CMS Physics Analysis Summary

Contact: cms-pag-conveners-top@cern.ch 2015/12/03

Comparisons of Theory Predictions with Data for $t\bar{t}$ Production in pp Collisions at $\sqrt{s}$= 8 TeV

The CMS Collaboration

Abstract

Studies are presented comparing $t\bar{t}$ differential distribution data collected with the CMS detector at $\sqrt{s} = 8$ TeV to state-of-the-art theoretical predictions. The predictions are accurate to next-to-leading order quantum chromodynamics and interfaced with the PYTHIA 8 and HERWIG++ event generators. The predictions of these simulation tools and a next-to-next-to-leading order calculation are compared to unfolded data distributions. All deviations of the theory predictions from the data remain inside the envelope considered for factorization and renormalization scale variations.

1 Introduction

Extracting the most information from top-quark measurements at hadron colliders requires a dedicated effort in theoretical modelling. In LHC Run I (data taking period with proton proton collisions at $\sqrt{s}$ = 7 and 8 TeV during the years 2010 and 2012), it is observed that the leading order (LO) generator MADGRAPH 5.1.3.30 [1] with up to three additional partons interfaced with PYTHIA 6.426 [2] using the Z2* underlying event tune [3][1] describes most of the differential $t\bar{t}$ distributions obtained from data reasonably well, except for the transverse momentum of the top quark, p_T^t [6]. A new generation of next-to-leading order (NLO) matrix element (ME) event generators interfaced with new parton-shower codes are expected to provide better modelling of signal and backgrounds, and reduce the dominant theoretical uncertainties with respect to LHC Run I.

In this document, we present comparisons of differential $t\bar{t}$ cross sections determined as a function of basic kinematic variables of the top quark and the $t\bar{t}$ system determined through kinematic fitting and reconstruction. Data distributions presented in [6] are compared to distributions obtained from MG5_aMC@NLO [7] and POWHEG v2 [8–10]. The comparisons are made for the following kinematic variables: top (or anti-top) quark transverse momentum (p_T^t), and rapidity (y_t), $t\bar{t}$ system transverse momentum ($p_T^{t\bar{t}}$), rapidity ($y_{t\bar{t}}$), and invariant mass ($m_{t\bar{t}}$) for the lepton+jets and dilepton channels separately. The normalized differential distributions are determined using the following formula:

$$\frac{1}{\sigma}\frac{d\sigma_i}{dX} = \frac{1}{\sigma}\frac{x_i}{\Delta_i^X \mathcal{L}} \tag{1}$$

where X is the kinematic variable, σ is the $t\bar{t}$ cross-section, Δ_i^X is the bin width of bin i, x_i is the measured number of events in data after background subtraction and corrected for detector efficiencies, acceptances, and bin-to-bin migrations, and finally $\mathcal{L}$ is the integrated luminosity. The unfolding procedure to the parton-level in the full phase space requires an extrapolation from the measured particles corrected for detector and reconstruction effects (particle-level) to the parton-level and is generator dependent. We use RIVET [11] for the comparisons between theory predictions and unfolded data. The RIVET framework is designed to be robust to run on any MC generator by allowing only the final state particles in the HepMC [12] data format. This ensures independence from unphysical information in the generators and direct theory versus data comparisons.

The top-quark final state at particle-level (see section 3) is implemented in a RIVET analysis module that is available in the code repository. Measurements at $\sqrt{s}$ = 8 TeV described above, are based on parton-level information in the full phase space using MADGRAPH +PYTHIA 6. To match the top-quark final state at particle-level distributions to the measurements unfolded to the parton-level, we applied a correction function to the particle-level distributions derived using the same MADGRAPH +PYTHIA 6 configuration that was used for the original measurement of the data points. Using the same MC configuration as used for the unfolding to correct back the parton-level to particle-level, the model dependence introduced in unfolding to parton-level and extrapolating the measurement to the full phase space is eliminated. The correction factors are described in section 3.

Finally, sections 4 and 5 show the results comparing the nominal theory predictions after corrections versus data, and a study of the effect in the predictions from renormalization and factorization scale variations in Matrix Elements.

[1] Z2* tune is based on the Z1 tune[4] but using the CTEQ6L PDF set [5] instead of the CTEQ5L PDF set.

