Combinations of inclusive $t\bar{t}$ production cross sections at the Tevatron and the LHC

Serban Protopopescu for the CDF, D0, ATLAS and CMS collaborations
Brookhaven National Laboratory, Upton, NY, USA

DOI: http://dx.doi.org/10.3204/DESY-PROC-2014-02/13

The results of combining inclusive $t\bar{t}$ production cross-section measurements from D0 and CDF at the Tevatron and from ATLAS and CMS at the LHC are presented.

1 Introduction

Calculations of inclusive $t\bar{t}$ production cross-sections are now available with next-to-next-to-leading order (NNLO) and next-to-next-to-leading logarithmic (NNLL) soft gluon resummation [1], for $p\bar{p}$ and pp production processes. It is worth noting that the dominant production mechanism in $p\bar{p}$ collisions at the Tevatron centre-of-mass energy ($\sqrt{s}$) 1.96 TeV is quark-anti-quark annihilation, while in pp collisions at LHC energies ($\sqrt{s}$=7 and 8 TeV) it is mainly gluon-gluon scattering. Precise measurements of these cross sections provide a significant test of the standard model (SM) and of PDFs. Combining measurements from different channels and experiments is central to achieve the best possible precision.

The experimental cross sections are extracted by measuring events in final states expected to have large contributions from $t\bar{t}$ pairs. The SM top-quark decays almost 100% of the time to $W + b$-quark. Final states are separated according to the W decay into either $e\nu_e, \mu\nu_\mu, \tau\nu_\tau, qq'$ from each top-quark. What is really measured in each final state is the cross section multiplied by the branching ratios of t and $\bar{t}$, and it is implicitly assumed that these are given by the W branching ratios.

Each experiment combines measurements from different final states to obtain a combined measurement. In turn, the combined measurements from each experiment are combined to give a final result at a given $\sqrt{s}$. The main issue in combining different measurements is how to handle correlations of systematic uncertainties. In all $\sigma_{t\bar{t}}$ measurements at the Tevatron and even more so at the LHC, the dominant sources of uncertainty are systematic. Two methods are used to combine measurements:

1. A best linear unbiased estimator (BLUE). This method requires the construction of a covariance matrix (including statistical and systematic uncertainties) with all correlations determined externally. A weight for each result is obtained by inverting the matrix. The results are then combined using these weights to obtain the best estimate. This method is used by CDF and to combine results from different experiments.

2. Construct a combined likelihood taking the product of likelihoods in each channel multiplied by a Gaussian term for each nuisance parameter (i.e. systematic uncertainty source)

centred at zero with width of the variance expected for each parameter. Correlated uncertainties are entered in terms of common parameters. Maximizing the likelihood automatically takes care of the correlations. This method is used to combine the measurements of each individual experiment by DØ, ATLAS and CMS.

2 Tevatron measurements

CDF combines four measurements using the BLUE method: two separate ones from single lepton (ℓ=e or μ)+jets channels, one from dilepton ($\ell\ell'$) channels and one from the all jets channel.

1. $\ell\ell'$ channel(DIL) [2]: ee, $\mu\mu$ and $e\mu$ events are counted when they have at least one b-jet in a data sample of 9.1 fb^{-1} of integrated luminosity ($\int Ldt$).
2. ℓ+jets without any b-jet requirement (LJ-ANN) [3]: use a neural network (NN) based on 7 kinematic variables to separate signal and background, $\int Ldt = 4.6$ fb^{-1}.
3. ℓ+jets requiring b-tag (LJ-SVX) [3]: extract signal by a maximum likelihood fit to events with a b-tag.
4. all jets channel(HAD) [4]: fit a reconstructed top mass from events with 6-8 jets,> 1 b-jet selected with a 13 variable NN, $\int Ldt = 2.9$ fb^{-1}.

