
DRAFT

CMS Physics Analysis Summary

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

2012/06/14

Head Id: 129678

Archive Id: 129658:129678

Archive Date: 2012/06/14

Archive Tag: trunk

Measurement of the top polarization in the dilepton final state

The CMS Collaboration

Abstract

A measurement of the top quark polarization in $t\bar{t} \rightarrow \ell^+\ell^-$ events is performed in a data sample corresponding to a total integrated luminosity of 5.0 fb^{-1} collected by the CMS experiment in pp collisions at a centre-of-mass energy of 7 TeV at the LHC. The measured value in the helicity basis is $P_n = -0.035 \pm 0.028 \pm 0.050$, in agreement with the standard model expectation.

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

PDFAuthor: Kevin Burkett, Oliver Gutsche, Sergo Jindariani, Vyacheslav Krutelyov, Jacob Linacre, Yanjun Tu
PDFTitle: Measurement of the top polarization in the dilepton final state
PDFSubject: CMS
PDFKeywords: CMS, physics, dilepton, top, polarization, polarisation, asymmetry

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

1 Introduction

In the standard model (SM), tops and anti-tops in $t\bar{t}$ pair production are produced unpolarized from QCD. A small polarization is expected in the SM from $t\bar{t}$ production through the Wtb coupling and the tops produced via this mechanism are left handed. In physics beyond SM, couplings of the top to new particles can alter its polarization. Thus net polarization of the tops produced in $t\bar{t}$ production would be a good way to separate the new physics from the SM.

The polarization of the top is reflected in the kinematic distributions of its daughters, because the top decays before hadronization effects can wash away this information. Among all the particles coming from the decay of the top, the charged lepton is most sensitive to the top's polarization [1]. The top polarization P_n along a chosen axis \hat{n} can thus be measured from the angular distribution of the charged leptons from the decays, measured in the top quark rest frame:

$$\frac{1}{\Gamma} \frac{d\Gamma}{d \cos \theta_{l,n}} = \frac{1}{2} (1 + P_n \cos \theta_{l,n}), \quad (1)$$

where $\theta_{l,n}$ is the direction of the lepton with respect to \hat{n} . In this analysis the helicity axis is used, where \hat{n} is given by the direction of the top in the $t\bar{t}$ CM frame.

This note presents a measurement of the top polarization in the helicity basis, using a data sample corresponding to an integrated luminosity of 5.0 fb^{-1} collected at $\sqrt{s} = 7 \text{ TeV}$ by the Compact Muon Solenoid (CMS) experiment at the LHC. A detailed description of the CMS detector can be found elsewhere [2]. Dilepton decays of the $t\bar{t}$ pair are used, and the top polarization is measured using the reconstructed objects. Because the reconstructed polarization is shaped by the reconstruction efficiency and resolution, we apply an unfolding technique to recover the parton-level distribution which can be compared with theoretical predictions.

2 Event samples, reconstruction, and preselection

The data used for this measurement were collected using one of the ee , $e\mu$, or $\mu\mu$ high- p_T double-lepton triggers. Muon candidates are reconstructed using two algorithms that require consistent signals in the tracker and muon systems: one matches the extrapolated trajectories from the silicon tracker to signals in the muon system (tracker-based muons), and the second performs a global fit requiring consistent patterns in the tracker and the muon system (globally fitted muons) [3]. Electron candidates are reconstructed starting from a cluster of energy deposits in the electromagnetic calorimeter. The cluster is then matched to signals in the silicon tracker. A selection using electron identification variables based on shower shape and track-cluster matching is applied to the reconstructed candidates [4]. Electron candidates within $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.1$ from a muon are rejected to remove candidates due to muon bremsstrahlung and final-state radiation. Both electrons and muons are required to be isolated from other activity in the event. This is achieved by imposing a maximum allowed value of 0.15 on the ratio of the scalar sum of track transverse momenta and calorimeter transverse energy deposits within a cone of $\Delta R < 0.3$ around the lepton candidate direction at the origin (the transverse momentum of the candidate is excluded), to the transverse momentum of the candidate.

