CMS Physics Analysis Summary

Contact: cms-pag-conveners-top@cern.ch 2012/09/20

First Measurement of the Cross Section Ratio $\sigma(t\bar{t}b\bar{b})/\sigma(t\bar{t}jj)$ in pp Collisions at $\sqrt{s} = 7$ TeV

The CMS Collaboration

Abstract

We present the first measurement of the cross section ratio $\sigma(t\bar{t}b\bar{b})/\sigma(t\bar{t}jj)$ in the dilepton decay mode, using a data sample corresponding to an integrated luminosity of 5 fb^{-1} collected in pp collisions at $\sqrt{s} = 7$ TeV with the CMS detector at the LHC. The cross section ratio $\sigma(t\bar{t}b\bar{b})/\sigma(t\bar{t}jj)$ is measured in the visible phase space corresponding to the detector acceptance, and corrected to particle level. The measurement is performed by means of a fit to the measured multiplicity distribution of b-tagged jets in a sample of dileptonic top quark pair candidate events with at least four reconstructed jets. The result of the measurement, $3.6 \pm 1.1(\text{stat.}) \pm 0.9(\text{syst.})\%$, is compared with various theory predictions.

1 Introduction

The Standard Model (SM) has so far been successful in describing the strong and electroweak interactions. With recent observations of a new particle with a mass around 125 GeV/c^2 [1, 2] and whose properties are consistent with that of the SM Higgs boson, the SM is closer to be completed. However, more data is needed to firmly establish that this particle is indeed the SM Higgs boson. In particular, it is necessary to determine the couplings of this new particle to fermions. Of particular interest is the couplings of this particle to the top quark. One of the most sensitive channels in the search for the SM Higgs boson, $H \to \gamma\gamma$, is expected to have top quark loops in the production and decay of the Higgs boson. In the SM, the Higgs boson is expected to couple to the top quark with a strength of about unity. Therefore, it is important to measure the direct coupling of the Higgs boson to the top quark and to check the consistency of the SM Higgs sector.

One of the most promising channels for a direct measurement of the top quark Yukawa coupling is the one where the Higgs boson is produced in association with a $t\bar{t}$ pair ($t\bar{t}H$). Assuming that the newly observed particle is the Higgs boson, it is expected to decay mostly to $b\bar{b}$, leading to a $t\bar{t}b\bar{b}$ final state. However, this final state has an irreducible non-resonant background from the production of a top quark pair in association with a $b\bar{b}$ pair, as predicted by higher order QCD. This final state has not yet been observed. In this note, we present the first measurement of the cross section ratio $\sigma(pp \to t\bar{t}b\bar{b})/\sigma(pp \to t\bar{t}jj)$, where the $t\bar{t}jj$ final state is defined by the presence of two additional jets in addition to two b quarks from the decays of the top quark pair as described in Section 2.

One of the primary motivations for measuring the cross section ratio rather than the absolute cross section is that many experimental uncertainties are expected to cancel. Luminosity uncertainty, jet reconstruction efficiency, lepton identification and isolation efficiencies will mostly cancel. The remaining systematic uncertainties are instead expected to be dominated by the knowledge of the b-jet tagging efficiency and corresponding false positive or mistagging rate. Moreover, many kinematic distributions are expected to be similar in $t\bar{t}jj$ and $t\bar{t}b\bar{b}$, leading to similar acceptances and eventually reduced systematic uncertainties.

Next-to-leading-order (NLO) QCD calculations are available for $t\bar{t}jj$ and $t\bar{t}b\bar{b}$ processes [3–8], but they suffer from large factorization and renormalization uncertainties [9, 10], due to the presence of two very different scales, the top quark mass, M_t, and the jet transverse momentum, p_T, in this process. However, the cross section ratio is expected to have a reduced dependence, and hence more predictive power. Therefore, this measurement can provide a good test of NLO QCD theory.

In this note, we present the first measurement of the cross section ratio $\sigma(t\bar{t}b\bar{b})/\sigma(t\bar{t}jj)$ in the dilepton decay mode, using a data sample corresponding to an integrated luminosity of 5.0 fb^{-1} collected in pp collisions with the CMS detector. The measurement was performed by means of a fit to the measured multiplicity distribution of b-tagged jets after requiring at least four reconstructed jets. The cross section ratio of $\sigma(t\bar{t}b\bar{b})/\sigma(t\bar{t}jj)$ is measured at the particle level in the visible phase space corresponding to the detector acceptance.

2 Signal Definition

The $t\bar{t}b\bar{b}$ signal is defined at generator level in a phase space region with at least 4 hard jets in the final state where there are two leptons from both W bosons in top quark decays. We look for leptons (electron or muon) from leptonic W boson decays in top quark pair events

with transverse momentum $p_T > 20$ GeV/c and absolute pseudo-rapidity $|\eta| < 2.5$. Electrons or muons originating from leptonic decays of taus produced in $W \to \tau\nu$ decays are included in the signal definition. Jets are obtained recombining all final state particles (after parton shower and hadronization) excluding neutrinos with the anti-k_T clustering algorithm [11] with clustering parameter $r = 0.5$ and are required to satisfy

$$p_T > 20 \text{ GeV/c}, \ |\eta| < 2.5, \ \Delta R(i,j) = \sqrt{\Delta\phi^2 + \Delta\eta^2} > 0.5 \quad (1)$$

where $\Delta R(i,j)$ is the minimal distance between any two jets. Jets which are close to a lepton within $\Delta R(l,j) < 0.5$ are removed. At the generator particle level, the b-jets are identified by the presence of a B-hadron as one of the jet constituents. A similar procedure is followed for c-jets. When a B-hadron is successfully matched, lighter quarks are not considered.

In this analysis, the $t\bar{t}jj$ final state is defined by the presence of two b quarks from the decays of the top quark pair, accompanied by two additional jets. If these two are b-(c-)jets, the event is classified as $t\bar{t}b\bar{b}$ ($t\bar{t}c\bar{c}$). The $t\bar{t}LF$ final state is defined as the events where the extra two jets are from Light Flavor (LF) jets such as gluon, u, d or s quark.

