VBF Higgs to Invisible - Towards Run II

HIG-14-038, AN-14-243

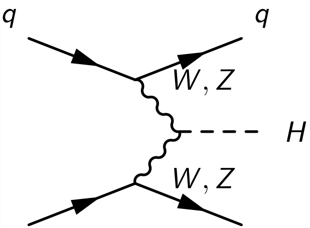
C. Asawatangtrakuldee, J. Brooke, D. Colling, G. Davies,
P. Dunne, A.M. Magnan, A. Nikitenko, J. Pela

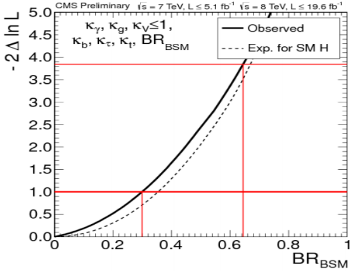
Reminder

- Motivation is the same as run 1:
  - Complemented by direct DM and indirect B(H→ inv) searches
- For Run 1 we had two sets of triggers
  - Prompt trigger used for published result: HIG-13-30
  - Parked triggers: analysis presented today
  - Parked analysis will be used as baseline for run 2

Overview

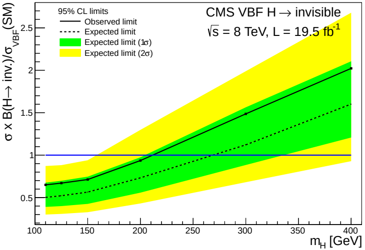
- Reminder of prompt analysis
- Details of parked analysis
  - Emphasis on changes from established prompt analysis
- Trigger plans for run 2





Prompt Analysis

- Standard object definition:
  - Details in backup
- Single bin counting experiment
  - Signal region chosen to eliminate QCD and be above trigger turn ons
- Major backgrounds use data driven estimates:
  - $Z \to \nu\nu, W \to \ell\nu$, QCD
- Minor backgrounds taken from MC:
  - VV, Wγ, $t\bar{t}$, single top
- Expected limit 49% at $m_H = 125$GeV



Data driven background estimation

W: $N_S = N_S^{MC} \frac{N_C^{Data} - N_C^{Bkg}}{N_C^{MC}}$

Z: $N_S^{Z\to\nu\nu} = \left(N_C^{Data} - N_C^{bkg}\right) \cdot \frac{\sigma(Z\to\nu\nu)}{\sigma(Z/\gamma^*\to\mu\mu)} \cdot \frac{\epsilon_S^{ZMC}}{\epsilon_C^{ZMC}}$

Parked Analysis

- All parked and prompt triggers are seeded by L1_ETM40
- Parked triggers have looser HLT thresholds:
  - this allows us to look at new phase space regions and analysis techniques
  - also increases QCD and PU backgrounds

Trigger Efficiency

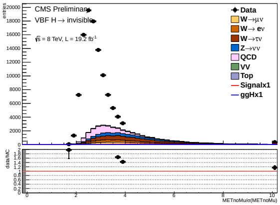
- Trigger efficiency has been measured including correlation between variables
  - This allows the trigger turn on region to be used
  - Control triggers will be required to do this for run 2

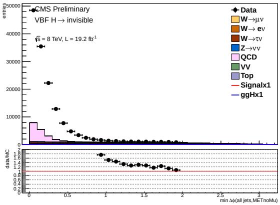
Signal region

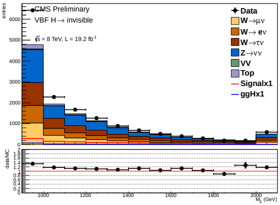
- QCD hard to model, signal region cuts are chosen to make remaining QCD small whilst enhancing real-MET using new Min$\Delta\phi$(jet,MET) variable
- The signal region has been reoptimised for the looser parked triggers
  - New region uses new variable has higher signal efficiency with much less QCD

Starting point for region choice

- Trigger turn ons and detector acceptance impose the following cuts:
  - $\eta_{j1} \cdot \eta_{j2} < 0$, $|\eta_{j1,2}| < 4.7$, jet 1 $p_T > 50$ GeV, $\Delta\eta_{jj} > 3.6$, jet 2 $p_T > 40$ GeV, METnomu > 90 GeV, $M_{jj} > 800$ GeV
- QCD in plots is VBF enriched MC doesn't model all QCD
- Following cuts added due to poor data-MC agreement from QCD contamination:
- $\frac{METnomu}{\sigma_{METnomu}} > 3.0$, Min$\Delta\phi$($all\ jets\ p_T > 30\ GeV$, $METnomu$) > 1.0, $M_{jj} > 1000$ GeV

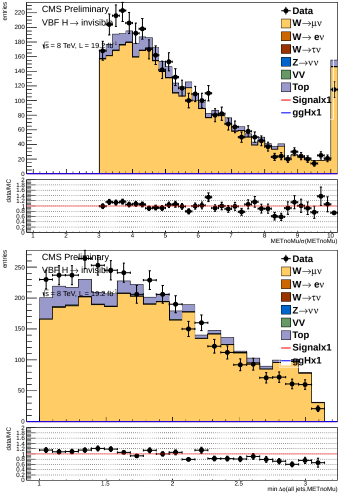






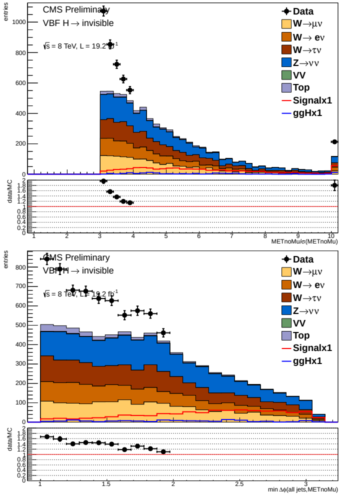
Signal region selection

- ▶ As in the prompt analysis we veto events with 'veto' electrons or muons
- ▶ Can't model QCD shape so cut hard to remove most QCD
  - Can then tolerate a larger uncertainty on QCD estimation
- ▶ Select region $\frac{METnoMU}{\sigma_{METnoMU}} > 4$ and $\mathrm{Min}\Delta\phi(all\ jets, METnomu) > 2.0$
- ▶ Signal contribution also large in this region of parameter space
  - We blind this region and use as a basis for signal region optimisation



Signal region selection

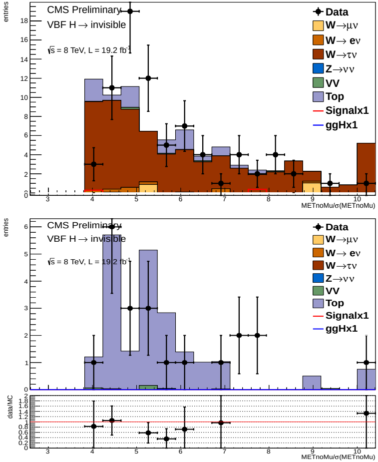
- We optimise by choosing the cut values with the best 95% C.L. expected limit
  - Limit calculation details later
- We scanned through jet 2 p_T, $\frac{METnoMU}{\sigma_{METnoMU}}$, $\mathrm{Min}\Delta\phi(all\ jets\, , METnomu)$ and M_{jj}
- Best limit was found for:
  - jet 2 $p_T > 45$ GeV
  - $\frac{METnoMU}{\sigma_{METnoMU}} > 4$
  - $\mathrm{Min}\Delta\phi(all\ jets, METnomu) > 2.3$
  - $M_{jj} > 1200$ GeV
- We defined this as our "signal region"
- Discrepancy outside signal region is from QCD



Top control region

- Top contribution to V+jets control regions is non-negligible
- Use method used for W backgrounds in prompt analysis
- Region: signal region with lepton veto replaced with requirement for 1 tight muon and 1 tight electron
  - Very few events in $e\mu$ region so also removed Min$\Delta\phi(all\ jets, METnomu)$ cut

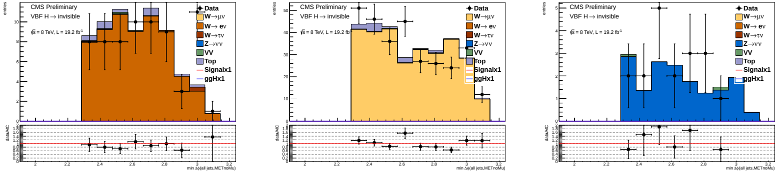
N_C^{data}	$21 \pm 4.6(\text{stat.})$
N_C^{bkg}	$0.3 \pm 0.1(\text{MC stat.})$
$N_S^{top\ MC}$	$5.3 \pm 1.3(\text{MC stat.})$
$N_C^{top\ MC}$	$24.6 \pm 4.0(\text{MC stat.})$
$\frac{N^{data}-N^{bkg}}{N_C^{top\ MC}}$	$0.8 \pm 0.2(stat.) \pm 0.1(MCstat.)$
N_S^{top}	$4.4 \pm 1.0(\text{stat.}) \pm 1.3(\text{MC stat.})$



$W \to e\nu$, $W \to \mu\nu$, $Z \to \nu\nu$

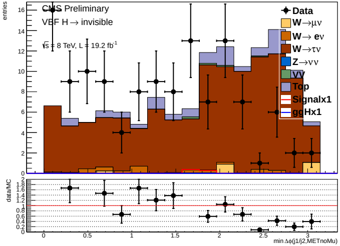
- Data-MC agreement good
- Same method used as for prompt analysis
- Details in backup

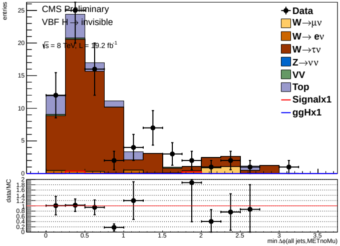
Background	Number of events
$W \to e\nu$	$57.4 \pm 7.3(\text{stat.}) \pm 5.9(\text{MC stat.})$
$W \to \mu\nu$	$101.8 \pm 6.1(\text{stat.}) \pm 8.3(\text{MC stat.})$
$Z \to \nu\nu$	$157.3 \pm 37.6(\mathit{datastat.}) \pm 18.2(\mathit{MCstat.})$