Event preselection is applied to reject events other than those from $t\bar{t}$ in the dilepton final state. Events are required to have two opposite-sign, isolated leptons (e^+e^- , $e^\pm\mu^\mp$, or $\mu^+\mu^-$). Both leptons must have transverse momentum $p_T > 20 \text{ GeV}/c$, and the electrons (muons) must have

42 $|\eta| < 2.5$ (2.4). The reconstructed lepton trajectories must be consistent with a common inter-
 43 action vertex. In the rare case ($< 0.1\%$) of events with more than two such leptons, the two
 44 leptons with the highest p_T are selected. Events with an e^+e^- or $\mu^+\mu^-$ pair with invariant
 45 mass between 76 and 106 GeV/c^2 or below 12 GeV/c^2 are removed to suppress Drell–Yan (DY)
 46 events ($Z/\gamma^* \rightarrow \ell^+\ell^-$) as well as low mass dilepton resonances. The jets and the missing trans-
 47 verse energy E_T^{miss} are reconstructed with a particle-flow technique [5]. The anti- k_T clustering
 48 algorithm [6] with a distance parameter of 0.5 is used for jet clustering. At least two jets with
 49 $p_T > 30 \text{ GeV}/c$ and $|\eta| < 2.5$, separated by $\Delta R > 0.4$ from leptons passing the analysis selection,
 50 are required in each event. At least one of these jets is required to be consistent with coming
 51 from the decay of heavy flavor hadrons and be identified as a b jet by the CSVM b-tagging
 52 algorithm [7], which relies on tracks with large impact parameters. The E_T^{miss} in the event is
 53 required to exceed 30 GeV , consistent with the presence of two undetected neutrinos.

54 Signal and background events are generated using the MADGRAPH 4.4.12 [8] and PYTHIA
 55 6.4.22 [9] event generators. For $t\bar{t}$ events, POWHEG with PYTHIA is used for the $t\bar{t} \rightarrow \ell^+\ell^-$
 56 component, while the remainder, denoted $t\bar{t} \rightarrow \text{other}$ is generated using MADGRAPH. The
 57 samples of DY with $M_{\ell\ell} > 50 \text{ GeV}/c^2$, diboson (WW, WZ, and ZZ only: the contribution from
 58 $W\gamma$ is assumed to be negligible), and single top quark events are generated using MADGRAPH.
 59 The DY event samples with $M_{\ell\ell} < 50 \text{ GeV}/c^2$ are generated using PYTHIA.

60 Events are then simulated using a GEANT4-based model [10] of the CMS detector, and finally
 61 reconstructed and analyzed with the same software used to process collision data. The cross
 62 section for $t\bar{t}$ production is taken from a recent CMS measurement [11], while next-to-leading
 63 order (NLO) cross sections are used for the remaining SM background samples.

64 With the steadily increasing LHC instantaneous luminosity, the mean number of interactions
 65 in a single bunch crossing also increased over the course of data taking, reaching about 15
 66 at the end of the 2011 running period. In the following, the yields of simulated events are
 67 weighted such that the distribution of reconstructed vertices observed in data is reproduced.
 68 The efficiency for events containing two leptons satisfying the analysis selection to pass at least
 69 one of the double-lepton triggers is measured to be approximately 100%, 95%, and 90% for
 70 the ee , $e\mu$, and $\mu\mu$ triggers, respectively [12], and corresponding weights are applied to the
 71 simulated event yields. In addition, b-tagging scale factors are applied to simulated events for
 72 each jet, to account for the difference between b-tagging efficiencies in data and simulation [7].

