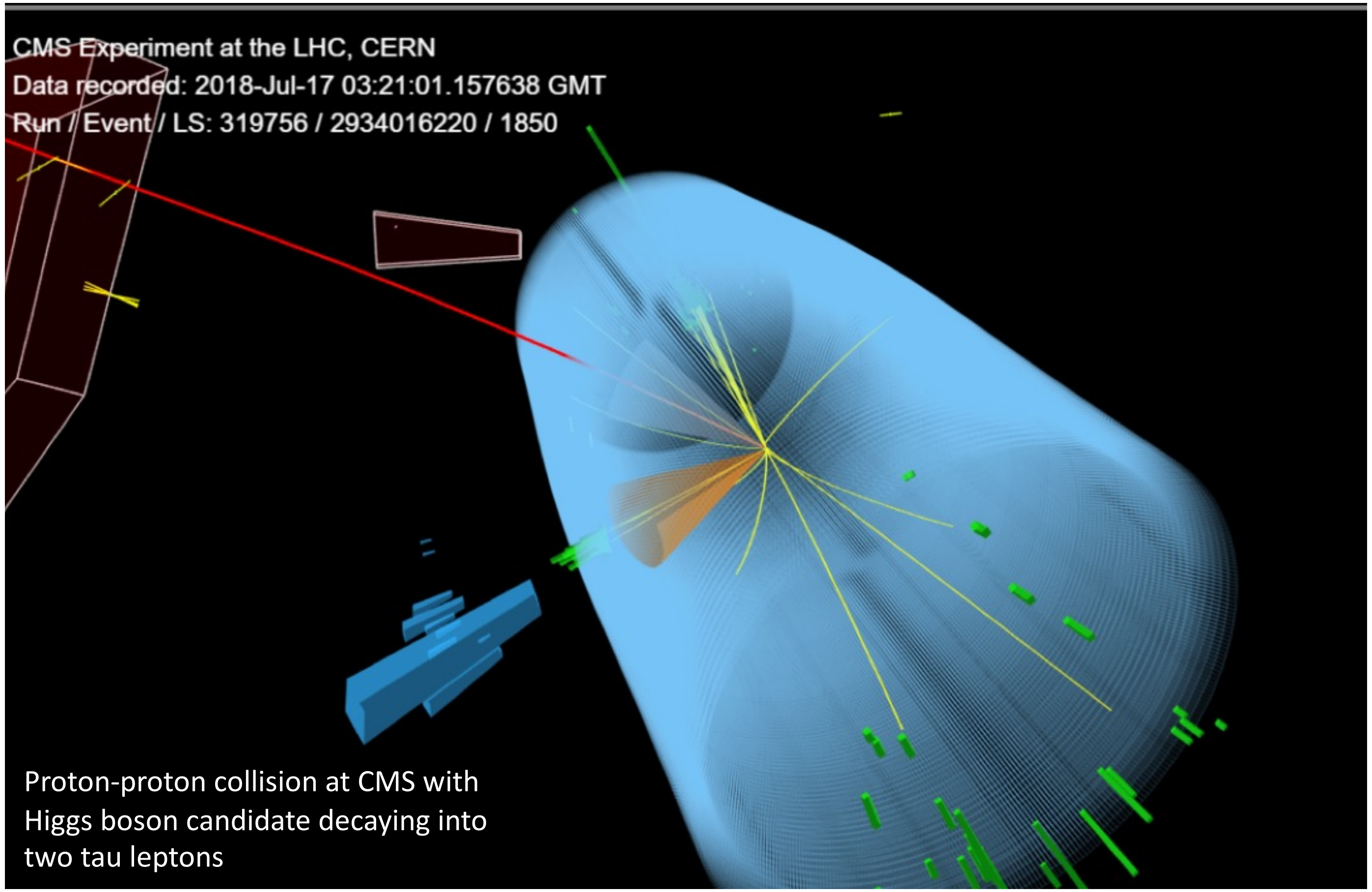
Université de Strasbourg

New method for calibrating tau leptons in the CMS experiment at the LHC

M2 PSA Internship defence

21/06/2022



Proton-proton collision at CMS with Higgs boson candidate decaying into two tau leptons

Océane PONCET

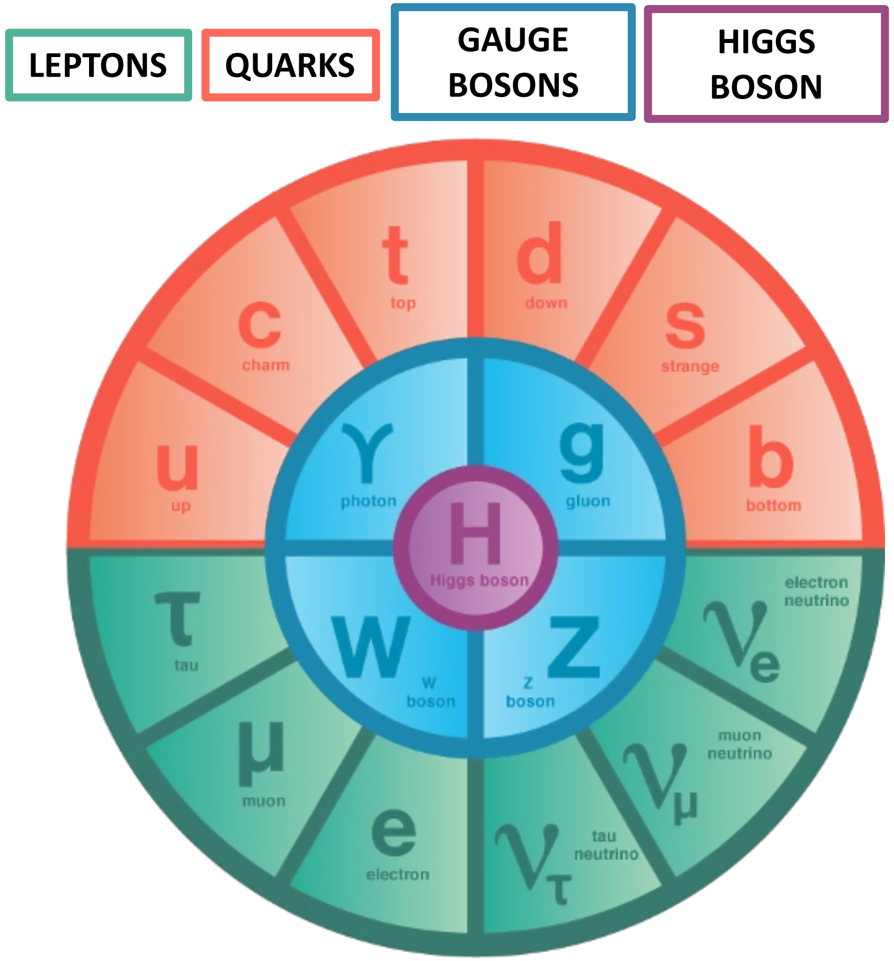
Directed by Saskia FALKE

iPHC Institut Pluridisciplinaire Hubert Curien Strasbourg

QMat Quantum Science and Nanomaterials

CMS Compact Muon Solenoid

τ leptons



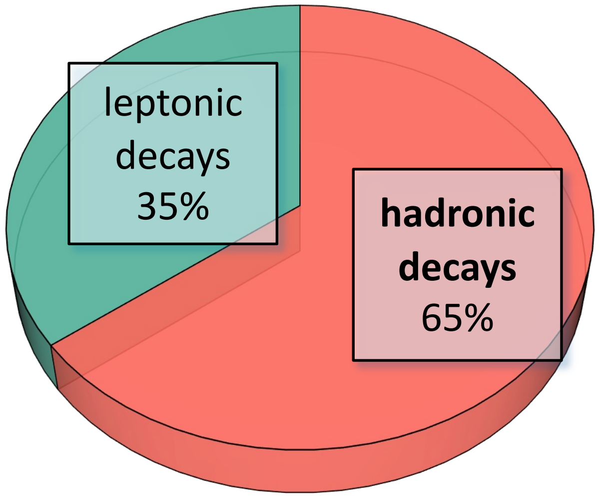
Standard Model (SM) of particle physics

Why τ leptons?

- $H \rightarrow \tau^+\tau^-$ allows to probe the CP nature in the interaction of the Higgs boson with fermions.
- Search of physics beyond the SM (BSM).

Challenge: τ leptons **decay** $\rightarrow$ need to be efficiently **reconstructed** from their decay products.

DECAY MODE OF τ LEPTONS :



Objective: develop a new method for calibrating hadronic tau leptons (τ_h) during their reconstruction.

Plan

I. Experimental setup

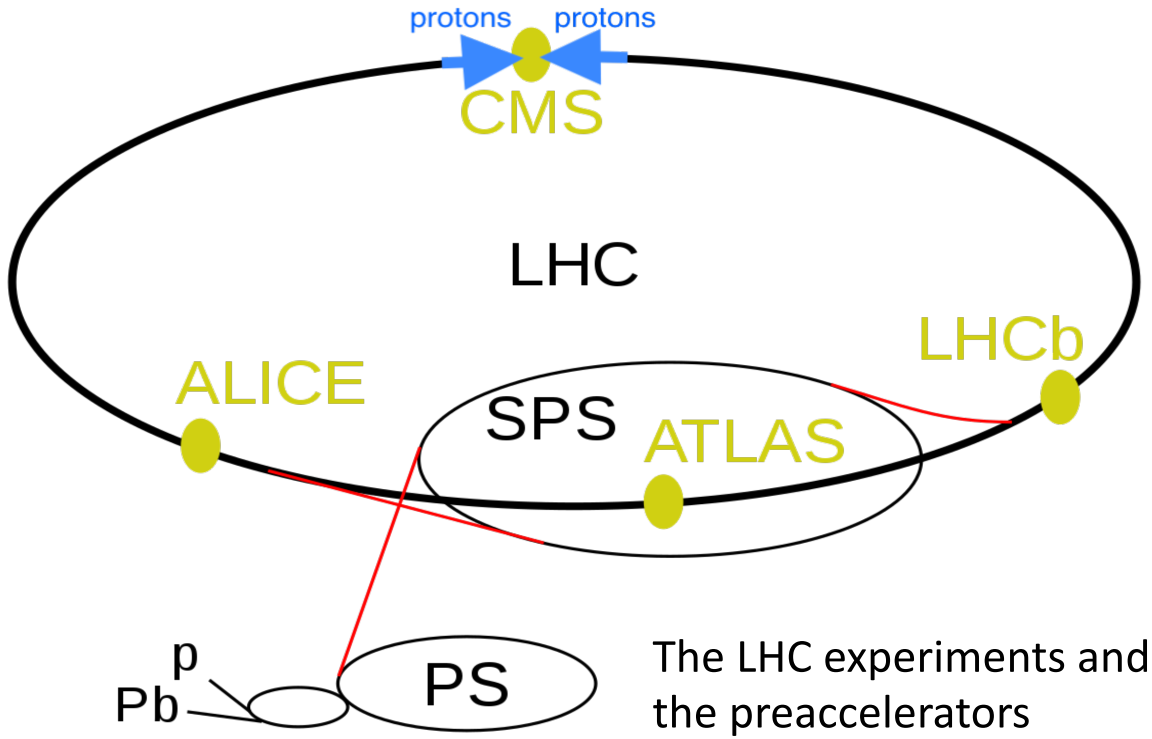
II. τ_h reconstruction and identification in CMS

III. Study of the tau energy scale fit

IV. The new method : simultaneous fit

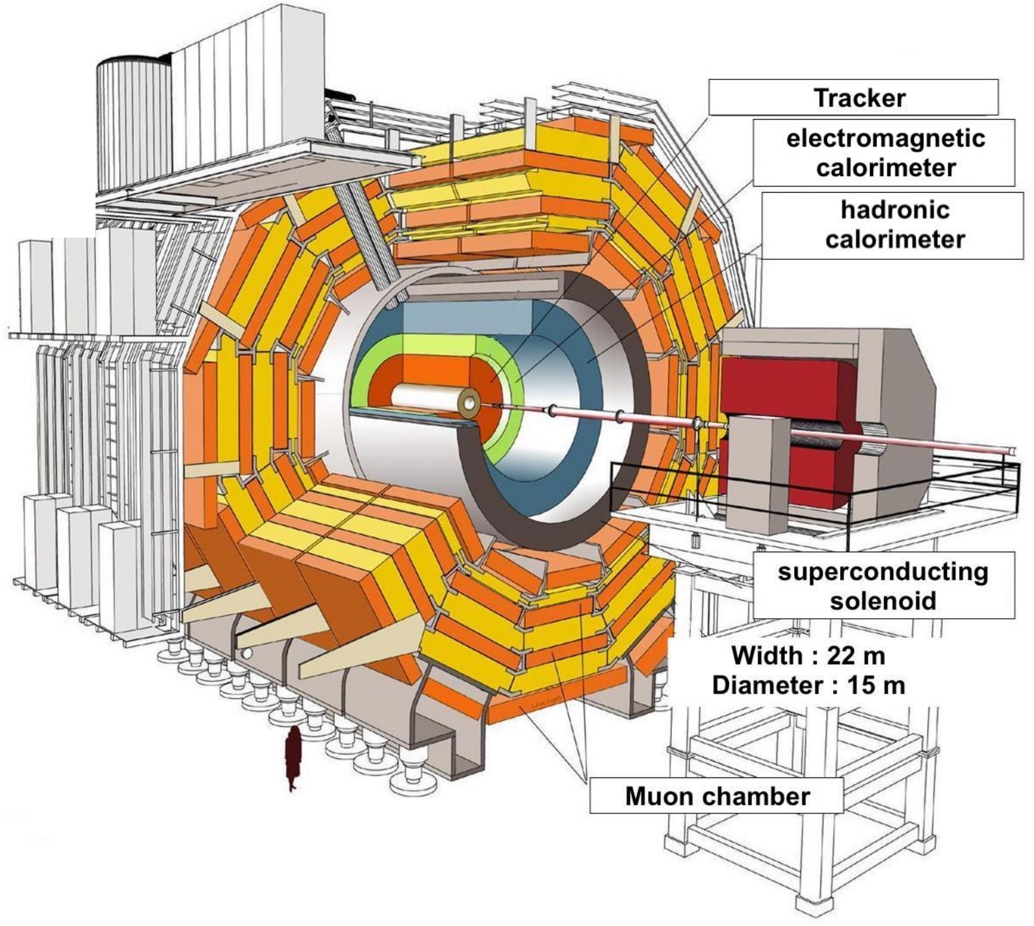
I. Experimental setup: CMS and LHC

LHC (Large Hadron Collider) particle accelerator on the French-Swiss border (CERN)



The LHC experiments and the preaccelerators

The particles are detected by the different layers of CMS (Compact Muon Solenoid).



