IMPERIAL

CMS Compact Muon Solenoid

An extended Higgs sector search to explain the g-2 anomaly - the Type-X 2HDM

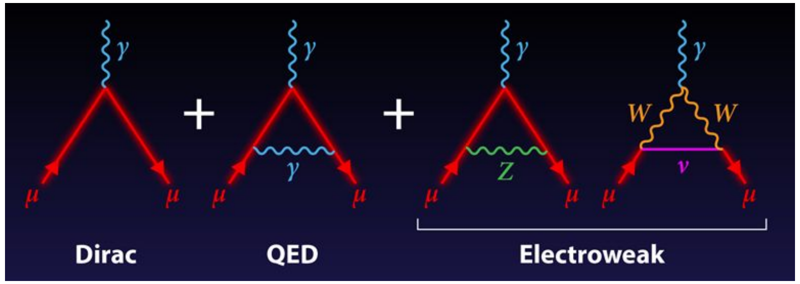
Daniel Winterbottom, David Colling, George Uttley, **Klitos Savva**
Imperial College London

IOP Institute of Physics

Motivation and Theory

IMPERIAL

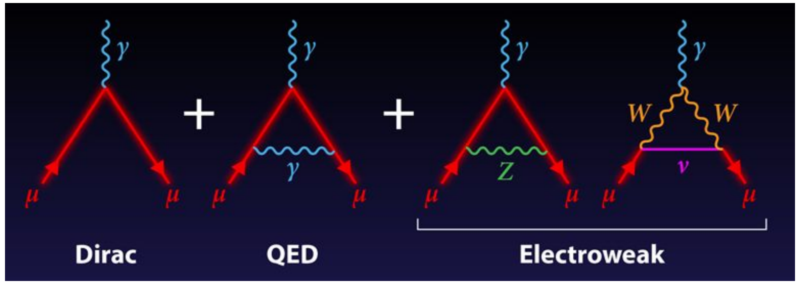
“While a comparison between the Fermilab result from Run-1/2/3 presented here, a_μ(FNAL), and the 2020 prediction yields a discrepancy of 5.0σ.” https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.131.161802



Motivation and Theory

IMPERIAL

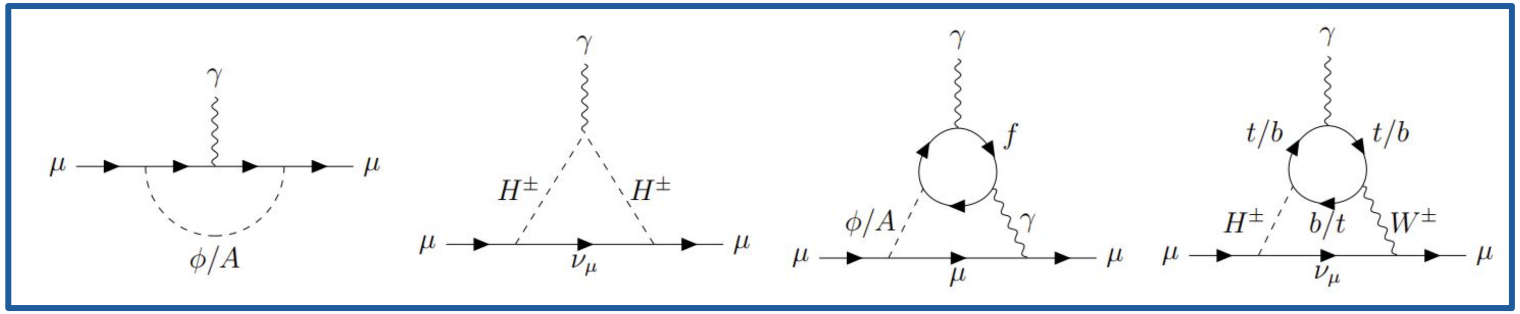
"While a comparison between the Fermilab result from Run-1/2/3 presented here, a_μ(FNAL), and the 2020 prediction yields a discrepancy of 5.0σ."   https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.131.161802



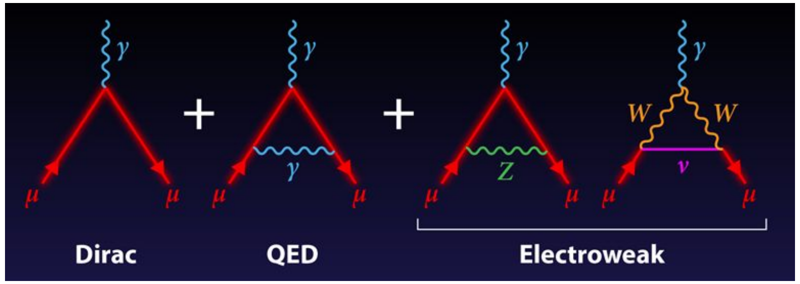
Two Higgs Doublet Models (2HDM)

- **Simplest** SM Higgs sector **extension**
- **Five** physical Higgs bosons:
  - **h** (Lighter CP-even boson)
  - **H** (Heavier CP-even boson)
  - **A** (CP-odd/Pseudoscalar boson)
  - **H±**

Any Higgs from a Two Higgs Doublet Model can contribute to this



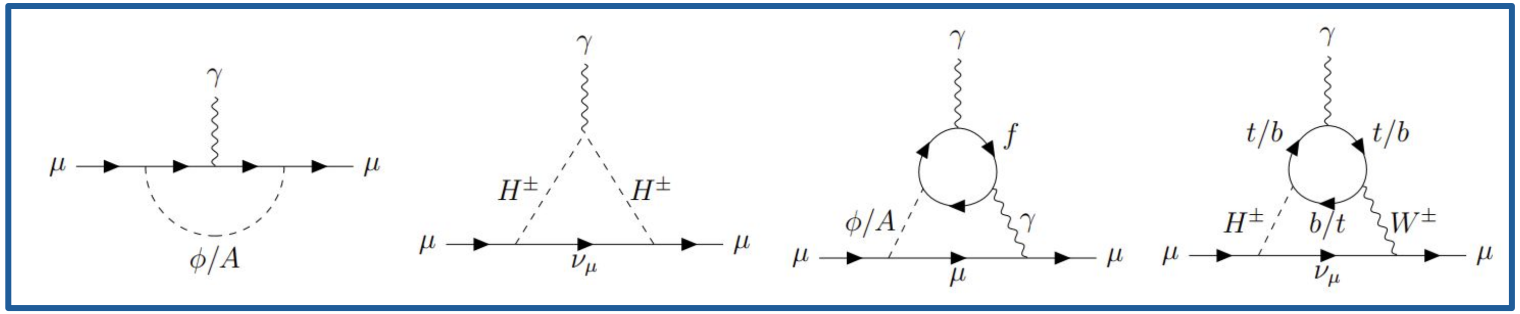
"While a comparison between the Fermilab result from Run-1/2/3 presented here, a_μ(FNAL), and the 2020 prediction yields a discrepancy of 5.0σ." https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.131.161802



Two Higgs Doublet Models (2HDM)

- **Simplest** SM Higgs sector **extension**
- **Five** physical Higgs bosons:
  - **h** (Lighter CP-even boson)
  - **H** (Heavier CP-even boson)
  - **A** (CP-odd/Pseudoscalar boson)
  - **H±**

Any Higgs from a Two Higgs Doublet Model can contribute to this



Normal Alignment Scenario → **H → ϕ**

$h_{obs} = h$

Inverted Alignment Scenario → **h → ϕ**

$h_{obs} = H$

What are the different types of 2HDM?

	Type I	Type II	Type X	Type Y
u	Φ_2	Φ_2	Φ_2	Φ_2
d	Φ_2	Φ_1	Φ_2	Φ_1
l	Φ_2	Φ_1	Φ_1	Φ_2

Two Higgs Doublet Models (2HDM)

IMPERIAL CMS

What are the different types of 2HDM?

