Study of Higgs boson pair production in the HH→bbττ channel with CMS Run 3 data

Elvira Martín Viscasillas

CMS

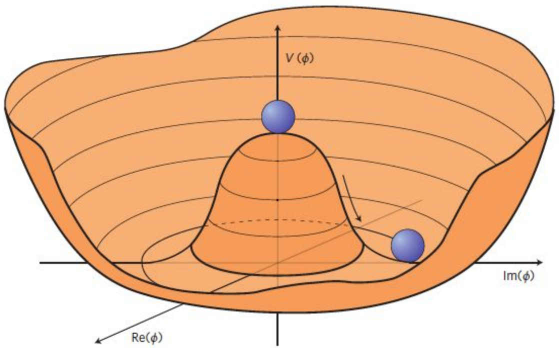
GOBIERNO DE ESPAÑA

MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES

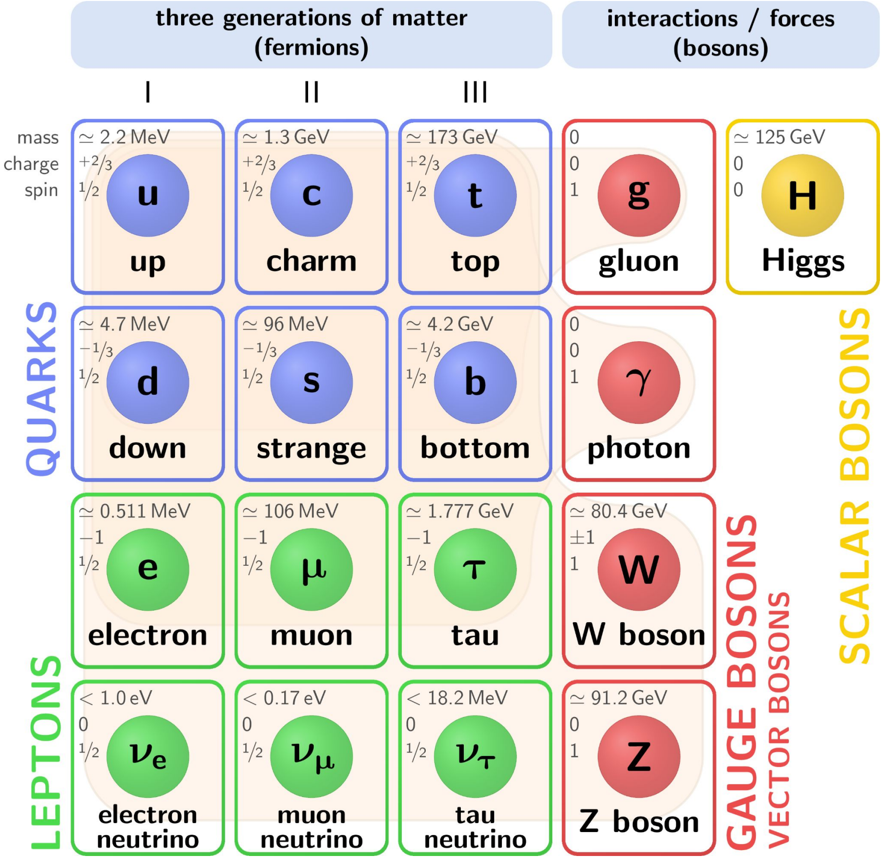
Ciemat
Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

Theoretical context

- The **Standard Model** of particle physics describes all elementary particles and three of the four fundamental interactions
- The electroweak interaction predicts the existence of **massless** $W\pm$ and Z bosons → Higgs mechanism explains the mass terms of gauge bosons through **electroweak symmetry breaking (EWSB)**
- The **Higgs potential** shape is the key of the **EWSB**



$$V(H) = \frac{1}{2} m_H^2 H^2 + \lambda_2 v H^3 + \frac{1}{4} \lambda_4 H^4$$



Exploring the Higgs Potential

- Two parameters describing the shape: Higgs boson **mass** and **self-interactions**
- **Higgs self-couplings not measured yet** ➔ Direct access via Di-Higgs production

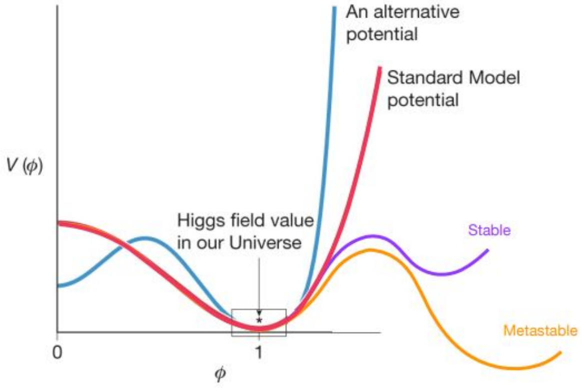
Mass term | **λ_3 ➔ HH production** | **λ_4 ➔ HHH production**

$$V(H) = \frac{1}{2}m_H^2 H^2 + \lambda_3 v H^3 + \frac{1}{4}\lambda_4 H^4$$

- In the **SM:** $\lambda_3 = \lambda_4$ and fully determined given m_H and VEV v

$$\lambda_3 = \frac{m_H^2}{2v^2} \approx 0.13$$

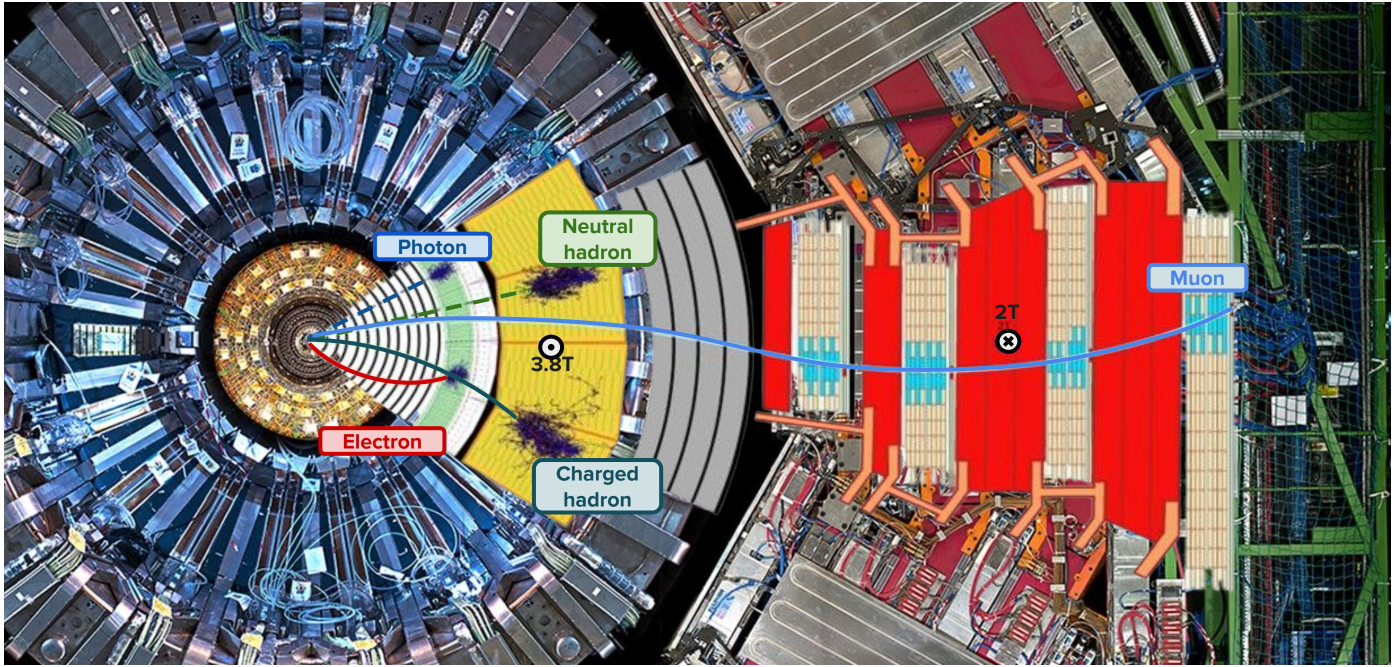
- **New physics** can affect the shape of the Higgs potential



Higgs self-interactions need to be measured as a powerful closure test of the EWSB and the SM

The LHC and the CMS experiment

- The **LHC** is a 27 km circular collider located at CERN
- **The Compact Muon Solenoid (CMS)** is one of the two general purpose experiments at the **LHC**
  - Cylindrical geometry detector composed by layers of subdetectors
  - Superconducting solenoid providing a magnetic field of 3.8T



Particles leave a signature in the subdetectors

Information coming from subdetectors is combined by Particle Flow algorithm

Reconstruction of the physics objects of the analysis (b-jets, τ...)

Di-Higgs production at LHC

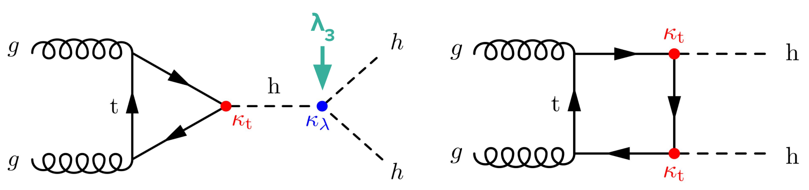
- Measurement of Higgs self-coupling is a main goal at the LHC
- The Di-Higgs cross section in the SM it's **~1000 times smaller** than Single-Higgs
- **New physics effects** can modify HH production rates and kinematics

Non-Resonant HH production (SM & BSM)

- **SM** HH production mechanisms ➔ **ggF** and **VBF** main production modes
- **BSM** physics effects parametrized by coupling modifiers (**$\kappa=\lambda/\lambda^{SM}$**): **$\kappa_\lambda$** , **$\kappa_t$**, **$\kappa_V$**, **$\kappa_{2V}$**

Gluon-Gluon Fusion

σ_{ggF} = 31.05 fb @ NNLO

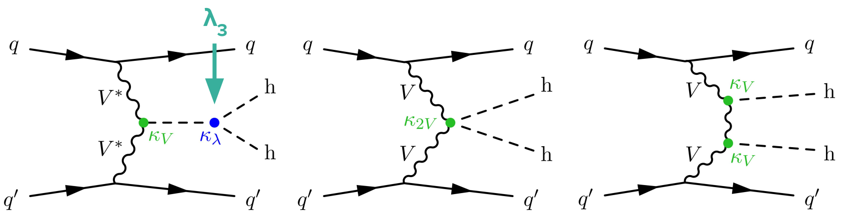


Destructive interference between triangle and box diagrams

➔ **Direct access to κ_λ (main sensitivity)**

Vector Boson Fusion

σ_{VBF} = 1.73 fb @ N3LO



Signature: 2 jets with high m_{jj} and large separation in η (VBF jets)

