Highlights from High pT at LHC

Andrea Rizzi University and INFN Pisa

Trento, 12th Sept 2017

High pt?

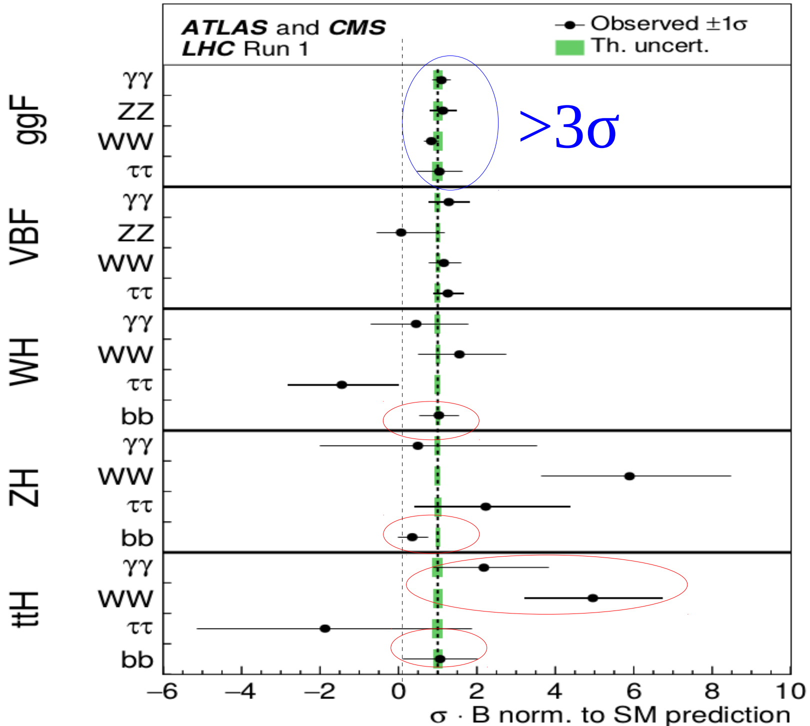
- Which topic to cover with this "high pt" talk?
  - ~~V+jets~~ → yesterday
  - ~~Top~~ → tomorrow
  - ~~BSM~~ → friday
  - So I'll cover Higgs and some di-boson
- Are Higgs and di-boson not "strong enough" for this workshop?
  - Higgs to bb and ttH
  - Boosted topologies of boson hadronic decays
  - VBF signatures (ok they are ewk "high pt processes" but with important "soft QCD" effects)
  - HH final states
  - In general backgrounds matters, and they are QCD (V+jets, top)

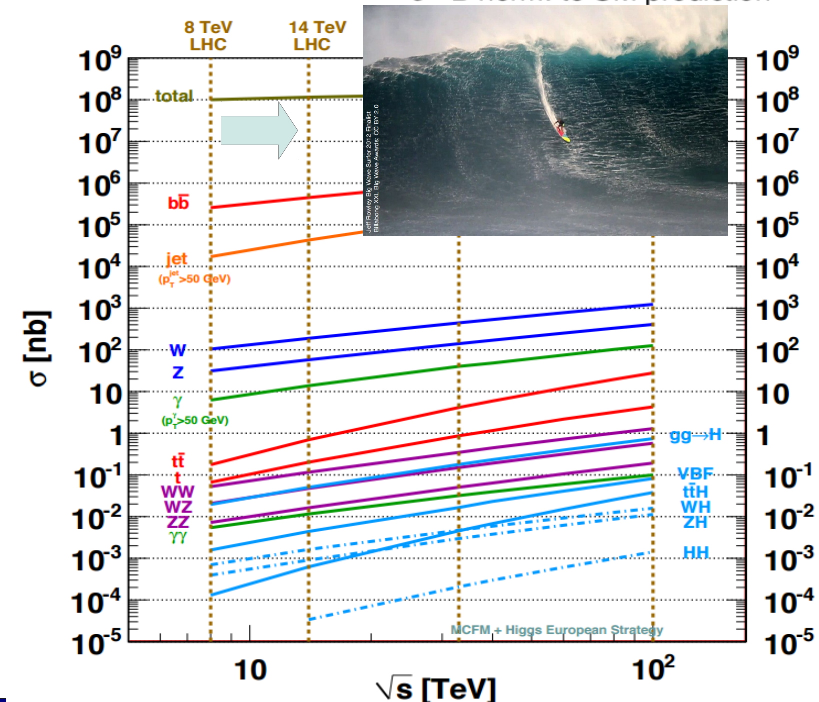
Outline

- Standard Model VH(bb) recent results
  - Details of the analysis
- Boosted Higgs
- VBF signatures
- di-boson
- HH (resonant and non resonant)

Higgs after Run 1

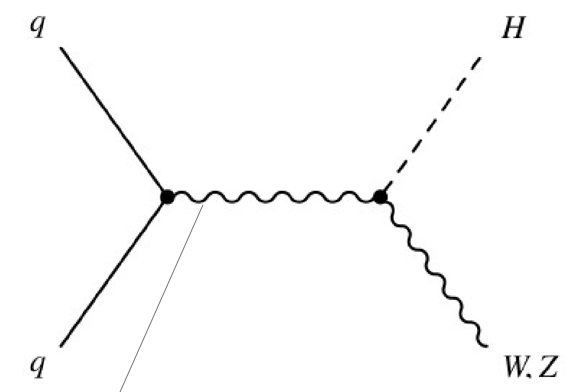
- The Higgs boson discovery was one of the major results of LHC Run1
- ATLAS and CMS measured different decay and production modes with results >3σ in most channels
- What was left after Run 1?
  - bb decay mode
  - ttH and VH associated productions
- So from Run 2 we would expect :
  - VH, with H→bb
  - ttH in the various decay modes
  - Higgs to muons?
- We surf the wave... but top does too!
  - And it is heavier!

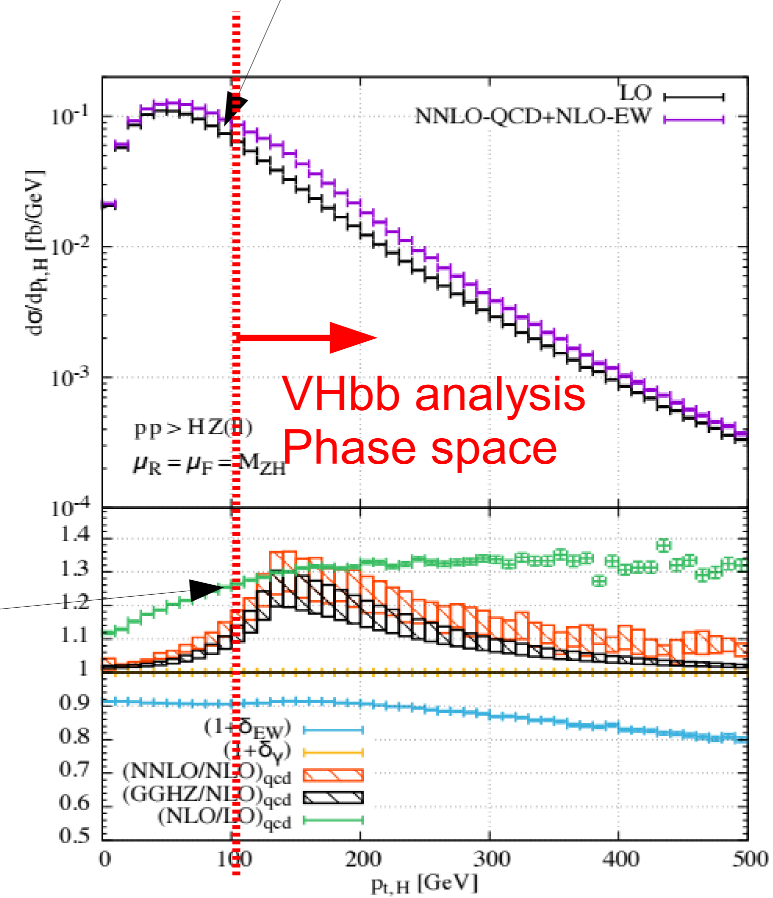


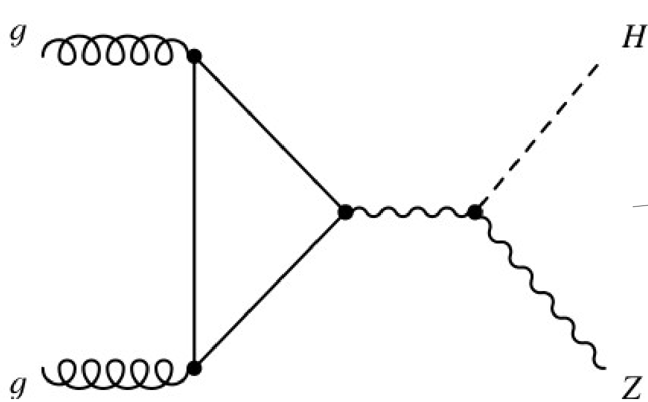


VH

- Dominant VH production at LHC via
  - qq → V* → VH
- Production via ggZH only few %
  - In the inclusive phase space
  - The picture changes in the typical VH(bb) analysis phase space
    - pT V > ~100-150 GeV
- Needed simulation of ggZH

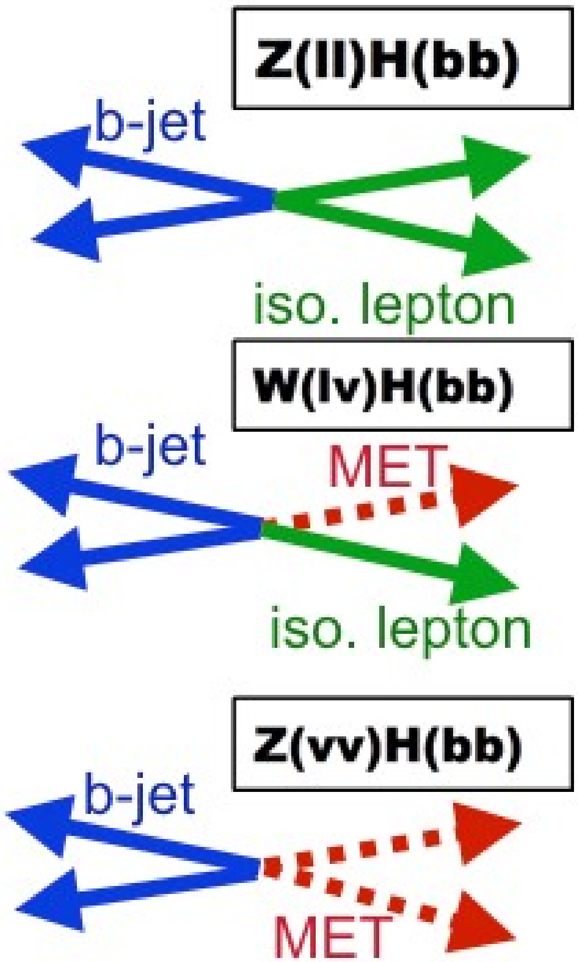






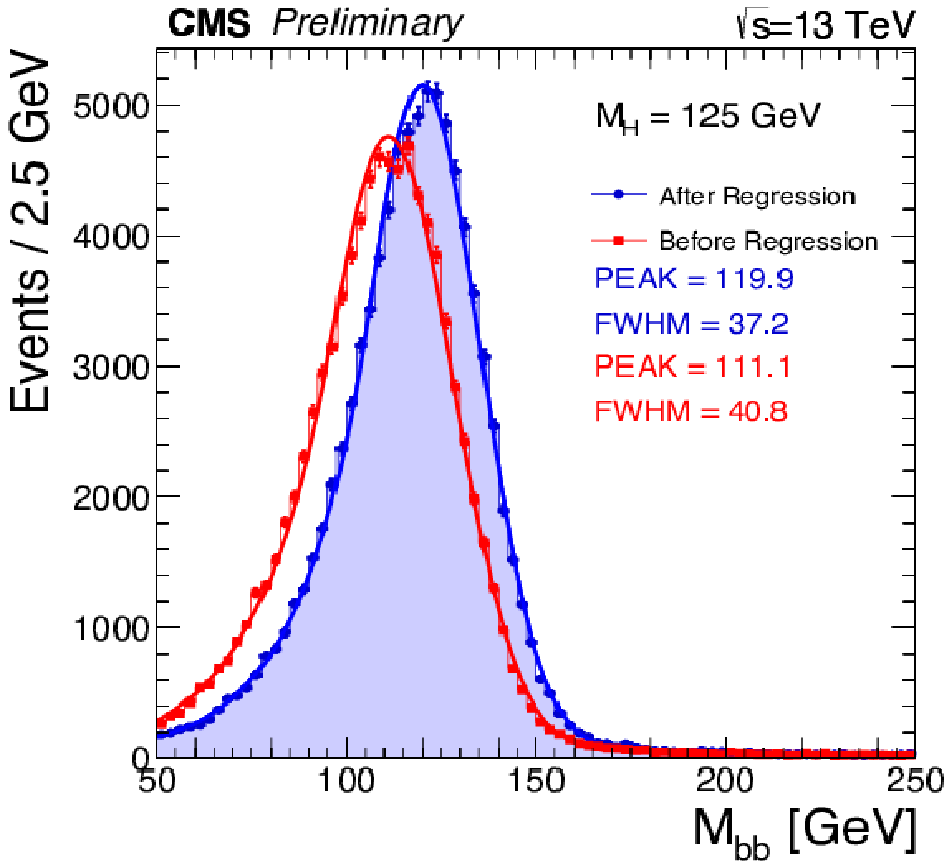
LHC VH(bb) search

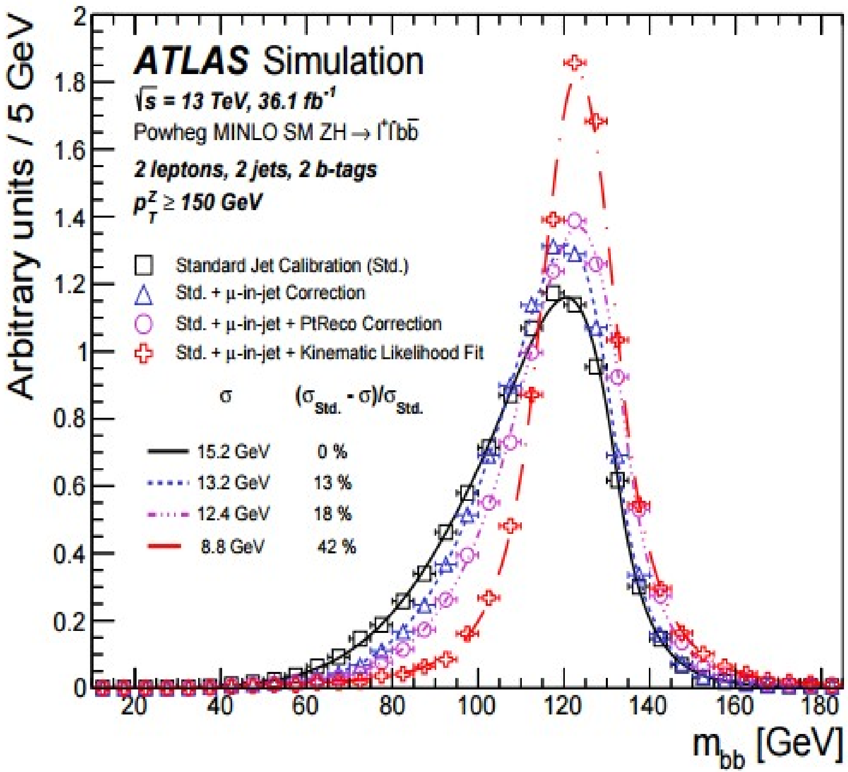
- Signal topology overview
  - Two b-tagged jets
  - 0,1,2 isolated leptons
  - or large missing energy
- Event features for S vs B separation:
  - dijet invariant mass
  - dijet system transverse momentum
  - Additional hadronic activity
- Main backgrounds:
  - V+jet
  - Ttbar & single top
  - diboson
  - QCD multi jet production



Di-jet mass

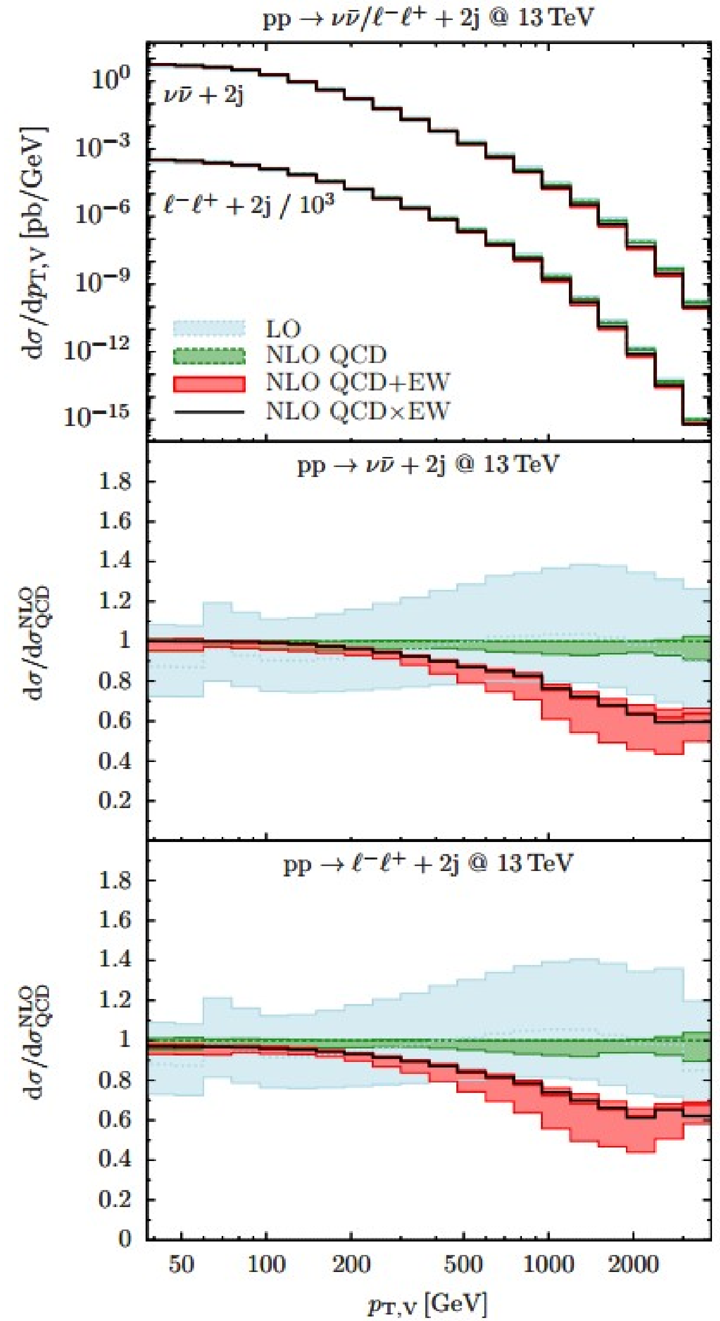
- B-jet energy is typically badly measured due to leptons+neutrinos in the decay
- Corrections can improve Mbb resolution
  - Additional constraint in $Z \rightarrow ll$ mode as the event is fully reconstructed





