A search for the dimuon decay of the Standard Model Higgs boson with the ATLAS detector

Dimitrii Krasnopevtsev*

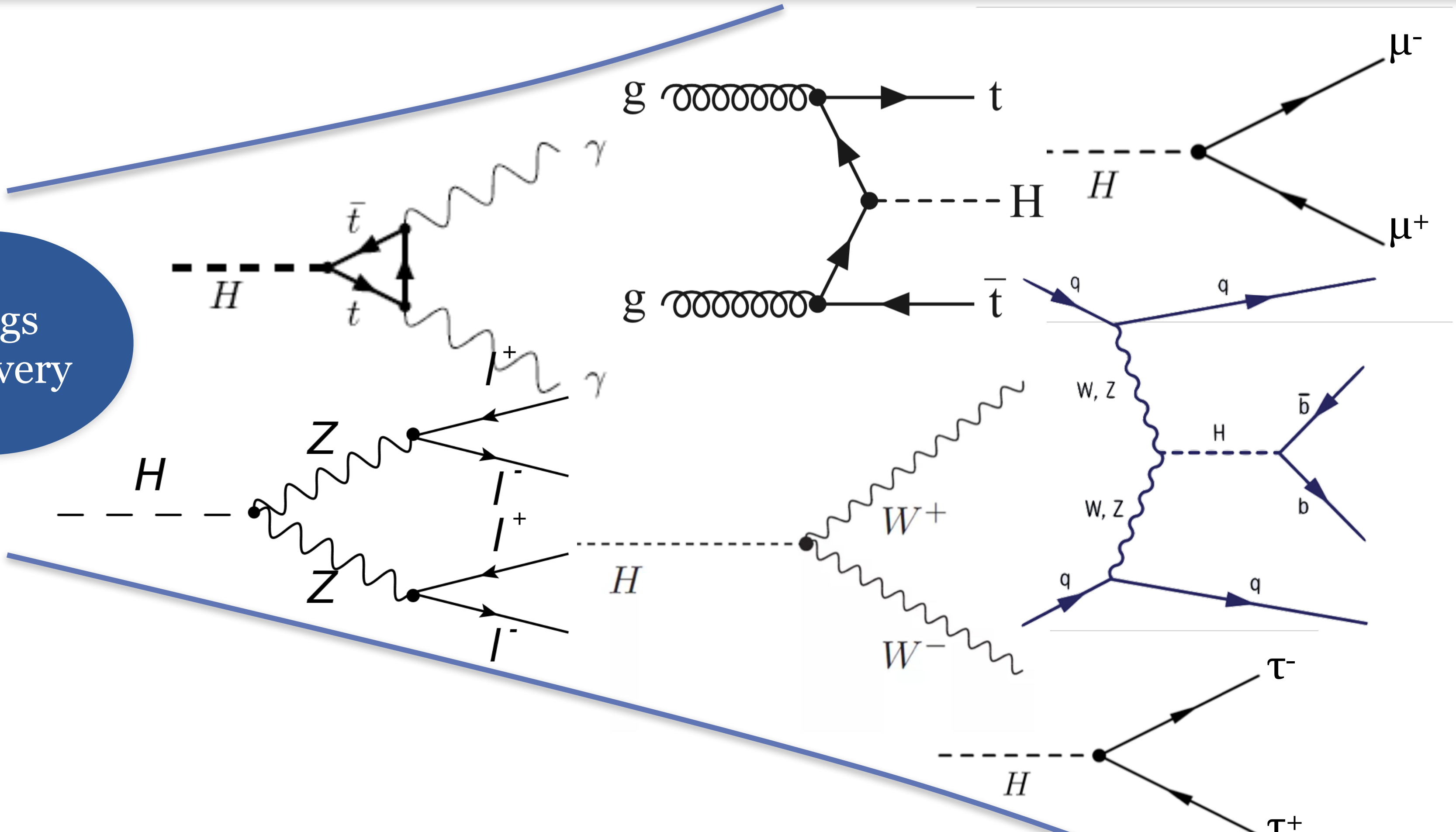
University of Science and Technology of China

Outline

- ❑Motivation
- ❑Analysis strategy
- ❑Event categorization
- ❑Signal and background models
- ❑Results in terms of signal strength
- ❑Prospects for HL LHC

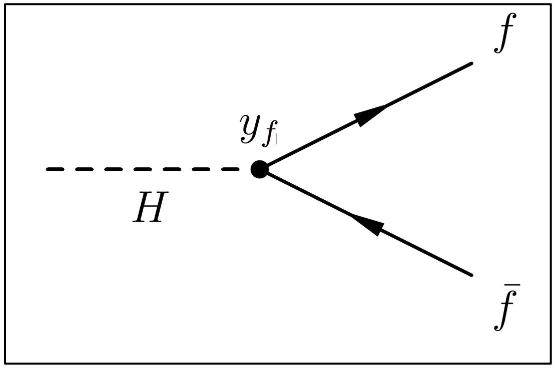
Higgs Bang

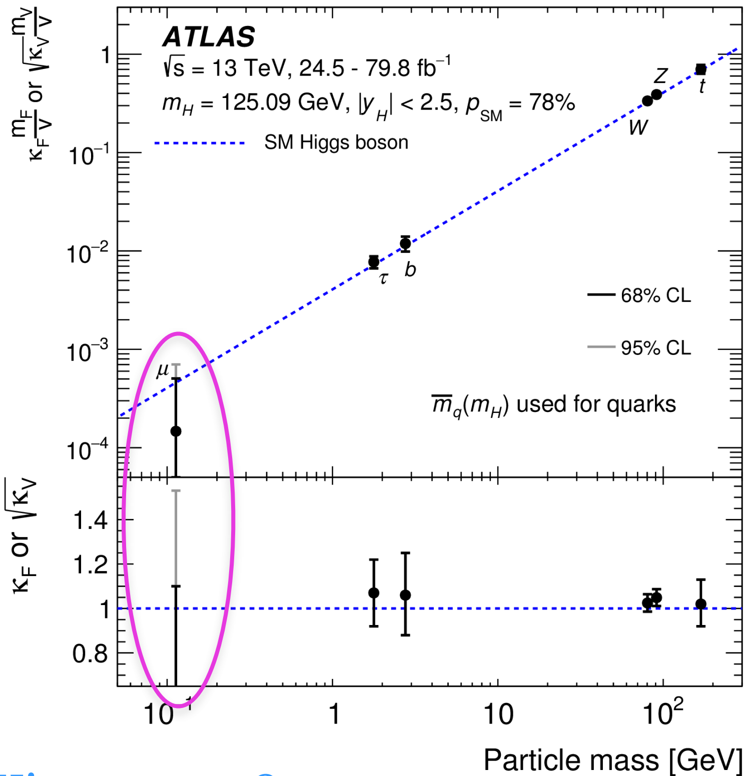
Higgs discovery



- ➢ Fast developing studies recent ≈10 years
  - pushed by new data from LHC and active competition
- ➢ Unique opportunities to learn analysis techniques

