Evidence for the dimuon decay of the Higgs boson with the ATLAS detector

Ehsan Mushajiang
on behalf of ATLAS Collaboration

Moriond EW 2026
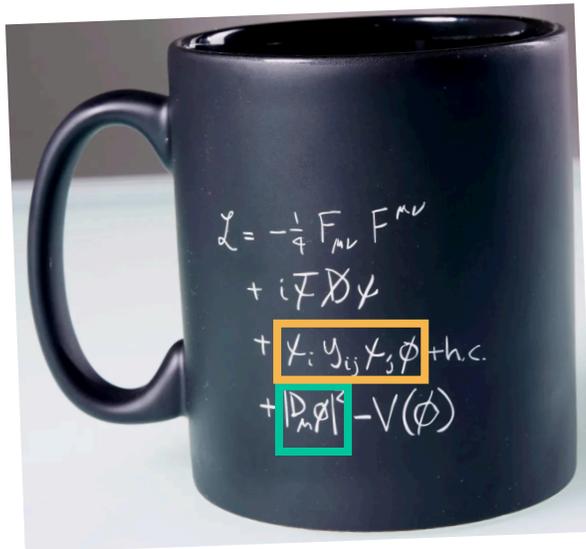
15-22 March 2026, La Thuile

- In the SM, W/Z bosons acquire mass through the (Brout - Englert - Higgs) **BEH mechanism** responsible for the Electroweak Symmetry Breaking.

- Fermions, acquire mass through **Yukawa interactions** with the Higgs boson.

❖ The heavier a particle is, the stronger it should be interacting with Higgs boson.
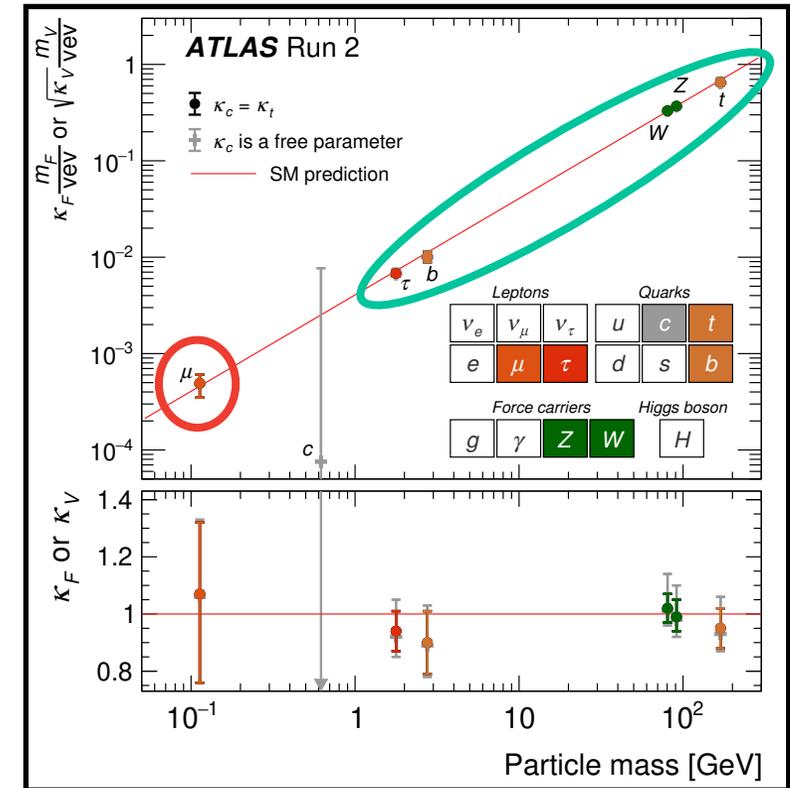
Previously in Run 2:

▸ ATLAS: obs.(exp.) significance $2.0\sigma(1.7\sigma)$.

▸ CMS: obs.(exp.) significance $3.0\sigma(2.5\sigma)$, **evidence!**

- **H→μ+μ-** offers the best opportunity to test the coupling for **2nd generation fermions** at the LHC!

● **Rare process**: branching ratio of (2.17 ± 0.04) × 10⁻⁴

● **Large background** from Drell-Yan μ⁺μ⁻ production
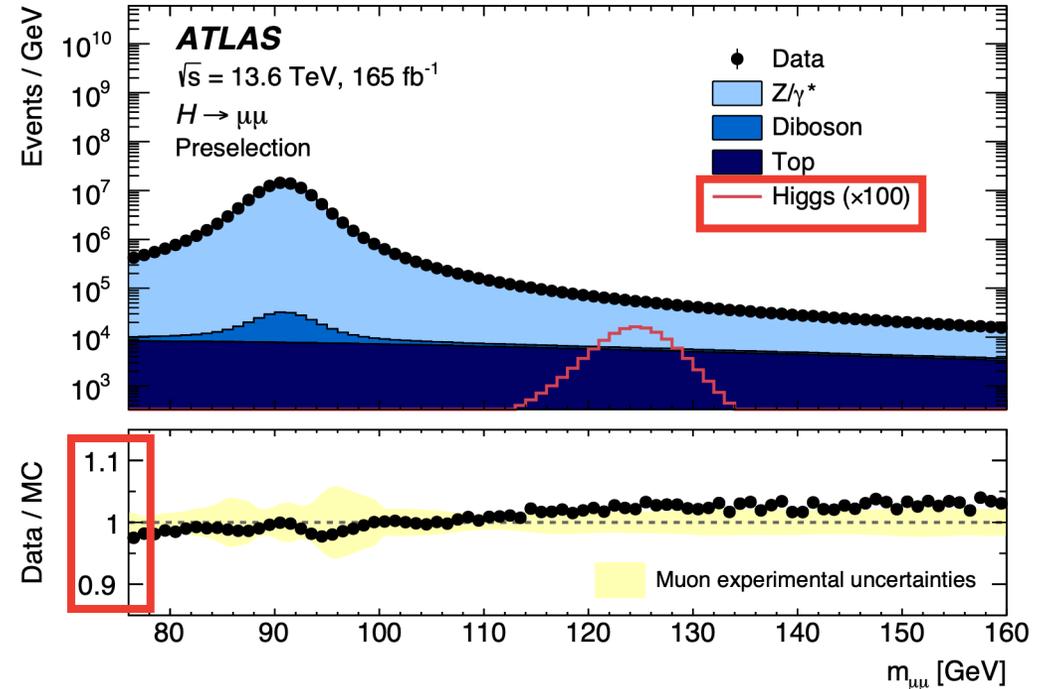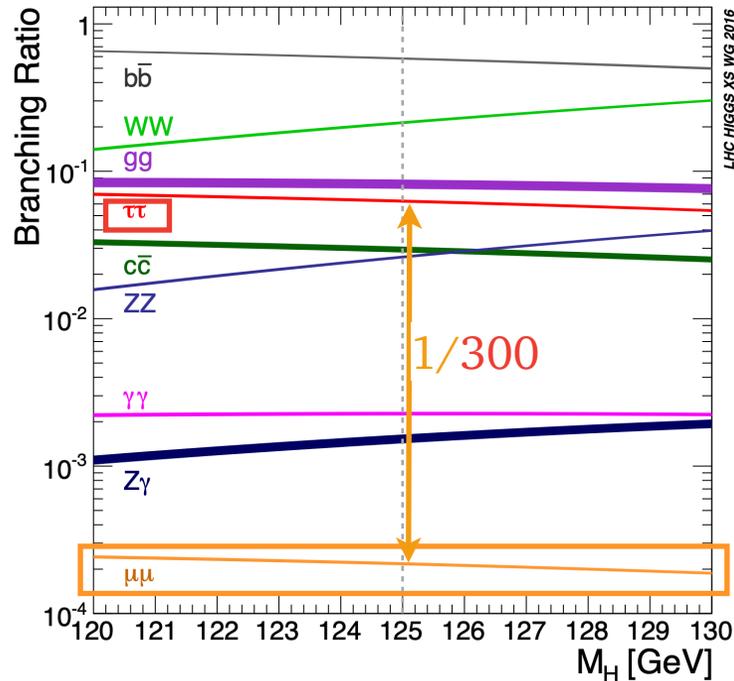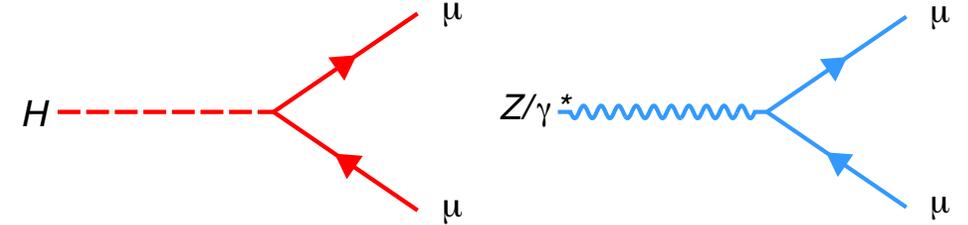
➡ Very small signal-to-background ratio, **typically O(0.1%)**

▸ Per-mille level precision on background modeling is required

Select events based on kinematic and quality requirements

Classification

Signal and background modeling using analytic functions

A profiled likelihood fit to extract signal strength
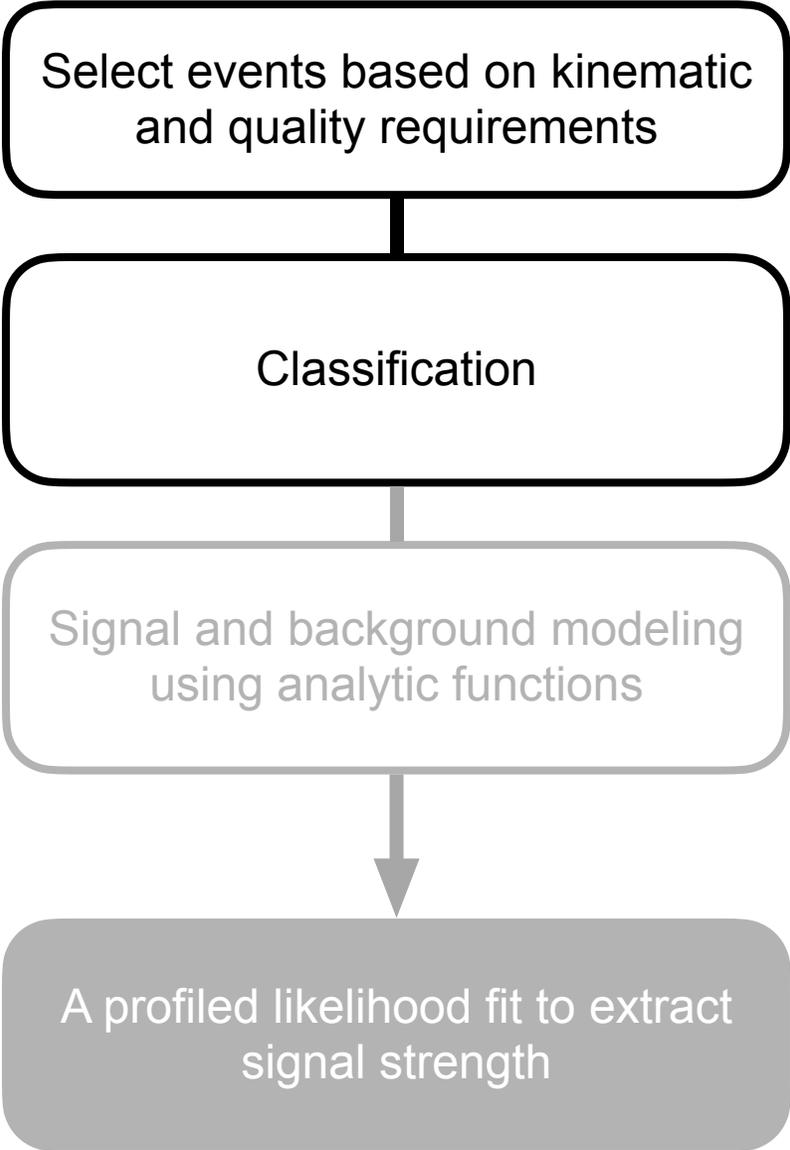
1. Sort events into mutually exclusive categories targeting the various Higgs boson production modes

