Measurement of the ratio of inclusive cross sections
\[ \sigma(p\bar{p} \to Z + 2 b\text{jets})/\sigma(p\bar{p} \to Z + 2 \text{jets}) \] in \( p\bar{p} \) collisions at \( \sqrt{s} = 1.96 \) TeV


2015 © American Physical Society

PHYSICAL REVIEW D 91, 052010 (2015)

1 LAFEX, Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, Brazil
2 Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil
3 Universidade Federal do ABC, Santo André, Brazil
4 University of Science and Technology of China, Hefei, People’s Republic of China
5 Universidad de los Andes, Bogotá, Colombia

(DO Collaboration)

1550-7998/2015/91(5)/052010(9) 052010-1 © 2015 American Physical Society
Charles University, Faculty of Mathematics and Physics, Center for Particle Physics, Prague, Czech Republic

Czech Technical University in Prague, Prague, Czech Republic

Institute of Physics, Academy of Sciences of the Czech Republic, Prague, Czech Republic

Universidad San Francisco de Quito, Quito, Ecuador

LPC, Université Blaise Pascal, CNRS/IN2P3, Clermont, France

LPSC, Université Joseph Fourier Grenoble 1, CNRS/IN2P3, Institut National Polytechnique de Grenoble, Grenoble, France

CPPM, Aix-Marseille Université, CNRS/IN2P3, Marseille, France

LAL, Université Paris-Sud, CNRS/IN2P3, Orsay, France

LPNHE, Universités Paris VI and VII, CNRS/IN2P3, Paris, France

CEA, Irfu, SPP, Saclay, France

IPHC, Université de Strasbourg, CNRS/IN2P3, Strasbourg, France

IPNL, Université Lyon 1, CNRS/IN2P3, Villeurbanne, France and Université de Lyon, Lyon, France

III. Physikalisches Institut A, RWTH Aachen University, Aachen, Germany

Physikalisches Institut, Universität Freiburg, Freiburg, Germany

Institut für Physik, Universität Mainz, Mainz, Germany

Ludwig-Maximilians-Universität München, München, Germany

Panjab University, Chandigarh, India

Delhi University, Delhi, India

Tata Institute of Fundamental Research, Mumbai, India

University College Dublin, Dublin, Ireland

Korea Detector Laboratory, Korea University, Seoul, Korea

CINVESTAV, Mexico City, Mexico

Nikhef, Science Park, Amsterdam, The Netherlands

Radboud University Nijmegen, Nijmegen, The Netherlands

Joint Institute for Nuclear Research, Dubna, Russia

Institute for Theoretical and Experimental Physics, Moscow, Russia

Moscow State University, Moscow, Russia

Institute for High Energy Physics, Protvino, Russia

Petersburg Nuclear Physics Institute, St. Petersburg, Russia

Institució Catalana de Recerca i Estudis Avançats (ICREA) and Institut de Física d’Altes Energies (IFAE), Barcelona, Spain

Uppsala University, Uppsala, Sweden

Taras Shevchenko National University of Kyiv, Kiev, Ukraine

Imperial College London, London SW7 2AZ, United Kingdom

The University of Manchester, Manchester M13 9PL, United Kingdom

University of Arizona, Tucson, Arizona 85721, USA

University of California Riverside, Riverside, California 92521, USA

University of Mississippi, University, Mississippi 38677, USA

University of Nebraska, Lincoln, Nebraska 68588, USA

Rutgers University, Piscataway, New Jersey 08855, USA

Princeton University, Princeton, New Jersey 08544, USA

State University of New York, Buffalo, New York 14260, USA
Studies of Z boson production in association with a bottom and an antibottom quark provide important tests of the predictions of perturbative quantum chromodynamics (pQCD) [1,2]. A good theoretical description of this process is essential since it forms a major background for a variety of physics processes, including standard model (SM) Higgs boson production in association with a Z boson, $ZH(H \rightarrow b\bar{b})$ [3], and searches for supersymmetric partners of the $b$ quark [4].

The ratio of $Z + b$ jet to $Z +$ jet production cross sections, for events with at least one jet, has been previously measured by the CDF [5,6] and D0 [7–9] Collaborations using Run II data. The ATLAS [10] and CMS [11] Collaborations have also studied $Z + b$ jet production at $\sqrt{s} = 7$ TeV.

