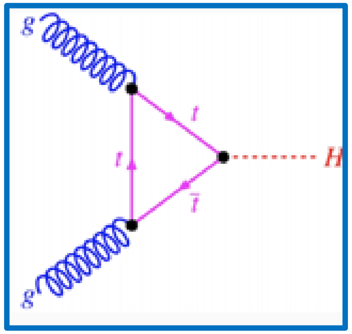


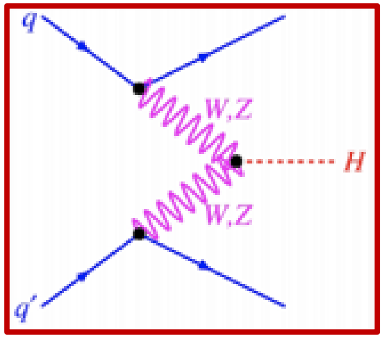


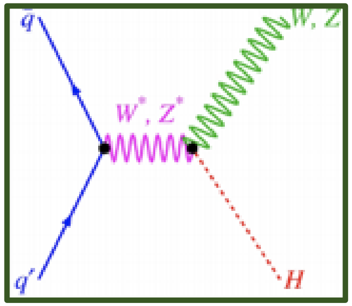
Higgs production modes at LHC

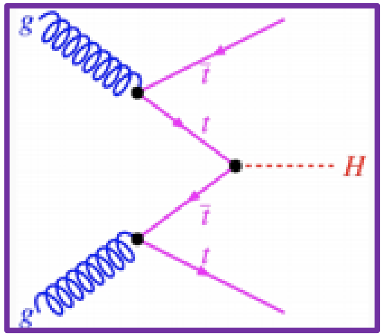
- **Gluon-gluon fusion → σ = 49 pb**
- **Vector boson fusion → σ = 4 pb**
- **Associated with a vector boson V (V=Z or V=W) → σ = 2 pb**
- **Associated with top quarks → σ = 1 pb**

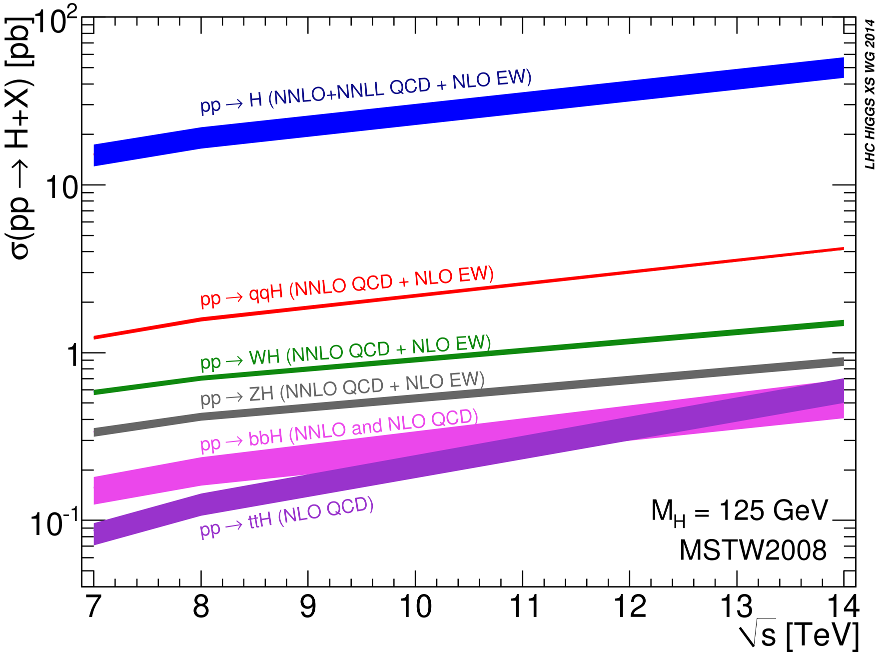
V leptonic decay → clear signature!









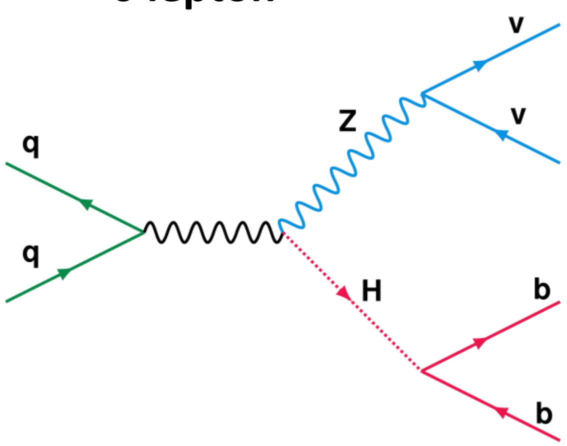


VH, H→$b\bar{b}$ channel

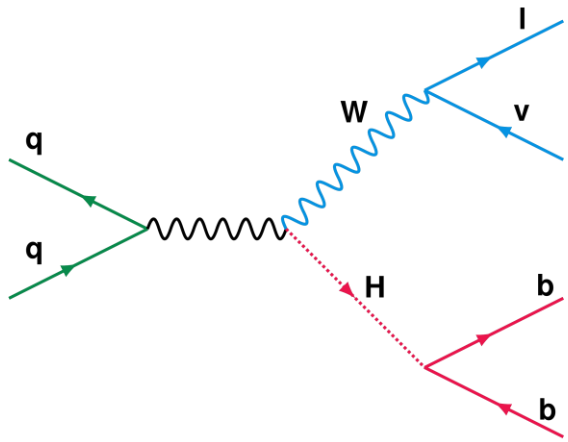
Three topologies studied:

- **H→$b\bar{b}$, Z→ $\nu\bar{\nu}$** ⟶ 0 charged leptons
- **H→$b\bar{b}$, W→ lν** (l=e,μ) ⟶ 1 charged lepton
- **H→$b\bar{b}$, Z→ $l\bar{l}$** (l=e,μ) ⟶ 2 charged leptons

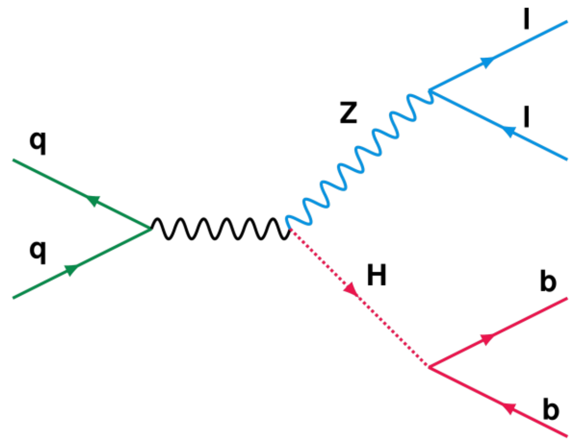
0 lepton



1 lepton



2 lepton

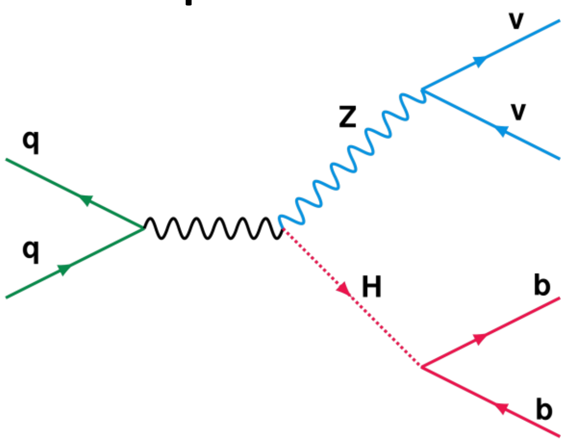


VH, $H \to b\bar{b}$ channel

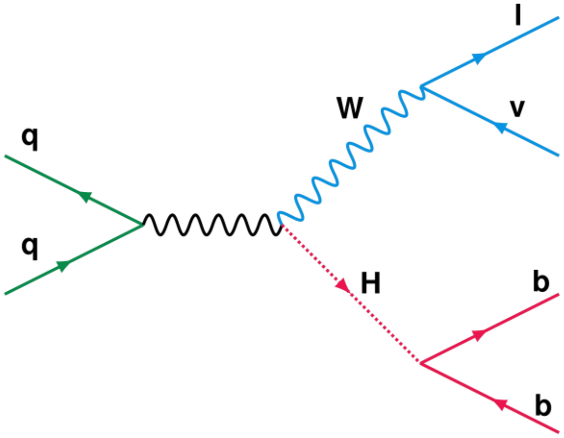
Three topologies studied:

- $\mathbf{H \to b\bar{b}, Z \to \nu\bar{\nu}}$ ⟶ 0 charged leptons
- $\mathbf{H \to b\bar{b}, W \to l\nu}$ (l=e,μ) ⟶ 1 charged lepton
- $\mathbf{H \to b\bar{b}, Z \to l\bar{l}}$ (l=e,μ) ⟶ 2 charged leptons