	Type I	Type II	Type X	Type Y
u	Φ_2	Φ_2	Φ_2	Φ_2
d	Φ_2	Φ_1	Φ_2	Φ_1
l	Φ_2	Φ_1	Φ_1	Φ_2

- **Quark couplings** ➤ (1/tan β)^2
- Suppressed "standard" production modes
- High tan β (>10) ➤ **High branching fraction of additional Higgs bosons to taus**

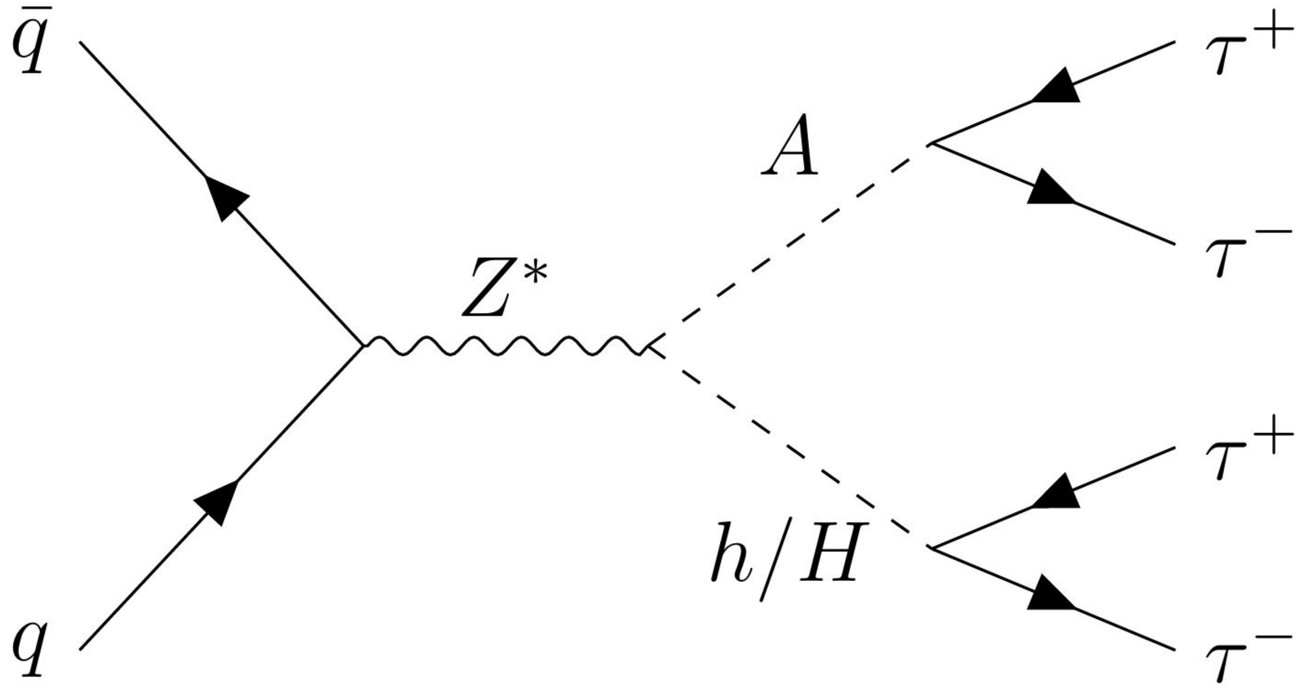
Why Type-X?

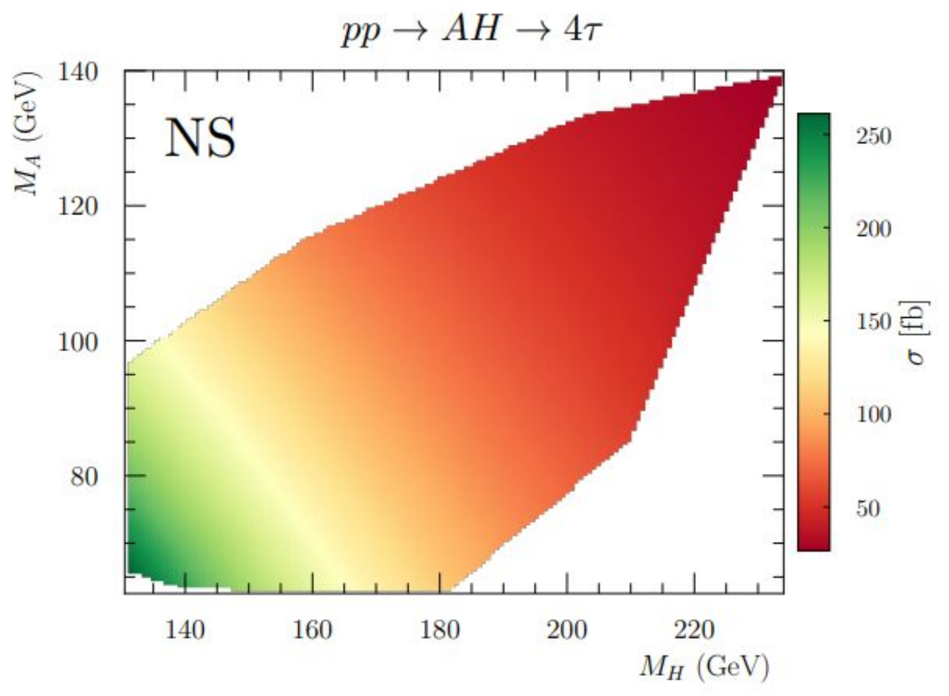
$$\tan\beta = \frac{\nu_2}{\nu_1}$$

	Type I	Type II	Type X	Type Y
g_h^u	c_α/s_β	c_α/s_β	c_α/s_β	c_α/s_β
g_h^d	c_α/s_β	$-s_\alpha/c_\beta$	c_α/s_β	$-s_\alpha/c_\beta$
g_h^l	c_α/s_β	$-s_\alpha/c_\beta$	$-s_\alpha/c_\beta$	c_α/s_β
g_H^u	s_α/s_β	s_α/s_β	s_α/s_β	s_α/s_β
g_H^d	s_α/s_β	c_α/c_β	s_α/s_β	c_α/c_β
g_H^l	s_α/s_β	c_α/c_β	c_α/c_β	s_α/s_β
g_A^u	$1/t_\beta$	$1/t_\beta$	$1/t_\beta$	$1/t_\beta$
g_A^d	$1/t_\beta$	t_β	$-1/t_\beta$	t_β
g_A^l	$1/t_\beta$	t_β	t_β	$-1/t_\beta$

Pinpointing a Process of Interest

IMPERIAL





Scenario	$\tan\beta$	m_A (GeV)	m_ϕ (GeV)	$m_{H^\pm}$ (GeV)
Normal	≥ 90	[62.5,145]	[130,245]	[95,285]
Inverted	≥ 120	[70,105]	[100,120]	[95,185]

https://journals.aps.org/prd/abstract/10.1103/PhysRevD.104.095008
https://link.springer.com/article/10.1007/JHEP11(2015)099
https://journals.aps.org/prd/pdf/10.1103/PhysRevD.104.053008

IMPERIAL

Preselection

2

3

4

IMPERIAL

Preselection

Reconstruction of ττττ (4τ) final state

States with highest BF considered

2

3

4

Preselection

Background Modelling

3

4

Background Modelling

Backgrounds modelled from Simulation (MC)

- Events with 0 jets misidentified as hadronic taus
- Mostly from all objects being reconstructed correctly from $ZZ \rightarrow 4L$
- A very small fraction from leptons misidentified as hadronic taus

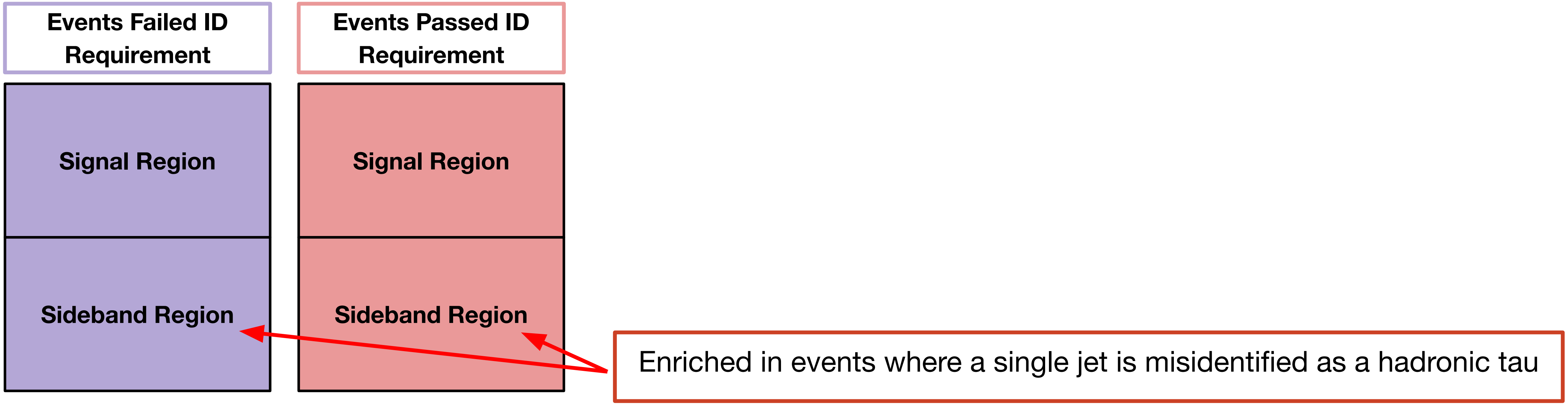
Backgrounds modelled with Data-driven methods

