Measurement of the inclusive $t\bar{t}$ production cross section in the lepton + jets channel in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector using support vector machines

Aaboud, M.; ATLAS Collaboration

DOI
10.1103/PhysRevD.108.032014

Publication date
2023

Document Version
Final published version

Published in
Physical Review D. Particles and Fields

License
CC BY

Citation for published version (APA):
Measurement of the inclusive $\bar{t}t$ production cross section in the lepton + jets channel in $pp$ collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector using support vector machines

M. Aaboud et al.*
(ATLAS Collaboration)

(Received 2 December 2022; accepted 14 March 2023; published 18 August 2023)

A measurement of the top quark pair-production cross section in the lepton + jets decay channel is presented. It is based on 4.6 fb$^{-1}$ of $\sqrt{s} = 7$ TeV $pp$ collision data collected during 2011 by the ATLAS experiment at the CERN Large Hadron Collider. A three-class, multidimensional event classifier based on support vector machines is used to differentiate $\bar{t}t$ events from backgrounds. The $\bar{t}t$ production cross section is found to be $\sigma_{\bar{t}t} = 168.5 \pm 0.7$(stat) $^{+6.2}_{-5.2}$(syst) $^{+3.4}_{-3.2}$(lumi) pb. The result is consistent with the Standard Model prediction based on QCD calculations at next-to-next-to-leading order.

DOI: 10.1103/PhysRevD.108.032014

I. INTRODUCTION

In the Standard Model of particle physics [1–3], the top quark ($t$) and the bottom quark ($b$) belong to a doublet representation of the weak-isospin SU(2). The top quark is the most massive of the known elementary particles. Because its mass is close to the electroweak symmetry breaking scale, it may play a fundamental role in the mechanism of breaking of the SU(2) symmetry of the electroweak interaction. Top quark production is also the dominant background in many analyses looking for physics beyond the Standard Model at high mass scales at the Large Hadron Collider (LHC), and a good understanding of top quark production is a necessary step in many “new physics” searches.

The cross section is one of the simplest observables that can be measured in the $\bar{t}t$ system. It allows one to make important comparisons with theoretical predictions available at next-to-next-to-leading order in perturbative QCD, including the soft-gluon resummation at next-to-next-to-leading-log order (NNLO + NNLL); see Ref. [4] and references therein. For $pp$ collisions at a center-of-mass energy of $\sqrt{s} = 7$ TeV, the predicted $\bar{t}t$ production cross section is $\sigma_{\bar{t}t}^{\text{NNLO+NNLL}} = 177_{-10}^{+10}$ pb. This theoretical value was calculated with the Top++ 2.0 program [5], including soft-gluon resummation, assuming a top quark mass value of 172.5 GeV, and using the PDF4LHC [6] procedures.

According to the Standard Model, top quarks from $pp$ collisions at the LHC are produced mostly via the strong interaction as $\bar{t}t$ pairs, with each top quark decaying into a $W$ boson and a $b$ quark nearly 100% of the time. The $\bar{t}t$ events in which one of the $W$ bosons decays into a quark pair and the other into a charged lepton and a neutrino are classified as “lepton + jets”, as such events contain an electron or muon or $\tau$ lepton, a neutrino and typically four hadronic jets (two of which originate from the $b$ quark and the $\bar{b}$ quark).

In this paper, a measurement of the top quark pair-production cross section at $\sqrt{s} = 7$ TeV using events with a single charged lepton (electron or muon) and jets in the final state is presented. The previously published result from the ATLAS Collaboration for the lepton + jets channel [7] uses 35 pb$^{-1}$ of data and obtains a precision of 12%. The most precise CMS $\bar{t}t$ cross-section measurement in the same channel [8] has a precision of 7%. In the dilepton channel, the best ATLAS result [9] achieves 3.5% precision, while CMS [10] reaches 3.6%. These ATLAS and CMS dilepton results have been combined [11], resulting in an uncertainty of 2.6% in $\sigma_{\bar{t}t}$.

The analysis presented in this paper is based on the full dataset collected with the ATLAS detector at the LHC in 2011, corresponding to an integrated luminosity of 4.6 fb$^{-1}$, and attains statistical and systematic uncertainties that are significantly lower than in previous ATLAS measurements in this final state. In an extension of the usual application of binary multivariate classifiers, this analysis uses a large number of variables to train three different support vector machines (SVMs). The three SVMs are used to define a three-dimensional space in which a multiclass event discriminator is constructed to identify the $\bar{t}t$ events through a simultaneous profile likelihood fit in four independent regions of this space.

*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.
II. ATLAS DETECTOR

The ATLAS detector is described in Ref. [12]. It is a multipurpose particle detector with forward-backward symmetry and a cylindrical geometry.¹ The inner tracking detectors are surrounded by a thin superconducting solenoid, electromagnetic and hadronic calorimeters, and a muon spectrometer with a magnetic field generated by three superconducting toroidal magnets of eight coils each. The inner-detector system (ID), in combination with the 2 T magnetic field from the solenoid, provides precision momentum measurements for charged particles within the pseudorapidity range |η| < 2.5. Moving radially outwards, it consists of a silicon pixel detector, a silicon microstrip detector, and a straw-tube tracker that also provides transition radiation measurements for electron identification. The calorimeter system covers the pseudorapidity range |η| < 4.9. A high-granularity liquid-argon (LAr) sampling calorimeter with lead absorber measures electromagnetic and hadronic showers within |η| < 3.2. In the region matched to the ID, |η| < 2.5, the innermost layer has fine segmentation in η to improve the resolution of the shower position and direction measurements. Hadronic showers are measured by an iron/plastic-scintillator tile calorimeter in the central region, |η| < 1.7, and by a LAr calorimeter in the end cap region, 1.5 < |η| < 3.2. In the forward region, measurements of both electromagnetic and hadronic showers are provided by a LAr calorimeter covering the pseudorapidity range 3.1 < |η| < 4.9. The muon spectrometer is instrumented with separate trigger and high-precision tracking chambers. It provides muon identification for charged-particle tracks within |η| < 2.7. The combination of all ATLAS detector systems provides charged-particle measurement along with lepton and photon measurement and identification in the pseudorapidity range |η| < 2.5. Jets are reconstructed over the full range covered by the calorimeters, |η| < 4.9.

A three-level trigger system [13] is used to select interesting events. The first-level (L1) trigger is implemented in hardware and uses a subset of detector information to reduce the event rate to a design value of at most 75 kHz. This is followed by two software-based trigger levels which together reduce the event rate to about 200 Hz.

An extensive software suite [14] is used in data simulation, in the reconstruction and analysis of real and simulated data, in detector operations, and in the trigger and data acquisition systems of the experiment.

¹ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the center of the detector and the z axis along the beam pipe. The x axis points from the IP to the center of the LHC ring, and the y axis points upwards. Cylindrical coordinates (r, φ) are used in the transverse plane, φ being the azimuthal angle around the z axis. The pseudorapidity is defined in terms of the polar angle θ as η = −ln tan(θ/2), and the distance ΔR in the η-φ space is defined as ΔR ≡ [(Δη)² + (Δφ)²]**(1/2).
as the sensitivity of a given jet to pileup. With this method [21], a contribution to the jet transverse momentum equal to the product of the jet area in the \( \eta-\phi \) plane and the average transverse energy density in a given event is subtracted. The effects of additional collisions in either the same bunch crossing or those adjacent in time are taken into account using corrections which depend on the average number of interactions per bunch crossing and the number of primary vertices. To avoid double counting of jets and electrons (which are reconstructed by independent algorithms), jets within \( \Delta R = 0.2 \) of a reconstructed electron are removed. For this analysis, only jets in the central region of the detector, \(|\eta| < 2.5\), and with transverse momentum \( p_T > 25 \) GeV are considered.

Identification of \( b \) jets.—The identification of “\( b \) jets” (jets arising from the decay of \( B \) hadrons) is performed with the MV1 algorithm [22], which combines the outputs of three different tagging algorithms into a multivariate discriminant. Jets are defined to be “\( b \) tagged” if the MV1 discriminant value is larger than a threshold (operating point) corresponding to \( 70\% \) efficiency to identify \( b \)-quark jets in simulated \( \bar{t}t \) events. Approximately \( 20\% \) of jets originating from charm quarks are identified as \( b \) jets, while light-flavor jets are mistagged as \( b \) jets at the \( 1\% \) level.

Missing transverse momentum.—The missing transverse momentum is calculated [23] as the complement of the vector sum, in the transverse plane, of calorimeter cell energies within \(|\eta| < 4.9\), after all corrections are applied to the associated physics objects (including jets, electrons, and muons). A correction for significant energy deposits not associated with high-\( p_T \) physics objects is also included. The magnitude of the missing transverse momentum vector is denoted by \( E_T^{\text{miss}} \), while its direction in the transverse plane is either denoted by an azimuthal angle \( \phi \) or inferred through its vector components \( E_T^{\text{miss}}_x \) and \( E_T^{\text{miss}}_y \).

