# Measurement of the Z boson plus two b-jets cross section in CMS with 100 pb$^{-1}$

**Aruna Kumar NAYAK**[*][†]

*Tata Institute of Fundamental Research, Mumbai*
*E-mail:* `aruna.nayak@cern.ch`

The cross section for production of Z bosons with two associated b-jets, and Z decaying to leptons, can be measured at the LHC with about 100 pb$^{-1}$ of data. We use simulated data to study possible strategies for an early measurement of this process with the CMS detector. The rate and kinematic properties of this final state needs to be well understood because it constitutes a large fraction of the total backgrounds to several of the Higgs discovery channels at the LHC.

---

[*]Speaker.
[†]On behalf of CMS Collaboration

## 1. Introduction

At the LHC, $Z/\gamma^*$ can be produced in association with two b-quarks, $gg/q\bar{q} \rightarrow b\bar{b}Z$, with a significant cross section [1], [2], [3]. The measurement of this process is an important test of QCD calculations and also can help to reduce the uncertainty of the Super Symmetry $b\bar{b}H$ production cross section calculation [4], since the two processes are produced via similar partonic processes at the initial state. This process is also a background to Higgs boson discovery channels at the LHC, like Standard Model $H \rightarrow ZZ \rightarrow 4\ell$ and associated production of SUSY Higgs boson $bb\varphi$, $\varphi \rightarrow \tau\tau$ ($\mu\mu$) ($\varphi = h$, $H$, $A$). In the following, the possibility of observing and measuring the cross section of the $b\bar{b}Z$ production at the LHC, using 100 pb$^{-1}$ of early data at CMS experiment, has been described which was studied in [5]. A related, but different cross section for $p\bar{p} \rightarrow Z + b$-jet has previously been measured at the Tevatron, both by D0 and CDF [6], [7].

## 2. Monte Carlo Signal and Background samples

The signal events $\ell\ell b\bar{b}$ (*Zbb* ) were generated at parton level using the leading order (LO) CompHEP generator [8] and hadronized with PYTHIA [9] with the following generator-level cuts : $p_t^b >$10 GeV/$c$, $|\eta^b| <$10, $m_{\ell\ell} > 40$ GeV/$c^2$ and $|\eta^\ell| <$2.5. The next-to-leading order cross-section $\sigma(\ell\ell b\bar{b}) = 45.9$ pb ($\ell$=e, $\mu$, $\tau$) was calculated with the MCFM program [10] applying the same generator-level cuts. The CTEQ6M parton density functions (pdf) and scale $\mu_R = \mu_F = M_Z$ were used. The backgrounds considered for this process were Drell-Yan $Z/\gamma^* \rightarrow \ell\ell$ ($\ell = e, \mu, \tau$) production in association with two or more light-quark and gluon jets (Z+jets), $\ell\ell c\bar{c}$+jets (Zcc+jets) and $t\bar{t}$+jets. Background samples were generated at the leading order (LO) with the ALPGEN generator [11]. The $t\bar{t}$+jets ALPGEN events were normalized to the NLO inclusive $t\bar{t}$ cross section 840 pb [12]. The Zcc+jets events were normalized to the NLO MCFM cross section of $\ell\ell c\bar{c}$, 13.29 pb, applying the ALPGEN production cuts: $p_T^c >$20 GeV/$c$, $|\eta^c| <$5, $m_{\ell\ell} > 40$ GeV/$c^2$. The same pdf and scale settings as for $\ell\ell b\bar{b}$ process were used. The Z+jets events were normalized to NLO MCFM cross section of $\ell\ell + 2$ jets, 714 pb, applying the ALPGEN production cuts: $p_T^j >$20 GeV/$c$, $|\eta^j| <$5, $m_{\ell\ell} > 40$ GeV/$c^2$. Signal and background samples were passed through the full simulation and reconstruction chain of CMS, under an imperfect calibration and alignment configuration assumed to be typical of the first 100 pb$^{-1}$ of data at CMS.
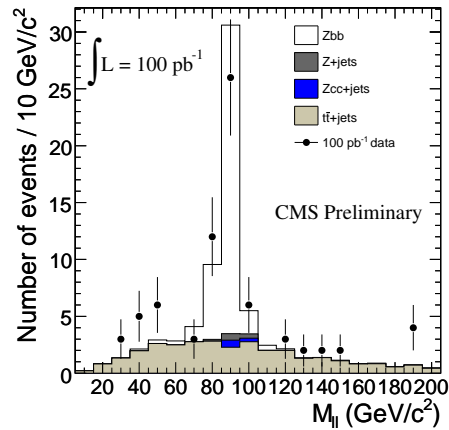
## 3. Event Selection