Motivation



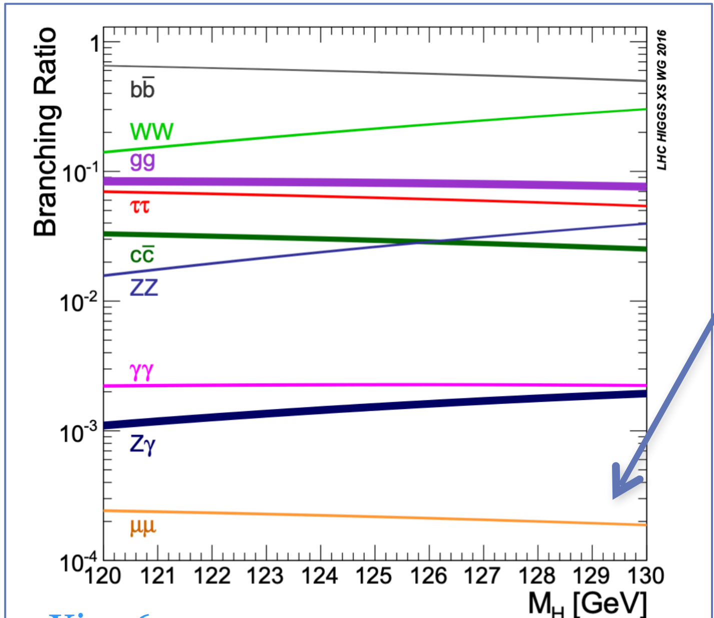


arXiv:1909.02845

➢ Process offers the best possibility to measure the Higgs boson coupling to the second-generation fermions at the LHC.

➢ Any deviation from the SM prediction could be a sign of new physics.

Challenges and previous results



arXiv:1610.07922

- ➢ Very small $B(H\to\mu^-\mu^+) = 2.2\times10^{-4}$ and large background rate from Drell-Yan - $Z/\gamma\to\mu^-\mu^+$
- ➢ Inclusive S/B is ~0.1%: very challenging analysis

Previous result: ICHEP2018, early Run 2 data:

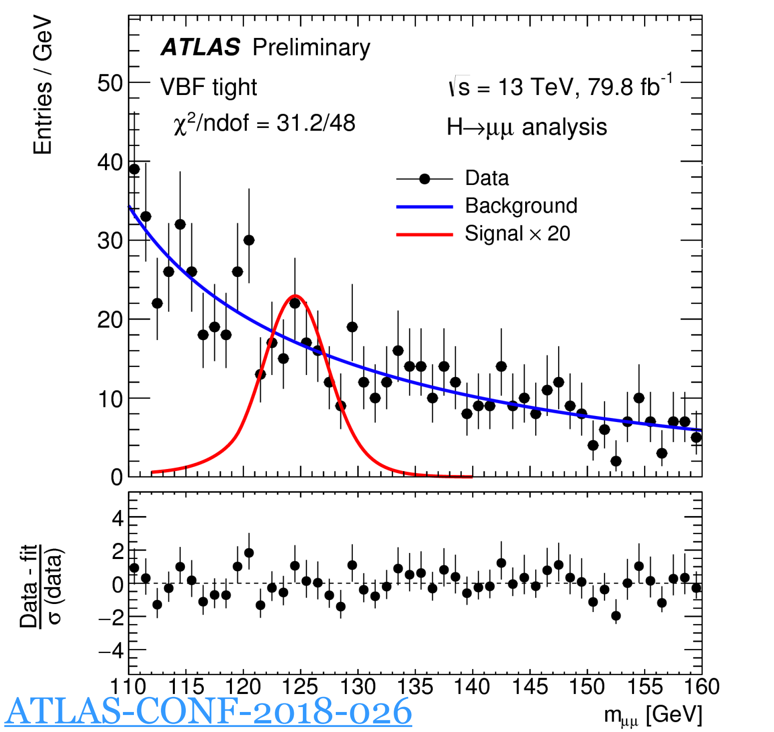
✓Expected significance: **0.95 σ**

✓Signal strength: **μ=0.22±1.05**

(μ=1.00±1.00 expected)

✓**95% CL limits: μ<2.23**

(μ<2.07 expected for μ=0)