2. Use Machine learning to enhance the signal-to-background ratio

ATLAS $\sqrt{s}$ = 13.6 TeV, 165 fb^{-1}, $H \rightarrow \mu\mu$

$S/\sqrt{B}$
S/B
B (×10^{-5})

Signal composition

VBF   2Jet   1Jet   0jet   VH   ttH

ggF+b$\bar{b}$H
VBF
VH
t$\bar{t}$H + tH

Background composition

Z/γ^*
Z-VBF
Diboson
Top

Analysis categories
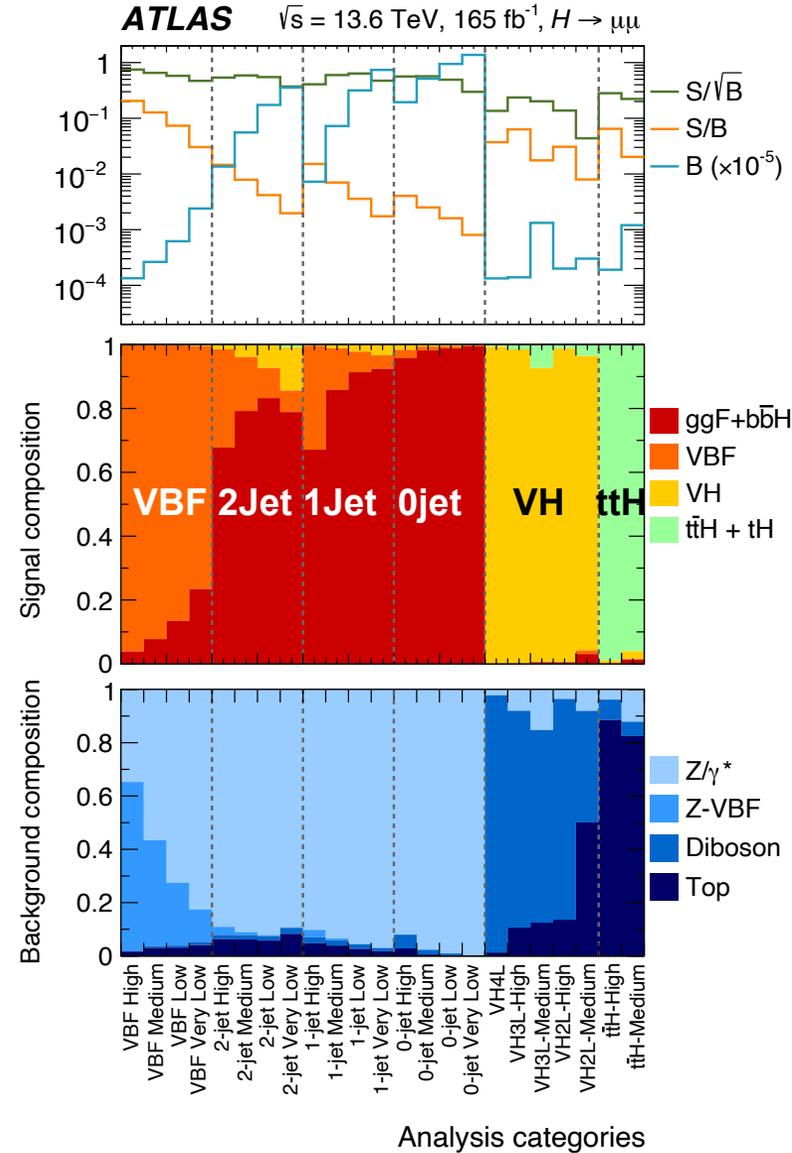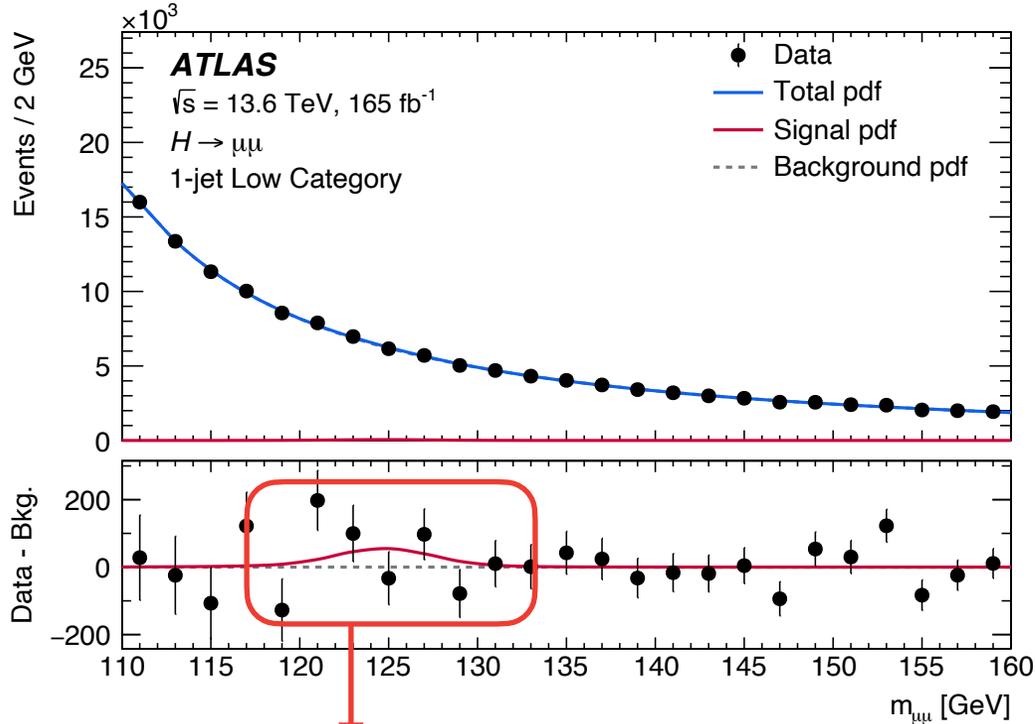
In each category:
    Data fluctuations > Higgs signals

- **Per-mille level precision** necessary to avoid biases in extracted signal yields

- Backgrounds are not modeled with MC templates but with analytic functions

  ‣ MC is not sufficiently good to model data out-of-the-box at the 0.1% level

  ‣ MC used to choose the function for each category

    - **Spurious signal**, bias between selected model and bkg. only template, assigned as background modeling uncertainty

  ‣ Shape parameters decided in the final fit to data

$$DY_{LS}(m_{\mu\mu}) = \sum_q \mathcal{L}_{q\bar{q}} \cdot \sigma_{q\bar{q}} \cdot m_{\mu\mu} \ , \ q = u, \ s, \ d, \ c, \ b$$

- "**core**×**empirical**" background models:

  - **Core function**: Capture the steepness of the DY background

  - **Empirical function**: handle the remaining bias between core function and Data

    ‣ Requires a large number of simulated events to select a function from the candidates

