The Higgs boson in the fermionic decay modes with the CMS experiment at LHC

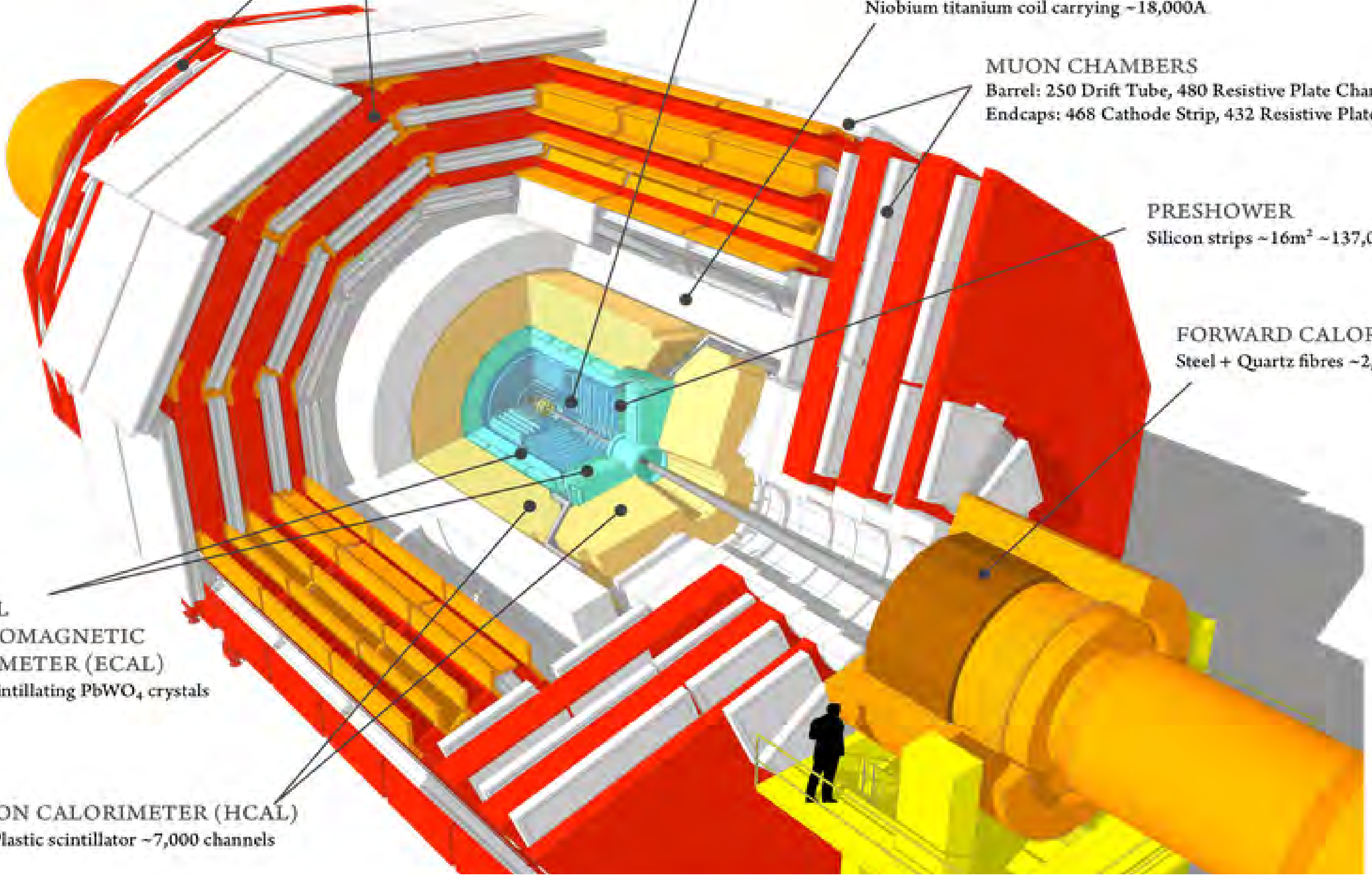
Andrea Rizzi, INFN e Universita' di Pisa

January 29th 2014, DESY

Outline

- Higgs search overview
- The Higgs to BB modes:
  - VH
  - VBF
  - ttH
- The Higgs to τ pairs
- MSSM and fermionic modes
- Some ideas for next runs

The CMS experiment



CMS DETECTOR

Total weight : 14,000 tonnes
Overall diameter : 15.0 m
Overall length : 28.7 m
Magnetic field : 3.8 T

STEEL RETURN YOKE
12,500 tonnes

SILICON TRACKERS
Pixel (100x150 μm) ~16m^2 ~66M channels
Microstrips (80x180 μm) ~200m^2 ~9.6M channels

SUPERCONDUCTING SOLENOID
Niobium titanium coil carrying ~18,000A

MUON CHAMBERS
Barrel: 250 Drift Tube, 480 Resistive Plate Chambers
Endcaps: 468 Cathode Strip, 432 Resistive Plate Chambers

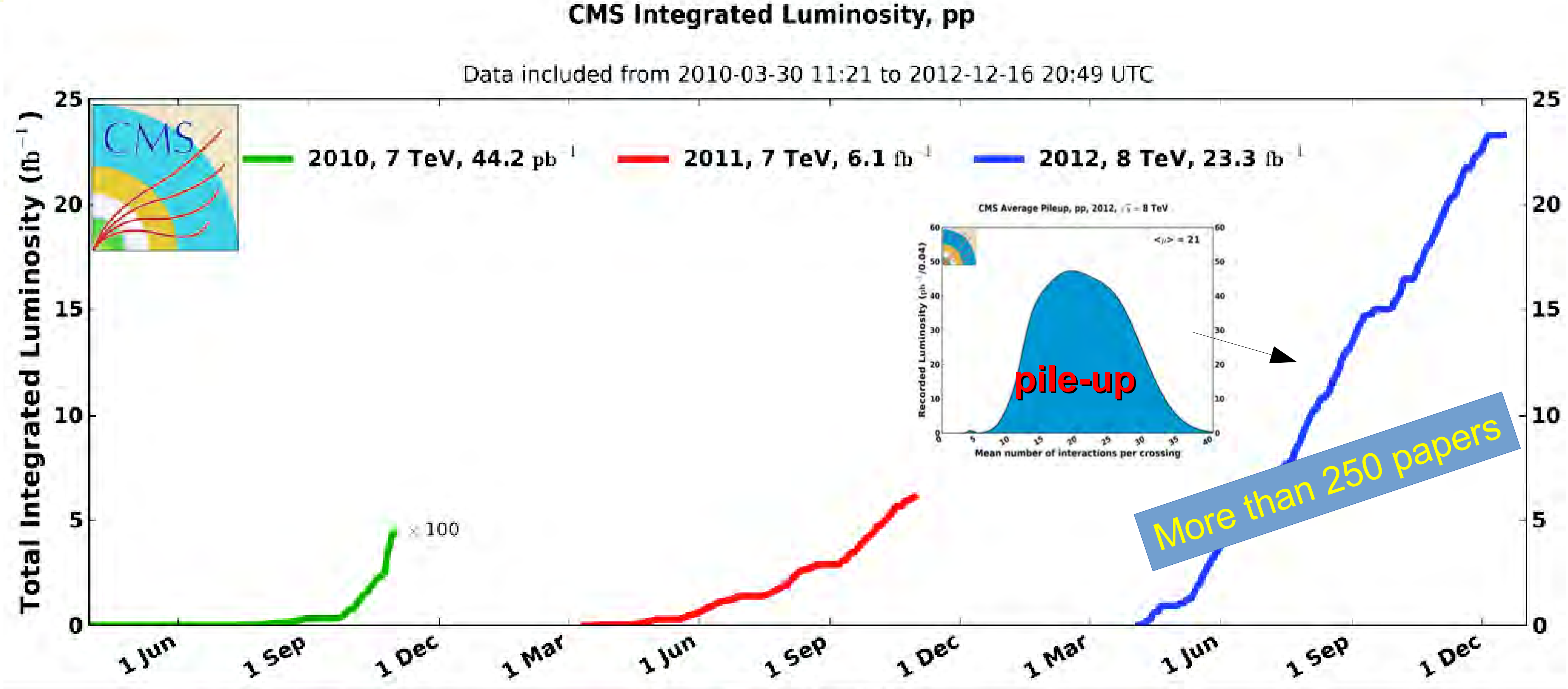
PRESHOWER
Silicon strips ~16m^2 ~137,000 channels

FORWARD CALORIMETER
Steel + Quartz fibres ~2,000 Channels

CRYSTAL ELECTROMAGNETIC CALORIMETER (ECAL)
~76,000 scintillating $PbWO_4$ crystals

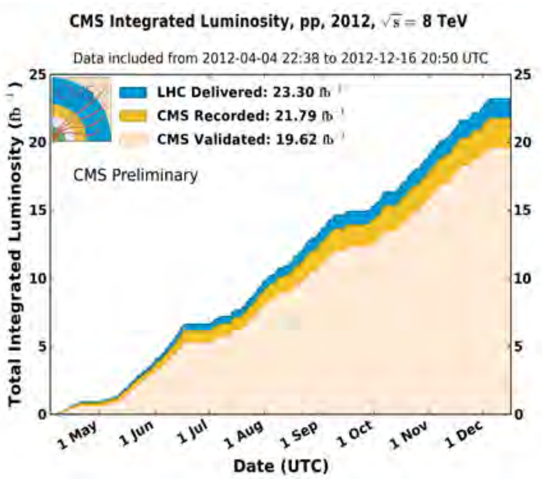
HADRON CALORIMETER (HCAL)
Brass + Plastic scintillator ~7,000 channels

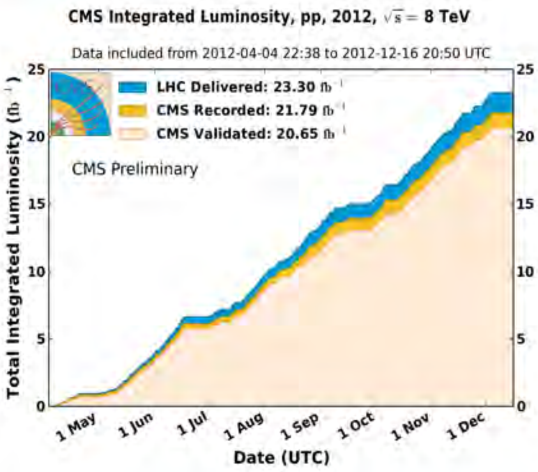
LHC 2011 & 2012



More than 250 papers

- <pileup> of 21 events
- Ddata taking efficiency ~ 93%
- Efficiency including validation
  - ~84% prompt reconstruction
  - ~89% reprocessing



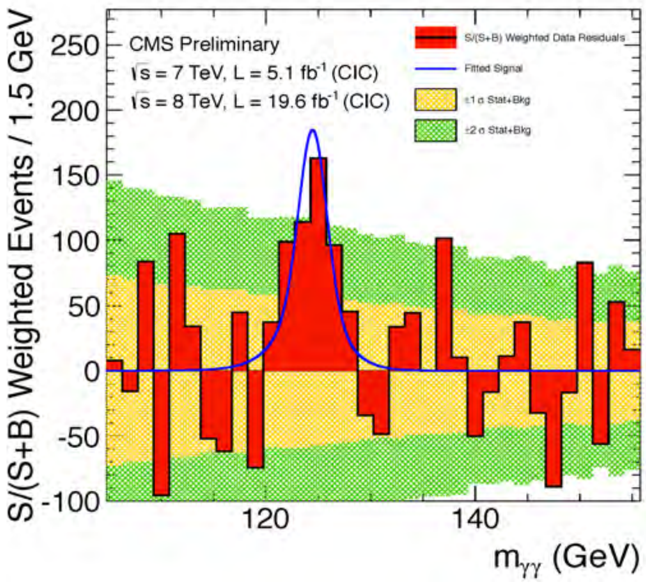


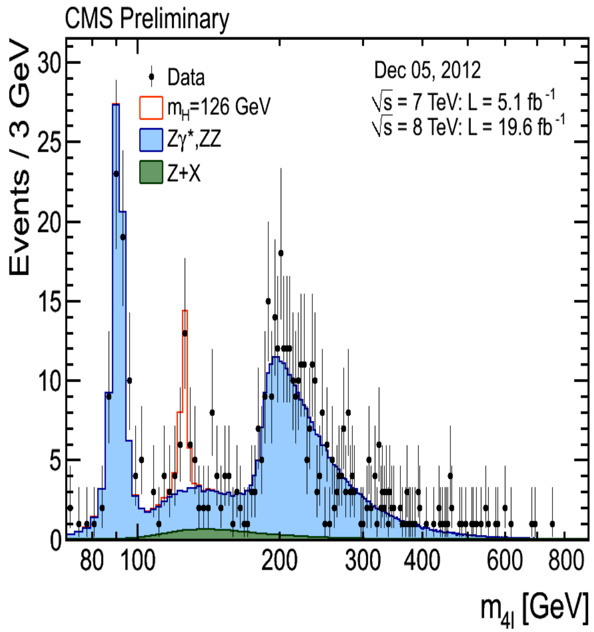
Higgs overview

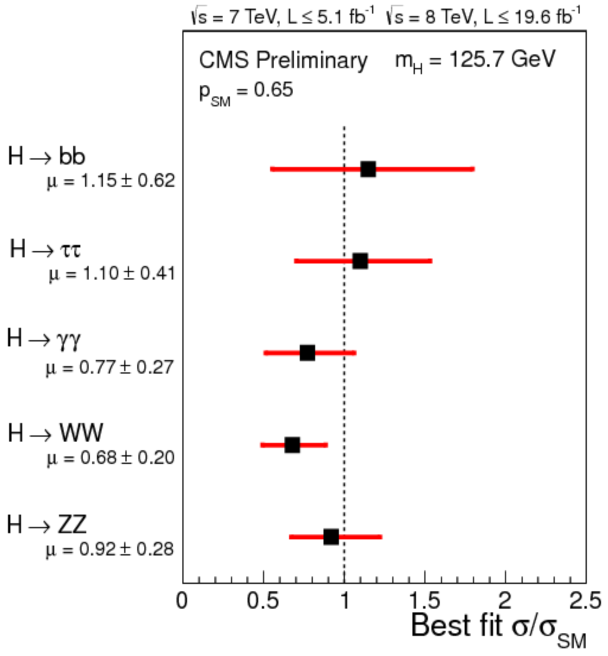
- Higgs signal well visible in gammagamma and ZZ
- Mass:

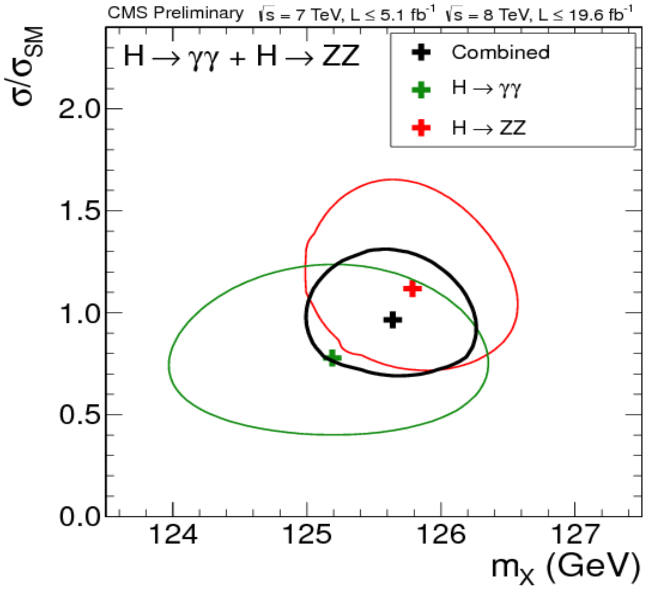
  125.7± 0.3(stat)± 0.3(syst)

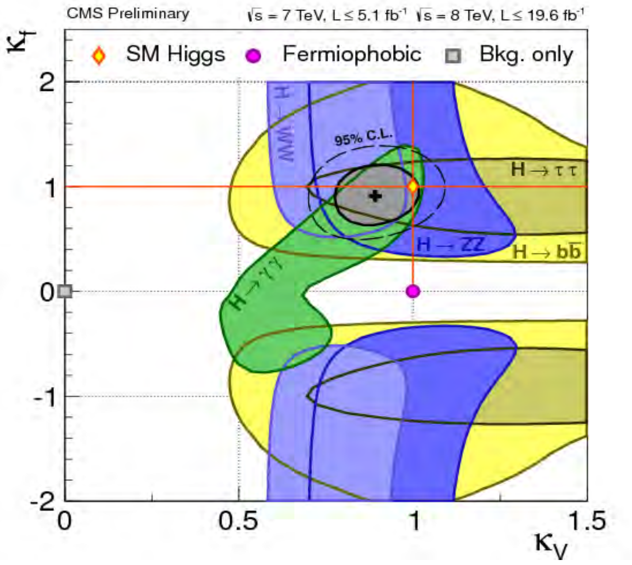
- How about fermions?



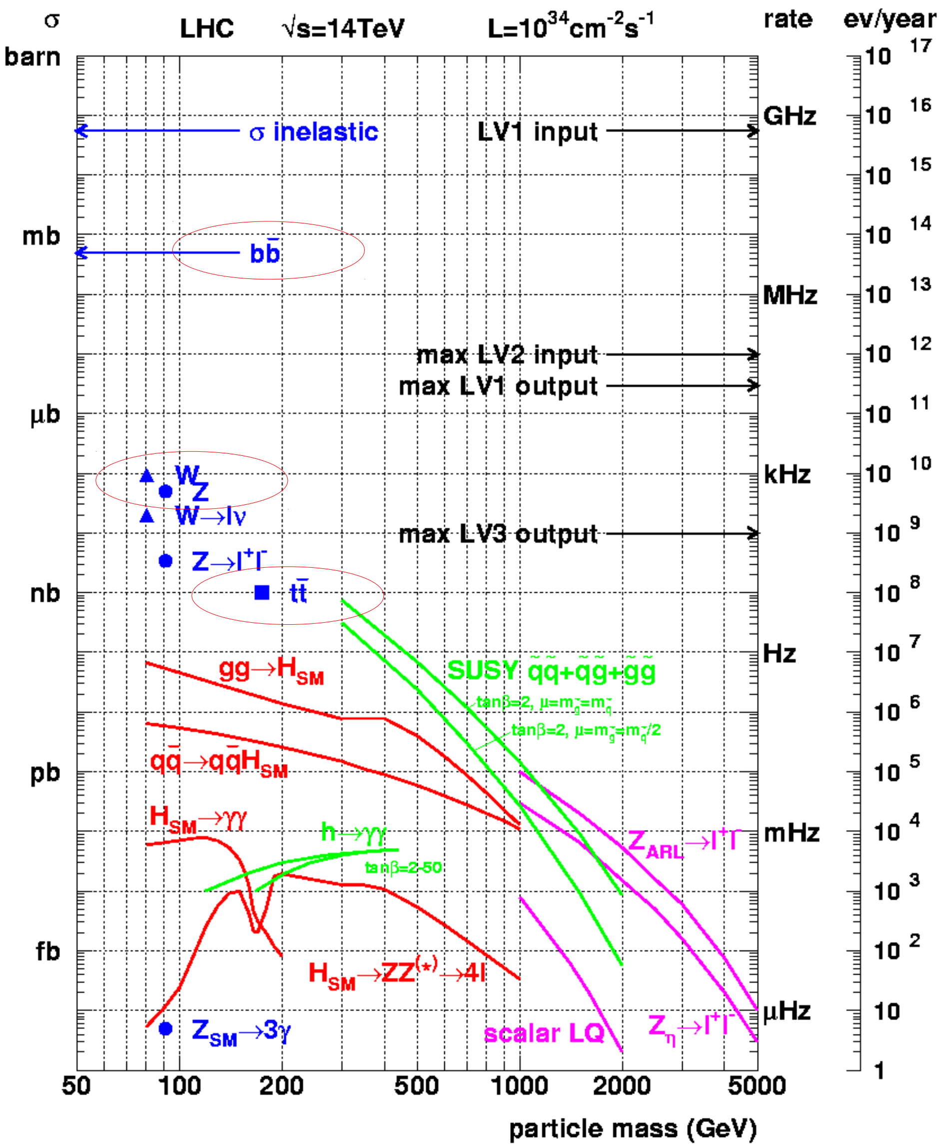


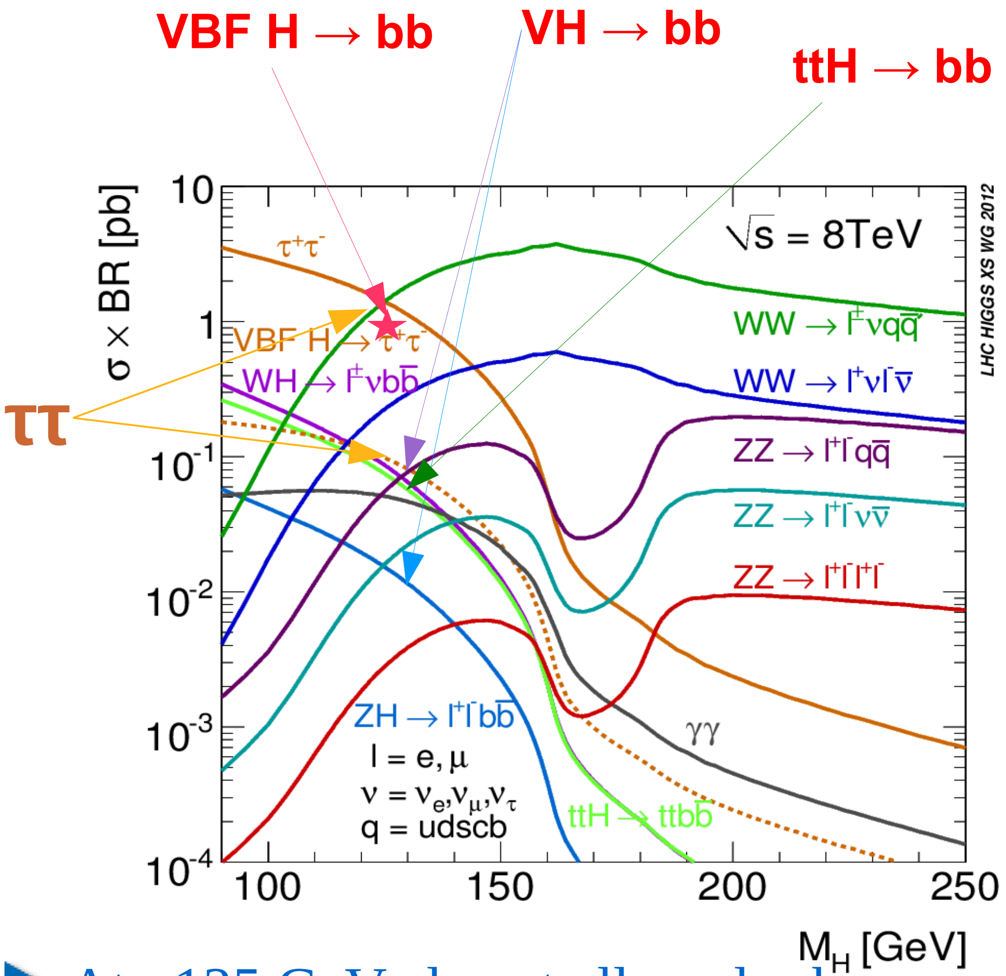






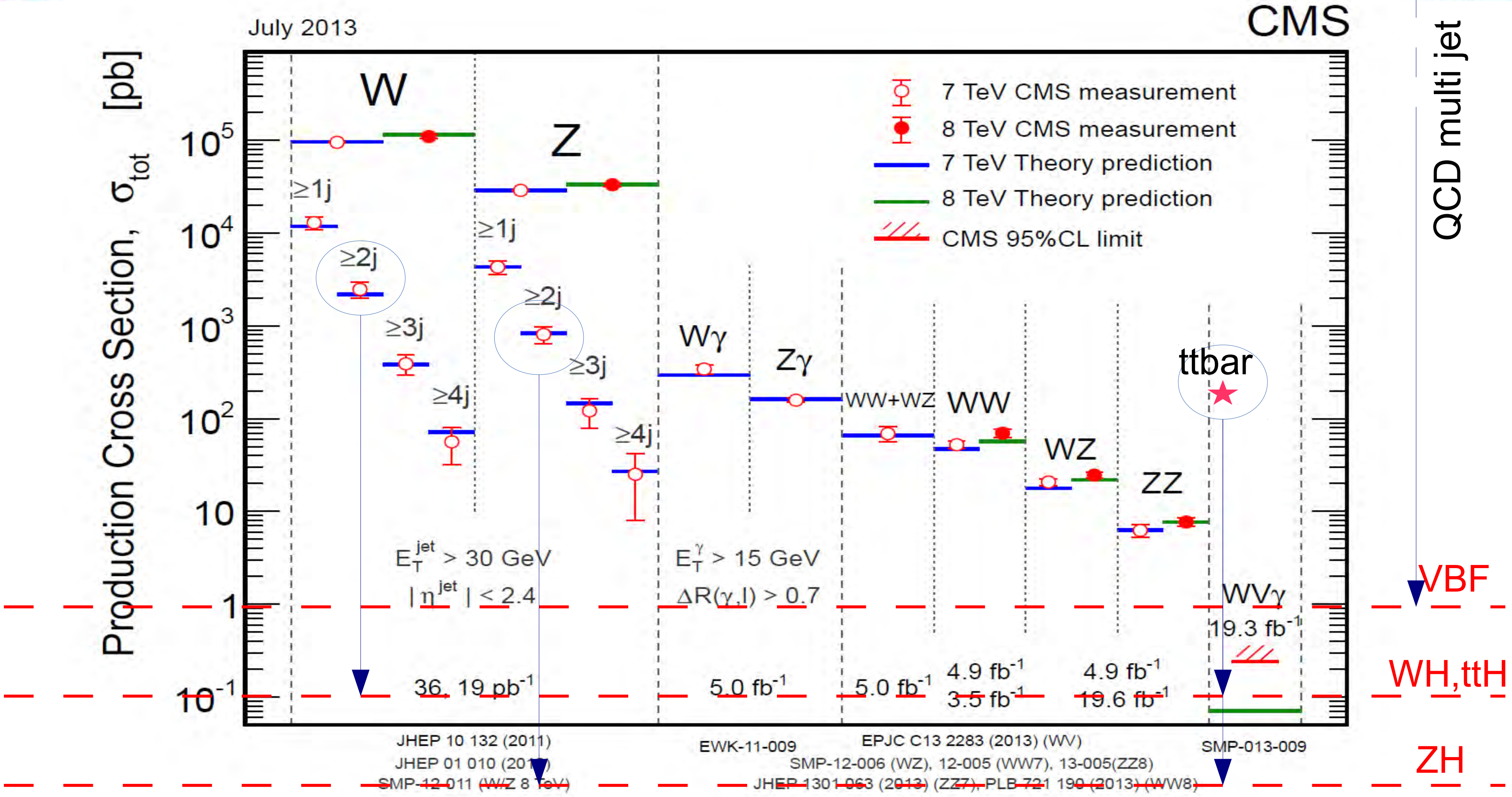
The cross sections





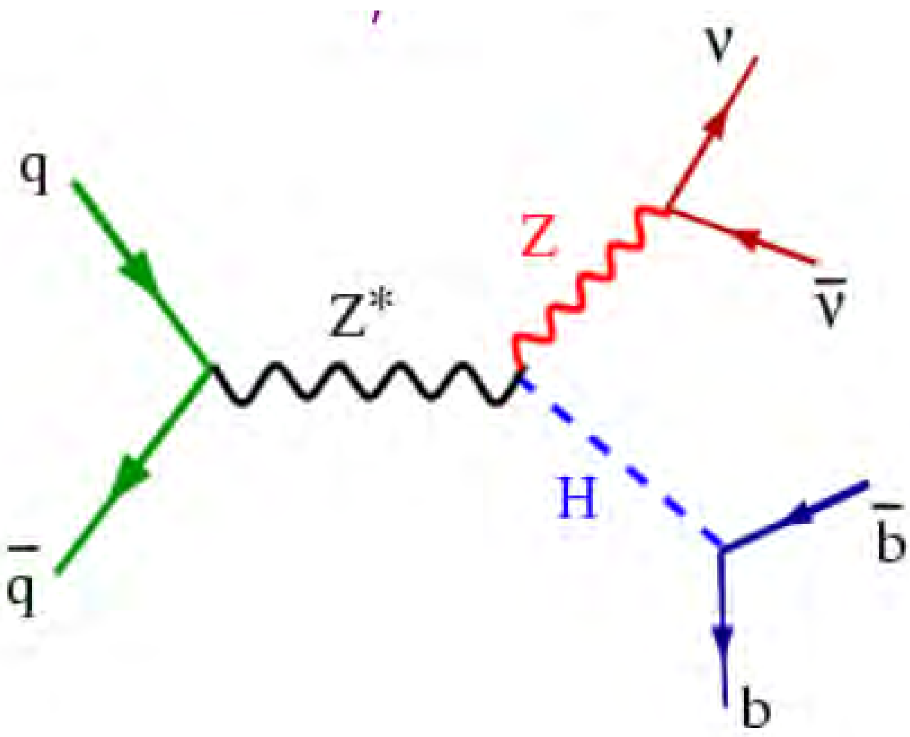
- At ~125 GeV almost all modes have a sizable BR, giving us a chance to make a measurement
- Despite the large BR, H → bb is studied only in associate productions that significantly reduce the σ x BR

