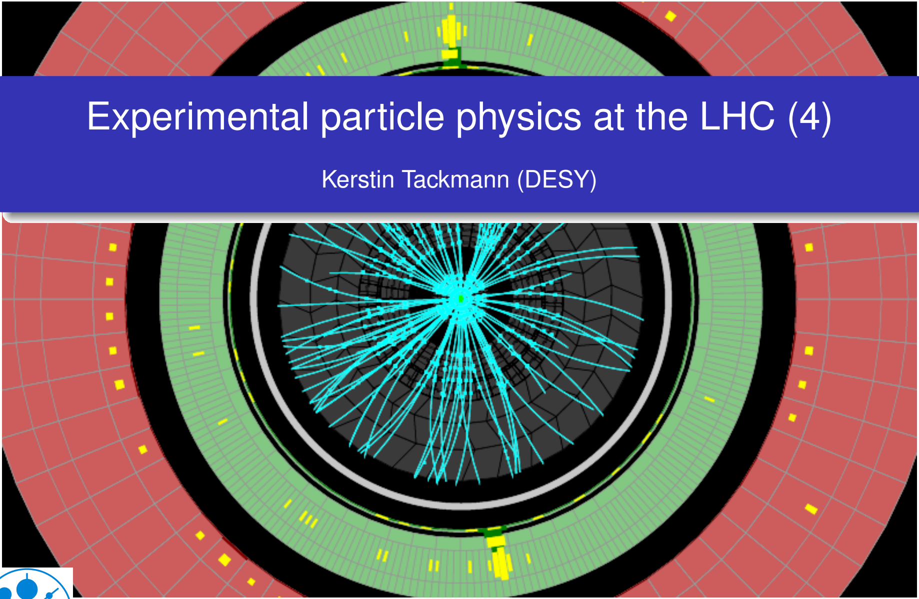
Experimental particle physics at the LHC (4)

Kerstin Tackmann (DESY)



GRK1504/2: Autumn Block Course 2016

Recap: cross section measurements (I)

$$\sigma = \frac{N_{\mathrm{meas}} - N_{\mathrm{bkgd}}}{\epsilon \cdot A \cdot \mathcal{B} \cdot \int \mathcal{L}\mathrm{d}t}$$

Experimental steps

- Estimate and subtract the background(s)
  - ⋆ Parameterize signal and background and fit to a distribution where signal and background look different ($m_{\gamma\gamma}$)
  - ⋆ Many analyses use control regions (in data) enriched in (a given type of) background to constrain contribution of background
- Correct for detector acceptance, and for efficiencies
  - ⋆ Detector and selection efficiencies as much as possible determined on control samples from data
  - ⋆ Acceptance corrections usually need to be based on theoretical predictions
  - ⋆ Fiducial measurements not or partially corrected for acceptance to reduce dependence on theoretical predictions
- If needed/wanted, correct for branching ratio(s)
- Determine the luminosity (see lecture 1)

Recap: cross section measurements (II)

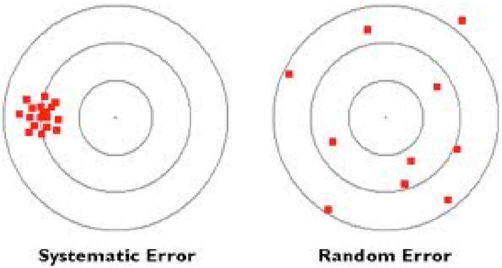
$$\sigma = \frac{N_{\mathrm{meas}} - N_{\mathrm{bkgd}}}{\epsilon \cdot A \cdot \mathcal{B} \cdot \int \mathcal{L}\mathrm{d}t}$$

Differential cross section in variable x: $\frac{\mathrm{d}\sigma}{\mathrm{d}x}$

- In practice: bin-averaged cross section $\frac{\Delta\sigma}{\Delta x}$
- Background estimation and subtraction, efficiency and acceptance corrections performed for every bin
- Requires correction of resolution effects in x: unfolding
  - ⋆ "Revert" the effect of imperfect detector resolution (migrations between bins)
  - ⋆ Different methods available and used (needs care to avoid introducing biases into the measurement)

Uncertainties

- Statistical uncertainties due to finite number of events
  - ⋆ In $H \rightarrow \gamma\gamma$, statistical uncertainties dominated by statistical uncertainties (fluctuations) in the background



- Systematic uncertainties related to analysis inputs, procedure, ...
  - ⋆ Understanding of detector and reconstruction
  - ⋆ Understanding of backgrounds
  - ⋆ ...

- Evaluation of systematic uncertainties usually requires dedicated study for each of the possible systematic uncertainties

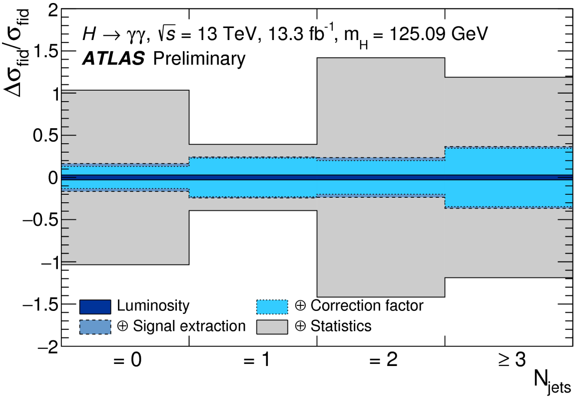
Uncertainties (non-differential measurements)

