Matrix Element Method for $H\to\tau\tau$

F. Beaudette, L.Cadamuro, O. Davignon,
G. Grasseau, L. Mastrolorenzo, G. Ortona,
P. Paganini, R.Salerno, T. Strebler

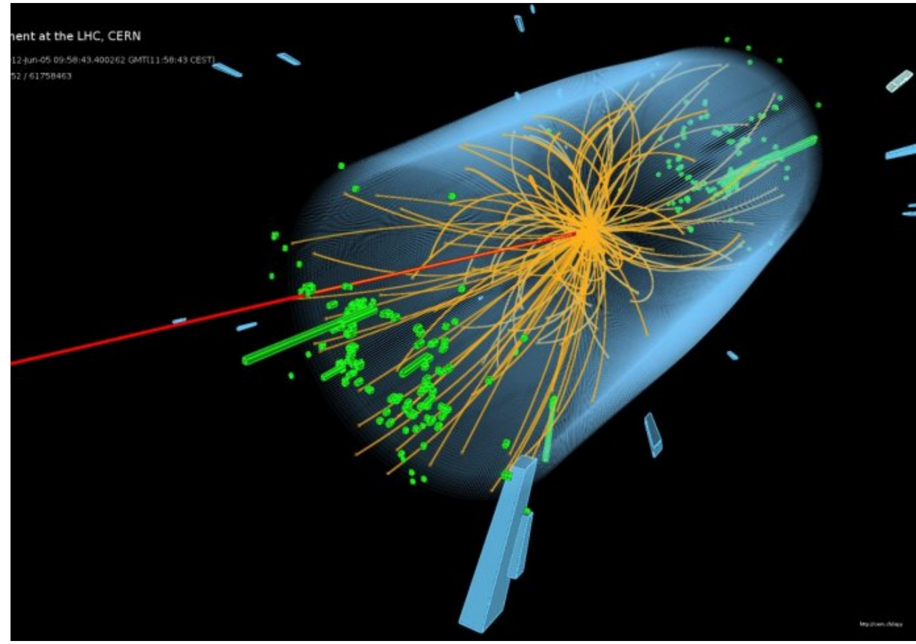
Laboratoire Leprince-Ringuet (LLR) - Ecole Polytechnique - CNRS-IN2P3

Preliminary results

MEM workshop UCL-CP3 - 01/12/2015

Introduction

- ATLAS-CMS Higgs couplings measurement combination recently brought evidence of H->$\tau\tau$ decay with a 5.5σ significance (exp. 5.0σ)



- Yet no independent discovery by either of the two experiments

Channel	References for individual publications		Signal strength [μ] from results in this paper (Section 5.2)		Signal significance [σ]	
	ATLAS	CMS	ATLAS	CMS	ATLAS	CMS
$H \to \tau\tau$	[58]	[59]	$1.41^{+0.40}_{-0.35}$	$0.89^{+0.31}_{-0.28}$	4.4	3.4
			$\left(^{+0.37}_{-0.33}\right)$	$\left(^{+0.31}_{-0.29}\right)$	(3.3)	(3.7)

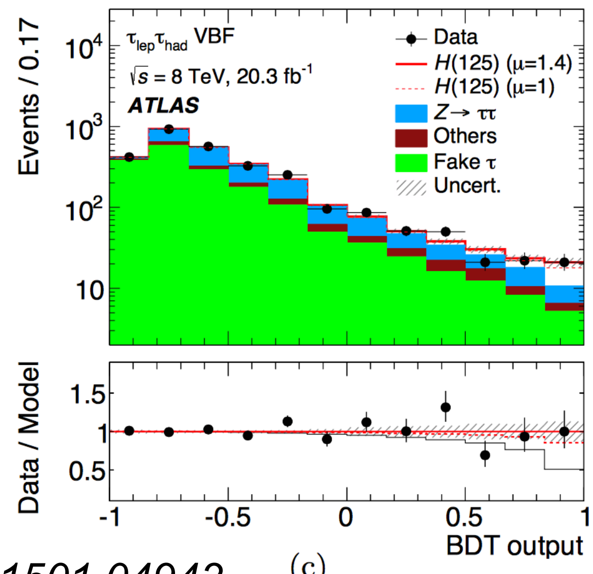
CMS-PAS-HIG-15-002

[58] arxiv:1501.04943
[59] arxiv:1401.5041

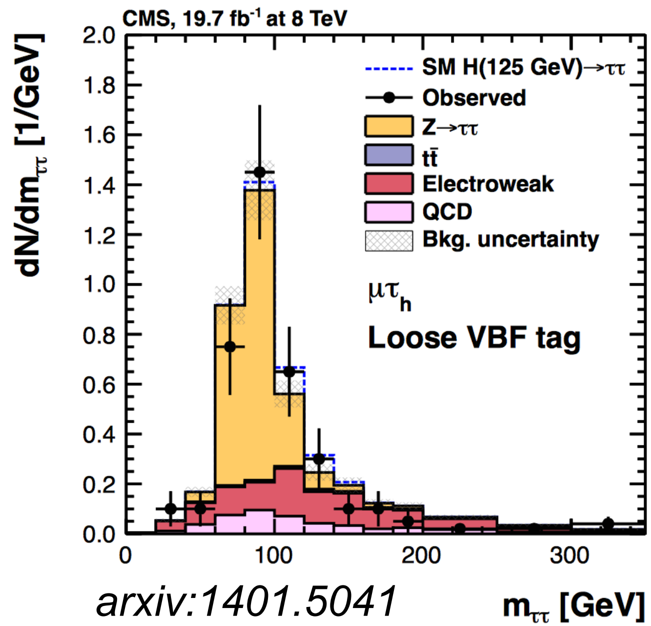
Introduction

- Di-τ mass reconstruction based on a likelihood approach in ATLAS (MMC) and CMS (SVFitMass)

- Signal extraction in ATLAS based on Boosted Decision Tree (BDT) while CMS analysis uses directly di-τ mass estimation



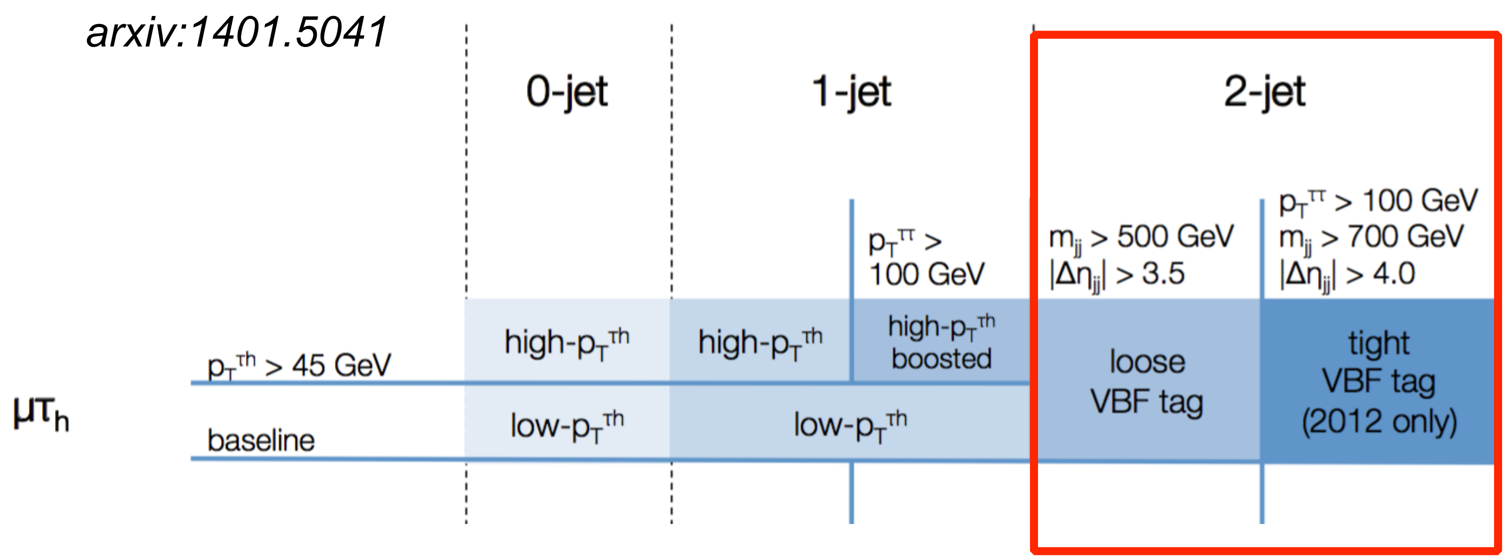
arxiv:1501.04943



arxiv:1401.5041

- New analysis method for H->$\tau\tau$, based on Matrix Element Method (MEM) will be presented here