2 Data, Monte Carlo Samples and Settings

Data samples corresponding to an integrated luminosity of 19.7 fb^{-1} taken at $\sqrt{s} = 8$ TeV are used. The differential distributions obtained from data are unfolded to the parton-level. A detailed description of the measurement can be found in [6]. Here only the most important aspects of unfolding and comparison with the theory are given. The results are presented in the full parton-level phase space of $t\bar{t}$ production. The default configuration for the generation of simulated $\text{t}\bar{\text{t}}$ signal events at $\sqrt{s} = 8$ TeV (which is also used for deriving the corrections to the data) is based on the Monte Carlo (MC) MADGRAPH 5.1.5.11 matrix element generator [1], MADSPIN [13] for the decay of heavy particles, PYTHIA 6.426 for parton showering, underlying event modelling and hadronization [2] using the $Z2^*$ tune [3], and TAUOLA [14] for decays of τ leptons. The same MC configuration is used for deriving the factors to correct the particle-level MC distributions to the parton-level as discussed in the next section. For the unfolding of detector effects, the particle-level events are passed to a full simulation of the CMS detector based on GEANT4 [15].

The additional MC distributions presented in this document are obtained with version 2.2.1 of the MG5_aMC@NLO and POWHEG v2 event generators. The heavy-quark process [16] as implemented in POWHEG v2 (Revision 2819) is used and showered with both PYTHIA 8.205 [17][2] and HERWIG++ 2.7.1 [19]. For MG5_aMC@NLO, samples for $t\bar{t}$ production are generated using the FXFX merging scheme [20], where one generates up to two additional partons at the ME level, along with PYTHIA 8. In addition, two MG5_aMC@NLO samples are generated for comparison: (1) an inclusive NLO MG5_aMC@NLO sample (with no matching or merging) showered and hadronized with HERWIG ++, and (2) a $\text{t}\bar{\text{t}}$ plus up to 3 additional jets at LO precision with MLM jet matching [21] as implemented in PYTHIA 8.

Some of the primary settings used for the MC sample generation are given in Table 1; a more detailed list of generator parameters are available in the Appendix. The top-quark mass value used in all the simulations is 172.5 GeV and the LO CTEQ6L1 [22] PDF is used for the shower and UE in the HERWIG++ samples while the corresponding PDF at Matrix Element is used again for the PYTHIA 8 showered samples. The factorization and renormalization scales are denoted by μ_F and μ_R, respectively. In the definitions of these scales, the transverse mass is used as $m_{\text{T}} = \sqrt{m_t^2 + P_{\text{T}}^2}$. The transverse mass of the top quark is shown with m_T^t. The h_{damp}, which is the model parameter that controls ME/Parton Shower (PS) matching in POWHEG and effectively regulates the high-p_{T} radiation, is set to the top-quark mass, i.e. 172.5 GeV. The q_{cut} parameter is the matching scale while x_{qcut} is the cutoff scale value for MLM matching in the k_T scheme. The smallest q_{cut} value is restricted to be above x_{qcut}. The q_{cut}^{ME} parameter is the cut applied to regulate multi-jet matrix elements in the FXFX matching procedure. The minimum p_T for the emission of light quarks is represented by $ptsqmin$.

[2] with matching performed using the `Main31` [18] user hook of PYTHIA 8 (see Appendix B.4)

Table 1: The MC samples used in the analysis.

ME generator	MADGRAPH +PYTHIA 6	MG5_aMC@NLO	POWHEG v2
ME mode	MLM	MLM	hvq
ME pQCD level	$t\bar{t}+0,1,2,3$ jets [LO]	$t\bar{t}+0,1,2,3$ jets [LO]	$t\bar{t}$ [NLO]
$\mu_R = \mu_F$	$\sqrt{m_t^2 + \sum_j p_T^2}$	$\sum_{t,\bar{t},\text{jets}} m_T/2$	m_T^t
PS	PYTHIA 6.426	PYTHIA 8.205	PYTHIA 8.205
Tune	Z2*	CUETP8M1 [23]	CUETP8M1
ME PDF	CTEQ6L1	NNPDF23 [24]	CT10 [25]
q_{cut}	40 GeV	40 GeV	-
Other	$x_{qcut} = 20$ GeV	$x_{qcut} = 20$ GeV	$h_{damp} = m_t$
			$ptsqmin$ = 0.8 GeV
			$pthard$ = 0

ME generator	POWHEG v2	aMC@NLO	MG5_aMC@NLO
ME mode	hvq	Inclusive	FxFx Merging
ME pQCD level	$t\bar{t}$ [NLO]	$t\bar{t}$ [NLO]	$t\bar{t}+0,1,2$ jets [NLO]
m_T^t	m_T^t	$\sum_{t,\bar{t}} m_T/2$	$\sum_{t,\bar{t},\text{jets}} m_T/2$
PS	HERWIG++ 2.7.1	HERWIG++ 2.7.1	PYTHIA 8.205
Tune	EE5C [19]	EE5C	CUETP8M1
ME PDF	CT10	CTEQ6L1	NNPDF3.0 [26]
q_{cut}	-	-	30 GeV
Other	$h_{damp} = m_t$,	-	$q_{cut}^{ME} = 10$ GeV
	$ptsqmin$ = 0.8 GeV		