- Events where 1 or more jets are misidentified as a hadronic taus
- No reliable QCD Simulation estimate for estimating misidentified hadronic taus
- Backgrounds modelled with a ML Fake Factor Method

Fake Factor Method

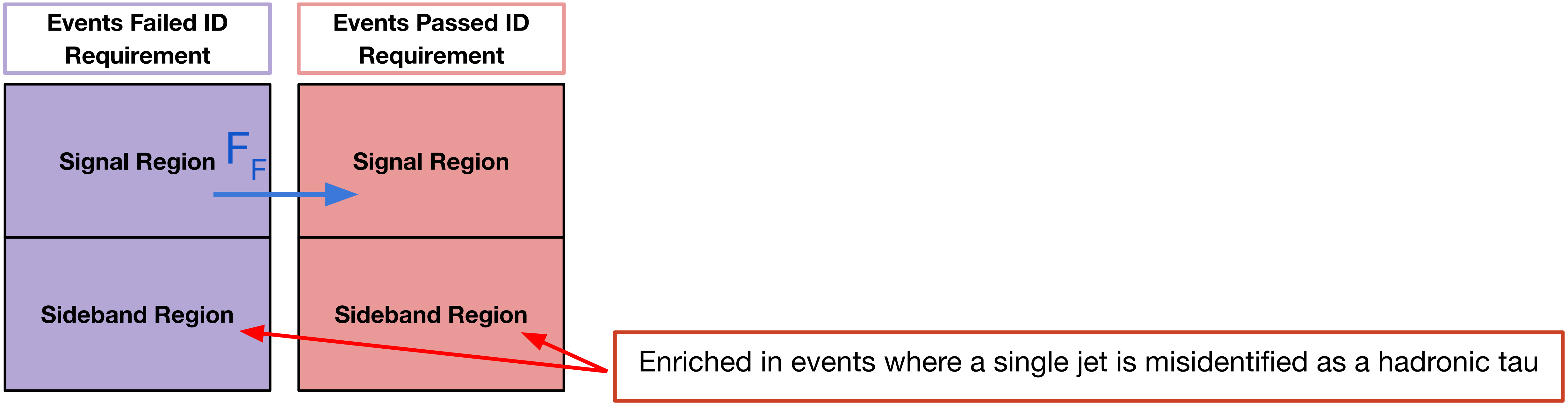
Modelling the background for a single jet misidentified as hadronic tau

$$F_F^i = \frac{N_{Pass}}{N_{Fail}}$$



Modelling the background for a single jet misidentified as hadronic tau

$$F_F^i = \frac{N_{Pass}}{N_{Fail}}$$

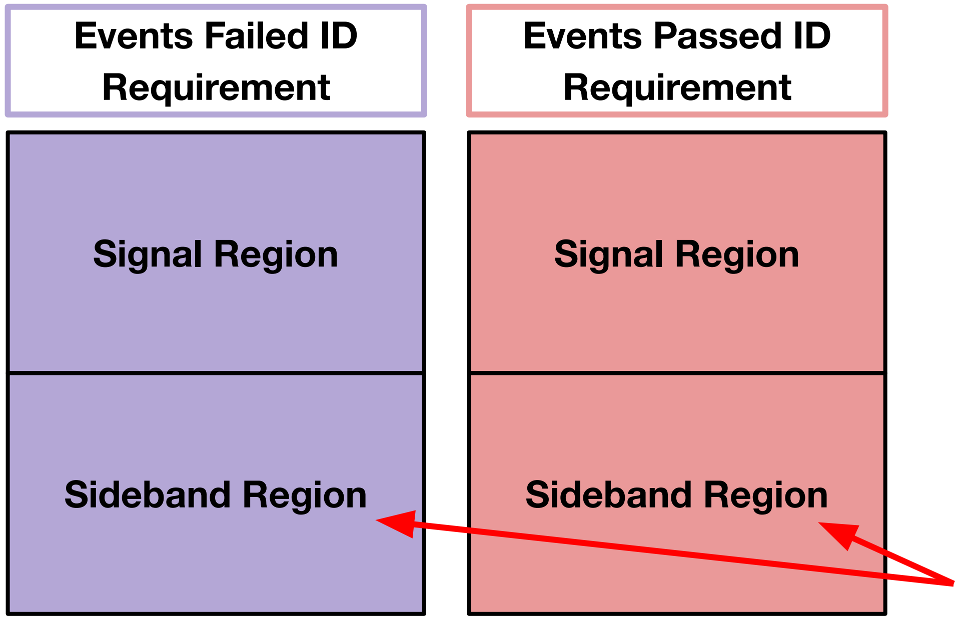


Fake Factor Method

IMPERIAL

Modelling the background for a single jet misidentified as hadronic tau

$$F_F^i = \frac{N_{Pass}}{N_{Fail}}$$



Enriched in events where a singlet jet is misidentified as a hadronic tau

What are the problems with this method?

- Needs to be done for any variables that parametrise the fake factor
  - In practise, this can only be done for only a few variables (Multi-dimensional reweighting with histograms is very limited)
- Variable choice is complex
  - Reweighting one variable often brings disagreement in others (Need corrections)
- Fake factor dependence on sideband region ➤ Method needs to be reproduced in additional sideband regions (ABCD Method)
- Have to fit each hadronic tau candidate separately

ML Fake Factor Method - BDT Reweighter

IMPERIAL

How can we improve the Fake Factor Method?

Reweighting distributions with a general-purpose ML techniques

- Train a binary discriminator
- Estimate ratio using classification probabilities

$$F_F^i = \frac{N_{Pass}}{N_{Fail}}$$

Multidimensional

What happens if the ratio is high or low?

High/Low ratio Regions

Important for reweighting **BUT** Easy to classify

How can we improve the Fake Factor Method?

BDT Reweighter

- Space of variables is split into a few large regions using decision trees (DTs)
  - DTs split the space of variables into regions (leafs) by checking simple conditions
- To best choose regions that need reweighting the algorithm looks to greedily optimise the symmetrized χ^2

$$F_F^i = \frac{N_{Pass}}{N_{Fail}}$$

$$\chi^2 = \sum_{\text{leaf}} \frac{(w_{\text{leaf, 1}} - w_{\text{leaf, 2}})^2}{(w_{\text{leaf, 1}} + w_{\text{leaf, 2}})^2}$$

