CMS Physics Analysis Summary

Contact: cms-pag-conveners-top@cern.ch 2012/03/01

Differential measurements of the charge asymmetry in top quark pair prodcution

The CMS Collaboration

Abstract

The difference in angular distributions between top quarks and antiquarks, commonly referred to as the charge asymmetry, is measured in a dataset corresponding to an integrated luminosity of $4.7\,\text{fb}^{-1}$, collected with the CMS detector at the LHC. In addition to the inclusive measurement, for the first time at the LHC also differential measurements of the charge asymmetry are performed. The invariant mass, the rapidity, and the transverse momentum of the $\mathrm{t\bar{t}}$ system are sensitive to the different processes contributing to the overall charge asymmetry and are therefore chosen as differentiating variables. The measured inclusive asymmetry of $A_C = 0.004 \pm 0.010(\text{stat.}) \pm 0.012\ (\text{syst.})$ and the measured differential asymmetries are consistent with the predictions from the standard model.

1 Introduction

The field of charge asymmetry in top quark pair production has attracted large attention in the recent past due to the published measurements of the CDF [1] and D0 [2] Collaborations at the Tevatron. The reported deviations from the standard model (SM) prediction on the order of two standard deviations, and even more in certain phase space regions, have triggered a large number of theoretical explanations that account these discrepancies to contributions from physics beyond the standard model.

The charge asymmetry occurs only in quark antiquark initial states. Since at the LHC the quarks in the initial state are mainly valence quarks, while the antiquarks are always sea quarks, the larger average momentum fraction of quarks leads to an excess of top quarks produced in the forward directions. This makes the difference of the rapidities of top quark and antiquark, $\Delta|y| = |y_t| - |y_{\bar{t}}|$, a suitable observable to measure the $t\bar{t}$ charge asymmetry, defined as

$$A_C = \frac{N^+ - N^-}{N^+ + N^-}\,, \tag{1}$$

where N^+ and N^- represent the number of events with positive and negative values in the sensitive variable, respectively.

Recently, CMS has published a first measurement of the charge asymmetry at the LHC and found $A_C = -0.013 \pm 0.028\ (\text{stat.})\ ^{+0.029}_{-0.031}\ (\text{syst.})$ [3], consistent with the SM predictions of $A_C(\mathit{theory}) = 0.0115 \pm 0.0006$ [4]. To shed light on the observed tension between the Tevatron results and the LHC results it is crucial to not only measure the inclusive asymmetry but to also measure the asymmetry as a function of variables that are suited to enhance the charge asymmetry in certain regions. We report on an update of the published CMS analysis and on the first differential measurements of the charge asymmetry in top quark pair production at the LHC, using the full 2011 dataset. We measure the charge asymmetry as a function of the rapidity, the transverse momentum, and the invariant mass of the $t\bar{t}$ system. Each of these variables is sensitive to a certain aspect of the $t\bar{t}$ charge asymmetry.

The rapidity of the summed four-vectors of top quark and antiquark in the laboratory frame, $|y_{t\bar{t}}|$, is sensitive to the ratio of the contributions from the $q\bar{q}$ and gluon-gluon initial state to top quark production. The gluon fusion process is dominant in the central region, while $t\bar{t}$ production through $q\bar{q}$ annihilation is more prominent for events with the $t\bar{t}$ pair at larger rapidities, which implies an enhancement of the charge asymmetry with increasing $|y_{t\bar{t}}|$.

The transverse momentum of the summed four-vectors of the $t\bar{t}$ pair in the laboratory frame, $p_{T,t\bar{t}}$, is sensitive to the ratio of the positive and negative contributions to the overall asymmetry. The interference between the Born and the box diagram leads to a positive contribution and the interference between initial and final state radiation (ISR and FSR) results in a negative contribution. The presence of additional hard radiation (ISR or FSR) implies on average higher transverse momenta of the $t\bar{t}$ system, $p_{T,t\bar{t}}$. Consequently, in events with large values of $p_{T,t\bar{t}}$ the negative contribution from the ISR-FSR interference is enhanced.