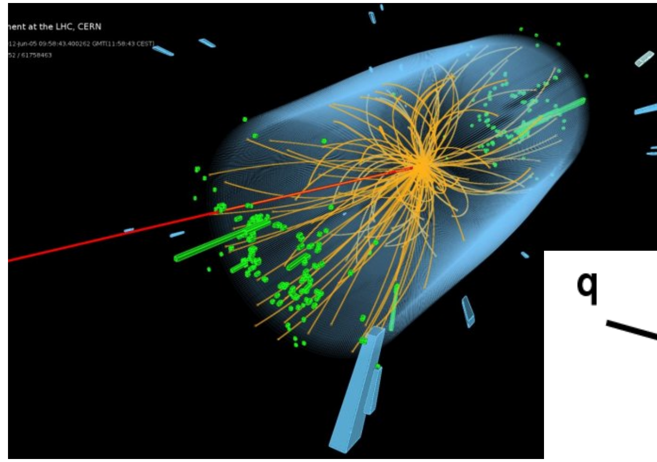
Introduction



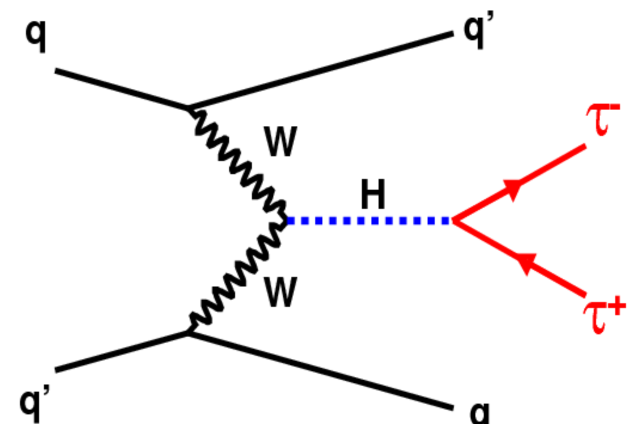
- 0-jet and 1-jet categories dominated by ggF production mode: di-τ mass estimator contains already most of the kinematic
- 2-jet categories dominated by VBF production mode: jet kinematics very powerful to increase discrimination => good target for MEM

MEM tools for H->$\tau\tau$ searches

MEM tools for H->$\tau\tau$ searches



- Observables **y**:
  - leptons
  - τ_h
  - jets
  - MET



- Phase-space point **x**:
  - τ (before decay)
  - quarks
  - neutrinos

- The final discriminant used is in principle

$$\mathcal{L}(\mathbf{y}) = \frac{w_S(\mathbf{y})}{w_S(\mathbf{y}) + w_B(\mathbf{y})}$$

with

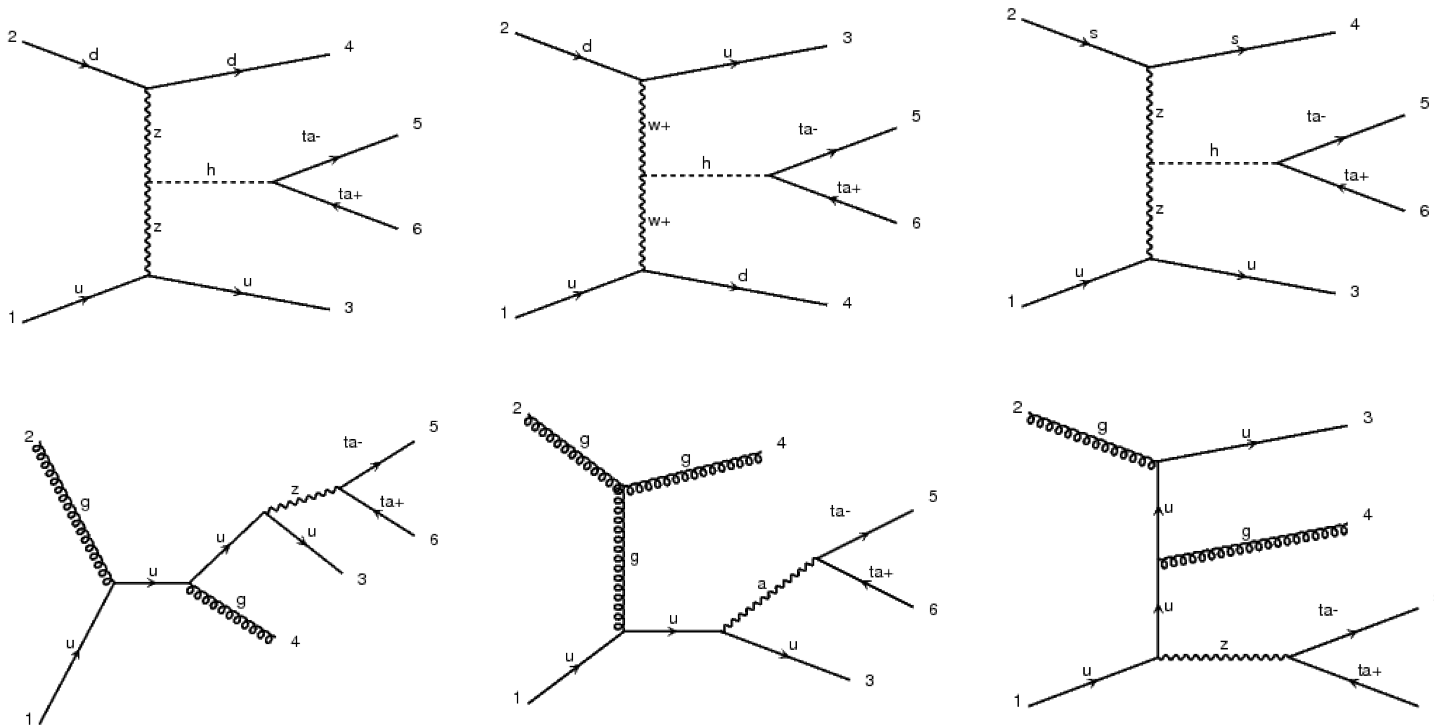
$$w_i(\mathbf{y}) = \frac{1}{\sigma_i} \sum_p \int d\mathbf{x} dx_a dx_b \frac{f(x_a, Q) f(x_b, Q)}{x_a x_b s} \delta^2(x_a P_a + x_b P_b - \sum p_k) |\mathcal{M}_i(\mathbf{x})|^2 W(\mathbf{y}||\mathbf{x})$$

MEM tools for H->$\tau\tau$ searches

- **Matrix Element computed at LO with MadGraph**

- Signal: VBF H->$\tau\tau$ (qq->qqH)

Irreducible background: DY+2jets (gg->qq̄Z, qg->qgZ, qq̄->ggZ, qq̄->qqZ)

