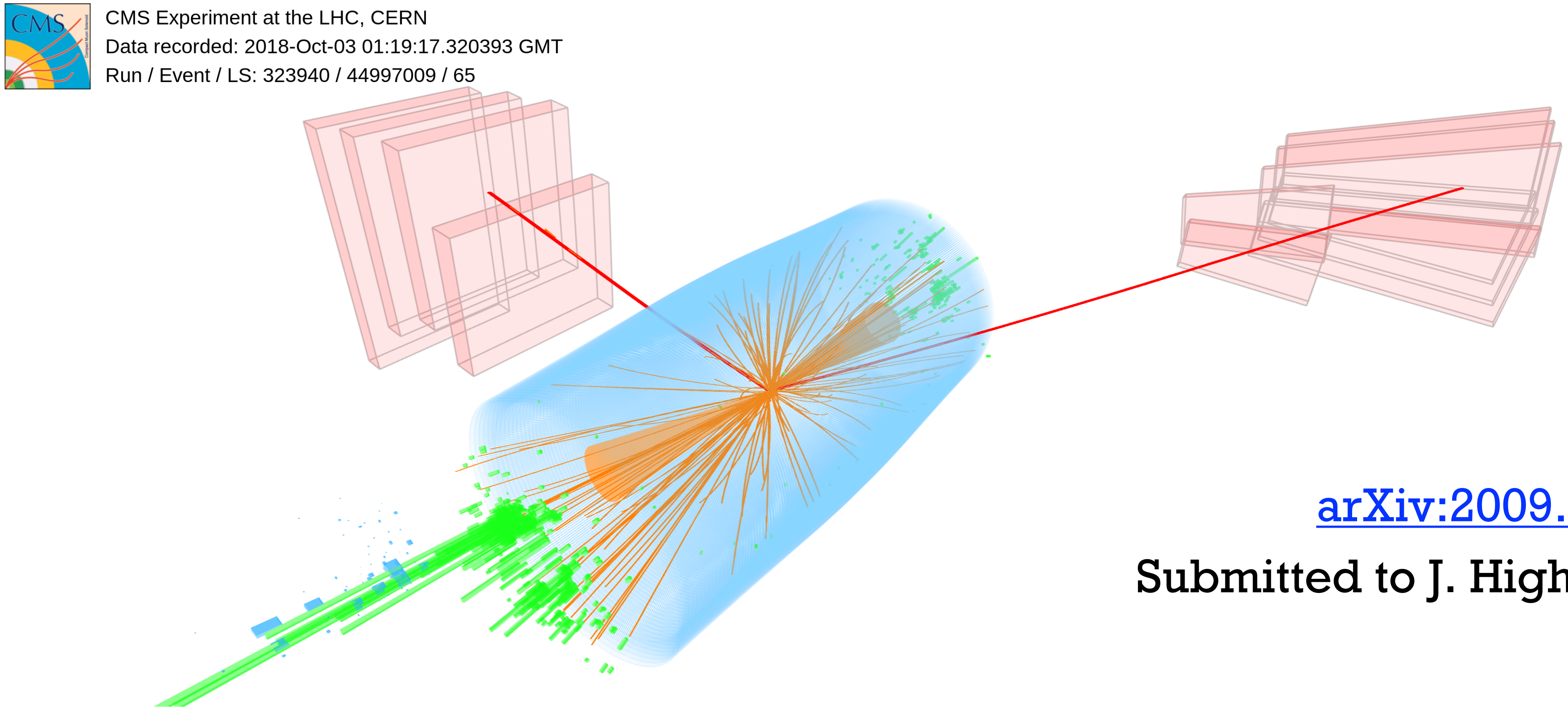
First evidence for Higgs boson decay to muons



[arXiv:2009.04363](https://arxiv.org/abs/2009.04363)

Submitted to J. High Energy Phys.

Stephane Cooperstein (UC San Diego)

LPHE Seminar, EPFL

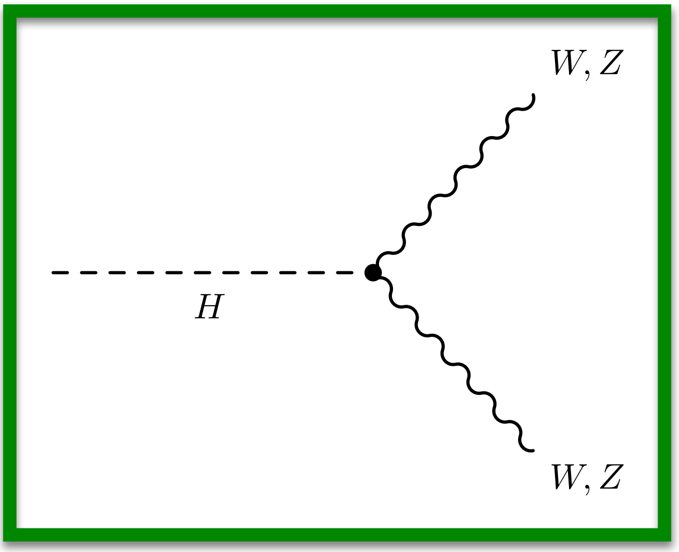
October 12, 2020

Higgs boson in the standard model

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + i \bar{\psi} \not{D} \psi + h.c. + \bar{\psi}_i y_{ij} \psi_j \phi + h.c. + \boxed{|D_\mu \phi|^2} - V(\phi)$$

- Mass generation for and Higgs interactions with gauge bosons is a direct consequence of electroweak symmetry breaking.

Bosonic Couplings



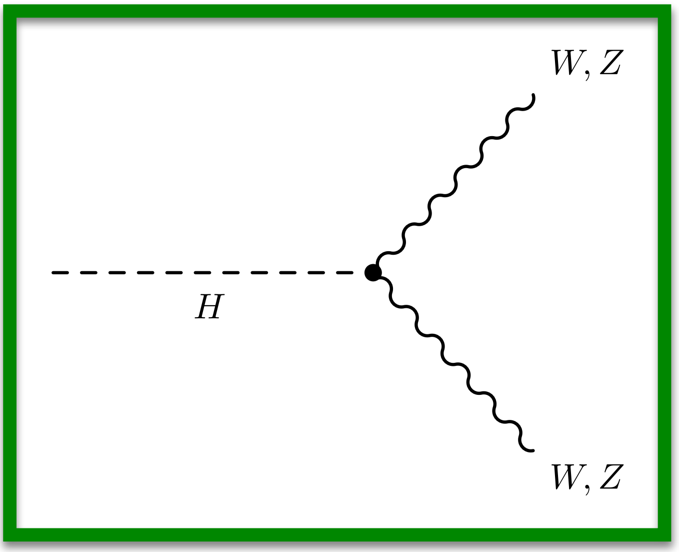
Higgs boson in the standard model

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\bar{\psi}\not{D}\psi + h.c. + \bar{\psi}_i y_{ij}\psi_j\phi + h.c. + |D_\mu\phi|^2 - V(\phi)$$

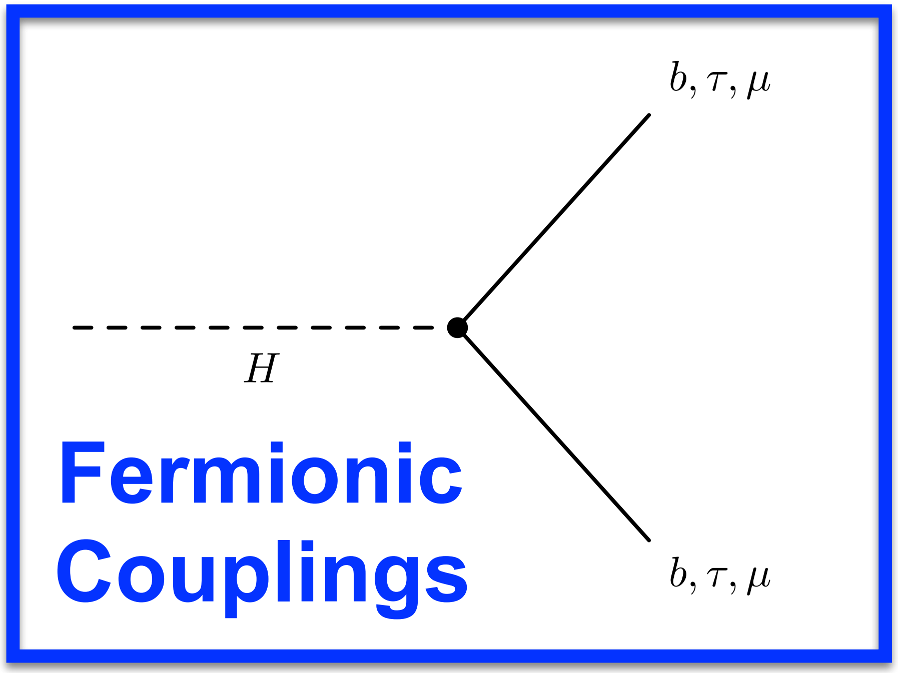
- Mass generation for and Higgs interactions with gauge bosons is a direct consequence of electroweak symmetry breaking.

- Yukawa interactions: fermion masses proportional to Higgs-fermion interaction strength.
  - Each Yukawa coupling Y_f is a free parameter of the SM.

$$L_{\text{Yuk}} = (1 + \frac{H}{v})m_f\bar{f}_L f_R$$



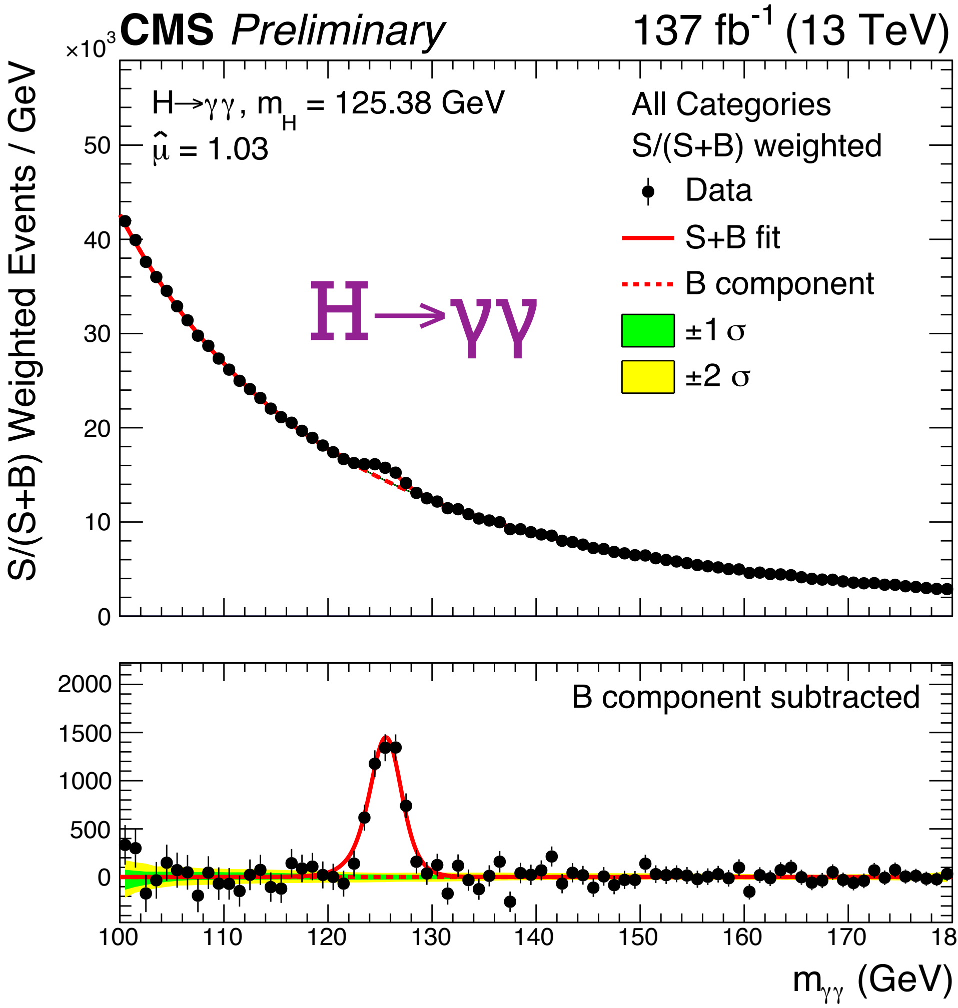
Bosonic Couplings



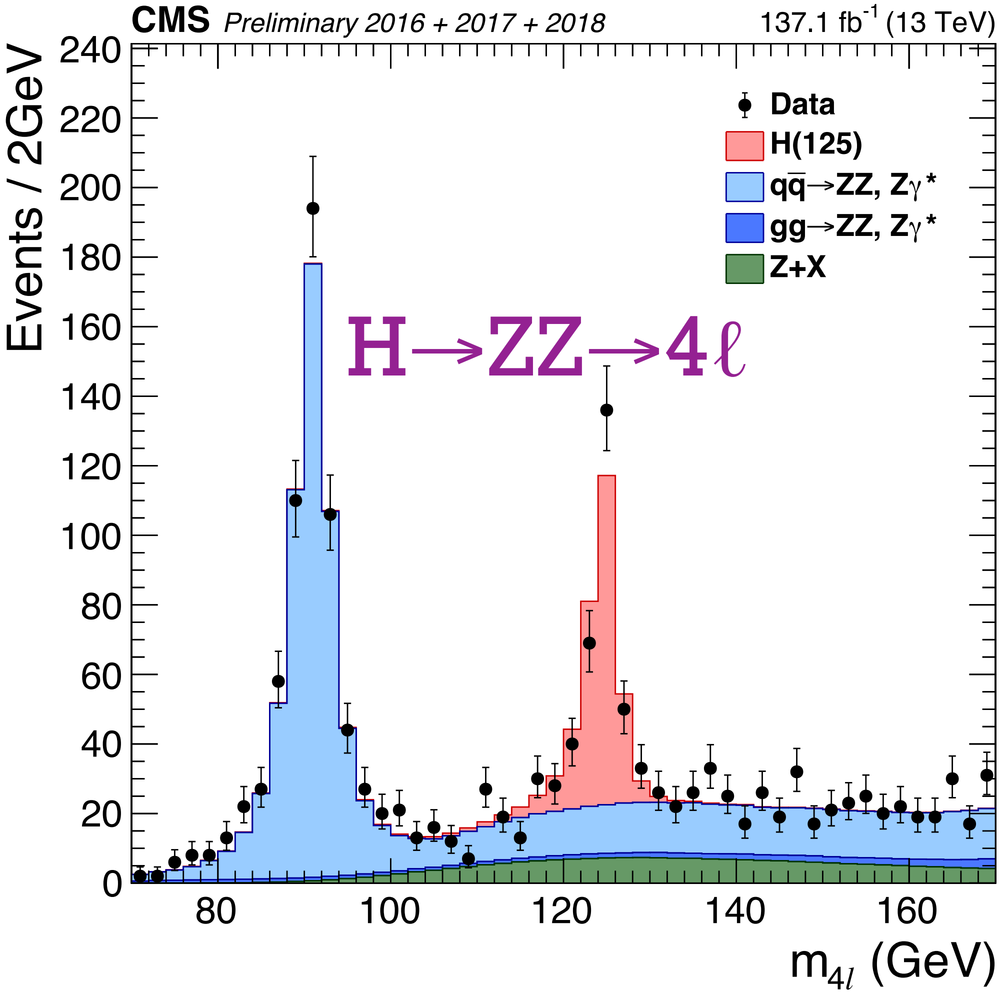
Fermionic Couplings

Higgs boson properties

- *We have come a long way in studying the properties of the Higgs boson in the eight years since its discovery.*
  - All measured properties (production rate, spin, CP, …) consistent with SM.
  - Mass measured at nearly per-mille level: $m_H = 125.38 \pm 0.14$ GeV (Phys. Lett. B 805 (2020) 135425).



CMS-PAS-HIG-19-015

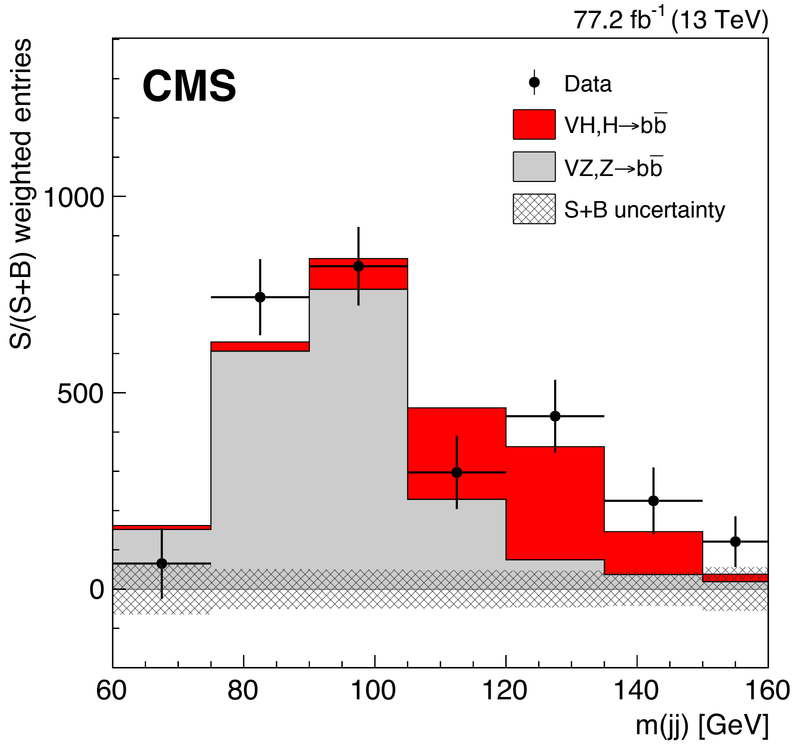


CMS-PAS-HIG-19-001

Couplings to third generation fermions

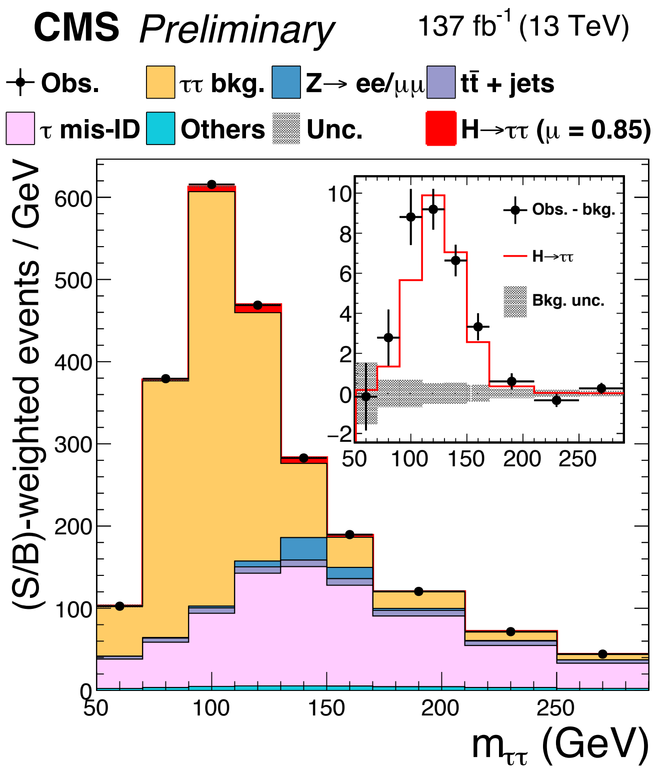
- In addition to gauge boson couplings, Higgs boson couplings to third generation fermions (t, τ, b) firmly established and consistent with SM.
- Yukawa interactions for third generation are clearly SM-like within the current experimental 10-20% precision.

$H\to b\bar{b}$



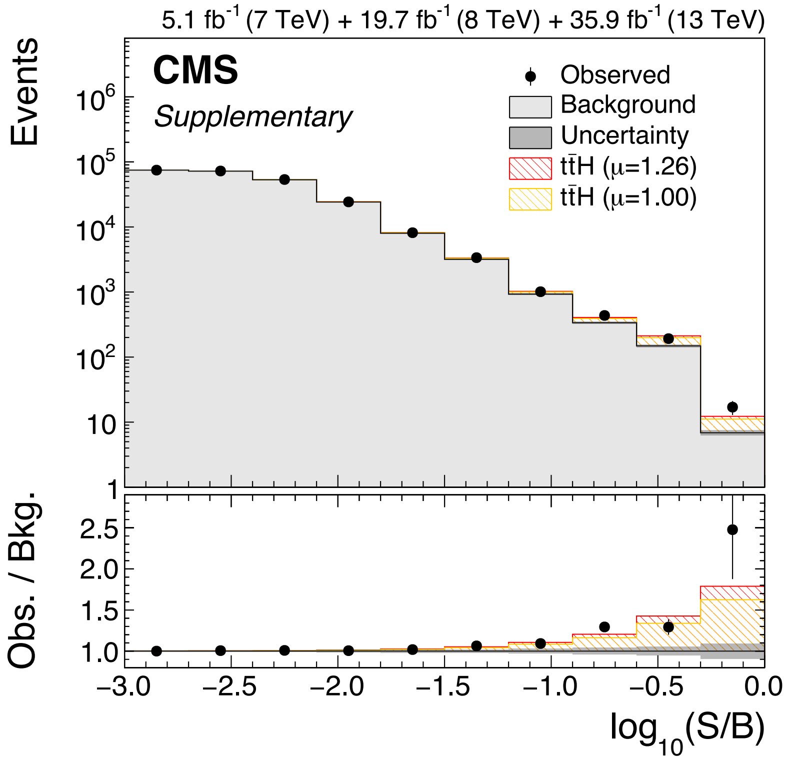
Phys. Rev. Lett. 121 (2018) 121801

$H\to\tau\tau$



CMS-PAS-HIG-19-010

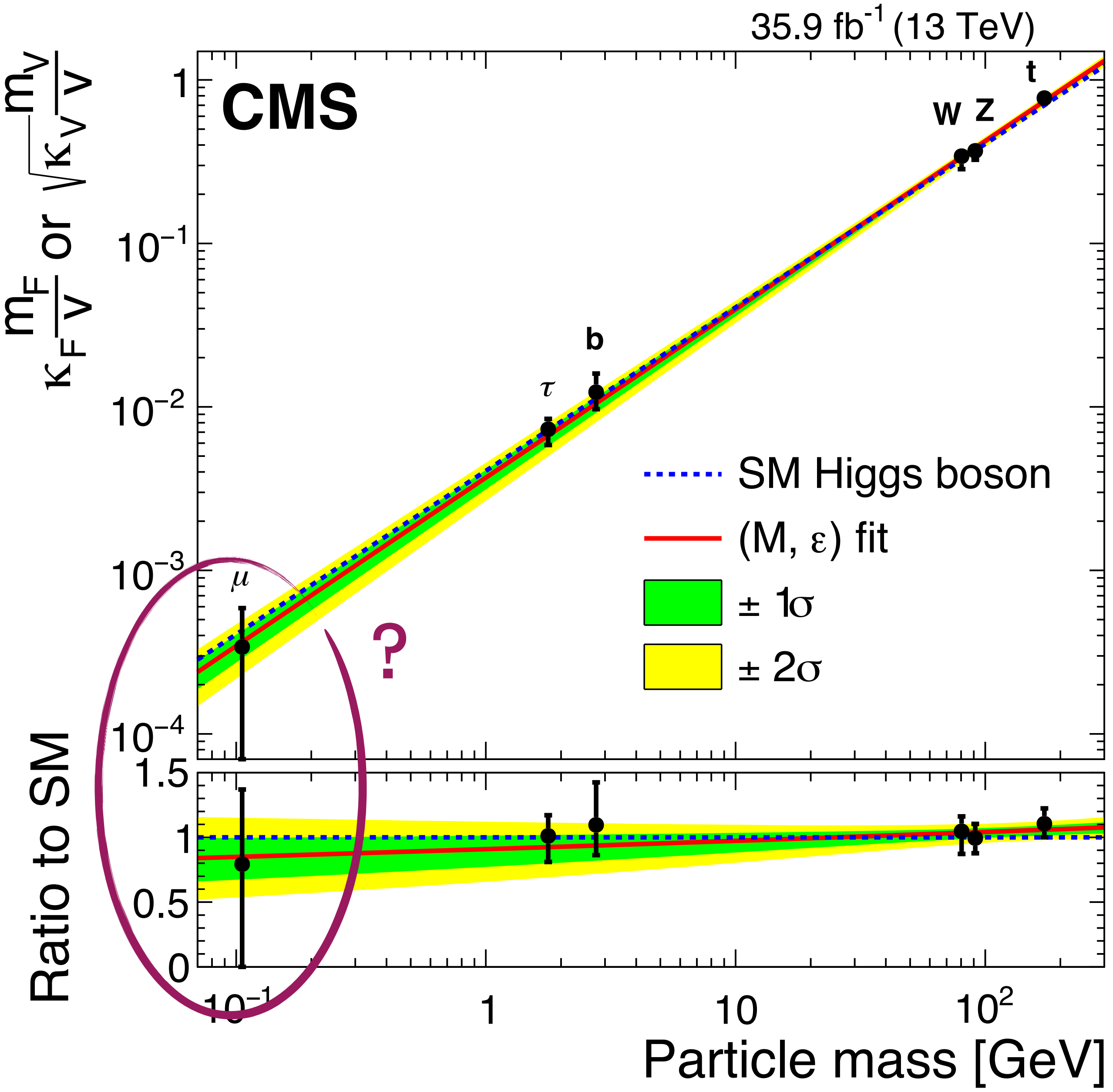
ttH



Phys. Rev. Lett. 120 (2018) 231801

Second generation couplings

- *Next frontier: Higgs boson couplings to second generation fermions*
  - **H→μμ is likely the only accessible probe of first or second generation couplings at the LHC.**
  - Extend probe of Higgs interactions by more than an order of magnitude in mass scale.



(from latest CMS Higgs couplings combination)

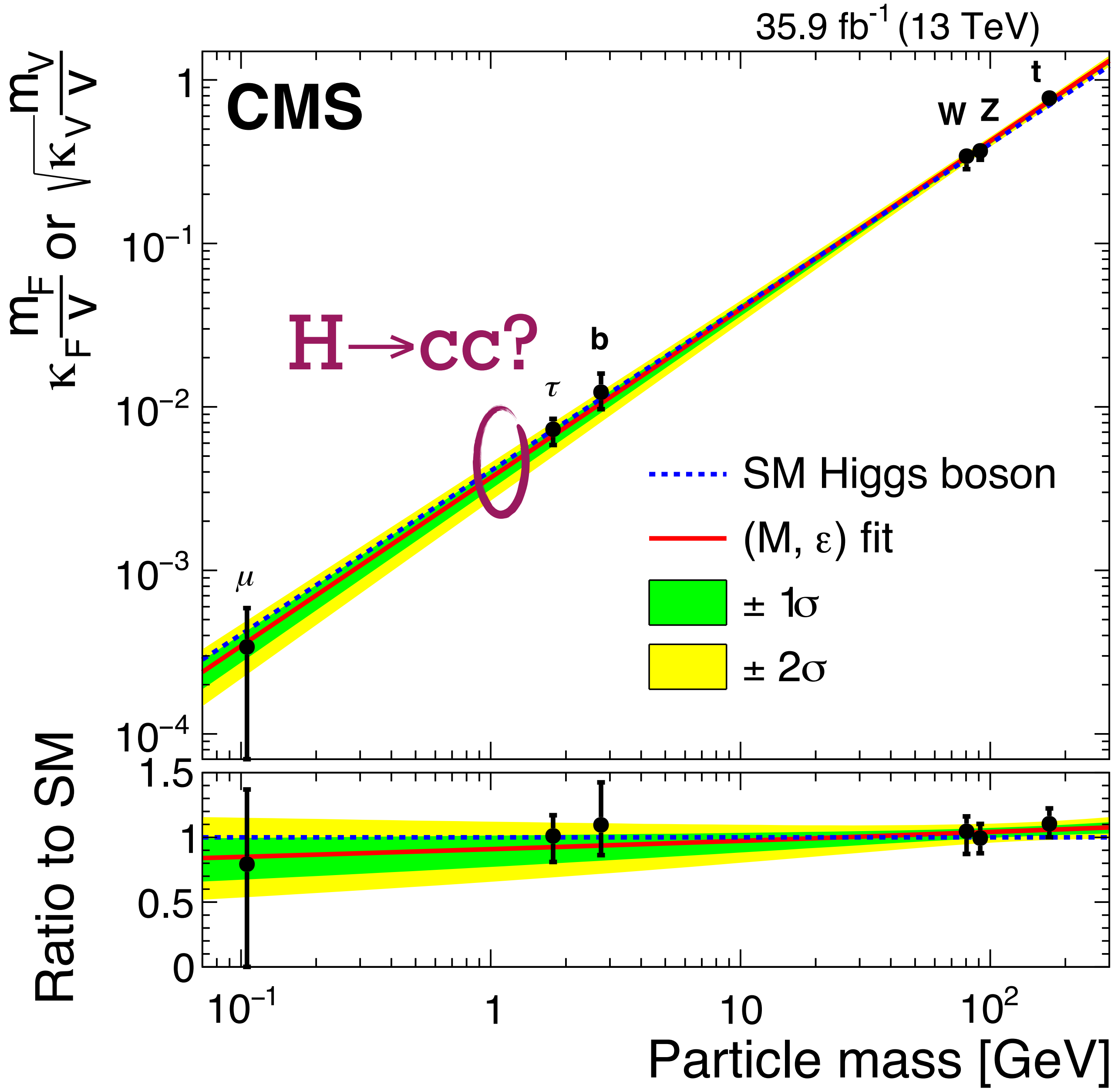
[Eur. Phys. J. C 79 (2019) 421]

Second generation couplings

- *What about H→cc?*
  - BR(H→cc) = 2.9%!

- Significant branching fraction, but very large backgrounds and difficult to isolate jets originating from charm quarks.

- Current limits are factor ~40 higher than SM with partial Run-2 data set [1].
  - ***Measuring H→cc will be very difficult, even with 3-4 ab^{-1} from HL-LHC.***

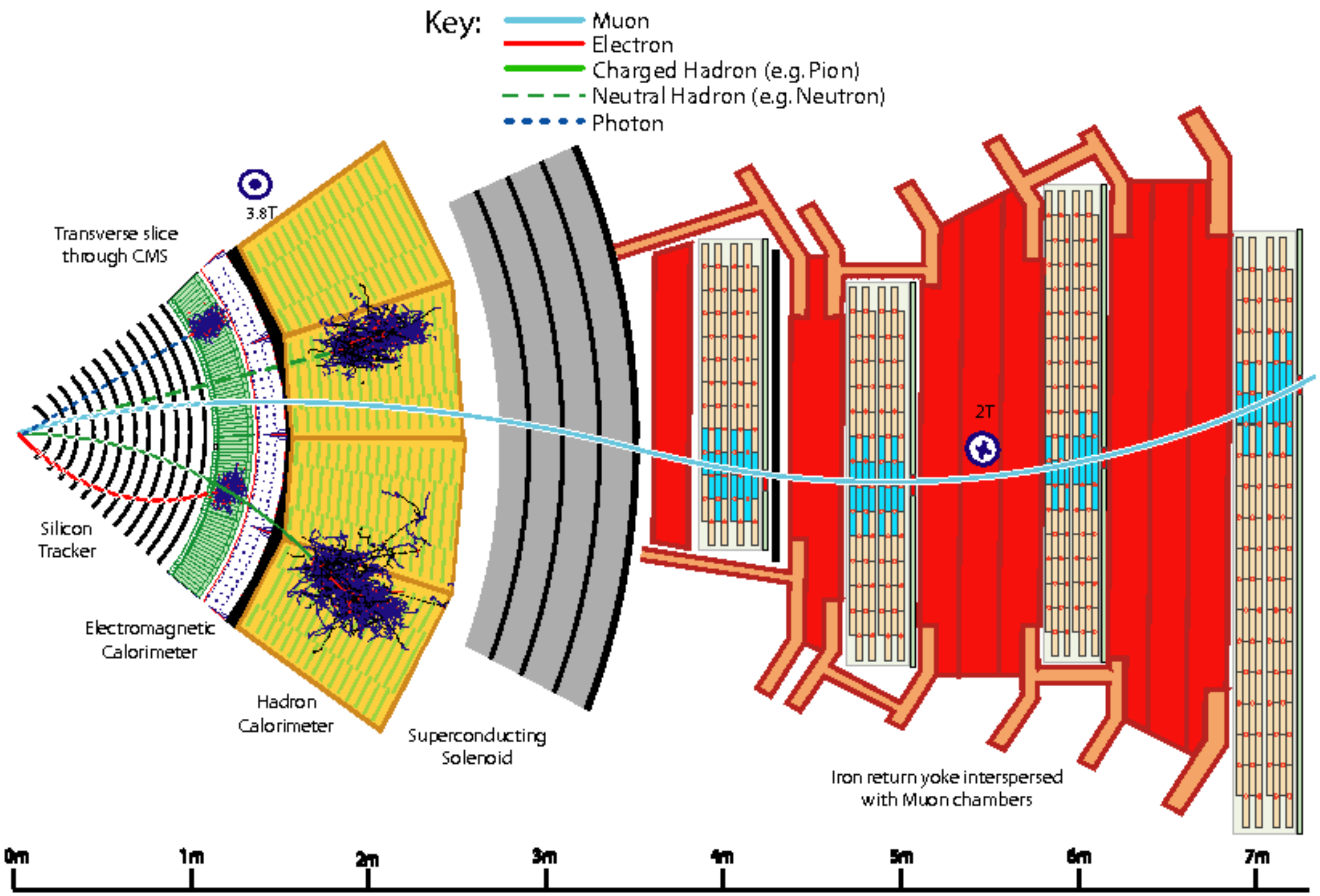


(from latest CMS Higgs couplings combination)

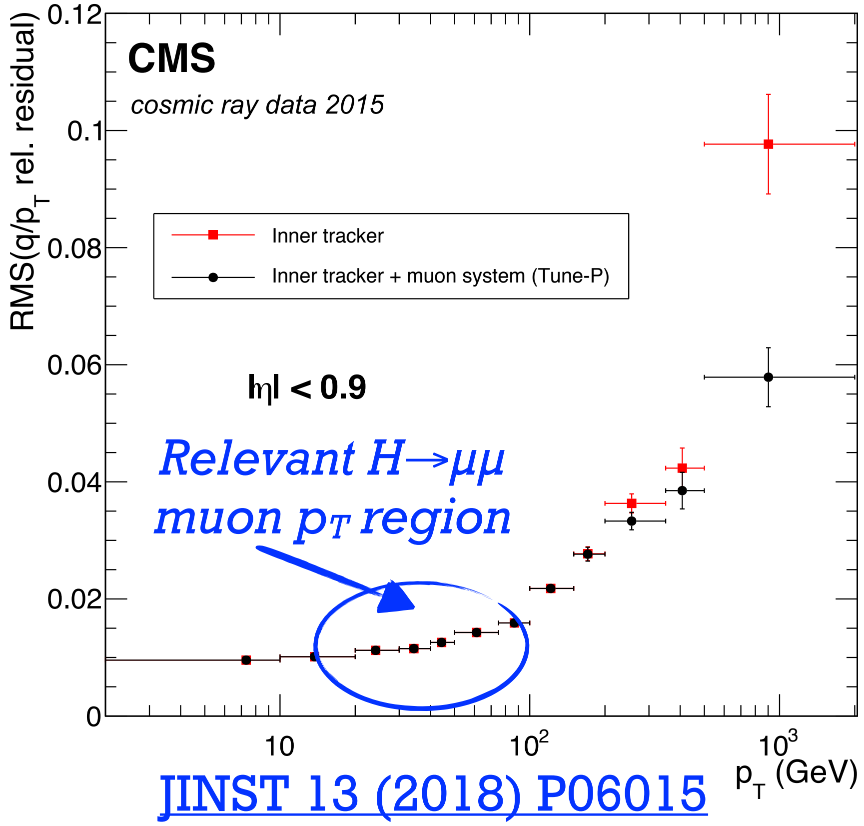
Eur. Phys. J. C 79 (2019) 421

[1] JHEP 03 (2020) 131

The CMS detector

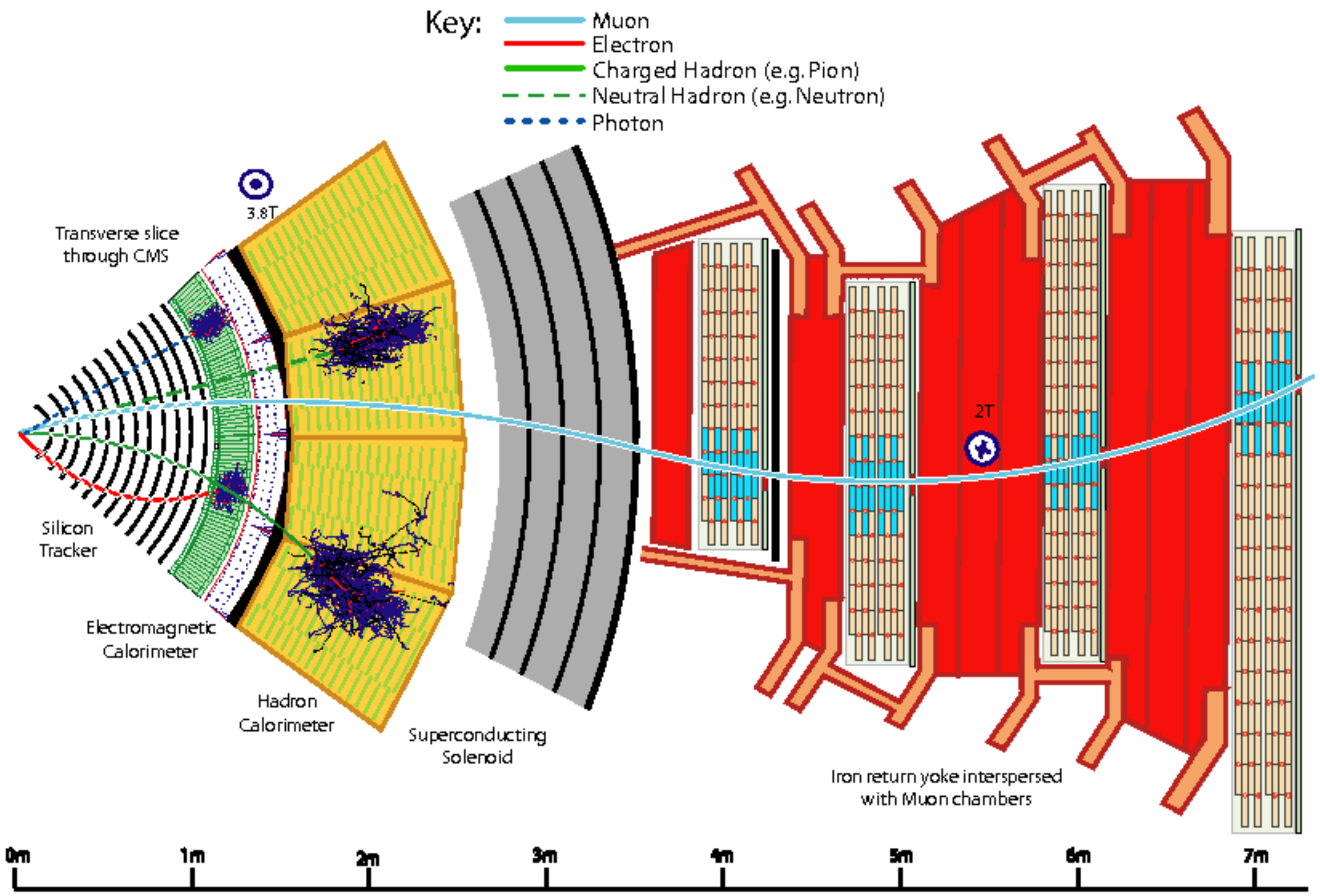


Muon momentum resolution vs $p_T(\mu)$

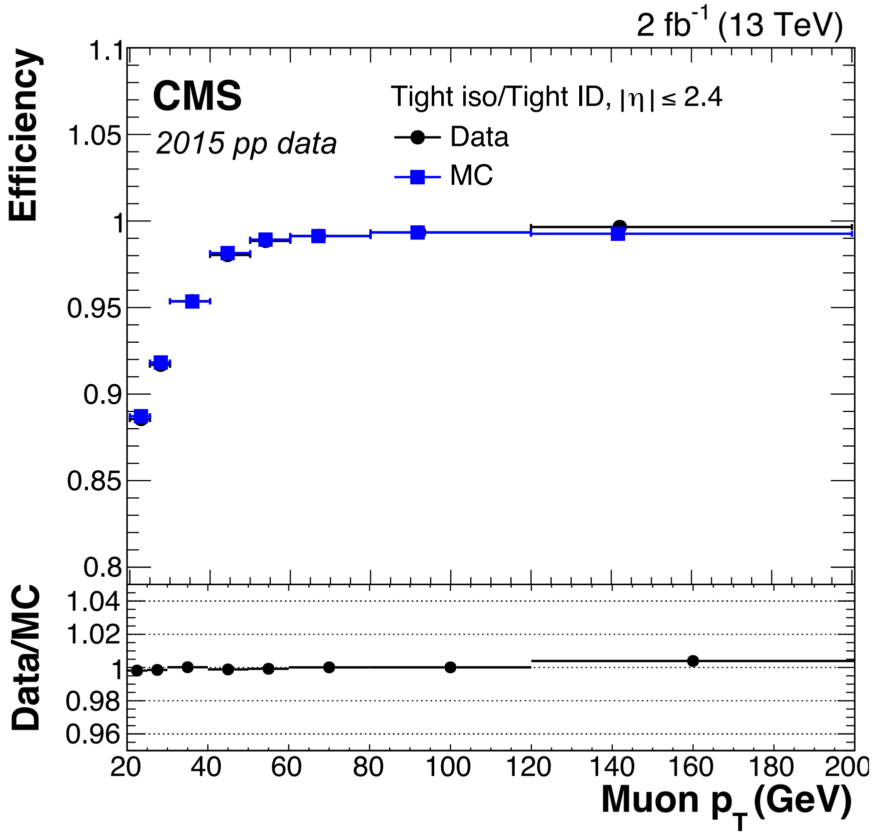


JINST 13 (2018) P06015

The CMS detector



Muon reconstruction efficiency vs $p_T(\mu)$



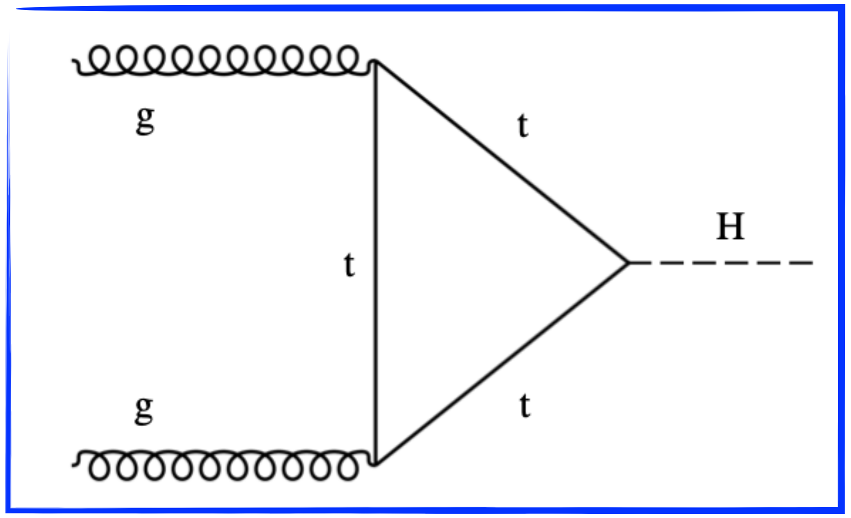
~95% selection efficiency for muons from H→μμ decays.

JINST 13 (2018) P06015

Expected signal rate

Gluon-gluon fusion production (ggH)

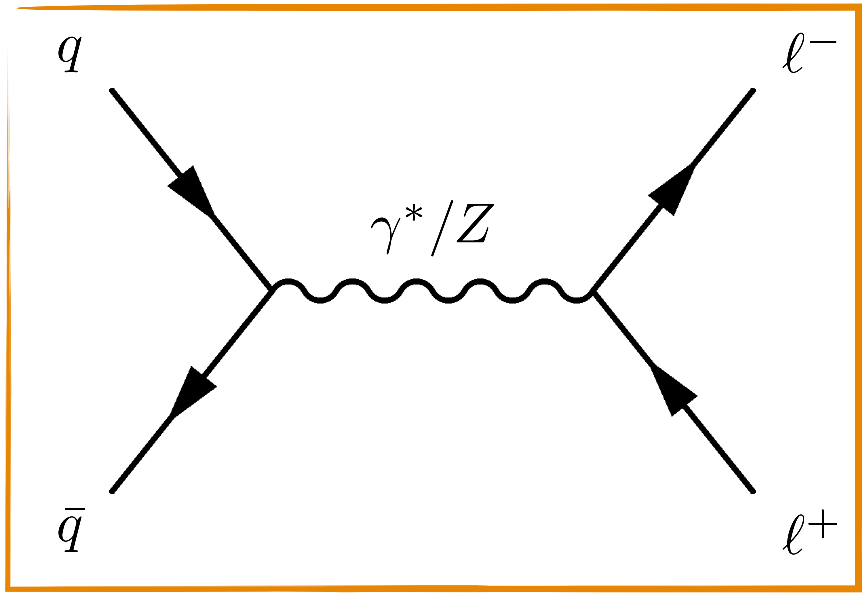
dominant Higgs boson production mode at the LHC



- Inclusive Higgs boson production cross section at 13 TeV: $\sigma_H \sim 50$ pb.
- Expected H→μμ branching fraction for m_H near 125 GeV: **BR(H→μμ) ~ $2.2*10^{-4}$**.
- $\Rightarrow \sigma_{eff}(H \rightarrow \mu\mu) \sim 0.01$ pb
- **Roughly 1k total signal events in Run-2 (137 fb^{-1}) dataset.**

The problem

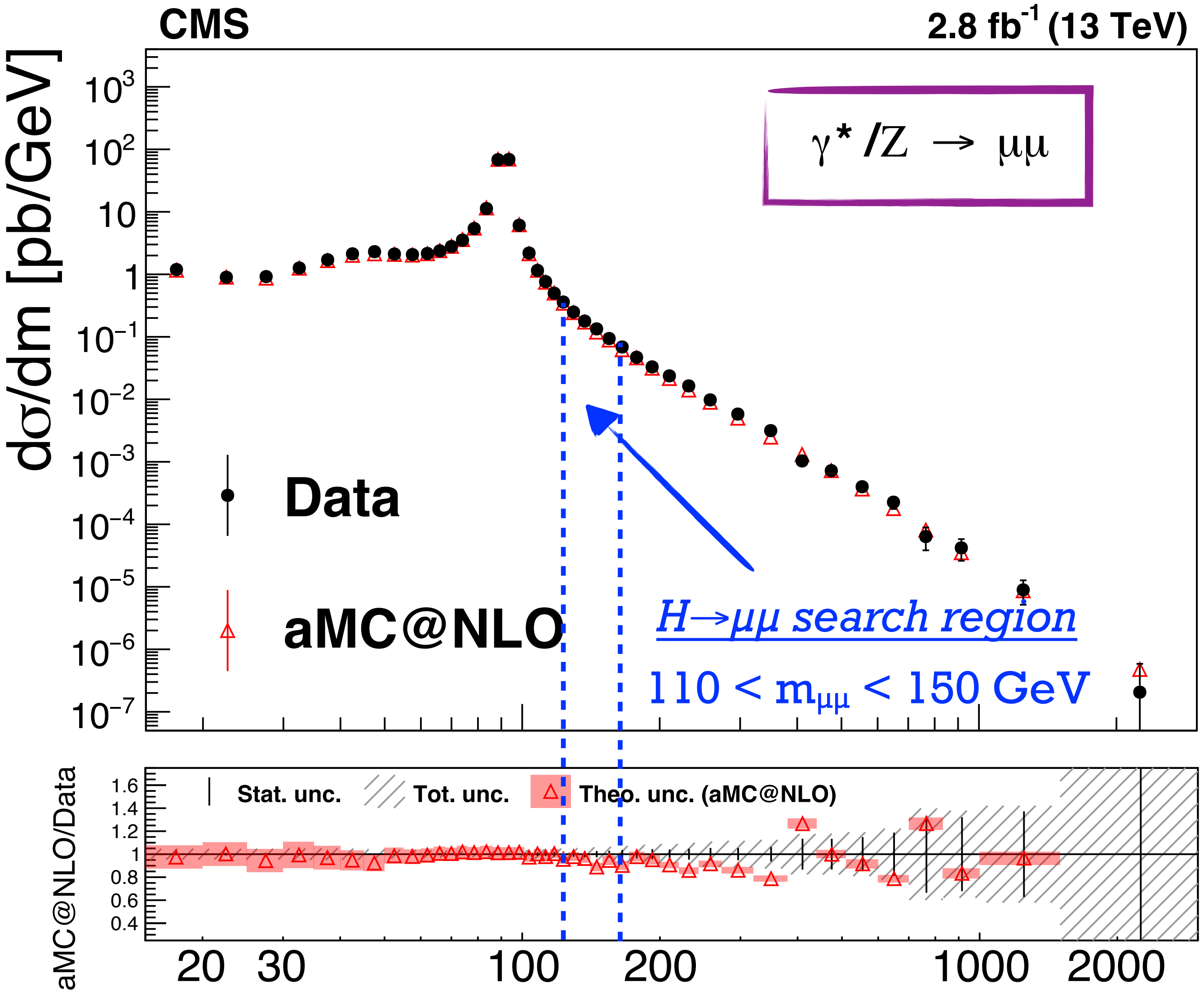
Drell-Yan (DY) production



Total effective cross section in H→μμ search region with $110 < m_{\mu\mu} < 150$ GeV:

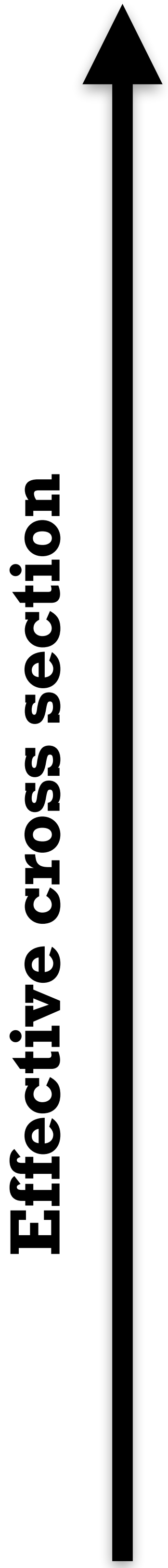
$\sigma_{\mathrm{eff}}(\mathrm{DY}) \sim 15$ pb

i.e. ~two million DY background events in H→μμ preselected 13 TeV data sample with $m_{\mu\mu}$ near 125 GeV.



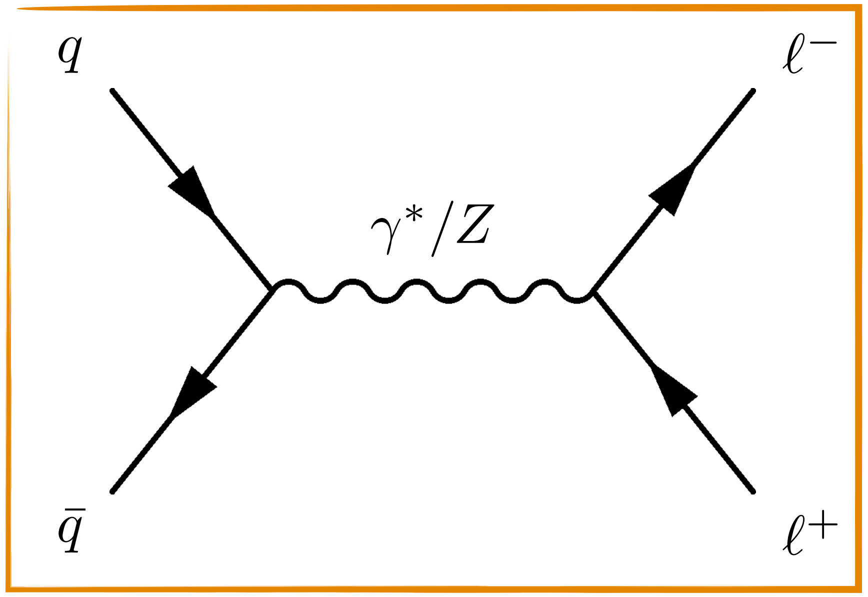
JHEP 12 (2019) 059

The challenge


Effective cross section

Drell-Yan (DY) production

$\sigma_{eff}(DY) \sim 15$ pb

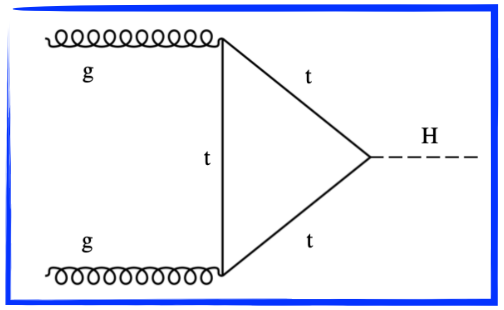


- S/B ~ one per mille with $m_{\mu\mu}$ near 125 GeV
- Large additional background rejection necessary to measure $H\rightarrow\mu\mu$ signal.