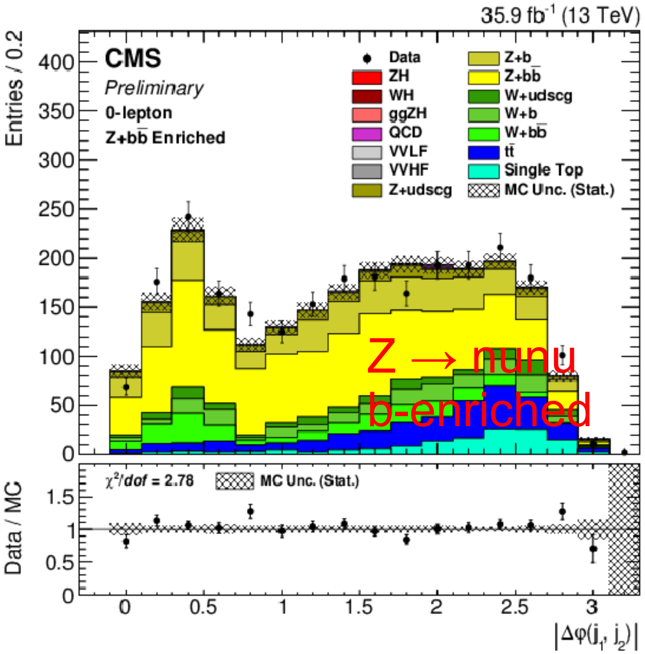
V+jet background

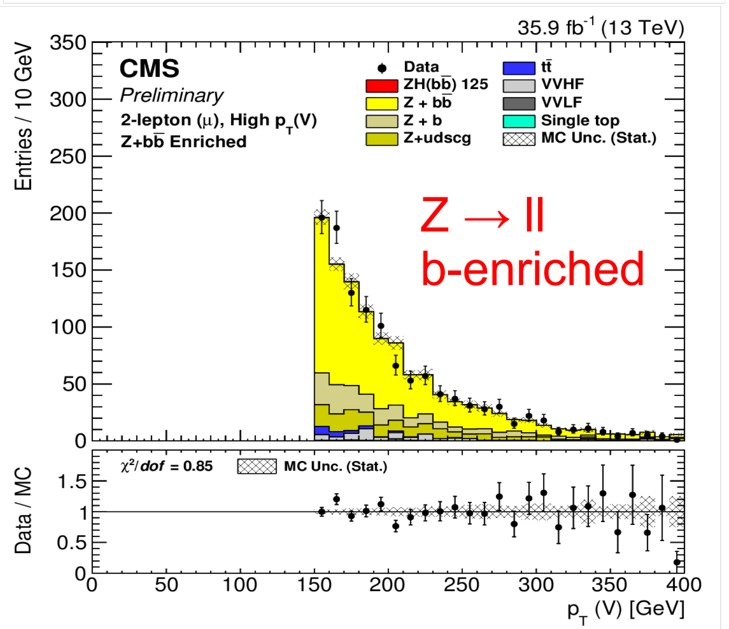
- Vector+jet is one of the dominant backgrounds
- Several difficulties in modelling this background
  - Important EWK and QCD corrections to the V-pt spectrum (see arxiv:1511.08692)
  - Experimental tags of 2 b-jets can have several origins (hard bb, gluon splitting, experimental mistags)
- Apply relevant corrections
  - Data to simulation (corners of phase space are not predicted as nicely as inclusive)
  - LO to NLO (when NLO fully simulated samples not available)
- Take into account V_pt uncertainties
  - V_pt is a crucial variable to separate signal from background

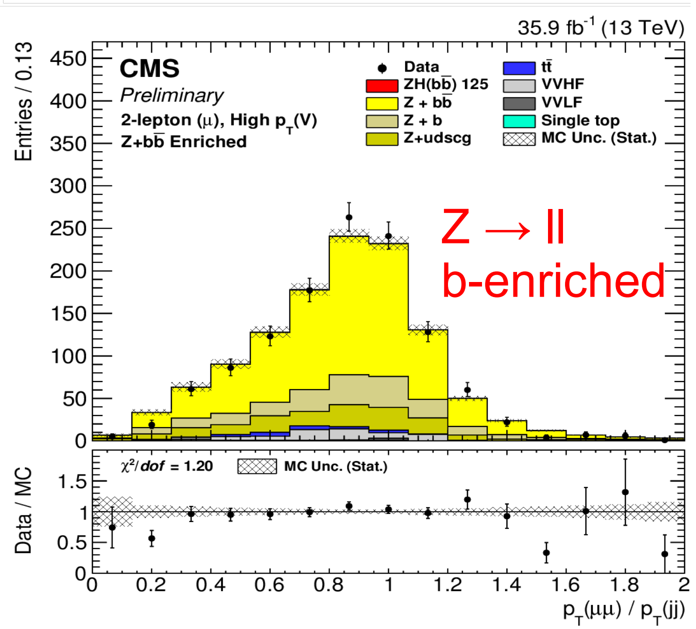


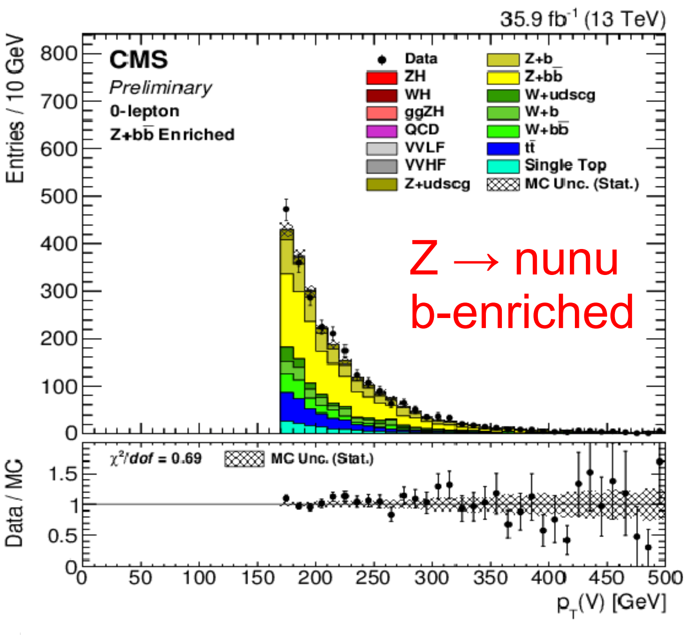
Background control regions

- Vector pt (and other important variables) checked after all corrections in dedicated control regions
- ttbar control region
- V+light control region
- V+heavy control region



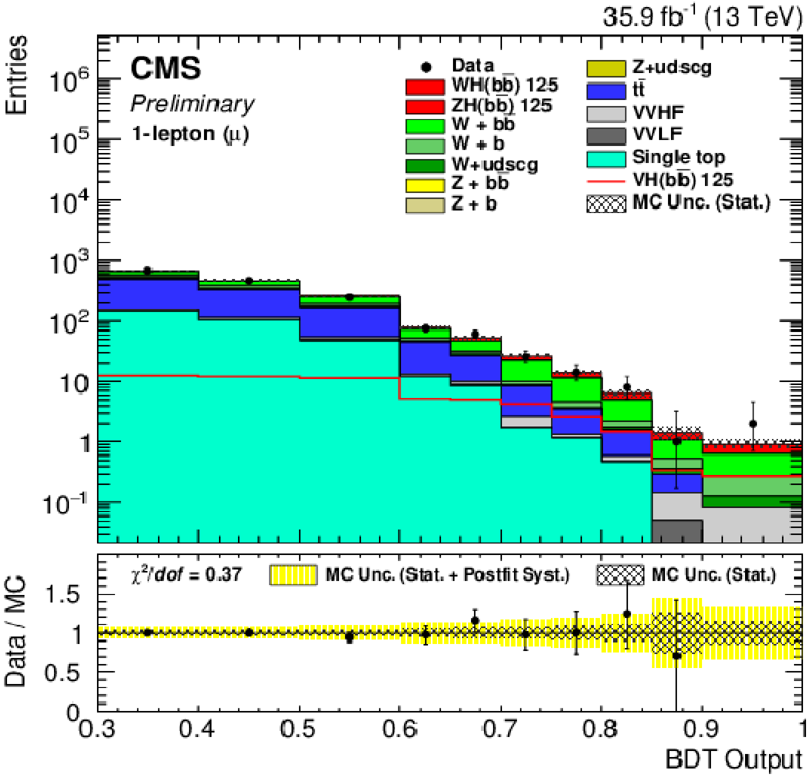






Signal vs background separation

- Several features can be combined to separate S-B
- Multivariate analysis based on Boosted Decision Tree used to maximize sensitivity



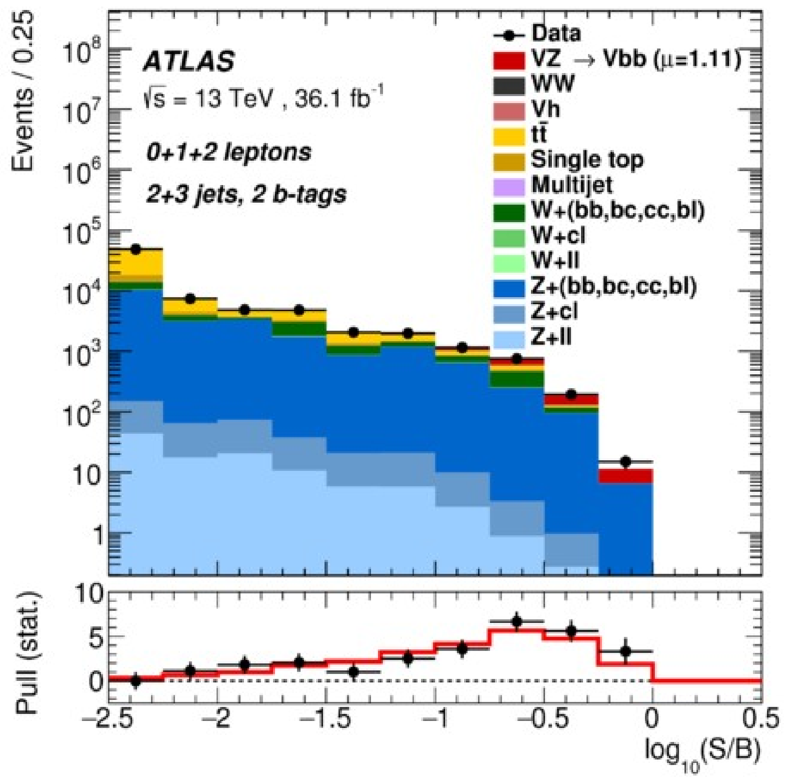
Variable	Channels utilizing
$M(\mathrm{jj})$: dijet invariant mass	All
$p_{\mathrm{T}}(\mathrm{jj})$: dijet transverse momentum	All
$p_{\mathrm{T}}(\mathrm{V})$: vector boson transverse momentum	All
$\mathrm{CMVA}_{\mathrm{max}}$: value of CMVA for the Higgs boson daughter with largest CSV value	2-lepton, 0-lepton
$\mathrm{CMVA}_{\mathrm{min}}$: value of CMVA for the Higgs boson daughter with second largest CSV value	All
$\mathrm{CMVA}_{\mathrm{add}}$: value of CMVA for the additional jet with largest CSV value	0-lepton
$\Delta\phi(\mathrm{V,H})$: azimuthal angle between V and dijet	All
$p_{\mathrm{T}}(j)$: transverse momentum of each Higgs boson daughter	2-lepton, 0-lepton
$p_{\mathrm{T}}(\mathrm{add.})$: transverse momentum of leading additional jet	0-lepton
$\lvert\Delta\eta(\mathrm{jj})\rvert$: difference in η between Higgs boson daughters	2-lepton, 0-lepton
$\Delta R(\mathrm{jj})$: distance in η–ϕ between Higgs boson daughters	2-lepton
N_{aj}: number of additional jets N.B. definition slightly different per channel	1-lepton, 2-lepton
$p_{\mathrm{T}}(\mathrm{jj})/p_{\mathrm{T}}(\mathrm{V})$: p_{T} balance between Higgs boson candidate and vector boson	2-lepton
: Z boson mass	2-lepton
SA5: number of soft activity jets with $p_{\mathrm{T}} > 5$ GeV	All
M_{t}: reconstructed top quark mass	1-lepton
$\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \ell)$: azimuthal angle between $E_{\mathrm{T}}^{\mathrm{miss}}$ and lepton	1-lepton
$E_{\mathrm{T}}^{\mathrm{miss}}$: missing transverse energy	1-lepton, 2-lepton
$m_{\mathrm{T}}(W)$: W transverse mass	1-lepton
: difference in ϕ between Higgs boson daughters	0-lepton
$\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \mathrm{jet.})$: azimuthal angle between $E_{\mathrm{T}}^{\mathrm{miss}}$ and the closest jet with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$	0-lepton

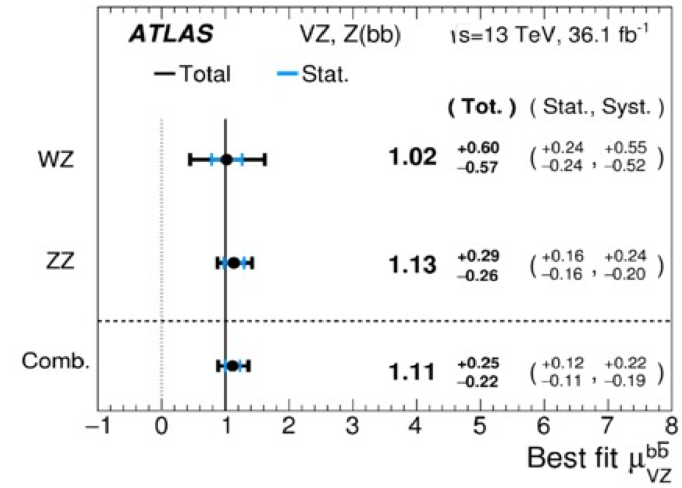
BDT? Can we trust BDT?

- Are BDT analysis reliable? (the problem is not the BDT)
  - The question should rather be, are our predictions of the backgrounds reliable
  - Can we really model all uncertainties?
- In general any background prediction is reliable only up to some given number of data events (real detector simulation is hard to model, qft calculations are limited in precision, etc.. etc...)
  - Typical “shape analysis” fits the data with signal and background models having some “uncertainty” parameterized as a scalar “nuisance parameter”
  - We need to be sure the background model have enough freedom in the fit
- So what? … cross check “candle” analysis!

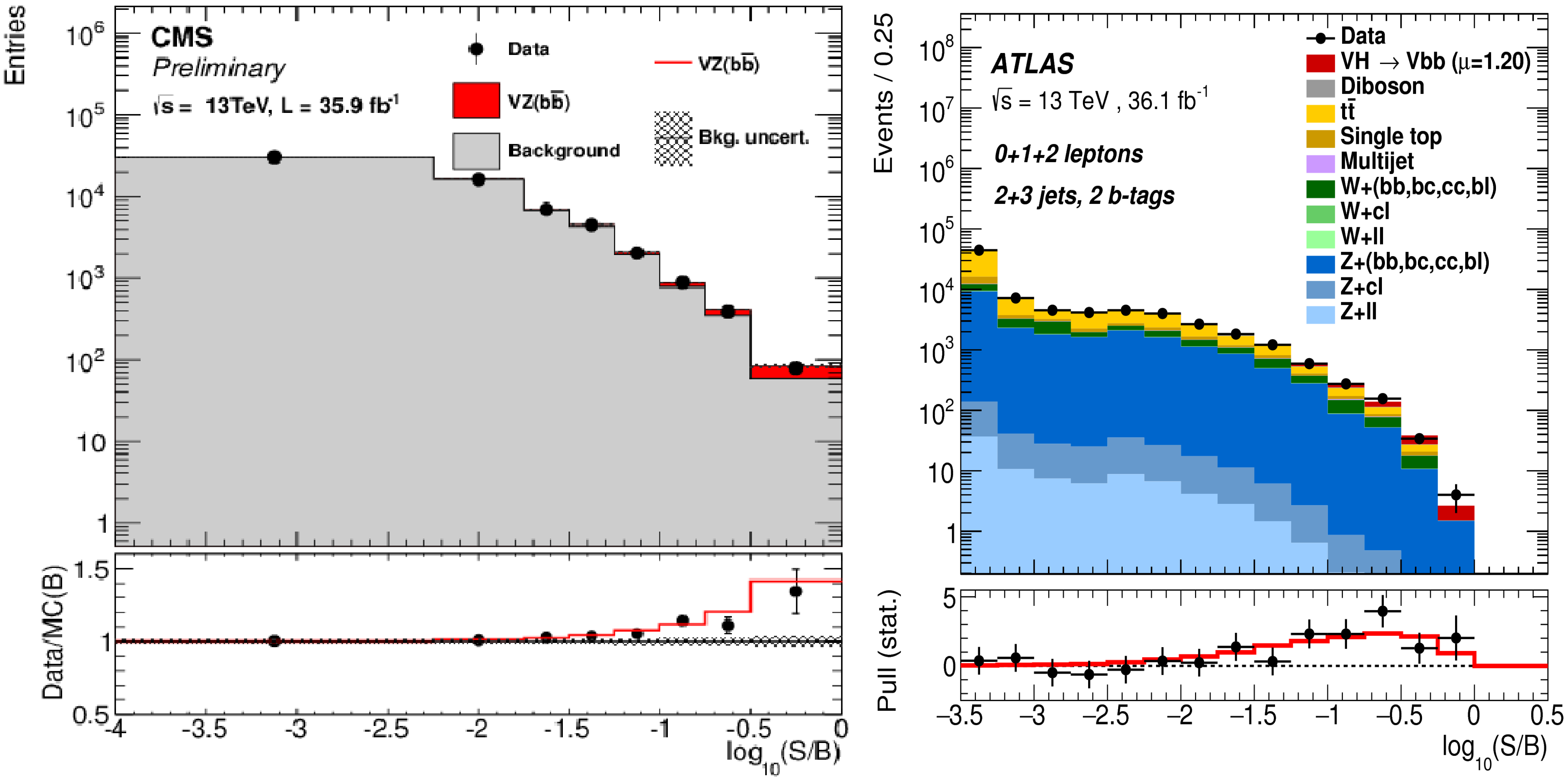
Diboson: V+Z(bb)

- Try to measure VZ (with Z → bb) instead of VH using the very same analysis technique
- Larger cross section
  - Observe with 5 sigma
- Expected theory value already cross checked in leptonic final states
- Main differences
  - Different peak value of the dijet mass (but reasonably close)
  - Peaking non b-jet signal (but b-tag is well studied in other processes)
  - Di-boson is a background for Higgs, but higgs is less relevant background for di-boson (assume SM higgs x-sec)





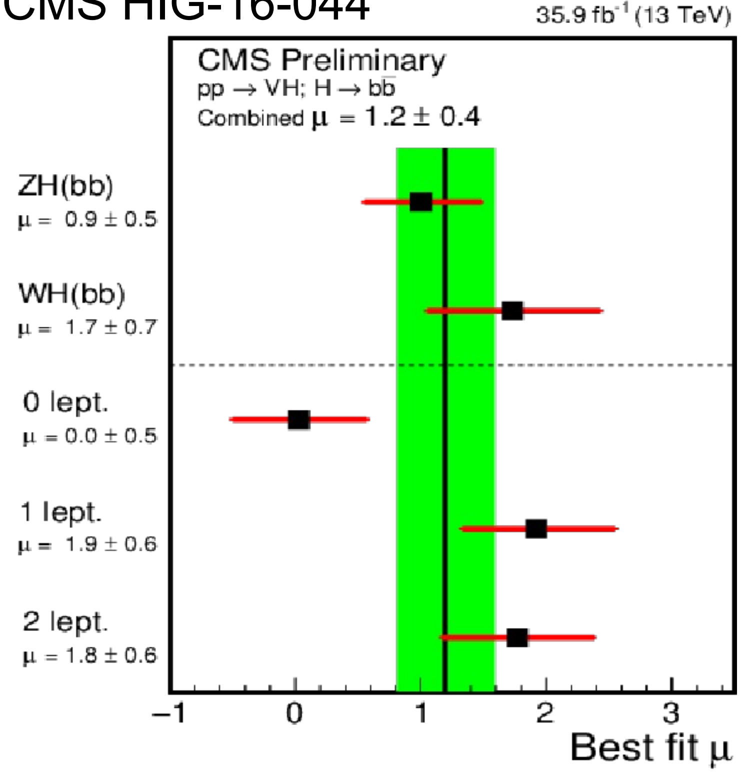
Higgs results

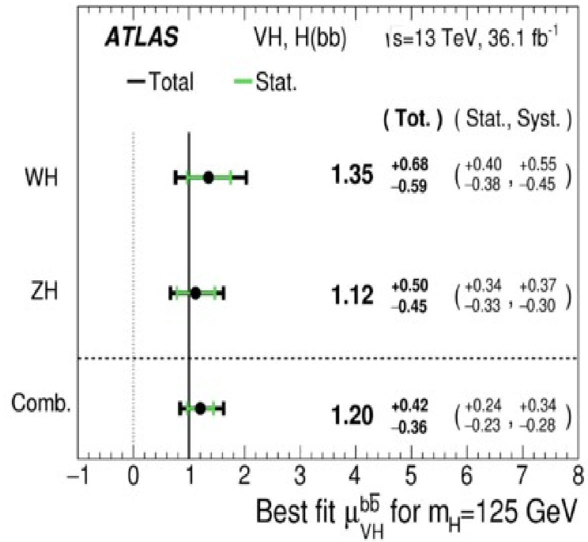


- Very nice agreement in control regions
- Signal enriched bins compatible with SM predictions

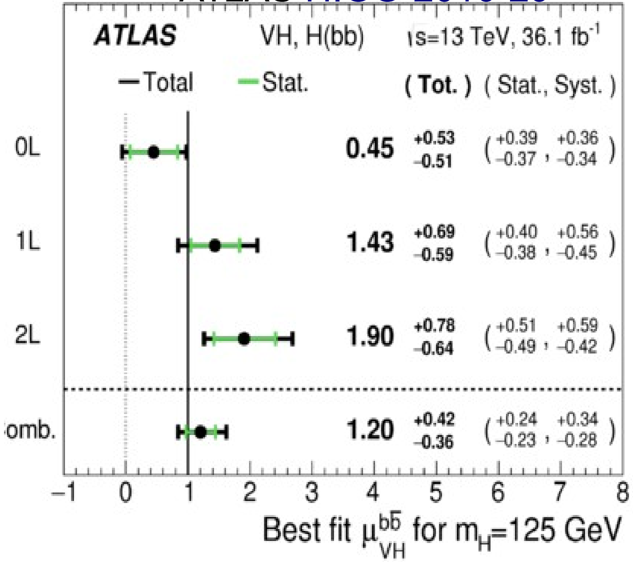
Evidence of higgs to bb

CMS HIG-16-044





ATLAS HIGG-2016-29



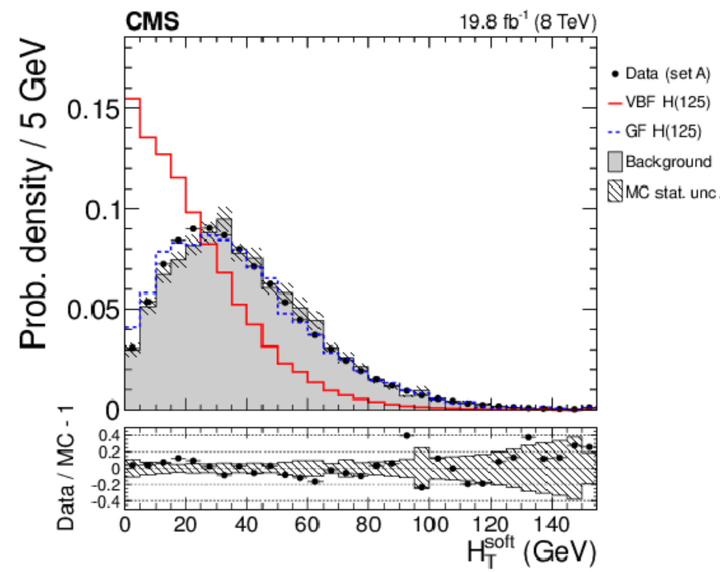
- Above 3 sigma in both experiments!

