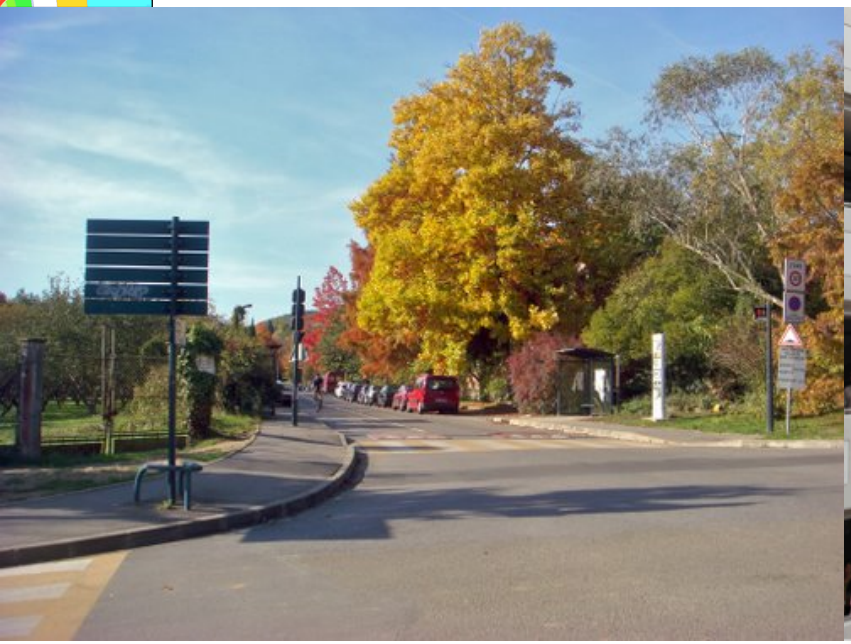
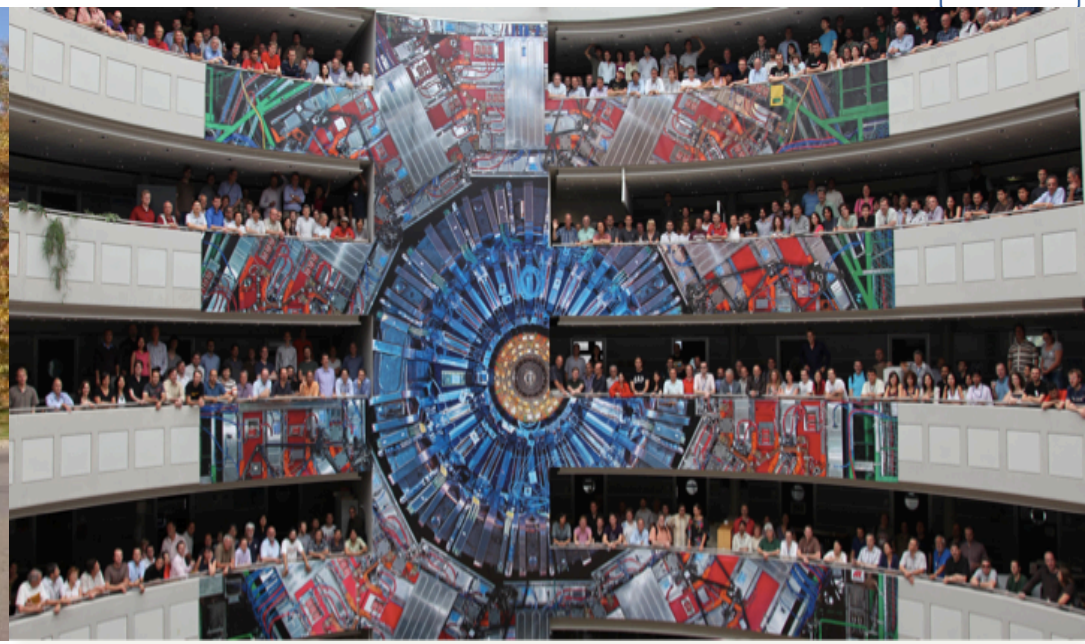




CMS Higgs search: ττ and bb final states

Michele de Gruttola

U of Florida – LPC Fermilab

on behalf of CMS collaboration

Higgs Hunting Workshop, Orsay 19/07/12

Outline

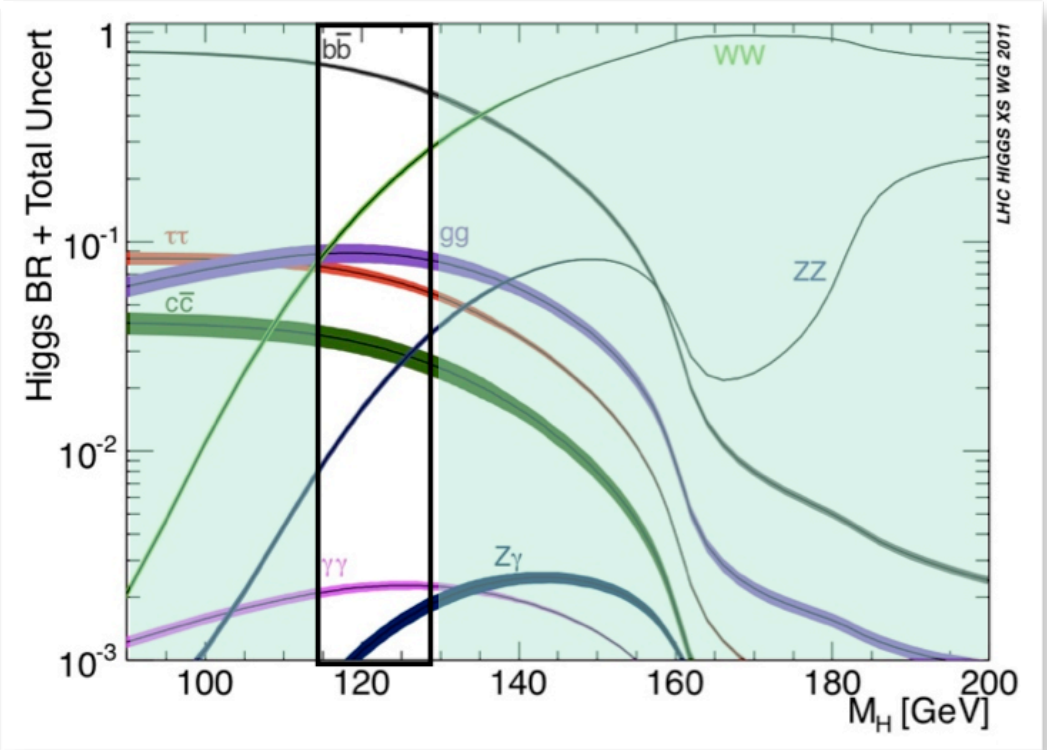
- Motivation
- Physics objects at CMS
  - b's and τ's
- VHbb, ttHbb, Hττ searches
  - searches in the context of MSSM have been presented in this workshop in A.Nayak's talk
  - VHbb presented also by P.Bortignon in this workshop

References

- PAS Hττ:
  - http://cdsweb.cern.ch/record/1460413?ln=en
- PAS VHbb
  - https://cdsweb.cern.ch/record/1460692?ln=en
- PAS ttHbb:
  - http://cdsweb.cern.ch/record/1460423?ln=en

- iCHEP talk of David Lopes Pegna
  - http://indico.cern.ch/contributionDisplay.py?contribId=437&confId=181298
- and Joshua Swanson
  - http://indico.cern.ch/contributionDisplay.py?contribId=436&confId=181298

Region of interest ;-)



@125:

BR(H→bb) ~ 58%

BR(H→WW) ~ 22%

BR(H→ττ) ~ 6%

BR(H→ZZ*) ~ 3%

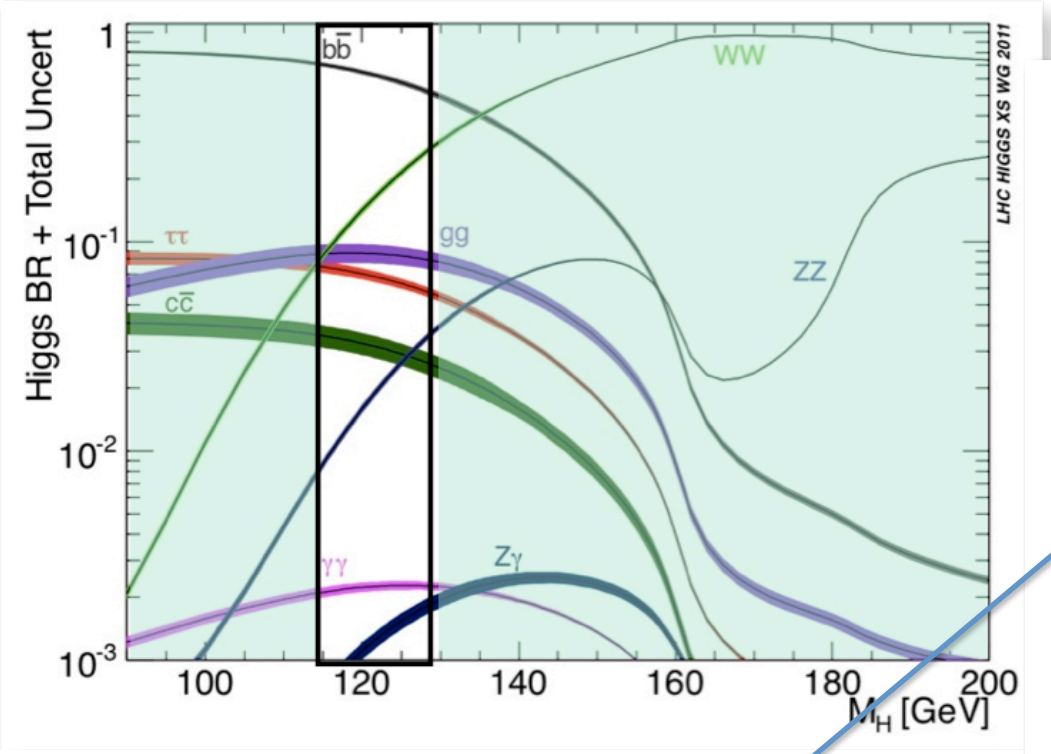
BR(H→γγ) ~ 0.22%

b's and tau's ← test the Higgs couplings to fermions vs. bosons

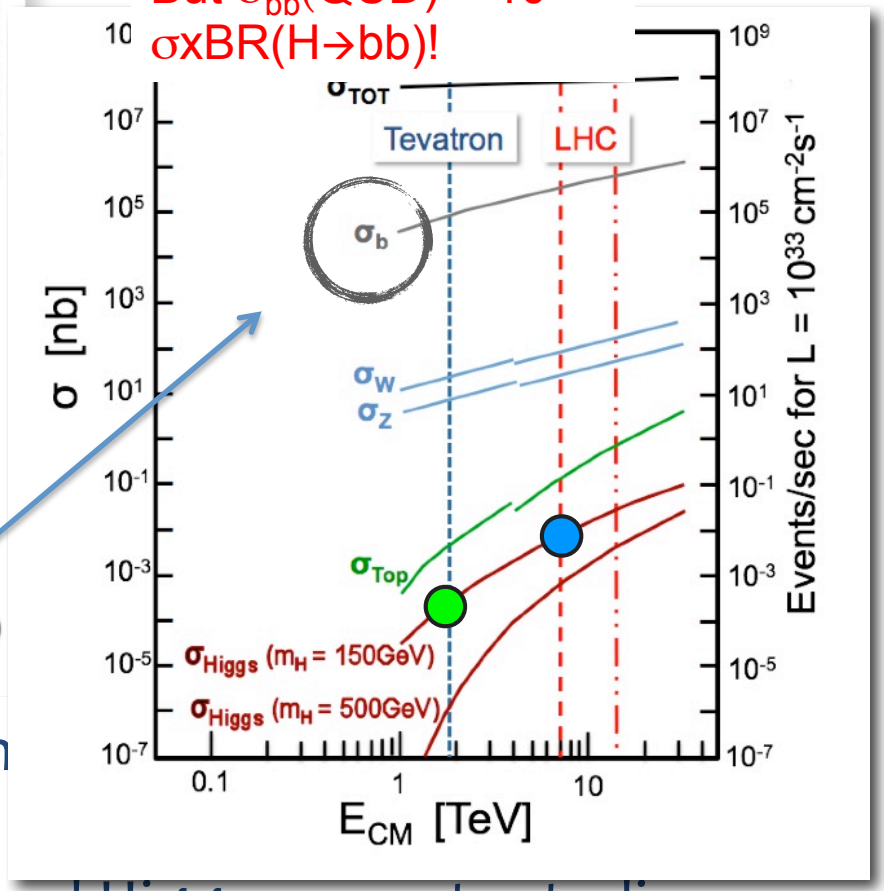
Vital part of the Higgs identification and Higgs property studies

But life is hard…

Region of interest ;-)



But σ_{bb}(QCD) ~ 10^7 σxBR(H→bb)!

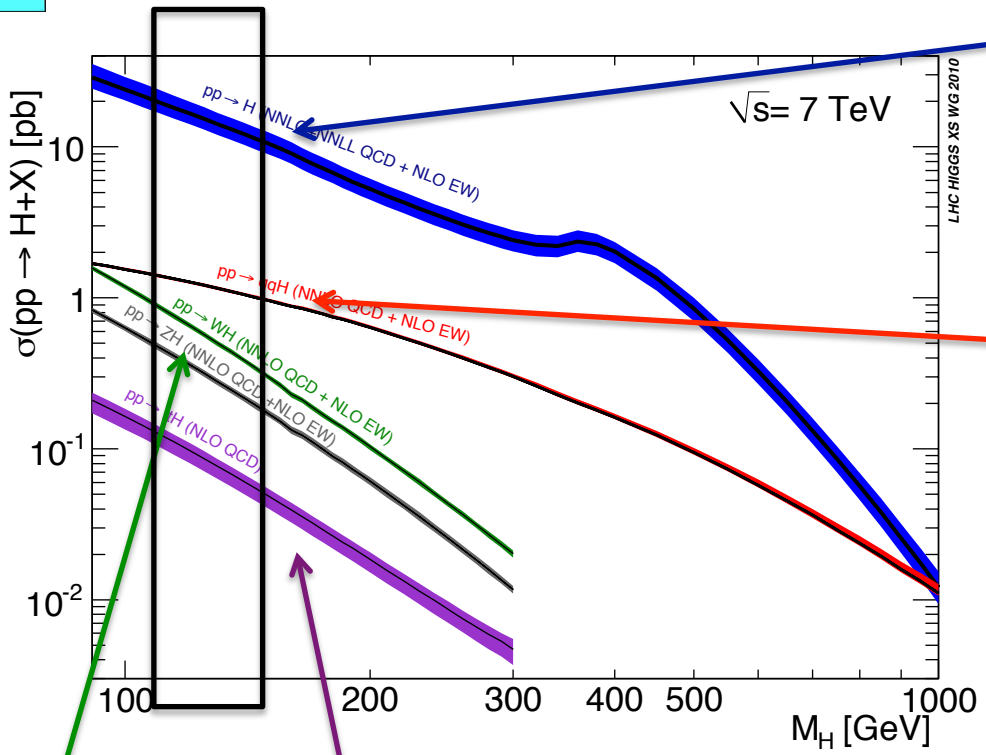


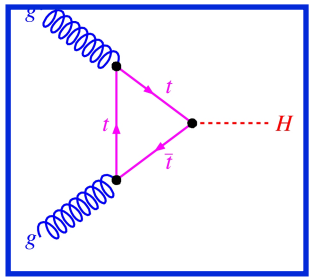
b's and tau's ← test the Higgs couplin

Vital part of the Higgs identification and Higgs property studies

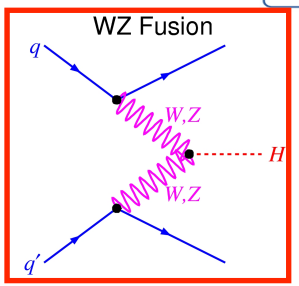
But life is hard…

func(Cross section,BRs,bkgs)





g-g fusion



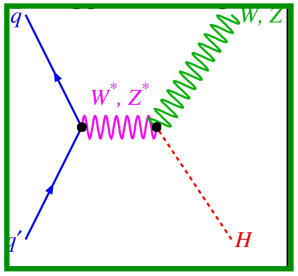
Vector boson fusion

H→ττ

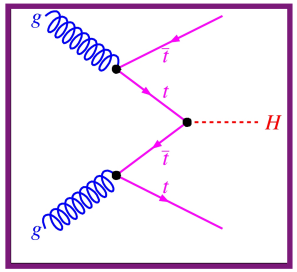
- gluon fusion difficult but possible(can use association with jets or go to high pt regime)
- VBF cleaner
- VH also added this year

very important for study Higgs properties!