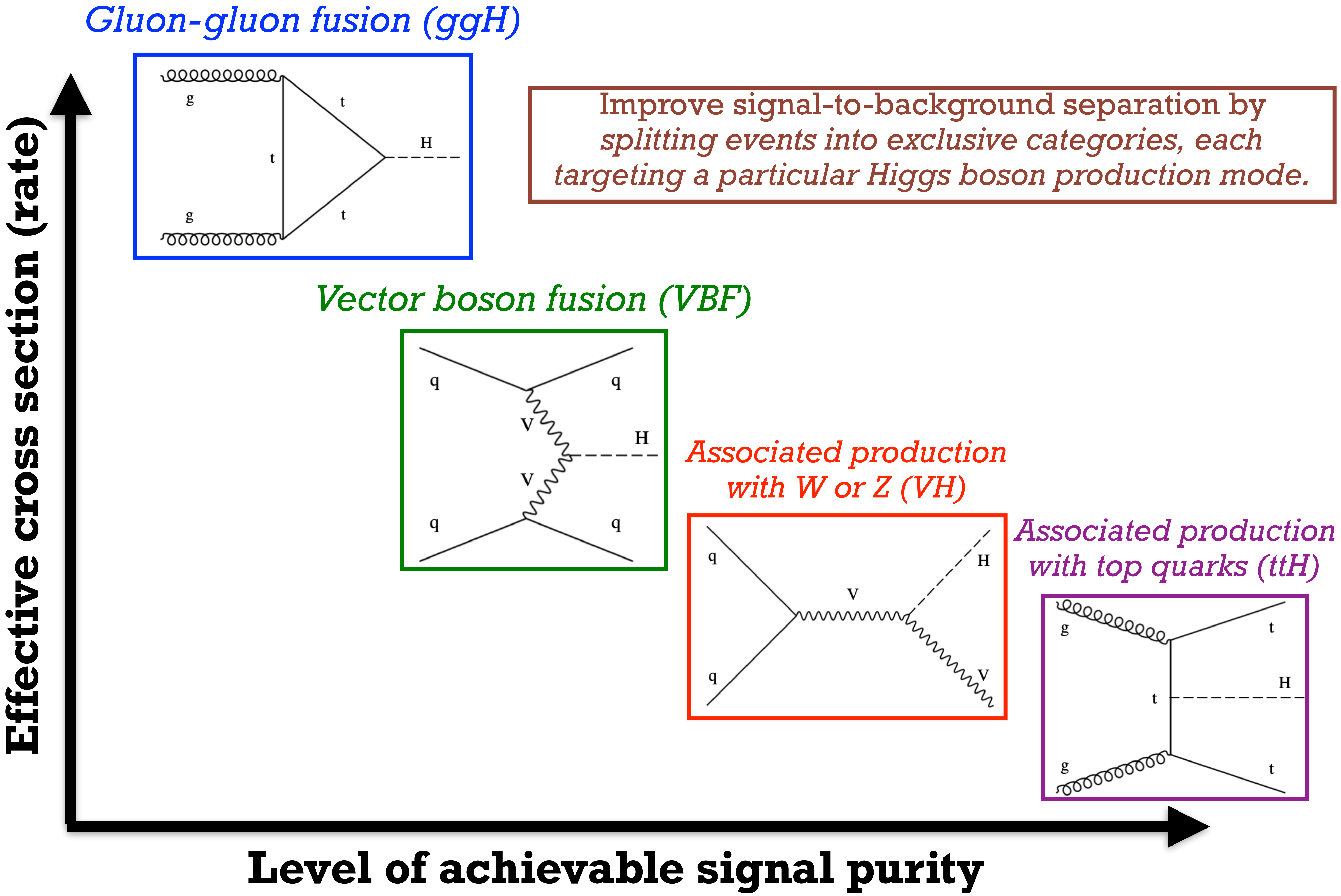
Three orders of magnitude more DY background than $H\rightarrow\mu\mu$ signal in preselected search region.

$H\rightarrow\mu\mu$ signal

$\sigma_{eff}(H\rightarrow\mu\mu) \sim 0.01$ pb



Higgs boson production modes at LHC



Gluon-gluon fusion (ggH)

Vector boson fusion (VBF)

Associated production with W or Z (VH)

Associated production with top quarks (ttH)

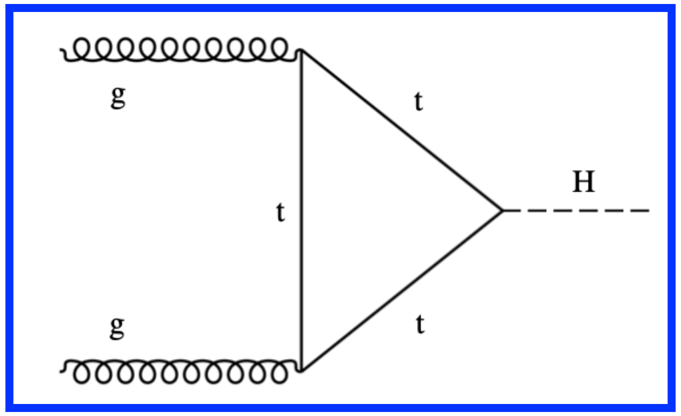
Effective cross section (rate)

Level of achievable signal purity

Improve signal-to-background separation by *splitting events into exclusive categories, each targeting a particular Higgs boson production mode.*

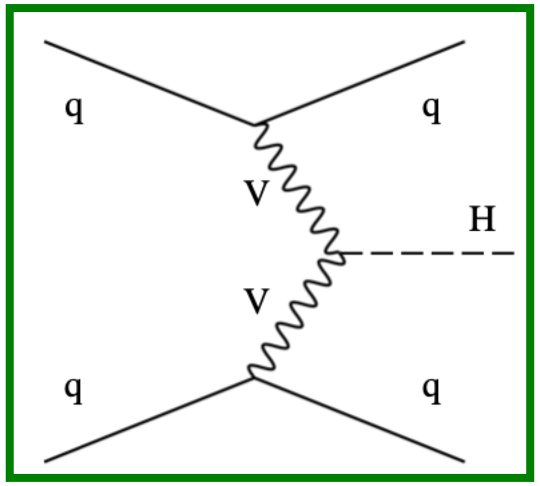
Higgs boson production modes at LHC

Gluon-gluon fusion (ggH)



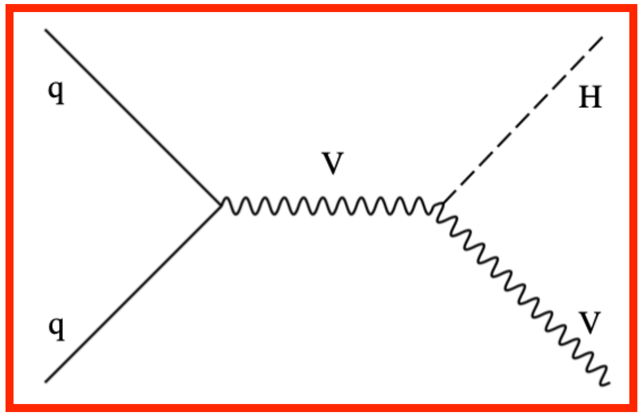
Significant improvements of existing strategy

Vector boson fusion (VBF)

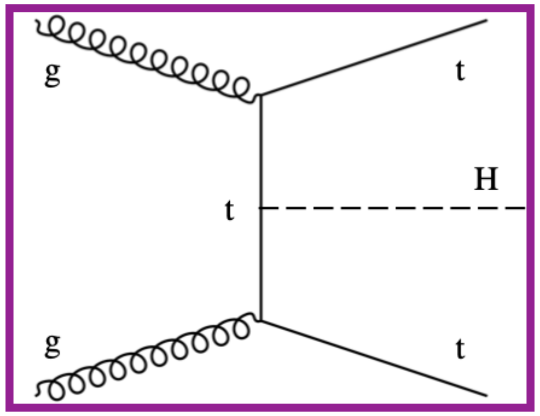


New categories

Associated production with W or Z (VH)



Associated production with top quarks (ttH)

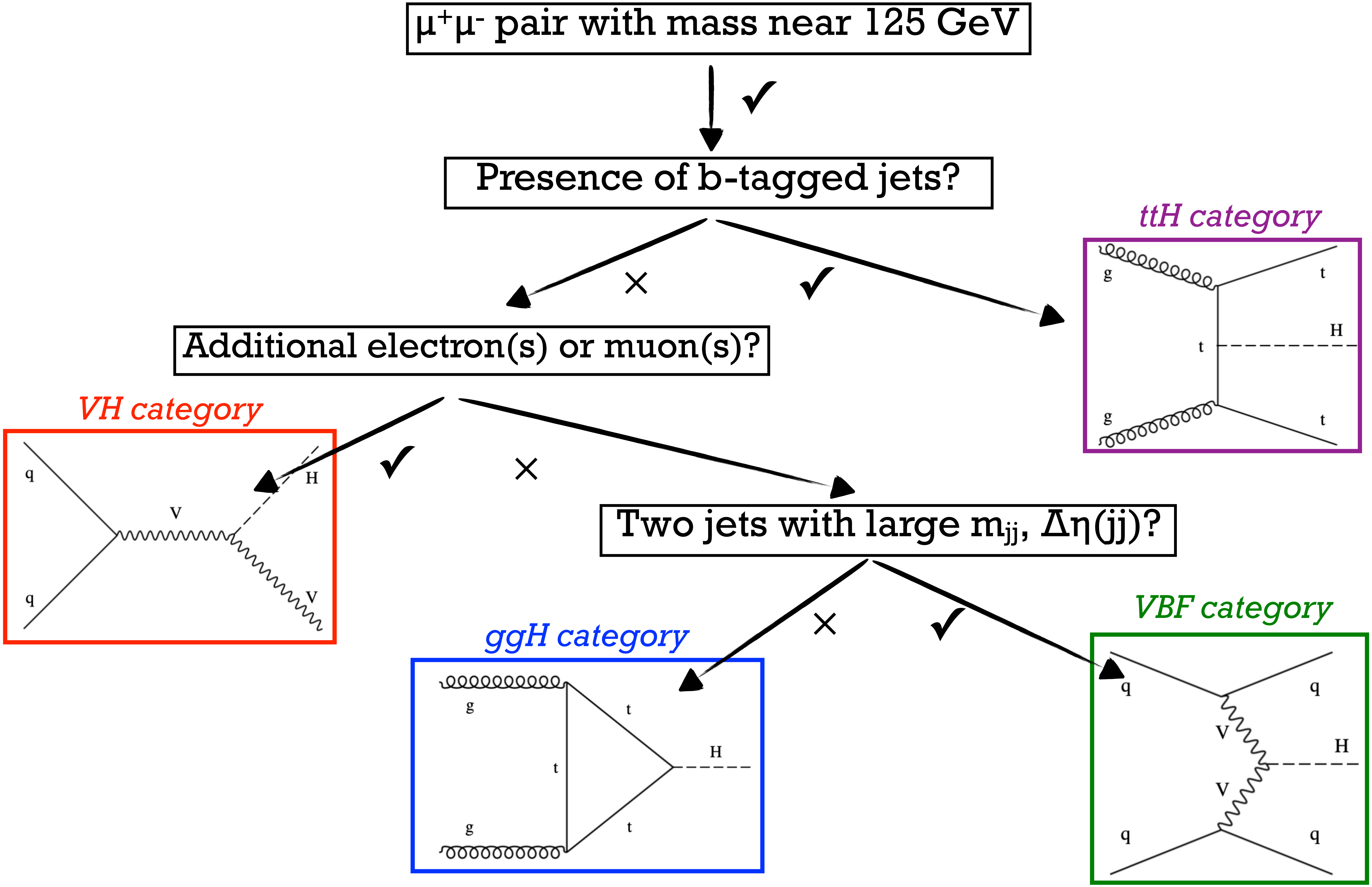


Completely redesigned strategy

Effective cross section (rate)

Level of achievable signal purity

Event selection



$\mu^+\mu^-$ pair with mass near 125 GeV

✓

Presence of b-tagged jets?

× ✓

ttH category

Additional electron(s) or muon(s)?

✓ ×

VH category

Two jets with large m_{jj}, $\Delta\eta(jj)$?

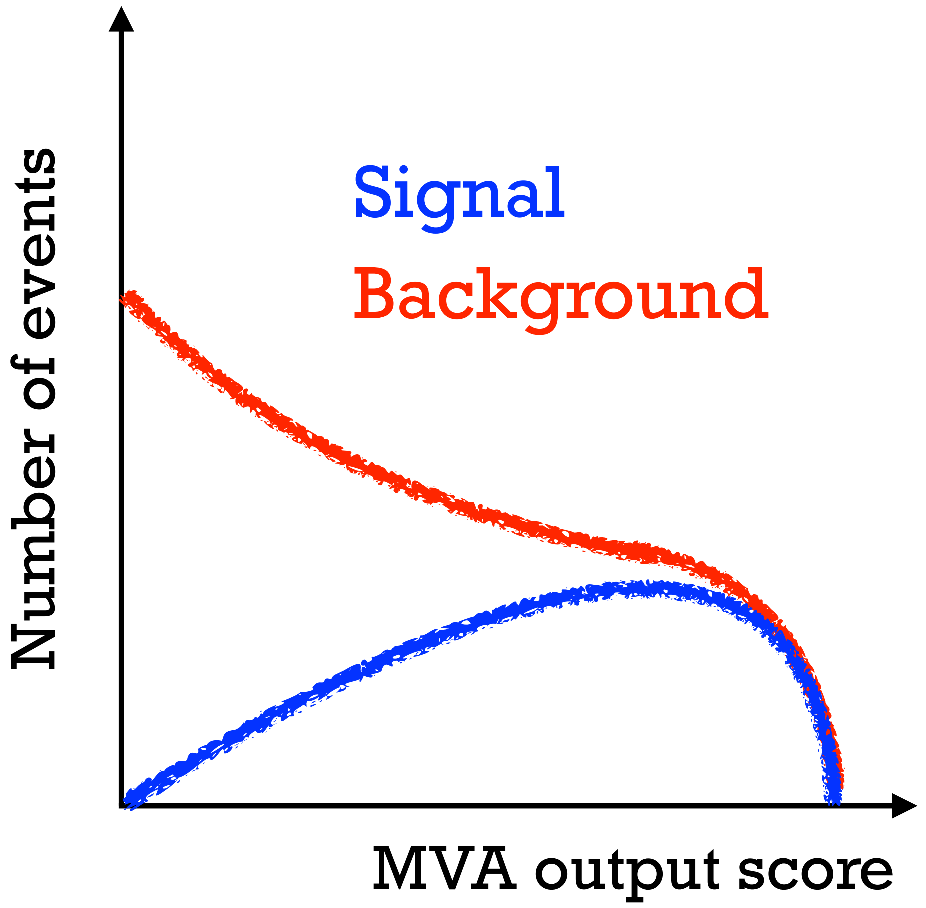
× ✓

ggH category

VBF category

The H→μμ search in a nutshell

- *Train a multivariate (MVA) classifier to separate signal from background*
  - Using kinematic input variables <u>uncorrelated with H candidate mass</u>.

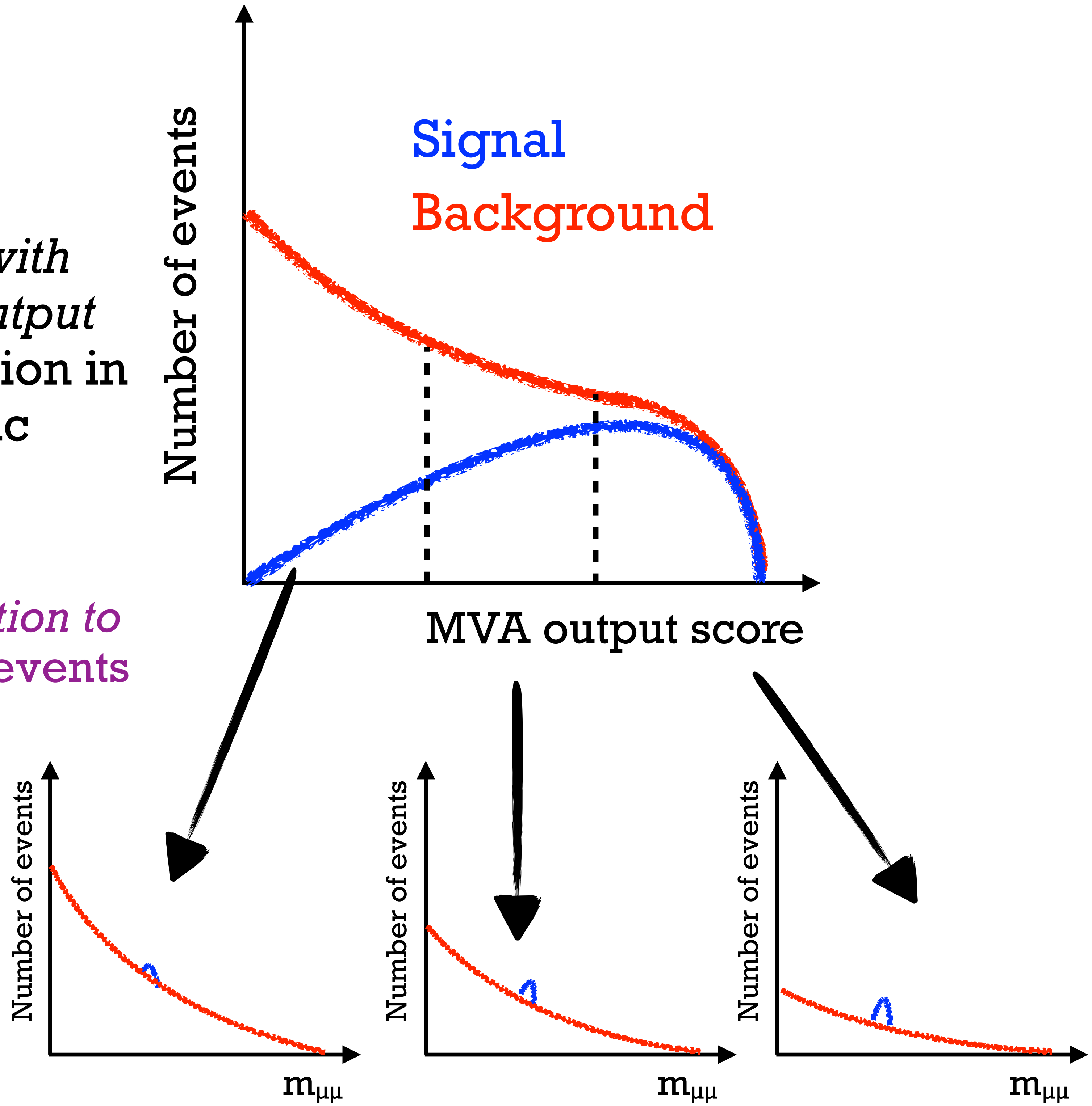


The H→μμ search in a nutshell

- *Train a multivariate (MVA) classifier to separate signal from background*
    - Using kinematic input variables uncorrelated with H candidate mass.

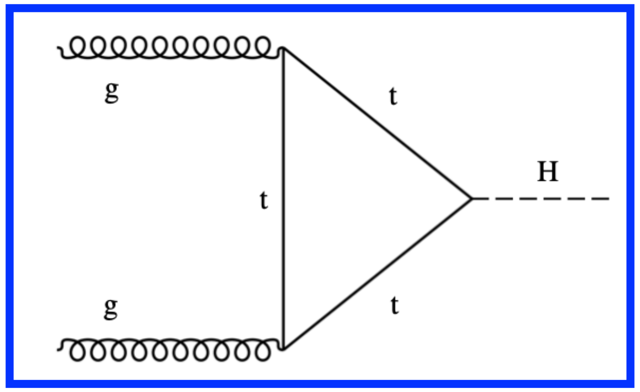
- ***Divide** events into subcategories with varying signal purity using MVA output **and fit*** the dimuon mass distribution in each subcategory with parametric functions.

- *Promote events with best mass resolution to high BDT score* by weighting signal events by $1/\sigma_m$ in BDT training.



MVA subcategories with varying signal purity

ggH candidate event

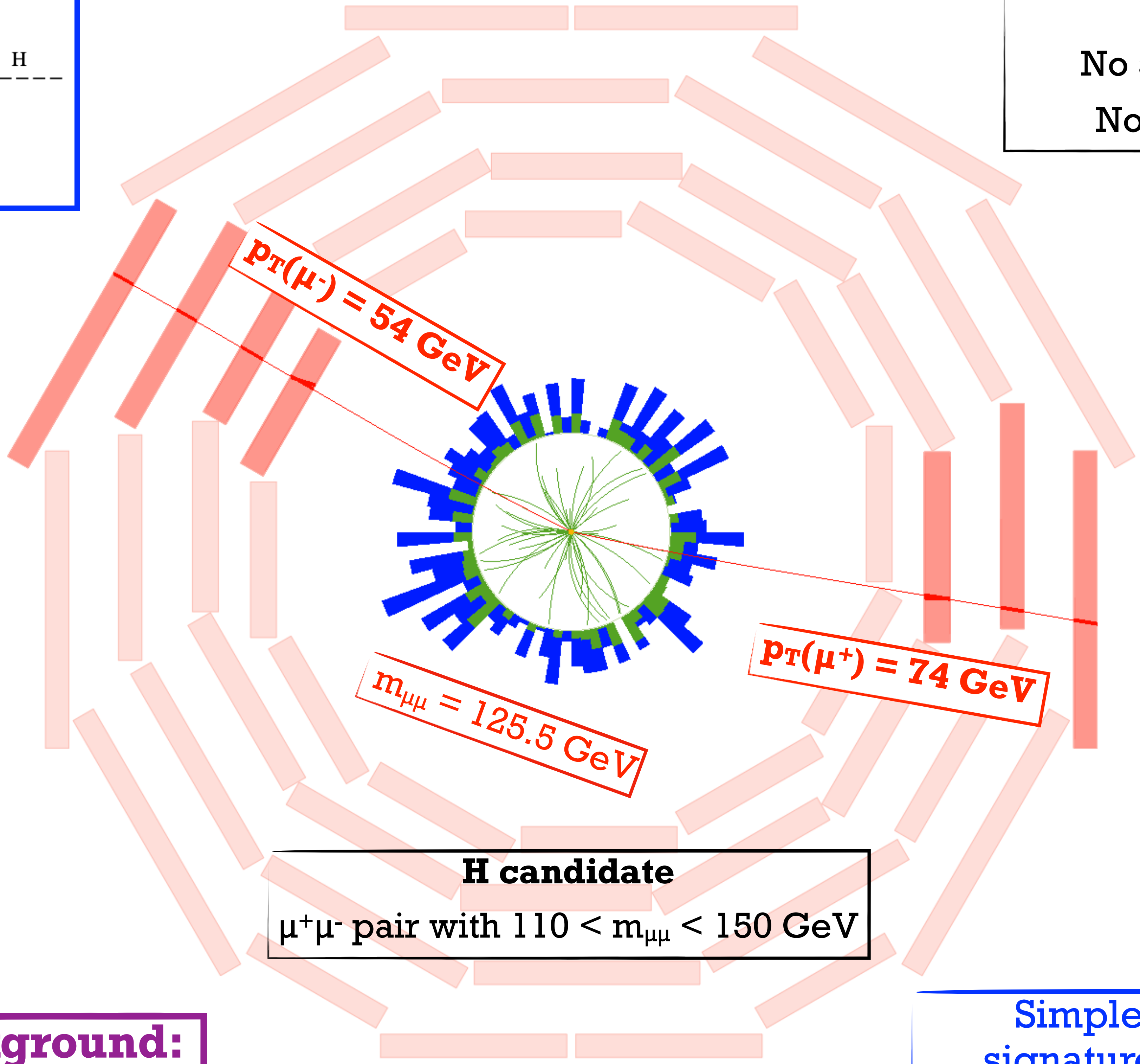


Veto exclusive categories:

No b jets

No additional leptons

No VBF-like jet pair



$p_T(\mu^-) = 54$ GeV

$m_{\mu\mu} = 125.5$ GeV

$p_T(\mu^+) = 74$ GeV

H candidate

$\mu^+\mu^-$ pair with $110 < m_{\mu\mu} < 150$ GeV

Dominant background:

Drell-Yan (DY)

Simple experimental signature, but difficult to distinguish from background.

ggH category BDT classifier

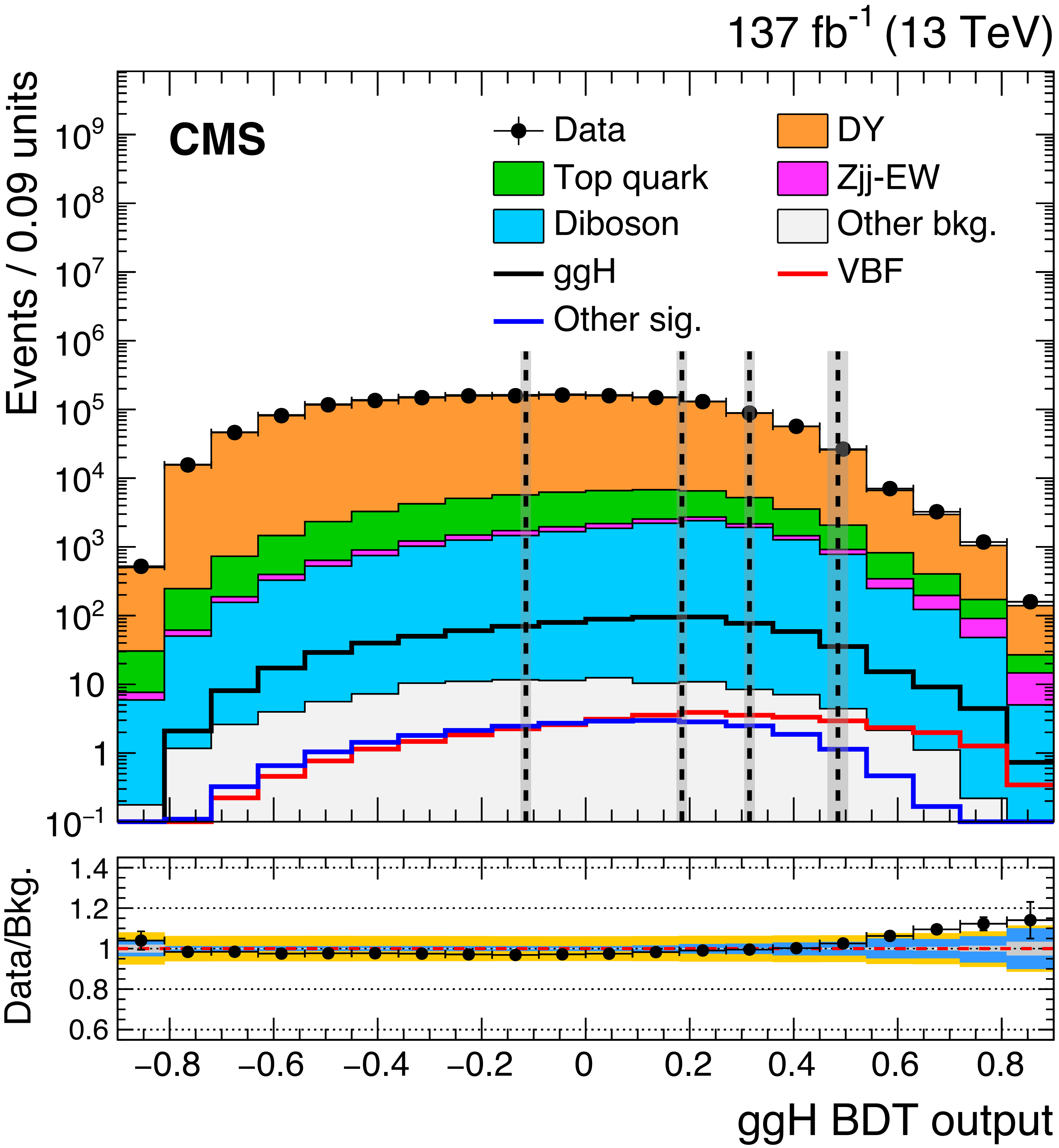
- *Train a Boosted Decision Tree (BDT) to discriminate signal from backgrounds.*
  - H candidate kinematic variables:
    - Dimuon p_T and rapidity, decay angles ϕ_{CS}, $\cos\theta_{CS}$
    - $\eta(\mu)$, $p_T(\mu)/m_{\mu\mu}$, …
  - Potential signature of initial state radiation:
    - Leading jet p_T and η.
    - Minimum angular separation between H candidate and jet.

Mass resolution per ggH subcategory

Event category	HWHM (GeV)
ggH-cat1	2.12
ggH-cat2	1.75
ggH-cat3	1.60
ggH-cat4	1.47
ggH-cat5	1.50

signal purity ↓

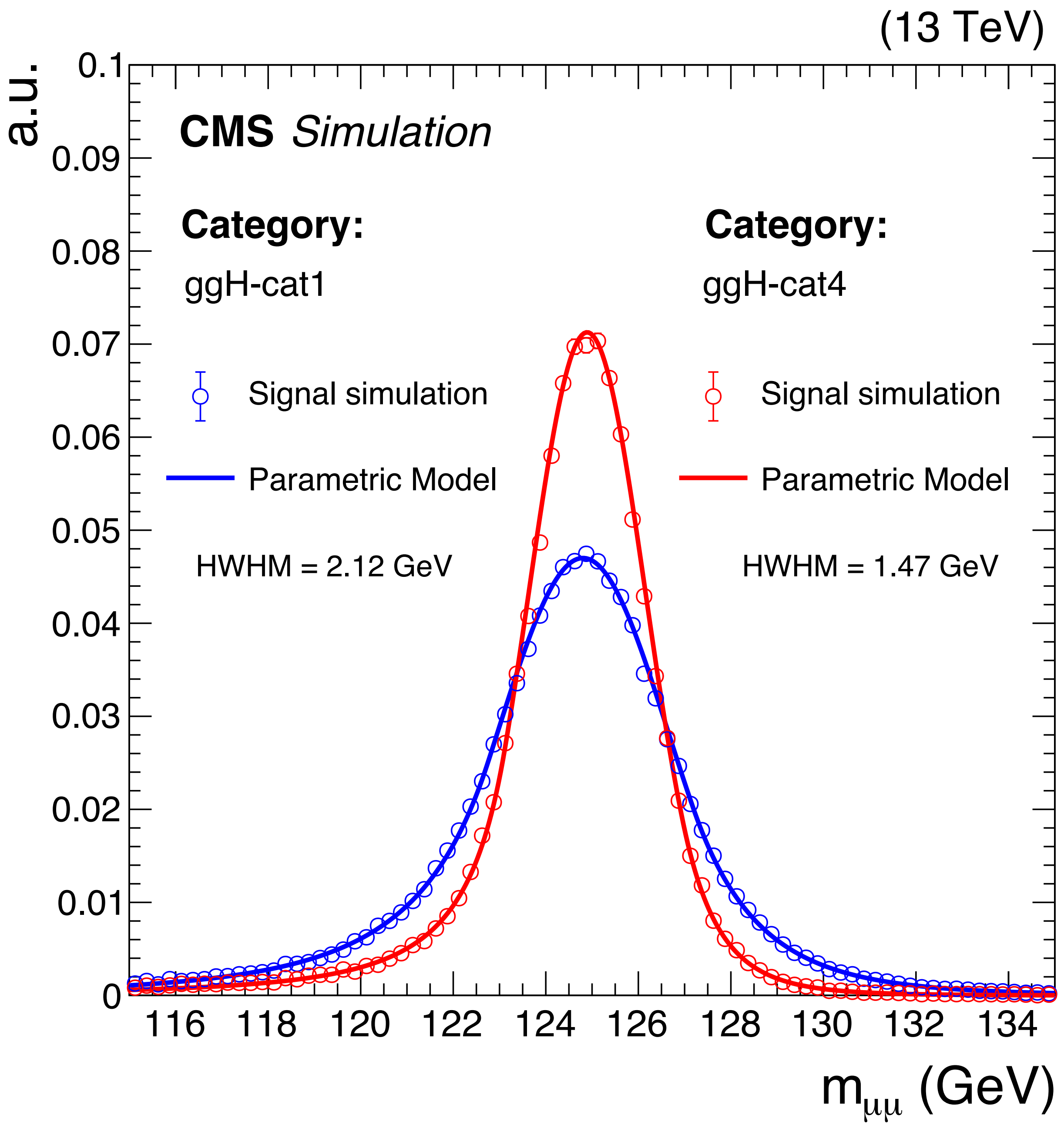
ggH category BDT score distribution



Dimuon mass resolution

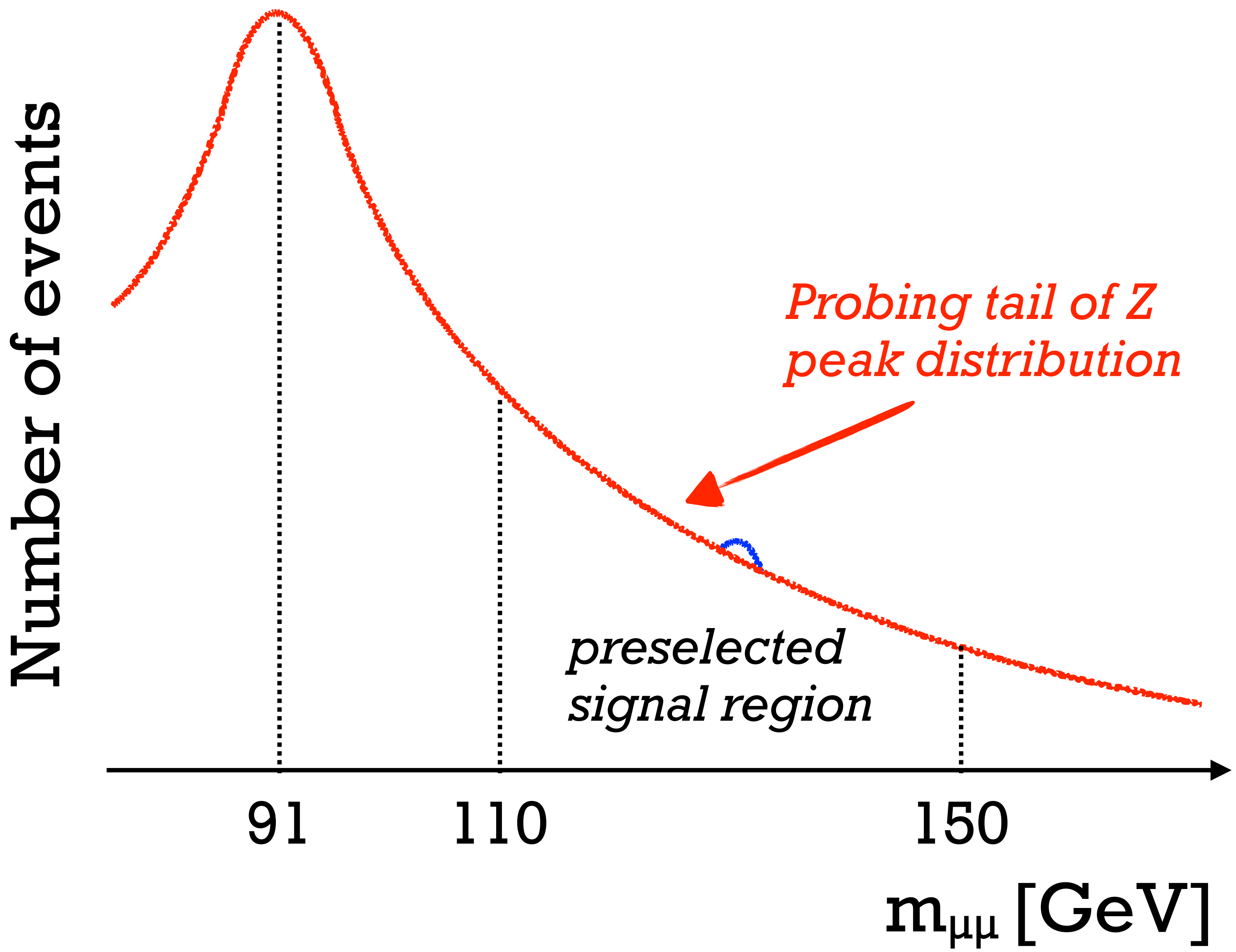
- Dimuon mass resolution ranges from 1 to 2% across subcategories, depending on muon p_T and η.

- Mass resolution further improved by:
  - recovery of photons from final state radiation (~3%)
  - constraining muon track to originate from interaction point (~4%)



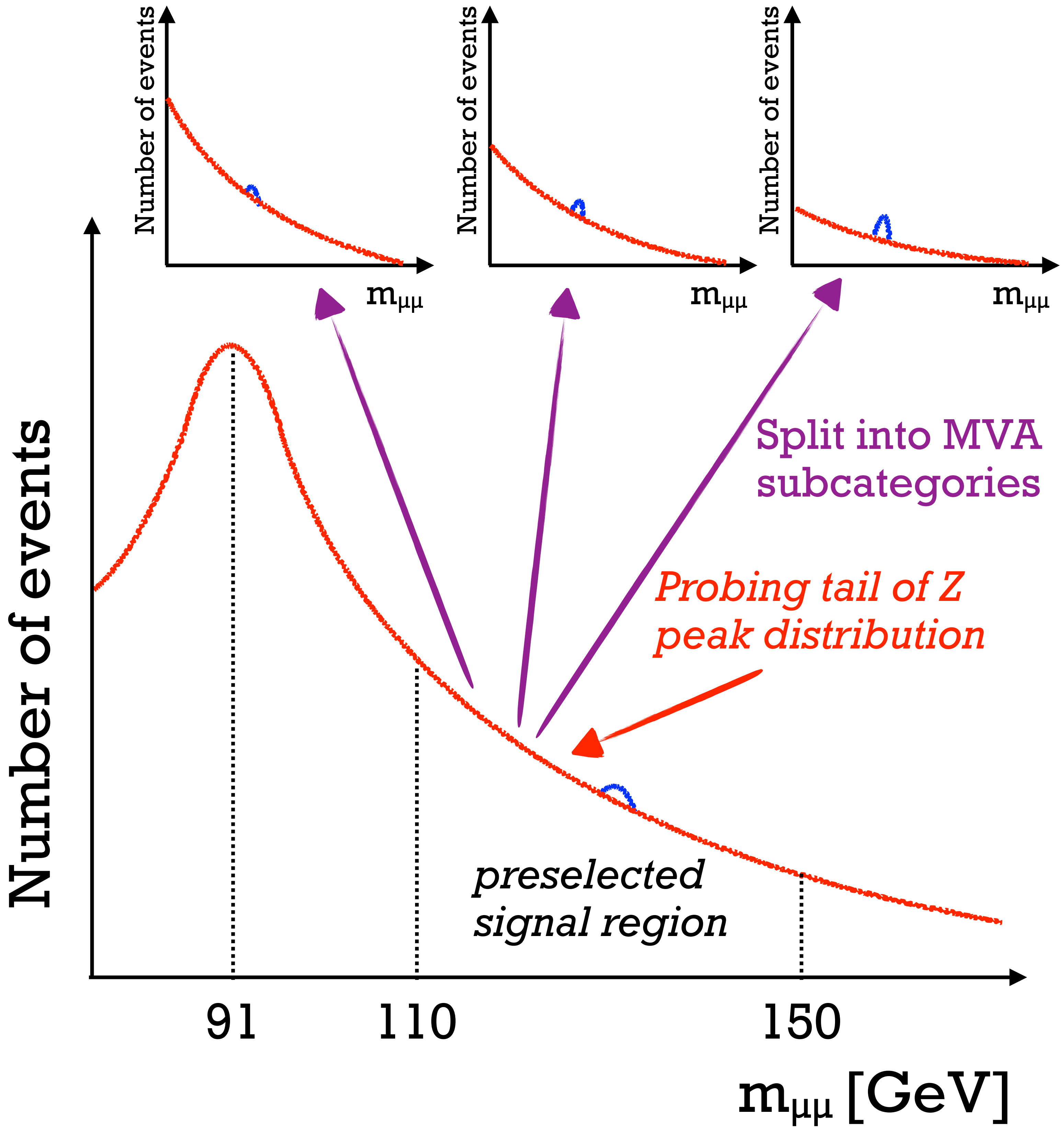
Background modeling strategy

- Looking for small resonant peak over large smoothly falling DY background.

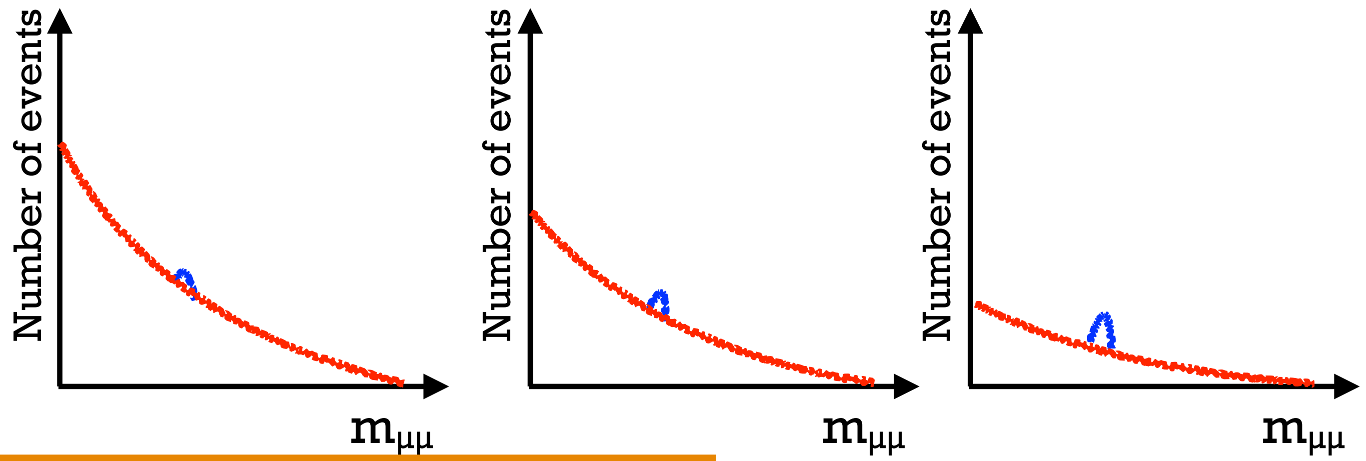


Background modeling strategy

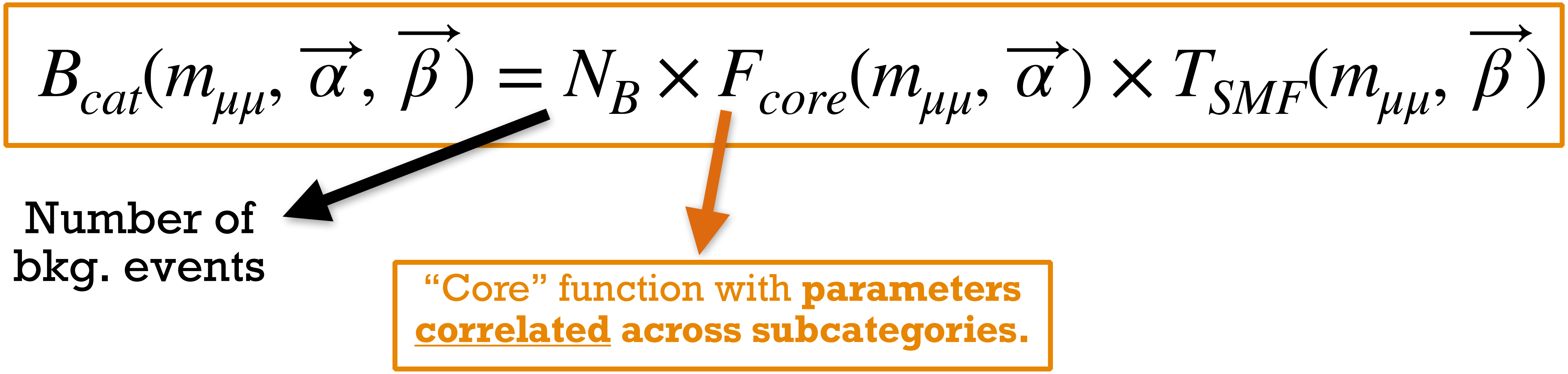
- Looking for small resonant peak over large smoothly falling DY background.
- *Background shape expected to be similar across subcategories, with minor variations from differing muon kinematics.*



Background modeling strategy



The “Core-PDF" method

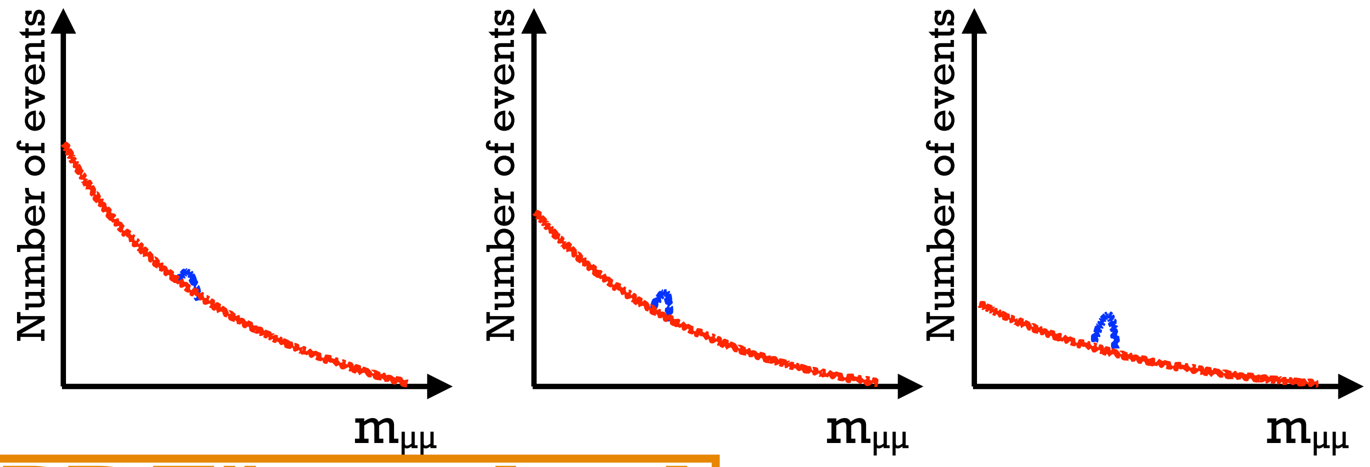


$$B_{cat}(m_{\mu\mu}, \vec{\alpha}, \vec{\beta}) = N_B \times F_{core}(m_{\mu\mu}, \vec{\alpha}) \times T_{SMF}(m_{\mu\mu}, \vec{\beta})$$

Number of bkg. events

“Core” function with **parameters** **<u>correlated</u> across subcategories.**

Background modeling strategy



The "Core-PDF" method

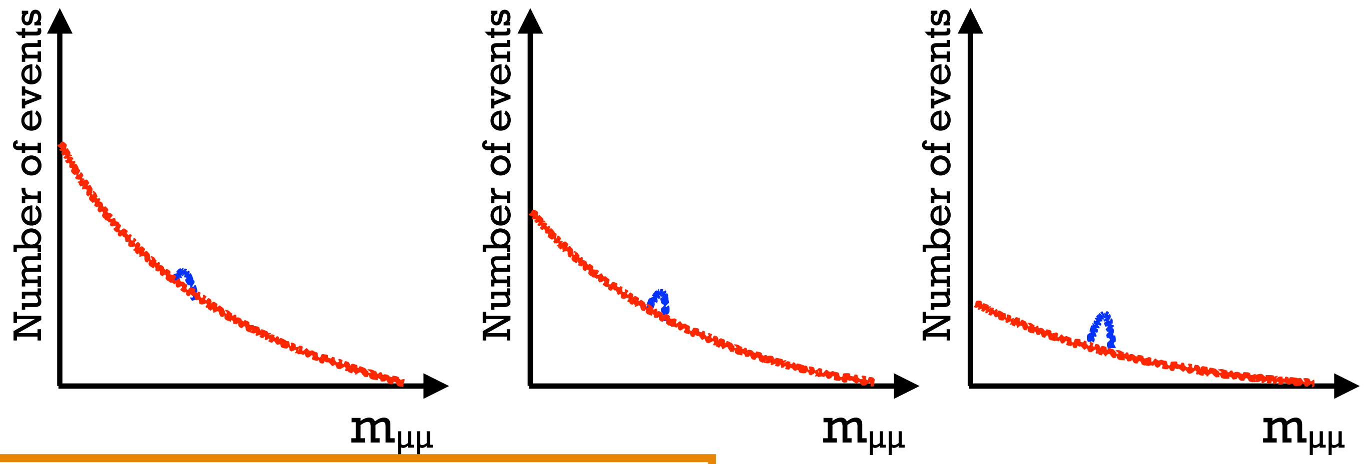
$$B_{cat}(m_{\mu\mu}, \overrightarrow{\alpha}, \overrightarrow{\beta}) = N_B \times F_{core}(m_{\mu\mu}, \overrightarrow{\alpha}) \times T_{SMF}(m_{\mu\mu}, \overrightarrow{\beta})$$

Number of bkg. events

"Core" function with **parameters <u>correlated</u> across subcategories.**

Per-category polynomial shape modifier to adjust for kinematic differences.

Background modeling strategy



The "Core-PDF" method

$$B_{cat}(m_{\mu\mu}, \vec{\alpha}, \vec{\beta}) = N_B \times F_{core}(m_{\mu\mu}, \vec{\alpha}) \times T_{SMF}(m_{\mu\mu}, \vec{\beta})$$

Number of bkg. events

"Core" function with **parameters correlated across subcategories.**

Per-category polynomial shape modifier to adjust for kinematic differences.

10% performance improvement and fewer total degrees of freedom with respect to previous strategy, while retaining negligible bias.

