CDF/ANAL/EXOTIC/CDFR/9955
Version 2.0
February 3, 2010

A search for the Higgs Boson in the All Hadronic Channel with Data Sample of 4 fb^{-1}

Ankush Mitra,Shang-yuu Tsai,Song-Ming Wang [1]
Academia Sinica, Taiwan

Abstract

This note documents the search for the Higgs boson in the all hadronic $W/ZH \rightarrow q\bar{q}b\bar{b}$ and $Hqq \rightarrow qqb\bar{b}$ mode. The search is performed on 4 fb^{-1} of data recorded at CDF. As no signal was observed, 95% confidence limits are quoted for Higgs masses from $100\,\mathrm{GeV/c^2}$ to$150\,\mathrm{GeV/c^2}$.

---

[1] mitra@fnal.gov,sytsai@fnal.gov, smwang@fnal.gov

Contents

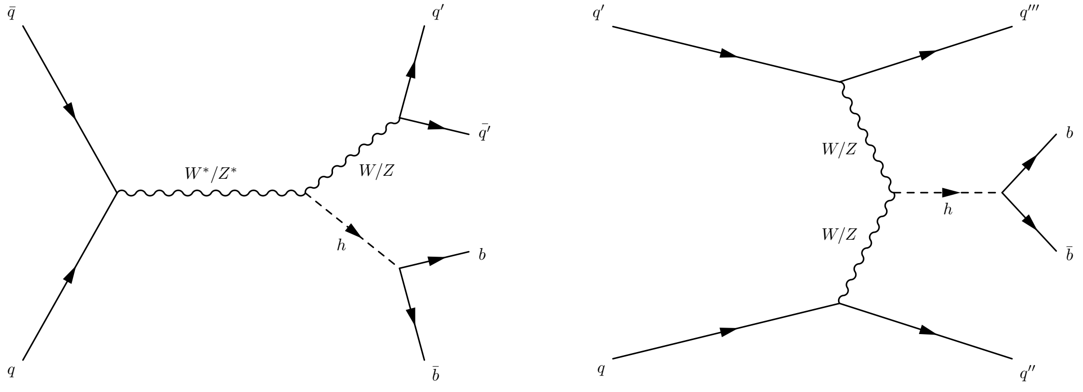
1 Introduction

The Higgs boson plays a central role in the Standard Model as it endows particles with mass. In the past, most of the searches for the Higgs boson at CDF have concentrated on channels using combinations of leptons, jets and missing-Et. Recently a search for the Higgs boson in the all hadronic channel was performed at CDF with 2 fb^{-1} of data sample [1] [2]. In that analysis the Higgs boson was searched in the production channels where the Higgs boson is produced in association with a W or Z boson, as shown in figure 1(a). The decay channels considered in the analysis are $H \rightarrow b\bar{b}$ and $W/Z \rightarrow q\bar{q}$. Thus the final state consists of at least four jets.

This note describes the continuation of the Higgs boson search in the all hadronic channel with a larger data sample of 4 fb^{-1}, which covers from p1 to p21 of the CDF datasets. Several improvements are added to this new analysis :

- Using data collected by a new VH_MULTIJET trigger [3] which has lower SumEt threshold applied at the CDF Level-2 trigger. This trigger was only implemented in May 2008, and thus only covers about 1 fb^{-1} of the data sample use for this analysis. The other 3 fb^{-1} of data sample is collected with the TOP_MULTI_JET which has a higher SumEt threshold at Level-2 The lower SumEt threshold for the VH_MULTIJET improved the trigger efficiency for the low mass Higgs compared to the TOP_MULTI_JET trigger.
- Inclusion of the vector boson fusion (VBF) production channel $qq \rightarrow qqH$ (see figure 1(b)), where the final state also consists of at least four jets.
- Adding a new double b-tag category. In the 2 fb^{-1} analysis we selected events with double tight SecVtx b-tag (SecVtx(tight)-SecVtx(tight)). In this new analysis we added the new SecVtx(tight)-JetProb(1%) double b-tag category. This addition can increase the Higgs signal acceptance by $\sim 40\%$. However the QCD multi-jet background is expected to increase by $\sim 55\%$.

In this 4 fb^{-1} analysis a Neural Net (NN) algorithm is used to discriminate the possible Higgs signal from the background.



(a) Production and decay mode of the $VH \rightarrow b\bar{b}qq$ channel

(b) Higgs boson production via vector boson fusion process.

Figure 1: Feynman diagrams for the two Higgs production channels studied in this ananlysis: Assocated Vector Boson Production & Vector Boson Fusion.

2 Data and Trigger

The data for this analysis was collected by two triggers: TOP_MULTI_JET & VH_MULTIJET . They are designed to select events with 4 high p_T jets with large Sum-Et which are characteristic of an

	L1	L2	L3	Integrated Luminosity / fb^{-1}
TOP_MULTIJET_v1-8	JET_10	L2_FOUR_JET15_SUMET125	4 Jet-15	0.965
TOP_MULTIJET_v-9	JET_20	L2_FOUR_JET15_SUMET175	4 Jet-15	1.255
TOP_MULTIJET_v-12	JET_20	L2_FOUR_JET15_SUMET175	4 Jet-15	0.619
VH_MULTIJET	JET_20	L2_THREEJET20_SUMET130	NULL	1.118

Table 1: Definition of the different versions of TOP_MULTI_JET and VH_MULTIJET and the recorded integrated luminosity. The integrated luminosity for TOP_MULTI_JET v-12 is the period before VH_MULTIJET was added to the CDF trigger table.

	L1 MC Scale Factor	L2 MC Scale Factor
TOP_MULTIJET_v1-8	Not Needed	0.963 ± 0.008
TOP_MULTIJET_v-9	$f(x) = A\left(1 - \exp^{-Bx}\right) + C$ $A = 19.23 \pm 0.01$ $B = 2.80E - 02 \pm 0.04E - 02$ $C = -18.24 \pm 0.01$	0.963 ± 0.008
TOP_MULTIJET_v-12	$f(x) = A\left(1 - \exp^{-Bx}\right)$ $A = 0.994 \pm 0.001$ $B = 0.0132 \pm 0.0003$	0.973 ± 0.06
VH_MULTIJET	$f(x) = A\left(1 - \exp^{-Bx}\right)$ $A = 0.9959 \pm 0.0009$ $B = 0.0144 \pm 0.0002$	0.953 ± 0.004

Table 2: The MC corrections for the multijet triggers. The overall correction is a product of the L1 & L2 scale factors. The x for the L1 scale factors is the total E_t of all jets above 15 GeV in the event [4].

all hadronic Higgs event. The first 2.8 fb^{-1} of CDF data was collected by TOP_MULTI_JET trigger and the remaining 1.1 fb^{-1} by VH_MULTIJET . The definition of the triggers and the recorded integrated luminosity are given in table 2.

The performance of the multijet triggers are senstive to the topology of the event which precludes the use of standard techniques of measuring the trigger turn-on from data; one has to rely on simulation. A study of the multijet triggers was performed and a set of corrections were derived to correct the MC [4] which are:

- The simulated Level-2 cluster energies had to be rescaled to compensate for a bug in the CDF-code [5].
- A trigger acceptance correction has to be applied. One correction corrects the L1_JET20 trigger response which is SumEt dependent and another L2 scale factor (Table 2).

2.1 Event Selection

2.1.1 Initial Event Selection

The following criteria was applied for events to be considered in this analysis:

- The event must be in the QCD with Silicon Good Run List V28
- The event must have fired the TOP_MULTI_JET or VH_MULTI_JET trigger. For MC, the corrected trigger decision is used.
- The $|Z|$ position of the highest P_T class($\geq$12) vertex must be less than 60 cm.

- A lepton veto is applied to remove events from $t\bar{t}$ decay and to ensure orthogonality with other Higgs searches. The Joint-Physics Lepton ID [6] cuts are used.
- The event must have 4 or 5 jets, where a jet is defined as:
  - 0.4 cone L7-corrected jet-Et $> 15.0\,\mathrm{GeV}$
  - 0.4 cone raw jet-Et $> 10.0\,\mathrm{GeV}$
  - Jet $|\eta| < 2.4$
- The event's MET (missing transverse energy) significance must be less than 6. This cut is used to reduce the contribution from Top quark pair production background.

2.1.2 Final Event Selection

For events which pass the intial event selection, they must fulfill the next set of criteria to be considered for the analysis.:
NB: Only the four leading jets are considered for the event selection. The state of fifth, sixth, etc jets have no affect.

- From the 4 leading jets, at least one or two jets must be tagged as a b-jet. One of the b-tagged jets must be tagged by tight SecVtx.
- Events with > 2 b-tagged jets (tagged by either tight SecVtx, or JetProb(1%)) from the 4 leading jets are rejected.
- The events SumEt (based on the sum of the transverse energy of the selected jets) must be greater than 220 GeV. This cut is to improve the signal to QCD multi-jet background significance.

Events which pass this final event selection are classified as:

- **One taggged events (1-Tag):** Events with exactly only one tight SecVtx tag. They are used to derive the background for two-tagged events. The reason to require only one b-tag is to reduce the chances of including the possible Higgs signal into the background prediction.
- **Two tagged events:** These are events with exactly 2 b tags. These events can be from these two categories :
  - SecVtx(tight)-SecVtx(tight) (SS)
  - SecVtx(tight)-JetProb(1%) (SJ)

We only consider an event for the SecVtx(tight)-JetProb(1%) double b-tag category if it fails the SecVtx(tight)-SecVtx(tight) requirement. Therefore the events in these two categories are exclusive.

2.1.3 Signal Regions for the Analysis

In this analysis we are searching for Higgs boson that is produced either through the associated production channel ($p\bar{p} \to W/Z + H$) or through the vector boson fusion channel ($p\bar{p} \to qqH$). For the associated production channel, there are two mass resonances in the final state. One is from the Higgs boson ($H \to b\bar{b}$), and the other is from the W or Z boson ($W \to qq'$, $Z \to q\bar{q}$). Whereas in the vector boson fusion channel, there is only one mass resonance from the Higgs boson decay. The other two jets in the final state are the two out-going quarks that have radiated off a pair of W or Z bosons that fused to form the Higgs boson. These two jets tend to go in the forward direction in the laboratory frame. These two jets are very far apart from each other and thus form a very board di-jet mass distribution. In this analysis for the Higgs signal, we assume that the two b-tagged jets are from the Higgs boson. The other two non b-tag jets (among the first four leading jets) are taken to be the

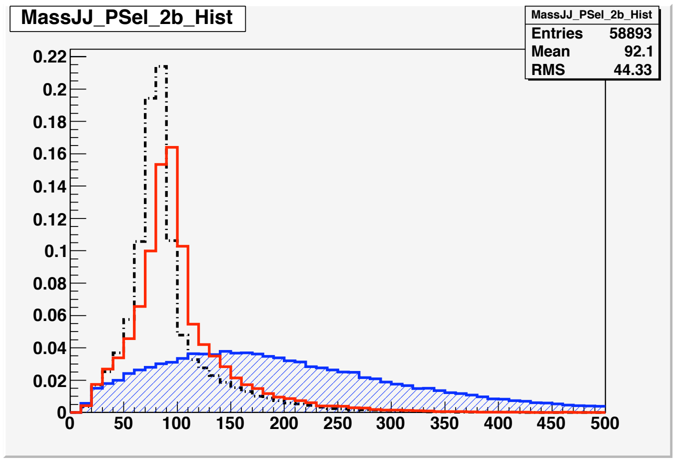


Figure 2: The di-jet invariant mass distributions of the two non b-tag jets for the associated productions and the vector boson fusion. (Black-dash) WH, (red-solid) ZH and (blue-hash) vector boson fusion productions.

jets coming from the W or Z boson decay for the associated production channel, or the two out-going quarks in the vector boson fusion channel. Figure 2 shows the di-jet invariant mass distribution of the two non b-tag jets for the associated productions and the vector boson fusion.

We make use of the two masses $M(bb)$ (invariant mass of the two b-tag jets) and $M(qq)$ (invariant mass of the two non b-tag jets) to define two signal regions in the $M(bb)$ vs $M(qq)$ mass plane, as shown in figure 3. The two signal regions are :

- WH/ZH signal region : $75 < M(bb) < 175$ GeV/c^2, $50 < M(qq) < 120$ GeV/c^2
- VBF signal region : $75 < M(bb) < 175$ GeV/c^2, $M(qq) > 120$ GeV/c^2

Both signal regions have a common $M(bb)$ mass range. This is because we are interested in searching for light Higgs boson with mass between 100 to 150 GeV/c^2.

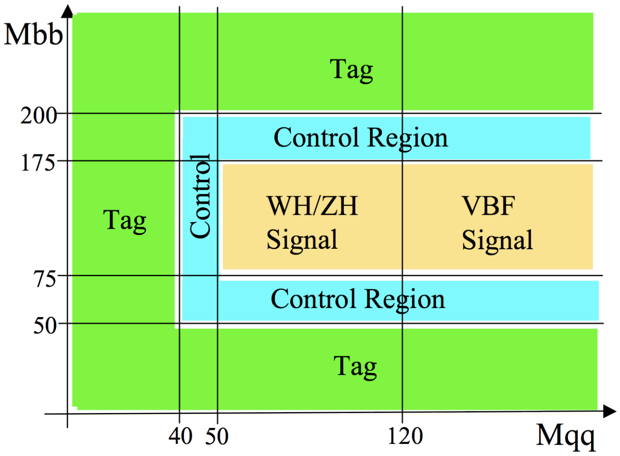


Figure 3: The definitions of TAG, CONTROL and signal regions in the $M(bb)$ vs $M(qq)$ mass plane.

2.1.4 Backgrounds

The backgrounds considered in this analysis are:

- QCD Multi-jets : this is the dominant background as it consists of about 98% of the total background
- Top quark pair production
- Single Top quark production
- W and heavy-flavor jets production
- Z plus jets production, with Z decays into $b\bar{b}$ and $c\bar{c}$.
- Diboson production : WW, WZ, ZZ

The QCD multi-jet will be estimated from a data driven based technique. The non-QCD backgrounds will be estimated from simulation.

2.2 Signal and background Monte-Carlo samples

The signal and background Monte-Carlo samples were generated using a combination of Pythia[7] and Alpgen [8] (See table 3). The generated events were then passed through the CDF detector simulation. For each sample, a set of corrections were applied:

- The simulated level 2 calorimeter clusters were corrected as described in CDF NOTE 9954 [4]. These corrected clusters were used to remake the trigger decision [4].
- An additional trigger acceptance correction was applied to the MC [4].
- A scale factor is applied to account for the difference in b-tagging efficiency measured in data and MC. The scale factor for a single tight SecVtx tag is 0.946 ± 0.036, and the scale factor for a 1% JetProb tag is 0.739 ± 0.044 [9]. Therefore the effective scale factor for SecVtx(tight)-SecVtx(tight) category is 0.895 ± 0.068, and the effective scale factor for SecVtx(tight)-JetProb(1%) category is 0.699 ± 0.068.

All the background events were generated as fully-inclusive except the signal and the Z+jets sample. The fully inclusive Z+jets cross-section is large and the majority of the events would not pass the trigger. A filter was used to select events which were likely to pass the trigger and event selection. The filter was devised at the generator level to select events with b/c partons, have 3 jets with $E_T > 5$ Gev and a Sum- $E_T > 60$ GeV. Further details on the filter can be found in section 10.1. The VH signal samples were generated with Pythia where the Higgs was forced to decay to $b\bar{b}$ and the W/Z was forced to decay to hadrons.

The Higgs signals (associated production and VBF) are all generated with Pythia and the goodrun listed used is up to period 17 (runlist_p0-p8_p10-p12_p15-p17_2008-12-17.txt) which consists of low, medium and high luminosity profiles. The $t\bar{t}$ MC samples used in this analysis (ttop75, otop49) also contain the low, medium and high luminosity profiles.

2.3 Expected Signal and Backgrounds

Tables 5 and 6 summarise the number of signal and background events expected in 4 fb^{-1} and the final number of events used in the analysis. $t\bar{t}$ is the second largest followed by Z+jets, single-top and wbb. However the largest background source is QCD which is not modelled by Monte-Carlo. Instead a data-driven model is derived to predict the QCD contribution.

Table 3: Generators used for Non-QCD background samples

Channel	Generator
$t\bar{t}$	Pythia
Single Top - S channel	Pythia
Single Top - T channel	Pythia
WW	Alpgen+Pythia
WZ	Alpgen+Pythia
ZZ	Alpgen+Pythia
$W + b\bar{b}$	Alpgen+Pythia
$W + c\bar{c}$	Alpgen+Pythia
$Z \to bb/c\bar{c} + jj$	Pythia

Table 4: The cross-sections for the non-QCD backgrounds and the sources used.

Process	Cross-section	Source
$t\bar{t}$	7.884 pb	CDF Note 9999 [10]
Single Top S channel	1.083 pb	CDF Note 9999 [10]
Single Top T channel	2.295 pb	CDF Note 9999 [10]
WW	12.4 pb	CDF Note 9999 [10]
WZ	3.7 pb	CDF Note 9999 [10]
ZZ	3.8 pb	CDF Note 9999 [10]
$W + b\bar{b}$	24.696 pb	(Alpgen+Pyhtia LO x 1.4 K_NLO)
$W + c\bar{c}$	40.642 pb	(Alpgen+Pythia LO x 1.4 K_NLO)
$Z \to b\bar{b}/c\bar{c} + jj$	700.26 pb	(Pythia LO x 1.4 K_NLO)

Table 5: Table of WH, ZH and VBF Signal sources. The last two columns on the right are the total number of events which pass the trigger, event selection, have two b-tags and have been scaled by the MC corrections.

	σ /pb	$\sigma\times$ BR /pb for $q\bar{q}b\bar{b}$	Number of VH/VBF events in 4 fb^{-1}	Used in analysis	
				SS	SJ
WH 100	0.2861	0.1571	621	5.9	2.2
WH 105	0.2446	0.1316	521	5.7	2.2
WH 110	0.2092	0.1089	431	5.4	2.0
WH 115	0.1788	0.0885	350	5.1	1.9
WH 120	0.1529	0.0702	278	4.5	1.7
WH 125	0.1324	0.0546	216	3.9	1.5
WH 130	0.1147	0.0409	162	3.3	1.2
WH 135	0.0993	0.0293	116	2.5	1.0
WH 140	0.0860	0.0200	79	1.9	0.7
WH 145	0.0753	0.0130	51	1.3	0.5
WH 150	0.0660	0.0078	31	0.8	0.3
ZH 100	0.1667	0.0946	374	4.4	1.5
ZH 105	0.1440	0.0801	317	4.1	1.5
ZH 110	0.1243	0.0669	265	4.0	1.4
ZH 115	0.1074	0.0550	218	3.7	1.3
ZH 120	0.0927	0.0440	174	3.3	1.2
ZH 125	0.0811	0.0346	137	2.8	1.0
ZH 130	0.0709	0.0261	103	2.3	0.8
ZH 135	0.0620	0.0189	75	1.8	0.6
ZH 140	0.0542	0.0130	52	1.3	0.5
ZH 145	0.0480	0.0086	34	0.9	0.3
ZH 150	0.0425	0.0052	21	0.6	0.2
VBF 100	0.0995	0.0808	320	3.4	1.2
VBF 105	0.0933	0.0742	294	3.5	1.3
VBF 110	0.0871	0.0671	265	3.6	1.3
VBF 115	0.0791	0.0579	229	3.4	1.3
VBF 120	0.0717	0.0486	192	3.2	1.2
VBF 125	0.0674	0.0411	163	2.9	1.1
VBF 130	0.0625	0.0329	130	2.5	0.9
VBF 135	0.0577	0.0251	100	2.1	0.7
VBF 140	0.0526	0.0181	72	1.6	0.6
VBF 145	0.0492	0.0126	50	1.2	0.4
VBF 150	0.0457	0.0080	32	0.8	0.3

Table 6: Table of Backgrounds: The number of expected events for each background source after passing the trigger requiements and are accepted by the offline selection cuts for the signal regions. The estimations are calculated for integrated luminosity of 4 fb^{-1}.

	WH/ZH Signal region		VBF Signal region	
	SS	SJ	SS	SJ
$t\bar{t}$	281.7	115.3	177.3	75.7
Single Top S channel	35.5	13.0	7.7	4.0
Single Top T channel	8.5	4.7	9.4	6.0
wbb	25.7	9.8	3.7	2.2
wcc	2.2	2.1	1.1	1.1
Z+jets	127.5	55.4	135.0	62.9
ww	1.2	2.2	0.8	1.2
wz	4.6	1.9	1.7	1.1
zz	5.7	4.4	2.8	1.5
Data	16857	9341	17776	9518

3 Jet Shapes

In the QCD multi-jet background, the jets in the final state can consist of a mixture of quark and gluon jets. However for the Higgs signal that we are searching for, the jets in the final state are mostly quark jets. Furthermore, in the signal final state, a large fraction of the jets that are not from the Higgs boson decay, are light-flavor quark jets. The width of light-flavor quark jets are usually smaller than jets from b-quark and gluon. Thus we make use of the jet width to help us in separating between the Higgs signal and QCD multi-jet background. The two jet shape variables that we use are:

$$\phi\text{-moment}(<\phi>) = \sqrt{\sum_{\text{towers}} \left(\left(\frac{E_t^{\text{tower}} \phi_{\text{tower}}}{E_t^{\text{jet}}} \right)^2 - \phi_{\text{jet}}^2 \right)} \tag{1a}$$

$$\eta\text{-moment}(<\eta>) = \sqrt{\sum_{\text{towers}} \left(\left(\frac{E_t^{\text{tower}} \eta_{\text{tower}}}{E_t^{\text{jet}}} \right)^2 - \eta_{\text{jet}}^2 \right)} \tag{1b}$$

3.1 Comparison of Jet Shape Variables Between Data and Simulation

Before we make use of the jet shape variables in the Higgs search analysis, we need to check if the simulation correctly describes the light-flavor quark jet's η-moment and ϕ-moment distributions. To perform this check we need to obtain a sample that has high purity of quark jets. This sample is obtained from $t\bar{t}$ production and in the lepton+jets decay channel. The light-flavor jets are selected from the hadronic decay of the W boson, which comes from the decay of the Top quark. The data sample to select the $t\bar{t}$ events are the high-Pt electron and muon datasets (from p0 to p21), and the MC samples are generated with Pythia (tt0p75+ot0p49). The $t\bar{t}$ candidates are selected with these cuts :

- 1 CEM electron ($Et > 20$ GeV), or 1 CMUP/CMX muon ($Pt > 20$ GeV)
- ≥ 4 jets ($Et(L7) > 20$ GeV, $\mid \eta \mid < 2.4$
- $MET > 25$ GeV, $Ht > 250$ GeV
- ≥ 1 tight SecVtx tagged b-jet

Similar cuts are used in the Top analysis documented in CDF 9462 [11]. In this note, the fraction of $t\bar{t}$ events with Njet=4,5 is $\sim$ 86%. The plots in figure 4 shows the comparison of the kinematic distributions of the selected $t\bar{t}$ candidates in the data to the $t\bar{t}$ events from simulation. One sees good agreement between data and simulation indicates that the fraction of real $t\bar{t}$ events in the data is high. For the jet shape studies, we select non b-tagged jets that are in the mass window $50 < M(qq) < 110$ GeV/c^2, which is consistent to the mass of the W boson.

The plots comparing the η-moment and ϕ-moment distributions between data and simulation for the $t\bar{t}$ events are shown in figure 5. Although the shapes between data and simulation are quite similar, however there is a small offset between them. We also found that other factors can affect the width of reconstructed jets. The variations of the average value of η-moment and ϕ-moment as a function of the jet's Et, η and the number of reconstructed vertices (NVtx), are shown in figure 6.

We mentioned earlier that $t\bar{t}$ production accounts about 86% of the events selected with NJets=4,5. In CDF note 9462, the remaining 14% events come from Wbb ($\sim$ 4.2%) production, Wcc ($\sim$ 2.3%) production , Wc, mistag, non-W, Z+jets, di-boson and single-top productions. The slight offset between data and MC as seen in Figure 5 could be due to the other non-$t\bar{t}$ sources that we had not included in. Just to see if adding any of the non-$t\bar{t}$ contributions would improve the agreement, we include three different MC sources:

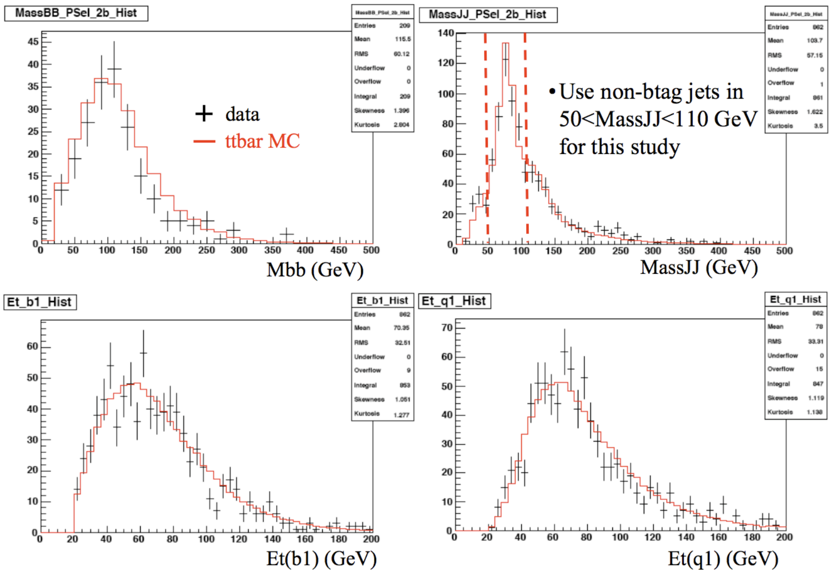


Figure 4: Comparison of the kinematic distributions of the selected $t\bar{t}$ candidates in the data to the $t\bar{t}$ events from simulation.

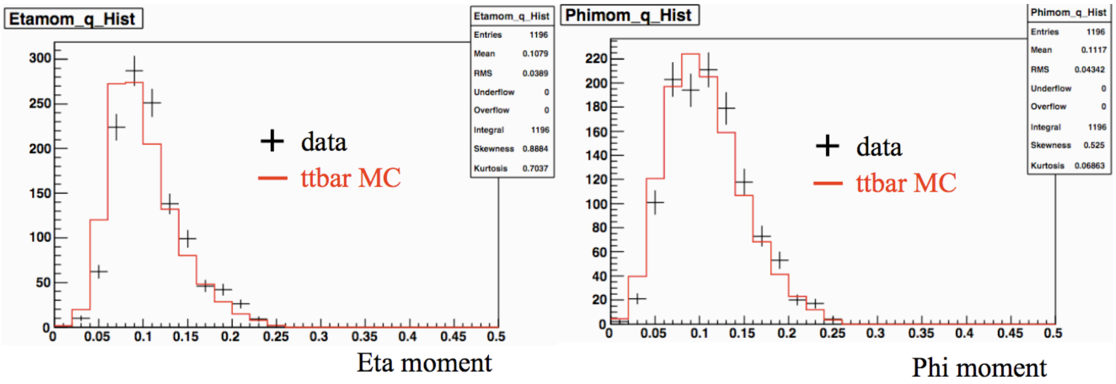


Figure 5: Comparing the η-moment and ϕ-moment distributions between data and simulation for selected $t\bar{t}$ events.

- $t\bar{t}$: 86%
- Wbb : 5.9%
- Wcc : 4.9%

We kept the relative fraction of the Wbb with respect to Wcc to be the same as in CDF note 9462. The plots in Figure 7 show the comparison of the η-moment and ϕ-moment distributions between data and simulation after adding Wbb and Wcc sources. The offset between data and simulation still exists.

3.1.1 Parameterizing the Jet Shape Dependence

We decide to parameterize the η-moment and ϕ-moment dependence on Et, η and the number of reconstructed vertices to remove such dependences. There are two reason to remove these dependencies:

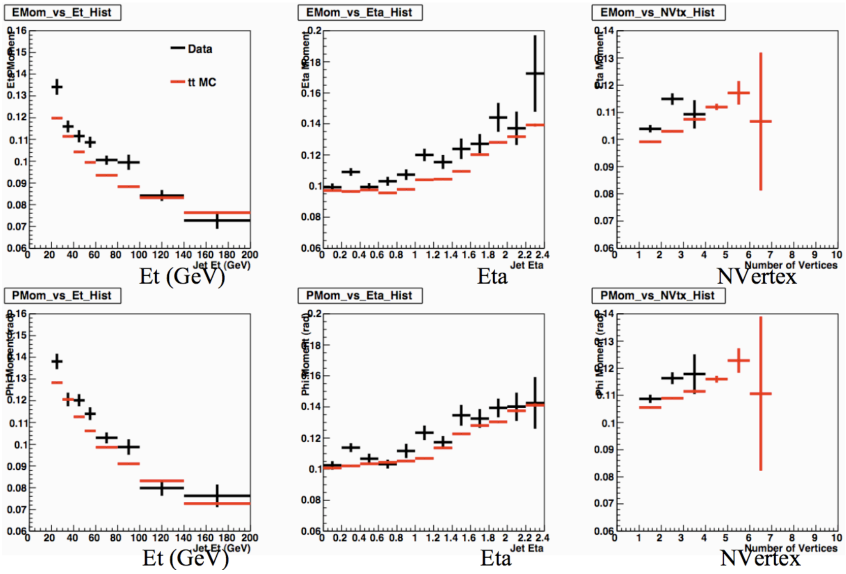


Figure 6: The variations of the average value of η-moment and ϕ-moment as function of the jet's Et, η and the number of reconstructed vertices (NVtx).

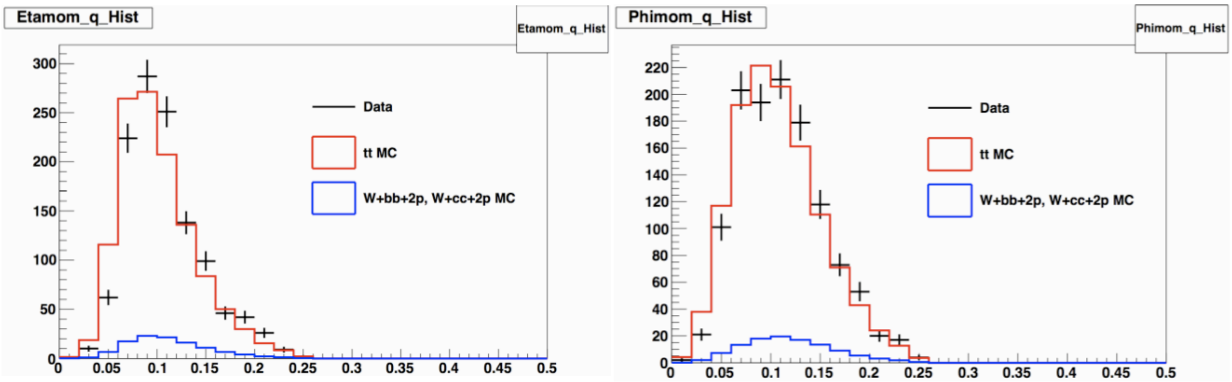


Figure 7: Comparing the η-moment and ϕ-moment distributions between data and simulation for selected $t\bar{t}$ events.

- The variation of η-moment and ϕ-moment as a function of these variables maybe different between data and simulation. Therefore the removal of these dependenices may improve the the agreement between data and simulation.
- The jet shape variables will be used in the Neural Net (section 5) to separate signal from background. If η-moment and ϕ-moment values are correlated with other jet kinematics, and if the correlations are not the same between data and simulation, then the NN training on the signal from the simulation may not be correct compared to a real Higgs signal in the data.