If heavy particles with different couplings to top quarks and antiquarks exist, one can expect an effect on the charge asymmetry with rising invariant mass of the $t\bar{t}$ system, $m_{t\bar{t}}$. But even in pure SM $t\bar{t}$ production, the charge asymmetry is expected to depend on $m_{t\bar{t}}$, since the contribution of $q\bar{q}$ initial state processes is enhanced for larger values of $m_{t\bar{t}}$.

2 Data and simulation

This analysis of $t\bar{t}$ events produced in proton-proton collisions at a centre-of-mass energy of 7 TeV is based on data taken with the CMS detector, corresponding to an integrated luminosity of 4.7 fb^{-1}. To translate the distributions measured with reconstructed objects to distributions for the underlying quarks, we use simulated data samples. Top-quark pair events are generated with two different generators, either with MADGRAPH version 5 [5] or withe next-to-leading order (NLO) generator POWHEG [6], where in both cases the parton shower is simulated using PYTHIA version 6.4 [7] and the MLM matching algorithm [8]. Also the t and tW channels of electroweak production of single top quarks are simulated using POWHEG. The production of weak vector bosons in association with jets (W+jets and Z+jets) is simulated using the same combination of MADGRAPH and PYTHIA as for the $t\bar{t}$ signal. Additional proton-proton interactions (pile-up) are overlaid on the simulated events as observed in the analyzed data.

3 Event selection and estimation of background

In this measurement we focus on $t\bar{t}$ events, where one of the W bosons from the decay of a top-quark pair subsequently decays into a muon or electron and the corresponding neutrino, and the other W boson decays into a pair of jets. We therefore select events containing one electron or muon and four or more jets, at least one of which is identified as originating from the hadronization of a b quark. For the reconstruction of electrons, muons, jets, and any imbalance in transverse momentum due to the neutrino, we use a particle-flow (PF) algorithm [9]. A detailed description of the applied selection criteria can be found in Ref. [3].

We find a total of 57 697 events, 24 705 in the electron+jets channel and 32 992 in the muon+jets channel. For the estimation of the background contributions we make use of the discriminating power of the imbalance in transverse momentum in an event, $E_{\mathrm{T}}^{\mathrm{miss}}$, and of M3, the invariant mass of the combination of three jets that corresponds to the largest vectorially summed transverse momentum, as described in detail in Ref. [3]. For the W +jets, Z+jets, and single-top-quark background processes, the respective simulated samples are used to model the shapes of the $E_{\mathrm{T}}^{\mathrm{miss}}$ and M3 distributions, while an approach based on data is pursued for the multijet background. Table 1 summarizes the results of the fits, along with their statistical uncertainties. Figure 1 shows the measured $E_{\mathrm{T}}^{\mathrm{miss}}$ and M3 distributions, with the individual simulated contributions normalized to the results from the fit.

Table 1: Results for the numbers of events for background and $t\bar{t}$ contributions from fits to data, along with their statistical uncertainties. The uncertainties quoted for the single-top-quark and Z+jets backgrounds are related to the constraints used as input for the likelihood fit.

process	electron+jets	muon+jets	total
single top (t + tW)	1054 ± 319	1358 ± 478	2412 ± 604
W+jets (+)	1839 ± 224	1832 ± 284	3671 ± 362
W+jets (−)	1469 ± 222	1342 ± 270	2811 ± 349
Z+jets	504 ± 145	566 ± 160	1070 ± 216
multijet	1169 ± 221	887 ± 204	2056 ± 301
total BG	6035 ± 521	5985 ± 670	12020 ± 849
$t\bar{t}$	18661 ± 386	26998 ± 464	45659 ± 604
observed data	24705	32992	57697

