The Higgs boson decaying to b-quarks

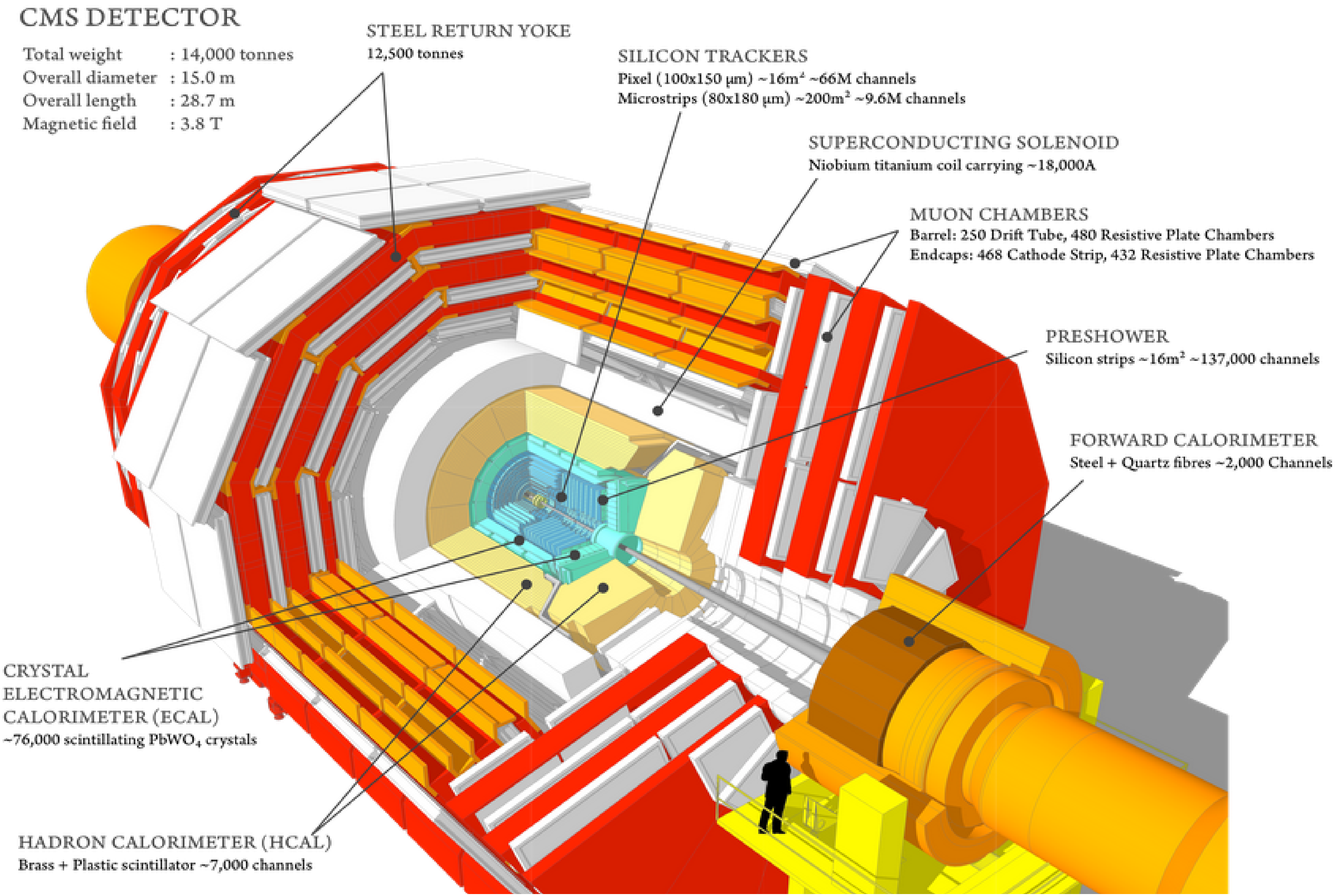
Andrea Rizzi, INFN e Universita' di Pisa

4 October 2013, Orsay-LAL, Paris

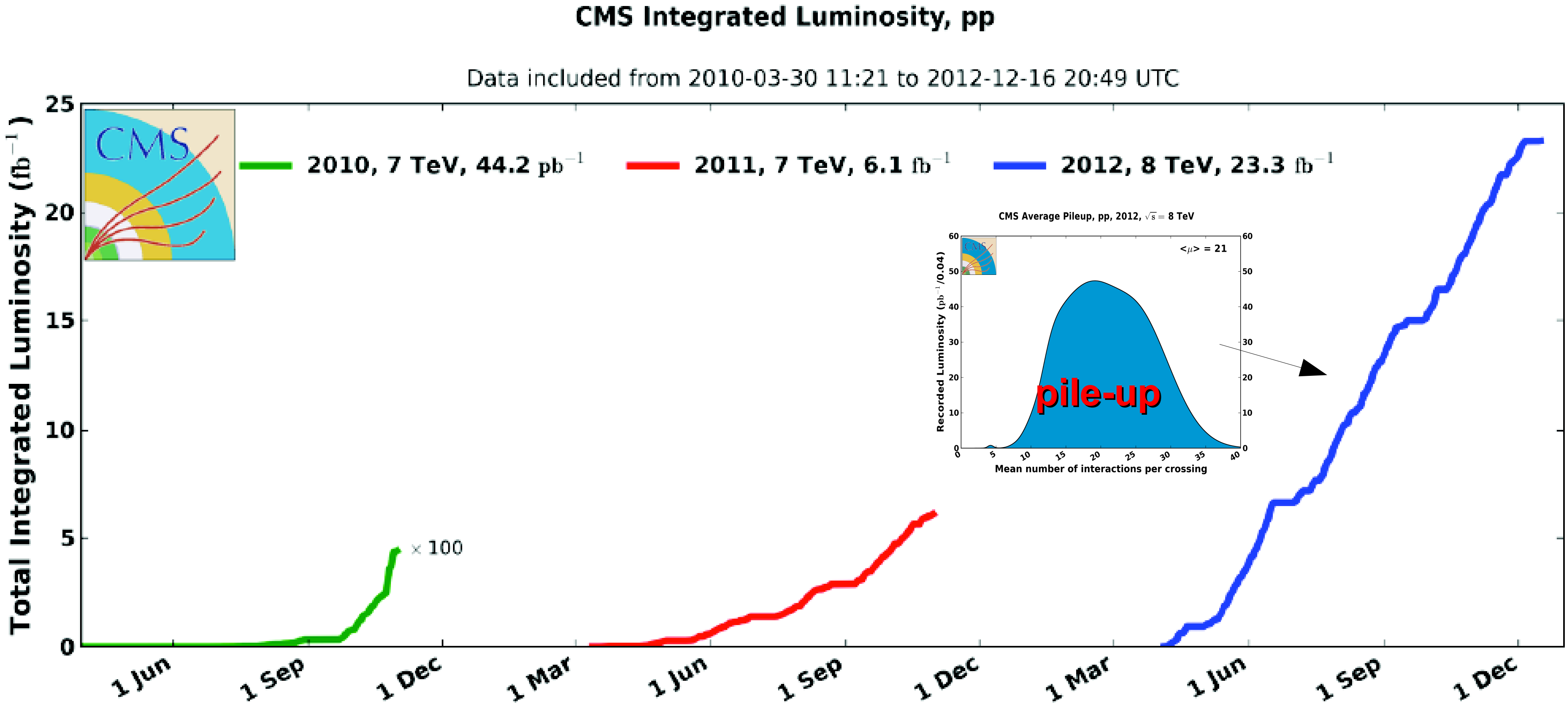
Outline

- CMS and LHC
- Overview of the Higgs results
- The Higgs to BB modes:
  - VH
  - VBF
  - ttH
- Discussion on next LHC runs

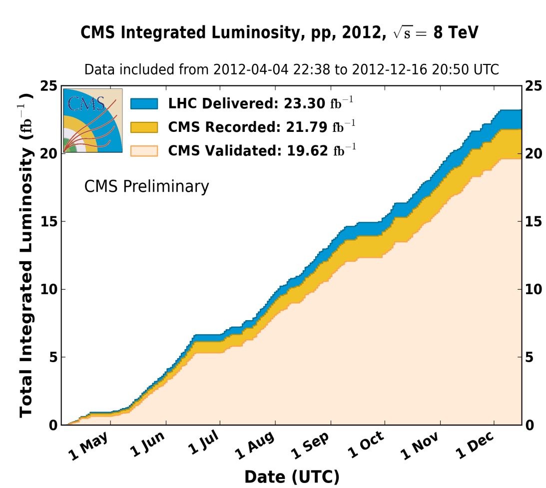
The CMS experiment

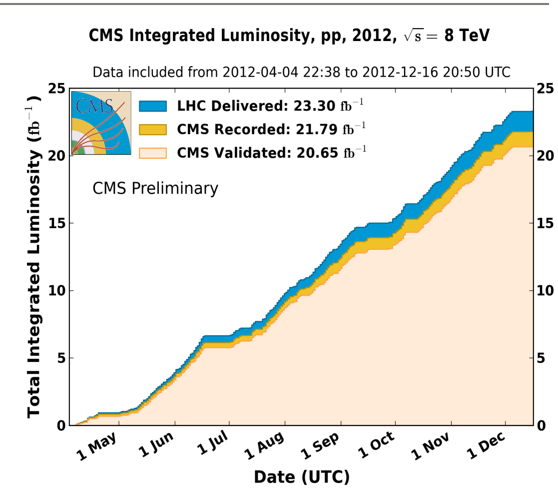


LHC 2011 & 2012



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



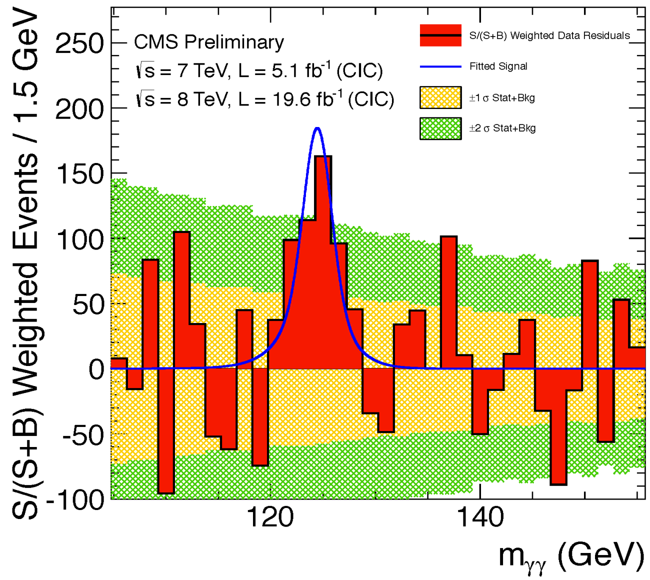


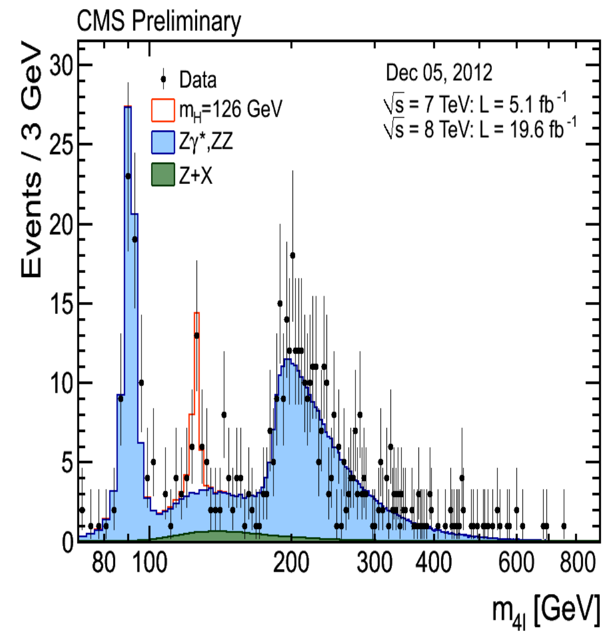
Higgs overview

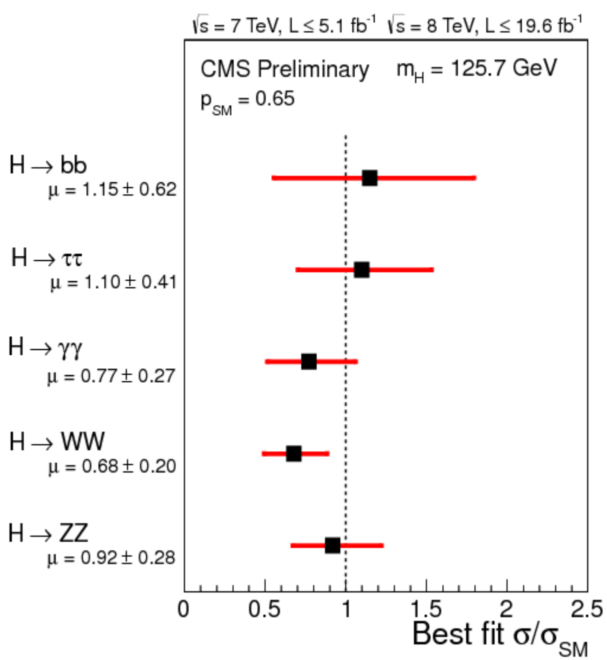
- Higgs signal well visible in gammagamma and ZZ
- Mass:

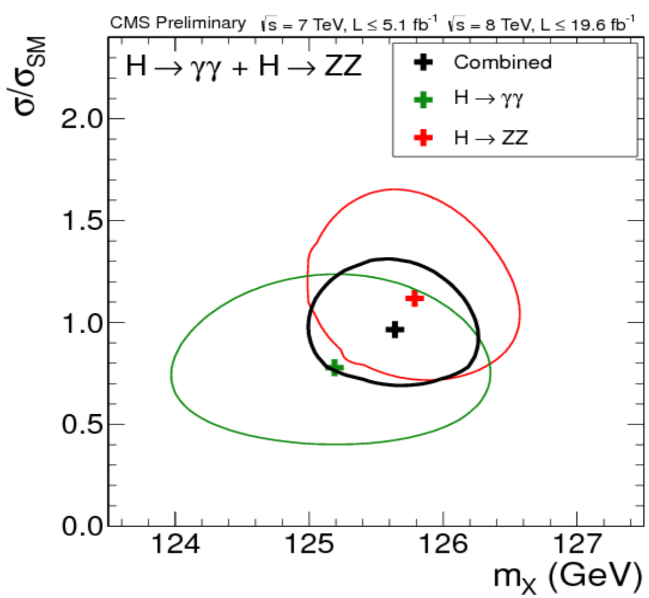
  125.7± 0.3(stat)± 0.3(syst)

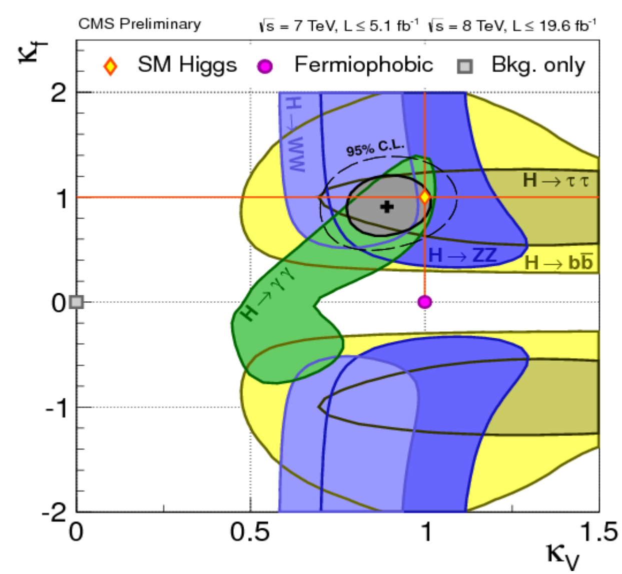
- How about fermions?