Source	Uncertainty on fiducial cross section (%)		
	Baseline	VBF-enhanced	single-lepton
Fit (stat.)	34.5	35.0	52.9
Fit (syst.)	9.0	11.1	9.3
Photon efficiency	4.4	4.4	4.4
Jet energy scale/resolution	-	9.4	-
Lepton selection	-	-	0.8
Pileup	1.1	2.0	1.4
Theoretical modelling	4.3	9.4	8.4
Luminosity	2.9	2.9	2.9

- Fit (stat.) statistical uncertainty, including contributions from floating the background parameters
- Fit (syst.) uncertainties on energy scale and resolution and background parametrization
- All others uncertainties on efficiency, acceptance and resolution corrections
  - ⋆ Theoretical modelling: Higgs production cross sections, Higgs kinematics, multiple parton interactions

Uncertainties (differential measurements)

- Uncertainties evaluated for each bin of a given observable
- Grouped into several categories:



Fiducial cross section measurements

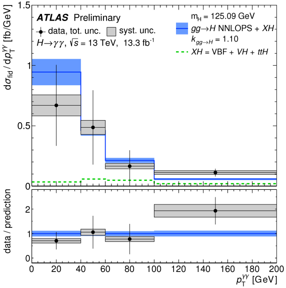
Fiducial cross sections with specific signatures and topologies

Fiducial region	Measured cross section (fb)	SM prediction (fb)	
Baseline	$43.2 \pm 14.9\,(\text{stat.}) \pm 4.9\,(\text{syst.})$	$62.8\,^{+3.4}_{-4.4}$	[N^{3}LO + XH]
VBF-enhanced	$4.0 \pm 1.4\,(\text{stat.}) \pm 0.7\,(\text{syst.})$	2.04 ± 0.13	[NNLOPS + XH]
single lepton	$1.5 \pm 0.8\,(\text{stat.}) \pm 0.2\,(\text{syst.})$	0.56 ± 0.03	[NNLOPS + XH]

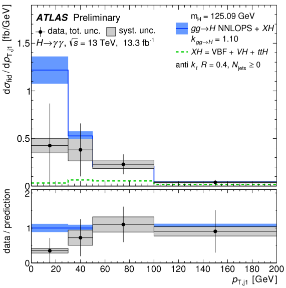
- Compared to theoretical predictions
  - $gg \to H$ N^{3}LO precision for total cross section, corrected for fiducial acceptance (with NNLOPS, with NNLO precision for total cross section) and $H \to \gamma\gamma$ branching ratio
  - VBF, VH, $t\bar{t}H$, ...: simulation samples reweighted to improved predictions for total cross sections
- Agreement with predictions to $1 - 2\,\sigma$

Differential cross section measurements (I)

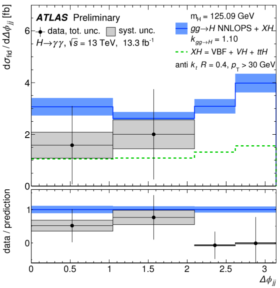
$p_T^{\gamma\gamma}$



p_T^{j1} (evts with $\geq$ 1 jet)

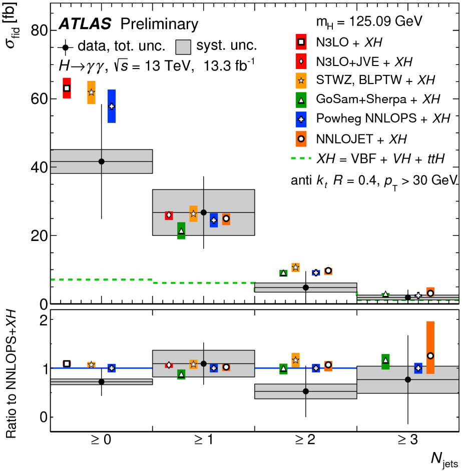


$\Delta\phi_{jj}$ (evts with $\geq$ 2 jets)



- Differential measurements presently dominated by statistical uncertainties
- Compared to predictions (NNLOPS for $gg \to H$, rescaled simulation for the other production processes)
- No significant disagreements between data and predictions within current uncertainties

Differential cross section measurements (II)



Inclusive jet cross sections (cross section for events with $\geq N$ jets) compared to a variety of theoretical predictions

- Analytical predictions for $gg \to H$ (e.g. N^{3}LO, STWZ/BLPTW)
- MC predictions (e.g. Powheg NNLOPS)

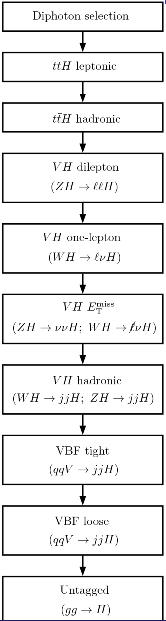
$H \to \gamma\gamma$ couplings measurement

- Measurements dedicated to understand Higgs boson production

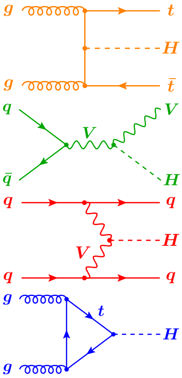
- Many of the steps are in common with cross section measurements
  - ⋆ Anything related to the two signal photons, general strategy for background subtraction
    - ▶ Will of course not discuss these again

- But some aspects come in addition to the cross section measurements
  - ⋆ Will concentrate on some of these

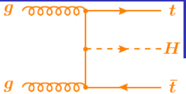
"Couplings analysis" in $H \to \gamma\gamma$