$w_{leaf,fail} >> w_{leaf,pass}$

Optimal for Reweighting

Solves a lot of the FF problems

Preselection

Background Modelling

Distributions

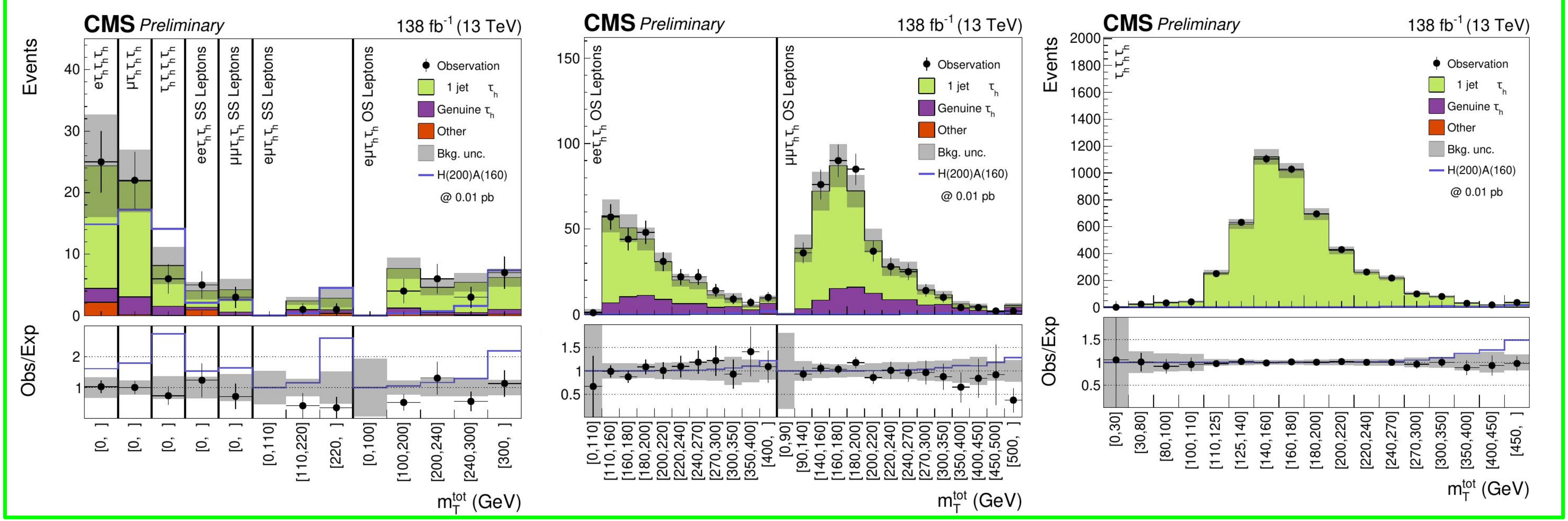
4

Analysis Distributions

IMPERIAL

- Showing every analysis bin and full uncertainty band after a background-only fit to data
- No signal is present and a good agreement between background and data is observed

$$m_T^{\text{tot}} = \sqrt{\sum_{i=1}^{N_\tau} m_T(\vec{p}_T^{\,\tau_i}, \vec{p}_T^{\,\text{miss}})^2 + \sum_{i,j=1;i\neq j}^{N_\tau} m_T(\vec{p}_T^{\,\tau_i}, \vec{p}_T^{\,\tau_j})^2}$$



Preselection

Background Modelling

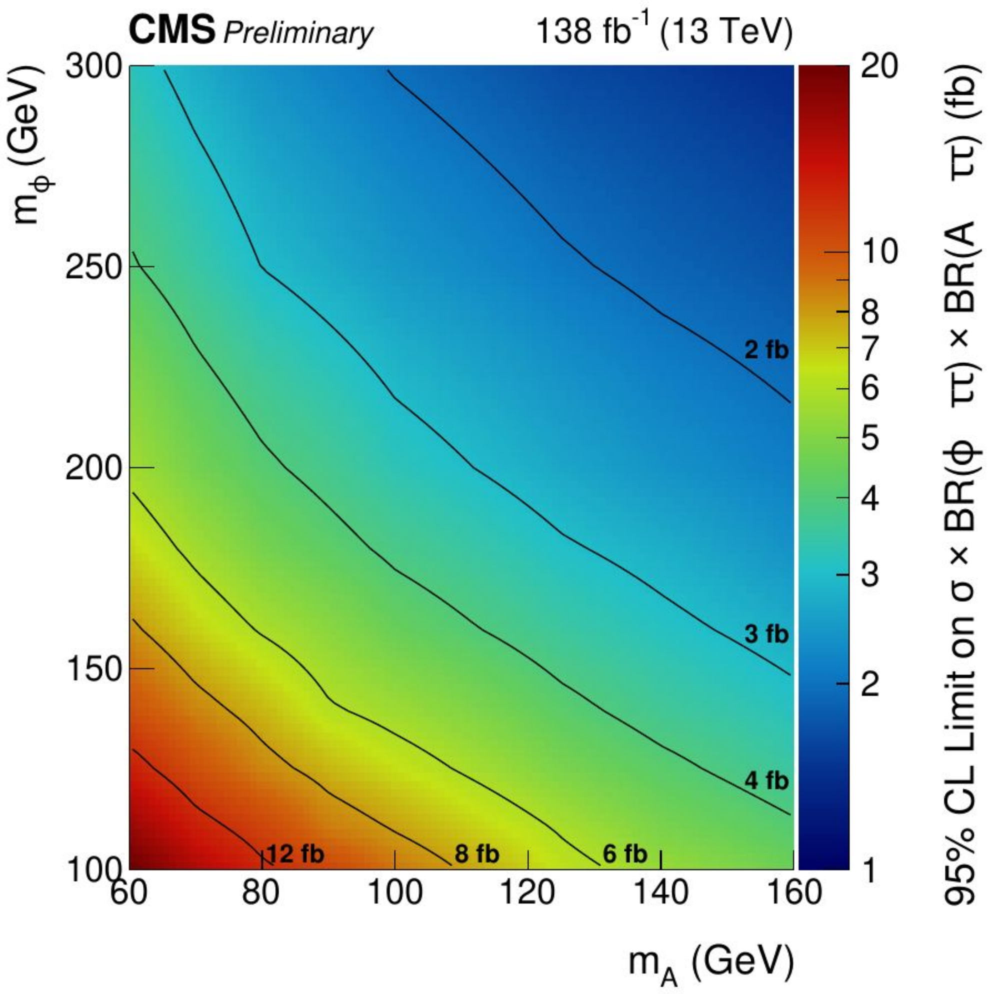
Distributions

Results

IMPERIAL

Model-Independent

- Set 95% CL limits on the cross-section multiplied by branching fractions of the $Z^* \rightarrow \phi A \rightarrow 4\tau$ process
- Predicted cross-sections ~ 10x higher than the limits at each mass point



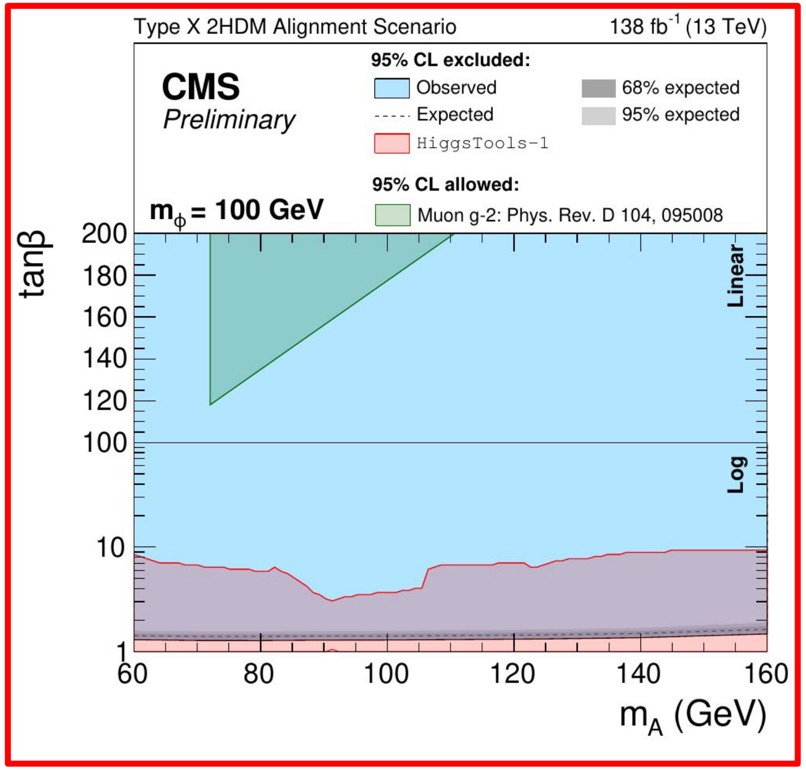
Results

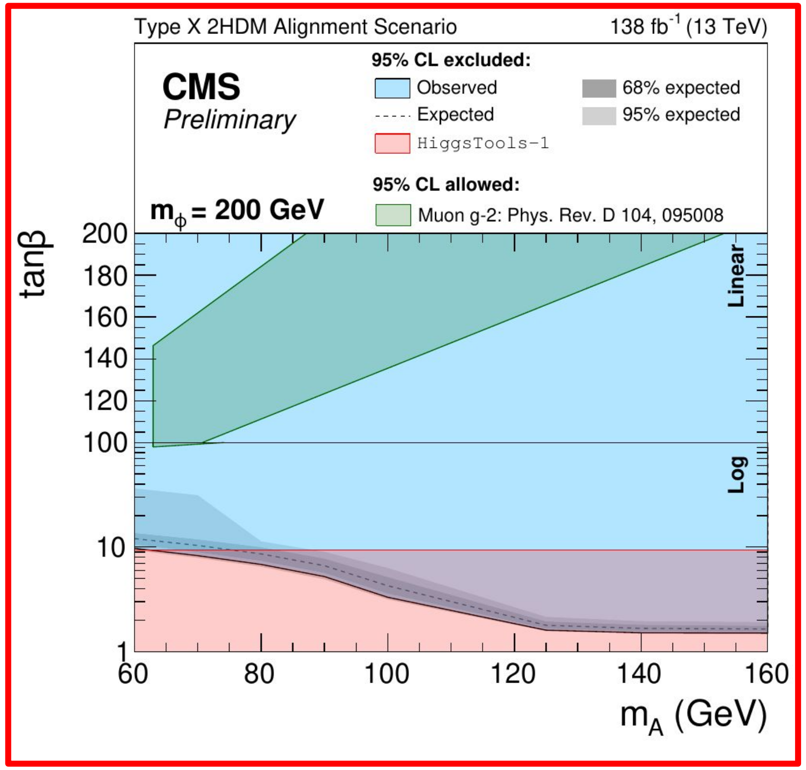
IMPERIAL

Model-Dependent

- Set 95% CL limits on the type-X 2HDM alignment scenario for different m_ϕ scenarios
- Type-X 2HDM excluded as a solution to the g-2 anomaly from the phenomenology papers