$W \to \tau\nu$ control region

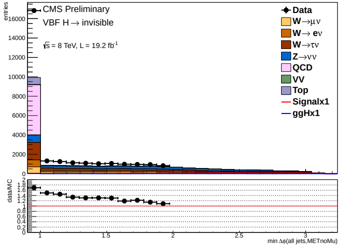
- Control regions for other W backgrounds:
  - signal region with lepton veto replaced with a requirement for a single lepton
- For $W \to \tau\nu$ there are not enough events in this region:
- Remove the $\text{Min}\Delta\phi(\textit{all jets}, \textit{METnomu})$ cut
- This leads to QCD contamination so we require:
  - $\text{Min}\Delta\phi(\textit{leading 2 jets}, \textit{METnomu}) > 1.0$
  - m_T of the lepton-MET system > 20 GeV
- We add a 20% systematic on the $W \to \tau\nu$ background
- Final $W \to \tau\nu$ estimate: $98.0 \pm 13.2(\textit{stat.}) \pm 12.6(\textit{MCstat.})$

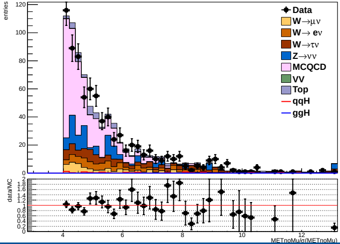




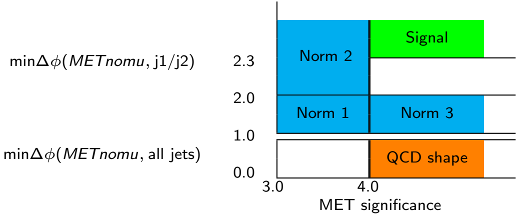
QCD background estimation: Shape

- Other background methods estimate unreconstructed contribution from reconstructed
- For QCD use MET near reconstructed jets to model MET from unreconstructed/mismeasured jets
- Use region with low $\mathrm{Min}\Delta\phi(\textit{all jets}, \textit{METnomu})$ but high $\mathrm{Min}\Delta\phi(\textit{leading jets}, \textit{METnomu})$
  - $\mathrm{Min}\Delta\phi(\textit{all jets}, \textit{METnomu}) < 1.0$
  - $\mathrm{Min}\Delta\phi(\textit{leading jets}, \textit{METnomu}) > 1.0$
- Has good shape agreement with enriched QCD MC
- Use shape from this region





QCD background estimation: Normalisation



- MET significance and min$\Delta\phi(METnomu, \text{all jets})$ correlated
  - Cannot use ABCD to normalise
- Normalisation shows strong dependence on cut variables
- Fit normalisation variation in norm 1
- Check consistency in norm 2 and 3
- Final prediction: $N_S^{QCD} = 17 \pm 14$

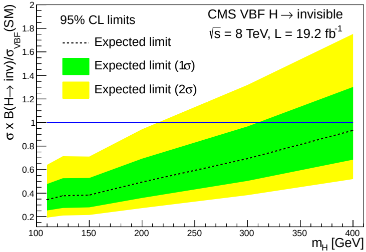
Results

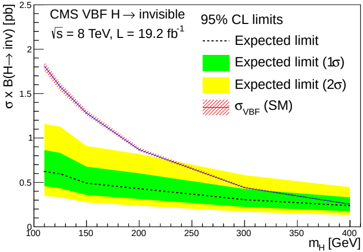
Background	$N_{est} \pm (stat) \pm (syst)$
$Z \to \nu\nu$	$157.3 \pm 37.6 \pm 38.3$
$W \to \mu\nu$	$101.8 \pm 6.1 \pm 11.9$
$W \to e\nu$	$57.4 \pm 7.3 \pm 7.0$
$W \to \tau\nu$	$98.0 \pm 13.2 \pm 25.4$
top	$4.4 \pm 1.0 \pm 1.4$
VV	$3.8 \pm 0.0 \pm 0.7$
QCD multijet	$17 \pm 0 \pm 14$
Total Background	$439.7 \pm 41.0 \pm 55.8$
Signal(VBF) 100% BF	$273.4 \pm 0.0 \pm 31.2$
Signal(ggH) 100% BF	$22.6 \pm 0.0 \pm 15.6$

- Uncertainties considered are mostly the same as the prompt analysis
  - Details in backup
- The size of each uncertainty is given as a percentage of total signal/background
- Data and MC statistics still largest single contribution

Expected limits

- Used Higgs combine package with Asymptotic CLs method
- Remaining backgrounds very signal like
  - Performed a single bin counting experiment
- Analysis blind so have expected limits only
- 95% C.L. Median limit on B(H→inv.) for $m_H = 125$ GeV is: 38%
- Prompt paper expected limit was 49%





Run 2 Trigger

- Current studies performed for baseline 2015 menu and to be used in MC production
- Efficiencies are for VBF Higgs to invisible process with $m_H = 125$ GeV
- Rates are calculated using minimum bias samples for L1 and full QCD for HLT
- Conditions used were 25ns bunch spacing and PU40

L1

- Start from baseline L1_ETM70 which is planned to be in trigger menu
  - find 7 kHz individual rate, 28% efficiency
  - run 1 ETM40 was 3kHz
- Systematic scan through other variables also finds:

1) Dijet30 + fwd/bck + $\Delta\eta(jj) > 3.5$ + Jet96
   - 4.6 kHz individual rate, 21% increase in efficiency over ETM70
2) Dijet30 + fwd/bck + $\Delta\eta(jj) > 3.5$ + ETM50
   - 5.0 kHz individual rate, 14% increase in efficiency over ETM70

Run 2 Trigger

HLT

- Use L1_ETM70 as seed
- Baseline HLT path in menu is HLT_PFMET170_NoiseCleaned with 9% efficiency
- Scan over variables also finds:
  - HLT_DiPFJetVBF40_DEta3p5_MJJ600_PFMETNoMu140
  - Additional rate 4.7 Hz total efficiency 10.5%
- We also proposed a prescaled control trigger for efficiency measurements:
  - HLT_DiPFJetVBF40_DEta3p5_MJJ600_PFMETNoMu80 seeded by L1_ETM50 with rate 0.5 Hz

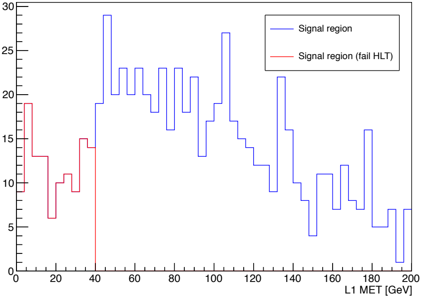
- For comparison run 1 prompt trigger had 8% efficiency at HLT for 1 Hz

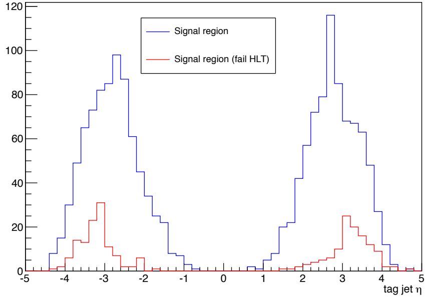
L1 MET Issues

- 2012 analysis showed significant loss of efficiency in trigger
  - MC signal yield after all cuts = **209 events**
  - MC signal yield after all cuts, *but no trigger* = **247 events**

Trigger loses 15% of potential signal

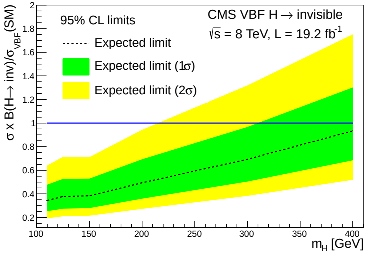
- Plots below show :
  - Events in signal region, no trigger & Events in signal region, failing trigger
  - Loss of efficiency entirely due to L1 MET cut
  - Tag jet η distribution shows majority of failing events have a jet in HF
  - **L1 MET calculation is configured to use only |η| < 3**

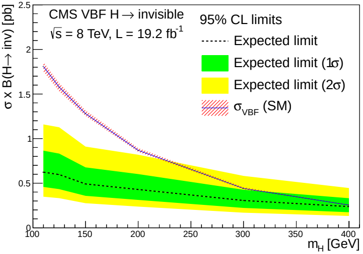




Summary

- New cut based analysis presented with less QCD and higher signal efficiency
  - Made possible by looser parked trigger thresholds
  - Expected limit 38% improved from 49% for prompt analysis
- This analysis will be used as our baseline for run 2 studies
- Run 2 trigger plans shown
  - We can match/improve on run 1 efficiency
  - Offline studies to follow





Backup

Documentation

AN-14-243

Available on the CMS information server **CMS NOTE AN-14-243**

The Compact Muon Solenoid Experiment

CMS Draft Note

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland

2014/12/01
Head Id: 269377
Archive Id: 266572:269470M
Archive Date: 2014/11/27
Archive Tag: trunk

Search for a Higgs boson decaying to invisible final states

Chayanit Asawatangtrakuldee[2], Jim Brooke[3], David Colling[1], Gavin Davies[1], Patrick Dunne[1], Anne-Marie Magnan[1], Alexander Nikitenko[1], and Joao Pela[1]

[1] Imperial College London (UK)
[2] Peking University, Beijing (China)
[3] University of Bristol (UK)

Abstract

In this note, investigations are made into improving the analysis for the search of a Higgs boson produced by Vector-Boson Fusion and decaying to invisible particles, compared to what was published in HIG-13-030 with the 8 TeV dataset, and in view of preparing for 13 TeV. The parked triggers are used instead of the prompt ones, which allow a small increase in statistics due to different requirements at HLT, in particular no requirement on the MET and looser thresholds on the jets pT. An improved cut-based selection is presented, with better rejection of the QCD multijet background. A BDT-based selection is also investigated. Both approaches are optimised in terms of expected 95%CL limits on the branching ratio of Higgs to invisible.

This box is only visible in draft mode. Please make sure the values below make sense.