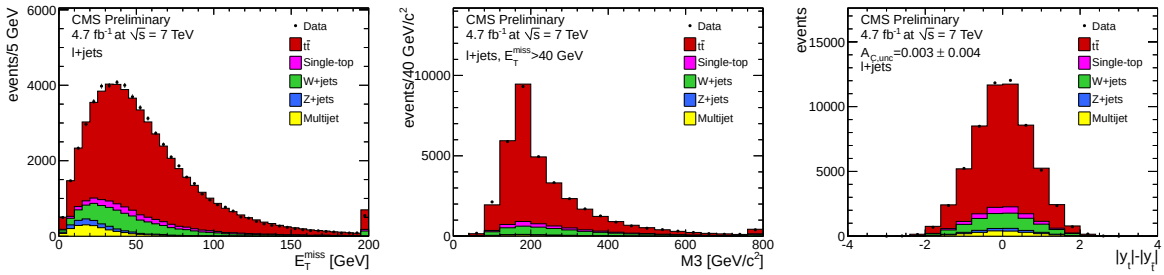

Figure 1: Comparison of the combined lepton+jets data with simulated contributions for the distributions in $E_{\mathrm{T}}^{\mathrm{miss}}$ (left) and M3 (middle), and $\Delta|y|$ (right). The simulated signal and background contributions are normalized to the results of the fits in Table 1.

4 Measurement of the $\mathrm{t\bar{t}}$ charge asymmetry

The measurement of the $\mathrm{t\bar{t}}$ charge asymmetry is based on the fully reconstructed four-momenta of t and $\bar{\mathrm{t}}$ in each event. We reconstruct the leptonically decaying W boson from the measured charged lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$ and associate the measured jets in the event with quarks in the $\mathrm{t\bar{t}}$ decay chain. For a detailed description of the reconstruction procedure see Ref. [3].

The reconstructed top quark and antiquark four-vectors are used to obtain the inclusive (see Fig. 1 (right)) and differential distributions of $\Delta|y|$ and the charge asymmetry is calculated by counting the entries with $\Delta|y| > 0$ and the entries with $\Delta|y| < 0$ and inserting these numbers into Eq. 1. In case of the differential measurements the asymmetries are calculated separately for the different bins in the differentiating variable V_d (where V_d is either $|y_{\mathrm{t\bar{t}}}|$, $p_{\mathrm{T,t\bar{t}}}$, or $m_{\mathrm{t\bar{t}}}$). Prior to this calculation, the reconstructed distributions have to be corrected for several effects, to be able to compare the resulting asymmetry with the predictions from theory on particle level. In a tiered procedure the measured distributions are corrected for background contributions, reconstruction effects, and selection efficiencies.

In the first correction step, the distributions of background processes are normalized to the estimated rates (see table 1) and subtracted from the data, assuming Gaussian uncertainties on the background rates as well as on statistical fluctuations in the background templates. The correlations among the individual background rates are taken into account.

The remaining background-free distributions are translated from the reconstruction level to the particle level after event selection, and from there to the particle level before event selection. The corrected distributions are then independent from the detector environment and analysis specifications. The corrections are achieved by applying a regularized unfolding procedure to the data [10] through a generalized matrix-inversion method. In this method, the disturbing effects are described by a smearing matrix S that translates the true spectrum $\vec{x}$ into the measured spectrum $\vec{w} = S\vec{x}$. The technical details on the applied unfolding procedure can be found in Ref. [3]. As reconstruction and selection effects factorize, the smearing matrix S can be calculated as the product of a migration matrix and a diagonal matrix with the efficiencies for each of the bins on the diagonal, and all other elements set to zero.

The number of bins and especially the bin ranges used for $\Delta|y|$ and the differentiating variables V_d has to be chosen with care. To stabilize the unfolding procedure it is desirable that the number of entries in each bin of the reconstructed distributions as well as in the unfolded distributions is approximately equal. For technical reasons the number of bins in the reconstructed spectra must be twice as high as the number of bins used for the unfolded spectra. We use 16 (8) bins for the reconstructed (unfolded) $\Delta|y|$ distribution and 6 (3) bins in the reconstructed (unfolded) V_d distributions. The ranges for the bins in the unfolded differentiating variables

