Search for CP Violation in H→ττ Decays at CMS

Alexei Raspereza

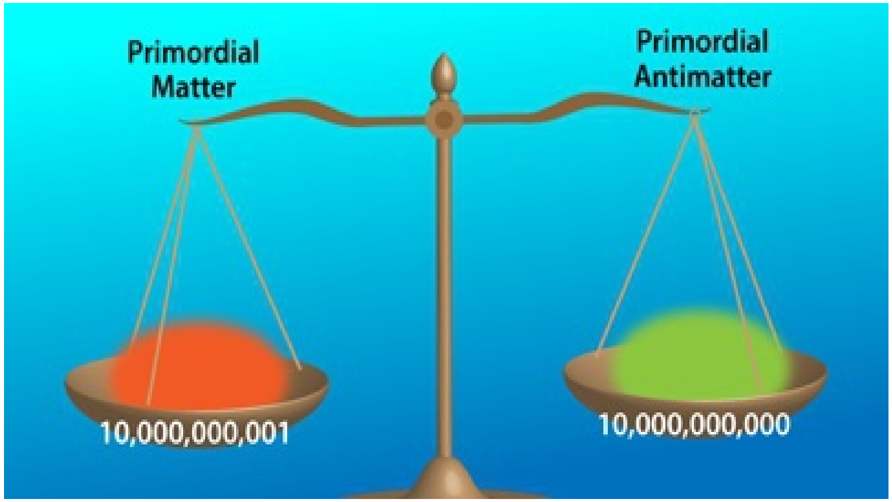
on behalf of the CMS Collaboration

CHEP-Yerevan, 11/09/2023

Introduction

- Origin of Baryon (matter-antimatter) Asymmetry in the Universe (BAU) remains one of the main riddles in contemporary cosmology
  - asymmetry of 10^{-10}-10^{-9} is required at the stage of baryogenesis to explain the observed dominance of matter over antimatter in nowadays Universe

- Sakharov conditions to generate BAU
  - baryon number violation
  - interactions out of thermal equilibrium
  - C and **CP violation**



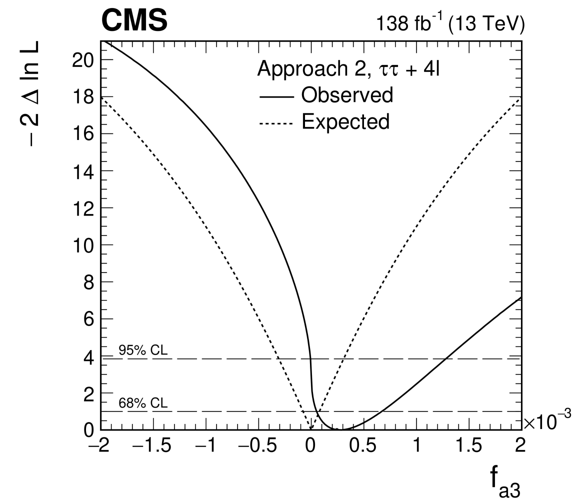
- Single complex phase in CKM matrix : source of CP violation in the SM
  - observed in decays of ***kaons, D*** and ***B mesons***
  - insufficient to explain the observed BAU

- Are there other sources of **CP violation**?
  - with the discovery of the Higgs boson, the scalar sector of theory has become new domain to search for **CP violation**

Higgs CP properties investigated at the LHC

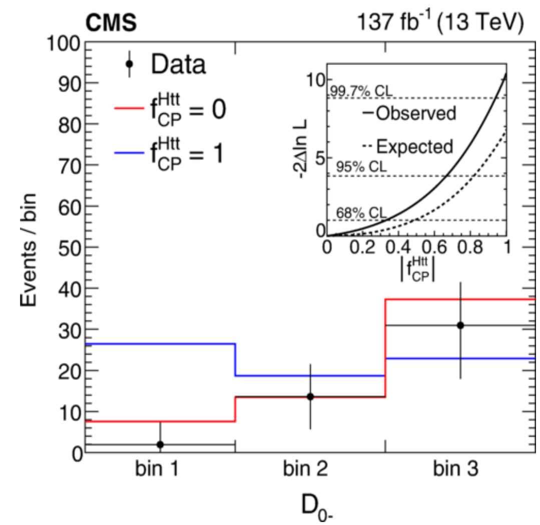
- CP properties of the HVV couplings extensively studied by CMS and ATLAS in production (VBF, VH) and decays (H→ZZ)
  → stringent upper limits are set on CP-odd terms of HVV couplings

Phys Rev D 108 (2023) 032013



- ttH production and H→ττ decays: excellent probes of the CP structure of Yukawa couplings
  → CP-odd term can contribute to coupling at tree level
  → measurements of ttH production at LHC disfavor pure CP-odd Htt coupling

PRL 125 061801



- This talk : measurements of the Higgs boson CP properties in H→ττ decays by CMS (**JHEP 06 (2022) 012**)

CP observable in H→ττ decays

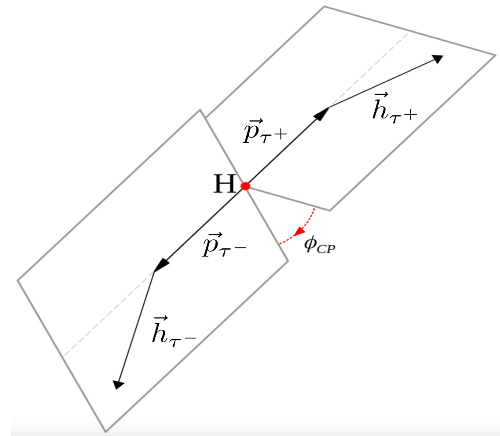
- Yukawa coupling (general case):

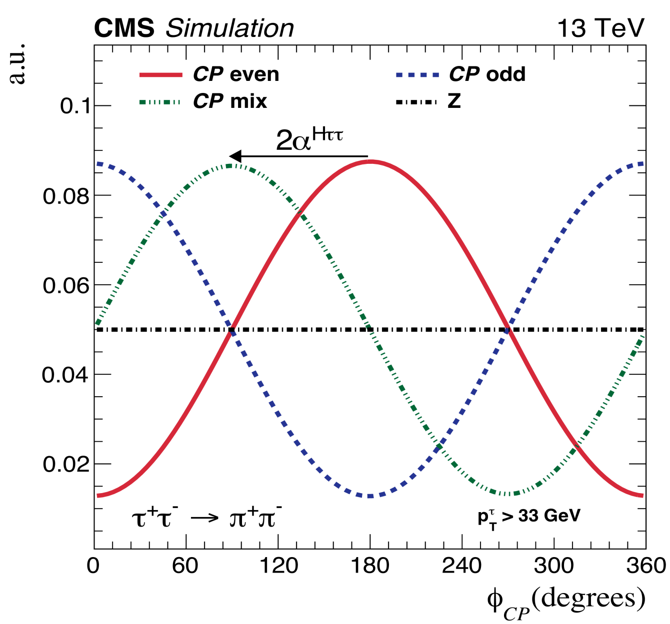
$$\mathcal{L}_Y = \frac{m_\tau}{v}\left(\kappa_\tau \bar{\tau}\tau + \tilde{\kappa}_\tau \bar{\tau} i\gamma_5 \tau\right) H$$

- CP-mixing angle : $\tan\left(\alpha^{\mathrm{H}\tau\tau}\right) = \frac{\tilde{\kappa}_\tau}{\kappa_\tau}$

- CP observable : angle between tau decay planes in the Higgs boson rest frame, ϕ_{CP}

- decay plane : plane spanned by
  - tau momentum, $\vec{p}_\tau$
  - polarimetric vector $\vec{h}_\tau$: most probable orientation of τ spin (defined by kinematics of τ decay products)





Information on CP mixing angle is encoded in the distribution of ϕ_{CP}

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}\phi_{CP}} \sim 1 - \frac{\pi^2}{16} b(x^+) b(x^-) \cos\left(\phi_{CP} - 2\alpha^{\mathrm{H}\tau\tau}\right)$$

$b(x^{\pm})$ - spectral functions, encapsulating dependence of the tau polarization on the momentum fraction $(x^{\pm})$ carried by the charged lepton or mesonic resonance

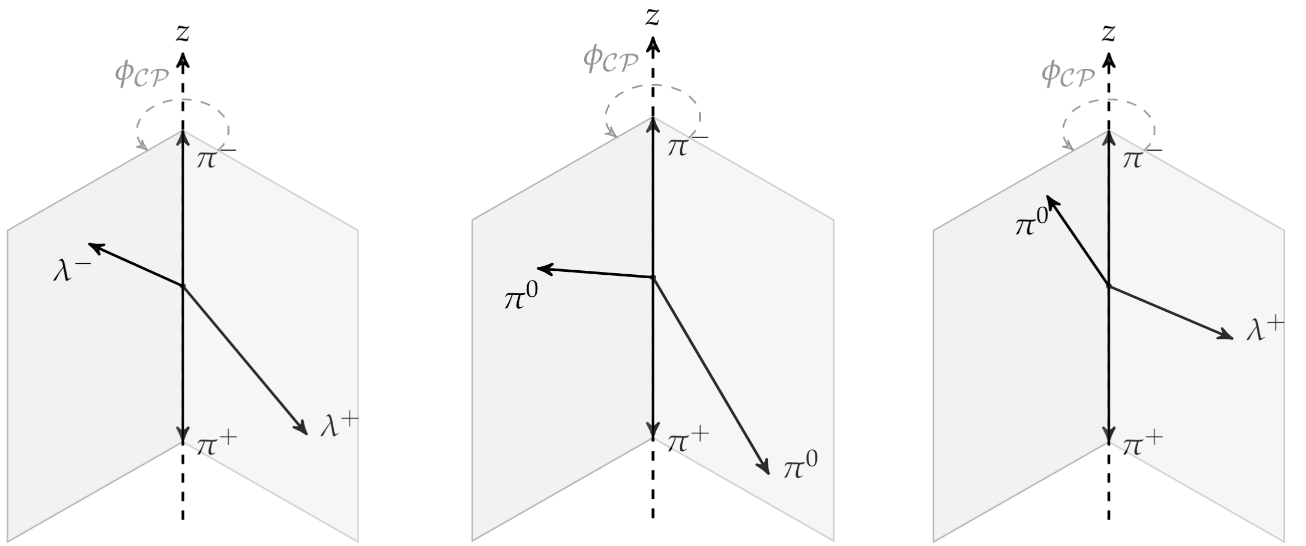
Reconstruction of τ decay planes

- exact reconstruction of tau decay planes in the Higgs rest frame (RF) is not feasible due to undetected neutrinos → use approximations

