DRAFT CMS Physics Analysis Summary

The content of this note is intended for CMS internal use and distribution only

2013/03/07

Head Id: 175386

Archive Id: 175385:175386 Archive Date: 2013/03/07 Archive Tag: trunk

Measurement of the W($\mu\nu$)bb̄ production cross section in pp collisions at $\sqrt{s}=7\,\text{TeV}$

The CMS Collaboration

Abstract

The production of W bosons in association with b-quarks is studied using proton-proton collisions at $\sqrt{s}=7$ TeV with the CMS experiment at the LHC in a data sample corresponding to an integrated luminosity of $5.0\,\mathrm{fb^{-1}}$. The Wb $\overline{\mathrm{b}}$ events are selected in the W $\rightarrow \mu\nu$ decay mode with a muon of $p_\mathrm{T}>25\,\mathrm{GeV}$ and $|\eta|<2.1$, and exactly two b-jets with $p_\mathrm{T}>25\,\mathrm{GeV}$ and $|\eta|<2.4$. The measured production cross section $0.53\pm0.05\,\mathrm{(stat.)}\pm0.09\,\mathrm{(syst.)}\pm0.06\,\mathrm{(theo.)}\pm0.01\,\mathrm{(lum.)}$ pb is found to be in good agreement with standard model predictions.

This box is only visible in draft mode. Please make sure the values below make sense.

PDFAuthor: Isobel Ojalvo, Maria Cepeda, Alexander Savin, Sridhara Dasu, Matthew

Herndon, Tom Perry

PDFTitle: Measurement of the W+bb cross section with the CMS detector

PDFSubject: SMP

PDFKeywords: CMS, physics, SMP

Please also verify that the abstract does not use any user defined symbols



- This Letter describes a study of the W boson production in association with two b-quarks,
- where W bosons are observed via their decays to muons and b-quarks are identified as two
- 3 separated jets. This production channel provides an important testing ground for standard
- 4 model (SM) predictions. Previous measurements have shown varying levels of agreement with
- $_{5}$ theoretical calculations [1–3], therefore performing the experimental measurements of the Wb $\overline{
 m b}$
- 6 cross section with high accuracy is important both for understanding the SM expectations and
- for constraining the Wbb background contribution in searches for new physics.
- The analysis uses a sample of proton-proton (pp) collisions at a center-of-mass energy of \sqrt{s}
- 9 7 TeV collected in 2011 with the CMS experiment at the LHC and corresponding to an integrated
- luminosity of $5.0 \, \text{fb}^{-1}$.

While the CMS detector is described in detail elsewhere [4], the key components for this anal-11 ysis are summarized here. The CMS experiment uses a right-handed coordinate system, with 12 the origin at the nominal interaction point, the x axis pointing to the center of the LHC ring, the y axis pointing up (perpendicular to the plane of the LHC ring), and the z axis along the counterclockwise-beam direction. The polar angle θ is measured from the positive z axis and 15 the azimuthal angle ϕ is measured in the x-y plane. The absolute value of the transverse 16 momentum (p_T) is calculated as $p_T = \sqrt{p_x^2 + p_y^2}$. A superconducting solenoid occupies the 17 central region of the CMS detector, providing an axial magnetic field of 3.8 T parallel to the 18 beam direction. The silicon pixel and strip tracker, the crystal electromagnetic calorimeter, 19 and the brass/scintillator hadron calorimeter are located within the solenoid. A quartz-fiber 20 Cherenkov calorimeter extends the coverage to $|\eta|$ < 5.0, where pseudorapidity is defined as 21 $\eta = -\ln[\tan{(\theta/2)}]$. Muons are measured in gas ionization detectors embedded in the steel 22 flux return yoke outside the solenoid. The first level of the CMS trigger system, composed of 23 custom hardware processors, is designed to select the most interesting events in less than 3 µs 24 using information from the calorimeters and muon detectors. The high-level-trigger processor 25 farm decreases the event rate from 100 kHz delivered by the first level trigger to a few hundred 26 hertz, before data storage. 27