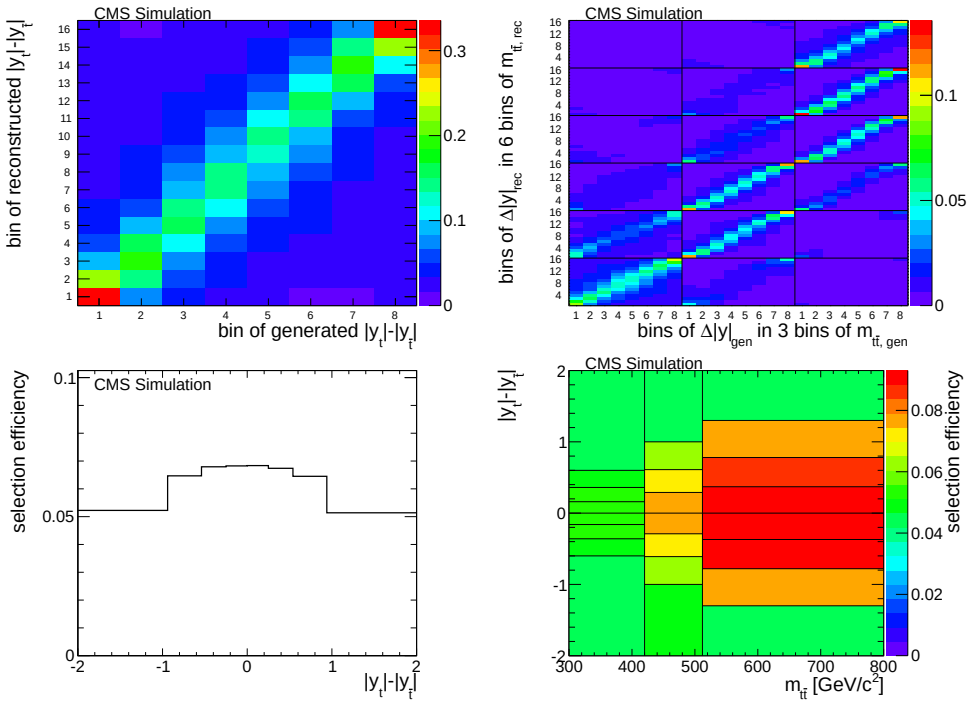

Figure 2: Migration matrix (upper row) between the generated and the reconstructed values in $\Delta|y|$, after the event selection (left) and for the measurement differential in $m_{t\bar{t}}$ (right). Selection efficiency (lower row) as a function of generated $\Delta|y|$, defined with respect to inclusive $t\bar{t}$ production (left) and for the measurement differential in $m_{t\bar{t}}$ (right)

are [0 – 0.41; 0.41 – 0.90; 0.90 – ∞] for $|y_{t\bar{t}}|$, [0 – 23; 23 – 58; 58 – ∞] for $p_{T,t\bar{t}}$ in GeV/c, and [0 – 420; 420 – 512; 512 – ∞] for $m_{t\bar{t}}$ in GeV/c^2.

We use separate migration matrices for the inclusive measurement and the three differential measurements, obtained from simulated $t\bar{t}$ events. Figure 2 shows the migration matrices for the inclusive measurement and for the differential measurement in $m_{t\bar{t}}$. While for the inclusive measurement the migration matrix describes the migration of selected events from true values of $\Delta|y|$ to different reconstructed values, for the migration matrices for the differential measurements not only the migration between bins of $\Delta|y|$ has to be taken into account, but also the migration between bins of V_d. The migration matrices for the differential measurements feature on large scale a grid of 6×3 bins in V_d with each of these bins hosting a 16×8 migration matrix describing the migration between different $\Delta|y|$ values. The values of $\Delta|y|$ and V_d affect the probability for an event to survive the event selection criteria. The $\Delta|y|$ and V_d dependent selection efficiencies are depicted in Fig. 2.