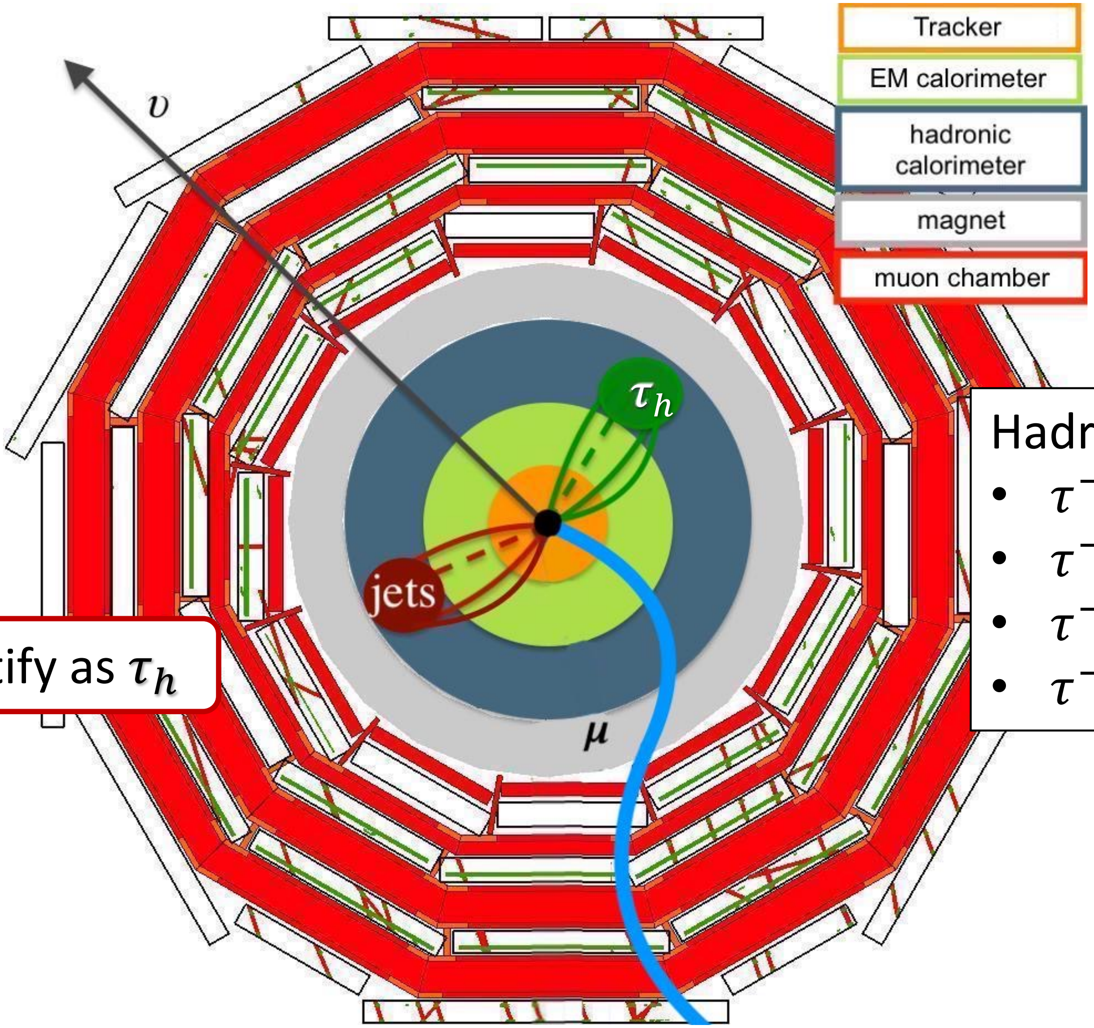
Particles are produced by proton-proton collisions at an energy of 13 TeV.

Schematic of the CMS detector

II.A. τ_h reconstruction and identification in CMS

Transverse view of CMS

Escape the detection (MET)



Misidentify as τ_h

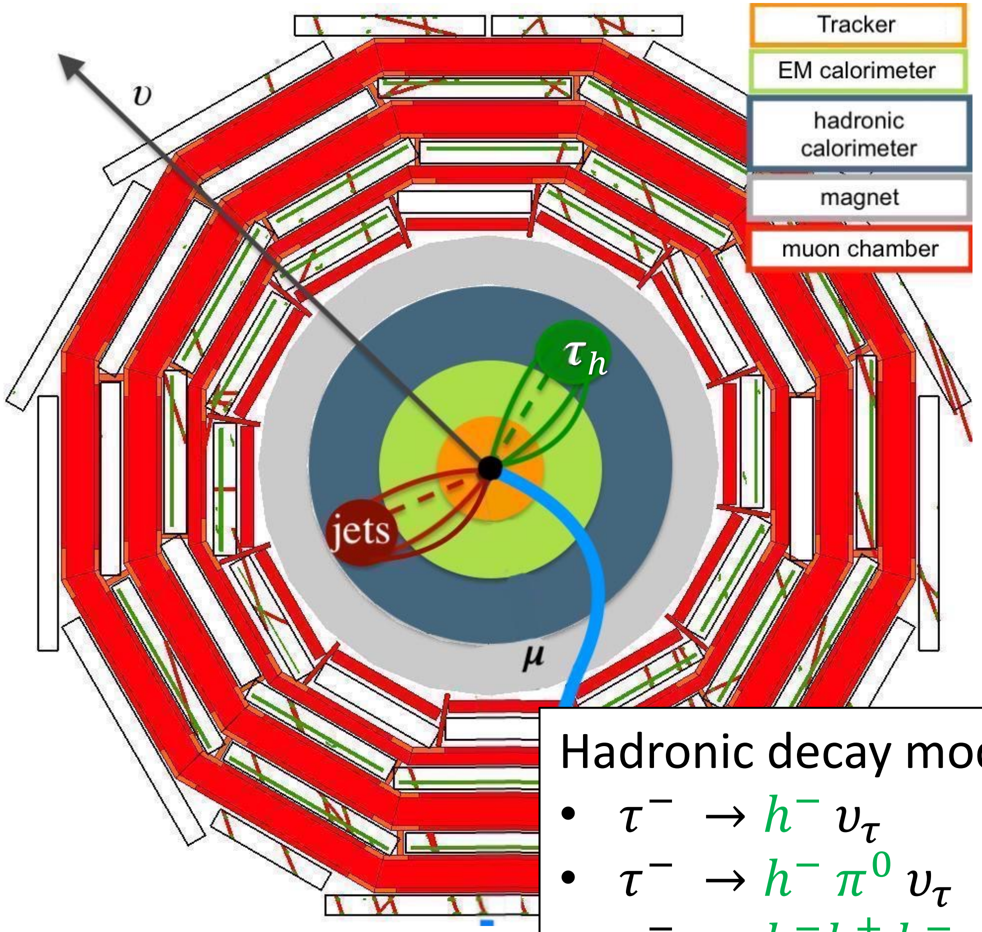
Hadronic decay modes:

- $\tau^- \rightarrow h^- \upsilon_\tau$
- $\tau^- \rightarrow h^- \pi^0 \upsilon_\tau$
- $\tau^- \rightarrow h^- h^+ h^-$
- $\tau^- \rightarrow h^- h^+ h^- \pi^0 \upsilon_\tau$

Precisely reconstructed

II.A. τ_h reconstruction and identification in CMS

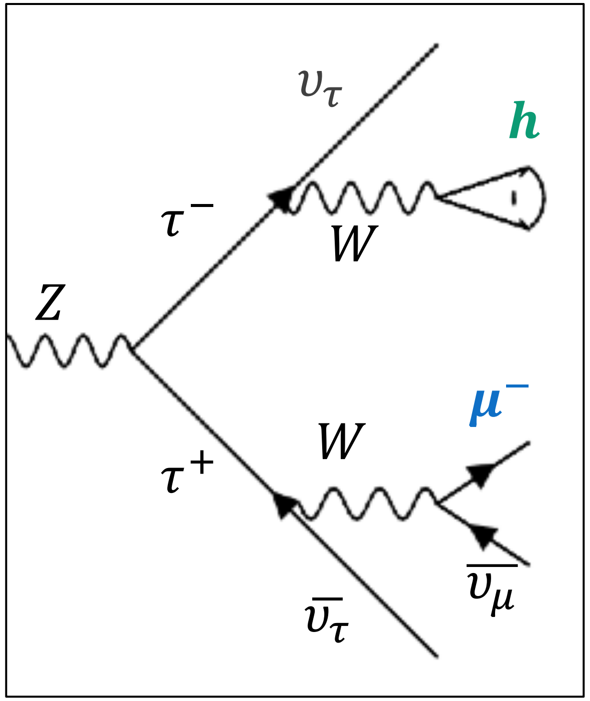
Transverse view of CMS



Hadronic decay modes:

- $\tau^- \rightarrow h^- \upsilon_\tau$
- $\tau^- \rightarrow h^- \pi^0 \upsilon_\tau$
- $\tau^- \rightarrow h^- h^+ h^-$
- $\tau^- \rightarrow h^- h^+ h^- \pi^0 \upsilon_\tau$

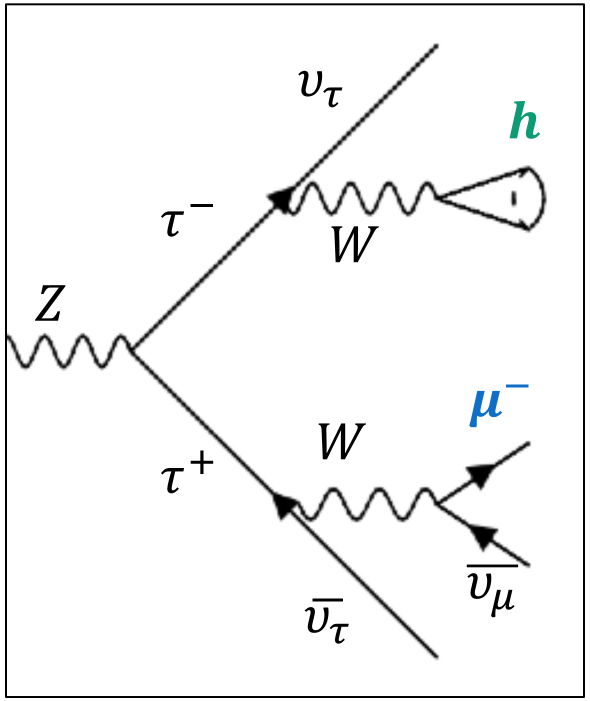
Calibration of τ_h



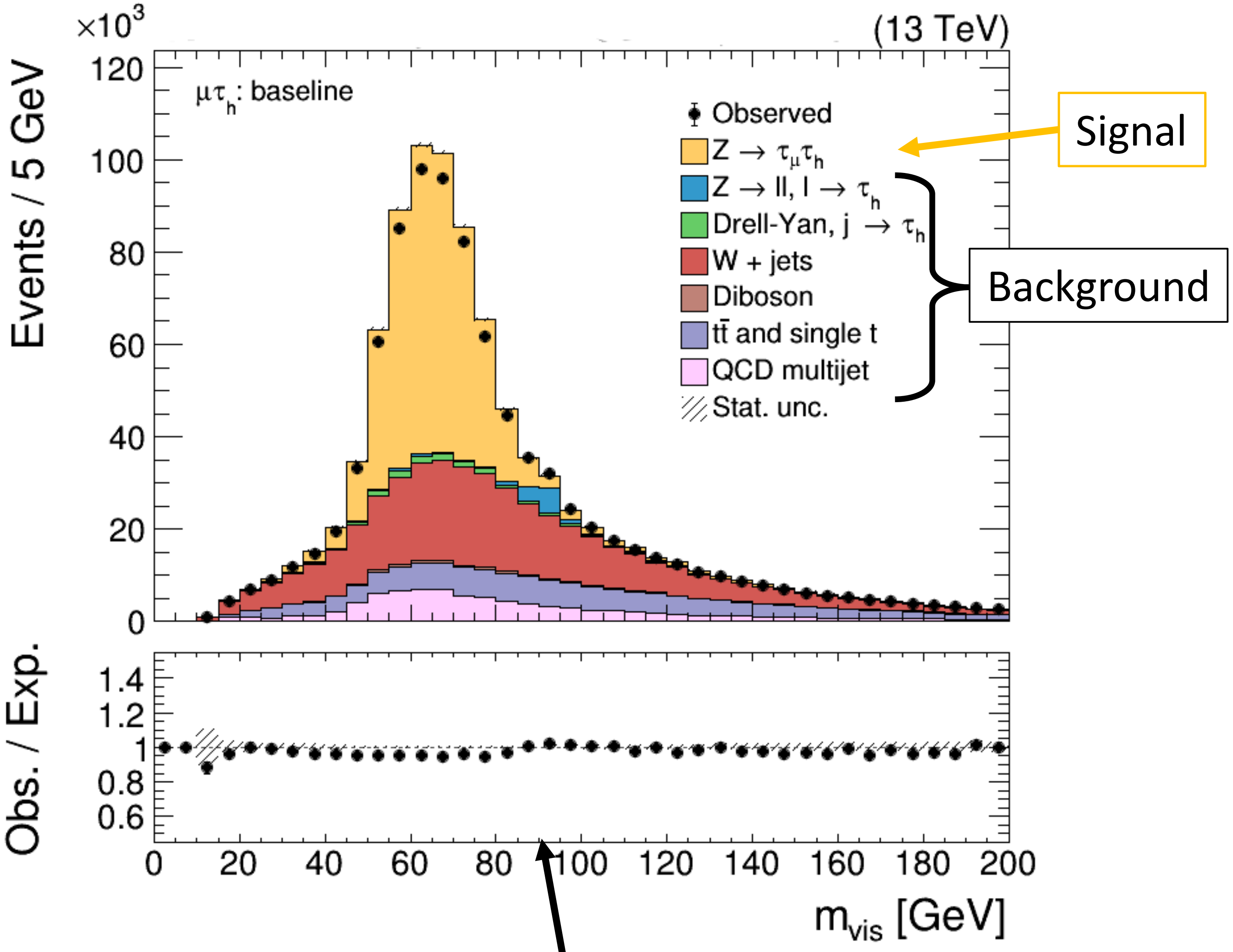
Event samples with $Z \rightarrow \tau_\mu \tau_h$ decay are used as reference process for the calibration.

