Towards the Search for the Standard Model Higgs boson produced in Vector Boson Fusion and decaying into a $\tau$ pair in CMS with 1 fb$^{-1}$: $\tau$ identifications studies.

CMS collaboration

Abstract

The Standard Model Higgs boson produced in vector boson fusion and decaying to a pair of $\tau$ leptons is an important channel in the search of the Higgs boson in the mass range between 115 and 145 GeV. A selection strategy for the $\tau\tau \rightarrow \ell\nu+\tau\text{-jet }\nu$ final state is presented. Some of the analysis methods that will be used to control the systematic uncertainties in the early searches at low luminosity of the order of 1 fb$^{-1}$ are also described.
1 Introduction

The Standard Model Higgs boson decay to a pair of $\tau$ leptons is an important channel in the search of the Higgs boson in the mass range between 115 and 145 GeV. The vector boson fusion (VBF) production process provides characteristic signatures of two forward-backward quark jets, which can be used to distinguish the Higgs boson signal from the background processes. The production cross section of the VBF process is the second highest at the LHC and the di-$\tau$ decay mode has a moderate branching ratio, which enables the observation of the Higgs boson with an accumulated luminosity of $\sim 30 \text{ fb}^{-1}$.

A particular signature of VBF events is the rapidity gap between the two outgoing quark jets due to the absence of the color exchange between the incoming quarks. This leaves an open space in the center of the detector enabling the observation of the isolated Higgs boson decay products. The forward jet tagging and the selection of the central rapidity gap are the key elements of the VBF Higgs boson analysis.

In the early studies at parton level [1, 2] and the studies with fast simulation of ATLAS and CMS [3, 4] and with full simulation of CMS [5], it was shown that the Higgs boson production in the VBF process decaying into a $\tau$ lepton pair could be a discovery channel with a luminosity of the order of $30 \text{ fb}^{-1}$.

The cross section measurement of $qqH, H \rightarrow \tau\tau, WW, \gamma\gamma$ channels will significantly extend the possibility of the Higgs boson coupling measurements [6, 7] and provide the possibility of an indirect measurement of the light Higgs boson width [6]. In the MSSM the $qqH(h), H(h) \rightarrow \tau\tau$ channels could be discovered in the largest region of the $M_A$-$\tan\beta$ parameter plane [1, 8, 9] in comparison with other modes, like $h \rightarrow \gamma\gamma, WW, ZZ$ or $\phi \rightarrow \tau\tau (\mu\mu) (\phi=h, H, A)$ in $pp \rightarrow b\bar{b}\phi$ production.

2 Selection strategy and background consideration

The goal of this analysis is to select events with a central lepton and a $\tau$-jet ($\tau_{\text{had}}$) plus two forward tagging jets.

The single isolated lepton triggers ($e, \mu$) and the combined lepton plus $\tau_{\text{had}}$ cross triggers ($e+\tau, \mu+\tau$) will be used to select $qq \rightarrow qqH, H \rightarrow \tau\tau$ events with the $\tau\tau \rightarrow \ell\nu\nu + \tau_{\text{had}}\nu$ final state. At low luminosities during the LHC startup the efficiency gain of the cross triggers will be small as the thresholds of the single $e$ and $\mu$ triggers will be low, so the cross triggers will be run only for commissioning purposes. At higher luminosities of the order of $2 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$ the combined $e(\mu)+\tau$ triggers will play an important role since they will allow to keep the lepton thresholds considerably lower than the thresholds of the single lepton triggers. It is also foreseen to have Level 1 topological triggers with an isolated lepton and two jets with a large rapidity separation which will allow to keep the lepton thresholds considerably lower than the single lepton thresholds, which will be very important in the high luminosity regime.

The off-line event analysis starts with the individual object reconstruction and identification. The reconstructed electrons and muons in each event that have passed the identification and isolation cuts are considered, and those events which have more than one lepton are rejected in order to remove background events with multiple leptons. Electrons with reconstructed $p_T > 10 \text{ GeV}/c$ and muons with $p_T > 5 \text{ GeV}/c$ are considered. The events with one remaining muon or electron of $p_T > 15 \text{ GeV}/c$ are selected.

