Measurements of jet observables sensitive to b-quark fragmentation in $t\bar{t}$ events at the LHC with the ATLAS detector

Aad, G.; ATLAS Collaboration

DOI
10.1103/PhysRevD.106.032008

Publication date
2022

Document Version
Final published version

Published in
Physical Review D. Particles, Fields, Gravitation, and Cosmology

License
CC BY

Citation for published version (APA):
https://doi.org/10.1103/PhysRevD.106.032008

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)
Measurements of jet observables sensitive to $b$-quark fragmentation in $t\bar{t}$ events at the LHC with the ATLAS detector

G. Aad et al.∗
(ATLAS Collaboration)

(Received 1 March 2022; accepted 8 July 2022; published 9 August 2022)

Several observables sensitive to the fragmentation of $b$ quarks into $b$ hadrons are measured using 36 fb$^{-1}$ of $\sqrt{s} = 13$ TeV proton-proton collision data collected with the ATLAS detector at the LHC. Jets containing $b$ hadrons are obtained from a sample of dileptonic $t\bar{t}$ events, and the associated set of charged-particle tracks is separated into those from the primary $pp$ interaction vertex and those from the displaced $b$-decay secondary vertex. This division is used to construct observables that characterize the longitudinal and transverse momentum distributions of the $b$ hadron within the jet. The measurements have been corrected for detector effects and provide a test of heavy-quark-fragmentation modeling at the LHC in a system where the top-quark decay products are color connected to the proton beam remnants. The unfolded distributions are compared with the predictions of several modern Monte Carlo parton-shower generators and generator tunes, and a wide range of agreement with the data is observed, with $p$ values varying from $5 \times 10^{-4}$ to 0.98. These measurements complement similar measurements from $e^+e^-$ collider experiments in which the $b$ quarks originate from a color singlet $Z/\gamma'$.

DOI: 10.1103/PhysRevD.106.032008

I. INTRODUCTION

The dominant paradigm for calculations of hadron collider physics observables makes use of quantum chromodynamics (QCD) factorization to separate the perturbative hard-scattering process from nonperturbative initial- and final-state effects. Historically, these effects were separated into three disjoint parts: a nonperturbative structure function representing the extraction of incoming parton configurations from the colliding hadrons, a perturbative hard process, and a nonperturbative fragmentation function mapping the outgoing hard-process partons to final-state hadrons and jets. Much of early QCD physics research was conducted within this paradigm [1,2].

This picture has been continually refined in the quest for increased precision and exclusive, realistic, Monte Carlo (MC) event modeling. In particular, modeling of both the initial and final states is now typically enhanced by perturbative analytic resummation or parton-shower algorithms, which go beyond fixed order in the QCD coupling to account for the enhanced effects of multiple QCD parton emissions in singular regions of emission phase space [3]. These can be considered enhancements of the hard process, which were previously absorbed—resulting in a loss of kinematic accuracy—into the parton distribution and fragmentation functions. The partonic content of the proton, from which the perturbative scattering is initiated, is described by the parton distribution functions (PDFs) [4], while the formation of hadrons below the hadronization scale $\sim \Lambda_{\text{QCD}}$ is described by the fragmentation functions [5,6]. Both are considered universal in hard-scattering processes (by virtue of the QCD factorization theorem) and are currently obtained by fitting nonperturbative low-scale functions to experimental data through their perturbative evolution to the hard-scattering scales.

The fragmentation of $b$ quarks is particularly interesting for several reasons. Bottom quarks play an important role in Large Hadron Collider (LHC) analyses, from top quark and Higgs boson measurements to new-particle searches. This makes it essential that the evolution and hadronization of heavy quarks be well understood. As a heavy quark whose mass plays a role in automatically regularizing the soft divergence in QCD splitting functions, the $b$ quarks’ evolution and hadronization also provide a fundamental test of perturbative QCD and the parton-shower formalism [3,7–11]. Finally, displaced weak-decay vertices make possible an unambiguous experimental connection between observed hadrons and their ancestor parton. Current MC event generator models [12,13] and their tuned parameters (tunes) [14–16] are based on $b$-fragmentation observables measured at $e^+e^-$ colliders such as LEP [17–19] and SLC [20,21], which are experimentally and theoretically straightforward to interpret due to the absence of QCD.

*Full author list given at the end of the article.

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI. Funded by SCOAP3.
initial-state radiation and an underlying event. However, the combination of perturbative and nonperturbative effects in fragmentation means that the more complex color flow in hadron-collider processes will affect observables related to fragmentation \[22,23\]; its measurement at the LHC complements those previously carried out at lepton colliders.

In this analysis, a measurement of observables sensitive to the fragmentation of $b$ quarks at the LHC is presented, using $t\bar{t}$ events observed in the ATLAS detector in the 2015 and 2016 data-taking runs. Choosing this process allows studies of $b$-jet fragmentation observables in which some fraction of the $b$ quarks are color connected to the initial state. The more complex QCD environment of a hadron collider may also affect fragmentation via effects such as QCD initial-state radiation, multiple partonic interactions, and color reconnection.

The observables used in this paper characterize the evolution of the $b$-quark momentum into the lower momentum of the weakly decaying $b$ hadron. This is achieved by comparing the momenta of reconstructed $b$ hadrons with their corresponding $b$ jets and with the event-wide momentum scale set by other top-quark decay products. While there is an immediate theoretical objection to making a one-to-one identification between quarks and jets—a colored quark cannot in isolation hadronize into a color-neutral jet—observables inspired by this simple picture can be well defined, and receive finite and calculable corrections from the initiator’s color connections \[24\]. In the $e^+e^-$ experiments where such observables were previously measured, incoming beam particles have a well-defined momentum, and hence there is no question as to the energy of the resulting $b$ quarks in a two-jet $e^+e^- \rightarrow bb$ process: at leading order, each must have the same energy as the incoming electron or positron beam. The fragmentation can hence be probed via the simple function $x_B = E_B/E_{beam}$—the energy of the $b$ hadron divided by that of the beam. Due to the nature of hadronic initial states, the partonic center-of-mass or energy scale is not known a priori for any particular collision event. For this reason, observables are defined using only the final-state particles rather than the beam energy.

It is experimentally challenging to determine whether electrically neutral particles originated from heavy hadron decays, the fragmentation products of the jet, the underlying event, or pileup interactions; as such, this study focuses on the charged constituents of the jets. Charged-particle tracks have better momentum and angular resolution than calorimeter clusters at the energies involved in this analysis and allow for the removal of contamination particles emitted from pileup interactions. To reconstruct these $b$ hadrons, inner-detector tracks are used to build secondary vertices displaced from the hard-scatter collision \[25\]. The constituents of secondary vertices correspond to the stable, charged $b$-hadron descendants. The $b$ hadron’s “charged momentum” is then calculated as the vectorial sum of these descendants. The $b$-hadron charged momentum is compared with various quantities that provide an energy scale for the surrounding environment.

One choice is to relate the $b$ hadron with the surrounding jet of particles, constructed via standard clustering techniques \[26\]. This is natural since the aim of all jet algorithms is to cluster flows of energy corresponding to well-separated initial partons in a theoretically safe way: a jet containing a $b$ hadron is a reasonably well-defined proxy for its initiating $b$ quark. Observables based on the ratio of $b$-hadron to $b$-jet momenta are by their nature most sensitive to the small-angle gluon emissions from the $b$ quark before hadronization, and therefore provide a view of $b$-quark fragmentation that is strongly complementary to that of the lepton-collider measurements, which rely on event-wide—rather than local—hadronic energy flow. Since the $b$ hadrons are reconstructed and identified through secondary vertices built from charged-particle tracks, the jet momentum is built from reconstructible charged constituents, putting the $b$ hadron and the jet on the same footing and allowing the $b$ hadron’s energy to be distinguished from the energy of the nearby hadronic activity. Two observables are measured in this vein: (1) the ratio of the $b$ hadron and jet charged momenta transverse to the incoming proton beams ($z_{T,b}$) and (2) the fraction of the jet charged momentum carried by the $b$-hadron decay products along the direction of the jet charged momentum ($z_{L,b}$). The analysis object-selection methodology is therefore focused on the identification of charged particles from primary and secondary vertices.