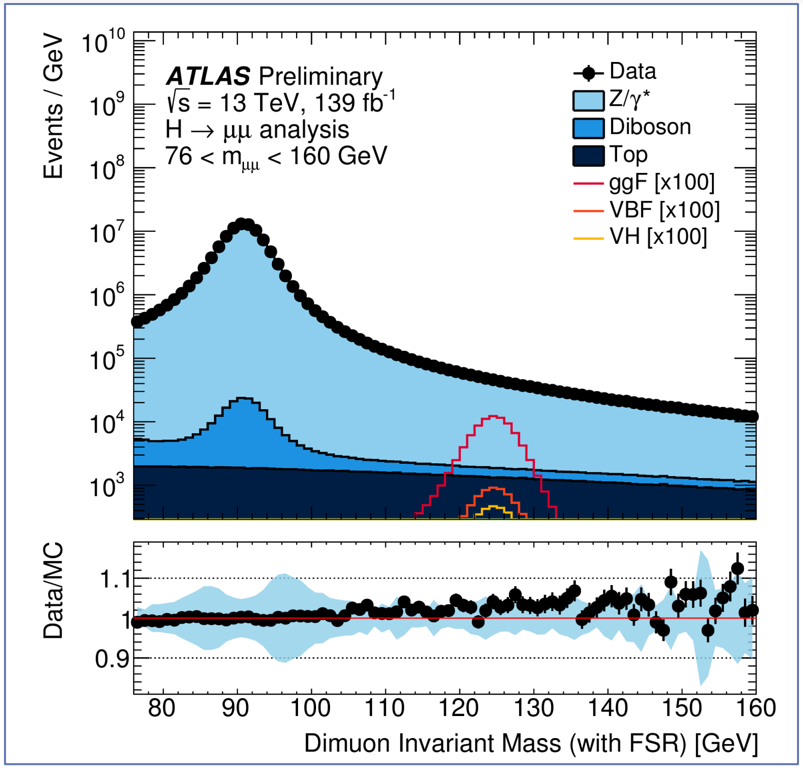


ATLAS-CONF-2018-026

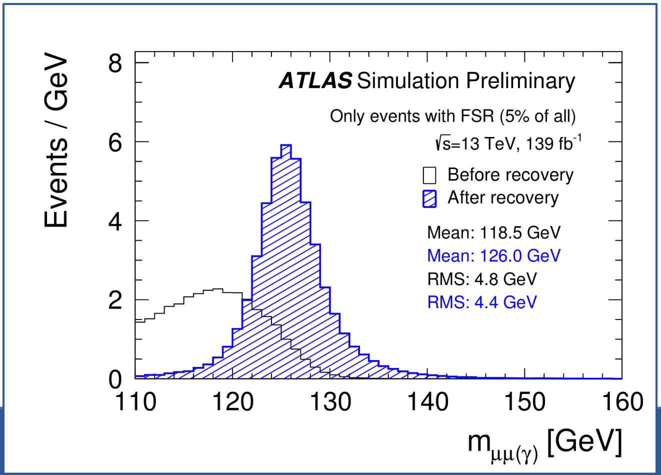
Analysis strategy

ATLAS-CONF-2019-028

- Loose di-muon selection to maximize efficiency
- Veto events with b-tagged jets to reject tt background
- Multivariate categorization to maximize sensitivity (NEW - classification for different jet multiplicities)
- Signal extraction from fit to $m_{\mu\mu(\gamma)}$ invariant mass in data (narrow resonance)
- Recover up to one FSR photon candidate to improve mass reconstruction (NEW)



3% RMS improvement for all ggF events



Events selection

Loosest unprescaled single muon triggers

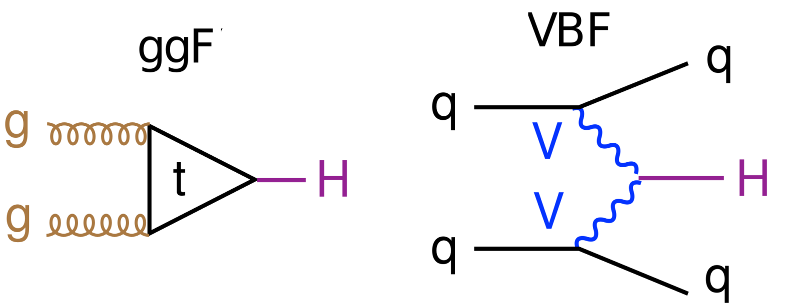
	Selection
Common	Primary vertex Two opposite-charge muons Muons: $\lvert\eta\rvert < 2.7$, $p_{\mathrm{T}}^{\mathrm{lead}} > 27GeV$, $p_{\mathrm{T}}^{\mathrm{sublead}} > 15GeV$ No b-tagged jets
Z region	$76 < m_{\mu\mu} < 106GeV$
Sideband region	$110 < m_{\mu\mu} < 120GeV$ or $130 < m_{\mu\mu} < 180GeV$
Fit region	$110 < m_{\mu\mu} < 160GeV$
Jets	$p_{\mathrm{T}} > 25GeV$ and $\lvert\eta\rvert < 2.5$ or with $p_{\mathrm{T}} > 30GeV$ and $2.5 < \lvert\eta\rvert < 4.5$

loose quality and isolation

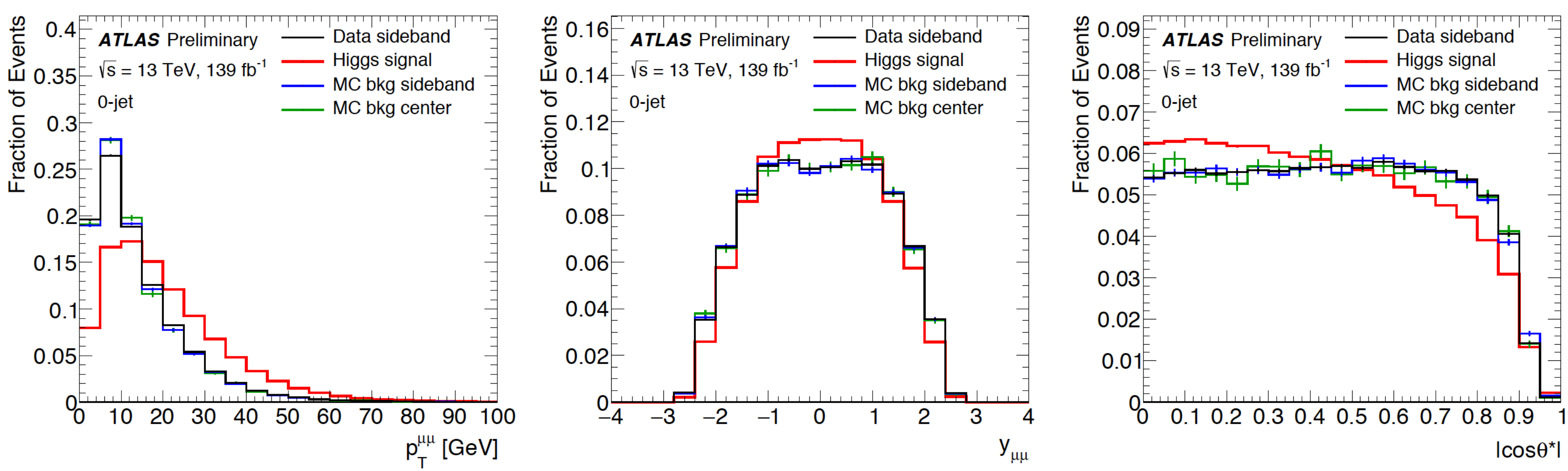
Allows to constrain the backgrounds at the same time as extracting the signal.