A number of Monte Carlo (MC) event generators are used to simulate the signals and backgrounds. The events with the W or Z boson production, or with tt are generated with MAD-GRAPH 5.1 [5]. Single top samples are generated at next-to-leading order (NLO) with 2.0 [6–8]. 30 Diboson (W⁺W⁻, WZ, ZZ) and QCD samples are generated with PYTHIA 6.4 [9]. For LO gen-31 erators, the default set of parton distribution functions (PDF) used to produce these samples 32 is CTEQ6L [10], while MSTWNNLO [11] is used for NLO generators. For all processes, the 33 detector response is simulated using a detailed description of the CMS detector, based on the GEANT4 package [12], and event reconstruction is performed with the same algorithms as used 35 for data. The simulated samples include additional interactions per bunch crossing (pileup) 36 where pileup distribution matches the data. Parton showering is simulated with PYTHIA us-37 ing the Z2 tune [13].

A complete reconstruction of the individual particles emerging from each collision event is obtained via a particle-flow (PF) technique [14, 15]. This approach uses the information from all CMS sub-detectors to identify and reconstruct individual particles in the collision event, classifying them into mutually exclusive categories: charged hadrons, neutral hadrons, photons, electrons, and muons.

Muons are reconstructed [16] by combining the information of the tracker and the muon spectrometer, The muon candidates are required to be compatible with the primary vertex of the event, which is chosen as the vertex with highest $\sum p_T^2$ of its associated tracks. The muon rela-

tive isolation is defined as

$$I^{\text{rel}} = \left(\sum_{i} p_{\text{T, charged}}^{i} + \sum_{j} E_{\text{T, neutral}}^{j}\right) / p_{\text{T}}^{\mu},$$

with the sums running over all the charged and neutral PF candidates, excluding the muon candidate itself, within a cone around the muon direction defined by $\Delta R < \Delta R_{\text{max}} = 0.4$, where $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}$, and E_{T} stands for the transverse energy.

These identified, isolated muons are then combined with the missing transverse energy \vec{E}_{T}^{miss} of the event to form a leptonic W candidate. The missing transverse energy \vec{E}_{T}^{miss} is defined as the negative vector sum of the transverse momenta of all reconstructed particles in the event, with $E_{T}^{miss} = |\vec{E}_{T}^{miss}|$, and is corrected using procedure described in Ref. [17].

Transverse mass is defined as

$$m_{\mathrm{T}} = \sqrt{2p_{\mathrm{T}}^{\ell}E_{\mathrm{T}}^{\mathrm{miss}}(1-\cos\Delta\phi)}$$
 ,

where $\Delta\phi$ is the difference in azimuth between $ec{E}_{
m T}^{
m miss}$ and $ec{P}_{
m T}^{\mu}$.

Jets are reconstructed from the PF candidates. The anti- k_T clustering algorithm [18] with distance parameter of 0.5 is used, as implemented in the FASTJET package [19, 20]. The jet energy is corrected for pileup in a manner similar to the correction of the energy inside a lepton isolation cone. Jet energy corrections are also applied as a function of the jet p_T and η [21].

Secondary vertexes (SV) are reconstructed inside each jet. This study makes use of the combined secondary vertex (CSV) b-tagging algorithm [22]; this algorithm makes use of the long lifetime and heavy-flavor of b-hadrons. The CSV algorithm employs likelihood ratio technique to optimize b-tagging discrimination, by combining information about impact parameter significance, secondary vertex, and jet kinematics in a likelihood ratio technique. The chosen tight b-jet selection criteria (CSV > 0.898) achieves approximately 15% efficiency for jets containing b-flavored hadrons, while suppressing light-quark jets by a factor of 100.

The Wb \overline{b} selected events are required to have an isolated muon with $I^{\rm rel}<0.12$, and $p_{\rm T}>0.12$, where both selected jets should pass the tight b-tagging CSV requirement and contain a secondary vertex.