The cross-section measurements, their combination and all their uncertainties are given in Table 1. It is worth noting that the measured integrated luminosity, which is a significant source of uncertainty in the $\ell\ell'$ and all hadronic channels measurements, is derived in the ℓ+jets channel using the expected $Z/\gamma*$ cross section rather than from inclusive $p\bar{p}$ measurements. Using the $Z/\gamma*$ measurements results in a luminosity uncertainty that is 2.5 times smaller. The dominant contribution to the CDF combination is the LJ-ANN measurement; adding the other measurements reduces the uncertainty by 10%. Combining the CDF measurements results in $\sigma_{t\bar{t}} = 7.63 \pm 0.50$ (statistical+systematic).

DØ combines two measurements:

1. ℓ+jets [5]: two methods were combined to extract the number of $t\bar{t}$ events from orthogonal data sets:
   (a) a 7 (3) kinematic variable multi-variate discriminant random-forest to separate signal and background in events with 0 (1) b-jets and 3 or > 3 jets;
   (b) a maximum likelihood fit in events with > 1 b-jets and 3 or > 3 jets.
2. $\ell\ell'$ [6]: fit 4 b-tagging NN discriminant distributions to signal and background templates, $e\mu$+1 jet and ee, $\mu\mu$ and $e\mu$+>1 jet.

The combination was carried out using a modified likelihood with nuisance parameters to take into account systematic uncertainties. The correlations are taken into account by using the same parameters for common systematics:

$$\mathcal{L} = \prod_i \prod_j P[n_{ij}|\mu_{ij}(\sigma_{t\bar{t}}, \nu_k)] \prod_k \mathcal{G}(\nu_k; 0, SD) \quad (1)$$

	DIL	LJ-ANN	LJ-SVX	HAD	CDF combined
Central value of $\sigma_{t\bar{t}}$	7.09	7.82	7.32	7.21	7.63
Sources of uncertainty					
Statistical	0.49	0.38	0.36	0.50	0.31
Detector model	039	0.11	0.34	0.41	0.17
Signal model	0.23	0.23	0.23	0.44	0.21
Jets model	0.23	0.23	0.29	0.71	0.21
Method to extract $\sigma_{t\bar{t}}$	0.00	0.01	0.01	0.08	0.01
Background model (theory)	0.01	0.13	0.29	0.00	0.10
Background model (data)	0.15	0.07	0.11	0.59	0.08
Normalization of $Z/\gamma*$ prediction	-	0.16	0.15	-	0.13
Luminosity inelastic $\sigma_{p\bar{p}}$	0.28	0.00	0.00	0.29	0.05
Luminosity detector	0.30	0.02	0.02	0.30	0.06
Total systematic uncertainty	0.67	0.41	0.61	1.18	0.39
Total uncertainty	0.83	0.56	0.71	1.28	0.50

Table 1: CDF measurements of $\sigma_{t\bar{t}}$ from $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV with absolute uncertainty components in pb and results of the channel combination.

where i refers to the channels, j to the bins of the NN distribution and k refers to the nuisance parameters. $P(n, \mu)$ is the probability of observing n events while expecting μ while $\mathcal{G}(\nu; 0, SD)$ is the Gaussian probability density with mean at zero and width one standard deviation (SD) of the systematic uncertainty ν. This formulation extracts the cross section and handles correlations of systematic uncertainties between channels automatically by maximizing the likelihood with $\sigma_{t\bar{t}}$ and ν_k as free parameters. Table 2 shows the statistical and systematic uncertainties for the dilepton channel and for the combination of dilepton with lepton+jets channel. The cross section measured with $\ell\ell'$ events is $7.36^{+0.90}_{-0.79}$ pb, the one with ℓ+jets events is $7.78^{+0.77}_{-0.64}$ pb and the combined result is $7.56^{+0.63}_{-0.56}$ pb. Combining the measurement from both channels improves the precision by 24% over the best single measurement (single-lepton channel).