The performance of the unfolding algorithm is tested in sets of pseudo experiments, each of which provides a randomly-generated sample distribution. The average asymmetries from 50 000 pseudo experiments for the inclusive as well as for the differential measurements agree well with the true asymmetries in the sample used to model the signal component and the pull distributions agree with expectations, indicating that the treatment of uncertainties is consistent with Gaussian behavior. To test the unfolding procedure for different asymmetries, we reweight the events of the default $t\bar{t}$ sample according to their $\Delta|y|$ value, to artificially introduce asymmetries between -0.2 and $+0.2$, and then perform pseudoexperiments for each of the reweighted distributions. For the differential measurements this test is performed in each of the three bins of V_d separately. We find a linear dependence of the ensemble mean on the input value.

5 Estimation of systematic uncertainties

The measured charge asymmetry A_C can be affected by several sources of systematic uncertainty. Influences on the direction of the reconstructed top-quark momenta can change the value of the reconstructed charge asymmetry. Systematic uncertainties with an impact on the differential selection efficiency can also bias the result, as well as variations in the background rates. To evaluate each source of systematic uncertainty, we perform a new background estimation and repeat the measurement on data using modified simulated samples. This procedure differs from the one used in the previous analysis [3], where pseudo experiments were performed using the modified templates to draw the pseudo data distributions. Although both methods are valid, the one used for this analysis has the advantage that it does not depend on the asymmetry of the default $\mathrm{t\bar{t}}$ simulation. The expected systematic uncertainty for each source is taken to be the shift in the values of the corrected asymmetry between the default measurement and the one using the shifted templates.

The corrections on jet-energy scale (JES) and jet-energy resolution (JER) are changed by ± 1 standard deviations of their η and p_{T}-dependent uncertainties to estimate their effects on the measurement. In addition the effect of variations in the frequency of occurrence of pile-up events is estimated. The effects of variations in the factorization and renormalization scales (Q^2) have been estimated for W+jets and $\mathrm{t\bar{t}}$ events. Possible mismodelling of the W+jets and multijet background has been estimated as well as possible η dependent variations of the b tagging efficiency and the lepton selection efficiency. The systematic uncertainties on the measured asymmetry from the choice of parton distributions functions (PDF) for the colliding protons are estimated using the CTEQ6.6 [11] PDF set and the LHAPDF [12] package. Variations in the matching threshold, causing the largest contribution to the total systematic uncertainty in the previous measurement, have no impact on the present measurement due to the usage of the NLO event generator POWHEG for modelling the $\mathrm{t\bar{t}}$ signal.

New with respect to the previous measurement is the estimation of the following systematic uncertainties. The uncertainty due to the choice of the event generator is estimated by using simulated events generated with MADGRAPH instead POWHEG for the unfolding procedure. In addition, a signal sample with a different hadronization and shower modelling has been used (MC@NLO [13] interfaced to HERWIG [14]) to estimate the systematic uncertainty due to this part of the event generation. The impact of statistical fluctuations in the bins of the migration matrices is evaluated by repeating the measurement with altered migration matrices, where each element is varied within its statistical uncertainties. In addition to these uncertainties we estimate the influence of possible dependencies of the asymmetry on one of the three differentiating variables. Therefore we perform pseudoexperiments with reweighted simulated signal samples and evaluate the differences between true and measured asymmetries in various reweighting scenarios. We take the average of the observed deviations and assign it as systematic uncertainty. The uncertainty due to variations in the amount of initial-state and final-state radiation (ISR/FSR) on the measurement is covered by the uncertainties due to the choice of the Q^2 scale and the "unfolding" systematic. The probability for additional radiation increases with decreasing Q^2 and vice verca. Due to the strong correlation between the amount of additional radiation and the transverse momentum of the $\mathrm{t\bar{t}}$ system, the variation of the generated asymmetry as a function of $p_{\mathrm{T,t\bar{t}}}$, as done in the estimation of the "unfolding" unertainty, is also suited to estimate the effects of variations in the amount of ISR/FSR on the measurement.

The contributions of the different sources of systematic uncertainties to the total uncertainty of the inclusive measurement are summarized in table 2. The systematic uncertainties on the

differential measurements are included in the error bars of the final differential distributions (see Fig. 3).

Table 2: Systematic uncertainties for the inclusive measurement of A_C. Listed are the shifts induced by systematic uncertainties in the measurement on data.