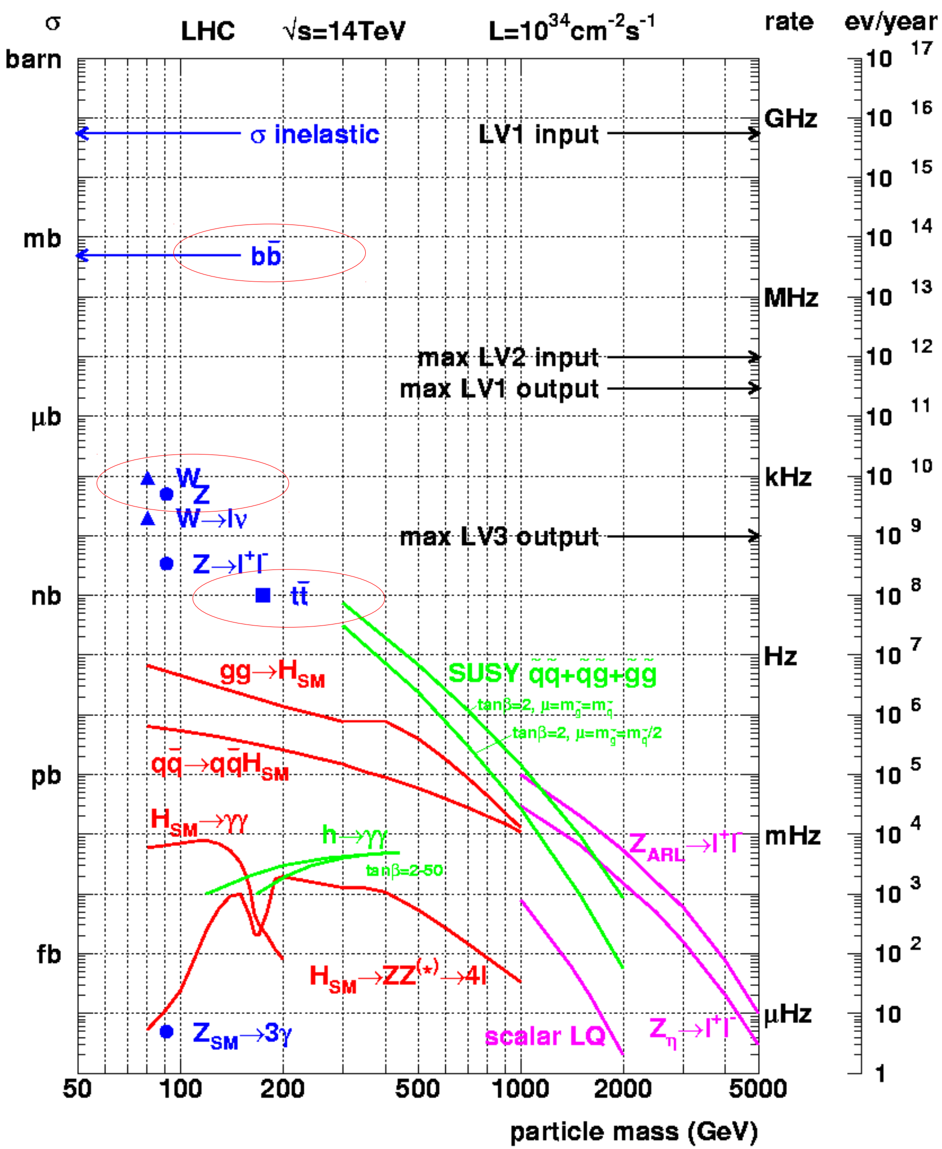


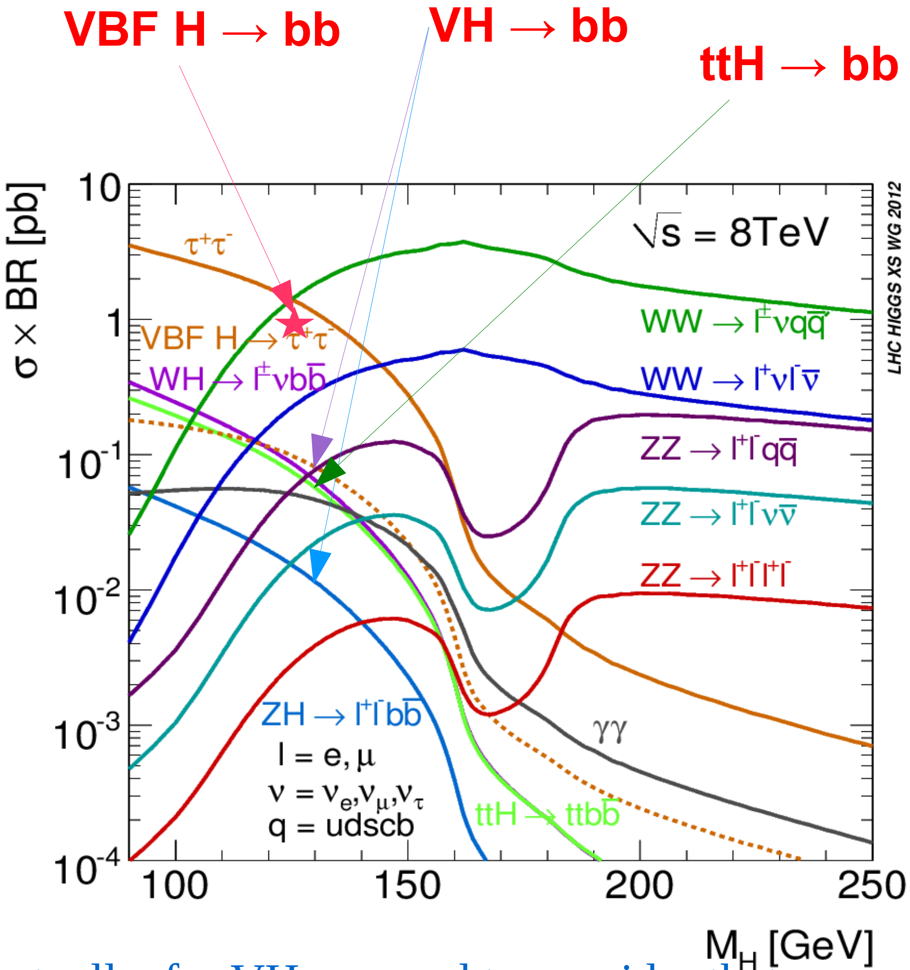






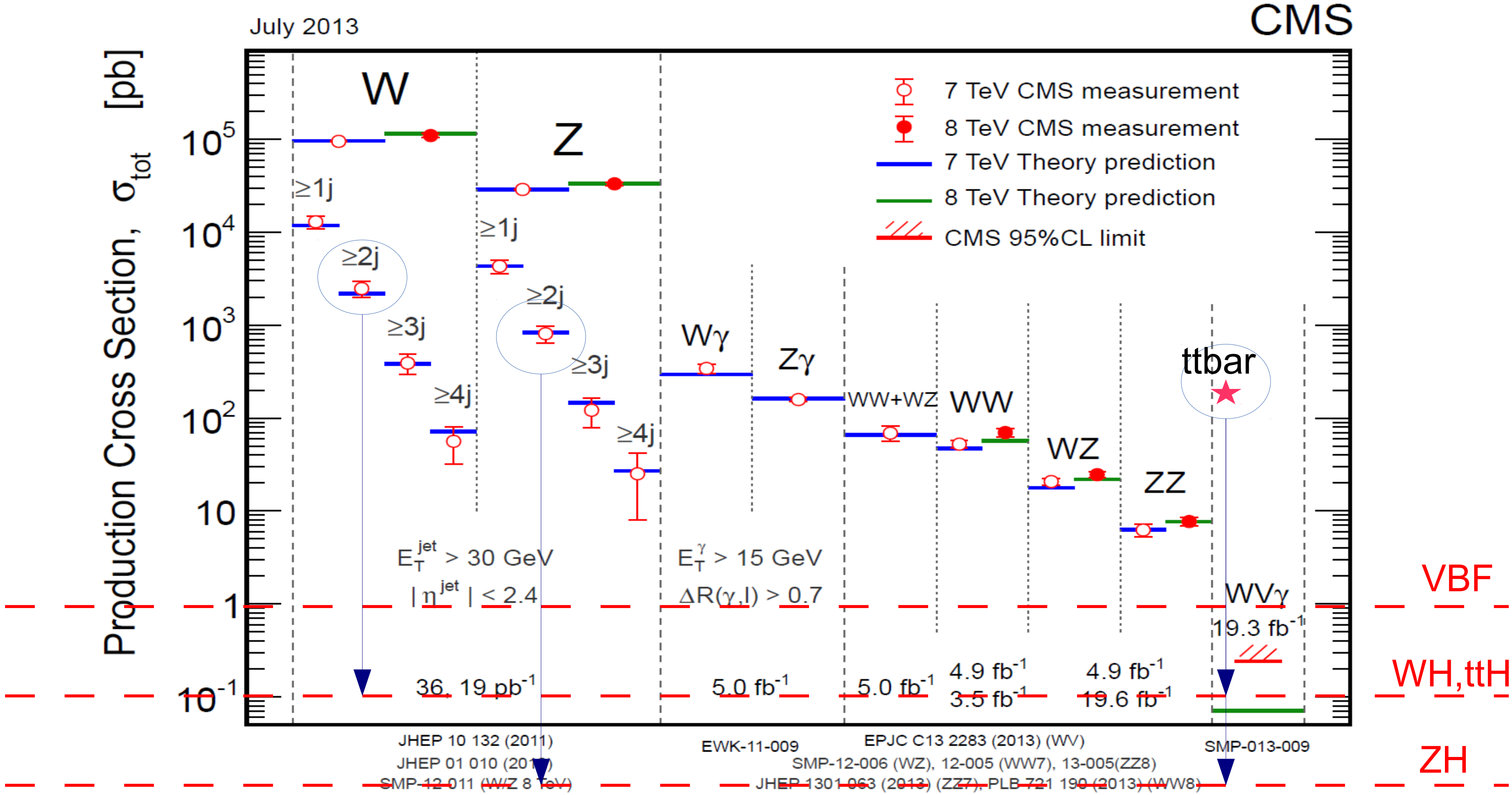
The cross sections





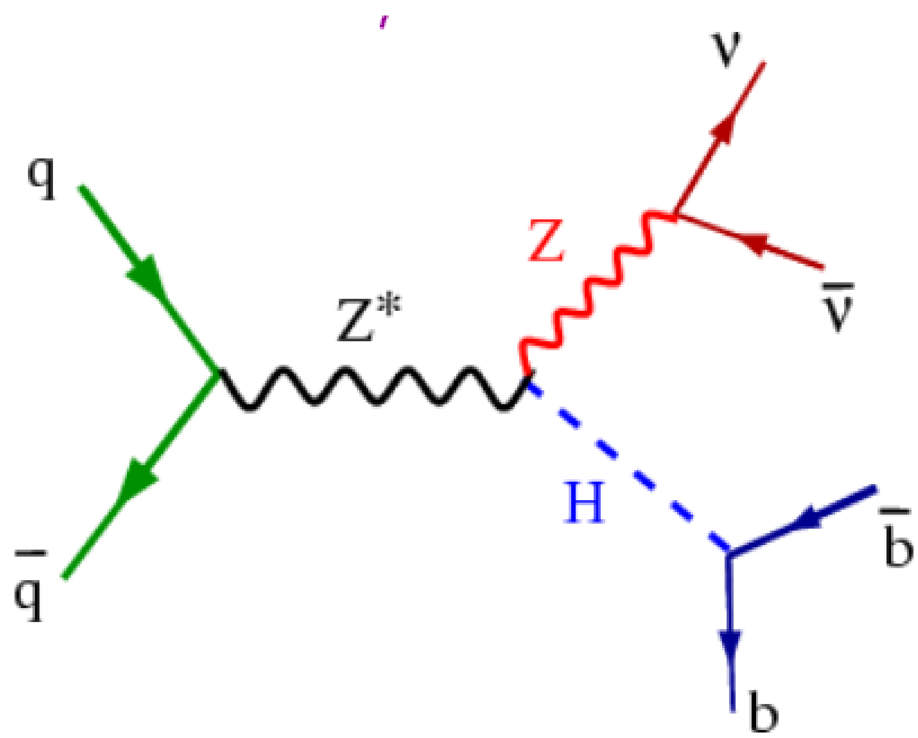
- Actually, for VH we need to consider that:
  - It is only studied in lnu,ll,nunu (x10%)
  - The relevant phase space is high pt (x 10%)
- In fact, also the backgrounds should be scaled down requiring an additional pair of jets (tt+jj, V+jj, bbjj, etc..)

Summary of SM measurements



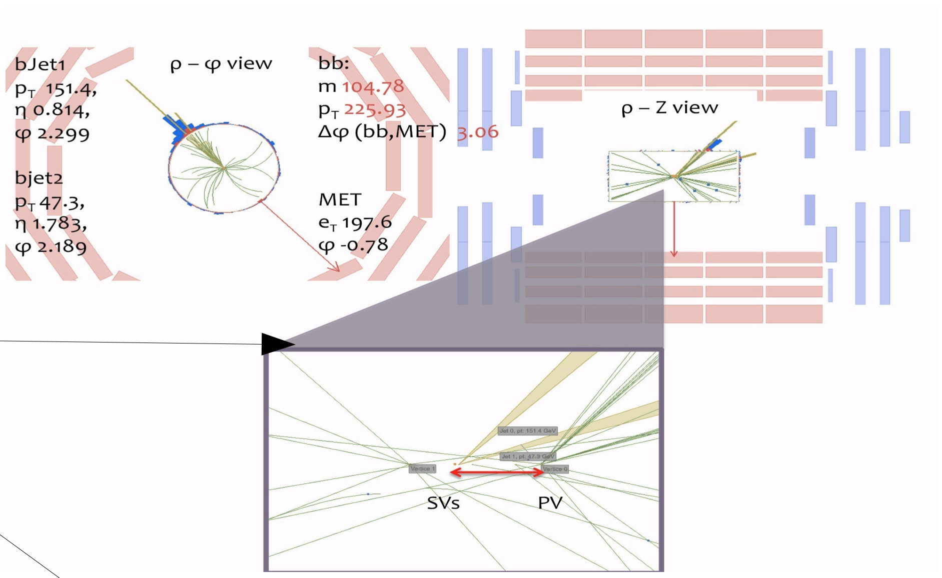
- About 3-4 orders of magnitude before cuts

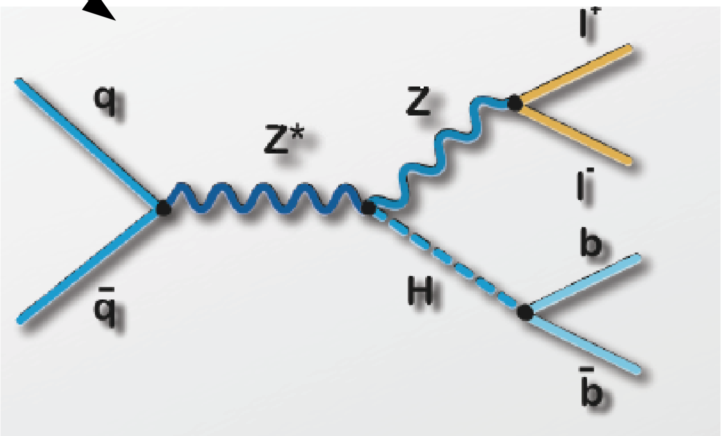
VH, H->bb



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

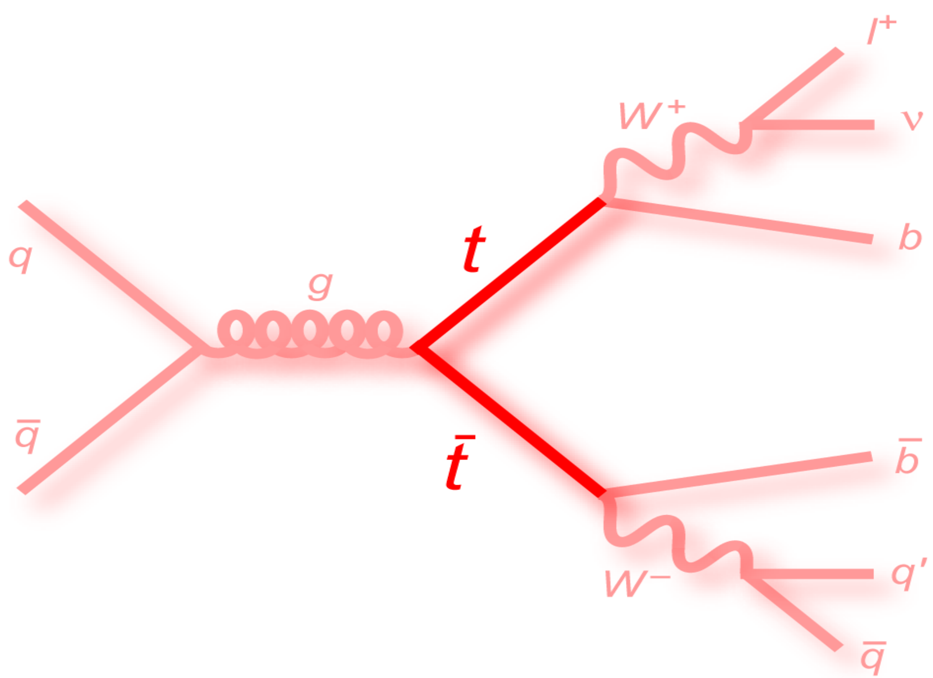




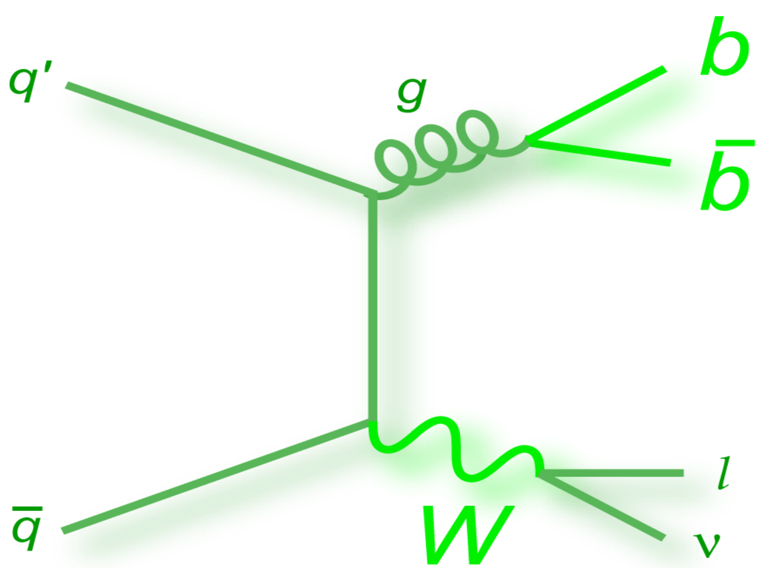
A ZH->llbb event candidate

Backgrounds

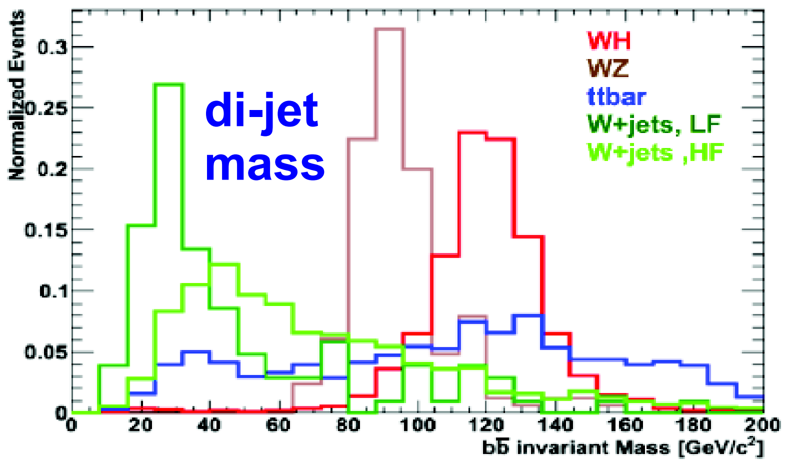
Reducible backgrounds
QCD, V+udscg (“light” jets)
ttbar and single top
=> reduced with b-tag, jet counting, additional leptons, lepton isolation