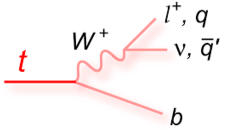
- Define exclusive event categories by specific signatures expected from the different production processes
  - ⋆ $t\bar{t}H$: events consistent with (semi)leptonic or hadronic $t\bar{t}$
  - ⋆ VH: events consistent with $Z \to \ell\ell$, $Z \to \nu\nu$, $W \to \ell\nu$ ($\ell = e, \mu$) or $V \to 2$jets
  - ⋆ VBF: events with 2 jets with topology consistent with VBF (large $\Delta\eta_{jj}$, large $p_{Tt}^{\gamma\gamma}$, jets and photons separated)
  - ⋆ $gg \to H$: kinematic separation ($p_{Tt}^{\gamma\gamma}$, η^{γ}) to increase sensitivity



$t\bar{t}H$-enriched categories (I)

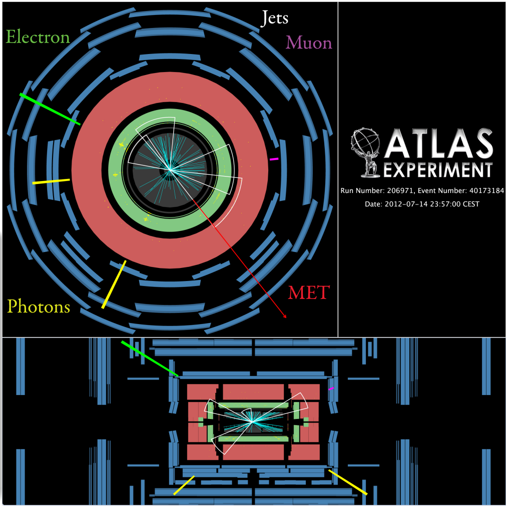


- Aim for high efficiency for $t\bar{t}H$, while suppressing other production modes



Search in two event categories

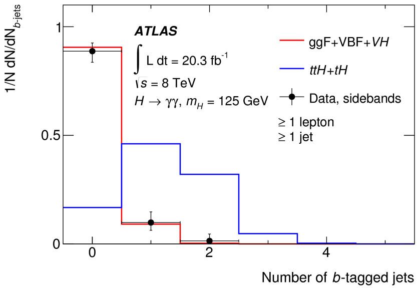
- Leptonic: 1 or 2 $t \to b\ell\nu$
  - ★ ≥1 electron or muon
  - ★ ≥1 b-tagged jet
  - ★ $E_T^{\mathrm{miss}} > 20$ GeV OR ≥1 b-tagged jet
- Fully hadronic: 2 $t \to bjj'$
  - ★ ≥ 5 jets (≥ 1 b-tagged)

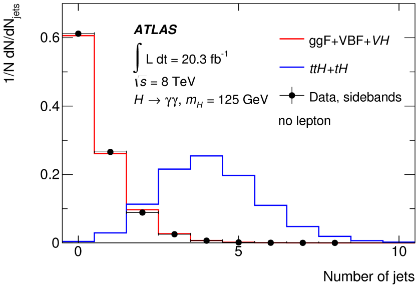


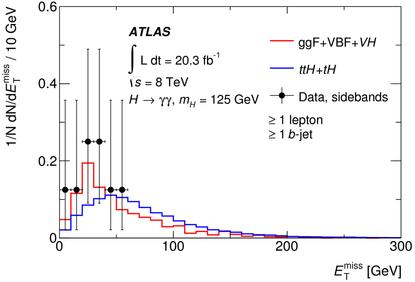
[ATLAS-CONF-2013-080]

$t\bar{t}H$-enriched categories (II)

- Selection requirements enrich in $t\bar{t}H$ events and suppress other Higgs production modes
- No discrimination between non-Higgs events and non-$t\bar{t}H$ events

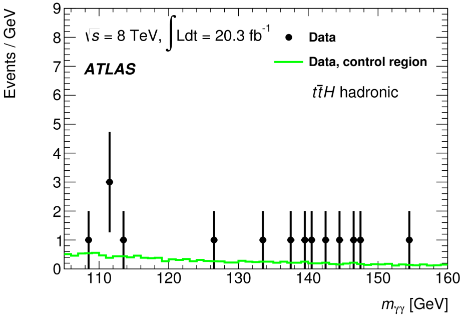






$t\bar{t}H$-enriched categories (III)

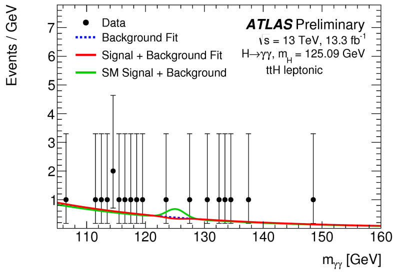
- Backgrounds $t\bar{t}$+photons, multi-jet/-photon
- Standard method to build background templates not applicable in $t\bar{t}H$
- Build control samples from data, e.g. $t\bar{t}H$ hadronic: revert identification or isolation requirements of photons, require 5 jets
  - ⋆ Note: plot from earlier analysis with different (control region) selection



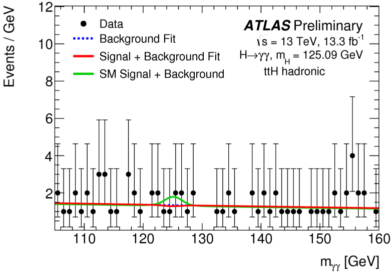
- Simple parametrization with exponential sufficient