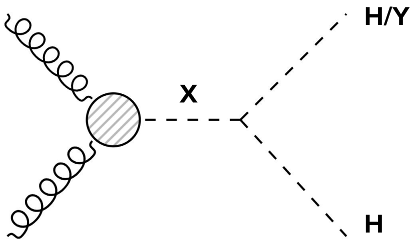
➔ **Direct access to κ_λ and κ_{2V}**

Di-Higgs production at LHC

- Measurement of Higgs self-coupling is a main goal at the LHC
- The Di-Higgs cross section in the SM it's **~1000 times smaller** than Single-Higgs
- **New physics effects** can modify HH production rates and kinematics

Resonant HH production (BSM)

- Higgs pair produced from heavy resonance X



- BSM physics effects parametrized by heavy resonance mass m_X

Rare production rate and complex signal kinematics → Experimental challenges

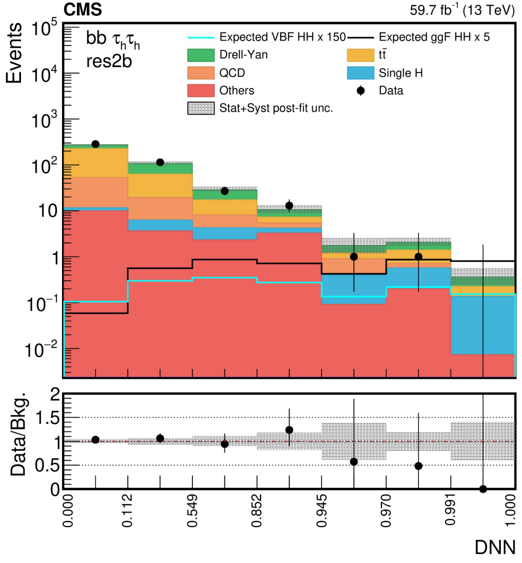
HH→bbττ analysis

HH→bbττ : sizeable branching ratio (58%×6%), lower background contamination → **One of the most sensitive channels**

Jaime's Thesis

Run-2 results:

- **Non-resonant ggF + VBF** Phys. Lett. B 842.137531
  - Inclusive (ggF + VBF) upper limit on signal strength:
    $\sigma^{HH} < 3.3 \times \sigma^{HH}(SM)$ @ 95% CL
  - VBF upper limit on signal strength: **$\sigma^{HH} < 124 \times \sigma^{HH}(SM)$ @ 95% CL**
  - Self coupling constrained to: **$-1.7 < \kappa_\lambda < 8.7$ (other couplings = 1) @ 95% CL**
  - C2V coupling (HHVV) constrained to:
    $-0.4 < \kappa_{2V} < 2.6$ (other couplings = 1) @ 95% CL
- **Resonant X→YH** JHEP 11 (2021) 057
- **Ongoing resonant X→HH** B2G-24-011 (not public yet)
  - With contributions of mine that I will highlight during the presentation



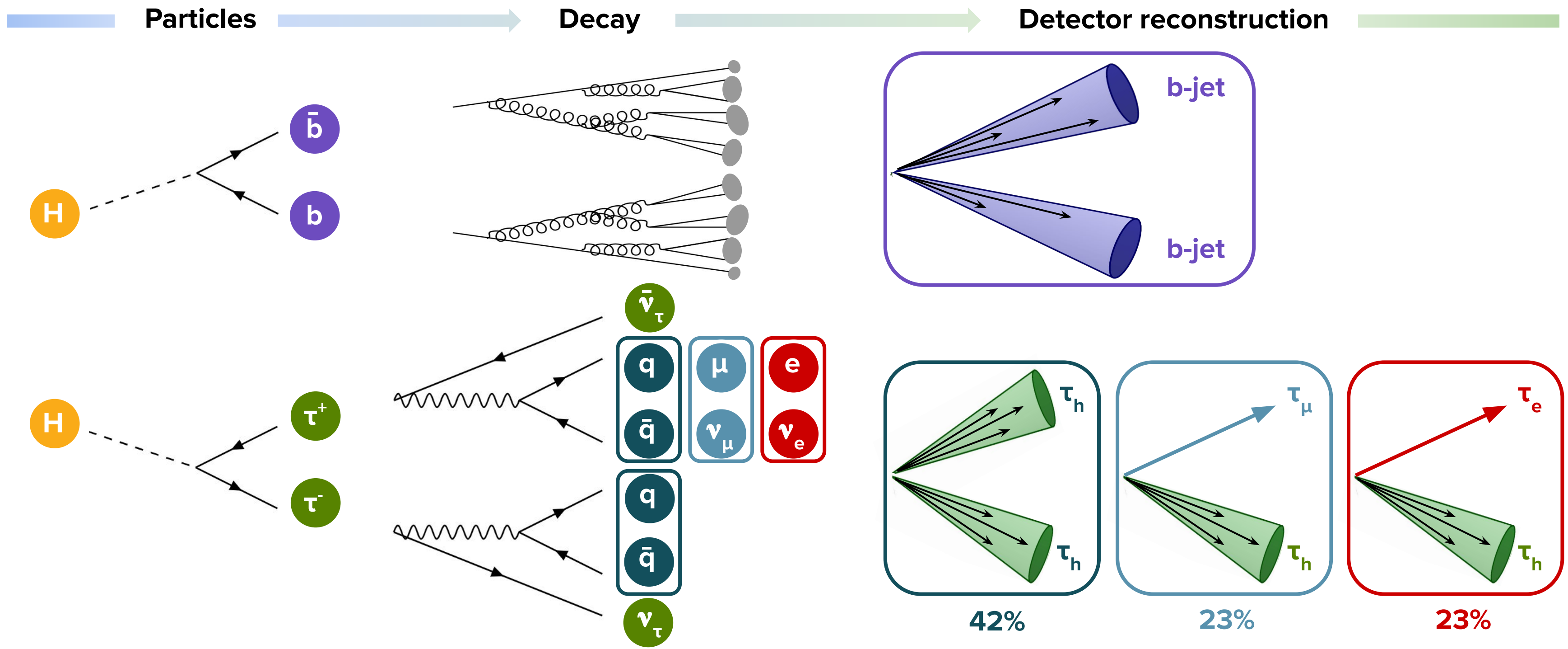
Current Run 3 effort

Comprehensive effort → working group collaboration for bbττ analysis: **UHH - LLR - UZH - KIT ...**

- **Early Run 3 publication** following the Run 2 strategy: non-resonant **2022&2023 analysis**, combined with Run 2
- Long term: **full Run 3** non-resonant & resonant analyses with re-analysis of Run 2 data

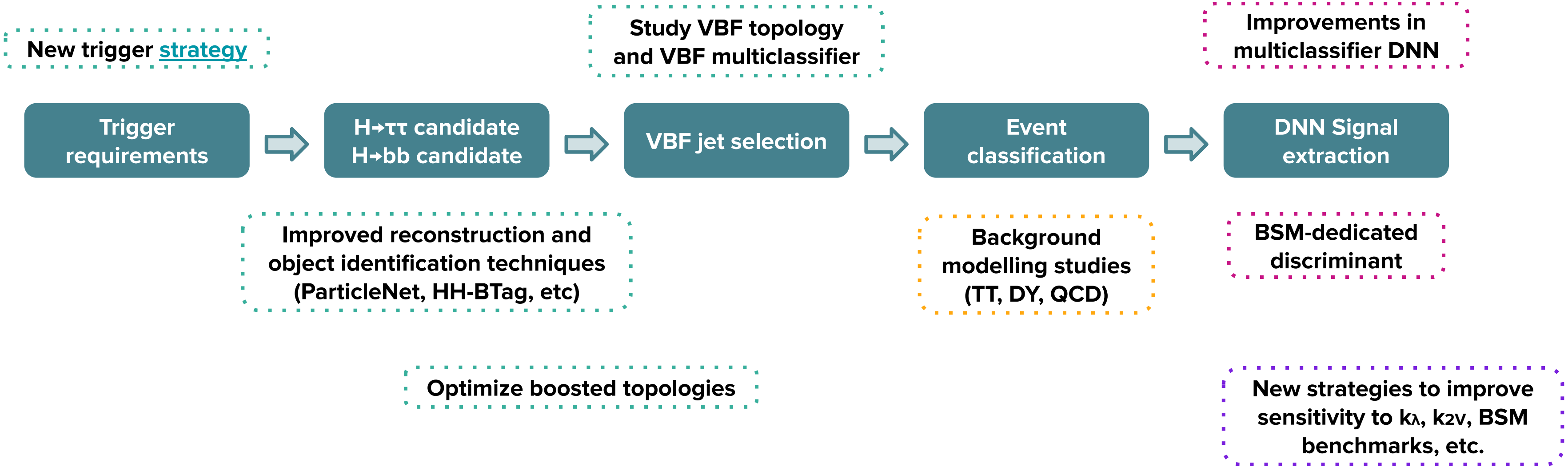
HH→bbττ final state

- Good compromise between branching ratio (58%×6%) and clean final state
- Searching for **Higgs boson pair production** decaying into 2 b quarks and 2 τ leptons



HH→bbττ analysis flow

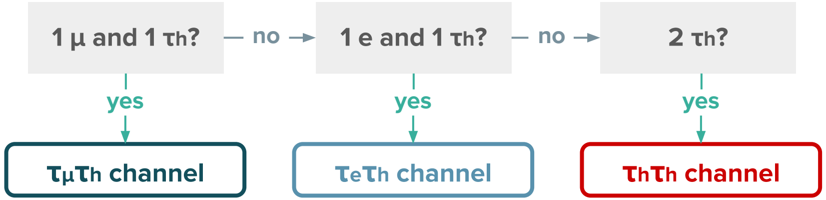
- Basis set relying on Run 2 analyses
- Many developments and **new ideas to optimize** the searches taking advantage of the **Run 3 improvements**:
  selections - **background modelling** - **signal extraction** - **interpretations**



$H\to\tau\tau$, $H\to bb$ and VBF jets candidates

$H\to\tau\tau$

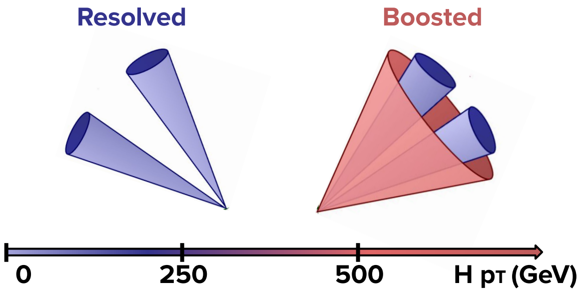
- Physics objects preselection and identification → Reconstruction of the lepton objects
- Three channels according to the final state:



- Kinematic selection depending on the **trigger strategy** followed (p_T thresholds)
- Machine learning techniques (SVFit, FastMTT) to reconstruct the ττ system

$H\to bb$

- b jet candidates are sorted using the **HH-btag algorithm** → The two jets with the highest score are selected as the $H\to bb$ candidate
- Two regimes considered

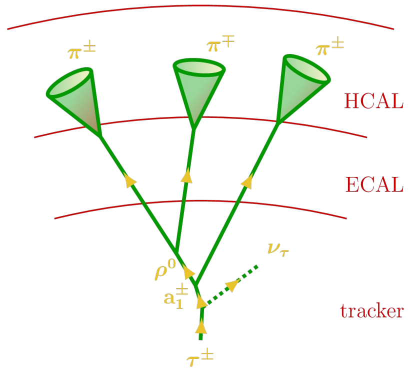


VBF jets

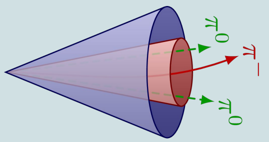
- VBF jets → among all jets not identified as the two b jets
- Two VBF jet candidates selected as the ones that, combined, give the highest invariant mass among all possible combinations.