The highest $E_T \tau_{\text{had}}$ candidate with $E_T > 30 \text{ GeV}$, which does not coincide with the identified
lepton within $\Delta R < 0.3$ is selected. The charge of the selected $\tau_{\text{had}}$ candidate is required to be opposite to that of the lepton. The $\tau_{\text{had}}$ identification is based on the matching of the calorimeter jet with the leading $p_T$ track, followed by the tracker isolation of the signal cone built around the leading track. The track-jet matching cone, the signal and isolation cones are calculated in $\eta - \phi$ space and have fixed values. The method to reject electrons faking $\tau$'s is described in details in Section 3.3. In addition, the one prong $\tau$'s with the impact point of the track on the calorimeter surface close to the $\eta$ gaps are rejected. The $\tau$ identification efficiency is $\simeq 36\%$ and the jet misidentification probability is $\simeq 2\%$ for a $\tau_{\text{had}}$ with $E_T > 40$ GeV. The contamination of muons faking a $\tau_{\text{had}}$ is negligible.

Among the reconstructed calorimeter jets, a pair of highest $E_T$ jets in opposite rapidity regions $(\eta_{j1} \times \eta_{j1} < 0)$ are selected as the two forward-backward jets that make the VBF signature in the signal event. Jets which coincide with the selected lepton or $\tau_{\text{had}}$ within $\Delta R < 0.5$ are excluded from this search. The jets are required to be within the detector acceptance of $|\eta| < 4.5$, and the minimum $E_T$ requirement was 30 GeV. A separation of the two jets in rapidity of $\Delta \eta_{j1,j2} > 2.5$ and a high di-jet effective mass, $M(j1,j2) > 400$ GeV are further required. The events with additional hadronic activity in the central region are vetoed to reject QCD background processes (see detailed description of the methods used in Section 3.1.1).

The invariant mass of the reconstructed di-$\tau$ system is calculated using combinations of the selected lepton, the $\tau_{\text{had}}$ and the missing $E_T$ ($E_T^{\text{miss}}$). Two different methods are tried as explained in detail in Section 3.2.2. Finally, the transverse invariant mass of the lepton-$E_T^{\text{miss}}$ system, $M_T(l,E_T^{\text{miss}})$, is considered for rejecting background processes involving $W$ boson. A cut at $M_T(l,E_T^{\text{miss}}) < 40$ GeV was used.

The background processes considered for the qqH, $H \rightarrow \tau\tau$ channel with a lepton and a $\tau_{\text{had}}$ in the final state are: $Z+2/3$jet ($Z \rightarrow \ell\ell$), $W+3/4$jet ($W \rightarrow \ell\nu$), $t\bar{t} \rightarrow WbWb$ and QCD di-jet events, with a real or a fake lepton. The dominant background is $Z+2/3$jet, $Z \rightarrow \tau\tau$ and to control the systematic uncertainties it is important to have a correct prediction of the di-$\tau$ mass shape of this background. The method described in Section 3.2.3 uses a combination of the real data and Monte-Carlo to model the di-$\tau$ mass shape. The contributions from the background processes will be calculated using the number of events obtained from the inclusive $W/Z+$jets and $t\bar{t}$ measurements and the event selection efficiencies in the signal region will be obtained from the Monte Carlo predictions. This will allow to take into account the jet energy scale and the $E_T^{\text{miss}}$ uncertainties as the cuts on $M(j1,j2)$ and $M_T(l,E_T^{\text{miss}})$ are applied in the signal selection. The QCD di-jet background will be determined from the data using opposite sign (OS) and same sign (SS) $\ell$-$\tau_{\text{had}}$ ratios measured with non-isolated leptons (the method is similar to one used by the D0 Collaboration [10]).

A profile likelihood method will be used to evaluate the upper limit on the number of the signal events. In the fit of the mass distribution the background variations will be restricted to the the prediction uncertainties and the uncertainty of the background mass shapes will be taken into account.

