Available on CMS information server

CMS CR -2012/185

The Compact Muon Solenoid Experiment

Conference Report

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland

10 August 2012 (v5, 18 September 2012)

Heavy flavour and vector bosons associated production

Carlotta Favaro for the CMS Collaboration

Abstract

The mechanism of production of heavy-flavoured mesons, containing b or c quarks, in association with vector bosons, W or Z, in the Standard Model is only partially understood. The study of events with one or two well-identified and isolated leptons accompanied by b-jets or secondary vertices is therefore crucial to refine the theoretical calculations in perturbative QCD, as well as validate associated Monte Carlo techniques. The deep understanding of these processes is furthermore required by Higgs and BSM analyses with similar final states. Using the LHC proton-proton collision data collected in 2010 and 2011 at a centre of mass energy of 7 TeV by the CMS detector, preliminary measurements of the Z+b(b) cross sections and angular correlations are presented. Finally, the study of the W+c production rate with respect to the W charge and W+light jets rates allows to probe the strange quark content of the proton. These results are also presented.

Presented at *QCD 12: 16th International Conference in Quantum ChromoDynamics*

Heavy flavor and vector boson associated production

Carlotta Favaro *on behalf of the CMS Collaboration*

Universität Zürich, Physik-Institut, Winterthurerstr. 190, 8057 Zürich, Switzerland

Abstract

The mechanism of production of heavy-flavoured mesons, containing b or c quarks, in association with vector bosons, W or Z/γ^*, in the Standard Model is only partially understood. The study of events with one or two well-identified and isolated leptons accompanied by b-jets or secondary vertices is therefore crucial to refine the theoretical calculations in perturbative QCD, as well as validate associated Monte Carlo techniques. The deep understanding of these processes is furthermore required by Higgs and BSM analyses with similar final states. Using the LHC proton-proton collision data collected in 2010 and 2011 at a centre of mass energy of 7 TeV by the CMS detector, preliminary measurements of the Z/γ^*+b(b) cross sections and angular correlations are presented. Finally, the study of the W+c production rate with respect to the W charge and W+light jets rates allows to probe the strange quark content of the proton. These results are also presented.

Keywords: LHC, CMS, Z, W, bottom, charm

1. Measurement of Z and heavy flavor associated production

1.1. Motivations

The associated production of a Z/γ^* boson and b-quarks is of great relevance at the Large Hadron Collider.
It represents a significant background to the search of the Standard Model Higgs boson produced in association with a vector boson and decaying into a $b\bar{b}$ pair, and it's a possible discovery channel for models with an extended Higgs sector, such as the two Higgs-doublet model.
Besides, the Z/γ^* + heavy flavor production is a good test of perturbative QCD and of Monte Carlo simulation. The description of the dynamics of b-quarks remains problematic, especially in the soft/collinear region, dominated by the gluon-splitting contribution. Two different approaches are implemented in the simulation of the Z/γ^* + b production: the four-flavor, which excludes the b-quarks from the proton parton density functions, and the five-flavor, which instead integrates the b-quarks in the PDFs. The measurement of the inclusive Z/γ^* + b-jet cross-section, and of the angular correlations in Z/γ^* + bb events can help identifying the modeling that provides the best description of data.