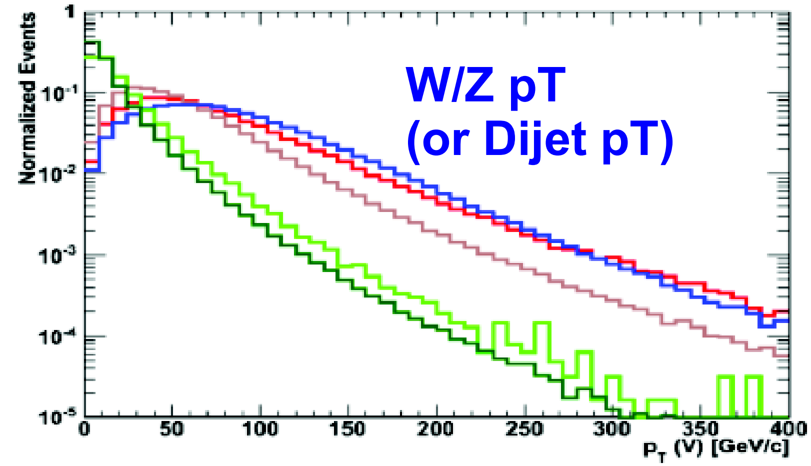


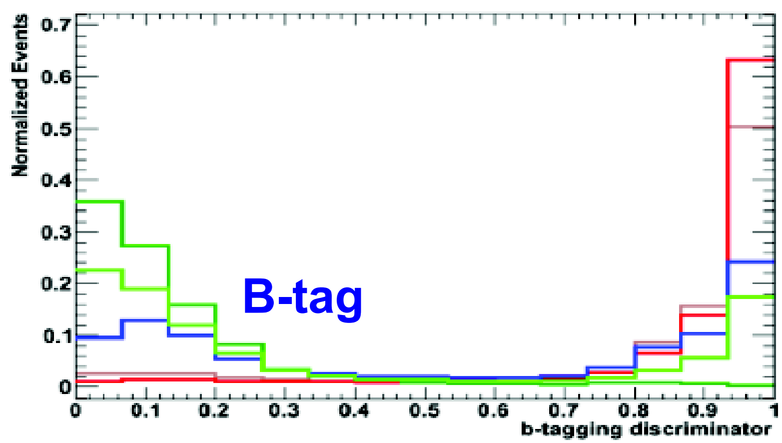
Less reducible backgrounds
V+bb
ZZ(bb), W(lν)Z(bb)
=> bb mass is the only handle

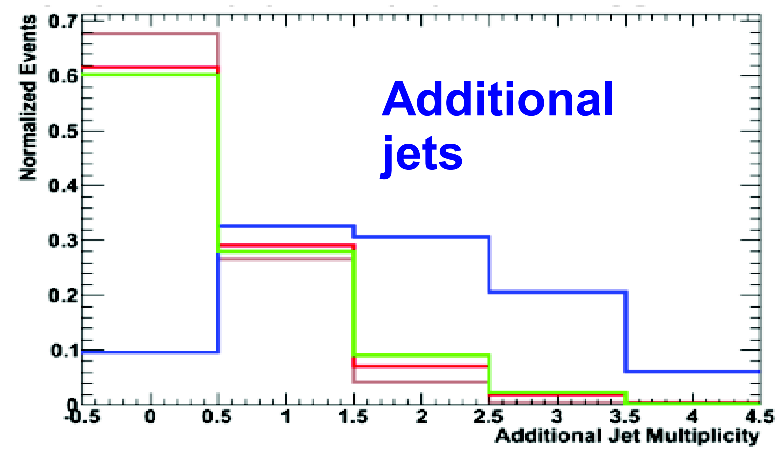


The main handles









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

Triggers

- 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)

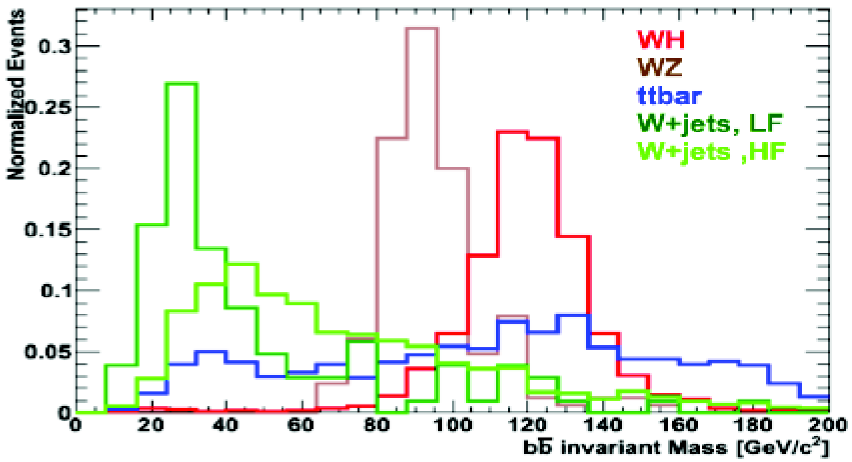
Mode	L1 Seed	HLT Trigger
W($\mu\nu$)H	SingleMu16(er)	IsoMu24(_eta2p1)
	SingleMu16(er)	Mu40(_eta2p1)
	SingleMu16(er)	IsoMu20(_eta201)_WCandPt80
Z($\mu\mu$)H	SingleMu16(er)	IsoMu24(_eta2p1)
	SingleMu16(er)	Mu40(_eta2p1)
W($e\nu$)H	SingleEG20 OR 22	Ele27_WP80
Z(ee)H	DoubleEG137	Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL
		_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

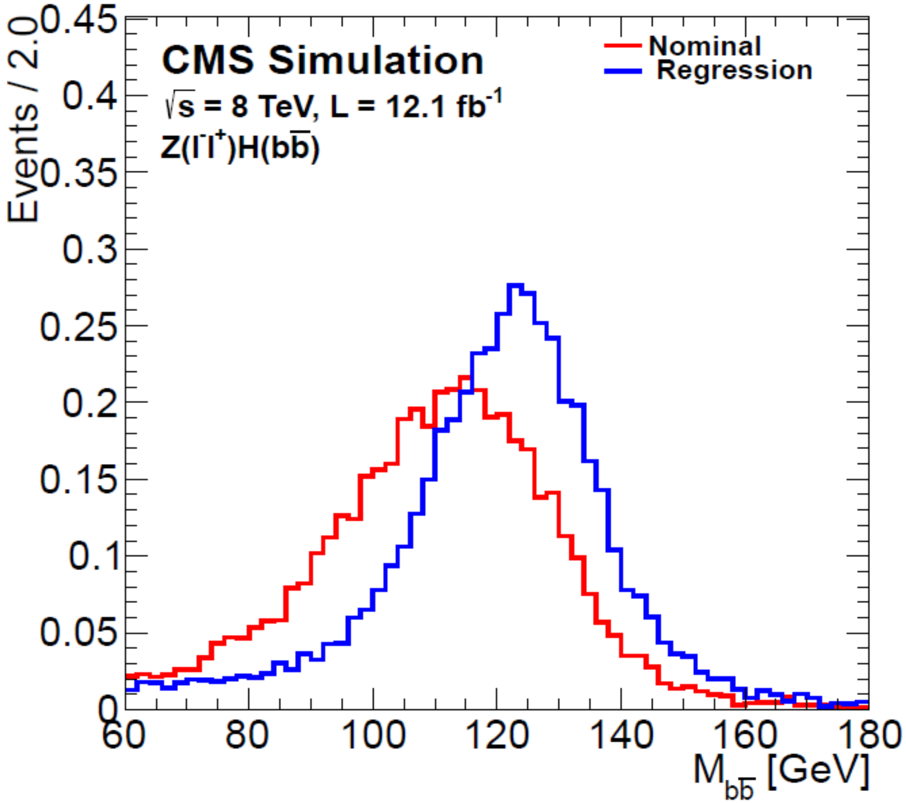
Analysis strategy

- 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

Jet energy regression

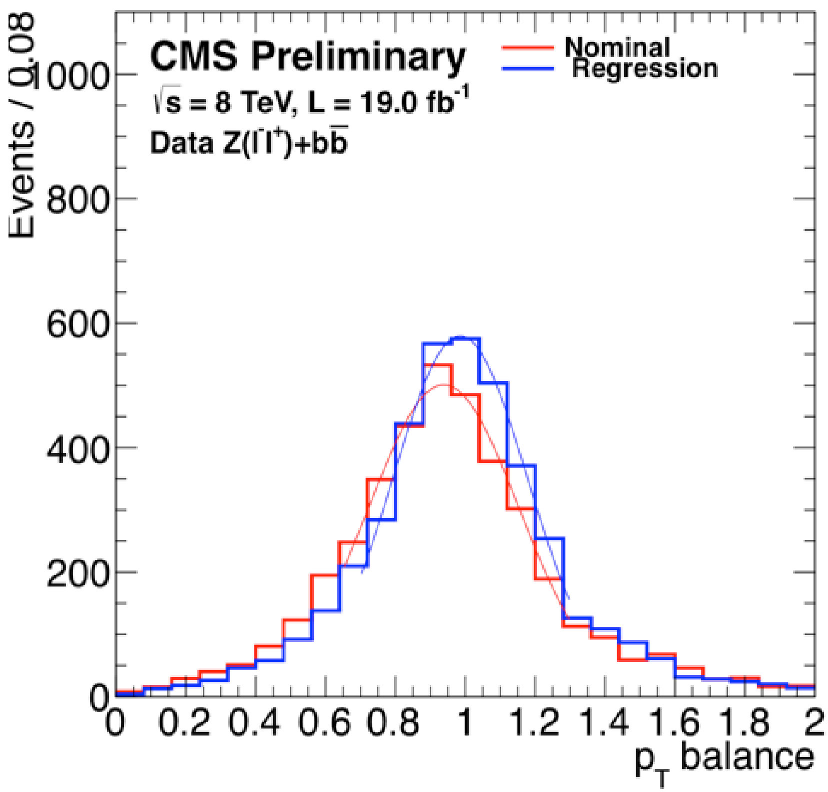
- The dijet mass is the most discriminating variable
- Its resolution depends on jets resolution
- b-jets are not like light jets
  - Presence of leptons and neutrinos
  - More massive (hence broader)
  - They can be "Tagged" with lifetime and secondary vertices
- Use a BDT regression in order to correct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

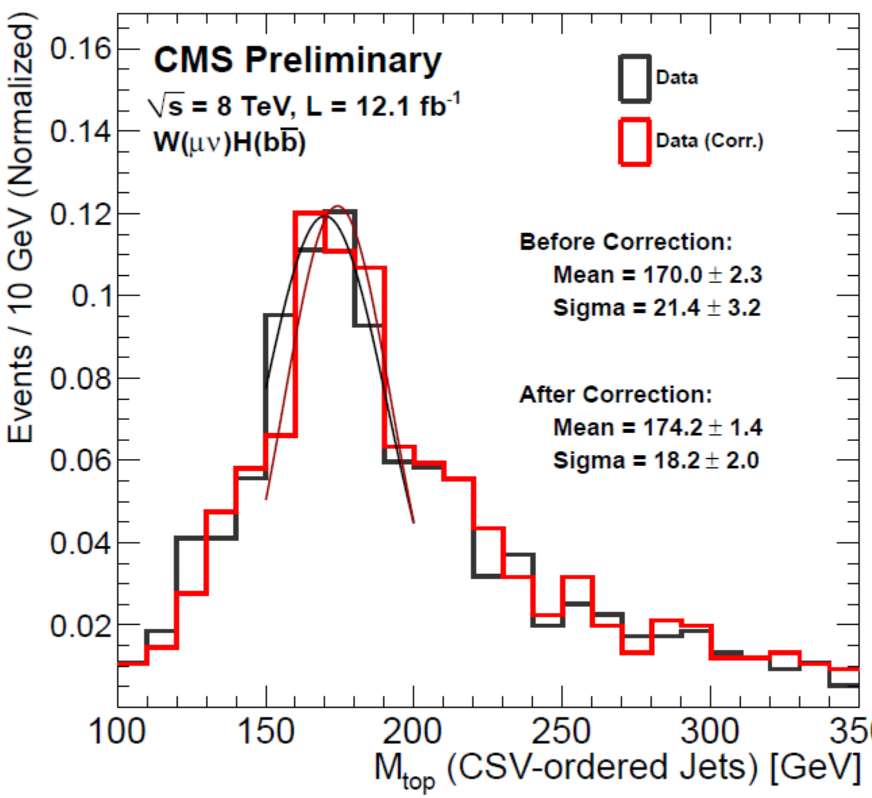




Jet energy regression

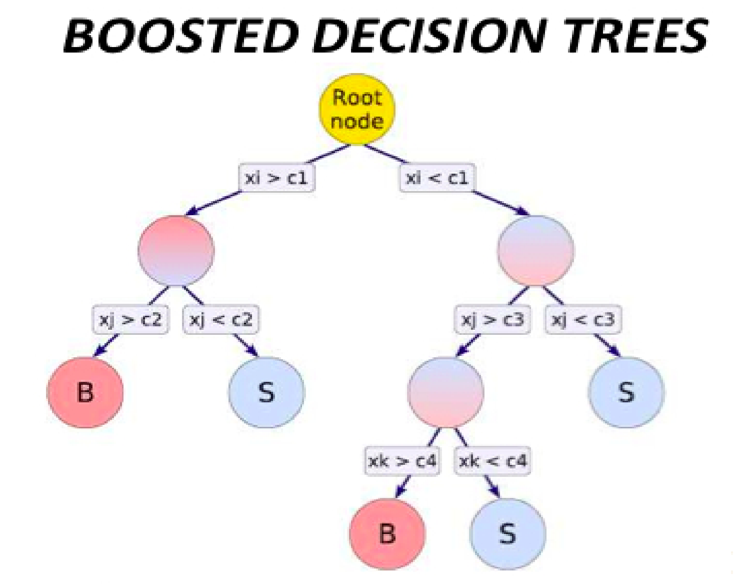
- The regression technique has been validated on data
  - pT balance in a Z+2b jets sample (Z->ll)
  - Top mass in a top enriched region
- In both cases the observed improvements matches the MC expectations





