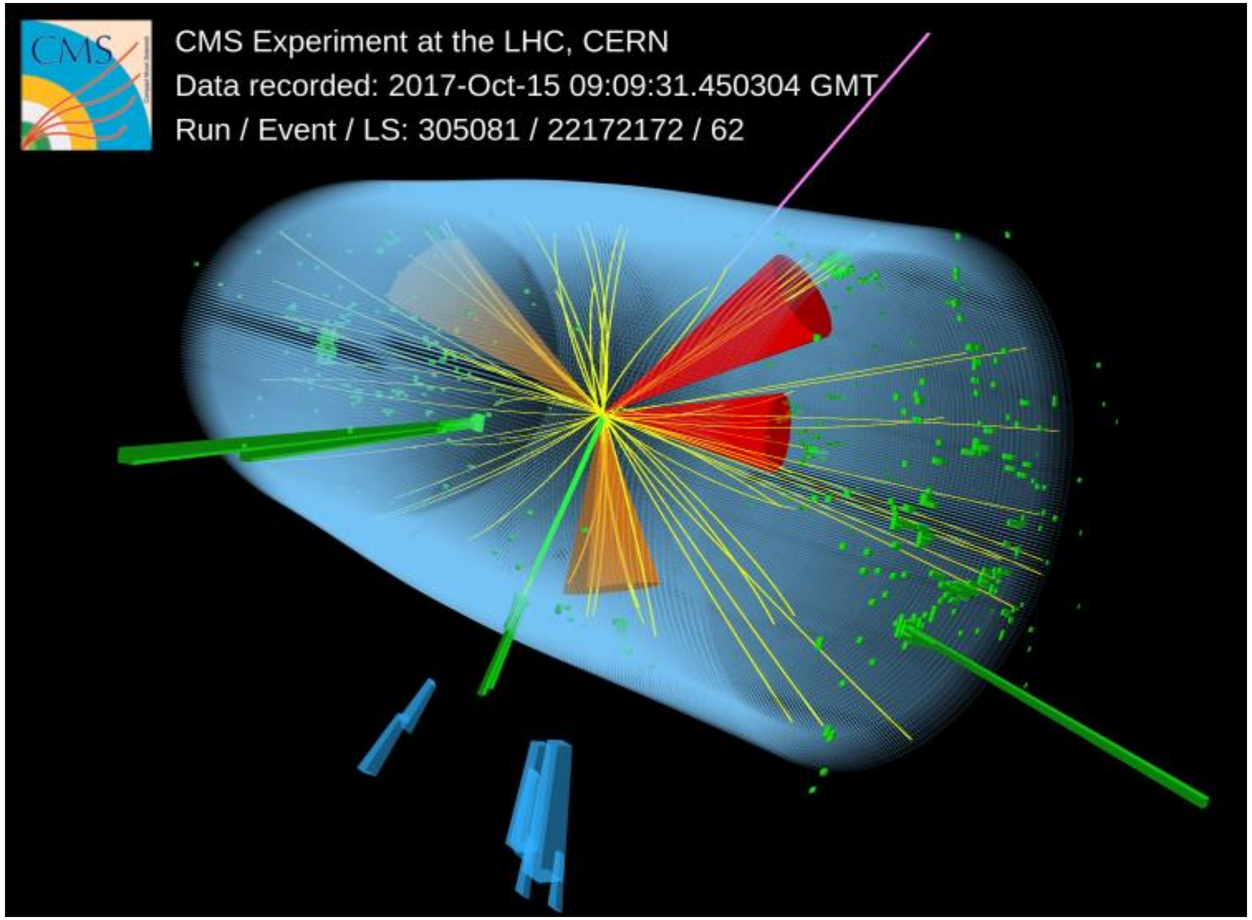
Measurements of $t\bar{t}H$ production and the CP structure of the Yukawa interaction between the Higgs boson and top quark in the diphoton decay channel

arXiv:2003.10866 [hep-ex]

João Seabra



What does the paper report?

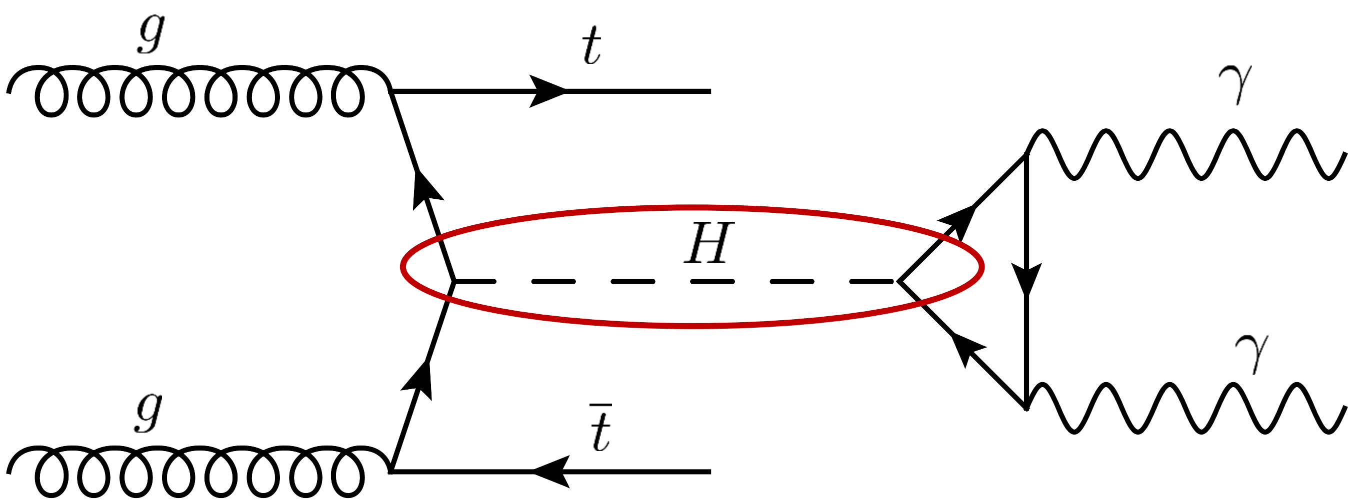
Measurement of the production rate of $t\bar{t}H$ with $H \to \gamma\gamma$

First observation of the tree level $t\bar{t}H$ coupling in a single decay channel of the Higgs boson

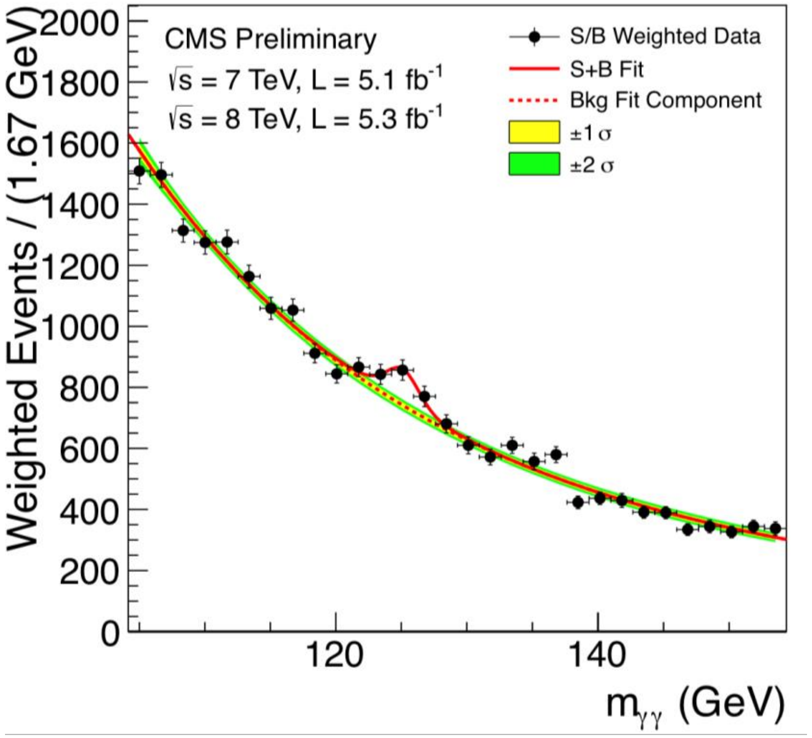
First test of the CP structure of the $t\bar{t}H$ coupling

- ✓ The results are based on a sample of proton-proton (pp) collisions at a center of mass energy $\sqrt{s} = 13\ \mathrm{TeV}$;
- ✓ The sample was collected between 2016 and 2018 using the CMS detector at the LHC, corresponding to an integrated luminosity of $137\ \mathrm{fb}^{-1}$.

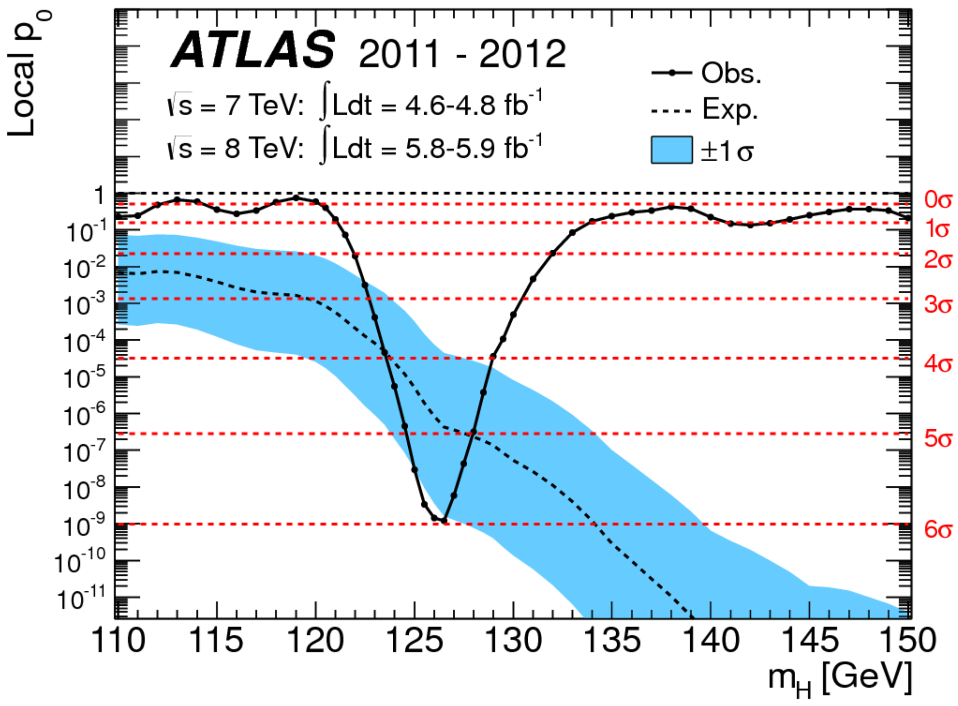
Introduction



While searching for the Standard Model (SM) Higgs boson, both ATLAS and CMS collaborations observed a new particle with mass 125 GeV.



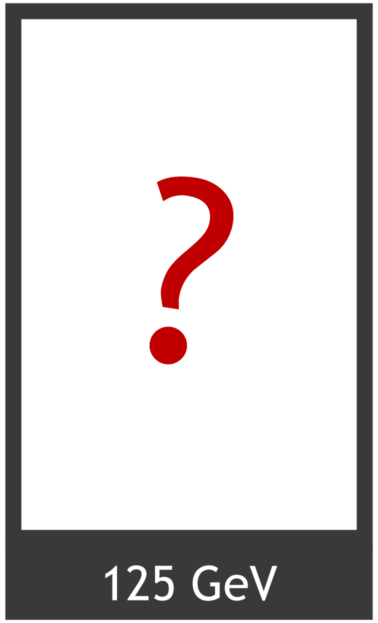
ATLAS Collaboration: Phys.Lett. B716 (2012)



CMS Collaboration: Phys. Lett. B 716 (2012)

Since that moment...

