PRAMANA
— journal of
physics

Vol. 79, No. 4
October 2012
pp. 911–915

Inclusive $b + Z$, $Z \to \mu^+\mu^-$, e^+e^- production at CMS

NATALIE HERACLEOUS, for the CMS Collaboration
I. Physikalisches Institut B, RWTH, Sommerfeldstr. 14, D-52074 Aachen, Germany
E-mail: heracleous@physik.rwth-aachen.de

Abstract. The production of Z bosons in association with at least one b-jet and Z decaying into muons or electrons is studied. This analysis is done by using proton–proton collisions delivered by the Large Hadron Collider (LHC) at a centre of mass energy of 7 TeV and with data recorded by the CMS detector in 2010, representing an integrated luminosity of 36 pb^{-1}. A final state of two well-identified muons or electrons and at least one b-tagged jet are required in order to perform this study. The event yields and shapes of kinematic variables are compared with Monte Carlo predictions from MADGRAPH. The ratio $\sigma(Z + b)/\sigma(Z + j)$ is found to be 0.054 ± 0.016 (0.046 ± 0.014) in the data for the Z $\to ee$ (Z $\to \mu\mu$) selection, compared to 0.043 ± 0.005 (0.047 ± 0.005) estimated from NLO theory predictions.

Keywords. b-jet; Z boson; Drell–Yan; ratio; CMS; Large Hadron Collider; heavy quarks.

PACS Nos 13.85.Qk; 14.65.Fy; 14.70.Hp

1. Introduction

The production mechanism of Z in association with b-quarks at the LHC is not yet well understood. The first test measures the corresponding cross-section in order to constrain the existing QCD predictions. In addition to this Standard-Model motivation, this process is a large background for many searches related to Higgs sectors and predictions beyond the Standard Model.

2. Data and Monte Carlo samples

The total integrated luminosity analysed is 35.9 ± 0.4 pb^{-1} [1]. Events are triggered using the loosest unprescaled single lepton trigger and the most recent alignment and energy calibration. The main backgrounds considered are the associated production of Z with charm jets ($Z + c$), Z with light jets ($Z + l$) where $l = u, d, s, g$ and the production of top pairs. The production of Z to ee and $\mu\mu$ in association with jets from any flavour is referred to as $Z + j$. All Monte Carlo samples used, have been generated with MADGRAPH [2] interfaced with PYTHIA 6.4 [3], reconstructed through the full simulation of the CMS

DOI: 10.1007/s12043-012-0402-6; ***e*Publication: 5 October 2012**

detector [4] based on GEANT4 [5], including tune $Z2$ and a pile-up scenario comparable to the data. All Monte Carlo samples are normalized to the total data-integrated luminosity using the NNLO cross-sections [6,7].

3. Event selection

In order for the two leptons to be selected as candidates, they should have opposite charges. Furthermore, both electrons are required to be reconstructed with $p_{\mathrm{T}} > 25\,\mathrm{GeV/c}$ and in the detector range of $\eta < 2.5$. For the muon pair these requirements are: $p_{\mathrm{T}} > 20\,\mathrm{GeV/c}$ and $\eta < 2.1$ [8]. Identification and isolation criteria for the leptons are also applied to ensure background rejection and good signal efficiency. Lepton isolation is defined using the sum of energy around the lepton in a cone size of $\Delta R < 0.3$ in the tracker, ECAL and HCAL detectors divided by the lepton p_{T}. For muons, this variable should be less than 0.15 while for the electrons cuts are applied on the individual isolation variables divided by the electron p_{T}. Electron identification and isolation criteria are chosen such that the rejection efficiency is 85%.

Jets are reconstructed using the particle flow (PF) algorithm [9]. Particles are clustered into jets using the anti-k_{T} algorithm [10]. Jets are required to have $p_{\mathrm{T}} > 25\,\mathrm{GeV/c}$ and to be separated from the selected lepton pair by at least $\Delta R = 0.5$. In addition, jets should be within the detector acceptance of $\eta < 2.1$ and fulfill loose jet identification criteria [11].

Identification of b-jets is done by exploiting b-hadron's long lifetime. The simple secondary vertex (SSV) algorithm discriminator [12] is computed from the three-dimensional flight distance significance. For the high-efficiency (HE) label, at least two tracks are required to be attached to the secondary vertex while for the high-purity (HP) one at least three tracks are required.

Finally, for further reduction of the $t\bar{t}$ + jets contribution, the invariant mass of the lepton pair should be in the range of 60 to $120\,\mathrm{GeV/c^2}$ and the transverse missing energy should be less than $40\,\mathrm{GeV/c}$.

4. Results

After applying all the cuts, a good agreement between data and MC is observed in both high-efficiency and high-purity selections, within 2σ at most of the expected statistical uncertainties on the data and without any rescaling of the MC to match the data. The event yields are summarized in table 1. The shapes of kinematic variables shown in figures 1 and 2 show good agreement between data and Monte Carlo predictions from MADGRAPH.