3 New Analysis Methods

The CMS Physics TDR study of the qqH, $H \rightarrow \tau\tau$ channel with a lepton and $\tau_{\text{had}}$ in the final state [5] has been extended in the prospect of the preparation of the analysis with the first 1 fb$^{-1}$ of LHC data. In comparison with the previous analysis a few new items of the analysis were developed and improved:
3.1 Selection of the central rapidity gap

3.1.1 Central rapidity gap veto methods

In the VBF Higgs boson searches at the LHC, a selection of events with a central rapidity gap between the two tagging jets allows to reduce the QCD $Z$+jets and other backgrounds, like $W$+jets and $t\bar{t}$, while keeping a high efficiency for the Higgs boson signal from the VBF production. The central jet veto was proposed and used in the first VBF Higgs boson studies [1, 2] and exploited in the recent ATLAS and CMS analyses [4, 5]. The central calorimeter jet veto technique is suffering from pile-up and electronic noise in the calorimeters which could create fake jets. The method to reduce the fake calorimeter jets using the information from the event vertex and tracks was proposed in [11] and successfully used in CMS analyses [5, 12].

We consider two methods to perform the hadron activity veto in the central rapidity region: the (traditional) central calorimeter jet veto (CJV) and the track counting veto (TCV). The idea of the track counting veto is inspired by the paper [13], where a method was proposed to distinguish between the gluon and vector boson fusion processes for the Higgs boson production. The performance of both methods is compared in terms of the signal efficiency and the QCD $Z$+jets background rejection.

3.1.2 Studies with pile-up

The fully simulated datasets from the VBF Higgs boson analysis ($H \rightarrow \tau\tau \rightarrow \ell + \text{jet}$) [5] for an instantaneous luminosity of $2 \times 10^{33}$ cm$^{-2}$s$^{-1}$ are used for this study, where pile-up events (4.3 events per crossing) are included in the simulation. At reconstruction level, the VBF selections used on the tagging jets are:

- $E_T^j > 40$ GeV, $|\eta^j| < 4.5$, $M_{j1j2} > 1000$ GeV, $|\Delta\eta^{j1j2}| > 4.2$, $\eta^{j1} \times \eta^{j2} < 0$, $\eta^{j1} \times \eta^{j2} < 0$, where $j1$ and $j2$ are two leading $E_T^j$ jets ordered in $E_T$.

The performance of the two methods, CJV and TCV was compared. The CJV rejects events with a third jet that satisfies

- $E_T^j > 10$ GeV, where $E_T$ is the raw, non-calibrated, energy of the jet
- fake jet rejection parameter $\alpha^j = \sum p_T^{\text{track}} / E_T^j > 0.1$ (see [5] for details).

The TCV rejects events which have a certain number of tracks within the tracker acceptance region, $|\eta| < 2.4$, that satisfy

- $p_T^{\text{track}} > p_T^{\text{cut}}$ GeV/c
- $\eta^{\text{track}} < \eta^{\text{track}} < \eta^{\text{track}} < \eta^{\text{track}} < \eta^{\text{track}}$.

The tracks used for the TCV are required to have $\geq 8$ hits and $\Delta Z(\text{track}, \text{vertex}) < 2$ mm. The lepton and $\tau$-had tracks are not included in the track count.

Fig. 1 shows the performance of the TCV algorithm, where the signal selection efficiency versus the background rejection efficiency is shown. For different cuts of the $p_T^{\text{track}}$ (1, 2 and 3 GeV/c), the efficiency is shown for different values of the tracking multiplicity. In each case, the smallest efficiency value correspond to a zero track multiplicity and the track multiplicity is increased in
steps of one. For comparison, the performance of the CJV algorithm based on calorimeter jets is also shown, which achieves a good performance: 80.0% efficiency for the signal and 39.7% efficiency for the background. The TCV algorithm can reach comparable discrimination power.

![Figure 1: The track counting veto (TCV) performances for different $p_T^{track}$ and track multiplicity thresholds compared with the performance of the CJV.](image)

Studies without pile-up and more relaxed VBF selections to be used in the early data taking, also show that the CJV and TCV algorithms can achieve similar performances (see Appendix).

With the full detector simulation it has been seen that both the central jet veto and the track counting algorithms can achieve similar performances. The robustness and the stability of the methods under a variation of the run and detector conditions will be tested with real data using $Z$+jets, $Z \rightarrow \mu\mu$ events. It is believed that the track counting algorithms which rely on a single sub-detector, the tracking system, could perform with higher reliability.