IV. EVENT SELECTION

This analysis considers the single-lepton decay channel for the \( \bar{t}t \) pair. The selected events are required to have exactly one lepton (either an electron or a muon), a large amount of \( E_T^{\text{miss}} \), and three or more hadronic jets. The number of \( b \)-tagged jets in an event must be two or less. Events must have at least one primary vertex, with five or more tracks with \( p_T > 150 \) MeV. If there is more than one primary vertex, the one with the largest \( \sum \text{tracks} p_T^2 \) is chosen. Events were collected using single-lepton triggers, and each lepton candidate must be matched to an appropriate lepton trigger. In the muon channel, events are selected with a single-muon trigger with a \( p_T \) threshold of 18 GeV. For the electron channel, a single-electron trigger is required with \( p_T > 20 \) GeV. This is increased to 22 GeV during high instantaneous luminosity periods. Three or more jets with \( p_T \) greater than 25 GeV are required in each event. A large amount of \( E_T^{\text{miss}} \) is required to select events containing a neutrino. For electron events the \( E_T^{\text{miss}} \) must be greater than 30 GeV, while for muon events the \( E_T^{\text{miss}} \) is required to be greater than 20 GeV.

In order to reduce the background due to multijet production containing misidentified or nonprompt leptons, an additional selection requirement is imposed. Typically, in an event arising from this background, the missing transverse momentum vector points in the same direction in the transverse plane as the charged-lepton candidate. Therefore, electron candidate events must pass a requirement that \( m_T(W) > 30 \) GeV, while muon candidate events must have \( m_T(W) + E_T^{\text{miss}} > 60 \) GeV. Here, \( m_T(W) \) is defined as

\[
m_T(W) = \sqrt{2p_T E_T^{\text{miss}} (1 - \cos \Delta \phi)},
\]

where \( \Delta \phi \) is the difference in \( \phi \) between the direction of the charged-lepton transverse momentum, \( p_T^l \), and the missing transverse momentum vector.

V. DATA SAMPLES

The data sample used in this analysis comes from the \( \sqrt{s} = 7 \) TeV \( pp \) collisions collected during LHC Run 1 in 2011, and was recorded during stable beam conditions with all relevant ATLAS subdetector systems operational. It corresponds to an integrated luminosity of 4.6 fb\(^{-1}\), with an uncertainty of 1.8% [24].

VI. SIGNAL AND BACKGROUND MODELING

Except for the background due to multijet production leading to misidentified leptons (which is estimated from the data), all signal and background samples are modeled using Monte Carlo (MC) simulated events in conjunction with factors to correct the simulations to data where required.

A. Multijet background/fake leptons

The background from multijet events, in which a jet is misidentified as a muon or electron, or where a nonprompt lepton within a jet passes the tight lepton selection requirements, is sizable because of the large multijet production cross section. Events from these two sources of multijet background are referred to as fake-lepton events. This background is estimated using the so-called “matrix method,” which is based on the measurement of the selection efficiencies of leptons using data event samples satisfying relaxed identification criteria [25,26]. Loose electrons are electrons satisfying the baseline selection criteria where the requirements on particle identification using transition radiation measurements and on the energy-to-momentum ratio \( (E/p) \) are eased, and no requirement on the isolation is imposed. For loose muons, isolation is not required, but all other selection criteria are applied.
The matrix method is based on the measurement of the efficiencies for real and fake leptons in the loose lepton selection to pass the tight selection criteria. The real-lepton efficiencies are measured in $Z \rightarrow \ell\ell$ data samples, while the fake-lepton efficiencies are determined in data control regions with selection requirements designed to enhance the multijet content (1 lepton, $\geq 1$ jet, and only a small amount of $E_T^{miss}$). These efficiencies depend on both the lepton kinematics and event characteristics. To account for this, event weights are computed from the efficiencies parametrized as a function of a number of observables, which are then used to reweight the sample of data events with lepton candidates that satisfy the loose but not the tight selection criteria. The sums of these weights provide estimates of the multijet background.

B. Monte Carlo samples

The samples used in this analysis were obtained from a simulation chain consisting of an event generator interfaced to a parton shower and hadronization model, the outputs of which were passed through a simulation of the ATLAS detector and trigger system [27], and then reconstructed with the same algorithms as the data. The ATLAS detector response was modeled with the ATLAS full simulation (FS) based on GEANT4 [28]. For $t\bar{t}$ samples used to evaluate the signal modeling systematic uncertainties, the ATLAS fast simulation AtlasFast-II (AF) [27,29] was used to model the response of the detector.

The $t\bar{t}$ signal was simulated using the NLO POWHEG-bvq (patch 4) matrix element generator [30] interfaced with PYTHIA 6.425 [31], using parameter values set according to the c variant of the Perugia 2011 tune [32] to model the underlying-event and parton shower. The NLO CT10 [33] parton distribution function (PDF) set was used for the NLO matrix element part, and the LO PDF set CTEQ6L1 [34] was used with PYTHIA. The top quark mass was fixed at 172.5 GeV. This sample is referred to as POWHEG+PYTHIA 6. In order to evaluate the dependence on the choice of parton shower and fragmentation models, additional samples of $t\bar{t}$ events were created. The MC@NLO+HERWIG sample was created with the NLO MC@NLO [35] generator interfaced with HERWIG [36] using the LO AUET2 tune [37]. The POWHEG+HERWIG sample was created with POWHEG interfaced to HERWIG using the LO AUET2 tune.

The largest backgrounds to the $t\bar{t}$ events in the selected sample are from $W + $ jets and $Z + $ jets production. These were simulated with the LO event generator ALPGEN 2.13 [38] with LO PDF set CTEQ6L1, and interfaced with HERWIG 6.52. ALPGEN calculates matrix elements (ME) for final states with up to six partons. The MLM [39] matching procedure was used to remove the overlaps between ME and parton shower products in samples with $N$ and $N + 1$ final-state partons. In addition to the inclusive parton-flavor processes, separate ME samples of $W + b\bar{b} + $ jets, $W + c\bar{c} + $ jets, $W + c + $ jets, and $Z + b\bar{b} + $ jets were generated. The double counting of $b$ and $c$ quarks in $W/Z + $ jets that occurs between the shower of the inclusive samples and the ME of heavy-flavor production was eliminated using an overlap-removal algorithm based on parton-to-jet $\Delta R$ matching [40]. The $W + $ jets and $Z + $ jets event samples are referred to as ALPGEN+HERWIG.

The single-top backgrounds were simulated at NLO using the MC@NLO generator with the NLO PDF set CTEQ6.6 [34], and interfaced with HERWIG, except the $t$-channel samples, which were modeled with the LO ACERMC 3.8 generator [41] interfaced with PYTHIA. Dibosons ($WW, WZ$, and $ZZ$) were generated with HERWIG using the LO PDF set CTEQ6L1. All samples generated with HERWIG for the parton shower evolution and hadronization used JIMMY 4.31 [42] for the underlying-event model.

The effects of pileup were modeled by overlaying simulated minimum-bias events on the hard-scattering events. The Monte Carlo events were then reweighted such that the distribution of the number of interactions per bunch crossing, $\langle \mu \rangle$, matched the shape and observed average of $\langle \mu \rangle = 9.1$ in the 2011 data.

C. Signal and background classes

The most challenging backgrounds (i.e., those which most resemble $t\bar{t}$) are single-top and $W/Z + b\bar{b} + $ jets. Therefore, the $t\bar{t}$ cross-section measurement is expected to be affected most by the modeling of these backgrounds. To improve discrimination between the $t\bar{t}$ signal and the different types of background, this analysis separates the background events into two classes and treats them independently. The “heavy” background class includes the Monte Carlo samples for single-top, $W + b\bar{b} + $ jets, and $Z + b\bar{b} + $ jets. All other types of background, including fake leptons, are assigned to the group designated as the “light” class. Table I summarizes the composition of the classes and lists the datasets which are used to model them.

<table>
<thead>
<tr>
<th>Class</th>
<th>Process</th>
<th>Source</th>
<th>Fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Light</td>
<td>$W + l + $ jets</td>
<td>ALPGEN+HERWIG</td>
<td>74.9%</td>
</tr>
<tr>
<td>Light</td>
<td>$Z + l + $ jets</td>
<td>ALPGEN+HERWIG</td>
<td>7.6%</td>
</tr>
<tr>
<td>Light</td>
<td>Dibosons</td>
<td>HERWIG</td>
<td>1.4%</td>
</tr>
<tr>
<td>Light</td>
<td>Fake $e$</td>
<td>Data</td>
<td>6.4%</td>
</tr>
<tr>
<td>Light</td>
<td>Fake $\mu$</td>
<td>Data</td>
<td>9.7%</td>
</tr>
<tr>
<td>Heavy</td>
<td>$W + b\bar{b} + $ jets</td>
<td>ALPGEN+HERWIG</td>
<td>43.3%</td>
</tr>
<tr>
<td>Heavy</td>
<td>$Z + b\bar{b} + $ jets</td>
<td>ALPGEN+HERWIG</td>
<td>6.3%</td>
</tr>
<tr>
<td>Heavy</td>
<td>Single-top</td>
<td>MC@NLO/ACERMC</td>
<td>50.4%</td>
</tr>
</tbody>
</table>
TABLE II. The observed and expected numbers of events in the selected sample is shown. The first two columns list the contributions by physics process, and the two rightmost columns present events by class. The heavy class includes the W/Z + bb + jets and single-top processes, while the light class includes all other backgrounds. Predicted values are rounded with respect to their individual uncertainties.