$t\bar{t}H$-enriched categories (IV)

Leptonic



Hadronic



	$t\bar{t}H$ leptonic	$t\bar{t}H$ hadronic
Expected N_H	1.4	2.0
Expected $N_{t\bar{t}H}$	1.3	1.8
Expected purity	89%	88%
Fitted N_H	-0.2	-0.2

VH-enriched categories (I)

Dilepton ($Z \to \ell\ell$)

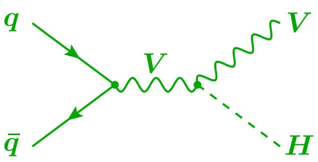
2 opposite-charge isolated electrons with $p_T^e > 10$ GeV or 2 opposite-charge isolated muons with $p_T^\mu > 10$ GeV with 70 GeV $< m_{\ell\ell} <$ 110 GeV

One lepton ($W \to \ell\nu$)

1 isolated electron $p_T^e > 10$ GeV or 1 isolated muon $p_T^\mu > 10$ GeV,
E_T^{miss} significance $\frac{E_T^{\text{miss}}}{\sqrt{\sum E_T}} > 4.5$, $p_T^{\gamma\gamma} > 60$ GeV

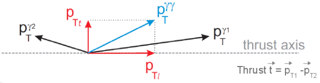
Missing transverse momentum ($W \to \ell\nu$, $Z \to \nu\nu$)

E_T^{miss} significance $\frac{E_T^{\text{miss}}}{\sqrt{\sum E_T}} > 7$, $p_T^{\gamma\gamma} > 90$ GeV



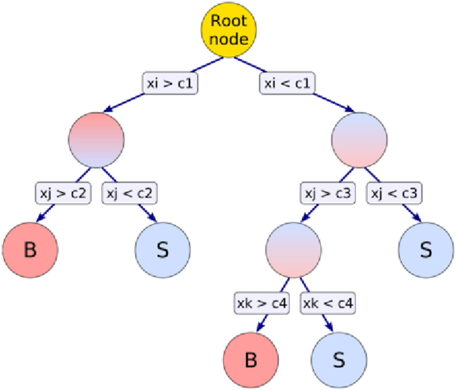
Dijet ($W \to jj$, $Z \to jj$)

≥ 2 jets with 50 GeV $< m_{jj} <$ 150 GeV,
BDT based on m_{jj}, p_{Tt}, $\cos\theta^*_{\gamma\gamma,jj}$



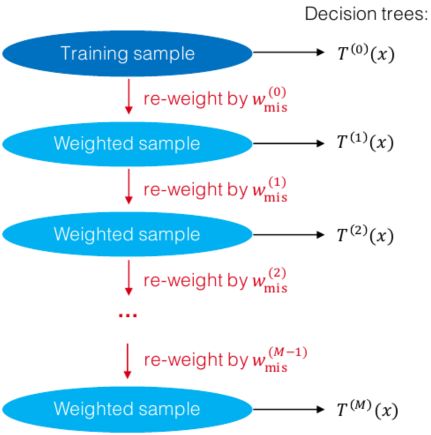
Interlude: (Boosted) decision trees (I)

- Decision tree: sequential application of cuts splits the data (into nodes), where the final decisions (nodes) calssify an event as signal or background
- First split training samples according to cut on best variable
- Continue splitting until min. number of events or max. purity reached



- Boosted decision tree: combination of many decision trees, with differently weighted events in each tree (trees themselves can be weighted)

Interlude: (Boosted) decision trees (II)



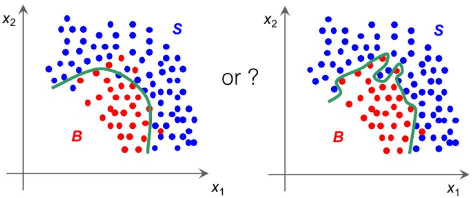
- Emphasize different feature in data sample, e.g. events that are hard to classify
- E.g. adaptove boosting reweights events misclassified by previous classifier by

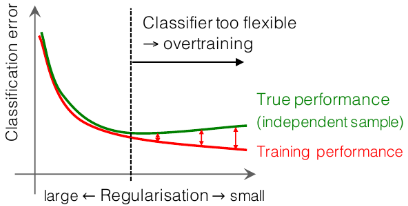
$$w_{\mathrm{mis}}^{(i)} = \frac{1 - f_{\mathrm{mis}}^{(i)}}{f_{\mathrm{mis}}^{(i)}}$$

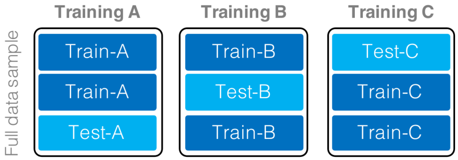
  with f_{mis} fraction of misclassified events
- Final decision obtained from weighted sum over all decision trees
- Different boosting algorithms exist

Interlude: (Boosted) decision trees (III)

- Danger: overtraining – tuning to statistically insignificant information in the training sample
- Reduced by increased size of training sample





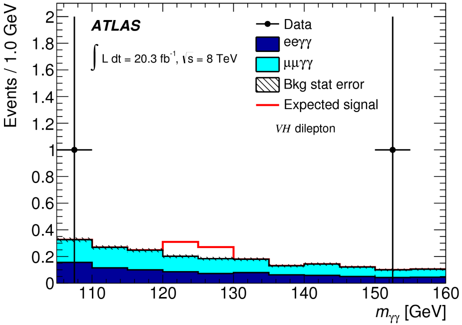


Cross validation

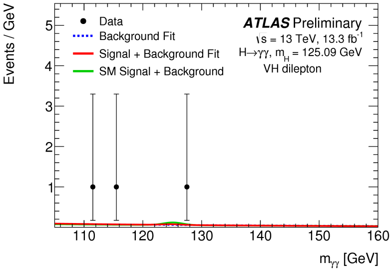
- Divide full data into n subsamples
- Train on all but subsample i
- Test performance on subsample i