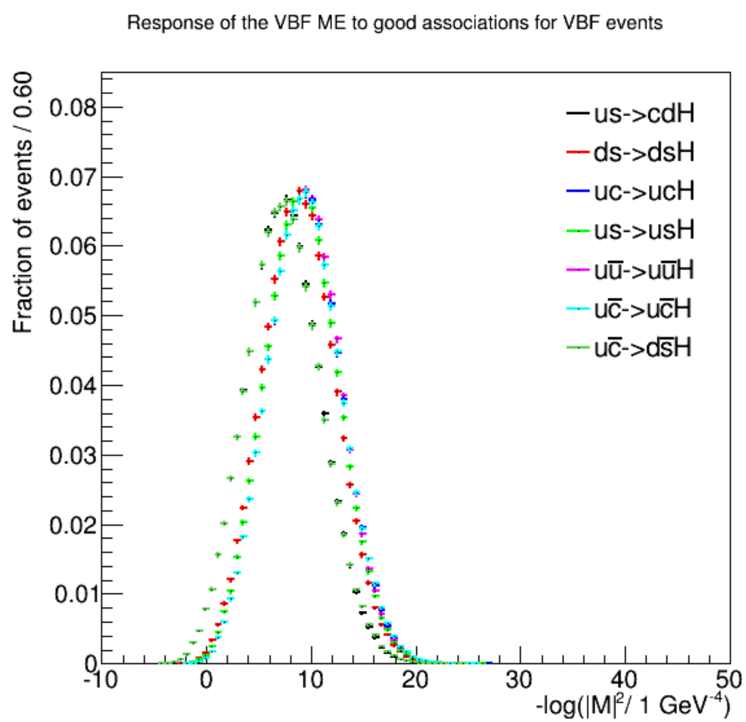


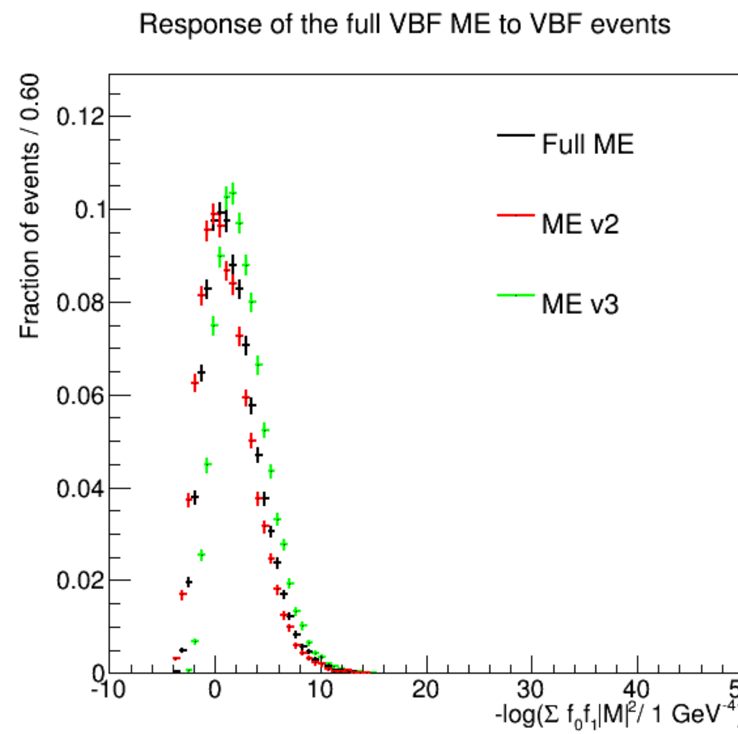
- A lot of contributing processes (48 VBF, 64 DY + permutations): can we discard some of them?

MEM tools for H->$\tau\tau$ searches

- **Matrix Element computed at LO with MadGraph**

- Compare response of different MEs with gen-level 4-momenta and put together MEs with similar responses





- ME v2 shown to have similar performance as full ME and reduces ME computation time by around a factor 10

MEM tools for H->$\tau\tau$ searches

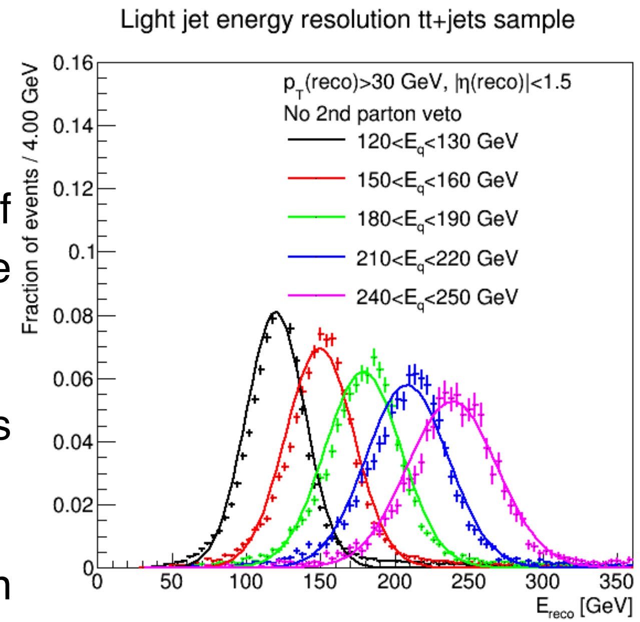
- **Transfer functions**

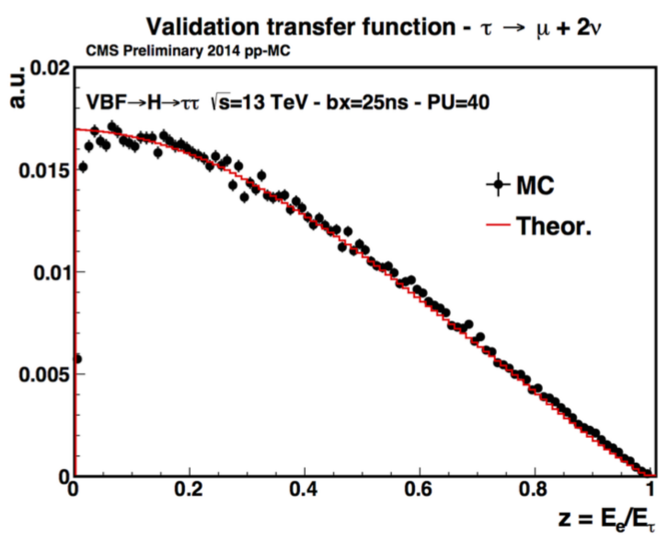
- Gaussians for energy of the jets

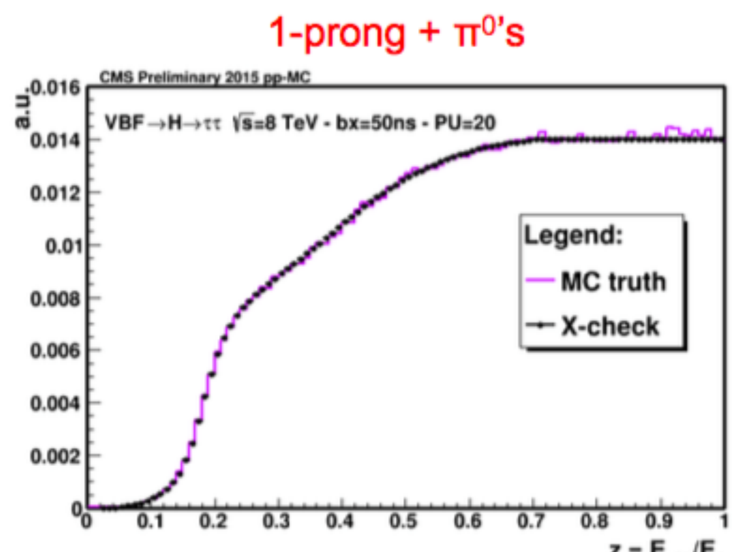
- For leptonic τ decay, analytic computation of differential decay width for the energy of the final state lepton

