Prospects for VBF $H \longrightarrow \tau\tau \longrightarrow \tau_{had} + l$ in CMS and commissioning of the Particle-Flow event reconstruction with data

Lorenzo Bianchini[1]

LLR, École Polytechnique

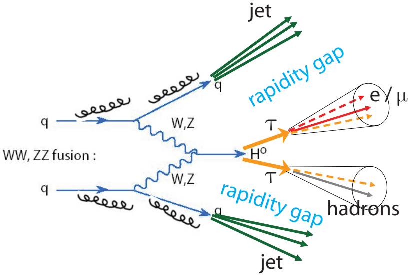
30 July 2010, Higgs Hunting 2010, Orsay

[1] on behalf of the CMS collaboration

Ref: *CMS PAS* HIG-08-006/008, PFT-10-02

The VBF $H \to \tau\tau \to l + \tau_{jet}$ channel

Second largest cross section and sizeable BR to τ's for $m_H \lesssim 145$ GeV; $l + \tau_{had}$ channel ($BR \sim 45\%$) has high trigger and reconstruction efficiency ($\epsilon\,(\tau_{jet}\mathrm{ID}) \gtrsim 45\%$).



- one $e/\mu + \tau_{jet}$ in the central region;
- 3ν from τ-decay $\Longrightarrow\ \vec{E}_T^{miss}$
- two *tagging* jets with large η separation;
- rapidity gap.

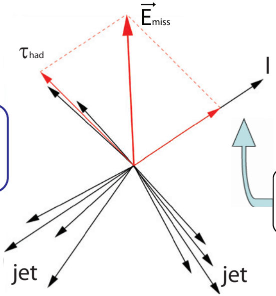
$\Longrightarrow$ the spectrum of $m_{\tau\tau}$ is sensitive to the presence of the Higgs boson.

- physics background: $Z \to \tau\tau$ + jets;
- reducible background: W + jets, QCD multi-jets, $t\bar{t}$, γ + jets;

Collinear approximation for $m_{\tau\tau}$ reconstruction

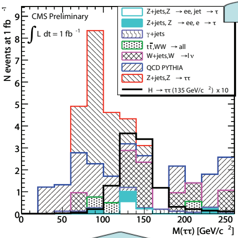
τ's are boosted => $\vec{p}_\nu \parallel \vec{p}_\tau$

(collinear approximation)



reconstruct ν's projecting $\vec{E}_{miss}$ onto the visible legs

Profile likelihood fits for 95% CL_s exclusion limit:

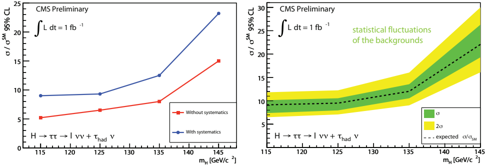


templates from data whenever possible:. Ex:

QCD => "SS" distribution

Z-> ττ => embedded-(Z->μμ)

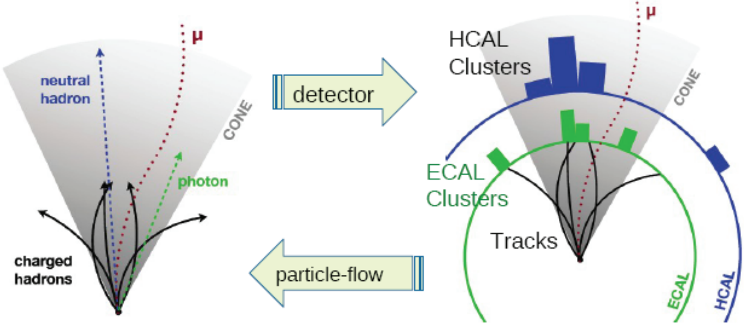
combined e+μ 95% CLs exclusion limits on σ/σ_{SM} with and without systematics



$(p_{\tau 1} + p_{\tau 2})^2$ using $\vec{E}_{miss}$ (collinear approximation)