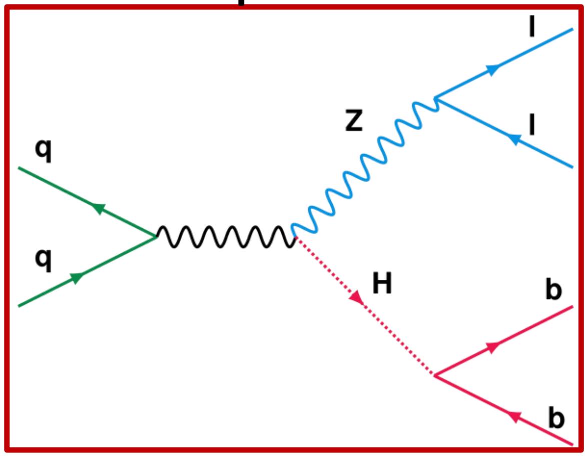
0 lepton

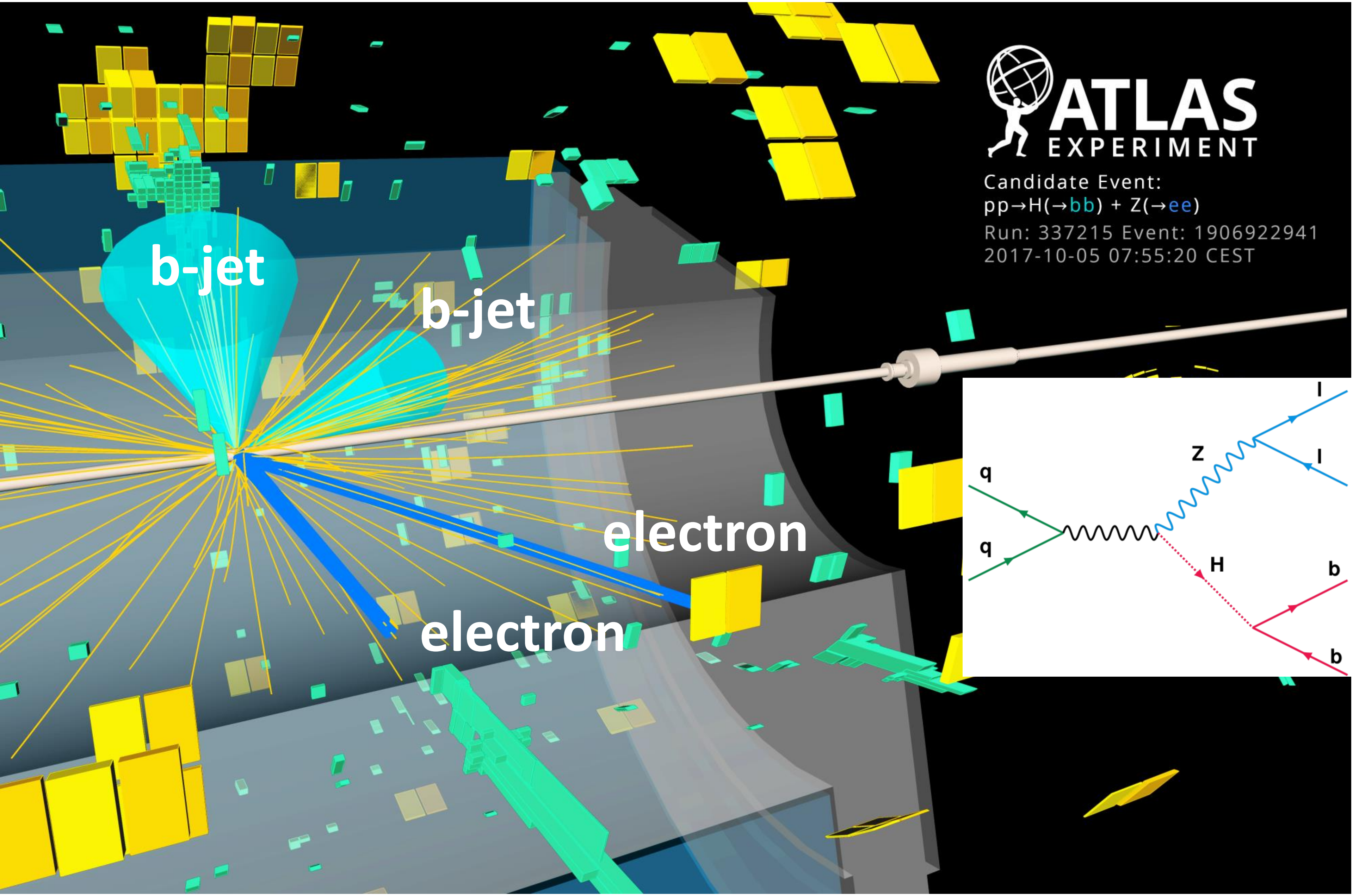


1 lepton

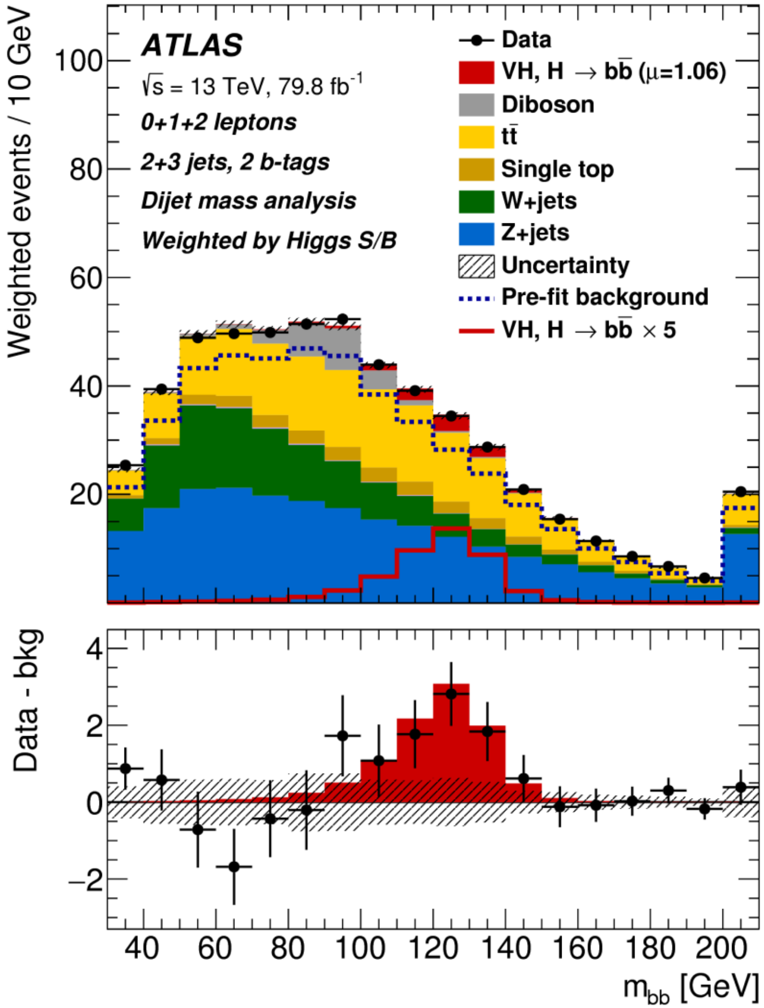


2 lepton




ATLAS
EXPERIMENT
Candidate Event:
pp→H(→bb) + Z(→ee)
Run: 337215 Event: 1906922941
2017-10-05 07:55:20 CEST
b-jet
b-jet
electron
electron
q
q
Z
H
l
l
b
b

Example of previous results



- Requiring leptons, MJ background is suppressed
- BUT bkg events are still there!

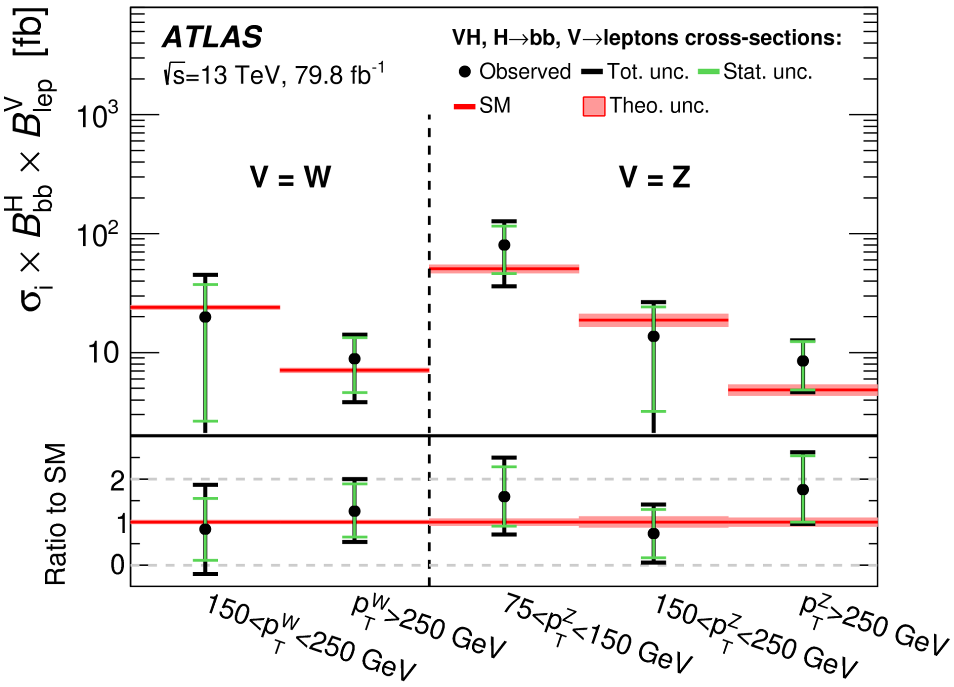
- **VH** = signal events
- **Diboson**, **Z+jets**, **W+jets**, **top** = background events

Thanks to advance analysis techniques (including mva)

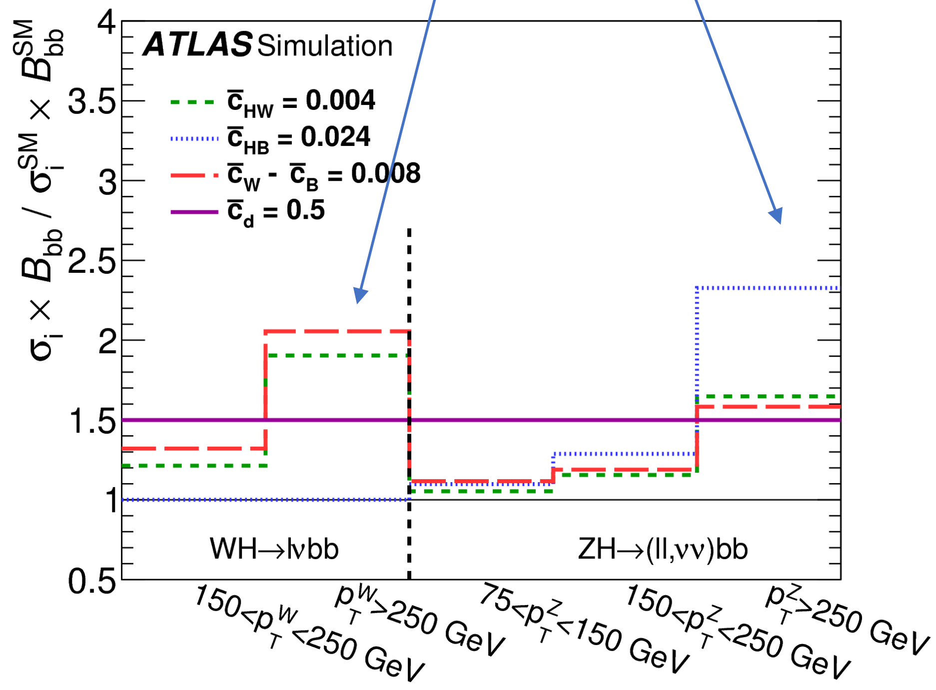
- **Observation** of **H→$b\bar{b}$** decay mode (obs. significance **5.4 σ**)
- **Observation** of **VH** production mode (obs. significance **5.3 σ**)

Cross section results

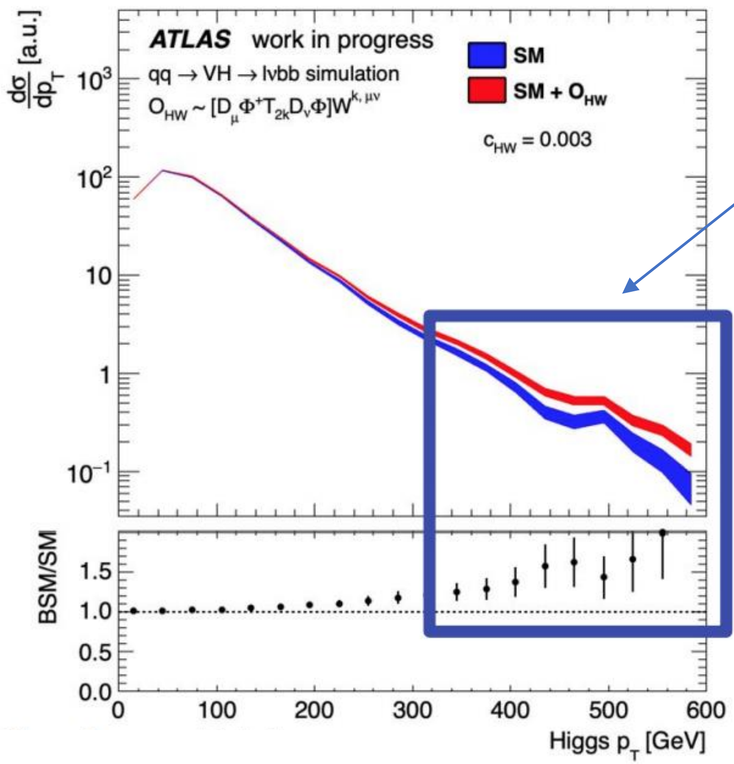
First differential cross section pp → VH measurements



High energy events more sensitive to Beyond Standard Model (BSM) effects



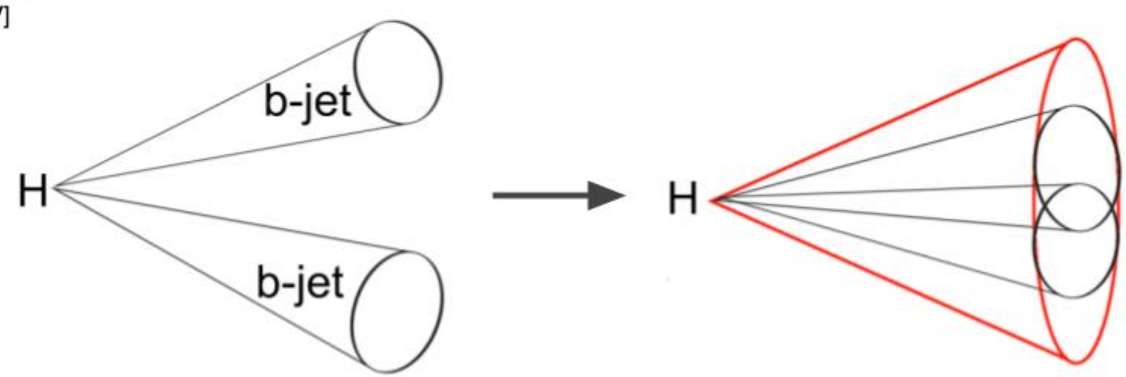
Why large-R jets?