PDFAuthor:	P. Dunne, A.-M. Magnan
PDFTitle:	Search for a Higgs boson decaying to invisible final states
PDFSubject:	CMS
PDFKeywords:	CMS, physics, Higgs boson, invisible

Please also verify that the abstract does not use any user defined symbols

HIG-14-038

CMS PAS HIG-14-038

DRAFT
CMS Physics Analysis Summary

The content of this note is intended for CMS internal use and distribution only

2014/12/02
Head Id: 269969
Archive Id: 269969P
Archive Date: 2014/12/02
Archive Tag: trunk

Search for invisible decays of Higgs bosons in the vector boson fusion production mode

The CMS Collaboration

Abstract

A search for invisible decays of Higgs bosons in the vector boson fusion (VBF) production mode is carried out using data recorded in 2012 at a centre-of-mass energy of 8 TeV by the CMS detector corresponding to an integrated luminosity of 19.2 inverse femtobarns. Limits are set on the production cross section times invisible branching fraction, as a function of the Higgs boson mass. Assuming standard model Higgs boson cross sections and acceptances, the observed (expected) upper limit on the invisible branching fraction at $m_H = 125$ GeV is found to be 0.XX (0.38) at 95% confidence level.

This box is only visible in draft mode. Please make sure the values below make sense.

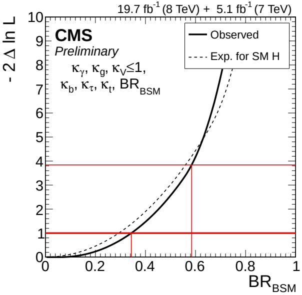
PDFAuthor:	CMS Collaboration
PDFTitle:	Search for invisible decays of Higgs bosons in the vector boson fusion production mode
PDFSubject:	CMS
PDFKeywords:	CMS, physics, Higgs

Please also verify that the abstract does not use any user defined symbols

Why Higgs to Invisible?

Experimental motivation

- Current measurements of the 125 GeV Higgs boson are compatible with Standard Model (SM) expectations
  - large uncertainties can still accommodate significant beyond the SM (BSM) properties
- Additional Higgs bosons with exotic decays are not excluded



Theoretical motivation

- Many BSM theories predict Higgs boson decays to invisible final states:
  - e.g. SUSY, extra dimensions, fourth-generation neutrinos
- These final state particles are often dark matter candidates

Parked Triggers

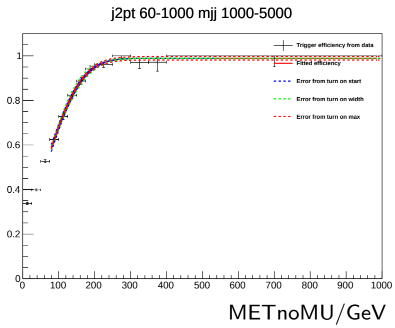
- Use already analysed prompt trigger for run A
- One parked trigger for runs B and C, another for run D
  - Parked trigger cuts are looser so prompt trigger not used where parked trigger is available
- All parked and prompt triggers are seeded by L1_ETM40
- Parked triggers have looser HLT thresholds
- This allows us to look at new regions of phase space and different analysis techniques
- Use prompt analysis as a base:
  - e.g. same objects and MC samples

HLT

Run period	MET cut	dijet p_T cut	dijet mass cut
A	METnoMuons>65 GeV	DiPFJet40	MJJ800
B&C	N/A	DiJet35	MJJ700
D	N/A	DiJet30	MJJ700

Trigger efficiency

- Variables used in prompt and parked triggers are highly correlated:
  - dijet mass, METnoMU, jet 2 p_T
- In the prompt analysis correlations were neglected as we cut to ensure trigger was $> 95\%$ efficient
- For the parked analysis we use a 2D binning in dijet mass and jet 2 p_T
  - MJJ: 0,600,800,900,1000,5000
  - Jet 2 p_T: 30,40,50,60,1000
- In each bin we fit the METnoMU trigger turn on using an error function
- We then combine the turn ons from runs A, BC and D weighted by luminosity and apply this to MC events



Prompt vs Parked selection

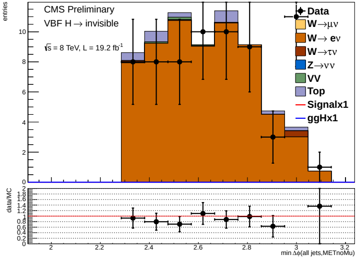
- Summary of differences in signal region selection

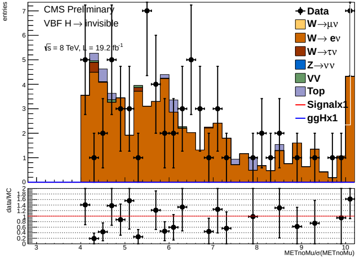
Variable	Prompt cut	Parked cut
Lepton veto	no veto e or μ	
$\eta_{j1,2}$	< 4.7	
$\eta_{j1} \cdot \eta_{j2}$	< 0	
jet 1 p_T	> 50 GeV	
jet 2 p_T	> 50 GeV	> 45 GeV
$\Delta\eta_{jj}$	> 4.2	> 3.6
M_{jj}	> 1100 GeV	> 1200 GeV
METnomu	> 130 GeV	> 90 GeV
Central jet veto	yes	no
$\Delta\phi_{jj}$	< 1.0	no cut
$\frac{METnoMU}{\sigma_{METnoMU}}$	no cut	> 4
Min$\Delta\phi(all\ jets, METnomu)$	no cut	> 2.3

$W \to e\nu$

- Data-MC agreement good
- Same method used as for prompt analysis

	Signal region	Control region
N^{data}	XXX	68 ± 8.2(stat.)
N^{bkg}	N/A	3.1 ± 1.5(stat.)
N^{WMC}	114.6 ± 8.9(stat.)	129.6 ± 8.1(stat.)
$(N^{data} - N^{bkg})/N_C^{WMC}$	0.50 ± 0.06(stat.)± 0.03(MC stat.)	
Final estimate	57.4 ± 7.3(stat.)± 5.9(MC stat.)	N/A

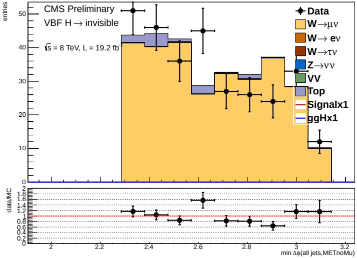


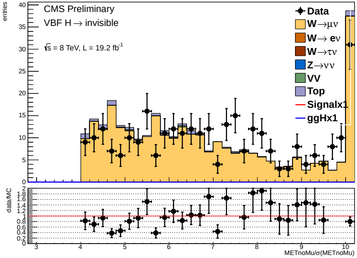


$W \to \mu\nu$

- Data-MC agreement good
- Same method used as for prompt analysis

	Signal region	Control region
N^{data}	XXX	300 ± 17.3(stat.)
N^{bkg}	N/A	12.7 ± 4.6(stat.)
N^{WMC}	142.1 ± 10.1(stat.)	401.1 ± 15.1(stat.)
$(N^{data} - N^{bkg})/N_C^{MC}$	0.72 ± 0.04(stat.)± 0.03(MC stat.)	
Final estimate	101.8 ± 6.1(stat.)± 8.3(MC stat.)	N/A

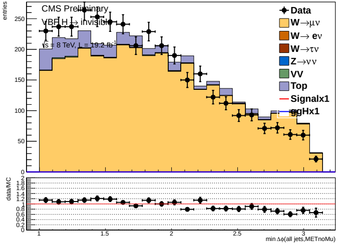


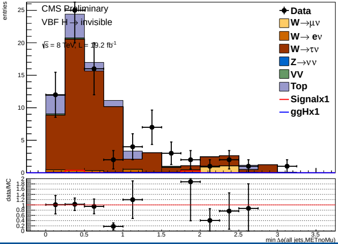


$W \to \tau\nu$ control region

- Difference in weight between control region and signal region cuts used to estimate error
- Measured in $W \to \mu\nu$ which has enough events to see data driven weight variation with $\mathrm{Min}\Delta\phi(\mathit{all\ jets}, \mathit{METnomu})$ cut
  - weight changes by 20% when loosening cut from 2.3 to 1.0.
  - We add a 20% systematic on the $W \to \tau\nu$ background

N_C^{data}	$76 \pm 8.7(stat.)$
N_C^{bkg}	$11.3 \pm 4.6(MCstat.)$
N_S^{MC}	$122.6 \pm 8.8(MCstat.)$
N_C^{MC}	$81.0 \pm 6.4(MCstat.)$
$\frac{N^{data} - N^{bkg}}{N_C^{MC}}$	$0.80 \pm 0.11(\mathrm{stat.}) \pm 0.08(\mathrm{MC\ stat.})$
N_S^W	$98.0 \pm 13.2(stat.) \pm 12.6(MCstat.)$

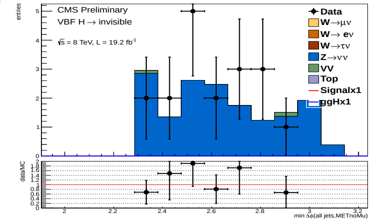


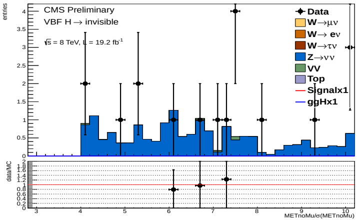


$Z \to \nu\nu$

- Data-MC agreement good for limited statistics
- Same method used as for prompt analysis

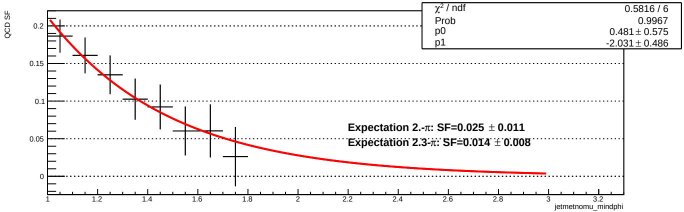
	Signal region	Control region
N^{data}	XXX	$18 \pm 4.2(stat.)$
N^{bkg}	N/A	$0.2 \pm 0.1(stat.)$
$N^{MC}(EWK)$	$7.9 \pm 0.2(stat.)$	$6.0 \pm 0.2(stat.)$
$N^{MC}(QCD)$	$29.5 \pm 3.0(stat.)$	$20.7 \pm 2.5(stat.)$
$\frac{N^{data}-N^{bkg}}{N^{MC}(EWK)+N^{MC}(QCD)}$	$0.67 \pm 0.16(stat.) \pm 0.06(MCstat.)$	
$FinalN^{Z\to\nu\nu}estimate$	$157.3 \pm 37.6(datastat.) \pm 18.2(MCstat.)$	N/A