H➔ττ: τ lepton objects

- Importance of accurate reconstruction and identification of analysis objects
- **τ lepton** decay almost immediately after its production



Hadronic decay τ_h



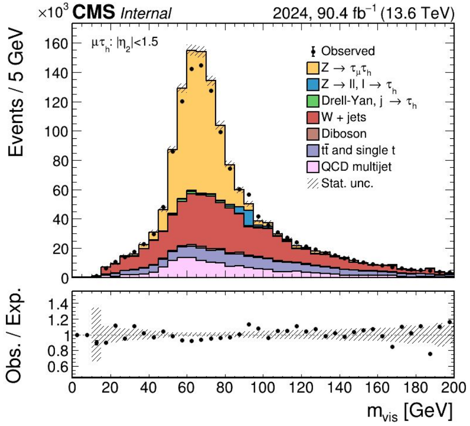
τ_h reconstructed as a composite object

- **Reconstruction** based on Particle Flow
- **Identification**: **DeepTau** ➔ Three different discriminators to identify τ_h against jets, electrons and muons

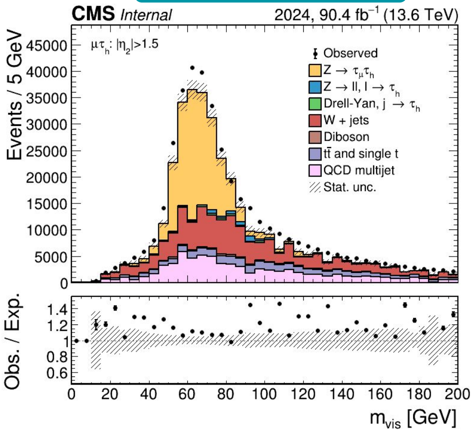
2024 Data/MC τ validation

- **As part of my contribution for CMS** : Validation work performed within the Tau Physics Object Groups (POG) ➔ τ lepton object performance for 2024 data
  - Studied for a well-understood sample: Z➔ττ in a well-understood decay selection: τμτh
- Scale factors to be provided to account for the data/MC discrepancy by the Tau POG

|η|(τ)<1.5



|η|(τ)>1.5



H→bb: The HH-btag algorithm

- **HH-BTag:** Neural network based method to improve the selection of b-jets from **H→bb decay in the bbττ final state** di-Higgs searches

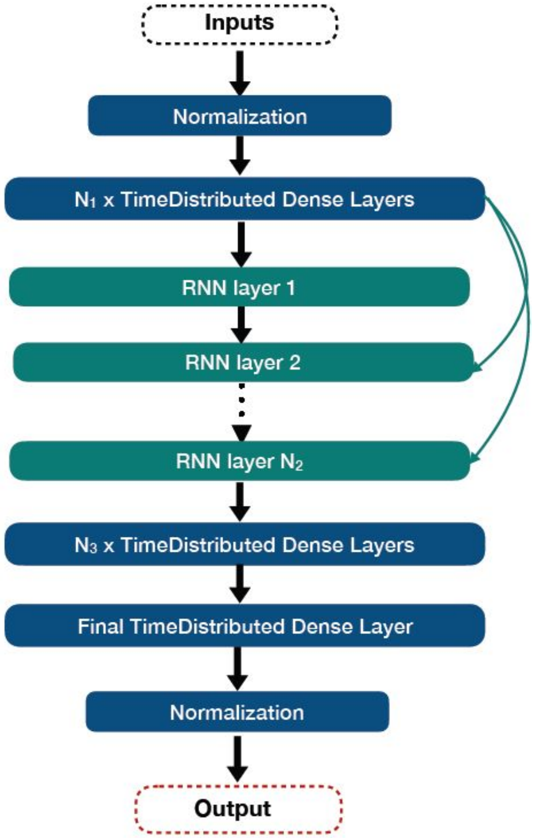
- Input information:
    - Score of the b-jet candidates given by the corresponding tagger
    - Kinematic variables of the b-jet candidate and the Hττ candidate
    - Global event variables

  → **Possible expansion to other di-Higgs final states**

- Developed and used in **bbττ Run 2 Non-Resonant analysis**, improving H→bb selection efficiency and mass resolution w.r.t. CMS standard b-tagging

- Retraining carried out targeting ongoing **bbττ Run 2 resonant analysis →** Performance shows improvements w.r.t. previous training

- **First version for Run 3** already developed



H→bb: The HH-btag algorithm

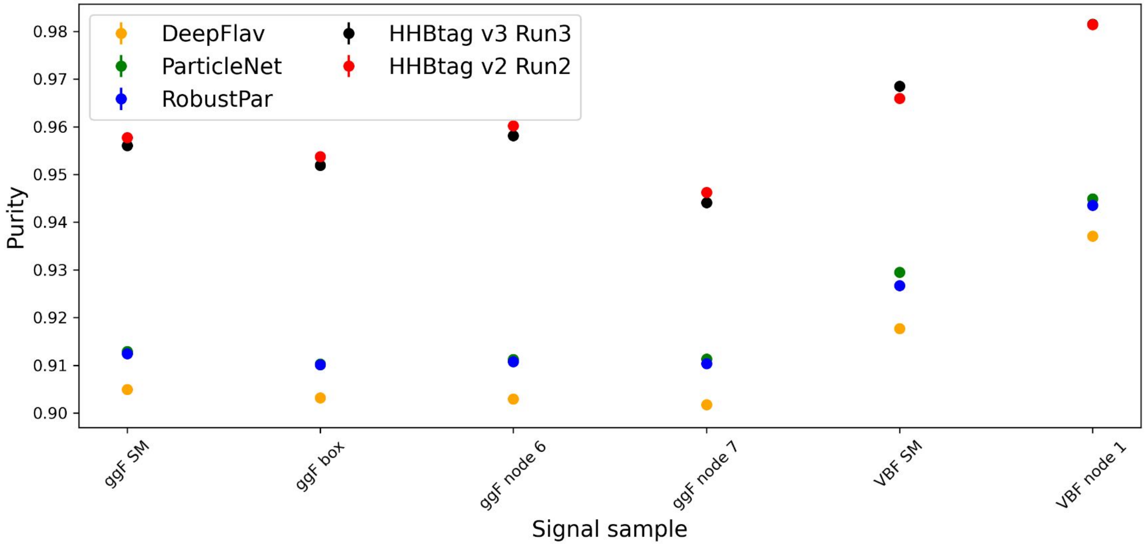
Metric to evaluate the tagger performance: **Purity**

Purity in the H→bb candidate selection obtained via:

- CMS standard b-taggers: **DeepFlav, ParticleNet** and **RobustParT**
- **HH-BTag v2**: HH-BTag training using UL Run 2 signal samples carried out for the ongoing Run 2 Resonant analysis
- **HH-BTag v3**: First version for Run 3 using 2022&2023 available signal samples

$$\textbf{Purity(classifier)} = \frac{N_{true}(\text{classifier})}{N(\text{classifier})}$$

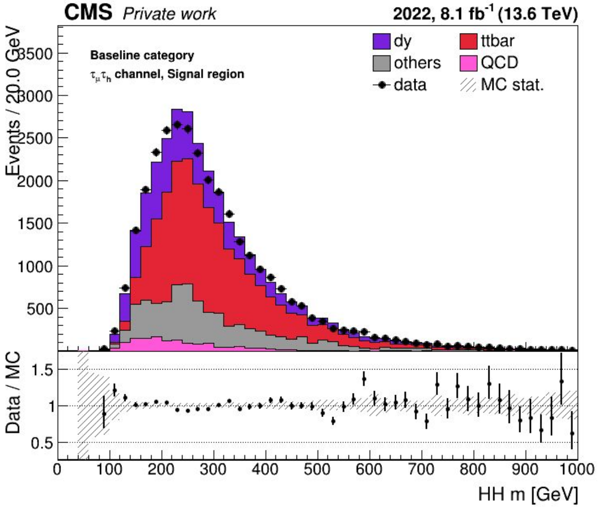
Purity results



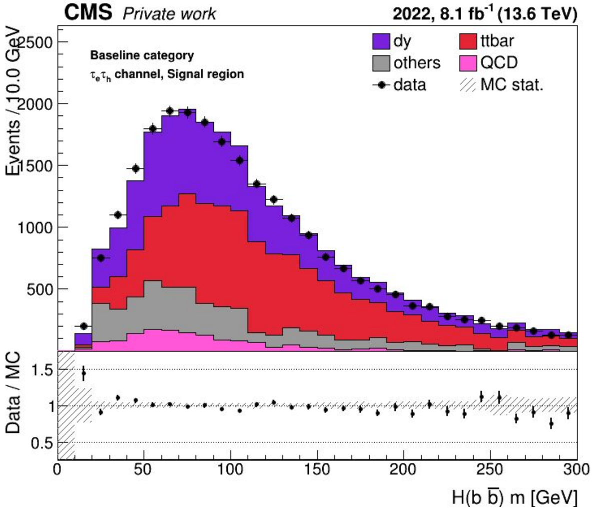
Baseline distributions

- **Baseline**: Events with a ττ and a bb pair candidates
- Distributions of the HH, H→bb, and H→ττ system masses after the baseline selection in the $\tau_\mu\tau_h$, $\tau_e\tau_h$ and $\tau_h\tau_h$ channels
- Results for a reduced dataset → cooperative work to produce results for the whole dataset
- Corrections to simulated events applied