The combined CDF and DØ measurements were in turn combined using the BLUE method, to obtain the best $\sigma_{t\bar{t}}$ value for the Tevatron [7]. CDF and DØ have weights of 60% and 40% respectively, while the correlation between them is 17%. The uncertainties for each experiment and their combination are given in Table 3. Quantities known to be correlated are assumed to be 100% correlated, which leads to an overestimate (and thus conservative estimate) of the overall uncertainty. The CDF measurement has a larger weight, due mainly to reducing the luminosity uncertainty in the ℓ+jets channel by using the $Z/\gamma*$ predicted cross section to calculate $\int Ldt$. The CDF, DØ and combined measurements are displayed in Fig 2. The Tevatron combination, $\sigma_{t\bar{t}} = 7.61 \pm 0.41$ pb, improves the precision by 20% over the CDF combined measurement. The combination is in very good agreement with the NNLO+NNLL SM prediction $7.34^{+0.23}_{-0.27}$ pb [1] for a top-quark mass of 172.5 GeV.

Source	$\ell\ell'$		$\ell\ell' + \ell j$	
Statistical	+0.50	-0.48	+0.20	-0.20
Muon identification	+0.11	-0.11	+0.07	-0.06
Electron identification and smearing	+0.24	-0.23	+0.13	-0.13
Signal model	+0.34	-0.33	+0.16	-0.06
Triggers	+0.19	-0.19	+0.05	-0.05
Jet energy scale	+0.13	-0.12	+0.04	-0.04
Jet reconstruction and identification	+0.21	-0.20	+0.12	-0.09
b-tagging	+0.06	-0.06	+0.16	-0.14
Background normalization	+0.29	-0.27	+0.11	-0.10
W+HF fraction	-	-	+0.12	-0.04
Luminosity	+0.57	-0.51	+0.48	-0.43
Other	+0.10	-0.10	+0.06	-0.06
Template statistics	+0.08	-0.08	+0.04	-0.04

Table 2: The DØ breakdown of uncertainties (in pb) in the $\ell\ell'$ channel and the combined $\ell\ell'$ and ℓj measurement using the nuisance parameter technique. The uncertainties show the change in the measured cross section when shifting the nuisance parameter by ± 1 standard deviation from its fitted mean.

	CDF	D0		Tevatron
Central value of $\sigma_{t\bar{t}}$ (pb)	7.63	7.56		7.60
Sources of uncertainty			Correlation	
Statistical	0.31	0.20	0	0.20
Detector model	017	0.22	0	0.13
Signal model	0.21	0.13	1	0.18
Jets model	0.21	0.11	0	0.13
Method to extract $\sigma_{t\bar{t}}$	0.01	0.07	0	0.03
Background model (theory)	0.10	0.08	1	0.10
Background model (data)	0.08	0.06	0	0.05
Normalization of $Z/\gamma*$ prediction	0.13	-	0	0.08
Luminosity inelastic $\sigma_{p\bar{p}}$	0.05	0.30	1	0.15
Luminosity detector	0.06	0.35	0	0.36
Total systematic uncertainty	0.39	0.56		0.36
Total uncertainty	0.50	0.59		0.41

Table 3: CDF and D0 measurements of $\sigma_{t\bar{t}}$ from $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV with absolute uncertainty components in pb and results of their combination. Systematic uncertainties known to be correlated are assumed to be 100% correlated.

3 LHC measurements

ATLAS combines six $\sqrt{s} = 7$ TeV measurements [8]:

1. three $\ell\ell'$ measurements [9]: ee,$\mu\mu$,$e\mu$ by counting events with ≥ 2 jets (with and without b-jets separately) in a data sample of $\int Ldt = 0.7$ fb^{-1}.

2. two ℓ+jets measurements [10]: e+jets, μ+jets using a 4 kinematic variable likelihood discriminant and no b-jet requirement ($\int Ldt = 0.7$ fb^{-1}).

