Analysis of the CP structure of the Higgs-tau Yukawa coupling

Merijn van de Klundert*, Elisabetta Gallo, Alexei Raspereza, Andrea Cardini, Oleg Filatov, Mareike Meyer

HELMHOLTZ RESEARCH FOR GRAND CHALLENGES

DESY.

Introduction and state-of-the-art

Scalar and pseudoscalar bosons

- Couplings of (speudo) scalar boson to fermions or gauge bosons can be:
  - CP-even (scalar), $J^{CP}=0^{++}$
  - CP-odd (speudoscalar), $J^{CP}=0^{+-}$
  - Mixture of even and odd couplings
- Standard model: all Higgs couplings (to fermions and bosons) are **CP-even**
  - Any deviation **unambiguous** sign new physics!
- Pure CP-odd couplings to gauge bosons excluded. Non-exhaustive refs:
  - arXiv:1903.06973, arXiv:1712.02304
- Note: CP-odd couplings only occur at tree level for fermion couplings!
  - Higgs-top Yukawa coupling: pure CP-odd coupling excluded with ~3 sigma (CMS and ATLAS) arXiv:2003.10866 arXiv:2004.04545
  - Higgs-tau: **first analysis of CP structure of Yukawa coupling by CMS!**
    - HIG-20-006, https://cds.cern.ch/record/2725571
    - More references to literature on above measurements

Theory background

ϕ_{cp}: angle between tau decay plane

- Parameterise CP even and odd couplings via mixing angle $\phi_{\tau\tau}$:

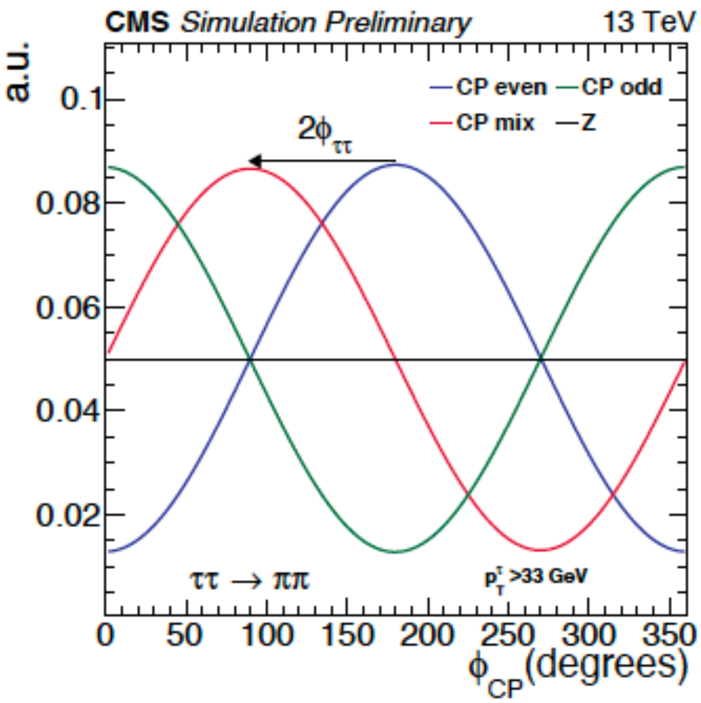
$$\mathcal{L}_Y = -\frac{m_\tau H}{v}\left(\kappa_\tau \bar{\tau}\tau + \tilde{\kappa}_\tau \bar{\tau} i\gamma_5 \tau\right) \quad \tan(\phi_{\tau\tau}) = \frac{\tilde{\kappa}_\tau}{\kappa_\tau},$$

- The CP information is transferred to correlations between transversal components tau spin

$$d\Gamma_{h\to\tau^+\tau^-} \sim 1 - s_z^- s_z^+ + |s_T^-||s_T^+|\cos(\varphi_s - 2\phi_\tau)$$

- This correlation can be probed via the angle ϕ_{cp} between the **tau decay planes**

$$d\Gamma_{h\to\tau^+\tau^-} \approx 1 - b(E_+)\, b(E_-)\frac{\pi^2}{16}\cos(\varphi^*_{CP} - 2\phi_\tau)$$



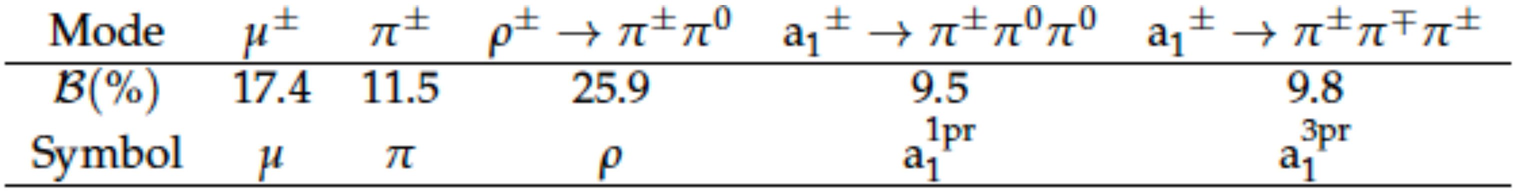
- Gen level distribution ϕ_{cp} for scalar, pseudoscalar and Z boson

⇒ **ϕ_{cp} discriminating variable for this analysis!**

Analysis strategy

In a nuthsell

- Utilise full Run 2 data set (137 fb^{-1})
- For PAS HIG-20-006 analyse most important decay modes (~50%)
  - Muon plus hadronic
  - Fully hadronic



Mode	$\mu^\pm$	$\pi^\pm$	$\rho^\pm \to \pi^\pm\pi^0$	$\mathrm{a}_1{}^\pm \to \pi^\pm\pi^0\pi^0$	$\mathrm{a}_1{}^\pm \to \pi^\pm\pi^\mp\pi^\pm$
$\mathcal{B}(\%)$	17.4	11.5	25.9	9.5	9.8
Symbol	μ	π	ρ	$\mathrm{a}_1^{1\mathrm{pr}}$	$\mathrm{a}_1^{3\mathrm{pr}}$

- Background extraction
- Signal-background distinction using machine learning methods
- Extract mixing angle $\phi_{\tau\tau}$ via combined template fit to signal and background distributions
- **First** review parts analysis not specific to CP-analysis

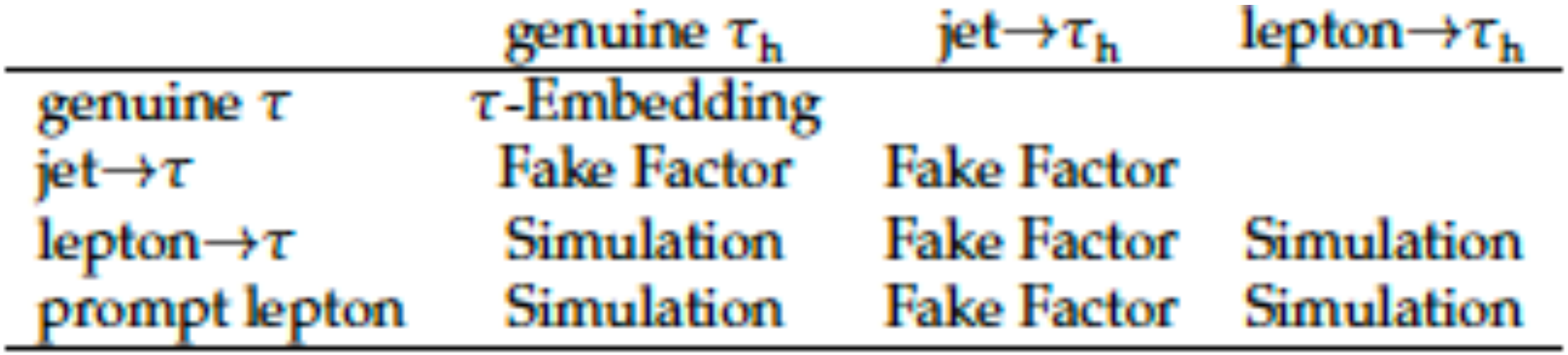
  => Analysis indebted to work STXS analysis **HIG-19-010** (full Run 2) and **HIG-18-032** (‘16/’17 with embedding and Machine learning techniques)
- **Next** focus on dedicated methods developed to optimise signal strength