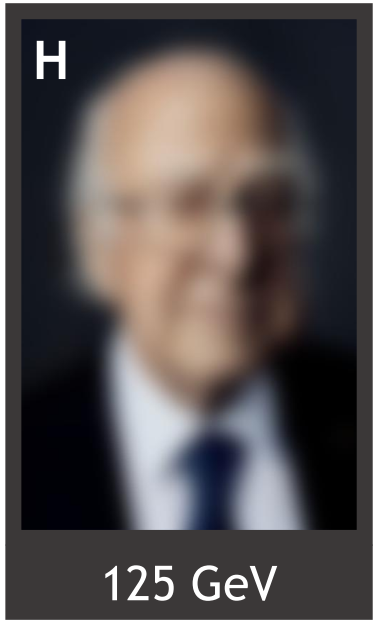
- ✓ Several studies of the properties of this new boson have been made at the LHC;



Since that moment...

- ✓ Several studies of the properties of this new boson have been made at the LHC;

- ✓ **So far, all measurements from ATLAS and CMS have shown that the properties of the boson discovered in 2012 are consistent with expectations for the SM Higgs boson;**

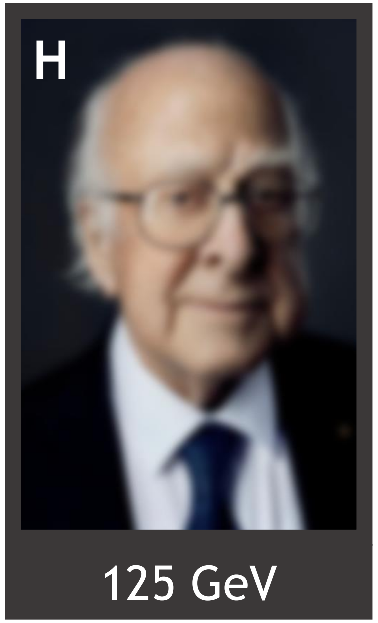


Since that moment...

- ✓ Several studies of the properties of this new boson have been made at the LHC;

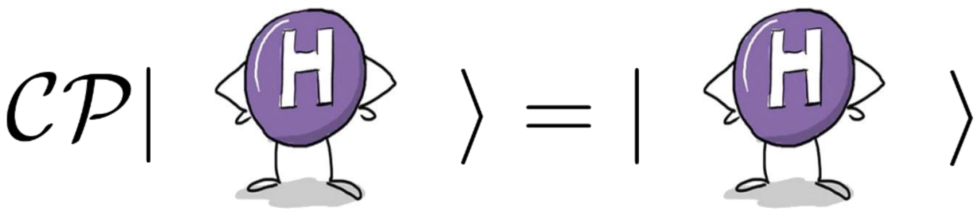
- ✓ **So far, all measurements from ATLAS and CMS have shown that the properties of the boson discovered in 2012 are consistent with expectations for the SM Higgs boson;**

- ✓ With the LHC collecting more data and the development of new analysis techniques, experimentalists have been able to measure more precisely the Higgs boson's properties, such as its behaviour under CP transformations...

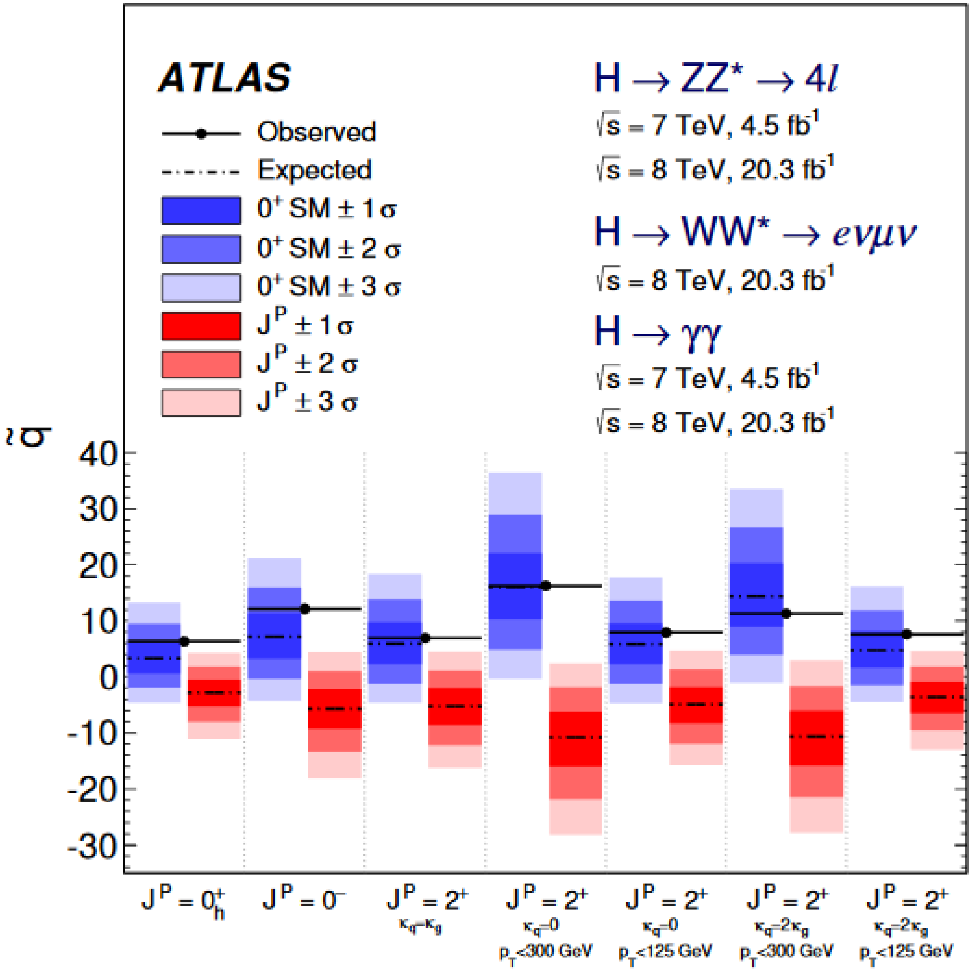


CP structure of the SM Higgs sector

- ✓ The SM Higgs boson is CP-even;



- ✓ Considering interactions between the Higgs boson and vector bosons, CMS and ATLAS have determined that the quantum numbers of the new 125 GeV boson are consistent with $J^{PC} = 0^{++}$;



M. Tanabashi *et al*. (Particle Data Group), Phys. Rev. D **98**, 030001 (2018) and 2019 update.

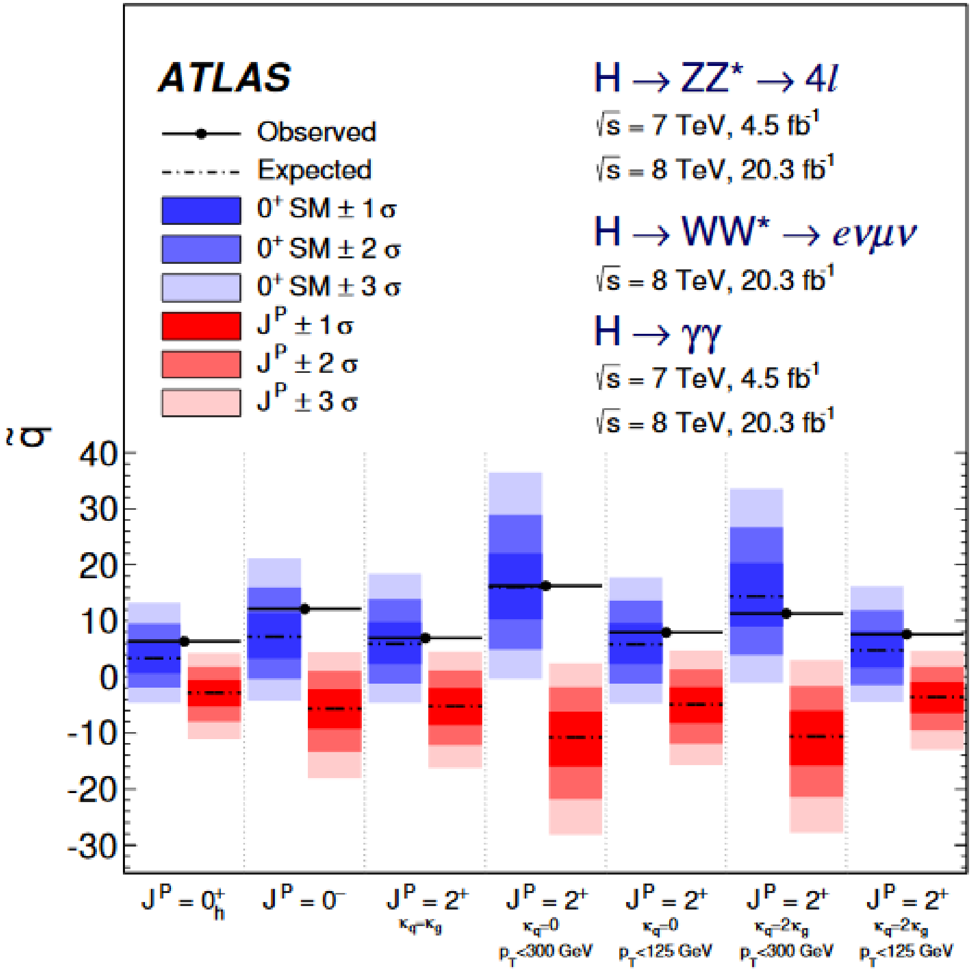
CP structure of the SM Higgs sector

- ✓ The SM Higgs boson is CP-even;

$\mathcal{CP}|$  $\rangle = |$  $\rangle$

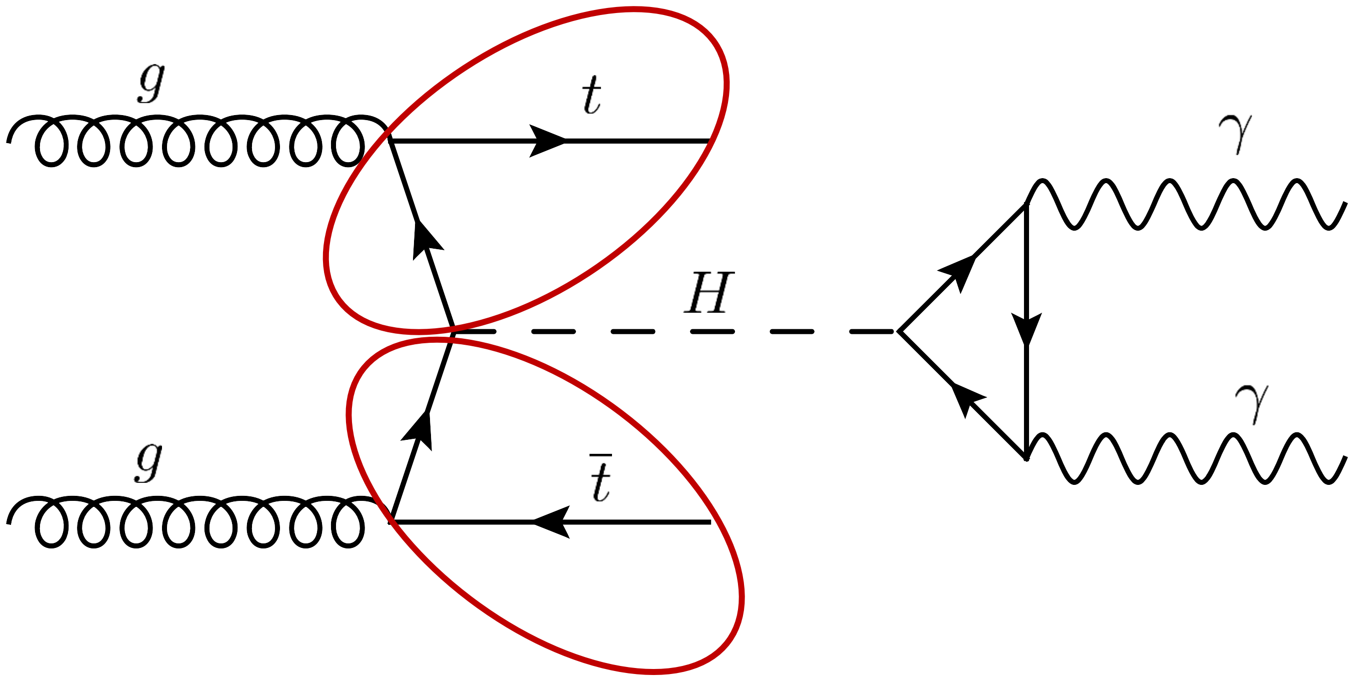
- ✓ Considering interactions between the Higgs boson and vector bosons, CMS and ATLAS have determined that the quantum numbers of the new 125 GeV boson are consistent with $J^{PC} = 0^{++}$;

Studies of the $t\bar{t}H$ Yukawa coupling provide an alternative and independent path for CP tests in the Higgs sector.