3. all jets [11]: extract signal by fitting a reconstructed top-quark mass with signal and background templates ($\int Ldt = 1.02$ fb^{-1}).

$$\mathcal{L} = \mathcal{L}_{\ell+jets}(\sigma_{t\bar{t}}, L, \vec{\nu}) \prod_{i\in ee,\mu\mu,e\mu} P[n_i|\mu_i(\vec{\nu})] \prod_{j\in \mathrm{all-hadbins}} P[n_j|s_j(\vec{\nu}) + b_j(\vec{\nu})] \prod_{k\notin \ell+\mathrm{jets}} \mathcal{G}(\nu_k; 0, SD) \quad (2)$$

where L is luminosity, μ expected number of events, s_j (b_j) expected number of signal (background) events and $\vec{\nu}$ are nuisance parameters. The main difference with respect to equation (1) is that the ℓ+jets likelihood is approximated by a multivariate Gaussian likelihood ($\mathcal{L}_{\ell+jets}$). Figure 1 shows the negative log likelihood for each channel and the combination as function of $\sigma_{t\bar{t}}/\sigma_{SM}$, with and without including systematic uncertainties. It shows that for the ℓ+jets channel the multivariate Gaussian is a very good approximation to the exact likelihood. The total number of parameters is 89 (including $\sigma_{t\bar{t}}$), 26 are shared between $\ell\ell'$ and ℓ+jets, and 12 are common to all channels. The value of $\sigma_{t\bar{t}}$ obtained from the combination is 177 ± 11 pb, an improvement in precision of 10% over the measurement from ℓ+jets alone, 179 ± 12 pb.

CMS combines seven $\sqrt{s} = 7$ TeV measurements [12]:

1. $\ell\ell'$ [13]: ee, $\mu\mu$ and $e\mu$ based on counting events with at least one identified b-jet ($\int Ldt = 1.14$ fb^{-1}).

2. ℓ+jets [14]: e+jets, μ+jets ($\int Ldt = 0.8$ fb^{-1}, 1.09 fb^{-1}). The number of $t\bar{t}$ events in data samples were extracted by maximizing a binned likelihood of secondary vertex mass distributions. Event samples were split by number of jets and identified b-jets.

3. all jets [15]: extract the number $t\bar{t}$ events in events with 6-8 jets and two identified b-jets by fitting a reconstructed top mass with signal and background templates ($\int Ldt = 1.09$ fb^{-1}).

4. $\mu + \tau$ [16]: count events with at least one identified b-jet after background subtraction and reducing the background by applying a series of cuts ($\int Ldt = 1.09$ fb^{-1}).

CMS obtains a cross section combining the above measurements using a binned maximum likelihood fit applied to a combined likelihood function similar to that used by DØ (Eq. 1), a product of the Poisson likelihood of observing a certain number of events given an expected number for each bin considered, multiplied by Gaussian distributions for the nuisance parameters. Counting experiments (ee, $\mu\mu$, $e\mu$ and $\mu\tau$) are entered as single bins. The hadronic analysis is based on an unbinned fit to the reconstructed top-quark mass distribution and is not binned. Instead, the results are parametrized so they can be input into the combined likelihood as a single bin. The CMS combined value of $\sigma_{t\bar{t}}$ at $\sqrt{s} = 7$ TeV is 166 ± 13.7 pb, an improvement