    ‣ Used 5 billion (NLO) Drell−Yan fully simulated sample

ATLAS Simulation
$\sqrt{s}$ = 13.6 TeV
$H \to \mu\mu$
1-jet Low Category

◇ Simulation
— Total pdf
···· Core pdf
— Signal pdf

χ^2/n.d.f. = 0.74
p(χ^2) = 80%

Core function
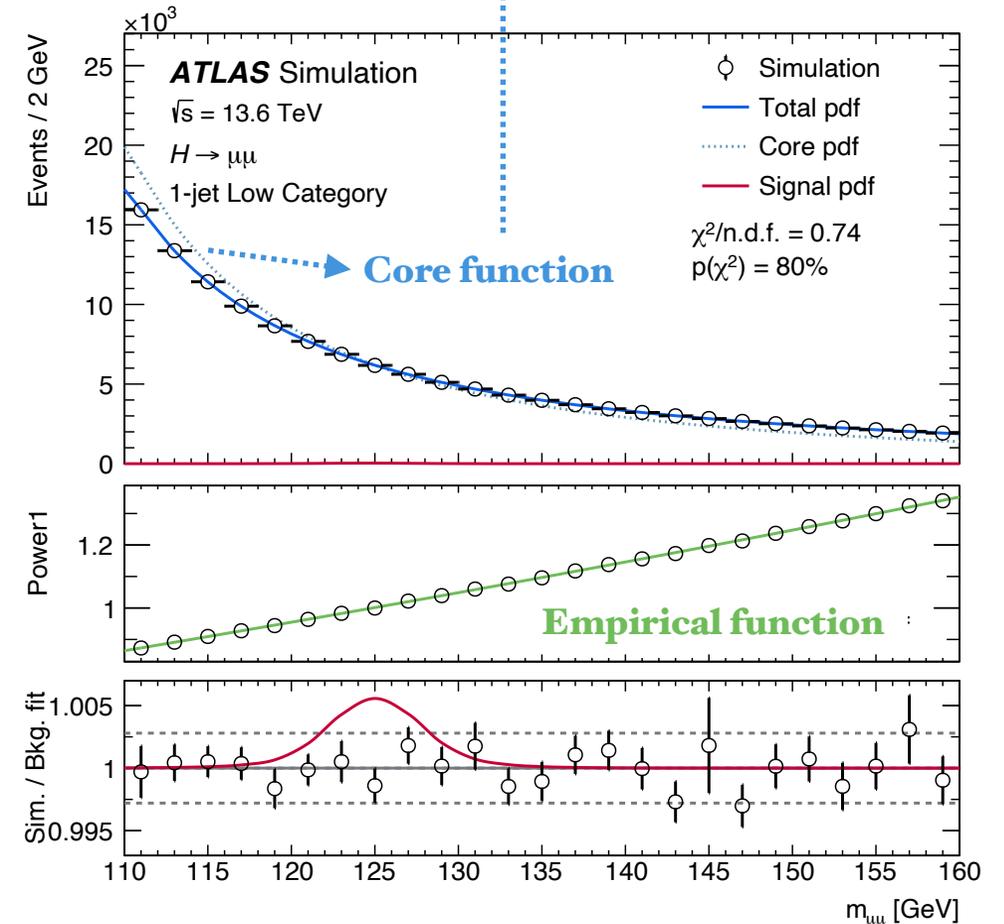
Empirical function

Fit across 43 categories

Background + Signal

Background
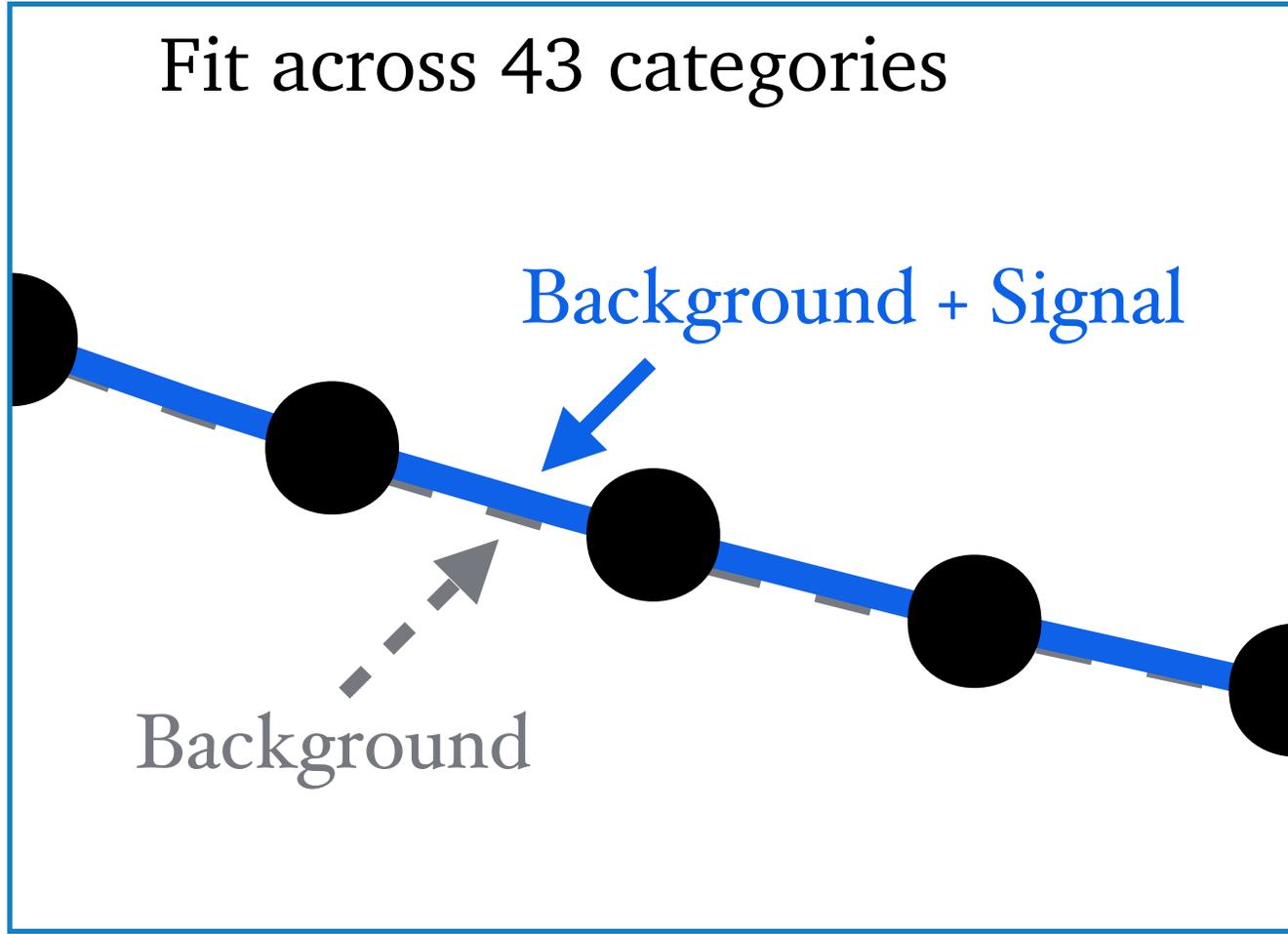
Combining Run 3 with Run 2

ATLAS
$\sqrt{s}$ = 13/13.6 TeV, 140/165 fb^{-1}
$H \rightarrow \mu\mu$
Inclusive, ln(1+S/B) weighted

- Data
- Total pdf
- Signal pdf
- Background pdf

Weighted events / 2 GeV

Data - Bkg. (weighted)

$m_{\mu\mu}$ [GeV]
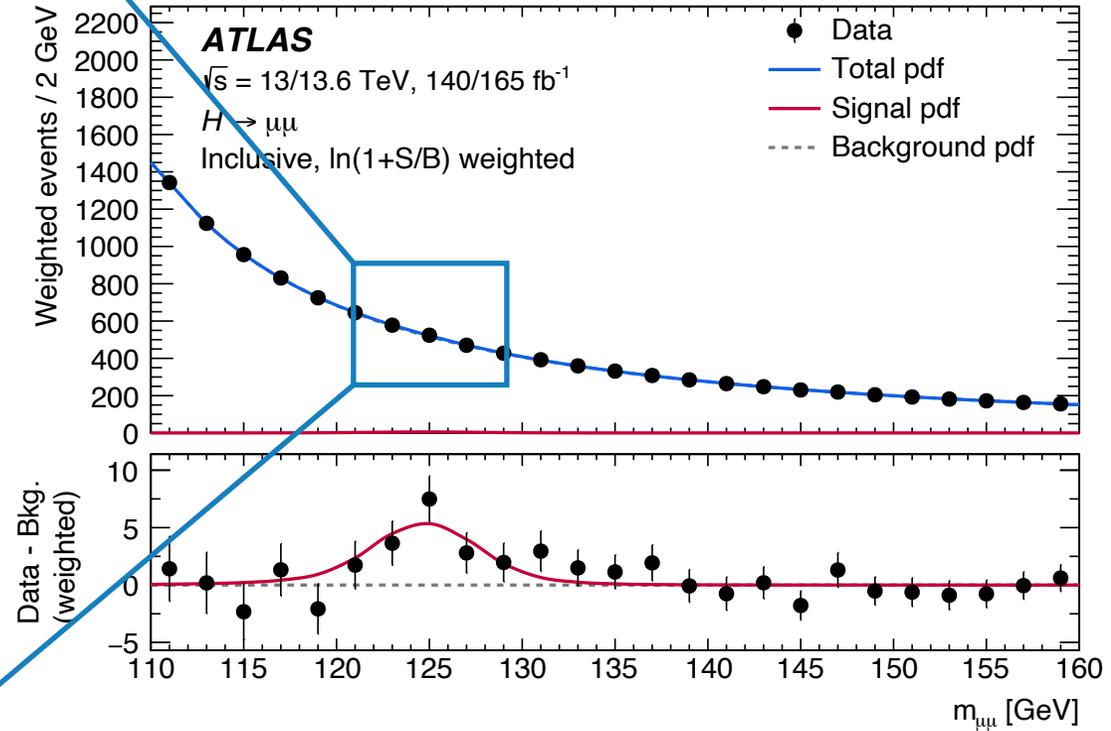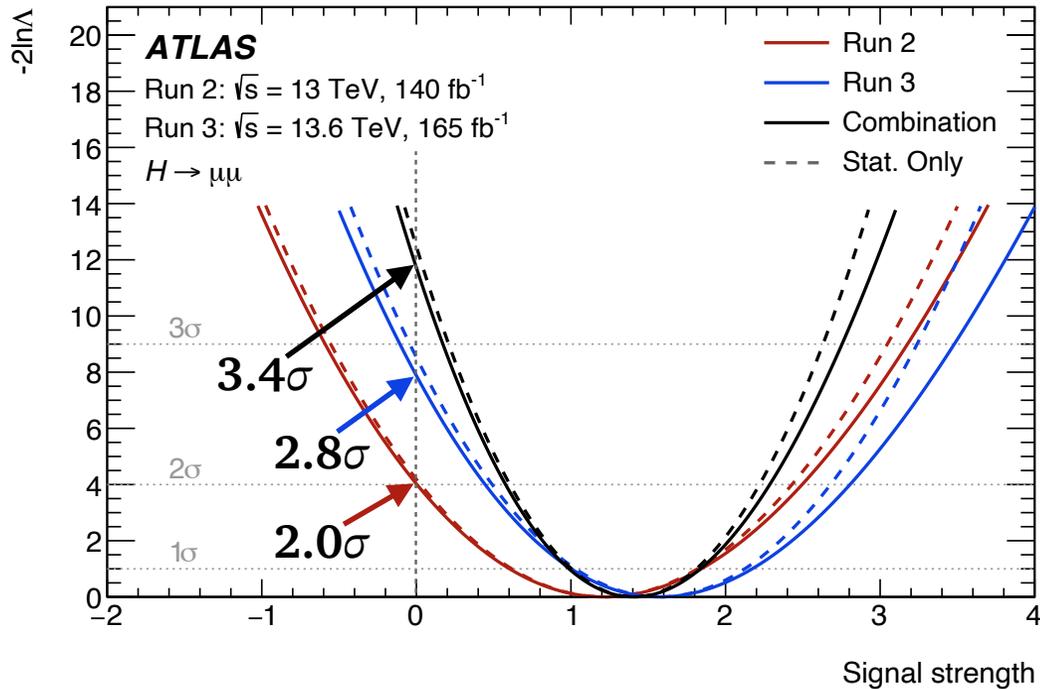
- Significance wrt bkg. Only hypothesis:

  - Expected: **2.5 σ**

    Combination improves Run 2 results **by 50%**

  - Observed: **3.4 σ**

  Clear evidence of H → μ⁺μ⁻ decay!

ATLAS EXPERIMENT
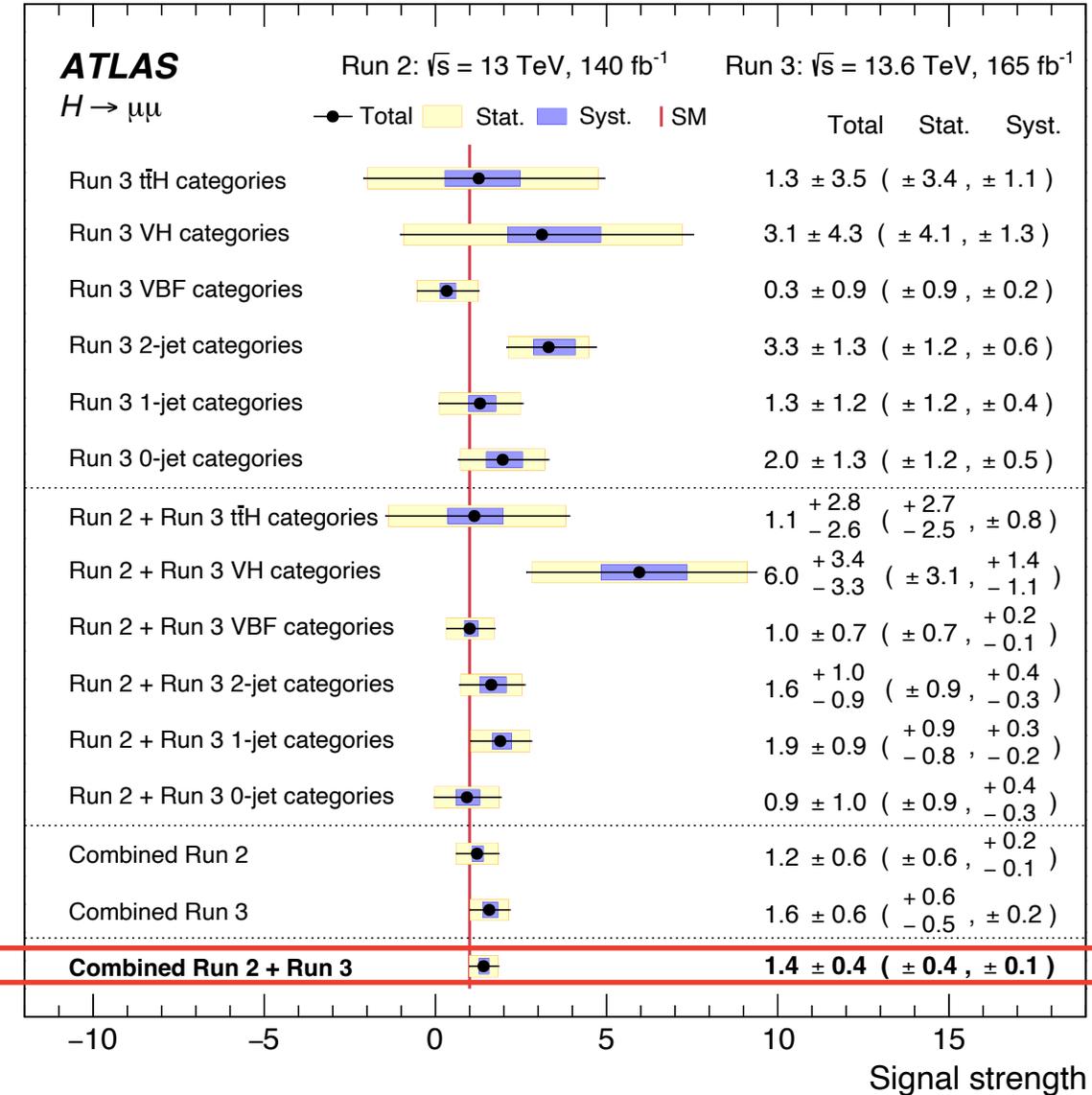
Run: 484466
Event: 3913031636
2024-09-13 05:46:24 CEST

Signal candidate

Evidence of
H → µ⁺µ⁻ decay!