A comparison between the kinematic properties of the $b$ hadron and the $t\bar{t}$ system is also made. This measurement is performed in $\mu\nu\ell\ell bb$ events, and, while the primary role of the electron and muon is to select a pure sample of $t\bar{t}$ events, the lepton momenta are correlated with the momenta of the top quarks from which they were produced. Although the leptons are far from being a perfect proxy for the top quarks, their momenta can be determined more precisely than those of the $b$ hadrons and therefore can be used to characterize the $b$ hadrons in the context of the $t\bar{t}$ system. The observable $\rho$ is defined as the ratio of the transverse momentum of the $b$ hadron to the average transverse momentum of the two leptons in the event, where “transverse” is defined relative to the colliding proton beams. As opposed to $z_{T,b}$ and $z_{L,b}$, $\rho$ is sensitive to radiation emitted in the top-quark decay, regardless of whether or not it was contained in the $b$ jet.

In addition to the above observables that focus on the $b$ hadron momentum, the number of stable, charged decay products of the $b$ hadron ($n_{prod}^{b}$) is measured. This observable is sensitive to the modeling of $b$-hadron production and decay.
For independence from detector biases, and for maximum impact on the development of both analytic and MC generator b-quark fragmentation modeling, the results have been corrected to a fiducial acceptance definition at stable-particle level. Particles are considered stable if they have a mean lifetime $\tau > 33$ ps ($\tau > 10$ mm). The unfolded data are then compared with predictions from several commonly used MC models and tunes, which are largely based on $e^+e^-$ $x_g$ distributions.

The organization of this paper is as follows. Section II gives an overview of the ATLAS detector, and Sec. III outlines the collision data and MC simulation used in this measurement. Section IV describes the object and event selection used in the experimental analysis as well as the fiducial definition to which the observed data are unfolded. The unfolding procedure and systematic uncertainties are presented in Secs. V and VI, respectively. Finally, the unfolded data are shown in Sec. VII and compared with a number of modern MC generators tuned to $e^+e^-$ collider data.

II. ATLAS DETECTOR

The ATLAS detector [27–29] is a multipurpose particle detector with a forward/backward-symmetric cylindrical geometry. The detector has a nearly $4\pi$ coverage in solid angle\(^1\) and consists of an inner tracking detector, electromagnetic and hadronic calorimeters, and a muon spectrometer. The inner detector is surrounded by a superconducting solenoid providing a 2T magnetic field and covers a pseudorapidity range of $|\eta| < 2.5$. The inner detector is composed of silicon pixel and microstrip detectors as well as a transition radiation tracker. The innermost pixel detector layer, called insertable B-layer (IBL) was installed before the start of run II. The high-granularity lead/liquid-argon (LAr) electromagnetic sampling calorimeters measure electromagnetic energies in the pseudorapidity region $|\eta| < 3.2$. Hadron energies are measured by a hadronic, steel/scintillator tile calorimeter with $|\eta| < 1.7$. The forward and end cap regions with $1.5 < |\eta| < 4.9$ are instrumented with LAr calorimeters for both the electromagnetic and hadronic measurements. Surrounding the calorimeters, the muon spectrometer consists of three large superconducting toroids with eight coils each. The muon spectrometer has a system of precision tracking chambers covering $|\eta| < 2.7$, consisting of monitored drift tubes and, in the forward region, cathode-strip chambers. In addition, it has fast trigger chambers covering $|\eta| < 2.4$, consisting of resistive-plate chambers in the barrel region and thin-gap chambers in the end caps. A two-level trigger system is used to select the events that are recorded [30]. The first-level trigger (L1) is implemented in hardware and utilizes partial detector information to accept events at a rate below 100 kHz from the 40 MHz bunch crossings. The high-level trigger is software-based and accepts events at a rate of 1 kHz.

III. DATA SAMPLE AND SIMULATION

The measurements in this analysis are based on data collected by the ATLAS detector in 2015 and 2016 at a center-of-mass energy of $\sqrt{s} = 13$ TeV. The recorded data correspond to an integrated luminosity of approximately 36.1 fb\(^{-1}\). Collision events are analyzed only if they satisfy the data quality criteria [31], the beam conditions were stable when they were recorded, and all subdetectors passed the requirements for operational status. An extensive software suite [32] is used in the reconstruction and analysis of real and simulated data, in detector operations, and in the trigger and data acquisition systems of the experiment.

This analysis requires at least one lepton to be identified in the trigger system for each candidate event. Depending on the data-taking period, different single-lepton trigger thresholds were used for electrons and muons [33,34] in order to maintain readout rates within specification over changes to the LHC instantaneous luminosity. For 2015 data, single-electron triggers with $p_T$ thresholds of 24, 60, and 120 GeV and single-muon triggers with $p_T$ thresholds of 20 and 50 GeV were used. For 2016 data, the $p_T$ thresholds were increased to 26, 60, and 140 GeV for electrons and 26 and 50 GeV for muons. The triggers with the with the lowest $p_T$ thresholds include isolation requirements, while for the triggers with higher thresholds those requirements are relaxed to increase the acceptance.

Several simulated event samples are used in estimating the Standard Model prediction for this measurement, particularly in obtaining the corrections for detector effects discussed in Sec. V. The $t\bar{t}$ signal process is simulated at next-to-leading order (NLO) in perturbative QCD using the HVQ program [35,36] implemented in the POWHEG BOX v2 [37,38] event generator with the NNPDF3.0 PDF sets [39]. Parton-level configurations obtained with POWHEG BOX are processed with the PYTHIA 8.230 [40] generator to model the parton shower (PS), hadronization and underlying event, using the A14 set of tuned parameters [16] and the NNPDF2.3 PDF set. The A14 tune of PYTHIA 8 is a tune of PS and multiple partonic interaction parameters that leaves the hadronization parameters at their default values and uses the Lund-Bowler fragmentation model [12]. The top-quark mass, $m_{\text{top}}$ is set to 172.5 GeV in all simulated
samples. The $h_{\text{damp}}$ parameter, which controls the transverse momentum of the first additional gluon emission beyond the Born configuration, is set equal to 1.5 times the mass of the top quark. The main effect of this choice is to regulate the high-$p_T$ emission against which the $t\bar{t}$ system recoils. The factorization and renormalization scales are set to $\sqrt{m_{\text{top}}^2 + p_T^2}$, where $p_T$ is the transverse momentum of the top quark.

The $t\bar{t}$ cross-section is normalized to $\sigma_{t\bar{t}} = 832_{-35}^{+46}$ pb, as calculated at next-to-next-to-leading order (NNLO) with next-to-next-to-leading logarithmic soft-gluon terms [41–45] using the TOP++ 2.0 program [46]. The PDF- and $\alpha_s$-induced uncertainties in this cross section are calculated using the PDF4LHC prescription [47] with the MSTW2008NNLO 68% C.L. PDF [48,49], CT10NNLO PDF [50,51], and NNPDF2.3 5f FEN PDF [52] and are added in quadrature with the uncertainties obtained from the independent variation of the factorization and renormalization scales.