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



Preselection cuts

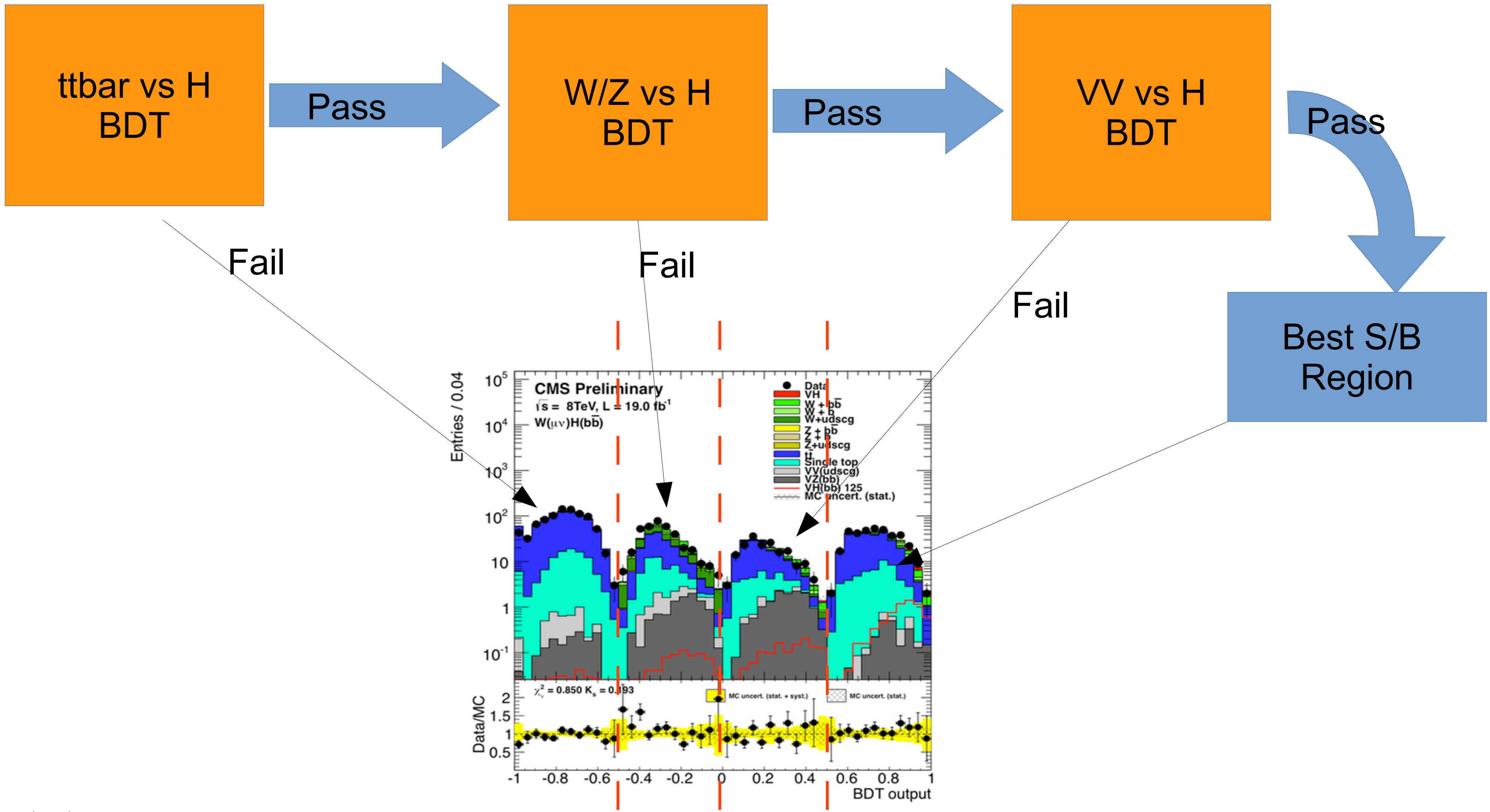
Variable	$W(\ell\nu)H$	$Z(\ell\ell)H$	$Z(\nu\nu)H$
$m_{\ell\ell}$	–	[75 – 105]	–
$p_T(j_1)$	> 30	> 20	> 60
$p_T(j_2)$	> 30	> 20	> 30
$p_T(jj)$	> 120	–	> 130
$m(jj)$	< 250	[80 – 150] (< 250)	< 250
$p_T(V)$	[120 – 170] (> 170)	[50 – 100] (> 100)	–
CSV_{max}	> 0.40	> 0.50 (> 0.244)	> 0.679
CSV_{min}	> 0.40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

BDT Input variables

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

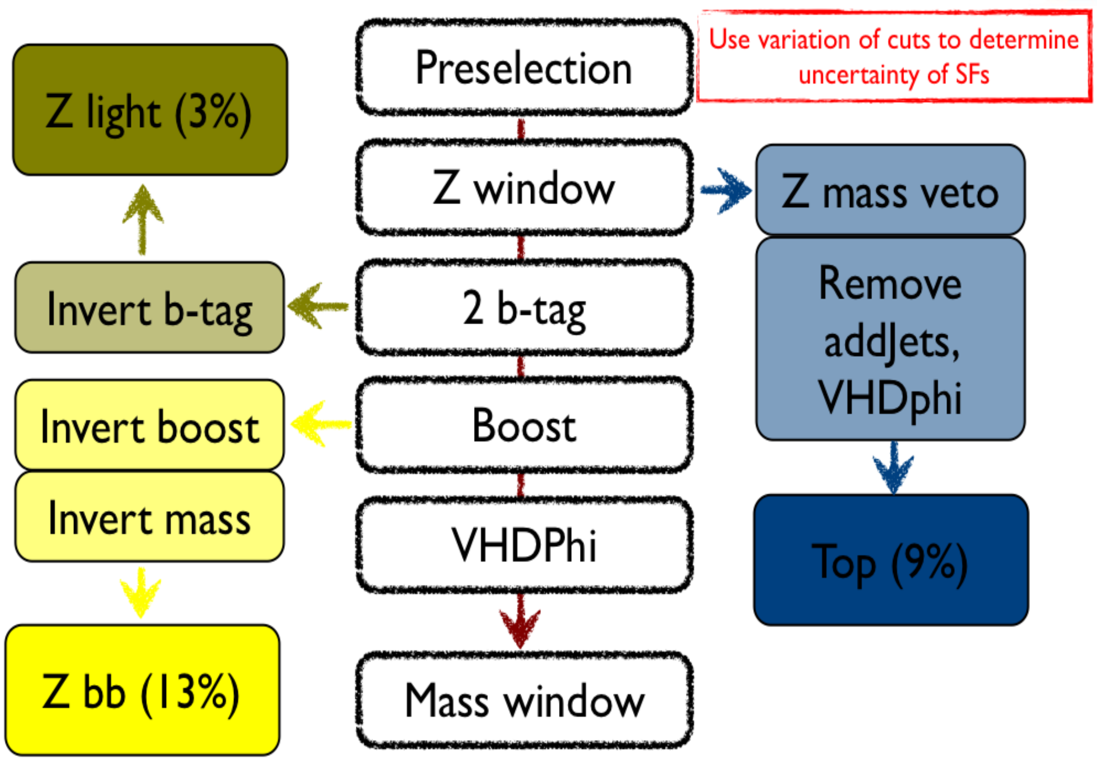
Multi BDT

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



Control regions

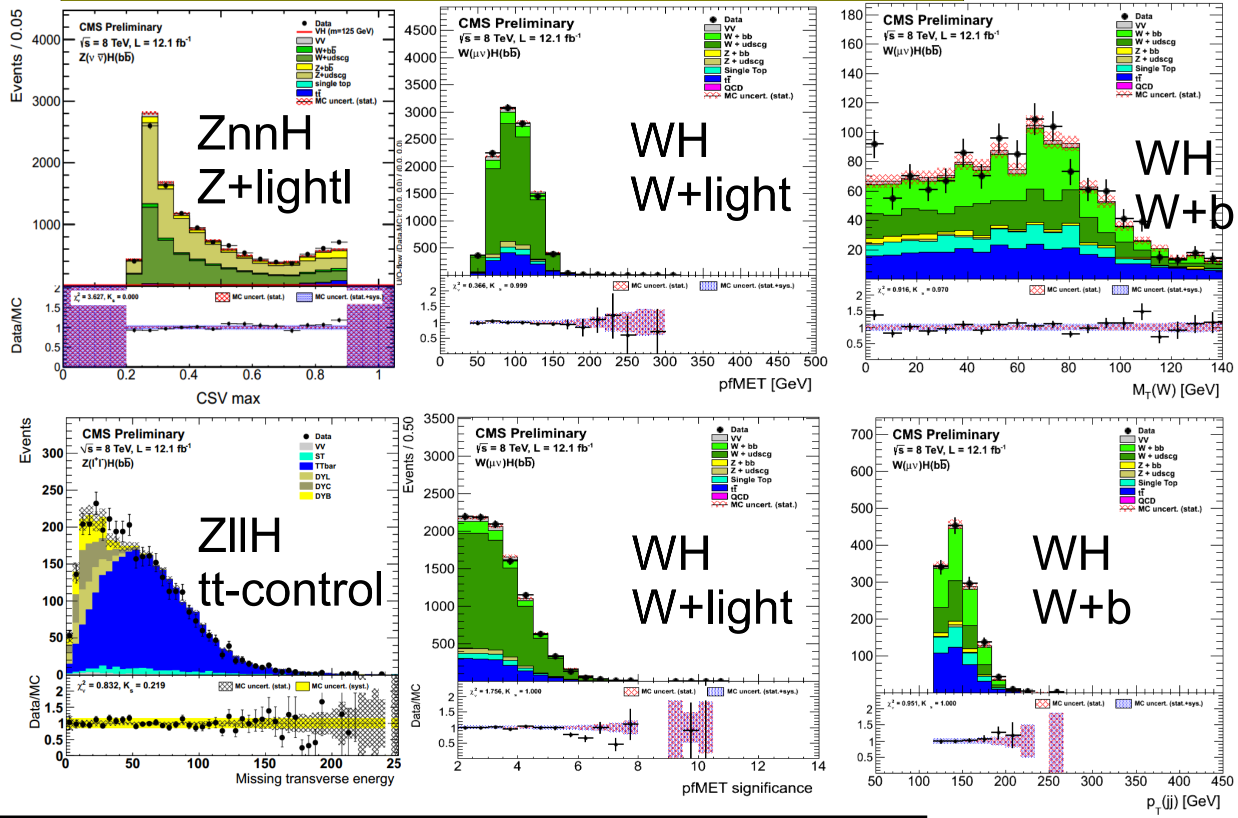
- Control regions are defined with several purpose:
  - Adjust MC prediction of main backgrounds (V+light,V+b,ttbar)
  - Verify BDT input variables distributions
  - Verify BDT input variable correlations
  - Verify BDT output distribution in signal free/depleted phase space
- Typical Control Region 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



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

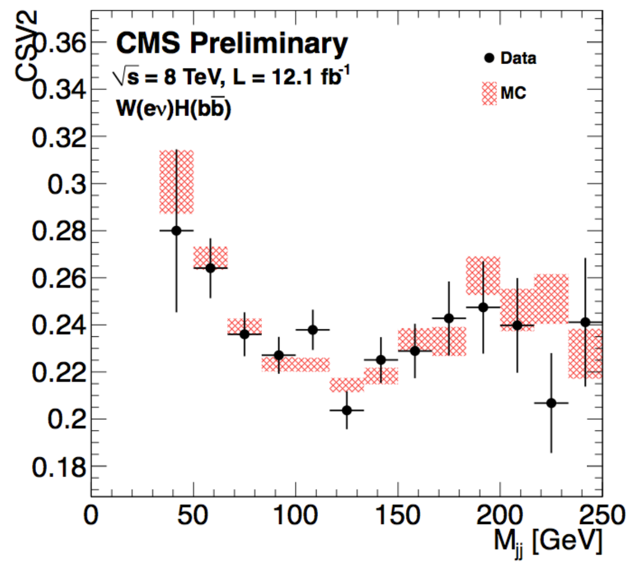


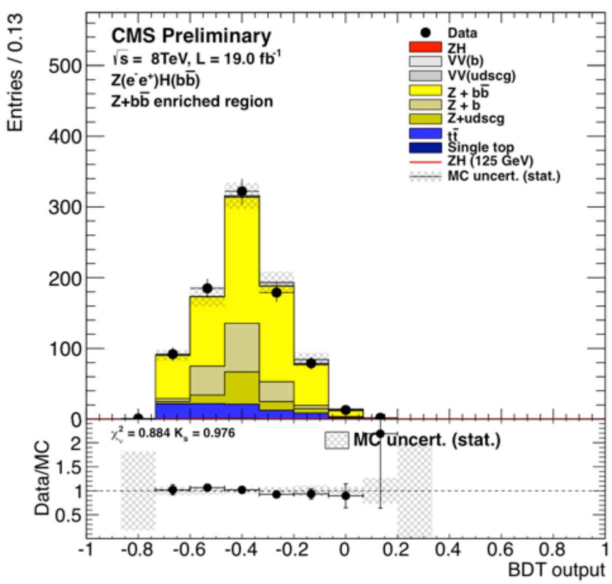
Scale Factors

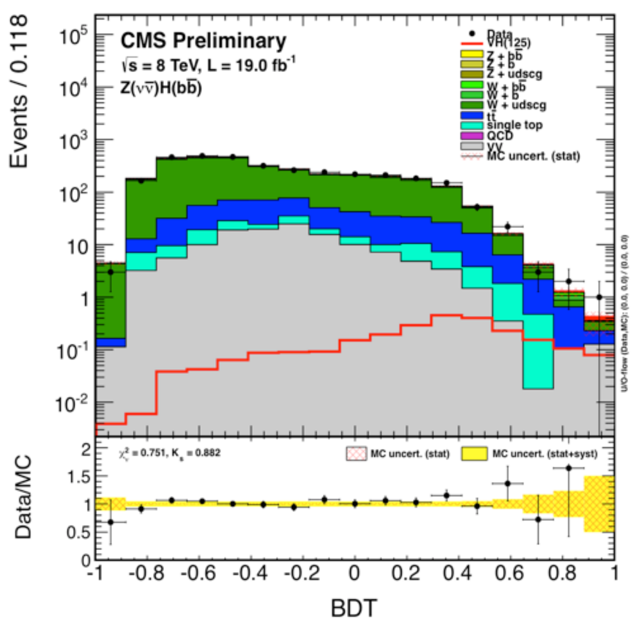
Process	$W(\ell\nu)H$	$W(\ell\nu)H$	$Z(\ell\ell)H$	$Z(\ell\ell)H$	$Z(\nu\nu)H$	$Z(\nu\nu)H$
Low p_T	7 TeV	8 TeV	7 TeV	8 TeV	7 TeV	8 TeV
W + udscg	$0.88 \pm 0.01 \pm 0.03$	$1.00 \pm 0.02 \pm 0.01$	–	–	$0.89 \pm 0.01 \pm 0.03$	$0.96 \pm 0.06 \pm 0.03$
$Wb\bar{b}$	$1.91 \pm 0.14 \pm 0.31$	$2.00 \pm 0.15 \pm 0.10$	–	–	$1.36 \pm 0.10 \pm 0.15$	$1.30 \pm 0.17 \pm 0.10$
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	–	–	$1.11 \pm 0.03 \pm 0.11$	$1.06 \pm 0.03 \pm 0.07$	$0.97 \pm 0.02 \pm 0.04$	$1.04 \pm 0.07 \pm 0.02$
$Zb\bar{b}$	–	–	$0.98 \pm 0.05 \pm 0.12$	$1.04 \pm 0.06 \pm 0.08$	$1.08 \pm 0.09 \pm 0.06$	$1.15 \pm 0.10 \pm 0.04$
$t\bar{t}$	$0.84 \pm 0.02 \pm 0.03$	$0.98 \pm 0.01 \pm 0.01$	$1.03 \pm 0.04 \pm 0.11$	$0.95 \pm 0.04 \pm 0.10$	$0.97 \pm 0.02 \pm 0.04$	$1.03 \pm 0.07 \pm 0.03$