<table>
<thead>
<tr>
<th>Process</th>
<th>Events</th>
<th>Class</th>
<th>Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>%t</td>
<td>86400 ± 5700</td>
<td>Signal</td>
<td>86400 ± 5700</td>
</tr>
<tr>
<td>W + jets</td>
<td>184000 ± 44000</td>
<td>Light</td>
<td>233000 ± 44000</td>
</tr>
<tr>
<td>Z + jets</td>
<td>19000 ± 4300</td>
<td>Heavy</td>
<td>21900 ± 1600</td>
</tr>
<tr>
<td>Dibosons</td>
<td>3200 ± 1600</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single-top</td>
<td>11040 ± 670</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fake leptons</td>
<td>37500 ± 8700</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>341000 ± 45000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observed</td>
<td>344 520</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The expected numbers of signal and background events in the selected sample are presented in Table II. The uncertainties shown include theoretical uncertainties in the production cross sections of the processes [4,43–47]. The W/Z + jets and diboson uncertainties include a contribution derived from an event yield comparison with SHERPA [48] Monte Carlo samples. The uncertainty in the number of events with fake leptons is estimated to be 20% for muons and 50% for electrons [49,50]. The observed number of events in data is in good agreement with the prediction.

VII. ANALYSIS METHOD

The SVM is a binary learning algorithm [51]. For any two classes of events, the signed distance from a hyperplane that separates the events is the SVM discriminant. For the analysis presented in this paper, a system of three support vector machines is used to create a three-dimensional multiclass event classifier to distinguish signal events from two classes of background (i.e., light and heavy). For events from any dataset, the distances from the three hyperplanes, trained to distinguish between signal vs light (SvL), signal vs heavy (SvH), and light vs heavy (LvH), are treated as the coordinates of points in a 3D decision space. The resulting templates of the prediction model are used in a binned likelihood fit to the analogous 3D distribution of the data events.

The SVM was chosen as the binary classifier because it is linear, it has firm mathematical foundations, and it offers a simple geometrical interpretation. Because the SVM method provides the solution to a straightforward convex optimization problem, the minimum it finds is a global one. The stopping point at the training stage is well defined, which therefore makes the method robust against overtraining. The method also works well in problems involving a large number of observables.

A. The SVM discriminant

Each event is described by N observables (i.e., features), and can be represented as a point, \( \vec{z} \), in an N-dimensional feature space. A linear binary classifier finds a hyperplane of dimension \( N - 1 \) that separates the two classes. Once the separating hyperplane is found, its reconstruction only requires the vectors that lie closest to the plane. These are the support vectors from whence the method derives its name. If the two classes to be discriminated are not linearly separable in the original feature space, this N-dimensional space can be mapped, \( \vec{z} \rightarrow \vec{\phi}(\vec{z}) \), into a higher-dimensional space in which the problem is linearly separable. Detailed knowledge of this mapping is not required when it is known how to calculate the inner product of the mapped vectors [52,53]. The distribution of classes in the mapped space can be probed directly by analyzing the multidimensional space which takes as its mathematical basis the SVM solutions for different class pairings.

The soft-margin SVMs [54] used in this analysis are constructed using a variant of the sequential minimal optimization algorithm [55] that includes the improvements suggested in Ref. [56].

In the case of a three-class problem like the one considered in this analysis, three different SVM classifiers are trained. Each SVM has the form

\[
f(\vec{z}) = \langle \vec{w} | \vec{z} \rangle - b = \sum_{i=1}^{SV} \lambda_i y_i K(\vec{v}_i, \vec{z}) - b,
\]

which is the generalized equation of a plane. The \( i \)th SVM (with \( i \in \{1, 2, 3 \} \)) has a normal vector given by \( |\vec{w}_i| \), and a constant offset \( b_j \) from the origin. The vectors \( \vec{v}_i \) are the support vectors from training (all other training vectors find their \( \lambda = 0 \)), the \( y \)'s are their “truth” values (\( \pm 1 \)), and the \( \lambda \)'s are parameters found in the training process along with \( b_j \). Hence, \( |\vec{w}_j| \) is a linear combination of training vectors mapped by \( \vec{\phi} \) to an alternative vector space. The bra-ket notation here serves as a reminder that these vectors belong to this mapped space. The inner product of two vectors in the mapped space given their nonmapped vectors \( \vec{x}_1 \) and \( \vec{x}_2 \) is determined via the kernel function \( K(\vec{x}_1, \vec{x}_2) \). The SVMs in this analysis use the Gaussian kernel:

\[
K(\vec{x}_1, \vec{x}_2) = \vec{\phi}(\vec{x}_1) \cdot \vec{\phi}(\vec{x}_2) = \exp(-|\vec{x}_1 - \vec{x}_2|^2/2\sigma^2).
\]

The width \( \sigma \) is an input parameter of the training process, along with an additional positive constant \( C \) which limits the range of the \( \lambda \)'s and is necessary for soft-margin SVMs.

In order to construct an orthonormal basis from the three trained SVMs, the Gram–Schmidt procedure [57] is used with their \( |\vec{w}_j| \) vectors.
Using this basis, 3-tuples (X, Y, Z) for a decision space are created:

\[
X(\vec{z}) = \frac{\langle z|w'_1 \rangle}{\sqrt{\langle w'_1|w'_1 \rangle}}, \quad Y(\vec{z}) = \frac{\langle z|w'_2 \rangle}{\sqrt{\langle w'_2|w'_2 \rangle}}, \quad Z(\vec{z}) = \frac{\langle z|w'_3 \rangle}{\sqrt{\langle w'_3|w'_3 \rangle}}.
\]

In this way, an input vector \( \vec{z} \) describing an event has coordinates in the XYZ space given by calculating \([X(\vec{z}), Y(\vec{z}), Z(\vec{z})]\). It is these new coordinates in the decision space which are used to describe all events, and this is the space in which the 3D templates of the likelihood function are created.

**B. Physics observables**

In this analysis, 21 physics observables are used to distinguish \( \pi^0 \) events from background events (see Table III). Twenty are kinematic variables, and one comprises the \( b \)-tagging information of the event. These include the electron or muon momentum, the number of jets in the event, the magnitude and direction of the missing transverse momentum vector, sums of the jet momenta components, the first five Fox–Wolfram moments (FWM), \( H_T \), the two largest eigenvalues of the normalized momentum tensor, and the mass of the lepton + jets system (\( m_{\ell} \)). The \( H_T \) is the scalar sum of \( E_{\text{miss}} \), electron \( p_T \) or muon \( p_T \), and the \( p_T \) of all jets passing the selection requirements.

Fox–Wolfram moments [58] were originally introduced for \( e^+e^- \) colliders. The FWMs correspond to a decomposition of the event’s phase space into Fourier modes on the surface of a sphere. They were modified for use at the Tevatron and the LHC to characterize the complex shapes of final states at hadron colliders [59,60]. They form a set of event shape variables, and the \( l \)th FWM (\( H_l \)) is defined in the following way:

\[
H_l = \frac{4\pi}{2l+1} \sum_{m=-l}^{l} \sum_{i=1}^{\text{jets}} \frac{E_T(i)}{E_T(\text{total})} Y_l^m(\theta_i, \phi_i)^2.
\]

The \( Y_l^m \)'s are the spherical harmonics, \( i \) runs over all selected jets in the event, and \( E_T(\text{total}) \) represents the sum of the transverse energy from selected jets. The angles \( \theta_i \) and \( \phi_i \) indicate the direction of the \( i \)th jet. This analysis makes use of \( H_1 \) through \( H_5 \).

The normalized momentum tensor uses the \( E_{\text{miss}} \) and the momenta of the lepton and up to five jets, and has the following form:

\[
P_{ij} = \frac{E_{T,\text{miss}}^\text{lep}}{|E_{T,\text{miss}}^\text{lep}|^2} \frac{E_{T,\text{miss}}}{|E_{T,\text{miss}}|^2} + \sum_{\text{jets}} p_i p_j |p|^2.
\]

Here \( i \) and \( j \) run over the \( x \), \( y \), and \( z \) components of momentum (for \( E_{\text{miss}}^\text{lep} \), only the \( x \) and \( y \) components are nonzero). The two largest eigenvalues of this “\( p \) tensor” are used as SVM inputs.

Because the lepton \( + \) jets decays are rotationally invariant in \( \phi \), some variables are calculated with respect to the lepton direction in the plane transverse to the beam. Hence, for the momenta of jets, \( p_{\perp} \) and \( p_{\parallel} \) denote the components which are parallel and perpendicular to the direction of the lepton in the transverse plane. Similarly, the \( \phi(E_{\text{miss}}) \) variable is then the angle between the transverse momentum of the lepton and the missing transverse momentum vector, and \( p_{\parallel} \) for the lepton corresponds to its entire transverse momentum.

The SVMs treat each variable as one of the coordinates of a point in a 21-dimensional space. The algorithm requires that each variable should fall roughly in the same numeric range so that all features contribute a similar weight when evaluating the distance from the separating hyperplane. The variables which have values outside the range of \((-1, +1)\) are transformed such that they approximately meet this requirement. All input variables and the values that were used to scale them are listed in Table III.