Modelling background processes

90% of backgrounds obtained in data-driven way

- Background processes:
  - Drell-Yan (leading semileptonic BG), QCD (leading hadronic BG), W+jets, ttbar, single top, diboson
- May also categorise backgrounds in genuine tau, jet-fakes, lepton-fakes, and prompt leptons
  - Backgrounds with 2 genuine taus obtained from tau embedding technique
  - Bg with QCD jet faking hadronic tau via fake-factor method
  - Remaining backgrounds via simulation

⇒ Overall, 90% of backgrounds obtained in data-driven manner!

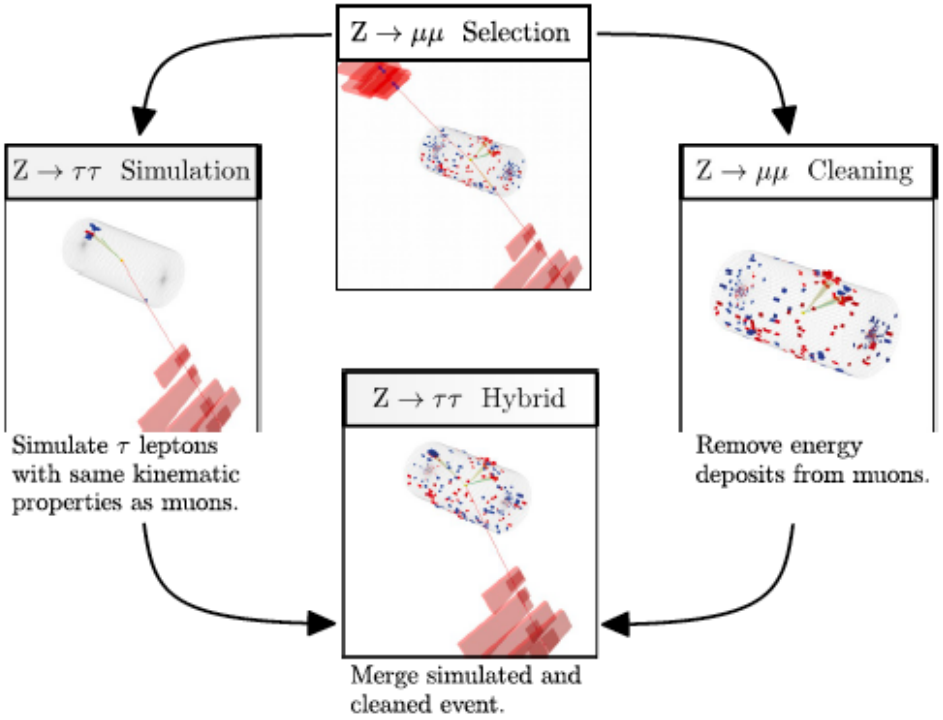


	genuine τ_h	jet$\to\tau_h$	lepton$\to\tau_h$
genuine τ	τ-Embedding		
jet$\to\tau$	Fake Factor	Fake Factor	
lepton$\to\tau$	Simulation	Fake Factor	Simulation
prompt lepton	Simulation	Fake Factor	Simulation

Modelling background processes

Tau embedding, arXiv:1903.01216

- Tau embedding relies on principle **lepton universality EWK processes**
- Exploit principle to model **genuine di-tau background:**
  - Select di-muon events in real data
  - Remove hits associated to muons
  - Simulate decaying Z-boson to **di-tau** with **identical kinematics as di-muon pair** (pt, invariant mass, eta-phi). **In empty detector**
  - Add the hits of tau decay products to the data event
- Obtain rate genuine tau background events, with fully data-driven underlying event!



Simulate τ leptons with same kinematic properties as muons.

Remove energy deposits from muons.

Merge simulated and cleaned event.

Modelling background processes

Fake factor method, https://arxiv.org/abs/1801.03535

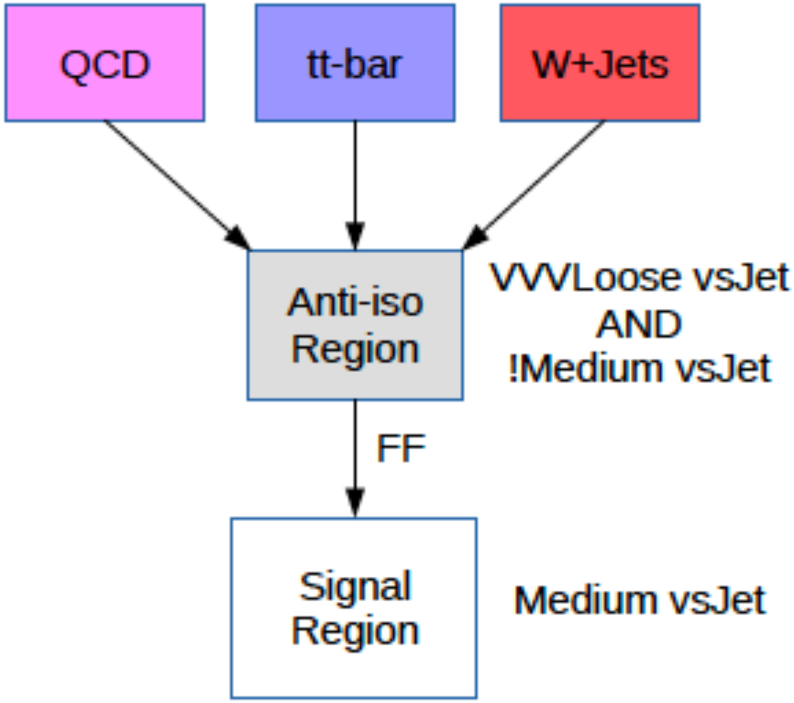
- Fake factor: data-driven approach to obtain contribution in signal region of **quark or gluon jet faking a tau lepton**

  - Define jet-enriched determination region, orthogonal to signal region
  - Determine rate of jets faking a hadronic tau lepton using tight and loose tau isolation criteria

  ⇒ Apply the fake factors to loose tau candidates in application region to obtain jet-fake rate!