Summary of SM measurements



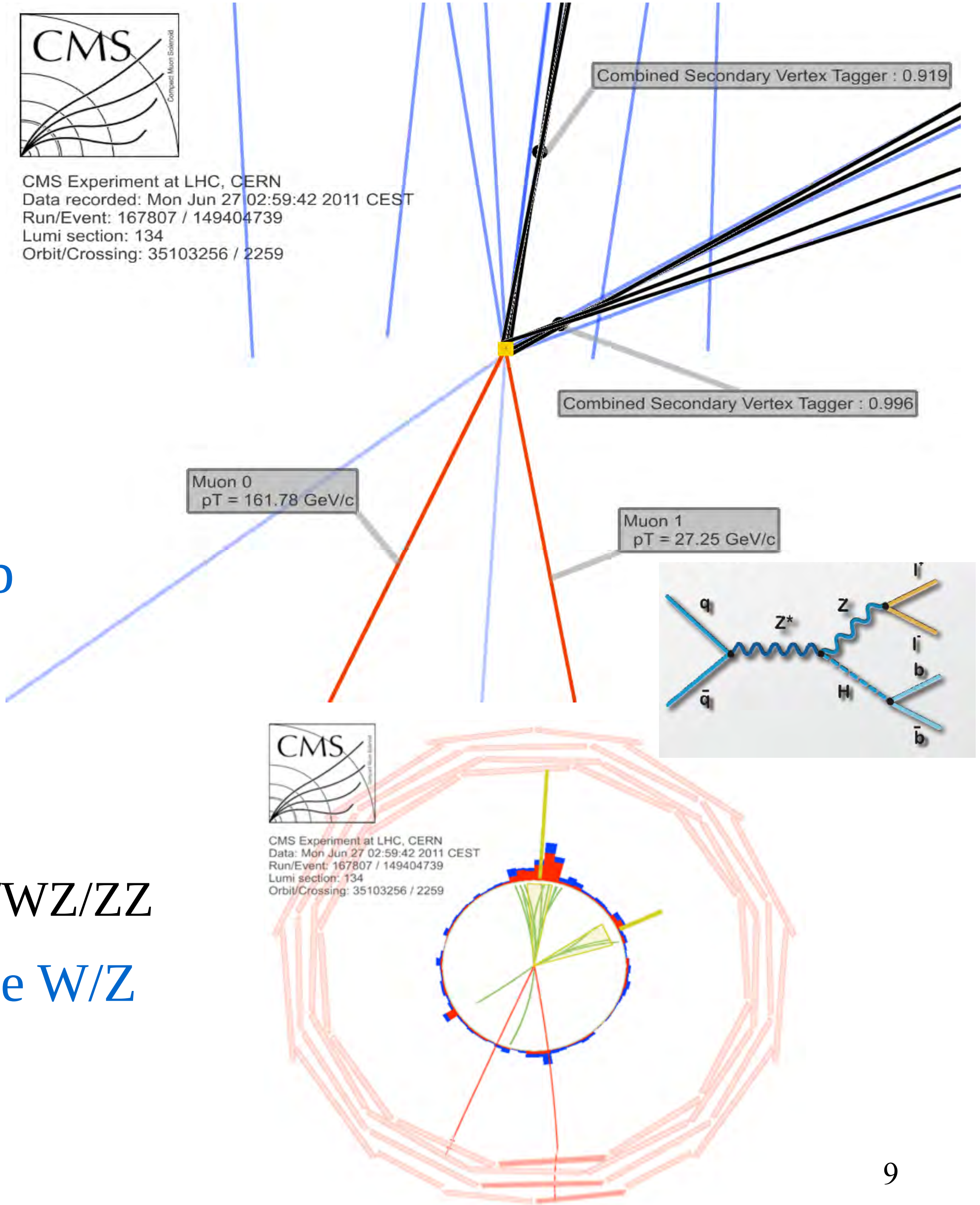
- Backgrounds production cross sections are about 3-4 orders of magnitude above signal (before any analysis cuts)
- Resonance peak in bb and ττ is not "narrow" (as for photon and Z)

VH, H->bb



VH → ll,lν,νν + *bb*

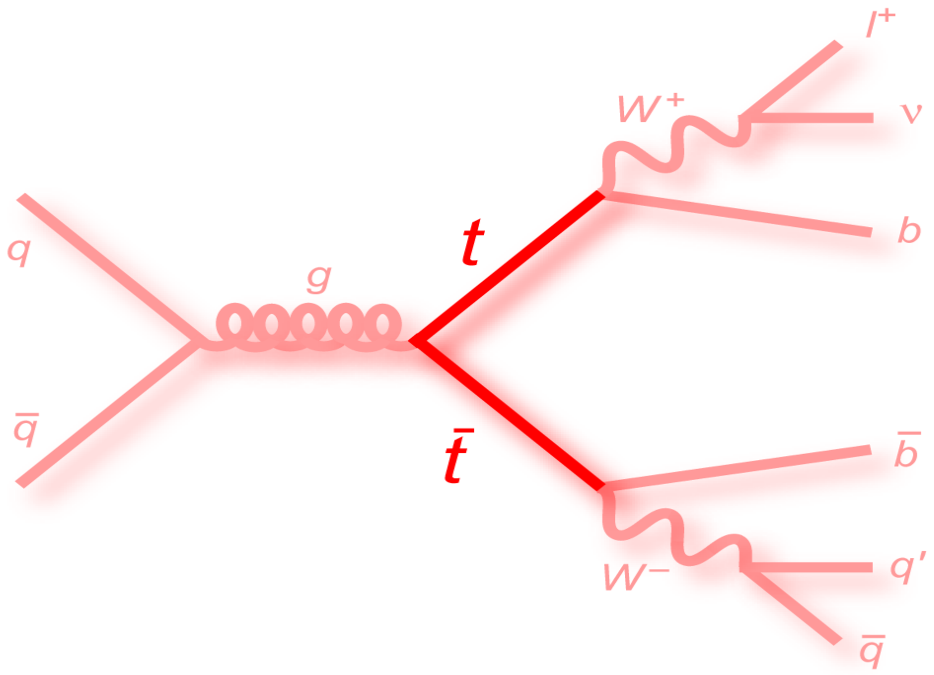
- Associated production of Higgs to a vector boson
  - Several modes considered:
    - W->lν (e, μ, τ)
    - Z-> νν
    - Z->ll (electrons or muons)
- Decay of the Higgs boson in bb
  - Use b-tagging to identify the jets coming from the Higgs decay
- Main backgrounds
  - V+(b)jets, ttbar, single top, WW/WZ/ZZ
- Trigger with the lepton(s) from the W/Z and/or MET+jets

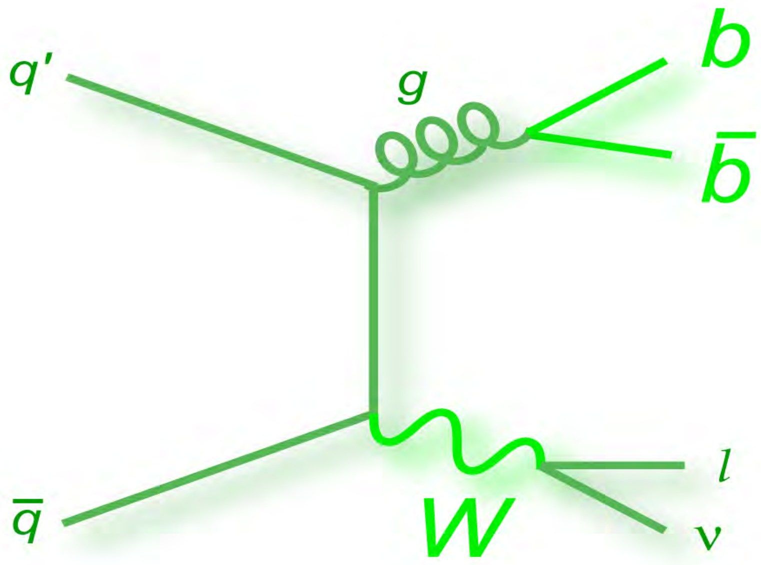


Backgrounds

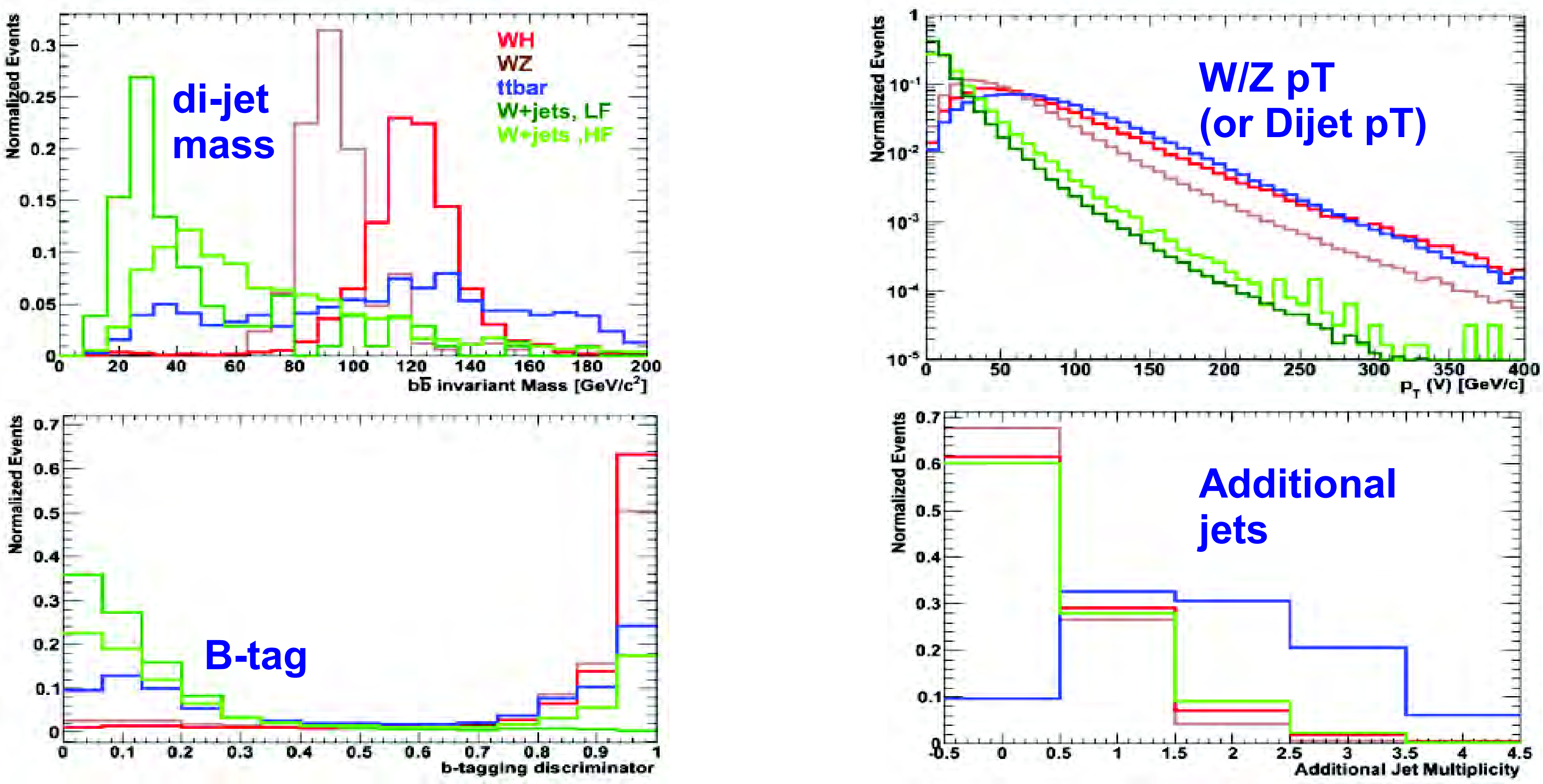
Reducible backgrounds
QCD, V+udscg (“light” jets)
ttbar and single top
=> reduced with b-tag, jet counting, additional leptons, lepton isolation

Less reducible backgrounds
V+bb
ZZ(bb), W(lν)Z(bb)
=> bb mass is the only handle





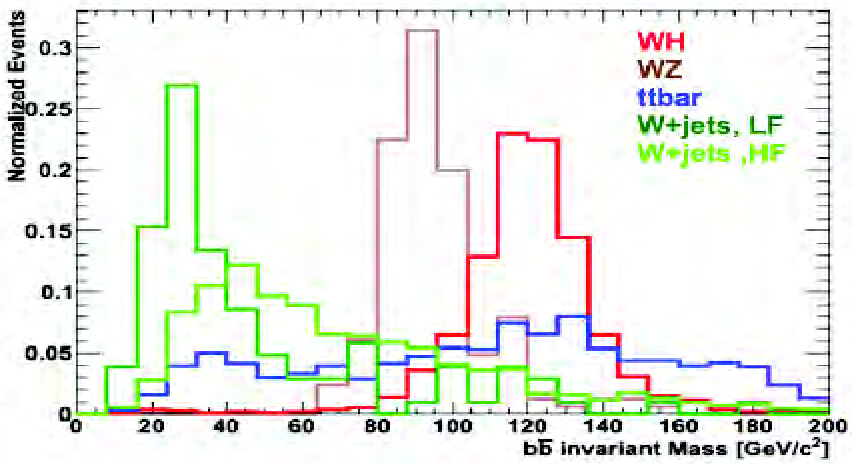
The main handles

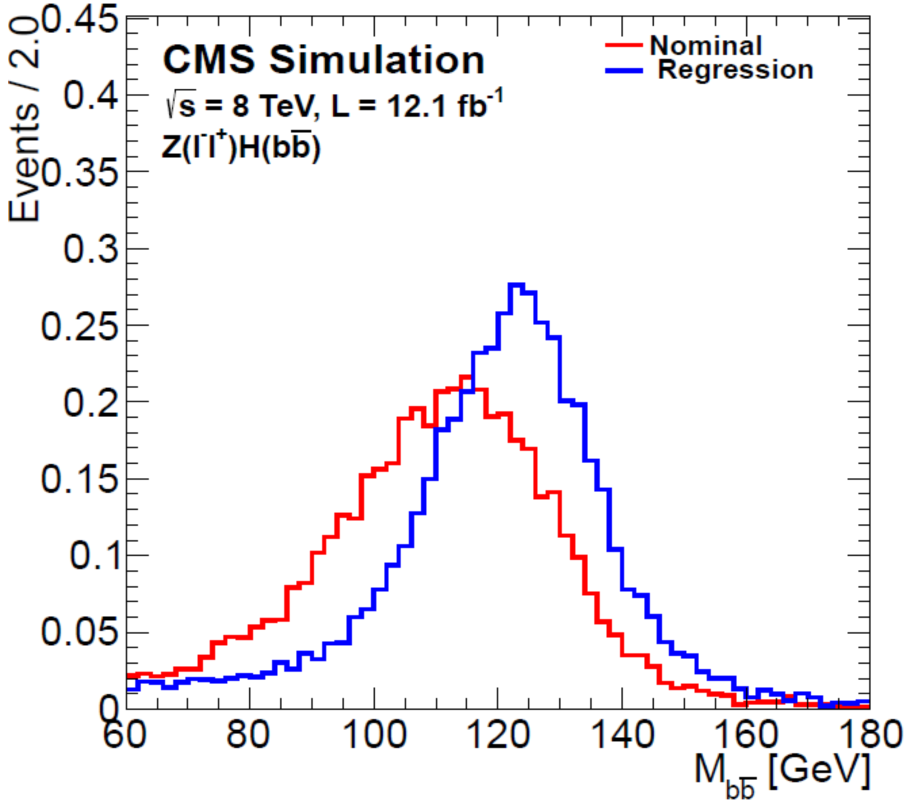


- Other important observables used in the analysis
  - MET, MET significance, MinDeltaPhi (Jet, MET)
  - DeltaPhi(W/Z,H)

Jet energy regression

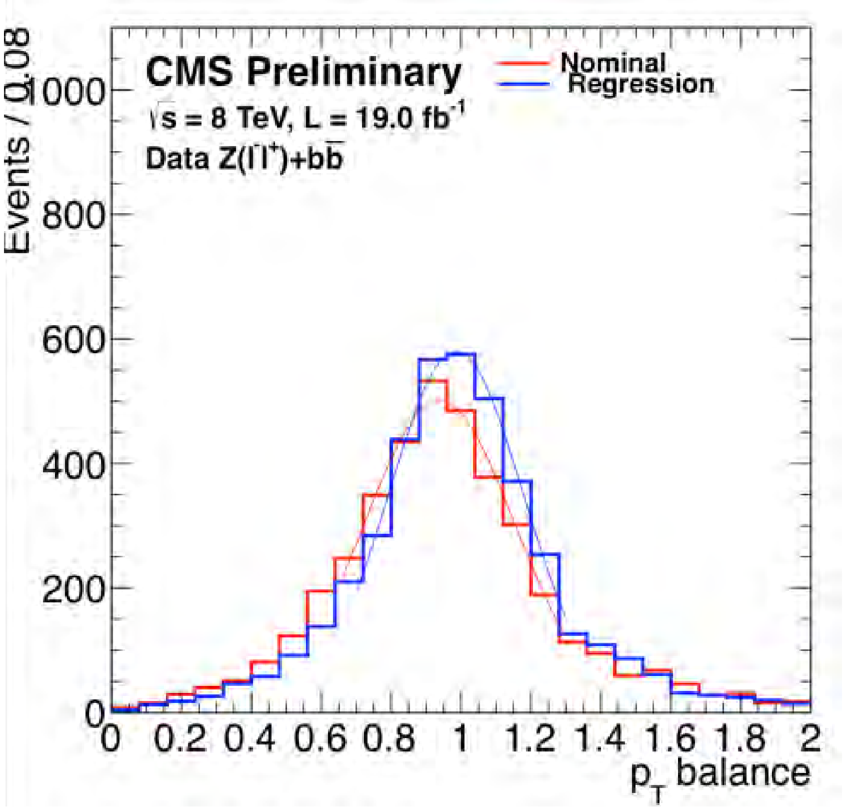
- The dijet mass is the most discriminating variable
- Its resolution depends on jets resolution
- b-jets are not like light jets
  - Presence of leptons and neutrinos
  - More massive (hence broader)
  - They can be “Tagged” with lifetime and secondary vertices
- Use a BDT regression in order to correct the jet energy exploiting jet and b-tag variables
  - ~ 15% improvement in mass resolution

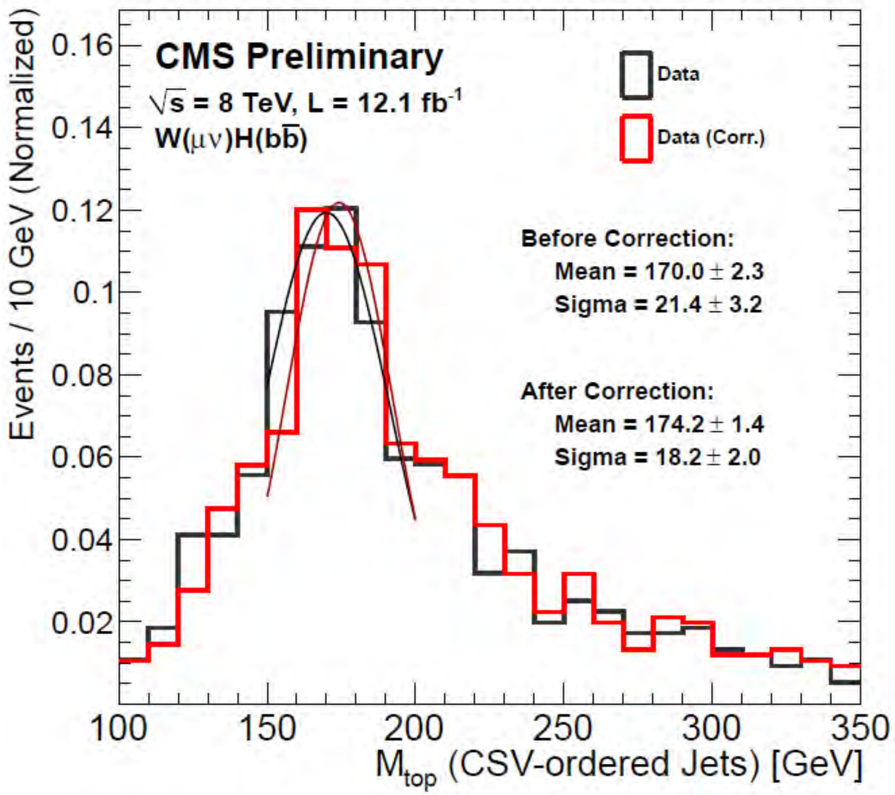




Jet energy regression

- The regression technique has been validated on data
  - pT balance in a Z+2b jets sample (Z->ll)
  - Top mass in a top enriched region
- In both cases the observed improvements matches the MC expectations