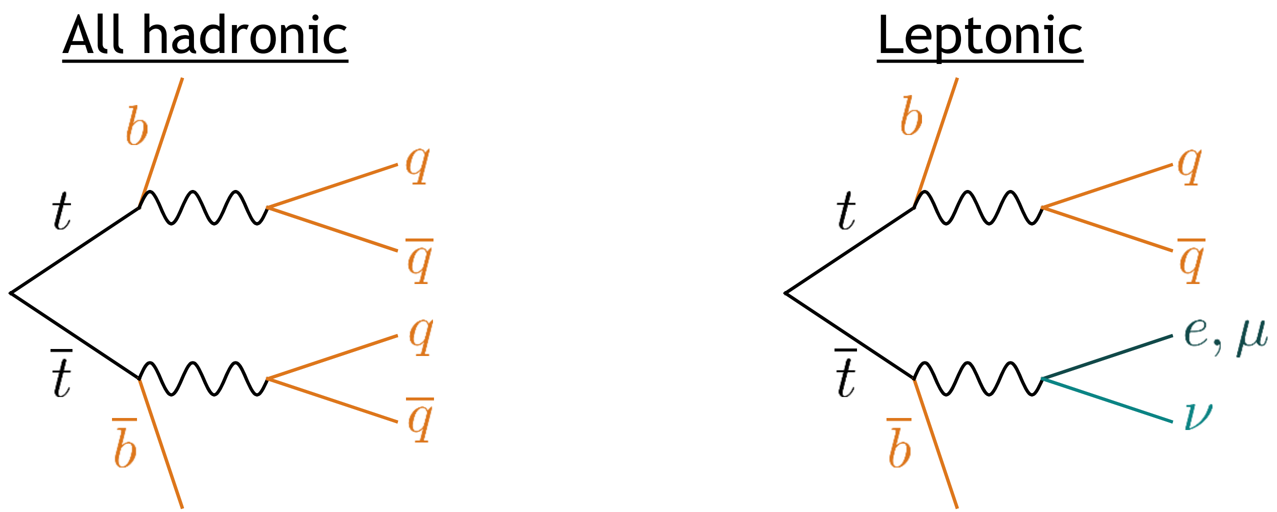
M. Tanabashi *et al*. (Particle Data Group), Phys. Rev. D **98**, 030001 (2018) and 2019 update.

Introduction



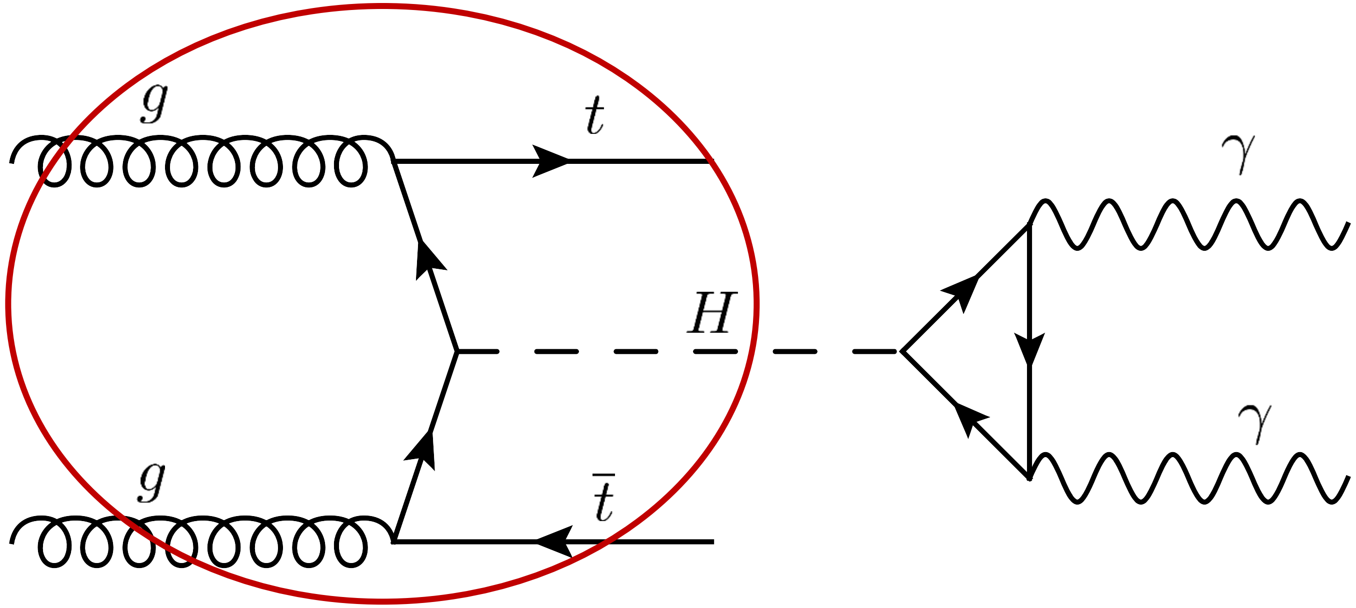
Top quark

- ✓ It is the heaviest particle of the SM;
- ✓ Mostly produced in pairs with both quarks decaying almost always in a W boson and a b-quark;
- ✓ According to the decays of the W bosons, three decay channels may be defined:



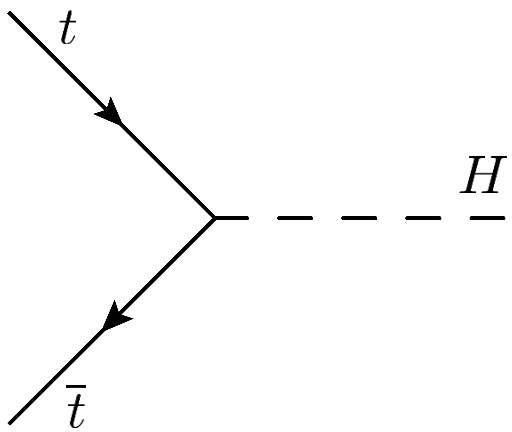
- ✓ b-jets are always present, so b-jet reconstruction plays an important role.

Introduction



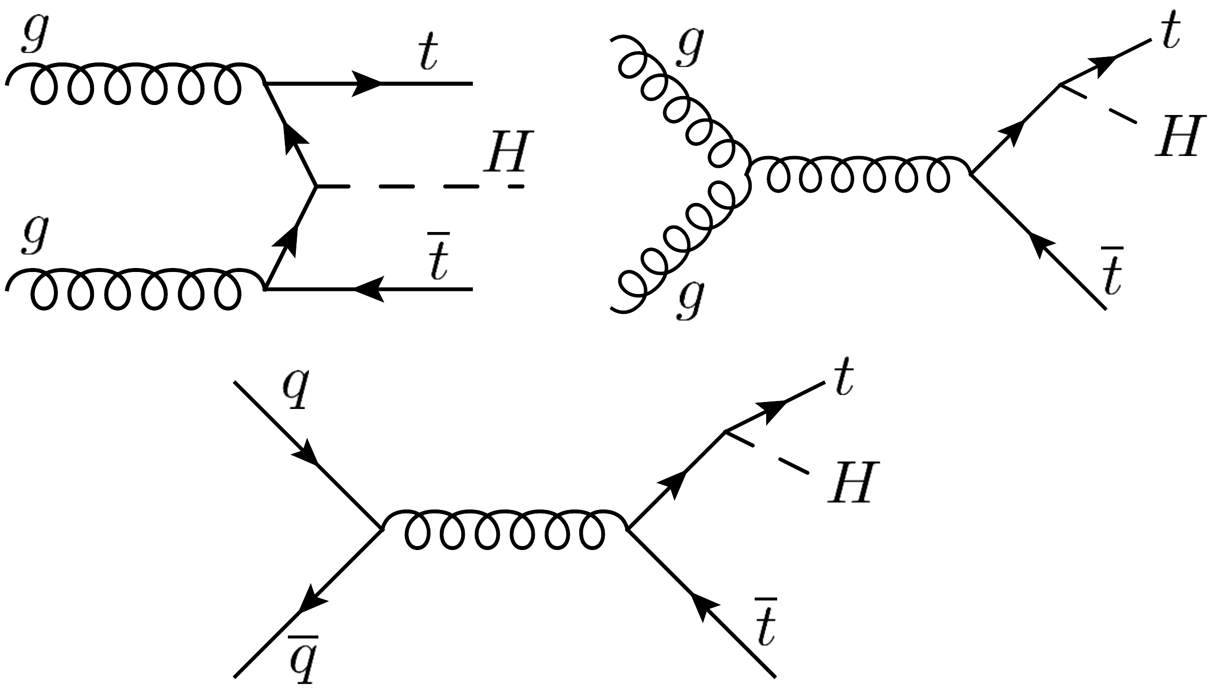
$t\bar{t}H$ Yukawa coupling

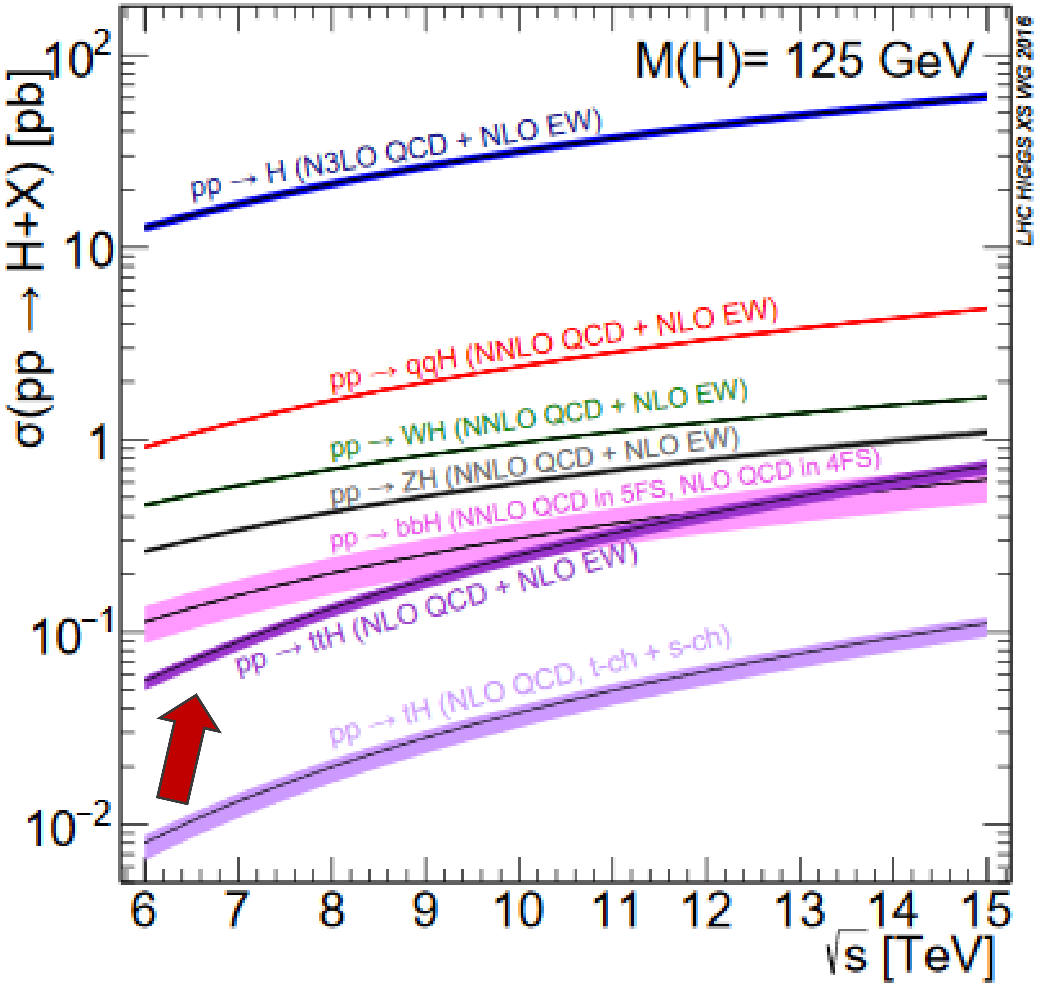
- ✓ It is the strongest Yukawa coupling of the SM (because the top-quark is the heaviest fermion);
- ✓ On-shell top quarks are too heavy to be produced in a Higgs boson decay, so we cannot test from it the measurement of a decay rate;
- ✓ One is still able to measure it directly by observing the production of a Higgs boson alongside a top quark and antiquark ($t\bar{t}H$ production).