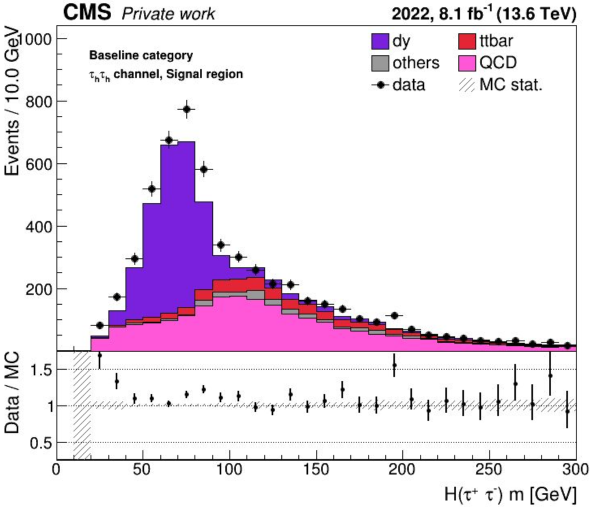
HH mass, $\tau_\mu\tau_h$



Hbb mass, $\tau_e\tau_h$



Hττ mass, $\tau_h\tau_h$

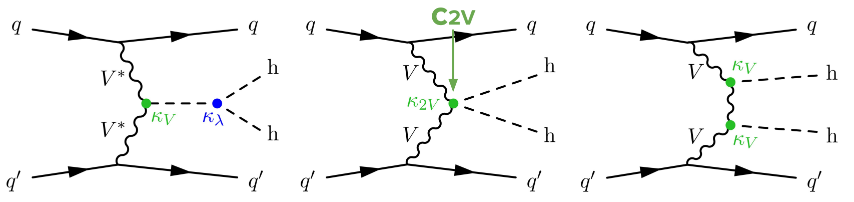


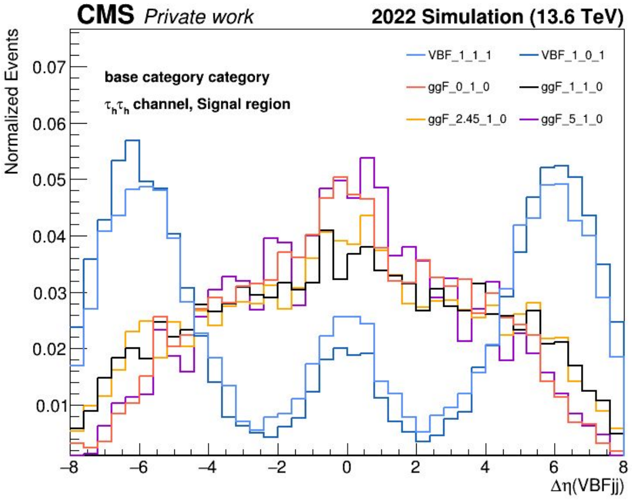
VBF topology

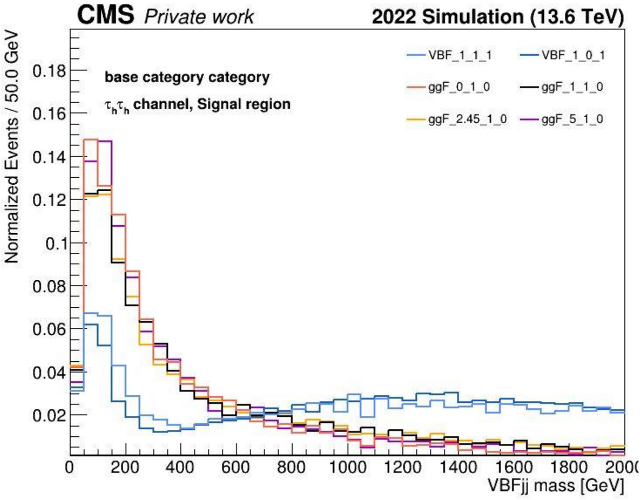
Ongoing effort to improve VBF analysis strategy

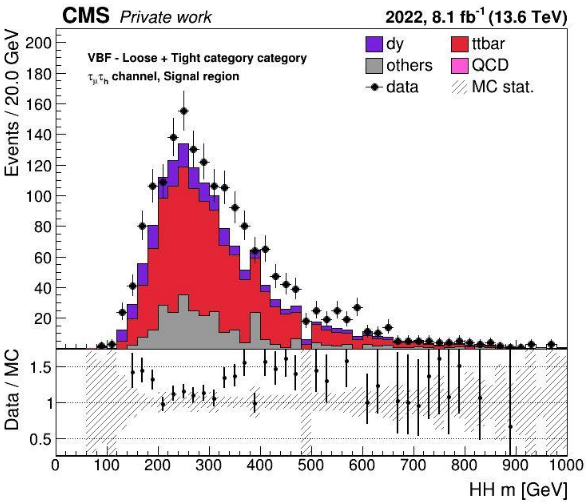
- Discriminant variables of the VBF topology:
  - Δη(jj): separation in pseudorapidity of the VBF jets
  - mjj: invariant mass of the VBF jets
- Optimization of VBF categorization
  - Current selection: mjj > 500 GeV & Δη(jj) > 3
- VBF signal extraction strategy

Unique handle to access **C2V** (VVHH coupling)





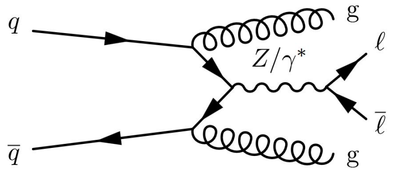




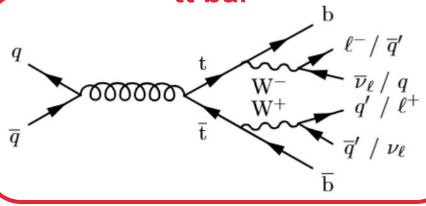
Background modelling

- All processes are obtained from MC simulation, except the QCD multijet background, estimated from data-driven method

Drell-Yan

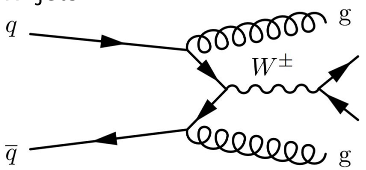


tt bar



Others

→ W+jets

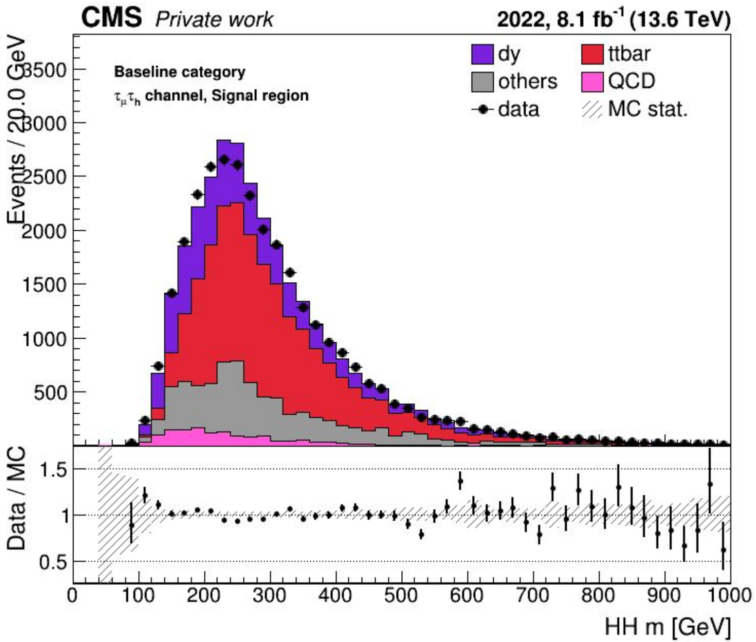


→ Single t, Di boson, Triboson, H, ...

Multi-jet QCD

→ **Data-driven** estimation in the signal region using ABCD method

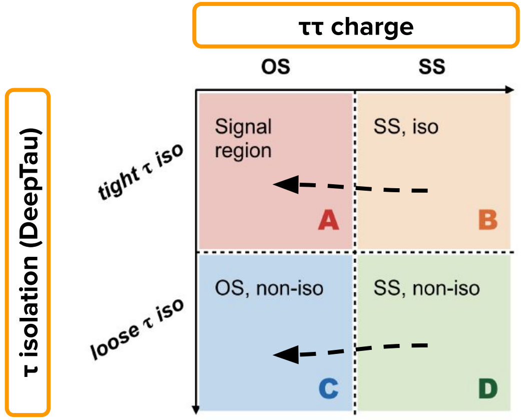
→ Two jets are misidentified as leptons



QCD estimation

- **QCD background** particularly important for $\tau_h\tau_h$ final state → **Data-driven** estimation in the signal region using **ABCD method**
- Subleading background, but one of the main ones in the **impact on the sensitivity** → Importance of developing studies on the **validity of the method**

ABCD method



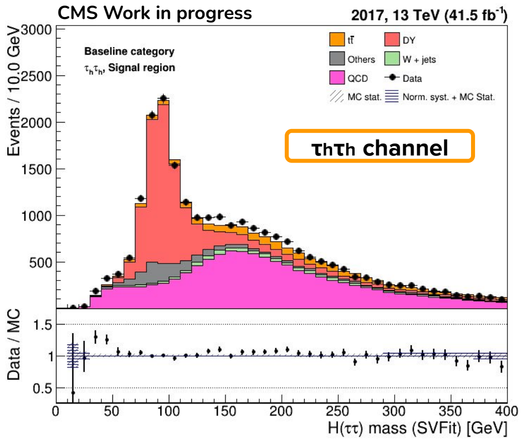
Yield in the signal region: A = C×B/D

Shape from C

Validity of the method

Several tests performed to study the validity and stability of the ABCD method

- Tested robustness of the method for different isolation definitions
- Tested on highly populated QCD regions



Validity and uncertainty calculation developed in the **bbττ Run 2 non-resonant analysis** → Method validated for **bbττ Run 2 resonant analysis**

Summary

My thesis focuses on the study of HH production with CMS data

- Large person power involved in bbττ analyses, 16 institutes working and synchronizing together
- Active collaboration within a working group involving 6 institutes
  - Deployment of the common framework (developed at CIEMAT)
  - Common analysis effort towards a 2022&2023 publication in 2025
- Contributions to **Resonant Run 2 X→HH→bbττ analysis** (not public yet, in review, targeting Moriond - B2G-24-011):
  - QCD validation tests
  - HH-BTag retraining and performance evaluation on the analysis samples
- **Run 3 HH→bbττ analysis → main thesis topic**
  - Continuous development of HH-BTag algorithm
  - Optimization of VBF strategy of the analysis
  - Development of new HH signal interpretations

Other contributions:

- Trigger shifts at CMS control room
- τ validation of 2024 data collected by CMS and continued involvement with Tau POG development work
- Conference talk: "Searches for Higgs boson pair production at CMS" at SUSY2024
- Conference talk: "Search for Higgs boson pair production in the bbττ final state at CMS" at XXXIX RSEF Physics Biennial

Thank you for your attention!