Thanks for your attention!

- Select events with **two opposite-charge high pT muons** + quality requirements

	Selection
Common preselection	Primary vertex Two opposite-charge muons Muons: $|\eta| < 2.5$, $p_{\mathrm{T}}^{\mathrm{lead}} > 27\,\mathrm{GeV}$, $p_{\mathrm{T}}^{\mathrm{sublead}} > 15\,\mathrm{GeV}$
Fit region	$m_{\mu\mu} = 110 - 160$ GeV
Jets	$p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and $|\eta| < 2.4$ or with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and $2.4 < |\eta| < 4.5$
b-tagged jets	$p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and $|\eta| < 2.4$ or with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and $2.4 < |\eta| < 2.5$ Tagging efficiency working point of 85%
$t\bar{t}H$ categories	At least one b-jet
VH 4-lepton category	Exactly two additional e or μ with $p_{\mathrm{T}} > 8\,\mathrm{GeV}$, $5\,\mathrm{GeV}$ (μ) / $7\,\mathrm{GeV}$ (e), no b-jets
VH 3-lepton categories	Exactly one additional e or μ with $p_{\mathrm{T}} > 15\,\mathrm{GeV}$, no b-jets
VH 2-lepton categories	No additional lepton, no b-jets, $E_{\mathrm{T}}^{\mathrm{miss}} > 120$ GeV
VBF and ggF categories	No additional lepton, no b-jets, $E_{\mathrm{T}}^{\mathrm{miss}} < 120$ GeV

FSR recovery

- Muons from the $H \to \mu\mu$ decay can lose their energy by Final State Radiation(FSR).

  - Large E_T, Small ΔR from μ

- Improve the invariant mass resolution by including FSR photons. (~4%)

- Currently applied to categories only contain two muons (ggF, VBF, $Z_{\nu\nu}H_{\mu\mu}$).

  - Far candidates are excluded with $\Delta R < 0.2$ cut to suppress photons from $H \to Z(\mu\mu)\gamma$ decay

  - Only near candidates are used, fake FSR suppressed via E_T cut: $E_T > 25\,\text{GeV} \cdot \Delta R(\mu, \gamma) + 3\,\text{GeV}$

ATLAS Simulation
$\sqrt{s} = 13.6$ TeV
$H \to \mu\mu$

Only events with FSR
(6.6% of all events with exactly two muons)

before FSR recovery
Mean: 111.8 GeV
RMS: 11.2 GeV

after FSR recovery
Mean: 125.6 GeV
RMS: 4.7 GeV

$m_{\mu\mu}$ [GeV]
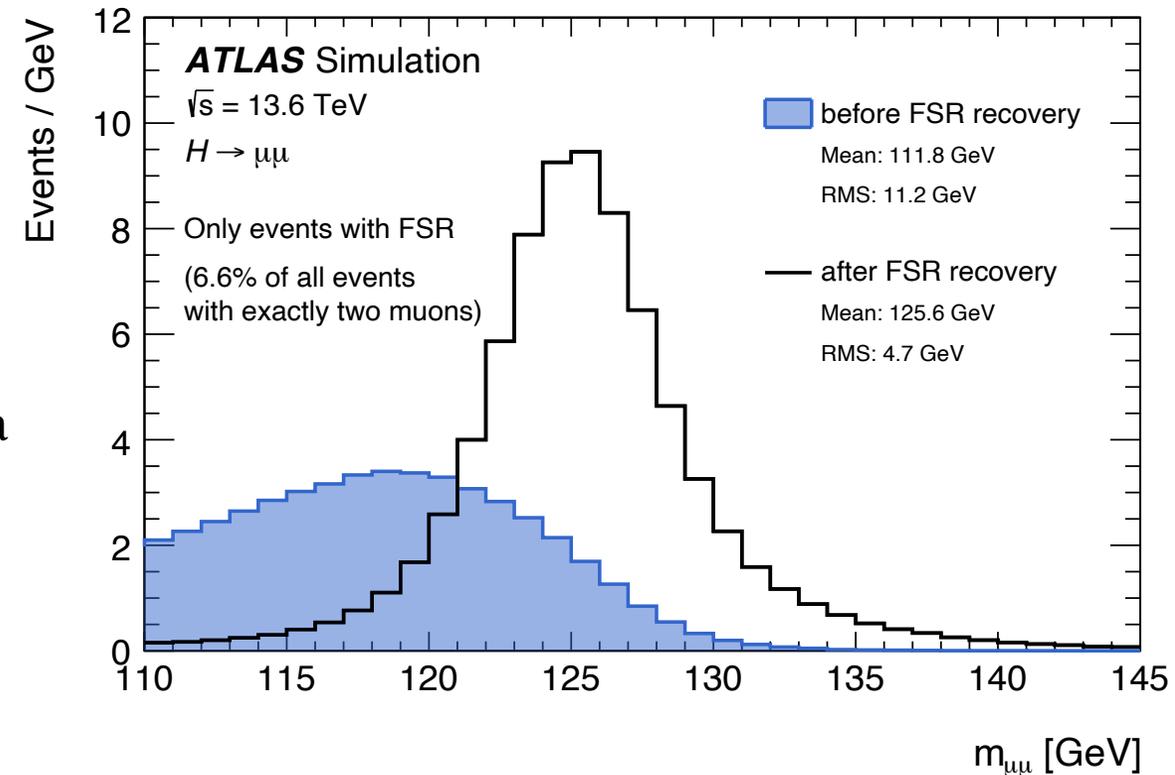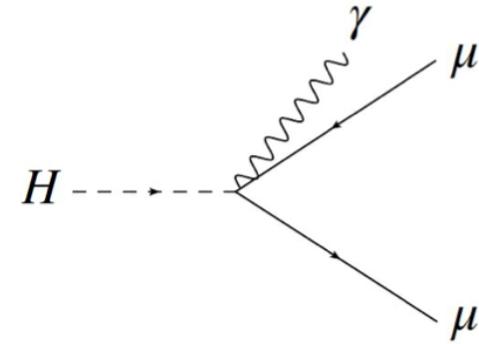
- The event reconstruction is improved by fitting the decay vertex of di-muon system.

  ○ A more precise value for $d_0 \rightarrow$ a better estimate of $m_{\mu\mu}$.

  ○ A narrower di-muon peaks: $Z(\mu\mu)$, $H(\mu\mu)$

    ▸ Invariant mass resolution improved ~2%.

ATLAS Simulation
$\sqrt{s} = 13.6$ TeV
$H \rightarrow \mu\mu$

standard reconstruction
Mean: 124.08 GeV
RMS: 3.73 GeV

+ vertex fitting
Mean: 124.09 GeV
RMS: 3.68 GeV

+ FSR recovery
Mean: 124.37 GeV
RMS: 3.53 GeV
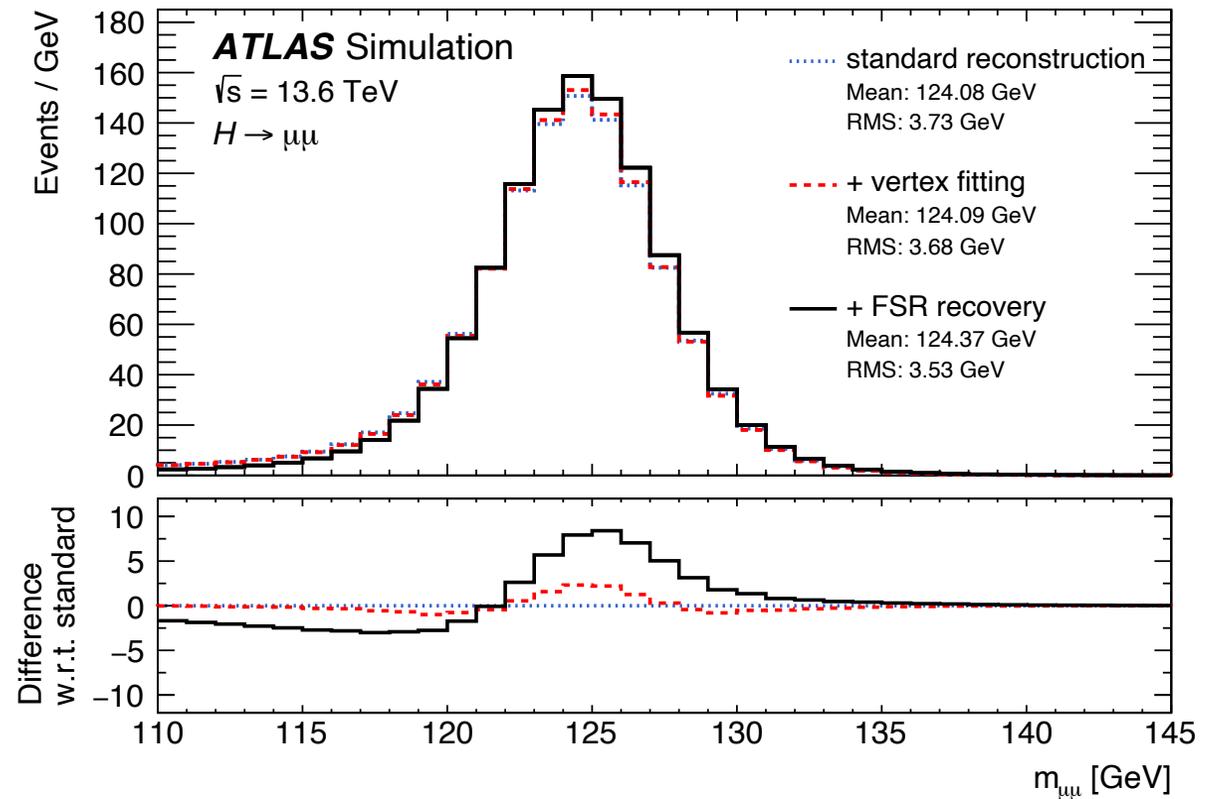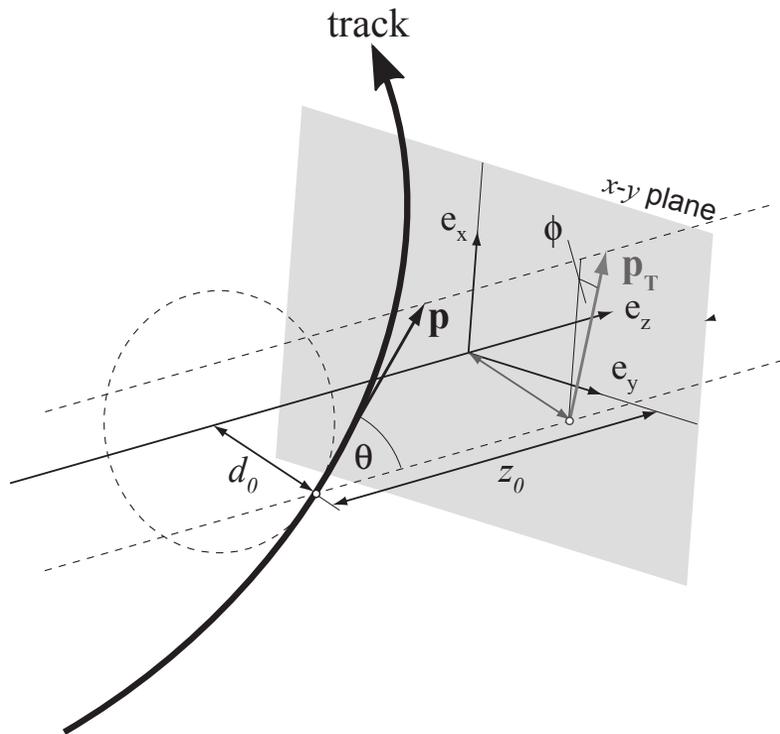
- ttH