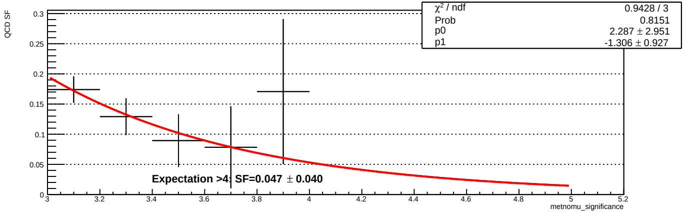


Normalisation variation

Jet met dphi



MET significance



QCD background estimation: Result and systematics

Region	Factor	Extrapolation mindphi>2.5	Extrapolation metsig>4
Norm 1	0.17 ± 0.02	0.014 ± 0.008	0.05 ± 0.04
Norm 1+2	0.12 ± 0.01	0.013 ± 0.004	0.01 ± 0.01
Norm 1+3	0.24 ± 0.03	0.03 ± 0.01	0.55 ± 0.06
Norm 2	0.06 ± 0.01	-	0.01 ± 0.02
Norm 3	0.5 ± 0.1	0.21 ± 0.11	-

- Good agreement in mindphi extrapolations
- Norm 3 agreement in metsig is poor
  - As norm 3 has low statistics and is an odd region: drop
- Use envelope of norm 1 scale factors
- Final prediction: $N_S^{QCD} = 17 \pm 14$

Systematics

- Changes are:
  - Top background now data driven so has data stat. error
  - New $W \to \tau\nu$ extrapolation error
  - QCD background error procedure now as described above
- We are reevaluating the $Z/\gamma^* \to \mu\mu$ to $Z \to \nu\nu$ with aMC@NLO
  - Waiting on MC jobs, very quick to update results

Systematics

- Uncertainties considered are mostly the same as the prompt analysis
  - Details in backup
- The size of each uncertainty is given as a percentage of total signal/background
  - Still statistically dominated

Source	Total background	Signal
Control region data stat.	9.30	0.00
$Z/\gamma^* \to \mu\mu$ to $Z \to \nu\nu$ extrapolation	7.16	0.00
MC stat.	5.54	3.82
Jet energy scale	4.94	10.70
$W \to \tau\nu$ control region extrapolation	4.46	0.00
Lepton ID efficiency	3.22	0.00
QCD normalisation	3.18	0.00
Jet energy resolution	2.86	1.81
Unclustered energy scale	2.28	1.64
Pileup weight	0.95	1.56
Luminosity	0.02	2.60
Theory Uncertainty	0.01	5.14

Source of gain

- Expected limits from prompt and parked analyses have been obtained from the prompt and parked data
- ggH signal and UES uncertainty not included
  - This explains 2% difference in parked cuts parked trigger number
- Data driven top control region used for both prompt and parked cuts
- Prompt trigger weights ignore correlations in turn on part of parked cut region

	Prompt trigger	parked trigger
Prompt cuts	45%	46%
Parked cuts	47%	40%

Interpretation

- Prompt cuts limits $\sim$ same as old card with both prompt and parked trigger
  - slight difference as paper limit had extra $W\gamma$
- The parked cuts give a worse limit with prompt trigger than with parked trigger
- Improvement comes from using the new phase space made available by the parked trigger

Uncertainty impact table - impacts larger than 0%

Expected limit with:	All Nuisances: 37.8%	No Nuisances: 14.3%
Nuisance	Removal effect (relative %)	Addition effect (relative %)
$Z \rightarrow \nu\nu$ data stat.:	-12.4%	65.8%
$Z \rightarrow \nu\nu$ extrapolation:	-8.3%	53.5%
JES:	-6.5%	12.2%
$W \rightarrow \tau\nu$ extrapolation:	-4.1%	24.5%
$Z \rightarrow \nu\nu$ MC stat.:	-2.6%	24.5%
$W \rightarrow \tau\nu$ MC stat.:	-1.5%	12.2%
$W \rightarrow \tau\nu$ data stat.:	-1.5%	13.6%
QCD normalisation:	-1.5%	7.5%
$W \rightarrow \mu\nu$ MC stat.:	-1.0%	6.1%
Muon ID efficiency:	-0.5%	6.8%
Tau ID efficiency:	-0.5%	5.5%
$W \rightarrow e\nu$ data stat.:	-0.5%	4.8%
$W \rightarrow \mu\nu$ data stat.:	-0.5%	3.4%
$W \rightarrow e\nu$ MC stat.:	-0.5%	3.4%
Electron ID efficiency:	0.0%	0.6%
PU weight:	0.0%	0.6%

Data Samples

Dataset/JSON	Int. Lumi [pb^{-1}]
/MET/Run2012A-22Jan2013-v1/AOD	889
/VBF1Parked/Run2012B-22Jan2013-v1/AOD	3871
/VBF1Parked/Run2012C-22Jan2013-v1/AOD	7152
/VBF1Parked/Run2012D-22Jan2013-v1/AOD	7317
Total analysed	19229
Cert_190456-208686_8TeV_22Jan2013ReReco_Collisions12_JSON.txt	19789

MC Samples-1

Dataset	σ [pb]	No. of Events	Eq. $\int L$ [fb^{-1}]
$(Z \to \nu\nu)$ + jets (50 < HT < 100 GeV)	381.2	4040980	10.6
$(Z \to \nu\nu)$ + jets (100 < HT < 200 GeV)	160.3	4416646	27.6
$(Z \to \nu\nu)$ + jets (200 < HT < 400 GeV)	41.49	5055885	122
$(Z \to \nu\nu)$ + jets (400 < HT < ∞ GeV)	5.274	1006928	191
$(W \to l\nu)$ + jets (inclusive)	37509(NNLO)	76102995	2.03
$(W \to l\nu)$ + 1 jet	5400	23141598	42.9
$(W \to l\nu)$ + 2 jet	1750	34044921	19.5
$(W \to l\nu)$ + 3 jet	519	15539503	29.9
$(W \to l\nu)$ + 4 jet	214	13382803	62.5
$(Z/\gamma \to ll)$ + jets (Mll > 50)	3503.71(NNLO)	30459503	8.7
$(Z/\gamma \to ll)$ + 1 jets (Mll > 50)	561	24045248	42.9
$(Z/\gamma \to ll)$ + 2 jets (Mll > 50)	181	21852156	121
$(Z/\gamma \to ll)$ + 3 jets (Mll > 50)	51.1	11015445	216
$(Z/\gamma \to ll)$ + 4 jets (Mll > 50)	23.04	6402827	278
EWK $(Z/\gamma \to ll)$ + 2 jets	0.888	2978717	3354
EWK $(W^+ \to l\nu)$ + 2 jets	6.48	8996164	1388
EWK $(W^- \to l\nu)$ + 2 jets	4.09	5994018	1466

MC Samples-2

Dataset	σ [pb]	No. of Events	Eq. $\int L$ [fb^{-1}]
WW	54.838(NLO)	10000431	182
WZ	33.21(NLO)	10000283	301
ZZ	17.654(NLO)	9799908	555
$W\gamma$	461.6	4802358	10.4
tt + jets	245.8(NNLO)	6923750	28.2
t (t-channel)	56.4(NLO)	3758227	66.6
t (tW-channel)	11.1(NLO)	497658	44.8
t (s-channel)	3.79(NLO)	259961	68.6
$\bar{t}$ (t-channel)	30.7(NLO	1935072	63.0
$\bar{t}$ (tW-channel)	11.1(NLO)	493460	44.5
$\bar{t}$ (s-channel)	1.76(NLO)	139974	79.5

Objects

PFMET

- Ignore muons
- Type0+1 corrections
- Smeared PFMET for MC

AK5 PFJets

- L1FastJet+L2+L3(+L2L3Residual) JEC
- "Loose" PF Jet ID
- Cleaned with veto leptons
- "Loose" PU jet ID
- Smeared jet collection for MC (JER is smeared to match data)

Veto leptons

- loose+PFiso muons $p_T > 10$ GeV, $|\eta| < 2.1$
- veto+PFiso electrons $p_T > 10$ GeV, $|\eta| < 2.4$

Tight leptons

- As veto leptons but "tight" ID and $p_T > 20$ GeV

Hadronic taus

- $p_T > 20$ GeV, $|\eta| < 2.3$,$d_Z < 0.2$ cm
- Tight ID, discriminant "byTightCombinedIsolationDeltaBetaCorr3Hits"
- Efficiency ~0.55, fake rate 0.02(barrel),0.03(endcap)

Data/MC reweighting

- We apply the following reweightings of the MC to match the data:
  - Trigger efficiency
  - Lepton efficiency: ID and isolation
  - Pileup distribution

Other approaches investigated and run 2 prospects

Initial investigations

- Planned to define a loose pre-selection and model QCD shape
- Several options for analysis strategy:
  - Rectangular cuts and counting experiment
  - Rectangular cuts and shape experiment
  - MVA and counting experiment
  - MVA and shape experiment
- Due to trigger conditions no appropriate QCD control region found - details later

Run 2 prospects

- Prescaled looser control triggers planned to enable better QCD control region
- Will reinvistigate shape analysis and MVA

Software framework strategy

Prompt analysis

- Two frameworks: Analyses A and B
- independent ntuples and analysis code

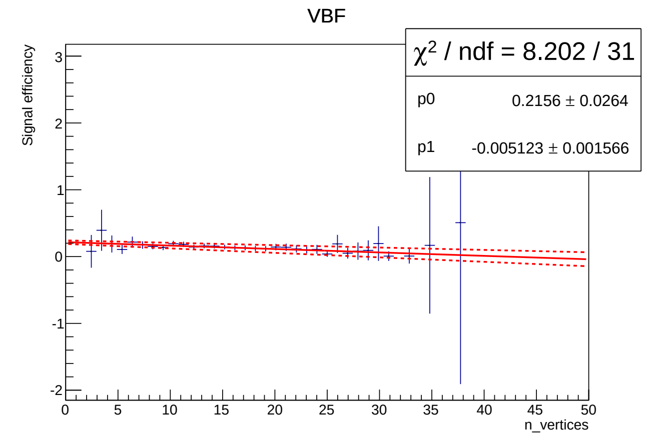
Parked analysis

- Insufficient manpower to maintain and develop two frameworks
- Moved to one fully developed framework
  - New framework is development of analysis B and uses same ntuples
- Synchronised yields in signal and control regions between new framework and old analyses A and B
- Repeated expected limit calculation from HIG-13-030 analysis with the new framework and parked data
  - Agrees with HIG-13-030 to within 2%, which is good given rereco, and change of global tag and triggers