Events categorization

- ➢ Target Higgs production via ggF and VBF in exclusive jet selections with ≥2/= 1/= 0 jets
- ➢ Production-driven BDT categorization for separate jet-topology channels to maximize sensitivity (XGBoost BDT trained with signal MC vs. side-band data)



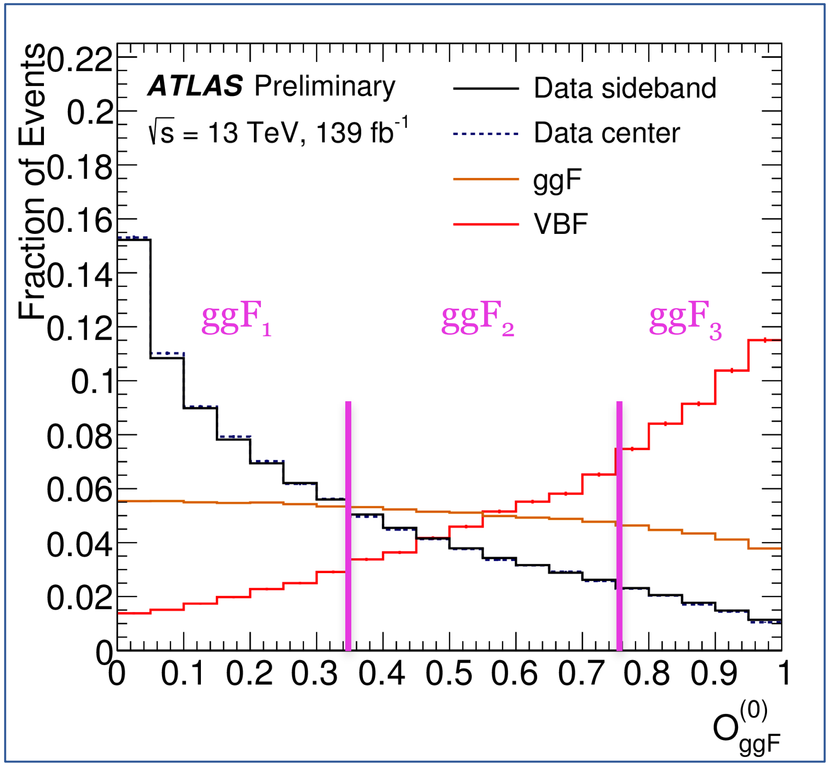
Unique signature

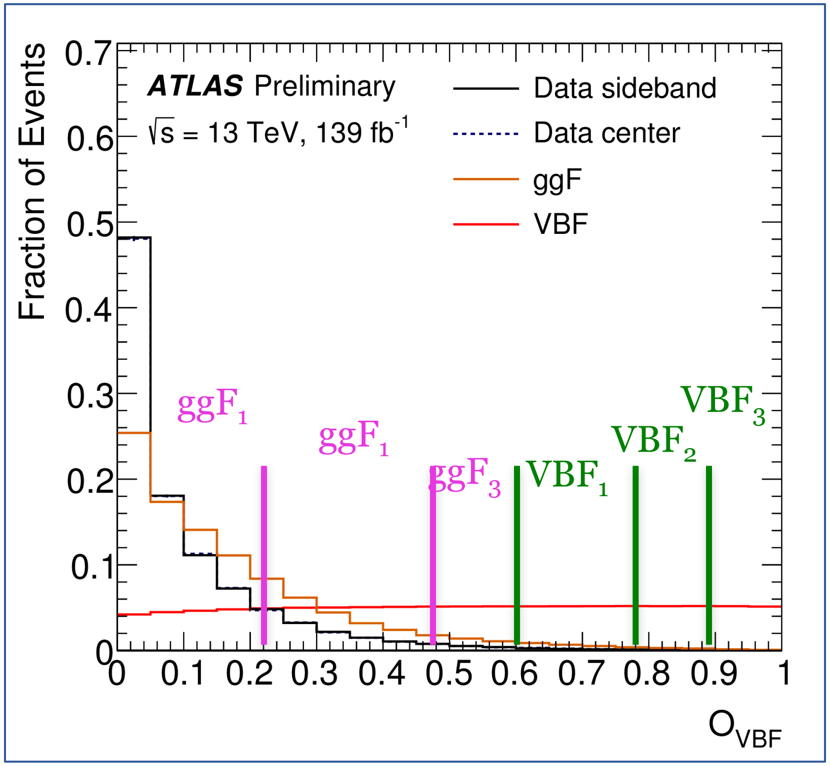


See backup for more details

Events categorization

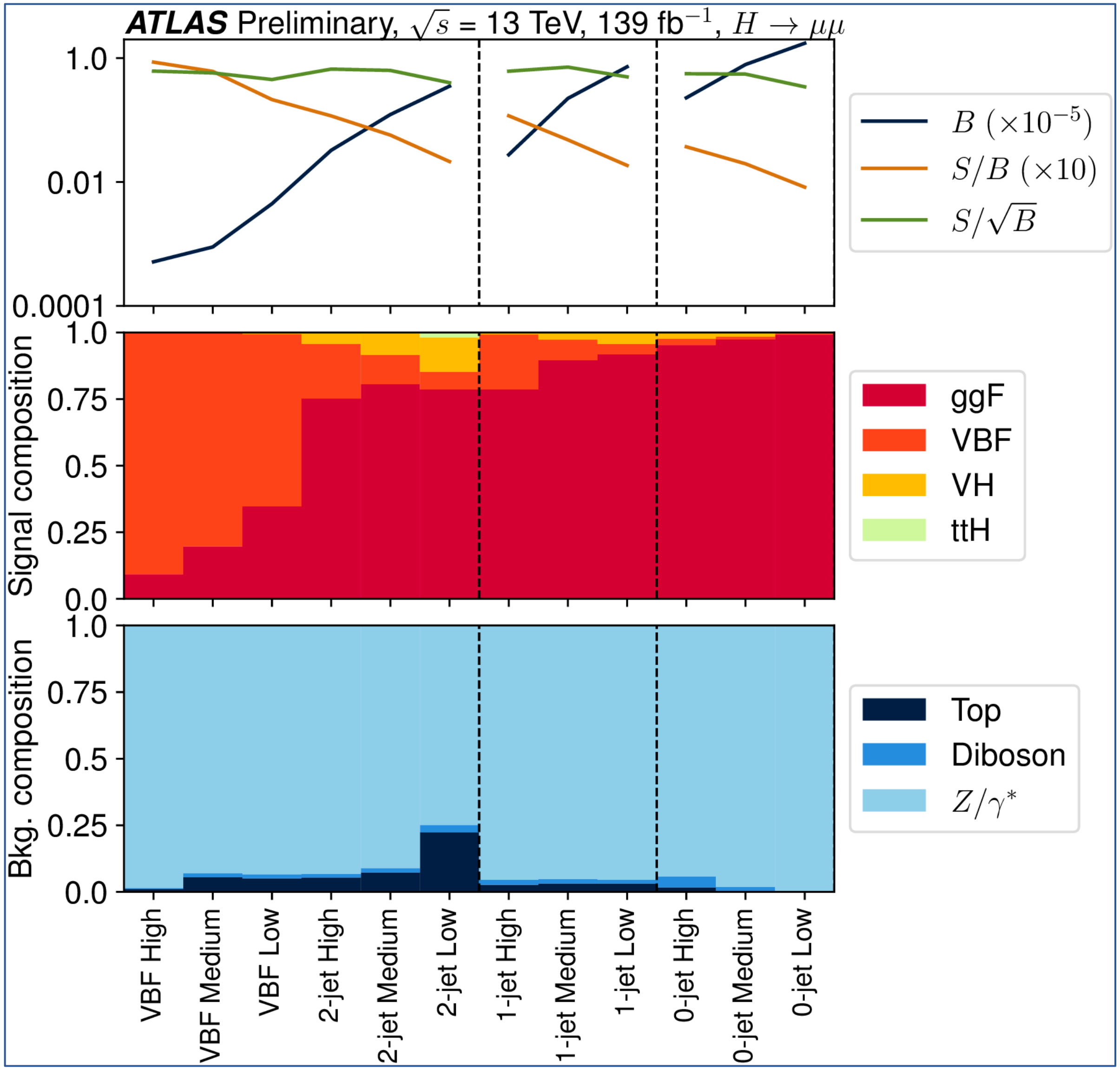
- ➢ Higgs classifier target ggF+VBF: di-muon variables and jet kinematics where available
- ➢ VBF classifier target VBF for 2 jets
- ➢ Optimize category boundaries to maximize expected significance