This article presents the ratio of $Z + 2b$ jets to $Z + 2$ jets inclusive production cross sections and is an extension of the previous D0 measurements utilizing similar event selections. The measurement of the ratio benefits from the cancellation of many systematic uncertainties, such as the uncertainty in luminosity and those related to lepton and jet identification, allowing a more precise comparison with theory. The remaining systematic uncertainties arise from the differences between $b$ jets and light jets. In the following, light-quark flavor ($u$, $d$, $s$) and gluon jets are referred to as “light jets.” The $Z + 2b$ jet production cross sections have been measured at CMS [12] and ATLAS [13] at $\sqrt{s} = 7$ TeV. The current measurement is based on the complete Run II data sample collected by the D0 experiment at the Fermilab Tevatron $p\bar{p}$ collider at a center-of-mass energy of $1.96$ TeV, and corresponds to an integrated luminosity of $9.7$ fb$^{-1}$.

We first briefly describe the main components of the D0 Run II detector [14,15] relevant to this analysis. The D0 detector has a central tracking system consisting of a silicon microstrip tracker (SMT) [16] and a central fiber tracker (CFT), both located within a 1.9 T superconducting solenoidal magnet, with designs optimized for tracking and vertexing at pseudorapidities $|\eta_{det}| < 3$ and $|\eta_{det}| < 2.5$, respectively [17]. A liquid argon and uranium calorimeter
has a central section (CC) covering pseudorapidities \(|\eta_{\text{det}}| \lesssim 1.1\), and two end calorimeters (EC) that extend coverage to \(|\eta_{\text{det}}| \approx 4.2\), with all three housed in separate cryostats [18]. An outer muon system, at \(|\eta_{\text{det}}| < 2\), consists of a layer of tracking detectors and scintillation counters in front of 1.8 T toroids, followed by two similar layers after the toroids. Luminosity is measured using plastic scintillator arrays located in front of the EC cryostats. The trigger and data acquisition systems are designed to accommodate the high instantaneous luminosities of Run II.

This analysis relies on all components of the D0 detector: tracking systems, the liquid-argon sampling calorimeter, muon system, and the ability to identify secondary vertices [14]. The SMT allows for precise reconstruction of the primary \(p\bar{p}\) interaction vertex and secondary vertices [17,19]. It also enables an accurate determination of the impact parameter, defined as the distance of closest approach of a track to the primary interaction vertex in the \(x-y\) plane. The impact parameter measurements of tracks, along with reconstructed secondary vertices, are important inputs to the \(b\)-jet tagging algorithm.

Events containing \(Z\) bosons decaying to \(\mu\mu\) or \(ee\) are collected using triggers based on single electrons or muons. For the off-line selection requirements discussed below, the triggers have an efficiency of approximately 100% for \(Z \rightarrow ee\) and more than 78% for \(Z \rightarrow \mu\mu\) decays depending on the transverse momentum of the muon. The \(Z + 2\) jet sample requires the presence of at least two jets in the event, while the \(Z + 2b\) jet sample requires at least two \(b\)-jet candidates, selected using a \(b\)-tagging algorithm [20].

An event is selected if it contains a \(p\bar{p}\) interaction vertex, reconstructed from at least three tracks, located within 60 cm of the center of the D0 detector along the beam axis. The selected events must also contain a \(Z\) boson candidate with a dilepton invariant mass \(70 < M_{\ell\ell} < 110\) GeV.

Dielectron (ee) events are required to have two electrons of transverse momentum \((p_T)\) greater than 15 GeV identified through electromagnetic (EM) showers in the calorimeter. The showers must have more than 97% of their energy deposited in the EM calorimeter, be isolated from other energy depositions, and have transverse and longitudinal energy profiles consistent with that expected for electrons. At least one electron must be identified in the CC, with \(|\eta_{\text{det}}| < 1.1\), and a second electron either in the CC or the EC, \(1.5 < |\eta_{\text{det}}| < 2.5\). Electron candidates in the CC are required to match central tracks or have a pattern of hits consistent with the passage of an electron through the central tracker. Electrons in the ECs are not required to have a track matched to them due to deteriorating tracking coverage for \(|\eta_{\text{det}}| > 2\). Due to the lack of track requirement in EC regions we do not apply any opposite sign requirement for the dielectron events.