73 3 Preselection yields and top polarization at reconstruction level

74 The observed and simulated yields after the event preselection are listed in Table 1, in which
 75 the categories $t\bar{t} \rightarrow \ell^+\ell^-$ and $\text{DY} \rightarrow \ell^+\ell^-$ correspond to dileptonic $t\bar{t}$ and DY decays, including
 76 τ leptons only when they also decay leptonically. All other $t\bar{t}$ decay modes are included in the
 77 category $t\bar{t} \rightarrow \text{other}$. The yields are dominated (92%) by top-pair production in the dilepton
 78 final state, with the largest background coming from single top production. The $t\bar{t} \rightarrow \ell^+\ell^-$
 79 yields are normalized such that the total simulated yield matches the data. Comparisons be-
 80 tween data and the simulation for the number of vertices and the number of b tagged jets are
 81 shown in Figure 1.

From Equation 1, the top polarization can be extracted from

$$P_n = \frac{N(\cos(\theta_l^+) > 0) - N(\cos(\theta_l^+) < 0)}{N(\cos(\theta_l^+) > 0) + N(\cos(\theta_l^+) < 0)},$$

82 where θ_l^+ is the production angle of the positively charged lepton in the rest frame of its parent

Table 1: The observed and simulated yields after the preselection described in the text. Uncertainties are statistical only. The systematic uncertainties on the simulated yields are given in Section 6. Where the simulated yields are zero, upper limits are given based on the weighted yield, had one of the simulated events passed the preselection.

Sample	ee	$\mu\mu$	$e\mu$	all
$t\bar{t} \rightarrow \ell^+\ell^-$	1791.7 ± 4.4	2127.3 ± 4.7	5069.4 ± 7.3	8988.5 ± 9.7
$t\bar{t} \rightarrow \text{other}$	32.5 ± 2.9	4.8 ± 1.1	53.3 ± 3.6	90.7 ± 4.8
W + jets	< 1.9	4.7 ± 3.3	4.7 ± 3.4	9.4 ± 4.7
DY \rightarrow ee	52.3 ± 5.8	< 0.6	< 0.6	52.3 ± 5.8
DY \rightarrow $\mu\mu$	< 0.6	72.8 ± 6.5	1.6 ± 0.9	74.4 ± 6.5
DY \rightarrow $\tau\tau$	17.6 ± 3.3	8.7 ± 2.2	18.7 ± 3.2	45.0 ± 5.1
Di-boson	10.6 ± 0.5	13.0 ± 0.5	24.0 ± 0.7	47.6 ± 1.0
Single top	84.9 ± 2.3	101.2 ± 2.4	252.1 ± 3.9	438.2 ± 5.1
Total (simulation)	1989.6 ± 8.8	2332.6 ± 9.3	5423.8 ± 10.3	9746.0 ± 16.4
Data	1961	2373	5412	9746