### 3.2 $Z(H) \rightarrow \tau\tau$ mass reconstruction and modeling of the di-$\tau$ mass shape from real $Z \rightarrow \mu\mu$ data

#### 3.2.1 $Z(H) \rightarrow \tau\tau$ mass reconstruction

The full mass of the $Z(H)$ boson can be reconstructed as the invariant mass of its decay products, two $\tau$ leptons in the $Z(H) \rightarrow \tau\tau$ process. For high $p_T$ $\tau$-leptons, the direction of the neutrinos from the $\tau$ decay can be approximated to be collinear with the visible decay products and their energies can be retrieved from the total $E_{T}^{miss}$ of the system. While this collinear approximation works in many cases and provides a satisfactory mass resolution, a fraction of the events are lost in the case where the approximation results in negative neutrino energy. This may arise from a particular event topology where the outgoing products are nearly back-to-back or simply due to a poor reconstruction of the $E_{T}^{miss}$.

Another method for reconstructing the di-tau invariant mass, particularly suited for early data-taking, is the reconstruction of the $\tau\tau$ mass using the visible decay products only, which is completely independent of the $E_{T}^{miss}$ reconstruction.
3.2 Z(H) → ττ mass reconstruction and modeling of the di-τ mass shape from real Z → μμ data

3.2.2 Higgs boson mass reconstruction performance

The distribution of the Higgs boson mass reconstructed using the two methods described in the previous section are shown in the Appendix. Each mass distribution was fitted with a gaussian function to estimate the mean and the width of the distribution. An asymmetric range was used for the full mass in order to fit the peak of the distribution. The fit parameters are listed in Table 1. The width of the distribution increases as the mass increases, while the mass resolution, \( \sigma/M \), remains similar. The mean value of the visible mass is \( \sim 0.55 \times M_H \) throughout the mass range considered. On the other hand, the fully reconstructed mass is greater than the actual value by 6-7\%, which is mainly due to the jet calibration factors used for the \( E_T^{miss} \) corrections in this study which over-estimates the \( p_T \) of the quark jets in \( qqH, H \rightarrow ττ \) events.

Table 1: Fit parameters for the reconstructed mass distribution for \( qq \rightarrow qqH, H \rightarrow ττ \) events for different Higgs boson mass scenarios.

<table>
<thead>
<tr>
<th>mass [GeV/c^2]</th>
<th>full mass, M(ττ)</th>
<th>visible mass, M(l+τhad)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean ( σ ) ( σ/M )</td>
<td>mean ( σ ) ( σ/M )</td>
</tr>
<tr>
<td>115</td>
<td>123.0 17.0 0.138</td>
<td>63.1 16.1 0.255</td>
</tr>
<tr>
<td>125</td>
<td>133.3 18.6 0.140</td>
<td>68.4 17.1 0.250</td>
</tr>
<tr>
<td>135</td>
<td>144.1 20.6 0.143</td>
<td>73.5 18.9 0.257</td>
</tr>
<tr>
<td>145</td>
<td>152.9 21.6 0.141</td>
<td>78.0 21.0 0.269</td>
</tr>
</tbody>
</table>

3.2.3 Measurement of the \( Z \rightarrow ττ \) mass shape from real \( Z \rightarrow μμ \) data

The di-τ invariant mass (\( M_{ττ} \)) distribution will be analyzed to search for the presence of the Higgs boson signal in the region above the di-τ mass peak corresponding to the Z boson background from the \( Z \rightarrow ττ \) decays. It is important to know the shape of the di-τ mass distribution from \( Z/γ* \rightarrow ττ \) decays in the region of the \( Z \) peak and above. The dominant uncertainties of the mass shape are expected to come from the modeling of the \( E_T^{miss} \) in the part related to pile-up, underlying event, calorimeter noise and the calorimeter response to multi-jets accompanying the \( Z \) boson production. To reduce uncertainties it will be crucial to obtain this part of the \( E_T^{miss} \) spectrum from real data. The uncertainties in the \( E_T^{miss} \) due to the lepton and \( τ_{had} \) momentum measurements are expected to be negligible.

A method to model the di-τ mass shape from real \( Z/γ* \rightarrow μμ \) events has been developed. As a first step, the muons are removed from the real event, thus the muon \( p_T \) and the energy deposition in the calorimeter are not counted in the \( E_T^{miss} \) calculation. As a second step, the di-τ events with the kinematics corresponding to the real muon kinematics are generated and their response in the detector is fully simulated. Finally, the real \( Z \rightarrow μμ \) event with the muons removed and the simulated di-τ event are superimposed to form one event, \( Z \rightarrow τμτμ \), and the di-τ mass is calculated in this event.