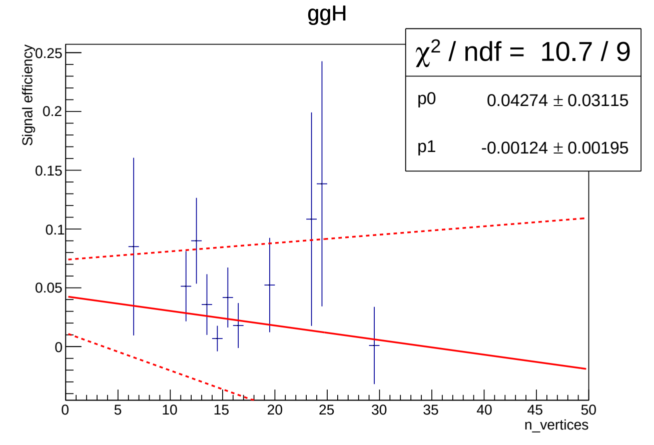
Software cross check

- Starting point region numbers also cross-checked in alternate software
  - Agreement is better than 0.5%

Signal efficiency as a function of PU

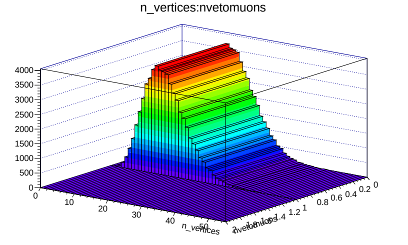


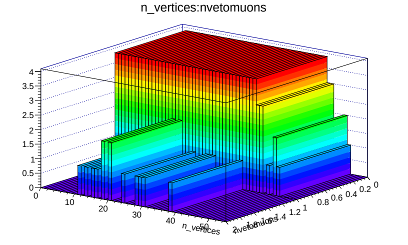
Signal efficiency as a function of PU



Veto muons in signal MC

- Veto muons don't have a dz or dxy cut
- Concern that we would be vetoing muons from a different vertex
- Muon veto efficiency turns out to be very high:
  - $\sim$10 signal MC events with a veto muon out of $\sim$55000
- nvetomuons doesn't seem correlated with PU

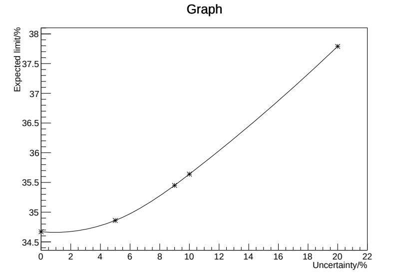




- Right plot is zoom of left

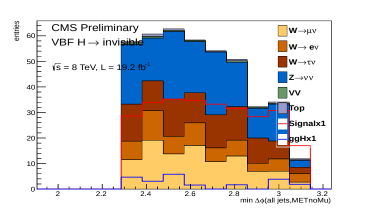
Effect of $Z/\gamma^* \to \mu\mu$ to $Z \to \nu\nu$ uncertainty

Uncertainty/%	Median expected limit/%
20	37.79
10	35.64
9	35.45
5	34.86
0	34.67

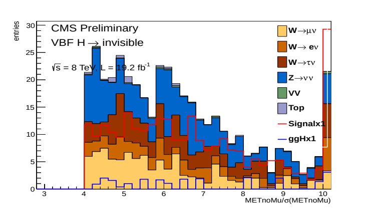


Signal Region Control Plots

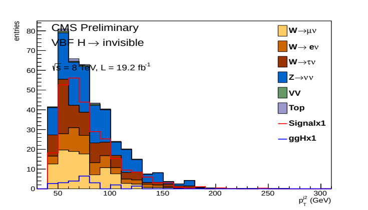
Min$\Delta\phi$(MET,all jets)