See backup for more details

Events categorization summary



S/B in range: 0.08%-12%

Signal composition

Background composition

Background model

Background model was developed with high-stat. dedicated fast Drell-Yan simulation (~100 ab^{-1})

Background PDF = DY×Empirical function

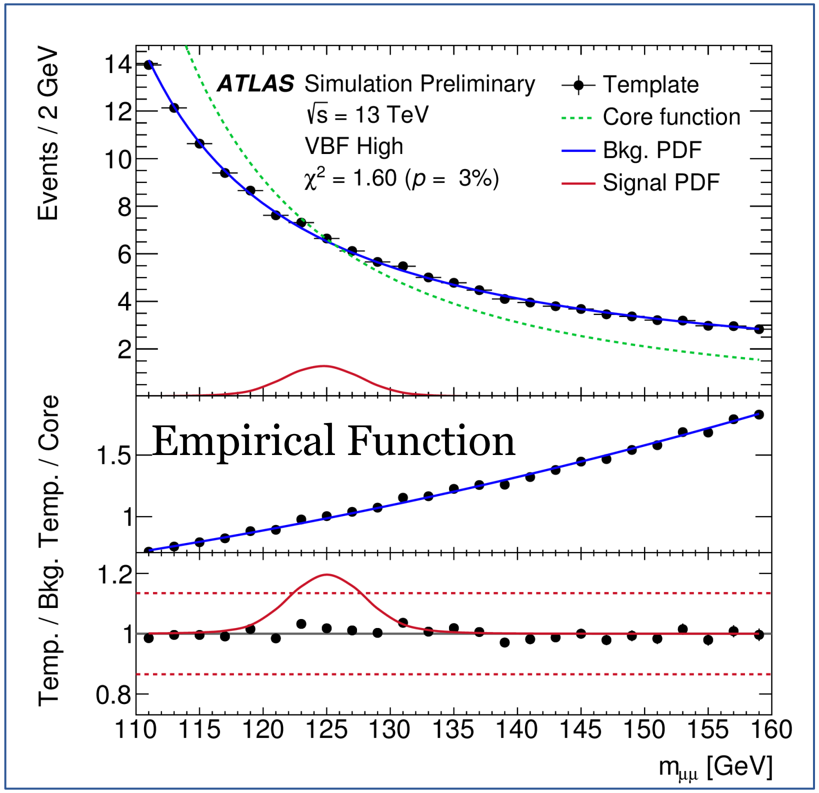
LO Drell-Yan Core function

✓NLO Powheg (0/1 jet) and LO Alpgen (2 jet).

✓without free parameters - same in each category.

Empirical function

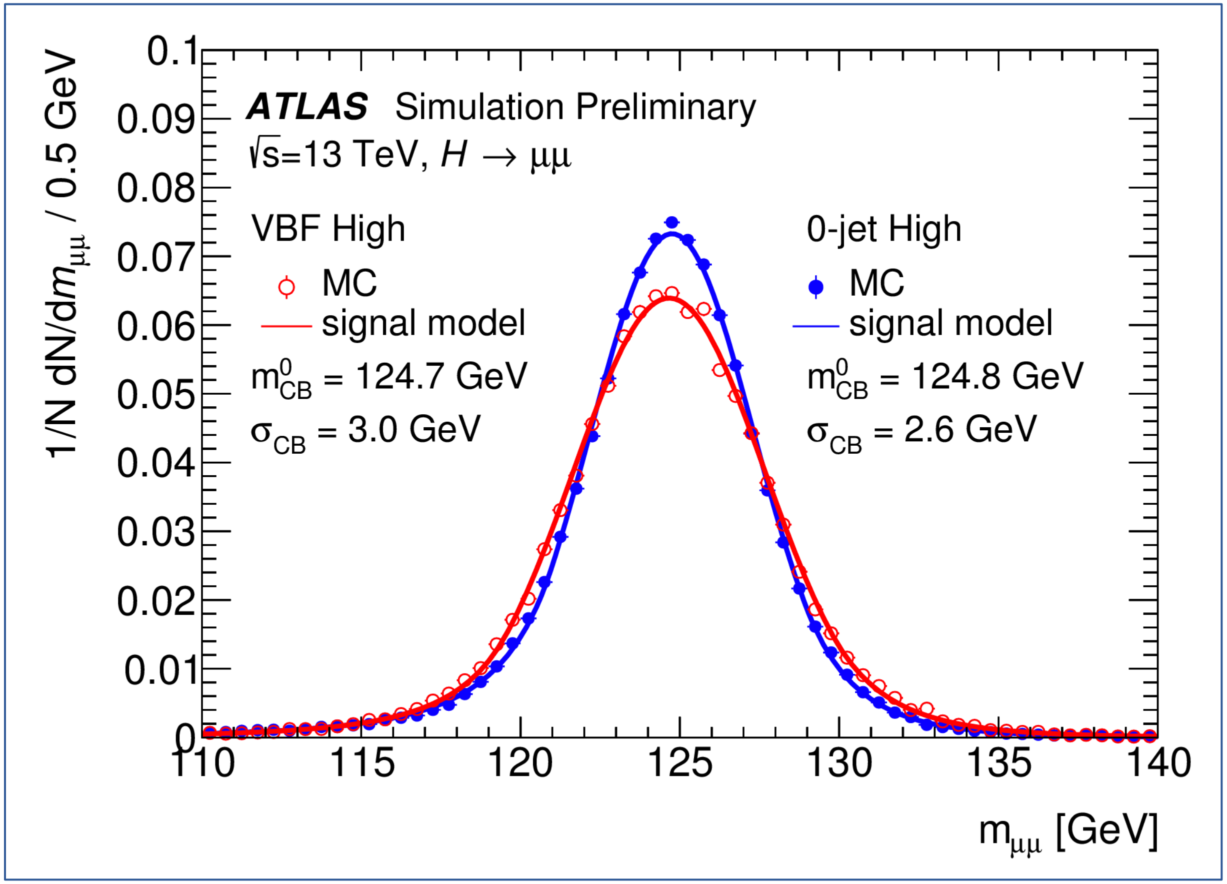
✓power-law or exponential of polynomial with parameters constrained by data in simultaneous fit - different in each category.