The dimuon (\(\mu\mu\)) event selection requires two oppositely charged muons detected in the muon system that are matched to reconstructed tracks in the central tracker with \(p_T > 15\) GeV and \(|\eta_{\text{det}}| < 2\). These muons must pass a combined tracking and calorimeter isolation requirement discussed in detail in Ref. [3]. Muons originating from cosmic rays are rejected by applying timing criteria using the hits in the scintillation counters and by limiting the measured displacement of the muon track with respect to the \(p\bar{p}\) interaction vertex [21].

A total of about 1.2 million \(Z\) boson candidate events are retained in the combined \(ee\) and \(\mu\mu\) channels with the above lepton selection criteria. The \(Z + 2\) jet sample is then selected by requiring at least two jets in the event with \(p_T^{\text{jet}} > 20\) GeV and \(|\eta_{\text{jet}}| < 2.5\). Jets are reconstructed from energy deposits in the calorimeter using an iterative midpoint cone algorithm [22] with a cone of radius \(\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta y)^2} = 0.5\) where \(\phi\) is the azimuthal angle and \(y\) is the rapidity. Jet energy is corrected for detector response, the presence of noise and multiple \(p\bar{p}\) interactions. We also correct the jet energy for the energy of those particles within the reconstruction cone that is deposited in the calorimeter outside the cone (and vice versa) [23].

To suppress background from top-antitop quark (\(t\bar{t}\)) production, events are rejected if the missing transverse energy is larger than 60 GeV, reducing the \(t\bar{t}\) contamination by a factor of two. These selection criteria retain an inclusive sample of 20,950 \(Z + 2\) jet event candidates in the combined \(ee\) and \(\mu\mu\) channels.

Processes such as diboson (WW, WZ, ZZ) production can contribute to the background when two leptons are reconstructed in the final state. Inclusive diboson production is simulated with the PYTHIA [24] Monte Carlo (MC) event generator. The \(Z + \text{jet}\), including heavy flavor jets, and \(t\bar{t}\) events are modeled by ALPGEN [25], which generates hard subprocesses including higher-order QCD tree-level matrix elements, interfaced with PYTHIA for parton showering and hadronization. The CTEQ6L1 [26] parton distribution functions (PDFs) are used in all simulations. The cross sections of the simulated samples are then scaled to the corresponding higher-order theoretical calculations. For the diboson and \(Z + 2\) jet processes, including the \(Z + b\bar{b}\) signal process and \(Z + c\bar{c}\) production, next-to-leading order (NLO) cross section predictions are taken from MCFM [27]. The \(t\bar{t}\) cross section is determined from NLO + NNLL (next-to-next-leading log) calculations [28]. To improve the modeling of the \(p_T\) distribution of the \(Z\) boson, simulated \(Z + 2\) jet events are also reweighted to be consistent with the measured \(p_T\) spectrum of \(Z\) bosons observed in data [29].

These generated samples are processed through a detailed detector simulation based on GEANT [30]. To model the effects of detector noise and pile-up events, collider data from random beam crossings with the same instantaneous luminosity distribution as for data are superimposed on simulated events. These events are then
reconstructed using the same algorithms as used for data. Scale factors, determined from data using independent samples, are applied to account for differences in reconstruction efficiency between data and simulation. The energies of simulated jets are corrected, based on their flavor, to reproduce the resolution and energy scale observed in data [23].

The background contribution from multijet events, in which jets are misidentified as leptons, is evaluated from data. This is performed using a multijet-enriched sample of events that pass all selection criteria except for some of the lepton quality requirements. In the case of electrons, the multijet sample is obtained by inverting the shower shape requirement and relaxing other electron identification criteria, while for the muon channel, the multijet sample consists of events with muon candidates that fail the isolation requirements. The normalization of the multijet background is determined from a simultaneous fit to the dilepton invariant mass distributions in different jet multiplicity bins.

Figures 1 and 2 show the dilepton invariant mass and leading jet $p_T$ distributions in data compared to the expectations from various processes. The dominant contribution comes from $Z +$ light jet production. The non-$Z +$ jet background fraction in the $ee$ channel is about 15%, and is dominated by multijet production. The muon channel has a higher purity with a background fraction of about 7%.

This analysis employs a two-step procedure to determine the $b$-quark content of jets in the selected events. First, a $b$-tagging algorithm is applied to jets to select a sample of $Z + 2$ jet events that is enriched in heavy flavor jets. After $b$ tagging, the relative light-, $c$-, and $b$-quark content is extracted by fitting templates built from a dedicated

**FIG. 1 (color online).** The invariant mass in (a) $Z \rightarrow \mu\mu$ and (b) $Z \rightarrow ee$ channels for data and background in events with a $Z$ boson candidate and at least two jets before $b$ tagging is applied.