We found that the dependencies of η-moment and ϕ-moment on jet's Et and η are not completely uncorrelated, therefore we performed the parameterizations in several steps:

- First we parameterize the Et and number of reconstructed vertex dependence. To parameterize the Et dependency, we choose jets that are located in the central region ($|\eta| < 1$) from events with $NVtx = 1$. As the statistic are not sufficient to make a good fit, we first parameterize the

Et dependency using central jets from events with $NVtx \geq 1$ (see figure 8 left plots) . We then take the shape of this fit to apply to the central jets from events with $NVtx = 1$. In this case we let the shape float during the fit in order to measure the pesdestal (figure 8 right plots). The parameterization of the Et dependence for simulation is shown in figure 9.

- Next we parameterize the NVtx dependency, as shown in figure 10.
- We then apply these two parameterizations to first remove the Et and NVtx dependencies. Once this is done we then parameterize the η dependency, which are shown in figure 11.

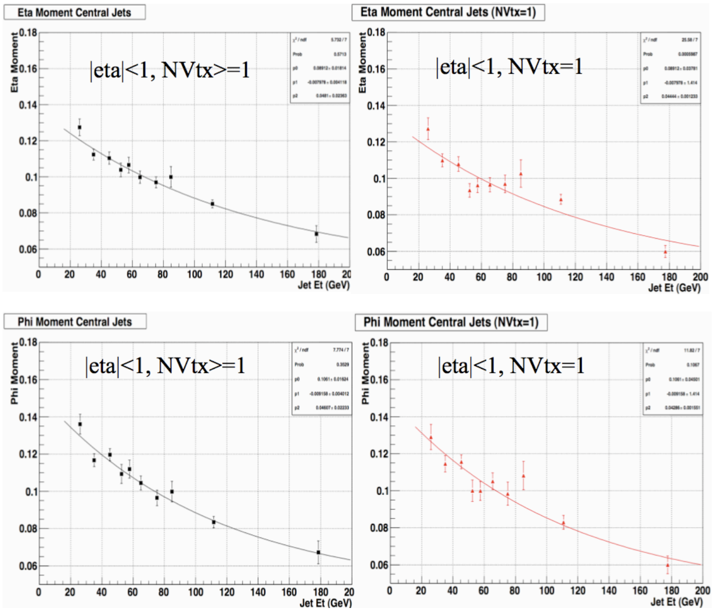


Figure 8: Parameterization of the jet Et dependence for η-moment and ϕ-moment of the non b-tagged jets in the selected $t\bar{t}$ candidates in data.

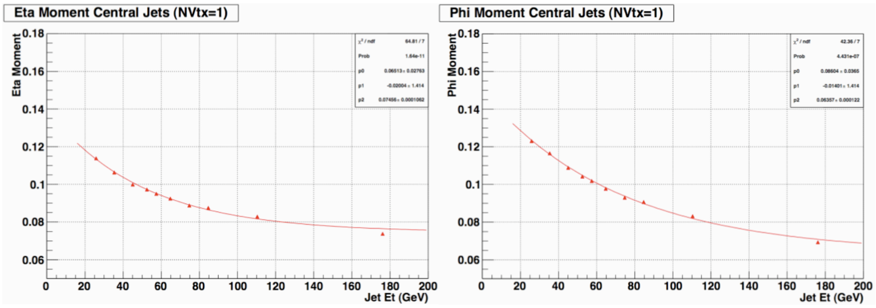


Figure 9: Parameterization of the jet Et dependence for η-moment and ϕ-moment of the non b-tagged jets in the selected $t\bar{t}$ candidates in simulation.

We make use of the parameterizations and equations 2,3 to remove the dependencies of η-moment and ϕ-moment on jet Et, η and NVtx.

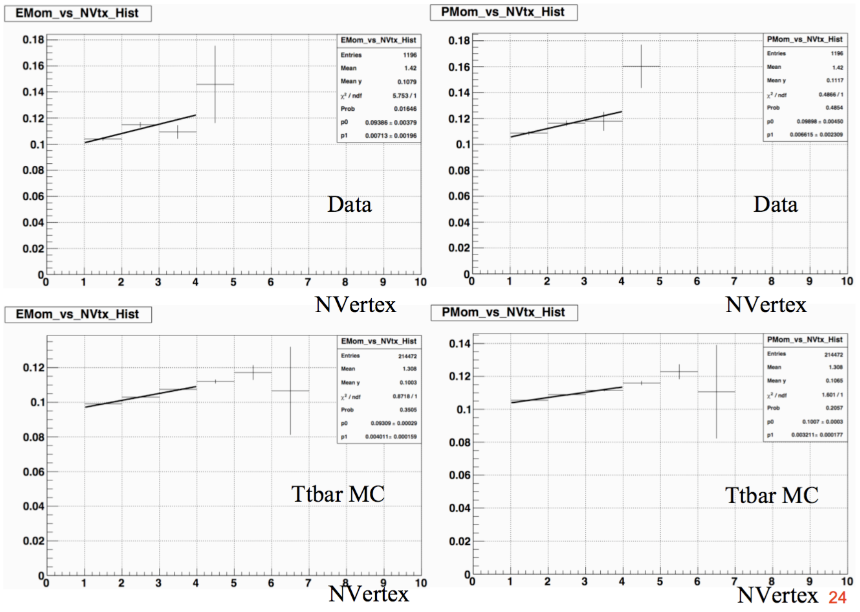


Figure 10: Parameterization of the NVtx dependency of the η-moment and ϕ-moment variables.

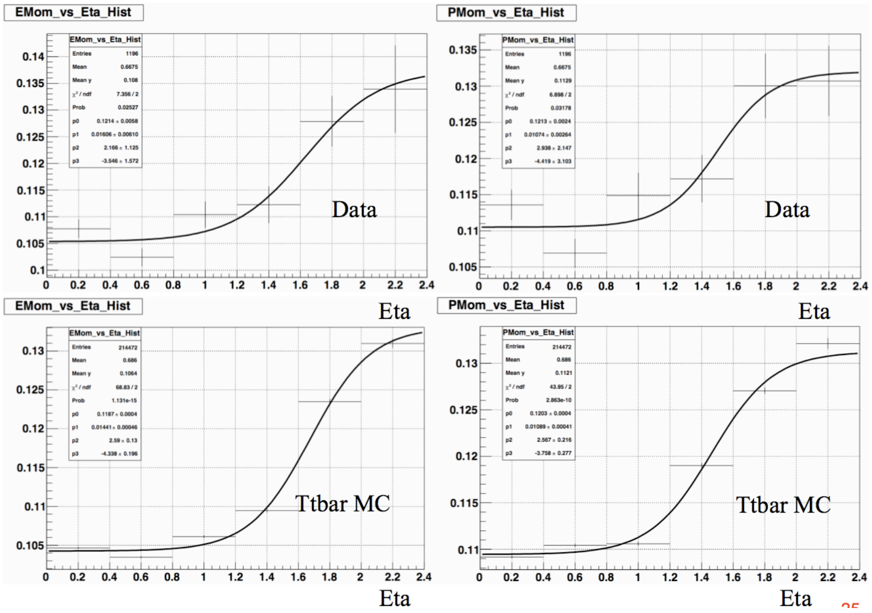


Figure 11: Parameterization of the η dependency of the η-moment and ϕ-moment variables.

$$\begin{aligned} <\eta/\phi>^{data}_{new} = <\eta/\phi>^{data}(x,y,z) \quad & \times \quad (f^{data}_{Et}(Et = 50GeV)/f^{data}_{Et}(Et = x)) \\ & \times \quad (f^{data}_{\eta}(\eta = 0)/f^{data}_{\eta}(\eta = y)) \\ & \times \quad (f^{data}_{NVtx}(NVtx = 1)/f^{data}_{NVtx}(NVtx = z)) \end{aligned} \tag{2}$$

$$\begin{aligned} <\eta/\phi>^{MC}_{new} = <\eta/\phi>^{MC}(x,y,z) & \times (f^{MC}_{Et}(Et=50GeV)/f^{MC}_{Et}(Et=x)) \\ & \times (f^{MC}_{\eta}(\eta=0)/f^{MC}_{\eta}(\eta=y)) \\ & \times (f^{MC}_{NVtx}(NVtx=1)/f^{MC}_{NVtx}(NVtx=z)) \\ & \times (f^{data}_{Et}(Et=50GeV)/f^{MC}_{Et}(Et=50)) \end{aligned} \quad (3)$$

The effects from each correction term in the equations are:

- "$f^{data,MC}_{Et}(Et=50GeV)/f^{data,MC}_{Et}(Et=x)$" corrects the jet's η-moment and ϕ-moment values at $Et=x$ to $Et=50$ GeV.

- "$f^{data,MC}_{\eta}(\eta=0)/f^{data,MC}_{\eta}(\eta=y)$" corrects the jet's η-moment and ϕ-moment values at $\eta=y$ to $\eta=0$.

- "$f^{data,MC}_{NVtx}(NVtx=1)/f^{data,MC}_{NVtx}(NVtx=z)$" corrects the jet's η-moment and ϕ-moment values in events with $NVtx=z$ to events with $NVtx=1$.

- "$f^{data}_{Et}(Et=50GeV)/f^{MC}_{Et}(Et=50))$" corrects the jet's η-moment and ϕ-moment values in MC at $Et=50$ GeV to the values measured in the data at $Et=50$ GeV.

The plots in figure 12 show η-moment and ϕ-moment as function of jet Et, η and NVtx after the dependencies removal. Figure 13 shows the η-moment and ϕ-moment distributions between data and simulation after the dependencies removal. The agreement between data ans simulation is much better now. Next we perform a shift scan to see if we can get an even better agreement between data and simulation. This is done by shifting the η-moment and ϕ-moment distributions in the simulation by some amount and calculate the chi-square between the data's distributions and the simulation's distributions. The plot of the chi-square per number-of-degree freedom is shown in figure 14. This study shows that we need to shift the η-moment (ϕ-moment) distribution in the simulation by $\sim +0.0025$ ($\sim +0.0015$). This additional offset is included into the simulations for this Higgs analysis. The plots of the η-moment and ϕ-moment distributions, after this additional offset, are shown in Figure 15. We use half the values of these additional offsets as the systematic uncertainty on the agreement of the η-moment and ϕ-moment distributions between data and MC. For η-moment (ϕ-moment) the uncertainty is ± 0.00125 (± 0.00075).

3.1.2 Jet Shape Studies Using Z+jets Sample

The jet shapes studies performed in earlier sections are based on selected $t\bar{t}$ candidate events. Unfortunately the statistic of the data sample is limited. To obtain a larger data sample, we turn to data events where jets are produced together with a Z boson. In this case the jets contain a mixture of quark and gluon jets. We assume that the Z+jet Monte Carlo simulation has the right proportion of quark and gluon jets as in the data. The data sample to select the Z+jets events are the high-Pt electron and muon datasets (from p0 to p21), and inclusive Z production events generated with Pythia. The Z boson is identified by selecting a pair of e^+e^- or $\mu^+\mu^-$ and the invariant mass of the pair of leptons lie in the mass window $84 < M(e^+e^-, \mu^+\mu^-) < 98$ GeV/c^2. The jet should have $Et > 20$ GeV and $\mid \eta \mid < 2.4$. The jets' η-moment and ϕ-moment dependencies on jet Et, η and NVtx are removed using the parameterizations measured from the $t\bar{t}$ samples. Figure 16 shows the η-moment and ϕ-moment distributions between data and simulation after the dependencies removal for Z+jets events. The simulation has an additional correction offset of +0.0025 (+0.0015) for the η-moment (ϕ-moment) distribution. The data and simulation are in good agreement.

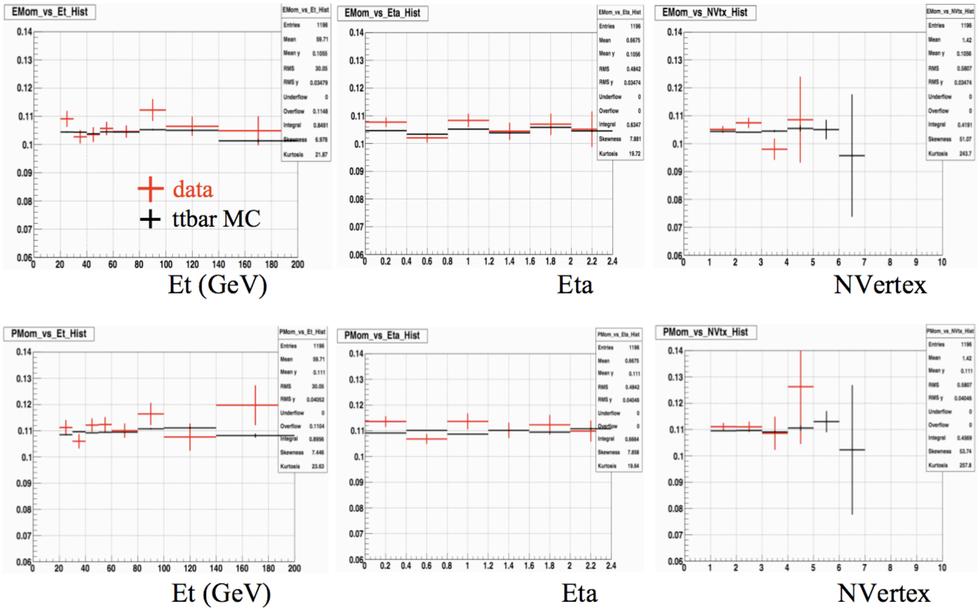


Figure 12: η-moment and ϕ-moment as function of jet Et, η and NVtx after dependencies removal.

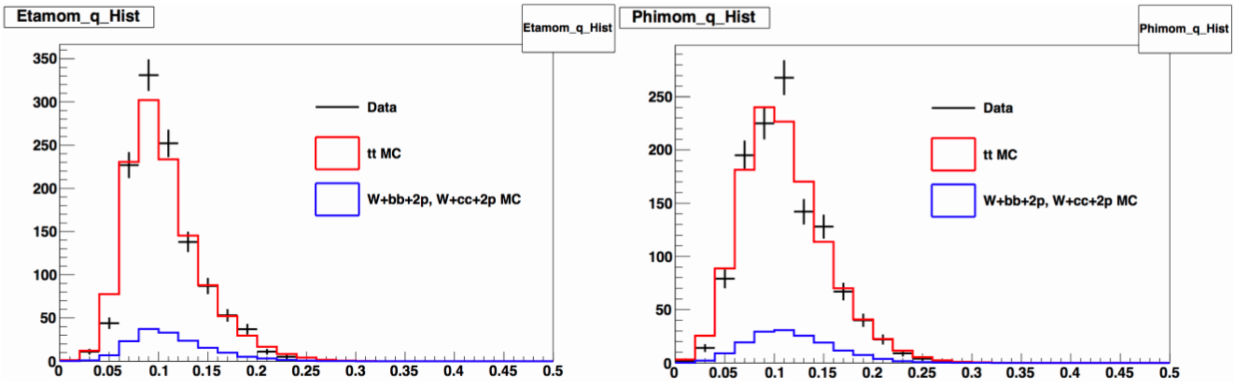


Figure 13: η-moment and ϕ-moment distributions between data and simulation after the dependencies removal.

3.1.3 Comparison of η-moment and ϕ-moment between $t\bar{t}$ Simulation and Higgs Simulations

In the earlier section we show that the jet η-moment and ϕ-moment of the $t\bar{t}$ candidates in the data agree with the $t\bar{t}$ simulation. We then check whether the same set of parameterizations can be applied onto the Higgs simulation samples to remove the jet Et, η and NVtx dependence of the jet η-moment and ϕ-moment. Figure 17 compares $t\bar{t}$ simulation to WH and ZH simualtions, and figure 18 compares the $t\bar{t}$ simulation to VBF simulation. The agreement between $t\bar{t}$ simulation to WH and ZH simualtions are quite good. However the agreement between $t\bar{t}$ simulation to VBF simulation is not well for jets in the central region ($\mid \eta \mid < 1$). Due to this difference between $t\bar{t}$ simulation and VBF simulation we decide to assign an additional uncertainty for VBF. Since for this search we consider jets in the region $\mid \eta \mid < 2.4$, we look at the difference between $t\bar{t}$ simulation and VBF simulation for non-btag jets in $\mid \eta \mid < 2.4$ to assign systematic uncertainty for the jet shapes. For the η-moment, the largest difference between $t\bar{t}$ and VBF is ~ 0.0050. For the ϕ-moment, the largest difference between

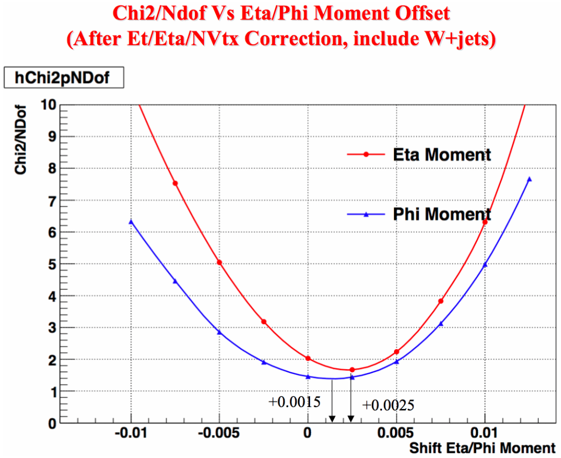


Figure 14: The chi-square per number-of-degree freedom between the data's η-moment and ϕ-moment distributions and the simulation's distributions after the dependencies removal.

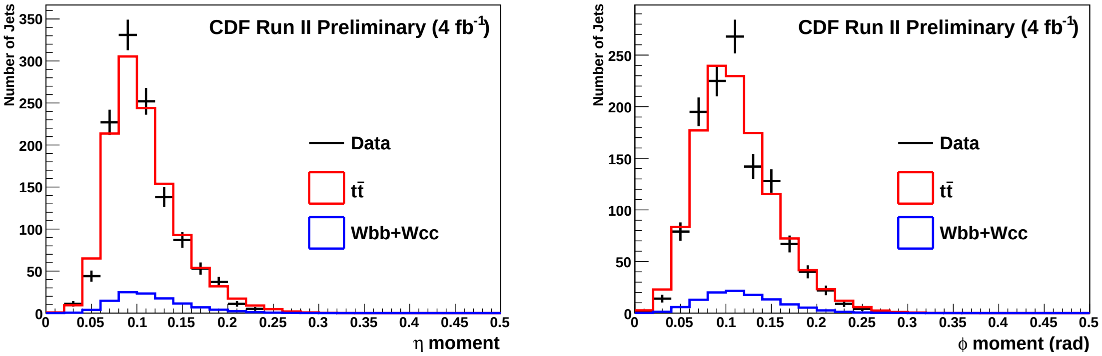


Figure 15: η-moment and ϕ-moment distributions between data and simulation after the dependencies removal. For the simulations, the η-moment (ϕ-moment) is shifted by an additional offset of $+0.0025$ ($+0.0015$) to have the best agreement with the data.

$t\bar{t}$ and VBF is ~ 0.0020. We use half of the differences as the uncertainty. Therefore for the η-moment (ϕ-moment) the additional uncertainty is ± 0.0025 (± 0.0010).

3.1.4 Systematic Uncertinties of the Jet Shape Description in Simulation

For the WH/ZH simulation we assign a systematic uncertainty of ± 0.00125 (± 0.00075) for the η-moment (ϕ-moment). This is the half the values of the additional offsets needed to get the $t\bar{t}$ simulation to best agree with the data, after the $Et/\eta/NVtx$ dependences removal. For VBF, due to the differences between $t\bar{t}$ and VBF in the central jets, we assign an additional systematic uncertainty (as mentioned in earlier section). The total systematic uncertainty on η-moment (ϕ-moment) for VBF is ± 0.0028 (± 0.00125).

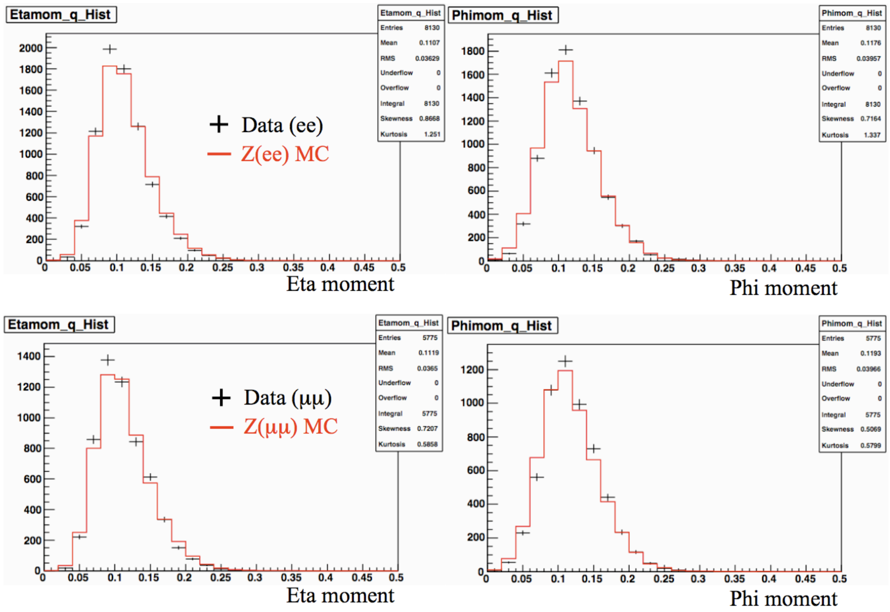


Figure 16: The η-moment and ϕ-moment distributions between data and simulation after the dependencies removal for Z+jets events.

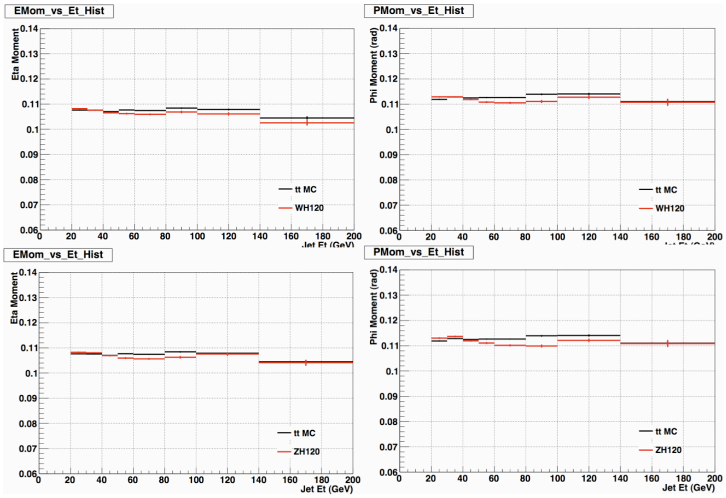


Figure 17: Comparison of the η-moment and ϕ-moment vs jet Et for simulated events betweem $t\bar{t}$ and WH or ZH.

4 Background Prediction: Tag-Rate-Function (TRF)

The critical component to this analysis is an accurate prediction of the QCD background as it is the dominant background. In this analysis a data driven model was devised to predict the two-tagged background from the background-rich one-tagged data. The assumption is the two-tagged background distribution is a scaled replica of the one-tagged distribution (figure 19). The scale factor which reduces the one-tag data is called the Tag Rate Function (TRF). The TRF is the probability of a jet being b-tagged in the event that already has one other jet tagged as a b-jet. The probability is measured in a kinematic region that has very little contribution from the Higgs signal. This measured

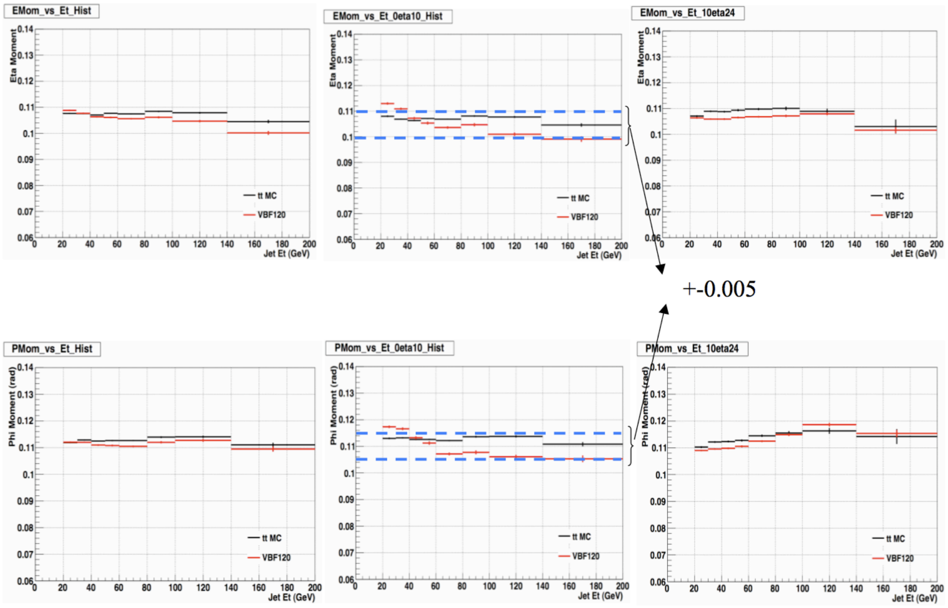
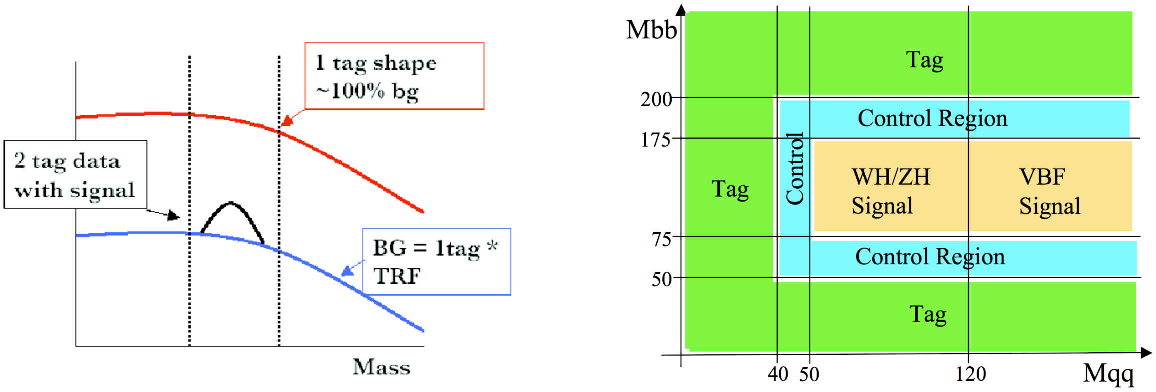


Figure 18: Compare the η-moment and ϕ-moment vs jet Et for simulated events betweem $t\bar{t}$ and VBF. (LEFT) jets in $0 <| \eta |< 2.4$, (MIDDLE) jets in $0 <| \eta |< 1.0$, (RIGHT) jets in $1.0 <| \eta |< 2.4$,



(a) TRF Principal: The red-line corresponds to the 1-tag distribution which is ≈100% background. The 2-tag background is assumed to be a scaled version of the 1-tag (background) distribution (blue line). The TRF is derived from the regions outside the signal peak.

(b) Mbb-Mqq plane to define the TAG and CONTROL regions. The default TRF is derived from the TAG -region (yellow). This is applied to the 1-Tag data in the signal region to predict the 2-Tag background. As a systematic, another TRF is derived from the CONTROL (CTRL) region. This is also applied to the 1-Tag data in the signal region to give an alternative background predicition. The difference of these two background predictions is applied as a systemtic error.

Figure 19: The Tag-Rate function and the m_{bb} - m_{qq} plane

probability is applied onto the one tagged events in the signal region to predict the double b-tagged QCD background. The key issue of this method is to make sure that the technique can correctly predict the shapes of the kinematic distributions of the double b-tagged QCD multi-jet events which will be used later in the NN training to separate the Higgs signal from the QCD background. This TRF method does not necessary predict the right normalization of the double b-tagged QCD background.

For each 1-Tagged event, one considers each permutation of the Tag jet (the b-tagged-jet) and the Probe jet. A probe jet is a taggable jet [2]. The TRF is parameterised as a function of three parameters which are:

- $\Delta\eta$ of Tag-probe jet pair ($\Delta\eta_{bb}$)
- P_T of the probe jet (E_T)
- pseudorapidity of the probe jet ($|\eta|$)

$$TRF(\Delta\eta_{bb}, E_t, |\eta|) = \frac{\text{Number of } 2-\text{Tag events}(\Delta\eta_{bb}, E_t, |\eta|)}{\text{Number of events with } \geq \text{ 1 tight SecVtx tagged jet}(\Delta\eta_{bb}, E_t, |\eta|)} \quad (4)$$

The TRFs are measured separately for SS and SJ double b-tagged categories. For the SJ category, events with only one tight SecVtx tagged jets are being considered in the measurement of the TRF(SJ).

The data used to derive the TRFs come from examining data outside the signal region (figure 19) in the M(bb)-M(qq) plane. This is defined by regions outside the $75 < M(bb) < 175$ GeV/c^2 and $M(qq) > 50$ GeV/c^2 mass window. Two regions are defined:

- **Tag**
- **Control**

The default TRF uses data from the TAG region. The additional CONTROL (CTRL) region is used to derive systematic errors (figure 19).

The TRF is applied to the one tagged events in signal region to predict the double b-tagged QCD background. However the one tagged events do contain contributions from non-QCD background (i.e. $t\bar{t}$, W plus heavy-flavor jets...). Since we are using simulation to predict the double b-tagged non-QCD background, there will be some double counting of the non-QCD events when we apply the TRF onto the one tagged events. To remove this double counting of the non-QCD background events, we first subtract the one tagged non-QCD background contributions and then add the double b-tagged non-QCD background predicted from simulation. Since $t\bar{t}$ and Z plus jets contributions to the non-QCD background are much larger than the other non-QCD background, we only perform the one tagged subtraction for $t\bar{t}$ and Z plus jets backgrounds.

$$\begin{aligned} \text{QCD Double Tagged Background} \quad &= \quad TRF \times (\text{one tagged data}) \\ &\quad - TRF \times (\text{one tagged } t\bar{t}) \\ &\quad - TRF \times (\text{one tagged } Z \ + \ \text{jets}) \end{aligned} \quad (5)$$

The kinematic distributions of the predicted double b-tagged events in the signal region are compared to the observed double b-tagged events for the SS category. These comparison plots are shown in figure 20-24. The shapes from the prediction agree quite well with the observed shapes. The plots in figure 20-22 are for the events in the WH/ZH signal region, and the plots in figure 24 are for events in the VBF signal region. The definitions for the variables shown in these plots are given in Section 5.