To reduce the background contributions from Z-boson production, the event is rejected if there is a second muon, without any requirements on the isolation and $p_{\rm T}$, which builds with the selected isolated muon a dimuon system with invariant mass $m_{\mu\mu} > 60\,{\rm GeV}$. The ${\rm t\bar{t}}$ background is reduced by requiring that there are no isolated electrons or muons with $p_T > 20\,{\rm GeV}$ in the event and no jets with $p_T > 25\,{\rm GeV}$ and $2.4 < |\eta| < 4.5$. To reduce contribution from QCD multijet events $m_T > 45\,{\rm GeV}$ is also required.

After all the selection requirements the significant background contributions are: tt, single top,
W+jets, Z+jets and QCD. Contributions of all these backgrounds are estimated in a simultaneous fit, which provides a final estimate for the signal and background yields.

With the exception of QCD, the shapes of the background distributions are taken from simulation. A shape for the QCD contribution is obtained directly from data from a background-enriched control region defined by all the selection requirements, but requiring the muon to be not isolated ($I^{rel} > 0.2$). The QCD uncertainty in the final fit is taken to be $\pm 50\%$; this uncertainty comes directly from the uncertainty on the fit in the QCD background-enriched control region.

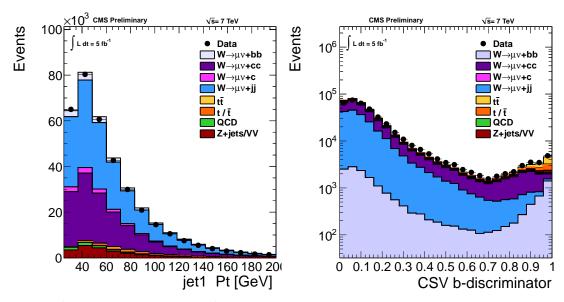


Figure 1: (left) The highest p_T jet before applying b-tagging. (right) The Combined Secondary Vertex b-discriminator is used to select b-Jets and reject light jets. the CSV b-discriminator for the high p_T jet.

The initial yields are taken either from data, in estimates based on the control regions, or from simulation, normalized to the NNLO predictions. The shapes and normalizations of the background distributions are validated in data with a set of control regions.

The W+jets (u,d,c,s,g) process, where the jets are not initiated by b-quarks, is the dominant background before applying the selection requirements on secondary vertex and b-tagging requirements. Figure 1(left) shows the p_T of the leading jet at this preselected stage. The CSV algorithm working point which provides maximum reduction of W+jets (u,d,c,s,g) is used. The CSV btagging discriminant for the leading jet is shown in figure 1(right). The presence of light and charm jets in the sample is almost negligible at the higher values of the discriminant. Furthermore, to increase the purity of the sample a secondary vertex is required to be reconstructed in each of the selected jets. These selection requirements have been validated in the tt and Z+jets control regions described below.

A single-top control region is defined by selecting events in which the W boson is accompanied by exactly one b-jet passing the described tagging criteria, and an additional forward jet with $|\eta| > 2.8$. No additional vetoes on extra light jets or leptons are applied. The simulation describes the single-top production control region well and therefore it is used to estimate the yield and shape of the distributions in the signal region.

The events selected for the $t\bar{t}$ control region pass the selection requirements, with no restrictions on number of leptons in the event. In addition to two highest-pT b-tag jets, the events are required to have at least two extra light jets. This higher jet multiplicity requirement selects a sample that is dominated by $t\bar{t}$ events. Fig. 2 (right) shows the invariant mass of the two highest- p_T additional jets (3rd and 4th highest- p_T in the event). In ttbar events this distribution reconstructs the mass of the hadronically decaying W boson. It is used in the final fit to extract the $t\bar{t}$ background normalization. The simulation describes data well both in the shape and normalization.

The Z+jets background estimate is validated in a control region where the standard selection is applied except that a second muon is required, $70 < m_{\mu\mu} < 100\,\text{GeV}$. Agreement between data and MC is observed in this region.