  - Hadronic channel: QCD background only. Semileptonic: weighted average for QCD, ttbar and W+jets

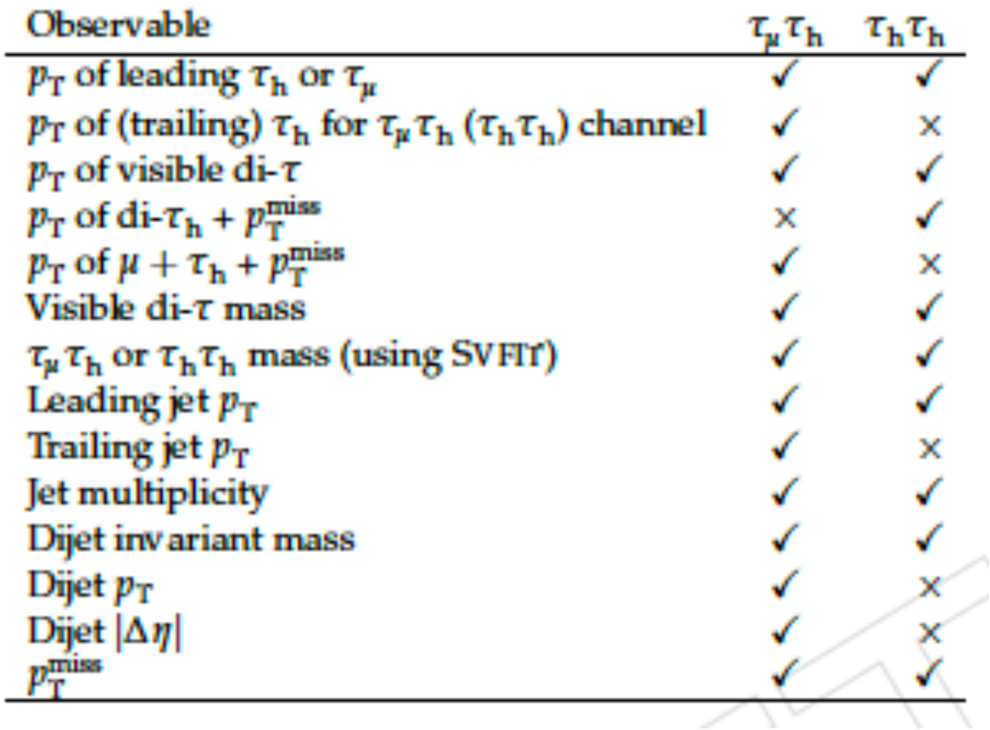
- Obtain rate of jet faking tau, with fully data-driven underlying event



Event categorisation

Separating signal from background events

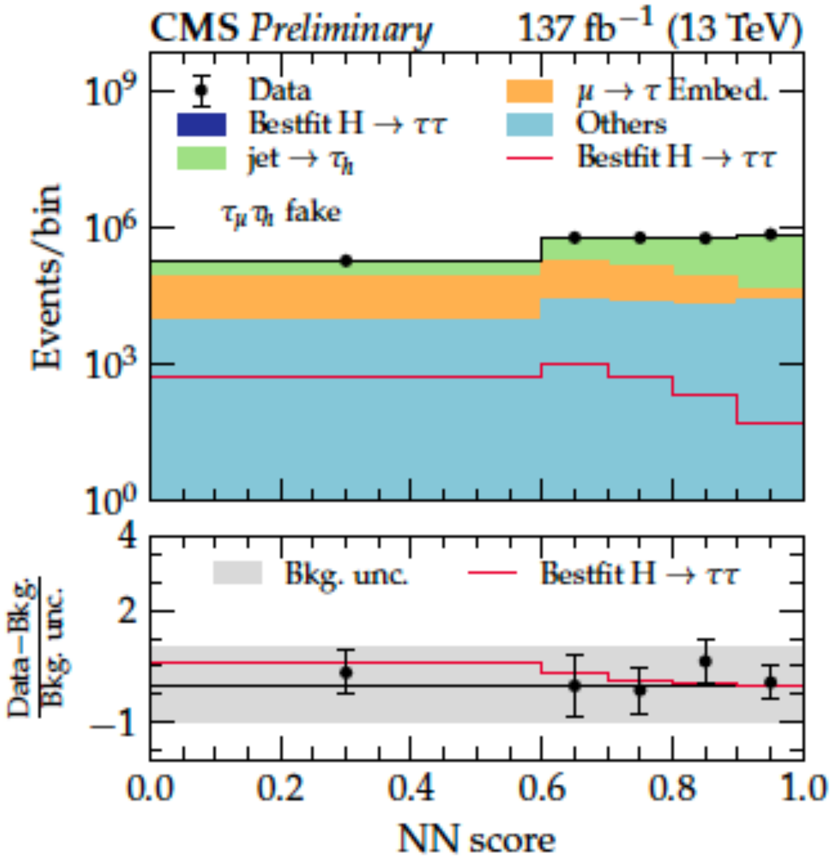
- Mu+tau channel: use neural net
- Hadronic channel: use Boosted Decision tree
- Input variables as in PAS-18-032
- Categorise events in 3 mutually exclusive cats:
  - Signal (ggH, VBF, VH)
  - Genuine tau pair
  - Jet fake (inc. prompt leptons and leptons faking hadronic tau)



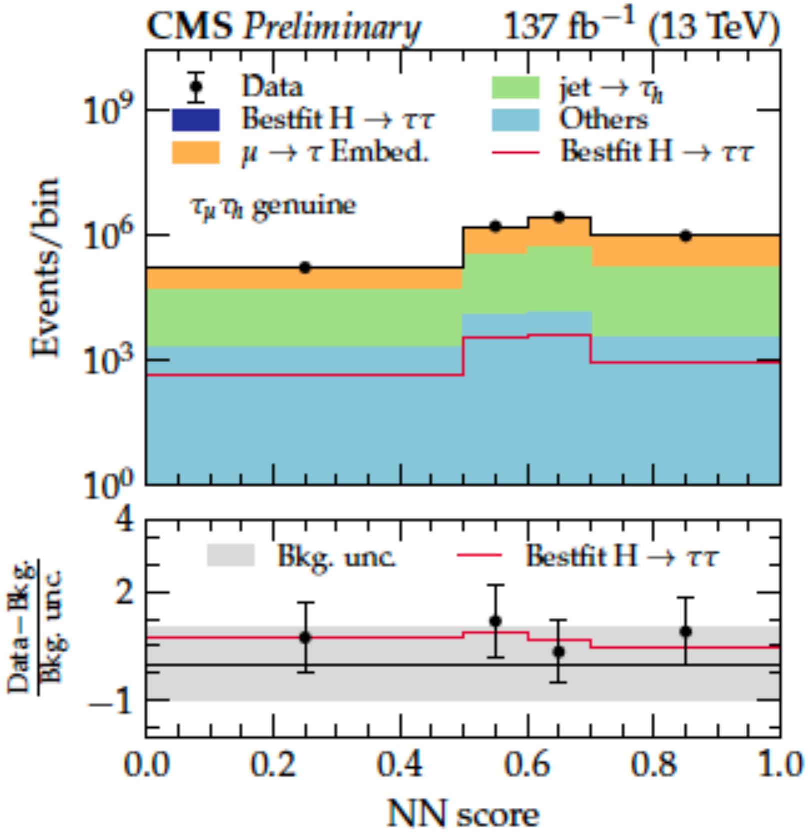
Observable	$\tau_\mu\tau_h$	$\tau_h\tau_h$
p_T of leading τ_h or τ_μ	✓	✓
p_T of (trailing) τ_h for $\tau_\mu\tau_h$ ($\tau_h\tau_h$) channel	✓	×
p_T of visible di-τ	✓	✓
p_T of di-τ_h + p_T^{miss}	×	✓
p_T of $\mu + \tau_h + p_T^{miss}$	✓	×
Visible di-τ mass	✓	✓
$\tau_\mu\tau_h$ or $\tau_h\tau_h$ mass (using SVFIT)	✓	✓
Leading jet p_T	✓	✓
Trailing jet p_T	✓	×
Jet multiplicity	✓	✓
Dijet invariant mass	✓	✓
Dijet p_T	✓	×
Dijet $\lvert\Delta\eta\rvert$	✓	×
p_T^{miss}	✓	✓