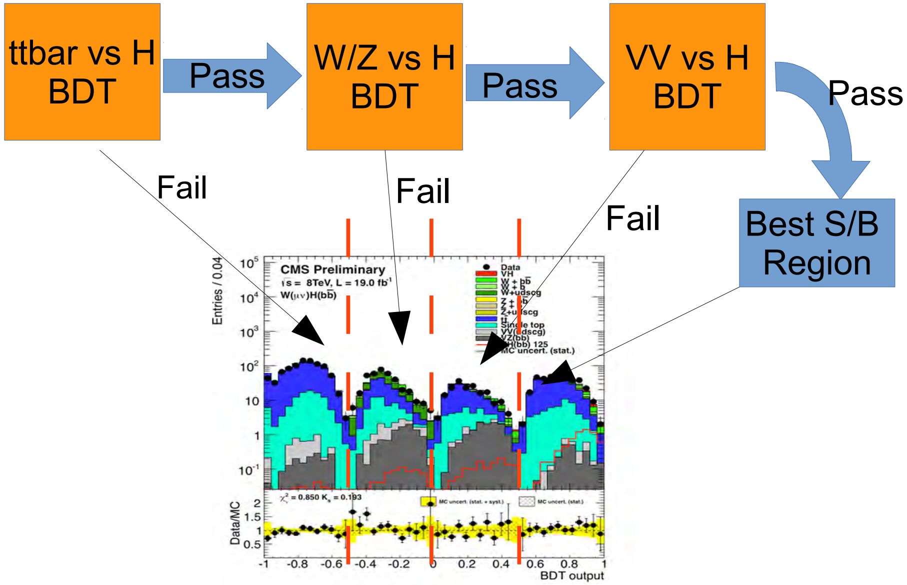




Analysis strategy

Multivariate analysis

- 2/3 categories per channels based on pT Z/W
- Loose preselection on b-tag and kinematics
- Intermediate BDT to better discriminate between different backgrounds
- Final BDT for shape fitting
- Shape uncertainties as templates from input systematic uncertainties



ttbar vs H BDT

Pass

W/Z vs H BDT

Pass

VV vs H BDT

Pass

Best S/B Region

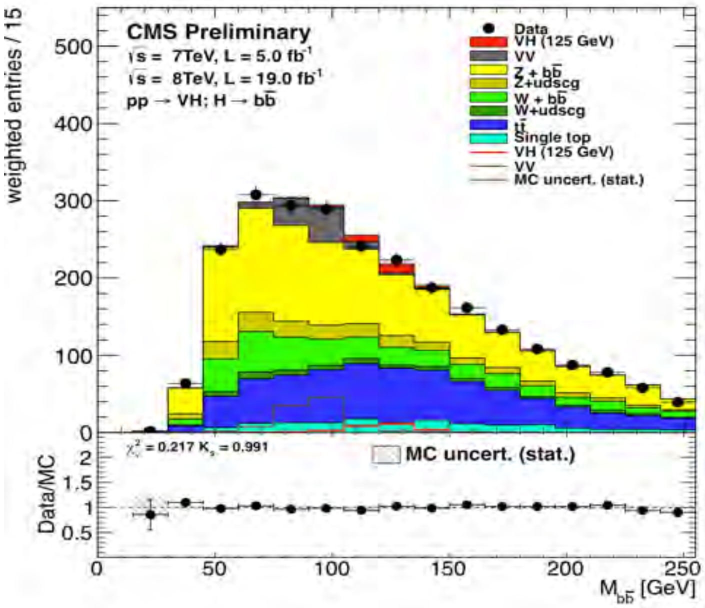
Fail

Fail

Fail

Cross check analysis

- 2/3 categories per channels based on pT Z/W
- Tighter selection on b-tag
- Invariant mass shape fit
- Shape uncertainties as templates from input systematic uncertainties
- Combined mass plot with S/B category weighting



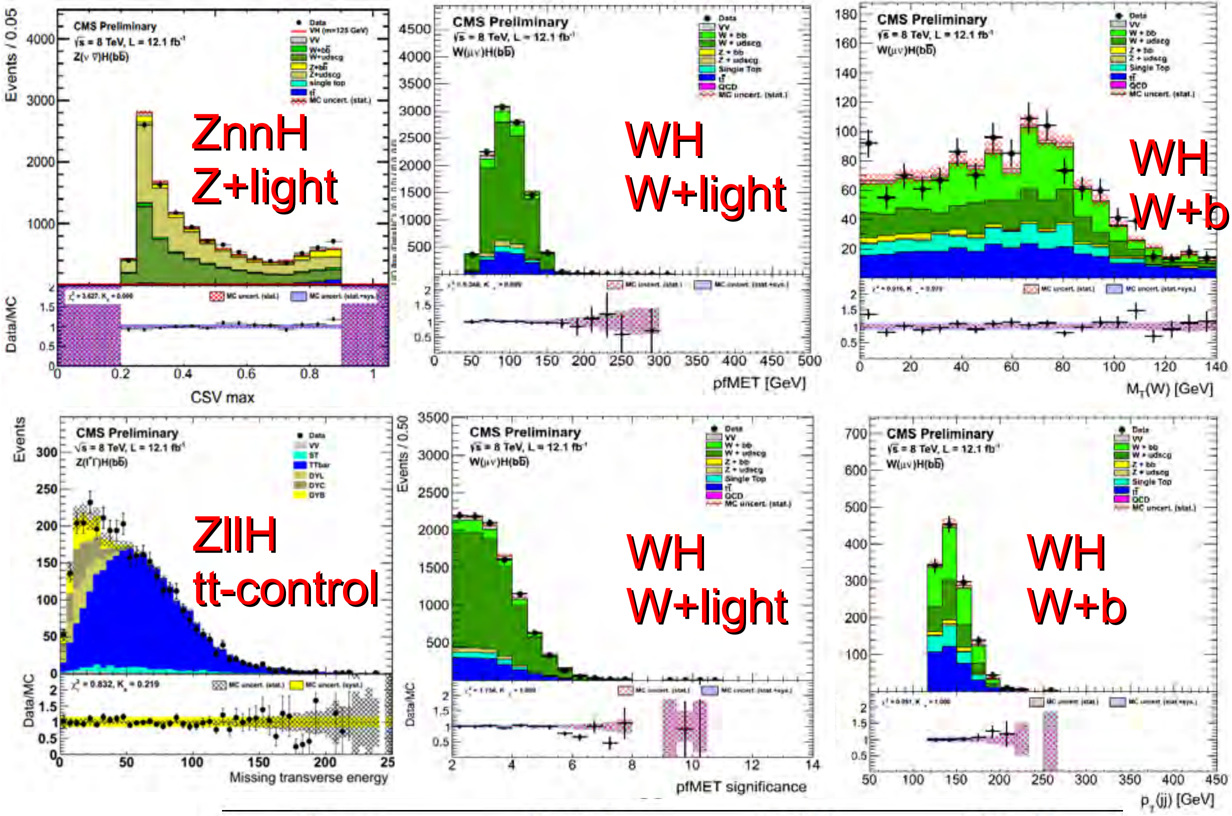
Control Regions – Background Normalization

- For each channel several control regions defined
- Shapes of all variables tested *data vs MC*
- Scale Factors for yields normalization
    - Used as starting value (with uncertainty) for nuisance parameters in the final fit
    - Only V+1b seem to be really mis-predicted by the MC

Scale Factors

Process	W$(\ell\nu)$H	Z$(\ell\ell)$H	Z$(\nu\nu)$H
High $p_T(V)$			
W + udscg	$1.04 \pm 0.01 \pm 0.07$	–	$0.93 \pm 0.02 \pm 0.03$
W + b	$2.46 \pm 0.33 \pm 0.22$	–	$2.12 \pm 0.22 \pm 0.10$
W + $b\bar{b}$	$0.77 \pm 0.25 \pm 0.08$	–	$0.71 \pm 0.25 \pm 0.15$
Z + udscg	–	$1.11 \pm 0.04 \pm 0.06$	$1.17 \pm 0.02 \pm 0.08$
Z + b	–	$1.59 \pm 0.07 \pm 0.08$	$2.13 \pm 0.05 \pm 0.07$
Z + $b\bar{b}$	–	$0.98 \pm 0.10 \pm 0.08$	$1.12 \pm 0.04 \pm 0.10$
$t\bar{t}$	$1.00 \pm 0.01 \pm 0.11$	$1.10 \pm 0.05 \pm 0.06$	$0.99 \pm 0.02 \pm 0.03$

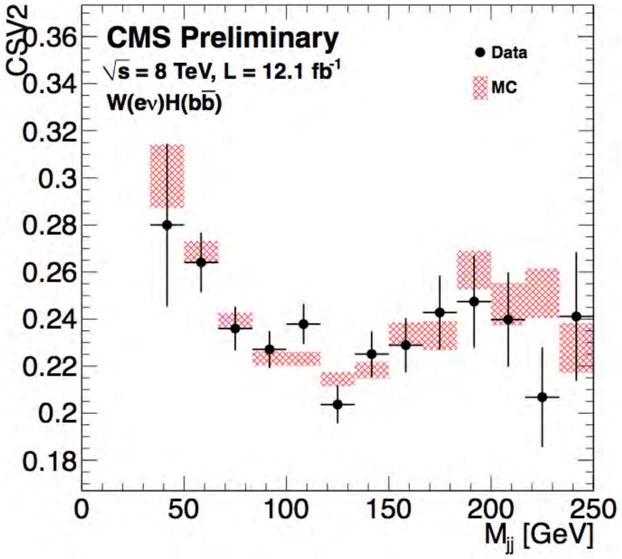
A small subset of checked variables

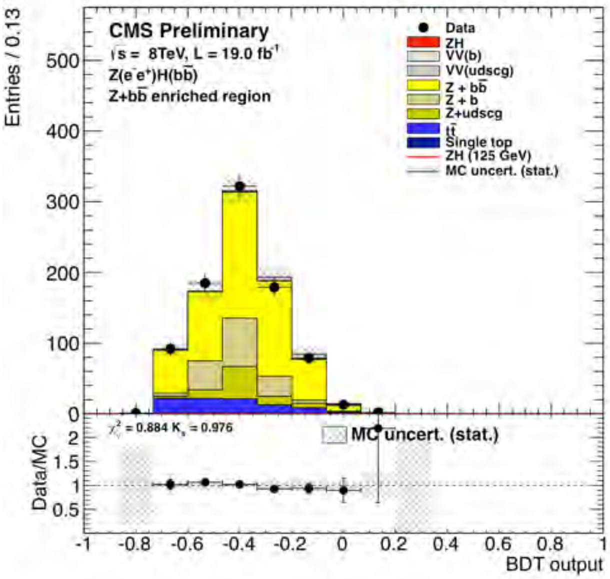


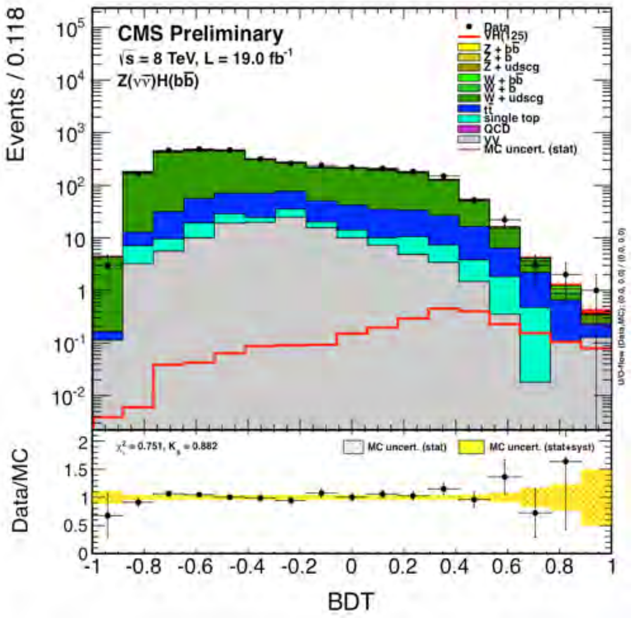
Process	W$(\ell\nu)$H	Z$(\ell\ell)$H	Z$(\nu\nu)$H
Low $p_T(V)$			
W + udscg	$1.03 \pm 0.01 \pm 0.05$	–	$0.83 \pm 0.02 \pm 0.04$
W + b	$2.22 \pm 0.25 \pm 0.20$	–	$2.30 \pm 0.21 \pm 0.11$
W + $b\bar{b}$	$1.58 \pm 0.26 \pm 0.24$	–	$0.85 \pm 0.24 \pm 0.14$
Z + udscg	–	$1.11 \pm 0.04 \pm 0.06$	$1.24 \pm 0.03 \pm 0.09$
Z + b	–	$1.59 \pm 0.07 \pm 0.08$	$2.06 \pm 0.06 \pm 0.09$
Z + $b\bar{b}$	–	$0.98 \pm 0.10 \pm 0.08$	$1.25 \pm 0.05 \pm 0.11$
$t\bar{t}$	$1.03 \pm 0.01 \pm 0.04$	$1.10 \pm 0.05 \pm 0.06$	$1.01 \pm 0.02 \pm 0.04$
Intermediate $p_T(V)$			
W + udscg	$1.02 \pm 0.01 \pm 0.07$	–	$0.93 \pm 0.02 \pm 0.04$
W + b	$2.90 \pm 0.26 \pm 0.20$	–	$2.08 \pm 0.20 \pm 0.12$
W + $b\bar{b}$	$1.30 \pm 0.23 \pm 0.14$	–	$0.75 \pm 0.26 \pm 0.11$
Z + udscg	–	–	$1.19 \pm 0.03 \pm 0.07$
Z + b	–	–	$2.30 \pm 0.07 \pm 0.08$
Z + $b\bar{b}$	–	–	$1.11 \pm 0.06 \pm 0.12$
$t\bar{t}$	$1.02 \pm 0.01 \pm 0.15$	–	$0.99 \pm 0.02 \pm 0.03$

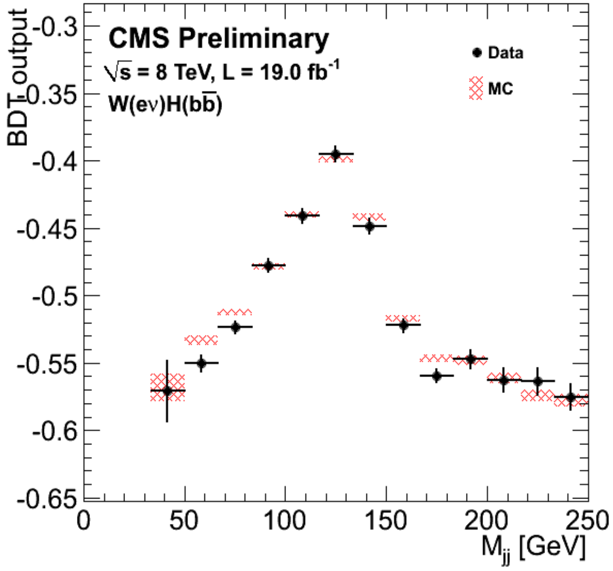
Control Regions - BDT

- Reliability of BDT from control regions
  - Correlations of input variables
  - Correlation of BDT output with input variables (e.g. *mass* vs BDT)
  - Output distribution of the BDT
- All data vs MC checks show excellent agreement



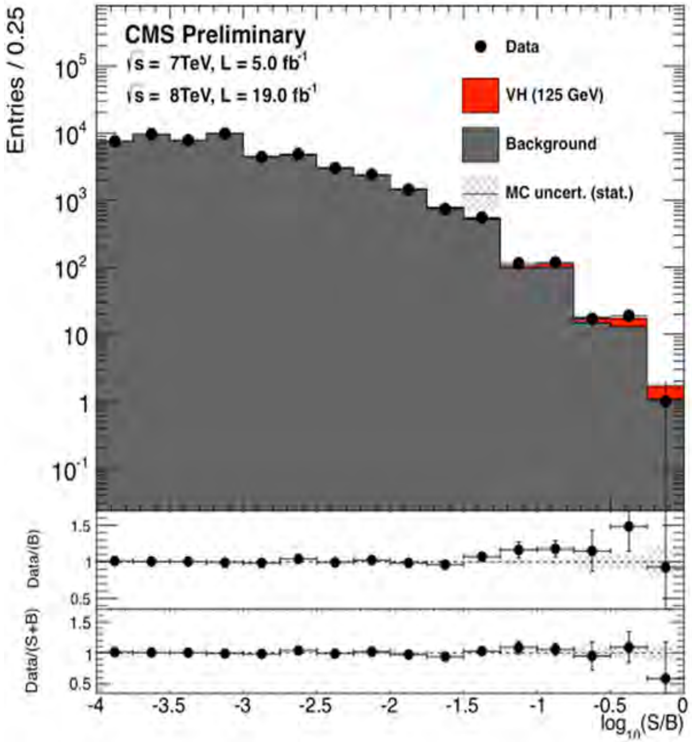


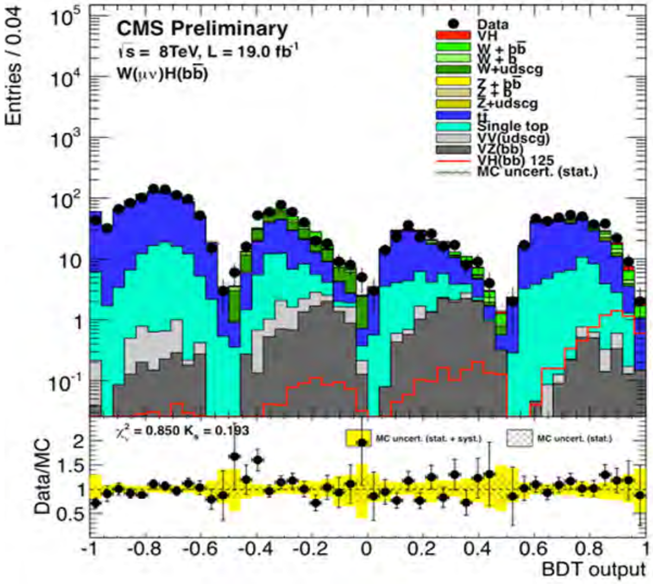


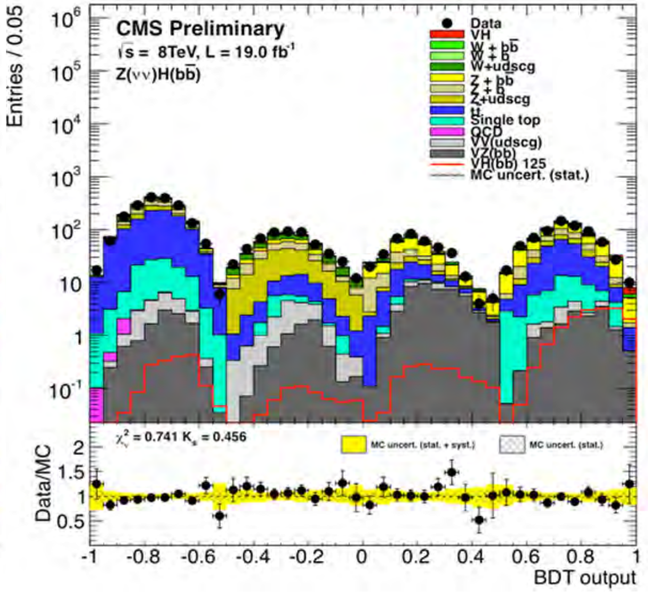


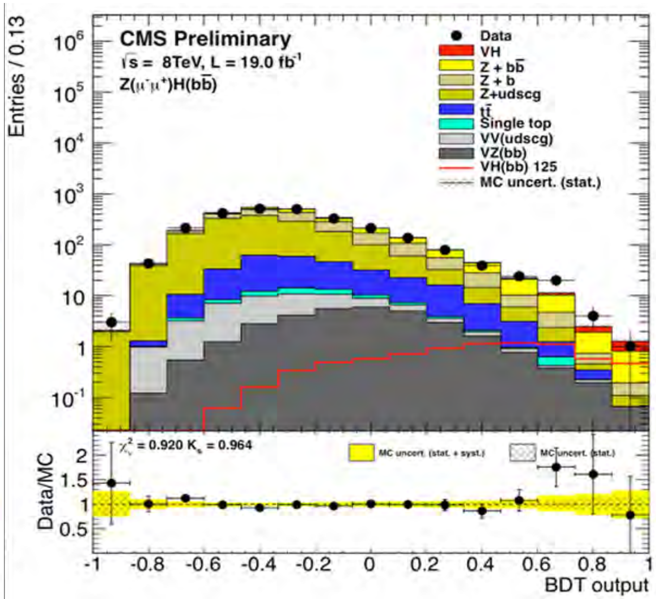
BDT output in signal region

- Each decay mode has an independently trained BDT
- To increase the sensitivity the analysis is divided into two pT bins and a low b-tag category is added
- The final result is obtained from a global fit with correlated nuisances



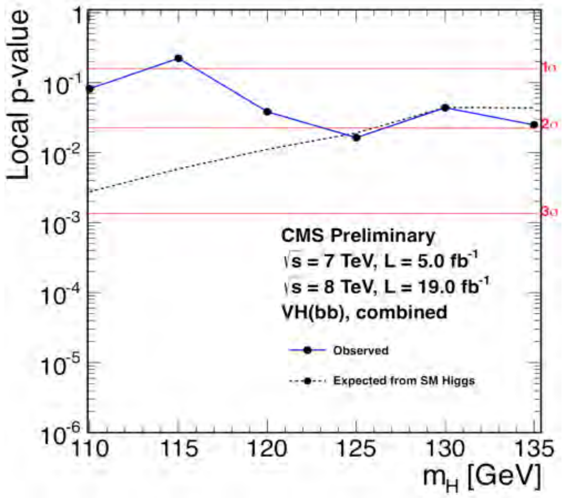


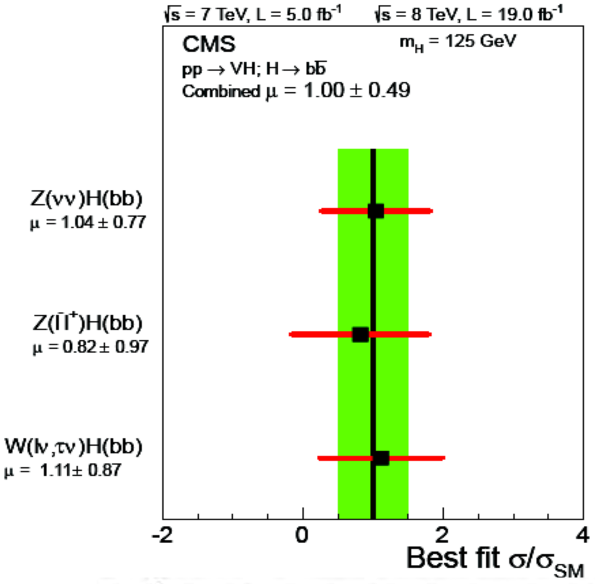


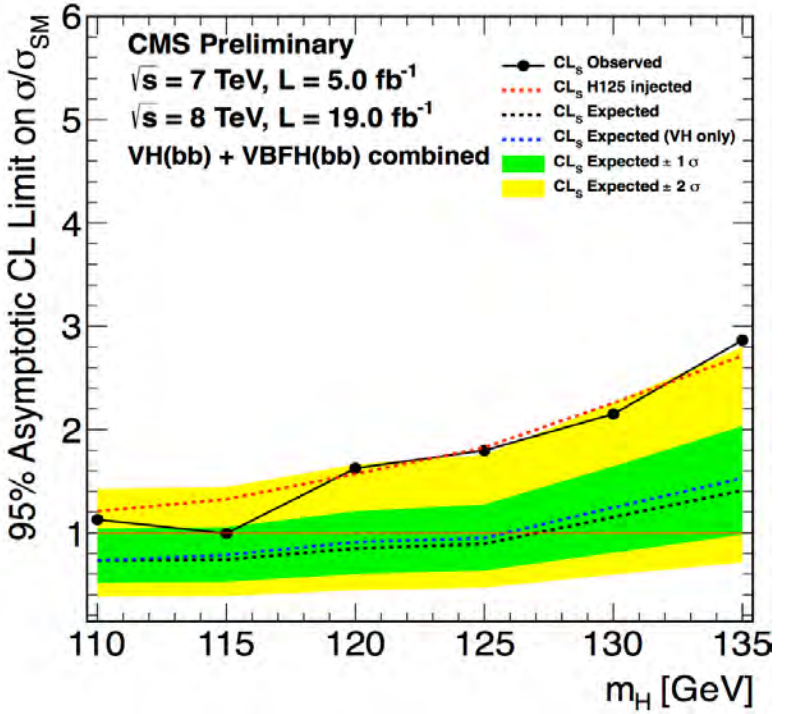


VH MVA Results

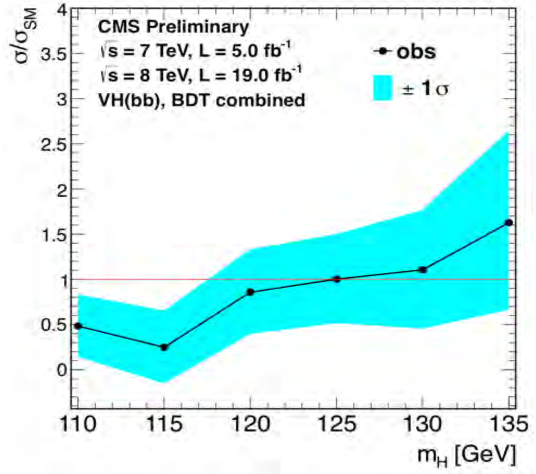
- Broad excess compatible with the 125 GeV boson
  - 2.1sigma obs, 2.1sigma exp

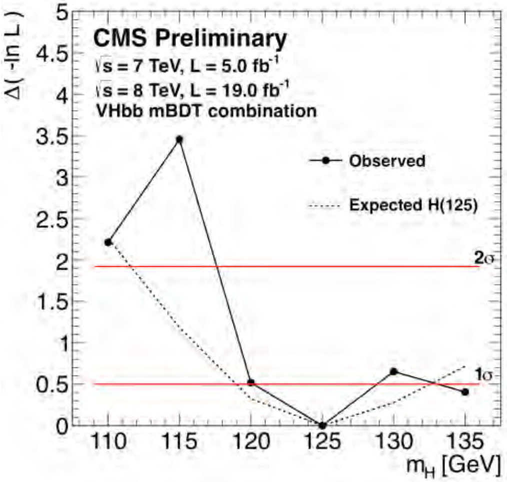






m_H [GeV]	110	115	120	125	130	135
σ/σ_{SM} (95% CL) median expected	0.73	0.79	0.91	0.95	1.25	1.53
σ/σ_{SM} (95% CL) observed	1.13	1.09	1.74	1.89	2.30	3.07