1.2. Z/γ^* + b and Z/γ^* + bb cross-section measurements

The inclusive cross-section of the Z/γ^* produced in association with at least one b-jet [1] as well as the inclusive Z/γ^* + bb production cross-section [2] are measured in a data sample corresponding to an integrated luminosity of 2.2 fb^{-1}, collected in 2011 with the CMS detector [3].
The Z/γ^* boson is reconstructed in its leptonic decays $Z/\gamma^* \to \mu\mu$ and $Z/\gamma^* \to ee$. The event sample is thus selected by requiring the presence of two isolated muons (electrons) with $p_T > 20$ GeV (25 GeV) and $|\eta| < 2.1$ (2.5). Only dilepton pairs fulfilling the condition $60 < M(ll) < 120$ GeV for the Z/γ^* + b-jet cross-section measurement, and by $76 < M(ll) < 106$ GeV for the Z/γ^* + bb analysis, are retained. Jets are reconstructed from particle-flow objets [8] clustered with the anti-k$_T$ algorithm [9], and are required to have $p_T > 25$ GeV, $|\eta| < 2.1$, and to be separated from each of the leptons by $\Delta R > 0.5$. The b fraction is enhanced by applying the Simple Secondary Vertex (SSV) b-tagging algorithm [10], whose discriminator is a function of the distance between the primary interaction and the secondary vertex from the heavy flavor decay. The algorithm is applied in its High Purity working point, that requires at least three tracks associated to the vertex, in the Z/γ^* + b cross-section estimation, and in its High Efficiency working point, requiring at least two tracks, in the Z/γ^* + bb measurement.

The main backgrounds to the Z/γ^* + b-jet signal are Z/γ^*+ light quark/gluon and charm jets and top pair production. A maximum-likelihood fit of the distribution of the secondary vertex mass for the leading b-jet is used to estimate the b-jet purity $\mathcal{P}$, while the fraction of top pair events $f_{t\bar{t}}$ is extracted by extrapolating into the signal region the number of observed events in the Z upper sideband. The Z/γ^* + b-jet cross-section is computed from the number of selected events $N(ll+b)$, taking into account the b-tagging efficiency ϵ_b, the lepton acceptance $\mathcal{A}_l$ and selection efficiency ϵ_l, the correction factor C_{hadron} for detector and reconstruction effects, and the integrated luminosity $\mathcal{L}$, as in the formula

$$\sigma(Z/\gamma^* + b) = \frac{N(ll+b) \cdot (\mathcal{P} - f_{t\bar{t}})}{\mathcal{A}_l \cdot C_{hadron} \cdot \epsilon_l \cdot \epsilon_b \cdot \mathcal{L}} \tag{1}$$

The b-tagging and lepton selection efficiency are estimated from Monte Carlo simulation, and rescaled to reproduce the efficiencies measured on data with the tag-and-probe technique. The

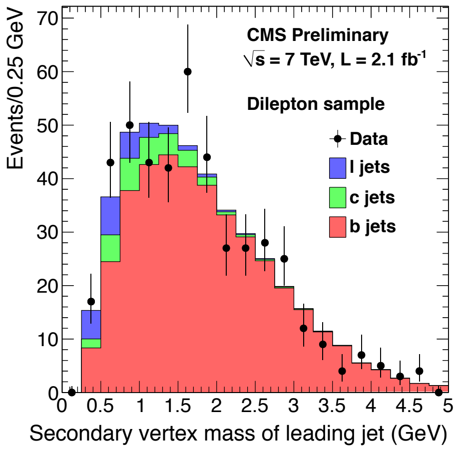

Figure 1: Secondary vertex invariant mass distribution for the leading b-jet in events with two leptons in the invariant mass range $76 < M(ll) < 106$ GeV and two b-jets. A maximum-likelihood fit of the vertex mass distribution is used to extract the b-jet purity $\mathcal{P}$.

correction for detector effects is instead evaluated on the simulation only.

The main systematic uncertainties are related to the b-tagging efficiency (10%) and b-jet purity (5%). The result is $\sigma(Z/\gamma^* + b) = 5.84 \pm 0.08(stat.) \pm 0.72(syst.)^{+0.25}_{-0.55}(theory)$ pb, referred to a phase space defined for the leptons by $60 < M(ll) < 120$ GeV, and for the jets by $p_T > 25$ GeV and $|\eta| < 2.1$. The result is 30% higher than the hadron-level-corrected NLO prediction of $\sigma(Z/\gamma^* + b) = 3.97 \pm 0.47$ pb by MCFM.