- For hadronic τ decay, analytic computation as well, assuming 2-body decay

- Direction of jets + 4-momentum of muon assumed to be perfectly reconstructed







MEM tools for H->$\tau\tau$ searches

- **Transfer functions**

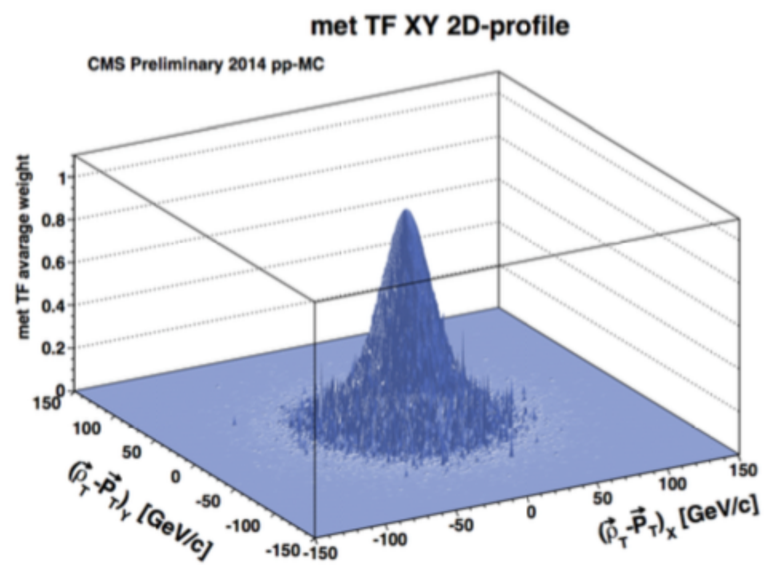
- For recoil + MET, standard MEM procedure used:

$$\vec{P}_T = \sum_{taus} \vec{P}_T + \sum_{quarks} \vec{P}_T$$ "True" recoil (dependent on integration point)

$$\vec{\rho}_T = \vec{E}_T^{miss} + \vec{p}_{T_{\tau-vis}} + \vec{p}_{T_\mu} + \sum_{recojets} \vec{p}_{T_{jet}}$$ Measured recoil

- Use 2D Gaussian to constrain recoil: MET covariance matrix estimated in situ

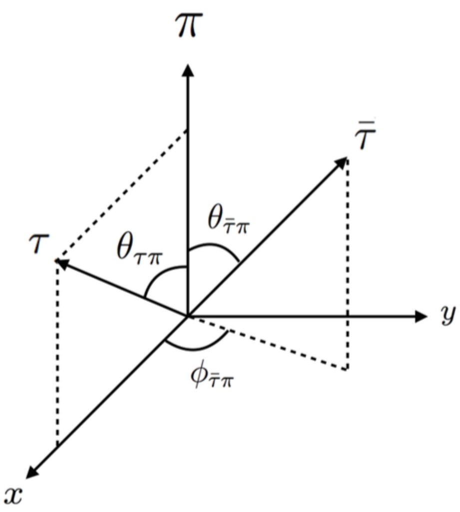
$$\mathcal{R}\left(\vec{\rho_T}|\vec{P_T}\right) = \frac{1}{2\pi\sqrt{|\mathbf{V}_{cov}|}} exp\left[-\frac{1}{2}\left(\vec{\rho}_T - \vec{P}_T\right)^T \mathbf{V}_{cov}^{-1}\left(\vec{\rho}_T - \vec{P}_T\right)\right]$$



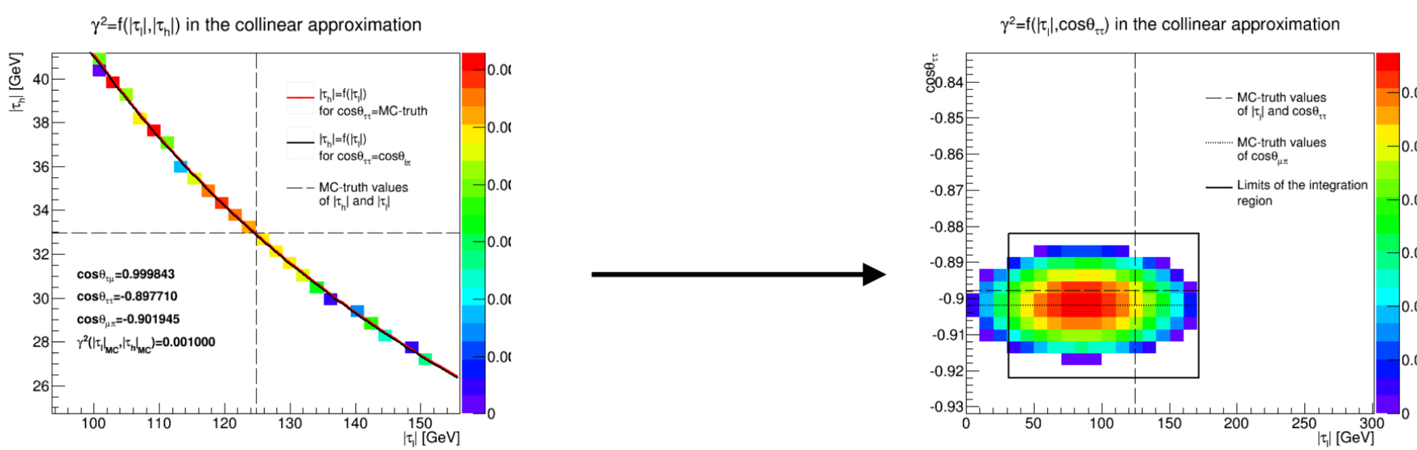
- Boost applied to 4-vectors for ME computation

MEM tools for H->$\tau\tau$ searches

- **Monte-Carlo integration (VEGAS)**
  - Because of mass-shell constraints for τ + narrow width of the Higgs, change of variables required for the integration
    => complex reconstruction of the di-τ system from those variables + measured observables



  - Non-trivial kinematic constraints taken into account to optimize integration



$d|\vec{\tau}_l|d|\vec{\tau}_h|dm^2_{\tau\bar{\tau}}$

$d|\vec{\tau}_l|d\cos\theta_{\tau\bar{\tau}}dm^2_{\tau\bar{\tau}}$

MEM tools for H->$\tau\tau$ searches

- **Monte-Carlo integration (VEGAS)**

  - 7 particles in the final state for VBF H->$\tau\tau$->$\mu\tau_h$: 21 dimensions with brute force MEM

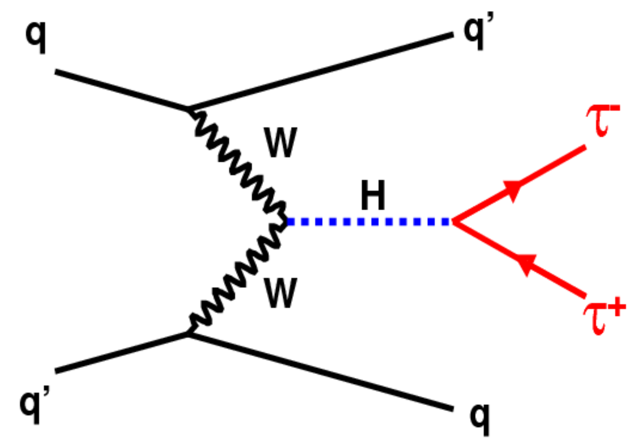
  - After dimensionality reduction, signal weight = 4 dimensions / background weight = 5 dimensions

Variable	***Integration boundaries***
2 * E(quark)	***95% CI of jet TF***
$d\lvert\vec{\tau}_l\rvert d\cos\theta_{\tau\bar{\tau}}$	***Kinematic constraints from τ decay + numeric evaluation of maximal phase space allowed (see prev. slide)***
$dm^2_{\tau\bar{\tau}}$	***For signal fixed at (125 GeV)^2 For background, integrated between m_{vis}^2 and (180 GeV)^2***

Application to VBF H->$\tau\tau$

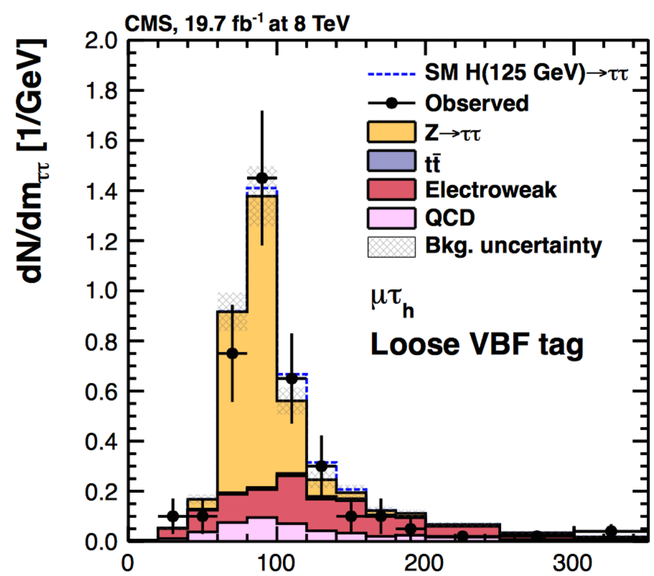
Application to VBF H->$\tau\tau$

- Validation of the MEM with H->$\tau\tau$ first done for VBF production mode in Run 1 data since CMS analysis already existing and can be used as benchmark