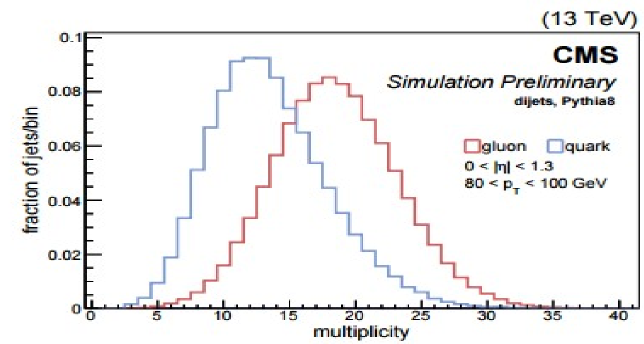
$m\text{H} = 125\,\text{GeV}$	Significance expected	Significance observed
0-lepton	1.5	0.0
1-lepton	1.5	3.2
2-lepton	1.8	3.1
All channels	2.8	3.3

$m\text{H} = 125\,\text{GeV}$	Significance expected	Significance observed	Signal strength observed
Run 1	2.5	2.1	$0.89^{+0.44}_{-0.42}$
Run 2	2.8	3.3	$1.19^{+0.40}_{-0.38}$
combined	3.8	3.8	$1.06^{+0.31}_{-0.29}$

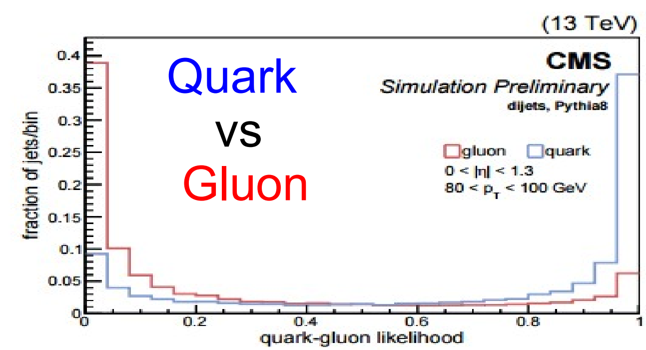
VBF Hbb

- Exploit VBF features to categorize events (with a mass-independent BDT)
  - Forward tag: quark vs gluon discrimination
  - Rapidity gap, m_jj mass, soft activity
- Fit (wide)peak on a smooth multi-jet qcd background
  - Validate with Z->bb (VBF “pre-selection” but no VBF-BDT requirements)

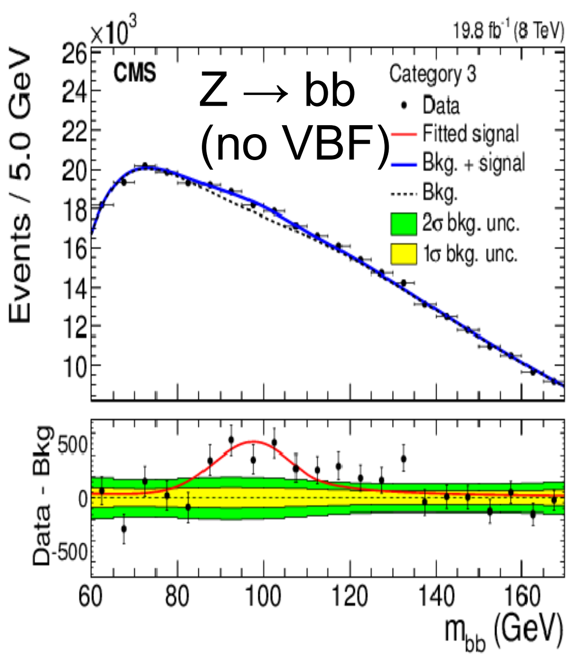


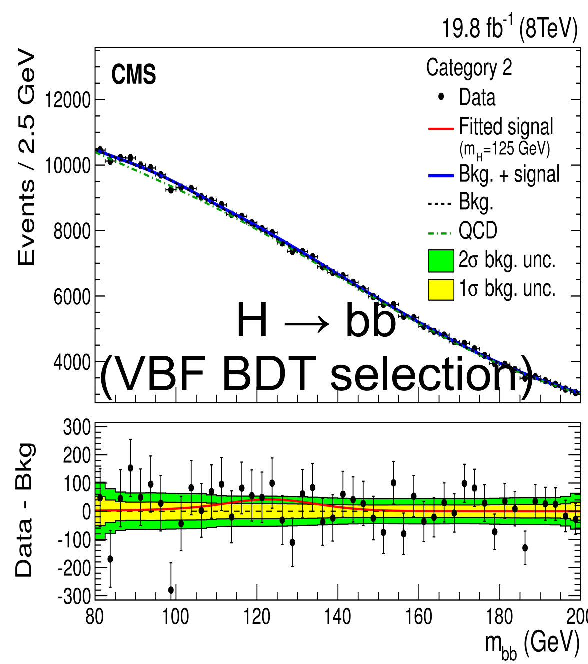


(b)



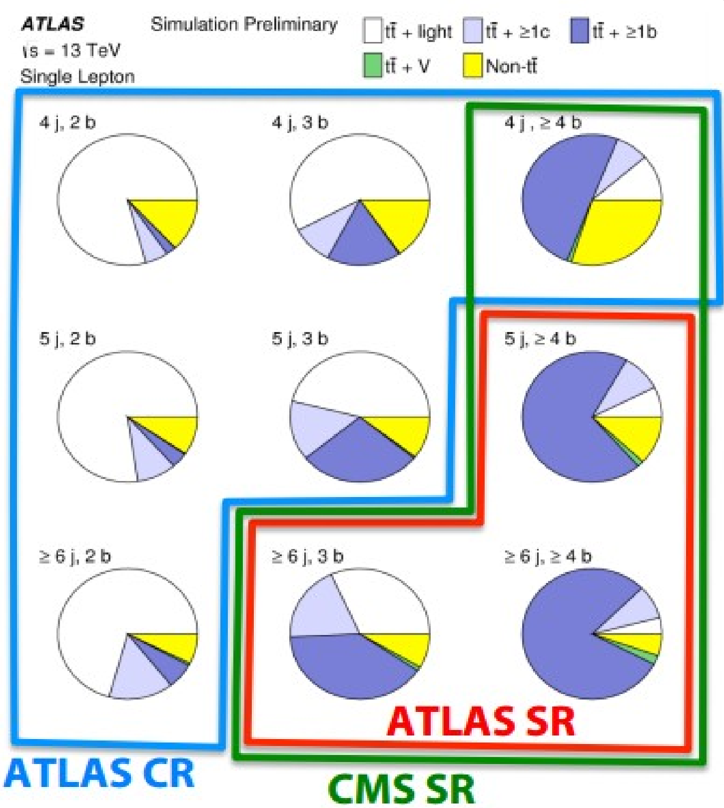
(d)

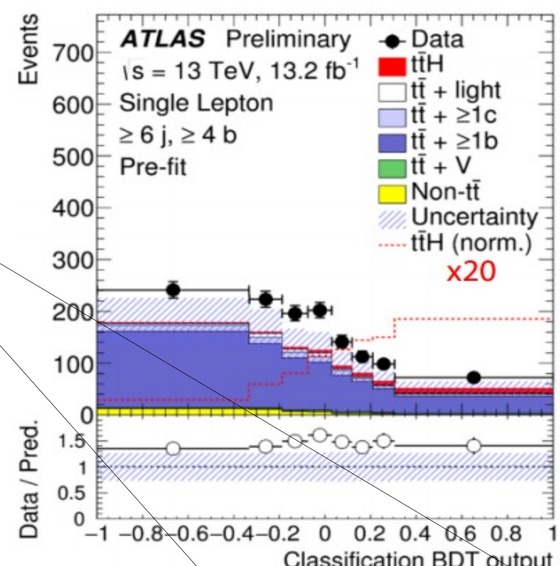


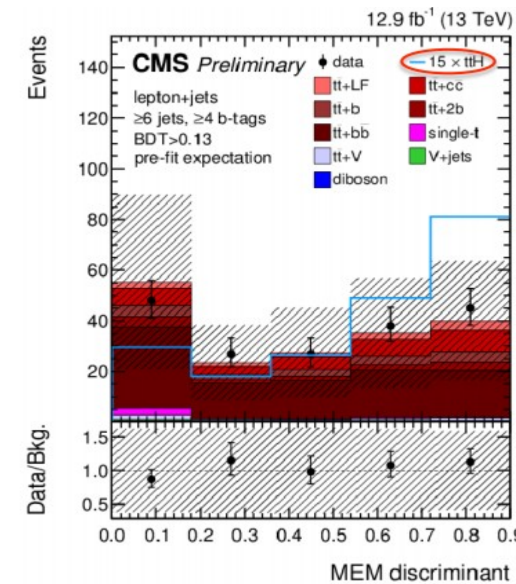


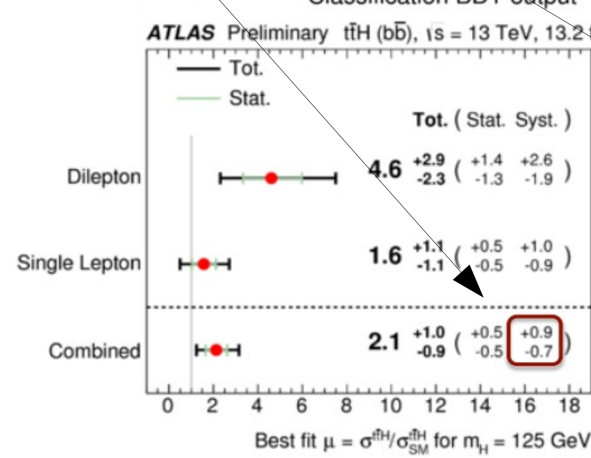
ttH, H→ bb

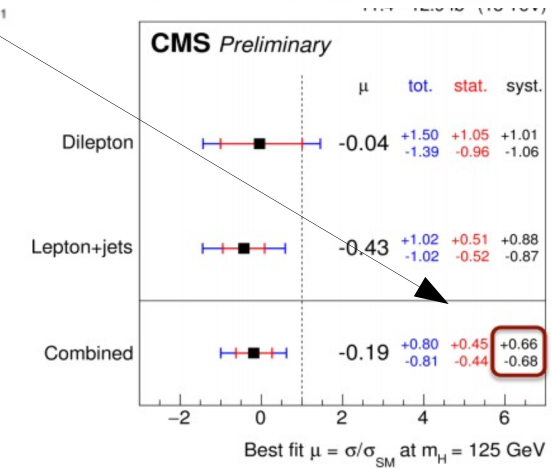
- Similar techniques in ATLAS and CMS
  - Categorize in N-jets and N-bjets, define CR/SR
  - Dominated by irreducible tt+bb background (no “mass peak”)
  - Difficulties in modelling collinear/soft bb production
- Systematics are the largest contribution to the final uncertainty





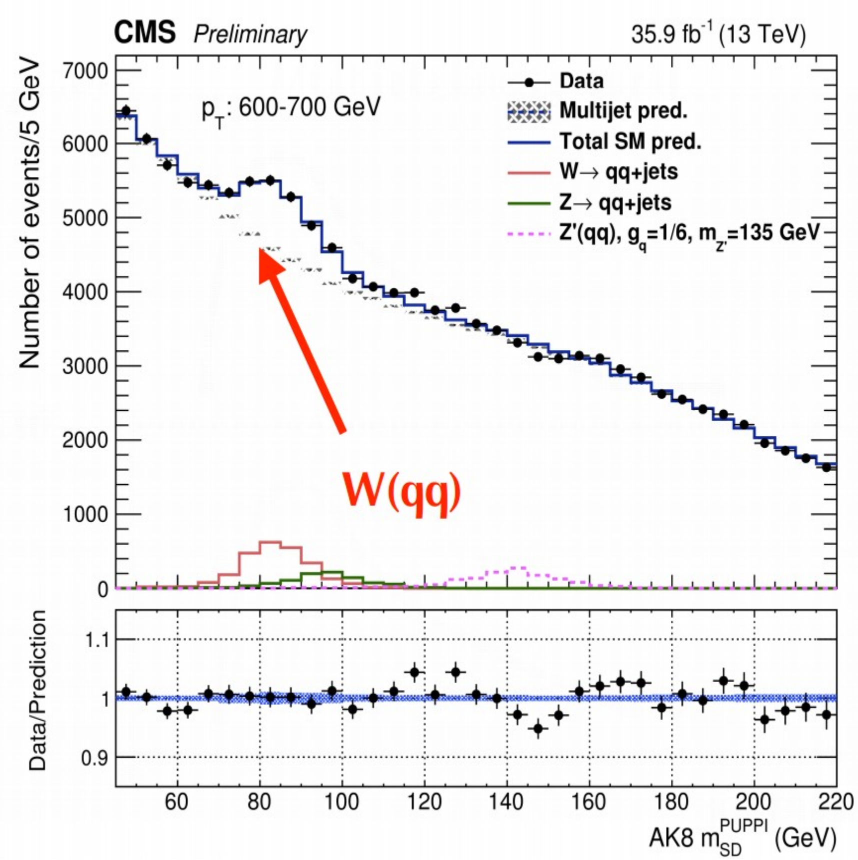






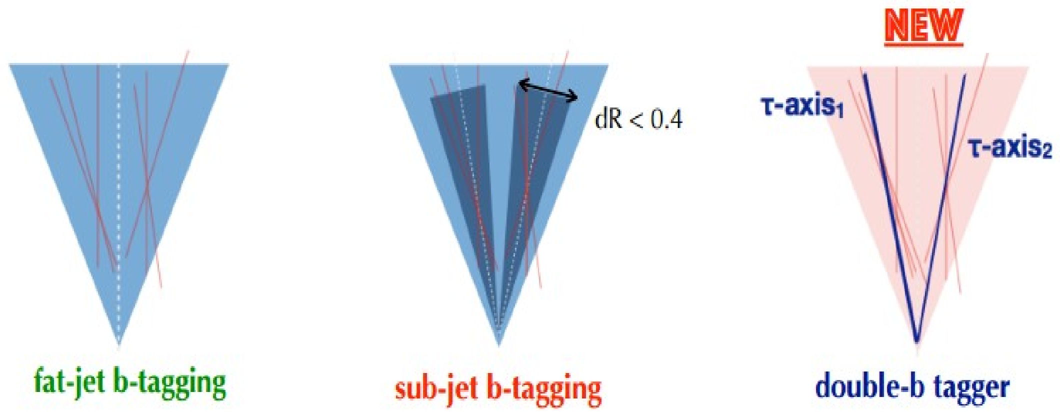
How about ggH, H → bb ?

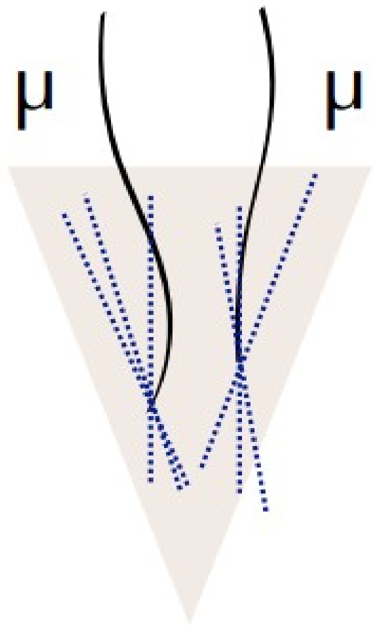
- Can we measure H → bb also in gluon fusion production?
  - QCD "bb" production is several orders of magnitude higher
  - How about the "boosted" regime ?
- Jet substructures techniques tested in the past years
  - Observe hadronic final state of V bosons and top
  - Mature for "measurements"
  - Recent studies on subjets b-tagging to distinguish from V-bosons
- New techniques to keep systematic uncertainty under control
  - Decorrelated Taggers
- All ready for boosted H → bb ?