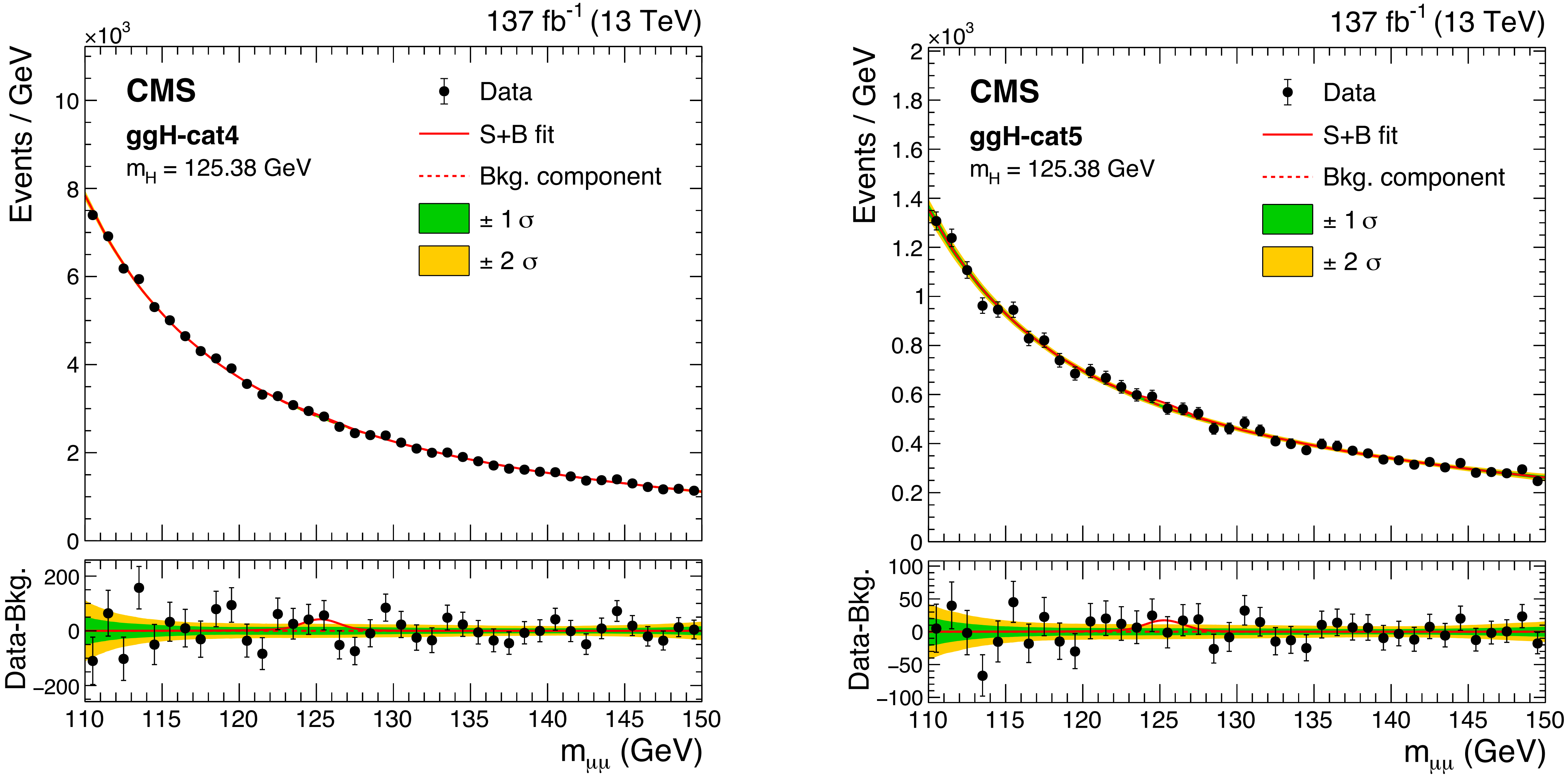
ggH category results

ggH category result

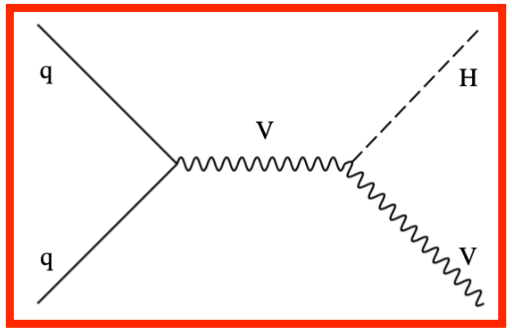
Observed (expected) significance: 1.0σ (1.6σ)

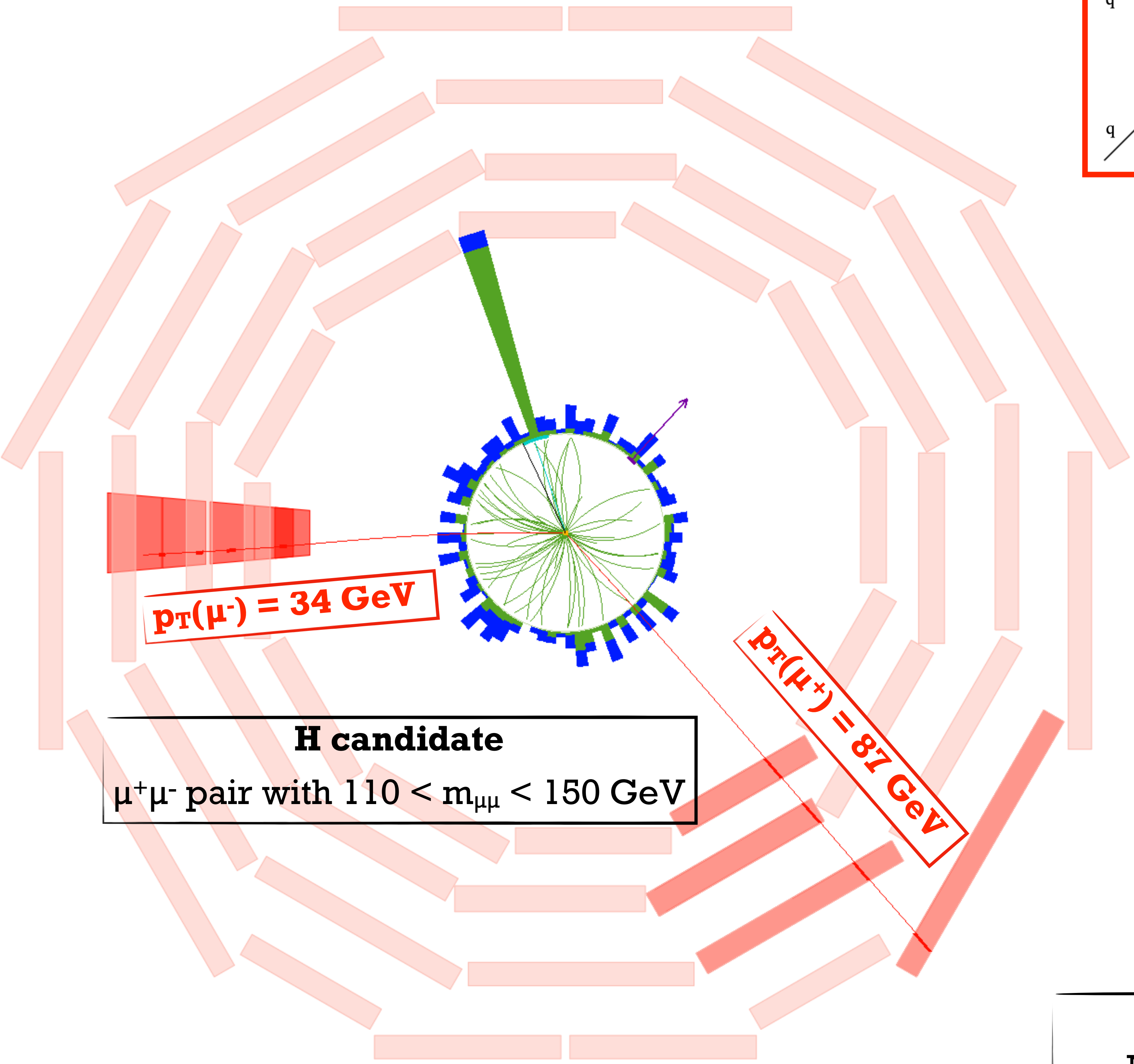
$$\mu = \frac{\sigma_H \cdot BR(H \to \mu\mu)}{\sigma_H \cdot BR(H \to \mu\mu)_{SM}} = 0.63^{+0.65}_{-0.65}$$

$m_{\mu\mu}$ distribution in the highest purity ggH subcategories



WH candidate event





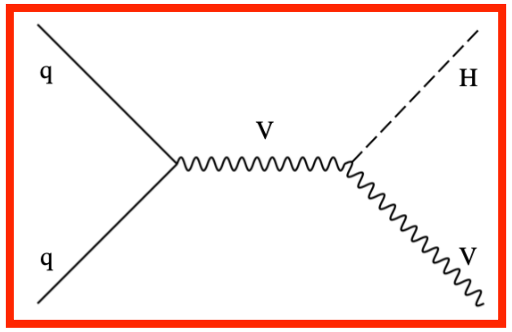
H candidate

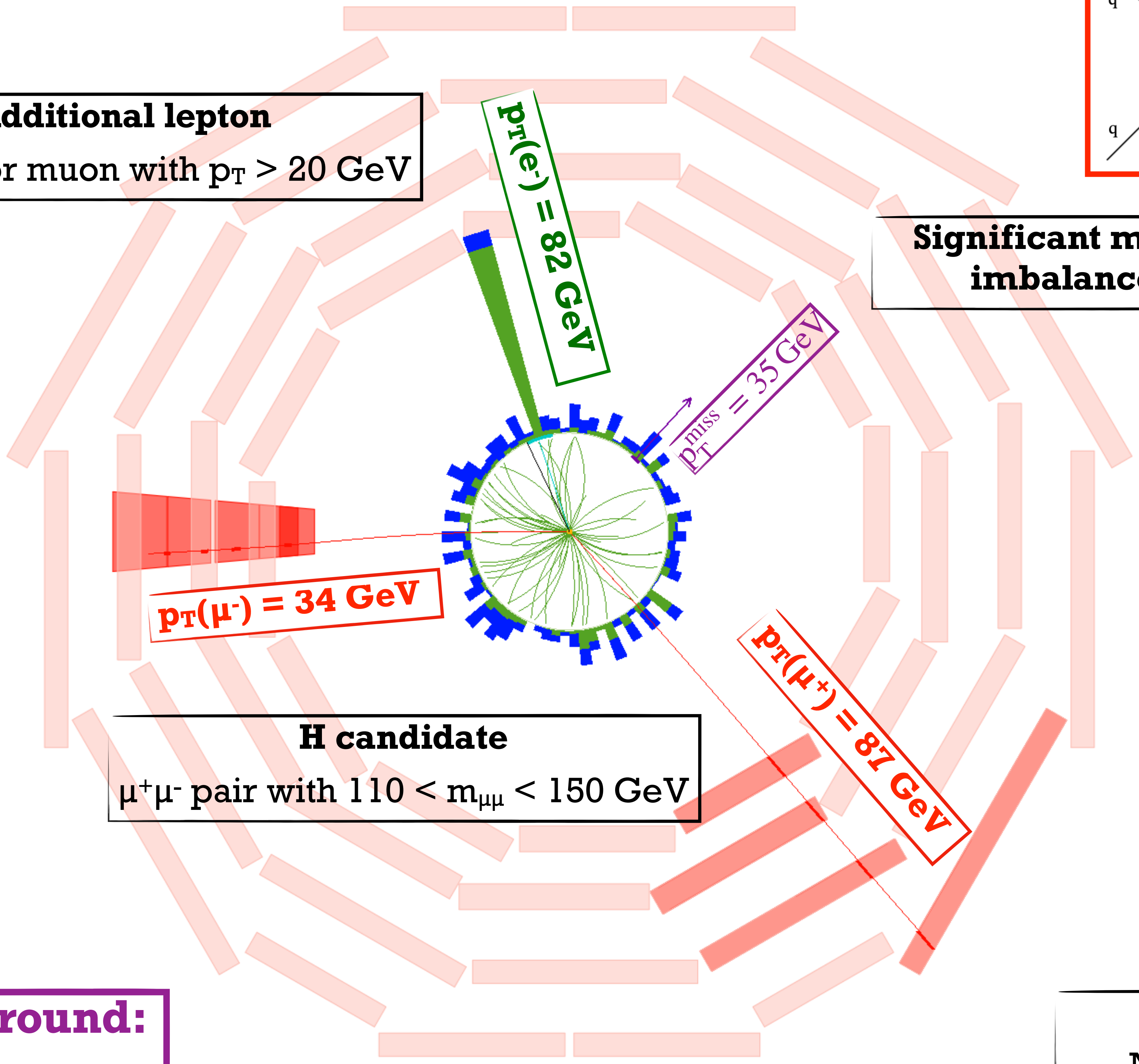
$\mu^+\mu^-$ pair with $110 < m_{\mu\mu} < 150$ GeV

b-jet veto

No jets in event passing b-tagging

WH candidate event





Additional lepton

Electron or muon with $p_T > 20$ GeV

Significant momentum imbalance p_T^{miss}

H candidate

$\mu^+\mu^-$ pair with $110 < m_{\mu\mu} < 150$ GeV

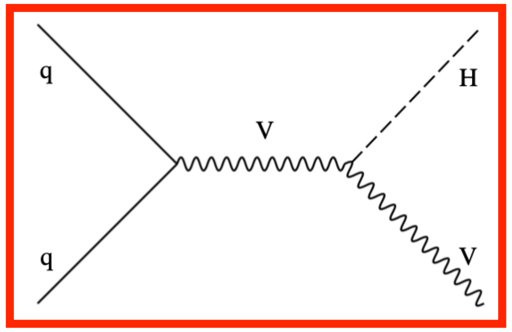
Dominant background:

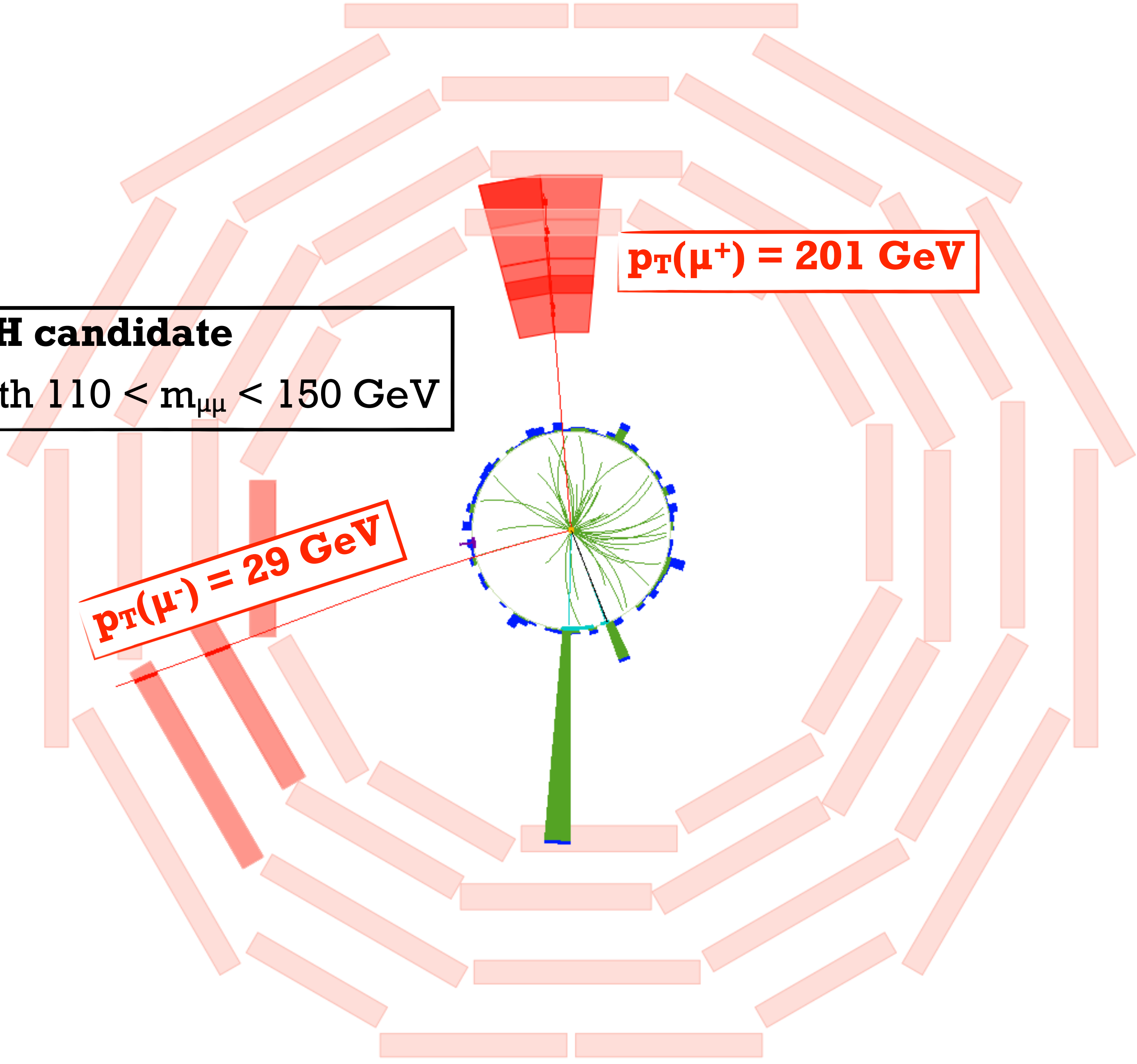
$WZ \to \ell\nu 2\mu$

b-jet veto

No jets in event passing b-tagging

ZH candidate event





$p_T(\mu^+)$ = 201 GeV

$p_T(\mu^-)$ = 29 GeV

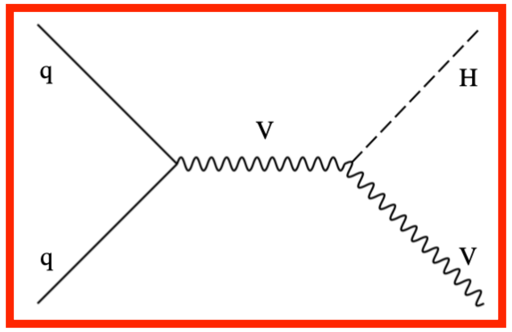
H candidate

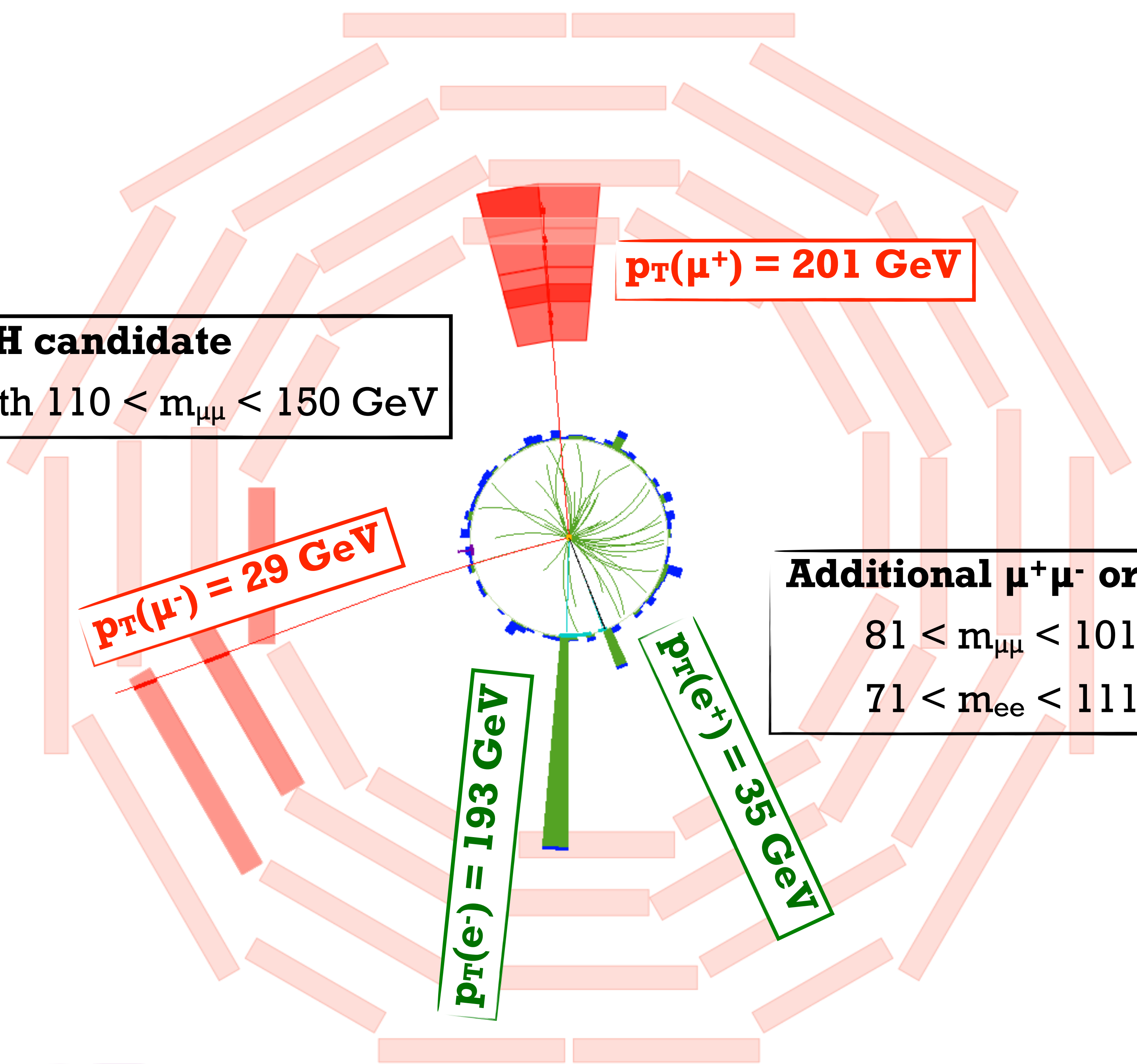
$\mu^+\mu^-$ pair with $110 < m_{\mu\mu} < 150$ GeV

b-jet veto

No jets in event passing b-tagging

ZH candidate event





H candidate

$\mu^+\mu^-$ pair with $110 < m_{\mu\mu} < 150$ GeV

Additional $\mu^+\mu^-$ or e^+e^- pair

$81 < m_{\mu\mu} < 101$ GeV

$71 < m_{ee} < 111$ GeV

Dominant background:

$ZZ \rightarrow 2\ell 2\mu$

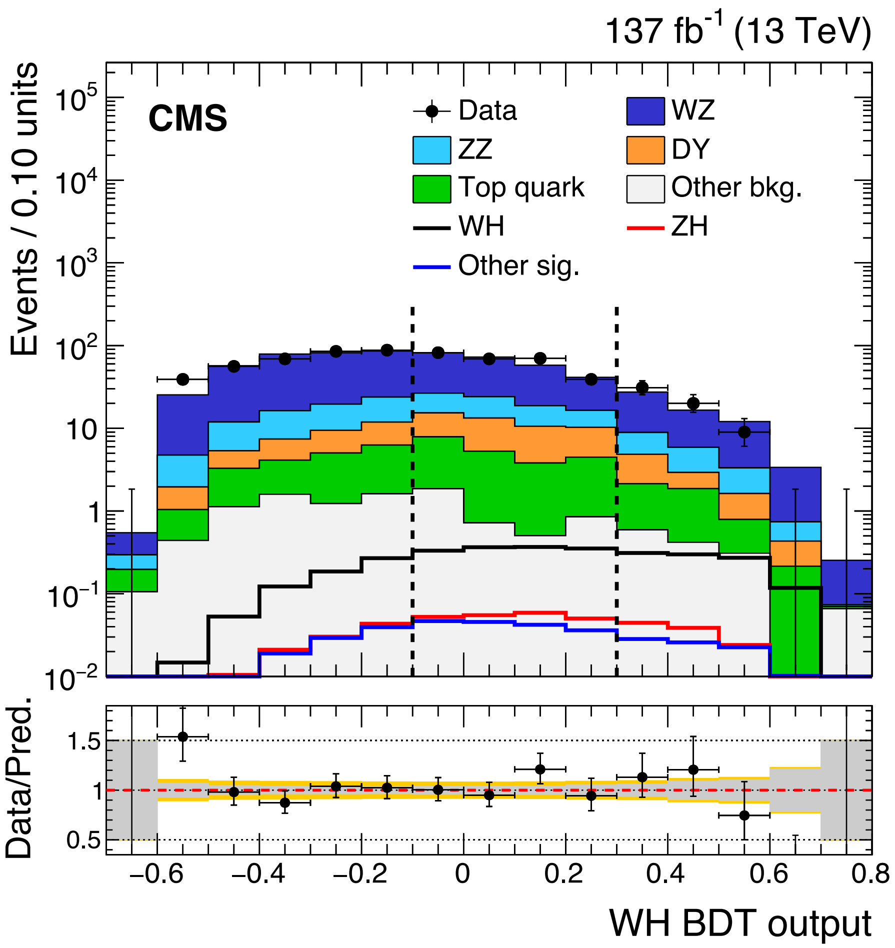
b-jet veto

No jets in event passing b-tagging

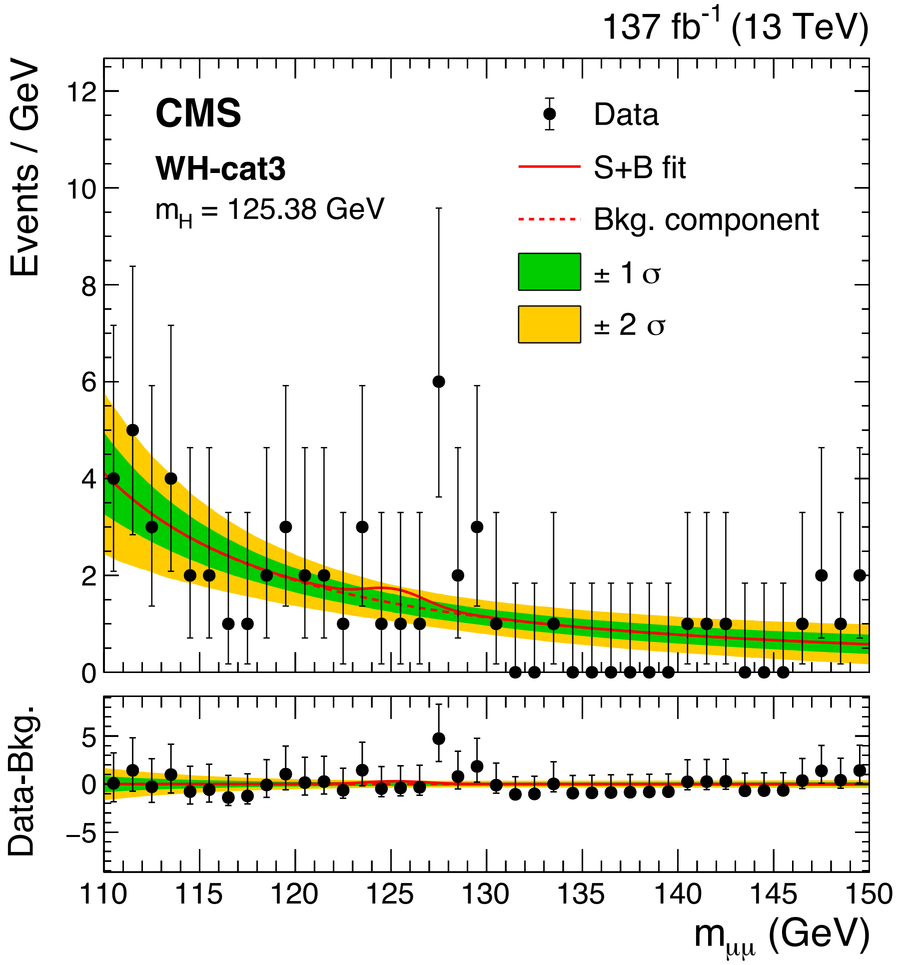
VH categories

- Inputs to WH and ZH BDT classifiers:
  - H candidate kinematic variables: $p_T(\mu\mu)$, $\Delta\phi(\mu\mu)$, …
  - WH system: $p_T(\ell_W)$, $\Delta\eta(\ell_W,H)$, $\Delta\phi(\ell_W,H)$, $m_T(\ell_W,p_T^{miss})$, …
  - ZH system: $p_T(Z)$, $\eta(Z)$, m_Z, $\Delta\eta(Z,H)$, $\Delta\phi(Z,H)$, $\cos\theta^*(Z,H)$, …

WH category BDT score



$m_{\mu\mu}$ distribution in the highest purity WH subcategory



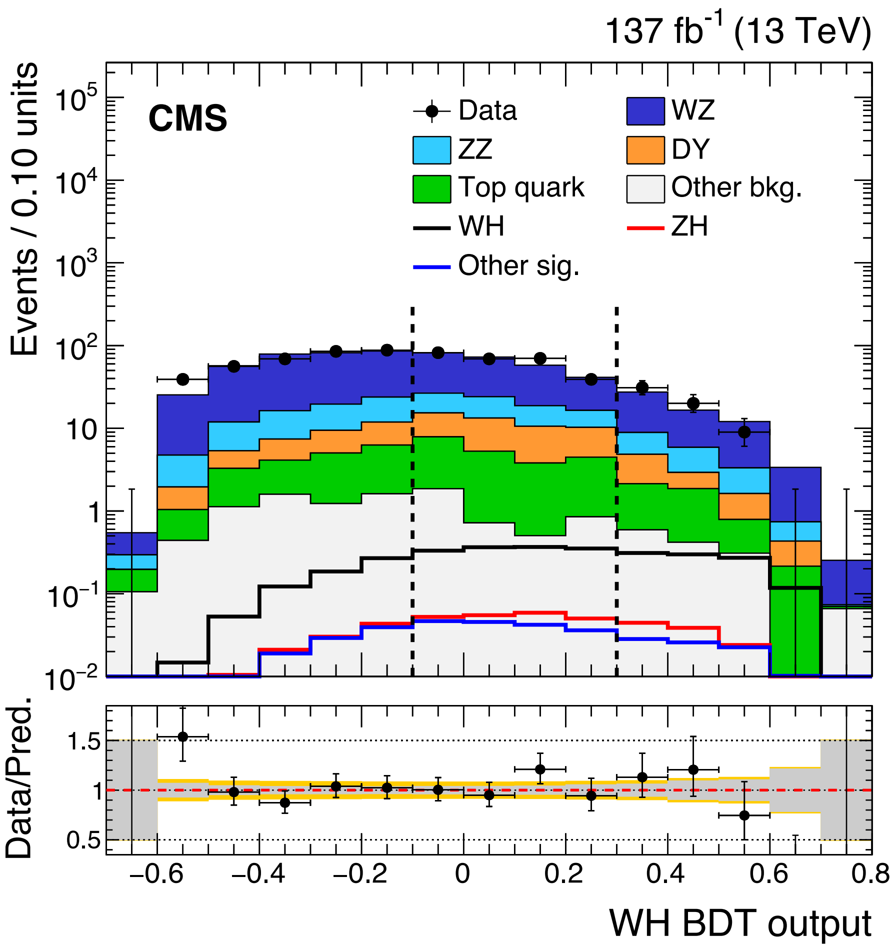
VH categories

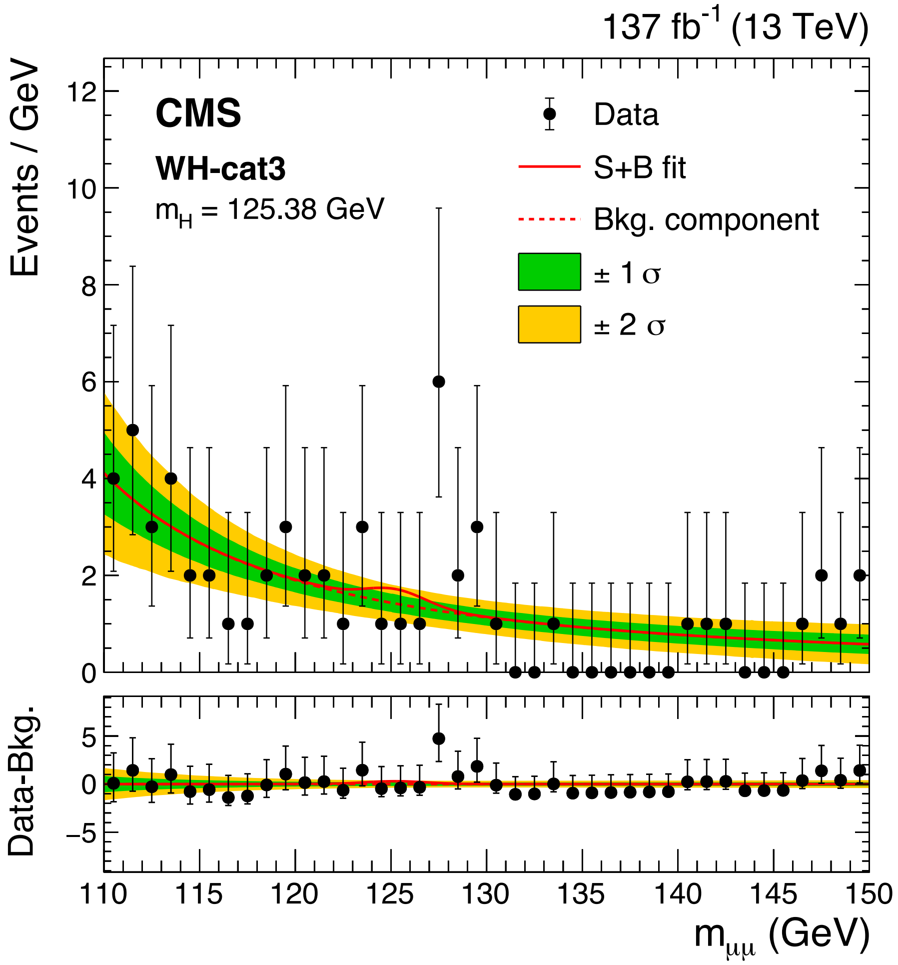
- Inputs to WH and ZH BDT classifiers:
  - H candidate kinematic variables: $p_T(\mu\mu)$, $\Delta\phi(\mu\mu)$, …
  - WH system: $p_T(\ell_W)$, $\Delta\eta(\ell_W,H)$, $\Delta\phi(\ell_W,H)$, $m_T(\ell_W,p_T^{miss})$, …
  - ZH system: $p_T(Z)$, $\eta(Z)$, m_Z, $\Delta\eta(Z,H)$, $\Delta\phi(Z,H)$, $\cos\theta^*(Z,H)$, …

Combined WH and ZH category result

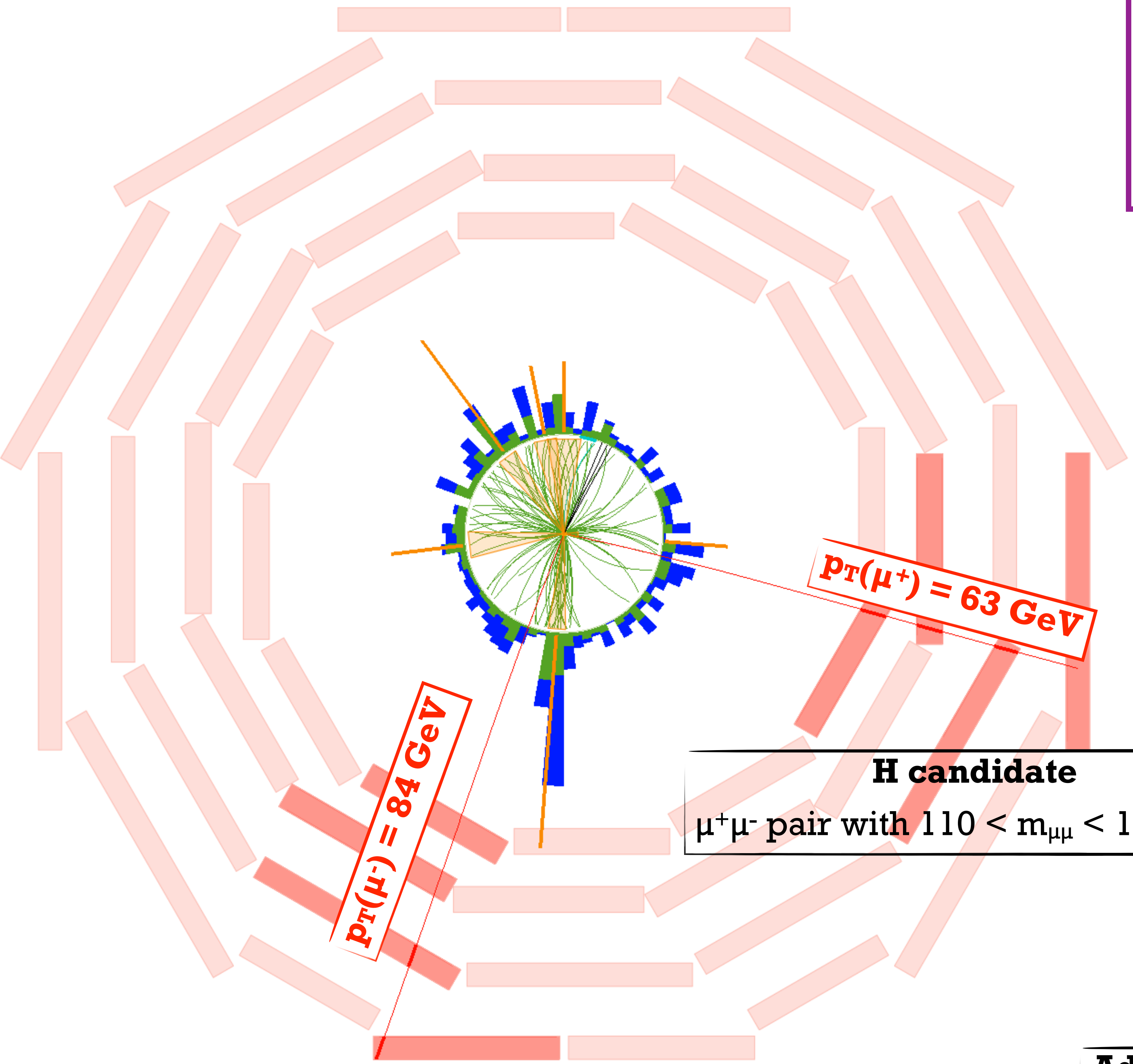
Observed (expected) significance: 2.0σ (0.4σ)

$$\mu = \frac{\sigma_H \cdot \mathrm{BR}(H \to \mu\mu)}{\sigma_H \cdot \mathrm{BR}(H \to \mu\mu)_{SM}} = 5.48^{+3.10}_{-2.83}$$



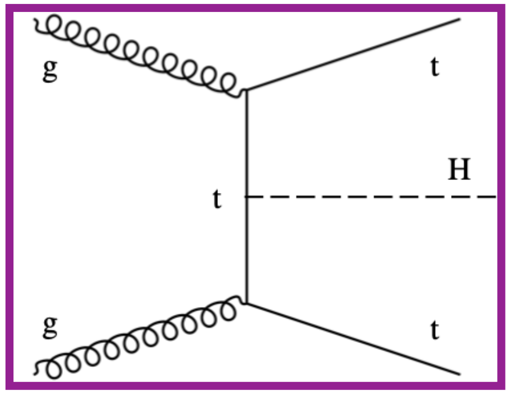


ttH (hadronic) candidate event



$p_T(\mu^+) = 63$ GeV

$p_T(\mu^-) = 84$ GeV

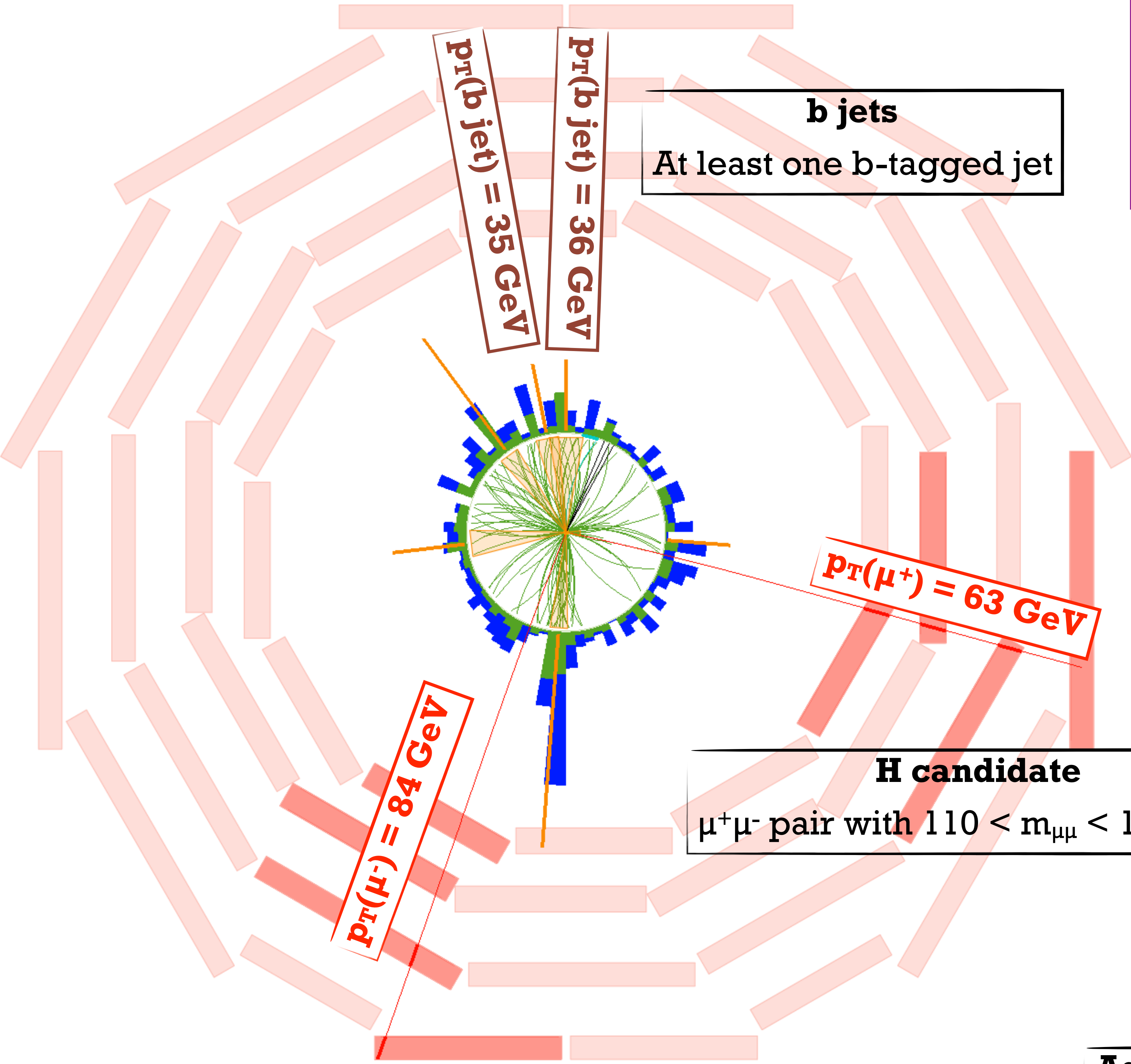


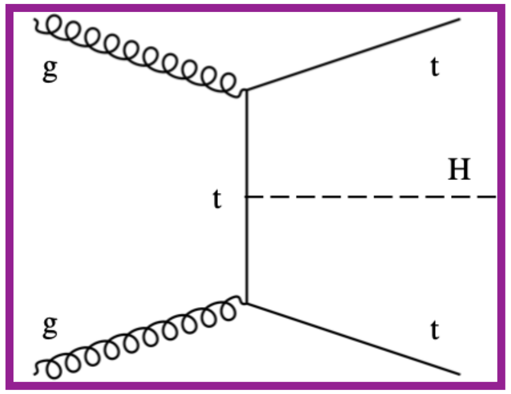
H candidate

$\mu^+\mu^-$ pair with $110 < m_{\mu\mu} < 150$ GeV

Additional lepton veto

ttH (hadronic) candidate event





b jets

At least one b-tagged jet

H candidate

$\mu^+\mu^-$ pair with $110 < m_{\mu\mu} < 150$ GeV

Additional lepton veto

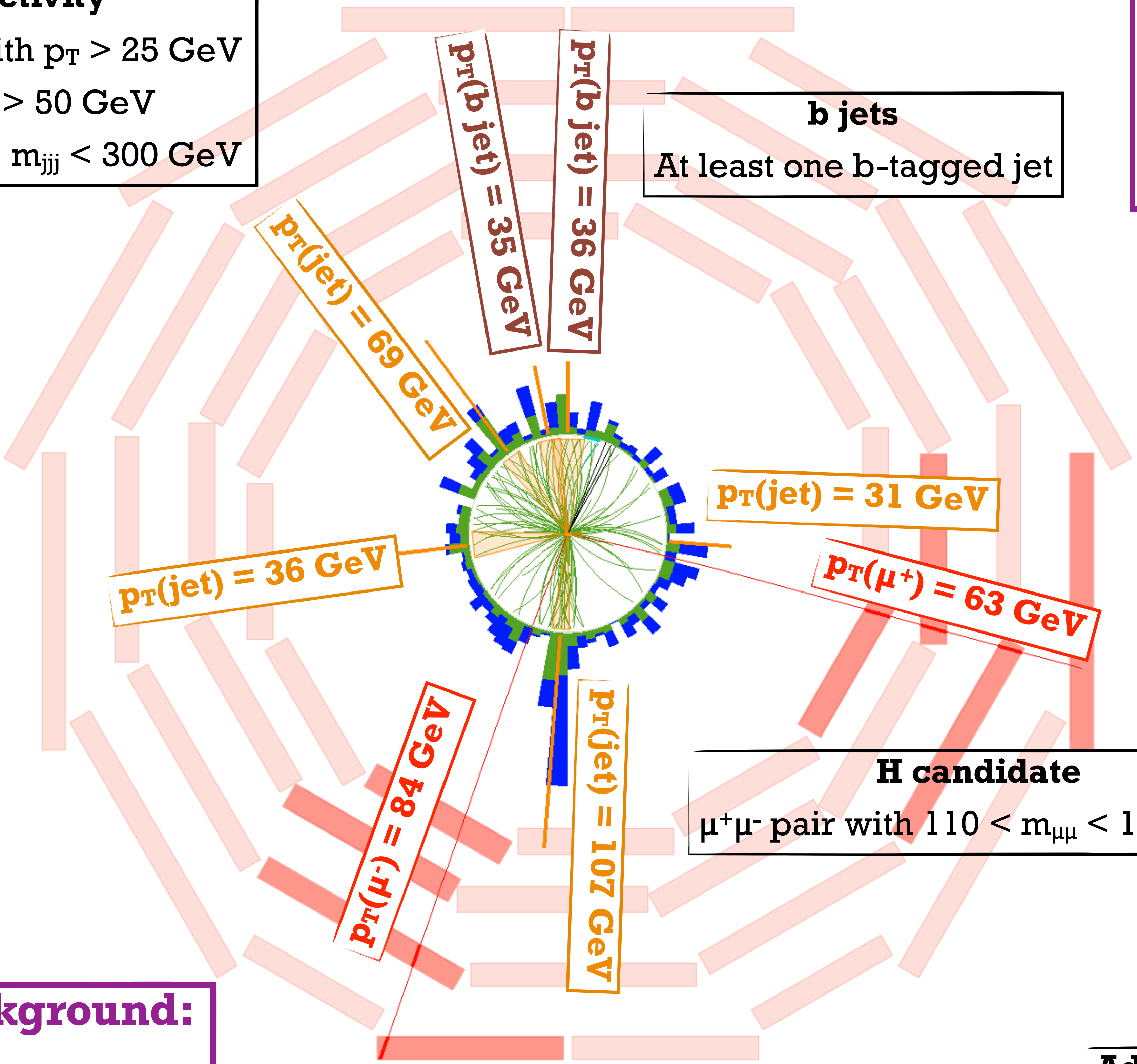
ttH (hadronic) candidate event

Hadronic activity

At least three jets with $p_T > 25$ GeV

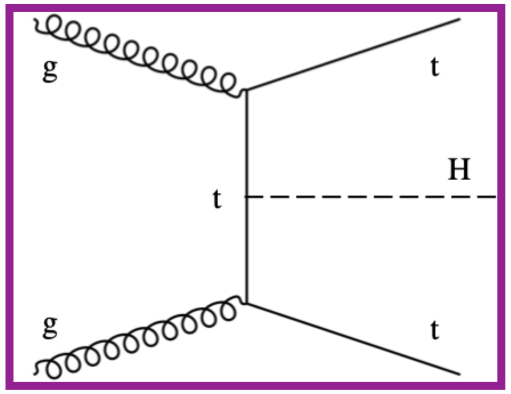
Leading jet $p_T > 50$ GeV

Jet triplet with $100 < m_{jjj} < 300$ GeV



b jets

At least one b-tagged jet



H candidate

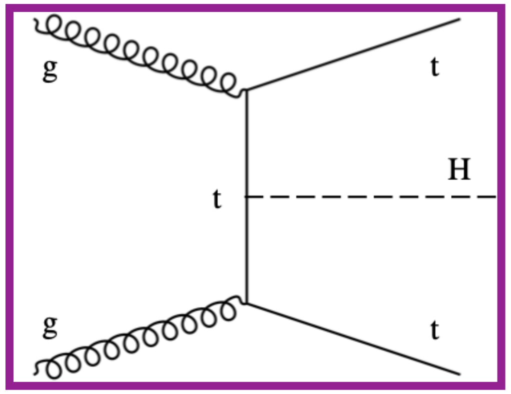
$\mu^+\mu^-$ pair with $110 < m_{\mu\mu} < 150$ GeV

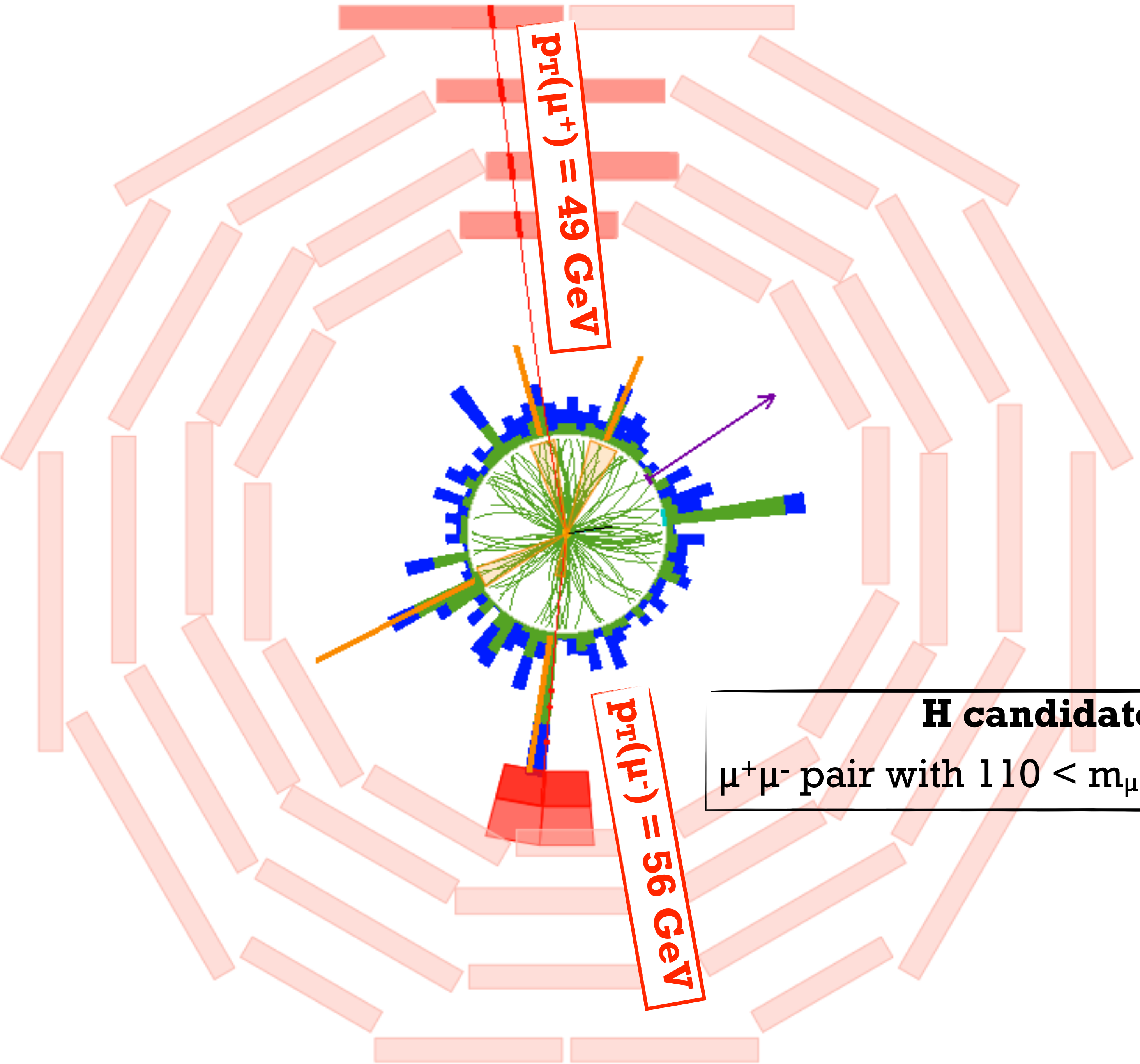
Dominant background:

tt+jets

Additional lepton veto

ttH (leptonic) candidate event



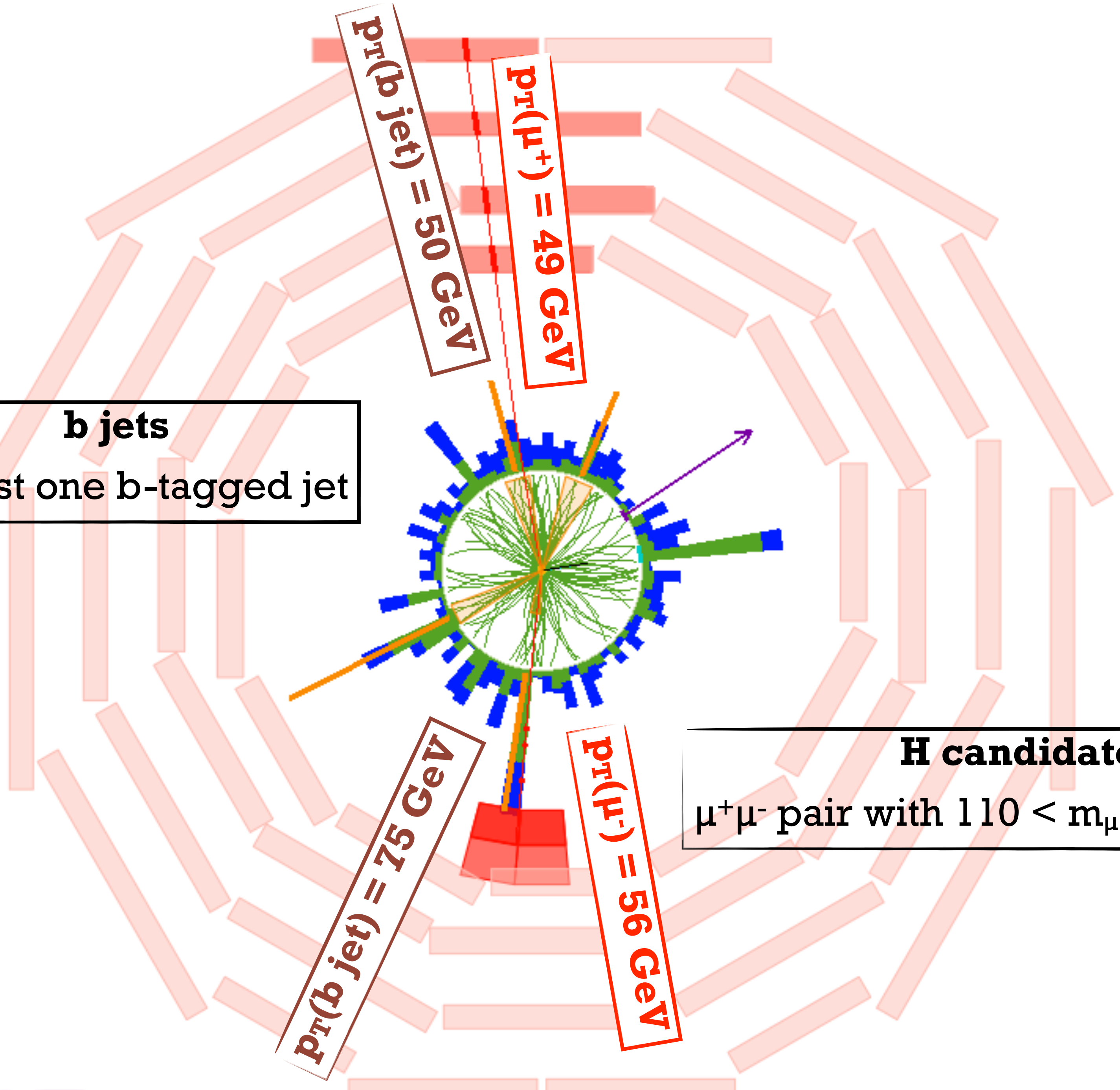


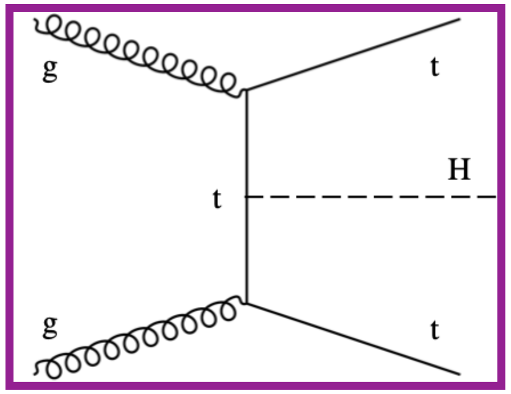
H candidate

$\mu^+\mu^-$ pair with $110 < m_{\mu\mu} < 150$ GeV

**Dileptonic ttH events also considered*

ttH (leptonic) candidate event





b jets

At least one b-tagged jet

H candidate

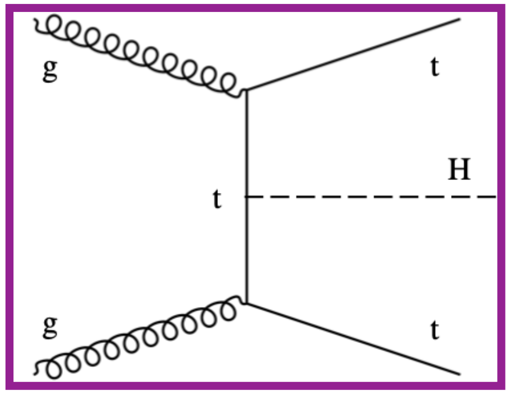
$\mu^+\mu^-$ pair with $110 < m_{\mu\mu} < 150$ GeV