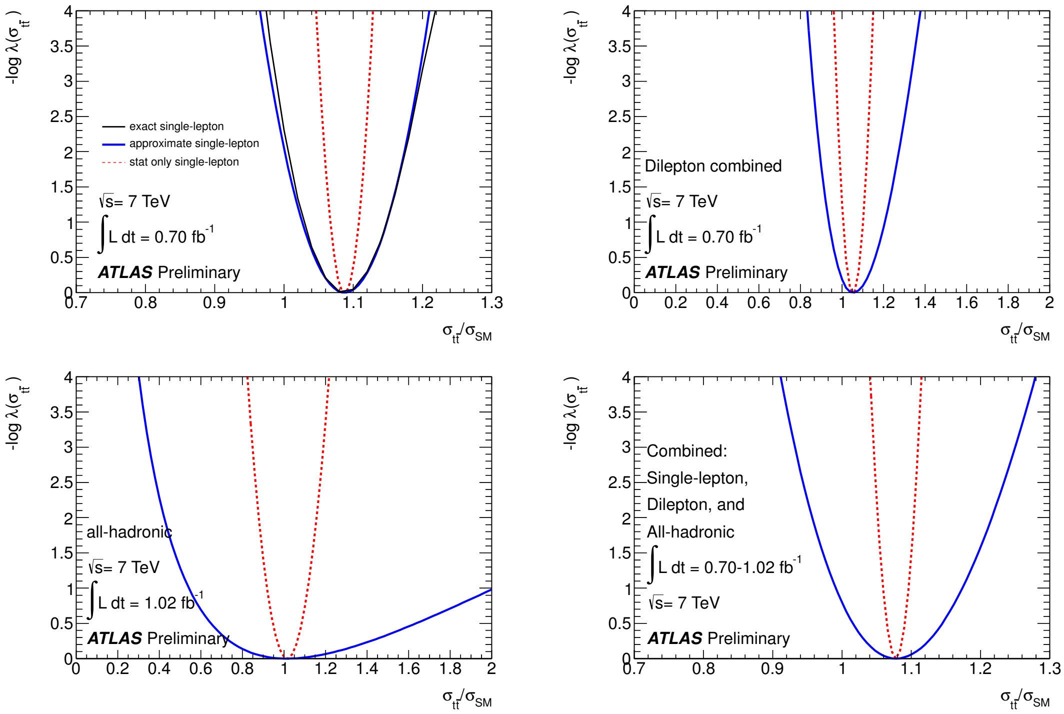

Figure 1: ATLAS -log(likelihood) for single lepton (systematic uncertainties do not include parameters common with dilepton and all-hadronic channels), dilepton combined, all-hadronic and all channels combined as function $\sigma_{t\bar{t}}/\sigma_{SM}$ [8].

	ATLAS	CMS		LHC
Central value of $\sigma_{t\bar{t}}$ (pb)	177.0	165.8		173.3
Sources of uncertainty			Correlation	
Statistical	3.2	2.2	0	2.3
Detector model	5.3	8.8	0	4.6
Jet energy scale	2.7	3.5	0	0.13
Signal model				
Monte Carlo	4.2	1.1	1	3.1
Parton shower	1.3	2.2	1	1.6
Radiation	0.8	4.1	1	1.9
PDF	1.9	4.1	1	2.6
Method to extract $\sigma_{t\bar{t}}$	2.4	n/e	0	1.6
Bacgkround model (theory)	1.6	1.6	1	1.6
Background model (data)	1.5	3.4	0	1.6
W leptonic branching ratio	1.0	1.0	1	0.08
Luminosity				
Bunch current	5.3	5.1	1	5.3
Luminosity measurement	4.3	5.9	0	3.4
Total systematic uncertainty	10.8	14.2		9.8
Total uncertainty	11.3	14.4		10.1

Table 4: ATLAS and CMS measurements of $\sigma_{t\bar{t}}$ from pp collisions at $\sqrt{s} = 7$ TeV with absolute uncertainty components in pb and results of their combination. Systematic uncertainties known to be correlated are assumed to be 100% correlated.