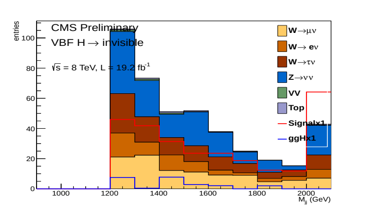
$METnoMU/\sigma_{METnoMU}$



$Jet2p_T$

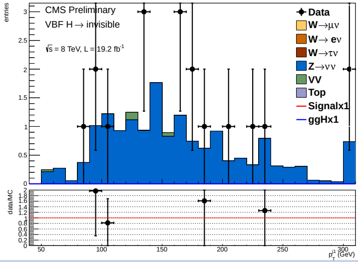


M_{jj}

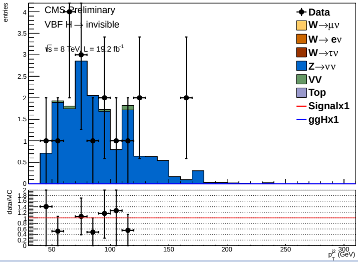


Control plots - mumu

Jet 1 pt

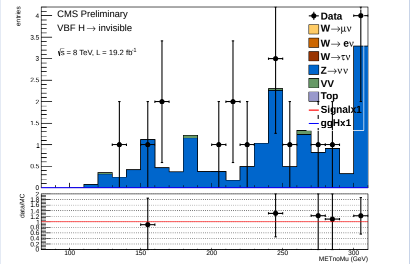


Jet 2 pt

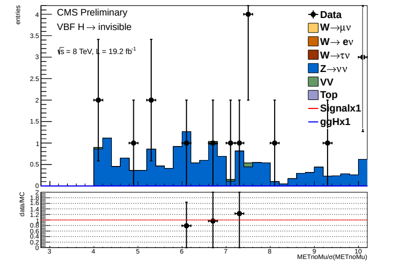


Control plots - mumu

METnomu

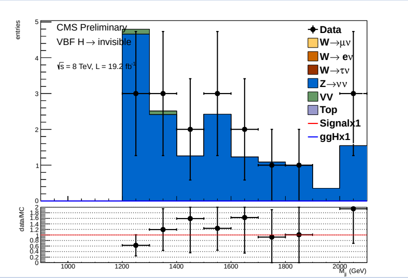


METnomusig

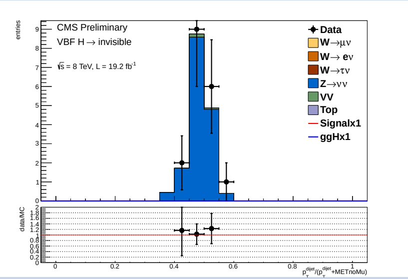


Control plots - mumu

Mjj

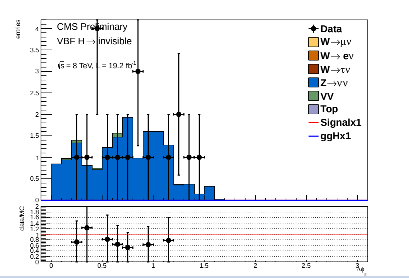


dijet-metnomu pt fraction

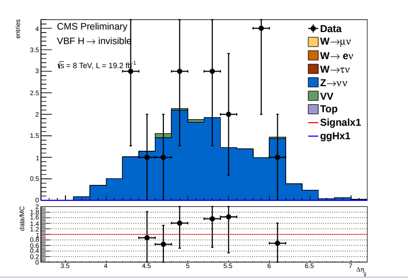


Control plots -mumu

Dphijj

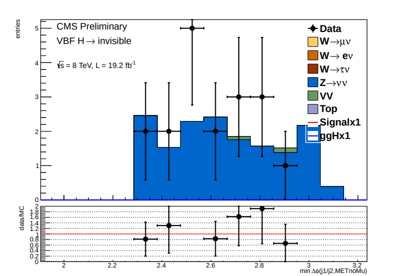


Detajj

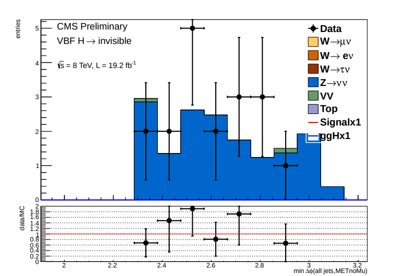


Control plots -mumu

Leading jets-met mindphi

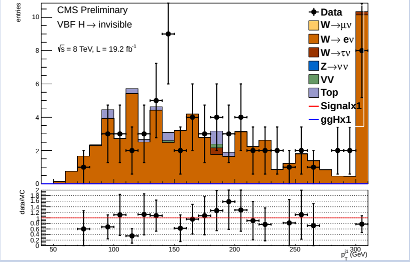


All jet-met mindphi

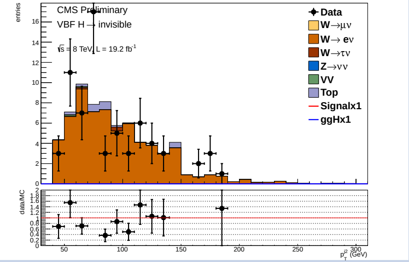


Control plots -enu

Jet 1 pt

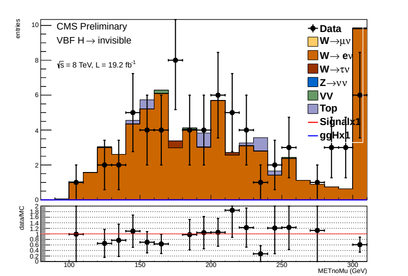


Jet 2 pt

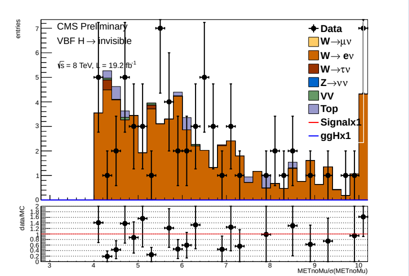


Control plots -enu

METnomu

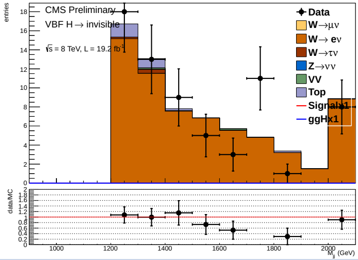


METnomusig

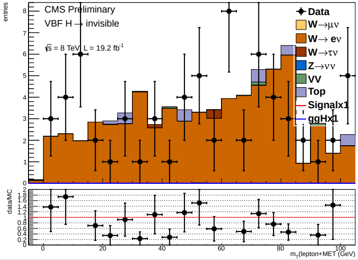


Control plots - enu

Mjj

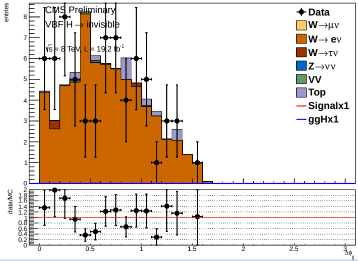


mt

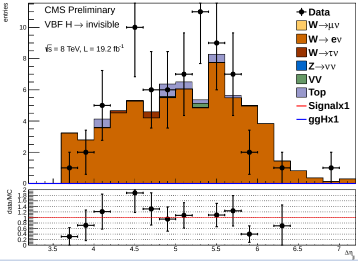


Control plots - enu

Dijet Dphi

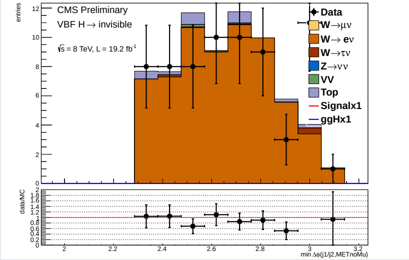


Detajj

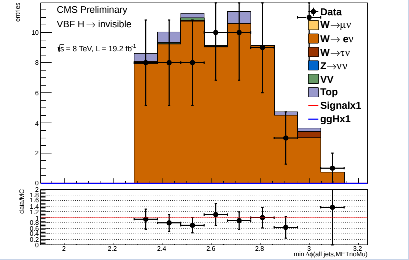


Control plots - enu

Leading jets-met mindphi

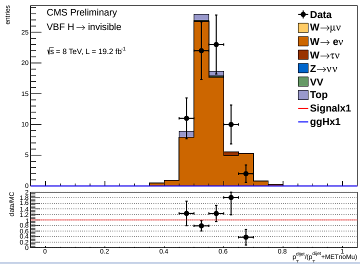


All jets-met mindphi



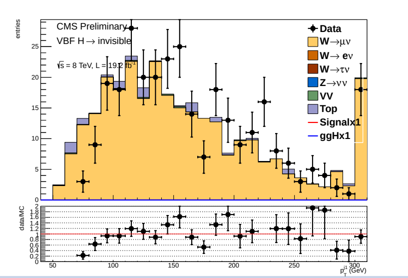
Control plots - enu

dijet-metnomu pt fraction

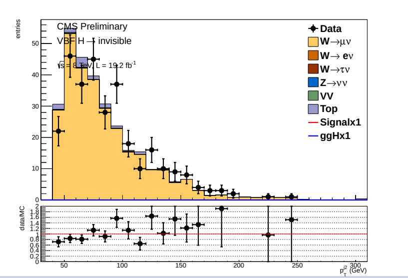


Control plots - munu

Jet 1 pt

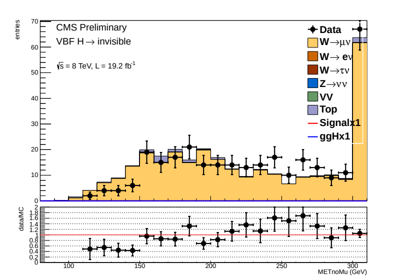


Jet 2 pt

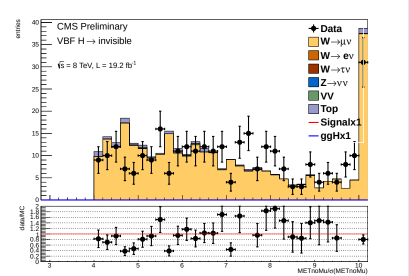


Control plots - munu

METnomu

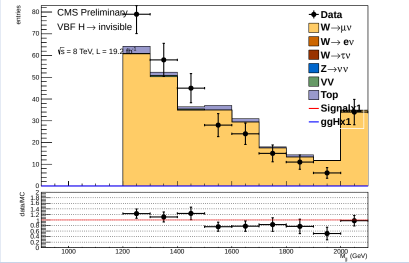


METnomusig

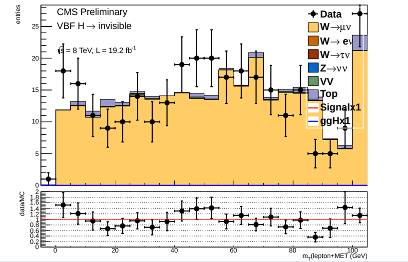


Control plots - munu

Mjj



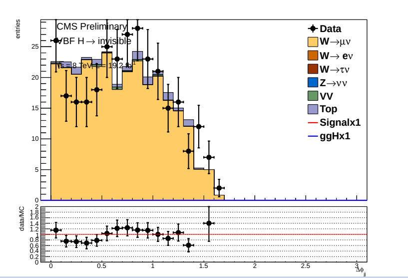
mt



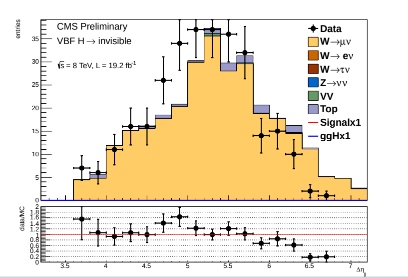
Imperial College London

Control plots - munu

Dijet Dphi

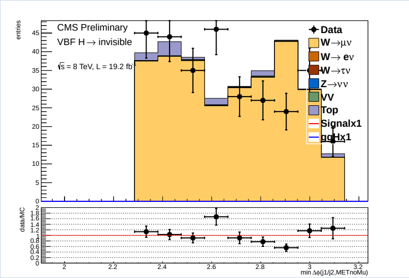


Detajj

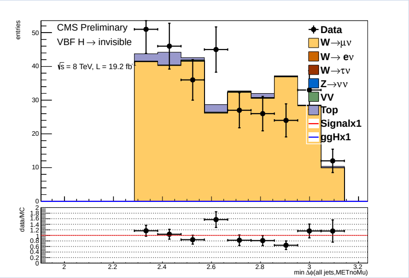


Control plots - munu

Leading jets-met mindphi

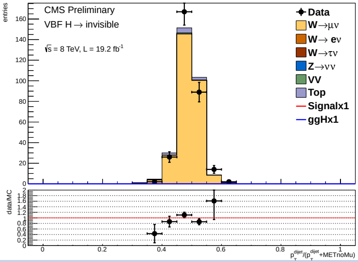


All jets-met mindphi



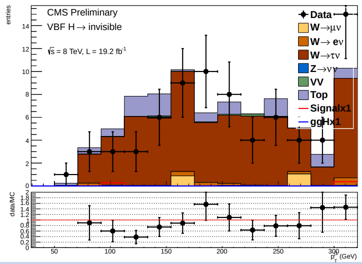
Control plots - munu

dijet-metnomu pt fraction

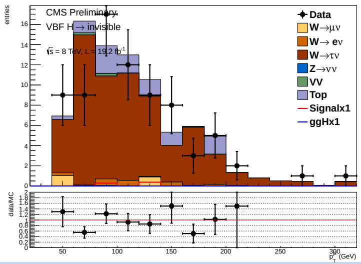


Control plots - taunu

Jet 1 pt

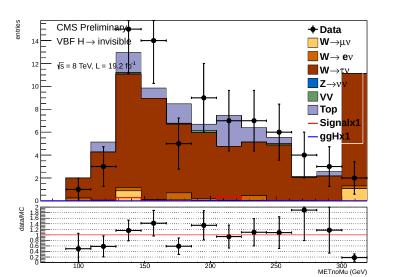


Jet 2 pt

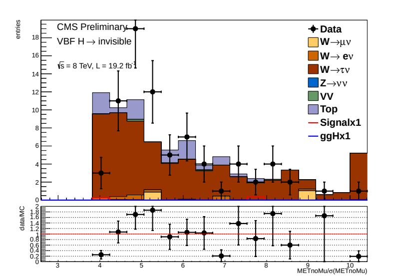