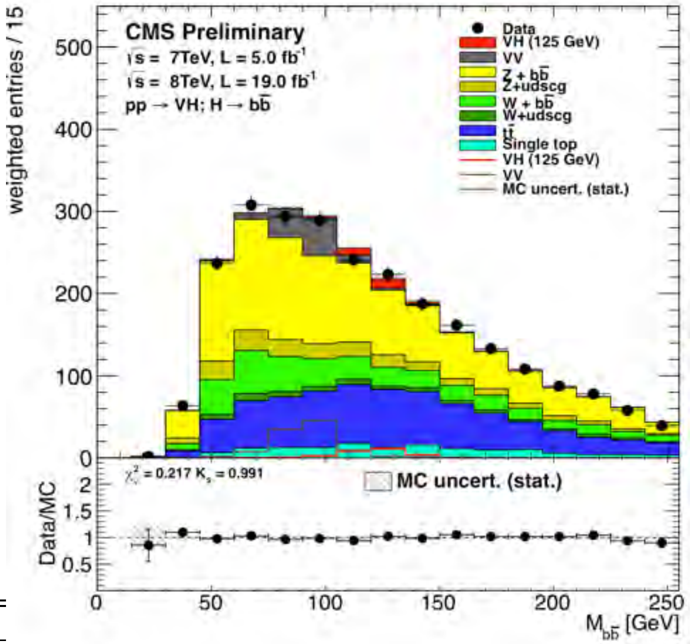


@125 GeV
sig = 2.1 std. dev.
mu = 1.0 + 0.5 – 0.5

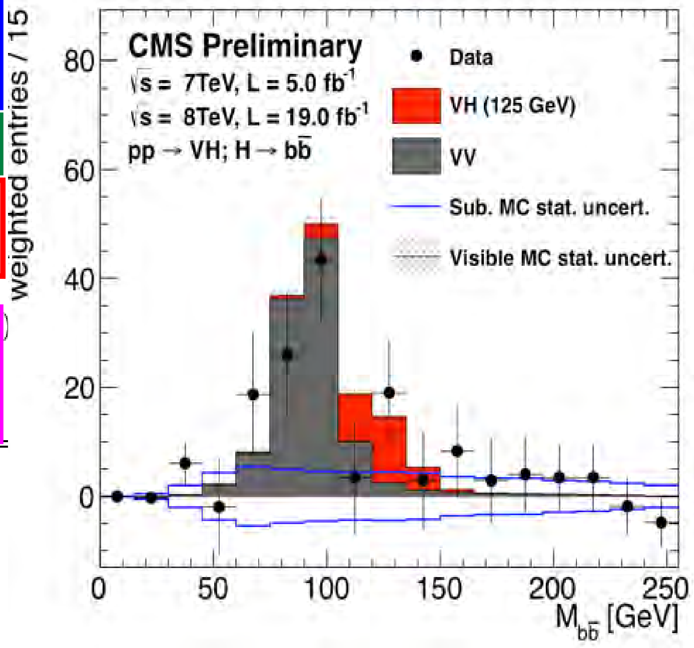
Mjj Analysis

While the main analysis is based on a BDT, a Cross-check analysis is implemented as *a shape analysis* on the dijet invariant mass selecting high S/B with:

→ Exploit the boost (pt binning)
→ Double asymmetric b-tagging
→ Topology: b2b, jet veto
→ QCD rejection

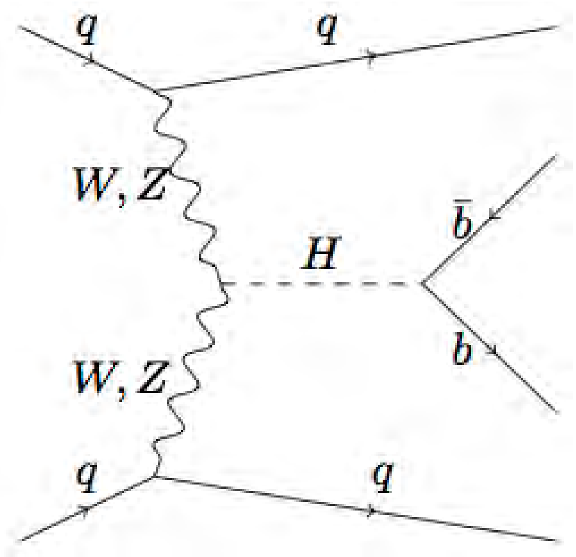


Variable	W($\mu\nu$)H	W($e\nu$)H	Z($\ell\ell$)H	Z($\nu\nu$)H
$m_{\ell\ell}$	–	–	$75 < m_{\ell\ell} < 105$	–
$p_T(j_1)$	> 30	> 30	> 20	> 60 (> 60, > 80)
$p_T(j_2)$	> 30	> 30	> 20	> 30
p_T(jj)	> 100	> 100	–	> 110 (> 140, > 190)
p_T(V)	$100-130$ ($130-180$, > 180)	$[100-150]$(> 150)	$[50-100]$ ($[100-150]$, > 150)	–
CSV1	CSVT	CSVT	CSVM	CSVT
CSV2	> 0.5	> 0.5	> 0.5	> 0.5
$\Delta\phi$(V, H)	> 2.95	> 2.95	–	> 2.95
ΔR(jj)	–	–	–(–, < 1.6)	–
N_{aj}	$= 0$	$= 0$	–	$= 0$
N_{al}	$= 0$	$= 0$	–	$= 0$
E_T^{miss}	> 45	> 45	$< 60.$	$[100-130]$ ($[130-170]$, > 170)
$\Delta\phi$(pfMET, J)	–	–	–	> 0.7 (> 0.7, > 0.5)
$\Delta\phi$(pfMET, trkMET)	–	–	–	< 0.5
$\Delta\phi$(pfMET, lep)	$< \pi/2$	$< \pi/2$	–	–



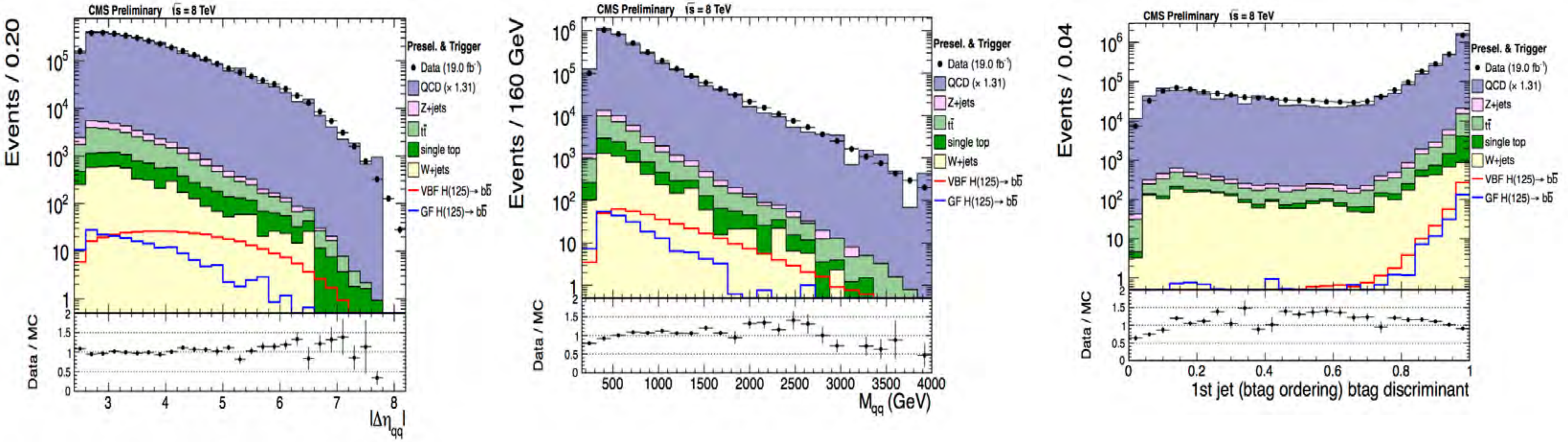
@125 GeV
sig = 1.1 std. dev.
Mu = 0.8 + 0.7 – 0.7

VBF, H->bb



VBF Hbb

- The well known VBF signature consists in an additional pair of forward-backward jets
- In the case of VBF,H->bb the final state is fully hadronic
  - Very large QCD background
- The discrimination is based on b-tag, rapidity gap and invariant mass of the light jets

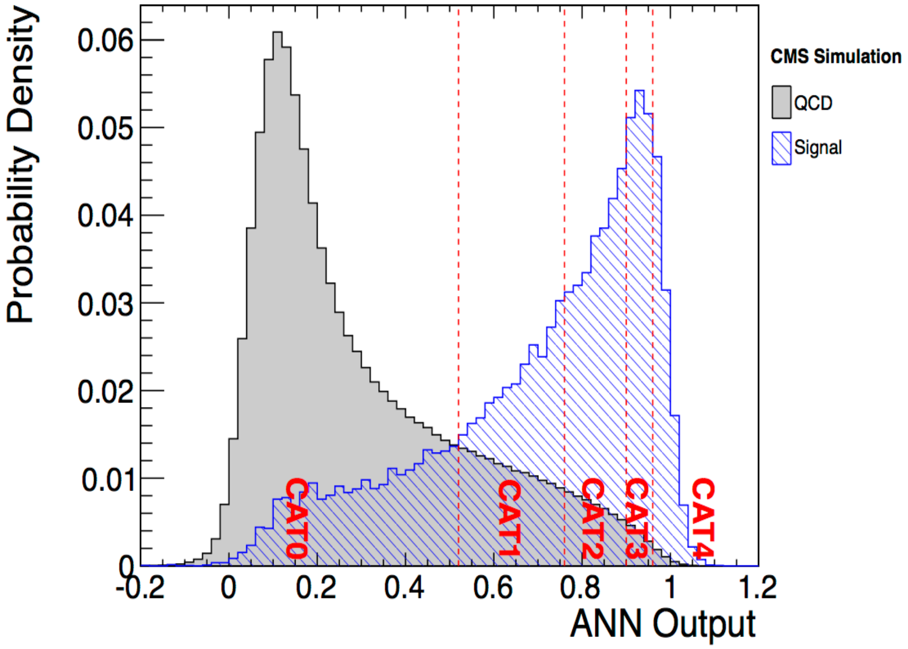


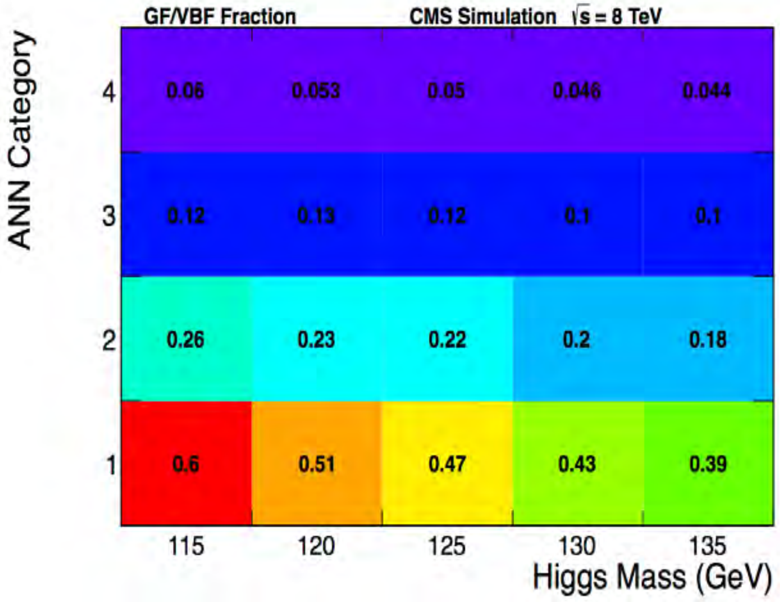
Analysis strategy

- Combine all discriminating variables into an MVA output
  - Do not use variables that correlates with Mbb
- Categorize events based on the MVA output
  - The MVA also separates gg->H from VBF H
- Fit a peaking signal on a smooth background

Inputs to the MVA:

- eta separation between the btag sorted qq jets.
- eta separation difference between the b-tag and eta sorted qq jets.
- invariant mass of the b-tag sorted qq jet pair
- average eta of the b-tag sorted qq jet pair system.
- CSV b-tagging output for the most b-tagged jet.
- SV b-tagging output for the second most b-tagged jet.
- quark/gluon discriminator for the third b-tagged jet.
- quark/gluon discriminator for the least b-tagged jet.
- eta of the third b-tagged jet.
- scalar p_T sum of the additional “soft” Track-Jets with $p_T > 1$ GeV.
- angular variables

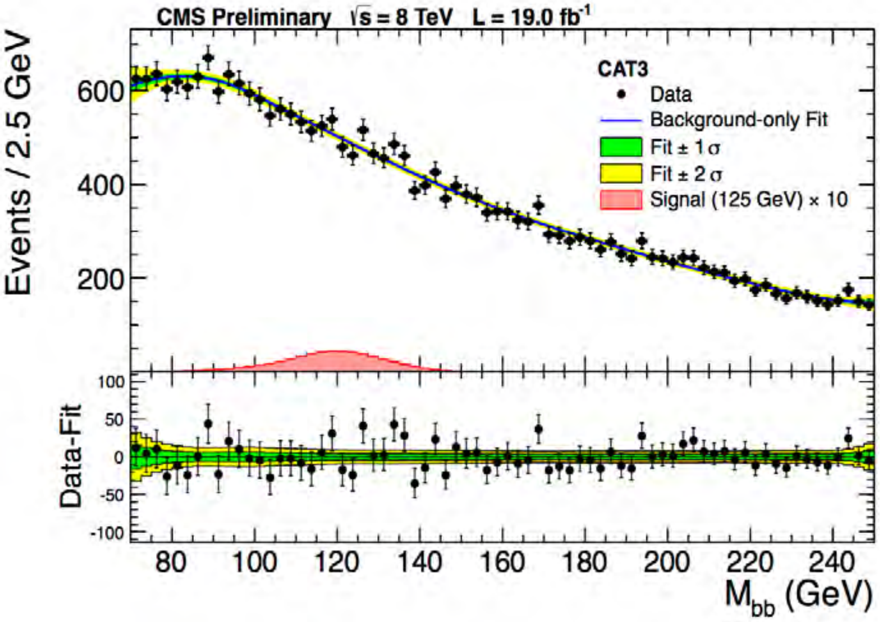


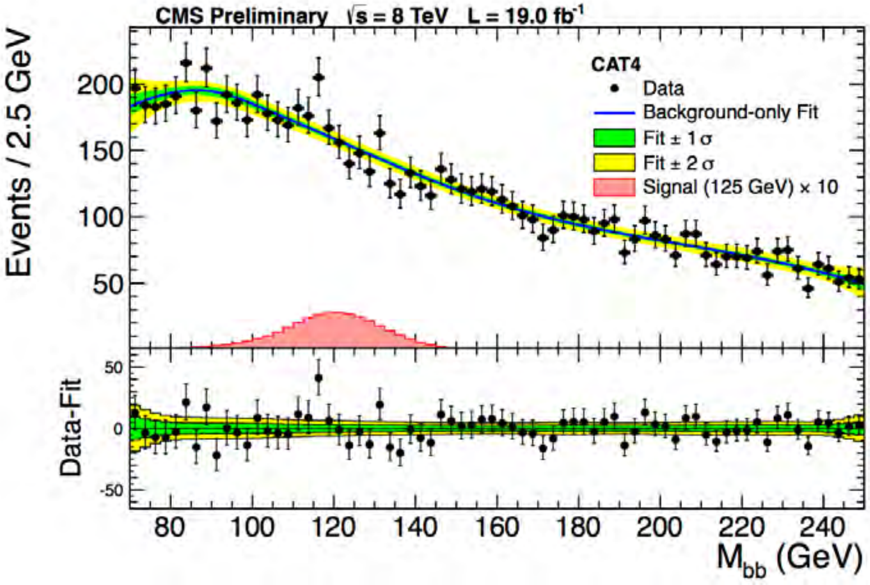


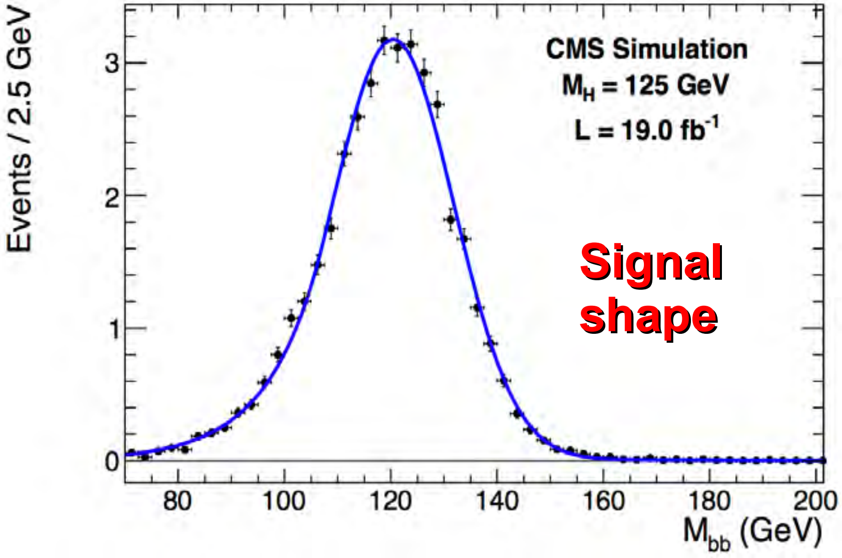
ANN Category	115	120	125	130	135
4	0.06	0.053	0.05	0.046	0.044
3	0.12	0.13	0.12	0.1	0.1
2	0.26	0.23	0.22	0.2	0.18
1	0.6	0.51	0.47	0.43	0.39

Fit in the bb invariant mass

- The mass fit is performed using generic templates (Bernstein polynomials) for the background
- The signal template shape is tuned on the MC (xtalball plus Bernstein)
- Reliability of the fit (bias, linearity) tested using different models and different signal injections
- Non QCD backgrounds templates taken from MC



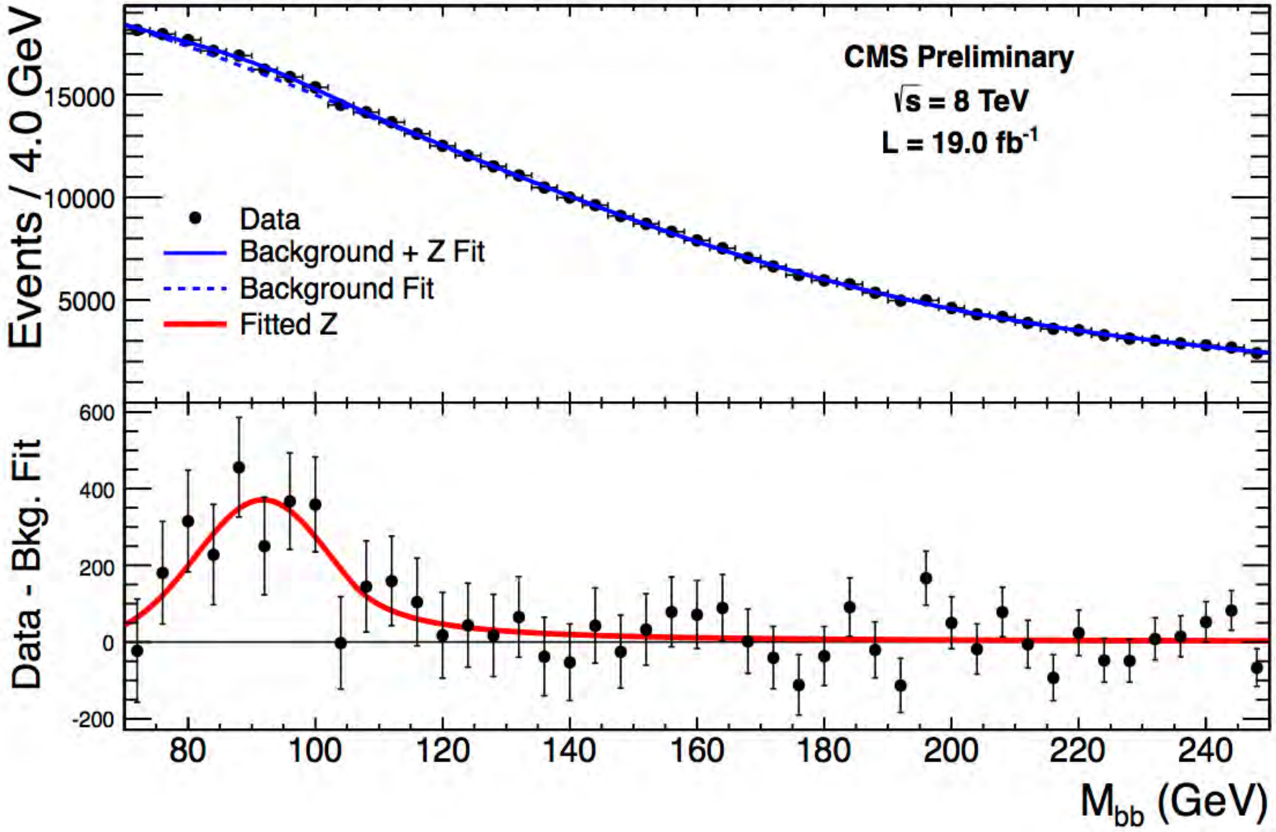




Signal shape

Z+jets cross check

- A cross check of the fitting machinery has been done without the MVA, targeting the Z+jets
- Excess due to Z correctly fitted on top of the very large background

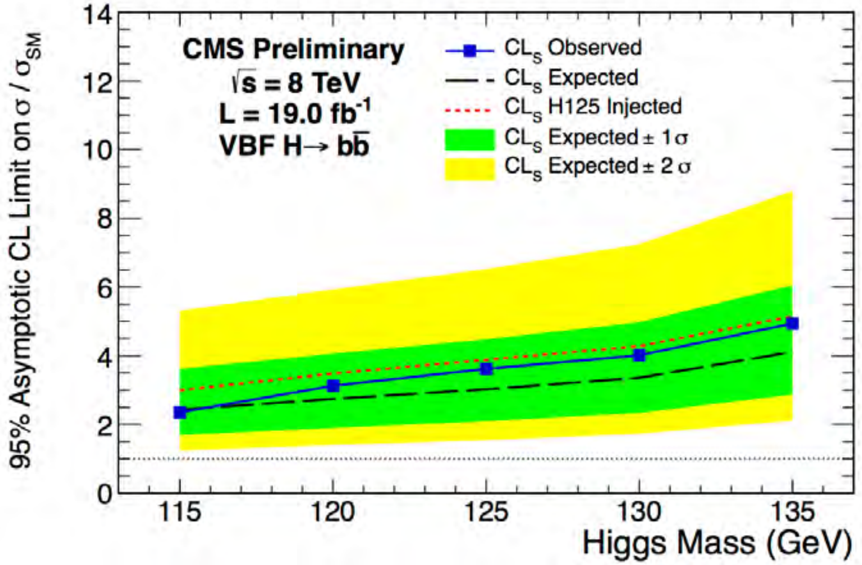


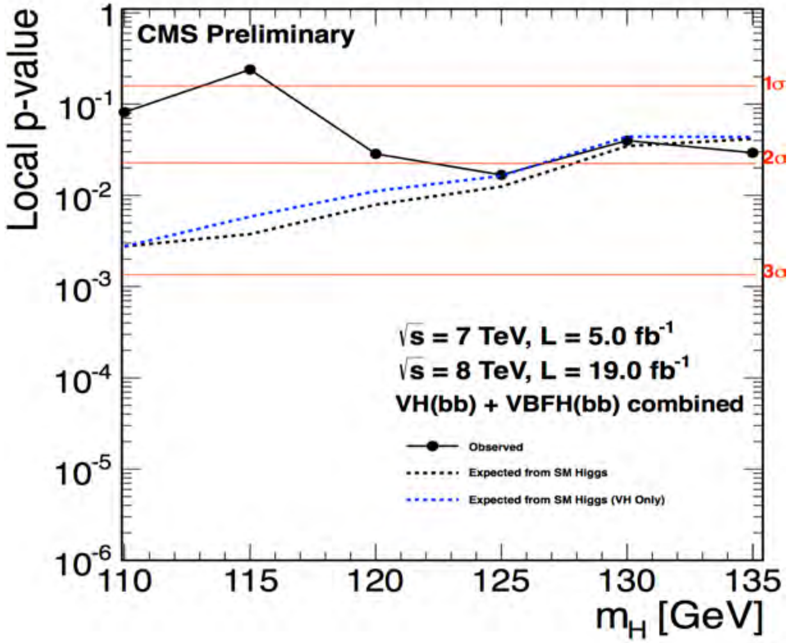
VBF preliminary results

- The first measurement at LHC of the VBF, H->bb is compatible with expectations
- Limits between 2 and 3 x SM were expected
- The observed value is compatible with the expectations for the 125 GeV Higgs boson

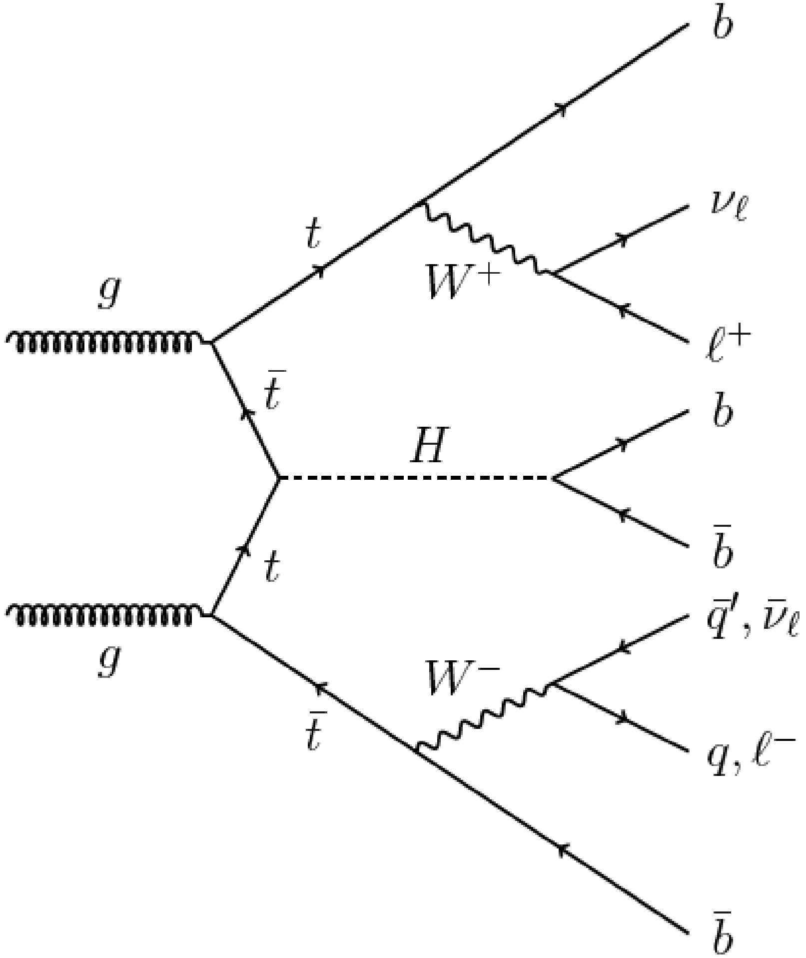
@125 GeV
Sig = 0.5 std. dev. (0.7 exp)
Mu = 0.7 + 1.4 – 1.4