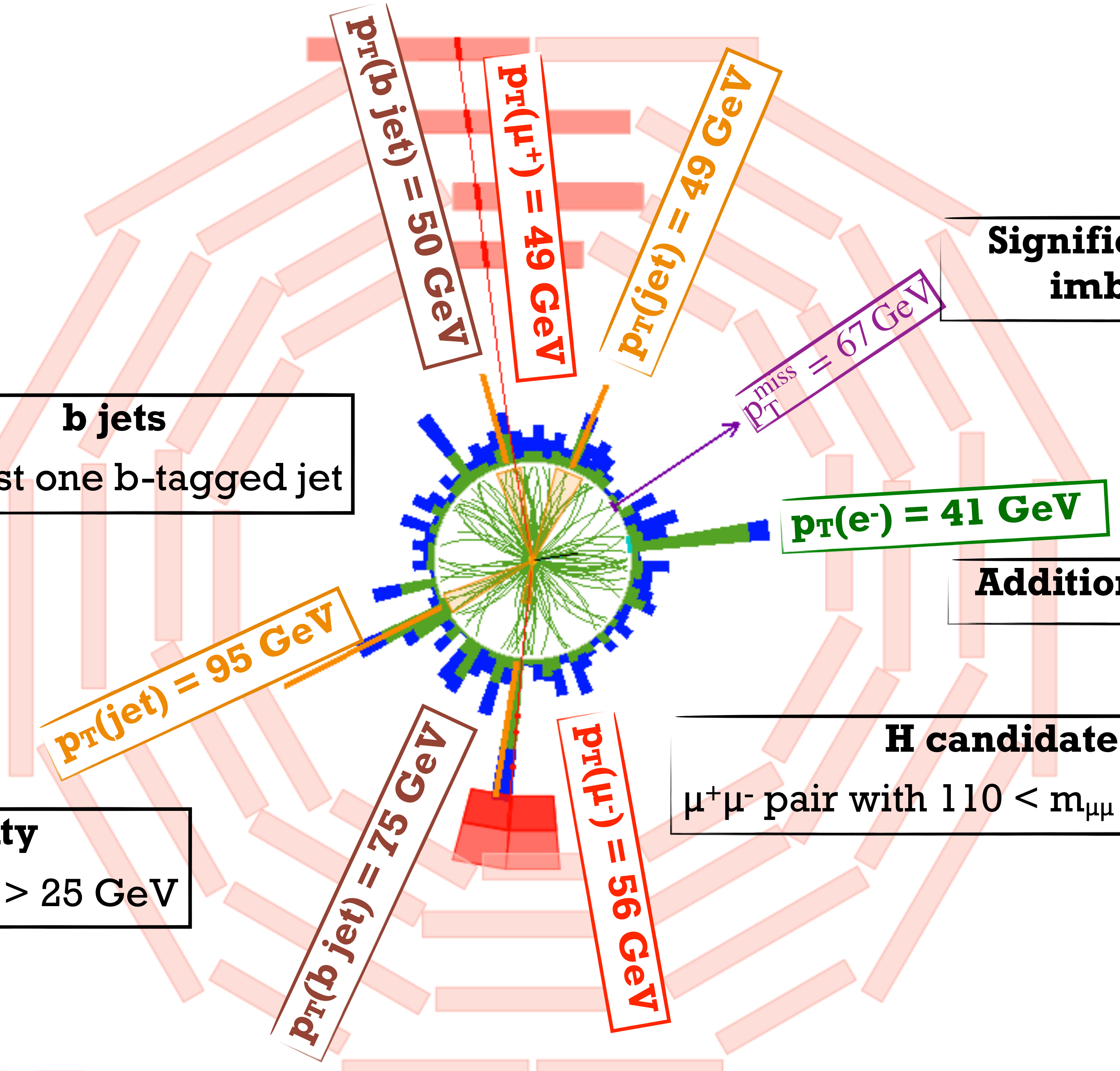
Dominant background:

tt, ttZ

**Dileptonic ttH events also considered*

ttH (leptonic) candidate event





Significant momentum imbalance p_T^{miss}

b jets

At least one b-tagged jet

Additional electron or muon

H candidate

$\mu^+\mu^-$ pair with $110 < m_{\mu\mu} < 150$ GeV

Hadronic activity

At least two jets with $p_T > 25$ GeV

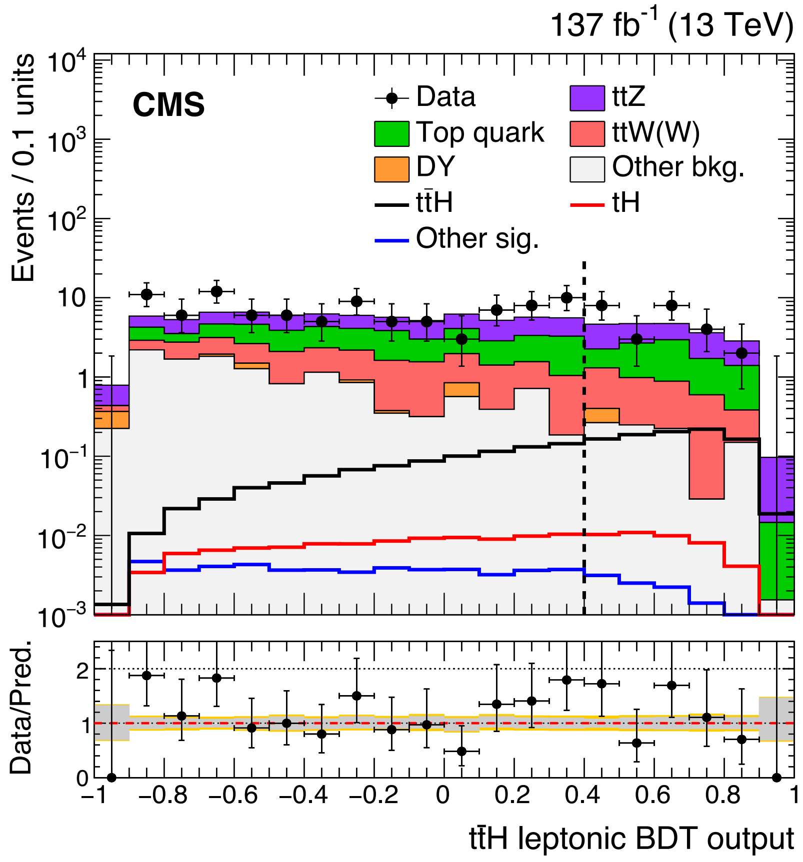
Dominant background:

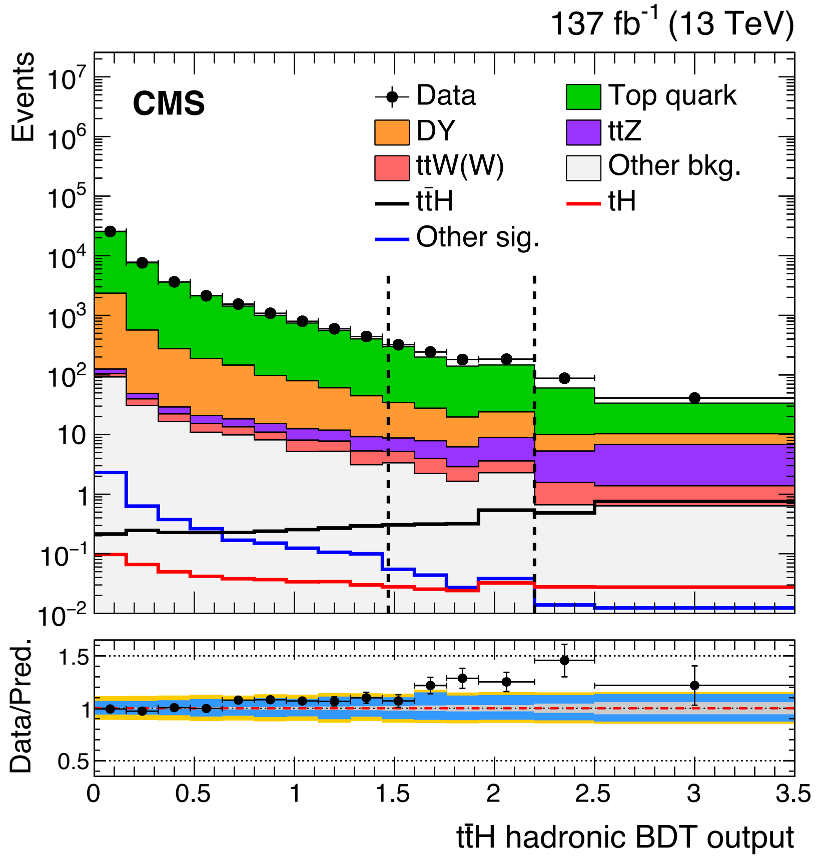
tt, ttZ

**Dileptonic ttH events also considered*

ttH category classifier

- Common ttH BDT inputs:
  - Dimuon p_T and rapidity, decay angles ϕ_{CS} , $\cos\theta_{CS}$, …
  - p_T^{miss}, H_T, jet multiplicity, …
- Specific inputs for hadronic category:
  - p_T, η of the three leading jets
  - top quark candidate: jet triplet with maximum Resolved Hadronic Top Tagger (RHTT) score
    - p_T(jjj), RHTT score, p_T-balance(H,jjj)
- Specific inputs for leptonic category:
  - ℓ^T: highest-p_T additional lepton
  - $\Delta\phi(\ell^T,H)$, mass(b jet, ℓ^T), $m_T(p_T^{miss},\ell^T)$





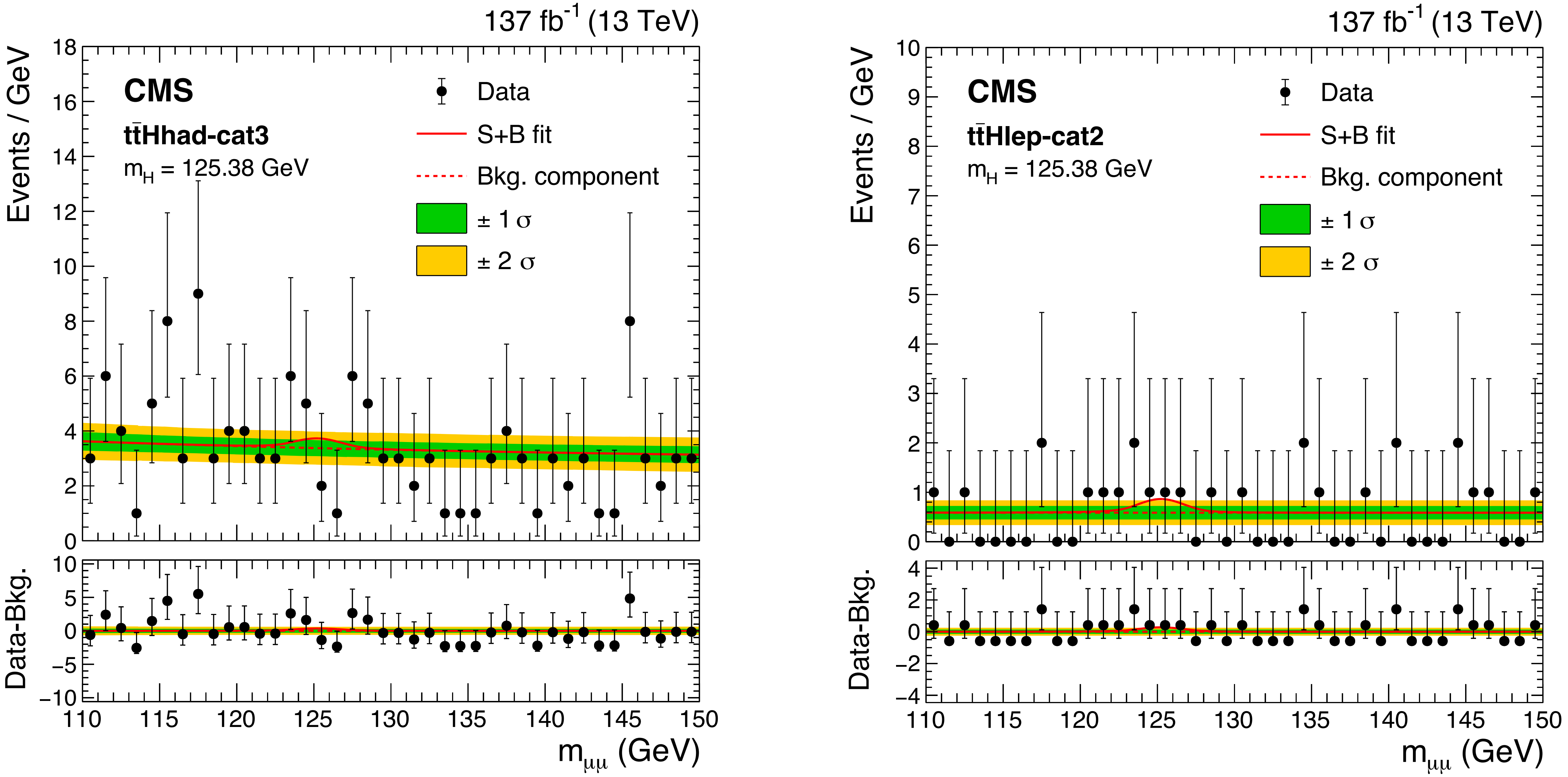
ttH category results

Combined ttH (hadronic + leptonic) category result

Observed (expected) significance: 1.2σ (0.5σ)

$$\mu = \frac{\sigma_{\rm H} \cdot {\rm BR(H \to \mu\mu)}}{\sigma_{\rm H} \cdot {\rm BR(H \to \mu\mu)}_{\rm SM}} = 2.32^{+2.27}_{-1.95}$$

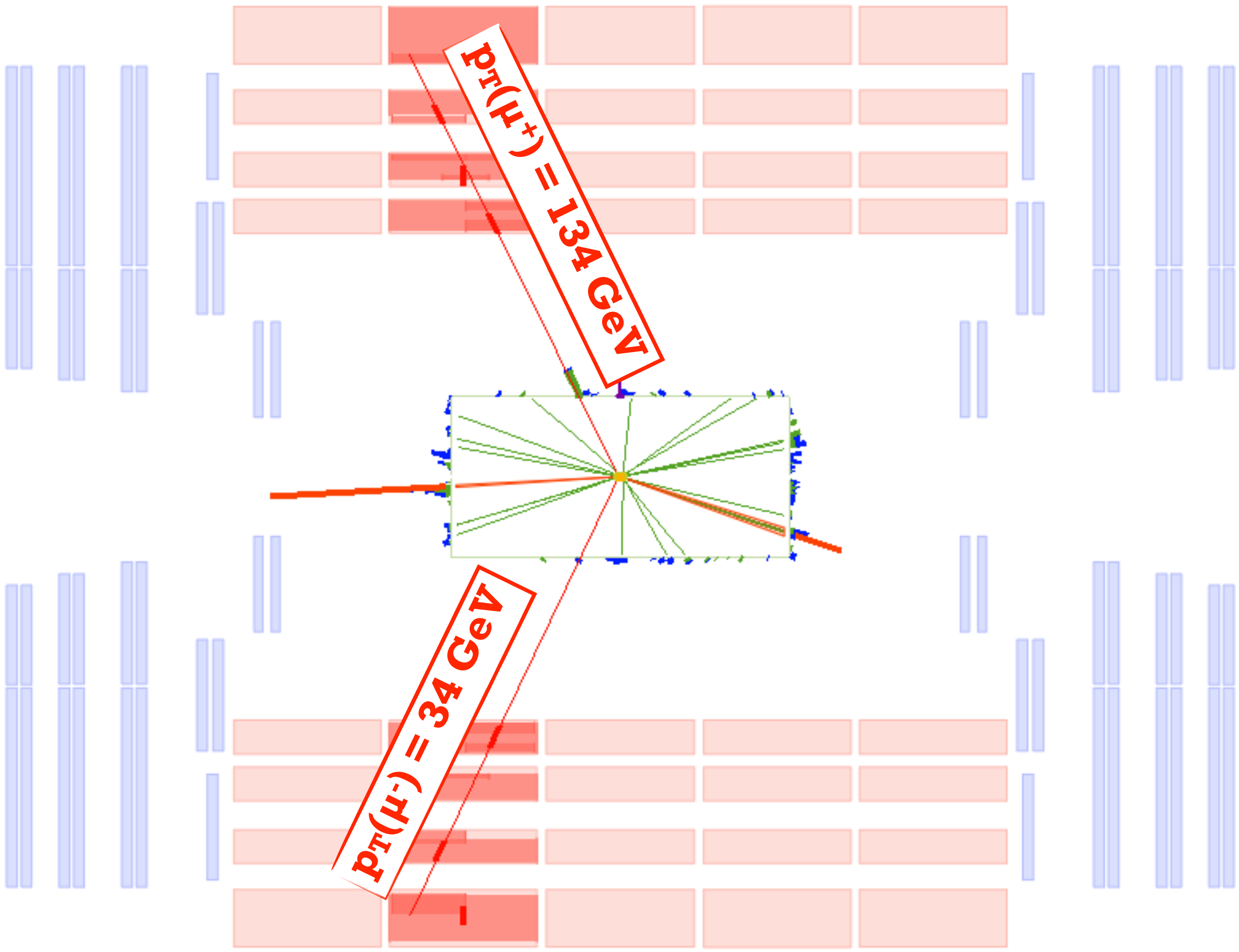
$m_{\mu\mu}$ distribution in the highest purity ttH subcategories

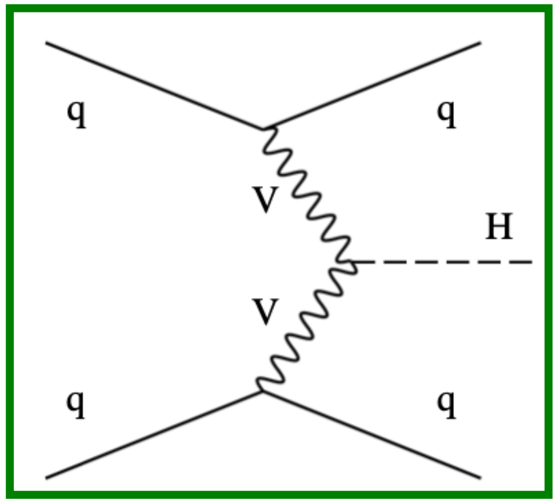


VBF candidate event

H candidate

$\mu^+\mu^-$ pair with $110 < m_{\mu\mu} < 150$ GeV



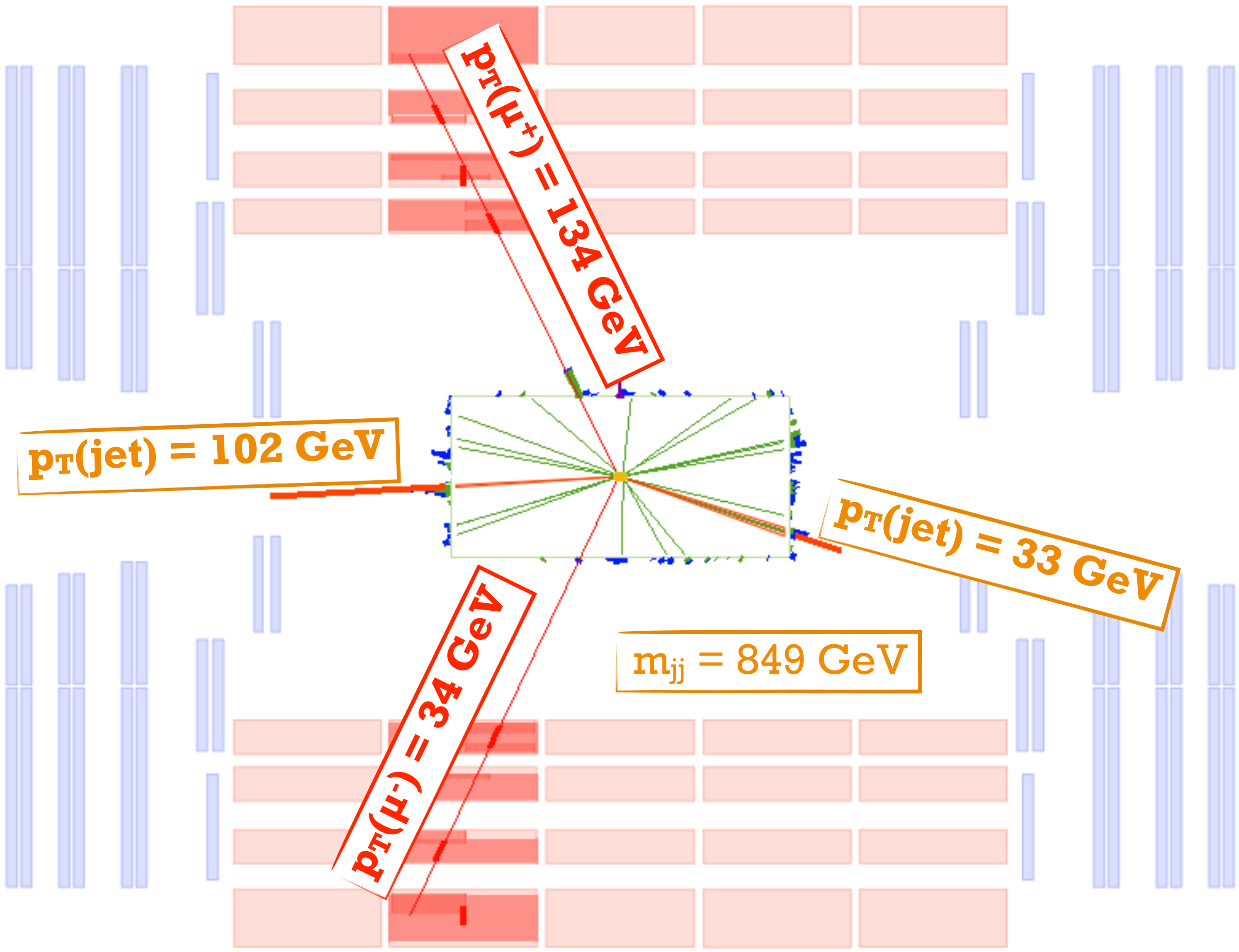


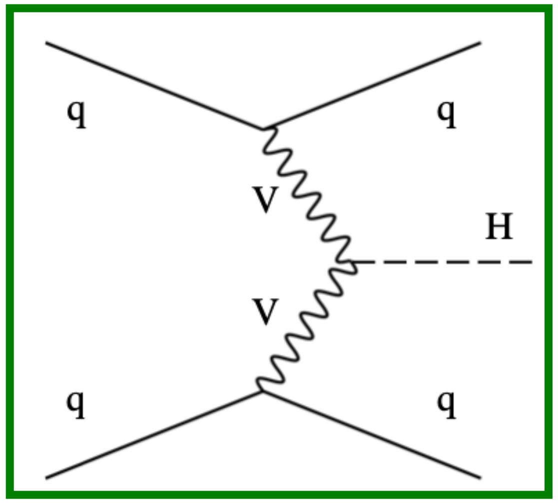
No additional leptons, no b jets.

VBF candidate event

H candidate

$\mu^+\mu^-$ pair with $110 < m_{\mu\mu} < 150$ GeV





VBF selections:

At least two jets with $p_T > 25$ GeV and $|\eta| < 4.7$,

leading jet $p_T > 35$ GeV,

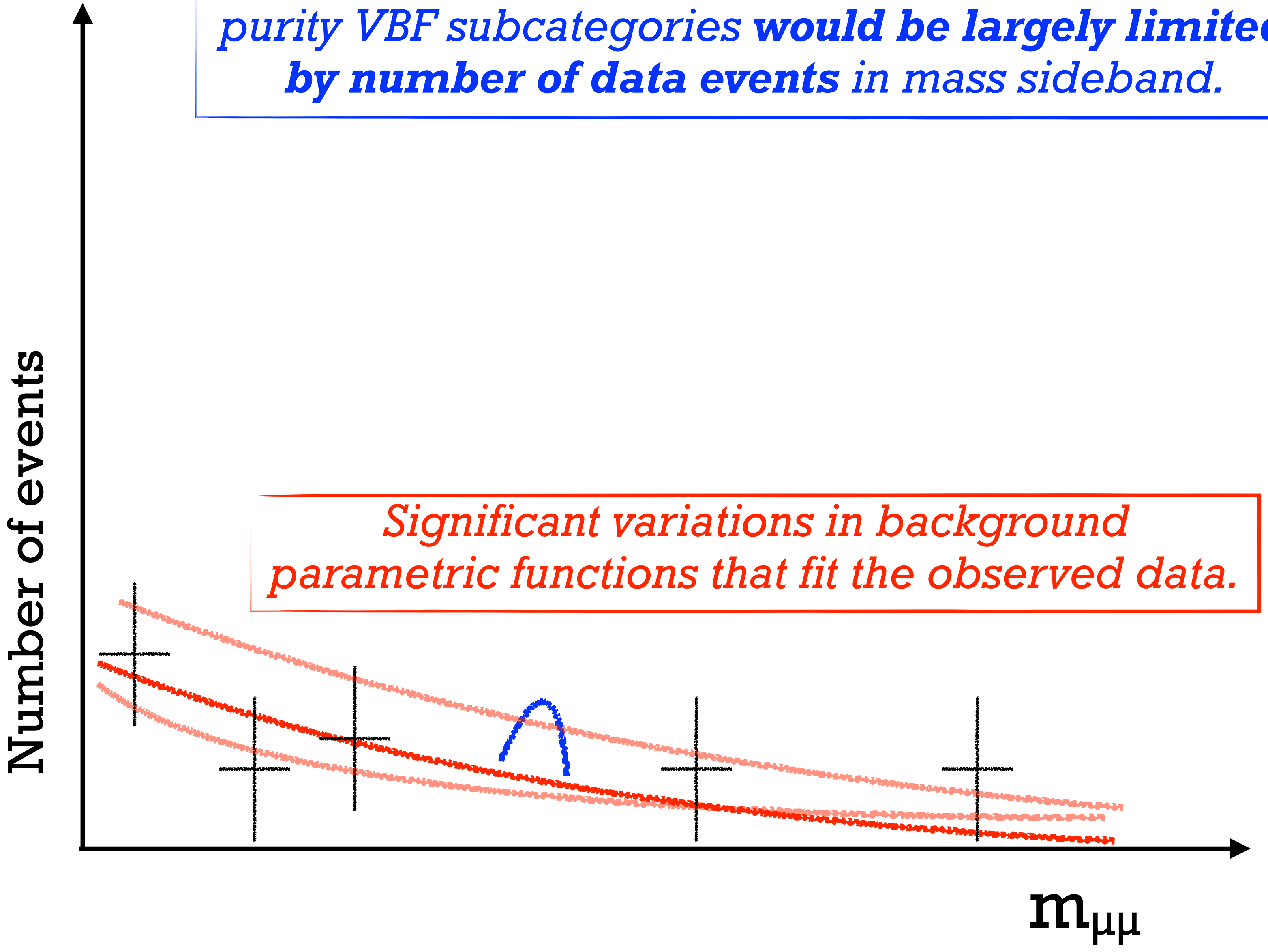
$m_{jj} > 400$ GeV and $|\Delta\eta(jj)| > 2.5$

Dominant backgrounds:

Drell-Yan (DY), EW Zjj

No additional leptons, no b jets.

VBF category: background modeling

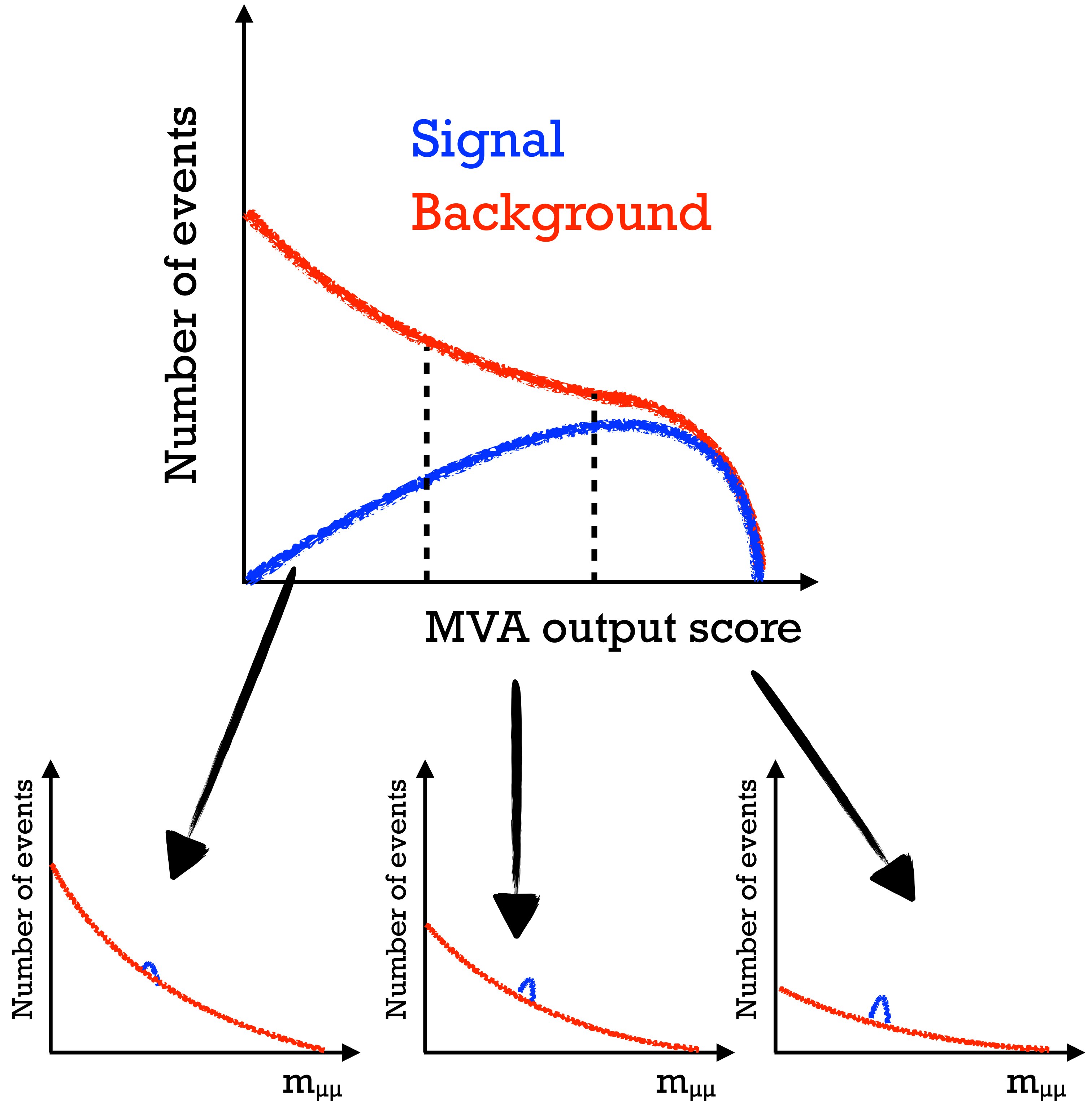


Data-driven background estimation *in high signal purity VBF subcategories* ***would be largely limited by number of data events*** *in mass sideband.*

Significant variations in background parametric functions that fit the observed data.

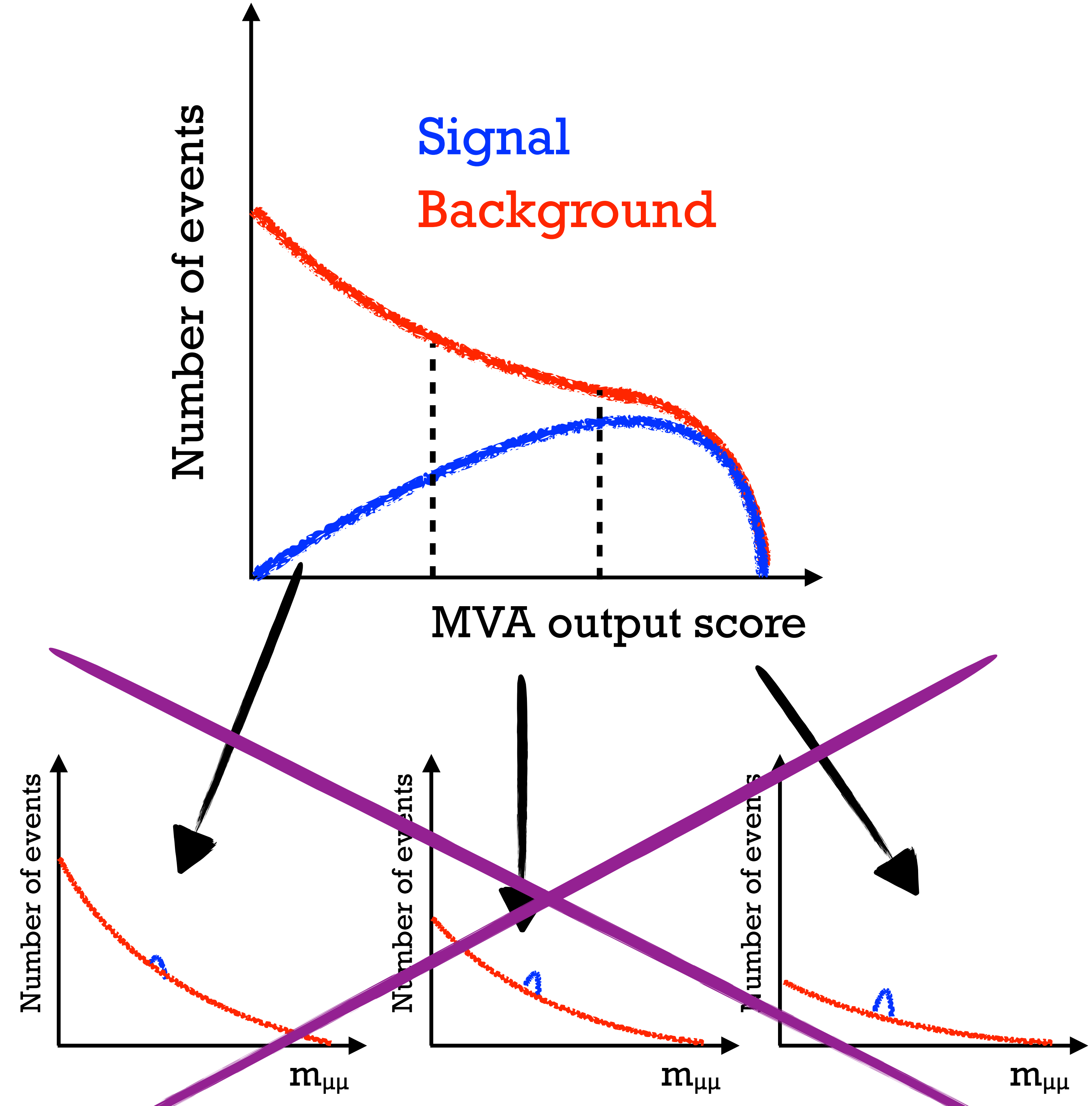
(representation of high purity VBF subcategory)

Revisiting the H→μμ strategy



MVA subcategories with varying signal purity

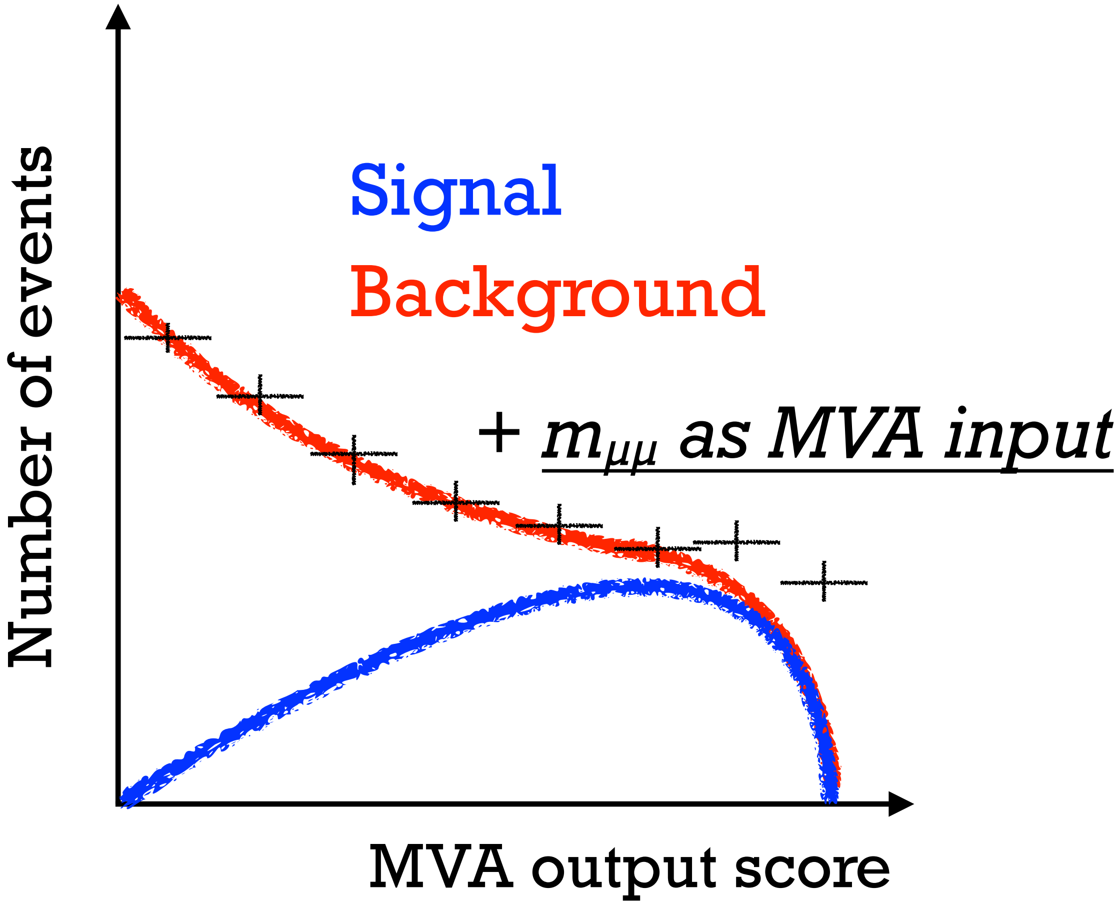
Revisiting the H→μμ strategy



~~MVA subcategories with varying signal purity~~

Revisiting the H→μμ strategy

- *New approach for the VBF category:*
  - Include **$m_{\mu\mu}$ as MVA input** variable.
  - **Fit MVA output directly** to extract H→μμ signal.
  - Take **background prediction from simulated samples.**
- Trading limited number of data events in mass sideband for systematic uncertainties in simulation.
- Following strategy similar to CMS EW Zjj measurement [1] and H→bb observation [2].



⇒ 20% improvement in performance in VBF category

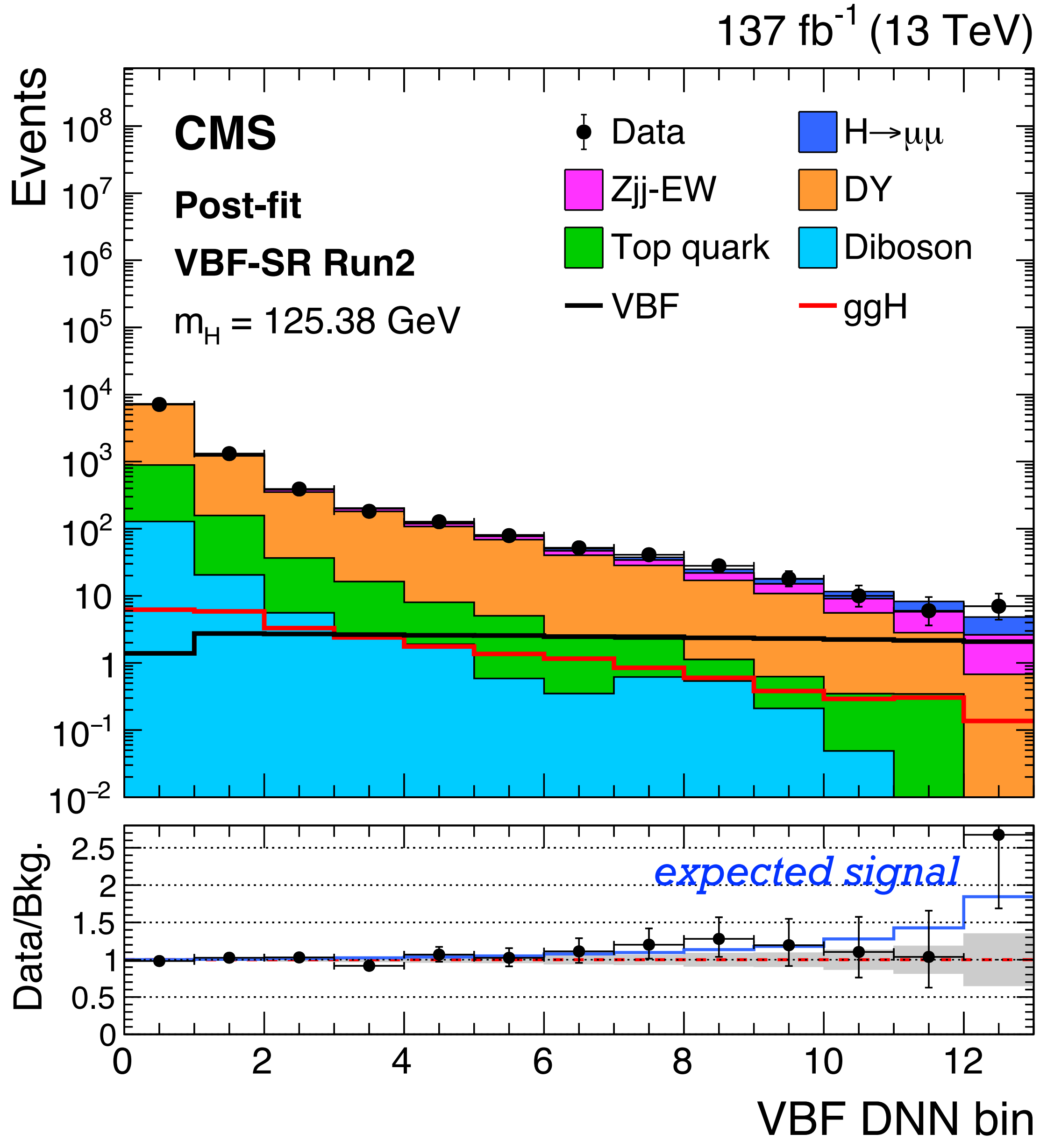
[1] Eur. Phys. J. C 78 (2018) 589

[2] Phys. Rev. Lett. 121 (2018) 121801

VBF DNN classifier

- *Train a deep neural network (DNN) including the H candidate mass $m_{\mu\mu}$ as an input.*

- DNN inputs targeting VBF H signal:
  - m_{jj}, $\Delta\eta(jj)$, $\Delta\phi(jj)$, min-$\Delta\eta(H,j)$, min-$\Delta\phi(H,j)$, p_T-balance(H,jj), H centrality, …
  - Suppressed hadronic activity in jet rapidity gap expected for VBF signal $\Rightarrow$ track-jet multiplicity and H_T in jet rapidity gap.



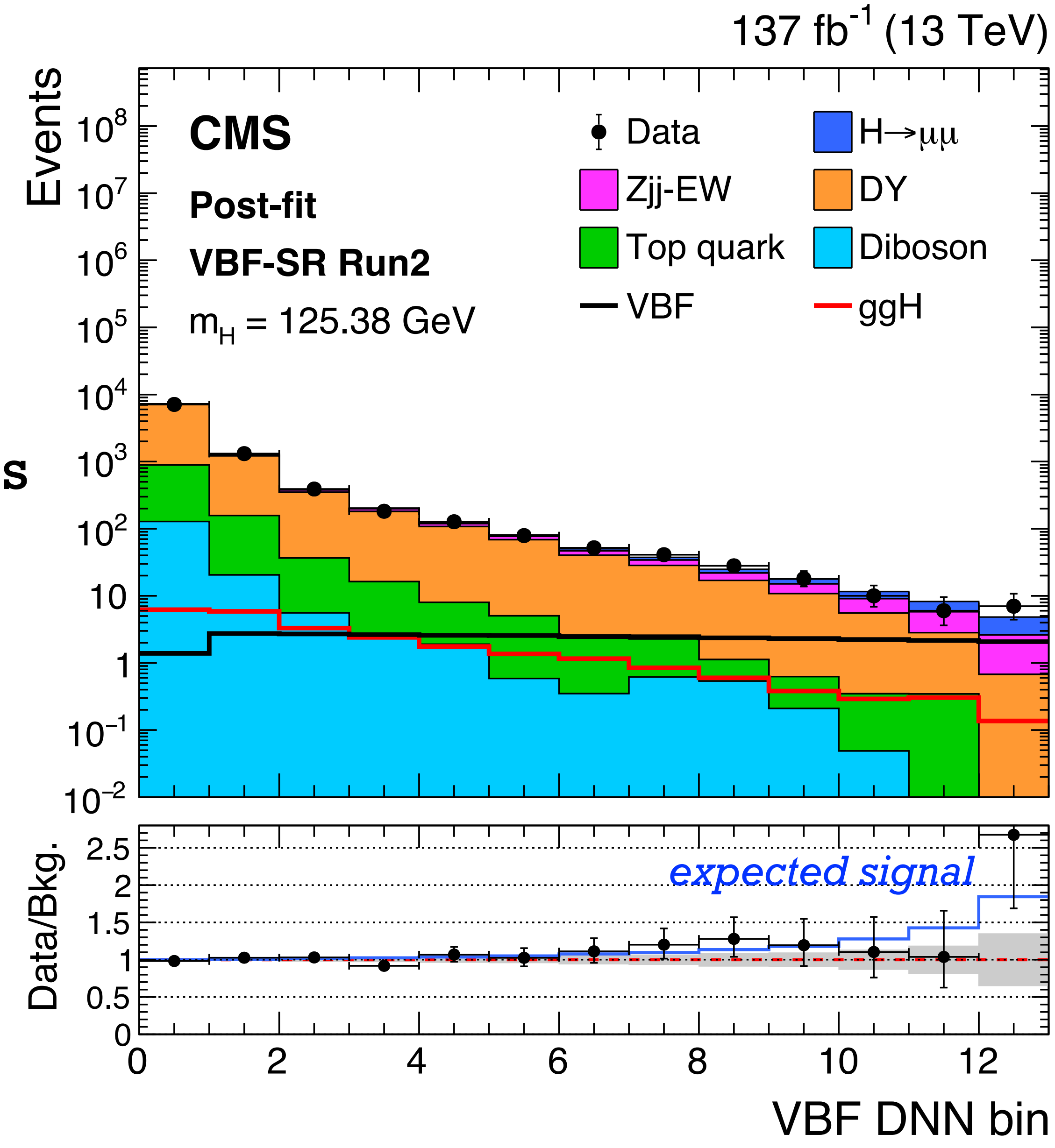
VBF category DNN score in signal region

VBF systematic uncertainties

Total systematic uncertainty impact <5%

- *Largest systematic uncertainties:*
    - VBF (H signal and Z background) parton shower modeling
    - Jet energy scale and resolution
    - DY contribution with one or more pileup jets
    - Statistical precision of simulated events
    - Theory uncertainties: missing higher order corrections, etc.

VBF category DNN score in signal region

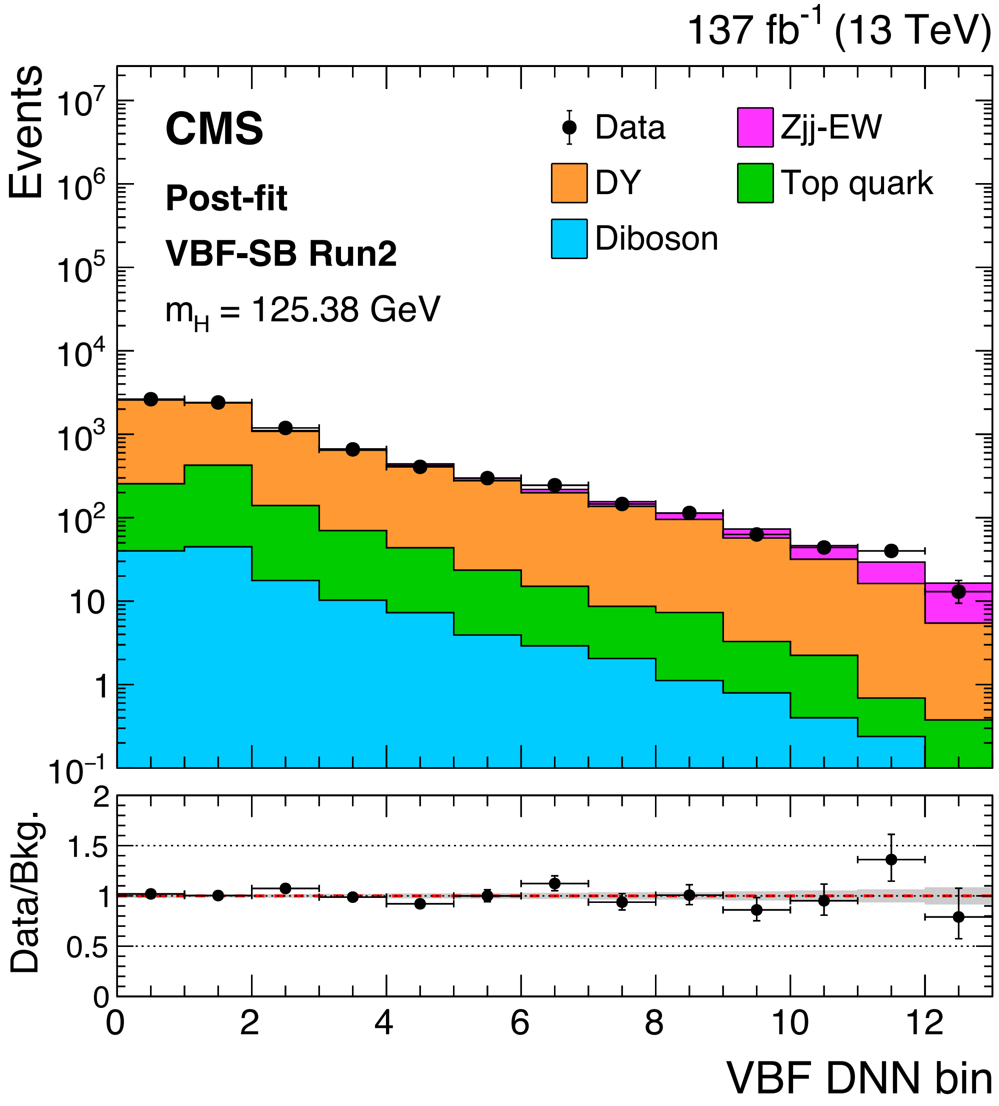


VBF category results

- Binned maximum-likelihood fit to DNN score simultaneously in the *signal region* and the *mass sideband region* to better constrain uncertainties.