Back to: VH-enriched categories (II)

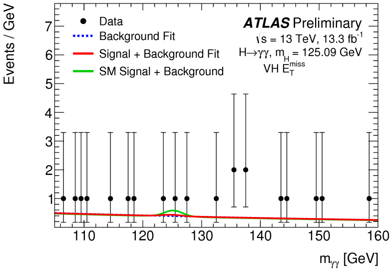
- Backgrounds dominated by $Z\gamma\gamma$ and $W\gamma\gamma$
- As for $t\bar{t}H$, standard method to build background templates not applicable for non-hadronic V decays



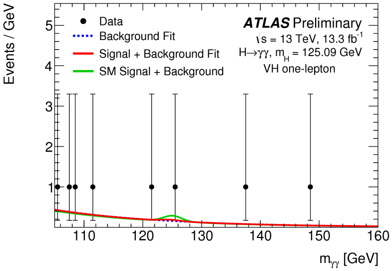
	dilep	one lep	$E_T^{\rm miss}$	dijet1	dijet2
Expected N_H	0.3	0.6	0.9	2.6	9.9
Expected purity ZH	95%	3%	56%	17%	11%
Expected purity WH	0%	84%	29%	28%	17%
Fitted N_H	0.07	0.12	0.18	1.0	4.7



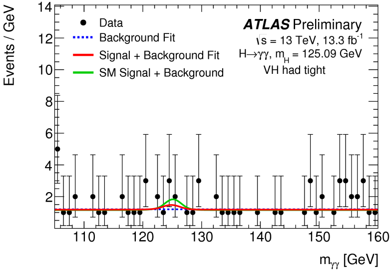
Events / GeV
Data
Background Fit
Signal + Background Fit
SM Signal + Background
ATLAS Preliminary
√s = 13 TeV, 13.3 fb⁻¹
H→γγ, m_H = 125.09 GeV
VH dilepton
m_γγ [GeV]



Events / GeV
Data
Background Fit
Signal + Background Fit
SM Signal + Background
ATLAS Preliminary
√s = 13 TeV, 13.3 fb⁻¹
H→γγ, m_H = 125.09 GeV
VH E_T^miss
m_γγ [GeV]



Events / GeV
Data
Background Fit
Signal + Background Fit
SM Signal + Background
ATLAS Preliminary
√s = 13 TeV, 13.3 fb⁻¹
H→γγ, m_H = 125.09 GeV
VH one-lepton
m_γγ [GeV]



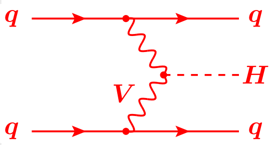
Events / GeV
Data
Background Fit
Signal + Background Fit
SM Signal + Background
ATLAS Preliminary
√s = 13 TeV, 13.3 fb⁻¹
H→γγ, m_H = 125.09 GeV
VH had tight
m_γγ [GeV]

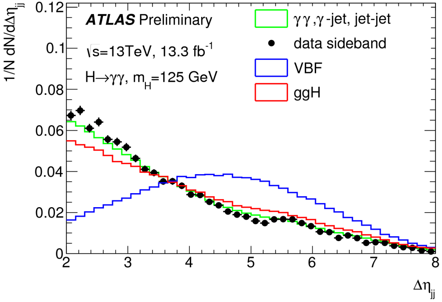
VBF-enriched categories (I)

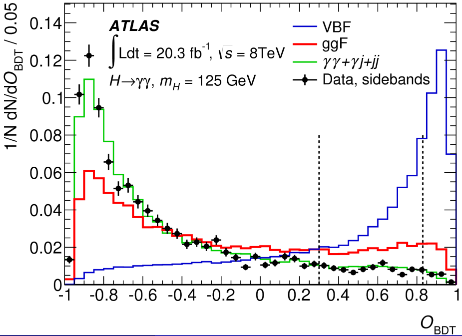
Select with 2 jets and VBF topology:

- 2 well-separated jets ($\Delta\eta_{jj}$, m_{jj})
- Boosted diphoton system ($p_{Tt}^{\gamma\gamma}$)
- Jet-photon separation ($\Delta\phi_{\gamma\gamma;jj}$, $\eta^* = \eta_{\gamma\gamma} - 1/2(\eta_{j1} + \eta_{j2})$, $\Delta R_{\min}^{\gamma j}$)