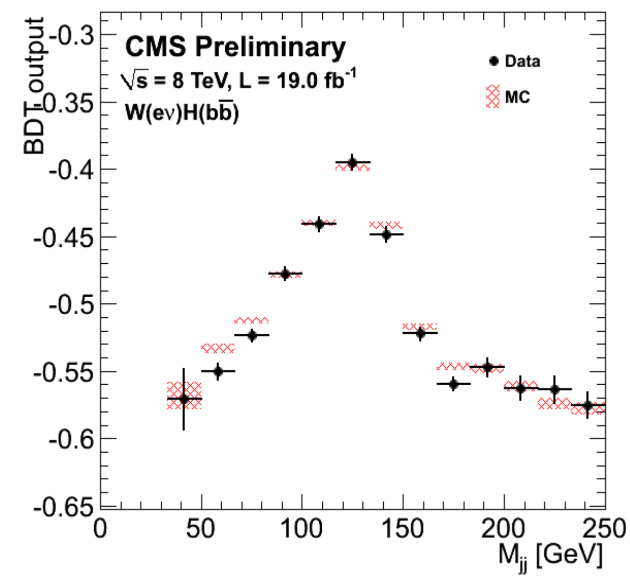
Control Regions - BDT

- 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



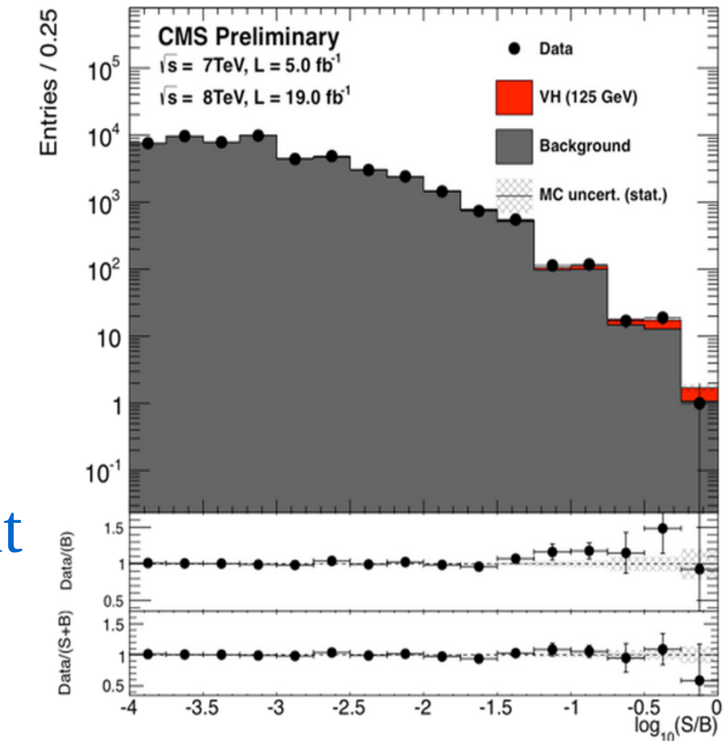


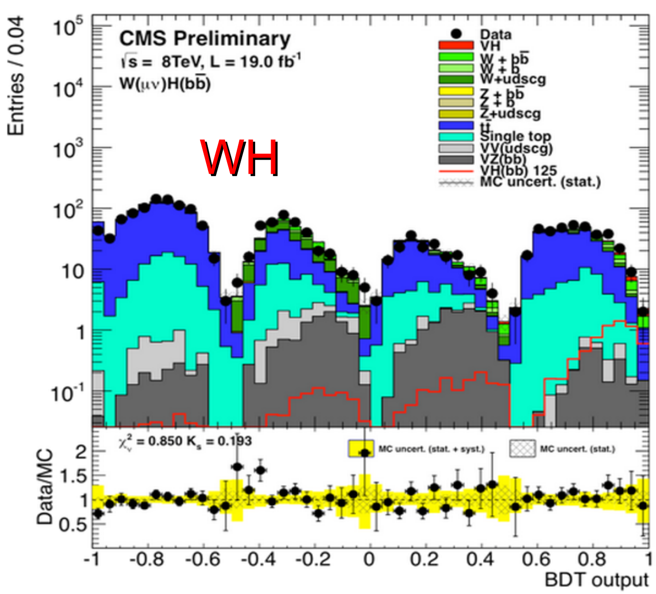


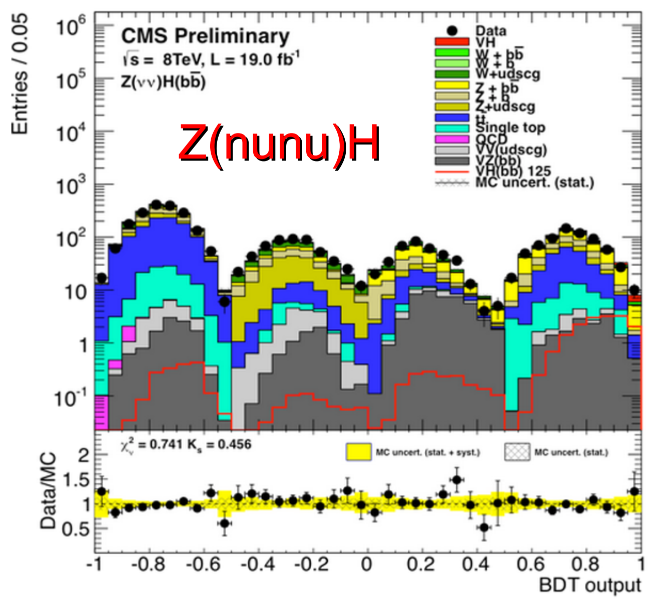


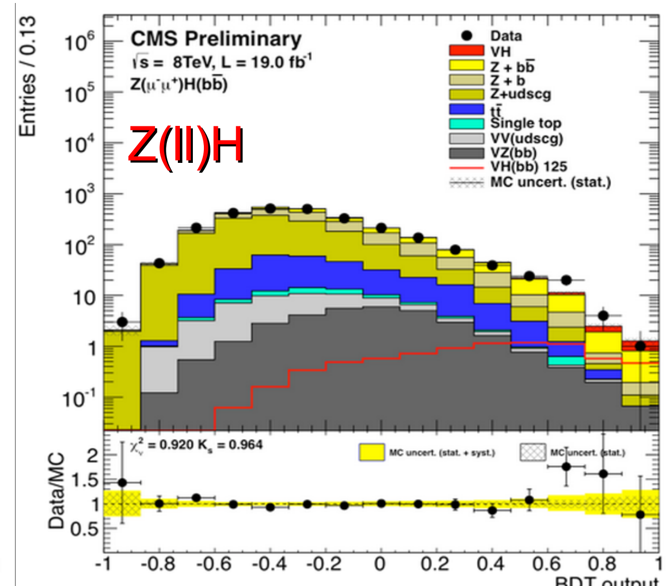
BDT output in signal region

- 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







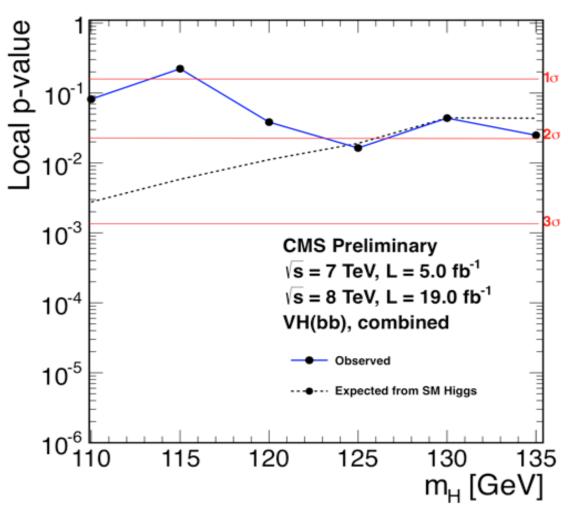


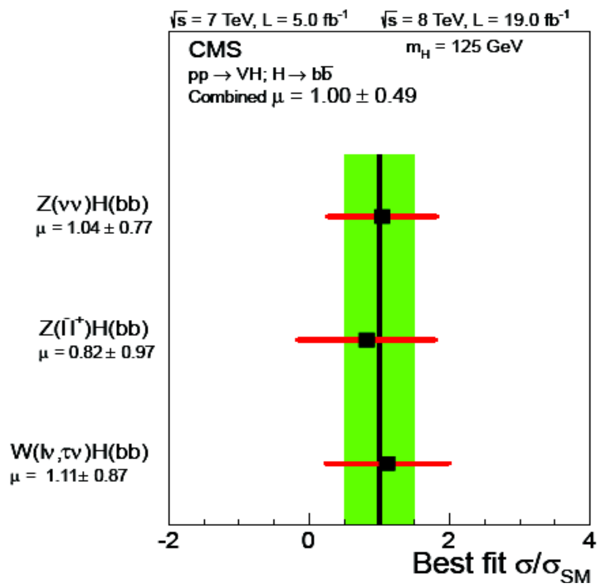
Systematic uncertainties

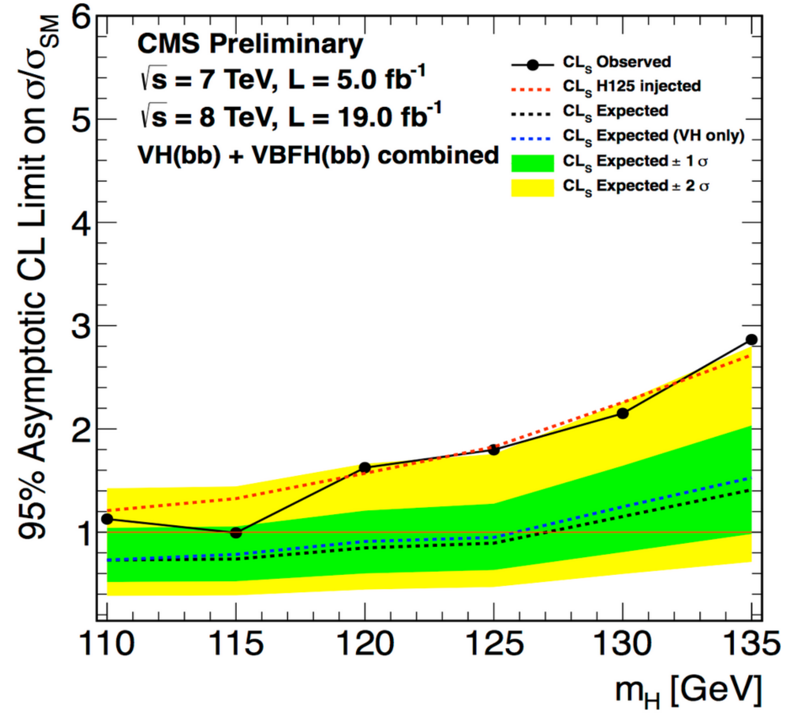
- The limit & significance are extracted with a shape analysis
- Systematic uncertainties are handled as nuisance parameters
- Where applicable a shape uncertainty is taken
    - B-tagging (doing discriminator re-shaping)
    - JEC/JER (variation within quoted uncertainties)
    - Background models (different generators)
    - Signal pt-spectrum (half size of NNLO QCD and NLO EWK corrections)
    - 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

VH MVA Results

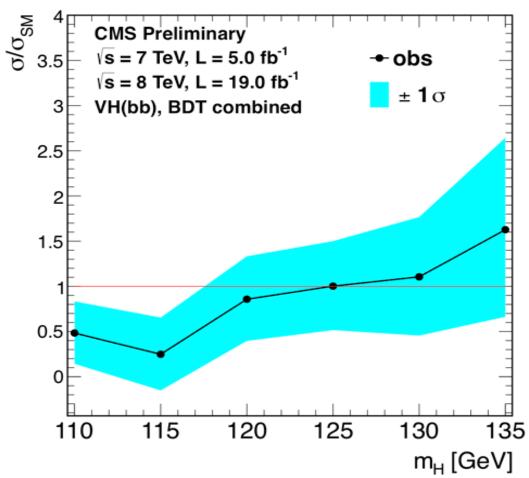
- Broad excess compatible with the 125 GeV boson
  - 2.1sigma obs, 2.1sigma exp

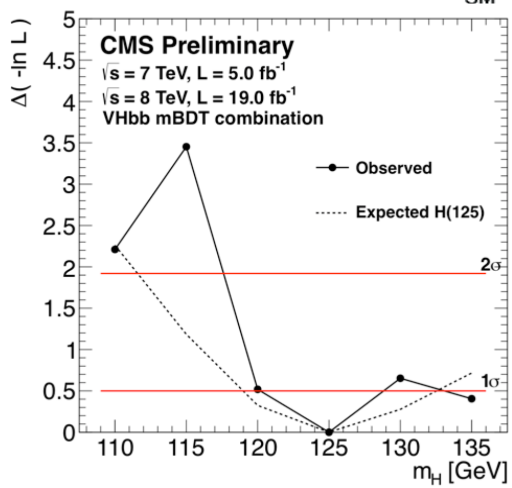






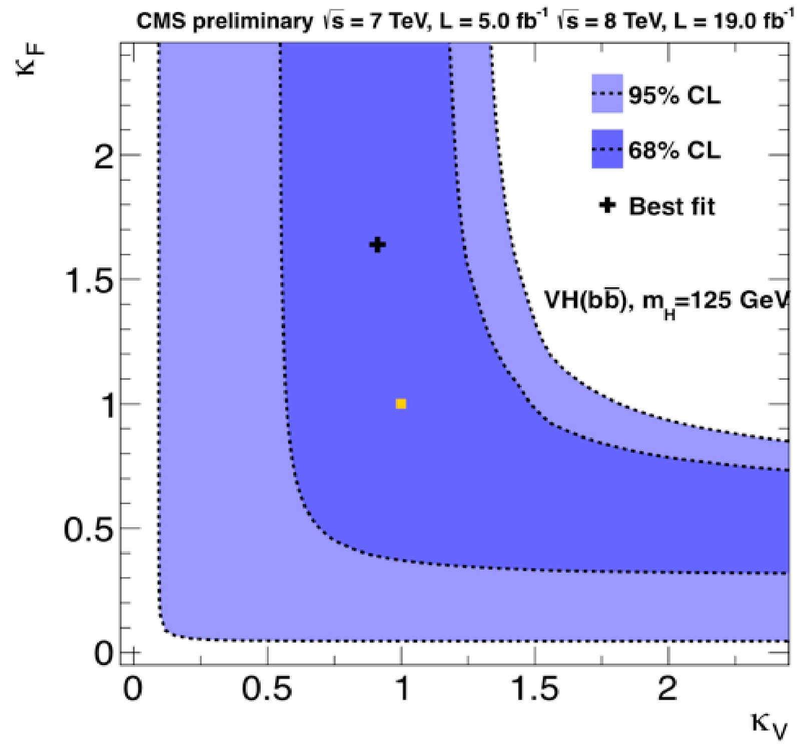
m_H [GeV]	110	115	120	125	130	135
σ/σ_{SM} (95% CL) median expected	0.73	0.79	0.91	0.95	1.25	1.53
σ/σ_{SM} (95% CL) observed	1.13	1.09	1.74	1.89	2.30	3.07

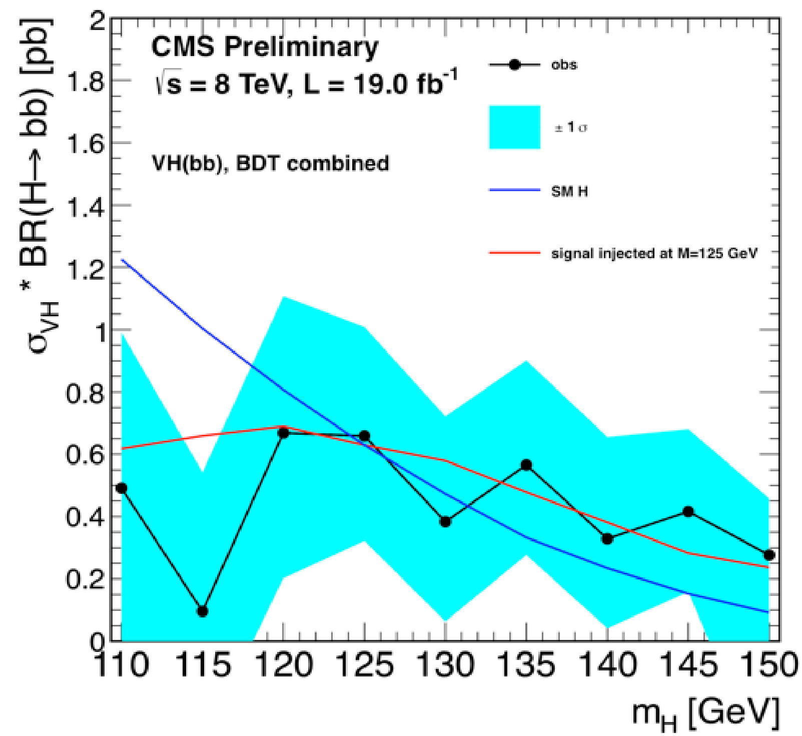




@125 GeV
sig = 2.1 std. dev.
mu = 1.0 + 0.5 – 0.5

More on VH results





- Result also interpreted in the kF – kV plane
- For 8TeV data the BDT analysis was actually extended to 150 GeV
  - The excess is broad (due to low mass resolution) but is compatible with 125GeV Higgs expectations
  - The fitted xsec decrease at higher mass