In order to evaluate the signal modeling uncertainties, alternative $t\bar{t}$ samples were produced. To evaluate the impact of initial-state radiation, two additional POWHEG+PYTHIA 8 samples were generated. In one sample, the factorization and renormalization scales in the matrix element are scaled down by a factor of 2, while at the same time setting $h_{\text{damp}} = 3m_{\text{top}}$ and using the VARI3c up variation of the A14 tune. The other sample uses factorization and renormalization scales increased by a factor of 2, while keeping $h_{\text{damp}}$ at the nominal value of 1.5$m_{\text{top}}$ and using the VARI3c down variation. The VARI3c tune variations correspond to a variation of $\alpha_s$ in the initial state. To account for final-state radiation (FSR) uncertainties, two POWHEG+PYTHIA 8 samples with the VARI2 up and down variations of the A14 tune are used. For VARI2, the values of both $\alpha_s$ in the final state and parameters sensitive to initial-state radiation are varied. A sample with POWHEG BOX interfaced with HERWIG 7.04 [53,54] with the H7UE tune [54] and the MMHT2014NNLO PDF [55] PDF set is used to probe the impact of using a different parton shower and hadronization model.

Single-top-quark production in the $tW$ channel is simulated using the POWHEG BOX v1 event generator with the CT10 PDF sets. It is interfaced with the PYTHIA 6.428 generator [56] to model the parton shower and hadronization, using the Perugia2012 set of tuned parameters [57] and the CTEQ6L1 PDF sets [58]. The total cross section for $tW$ production is calculated at NLO with next-to-next-to-leading logarithm (NNLL) soft-gluon corrections [59]. The interference between $t\bar{t}$ and $tW$ production leads to an ambiguity in the definition of these processes starting at NLO. This ambiguity is removed from the $tW$ sample by via the diagram-removal scheme [60]. An alternative $tW$ sample with the diagram-subtraction scheme is also used.

The $Z/\gamma + \text{jets}$ process is simulated using SHERPA 2.2.1 [61] with the NNPDF3.0 PDF set. The matrix elements are calculated using COMIX [62] and OPENLOOPS [63], for up to two parton at NLO and up to four partons at leading order (LO) in the QCD coupling. The MEPS@NLO prescription [64] is used to merge the matrix element and the parton shower contributions. The total cross section is normalized to the NNLO calculation [65]. Electroweak diboson processes are simulated using SHERPA 2.1.1 with the CT10 PDF sets. The matrix elements are calculated using COMIX and OPENLOOPS. The matrix elements for the $WW$ and $WZ$ processes are calculated with no additional partons at NLO. Matrix elements for the $ZZ$ process are calculated with up to one additional parton at NLO. The matrix-element calculations for all three diboson processes ($WW$, $WZ$, and $ZZ$) are performed for up to three additional partons at LO.

All nominal simulated samples are processed through a simulation of the ATLAS detector [66] implemented using the GEANT4 [67] framework. A “fast simulation,” using parametrized showers in the calorimeter-interaction modeling [68], is used for $t\bar{t}$ samples with variations modeling systematic uncertainties. Additional inclusive $pp$ interaction events generated using PYTHIA 8.186 [69] and the A3 set of tuned parameters [70] are overlaid on all simulated signal and background events to simulate the effect of multiple $pp$ interactions in each bunch crossing (pileup). For all samples except the ones generated with SHERPA, the decays of $b$ and $c$ hadrons are modeled using the EVTGEN 1.6.0 program [71].

The same reconstruction algorithms and analysis procedures are applied to both data and MC simulation, allowing direct comparisons of the respective reconstruction-level quantities and thus extraction of detector-effect corrections.

IV. ANALYSIS

To obtain a high-purity sample of reconstructed $b$ jets, this analysis uses dilepton $t\bar{t}$ ($e\mu\nu\nu\nu\nu$) events, which provide a high cross-section source of $b$ jets with relatively low contamination expected from light-flavor jets. The $e\mu$ channel is used exclusively, as it has particularly small contributions from non-$t\bar{t}$ processes. Events with an electron, a muon, and exactly two reconstructed jets are selected.

Biases in reconstruction-level observables introduced by the detector and reconstruction algorithms are corrected to a fiducial particle-level event and object selection, designed to closely match the procedure followed for data events. Systematic uncertainties due to possible discrepancies between the observed and predicted data are taken into account by introducing variations of the predictions for background processes and of the response matrices for the $t\bar{t}$ signal; the considered sources of systematic uncertainty are discussed in Sec. VI. These are propagated through a fully Bayesian unfolding procedure [72] into the final unfolded distributions, as described in Secs. V and VI.
A. Object definitions and selection

Leptons: To reduce the number of background events with nonprompt leptons, tight identification criteria for the reconstructed electrons and medium criteria for the muons are imposed [73–75]. Electrons are additionally required to fall within the fiducial volume of the electromagnetic calorimeter ($|\eta| < 2.47$, excluding $1.37 < |\eta| < 1.52$), and muons must be central ($|\eta| < 2.5$). The invariant mass of the two leptons, $m_{\ell\ell}$, has to be larger than 15 GeV. At least one selected lepton has to be matched to a lepton that fulfilled one of the trigger decisions discussed in Sec. III. Both the electrons and muons must have $p_T > 25$ GeV to avoid large trigger-efficiency uncertainties for leptons with low $p_T$, and be isolated from significant energy deposits in the calorimeters and from high-momentum tracks. For 2016 data, at least one lepton with $p_T > 27$ GeV is required in order to account for the higher trigger threshold. Differences in reconstruction and identification efficiencies between data and simulated events are accounted for by applying scale factors derived in $Z \rightarrow \ell\ell$ events, and the lepton energy and momentum scale and resolutions are calibrated in data by inspecting the dilepton mass spectrum near to the $Z$ peak [73,74].

The particle-level fiducial lepton definition requirements are $p_T > 25$ GeV and $|\eta| < 2.5$, and the leptons must not have final-state hadron ancestors (i.e., they are “prompt”). The four-momenta of photons that do not originate from hadron decays and are within a cone of size $\Delta R = 0.1$ around the lepton direction are added to the lepton four-momentum. This is later referred to as “photon dressing,” and produces “dressed leptons.”

Jets: Detector-level jets are built from topological clusters of energy deposits in calorimeter cells calibrated to the electromagnetic interaction scale, using the anti-$k_T$ algorithm with a radius parameter of $R = 0.4$ [76] implemented in FastJet [77]. The resulting jets’ transverse momenta are further corrected to the corresponding particle-level jet $p_T$ using simulation and in situ methods [78]. After these calibrations, jets with $p_T < 30$ GeV or $|\eta| > 2.5$ are removed. A multivariate discriminant method (JVT) [79] is used to remove jets with $p_T < 60$ GeV and $|\eta| < 2.4$ that have large estimated energy fractions from pileup collision vertices; above this $p_T$ scale, this requirement is unnecessary. Simulated jets are corrected for JVT efficiency differences relative to the collision data.

In order to avoid double-counting of energy deposits from leptons in calorimeter jets, those jets that are reconstructed very near to selected leptons, with $\Delta R_{\ell,jet,lepton} < 0.2$, are removed because they are the result of calorimeter deposits from the leptons themselves, where $\Delta R_{\ell,jet,lepton} = \sqrt{\Delta \phi^2 + \Delta \eta^2}$. On the other hand, leptons that are reconstructed within the range $0.2 < \Delta R_{\ell,jet,lepton} < 0.4$ are known to mostly come from heavy-hadron decays and are not considered prompt-lepton candidates [73–75].

Charged-particle tracks and secondary vertexing:
Reconstructed tracks are limited to the fiducial volume of the inner tracker, $|\eta| < 2.5$, and are required to have $p_T > 500$ MeV. Two track-identification working points are used: tight-primary and tight-secondary. The tight-primary working point is optimized to achieve a high efficiency for tracks originating from the $pp$ collision vertex while rejecting tracks originating from random crossings of charged particles from pileup vertices [80]; tight-primary tracks must be constituents of the reconstructed hard-scatter vertex. In order to achieve a higher efficiency for charged particles from $b$-hadron decays, tight-secondary tracks differ from tight-primary tracks in two main ways: (1) tight-secondary tracks are not required to be constituents of the hard-scatter vertex but must have a transverse impact parameter satisfying $|d_0| < 3.5$ mm, and (2) while tight-primary tracks must have at least nine hits in the silicon pixel or microstrip layers, including at least one IBL or B-layer pixel hit, tight-secondary tracks need only seven silicon hits and are not required to have an IBL or B-layer hit [25,81,82].