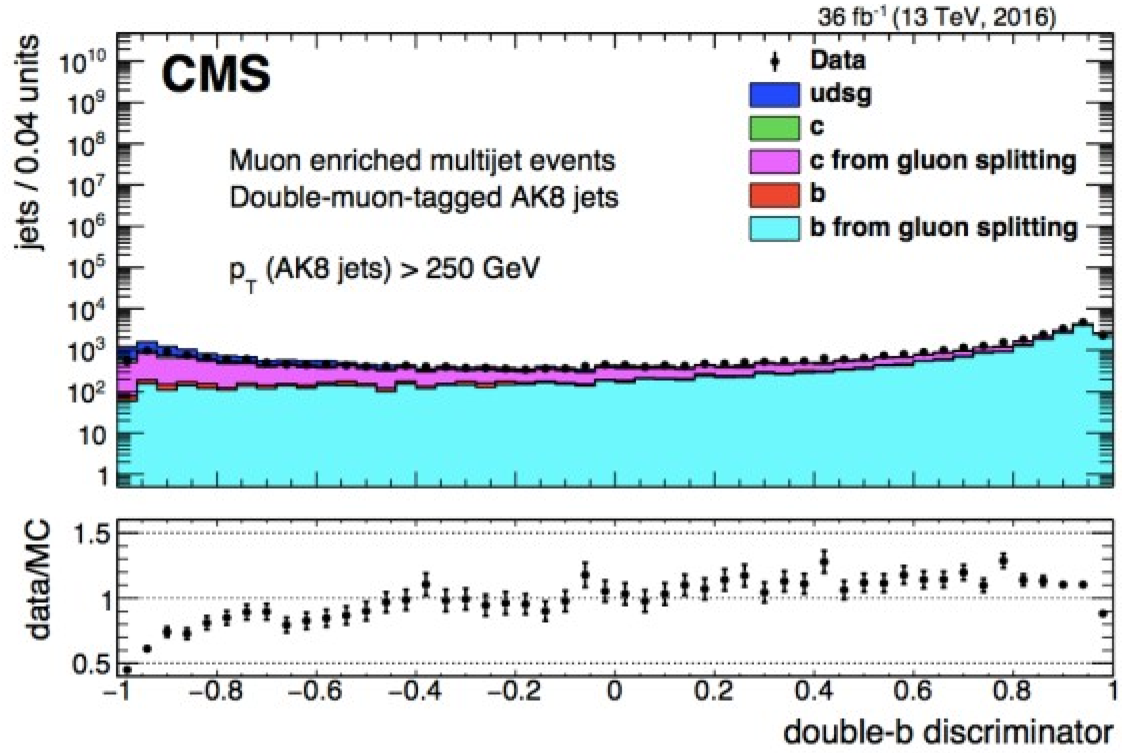


Dedicated bb-tagger

- Multivariate bb-in-a-jet tagger exploiting substructure information (Nsubjetiness axis)
- Measure performance on data with two-muons-in-jet sample (as a proxy for $g \rightarrow bb$)

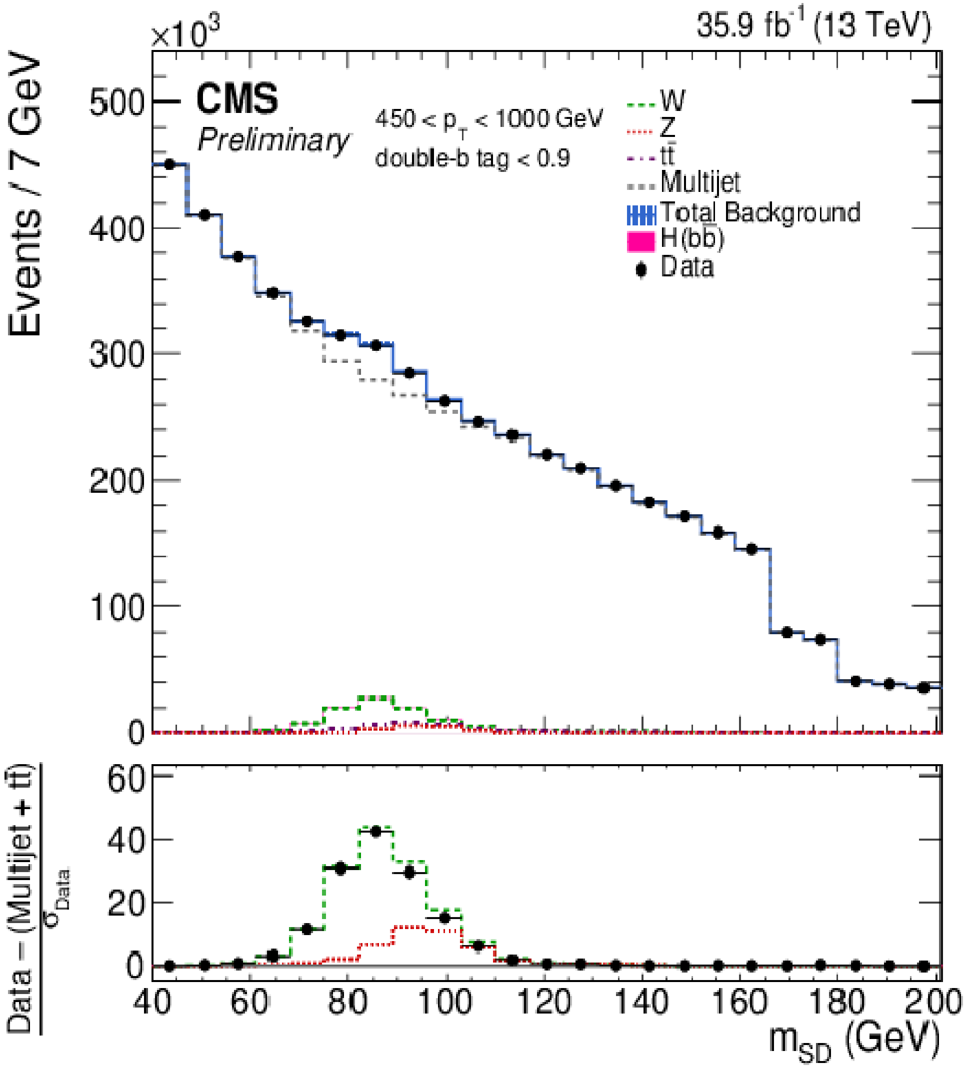




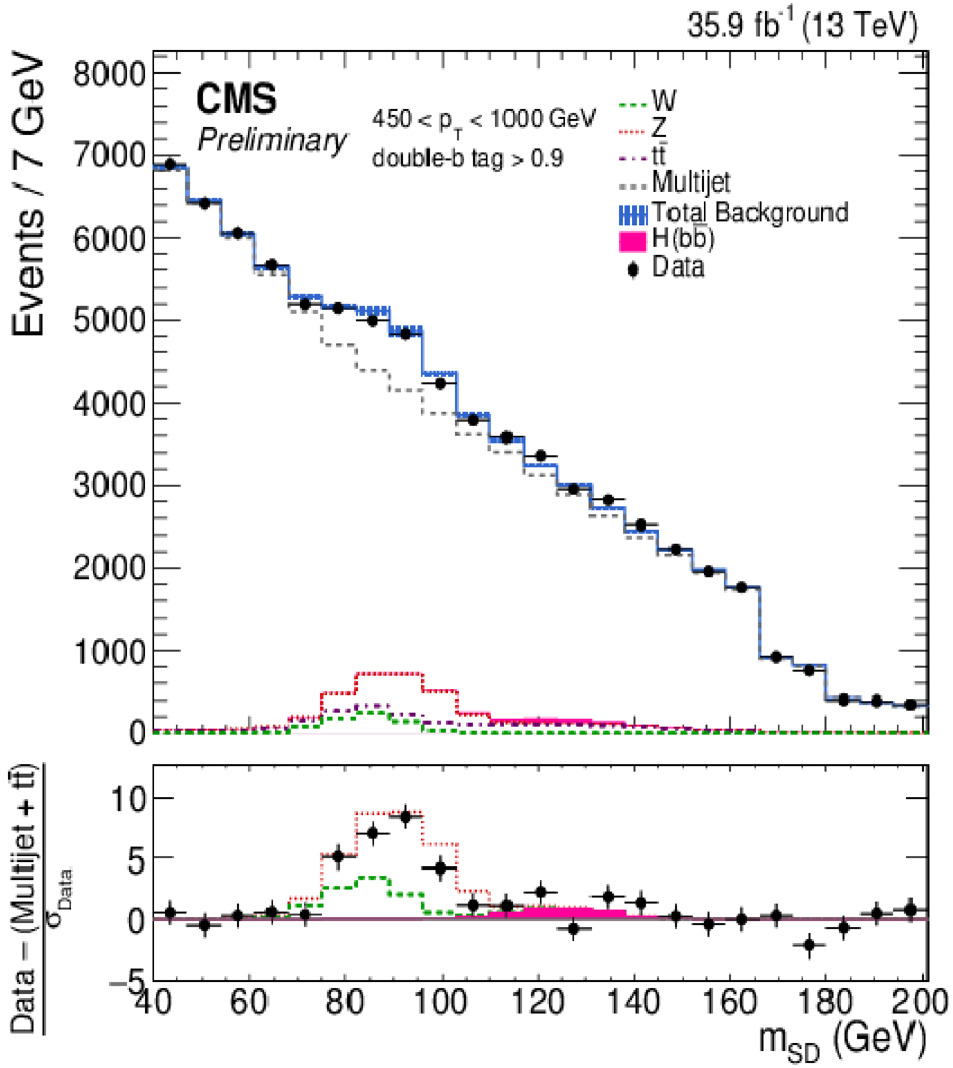


Boosted Higgs to bb

Anti b-tag (control region)



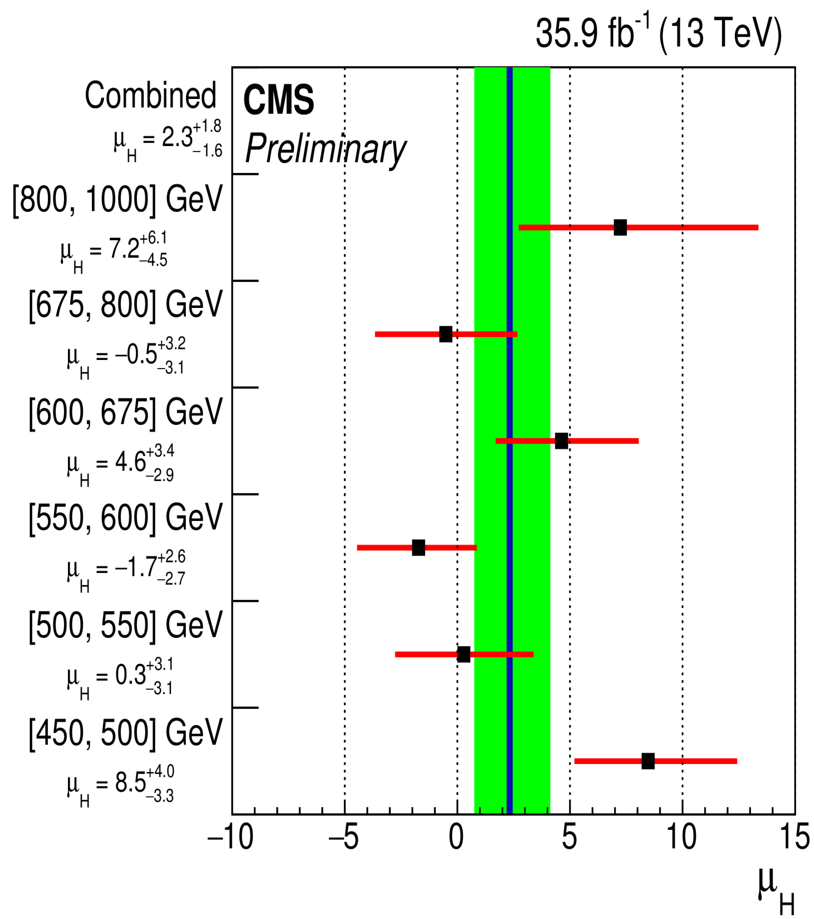
B-tag (signal region)

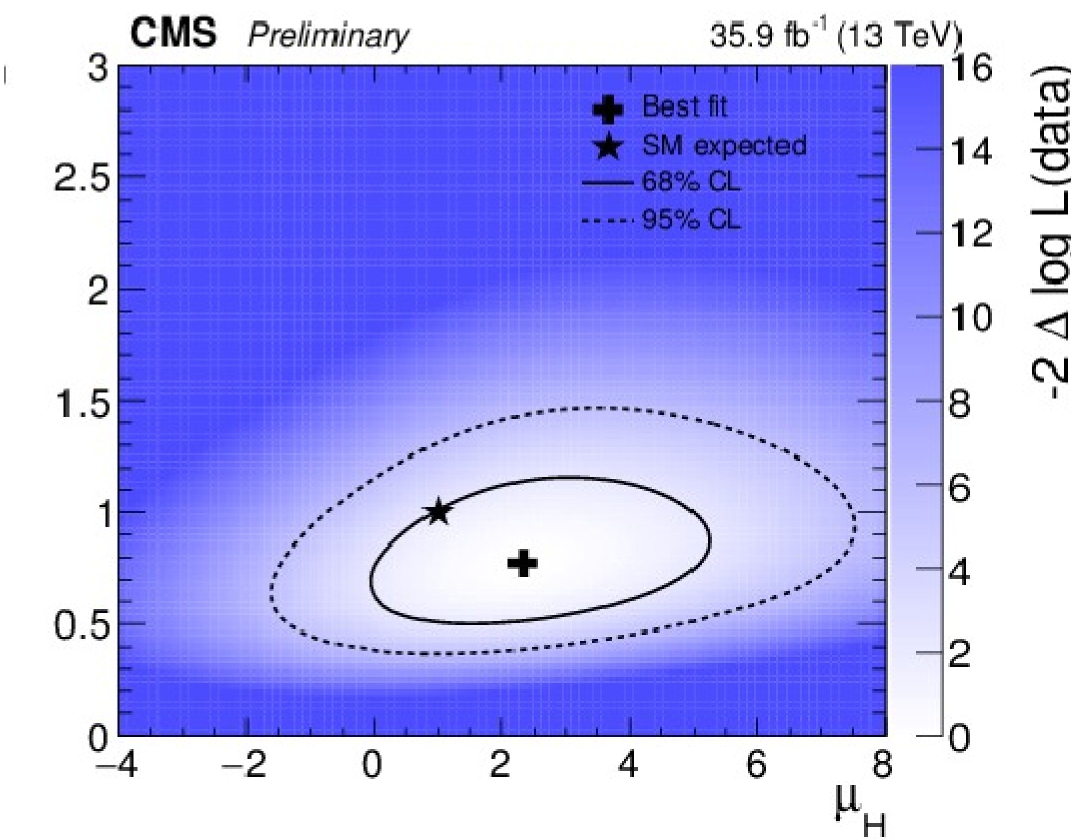


Boosted H->bb results

CMS-PAS-HIG-17-010

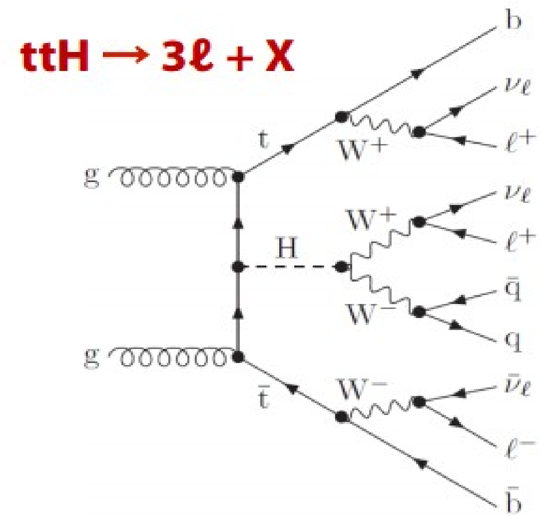
- Sensitivity at 1 sigma level with 36/fb
  - To be continued at high luminosity....



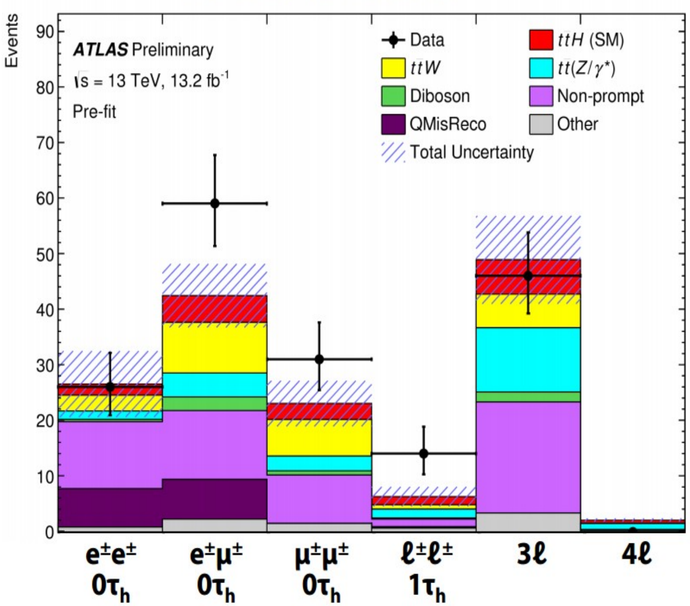


ttH - multileptons

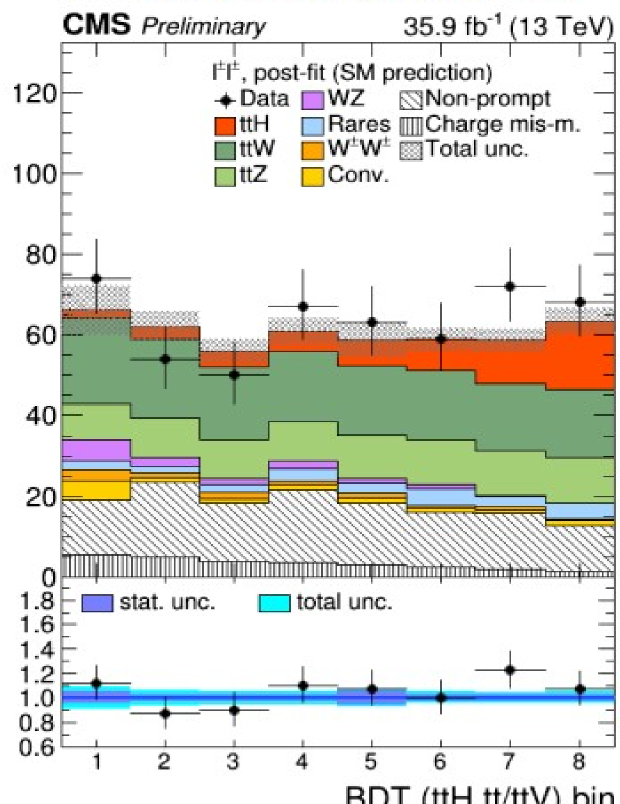
- Decay of H via W/Z
  - Low branching ratios, but cleaner states
  - Look for >3leptons or 2 same-sign lept
- Backgrounds: tt+vector boson, tt+jets, diboson
- Reaching 2/3 sigma sensitivity
  - First evidence of ttH?

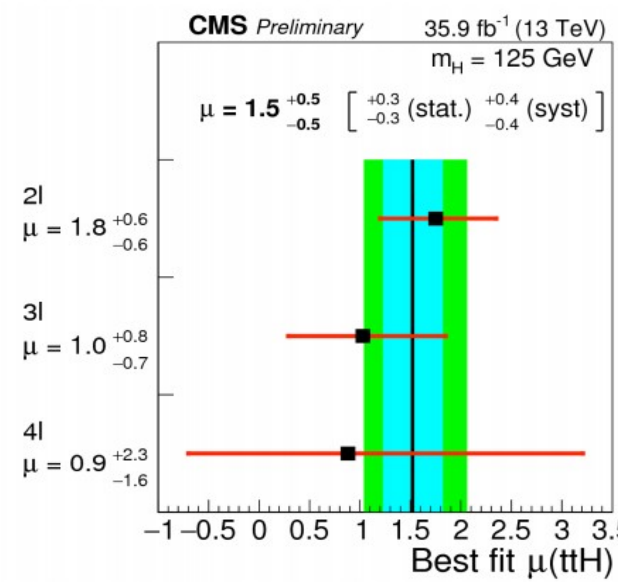


ATLAS data



SR bins (all cat. combined)