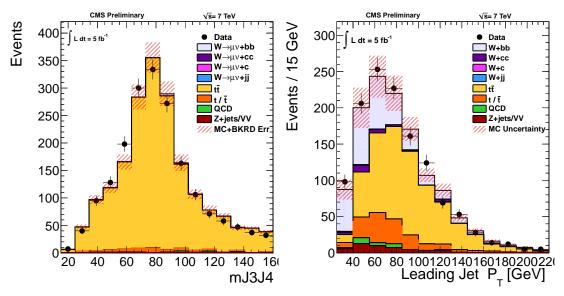


Figure 2: (right) p_T distribution of the leading- p_T jet, normalized to the result of the binned maximum likelihood fit. (left) The $t\bar{t}$ control region invariant mass of the two additional light jets in the $t\bar{t}$ control region, also normalized to the results of the fit.

One of the dominant systematic uncertainties comes from the background normalization, the uncertainties on the b-tagging is taken from Ref. [23]. The jet (muon) energy scales are allowed to vary up and down by one standard deviation (0.5%) and is added to the fit as a binned shape variation. The uncertainty associated to pileup description in MC is estimated by shifting the overall mean of the number of vertexes up or down by 0.6 bunch crossing. To account for $E_{\rm T}^{\rm miss}$ uncertainty the component of $E_{\rm T}^{\rm miss}$ that is not clustered in jets is scaled by $\pm 10\%$. Uncertainties due to muon efficiency estimation (triggering, identification, isolation) are found to be 1%. The luminosity uncertainty, 2.2%, is taken from Ref. [24]

In summary, the systematic uncertainties considered in this analysis are: background contributions from non Wbb processes, b-tagging efficiencies and inefficiencies, total Jet Energy Scale, Muon Energy Scale, Pile Up Multiplicity, Unclustered Energy to account for MET uncertainties, uncertainties due to trigger simulation and lepton reconstruction, and QCD Yield uncertainty.

The final yields are extracted via a binned maximum likelihood fit. To constrain the most prominent backgrounds and reduce the final systematic uncertainty the fit is performed simultaneously on the p_T of the leading jet (J_1) in the signal region after all selection requirements have been applied, and on the m_{ij} distribution obtained from the $t\bar{t}$ control region. The leading jet (J_1) p_T is chosen as final fit variable due to its discrimination power against top related backgrounds. Figure 2 shows the two fitted distributions, leading jet (J_1) p_T in the signal region (left) and W Mass in the $t\bar{t}$ control region (right), normalized to the results of the fit.

The statistical and systematic uncertainties are introduced in the form of nuisance parameters via log-normal distributions around the estimated central values. The fitted yields for each one of the processes can be found in Table 1, compared to the MC predictions. To calculate the final systematic and statistical uncertainties, the statistical and systematic errors are calculated simultaneously then the systematic nuisance parameters are removed, the fit is re-run and the statistical uncertainty is obtained. The total systematic error is calculated by subtracting in quadrature the statistical uncertainty from the total error.

The observed number of events in data after selection in the signal region is $N(S+B)_{data} = 1230 \pm 35$. After background subtraction, the number of signal events is 300^{+63}_{-56} .

Process	Wbb	W+light	Wcc	Z+jets	$t\bar{t}$	Single Top	VV	QCD
MC prediction	332	1.5	21.	30.9	596	160	19	33
	±66	± 0.2	$\pm 4.$	±2.9	±35	± 13	±3	±17
Fitted	300	1	20	32	647	170	17	33.
Yields	-56/+63	±1	± 4	±3	±52	± 13	±3	±16

Table 1: Comparison of the expected (before the fit) and measured (after the fit) yields for each of the processes.