Control plots - taunu

METnomu

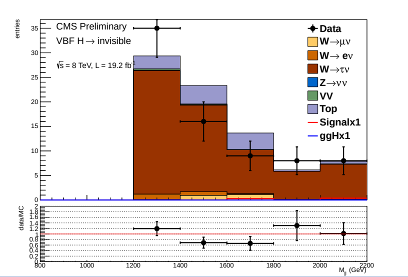


METnomusig

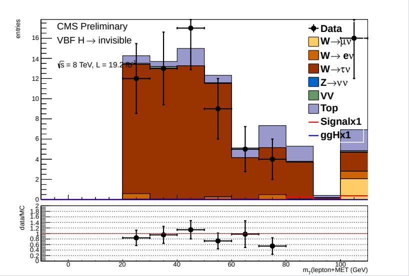


Control plots - taunu

Mjj



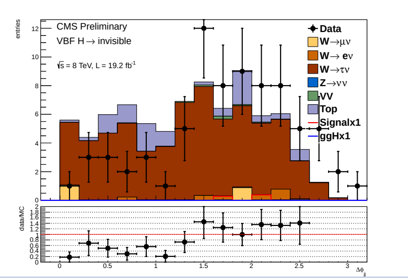
mt



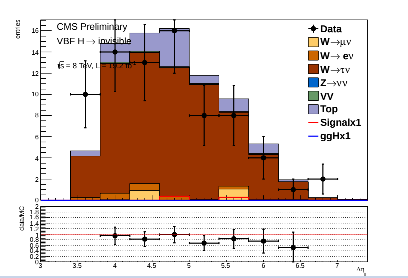
Imperial College London

Control plots - taunu

Dijet Dphi

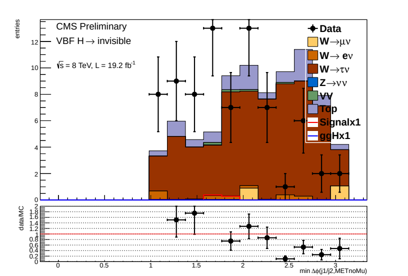


Detajj

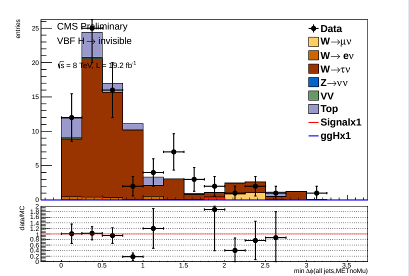


Control plots - taunu

Leading jets-met mindphi

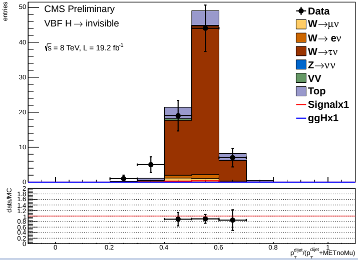


All jets-met mindphi



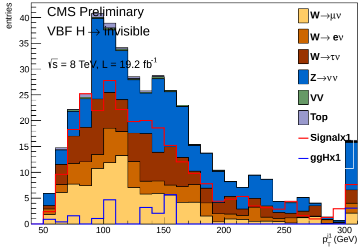
Control plots - taunu

dijet-metnomu pt fraction

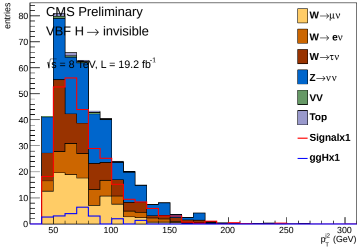


Control plots - nunu

Jet 1 pt

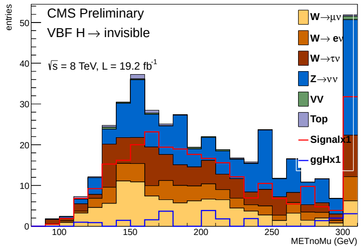


Jet 2 pt

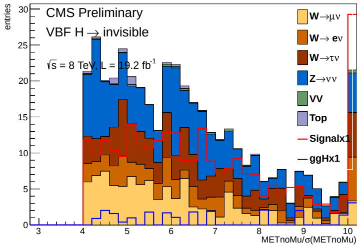


Control plots - nunu

METnomu

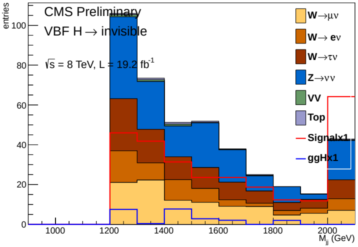


METnomusig



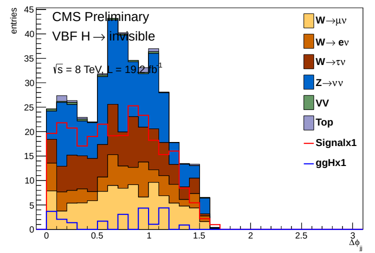
Control plots - nunu

Mjj

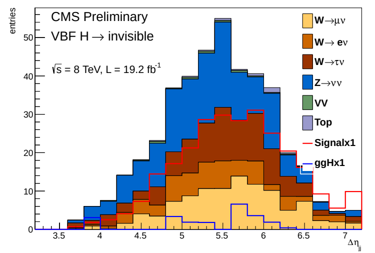


Control plots - nunu

Dijet Dphi

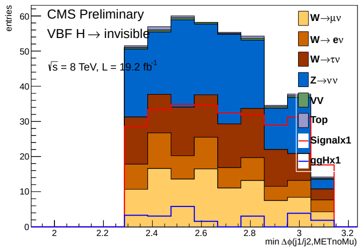


Detajj

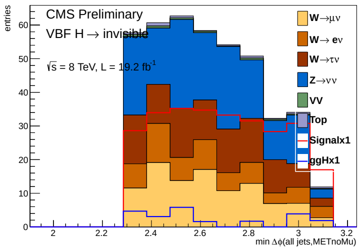


Control plots - nunu

Leading jets-met mindphi

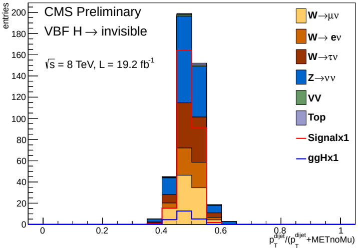


All jets-met mindphi



Control plots - nunu

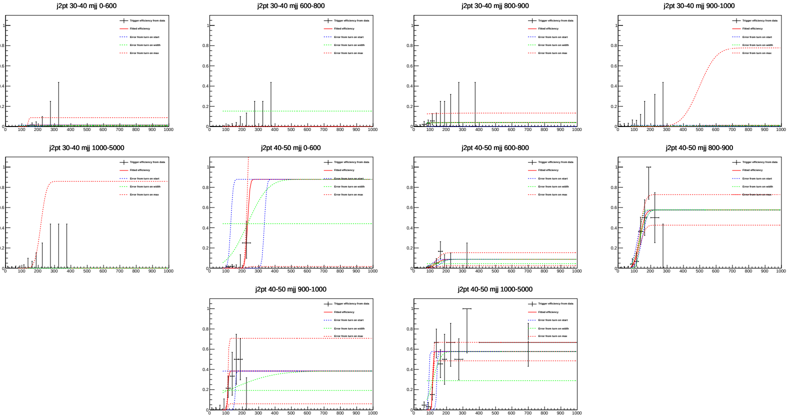
dijet-metnomu pt fraction



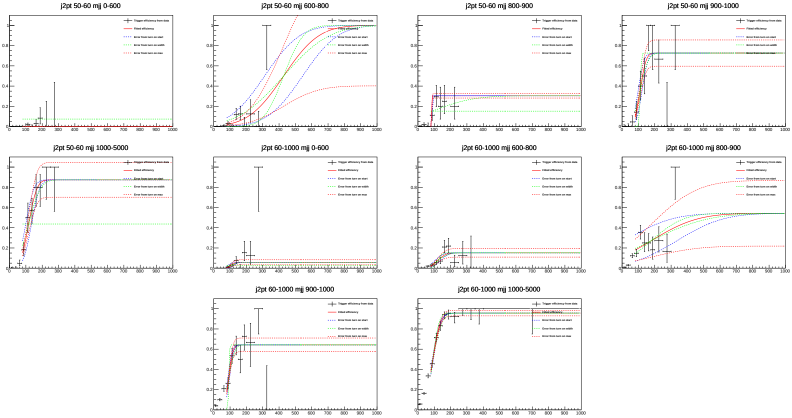
Trigger efficiency error

- Bin used with largest uncertainty picked for each run period
- Worst case scenario assumed of all bins having this uncertainty gives 2.3% uncertainty
- Error cancels in all data driven backgrounds
- Only affects signal and VV
- Small compared to other uncertainties and doesn't affect limit
- Treated as negligible

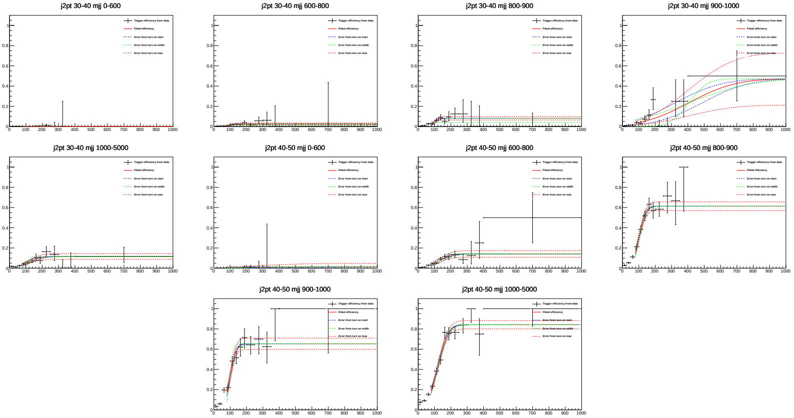
Trigger efficiency-Run A-1



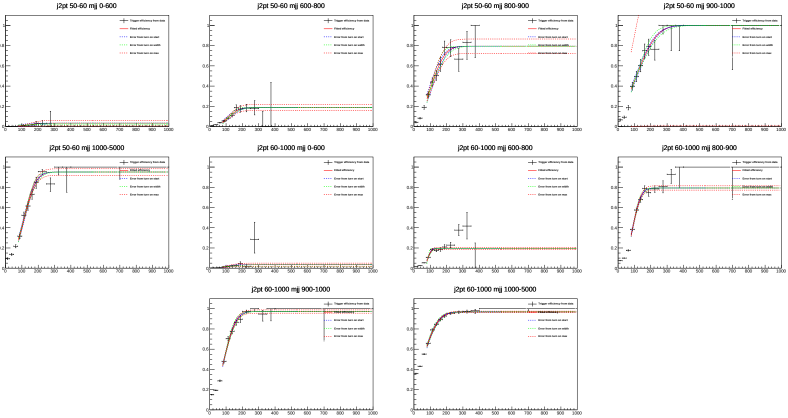
Trigger efficiency-Run A-2



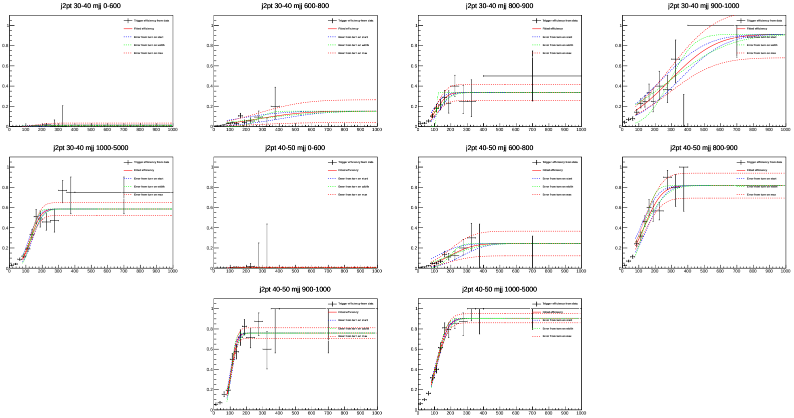
Trigger efficiency-Run BC-1



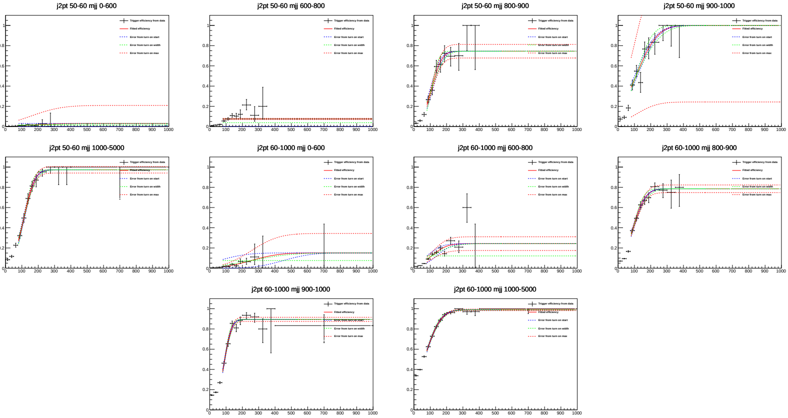
Trigger efficiency-Run BC-2



Trigger efficiency-Run D-1



Trigger efficiency-Run D-2



QCD options tried

Several methods tried to model QCD

Standard MC

- doesn't have enough events

Private VBF+MET enriched QCD MC sample

- Can only enrich in events with real met
- Can't model met from mismeasurement

Data-driven shape using different jet pairs in the event

- Jet kinematics are very biased
- Ordering in p_T and angle have been tried
- Reweighting individual distributions to fix others has been tried