The procedure to extract the Z/γ^* + bb inclusive cross-section is similar to the one described above. The requirement of the presence of two b-jets suppresses the contamination from Z/γ^* + light quark/gluon and charm jets and increases the b-jet purity of the sample. The latter is estimated through a simultaneous maximum-likelihood fit of secondary vertex mass distribution the leading and sub-leading b-tagged jets, as shown in Fig.1, and is estimated to be about 84%. In order to reduce the top background, a cut is applied on the missing transverse energy $\not{E}_T < 50$ GeV. The remaining top fraction, about 20%, is evaluated through a maximum-likelihood fit of the dilepton invariant mass distribution. A small contamination from diboson production is instead estimated from Monte Carlo simulation.
The hadron-level cross-section of Z/γ^* + bb is extracted by applying a set of unfolding matrices, accounting for bin-to-bin migrations in the b-jet multiplicity due to dilepton efficiency and acceptance, b-tagging efficiency, and detector and reconstruction effects. The result is $\sigma(Z/\gamma^* + bb) = 0.37 \pm 0.02(stat.) \pm 0.07(syst.) \pm 0.11(theory)$ pb, in good agreement with the tree-level prediction of $\sigma(Z/\gamma^* + bb) = 0.33 \pm 0.01$ pb by MADGRAPH [6].
Some discrepancies are though observed in the event kinematics, as shown in Fig. 2.

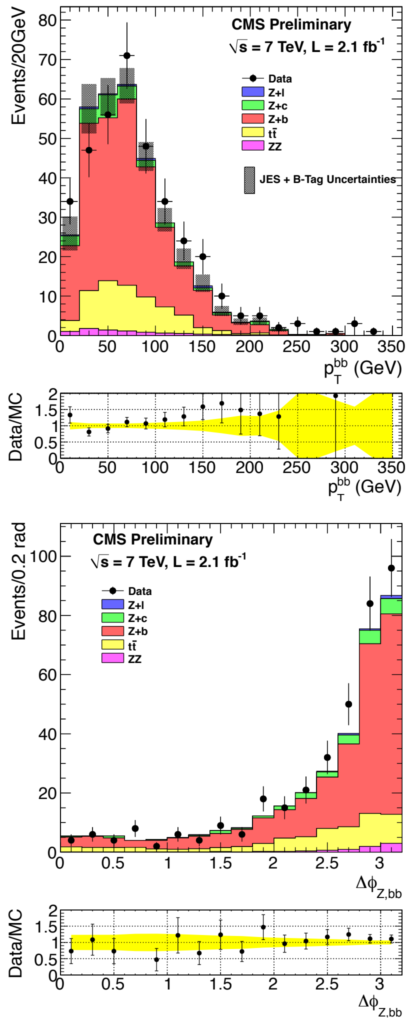

Figure 2: Transverse momentum of the b-jet pair (*top*) and transverse angular separation between the Z/γ^* and the b-jet pair (*bottom*), in events with two b-jets. The yellow band in the bottom part represents the statistical uncertainty on the Monte Carlo yields.

1.3. *Angular correlation between B-hadrons produced in association with a Z boson*

The normalized differential cross-section of Z/γ^*+BB-hadrons as function of the angular separation between the B-hadrons, defined by the variable $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$, is measured using a data sample corresponding to 4.6 pb^{-1} collected in 2011 [4].
A sample of Z/γ^* events is selected using the same criteria as for the Z/γ^*+b-jet cross-section measurement. The identification of the B-hadrons is performed through the reconstruction of the displaced secondary vertex from their decay, with no use of jets, providing unprecedented sensitivity to B pair production at small angular separation. The measurement refers to the kinematic region defined by two B-hadrons with $p_T > 15$ GeV and $|\eta| < 2$ and two oppositely charged leptons with $p_T > 20$ GeV, $|\eta| < 2.4$ and invariant mass $60 < M(ll) < 120$ GeV.