- Effects of BMS more evident at high energy.
- High energy regime → production of particle with high transverse momenta (p_T) much larger than their mass
- Boosted particles → decay products highly collimated → need large-R jet to reconstruct the event

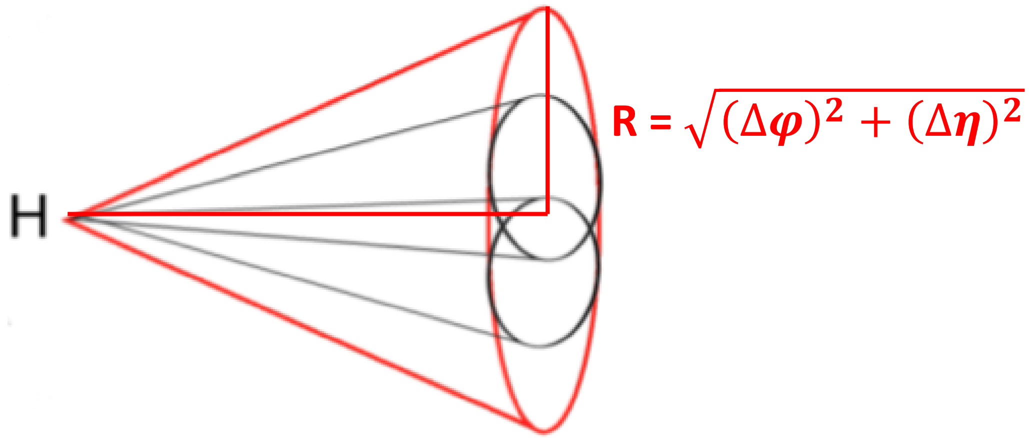
$$R = \sqrt{(\Delta\varphi)^2 + (\Delta\eta)^2}$$

High energy regime



Higgs boson reconstruction

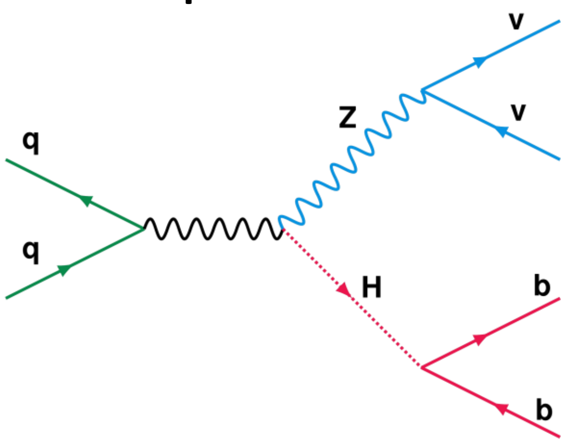
- **Higgs candidate** reconstructed with a **large-R** (R=1.0) jet



- To reconstruct the **Higgs decay products ($b\bar{b}$)**
  - select 2 small-R leading jets inside the large-R jet
  - apply b-tagging algorithm

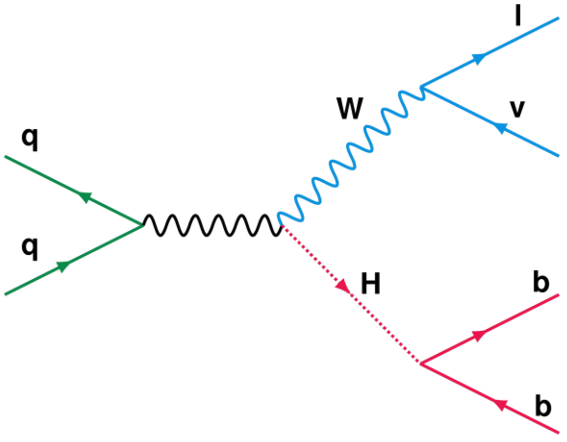
Event selection

0 lepton



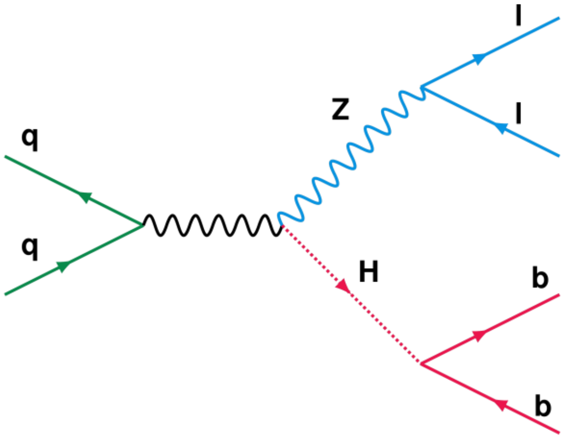
- $E_T^{miss} > 250$ GeV
- 0 charged leptons
- Angular cuts for multi-jet background rejection

1 lepton



- $p_T^W > 250$ GeV
- 1 charged lepton

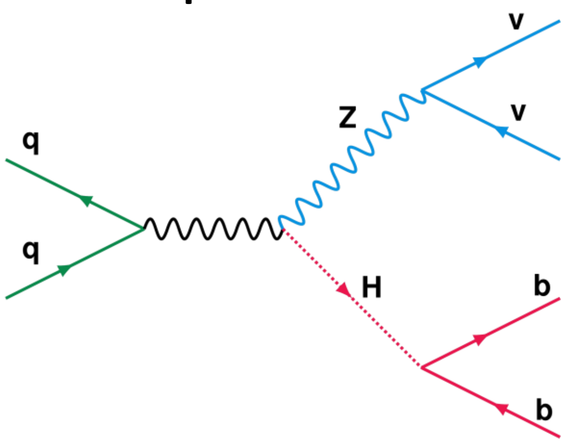
2 lepton



- $p_T^Z > 250$ GeV,
- 2 charged leptons
- Z boson mass:
  61 GeV $< m_{ll} <$ 116 GeV

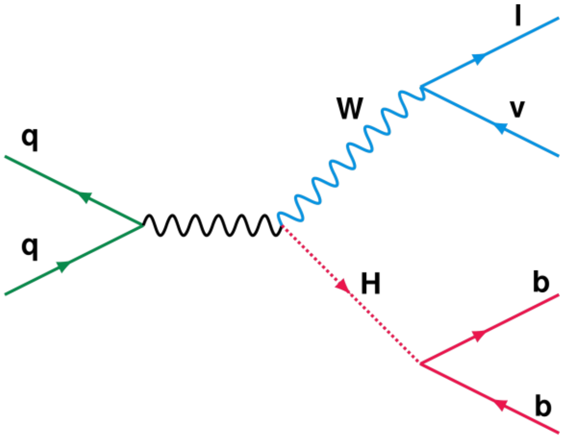
Event selection

0 lepton



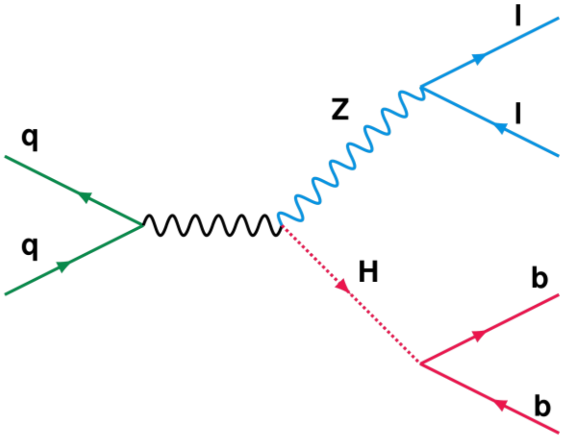
- $E_T^{miss} > 250$ GeV
- 0 charged leptons
- Angular cuts for multi-jet background rejection