Event categorisation

Left: jet fakes



Right: genuine tau pair category

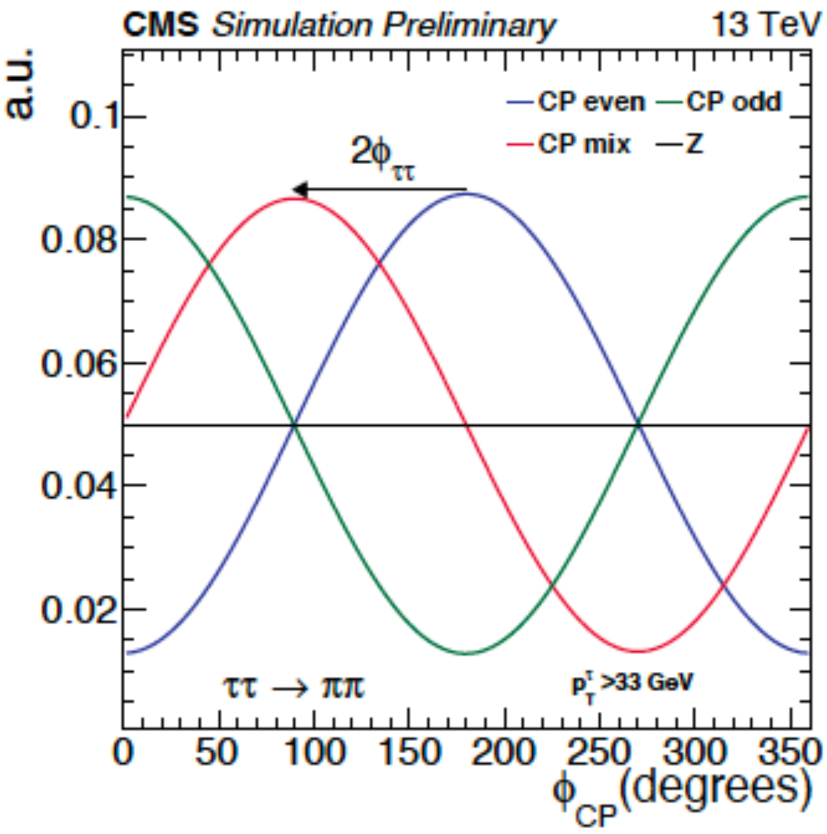


Extracting and optimising sensitivity

- Brief recap:
  - A measurement of ϕ_{CP} is sensitive to $\phi_{\tau\tau}$.
  - We follow background treatment and event categorisation largely conform the STXS measurement

⇒ Next: review how **experimentally** assess ϕ_{CP}

⇒ Methods dedicatedly developped for analysis to **optimise analysis sensitivity**

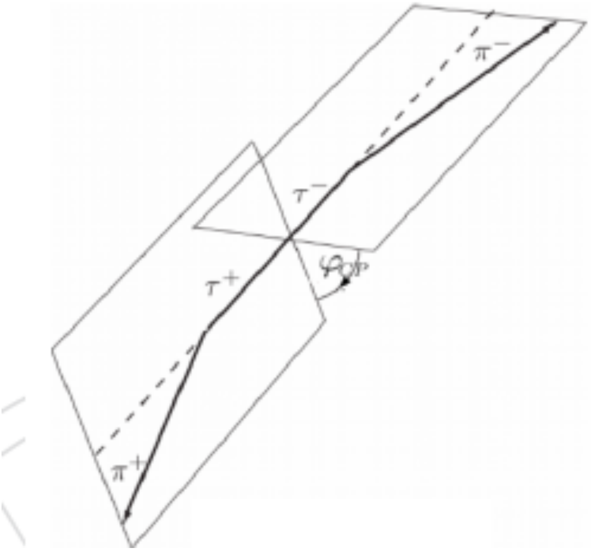


Experimentally extracting $\phi_{\tau\tau}$

Extracting tau decay planes

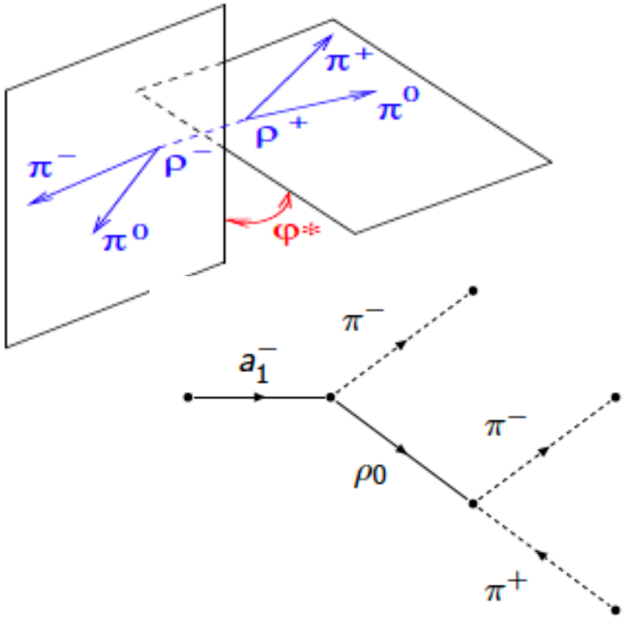
Mode	$\mu^\pm$	$\pi^\pm$	$\rho^\pm \to \pi^\pm\pi^0$	$a_1{}^\pm \to \pi^\pm\pi^0\pi^0$	$a_1{}^\pm \to \pi^\pm\pi^\mp\pi^\pm$
$\mathcal{B}(\%)$	17.4	11.5	25.9	9.5	9.8
Symbol	μ	π	ρ	a_1^{1pr}	a_1^{3pr}

- Decay to muon or single charged pion
  - Use Impact parameter method



- IP method: arXiv:1108.0670

- Decay to rho, $a_1{}^{1p}$, $a_1{}^{3p}$
  - Use neutral pion method



- DP Method: arXiv:0307.331

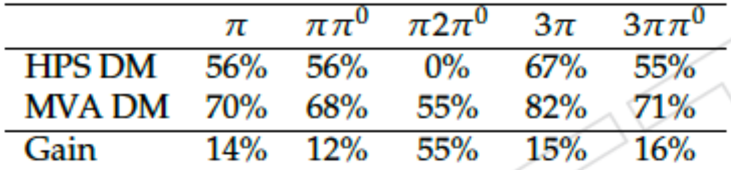
Optimisation analysis sensitivity

Focus on additional corrections applied to optimise signal strength

- Decay mode identification important, migrations will lead to incorrect ϕ_{cp} estimates
- Per default, decay mode given by HPS (hadron-plus-strip) algorithm
- Dedicated MVA developpd for enhanced decay mode distinction (on top of deeptau discriminant,

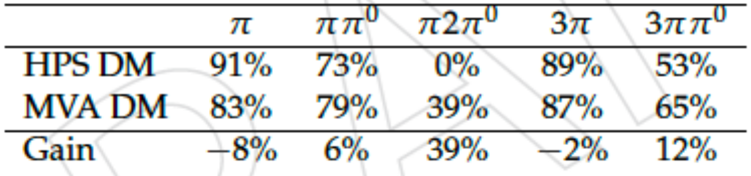
  CMS-DP-2019-033)
  - Inputs: kinematics tau decay products and HPS decay mode
- Substantial gain in purity and efficiency
- Improves signal sensitivity by O(15%)

Table 17: Comparing purity of the HPS and MVA DM.