$\tau^{\pm} \to \pi^{\pm}\nu, \ell^{\pm}\nu\nu$:
spanned by impact parameter vector $\lambda^{\pm}$ and momentum of $\pi^{\pm}(\ell^{\pm})$

$\tau^{\pm} \to \rho^{\pm}(\pi^{\pm}\pi^{0})\nu,\ a_1^{\pm}(\pi^{\pm}\pi^{0}\pi^{0})\nu$:
spanned by momenta of charged pion and neutral pion system

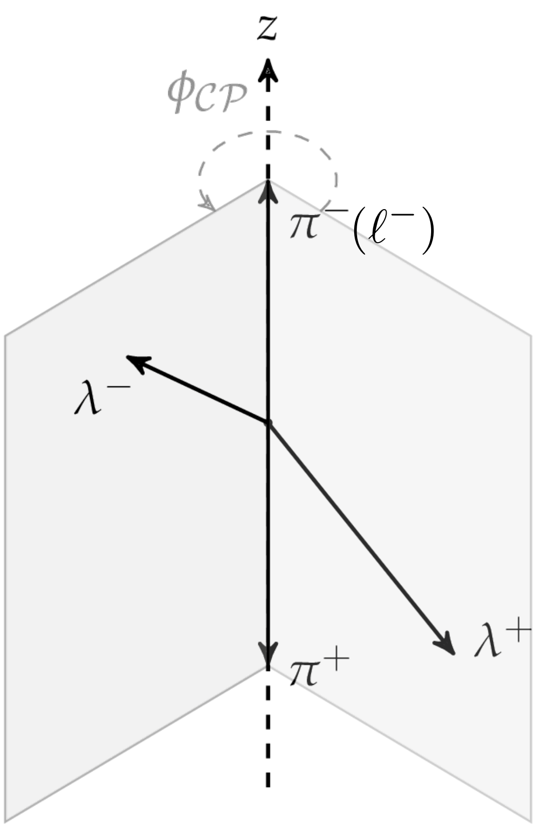
- decay planes are reconstructed in the RF of $\pi^{+}\pi^{-}(\pi^{\pm}\ell^{\mp})$



Impact parameter (IP) method

- IP method is applied in case of the following decays:

$$\tau^+\tau^- \to (\pi^+\nu)(\pi^-\nu),$$
$$\tau^\pm\tau^\mp \to (\pi^\pm\nu)(\ell^\mp\nu\nu)$$



- Impact parameter vectors are boosted in RF of $\pi^+\pi^-(\pi^\pm\ell^\mp)$

$$\vec{P}^+ + \vec{P}^- = 0$$

- acoplanarity angle

$$\phi^* = \arccos\left(\vec{\lambda}_\perp^+ \cdot \vec{\lambda}_\perp^-\right)$$

- introduce observable

$$\mathcal{O}^* = \vec{P}^- \cdot (\vec{\lambda}_\perp^+ \times \vec{\lambda}_\perp^-)$$

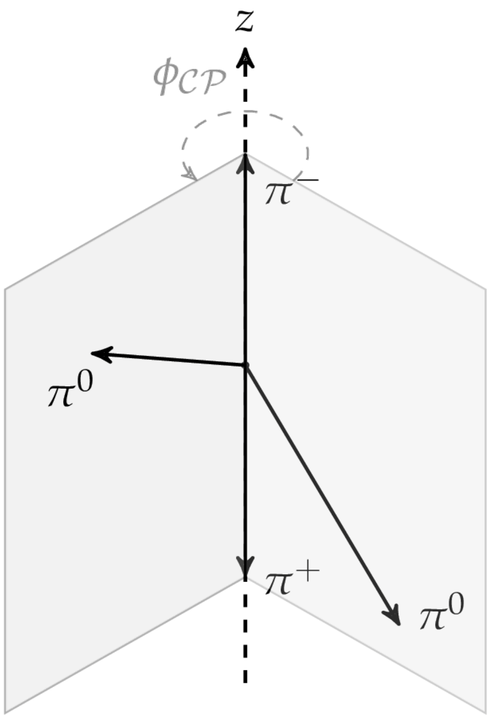
- ϕ_{CP} is defined in the range [0,2π]

$$\mathcal{O}^* > 0 : \ \phi_{CP} = \phi$$

$$\mathcal{O}^* < 0 : \ \phi_{CP} = 2\pi - \phi^*$$

Neutral-Pion (NP) method

- NP method targets primarily case when both τ leptons decay via $\tau^{\pm} \to \rho^{\pm}(\pi^{+}\pi^{0})\nu$



- 4-vectors of charged and neutral pions are boosted in RF of $\pi^{+}\pi^{-}$
- acoplanarity angle ϕ^{*} : angle between $\rho^{\pm}$ decay planes
- introduce variable

  $y^{\pm} = E_{\pi^{\pm}} - E_{\pi^{0}},\ y = y^{+}y^{-}$
- ϕ_{CP} is defined in the range [0,2π]

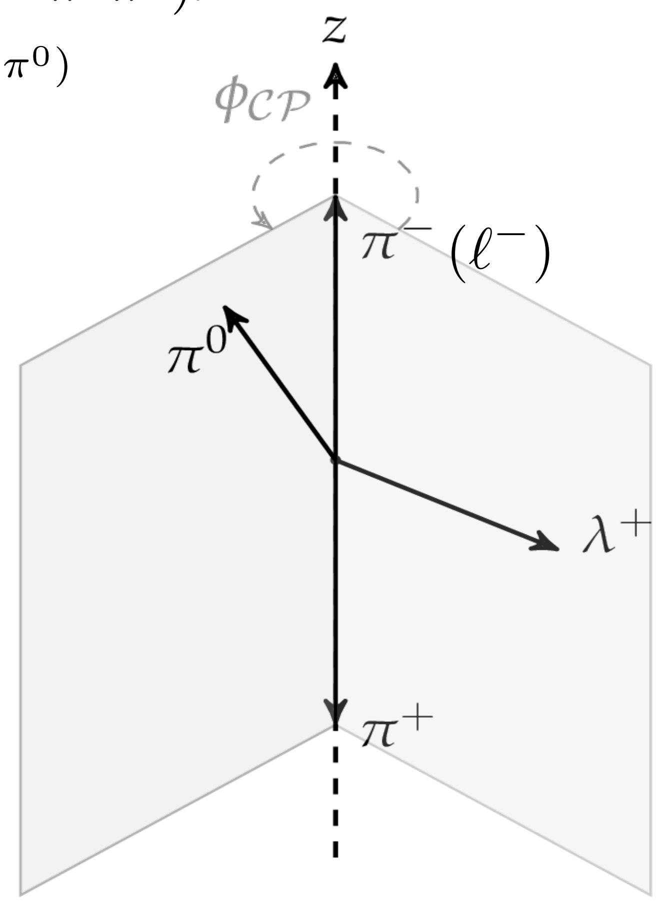
  $y > 0 :\ \phi_{CP} = \phi^{*}$

  $y < 0 :\ \phi_{CP} = 2\pi - \phi$

- Method is extended to the case of $\tau^{\pm} \to a_1^{\pm}(\pi^{+}\pi^{0}\pi^{0})\nu$ decay
  - electromagnetic constituents of τ are combined together to form $\pi^{0}\pi^{0}$-system that is treated analogously to π^{0} in $\tau^{\pm} \to \rho^{\pm}(\pi^{+}\pi^{0})\nu$ decay

Mixed (IP-NP) method

- Mixed method is used when:
  - one τ decay via $\tau^{\pm} \to \pi^{\pm}\nu, \ell^{\pm}\nu\nu$
    $\to$ decay plane is reconstructed with $\lambda^{\pm}$
  - while another – via $\tau^{\pm} \to \rho^{\pm}(\pi^{+}\pi^{0})\nu, a_1^{\pm}(\pi^{+}\pi^{0}\pi^{0})\nu$
    $\to$ decay plane is reconstructed with $\vec{p}_{\pi^0(\pi^0)}$



- definition of ϕ_{CP} via ϕ^* depend on variable

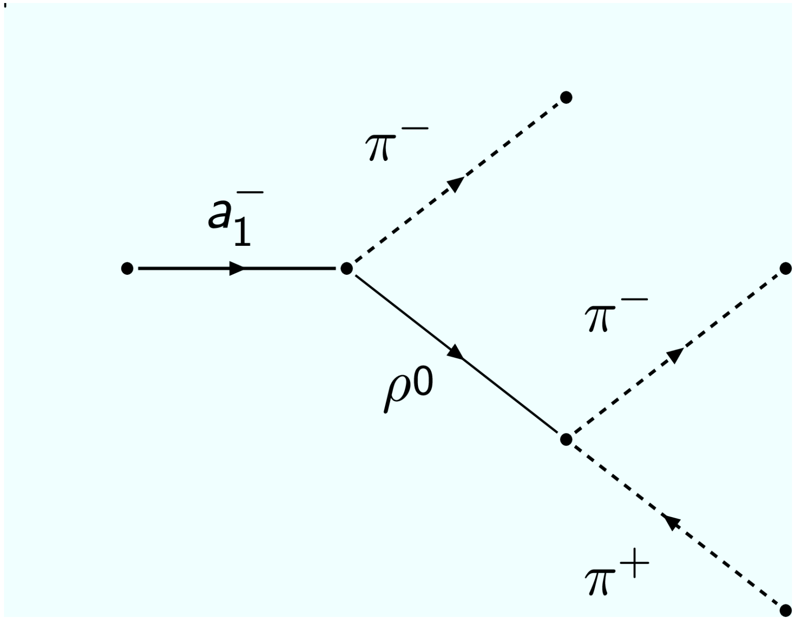
$$y^{\pm} = E_{\pi^{\pm}} - E_{\pi^0(\pi^0)}$$

$$\Rightarrow \quad \begin{array}{l} y^{\pm} > 0 : \ \phi_{CP} = \phi^* \\ y^{\pm} < 0 : \ \phi_{CP} = 2\pi - \phi \end{array}$$