1 lepton



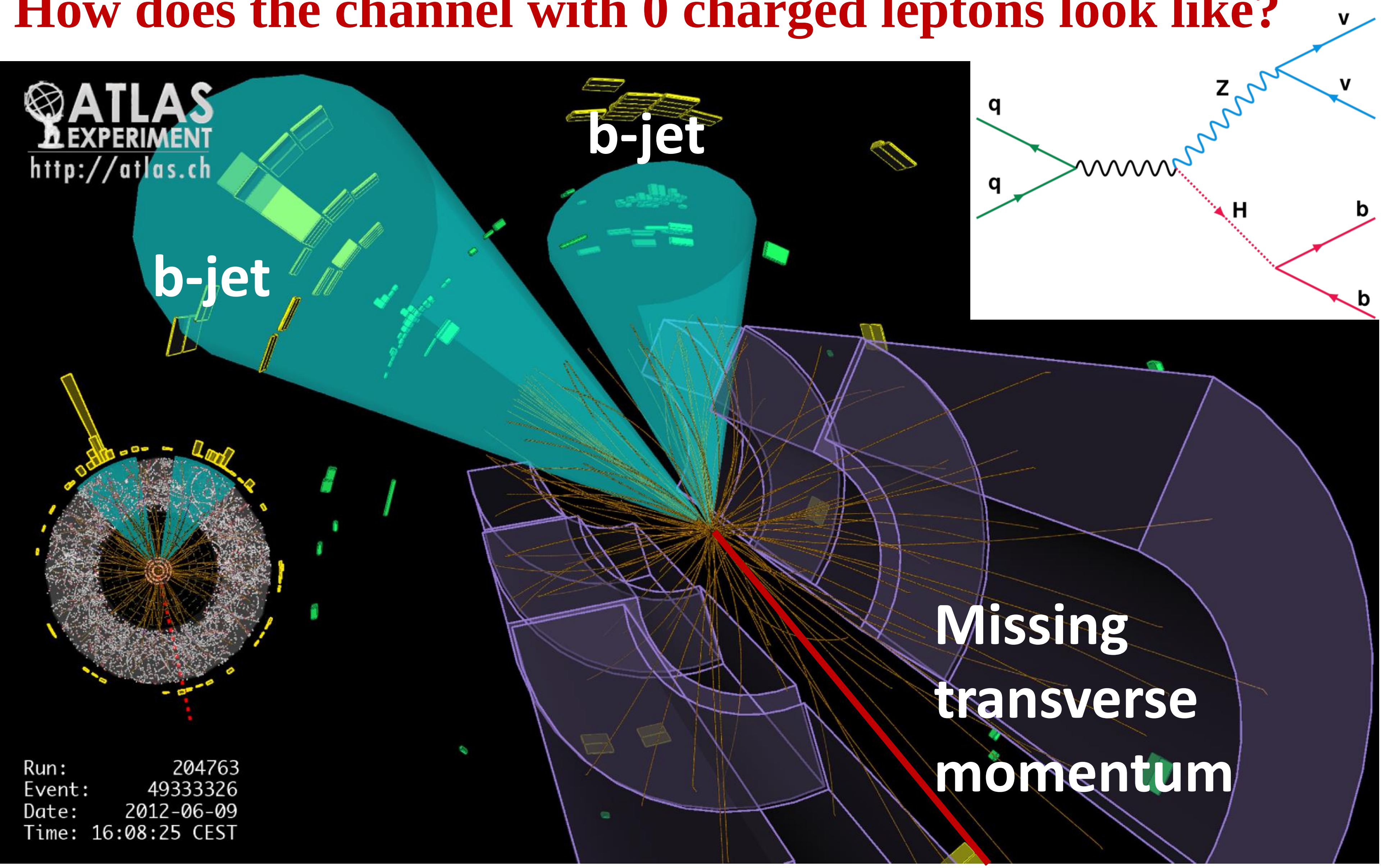
- $p_T^W > 250$ GeV
- 1 charged lepton

2 lepton

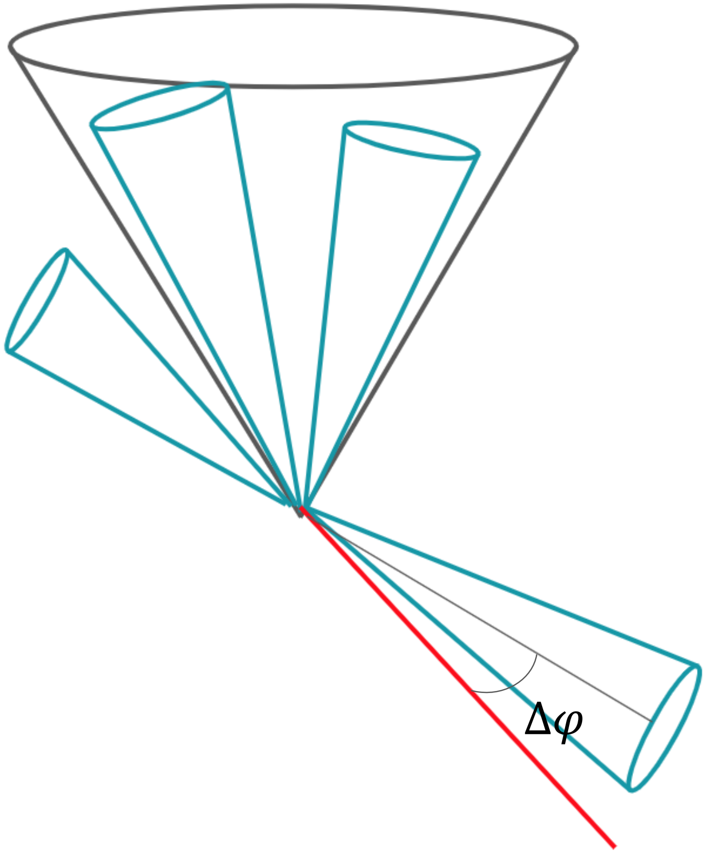


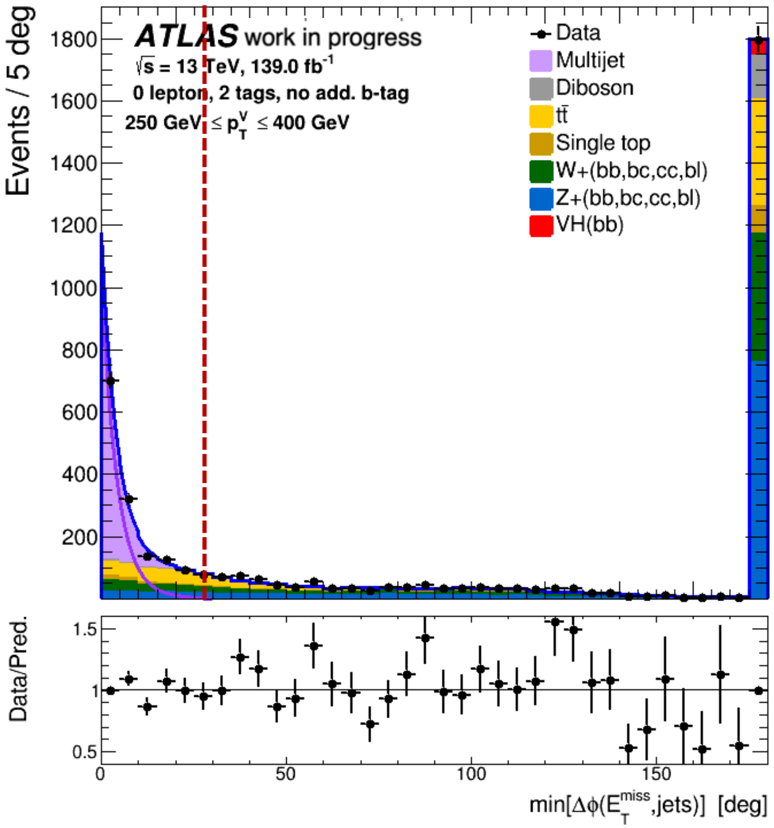
- $p_T^Z > 250$ GeV,
- 2 charged leptons
- Z boson mass:
  61 GeV < m_{ll} < 116 GeV

How does the channel with 0 charged leptons look like?



Multi-jet (MJ) background in 0L

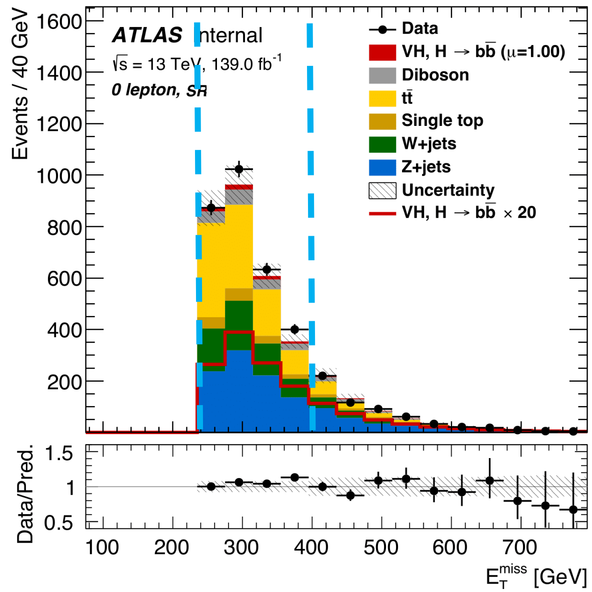


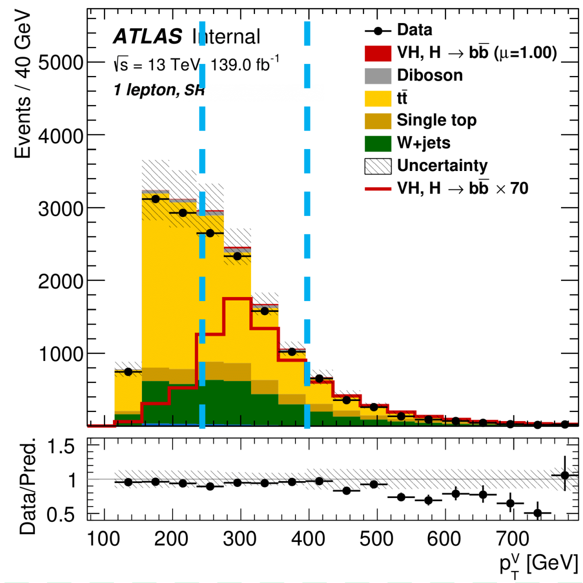


- MJ bkg estimated with a **data-driven method**
  - **MJ bkg** fitted with an exponential decay (**violet line**)
- $\min[\Delta\varphi(E_T^{miss}\ ,\ \text{jets})] > 30°$
  - Remaining MJ fraction negligible
  - 7% signal loss

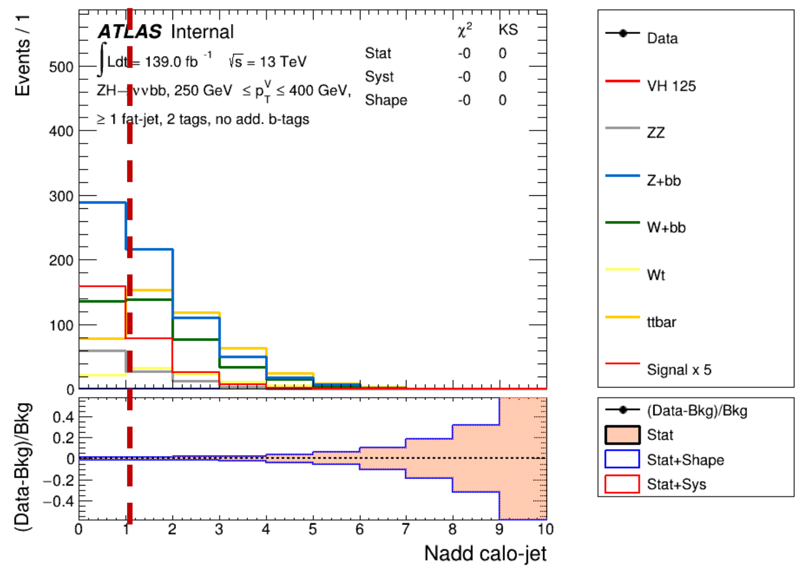
How to improve the analysis?

- Split events in **2 bins in ptv**:
  - 250 GeV < p_T^V < 400 GeV;
  - p_T^V > 400 GeV





- Split events in **number of additional jets**
  - 0 add. jets
  - 1 or more add. jets



Applying SR splitting:

- **16% gain in 0L**
- **30% gain in 1L**

Statistical treatment

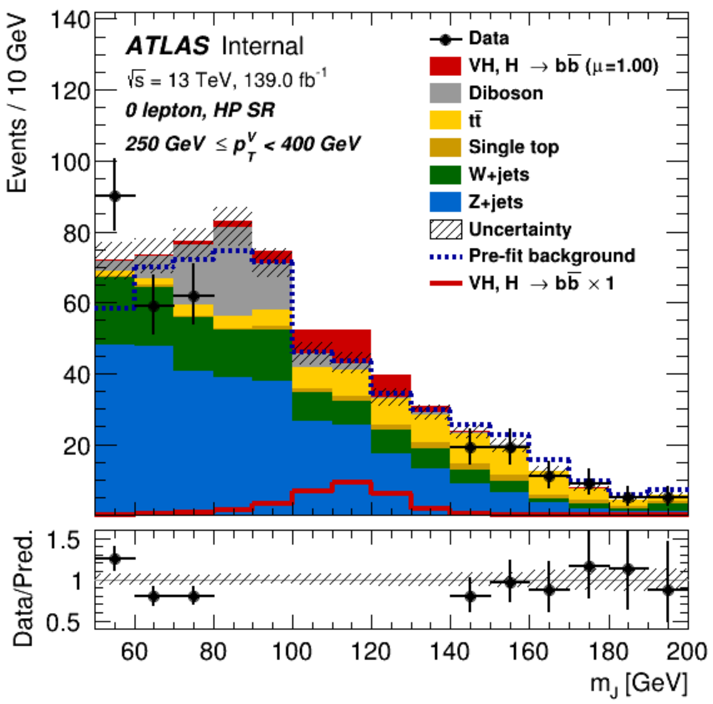
- Invariant mass of the large-R jet constituents (**mJ**) is the **final discriminant**
- Profile likelihood fit is used to extract
  - signal strength μ

$$\mu = \frac{(\sigma \times \mathrm{BR}(H \to b\bar{b}))_{\mathrm{measured}}}{(\sigma \times \mathrm{BR}(H \to b\bar{b}))_{\mathrm{expected}}(\mathrm{SM})}$$

  - main background normalization

- Systematic **uncertainties** (experimental and modelling) described in the likelihood $\mathcal{L}(\mu, \vartheta)$ with **nuisance parameters** ϑ
  - Modelling uncertainties still missing → studies are on going!