Variable	Description
$p_T^{\mu_1}$	Transverse momentum of the leading muon
$p_T^{\mu_2}$	Transverse momentum of the sub-leading muon
η^{μ_1}	Pseudorapidity of the leading muon
η^{μ_2}	Pseudorapidity of the sub-leading muon
ϕ^{μ_1}	Azimuthal angle of the leading muon
ϕ^{μ_2}	Azimuthal angle of the sub-leading muon
N_μ	Number of muons
$\cos\theta^*$	μ^- decay angle in the Collins–Soper frame
$p_T^{j_1}$	Transverse momentum of the leading jet
$p_T^{j_2}$	Transverse momentum of the sub-leading jet
η^{j_1}	Pseudorapidity of the leading jet
η^{j_2}	Pseudorapidity of the sub-leading jet
$N_{\text{Track}}^{j_1}$	Number of tracks in the leading jet with $p_T > 0.5$ GeV
$N_{\text{Track}}^{j_2}$	Number of tracks in the sub-leading jet with $p_T > 0.5$ GeV
m^{jj}	Invariant mass of the dijet system
H_T^j	Scalar sum of jet transverse momentum
$N_{j-\text{central}}$	Number of central ($|\eta| < 2.5$) jets
$N_{b-\text{jet}}$	Number of b-jets with 85% working point
E_T^{miss}	Missing transverse energy
N_e	Number of electrons
$p_T^{\ell_3}$	Transverse momentum of the third lepton
$p_T^{\ell_4}$	Transverse momentum of the fourth lepton
m^{sub}	Mass of leading non-Higgs-boson-candidate muon and opposite-sign Higgs-boson-candidate muon
$m^{\ell_3\ell_4}$	Mass of third and fourth leptons
$\Delta\phi(\ell_3, \ell_4)$	Azimuthal angle between the third and fourth leptons
$\Delta\eta(\ell\ell, \ell_3)$	Pseudorapidity separation between the dimuon system and third lepton
$m_T^{\text{t-Lep}}$	Transverse mass of leptonic top-quark decay
$m^{\text{t-Had}}$	Mass of hadronic top-quark decay
$m_T^{W\text{-Lep}}$	Transverse mass of leptonic W-boson decay
$m^{W\text{-Had}}$	Mass of hadronic W-boson decay

- ZH4L

Variable	Description
$p_T^{\text{pair}_1}$	Transverse momentum of the first pair of leptons
$p_T^{\text{pair}_2}$	Transverse momentum of the second pair of leptons
$p_T^{\text{pair}_1 \ell_1}$	Transverse momenta of the leading lepton of the first pair
$\Delta R(\text{pair}_1, \text{pair}_2)$	ΔR between the first and second pair
$\Delta R(\ell_1, \ell_2)$	ΔR between the leading leptons of each pair
$\alpha(\text{pair}_1, \text{pair}_2)$	Opening angle between the lepton pairs
$\log_{10}(-u), \log_{10}(-t)$	Mandelstam $u-$ and t-variables
H_T^ℓ	Scalar sum of transverse momentum of the four leptons
$p_T^{\ell_i} \ (i = 2, 3, 4)$	p_T of the second, third and fourth lepton
$|\eta_{\text{forward}}|$	Pseudorapidity of most forward lepton pair
$\alpha_{CMF}(\ell_1^{\text{pair}_1}, \ell_2^{\text{pair}_1})$	Opening angle between leptons of Z-like pair (or pair$_1$) in center-of-mass frame
$\theta_{CMF}(\text{pair}_2)$	Polar angle of H-boson candidate pair (or pair$_2$) in center-of-mass frame

- WH3L

Variable	Description
$p_T^{\ell_W}$	Transverse momentum of the W-boson lepton
E_T^{miss}	Missing transverse energy
m_T^W	Transverse mass of the W boson
N_j	Number of jets
N_μ	Number of muons
$\Delta R(\mu, \mu)$	ΔR between the Higgs-boson candidate muons
$\cos\theta^*$	μ^- decay angle in the Collins–Soper frame
$m^{3\ell}$	Invariant mass of the three leptons
$p_T^{j_1}$	Transverse momenta of the leading jet
$\Delta\phi(\ell\ell, \ell_W)$	Azimuthal angle between the muon pair and the W-boson lepton
$\Delta\eta(\ell\ell, \ell_W)$	Pseudorapidity separation between muon pair and W-boson lepton
$\Delta\phi(\ell\ell, E_T^{\text{miss}})$	Azimuthal angle between the muon pair and the missing transverse momentum

- VBF and ggF2/1/0- Jets

Variable	Description	$n_j = 0$	$n_j = 1$	$n_j \geq 2$ and VBF
$p_T^{\mu\mu}$	Transverse momentum of muon pair	x	x	x
$Y^{\mu\mu}$	Rapidity of muon pair	x	x	x
$\cos\theta^*$	μ^- decay angle in the Collins–Soper frame	x	x	x
$p_T^{j_1}$	Transverse momentum of leading jet		x	x
η^{j_1}	Pseudorapidity of leading jet		x	x
$\Delta\phi(\mu\mu, j_1)$	Azimuthal angle between the muon pair and the leading jet		x	x
$N_{\text{Track}}^{j_1}$	Number of tracks in leading jet with $p_T > 0.5$ GeV		x	x
$p_T^{j_2}$	Transverse momentum of sub-leading jet			x
η^{j_2}	Pseudorapidity of sub-leading jet			x
$\Delta\phi(\mu\mu, j_2)$	Azimuthal angle between the muon pair and the sub-leading jet			x
$N_{\text{Track}}^{j_2}$	Number of tracks in sub-leading jet with $p_T > 0.5$ GeV			x
p_T^{jj}	Transverse momentum of dijet system			x
Y^{jj}	Rapidity of dijet system			x
$\Delta\phi(\mu\mu, jj)$	Azimuthal angle between the muon pair and the dijet system			x
m^{jj}	Invariant mass of dijet system			x
E_T^{miss}	Missing transverse energy			x
H_T^j	Scalar sum of jet transverse momentum			x

- $Z(\nu\nu)H(\mu\mu)$

| Variable	Description
$p_T^{\mu_1}$	Transverse momentum of leading muon
$p_T^{\mu_2}$	Transverse momentum of sub-leading muon
E_T^{miss}	Missing transverse energy
N_j	Number of jets
$\Delta R(\mu, \mu)$	ΔR between the Higgs-boson candidate muons
$\Delta\phi(\ell\ell, E_T^{\text{miss}})$	Azimuthal angle between the muon pair and the missing transverse momentum
V_T/H_T	Ratio of vectorial V_T and scalar H_T sum of reconstructed objects in event
$\cos\theta^*$	μ^- decay angle in the Collins–Soper frame
$\sigma(E_T^{\text{miss}})$	Missing transverse energy significance

Categorization workflow

- Preselections:
  - Muons: $|\eta| < 2.5$, Leading (Sub) $p_T > 27$ (15) GeV
  - Jets: $p_T > 25$ GeV and $|\eta| < 2.4$, or $p_T > 30$ GeV and $|\eta| < 4.5$
  - B-tagging efficiency: GN2 WP of 85%
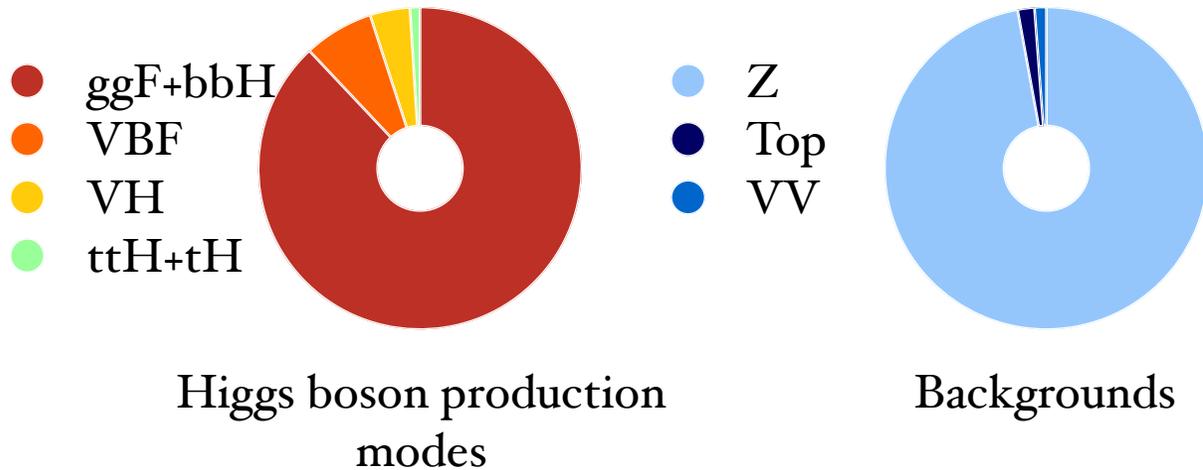
Higgs boson production modes

Backgrounds

- ggF+bbH
- VBF
- VH
- ttH+tH

- Z
- Top
- VV

- Sort events into mutually exclusive categories, targeting the various Higgs boson production modes

- Category boundaries optimized for maximize sensitivity
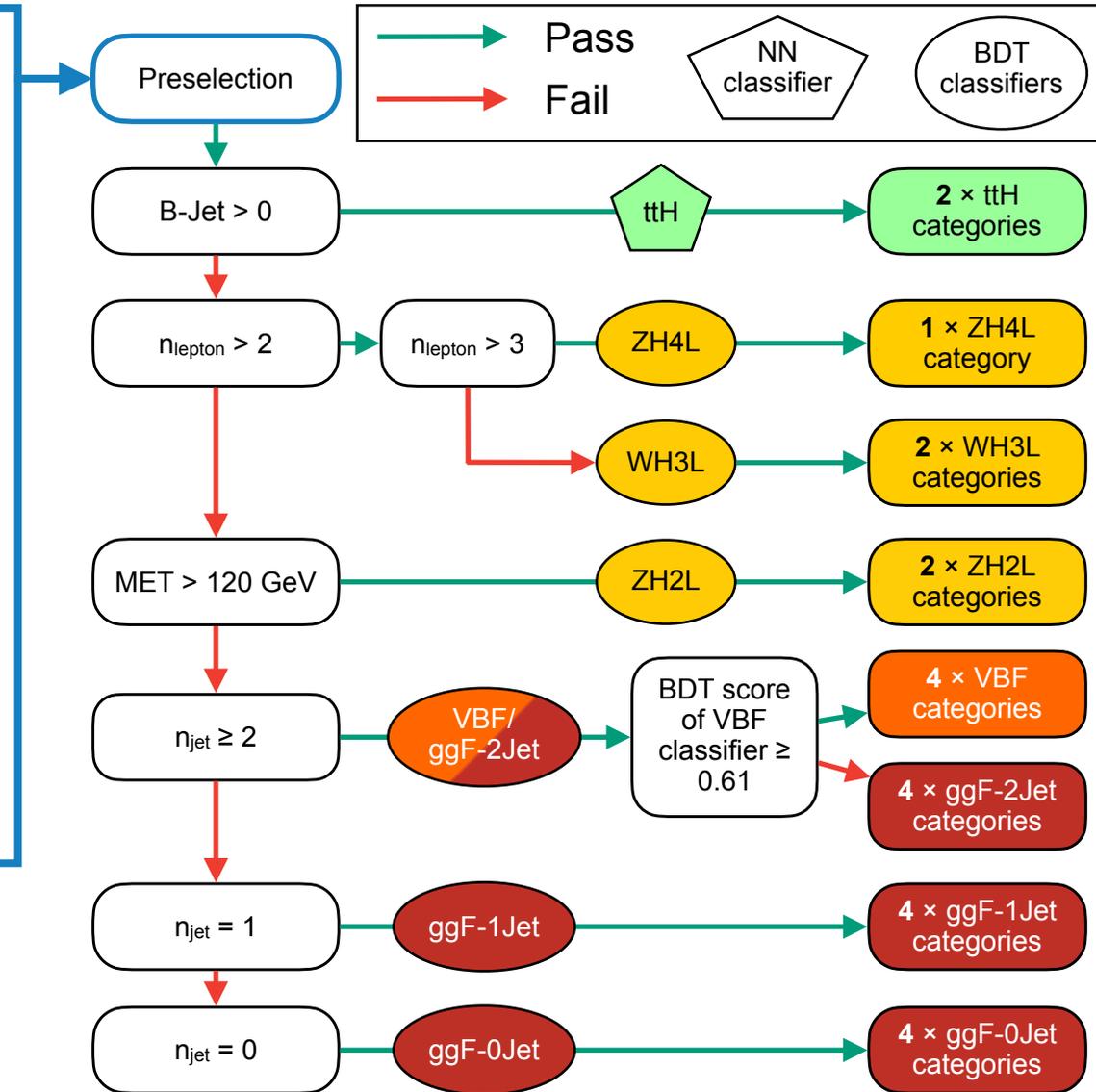
- Tagging the ttH and VH events with the decay products of top/Z/W.