3 Data and Monte-Carlo Samples

This analysis is based on proton-proton (pp) collisions at $\sqrt{s} = 7$ TeV using the complete dataset recorded in 2011, corresponding to an integrated luminosity of 5.0 ± 0.1 fb^{-1}. Events are required to pass the double muon (electron) trigger with p_T threshold up to 17 GeV/c depending on run ranges for the $\mu^{\pm}\mu^{\mp}$ ($e^{\pm}e^{\mp}$) final state. In the $e^{\pm}\mu^{\mp}$ final state, events are required to pass the electron and muon cross trigger.

The simulated data samples for Standard Model backgrounds and signal were generated based on PYTHIA (v. 6.424) [12], MADGRAPH (v. 5.1.1.0) [13] or POWHEG (r1380) [14]. Several approximate next-to-next-to-leading-order (NNLO) calculations are available in the literature, in particular those of Kidonakis [15], Ahrens et al. [16] and the HATHOR program [17]. The $t\bar{t}$ normalization in this analysis is based on the approximate NNLO result of 165.5 pb [15]. The acceptance for selecting $t\bar{t}b\bar{b}$ and $t\bar{t}jj$ events is modeled using the MADGRAPH event generator with matrix elements (ME) providing up to four additional partons including b quarks. The generated events are interfaced with PYTHIA to provide the showering of the partons and to perform the matching of the soft radiation with the contributions from the ME. Tau decays are handled with TAUOLA (v. 2.75) [18]. The CMS detector response is simulated using GEANT4 (v. 9.4) [19].

The W+jets and Z+jets samples are simulated in MADGRAPH and normalized to a cross section of 31.3 ± 1.6 nb and 3.04 ± 0.13 nb, respectively. The total inclusive NNLO cross section is calculated with the FEWZ program [20]. Di-boson production is simulated with PYTHIA. The electro-weak production of single top quarks ($pp \to tW$ and $pp \to \bar{t}W$) is simulated in POWHEG with a cross section of 15.7 ± 0.4 pb calculated at approximate NNLO [21].

Because of the large LHC luminosity, there is typically more than one interaction per bunch crossing (so-called pile-up), with a maximum value of more than 20 in the 2011 data taking. The simulated multiplicity of these additional pile-up interactions is re-weighted in order to match the one observed in the data.

4 Event selection and background estimation

The physics objects used in this analysis are reconstructed through the particle-flow algorithm [22, 23]. Muon candidates are reconstructed by combining information from the inner tracker and from the muon system [24]. Muon candidates are required to have $p_T > 20$ GeV/c and $|\eta| < 2.4$. Additionally, the track associated with the muon candidate is required to have a minimum number of hits in the silicon tracker, to be consistent with originating from the beam spot, and to have a high-quality global fit including a minimum number of hits in the muon detector.

Electron candidates are reconstructed by a combination of a track and energy deposits in the ECAL, from the electron itself and from possible Bremsstrahlung photons radiated by the electron in the tracker material on its way to the ECAL. Electron candidates are required to have $p_T > 20$ GeV/c and $|\eta| < 2.4$. The electron candidate track is required to be consistent with originating from the beam spot. Requirements on electron identification variables based on shower shape and track-cluster matching are applied to the reconstructed candidates. The criteria is optimized to select electrons from W decays and reject fakes from jets or photon conversions. In addition, electrons from photon conversions are rejected.

In order to reduce the QCD background coming mainly from muons or electrons originating from semi-leptonic b decays, isolation criteria based on the particle-flow objects are applied. The isolation criteria are defined as the sum of the transverse momenta of the charged hadrons, neutral hadrons and photons in a cone around the lepton direction:

$$I_{rel} = \frac{\sum p_T^{charged\ hadron} + \sum p_T^{neutral\ hadron} + \sum p_T^{photons}}{p_T^{lepton}}, \text{ for } \Delta R < 0.3 \quad (2)$$

I_{rel} is required to be smaller than 0.20 and 0.17 for muons and electrons, respectively.

The efficiencies for the above lepton-identification requirements, calculated as a function of the p_T and η of the leptons using a tag-and-probe method in Z boson candidates, are in the range of 92–97% for muons and 82–97 % for electrons. The efficiencies measured in data are found to be very close to the estimates from the simulation. Based on an overall comparison of lepton-selection efficiencies in data and simulation, residual corrections are applied in simulation.

Neutrinos can not be measured directly, but their presence can be inferred through the missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$), which represents the momentum imbalance in the transverse plane. The $E_{\mathrm{T}}^{\mathrm{miss}}$ vector in the particle-flow algorithm is reconstructed as the opposite of the vector sum of the transverse momentum of all particle-flow particles reconstructed in the event [25].

Jets are reconstructed with anti-k_{T} clustering algorithm [11] (with clustering parameter equal to 0.5), using particle-flow objects as input. Jets are required to have $p_T > 30$ GeV/c and $|\eta| < 2.5$.

In order to identify jets from heavy flavor, we apply the Combined Secondary Vertex (CSV) algorithm [26], which uses secondary vertices together with track-based lifetime information, with medium working point (CSVM) or tight working point (CSVT) corresponding to about 60% or 45% b-tag efficiency, respectively. The mistagging rate to identify light-flavor jets as b-jets is about 1% for CSVM and 0.1% for CSVT.

The event selection proceeds in five steps, labeled in the following S1 to S5, described in the following paragraphs.

S1 Events are required to have at least one pair of isolated opposite sign leptons. Dilepton

candidate events with $M_{ll} < 12$ GeV/c^2 are removed to suppress events from multijet processes, at essentially no penalty for the collected signal.

S2 The invariant mass of the lepton pair is required to be outside a $\pm$ 12 GeV/c^2 window centered at the mass of the Z boson for the $\mu^+\mu^-$ and e^+e^- final states to remove the large background from $Z \to l^+l^-$ decays.