II.B. Calibration

m_{vis} = invariant mass of visible decay products



$$m_{vis} = \sqrt{(\sum E_h + E_\mu)^2 - (\sum \vec{p_h} + \vec{p_\mu})^2}$$



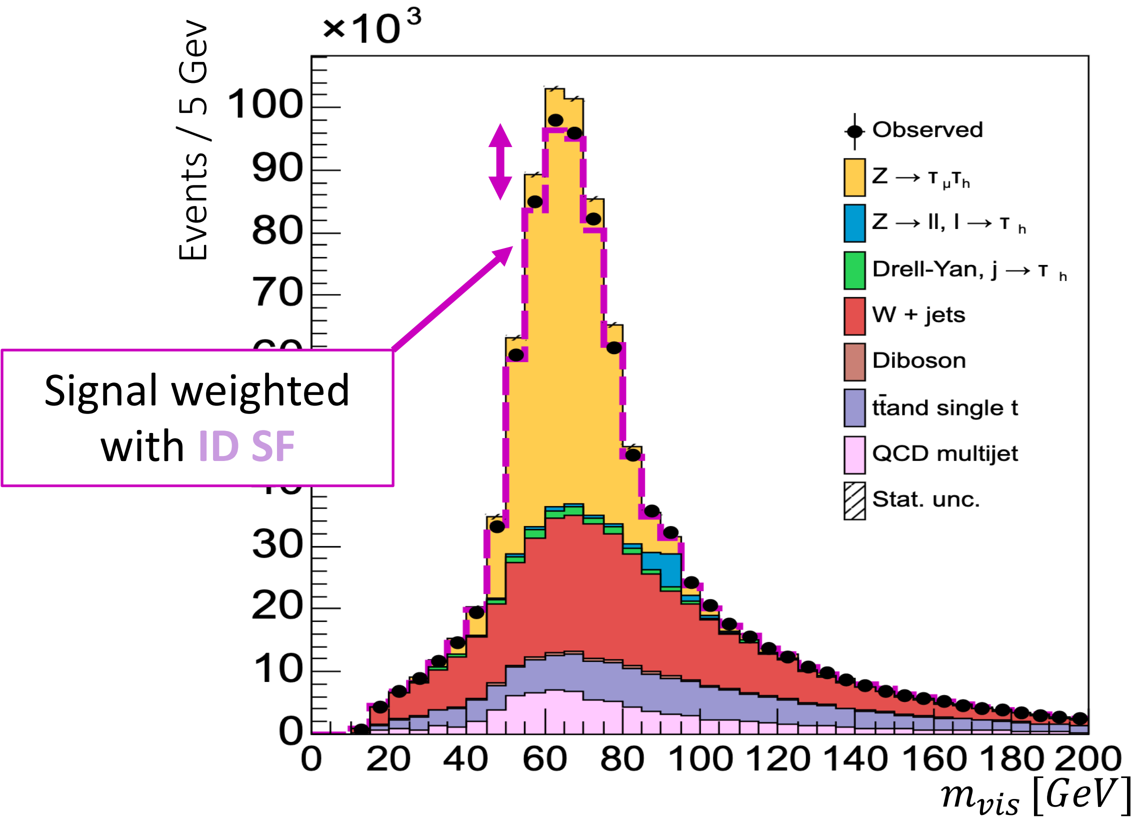
$m_Z \approx 91\ GeV$

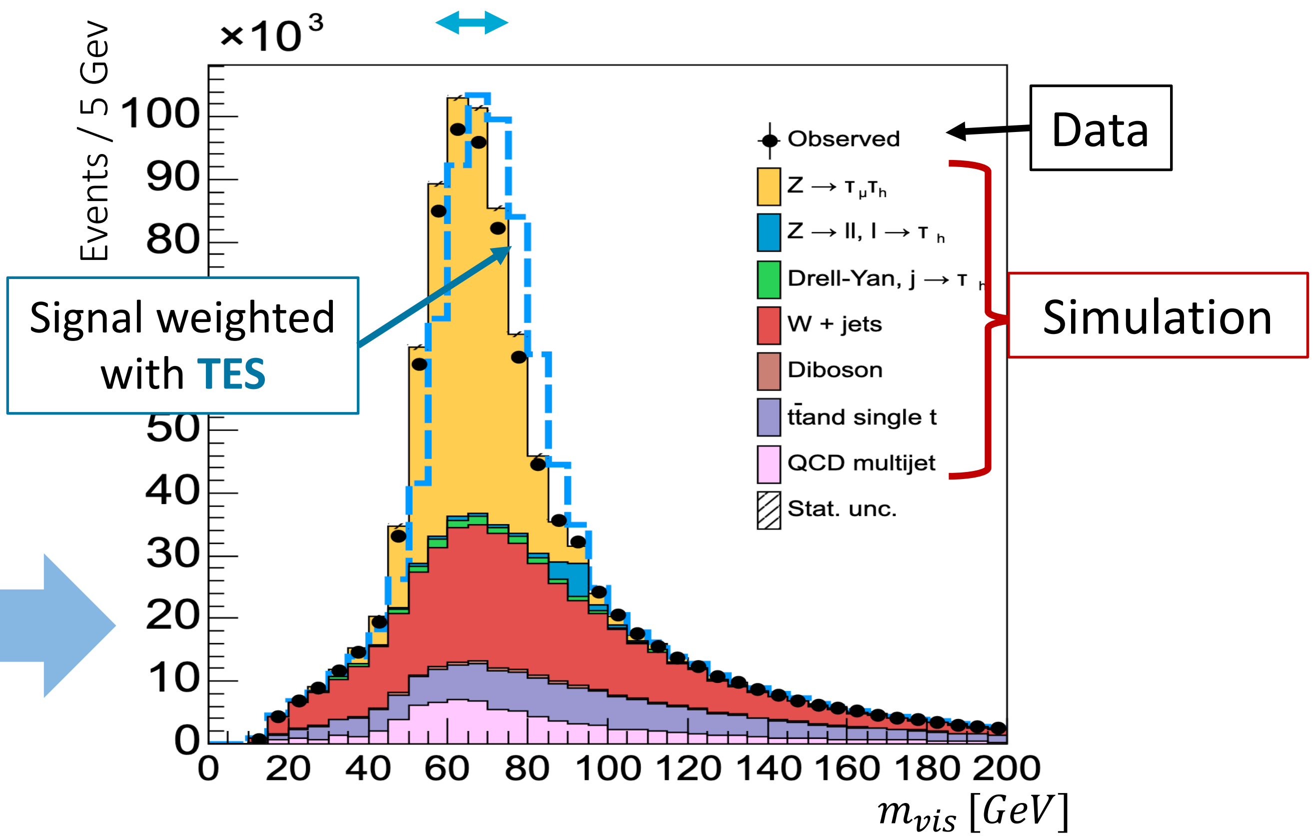
II.B. Calibration

Correction factors adjust the m_{vis} distribution of the simulation on the data.

τ_h **identification efficiency scale factor (ID SF) adjust the normalization** of the signal.

τ_h **energy scale (TES) adjust the energy** of the signal.





II.B. Calibration : Maximum likelihood fit

Correction factors adjust the m_{vis} distribution of the simulation on the data.

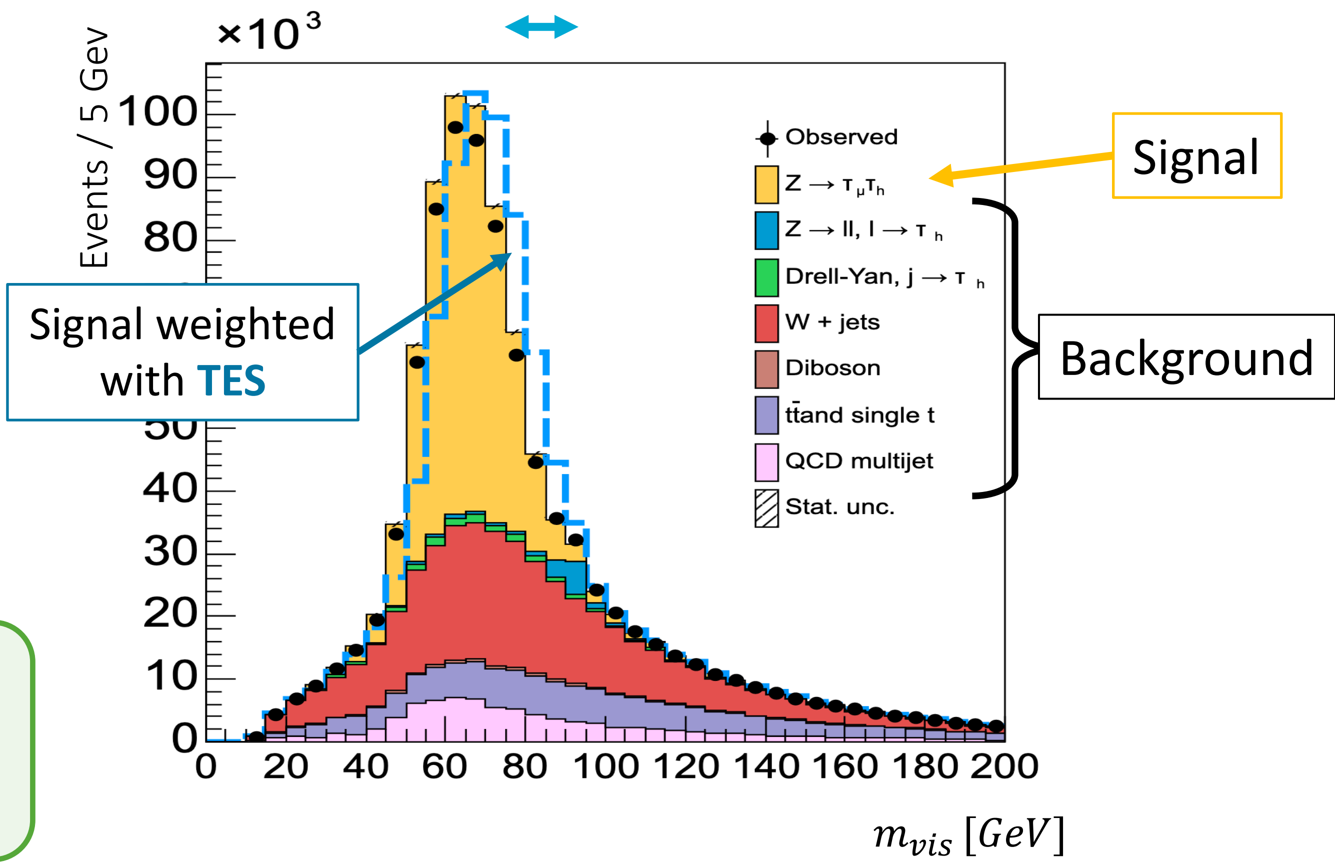
The ID SF and the TES are measured with **maximum likelihood fit** that maximize the likelihood function:

$$\mathcal{L}(\text{TES}, \theta) = p(\text{data}|\text{TES}, \theta) \prod_{i=1}^{n} p(\bar{\theta}_i|\theta_i)$$

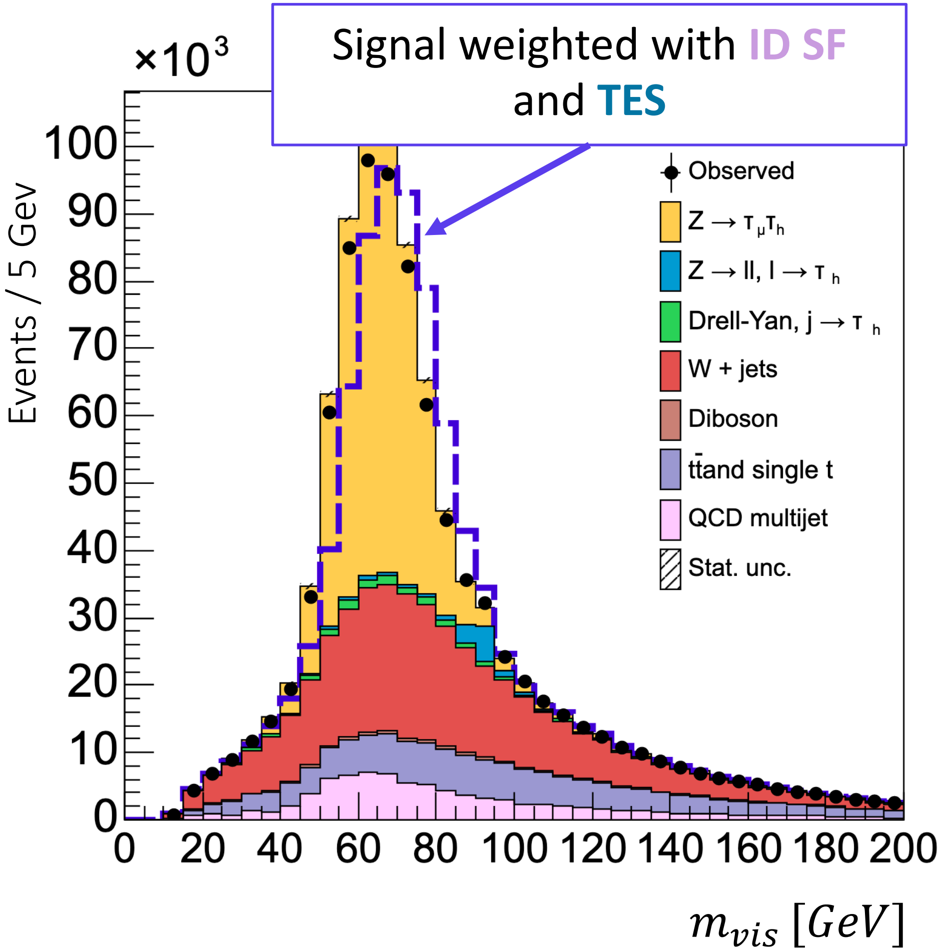
Probability of getting the data given the TES and θ.

Likelihood of the mean value $\bar{\theta}_i$ for each value of the syst. uncert. θ_i.

τ_h energy scale (TES)



II.B. Calibration : simultaneous fit



Current approach in CMS:
The ID SF and the TES are fitted **separately** on the same distributions.

Likelihood : $\mathcal{L}_{1,DM}$(**ID SF**$_{\mathbf{DM}}$, syst. uncert.$_{DM}$)
et $\mathcal{L}_{2,DM}$ (**TES**$_{DM}$, syst. uncert.$_{DM}$, **ID SF**$_{\mathbf{DM}}$ uncert.)