---

[2] Jet-Et > 15 GeV, $|\eta| < 2.4$, Number of good SecVtx (JetProb) tracks ≥ 2 when considering if the jet is SecVtx (JetProb) tagged

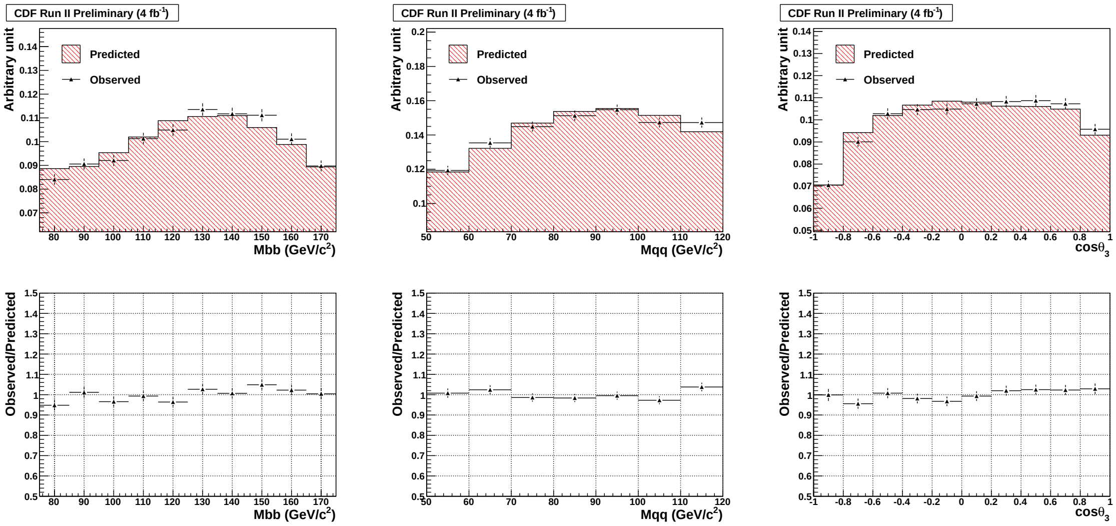


Figure 20: The kinematic distributions of the predicted double b-tagged events in the WH/ZH signal region are compared to the observed double b-tagged events for the SS category. The red hashed histograms are the predicted double b-tagged events, and the black points are the observed double b-tagged events.

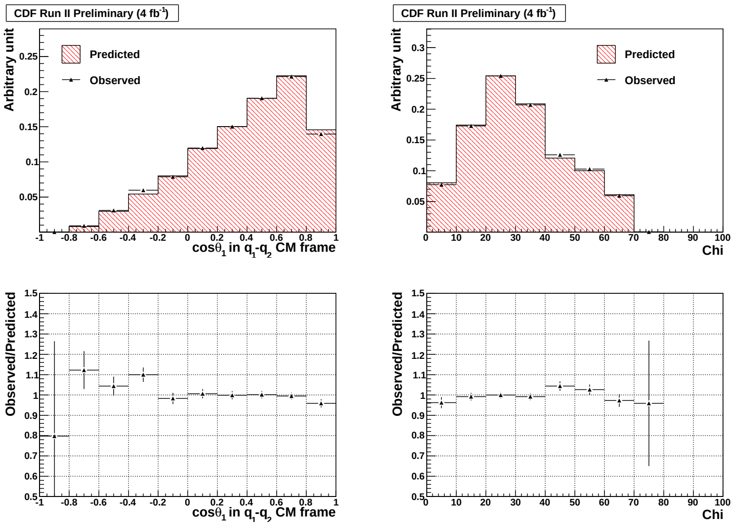


Figure 21: The kinematic distributions of the predicted double b-tagged events in the WH/ZH signal region are compared to the observed double b-tagged events for the SS category. The red hashed histograms are the predicted double b-tagged events, and the black points are the observed double b-tagged events.

4.1 Tunning the Modeling of the Mass $M(qq)$ and Jet Shape Variables

Initially the TRF does not predict very well the shapes of the mass $M(qq)$ and jet shape variables (η moment and ϕ moment). This is because these variables are not included in the parameterization of the TRF. Figure 25 show the predicted shapes of the mass $M(qq)$ and jet shape variables using the TRF, which do not match well to the observed shapes. To correct for this, we measure the correction

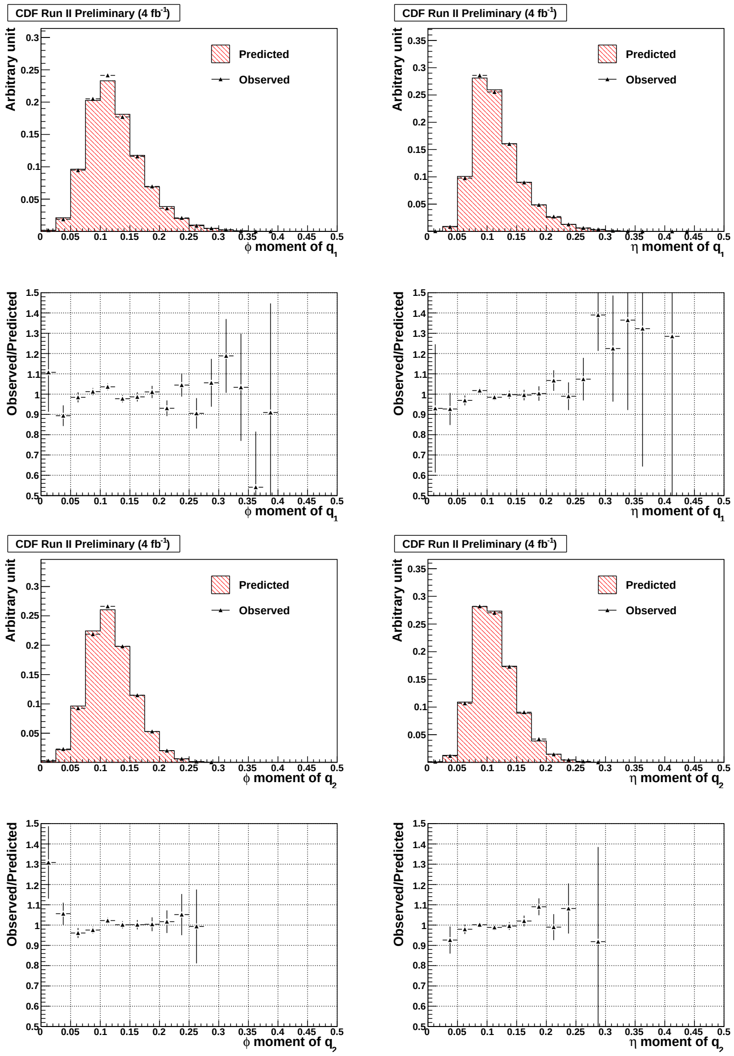


Figure 22: The kinematic distributions of the predicted double b-tagged events in the WH/ZH signal region are compared to the observed double b-tagged events for the SS category. The red hashed histograms are the predicted double b-tagged events, and the black points are the observed double b-tagged events.

functions by applying the TRF to the one tagged events in the TAG region and measure the ratio of the predicted double b-tagged events to the observed b-tagged events as a function of the mass $M(qq)$ and η moment. We do not measure a separate correction function for ϕ moment because the η moment variable is highly correlated with the ϕ moment. These correction functions are then applied in the signal region when we are predicting the double b-tagged events in the signal region. As a systematic check, we also measured a new set of correction functions in the CONTROL region and apply these new correction functions in the signal region. The difference in the predictions between using the correction functions mesured from the TAG region and from the CONTROL region is the uncertainties of tuning the modeling of the mass $M(qq)$ and the jet shape variables.

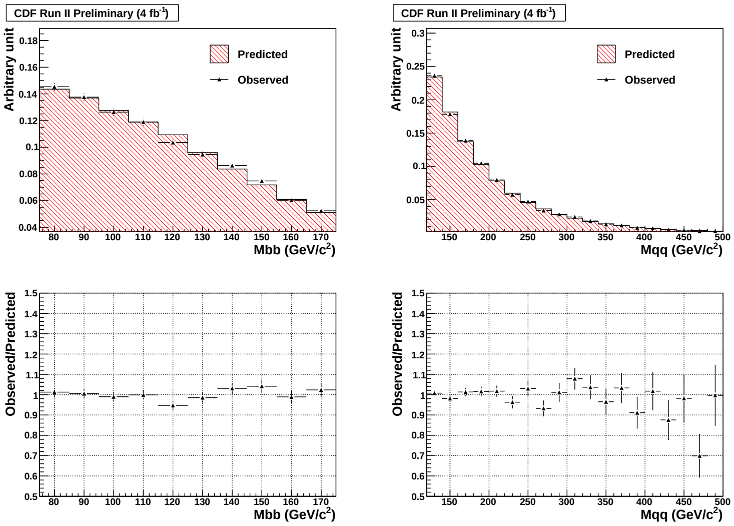


Figure 23: The kinematic distributions of the predicted double b-tagged events in the VBF signal region are compared to the observed double b-tagged events for the SS category. The red hashed histograms are the predicted double b-tagged events, and the black points are the observed double b-tagged events.

4.2 Influence of the Higgs Signal to the TRF

If a Higgs signal exists, it would contribute to the TRF. At the predicted Higgs cross-section, the contribution of the Higgs signal to the TRF would be minimal. However when the limit is calculated, the cross-section is inflated and so the influence of the Higgs signal must be considered. Using the same prescription used by the non-QCD sources, the ($n\times$ Higgs signal) and background is:

$$\begin{aligned}(n \times \text{Higgs Signal}) + \text{Background} &= n(\text{2 Tag Higgs data}) \\ &\quad +TRF \times (\text{1-Tag Data - } n \text{ 1-Tag Higgs}) \\ &= n(\text{2 Tag Higgs - } TRF\times \text{ 1-Tag Higgs}) \\ &\quad +(\text{TRF 1-Tag Data}) \qquad (6)\end{aligned}$$

The Higgs signal is reduced by its 1-Tag contribution.

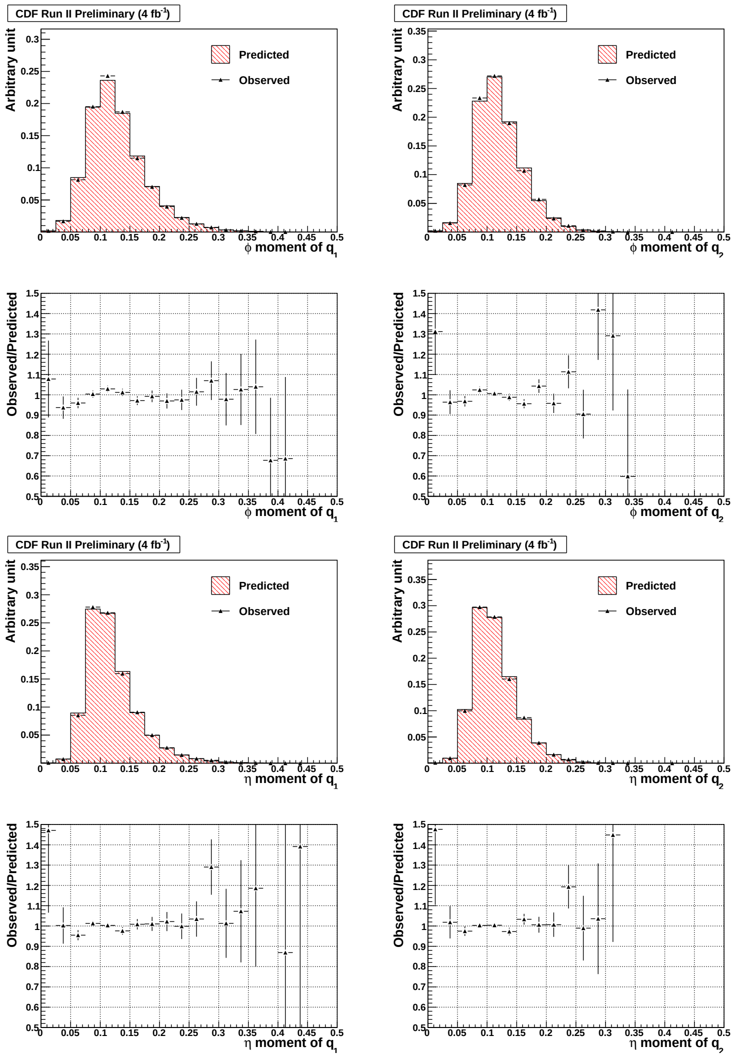


Figure 24: The kinematic distributions of the predicted double b-tagged events in the VBF signal region are compared to the observed double b-tagged events for the SS category. The red hashed histograms are the predicted double b-tagged events, and the black points are the observed double b-tagged events.

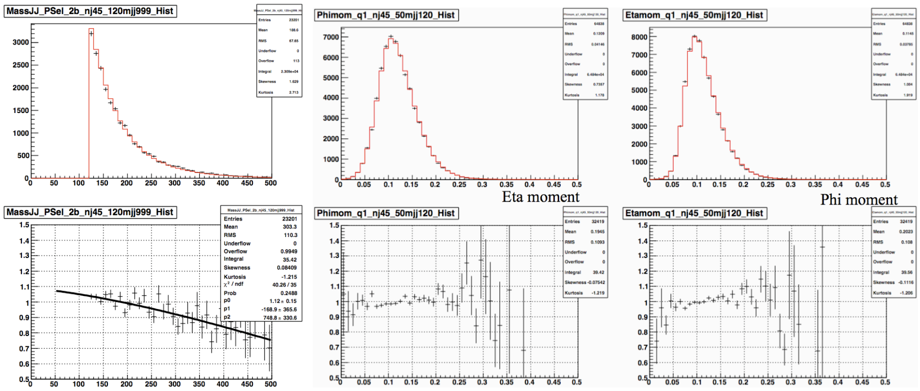


Figure 25: The predicted shapes of the (LEFT) mass $M(qq)$, jet shape variables (MIDDLE) η moment and (RIGHT) ϕ moment using the TRF in the signal region, before applying the correction functions measured in the TAG region. The red histograms are the predicted double b-tagged events, and the black crosses are the observed double b-tagged events.

5 Neural Network Training

For the Higgs analysis, a multivariate discriminant has the ability to combine the information from several variables. This improves the ability to separate a Higgs signal from background events far greater than a standard cuts analysis. The TMVA package [12] allows one to evaluate several multivariate classifiers. For this analysis, we considered a Log-Likelighood, Boosted Decision Tree & an Artificial Neural Network [3]. After training with different variables and examining the signal efficiency/background rejection, expected limit & overtraining results, the Neural Net performed best & was selected as the classifier for this analysis. The settings for the neural work were:

- Neuron Type : tanh
- Number of Training Cycles : 500
- Number of Hidden Layers : 1 Hidden layer with N-inputs+1 nodes.
- 13,000/26,000 training & testing events are used for signal & background for MH100,MH140/MH120.

As the background is dominated by QCD, the 1-Tag background, weighted by the TRF, is used as the background sample for the NN. As we wish to keep the NN training, testing and final analysis events separate, the samples were divided as follows:

- 10% of the 1-Tag background sample from Signal region used only for training & testing Neural Net.
- The remaining 90% of the 1-Tag background sample from the Signal region was only used for the analysis.
- A dedicated Higgs signal sample for NN training & testing
- A dedicated Higgs signal sample for the analysis which is statistically independent to the NN training/testing samples.
- The two tagged events are only ever used in the analysis.

The Neural Net was trained at three target Higgs masses; 100 GeV, 120 Gev and 140 GeV. These three trained neural nets were used to search for a Higgs boson between 100 GeV to 150 GeV. For each mass point, the closest trained neural net was used as follows:

- 100,105,110 GeV Higgs used Neural Net trained on 100 GeV Higgs sample
- 115,120,125,130 GeV Higgs used Neural Net trained on 120 GeV Higgs sample
- 135,140,145,150 GeV Higgs used Neural Net trained on 140 GeV Higgs sample

The Neural Nets were trained using SS data as the kinematics are not affected by the different b-tagging categories (figure 26)and it has a larger $S/\sqrt{B}$ compared to SJ data.

5.1 VH Neural Net Training

For the VH neural net training, the samples used were:

- Signal : the WH and ZH training samples are summed together.
- Background: the 10% 1-Tag background

---

[3] We followed TMVAs recommendation of the Multi-layer Perceptron algorithm for the artificial neural network

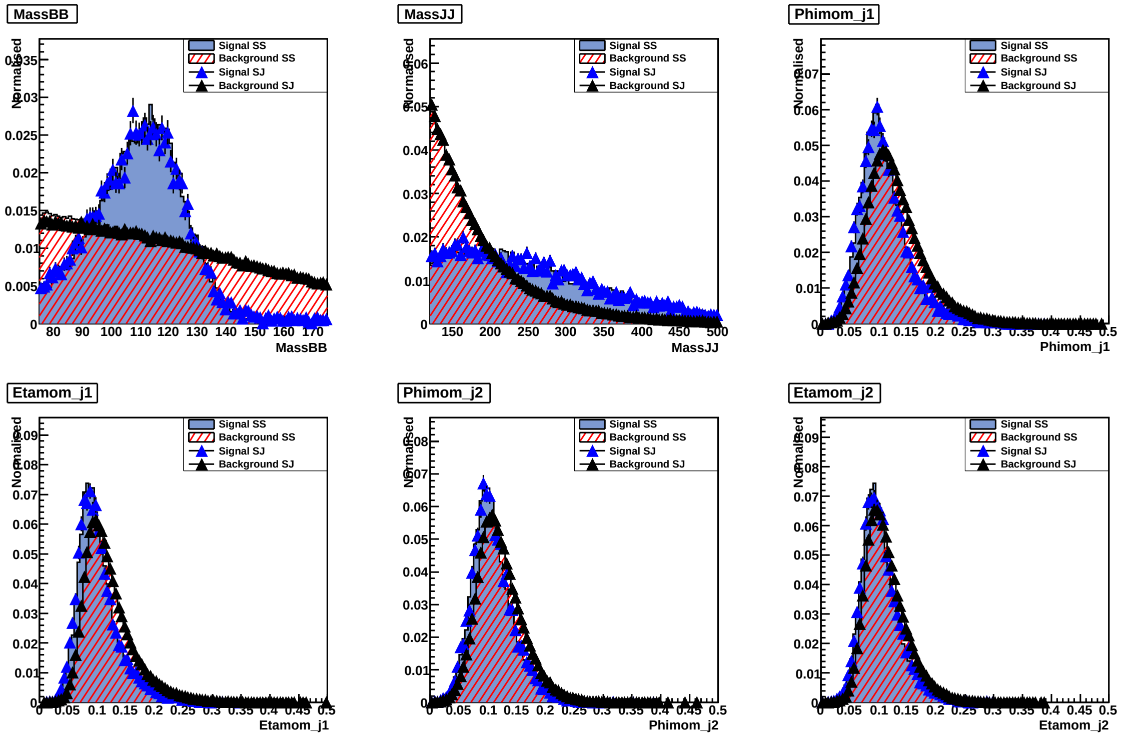


Figure 26: Comparisons of kinematics of SS to SJ. The different b-tagging categories does not alter any of the kinematic distributions. Thus a neural net trained on SS samples is suitable for SJ. In this figure, the plots are for VBF. Analogous plots for VH also show the kinetmatics do not depend on the b-tagging category.

For both samples, the VH signal window cut was applied:

$$\begin{aligned} \text{VH Signal Window} \quad : \quad & 75 < M(b\bar{b}) < 175\text{GeV} \\ & 50 < M(q\bar{q}) < 120\text{GeV} \end{aligned} \tag{7}$$

The selection of variables for the neural net training must fulfil two criteria:

- the variable must give good background-signal separation.
- The variable must be well modelled by the Tag-Rate-Function (TRF).

A initial list of approximately 40 variables was drawn and were judged by the two criteria. After this first pass, the number of suitable variables reduced to 20. The next step was to examine which variables help the Neural Net to separate the signal from background. An initial neural net trained with just M_{bb} and M_{jj} (the Higgs & W/Z mass resonance). Then an additional variable was added from the reduced list until all 20 variables were added. At each stage, the signal efficiency for different background rejection rates and the expected (statistics-only) limit was recorded. The final list of training variables for the VH Neural Net are:

- Mass of the two b-tagged jets ($M(bb)$)
- Mass of the two non b-tagged jets ($M(qq)$)
- The helicity angle of the leading non b-jet ($\cos(q_1 cm)$)

- cosine of the leading-jet scattering angle in the 4-jet rest-frame ($\cos(\theta_3)$)
- Chi
- Eta-Moment of leading non b-jet ($< \eta > j_1$)
- Phi-Moment of leading non b-jet ($< \phi > j_1$)
- Eta-Moment of second leading non b-jet ($< \eta > j_2$)
- Phi-Moment of second leading non b-jet ($< \phi > j_2$)

$\cos(q_1 cm)$ is the cosine helicity angle of the leading non b-jet (q_1). The helicity angle $q_1 cm$ of the leading non b-jet q_1 is defined to be the angle between the mementum of q_1 in the $q_1 - q_2$ rest frame and the total momentum of $q_1 - q_2$ in the lab frame.

The definition for $\cos(\theta_3)$ can be found in [13]

The *Chi* variable is defined as:

$$Chi = \mathrm{Min}(Chi_W, Chi_Z) \tag{8a}$$

$$Chi_{W/Z} = \sqrt{(M_{W/Z} - M_{qq})^2 + (M_H - M_{bb})^2} \tag{8b}$$

The jet-moments were defined in equation 1.

Figure 27 shows the signal & background plots for the selected variables and the modelling by the Tag-Rate-Function (TRF) are shown in figures 28 and 29.

The neural net is trained with the selected 9 variables with the settings given in section 5. For any multivariate analysis, one must ensure their classifier is not overtrained. Otherwise it would lead to an overoptimistic performance of the classifier. A measure of the over-training is to compare the Neural Net distribution using events used for training & events from an independent test sample. The TMVA package provides such a test and reports the Kolmargorov-Smirnoff probabilties for the signal & background. For the VH training, the values are 99.4% for the background and 31.1% for a MH120 VH signal (figure 33(a)). The Neural Net distributions for a VH (MH120) signal is shown in figure 47.

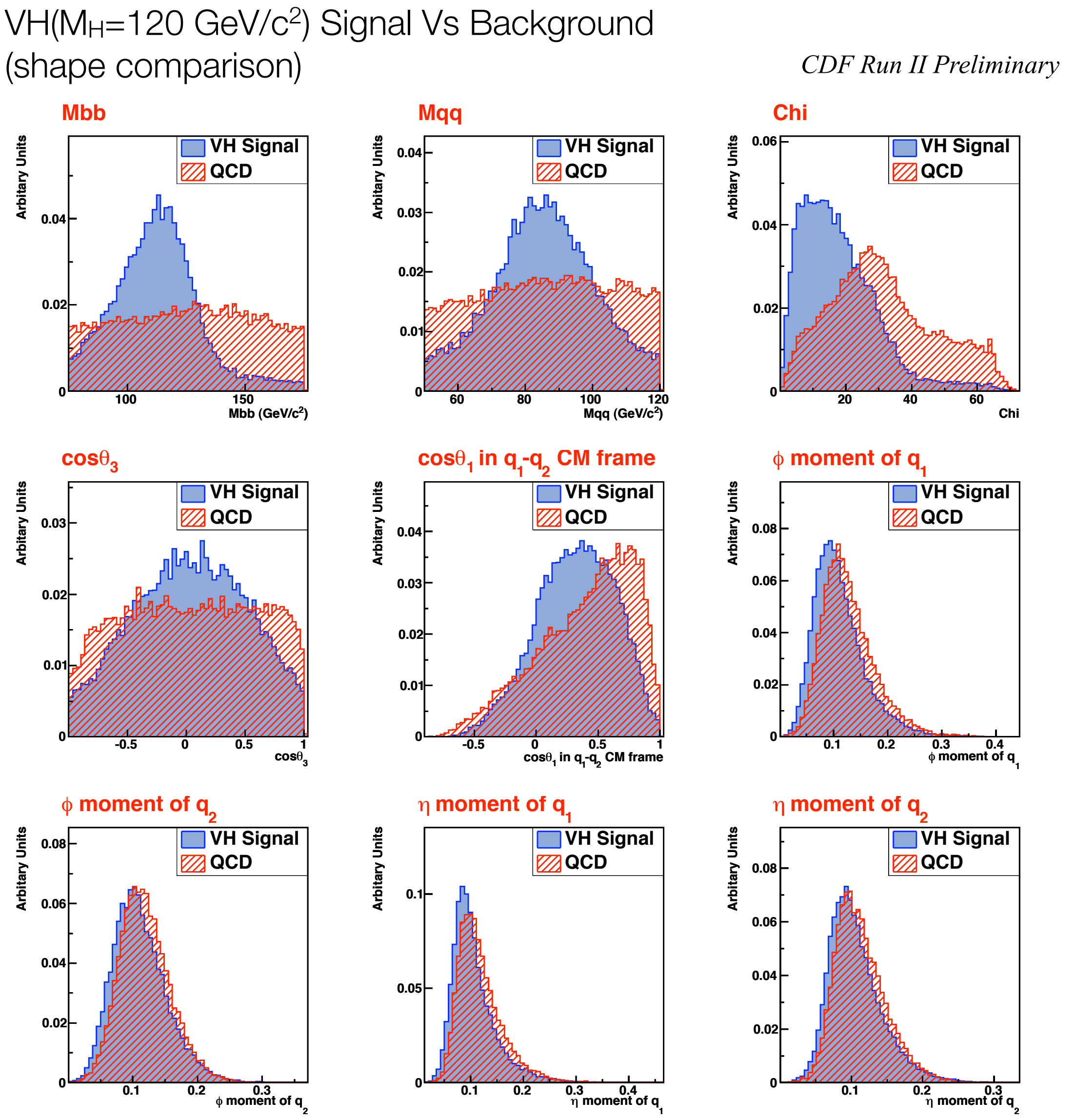


Figure 27: The training variables used for the VH analysis. VH(MH120) is used as the signal.

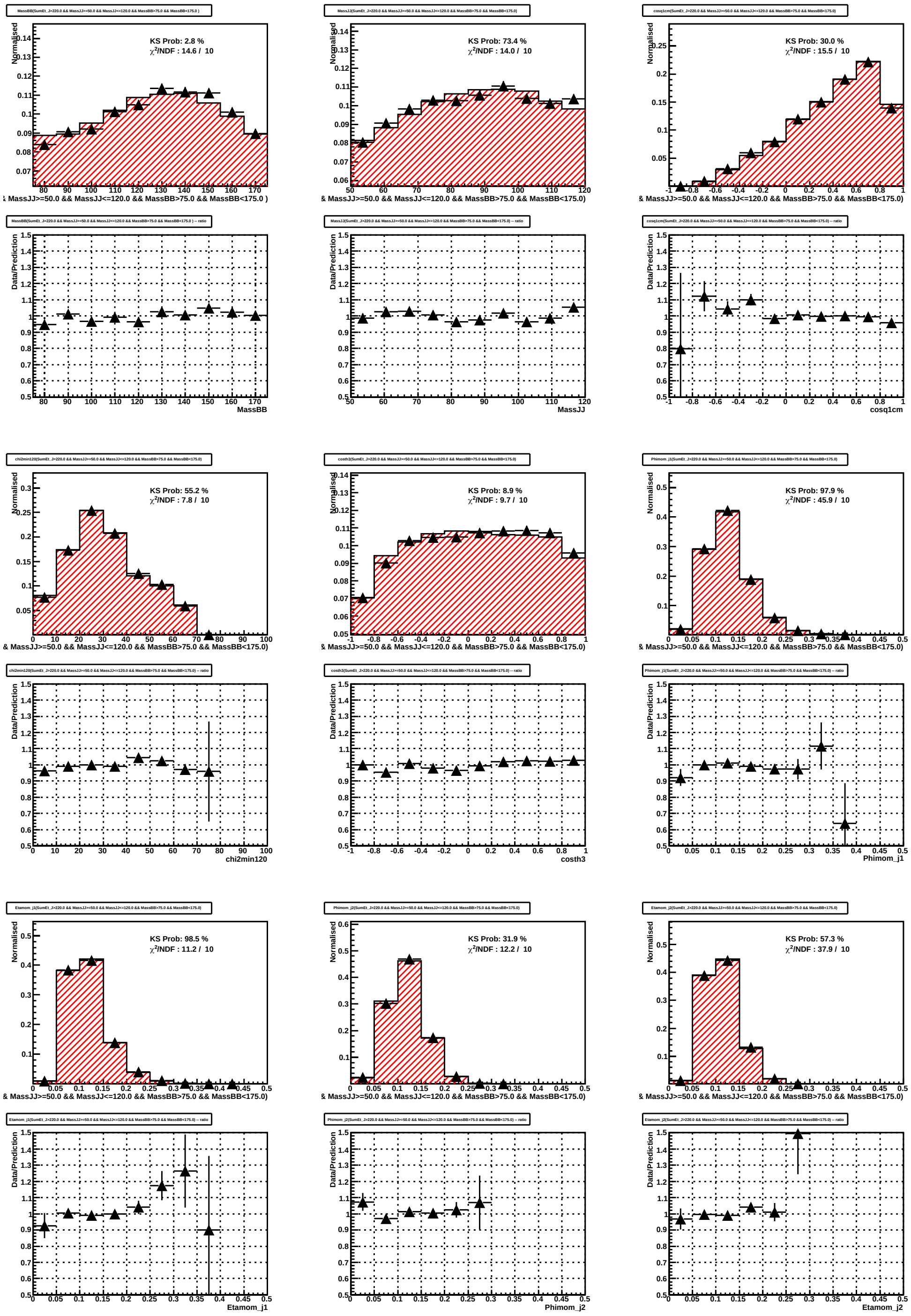


Figure 28: The SS TRF prediction of the training variables for the VH channel. The red shaded plots are the TRF prediction and the black points are the data for the VH signal region. As the TRF only predicts the shape, the histograms are normalised to unit area. Below the histograms are the ratio plots of Data/prediciton. Variables which are modelled by the TRF have a flat ratio.

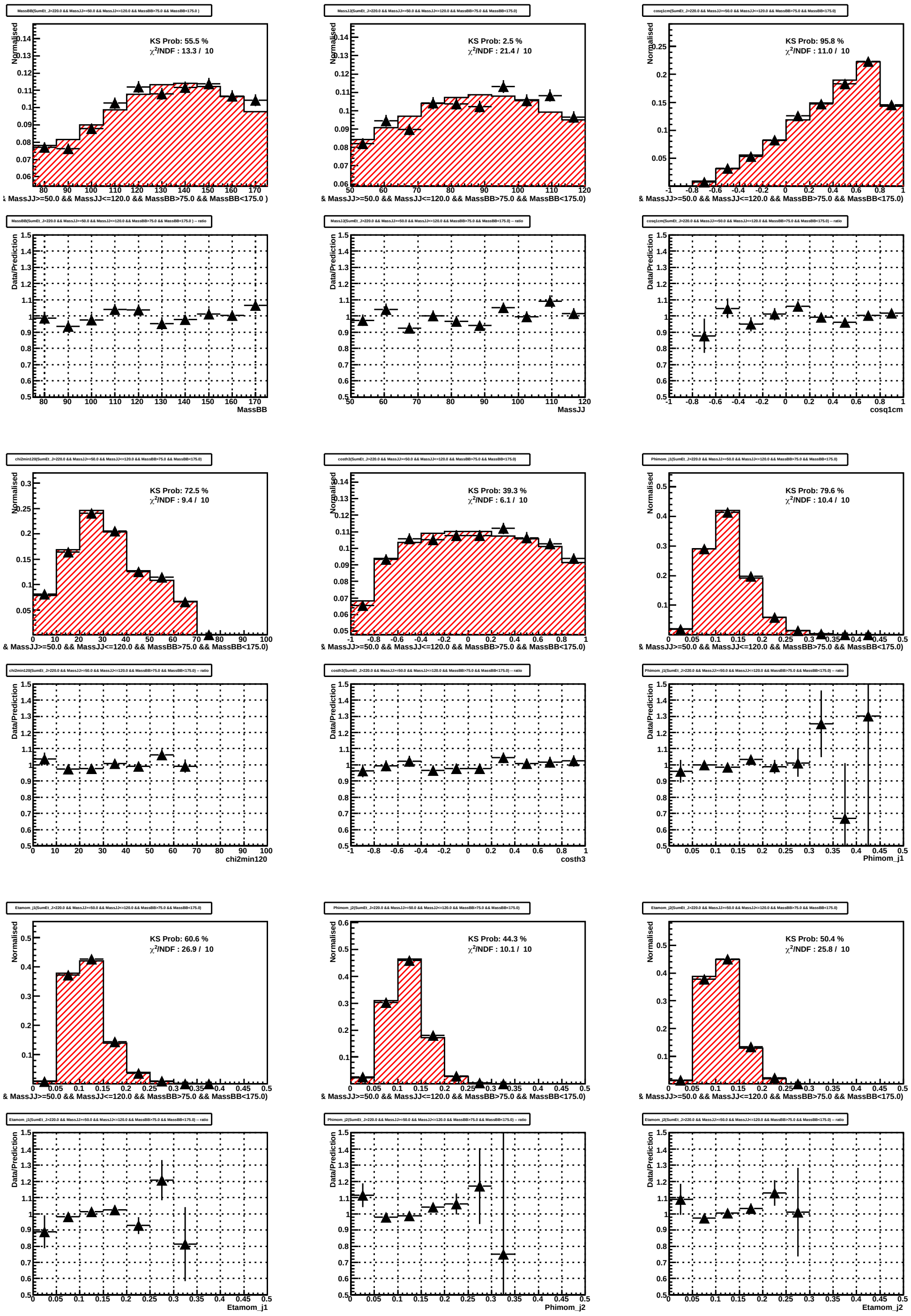


Figure 29: The SJ TRF prediction of the training variables for the VH channel. The red shaded plots are the TRF prediction and the black points are the data for the VH signal region. As the TRF only predicts the shape, the histograms are normalised to unit area. Below the histograms are the ratio plots of Data/prediciton. Variables which are modelled by the TRF have a flat ratio.