Association with Z/W



Association with tt^bar

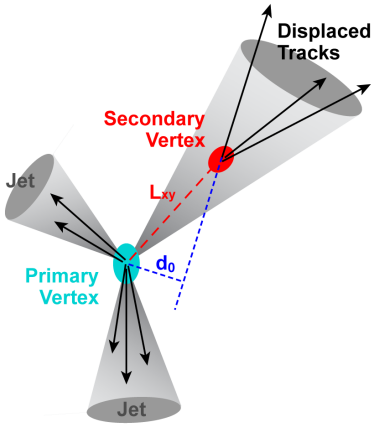


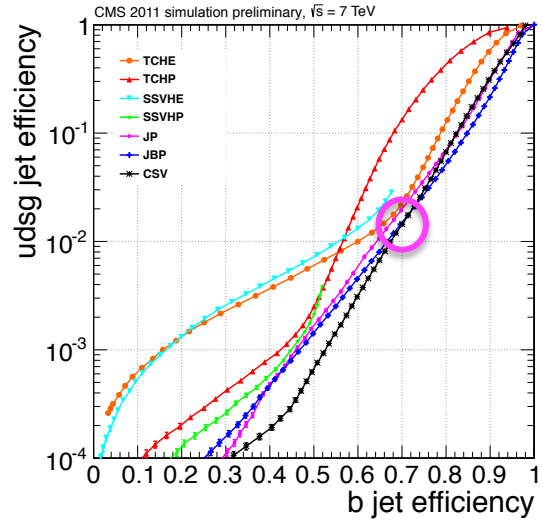
H→bb

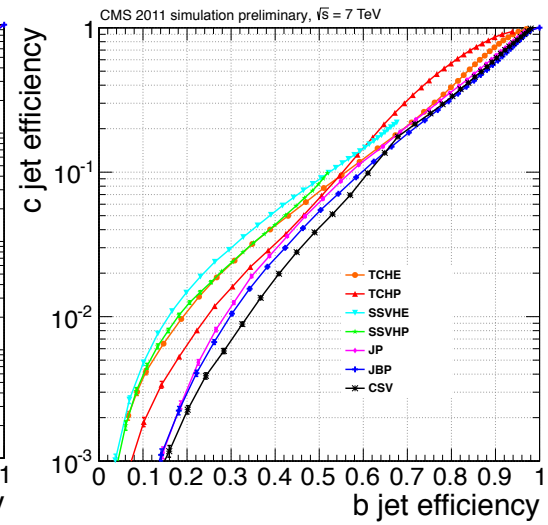
overwhelming QCD: need to go with associated production,

- VH - can use lepton/MET and topology
- ttH - even less rate, look at topology, but ttbar bkg well understood

B-jets at CMS

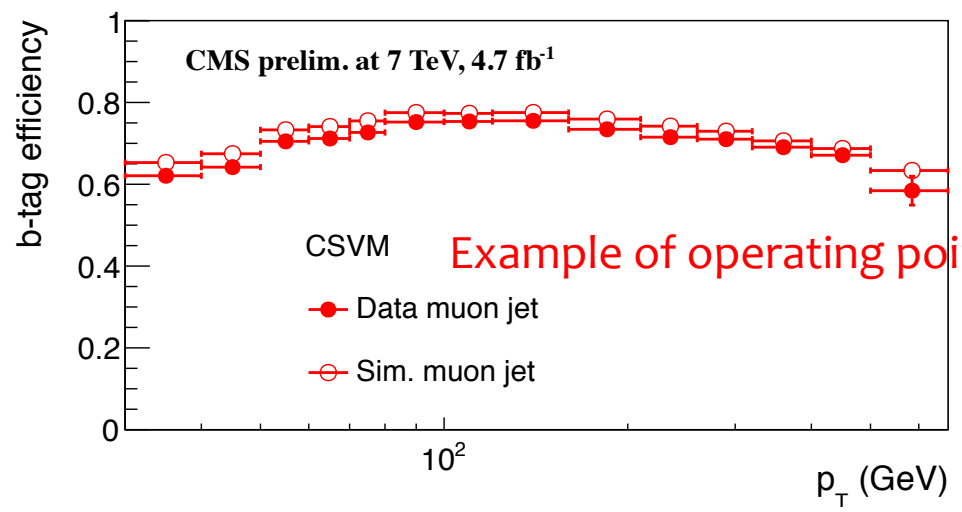






- Many algorithms deployed at CMS
- Best separation from udscg and c jets
  - **CSV**: likelihood tagger using many jet properties: secondary vertex (if any), tracks impact parameters, etc.

Eff of ~70% for a fake-rate of ~2%

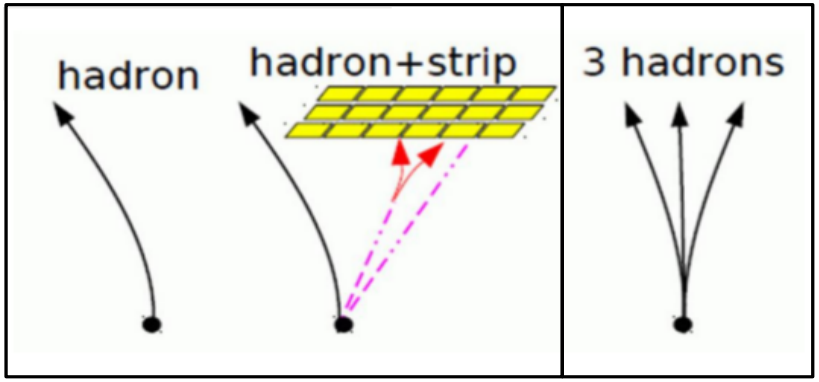


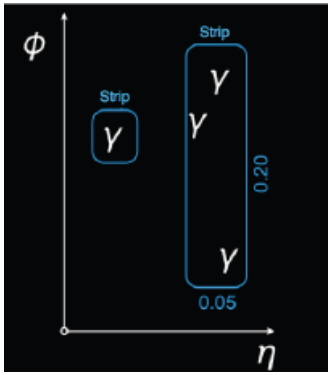
eff and fake-rate from ttbar samples and muon plus jets

Taus at CMS

Hadronic Tau identification:

- Reconstruct individual decay modes
- Charged hadrons + electromagnetic obj arranged in strips or single photons



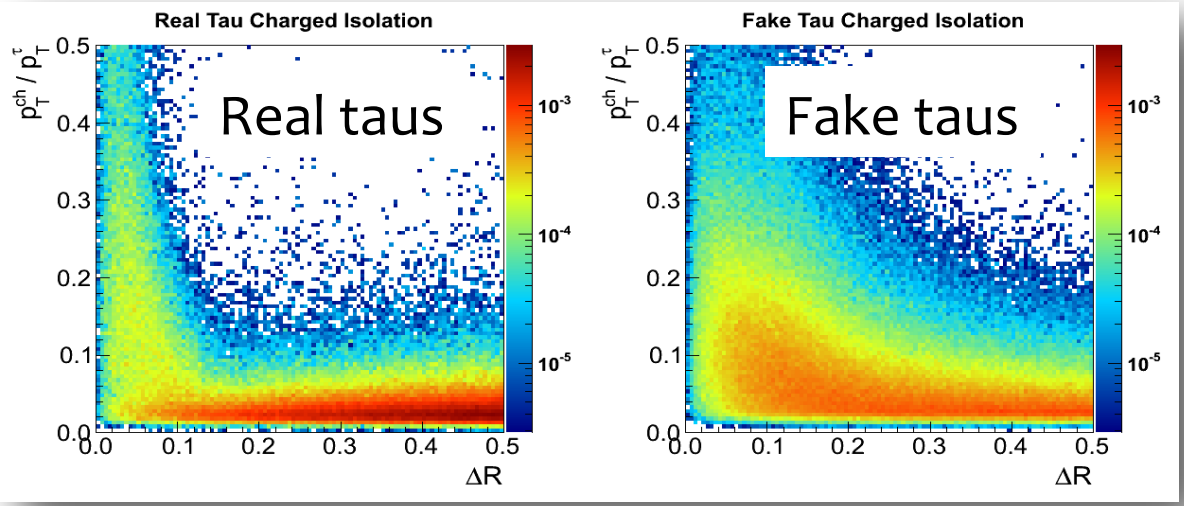


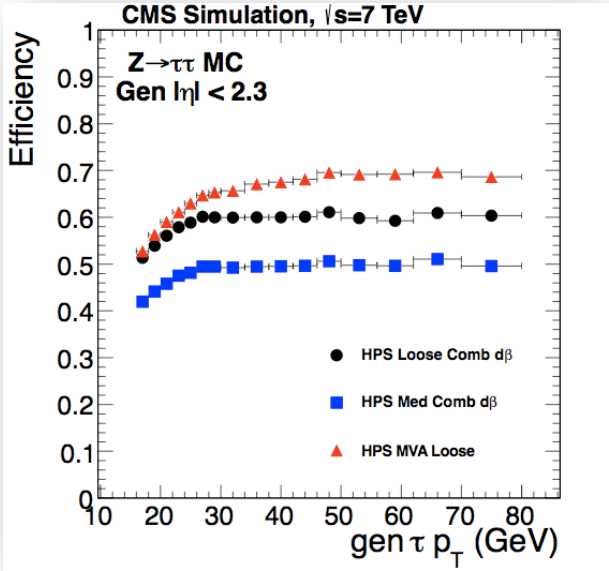
$\tau \to \pi\nu$	$\tau \to \rho\nu$	$\tau \to a_1\nu$
$\pi^\pm$	$\rho^\pm \to \pi^\pm \pi^0$	$a_1 \to \pi^\pm \pi^0 \pi^0$ $a_1 \to \pi^- \pi^+ \pi^-$ $a_1 \to \pi^+ \pi^- \pi^+$

Tau Isolation:

- Multivariate discriminator using sum of energy deposits in dR rings around the tau (from 0.1 to 0.5)

eff ~62% for a fake rate of ~6%

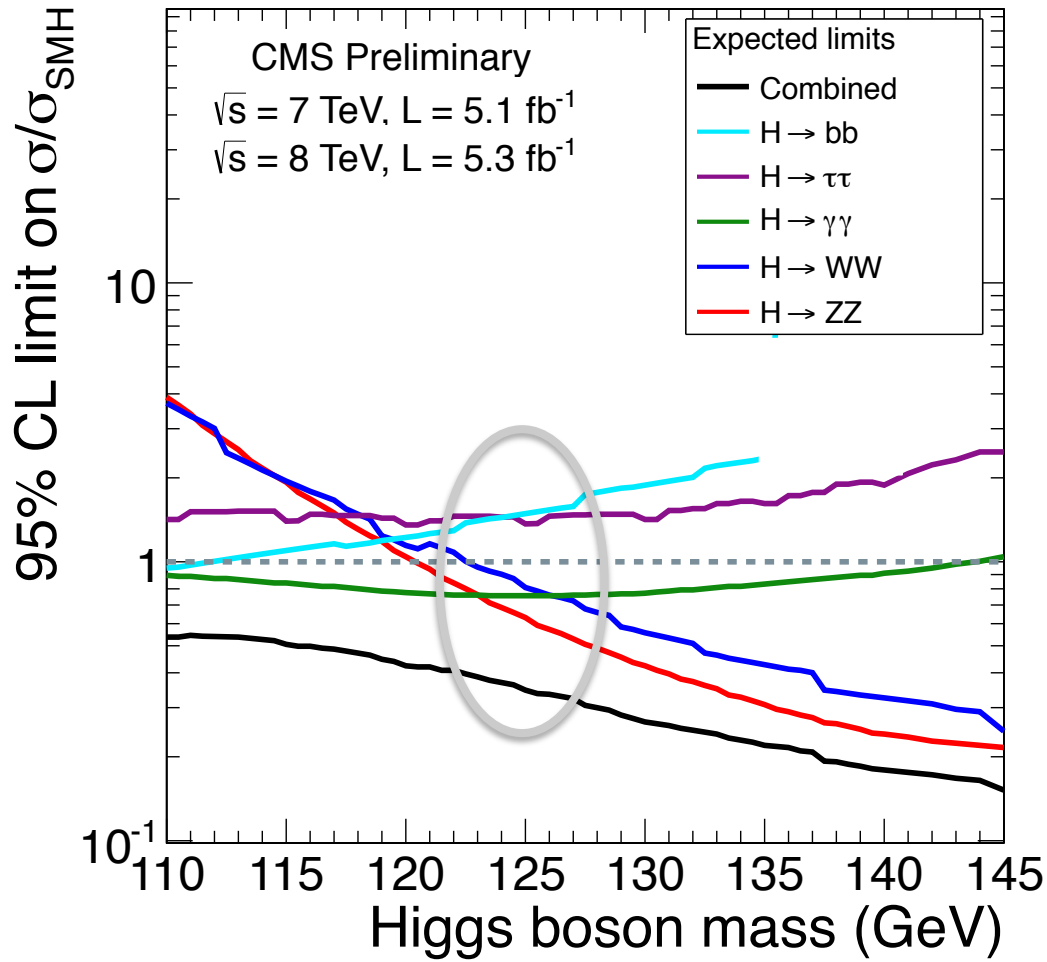




SM Higgs search

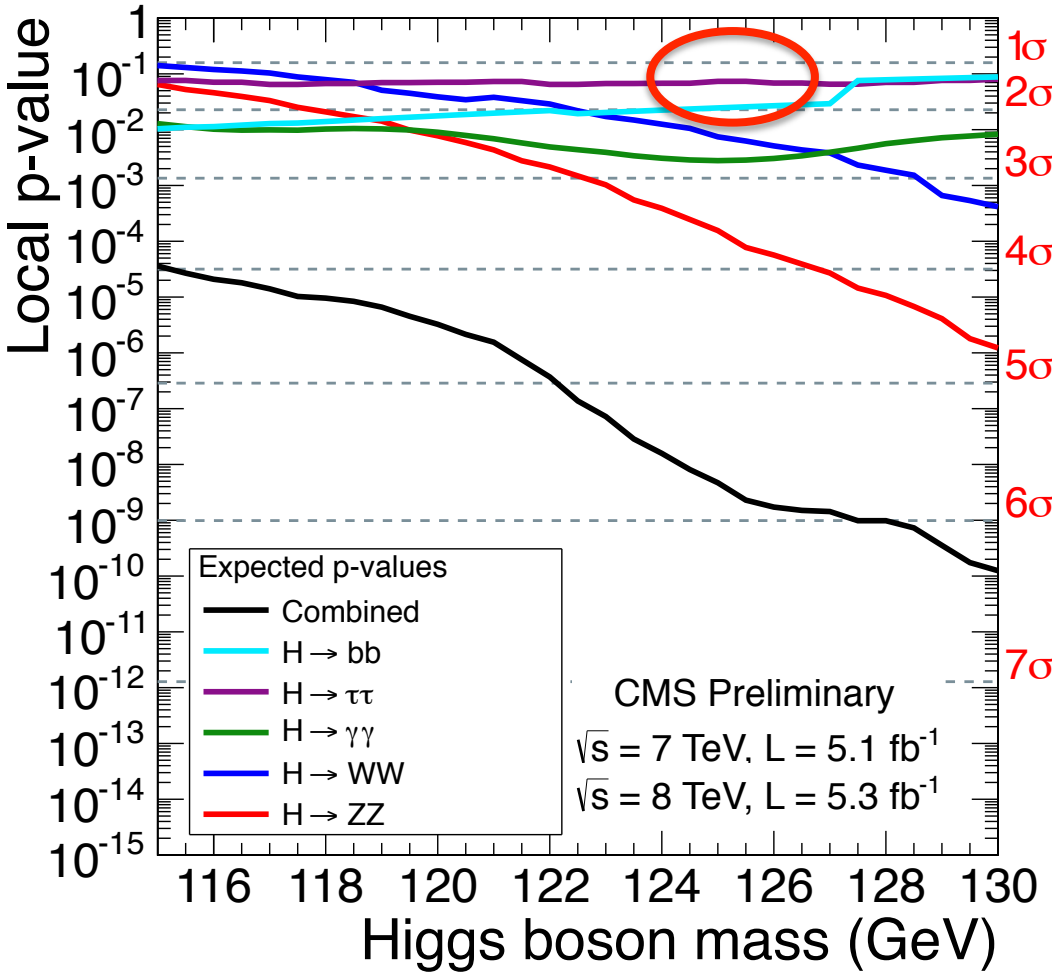
Expected sensitivity at low mass

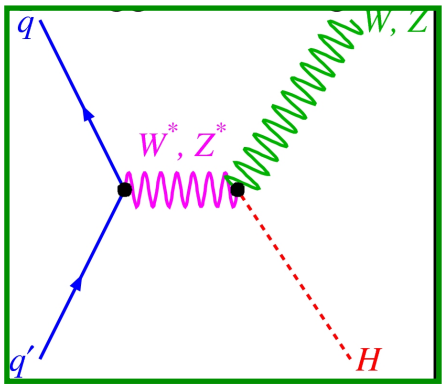
- 125 mass: b and τ have similar sensitivities for exclusion!
  - b and τ are ~2 less sensitive that ZZ, γγ and WW
  - But with them we could really test if the new particle is SM