https://journals.aps.org/prd/abstract/10.1103/PhysRevD.104.095008
https://link.springer.com/article/10.1007/JHEP11(2015)099
https://journals.aps.org/prd/pdf/10.1103/PhysRevD.104.053008





Conclusion

IMPERIAL

- New analysis searching for $Z^* \rightarrow \phi A \rightarrow 4\tau$ presented
    - Motivated by the muon g-2 anomaly
- Several final states reconstructed to capture enough of the signal
- An improved ML Fake Factor method is used to model the background from jets misidentified as hadronic taus
- Analysis excludes the Type-X 2HDM as a solution to the g-2 anomaly while also excluding the whole model in the mass ranges scanned

IMPERIAL

Thank you !

IMPERIAL

CMS

Compact Muon Solenoid

Backup

IOP Institute of Physics

Signal Modelling

IMPERIAL

- Physical Basis of parameters:

  m_h , m_H , m_A , $m_{H\pm}$, tan β , cos(β-α), m_{12}^2, λ_6 , λ_7

  Removed by a soft Z_2 symmetry

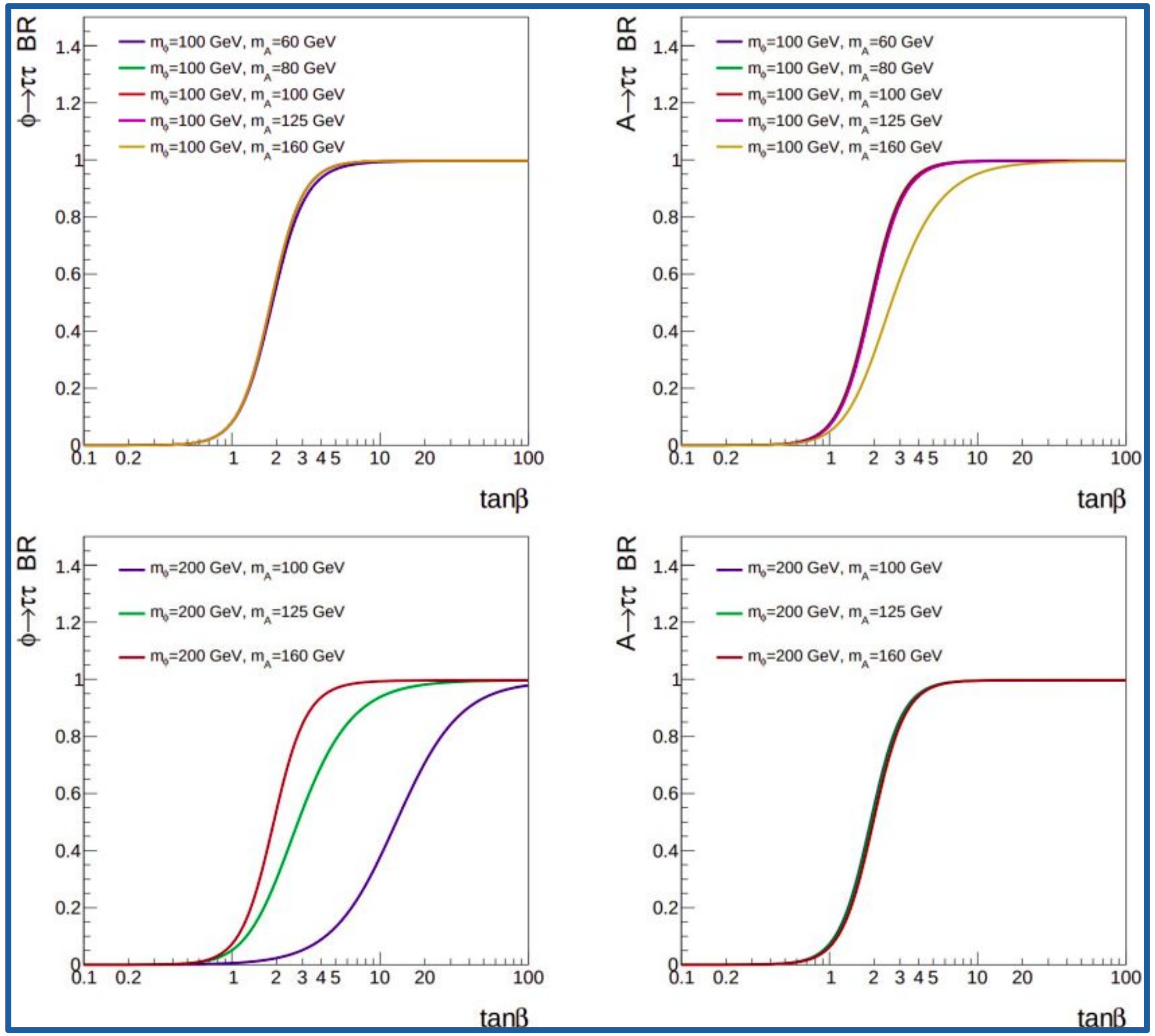
  Sensitive in this analysis

- The production is modelled in **MADGRAPH5 aMC@NLO** and the decays in PYTHIA
  - Production independent of tan β
- The cross sections are determined also from **MADGRAPH5 aMC@NLO**
- 2D Grid of masses for A and ϕ
  - m_A : 60, 70, 80, 90, 100, 125, 140, 160
  - m_ϕ : 100, 110, 125, 140, 160, 180, 200, 250, 300
- If m_ϕ > 125 GeV, use normal alignment scenario m_h = 125 GeV
- If m_ϕ < 125 GeV, use inverted alignment scenario, m_H = 125 GeV
- Remaining parameters are negligible

Additional Higgs BRs

IMPERIAL

- ❖ Higgs BRs calculated with 2HDECAY
  - https://arxiv.org/abs/1810.00768
- ❖ BFs for the Type-X 2HDM in the alignment limit shown
- ❖ When H and A become too separated the BF has to compete with H → ZA
- ❖ BR drops sharply between 1 and 2 in tan β



Simulation Corrections & Uncertainty Model

IMPERIAL

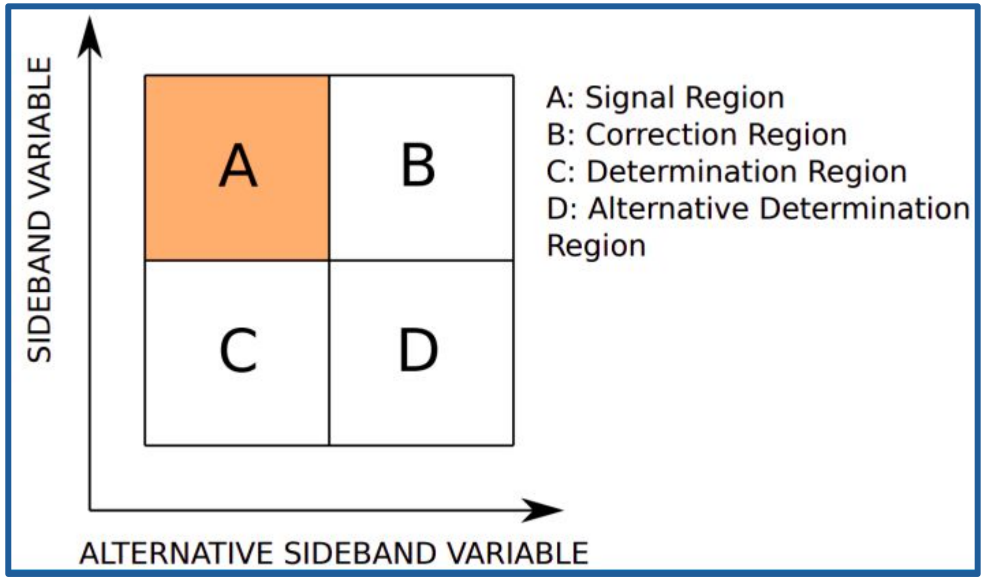
- Corrections applied for:
  - Pileup
  - B tagging
  - Electron and muon tracking/reconstruction
  - Electron, muon, hadronic tau IDs
  - Electron and muon isolation
  - Single electron, single muon and double-tau triggers
- General MC Shape Uncertainties:
  - Lepton ID efficiencies
  - Lepton trigger efficiencies
  - Electron, Tau energy scales
  - Jet energy scale, resolution
  - MET unclustered uncertainty
- Specific MC Shape Uncertainties:
  - $ZZ \rightarrow 4L$ k factor uncertainty
  - Signal theory uncertainties
    - QCD scale ~2%
    - PDF variations ~6%
    - α_s 1%
- Simulation Normalisation Uncertainties:
  - Luminosity
  - B tagging efficiency / Misidentification rate
  - Prefiring