Using the results above, an important quantity for the treatment of heavy flavours in matrix element calculations and test of b-quark content of PDFs can be extracted, which is the ratio $\sigma(Z+b)/\sigma(Z+j)$. We calculate $\mathcal{R}$ as

$$\mathcal{R} = \frac{N^{\mathrm{data}}_{\mathrm{Z}+b} \times \mathcal{P} - N^{\mathrm{MC}}_{\mathrm{t\bar{t}}}}{\varepsilon_{\mathrm{MC}} \times N^{\mathrm{data}}_{\mathrm{Z}+j}}$$

Table 1. Number of events selected in data and MC samples, and purity and significance in $Z + b$ expected from MC. Only statistical uncertainties are shown. No data/MC scaling factor is applied on the MC.

Selection	Data	$t\bar{t}$	$Z + l$	$Z + c$	$Z + b$
HE *eeb*	54 ± 7	2.0 ± 0.1	13.4 ± 0.5	14.5 ± 0.7	35.5 ± 1.1
HP *eeb*	29 ± 5	1.60 ± 0.09	0.8 ± 0.1	5.3 ± 0.4	25.1 ± 0.9
HE $\mu\mu b$	91 ± 10	3.0 ± 0.1	18.5 ± 0.6	20.3 ± 0.9	54.4 ± 1.4
HP $\mu\mu b$	36 ± 6	2.4 ± 0.1	1.3 ± 0.2	7.2 ± 0.5	37.5 ± 1.2

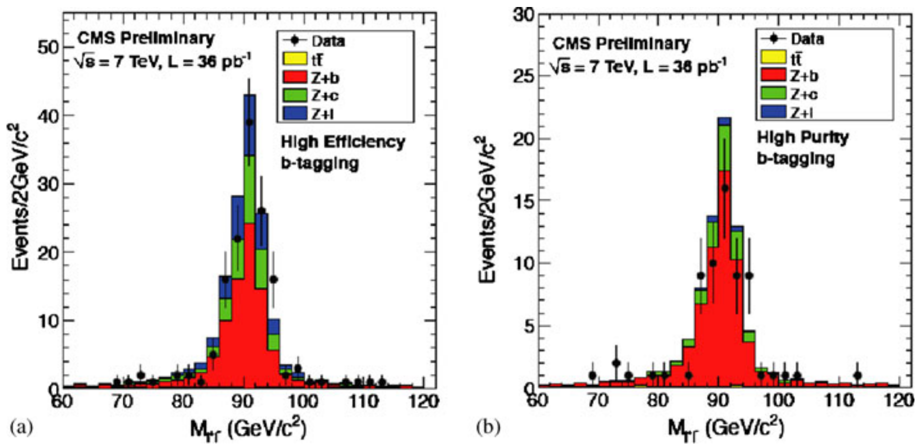

Figure 1. Invariant mass for the (**a**) high-efficiency and (**b**) high-purity selections.

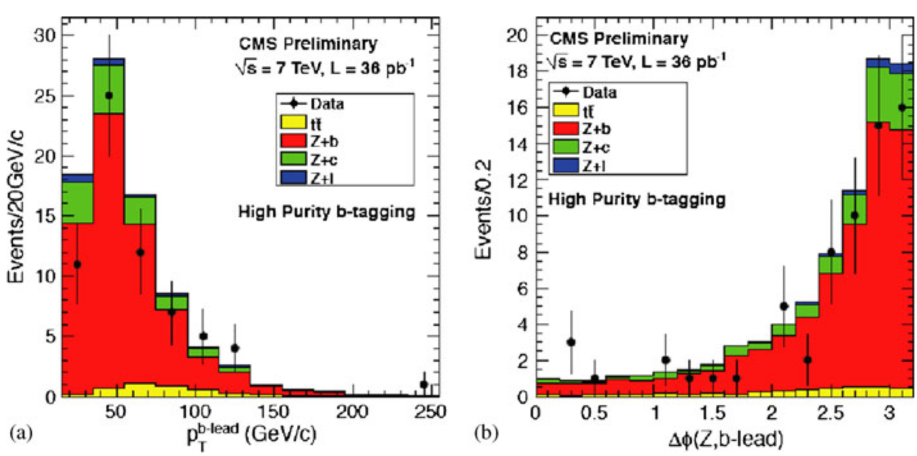

Figure 2. Leading b-jet p_{T} (**a**) and $\Delta\phi$ between the lepton pair and the leading b-jet (**b**) after the high-purity selection.

