CDF note 10010

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

The CDF Collaboration
URL http://www-cdf.fnal.gov

We report the results of a search for the standard model Higgs boson, in which either the Higgs boson is produced together with a W or Z boson (associated production, $p\bar{p} \to W/ZH$), or produced via the fusion of two W or Z bosons (vector boson fusion, $p\bar{p} \to qqH$). The decay channels considered in this analysis are that the W and Z bosons in the associated production decay hadronically, and the Higgs boson decays into a $b\bar{b}$ pair. The final state signature consists of four or more jets, with at least two b-jets. The search is performed on 4 fb^{-1} of data recorded at CDF. The data are in agreement with the background model and we set 95% confidence limits on the Higgs production as function of the Higgs mass.

Preliminary Results for Winter 2010 Conferences

Contents

I. Introduction 3

II. Data Sample, Event Selections and Backgrounds 3
A. Signal Regions for the Analysis 4
B. Background and Signal 4
C. Expected Signal and Backgrounds 5

III. QCD Multi-jet Background Prediction 6
A. Tunning the Modeling of m_{qq} and Jet Shape Variables 6
B. Jet Shape Variables 10
C. Influence of the Higgs Signal to the TRF 10

IV. Neural Network Training 11

V. Systematics 14

VI. Results 15

VII. Conclusions 22

VIII. Appendix 22
A. Z+jets Generator Level Filter 22
B. Unblinded Signal Region for Other Higgs Masses 22

References 33

I. INTRODUCTION

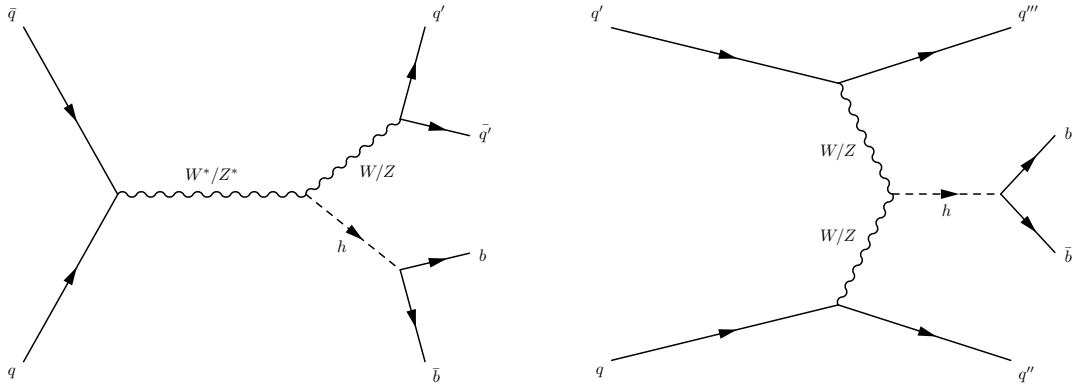
The Higgs boson plays a central role in the standard model (SM) as it endows particles with mass. At the Tevatron, the SM Higgs boson can be produced via several channels. The production channel which has the largest cross section is gluon fusion. For a low mass Higgs ($m_H \lesssim 135$ GeV/c^2), the Higgs boson decays primarily into a pair of b-quarks. Thus the final state signature for a low mass Higgs search via the gluon fusion process would be a pair of b-jets. But this particular search channel suffers from a huge background from QCD multi-jet production.

Alternative Higgs production processes such as associated production ($p\bar{p} \to W/ZH$) and vector boson fusion ($p\bar{p} \to qqH$) have production cross-sections about an order of magnitude lower than the gluon fusion process. But their decay kinematics allow one to improve upon the Higgs signal over the background significance.

Previous searches at CDF for a low mass Higgs boson have concentrated on final state signatures using combinations of leptons, jets and missing-Et. Recently a search for the Higgs boson in the all hadronic channel was performed using 2 fb^{-1} of CDF data [1]. In that analysis the Higgs boson was searched for in the associated production mode where $H \to b\bar{b}$ and $W/Z \to q\bar{q}$ as shown in Fig. 1(a). Thus the final state consists of at least four jets, with at least two b-jets.

This analysis extends the Higgs boson search in the all hadronic channel with a larger data sample of 4 fb^{-1} and the inclusion of several new innovations. The analysis now searches for the Higgs boson in two production channels: associated production as in the 2 fb^{-1} analysis (Fig 1(a)) and vector boson fusion (Fig. 1(b))

CDF is a general-purpose detector that is described in detail in reference [2]. The components relevant to this analysis are briefly described here. The charged-particle tracking system is closest to the beam pipe, and consists of a multi-layer silicon detectors (SVX) [3] and a large open-cell drift chamber covering the pseudorapidity region $|\eta| < 1$ [4]. The silicon detectors allow a precise measurement of a track's impact parameter with respect to the primary vertex in the plane transverse to the beam direction. The CDF central tracking region covers the pseudorapidity region $|\eta| \leq 1$ and is used to reconstruct charged particle momenta and the collision vertex. The tracking system is enclosed in a superconducting solenoid, which in turn is surrounded by a calorimeter. The CDF calorimeter system is organized into electromagnetic and hadronic sections segmented in a projective tower geometry, and covers the region $|\eta| < 3.6$. The electromagnetic calorimeters utilize a lead-scintillator sampling technology, whereas the hadron calorimeters use iron-scintillator technology. The central muon-detection system is located outside of the calorimeter and covers the range $|\eta| < 1$.



(a)Production and decay mode of the $W/ZH \to b\bar{b}qq$ channel

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

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

II. DATA SAMPLE, EVENT SELECTIONS AND BACKGROUNDS

The data for this analysis were collected by CDF's multi-jet triggers. They were designed to select events with multiple high E_T jets with large Sum-Et, which is the characteristic of an all-hadronic Higgs event. The first 2.8 fb^{-1} of CDF data were collected by a multi-jet trigger that requires at least four jets with $E_T > 15$ GeV and Sum-Et> 175 GeV. The remaining 1.1 fb^{-1} data was collected by another trigger that requires at least three jets with $E_T > 20$ GeV and Sum-Et> 130 GeV. The new trigger was designed to increase the acceptance for low mass Higgs boson.

The events selected for the analysis have to pass two set of selection criteria which are defined below.

- Initial Selection
  - The event must pass the multijet trigger
  - The detector components essential to the analysis are operating and producing good data during the data taking period.
  - The event's reconstructed primary vertex to be in the luminous region ($|V_Z| < 60$ cm, where V_Z is the Z position of the reconstructed primary vertex along the beam axis.
  - The event should have at least four or five jets with $E_T > 15$ GeV and within the pseudorapidity region $|\eta| < 2.4$. Jets are reconstructed from the calorimeter towers using a cone algorithm with fixed radius $\Delta R \equiv \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.4$ in $\eta - \phi$ space [5]. The jet E_T measurements are corrected for detector effects [**?**].
  - The missing transverse energy significance [6] should be small to reduce the contribution from the top quark pair production background.
- Final Selection
  - Two of the four leading jets are required to be identified/tagged as b-jets. CDF uses two algorithms to identify b-jets: SecVtx [7] and Jet-Probability [8]. For this analysis we consider two combinations. Both b-jets are tagged by SecVtx (SS) or one b-jet is tagged by SecVtx and the other by Jet-Probability (SJ). Jets from b partons are characterized by their secondary decays which are displaced from the primary vertex; thus their tracks have a large impact parameter. Light-flavor jets appear to come from the primary interaction and their tracks' impact parameter are consistent with the primary vertex (within the resolution of the tracking detector).

    The secvtx algorithm looks for tracks within jets that form a secondary vertex significantly displaced from the primary interaction point. The jet-probability algorithm examines the impact parameter of the tracks within a jet and computes a probability that the jet is a b jet.
  - The event 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.

A. Signal Regions for the Analysis

In this analysis we are searching for a Higgs boson that is produced through the associated production channel ($p\bar{p} \to W/Z + H$) or the vector boson fusion channel ($p\bar{p} \to qqH$). For the associated production channel, there are two di-jet 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 ($W \to qq'$, $Z \to q\bar{q}$). In the vector boson fusion channel there is only one di-jet 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 (lab) frame, forming jets very far apart from each other, which result in a very board di-jet mass distribution. In this analysis, we assume that the two b-tagged jets are from the Higgs boson and other two non b-tag jets (among the first four leading jets) are from the W or Z boson decay in the associated production channel, or from the two out-going quarks in the vector boson fusion channel.

The invariant mass of the two b-tag jets (M_{bb}) and the invariant mass of the two non b-tag jets (M_{qq}) are used to define two signal regions in the M_{bb} vs M_{qq} mass plane, as shown in Fig. 2. 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 low mass Higgs boson with mass between 100 to 150 GeV/c^2.

B. Background and Signal

The backgrounds that have a similar final state signature as the all hadronic Higgs signal are QCD multi-jet production, top quark pair production, single top quark production, W plus heavy-flavor jets production, Z plus

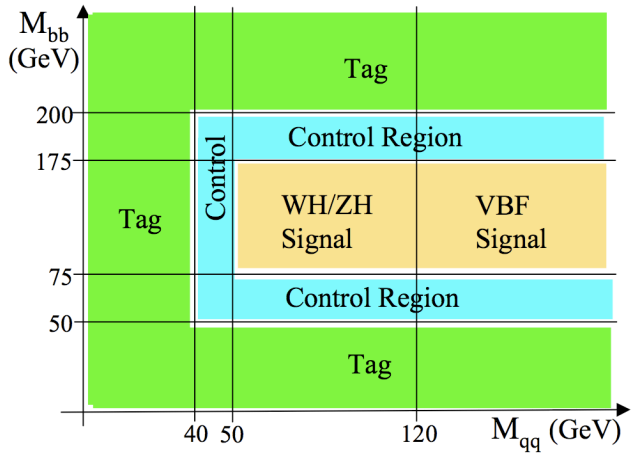


FIG. 2: The definitions of TAG, CONTROL and signal regions in the M_{bb} vs M_{qq} mass plane.

jets production where Z decays into a pair of b or c quarks, and diboson productions (WW, WZ, ZZ). The QCD multi-jet production is the dominant background as it consists of about 98% of the total background Its contribution is estimated from a data driven based technique, which is described in later section of the paper. The non-QCD backgrounds are estimated from simulation. The W plus heavy-flavor jets production is simulated with the ALPGEN [9] generator to simulate the W boson plus parton production, with PYTHIA [10] used to model parton showers. The other non-QCD backgrounds and the SM Higgs signal samples were simulated with the PYTHIA generator. All the non-QCDbackground samples are generated using the CTEQ5L PDFs [11].

C. Expected Signal and Backgrounds

The number of signal and background events which are used in the analysis, after passing the trigger and event selection, are given in tables I and II.

TABLE I: *CDF Run II Preliminary* (4 fb^{-1}) Expected number of signal events which pass the trigger, event selection, have two b-tags (SecVtx-SecVtx [SS] or SecVtx-SetProbJ [SJ]).

Signal regions	WH		ZH		VBF	
Higgs Mass (GeV/c^2)	SS	SJ	SS	SJ	SS	SJ
100	5.9	2.2	4.4	1.5	3.4	1.2
110	5.4	2.0	4.0	1.4	3.6	1.3
120	4.5	1.7	3.3	1.2	3.2	1.2
130	3.3	1.2	2.3	0.8	2.5	0.9
140	1.9	0.7	1.3	0.5	1.6	0.6
150	0.8	0.3	0.6	0.2	0.8	0.3

TABLE II: *CDF Run II Preliminary* (4 fb^{-1}) Expected number of background events which pass the trigger, event selection and have two b-tags (SecVtx-SecVtx [SS] or SecVtx-SetProbJ [SJ]).