Expected sensitivity at low mass

- 125 mass: b and τ have similar sensitivities for exclusion!
  - b and τ are ~2 less sensitive that ZZ, γγ and WW
  - But with them we could really test if the new particle is SM
- If excess: p-values complement the picture
  - b and τ needs more data to see a significant excess





VH → Vbb

V→lν,ll,νν

Most sensitive channel with b as final state, also interesting to compare with Tevatron results

VHbb: Search strategy

First CMS VHbb 2011 analysis: Phys. Lett. B 710(2012) 284-306

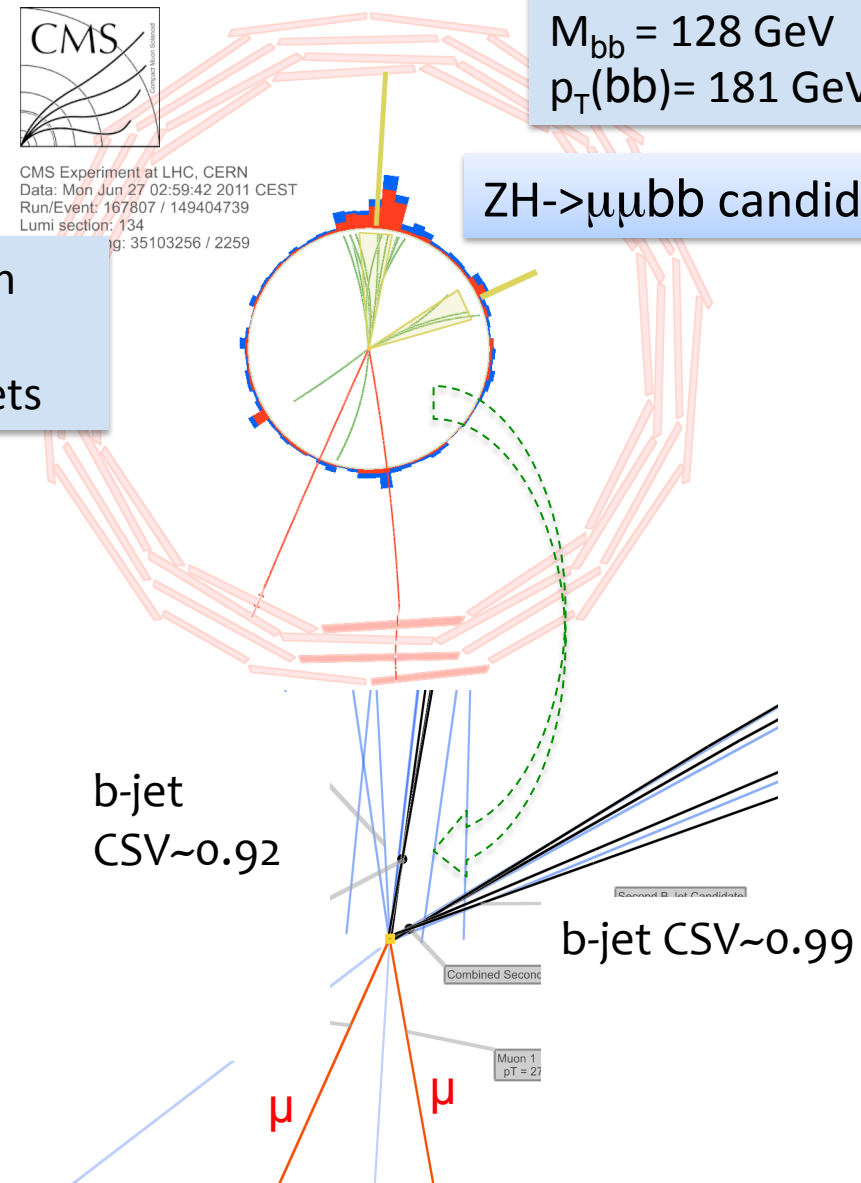
- $5fb^{-1}$@7+$5fb^{-1}$@8 TeV fb^{-1} analysis, 5 channels:
  - Z(ll)Hbb
  - Z(νν)Hbb
  - W(lν)Hbb

Associated Production → final states with leptons, MET and b-jets

General strategy:

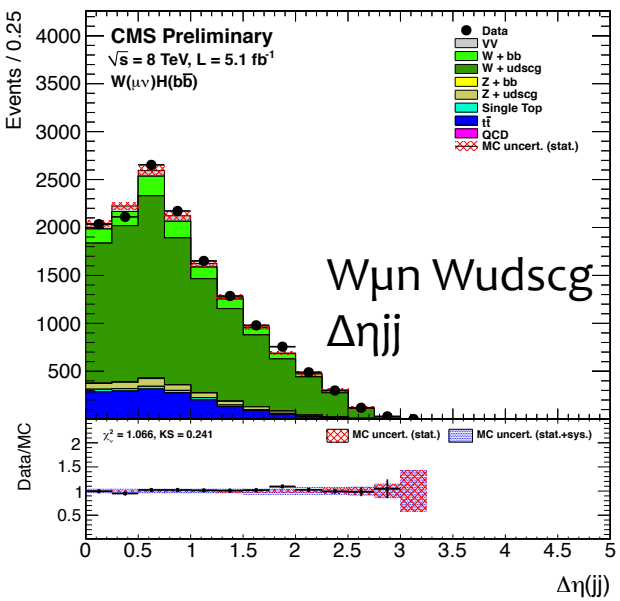
(a) High boosted vector boson and dijet,
(b) 2 b-tagged jets,
(c) back-to-back V & H
(d) Reconstruct m_{bb}

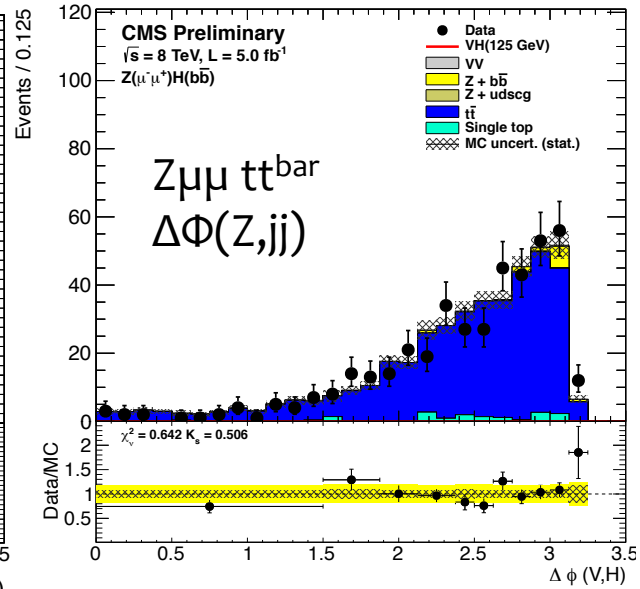
Main backgrounds (V+jets, tt^{bar})estimated from data in control regions

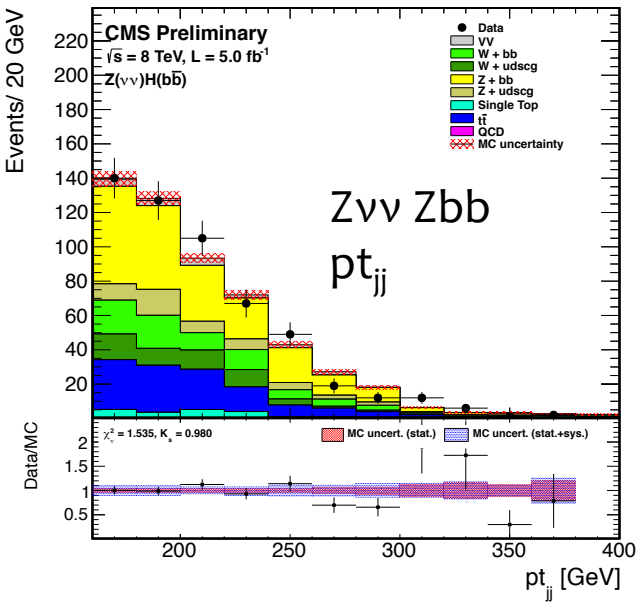


VHbb: control regions

- Background calibrated in control regions (CRs)



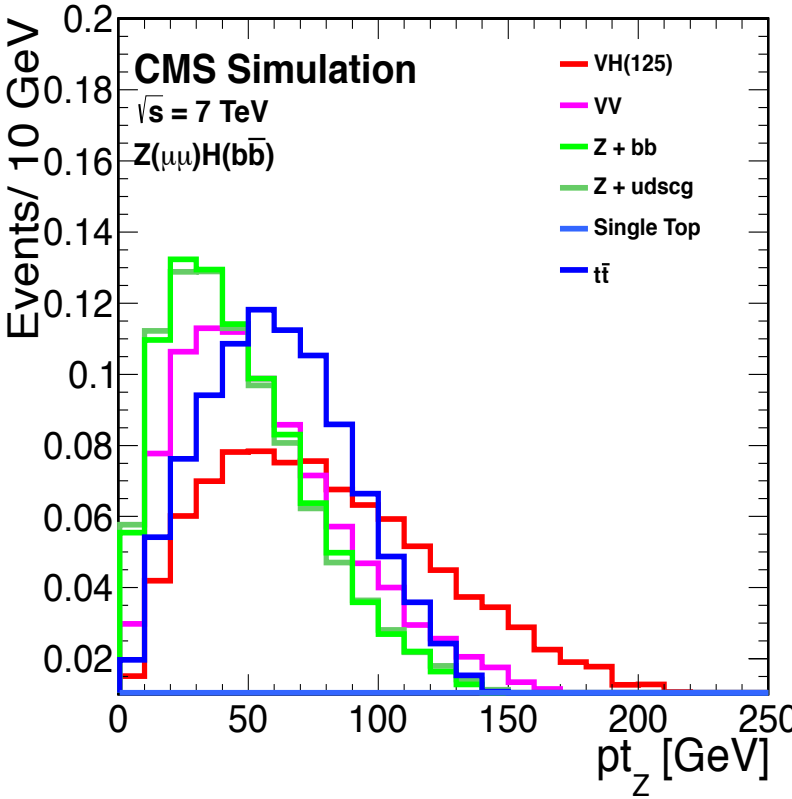




Examples of variables used in the analysis

VHbb

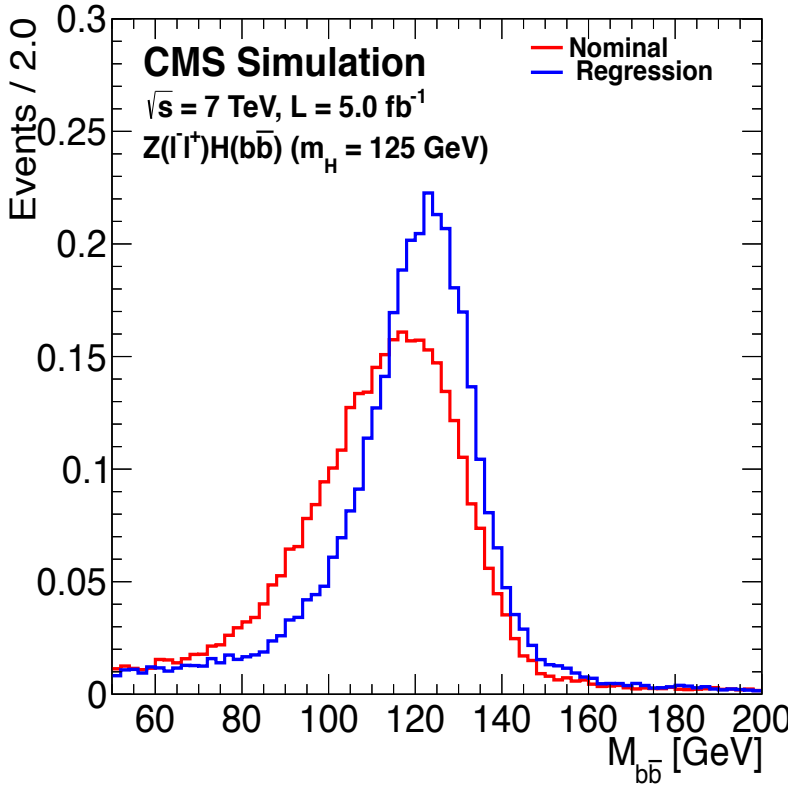
- Analysis presented at Jul4/ iCHEP shows several improvements:
  - add events at medium boost



Channel	Medium boost	High boost
ZllH	50<Zpt<100	Zpt>100
WlnH	120< Wpt<170	Wpt>170
ZnnH	120<Zpt<160	Zpt>160

VHbb

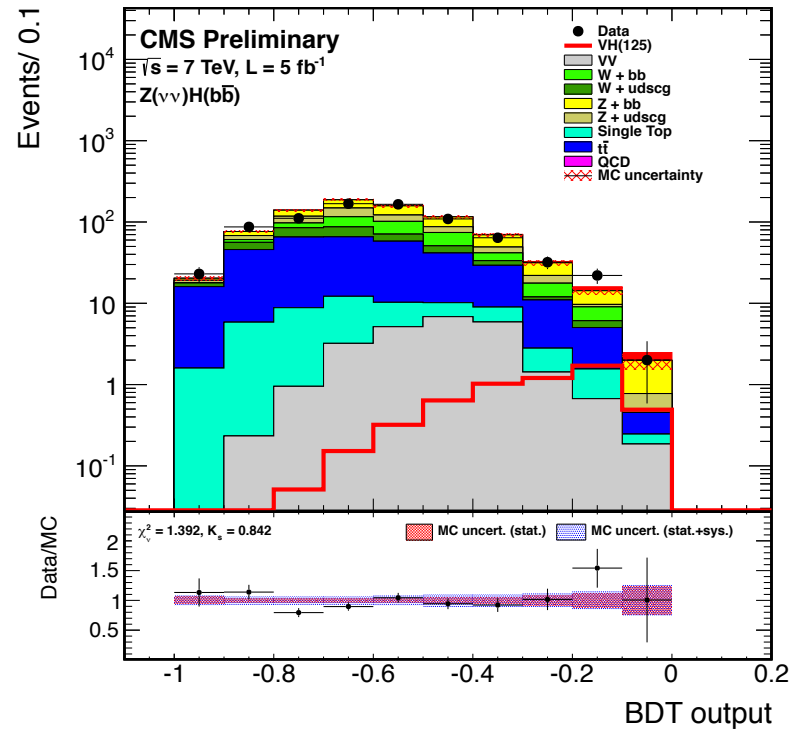
- Analysis presented at Jul4/ iCHEP shows several improvements:
  - add events at medium boost
  - (b-)Jet energy reconstruction using mva regression



Extensively validated in data using V + dijets balance in Z(ll) + bb, ttbar and single top events in channels with MET

VHbb

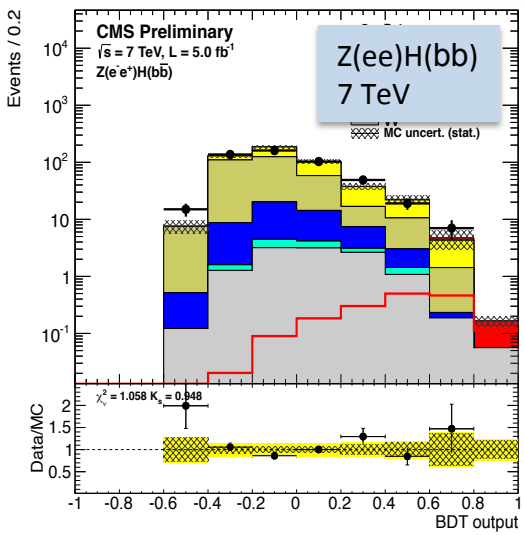
- Analysis presented at Jul4/ iCHEP shows several improvements:
  - add events at medium boost
  - (b-)Jet energy reconstruction using mva regression
  - Fit full shape of final MVA discriminator

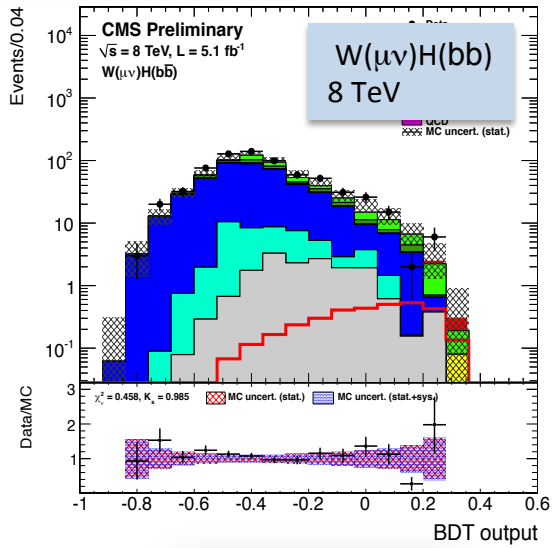


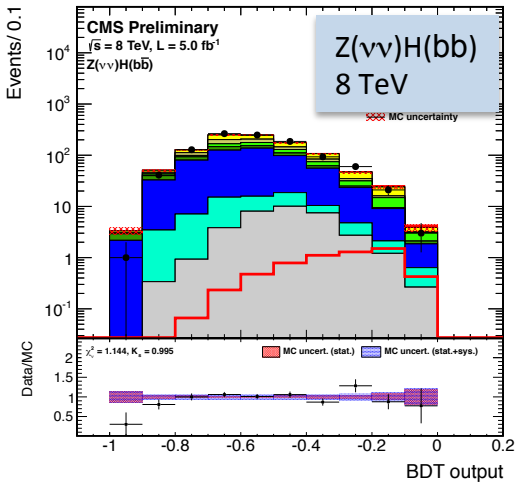
Gain 50% of sensitivity already with 2011 dataset