  ▸ **ttH categories**: events with at least one b-tagged jet

    - Used distance correlation neural-network classifier to avoid correlation to dimuon mass

  ▸ **ZH4L category:** events with 4-leps

    - Dilepton pair with mass closest to the Z boson is labeled as the Z-boson candidate

$$\chi^2 = a(\frac{m_{l_1l_2} - m_H}{\sigma_H})^2 + (\frac{m_{l_3l_4} - m_Z}{\sigma_Z})^2$$

  ▸ **WH3L category**: events with 3-leps

    - Pairing the leptons

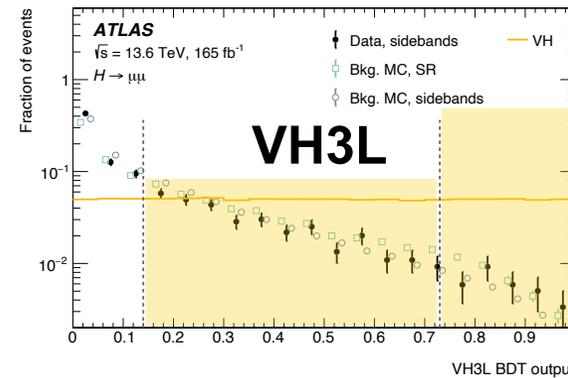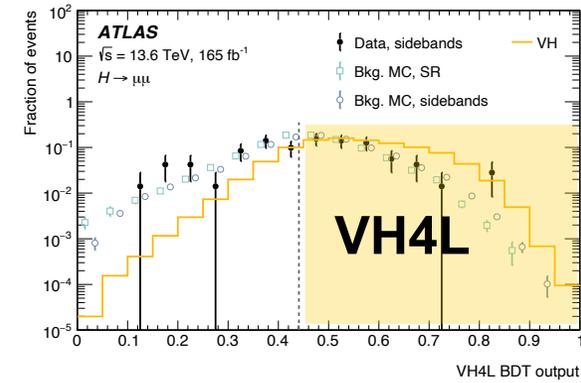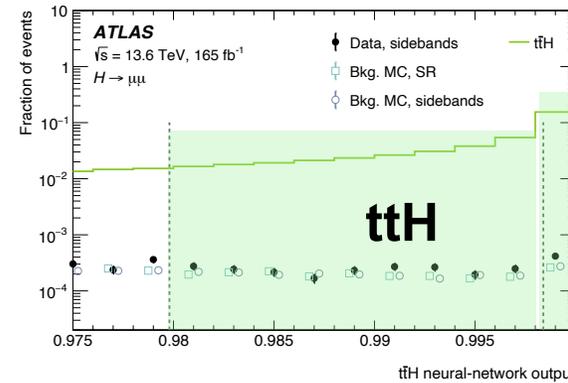$$\chi^2 = (\frac{m_{\mu_1\mu_2} - m_H}{\sigma_H})^2 + (\frac{m^T_{\mu_3,MET} - m^T_W}{\sigma^T_W})^2$$

  ▸ **ZH2L category**: events with 2-leps and large Missing E_T (>120 GeV)

    - Targeting Z boson's decay to neutrinos
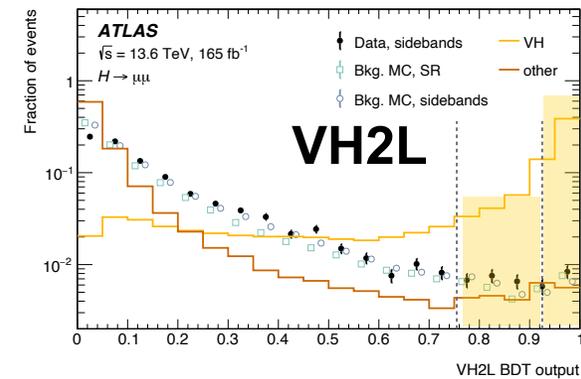
- Remaining events from the ttH and VH selections.
  → no extra leptons, no b-tagged jet, Missing E_T <120 GeV

  ▸ **VBF categories**: events with at least 2 jets

    - 2-D boundary optimization with ggF 2-jets categories

  ▸ **ggF 2/1/0-jet categories**:

    - Exploit boost of dimuon pT, central η etc.

- All the classifiers used multi-fold approach to maximum use of the MC events and to avoid from over training.
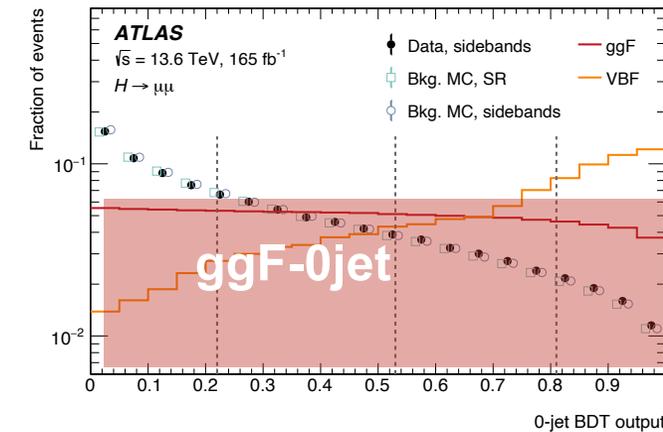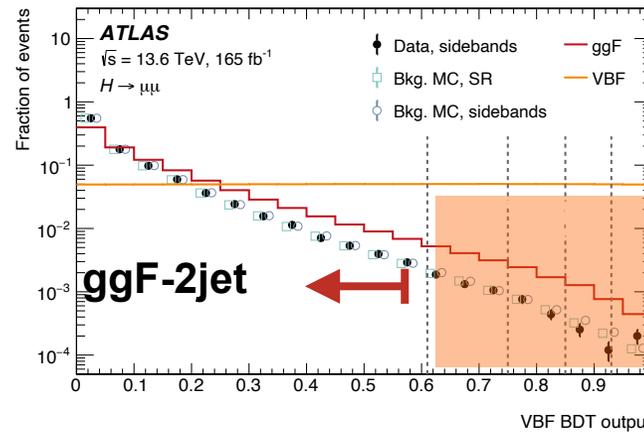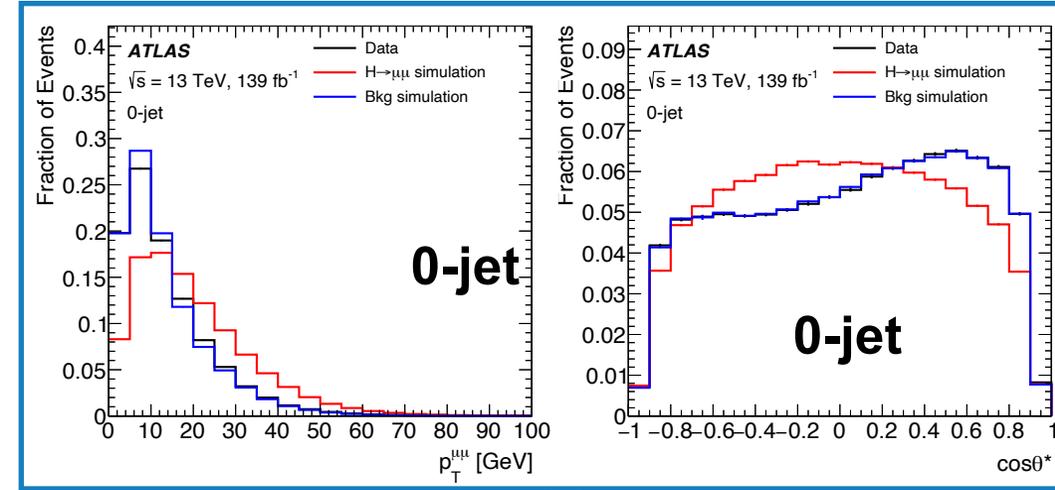
Events	25%	25%	25%	25%	Final
1st fold	Training	Training	Validation	Test	Test
2st fold	Test	Training	Training	Validation	Test
3st fold	Validation	Test	Training	Training	Test
4st fold	Training	Validation	Test	Training	Test

- Empirical function Determine by **spurious signal test**:

  ▸ Fit **S+B model** to a **bkg. only MC template**, maximum value of fitted signal yield in range [120, 130] GeV called "spurious signal" (SS).

- Tested candidates
  - EpolyN → $exp(a_0 + a_1 m_{\mu\mu} + \ldots + a_N m_{\mu\mu}^N)$
  - PowerN → $m_{\mu\mu}^{(a_0 + a_1 m_{\mu\mu} + \ldots + a_N m_{\mu\mu}^N)}$

▸ Require:

  - SS < 20% of the uncertanty from the data statistical fluctuation

  - χ^2 probability of the bkg. only fit > 1%

▸ Select the model with the **smallest number of free parameters** then with the **smallest SS** among the functions met the above requirements.