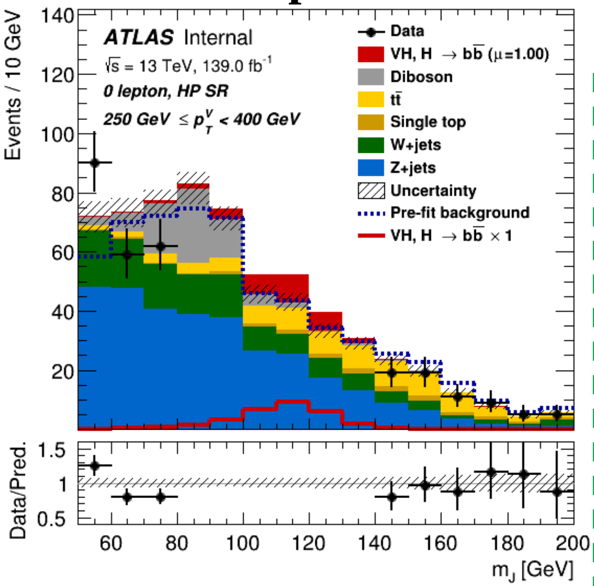
Signal region

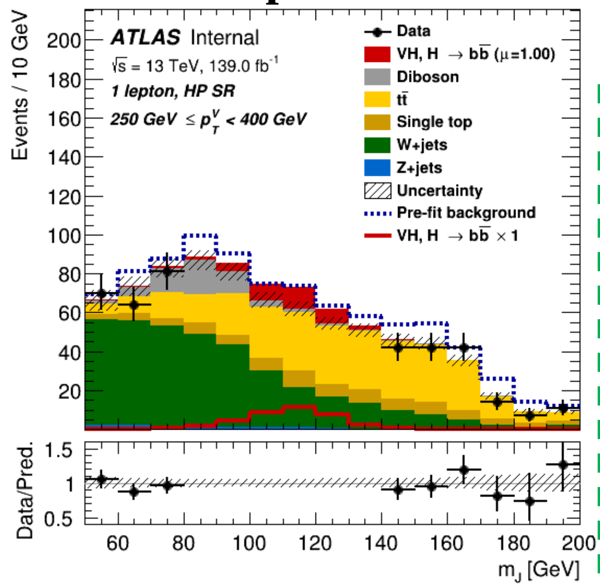


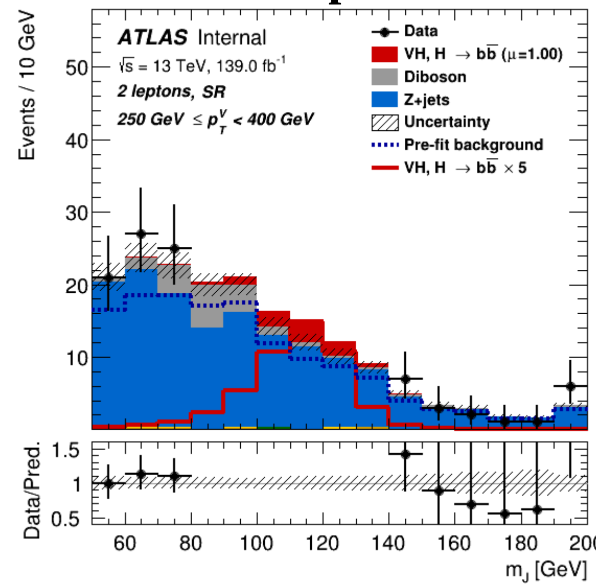
mJ distributions

0-lepton / 1-lepton / 2-lepton

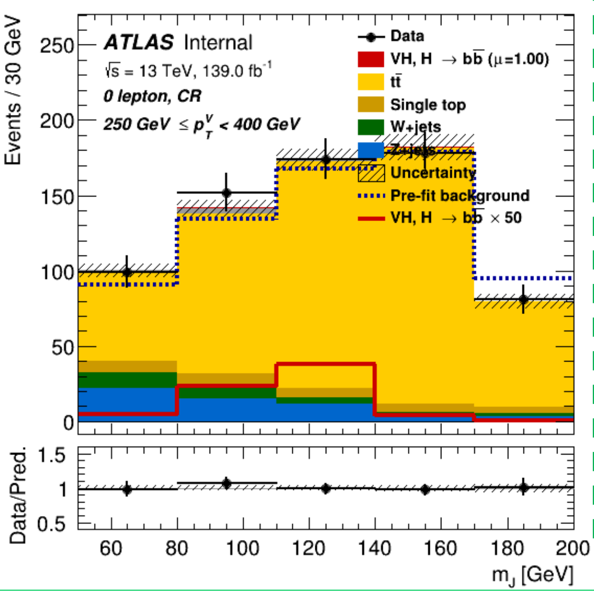
SR

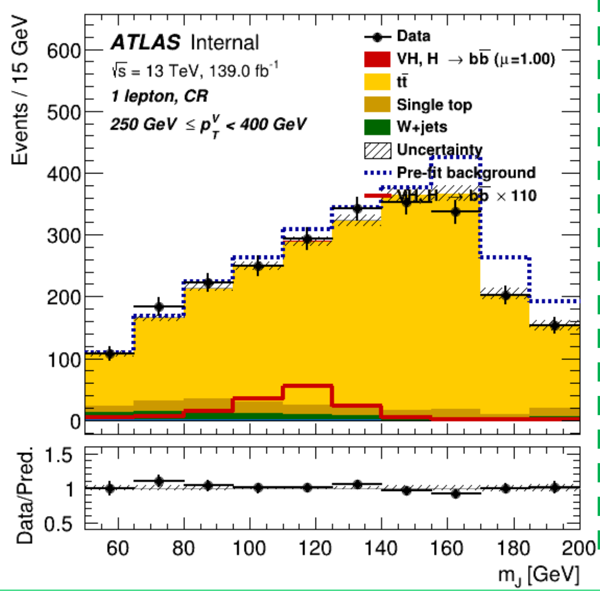






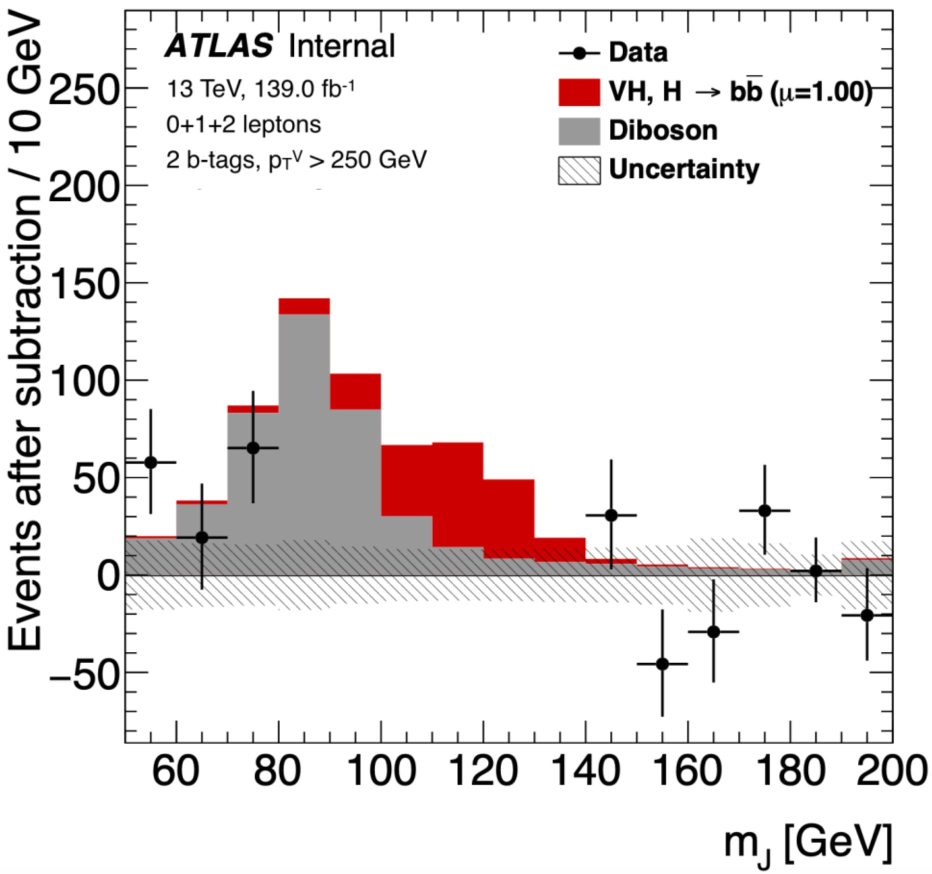
CR





	0L		1L		2L
	# add. small-R jets		# add small-R jets		
	0	>=1	0	>=1	
250-400 GeV	SR	SR	SR	SR	SR
	CR		CR		
>400 GeV	SR	SR	SR	SR	SR
	CR		CR		

Fit results



- Expected significance: **3.09 $\boldsymbol{\sigma}$**

Source of uncertainty	Signed impact	Avg. impact
Total	+0.369 / -0.337	0.353
Statistical	+0.243 / -0.238	0.240
Systematic	+0.277 / -0.239	0.258

- Expected uncertainties on signal strength: $\boldsymbol{\sigma_{\mu}}$~ **0.35**
  - Stat. uncertainties ≈ Sys. uncertainties

Next steps & conclusions

- Conclusions:
  - During Run 2 period, precise measurements on the Higgs properties
  - Personal contributions
    - **Observation** of **VH** production and **$H \rightarrow b\bar{b}$** decay mode (arXiv:1808.08238)
    - First measurement of **differential cross section** pp →VH (arXiv:1903.04618)
    - Search for **Higgs** candidate in **high p_T regime** (work in progress)

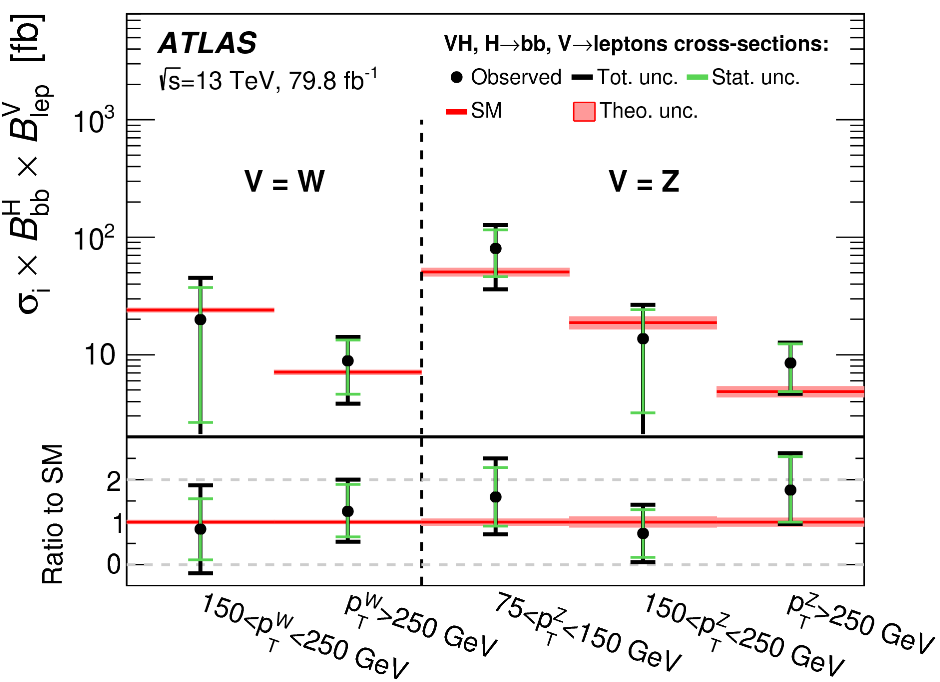
- My next steps in the high p_T regime:
  - Complete **systematics uncertainties** evaluation
  - Corroborate **statistical treatment** of the analysis
  - **Cross sections** measurements in the high p_T region (>400 GeV)
  - Interpret results in terms of **BSM scenario**
  - **Combine results** with previous analysis
  - **Write thesis!**

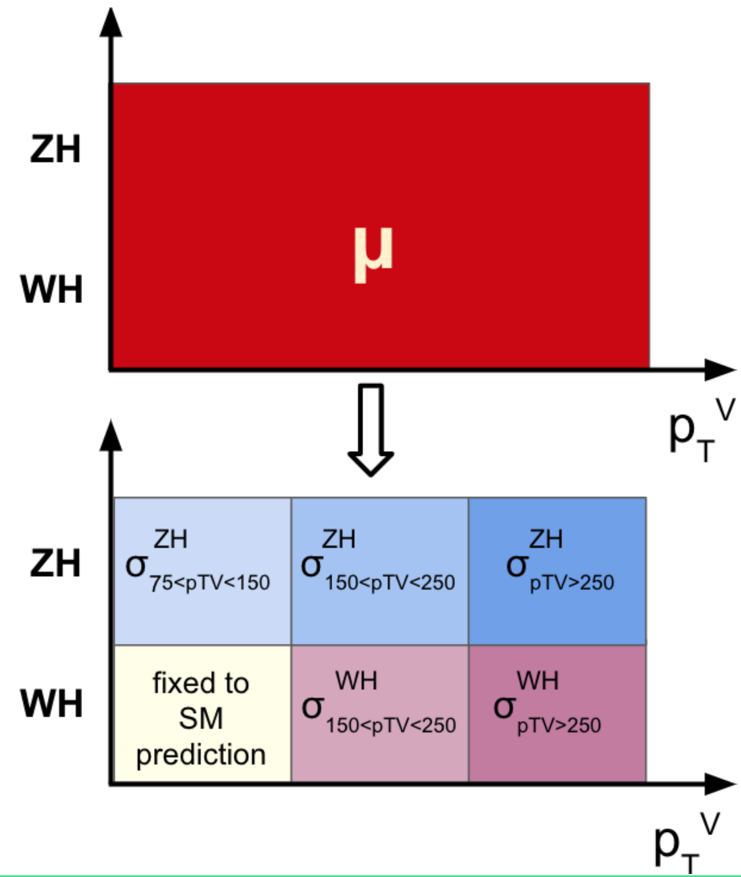
BACKUP SLIDES

Simplified template cross-section

- Same event selection and classification
- Splitted signal:
  - Poduction mode→ ZH o WH
  - p_T^V → cut at 150 GeV and 250 GeV

Differential cross section pp → VH measurement





Single fit results

Fit	Pre-fit Asimov
0L+1L+2L	3.09
0L	2.04
1L	1.94
2L	1.26

Uncertainties

- Two sources of uncertainties:
  - Statistical uncertainties
  - Systematics uncertainties
- Source of systematics:
  - experimental uncertainties
  - modelling of simulated data
    - Still missed in the fit results
    - Studies on going!
- Acceptance ratio sys introduced in the default fit model to mimic modeling uncertainties.