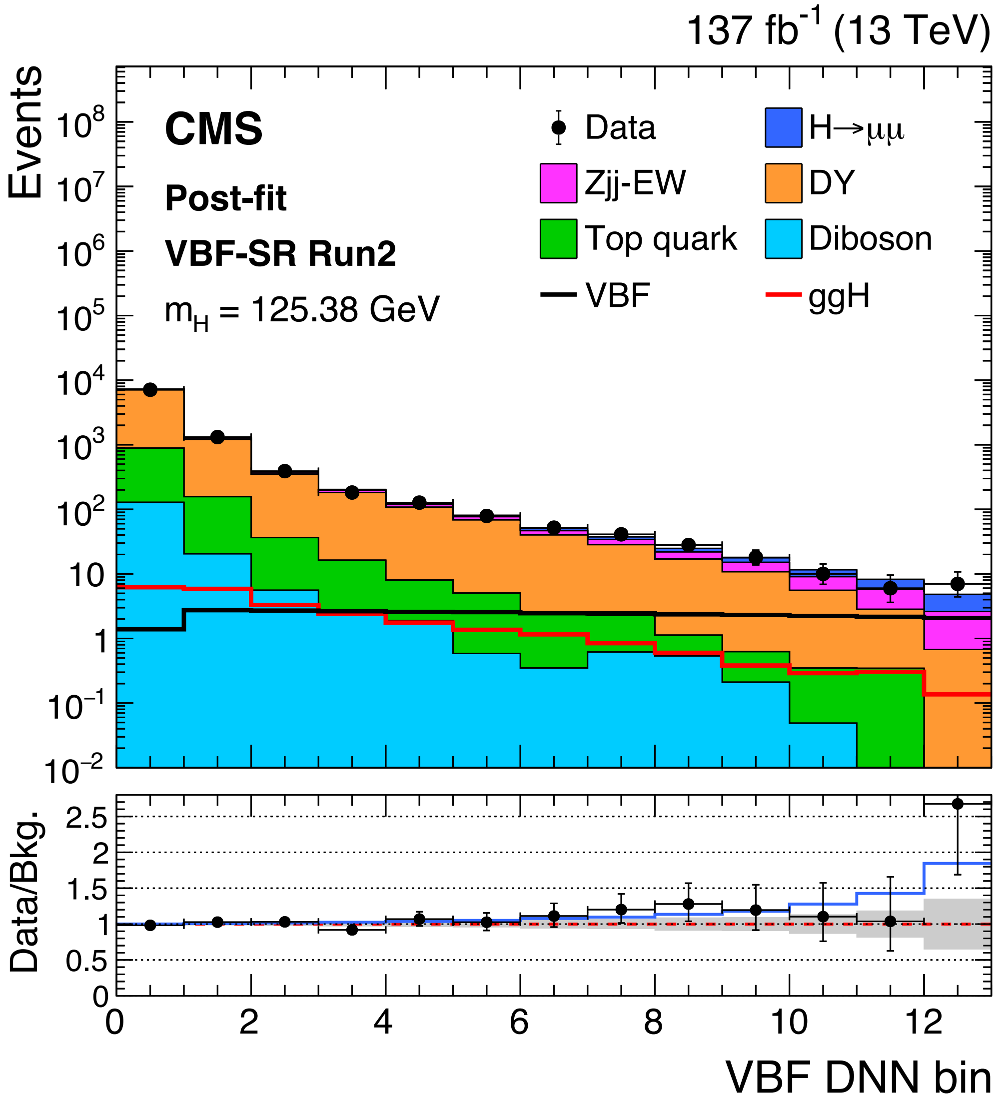
Mass sideband region:

$m_{\mu\mu}$ in [105,115] + [135,150] GeV



Signal region:

115 < $m_{\mu\mu}$ < 135 GeV



VBF category results

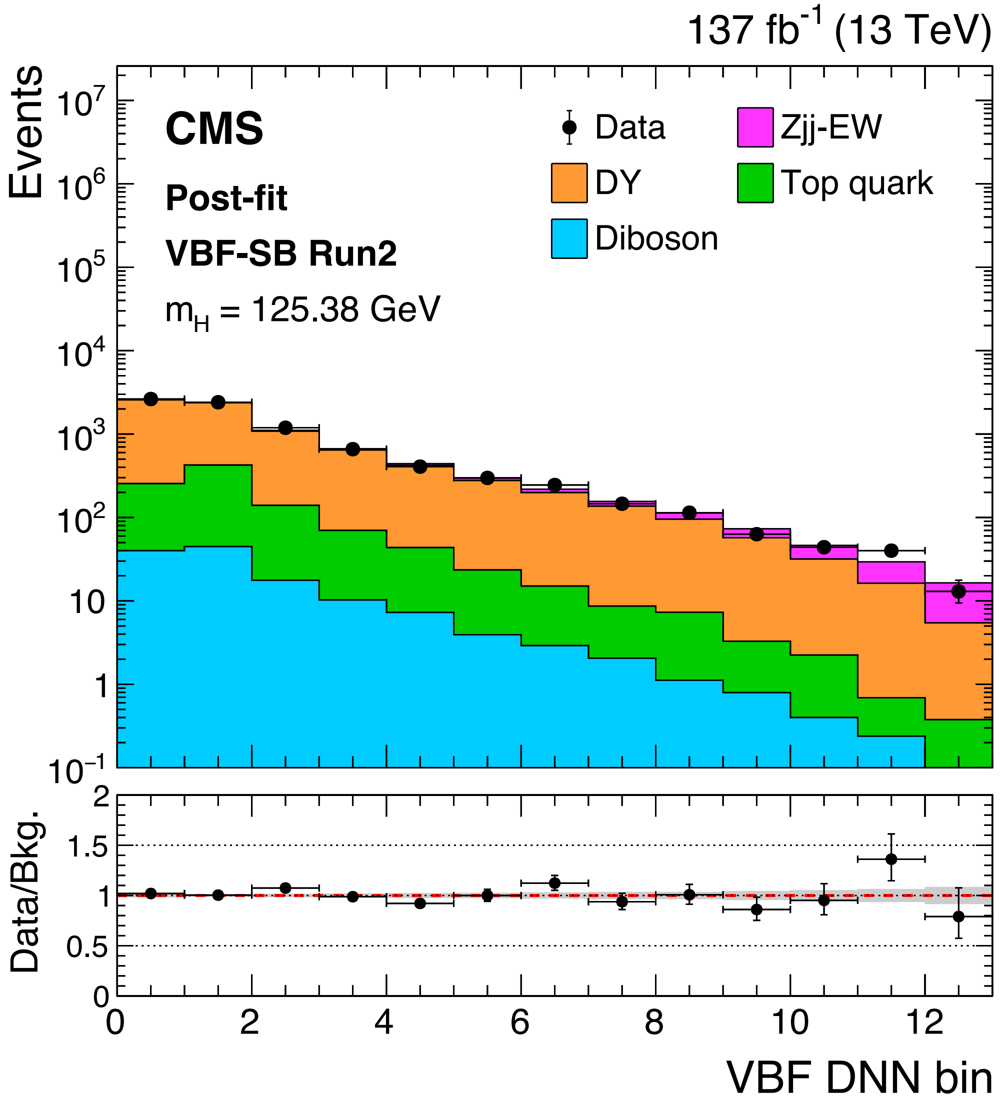
VBF category result

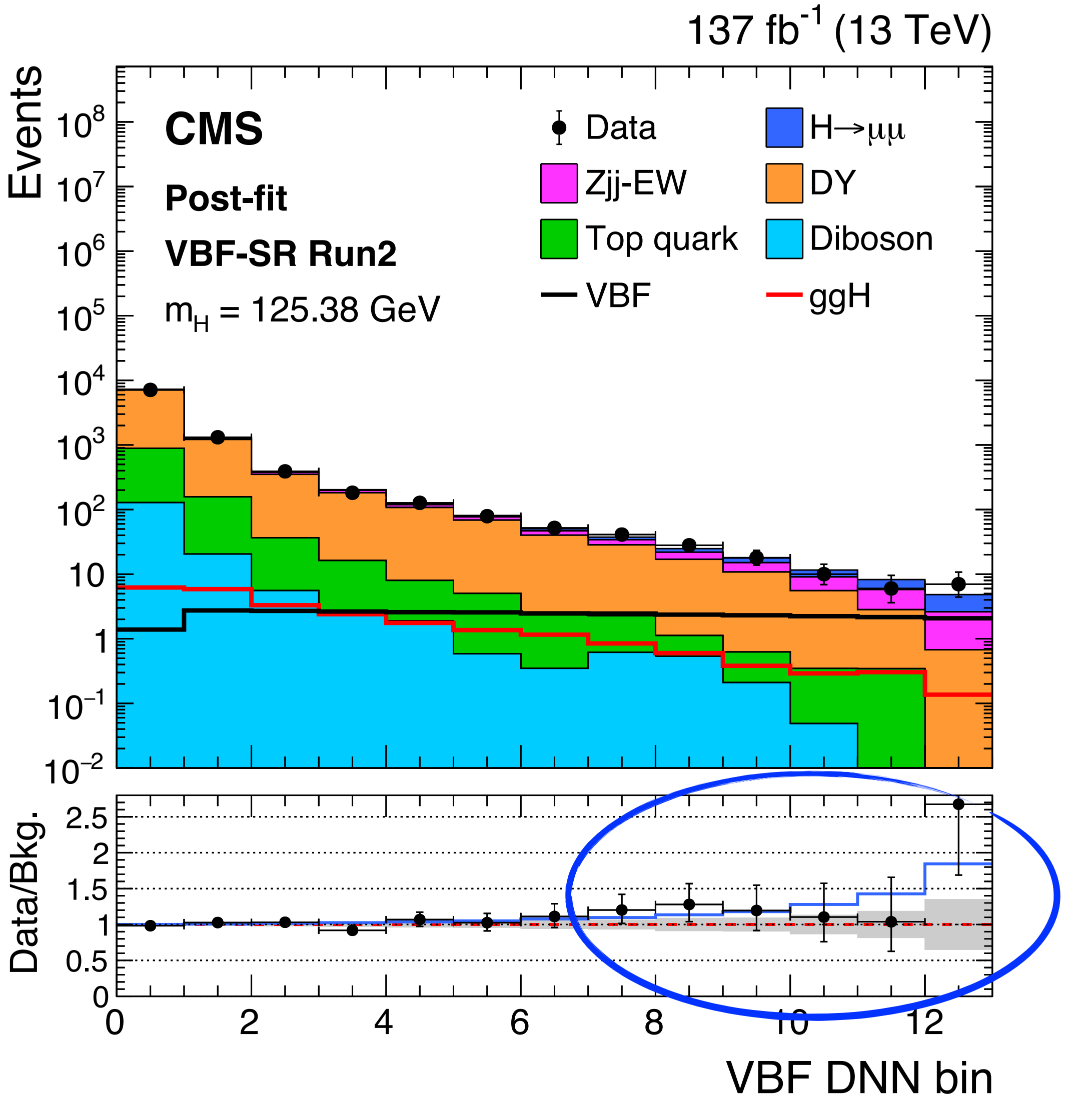
Observed (expected) significance: 2.4σ (1.8σ)

$$\mu = \frac{\sigma_{\mathrm{H}} \cdot \mathrm{BR}(\mathrm{H} \to \mu\mu)}{\sigma_{\mathrm{H}} \cdot \mathrm{BR}(\mathrm{H} \to \mu\mu)_{\mathrm{SM}}} = 1.36^{+0.69}_{-0.61}$$

Mass sideband region:
$m_{\mu\mu}$ in [105,115] + [135,150] GeV

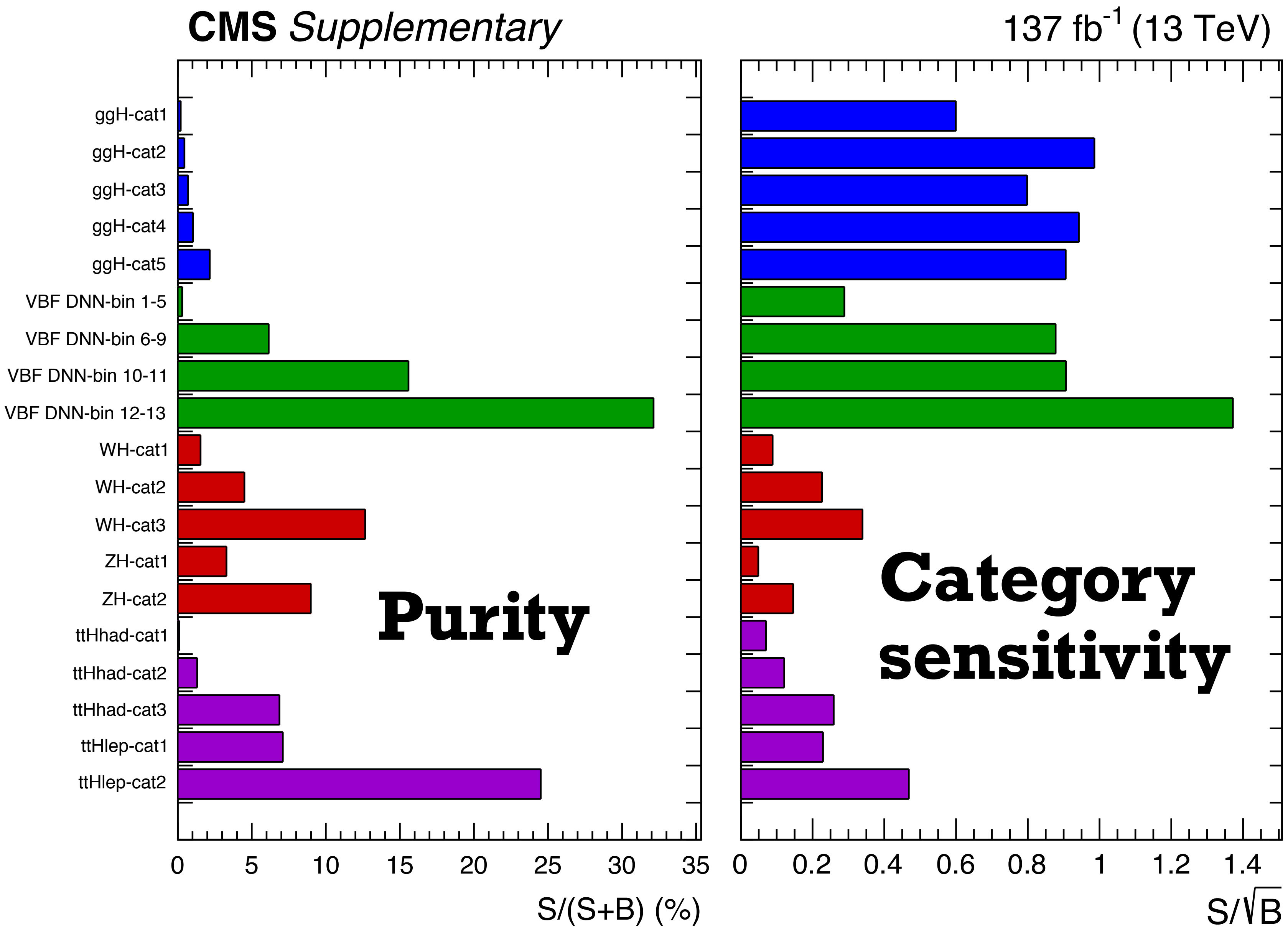
Signal region:
115 < $m_{\mu\mu}$ < 135 GeV





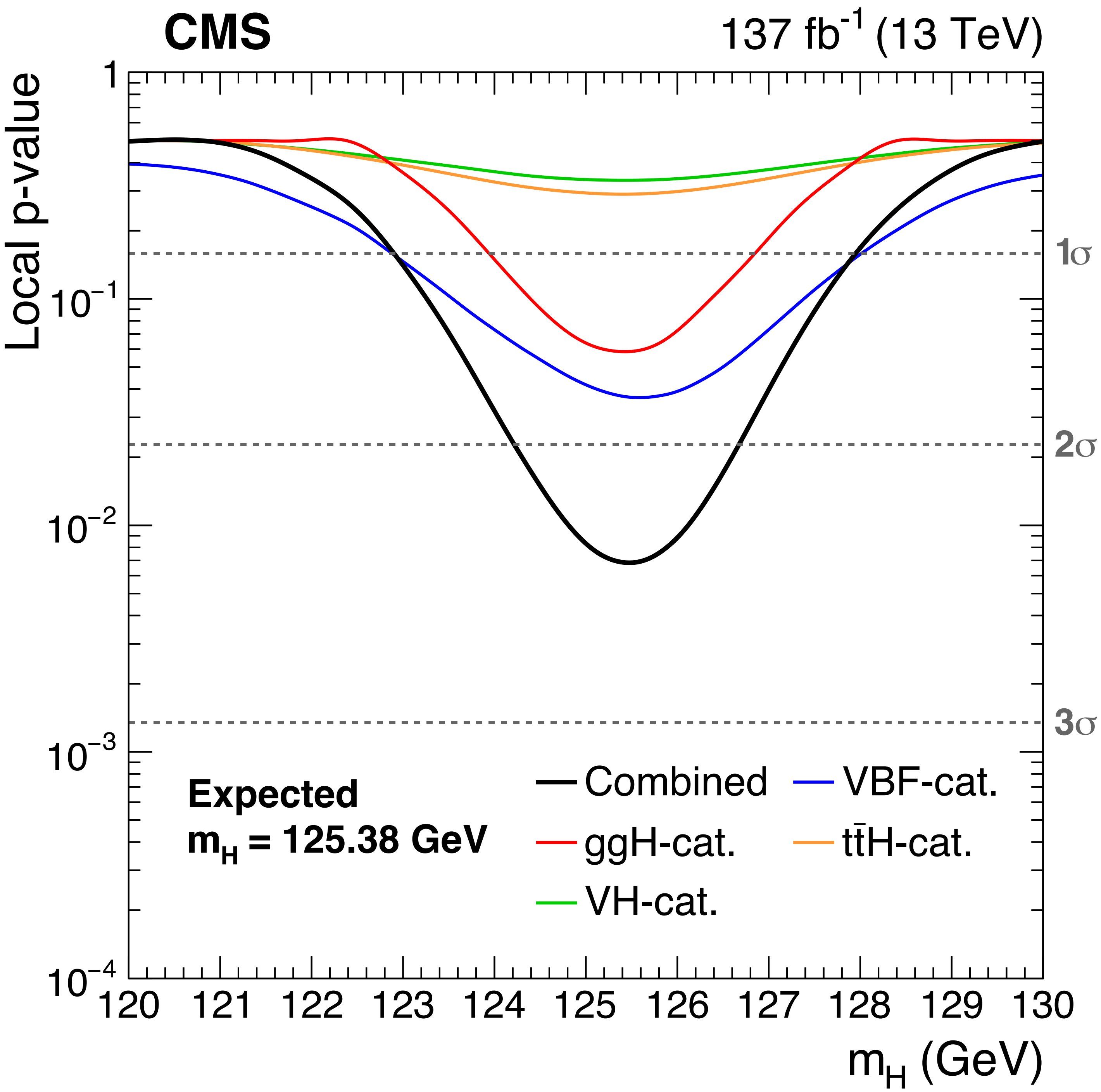
Putting it all together

- Sensitivity of each category is a balance between purity and signal yield.
- ggH and VBF category sensitivities are comparable, with VBF slightly better.
- ttH and VH categories are strongly limited by small signal yield.



Putting it all together: expected

- Simultaneous fit of all categories to extract combined H→μμ signal.
- *VBF category* has the best expected sensitivity, followed by the *ggH category*.



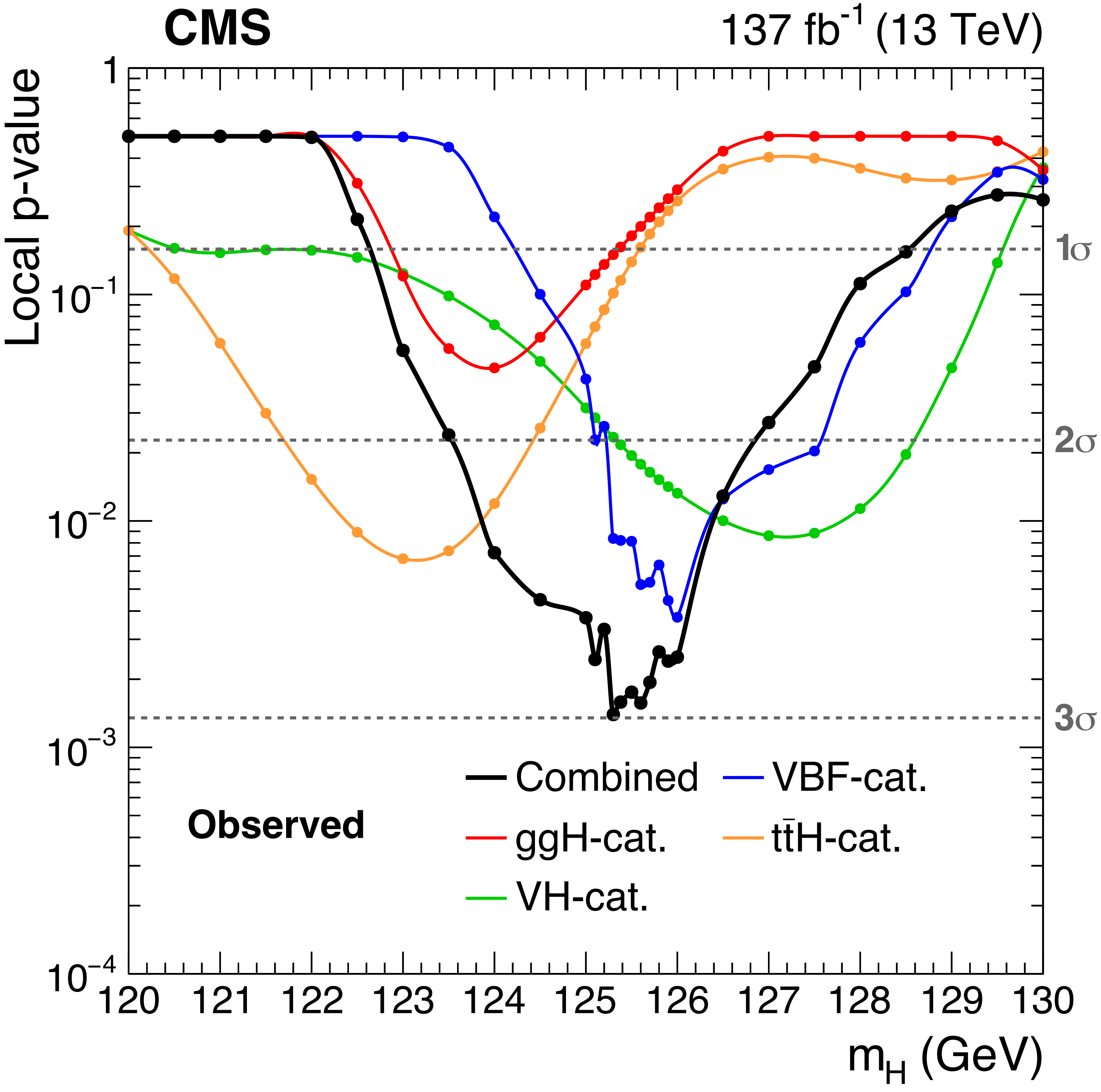
Expected (m_H = 125.38 GeV) combined significance: 2.5σ

Putting it all together: observed

- Excess over background-only prediction observed at m_H = 125.38 GeV with a statistical significance of 3.0σ.

- *This constitutes the first evidence for the H→μμ decay.*

- Fluctuations in the observed p-value arising from discrete nature of the mass profiling in the VBF category.



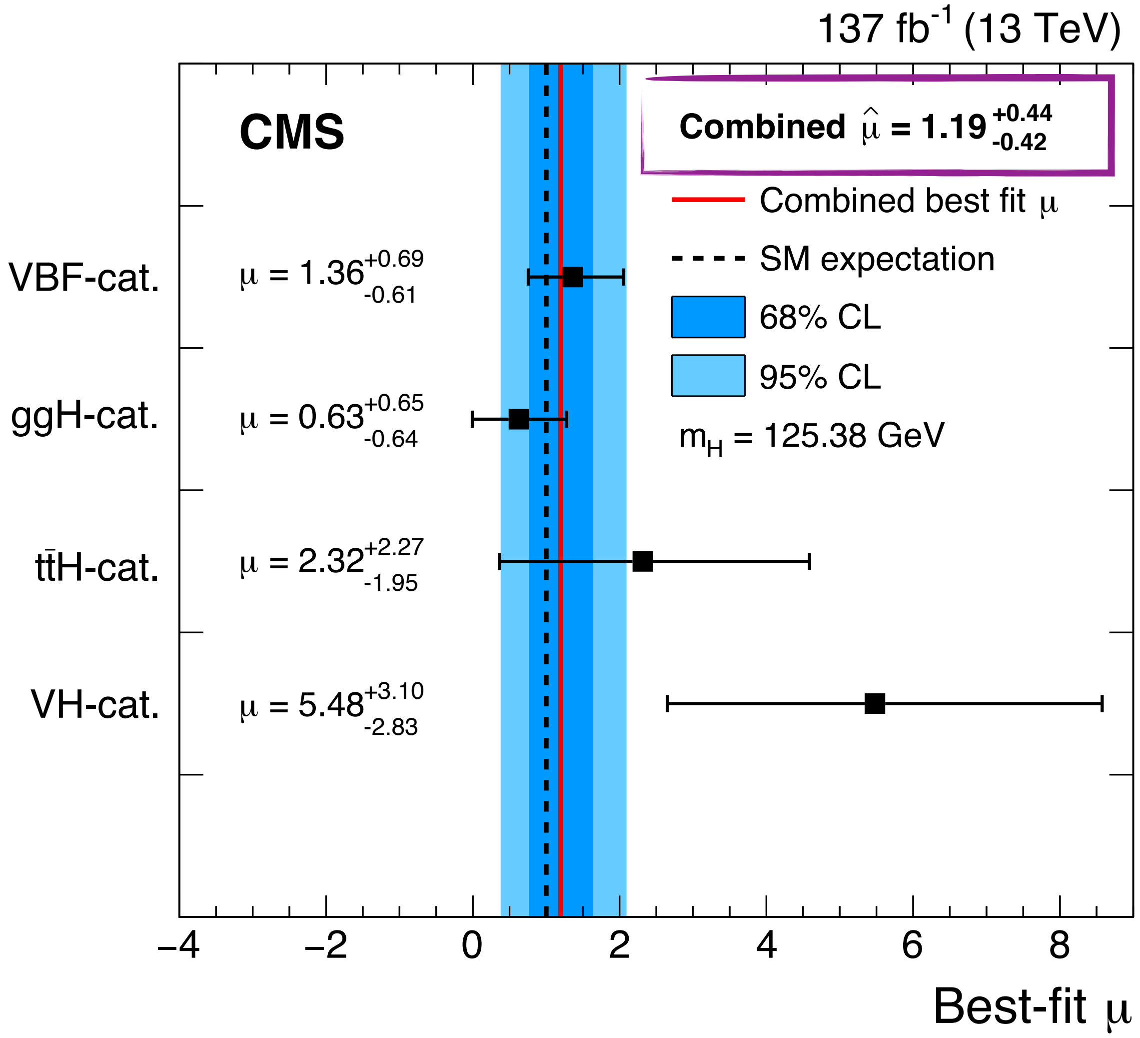
Observed (m_H = 125.38 GeV) combined significance: 3.0σ

Measured signal properties

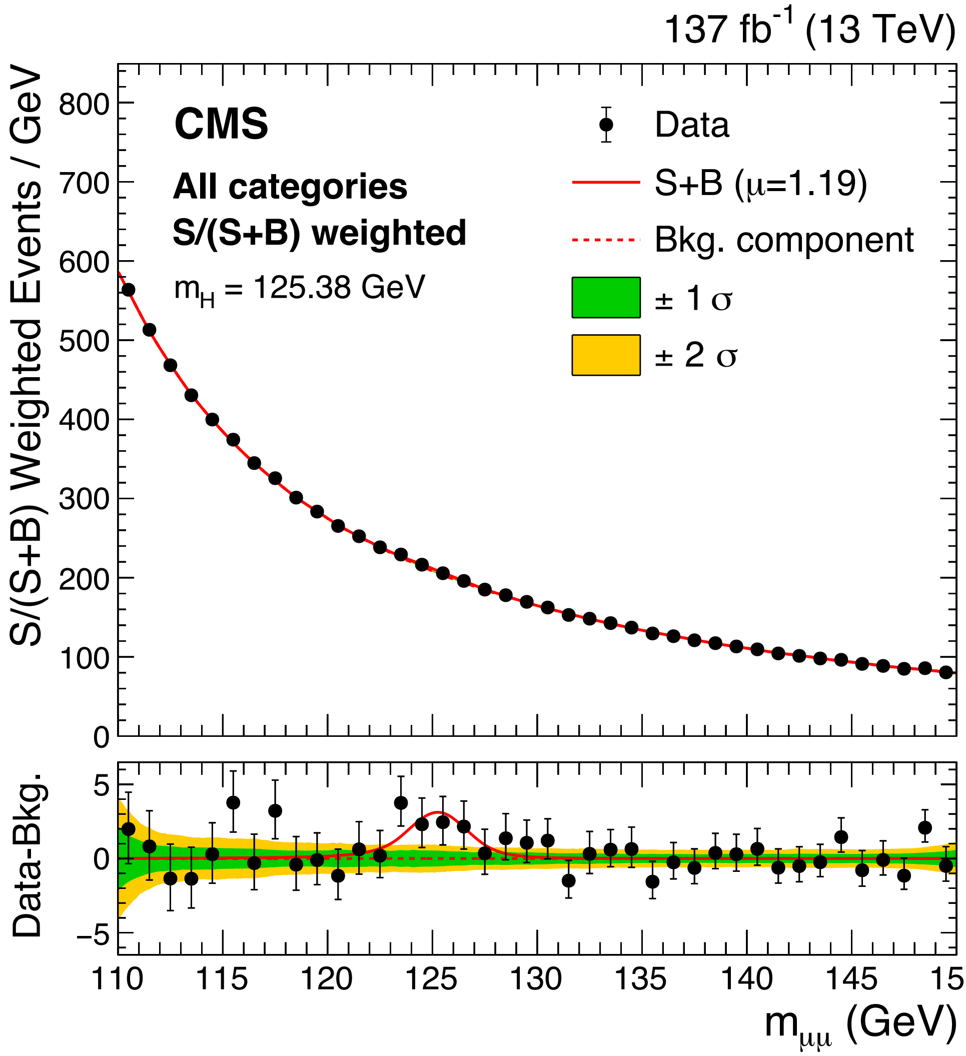
- *The observed signal is well compatible, within uncertainties, with the SM expectation for the Higgs boson interaction with the muon.*

- Dominant uncertainties are statistical ⇒ with more data, we will more precisely test this interaction.

Signal strength per category



Visualization of the observed excess vs. $m_{\mu\mu}$



Direct constraints on H-μ interaction

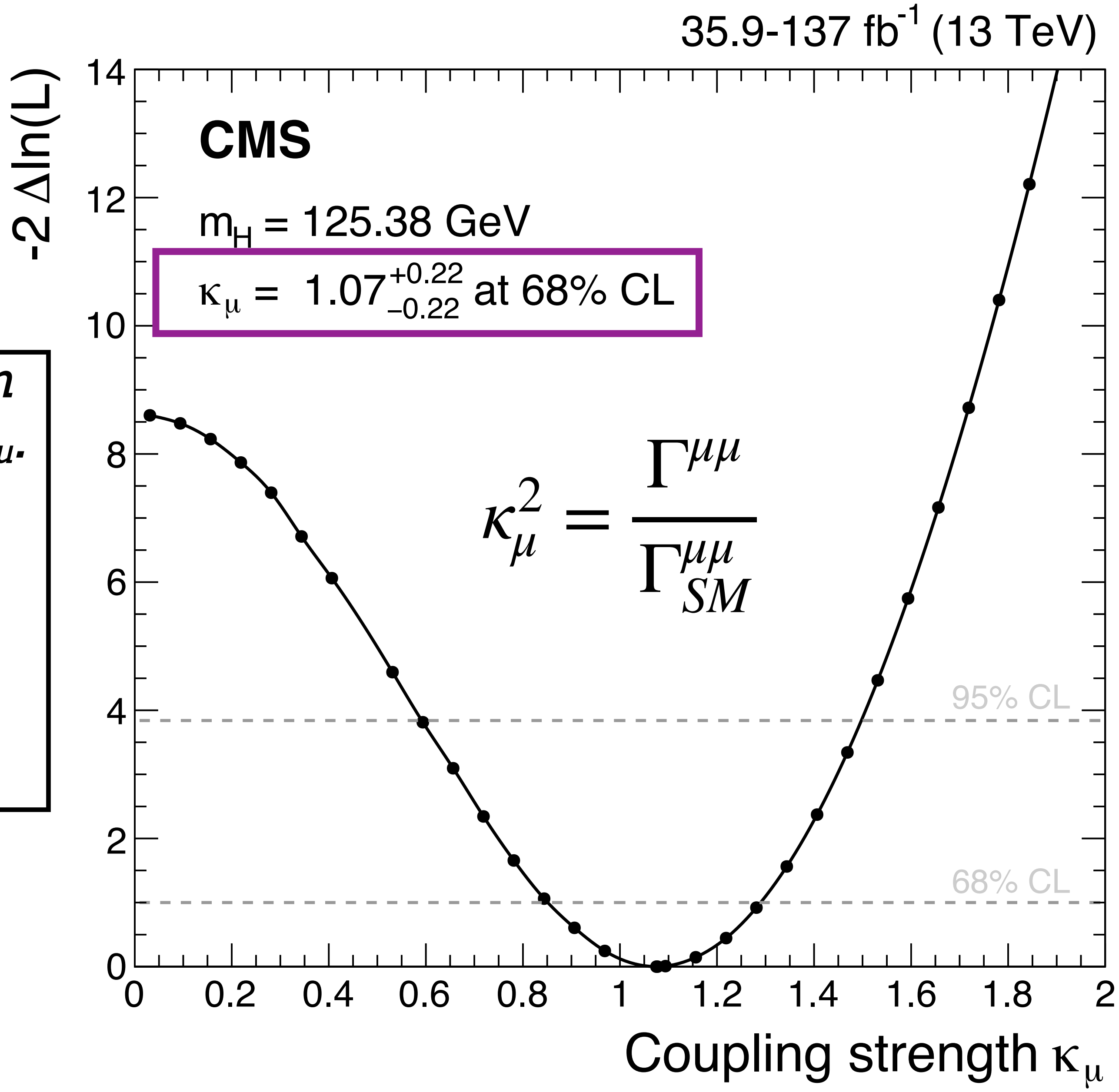
Direct access to H-μ interaction

$$\mu = \sigma_H \cdot \mathrm{BR}(H \rightarrow \mu\mu)$$

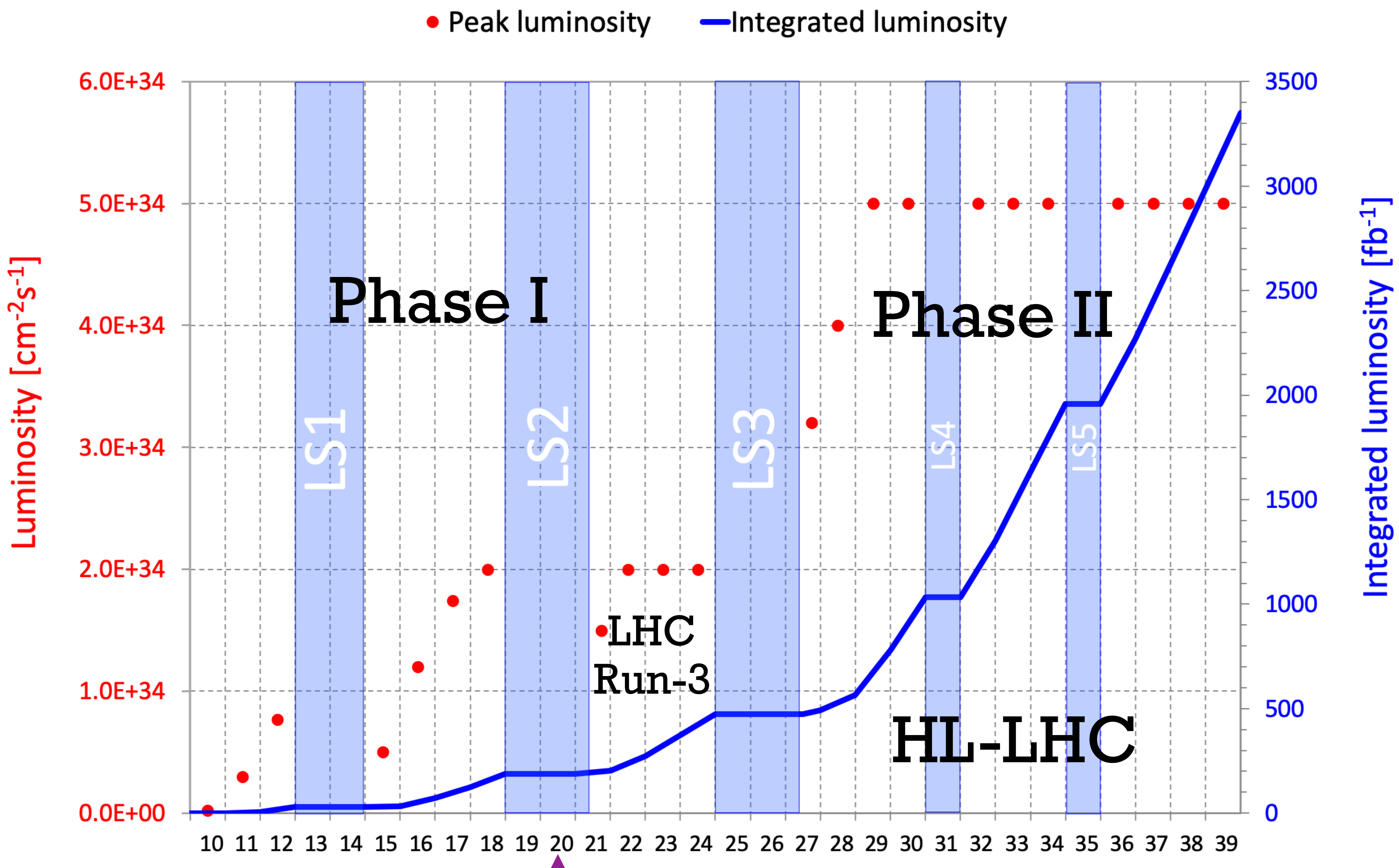
Depends on Higgs boson interaction strength with t, W, Z, …

- *Combine with other Higgs channels to obtain direct constraints on H-μ coupling strength κ_μ.*
- *Coupling strength of Higgs boson to muon measured with nearly 20% precision at 68% confidence level.*

H-μ coupling strength from combination with measurements of other Higgs channels (EPJC 79 (2019) 421)



(HL) LHC schedule



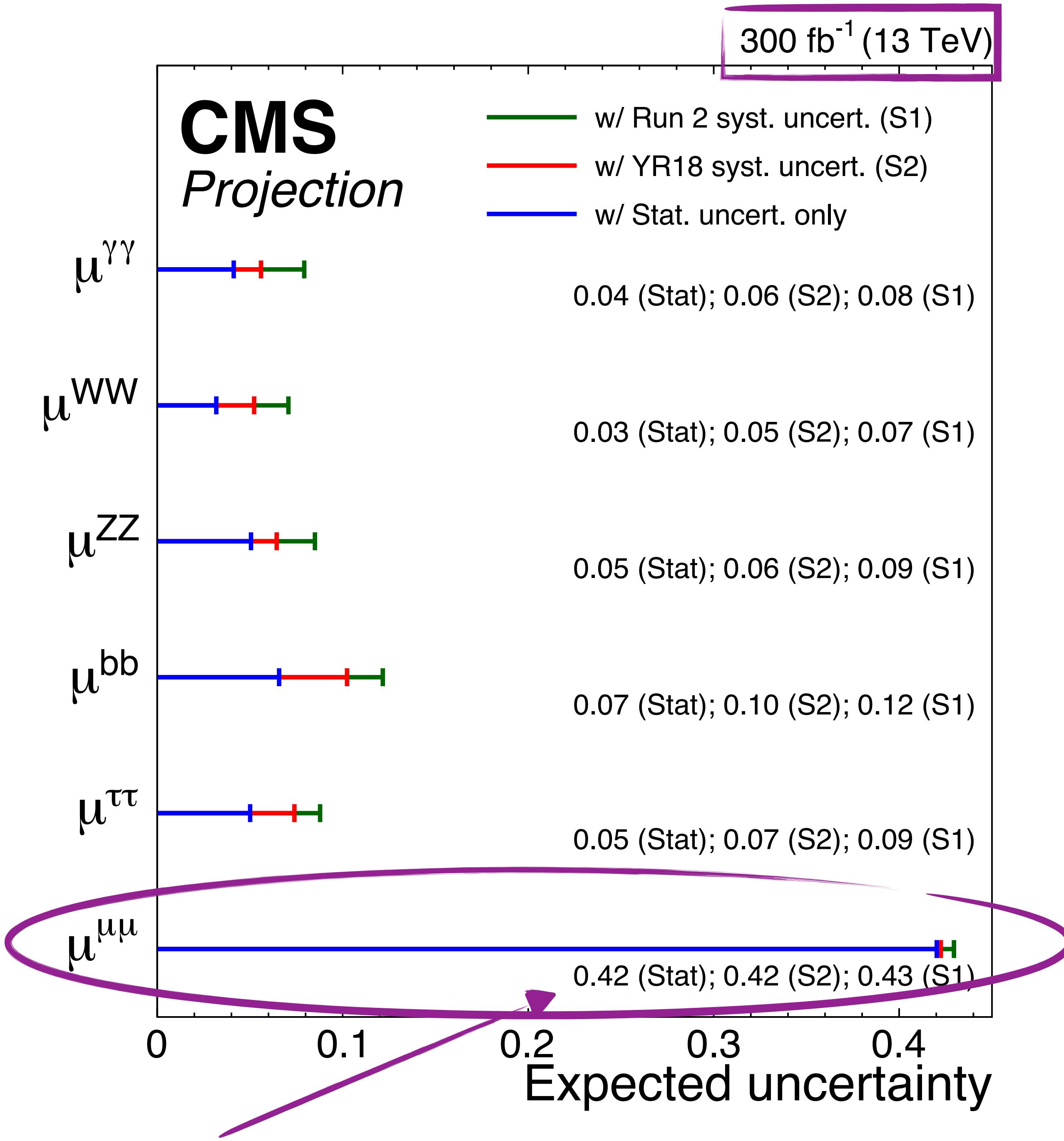
We are here (few % of final dataset)

Future prospects for H→μμ

- CMS has released projections of expected precision of Higgs measurements up to end of HL-LHC.
  - Including H→μμ, based on projections from previous CMS H→μμ publication [1].
- *The future is hard to predict!*
  - Cannot extrapolate ingenuity, new ideas, new methods, precise performance and usage of new technologies, etc.
- So please keep this in mind…

CMS-PAS-FTR-18-011

(based on previous H→μμ published analysis)



We are already here with 137 fb^{-1} !

[1] Phys. Rev. Lett. 122 (2019) 021801

H→μμ with Run-3 data

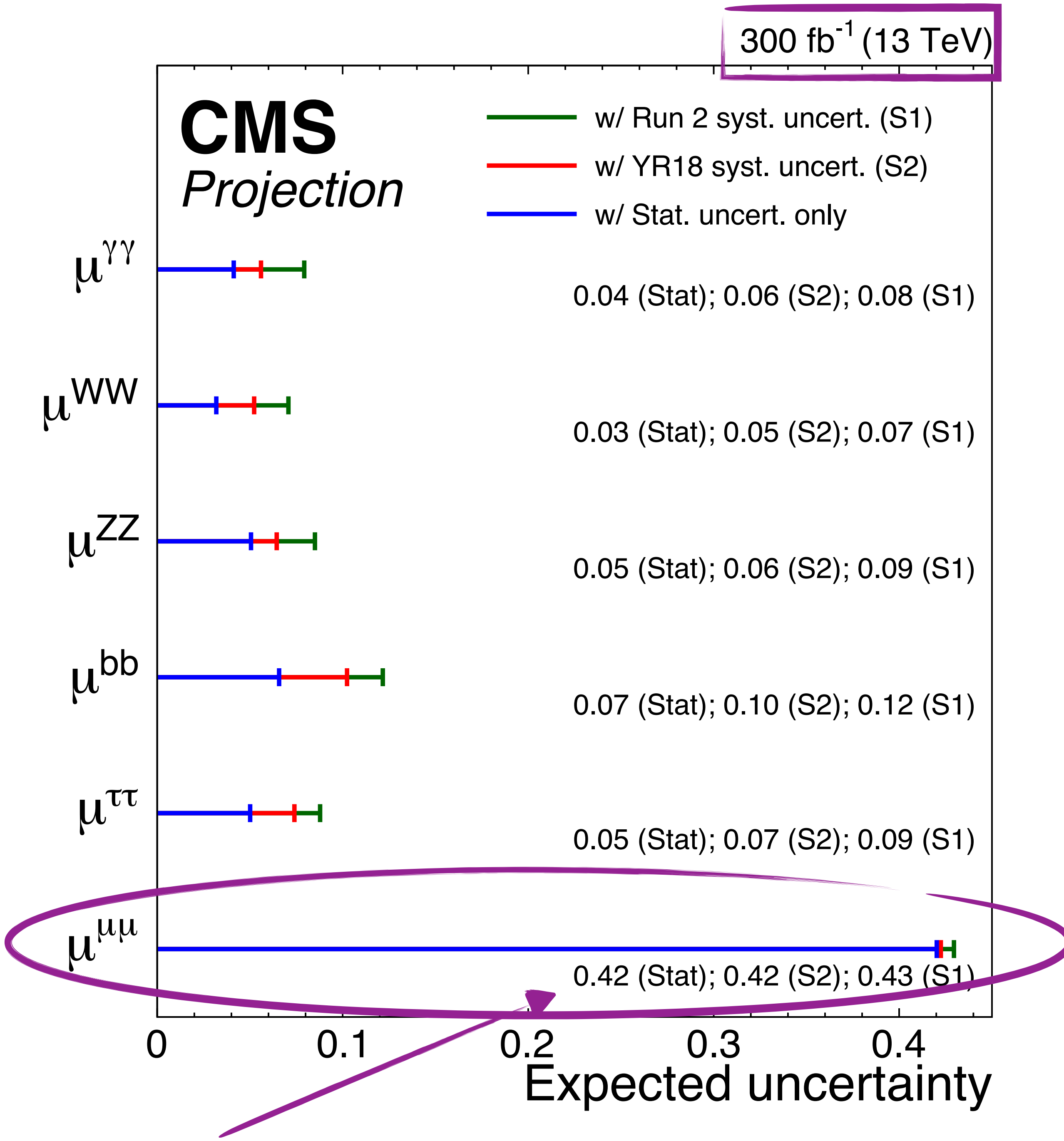
- LHC Run-3 (2022-2024) expected to deliver about 200 fb^{-1} of 13 or 14 TeV data.
  - Data conditions similar to those at end of Run-2.

- Some small upgrades to CMS detector, but similar performance expected overall.

- Assuming no significant improvements to the analysis, expect *roughly 4σ sensitivity to H→μμ including Run-3 data*.

CMS-PAS-FTR-18-011

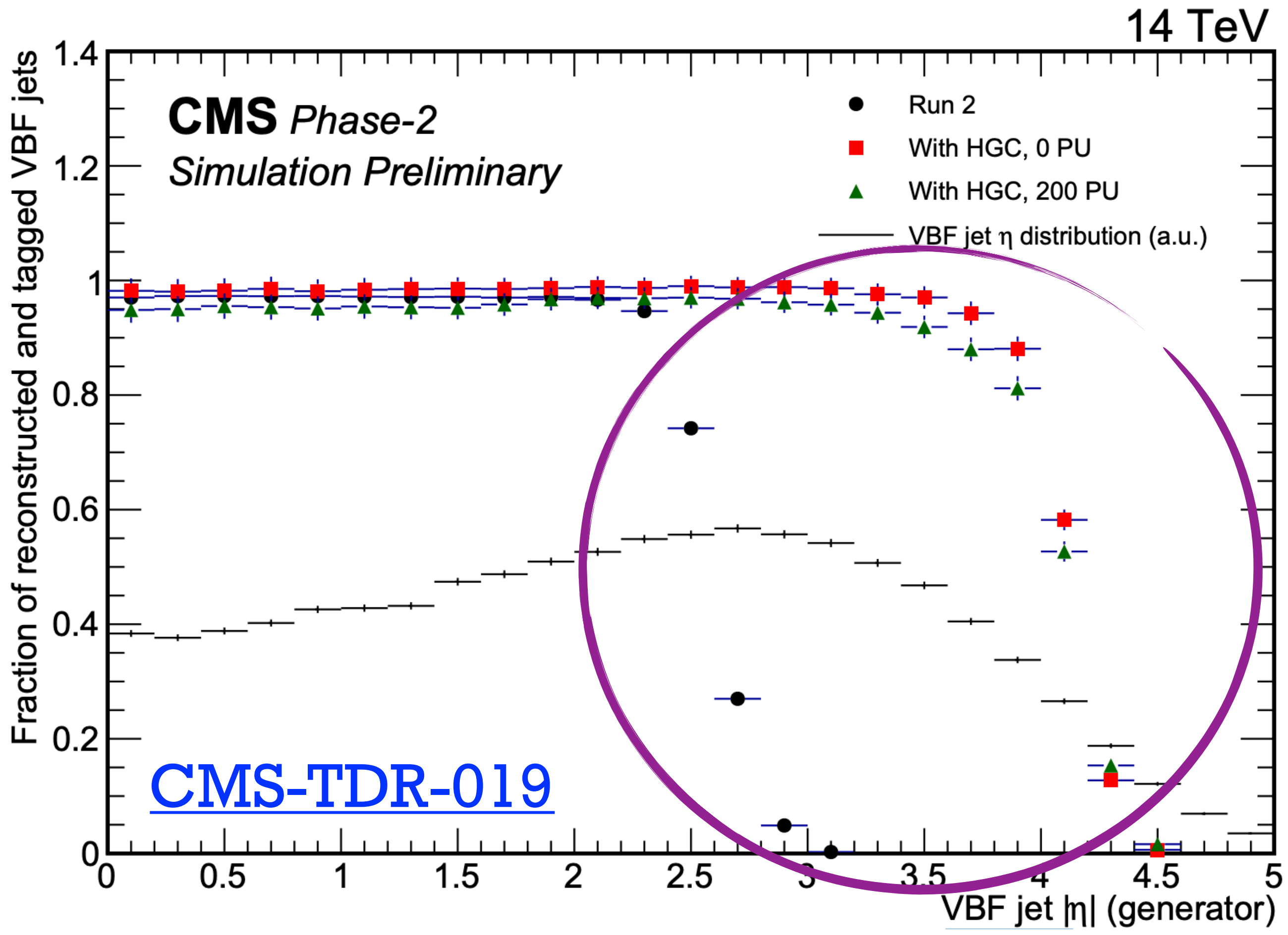
(based on previous H→μμ published analysis)



We are already here with 137 fb^{-1} !

H→μμ at the HL-LHC

- CMS detector will largely be fully redesigned, improved.
    - Tracking coverage extended in η from 2.4 to 4.0.
    - High-granularity forward calorimetry with improved resolution.
    - Improved muon resolution and efficiency, η(μ) extended from 2.4 to 2.8.