3 The Particle-Level Top Quark and RIVET Implementation

In this analysis, the top-quark final state at particle-level is defined using the following particle definitions:

- Dressed lepton : An object constructed by applying the anti-k_T jet clustering [27] with a distance parameter of $R = 0.1$ to the collection of final state electrons, muons and photons that do not originate from hadron decays. The clustered objects that pass the acceptance cuts are considered.
- Prompt neutrino : A neutrino that does not originate from a hadronic decay.
- Ghost B hadron : A weakly decaying hadron with B flavor and its decay products are selected. Since both the B hadron and its decay products are included, the momenta of the B hadron is rescaled by a very small factor (10^{-20}) to be used in the jet clustering. This avoids double-counting, but retains information about the origin of the decay products.
- Jet : A jet is constructed applying the anti-k_T jet clustering algorithm with a distance parameter of $R = 0.5$ to the collection of all particles except neutrinos and constituents of dressed leptons. For the generator level b-jet finding, ghost B hadrons are also added to the jet clustering. Jets are required to pass acceptance cuts of $p_T > 30$ GeV and $|\eta| < 2.4$.
- b-jets : b-quark jets are defined as the jets with a Ghost B hadron among their constituents.

Based on the particle-level objects defined above, W bosons and top quarks are reconstructed using W boson and top-quark mass constraints, respectively, as described below

- W boson pair final state at particle-level : Combinations of dressed lepton and prompt neutrinos are made to reconstruct the W boson final state at particle-level that decays to a lepton-neutrino pair. The same is done with jets without B hadrons to reconstruct the W boson decaying hadronically. The event is considered as lepton + jet channel if only one lepton + neutrino pair exists. If there are two lepton + neutrino pairs in the event, the event is considered in the dilepton channel. The pair that gives the minimum of $|M_W^a - M_{\mathrm{W}}| + |M_W^b - M_{\mathrm{W}}|$ is defined as the W boson pair final state at particle-level. Here, M_{W} is the world average W boson mass value [28].
- Top-quark pair final state at particle-level : Combinations of particle-level W bosons and b jets are considered to reconstruct the top and the anti-top quarks in the event. The pair that yields the minimum of $|M_{top}^a - M_{\mathrm{t}}| + |M_{top}^b - M_{\mathrm{t}}|$ is defined as the top-quark pair final state at particle-level, and M_{t} is the top-quark mass used as input in the simulations.

Finally, decay channel dependent selections are made to mimic those applied on the reconstructed data. The existence of only one lepton with $p_{\mathrm{T}} > 33$ GeV and $|\eta| < 2.1$ for the lepton+jet channel, and two opposite charge leptons with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.4$ for the dilepton channel are required.

Using the particle-level definition above, bin-by-bin correction factors are derived using MADGRAPH +PYTHIA 6 as

$$C_i = \frac{\sigma_i^{\mathrm{parton}}(\text{full phase space})}{\sigma_i^{\mathrm{particle}}(\text{fiducial volume})}, \tag{2}$$

where σ_i is the cross-section in bin i, "parton" refers to top quarks or anti quarks before the decay, "particle" refers to top quarks or anti quarks as defined in the RIVET implementation above. The correction factors are shown in Tables 2-5.

The correction factors are given in Tables 2-5 along with their statistical error. Note that the scale factors are larger than one because of the large corrections extrapolating top-quark final state at particle-level in the fiducial volume to the full phase space. In the tables, it can be seen that the correction factors in the dilepton channel are smaller due to the looser cuts compared to the ones in the lepton+jets channel. These correction factors are used to correct the MC distributions to parton-level in order to make the comparisons with the data.

Table 2: The particle to parton-level correction factors with the corresponding statistical errors for top-quark p_{T} and y bins in the l+jets channel.

p_{T}^{t} (GeV)	$C_i(p_T^t)$	y_t	$C_i(y_t)$
$[0, 60)$	5.474±0.005	$[-2.5, -1.6)$	10.004±0.022
$[60, 100)$	4.941±0.004	$[-1.6, -1.2)$	4.546±0.007
$[100, 150)$	4.173±0.003	$[-1.2, -0.8)$	3.828±0.004
$[150, 200)$	3.391±0.003	$[-0.8, -0.4)$	3.601±0.004
$[200, 260)$	2.786±0.003	$[-0.4, 0.0)$	3.522±0.003
$[260, 320)$	2.371±0.004	$[0.0, 0.4)$	3.525±0.003
$[320, 400)$	2.194±0.005	$[0.4, 0.8)$	3.601±0.004
$[400, 500)$	2.197±0.009	$[0.8, 1.2)$	3.808±0.004
		$[1.2, 1.6)$	4.532±0.007
		$[1.6, 2.5)$	9.995±0.022

Table 3: The particle to parton-level correction factors for $p_{\mathrm{T}}^{t\bar{t}}$, $y_{t\bar{t}}$, and $m_{t\bar{t}}$ bins in the l+jets channel.