The measured observables require identification of the charged decay products of $b$-hadron candidates, which are reconstructed from inner-detector tracks matched to calorimeter jets by identifying displaced secondary vertices. The matching of detector-level tracks to jets uses a cone whose angular radius shrinks with increasing jet $p_T$ from $\Delta R = 0.42$ for 30 GeV jets to an asymptotic minimum of $\Delta R = 0.24$:

$$\Delta R_{\text{max}} = 0.24 + e^{-[1.22+(1.64 \times 10^{-2})p_T/\text{GeV}]}.$$  

This matching was optimized to maximize discrimination between jets that contain $b$ hadrons and those that do not. Secondary vertex (SV) construction from these tight-secondary tracks is performed with the single secondary vertex finder algorithm [25]; secondary vertices are required to have at least three associated tracks, since those with two or fewer were found to be dominated by cases in which at least one track was either not reconstructed or not properly matched to the secondary vertex.

The jet charged momentum is constructed using tracks from the hard-scatter primary interaction vertex (PV) [83] in addition to those from the $b$-hadron candidate secondary vertex. PV tracks are matched to calorimeter jets via ghost association [84]; in this procedure, tracks are used as infinitesimal-energy ("ghost") inputs to jet reconstruction, and each track is considered matched to the jet of which its ghost is a constituent. The total jet charged momentum is built from the union of PV and SV tracks.

$b$ tagging: Jets that are likely to contain a $b$ hadron are tagged using a multivariate discrimination algorithm (MV2\(c)\)) based on inputs from charged-track impact parameter and secondary vertex reconstruction with respect to the hard-scatter PV. The hard-scatter PV is selected as the
one with at least two associated tracks and the highest \( \sum p^2_T \), where the sum extends over all tracks with \( p_T > 400 \text{ MeV} \) matched to the vertex. The tagger performance was optimized using \( b \) jets, \( c \) jets, and light-flavor jets in \( t\bar{t} \) events to maintain a high tagging efficiency for \( b \) jets while maximally rejecting lighter-flavor jets. A working point with 70\% \( b \)-jet tagging efficiency with an approximate light-flavor jet (charm-jet) rejection factor of 400 (10) is used.

At particle level, jets are formed by clustering all final-state particles within the calorimeter acceptance \( |\eta| < 4.9 \), except neutrinos and prompt electrons and muons. Like detector-level jets, they are built using the anti-\( k \)-algorithm with \( R = 0.4 \). Fiducial jets are required to have \( p_T > 30 \text{ GeV} \) and \( |\eta| < 2.5 \). Weakly decaying \( b \) hadrons with \( p_T > 5 \text{ GeV} \) are matched to jets by finding the closest jet within \( \Delta R < 0.3 \) to each \( b \) hadron. The jets are then labeled as having \( b \)-quark flavor if they have exactly one associated weakly decaying \( b \) hadron. Jets with two associated \( b \) hadrons are explicitly excluded since they have very different features; the fraction of such jets predicted in simulation was found to be below the per-mille level.

### B. Detector event and probe-jet selection

In the detector-level selection, at least one lepton is required to be identified in the trigger system for each candidate event, using a logical OR of the most inclusive electron and muon triggers. Scale factors are applied to the simulation in order to correct for known differences in trigger efficiencies between the simulation and collision data [30], with residual uncertainties accounted for as described in Sec. VI.C.

In order to reject backgrounds with fewer than two prompt leptons (e.g., multijet and \( W + jets \) events), exactly two reconstructed leptons with opposite charge are required. Contributions from backgrounds with \( Z \) bosons are suppressed by requiring one lepton to be an electron and the other to be a muon. Exactly two reconstructed jets (after the lepton overlap removal is applied) are required in order to reject \( t\bar{t} \) events with light-flavor jets from initial- or final-state radiation. These two jets must be separated by \( \Delta R > 0.5 \) to avoid \( b \)-hadron decay products being matched to a jet that does not contain the initial \( b \) hadron.

A “tag-and-probe” technique, described below, is employed to measure the observables of interest for the selected jets. If the leading jet in the event is \( b \) tagged and the subleading jet has both \( |\eta| < 2.1 \) and a reconstructed secondary vertex with at least three associated tracks, then the subleading jet is considered a “probe-jet” and its observables are measured. The more stringent \( \eta \) requirement for probe jets ensures that the full jet area falls inside the ATLAS tracker volume. This process is then repeated with the roles of the leading and subleading jets reversed, and the fragmentation observables for all probe jets are measured. Using this procedure it is possible for both jets, only one jet, or even neither jet to be measured, but care is taken to not bias the measurement. After an event preselection requiring exactly two jets and at least one \( b \)-tagged jet, both jets are valid probe jets in about 25\% of \( t\bar{t} \) events, and only one jet is a valid probe jet about 17\% of the time. There is a small correlation between the jets’ probability of being measured, but this correlation is well modeled by the simulation.

### C. Particle-level fiducial event selection and observables

For the fiducial event selection, particle-level \( t\bar{t} \) events must contain exactly one prompt electron and one prompt muon, each with \( |\eta| < 2.5 \) and \( p_T > 25 \text{ GeV} \) after photon dressing. The leptons must have opposite charge. In addition, exactly two \( b \)-tagged jets with \( |\eta| < 2.1 \), \( p_T > 30 \text{ GeV} \), and \( \Delta R(jet,jet) > 0.5 \) are required. If a particle-level \( t\bar{t} \) event passes the above selection, then both jets are considered as possible fiducial probe jets.

Fiducial observables reflect their detector-level counterparts as closely as possible to reduce dependence on the simulation during unfolding. They are calculated using stable, charged particles with \( p_T > 500 \text{ MeV} \), which are called “fiducial charged particles.” Each fiducial probe jet must contain exactly one weakly decaying \( b \) hadron with \( p_T > 5 \text{ GeV} \) and at least three fiducial-charged-particle decay products. The charged momentum of the \( b \) hadron, \( \vec{p}_{ch}^b \), is then defined as the momentum sum of all fiducial charged particles from the \( b \)-hadron decay, and the \( b \)-jet charged component, \( \vec{p}_{ch}^b \), is built from all fiducial charged particles that either are jet constituents or originate from the \( b \)-hadron decay. The dressed leptons in the event are used to calculate the value of the \( \rho \) observable defined in Sec. I and again below.

To summarize, the four observables measured are the ratio of each \( b \) hadron and jet charged momenta transverse to the incoming proton beams,

\[
\zeta_{T,b} = \frac{p_{T,b}}{p_{T,jet}},
\]

the fraction of the jet charged momentum carried by the \( b \)-hadron decay products along the direction of the jet charged momentum,

\[
\zeta_{L,b} = \frac{\vec{p}_{b} \cdot \vec{p}_{ch}}{|p_{ch}^b|^2},
\]

the ratio of the transverse momentum of the \( b \) hadron to the average transverse momentum of the two leptons in the event,

\[
\rho = \frac{2p_{T,b}}{p_T + p_T'},
\]

032008-6
and the number of fiducial charged particles produced in the \(b\)-hadron decay, \(n_{b}^{\text{ch}}\).

### D. Sources of background

Several sources of background jets are expected in 13 TeV \(pp\) collisions. A “background probe jet” here is defined as a jet which passes the detector-level requirements but fails the fiducial definition outlined in Secs. IVA and IV C. Such jets can originate from \(t\bar{t} \rightarrow e\mu b\bar{b}\) events in which one \(b\) jet from a top-quark decay lies outside the detector acceptance but another jet (often from initial- or final-state radiation) is reconstructed as a valid detector-level probe. They can also arise from \(t\bar{t} \rightarrow e\mu b\bar{b}\) events that fail the fiducial event definition but pass the detector-level cuts. Both sources of \(t\bar{t}\) induced backgrounds are estimated using MC simulation.