Systematic uncertainty	inclusive A_C
JES	0.002
JER	0.002
Pileup	0.001
Generator	0.001
Migration matrix	0.002
Unfolding	0.008
W+jets	0.004
Multijet	0.001
Lepton ID/sel. efficiency	0.006
Q^2 scale	0.002
Hadronization	0.001
PDF	0.002
Total	0.012

6 Results

Table 3 gives the values of the measured inclusive asymmetry at the different stages of the analysis. The unfolded $\Delta|y|$ distribution, shown in Fig. 3 (upper left), is used to calculate the corrected asymmetry.

Table 3: The measured inclusive asymmetry at the different stages of the analysis and the corresponding theory prediction from the SM.

Uncorrected	0.003 ± 0.004 (stat.)
BG-subtracted	0.001 ± 0.005 (stat.)
Final corrected	0.004 ± 0.010 (stat.) ± 0.012 (syst.)
Theory prediction (SM)	0.0115 ± 0.0006

The results of the three differential measurements can be found in Fig. 3 (upper right, lower left and right). The measured values are compared to predictions from SM calculations [15] and to predictions from an effective field theory [16, 17]. The latter theory is capable of explaining the CDF results by introducing an anomalous effective axial-vector coupling to the gluon at the one-loop level. The gluon-quark vertex is treated in the approximation of an effective field theory with a scale for new physics contributions on the order of 1 TeV. In case of $p_{T,t\bar{t}}$, for technical reasons no theory predictions are available. Instead, we compare the measured asymmetries with the predictions obtained from POWHEG simulation. Within the uncertainties all measured values are consistent with the SM predicted values.

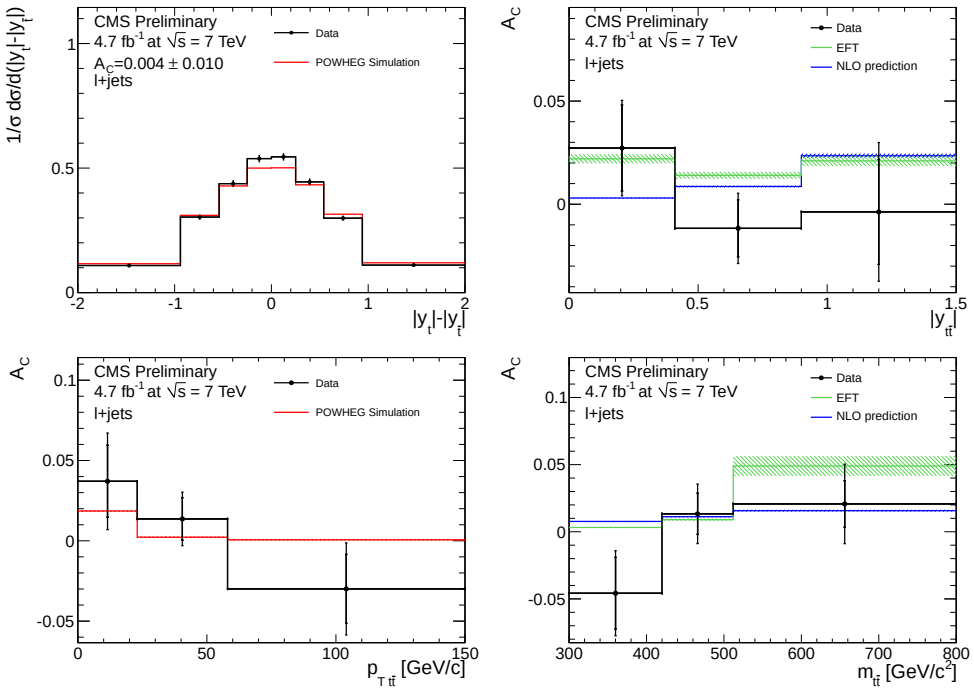

Figure 3: Unfolded inclusive $\Delta|y|$ distribution (upper left), corrected asymmetry as a function of $|y_{t\bar{t}}|$ (upper right), $p_{T,t\bar{t}}$ (lower left), and $m_{t\bar{t}}$ (lower right). The measured values are compared to NLO calculations for the SM [15] and to predictions of an effective field theory (EFT) [17]. The error bars on the differential asymmetry values indicate the statistical and systematic uncertainties.