S3 Signal events typically contain large missing transverse energy due to the presence of neutrinos from the W boson decays. On the other hand, background processes such as Z/γ^*+jets or QCD multijet production typically do not have large MET. Therefore, the $E_{\mathrm{T}}^{\mathrm{miss}}$ from these backgrounds is expected to be very small or ideally zero. The QCD multijet background contribution is negligible once two isolated leptons are required. In order to suppress the remaining background contribution from Z/γ^*+jets production, the requirement $E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV/c is applied for the $\mu^+\mu^-/e^+e^-$ final state. For the $e^\pm\mu^\mp$ final state, no $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement is applied in order to improve the signal-to-background ratio.

S4 The event is required to have at least four reconstructed jets.

S5 The event should have at least two b-tagged jets, using either the medium or the tight working point of the b-tagging algorithm.

Since the b-tagging is a key element of the analysis, a detailed study of data and simulation comparison is performed, and corrections are applied when needed. The b-tagging efficiencies from data and MC simulation are generally not the same [26]. In order to take the difference into account, events from MC simulation are re-weighted by applying a scale factor. The b-tagging efficiency, which has a dependency on event topology, is obtained after the four jet selection (S4) as a function of p_T and η from the signal simulation. It was observed that the b-tagging efficiency is different depending on the origin of b-jets, whether or not the b-jet is from a top decay. We take into account the difference as an additional systematic uncertainty on b-tagging scale factor variation which is described in detail in Section 8.

The data and MC simulation b-tagging probabilities for each event are obtained taking into account jet flavor dependency for the scale factor and b-tagging efficiency. Once the probability for a given number of b-tagged jets is calculated, each event is re-weighted by the ratio of data and MC probability.

To estimate the background from Z/γ^*+jets events contributing to $\mu^+\mu^-$ and e^+e^- final states, we use the method described in Refs. [27, 28]. The number of Z/γ^*+jets events in data that pass the Z-boson veto on M_{ll} can be estimated from the number of events in data with a dilepton invariant mass within $76 < M_{ll} < 106$ GeV/c, scaled by the ratio ($\mathrm{R}_{out/In}$) of events that fail and that pass this selection, which is estimated from the simulated Z/γ^*+jets event sample.

The multijet background contribution is estimated directly from the data. It is mostly from semileptonic decays of $b\bar{b}$. Differential distributions of this background are obtained by inverting the isolation cut for both leptons. The number of the events in this selection is obtained at each step with the two leptons not fulfilling the isolation criteria. We extrapolate the multijet background estimate from the non-isolated region to the isolated region of the phase space by applying a normalization factor obtained in like-sign dilepton sample, where a cascade decay of one of the b quarks can lead to a like-sign muon pair.

5 Comparison between Data and Monte-Carlo

Figure 1 shows the distributions of the jet transverse momentum after requiring at least four jets (S4). The kinematic distributions of $t\bar{t}b\bar{b}$ and $t\bar{t}jj$ are shown to be similar to each other. Therefore, it is expected that the uncertainty due to the jet energy scale would be relatively small in the measurement of the ratio. The b-tagged jet multiplicity distributions for both the medium and the tight working points of the b-tagging algorithm and after requiring the presence of at least four jets (S1-S4), are presented in Fig 2. These distributions will be used to measure the cross section ratio $\sigma(t\bar{t}b\bar{b})/\sigma(t\bar{t}jj)$, as discussed in the next section. A comparison of the simulated and observed event yields after the full event selection (S1-S5) is shown in Tables 1 and 2. After the full event selection has been applied, the event sample is found to be dominated by $ttjj$ events. There is a good agreement between data and the Monte Carlo simulation plus multijet background, both for the normalization and for the shape of the jet p_T and the b-jet multiplicity distributions after requiring the presence of at least four jets (S1-S4).

Table 1: Expected and observed event counts after applying the final event selection, using the "medium" working point of the b-tagging algorithm (CSVM) and summing over $ee/\mu\mu/e\mu$ final states. Only statistical uncertainties are quoted.

Final state	ee	$\mu\mu$	$e\mu$	All
$t\bar{t}+b\bar{b}$	2.8±0.2	4.0±0.3	10.0±0.4	16.6±0.6
$t\bar{t}+c\bar{c}$	4.5±0.3	5.6±0.3	14.3±0.5	24.4±0.7
$t\bar{t}+LF$	110	132	367	608
$t\bar{t}$ others	19.5±0.6	26.1±0.7	69.3±1.2	115
multijet	2.4±1.5	2.9±1.7	4.5±2.1	9.8±3.1
$W+jets$	<0.1	<0.1	<0.1	<0.1
VV	<0.1	<0.1	0.1±0.1	0.1±0.1
Single top-tW	2.5±0.4	3.7±0.5	8.4±0.7	14.6±0.9
$Z/\gamma^* \to ll$	0.7±0.5	0.2±0.3	<0.1	0.9±0.6
Background total	139	170	463	772
Data	145	176	463	784

6 Fitting Method

The fraction of $t\bar{t}b\bar{b}$ events with respect to $t\bar{t}jj$ events is obtained from data by fitting the b-tagged jet multiplicity. Figure 3 shows the normalized distribution of the b-tagged jet multiplicity obtained using the CSVM and CSVT b-tagging algorithms for $t\bar{t}b\bar{b}$, $t\bar{t}c\bar{c}$ and $t\bar{t}LF$ separately after the four jet selection (S4) in the defined visible phase space. As shown in this figure, the shape of $t\bar{t}b\bar{b}$ can be distinguished from $t\bar{t}LF$ and $t\bar{t}c\bar{c}$ as $t\bar{t}b\bar{b}$ events tend to have higher b-jet multiplicities. The shape of $t\bar{t}c\bar{c}$ and $t\bar{t}LF$ are found to be similar to each other and therefore will not be differentiated in the analysis.