Signal regions	VH (WH+ZH)		VBF	
Backgrounds	SS	SJ	SS	SJ
$t\bar{t}$	239	98	150	64
single-top	36.2	14.6	14.4	8.4
$W + bb/cc$	27.9	12.0	4.8	3.3
$Z(bb, cc)$+jets	128	55.4	135	63
Diboson	10.1	7.5	4.7	3.4
Data	16857	9341	17776	9518

III. QCD MULTI-JET BACKGROUND PREDICTION

To estimate the QCD multi-jet background, a model is constructed from the data using the background-dominated sample with only one jet tagged by the SECVTX algorithm (one-tagged events) Each of the additional jets, called a probe jet, is weighted by the probability for it to receive a b-tag, called the tag rate function (TRF). For an event to contribute to the background model in the signal region, the invariant mass of the tagged jet and the probe jet must be in the range $75 < m_{bb} < 175$ GeV/c^2 and the mass of the other two leading jets, m_{qq}, in the event must exceed 50 GeV/c^2, as shown in Fig 2. The data used to derive the TRFs come from examining the data outside the signal region in the M_{bb}-M_{qq} plane where at least one jet is tagged by the SECVTX algorithm. Two regions are defined: TAG and CONTROL. The default TRF was measured on the data from the TAG region. The additional CONTROL region is used to derive the systematic errors. The tag rate is measured as a function of three variables: the p_T of the proble jet, ΔR between the probe jet and the tagged jet, and the pseudorapidity of the probe jet ($|\eta|$). The TRF for SS and SJ b-tag categories are measured separately. For the SJ category, events with only one tight SecVtx tagged jets are considered in the measurement of the TRF(SJ). 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 neural net training to separate the Higgs signal from the QCD background. This TRF method does not necessarily predict the right normalization of the double b-tagged QCD background.

The TRF is applied to exactly one-tagged events in the 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,etc).

$$\begin{aligned}\text{Predicted two tag data} &= \text{one tagged Data} \times TRF \\ &= [\text{one tagged } QCD + \text{one tagged } t\bar{t} \\ &\quad + \text{one tagged } Z + \text{jets} + \cdots] \times TRF \end{aligned} \tag{1}$$

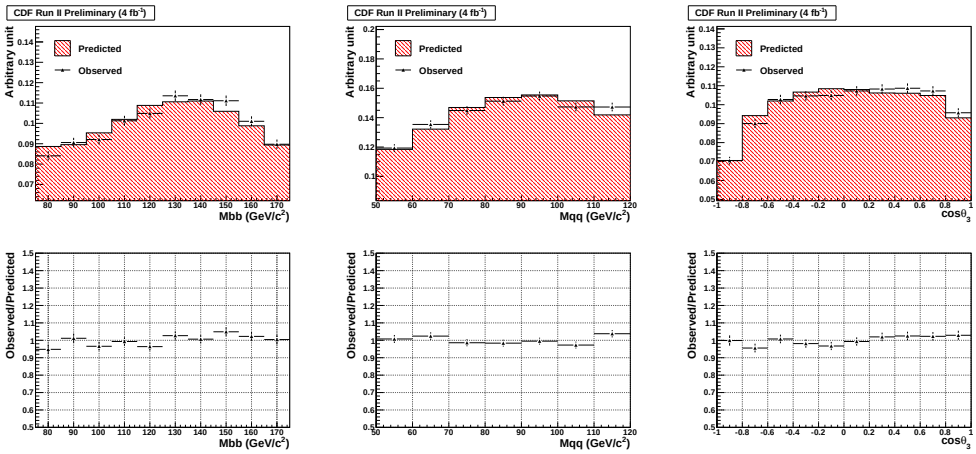
As we are using the simulation to predict the double b-tagged non-QCD background, there will be some double counting of the non-QCD events when we apply the TRF onto the one-tagged data. This 'double-count' of the non-QCD contribution is removed by subtracting the one-tagged non-QCD background contributions from the prediction. As $t\bar{t}$ 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 the $t\bar{t}$ and Z plus jets backgrounds.

$$\begin{aligned}\text{QCD Double Tagged Background} &= TRF \times (\text{one-tagged data}) \\ &\quad - TRF \times (\text{one-tagged } t\bar{t}_{MC}) \\ &\quad - TRF \times (\text{one-tagged } Z + \text{jets}_{MC}) \end{aligned} \tag{2}$$

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 Figs. 3-7. The shapes from the prediction agree quite well with the observed shapes. The plots in Figs. 3-5 are for the events in the WH/ZH signal region, and the plots in Figs. 6-7 are for events in the VBF signal region. The definitions for the variables shown in these plots are given in Section IV.

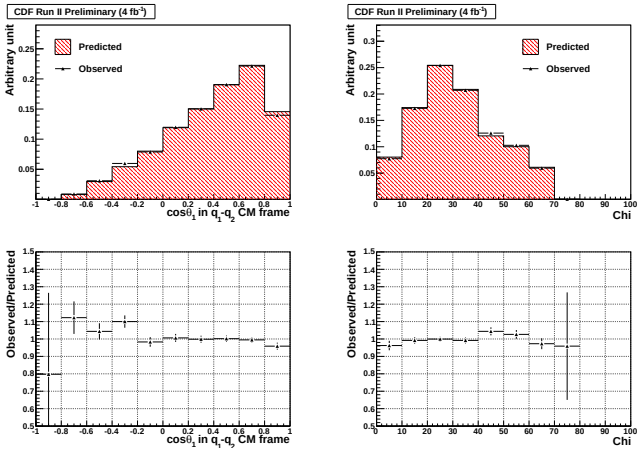
A. Tunning the Modeling of m_{qq} and Jet Shape Variables

The TRF is explicitly not parameterised as a function of m_{qq} and the jet shape variables η-moment and ϕ-moment (described in section III B). As the TRF does not model the correlations between these variables, there is a small shape difference between the prediction & observation for these variables. A correction is applied by scaling the prediction by the ratio of the observed m_{qq} and η/ϕ-moment distributions to the predicted distributions in the *tag* region but with $m_{qq} > 50$ GeV/c^2. As a systematic check, an alternate scaling correction to the TRF is obtained by repeating the study in the CONTROL region with $m_{qq} > 50$ GeV/c^2. The corrections are smooth functions of m_{qq} η-moment and ϕ-moment. For η-moment and ϕ-moment the correction is in the order of 5%, whereas for m_{qq}, the correction decreases with m_{qq} from $\sim +5\%$ at $m_{qq} = 50$ GeV/c^2 to $\sim -25\%$ at $m_{qq} = 500$ GeV/c^2.



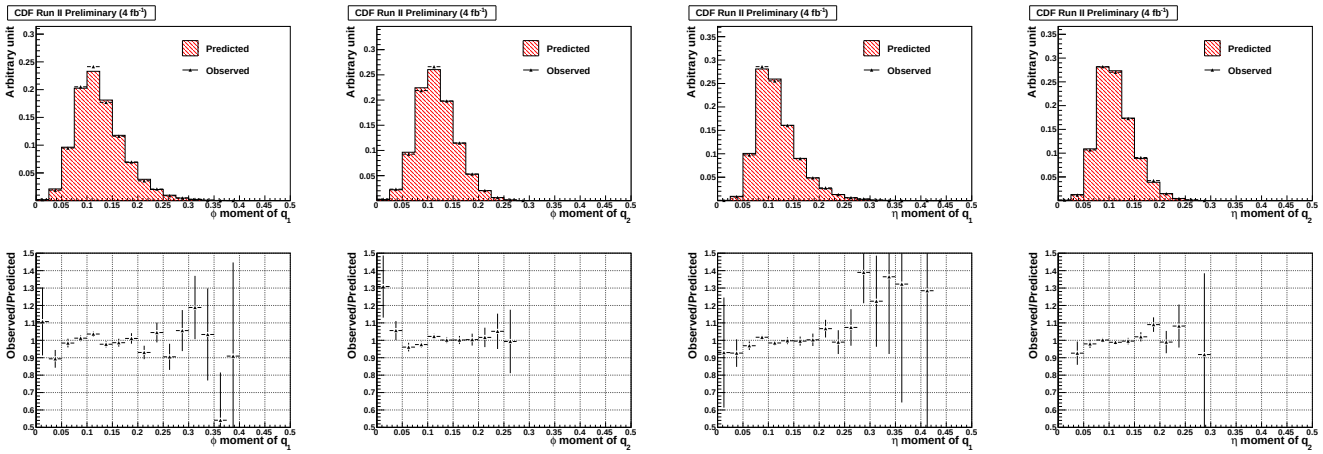
(a)M_{bb} (b)M_{qq} (c)$\cos(\theta_3)$

FIG. 3: 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.



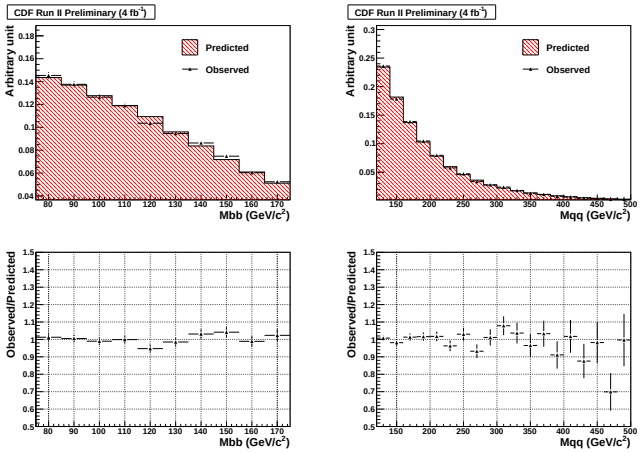
(a)Helcity Angle (b)Chi

FIG. 4: 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.



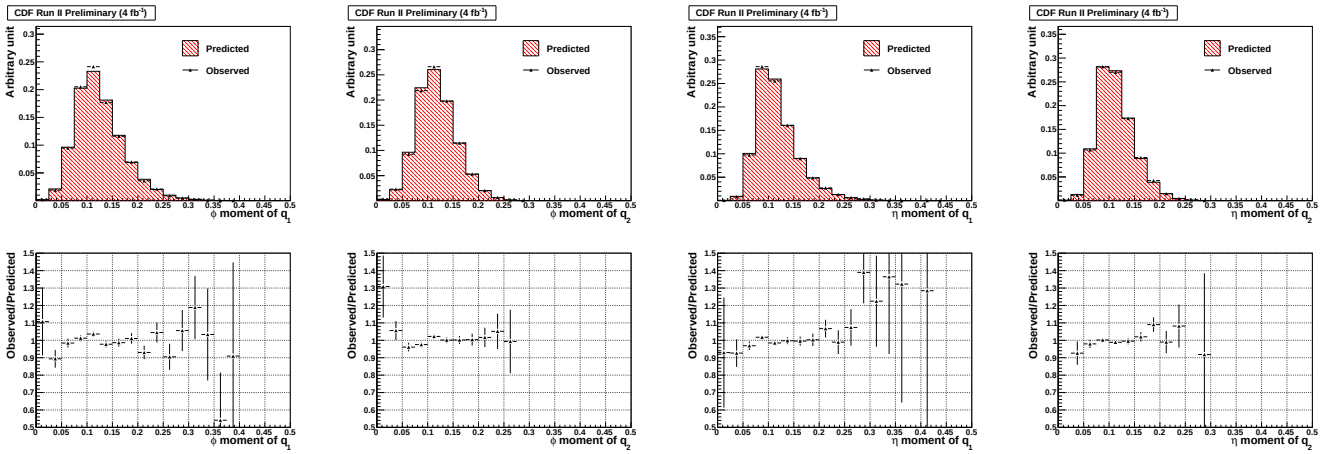
(a)ϕ moment of q_1 (b)ϕ moment of q_2 (c)η moment of q_1 (d)η moment of q_2

FIG. 5: 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.



(a)M_{bb} (b)M_{qq}

FIG. 6: 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.



(a)ϕ moment of q_1 (b)ϕ moment of q_2 (c)η moment of q_1 (d)η moment of q_2

FIG. 7: 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.