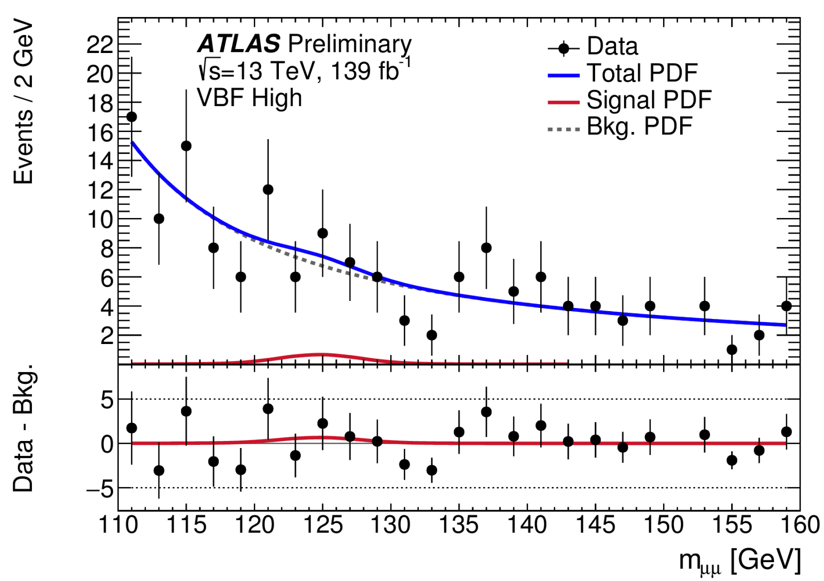
See backup for more details

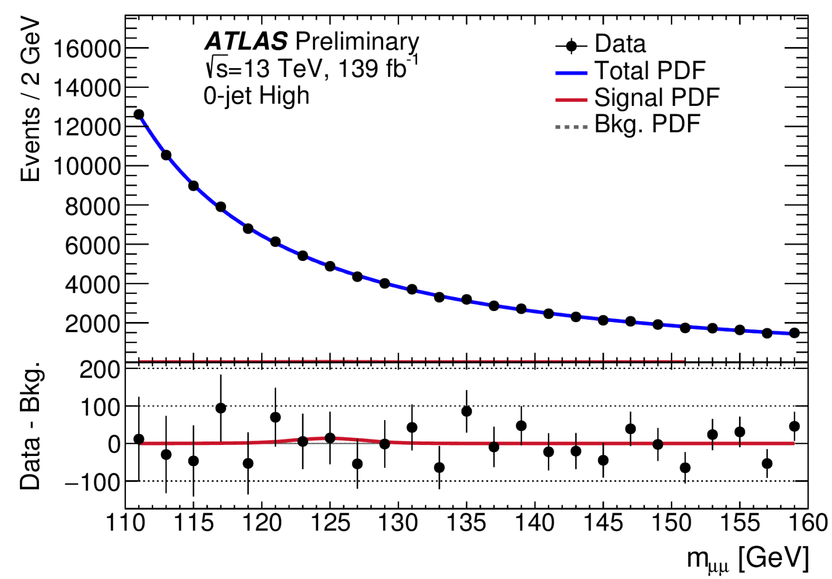
Signal model

- Signal is predicted to be a narrow resonance with a width of 4.1 MeV for m_H = 125.09.
- Signal model fitted in each category to signal MC (double-sided Crystal Ball function)
- Individual simultaneous S + B fit in all 12 categories



Signal and background yields





Category	Data	S_{SM}	S	B	$S/\sqrt{B}$	S/B [%]
VBF High	40	4.5	2.3	34	0.39	6.6
VBF Medium	109	5.5	2.8	100	0.28	2.8
VBF Low	450	9.6	4.9	420	0.24	1.2
2-jet High	3400	38	19	3440	0.33	0.6
2-jet Medium	13938	70	35	13910	0.30	0.3
2-jet Low	40747	75	38	40860	0.19	0.1
1-jet High	2885	32	16	2830	0.31	0.6
1-jet Medium	24919	107	54	24890	0.35	0.2
1-jet Low	77482	134	68	77670	0.24	0.1
0-jet High	24777	85	43	24740	0.27	0.2
0-jet Medium	85281	155	79	85000	0.27	0.1
0-jet Low	180478	144	73	180000	0.17	<0.1