The events were selected by the CMS Level-1 (L1) and High-Level (HLT) single isolated electron and muon triggers designed for the low luminosity period (L=$10^{32}$cm$^{-2}$s$^{-1}$). In offline, Events with at least two isolated and oppositely charged electrons or muons, with $p_T^{e,\mu}$ greater than 20 GeV, $|\eta^e| < 2.5$ and $|\eta^\mu| < 2.0$, were selected to form $Z$ boson candidate. Events are selected with two or more jets, reconstructed using iterative cone algorithm of cone size 0.5, with corrected jet $E_T$ greater than 30 GeV and jet $|\eta| < 2.4$. The $\eta$ cut ensured good quality b-tagging. The Figure 1 (Left) shows the distribution of di-lepton invariant mass for the electron final state events with at least two jets passing lepton and jet selections. The b-tagging is an important and effective tool to ensure the purity of *Zbb* events and reduce the $Z + jets$ and Zcc+jets backgrounds since
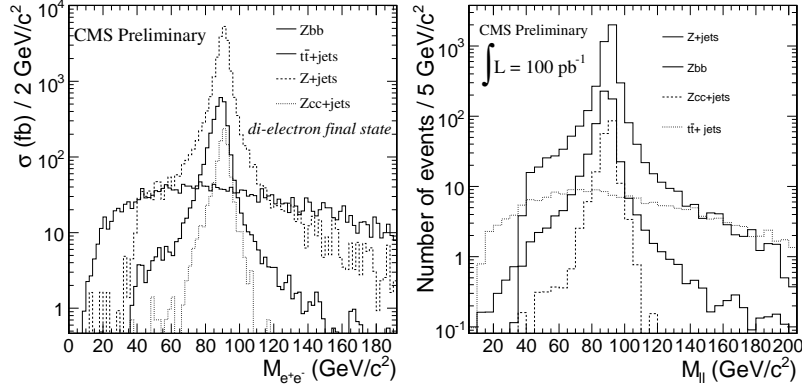
**Figure 1:** Left : Di-lepton invariant mass distribution for events passing lepton and jet selections as described in text. Right : Di-lepton invariant mass for events after all selections except b-tagging.

the b-tag discriminator values for light quarks, gluon and c-quark jets tends to be lower compared to that of the b-quark jets. The events were double b-tagged using 'track counting' b-tagging, which uses the 3D track impact parameter significance of the 3rd highest significance track in the jet as the b-tagging discriminator. A b-tag discriminator value of 2.5 has been used in the analysis which ensured high purity of b-tagged jets. The possibility of applying selections on the amount of missing transverse energy in the event, $E_T^{miss}$, is exploited to suppress $t\bar{t}$+jets background events since $t\bar{t}$+jets events contain neutrinos from $W \rightarrow \ell\nu$ decays and has higher $E_T^{miss}$ compared to *Zbb* events. A selection cut on $E_T^{miss}$ less than 50 GeV is applied on $E_T^{miss}$, where $E_T^{miss}$ is reconstructed using calorimeter tower information and calibrated using jet energy corrections and muon corrections. The dilepton invariant mass distributions of the signal and various background events passing all selection criteria are shown in Figure 2. The events are scaled to 100 pb$^{-1}$ of data.