- A combination with VH result is also performed



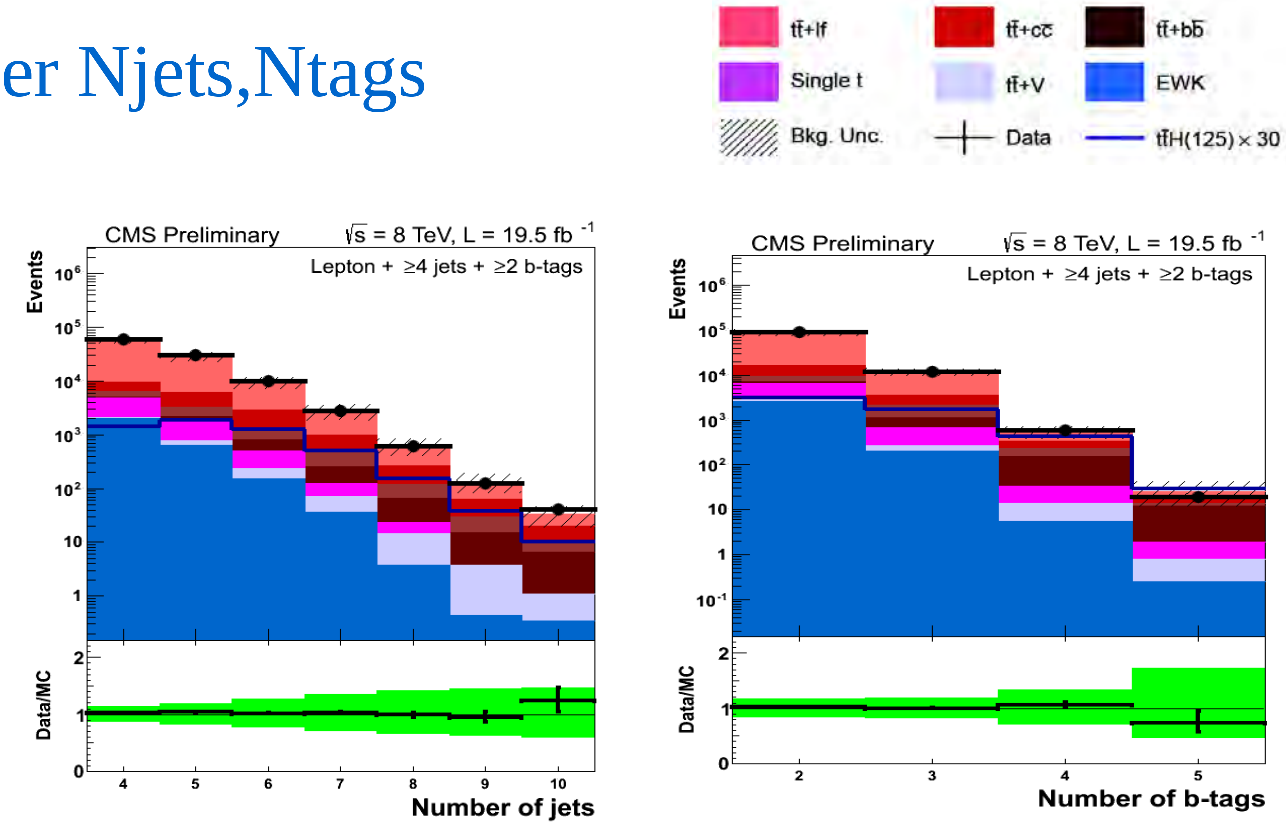


ttH (H->bb)



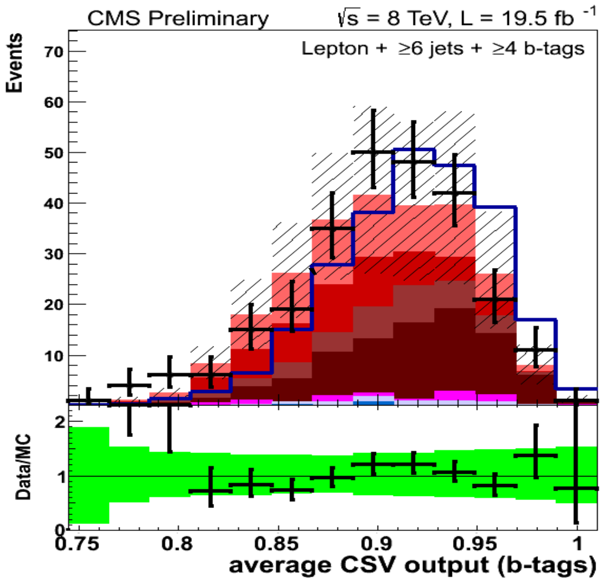
ttH, H to bb

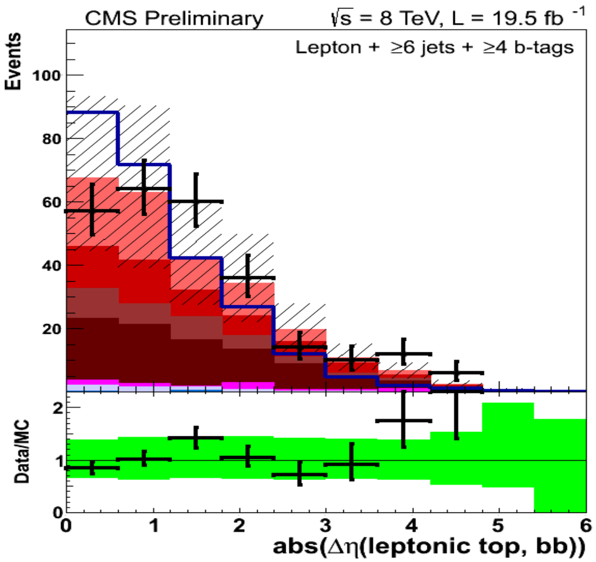
- Two modes studied: semi-leptonic and dileptonic
- Signal to background ratio rapidly increasing with
  - Total number of jets (expect 6 or 4 jets in final state)
  - Number of b-tagged jets (4 b in final state)
- Analysis categorized per Njets,Ntags
  - Low Njets,Ntags useful for backgrounds normalization
  - High Njets,Ntags are the signal region
- tt+bb background is basically irreducible

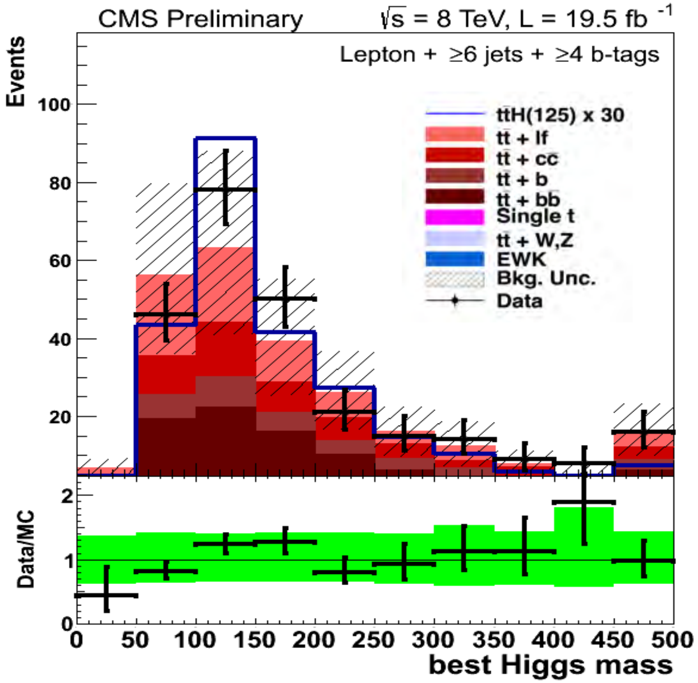


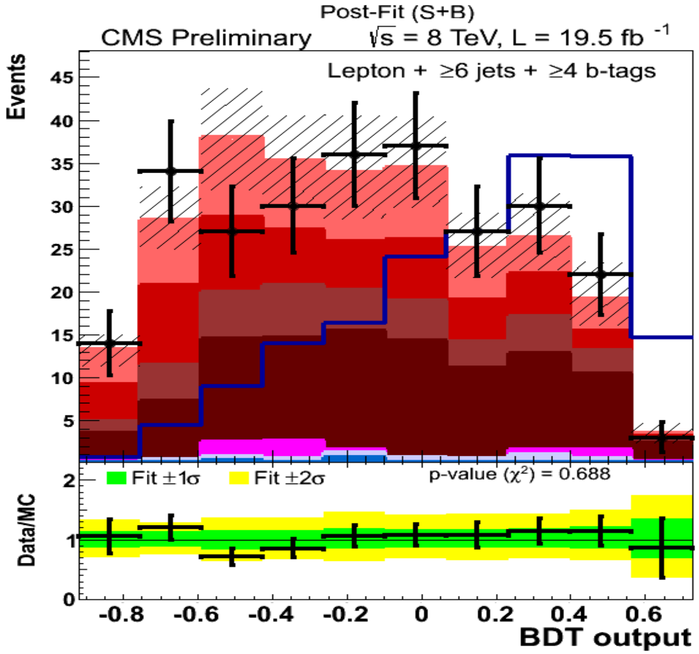
ttH, H to bb

- Several mildly discriminating variables
- Use BDT to combine
- An “Higgs mass” only defined in many jets/tags cat.





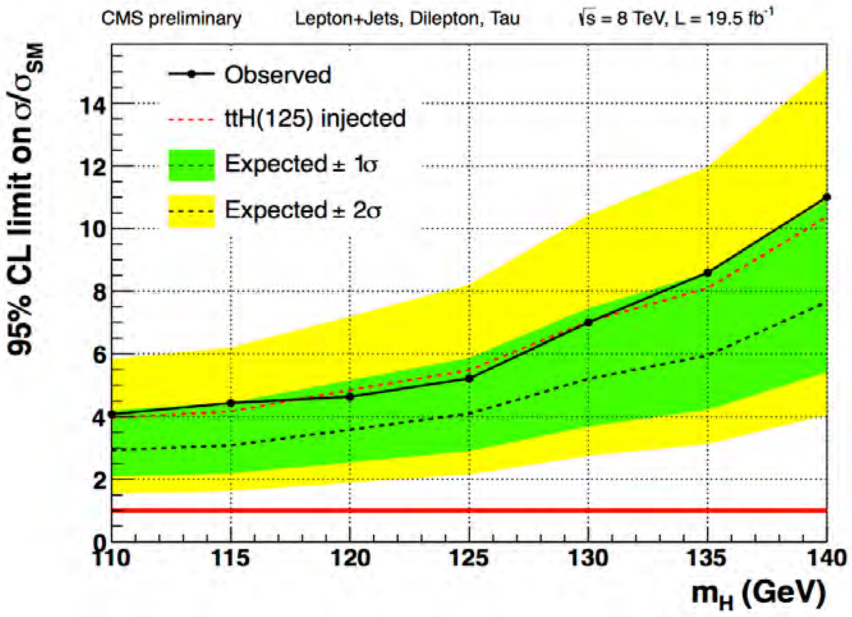


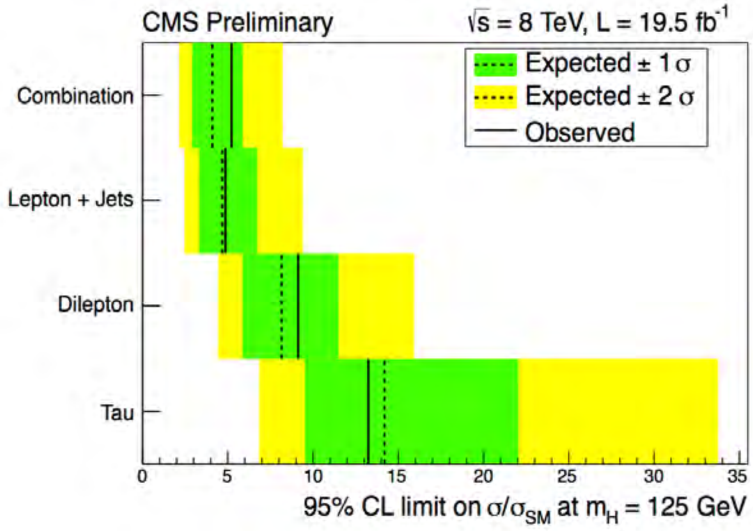


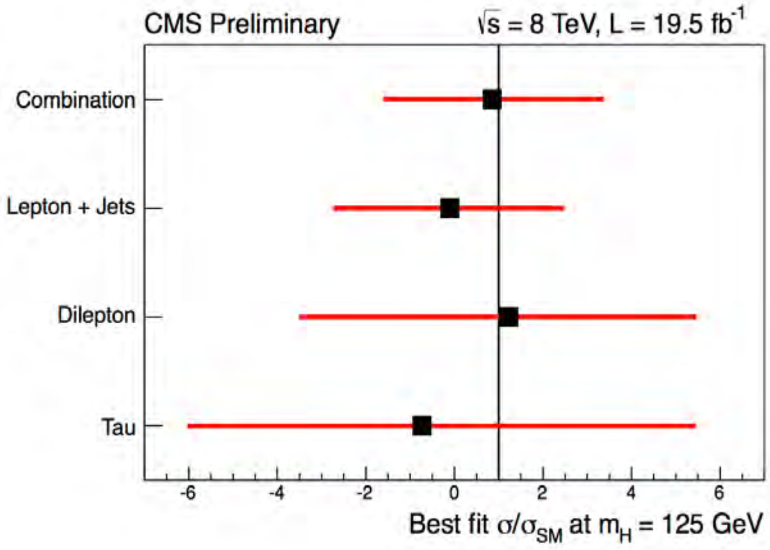
$\geq$ 6 jets, 3 tags	$\geq$ 6 jets, $\geq$ 4 tags
H_0	$(\Sigma$ jet $p_T)/(\Sigma$ jet E$)$
sphericity	ave ΔR(tag,tag)
$(\Sigma$ jet $p_T)/(\Sigma$ jet E$)$	product($\Delta\eta$(leptonic top, bb), $\Delta\eta$(hadronic top, bb))
max $\Delta\eta$ (jet, ave jet η)	closest tag mass
$\sum p_T$(jets,lepton,MET)	max $\Delta\eta$ (tag, ave tag η)
ave CSV (tags)	ave CSV (tags)
second-highest CSV (tags)	third-highest CSV (tags)
third-highest CSV (tags)	fourth-highest CSV (tags)
fourth-highest CSV (jets)	best Higgs boson mass
ttbb/ttH BDT	ttbb/ttH BDT

ttH (Hbb and Hττ)

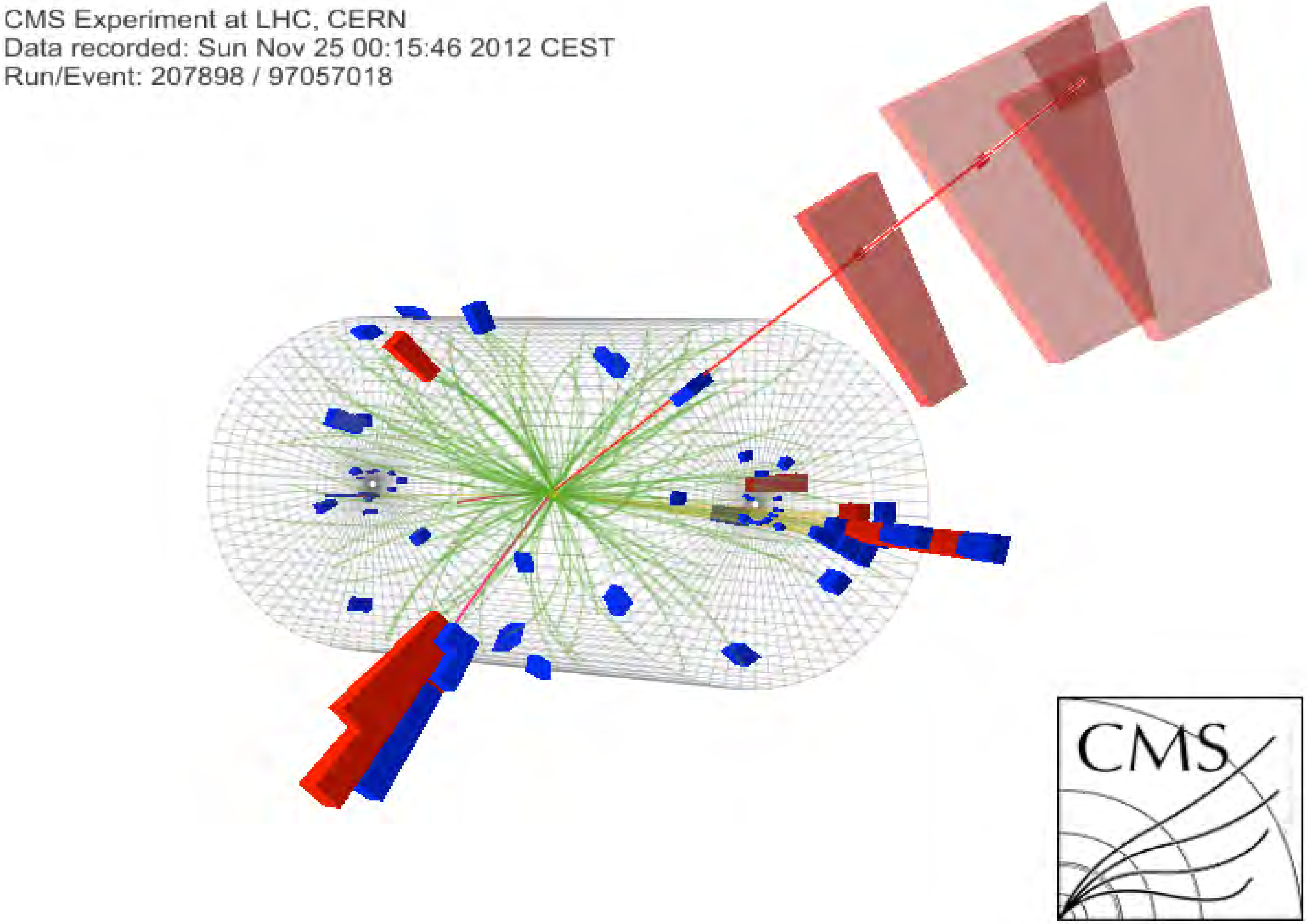
- Updated result with full 2012 luminosity presented in combination with ttH → ττ
- Sensitivity to 3-8 times the SM
- Slight excess observed, compatible with SM Higgs at 125GeV



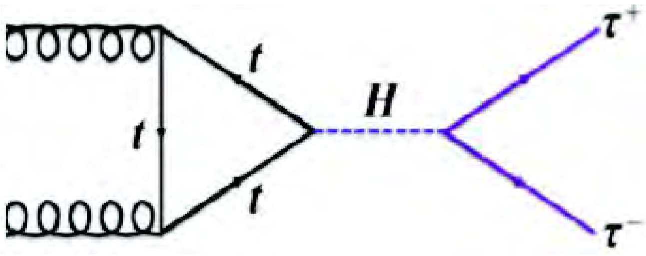


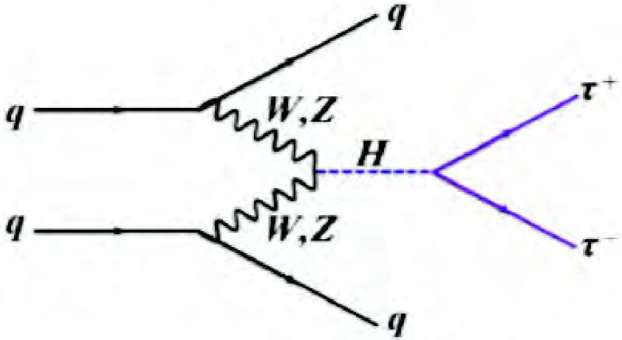


$H \to \tau\tau$


CMS Experiment at LHC, CERN
Data recorded: Sun Nov 25 00:15:46 2012 CEST
Run/Event: 207898 / 97057018
CMS

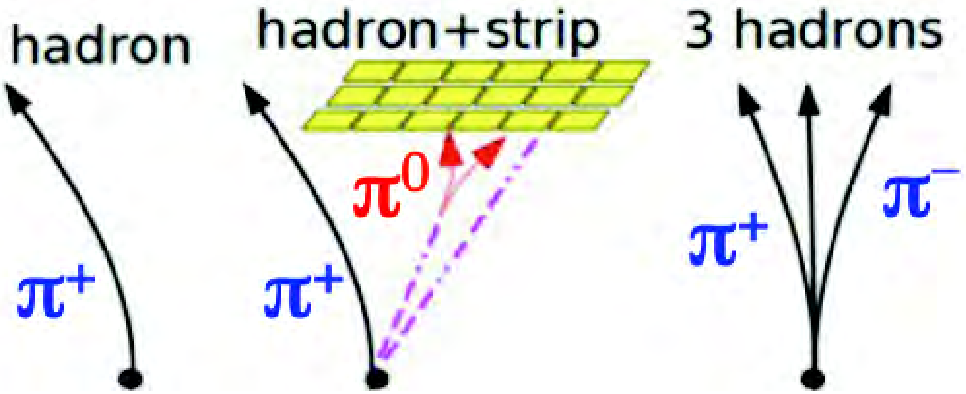
H → ττ





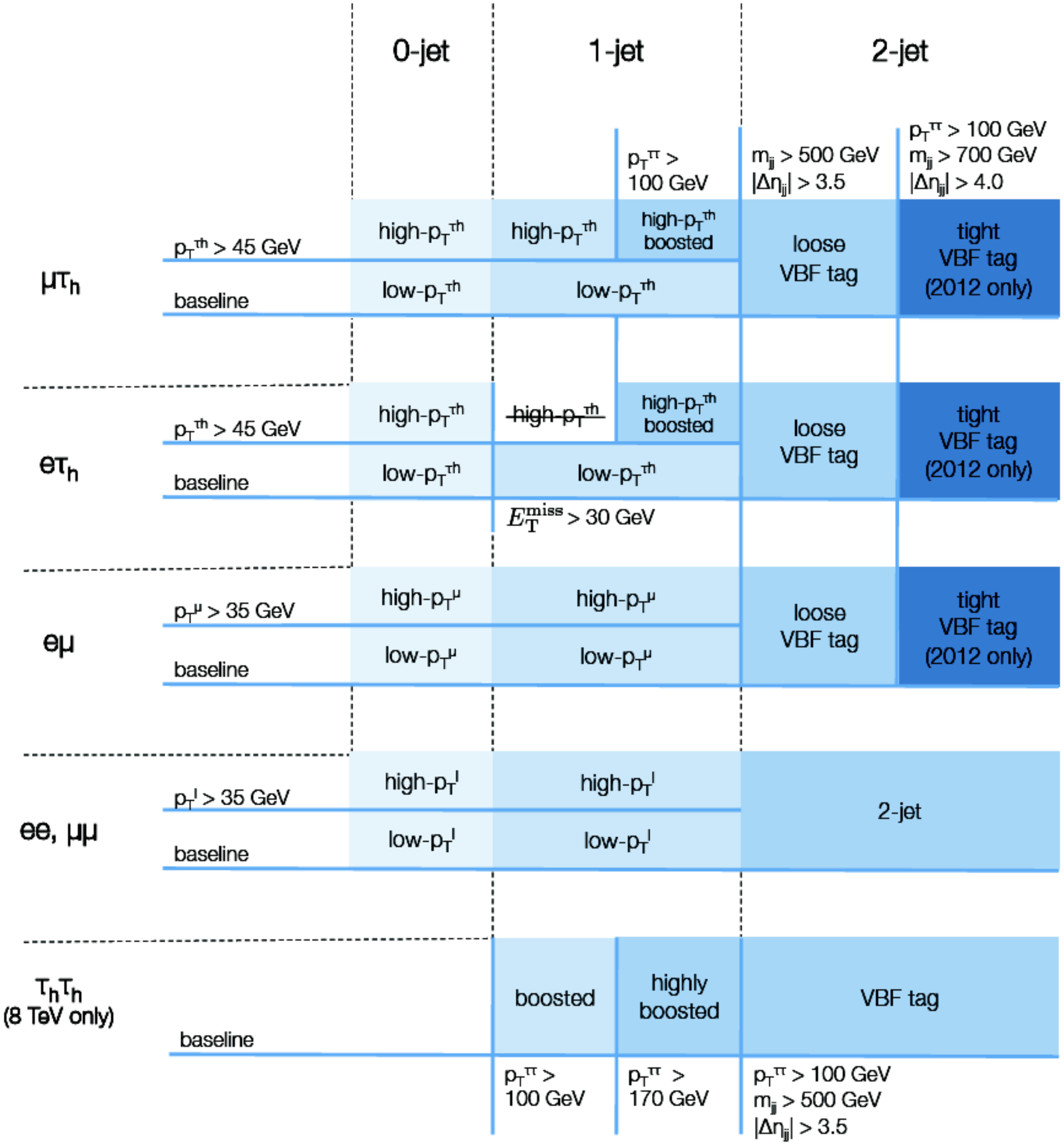
- All τ decay modes covered
  - had+had 42%
  - e/mu+had ~23%
  - e+mu 6%
  - ee/mumu 3%
- Production mechanism separated with
  - 0/1/2 jets and VBF cuts
  - WH / ZH final state with additional e/mu