B. Jet Shape Variables

The multi-jet QCD background are a mixture of quark and gluon jets. Whereas the Higgs signal, the jets not from the HIggs boson decay are light flavor quark jets. The width of light-flavor quark jets are usually smaller than jets from a b-quark or gluon. Thus the jet's width is used in separating the Higgs signal from the QCD multi-jet background. The two jet shape variables which measure the jet's width are:

$$\phi\text{-moment}(\langle\phi\rangle) = \sqrt{\sum_{\text{towers}} \left(\left(\frac{E_t^{\text{tower}}}{E_t^{\text{jet}} \phi_{\text{tower}}} \right)^2 - \phi_{\text{jet}}^2 \right)} \tag{3a}$$

$$\eta\text{-moment}(\langle\eta\rangle) = \sqrt{\sum_{\text{towers}} \left(\left(\frac{E_t^{\text{tower}}}{E_t^{\text{jet}} \eta_{\text{tower}}} \right)^2 - \eta_{\text{jet}}^2 \right)} \tag{3b}$$

The towers are the calorimeter towers that are used to form the jets.

The simulation of jet shape variables for light-flavor quarks was validated by studying the jets from $t\bar{t}$ production events in the decay mode $t\bar{t} \rightarrow bW^+\bar{b}W^- \rightarrow bl^+\nu b^- qq'$; assuming the non b-tagged jets are from decays of W boson. The selected $t\bar{t}$ candidate events in the data consist of $\sim 86\%$ from $t\bar{t}$ production, $\sim 7\%$ from Wbb and Wcc productions. The remainder are Wc, non-W, mis-tag (wrong identification of b-jets), Z+jets, single-Top quark, and diboson productions. The width of the jet shapes show dependencies on the jet's transverse energy, pseudorapidity, and the number of reconstructed vertices in the event. These dependencies are parameterized separately in the data and simulation to remove these dependencies. The plots in Fig. 8 are the η/ϕ-moment distributions after the dependencies removal. The $t\bar{t}$ is normalized to 86% of the total contribution and the Wbb+Wcc is normalized to the remaining 14% contribution. The jet shape distributions for light-flavor quarks show good agreement between data and simulation.

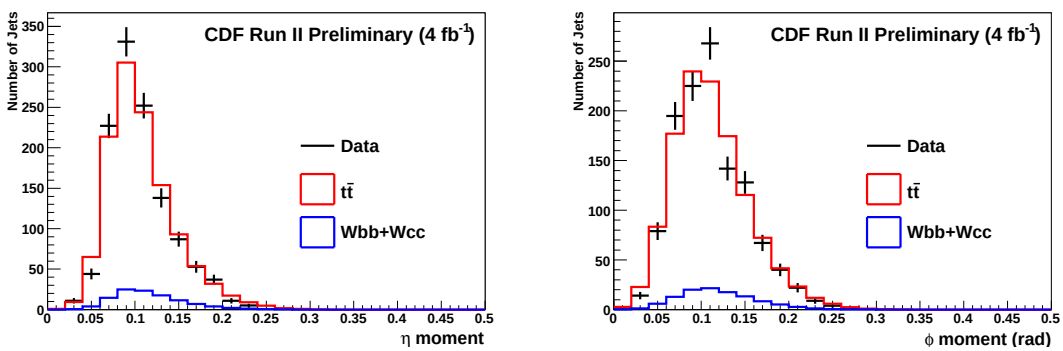


FIG. 8: η-moment and ϕ-moment distributions between data and simulation of non-b tagged jets in $t\bar{t}$ candidate events.

C. Influence of the Higgs Signal to the TRF

If a Higgs signal existed, 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. The ($n\times$ Higgs signal) contribution is removed from the background prediction using the same prescription to remove the non-QCD sources.

$$\begin{aligned}(n \times \text{Higgs Signal}) + \text{Background} &= n(\text{2 Tag Higgs data}) \\ &\quad + TRF \times (\text{1-Tag Data - } n \text{ 1-Tag Higgs}) \\ &= n(\text{2 Tag Higgs - } TRF \times \text{ 1-Tag Higgs}) \\ &\quad + (\text{TRF 1-Tag Data})\end{aligned} \tag{4}$$

By correcting for the 1-Tag Higgs contribution, the Higgs signal is reduced by its 1-Tag contribution.

IV. NEURAL NETWORK TRAINING

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. For this analysis, we considered an Artificial Neural Network (NN) [12]. As the background is dominated by QCD, the 1-Tag background, weighted by the TRF, is used as the background sample for the NN.

For the analysis, we searched for a Higgs boson between 100 GeV/c^2, to 150 GeV/c^2, in 10 GeV/c^2, steps. The Neural Net was trained at three target Higgs masses: 100 GeV/c^2, 120 GeV/c^2 and 140 GeV/c^2. The trained Neural Nets are applied to their target Higgs mass and the next higher mass point.

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

A separate neural net is trained for the VH and VBF channels. The VH neural net training was done in the VH signal region defined as:

$$\text{VH Signal Window} \;:\; \begin{aligned} 75 < M_{bb} < 175\text{GeV} \\ 50 < M_{qq} < 120\text{GeV} \end{aligned} \tag{5}$$

and the VBF Neural Net training was done in the VBF region:

$$\text{VBF Signal Window} \;:\; \begin{aligned} 75 < M_{bb} < 175\text{GeV} \\ M_{qq} > 120\text{GeV} \end{aligned} \tag{6}$$

After experimenting with combinations of variables for the training, we decided on the follwoing set of variables for the VH training:

- 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 ($\cos\theta^*_{q1}$) [13]
- cosine of the leading-jet scattering angle in the 4-jet rest-frame ($\cos(\theta_3)$) [14]
- Chi which is defined as:

$$Chi = \text{Min}(Chi_W, Chi_Z) \tag{7}$$

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

- Eta-Moment of leading non b-jet ($\langle\eta\rangle_1$)
- Phi-Moment of leading non b-jet ($\langle\phi\rangle_1$)
- Eta-Moment of second leading non b-jet ($\langle\eta\rangle_2$)
- Phi-Moment of second leading non b-jet ($\langle\phi\rangle_2$)

and the following set of variables for the VBF training:

- Mass of the two b-tagged jets (M_{bb})
- Mass of the two non b-tagged jets (M_{qq})
- Eta-Moment of leading non b-jet ($\langle\eta\rangle 1$)
- Phi-Moment of leading non b-jet ($\langle\phi\rangle_1$)
- Eta-Moment of second leading non b-jet ($\langle\eta\rangle_2$)
- Phi-Moment of second leading non b-jet ($\langle\phi\rangle_2$)

Figs. 9 and 10 show the signal and background plots for the selected variables for the VH and the VBF channel respectively.

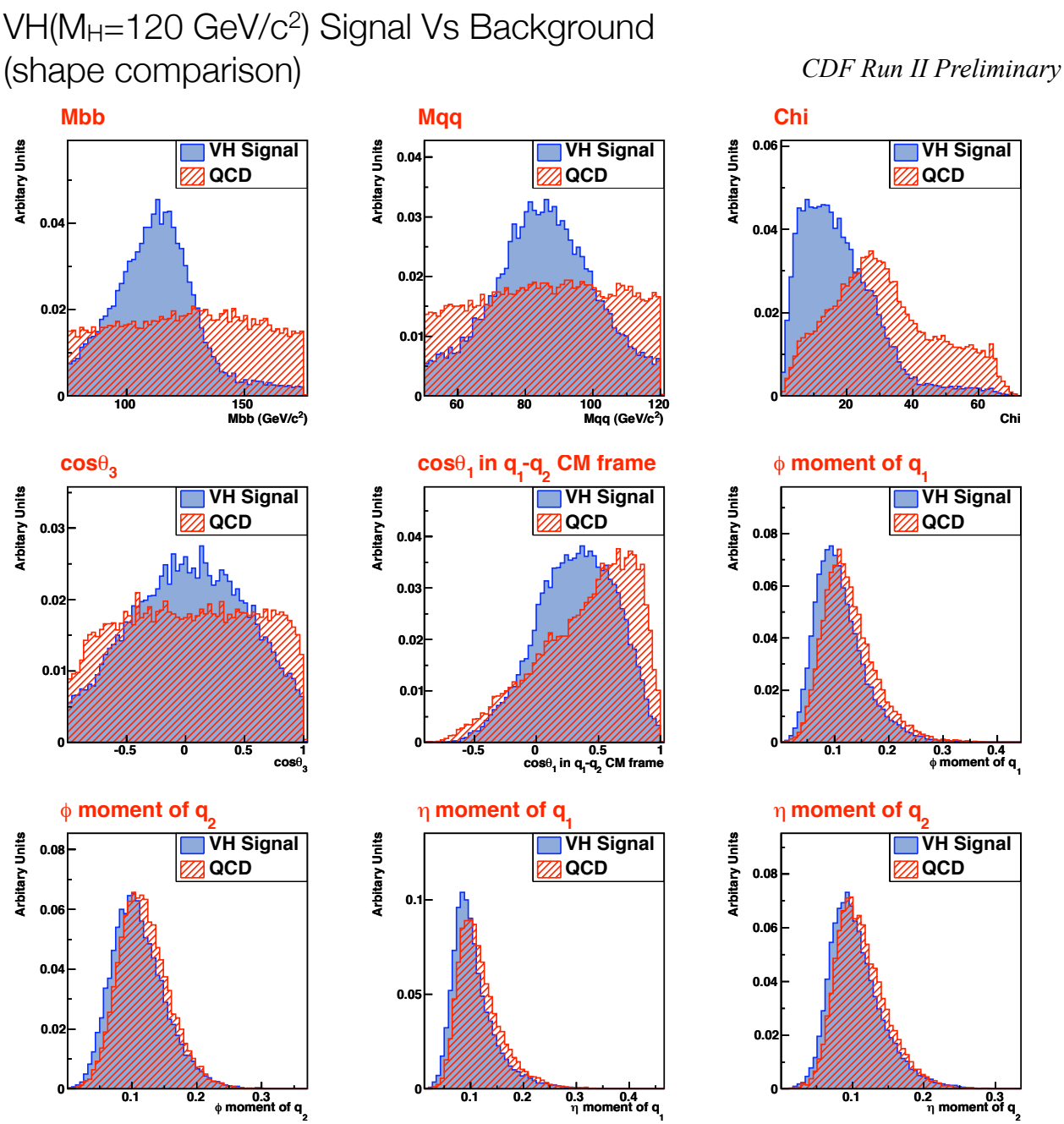


FIG. 9: The training variables used for the VH analysis. VH(MH120) is used as the signal. The VH signal and QCD background have been normalised to unit area.

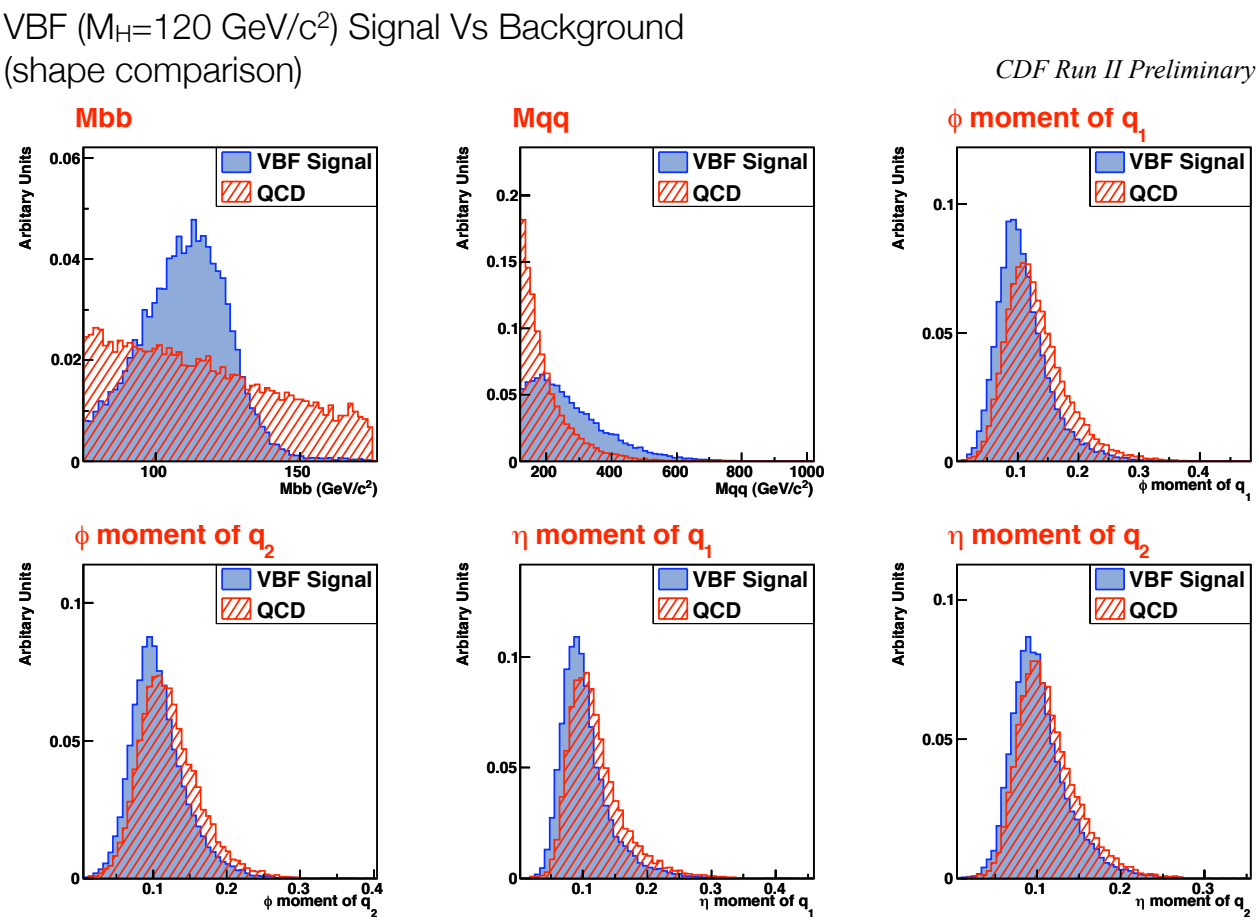


FIG. 10: The training variables used for the VBF analysis. VBF(MH120) is used as the signal. The VBF signal and QCD background have been normalised to unit area.