in precision of 3.7% over the ℓ+jets cross section, 164 ± 14.2 pb. The combined $\sigma_{t\bar{t}}$ from ATLAS and CMS were in turn combined using the BLUE method as the different treatment of systematic uncertainties makes the likelihood approach difficult to implement [19]. Table 4 shows a detailed composition of the uncertainties and their correlations used in the combination. Note that whenever a systematic uncertainty is known to be correlated, the correlation is set to 100% (as it was for the CDF-DØ combination in Section 2). The largest uncertainty components are the beam-bunch current (5.3 pb) and the sum of all terms that depend on Monte Carlo simulation (5.0 pb). Both of these are correlated between experiments, together they contribute 7.3 pb uncertainty to the result 173.3 ± 10.1 pb. Combining ATLAS and CMS reduces the uncertainty over the best combined measurement from a single experiment, 177.0 ± 11.3 pb, by 10%. Figure 2 displays the cross-section measurements of each experiment that are used in the combinations and the results of combining the measurements. The best LHC value at $\sqrt{s} = 7$ TeV, 173.3 ± 10.1 pb is in very good agreement with the SM prediction of $172.0^{+6.4}_{-7.5}$ pb[1]. It should be noted that since the combination, CMS has published cross-section measurements in $\ell\ell'$ [17] and ℓ+jets [18] channels using 2.3 fb^{-1} of data with improved precision.

The LHC measurements of $\sigma_{t\bar{t}}$ at $\sqrt{s} = 8$ TeV have not been combined yet. There is an ATLAS measurement in the $e\mu$ channel using all the data collected in 2012 ($\int L dt = 20.3$ fb^{-1}) requiring at least one identified b-jet, 238 ± 11 pb [20], and a CMS measurement using all three $\ell\ell'$ channels with identified b-jets but only a subset of the 2012 data ($\int L dt = 2.4$ fb^{-1}),

227 ± 15 [21]. It is worth noting that the largest single source of uncertainty in these LHC measurements is the integrated luminosity. These measurements are in very good agreement with the SM prediction of $245.8^{+8.8}_{-10.6}$ pb [1]. There are also measurements by ATLAS and CMS using the ℓ+jets channels, 241 ± 32 [22] and 228 ± 32 [23], which are in good agreement with $\ell\ell'$ measurements. These measurements do not have an impact on a combined result due to their larger systematic uncertainties.

4 Conclusions

All the various measurements of $\sigma_{t\bar{t}}$ at the Tevatron and LHC are consistent with each other and are in good agreement with SM expectations. By combining the measurements from different channels and experiments the value of $\sigma_{t\bar{t}}$ has been determined with a precision of 5-6% in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV [7] and in pp collisions at $\sqrt{s} = 7$ TeV [19] of 7.60 ± 0.41 pb and 173.3 ± 10.1 pb respectively. Similar precision has been achieved using just the $e\mu$ channel at $\sqrt{s} = 8$ TeV [20], 238 ± 11 pb. These measurements are approaching the precision of the NNLO+NNLL theoretical predictions, $7.34^{+0.23}_{-0.27}$ pb, $172.0^{+6.4}_{-7.5}$ pb and $245.8^{+8.8}_{-10.6}$ pb, and are in excellent agreement with them.

5 Acknowledgement

This work was supported by the US Department of Energy under Contract No. DE-AC02-98CH10886.