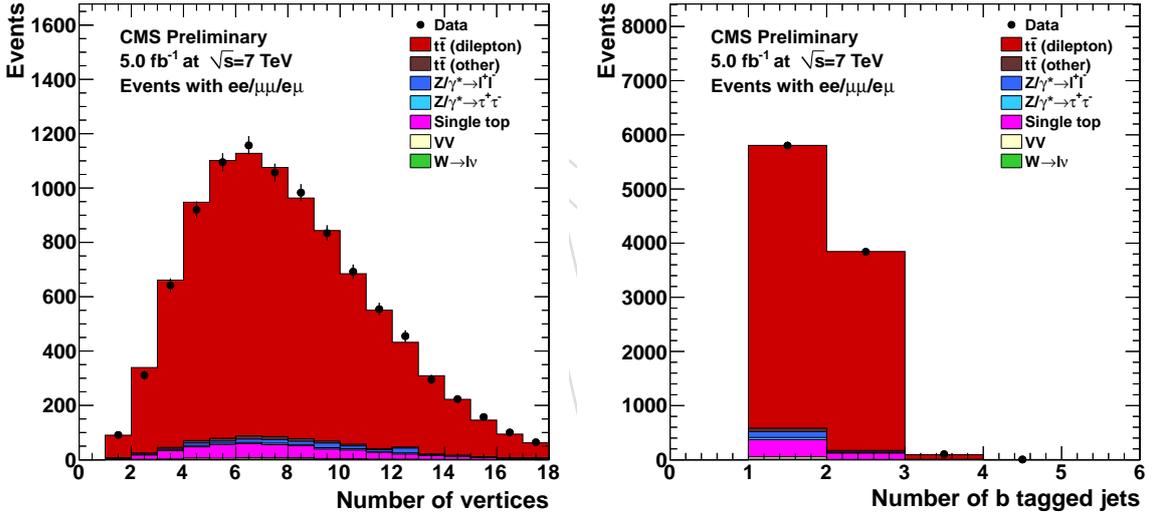


Figure 1: Comparison of data and simulation for number of vertices (left) and number of b tagged jets (right).

83 top, with respect to the direction of the parent top in the $t\bar{t}$ rest frame. Equivalently, P_n can be
 84 measured using θ_l^- , but in this analysis θ_l^+ is always used.

85 Measurement of this observable requires reconstruction of the $t\bar{t}$ system. Each event has two
 86 neutrinos, and there is also ambiguity in combining b jets with leptons. The analytical matrix
 87 weighting technique [13] is used to find a probable solution. Each event is reconstructed using
 88 a range of possible M_t values from 100-300 GeV/ c^2 in 1 GeV/ c^2 steps. The M_t hypothesis with
 89 the maximum averaged weight over possible solutions is taken, and the $t\bar{t}$ kinematics are then
 90 taken from the solution with largest weight. Approximately 17% of events have no solution,
 91 and are thus not used in the measurement of the top polarization.

92 A comparison of the $\cos(\theta_l^+)$ distributions in data and simulation is shown in Figure 2. The

93 resulting value of P_n at reconstruction level is 0.083 ± 0.011 in data and 0.103 ± 0.002 in the
 94 simulation, where the uncertainties are statistical only. The remainder of this note focuses on
 95 making the necessary corrections to obtain P_n at parton-level, due to the presence of back-
 96 ground events and resolution and acceptance effects.

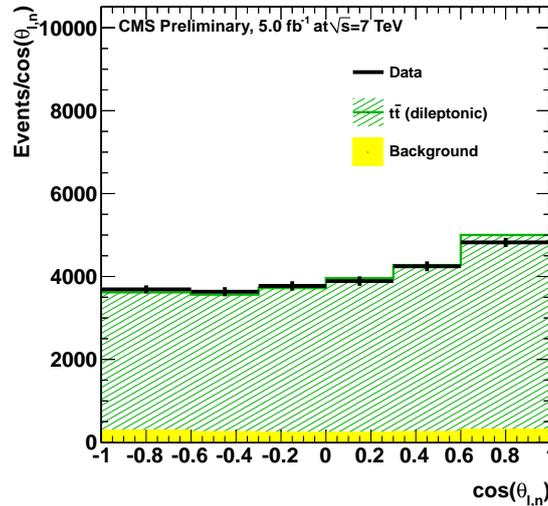


Figure 2: The reconstruction level $\cos(\theta_{l,n}^+)$ distribution in the preselection region, for data and the simulation. The simulated events are divided into $t\bar{t} \rightarrow \ell^+\ell^-$ and background, where the background consists of the categories other than $t\bar{t} \rightarrow \ell^+\ell^-$ in Table 1.

97 4 Background estimation

98 The simulation is used to predict the background event yields and shapes. We use methods
 99 based on data to cross-check these estimates for the background contributions from events with
 100 misidentified leptons and from $DY \rightarrow ee/\mu\mu$ events. However, the dependence of the measured
 101 top polarization on the background normalization is small, and in Section 6 the systematic
 102 uncertainty is estimated based on changing the normalization of each component by up to
 103 100%.