Non-\(t\bar{t}\) processes with at least one prompt electron, at least one prompt muon, and jets are another source of background probe-jets and are also evaluated with MC simulation. In this analysis, the \(tW\), Drell-Yan, and diboson production processes are identified as non-\(t\bar{t}\) backgrounds with prompt leptons.

Finally, backgrounds with fewer than two prompt leptons (dominated by \(t\bar{t}\) and single-top-quark production) are expected to account for less than 1% of selected events and are estimated from the MC simulation.

### E. Comparison between prediction and observed data

Table I shows the numbers of selected events and probe-jet yields. The uncertainties on the predicted yields include all components discussed in Sec. VI. Figures 1–3 compare relevant probe-jet observables between the MC prediction and the observed data. In both cases, all uncertainties described in Sec. VI are included in the prediction uncertainty. Agreement within the estimated uncertainties is observed between the simulated and observed events, showing that effects from the underlying event, pileup, track reconstruction, and secondary-vertex finding are well modeled by the simulation for the purpose of \(b\)-hadron and \(b\)-jet measurements.

**Table I.** Summary of expected and observed event and probe-jet yields. The uncertainties on the predicted yields include all components discussed in Sec. VI. Note that the sums of column entries may not agree exactly with reported total yields due to rounding.

<table>
<thead>
<tr>
<th>Process</th>
<th>Predicted yields</th>
<th>Observed yields</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fiducial (t\bar{t})</td>
<td>(44000 \pm 9000)</td>
</tr>
<tr>
<td></td>
<td>Nonfiducial (t\bar{t})</td>
<td>(6700 \pm 1500)</td>
</tr>
<tr>
<td>Total (t\bar{t})</td>
<td>(76000 \pm 12000)</td>
<td>(51000 \pm 9000)</td>
</tr>
<tr>
<td>Single top</td>
<td>(4400 \pm 1500)</td>
<td>(1580 \pm 600)</td>
</tr>
<tr>
<td>(Z +)jets</td>
<td>(125 \pm 45)</td>
<td>(13.0 \pm 5.1)</td>
</tr>
<tr>
<td>Diboson</td>
<td>(90 \pm 34)</td>
<td>(9.7 \pm 3.9)</td>
</tr>
<tr>
<td>Total non-(t\bar{t})</td>
<td>(4600 \pm 1600)</td>
<td>(1600 \pm 600)</td>
</tr>
<tr>
<td>(b) jets</td>
<td>(52000 \pm 9000)</td>
<td>(1600 \pm 600)</td>
</tr>
<tr>
<td>(c) jets</td>
<td>(180 \pm 60)</td>
<td>(1580 \pm 900)</td>
</tr>
<tr>
<td>Other jets</td>
<td>(250 \pm 70)</td>
<td>(250 \pm 70)</td>
</tr>
<tr>
<td>Total prediction</td>
<td>(81000 \pm 13000)</td>
<td>(53000 \pm 9000)</td>
</tr>
</tbody>
</table>

**FIG. 1.** Comparison of detector-level probe-jet observable distributions between simulation and collision data: (a) probe-jet \(\eta\) and (b) \(p_T\). The nominal non-\(t\bar{t}\) background and fiducial \(t\bar{t} \rightarrow e\mu b\bar{b}\) predictions are shown in addition to the total prediction; the fiducial probe-jet histogram is not stacked on top of the non-\(t\bar{t}\) background, in order to show the expected fraction of fiducial events. All systematic uncertainties are included in the uncertainty on the total prediction. The first and last histogram bins do not include the underflow and overflow entries.
V. UNFOLDING PROCEDURE

The fully Bayesian unfolding technique is used to extract the posterior probability of the particle-level differential cross sections given the observed data \([72]\). A likelihood of the data given a prediction is constructed with the particle-level signal cross sections \(\vec{\sigma}_p\) as the parameters of interest. Systematic uncertainties, which are described in more detail in Sec. VI, are included in the likelihood as nuisance parameters (NPs), whose vector of values is denoted by \(\vec{\lambda}\).

To build the likelihood function, the predicted numbers of signal and background events is calculated in each bin of the detector-level observables (defined in Sec. IV C), given a set of \(\vec{\sigma}_p\) and \(\vec{\lambda}\) values. The total predicted event count, \(x_i\), in bin \(i\) of a detector-level observable is first constructed as the luminosity, \(L(\vec{\lambda})\), times the predicted background cross section, \(b_i(\vec{\lambda})\), plus the response matrix times the particle-level signal cross sections, \(\sum_j M_{ij}(\vec{\lambda})\sigma_j^p\), where \(j\) runs over the particle-level bin indices.

---

FIG. 2. Comparison of detector-level probe-jet observable distributions between simulation and collision data: (a) the probe-jet total charged \(p_T\) and (b) the number of tracks matched to the jet that originate from the primary vertex. The nominal non-\(t\bar{t}\) background and fiducial \(t\bar{t} \rightarrow e\mu b\bar{b}\) predictions are shown in addition to the total prediction; the fiducial probe-jet histogram is not stacked on top of the non-\(t\bar{t}\) background, in order to show the expected fraction of fiducial events. All systematic uncertainties are included in the uncertainty on the total prediction. The first and last histogram bins do not include the underflow and overflow entries.

FIG. 3. Comparison of detector-level probe-jet observable distributions between simulation and collision data: (a) the invariant mass of all track constituents of the jet secondary vertex and (b) the relative momentum of the secondary vertex transverse to the jet charged momentum, \(z_{ch\,rel} = |(\vec{p}_{ch\,b} \times \vec{p}_{ch\,jet})|/|\vec{p}_{ch\,jet}|^2\). The nominal non-\(t\bar{t}\) background and fiducial \(t\bar{t} \rightarrow e\mu b\bar{b}\) predictions are shown in addition to the total prediction; the fiducial probe-jet histogram is not stacked on top of the non-\(t\bar{t}\) background, in order to show the expected fraction of fiducial events. All systematic uncertainties are included in the uncertainty on the total prediction. The first and last histogram bins do not include the underflow and overflow entries.
\[ x_i(\vec{\sigma}^p, \vec{\lambda}) = \mathcal{L}(\vec{\lambda}) \cdot \left( b_i(\vec{\lambda}) + \sum_{j} M_{ij}^{p,d}(\vec{\lambda}) \sigma_j^p \right), \]

where the response matrices \( M_{ij}^{p,d}(\vec{\lambda}) \), background predictions \( b_i(\vec{\lambda}) \), and their related uncertainties are derived from simulation.