V. SYSTEMATICS

We consider systematics which affect the shapes and normalisation of the Neural Net distributions for the signal and background. The systematics which are considered are:

- QCD Modelleing, which is divided into three categories:
  - Interpolation Uncertainty
  - M_{qq} Tuning Uncertainties
  - Jet Shape Tuning Uncertainties
- Jet Energy Scale
- Initial and fnal state radiation (ISR/FSR).
- Luminosity
- b-tag scale factor
- Parton distribution uncertainties.

The nominal background predicition uses the TRF derived from the TAG region which is interpolated into the signal region. The systematic error for this interpolation is taken from the background prediction using the TRF derived from the CTRL region. The difference between these two background shapes is used as the interpolation shape systematic.

Section III A discussed the tuning required for TRF prediction of m_{qq} and the jet shapes. The tuning used functions derived from the TAG region and CTRL region. The nominal correction used the TAG derived tuning function. As a measure of the systematic error, the CTRL region tuning function is used. The difference in the neural net output shape from using the two tuning functions is used as the m_{qq} & jet-shapes shape systematic.

The Jet Energy Scale affects the shape of the Jet E_t related quantities. The training variables which are affected most are M_{bb} and M_{qq}. Although the Jet Energy Scale systematics affects the M_{bb} and M_{qq} distributions, the Neural Net shape is less affected. The Jet Shape Systematic also affect the acceptance on account of the SumEt>220.0 GeV event selection cut, which gives rise to a ±7% rate systematic error in addition to the shape errors.

The initial and final state radiation systematics also include a shape systematic as well as a ±2% rate systematic for the VH and ±3% for the VBF Higgs signal.

Other rate systematics include a ±2% PDF rate systematic and the 6% luminosity uncertainty. For B-Tagging, all MC SecVtx-SecVtx events have to apply a scale factor of 0.895 with a 7.6% rate systematic. For the SecVtx-JetProb events, the MC scale factor is 0.699 with a 9.7% rate systematic. Finally there is a 4% rate systematic for all signal samples according to our trigger study.

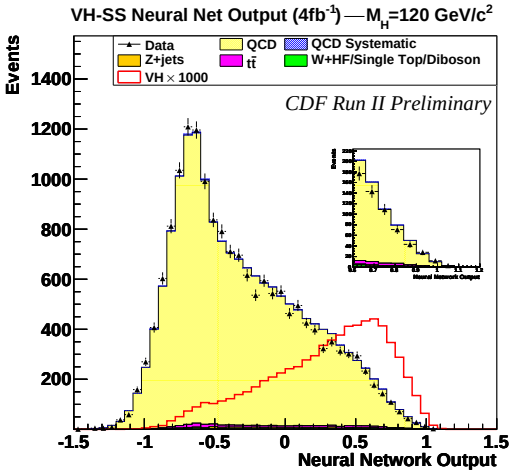
Table III summarises all the rate uncertainties and shape systematics which are applied in this analysis.

TABLE III: Summary of all Systematic Uncertainties used to calculate the limit

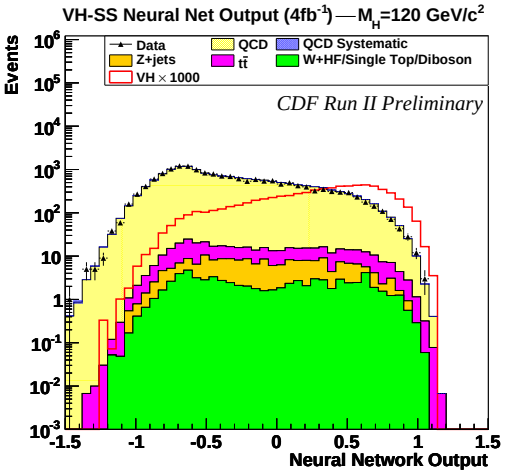
Systematics	Rate or Shape (or both) systematic
Jet Energy Correction	±7% Rate Shape
PDF	± 2% Rate
B-tagging (SecVtx+SecVtx)	7.6% Rate
B-tagging (SecVtx+JetProb)	9.7% Rate
Luminosity	6% Rate
ISR/FSR	± 2% for VH Rate ± 3% for VBF Rate Shape for VH & VBF
Trigger	± 4% Rate
Interpolation	Shape
M_{qq} Tuning	Shape
Jet Shape Tuning	Shape

VI. RESULTS

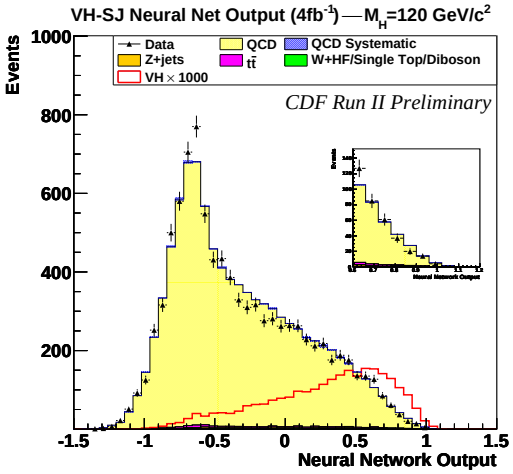
Figs. 11-12 show the unblinded data for the 120 GeV Higgs for the VH and VBF channels respectively. Over the whole neural net range, the data is in good agreement with the expected background with no clear evidence of a Higgs signal.



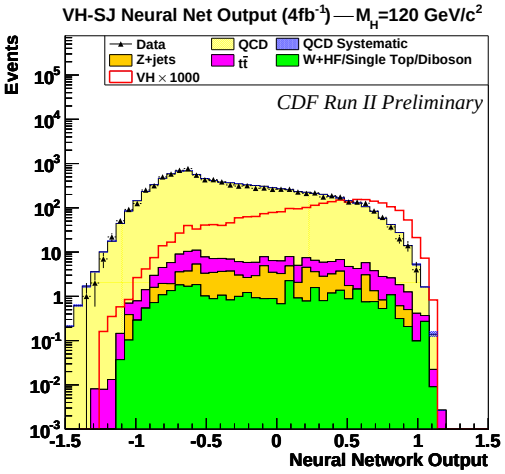
(a)VH120 SS Linear



(b)VH120 SS Log

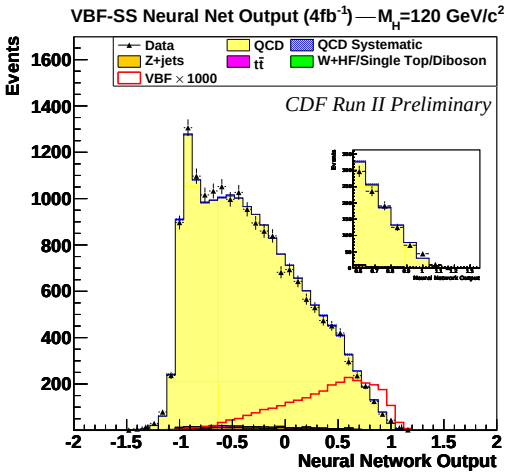


(c)VH120 SS Linear

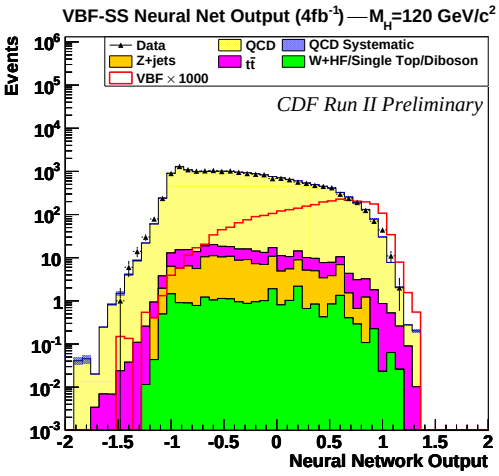


(d)VH120 SS Log

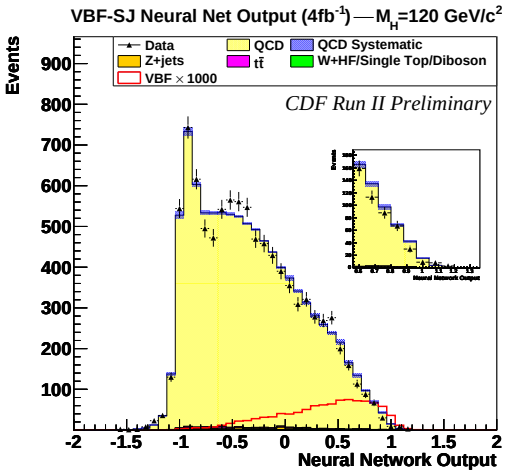
FIG. 11: The Neural Net distributions for the VH channel for Higgs masses of 120 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel.



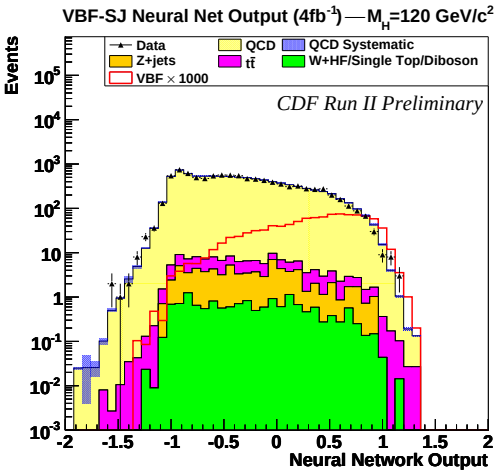
(a)VBF120 SS Linear



(b)VBF120 SS Log



(c)VBF120 SS Linear



(d)VBF120 SS Log

FIG. 12: 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.

Since no clear sign of a Higgs signal was observed in the data; 95% confidence limits (CL) are quoted. Table IV has the limits from the combination of all 4 channels and is plotted in Fig. 13. The limits for each individual channel are quoted in tables V-VIII with plots of the limits in Fig. 14- 17. All the limits in the tables are normalised to the expected Higgs signal cross-section.

TABLE IV: Expected and observed 95% CL upper limits for the combined VH and VBF channels using 4fb^{-1} of $p\bar{p}$ data collected by the CDF detector. The limits are normalised to the expected Higgs cross-section.