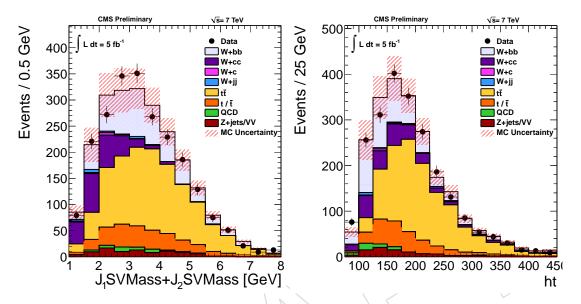


Figure 3: The distribution of J_1+J_2 SV mass (right) and h_T of the system (left) in the alternate two medium b-Tag selection, normalized to the results of the cross-check fit.

To show the stability of the $Wb\bar{b}$ fit result a separate study was performed with two selected b-tagged jets that require each jet to have a b-discriminator value greater than 0.679 rather than the requirement of 0.898. This looser b-tagging working point corresponds to an efficiency of 20% for jets containing b-flavored hadrons, while suppressing light-quark jets by a factor of 30. With the exception of the modification to the b-discriminator CSV variable, all other selections for the signal and control region remain unchanged. The $Wc\bar{c}$ contribution is non-negligible with this selection, therefore, the sum of the invariant masses of the secondary vertex found in each selected jet is used to distinguish between $Wb\bar{b}$ and $Wc\bar{c}$. H_T is used to distinguish W+jets from top contributions. The $Wb\bar{b}$ signal is extracted via a two dimensional fit to H_T versus the secondary vertex mass of the highest p_T + second p_T jet. An equivalent $t\bar{t}$ control region to the one described in the tighter selection, based on the reconstruction of the W mass using two light jets, is also used in this case. The variables J_1+J_2 SV mass and H_T are shown in Figure 3, with yields normalized to the results of the fit. This alternative fit method is found to be consistent with the primary fit quoted in this analysis within 4%.

The Wbb cross section within the reference fiducial phase space is determined using the following Equation:

$$\sigma(pp \to Wb\overline{b} + X) = \frac{N_{Data} - N_{Bckg}}{\mathcal{L}_{int} \epsilon_{sel}},$$

where the efficiency of the selection requirements $\epsilon_{sel}=10.4\pm1.0\%$ is estimated with MAD-GRAPH. The associated errors correspond to PDFs and to the factorization/matching scales.

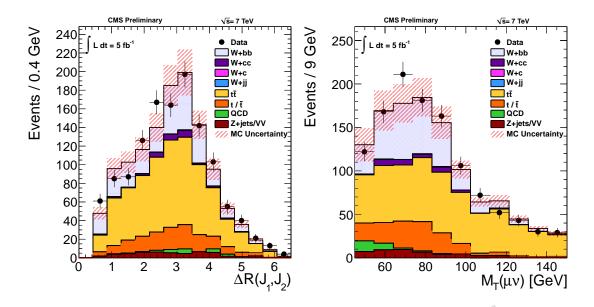


Figure 4: (left) The Δ R between the two selected b-jets (right) the m_T distribution, normalized to the results of the fit.

The measured cross section is

$$\sigma(pp \to Wb\bar{b}) Br(W \to \mu\nu) = 0.53 \pm 0.05 \text{ (stat.)} \pm 0.09 \text{ (syst.)} \pm 0.06 \text{ (theo.)} \pm 0.01 \text{ (lum.) pb.}$$

This cross section is calculated at the level of final state particles, by requiring a muon with $p_T > 25$ GeV and $|\eta| < 2.1$ and exactly two jets, reconstructed using the anti-kt jet algorithm with distance parameter 0.5, with $p_T > 25$ GeV and $|\eta| < 2.4$ and each containing at least one b-hadron with $p_T > 5$ GeV. Events with extra jets are vetoed. A correction factor $C_{b \to B} \approx 0.92 \pm 0.01$ to extrapolate from the final-state-particles level to the parton-level cross section is estimated with MADGRAPH, where at the parton level the events are required to have a muon with $p_T > 25$ GeV and $|\eta| < 2.1$ and exactly two parton-jets with $p_T > 25$ GeV and $|\eta| < 2.4$ each containing a b parton just before hadronization. This factor has been computed in the 4-flavor and 5-flavor schemes, and the difference between said calculations is assigned as systematic uncertainty to the measurement.