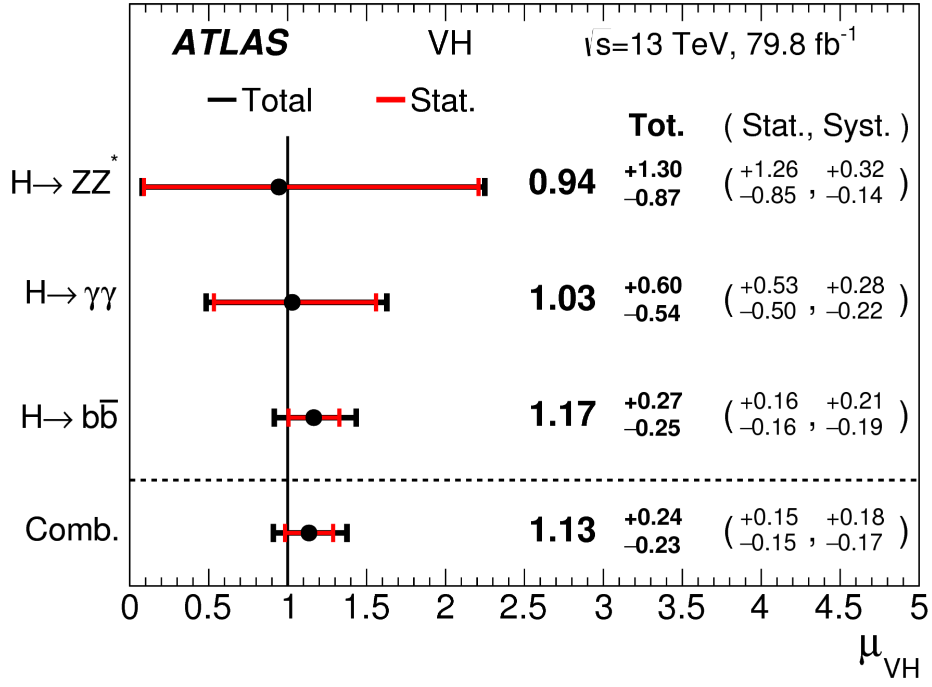
Systematics uncertainties

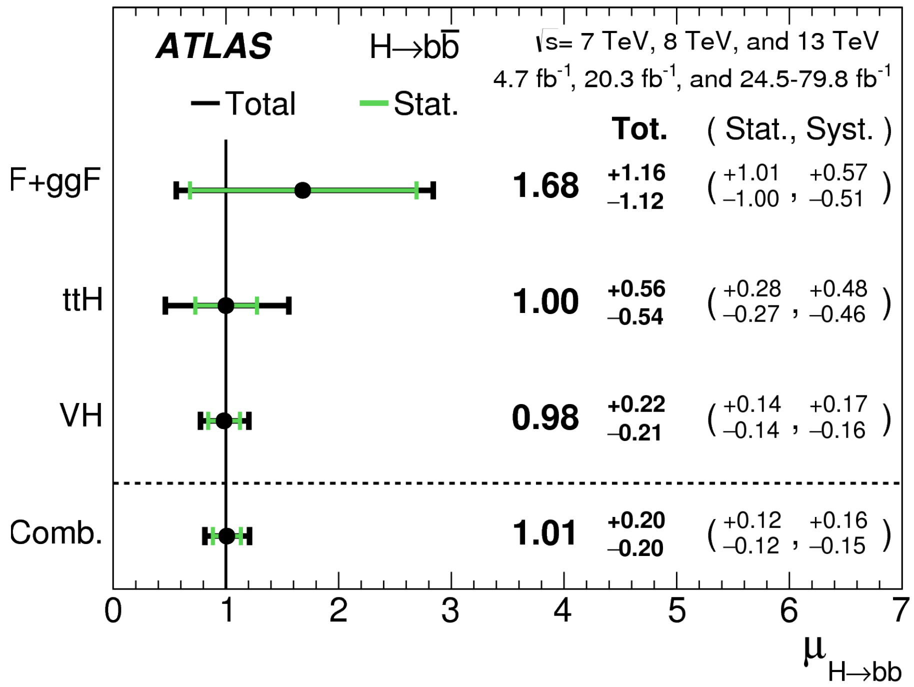
Source of uncertainty		Signed impact	Avg. impact
Total		+0.369 / -0.337	0.353
Statistical		+0.243 / -0.238	0.240
Systematic		+0.277 / -0.239	0.258
Experimental uncertainties			
small-R jets		+0.055 / -0.049	0.052
large-R jets		+0.117 / -0.087	0.102
$E_{\mathrm{T}}^{\mathrm{miss}}$		+0.007 / -0.007	0.007
Leptons		+0.023 / -0.016	0.020
b-tagging	b-jets	+0.059 / -0.026	0.042
	c-jets	+0.005 / -0.004	0.004
	light-flavour jets	+0.007 / -0.007	0.007
	extrapolation	+0.009 / -0.009	0.009
Pile-up		+0.009 / -0.006	0.008
Luminosity		+0.020 / -0.009	0.015
Acceptance ratios			
SR to CR		+0.063 / -0.041	0.052
HP to LP		+0.098 / -0.060	0.079
med. to high p_T^V		+0.032 / -0.030	0.031
Theoretical and modelling uncertainties			
Signal		+0.073 / -0.037	0.055
Floating normalisations		+0.099 / -0.090	0.094
Z + jets		+0.010 / -0.010	0.010
W + jets		+0.011 / -0.011	0.011
$t\bar{t}$		+0.001 / -0.001	0.001
Single top quark		+0.003 / -0.003	0.003
Diboson		+0.038 / -0.030	0.034
Multijet		+0.000 / -0.000	0.000
MC statistical		+0.102 / -0.101	0.102

Dominant uncertaintes:

- Large-R jets unc.
- MC stat.

Observation VH e $H \to b\bar{b}$





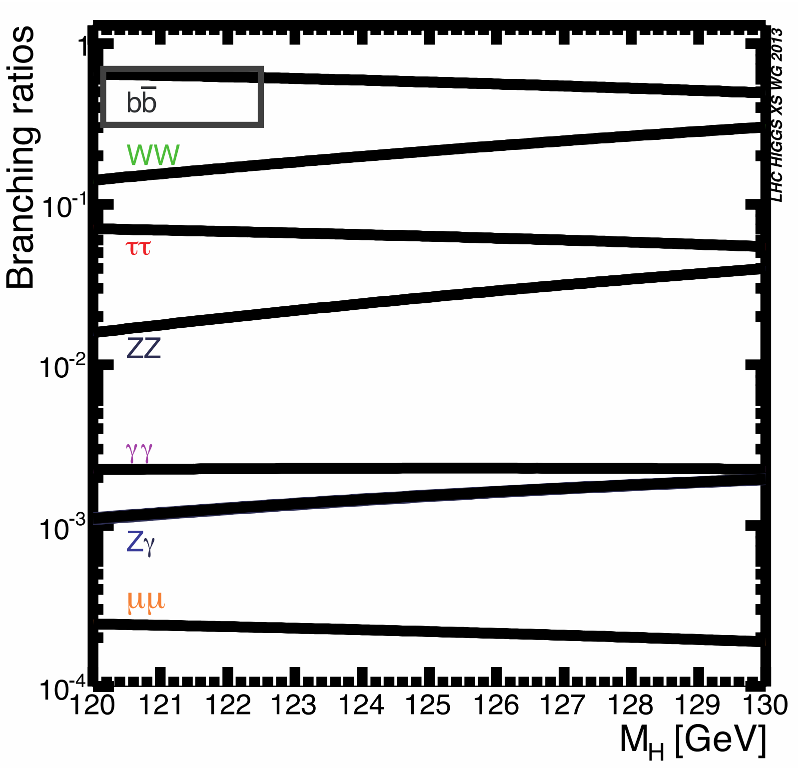
Combination with other decay modes.
Observed significance: **5.3 σ**

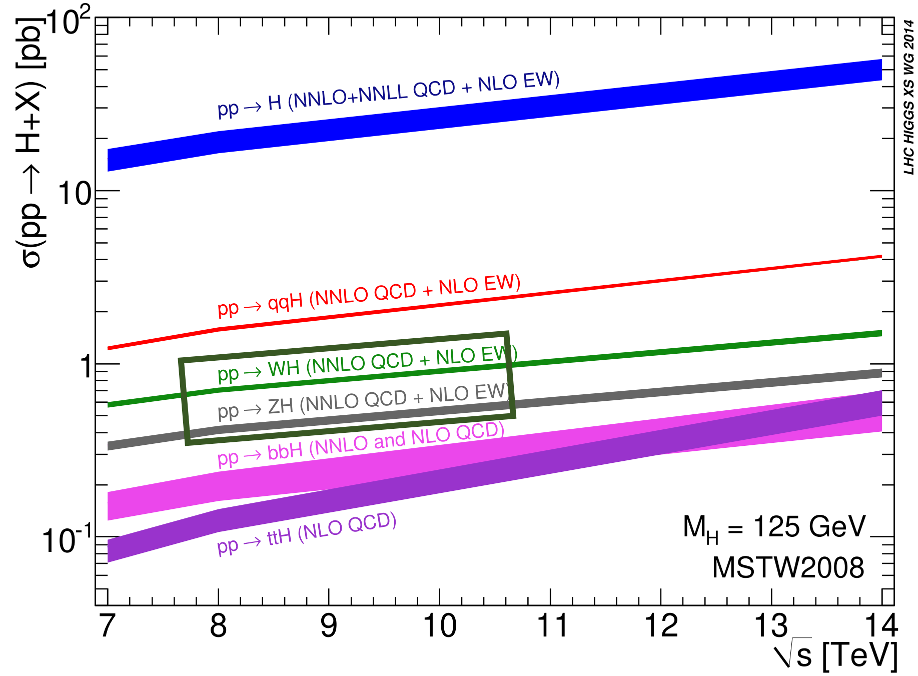
Combination with other production mode.
Observed significance: **5.4 σ**

First **observation** of **VH** production mode

First **observation** of **$H \to b\bar{b}$** decay mode

Why VH, H→$b\bar{b}$ channel?