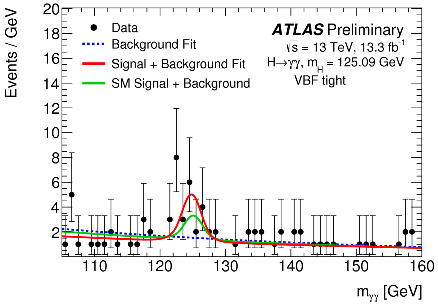
→ Combined in a BDT

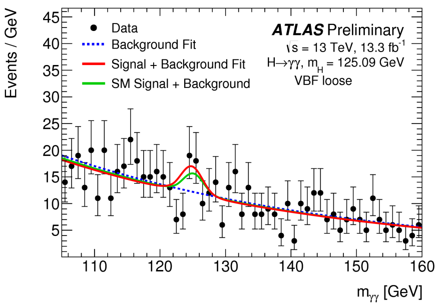






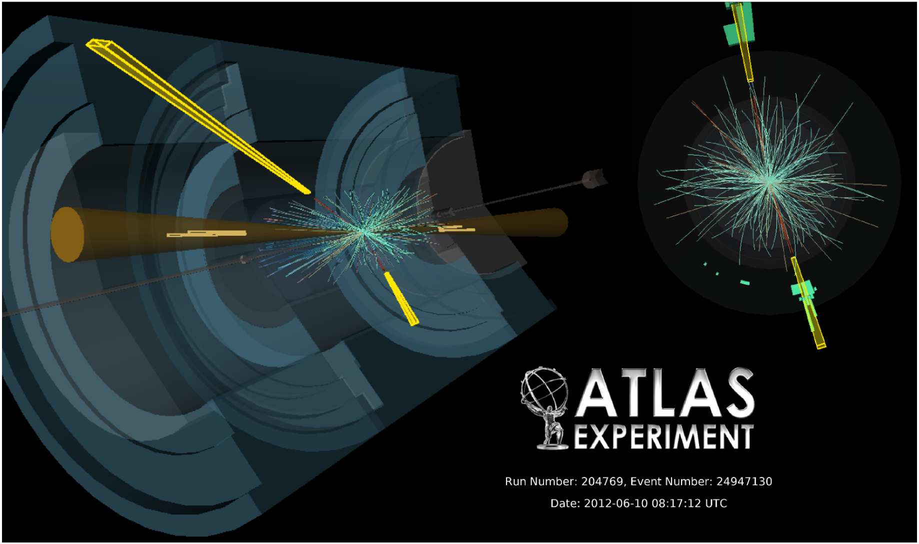
VBF-enriched categories (II)





	tight	loose
Expected N_H	8.1	17
Expected purity	76%	52%
Fitted N_H	13	21

2-Jets candidate



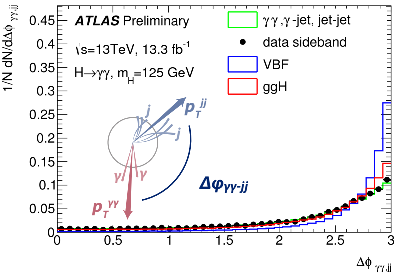
[Phys. Lett. B 726 (2013)]

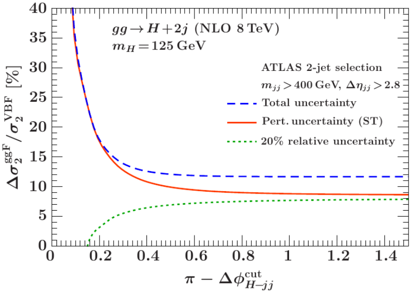
VBF-enriched categories: theoretical uncertainties

- VBF categories selected using a BDT based on 6 kinematic variables
- $\Delta\phi_{\gamma\gamma jj}$ sensitive to presence of third jet ($\rightarrow$ indirect jet veto)
- Substantial theoretical uncertainties on acceptance for $gg \rightarrow H$ for this specific region of phase space
  - ⋆ Due to strong restriction of additional QCD radiation

	σ
VBF tight	52%
VBF loose	25%

[8 TeV analysis]





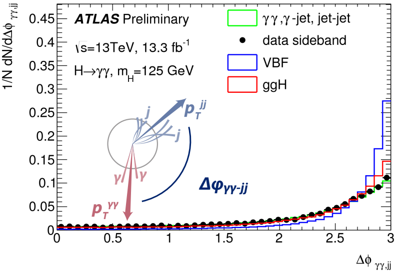
[Phys.Rev. D87 (2013) no.9, 093008]

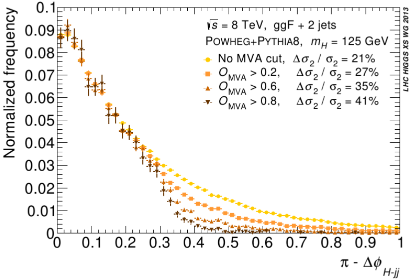
VBF-enriched categories: theoretical uncertainties

- VBF categories selected using a BDT based on 6 kinematic variables
- $\Delta\phi_{\gamma\gamma jj}$ sensitive to presence of third jet ($\rightarrow$ indirect jet veto)
- Substantial theoretical uncertainties on acceptance for $gg \rightarrow H$ for this specific region of phase space
  - ⋆ Due to strong restriction of additional QCD radiation

	σ
VBF tight	52%
VBF loose	25%

[8 TeV analysis]





Untagged categories (I)

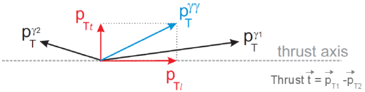
Remaining events are categorized based on photon η and p_{Tt}

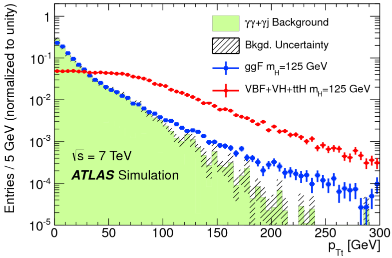
- Separate events with better and worse signal-to-background
- Separate events with better and worse signal resolution