3.2.4 Test of the method for the \( ττ \rightarrow μνν + μνν \) and \( ττ \rightarrow ℓνν + τ_{had}ν \) final states

Figure 2 (left) shows the reconstructed di-τ mass distribution for the \( ττ \rightarrow μνν + μνν \) final state. Also shown is the mass distribution obtained with fake \( Z \rightarrow ττ \) events obtained from Drell-Yan \( Z \rightarrow μμ \) events, which pass the selection cuts \( p_T^{μ} >10 \text{ GeV}, \vert \eta^{μ} \vert <2.4 \) and 70 GeV \( < M_{μμ} < 110 \text{ GeV} \). A good agreement between the di-τ mass shapes is obtained. The distributions are normalized to the expected number of \( Z \rightarrow μμ \) events with a luminosity of 130 pb\(^{-1}\). Fig. 2
(right) shows the reconstructed di-$\tau$ mass distribution for the $\tau\tau \rightarrow \ell\nu\nu + \tau_\text{had}\nu$ final state for a sample of Z+jets events, with two jets that pass the VBF selections described in Section 2. The distributions are normalized to the expected number of $Z \rightarrow \mu\mu + \text{jets}$ events with a luminosity of 1 fb$^{-1}$.

$$\int L \, dt = 130 \, \text{pb}^{-1}$$

Figure 2: The reconstructed di-$\tau$ mass distribution for real and fake $Z \rightarrow \tau\tau$ events for (left) $\tau\tau \rightarrow \mu\nu\nu + \mu\nu\nu$ final state from inclusive Drell-Yan events and (right) $\tau\tau \rightarrow \ell\nu\nu + \tau_\text{had}\nu$ final state from Z+jets events.

### 3.3 Rejection of electrons faking a $\tau_\text{had}$

Electrons which are characterized by a calorimeter object with a single matching track are likely to pass the track-based $\tau$ identification criteria. In the CMS Physics TDR Vol.2 two variables were proposed in the CMS Physics TDR in order to reduce the contamination of electrons faking a $\tau_\text{had}$: $E_{\text{T}}^{\text{HT}}$ and $E_{\text{T}}^{\text{HT}}/p_{\text{T}}^{\text{tr}}$, where $E_{\text{T}}^{\text{HT}}$ is the transverse energy of the hottest HCAL tower inside the $\tau_\text{had}$. New variables based on sum of the transverse energy of several HCAL towers, $E_{\text{T}}^{\text{HT}}/n_{\times n}$ ($n = 2,3$), around the leading track impact point at the calorimeter surface are now also considered. Fig. 3 shows the selection efficiency for electrons from Z+jets, $Z \rightarrow \mu\mu$ events and for $\tau_\text{had}$ from Z+jets, $Z \rightarrow \tau\tau$ events for different cuts on the electron rejection variables.

For a fixed $\tau_\text{had}$ efficiency, the variable $E_{\text{T}}^{\text{HT}}/p_{\text{T}}^{\text{tr}}$ suppresses the electron contribution by half compared to the $E_{\text{T}}^{\text{HT}}/p_{\text{T}}^{\text{tr}}$ variable. A further factor two reduction is achieved with the sum of HCAL towers, $E_{\text{T}}^{\text{HT}}/p_{\text{T}}^{\text{tr}}$. With real Z+jets data it will be possible to cross check the electron rejection power of this variable. The threshold value was set to 0.1, which gives a reasonable efficiency for $\tau$ jets above 85%. The efficiency for the real $\tau$ jets to pass the electron rejection criteria can be measured from real data using a sample of single isolated tracks.
3.4 Estimation of the Z+jets background contribution from double parton scattering

An additional Z+jets background can be originated from double parton interactions (DPI) in a proton-proton collision when the Z boson is produced in one parton-parton interaction and QCD di-jets are produced in the second parton-parton interaction. In that case two choices of the tagging jets are possible: (a) one tagging jet is from QCD di-jet production and the second one is from the Drell-Yan production and (b) two tagging jets are both selected from the QCD di-jet production.