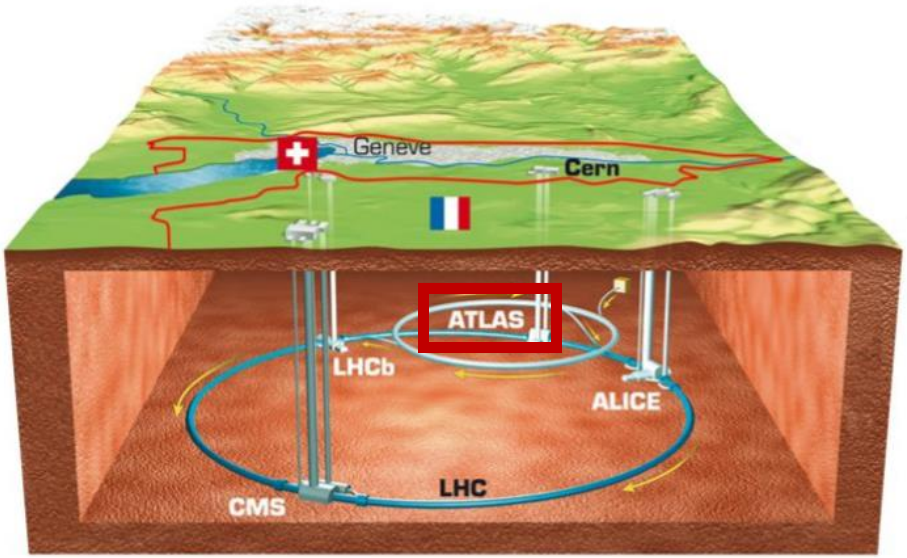


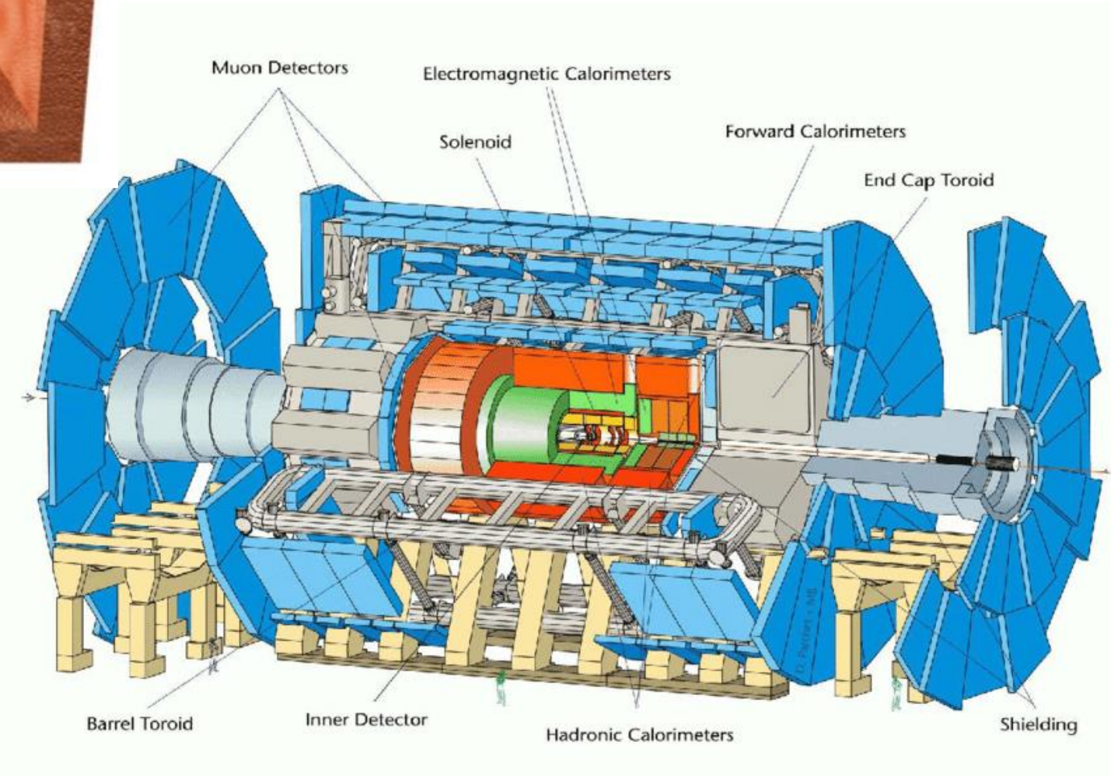


- BR (H→$b\bar{b}$) ~ 58%
  - **<u>Caveat</u>**: Multi-jet background
- First direct evidence of **Higgs coupling with quarks**

- Associated production with a vector boson V (V= Z o V=W),
  - V leptonic decay → clear signature!

ATLAS detector

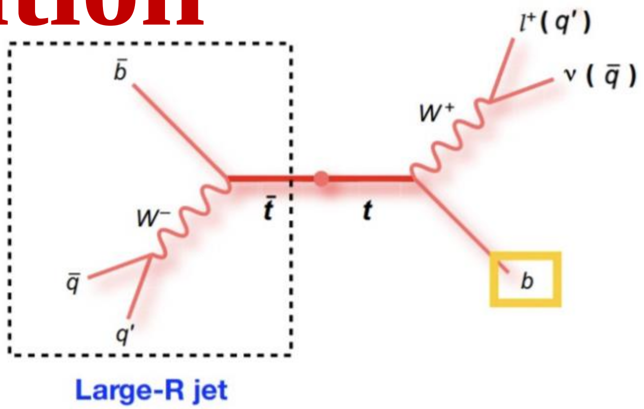




ATLAS is formed by

- **Inner Detector**
- **Electromagnetic calorimeter**
- **Hadronic calorimeter**
- **Muon spectrometer**

Event categorization



- Events splitted in **2 bin in ptv**:
  - 250 GeV < p_T^V < 400 GeV;
  - p_T^V > 400 GeV

- In 0- and 1-lepton channel:
  - **SR:** 0 *b*-tagged VR track jets <u>not</u> matched to the large-R jet
  - **top CR:** ≥ 1 *b*-tagged VR track jets <u>not</u> matched to the large-R jet

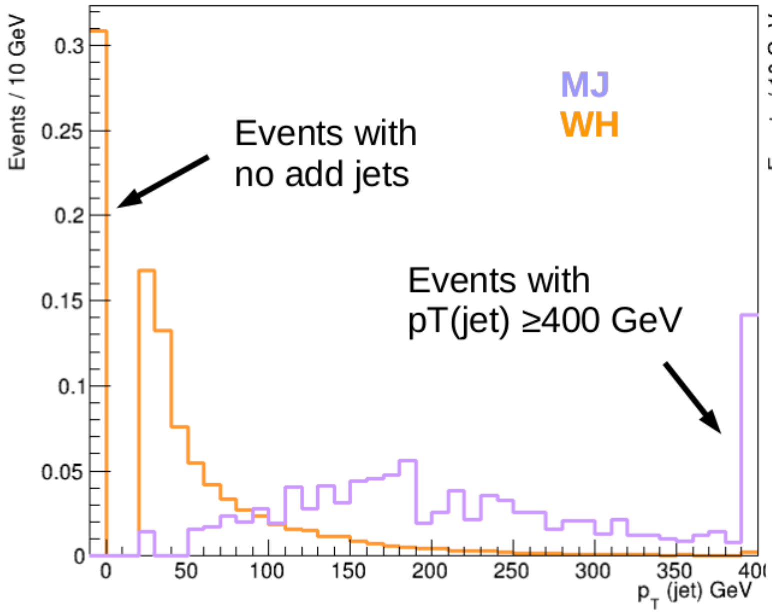
- In 0- and 1-lepton, SR splitted in:
  - **High purity SR**: 0 add calo small-R add. jet with p_T > 30 GeV
  - **Low purity SR**: ≥ 1 add calo small-R add. jet with p_T > 30 GeV

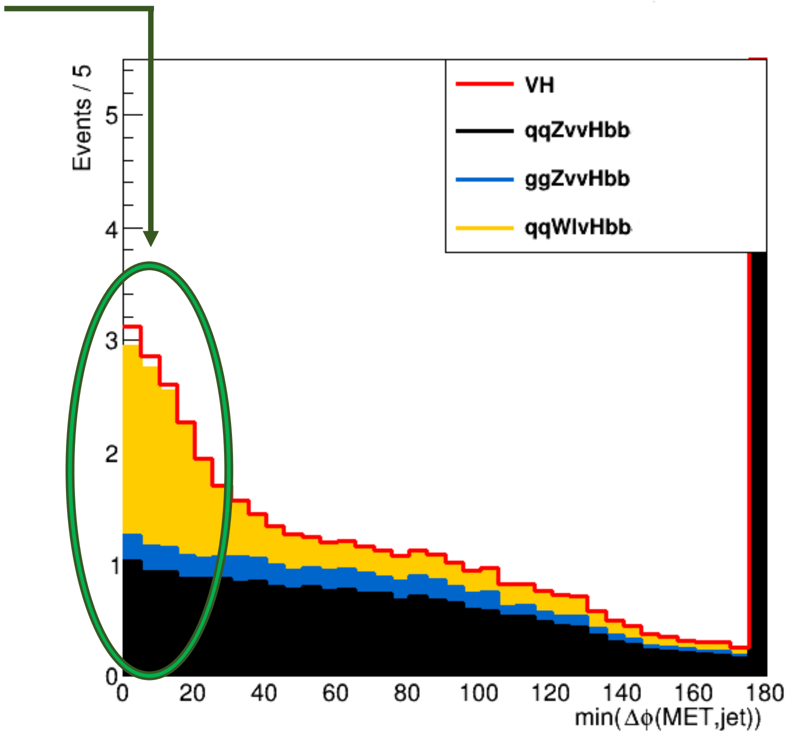
10 SRs e 4 CRs

<table>
<tr><td rowspan="3"></td><td colspan="2">0L</td><td colspan="2">1L</td><td rowspan="3">2L</td></tr>
<tr><td colspan="2"># add. small-R jets</td><td colspan="2"># add small-R jets</td></tr>
<tr><td>0</td><td>>=1</td><td>0</td><td>>=1</td></tr>
<tr><td rowspan="2">250-400 GeV</td><td>SR</td><td>SR</td><td>SR</td><td>SR</td><td rowspan="2">SR</td></tr>
<tr><td colspan="2">CR</td><td colspan="2">CR</td></tr>
<tr><td rowspan="2">>400 GeV</td><td>SR</td><td>SR</td><td>SR</td><td>SR</td><td rowspan="2">SR</td></tr>
<tr><td colspan="2">CR</td><td colspan="2">CR</td></tr>
</table>

Multi-jet (MJ) background in 0L

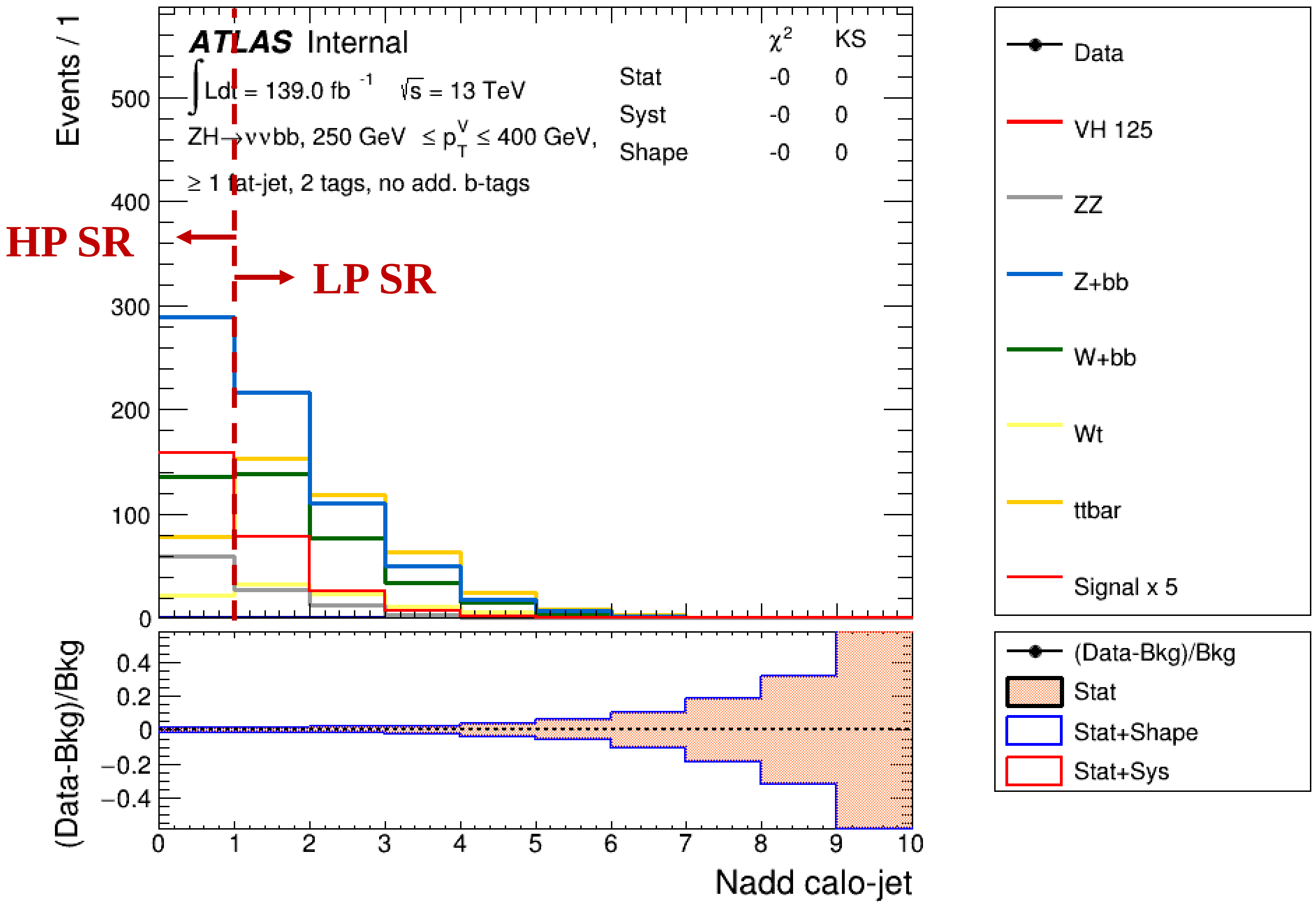
- MJ bkgs enters due to the mis-measurements of the energy of a jet → E_T^{miss} and mis-measured jets tends to be aligned
- $\min[\Delta\varphi(E_T^{miss}$, small-R jets$)] > 30°$
  - BUT 22% signal loss due to WH sample (tau misidentification)
- pT(jet) can be used to discriminate WH and MJ events