	π	$\pi\pi^0$	$\pi 2\pi^0$	3π	$3\pi\pi^0$
HPS DM	56%	56%	0%	67%	55%
MVA DM	70%	68%	55%	82%	71%
Gain	14%	12%	55%	15%	16%

Table 18: Comparing efficiency of the HPS and MVA DM.

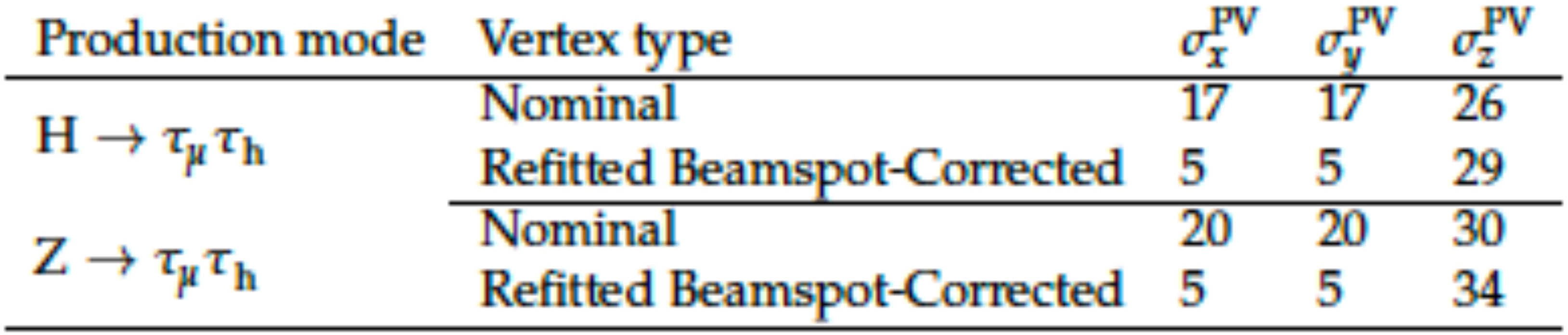


	π	$\pi\pi^0$	$\pi 2\pi^0$	3π	$3\pi\pi^0$
HPS DM	91%	73%	0%	89%	53%
MVA DM	83%	79%	39%	87%	65%
Gain	−8%	6%	39%	−2%	12%

Improvements related to IP method

Improvements in Primary Vertex (PV) estimates

- Two improvements in determination PV location:
  - Remove tracks associated to tau decay products. If boosted Higgs, non-zero impact parameters may pull PV
  - Add beam spot information in fit of PV
- Resolution in transversal plane increases by factor 3 (!)



Production mode	Vertex type	σ_x^{PV}	σ_y^{PV}	σ_z^{PV}
$H \to \tau_\mu \tau_h$	Nominal	17	17	26
	Refitted Beamspot-Corrected	5	5	29
$Z \to \tau_\mu \tau_h$	Nominal	20	20	30
	Refitted Beamspot-Corrected	5	5	34

Helical extrapolation

3-d vs 2-d extrapolation tracks

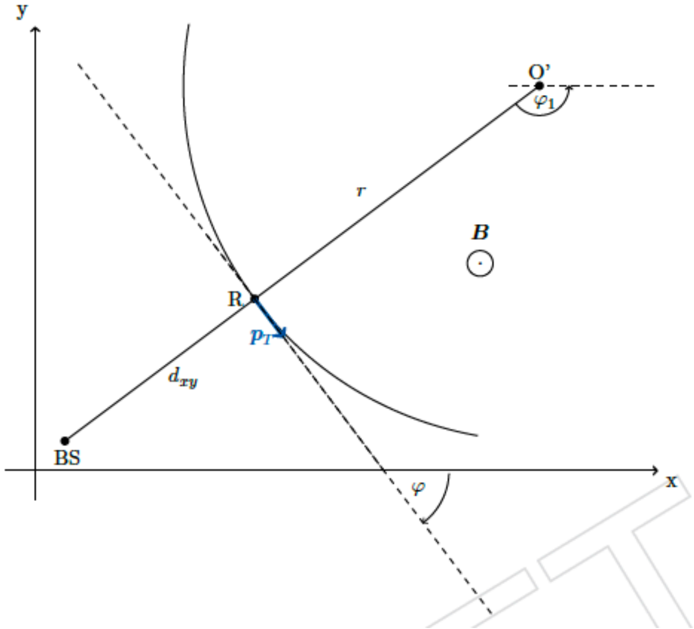
- Per default, track extrapolation to find PCA (point closest approach) performed in trasnversal plane

- Using helical, 3-dimensional approach has 2 profound advantages:
  - IP estimate better for tracks with high eta values
  - Can propagate uncertainties in track and PV in consistent manner

  ⇒ Define an impact-parameters significance as |IP|/sigma(IP)

  ⇒ throughout analysis require
  |IP|/sigma IP> 1.5

- Lead to improvement sensitivity O(15%)



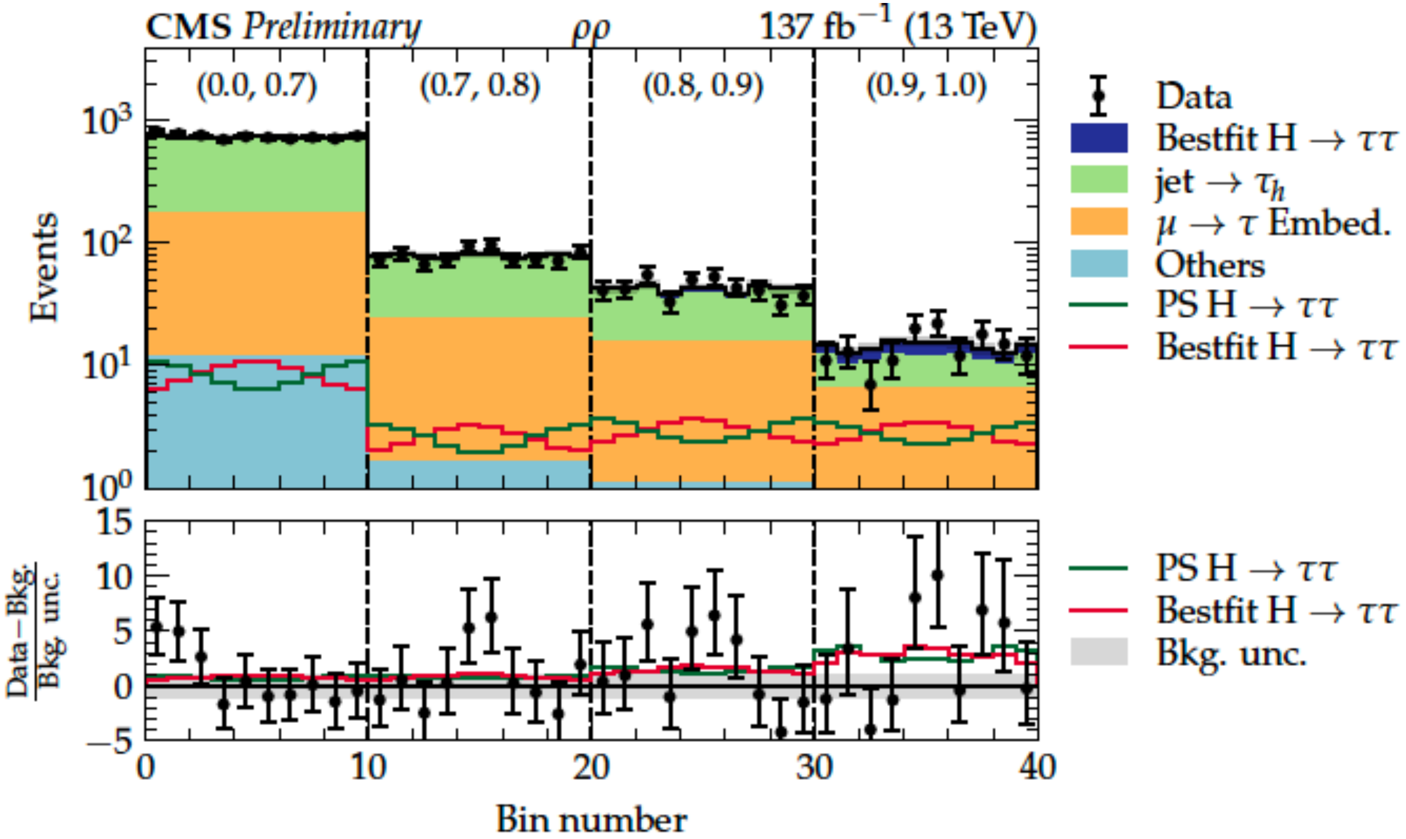
More details in PAS,
https://cds.cern.ch/record/2725571