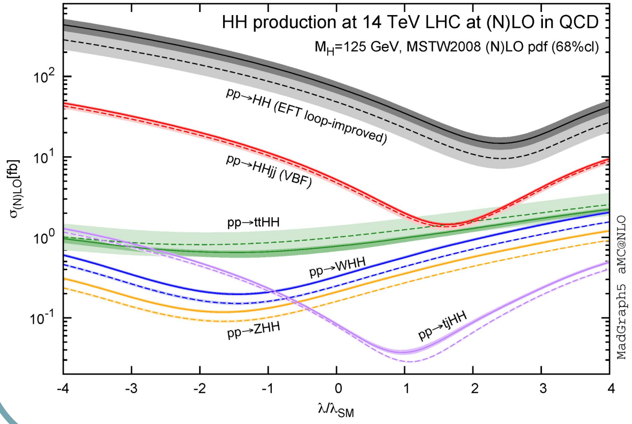
Measuring couplings modifiers

- Self-couplings can be constrained through total HH cross section and differential distributions

→ **Experimental sensitivity to κλ and κ2V depends on HH kinematics → directly related to the mHH shape**

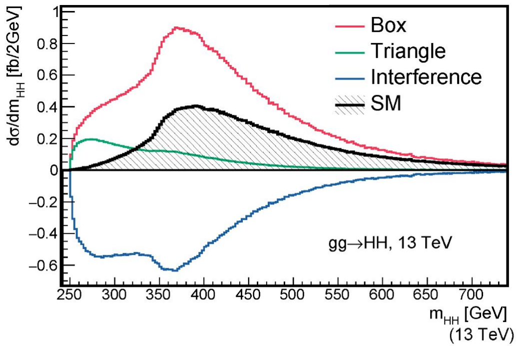
Total HH cross section

Dependence with κλ Phys. Lett. B 732

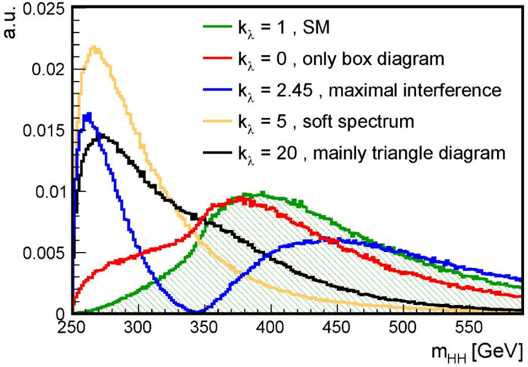


Differential cross section dσ/dmHH

Symmetry 2022, 14(7), 1467



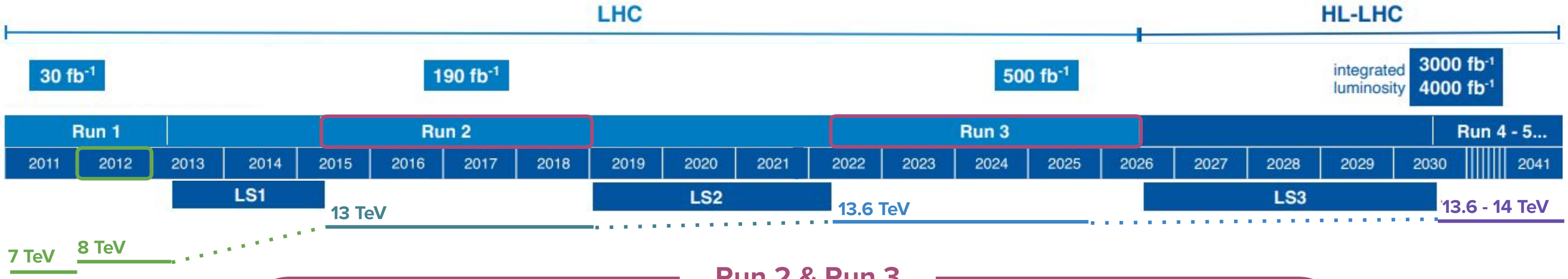
Anomalous κλ values alter the shape of the mHH → mHH largely differs for various coupling hypotheses



The Higgs at the LHC

[CMS Lumi Public Results]

- The **LHC** is a 27 km circular collider located at CERN
- Protons or heavy ions circulate in opposite directions, colliding in four interaction points where the experiments (CMS, ATLAS, LHCb, ALICE) are located

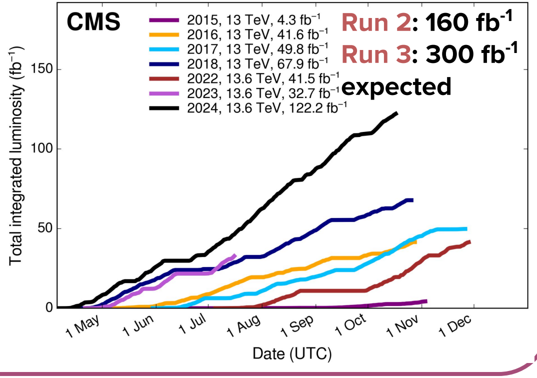


4 July 2012
Higgs boson discovery by CMS and ATLAS

Run 2 & Run 3

Measurement of the Higgs boson properties
➔ The Higgs boson so far behaves in a **very SM-like manner**
➔ Uncertainties still large, more data is needed to really understand the nature of the boson

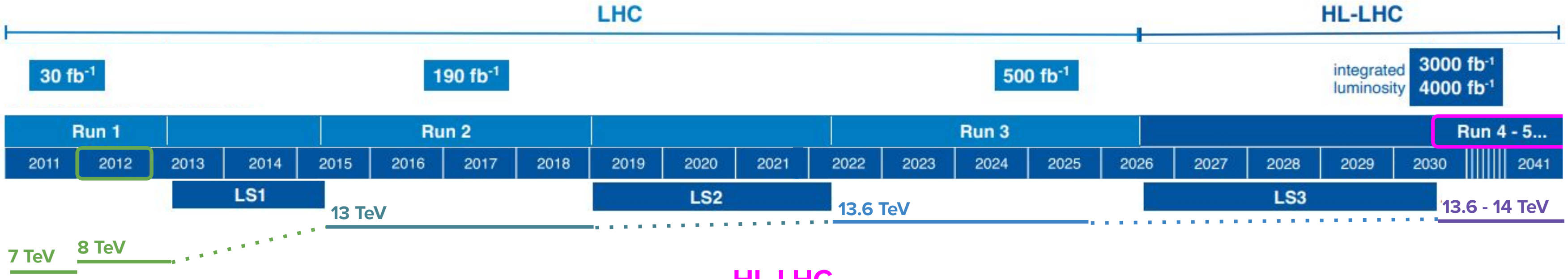
More questions still unanswered:
➔ **Shape of potential: Higgs self-couplings**



The Higgs at the LHC

CMS-PAS-HIG-20-011

- The **LHC** is a 27 km circular collider located at CERN
- Protons or heavy ions circulate in opposite directions, colliding in four interaction points where the experiments (CMS, ATLAS, LHCb, ALICE) are located



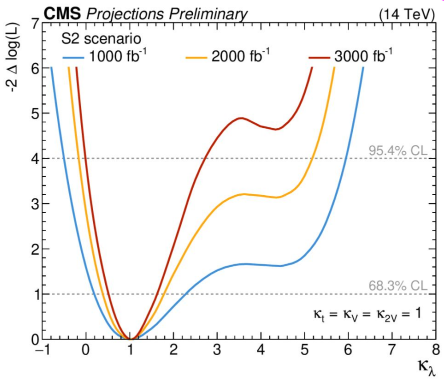
4 July 2012
Higgs boson discovery by CMS and ATLAS

HL-LHC

Total integrated luminosity: 3000 fb-1
Factor ~10 higher than LHC luminosity

Large impact of the high luminosity
➔ Increase the potential for new discoveries

➔ Potential to observe HH on HL-LHC



Di-Higgs phenomenology

- Rich phenomenology with many final states accessible at LHC
  ➜ **There is not a single golden channel**

- Significant **experimental challenges** due to their rare production rate and complex final states
- To achieve good sensitivity ➜ **compromise** between
  - **Branching Ratio (BR)**
  - **Final state signal purity**

➜ Escaping gradually these two constraints thanks to improving reconstruction techniques and identification methods

"Big 3" HH analyses

HH➜4b : Largest BR, challenging due to high b-jet multiplicity and QCD background
HH➜2b2τ : sizeable branching ratio, lower QCD background
HH➜2b2γ : rare process but clean signature due to photons

HH Branching Ratios

Increasing statistics

Decreasing background complexity

	bb	WW	$\tau\tau$	ZZ	$\gamma\gamma$
bb	34%				
WW	25%	4.6%			
$\tau\tau$	7.3%	2.7%	0.39%		
ZZ	3.1%	1.1%	0.33%	0.069%	
$\gamma\gamma$	0.26%	0.10%	0.028%	0.012%	0.0005%

Public result by CMS

Basis for Run-2 HH→bbττ searches

Analysis strategies

Target
- Events with at least 2 b-jets tagged (DeepJet) and 2τ tagged (DeepTau)
- Covering 87.6% of the full di-τ decay modes by considering three di-τ final states: $e\tau_h$, $\mu\tau_h$, $\tau_h\tau_h$

Trigger strategy
Single electron, single muon, electron or muon + hadronic τ, di-τ_h

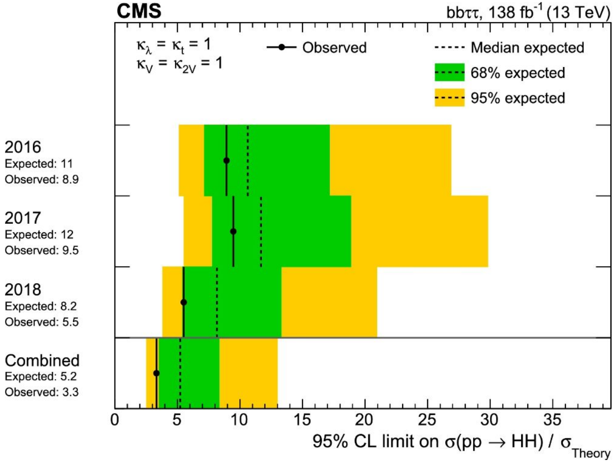
Main backgrounds
- Large $t\bar{t}$ and DY → Simulation
- QCD subleading → Data driven method

Event classification
Classification of the events in different categories to maximize the sensitivity

Signal extraction
Fit signal vs background NN output

Main uncertainties
- Statistically dominated
- Theory uncertainties
- Main experimental systematic: QCD background modelling

Phys. Lett. B 842.137531



Inclusive (ggF + VBF)

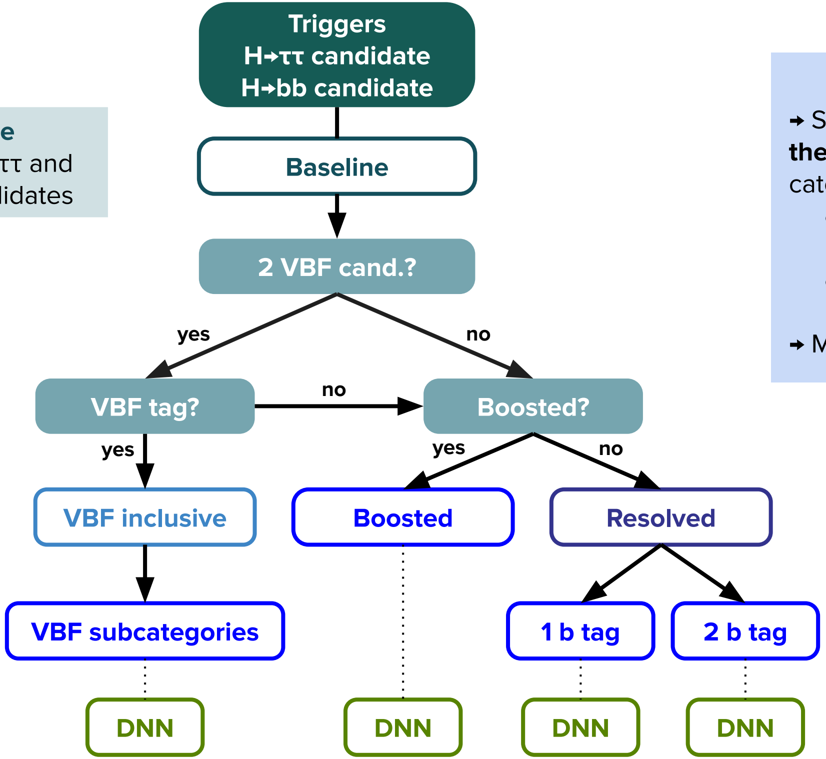
Expected: $\sigma^{HH} < 5.2\times\sigma^{HH}$(SM) @95% CL
Observed: $\sigma^{HH} < 3.3\times\sigma^{HH}$(SM) @95% CL