Mjj Analysis

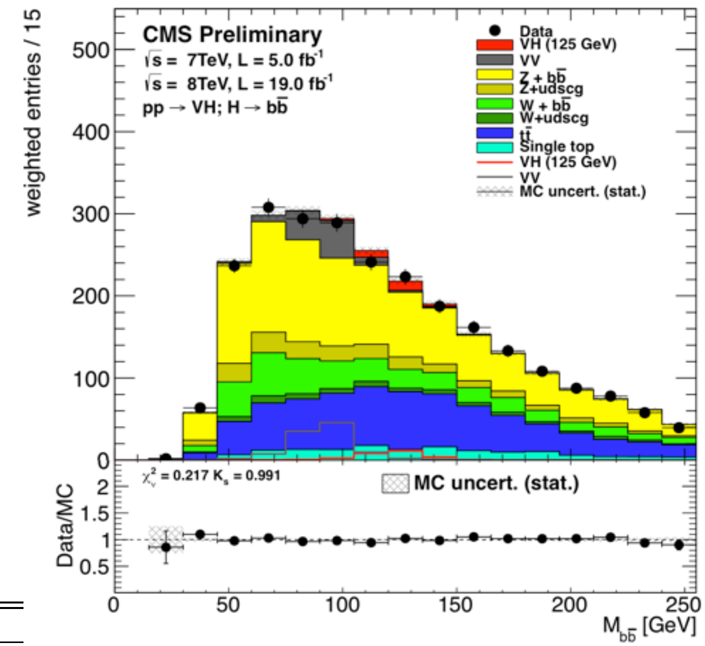
While the main analysis is based on a BDT, a cross-check analysis is implemented as *a shape analysis* on the dijet invariant mass selecting high S/B with:

→ Exploit the boost (pt binning)

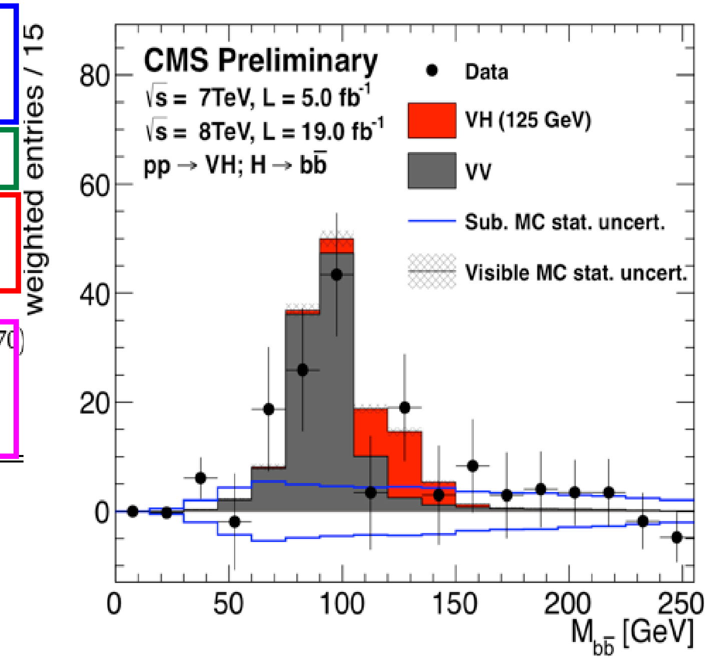
→ Double asymmetric b-tagging

→ Topology: b2b, jet veto

→ QCD rejection

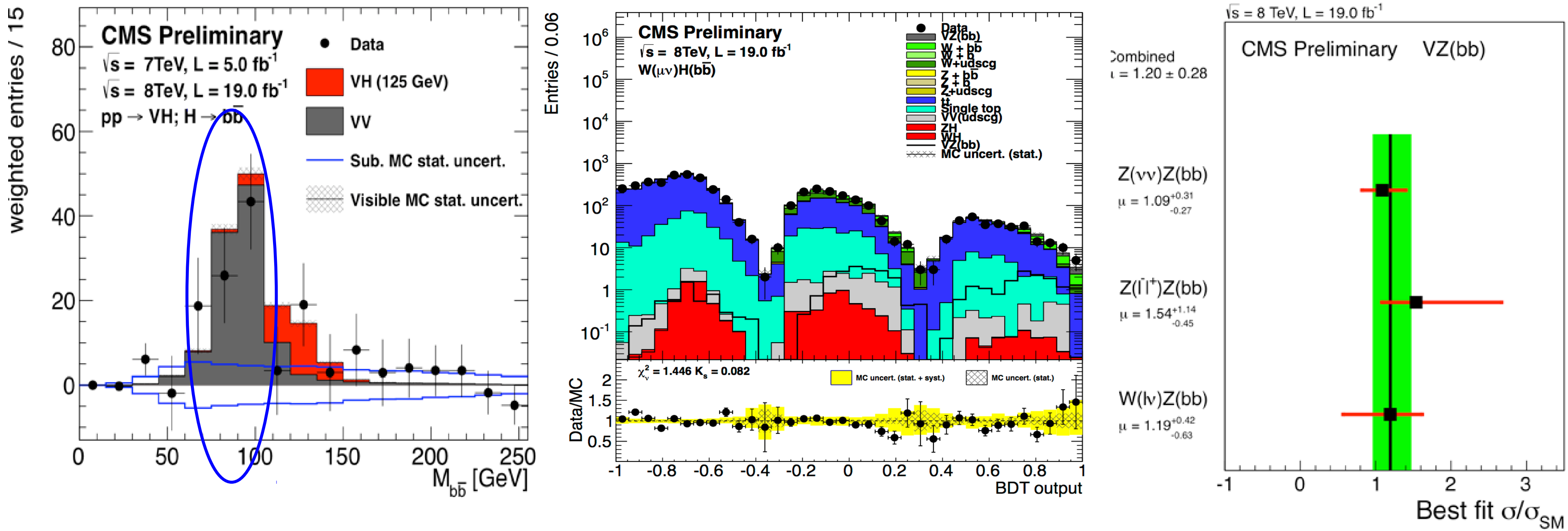


Variable	$W(\mu\nu)H$	$W(e\nu)H$	$Z(\ell\ell)H$	$Z(\nu\nu)H$
$m_{\ell\ell}$	–	–	$75 < m_{\ell\ell} < 105$	–
$p_T(j_1)$	> 30	> 30	> 20	> 60 $(> 60, > 80)$
$p_T(j_2)$	> 30	> 30	> 20	> 30
$p_T(jj)$	> 100	> 100	–	> 110 $(> 140, > 190)$
$p_T(V)$	$100-130$ $(130-180, > 180)$	$[100-150]$ (> 150)	$[50-100]$ $([100-150], > 150)$	–
CSV1	CSVT	CSVT	CSVM	CSVT
CSV2	> 0.5	> 0.5	> 0.5	> 0.5
$\Delta\phi(V,H)$	> 2.95	> 2.95	–	> 2.95
$\Delta R(jj)$	–	–	$-(-, < 1.6)$	–
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)	$< \pi/2$	$< \pi/2$	–	–



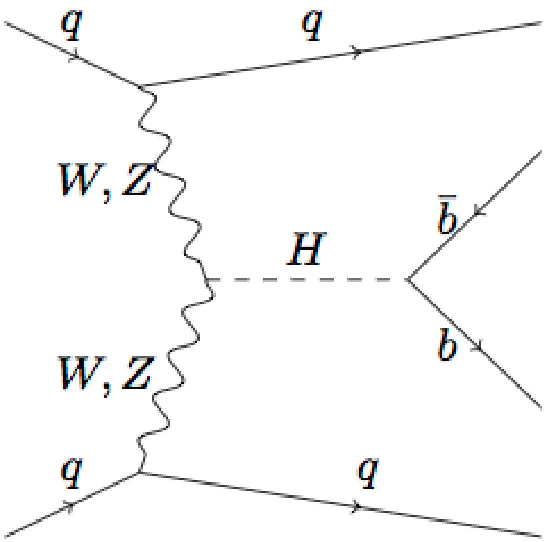
@125 GeV
sig = 1.1 std. dev.
Mu = 0.8 + 0.7 – 0.7

Di-boson cross section



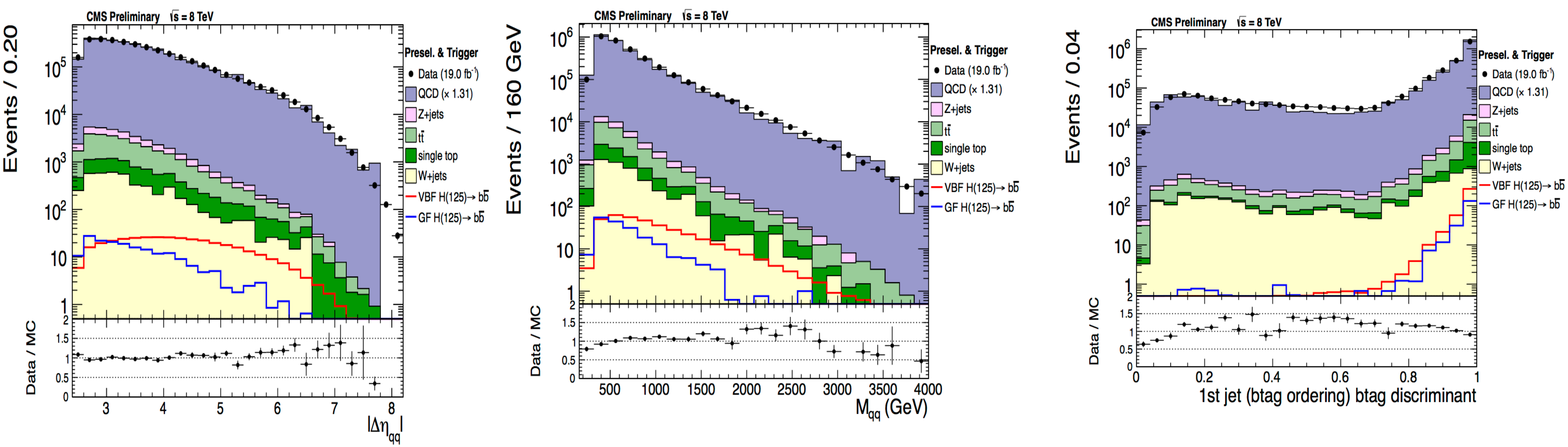
- We can validate the whole analysis chain targeting VZ(bb) instead of VH(bb)
- Testing both the (multi)BDT technique and the simple Mjj
- Results compatible with SM expectation
  - >6 sigma for BDT, ~4 sigma for Mjj

VBF, H->bb



VBF Hbb

- The VBF signature is the usual 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

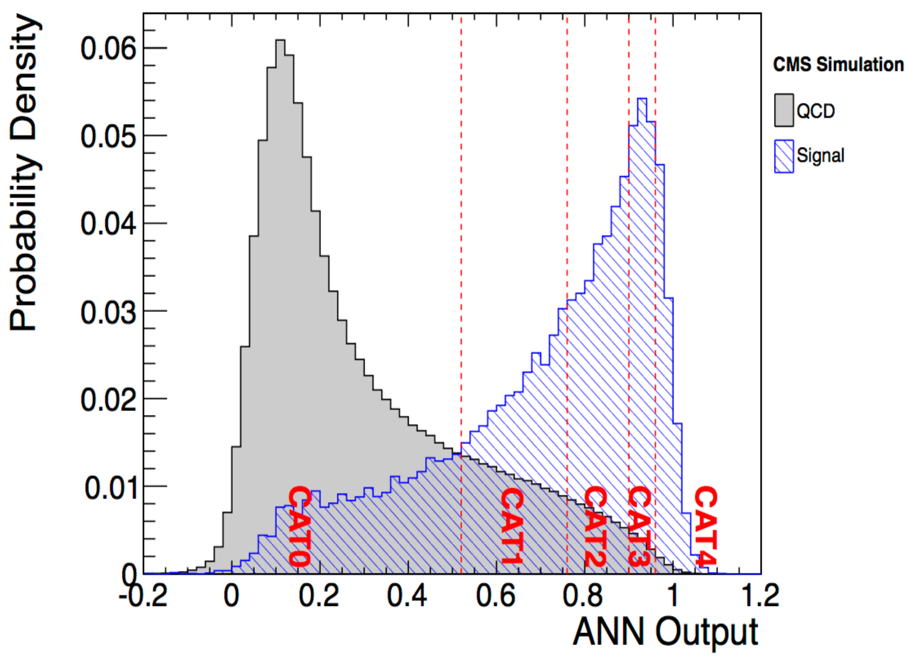


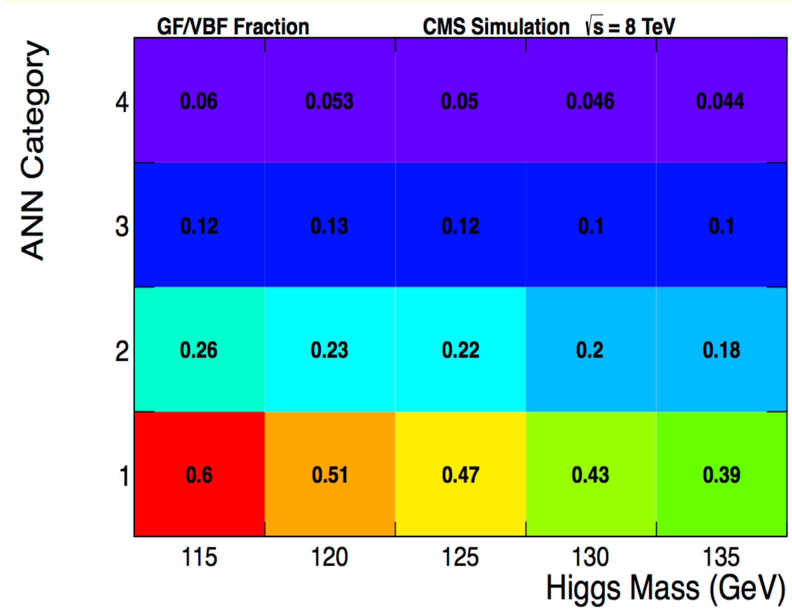
Analysis strategy

- Combine all discriminating variables into an MVA output
  - Do not use variables highly correlated with b-bbar invariant mass
- 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

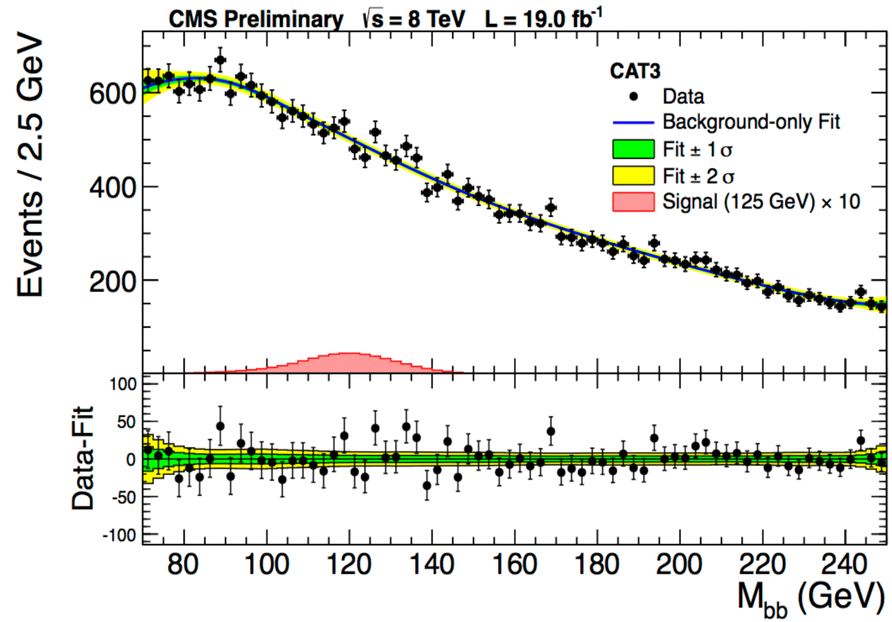


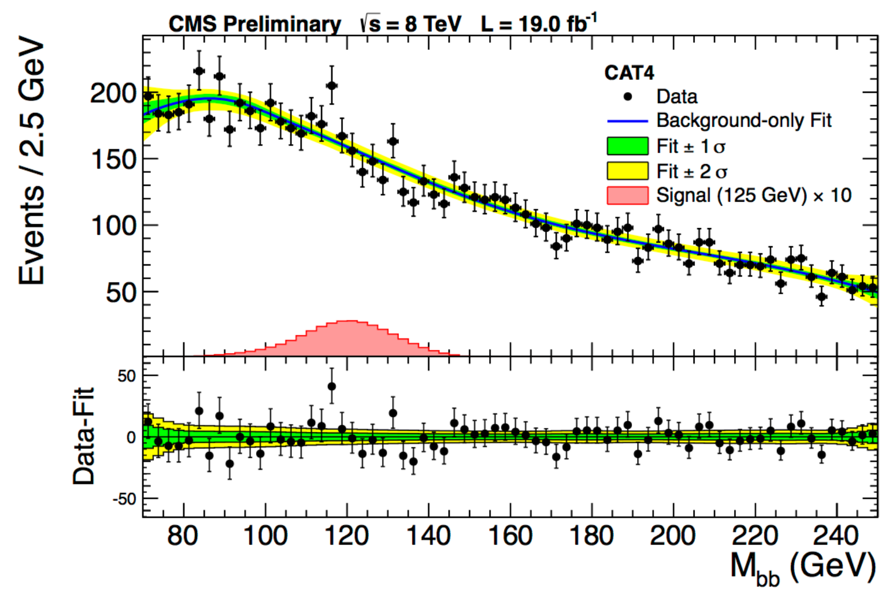


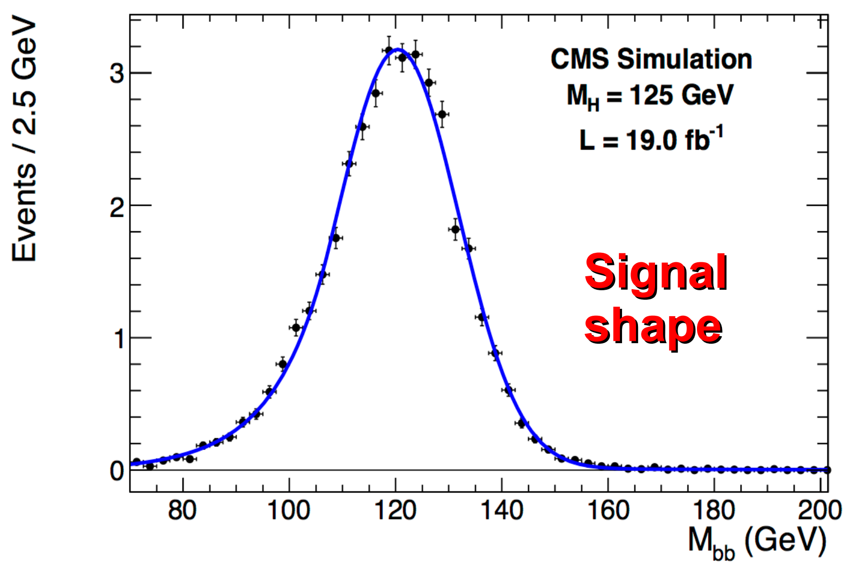
Cat. 0	Cat. 1	Cat. 2	Cat. 3	Cat. 4
$ANN < 0.52$	$0.52 \leq ANN < 0.76$	$0.76 \leq ANN < 0.90$	$0.90 \leq ANN < 0.96$	$ANN \geq 0.96$