$t\bar{t}H$ production

- ✓ Examples of Leading order Feynman diagrams:

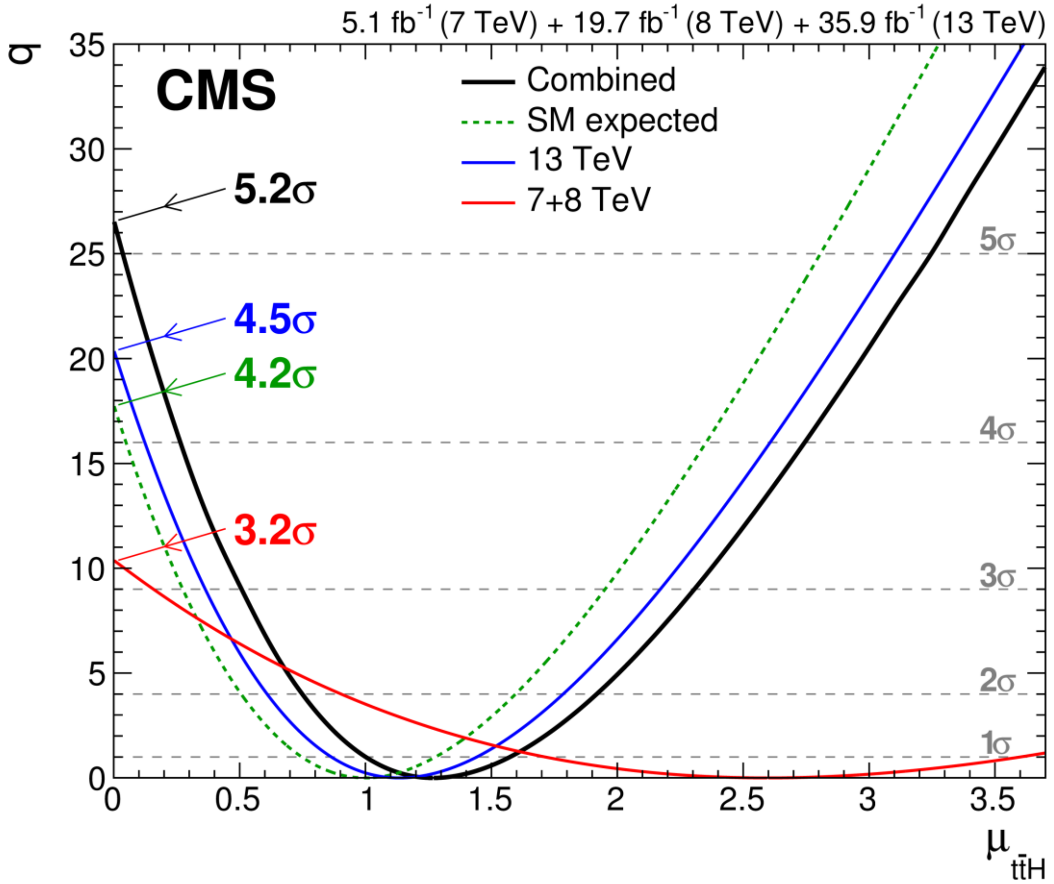




- ✓ Only 1% of all Higgs bosons are produced in association with a top quark pair.

Challenging to observe... but possible!

- ✓ In 2018, both ATLAS and CMS announced the observation of $t\bar{t}H$ production with a cross section compatible with the SM expectation;
- ✓ A combined analysis of pp collision data at center-of-mass energies of $\sqrt{s} = 7,\ 8$ and $13\ \mathrm{TeV}$ was required;
- ✓ To maximize sensitivity, the results of independent searches for $t\bar{t}H$ production using different decay channels of the SM Higgs boson were also combined.



CMS Collaboration: PhysRevLett.120.231801

Challenging to observe… but possible!

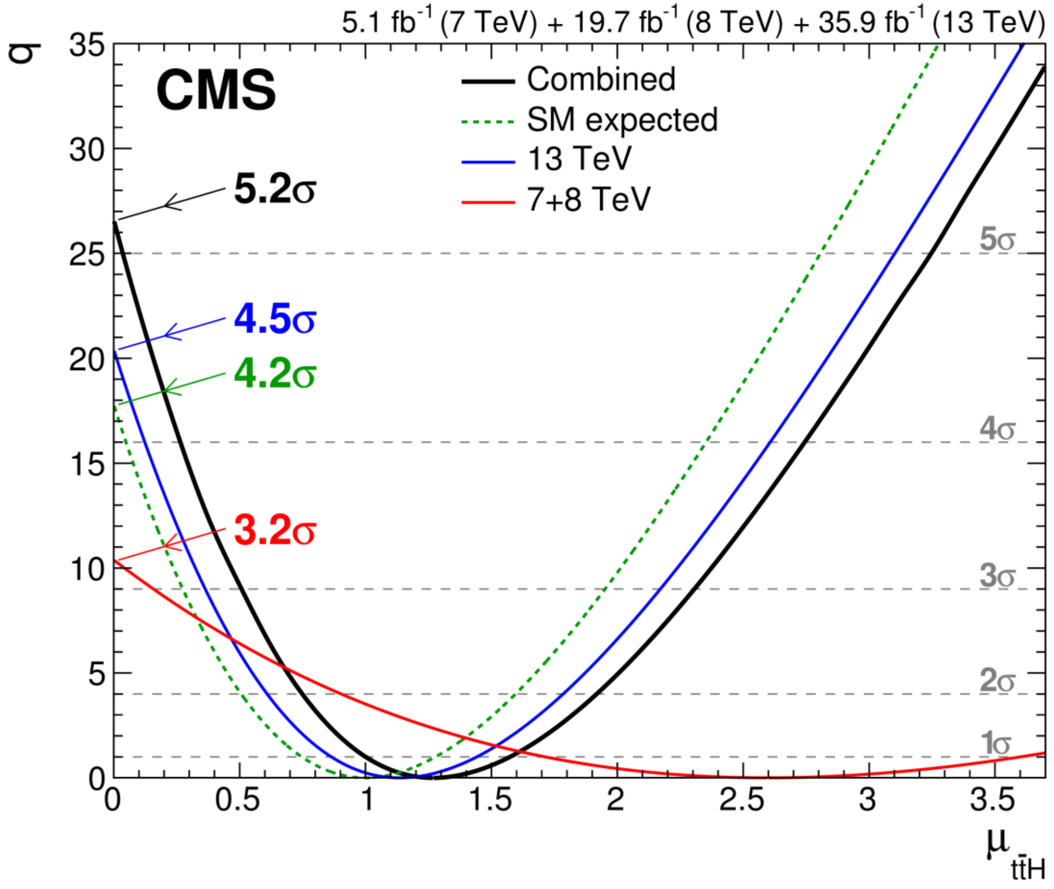
- ✓ In 2018, both ATLAS and CMS announced the observation of $t\bar{t}H$ production with a cross section compatible with the SM expectation;

- ✓ A combined analysis of pp collision data at center-of-mass energies of $\sqrt{s} = 7,\ 8$ and $13\ \mathrm{TeV}$ was required;

- ✓ To maximize sensitivity, the results of independent searches for $t\bar{t}H$ production using different decay channels of the SM Higgs boson were also combined.

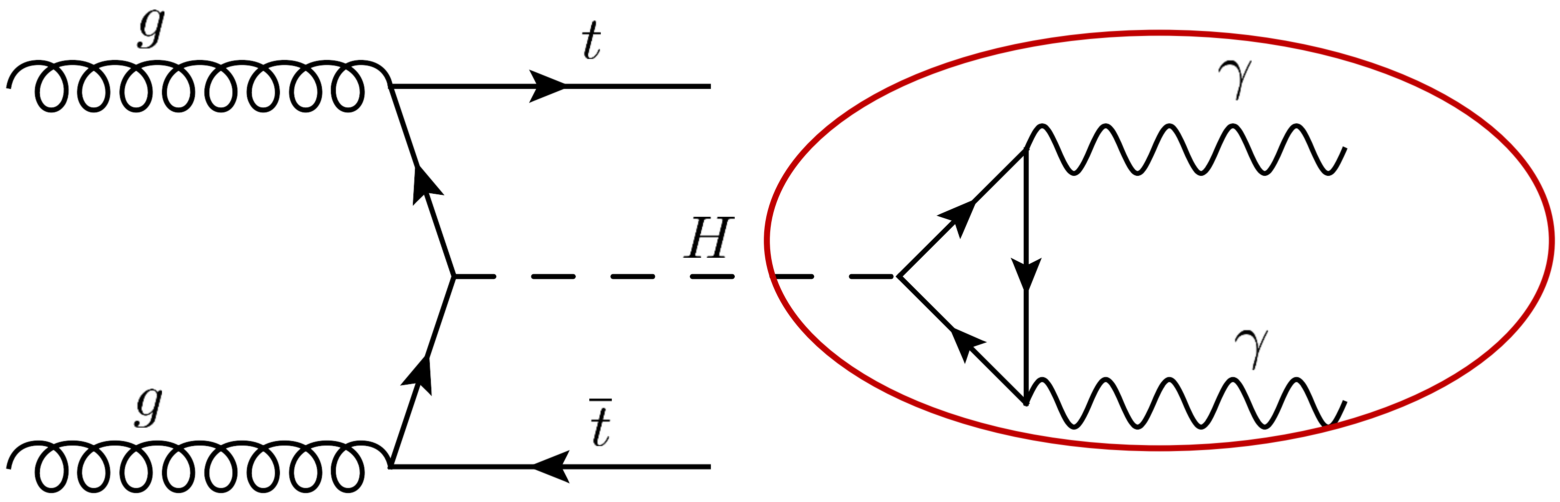
In the paper we are going to analyse, only one decay channel was used:

$$\boldsymbol{H \to \gamma\gamma}$$

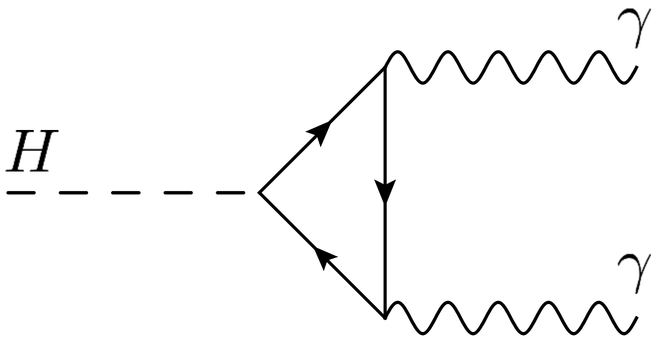


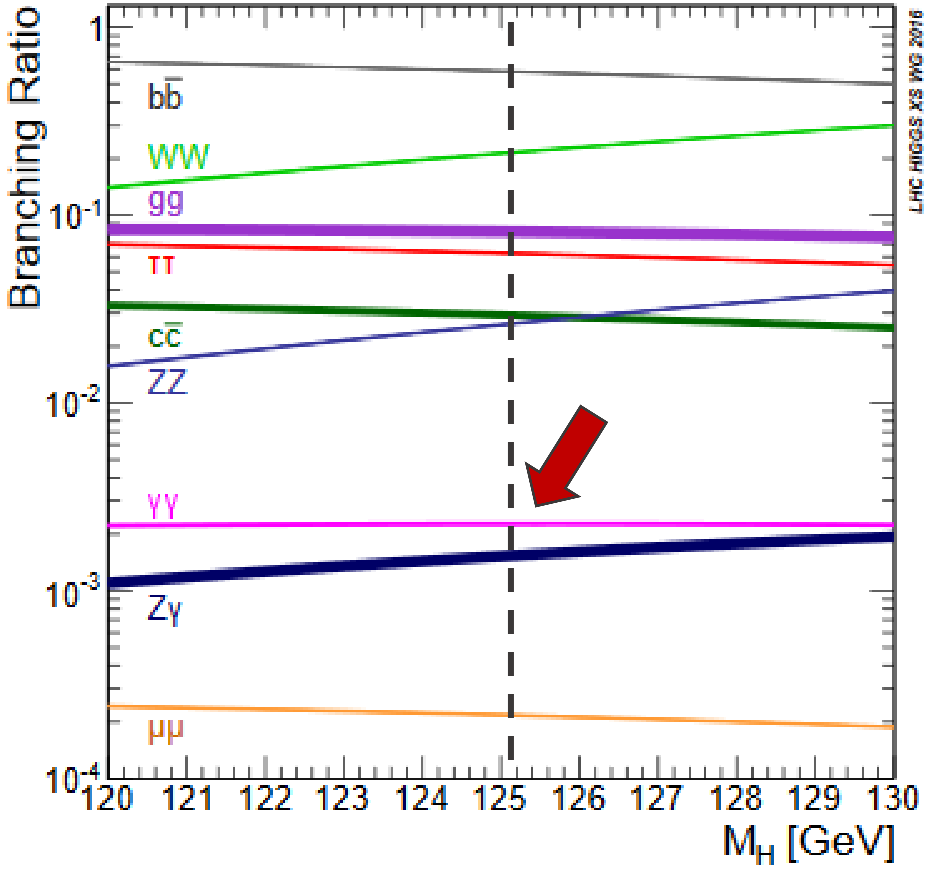
CMS Collaboration: PhysRevLett.120.231801