Unrolled phi-CP distributions

rho+rho channel. Resolution ~ 1.1 sigma

- Observe s/b improvement owing to BDT
- Backgrounds with genuine taus expected to be flat in ϕ_{cp}

$\Rightarrow$ Enhance sensitivity by merging bins. Jet fakes: symmetrise around $\phi_{cp}=\pi$

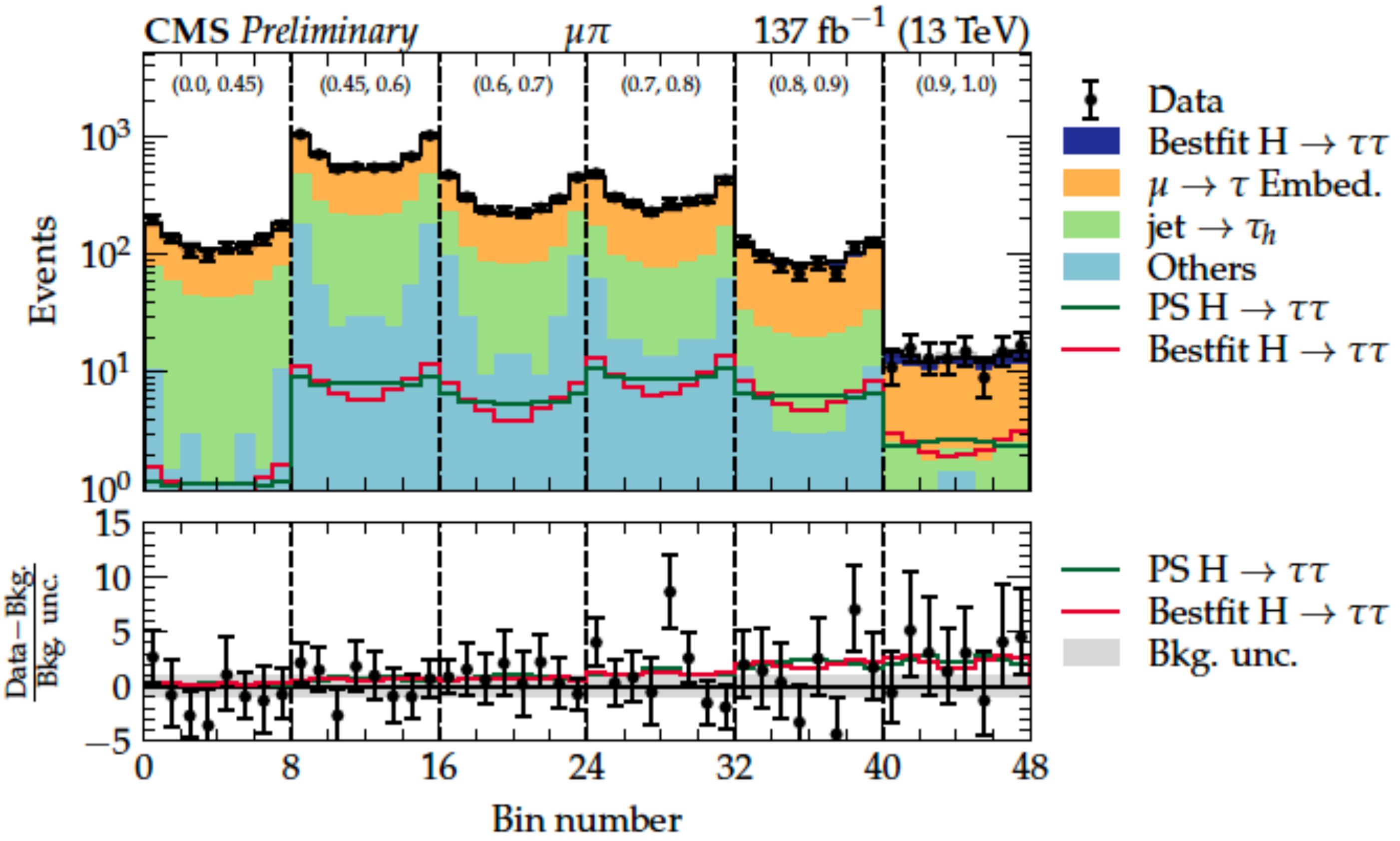


Unrolled phi-CP distributions

mu+pi channel. Resolution: ~1.0 sigma

- Using IP method twice results in correlated PV smearing effects

  ⇒ Only symmetrise bins in $\phi_{cp}=\pi$



Results

Combined Negative log-likelihood fit

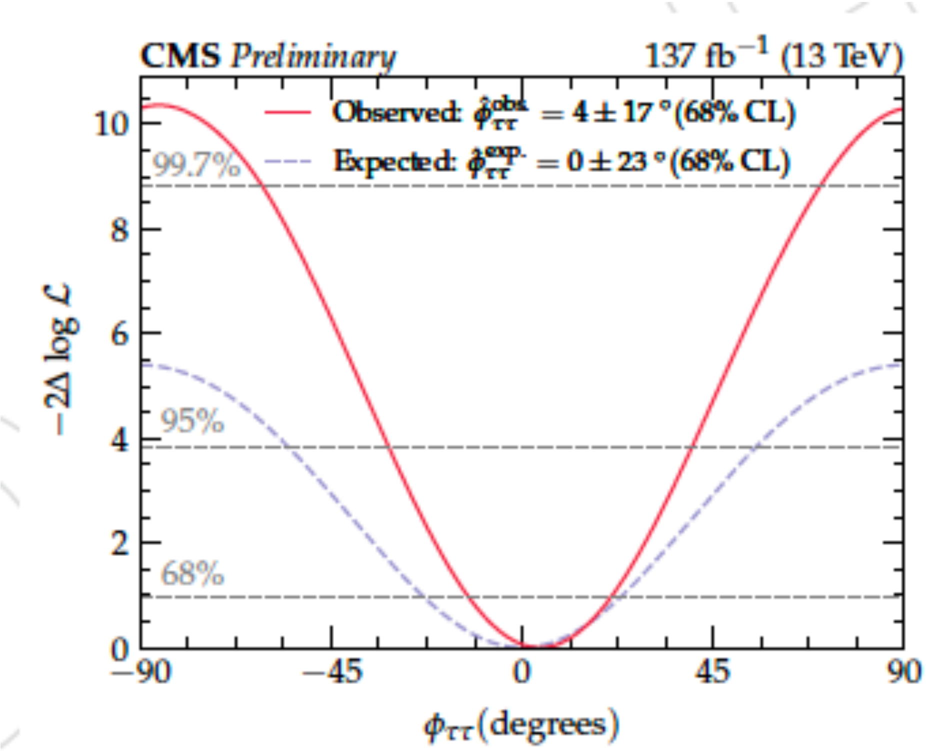
- Obtain observed (expected) sensitivity of **3.2 (2.3) sigma**
- Mixing angle (68% CL level)

$$\phi_{\tau\tau} = (4 \pm 17\,(\text{stat}) \pm 2(\text{bin-by-bin}) \pm 1\,(\text{syst}) \pm 1(\text{theory}))^{\circ}.$$