(1.1 ± 0.1) signal events expected in 1 fb^{-1} at $\sqrt{s} = 14$ TeV

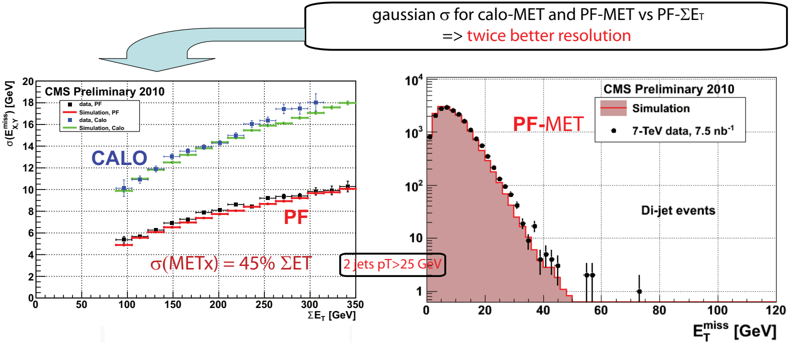
The Particle-flow approach



Output: a list of particle candidates ⇒ build higher level objects (jets, taus, MET, isolation deposits, ...)

Deployed for the first time at a hadron collider!

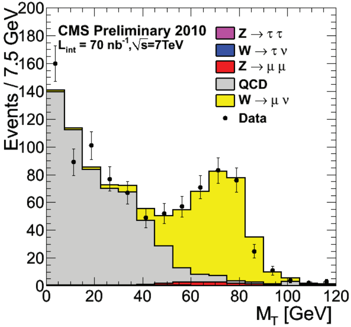
Commissioning of MET (1)



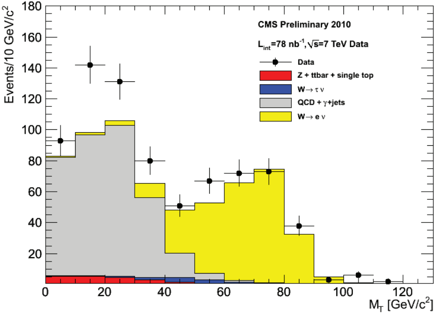
A factor 2 better resolution w.r.t calo-MET is found in di-jet as well as minimum-bias events

⇒ we can improve the H/Z peaks separation.

1 μ p_T >25 GeV, no isolation



1 ele p_T >20 GeV, isolated

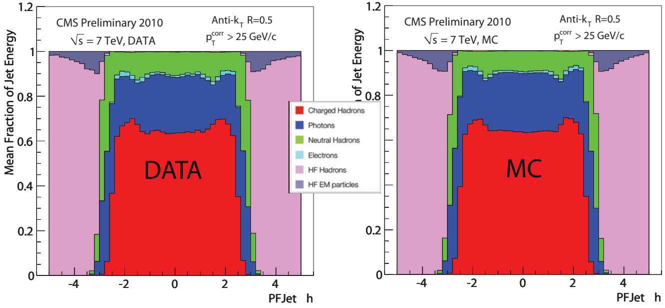


Good data/MC agreemenet for the signal
$\Rightarrow$ reliability of PF MET and lepton reconstruction.

Commissioning of τ_{had}-ID and jets

PFlow jets $\Rightarrow$ run the jet-algo over the list of PFlow candidates.

Select di-jet events pT>25 GeV/c



τ_{jet}-ID relies on the particle content of τ_{had} candidates. Fake rates well reproduced by the simulation (shown by M.Bluj)

$\Rightarrow$ good knowledge of τ_{had}-fake rate needed to handle the reducible bkg.

Conclusions

- Prospects on the exclusion limits on a SM-like Higgs boson produced in VBF and decaying to a tau-pair in the CMS experiment with 1 fb^{-1} at $\sqrt{s} = 14$ TeV have been reviewed;
- the signal detection will benefit from a global particle-flow event reconstruction, since the performances in terms of jet, MET and taus reconstruction will be significantly improved;
- all the building bricks as well as higher-level objects of the particle-flow reconstruction have been commissioned and validated with LHC data.

What next?