Signal Strength

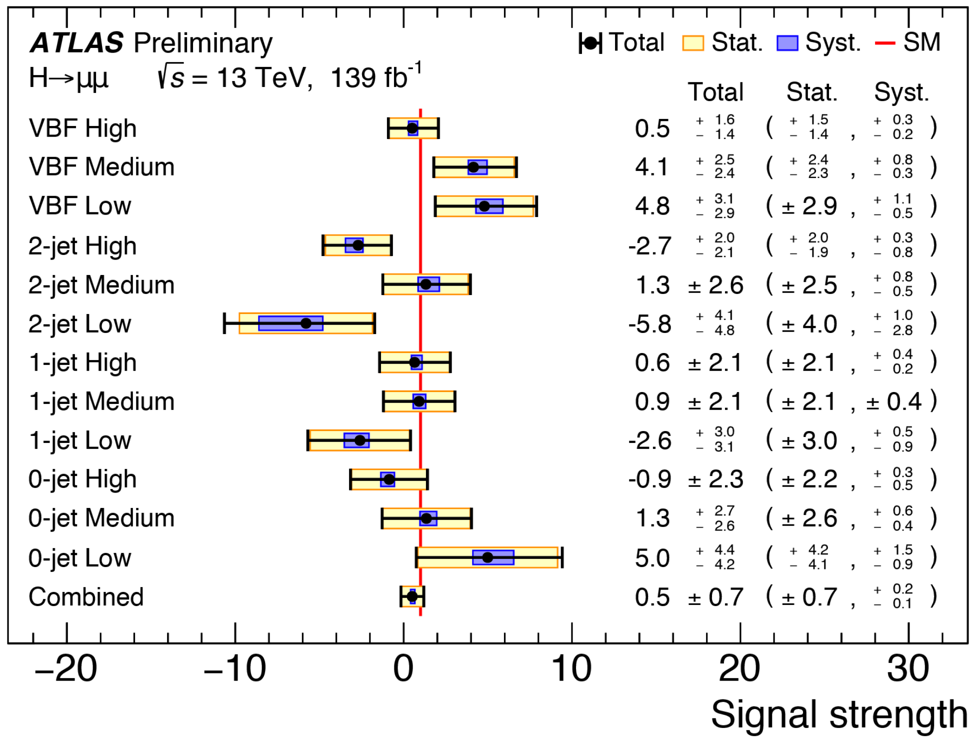
Uncertainties on μ:

➢data statistics: 0.7 - ***dominant***

➢Theory signal systematics: 0.08

➢Experimental signal systematics: 0.07

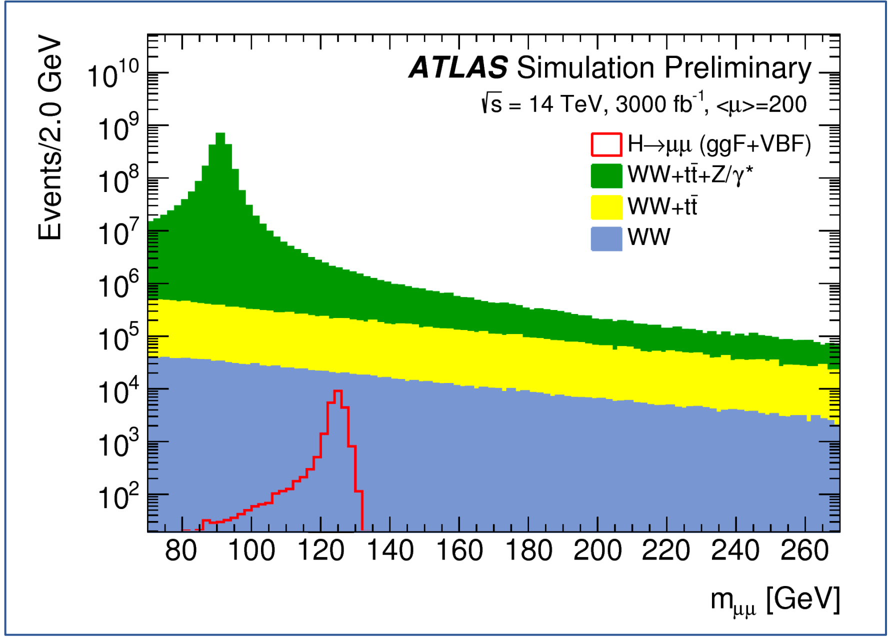
➢Background modelling: 0.06



Compatibility of the measured signal strengths between the 12 categories is at the level of 33%

Prospects HL LHC

- ✓ The H→μμ signal from ggF and VBF is expected to be observed with a significance of more than 9σ.
- ✓ The uncertainty on the Higgs production cross section times the branching ratio to dimuons is expected to be around 13%.

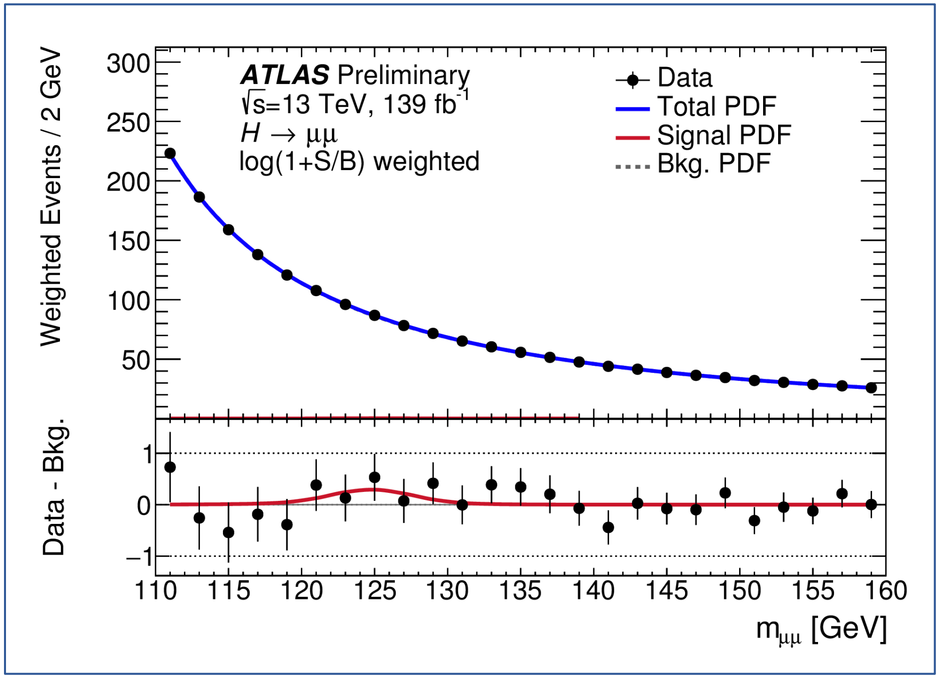


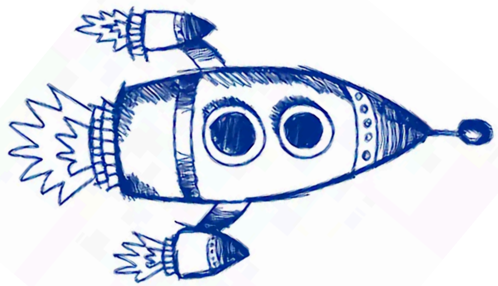
Category	S	**VBF**	B	**FWHM [GeV]**	σ_G [GeV]	$S/\sqrt{S+B}$
VBF-like	386	197	19430	4.37	1.88	2.75
low p_T, central	921	11	350500	3.21	1.37	1.55
med p_T, central	2210	84	300500	3.08	1.32	4.01
hi p_T, central	1810	242	211800	3.50	1.56	3.91
low p_T, non central	2460	28	1740500	4.11	1.79	1.86
med p_T, non central	5860	230	1483600	4.24	1.80	4.80
hi p_T, non central	4380	588	829000	4.70	1.92	4.80
Total	18020	1380	4935500	3.93	1.69	9.53