Reconstruction of decay plane in $\tau \to a_1 \nu$

- Decay $\tau^- \to a_1^- \nu \to \pi^- \pi^- \pi^+ \nu$ proceeds via intermediate ρ^0 resonance :

  $a_1^- \to \rho^0 \pi^- \to (\pi^+ \pi^-) \pi^-$



- Most of tau spin information is carried by ρ^0 resonance which is treated as tau polarimeter

  $\to (\pi^+ \pi^-)$- pair with mass closest to m_{ρ^0} is selected as ρ^0 candidate

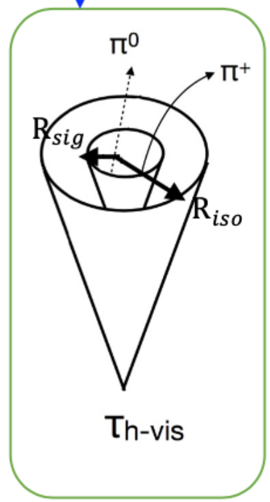
  $\to$ decay plane is defined as a plane spanned by π^+ and π^- from $\rho^0 \to \pi^+ \pi^-$

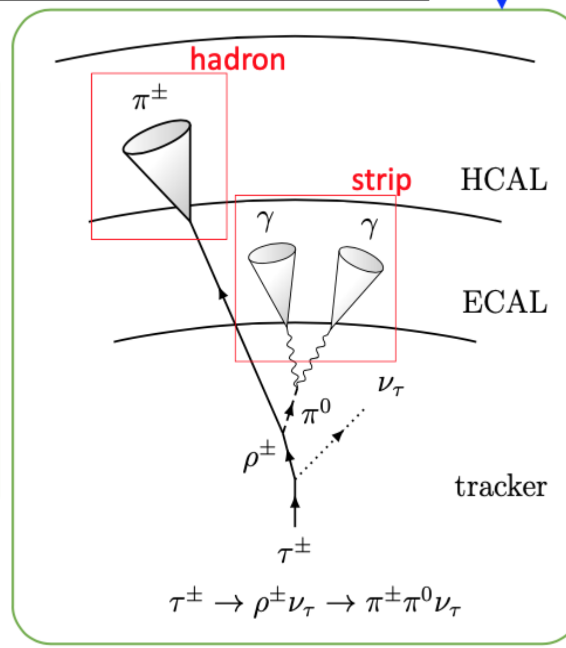
Reconstruction of hadronic tau decays

Hadron+strip (HPS) algorithm is used to reconstruct hadronic tau decays (τ_h)

- inputs : reconstructed PF candidates (γ, $e^{\pm}$, $\mu^{\pm}$, $h^{\pm}$, h^0)
- seeded by AK4 jets
- e/γ within $\eta\times\phi$ strips of dynamic size are collected and considered as π^0 candidates
- strips are combined with PF candidates identified as charged hadrons ($h^{\pm}$) to form four possible configurations
  1. $h^{\pm}$ and no strips
  2. $h^{\pm}$ + one or two strips
  3. $h^{\pm}h^{\pm}h^{\mp}$ and no strips
  4. $h^{\pm}h^{\pm}h^{\mp}$ + one strip

Decay mode	Resonance	$\mathcal{B}$ (%)
Leptonic decays		35.2
$\tau^- \to e^- \bar{\nu}_e \nu_\tau$		17.8
$\tau^- \to \mu^- \bar{\nu}_\mu \nu_\tau$		17.4
Hadronic decays		64.8
$\tau^- \to h^- \nu_\tau$		11.5
$\tau^- \to h^- \pi^0 \nu_\tau$	$\rho(770)$	25.9
$\tau^- \to h^- \pi^0 \pi^0 \nu_\tau$	$a_1(1260)$	9.5
$\tau^- \to h^- h^+ h^- \nu_\tau$	$a_1(1260)$	9.8
$\tau^- \to h^- h^+ h^- \pi^0 \nu_\tau$		4.8
Other		3.3

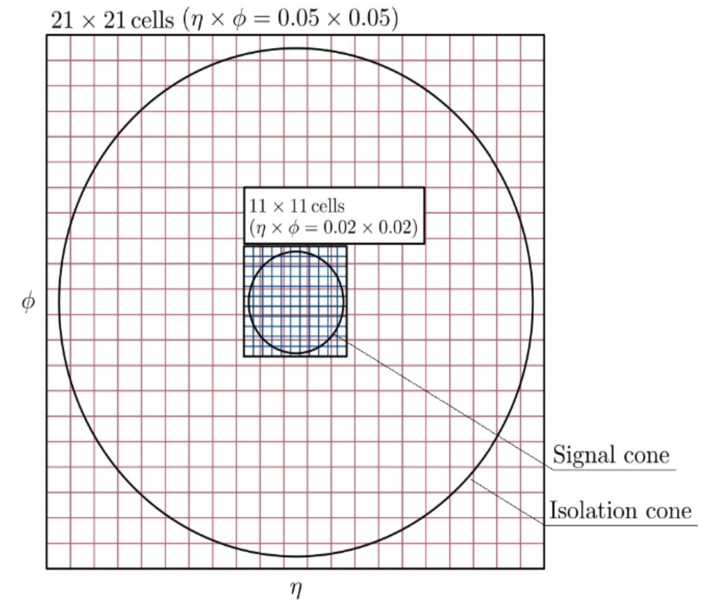


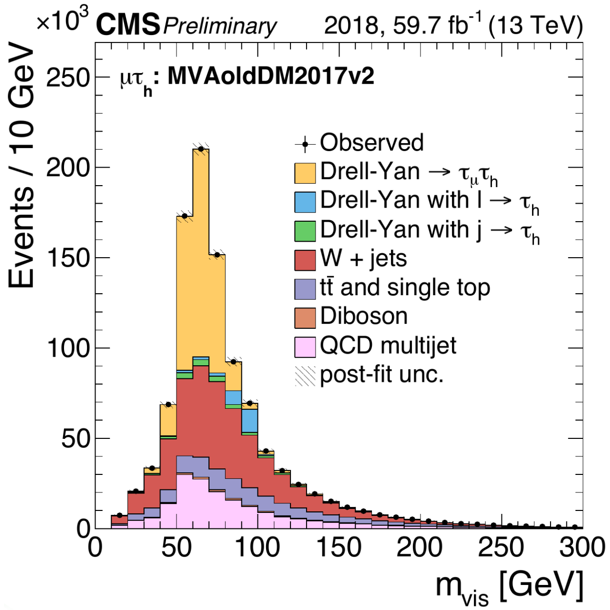


- all constituents must lie within a cone $\Delta R_{\text{sig}} = 3/p_{\text{T}}^{\tau}\,[\text{GeV}]$ (with ΔR_{sig} limited to range 0.05-0.1)

Discrimination of genuine τ_h against fakes

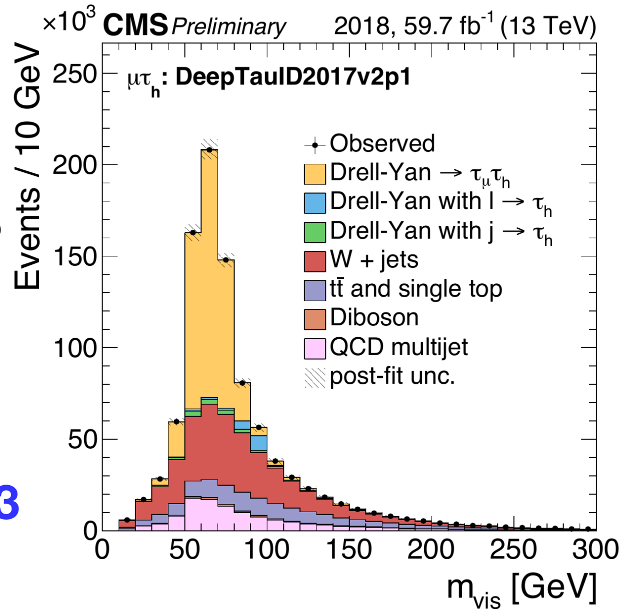
- Hadronic tau decays can be mimicked by QCD jets, electrons and muons
- Deep convolutional neural network (DeepTau) developed to discriminate τ_h against fakes
- Network combines
  - low-level information : PF candidates per cell in the fine/coarse $\eta \times \phi$ grid around τ axis
  - high-level information : kinematics of constituents, isolation variables, PU density





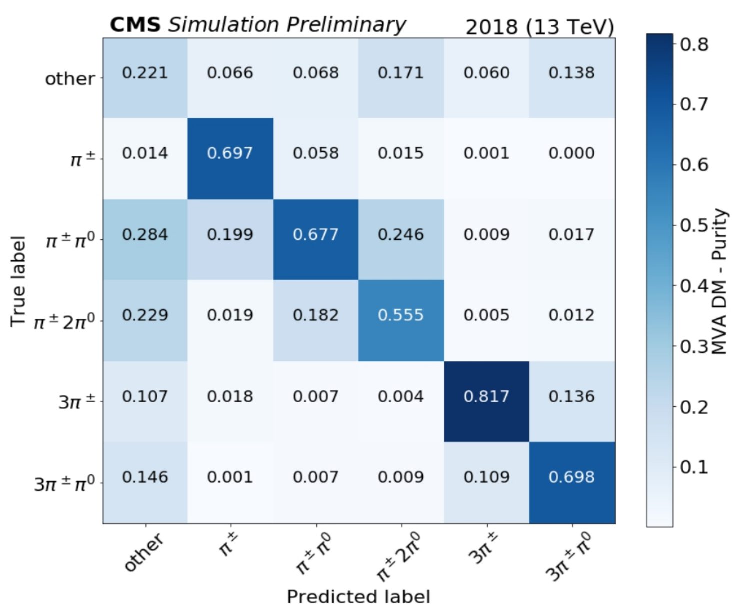
DeepTau:
increases genuine τ ID efficiency by ~ 20%, reduces fake rate by ~23%