▸ SS taken as background model systematic.
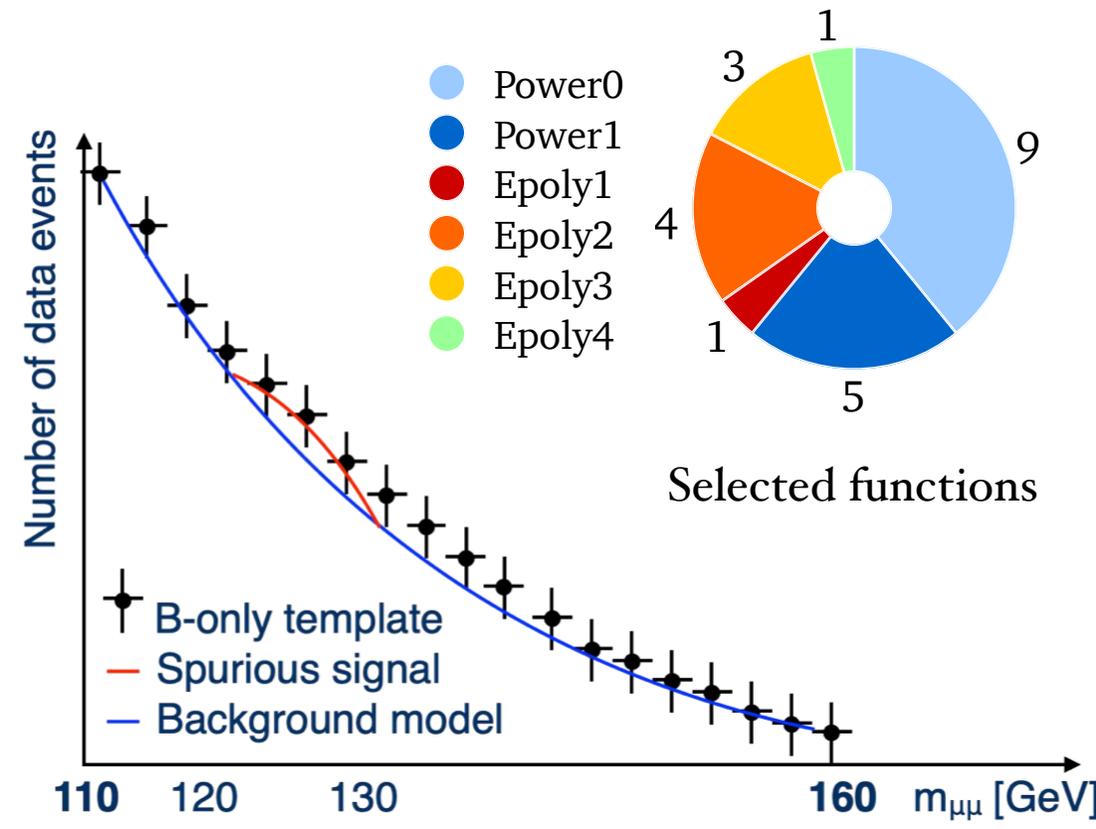
Selected functions

- Best fit value of the signal strength:
  (m_H = 125.09 GeV)

  ▸ Run3: $\mu(\sigma/\sigma_{SM}) = 1.6 \pm 0.6$

- Significance wrt bkg. Only hypothesis:

  - Observed: **2.8 σ**

  - Expected: **1.8 σ**

Uncertainty source	$\Delta\mu$	
Statistical uncertainty	-0.55	+0.55
Systematic uncertainty	-0.18	+0.24
Spurious signal	-0.14	+0.15
Theory	-0.08	+0.13
Luminosity	-0.04	+0.08
Muon	-0.03	+0.07
Jets, flavor tagging	-0.02	+0.03
Other	-0.04	+0.07
Total	-0.57	+0.60

ATLAS
$\sqrt{s}$ = 13.6 TeV, 165 fb^{-1}
$H \to \mu\mu$
Inclusive, ln(1+S/B) weighted

Data
Total pdf
Signal pdf
Background pdf
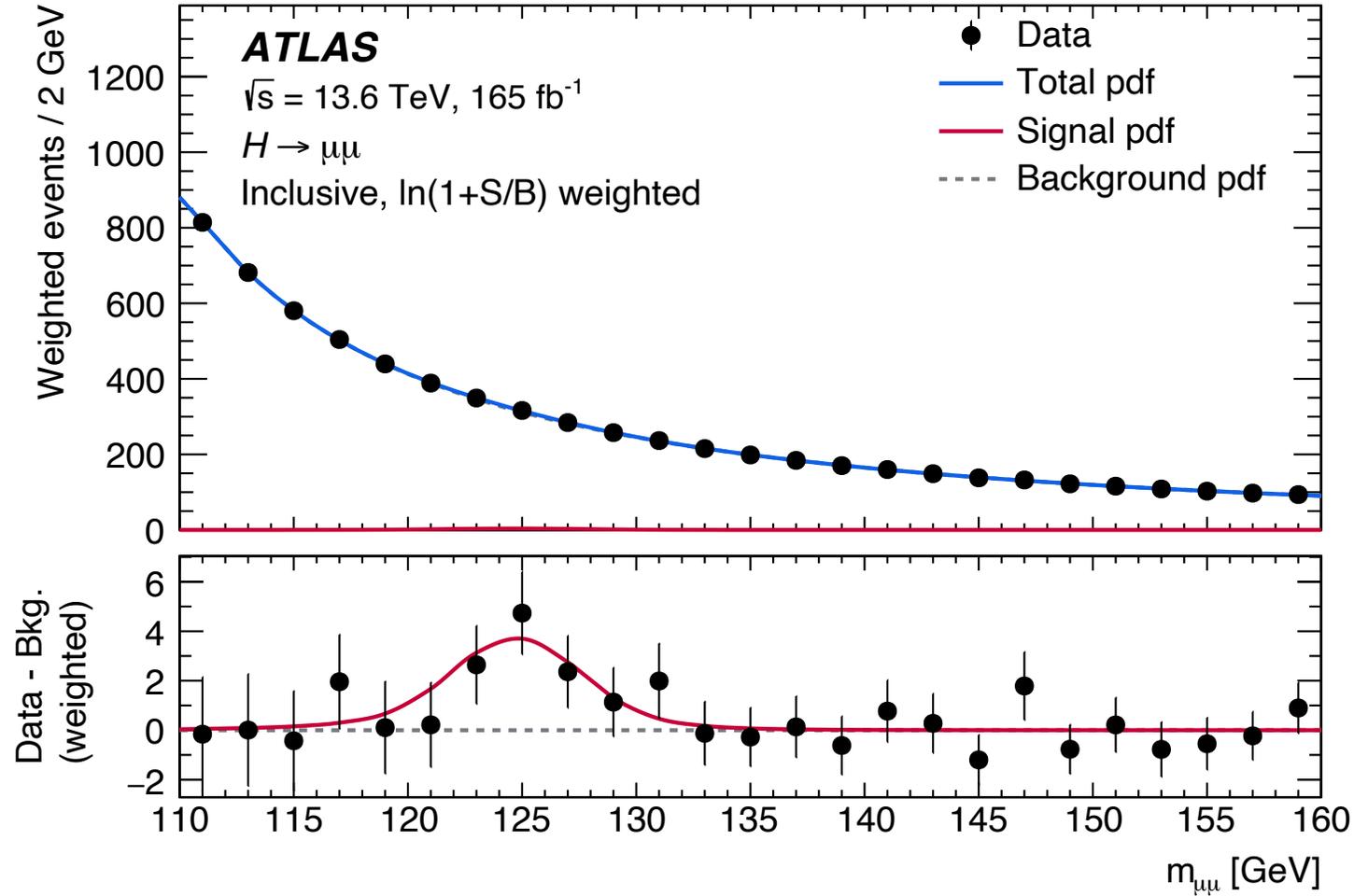
- Best fit value of the signal strength:
  (m_H = 125.09 GeV)