Fit in the bb invariant mass

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







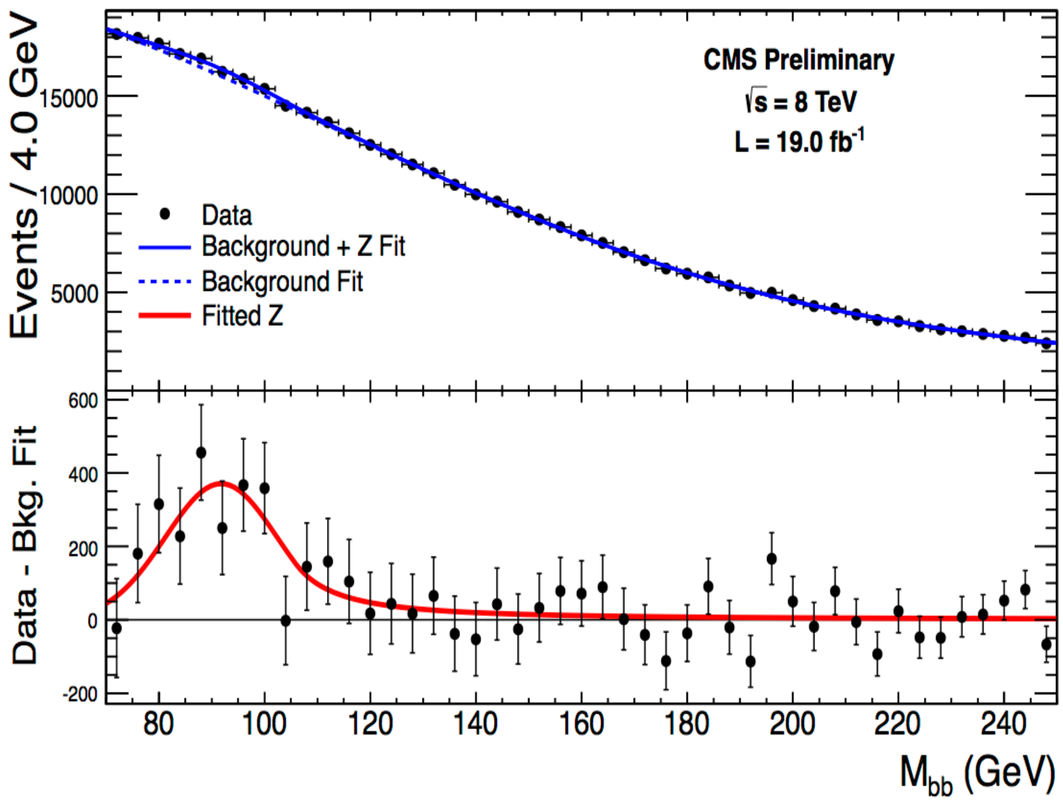
Z+jets cross check

- A cross check of the fitting machinery has been done to see the Z+jet candle
- Two versions tested:
  - Looser preselection cuts
  - Higgs like selection (no MVA, but b-tag cuts)
- Excess due to Z correctly fitted on top of the very large background

Results:

- 8 sigma excess for loose sel
- 2.5 sigma excess for higgs like

Both in agreement with SM expectations



Systematic uncertainties

- Dominant background (QCD) is completely data driven
- MC uncertainties mostly for signal acceptance
- Total effect of systematics on the final result is about 15%

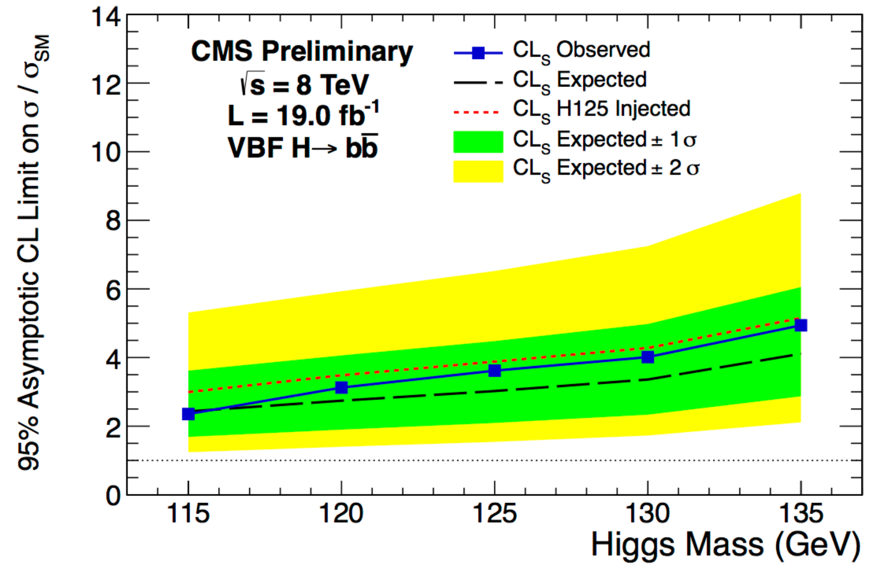
Source	Uncertainty
Background fit	depending on the statistics of each category
Z+jets cross section	$\pm 20\%$
top cross section	$\pm 20\%$
Signal and Z peak position (JES)	$\pm 1.5\%$
Signal and Z resolution	$\pm 10\%$
Luminosity	$\pm 4.4\%$
Trigger efficiency	$\pm 5 - 8\%$
Signal acceptance due to JES	$\pm 10\%$
Signal acceptance due to JER	$\pm 2\%$
VBF cross section	$\pm 3\%$
VBF Monte Carlo acceptance	$\pm 10\%$
PDF	$\pm 5\%$
VBF ANN shape due to b-tag	$\pm 2\%$
VBF ANN shape due to quark-gluon discriminator	$\pm 2\%$
VBF ANN shape due to UE modeling	$-8 - +2\%$
GF cross section	$\pm 15\%$
GF Monte Carlo acceptance	$\pm 50\%$
GF ANN shape	$\pm 50\%$

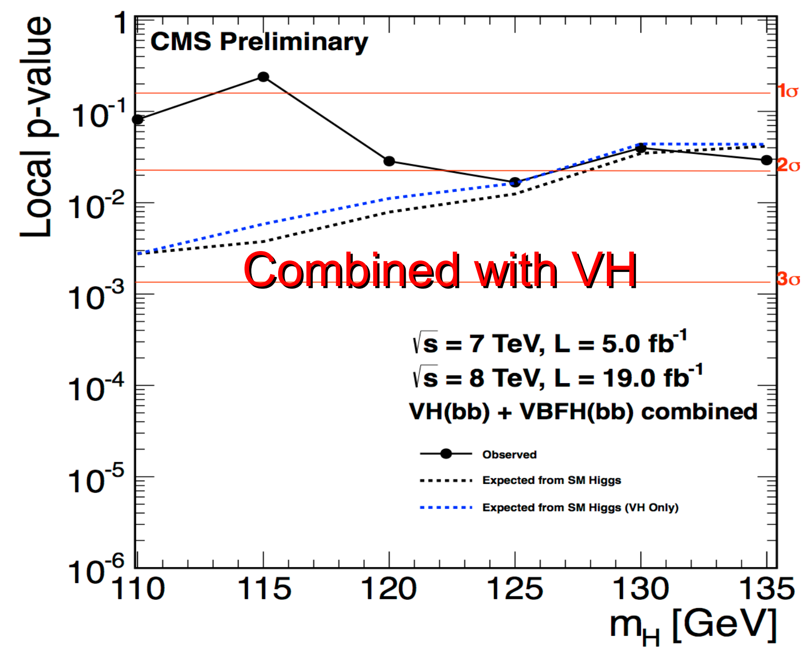
VBF results

- 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

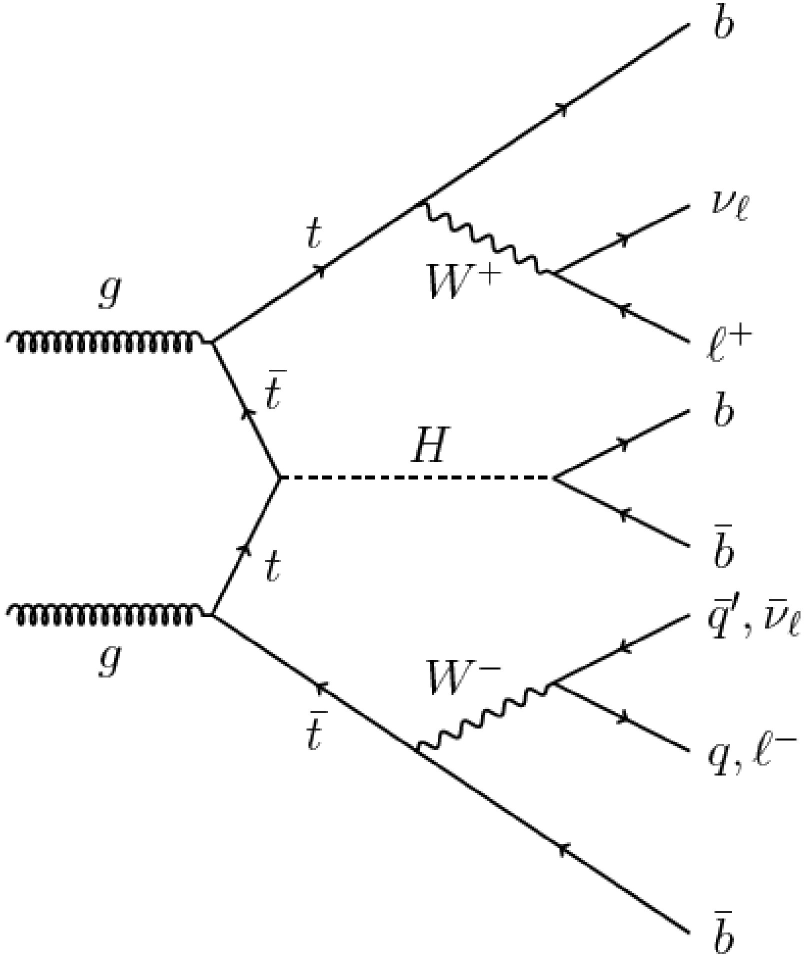
@125 GeV
Sig = 0.5 std. dev. (0.7 exp)
Mu = 0.7 + 1.4 – 1.4

- A combination with VH result is also performed
  - Relative weight of the VBF is about 10%



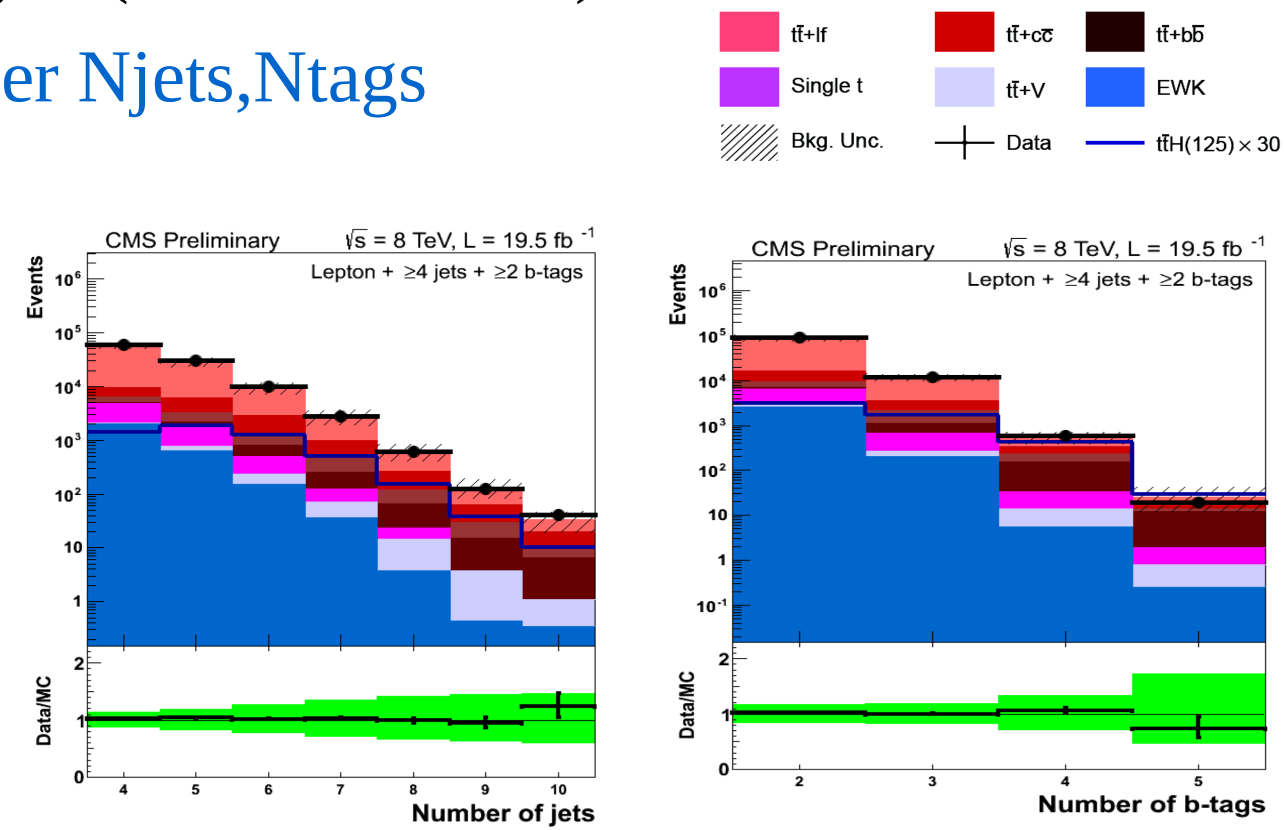


ttH (H->bb)



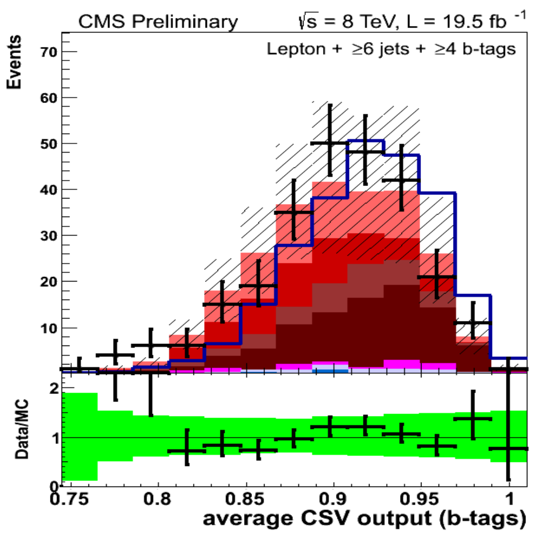
ttH, H to bb

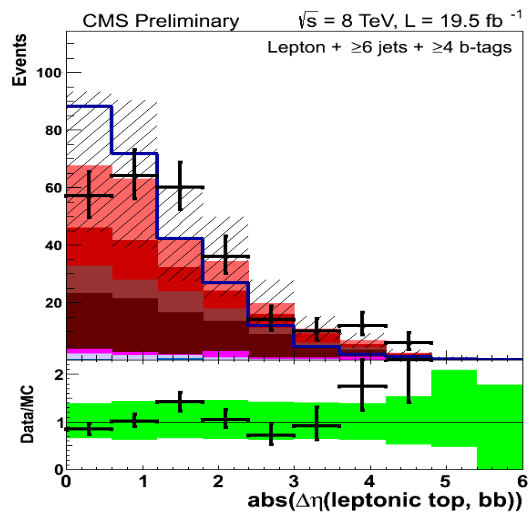
- Two modes studied (for bb): semi-leptonic and dileptonic
- Signal to background ratio rapidly increasing with
  - Total number of jets (expect 6 or 4 jets in final state)
  - Number of b-tagged jets (4 b in final state)
- Analysis categorized per Njets,Ntags
  - Low Njets,Ntags useful for backgrounds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