5.2 VBF Neural Net Training

The VBF neural net is trained with the following samples:

- Signal : dedicated VBF training sample
- Background: the 10% 1-Tag background

For both samples, the VBF signal window cut was applied:

$$\begin{aligned}\text{VBF Signal Window} \quad : \quad & 75 < M(b\bar{b}) < 175\text{GeV} \\ & M(q\bar{q}) > 120\text{GeV}\end{aligned} \tag{9}$$

Due to the different underlying physics process of the Vector Boson Fusion (VBF) compared to VH channel, we expect a different set of variables will be optimal for the VBF channel. The same variable selection procedure for VH was adopted for VBF. This resulted in 6 training variables for VBF:

- The invariant mass of the two b-tagged jets M_{bb},
- the invariant mass of the two non b-tagged jets M_{qq},
- the four jet shape (moment) variables of the two non b-tagged jets: $\langle\eta_1\rangle$, $\langle\eta_2\rangle$, $\langle\phi_1\rangle$ and $\langle\phi_2\rangle$

Fig. 30 shows the signal and background plots for the selected variables and Figs. 31 and 32 shows the TRF modelling of the NN training variables and Fig. 33(b) shows the overtraining check .

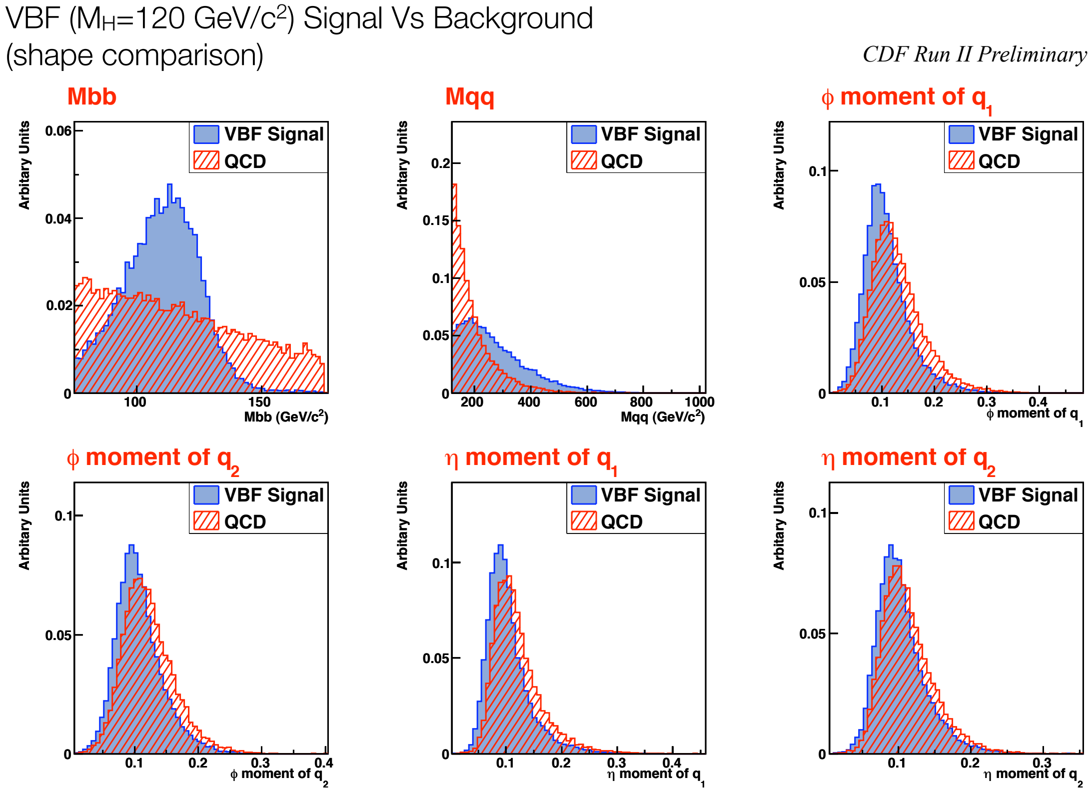


Figure 30: The training variables used for the VBF analysis. VBF(MH120) is used as the signal.

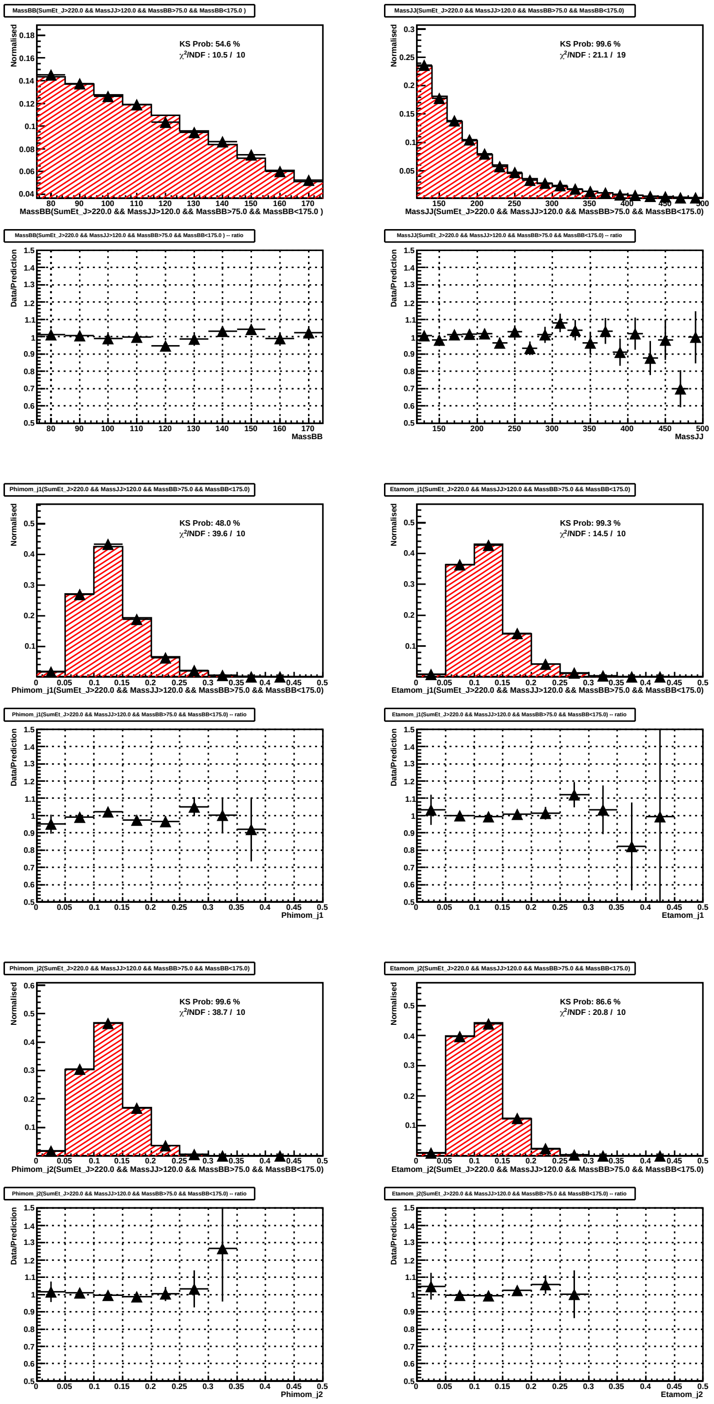


Figure 31: Training variables for the SS VBF channel. Each variable consists of two plots. The top plots are the overlay of the 2-tag data (black points) and TRF-predicted 2-tag background (Red shaded histograms) and the bottom ones are the ratio of the 2-tag data divided by the TRF-predicted 2-tag background. One can see that all the TRF-predicted 2-tag distributions are in good agreement with the observed 2-tag data.

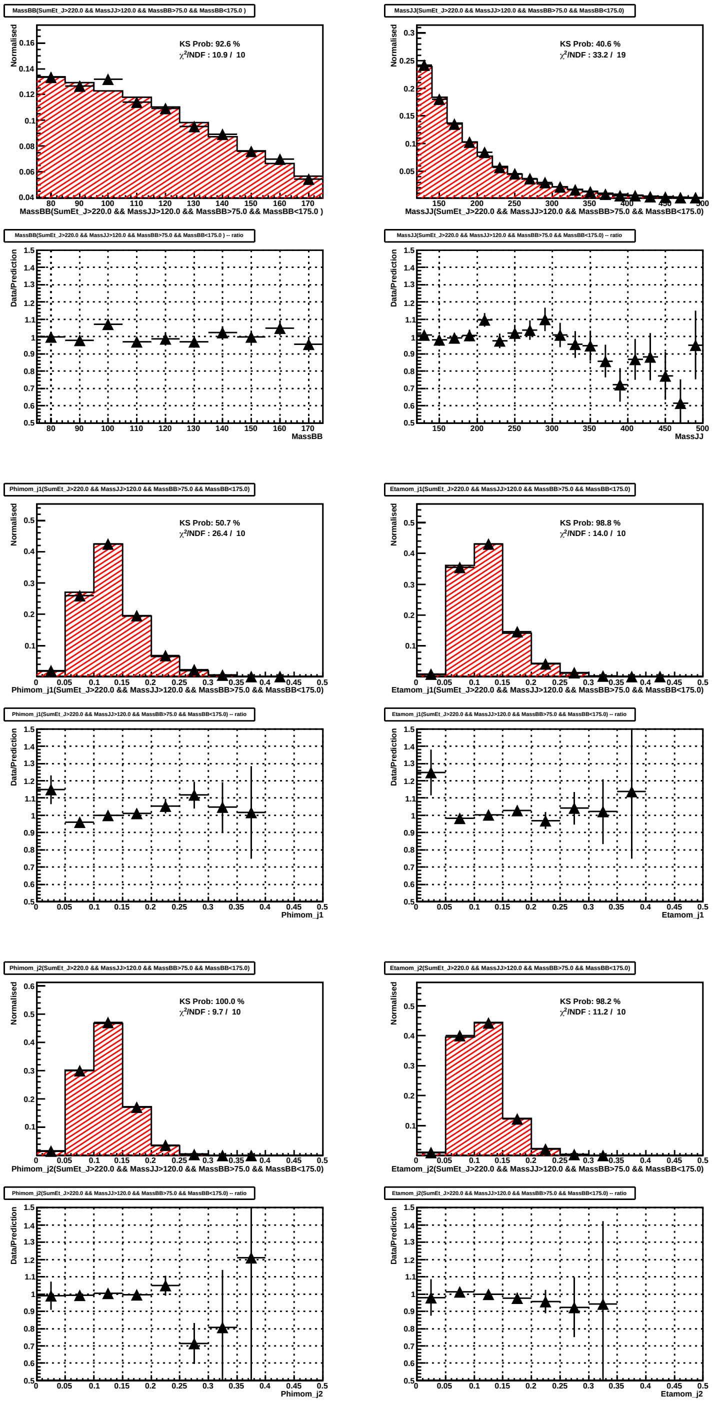
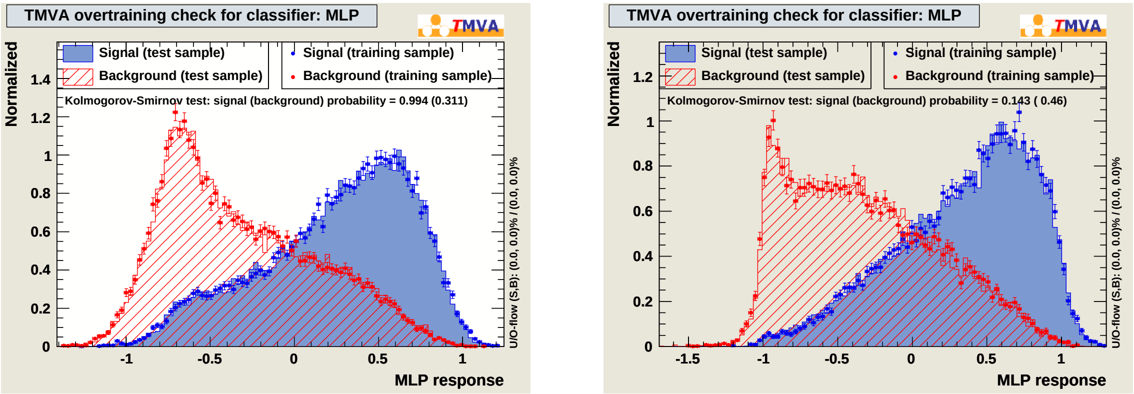


Figure 32: Training variables for the SJ VBF channel. Each variable consists of two plots. The top plots are the overlay of the 2-tag data (black points) and TRF-predicted 2-tag background (Red shaded histograms) and the bottom ones are the ratio of the 2-tag data divided by the TRF-predicted 2-tag background. One can see that all the TRF-predicted 2-tag distributions are in good agreement with the observed 2-tag data.



(a) Overtraining check for VH

(b) Overtraining check for VBF

Figure 33: Overtraining checks for VH & VBF trained for MH120. Both trained neural nets give good Kolmorogov-Smirnoff probabilities which indicates no overtraining.

6 Systematics

6.1 TRF Systematics

For the TRF predicited backgrounds, three sources of systematic error were considered:

- Interpolation Uncertainty
- Mass $M(qq)$ and Jet Shape Tuning Uncertainties
- Mis-Modeling Uncertainty

All three systematics were taken as shape systematics. Normalisation errors were not considered as the background prediction will be floated to best fit the data.

6.1.1 TRF Interpolation Uncertainty

The nominal background predicition used the TRF derived from the Tag regions (TRF(Tag)) which is interpolated into the signal region. The systematic error for this interpolation is taken from the background prediction using the TRF(Ctrl). The difference of these two background shapes is used as the extrapolation shape systematic. The purpose of choosing to measure the TRF in the Ctrl region because the kinematic in this region is much closer to the signal region compared to the Tag region. Figure 34 shows the Neural Net shapes using the nominal TRF(Tag) and TRF(Ctrl).

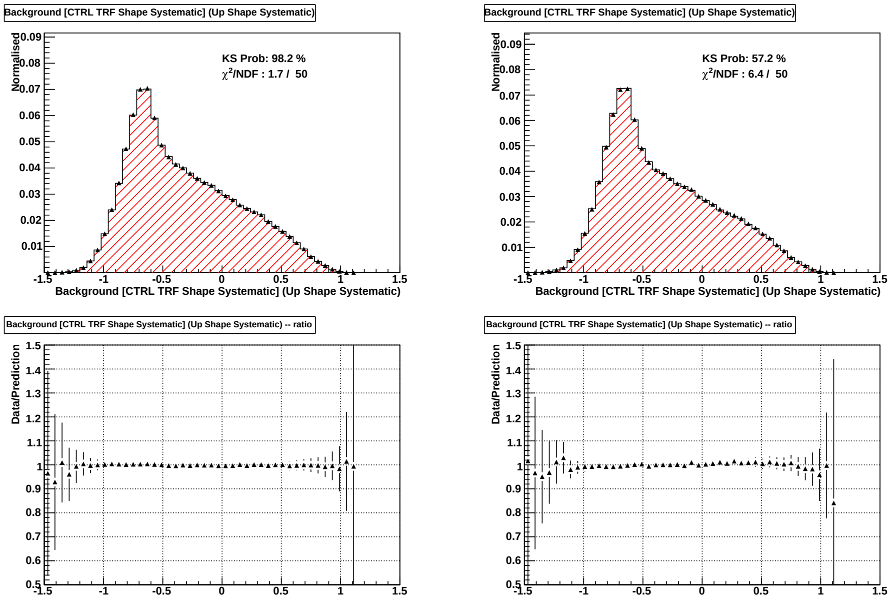
6.1.2 Mass $M(qq)$ and Jet Shape Tuning Uncertainties

In section 4.1 it was shown that the predicted mass $M(qq)$ and jet shape (η moment and ϕ moment) distributions are tuned using correction functions measured in the Tag region. To estimate the uncertainties due to these tuning corrections, we measured another set of correction functions in the Ctrl region. The difference in the NN output due to the different sets of correction functions are then taken as the effect of the uncertainties from these tuning corrections. Figures 35 and 36 show the background predictions using the nominal tunings derived from the Tag region and the alternative tunings derived from the Ctrl region. The ratio plots shows the background shape are not affected by the different tunings.

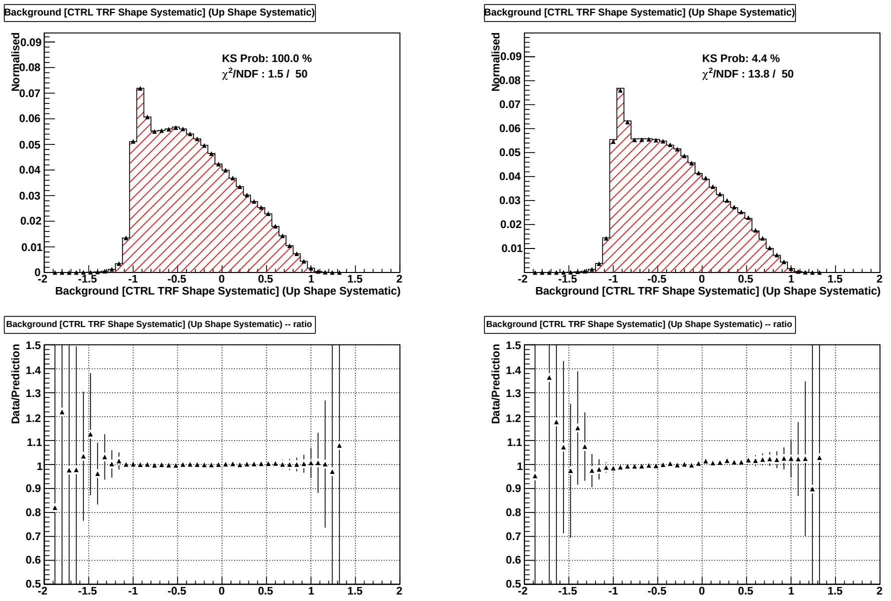
6.1.3 Mis-Modeling Uncertainty

We investigated the uncertainty of the QCD background prediction due to the limitation of the three-dimensional TRF parameterization to describe all the quantities that affect the NN shape. To perform this study we selected a kinematic region that has very little contribution from the Higgs signal and yet also closely resembles the QCD multi-jet background in the signal region. The kinematic region we selected are events with at least six jets. In this jet multiplicity region, we measured a TRF (also parameterized with respect to the three parameters mentioned earlier) in the mass regions $M(bb) > 50$ GeV/c^2 and $M(qq) > 40$ GeV/c^2. We then use this TRF to predict the double b-tagged events in the same signal region in the $M(bb)$ vs $M(qq)$ mass plane as the analysis, except with $NJet \geq 6$. The difference between the observed double b-tagged and the predicted double b-tagged is taken as the mis-modeling uncertainty using the TRF method.

Figure 37 shows the TRF predictions and the 2-Tag events from this $NJet \geq 6$ bin. Figure 38 shows the ratio of the 2-Tag events with respect to the TRF predictions for the combination of the SS (SecVtx-SecVtx) and SJ (SecVtx-JetProb) channels in the $NJet \geq 6$ bin. We fit a flat (slope= 0) line to the combination and the χ^2 per degree of freedom suggests that the distribution is consistent with no systematic shape and the fluctuations are mostly due to statistics. Therefore we drop the the mis-modeling uncertainty from the systematics.

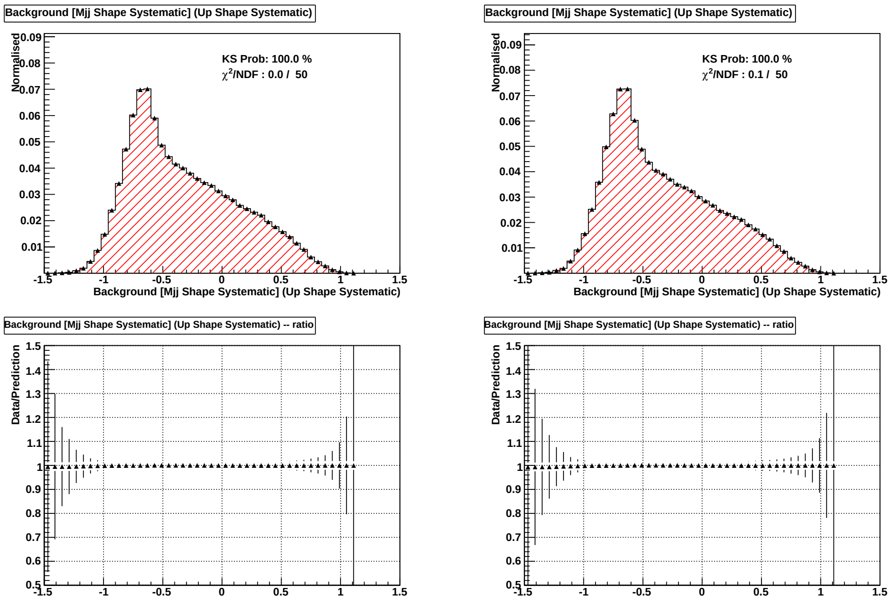


(a) VH SS (b) VH SJ

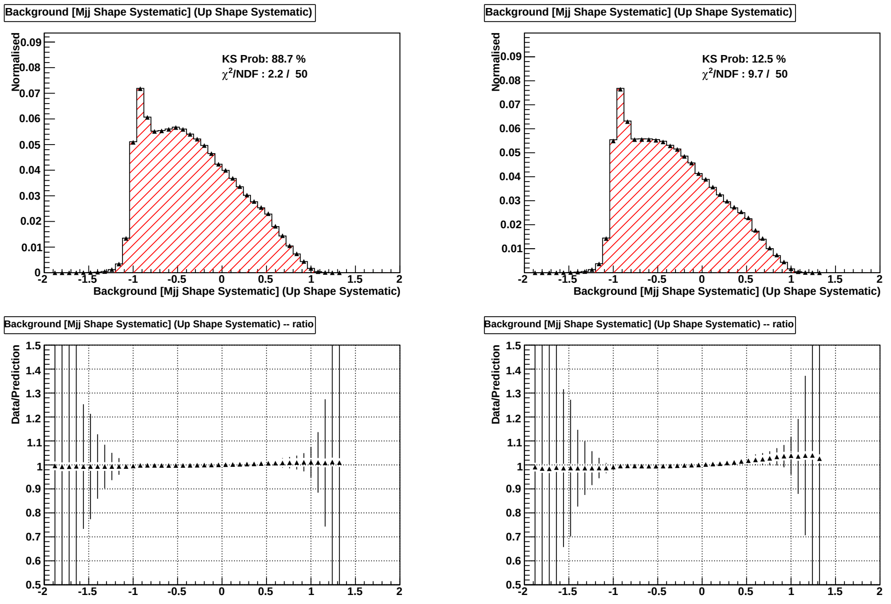


(c) VBF SS (d) VBF SJ

Figure 34: TRF Interpolation Systematics: Each plot compares the background prediction using the nominal TRF derived from the TAG region (shaded red) and the TRF derived from the CTRL region (black points). As the TRF only predicts shapes, both predictions have been normalised to unit area. Below each histogram is a ratio of the CTRL TRF prediction / nominal TRF prediction. The ratio of the two shapes is flat which implies a small shape systematic.

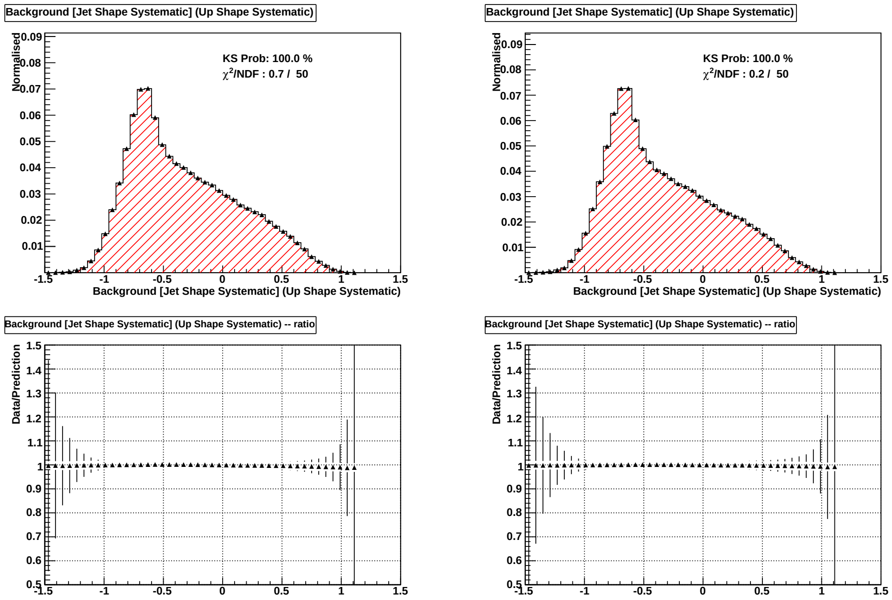


(a) VH SS (b) VH SJ

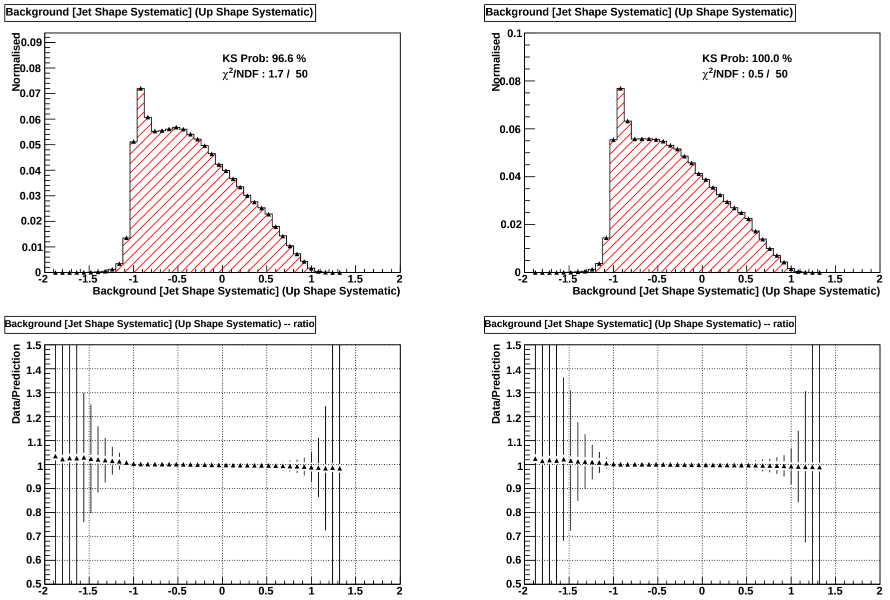


(c) VBF SS (d) VBF SJ

Figure 35: $M(qq)$ Tuning : The background predictions using the nominal $M(qq)$ tuning (red shaded) and the alternative tuning derived from the CTRL region (black points). As the TRF only predicts shapes, both predictions have been normalised to unit area. The ratio plots of CTRL region $M(qq)$ tuning / TAG region $M(qq)$ tuning are drawn below the histograms. The flat ratio shows little change in the background shape between the two $M(qq)$ tunings.

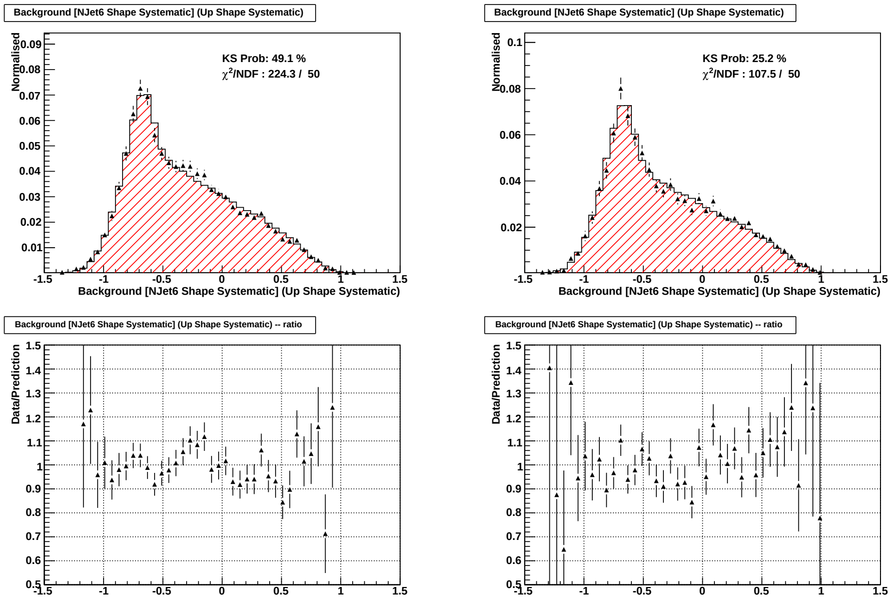


(a) VH SS (b) VH SJ

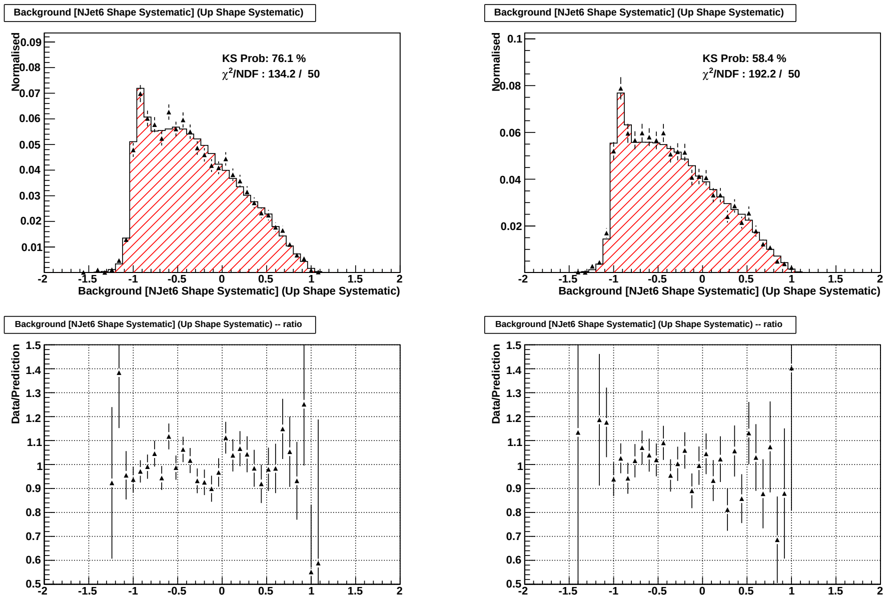


(c) VBF SS (d) VBF SJ

Figure 36: Jet-Shape Tuning : The background predictions using the nominal Jet-Shape tuning and the alternative tuning derived from the CTRL region. As the TRF only predicts shapes, both predictions have been normalised to unit area. The ratio plots of CTRL region Jet-Shape tuning / TAG region Jet-Shape tuning are drawn below the histograms. The flat ratio shows little change in the background shape between the two Jet-Shape tunings.