### Table III. List of the 21 variables used as input to the SVMs. The variables were divided by the given values to make them all of similar magnitude.

<table>
<thead>
<tr>
<th>Number</th>
<th>Feature</th>
<th>Divided by</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( E_{\text{miss}} ) [GeV]</td>
<td>250</td>
</tr>
<tr>
<td>2</td>
<td>( \phi(E_{\text{miss}}) ) [radians]</td>
<td>2( \pi )</td>
</tr>
<tr>
<td>3</td>
<td>Lepton ( E ) [GeV]</td>
<td>400</td>
</tr>
<tr>
<td>4</td>
<td>Lepton ( p_T ) [GeV]</td>
<td>400</td>
</tr>
<tr>
<td>5</td>
<td>Lepton ( p_T ) [GeV]</td>
<td>400</td>
</tr>
<tr>
<td>6</td>
<td>Mass(lepton + jets) [GeV]</td>
<td>750</td>
</tr>
<tr>
<td>7</td>
<td>Fox–Wolfram moment 1</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>Fox–Wolfram moment 2</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>Fox–Wolfram moment 3</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>Fox–Wolfram moment 4</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>Fox–Wolfram moment 5</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>Sum all jets ( E_T ) [GeV]</td>
<td>500</td>
</tr>
<tr>
<td>13</td>
<td>Sum all jets ( E ) [GeV]</td>
<td>750</td>
</tr>
<tr>
<td>14</td>
<td>Sum all jets ( p_T ) [GeV]</td>
<td>750</td>
</tr>
<tr>
<td>15</td>
<td>Sum all jets ( p_{\perp} ) [GeV]</td>
<td>750</td>
</tr>
<tr>
<td>16</td>
<td>Sum all jets ( p_\parallel ) [GeV]</td>
<td>750</td>
</tr>
<tr>
<td>17</td>
<td>( H_T ) [GeV]</td>
<td>500</td>
</tr>
<tr>
<td>18</td>
<td>( p )-tensor eigenvalue 1</td>
<td>1</td>
</tr>
<tr>
<td>19</td>
<td>( p )-tensor eigenvalue 2</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>Number of jets</td>
<td>10</td>
</tr>
<tr>
<td>21</td>
<td>Number of ( b )-tags</td>
<td>10</td>
</tr>
</tbody>
</table>
**C. SVM training**

The SVMs are trained to separate three classes of events: the light and heavy backgrounds and the $\bar{t}t$ signal. In order to train the SVMs, the Monte Carlo simulation samples and a data sample representing the multijet background are split into two subsamples. For training purposes, events from each class are randomly selected from those passing the selection requirements. The remaining events are used to test how well the trained SVMs perform. Also, it is only these remaining events that are utilized in the subsequent analysis.

The training process aims to find the set of support vectors that forms the optimal decision plane in the mapped space induced by the kernel function. As described in Sec. VII A, there are two free parameters that need to be specified when training. These are the $\sigma$ parameter of the Gaussian kernel and $C$, the positive constant which constrains the $\lambda$'s in the solution. A search grid over the values of these parameters was implemented, and the performance of a given training was then evaluated based upon the area of the resulting receiver operating characteristic (ROC) curve created with the events not used in the training. As a result of this study, the values of 1.2 for $\sigma$ and 2.0 for $C$ were settled upon. The $\bar{t}t$ signal class and the two background classes, light and heavy, each used 8,000 events for training, which is a small fraction of the total available events. Increasing the number of training events was not found to improve performance. It was also verified that the trained SVMs were not overtrained (i.e., that their discriminant distributions generalize well from the training set to the full class dataset).

**D. Class templates**

Different physics processes can be distinguished by their distinctly different distributions in the XYZ decision space. These distributions are obtained by applying the Gram–Schmidt procedure to each event’s SVM output values. Histogramming the different physics processes in the resulting 3D decision space creates probability distribution functions (i.e., templates) that can then be used in the likelihood fit.

Figure 1(a) shows a contour at a fixed value of the template function from each of the classes. This highlights the different regions in 3D decision space where the different event types congregate. As an alternative way of illustrating these distributions, Fig. 1(b) shows a sampling of 10,000 events from each class. Also shown are the three trained decision planes, which serve to demonstrate the nonorthogonal nature of the basis defined by the SVMs. The SvL plane at $X = 0$ is seen to separate signal from light as it extends downwards in Z. Similarly, the SvH and LvH planes separate their training classes. The multiple band structure in the templates arises because of training with the number of $b$ tags, which is a strong discriminant and is discrete. These template bands represent groupings of events with 0, 1, or 2 $b$ tags.

---

**FIG. 1.** Distribution of the three classes of events shown in the 3D orthonormalized decision space. (a) Contours from each class are shown together. (b) A view of the SVM decision space is shown with a random sampling of 10,000 events from each class. The three planes corresponding to each of the three trained SVMs are also depicted.
In order to minimize the potential effect of small fluctuations in the modeling, a small set of wider bins was constructed. The full 3D XYZ decision space was organized into four quadrants by dividing the space at \( Y = 0 \) and \( Z = -0.01 \). Each quadrant was then further divided into bins along the X axis. The quadrants are designated YZ, Yz, yZ, and yz, where the capitalization of the letters indicates where the quadrant is located (e.g., Yz means \( Y > 0.00 \) and \( Z \leq -0.01 \)). The division points for X were chosen to keep a minimum of approximately 1,000 events in each bin while preserving the shapes of the distributions. It is these four binned distributions from the quadrants that are used for the final profile likelihood fit.

### E. Cross-section measurement

A binned profile likelihood function is used in a fit to determine the \( t\bar{t} \) cross section from the data. In the likelihood fit, four templates are used: \( t\bar{t} \), \( W/Z \), single-top, and fake lepton. In evaluating the systematic uncertainties, particularly with respect to the modeling of \( t\bar{t} \), it was observed that a large uncertainty occurs because of the similarity between the final states of \( t\bar{t} \) events and single-top background events. To alleviate this effect the single-top backgrounds, which arise from electroweak processes rather than the strong interactions responsible for the production of \( t\bar{t} \) pairs, are combined into a single template normalized to their predicted cross sections. The \( W/Z + b\bar{b} + \text{jets}, the light-flavor W/Z + jets, and the diboson backgrounds are combined into a W/Z template. The normalizations of the \( t\bar{t} \), \( W/Z \), and fake-lepton templates are free parameters of the fit.

The grouping of physics processes used when constructing the templates for the fit can differ from the class definitions used for training. At the training stage, the events are arranged in order to create SVMs that can distinguish between the \( t\bar{t} \) signal, the light backgrounds, and the heavy backgrounds. After training, each physics process can be reassigned to templates. The chosen allocation of the physics processes to four templates (\( t\bar{t} \), \( W/Z \), single-top, and fake lepton) results in smaller expected uncertainties in the \( t\bar{t} \) cross section.

The likelihood function uses the templates (projections onto the X axis from each of the quadrants) that have been built in the XYZ decision space. Each template has an associated strength parameter \( \theta \) in the likelihood. The maximum value of the likelihood is obtained in determining the central values of the \( \theta \) parameters. The systematic uncertainties of the fit results are also included in the likelihood as nuisance parameters (NPs, \( \alpha \)'s, or collectively \( \alpha \)) with Gaussian constraints. Each template is a function of the nuisance parameters in the likelihood, which is then able to capture the effects due to each source of systematic uncertainty.

The likelihood of an unknown sample for an \( n_t \) template problem is defined as

\[
L = \prod_{i}^{\text{bins}} P(\mathcal{E}_i, \alpha_i) \times \prod_{j}^{\text{NPs}} G(\alpha_j, \sigma_j).
\]

Here \( G(\alpha_j, \sigma_j) \) is the Gaussian constraint for the \( j \)th NP, \( \alpha_j \), with the corresponding uncertainty \( \sigma_j \); and \( P(\mathcal{E}_i, \alpha_i) \) is the Poisson probability mass function for the \( i \)th bin given the observed number of events, \( \alpha_i \), and the expected number of events, \( \mathcal{E}_i \):

\[
\mathcal{E}_i = \mathcal{E}(\vec{\theta}, \vec{\alpha})_i = \sum_j^{\text{np}} \theta_j N(\vec{\alpha})_j T_j(i, \vec{\alpha}).
\]

The templates are constructed such that \( T_j(i, \vec{\alpha}) \) gives the fractional number of events in template \( T_j \)'s \( i \)th bin. Consequently, the sum over all bins of a given template is equal to 1. The \( N(\vec{\alpha})_j \) are defined to be the total number of events expected from the \( j \)th template assuming an integrated luminosity \( \mathcal{L} \). To calculate this value, the following sum over all modeled processes belonging to the \( j \)th template is computed:

\[
N(\vec{\alpha})_j = \sum_k^{\text{processes}} \sigma_k \epsilon_k(\vec{\alpha})_k \mathcal{L}.
\]

The \( \sigma_k \) and \( \epsilon_k \) are the cross section and acceptance for the \( k \)th physics process. They are derived from MC simulation. For the multijet background, \( N_j \) is taken from the fake-lepton estimate.