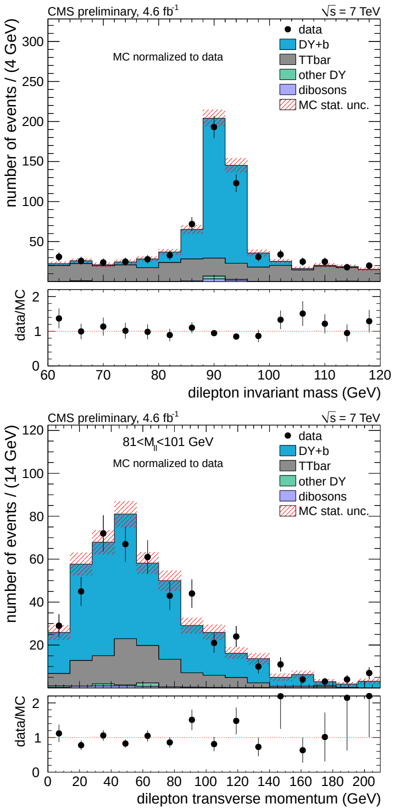

Figure 3: *Top* Dilepton invariant mass distribution in events with two B candidates, used to extract the signal yield and reject the top contamination through an extended maximum likelihood fit. *Bottom* Dilepton transverse momentum in the signal-enhanced region.

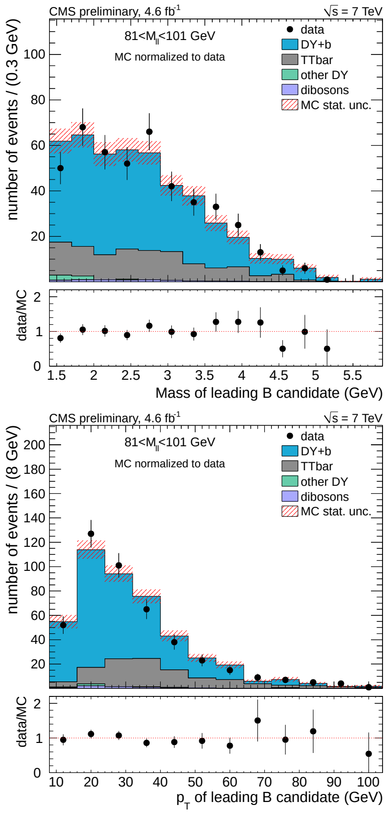

Figure 4: Secondary vertex invariant mass (*top*) and transverse momentum (*bottom*) for events with two B candidates and two leptons in the signal-enhanced invariant mass range $81 < M(ll) < 101$ GeV. The red bands show the statistical uncertainty of the Monte Carlo simulation.

The distributions of the most relevant kinematic variables are in good agreement between the data and the simulation, as it is shown in Fig. 3 and 4.

The main background is top pair production, which represents 50% of the total selected sample. The signal yield in each ΔR bin N_i^{data} is extracted through an extended unbinned maximum likelihood fit of the dilepton invariant mass distribution. N_i^{data} is corrected for the dilepton selection efficiency and acceptance $\epsilon_i^{\ell} \cdot \mathcal{A}^{\ell}$, for the B purity $\mathcal{P}_i$ and for the B pair identification efficiency ϵ_i^{2B}. Thanks to the very good resolution in the B-hadron flight directions, provided by the excellent tracking performance at CMS, no unfolding procedure is required to extract the hadron-level cross-section. The dilepton selection efficiency is evaluated from Monte Carlo simulation, and rescaled to match the efficiency measured in data. The B reconstruction efficiency and B purity can be estimated from simulation, given the good description of the event dynamics. The ΔR distribution is normalized by the total integral $\sigma_{visible}$. The dominant systematic uncertainties are related to the B identification efficiency (9%) and to the limited size of the Monte Carlo samples (from 6 to 10%). The result is shown in Fig.5. The observed curve is compared to the tree-level prediction by MADGRAPH in the five-flavor scheme, which includes the contribution of B hadron pairs and Z originating from separate partonic interactions, and to the NLO prediction by aMC@NLO [7]. The comparisons show a fair agreement, although data seems to suggest a flatter trend.