  ‣ Run3 + Run2: $\mu(\sigma/\sigma_{SM})$ = 1.4 ± 0.4

Uncertainty source	$\Delta\mu$	
Statistical uncertainty	-0.40	+0.41
Systematic uncertainty	-0.12	+0.15
Spurious signal	-0.09	+0.09
Theory	-0.07	+0.10
Luminosity	-0.02	+0.04
Muon	-0.02	+0.04
Jets, flavor tagging	-0.01	+0.01
Other	-0.02	+0.03
Total	-0.42	+0.43
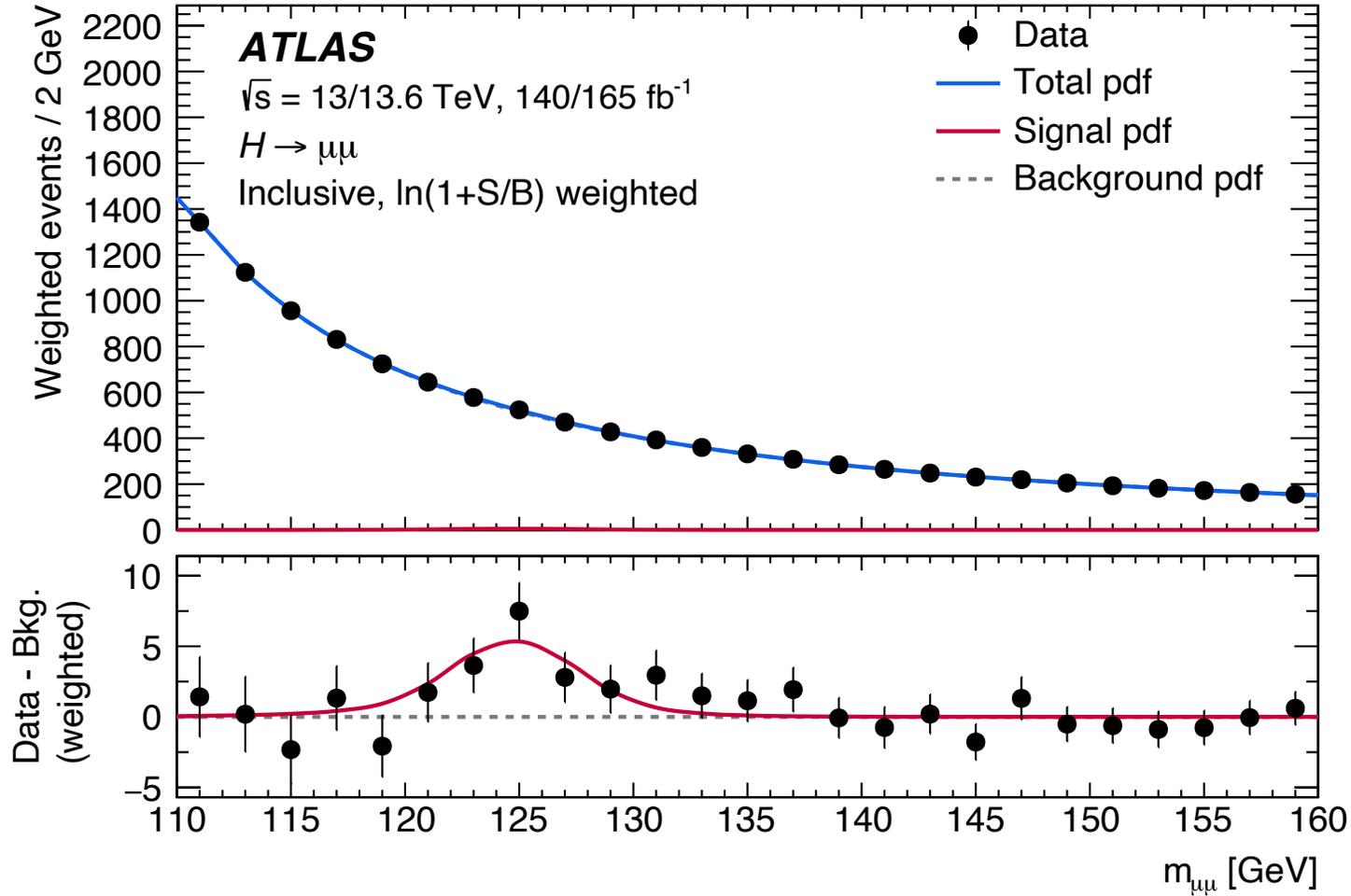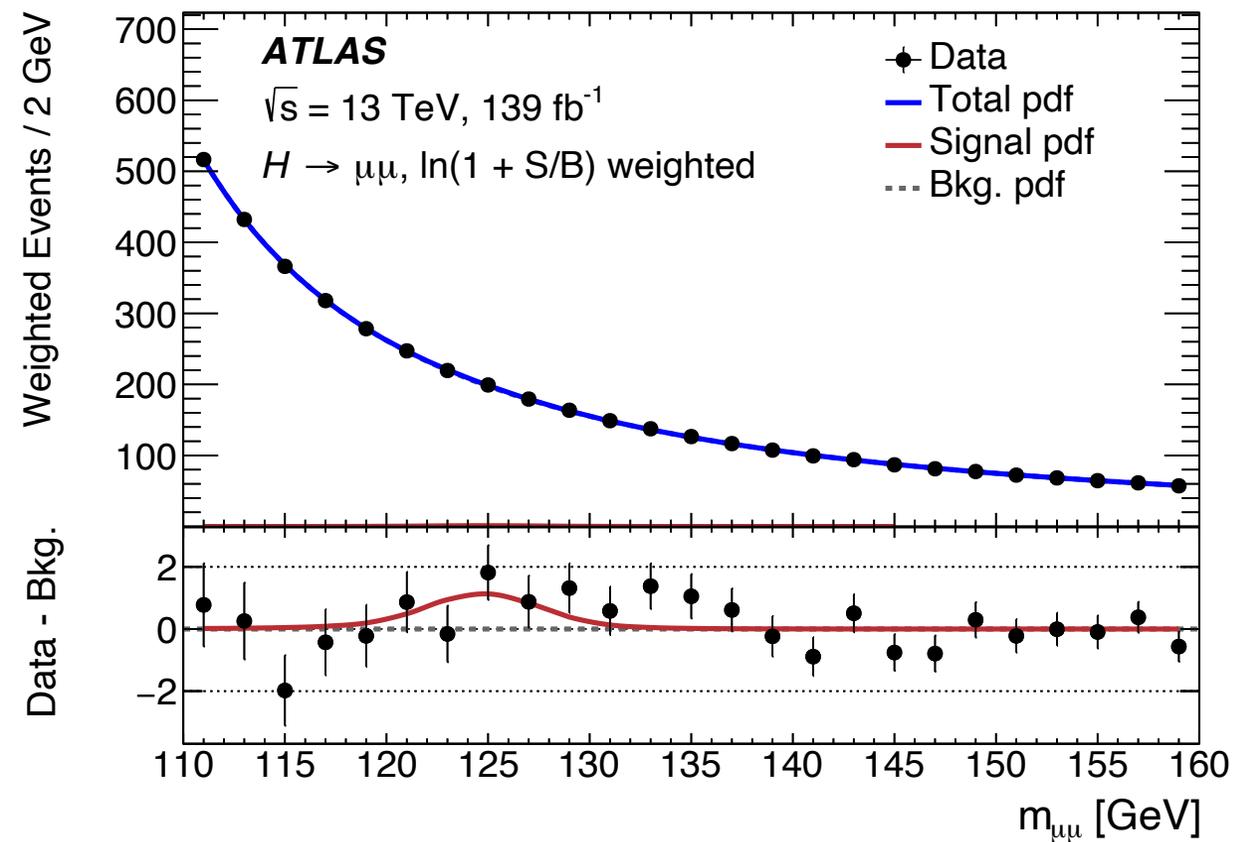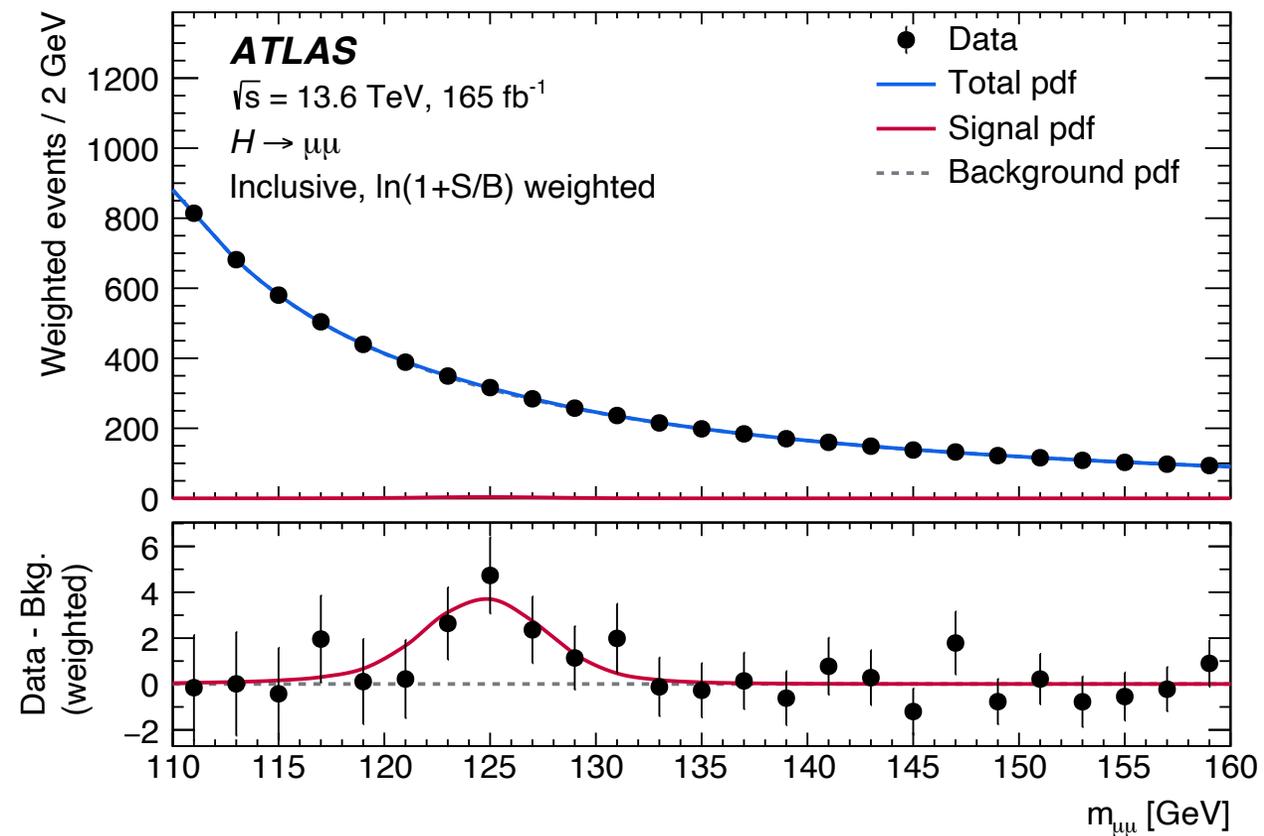
- The expected contribution of the SM Higgs boson production process

Category	$gg\text{F} + b\bar{b}H$	VBF	WH	ZH	$t\bar{t}H + tH$
$t\bar{t}H$-High	0.1%	0.0%	0.5%	0.5%	98.9%
$t\bar{t}H$-Medium	1.4%	0.1%	1.2%	1.3%	96.0%
VH4L	0.0%	0.0%	0.0%	99.3%	0.7%
VH3L-High	0.1%	0.0%	95.1%	3.3%	1.5%
VH3L-Medium	0.5%	0.2%	83.7%	8.3%	7.3%
VH2L-High	0.4%	0.1%	26.1%	72.0%	1.3%
VH2L-Medium	3.1%	1.1%	29.7%	62.5%	3.5%
VBF High	3.9%	96.1%	0.0%	0.0%	0.0%
VBF Medium	7.7%	92.2%	0.0%	0.0%	0.0%
VBF Low	13.4%	86.4%	0.0%	0.1%	0.1%
VBF Very Low	23.4%	76.2%	0.1%	0.1%	0.2%
2-jet High	67.8%	30.7%	0.7%	0.6%	0.2%
2-jet Medium	79.2%	16.9%	2.2%	1.5%	0.2%
2-jet Low	83.2%	9.4%	4.4%	2.7%	0.3%
2-jet Very Low	78.9%	6.7%	8.4%	5.2%	0.8%
1-jet High	67.1%	32.5%	0.2%	0.2%	0.0%
1-jet Medium	85.8%	13.0%	0.7%	0.5%	0.0%
1-jet Low	91.3%	6.5%	1.4%	0.8%	0.0%
1-jet Very Low	92.4%	4.2%	2.2%	1.1%	0.0%
0-jet High	95.9%	2.5%	0.7%	0.9%	0.0%
0-jet Medium	98.3%	1.1%	0.3%	0.3%	0.0%
0-jet Low	98.8%	0.7%	0.2%	0.3%	0.0%
0-jet Very Low	99.5%	0.3%	0.1%	0.1%	0.0%

- The number of observed events and the fitted number of signal (S) and background (B) events

Category	Data	S	B	$S/\sqrt{B}$	S/B [%]
$t\bar{t}H$-High	12	1.9 ± 0.7	15.5 ± 2.3	0.49	12.5
$t\bar{t}H$-Medium	117	3.9 ± 1.4	115 ± 7	0.36	3.4
VH4L	11	0.78 ± 0.29	12.2 ± 1.8	0.22	6.4
VH3L-High	25	1.4 ± 0.5	17.4 ± 2.9	0.33	8.0
VH3L-Medium	143	3.7 ± 1.4	136 ± 10	0.31	2.7
VH2L-High	19	1.0 ± 0.4	18.3 ± 2.9	0.23	5.3
VH2L-Medium	30	0.38 ± 0.14	31.7 ± 3.1	0.07	1.2
VBF High	9	4.3 ± 1.6	10.5 ± 1.8	1.34	41.2
VBF Medium	28	5.3 ± 2.0	25.8 ± 2.7	1.04	20.5
VBF Low	69	7.2 ± 2.7	62 ± 4	0.91	11.6
VBF Very Low	217	11 ± 4	225 ± 8	0.76	5.1
2-jet High	1 399	31 ± 12	$1\,367 \pm 25$	0.84	2.3
2-jet Medium	5 657	69 ± 26	$5\,560 \pm 50$	0.92	1.2
2-jet Low	17 684	110 ± 40	$17\,300 \pm 70$	0.87	0.7
2-jet Very Low	35 147	110 ± 40	$35\,160 \pm 140$	0.59	0.3
1-jet High	708	17 ± 6	710 ± 16	0.65	2.4
1-jet Medium	7 166	80 ± 30	$7\,140 \pm 70$	0.95	1.1
1-jet Low	31 761	180 ± 70	$31\,510 \pm 120$	1.00	0.6
1-jet Very Low	73 578	200 ± 80	$73\,330 \pm 200$	0.75	0.3
0-jet High	19 445	120 ± 50	$19\,260 \pm 90$	0.89	0.6
0-jet Medium	50 742	200 ± 80	$50\,830 \pm 190$	0.90	0.4
0-jet Low	94 032	240 ± 90	$93\,770 \pm 210$	0.78	0.3
0-jet Very Low	136 762	170 ± 60	$136\,510 \pm 290$	0.47	0.1

ATLAS $\sqrt{s}$ = 13 TeV, 139 fb^{-1} $H \to \mu\mu$

	Total	Stat.	Syst.
VH and ttH categories	5.0 ± 3.5	(± 3.3 ,	± 1.1)
ggF 0-jet categories	-0.4 ± 1.6	(± 1.5 ,	± 0.3)
ggF 1-jet categories	2.4 ± 1.2	(± 1.2 ,	± 0.3)
ggF 2-jet categories	-0.6 ± 1.2	(± 1.2 ,	± 0.3)
VBF categories	1.8 ± 1.0	(± 1.0 ,	± 0.2)
Combined	1.2 ± 0.6	(± 0.6 ,	$^{+0.2}_{-0.1}$)

Signal strength

CMS *Preliminary* 137 fb^{-1} (13 TeV)

Combined $\hat{\mu}$ = 1.19$^{+0.44}_{-0.42}$

- Combined best fit μ
- SM expectation
- 68% CL
- 95% CL

m_H = 125.38 GeV

ggH-cat.	$\mu = 0.63^{+0.65}_{-0.64}$
VBF-cat.	$\mu = 1.36^{+0.69}_{-0.61}$
t̄tH-cat.	$\mu = 2.32^{+2.27}_{-1.95}$
VH-cat.	$\mu = 5.48^{+3.10}_{-2.83}$

Best-fit μ