$\Rightarrow$ Fully consistent with SM

$\Rightarrow$ Excludes part nMSSM phase space ($\phi_{\tau\tau} <= 27^{\circ}$)

- Measurement **statistically dominated**



Outlook analysis

- Outlook for Run-2 paper:
  - Add e+t channel (~10% expected enhancement)
  - Replace neutral pion with polarimetric method for 3p*3p channel
- Run 3/phase-2: apply regressive machine learning algorithms for ϕ_{CP} determination

Event selection

Kinematic cuts in a birds eye view:

- **Muon-hadronic channel**
  - Utilise single-muon trigger and a paired muon-tau trigger
  - Offline muon: minimal pt of 20 (21) GeV in 2016 (2017, 2018)
  - Offline tau: pt > 25 (32) GeV in 2016 (2017, 2018)
  - Muon: eta< 2.1. Tau: eta< 2.3
  - Cut on transverse mass to suppress W+Jet background:

$$m_{T} \equiv \sqrt{2 p_{T}^{\mu} p_{T}^{\mathrm{miss}} [1 - \cos(\Delta\phi)]} < 50\,\mathrm{GeV},$$

  - Veto b-jets

- **Hadronic channel**
  - Use hadronic di-tau trigger
  - Offline, select tau pairs with pt> 40 GeV
  - Eta<2.1

Common selections:

- Loose cut on impact parameters muon and tau
- Multiple pairs: select pair with most isolated objects
- Further selections and cuts in PAS…

2-dimensional scan of reduced couplings

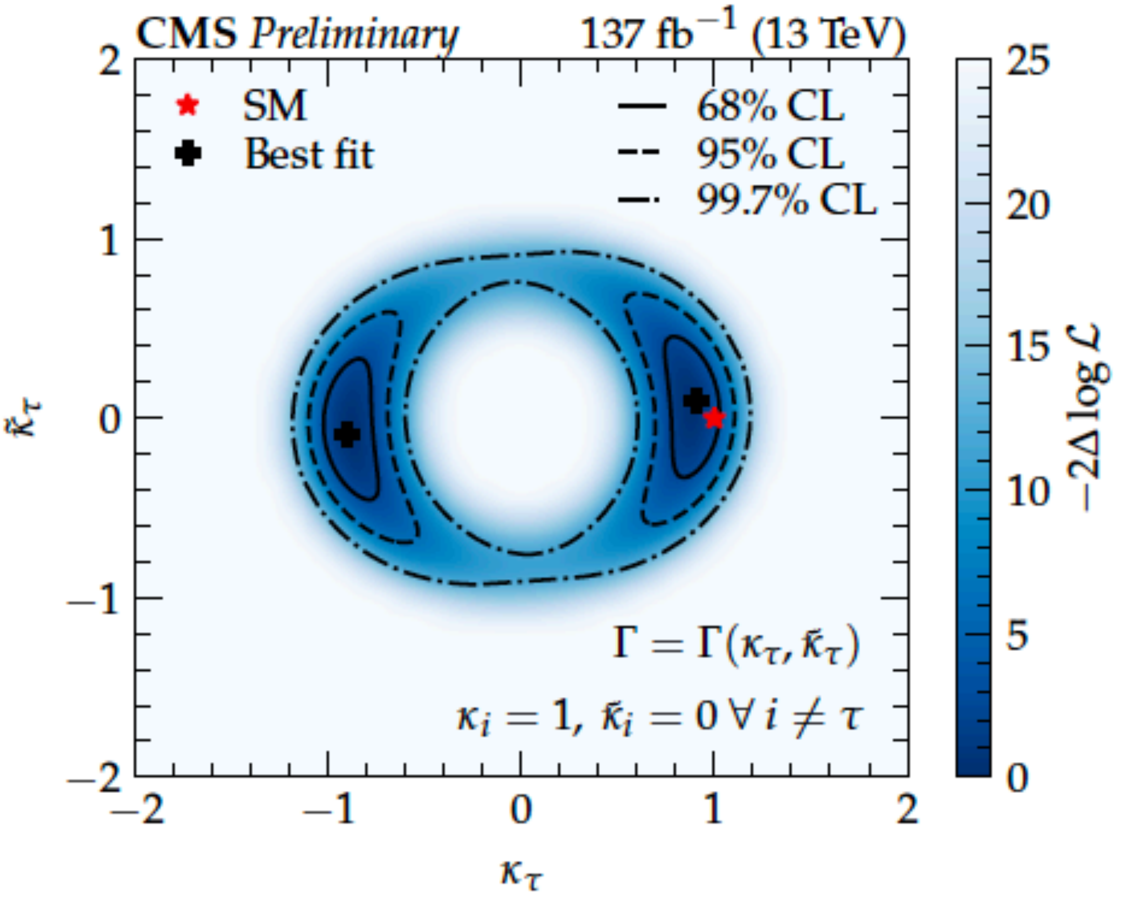
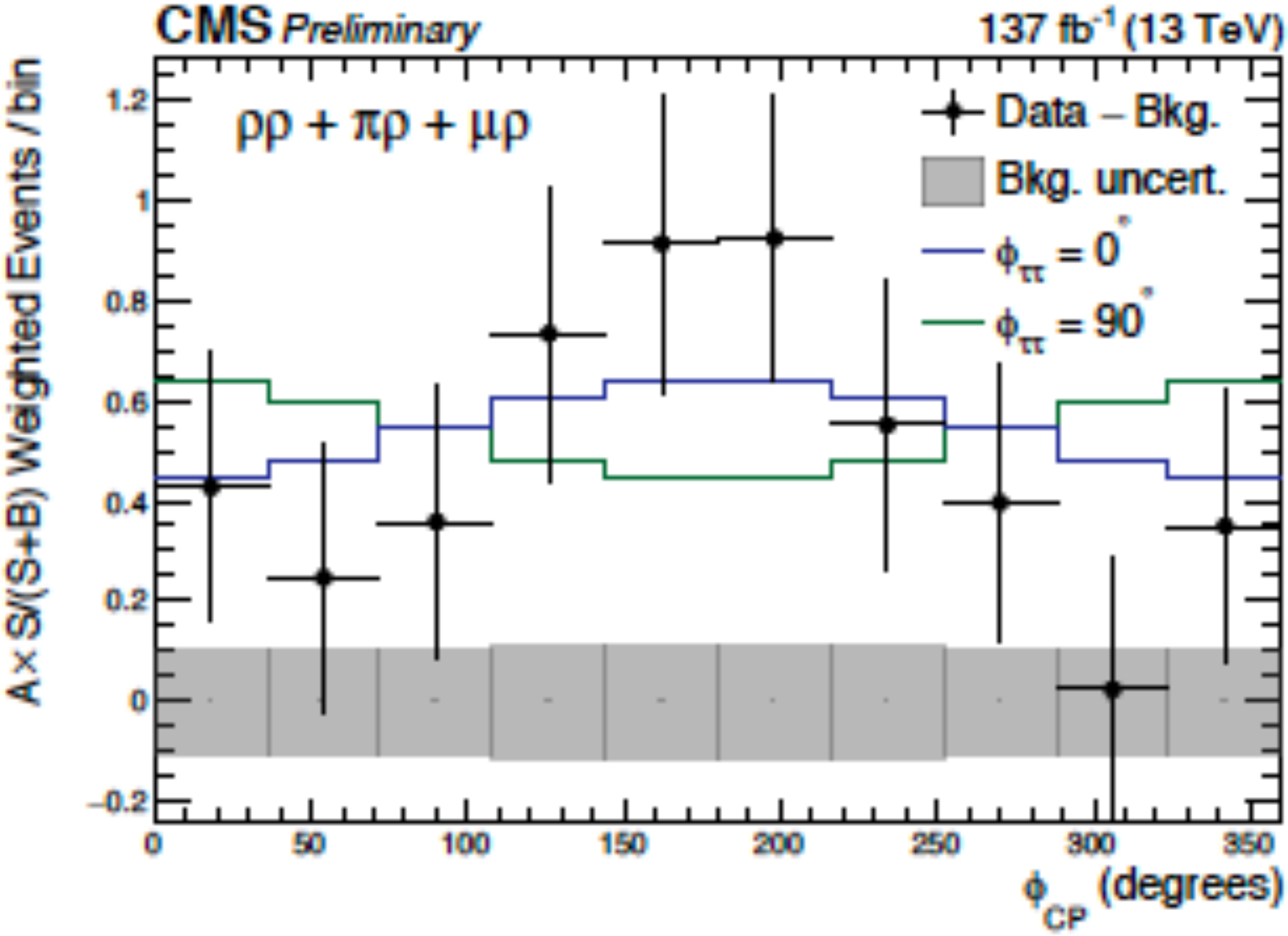