A maximum-likelihood fit is performed to extract the values of the \( \theta \) and \( \alpha \) parameters. The 1\( \sigma \) uncertainty for a given parameter is taken to be the change in the value of that parameter which causes \( \ln L \) to decrease by 0.5 away from \( \ln L_0 \), when \( \ln L \) is maximized with respect to all other free parameters and where \( \ln L_0 \) is the global maximum. All \( \theta \) and \( \alpha \) parameters have both their +1\( \sigma \) and −1\( \sigma \) uncertainties determined in this way.

The \( \theta \) for the \( t\bar{t} \) class, multiplied by the assumed cross section, gives the measured value of the \( t\bar{t} \) cross section. Similarly, the uncertainty in \( \theta t \) from the fit, multiplied by the assumed cross section, gives the uncertainty in the \( \sigma_{t\bar{t}} \) measurement.

### F. Systematic uncertainties

All systematic uncertainties were evaluated using the profile likelihood fit. The systematic effects are incorporated into the templates, and each template is associated with appropriate nuisance parameters (\( \alpha \)'s) in the likelihood. A nuisance parameter that takes a value of 0 in the fit keeps the nominal template, while a value of +1 or −1 changes the template to look like the +1\( \sigma \)/−1\( \sigma \) effect. Templates at intermediate values of the \( \alpha \)'s are linearly interpolated. A Gaussian constraint is also applied to each \( \alpha \) in order to propagate its controlled uncertainty when the
TABLE IV. Summary table of the measurement uncertainties. Because the profile likelihood fit accounts for correlations, the total error is not simply the components added in quadrature. Individual effects were obtained by leaving groupings of NPs out of the fit, and calculating the square of each effect as the difference of the squares of the total error and the residual error.

<table>
<thead>
<tr>
<th>Source</th>
<th>$-1\sigma$</th>
<th>$+1\sigma$</th>
<th>$-1\sigma$</th>
<th>$+1\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Object Modeling</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leptons</td>
<td>$-3.1$</td>
<td>$+3.3$</td>
<td>$-1.8$</td>
<td>$+2.0$</td>
</tr>
<tr>
<td>Jets</td>
<td>$-2.9$</td>
<td>$+3.0$</td>
<td>$-1.7$</td>
<td>$+1.8$</td>
</tr>
<tr>
<td>$b$-tag</td>
<td>$-1.9$</td>
<td>$+2.0$</td>
<td>$-1.1$</td>
<td>$+1.2$</td>
</tr>
<tr>
<td>$c$-tag</td>
<td>$-0.4$</td>
<td>$+0.4$</td>
<td>$-0.3$</td>
<td>$+0.3$</td>
</tr>
<tr>
<td>Mistag</td>
<td>$-0.3$</td>
<td>$+0.3$</td>
<td>$-0.2$</td>
<td>$+0.2$</td>
</tr>
<tr>
<td>Missing transverse momentum</td>
<td>$-0.2$</td>
<td>$+0.2$</td>
<td>$-0.1$</td>
<td>$+0.1$</td>
</tr>
<tr>
<td><strong>$\bar{t}t$ Modeling</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generator</td>
<td>$-1.6$</td>
<td>$+1.8$</td>
<td>$-1.0$</td>
<td>$+1.1$</td>
</tr>
<tr>
<td>Shower/hadronization</td>
<td>$-2.4$</td>
<td>$+2.6$</td>
<td>$-1.4$</td>
<td>$+1.5$</td>
</tr>
<tr>
<td>Renormalization/factorization</td>
<td>$-1.4$</td>
<td>$+1.4$</td>
<td>$-0.8$</td>
<td>$+0.9$</td>
</tr>
<tr>
<td>ISR/FSR</td>
<td>$-2.4$</td>
<td>$+2.5$</td>
<td>$-1.4$</td>
<td>$+1.5$</td>
</tr>
<tr>
<td>Underlying event</td>
<td>$-0.7$</td>
<td>$+0.8$</td>
<td>$-0.4$</td>
<td>$+0.5$</td>
</tr>
<tr>
<td>Color reconnection</td>
<td>$-0.5$</td>
<td>$+0.5$</td>
<td>$-0.3$</td>
<td>$+0.3$</td>
</tr>
<tr>
<td>PDF</td>
<td>$-1.8$</td>
<td>$+1.9$</td>
<td>$-1.1$</td>
<td>$+1.1$</td>
</tr>
<tr>
<td><strong>Background Modeling</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$W/Z$ reweighting</td>
<td>$-1.0$</td>
<td>$+1.0$</td>
<td>$-0.6$</td>
<td>$+0.6$</td>
</tr>
<tr>
<td>$W/Z$/fakes relative normaliz.</td>
<td>$-1.2$</td>
<td>$+1.2$</td>
<td>$-0.7$</td>
<td>$+0.7$</td>
</tr>
<tr>
<td>Heavy-flavor fraction</td>
<td>$-1.1$</td>
<td>$+1.2$</td>
<td>$-0.7$</td>
<td>$+0.7$</td>
</tr>
<tr>
<td>Single-top</td>
<td>$-1.0$</td>
<td>$+1.0$</td>
<td>$-0.6$</td>
<td>$+0.6$</td>
</tr>
<tr>
<td><strong>Other</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data statistics</td>
<td>$-0.7$</td>
<td>$+0.7$</td>
<td>$-0.4$</td>
<td>$+0.4$</td>
</tr>
<tr>
<td>Template statistics</td>
<td>$-1.0$</td>
<td>$+1.0$</td>
<td>$-0.6$</td>
<td>$+0.6$</td>
</tr>
<tr>
<td>Luminosity</td>
<td>$-3.2$</td>
<td>$+3.4$</td>
<td>$-1.9$</td>
<td>$+2.0$</td>
</tr>
<tr>
<td>Total</td>
<td>$-6.8$</td>
<td>$+7.1$</td>
<td>$-4.0$</td>
<td>$+4.2$</td>
</tr>
</tbody>
</table>

data have no preference for that systematic effect. The profile likelihood fit then provides a simultaneous measurement of the $\theta$ and $\alpha$ parameters. In this way, the systematic effects are converted into a statistical framework that properly takes into account correlations and which can potentially lower the uncertainties in the measurement.

The individual effects of various sources of systematic uncertainty are displayed in Table IV. They are obtained by leaving groupings of nuisance parameters out of the fit, and calculating the square of each effect as the difference of the squares of the total error and the residual error.

**Object modeling.**—Systematic uncertainties in the lepton selection arise from uncertainties in lepton identification, reconstruction, and triggering. These are evaluated by applying tag-and-probe methods to $Z \rightarrow \ell\ell'$ events [16]. Uncertainties due to the energy scale and resolution are also considered for electrons and muons. These effects are evaluated by assigning each of them a separate nuisance parameter in the likelihood so as to allow the error source to be shifted both upwards and downwards by its uncertainty. The resulting systematic effects are summarized in Table IV as Leptons.

For jets, the main source of uncertainty is the jet energy scale (JES). The JES and its uncertainty are evaluated using a combination of test-beam data, LHC collision data, and simulation [21]. As a result of the in situ analyses for the calibration of the full 2011 dataset, the correlations between various JES uncertainty components are encoded in 21 subcomponents. These include statistical and method uncertainties, detector uncertainties, modeling and theory uncertainties, mixed detector and modeling uncertainties, and pileup. The JES uncertainty is evaluated by assigning a separate NP to each of these 21 JES subcomponents. The jet energy resolution is separated by process ($\bar{t}t$, single-top, and $W/Z +$ jets) and is assigned three corresponding NPs. These extra degrees of freedom allow differences in the kinematics and prevalence of $b$-quark, light-quark, and gluon jets in these processes to be better represented in the profile likelihood. The resulting uncertainties in $\sigma_{\bar{t}t}$ from these sources are indicated in Table IV as Jets.

The jet-flavor-dependent efficiencies of the $b$-tagging algorithm are calibrated using dijet events, and dilepton $\bar{t}t$ events from data. Differences in the $b$-tagging efficiency as well as $c$-jet and light-jet mistag rates between data and simulation are parametrized using correction factors, which are functions of $p_T$ and $\eta$ [22]. The $b$-tag systematic uncertainties were evaluated by constructing nine NPs that correspond to unique bins in jet $p_T$, as the uncertainties at low and high jet $p_T$ should be largely uncorrelated. Single NPs were used for the $c$-tag and Mistag systematic uncertainties. These systematic effects appear in Table IV as three uncertainties labeled $b$-tag, $c$-tag, and Mistag.

During the variation of nuisance parameters related to jets and leptons, the $E_T^{\text{miss}}$ is recalculated in accordance with the changes caused by those systematic effects. In this way, the jet and lepton uncertainties are propagated to the $E_T^{\text{miss}}$ uncertainty. However, the $E_T^{\text{miss}}$ uncertainty due to calorimeter cells not assigned to any other physics object is evaluated individually. Also, an additional 6.6% uncertainty due to pileup is applied to $E_T^{\text{miss}}$. Both of these are given separate NPs in the profile likelihood, and they are listed in Table IV under Missing transverse momentum.