Fitting Regions and Variables

IMPERIAL

Fitting Region

- In region C and D, the nominal total charge selection is inverted
- B and D have at least one of the other τ_h candidates failing the tau ID
- All years and τ_h candidates in the event fitted together. Each channel fit separately



Variables

- The HPS decay mode of the hadronic tau candidate
- p_T of the hadronic tau candidate
- The ratio of the p_T of the matched jet to the p_T of the hadronic tau candidate
- η of the hadronic tau candidate
- Where the candidate passes a double tau trigger leg
- The charge of the hadronic tau candidate
- The total charge of the combined objects
- The DeepTauVsJetsID of the other hadronic tau candidates
  - Sorted by p_T
- Year of data taking
- p_T rank of the hadronic tau in the event

Applying Fake Factors

IMPERIAL

- We now have fake factors for all hadronic taus in the event
- If we just use the fake factor on the leading hadronic tau applied to events where it fails the tau ID, we miss events where the leading hadronic tau is genuine and another is a fake

Example: τ_h τ_h channels

- By applying fake factors to a specific hadronic tau that fails the ID, not all of the possible combinations of jets misidentified as hadronic taus are covered
- Need to subtract off the "double" region to not over count events where there are two jets misidentified as hadronic taus

Region	$\tau_h^1(\tau)\tau_h^2(j)$	$\tau_h^1(j)\tau_h^2(\tau)$	$\tau_h^1(j)\tau_h^2(j)$
R_1 (from τ_h^1)	0	1	1
R_2 (from τ_h^2)	1	0	1
R_{12} (from both)	0	0	1
$R_1 + R_2 - R_{12}$	1	1	1

Formulas

$\mu\mu\tau_h\tau_h$, $ee\tau_h\tau_h$ and $e\mu\tau_h\tau_h$

$$R_1 + R_2 - R_{12}$$

$\mu\tau_h\tau_h\tau_h$, $e\tau_h\tau_h\tau_h$ and $\tau_h\tau_h\tau_h$

$$R_1 + R_2 + R_3 - R_{12} - R_{13} - R_{23} + R_{234}$$

$\tau_h\tau_h\tau_h\tau_h$

$$R_{12} + R_{13} + R_{14} + R_{23} + R_{24} + R_{34}$$
$$- 2(R_{123} + R_{124} + R_{134} + R_{234}) + 3R_{1234}$$

Fake Factor Uncertainties

IMPERIAL CMS

Non Closure Uncertainty

Non-closure uncertainty accounts for any mismodelling caused by the BDT reweighting fit

- Histograms for all variables are drawn in the fitted pass and fail tau ID regions weighted by the derived fake factors (rebinned to minimise fluctuations)
- Histograms are calculated from the difference between two histograms in each variable
- For every event, the difference in each variable is found
- The largest fractional shift is taken as an uncertainty on that fake factor weight (symmetrised).

C to A & B to A Uncertainties

Covers any mismodelling due to the derivation of the fake factors being partly performed in a region with sideband or alternative sideband variables not being that of the signal region

- Find the largest shift in the algorithm by sampling different values in the sideband and alternative sideband variables in B and C
- Decorrelated in each combination of sideband variables

Subtraction Uncertainties

Covers any mismodelling due to the purification of the fitting regions

- Non-closure uncertainty on the method compared to the histogram subtraction in each variable
- Shifts from adjusting the MC “subtracted” yields up and down by 10%

Flavour Dependence - Study

IMPERIAL

- Jets misidentified as hadronic taus in this search come from 5 main processes:
  - TTBar
  - Drell-Yan
  - Diboson
  - W+Jets
  - QCD
- Out of these only TTBar significantly contributes heavy flavours of jets
- Fake Factors previously have been shown to be consistent between light flavour quark processes
- Questions:
  - How much does the heavy-to-light flavour jet ratio change between the determination and signal region?
  - How much does this change affect the closure of the fake factors?
- Performed a study on MC to answer these questions. Note we have no QCD MC, but the addition of QCD in this study would only make the fraction changes smaller, making the example less extreme

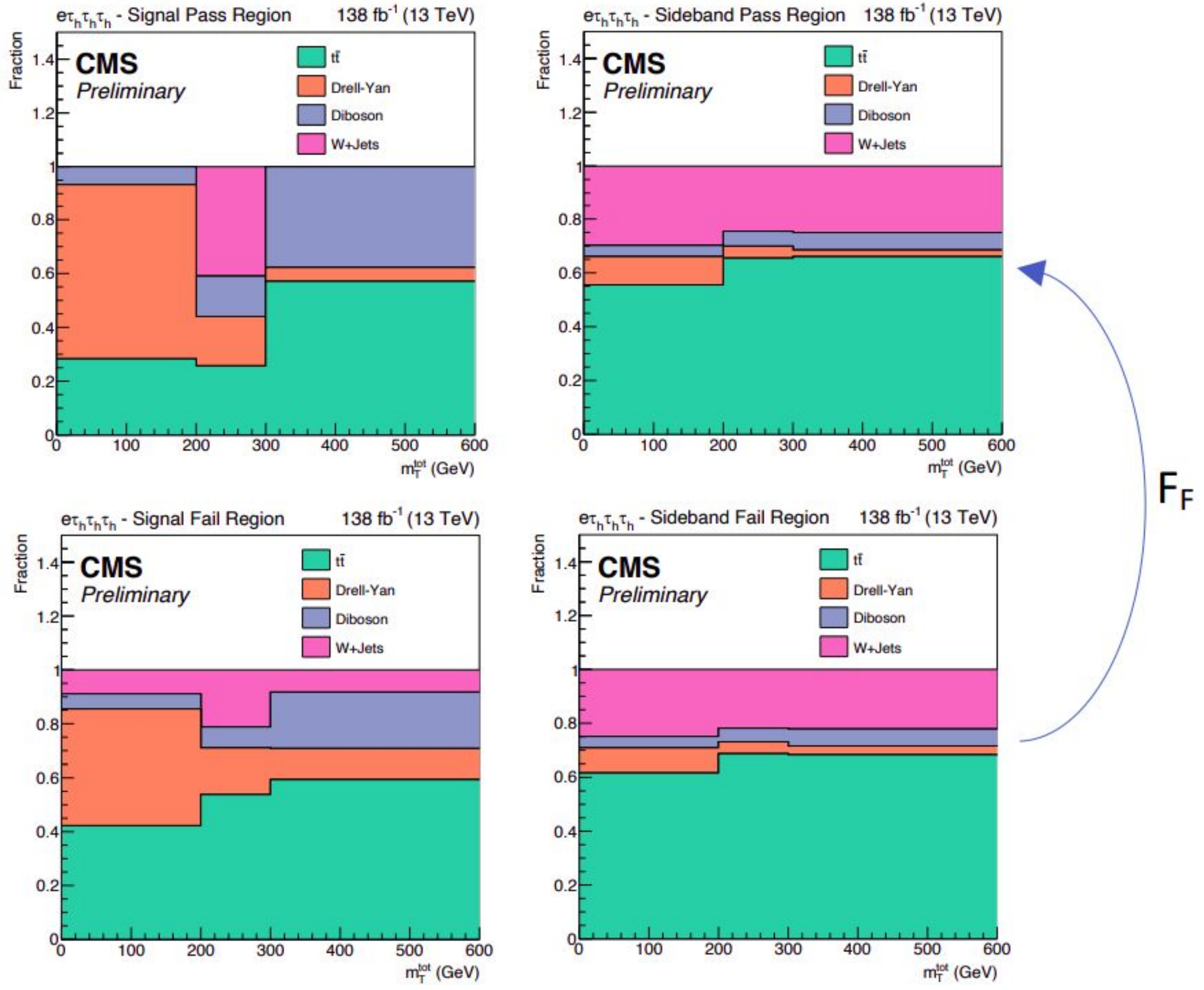
Fractions of jets misidentified as τ_h in $e\tau_h\tau_h\tau_h$