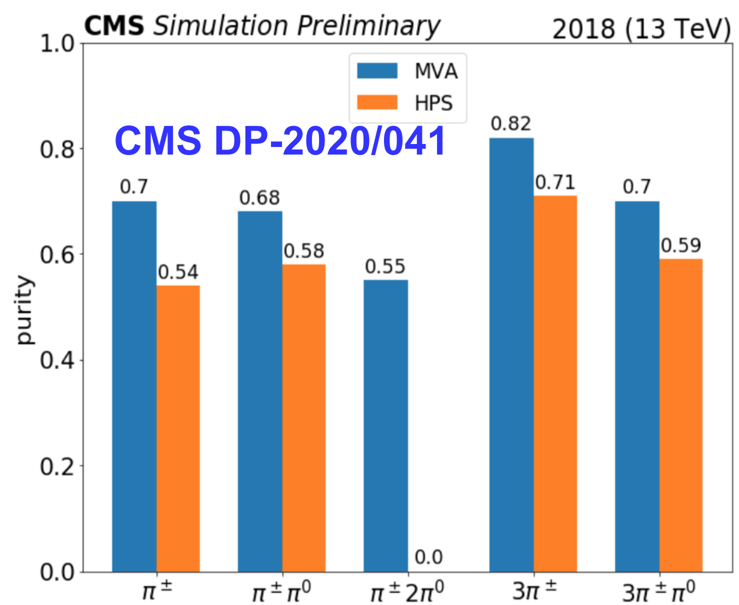
JINST 17 (2022) P07023



Identification of τ_h decay modes

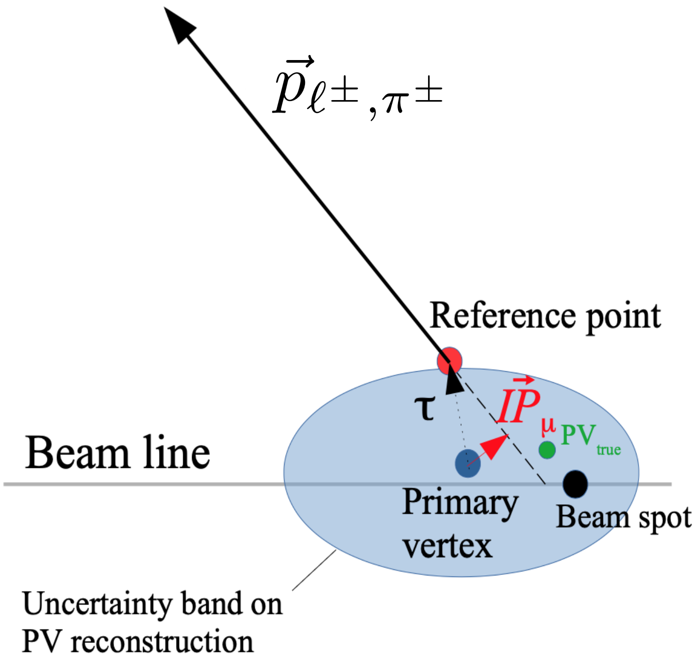
- HPS is not optimized to count π^0's in hadronic decays of tau leptons
- To improve decay mode identification two independent BDT classifiers are trained for 1 and 3-prong τ decays
  - Inputs:
    - → invariant mass of π^0 and ρ candidates
    - → kinematics of τ constituents
    - → angular separation between τ constituents
- Purity has improved by 10-20% compared to HPS algorithm
- Classifier provides also access to a_1(1-prong) decay mode



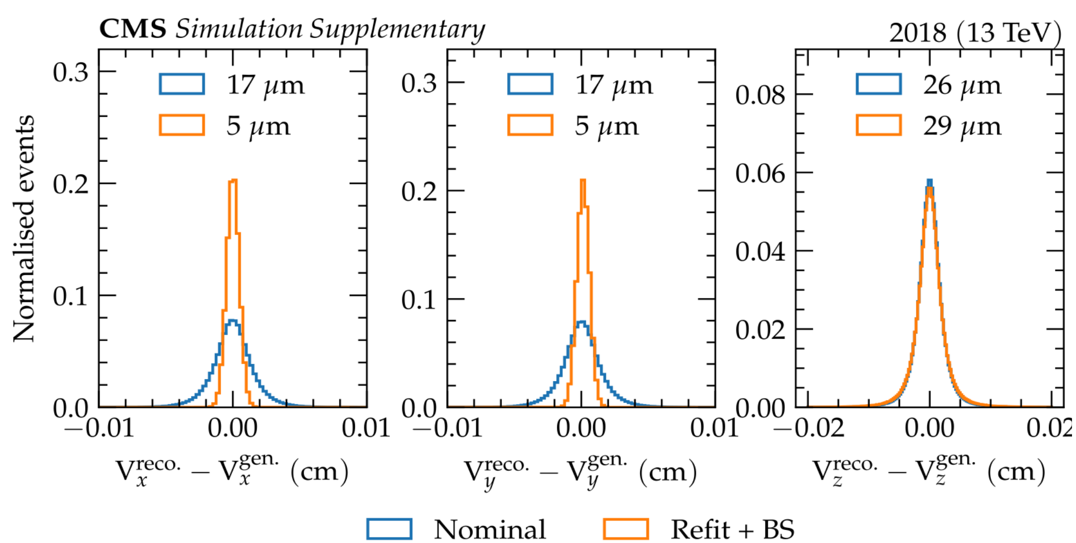


Reconstruction of impact parameters

- impact parameters (IP) reconstruction relies on precise measurement of primary vertex (PV)



- several PV choice have been studied
- PV reconstruction is improved by:
  - excluding tracks, originating from τ-lepton decay, from vertex fit
  - including beam spot (BS) constraint in the vertex fit



- with BS constraint precision in the reconstruction of transverse IP is improved by a factor of 3!

Analysis Overview

- Analysis is performed on 137 fb^{-1} of data collected with CMS at center-of-mass energy of 13 TeV
- Production mechanisms considered : $\mathrm{gg} \rightarrow \mathrm{H},\ \mathrm{qqH},\ \mathrm{VH}$
- Final states exploited : $\tau_\mu \tau_\mathrm{h},\ \tau_\mathrm{e} \tau_\mathrm{h},\ \tau_\mathrm{h} \tau_\mathrm{h}$

$$(\mathrm{e}, \mu, \pi, \rho, \mathrm{a}_1^{1\mathrm{pr}}, \mathrm{a}_1^{3\mathrm{pr}}) \times (\pi, \rho, \mathrm{a}_1^{1\mathrm{pr}}, \mathrm{a}_1^{3\mathrm{pr}})$$

- Triggers and offline selection

Channel	Year	Trigger requirement	Offline p_T (GeV)
$\tau_\mathrm{h}\tau_\mathrm{h}$	All years	$\tau_\mathrm{h}(35)$ & $\tau_\mathrm{h}(35)$	$p_\mathrm{T}^{\tau_\mathrm{h}} > 40$
$\tau_\mu\tau_\mathrm{h}$	2016	$\mu(22)$, $\mu(19)$ & $\tau_\mathrm{h}(20)$	$p_\mathrm{T}^{\mu} > 20$, $p_\mathrm{T}^{\tau_\mathrm{h}} > 25$
	2017, 2018	$\mu(24)$, $\mu(20)$ & $\tau_\mathrm{h}(27)$	$p_\mathrm{T}^{\mu} > 21$, $p_\mathrm{T}^{\tau_\mathrm{h}} > 32$
$\tau_\mathrm{e}\tau_\mathrm{h}$	2016	$\mathrm{e}(25)$	$p_\mathrm{T}^{\mathrm{e}} > 26$
	2017	$\mathrm{e}(27)$, $\mathrm{e}(24)$ & $\tau_\mathrm{h}(30)$	$p_\mathrm{T}^{\mathrm{e}} > 25$, $p_\mathrm{T}^{\tau_\mathrm{h}} > 35$
	2018	$\mathrm{e}(32)$, $\mathrm{e}(24)$ & $\tau_\mathrm{h}(30)$	$p_\mathrm{T}^{\mathrm{e}} > 25$, $p_\mathrm{T}^{\tau_\mathrm{h}} > 35$

What are the most important backgrounds?

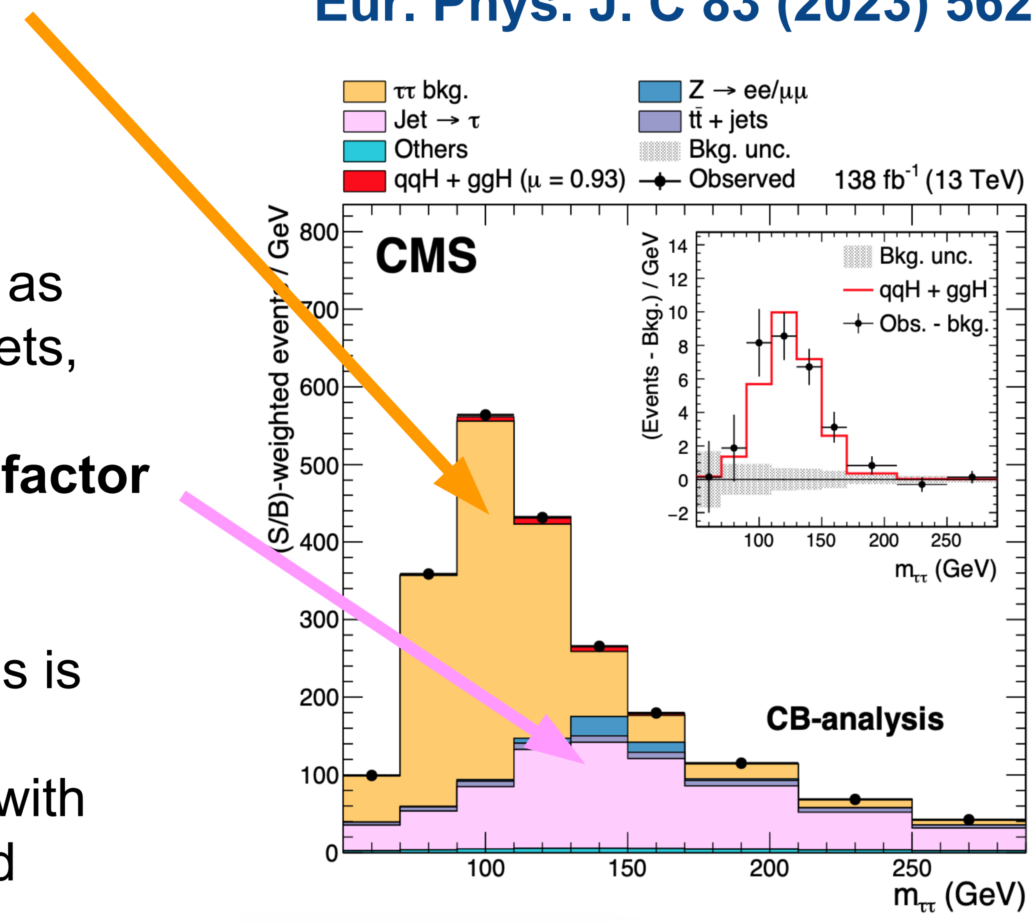
- S/B weighted summary plot of the di-τ mass distribution illustrates what major backgrounds are