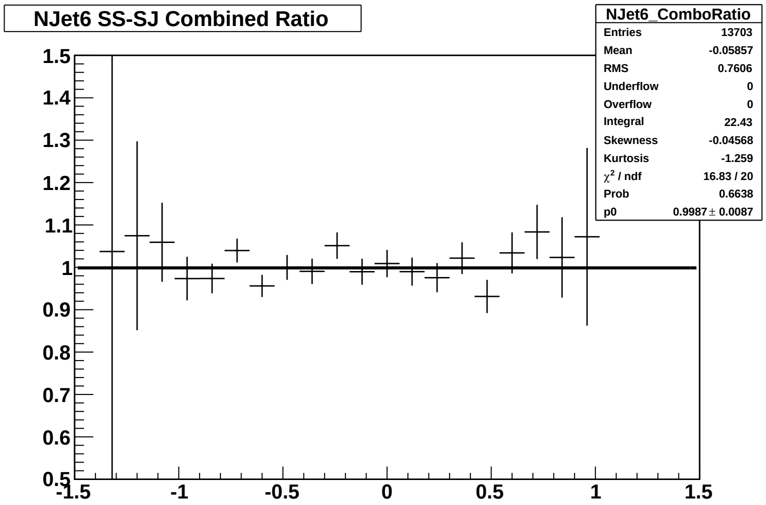


(a) VH SS (b) VH SJ

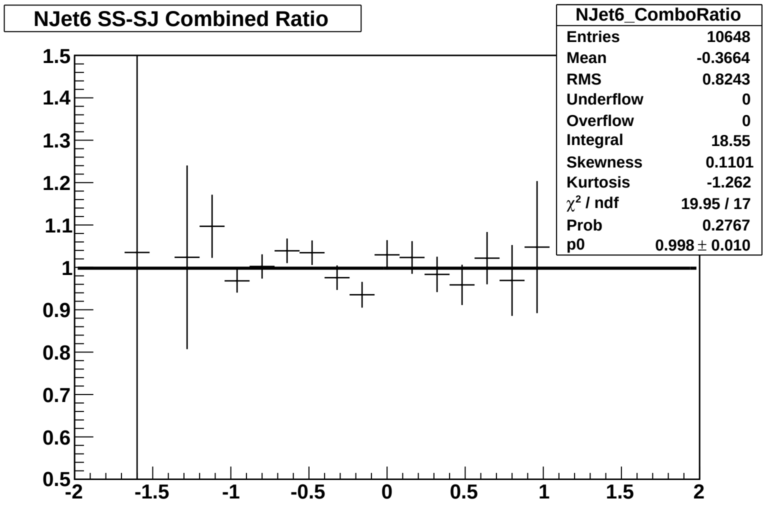


(c) VBF SS (d) VBF SJ

Figure 37: Mis-Modelling Systematic: The background prediction using a TRF derived from the $NJet \geq 6$ control region. It is compared to the 2-Tag events in the $NJet \geq 6$ Jets bin. As the TRF only predicts shapes, both predictions have been normalised to unit area. The variation observed for this systematic is far greater than any other systematic and is the dominant systematic error for this analysis.



(a) VH



(b) VBF

Figure 38: Combining SS and SJ channels for the mis-model uncertainty. A fit of a flat line to the combination has been done for both the VH and the VBF channels. The χ^2 per degree of freedom (16.8/20 for VH and 20/17 for VBF, both ~ 1) show that the distribution is consistent with the case of no systematic shapes. The fluctuation seen in the figure is due to statistics only.

6.2 Signal Systematics

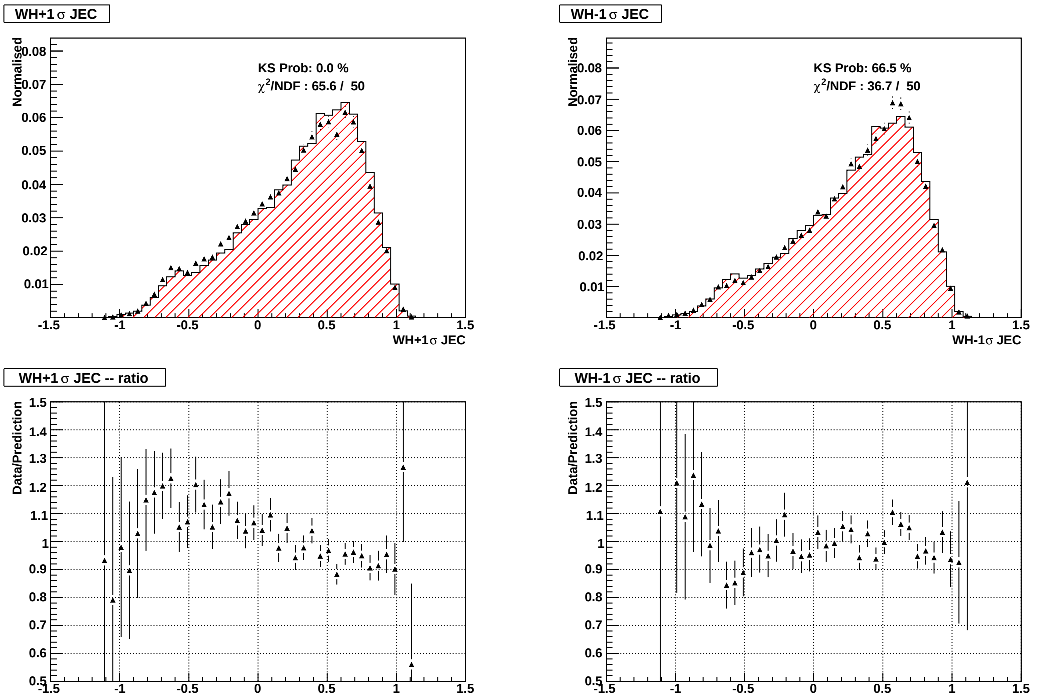
6.2.1 Jet Energy Scale

The Jet Energy Scale affects the Jet E_t related quantities. The training variables which are affected most are the $M(bb)$ and $M(qq)$. The effect upon the Neural Net distribution is shown in figures 39 and 40. Although the Jet Energy Scale systematics affects the $M(bb)$ & $M(qq)$ distribution, the Neural Net shape is less affected.

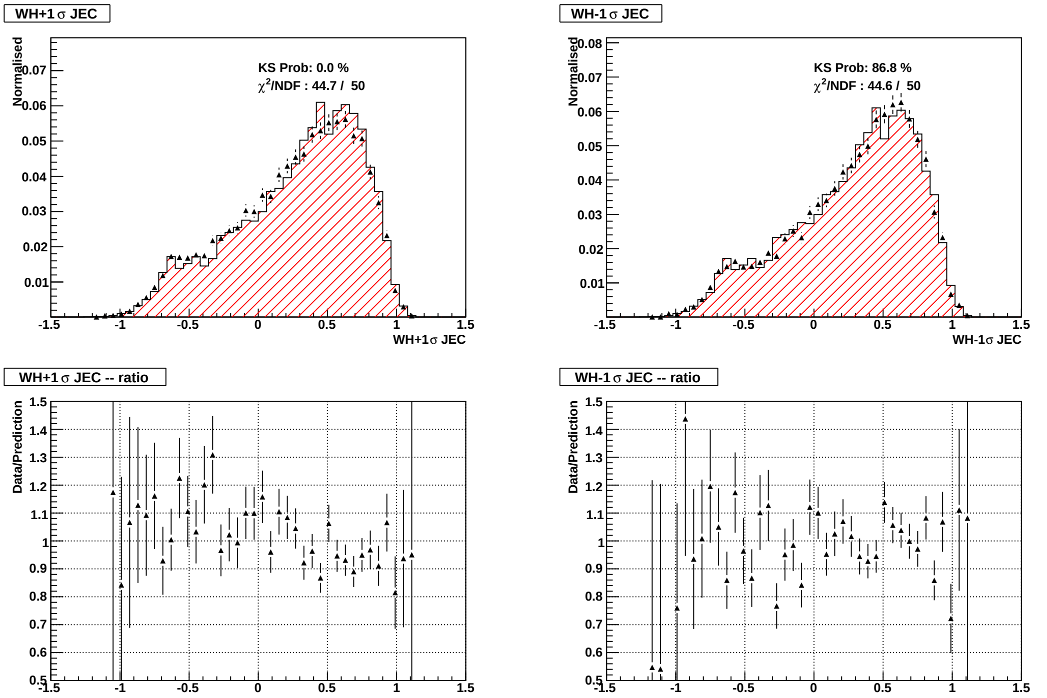
The Jet Shape Systematic does affect the acceptance on account of the SumEt>220.0 GeV event selection cut. Thus the Jet-Energy Correction carries a ±7% rate systematic error in addition to the shape errors.

6.2.2 ISR/FSR Systematics

The ISR/FSR systematics for the Higgs signal was measured using the prescription defined in May 27 2005 Joint Physics Meeting [14]. The effect of the systematic upon the neural network shape for WH120 & VBF120 signal are shown in figures 41 and 42. This is included as a shape systematic in the limit calculation. In addition to the shape systematic, the change in ISR/FSR affects the acceptance. There is a ±2% rate systematic for VH and ±3% for the VBF Higgs signal.

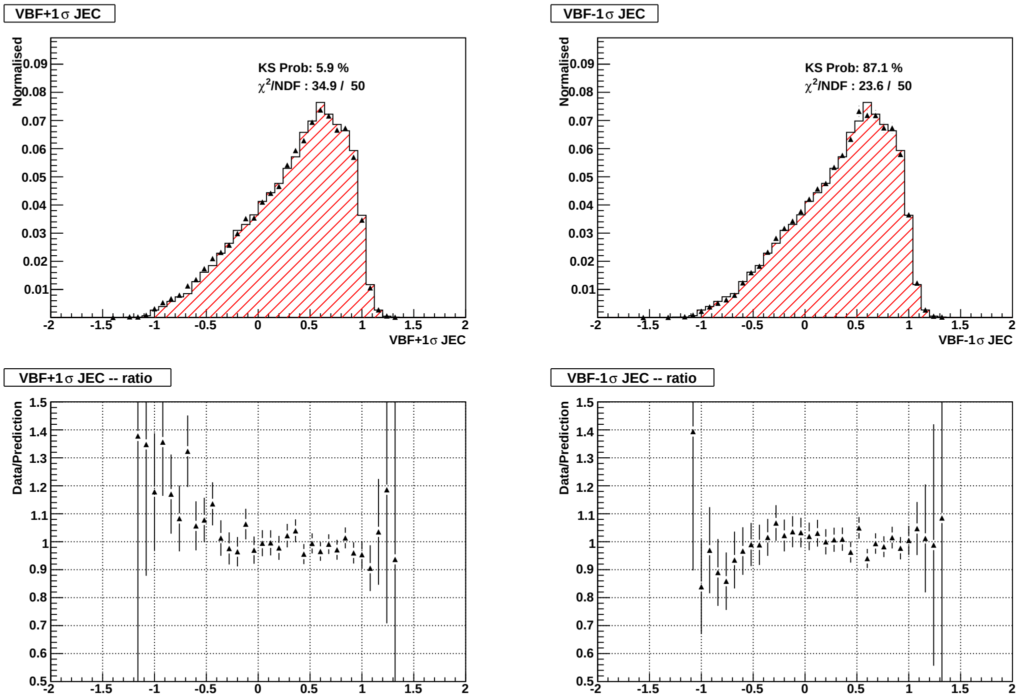


(a) Jet Energy Correction $+1\sigma$ for SS WH120 sample (b) Jet Energy Correction -1σ for SS WH120 sample

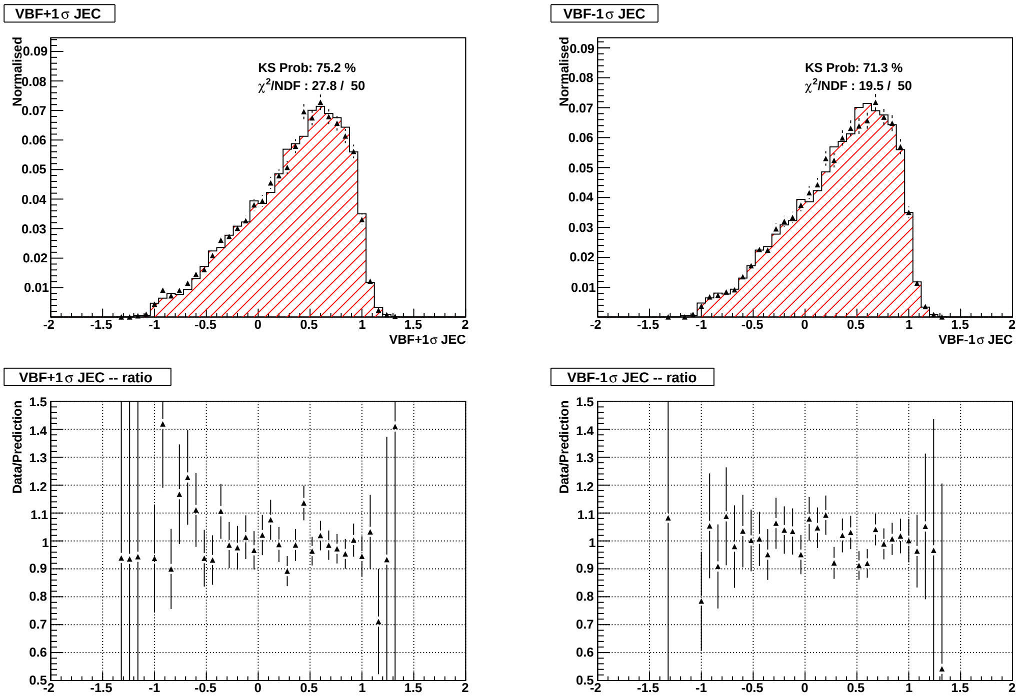


(c) Jet Energy Correction $+1\sigma$ for SJ WH120 sample (d) Jet Energy Correction -1σ for SJ WH120 sample

Figure 39: Jet Energy Correction Systematic : The variation of the WH120 NN output for $\pm1\sigma$ variations of the Jet Energy Correction. The red histograms are the nominal signal distribution. Black points are the jet energy scale systematics with the jet energy scale shifted by $\pm1\sigma$. The NN distributions have been normalised to unit area to measure variations in shape. The ratio plots show shifted Jet Energy Correction / Nominal Jet Energy Correction.

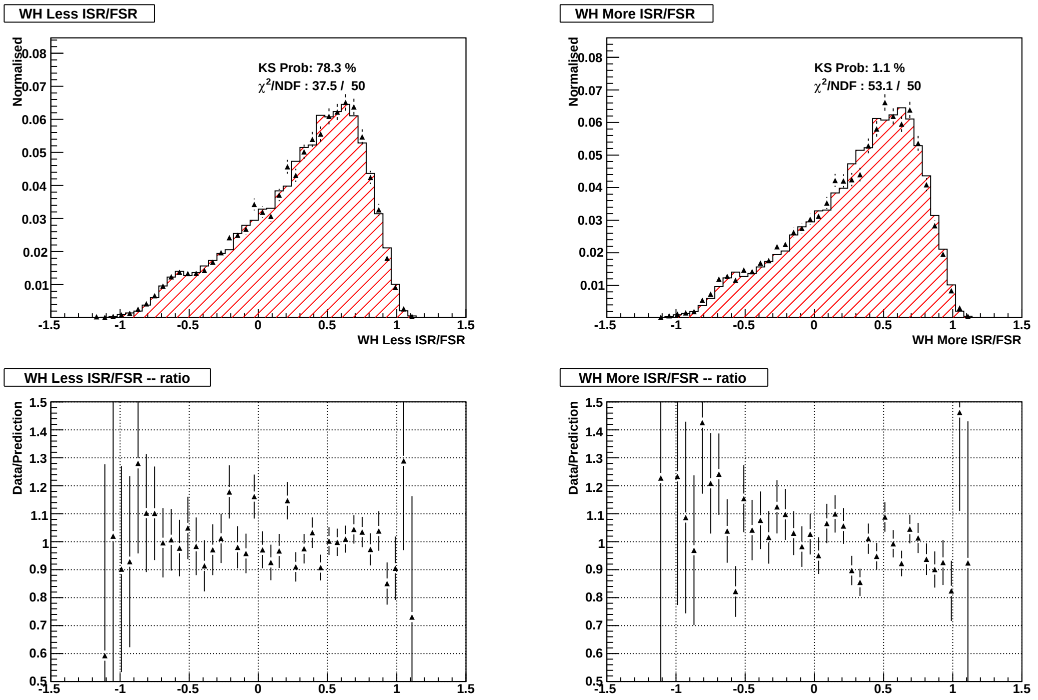


(a) VBF SS: $+1\sigma$ Jet Energy Correction (b) VBF SS -1σ Jet Energy Correction

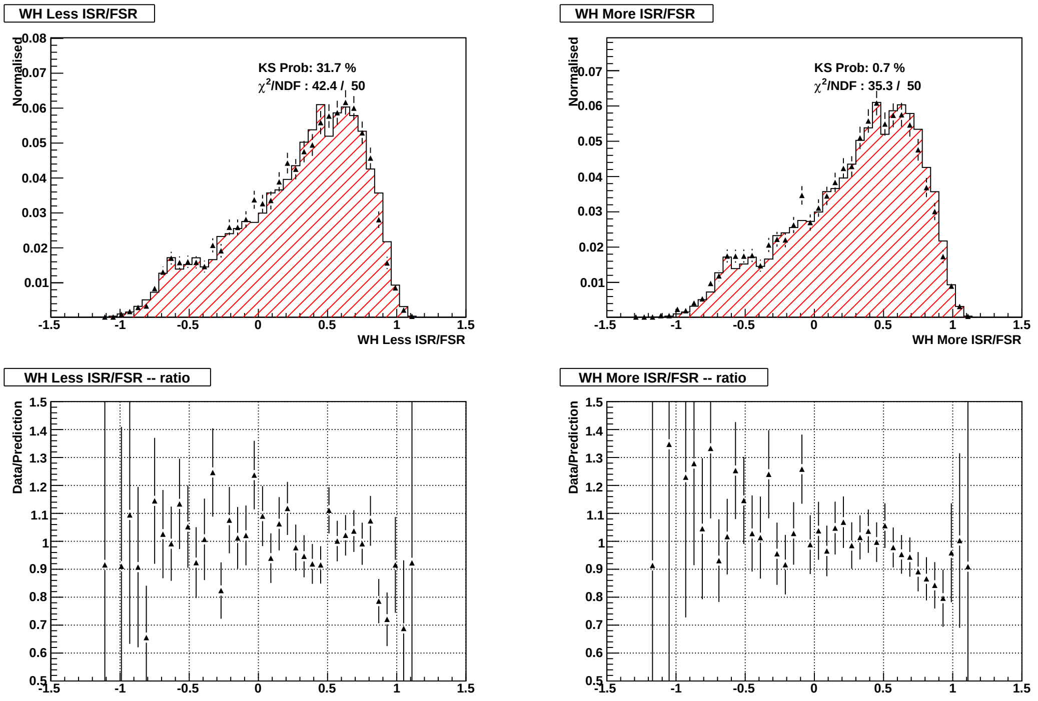


(c) VBF SJ $+1\sigma$ Jet Energy Correction (d) VBF SJ -1σ Jet Energy Correction

Figure 40: Jet Energy Correction Systematic : The variation of the VBF120 NN output for $\pm 1\sigma$ variations of the Jet Energy Correction. The red histograms are the nominal signal distribution. Black points are the jet energy scale systematics with the jet energy scale shifted by $\pm 1\sigma$. The NN distributions have been normalised to unit area to measure variations in shape. The ratio plots show shifted Jet Energy Correction / Nominal Jet Energy Correction.

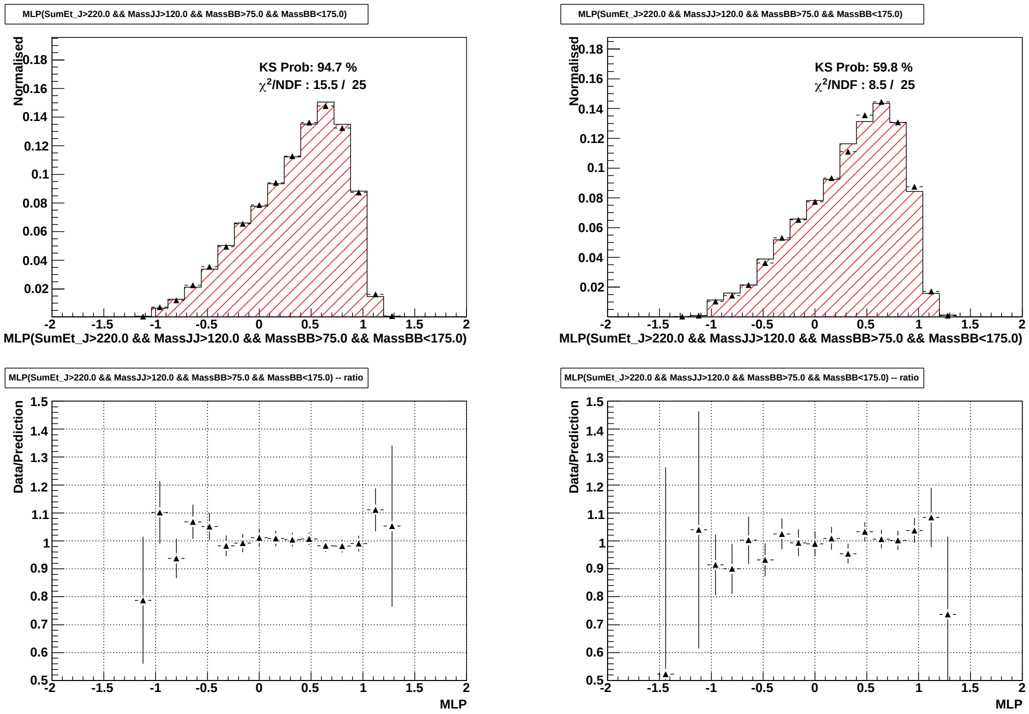


(a) WH120 SS Less ISR/FSR (b) WH120 SS More ISR/FSR



(c) WH120 SJ Less ISR/FSR (d) WH120 SJ More ISR/FSR

Figure 41: Signal NN output for the WH120 channel. The red histogram is the nominal signal distribution. the black points are the ISR/FSR systematic. The bottom plot is the ratio of the systematic with respect to the nominal distribution.



(a) VBF120 SS ISR/FSR

(b) VBF120 SJ ISR/FSR

Figure 42: Signal NN output for the SS VBF channel. The red histogram is the nominal signal distribution. the black points are the ISR/FSR systematic. The bottom plot is the ratio of the systematic with respect to the nominal distribution.

6.3 Non-QCD Background Cross-Section Systematics

Monte-Carlo simulation is used to model all the non-QCD backgrounds. For each non-QCD background, there is an uncertainty for the cross-section. This uncertainty is applied as a rate systematic for the limit setting calculation (Table 2.2). The applied uncertainties are:

$t\bar{t}$ & Single-Top A rate uncertainty of 10% is applied

WW/WZ/WW A rate uncertainty of 6% is applied

W+Jets & Z+Jets backgrounds A uniform 50% rate uncertainty is applied

6.4 Signal Rate Systematics

In addition the mentioned signal shape systematics, there are several rate signal systematic:

PDF Using the prescription defined in CDF Joint Physics Meeting (April 2nd 2004) [15] [16], there is a ±2% rate systematic.

B-Tagging All MC SS events have to apply a scale factor of 0.895 with a 7.6% rate systematic. For SJ events, the MC scale factor is 0.699 with a 9.7% rate systematic.

Luminosity The standard 6% luminosity uncertainty is applied.

6.5 Trigger

From the studies described in section & CDF Note 9954 [4], there is a 4% rate systematic for all signal samples.

6.6 Summary of all Systematic

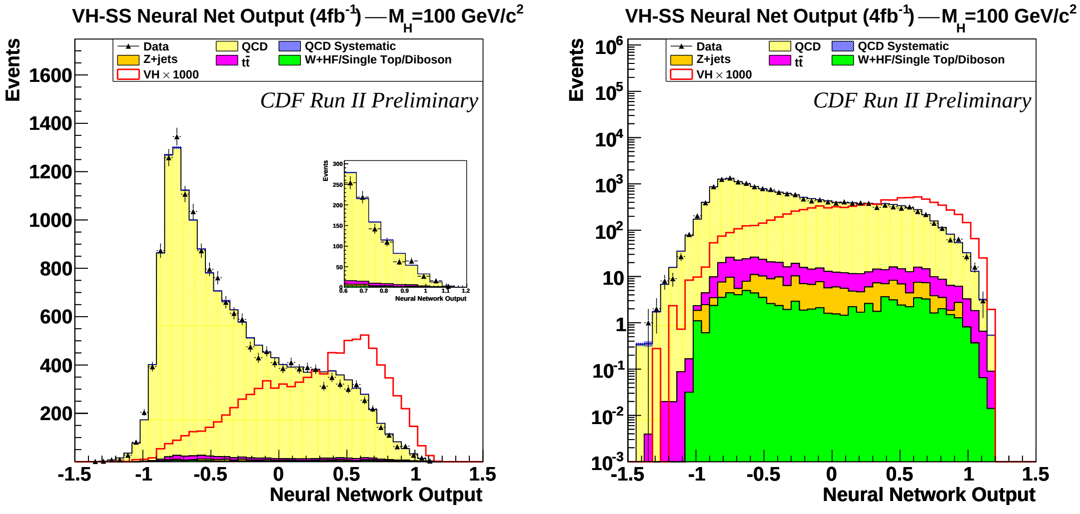
Table 7 summarises all the rate uncertainties & shape systematics which are applied to calculate the limit.

Table 7: Summary of all Systematic Uncertainties used to calculate the limit

Jet Energy Correction	±7% Rate Shape
PDF	± 2% Rate
SecVtx+SecVtx	7.6% Rate
SecVtx+JetProb	9.7% Rate
Luminosity	6% Rate
ISR/FSR	± 2% for VH Rate ± 3% for VBF Rate Shape for VH & VBF
Jet Moment	Shape
Trigger	± 4% Rate
QCD Interpolation	Shape
QCD MJJ Tuning	Shape
QCD Jet Moment Tuning	Shape
$t\bar{t}$ & single-top cross-section	± 10% Rate
Diboson (WW/WZ/ZZ) cross-section	± 6% Rate
W+Jets & Z+Jets cross-section	± 50% Rate

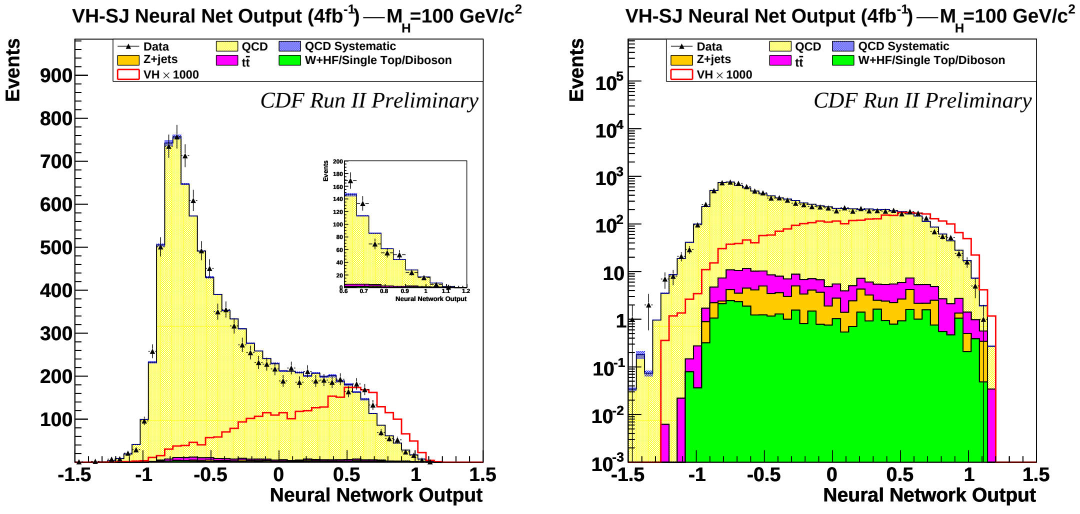
7 Unblinded Signal Region

Figures 43-64 show the unblinded data for the VH and VBF channels. Over the whole neural net range, the data is in good agreement with the expected background with no evidence of a Higgs signal.