CMS-TDR-019

Projections: 3 ab^{-1} of HL-LHC data

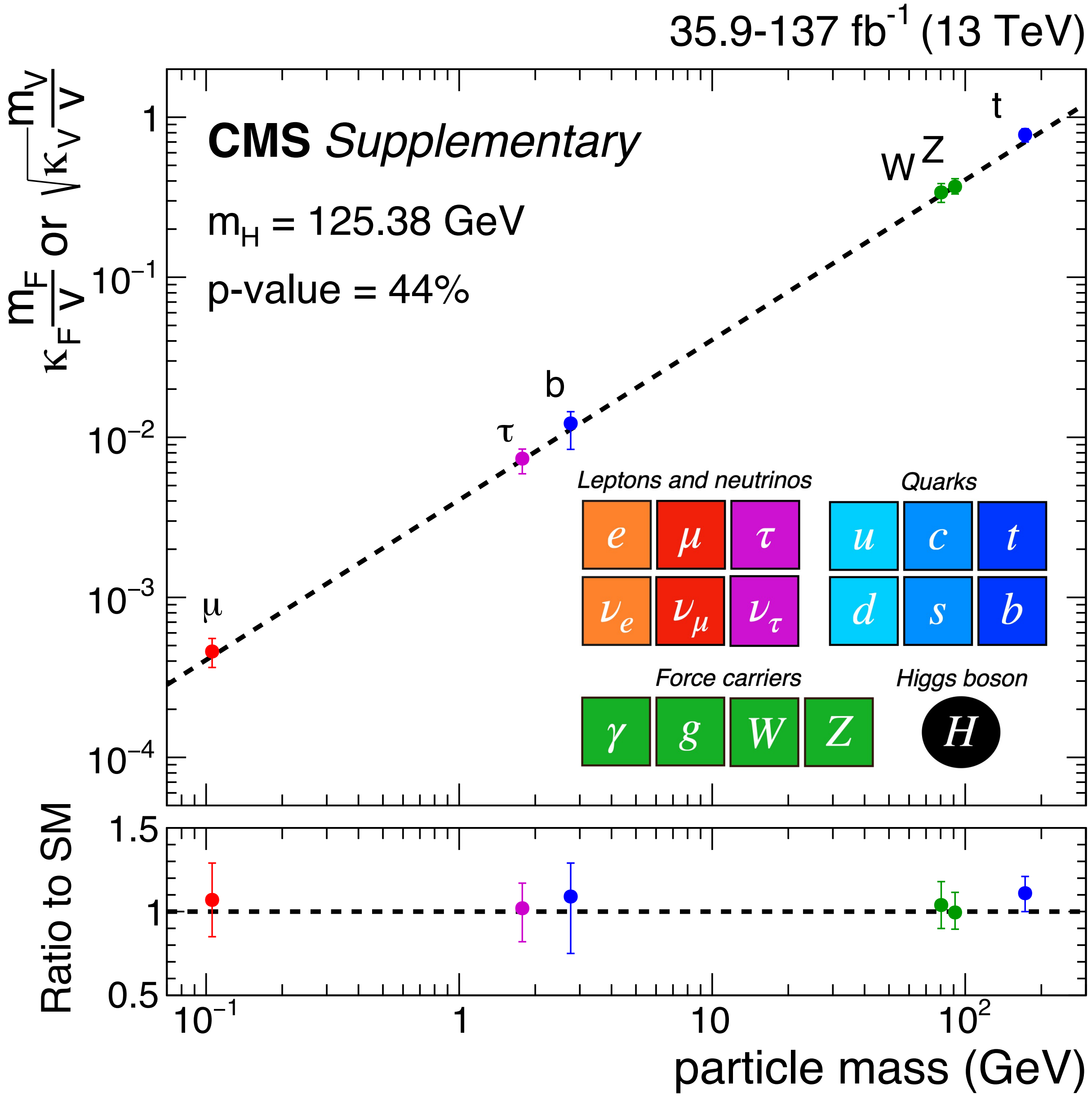
Experiment	CMS	
Process	Combination	
Scenario	S1	S2
Total uncertainty	13%	10%
Statistical uncert.	9%	9%
Experimental uncert.	8%	2%
Theory uncer.	5%	3%

arXiv:1902.00134

- Currently projecting ~10% uncertainty on H→μμ, i.e. ~5% precision on κ_μ.
    - Note that projection only considers increased dataset and improved mass resolution.

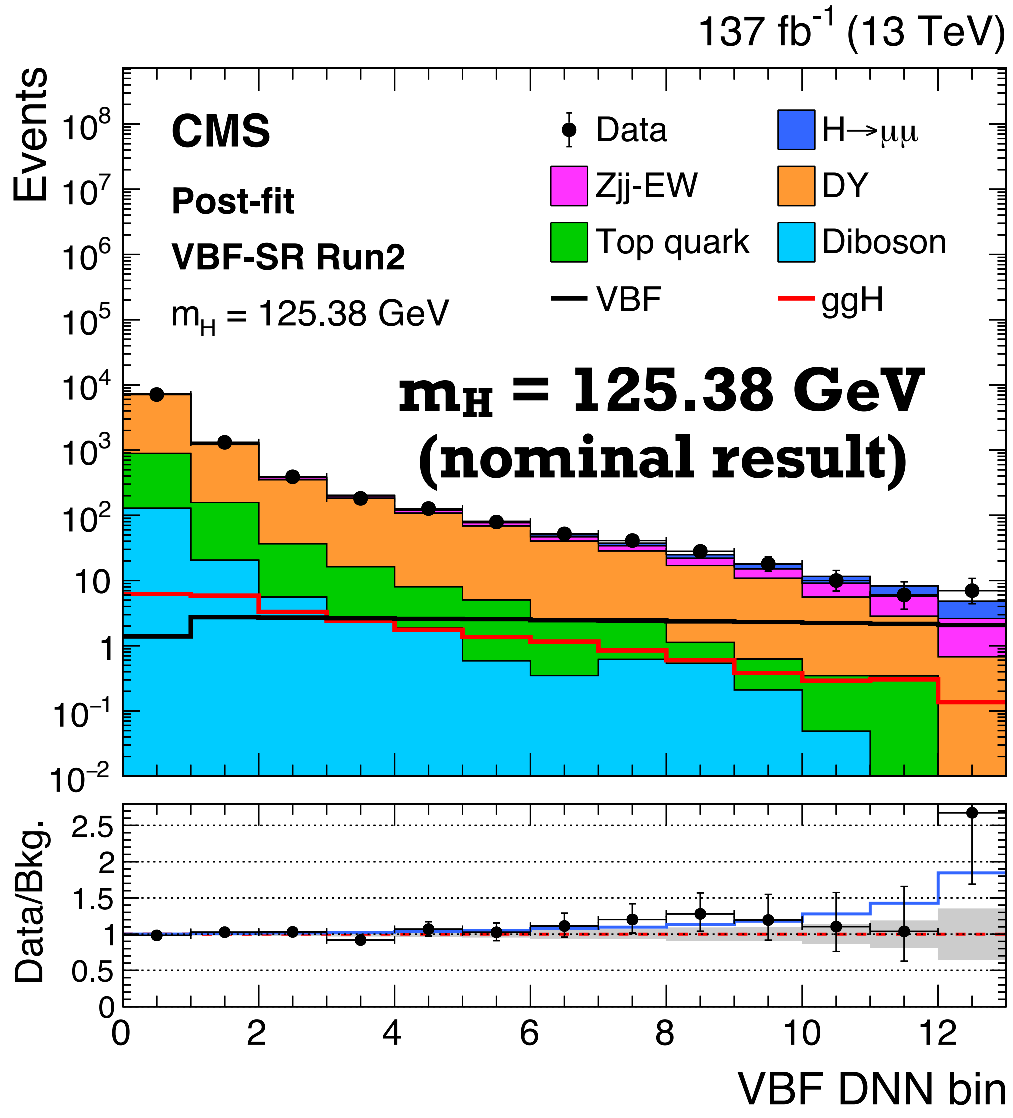
Summary

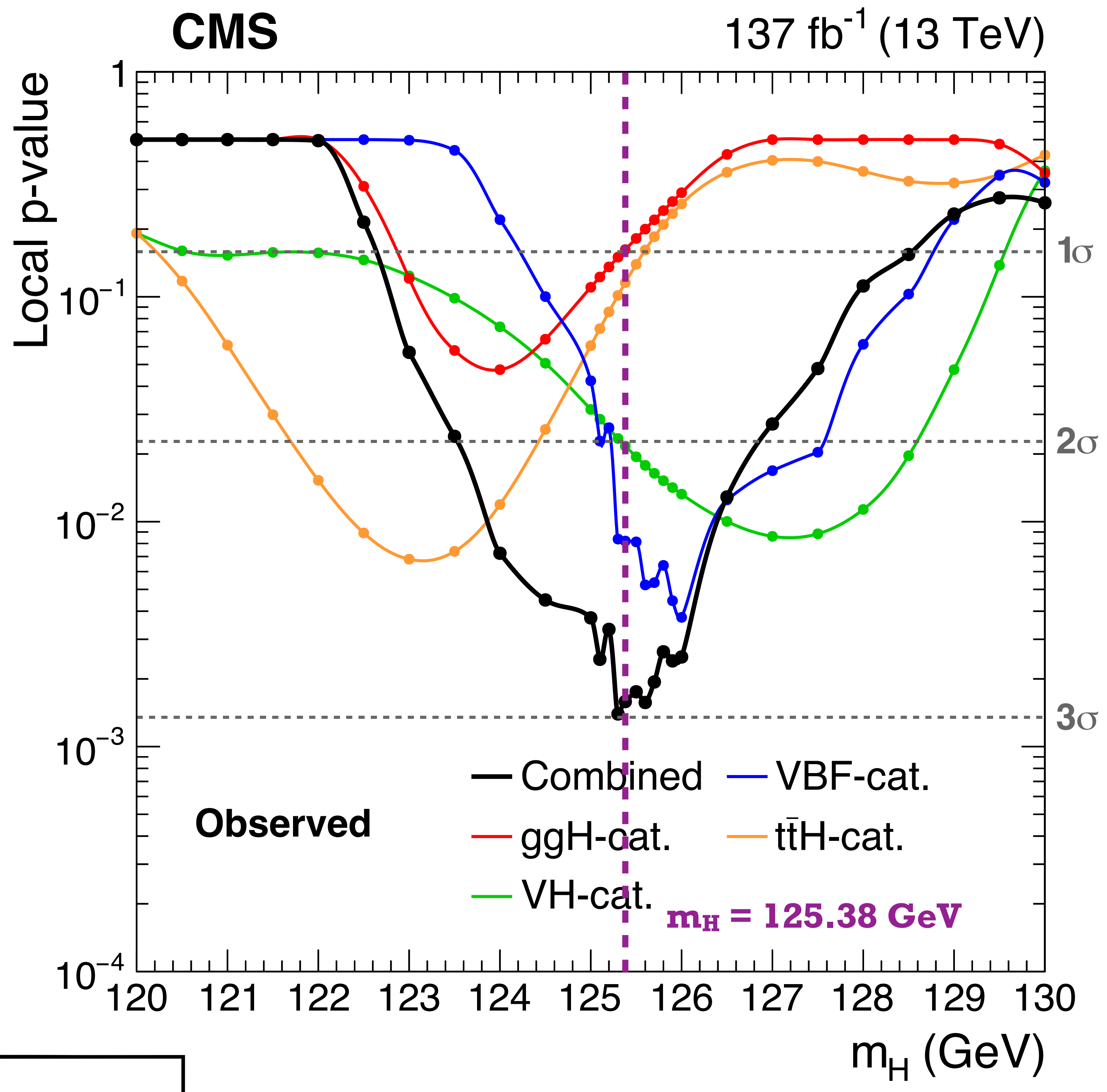
- Measurement of H→μμ performed by CMS with 137 fb^{-1} of 13 TeV Run-2 data.
- Observed (expected) significance: 3.0 (2.5)σ.
- First evidence for H→μμ decay.
- First evidence of Higgs boson interaction with second generation fermions.
- *Remarkable success of the standard model (unfortunately) continues!*



Additional Material

VBF DNN distribution evolution with mass



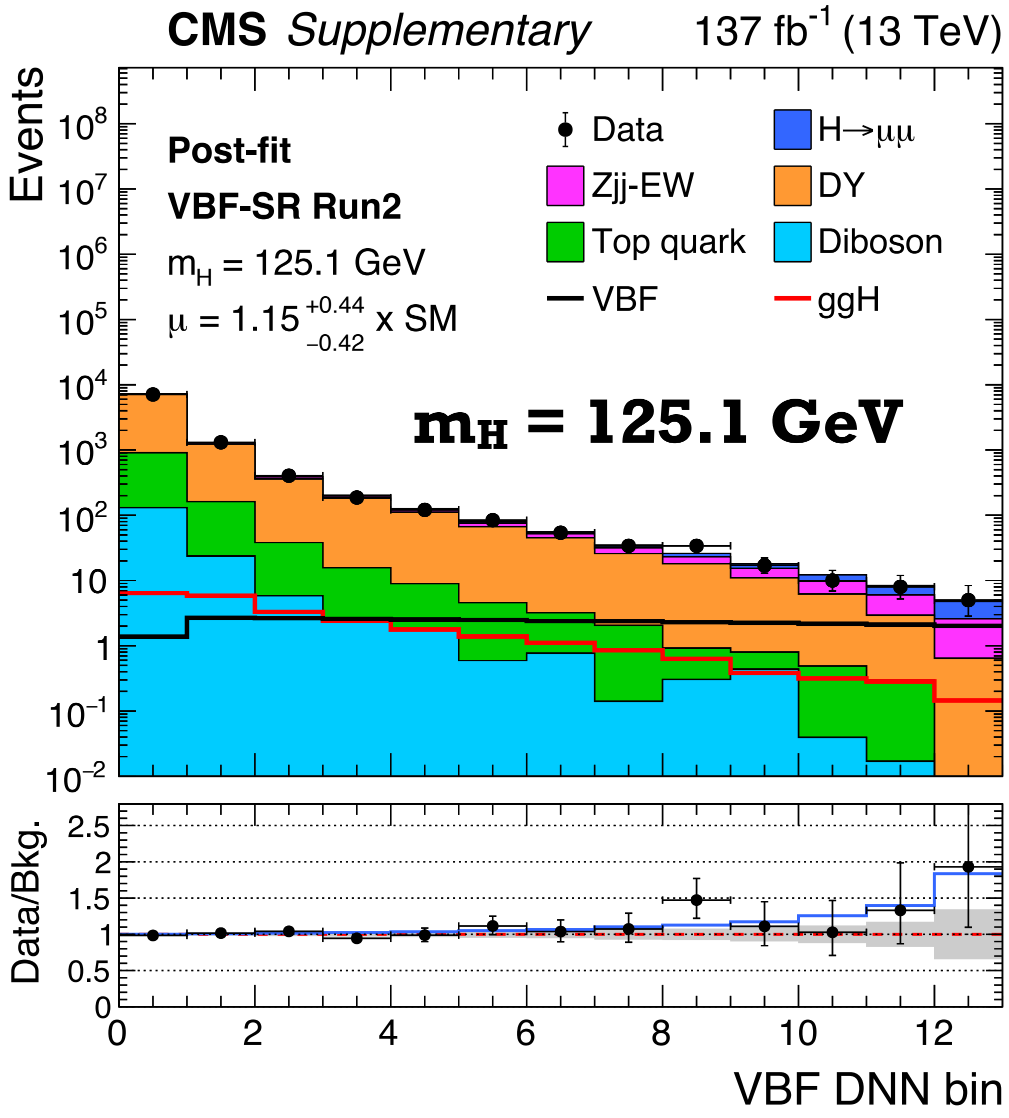


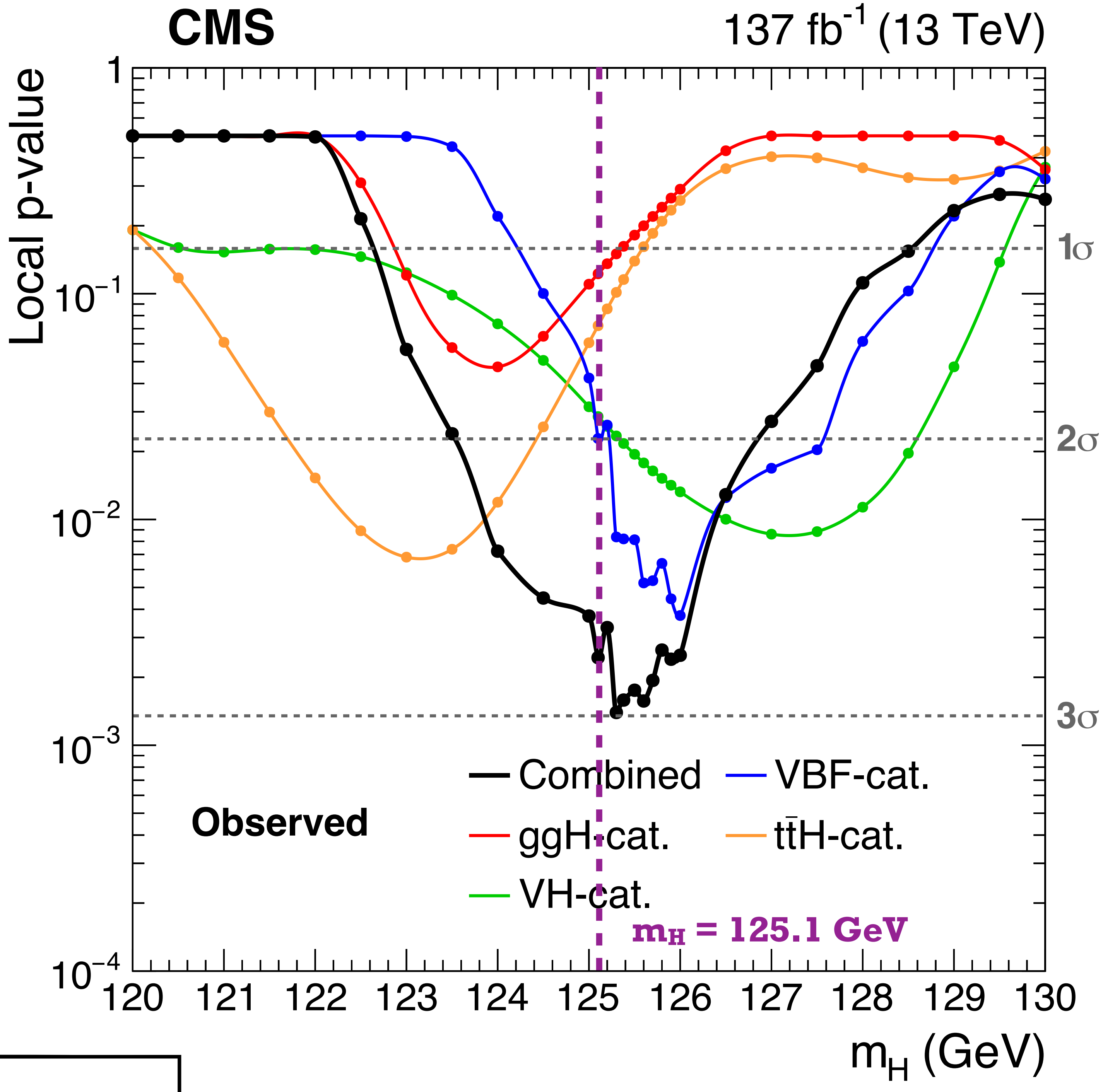
Small change in m_H hypothesis leads to:

⇒ ***shuffling of a few data events at high DNN score.***

⇒ ***discrete jumps in observed p-value at few percent level.***

VBF DNN distribution evolution with mass



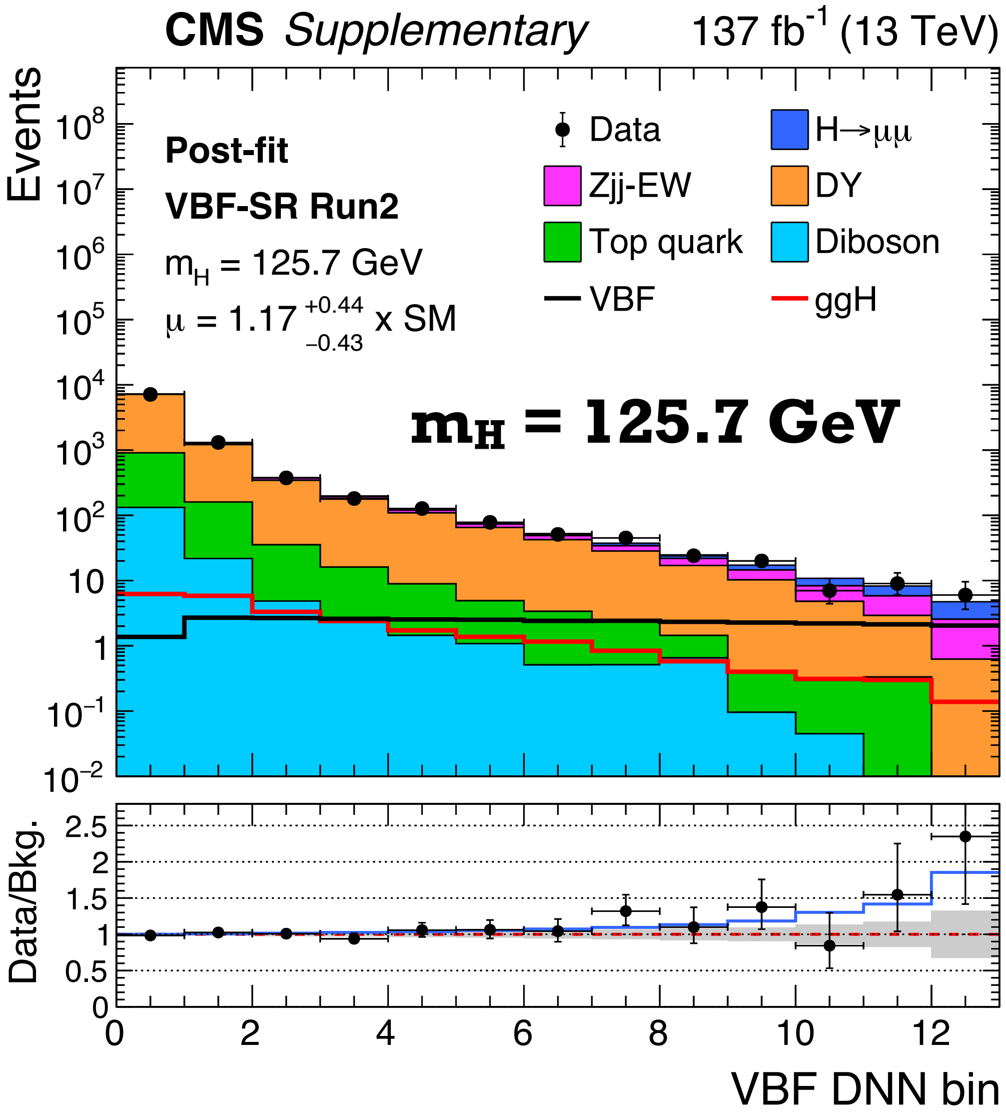


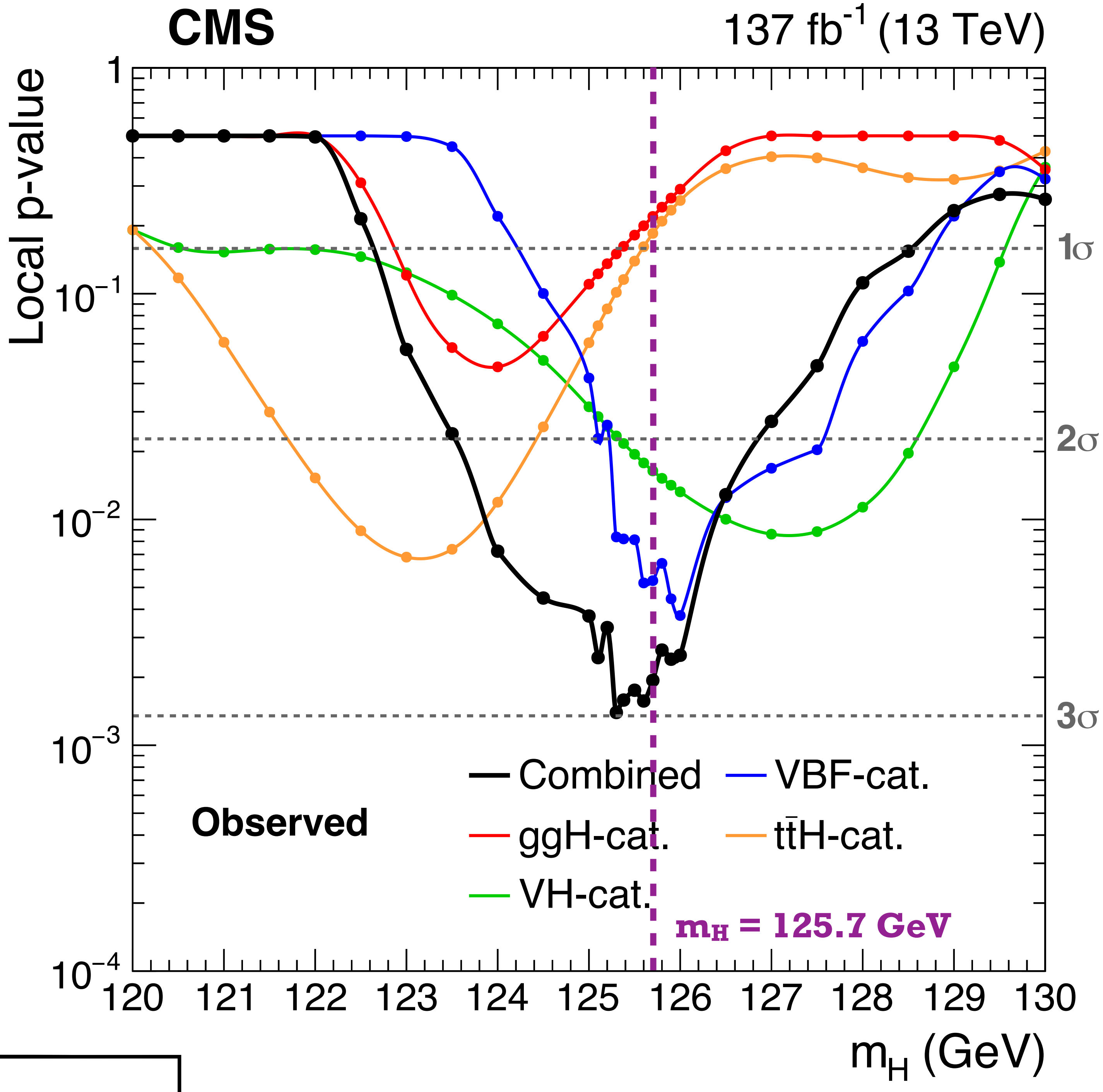
Small change in m_H hypothesis leads to:

⇒ shuffling of a few data events at high DNN score.

⇒ discrete jumps in observed p-value at few percent level.

VBF DNN distribution evolution with mass





Small change in m_H hypothesis leads to:

⇒ ***shuffling of a few data events at high DNN score.***

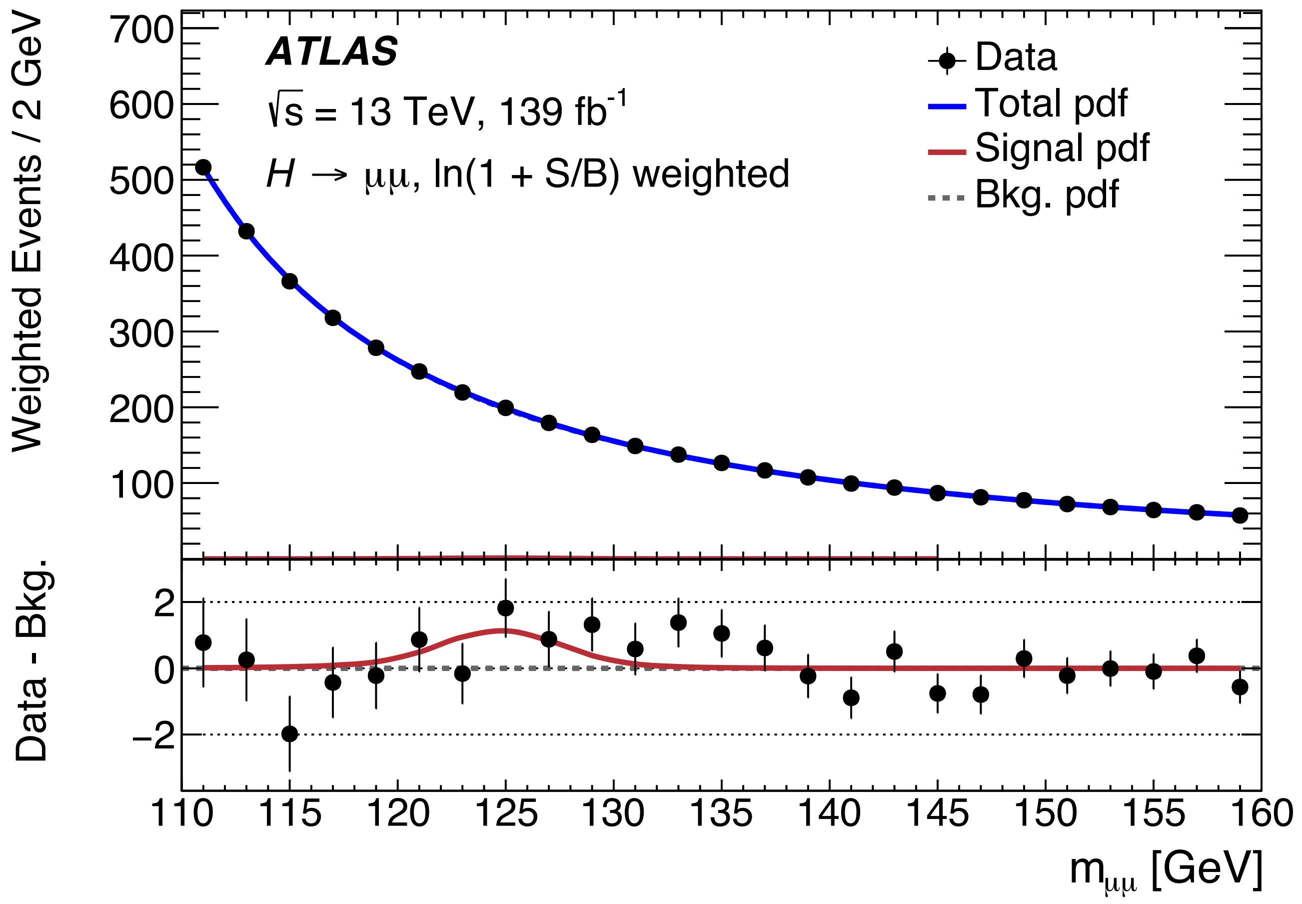
⇒ ***discrete jumps in observed p-value at few percent level.***

New ATLAS result

- Both CMS and ATLAS results include exclusive ttH and VH categories, as well as VBF-targeted region.
- However, many details are sufficiently different to make a direct comparison of these two new results quite complicated.

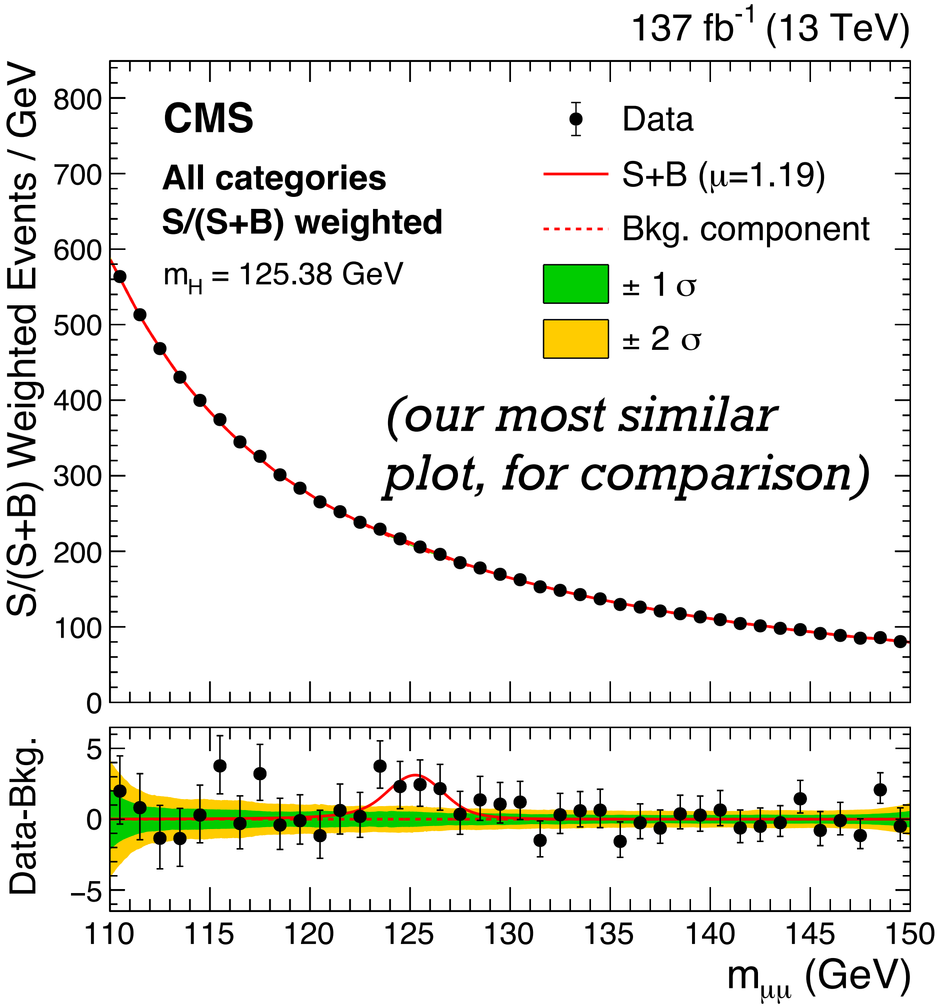
Obs. (exp.) significance: 2.0 (1.7) s.d.

Signal strength μ = 1.2 ± 0.6



Obs. (exp.) significance: 3.0 (2.5) s.d.

Signal strength μ = 1.2 ± 0.4



(our most similar plot, for comparison)

[1] arXiv:2007.07830, submitted to PLB

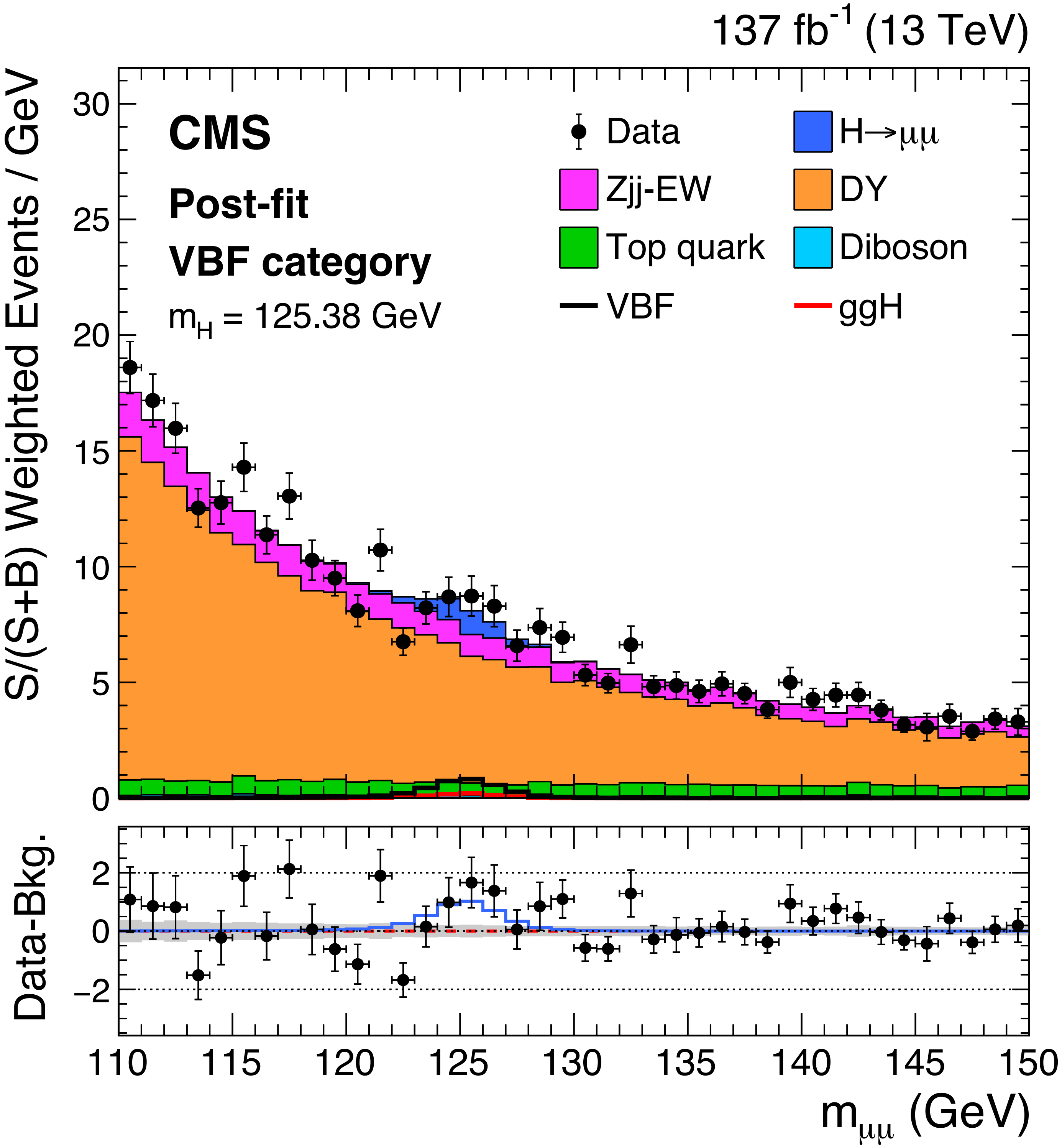
ATLAS H→μμ categories

Category	Data	S_{SM}	S	B	$S/\sqrt{B}$	S/B [%]	σ [GeV]
VBF Very High	15	2.81 ± 0.27	3.3 ± 1.7	14.5 ± 2.1	0.86	22.6	3.0
VBF High	39	3.46 ± 0.36	4.0 ± 2.1	32.5 ± 2.9	0.71	12.4	3.0
VBF Medium	112	4.8 ± 0.5	5.6 ± 2.8	85 ± 4	0.61	6.6	2.9
VBF Low	284	7.5 ± 0.9	9 ± 4	273 ± 8	0.53	3.2	3.0
2-jet Very High	1030	17.6 ± 3.3	21 ± 10	1024 ± 22	0.63	2.0	3.1
2-jet High	5433	50 ± 8	58 ± 30	5440 ± 50	0.77	1.0	2.9
2-jet Medium	18 311	79 ± 15	90 ± 50	$18\,320 \pm 90$	0.66	0.5	2.9
2-jet Low	36 409	63 ± 17	70 ± 40	$36\,340 \pm 140$	0.37	0.2	2.9
1-jet Very High	1097	16.5 ± 2.4	19 ± 10	1071 ± 22	0.59	1.8	2.9
1-jet High	6413	46 ± 7	54 ± 28	6320 ± 50	0.69	0.9	2.8
1-jet Medium	24 576	90 ± 11	100 ± 50	$24\,290 \pm 100$	0.67	0.4	2.7
1-jet Low	73 459	125 ± 17	150 ± 70	$73\,480 \pm 190$	0.53	0.2	2.8
0-jet Very High	15 986	59 ± 11	70 ± 40	$16\,090 \pm 90$	0.55	0.4	2.6
0-jet High	46 523	99 ± 13	120 ± 60	$46\,190 \pm 150$	0.54	0.3	2.6
0-jet Medium	91 392	119 ± 14	140 ± 70	$91\,310 \pm 210$	0.46	0.2	2.7
0-jet Low	121 354	79 ± 10	90 ± 50	$121\,310 \pm 280$	0.26	0.1	2.7
VH4L	34	0.53 ± 0.05	0.6 ± 0.3	24 ± 4	0.13	2.6	2.9
VH3LH	41	1.45 ± 0.14	1.7 ± 0.9	41 ± 5	0.27	4.2	3.1
VH3LM	358	2.76 ± 0.24	3.2 ± 1.6	347 ± 15	0.17	0.9	3.0
$t\bar{t}H$	17	1.19 ± 0.13	1.4 ± 0.7	15.1 ± 2.2	0.36	9.2	3.2

VBF channel S/S+B-weighted mass plot

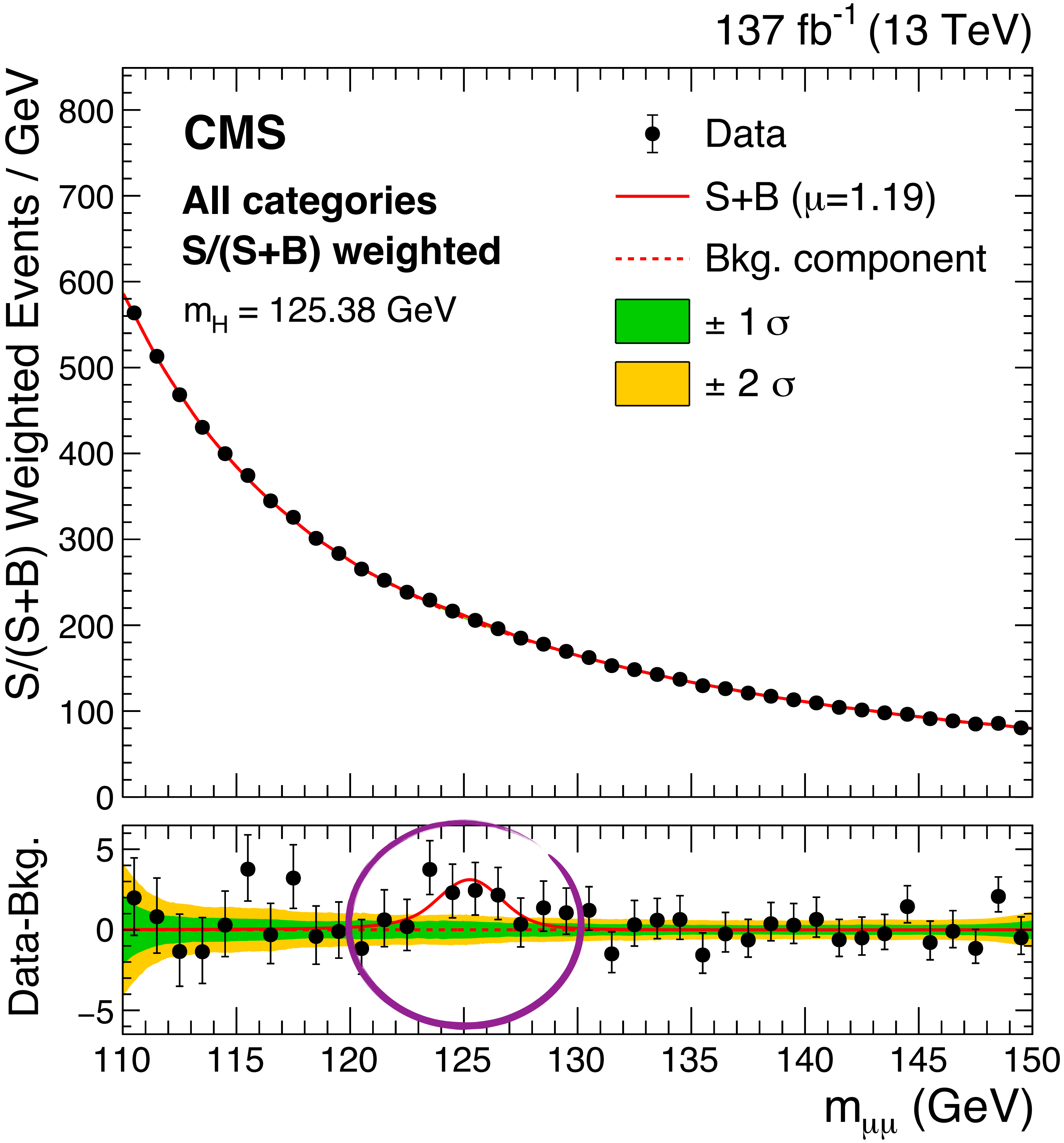
- For VBF channel, assign per-event S/S+B calculated as a function of mass-decorrelated (m_H fixed to 125 GeV) DNN score
  - This binned $m_{\mu\mu}$ histogram is interpolated with a spline and merged with the other channels for the combined mass plot.

- *Note that this is just a way to visualize the observed excess, not the fit result itself.*



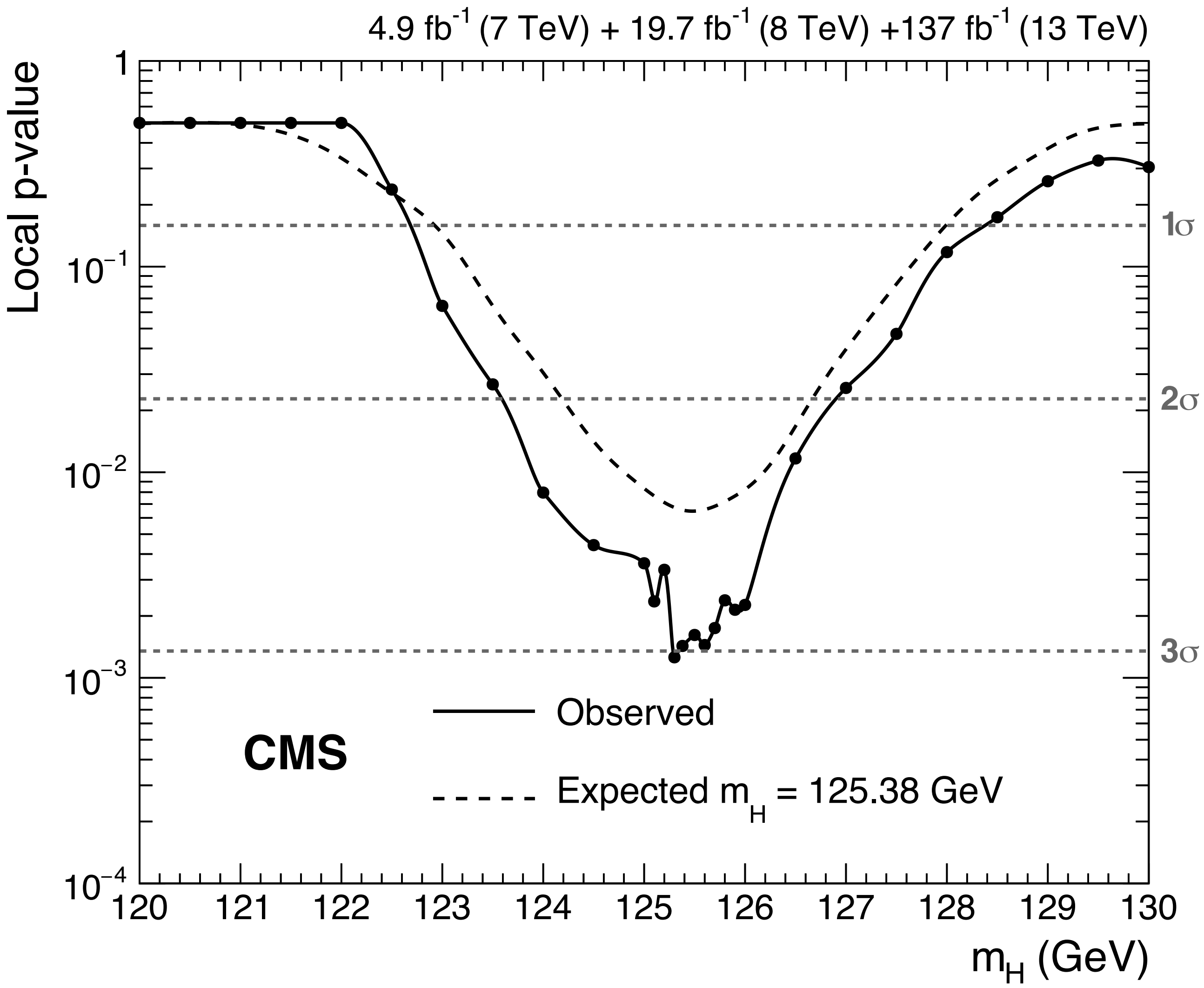
Combined mass distribution

- Weight $m_{\mu\mu}$ distribution in each ggH, VH, and ttH category by S/S+B within signal HWHM.
- Interpolate binned VBF category $m_{\mu\mu}$ histogram (see previous slide) with spline and merge with other categories.
- *Note that this is just a way to visualize the observed excess, not the fit result itself.*



Combination with Run-1

- Combination performed with CMS Run-1 H→μμ search.
    - Full p-value scan vs. m_H.
    - Run-1 adjusted to m_H = 125.38 GeV signal hypothesis.
- **Observed** (expected) **significance 2.98σ** (2.48σ).
- *Local minimum at m_H = 125.3 GeV is 3.02σ.*



Production category	Observed (expected) signif.	Observed (expected) UL on μ
VBF	2.40 (1.77)	2.57 (1.22)
ggH	0.99 (1.56)	1.77 (1.28)
$t\bar{t}$H	1.20 (0.54)	6.48 (4.20)
VH	2.02 (0.42)	10.8 (5.13)
Combined $\sqrt{s} = 13$ TeV	2.95 (2.46)	1.94 (0.82)
Combined $\sqrt{s} = 7, 8, 13$ TeV	2.98 (2.48)	1.93 (0.81)

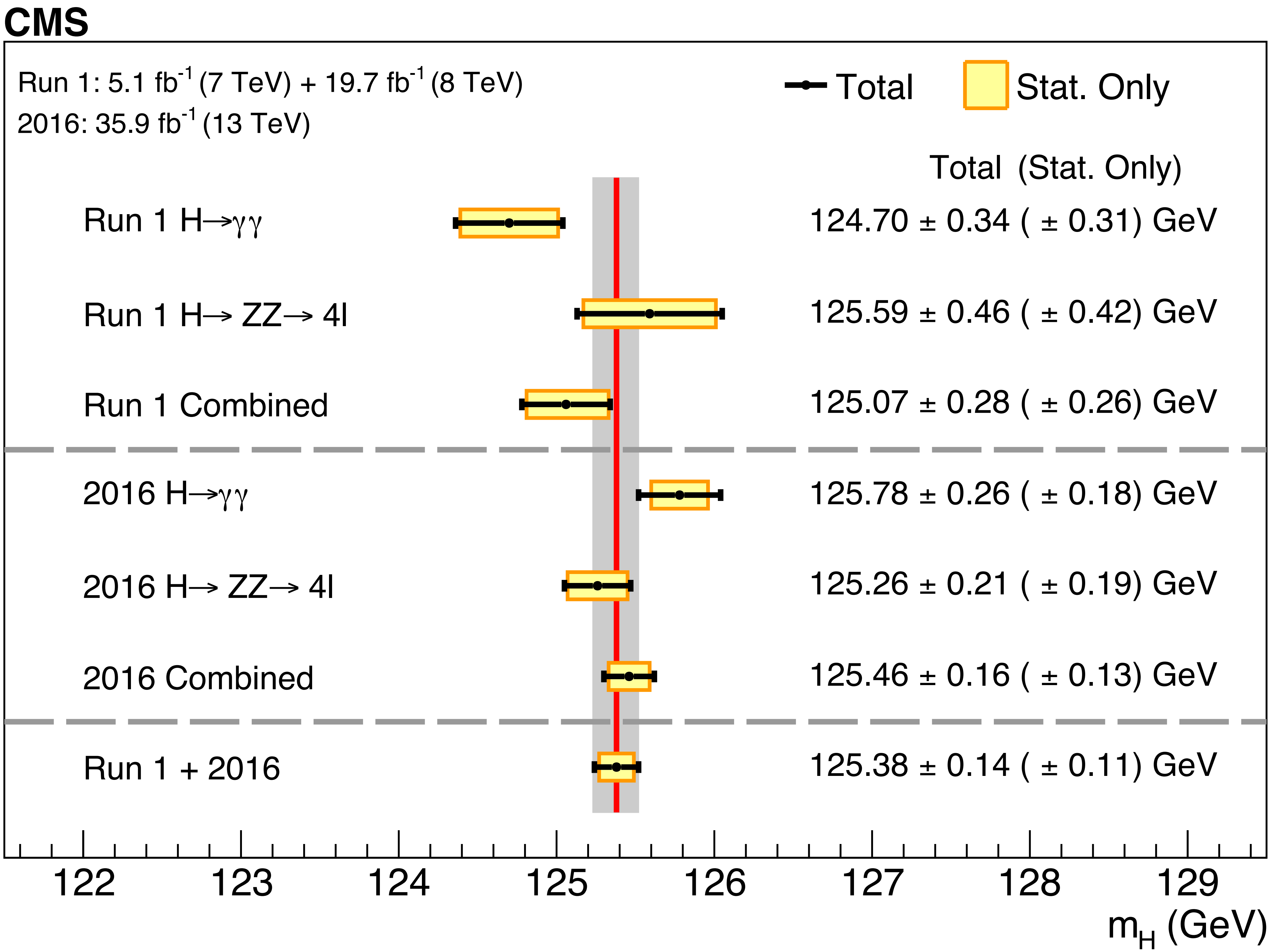
Systematic uncertainties

- Precision in all channels dominated by limited amount of data.
  - Largest impact from systematics in VBF category (<5%).

- Largest systematic uncertainty impacts from limited MC statistics in VBF category and VBF (signal and EW Zjj) parton shower modeling.