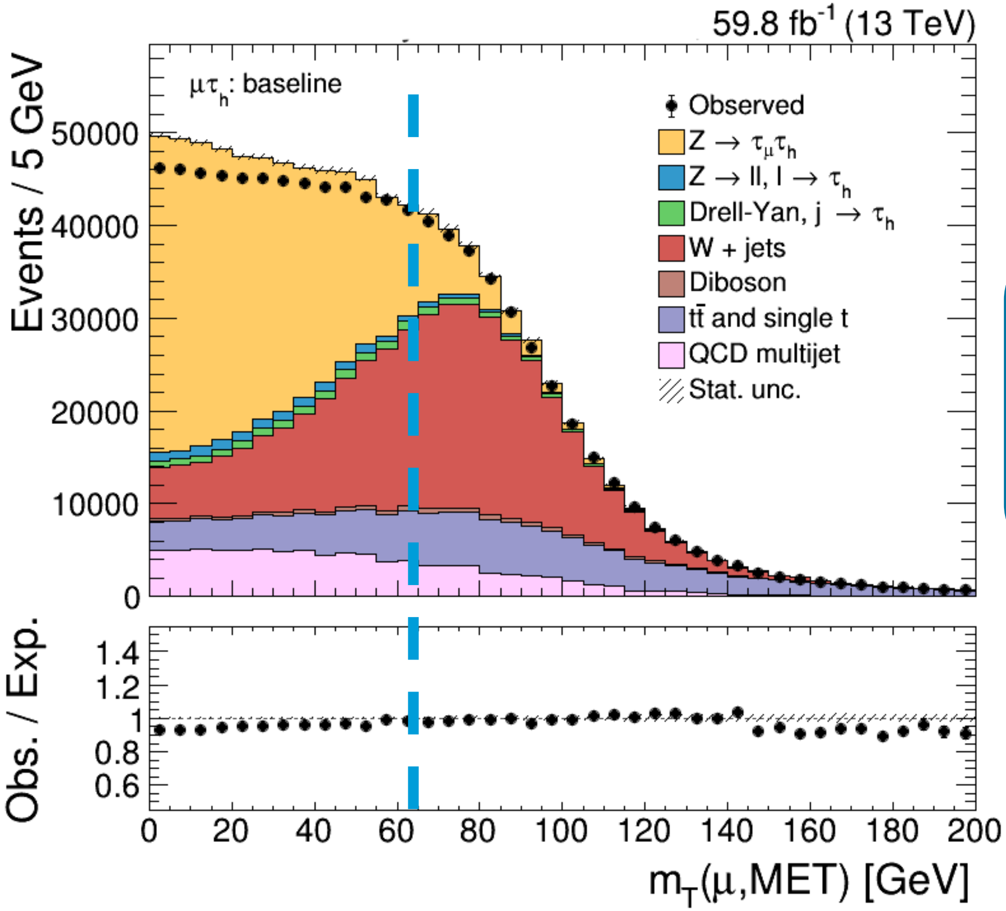
New method :
Simultaneous fit of the ID SF and the **TES.**

Likelihood : $\mathcal{L}_{DM}$(**ID SF**$_{\mathbf{DM}}$,**TES**$_{DM}$, syst. uncert.$_{DM}$)

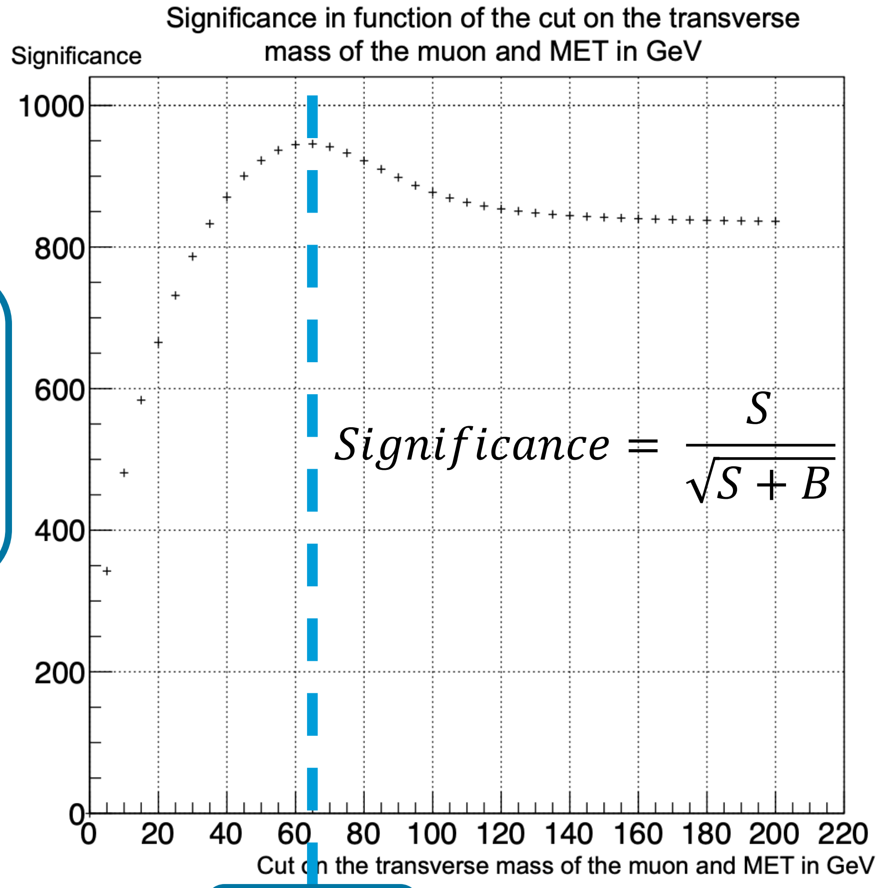
Goal :
↪ simplification of the interpretation of the results
↪ possible reduction of uncertainties

III.A. Study of the tau energy scale fit : Minimization of the background

Why ? To optimize the fit and reduce the uncertainties.
How ? Applying cuts on variables that **discriminates** the background and signal easily → $m_T(\mu, MET)$



The cut that **maximizes** this **significance** **minimizes** the **statistical uncertainties**

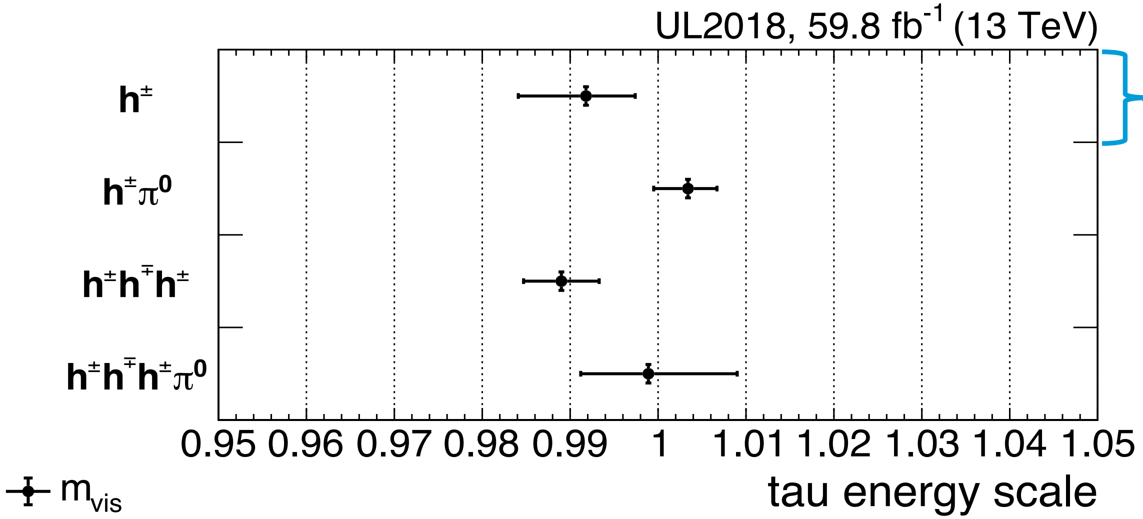


$$Significance = \frac{S}{\sqrt{S+B}}$$

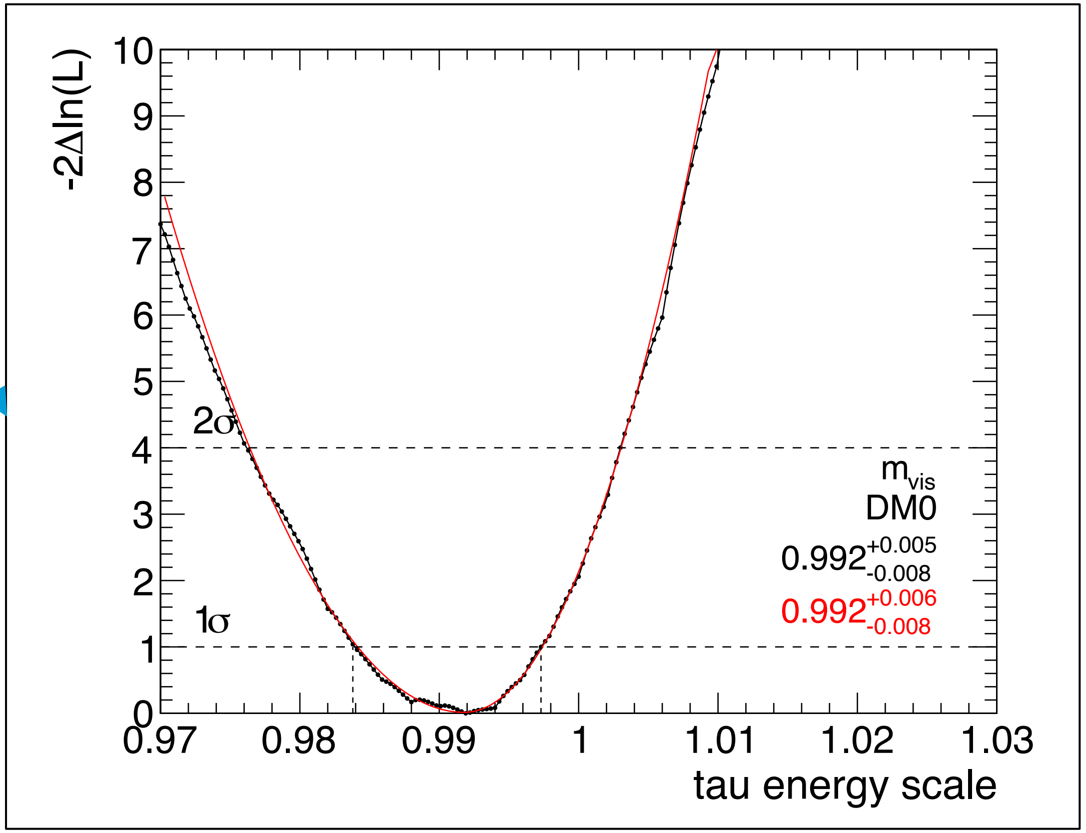
65 GeV

III.B. Study of the tau energy scale fit : the TES fit

Likelihood to maximize :
$\mathcal{L}_{DM}$ (**TES**, syst. uncert.$_{DM}$, $\mathbf{ID\ SF_{DM}}$ uncert.)



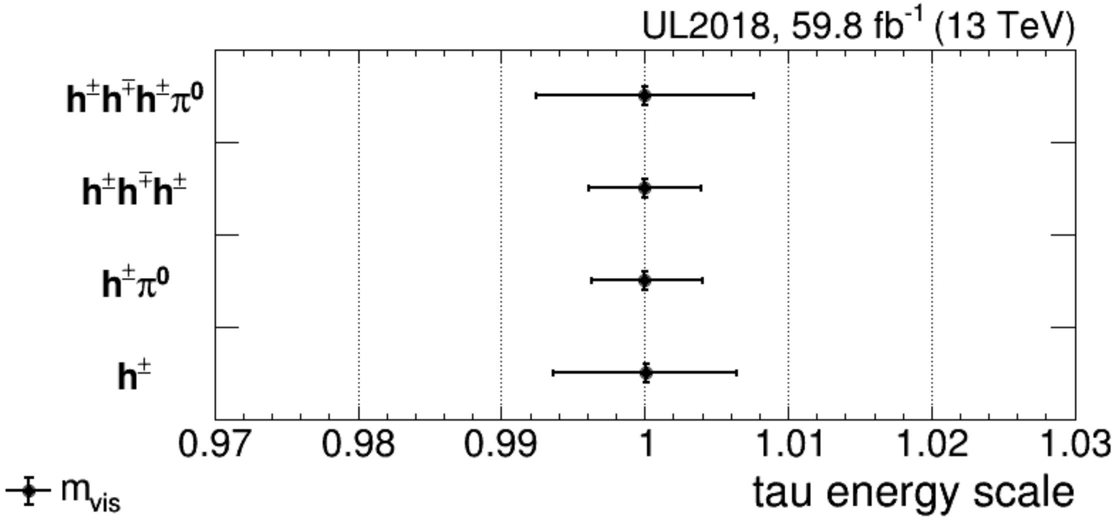
The **negative log-likelihood profile** −2ΔL as a function of the scanned value of the TES. Example for DM0: $h^{\pm}$



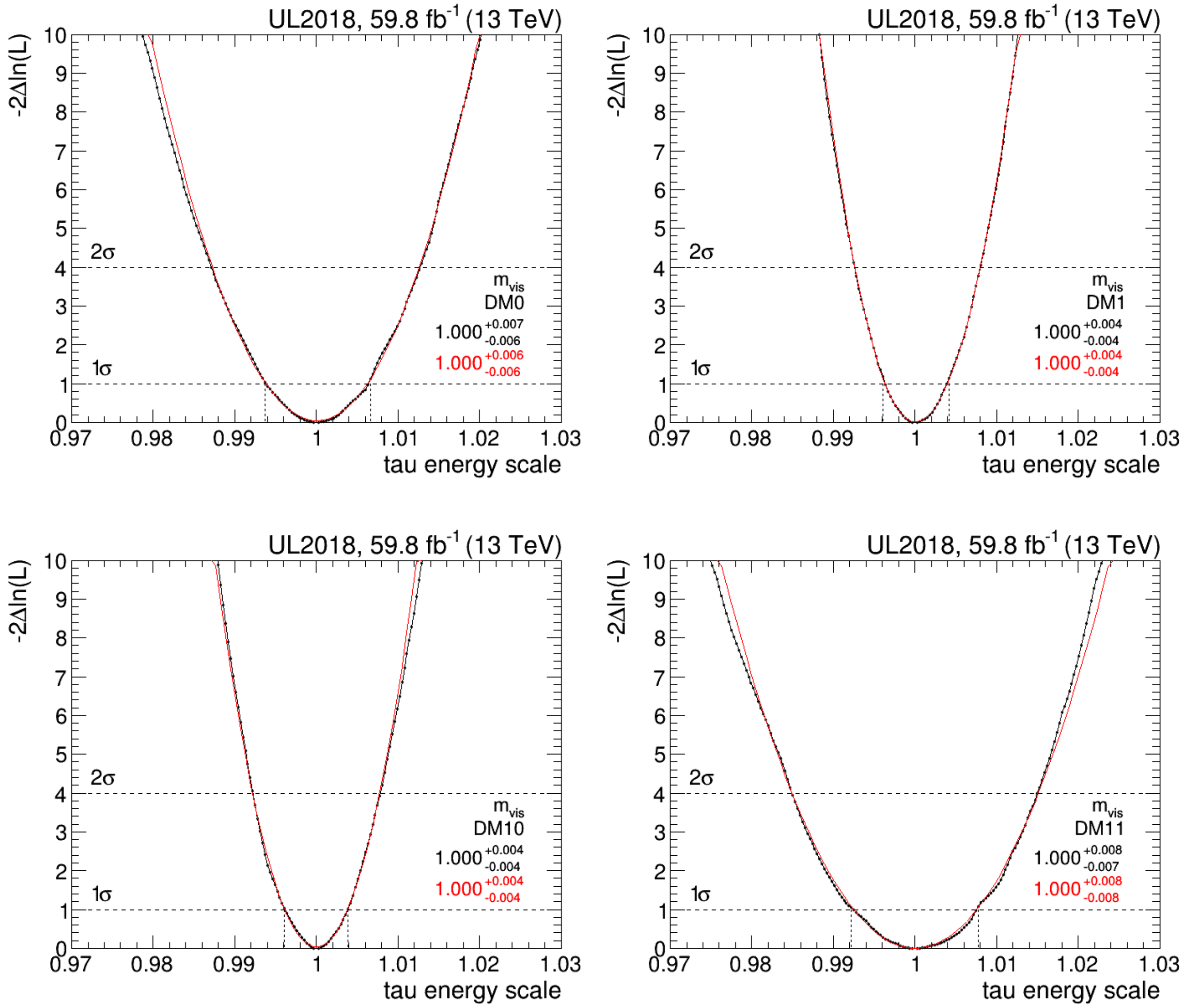
- The likelihood profiles are **fitted** with asymmetric **parabolae** (in red)
- Agreement with the current results of CMS.