2. Measurement of *W* boson and charm associated production

At the LHC, the $pp \to W+c+X$ final state, mostly produced by the $\bar{s}g \to W^+\bar{c}$ and $sg \to W^-c$ processes, is a probe of the strange and antistrange quark parton density functions of the proton at the electroweak scale. The processes $\bar{d}g \to W^+\bar{c}$ and $dg \to W^-c$ participate as well, with contributions of 5% and 15% respectively. Assuming s and $\bar{s}$ having approximately the same parton density function, the ratio $R_c^{\pm} = \sigma(W^+\bar{c})/\sigma(W^-c)$ is thus predicted to be about 0.9. The measurement of the ratios $R_c^{\pm}$, and $R_c = \sigma(W+c)/\sigma(W+jets)$, which are free from many theoretical and experimental uncertainties, is performed using a data sample corresponding to an integrated luminosity of 36 pb^{-1} collected in 2010 [5]. A sample of W + jets events is selected by reconstructing the W boson through its leptonic decay $W \to \mu\nu$, and by requiring the presence of an additional jet with $p_T > 20$ GeV and $|\eta| < 2.1$. The W

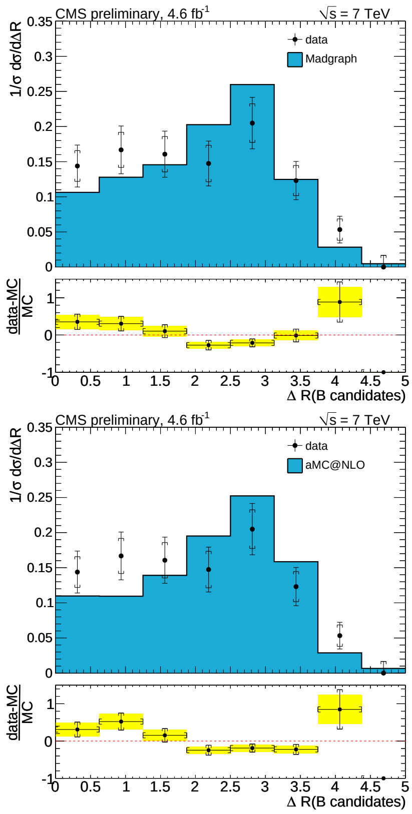

Figure 5: *Top*: Normalized production cross section as function of ΔR. The statistical errors in data are shown by the square brackets, while bars terminated by the straight segment represents the sum of systematic and statistical uncertainties. The `MADGRAPH` prediction in the five-flavor schema is also shown. Distributions are normalized to one. *Bottom*: as for top but with the `aMC@NLO` distribution. The yellow band in the bottom part represents the systematic uncertainty while the line segments represent the statistical uncertainties.

+ c fraction is enhanced by requesting the presence of a displaced secondary vertex with two or more associated tracks, as in the Simple Secondary Vertex High Efficiency (SSVHE) algorithm. The discriminator D_{SSVHE} is a function of the vertex decay length significance, and can assume negative values due to detector resolution effects. The main backgrounds are top production (20%) and W + light quark/gluon jets (25% of the selected sample). The signal yield is extracted through a maximum likelihood fit of the D_{SSVHE} discriminant, extended to the negative part of the distribution, as shown in Fig.6. Additional components, such as W + b and Drell-Yan, are estimated from Monte Carlo simulation. The obtained yields are $N(W^+\overline{c}) = 252.7 \pm 39.6(stat.)$ and $N(W^-c) = 273.5 \pm 38.0(stat.)$, to be corrected for the selection efficiency and acceptance, and for the efficiency of charm identification. The main systematic uncertainties on $R_c^{\pm}$ are related to the PDFs (2.2%), to pile-up effects (1.8%) and to the templates used to describe the top background (1.7%). The R_c measurement is mostly affected by the uncertainties on the secondary vertex reconstruction (14.1%)