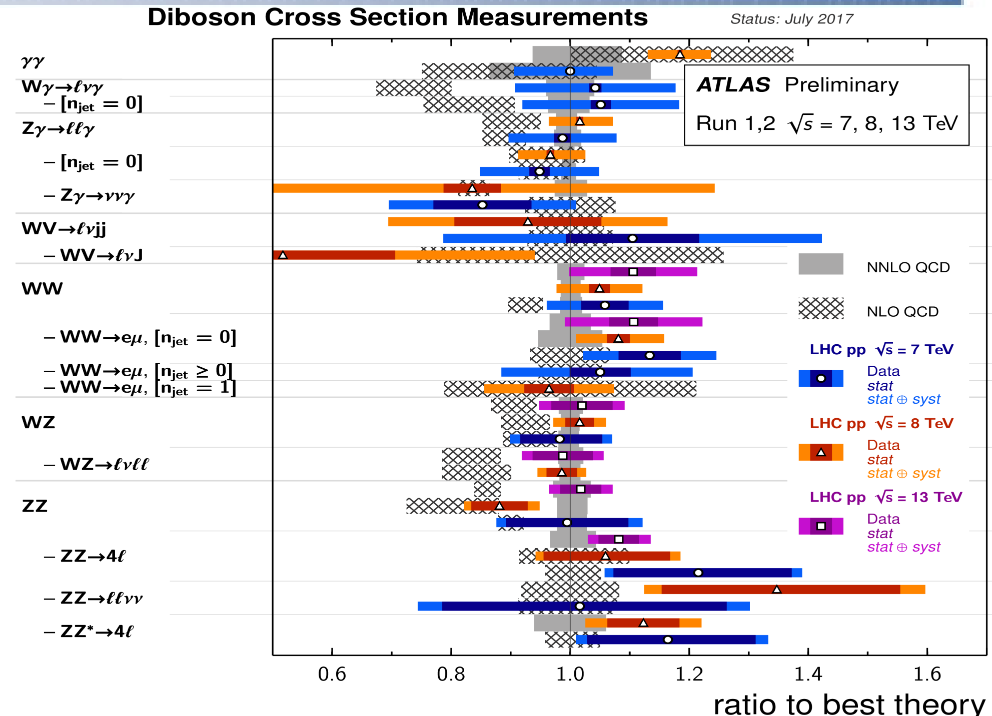


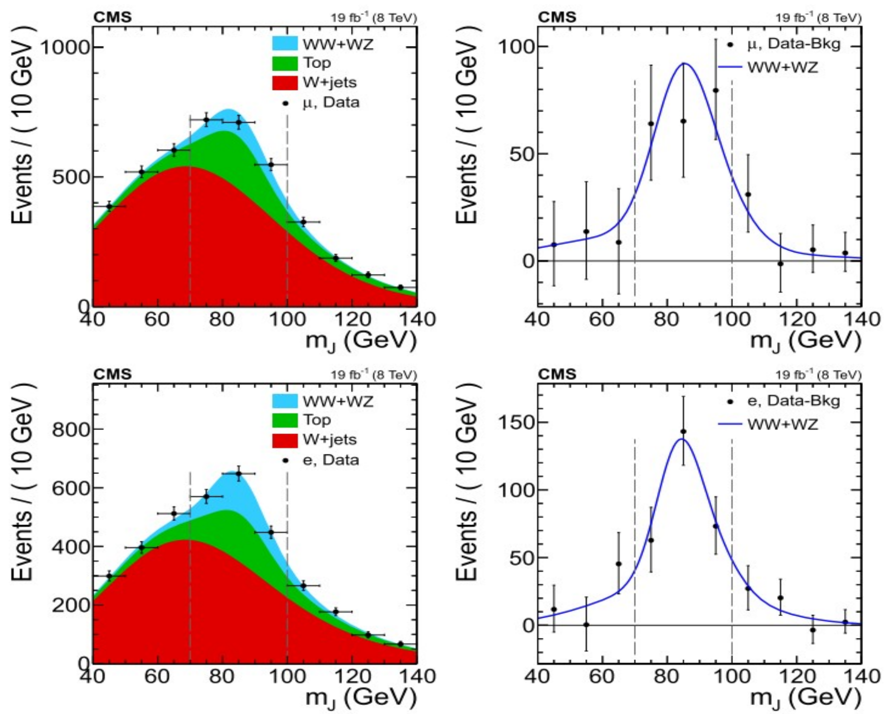


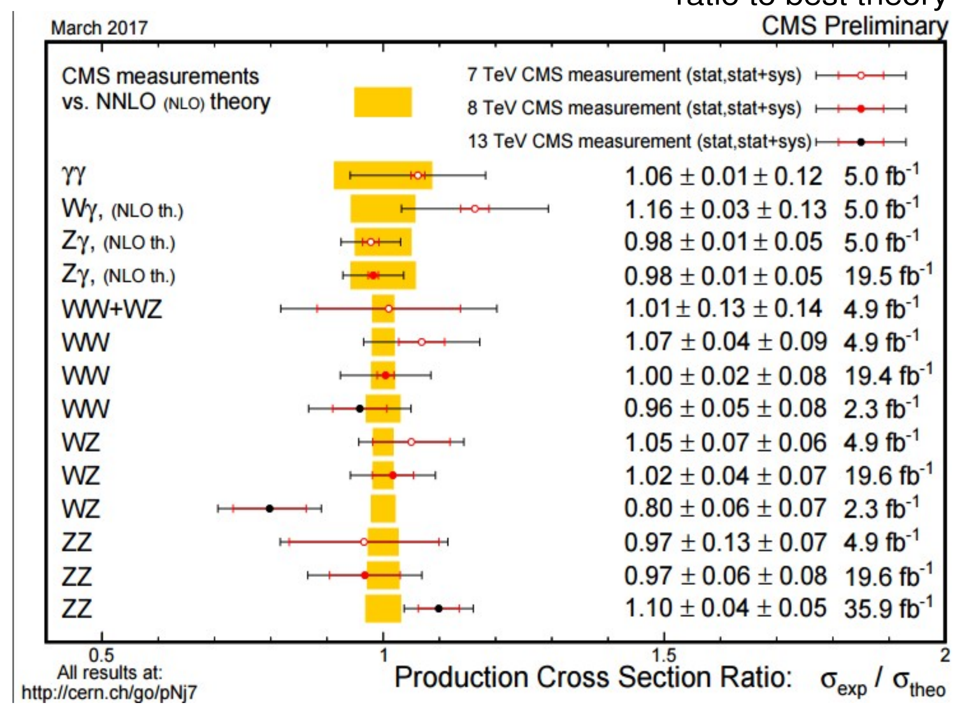
ATLAS: mu=2.5+1.3-1.1

Di-boson

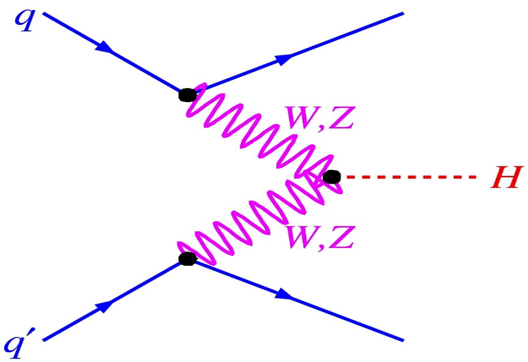
- Various multilepton di-boson analysis repeated at 13 TeV
  - Nice agreement with NNLO QCD
- Boosted analysis performed in final state with lepton+fat-jet
  - Allow to measure W pt in WW channel
  - Probe high pt region where all lepton analysis are limited by sigma*BR

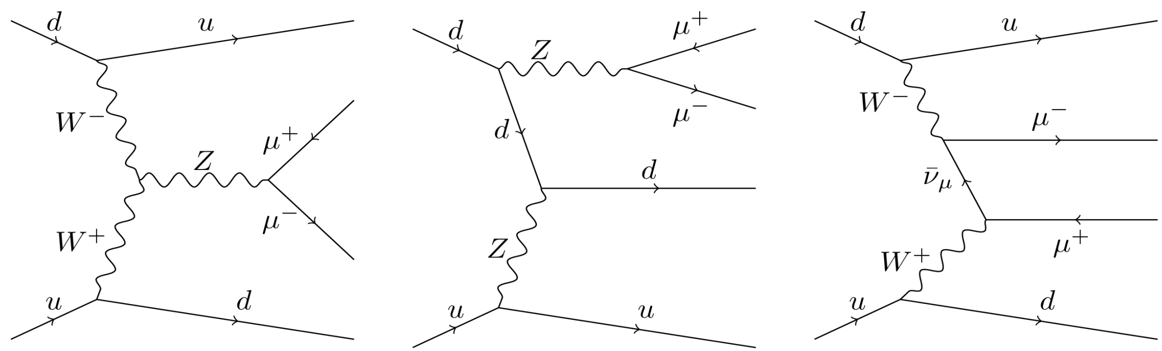




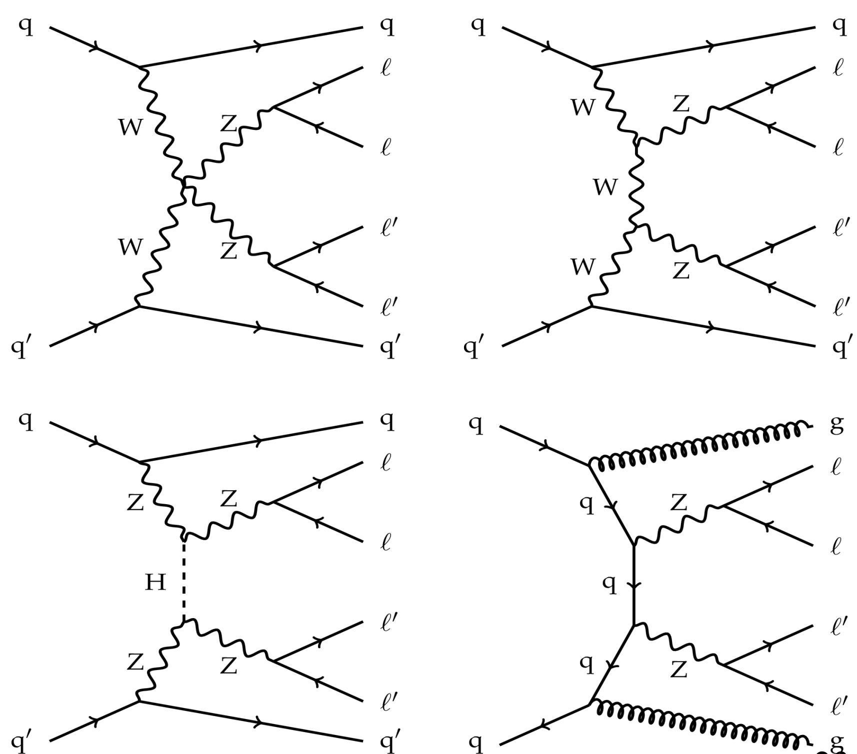


VBF/VBS and alike



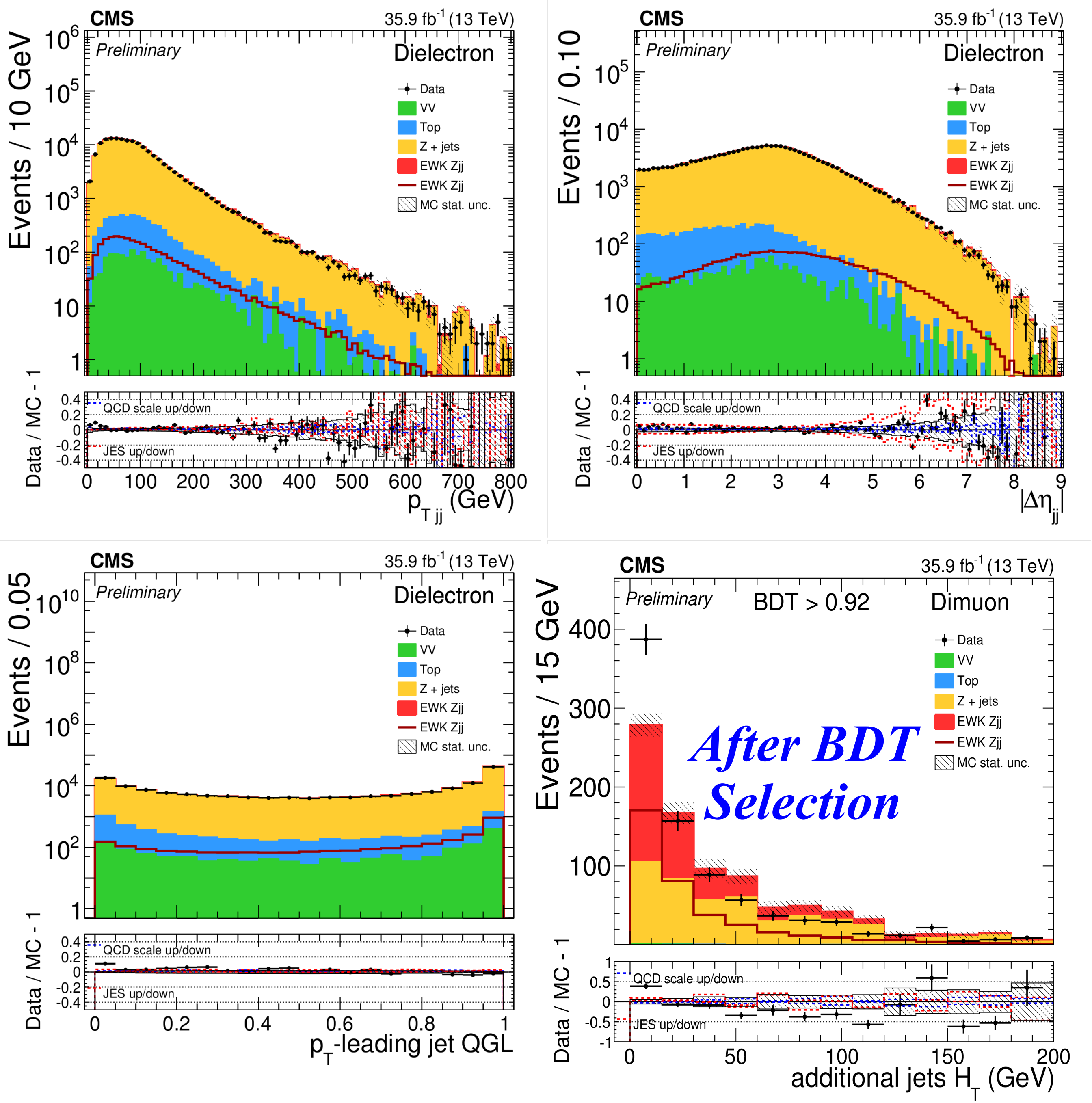


- Same experimental techniques as for VBF Higgs tagging
- Associated production of Z,W, diboson with two jets
- Hard process is typically purely EWK
- QCD play a crucial role in
  - Identification of the two q-jets
  - Background modeling
  - Underlying event activity



QCD features (e.g. Z+jj)

- V+jets is the dominant background
- Rapidity gap distinguish S vs B
- Quard vs Gluon discriminator
- Soft-activity variables (e.g. HT of additional jets)



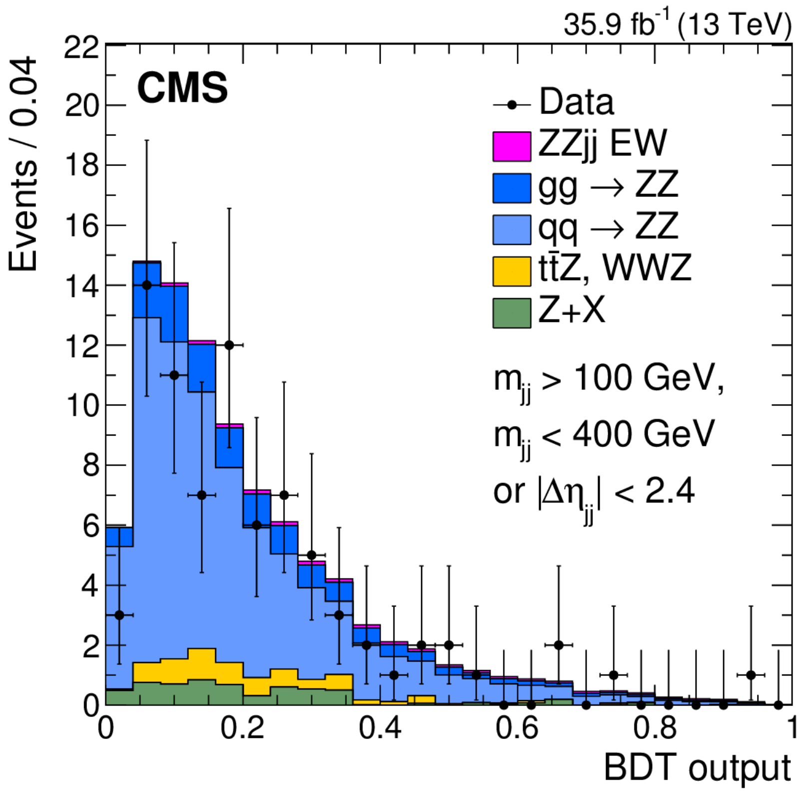
Results of “VBF” analyses

- EWK Zjj production

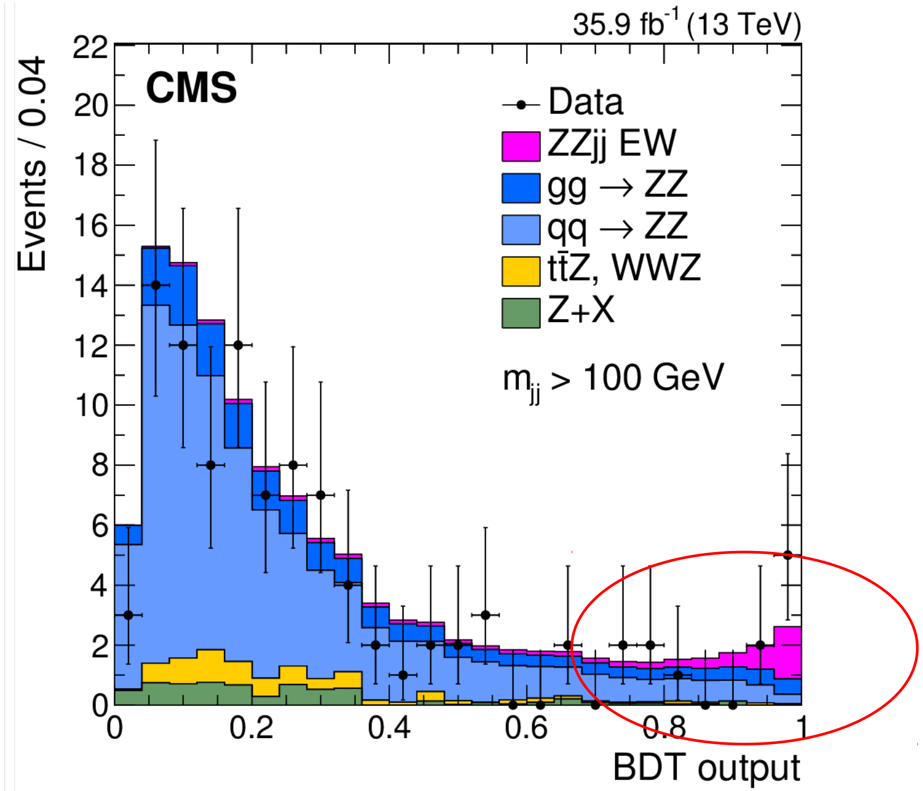
$$\sigma_{\mathrm{EW}}(\ell\ell\mathrm{jj}) = 552 \pm 19\,(\mathrm{stat}) \pm 55\,(\mathrm{syst})\ \mathrm{fb}$$

- EWK diboson production (VBS):

Control region



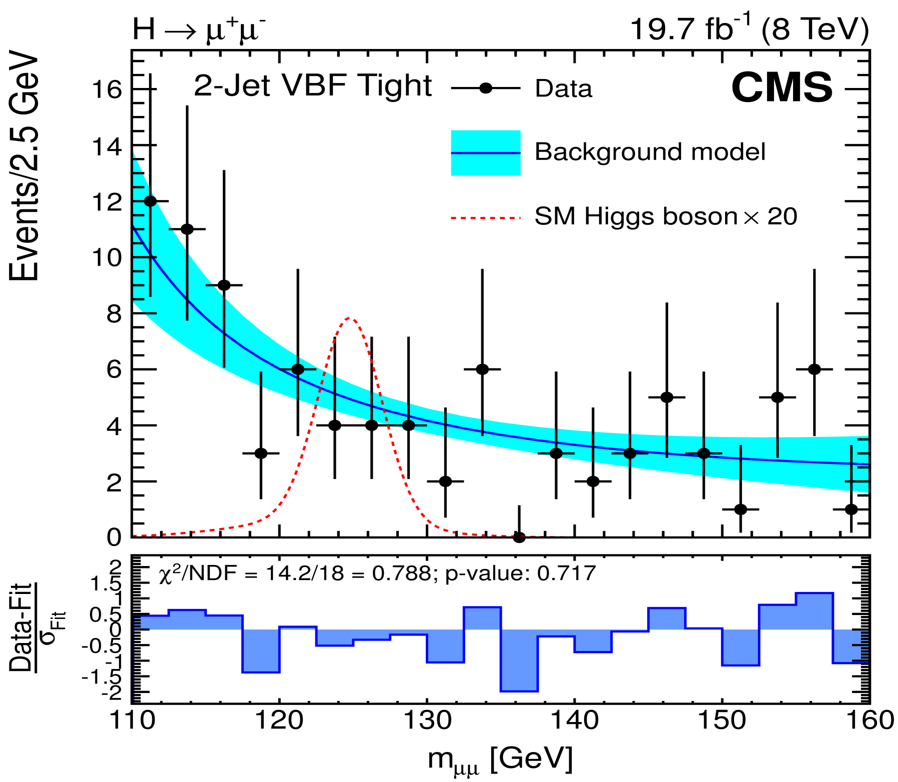
Signal Region

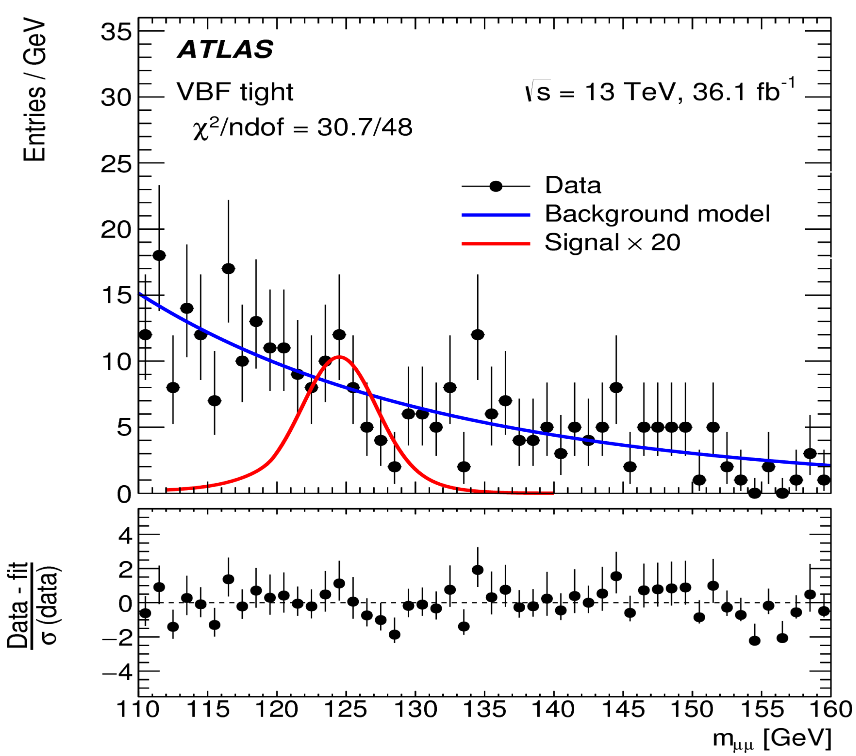


Another interesting VBF

- $H \to \mu\mu$
- VBF signature is the most sensitive and have largest S/B
  - Need precise prediction of V+jets to use multivariate VBF techniques