Introduction



$H \to \gamma\gamma$





- ✓ Despite the small branching fraction predicted by the SM ($\approx 0.2\%$) it was one of the most important channels for the discovery of the Higgs boson;

- ✓ It is possible to measure the energies of photons very precisely, so the mass of a diphoton pair from a Higgs boson can be determined to great precision.

arXiv:2003.10866 [hep-ex]

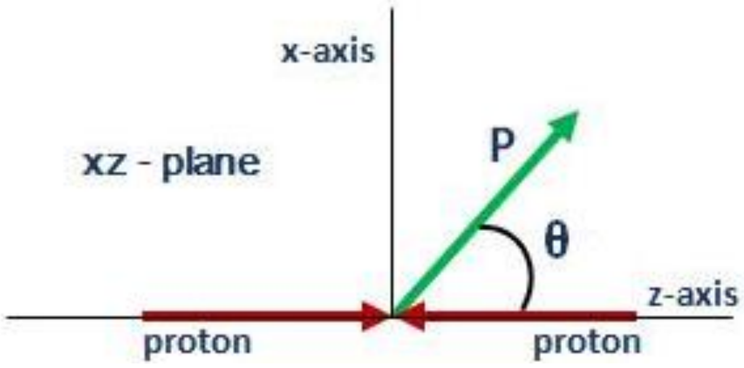
Paper’s presentation

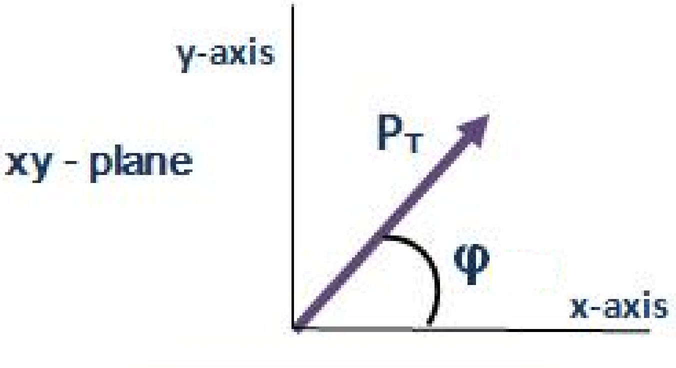
How did experimentalists achieve those results?

Recap: Kinematic variables

✓ Considering the four-vector (E, p_x, p_y, p_z), it is useful to define the following variables:

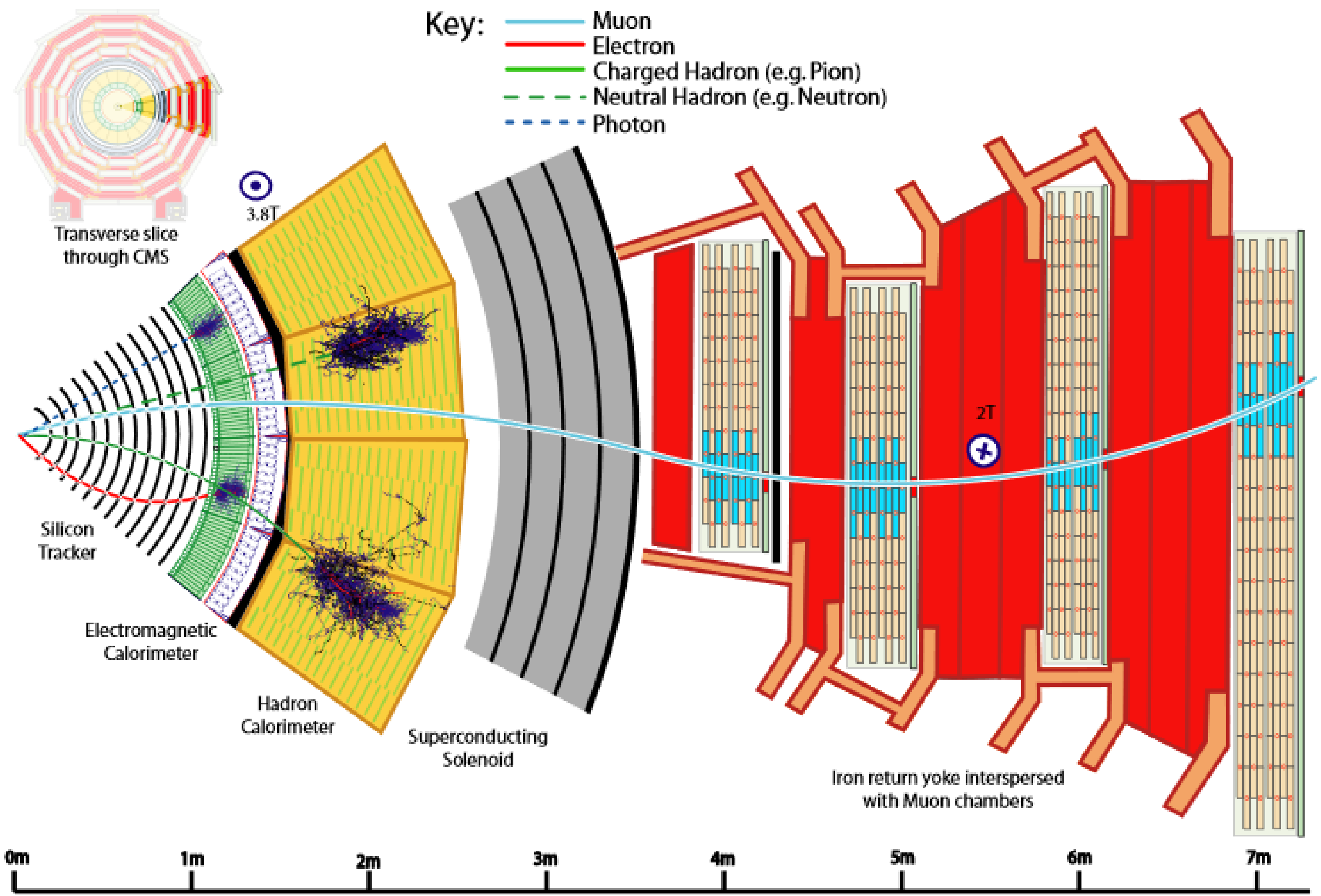
- Transverse momentum: $p_T = \sqrt{p_x^2 + p_y^2}$
- Azimuthal angle: φ
- Rapidity: $y = \frac{1}{2}\log\left(\frac{E+p_z}{E-p_z}\right)$
- Pseudo-rapidity: $\eta = \frac{1}{2}\log\left(\frac{|\vec{p}|+p_z}{|\vec{p}|-p_z}\right) = -\log\left(\tan\frac{\theta}{2}\right)$





Adapted from: https://www.lhc-closer.es/taking_a_closer_look_at_lhc/0.momentum

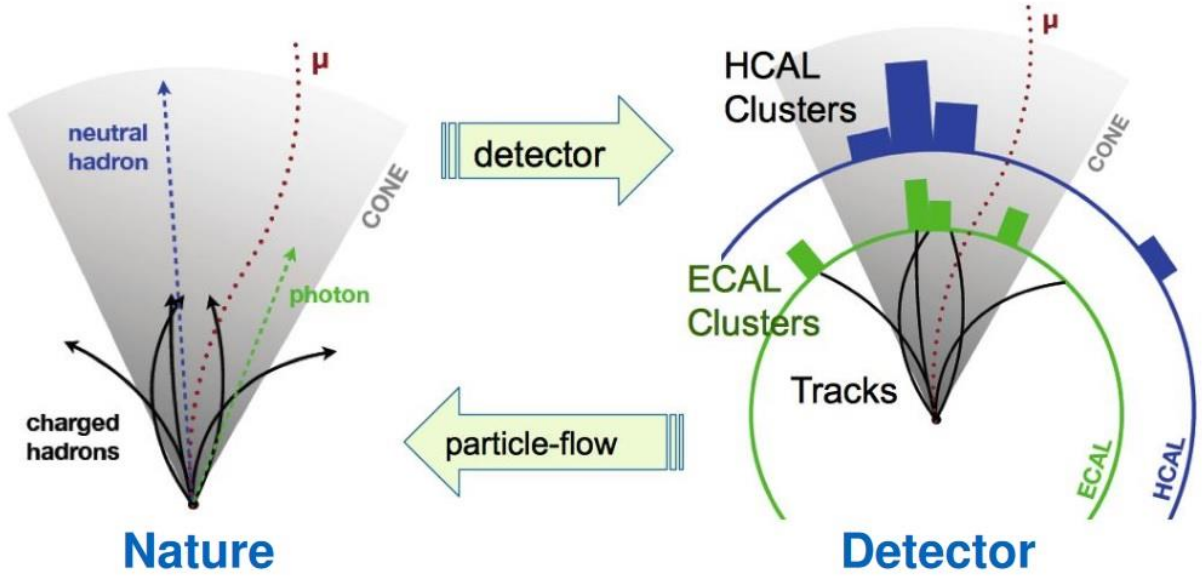
The CMS detector



CMS collaboration, JINST 12 (2017) P10003

Event reconstruction

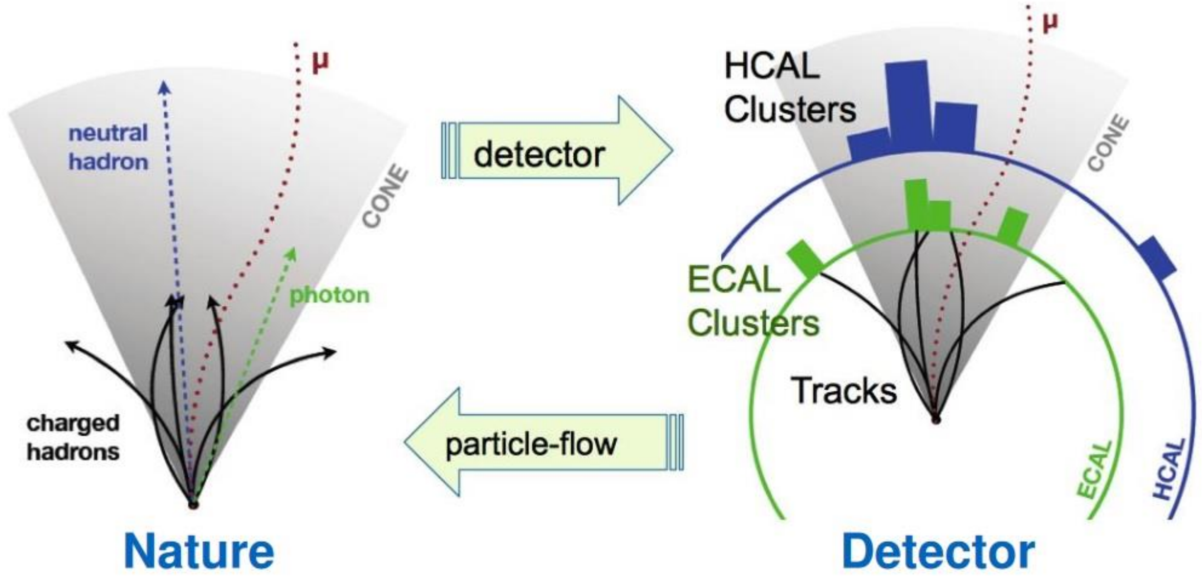
- ✓ The particle-flow (PF) algorithm converts detector signals back to physical objects;
- ✓ The anti-k_t algorithm is used to build jets with a distance parameter of 0.4;
- ✓ The reconstructed vertex with the largest value of summed physics-object p_T^2 corresponds to the primary pp interaction vertex;



https://indico.in2p3.fr/event/13763/contributions/15192/attachments/12667/15553/4_JuskaPekkanen.pdf