The contribution from the double-parton interaction was estimated at particle level following the procedure provided by T. Sjostrand [14], which is described in the Appendix. Events are selected which have at least two leading jets reconstructed with a cone algorithm (cone size 0.5) that satisfy the following requirements: \( E_T^j > 40 \text{ GeV}, \eta_j < 5.0, M_{j1j2} > 1000 \text{ GeV}, |\Delta \eta^{j1j2}| > 4.2, \eta^{j1} \times \eta^{j2} < 0 \), where \( j1 \) and \( j2 \) are the two leading \( E_T^j \) jets ordered in \( E_T \). After applying these cuts the cross section of Z+jets from DPI is \( \simeq 15 \% \) of the Z+jets background from a single parton interaction. The possibility of using Z+2jets, \( Z \rightarrow \mu\mu \) events with VBF jet selections and studying the unbalance in \( \vec{p}_T^\tau \) between the Z boson and the di-jets to control and measure the Z+jets background from the double parton is under investigation.

4 Appendix

4.1 Central rapidity gap veto: studies with no pileup

The efficiency of the central rapidity gap selection was evaluated for two sets of the VBF requirements:

- soft VBF selections: \( M_{j1j2} > 400 \text{ GeV}/c^2, |\Delta \eta^{j1j2}| > 2.5 \)
- hard VBF selections: \( M_{j1j2} > 800 \text{ GeV}/c^2, |\Delta \eta^{j1j2}| > 3.5 \).
Figure 4 shows the efficiency of selecting qqH, H → ττ signal events versus the efficiency for the Z+jets, Z → ℓℓ background events using the two central rapidity gap selection methods, CJV (left) and TCV (right). The jet $E_T$ thresholds were varied for the CJV using the calibrated and the non-calibrated values, both providing very similar performances. For the TCV, the veto requirement on the minimum number of tracks were varied for different track $p_T$ thresholds. The highest $p_T$ cut of 3 GeV/c provided the best performance in keeping the background efficiency low. Requiring no more than three tracks with $p_T > 3$ GeV/c gives rise to a similar signal and background efficiency to the CJV with the default threshold of $E_{raw}>10$ GeV. The two sets of VBF selection requirements on the leading jets were tried for both the CJV and the TCV. It can be clearly seen that tightening the VBF requirements significantly reduces the background efficiency, providing a better central rapidity gap selection performance.

![Figure 4](image.png)

Figure 4: The signal efficiency (qqH, H → ττ, $M_H = 135$ GeV/$c^2$) vs the background efficiency (Z+jets, Z → ℓℓ) for CJV (left) and TCV (right) with VBF requirements.

4.2 The H → ττ mass distributions

The distribution of the Higgs boson mass reconstructed using the two of the methods described in the Section 3.2.1 are shown in Fig. 5 for the qqH, H → ττ events with four different masses.

4.3 Estimation of the Z+jets background from double parton scattering

The contribution from double-parton interactions was estimated at particle level with PYTHIA 6.4. Drell-Yan and QCD di-jet events were generated separately. The Drell-Yan production was generated with full underlying event (UE) using Tune DWT [15], while for the QCD di-jet production the UE was switched off. Drell-Yan events with di-lepton mass $m_{ℓℓ} > 70$ GeV/$c^2$ and QCD di-jet events with $p_T > 20$ GeV/c were generated. The NLO cross section $2 \times 10^6$ fb for the Drell-Yan production and the PYTHIA cross section $8.2 \times 10^{11}$ fb for the QCD di-jet production were used in the estimates presented in this section. At the second step two events (Drell-Yan and QCD di-jets) were mixed together and analyzed as one event. Jets were found at particle level by a simple cone algorithm, with cone size 0.5, implemented in the PYTHIA PYCELL routine.
4.3 Estimation of the Z+jets background from double parton scattering

Figure 5: The reconstructed mass distribution using (left) visible decay products and (right) and the collinear approximation of the $E_T^{miss}$, for $q\bar{q}H, H \rightarrow \tau\tau$ events. All selection cuts are applied, and the distributions are normalized to the expected number of events with a luminosity of $1 \text{ fb}^{-1}$.