III.C. Study of the tau energy scale fit : Asimov dataset

Asimov data set = "perfect data set" = all observed quantities are set equal to their **expected values** and there **are no statistical fluctuations**.



Useful to study the source of **instability** in the fit and **uncertainties**.



III.D. Study of the tau energy scale fit : uncertainties

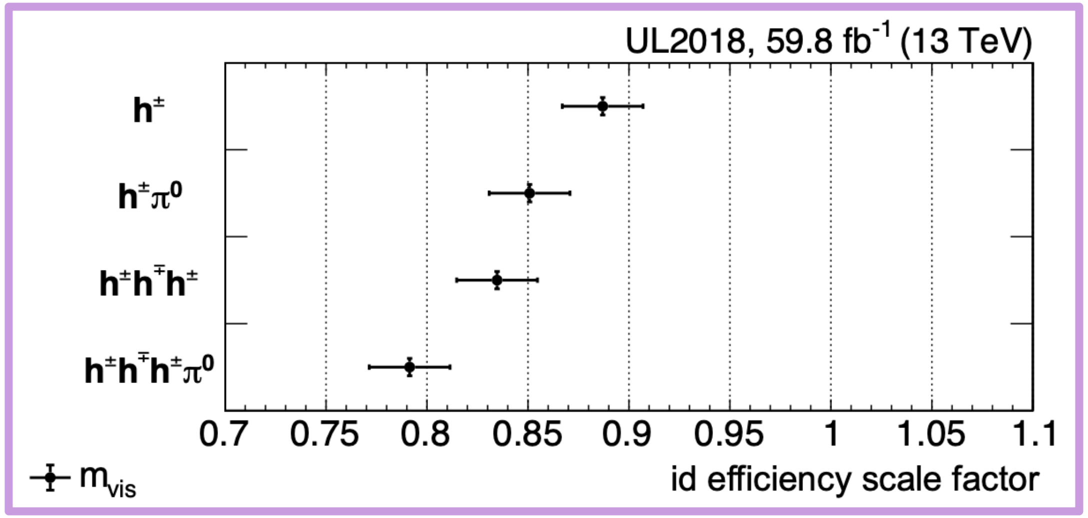
Total uncertainties on the TES: $\sigma_{tot}^2 = \sigma_{stat}^2 + \sigma_{syst}^2 + \sigma_{simul}^2$

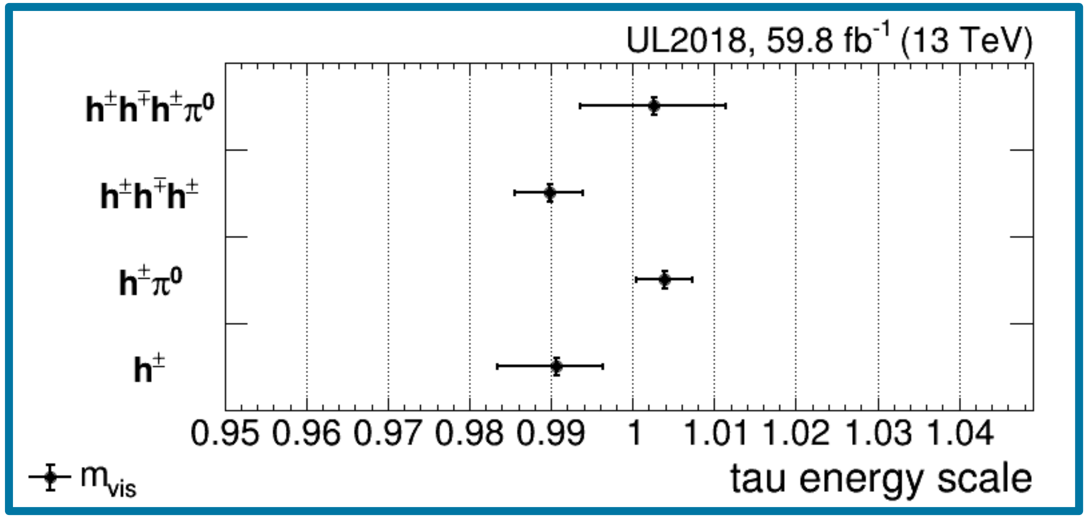
Decay mode (DM)	DM0 : $h^{\pm}$	DM1 : $h^{\pm}\,\pi^0$	DM10 : $h^{\pm}h^{\mp}h^{\pm}$	DM11 : $h^{\pm}h^{\mp}h^{\pm}\,\pi^0$
Total uncer. (σ_{tot}^2)	-0.0060 +0.0069	-0.0036 +0.0045	-0.0045 +0.0039	-0.0078 +0.0078
Syst. uncertainty (σ_{syst}^2)	**-0.0043** **+0.0055**	**-0.0029** **+0.0036**	**-0.0030** +0.0021	-0.0046 +0.0041
Stat uncert. on the simulation (σ_{simul}^2)	-0.0038 +0.0035	-0.0017 +0.0027	-0.0029 **+0.0025**	**-0.0054** **+0.0051**
Stat. uncer (σ_{stat}^2)	-0.0018 +0.0024	-0.0012 +0.0020	-0.0015 +0.0021	-0.0033 +0.0042

Systematic uncertainties and statistical uncertainty on the simulation are dominant.

IV.A. The new method : Simultaneous fit of the TES and the ID SF

Likelihood : $\mathcal{L}_{DM}(ID\ SF_{DM}, \mathbf{TES}_{DM}, \text{syst. uncert.}_{DM})$





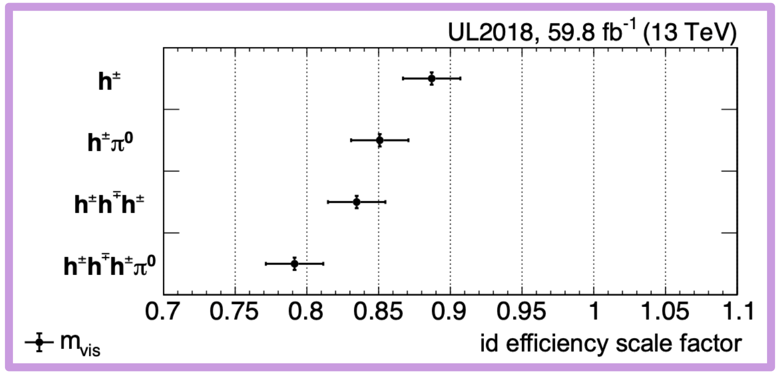
Simultanous fit by decay mode (DM):

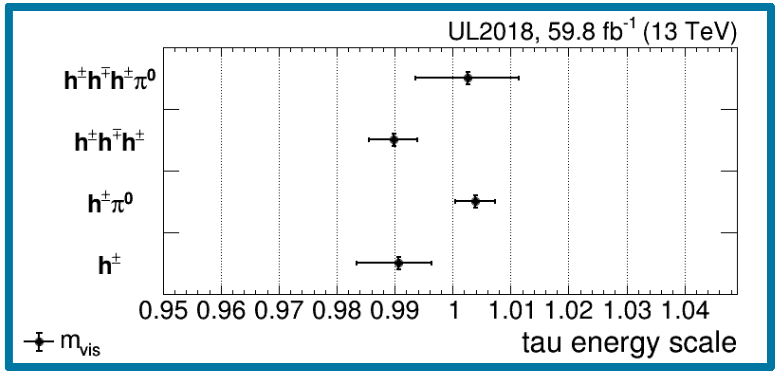
- The values of the ID SF are different of 1 including the error bars. It shows the **importance of the calibration** step.
- The values of the ID SF seem to **decrease linearly** as a function of the number of particles in the decay mode.
- Results are **consistent** with the previous ones obtained with a separated fit.

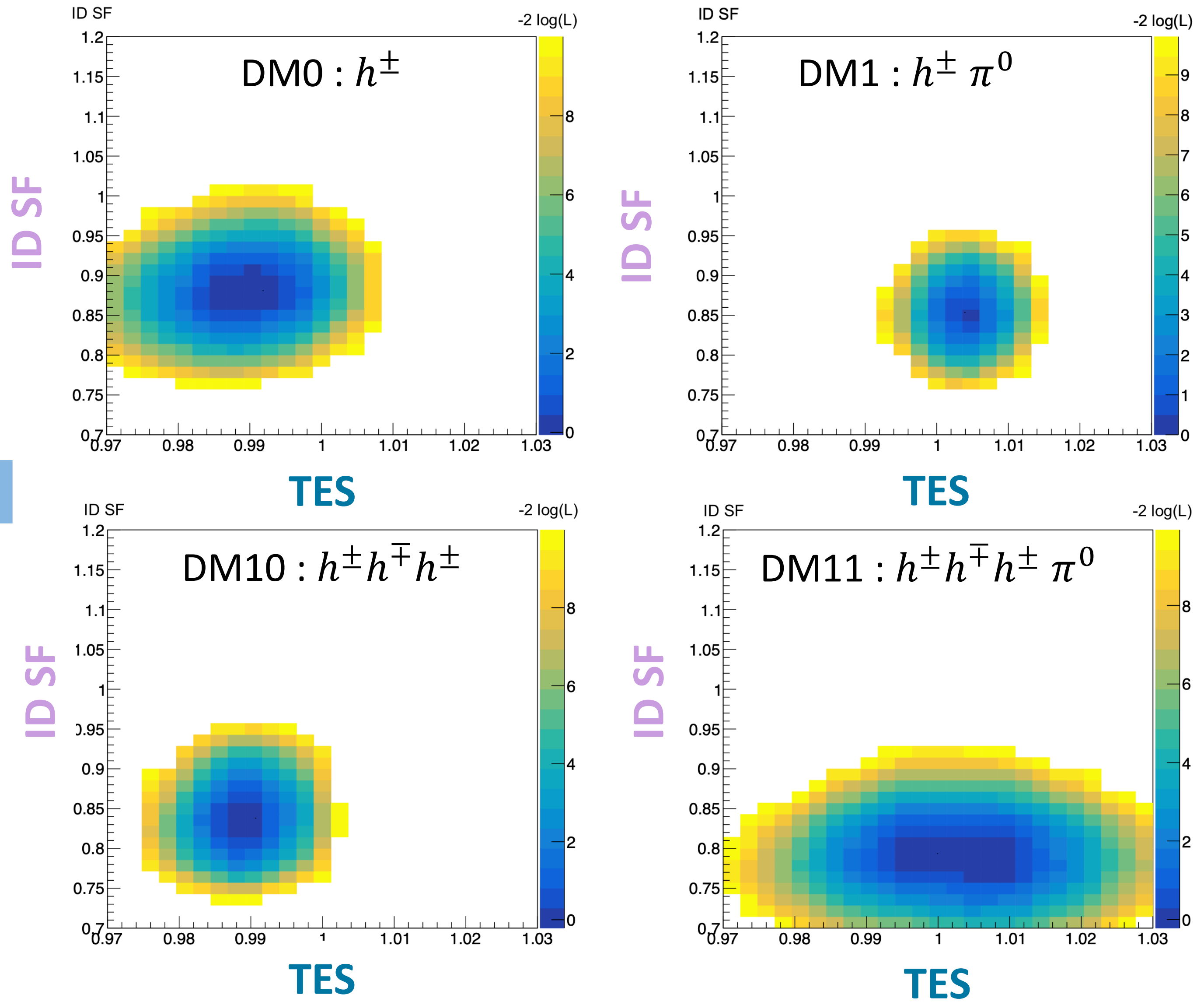
IV.A. The new method : Simultaneous fit of the TES and the ID SF

Likelihood :

$\mathcal{L}_{DM}(ID\ SF_{DM}, \mathbf{TES}_{DM}, \text{syst. uncert.}_{DM})$







2D log likelihood profiles for each decay mode (DM)

IV.B. New method : Combined fit of the TES and the ID SF in several regions

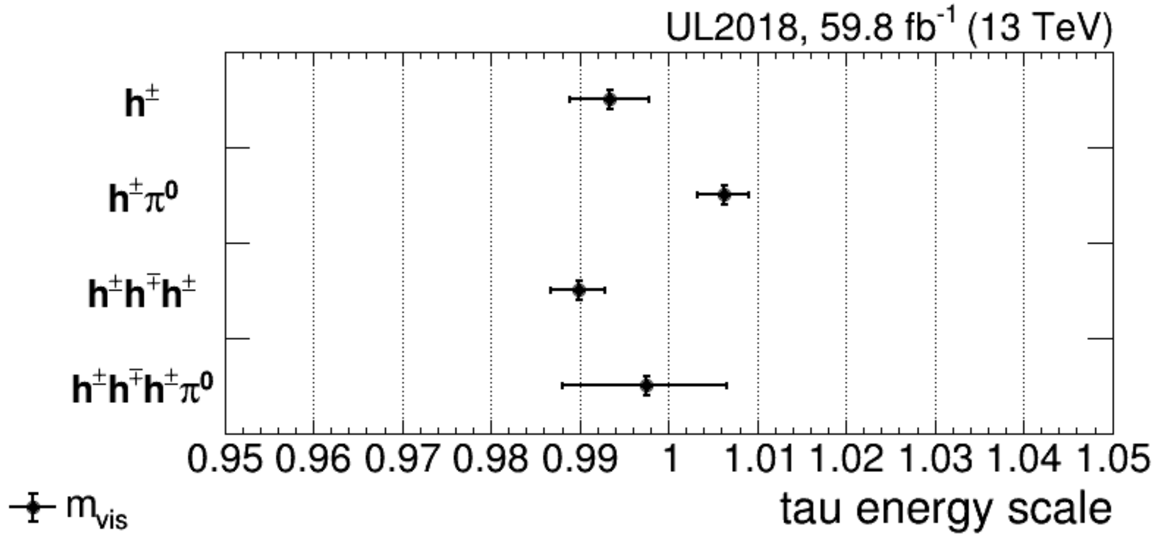
The TES and the ID SF depend on DM and kinetic variables such as $p_T(\tau_h)$.