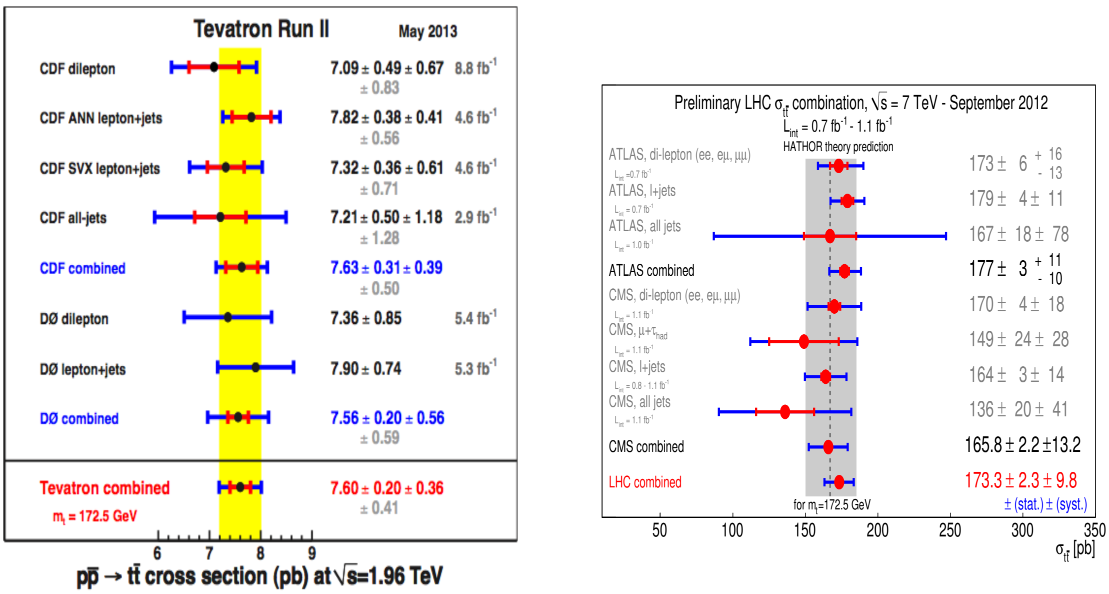

Figure 2: $\sigma_{t\bar{t}}$ measurements in the channels used for combinations and the combined results. Left CDF and DØ [7] , right ATLAS and CMS [19]. The theory band corresponds to the value predicted by NLO calculations, not the more precise NNLO+NNLL calculations referred to in the text.

6 Bibliography

References

[1] M. Czakon, P. Fiedler and A. Mitov, Phys. Rev. Lett. **110** (2013) 25, 252004, arXiv:1303.6254 [hep-ph].

[2] T. Aaltonen *et al.* [CDF Collaboration], Phys. Rev. D **88** (2013) 091103, arXiv:1304.7961 [hep-ex].

[3] T. Aaltonen *et al.* [CDF Collaboration], Phys. Rev. Lett. **105** (2010) 012001, arXiv:1004.3224 [hep-ex].

[4] T. Aaltonen *et al.* [CDF Collaboration], Phys. Rev. D **81** (2010) 052011, arXiv:1002.0365 [hep-ex].

[5] J. P. Lees *et al.* [BaBar Collaboration], Phys. Rev. D **84** (2011) 012001, arXiv:1105.5159 [hep-ex].

[6] V. M. Abazov *et al.* [D0 Collaboration], Phys. Lett. B **704** (2011) 403, arXiv:1105.5384 [hep-ex].

[7] CDF Collaboration, CDF Conference Note 10926; DØ Collaboration, DØ Note 6363-CONF.

[8] ATLAS Collaboration, ATLAS-CONF-2012-024.

[9] L. Wang and X. -F. Han, JHEP **1205** (2012) 088, arXiv:1203.4477 [hep-ph].

[10] ATLAS Collaboration, ATLAS-CONF-2011-121.

[11] ATLAS Collaboration, ATLAS-CONF-2011-031.

[12] CMS Collaboration, CMS-PAS-TOP-11-024.

[13] CMS Collaboration, CMS-PAS-TOP-11-005.

[14] CMS Collaboration, CMS-PAS-TOP-11-003.

[15] CMS Collaboration, CMS-PAS-TOP-11-007.

[16] CMS Collaboration, CMS-PAS-TOP-11-006.

[17] CMS Collaboration, JHEP 1211 (2012) 067.

[18] S. Chatrchyan *et al.* [CMS Collaboration], Phys. Lett. B **720** (2013) 83, arXiv:1212.6682 [hep-ex].

[19] ATLAS Collaboration, ATLAS-CONF-2012-134; CMS Collaboration, CMS-PAS-TOP-12-003.

[20] ATLAS Collaboration, ATLAS-CONF-2013-097.

[21] CMS Collaboration, CMS-PAS-TOP-12-007.

[22] ATLAS Collaboration, ATLAS-CONF-2011-149.

[23] CMS Collaboration, CMS-PAS-TOP-12-006.