104 A misidentified lepton is defined as a lepton candidate not originating from a prompt decay,
 105 such as a lepton from semileptonic b or c decays, a muon from a pion or kaon decay, an uniden-
 106 tified photon conversion, or a pion misidentified as an electron. The background from events
 107 with misidentified leptons is predicted based on the number of events in data with a candidate
 108 lepton that can pass only loosened selection criteria [14]. Using a measurement of the fraction
 109 of such “loose” leptons that go on to pass the selection requirements, the number of misiden-
 110 tified leptons in the event sample can be estimated. The resulting prediction is 138^{+281}_{-138} events,
 111 including both statistical and systematic uncertainties. The simulated yield is 100.1 ± 6.7 , in
 112 reasonable agreement.

113 The estimation method for $DY \rightarrow ee/\mu\mu$ events [15] is based on counting the number of Z
 114 candidates in the Z veto region (after subtracting the number of non Z events estimated using
 115 the number of $e\mu$ events), and multiplying this number by the ratio of simulated DY yields
 116 outside to inside the Z veto region. The result for the preselection region is an estimate of
 117 142.4 ± 15.0 $DY \rightarrow ee/\mu\mu$ events. The result is thus consistent with the simulated prediction of
 118 126.7 ± 8.7 from Table 1.

5 Unfolding

The measured distribution is distorted from the true underlying distribution by the limited acceptance of our detector and the finite resolution of the measurement. The unfolding procedure allows to correct the data for these effects, yielding the “parton-level” $\cos(\theta_l^+)$ distribution and top polarization. This distribution represents the differential cross-section in $\cos(\theta_l^+)$, and is normalized to unity.

Reconstruction and identification requirements and the kinematic fitter are known to smear out the true kinematics of reconstructed leptons and top quarks. In addition to these smearing effects, the true distribution is also modified by the event selection. If the selection is biased with respect to $\cos(\theta_l^+)$, such bias would cause a change in the observed polarization.

In general, the background-subtracted measured distribution \vec{b} is related to the underlying parton-level distribution \vec{x} by the matrix equation $\vec{b} = SA\vec{x}$, where A is a diagonal matrix describing the acceptance in each bin of the measured distribution, and S is a non-diagonal smearing matrix describing the migration of events between bins due to the detector resolution and reconstruction techniques.

Choice of a binning scheme for the distribution is motivated by the following considerations. Very fine binning would result in large bin-to-bin oscillations caused by statistical fluctuations, while having very few bins is sub-optimal due to reduced information about the smearing. Based on unfolding studies, we find that for our level of statistics, the use of six bins is optimal. The bin size is variable and is chosen to ensure similar level of statistics in each bin of the distribution. A summary of the binning is provided in Table 2.

Table 2: Binning used in the distributions of $\cos(\theta_l^+)$.

B1	B2	B3	B4	B5	B6
[-1.0,-0.6]	[-0.6,-0.3]	[-0.3,-0.0]	[0.0, 0.3]	[0.3, 0.6]	[0.6, 1.0]

The A and S matrices are modeled using the NLO POWHEG-PYTHIA $t\bar{t}$ sample, and are shown in Fig. 3. The smearing effects are quite large due to the uncertainties of top reconstruction. However, most of the large values lie close to the diagonal, meaning there is little extreme smearing between far-apart bins. The distribution is roughly symmetric around the diagonal, indicating that the smearing does not generate artificial asymmetry in reconstructed data, but rather dilutes any existing asymmetry in the true distribution.

We employ a regularized ‘unfolding’ algorithm based on singular value decomposition (SVD) [16], which is implemented in the RooUnfold package. Effects of large statistical fluctuations in the algorithm are greatly reduced by introducing a regularization term to the unfolding procedure. Regularization strength is defined by the parameter k , which in this measurement is equal to 3. This choice of k was motivated by the effort to maintain balance between the statistical uncertainty of the method and the size of the bias introduced by the unfolding procedure. This is a conservative choice leading to a slightly larger statistical uncertainty compared to choosing a smaller k value, but also reduces the degree to which the corrections tend to bias the result to the response model. The full covariance matrix is used in the evaluation of the statistical uncertainty of the measured asymmetry.