Event reconstruction

- ✓ Charged hadrons originating from additional pp interactions are removed from the analysis;
- ✓ Jets from the hadronization of bottom-quarks are tagged by a secondary vertex algorithm, based on the score of a deep neural network (DNN);
- ✓ Higgs boson candidates are built from pairs of photon candidates, which are reconstructed from energy clusters in the electromagnetic calorimeter not linked to charged-particle tracks.



https://indico.in2p3.fr/event/13763/contributions/15192/attachments/12667/15553/4_JuskaPekkanen.pdf

Event generation and selection

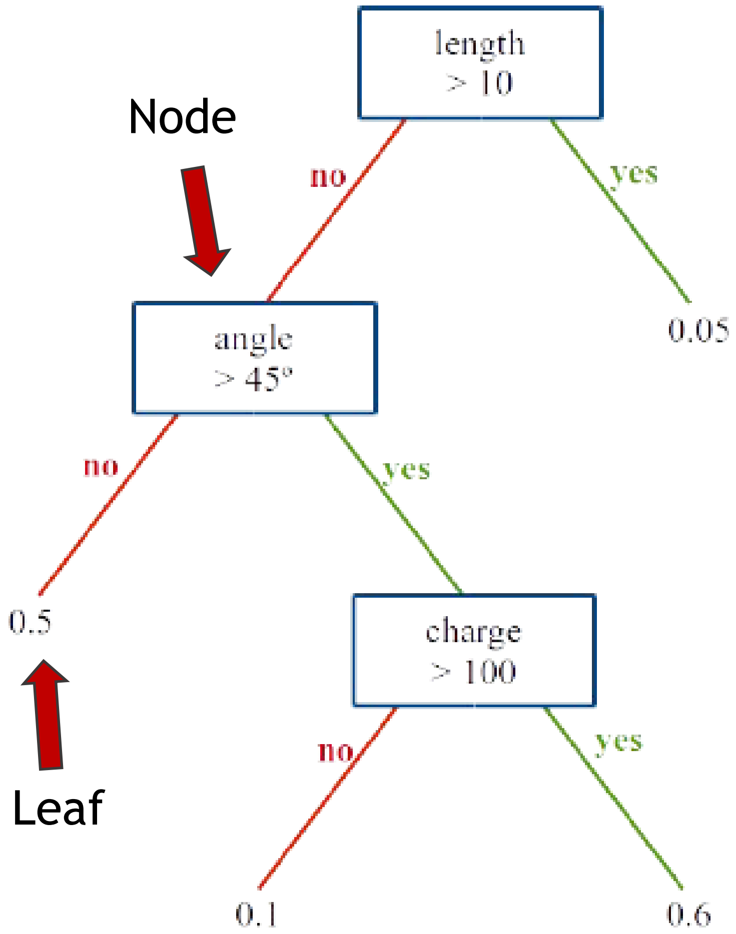
- ✓ Signal and background processes are generated with Monte Carlo programs;
- ✓ The trigger selects diphoton events with a loose calorimetric identification and asymmetric photon transverse energy thresholds of 30 and $18\ (22)\ \mathrm{GeV}$ for the data collected during 2016 (2017/2018);
- ✓ The trigger efficiency is measured to be $> 95\%$.
- ✓ Photons are further required to satisfy an identification criterion based on a boosted decision tree (BDT) classifier trained to separate photons from jets;
- ✓ One requires $100 < m_{\gamma\gamma} < 180\ \mathrm{GeV}$, $p_T/m_{\gamma\gamma} > 1/3$ and $1/4$ for the leading (in p_T) and subleading photons, respectively;

Separating signal from background

- ✓ Events are further divided in two channels:
  - **Leptonic channel**
    - At least one top quark decays leptonically;
    - At least one jet with $p_T > 25\ \mathrm{GeV}$ and $|\eta| < 2.4$;
    - At least one electron (muon) of $p_T > 10\ (5)\ \mathrm{GeV}$ and $|\eta| < 2.4$.
  - **Hadronic channel**
    - At least three jets with one of them b-tagged;
    - No isolated leptons.
- ✓ For each channel of events, a boosted decision tree (BDT) is used to distinguish between signal and background events;
- ✓ The same algorithm will be important later for our CP study;
- ✓ So... **What is a Boosted Decision Tree?**

First, what is a decision tree?

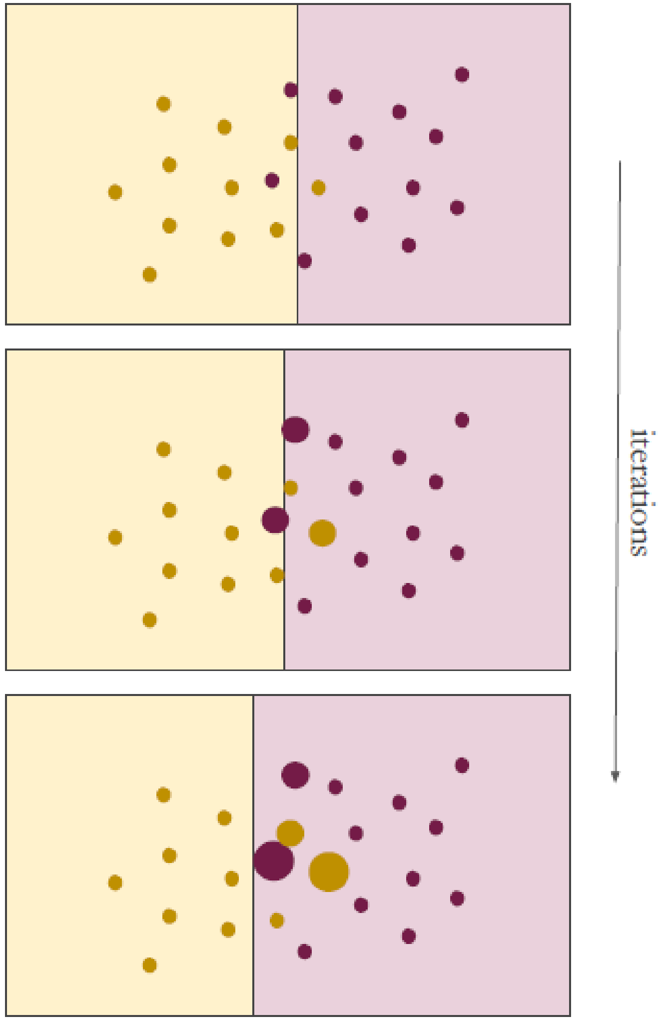
- ✓ A machine learning algorithm which takes a set of input features (length, angle and charge) and splits data recursively based on those features;
- ✓ The data is split based on a value of one of the input features at each node;
- ✓ Each split is chosen to maximize information gain;
- ✓ The splits are created recursively until a stop condition is met (for example, if the maximum depth of tree is reached or if there is no more information gain);
- ✓ Leaves usually represent a probability;



indico.fnal.gov/event/15356/contributions/31377/attachments/19671/24560/DecisionTrees.pdf

Boosted decision tree

- ✓ Boosting is a method of combining weak learners (e.g. trees) into a strong classifier;
- ✓ Each tree is created iteratively;
- ✓ After each iteration, all events are weighted based on its misclassification;
- ✓ The goal is to minimize: $O(x) = \sum_i l(\hat{y}_i, y_i) + \sum_t \Omega(f_t)$
- ✓ $l(\hat{y}_i, y_i)$ is a loss function – distance between the prediction $\hat{y}_i$ and the real value of the i-th sample y_i;
- ✓ $\Omega(f_t)$ is a regularization function – it prevents overfitting by penalizing the complexity of the t-th tree;



indico.fnal.gov/event/15356/contributions/31377/attachments/19671/24560/DecisionTrees.pdf

More about separating signal from background

- ✓ Some input features of the BDT that are used to separate signal from background are: kinematic properties of jets, leptons, photons and diphotons (without $m_{\gamma\gamma}$); jet and lepton multiplicity; b-tagging scores of jets; etc.

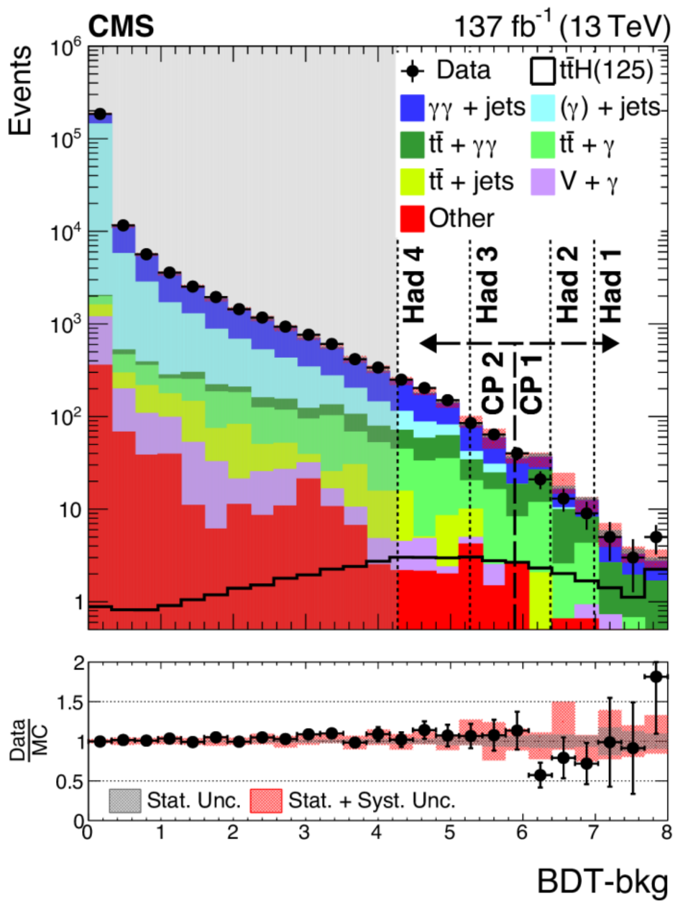
- ✓ BDT's results are enhanced with a DNN that can exploit other features, such as the full four-vectors of each jet;

- ✓ The score of the BDT has been validated by comparing the distributions in data and Monte Carlo simulations in the region $100 < m_{\gamma\gamma} < 120$ GeV or $130 < m_{\gamma\gamma} < 180$ GeV, as well as in dedicated control regions which target $t\bar{t} + Z$.