The background cross-section predictions \( b_i(\vec{\lambda}) \) are defined as

\[ b_i(\vec{\lambda}) = b_{i,0} + \sum_{k \in \text{NPs}} \lambda_k (b_{i,k} - b_{i,0}), \]

where \( b_{i,0} \) predicts the nominal background in bin \( i \), and \( b_{i,k} \) predicts the background in bin \( i \) with the 1\( \sigma \) variation corresponding to nuisance parameter \( k \). Detector response matrices are computed similarly:

\[ M_{ij}^{p,d}(\vec{\lambda}) = M_{ij,0}^{p,d} + \sum_{k \in \text{NPs}} \lambda_k (M_{ij,k}^{p,d} - M_{ij,0}^{p,d}), \]

where \( M_{ij,0}^{p,d} \) is the nominal probability of a jet in particle-level bin \( j \) to be observed in detector-level bin \( i \), and \( M_{ij,k}^{p,d} \) is the corresponding probability for systematic variation \( k \). A likelihood is then constructed as the product of Poisson probabilities over all detector-level bins \( i \) as a function of the model parameters, \( \vec{\sigma}^p \) and \( \vec{\lambda} \):

\[ L(\vec{d}|\vec{\sigma}^p, \vec{\lambda}) = \prod_{i} \text{Poisson}(d_i, x_i(\vec{\sigma}^p, \vec{\lambda})). \]

Finally, Gaussian priors with \( \mu = 0 \) and \( \sigma = 1 \) are imposed for the nuisance parameters corresponding to systematic variations; the one exception is the luminosity, for which a log-normal prior is used with \( \mu = 0 \) and \( \sigma = 0.021 \), corresponding to the luminosity uncertainty described in Sec. VI. A flat, non-negative prior is imposed on signal cross sections, \( \vec{\sigma}^p \). The posterior probability distribution given the observed data is then

\[ P(\vec{\sigma}^p, \vec{\lambda}|\vec{d}) \propto L(\vec{d}|\vec{\sigma}^p, \vec{\lambda}) \cdot \prod_{k \in \text{NPs}} \text{Prior}_k(p_k). \]

The maximum of this distribution over all parameters of interest and nuisance parameters is determined using gradient ascent. The distribution is then marginalized by integrating over the nuisance parameters via sampling with a Hamiltonian Markov chain Monte Carlo approach [85].

The dominant causes of imperfect detector response are tracking reconstruction inefficiencies, \( b \)-hadron daughter tracks not being correctly matched to the reconstructed secondary vertex, and tracks from pileup particles being incorrectly matched to the hard-scatter vertex. This results in a shift in the observables’ mean value of about -0.2 for \( z_{L,k}^{ch} \) and \( z_{L,b}^{ch} \), and about -1 for \( n_{b}^{ch} \), which are accounted for during unfolding.

The unfolding procedure was found to correctly reproduce particle-level spectra from detector-level observations. To check this, a “stress test” of the unfolding was performed: detector-level pseudodata from HERWIG 7.0.4, SHERPA 2.1.1, and POWHEG+PYTHIA 8 with the A14 tune \( \alpha_s \) variations were unfolded to the particle level. In all cases the central values of the particle-level posterior distributions were in excellent agreement with the true particle-level spectra, well within the total uncertainty of the unfolded pseudodata.

VI. SYSTEMATIC UNCERTAINTIES

Systematic uncertainties associated with the detector, reconstruction, and simulation are treated in this analysis by calculating the impact of each uncertainty on the background prediction and detector response matrix and including these variations in the unfolding model laid out in Sec. V. Here, the sources of systematic uncertainty considered in the unfolding of the measured observables are outlined.

A. Theory and modeling uncertainties

Six variations of the nominal \( \vec{t} \) model are considered. Their effects on both the response matrix and the predicted nonfiducial \( \vec{t} \) background are taken into account. In addition to the sources of uncertainty listed below, an uncertainty in the \( \vec{t} \) normalization in this phase space is considered. Recent experimental measurements in the exclusive \( \vec{t} + 0 \)-jet phase space [86] quote uncertainties at the 10% level, so a 10% uncertainty is assigned to the \( \vec{t} + 0 \)-jet cross section. Uncertainties from the \( \vec{t} \) matrix-element calculation and matching scheme were found to be small, as expected, and are neglected in the final analysis. The six uncertainties on the nominal \( \vec{t} \) model are:

(i) Parton shower and hadronization: POWHEG+HERWIG 7.0.4 is used to build an alternative prediction and its deviation from the nominal prediction is taken as an uncertainty.

(ii) Initial-state radiation: VAR3 of the A14 PYTHIA 8 tune corresponds to variations of \( \alpha_s(m_Z) \) between 0.115 and 0.140. This parameter is varied together with the hadronization and factorization scales in the matrix element and the \( h_{damp} \) value as described in Sec. III, to increase and decrease the amount of initial-state radiation in \( \vec{t} \) events.

(iii) Final-state radiation: The nominal POWHEG+PYTHIA 8 A14 sample is compared with two samples generated with POWHEG+PYTHIA 8 using the VAR2 eigentune variations. The latter include the effects of varying several parameters in PYTHIA 8, the largest of which is due to varying of the amount of final-state radiation by shifting \( \alpha_s^{FSR} \) to 0.111 and 0.139, respectively.
(iv) Production and decay fractions of $b$ hadrons in $t\bar{t}$ events: To assess uncertainties in the $b$-hadron species production fractions, the procedure developed in a recent ATLAS measurement of the top-quark mass [87] is followed. The $b$-hadron species production fractions and decay fractions in nominal simulation are reweighted to experimental world averages [88], and effect on the detector response matrix is taken as an uncertainty.

(v) Modeling of SV track multiplicity: The secondary-vertex track multiplicity distribution in probe jets is compared between the nominal prediction and data, as shown in Fig. 4(d). Weights are derived from any observed differences and applied to the simulation to achieve agreement with the data. The effect of this reweighting is propagated into the response matrix as a systematic uncertainty of the detector response.

(vi) Top-quark and $b$-jet $p_T$ modeling: The jet charged $p_T$ spectrum for the nominal prediction is compared with the observed data, as shown in Fig. 2(a). Weights are derived from the observed differences, and the effect of this reweighting is propagated into the response matrix and treated as a systematic variation.

B. Non-$t\bar{t}$ modeling uncertainties

In addition to uncertainties in the prediction of $t\bar{t}$ events, a 30% uncertainty is assessed for the normalization of three non-$t\bar{t}$ background processes with two prompt leptons: $tW$, Drell-Yan, and diboson production. The largest of these backgrounds is from $tW$ production, for which the difference between the diagram removal and diagram subtraction schemes for calculating interference between single-top-quark and $t\bar{t}$ production is also considered [60].
C. Detector uncertainties

The systematic uncertainties related to imperfect understanding of the detector in data have an impact on the estimated background yield and on the signal detector response. The jet energy scale uncertainty consists of 18 eigenvector components and is derived with a combination of test-beam data, in situ measurements and simulation at 13 TeV \cite{89}. Further contributions originate from the $\eta$-intercalibration, jet-flavor composition and response, single-particle response and pileup effects. An in situ measurement of the jet response in dijet events \cite{78} was used to estimate the systematic uncertainty due to the jet energy resolution. An additional uncertainty from the efficiency of the JVT in removing pileup jets is evaluated.

Since measured observables are built from charged-particle tracks, only the overall event reconstruction efficiency is significantly affected by calorimeter jet uncertainties.

The uncertainties related to lepton identification, reconstruction, isolation and trigger efficiencies are taken into account by a variation of the corresponding scale factors in

![Graphs showing comparisons of particle-level observable distributions between MC and unfolded data.](image)

FIG. 5. Comparison of particle-level observable distributions between MC and unfolded data: (a) $z_{chL}$, (b) $z_{chT}$, (c) $\rho$, and (d) $n_{ch}$. The plotted data points correspond to the maximum likelihood for the particle-level cross section, and the gray uncertainty band in the lower panel of each figure shows the total uncertainty on the measured cross sections.
simulation within their assigned uncertainties. Additional uncertainties for the modeling of the lepton energy/momentum resolution and scale, obtained from measurements in data using \( J/\psi \to \ell\ell \) and \( Z \to \ell\ell \) events [73,74] are also considered. Their impacts on the predicted background yields and detector response are taken into account.

To correct for different \( b \)-tagging efficiencies in simulation and data, scale factors are applied to the simulation. Different scale factors have been obtained for \( b \) jets, \( c \) jets, and light-parton jets from \( t\bar{t} \) data, dijet data, and simulated events [82,90,91], which are then combined to give nominal scale factors and associated uncertainties. The total uncertainty due to the \( b \)-tagging efficiency and fake rate in this analysis was obtained by varying these scale factors within their respective uncertainties.