↪ **New regions** of the fit has been created.

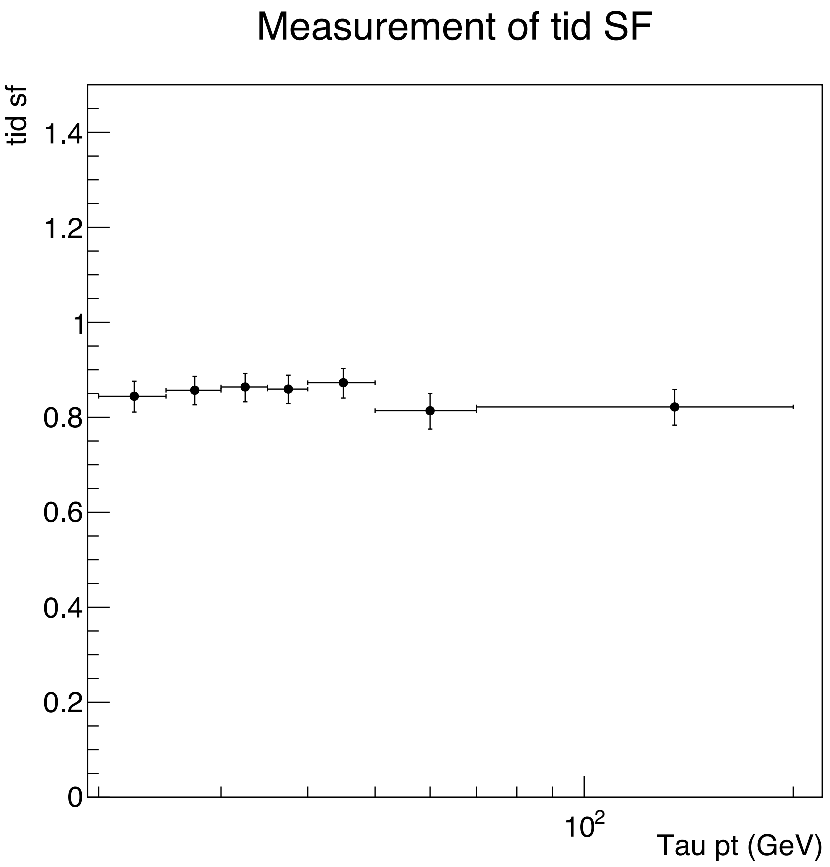
The common uncertainties can be constrained to stabilize the fit.

↪ The fit is done **simultaneously on the regions**.

The TES are fitted by decay mode (DM).



The ID SF are fitted by $p_T(\tau_h)$ regions.



Likelihood : $\mathcal{L}(ID\ SF_{p_{T1}} \ldots ID\ SF_{p_{T7}}, \mathbf{TES}_{DM0} \ldots \mathbf{TES}_{DM11}, syst.\ \text{uncert.})$

Conclusion

Summary

Study of the TES fit shows that the **fit are dominated by the systematic uncertainties** and mostly by **the statistical uncertainty of the simulation** → will be improved in next data taking.

For the **1st time**, a **simultaneous fit of the TES and the ID SF** has been implemented.

- By decay mode (DM) → It gives better interpretation of the results and of the uncertainties.
- By kinematic **regions** and DM **fitted simultaneously** → implemented in a flexible way such that it is possible to easily add or modified the regions of the fit.

Outlook

The profile likelihood scans of the ID SF are smooth, but the ones of the TES show fluctuations. Further investigations are needed to find how to reduce them.

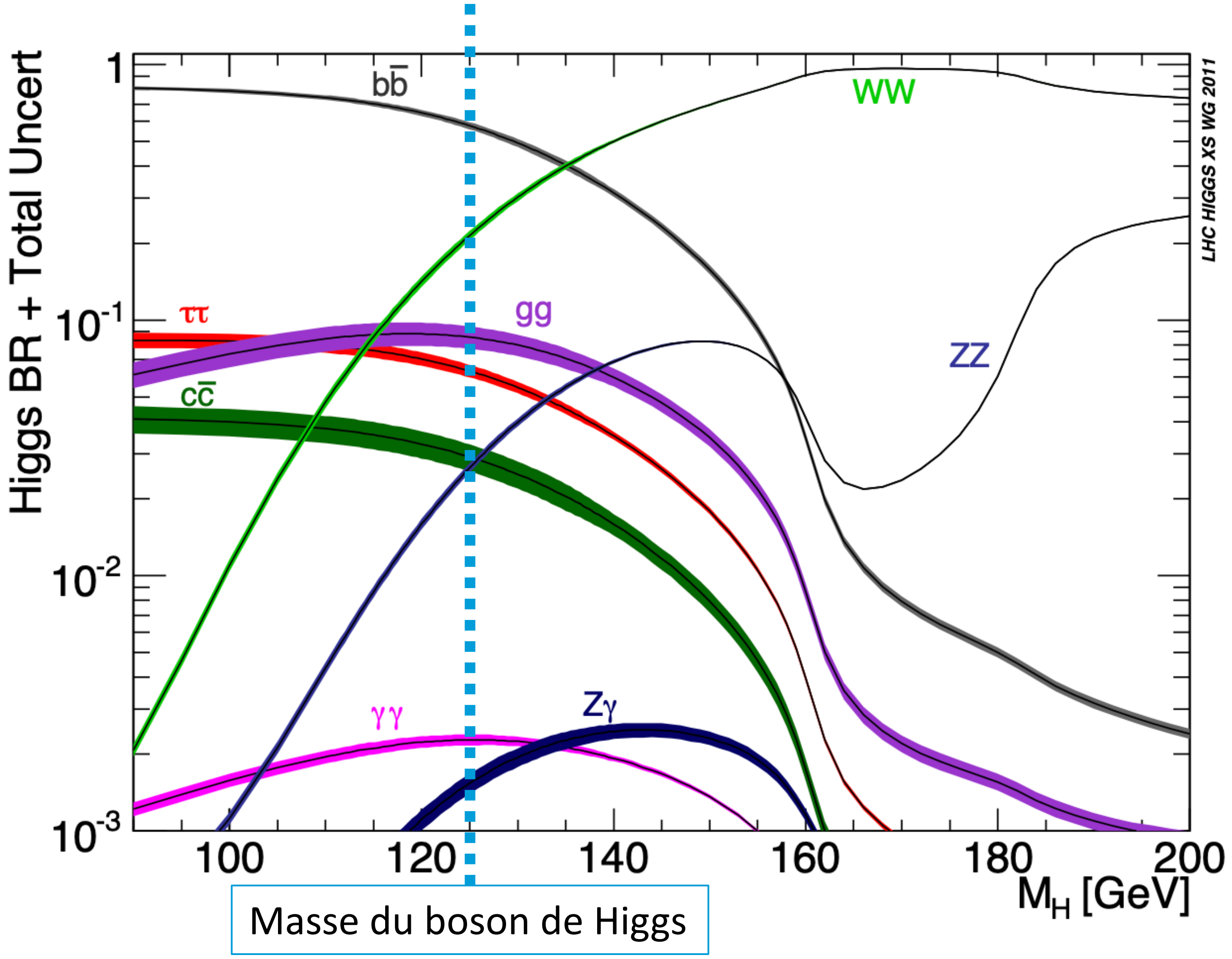
This new method will be useful for **analysis using** $\boldsymbol{\tau_h}$ such as the search of charge-parity (CP) violation in $H \rightarrow \tau\tau$ decay and for the new data taking.

Questions ?

Backup

Higgs boson decay

Branching Ratio: The probability that a particle decays in one decay mode out of all of its decay modes.



Vocabulary

- **Electron volt (eV) :** 1 eV = 1,602 176 634 × 10^{-19} J → 1 TeV = 10^{12} eV = 1,602 177 × 10^{-7} J
- **hadrons :** composite particle made up of quarks (and gluons) governed by the strong interaction (e.g. proton, neutrons)
- **Transverse momentum $\boldsymbol{p_T}$** : momentum in the transverse plane of the beam .
- **Invariante mass** : $\mathrm{M} = \sqrt{(\sum_i E_i)^{2} - (\sum_i p_i)^{2}}$
- $m_T(\mu, MET) = \sqrt{\left(E_{T,\mu} + E_{T,MET}\right)^2 - \left(\overrightarrow{p_{T,\mu}} + \overrightarrow{p_{T,MET}}\right)^2}$

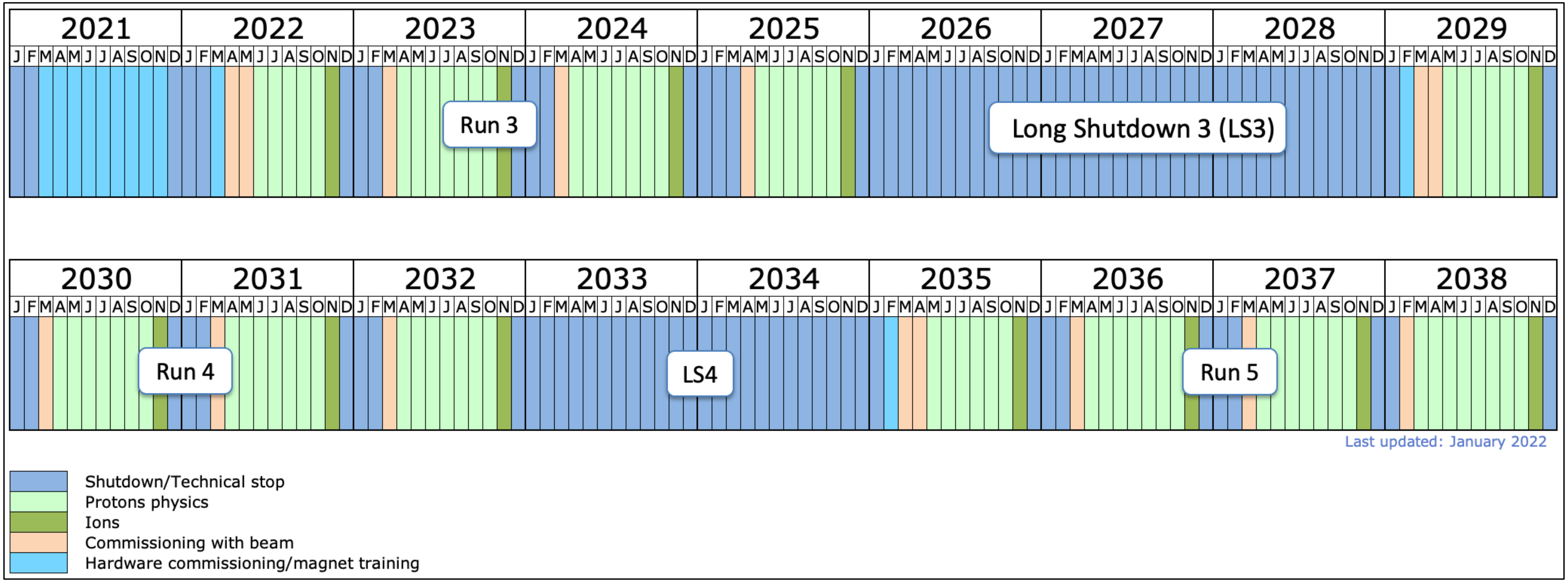
Systematci uncertainties

Tau ES systematic uncertainties

DY = Drell-Yan MC (ZTT + ZL + ZJ)
ZTT = DY, real τ_h
ZL = DY, $\ell \to \tau_h$ fake
ZJ = DY, $j \to \tau_h$ fake
ttbar = TTT + TTL + TTJ

nuissance parameter	distribution	uncertainty	applied to
luminosity	lnN	±2.5%	all, except QCD
muon efficiency	lnN	±2%	all, except QCD
tau ID	**shape**	**from recommendation**	ZTT, TTT
DY cross section	lnN	±2%	DY
ttbar cross section	lnN	±6%	ttbar
single top cross section	lnN	±5%	single top
diboson cross section	lnN	±5%	diboson
W + jets normalization	lnN	**±8%**	WJ
QCD normalization	lnN	**±10%**	QCD
$j \to \tau_h$ fake rate	lnN	**±15%**	ZJ, WJ, QCD, TTJ, STJ
$j \to \tau_h$ fake energy scale	shape	**±5% on $j \to \tau_h$ energy**	ZJ, W, TTJ
$\ell \to \tau_h$ fake rate	**shape**	**from recommendation**	ZL, TTL
$\ell \to \tau_h$ fake energy scale	shape	±2% on $\ell \to \tau_h$ energy	ZL, TTL
Z p_T reweighting	shape	apply weight ±10%	DY
bin-by-bin	shape		all

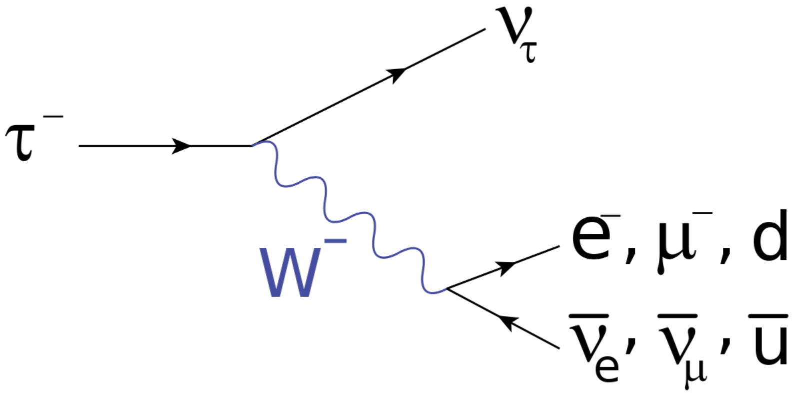
LHC Schedule



τ leptons

Table 1: Decays of τ leptons and their branching fractions ($\mathcal{B}$) in % [59]. The known intermediate resonances of all the listed hadrons are indicated where appropriate. Charged hadrons are denoted by the symbol $h^{\pm}$. Although only τ^- decays are shown, the decays and values of the branching fractions are identical for charge-conjugate decays.