IMPERIAL

Signal | Sideband

Pass tau ID

Fail tau ID

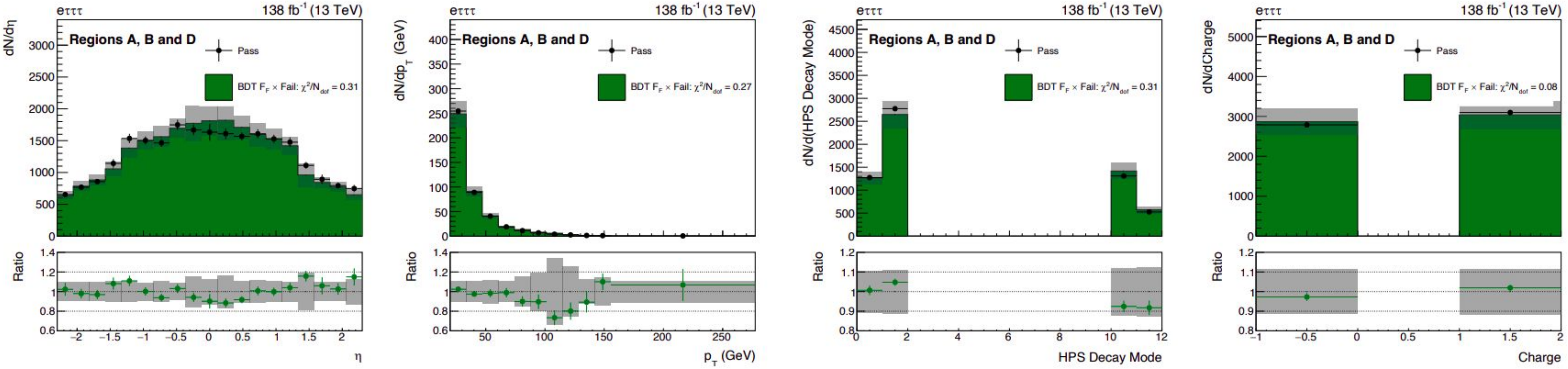


- This is the most extreme sample:
  - Go from approximately 60% TTBar in the fail and pass tau ID regions in the sideband
  - To about 40% and 25% in the fail and pass tau ID regions respectively in the signal region
- Use this as our studied example
  - Derive F_F from this MC sideband region
  - Check closure applied to signal region

$e\tau_h\tau_h\tau_h$ closure in sideband

IMPERIAL

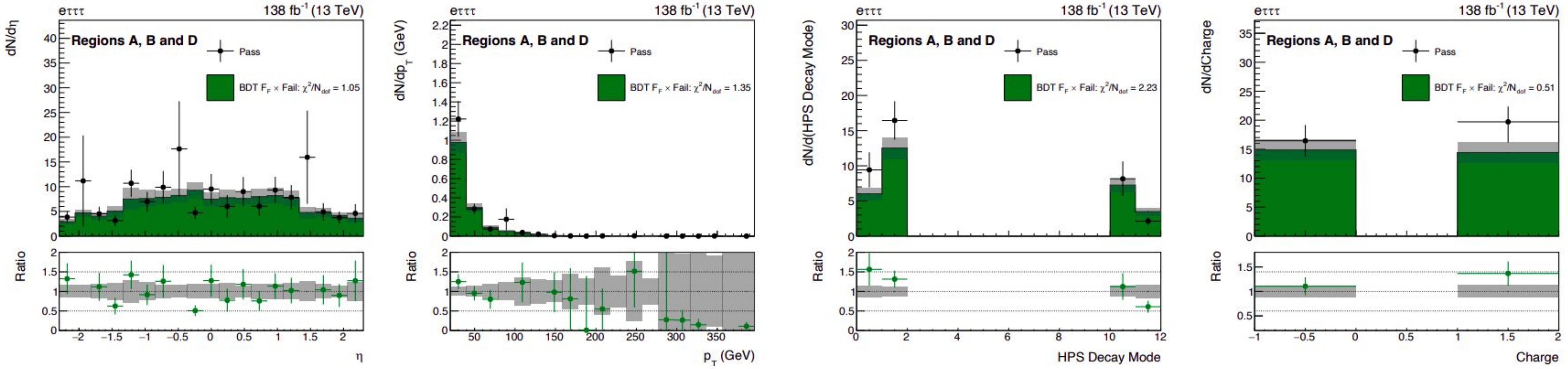
- Good closure



$e\tau_h\tau_h\tau_h$ closure in sideband

IMPERIAL

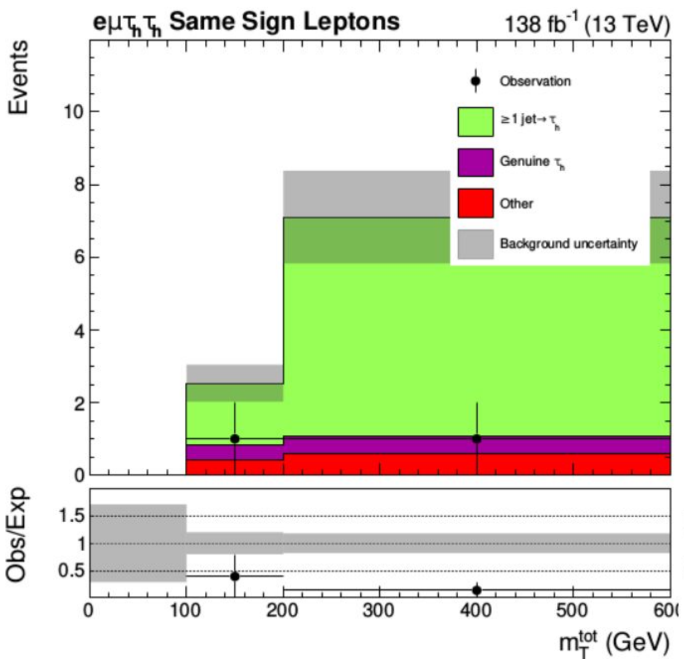
- Closure still reasonable in signal region with the different heavy jet flavour composition
- This is the extreme case, not including any QCD, if QCD present (which there is), this effect is even smaller
- No major modelling bias for jets misidentified as hadronic taus from the jet flavour composition in this analysis



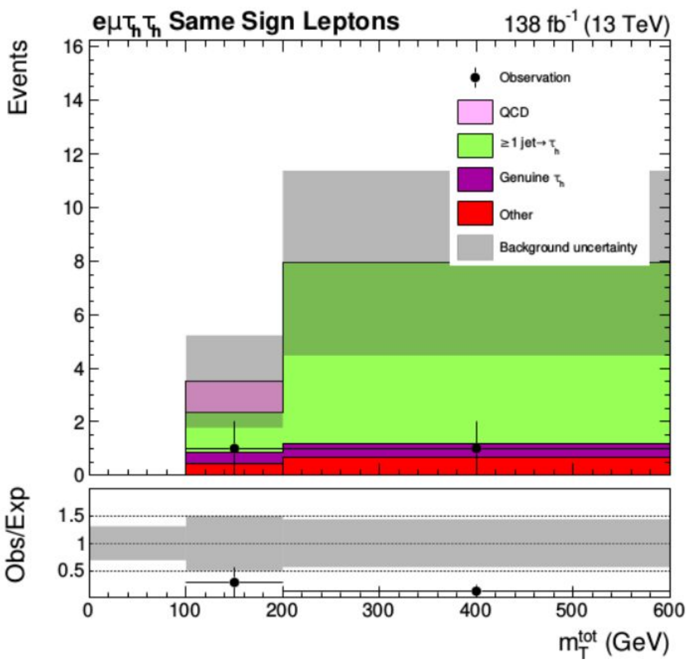
Check of $e\mu\tau_h\tau_h$ with MC

- Set up a second method for modelling jets misidentified as hadronic taus to check background predictions
- Use MC for events with 1 or more jets misidentified as hadronic taus for W, DY, TTbar, Diboson, Triboson, EWK-W, EWK-Z, single-top
- Use a basic ABCD method for QCD from qsum != 0 region, with transfer factor taken from the region where one or more object is anti-isolated
- Anti-isolated:
  - Leptons: Isolation > 0.15
  - Hadronic Taus: !Loose vsJets WP && vsJets score > 0.1
- Prefit with statistical uncertainties only shown
- Predictions are approximately equivalent