resolution of the core function

Summary

- Significance: 0.8σ (1.5σ) obs. (exp.)
- 50% improvement in sensitivity over 80 fb^{-1} result
- Signal strength: μ= 0.5 ±0.7
- Upper limit: BR($H \to \mu^-\mu^+$) < $1.7\times\sigma^{SM}$ (for expected $<1.3\times\sigma^{SM}$)



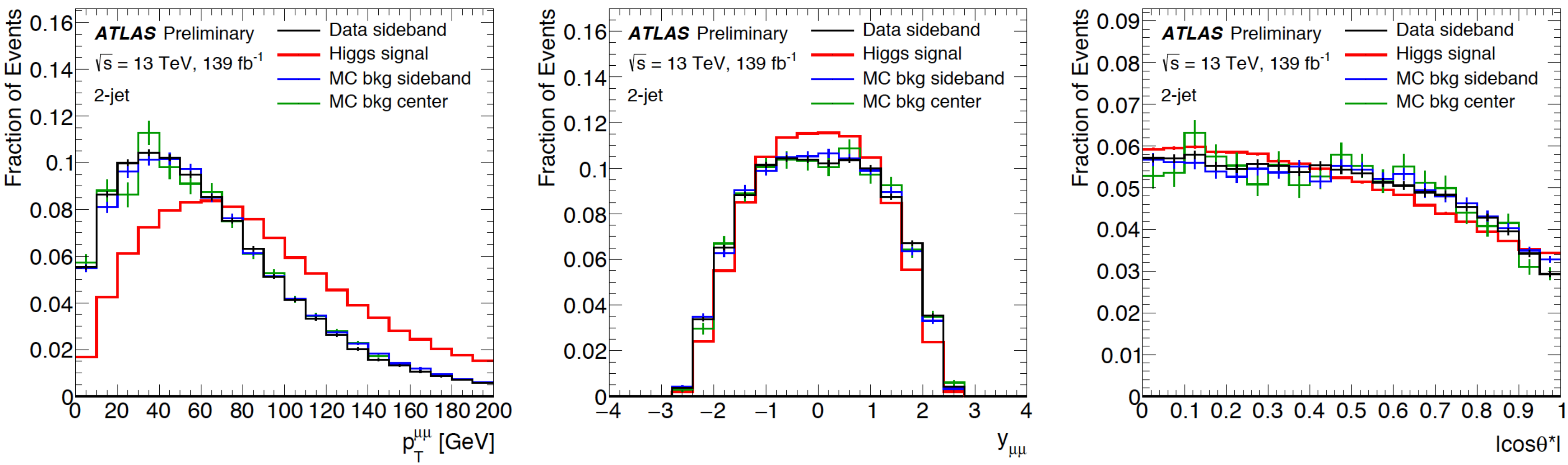


Stay tuned for future $H \to \mu^-\mu^+$ results!

Backup

Training variables

- 0-jet: pTμμ, Yμμ, |cos(θ*)|
- 1-jet: 0-jet variables + pTj1, ηj1, Δφj1,μμ
- ≥ 2-jet: 1-jet variables + pTj2, ηj2, Δφj2,μμ, pTjj, Yjj, Δφjj,μμ, mμμ, MET



Empirical function for background model

Function	Expression
PowerN	$m_{\mu\mu}^{(a_0+a_1 m_{\mu\mu}+a_2 m_{\mu\mu}^2+...+a_N m_{\mu\mu}^N)}$
EpolyN	$\exp(a_1 m_{\mu\mu} + a_2 m_{\mu\mu}^2 + ... + a_N m_{\mu\mu}^N)$

Category	Empirical Function	max(SS/ δS)[%]	max(SS/S_{SM})[%]
VBF High	Power0	10.6	14.7
VBF Medium	Epoly2	0.51	1.3
VBF Low	Power1	3.6	7.5
2-jet High	Epoly2	8.7	16.3
2-jet Medium	Epoly4	1.2	3.9
2-jet Low	Epoly3	-8.2	-33.2
1-jet High	Power1	6.1	12.1
1-jet Medium	Epoly3	-8.1	-19.8
1-jet Low	Epoly3	-2.5	-5.8
0-jet High	Power1	14.6	26.5
0-jet Medium	Epoly3	-11.6	-39.0
0-jet Low	Epoly3	-18.5	-74.2

The functions are selected so that the spurious signal systematics become < 20% of the statistical uncertainty in each category.

BDT categories boundaries optimization

Category	0-jet	1-jet	VBF	2-jet $O_{\mathrm{VBF}} < 0.60$
High	$O^0_{ggF} \geq 0.75$	$O^1_{ggF} \geq 0.78$	$O_{\mathrm{VBF}} \geq 0.89$	$O^2_{ggF} \geq 0.48$
Medium	$0.35 \leq O^0_{ggF} < 0.75$	$0.38 \leq O^1_{ggF} < 0.78$	$0.77 \leq O_{\mathrm{VBF}} < 0.89$	$0.22 \leq O^2_{ggF} < 0.48$
Low	$O^0_{ggF} < 0.35$	$O^1_{ggF} < 0.38$	$0.60 \leq O_{\mathrm{VBF}} < 0.77$	$O^2_{ggF} < 0.22$

To maximize the total number counting significance

$$Z_{\mathrm{tot}} = \sqrt{\sum_i Z_i^2},$$

$$Z_i = \sqrt{2\left((s_i + b_i)\log\left(\frac{s_i + b_i}{b_i}\right) - s_i\right)},$$

ATLAS vs CMS results

	ATLAS	CMS
significance (expected)	0.8σ (1.5σ)	0.9σ (1.0σ)
signal strength	0.5 ±0.7(stat) +0.2/-0.1(syst)	1.0 ± 1.0 (stat) ± 0.1 (syst)
upper limit	<3.8 × 10⁻⁴	<6.4 × 10⁻⁴