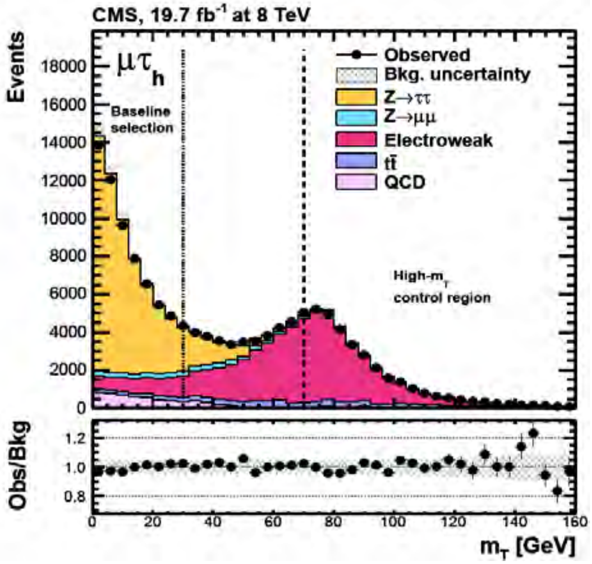
- Hadronic τ reconstructed with Particle Flow algorithms
  - 1 prong
  - 1 pront + pi0
  - 3 prongs



$H \rightarrow \tau\tau$

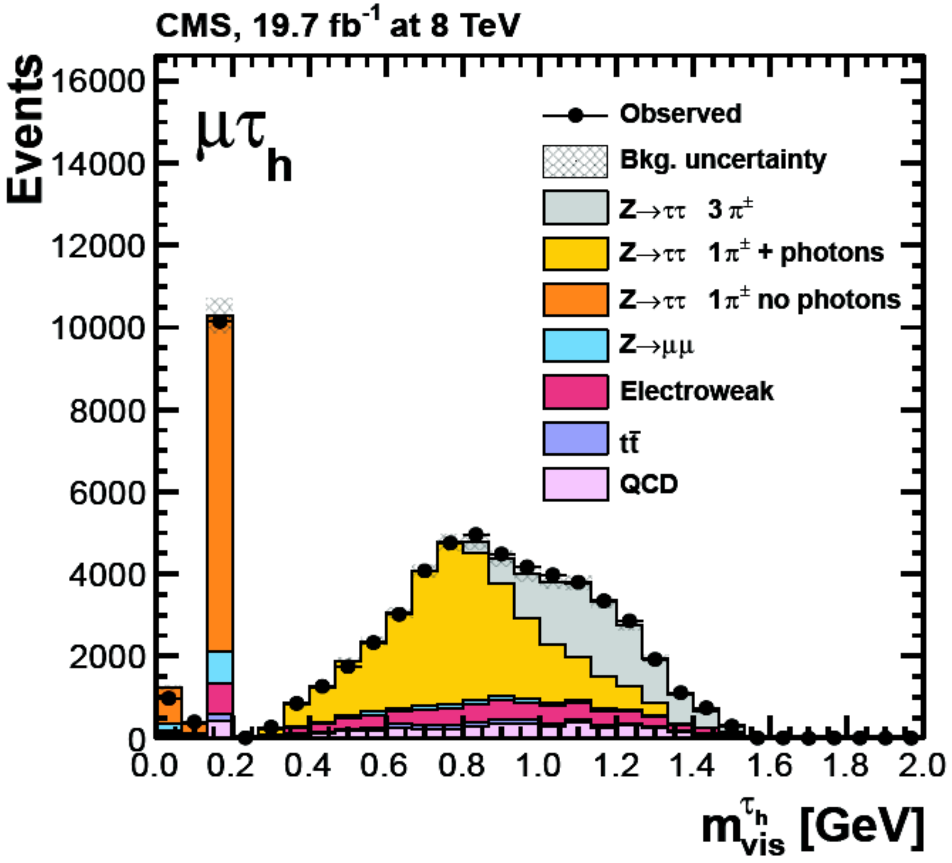
- Several regions defined in order to
  - Control backgrounds
  - Increase sensitivity
  - Distinguish different production mechanism
- E.g. EWK background controlled requiring large mT





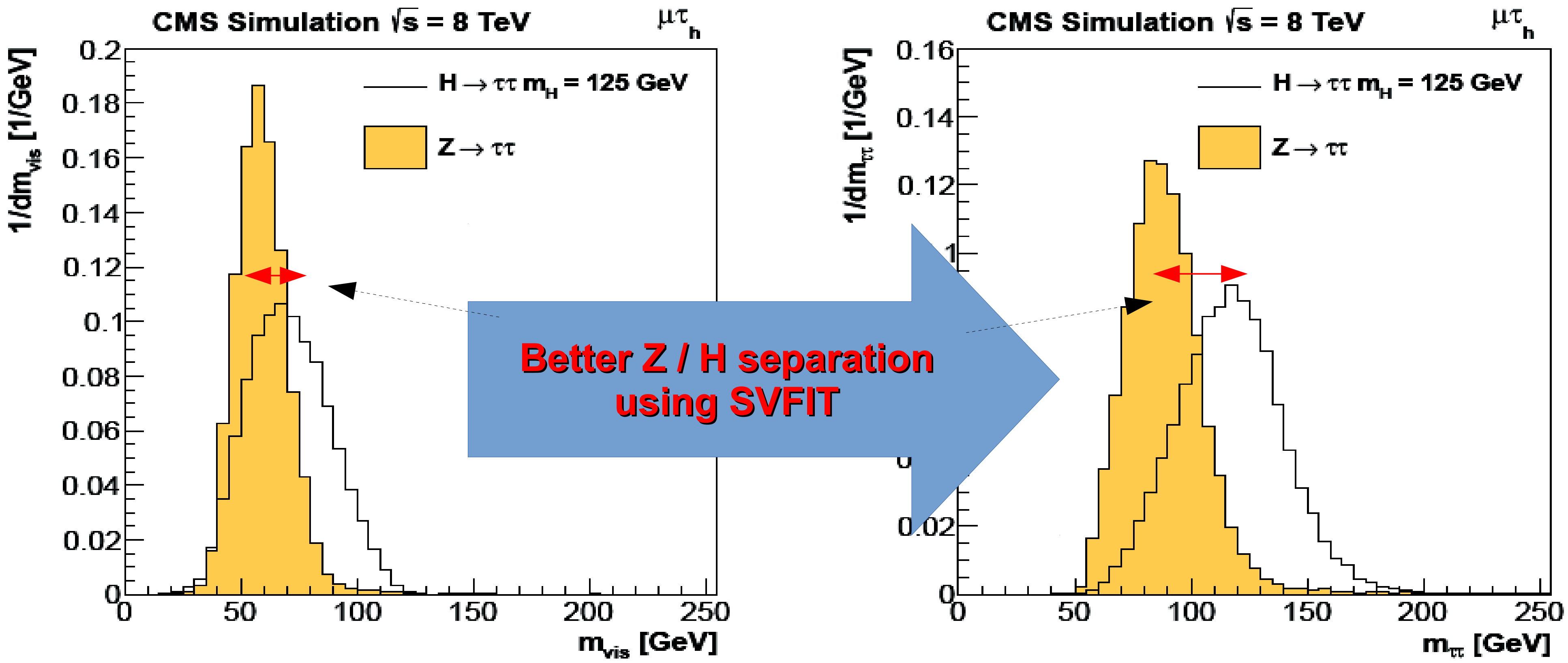
Hadronic τ reconstruction

- τ reconstruction in CMS is based on Particle Flow techniques
  - Exploit combined information of tracker, calorimeters, muon det.
  - Typical performance for hadronic τ:
    - 60% efficiency
    - 1% fake rate (Jets)
- Visible τ mass used to validate MC energy scale
  - Testing on Z → ττ sample
  - Agreement to 3% level



Di-τ mass reconstruction

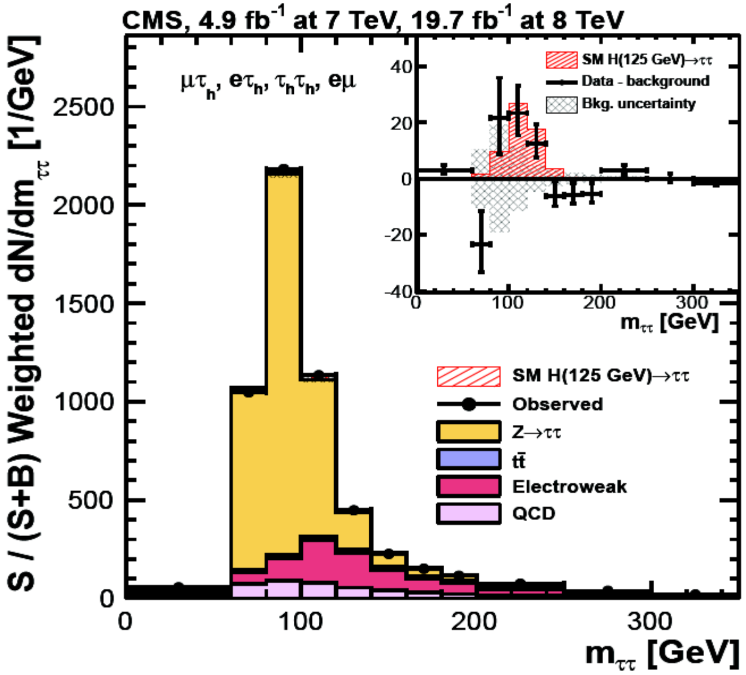
- Dedicated algorithm used to reconstruct the invariant mass
  - Presence of neutrinos spoils the resolution of the “visible mass”
  - Use MET and τ decay products kinematic variables in a dedicated (SVFIT) algorithm
    - better response and resolution
    - Better separation of Higgs from Z



Better Z / H separation using SVFIT

H → ττ

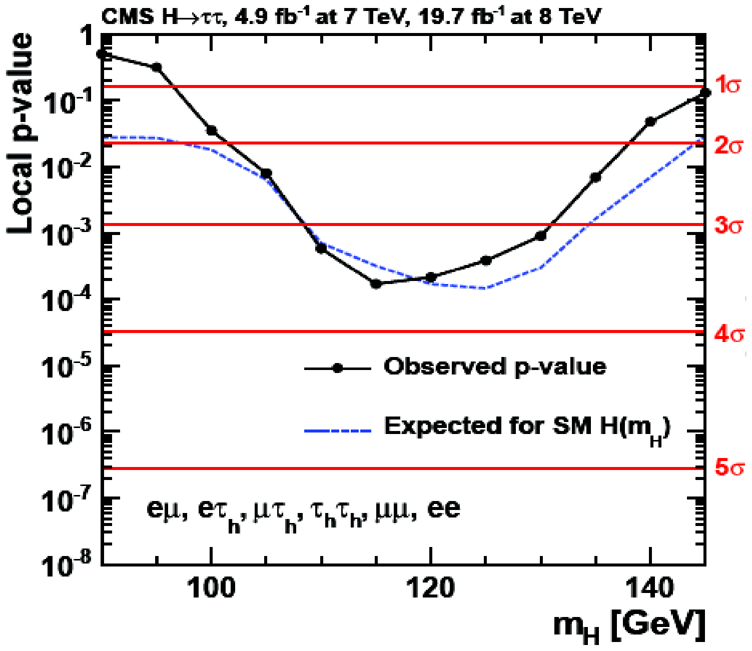
- Results in dilepton final state
  - Fit di-τ mass in all modes but in ee and mumu (BDT instead)
  - Excess of events near 120 GeV
  - Compatible with 125GeV Higgs boson
  - More than 3 sigma expected and observed

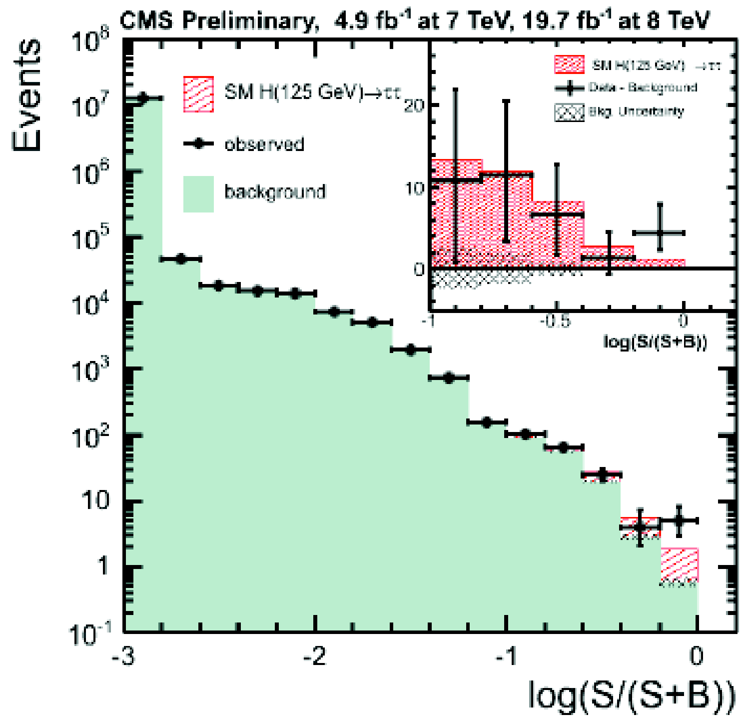


@125GeV

3.6σ exp

3.4σ obs





ττ – combination

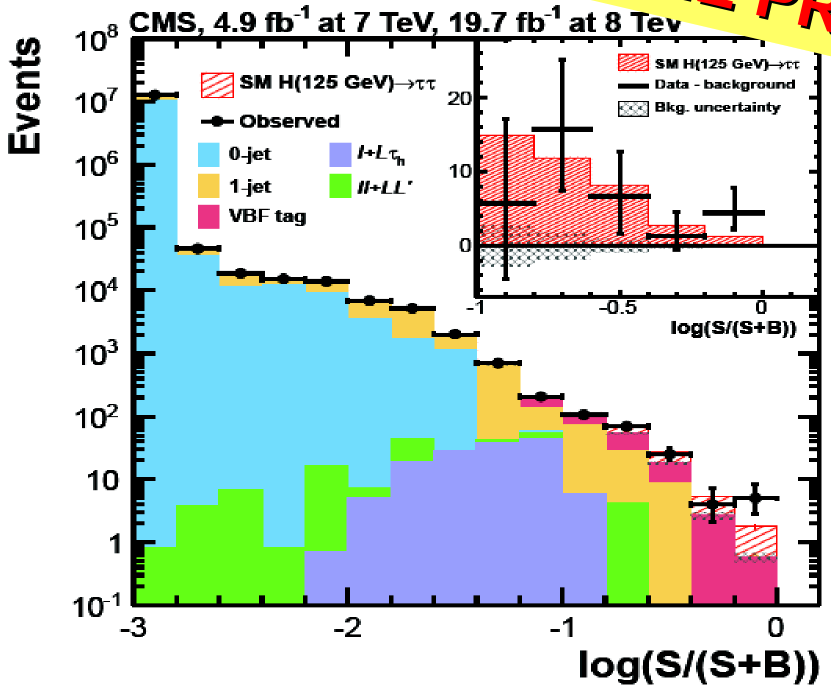
HOT OFF THE PRESS

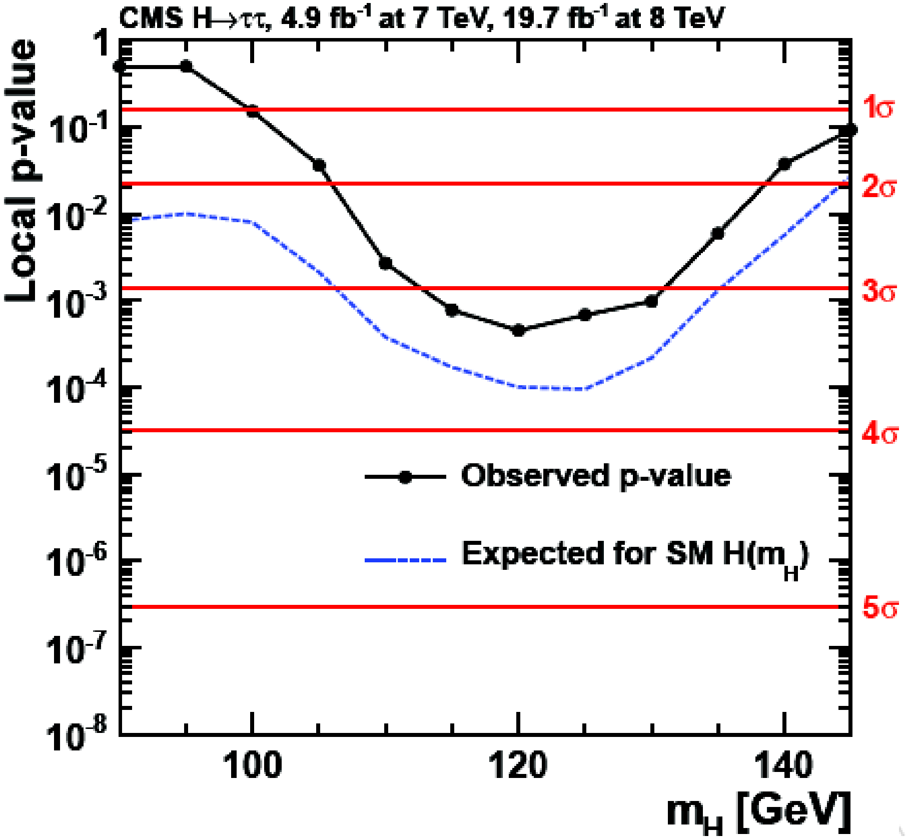
- Combining all H → ττ channels

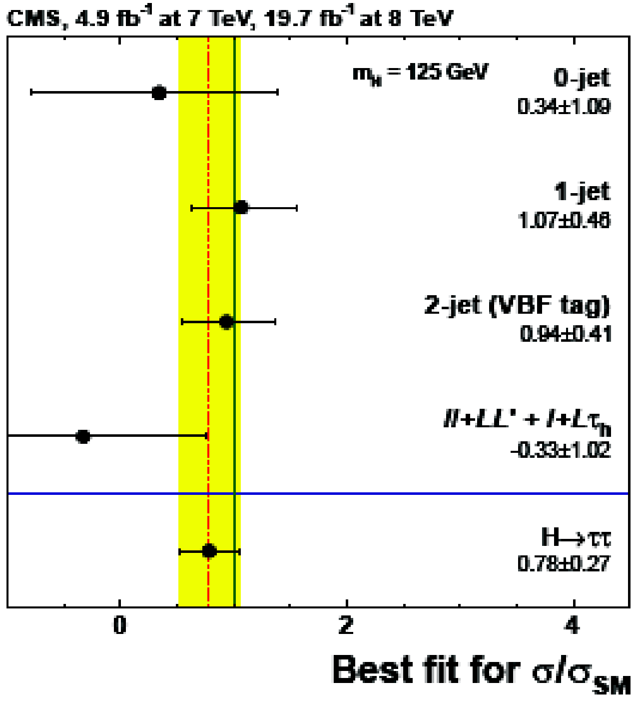
@125GeV

3.7σ exp

3.2σ obs





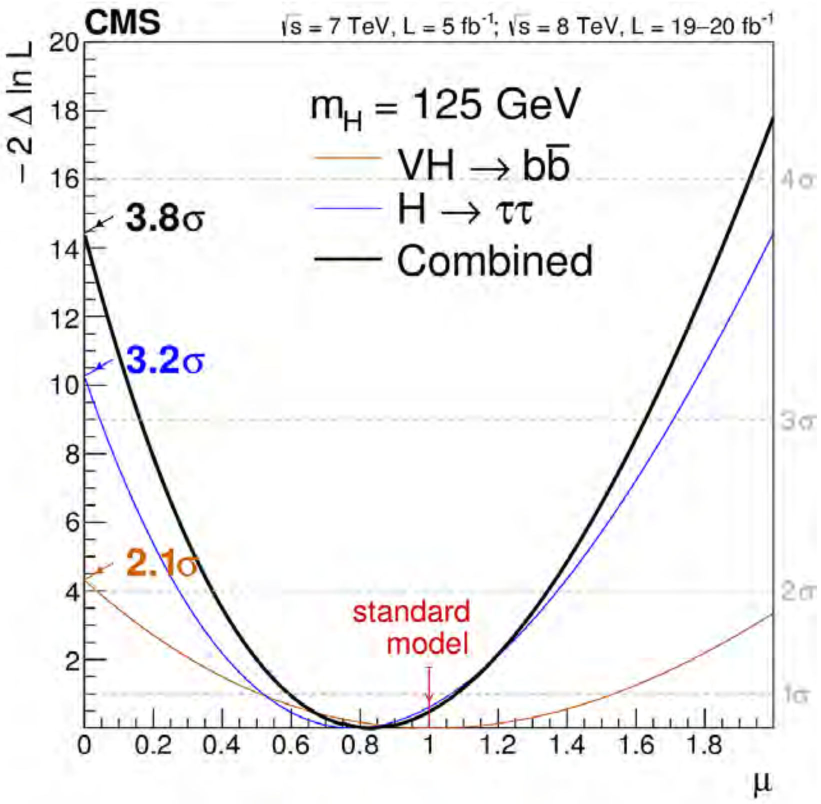


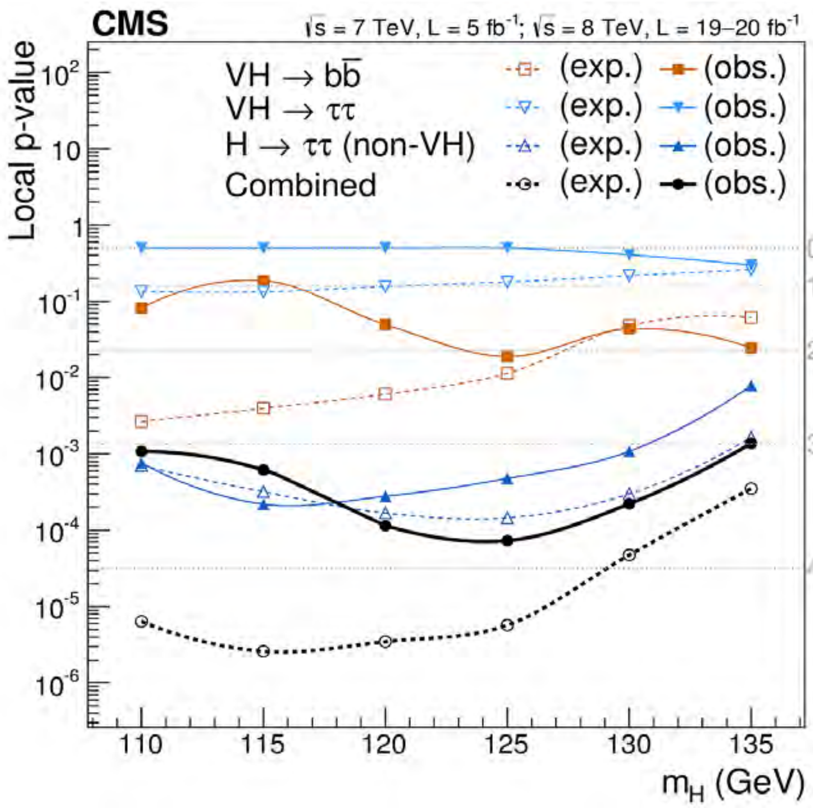
Combining τ and B

HOT OFF THE PRESS

- A combination of the most sensitive Hbb and Hττ channels has been prepared
- Measured fermion coupling

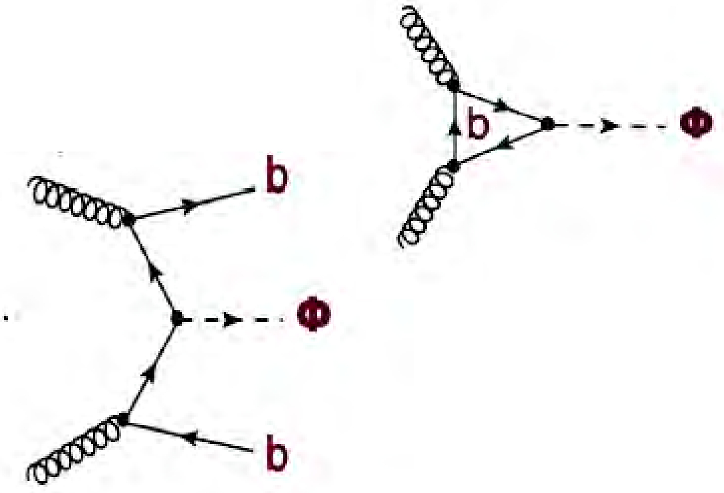
Channel	Significance (σ)		Best-fit
($m_H = 125$ GeV)	Expected	Observed	μ
$VH \to b\bar{b}$	2.3	2.1	1.0 ± 0.5
$H \to \tau\tau$	3.7	3.2	0.78 ± 0.27
Combined	4.4	3.8	0.83 ± 0.24