- Background with genuine tau lepton pairs ($Z \rightarrow \tau\tau$, top pairs, dibosons, ...)
  → estimated from data using **tau-embedding** technique

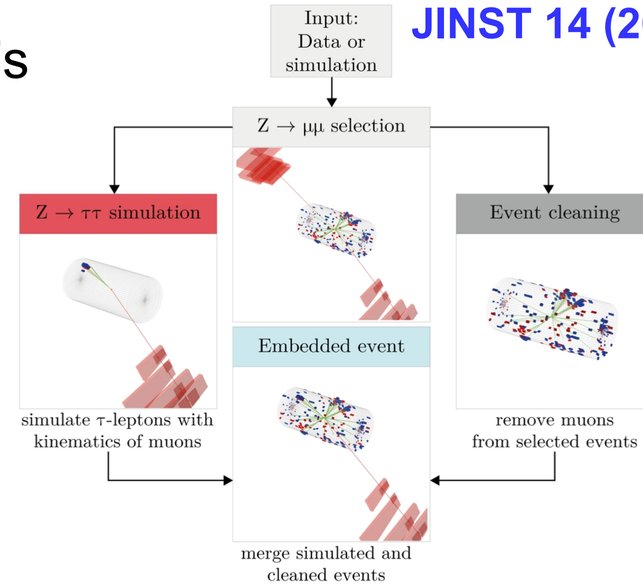
- Processes with jets misidentified as hadronic tau decays (QCD multijets, W+jets, ...)
  → estimated from data with **fake factor** method

- Contribution of other backgrounds is subdominant
  → evaluated with MC simulation with appropriate corrections applied

Eur. Phys. J. C 83 (2023) 562



Data driven background modeling



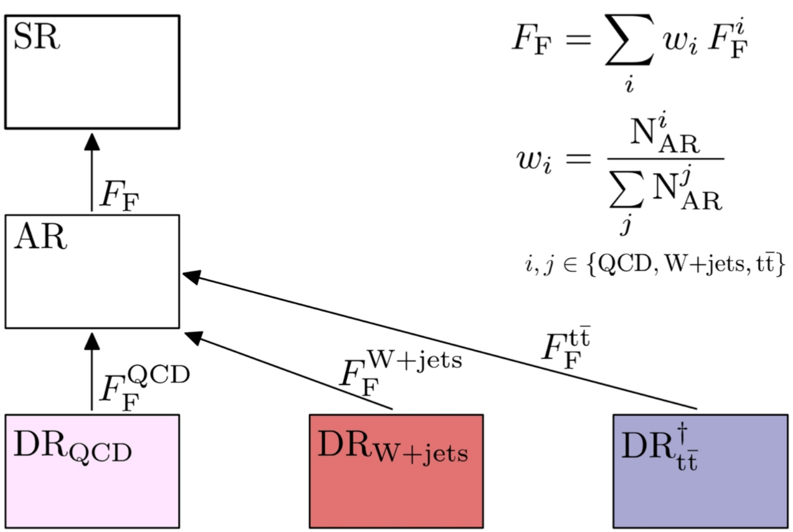
JINST 14 (2019) P06032

- Irreducible background with 2 genuine τ's (Z→ττ, ttbar, dibosons) is modeled with τ-embedding

→ replace muons in selected $\mu^+\mu^-$ events from data by simulated τ decays

→ data-driven description of detector activity except for τ decays

JHEP 09 (2018)007



$$F_F = \sum_i w_i \, F_F^i$$

$$w_i = \frac{N^i_{AR}}{\sum_j N^j_{AR}}$$

$$i, j \in \{\mathrm{QCD}, \mathrm{W+jets}, t\bar{t}\}$$

†Taken from simulation

- Background with jet→τ fakes (QCD, W+jets, ttbar)

→ compute jet→τ mis-id probabilities (fake factors) in determination regions (DR)

→ weight events in application region (AR) with fake factors to estimate background in the signal region (SR)

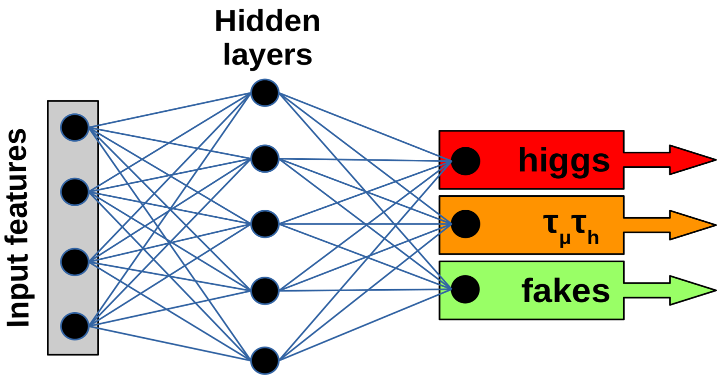
- **in total ~90% of background is estimated from data**

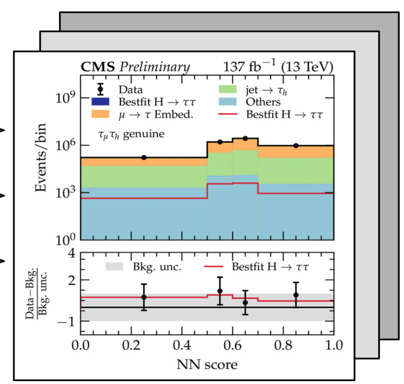
Signal vs background discrimination

- Multivariate discrimination between signal and background
  - kinematic properties of tau leptons, accompanying jets and missing transverse momentum are exploited
- Multi-classifiers are developed based on state-of-art ML techniques

  → $\tau_h \tau_h$: boosted decision trees implemented with XGBoost

  → $\tau_\ell \tau_h$: feed forward neural network
- Events classified into 3 categories
  - Higgs signal
  - events with genuine tau leptons
  - events with tau leptons mimicked by hadronic jets





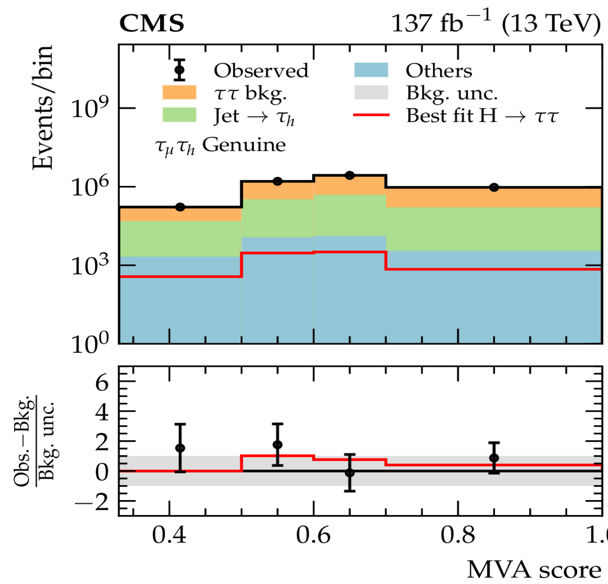
Observable	$\tau_\ell \tau_h$	$\tau_h \tau_h$
p_T of leading τ_h	✓	✓
p_T of trailing τ_h	—	✓
p_T of τ_ℓ	✓	—
p_T of visible di-τ	✓	✓
p_T of di-τ_h + p_T^{miss}	—	✓
p_T of $\tau_\ell \tau_h$ + p_T^{miss}	✓	—
Visible di-τ mass	✓	✓
Di-τ mass (using SVFIT)	✓	✓
Leading jet p_T	✓	✓
Trailing jet p_T	✓	
Jet multiplicity	✓	✓
Dijet invariant mass	✓	✓
Dijet p_T	✓	
Dijet $\|\Delta\eta\|$	✓	
p_T^{miss}	✓	✓