Central Both photons have $|\eta| < 0.95$

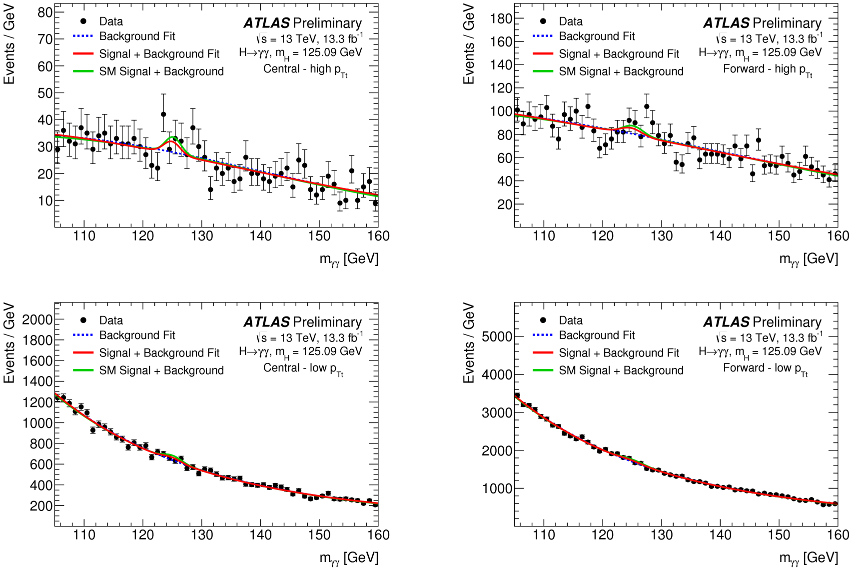
Forward At least one photon has $|\eta| >$0.95

Central and forward divided into $p_{Tt} < 70\,\text{GeV}$ (low) and $p_{Tt} > 70\,\text{GeV}$ (high)



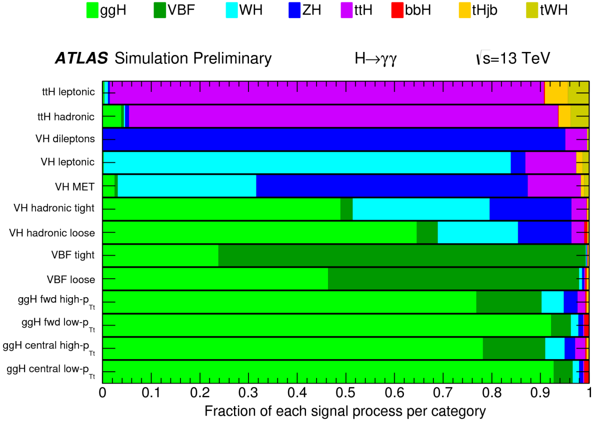


Untagged categories (II)



Categories ↔ production processes

Event categories are enriched in events from a given production process



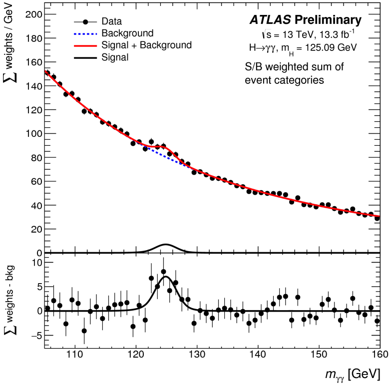
Then "unfold" to signal strength per production process μ_i by one global fit:

$$N_k^{\mathrm{sig}} = \sum_i \mu_i \sigma_i^{\mathrm{SM}} \cdot \mathcal{B}(H \to \gamma\gamma)^{\mathrm{SM}} \cdot \epsilon_{ik} \cdot A_{ik} \cdot \int L\,\mathrm{d}t$$

with $\mu_i = \sigma_i \cdot \mathcal{B}(H \to \gamma\gamma) / \sigma_i^{\mathrm{SM}} \cdot \mathcal{B}(H \to \gamma\gamma)^{\mathrm{SM}}$

Effect of categorization on analysis (I)

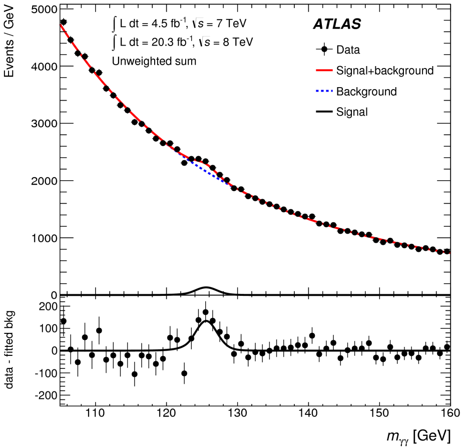
- Weighting each event by the expected signal-to-background ratio of its category



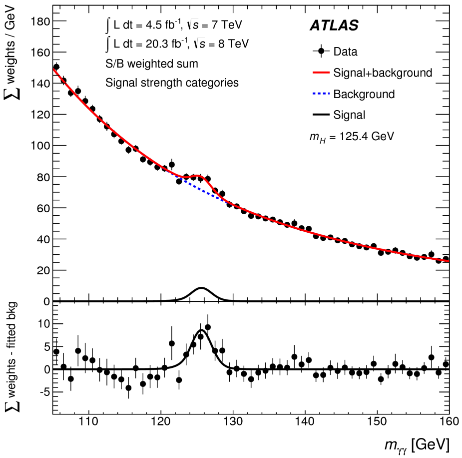
- Visual impression of effect of event categorization on analysis
  - ⋆ Gain beyond quantities that rely on the separation of production modes