We verify that the unfolding procedure is able to correctly unfold distributions with different levels of asymmetry. In order to do this, we re-weight generated $t\bar{t}$ events according to a linear function of $\cos(\theta_l^+)$: $\text{weight}=1+K \cos(\theta_l^+)$. The parameter K is varied between -0.5 and 0.5 in steps of 0.2, introducing a polarization of up to 40%, far more than is expected in $t\bar{t}$ events. For

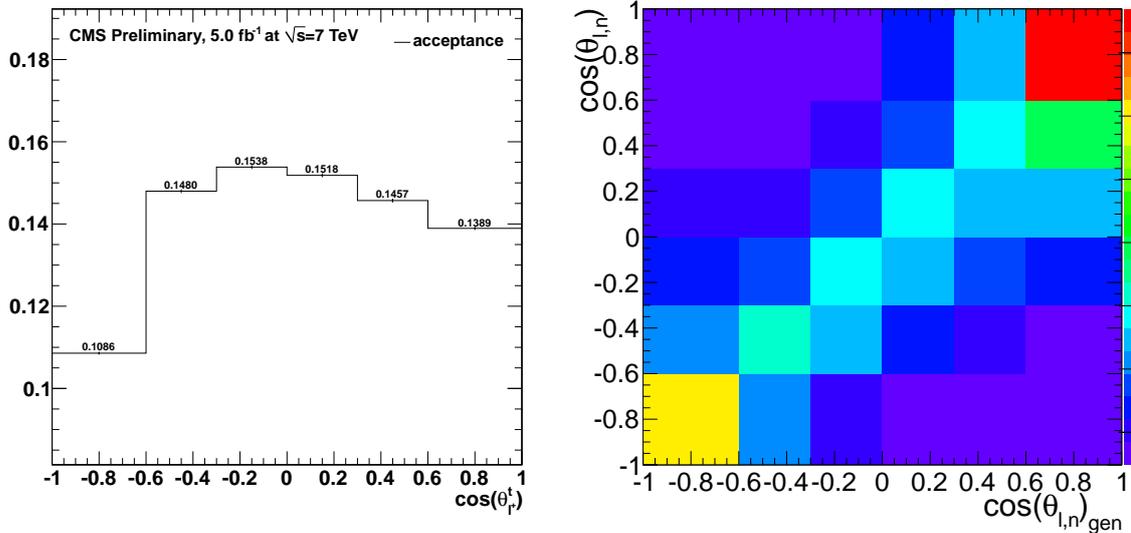


Figure 3: Acceptance matrix bins (left) and smearing effects (right).

160 each value of K , we generate 2000 pseudo-experiments, in which the number of events in each
 161 bin of the distribution is fluctuated according to Poisson statistics, and then the distribution is
 162 unfolded. The average value of the asymmetry in 2000 pseudo-experiments is compared to the
 163 original true-level value. We find a linear behavior of this distribution, suggesting that non-SM
 164 asymmetry values will also be measured correctly. The offsets and slopes obtained in the linear
 165 function fit are -0.004 ± 0.009 and 1.031 ± 0.053 , respectively. We also look at the distribution
 166 of the pulls in the set of pseudo-experiments and fit it to gaussian function. We find a small bias
 167 leading to asymmetry changes of up to 1%, and assign it as an additional systematic uncertainty
 168 associated with unfolding bias. The width of the gaussian function obtained in the fit is 0.9,
 169 indicating that we slightly over-estimate the measured statistical uncertainty. No correction for
 170 this is made.

171 6 Systematic uncertainties

172 The systematic uncertainty associated with the jet and E_T^{miss} energy scale can directly affect the
 173 shape of the asymmetry distributions. We evaluate this uncertainty, assuming a 7.5% uncer-
 174 tainty to the hadronic energy scale after jet corrections have been applied [17].