PhD thesis of Andrea Cardini

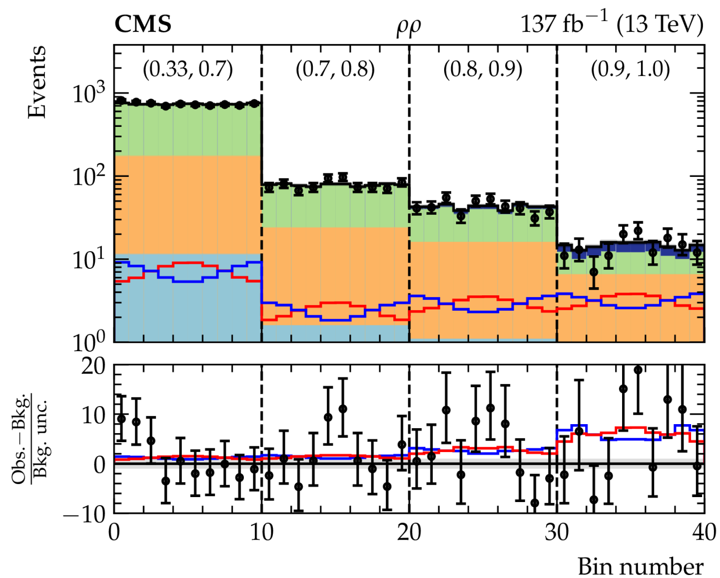
doi:10.3204/PUBDB-2021-03550

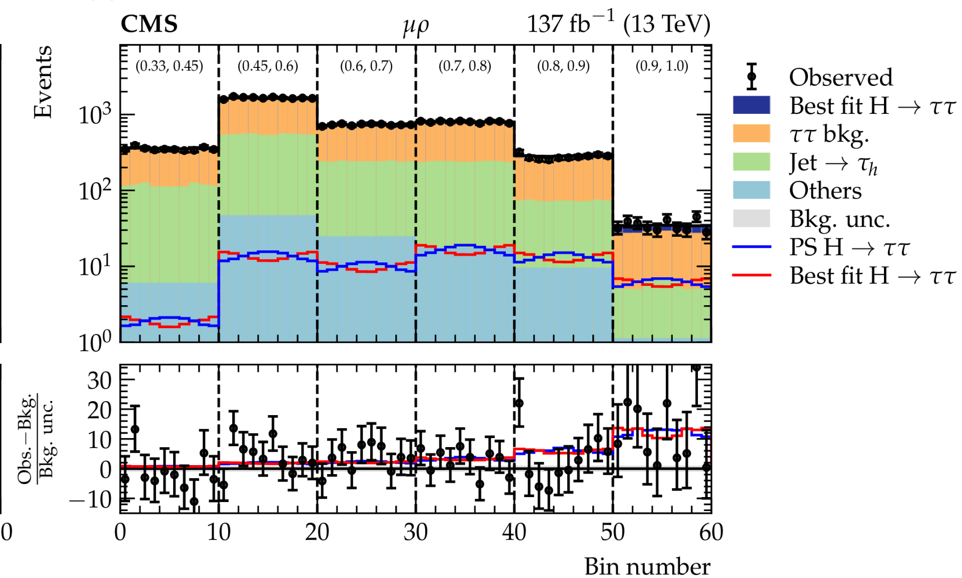
Statistical inference

- Simultaneous fit in all channels/categories:
  - **MVA score** distribution in background categories (constrains uncertainties in background estimates)
  - two-dimensional distribution of **MVA score** vs. $\boldsymbol{\phi_{CP}}$ in signal categories



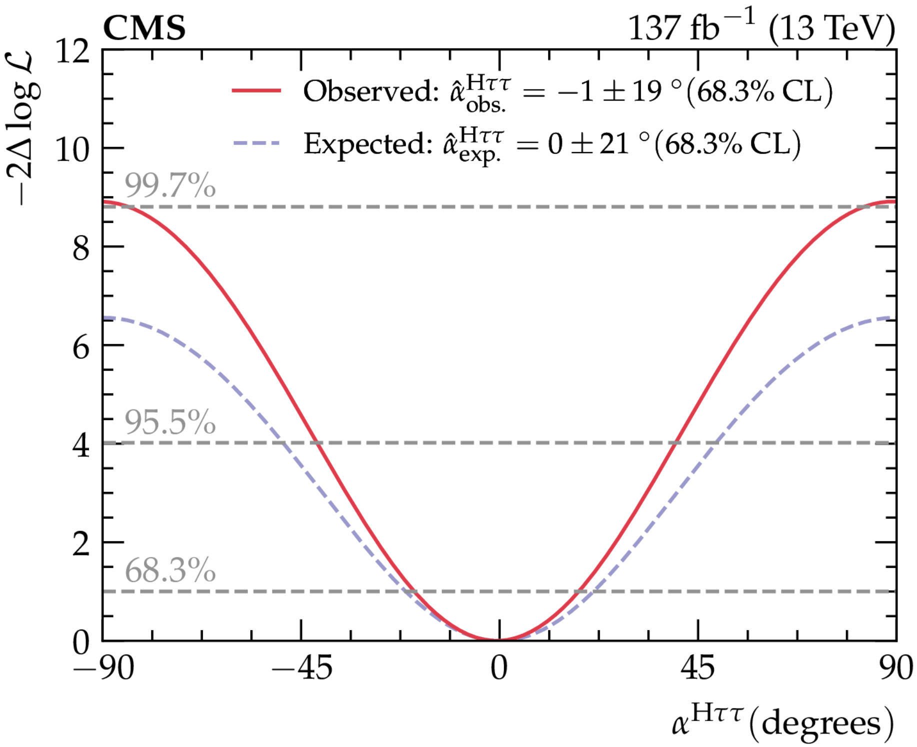
- freely floating parameters in the fit: $\alpha^{H\tau\tau}$ and signal strength modifiers μ_{ggH} and μ_{VH+qqH}

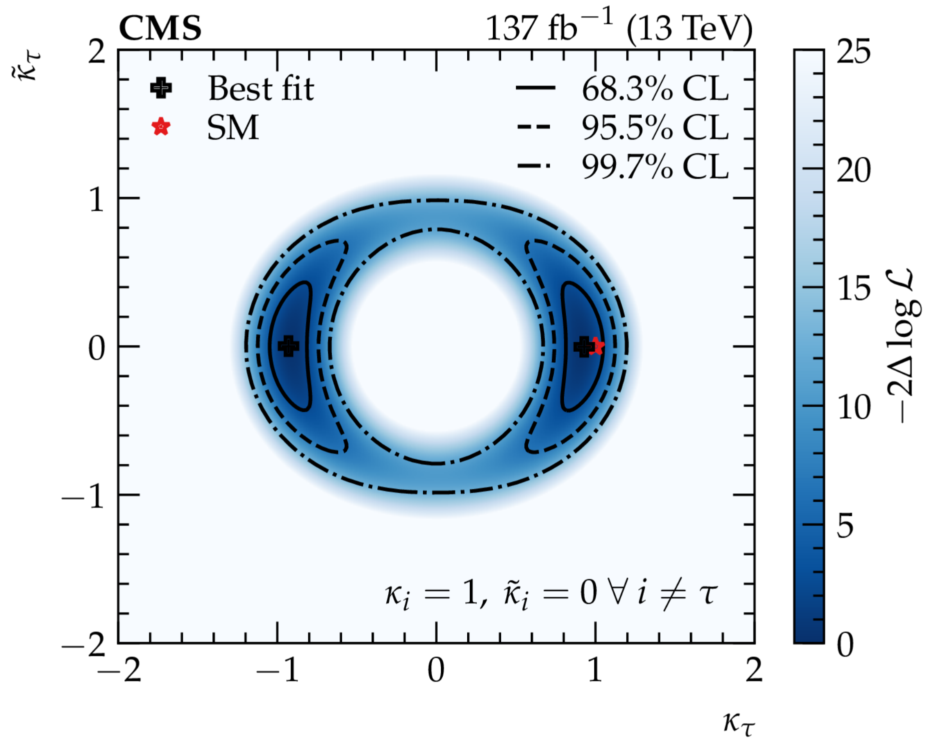




Results

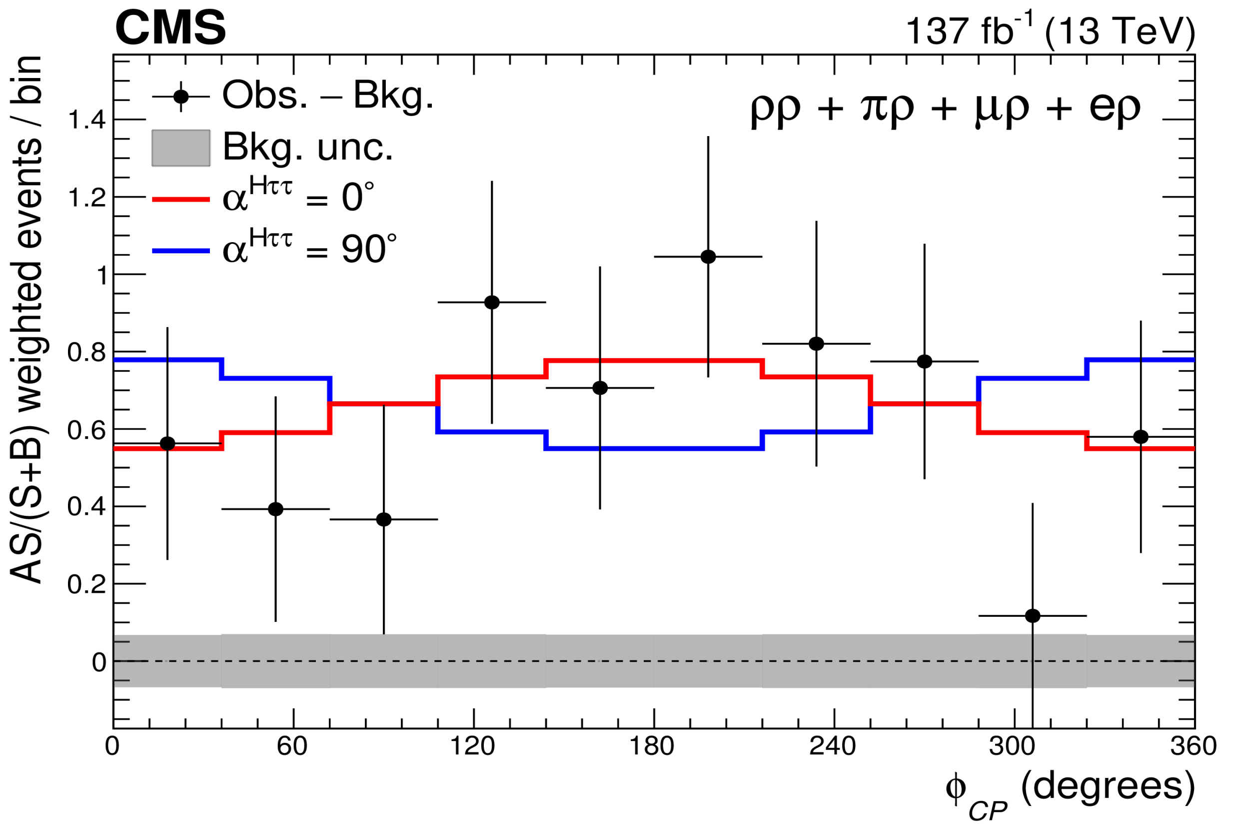
- Measured value of $\alpha^{H\tau\tau}$ = (-1±19(stat)±1(sys)±2(bbb)±1(theo))°
- Results are also interpreted in terms of constraints on $(\kappa_\tau, \tilde{\kappa}_\tau)$ assuming SM predictions for all other Higgs boson couplings