- Standard analysis defines VBF categories based on kinematic of two forward jets and uses SVFitMass distribution for signal extraction

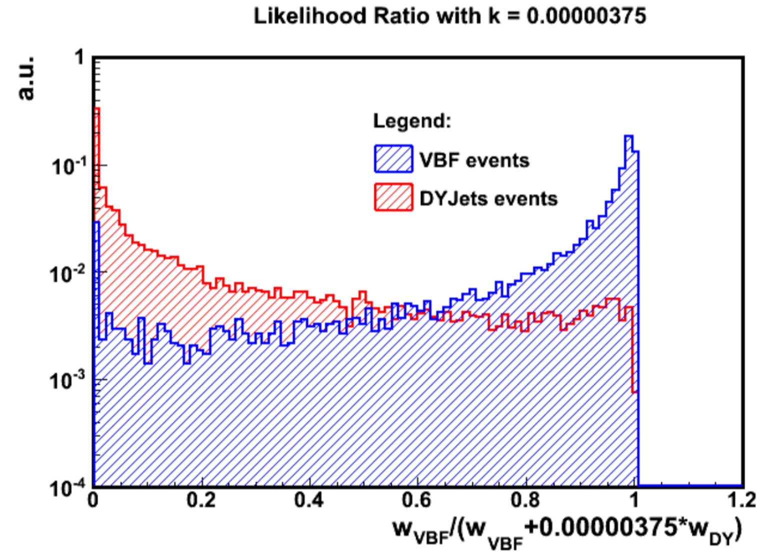
- Main background from DY+2 jets production: for now, only background considered in MEM integration (ongoing studies to include W+jets background)

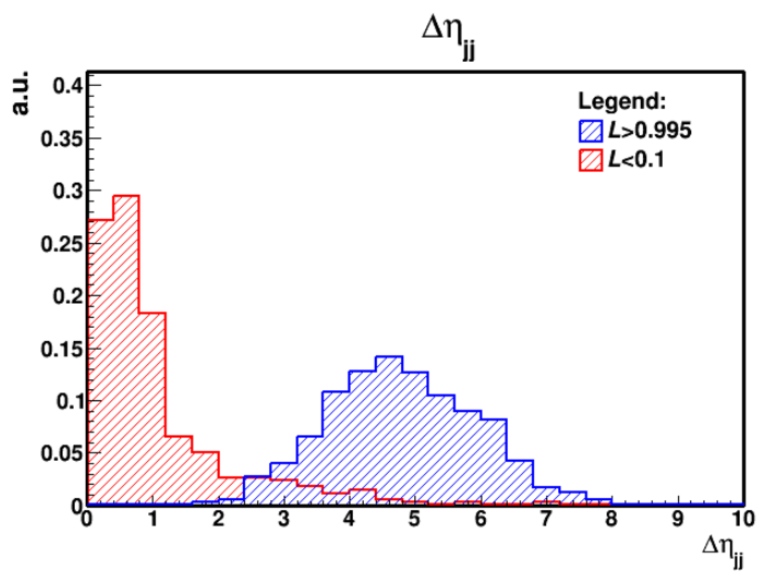


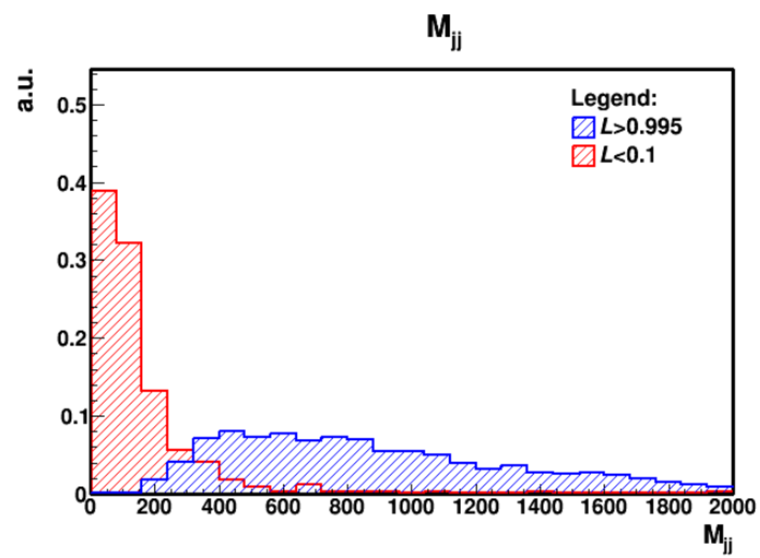
arxiv:1401.5041

Application to VBF H->$\tau\tau$

- Since kinematics of the process well-known, possible to cross-check the correlations between the MEM likelihood ratio and the relevant observables
- As expected, signal-like events show large Δη(jj) and m(jj)



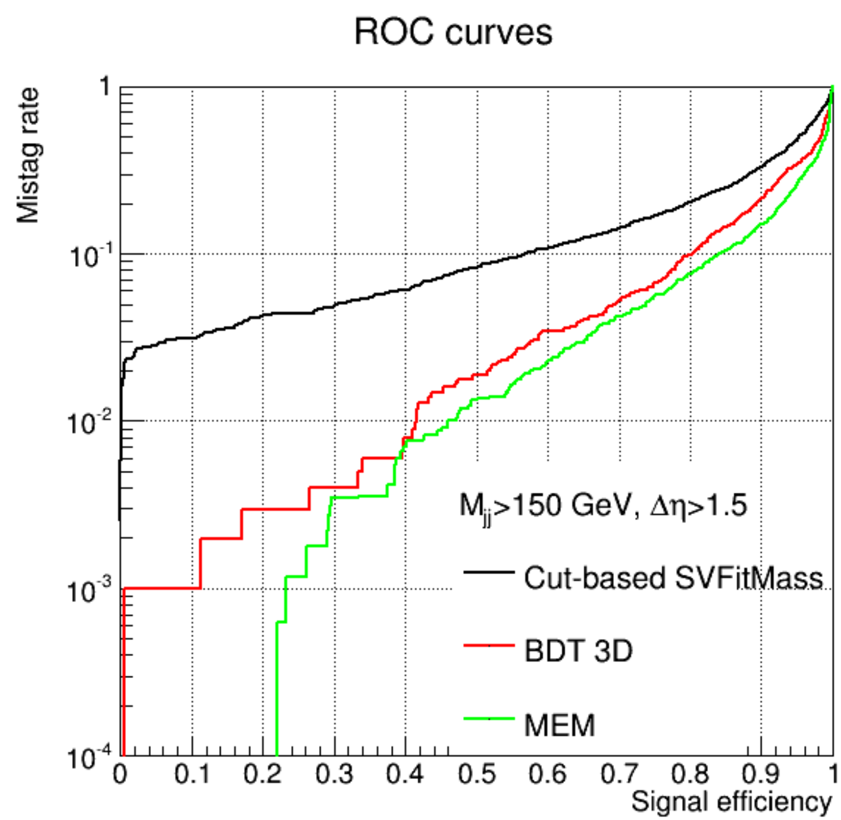




VBF sample

Application to VBF H->$\tau\tau$

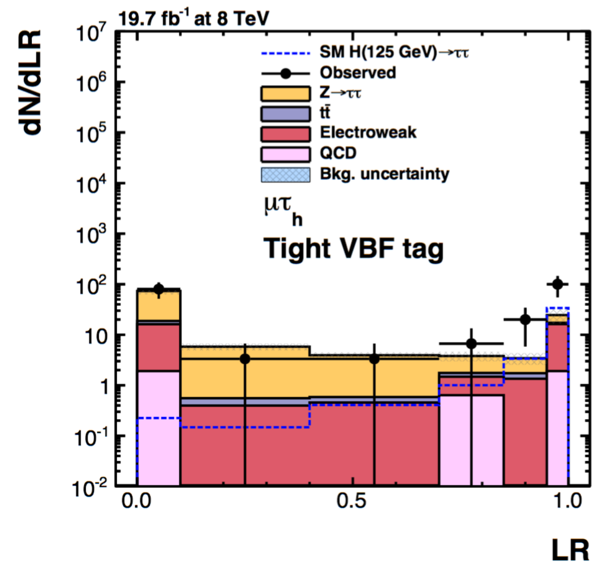
- Performance of the MEM compared to SVFitMass alone and BDT using SVFitMass + Mjj + Δη
  => with current MEM settings, significant improvement with respect to default analysis