Figure 9: Two-dimensional scan of the (reduced) CP-even (κ) and CP-odd ($\tilde{\kappa}$) τ Yukawa couplings.

“propaganda plot”

Provides explanation observed significance higher than expected

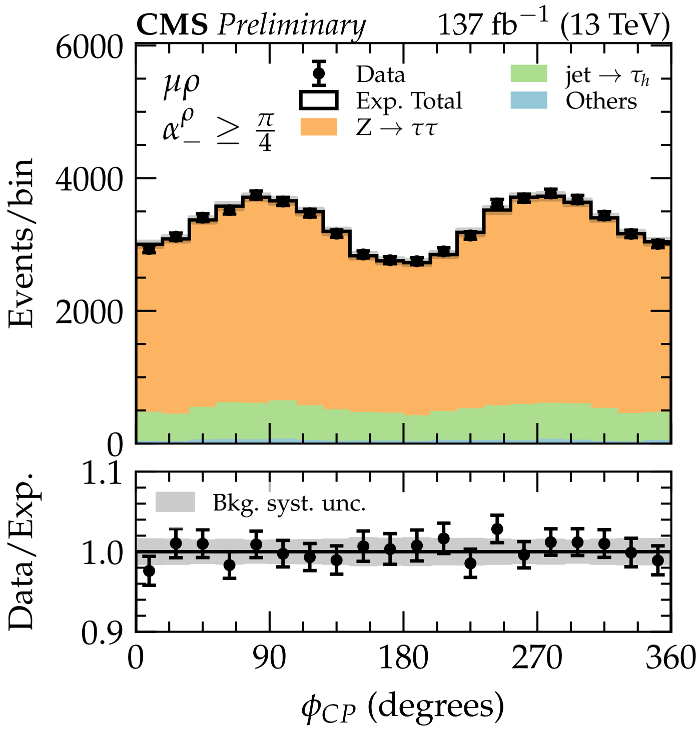


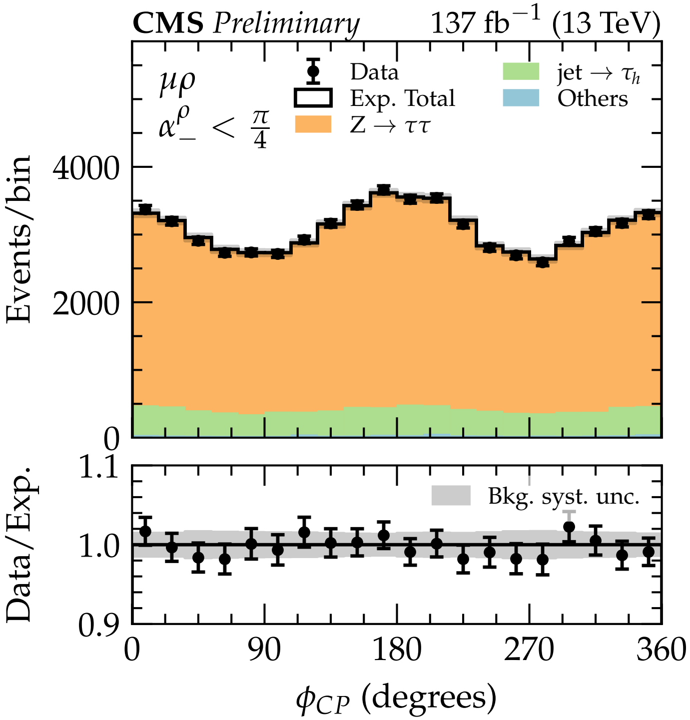
Mu+rho channel, Drell-Yan evts, phi_CP spectrum

Decompose using variable alpha

Great handle for data-mc validation!

$$\cos\alpha_- = \left| \frac{\hat{\mathbf{e}}_z \times \hat{\mathbf{p}}_{L-}}{|\hat{\mathbf{e}}_z \times \hat{\mathbf{p}}_{L-}|} \cdot \frac{\hat{\mathbf{n}}_- \times \hat{\mathbf{p}}_{L-}}{|\hat{\mathbf{n}}_- \times \hat{\mathbf{p}}_{L-}|} \right|$$





More supplementary results at: http://cms-results.web.cern.ch/cms-results/public-results/preliminary-results/HIG-20-006/index.html#AddFig

We construct a 4-component vector in the laboratory frame as $\lambda^{\pm} = (0, \mathbf{j}^{\pm})$. These four vectors $\lambda^{\pm}$ are boosted in the ZMF and denoted $\lambda^{*\pm}$. We also boost the respective charged pion four vectors to the ZMF, denoted $q^{*\pm}$. Then we take the transverse components of $\lambda^{*\pm}$ w.r.t. $q^{*\pm}$. We normalise the vectors to obtain unit vectors $\hat{\lambda}_{\perp}^{*+}$ and $\hat{\lambda}_{\perp}^{*-}$. From these vectors we reconstruct the angles ϕ^* and O^* as:

$$\begin{aligned} \phi^* &= \arccos(\hat{\lambda}_{\perp}^{*+} \cdot \hat{\lambda}_{\perp}^{*-}) \\ O^* &= \hat{q}^{*-} \cdot (\hat{\lambda}_{\perp}^{*+} \times \hat{\lambda}_{\perp}^{*-}), \end{aligned} \tag{4}$$

From ϕ^* and O^* we reconstruct ϕ_{CP} on a range $[0, 360^{\circ}]$ as:

$$\phi_{\mathrm{CP}} = \begin{cases} \phi^* & \text{if } O^* \geq 0 \\ 360^{\circ} - \phi^* & \text{if } O^* < 0 \end{cases} \tag{5}$$