**FIG. 2 (color online).** The leading jet $p_T$ in the (a) $Z \rightarrow \mu\mu$ and (b) $Z \rightarrow ee$ channels for data and background in events with a $Z$ boson candidate and at least two jets before $b$ tagging is applied.
discriminant that provides an optimized separation between the three components.

Jets considered for $b$-jet tagging are subject to a preselection requirement, called taggability, to decouple the intrinsic performance of the $b$-jet tagging algorithm from effects related to the track reconstruction efficiency. For this purpose, the jet is required to have at least two associated tracks with $p_T > 0.5$ GeV, the leading track must have $p_T > 1$ GeV, and each track must have at least one SMT hit. This requirement has a typical efficiency of 90% per jet.

The $b$-jet tagging algorithm is based on a multivariate analysis (MVA) technique [31]. This algorithm, MVA$_{bl}$, discriminates $b$ jets from light flavor jets utilizing the relatively long lifetime of the $b$ hadrons when compared to their lighter counterparts [20]. Events with at least two jets tagged by this algorithm are considered.

The MVA$_{bl}$ discriminant combines various properties of the jet and associated tracks to create a continuous output that tends towards unity for $b$ jets and zero for light jets. Inputs include the number of secondary vertices and the charge track multiplicity, invariant mass of the secondary vertex ($M_{SV}$), decay length and impact parameter of secondary vertices, the multiplicity of charged tracks associated with them, and the jet lifetime probability (JLIP), which is the probability that tracks associated with the jet originate from the $p \bar{p}$ interaction vertex [20]. Events are retained for further analysis if they contain at least two jets with an MVA$_{bl}$ output greater than 0.15. After these requirements, 241 $Z + 2$ jet events are selected with at least two $b$-tagged jets, where only the two highest $p_T$ tagged jets are examined in the analysis and the electron and muon channels are combined. The efficiency for tagging two $b$ jets in data is 33%. In the MC correction, factors are applied to account for differences with data [20]. The background contamination from diboson, multijet, and top production after $b$-tagging, for the electron and muon channels combined are 8%, 2% and 15%, respectively.

To determine the fraction of events with 2$b$ jets, a dedicated discriminant, $D_{ML}$, is employed [8,32]. It is a combination of the two most discriminating MVA$_{bl}$ inputs, $M_{SV}$ and JLIP: $D_{ML} = 0.5 \times (M_{SV}/5 \text{ GeV}) - \ln(\text{JLIP})/20$.

To measure the fraction of events with different jet flavors in the selected sample, we count the number of events as a function of the $D_{ML}$ of the two leading jets $N(D_{ML1}, D_{ML2})$ and then perform a two-dimensional binned maximum likelihood fit to that distribution. The data sample with two heavy flavor tagged jets is fitted to templates consisting mainly of 2$b$ jet, 2$c$ jet, and light flavor jet events, as obtained from ALPGEN+PYTHIA simulated samples. We also compared the shapes of the templates from SHERPA simulated samples and found the templates to be consistent for the two models. Before the fit, all non-$Z +$ jet background contributions, estimated from simulated samples after the MVA$_{bl}$ requirement, are subtracted from the data leaving 180 $Z + 2$ jet events in the combined $ee$ and $\mu\mu$ channel. Next, we measure the jet flavor fractions in the dielectron and dimuon samples combined, yielding the $2b$ jet flavor fraction ($f_{bb}$) of $0.64 \pm 0.08(\text{stat})$ and the $2c$ jet flavor fraction of $0.32 \pm 0.08(\text{stat})$. Figure 3 shows the one-dimensional projection onto the highest-$p_T$ jet and the second-highest-$p_T$ jet $D_{ML}$ axis of the two-dimensional fit.