The cross section for double parton interactions was evaluated with the factorization formula

$$\sigma_{A,B} = \frac{m}{2} \frac{\sigma_A \times \sigma_B}{\sigma_{\text{eff}}},$$

where $m=1$, for indistinguishable parton processes and $m=2$ for distinguishable parton processes (in our case we use $m=2$). In the experimental study of double parton collisions CDF quotes $\sigma_{\text{eff}}=14.5 \text{ mb}$ [16]. For LHC energies the value of $\sigma_{\text{eff}}=20 \text{ mb}$ according to [14] was used. It gives the $\sigma_{A,B}=8.2 \times 10^{4} \text{ fb}$ ($A=$Drell-Yan, $B=$QCD di-jets). More pessimistic value of $12 \text{ mb}$ quoted in [17] will double our estimates of the Z+jets background from the double parton interactions. The longitudinal correlations in the double-parton structure functions neglected in the above formula can have a sizable effect at LHC according to references [18, 19].

The Z+jets background from double parton collisions is compared to the QCD Z+jets background from one parton-parton collision. Events were selected with at least two leading $E_T$ jets that satisfy the following requirements: $E_T^j > 20 \text{ GeV}, \eta^j < 5.0, M_{j1j2} > 1000 \text{ GeV}, |\Delta\eta^{j1j2}| > 4.2, \eta^{j1} \times \eta^{j2} < 0$, where $j1$ and $j2$ are two leading $E_T$ jets ordered in $E_T$. The double parton scattering events where the two leading jets were both originated from the Drell-Yan production (the fraction of such events is $\simeq 20\%$) were excluded from the consideration to avoid double counting.

Table 2 shows the initial cross sections for the Z+jets background from one and two parton-parton interactions and the cross sections after VBF cuts. After selections the contribution from double parton interactions is $\simeq 40\%$ (320 fb) of the Z+jets background from one parton-parton interactions (770 fb). The fraction of the DPI events where one tagging jet is selected from the Drell-Yan process and the other from the QCD di-jet process is $\simeq 70\%$ while in 30% of the DPI events both tagging jets are selected from the QCD di-jet process.
Table 2: The initial cross sections for Z+jets background from one and two parton-parton interactions and the cross sections after VBF cuts in fb.

<table>
<thead>
<tr>
<th>interaction process</th>
<th>one parton-parton exclusive $\ell\ell+2j$</th>
<th>one parton-parton inclusive $\ell\ell+3j$</th>
<th>two parton-parton Drell-Yan</th>
<th>two parton-parton QCD di-jets</th>
</tr>
</thead>
<tbody>
<tr>
<td>no cuts</td>
<td>$1.0 \times 10^3$</td>
<td>$2.0 \times 10^3$</td>
<td>$8.2 \times 10^4$</td>
<td></td>
</tr>
<tr>
<td>$\geq 2$ jets, $E_T^j &gt; 20$ GeV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta\eta_{j1,j2} &gt; 4.2$, $\eta_{j1} \times \eta_{j2} &lt; 0$</td>
<td>$2.4 \times 10^2$</td>
<td>$5.3 \times 10^2$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M_{j1,j2} &gt; 1000$ GeV/$c^2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 6 (left) shows the angle in the transverse plane between the two tagging jets ($\Delta\phi_{j1,j2}$) for the Z+jets background from single and double parton-parton interactions and the Higgs signal. Figure 6 (right) shows the $\Delta\phi_{j1,j2}$ distribution for the case where one of the tagging jets comes from the Drell-Yan process and the other from the QCD di-jet event and the case where both tagging jets come from the QCD dijet-event. In the first case there is no correlation between two tagging jets, as expected, while in the second case the tagging jets are mainly in a back-to-back topology. All distributions are normalized to unity.

Fig. 7 shows the transverse energy of the tagging jets from the Z+jets background and the Higgs signal. The Z+jets background from DPI can be largely suppressed with cut on the tagging jet energy of $E_T^j > 40$ GeV, which was used in the full simulation analysis [5]. After this selection, the cross section of Z+jets from DPI is $\approx 100$ fb and the cross section of the Z+jets background from a single parton interaction is $\approx 700$ fb. The relative contribution from DPI is reduced to $\approx 15\%$ and the signal event rate is reduced by 20% if the tagging jet $E_T^j$ selection is increased from 20 to 40 GeV.
4.3 Estimation of the Z+jets background from double parton scattering

Figure 7: The transverse energy of the tagging jets from Z+jets backgrounds and the signal VV → H; distributions are normalized to unit

References


[14] p. c. T. Sjostrand. The possibility to superimpose two hard interactions in one proton-proton collision is realized in PYTHIA8 version.