(a) VH100 SS Linear

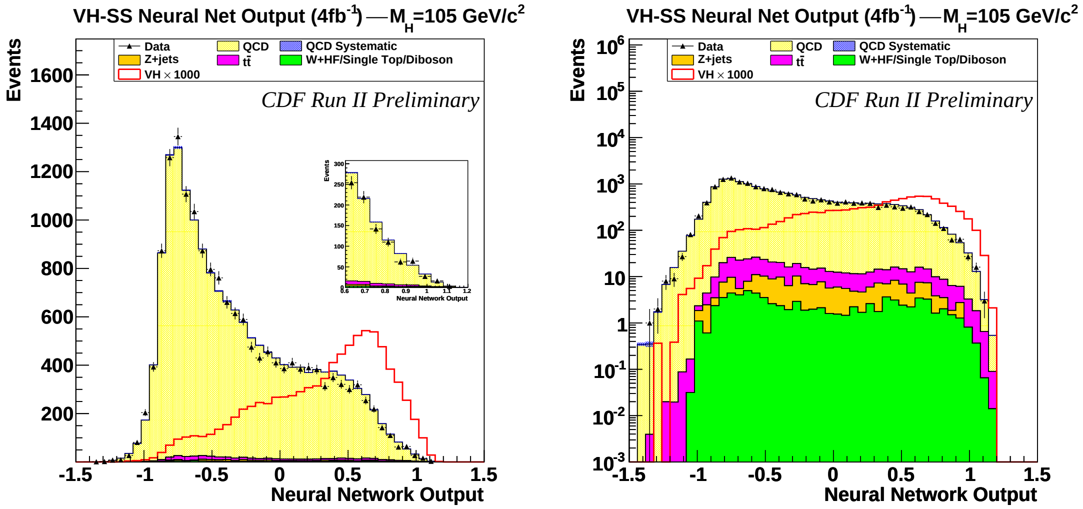
(b) VH100 SJ Log



(c) VH 100 SJ Linear

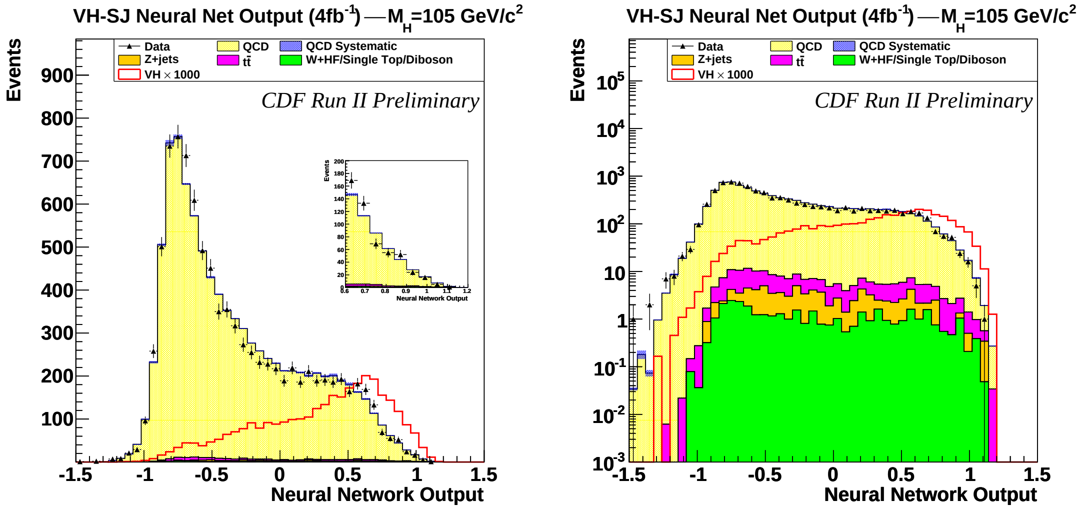
(d) VH 100 SJ Log

Figure 43: The Neural Net distributions for the VH channel for Higgs masses of 100 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel. The inset in the linear scale plot highlights the signal region which shows the data is consistent with the background prediction.



(a) VH105 SS Linear

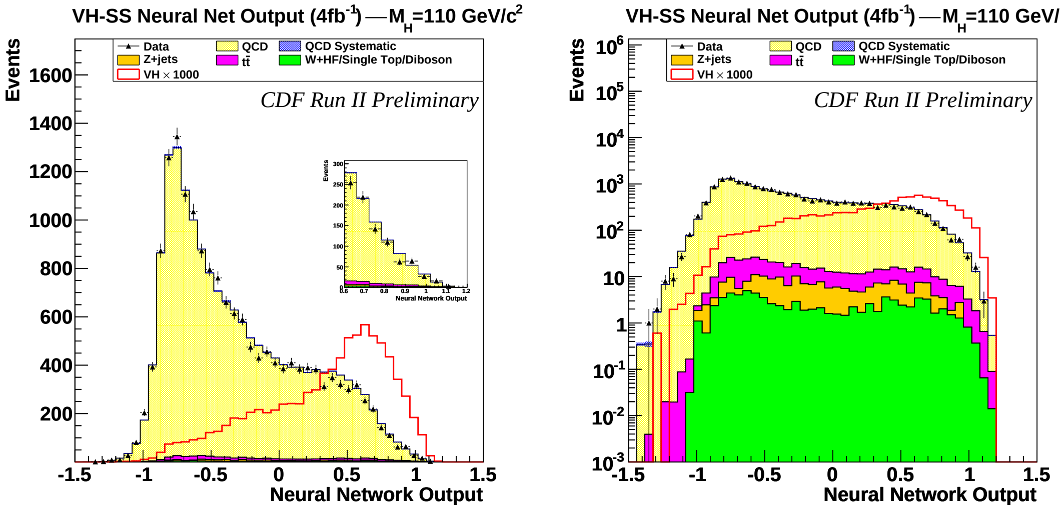
(b) VH105 SJ Log



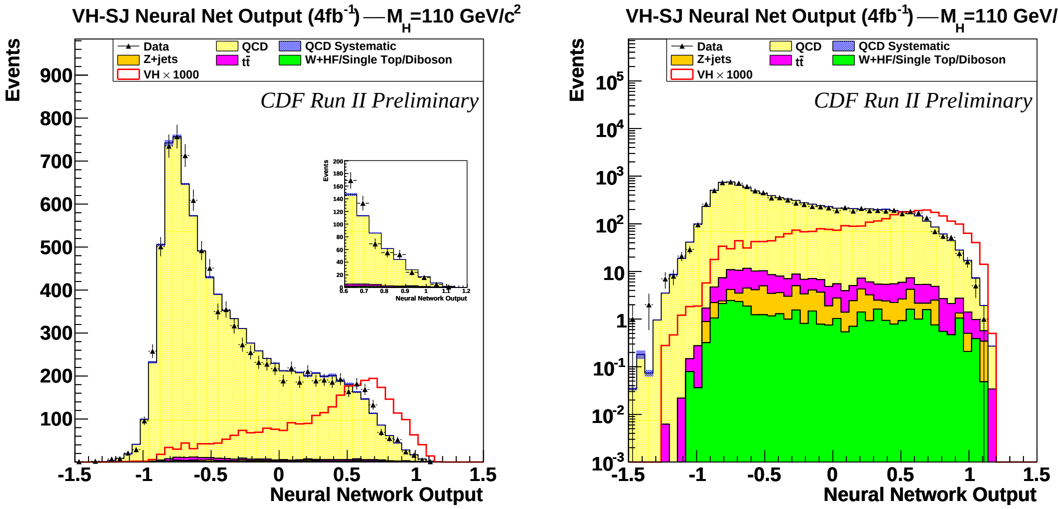
(c) VH 105 SJ Linear

(d) VH 105 SJ Log

Figure 44: The Neural Net distributions for the VH channel for Higgs masses of 105 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel. The inset in the linear scale plot highlights the signal region which shows the data is consistent with the background prediction.

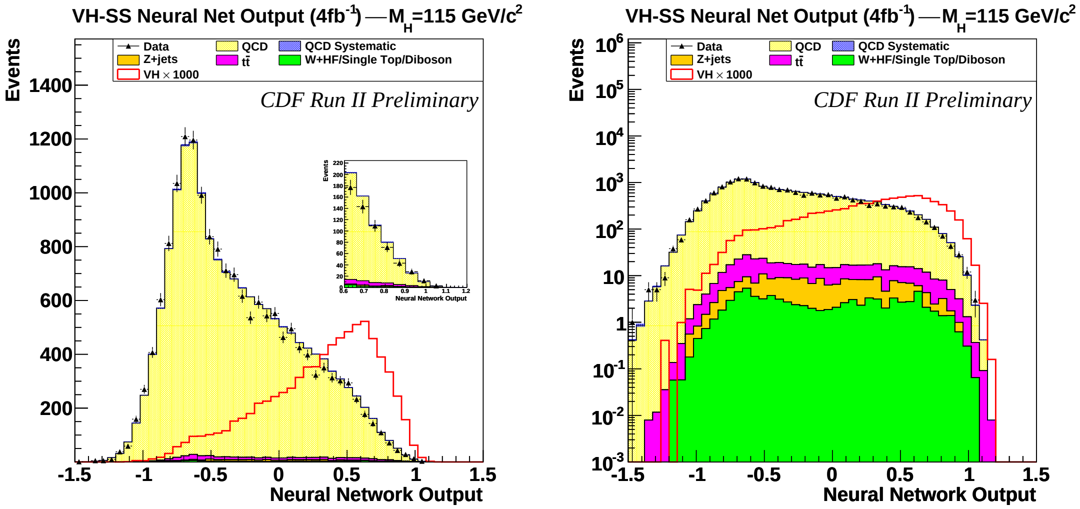


(a) VH110 SS Linear (b) VH110 SJ Log



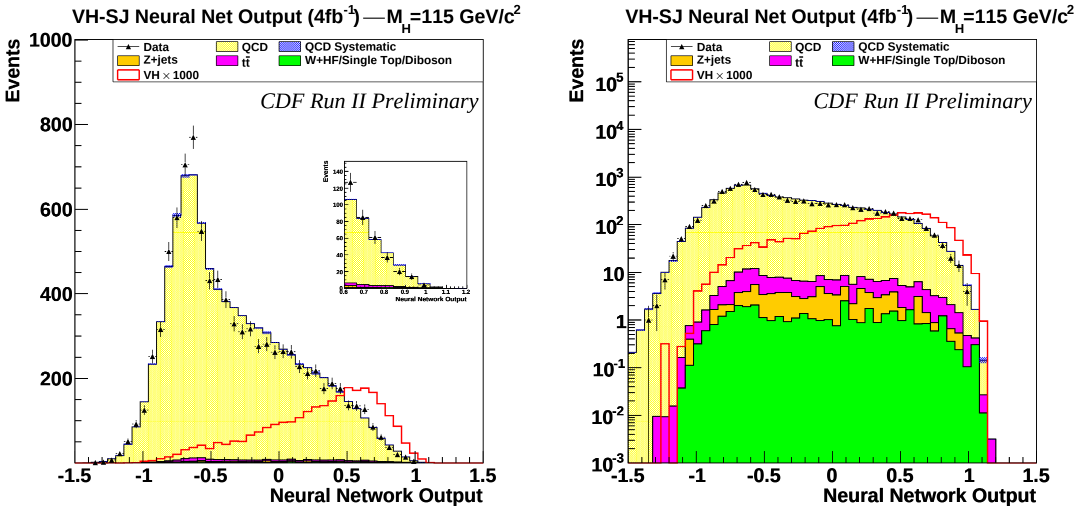
(c) VH 110 SJ Linear (d) VH 110 SJ Log

Figure 45: The Neural Net distributions for the VH channel for Higgs masses of 110 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel. The inset in the linear scale plot highlights the signal region which shows the data is consistent with the background prediction.



(a) VH115 SS Linear

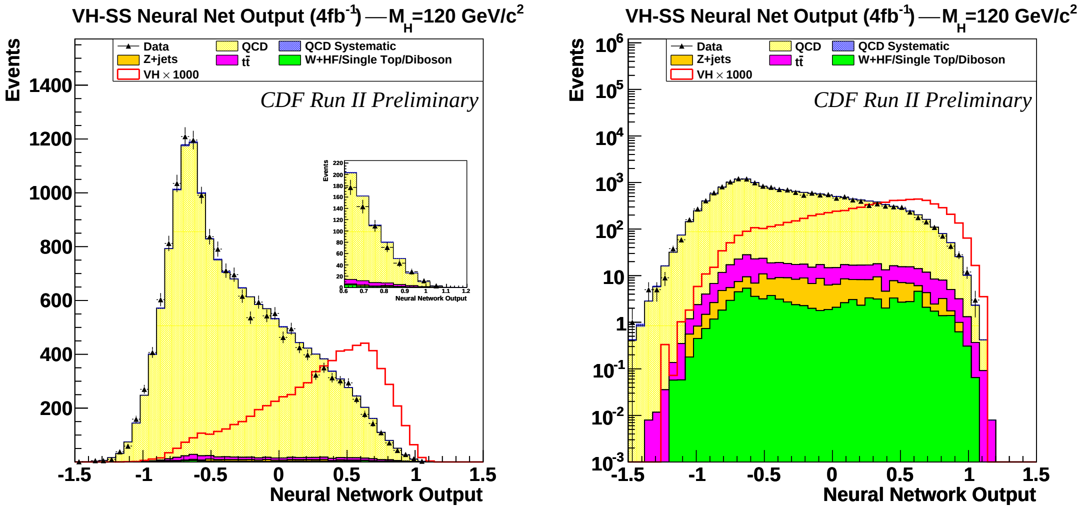
(b) VH115 SJ Log



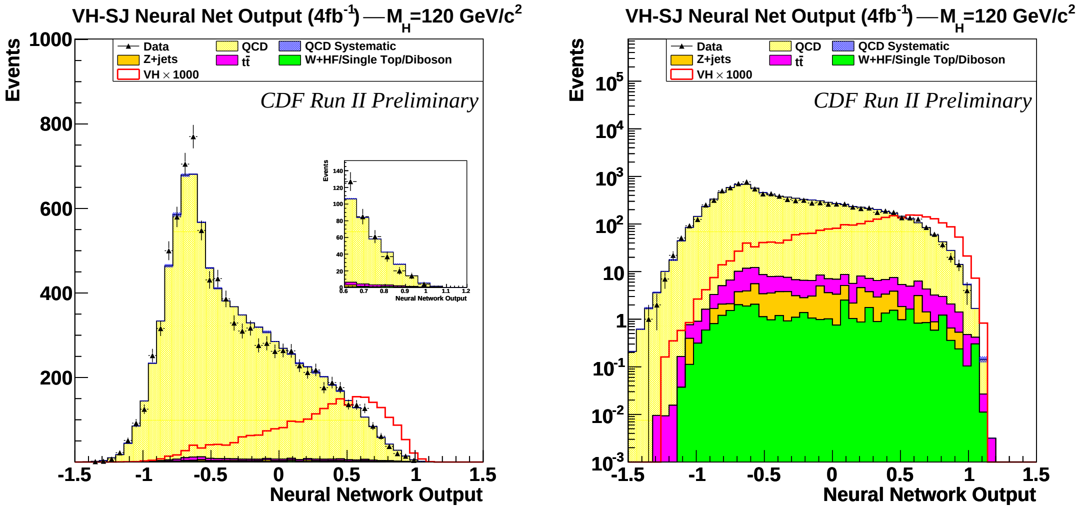
(c) VH 115 SJ Linear

(d) VH 115 SJ Log

Figure 46: The Neural Net distributions for the VH channel for Higgs masses of 115 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel. The inset in the linear scale plot highlights the signal region which shows the data is consistent with the background prediction.

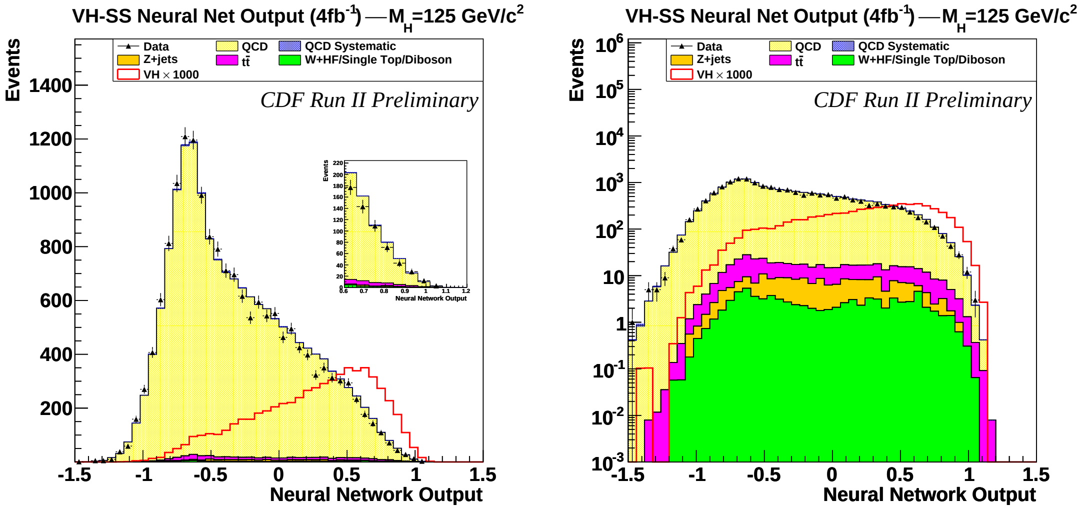


(a) VH120 SS Linear (b) VH120 SJ Log



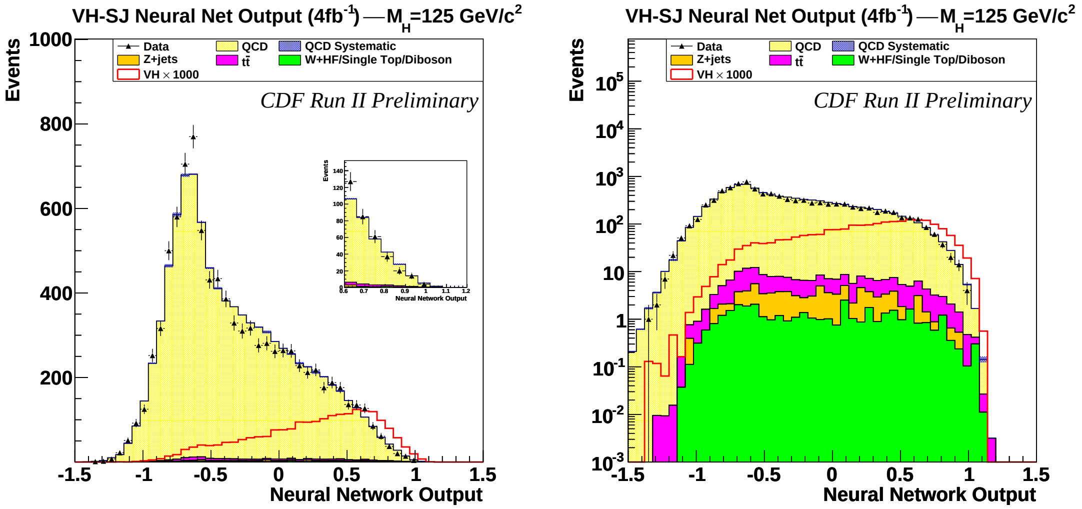
(c) VH 120 SJ Linear (d) VH 120 SJ Log

Figure 47: The Neural Net distributions for the VH channel for Higgs masses of 120 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel. The inset in the linear scale plot highlights the signal region which shows the data is consistent with the background prediction.



(a) VH125 SS Linear

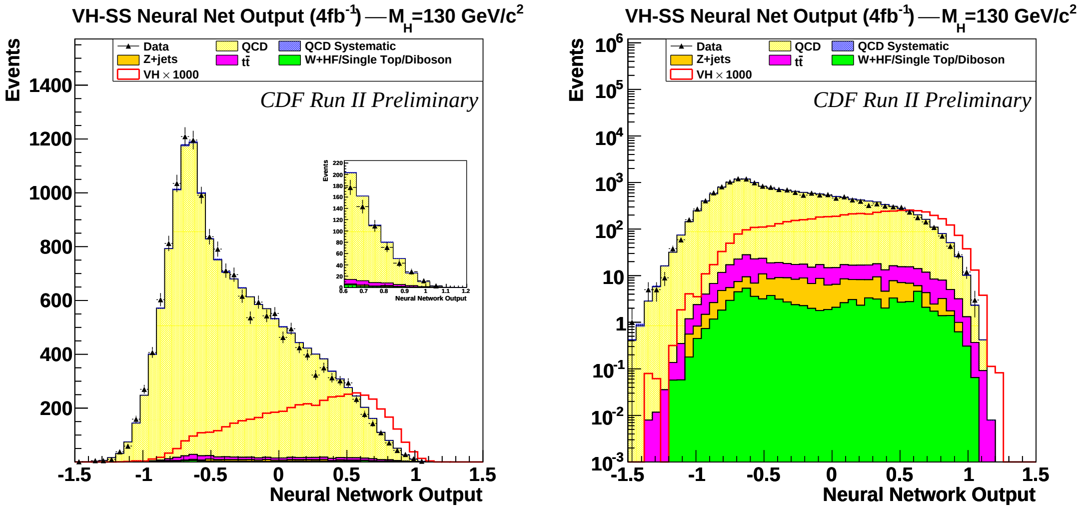
(b) VH125 SJ Log



(c) VH 125 SJ Linear

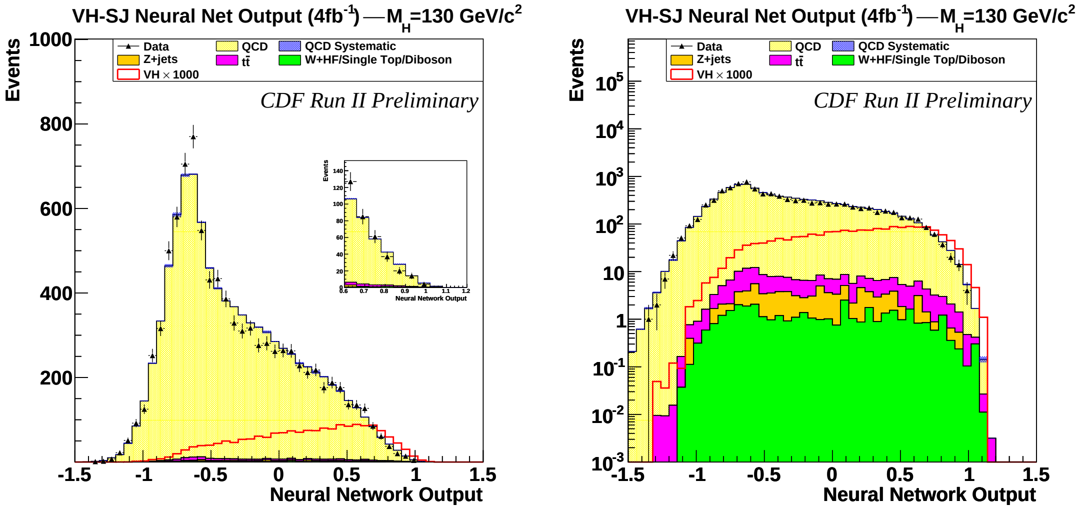
(d) VH 125 SJ Log

Figure 48: The Neural Net distributions for the VH channel for Higgs masses of 125 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel. The inset in the linear scale plot highlights the signal region which shows the data is consistent with the background prediction.



(a) VH130 SS Linear

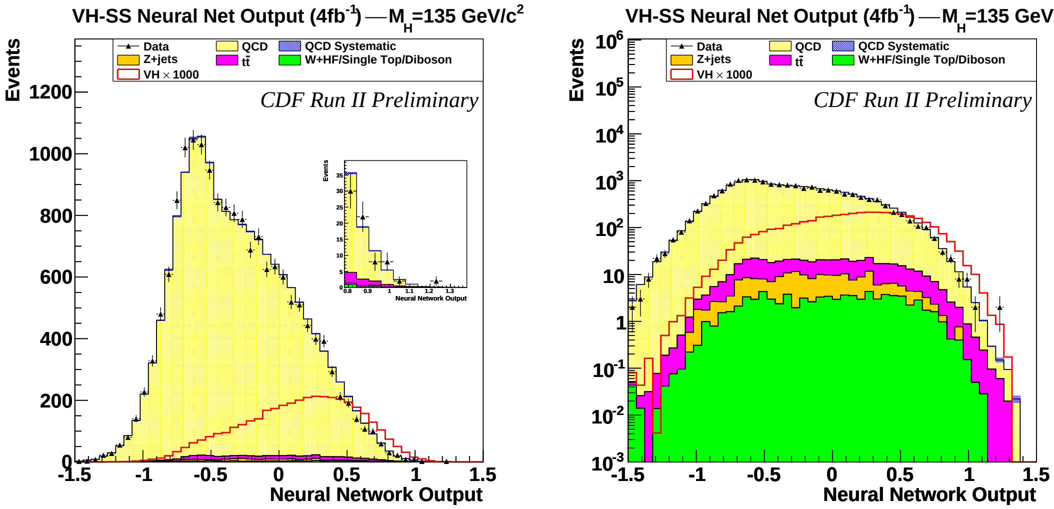
(b) VH130 SJ Log



(c) VH 130 SJ Linear

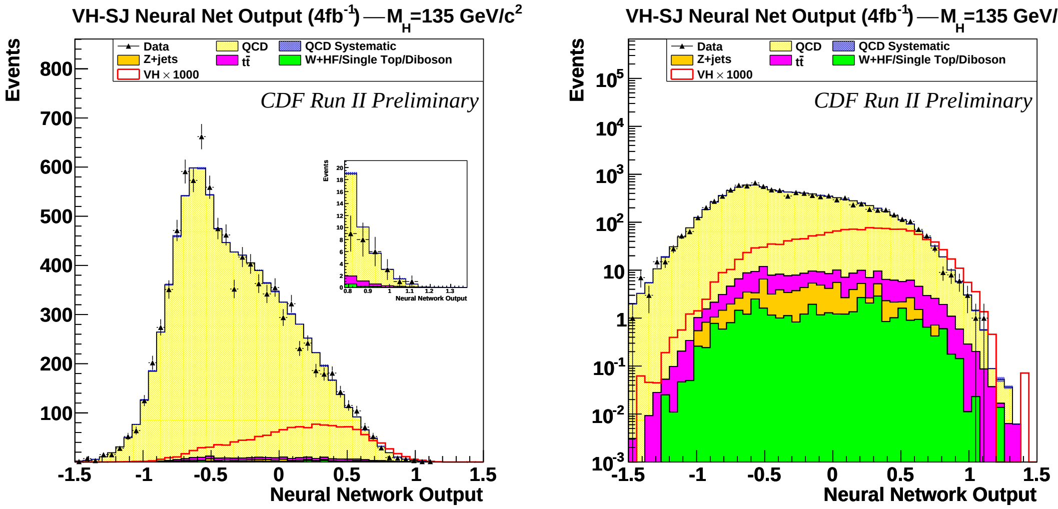
(d) VH 130 SJ Log

Figure 49: The Neural Net distributions for the VH channel for Higgs masses of 130 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel. The inset in the linear scale plot highlights the signal region which shows the data is consistent with the background prediction.



(a) VH135 SS Linear

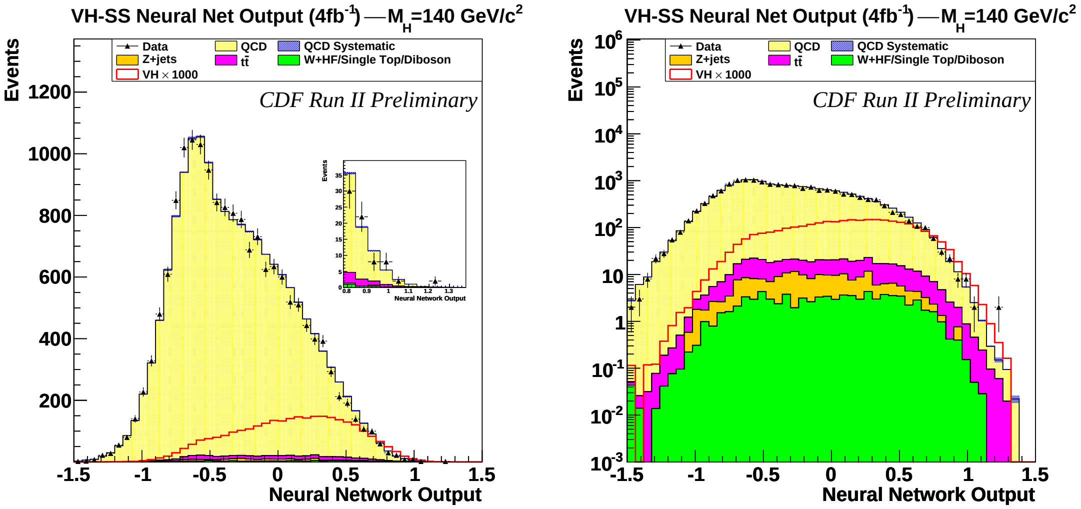
(b) VH135 SJ Log



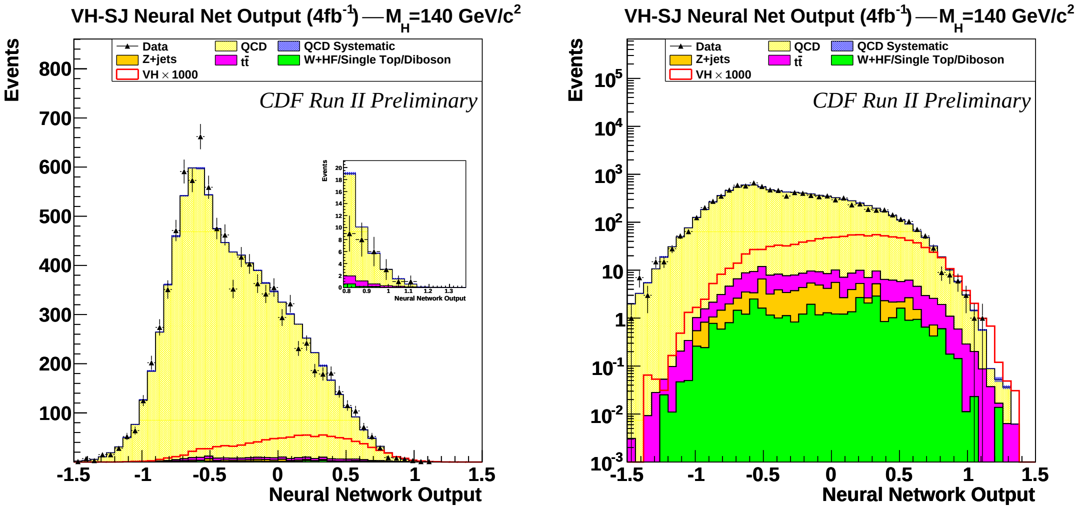
(c) VH 135 SJ Linear

(d) VH 135 SJ Log

Figure 50: The Neural Net distributions for the VH channel for Higgs masses of 135 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel. The inset in the linear scale plot highlights the signal region which shows the data is consistent with the background prediction.

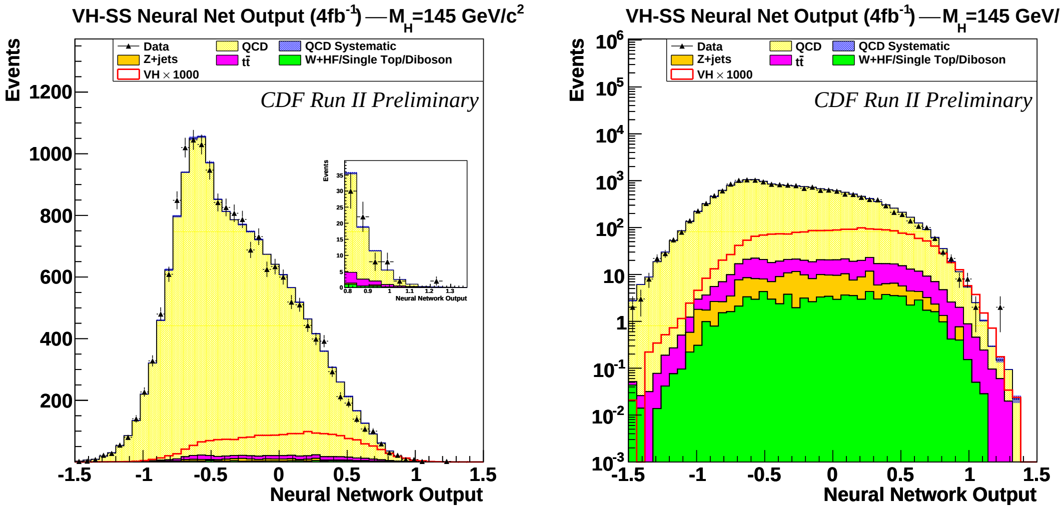


(a) VH140 SS Linear (b) VH140 SJ Log



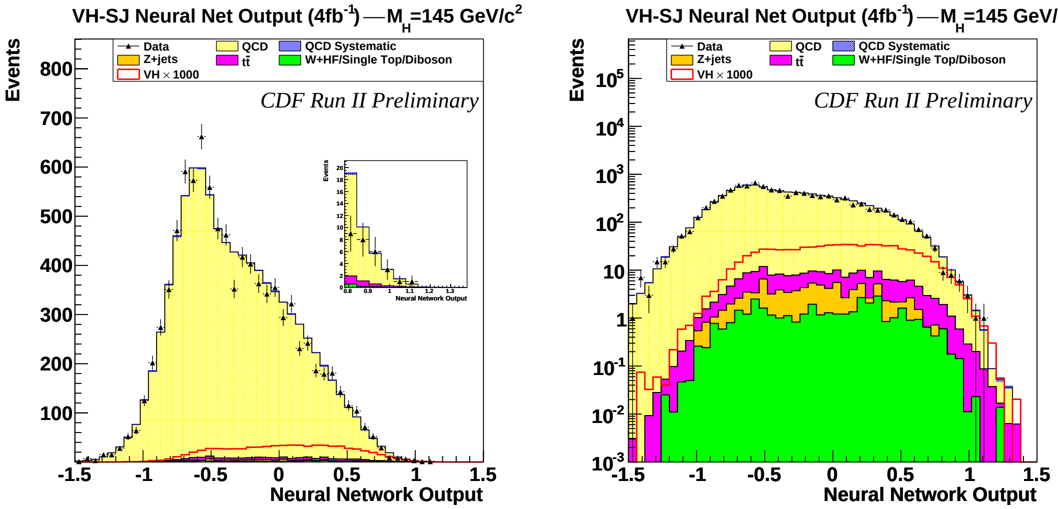
(c) VH 140 SJ Linear (d) VH 140 SJ Log

Figure 51: The Neural Net distributions for the VH channel for Higgs masses of 140 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel. The inset in the linear scale plot highlights the signal region which shows the data is consistent with the background prediction.