The fit function used to extract the fraction can be expressed in the following way:

$$f(k,R) = k \cdot N_{t\bar{t}jj} \cdot [R \cdot N^{norm}_{t\bar{t}b\bar{b}} + (1-R) \cdot N^{norm}_{t\bar{t}LF/t\bar{t}c\bar{c}}] + k \cdot N^{MC}_{bkg} + N^{Data-driven}_{bkg} \quad (3)$$

where R is the fraction of $t\bar{t}b\bar{b}$ events with respect to the number of $t\bar{t}jj$ events at reconstruction level. $N_{t\bar{t}jj}$ is the number of events for $t\bar{t}jj$ and $N^{norm}_{t\bar{t}b\bar{b}}$ and $N^{norm}_{t\bar{t}LF}$ are the normalized distributions

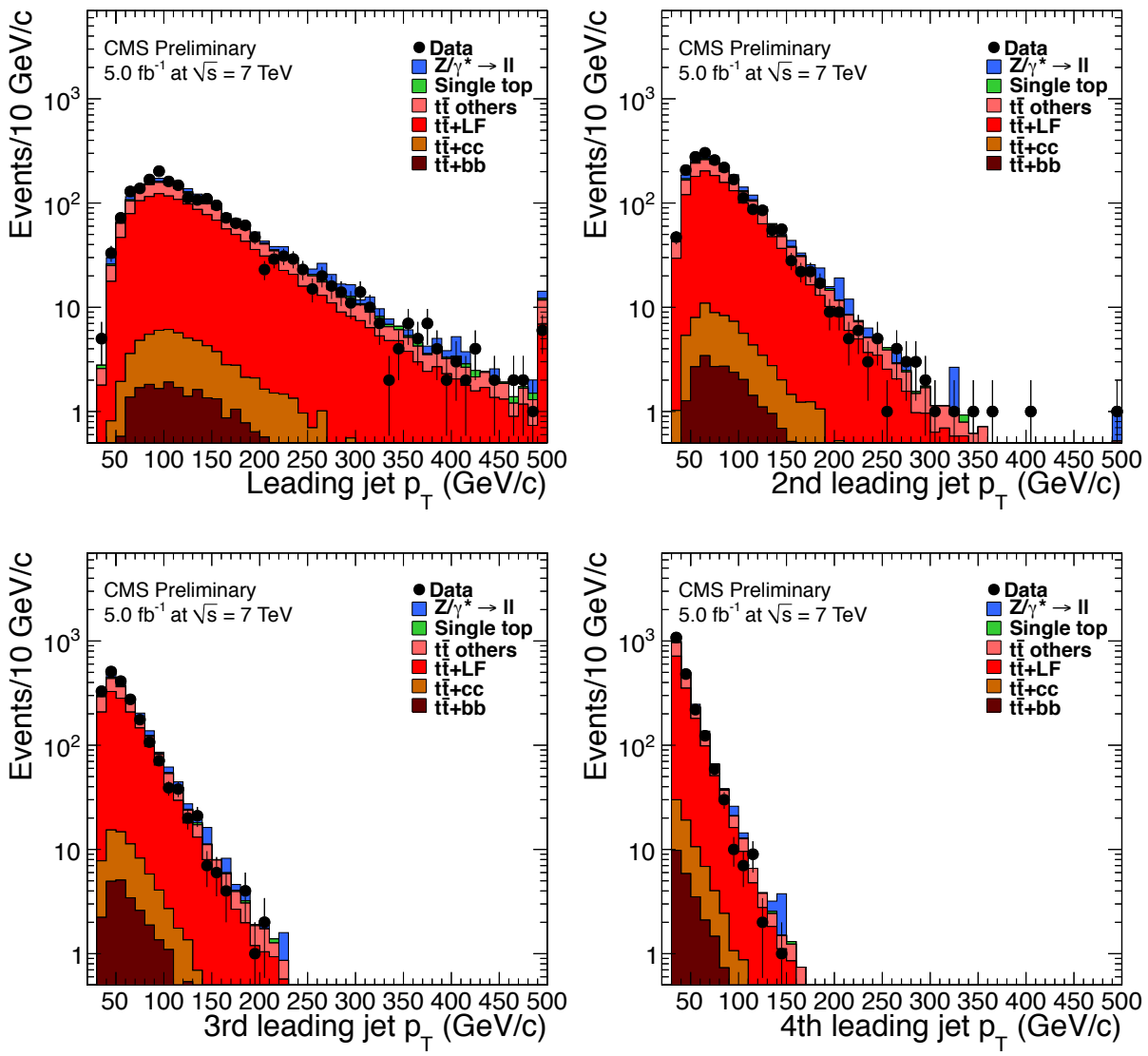

Figure 1: Distributions of the transverse momentum of the 1st, 2nd, 3rd, and 4th leading jet after the requirement of at least four reconstructed jets. The distributions correspond to the sum of the ee, $\mu\mu$ and $e\mu$ final states. Data are compared with Monte Carlo simulations where the $t\bar{t}b\bar{b}$ process is required to have two additional b-jets, $t\bar{t}c\bar{c}$ to have at least two c-jets and $t\bar{t}LF$ events are from Light Flavor (LF) jets such as gluon, u, d or s quark. The contribution labelled "$t\bar{t}$ others" is from top quark pair events that do not pass the visible phase space definition.