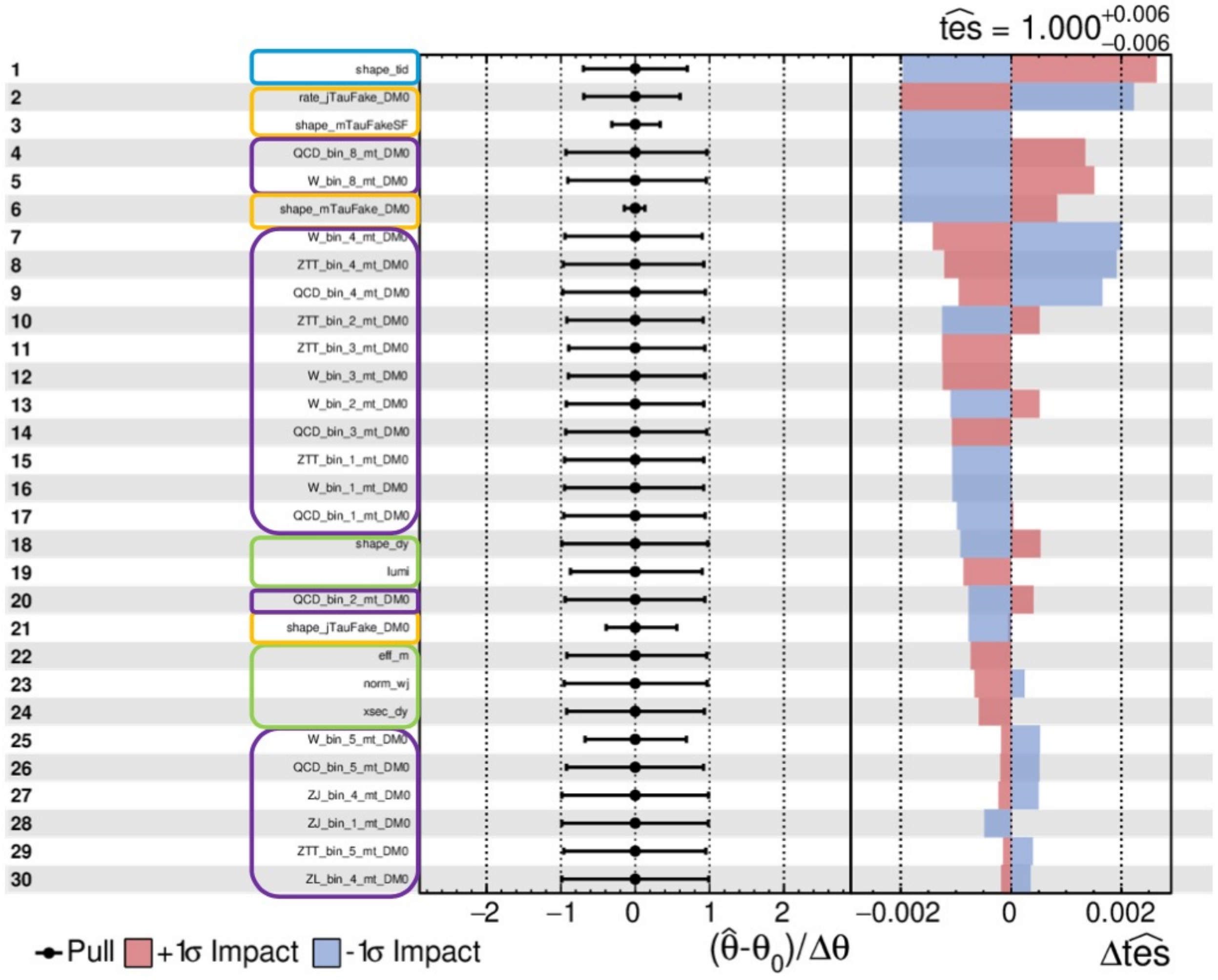
Decay mode	Resonance	$\mathcal{B}$ (%)	
Leptonic decays		35.2	
$\tau^- \to e^- \bar{\nu}_e \nu_\tau$			17.8
$\tau^- \to \mu^- \bar{\nu}_\mu \nu_\tau$			17.4
Hadronic decays		64.8	
$\tau^- \to h^- \nu_\tau$			11.5
$\tau^- \to h^- \pi^0 \nu_\tau$	$\rho(770)$		25.9
$\tau^- \to h^- \pi^0 \pi^0 \nu_\tau$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



m_T = 1 777 MeV.c^{-2}
lifetime τ_τ = 2.8×10^{-13} s

III.D. Study of the tau energy scale fit : Systematic uncertainties

Total uncertainties on the TES: $\sigma_{tot}^2 = \sigma_{stat}^2 + \sigma_{syst}^2$ ← **dominant**



Main source of systematic uncertainties :

⇒ ID SF uncert. : should be reduced with combine fit of the TES and the ID SF

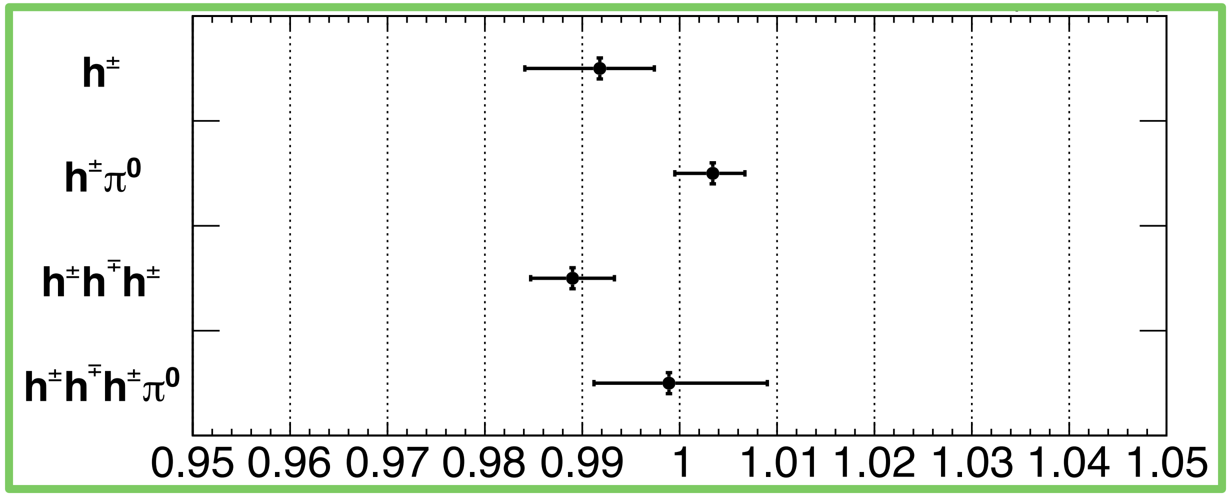
⇒ Fake tau: need further investigations

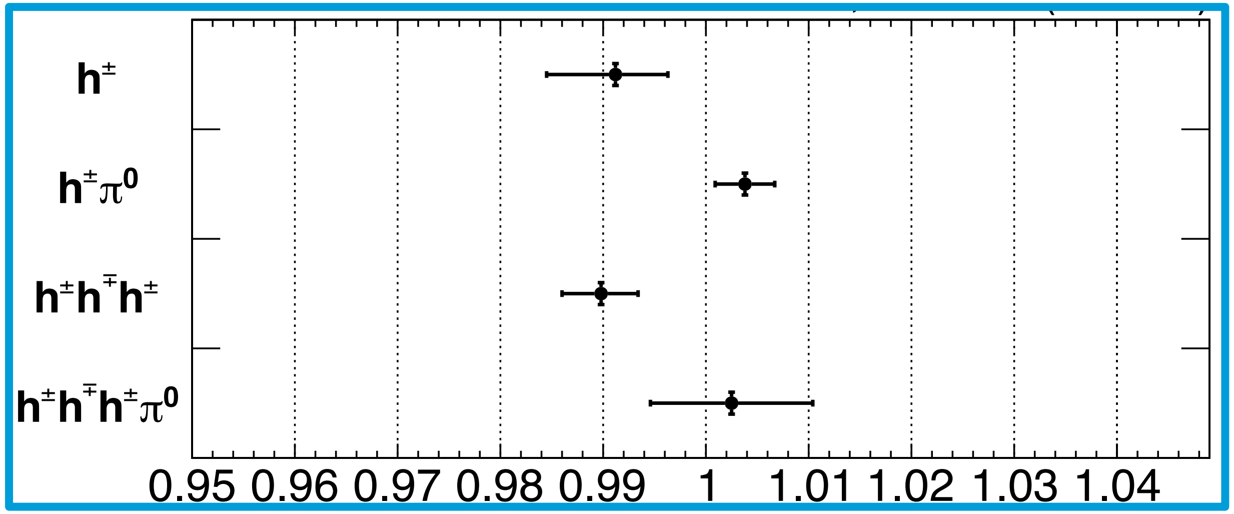
⇒ BBB uncert. : related to the statistics of the Monte Carlo

⇒ common uncertainties: constrain them with combine fit on regions

The TES fit by DM

Separate fit of the energy scale and the identification efficiency





	Séparé	Simultané
$h^{\pm}$	$0.9918^{-0.0077}_{\mathbf{+0.0056}}$	$0.9907^{\mathbf{-0.0073}}_{+0.0057}$
$h^{\pm}\pi^0$	$1.0034^{-0.0039}_{\mathbf{+0.0033}}$	$1.0039^{\mathbf{-0.0034}}_{+0.0034}$
$h^{\pm}h^{\mp}h^{\pm}$	$0.9890^{-0.0043}_{+0.0043}$	$0.9898^{-0.0043}_{\mathbf{+0.0041}}$
$h^{\pm}h^{\mp}h^{\pm}\pi^0$	$0.9989^{\mathbf{-0.0077}}_{+0.0101}$	$1.0025^{-0.0089}_{\mathbf{+0.0089}}$

Simultaneous fit of the energy scale and the identification efficiency

Searching for charge-parity (CP) violation in $H \rightarrow \tau^{+}\tau^{-}$ decay

CP properties :

CP properties of Higgs bosons = behavious under **C**harge and **P**aritey transformation.

C symmetry transforms the particle into its **antiparticule**:

→ No effect because the Higgs bson is neutral

P symmetry inverse the **spatial coordinates** :

→ If not effect, H is **scalar** or **CP even (SM) (pair)**

→ If transformed, H is **pseudo scalar or CP odd (impair)**

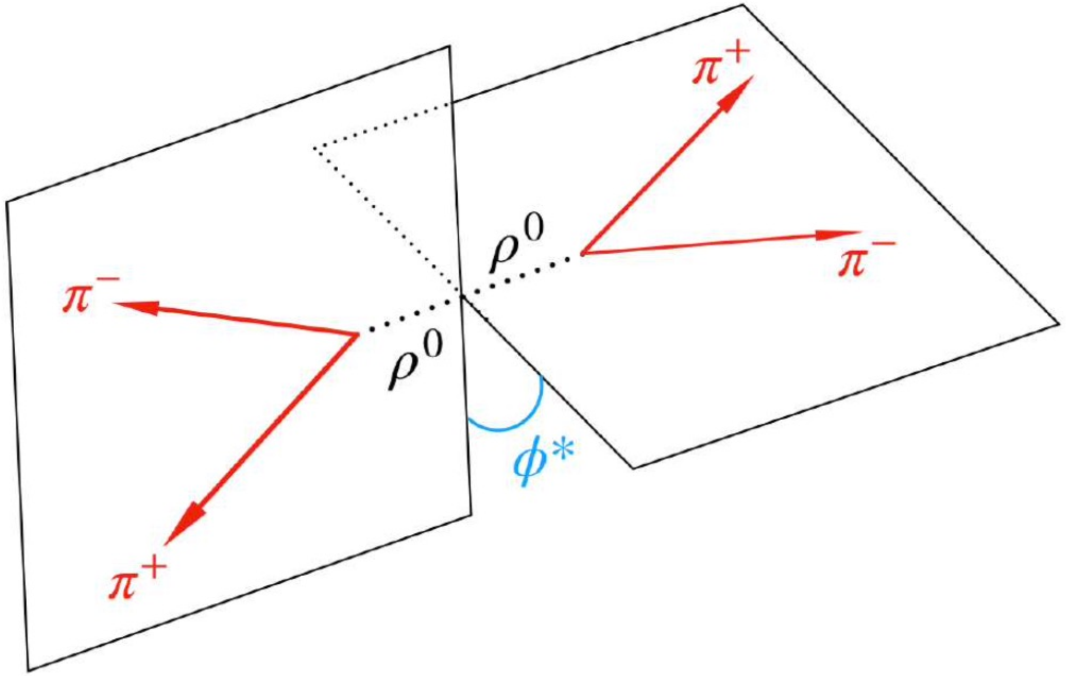
The Higgs boson transmits its CP properties to fermions through the Yukawa coupling: different for each fermion.

To measure the CP properties of the Higgs boson we need to measure Φ_{τ} which is senstive to CP properties.