**Modeling of $\bar{t}t$ events.**—Systematic uncertainties due to the choice of $\bar{t}t$ MC generator are evaluated by taking the full difference between POWHEG-HERWIG (AF) and MC@NLO-HERWIG (AF). The systematic uncertainty due to the choice of parton shower model is taken as the full difference between POWHEG-HERWIG (AF) and POWHEG-PYTHIA 6 (AF). These are listed as Generator and Shower/hadronization in Table IV, respectively.

The systematic error due to uncertainties in the modeling of initial- and final-state radiation (ISR/FSR) is evaluated using ALPGEN interfaced to PYTHIA 6. Monte Carlo samples were created in which the parameter that controls the
amount of ISR/FSR in ALPGEN was either halved or doubled. Half of the spread between the ALPGEN samples with raised and lowered ISR/FSR parameter values is taken as the systematic error.

The uncertainty due to renormalization and factorization scales is evaluated with two modified samples, generated with MC@NLO interfaced with HERWIG, in which parameters controlling the renormalization and factorization scales, introduced to cure the ultraviolet and infrared divergences in ME calculations, are simultaneously either halved or doubled. The full difference between the two samples is taken as the Renormalization/factorization error.

Each of the major $t\bar{t}$ modeling systematic uncertainties (Generator, Shower/hadronization, ISR/FSR, and Renormalization/factorization) is given a shape NP in each quadrant. The uncertainty in the normalization of events is assigned two NPs. One of these is used to track the migration of events between quadrants, and it mirrors the change in the relative normalization of the quadrants as observed when comparing the nominal and systematically shifted samples. The second NP is taken as an overall normalization error which corresponds to the normalization of events between fake-electron and fake-muon events.

The underlying-event modeling error is evaluated by comparing two different $t\bar{t}$ MC event samples produced with varied parameters in POWHEG+PYTHIA 6. One was generated with the Perugia 2011 central tune, and the other with Perugia 2011 upHI [32]. Both of these samples use the P2011 CTEQ5L PYTHIA tune, and not P2011C CTEQ6L1, which applies to the nominal $t\bar{t}$ MC sample. Their full difference is used as the measurement uncertainty of the underlying event, using a single NP in the profile likelihood.

All particles in the final state from the LHC $pp$ collisions must be color singlets. Different schemes for the color reconnection (CR) of the beam remnant and other outgoing hard collision objects are examined. The $t\bar{t}$ cross-section uncertainty due to this effect is estimated by comparing two different $t\bar{t}$ MC samples produced with POWHEG+PYTHIA 6. A reference sample was obtained using the Perugia 2011 central tune. The other sample was generated with Perugia 2011 noCR [32], and has modified color reconnection parameters. These samples use the P2011 CTEQ5L PYTHIA tune. The full difference between these two samples is taken as the CR uncertainty, using a single NP.

To estimate the uncertainty due to the choice of parton distribution function, the CT10 PDF set parametrization is examined using its 26 upwards and downwards systematic variations. Each of the 26 CT10 eigenvector components is assigned a separate NP in the profile likelihood.

**Background modeling.**—To estimate the error due to the shape of the $W/Z + jets$ backgrounds and to assess the effect of any mismodeling, background samples are reweighted to match data for each of the following variables, taken one at a time, in a signal-depleted control region: lepton $E$, $E_T^{miss}$, $m_{T}\ell$, $\sum_{jets} p_T$, and $\sum_{jets} p_T^\perp$. This control region was defined as events matching the nominal selection, but containing exactly three jets, none of which are $b$ tagged. The reweighting functions were applied only to $W/Z + jets$ samples (leaving $t\bar{t}$, single-top, and fake lepton untouched). Five NPs in the profile likelihood implement these functions such that the NPs turn the reweighting effects on and off, each according to the differences seen for these five variables in the data. In Table IV, these effects appear under the heading $W/Z$ reweighting. For the single-top shape, ACRMC samples with raised and lowered ISR/FSR parameter values are compared, and a single NP is assigned to this systematic uncertainty.

The effects due to the uncertainty of the single-top, $W + jets$, and $Z + jets$ cross sections are investigated by varying these cross sections within their theoretical errors. For the $W + jets$ background, a 4% uncertainty is applied to the inclusive $W$ boson cross section, with an additional 24% uncertainty added in quadrature at each ascending jet multiplicity [39,61]. This method is also applied to the $Z + jets$ cross sections. To evaluate the systematic uncertainty in the $t\bar{t}$ cross section due to the theoretical uncertainties in the single-top cross section, the single-top cross section is varied in accordance with the theoretical results, taken from Refs. [45–47]. The relative normalization within the $W/Z + jets$ MC sample is varied by raising and lowering the corresponding nominal relative yields of each jet multiplicity by their respective errors. Similarly, the relative normalization between fake-electron and fake-muon events is varied by raising and lowering their nominal predictions by their errors. The resulting effects are evaluated using appropriate NPs added in the profile likelihood fit. This uncertainty is quoted as $W/Z$ and fakes relative normalization in Table IV. Also, the uncertainty due to variations of the $W + jets$ heavy-flavor fraction is included via three NPs in the profile likelihood. These NPs place an additional 25% uncertainty on each of the assumed $W + b\bar{b} + jets$, $W + c\bar{c} + jets$, and $W + c + jets$ cross sections. Table IV summarizes these in the row labeled Heavy-flavor fraction.

**Template statistics / luminosity.**—For the profile likelihood fit, an additional fit parameter is introduced for each bin. These parameters are used to represent the Poisson fluctuation of the predicted number of events in each bin as estimated from the size of the Monte Carlo samples. The error propagated to $\theta_{ij}$ from these additional parameters is then an appropriate representation of the MC statistical error.

The integrated luminosity measurement has an uncertainty of 1.8% [24], and therefore each physics process is...
assigned an uncertainty of this magnitude. This systematic error is controlled by a single nuisance parameter in the likelihood.

The total measurement uncertainty, including individual groups of contributions, is listed in Table IV. The largest uncertainties are due to the lepton selection and luminosity, followed by the uncertainties due to JES, b-tagging, ISR/FSR, and other t\bar{t} modeling.

**Beam energy.**—The LHC beam energy during the 2012 $\sqrt{s} = 8$ TeV pp run was measured to be within 0.1% of the nominal value of 4 TeV per beam, using the revolution frequency difference of protons and lead ions during

![Graphs showing data and fit predictions](image)

**FIG. 2.** The fitted yields of signal and background processes compared with data, shown in four YZ quadrants divided along the X axis, as used in the fit. They are labeled quad1YZ, quad2Yz, quad3yZ, and quad4yz (the boundary letters are appended for easy reference). The lower panel shows the ratio of data to fit prediction. The shaded regions correspond to the statistical and systematic uncertainties. The first and last bins also contain any events found outside the range of the horizontal axis.
FIG. 3. The data distributions of six selected input variables are shown with their post-fit predictions in the selected sample. The predicted signal fraction is 24.8%. Shown are (a) the number of jets, (b) the number of $b$-tagged jets, (c) $H_T$, (d) the 4th Fox–Wolfram moment, (e) $E_{\text{miss}}$, and (f) the mass of the lepton and jets. Data are shown with the overlaid dots. The predicted events are shown for each of the templates used in modeling the data. The statistical and systematic error bands are given by the shaded regions. The first and last bins contain events found outside the range of the horizontal axis.
FIG. 4. The data distributions of six selected input variables are shown with their post-fit predictions in the selected sample for the signal-rich region with $X > 0$, $Y > 0$, and $Z < 0$. The predicted signal fraction in this region is 79.3%. Shown are (a) the number of jets, (b) the number of $b$-tagged jets, (c) the 4th Fox–Wolfram moment, (d) $E_T^{\text{miss}}$, and (e) the mass of the lepton and jets. Data are shown with the overlaid dots. The predicted events are shown for each of the templates used in modeling the data. The statistical and systematic error bands are given by the shaded regions. The first and last bins contain events found outside the range of the horizontal axis.
FIG. 5. The data distributions of six selected input variables are shown with their post-fit predictions in the selected sample for the background-rich region with $X < 0$, $Y < 0$, and $Z > 0$. The predicted signal fraction in this region is 3.1%. Shown are (a) the number of jets, (b) the number of $b$-tagged jets, (c) $H_T$, (d) the 4th Fox–Wolfram moment, (e) $E_T^{\text{miss}}$, and (f) the mass of the lepton and jets. Data are shown with the overlaid dots. The predicted events are shown for each of the templates used in modeling the data. The statistical and systematic error bands are given by the shaded regions. The first and last bins contain events found outside the range of the horizontal axis.
$p + Pb$ runs in early 2013 combined with the magnetic model errors [62]. A similar uncertainty in the beam energy is applicable to the 2011 LHC run. The approach used in Ref. [63] was therefore applied to the measurement using the $\sqrt{s} = 7$ TeV dataset. The uncertainty in the $t\bar{t}$ theoretical cross section due to this energy difference was calculated to be 0.27%, using the Top++ 2.0 program [5] and assuming that the relative change of the $t\bar{t}$ cross section for a 0.1% change in $\sqrt{s}$ is as predicted by the NNLO + NNLL calculation. It is negligible compared to other sources of systematic uncertainty.