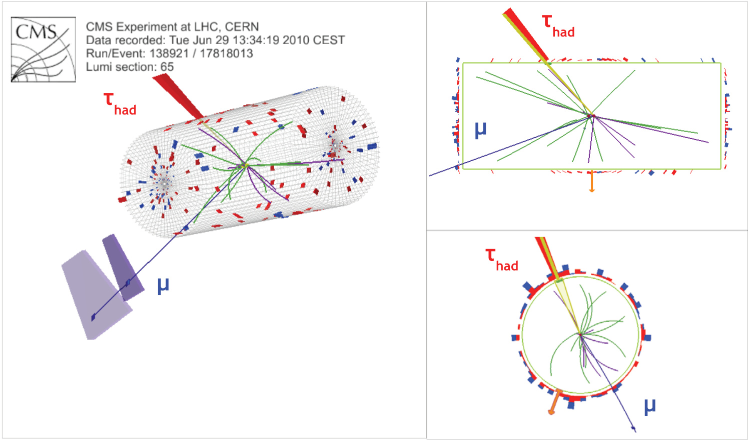
- test the *ensemble* in physics analysis ...

Conclusions

- Prospects on the exclusion limits on a SM-like Higgs boson produced in VBF and decaying to a tau-pair in the CMS experiment with 1 fb^{-1} at $\sqrt{s} = 14$ TeV have been reviewed;
- the signal detection will benefit from a global particle-flow event reconstruction, since the performances in terms of jet, MET and taus reconstruction will be significantly improved;
- all the building bricks as well as higher-level objects of the particle-flow reconstruction have been commissioned and validated with LHC data.

What next?

- test the *ensemble* in physics analysis ...
- ... and be ready, as the first $(Z \longrightarrow)\ \tau_{had} + l$ candidates start to appear!



$\mu\ p_T$ = 22.8 GeV/c

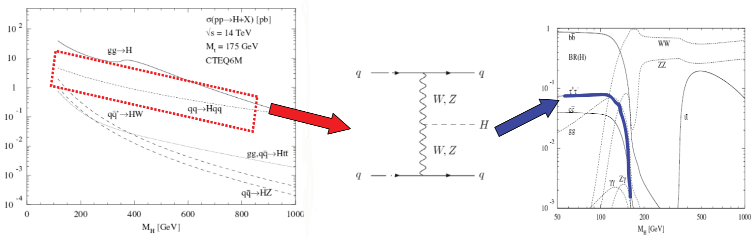
$\tau_{had}\ E_T$ = 32.9 GeV

Vis. Mass = 60.8 GeV/c^2
$M_T(\mu,MET)$ = 10.1 GeV

Back-up

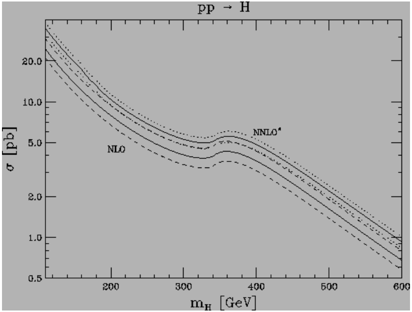
The VBF $H \to \tau\tau \to l + \tau_{jet}$ at the LHC

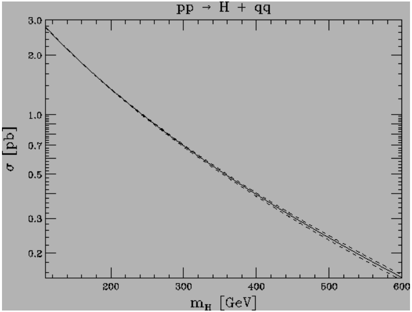
Proposed about 10 years ago as a discovery channel for $\sim 30\ \text{fb}^{-1}$ (Zeppenfeld *et al.*). Here, CMS prospects for 1 fb^{-1} at 14 TeV.



- $\sigma(qqH) = O(4\ \text{pb})$, second largest at the LHC;
- sizeable BR: $3\% < \text{BR}\,(H \to \tau\tau) < 7\%$ for $115 < m_H < 145$ GeV;
- $\text{BR}\,(\tau\tau \to l + \tau_{jet})$, $l = e, \mu, \approx 45\%$;
- $l + \tau_{jet} \Rightarrow$ single-lepton trigger or combined-trigger with τ;
- $\epsilon\,(\tau_{jet}\,\text{ID}) \gtrsim 45\% \rightleftharpoons$ QCD τ fake-rate $\lesssim 2\%$.