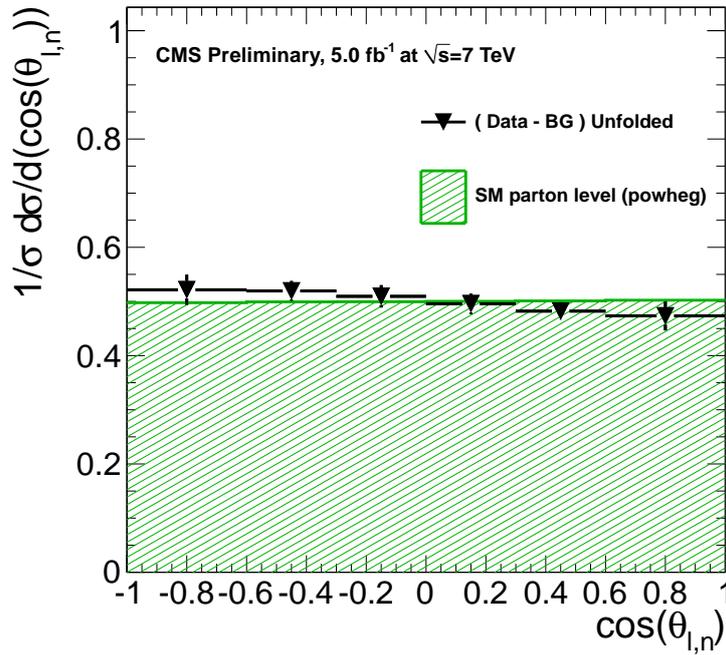
175 There are also a number of systematic uncertainties due to the background subtraction and
 176 unfolding used to produce the parton level measurement. Given the uncertainties of the data
 177 driven background estimates described in Section 4, the uncertainty associated with the back-
 178 ground estimate is estimated by varying the backgrounds from DY and misidentified leptons
 179 by 100%. In addition, we vary the single top background by 50%. The systematic uncer-
 180 tainty from the $t\bar{t}$ modeling is estimated by applying unfolding derived using MC@NLO $t\bar{t}$
 181 to POWHEG-PYTHIA events, and taking the difference in the result compared to that for the
 182 POWHEG-PYTHIA derived unfolding. We also assess the systematics due to the shower match-
 183 ing, the top mass, the b -tagging, the trigger efficiency, the lepton ID efficiency and the pile-up
 184 reweighting. There is a systematic uncertainty of 0.9% to account for the unfolding bias. The
 185 systematic uncertainties on the unfolded P_n measurement are summarized in Table 3, combin-
 186 ing to a total systematic uncertainty of 0.050.

Table 3: Systematic uncertainties.

	JES	BG	modeling	unfolding	top mass	b -tagging	Trigg(lep ID)	PU	Total
P_n	0.043	0.009	0.014	0.009	0.019	0.001	0.000	0.002	0.050

187 7 Results and summary

188 The background-subtracted and unfolded $\cos(\theta_l^+)$ distribution for $t\bar{t} \rightarrow \ell^+\ell^-$ events is shown
 189 in Figure 4. The measured value of P_n at parton-level is $-0.035 \pm 0.028 \pm 0.050$, while the $t\bar{t}$
 190 parton level prediction obtained from POWHEG-PYTHIA simulation is 0.003 ± 0.0004 .

Figure 4: Background-subtracted and unfolded $\cos(\theta_l^+)$ distribution.

191 The measured value of P_n is consistent with the POWHEG-PYTHIA value, and we thus observe
 192 no significant deviation from the SM expectation.

193 8 Acknowledgements

194 The authors would like to thank Dr. Tao Liu for fruitful discussions and helpful correspondence
 195 on the theory part. The authors are also very grateful to T. Schwarz and D. Mielicki for fruitful
 196 discussions.