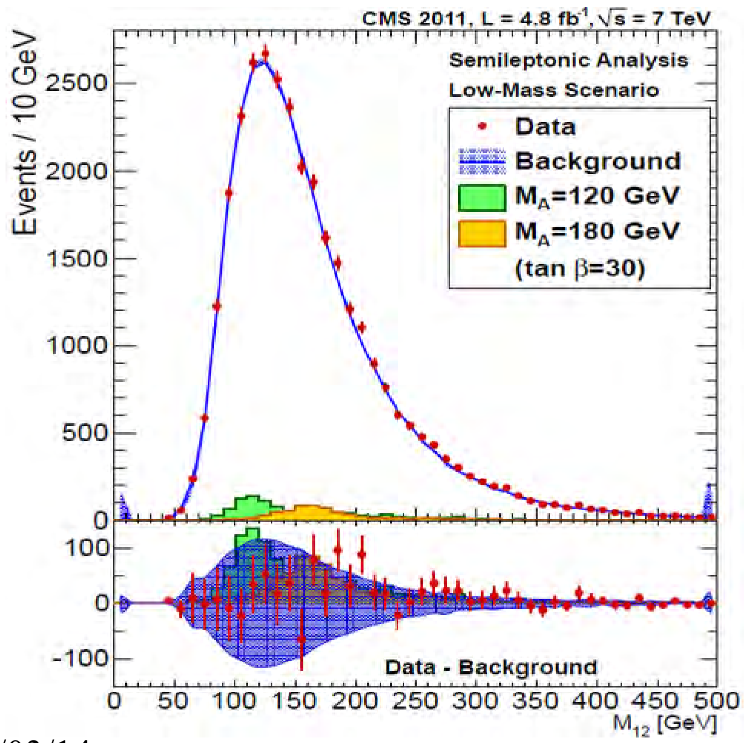


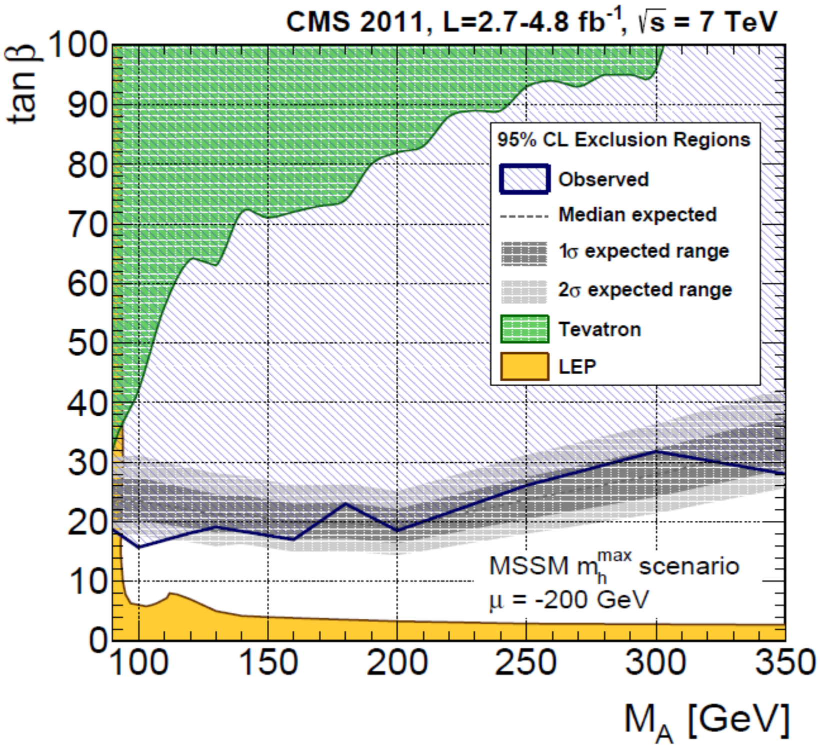


MSSM (Hbb)

- In MSSM at high tan(β)
  - The associated b-production is enanched
  - The decay to τ and b is favoured
- Dedicated H → bb search with pure QCD background

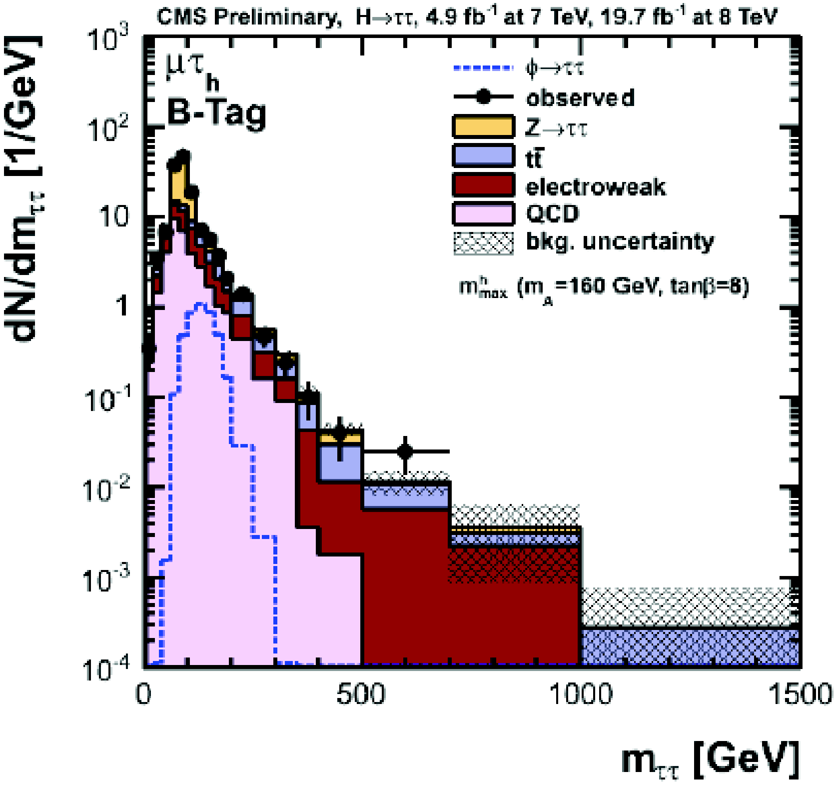


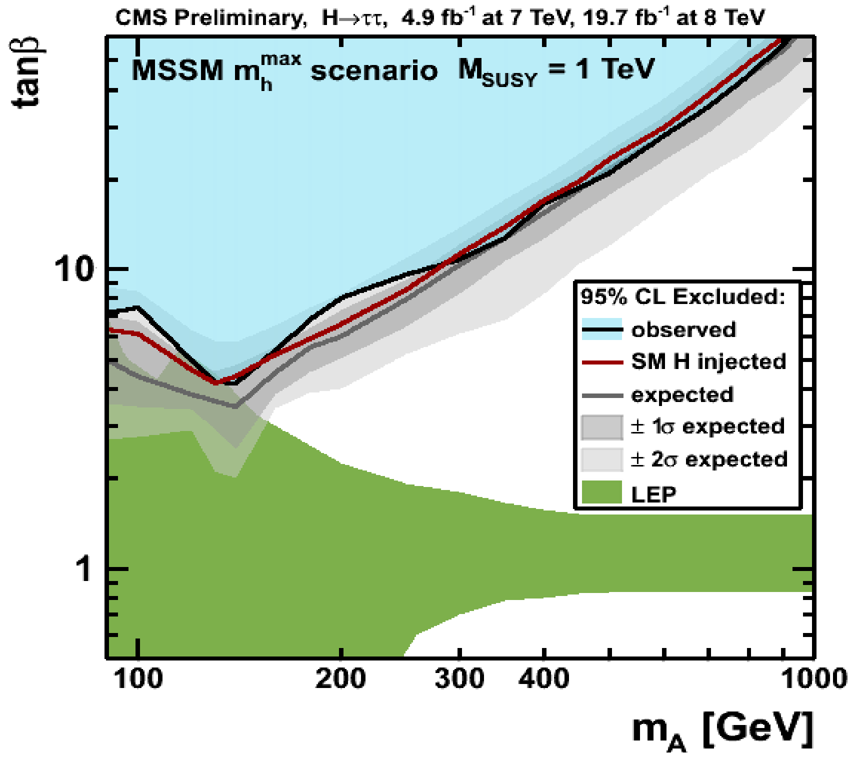




MSSM (Hττ)

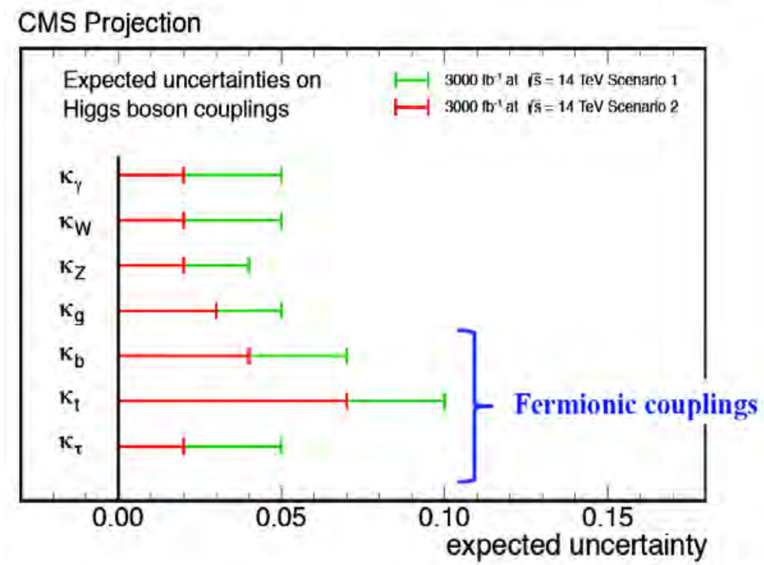
- Same search for H → τ τ
  - Including 2011 and 2012 data
- Use additional jet with b-tag category to increase S/B
- Large MSSM phase space excluded





Ideas and challenges for next runs

- What's next? Is the game over? (of course not!)
- Now that we know that this new particle exists and we know that it mostly decays to fermions we can use it as a probe for new physics (this is doable, and being done, with run1 data)
  - Search the H→bb or H->ττ resonance in SUSY final state
  - Search resonances of X->HH (w/ final state 4b, ττbb, γγbb...)
  - Search for H→bb in high pt jets (boosted search, subjets techniques)
- Precision measurement of SM Higgs and measurement of self coupling with larger datasets
  - Preparation for next run (300/fb,14TeV)
    - Boosted scenarios
    - Larger PU
    - Higher energies
  - Longer term (3000/fb)



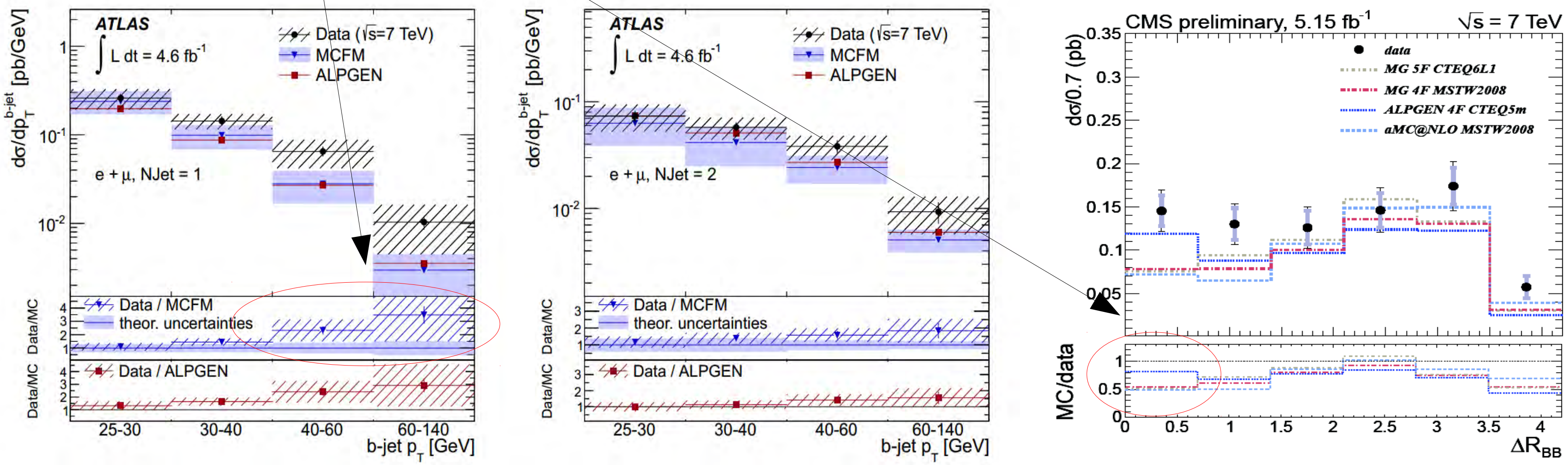
Conclusions

- The Higgs to b-quarks decay is being studied in at least three different channels
  - Best sensitivity in VH ~ 2σ
  - CMS recently added VBF to the family of Hbb studies
- Higgs to τ pairs studied in all production modes
  - Combined result well above 3σ sensitivity
- Combined fermion search is well compatible with Standard Model Higgs-to-fermions coupling (0.83+-0.24)
- A lot of road ahead to achieve precision measurements and using the freshly discovered particle as a “tool” to search for new physics

Back up

Are we ready for the 100/fb and above?

- What we may need from theorists:
  - Background uncertainties are probably more relevant than those on the signal
    - ..but a precise understanding of the pt spectrum for VH is needed
  - tt+jj and tt+bb backgrounds are important for ttH
    - In particular the “tt+1b” (gluon splitting with 1 soft or collinear b) has large uncertainties
  - We would benefit from more studies of NLO generators and gluon splitting tuning in generators (in general, not just in tt+b)
  - The 1b and/or small angle regions showed disagreement in recent measurement from Atlas and CMS



Are we ready for the 100/fb and above?

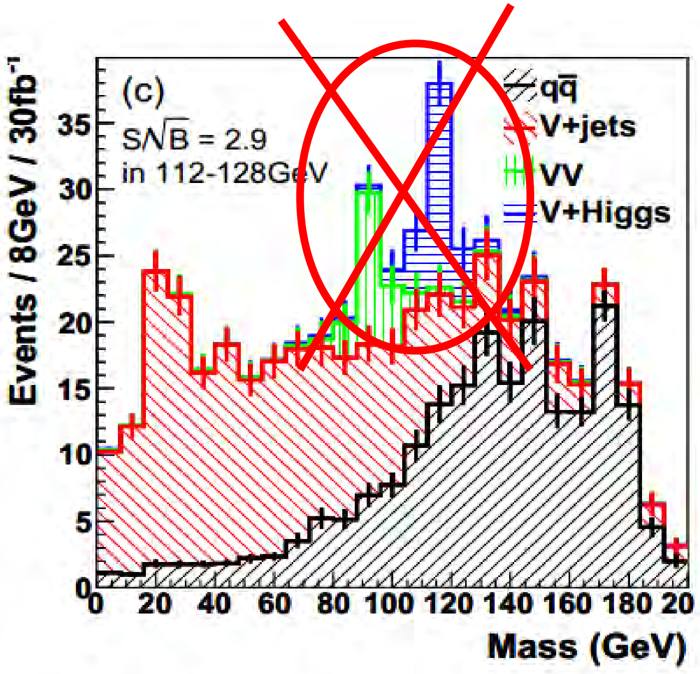
- Luminosity scaling
  - In VH, S/B is at most ~ 1/6
  - MC predictions becoming systematically limited?
    - More stat in the sidebands
    - Less extrapolations
    - Use generic templats (smooth shapes) instead of MC shapes
  - 450M MC events used for 20/fb, we cannot produce a factor of 10 more....

Scaling with sqrt(s) and PU

- ttbar xsec grows faster than VH xsec!
  - Already seen in 7->8 TeV
  - Z->nunu & W->ln have large ttbar background
  - “additional jets” used to cut ttbar are affected by PU
  - Z->ll on the other hand stays clean
- ttH xsec grows faster than ttbar
  - ttH should increase the sensitivity
- VBF, H->bb
  - More rapidity gap for the tag jets
  - ...but also more QCD
  - Trigger becoming really a challenge?

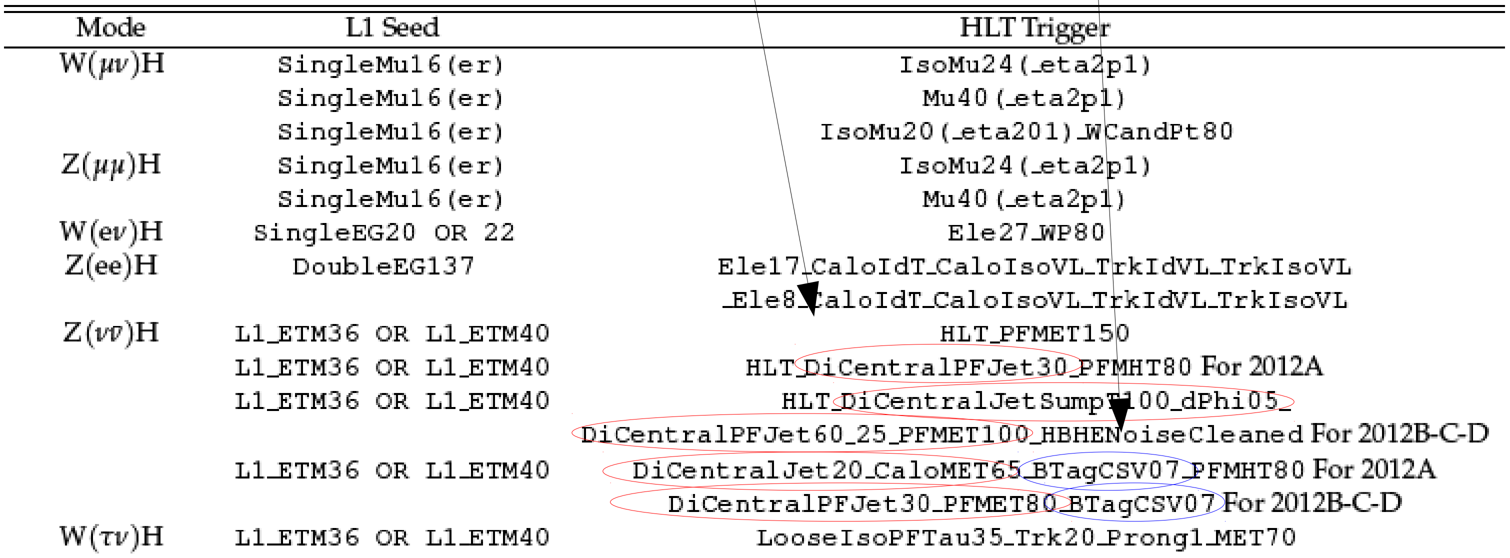
Substructures

- And how about substructures?
  - Jet merging really happens only for pT > 400 GeV
  - No benefit from substructure in current regime (jets are always well separated)
    - The few GeV resolution seen at 200 GeV in theory papers is not there in full simulation studies
  - On the other hand, at 13 TeV
    - Larger number of high boost events
    - The fraction of merged jets could be significant
    - Substructure are likely need in the high boost regime



Triggers VHbb

- Triggers are mostly based on the W/Z
  - i.e. leptons and MET
- Higgs decay product (di-jets or even btag) are only exploited for the medium-low pT region of ZH->nunubb
- All efficiencies are data driven (turn-on curves from prescaled triggers)



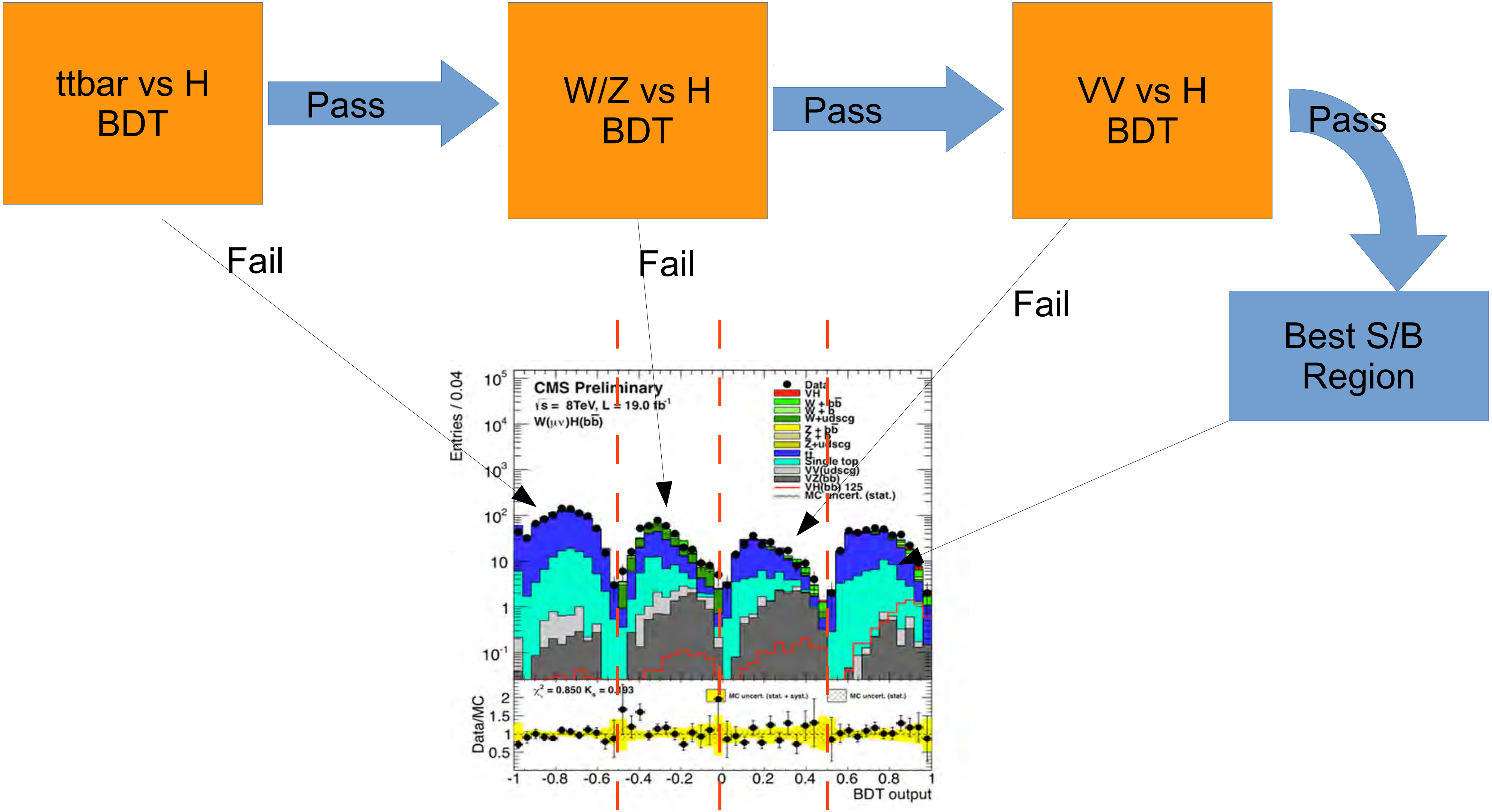
Mode	L1 Seed	HLT Trigger
W($\mu\nu$)H	SingleMu16(er)	IsoMu24(_eta2p1)
	SingleMu16(er)	Mu40(_eta2p1)
	SingleMu16(er)	IsoMu20(_eta201)_WCandPt80
Z($\mu\mu$)H	SingleMu16(er)	IsoMu24(_eta2p1)
	SingleMu16(er)	Mu40(_eta2p1)
W($e\nu$)H	SingleEG20 OR 22	Ele27_WP80
Z(ee)H	DoubleEG137	Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL _Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL
Z($\nu\bar{\nu}$)H	L1_ETM36 OR L1_ETM40	HLT_PFMET150
	L1_ETM36 OR L1_ETM40	HLT_DiCentralPFJet30_PFMHT80 For 2012A
	L1_ETM36 OR L1_ETM40	HLT_DiCentralJetSumpT100_dPhi05_ DiCentralPFJet60_25_PFMET100_HBHENoiseCleaned For 2012B-C-D
	L1_ETM36 OR L1_ETM40	DiCentralJet20_CaloMET65_BTagCSV07_PFMHT80 For 2012A DiCentralPFJet30_PFMET80_BTagCSV07 For 2012B-C-D
W($\tau\nu$)H	L1_ETM36 OR L1_ETM40	LooseIsoPFTau35_Trk20_Prong1_MET70

Analysis strategy

- Each mode (ll,lnu,nunu) has a dedicated analysis optimization, but the overall schema is common
  - Categorize the analysis in pT bins (3 bins with boundaries optimized in each analysis, typically around 100~200 GeV)
  - Use a jet energy regression to improve the signal shape
  - Estimate the backgrounds in control regions
  - Train an MVA with all discriminating variables (including the mass)
  - Shape fit on the MVA output
- As cross check also a non MVA analysis has been performed
  - Keep pT categories
  - Cut based selection on b-tag and few other variables
  - Use di-jet mass for the shape fit

Multi BDT

- Use 3 dedicated BDT to categorize the events
- Glue together the “overall BDT” for the 4 resulting categories



ttbar vs H BDT

Pass

W/Z vs H BDT

Pass

VV vs H BDT

Pass

Best S/B Region

Fail

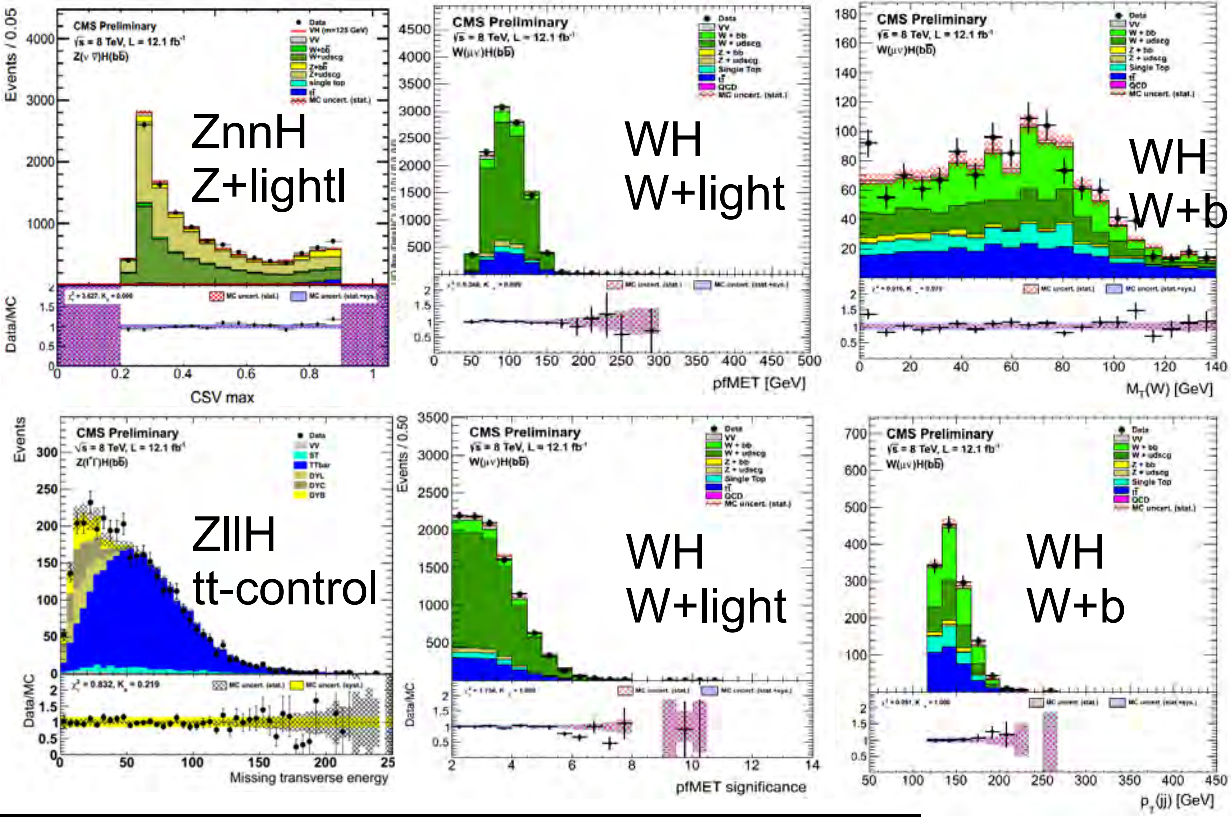
Fail

Fail

Control Regions – Scale Factors

- For each channel several control regions defined
- Shapes of all variables tested *data vs MC*
- Scale Factors for yields normalization
  - Used as starting value (with uncertainty) for nuisance parameters in the final fit