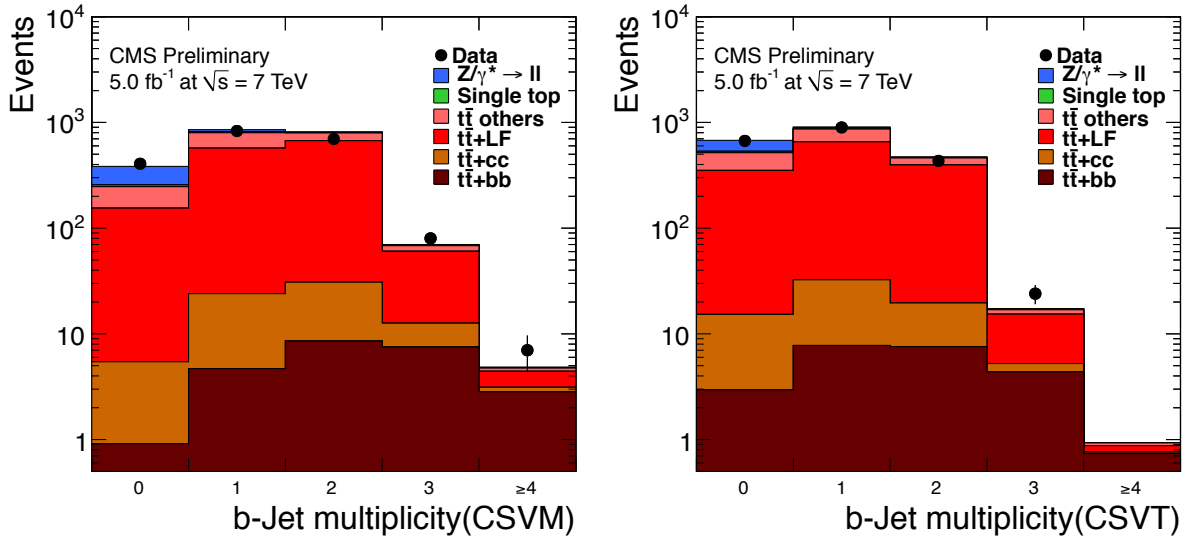

Figure 2: Distributions of the b-jet multiplicity for the "medium" and "tight" working point of the b-tagging algorithm (CSVM and CSVT), corresponding to a light flavor mistagging efficiency of 1 % and 0.1 %, respectively after the requirement of at least four reconstructed jets. The distributions correspond to the sum of the ee, $\mu\mu$ and $e\mu$ final states. Data are compared with Monte Carlo simulations where the $t\bar{t}b\bar{b}$ process is required to have two additional b-jets, $t\bar{t}c\bar{c}$ to have at least two c-jets and $t\bar{t}LF$ events are from Light Flavor (LF) jets such as gluon, u, d or s quark. The contribution labelled "$t\bar{t}$ others" is from top quark pair events that do not pass the visible phase space definition.

Table 2: Expected and observed event counts after applying the final event selection, using the "tight" working point of the b-tagging algorithm (CSVT) and summing over $ee/\mu\mu/e\mu$ final states. Only statistical uncertainties are quoted.

Final state	ee	$\mu\mu$	$e\mu$	All
$t\bar{t}+b\bar{b}$	1.8±0.2	2.9±0.2	7.1±0.4	11.8±0.5
$t\bar{t}+c\bar{c}$	2.2±0.2	2.8±0.2	7.0±0.4	12.0±0.5
$t\bar{t}+LF$	64.3±1.1	77.4±1.2	218	360
$t\bar{t}$ others	10.3±0.4	13.7±0.5	37.6±0.9	61.6±1.1
multijet	0.9±1.0	1.3±1.1	2.1±1.4	4.3±2.1
$W+jets$	<0.1	<0.1	<0.1	<0.1
VV	<0.1	<0.1	0.1±0.1	0.1±0.1
Single top-tW	1.6±0.3	2.0±0.3	4.4±0.5	8.0±0.7
$Z/\gamma^* \to ll$	0.1±0.2	0.2±0.3	<0.1	0.2±0.3
Background total	79.3±1.6	97.4±1.8	269	446
Data	77	98	281	456

for the combination of $t\bar{t}LF$ and $t\bar{t}c\bar{c}$, respectively. The factor by which the total MC simulation is scaled in order to match the data is denoted k. The background contributions obtained from simulation and data are labelled N_{bkg}^{MC} and $N_{bkg}^{Data-driven}$, respectively.

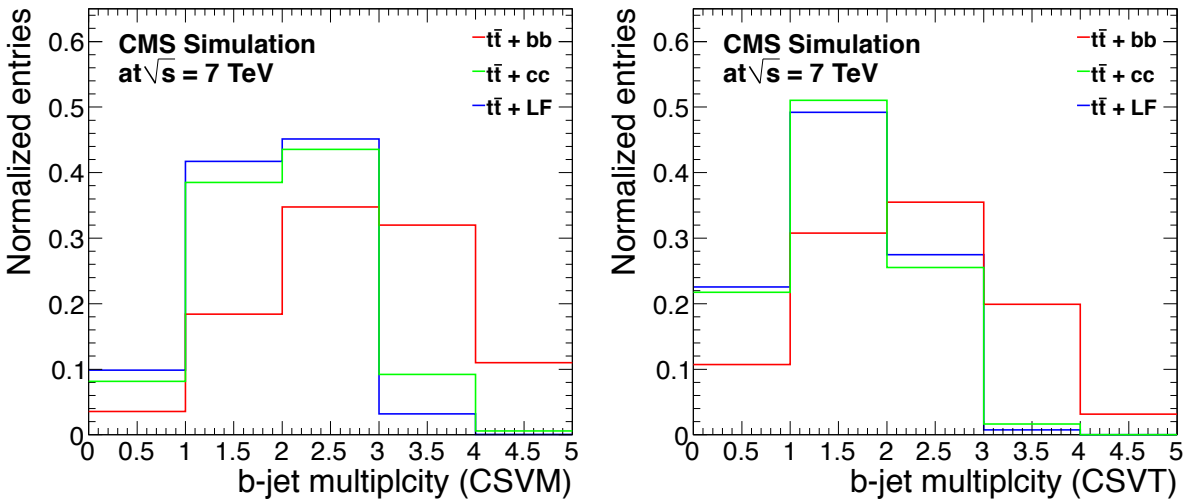

Figure 3: Normalized distribution of the b-tagged jet multiplicity using the "medium" and "tight" working point of the b-tagging algorithm (CSVM and CSVT), after the requirement of at least four reconstructed jets, shown for simulated $t\bar{t}b\bar{b}$, $t\bar{t}c\bar{c}$ and $t\bar{t}LF$ events.