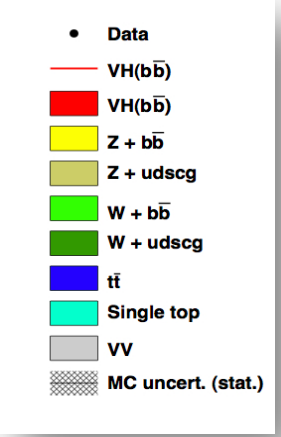
VHbb: Results

Examples of final BDT distributions



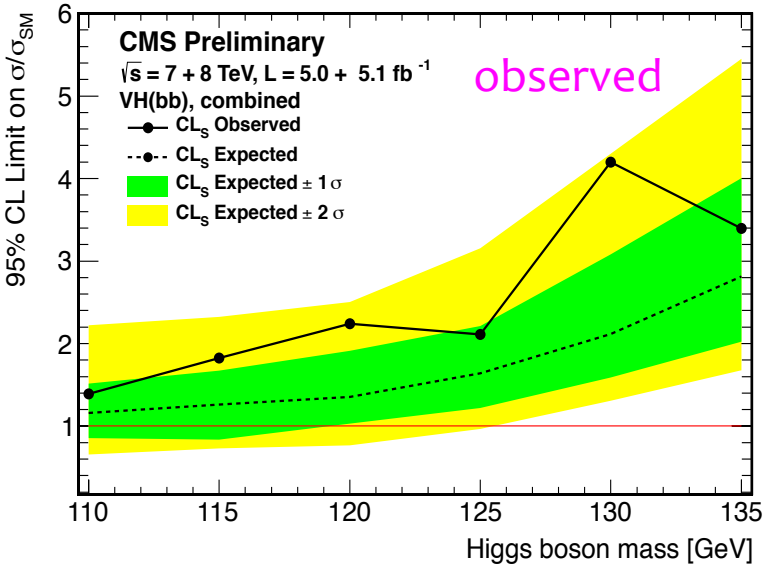






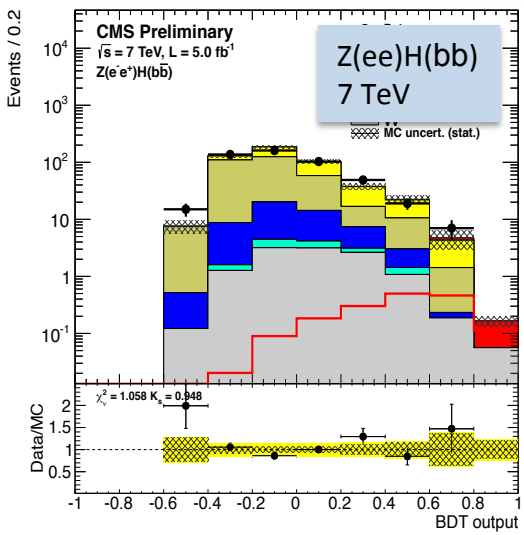
limits at 125:

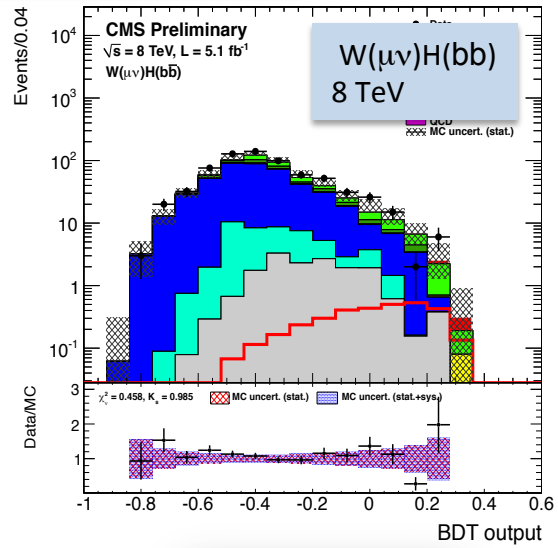
- exp(obs) 1.6(2.1)
  - exp comparable with Tevatron
- Compatible with either background or signal from a 125 GeV Higgs

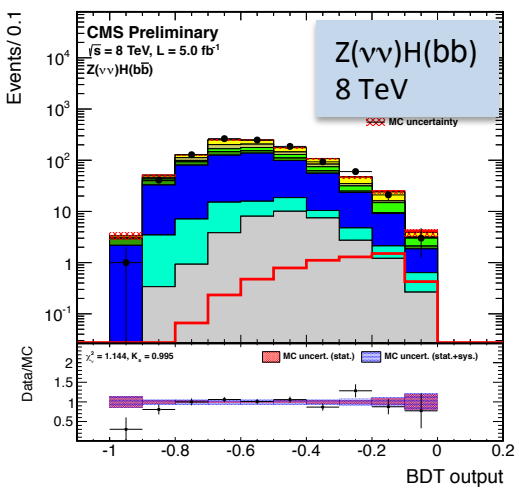


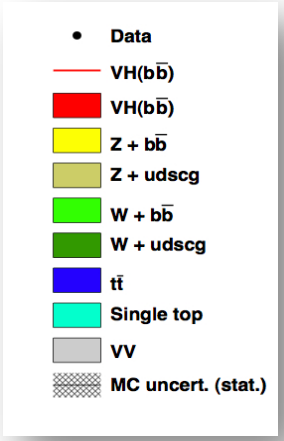
VHbb: Results

Examples of final BDT distributions



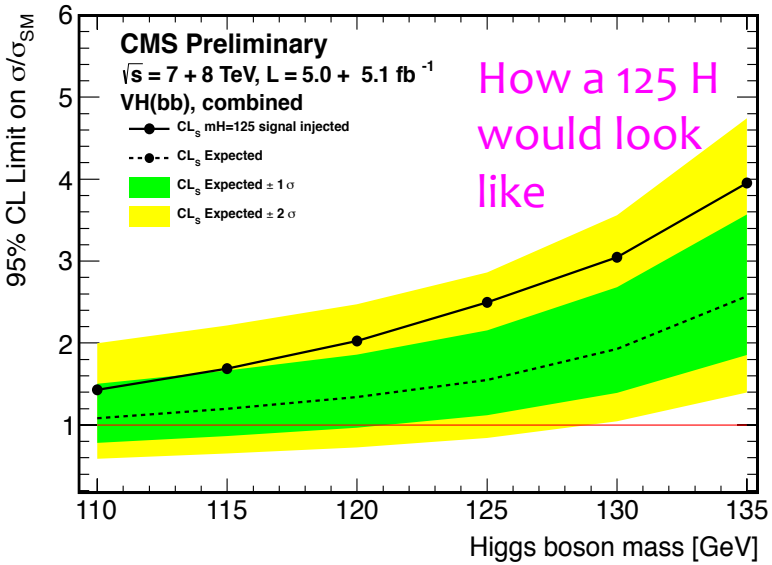






limits at 125:

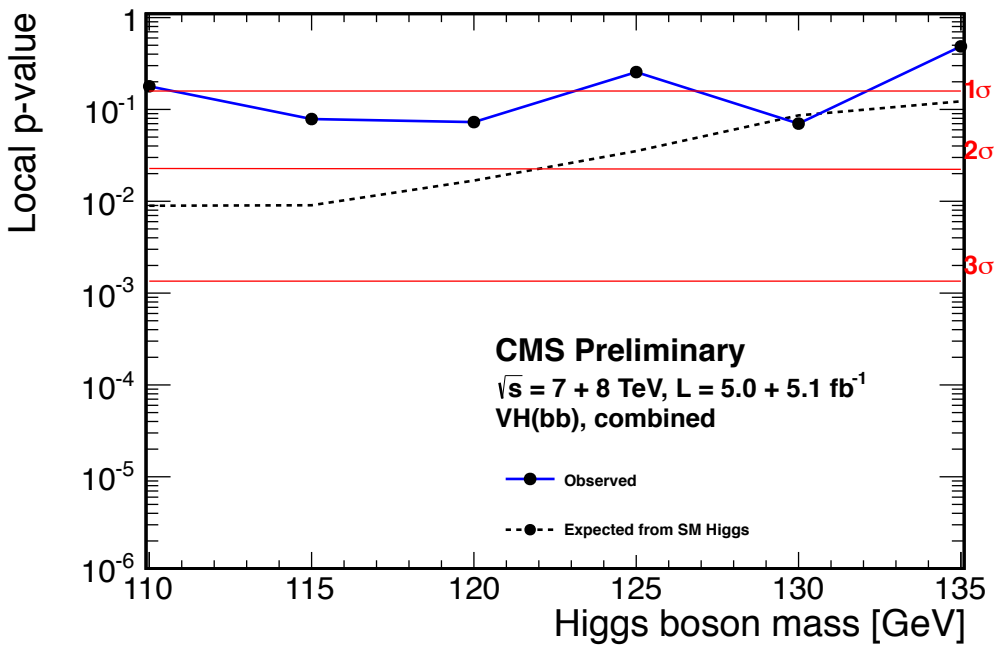
- exp(obs) 1.6(2.1)
  - exp comparable with Tevatron
- Compatible with either background or signal from a 125 GeV Higgs

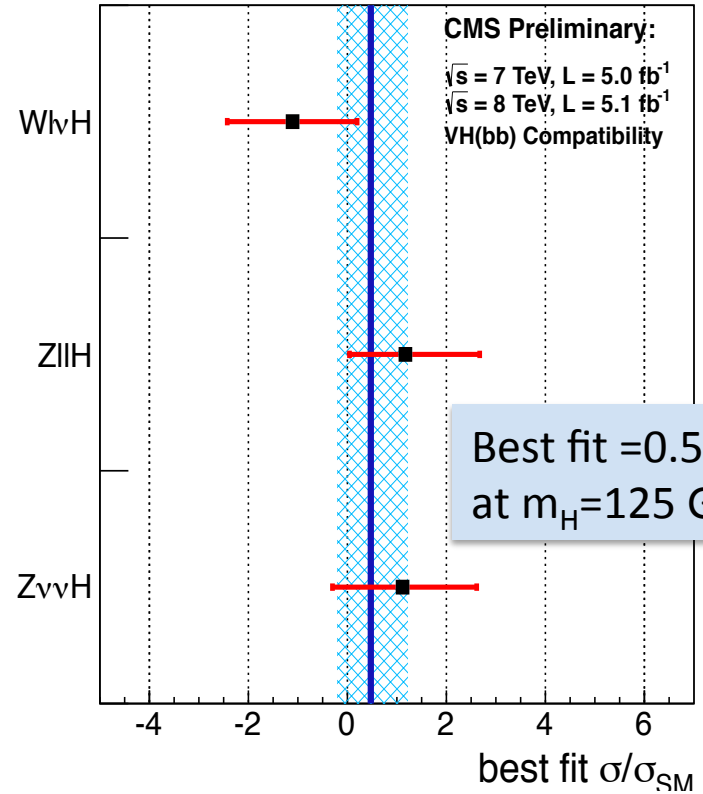


VHbb signal strength and p-value

Observed limits:

- Compatible with either background or signal from a 125 GeV Higgs





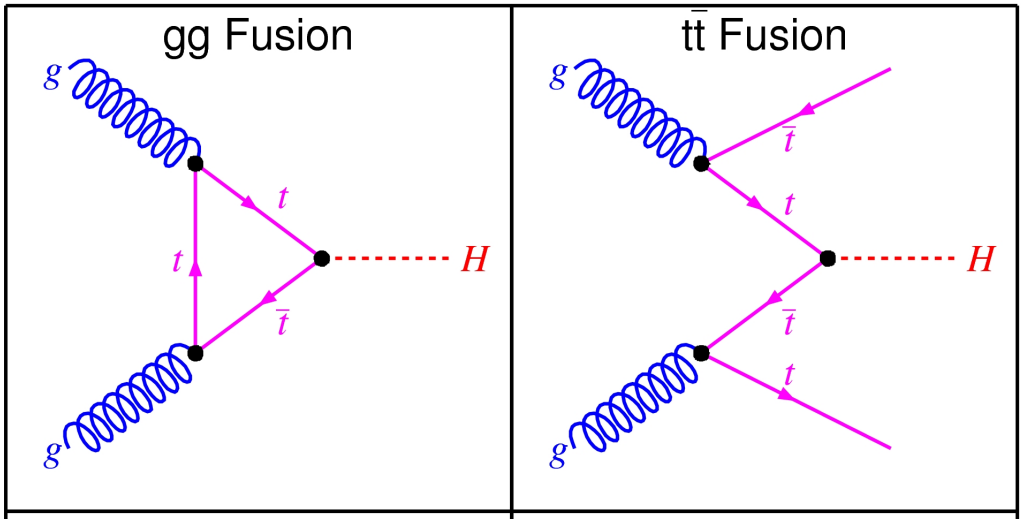
Best fit $=0.5 \cdot \sigma_{SM}$ at m_H=125 GeV

Some excess compared to background predictions (significance ~1σ everywhere)

Brand new

gg→tt(H)bb

Higgs fermions coupling both in production and decay

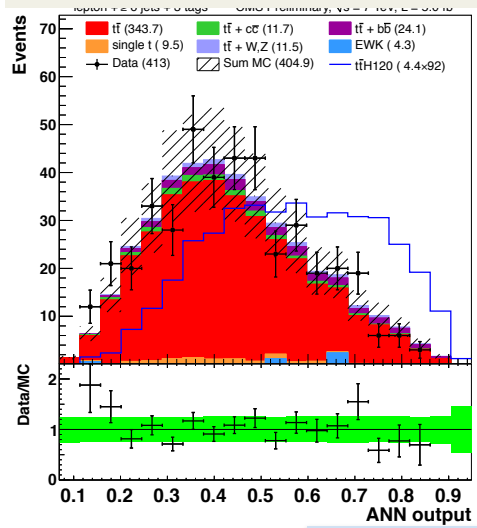


ttH, H→bb

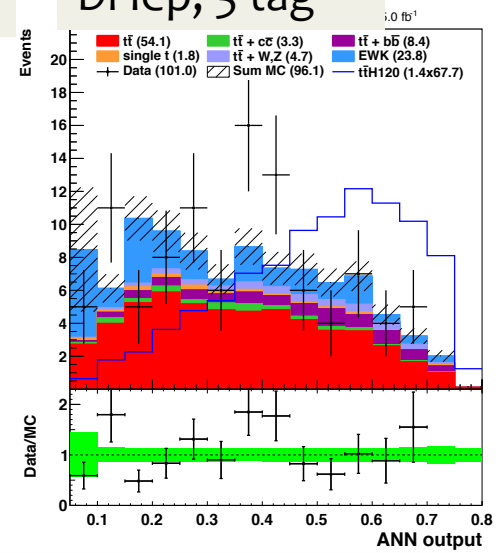
Strategy:

- Separate events by
  - top decay mode
    - di-lep. vs lep+jets (most sensitive)
  - number of jets and b-tags
- fit MVA shapes in each event category
  - low S/B categories used to constrain the background, mainly ttbar +(b)jets
  - Most relevant variables are b-tag, kinematic and angular correlation (for example minΔR beetween b-jets pair)
- 2011 7 TeV data results public!