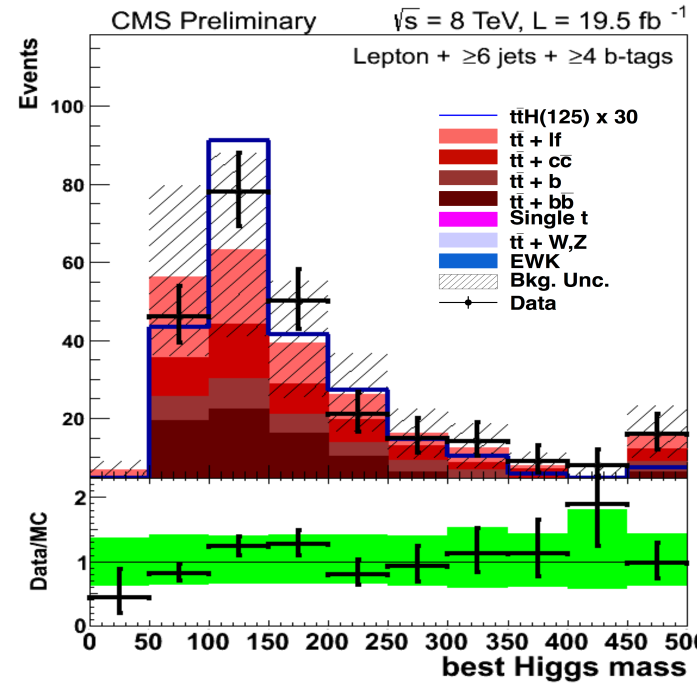


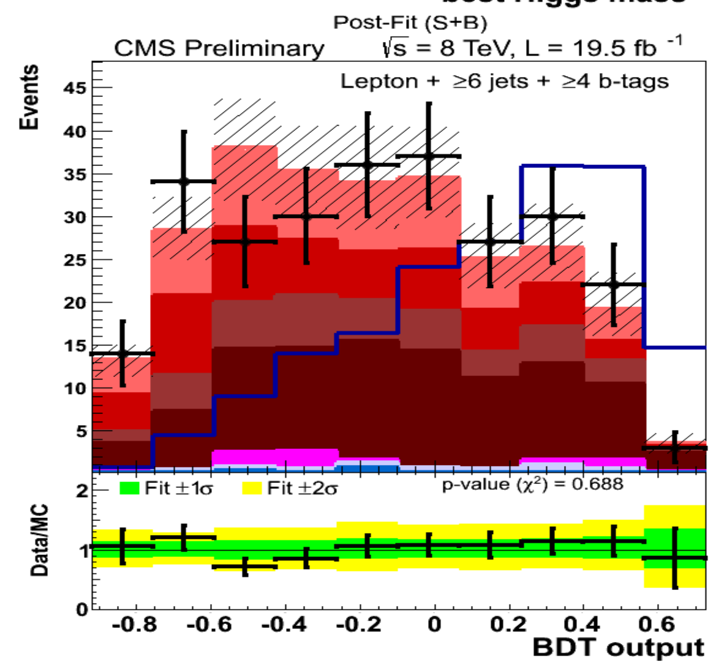
ttH, H to bb

- 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

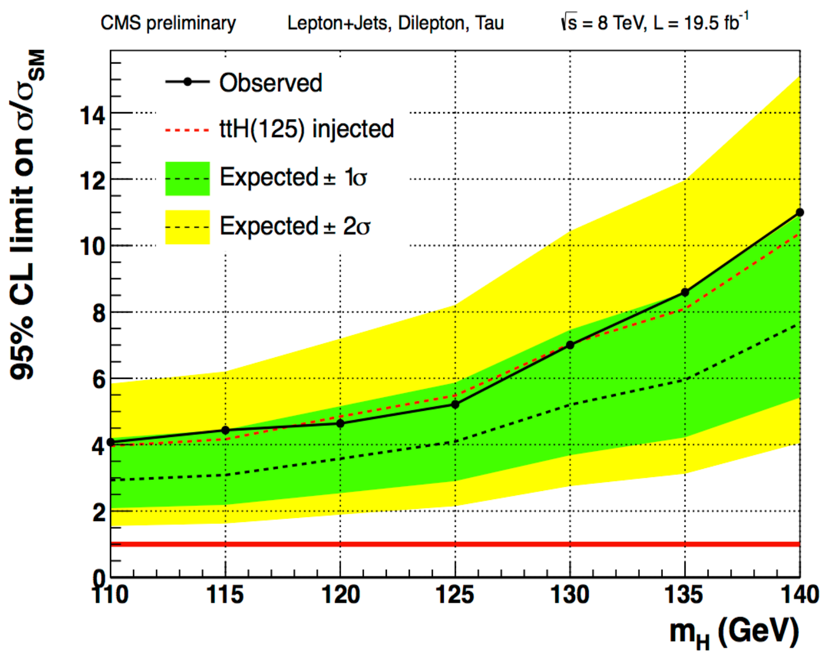
Systematics

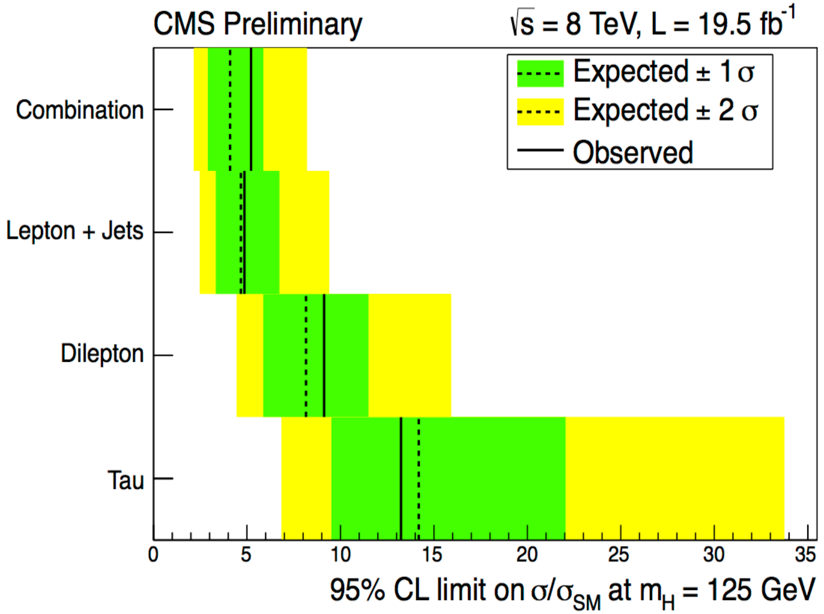
- Dominant systematics:
  - tt+bb normalization
  - B-tag shape uncertainties
  - Jet Energy Scale

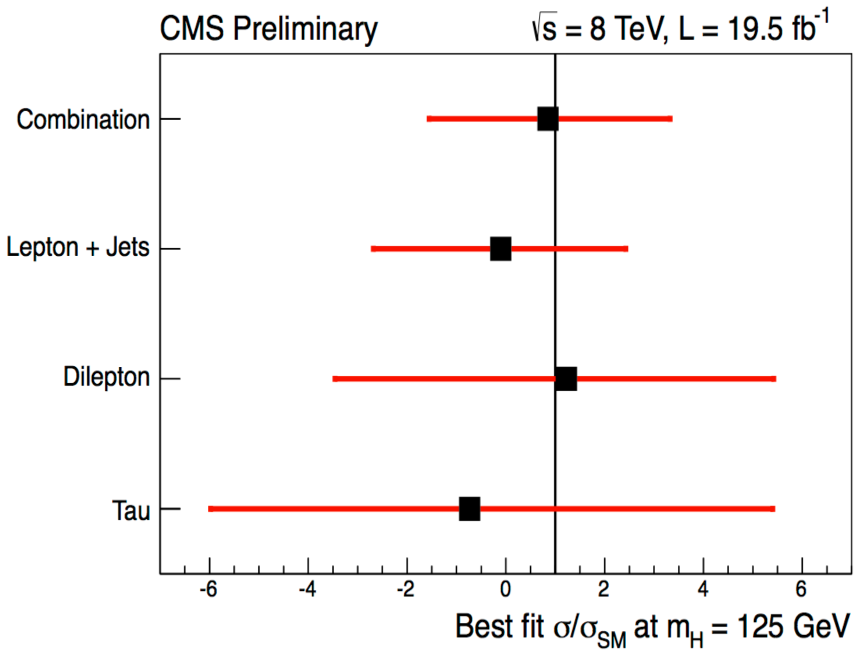
Uncertainties of the sum of $t\bar{t}$+lf, $t\bar{t}$+b, $t\bar{t}+b\bar{b}$, and $t\bar{t}+c\bar{c}$ events with $\geq$ 6 jets and $\geq$ 4 b-tags		
Source	Rate	Shape?
QCD Scale (all $t\bar{t}$+hf)	35%	No
QCD Scale ($t\bar{t}+b\bar{b}$)	17%	No
b-Tag bottom-flavor contamination	17%	Yes
QCD Scale ($t\bar{t}+c\bar{c}$)	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

ttH (Hbb and Htautau)

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



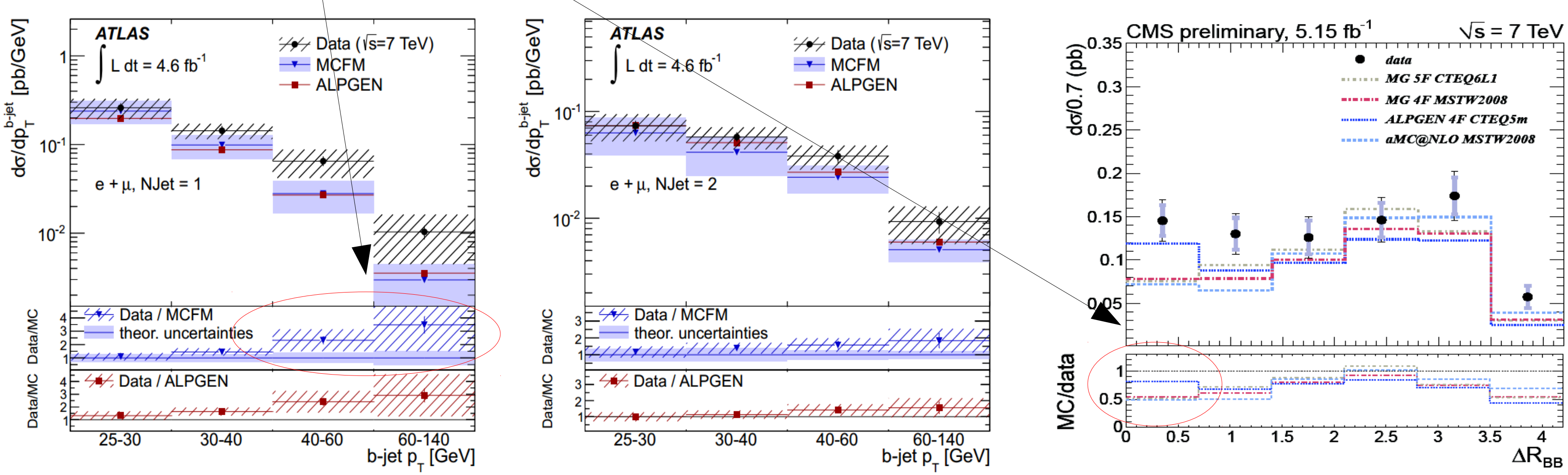




What about the future?
13 TeV, hundreds of 1/fb, high pileup....

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?

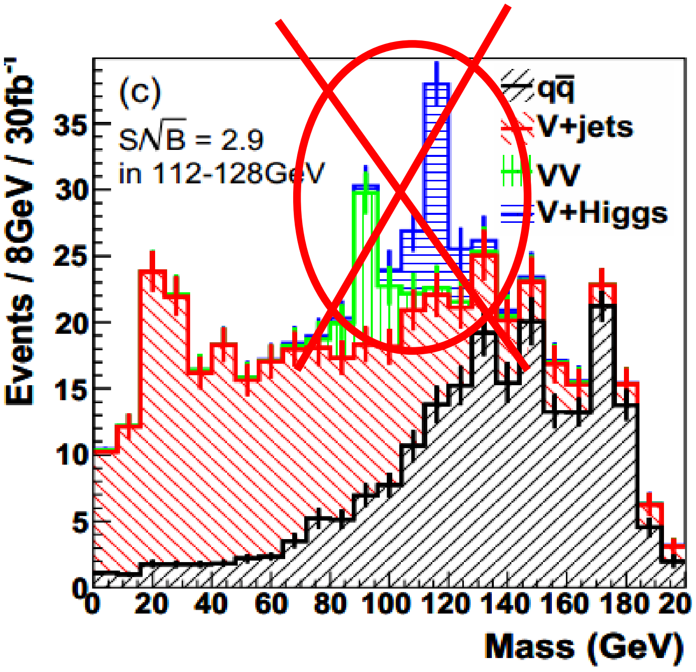
- Luminosity scaling
  - In VH, S/B is at most ~ 1/6
  - MC predictions becoming systematically limited?
    - More stat in the sidebands
    - Less extrapolations
    - Use generic templats (smooth shapes) instead of MC shapes
  - 450M MC events used for 20/fb, we cannot produce a factor of 10 more....

Scaling with sqrt(s) and PU

- ttbar cross section grows faster than VH one!
  - Already seen in 7->8 TeV transition
  - 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 cross section grows faster than ttbar one
  - 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



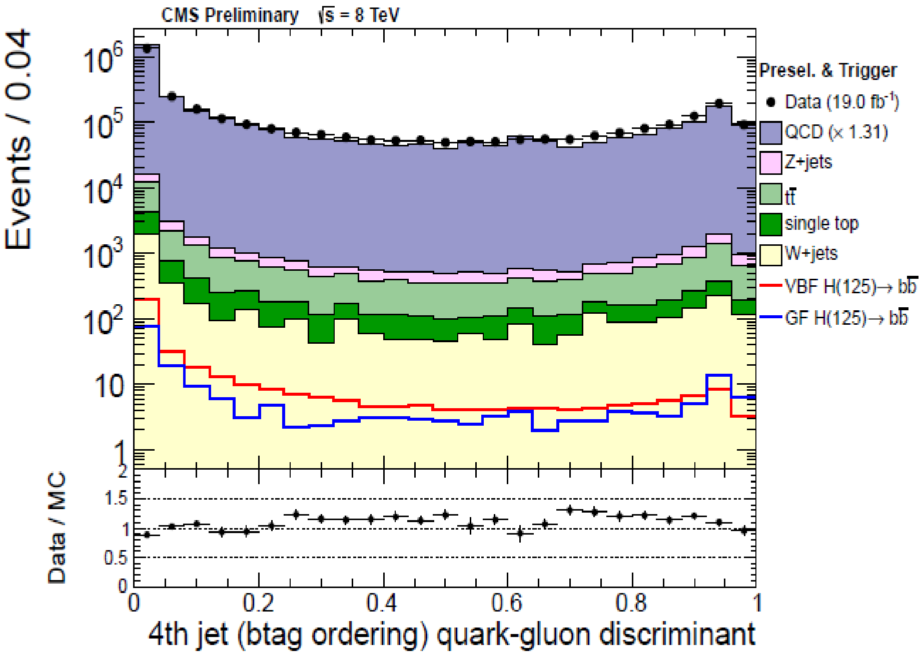
Conclusions

- The Higgs to b-quarks problem can be studied in at least three different channels
- CMS recently added VBF to the family of Hbb studies
- The VH mode is on track to have a 3 sigma evidence with first data at 13 TeV
- ttH and VBF not yet reaching SM sensitivity but can likely get there with ~100/fb (and some work to control the systematics)

Back up

Quark-gluon discriminator

- Uses jet properties to distinguish quark jets from gluons jets
  - RMS of the constituents in eta-phi plane
  - Asymmetry of the constituents wrt the center of the jet
  - Number of constituents
  - Max energy fraction carried by a single const.
- Validated on dijet production
- Used also in measurement of EWK produced Z+jj



Gluon splitting

- Studied already in 2010 on pure QCD!
- Ratio between "back to back" b-bbar production and small angle region
- Very different predictions from different generators, none of them really doing a good job