Effect of categorization on analysis (II)

Unweighted

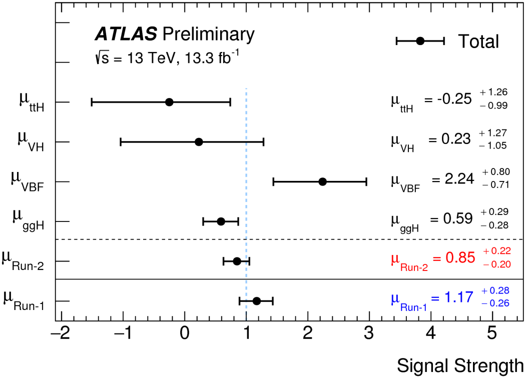


Weighted



Signal strength measurements

Results presented at ICHEP 2016



Caveat: $\mu_{\mathbf{Run-1}}$ was derived assuming an older prediction for the Higgs production cross section than $\mu_{\mathbf{Run-2}}$, where the gluon fusion production cross section is larger by approximately 10%

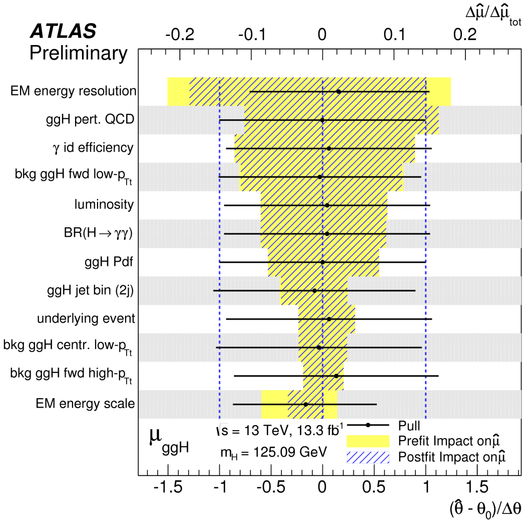
Cross section measurements

- Avoid problems like this by measuring cross sections

$$\begin{aligned}
\sigma_{ggH} \times \mathcal{B}(H \to \gamma\gamma) &= 65\ ^{+32}_{-31}\ \text{fb} \\
\sigma_{\text{VBF}} \times \mathcal{B}(H \to \gamma\gamma) &= 19.2\ ^{+6.8}_{-6.1}\ \text{fb} \\
\sigma_{VH} \times \mathcal{B}(H \to \gamma\gamma) &= 1.2\ ^{+6.5}_{-5.4}\ \text{fb} \\
\sigma_{t\bar{t}H} \times \mathcal{B}(H \to \gamma\gamma) &= -0.3\ ^{+1.4}_{-1.1}\ \text{fb}
\end{aligned}$$

- In future, move to measuring cross sections instead of signal strength(s)

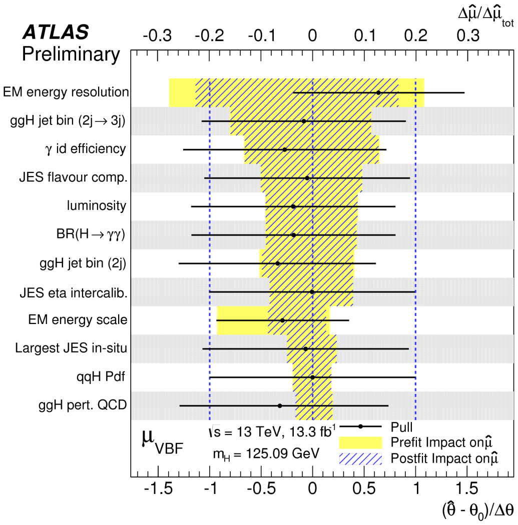
Systematic uncertainties: $gg \to H$



→ What is the impact of a given uncertainty on the measurement?

- Dominant experimental: energy resolution, photon id efficiency, background parametrization
- Dominant theoretical: $gg \to H$ missing higher orders, $\mathcal{B}(H \to \gamma\gamma)$, $gg \to H$ Pdf
  - ⋆ Substantially smaller for cross section measurement

Systematic uncertainties: VBF



Other uncertainties are important (due to use of jets)

- Dominant experimental: energy resolution, photon id efficiency, jet energy scale
- Dominant theoretical: $gg \to H$ missing higher orders in VBF phase space, $\mathcal{B}(H \to \gamma\gamma)$
  - ★ Phase space-specific uncertainties are not reduced in measurement of total VBF cross section

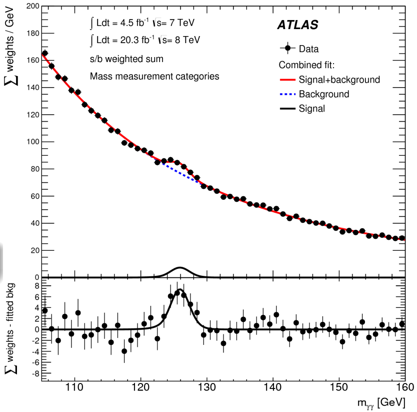
Bonus: Mass measurement

- Mass measured from Higgs peak position

- Dedicated event categorization: 10 categories according to η^γ, converted/unconverted γ and p_{Tt}
  - ⋆ Splitting motivated by systematic uncertainties

m_H =125.98±0.42(stat)±0.28(syst) GeV

- Dominant systematic uncertainty from energy scale



Mass measurement: systematic uncertainties