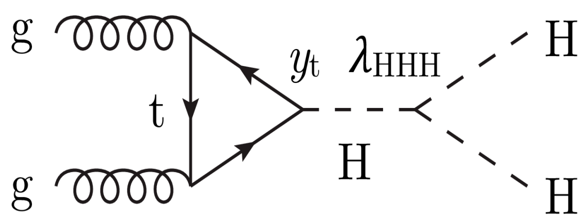
	S	B	$S/\sqrt{B}$
Central low $p_T^{\mu\mu}$	11	8000	0.12
Noncentral low $p_T^{\mu\mu}$	32	38 000	0.16
Central medium $p_T^{\mu\mu}$	23	6400	0.29
Noncentral medium $p_T^{\mu\mu}$	66	31 000	0.37
Central high $p_T^{\mu\mu}$	16	3300	0.28
Noncentral high $p_T^{\mu\mu}$	40	13 000	0.35
VBF loose	3.4	260	0.21
VBF tight	3.4	78	0.38

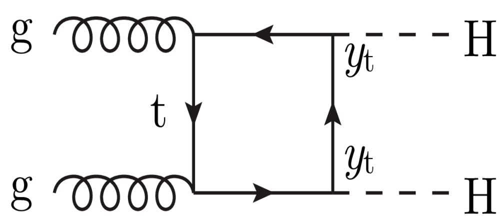




HH

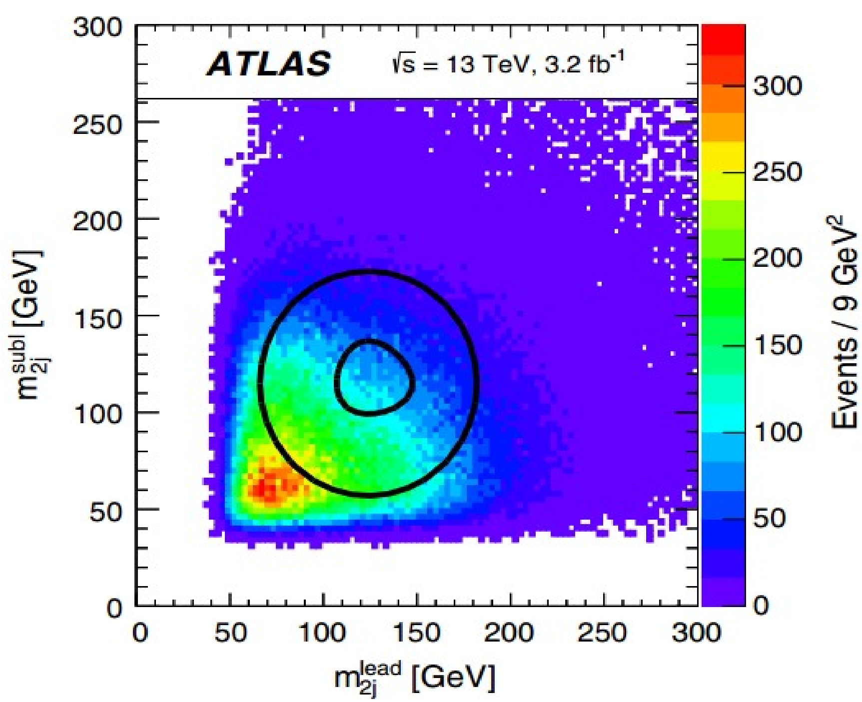
- Higgs boson pairs interesting both for SM and exotics

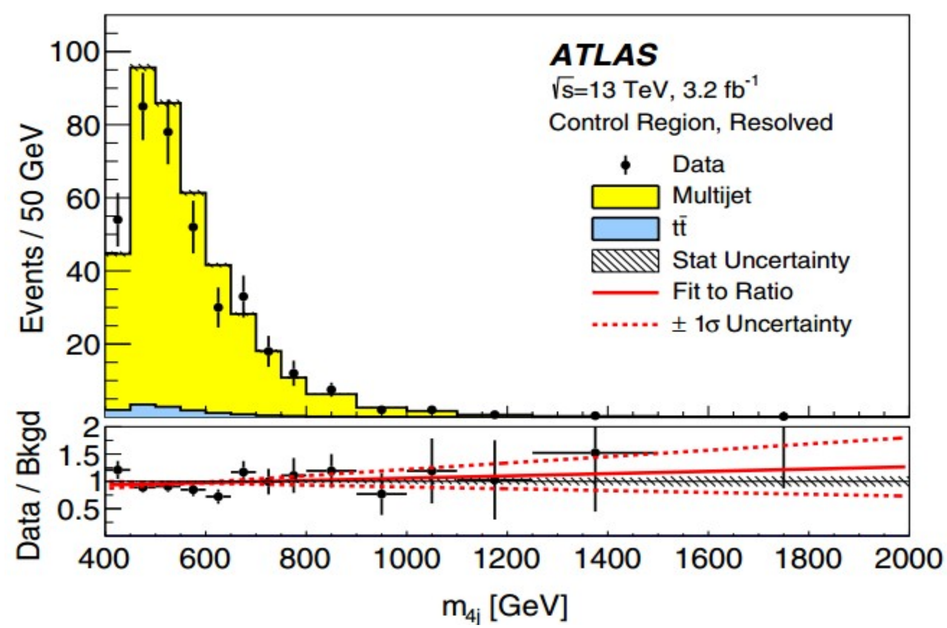




- SM sensitivity is still far away
- bb+bb mode has largest BR
  - Best sensitivity for m_res> 500GeV

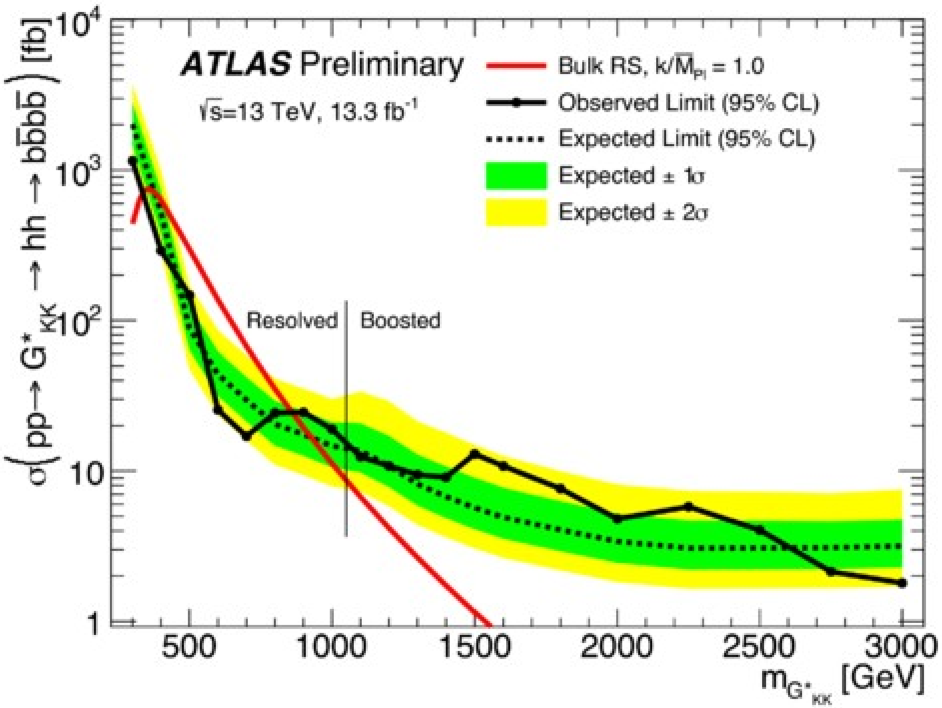
Sample	2015 Signal Region	2016 Signal Region
Multijet	$1\,131 \pm 68$	$3\,670 \pm 200$
$t\bar{t}$	57 ± 34	190 ± 110
Total	$1\,189 \pm 76$	$3\,860 \pm 230$
Data	1 231	3 990
SM hh	0.47 ± 0.12	1.5 ± 0.4
G^*_{KK} (800 GeV), $k/\bar{M}_{Pl} = 1$	8 ± 3	24 ± 8

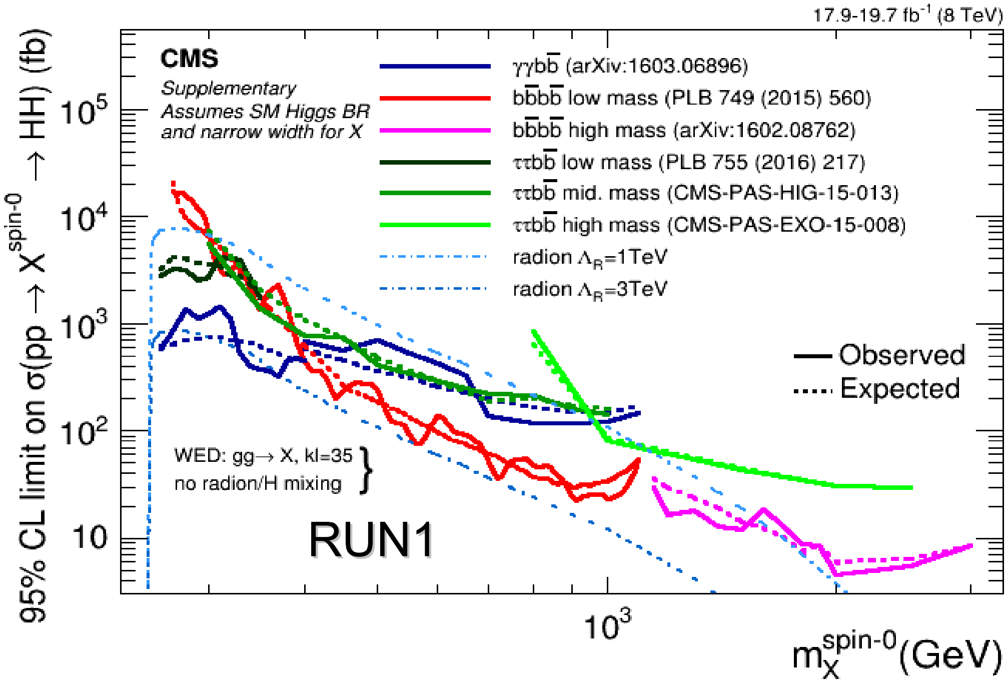


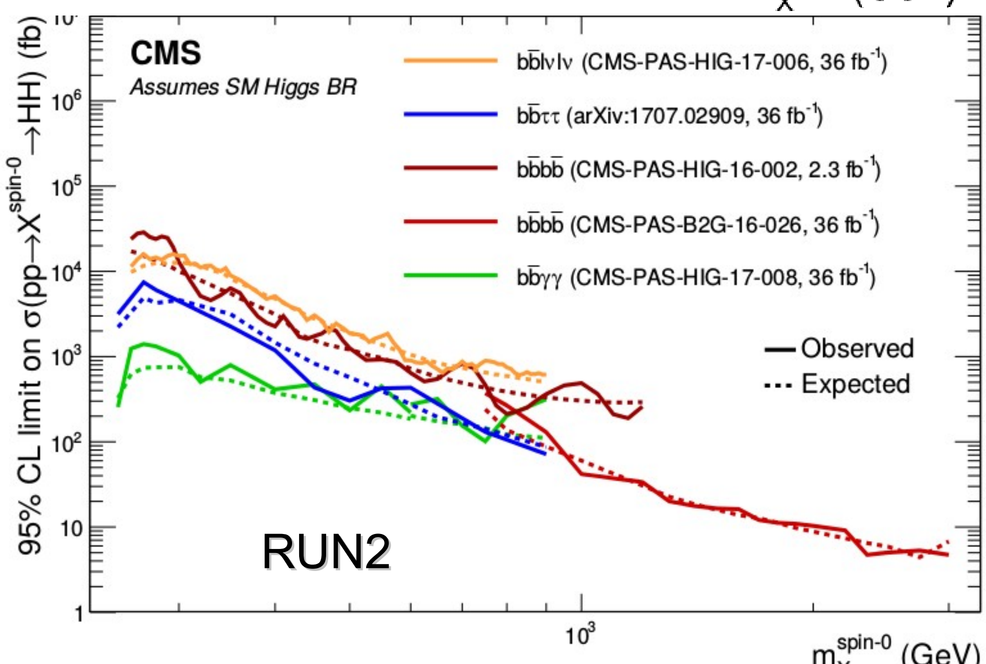


HH results

- Resonant analysis: low mass – resolved, high mass – boosted







- Non-resonant SM, ATLAS 4b:

$$\sigma(pp \to hh \to b\bar{b}b\bar{b}) < 330 \text{ fb},$$

- 30x the SM prediction:

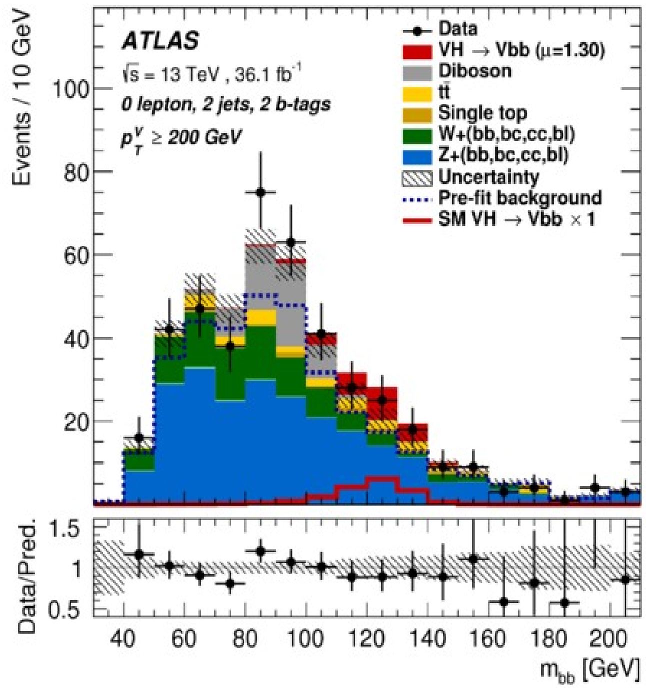
$$\sigma(pp \to hh \to b\bar{b}b\bar{b}) = 11.3^{+0.9}_{-1.0} \text{ fb}$$

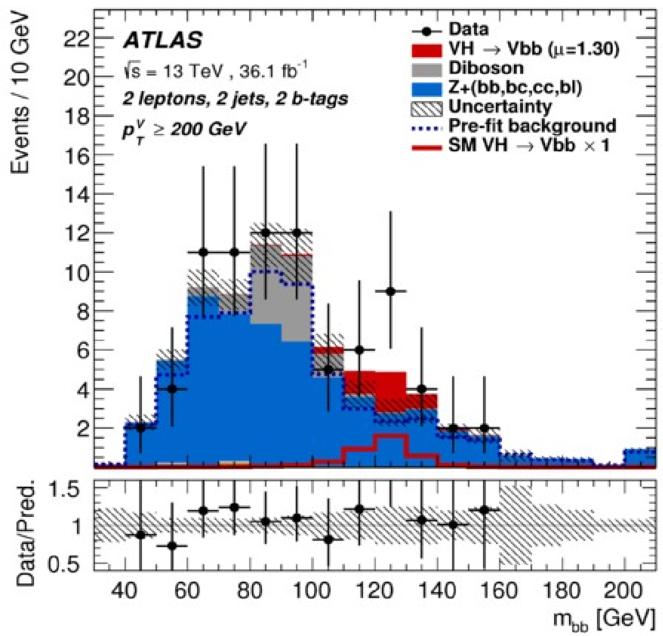
Conclusions

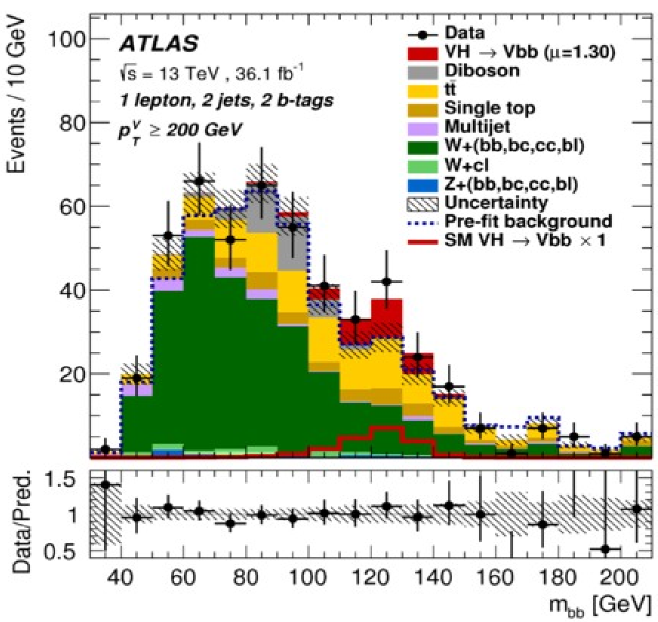
- Many measurements ongoing at high-pt scale in Run2
- Current understanding of hard and soft QCD processes is often a limiting factor
- Vector plus jet is an everywhere background
  - Lot of theory calculation available in large regions of the phase space with NNLO accuracy
  - Large uncertainties in heavy flavour (gluon splitting) realm
- Data driven background models often(always) used for multijet processes
- Long term targets/more precise measurements will require perfect understanding of key kinematic distributions (perhaps it is not just a matter of how many Ns in front of LO)
- Stay tuned for $H \rightarrow$ mumu