Higgs mass (GeV/c^2)	-2σ	-1σ	Median (σ/σ_H)	$+1\sigma$	$+2\sigma$	Observed
100	9.2	13.0	18.7	26.6	37.3	10.2
110	8.5	11.6	16.7	24.2	34.4	10.3
120	10.0	13.8	19.9	28.4	39.6	10.4
130	18.1	24.3	34.8	50.9	73.2	17.3
140	28.2	41.4	60.2	85.8	119.0	34.0
150	73.0	108.7	161.7	235.2	331.7	111.2

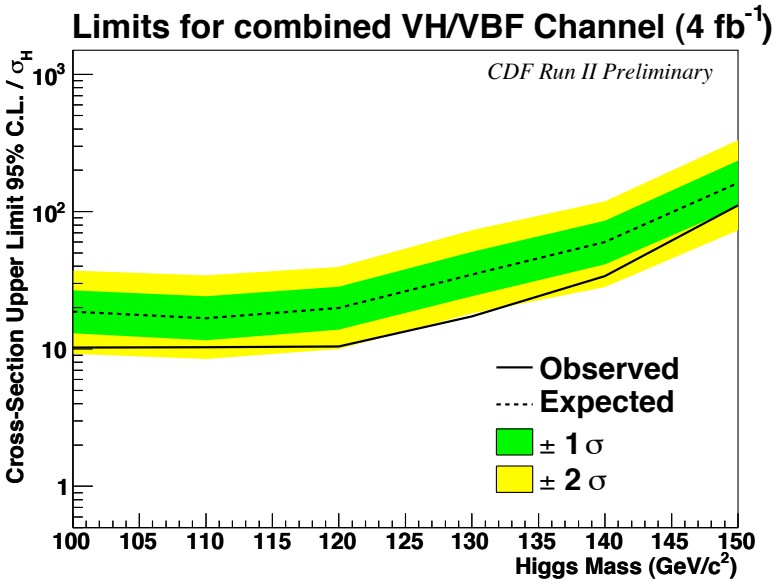


FIG. 13: Limits for combined VH and VBF channels: The expected and observed limits are plotted as a function of the Higgs mass. The limits are normalised to the expected Higgs cross-section.

TABLE V: Expected and observed 95% CL upper limits for the VH-SS channel using 4fb^{-1} of $p\bar{p}$ data collected by the CDF detector. The limits are normalised to the expected Higgs cross-section.

Higgs mass (GeV/c^2)	-2σ	-1σ	Median (σ/σ_H)	$+1\sigma$	$+2\sigma$	Observed
100	13.3	17.8	24.1	32.7	43.7	11.6
110	10.9	14.6	19.9	26.9	36.1	11.0
120	14.1	18.3	25.3	35.5	49.4	12.3
130	25.9	34.0	46.9	65.3	90.1	21.5
140	52.9	69.0	94.6	131.7	181.9	46.2
150	175.0	231.5	321.0	449.1	622.3	153.4

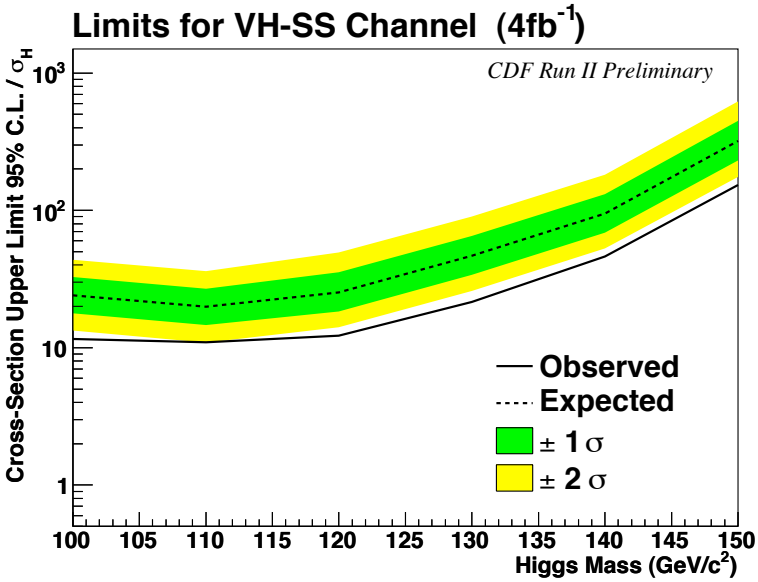


FIG. 14: Limits for VH-SS channel: The expected and observed limits are plotted as a function of the Higgs mass. The limits are normalised to the expected Higgs cross-section.

TABLE VI: Expected and observed 95% CL upper limits for the VH-SJ channel using 4fb^{-1} of $p\bar{p}$ data collected by the CDF detector. The limits are normalised to the expected Higgs cross-section.

Higgs mass (GeV/c^2)	-2σ	-1σ	Median (σ/σ_H)	$+1\sigma$	$+2\sigma$	Observed
100	30.8	41.9	58.3	81.2	110.9	39.2
110	24.5	32.3	45.2	63.9	89.3	39.2
120	28.9	37.4	51.2	71.3	98.5	43.7
130	52.1	69.3	96.1	134.5	185.8	66.4
140	98.0	130.6	181.5	254.6	352.8	120.2
150	333.9	452.9	634.9	892.0	1234.3	564.4

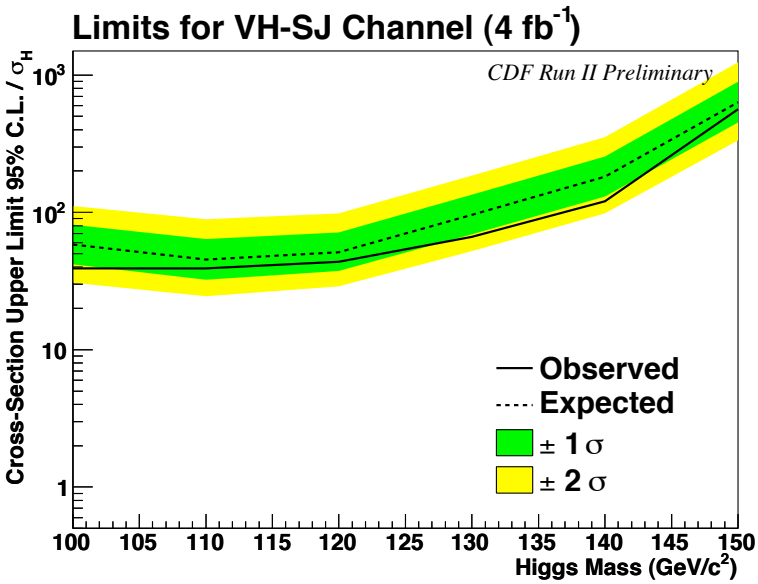


FIG. 15: Limits for VH-SJ channel: The expected and observed limits are plotted as a function of the Higgs mass. The limits are normalised to the expected Higgs cross-section.

TABLE VII: Expected and observed 95% CL upper limits for the VBF-SS channel using 4fb^{-1} of $p\bar{p}$ data collected by the CDF detector. The limits are normalised to the expected Higgs cross-section.

Higgs mass (GeV/c^2)	-2σ	-1σ	Median (σ/σ_H)	$+1\sigma$	$+2\sigma$	Observed
100	22.8	29.9	41.6	59.1	82.6	47.1
110	27.2	36.6	51.8	73.7	103.2	54.4
120	25.8	35.1	49.3	69.2	92.3	47.6
130	41.4	57.5	80.9	112.8	154.5	71.6
140	56.1	75.3	105.9	149.5	208.3	105.1
150	123.1	166.1	233.5	329.6	459.0	224.9

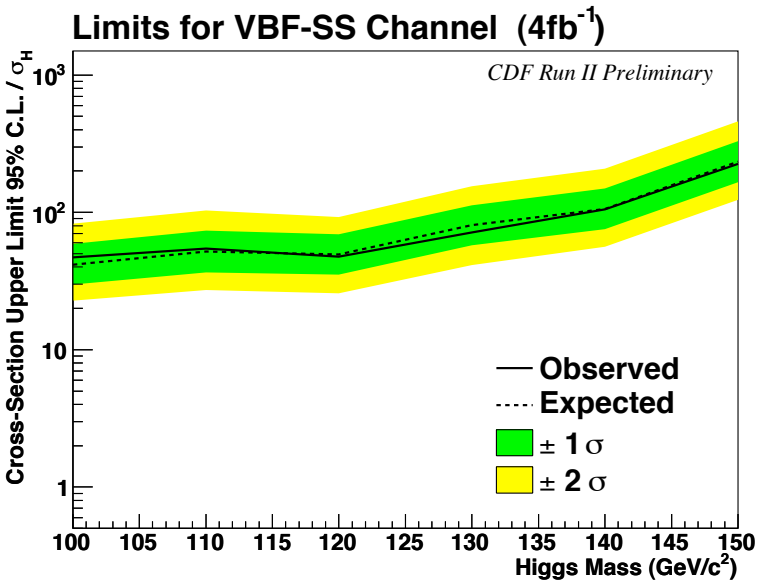


FIG. 16: Limits for VBF-SS channel: The expected and observed limits are plotted as a function of the Higgs mass. The limits are normalised to the expected Higgs cross-section.

TABLE VIII: Expected and observed 95% CL upper limits for the VBF-SJ channel using 4fb^{-1} of $p\bar{p}$ data collected by the CDF detector. The limits are normalised to the expected Higgs cross-section.

Higgs mass (GeV/c^2)	-2σ	-1σ	Median (σ/σ_H)	$+1\sigma$	$+2\sigma$	Observed
100	52.1	71.1	100.3	142.0	197.3	75.8
110	67.9	91.5	129.4	183.9	253.1	92.8
120	71.2	95.2	133.0	186.9	255.7	94.0
130	100.6	138.4	196.7	279.2	389.3	133.1
140	123.9	173.9	244.4	338.6	460.1	212.5
150	279.5	385.7	544.5	761.7	1046.6	436.8

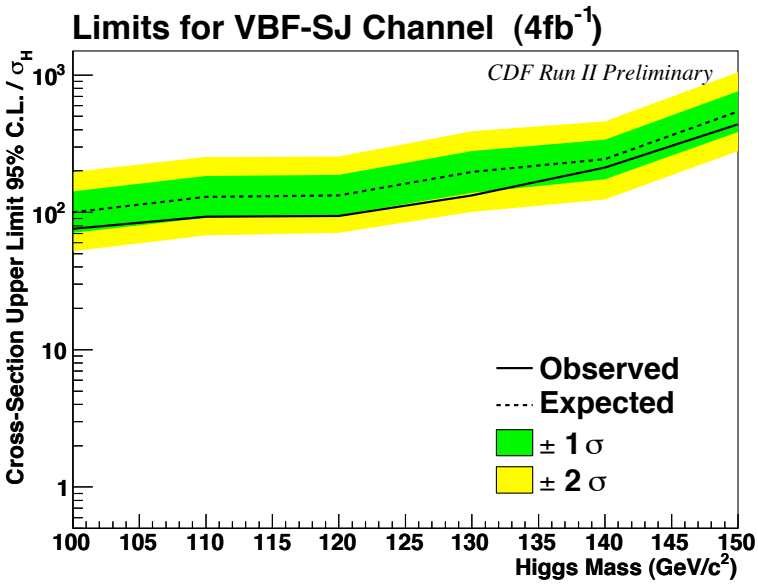


FIG. 17: Limits for VBF-SJ channel: The expected and observed limits are plotted as a function of the Higgs mass. The limits are normalised to the expected Higgs cross-section.

VII. CONCLUSIONS

We have presented the search for the Standard Model Higgs boson in the all hadronic mode using 4 fb^{-1} of data collected by the CDF detector. 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.

VIII. APPENDIX

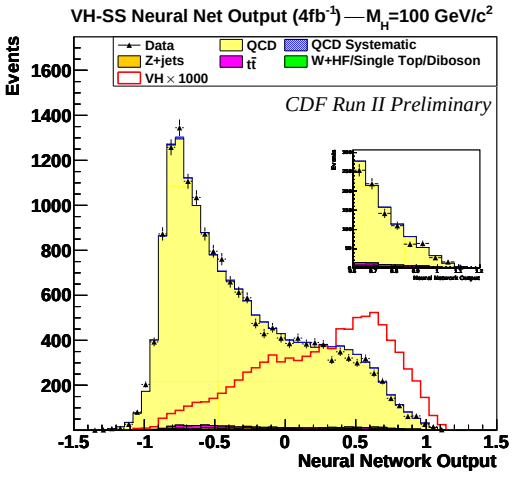
A. Z+jets Generator Level Filter

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.