Uncertainty source	$\Delta\mu$	
Post-fit uncertainty	$+0.44$	-0.42
Statistical uncertainty	$+0.41$	-0.40
Systematic uncertainty	$+0.17$	-0.16
Experimental uncertainty	$+0.12$	-0.11
Theoretical uncertainty	$+0.10$	-0.11
Size of simulated samples	$+0.07$	-0.06

Summary of CMS Higgs boson mass measurements



Phys. Lett. B 805 (2020) 135425

Background modeling: Core-PDF method

- Background composition is quite stable across categories and dominated by DY (>90%).
- Core background function built as discrete profile of two physics-inspired (Breit-Wigner, FEWZ) and an agnostic (sum of exponentials) function.
- Bias studied against multiple physics-inspired and agnostic background functions and always < 20% (negligible impact on result).
- No prior assumption on background shape or normalization

$$\boxed{B_{cat}(m_{\mu\mu}, \overrightarrow{\alpha}, \overrightarrow{\beta}) = N_B \times F_{core}(m_{\mu\mu}, \overrightarrow{\alpha}) \times T_{SMF}(m_{\mu\mu}, \overrightarrow{\beta})}$$

Number of bkg. events

Core-PDF, defined as a **discrete profile** of three functions. Associated **parameters** are **correlated across categories**

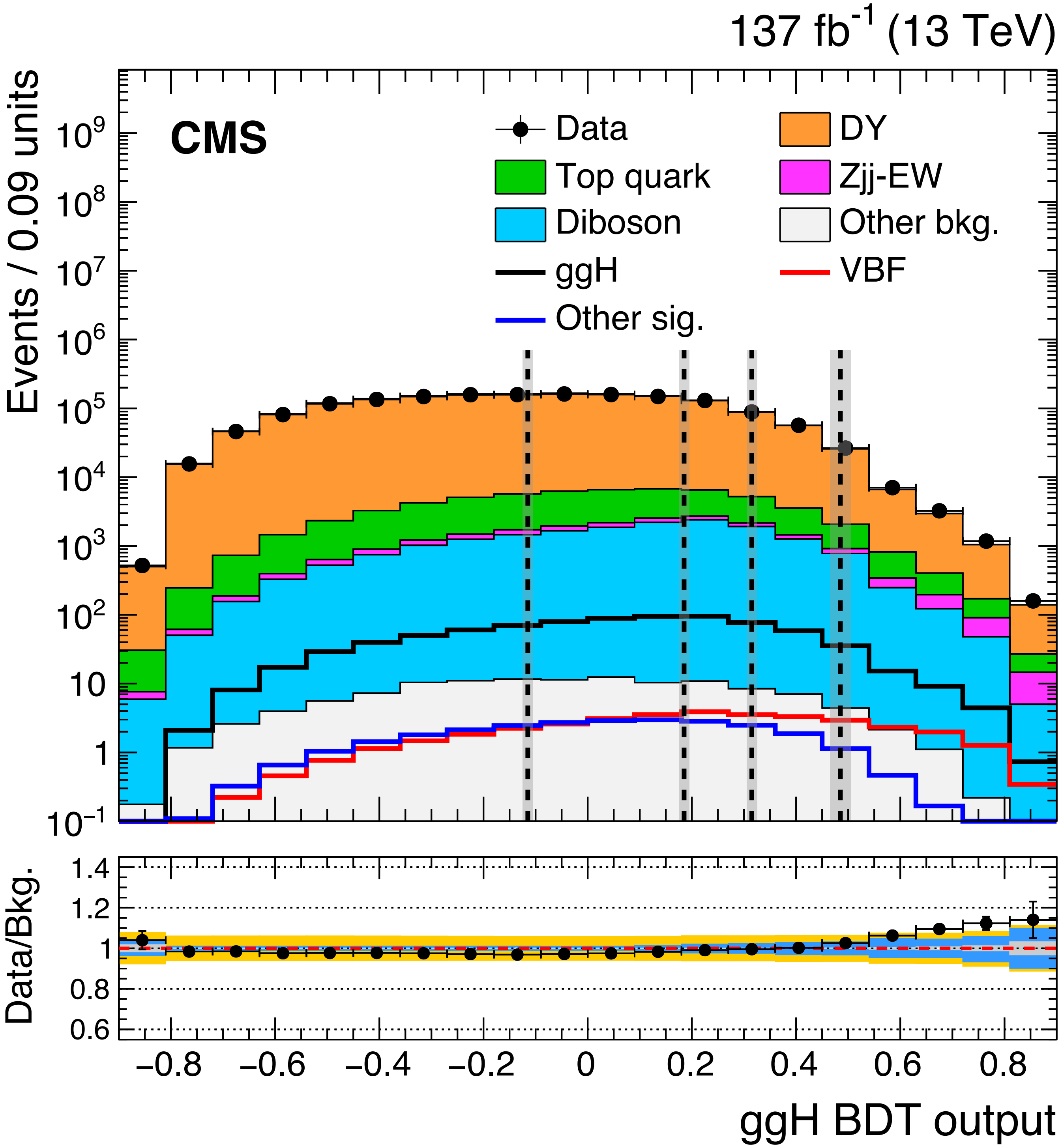
Per-category **polynomial shape modifier** use to “morph” the core component

*method in ggH channel only

ggH category details

- Consider all events not selected by exclusive categories.
    - About 96% of total inclusive events selected in ggH category.
- Largest signal yield, but smallest S/B ~ 0.1%.
- Train mass-decorrelated BDT based on muon kinematics + possible jet kinematics (to pick up also ggH+X, residual VBF, hadronic VH)

BDT subcategory boundaries optimized iteratively based on significance from full fit to MC.



Dimuon system
- **$p_T(\mu\mu)$**
- $y(\mu\mu)$
- **Colin-Soper angles**

Leading jet
- $p_T(j_1)$
- **$\eta(j_1)$**
- **$\Delta\eta(\mu\mu,j_1)$**
- $\Delta\phi(\mu\mu,j_2)$

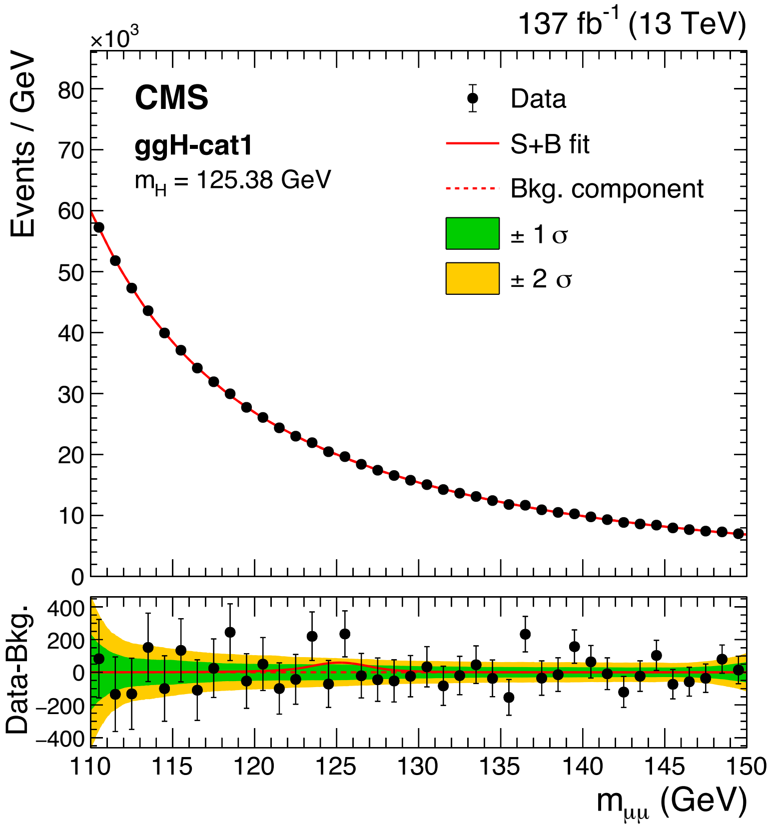
Dijet system
- **m(jj)**
- $\Delta\eta(jj)$
- **$\Delta\phi(jj)$**
- Zeppenfeld
- $p_T(j_2)$
- **min-$\Delta\eta(\mu\mu,j)$**
- **min-$\Delta\phi(\mu\mu,j)$**

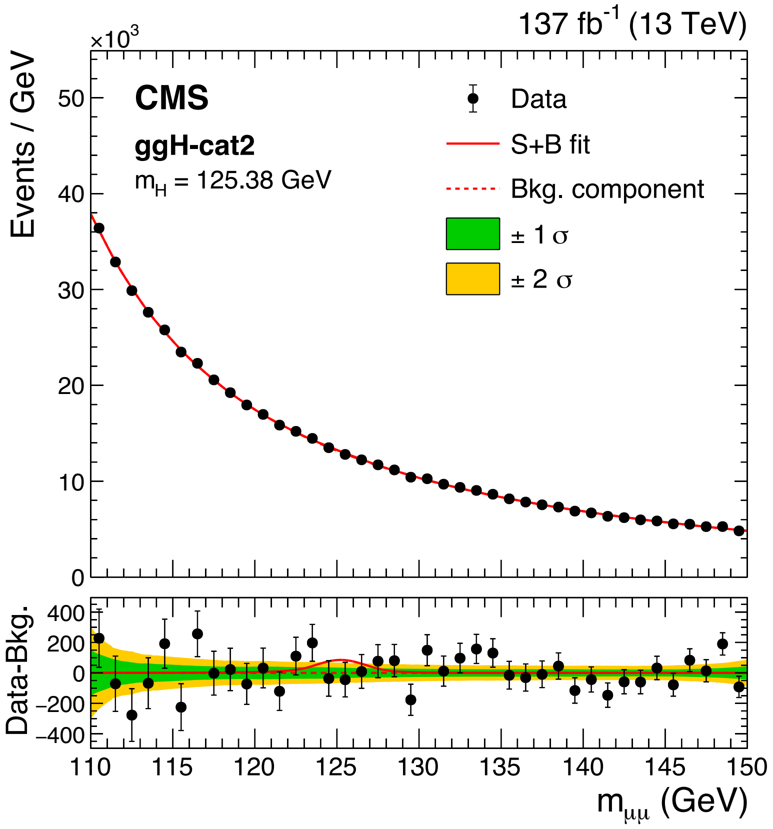
Single muon
- $p_T(\mu)/m(\mu\mu)$
- $\eta(\mu)$

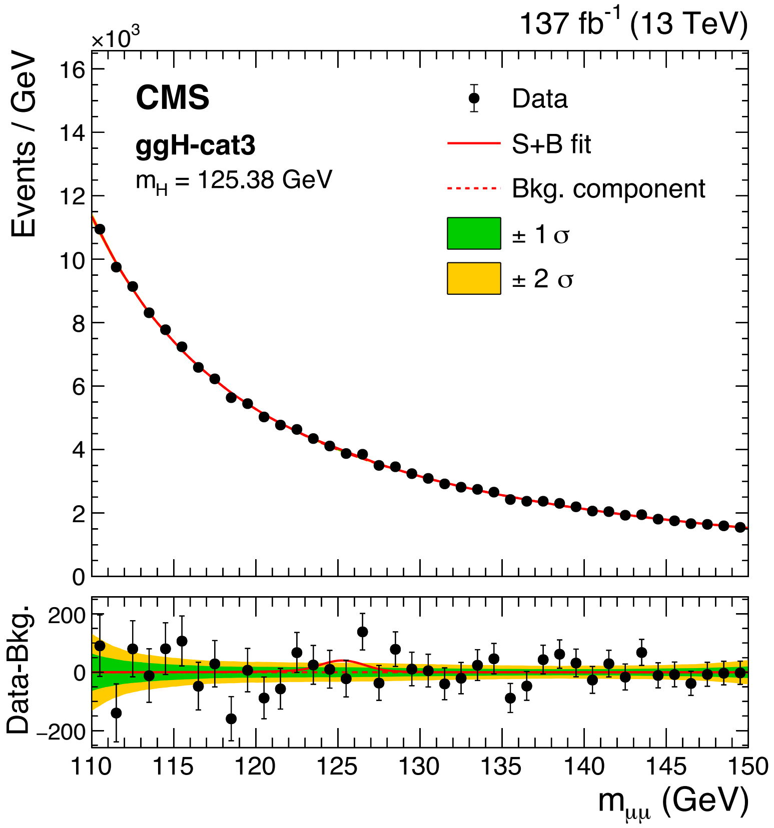
Event variables
- **N_{jets}**

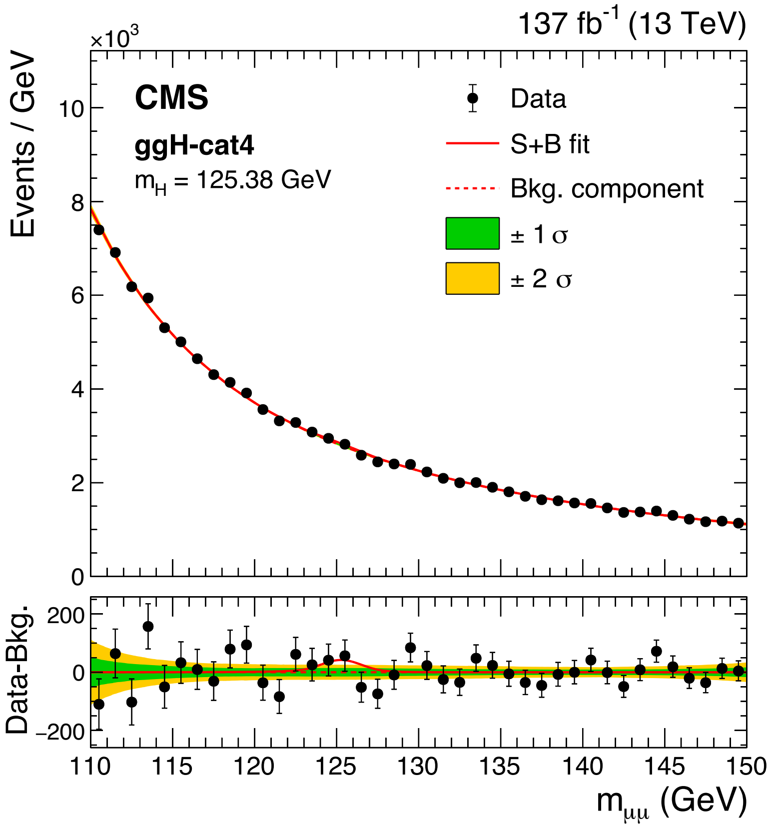
ggH channel: fitted distributions

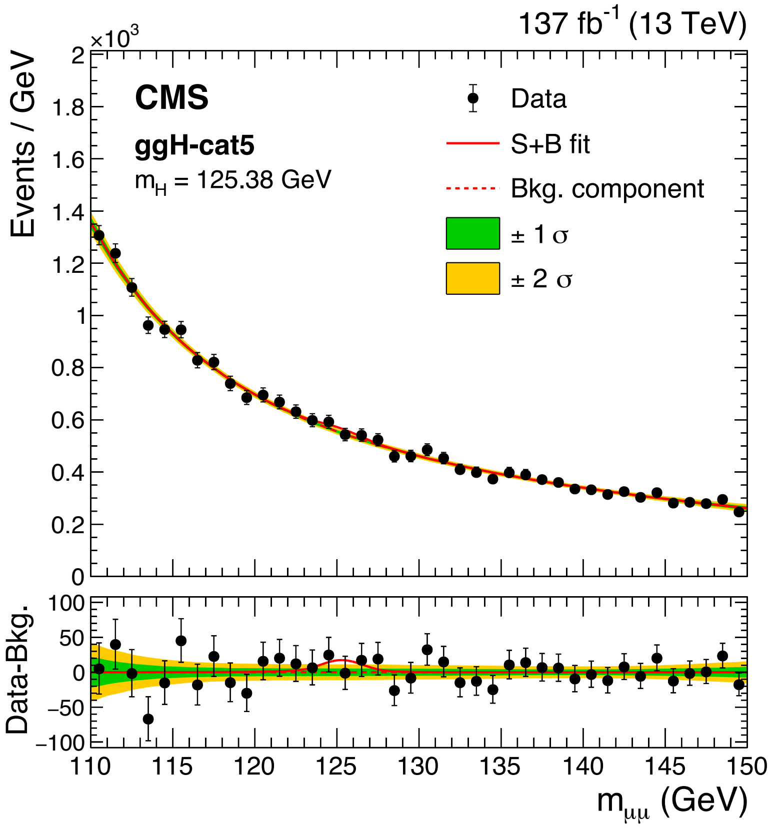
- S+B background fit describes the data well throughout the $m_{\mu\mu}$ spectrum in all categories.





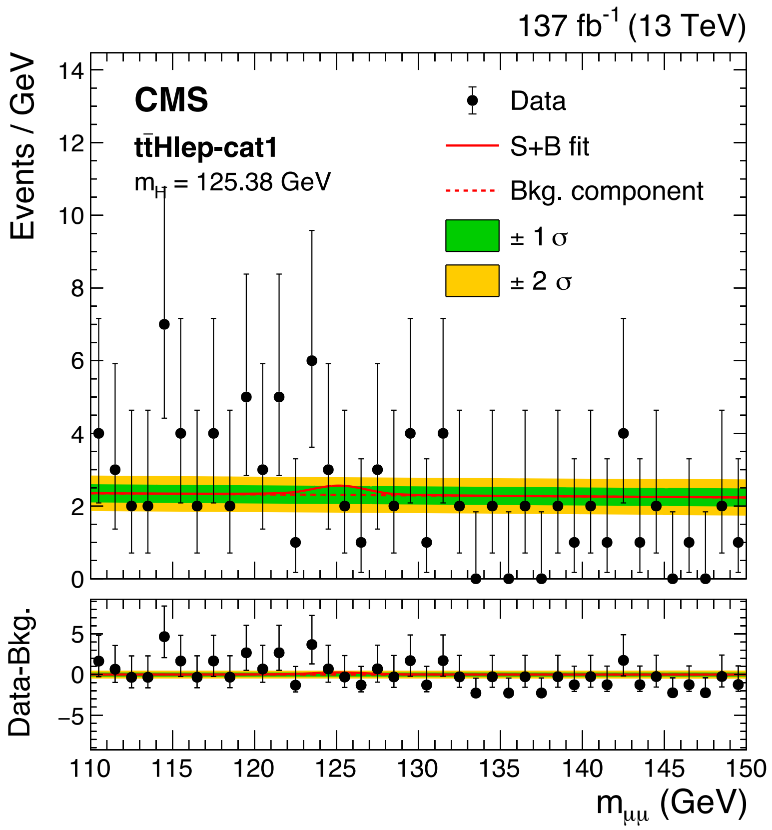


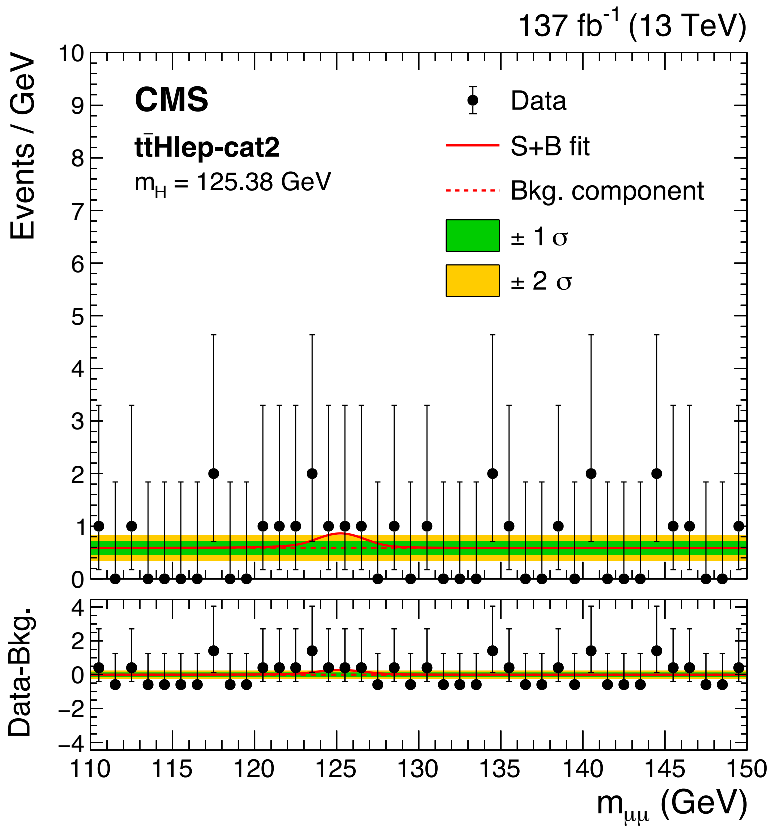


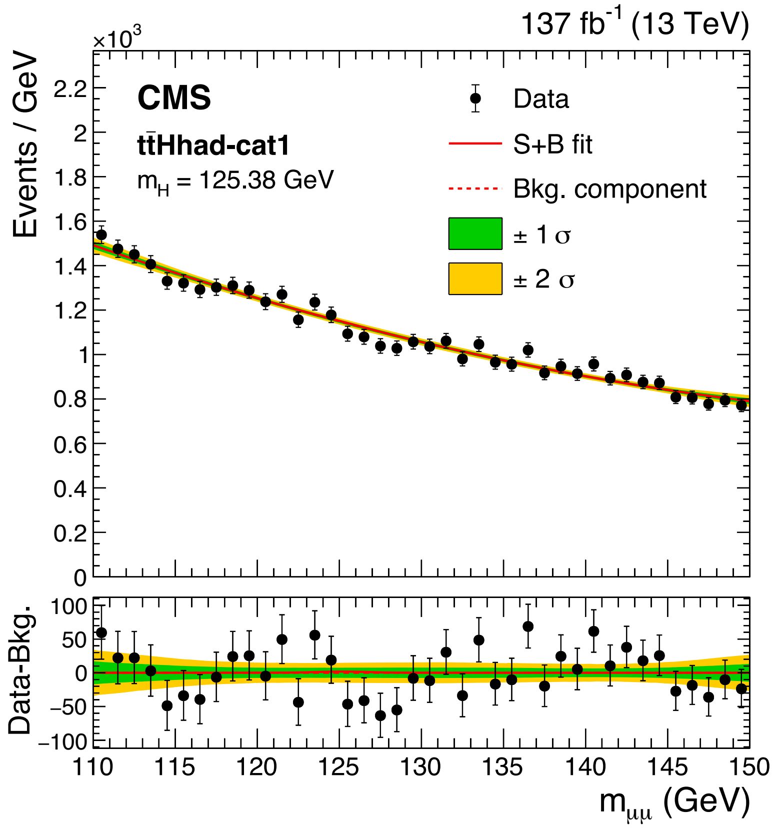


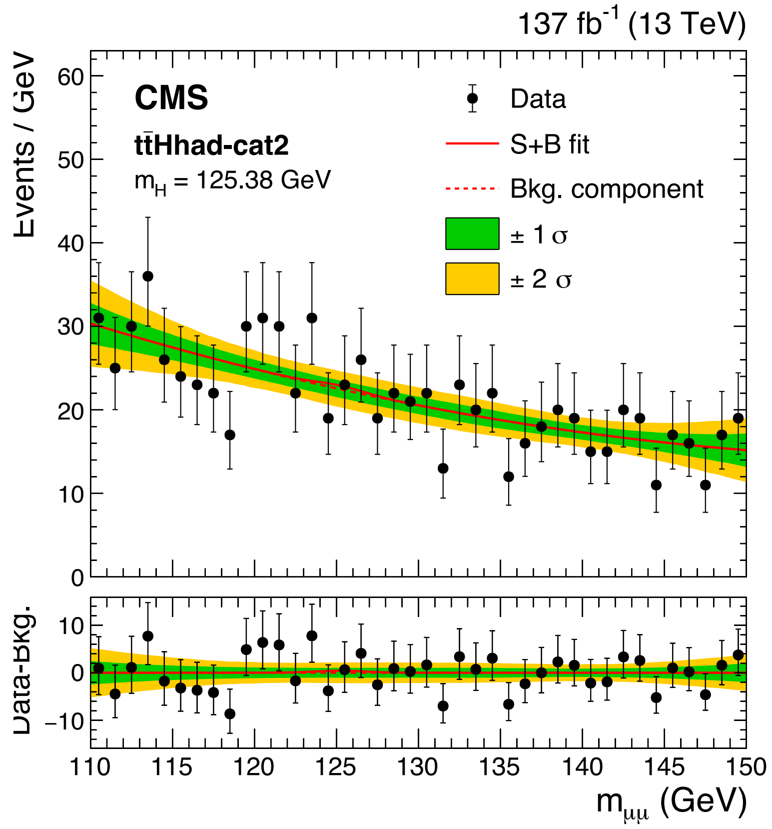
ttH channel: fitted distributions

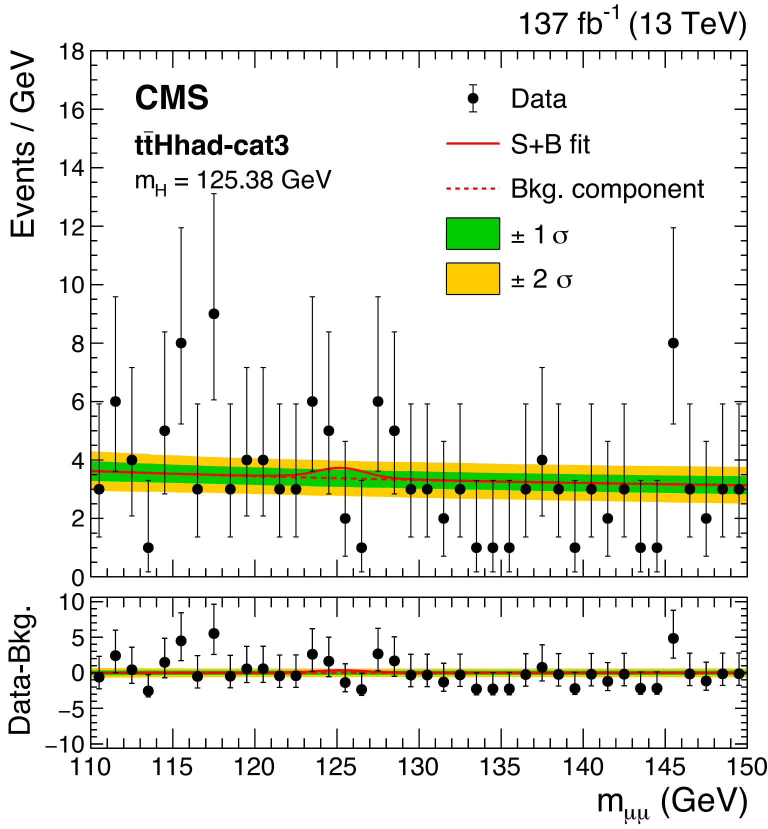
- Background fit with simple exponential (ttH-lep) or polynomial (ttH-had).
  - Bias checked following similar procedure as in ggH channel.
- Categories optimized following same strategy as ggH channel.







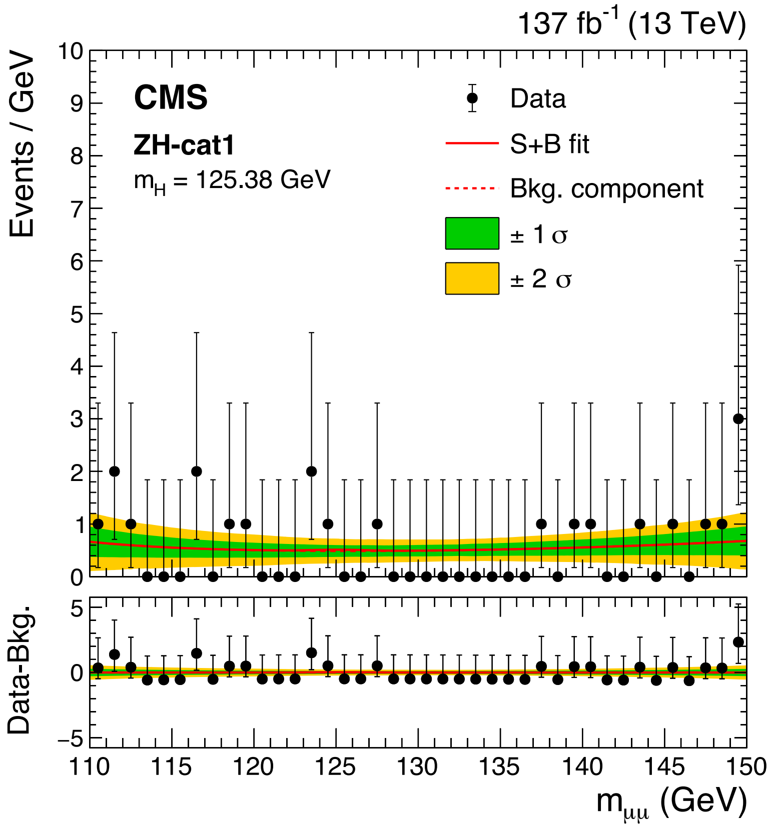


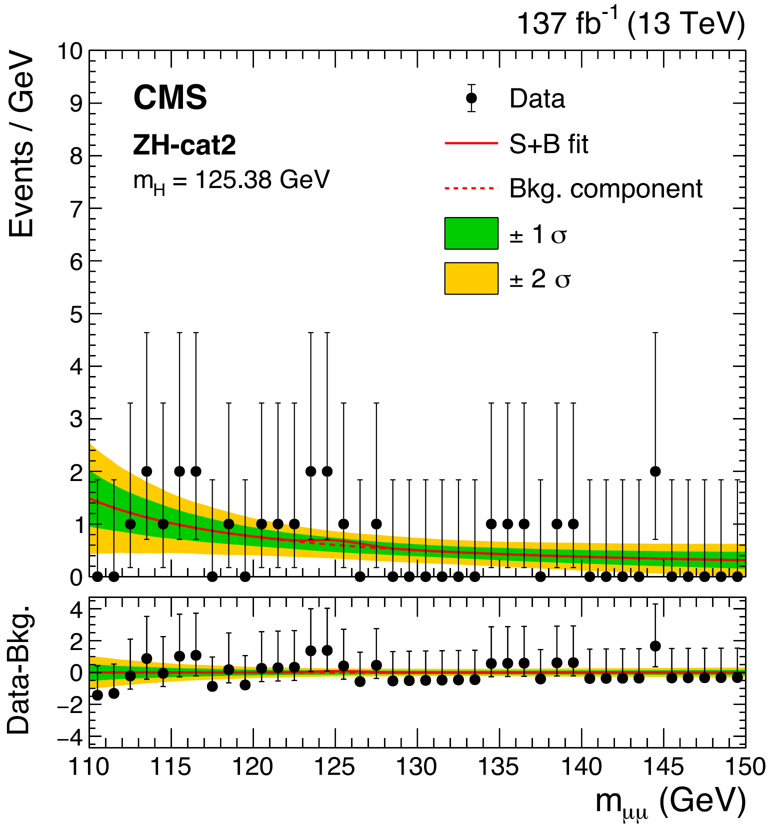


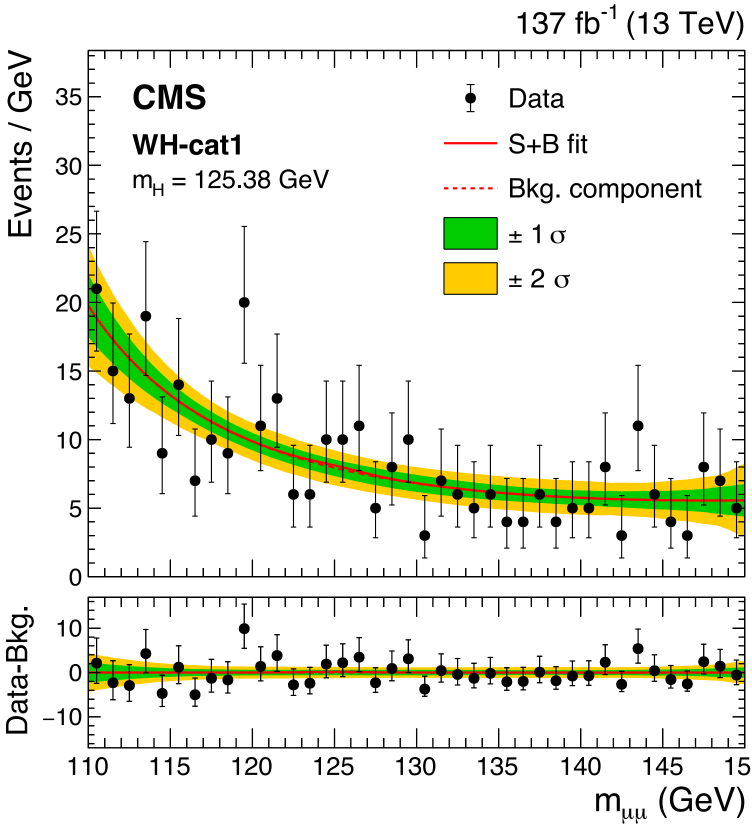
VH channel: fitted distributions

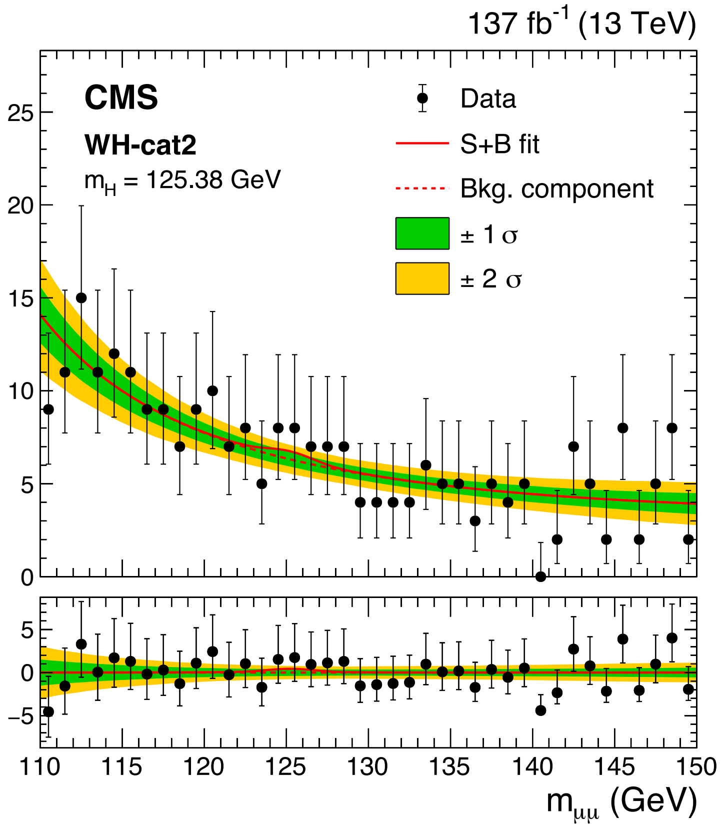
- Background fit with BWZ function.
  - Choice of functions and bias studies similar to other channels.

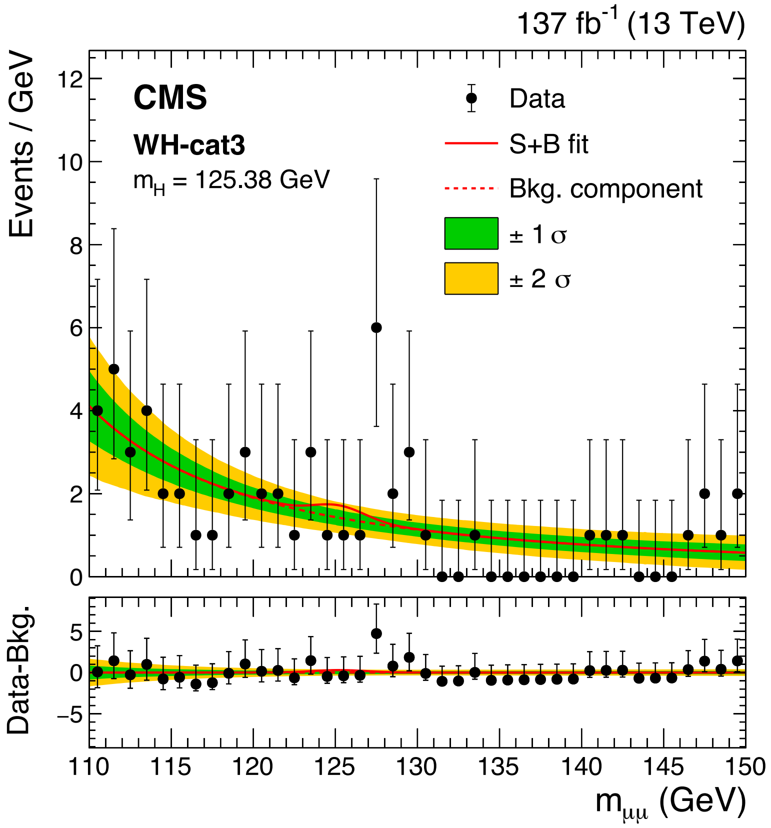
- Small excesses in data near 125 GeV, but consistent with expectation within (large) statistical uncertainties.



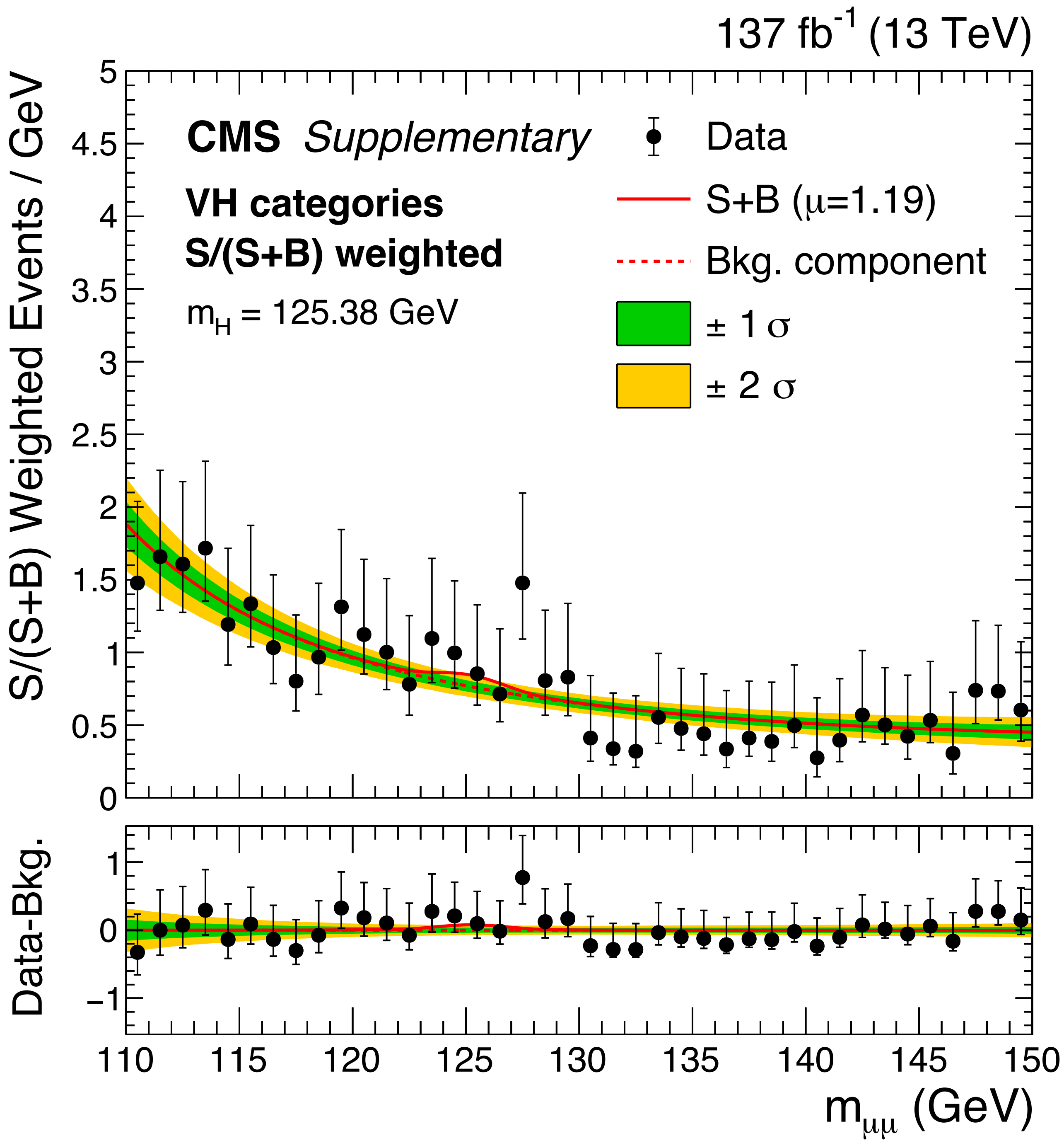


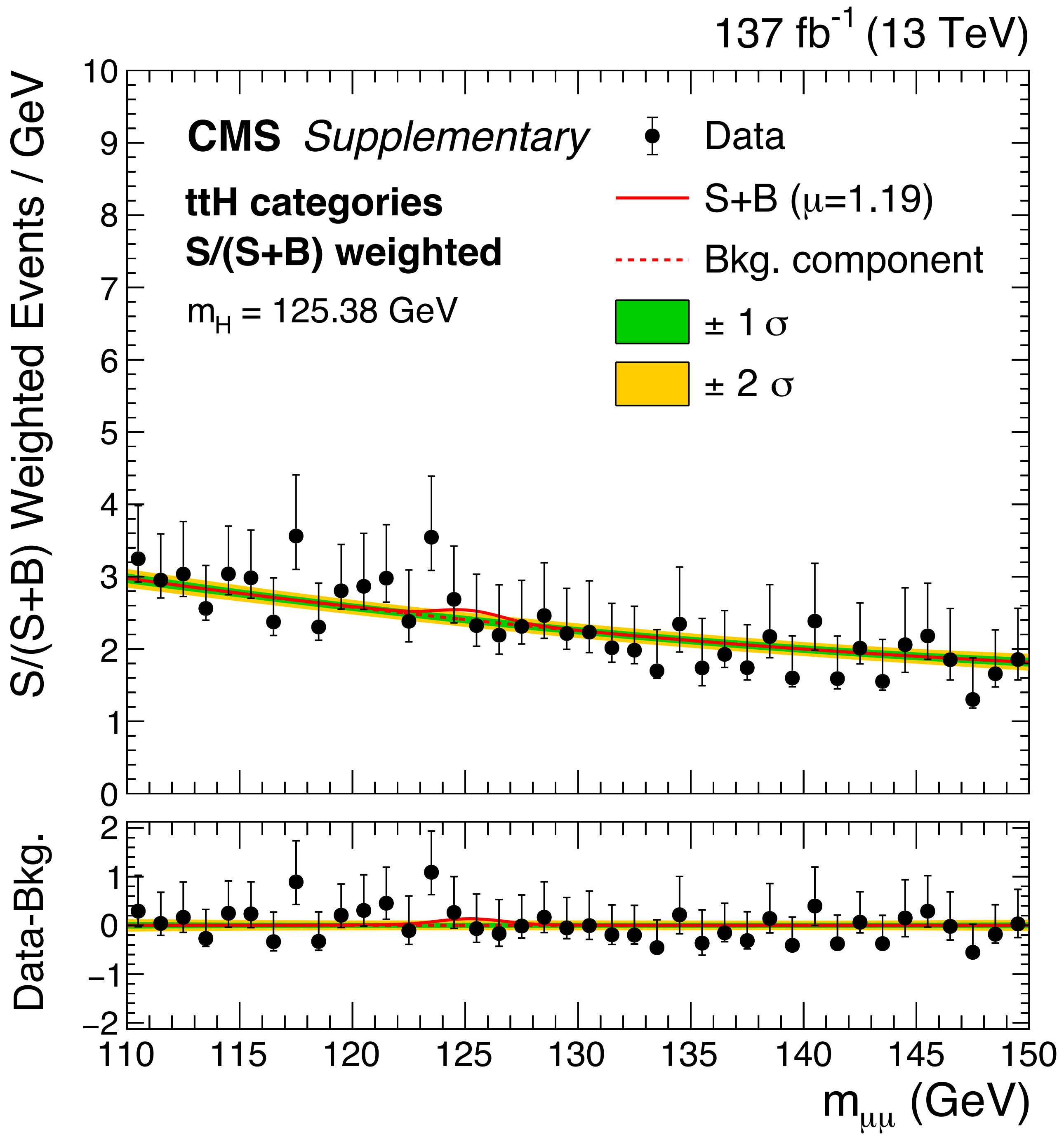




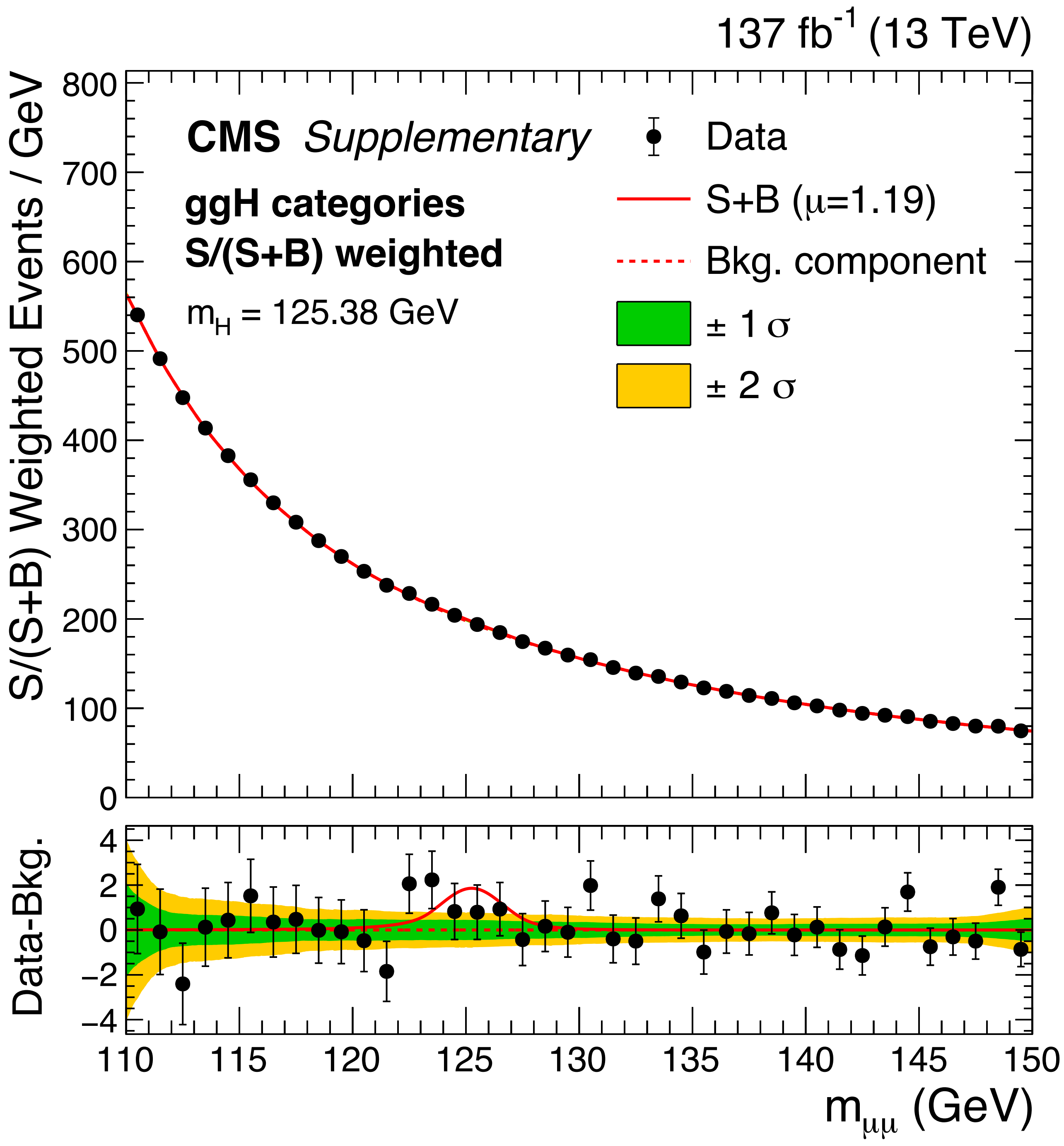


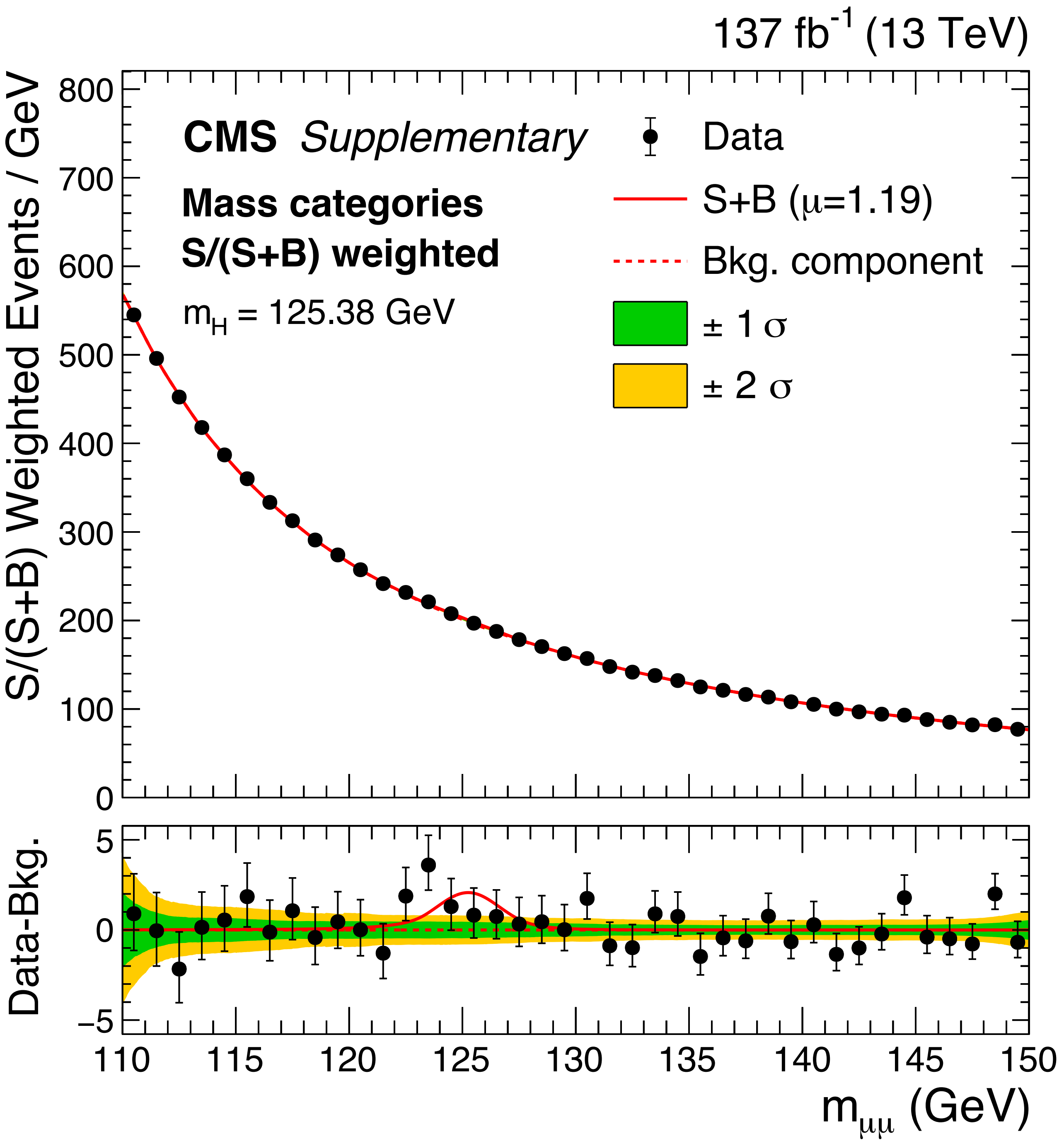
S/S+B-weighted plots per category





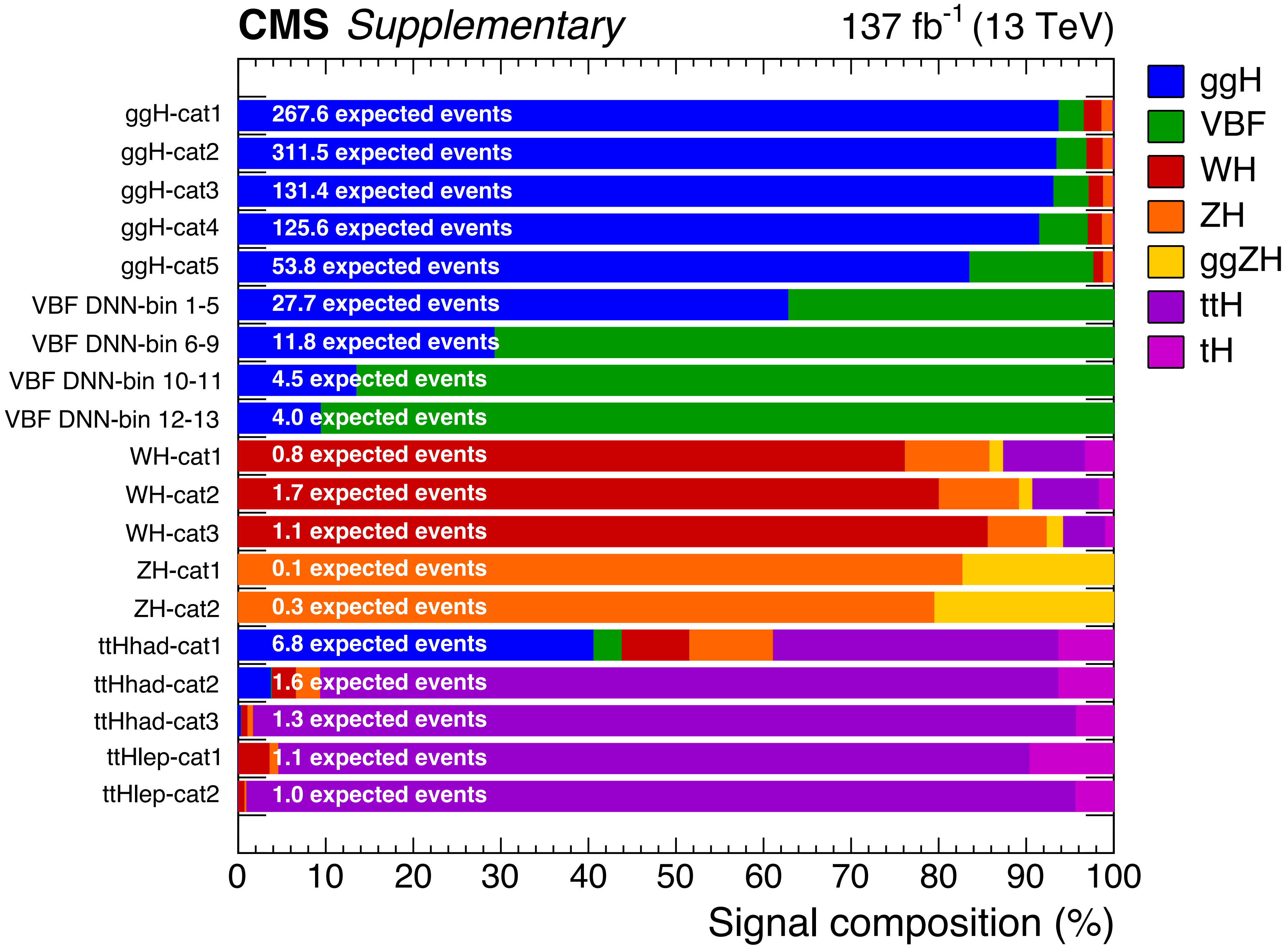
S/S+B-weighted plots per category





Signal composition across categories

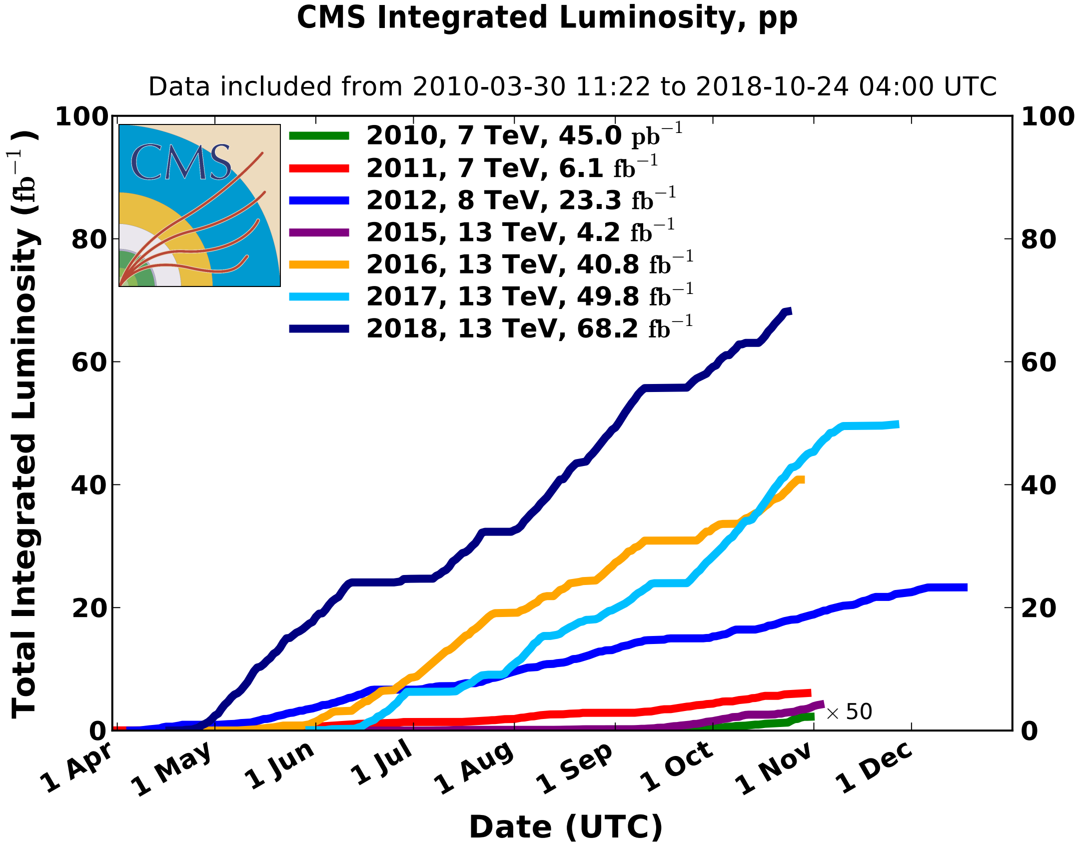
- Relatively high purity in targeted Higgs boson production mode achieved in each category.