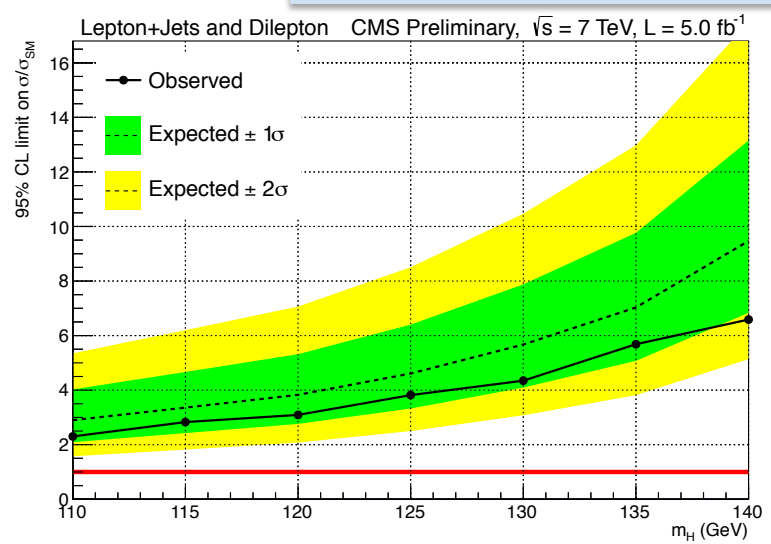
Lep+jets, 6 jets, 3 tag



Di-lep, 3 tag

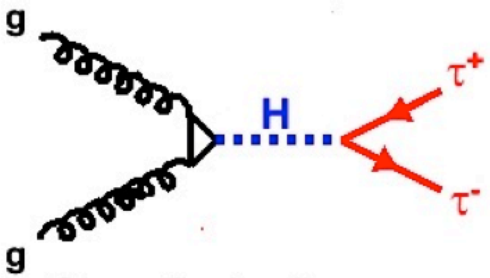


No excess: exp(obs) limit at mH(125) = 4.6(3.8)

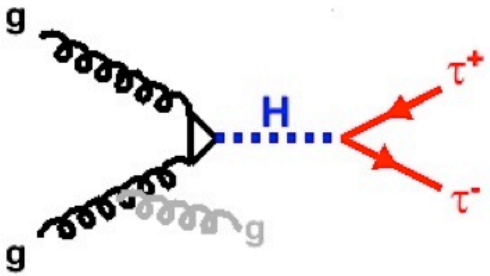


$H \rightarrow \tau\tau, \rightarrow \mu\tau_h, e\tau_h, e\mu, \mu\mu$

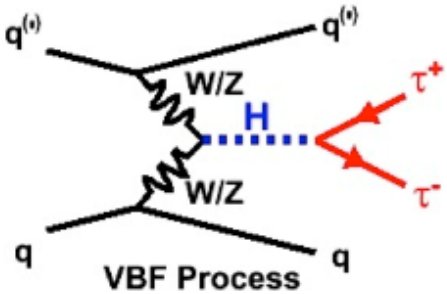
- Sensitive to all production modes
- Probes coupling to leptons
- Enhanced σ x BR in MSSM



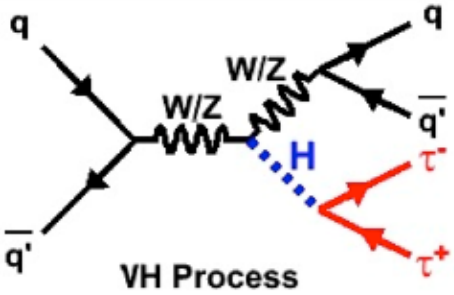
Gluon Fusion Process



Boosted Gluon Fusion Process



VBF Process



VH Process

$H \to \tau\tau$: analysis strategy

- Search performed in 4 tau-pair final states: **$\mu\tau_h$, $e\tau_h$, $e\mu$, $\mu\mu$**
- Analysis divided into 5 categories: mass resolution, S/B
- All categories are fit simultaneously

Jets p_T > 30 GeV →

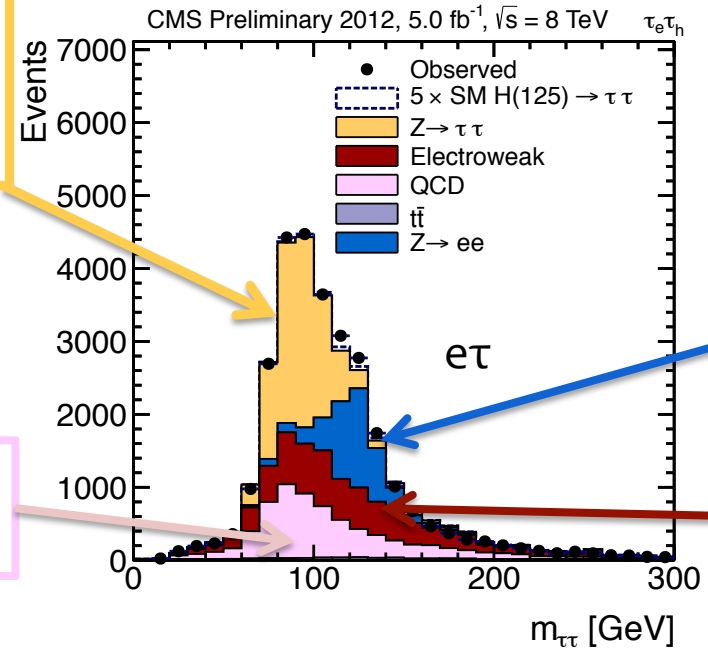
0 Jet, Low p_T High background	**1 Jet, Low p_T** Enhancement from jet requirement
0 Jet, High pT Lepton p_T spectrum harder from H	**1 Jet, High p_T** Enhancement from p_T and jet requirement

VBF
2 jets, no jets in rapidity gap
MVA based selection

categorization based on τ_h p_T for $\mu\tau_h$, $e\tau_h$; μ p_T for $e\mu$; leading μ p_T for $\mu\mu$

H→ττ: backgrounds

DY→ττ – Efficiency measured using τ embedded μμ events

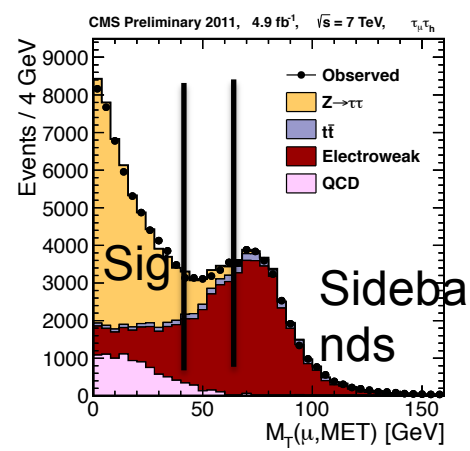


DY→ll – from MC but correcting the l→ τ fake rate

QCD - estimated from same sign data

EWK – mostly W+jets: using M_T sidebands and angles between MET and τs

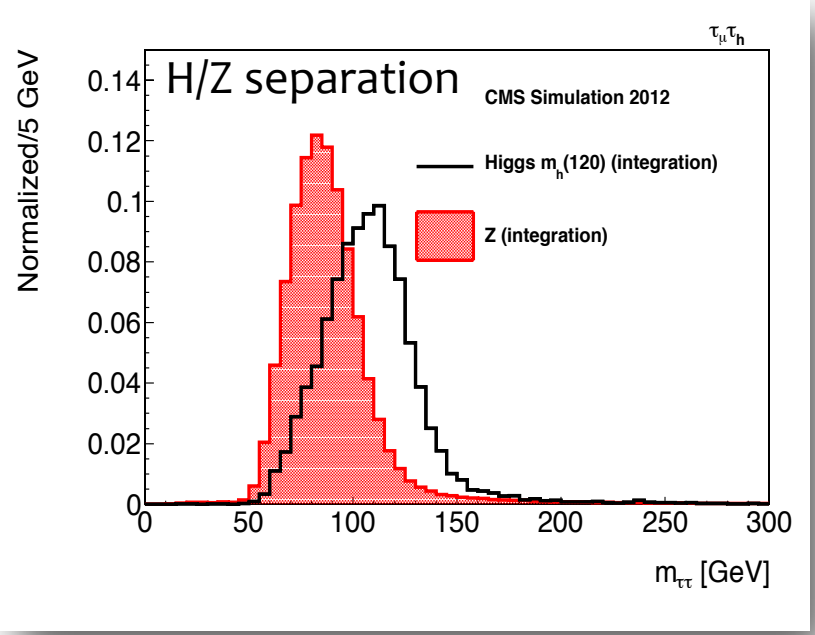
Backgrounds measured as much as possible form data

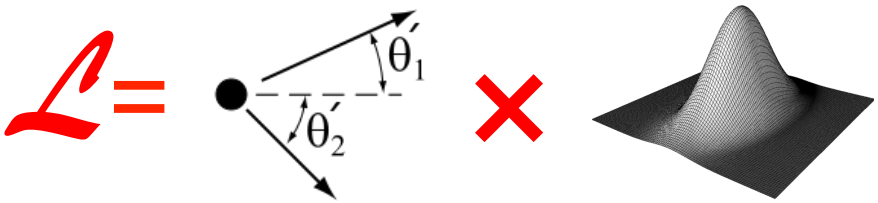


H→ττ

2011 analysis in PLB: Phys. Lett. B 713(2012) 68-90

- New
  - Improved tau ID and new mass reconstructions (20% better mass resolution)

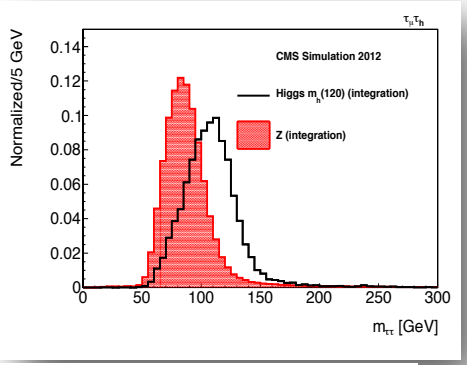




- SVFit
  - Event-by-event estimator of true m(ττ) likelihood
    - Matrix Element used for τ→lνν
    - Phase-Space is used for τ→π
- Mass peaks at true value
  - Better separation of H from Z

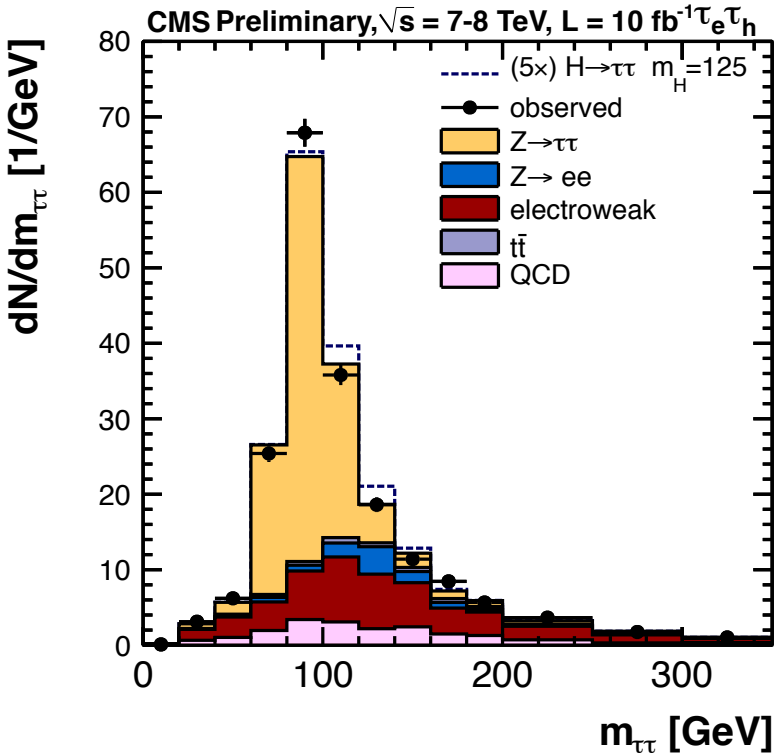
H→ττ

2011 analysis in PLB: Phys. Lett. B 713(2012) 68-90



- New
  - Improved tau ID and new mass reconstructions (20% better mass resolution)
  - New event categorization: lower jet p_T thresholds, rely also on p_T of the tau, adding VH modes

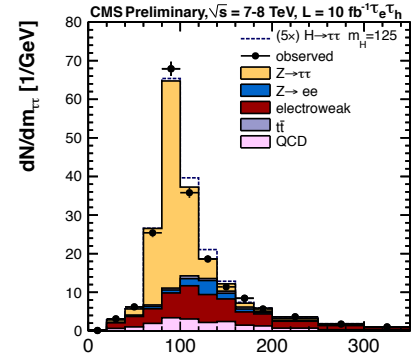
High/low pt split makes the 1 jet category more powerful

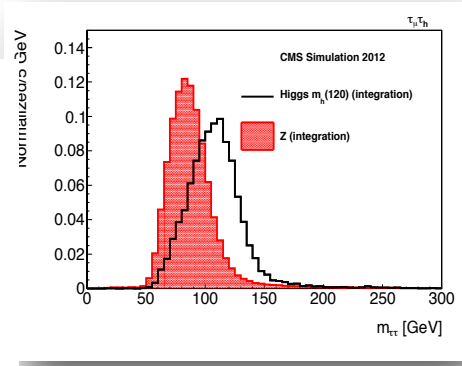


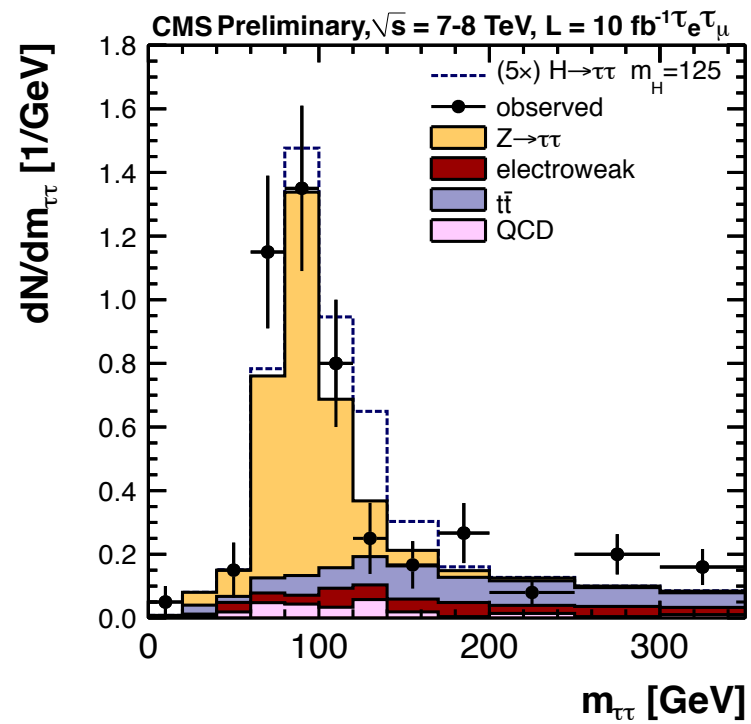
H→ττ

2011 analysis in PLB: Phys. Lett. B 713(2012) 68-90

- Since then
  - Improved tau ID and new mass reconstructions (20% better mass resolution)
  - New event categorization: lower jet p_T thresholds, rely also on p_T of the tau, adding more VH modes
  - MVA analyis for VBF category

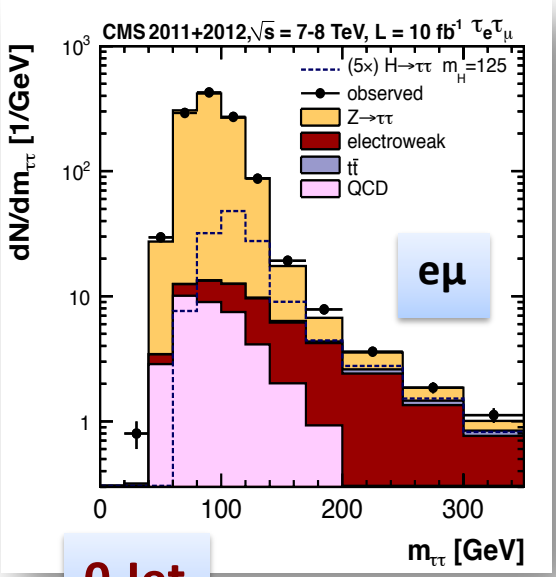






Total of ~70% improvement already with 2011 data alone

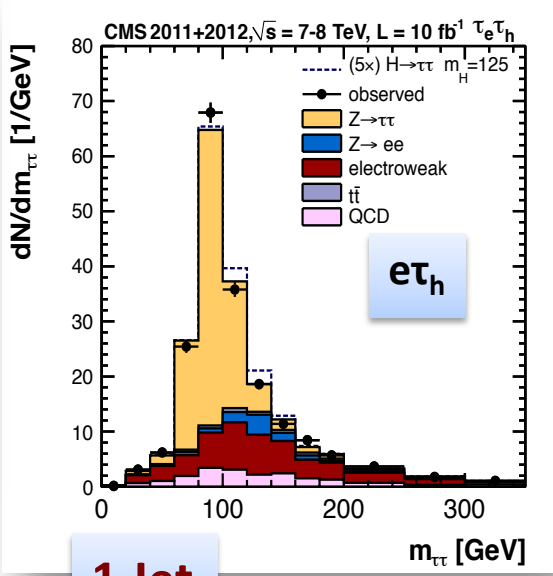
$m_{\tau\tau}$ in Event Categories



0 Jet

Constrains energy scales and efficiencies

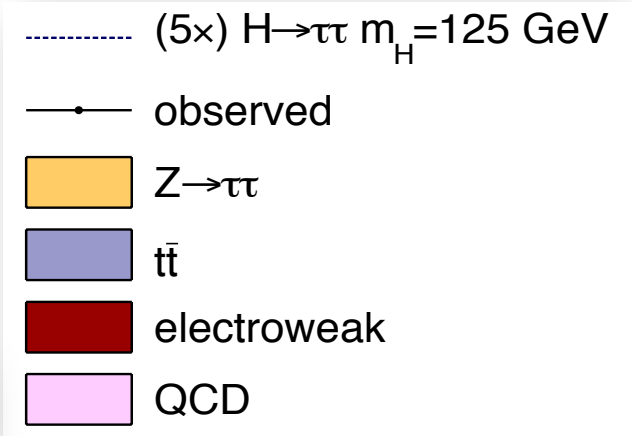
- Large Drell-Yan background
- Sensitivity boosted by low/high p_T split