Event classification

- Current classification of the events in different categories to maximize the sensitivity (**Run 2 strategy**)

Baseline
Events with a ττ and a bb pair candidates

VBF inclusive
➔ 2 VBF jet candidates with mjj>500 GeV and Δη>3
➔ At least one b jet must satisfy the medium b tag WP



Final analysis categories
➔ Signal region defined through an **elliptical cut in the mττ-mbb plane** for resolved and boosted categories

- **Boosted**: Fat jet with 2 loose b tag subjets matched to the selected b jets
- **Resolved 1-b & 2-b**: one or two jets satisfying the medium b tag WP

➔ Multi-class classification for **VBF categories**

DNN for signal extraction to improve the analysis sensitivity
➔ Input features based on kinematics and topology of the event

HH→2b2τ non-resonant

Phys. Lett. B 842.137531

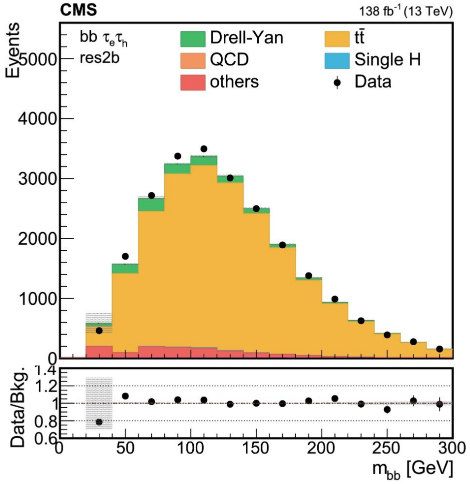
- 2b2τ analysis studying the two main production modes: ggF and VBF
  - Signature of VBF: Two additional forward jets with high invariant mass

Data/MC results

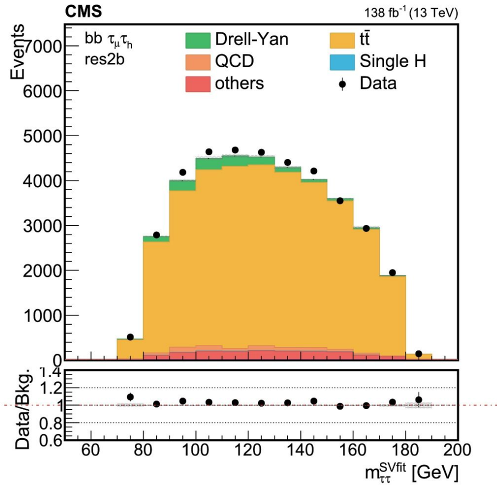
- **Events classified in 8 mutually exclusive categories:**
  - **boosted → Fat jet with 2 loose b-tag sub-jets**
  - **resolved 1b, resolved 2b → 1 or 2 jets passing medium b-tag WP**
  - **VBF categories → mjj>500 GeV and Δηjj > 3. Split into five subcategories (VBF, ggF, ttH, TT, DY) using a multi-classifier**

Data/MC distributions- res 2b

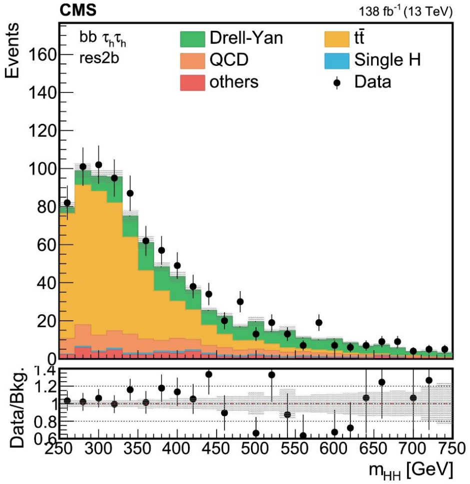
bb mass - eτh channel



ττ mass - μτh channel



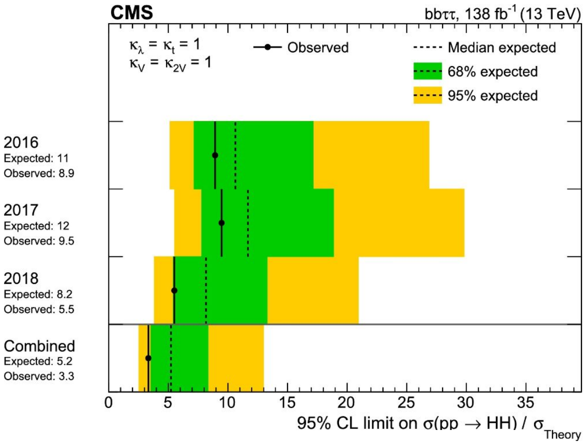
HH mass - τhτh channel



HH→2b2τ non-resonant

Phys. Lett. B 842.137531

Limits on signal strength



Inclusive (ggF + VBF)

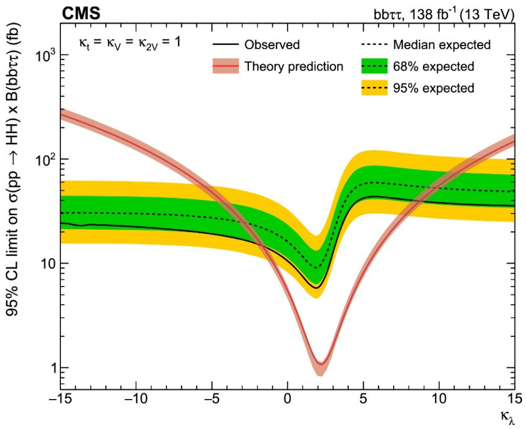
Expected: $\sigma^{HH} < 5.2\times\sigma^{HH}$(SM) @95% CL
Observed: $\sigma^{HH} < 3.3\times\sigma^{HH}$(SM) @95% CL

Only VBF

Expected: $\sigma^{HH} < 154\times\sigma^{HH}$(SM) @95% CL
Observed: $\sigma^{HH} < 124\times\sigma^{HH}$(SM) @95% CL

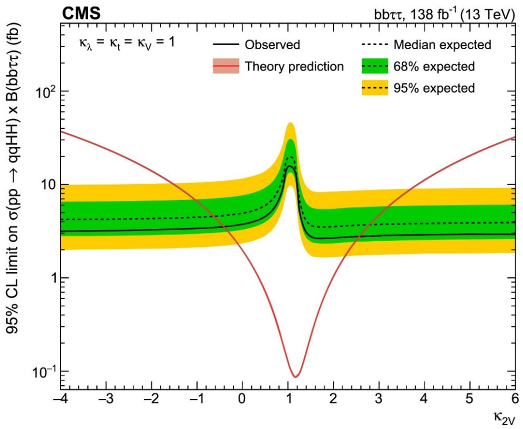
Limits on couplings

Upper limits on the σ^{HH} inclusive (ggF + VBF) × BR(bbττ) as function of κ_λ
→ Other couplings set to SM value



Expected at 95% CL: $-2.9<\kappa_\lambda<9.8$
Observed at 95% CL: $-1.7<\kappa_\lambda<8.7$

Upper limits on the VBF only σ^{HH} × BR(bbττ) as function of κ_{2V}
→ Other couplings set to SM value



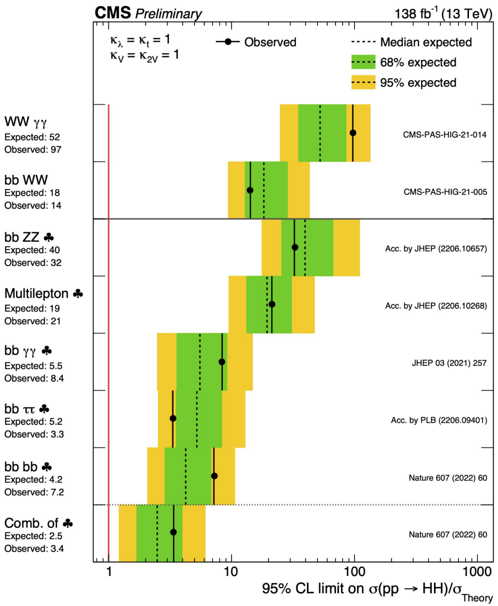
Expected at 95% CL: $-0.6<\kappa_{2V}<2.8$
Observed at 95% CL: $-0.4<\kappa_{2V}<2.6$

Non-resonant HH Run 2 combination

Nature 607 (2022) 60

- Similar sensitivity between **boosted HH→4b**, **HH→2b2τ** and **HH→2b2γ**
- Maximal sensitivity obtained through combination → **most restrictive upper limits** on the Di-Higgs cross section

Limits on signal strength

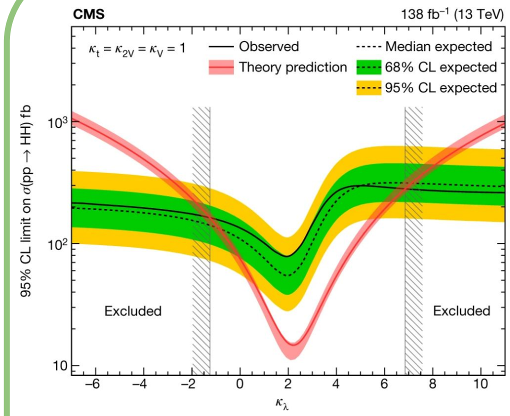


$\sigma^{HH} < 3.4\ (2.5) \times \sigma^{HH}(SM)$ Obs (Exp)

x5 better than 36fb^{-1} combination

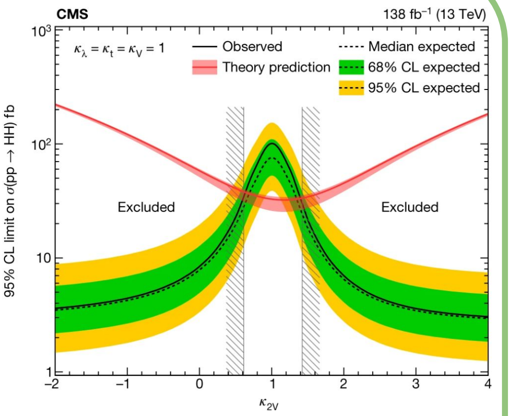
Much better than predicted by simple luminosity scaling thanks to improved reconstruction/identification techniques and analysis techniques → **Crucial role of developing analysis strategies**

Limits on couplings



-1.24<κ_λ<6.49

Approaching the exclusion of κ_λ = 0



0.67<κ_{2V}<1.38

κ_{2V} = 0 excluded at >5σ

JHEP 11 (2021) 057

HH→2b2τ resonant

- Search for a heavy Higgs boson H decaying into h(125) and another hs: H→hhs production decaying into **2b2τ** (**h→2τ, hs→2b**)