7 Conclusion

An inclusive and three differential measurements of the charge asymmetry in $t\bar{t}$ production using data corresponding to an integrated luminosity of $4.7\,\text{fb}^{-1}$ have been reported. Events with top-quark pairs decaying in the lepton+jets channel were selected and a full $t\bar{t}$ event reconstruction was performed to determine the four-momenta of the top quarks and antiquarks. The measured distributions of the sensitive observable were then corrected for acceptance and reconstruction effects. The measured value for the inclusive asymmetry as well as the measured asymmetry as a function of three differentiating variables, the rapidity, the transverse momentum, and the invariant mass of the $t\bar{t}$ pair is in agreement with the predictions and no hints for contributions from physics beyond the standard model have been found.

8 Acknowledgment

We thank E. Gabrielli, A. Giammanco, A. Racioppi, M. Raidal, G. Rodrigo, and J.H. Kühn for very fruitful discussions and communication.

References

[1] CDF Collaboration, "Evidence for a mass dependent forward-backward asymmetry in top quark pair production", *Phys. Rev. D* **83** (2011) 112003, arXiv:1101.0034. doi:10.1103/PhysRevD.83.112003.

[2] D0 Collaboration, "Forward-backward asymmetry in top quark-antiquark production", *Phys. Rev. D* **84** (2011) 112005. doi:10.1103/PhysRevD.84.112005.

[3] CMS Collaboration Collaboration, "Measurement of the charge asymmetry in top-quark pair production in proton-proton collisions at sqrt(s) = 7 TeV", (2011). arXiv:1112.5100. Submitted to *Phys. Lett. B*.

[4] J. H. Kühn and G. Rodrigo, "Charge asymmetries of top quarks at hadron colliders revisited", (2011). arXiv:1109.6830.

[5] J. Alwall, M. Herquet, F. Maltoni et al., "MadGraph 5: going beyond", *JHEP* **06** (2011) 128, arXiv:1106.0522. doi:10.1007/JHEP06(2011)128.

[6] S. Alioli, P. Nason, C. Oleari et al., "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **1006** (2010) 043, arXiv:1002.2581. doi:10.1007/JHEP06(2010)043.

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

[8] M. L. Mangano, M. Moretti, F. Piccinini et al., "Matching matrix elements and shower evolution for top- quark production in hadronic collisions", *JHEP* **01** (2007) 013, arXiv:hep-ph/0611129. doi:10.1088/1126-6708/2007/01/013.

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

[10] V. Blobel, "An unfolding method for high energy physics experiments", arXiv:hep-ex/0208022.

[11] P. M. Nadolsky, H.-L. Lai, Q.-H. Cao et al., "Implications of CTEQ global analysis for collider observables", *Phys. Rev. D* **78** (2008) 013004, arXiv:0802.0007. doi:10.1103/PhysRevD.78.013004.

[12] WBG Collaboration, "The Les Houches Accord PDFs (LHAPDF) and Lhaglue", (2005). arXiv:hep-ph/0508110.

[13] S. Frixione and B. R. Webber, "Matching NLO QCD computations and parton shower simulations", *JHEP* **0206** (2002) 029, arXiv:hep-ph/0204244.

[14] G. Corcella, I. Knowles, G. Marchesini et al., "HERWIG 6.5 Release Note", 2002.

[15] G. Rodrigo *Private Communication* (2012).

[16] E. Gabrielli, M. Raidal, and A. Racioppi, "Implications of the effective axial-vector coupling of gluon on top-quark charge asymmetry at the LHC", (2011). arXiv:1112.5885. Submitted to Phys. Rev. D.

[17] E. Gabrielli, A. Giammanco, M. Raidal et al., "Effective axial-vector coupling of gluon and top quark charge asymmetry at the LHC", (2011). To appear in the proceedings of the LesHouches 2011 workshop.