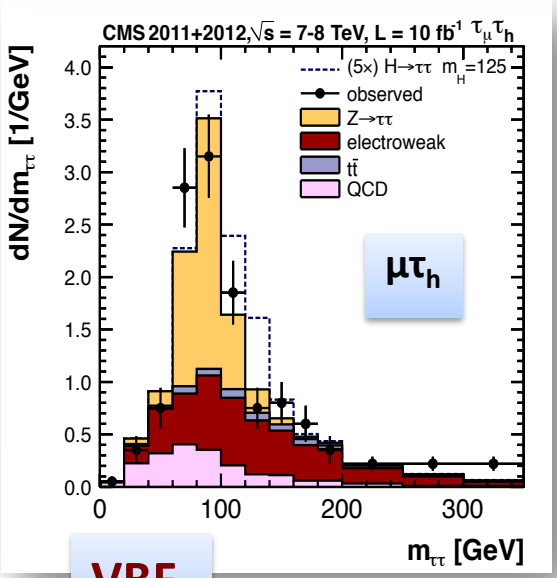


1 Jet

Enhanced sensitivity to gluon fusion

- Improved mass resolution
- Increased sensitivity by splitting into low and high p_T categories





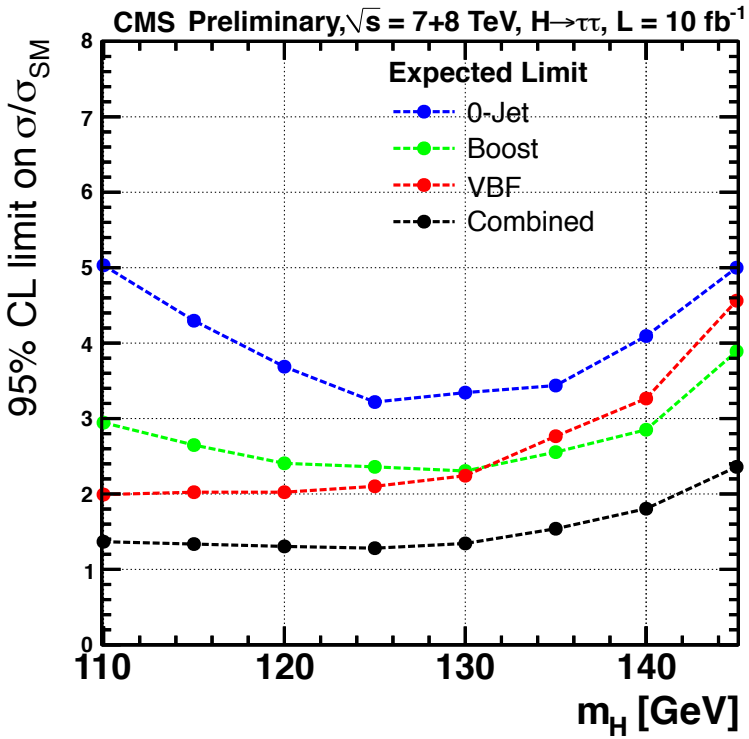
VBF

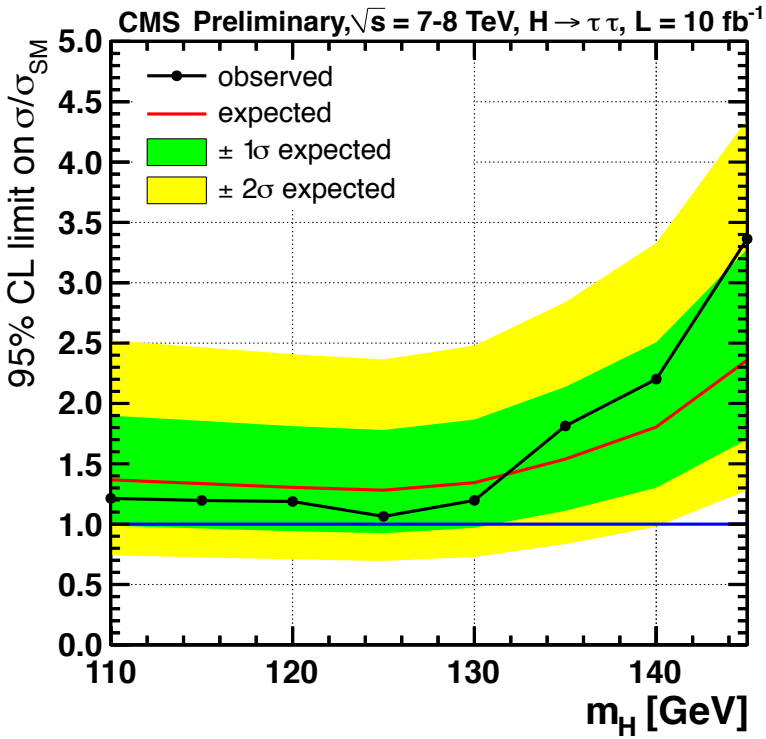
Enhanced sensitivity to VBF production

- Highest sensitivity for $m_H < 130$ GeV

H→ττ: Results

- Sensitivity of new analysis very close to $1\times\sigma_{SMH}$
- No excess seen from SM background-only expectation
  - Observed limit of 1.06 x SM at m_H = 125 GeV (exp 1.3)
  - Bad luck or non SM Higgs?
    - driven by the VBF (most sensible) category
    - Need to reach 1×σSM sensitivity in all cathegories: feasable for the end of the year!

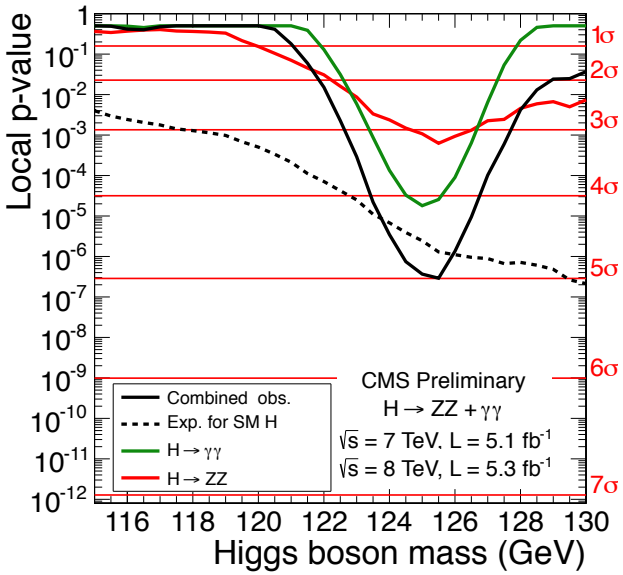


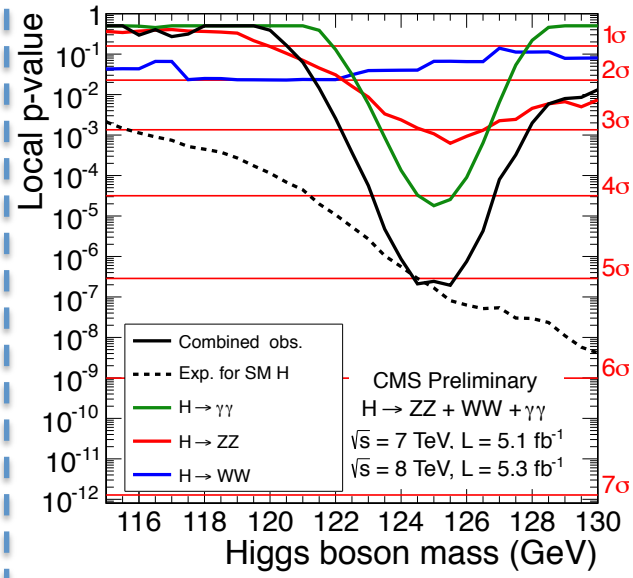


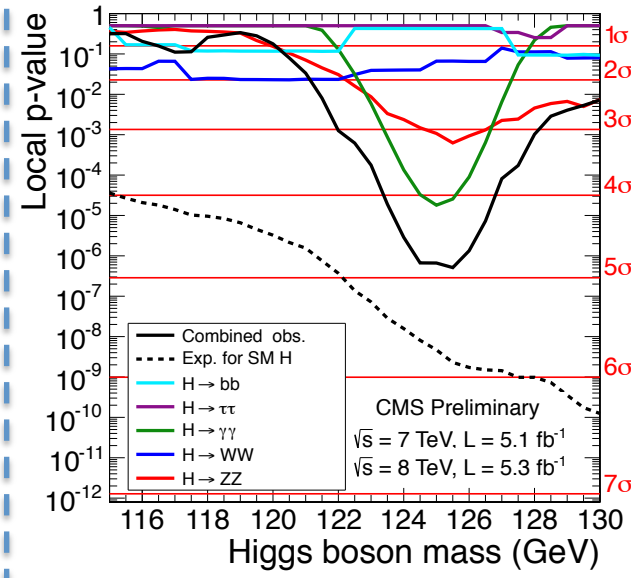
H→ ff in the context of the whole SM Higgs search

more in M.Chen talk

P-values evolution







- γγ and ZZ combination
- 5σ excess (exp. 4.7)

Discovery of a new boson!

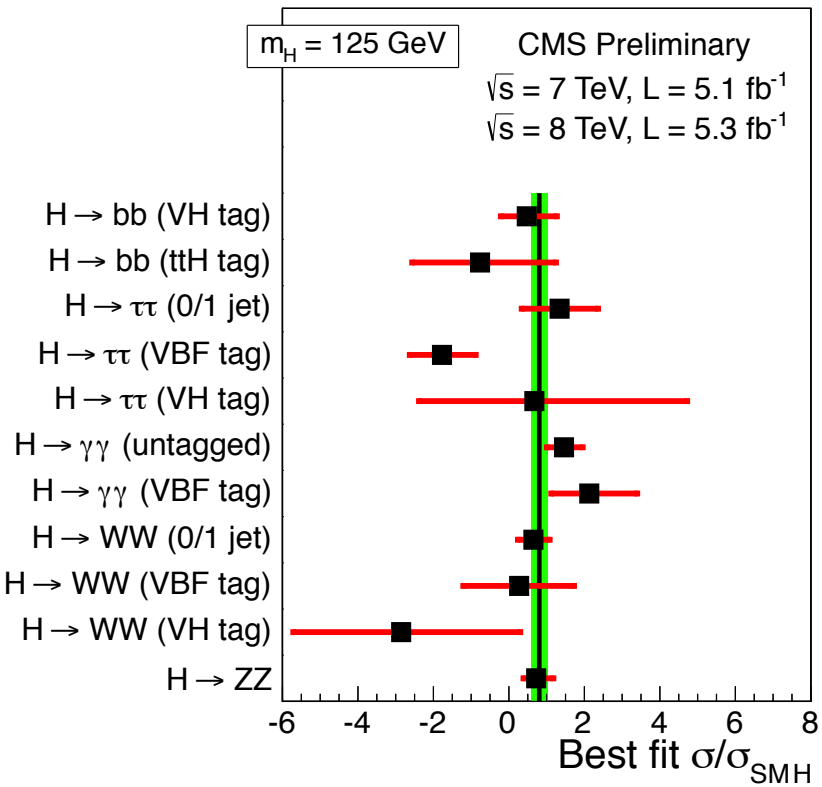
- Adding WW
- 5.1 σ excess (exp. 5.2)

WW/ZZ ok

- Adding bb and ττ
- 4.9 σ excess (exp. 5.9)

Still compatible with SM, of course worth keeping an eye here

Is this "the" Higgs boson?

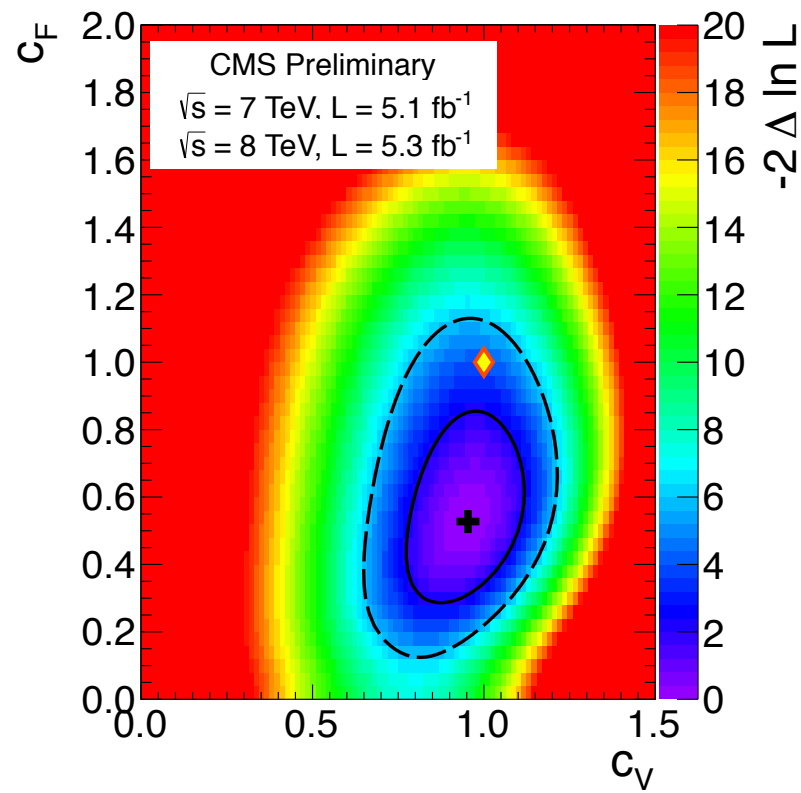


Good agreement among modes:

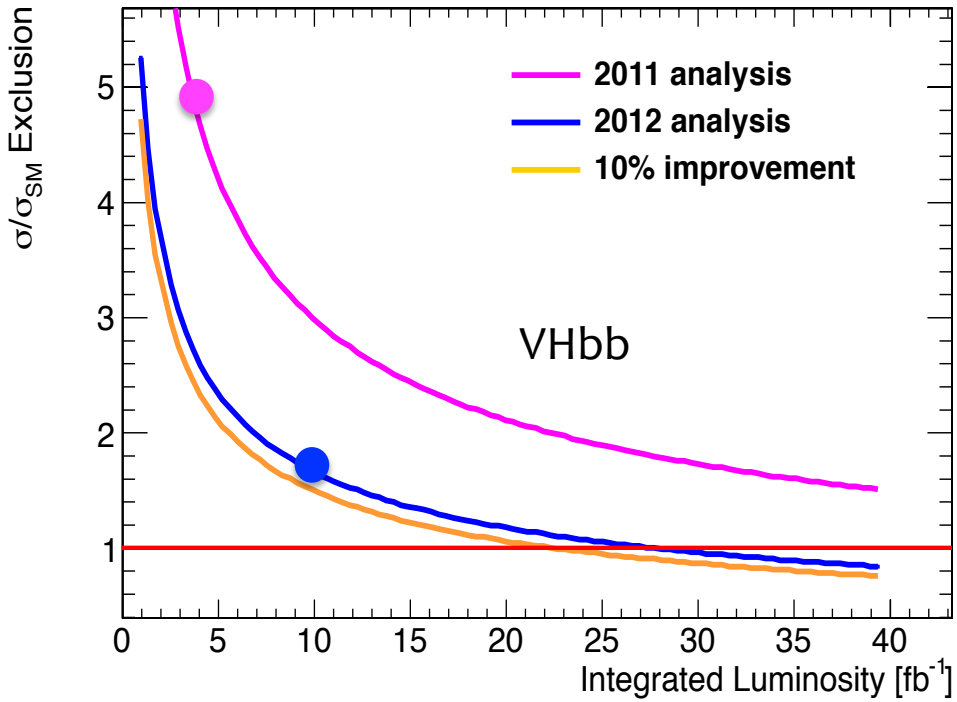
- Exception: γγ large and ττ small

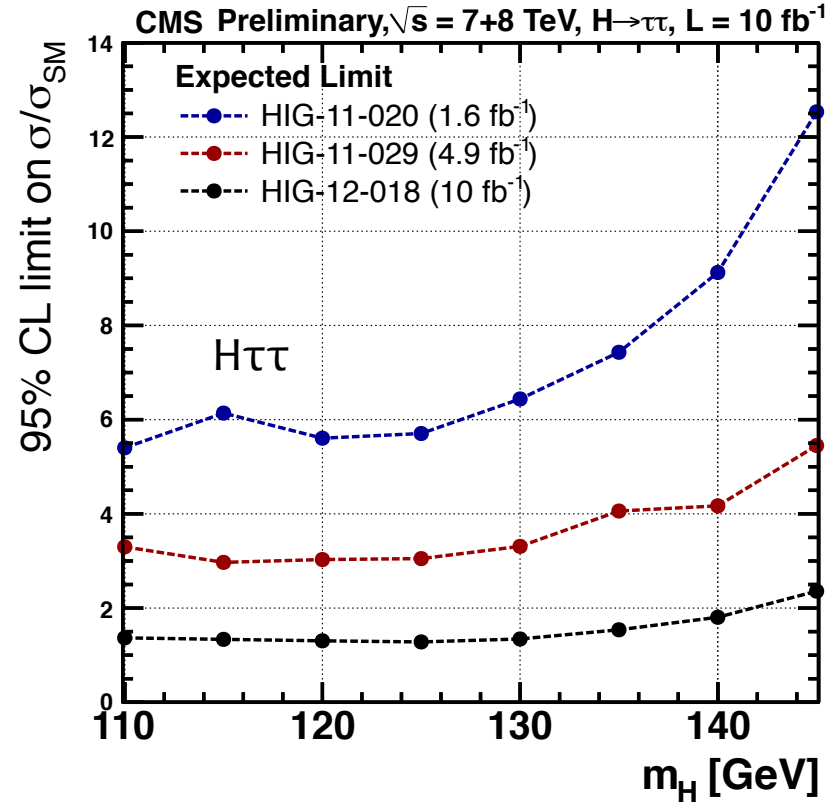
Grouping coupling in fermionic and vectorial (C_F, C_V)