VBF enriched QCD MC

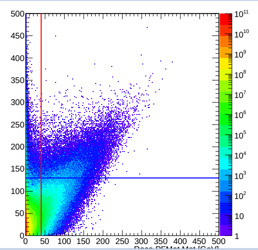
MC Filter: Vectorial sum of neutrino E_T

- $\sum E_{\perp}(\vec{\nu}) > 40 \; GeV$

MC Filter: Dijet Filter

- Select jets with:
  - $p_{\perp} > 20 \; GeV$
  - $|\eta| < 5.0$
- From selected jets at least one pair with:
  - $m_{jj} > 700 \; GeV$
  - $\Delta\eta > 3.2$

QCD Inc 80-600 *GeV*

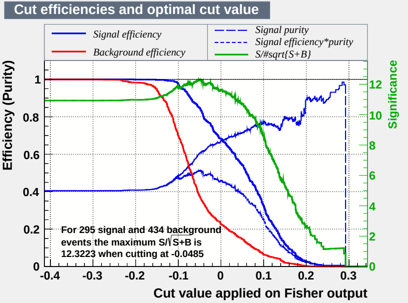


There 2 distict populations of events: real and fake met.

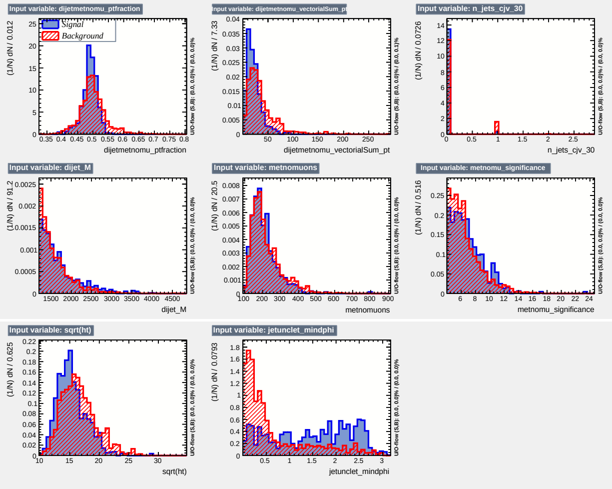
Sample	Ev. Gen.	Filter Eff.	Events	XS [*pb*]	Eq. Lumi. [fb^{-1}]
QCD-Pt-80to120	39376000000	0.000049	1614416	1033680	38.09
QCD-Pt-120to170	7000000000	0.000283	2051000	156293.3	44.79
QCD-Pt-170to300	1375000000	0.000987	1391500	34138.15	40.28
QCD-Pt-300to470	80000000	0.002659	207840	1759.549	45.47
QCD-Pt-470to600	25000000	0.004127	104675	113.8791	219.53

BDT Study

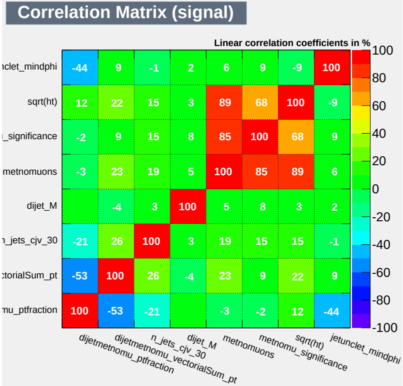
- Had a quick look at MVA analysis
- Started from cut based signal region
  - Only region with negligible QCD
- Best expected limit obtained 37%
  - Does not take into account any increased systematic
  - Therefore unlikely to be worthwhile
- New variables could make MVA worthwile
- Ability to model QCD would enable looser starting selection which may make MVA worthwhile

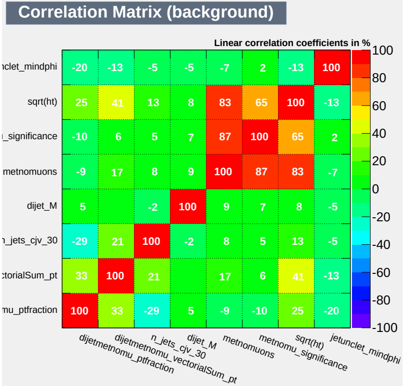


MVA Details


Input variable: dijetmetnomu_ptfraction
Signal
Background
(1/N) dN / 0.012
dijetmetnomu_ptfraction
Input variable: dijetmetnomu_vectorialSum_pt
(1/N) dN / 7.33
dijetmetnomu_vectorialSum_pt
Input variable: n_jets_cjv_30
(1/N) dN / 0.0726
n_jets_cjv_30
Input variable: dijet_M
(1/N) dN / 91.2
dijet_M
Input variable: metnomuons
(1/N) dN / 20.5
metnomuons
Input variable: metnomu_significance
(1/N) dN / 0.516
metnomu_significance
Input variable: sqrt(ht)
(1/N) dN / 0.625
sqrt(ht)
Input variable: jetunclet_mindphi
(1/N) dN / 0.0793
jetunclet_mindphi

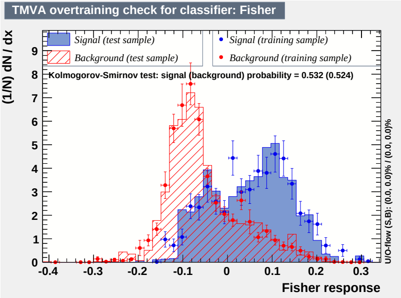
MVA Details

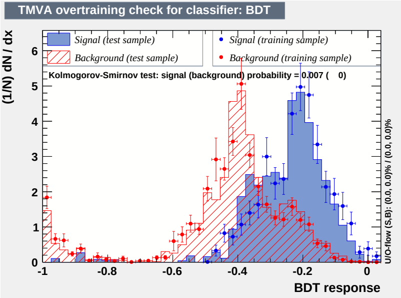

Correlation Matrix (signal)
Linear correlation coefficients in %
nclet_mindphi -44 9 -1 2 6 9 -9 100
sqrt(ht) 12 22 15 3 89 68 100 -9
_significance -2 9 15 8 85 100 68 9
metnomuons -3 23 19 5 100 85 89 6
dijet_M -4 3 100 5 8 3 2
n_jets_cjv_30 -21 26 100 3 19 15 15 -1
ctorialSum_pt -53 100 26 -4 23 9 22 9
mu_ptfraction 100 -53 -21 -3 -2 12 -44
dijetmetnomu_ptfraction
dijetmetnomu_vectorialSum_pt
n_jets_cjv_30
dijet_M
metnomuons
metnomu_significance
sqrt(ht)
jetunclet_mindphi
100
80
60
40
20
0
-20
-40
-60
-80
-100


Correlation Matrix (background)
Linear correlation coefficients in %
nclet_mindphi -20 -13 -5 -5 -7 2 -13 100
sqrt(ht) 25 41 13 8 83 65 100 -13
_significance -10 6 5 7 87 100 65 2
metnomuons -9 17 8 9 100 87 83 -7
dijet_M 5 -2 100 9 7 8 -5
n_jets_cjv_30 -29 21 100 -2 8 5 13 -5
ctorialSum_pt 33 100 21 17 6 41 -13
mu_ptfraction 100 33 -29 5 -9 -10 25 -20
dijetmetnomu_ptfraction
dijetmetnomu_vectorialSum_pt
n_jets_cjv_30
dijet_M
metnomuons
metnomu_significance
sqrt(ht)
jetunclet_mindphi
100
80
60
40
20
0
-20
-40
-60
-80
-100

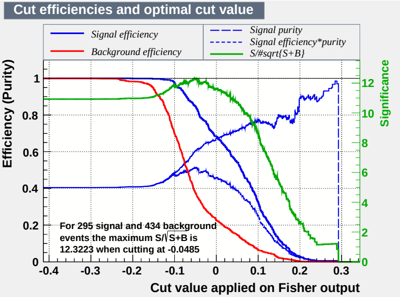
Imperial College London

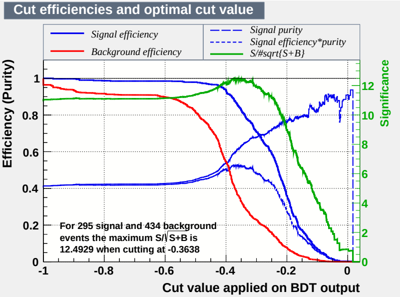
MVA Details





MVA Details


Cut efficiencies and optimal cut value
Signal efficiency
Background efficiency
Signal purity
Signal efficiency*purity
S/#sqrt{S+B}
Efficiency (Purity)
Significance
For 295 signal and 434 background events the maximum S/√S+B is 12.3223 when cutting at -0.0485
Cut value applied on Fisher output


Cut efficiencies and optimal cut value
Signal efficiency
Background efficiency
Signal purity
Signal efficiency*purity
S/#sqrt{S+B}
Efficiency (Purity)
Significance
For 295 signal and 434 background events the maximum S/√S+B is 12.4929 when cutting at -0.3638
Cut value applied on BDT output

No systematics limits

- Prompt limit with no systematics 16.6%
- Parked limit with no systematics 14.3%