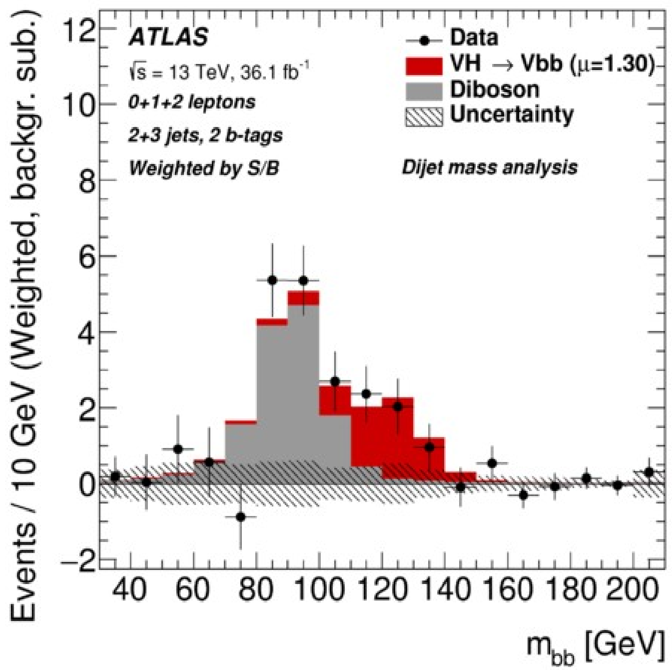
backup

Hbb mass

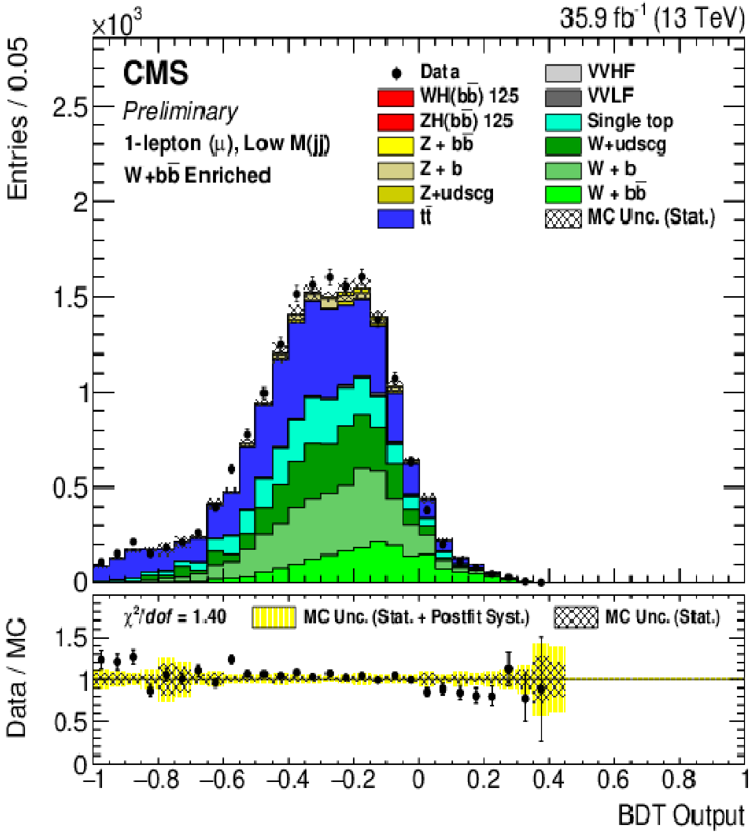

ATLAS
√s = 13 TeV , 36.1 fb-1
0 lepton, 2 jets, 2 b-tags
p_T^V ≥ 200 GeV
Data
VH → Vbb (μ=1.30)
Diboson
tt
Single top
W+(bb,bc,cc,bl)
Z+(bb,bc,cc,bl)
Uncertainty
Pre-fit background
SM VH → Vbb × 1
Events / 10 GeV
Data/Pred.
m_bb [GeV]

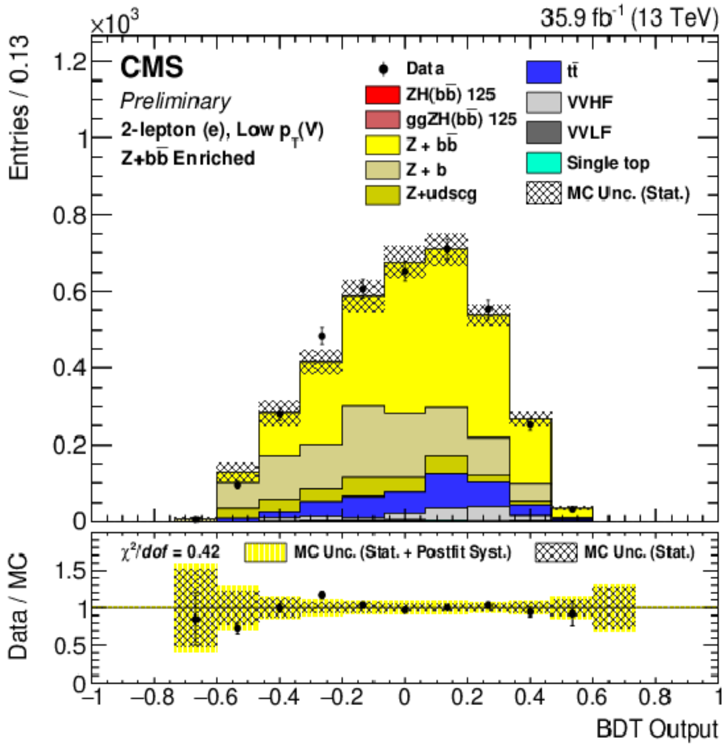

ATLAS
√s = 13 TeV , 36.1 fb-1
2 leptons, 2 jets, 2 b-tags
p_T^V ≥ 200 GeV
Data
VH → Vbb (μ=1.30)
Diboson
Z+(bb,bc,cc,bl)
Uncertainty
Pre-fit background
SM VH → Vbb × 1
Events / 10 GeV
Data/Pred.
m_bb [GeV]


ATLAS
√s = 13 TeV , 36.1 fb-1
1 lepton, 2 jets, 2 b-tags
p_T^V ≥ 200 GeV
Data
VH → Vbb (μ=1.30)
Diboson
tt
Single top
Multijet
W+(bb,bc,cc,bl)
W+cl
Z+(bb,bc,cc,bl)
Uncertainty
Pre-fit background
SM VH → Vbb × 1
Events / 10 GeV
Data/Pred.
m_bb [GeV]

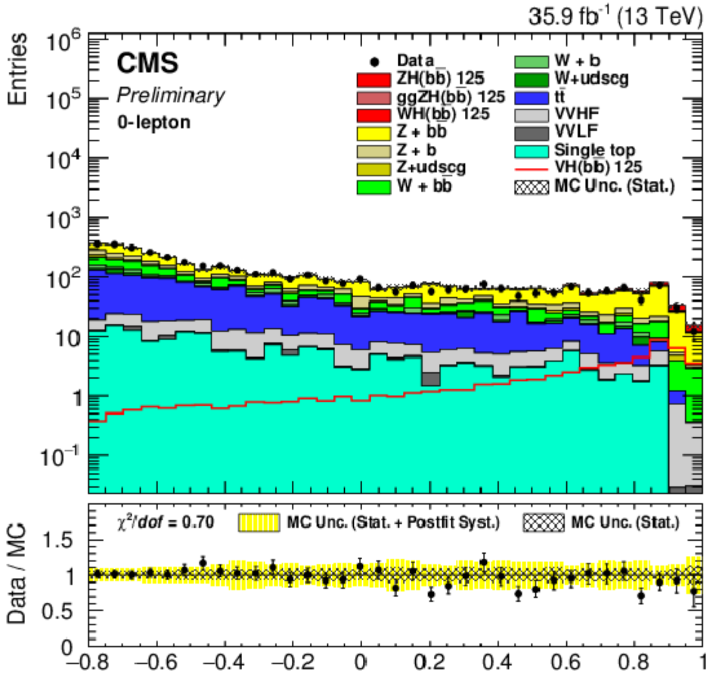

ATLAS
√s = 13 TeV, 36.1 fb-1
0+1+2 leptons
2+3 jets, 2 b-tags
Weighted by S/B
Dijet mass analysis
Data
VH → Vbb (μ=1.30)
Diboson
Uncertainty
Events / 10 GeV (Weighted, backgr. sub.)
m_bb [GeV]

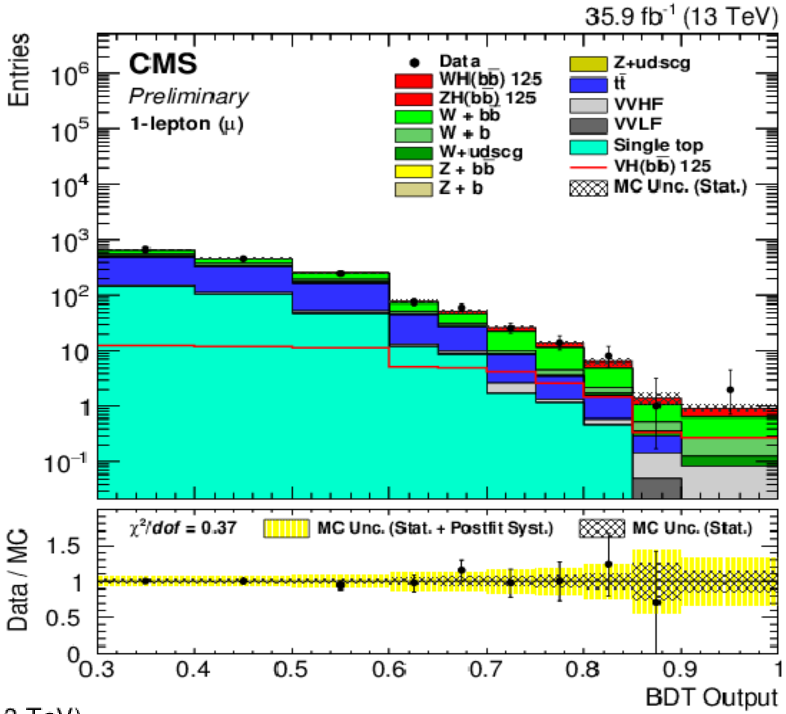
BDT in control regions

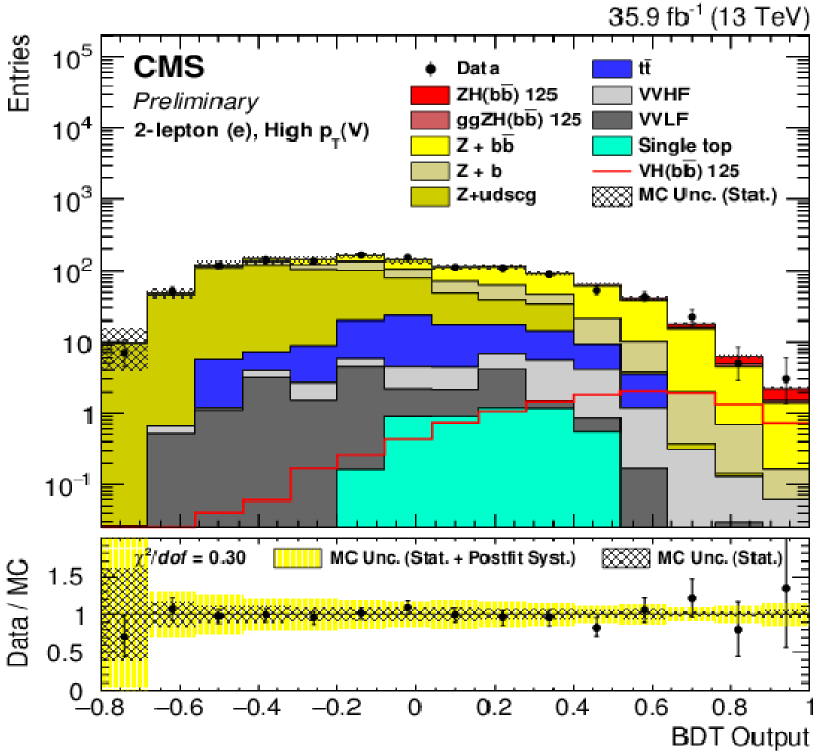


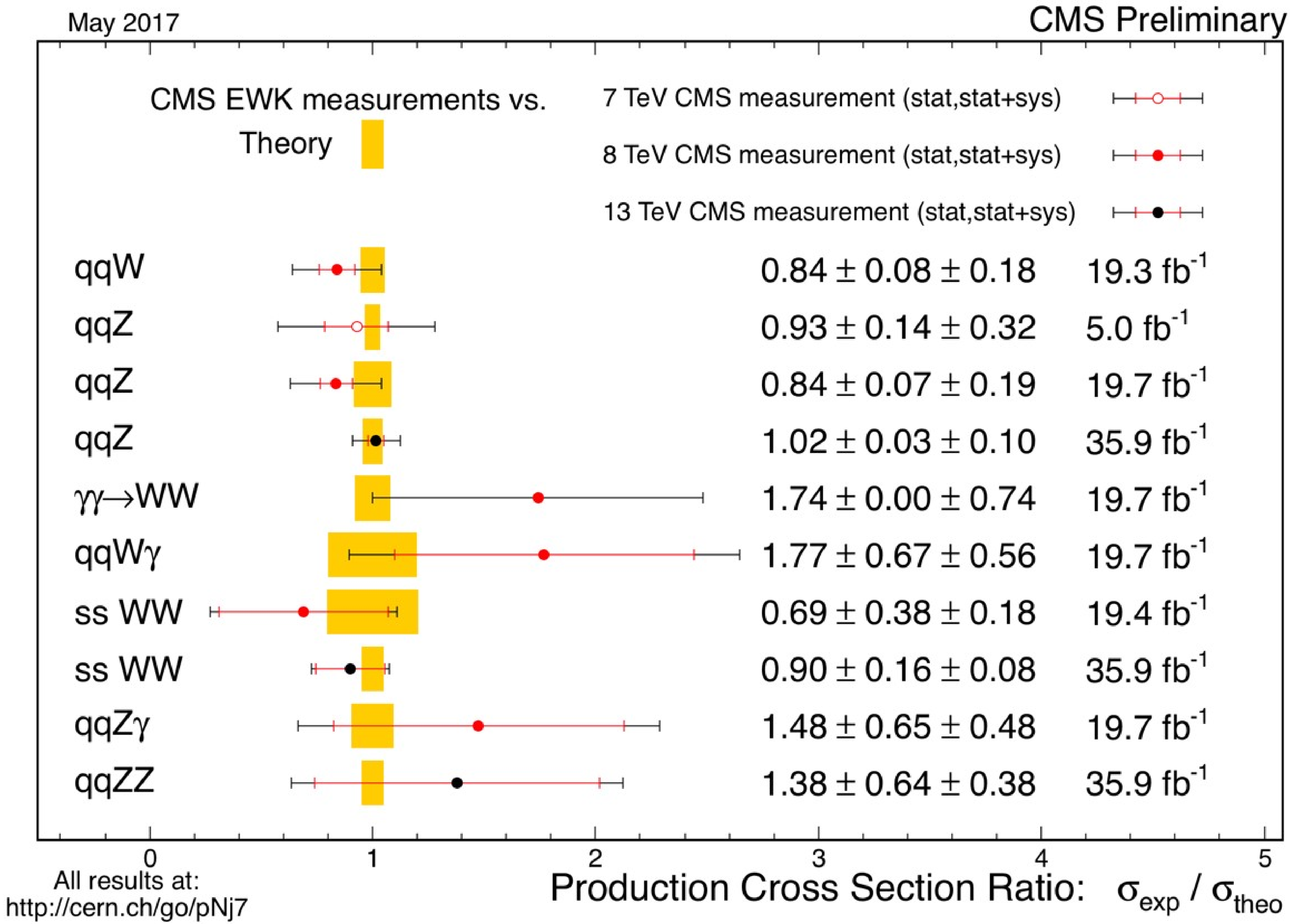


Hbb BDT

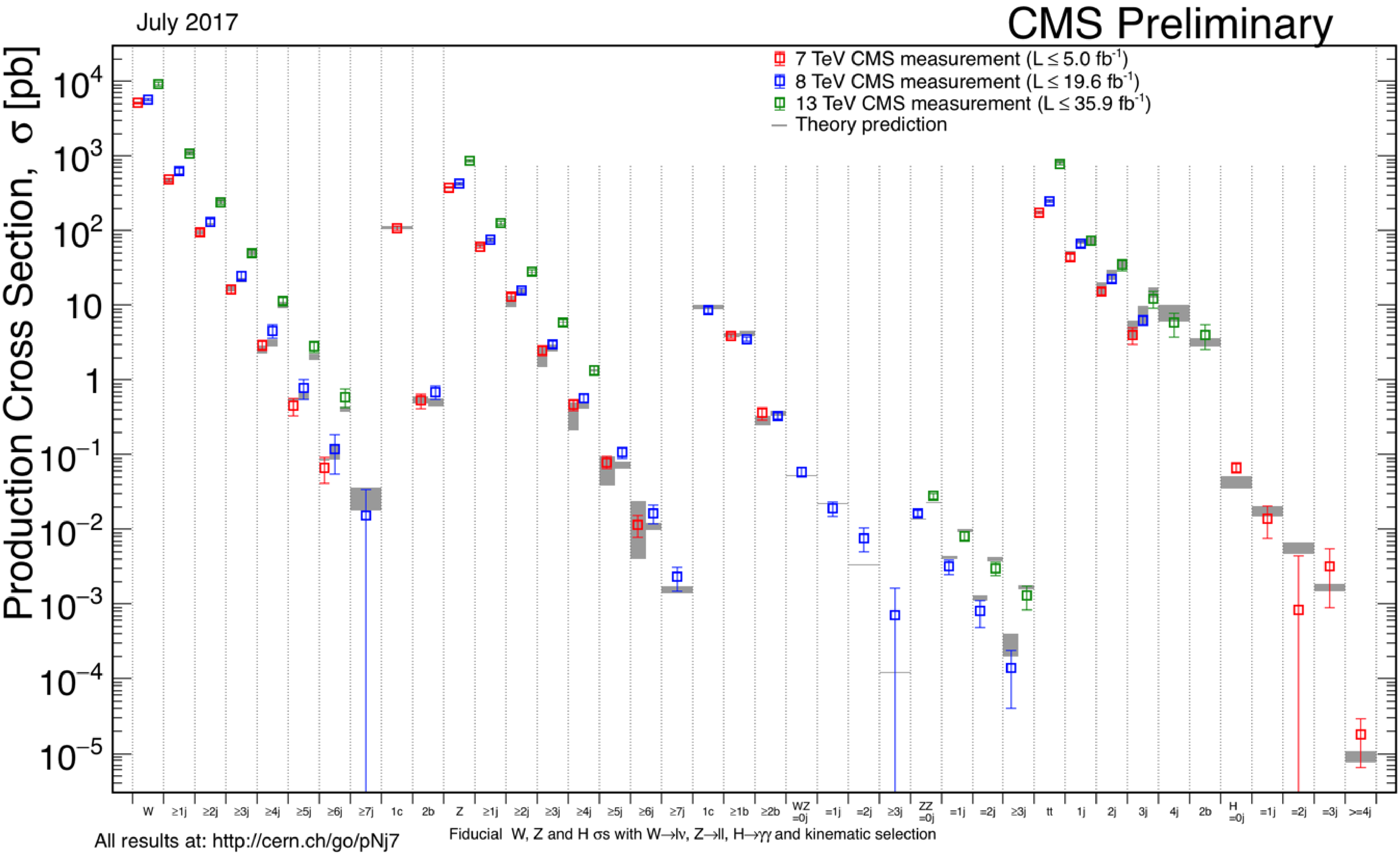

35.9 fb-1 (13 TeV)
CMS
Preliminary
0-lepton
Data
ZH(bb̄) 125
ggZH(bb̄) 125
WH(bb̄) 125
Z + bb̄
Z + b
Z+udscg
W + bb̄
W + b
W+udscg
tt̄
VVHF
VVLF
Single top
VH(bb̄) 125
MC Unc. (Stat.)
Entries
χ²/dof = 0.70
MC Unc. (Stat. + Postfit Syst.)
Data / MC


35.9 fb-1 (13 TeV)
CMS
Preliminary
1-lepton (μ)
Data
WH(bb̄) 125
ZH(bb̄) 125
W + bb̄
W + b
W+udscg
Z + bb̄
Z + b
Z+udscg
tt̄
VVHF
VVLF
Single top
VH(bb̄) 125
MC Unc. (Stat.)
Entries
χ²/dof = 0.37
MC Unc. (Stat. + Postfit Syst.)
Data / MC
BDT Output