- Assume SM Higgs@125
- Agreement within 95% CL
- Tensions need to be followed with more data
  - bb and ττ will tell us



H→bb and H→ττ evolution





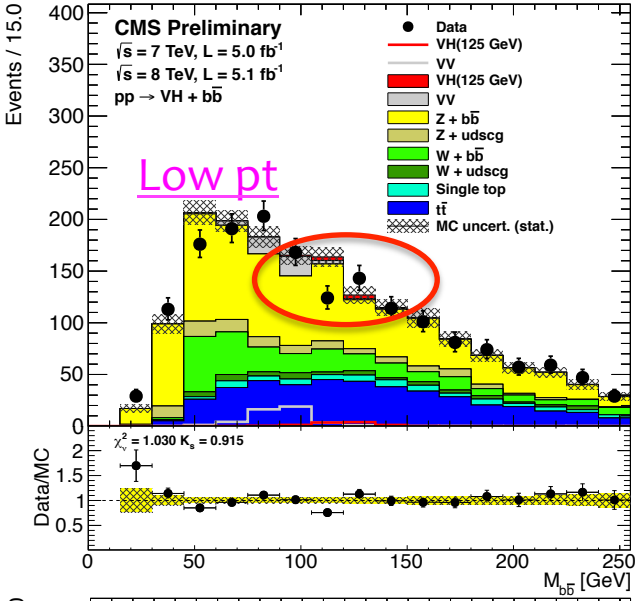
- We improve better than 1/(√∫Ldt)
- LHC will allows us to reach SM sensitivity at 20 fb^{-1}
- ttH cross section increases x1.5 from 7 to 8 TeV (could reach SM sensitivity with L ~30 fb^{-1} at 8 TeV)

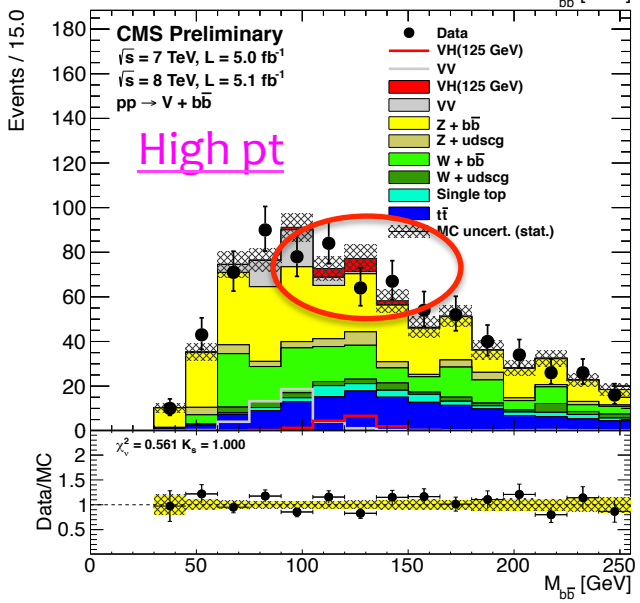
Conclusions

- CMS updated searches of SM Higgs in bb and ττ final states
- No significance excess seen in ττ and slight excess in bb
- New data are needed to conclude it was under-fluctuation or hints of new mechanism
  - We really need to reach SM sensitivity in all subchannels
  - Good news is that will very likely be able to do that with all 2012 LHC data
  - bb and ττ are vital to resolve the puzzle

BACK UP

Ok the BDT… but the mass?





mass distribution after tight cuts… we have to!

At the level to observe also VV(bb)

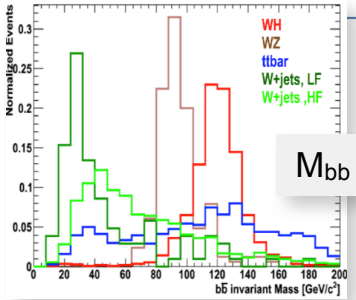
Systematics put limits here, but with more data the situation will improve

VHbb: Multi-Variate Analysis

Apply loose preselection cuts
and exploit MVA discrimination

input variables to a "Boosted Decision Tree" algorithm

Analysis on shapes of the discriminator

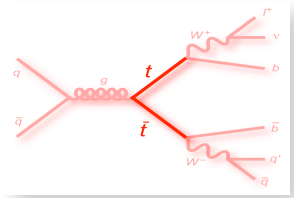


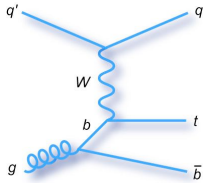
Variable
p_{Tj}: transverse momentum of each Higgs daughter
$m(jj)$: dijet invariant mass
$p_T(jj)$: dijet transverse momentum
$p_T(V)$: vector boson transverse momentum (or pfMET)
CSV_{max}: value of CSV for the b-tagged jet with largest CSV value
CSV_{min}: value of CSV for the b-tagged jet with second largest CSV value
$\Delta\phi(V,H)$: azimuthal angle between V (or E_T^{miss}) and dijet
$\lvert\Delta\eta(jj)\rvert$; difference in η between Higgs daughters
$\Delta R(j1,j2)$; distance in η–ϕ between Higgs daughters (not for $Z(\ell\ell)H$)
N_{aj}: number of additional jets ($p_T > 30$ GeV, $\lvert\eta\rvert < 4.5$)
$\Delta\phi(E_T^{miss}, jet)$: azimuthal angle between E_T^{miss} and the closest jet (only for $Z(\nu\nu)H$)
$\Delta\theta_{pull}$: color pull angle [62] (not for $Z(\ell\ell)H$)

VHbb: Multi-Variate Analysis

Apply loose preselection cuts and exploit MVA discrimination

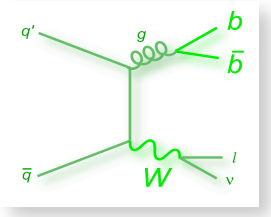
input variables to a "Boosted Decision Tree" algorithm





Reducible backgrounds QCD, V+lights, ttbar, single top

Analysis on shapes of the discriminator



Variable	W($\ell\nu$)H	Z($\ell\ell$)H	Z($\nu\nu$)H
$m_{\ell\ell}$	–	$75 < m_{\ell\ell} < 105$	–
$p_T(j_1)$	> 30	> 20	> 80
$p_T(j_2)$	> 30	> 20	> 20
$p_T(jj)$	> 120	–	$120 - 160\,(> 160)$
$m(jj)$	< 250	$80 < m(jj) < 150$ (–)	< 250
$p_T(V)$	$120 - 170\,(> 170)$	$50 - 100\,(> 100)$	–
CSV_{max}	> 0.40	$0.50\,(0.244)$	> 0.50
CSV_{min}	> 0.40	0.244	> 0.50
N_{al}	$= 0$	–	$= 0$
$\Delta\phi(E_T^{miss}, jet)$	–	–	> 0.5
E_T^{miss}	> 35 (elec)	–	$120 - 160\,(> 160)$
BDT	full distribution	full distribution	full distribution

Less reducible V+bb, ZZ(bb), WZ(bb)

VHbb: Bkg Control Regions

Definition of Control Regions (CR) crucial element of the analysis

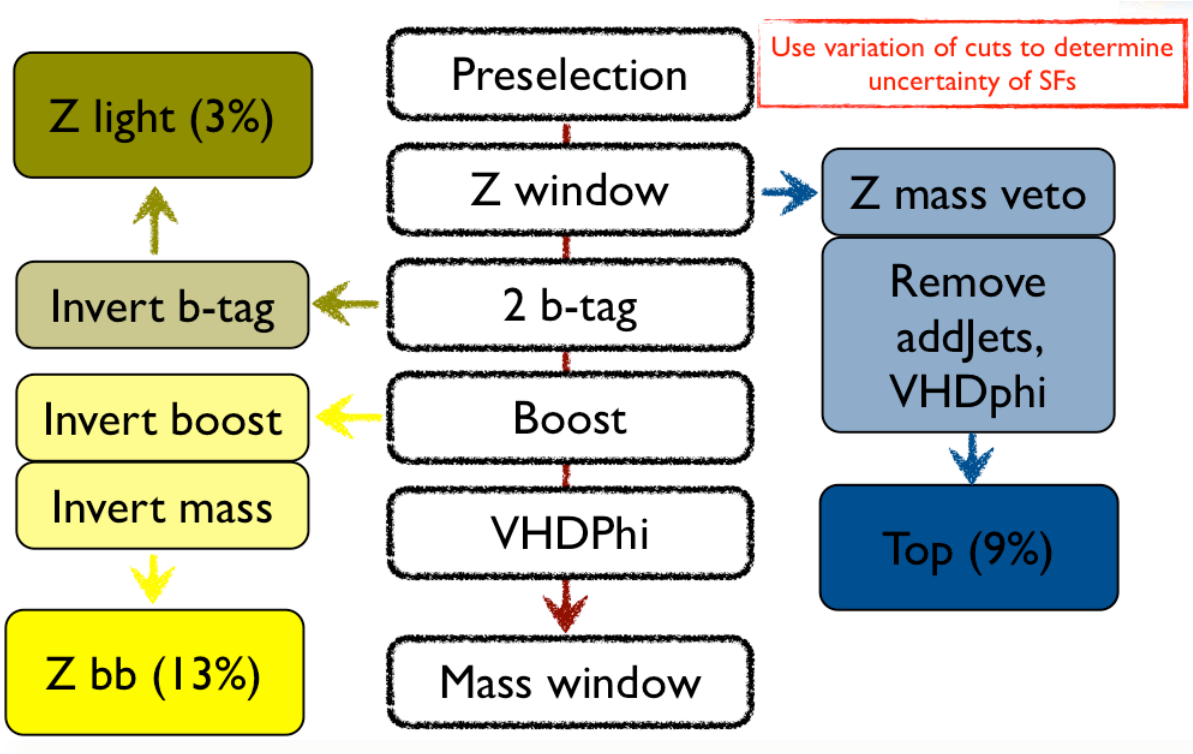
Define several CR enriched in different background components:

Control regions cuts as close as possible to the signal region

Renormalize MC simulation yields and extrapolate to signal region

Account for associated stat. and syst. uncertainties

Example: Zee control region definition



(%)= uncertainty

H->bb systematics

Source	Range
Luminosity	2.2%
Lepton efficiency and trigger (per lepton)	3%
$Z(\nu\nu)H$ triggers	2%
Jet energy scale	2–3%
Jet energy resolution	3–6%
Missing transverse energy	3%
b-tagging	3–15%
Signal cross section (scale and PDF)	4%
Signal cross section (p_T boost, EWK/QCD)	5–10% / 10%
Signal Monte Carlo statistics	1-5%
Backgrounds (data estimate)	$\approx$ 10%
Diboson and single-top (simulation estimate)	30%

4.4% at 8 TeV

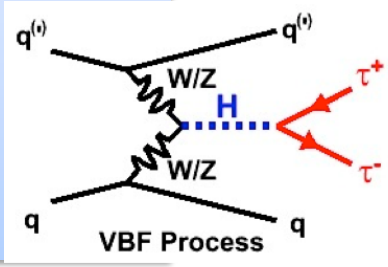
M. Ciccolini et al., "Strong and electroweak corrections to the production of Higgs+2jets via weak interactions at the LHC", Phys. Rev. Lett. 99 (2007) 161803, doi:10.1103/PhysRevLett.99.161803, arXiv:0707.0381.

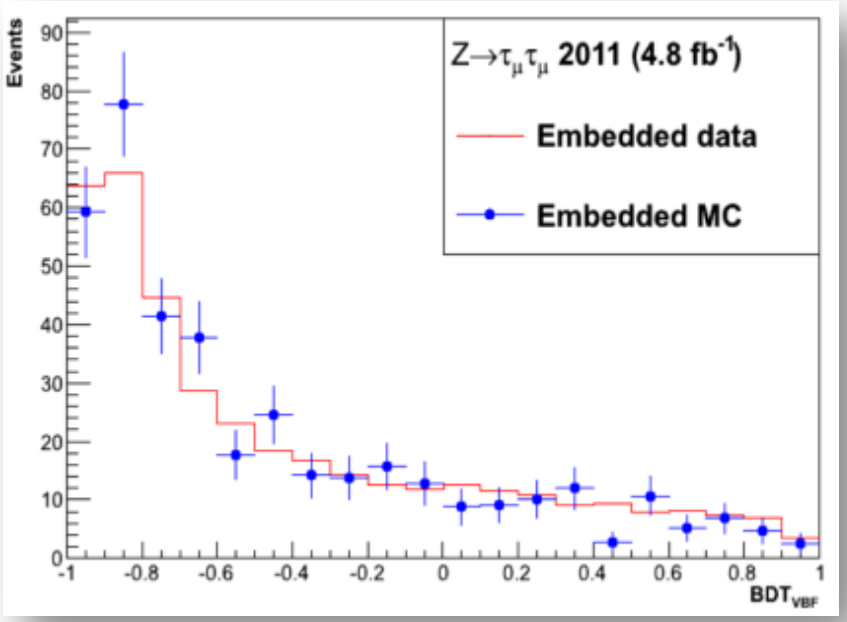
M. Ciccolini, A. Denner, and S. Dittmaier, "Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC", Phys. Rev. D 77 (2008) 013002, doi:10.1103/PhysRevD.77.013002, arXiv:0710.4749.

A. Denner, S. Dittmaier, S. Kallweit et al., "Electroweak corrections to Higgs-strahlung off W/Z bosons at the Tevatron and the LHC with HAWK", (2011). arXiv:1112.5142.