Due to the small statistics in the bins with 3 and 4 b-tagged jets, a likelihood fitting method is employed. Figure 4 shows the b-tag multiplicity distributions where the simulation has been scaled to the fit result. The fit is performed simultaneously in the ee,$\mu\mu$ and $e\mu$ final states and after the requirement of at least two b-tagged jets (S1-S5). The ratio, R at reconstruction level is measured as 7.9 $\pm$ 2.5 % (8.0 $\pm$ 3.3 %) using CSVM (CSVT) b-tagging algorithm.

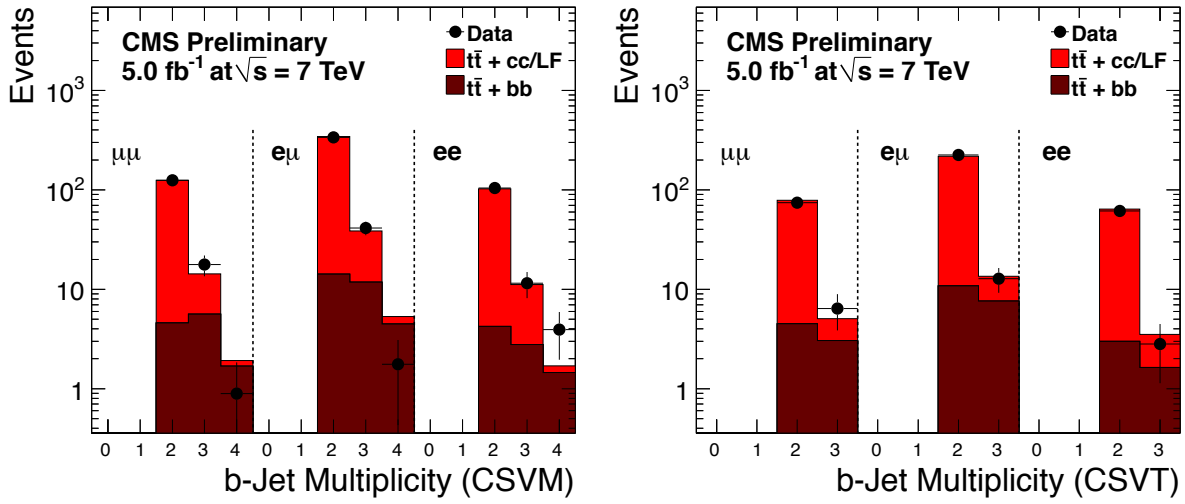

Figure 4: Distribution of the b-tagged jet multiplicity using the "medium" and "tight" working point of the b-tagging algorithm (CSVM and CSVT) for data and simulated samples of $t\bar{t}b\bar{b}$ and $t\bar{t}c\bar{c}$+$t\bar{t}LF$ events, where the normalizations of the latter were obtained from the fit to the data, as described in the text.

7 Correction to Particle Level

The ratio R which is determined by the fitting procedure discussed in the previous section is defined at the reconstruction level. In order to be able to compare with theory predictions or with other experiments, a correction factor to the particle level definition of the cross section ratio within the visible phase space (see Section 2) needs to be applied. The acceptance and efficiency ratio of $t\bar{t}jj$ and $t\bar{t}b\bar{b}$ should be applied to the measured ratio at reconstruction level. Since the ratio is already defined in terms of the visible phase space, no acceptance correction is needed, and the correction to the particle level only needs to consider the ratio of selection efficiencies for $t\bar{t}jj$ and $t\bar{t}b\bar{b}$ events. The efficiency $\epsilon_{t\bar{t}jj}$ is defined as the number of reconstructed $t\bar{t}jj$ events divided by the number of $t\bar{t}jj$ events at the particle level, both defined in the visible phase space. Similarly, the efficiency $\epsilon_{t\bar{t}b\bar{b}}$ is defined. The definitions of $t\bar{t}jj$ and $t\bar{t}b\bar{b}$ events are described in Section 2. The efficiency and the efficiency ratio for each selection step are shown in the Table 3. As shown in the table, the correction factor after the final selection (S5) is found to be 0.46 and 0.37 for CSVM and CSVT selection, respectively. The correction factor is applied to the fraction, R at the reconstruction level in the following way:

$$\sigma(t\bar{t}b\bar{b})/\sigma(t\bar{t}jj) = R \times \epsilon_{t\bar{t}jj}/\epsilon_{t\bar{t}b\bar{b}} \tag{4}$$

8 Estimation of Systematic Uncertainties

Normalization uncertainties, as the one related to the measurement of the luminosity, are expected to cancel in the ratio. The lepton efficiency scale factor is very close to unity, and also its variation as a function of kinematics would not affect the measured b-tagged jet multiplicity distribution. Since the kinematic distributions of jets in $t\bar{t}jj$ and $t\bar{t}b\bar{b}$ are similar, the jet energy scale uncertainty is expected to be relatively small in the measurement of the ratio. The main residual source of systematic uncertainty is expected to arise from the knowledge of the

Table 3: Selection efficiencies for $t\bar{t}jj = t\bar{t}b\bar{b} + t\bar{t}c\bar{c} + t\bar{t}LF$ and $t\bar{t}b\bar{b}$ events at various steps of the event selection (see text), as well as their ratios, shown for both working points of the b-tagging algorithm.

Selection	S1	S2	S3	S4	S5 (CSVM)	S5 (CSVT)
$\epsilon_{t\bar{t}jj}$	71%	63%	59%	30%	15%	8.6%
$\epsilon_{t\bar{t}b\bar{b}}$	73%	64%	60%	40%	32%	23%
$\epsilon_{t\bar{t}jj}/\epsilon_{t\bar{t}b\bar{b}}(\pm 2\%)$	0.98	0.99	0.98	0.74	0.46	0.37

b-tagging efficiency and mistagging rate, which are expected to change the shape of the measured b-jet multiplicity distribution. The dependence of the correction factor to particle level on the assumptions made in the MC simulation is another source of systematic uncertainty. Therefore, we take into account the following systematic uncertainties.