pp cross sections at 7 TeV





Simulated samples $\sqrt{s} = 14$ TeV

- *qqH*: four mass points (115,125,135,145 GeV/c^2) generated with PYTHIA, τ decay simulated with TAUOLA;
- QCD $2\tau + 2/3j$: generated with ALPGEN with CTEQ5L pdf, interfaced to PYTHIA for parton-shower and hadronization (MLM prescription);
- EWK $2\tau + 2/3j$: generated with MadGraph with CTEQ5L pdf, interfaced to PYTHIA for parton-shower and hadronization (MLM prescription);
- $W + jets$: generated with ALPGEN with CTEQ5L pdf, interfaced to PYTHIA for parton-shower and hadronization (MLM prescription);
- $t\bar{t}$: generated with PYTHIA, TopRex, ALPGEN, *CompHEP* and MadGraph, all leptonic decays included;

Event selection

Select a signal-enriched sample by requiring (*CMS PAS* **HIG-08-008**):

1. the VBF signature
   - two tagging jets:
     $p_T^{j1,j2} > 30$ GeV$/c$, $\eta_{j1} \times \eta_{j2} < 0$, $|\Delta\eta_{j1j2}| > 2.5$, $M_{j1j2} > 400$ GeV$/c^2$
   - central jet veto:
     veto a third jet if $\eta^{j1} - 0.5 < \eta^{j3} < \eta^{j2} + 0.5$ and $E_T^{j3} > 10$ GeV
2. the H decay products
   - one identified lepton (*e* or *mu*):
     $|\eta_l| < 2.5$, isolated
   - one identified τ_{jet}:
     $|\eta_\tau| < 2.5$, $E_T > 30$ GeV
   - di-tau reconstructed in collinear approximation

Systematic uncertainties expected with 1 fb^{-1}

- Strategies for the VBF $H \to \tau\tau \to l + \tau_{jet}$ channel with 1 fb^{-1} at $\sqrt{s} = 14$ TeV can be found in **CMS PAS HIG-08-008**.
- Data-driven modeling of the $M_{\tau\tau}$ spectrum:
  - $Z \to \tau\tau$: from real $Z \to \mu\mu$ events with τ *embedding* (and with only marginal systematics);
  - QCD: from data, using ABCD methods on OS/SS samples;
  - W and $t\bar{t}$: from their leptonic channels with τ embedding (for real taus), from MC simulation and measured tau fake-rates (otherwise);
  - τ-ID efficiency: from $Z \to \tau\tau$ events with $\sim 5\%$ systematic;
  - jet$\to \tau$ fake-rate: from Z + jets ($Z \to \mu\mu$) with 10% systematic;
  - e,μ trigger&offline efficiencies, $e \to \tau$ fake-rate: from $Z \to ee$ with $< 1\%$ systematic using tag-and-probe.
  - Jet veto for $Z \to \tau\tau$: from $Z \to \mu\mu$ events with 5% systematic.
  - JES and MET scale: from γ+jet, $Z \to \mu\mu$, with $\sim 5\%$ systematic.

Event yields expected in 1 fb^{-1} at $\sqrt{s} = 14$ TeV

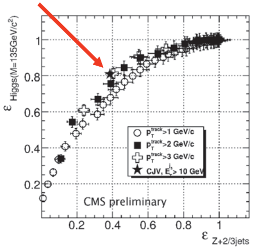
process	$qqH(135)$	$Z\tau\tau$	QCD	W	$t\bar{t}$	$\gamma + jet$	Zee
exp	1.1	13.2	13.4	11.1	3.5	2.3	$0.9 + 0.8$
tot. uncert %	12%	29%	33%	30%	25%	28%	$44\% \oplus 23\%$

From 14 TeV $\rightarrow$ 7 TeV the VBF cross section decreases by a factor ~ 3.5.
By rescaling the expected cross sections, we expect a factor ~ 2.5 increase in the σ/σ_{SM} upper bounds for a given mass.

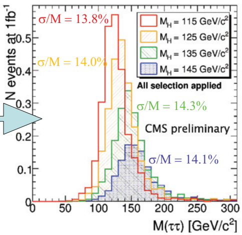
Basic startegies for event selection and background control

- VBF cuts + central jet veto,
- $m_{\tau\tau} = (p_{\tau 1} + p_{\tau 2})^2$ using $\vec{E}_T^{miss}$ (collinear approx.),
- $Z \to \tau\tau$ spectrum from an embedding of simulated τ_{jet} in real $Z \to \mu\mu$ events,
- data-driven control of QCD-multijets bkg.