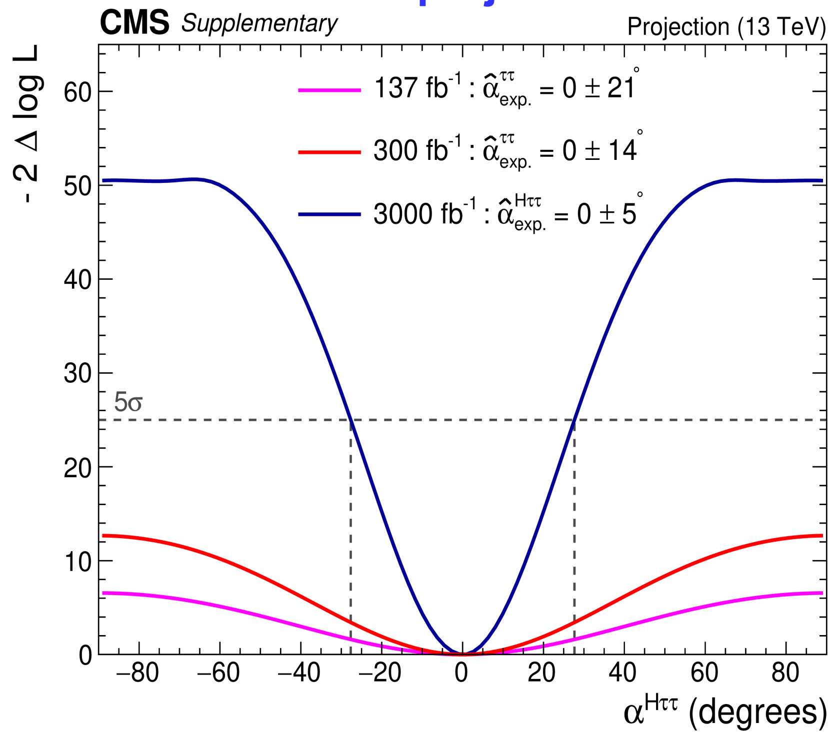
Combined ϕ_{CP} distribution

- To illustrate the results, four most sensitive channels (ρρ, πρ, μρ and eρ) are combined into a A×S/(S+B) weighted plot of ϕ_{CP}
- Data show clear preference for CP-even hypothesis
- Hypothesis of pure CP-odd coupling is rejected at 3σ (expected exclusion is 2.6σ)

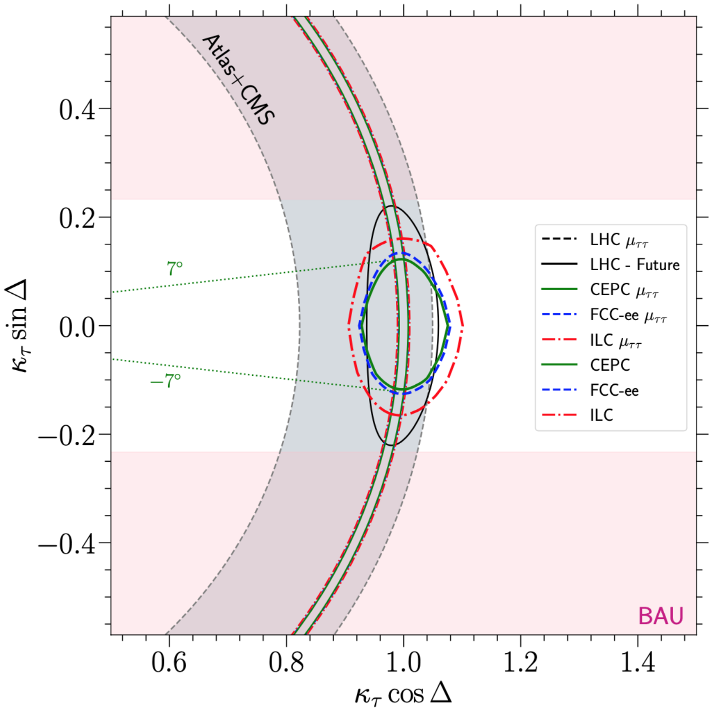


Projection of sensitivity

CMS projection



Phys. Rev. D 103, 095027 (2021)



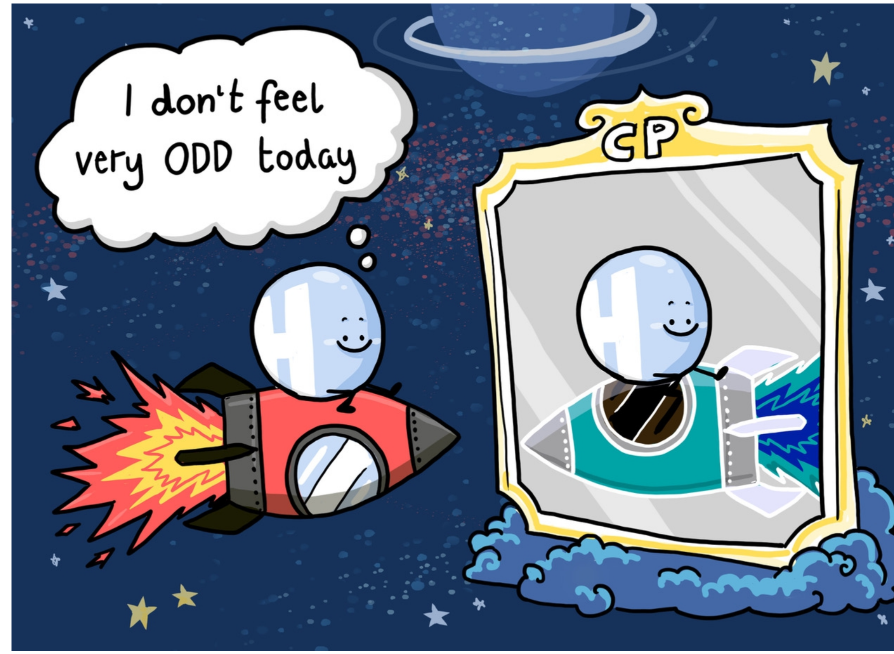
with additional LHC data and improved analysis techniques, precision of better than 10° in CP-mixing angle is plausible

→ good target for future analysis with Run 3 and HL-LHC data

→ sensitive probe of the lepton flavored electroweak baryogenesis models (see for example Phys. Rev. D 103, 095027 (2021), arXiv:1206.2942 and Phys. Rev. D 96, 115034 (2017))

Summary

- The CP structure of the Yukawa coupling between Higgs and tau leptons is investigated by the CMS collaboration with 137 fb^{-1} of LHC Run 2 data
- CP-mixing angle in the H$\tau\tau$ coupling measured to be **$\alpha^{H\tau\tau} = (-1 \pm 19)^{\circ}$** and is consistent with the SM prediction of pure CP-even coupling
- Hypothesis of pure CP-odd coupling is rejected at 3σ level
- Uncertainty of measurement is statistics dominated!



Courtesy of Anna Penkner and Renate Pommerening
Designdoppel GbR

- With additional LHC data expected in Run 3 and after HL upgrade, uncertainty in the CP-mixing angle can be reduced well below 10°
→ sensitive probe of the electroweak scale baryogenesis models

Backup

Combined ϕ_{CP} distribution

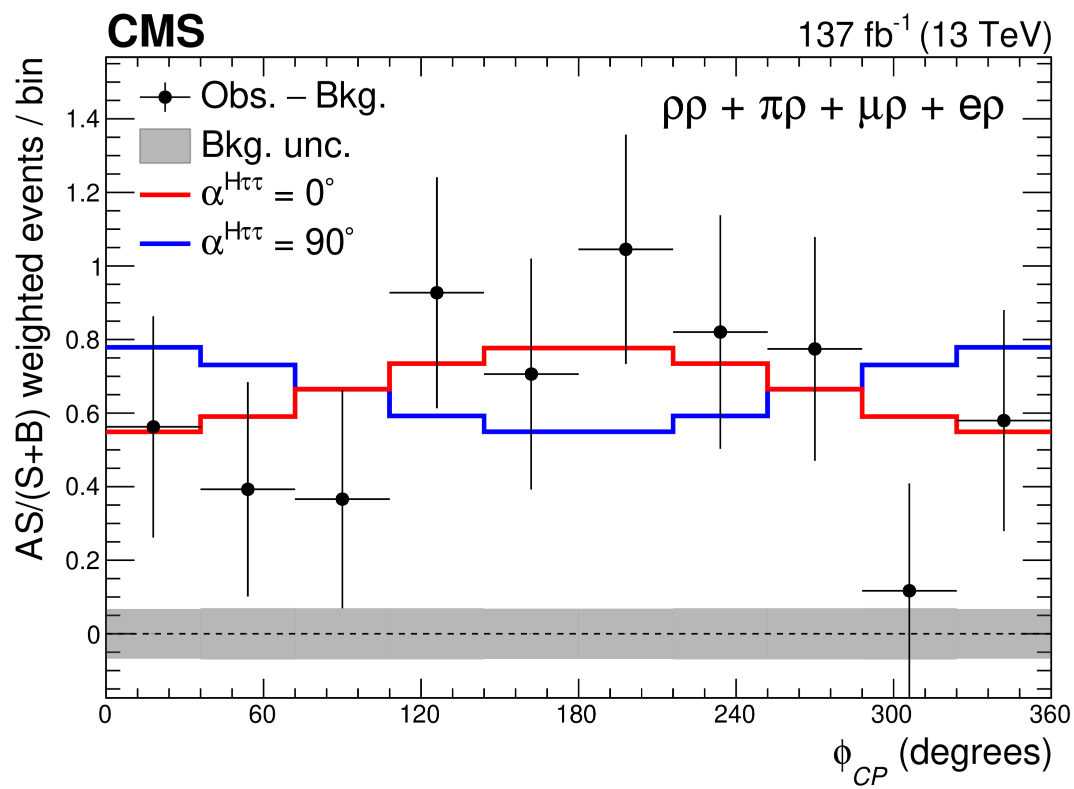
- To illustrate the results, four most sensitive channels (ρρ, πρ, μρ, eρ) are combined into a plot of ϕ_{CP}
- Each BDT/DNN score bin in 2D distribution is weighted by