Pileup

The number of pile-up interactions in data is estimated from the measured bunch-by-bunch luminosity times the total inelastic cross section. The systematic uncertainty on the number of pile-up events in simulation is estimated by varying this cross section within $\pm$ 5 %.

Jet energy scale

An uncertainty on the jet energy scale [29] affects the jet selection and can induce a possible bias on the reconstruction of the $E_{\mathrm{T}}^{\mathrm{miss}}$. In order to estimate the effect of the jet energy scale uncertainty on the event selection, the jet energy scale is varied according to its uncertainty, which depends on the jet kinematics, and the corresponding variation is simultaneously propagated to the value of the $E_{\mathrm{T}}^{\mathrm{miss}}$ [29]. The effect on the normalized b-jet multiplicity distribution is of order 2-3 %.

b-tagging scale factor

The systematic uncertainties associated to the b-tagging scale factors for b and light jets are varied separately, taking into account the statistical uncertainty on the MC b-tag efficiencies [26].

The different decay kinematics of b-jets originated from top quark decays and from those originated from other processes (mainly gluon splitting) reflects on a different b-tagging efficiency in both cases. Since Ref. [26] is based on b-jets from top quark decays, an additional systematic error is introduced accounting for this difference. Conservatively, the difference found in simulation between the efficiency in both type of b-jets is taken as systematic error.

MC generator

MADGRAPH or POWHEG have been compared in order to evaluate the impact on the applied correction to particle level.

Q^2 scale uncertainty

The $t\bar{t}b\bar{b}$ cross section strongly depends on the renormalization and factorization scale [9, 10]. However, we expect this dependency to be reduced in the cross section ratio $sttbb/sttjj$. Dedicated simulation samples were used to estimate this uncertainty, where the factorization and renormalization scales Q were varied simultaneously by factors 2 and 0.5.

Table 4: Break-down of the contributions of the various sources of systematic uncertainty considered to the measurement of $\sigma(t\bar{t}b\bar{b})/\sigma(t\bar{t}jj)$ cross section ratio, shown for both working points of the b-tagging algorithm.

Source	$\sigma(t\bar{t}b\bar{b})/\sigma(t\bar{t}jj)$ (%) (CSVM)	$\sigma(t\bar{t}b\bar{b})/\sigma(t\bar{t}jj)$ (%) (CSVT)
Pile-up	0.5	0.5
JES	3.0	2.0
b-tag (heavy flavor)	6.0	4.0
b-tag (light flavor)	+23 -19	+18 -15
MC gen.	3.0	3.0
Q^2 scale	6.0	6.0
Total uncertainty	+25 -21	+20 -17

However, due to the limited statistics of these samples, the quoted systematic uncertainty is probably somewhat over-estimated.

The uncertainties considered in this analysis are summarized in the Table 4. It was found that the mistag, b-tag and Q^2 systematics are the dominating systematic uncertainties.

9 Result

After correcting the acceptance ratio and taking into account the systematic uncertainty, we measure the cross section ratio, $\sigma(t\bar{t}b\bar{b})/\sigma(t\bar{t}jj)$ in the visible phase space using both the medium and tight working points of the CSV b-tagging algorithm. The measured cross section ratio is $3.6 \pm 1.1(\text{stat.}) \pm 0.9(\text{syst.})\%$ for CSVM and $2.9 \pm 1.2(\text{stat.}) \pm 0.6(\text{syst.})\%$ for CSVT. The results from both selections agree with each other within the statistical and systematic uncertainties, given that the correlation between medium and tight working point of the b-tagging algorithm after the requirement of at least two b-tagged jets is not strong. The result with CSVM provides the better precision and is therefore taken as the final result.

$$\sigma(t\bar{t}b\bar{b})/\sigma(t\bar{t}jj) = 3.6 \pm 1.1(\text{stat.}) \pm 0.9(\text{syst.})\% \tag{5}$$

This number can be compared to, and is somewhat larger than the predictions using MADGRAPH and POWHEG, which are 1.2 % and 1.3 %, respectively.

The present result can not be directly compared with NLO QCD calculations, since a further correction of the measured cross section ratio from the particle to the parton level would be required. Predictions at NLO QCD exist for both the $t\bar{t}jj$ and $t\bar{t}b\bar{b}$ cross sections. According to Ref. [4], the cross section $\sigma(t\bar{t}jj)$ is 9.82 pb for $p_T(j) > 50$ GeV/c, $|y(j)| < 2.5$ and $\Delta R(j,j) > 0.5$. Likewise, the authors of Ref. [4] predict the cross section $\sigma(t\bar{t}b\bar{b})$ to be 458.3 fb for $p_T(b) > 20$ GeV/c, $|y(b)| < 2.5$ and $\Delta R(b,b) > 0.7$. The resulting ratio of 4.7 % can however not be compared with this measurement, because of the missing hadron-to-parton correction, as well as the different p_T requirements.

10 Conclusion

We performed the first measurement of the cross section ratio $\sigma(t\bar{t}b\bar{b})/\sigma(t\bar{t}jj)$ in the dilepton decay mode, using a data sample corresponding to an integrated luminosity of 5.0 fb^{-1}. The

cross section ratio of $\sigma(t\bar{t}b\bar{b})/\sigma(t\bar{t}jj)$ is measured in the visible phase space corresponding to the detector acceptance, and corrected to the particle level. The measurement was performed by means of a fit to the measured multiplicity distribution of b-tagged jets in a sample of dileptonic top pair candidate events with at least four reconstructed jets and at least two b-tagged jets. The cross section ratio is $\sigma(t\bar{t}b\bar{b})/\sigma(t\bar{t}jj) = 3.6 \pm 1.1(\text{stat.}) \pm 0.9(\text{syst.})\%$. This is the first measurement of the cross section ratio $\sigma(t\bar{t}b\bar{b})/\sigma(t\bar{t}jj)$.

References

[1] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys. Lett. B* **716** (2012) 1, doi:10.1016/j.physletb.2012.08.020.