197 References

- 198 [1] D. Krohn, T. Liu, J. Shelton et al., "A Polarized View of the Top Asymmetry", *Phys.Rev.*
 199 **D84** (2011) 074034, doi:10.1103/PhysRevD.84.074034, arXiv:1105.3743.
- 200 [2] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004,
 201 doi:10.1088/1748-0221/3/08/S08004.

- 202 [3] CMS Collaboration, "Performance of muon identification in pp collisions at $\sqrt{s} = 7$ TeV",
203 CMS Physics Analysis Summary CMS-PAS-MUO-10-002, (2010).
- 204 [4] CMS Collaboration, "Electron Reconstruction and Identification at $\sqrt{s} = 7$ TeV", CMS
205 Physics Analysis Summary CMS-PAS-EGM-10-004, (2010).
- 206 [5] CMS Collaboration, "Commissioning of the Particle-Flow Reconstruction in
207 Minimum-Bias and Jet Events from pp Collisions at 7 TeV", CMS Physics Analysis
208 Summary CMS-PAS-PFT-10-002, (2010).
- 209 [6] M. Cacciari, G. Salam, and G. Soyez, "The anti- k_T jet clustering algorithm", *JHEP* **04**
210 (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.
- 211 [7] CMS Collaboration, "Measurement of the b-tagging efficiency using $t\bar{t}$ events", CMS
212 Physics Analysis Summary CMS-PAS-BTV-11-003, (2011).
- 213 [8] J. Alwall, P. Demin, S. de Visscher et al., "MadGraph/MadEvent v4: the new web
214 generation", *JHEP* **09** (2007) 028, doi:10.1088/1126-6708/2007/09/028,
215 arXiv:0706.2334.
- 216 [9] T. Sjöstrand, S. Mrenna, and P. Skands, "PYTHIA 6.4 physics and manual", *JHEP* **05**
217 (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:0706.2334.
- 218 [10] S. Agostinelli et al., "GEANT4 – a simulation toolkit", *Nucl. Instr. and Meth. A* **506** (2003)
219 250, doi:10.1016/S0168-9002(03)01368-8.
- 220 [11] CMS Collaboration, "Measurement of the $t\bar{t}$ production cross section in pp collisions at
221 7 TeV in lepton + jets events using b-quark jet identification", *Phys. Rev. D* **84** (2011)
222 092004, doi:10.1103/PhysRevD.84.092004, arXiv:1108.3773.
- 223 [12] CMS Collaboration, "Search for heavy, top-like quark pair production in the dilepton
224 final state in pp collisions at $\sqrt{s} = 7$ TeV", *Submitted to PLB* (2012) arXiv:1203.5410.
- 225 [13] CMS Collaboration Collaboration, "Measurement of the t t-bar production cross section
226 and the top quark mass in the dilepton channel in pp collisions at sqrt(s) =7 TeV", *JHEP*
227 **1107** (2011) 049, arXiv:1105.5661.
- 228 [14] CMS Collaboration, "Search for new physics with same-sign isolated dilepton events
229 with jets and missing transverse energy at the LHC", *JHEP* **6** (2011) 077,
230 doi:10.1007/JHEP06(2011)077, arXiv:1104.3168.
- 231 [15] CMS Collaboration Collaboration, "First Measurement of the Cross Section for
232 Top-Quark Pair Production in Proton-Proton Collisions at sqrt(s)=7 TeV", *Phys.Lett.*
233 **B695** (2011) 424–443, arXiv:1010.5994.
- 234 [16] A. Hocker and V. Kartvelishvili, "SVD approach to data unfolding", *Nucl.Instrum.Meth.*
235 **A372** (1996) 469–481, doi:10.1016/0168-9002(95)01478-0,
236 arXiv:hep-ph/9509307.
- 237 [17] CMS Collaboration, "Determination of the Jet Energy Scale in CMS with pp Collisions at
238 $\sqrt{s} = 7$ TeV", CMS Physics Analysis Summary CMS-PAS-JME-10-010, (2010).