Uncertainties due to differences between the simulation and data in the detector alignment, track reconstruction efficiency, track fake rate, and track impact parameter (\( d_0 \) and \( z_0 \)) resolutions are determined using \( Z \to \mu\mu \), dijet, and minimum-bias events [80,92]. The distribution of the average number of interactions per bunch crossing (\( \langle n_{\text{PU}} \rangle \)) is altered to assess the effect of possible mismodeling of additional pileup vertices on the measurement [93,94]. The uncertainty in the combined 2015–2016 integrated

**TABLE II. Summary of observed \( b \)-jet spectra in particle-level bins for the four observables.**

<table>
<thead>
<tr>
<th>( \chi_{L,b}^{\text{ch}} ) range</th>
<th>0.0–0.4</th>
<th>0.4–0.5</th>
<th>0.5–0.6</th>
<th>0.6–0.7</th>
<th>0.7–0.8</th>
<th>0.8–0.9</th>
<th>0.9–1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction</td>
<td>0.041 ± 0.015</td>
<td>0.034 ± 0.012</td>
<td>0.068 ± 0.016</td>
<td>0.107 ± 0.019</td>
<td>0.174 ± 0.018</td>
<td>0.283 ± 0.024</td>
<td>0.293 ± 0.033</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \chi_{T,b}^{\text{ch}} ) range</th>
<th>0.0–0.4</th>
<th>0.4–0.5</th>
<th>0.5–0.6</th>
<th>0.6–0.7</th>
<th>0.7–0.8</th>
<th>0.8–0.9</th>
<th>0.9–1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction</td>
<td>0.048 ± 0.017</td>
<td>0.028 ± 0.011</td>
<td>0.083 ± 0.016</td>
<td>0.114 ± 0.018</td>
<td>0.157 ± 0.018</td>
<td>0.289 ± 0.026</td>
<td>0.281 ± 0.033</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \rho ) range</th>
<th>0.0–0.4</th>
<th>0.4–0.7</th>
<th>0.7–1.0</th>
<th>1.0–1.3</th>
<th>1.3–1.6</th>
<th>1.6–2.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction</td>
<td>0.323 ± 0.022</td>
<td>0.323 ± 0.012</td>
<td>0.175 ± 0.008</td>
<td>0.093 ± 0.010</td>
<td>0.047 ± 0.006</td>
<td>0.038 ± 0.012</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( n_{b}^{\text{ch}} ) range</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7–8</th>
<th>9–10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction</td>
<td>0.185 ± 0.032</td>
<td>0.255 ± 0.032</td>
<td>0.190 ± 0.033</td>
<td>0.176 ± 0.027</td>
<td>0.164 ± 0.034</td>
<td>0.030 ± 0.006</td>
</tr>
</tbody>
</table>

**FIG. 6.** Sources of uncertainty in the \( b \)-jet fractions in each particle-level observable bin for (a) \( \chi_{L,b}^{\text{ch}} \), (b) \( \chi_{T,b}^{\text{ch}} \), (c) \( \rho \), and (d) \( n_{b}^{\text{ch}} \).
luminosity is 2.1% [94], obtained using the LUCID-2 detector [93] for the primary luminosity measurements.

**D. Uncertainty pruning**

Systematic uncertainties with a negligible impact on the background prediction and response matrices are not included in the unfolding procedure; variations that change the predicted background by less than 5% in all detector-level observable bins and alter the response matrix by less than 0.01% in all bins are removed. Tracking uncertainties from the alignment of the inner detector, and uncertainties in the lepton momentum scales and reconstruction efficiencies, the JVT efficiency for hard-scatter jets, the charm- and light-flavor jet-$b$-tagging mistag rates, the normalization of the $Z$ + jets and diboson production cross sections, and the predicted yield of events with only one prompt lepton are all found to have only a small impact on the unfolding ingredients and are therefore neglected.

**VII. RESULTS**

Detector-level observable spectra for the collected data and the prediction, both before and after unfolding, are shown in Fig. 4; all uncertainties in the background prediction and signal detector response, as described in Sec. VI, are included in the uncertainty bands. Reasonable agreement between the detector-level data and simulation is observed for $z_{ch,L}^b$, $z_{ch,T}^b$, and $\rho$. However, the prior prediction shows slightly fewer SV tracks than appear in the data, which motivates the data-based $b$-hadron production and decay fraction uncertainty described in Sec. VI. Based on the level of agreement between the data and the posterior detector-level distribution, it is concluded that the model
used to unfold the detector effects is sufficient to describe the observed data. The large difference between the prior and posterior uncertainty bands for the detector-level observables is driven by correlations between parameters in the posterior model, which are introduced during the fit to data.

Posterior distributions for the particle-level cross sections are extracted using the observed detector-level data, predicted detector response matrices, and predicted detector-level backgrounds. In all figures, data points correspond to the maximum-likelihood particle-level cross-sections, and the uncertainty bands are the 16%–84% quantiles of the posterior distribution. Figure 5 compares the unfolded particle-level observable spectra with MC predictions from POWHEG+PYTHIA 8, POWHEG+HERWIG 7, and SHERPA 2, which are commonly used in other ATLAS data analyses. The unfolded spectra are summarized in Table II. Figure 6 shows the systematic uncertainties of the fiducial b-jet fractions broken down by source; the statistical uncertainty is dominant in the middle bins of the \( z_{L,b} \) and \( z_{T,b} \) distributions but subdominant elsewhere.

The effect on these observables of changing the values of nonperturbative MC parameters, which were largely tuned to the LEP and SLC \( e^+e^- \rightarrow b\bar{b} \) data, is explored. Figure 7 compares the data with predictions using various choices of \( \alpha_{\text{FSR}} \) and \( r_B \), a parameter rescaling the b-quark mass, in the PYTHIA 8 parton shower with the Lund-Bowler fragmentation model [12]. The \( \alpha_{\text{FSR}} \) values 0.127, 0.111, and 0.139 are those used in the nominal and down/up variations of the A14 tune, respectively; in the original A14 tune, \( r_B \) is taken from the Monash tune and set to 0.855 [15]. In the context of measuring the top-quark mass, \( r_B \) was recently refitted to the electron-positron collider data with the nominal \( \alpha_{\text{FSR}} = 0.127 \) value from the A14 tune; a value of \( r_B = 1.050 \) was found to be optimal, and this choice is also compared with

FIG. 8. Comparison of particle-level observables between PYTHIA8 tunes and unfolded data for (a) \( z_{L,b}^{ch} \), (b) \( z_{T,b}^{ch} \), (c) \( \rho \), and (d) \( n_b^{ch} \).
data in Fig. 7 [87]. For $z_{c,h}^{L,b}$, the A14 PYTHIA 8 tune and its $\alpha_{FSR}^{b}$ variations accurately describe the data and provide a reasonable uncertainty envelope for $b$-quark fragmentation. Figure 8 shows alternative PYTHIA 8 setups that are in common use: aMC@NLO+PYTHIA, the Monash tune of PYTHIA interfaced to POWHEG, and the Monash tune together with the Peterson fragmentation model [6] again interfaced to POWHEG. The Peterson model is found to be in poor agreement with the unfolded data for the $z$ and $\rho$ observables.

Comparisons with POWHEG+HERWIG 7 and SHERPA predictions are shown in Figs. 9 and 10, respectively. While HERWIG 7.0.4 appears to overestimate $b$ hadrons with soft fragmentation, developments introduced in version 7.1 give significantly better agreement with LHC data. There has also been substantial progress in recent versions of SHERPA: the choice of parameter values used by ATLAS in SHERPA 2.2.1 showed significant mismodeling of $n_{ch}^{b}$, but in versions 2.2.8 and 2.2.10 this is largely corrected. Two SHERPA 2.2.8 curves are shown: in the first curve the default settings are used, and in the second the settings recommended in Ref. [95] are used. In the second, the shower evolution variable in the splitting functions involving heavy-flavor quarks was tuned according to LHC $Z + b\bar{b}$ production measurements. The prediction from the default settings is in much better agreement with the data.