[2] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett. B* **716** (2012) 30, doi:10.1016/j.physletb.2012.08.021.

[3] G. Bevilacqua, M. Czakon, C. G. Papadopoulos et al., "Assault on the NLO Wishlist: pp $\to$ tt bb", *JHEP* **0909** (2009) 109, doi:10.1088/1126-6708/2009/09/109, arXiv:0907.4723.

[4] G. Bevilacqua, M. Czakon, C. G. Papadopoulos et al., "Hadronic top-quark pair production in association with two jets at next-to-leading order QCD", *Phys. Rev. D* **84** (2011) 114017, doi:10.1103/PhysRevD.84.114017.

[5] M. Worek, "Next-to-Leading Order QCD Corrections to $t\bar{t}b\bar{b}$ Production at the LHC", *Acta Phys.Polon.B* **40** (2009) 2937, doi:10.1088/1126-6708/2008/08/108.

[6] G. Bevilacqua, M. Czakon, C. G. Papadopoulos et al., "Dominant QCD Backgrounds in Higgs Boson Analyses at the LHC: A Study of $pp \to t\bar{t} + 2$ Jets at Next-to-Leading Order", *Phys. Rev. Lett.* **104** (2010) 162002, doi:10.1103/PhysRevLett.104.162002.

[7] A. Bredenstein, A. Denner, S. Dittmaier et al., "NLO QCD corrections to top anti-top bottom anti-bottom production at the LHC: 2. full hadronic results", *JHEP* **1003** (2010) 021, arXiv:1001.4006.

[8] M. Worek, "On the next-to-leading order QCD K-factor for top anti-top bottom anti-bottom production at the Tevatron", *JHEP* **1202** (2011) 043, doi:10.1007/JHEP02(2012)043, arXiv:1112.4325.

[9] A. Bredenstein, A. Denner, S. Dittmaier et al., "Next-To-Leading Order QCD Corrections to $pp \to t\bar{t}b\bar{b} + X$ at the LHC", *Phys. Rev. Lett.* **103** (2009) 012002, doi:10.1103/PhysRevLett.103.012002.

[10] A. Bredenstein, A. Denner, S. Dittmaier et al., "Next-To-Leading Order QCD Corrections to $pp \to t\bar{t}b\bar{b} + X$ at the LHC", *JHEP* **08** (2008) 108, doi:10.1088/1126-6708/2008/08/108.

[11] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_T jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[12] T. Sjöstrand, S. Mrenna, and P. Skands, "PYTHIA 6.4 physics and manual", *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175.

[13] F. Maltoni and T. Stelzer, "MadEvent: Automatic event generation with MadGraph", *JHEP* **02** (2003) 027, doi:10.1088/1126-6708/2003/02/027, arXiv:hep-ph/0208156.

[14] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method", *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.

[15] N. Kidonakis, "Next-to-next-to-leading soft-gluon corrections for the top quark cross section and transverse momentum distribution", *Phys. Rev.* **D82** (2010) 114030, doi:10.1103/PhysRevD.82.114030, arXiv:1009.4935.

[16] V. Ahrens et al., "Renormalization-Group Improved Predictions for Top-Quark Pair Production at Hadron Colliders", *JHEP* **09** (2010) 097, doi:10.1007/JHEP09(2010)097, arXiv:1003.5827.

[17] M. Aliev et al., "HATHOR: HAdronic Top and Heavy quarks cross section calculatoR.", *Comput. Phys. Commun.* **182** (2011) 1034, doi:10.1016/j.cpc.2010.12.040, arXiv:1007.1327.

[18] N. Davidson et al., "Universal Interface of TAUOLA Technical and Physics Documentation", *Computer Physics Communications* **183, Issue 3** (2010) 821, doi:10.1016/j.cpc.2011.12.009, arXiv:1002.0543.

[19] J. Allison et al., "Geant4 developments and applications", *IEEE Trans. Nucl. Sci.* **53** (2006) 270, doi:10.1109/TNS.2006.869826.

[20] K. Melnikov and F. Petriello, "Electroweak gauge boson production at hadron colliders through $O(\alpha_s^2)$", *Phys. Rev. D* **74** (2006) 114017, doi:10.1103/PhysRevD.74.114017, arXiv:hep-ph/0609070.

[21] N. Kidonakis, "Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^-", *Phys. Rev. D* **82** (2010) 054018, doi:10.1103/PhysRevD.82.054018, arXiv:1005.4451.

[22] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and $E_{\mathrm{T}}^{\mathrm{miss}}$", **PAS PFT-09-001** (2009).

[23] CMS Collaboration, "Particle-flow commissioning with muons and electrons from J/ψ and W events at $\sqrt{s}$ = 7 TeV", **PAS PFT-10-003** (2010).

[24] CMS Collaboration, "Performance of CMS muon reconstruction in pp Collision at $\sqrt{s} = 7$ TeV", arXiv:1206.4071.

[25] CMS Collaboration, "Missing transverse energy performance of the CMS detector", *J. Instrum.* **6** (2011) P09001, doi:doi:10.1088/1748-0221/6/09/P09001.

[26] CMS Collaboration, "*b*-Jet Identification in the CMS Experiment", **PAS BTV-11-004** (2011).

[27] CMS Collaboration, "First Measurement of the Cross Section for Top-Quark Pair Production in Proton-Proton Collisions at $\sqrt{s}$ = 7 TeV", *Phys. Lett. B* **695** (2011) 424, doi:10.1016/j.physletb.2010.11.058, arXiv:1010.5994.

[28] CMS Collaboration, "Measurement of the $\mathrm{t\bar{t}}$ production cross section and the top quark mass in the dilepton channel in pp collisions at $\sqrt{s} = 7$ TeV", *JHEP* **07** (2011) 049, doi:10.1007/JHEP07(2011)049, arXiv:1105.5661.

[29] CMS Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS", *J. Instrum.* **6** (2011) P11002, doi:doi:10.1088/1748-0221/6/11/P11002.