**Figure 2:** Di-lepton invariant mass for events after all selections. The events are normalized to 100 pb$^{-1}$ of data. The dots are result of one random MC experiment with 100 pb$^{-1}$ of data. Nearly 46 events are expected in the mass range 75-105 GeV with S/B $\sim 3.6$

## 4. Background Estimation and Systematic Uncertainties

The contribution of the $t\bar{t}$+jets background events in the signal region of the di-lepton invariant mass distribution can be estimated from the data, by extrapolating sidebands in the di-lepton invariant mass spectrum. The following method is used to estimate the $t\bar{t}$+jets background using the relation $N_Z(t\bar{t}) = (\varepsilon_Z(t\bar{t})/\varepsilon_{noZ}(t\bar{t})) \times N_{noZ}(t\bar{t})$, where $N_Z(t\bar{t})$=8 is the expected number of $t\bar{t}$+jets events in the signal region within the $m_{\ell\ell}$ mass window of 75-105 GeV/$c^2$, $N_{noZ}(t\bar{t})$=27 is the measured number of $t\bar{t}$ events outside the signal region, $\varepsilon_Z(t\bar{t})$ is the selection efficiency for $t\bar{t}$+jets events in signal region, $\varepsilon_{noZ}(t\bar{t})$) is the selection efficiency for $t\bar{t}$+jets events outside the signal region. $\Delta N_Z(t\bar{t})$ is the uncertainty of the expected number of $t\bar{t}$+jets events in the signal region given by $\Delta N_Z(t\bar{t})/N_Z(t\bar{t}) = 1/\sqrt{N_{noZ}(t\bar{t})}$. The uncertainty on the ratio $\varepsilon_Z(t\bar{t})/\varepsilon_{noZ}(t\bar{t})$ is negligible in comparison to the statistical uncertainty of the number of events outside the signal region.

The following equations are used to determine the number of $Zbb$ events ($N_{Zbb}$) and it's uncertainty $\Delta N_{Zbb}$ for 100 pb$^{-1}$, after all selections except double b-tagging.

$$N_Z^{before\ b-tag} = N_{Zjj} + N_{Zcc} + N_{Zbb} \tag{4.1}$$

$$N_Z^{after\ b-tag} = \varepsilon_\ell \times N_{Zjj} + \varepsilon_c \times N_{Zcc} + \varepsilon_b \times N_{Zbb} \tag{4.2}$$

where, $N_Z^{before\ b-tag}$ = 4644 is the measured number of $Z/\gamma^* \to \ell\ell$ events in the di-lepton mass window between 75 and 105 GeV/$c^2$ after all selections have been applied except double b-tagging. It receives negligible contribution (less than 1%) from the $t\bar{t}$ background, as one can see from Figure 1 (Right). $N_Z^{after\ b-tag}$ = 38 is the measured number of $Z/\gamma^* \to \ell\ell$ events after all selections including double b-tagging. It is defined after subtraction of the $t\bar{t}$ background described above. $N_{Zjj}$ is the unknown number of $\ell\ell$+jets (jet=u,d,s,g) events before double b-tagging. $N_{Zcc}$ is the unknown number of $\ell\ell c\bar{c}$ events before double b-tagging. $N_{Zbb}$ is the unknown number of $\ell\ell b\bar{b}$ events before double b-tagging. $\varepsilon_b$ is the efficiency of double b-tagging for $Zbb$ signal events which is determined from the Monte-Carlo simulation tuned to reproduce the b-tagging efficiency measured from the data, $\varepsilon_c$ is the efficiency of double b-tagging for $Zcc$+jets background events which is determined from the Monte Carlo, $\varepsilon_\ell$ is the efficiency of double b-tagging for $Z$+jets (jet=u,d,s,g) background events which is determined from the Monte-Carlo simulation tuned to reproduce the mistagging efficiency measured from the data.

The number of unknown parameters are reduced to two, following the D0 analysis approach [7], by using the theoretical ratio of cross sections and ratio of selection efficiencies

$$R = \frac{\sigma(Zcc)}{\sigma(Zjj)} \times \frac{\varepsilon_{Zcc}^{sel}}{\varepsilon_{Zjj}^{sel}} = 0.046 \pm 0.002 \tag{4.3}$$