(a) VH145 SS Linear

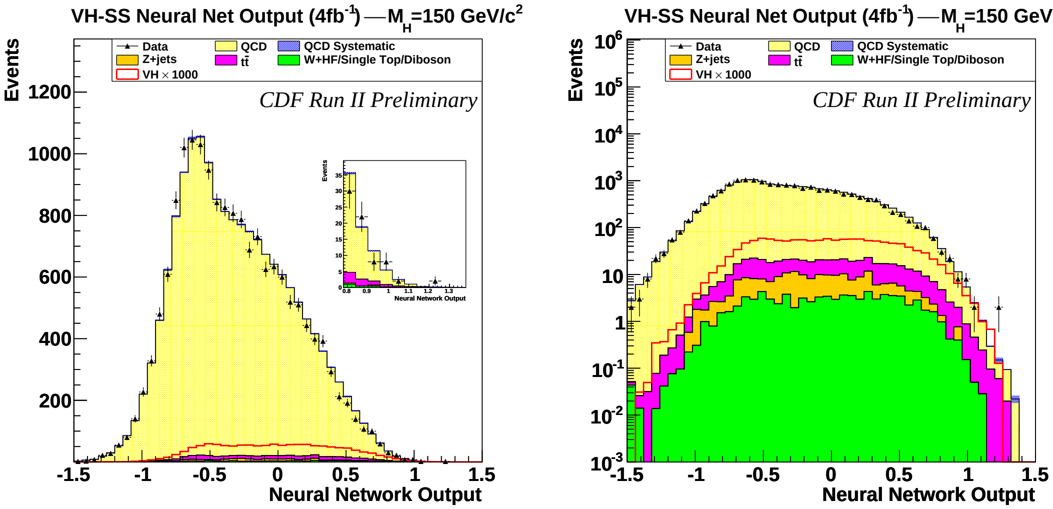
(b) VH145 SJ Log



(c) VH 145 SJ Linear

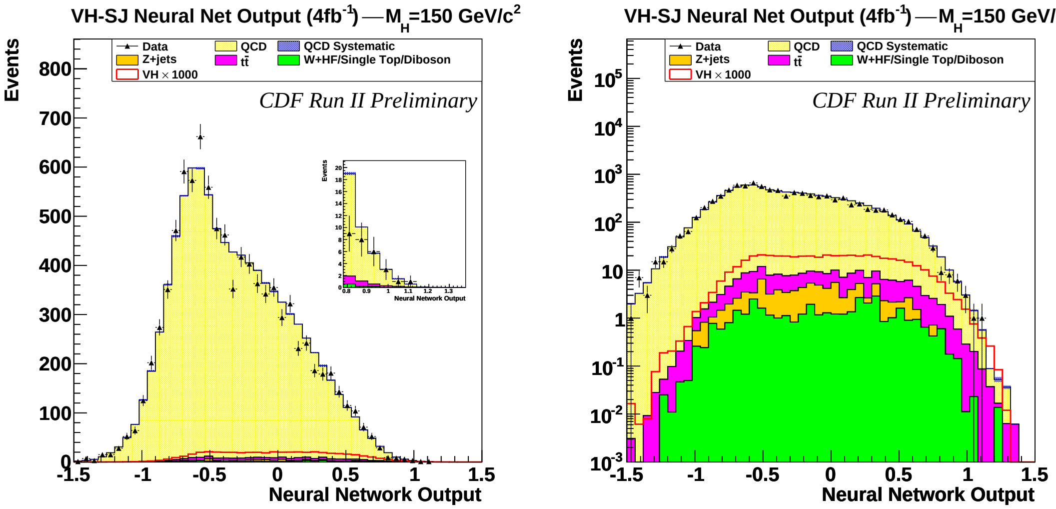
(d) VH 145 SJ Log

Figure 52: The Neural Net distributions for the VH channel for Higgs masses of 145 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel. The inset in the linear scale plot highlights the signal region which shows the data is consistent with the background prediction.



(a) VH150 SS Linear

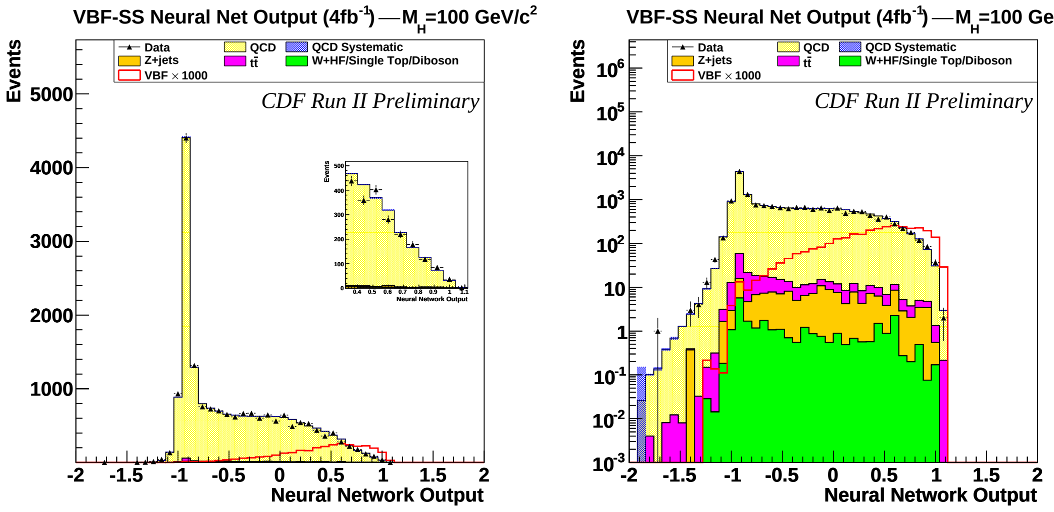
(b) VH150 SJ Log



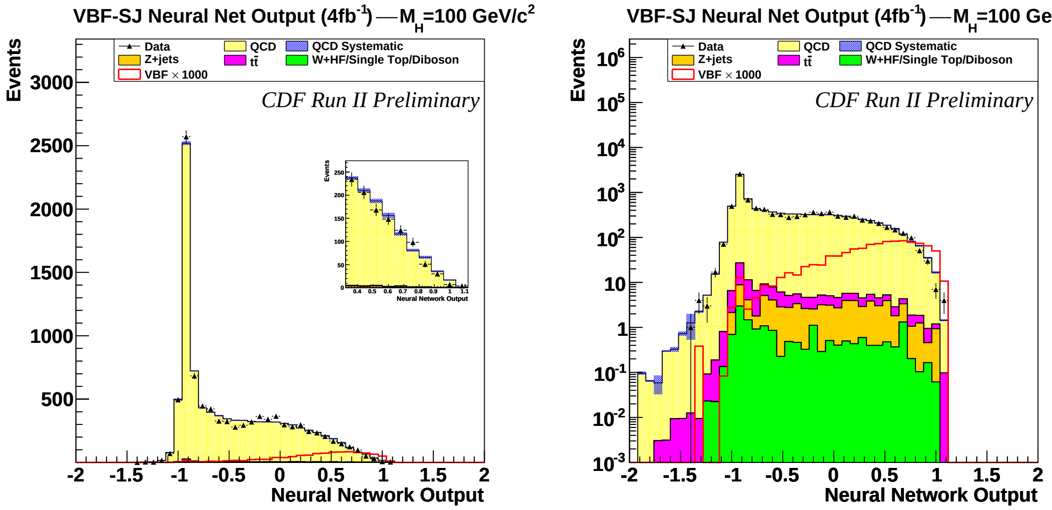
(c) VH 150 SJ Linear

(d) VH 150 SJ Log

Figure 53: The Neural Net distributions for the VH channel for Higgs masses of 150 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel. The inset in the linear scale plot highlights the signal region which shows the data is consistent with the background prediction.

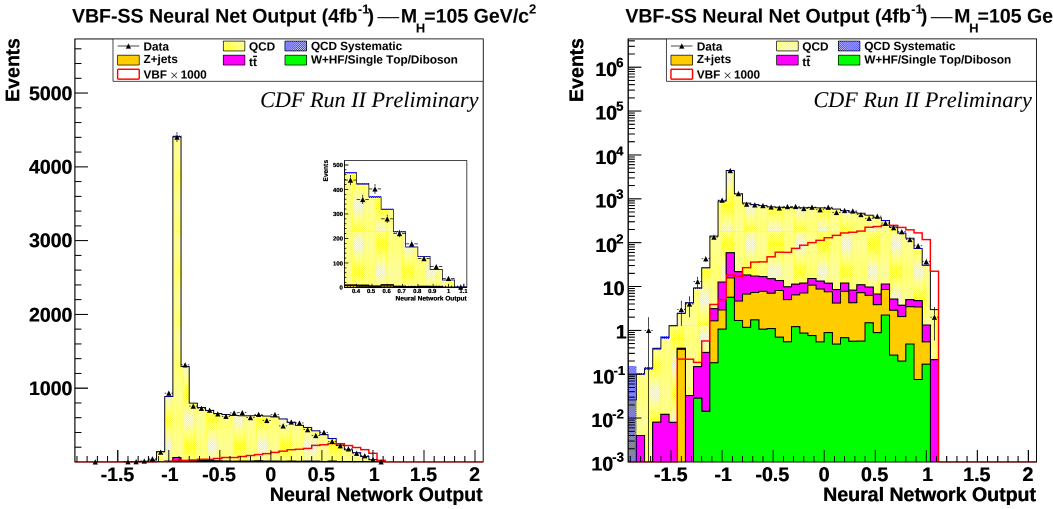


(a) VBF 100 SS Linear (b) VBF 100 SJ Log

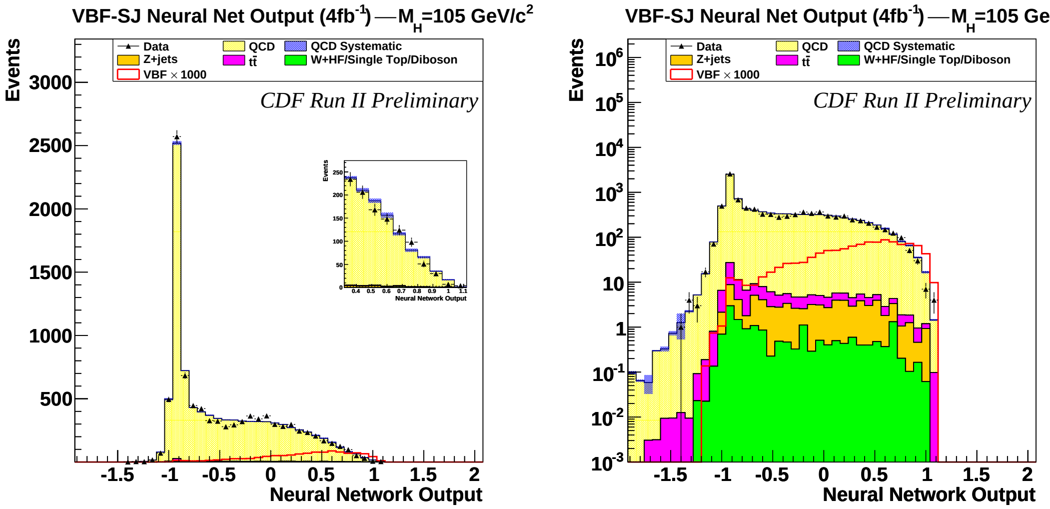


(c) VBF 100 SJ Linear (d) VBF 100 SJ Log

Figure 54: The Neural Net distributions for the VBF channel for Higgs masses of 100 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel. The inset in the linear scale plot highlights the signal region which shows the data is consistent with the background prediction.

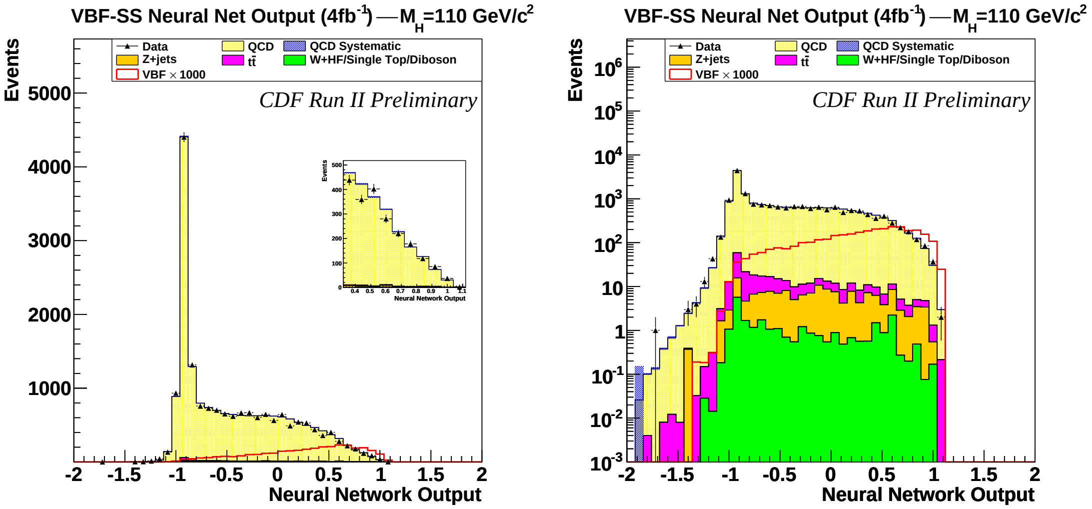


(a) VBF 105 SS Linear (b) VBF 105 SJ Log

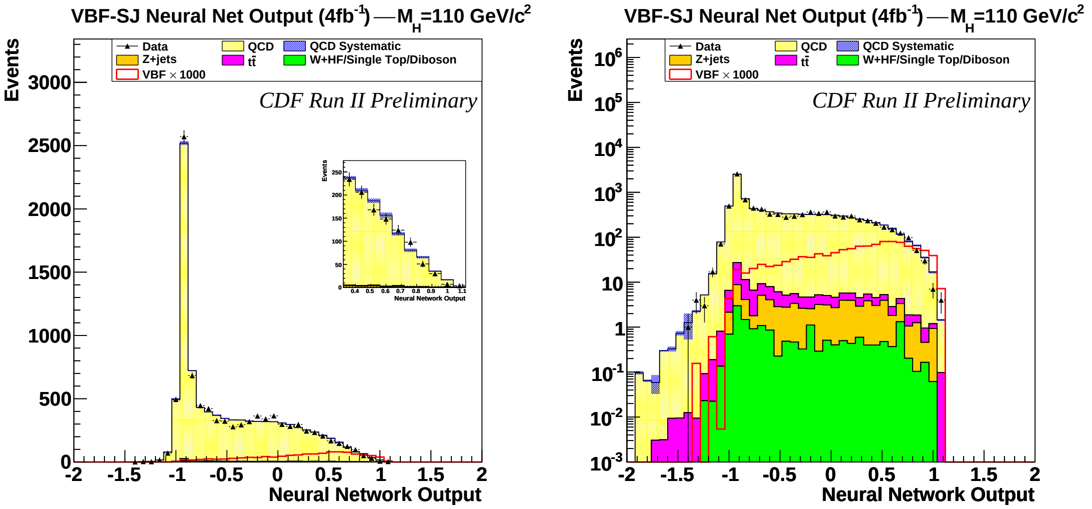


(c) VBF 105 SJ Linear (d) VBF 105 SJ Log

Figure 55: The Neural Net distributions for the VBF channel for Higgs masses of 105 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel. The inset in the linear scale plot highlights the signal region which shows the data is consistent with the background prediction.

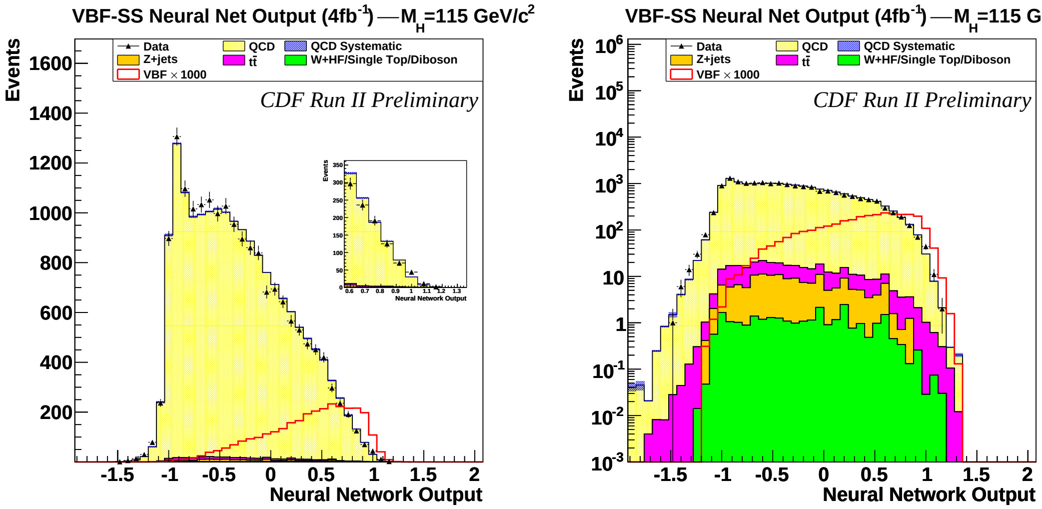


(a) VBF 110 SS Linear (b) VBF 110 SJ Log

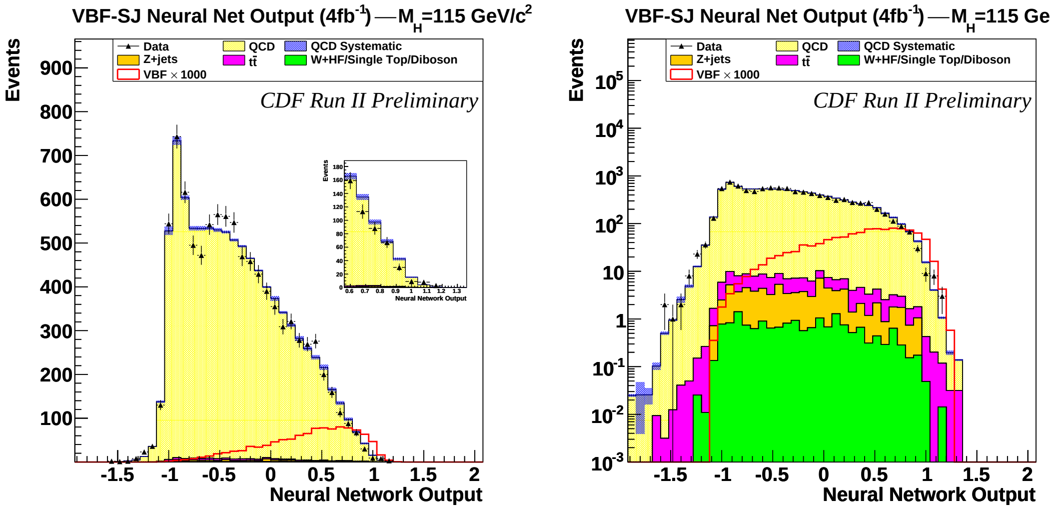


(c) VBF 110 SJ Linear (d) VBF 110 SJ Log

Figure 56: The Neural Net distributions for the VBF channel for Higgs masses of 110 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel. The inset in the linear scale plot highlights the signal region which shows the data is consistent with the background prediction.

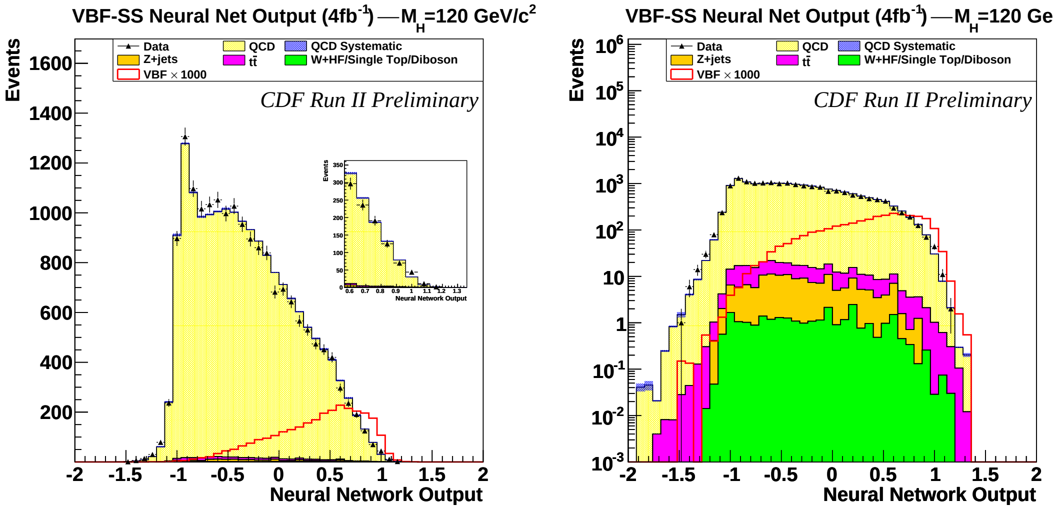


(a) VBF 115 SS Linear (b) VBF 115 SJ Log

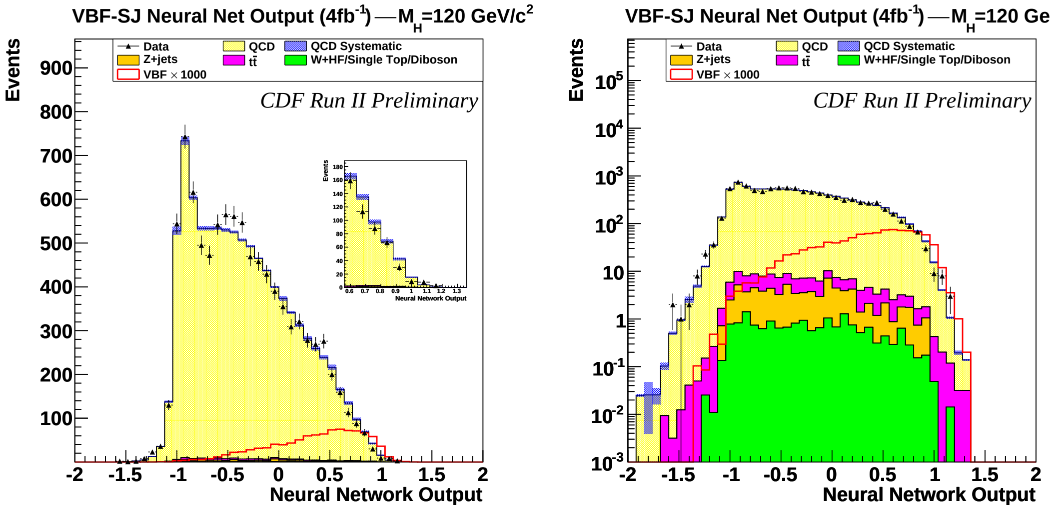


(c) VBF 115 SJ Linear (d) VBF 115 SJ Log

Figure 57: The Neural Net distributions for the VBF channel for Higgs masses of 115 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel. The inset in the linear scale plot highlights the signal region which shows the data is consistent with the background prediction.

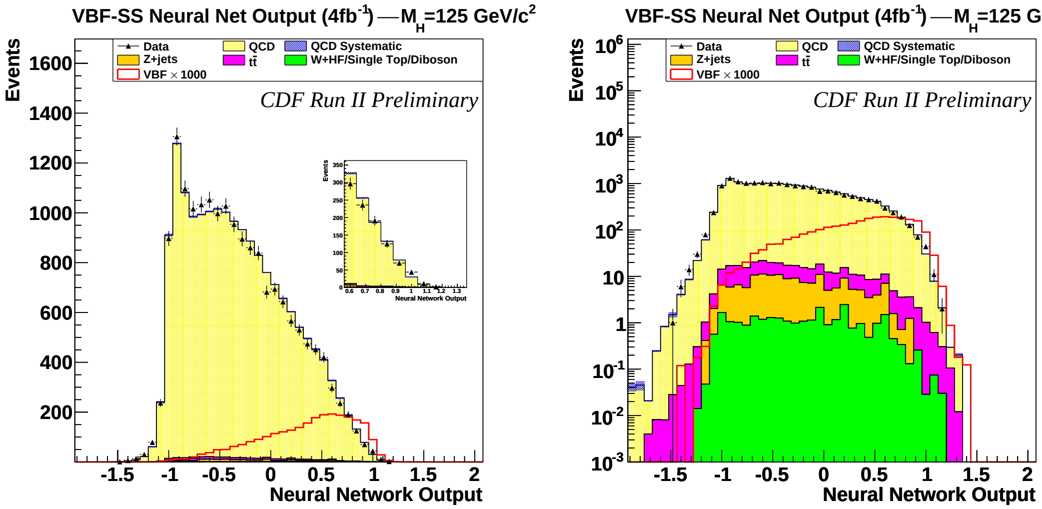


(a) VBF 120 SS Linear (b) VBF 120 SJ Log



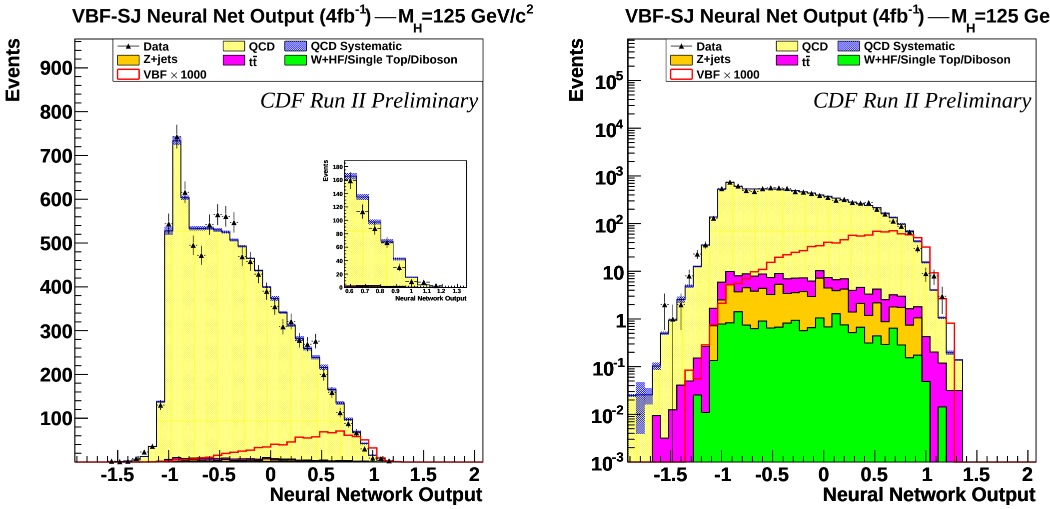
(c) VBF 120 SJ Linear (d) VBF 120 SJ Log

Figure 58: The Neural Net distributions for the VBF channel for Higgs masses of 120 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel. The inset in the linear scale plot highlights the signal region which shows the data is consistent with the background prediction.



(a) VBF 125 SS Linear

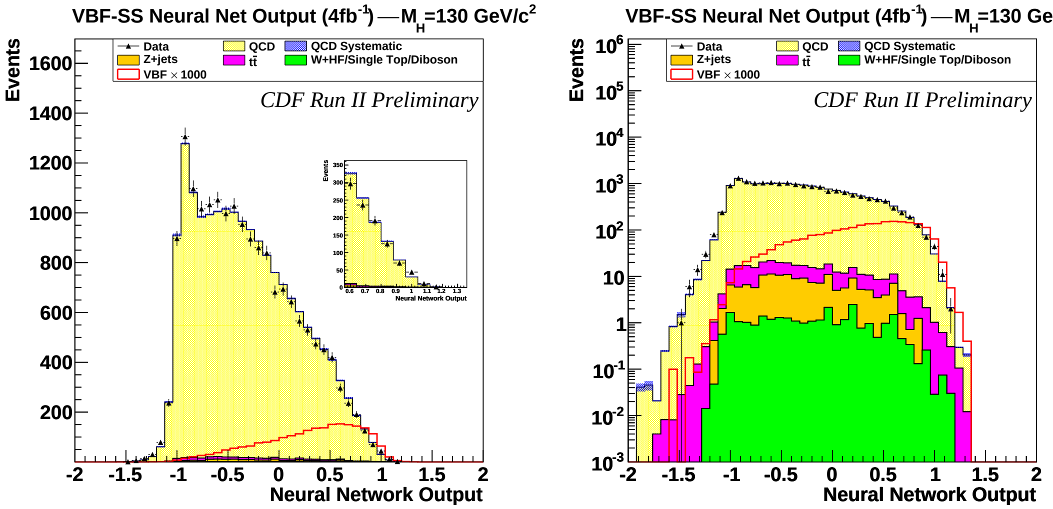
(b) VBF 125 SJ Log



(c) VBF 125 SJ Linear

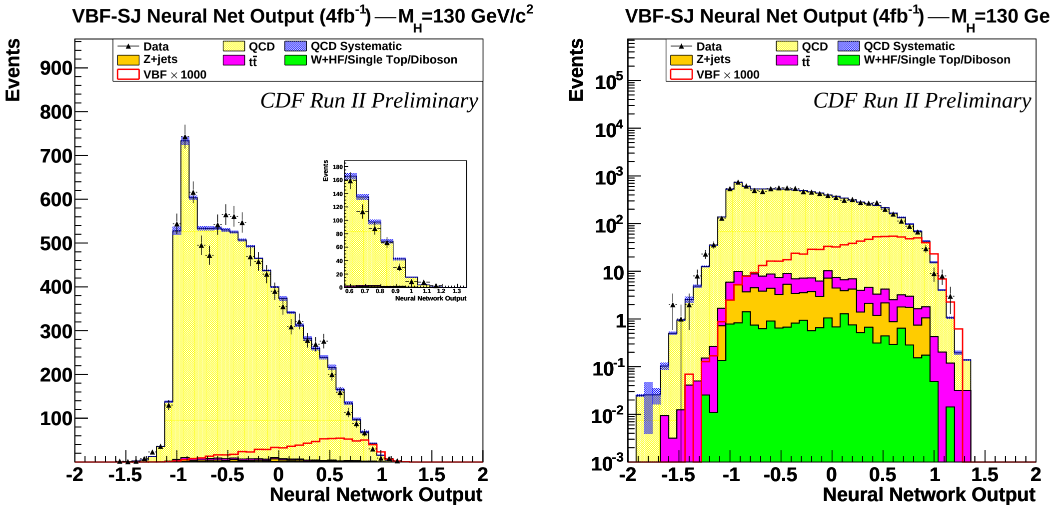
(d) VBF 125 SJ Log

Figure 59: The Neural Net distributions for the VBF channel for Higgs masses of 125 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel. The inset in the linear scale plot highlights the signal region which shows the data is consistent with the background prediction.