Corresponding $p$ values are calculated by approximating the posterior probability distribution over the parameters of interest as a multivariate normal distribution and taking the log-likelihood of a given prediction as the test statistic. The $p$ values of the comparisons between unfolded data and generator predictions are presented in Table III. POWHEG+PYTHIA 8 A14 and its variations predict the data reasonably well, as do aMC@NLO+PYTHIA 8, POWHEG+PYTHIA 8 Monash, HERWIG 7.1.3, and SHERPA 2.2.8 and 2.2.10.
FIG. 10. Comparison of particle-level observables between SHERPA versions and unfolded data for (a) $z_{L,b}^{ch}$, (b) $z_{T,b}^{ch}$, (c) $\rho$, and (d) $n_{ch}^{bb}$.

TABLE III. Summary of $p$ values for various MC generator configurations based on the observed data distributions.

<table>
<thead>
<tr>
<th>Generator configuration</th>
<th>$z_{L,b}^{ch}$ $p$ value</th>
<th>$z_{T,b}^{ch}$ $p$ value</th>
<th>$\rho$ $p$ value</th>
<th>$n_{ch}^{bb}$ $p$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>POWHEG+PYTHIA 8 A14</td>
<td>0.24</td>
<td>0.85</td>
<td>0.19</td>
<td>0.98</td>
</tr>
<tr>
<td>POWHEG+PYTHIA 8 A14 $\alpha_{FSR} = 0.139$</td>
<td>0.09</td>
<td>0.33</td>
<td>0.28</td>
<td>0.98</td>
</tr>
<tr>
<td>POWHEG+PYTHIA 8 A14 $\alpha_{FSR} = 0.111$</td>
<td>0.04</td>
<td>0.32</td>
<td>0.07</td>
<td>0.98</td>
</tr>
<tr>
<td>POWHEG+PYTHIA 8 A14 $r_B = 1.05$</td>
<td>0.09</td>
<td>0.48</td>
<td>0.28</td>
<td>0.98</td>
</tr>
<tr>
<td>aMC@NLO+PYTHIA 8 A14</td>
<td>0.29</td>
<td>0.92</td>
<td>0.21</td>
<td>0.98</td>
</tr>
<tr>
<td>POWHEG+PYTHIA 8 Monash</td>
<td>0.13</td>
<td>0.57</td>
<td>0.32</td>
<td>0.97</td>
</tr>
<tr>
<td>POWHEG+PYTHIA 8 Monash Peterson</td>
<td>0.01</td>
<td>0.02</td>
<td>0.10</td>
<td>0.96</td>
</tr>
<tr>
<td>POWHEG+HERWIG 7.0.4</td>
<td>0.16</td>
<td>0.46</td>
<td>0.11</td>
<td>0.98</td>
</tr>
<tr>
<td>POWHEG+HERWIG 7.1.3</td>
<td>0.23</td>
<td>0.71</td>
<td>0.16</td>
<td>0.96</td>
</tr>
<tr>
<td>SHERPA 2.2.1</td>
<td>0.08</td>
<td>0.42</td>
<td>0.53</td>
<td>0.01</td>
</tr>
<tr>
<td>SHERPA 2.2.8 ($Z + bb$ tune)</td>
<td>0.0005</td>
<td>0.005</td>
<td>0.19</td>
<td>0.48</td>
</tr>
<tr>
<td>SHERPA 2.2.8</td>
<td>0.10</td>
<td>0.47</td>
<td>0.11</td>
<td>0.61</td>
</tr>
<tr>
<td>SHERPA 2.2.10</td>
<td>0.07</td>
<td>0.53</td>
<td>0.07</td>
<td>0.40</td>
</tr>
</tbody>
</table>
SHERPA 2.2.8 with modified $g \to b\bar{b}$ splitting functions and POWHEG+PYTHIA 8 with the Peterson fragmentation model are strongly disfavored by the data for the $z$ observables, as is SHERPA 2.1.1 for $n_{ch}$. POWHEG+PYTHIA 8 A14 with $\alpha_{FSR} = 0.111$ and POWHEG+HERWIG 7.0.4 are in mild disagreement with the data for the $z$ observables and $\rho$.

VIII. CONCLUSION

Due to their ubiquity in decays of top quarks and Higgs bosons, $b$ quarks have a substantial role to play in the analysis of LHC collision data and are likely to continue to be invaluable in future collider experiments. It is therefore important that the evolution and hadronization of heavy quarks be well understood. As such, measurements of several observables sensitive to the fragmentation of $b$ quarks have been carried out with the ATLAS detector, using $b$ quarks from top-quark decays in $t\bar{t} \to e\nu b\bar{b}$ events in 36 fb$^{-1}$ of LHC $pp$ collision data at $\sqrt{s} = 13$ TeV. They complement similar measurements from $e^+e^-$ collider experiments in which the $b$ quarks originate from a color-singlet $Z/\gamma^*$ by probing the fragmentation process in a new environment as well as introducing new observables that are sensitive to the local hadronic energy flow. Since they provide a unique test of heavy-quark-fragmentation modeling at the LHC, the measurements were corrected for detector effects and compared with state-of-the-art MC event generator predictions. With a few exceptions (e.g., POWHEG+PYTHIA 8 Monash Peterson, POWHEG+PYTHIA 8 A14 $\alpha_{FSR} = 0.111$, and SHERPA 2.2.8 tuned to $Z + b\bar{b}$ measurements), generators tuned to a combination of lepton- and hadron-collider measurements yield predictions that are found to agree with the observed LHC data.

ACKNOWLEDGMENTS

We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently. We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWF and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; ANID, Chile; CAS, MOST and NSFC, China; Minciencias, Colombia; MEYS CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS and CEA-DRF/IRFU, France; SRSNFG, Georgia; BMBF, HGF and MPG, Germany; GSI, Greece; RGC and Hong Kong SAR, China; ISF and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands; RCN, Norway; MEiN, Poland; FCT, Portugal; MINE/IFA, Romania; JINR; MES of Russia and NRC KI, Russian Federation; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, Slovenia; DSI/NRF, South Africa; MICINN, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; MOST, Taiwan; TAEK, Turkey; STFC, United Kingdom; DOE and NSF, United States of America. In addition, individual groups and members have received support from BCKDF, CANARIE, Compute Canada and CRC, Canada; COST, ERC, ERDF, Horizon 2020 and Marie Skłodowska-Curie Actions, European Union; Investissements d’Avenir Labex, Investissements d’Avenir Idex and ANR, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programs cofinanced by EU-ESF and the Greek NSRF, Greece; BSF-NSF and GIF, Israel; Norwegian Financial Mechanism 2014-2021, Norway; NCN and NAWA, Poland; La Caixa Banking Foundation, CERCA Programme Generalitat de Catalunya and PROMETEO and GenT Programmes Generalitat Valenciana, Spain; Göran Gustafssons Stiftelse, Sweden; The Royal Society and Leverhulme Trust, United Kingdom.

The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK) and BNL (USA), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in Ref. [32].

40. P. Bärrreuther, M. Czakon, and A. Mitov, Percent-Level-Precision Physics at the Tevatron: Next-to-Next-to-Leading Order QCD Corrections to $q\bar{q} \rightarrow t\bar{t} + X$, Phys. Rev. Lett. 109, 132001 (2012).
45. M. Botje et al., The PDF4LHC working group interim recommendations, arXiv:1101.0538.


MEASUREMENTS OF JET OBSERVABLES SENSITIVE TO B- …

PHYs. REV. D 106, 032008 (2022)

MEASUREMENTS OF JET OBSERVABLES SENSITIVE TO B-…

PHYS. REV. D 106, 032008 (2022)
Measurements of jet observables sensitive to $B^-$... Phys. Rev. D 106, 032008 (2022)