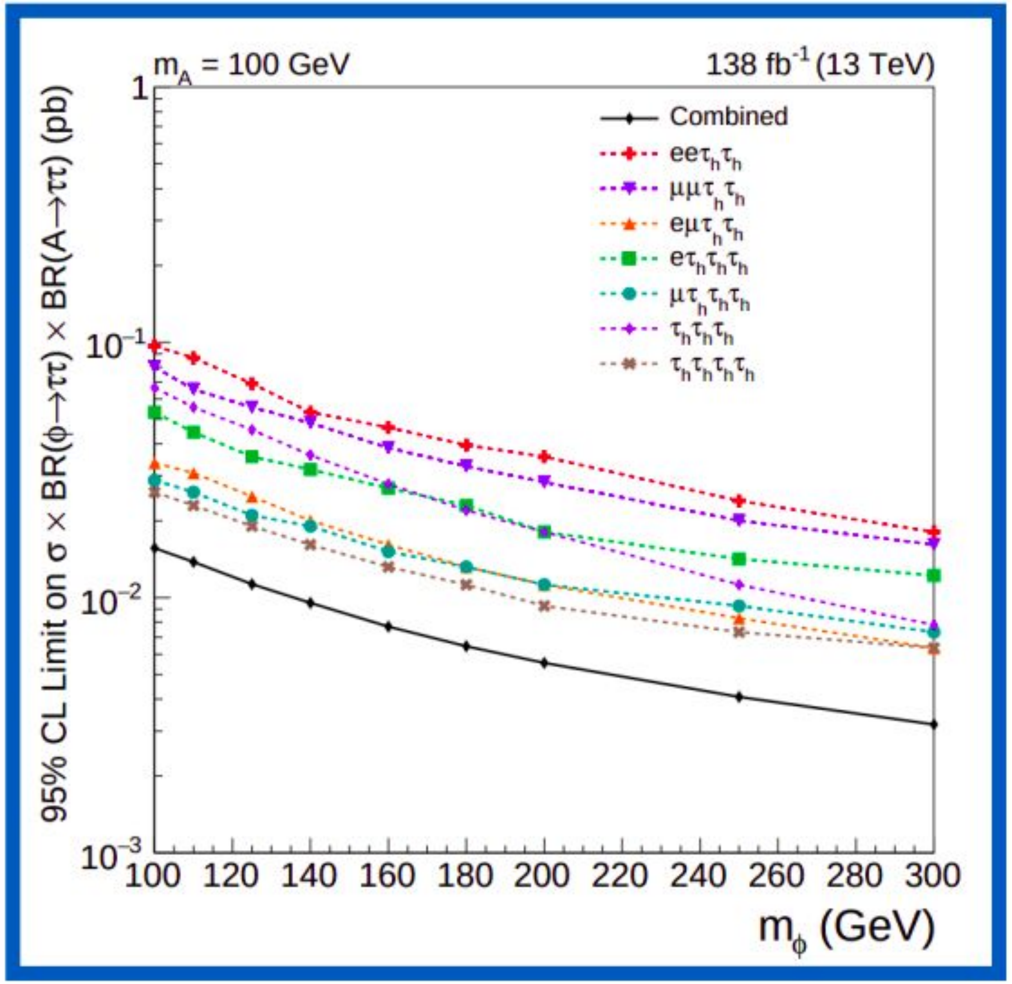
F_F Method

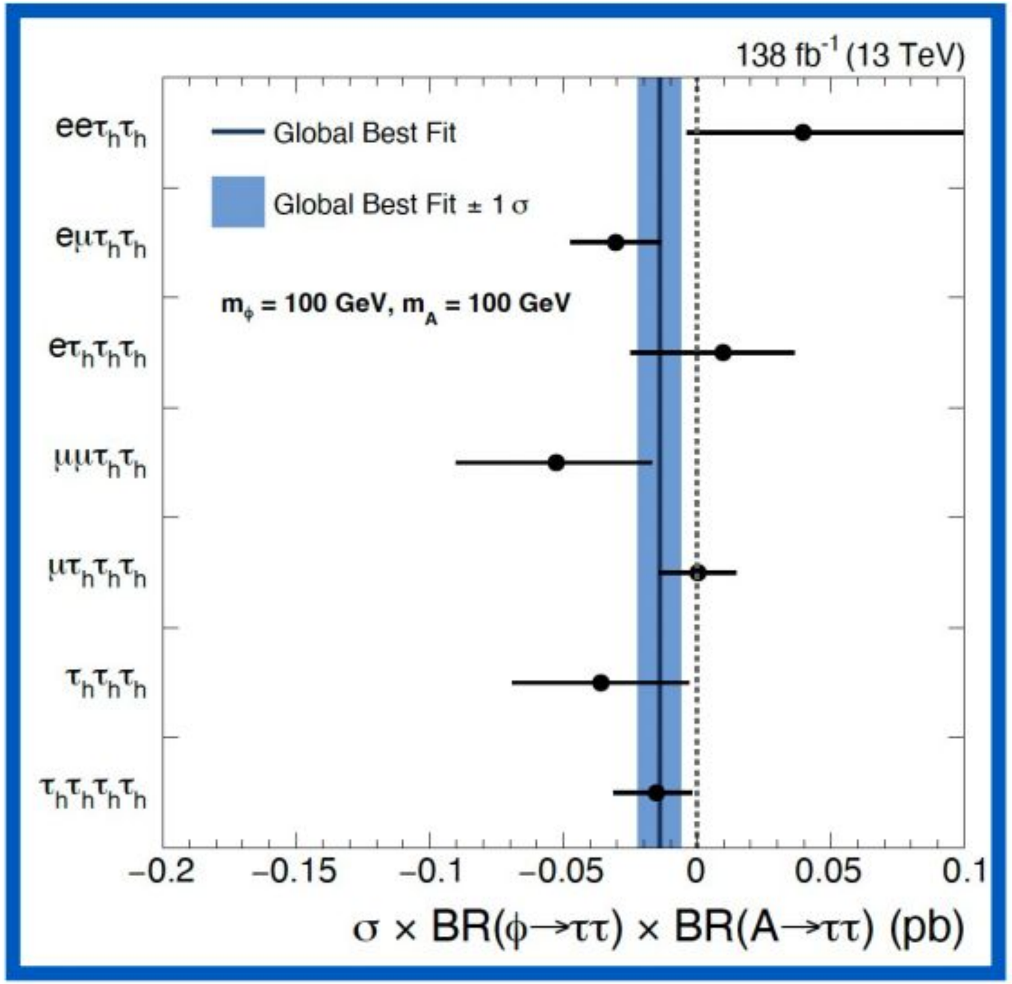


MC with QCD ABCD Method



Channel Sensitivity Analysis





Signal Widths

IMPERIAL

m_A=200 GeV, m_H=200 GeV

tan(β)	Type X Widths		Type II Widths	
	H	A	H	A
10	5.48E-03	4.27E-02	3.67E-01	3.73E-01
20	1.51E-01	1.50E-01	1.63E+00	1.66E+00
30	3.41E-01	3.83E-01	3.92E+00	3.73E+00
40	6.27E-01	6.82E-01	5.97E+00	6.06E+00
50	9.92E-01	1.07E+00	-8.73E+00	3.11E+00
60	1.36E+00	1.54E+00	-2.28E+02	-5.61E+01
70	1.97E+00	2.10E+00	-3.09E+03	-8.03E+02
80	2.59E+00	2.75E+00	-5.85E+04	-1.66E+04
90	3.31E+00	3.49E+00	-3.44E+06	-1.45E+06
100	4.15E+00	4.33E+00	-2.96E+11	-2.96E+11

- Table of widths for a signal hypothesis for Type-X and Type-II 2HDMs
- Type-II widths grow much quicker than Type-X, this is because the widths grow with the coupling enhancements
- In Type-II there are double tan β enhancements than in Type-X
- You can go higher in tan β for Type-X rather than Type-II