and by replacing $N_{Zcc}$ by $R \times N_{Zjj}$. The uncertainty of the $\frac{\sigma(Zcc)}{\sigma(Zjj)}$ ratio due to $\mu_R$, $\mu_F$ scale variation and JES and MET scale are considered. The equations are solved to evaluate the value of $N_{Zbb}$. The uncertainty on $N_{Zbb}$, $\Delta N_{Zbb}$, is than calculated due to the uncertainties of $N_Z^{after\ b-tag}$ from $t\bar{t}$ background subtraction ($\delta N_Z^{after\ b-tag}$=4.0%), the uncertainty of $R$ and the uncertainties of $\varepsilon_b$, $\varepsilon_c$ and $\varepsilon_\ell$. From the estimated number of $N_{Zbb}$ events after all selections except b-tagging, the number of $Zbb$ events before the lepton and jet selections and cut on $E_T^{miss}$ can be evaluated. The

systematic uncertainties due to Jet Energy Scale, $E_T^{miss}$ scale, lepton selections, b-tagging, Monte Carlo dependence on $p_T$ and $\eta$ of jets and luminosity measurement have been taken into account according to the values foreseen to be achieved in CMS with data corresponding to 100 pb⁻¹ of integrated luminosity.

The statistical uncertainty is defined as $\Delta N_{sel}/N_{sel}=1/\sqrt{N_{sel}}$, where $N_{sel}$ is the measured number of events after all selections (46), thus it is $\delta N_{sel}$=14.7%.

Finally, the $\ell\ell b\bar{b}$ ($\ell$=e, $\mu$) cross-section is expected to be measured with an accuracy of

$$\delta\sigma \;=\; \pm\; {}_{25\%}^{21\%}\,(syst) \;\pm\; 15\%\,(stat.) \tag{4.4}$$

## 5. Conclusion

The process $pp \rightarrow b\bar{b}Z$ at LHC needs to be well understood because of its importance as a background to SM and SUSY Higgs boson discovery channels. The possibility of measuring the $b\bar{b}Z, Z \rightarrow \ell\ell$ process at CMS with the first 100 pb⁻¹ of data has been studied with a robust selection of leptons and b-jets. Possible methods of measuring backgrounds from data have been discussed. The statistical and systematic uncertainties have been discussed.

## References

[1] J.M.Campbell, R.K.Ellis, *Radiative corrections to Z b anti-b production, Phys.Rev.***D62**(2000) 114012, [hep-ph/0006304]

[2] J.M.Campbell, R.K.Ellis and D. Rainwater *Next-to-leading order QCD predictions for W + 2jet and Z +2jet production at the CERN LHC, Phys.Rev.***D68**(2003) 094021, [hep-ph/0308195]

[3] F.Maltoni, T.McElmurry and S. Willenbrock, *Inclusive production of a Higgs or Z boson in association with heavy quarks, Phys.Rev.***D72**(2005)074024,[hep-ph/0505014]

[4] J.Campbell et al., *Higgs Boson Production in Association with Bottom Quarks*, [hep-ph/0405302]

[5] CMS Collaboration, *Measurement of the b̄bZ, Z → ℓℓ cross section with 100 pb⁻¹ of early CMS data at the LHC*, **CMS PAS EWK-08-001**(2008)

[6] CDF Collaboration, A. Abulencia et al, *Measurement of the b jet cross-section in events with a Z boson in pp̄ collisions at $\sqrt{s}$ = 1.96-TeV, Phys. Rev.***D74**(2006), 032008, [hep-ex/0605099]

[7] D0 Collaboration, V.M. Abazov et al., *A measurement of the ratio of inclusive cross sections $\sigma(pp \rightarrow Z+b-jet)/\sigma(pp \rightarrow Z+jet)$ at $\sqrt{s}$ = 1.96 TeV, Phys.Rev.Lett.***94**(2005), 161801, [hep-ex/0410078]

[8] A.Pukhov et al. *CompHEP: A package for evaluation of Feynman diagrams and integration over multi-particle phase space.*, [hep-ph/9908288]

[9] T.Sjostrand et al. *PYTHIA 6.3: Physics and manual* [hep-ph/0308153]

[10] J.Campbell, and R.K.Ellis, *Next-to-leading order corrections to W + 2jet and Z + 2jet production at hadron colliders, Phys.Rev.***D65**(2002) 113007, [hep-ph/0202176]

[11] M.L.Mangano et al., *ALPGEN, a generator for hard multiparton processes in hadronic collisions, JHEP* **07**(2003) 001, [hep-ph/0206293]

[12] Beneke, M., et al. *Top quark physics*, [hep-ph/0003033]