The measured values can be compared to the NLO cross section 0.52 ± 0.03 pb calculated with MCFM [25, 26]. The MCFM cross section is calculated using the MSTW2008NNLO [11] PDF, and $\mu_F = \mu_R = m_W + 2m_b$, and assumes massive bs. The cross section error is estimated by varying the μ_{FR} simultaneously up and down by a factor two.

The distributions of the angle ($\Delta R = \sqrt{\Delta \phi^2 + \Delta \eta^2}$) between the two selected b-jets and the transverse mass ($m_{\rm T}$) of the W are presented in Fig 4 (left), compared to the Monte Carlo predicted shapes with yields normalized to the fit results. Furthermore, in Fig 5 (right) the invariant mass and the transverse momentum of the two selected b-jets are presented. The data is well described by the simulation in both cases.

In summary, we have presented a measurement of the Wb \overline{b} production cross section in proton-proton collisions at 7 TeV. The Wb \overline{b} have been selected in the W $\rightarrow \mu\nu$ decay mode with a muon of $p_T > 25$ GeV and $|\eta| < 2.1$, and two b-jets with $p_T > 25$ GeV and $|\eta| < 2.4$. The data sample corresponds to an integrated luminosity of 5.0 fb $^{-1}$. The measured cross section $\sigma(\text{Wb}\overline{b})\text{Br}(\text{W} \rightarrow \mu\nu) = 0.53 \pm 0.05$ (stat.) ± 0.09 (syst.) ± 0.06 (theo.) ± 0.01 (lum.) pb for exclusive production of a W boson is association with two b jets is in agreement with SM predictions.

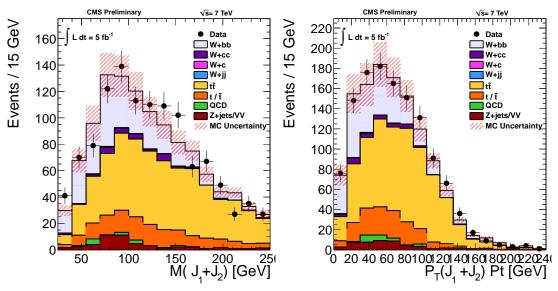


Figure 5: (left) The invariant mass of the two selected b-jets and (right) the $p_T(JJ)$ distribution, normalized to the results of the fit.

This results is approaching the precision of theoretical predictions at NNLO, allowing a sensitive test of perturbative calculations in the SM.

References

- [1] CDF Collaboration Collaboration, "First Measurement of the *b*-Jet Cross Section in Events with a *W* Boson in $p\overline{p}$ Collisions at $\sqrt{s}=1.96$ TeV", *Phys. Rev. Lett.* **104** (Apr, 2010) 131801. doi:10.1103/PhysRevLett.104.131801.
- 184 [2] D0 Collaboration, "Measurement of the $p\bar{p} \to W + b + X$ production cross section at $\sqrt{s} = 1.96$ TeV", *Phys. Lett. B* **718** (2013).
- 186 [3] ATLAS Collaboration, "Measurement of the cross section for the production of a W boson in association with b jets in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector", Phys. Lett. B **707** (2012).
- ¹⁸⁹ [4] CMS Collaboration, "The CMS experiment at the CERN LHC", JINST **3** (2008) S08004. doi:10.1088/1748-0221/3/08/S08004.
- [5] J. Alwall, M. Herquet, F. Maltoni et al., "MadGraph 5: going beyond", JHEP 06 (2011)
 128, arXiv:1106.0522. doi:10.1007/JHEP06 (2011) 128.
- [6] S. Alioli, P. Nason, C. Oleari et al., "NLO vector-boson production matched with shower in POWHEG", JHEP 07 (2008) 060, arXiv:0805.4802.
 doi:10.1088/1126-6708/2008/07/060.
- P. Nason, "A new method for combining NLO QCD with shower Monte Carlo algorithms", JHEP 11 (2004) 040, arXiv:hep-ph/0409146.
 doi:10.1088/1126-6708/2004/11/040.
- [8] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method", JHEP 11 (2007) 070, arXiv:0709.2092. doi:10.1088/1126-6708/2007/11/070.