A small subset of checked variables



Scale Factors

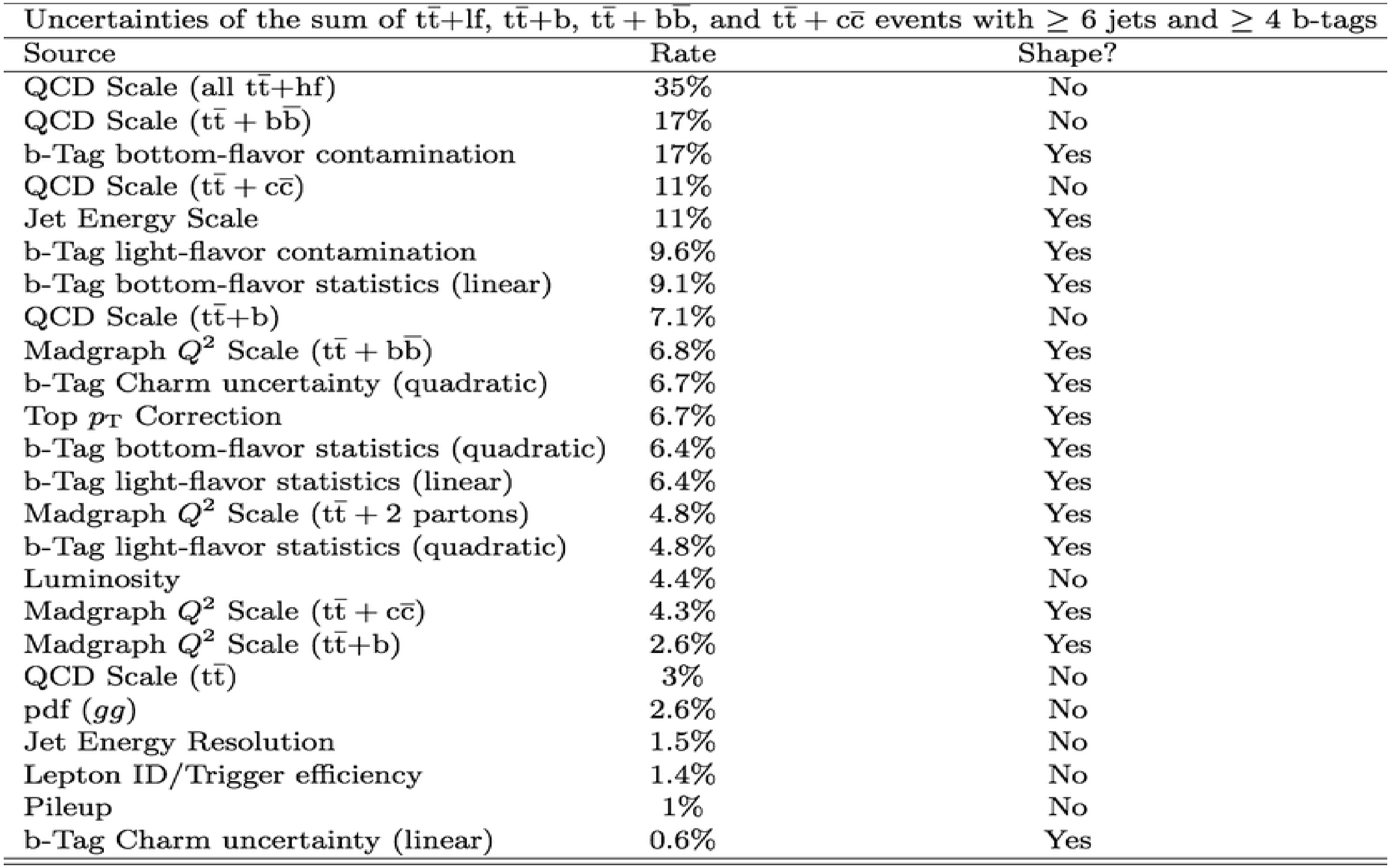
Process	W(ℓν)H	W(ℓν)H	Z(ℓℓ)H	Z(ℓℓ)H	Z(νν)H	Z(νν)H
Low p_T	7 TeV	8 TeV	7 TeV	8 TeV	7 TeV	8 TeV
W + udscg	$0.88 \pm 0.01 \pm 0.03$	$1.00 \pm 0.02 \pm 0.01$	–	–	$0.89 \pm 0.01 \pm 0.03$	$0.96 \pm 0.06 \pm 0.03$
Wb$\bar{b}$	$1.91 \pm 0.14 \pm 0.31$	$2.00 \pm 0.15 \pm 0.10$	–	–	$1.36 \pm 0.10 \pm 0.15$	$1.30 \pm 0.17 \pm 0.10$
Z + udscg	–	–	$1.11 \pm 0.03 \pm 0.11$	$1.06 \pm 0.03 \pm 0.07$	$0.87 \pm 0.01 \pm 0.03$	$1.15 \pm 0.07 \pm 0.03$
Zb$\bar{b}$	–	–	$0.98 \pm 0.05 \pm 0.12$	$1.04 \pm 0.05 \pm 0.08$	$0.96 \pm 0.02 \pm 0.03$	$1.12 \pm 0.10 \pm 0.04$
t$\bar{t}$	$0.93 \pm 0.02 \pm 0.05$	$1.07 \pm 0.01 \pm 0.01$	$1.03 \pm 0.04 \pm 0.11$	$0.95 \pm 0.04 \pm 0.10$	$0.97 \pm 0.02 \pm 0.04$	$1.05 \pm 0.07 \pm 0.03$
High p_T	7 TeV	8 TeV	7 TeV	8 TeV	7 TeV	8 TeV
W + udscg	$0.79 \pm 0.01 \pm 0.02$	$0.94 \pm 0.02 \pm 0.01$	–	–	$0.78 \pm 0.02 \pm 0.03$	$0.95 \pm 0.05 \pm 0.02$
Wb$\bar{b}$	$1.49 \pm 0.14 \pm 0.19$	$1.72 \pm 0.16 \pm 0.08$	–	–	$1.48 \pm 0.15 \pm 0.20$	$1.27 \pm 0.18 \pm 0.10$
Z + udscg	–	–	$1.11 \pm 0.03 \pm 0.11$	$1.06 \pm 0.03 \pm 0.07$	$0.97 \pm 0.02 \pm 0.04$	$1.04 \pm 0.07 \pm 0.02$
Zb$\bar{b}$	–	–	$0.98 \pm 0.05 \pm 0.12$	$1.04 \pm 0.06 \pm 0.08$	$1.08 \pm 0.09 \pm 0.06$	$1.15 \pm 0.10 \pm 0.04$
t$\bar{t}$	$0.84 \pm 0.02 \pm 0.03$	$0.98 \pm 0.01 \pm 0.01$	$1.03 \pm 0.04 \pm 0.11$	$0.95 \pm 0.04 \pm 0.10$	$0.97 \pm 0.02 \pm 0.04$	$1.03 \pm 0.07 \pm 0.03$

Systematic uncertainties VHbb

- The limit & significance are extracted with a shape analysis
- Systematic uncertainties are handled as nuisance parameters
- Where applicable a shape uncertainty is taken
  - B-tagging (doing discriminator re-shaping)
  - JEC/JER (variation within quoted uncertainties)
  - Background models (different generators)
  - Signal pt-spectrum (NNLO QCD and NLO EWK)
  - Trigger (measured turn-on uncertainties)
  - MC normalization (control region SF uncertainties)
  - Diboson and single top yields (xsec uncertainty)
- Different choices of nuisance parameterization tested to verify robustness of the shape analysis
- No particular concerns from post-fit nuisance pulls

Systematics ttH

- Dominant systematics:
  - tt+bb normalization
  - B-tag shape uncertainties
  - Jet Energy Scale

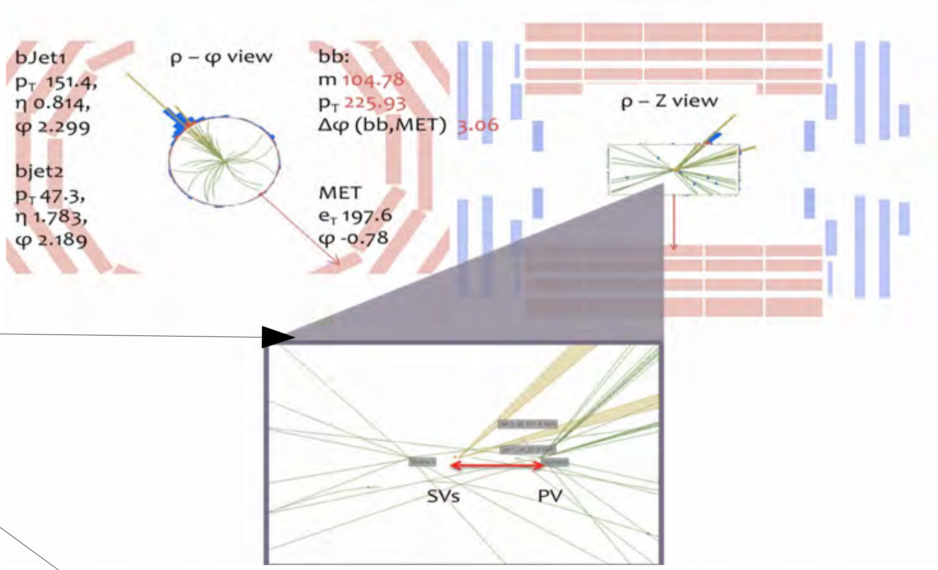


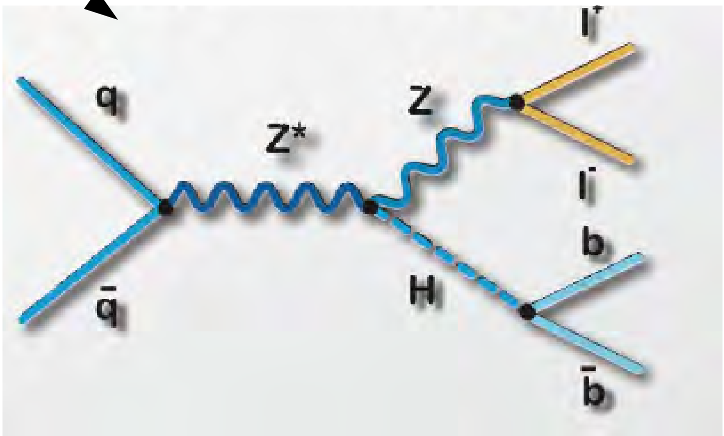
Uncertainties of the sum of $t\bar{t}$+lf, $t\bar{t}$+b, $t\bar{t}+b\bar{b}$, and $t\bar{t}+c\bar{c}$ events with ≥ 6 jets and ≥ 4 b-tags

Source	Rate	Shape?
QCD Scale (all $t\bar{t}$+hf)	35%	No
QCD Scale ($t\bar{t}+b\bar{b}$)	17%	No
b-Tag bottom-flavor contamination	17%	Yes
QCD Scale ($t\bar{t}+c\bar{c}$)	11%	No
Jet Energy Scale	11%	Yes
b-Tag light-flavor contamination	9.6%	Yes
b-Tag bottom-flavor statistics (linear)	9.1%	Yes
QCD Scale ($t\bar{t}$+b)	7.1%	No
Madgraph Q^2 Scale ($t\bar{t}+b\bar{b}$)	6.8%	Yes
b-Tag Charm uncertainty (quadratic)	6.7%	Yes
Top p_{T} Correction	6.7%	Yes
b-Tag bottom-flavor statistics (quadratic)	6.4%	Yes
b-Tag light-flavor statistics (linear)	6.4%	Yes
Madgraph Q^2 Scale ($t\bar{t}+2$ partons)	4.8%	Yes
b-Tag light-flavor statistics (quadratic)	4.8%	Yes
Luminosity	4.4%	No
Madgraph Q^2 Scale ($t\bar{t}+c\bar{c}$)	4.3%	Yes
Madgraph Q^2 Scale ($t\bar{t}$+b)	2.6%	Yes
QCD Scale ($t\bar{t}$)	3%	No
pdf (gg)	2.6%	No
Jet Energy Resolution	1.5%	No
Lepton ID/Trigger efficiency	1.4%	No
Pileup	1%	No
b-Tag Charm uncertainty (linear)	0.6%	Yes

VH, H to bb

- Associated production of Higgs to a vector boson
  - Several modes considered:
    - W->lnu (electron or muon)
    - Z-> nunu
    - Z->ll (electrons or muons)
- Decay of the Higgs boson in bb
  - Use b-tagging to identify the jets coming from the Higgs decay
- Backgrounds:
  - V+b-jets, ttbar, single top, VV
- Trigger with the lepton(s) from the V and/or MET

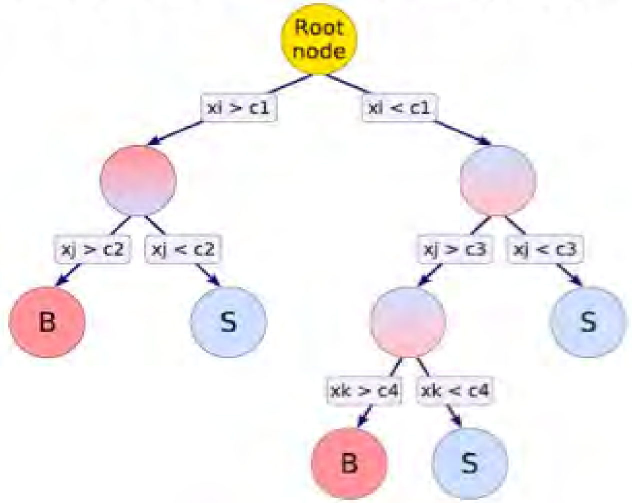




A ZH->llbb event candidate

Multi-Variate Analysis

- Apply loose preselection cuts and let and MVA increase the S/B
- Use a dozen input variables to train a Bosted Decision Tree
- Optionally train different BDTs for different backgrounds and split the final BDT in different regions



BOOSTED DECISION TREES

Preselection cuts

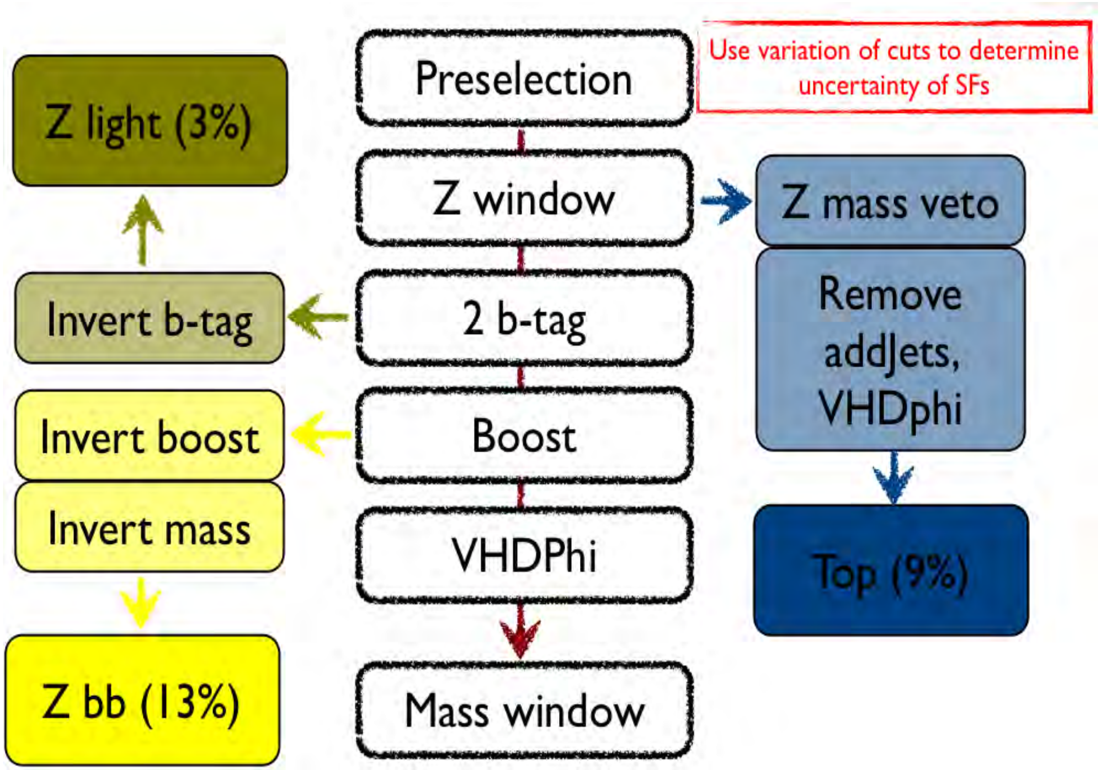
Variable	W($\ell\nu$)H	Z($\ell\ell$)H	Z($\nu\nu$)H
$m_{\ell\ell}$	–	[75 – 105]	–
$p_T(j_1)$	> 30	> 20	> 60
$p_T(j_2)$	> 30	> 20	> 30
$p_T(\mathrm{jj})$	> 120	–	> 130
$m(\mathrm{jj})$	< 250	[80 – 150] (< 250)	< 250
$p_T(\mathrm{V})$	[120 – 170] (> 170)	[50 – 100] (> 100)	–
$\mathrm{CSV}_{\mathrm{max}}$	> 0.40	> 0.50 (> 0.244)	> 0.679
$\mathrm{CSV}_{\mathrm{min}}$	> 0.40	> 0.244	> 0.244
$\mathrm{CSV}_{\mathrm{min}}^{loose}$	– (< 0.40)	–	– (< 0.244)
N_{al}	$= 0$	–	$= 0$
E_T^{miss}	> 45 (elec)	–	[130 – 170] (> 170)
$\Delta\phi(\mathrm{E_T^{miss}}, \mathrm{jet})$	–	–	> 0.5
$\Delta\phi(\mathrm{E_T^{miss}}, \mathrm{E_T^{miss\,(trks)}})$	–	–	< 0.5
$\Delta\phi(\mathrm{V}, \mathrm{H})$	–	–	> 2.0

BDT Input variables

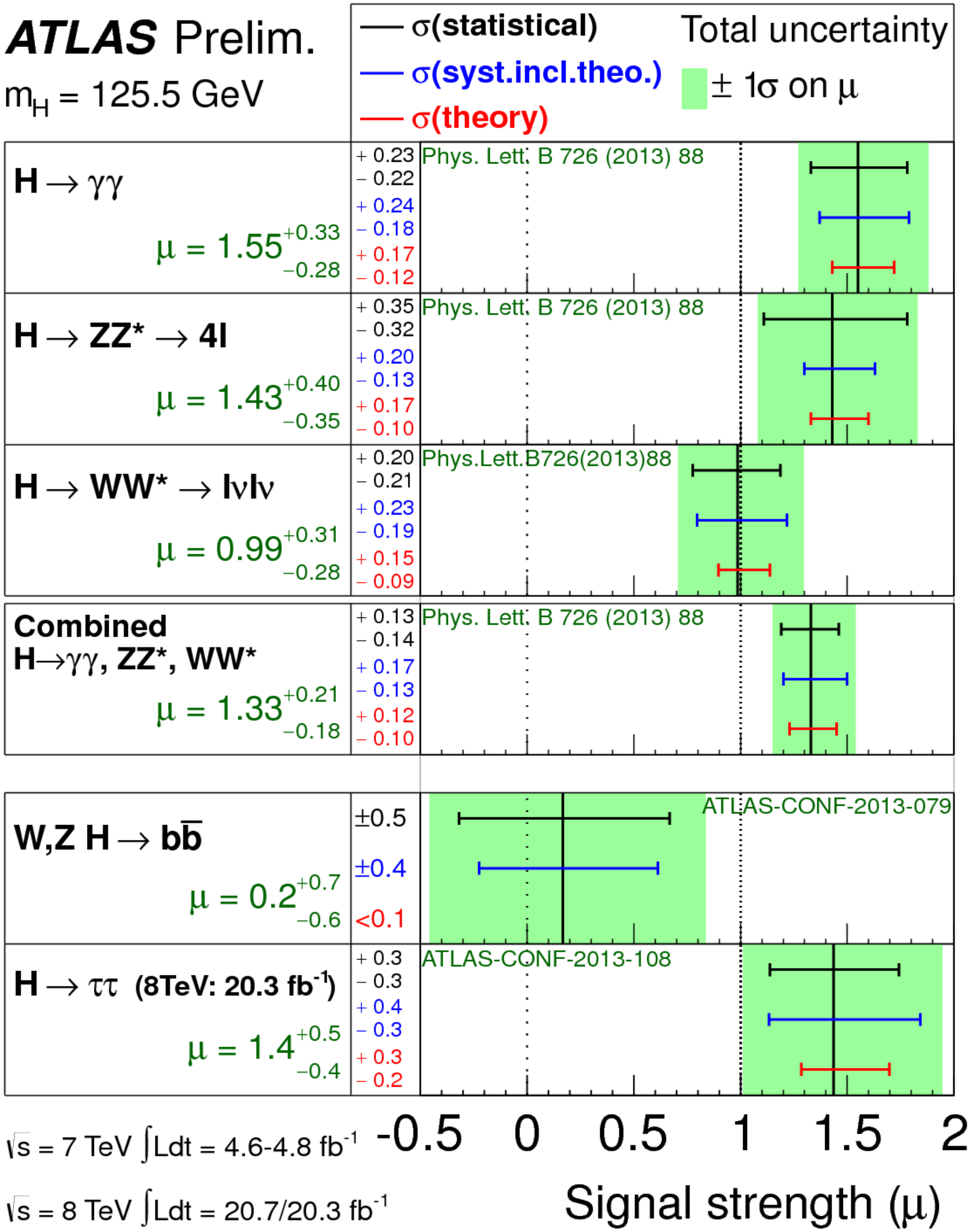
Variable
p_{Tj}: transverse momentum of each Higgs daughter
$m(\mathrm{jj})$: dijet invariant mass
$p_T(\mathrm{jj})$: dijet transverse momentum
$p_T(\mathrm{V})$: vector boson transverse momentum (or E_T^{miss})
$\mathrm{CSV}_{\mathrm{max}}$: value of CSV for the Higgs daughter with largest CSV value
$\mathrm{CSV}_{\mathrm{min}}$: value of CSV for the Higgs daughter with second largest CSV value
$\Delta\phi(\mathrm{V}, \mathrm{H})$: azimuthal angle between V (or E_T^{miss}) and dijet
$\|\Delta\eta(\mathrm{jj})\|$: difference in η between Higgs daughters
$\Delta R(\mathrm{jj})$: distance in η–ϕ between Higgs daughters
N_{aj}: number of additional jets
$\Delta\phi(\mathrm{E_T^{miss}}, \mathrm{jet})$: azimuthal angle between E_T^{miss} and the closest jet (only for Z($\nu\nu$)H)
$\Delta\theta_{\mathrm{pull}}$: color pull angle [35]

Control regions

- Cotrol regions are defined with several purpose:
  - Adjust MC prediction of main backgrounds (V+light,V+b,ttbar)
  - Verify BDT input variables distributions
  - Verify BDT input variable correlations
  - Verify BDT output distribution in signal free/depleted phase space
- Typical Control Region definition:
  - Same preselection as for signal
  - Invert some cuts
  - and/or apply mass window veto
- Perform a simultaneous fit of highly discriminating variables (e.g. btag) to extract data/MC scale factors



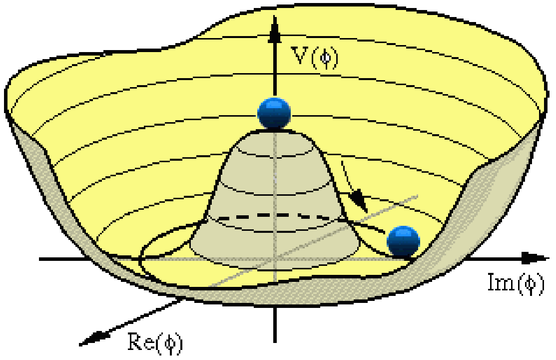
Comparison with atlas



- Tau Tau
  - Atlas 4.1 obs , 3.2 exp
  - CMS(all) 3.2 obs, 3.7 exp
  - CMS(noVH) 3.4obs, 3.6exp
- VH, BB
  - Atlas mu=0.2+0.7-0.6
  - CMS mu=1.0+-0.5

The Higgs Mechanism

- The Standard Model is a gauge theory with massless fermions
- Mass term cannot be simply “added” to the Lagrangian (not symmetric under the gauge groups)
- The Higgs mechanism allows to naturally break the electroweak symmetry
  - Introduce a complex doublet and a potential with minimum at non-zero value
  - Introduce Yukawa couplings of the new field to the fermions
  - This additional “Symmetry Breaking Sector” produces:
    - Masses for fermions and gauge bosons
    - A new particle, the Higgs boson, coupled to fermions and gauge bosons with strength proportional to their mass
- The Higgs boson mass is not predicted
  - But constraints comes from precision measurements of SM processes being the only free parameter



$$\mathcal{L}_{SBS} = (D_\mu\phi)^\dagger(D^\mu\phi) - V(\phi) + \mathcal{L}_{YK}$$