Hττ: VBF MVA

- Multivariate discriminator
  - 8 input variables
  - Replaces traditional cut-based VBF selection using m_{jj} and $\Delta\eta_{jj}$
  - Working point optimized for best sensitivity
    - 15-20% improvement over previous cut-based selection
- MVA output validated on Z→μμ data
- Events with 2 jets $p_T > 30$ GeV



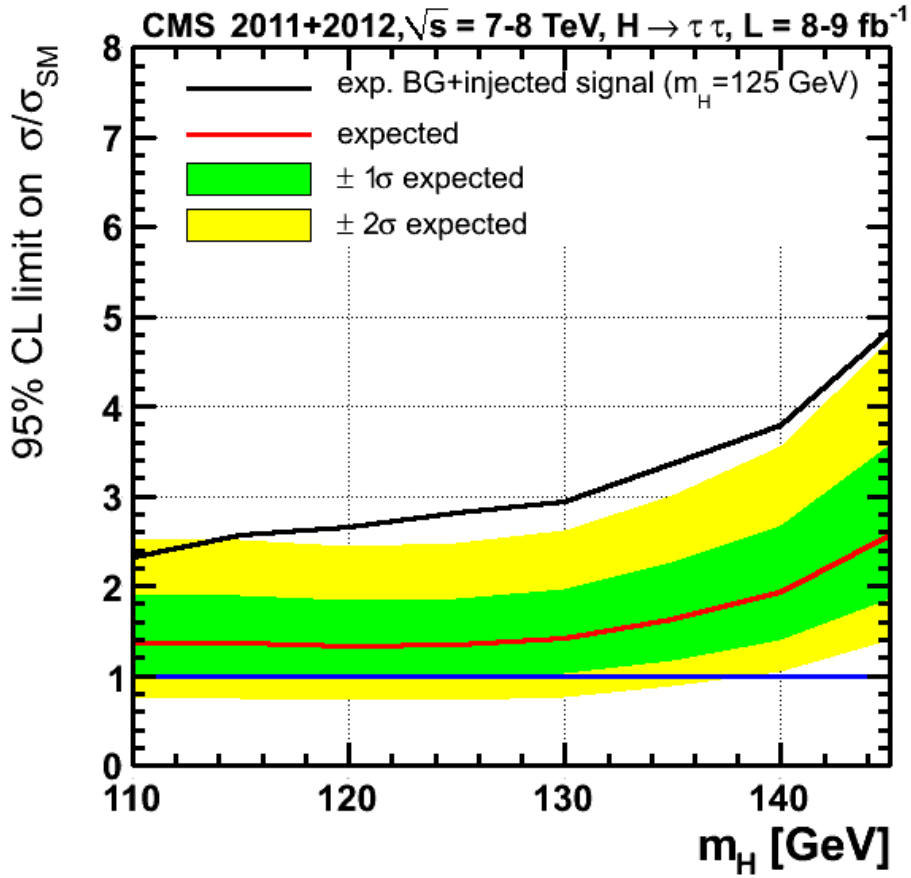


MVA input variables

1. m(jj)
2. Δη(jj)
3. Δφ(jj)
4. Δφ(ττ,jj)
5. p_T(jj)
6. p_T(ττ) (including MET)
7. Visible p_T(ττ)
8. Δη(ττ,jet)

Signal Injection

- Signal injected for a 125 GeV Higgs Boson
- Median result from 100 toy experiments

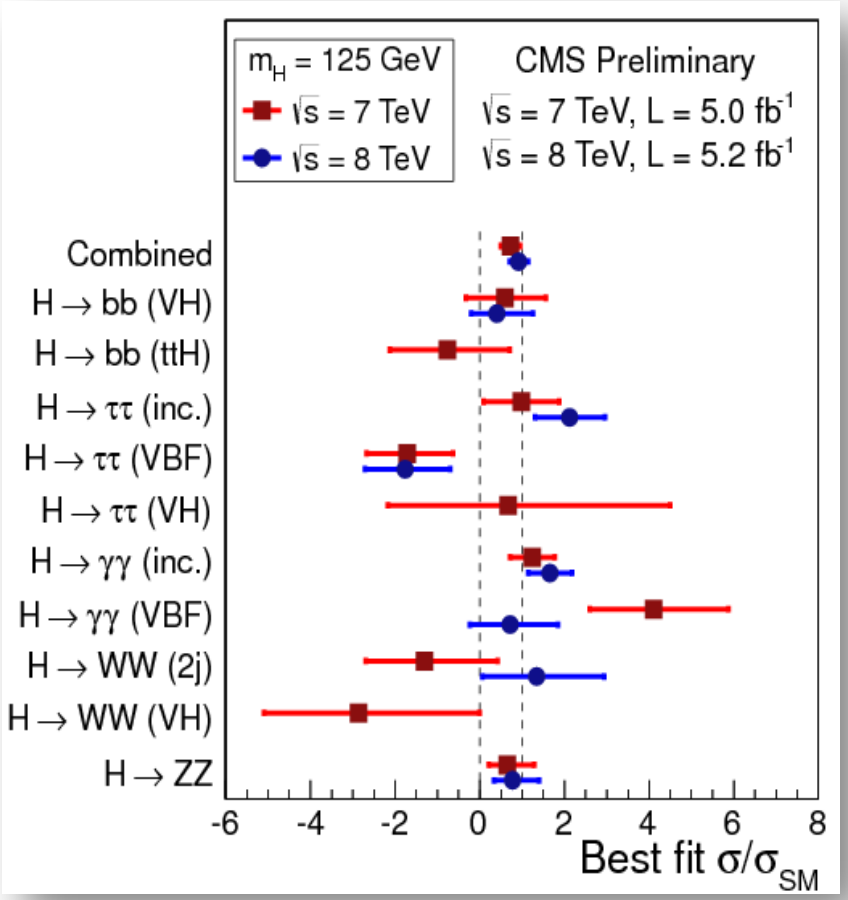


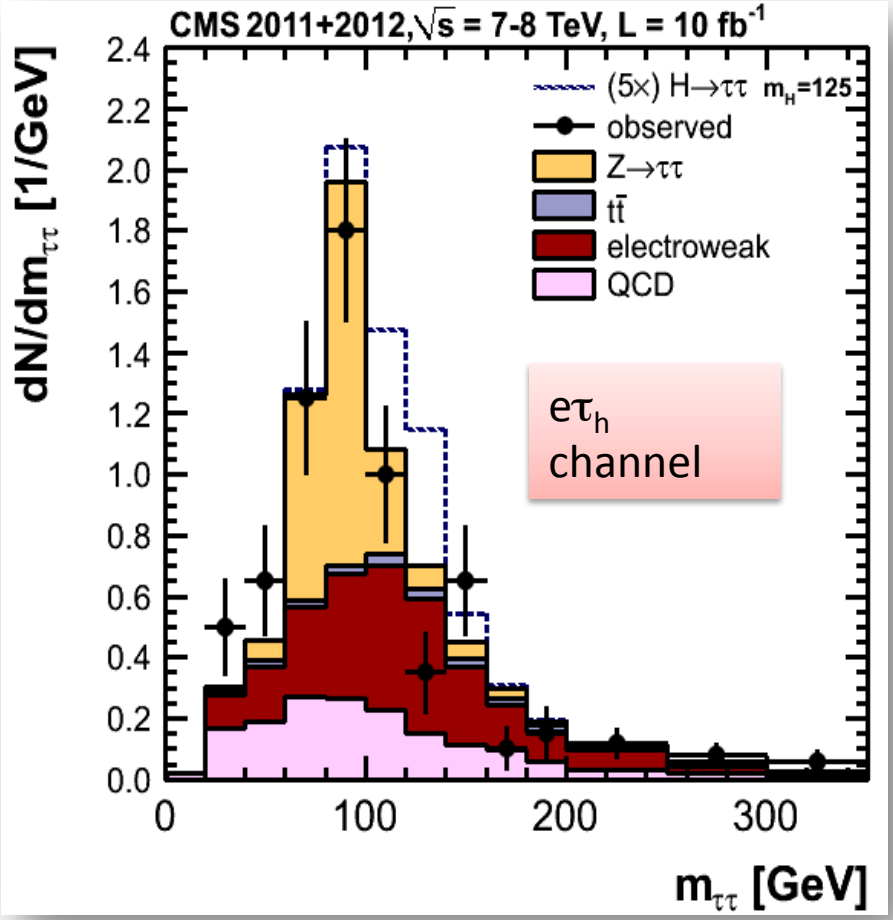
H->ττ

Experimental Uncertainties		Propagation into Limit Calculation		
Uncertainty	Uncert.	*0-Jet*	*Boost*	*VBF*
Electron ID & Trigger (*)	$\pm 2\%$	$\pm 2\%$	$\pm 2\%$	$\pm 2\%$
Muon ID & Trigger (*)	$\pm 2\%$	$\pm 2\%$	$\pm 2\%$	$\pm 2\%$
Tau ID & Trigger (*)	$\pm 7\%$	$\pm 7\%$	$\pm 7\%$	$\pm 7\%$
JES (Norm.) (*)	$\pm 2.5 - 5\%$	$\mp 1\%$	$\pm 5\%$	$\pm 10\%$
b-Tag Efficiency (*)	$\pm 10\%$	$\mp 1\%$	$\mp 2\%$	$\mp 2\%$
Mis-Tagging (*)	$\pm 30\%$	$\mp 1\%$	$\mp 1\%$	$\mp 1\%$
Norm. $Z \to \tau\tau$	$\pm 3\%$	$\pm 3\%$	$\pm 5\%$	$\pm 13\%$
Norm. $t\bar{t}$ (*)	$\pm 10 - 30\%$	$\pm 10\%$	$\pm 12\%$	$\pm 30\%$
Norm EWK	$\pm 30\%$	$\pm 30\%$	$\pm 15 - 30\%$	$\pm 30 - 100\%$
Norm Fakes	$\pm 10 - 30\%$	$\pm 10\%$	$\pm 10\%$	$\pm 30\%$
Lumi (Signal & EWK)	$\pm 2.2(5)\%$	$\pm 2.2(5)\%$	$\pm 2.2(5)\%$	$\pm 2.2(5)\%$
Norm. $W + jets$	$\pm 10 - 30\%$	$\pm 10\%$	$\pm 10 - 30\%$	$\pm 30\%$
Norm. Z: l fakes τ_h	$\pm 20 - 100\%$	$\pm 20 - 30\%$	$\pm 20 - 100\%$	$\pm 30\%$
Norm. Z: jet fakes τ_h	$\pm 20\%$	$\pm 20\%$	$\pm 20\%$	$\pm 30\%$

Theory Uncertainties (SM)		Propagation into Limit Calculation		
Uncertainty	Uncert.	*0-Jet*	*Boost*	*VBF*
PDF (*)	-	$\pm 2 - 8\%$	$\pm 2 - 8\%$	$\pm 2 - 8\%$
$\mu_r/\mu_f(gg \to H)$ (*)	-	$\pm 8\%$	$\pm 10\%$	$\pm 30\%$
$\mu_r/\mu_f(qq \to H)$ (*)	-	$\pm 3.5\%$	$\pm 4\%$	$\pm 10\%$
$\mu_r/\mu_f(qq \to VH)$ (*)	-	$\pm 4\%$	$\pm 4\%$	$\pm 4\%$
UE & PS (*)	-	$\mp 4\%$	$\pm 4\%$	$\pm 4\%$

VBF ττ low μ values

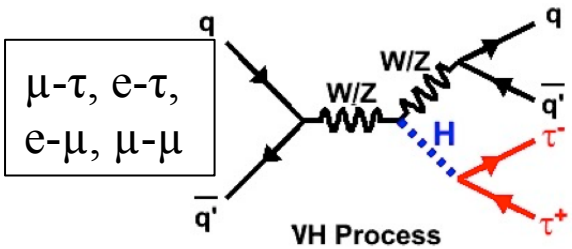




- Low values for best fit σ/σSM are found for the VBF ττ channel in 2011 and 2012
  - Traced to deficit of events in the 120-140 GeV bin in data as seen above
    - About 1σ low in 8 TeV data and a bit lower in 7 TeV
  - NB that the data generally appears to agree well with the expectation in the plot overall

$\tau\tau$ Associated Production ($W/Z \rightarrow jj$)

- New category for 2012
  - 2 Jets p_T > 30 GeV,
  - 70 GeV < m_{jj} < 120 GeV
  - $p_T(jj)$ > 150 GeV
  - no b-Tagged Jets
  - Not in VBF Category
- 25% Systematic for ggH falling in this category, 8% pdf uncertainty for VH
- Adds 5-10% to overall limit compared to 2011 analysis.

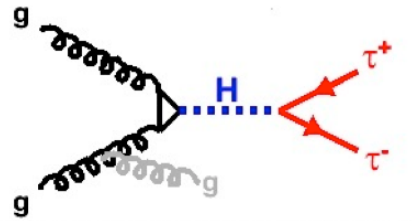


$\tau\tau$ Associated Production (W/Z→leptons)

- publication on 2011 data
  - WH/ZH analyses
  - Add 10-15% to 2011 H→τ τ Analysis
- New for ICHEP:
  - new object id
  - Add 2012 data available up to cutoff
  - Looking to add eeτ
  - Full mass reco for ZH

ττ 1 Jet Category

μ-τ, e-τ, e-μ, μ-μ



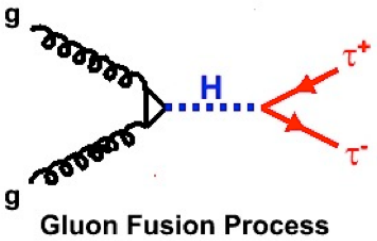
Boosted Gluon Fusion Process

- The boosted category from 2012 gave excellent S/B
  - large systematic on the Higgs p_T (25%)
- Main advantage from improved mass resolution
  - Can be achieved with lower jet p_T
- Definition
  - ≥ 1 Jet with p_T > 30 GeV
  - No B-Tagged Jet
  - Not in VBF or VH Category
- 30 GeV 1 Jet category gives ~10% better limit compared to 2011 categories

Other Categories

0-Jet

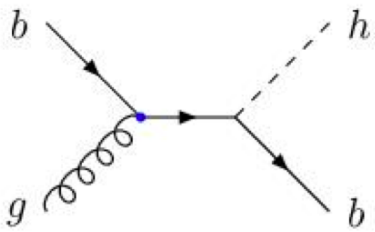
μ-τ, e-τ, e-μ, μ-μ



- Events not passing any other category
- Large signal contribution
- more sensitive then in 2011 due to improved mass resolution

B-Jet

μ-τ, e-τ, e-μ, μ-μ

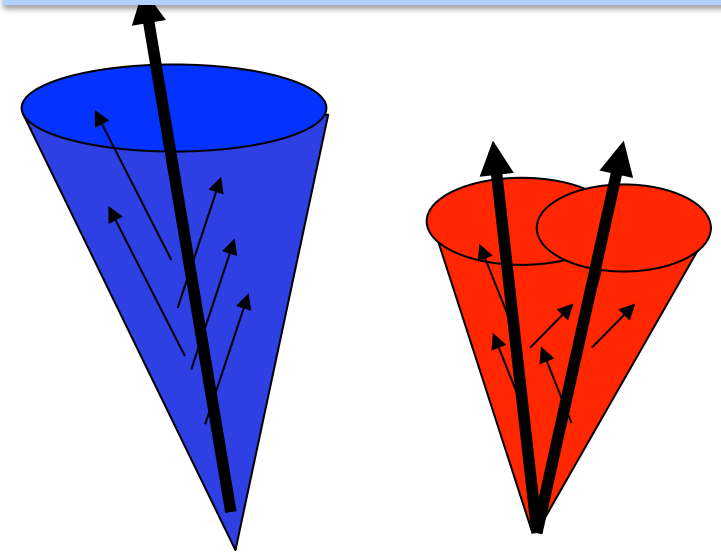


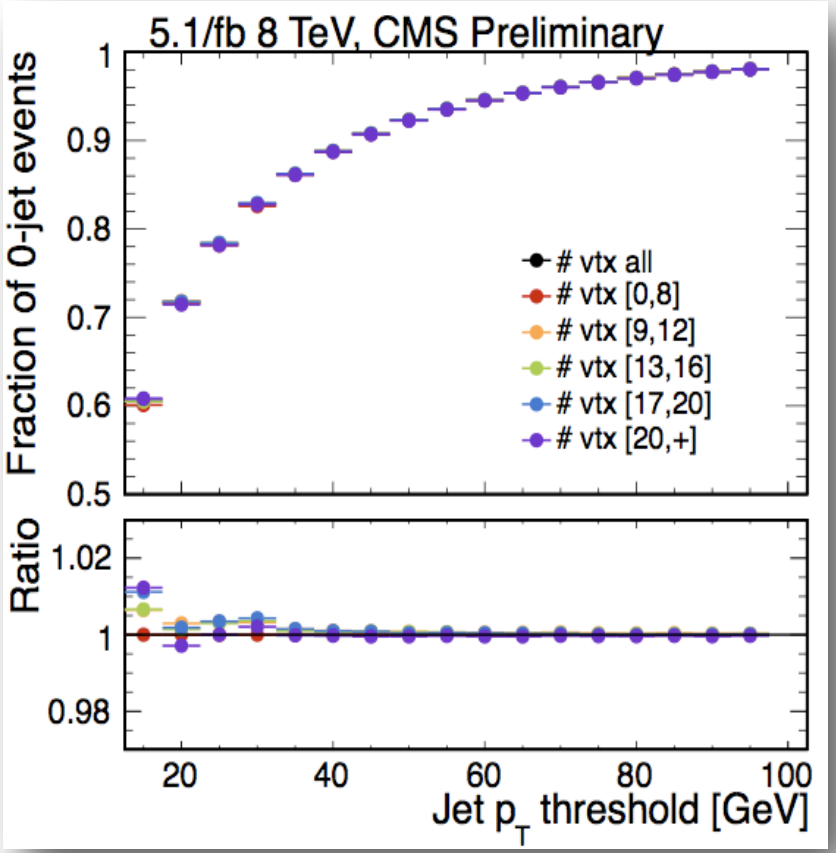
- Traditional MSSM category
- ≥ 1 b tagged jet with p_T > 20 GeV

Joshua Swanson University of Wisconsin - Madison

Jet Identification

- Jet reconstruction
  - Reconstruction with particle flow objects





- Pileup jets structure differs wrt regular jets:
  - Pileup jets originate from several overlapping jets which merge together
  - Likelihood grows rapidly with high pileup
- Discriminant exploits shape and tracking variables
  - discrimination both inside and outside tracker acceptance