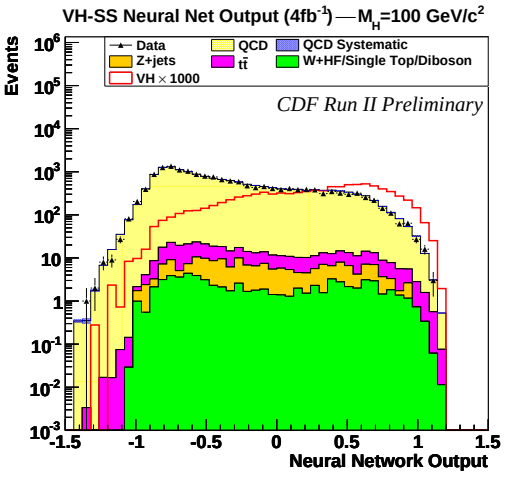
- At generator level, select events with ≥ 1 b or c parton.
- ≥ 3 jets with $E_T > 5\,\text{GeV}$. The jets were defined by cone sizes of 0.4, 0.7 and 1.0.
- The Sum- E_T for the 0.4, 0.7 and 1.0 jets are calculated. Accept the event if any of these sums $\geq 60\,\text{GeV}$

B. Unblinded Signal Region for Other Higgs Masses

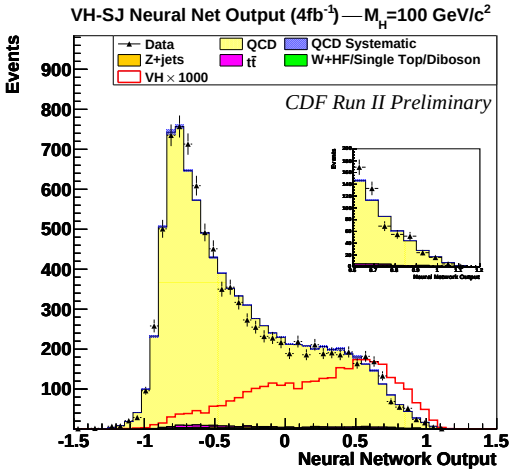
Figs. 18-27 show the unblinded data for the Higgs masses 100, 110, 130, 140 and 150 GeV for the VH and VBF channels respectively.



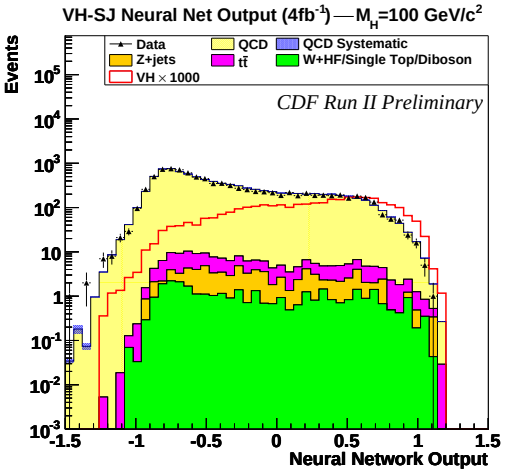
(a)VH100 SS Linear



(b)VH100 SS Log

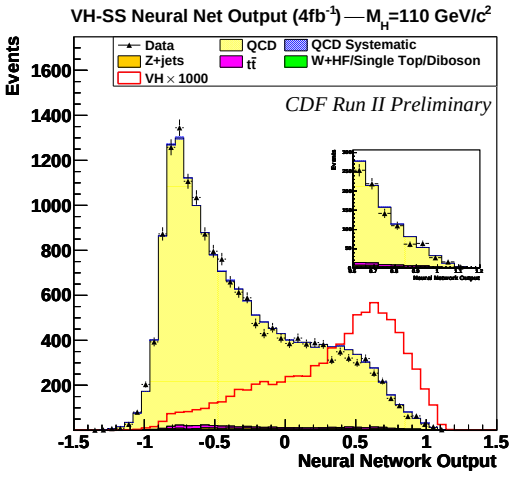


(c)VH100 SS Linear

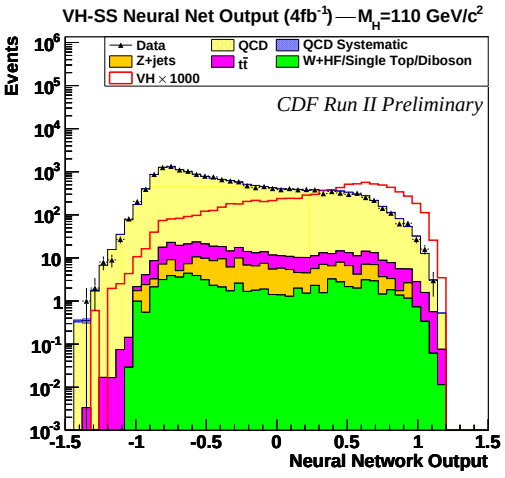


(d)VH100 SS Log

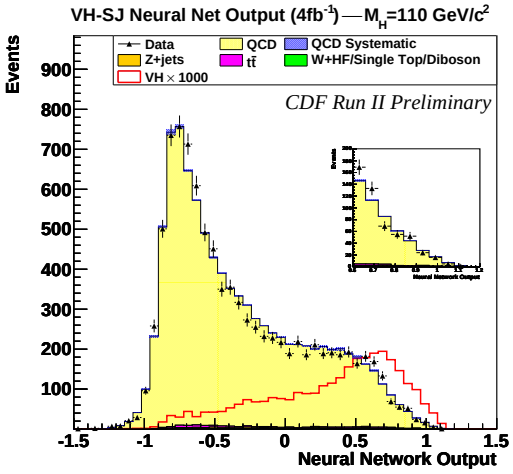
FIG. 18: The Neural Net distributions for the VH channel for Higgs masses of 100 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel.



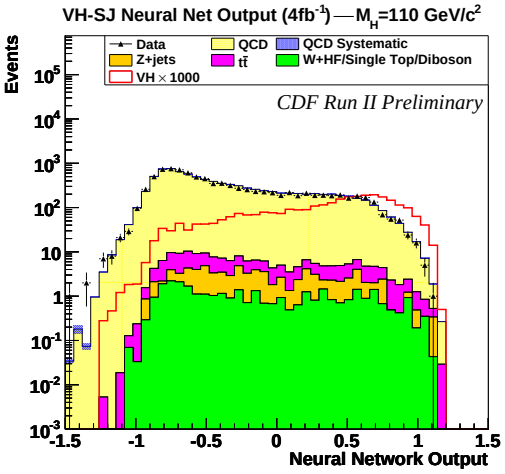
(a)VH110 SS Linear



(b)VH110 SS Log

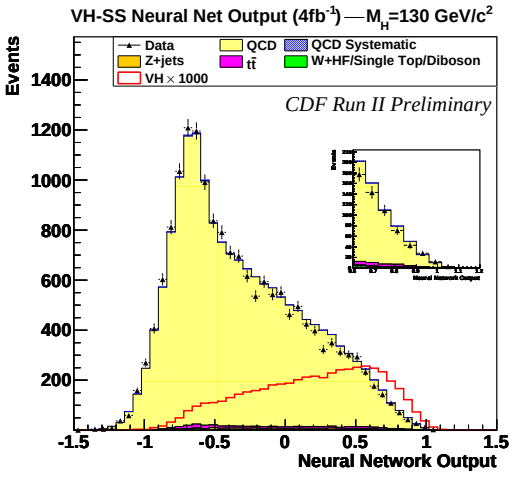


(c)VH110 SS Linear

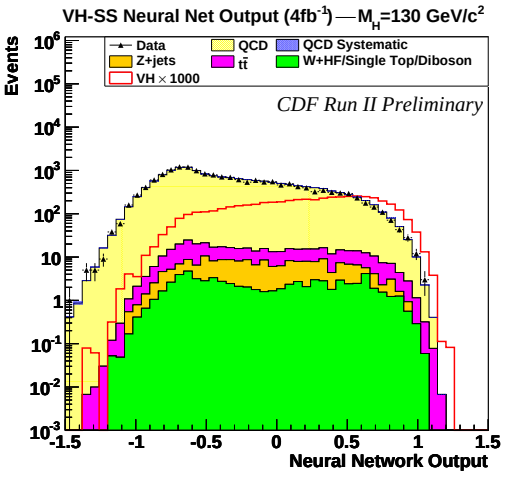


(d)VH110 SS Log

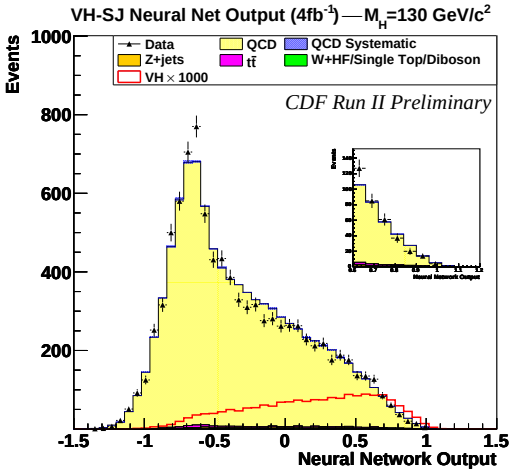
FIG. 19: The Neural Net distributions for the VH channel for Higgs masses of 110 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel.



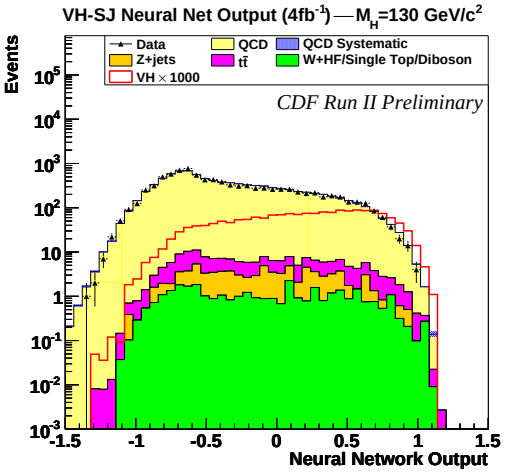
(a)VH130 SS Linear



(b)VH130 SS Log

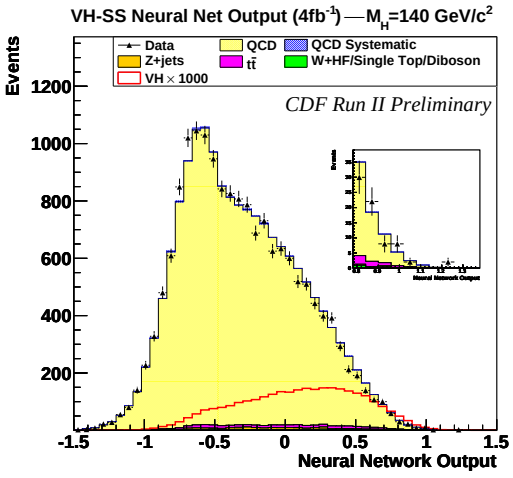


(c)VH130 SS Linear

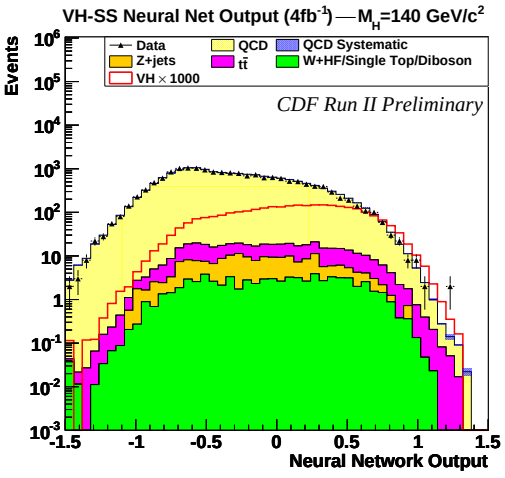


(d)VH130 SS Log

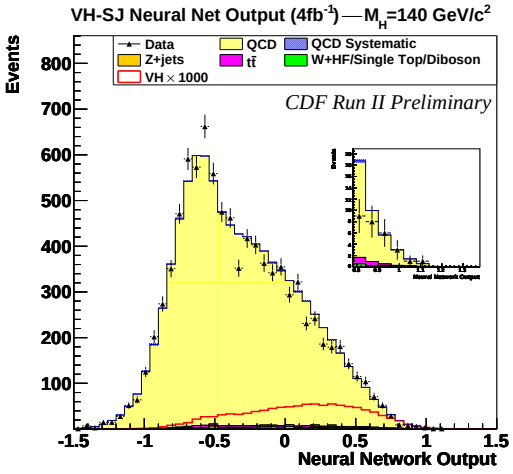
FIG. 20: The Neural Net distributions for the VH channel for Higgs masses of 130 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel.