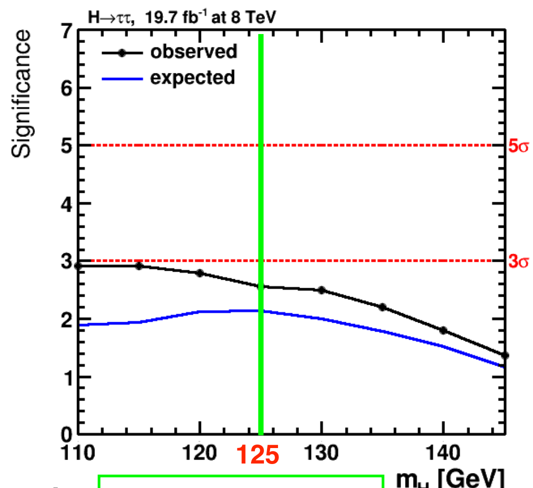


Application to VBF H->$\tau\tau$

Results from L. Mastrolorenzo's PhD thesis

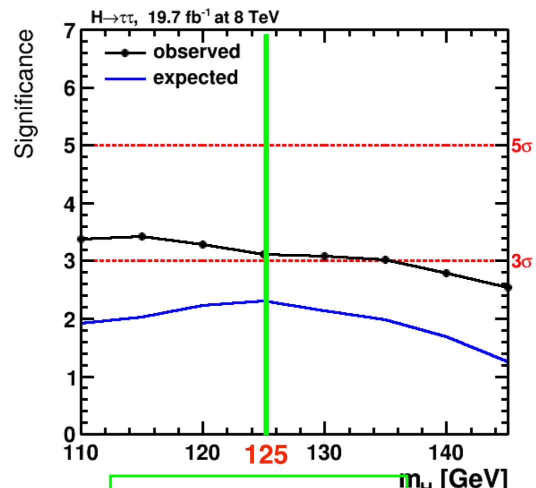
- Full CMS H->$\tau\tau$ analysis set-up reproduced including MEM (only in VBF categories)
- Evaluation of the final limits already performed in $\mu\tau$ channel
- If improvement confirmed in other channels, can use the MEM for Run 2 analysis and contribute to the discovery of H->$\tau\tau$ independently by CMS





Analysis w/o MEM

Expected: 2.14
Observed: 2.56



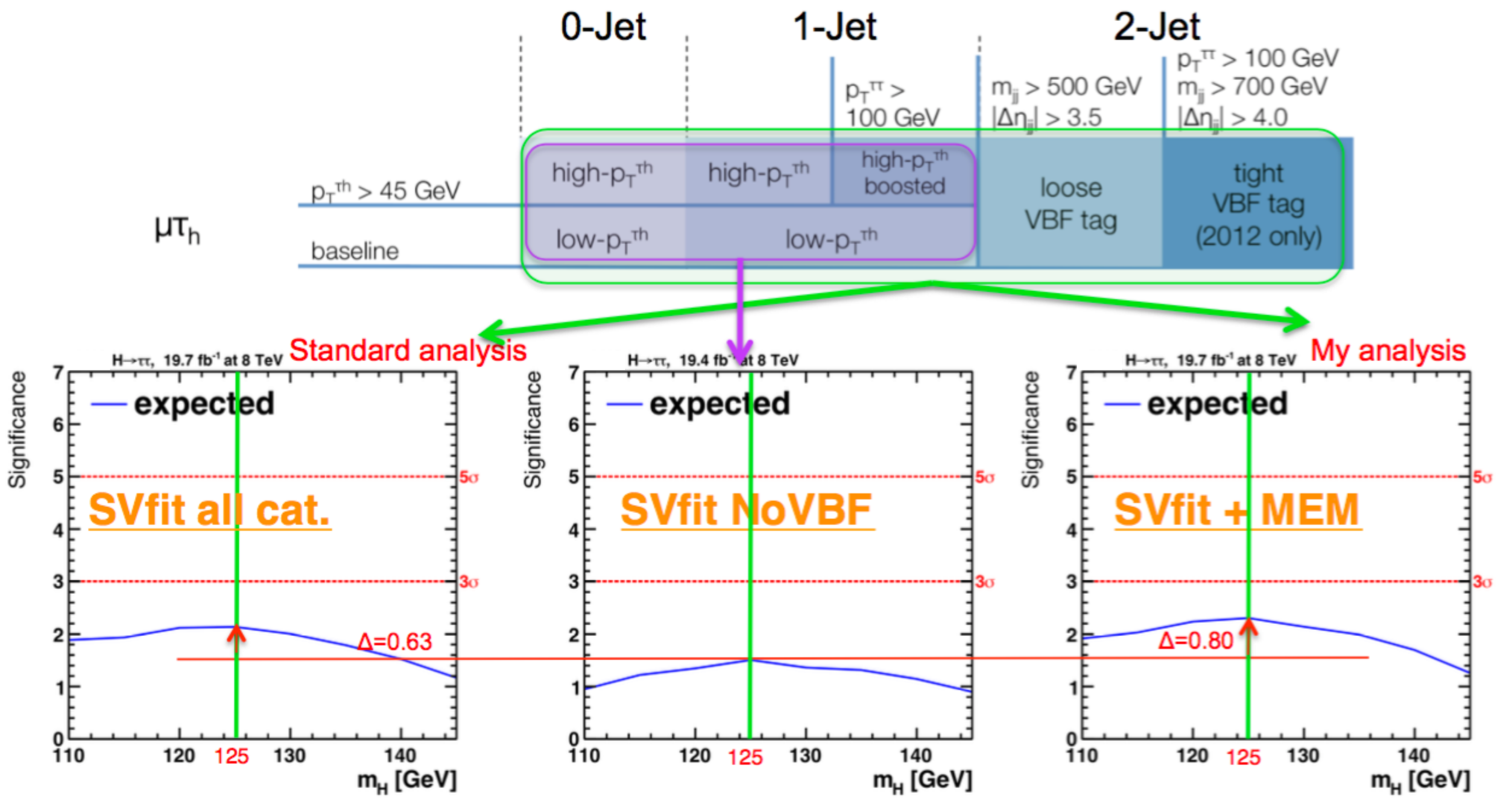
Expected: 2.31
Observed: 3.12

Analysis with MEM in VBF categories (no MEM in ggF-dominated categories)