**VIII. RESULTS**

The top quark pair-production cross section for $pp$ collisions at a center-of-mass energy of $\sqrt{s} = 7$ TeV in the lepton + jets channel is found to be

$$\sigma_{t\bar{t}} = 168.5 \pm 0.7\text{(stat)}^{+6.2}_{-5.0}\text{(syst)}^{+3.4}_{-3.2}\text{(lumi)} \text{ pb}.$$  

This result includes all systematic uncertainties as evaluated with the profile likelihood fit, with the statistical and luminosity errors listed separately.

Figure 2 shows a comparison between the observed and fitted numbers of events in each of the quadrants. An uncorrelated $\chi^2$ test was used to check that there is good agreement between the data and the fit results within the combined statistical and systematic error bands.

Comparison plots between data and the fit prediction are shown for a few selected input variables in Fig. 3 for all events. Analogous comparisons in signal-rich and background-rich regions of the XYZ space are shown in Figs. 4 and 5, respectively. The signal-rich region is defined by $X > 0$, $Y > 0$, and $Z < 0$, while the background-rich region lies in the opposite octant of XYZ space, and has $X < 0$, $Y < 0$, and $Z > 0$. The $X$ dimension corresponds to the signal vs light decision hyperplane, while the $Y$ and $Z$ dimensions are linear combinations of the other SVM hyperplanes and are the directions orthogonal to $X$.

Based on a correlated $\chi^2$ test, the data and the fit agree well within the combined statistical and systematic error bands for all 21 variables.

The measured $t\bar{t}$ cross section is in good agreement with the theoretical predictions based on the NNLO + NNLL calculations of $\sigma_{t\bar{t}}^{\text{NNLO+NNLL}} = 177^{+5}_{-6}\text{(scale)}\pm 9\text{(PDF}+\alpha_s)\text{pb} = 177^{+10}_{-11}\text{pb}$ for $pp$ collisions at a center-of-mass energy of $\sqrt{s} = 7$ TeV and a top quark mass of 172.5 GeV [4].

The ATLAS measurement of the $t\bar{t}$ cross section at 7 TeV in the dilepton channel [9] is $\sigma_{t\bar{t}} = 182.9 \pm 3.1\text{(stat)} \pm 4.2\text{(syst)} \pm 3.6\text{(lumi)} \text{ pb}$. Depending upon the assumptions made for the systematic uncertainty correlations between these two measurements, the significance of their discrepancy was found to be in the 1.9$\sigma$ to 2.1$\sigma$ range.

**A. Top quark mass dependence**

The result of the profile likelihood fit depends on the assumed mass of the top quark through differences in $t\bar{t}$ acceptance owing to lepton kinematics, and also from minor variations in the shape of the discriminant. The analysis in this paper assumes a top quark mass of $m_{\text{ref}} = 172.5$ GeV. The 2014 average of Tevatron and LHC Run 1 measurements of the top quark mass [64] gives a value of $m_t = 173.34 \pm 0.27\text{(stat)} \pm 0.71\text{(syst)}$ GeV. The current ATLAS average from 2019 [65] yields $m_t = 172.69 \pm 0.25\text{(stat)} \pm 0.41\text{(syst)}$ GeV.

The dependence of the $t\bar{t}$ cross section on the mass of the top quark was determined through alternative profile likelihood fits that assume different top quark masses. Monte Carlo samples for both $t\bar{t}$ and single-top with top quark mass values of 165.0, 170.0, 172.5, 175.0, 177.5, and 180.0 GeV were employed to measure the $t\bar{t}$ cross section assuming each of these masses. These measurements were then fitted to obtain the $t\bar{t}$ cross section’s top quark mass dependence. When constrained to go through this measurement’s central value at 172.5 GeV, the best-fit second-order polynomial for the $t\bar{t}$ cross section as a function of $\Delta m_t = m_t - m_{\text{ref}}$ is $\sigma_{t\bar{t}}(\Delta m_t) = 0.016 \cdot \Delta m_t^2 - 0.75 \cdot \Delta m_t + 168.5$ pb, with $\Delta m_t$ in GeV.

**IX. SUMMARY**

A measurement of the top quark pair-production cross section in the lepton + jets channel was performed with the ATLAS experiment at the LHC, using a multivariate technique based on support vector machines. The measurement was obtained with a three-class, multidimensional event classifier. It is based on 4.6 $fb^{-1}$ of data collected during 2011 in $pp$ collisions at $\sqrt{s} = 7$ TeV. The $t\bar{t}$ cross section is found to be $\sigma_{t\bar{t}} = 168.5 \pm 0.7\text{(stat)}^{+6.2}_{-5.0}\text{(syst)}^{+3.4}_{-3.2}\text{(lumi)} \text{ pb}$, which has a relative uncertainty of $-4.0 \pm 4.2\%$. This measurement is consistent with the Standard Model prediction based on QCD calculations at next-to-next-to-leading order.

**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; CNPq and FAPESP, Brazil; NSERC, NFC and CFI, Canada; CERN; ANID, Chile; CAS, MOST and NSFC, China; Minciencias, Colombia; MEYS CR, Czech Republic; DNRF and DN, Denmark; IN2P3-CNRS and CEA-DRF/IFR, France; SRNSFG, Georgia; BMBF, HGF and MPG, Germany; GSRI, Greece; RGC and Hong Kong SAR, China; ISF and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS,
APPENDIX: FIT VISUALIZATION

The fit to the data can be difficult to visualize because of the 3D nature of the decision space. A series of 2D and 1D projections of the 3D space have been created in order to better illustrate its characteristics.

Projections are provided in the XY, XZ, and YZ planes to give a qualitative comparison of the fit results with the data in XYZ space, and also to help visualize the definitions of the signal-rich and background-rich regions employed in Figs. 4 and 5. The 2D projections of the 3D decision space use the standard fit results in XYZ space, but project the constant binning obtained with cubes of edge length 0.008 in that space. On the right side of Fig. 6 the 2D projection plots show the fractional relative difference between the number of events predicted by the fit and the observed data. The 2D plots on the left side of Fig. 6 show the class composition of each bin when projected onto the XY, XZ, and YZ planes. The rectangle representing each 2D bin is colored from top to bottom such that the colors that fill it are in the same ratio as the predictions for each class in that bin. Contour lines, which each represent the concentration of events in the plane, are drawn on top of the colored bins. Contours exist for each of the following fractional values of the maximal 2D bin height: (0.05, 0.10, 0.20, 0.40, 0.60, 0.80, 0.90, 0.95).

The 1D projections of the 3D decision space onto the X, Y, and Z axes are shown in Fig. 7 for the results of the fit to the data. For visualization purposes, the full 3D decision space is projected onto each of these axes. These projections have the systematic error bands superimposed, and also include the ratio of data to fit prediction.
FIG. 6. Left: These 2D projections show the composition of each bin according to the template fit results. Bins are shaded in the same ratio as the fit prediction. The contours drawn on top of the bins represent the overall concentration of events. Contours are provided at 5%, 10%, 20%, 40%, 60%, 80%, 90%, and 95% of the maximal 2D bin height. Right: The 2D bins are shaded in accordance with the fractional difference between the observed data and the number of events predicted by the fit. The contours shown are the same as in the left column and provide a reference.
FIG. 7. The 1D projections of the data and expected events after the likelihood fit. X, Y, and Z form an orthogonal basis in the decision space defined by the three trained SVMs. X corresponds to the signal vs light plane. Y and Z are the directions orthogonal to X found with the remaining signal vs heavy and light vs heavy SVMs. The projections use the fit parameters obtained with the profile likelihood fit to four quadrants, but project a constant binning obtained with cubes of edge length 0.008 in XYZ space. In each of these plots, the prediction is broken down by the templates used in the fit. The first and last bins contain events found outside the range of the horizontal axis.
MEASUREMENT OF THE INCLUSIVE $i\bar{t}$ ...