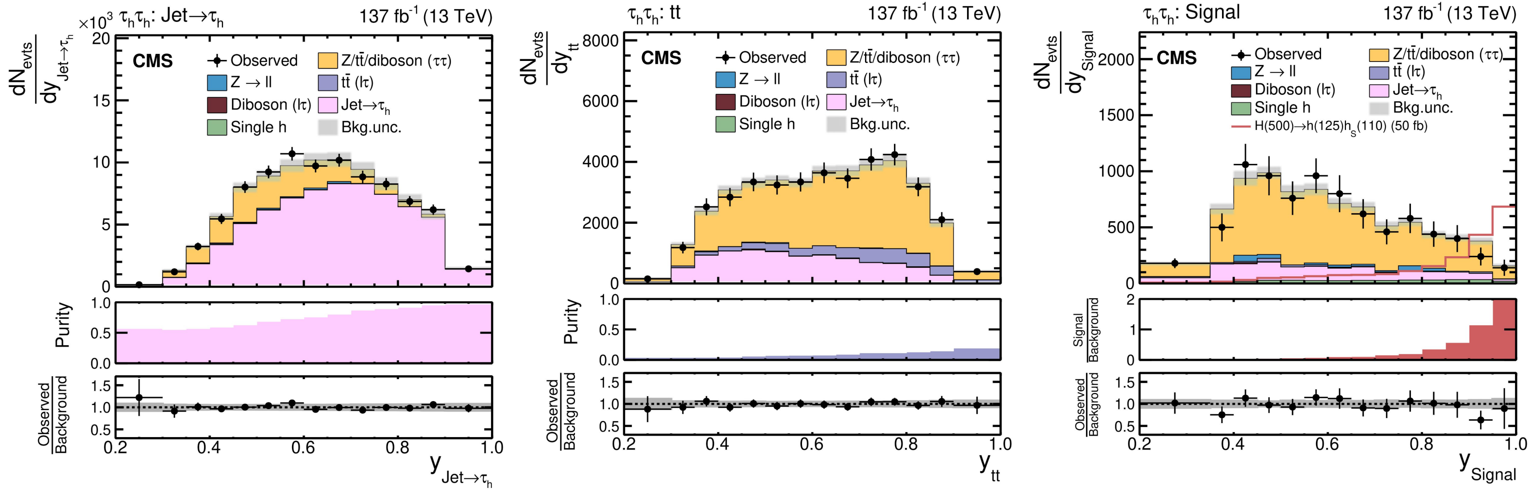
Data/MC results

Events classified in 5 categories for each final state (eτh, μτh, τhτh) and each data taking period (2016, 2017, 2018) using NN multiclassifier:

- **Signal category**
- **Background enriched categories:**
  - **→ ττ**
  - **→ tt̄**
  - **→ Jet misidentified as τh (Jet→τh)**
  - **→ Remaining minor background processes (misc)**

τhτh event categories (all years)

Jet→τh category | **tt̄ category** | **Signal category**



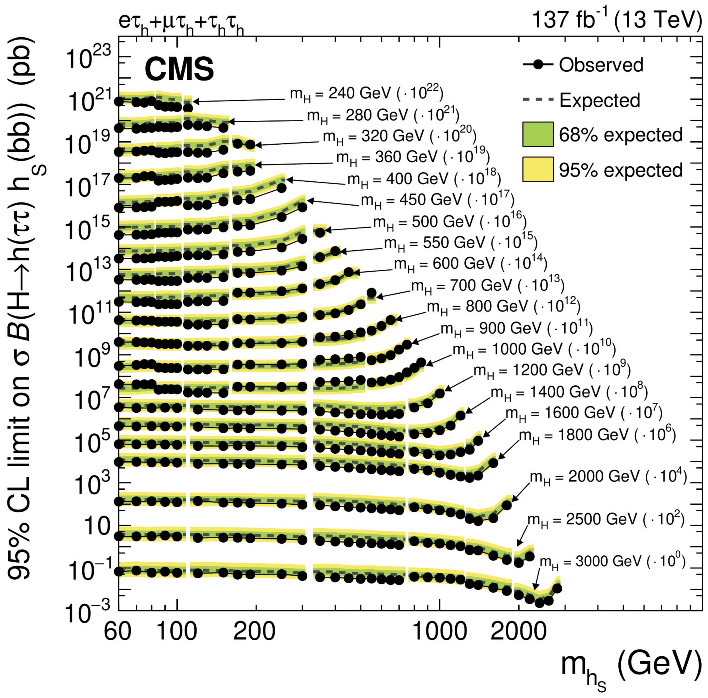
HH→2b2τ resonant

JHEP 11 (2021) 057

- Search for a heavy Higgs boson H decaying into h(125) and another hs: H→hhs production decaying into **2b2τ** (**h→2τ, hs→2b**)
  - H mass range 240-3000 GeV
  - hs mass range 60-2800 GeV

  → No signal has been observed

Limits on σ×BR(H→h(ττ)hs(bb))



Model independent @95% CL upper limits on σ×BR(H→h(ττ)hs(bb)) set with a sensitivity ranging from 125 fb (mH=240GeV, mhs=85GeV) to 2.7 fb (mH=1000GeV, mhs=350GeV)

Run 3 improvements

Improve statistics

Run 3 is underway, with higher energy and more data to analyze to overcome statistical limitations

Analysis techniques

Development of more advanced analysis techniques by taking advantage of new tools and techniques becoming available

Reconstruction and identification

Improved reconstruction and object identification techniques: ParticleNet and extensive and enhanced use of Machine Learning

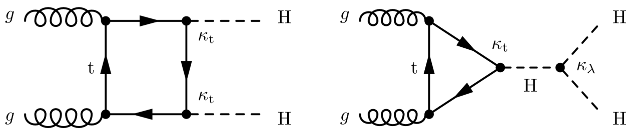
Trigger strategies

Improved for Run 3 based on improved object identification: new b-tagging and τ-tagging algorithms (ParticleNet and DeepTau) → Improvements are expected for all HH searches targeting bb or $\tau\tau$ final states

Extended interpretations

Exploration of extended interpretations, such as SMEFT, HEFT

Run-3 HH→bbττ effort



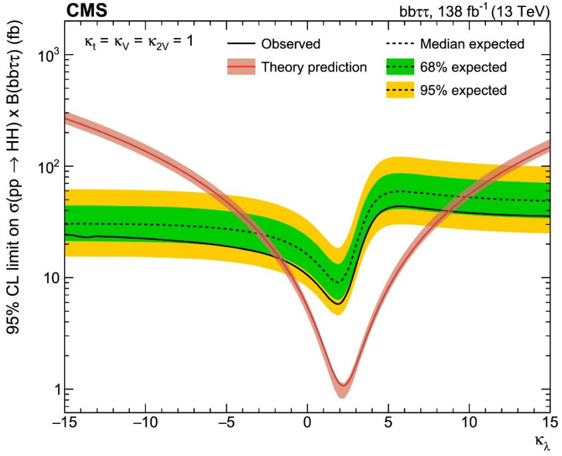
Comprehensive bbττ analysis effort: targeting HH, ZH, ZZ

- ggF + VBF in both **resolved & boosted** topologies
- Cross-section and couplings scans **kλ, k2V**
- Exploration of **extended interpretations**: SM + SMEFT/HEFT

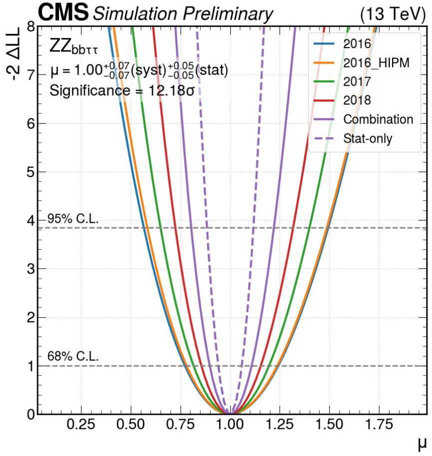
➜ Early publication: **2022&2023 analysis, combined with Run-2 at datacard-level**

➜ Long term: **full Run-3 nonresonant & resonant analyses with re-analysis of Run-2 data**

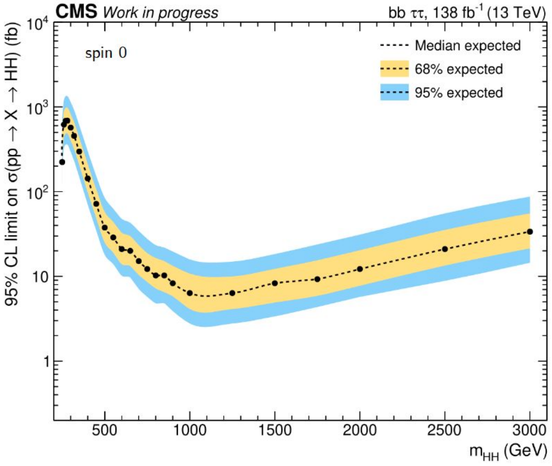
HH non-resonant



ZZ/ZH



Ongoing HH resonant



➜ **For now prioritizing HH nonresonant early publication**

Frameworks & groups

Targeting early nonresonant publication

- CCLUB **(CIEMAT - CEA - LLR - UZH - Milano-Bicocca - Colorado-Boulder)**
- Bamboo **(UCLouvain)**
- ColumnFlow **(UHH - KBFI - LIP)**
- FLAF **(Texas A&U - Pisa) ➔** Using **HLepRare Run-3 skims**
- Run 3 framework setup using coffea **(CMU)**
- **PKU**

Further efforts targeting resonant production

- CROWN **(KIT) (X➔HY)**
- **University of Wisconsin-Madison (High mass boosted)**

Person power: ~16PhD + ~11 Postdoc + ~2 Master students (plus staff/seniors in each institute)

CMS standard b-taggers

- **DeepFlav**
  - Multi-classification deep-neural-network algorithm employing low-level properties of several charged and neutral particle-flow jet constituents, supplemented with properties of secondary vertices associated with a jet.
  - State-of-the-art tagger during Run 2 for heavy flavor tagging

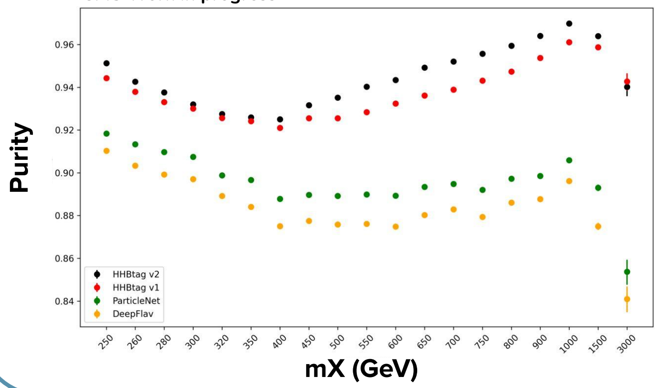
- **ParticleNet**
  - Graph Neural Network architecture ➔ Dynamic Graph Convolutional Neural Network based jet tagging algorithm. Instead of treating the jet as a collection of ordered constituents like DeepJet, a jet is considered as an unordered set of its constituent particles or a “particle cloud”
  - Recommended during Run 3 for heavy flavor tagging