Application to VBF H->$\tau\tau$

Results from
L. Mastrolorenzo's PhD thesis

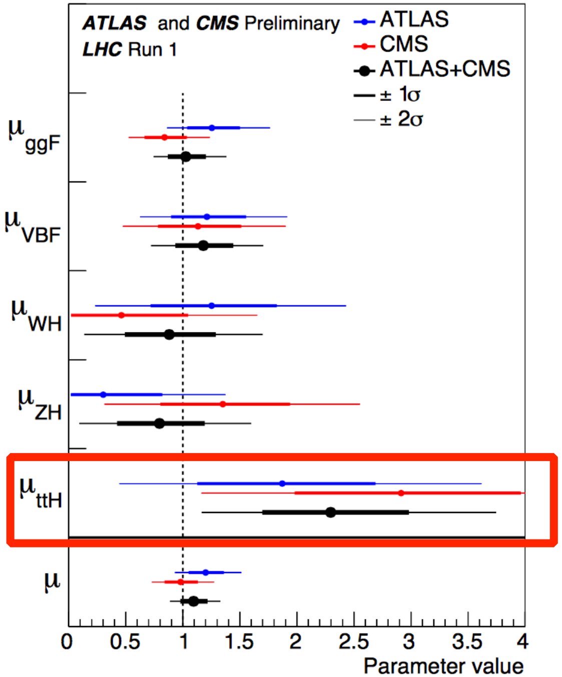
Improvement in the VBF only categories



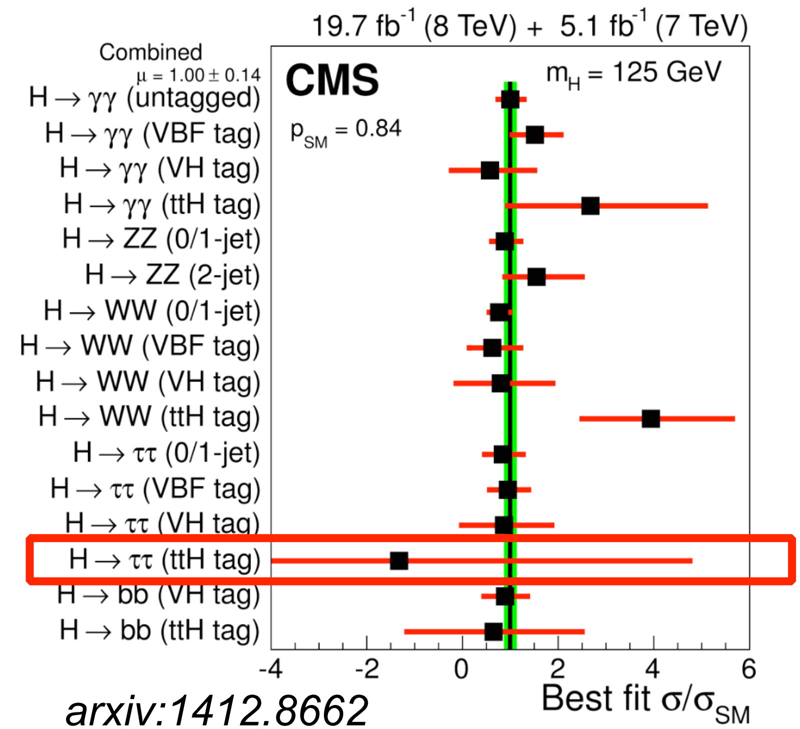
Improvement in the VBF only categories ~27%!!

Prospects for ttH H->$\tau\tau$

Prospects for ttH H->$\tau\tau$



CMS-PAS-15-002



arxiv:1412.8662

- ATLAS-CMS combination confirmed observed excess in ttH production mode
- Interesting channel to look at for Run 2: direct measurement of top-Higgs coupling
- Among all the ttH channels looked for by CMS at Run 1, H->$\tau\tau$ has the less precise measurement: large margin for improvement with MEM

Prospects for ttH H->$\tau\tau$

- Presence of τ_h in final-state can be used to decouple from ttH multi-lepton analysis

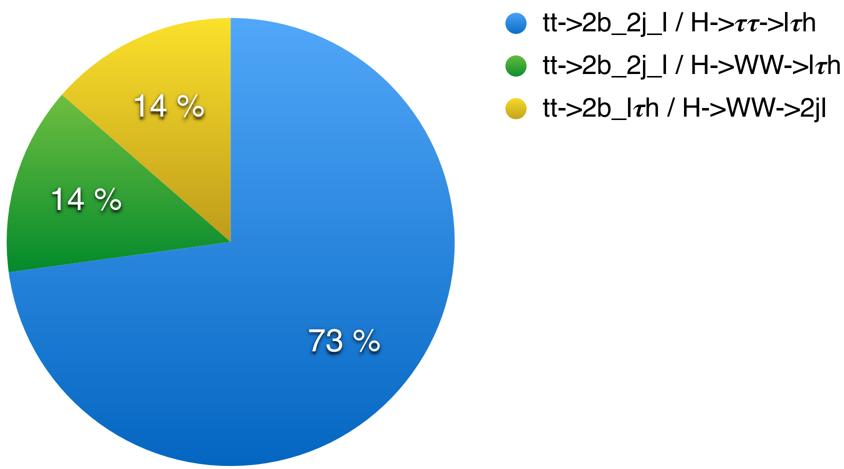
- Multiple leptons in addition can still be used to reduce the background (possibility to require 2l OS / 2l SS / 3l)

- Preselections: ≥ 2 leptons, ≥ 1 τ_h, ≥ 3 jets

- **2 lep SS + 1τ_h + jets:**

7.9% ttH, H->$\tau\tau$

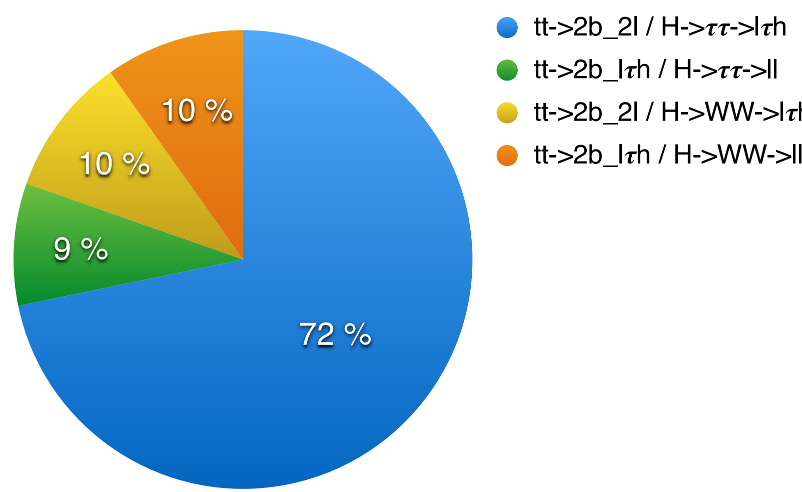
ttH repartition



- **3 lep + 1τ_h + jets:**

2.8% ttH, H->$\tau\tau$

ttH repartition

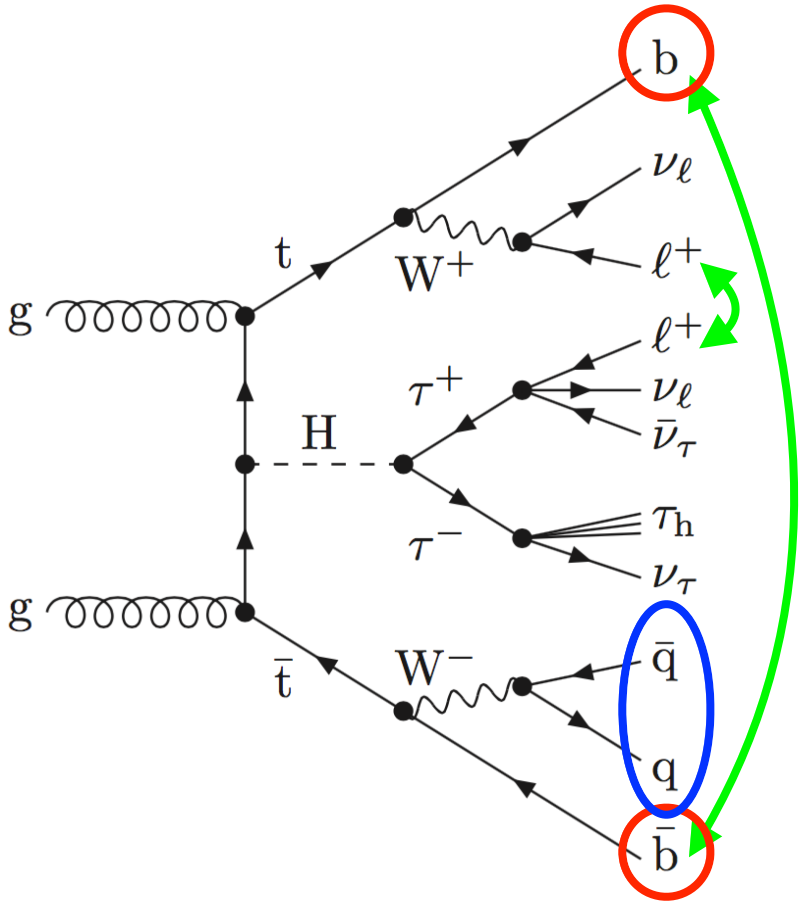


Prospects for ttH H->$\tau\tau$

- Although final-state more complex than VBF, possibility to rely on b-tagging to identify b-jets + **sum over ambiguous permutations** with MEM