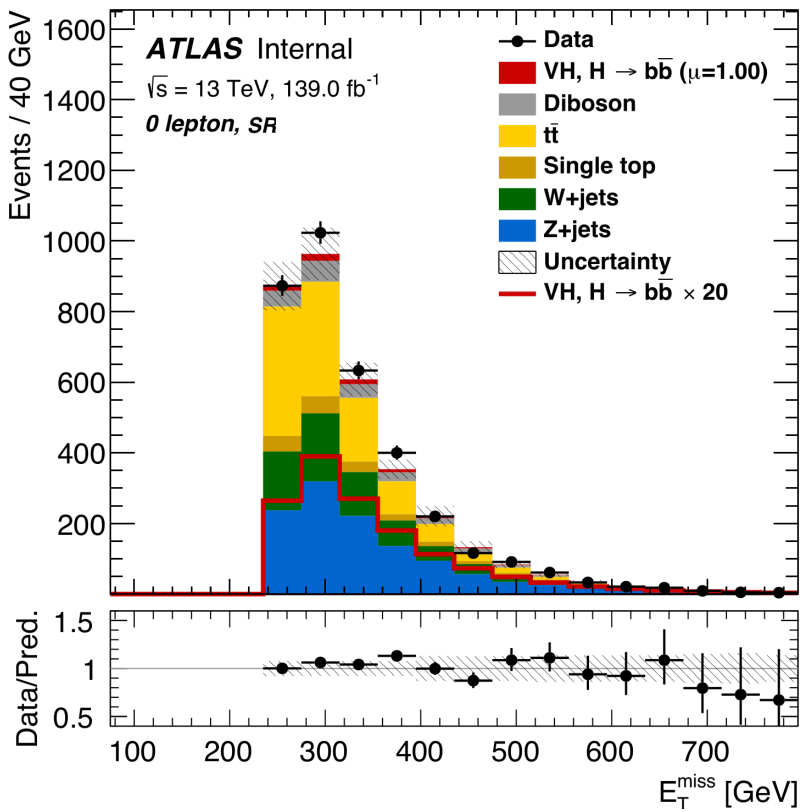
SR splitting

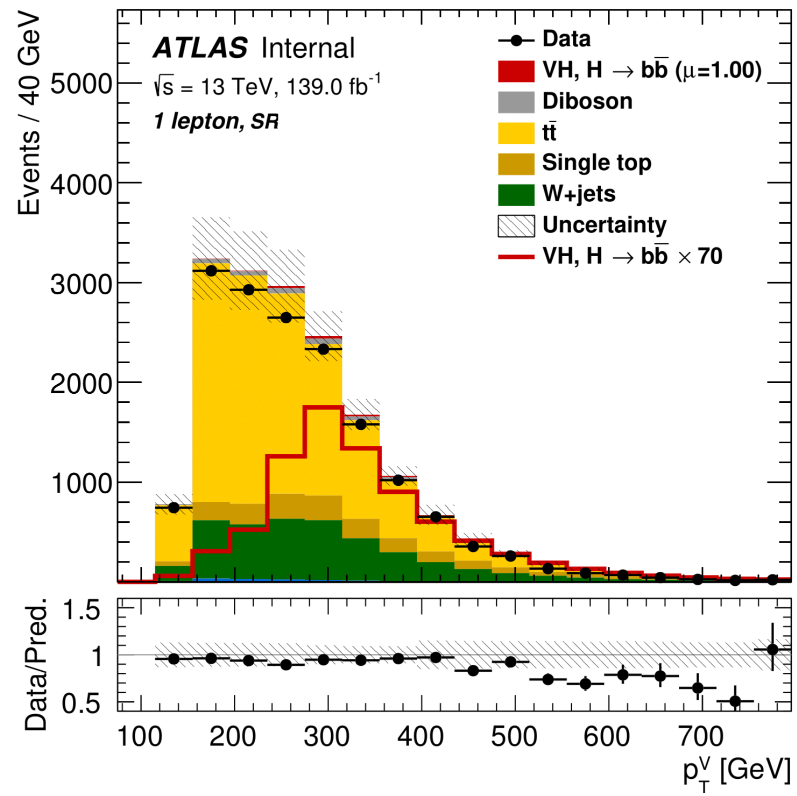
- Top process one of the main bkgs
  - Characterized by more jets than signal events



Applying SR splitting:
- **16% gain in 0L**
- **30% gain in 1L**

Main background



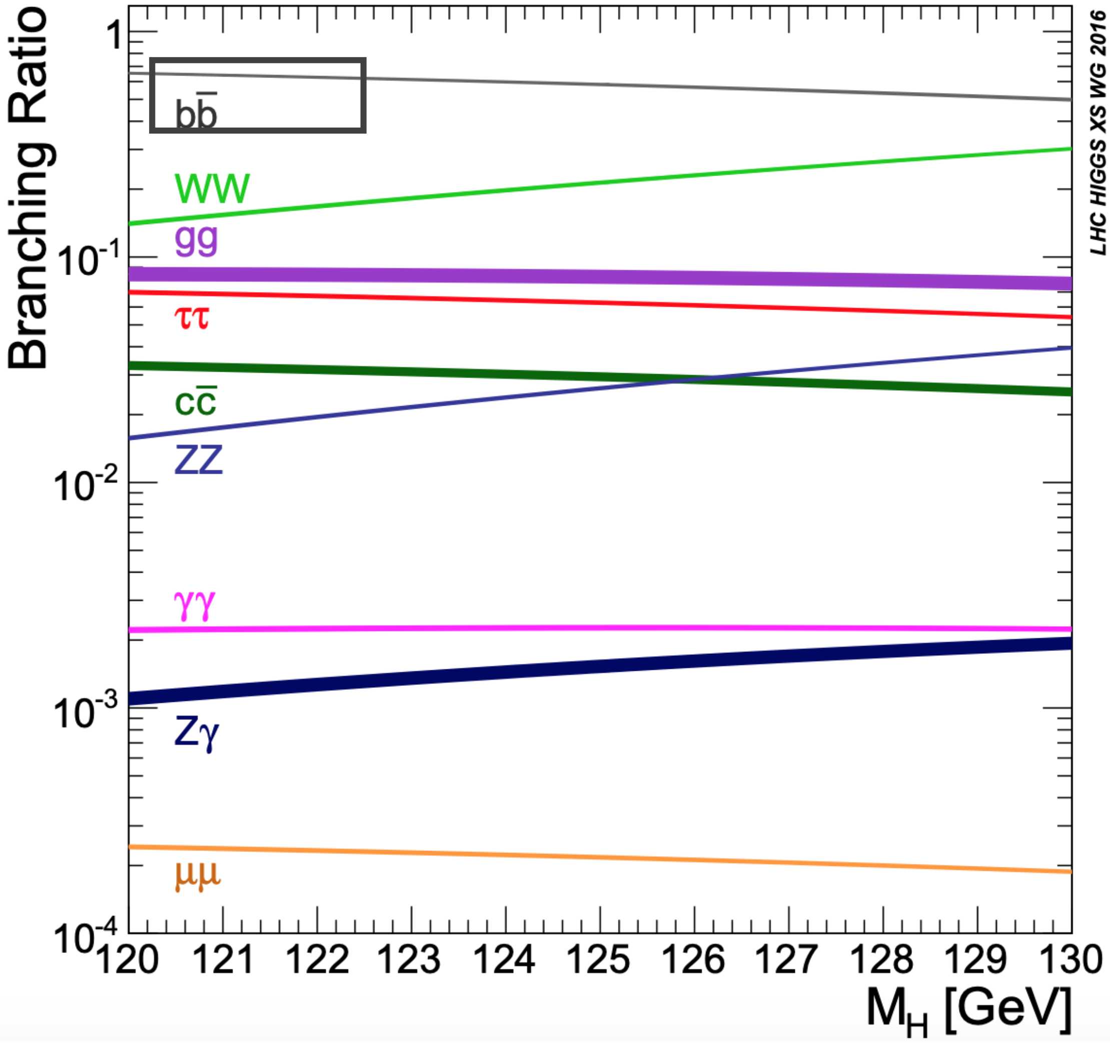


- **Z+jets** e **W+jets** → estimated from simulation
- Top (**ttbar** e **single-top**) → dedicated control regions
- Diboson (**WZ, ZZ**) → final state similar to VH, estimated from simulation
- **Multi-jet** → suppress with angular cuts, estimated with data-driven methods

Why is the H→$b\bar{b}$ channel interesting?

- To establish the **fate of the Higgs** boson due to the expected BR.
- To control the **Higgs Yukawa sector**
- To **extract the total width** (not directly measurable at the LHC)
  - Only **ratio of BR** (couplings) are truly **model independent** at the LHC
  - **Absolute coupling** measurement requires **assumptions** on the **total width** → a term accounting for 58% of the total has a dominant effect on all the coupling determination

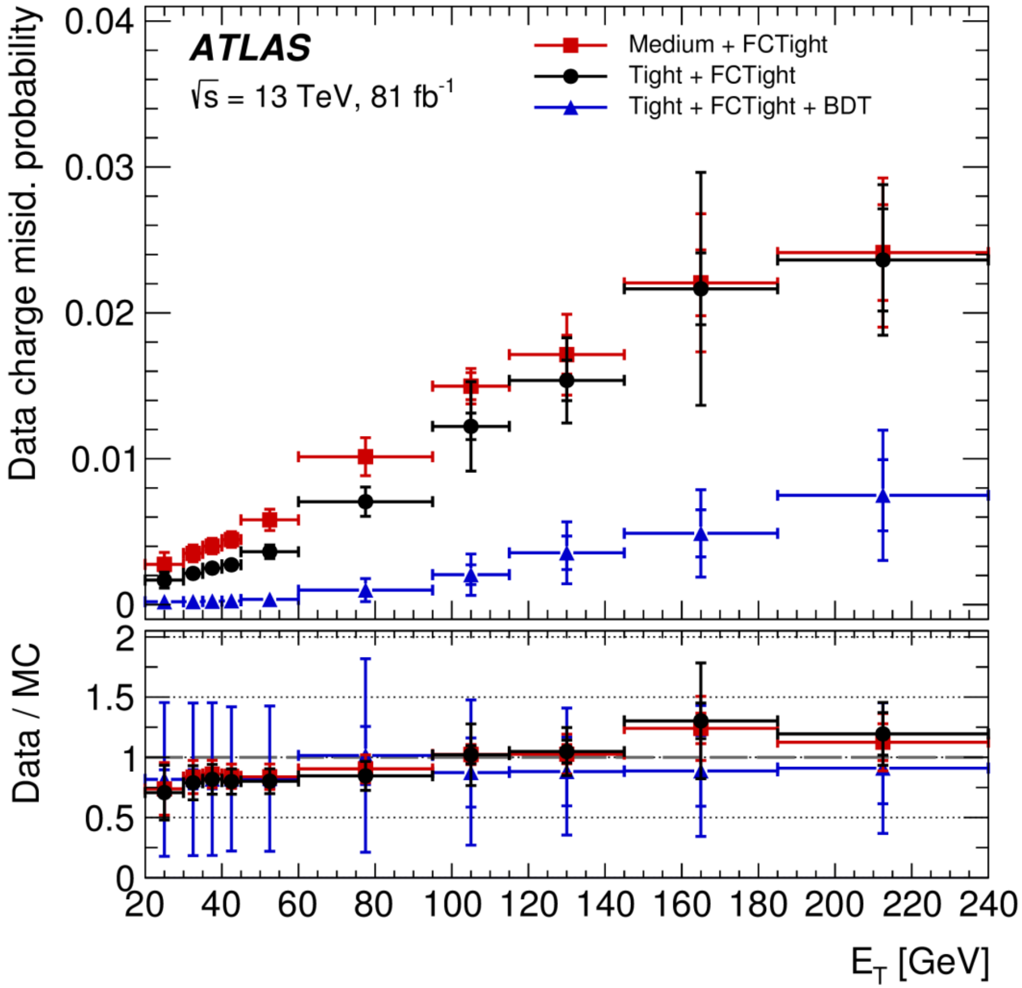
Higgs decay modes



- BR ($H \rightarrow b\bar{b}$) ~ 58%

- Why it took so long to find the largest Higgs boson decay mode?

Charge misidentification probability



P(mis. id.) ~ 3-4%

Efficiency = 0.97*0.97 = 0.94
→ 6% of signal loss