( ATLAS Collaboration)

1 Department of Physics, University of Adelaide, Adelaide, Australia
2 Physics Department, SUNY Albany, Albany New York, USA
3 Department of Physics, University of Alberta, Edmonton AB, Canada
4 Department of Physics, Ankara University, Ankara, Türkiye
5 LAPP, Univ. Savoie Mont Blanc, CNRS/IN2P3, Annecy, France
6 Department of Physics, University of Arizona, Tucson Arizona, USA
7 Department of Physics, University of Texas at Arlington, Arlington Texas, USA
8 Physics Department, National and Kapodistrian University of Athens, Zografou, Greece
9 Institute of Physics, University of Texas at Austin, Austin Texas, USA
10 Institute of Physics, National Technical University of Athens, Zografou, Greece
11 Institute of Physics, Bogazici University, Istanbul, Türkiye
12 Institute of Physics, University of Physics, Istanbul, Türkiye
13 Institute of Physics, TOBB University of Economics and Technology, Ankara, Türkiye
14 IAP, Univ. Savoie Mont Blanc, CNRS/IN2P3, Annecy, France
15 High Energy Physics Division, Argonne National Laboratory, Argonne Illinois, USA
16 Department of Physics, University of Arizona, Tucson Arizona, USA
17 Department of Physics, University of Texas at Arlington, Arlington Texas, USA
18 Physics Department, National and Kapodistrian University of Athens, Zografou, Greece
19 Institute of Physics, University of Texas at Austin, Austin Texas, USA
20 Institute of Physics, Bogazici University, Istanbul, Türkiye
21 Institute of Physics, University of Physics, Istanbul, Türkiye
22 Institute of Physics, TOBB University of Economics and Technology, Ankara, Türkiye
23 Institute of Physics, Bogazici University, Istanbul, Türkiye
24 Institute of Physics, Azerbaijan Academy of Sciences, Baku, Azerbaijan

MEASUREMENT OF THE INCLUSIVE $t\bar{t}$ PHYS. REV. D 108, 032014 (2023)
Graduate School of Science, Kobe University, Kobe, Japan

AGH University of Science and Technology, Faculty of Physics and Applied Computer Science, Krakow, Poland

Marian Smoluchowski Institute of Physics, Jagiellonian University, Krakow, Poland

Institute of Nuclear Physics, Polish Academy of Sciences, Krakow, Poland

Faculty of Science, Kyoto University, Kyoto, Japan

Kyoto University of Education, Kyoto, Japan

Research Center for Advanced Particle Physics and Department of Physics, Kyushu University, Fukuoka, Japan

Instituto de Física La Plata, Universidad Nacional de La Plata and CONICET, La Plata, Argentina

Physics Department, Lancaster University, Lancaster, United Kingdom

Oliver Lodge Laboratory, University of Liverpool, Liverpool, United Kingdom

Department of Experimental Particle Physics, Jožef Stefan Institute and Department of Physics, University of Ljubljana, Ljubljana, Slovenia

School of Physics and Astronomy, Queen Mary University of London, London, United Kingdom

Department of Physics, Royal Holloway University of London, Egham, United Kingdom

Department of Physics and Astronomy, University College London, London, United Kingdom

Louisiana Tech University, Ruston Los Angeles, USA

Fysiska institutionen, Lunds universitet, Lund, Sweden

Centre de Calcul de l’Institut National de Physique Nucléaire et de Physique des Particules (IN2P3), Villeurbanne, France

Departamento de Física Teorica C-15 and CIAFF, Universidad Autónoma de Madrid, Madrid, Spain

Institut für Physik, Universität Mainz, Mainz, Germany

School of Physics and Astronomy, University of Manchester, Manchester, United Kingdom

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

Department of Physics, University of Massachusetts, Amherst Massachusetts, USA

School of Physics, McGill University, Montreal Quebec, Canada

School of Physics, University of Melbourne, Victoria, Australia

Department of Physics, University of Michigan, Ann Arbor Michigan, USA

Department of Physics and Astronomy, Michigan State University, East Lansing Michigan, USA

Group of Particle Physics, University of Montreal, Montreal Quebec, Canada

Fakultät für Physik, Ludwig-Maximilians-Universität München, München, Germany

Max-Planck-Institut für Physik (Werner-Heisenberg-Institut), München, Germany

Nagasaki Institute of Applied Science, Nagasaki, Japan

Graduate School of Science and Kobayashi-Maskawa Institute, Nagoya University, Nagoya, Japan

Department of Physics and Astronomy, University of New Mexico, Albuquerque New Mexico, United States of America

Institute for Mathematics, Astrophysics and Particle Physics, Radboud University/Nikhef, Nijmegen, Netherlands

Nikhef National Institute for Subatomic Physics and University of Amsterdam, Amsterdam, Netherlands

Department of Physics, Northern Illinois University, DeKalb Illinois, United States of America

Department of Physics, New York University, New York New York, United States of America

Ochanomizu University, Otsuka, Bunkyo-ku, Tokyo, Japan

Ohio State University, Columbus Ohio, USA

Faculty of Science, Okayama University, Okayama, Japan

Homer L. Dodge Department of Physics and Astronomy, University of Oklahoma, Norman Oklahoma, USA

Department of Physics, Oklahoma State University, Stillwater Oklahoma, USA

Palacký University, Joint Laboratory of Optics, Olomouc, Czech Republic

Institute for Fundamental Science, University of Oregon, Eugene, Oregon, USA

Graduate School of Science, Osaka University, Osaka, Japan

Department of Physics, University of Oslo, Oslo, Norway

Department of Physics, Oxford University, Oxford, United Kingdom

LPNHE, Sorbonne Université, Université Paris Cité, CNRS/IN2P3, Paris, France

Department of Physics, University of Pennsylvania, Philadelphia Pennsylvania, USA

Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh Pennsylvania, USA

Laboratório de Instrumentação e Física Experimental de Partículas—LIP, Lisboa, Portugal

Departamento de Física, Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal

Departamento de Física, Universidade de Coimbra, Coimbra, Portugal

Centro de Física Nuclear da Universidade de Lisboa, Lisboa, Portugal
MEASUREMENT OF THE INCLUSIVE $t\bar{t}$)|(PHYS. REV. D 108, 032014 (2023)

132aDepartamento de Física, Universidade do Minho, Braga, Portugal
132bDepartamento de Física Teórica y del Cosmos, Universidad de Granada, Granada (Spain), Spain
132cDep Física and CEFITEC of Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal
133Institute of Physics of the Czech Academy of Sciences, Prague, Czech Republic
133Czech Technical University in Prague, Prague, Czech Republic
135Charles University, Faculty of Mathematics and Physics, Prague, Czech Republic
136Particle Physics Department, Rutherford Appleton Laboratory, Didcot, United Kingdom
137IRFU, CEA, Université Paris-Saclay, Gif-sur-Yvette, France
138Santa Cruz Institute for Particle Physics, University of California Santa Cruz, Santa Cruz California, USA
139Dep Física, Pontificia Universidad Católica de Chile, Santiago, Chile
139bDepartamento de Física, Universidad Técnica Federico Santa María, Valparaíso, Chile
140Universidade Federal de São João del Rei (UFSJ), São João del Rei, Brazil
141Department of Physics, University of Washington, Seattle Washington, USA
142Department of Physics and Astronomy, University of Sheffield, Sheffield, United Kingdom
143Department of Physics, Shinshu University, Nagano, Japan
144Department Physik, Universität Siegen, Siegen, Germany
145Department of Physics, Simon Fraser University, Burnaby British Columbia, Canada
146SLAC National Accelerator Laboratory, Stanford California, USA
147Department of Physics, Royal Institute of Technology, Stockholm, Sweden
148Departments of Physics and Astronomy, Stony Brook University, Stony Brook New York, United States of America
149Department of Physics and Astronomy, University of Sussex, Brighton, United Kingdom
150School of Physics, University of Sydney, Sydney, Australia
151Institute of Physics, Academia Sinica, Taipei, Taiwan
152aE. Andronikashvili Institute of Physics, Iv. Javakhishvili Tbilisi State University, Tbilisi, Georgia
152bHigh Energy Physics Institute, Tbilisi State University, Tbilisi, Georgia
153Department of Physics, Technion, Israel Institute of Technology, Haifa, Israel
154Raymond and Beverly Sackler School of Physics and Astronomy, Tel Aviv University, Tel Aviv, Israel
155Department of Physics, Aristotle University of Thessaloniki, Thessaloniki, Greece
156International Center for Elementary Particle Physics and Department of Physics, University of Tokyo, Tokyo, Japan
157Graduate School of Science and Technology, Tokyo Metropolitan University, Tokyo, Japan
158Department of Physics, Tokyo Institute of Technology, Tokyo, Japan
159Department of Physics, University of Toronto, Toronto ON, Canada
160aTRIUMF, Vancouver BC, Canada
160bDepartment of Physics and Astronomy, York University, Toronto Ontario, Canada
161Division of Physics and Tomonaga Center for the History of the Universe, Faculty of Pure and Applied Sciences, University of Tsukuba, Tsukuba, Japan
162Department of Physics and Astronomy, Tufts University, Medford Massachusetts, USA
163Department of Physics and Astronomy, University of California Irvine, Irvine California, USA
164Department of Physics and Astronomy, University of Uppsala, Uppsala, Sweden
165Department of Physics, University of Illinois, Urbana Illinois, USA
166Instituto de Física Corpuscular (IFIC), Centro Mixto Universidad de Valencia—CSIC, Valencia, Spain
167Department of Physics, University of British Columbia, Vancouver British Columbia, Canada
168Department of Physics and Astronomy, University of Victoria, Victoria British Columbia, Canada
169Fakultät für Physik und Astronomie, Julius-Maximilians-Universität Würzburg, Würzburg, Germany
170Department of Physics, University of Warwick, Coventry, United Kingdom
171Waseda University, Tokyo, Japan
172Department of Particle Physics and Astrophysics, Weizmann Institute of Science, Rehovot, Israel
173Department of Physics, University of Wisconsin, Madison Wisconsin, USA
174Fakultät für Mathematik und Naturwissenschaften, Fachgruppe Physik, Bergische Universität Wuppertal, Wuppertal, Germany
175Department of Physics, Yale University, New Haven Connecticut, USA
176Yerevan Physics Institute, Yerevan, Armenia

aDeceased.
bAlso at Department of Physics, King’s College London, London, United Kingdom.