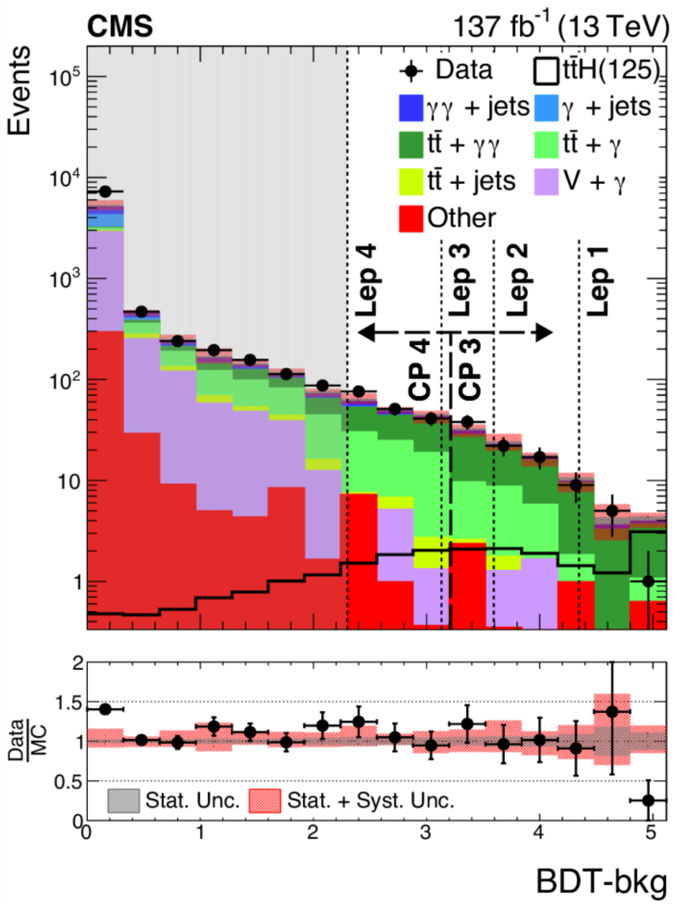
Distribution of BDT's output

- ✓ Events shown are taken from the regions of $m_{\gamma\gamma}$ where the BDT scores are validated;
- ✓ Events are rejected (grey shaded region) or further divided in eight (four) categories to maximize expected significance (sensitivity to CP structure of the Yukawa coupling to top quarks);
- ✓ Statistical (statistical + systematic) background uncertainties are represented by the black (red) shaded bands.

Hadronic channel



Leptonic channel



Distribution of BDT’s output

	Total	$t\bar{t}$H (%)	tH (%)	ggH (%)	VH (%)	VBF (%)	$b\bar{b}$H (%)
Had1	5.8	89.1	6.8	3.3	0.8	<0.1	0.1
Had2	4.2	82.9	6.8	8.7	1.4	0.2	0.1
Had3	11.6	78.6	7.2	10.3	3.5	0.3	0.1
Had4	13.6	65.4	7.7	19.3	6.9	0.7	0.1
Lep1	5.8	90.6	7.9	0.5	1.0	<0.1	<0.1
Lep2	4.9	90.0	6.7	0.4	2.9	<0.1	<0.1
Lep3	3.5	86.2	7.4	0.4	6.0	<0.1	<0.1
Lep4	5.7	78.1	8.2	1.1	12.7	<0.1	<0.1
Total	55.1	79.5	7.4	8.2	4.7	0.3	<0.1

Expected number of Higgs boson events in the hadronic and leptonic channels per category and the fractional contribution per production modes of the Higgs boson.

Measuring $\sigma_{t\bar{t}H}\mathcal{B}_{\gamma\gamma}$ and $\mu_{t\bar{t}H}$

- ✓ A simultaneous binned maximum likelihood fit is performed to the eight categories of events to extract $\sigma_{t\bar{t}H}\mathcal{B}_{\gamma\gamma}$ and $\mu_{t\bar{t}H}$;

- ✓ All other Higgs production modes in the fit are constrained to their SM predictions;

- ✓ The dominant theoretical uncertainty (8%) in $\mu_{t\bar{t}H}$ arises from the SM prediction of $\sigma_{t\bar{t}H}$. The theoretical uncertainties in parton distribution functions, QCD coupling, parton showers and underlying event, and $\mathcal{B}_{\gamma\gamma}$ each affect $\mu_{t\bar{t}H}$ by $2-5\%$;

- ✓ The main experimental uncertainties that affect $\mu_{t\bar{t}H}$, with effects of $2-6\%$, are related to the bottom quark and photon identification, jet energy scale and resolution, and the integrated luminosity;

- ✓ Other systematic uncertainties, such as the trigger efficiency, have a $< 2\%$ effect on measurements of $\sigma_{t\bar{t}H}\mathcal{B}_{\gamma\gamma}$ and $\mu_{t\bar{t}H}$.

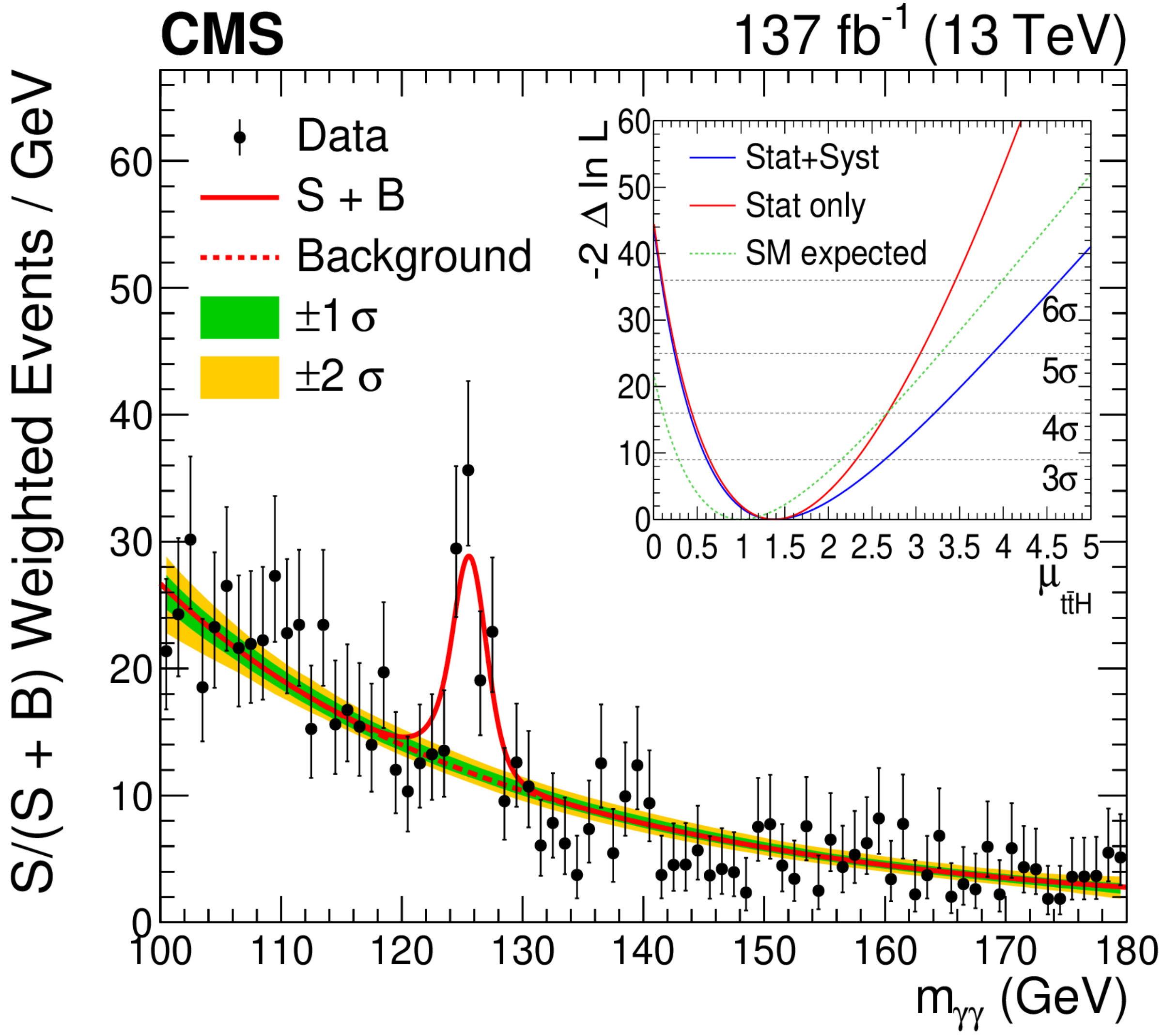
Results

✓ **Data and fit results:**

$$\sigma_{t\bar{t}H}\mathcal{B}_{\gamma\gamma} = 1.56^{+0.34}_{-0.32}\ \text{fb}$$
$$= 1.56^{+0.33}_{-0.30}\ (\text{stat})^{+0.09}_{-0.08}\ (\text{syst})\ \text{fb}$$

✓ **SM prediction:**

$$\sigma_{t\bar{t}H}\mathcal{B}_{\gamma\gamma} = 1.13^{+0.08}_{-0.11}\ \text{fb}$$



Results

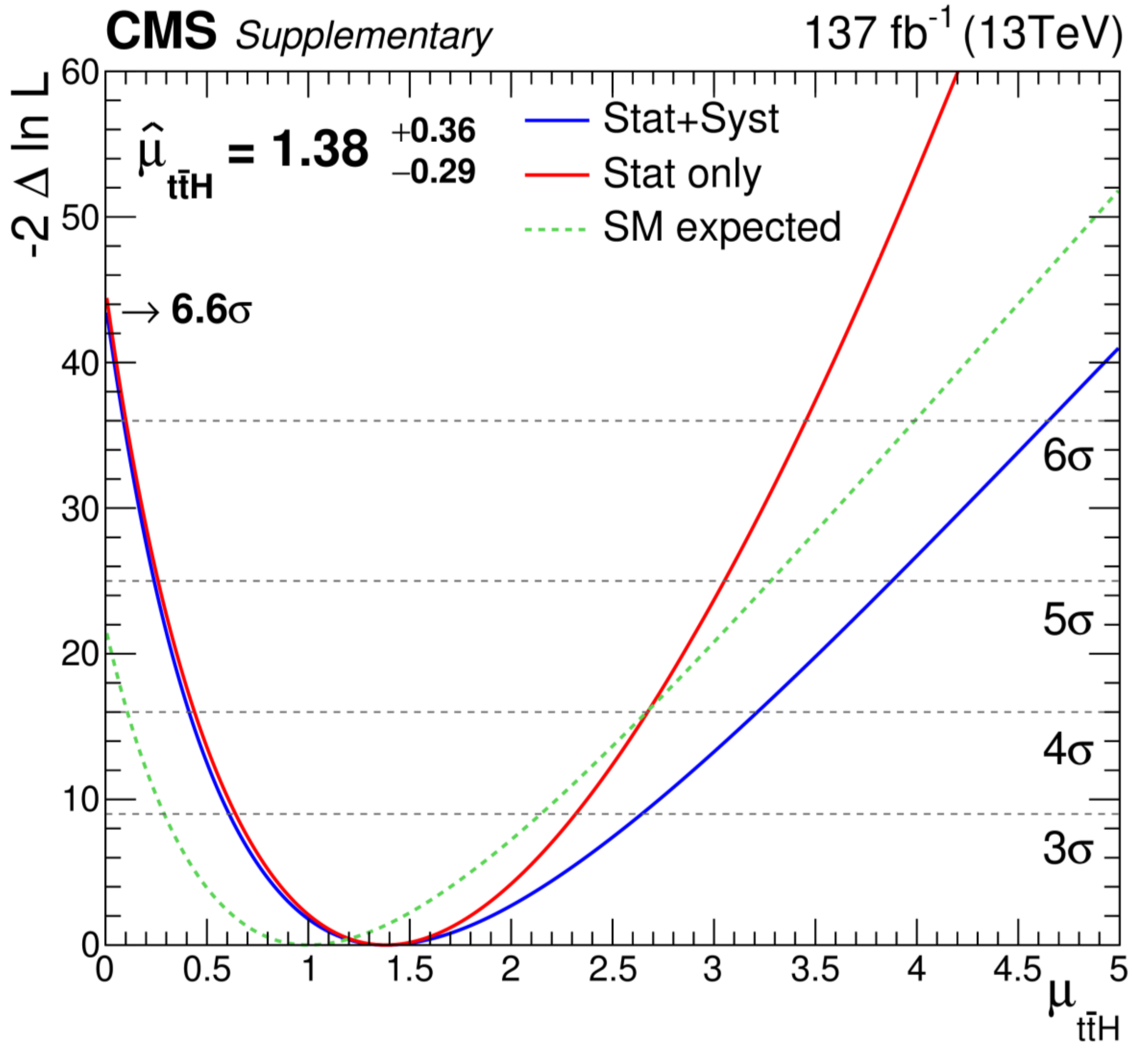
✓ **Signal strength:**

$$\mu_{t\bar{t}H} = \frac{[\sigma_{t\bar{t}H}\mathcal{B}_{\gamma\gamma}]_{\text{data}}}{[\sigma_{t\bar{t}H}\mathcal{B}_{\gamma\gamma}]_{\text{SM}}}$$

$$= 1.38^{+0.36}_{-0.29}\ \text{fb}$$

$$= 1.38^{+0.29}_{-0.27}\ (\text{stat})^{+0.21}_{-0.11}\ (\text{syst})\ \text{fb}$$

✓ The observed significance relative to the background-only hypothesis is 6.6σ, while the expected significance assuming the SM Higgs boson is 4.7σ.