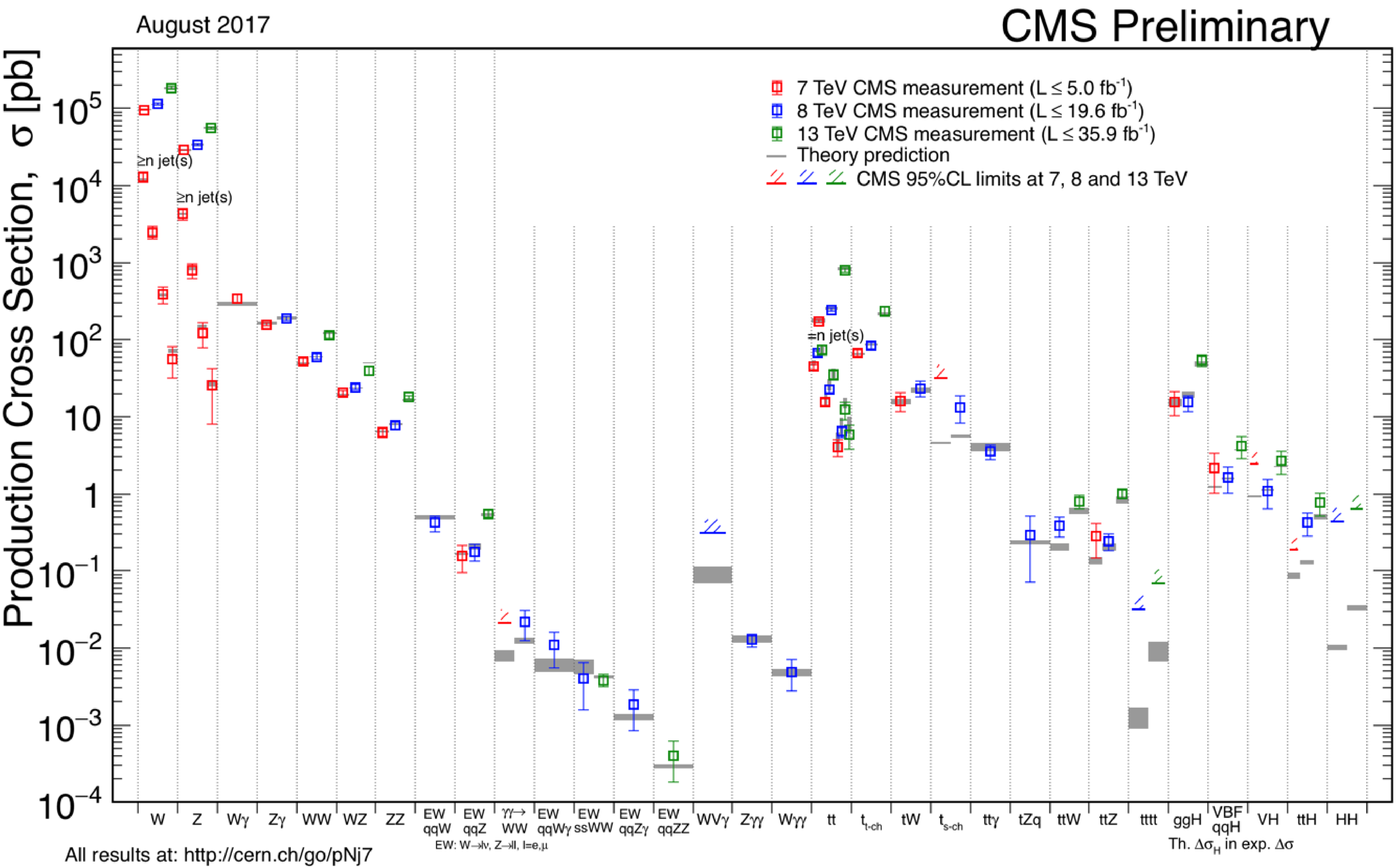

35.9 fb-1 (13 TeV)
CMS
Preliminary
2-lepton (e), High pT(V)
Data
ZH(bb̄) 125
ggZH(bb̄) 125
Z + bb̄
Z + b
Z+udscg
tt̄
VVHF
VVLF
Single top
VH(bb̄) 125
MC Unc. (Stat.)
Entries
χ²/dof = 0.30
MC Unc. (Stat. + Postfit Syst.)
Data / MC
BDT Output



May 2017
CMS Preliminary
CMS EWK measurements vs.
Theory
7 TeV CMS measurement (stat,stat+sys)
8 TeV CMS measurement (stat,stat+sys)
13 TeV CMS measurement (stat,stat+sys)
qqW
qqZ
qqZ
qqZ
γγ→WW
qqWγ
ss WW
ss WW
qqZγ
qqZZ
0.84 ± 0.08 ± 0.18
0.93 ± 0.14 ± 0.32
0.84 ± 0.07 ± 0.19
1.02 ± 0.03 ± 0.10
1.74 ± 0.00 ± 0.74
1.77 ± 0.67 ± 0.56
0.69 ± 0.38 ± 0.18
0.90 ± 0.16 ± 0.08
1.48 ± 0.65 ± 0.48
1.38 ± 0.64 ± 0.38
19.3 fb-1
5.0 fb-1
19.7 fb-1
35.9 fb-1
19.7 fb-1
19.7 fb-1
19.4 fb-1
35.9 fb-1
19.7 fb-1
35.9 fb-1
0
1
2
3
4
5
All results at:
http://cern.ch/go/pNj7
Production Cross Section Ratio: σexp / σtheo

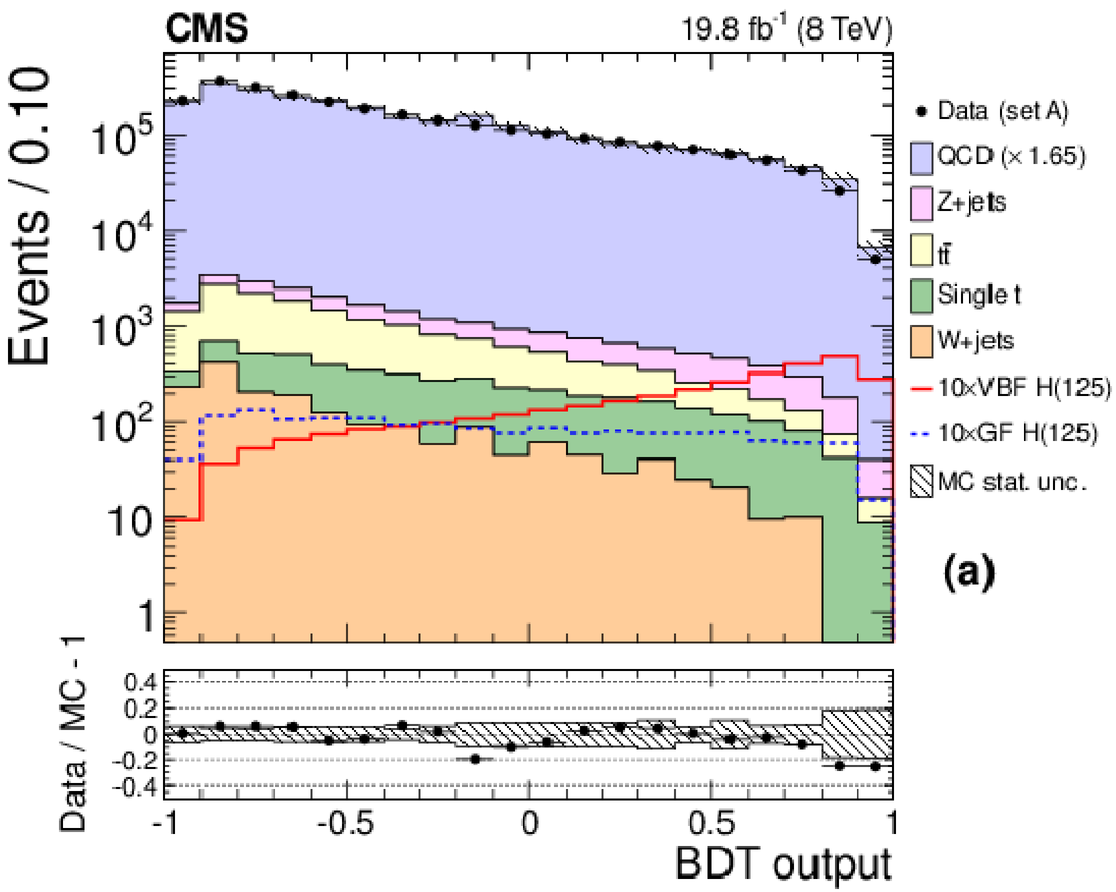


July 2017
CMS Preliminary
Production Cross Section, σ [pb]
7 TeV CMS measurement (L ≤ 5.0 fb⁻¹)
8 TeV CMS measurement (L ≤ 19.6 fb⁻¹)
13 TeV CMS measurement (L ≤ 35.9 fb⁻¹)
Theory prediction
W ≥1j ≥2j ≥3j ≥4j ≥5j ≥6j ≥7j 1c 2b Z ≥1j ≥2j ≥3j ≥4j ≥5j ≥6j ≥7j 1c ≥1b ≥2b WZ =0j =1j =2j ≥3j ZZ =0j =1j =2j ≥3j tt 1j 2j 3j 4j 2b H =0j =1j =2j =3j >=4j
Fiducial W, Z and H σs with W→lν, Z→ll, H→γγ and kinematic selection
All results at: http://cern.ch/go/pNj7

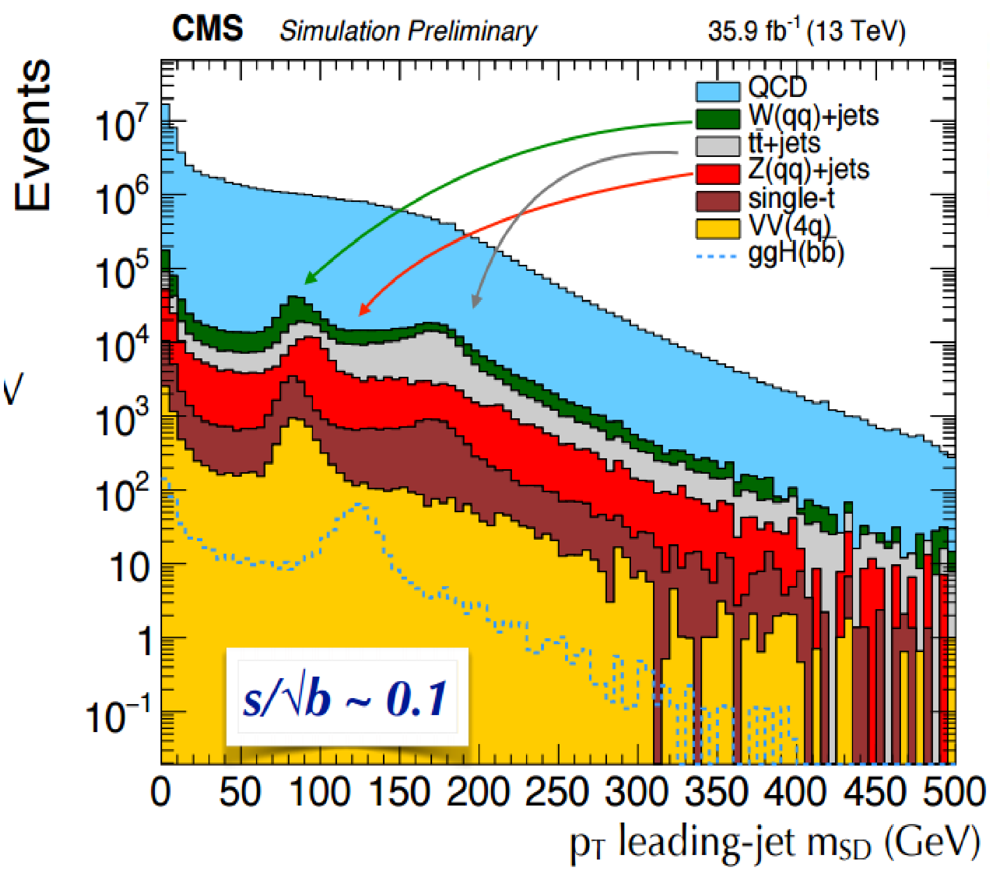


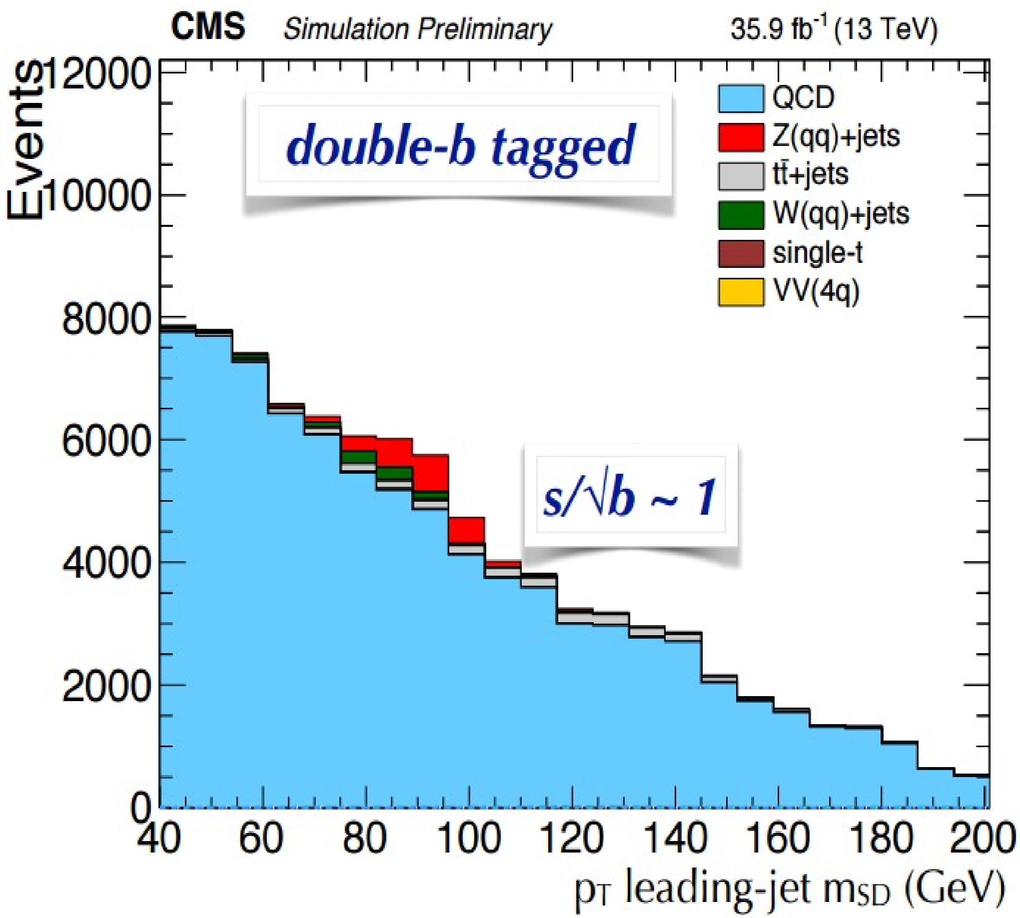
August 2017
CMS Preliminary
Production Cross Section, σ [pb]
7 TeV CMS measurement (L ≤ 5.0 fb-1)
8 TeV CMS measurement (L ≤ 19.6 fb-1)
13 TeV CMS measurement (L ≤ 35.9 fb-1)
Theory prediction
CMS 95%CL limits at 7, 8 and 13 TeV
≥n jet(s)
≥n jet(s)
=n jet(s)
W
Z
Wγ
Zγ
WW
WZ
ZZ
EW qqW
EW qqZ
γγ→WW
EW qqWγ
EW ssWW
EW qqZγ
EW qqZZ
WVγ
Zγγ
Wγγ
tt
t-ch
tW
ts-ch
ttγ
tZq
ttW
ttZ
tttt
ggH
VBF qqH
VH
ttH
HH
EW: W→lν, Z→ll, l=e,μ
Th. ΔσH in exp. Δσ
All results at: http://cern.ch/go/pNj7

VBF Hbb BDT


CMS
19.8 fb^{-1} (8 TeV)
Events / 0.10
Data / MC - 1
BDT output
Data (set A)
QCD (× 1.65)
Z+jets
tt̄
Single t
W+jets
10×VBF H(125)
10×GF H(125)
MC stat. unc.
(a)

Boosted Hbb


CMS Simulation Preliminary
35.9 fb^{-1} (13 TeV)
Events
QCD
W(qq)+jets
$t\bar{t}$+jets
Z(qq)+jets
single-t
VV(4q)
ggH(bb)
$s/\sqrt{b}$ ~ 0.1
p_T leading-jet m_{SD} (GeV)

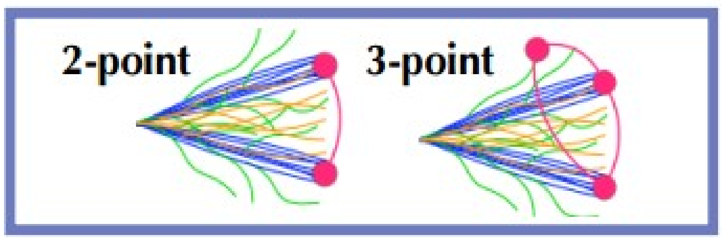

CMS Simulation Preliminary
35.9 fb^{-1} (13 TeV)
Events
double-b tagged
QCD
Z(qq)+jets
$t\bar{t}$+jets
W(qq)+jets
single-t
VV(4q)
$s/\sqrt{b}$ ~ 1
p_T leading-jet m_{SD} (GeV)

Jet Substructure

ArXiv:1609.07483

- Measures the degree to which a jet can be considered as composed of N prongs
- **Energy correlation functions** are sensitive to N-point correlations in a jet
  - A 2-pronged jet will have $e_3 < e_2$

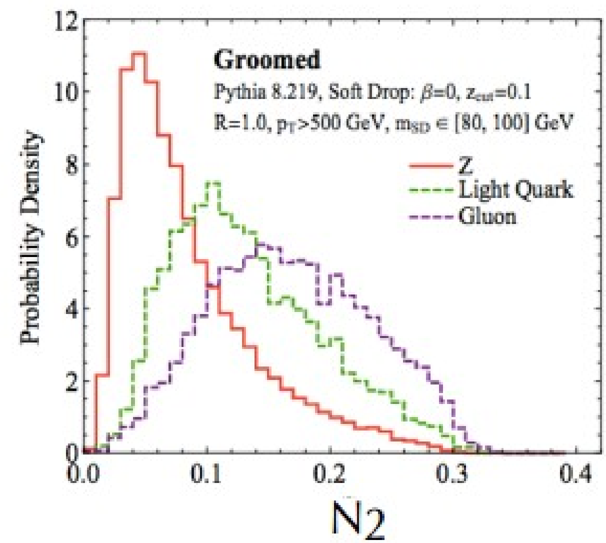
$$N_2^{\beta} = \frac{{}_2e_3^{\beta}}{({}_1e_2^{\beta})^2}$$



$${}_1e_2^{\beta} = \sum_{1 \leq i<j \leq n_j} z_i z_j \Delta R_{ij}^{\beta}$$

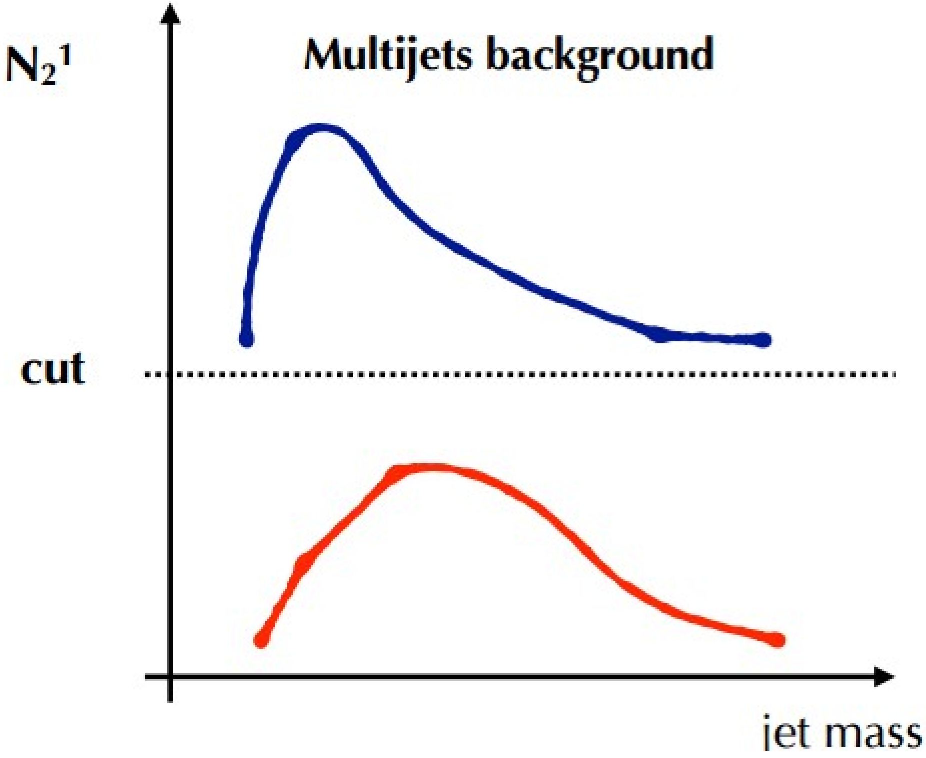
$${}_2e_3^{\beta} = \sum_{1 \leq i<j<k \leq n_j} z_i z_j z_k \min\left\{\Delta R_{ij}^{\beta}\Delta R_{ik}^{\beta}, \Delta R_{ij}^{\beta}\Delta R_{jk}^{\beta}, \Delta R_{ik}^{\beta}\Delta R_{jk}^{\beta}\right\}$$

β=1



Jet Substructure

ArXiv:1603.00027
CMS-PAS-B2G-17-001
CMS-PAS-EXO-17-001



$N^1{}_2$ sculpts jet mass distribution