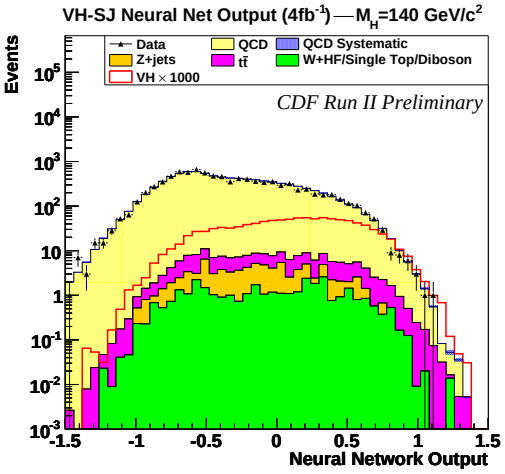
(a)VH140 SS Linear



(b)VH140 SS Log

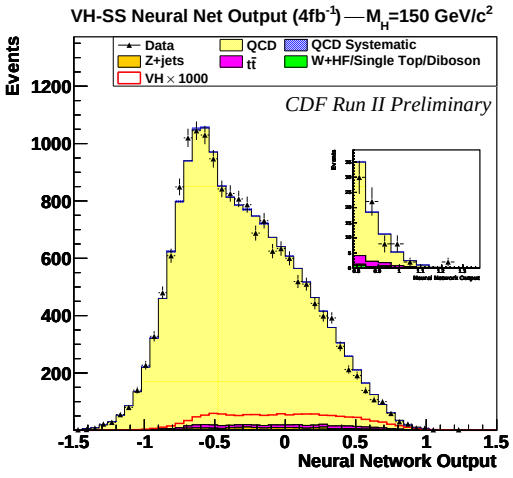


(c)VH140 SS Linear

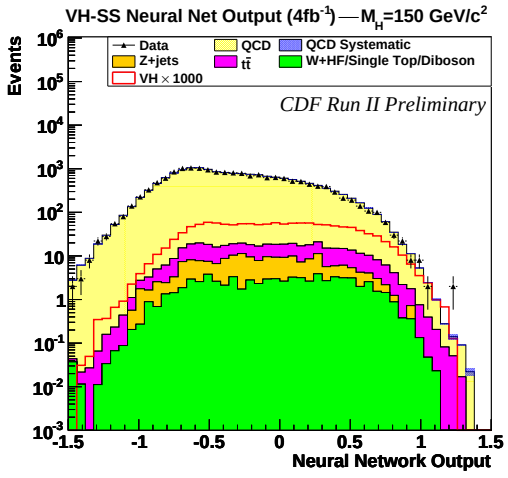


(d)VH140 SS Log

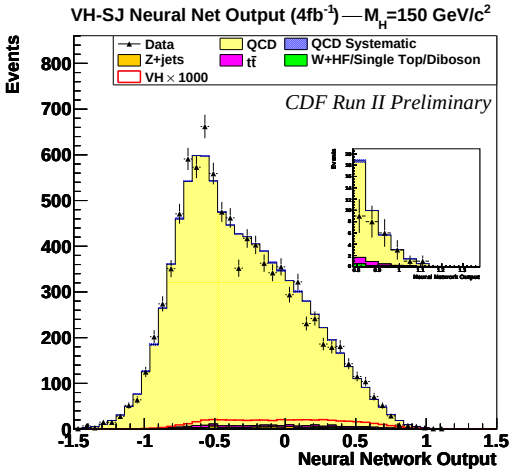
FIG. 21: The Neural Net distributions for the VH channel for Higgs masses of 140 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel.



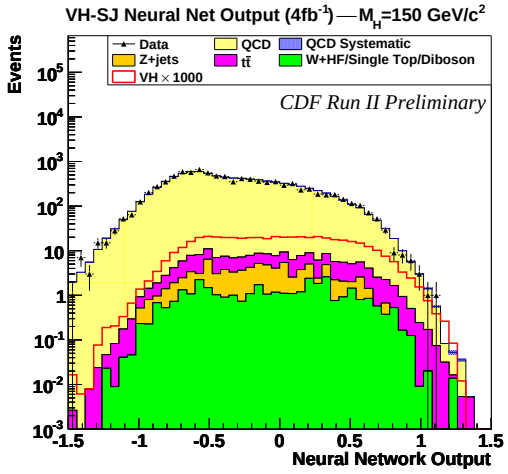
(a)VH150 SS Linear



(b)VH150 SS Log

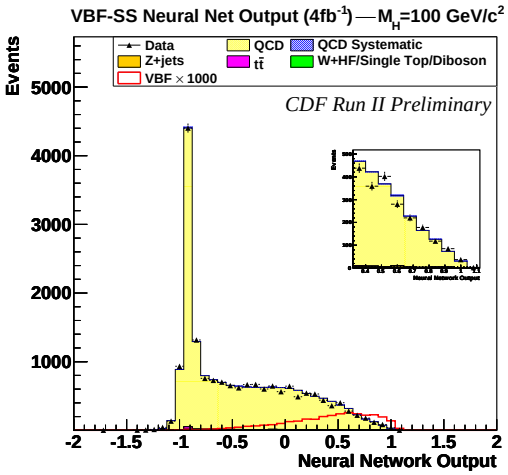


(c)VH150 SS Linear

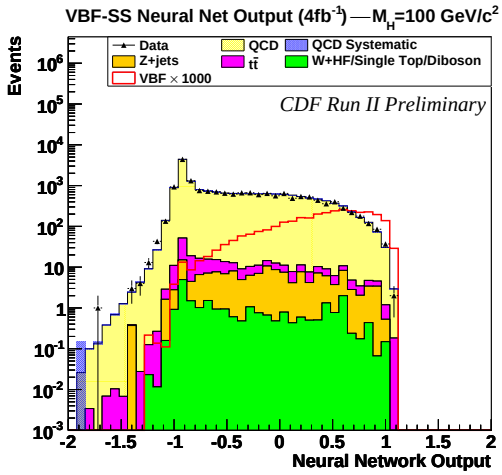


(d)VH150 SS Log

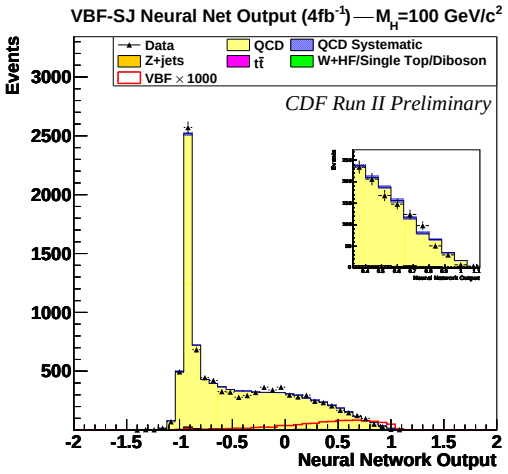
FIG. 22: The Neural Net distributions for the VH channel for Higgs masses of 150 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel.



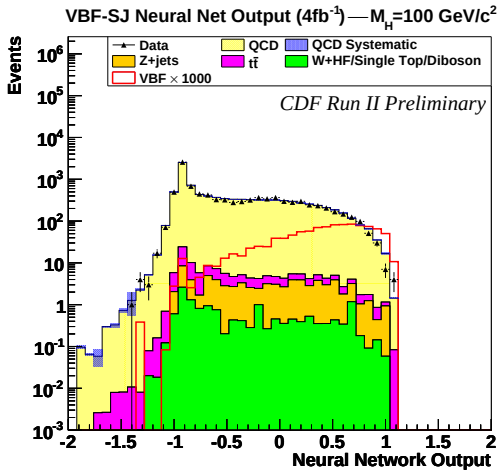
(a)VBF100 SS Linear



(b)VBF100 SS Log

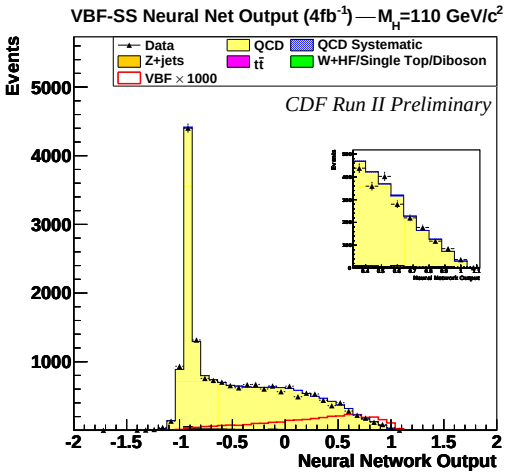


(c)VBF100 SS Linear

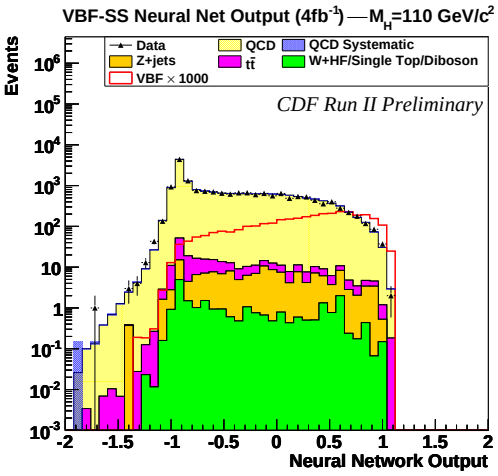


(d)VBF100 SS Log

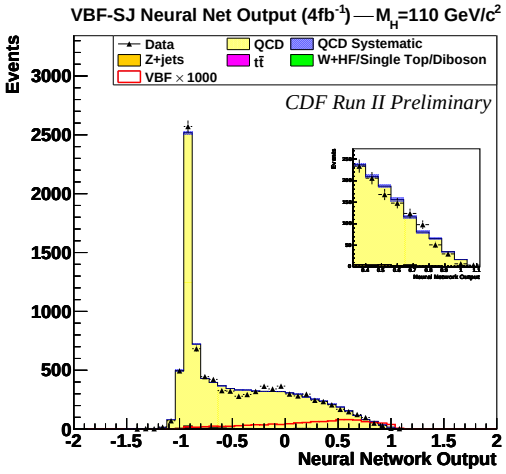
FIG. 23: The Neural Net distributions for the VBF channel for Higgs masses of 100 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel.