Searching for charge-parity (CP) violation in $H \to \tau^+\tau^-$ decay

For a coupling to a pair of taus the Yukawa coupling is written:

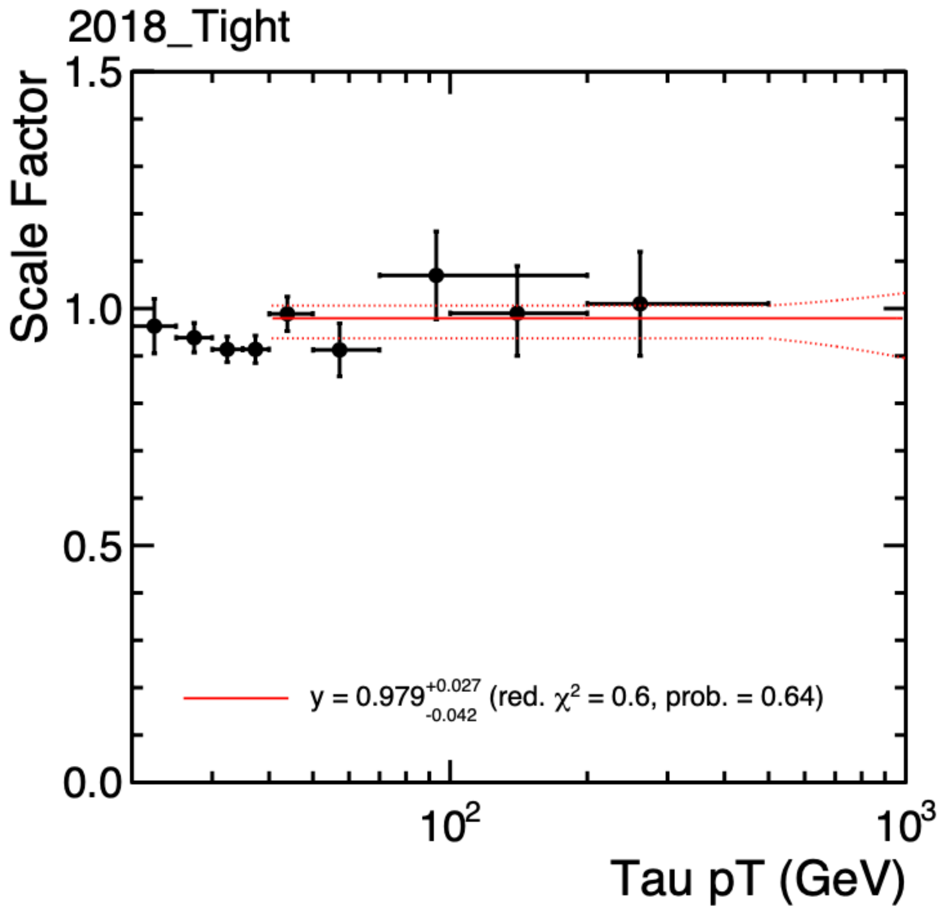
$$L_Y = -\frac{m_\tau}{\nu}\kappa_\tau(\cos\Phi_\tau\bar{\tau}\tau + \sin\Phi_\tau\bar{\tau}\tau i\gamma_5\tau)h$$



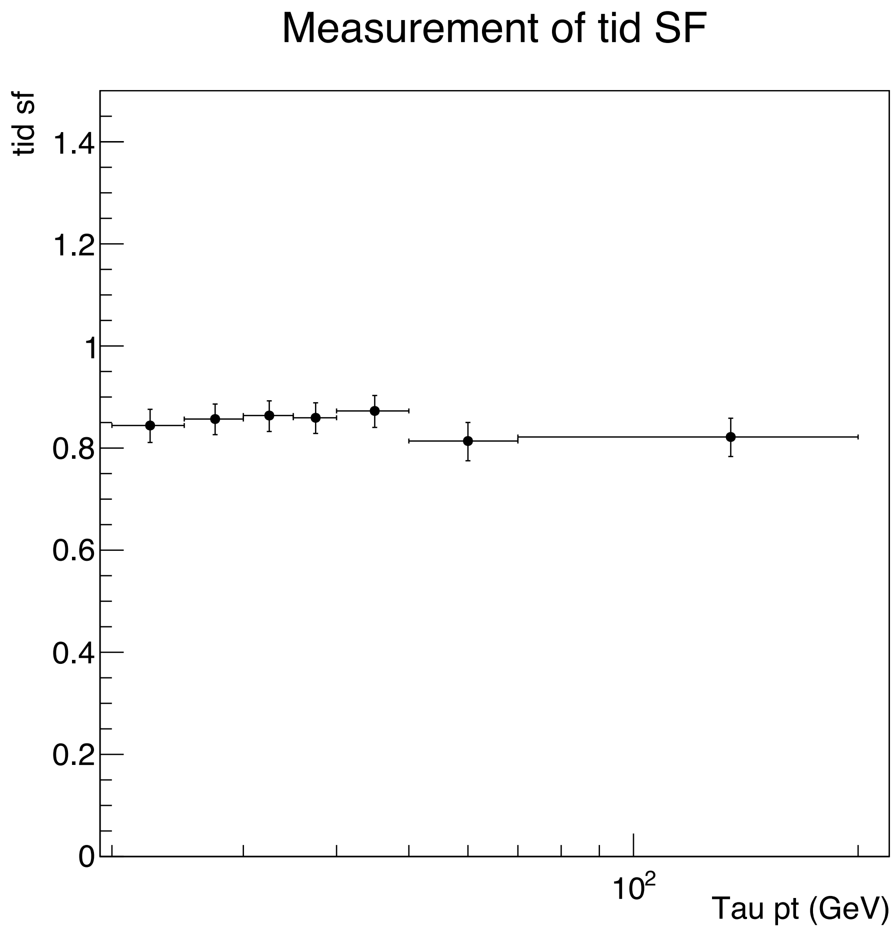
ν : vacuum expectation value
κ_τ : CP-pair coupling strength
Φ_τ : mixing angle
τ : tau lepton field
γ_5 : Dirac matrix
h :

To measure the CP properties of the Higgs boson we need to measure $\boldsymbol{\Phi_\tau}$ the angle between the planes.

Tau ID SF CMS results

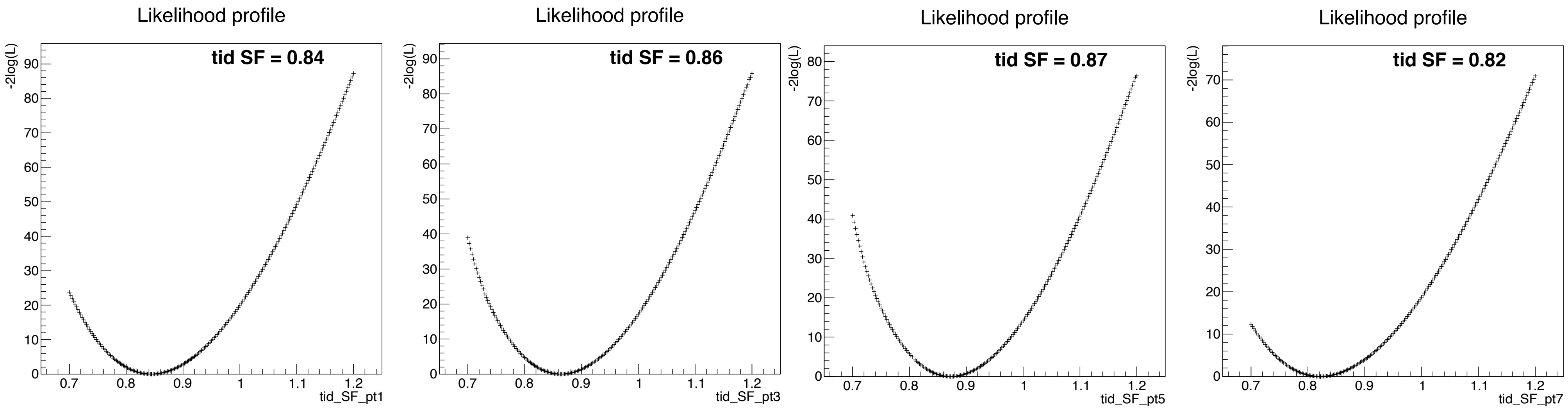


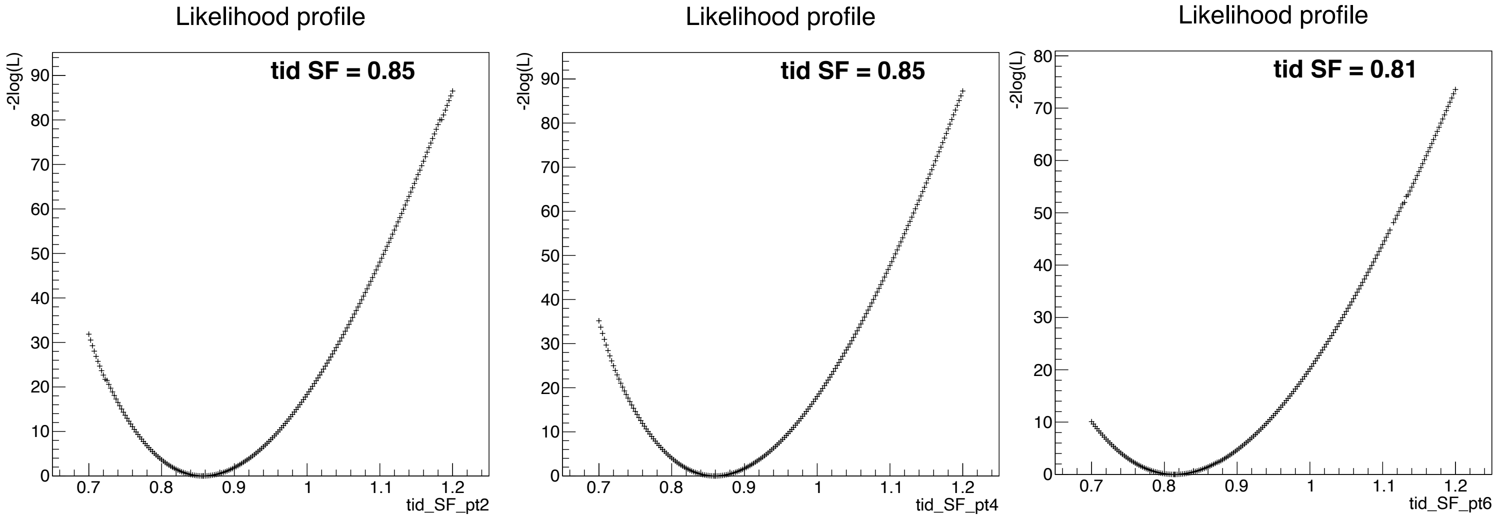
CMS results



Simultanoues fit results

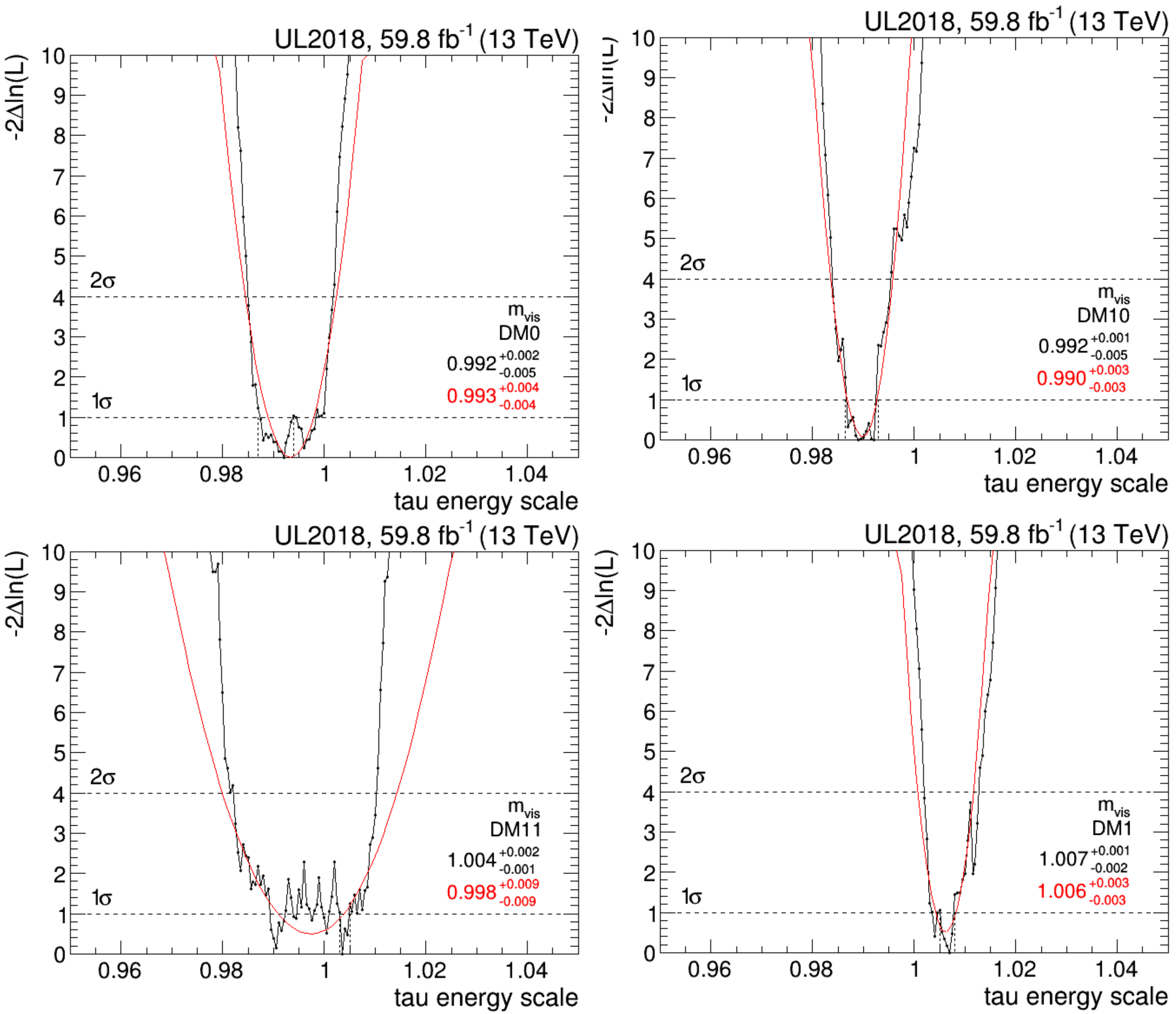
Tau ID SF log likelihood profiles





p_{τ_h}:[20,25],[25,30],[30,35],[35,40],[40,50],[50,70],[70,∞)GeV

TES log likelihood profiles



Blibliographie

- Images :
  - https://fr.wikipedia.org/wiki/Modèle_standard_de_la_physique_des_particules#/media/Fichier:Standard_Model_of_Elementary_Particles-fr.svg
  - https://fr.wikipedia.org/wiki/CMS_(expérience)#/media/Fichier:Schema_transverse_cms.png
  - https://structurae.net/fr/ouvrages/large-hadron-collider-lhc
  - https://cms.cern/detector
  - https://indico.cern.ch/event/981823/contributions/4295585/attachments/2250510/3819690/poster_434.pdf
  - https://en.wikipedia.org/wiki/Tau_(particle)
  - https://lhc-commissioning.web.cern.ch/schedule/LHC-long-term.htm