CJV



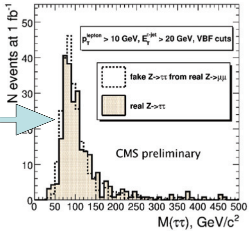
di-tau mass from H->ττ in coll. approx.



m(ττ) resolution dominated by MET resolution

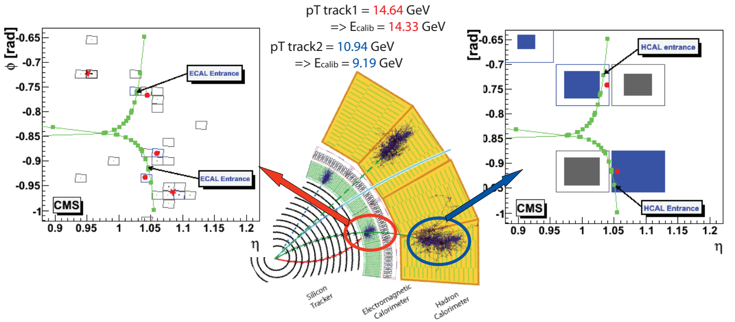
τ-embedding technique

di-tau mass from real and fake Z->ττ



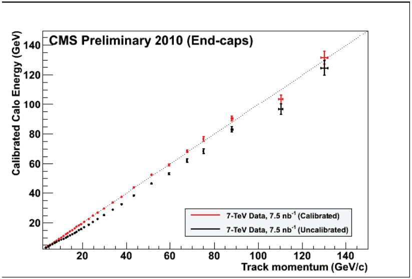
Validation of the link algorithm and calibration

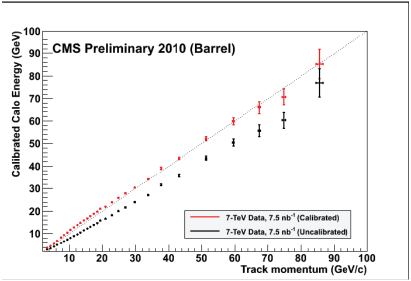
At the heart of the algorithm is the track-cluster matching.
It relies on a proper simulation of the detector ⇒ it needs to be commissioned!



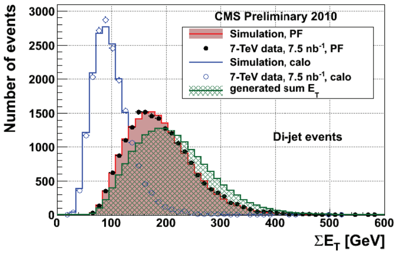
- Track-cluster link distance measured from single isolated tracks agrees well with simulation
- Calibrated calo-energy matches track momentum over two order of magnitude ($\sim$ 5% deviations at low p_T will disappear with data-driven calibartions)

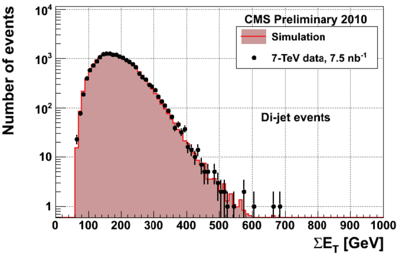
Validation of the calibration procedure with data



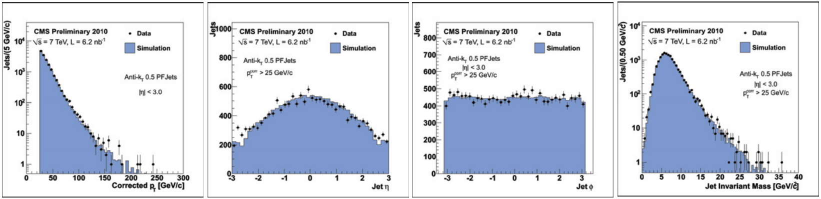


$\sum E_T$ in di-jets events





Commissioning of PF jets



Select di-jet events p_T>25 GeV/c, no Jet-ID => general agreement on jets variables