- [9] T. Sjöstrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 physics and manual", *JHEP* **05** (2006) 026, arXiv:hep-ph/0603175. doi:10.1088/1126-6708/2006/05/026.
- [10] H.-L. Lai, J. Huston, Z. Li et al., "Uncertainty induced by QCD coupling in the CTEQ global analysis of parton distributions", *Phys. Rev. D* **82** (2010) 054021, arXiv:1004.4624. doi:10.1103/PhysRevD.82.054021.
- 207 [11] A. D. Martin, W. J. Stirling, R. S. Thorne et al., "Parton distributions for the LHC", Eur. 208 Phys. J. C 63 (2009) 189, arXiv:0901.0002. 209 doi:10.1140/epjc/s10052-009-1072-5.
- [12] GEANT4 Collaboration, "GEANT4—a simulation toolkit", Nucl. Instrum. Meth. A **506** (2003) 250. doi:10.1016/S0168-9002(03)01368-8.
- [13] R. Field, "Early LHC Underlying Event Data Findings and Surprises", arXiv:arXiv:1010.3558 [hep-ph].
- ²¹⁴ [14] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET", CMS Physics Analysis Summary CMS-PAS-PFT-09-001, (2009).
- 216 [15] CMS Collaboration, "Commissioning of the Particle–Flow reconstruction in Minimum–Bias and Jet Events from pp Collisions at 7 TeV", CMS Physics Analysis Summary CMS-PAS-PFT-10-002, (2010).
- [16] CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7 \text{ TeV}$ ", JINST 7 (2012) P10002, arXiv:1206.4071. doi:10.1088/1748-0221/7/10/P10002.
- [17] CMS Collaboration, "Measurements of Inclusive W and Z Cross Sections in pp Collisions at $\sqrt{s} = 7 \text{ TeV}$ ", JHEP 2 (2011) 2-40, arXiv:hep-ph/10122466. doi:10.1007/JHEP01(2011)080.
- [18] M. Cacciari, G.P. Salam, and G. Soyez, The anti-kt jet clustering algorithm, JHEP 0804, (2008) 063.
- ²²⁷ [19] M. Cacciari and G.P. Salam,, "Pileup subtraction using jet areas (hep-ph/0707.1378)", ²²⁸ Physics Letters **B** 659 (2008) 119–126. doi:10.1016/j.physletb.2007.09.077.
- [20] M. Cacciari, G.P. Salam and G. Soyez, "The catchment area of jets (hep-ph/0802.1188)", JHEP 0804 (2008) 005. doi:10.1088/1126-6708/2008/04/005.
- ²³¹ [21] CMS Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS", JINST **6** (2011) P11002.

 doi:10.1088/1748-0221/6/11/P11002.
- [22] C. Weiser, "A Combined Secondary Vertex Based B-Tagging Algorithm in CMS", CMS
 NOTE 014 (2006) 17.
- ²³⁶ [23] CMS Collaboration, "Commissioning of b-jet identification with pp collisions at sqrt(s) = 7 TeV", CMS PAS BTV **10-001** (2010).
- Proceedings of the workshop: *HERA and the LHC workshop series on the implications of HERA for LHC physics* , **arXiv:0903.3861**.

[25] J. M. Campbell and R. Ellis, "MCFM for the Tevatron and the LHC", Nucl. Phys. Proc.
 Suppl. 205 (2010) 10, arXiv:1007.3492.
 doi:10.1016/j.nuclphysbps.2010.08.011.

[26] S. Badger, J. M. Campbell, and R. Ellis, "QCD corrections to the hadronic production of a heavy quark pair and a W-boson including decay correlations", *JHEP* **03** (2011) 027, arXiv:1011.6647. doi:10.1007/JHEP03 (2011) 027.