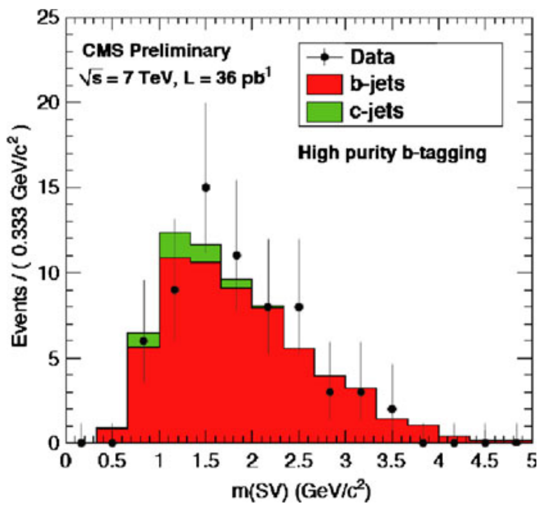

Figure 3. Binned likelihood fit of the mass of the secondary vertex in data events for HP selection.

where N^{data}_{Z+b} and N^{data}_{Z+j} are the number of data events remaining after the $Z + b$ (and MET) and $Z + j$ selections, $N^{\text{MC}}_{\text{t}\bar{\text{t}}}$ is the number of $t\bar{t}$ events expected from MC after the $Z + b$ and MET selection. ε_{MC} contains the b-tagging, MET selection efficiency and b-jet acceptance effects. The values are found to be 0.43 ± 0.01(stat) $\pm$ 0.09(syst) for HE and 0.30 ± 0.01(stat) $\pm$ 0.06(syst) for HP. A systematic uncertainty of 20% is assigned to these numbers, which is dominated by the uncertainty on the b-tagging efficiency (15%) [12]. The b-jet acceptance assumed here concerns the fraction of events with a real b-jet in which at least one jet lies within the acceptance cuts, but no b-jet does. The b-jet acceptance was found to be about 18% in the MC $Z + b$ sample. $\mathcal{P}$ is the purity in b-jet which contains both $Z + b$ and $t\bar{t}$ components. The purity is extracted by a binned likelihood fit to the mass of the secondary vertex in the data, using MC template functions for the b, c and light contributions (figure 3). The purity in b-jets is found to be 55 ± 9% (88 ± 11%) for the HE (HP) selection, which is in good agreement with the MC estimate

Table 2. Extracted ratio $\mathcal{R}$ from data, and MADGRAPH+PYTHIA samples including a prediction using NLO theory (MCFM).

Sample	$\mathcal{R}(Z \to ee)$ (%)
Data HE	4.3 ± 0.6(stat) ± 1.1(sys)
Data HP	5.4 ± 1.0(stat) ± 1.2(sys)
MADGRAPH	5.1 ± 0.2(stat) ± 0.2(sys) ± 0.6(th)
MCFM	4.3 ± 0.5(th)
Sample	$\mathcal{R}(Z \to \mu\mu)$ (%)
Data HE	5.1 ± 0.6(stat) ± 1.3(sys)
Data HP	4.6 ± 0.8(stat) ± 1.1(sys)
MADGRAPH	5.3 ± 0.1(stat) ± 0.2(sys) ± 0.6(th)
MCFM	4.7 ± 0.5(th)

of $57 \pm 3\%$ ($82 \pm 4\%$). Results for the ratio are given in table 2 with the NLO predictions from MCFM.

5. Conclusions

First observation of Z+b events is made at $\sqrt{s} = 7$ TeV using 36 pb^{-1} of CMS data. The ratio $\sigma(Z + b)/\sigma(Z + j)$ is found to be 0.054 ± 0.016 in the data for the Z$\to$ ee and 0.046 ± 0.014 for the Z$\to$ $\mu\mu$ selection. The estimation from NLO theory predictions is 0.043 ± 0.005 (0.047 ± 0.005) using the same kinematic selections on the leptons and jets as the selections applied to the data.

References

[1] The CMS Collaboration, *Measurement of CMS luminosity*, CMS PAS, EWK-10-004 (2010)
[2] J Alwall *et al*, *J. High Energy Phys.* **09**, 028 (2007)
[3] T Sjostrand, S Mrenna and P Z Skands, *J. High Energy Phys.* **05**, 026 (2006)
[4] The CMS Collaboration, *J. Instrum.* **3**, S08004 (2008)
[5] S Agostinelli *et al*, *Nucl. Instrum. Methods* **A506**, 250 (2003)
[6] F Maltoni, T McElmurry and S Willenbrock, *Phys. Rev.* **D72**, 074024 (2005), arXiv: hep-ph/0505014
[7] J M Campbell, R K Ellis, F Maltoni *et al*, *Phys. Rev.* **D73**, 054007 (2006), arXiv: hep-ph/0510362
[8] The CMS Collaboration, *Study of W and Z boson production at 7 TeV*, CMS PAS, EWK-10-005 (2011)
[9] The CMS Collaboration, *Commissioning of the particle-flow event reconstruction with the first LHC collisions recorded in the CMS detector*, CMS PAS, PFT-10-001 (2010)
[10] M Cacciari, G Salam and G Soyez, *J. High Energy Phys.* **04**, 063 (2008)
[11] The CMS Collaboration, *Jets in 0.9 and 2.36 TeV pp collisions*, CMS PAS, JME-10-001 (2010)
[12] The CMS Collaboration, *Performance of b-jet identification in CMS*, CMS PAS, BTV-10-001 (2011)