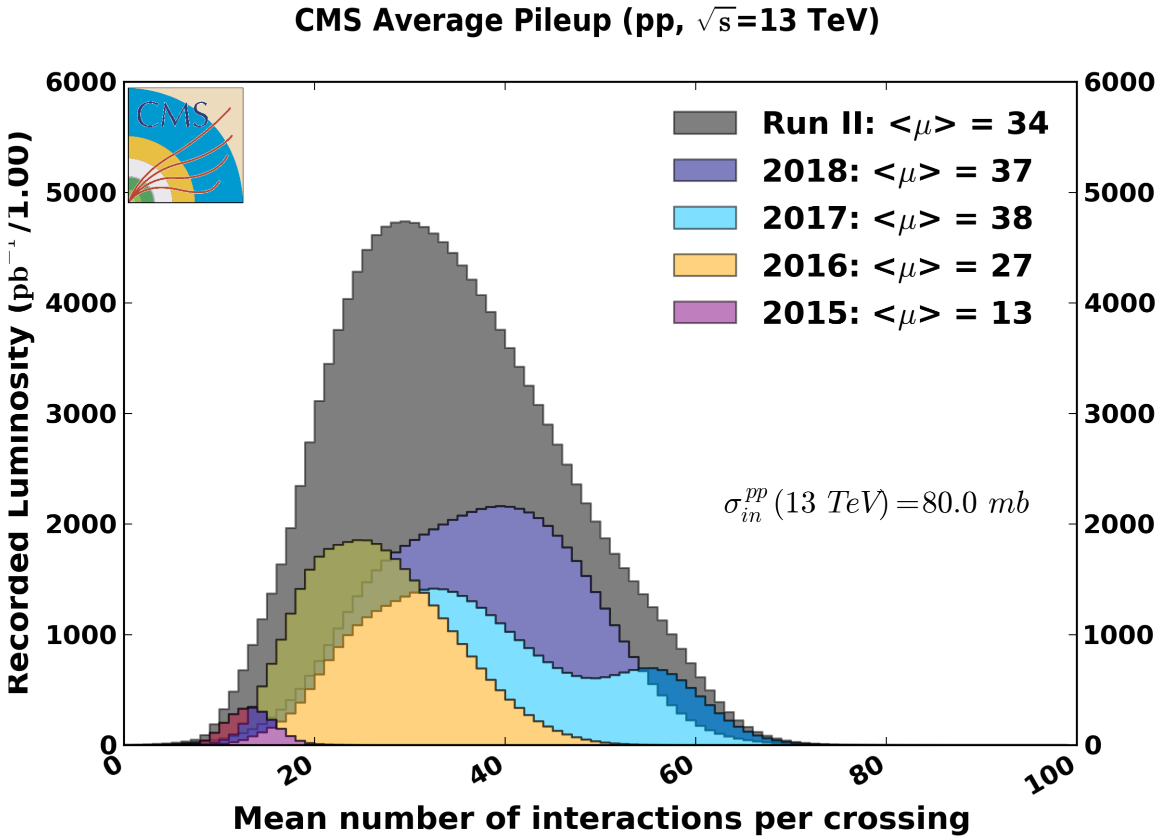


The data set

- *This result focuses on 137 fb^{-1} of 13 TeV data collected by CMS from 2016 to 2018.*
    - Including new analysis of recalibrated 2016 data, which had been used for previous CMS H→μμ search.

- With excellent LHC performance comes the challenge of high rates and many simultaneous collisions (pileup).





The CMS detector

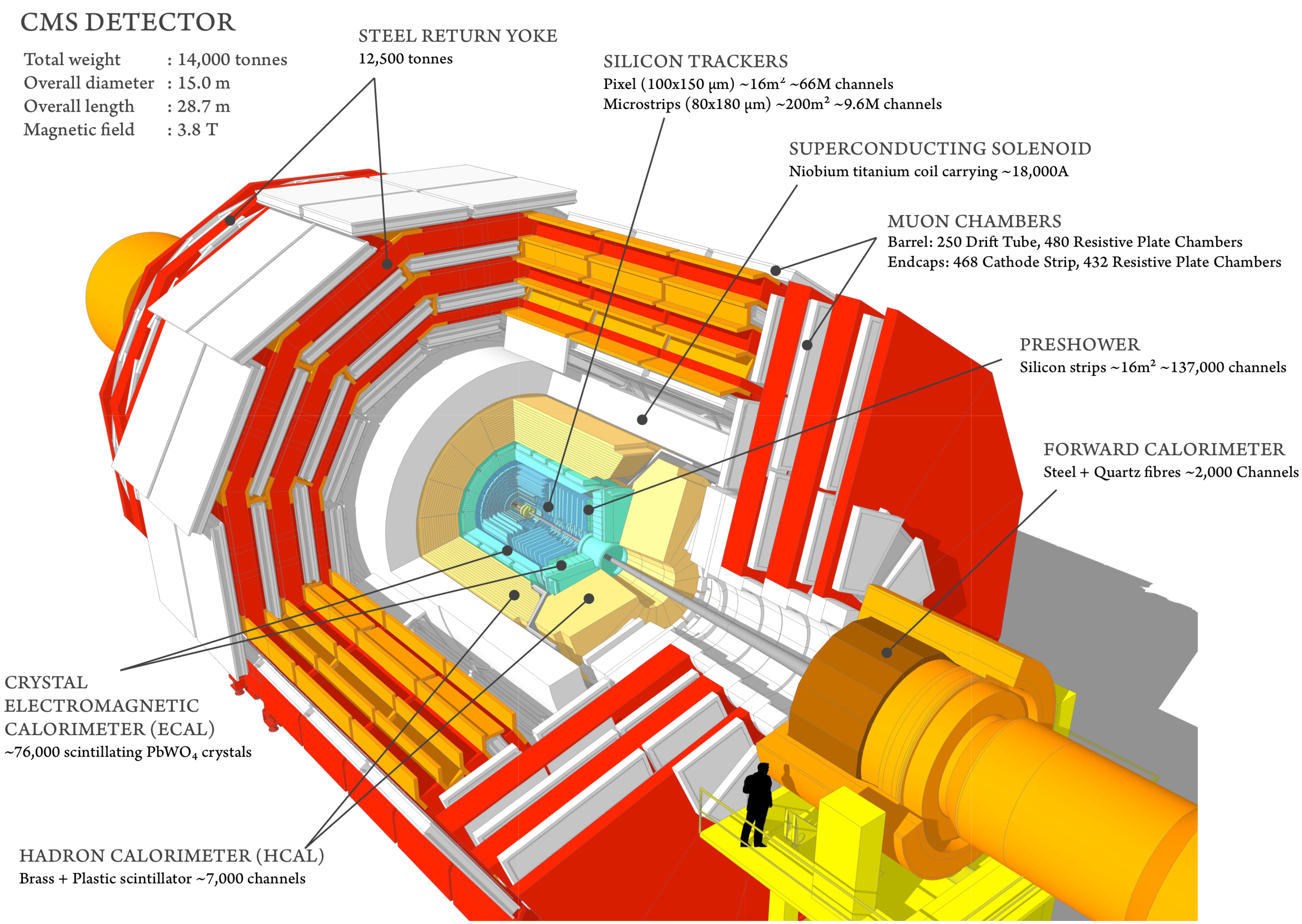
- 3.8T **superconducting solenoidal magnet** with 6m diameter.

- **Tracker System**: silicon strip+ pixel system which reconstructs the trajectories of charged particles.

- **Electromagnetic calorimeter** (ECAL): scintillator made from lead tungstate crystals sensitive to energy deposits from electrons and photons.

- **Hadronic calorimeter** (HCAL): brass scintillator sensitive to energy deposits from hadrons, mainly pions and kaons.

- Gas ionization chambers for **muon detection**.



VBF category details

Observable	VBF-SB	VBF-SR
Number of loose (medium) b-tagged jets	$\leq$1 (0)	
Number of selected muons	=2	
Number of selected electrons	=0	
Jet multiplicity ($p_T > 25\,\text{GeV}$, $\lvert\eta\rvert < 4.7$)	$\geq$2	
Leading jet p_T	$\geq$35 GeV	
Dijet mass (m_{jj})	$\geq$400 GeV	
Pseudorapidity separation ($\lvert\Delta\eta_{jj}\rvert$)	$\geq$2.5	
Dimuon invariant mass	$110 < m_{\mu\mu} < 115\,\text{GeV}$ or $135 < m_{\mu\mu} < 150\,\text{GeV}$	$115 < m_{\mu\mu} < 135\,\text{GeV}$

DNN bin	Total signal	VBF (%)	ggH (%)	Bkg. $\pm$ ΔB	Data	S/(S+B) (%)	S/$\sqrt{\text{B}}$
1–3	19.5	30	70	8890 $\pm$ 67	8815	0.22	0.21
4–6	11.6	57	43	394 $\pm$ 8	388	2.86	0.58
7–9	8.43	73	27	103 $\pm$ 4	121	7.56	0.83
10	2.30	85	15	15.1 $\pm$ 1.4	18	13.2	0.59
11	2.15	88	12	9.1 $\pm$ 1.2	10	19.1	0.71
12	2.10	87	13	5.8 $\pm$ 1.1	6	26.6	0.87
13	1.87	94	6	2.6 $\pm$ 0.9	7	41.8	1.16

ggH category details

Observable	Selection
Number of loose (medium) b-tagged jets	$\leq$1 (0)
Number of selected muons	=2
Number of selected electrons	=0
VBF selection veto	if $N_{\text{jets}} \geq 2$ $m_{\text{jj}} < 400\,\text{GeV}$ or $\lvert\Delta\eta_{\text{jj}}\rvert < 2.5$ or $p_{\text{T}}(\text{j}_1) < 35\,\text{GeV}$

Event category	Total signal	ggH (%)	VBF (%)	Other (%)	HWHM (GeV)	Bkg. @HWHM	Data @HWHM	S/(S+B) (%) @HWHM	$S/\sqrt{B}$ @HWHM
ggH-cat1	268	93.7	2.9	3.4	2.12	86 360	86 632	0.20	0.60
ggH-cat2	312	93.5	3.4	3.1	1.75	46 350	46 393	0.46	0.98
ggH-cat3	131	93.2	4.0	2.8	1.60	12 660	12 738	0.70	0.80
ggH-cat4	126	91.5	5.5	3.0	1.47	8260	8377	1.03	0.96
ggH-cat5	53.8	83.5	14.3	2.2	1.50	1680	1711	2.16	0.91

ttH category details

Observable	$t\bar{t}$H hadronic	$t\bar{t}$H leptonic
Number of b quark jets	$>$0 medium or $>$1 loose b-tagged jets	
Number of leptons ($N(\ell = \mu, e)$)	=2	=3 or 4
Lepton charge ($q(\ell)$)	$\sum q(\ell) = 0$	$N(\ell) = 3\ (4) \rightarrow \sum q(\ell) = \pm 1\ (0)$
Jet multiplicity ($p_T > 25\,\mathrm{GeV}$, $\lvert\eta\rvert < 4.7$)	$\geq$3	$\geq$2
Leading jet p_T	$>$50 GeV	$>$35 GeV
Z boson veto	—	$\lvert m_{\ell\ell} - m_Z\rvert > 10\,\mathrm{GeV}$
Low-mass resonance veto	—	$m_{\ell\ell} > 12\,\mathrm{GeV}$
Jet triplet mass	$100 < m_{jjj} < 300\,\mathrm{GeV}$	—

Event category	Total signal	$t\bar{t}$H (%)	ggH (%)	VH (%)	Other (%)	HWHM (GeV)	Bkg. fit function	Bkg. @HWHM	Data @HWHM	S/(S+B) (%) @HWHM	$S/\sqrt{B}$ @HWHM
$t\bar{t}$Hhad-cat1	6.87	32.3	40.3	17.2	10.2	1.85	Bern(2)	4298	4251	1.07	0.07
$t\bar{t}$Hhad-cat2	1.62	84.3	3.8	5.6	6.2	1.81	Bern(2)	82.0	89	1.32	0.12
$t\bar{t}$Hhad-cat3	1.33	94.0	0.3	1.3	4.4	1.80	S-Exp	12.3	12	6.87	0.26
$t\bar{t}$Hlep-cat1	1.06	85.8	—	4.7	9.5	1.92	Exp	9.00	13	7.09	0.22
$t\bar{t}$Hlep-cat2	0.99	94.7	—	1.0	4.3	1.75	Exp	2.08	4	24.5	0.47

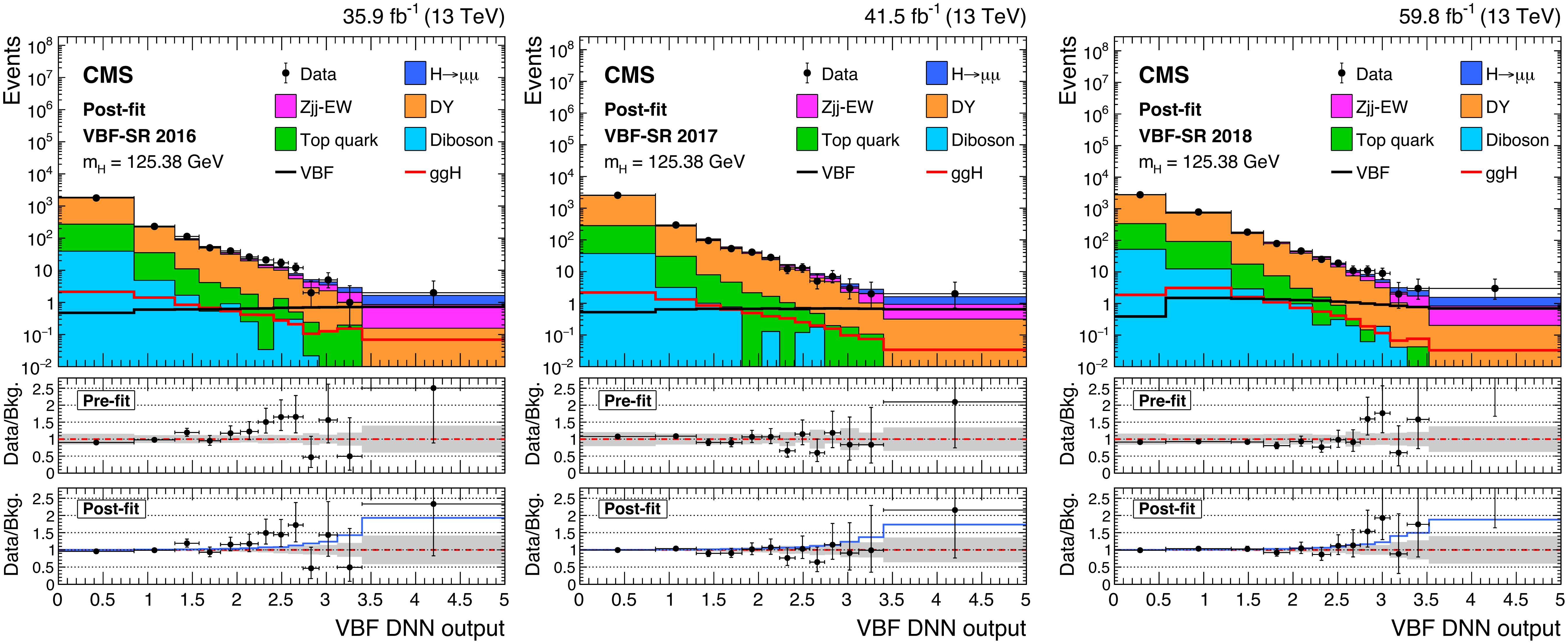
VH category details

Observable	WH leptonic		ZH leptonic	
	$\mu\mu\mu$	$\mu\mu e$	4μ	$2\mu 2e$
Number of loose (medium) b-tagged jets	$\leq$1 (0)	$\leq$1 (0)	$\leq$1 (0)	$\leq$1 (0)
Number of selected muons	=3	=2	=4	=2
Number of selected electrons	=0	=1	=0	=2
Lepton charge ($q(\ell)$)	$\sum q(\ell) = \pm 1$		$\sum q(\ell) = 0$	
Low-mass resonance veto	$m_{\ell\ell} > 12\,\text{GeV}$			
N($\mu^+\mu^-$) pairs with $110 < m_{\mu\mu} < 150\,\text{GeV}$	$\geq$1	=1	$\geq$1	=1
N($\mu^+\mu^-$) pairs with $\lvert m_{\mu\mu} - m_Z \rvert < 10\,\text{GeV}$	=0	=0	=1	=0
N(e^+e^-) pairs with $\lvert m_{ee} - m_Z \rvert < 20\,\text{GeV}$	=0	=0	=1	=1

Event category	Total signal	WH (%)	qqZH (%)	ggZH (%)	$t\bar{t}$H+tH (%)	HWHM (GeV)	Bkg. fit function	Bkg. @HWHM	Data @HWHM	S/(S+B) (%) @HWHM	$S/\sqrt{B}$ @HWHM
WH-cat1	0.82	76.2	9.6	1.6	12.6	2.00	BWZγ	32.0	34	1.54	0.09
WH-cat2	1.72	80.1	9.1	1.5	9.3	1.80	BWZ	23.1	27	4.50	0.23
WH-cat3	1.14	85.7	6.7	1.8	4.8	1.90	BWZ	5.48	4	12.6	0.35
ZH-cat1	0.11	—	82.8	17.2	—	2.07	BWZ	2.05	4	3.29	0.05
ZH-cat2	0.31	—	79.6	20.4	—	1.80	BWZ	2.19	4	8.98	0.14

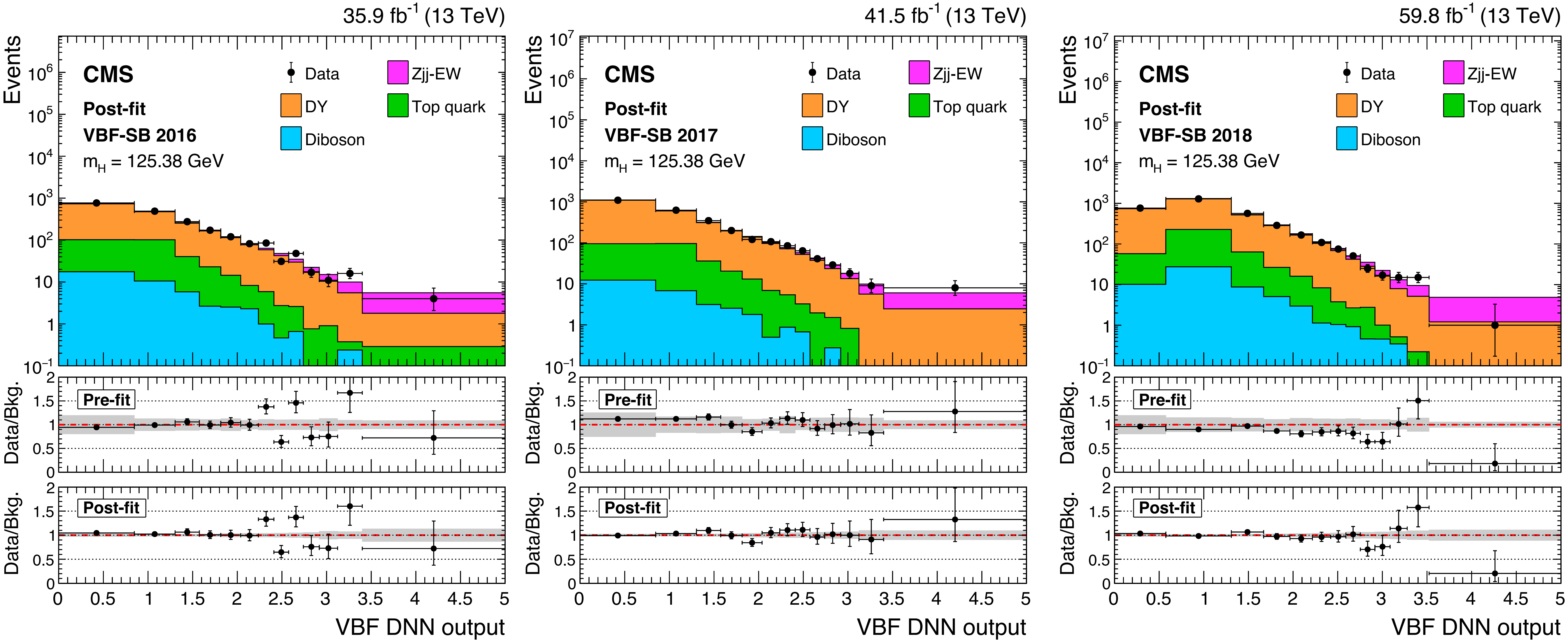
VBF DNN distributions per year

VBF signal region,
115 < m(μμ) 135 GeV



VBF DNN distributions per year

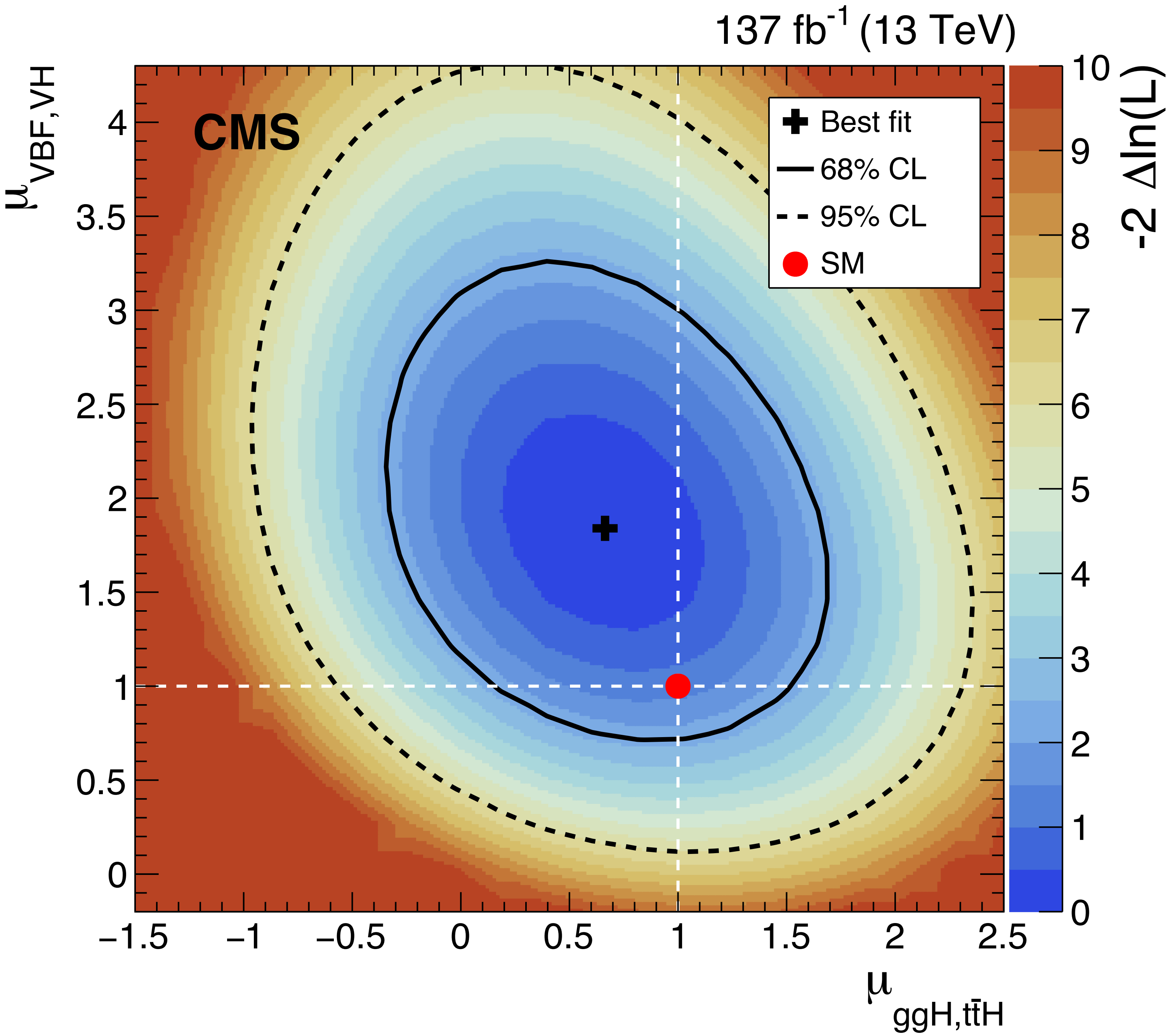
VBF mass sideband,
110 < m(μμ) < 115 GeV,
135 < m(μμ) < 150 GeV



VBF+VH vs. ggH+ttH signal strength

- Independent fit for Higgs boson production channels sensitive to vector boson couplings (VBF, VH) and top quark coupling (ggH, ttH).

- Result well consistent with SM within uncertainties.

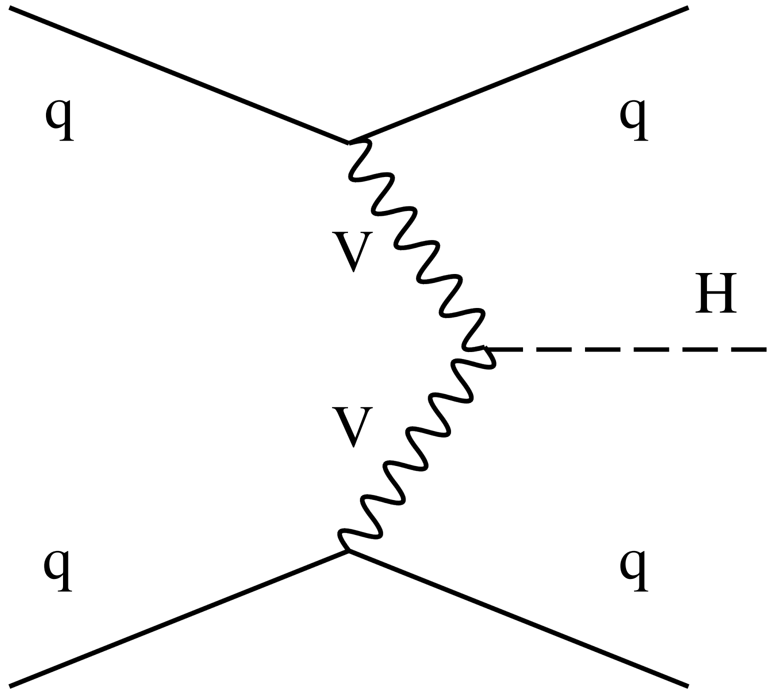


VBF H and EW Zjj modeling

- VBF process incoming and outgoing quark lines are color-connected (pure EW exchange).

- Until recently, PYTHIA did not account for this effect (default “global recoil” scheme).
    - “Dipole recoil” scheme recently implemented into PYTHIA that takes this effect into account.
    - Herwig7 default angular-ordered shower considers effect similar to PYTHIA with dipole recoil.

- Private VBF H→μμ signal samples generated with PYTHIA dipole recoil (nominal prediction) and HERWIG7 (to assess systematic uncertainty).

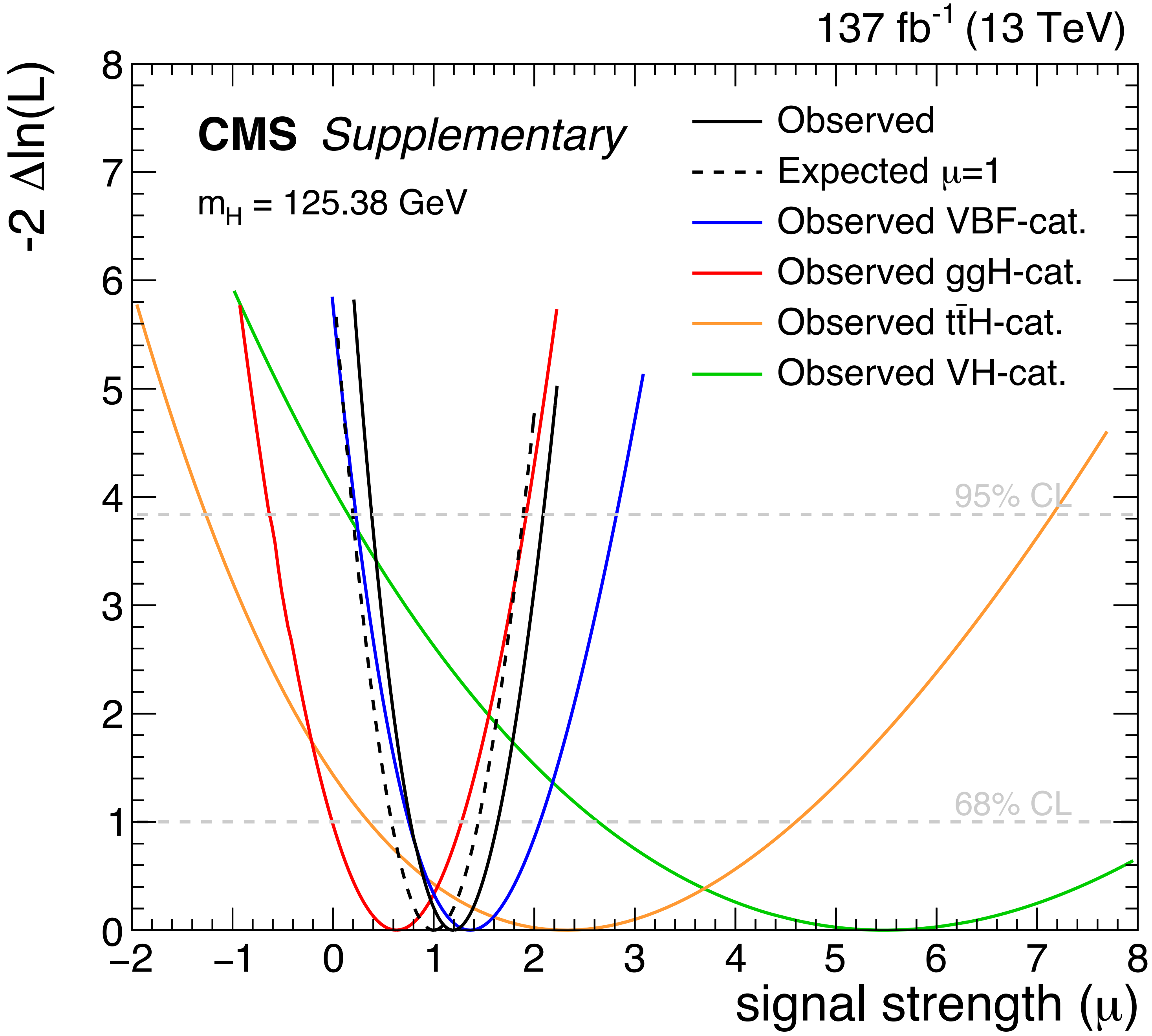


*discussed within VBF HXSWG and recent dedicated theory paper arXiv:2003.12435

Likelihood scans

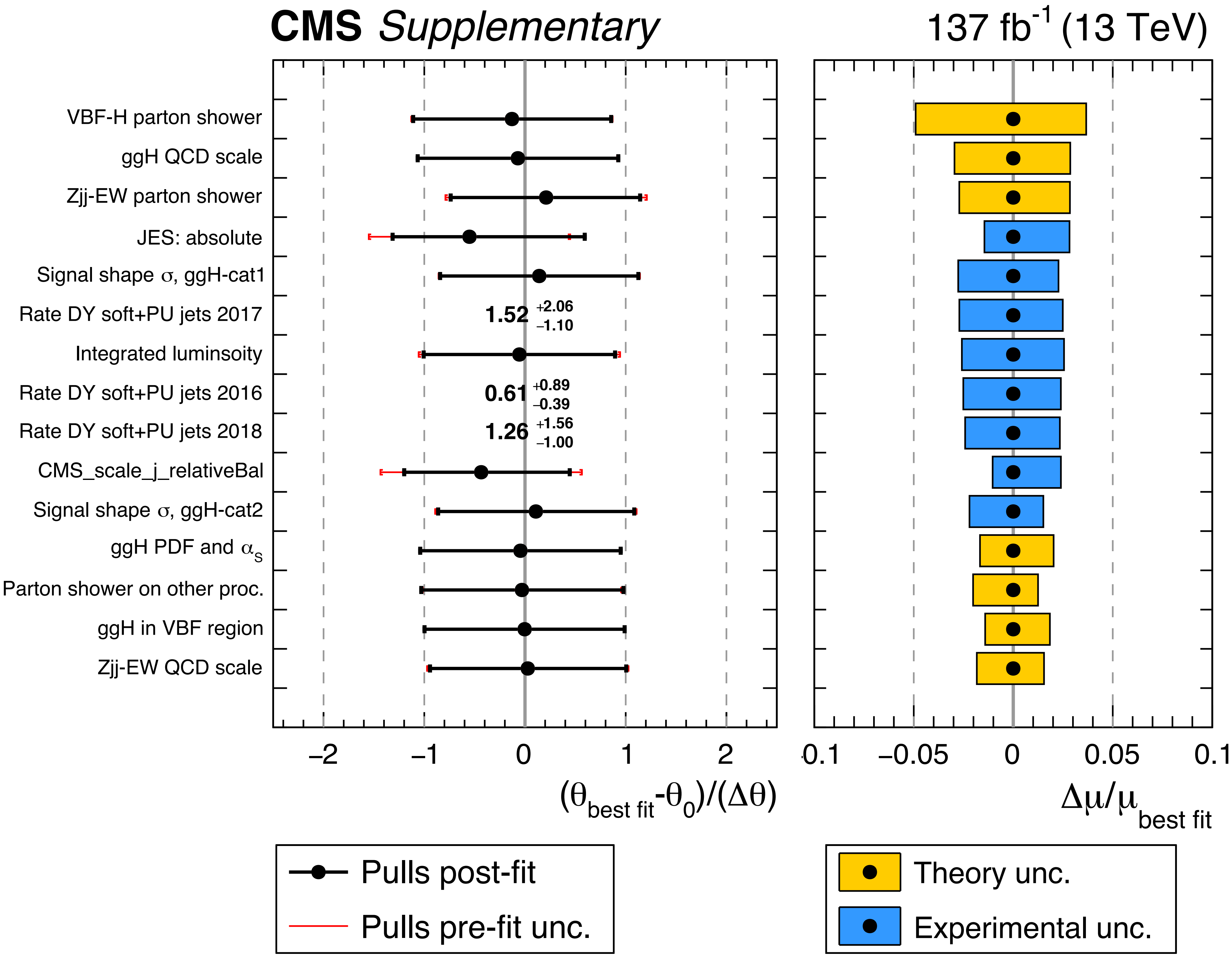
- Observed log-likelihood ratios as a function of signal strength for the combined result as well as the individual channels.

- For the combination, also the expected likelihood shape with μ = 1 signal injected.

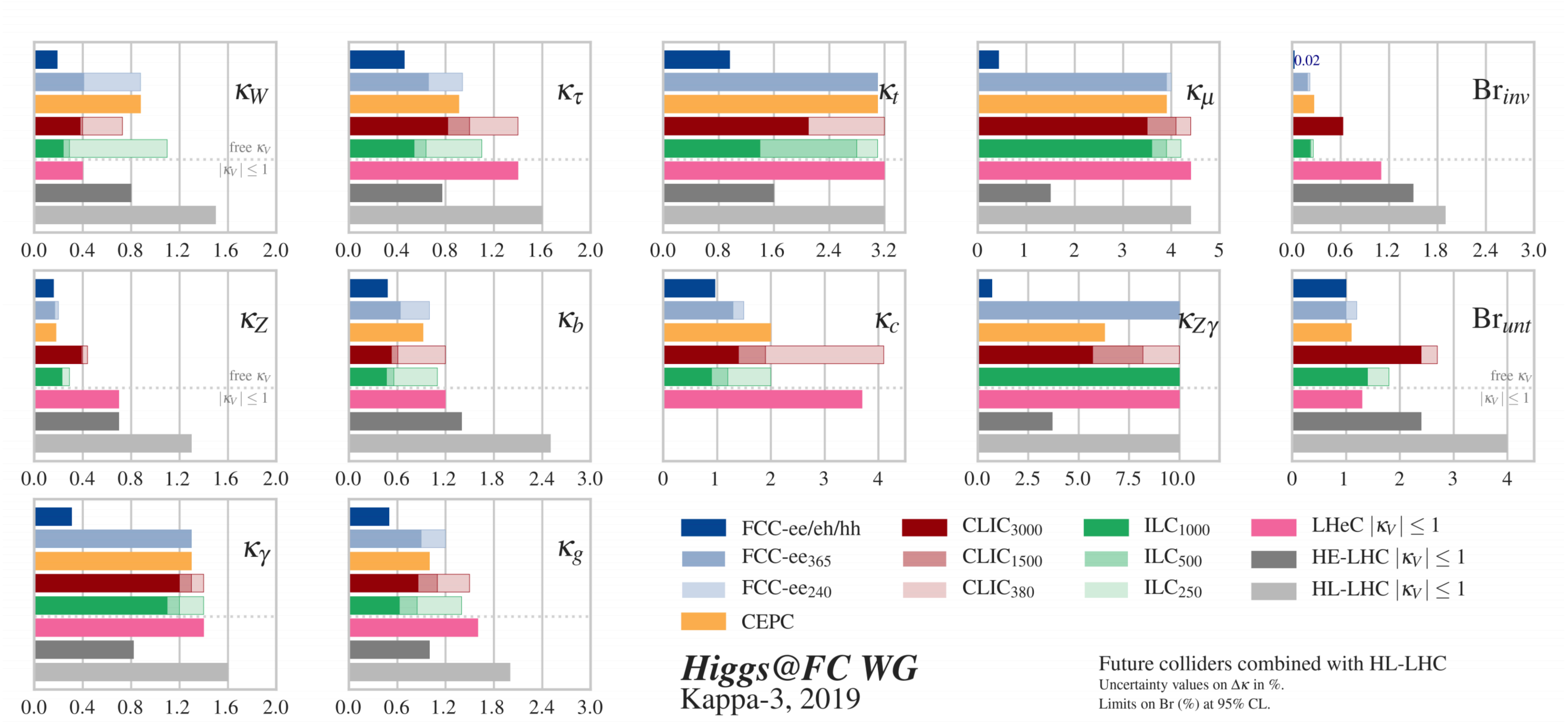


Leading systematic uncertainties

Overall impact of systematics uncertainties on measurement is a few percent.



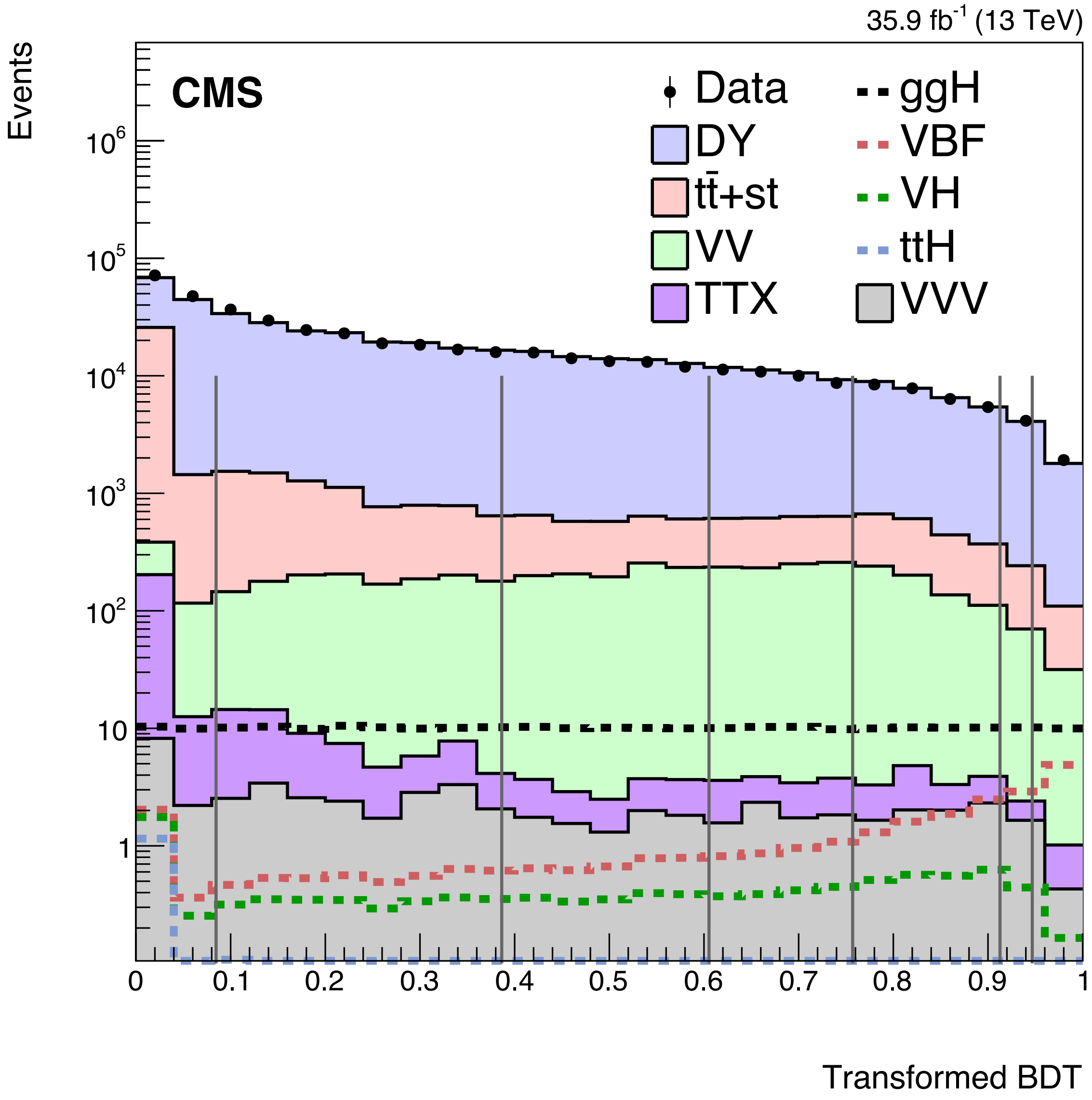
Higgs couplings at future colliders

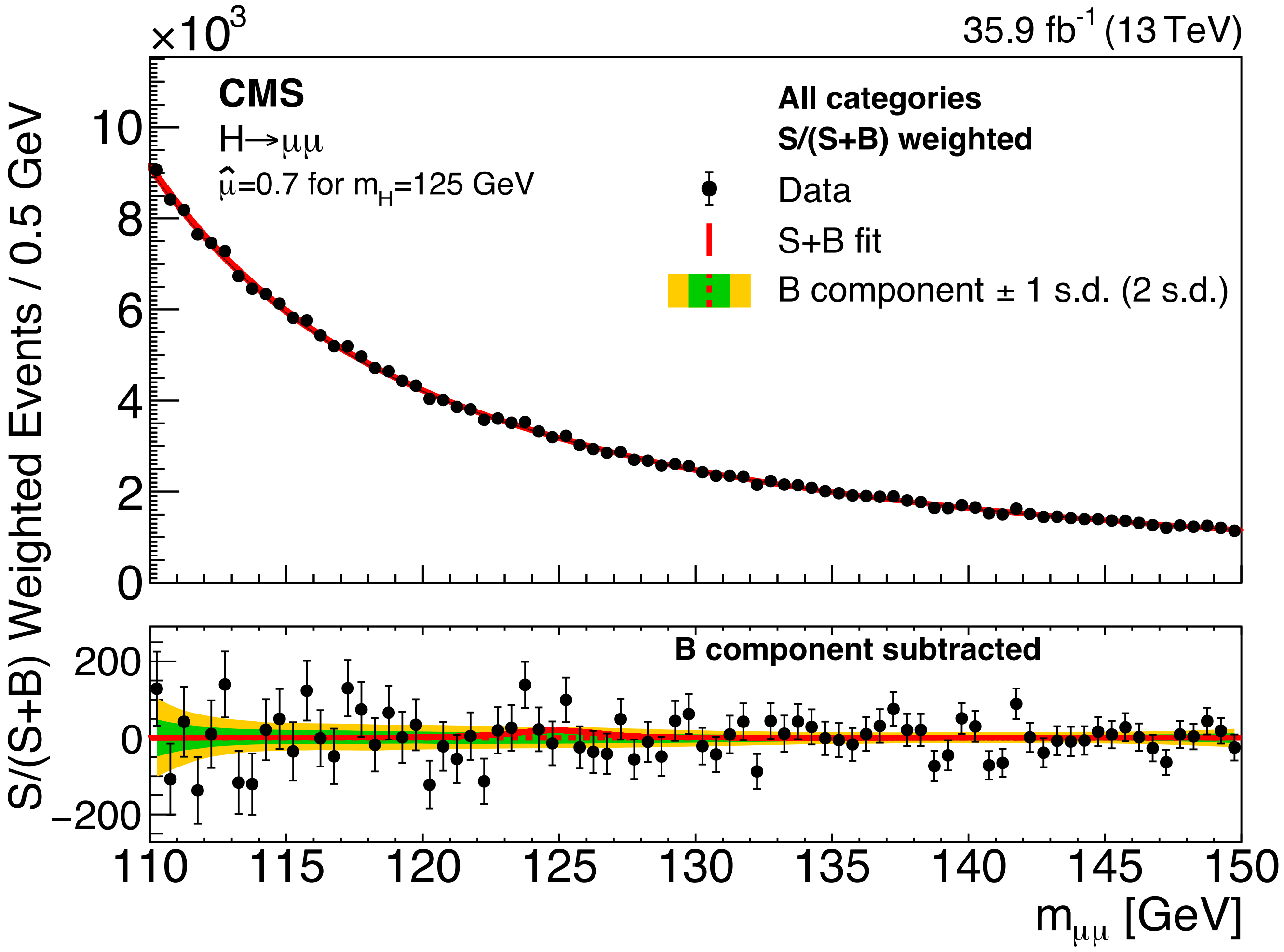


From Physics Briefing Book

Previous CMS result

- With 35.9 fb^{-1} 13 TeV data:
  - Observed (expected) significance 0.6 (0.9)σ, signal strength $\mu = 0.7 \pm 1.0$
- Combined with 7 and 8 TeV data:
  - Observed (expected) significance 0.9 (1.0)σ, signal strength $\mu = 1.0 \pm 1.0$





Phys. Rev. Lett. 122 (2019) 021801

Simulated samples and data sets

- SingleMuon primary data sets used by all channels.

- Background simulation (specifically requested for H→μμ):
  - DY Madgraph samples at NLO with m(μμ) [105,160] GeV filter, including set of VBF-enriched ($m_{jj,GEN}$ > 350 GeV) samples.
  - VBF Z Madgraph samples at LO with Herwig PS* and m(μμ) [105,160] GeV filter.
    - Detailed studies of stability of prediction vs. Madgraph version and $p_T(j)$ minimum threshold.

- Signal simulation:
  - ggH signal: MG_aMC at NLO with up to two partons in final state at ME level
  - VBF signal: POWHEG interfaced with PYTHIA parton shower with dipole recoil shower (more details in backup)
  - VH, ttH signal: POWHEG + PYTHIA PS

*as studied extensively in CMS VBF Z measurement:
Eur. Phys. J. C 78 (2018) 589

*highlighting just the most relevant MC samples used

Physics objects and corrections

Muons

- Muons passing medium ID and loose PF isolation.
- $p_T > 26\text{-}29\ (20)$ GeV, $|\eta| < 2.4$
- Momentum scale and resolution calibration with Rochester corrections.
- FSR recovery: strategy inspired by H(4l) and optimized for $H \rightarrow \mu\mu$.
  - Negligible $H \rightarrow Z\gamma$
- GeoFit correction: using beam spot as additional track constraint to improve resolution.
  - 3-10% improvement in $\sigma(m_{\mu\mu})$.

Jets

- AK4 CHS jets with $p_T > 25$ GeV and $|\eta| < 4.7$
- Loose (tight) jet ID in 2016 (2017/2018).
- Loose pileup jet ID for jets with $p_T < 50$ GeV.
- Dedicated 2017 pileup jet ID training $\Rightarrow$ 15% improvement in 2017 signal efficiency.
- Latest JEC applied, JER not applied but used for systematic uncertainty.

Electrons

- $p_T > 20$ GeV and $|\eta| < 2.5$
- Passing MVA ID WP90

B-tagging

- $p_T > 25$ GeV and $|\eta| < 2.5$
- Passing DeepCSV loose or medium WPs.

Other experimental and theory corrections

Pileup re-weight

- Applied taking the certified pileup profile in data
- **Uncertainty** estimated by varying the minimum bias cross section

L1 prefiring

- In 2016/17, **mis-timed ECAL TPs** lead to inefficiency in the L1 trigger
- **Corrections** measured from a set of **unprefirable events** by JetMET
- **Corrections applied** in the analysis and **only relevant** for **VBF channel**

DY $p_T(Z)$ spectrum

- DY MC poorly describe the data for $p_T(\mu\mu) <$ 40 GeV
- This is due to **missing resummation** effects [10.1007/JHEP12(2019)061]
- **Reweighting** performed using **data** with **70 < m(μμ) < 110 GeV**, as a function of $\mathbf{p_T(\mu\mu)}$ and $\mathbf{N_{jets}}$

ggH $p_T(H)$ spectrum

- **Reweighed** to **NNLOPS** in bins of N_{jets} at the generator level