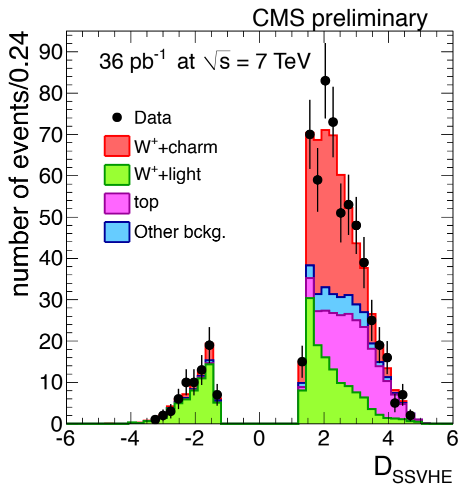

Figure 6: Fit of the Simple Secondary Vertex b-tagging discriminant (D_{SSVHE}) to extract the signal yield, for the $W^+\overline{c}$ sample. Negative values are mostly due to resolution effects. A similar distribution is observed for the charge conjugate sample.

and on the top background templates (6.2%). The final results are $R_c^{\pm} = 0.92 \pm 0.19(stat.) \pm 0.04(syst.)$, which is consistent with the expectation, and $R_c = 0.143 \pm 0.015(stat.) \pm 0.024(syst.)$, in good agreement with the NLO prediction $R_c = 0.13 \pm 0.02$ (`MCFM` with CT10 PDFs).

3. Conclusions

The cross-section of Z/γ^* produced in association with at least one or two b-jets are measured at CMS with a data sample corresponding to an integrated luminosity of 2.2 fb^{-1}. The results are found to be higher than the tree-level and NLO predictions. In addition, some discrepancies are observed in the event dynamics. The normalized differential cross-section of Z/γ^* + BB as function of the angular separation between B-hadrons is measured with 4.6 fb^{-1}. Despite a fair agreement within the experimental errors, the distribution is observed to be flatter than the tree-level and NLO predictions. Finally, the W+c production is studied with 36 pb^{-1}. The results are in good agreement with the expectations.

4. Bibliography

[1] CMS Collaboration, “Measurement of the Z/γ^* + b-jet cross-section in pp collisions at $\sqrt{s}$ = 7 TeV”, arXiv:1204.1643v1, submitted to *JHEP*

[2] CMS Collaboration, “Measurement of the Z/γ^* + bb-jets cross-section in pp collisions at $\sqrt{s}$ = 7 TeV with the CMS detector”, *CMS PAS* SMP-12-003 (2012)

[3] CMS Collaboration, “The CMS experiment at the CERN LHC”, JINST 3:S08004, 2008.

[4] CMS Collaboration, “Angular correlation between B-hadrons produced in association with a Z boson in pp collisions at $\sqrt{s}$ = 7 TeV”, *CMS PAS* EWK-11-015 (2012)

[5] CMS Collaboration, “Measurement of associated charm production in W final states at $\sqrt{s}$ = 7 TeV ”, *CMS PAS* EWK-11-003

[6] J. Alwall, P. Demin, Pavel, S. de Visscher, R. Frederix, M. Herquet et al., “MadGraph/MadEvent v4: The New Web Generation”, *JHEP* 0709 (2007) 028, arXiv:0706.2334

[7] R. Frederix et al., “W and Z/gamma* boson production in association with a bottom-antibottom pair”, *JHEP* 09 (2011) 061, arXiv:1106.6019
[8] 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)
[9] M. Cacciari, G. Salam, and G. Soyez, “The anti-kt jet clustering algorithm”, *JHEP* 04 (2008) 063
[10] CMS Collaboration, “Commissioning of b-jet identification with pp collisions at $\sqrt{s}$ = 7 TeV”, *CMS PAS* BTV-10-001 (2010)