(a) VBF 130 SS Linear

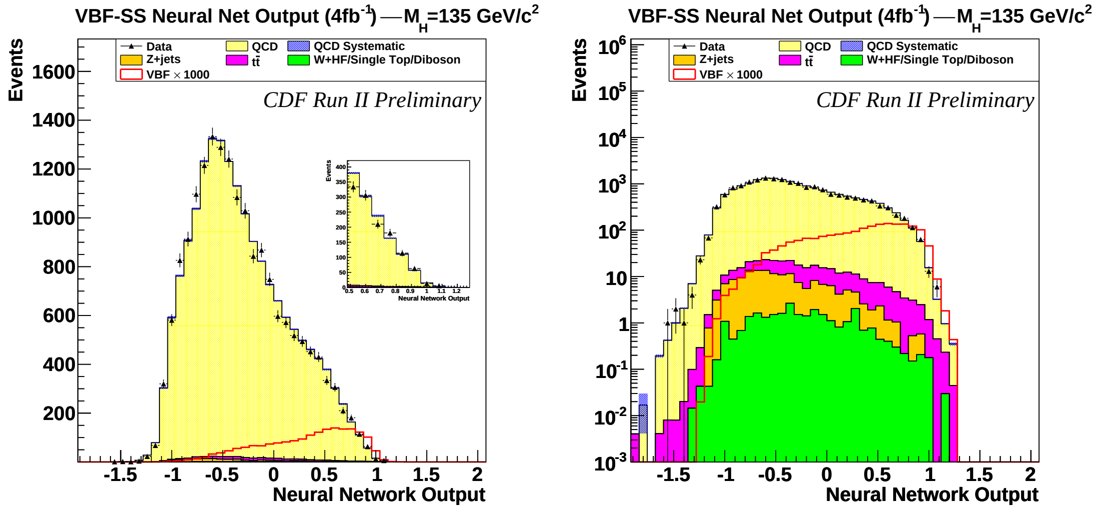
(b) VBF 130 SJ Log



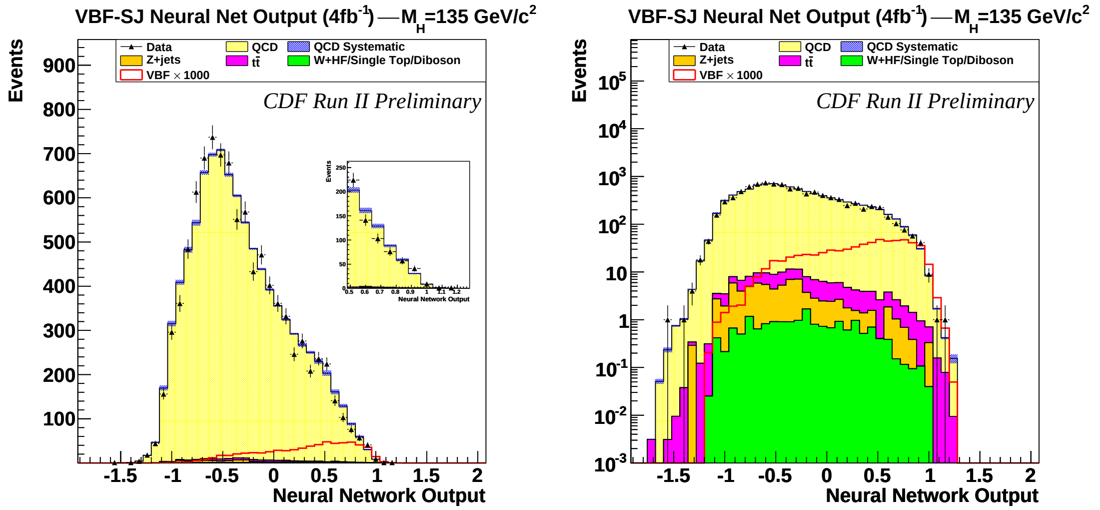
(c) VBF 130 SJ Linear

(d) VBF 130 SJ Log

Figure 60: The Neural Net distributions for the VBF channel for Higgs masses of 130 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel. The inset in the linear scale plot highlights the signal region which shows the data is consistent with the background prediction.

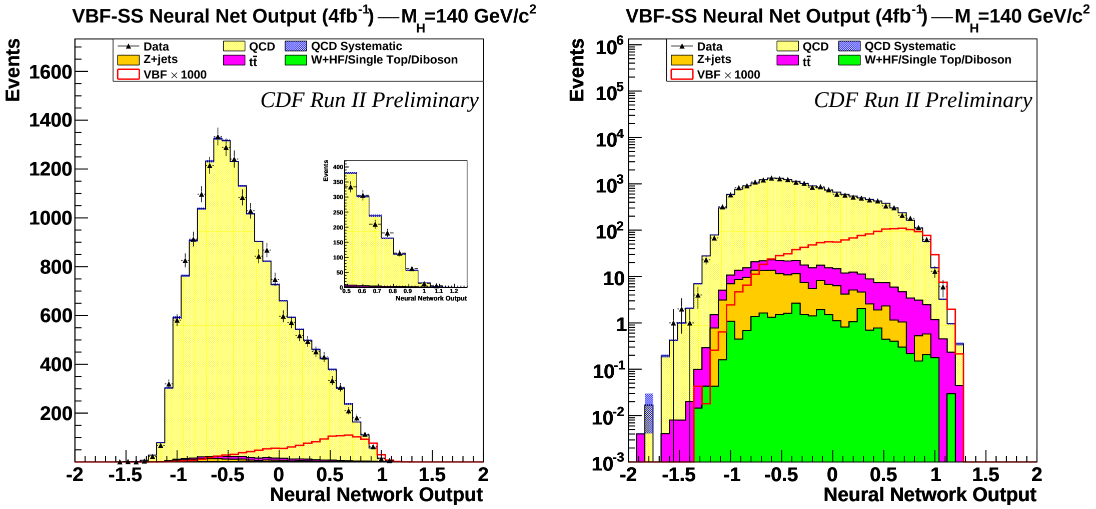


(a) VBF 135 SS Linear (b) VBF 135 SJ Log



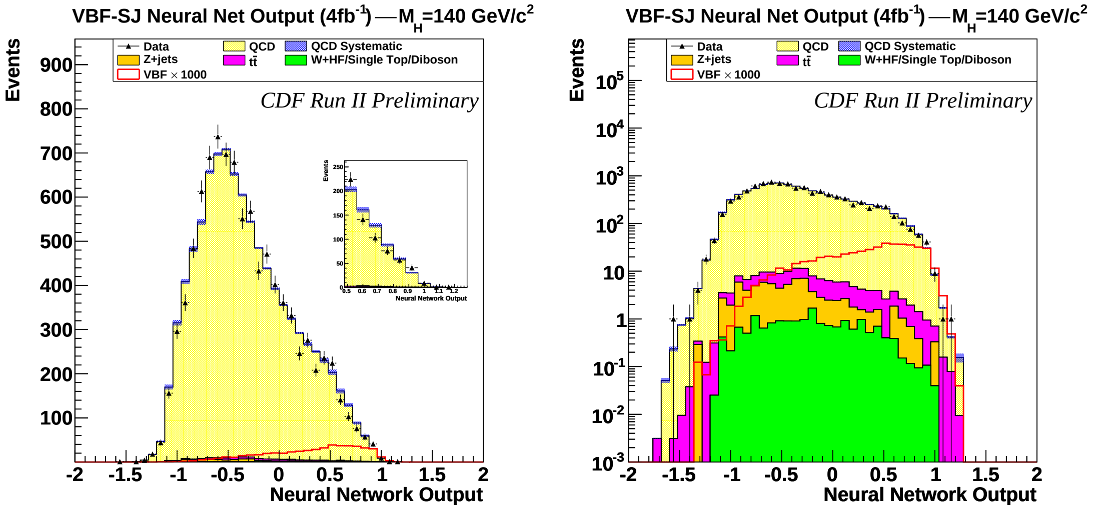
(c) VBF 135 SJ Linear (d) VBF 135 SJ Log

Figure 61: The Neural Net distributions for the VBF channel for Higgs masses of 135 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel. The inset in the linear scale plot highlights the signal region which shows the data is consistent with the background prediction.



(a) VBF 140 SS Linear

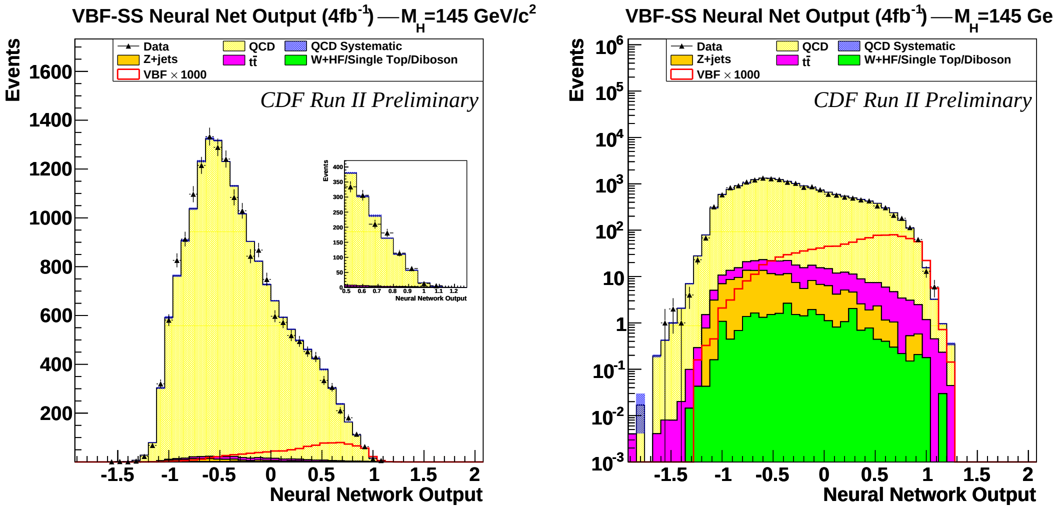
(b) VBF 140 SJ Log



(c) VBF 140 SJ Linear

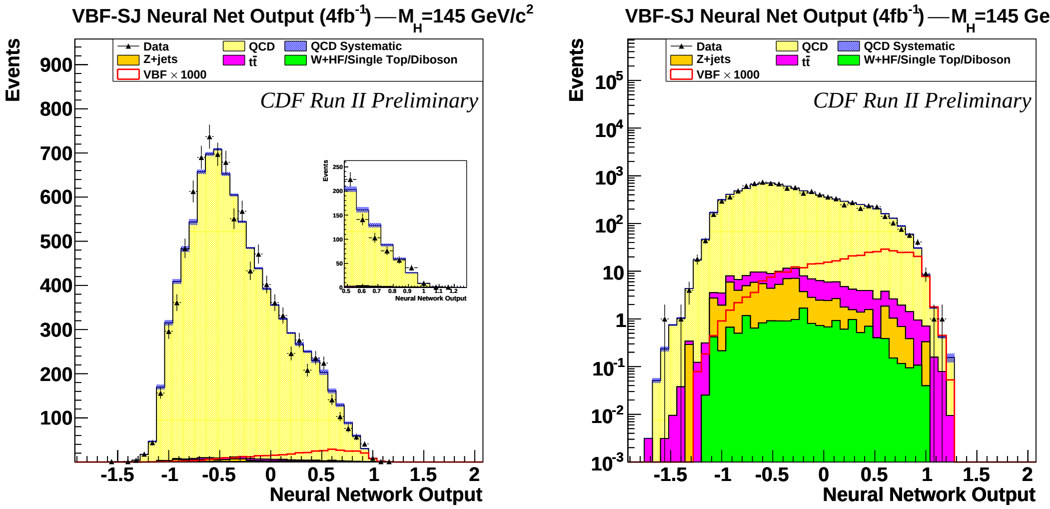
(d) VBF 140 SJ Log

Figure 62: The Neural Net distributions for the VBF channel for Higgs masses of 140 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel. The inset in the linear scale plot highlights the signal region which shows the data is consistent with the background prediction.



(a) VBF 145 SS Linear

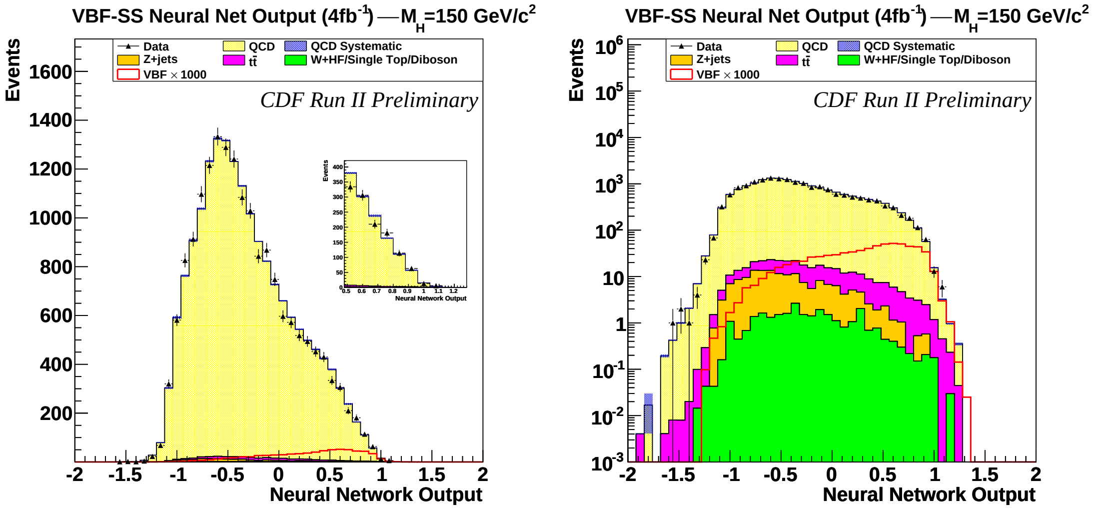
(b) VBF 145 SJ Log



(c) VBF 145 SJ Linear

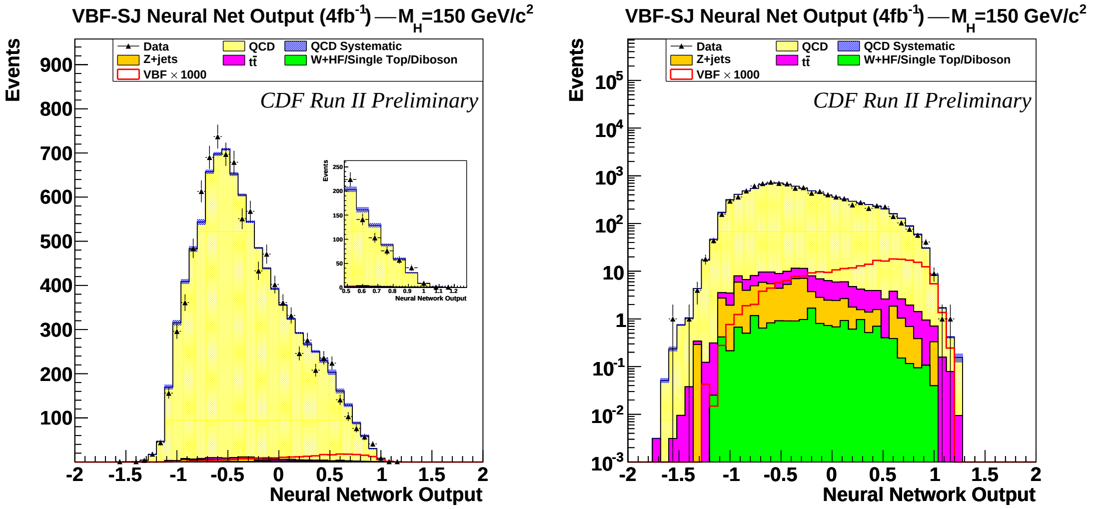
(d) VBF 145 SJ Log

Figure 63: The Neural Net distributions for the VBF channel for Higgs masses of 145 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel. The inset in the linear scale plot highlights the signal region which shows the data is consistent with the background prediction.



(a) VBF 150 SS Linear

(b) VBF 150 SJ Log



(c) VBF 150 SJ Linear

(d) VBF 150 SJ Log

Figure 64: The Neural Net distributions for the VBF channel for Higgs masses of 150 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel. The inset in the linear scale plot highlights the signal region which shows the data is consistent with the background prediction.

8 Results

As one can see from Figs. 47 and 58, no clear sign of a Higgs signal was observed in the data; 95% confidence limits (CL) are quoted. Table 8 has the limits from the combination of all 4 channels and is plotted in figure 65. The limits for each individual channel are quoted in tables 9, 10, 11 and 12 with plots of the limits in figure 66. All the limits in the tables are normalised to the expected Higgs signal cross-section.

Table 8: Summary of MCLimit for each Higgs mass from combining all VH & VBF channels

Higgs mass	-2σ	-1σ	Median	$+1\sigma$	$+2\sigma$	Observed
100	9.1	12.8	18.8	27.2	38.5	10.1
105	8.7	12.1	17.4	25.2	35.8	9.9
110	8.0	11.7	17.1	24.5	34.2	10.2
115	8.8	12.2	17.8	25.9	36.9	9.1
120	9.3	13.7	20.0	28.5	39.5	10.5
125	13.5	18.7	27.3	39.8	57.0	13.8
130	17.0	24.4	36.1	52.8	75.4	17.2
135	19.6	28.6	41.9	59.7	82.7	22.7
140	26.7	40.7	60.4	86.6	120.2	35.2
145	43.4	63.5	95.7	142.1	205.3	55.8
150	73.8	109.9	164.1	240.3	341.9	101.0

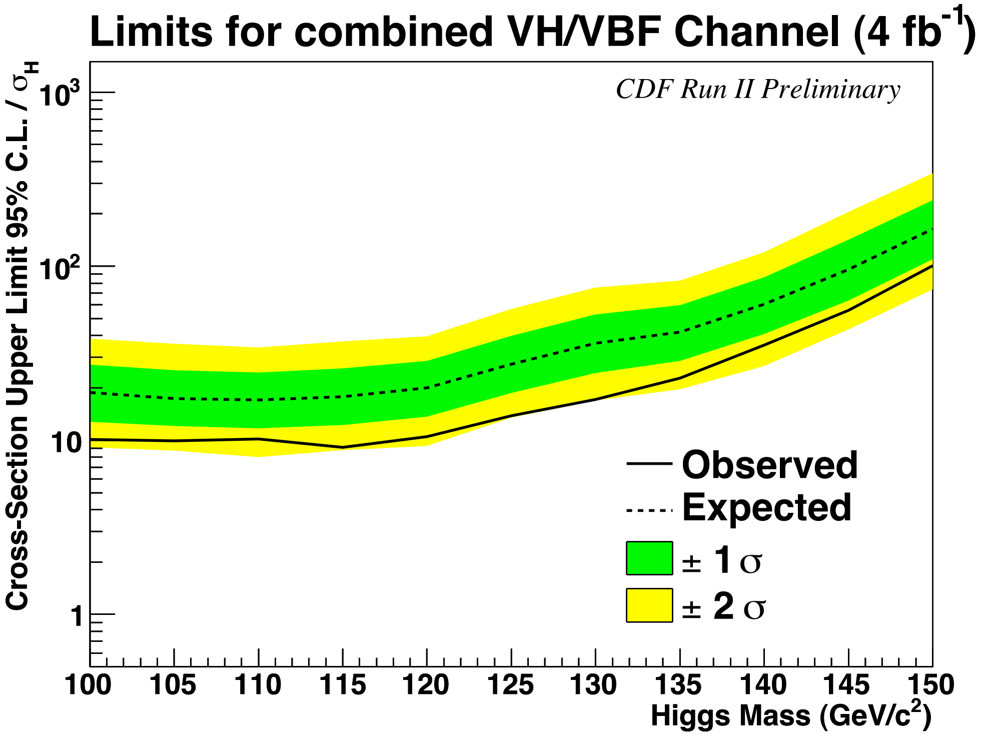


Figure 65: Limits for combined VH & VBF channels: The expected & observed limits are plotted as a function of the Higgs mass. The limits are normalised to the expected Higgs cross-section.

Table 9: Summary of MCLimit for each Higgs mass for VH-SS Channel

Higgs mass	-2σ	-1σ	Median	$+1\sigma$	$+2\sigma$	Observed
100	12.9	17.4	24.2	33.7	46.5	11.2
105	11.7	15.7	21.6	29.4	39.6	10.9
110	10.9	14.8	20.7	28.8	39.5	10.6
115	12.1	15.8	22.1	31.6	44.5	10.4
120	14.0	18.5	25.6	35.5	48.7	11.9
125	19.7	26.4	36.7	50.9	69.8	15.9
130	25.8	34.3	47.6	66.5	91.9	20.9
135	32.0	43.0	59.5	82.7	113.6	28.9
140	49.5	67.1	94.5	133.6	185.7	44.3
145	92.9	122.4	169.4	236.8	327.7	77.5
150	179.6	232.2	322.5	457.9	643.9	147.3

Table 10: Summary of MCLimit for each Higgs mass for VH-SJ Channel

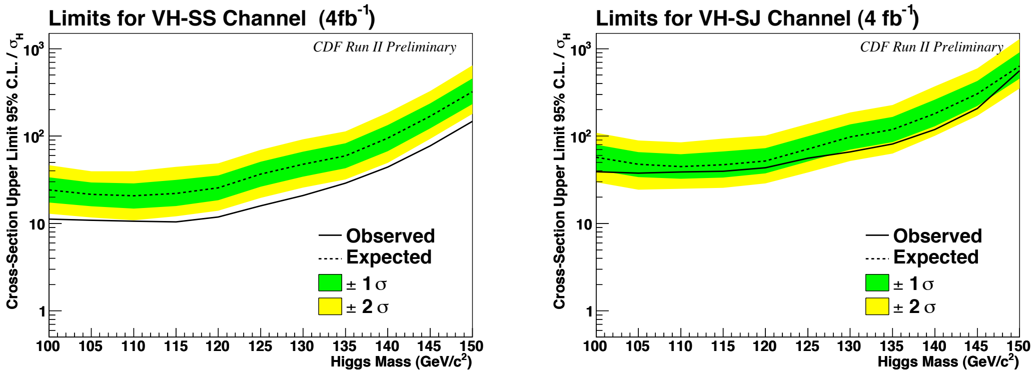
Higgs mass	-2σ	-1σ	Median	$+1\sigma$	$+2\sigma$	Observed
100	29.7	40.9	57.4	79.8	109.1	39.0
105	24.4	33.9	47.5	65.7	89.2	37.6
110	24.9	32.6	44.8	62.0	85.0	38.9
115	25.6	33.6	46.9	66.7	93.7	39.7
120	28.7	37.5	51.8	72.9	101.6	43.4
125	38.4	51.1	71.1	99.5	137.8	55.8
130	51.8	70.2	97.8	136.0	186.6	66.1
135	63.0	85.5	119.2	166.0	227.5	81.1
140	100.6	129.3	180.7	260.6	372.0	119.3
145	171.0	221.5	304.7	427.3	596.0	206.6
150	351.7	453.6	632.4	907.3	1291.1	569.1

Table 11: Summary of MCLimit for each Higgs mass for VBF-SS Channel

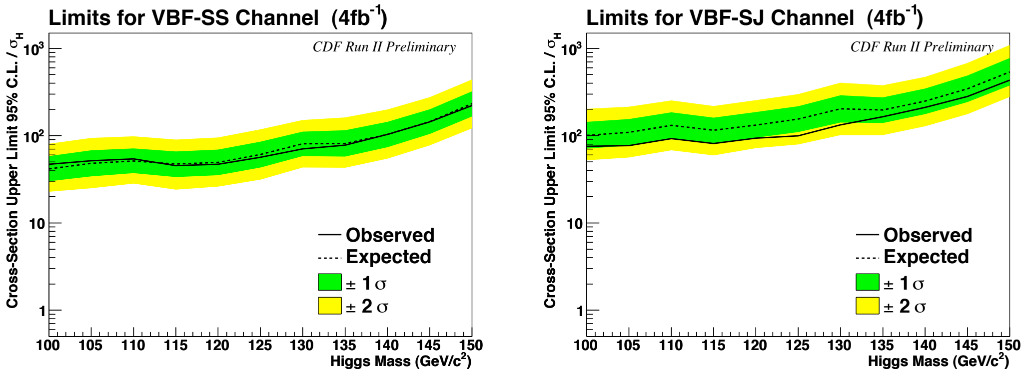
Higgs mass	-2σ	-1σ	Median	$+1\sigma$	$+2\sigma$	Observed
100	22.8	30.0	41.5	58.3	80.8	46.9
105	24.9	34.2	48.3	68.0	93.9	51.5
110	28.2	37.2	51.3	71.4	98.2	54.3
115	24.0	33.5	47.3	66.0	90.3	45.2
120	26.1	35.3	49.4	69.2	95.5	47.0
125	31.4	43.2	60.9	85.6	117.9	56.5
130	43.1	58.4	80.9	111.3	150.8	70.6
135	42.8	57.6	81.3	115.4	161.9	77.8
140	54.6	73.7	103.1	144.3	199.0	103.1
145	77.4	104.7	145.0	201.2	275.4	143.7
150	12311	166.1	231.9	321.7	439.6	220.8

Table 12: Summary of MCLimit for each Higgs mass for VBF-SJ Channel

Higgs mass	-2σ	-1σ	Median	$+1\sigma$	$+2\sigma$	Observed
100	52.1	70.8	100.7	143.9	202.5	75.5
105	56.1	77.0	109.3	155.0	215.2	76.7
110	67.9	92.6	130.9	184.9	253.0	92.5
115	59.3	82.0	115.2	160.4	219.3	81.2
120	71.7	95.0	132.3	186.0	254.2	93.7
125	79.5	110.0	155.2	217.3	299.5	99.7
130	101.1	140.9	202.5	289.2	403.9	132.5
135	101.6	139.5	196.7	275.2	378.7	164.3
140	128.2	176.8	248.6	345.8	472.5	210.5
145	175.8	242.0	344.7	489.3	681.9	280.5
150	277.3	374.8	537.9	773.5	1094.6	432.7



(a) VH-SS Limits (b) VH-SJ Limits



(c) VBF-SS Limits (d) VBF-SJ Limits

Figure 66: Higgs Limits for each channel. The expected & observed limit for each channel is calculated. The observed (measured) limit for each channel is consistent with background-only hypothesis to $\pm 2\sigma$. From the 4 channels, the VH-SS channel is the most sensitive.

9 Conclusions

A search for the Higgs boson in the all hadronic mode was conducted using 4 fb^{-1} of data. A Neural Network was used to separate the background events from the signal. As the presence of a Higgs signal was not observed, 95% confidence limits were calculated. The median expected limit for Higgs mass 120 GeV is at ~ 20 while the observed limit is at ~ 10.

10 Appendix

10.1 Z+jets Generator Level Filter

The large cross-section for Z+jets would produce an extremely large number of events; of which many would be rejected by the trigger. A filter was devised to select events which were likely to pass the trigger.

- At generator level, select events with ≥ 1 b or c parton.
- ≥ 3 jets with $E_T > 5\,\text{GeV}$. The jets were defined by cone sizes of 0.4, 0.7 and 1.0.
- The Sum- E_T for the 0.4, 0.7 and 1.0 jets are calculated. Accept the event if any of these sums $\geq 60\,\text{GeV}$

References

[1] Rong-Shyang Lu, Ankush Mitra, Song-Ming Wang, Aart Heijboer, Joe Kroll, and Daniel Whiteson. A search for the standard model higgs boson in the all-hadronic channel using a matrix element discriminant. CDF/ANAL/CDF/CDFR 9286, Academia Sinica and University of Pennsylvania and University of California,Irvine, April 2008.

[2] Rong-Shyang Lu, Ankush Mitra, Song-Ming Wang, Aart Heijboer, Joe Kroll, and Daniel Whiteson. A search for the higgs boson in the all hadronic channel using a boosted descision tree. CDF/ANAL/CDF/CDFR 9288, Academia Sinica and University of Pennsylvania and University of California,Irvine, June 2008.

[3] Yen-Chu Chen, Ankush Mitra, Shang-Yuu Tsai, Song-Ming Wang, and Rong-Shyang Lu. New multijet trigger. CDF/PHYS/TRIGGER/CDFR 9326, Institute of Physics, Academia Sinica, Taiwan and Department of Physics, National Taiwan University, Taiwan, 15 May 2008.

[4] Ankush Mitra, Shang-Yuu Tsai, and Song-Ming Wang. Study of multijet triggers for the all-hadronic higgs search. CDF/PHYS/TRIGGER/CDFR 9954, Institute of Physics, Academia Sinica, Taiwan, October 2009.

[5] Ankush Mitra. Cdf joint physics meeting: Calorimeter trigger energy scale, May 2009.

[6] T. Spreitzer, C Mills, and J Incandela. Electron identification in offline release 6.1.2. CDF/DOC/ELECTRON/CDFR/ 7950, University of Toronto and University of California,Santa Barbara, May 2006.

[7] Torbjorn Sjostrand, Leif Lonnblad, and Stephen Mrenna. PYTHIA 6.2: Physics and manual. 2001. hep-ph/0108264.

[8] Michelangelo L. Mangano, Mauro Moretti, Fulvio Piccinini, Roberto Pittau, and Antonio D. Polosa. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *JHEP*, 07:001, 2003.

[9] Sebastian Grinstein and D Sherman. Secvtx scale factors and mistag matrices for the 2007 summer conferences. CDF/DOC/SEC_VTX/CDFR/ 8910, Harvard University, July 2007.

[10] D Benjamin, M Herndon, E James, T Junk, B Kilminster, N Krumnack, and W.-M. Yao. Combined upper limit on standard model higgs boson production for HCP 2009. CDF/PUB/EXOTIC/PUBLIC 9999, 17 November 2009.

[11] Thomas Schwarz, Andrew Ivanov, and Robin Erbacher. Measurement of the top pair cross section in the lepton plus jets decay channel with 2.7 fb^{-1}. CDF/PUB/TOP/PUBLIC/ 9462, University of California, Davis, July 2008.

[12] A. Hocker et al. TMVA: Toolkit for multivariate data analysis. 2007.

[13] S. Geer and T. Asakawa. Analysis of multijet events produced at high energy hadron colliders. *Phys. Rev. D*, 53(9):4793–4805, May 1996.

[14] 2005. ISR/FSR Prescription : http://www-cdf.fnal.gov/internal/physics/joint_physics/agenda/20050527-minutes.html.

[15] Charles Plager. Pdf uncertainties. web. http://www-cdf.fnal.gov/internal/physics/joint_physics/instructions/PDFUncertainties/pdf.html.

[16] Oscar Gonzalez and Carsten Rott. Uncertainties due to the pdfs for the gluino-sbottom search. CDF/PHYS/EXOTIC/CDFR 7051, Purdue University, West Lafayette, IN, 47907, June 2004.