CP structure of the $t\bar{t}H$ coupling

- ✓ The coupling's amplitude can be parametrized as

$$\mathcal{A}(t\bar{t}H) = -\frac{m_t}{v}\overline{\psi}_t\left(\kappa_t + i\widetilde{\kappa}_t\gamma_5\right)\psi_t$$

- ✓ In the SM, $\kappa_t = 1$ and $\widetilde{\kappa}_t = 0$;

- ✓ The CP structure is measured with

$$f_{\rm CP}^{\rm Htt} = \frac{|\widetilde{\kappa}_t|^2}{|\kappa_t|^2+|\widetilde{\kappa}_t|^2}\,{\rm sign}(\widetilde{\kappa}_t/\kappa_t)$$

- ✓ $f_{\rm CP}^{\rm Htt} \neq 0$ ➡ **New Physics!**

$\overline{\psi}_t, \psi_t \rightarrow$ Dirac spinors

$m_t \rightarrow$ Mass of the top quark

$v \rightarrow$ Vacuum expectation value

$\kappa_t \rightarrow$ CP-even Yukawa coupling

$\widetilde{\kappa}_t \rightarrow$ CP-odd Yukawa coupling

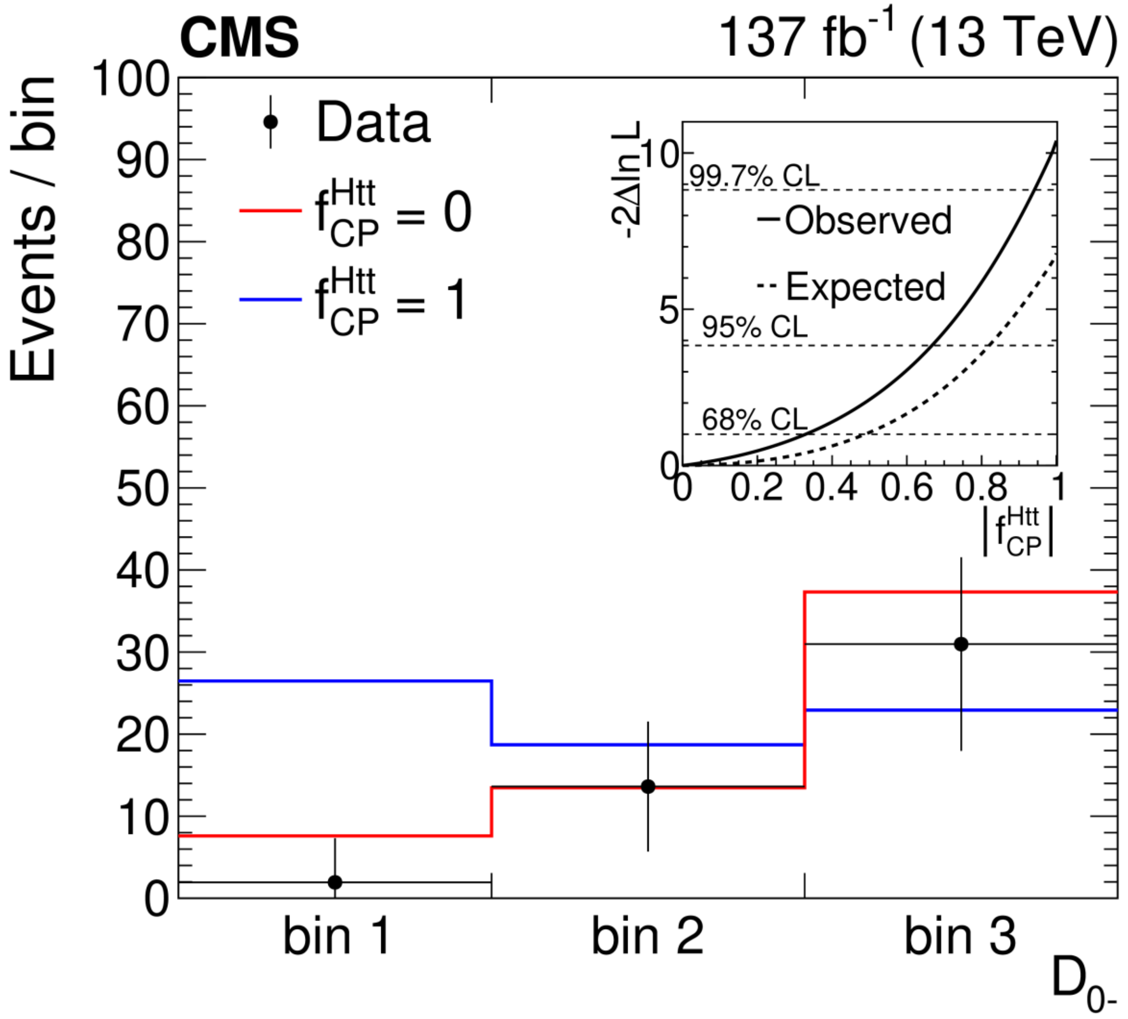
arXiv:2003.10866 [hep-ex]

Analysis of the CP structure of the $t\bar{t}H$ process

- ✓ An optimal analysis of the CP structure in the $t\bar{t}H$ process can be performed with two observables:
  - $\mathcal{D}_{0-} \rightarrow$ optimal to separate the amplitudes squared representing scalar and pseudoscalar contributions;
  - $\mathcal{D}_{\mathrm{CP}} \rightarrow$ optimal to separate interference between scalar and pseudoscalar amplitudes;
- ✓ A BDT is employed to distinguish CP-even and CP-odd contributions. Its output is just $\mathcal{D}_{0-}$.
- ✓ The flavour of light jets is not tagged. Then, we cannot get $\mathcal{D}_{\mathrm{CP}}$ and it is not possible to measure the relative sign between κ_t and $\widetilde{\kappa}_t$;
- ✓ Simulation shows that $\mathcal{D}_{0-}$ has negligible correlation with the BDT- background discriminant;
- ✓ The $\mathcal{D}_{0-}$ discriminant is included in the measurement of $f_{\mathrm{CP}}^{\mathrm{Htt}}$, which is dominated by statistical uncertainties.

More results

- ✓ The leptonic and hadronic channels and the categories of the BDT- background discriminant are combined in the mass range $115 < m_{\gamma\gamma} < 135\ \text{GeV}$.



More results

- **Fit results:**

$$f_{CP}^{t\bar{t}H} = 0.00 \pm 0.33\,, \ \ |f_{CP}^{t\bar{t}H}| < 0.67 \ (95\% \ \mathrm{CL})$$

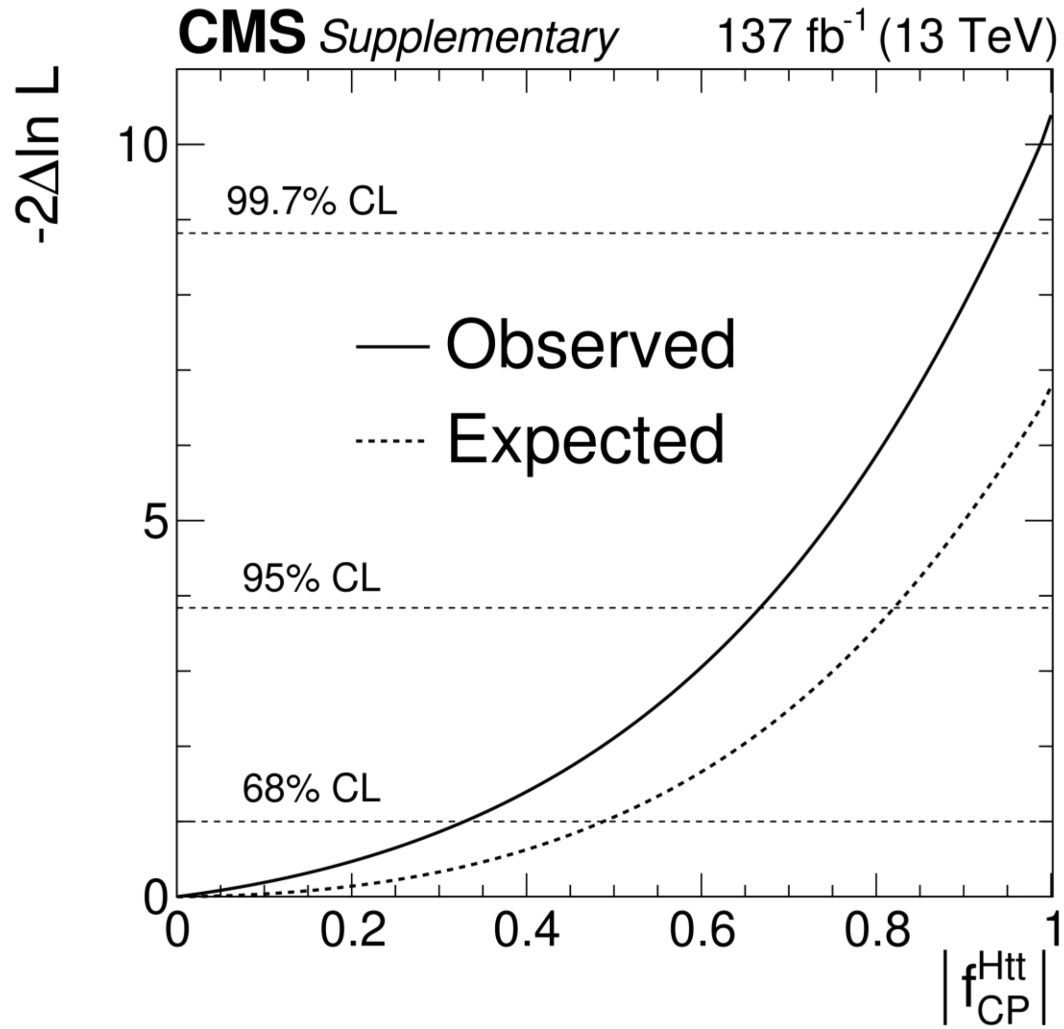
The pure pseudoscalar model of CP structure of the $t\bar{t}H$ coupling is excluded at 3.2σ.

- **Expected contraints:**

$$f_{CP}^{t\bar{t}H} = 0.00 \pm 0.49 \ (68\% \ \mathrm{CL})$$

$$|f_{CP}^{t\bar{t}H}| < 0.82 \ (95\% \ \mathrm{CL})$$

2.6σ exclusion of the $f_{CP}^{t\bar{t}H} = 1$ model.



Conclusions

- ✓ Observing the production of the Higgs boson alongside a top quark pair is crucial in order to probe the $t\bar{t}H$ Yukawa coupling;

- ✓ The paper reports the first observation of $t\bar{t}H$ production in a single decay channel ($H \to \gamma\gamma$) and the first measurement of the CP structure of the $t\bar{t}H$ coupling;

- ✓ The product of the $t\bar{t}H$ cross section $\sigma_{t\bar{t}H}$ with the $H \to \gamma\gamma$ branching fraction $\mathcal{B}_{\gamma\gamma}$ is measured to be $\sigma_{t\bar{t}H}\mathcal{B}_{\gamma\gamma} = 1.56^{+0.34}_{-0.32}$ fb, corresponding to a signal strength $\mu_{t\bar{t}H} = 1.38^{+0.36}_{-0.29}$ fb, with a significance of 6.6σ;

- ✓ Data disfavour the CP-odd model of the $t\bar{t}H$ coupling at 3.2σ and a fractional CP-odd contribution is constrained to be $|f_{\mathrm{CP}}^{t\bar{t}H}| < 0.67$ at 95% CL;

- ✓ All results are compatible with the SM predictions.

Questions?

Thank you!