The fraction of 2$b$ jets measured in the heavy flavor enriched sample is combined with the corresponding event acceptances to determine the ratio, $R$, of the cross sections,

$$R = \frac{\sigma(Z + 2b \text{ jets})}{\sigma(Z + 2 \text{ jets})} = \frac{N_{bb}f_{bb}}{N_{incl}\epsilon_{tag}} \times \frac{A_{incl}}{A_{bb}},$$

FIG. 3 (color online). The one dimensional projection onto (a) the highest-$p_T$ jet and (b) the second highest-$p_T$ jet $D_{ML}$ axis of the two-dimensional fit. The distributions of the $b$, $c$, and light jets are normalized by the fractions found from the fit.
Table I. The ratio of integrated cross sections, \( \sigma(p \bar{p} \rightarrow Z + 2b \text{jet})/\sigma(p \bar{p} \rightarrow Z + 2 \text{jet}) \) and total systematic uncertainties (\( \delta_{\text{syst}} \)). The column \( \delta_{\text{stat}} \) shows the total statistical uncertainty obtained by adding \( \delta_{\text{stat}} \) and \( \delta_{\text{syst}} \) in quadrature. The last three columns show theoretical predictions obtained using NLO QCD with scale uncertainties and two MC event generators, PYTHIA and ALPGEN.

\[
\begin{array}{cccc}
\sigma(p \bar{p} \rightarrow Z + 2b \text{jet})/\sigma(p \bar{p} \rightarrow Z + 2 \text{jet}) & \delta_{\text{stat}} & \delta_{\text{syst}} & \delta_{\text{tot}} \\
\text{Data} & (2.36 \pm 0.32 \pm 0.35) \times 10^{-2} & 0.47 \times 10^{-2} \\
\text{NLO QCD(MSTW)} & (1.76 \pm 0.26) \times 10^{-2} & 2.42 \times 10^{-2} \\
\text{PYTHIA} & & 2.21 \times 10^{-2} \\
\text{ALPGEN} & & & \\
\end{array}
\]
in a restricted phase space of leptons with $p_T^l > 15$ GeV, $|\eta^l| < 2.0$ and with two jets limited to $p_T^{jets} > 20$ GeV and $|\eta^{jets}| < 2.5$. Measurements are based on the full data sample collected by the D0 experiment in Run II of the Tevatron, corresponding to an integrated luminosity of 9.7 fb$^{-1}$ at a center-of-mass energy of 1.96 TeV. The measured integrated ratio of $0.0236 \pm 0.0032$ (stat) $\pm 0.0035$ (syst) is in agreement with the theoretical predictions within uncertainties.

We thank John Campbell and Doreen Wackeroth for valuable discussions, and the staffs at Fermilab and collaborating institutions, and acknowledge support from the Department of Energy and National Science Foundation (United States of America); Alternative Energies and Atomic Energy Commission and National Center for Scientific Research/National Institute of Nuclear and Particle Physics (France); Ministry of Education and Science of the Russian Federation, National Research Center “Kurchatov Institute” of the Russian Federation, and Russian Foundation for Basic Research (Russia); National Council for the Development of Science and Technology and Carlos Chagas Filho Foundation for the Support of Research in the State of Rio de Janeiro (Brazil); Department of Atomic Energy and Department of Science and Technology (India); Administrative Department of Science, Technology and Innovation (Colombia); National Council of Science and Technology (Mexico); National Research Foundation of Korea (Korea); Foundation for Fundamental Research on Matter (The Netherlands); Science and Technology Facilities Council and The Royal Society (United Kingdom); Ministry of Education, Youth and Sports (Czech Republic); Bundesministerium für Bildung und Forschung (Federal Ministry of Education and Research) and Deutsche Forschungsgemeinschaft (German Research Foundation) (Germany); Science Foundation Ireland (Ireland); Swedish Research Council (Sweden); China Academy of Sciences and National Natural Science Foundation of China (China); and Ministry of Education and Science of Ukraine (Ukraine).

[17] We use a standard right-handed coordinate system. The nominal collision point is the center of the detector with coordinates (0, 0, 0). The direction of the proton beam is the positive +z axis. The +x axis is horizontal, pointing away from the center of the Tevatron ring. The +y axis points vertically upwards. The polar angle, $\theta$, is defined such that $\theta = 0$ is the +z direction. The rapidity is defined as $y = -\ln [(E + p_z)/(E - p_z)]$, where $E$ is the energy and $p_z$ is the momentum component along the proton beam direction. Pseudorapidity is defined as $\eta = -\ln (\tan \theta/2)$. $\eta$ is defined as the azimuthal angle in the plane transverse to the proton beam direction. Also, $\eta_{el}$ and $\eta_{jet}$ are the pseudorapidity and the azimuthal angle measured with respect to the center of the detector.
[19] The primary $p\bar{p}$ interaction vertex is that found to be the most likely collision point, among possibly several collisions within a specific beam crossing, from which selected objects emanate. The algorithm for defining it can be found in [20].