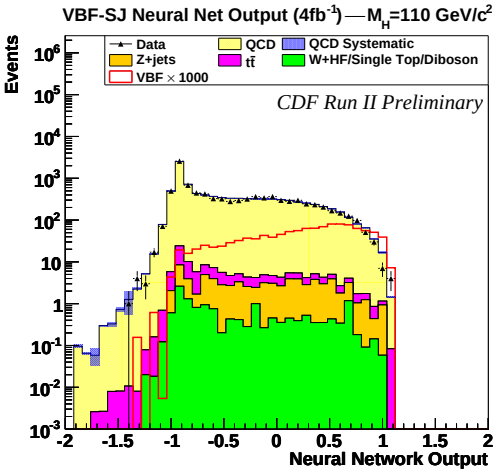
(a)VBF110 SS Linear



(b)VBF110 SS Log

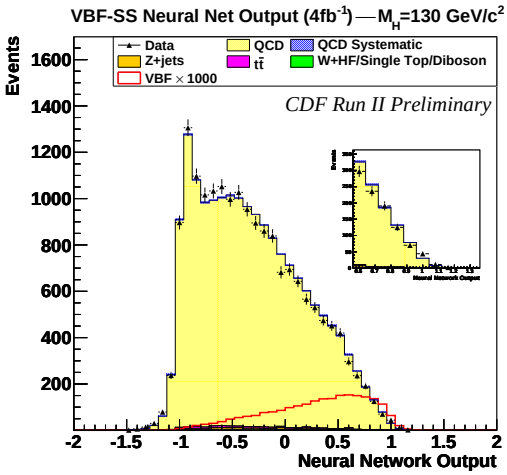


(c)VBF110 SS Linear

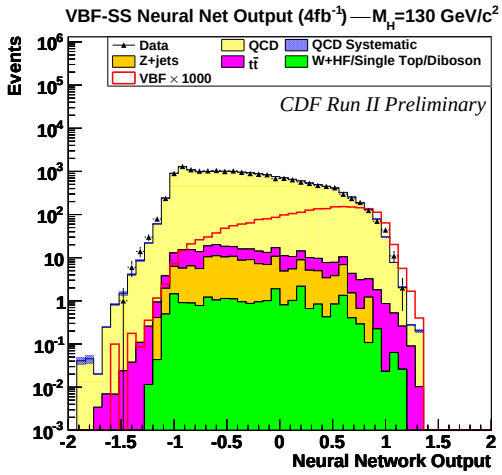


(d)VBF110 SS Log

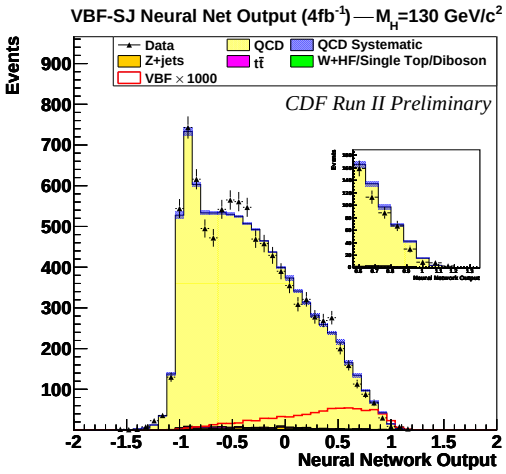
FIG. 24: The Neural Net distributions for the VBF channel for Higgs masses of 110 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel.



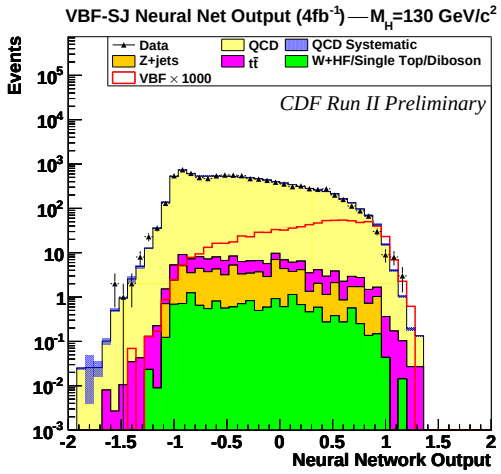
(a)VBF130 SS Linear



(b)VBF130 SS Log

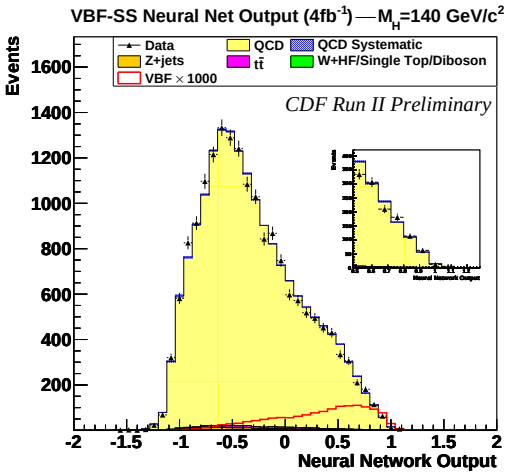


(c)VBF130 SS Linear

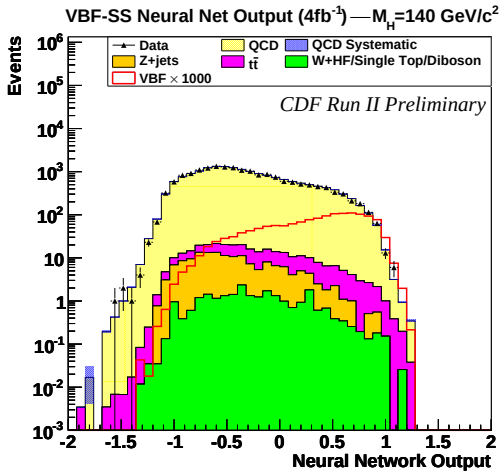


(d)VBF130 SS Log

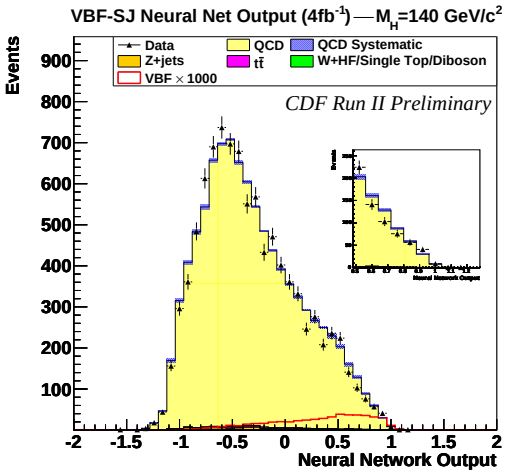
FIG. 25: The Neural Net distributions for the VBF channel for Higgs masses of 130 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel.



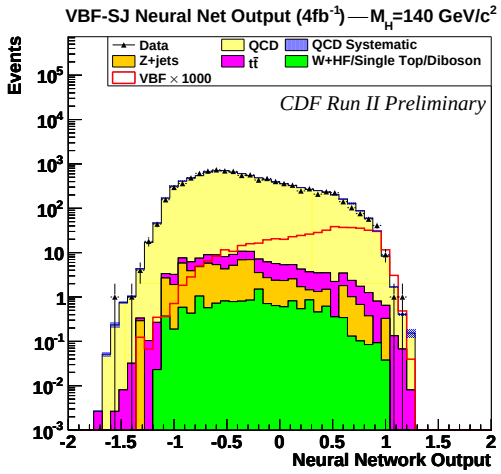
(a)VBF140 SS Linear



(b)VBF140 SS Log

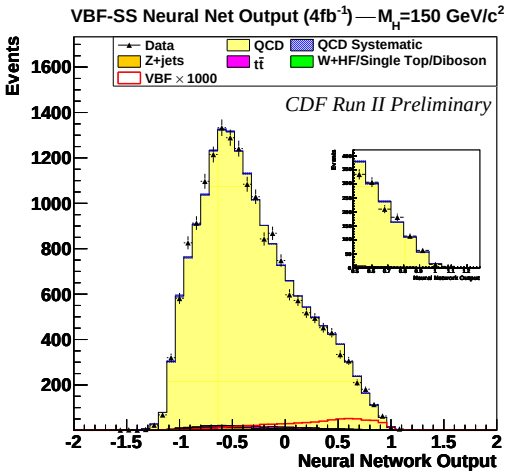


(c)VBF140 SS Linear

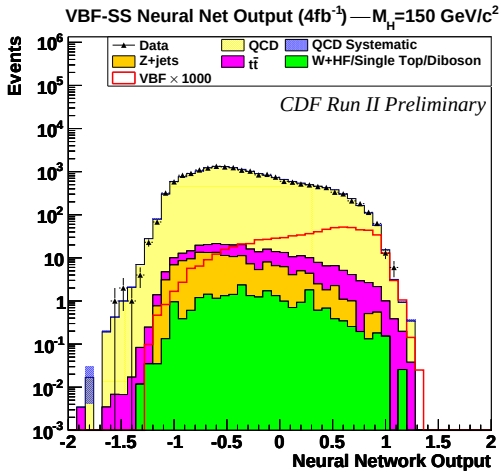


(d)VBF140 SS Log

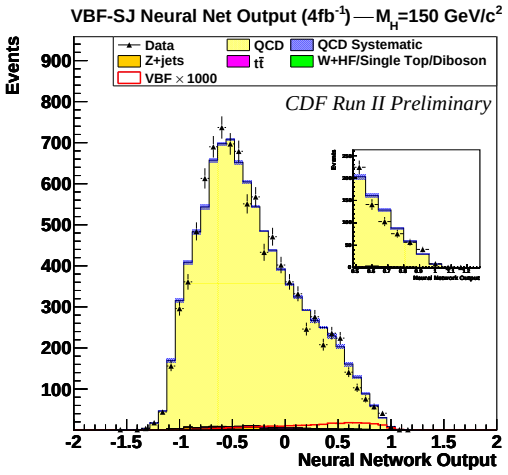
FIG. 26: The Neural Net distributions for the VBF channel for Higgs masses of 140 GeV are shown on a linear scale (left) and log scale (right). The upper plots are for the SecVtx-SecVtx channel and the bottom plots are for the SecVtx-JetProb channel.



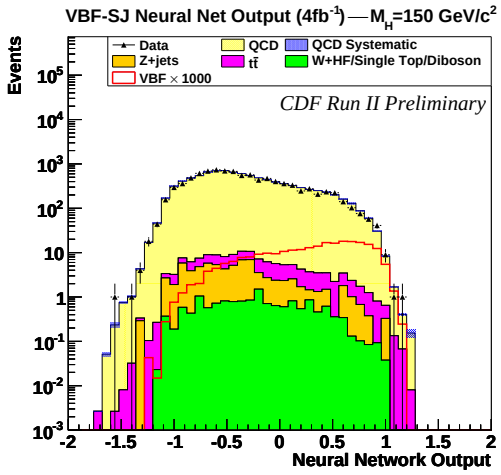
(a)VBF150 SS Linear



(b)VBF150 SS Log



(c)VBF150 SS Linear



(d)VBF150 SS Log

FIG. 27: 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.

[1] CDF Collaboration. A search for the associated production of the standard-model higgs boson in the all-hadronic channel. *Fermilab-Pub-09-341-E. arXiv: 0907.0810*, 2009.

[2] CDF Collaboration. *FERMILAB-PUB-96/390-E.*

[3] A.Sill et al. *Nucl. Instrum. and Methods A*, 447:1, 2000.

[4] CDF uses a cylindrical coordinate system in which θ is the polar angle to the proton beam, ϕ is the azimuthal angle about the beam axis, and pseudorapidity is defined as $\eta = -\ln \tan(\theta/2)$. The transverse energy and transverse momentum are defined as $E_T = E\sin\theta$ and $p_T = p\sin\theta$, where E is energy measured in the calorimeter and p is momentum measured by the tracking system.

[5] F. Abe et al. Topology of three-jet events in $p\bar{p}$ collisions at $s = 1.8$ tev. *Phys. Rev. D*, 45(5):1448–1458, Mar 1992.

[6] The missing transverse energy significance is defined as $\not\!\!E_T / \sum E_T$, where $\not\!\!E_T$ is the missing transverse energy and $\sum E_T$ the total transverse energy of the event which passes the event selection.

[7] D. Acosta et al. Measurement of the $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV using lepton + jets events with secondary vertex b -tagging. *Physical Review D (Particles and Fields)*, 71(5):052003, 2005.

[8] A. Abulencia et al. Measurement of the $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV using lepton + jets events with jet probability b-tagging. *Physical Review D (Particles and Fields)*, 74(7):072006, 2006.

[9] 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.

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

[11] H. L. Lai et al. Global QCD analysis of parton structure of the nucleon: CTEQ5 parton distributions. *Eur. Phys. J.*, C12:375–392, 2000.

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

[13] The helicity angle θ^*_{q1} of the leading non b-jet q_1 is defined to be the angle between the momentum of q_1 in the $q_1 - q_2$ rest frame and the total momentum of $q_1 - q_2$ in the lab frame.

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