- **RobustParT**
  - A ParticleTransformer model specific for the classification of AK4 jets. The transformer model introduces pairwise "interaction" features between all input jet constituents and secondary vertices. These additional layer of inputs give better view of the internal relations of the jet constituents, thus improving the performance of the model.
  - In addition, an Adversarial Training (AT) is used to enhance the robustness of the model against the mismodeling of our Monte-Carlo (MC) simulation. AT performs a distortion of our inputs features with respect to the loss function of the neural network. This allows our model to learn how to classify the jet flavour in a region around the jet input features distributions observed on our MC simulation, later reducing the impact of the mismodeling.
  - A combination of these two approaches is used to preserve the performance and improve the robustness of heavy flavor tagging and the tagger is called RobustParT

H→bb: The HH-btag algorithm

$$\text{Purity(classifier)} = \frac{N_{true}\text{(classifier)}}{N\text{(classifier)}}$$

Performance of different taggers



- **HH-BTag v2 (retrained for ongoing resonant analysis)**
- **HH-BTag v1 (used in non-resonant analysis)**
- **ParticleNet**
- **DeepJet**

Improved performance of HH-BTag

N_{true}(classifier): The number of events in which the selection of the b-jet pair candidates made by each classifier matches the **ground truth*** definition

N(classifier): Total number of events where a candidate is reconstructed (all the events by requiring two jet candidates reconstructed)

***Ground truth definition**

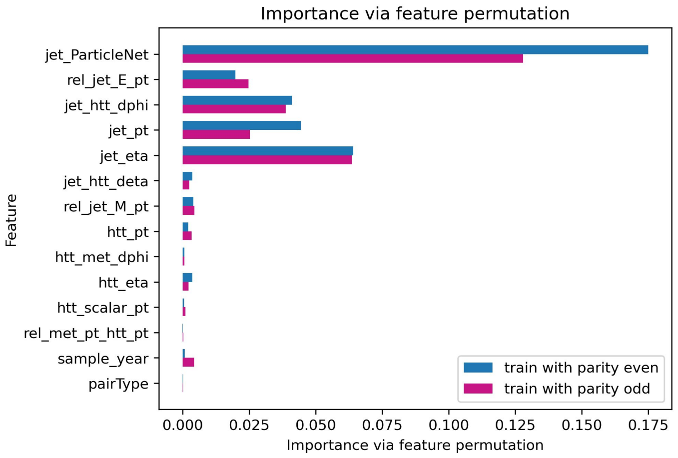
→ Two most energetic b-gen jets to perform the match

→ The true b-jet label is assigned to the reconstructed b-jet if a matching b-gen is found within a cone of size ΔR<0.5 around the direction of the b-reco

H→bb: The HH-btag algorithm

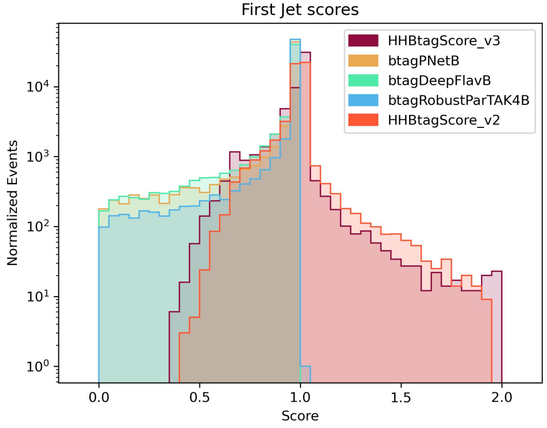
Feature importance

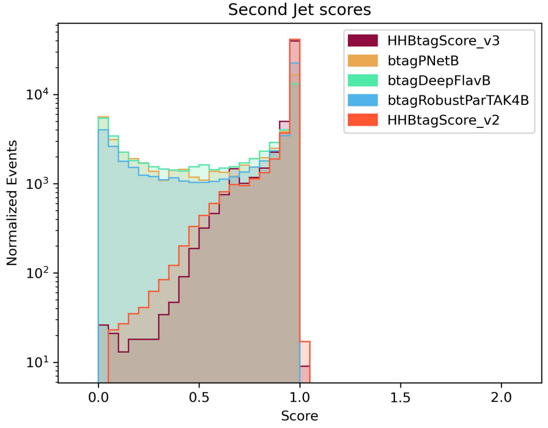
- Study of the feature importance computed via **feature permutation** on the training dataset (**parity even** / **parity odd**)
  - Overall dominance of the **b-tag** feature importance
  - Kinematics of the jet



Scores distributions

- **ggF 2022preEE** signal sample
- Jets **ordered according to the corresponding tagger**
- Requiring **at least 2 reco jets**



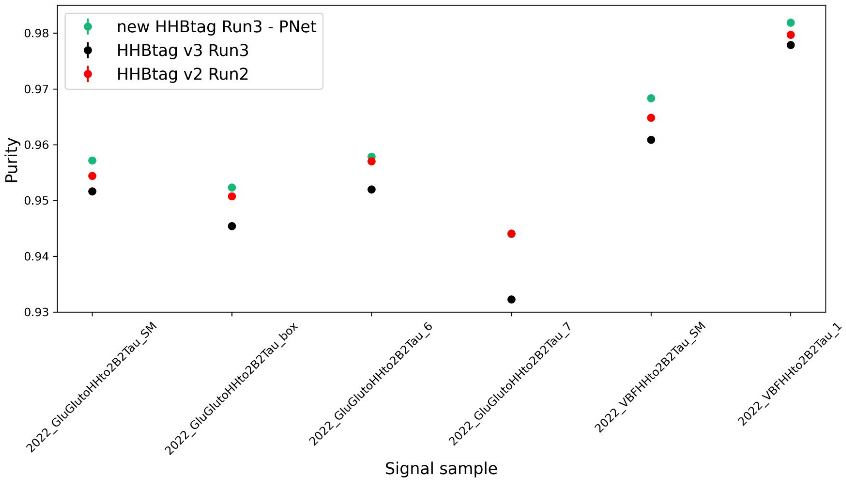


Purity studies

True b-jet label definition: GenJet Hbb:

- **v2**: Was selected by finding the two b-jets with invariant mass closest to 125 GeV
- **New definition**: Two GenJets that matches two GenPart required to be pdgId == 5 & MotherpdgId == 25 within a cone ΔR(genJet,genPart)<0.4

2022preEE



- **new HHBtag: v3.1** → Purity using new true b-jet label definition
- **HHBtag Run3: v3.0** → Purity using original true b-jet label definition
- **HHBtag Run2: v2** → Purity using original true b-jet label definition

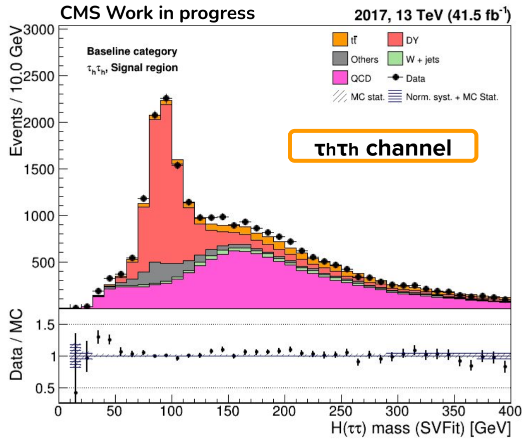
QCD estimation

- **QCD background** particularly important for τhτh final state → **Data-driven** estimation in the signal region using **ABCD method**
- Subleading background, but one of the main ones in the **impact on the sensitivity** → Importance of developing studies on the **validity of the method**

Validity of the method

Several tests performed to study the validity and stability of the ABCD method

- Tested robustness of the method for different isolation definitions → calculation of the associated systematic uncertainty
- Comparison with direct Data-MC subtraction in sideband region
- Study of the shape modelling
- Tested on highly populated QCD regions



Validity and uncertainty calculation developed in the **bbττ Run 2 non-resonant analysis** → Method validated for **bbττ Run 2 resonant analysis**

Object corrections status

2022 pre/postEE and 2023 pre/postBPix corrections available and implemented:

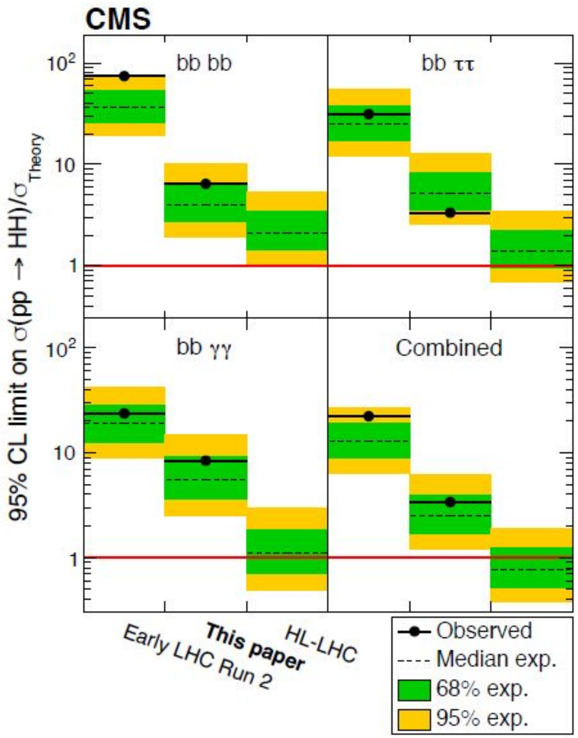
- **BTV** ➔ Run-3 Signal SF from X(bb) analysis available. Corrections for background to be done.
- **EGM** ✅
- **LUM** ✅
- **MUO** ✅
- **TAU** ✅
- **JME** ✅
  - **JECs**
  - **JERs**
  - **Jet Veto Maps**
- **MET** ✅
  - **JERs propagation**
  - **MET Filters**

Prospects for HH measurements

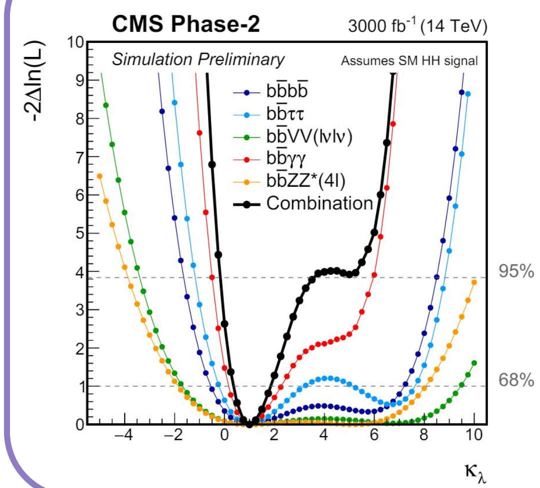
[Nature 607 (2022) 60](#)
[YR2018 (ECFA)](#) [Snowmass ATLAS+CMS](#)

- Large impact of the high luminosity that allows to extend the Di-Higgs production and decays modes accessible at LHC
- Many new developments on reconstruction and identification methods (triggers, machine learning based taggers)

➔ Potential to observe HH on HL-LHC (ATLAS+CMS)



Expected likelihood scan (assuming HH signal)



Prospects for λHHH (assuming no HH signal)