- Categorization based on **b-tagging requirements** (2 CSV-medium tagged / 1 CSV-medium + 1 CSV-loose)

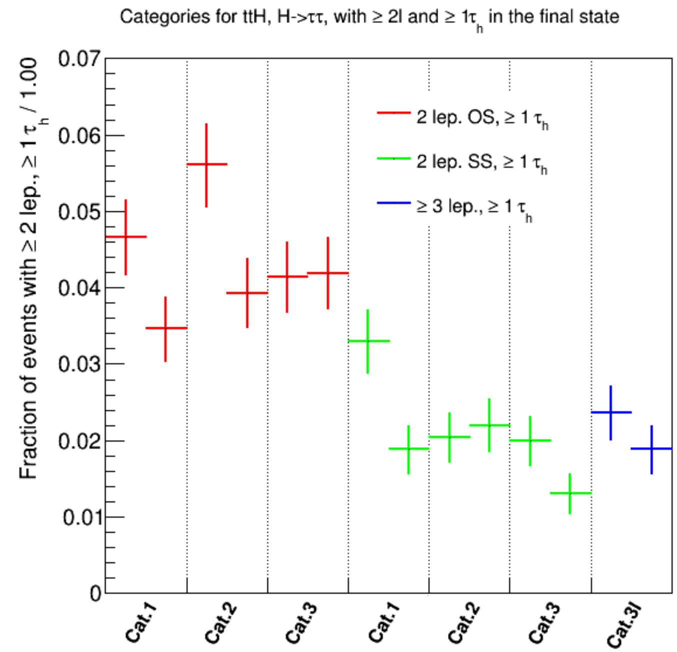
- Further categorization based on compatibility of untagged jets with W mass (M(jj) = M_W ± 20 GeV)
  => if no W-tagged pair of jets, **possibility to integrate over the missing jet direction**



Prospects for ttH H->$\tau\tau$

- Cat. 1: ≥ 2 leptons, ≥ 1 τ_h, ≥ 4 jets, W-tagged pair of light jets
  Cat. 2: ≥ 2 leptons, ≥ 1 τ_h, ≥ 4 jets, no W-tagged pair of light jets
  Cat. 3: ≥ 2 leptons, ≥ 1 τ_h, 3 jets

- Use of narrow-width approximation for top, W and Higgs to keep the number of dimensions for integration as low as possible (VBF H->$\tau\tau$ 4 dim.)



In each category, 2 CSVM / 1 CSVM + 1 CSVL distinction

2 lep. Cat. 1	*2 lep. Cat. 2*	*2 lep. Cat. 3*	*3 lep.*
5 dim.	***7 dim.***	***7 dim.***	***6 dim.***

- Performance on various backgrounds under study (tt+jets, ttW, ttZ)

Goals for 2016

- **GPU implementation**
  Should reduce computing time (~1 week for full VBF analysis with MPI implementation, 250k h of computation time over the last 2 months)
  See G. Grasseau's presentation

- **VBF H->$\tau\tau$**
  Implementation for Run 2 analysis ongoing
  Ongoing studies to include W+jets background in MEM computation
  Possible spin/CP measurements

- **ttH H->$\tau\tau$**
  Method already implemented
  Performance evaluation ongoing (requires dedicated background evaluation as low statistics available)

Conclusion

- MEM has been successfully implemented for H->$\tau\tau$ analyses: first implementation of MEM for di-τ resonance search
- Shows very promising results for VBF H->$\tau\tau$ analysis and could lead to the 5σ discovery of H->$\tau\tau$
- Could also provide good sensitivity to deal with complex final state in the ttH H->$\tau\tau$ analysis

Back-up

Principles of the Matrix Element Method

$$w_i(\mathbf{y}) = \frac{1}{\sigma_i}\sum_p \int d\mathbf{x}dx_a dx_b \frac{f(x_a,Q)f(x_b,Q)}{x_a x_b s}\delta^2(x_a P_a + x_b P_b - \sum p_k)|\mathcal{M}_i(\mathbf{x})|^2 W(\mathbf{y}||\mathbf{x})$$

$\frac{1}{\sigma_i}$ normalization coefficient (independent of **y**)

$\sum_p$ sum over all the potential associations between the reconstructed objects and the final-state particules + over the potential processes

$\int d\mathbf{x}dx_a dx_b$ integration over the phase space of the final-state particles **x** and over the momentum fractions x_a, x_b of the incoming partons

$f(x_a,Q)f(x_b,Q)$ PDFs of the incoming partons, evaluated with LHAPDF

$\delta^2(x_a P_a + x_b P_b - \sum p_k)$ δ-function enforcing the conservation of energy and longitudinal momentum between the incoming partons and the final-state particles

Principles of the Matrix Element Method

$$w_i(\mathbf{y}) = \frac{1}{\sigma_i}\sum_p \int d\mathbf{x} dx_a dx_b \frac{f(x_a,Q)f(x_b,Q)}{x_a x_b s}\delta^2(x_a P_a + x_b P_b - \sum p_k)|\mathcal{M}_i(\mathbf{x})|^2 W(\mathbf{y}||\mathbf{x})$$

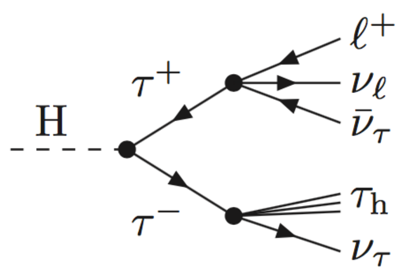
$|\mathcal{M}_i(\mathbf{x})|^2$ matrix element (ME) squared of the process *i* at LO (for instance ud->udH,H->$\tau\tau$)

$W(\mathbf{y}||\mathbf{x})$ transfer function = probability of measuring **y** given a point **x** in the phase space of the final-state particles

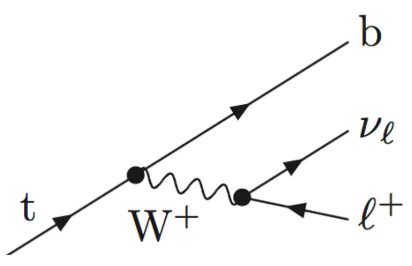
describes the decay of the unstable final-state particles of the hard-scattering (τ)

+ takes into account the resolution of the detector on the energy of the jets and on the MET

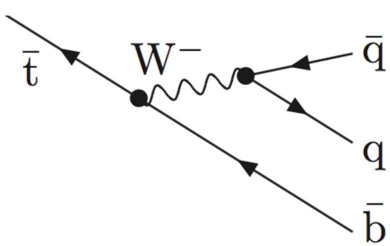
Principles of the Matrix Element Method



- **Higgs (Z) decay to $\tau\tau$**
  => 2 integrations over $d|\vec{\tau}_l|d\cos\theta_{\tau\bar{\tau}}$
  + optional integration over $dm^2_{\tau\bar{\tau}}$



- **Leptonic top decay**
  Enu determined from lepton momentum + M_W constraint
  Eb determined from W momentum + M_t constraint
  => 2 integrations over neutrino direction



- **Hadronic top decay**
  Eqbar determined from q momentum + M_W constraint
  Eb determined from W momentum + M_t constraint
  => 1 integration over Eq