$$A \times S/(S+B)$$

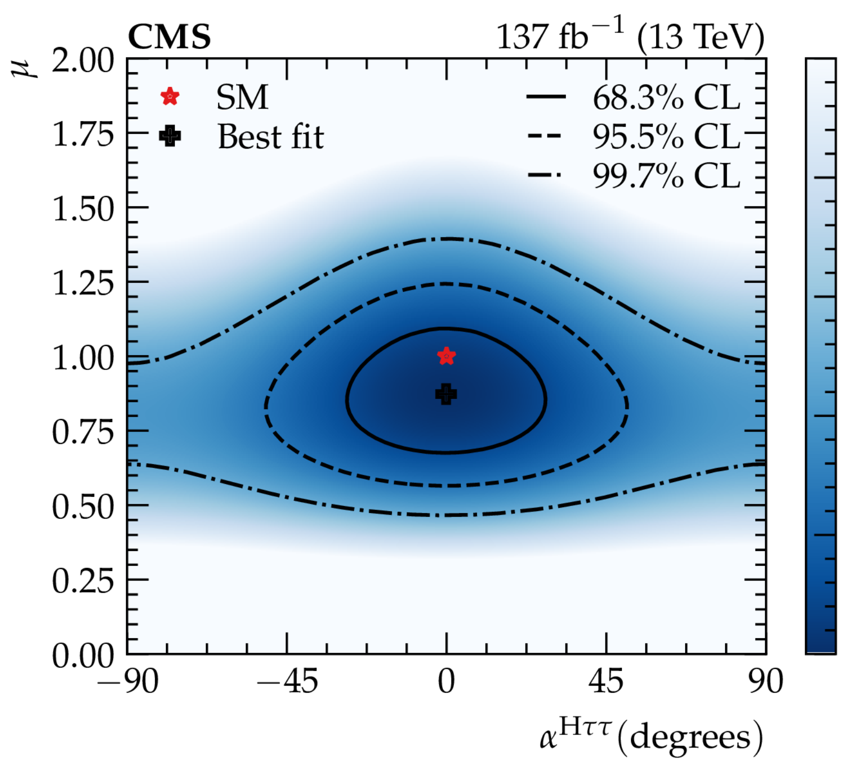
S : expected signal yield (in a given BDT/NN bin)
B : expected background yield
A : average asymmetry

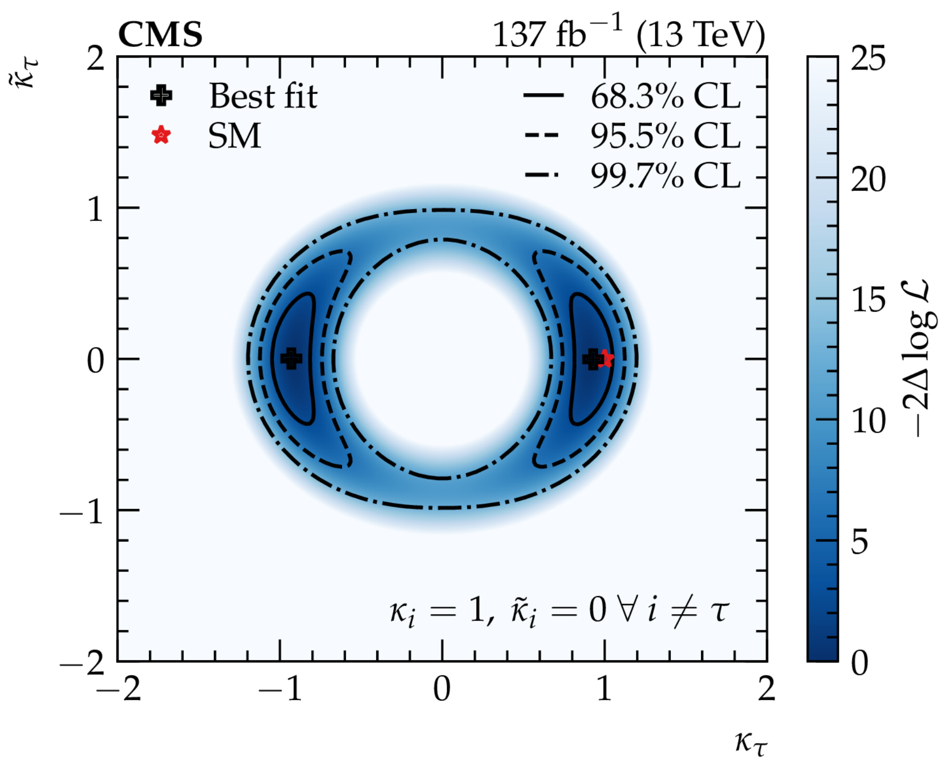
$$A = \frac{1}{N_{\text{bins}}} \sum_{\text{bins}} \frac{|N^{\text{CP-even}} - N^{\text{CP-odd}}|}{N^{\text{CP-even}} + N^{\text{CP-odd}}}$$

- Data show clear preference for CP-even hypothesis
- Hypothesis of pure CP-odd coupling is rejected at 3σ (expected exclusion is 2.6σ)



Two-dimensional likelihood scans





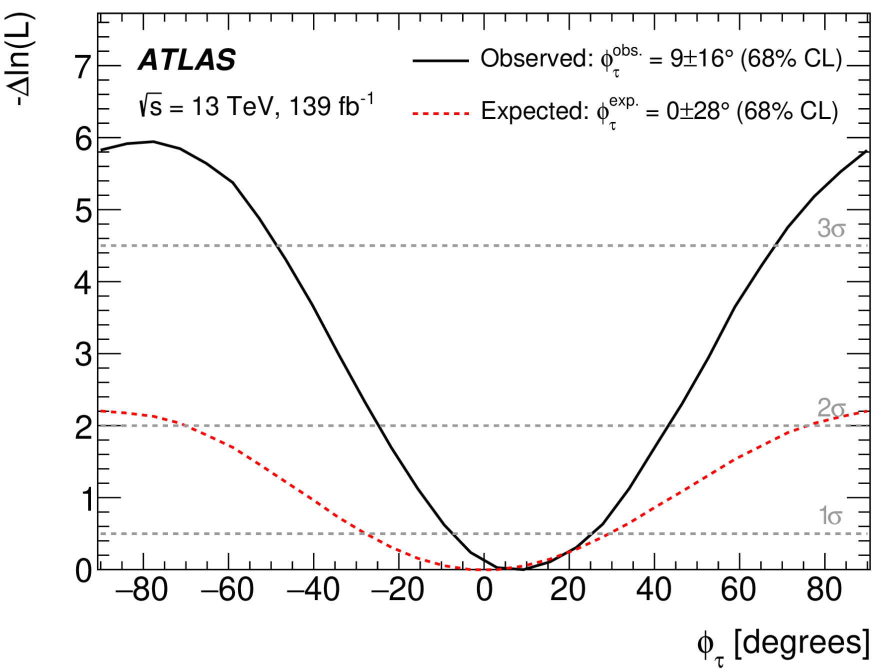
- Results of the analysis are interpreted in terms of 2D scans
  - H→ττ signal strength modifier (μ) vs CP mixing angle ($\alpha^{H\tau\tau}$)
  - CP-odd ($\tilde{\kappa}_\tau$) vs CP-even (κ_τ) components of the Hττ Yukawa coupling
- best fit values are consistent with the SM predictions within 1 st.d.

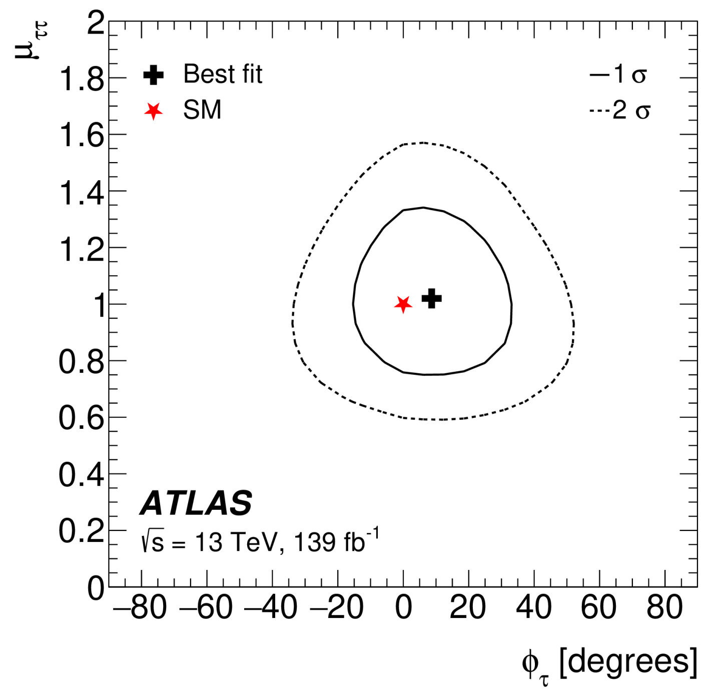
CP H→ττ analysis by ATLAS

- CP H→ττ analysis in ATLAS (**Eur. Phys. J. C 83 (2023)**) is similar in spirit to the CMS analysis
  - di-tau decay modes considered:

$$\{e^{\pm}, \mu^{\pm}, \pi^{\pm}, \rho^{\pm}, a_1^{\pm}\} \times \{\pi^{\mp}, \rho^{\mp}, a_1^{\mp}\}$$

  - dedicated event categorization targeting VBF and boosted production mechanisms





Combined ϕ_{CP} distribution by ATLAS

- Observed results in the ATLAS H analysis are illustrated by the combined postfit distribution of CP observable
  - Events are weighted by log(1+S/B) of the respective event categories and decay modes