$p_{\mathrm{T}}^{t\bar{t}}$ (GeV)	$C_i(p_T^{t\bar{t}})$	$y_{t\bar{t}}$	$C_i(y_{t\bar{t}})$	$m_{t\bar{t}}$ (GeV)	$C_i(m_{t\bar{t}})$
$[0, 20)$	6.053±0.007	[-2.5,-1.3)	11.932±0.036	$[345, 400)$	7.145±0.009
$[20, 45)$	3.777±0.004	[-1.3,-0.9)	4.804±0.009	$[400, 470)$	5.637±0.006
$[45, 75)$	3.562±0.004	[-0.9,-0.6)	3.783±0.006	$[470, 550)$	4.050±0.005
$[75, 120)$	3.601±0.005	[-0.6,-0.3)	3.391±0.005	$[550, 650)$	3.026±0.004
$[120, 190)$	3.569±0.006	[-0.3,0.0)	3.227±0.004	$[650, 800)$	2.326±0.004
$[190, 300)$	3.410±0.009	[0.0,0.3)	3.219±0.004	$[800, 1100)$	1.774±0.004
		$[0.3, 0.6)$	3.383±0.005	$[1100 - 1600)$	1.235±0.006
		$[0.6, 0.9)$	3.774±0.006		
		$[0.9, 1.3)$	4.788±0.009		
		$[1.3, 2.5)$	11.905±0.035		

Table 4: The particle to parton-level correction factors with the corresponding statistical errors for top-quark p_{T} and y bins in the dilepton channel.

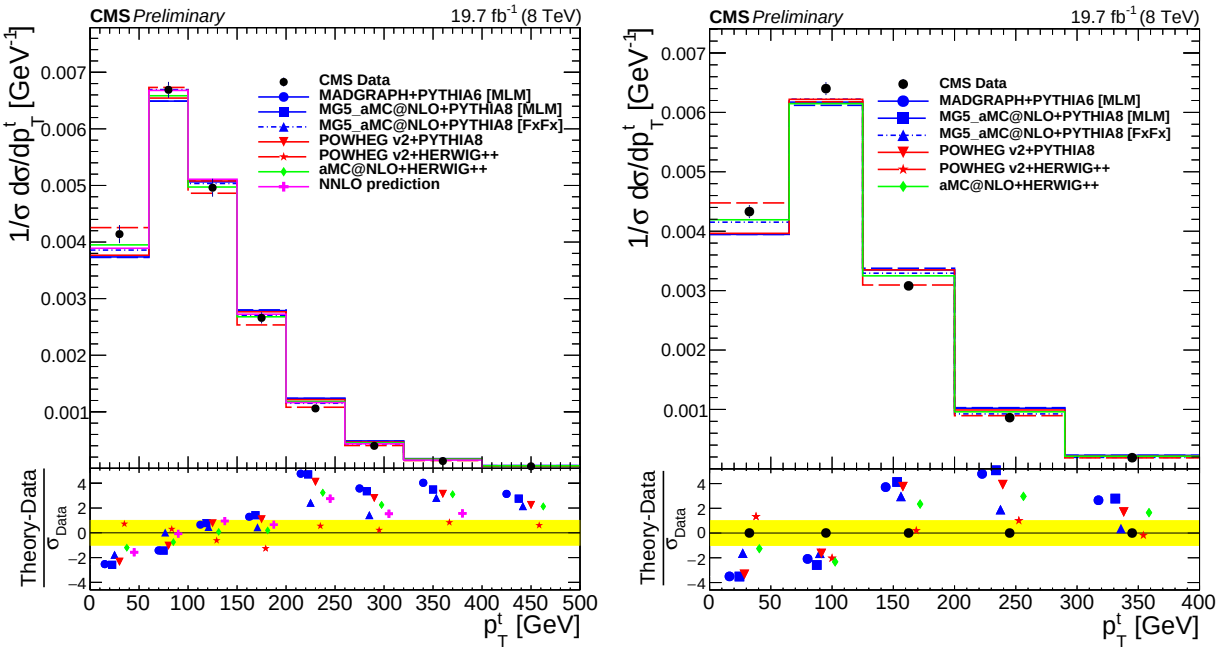
p_{T}^{t} (GeV)	$C_i(p_T^t)$	y_t	$C_i(y_t)$
$[0, 65)$	2.268±0.003	$[-2.5, -1.6)$	3.993±0.014
$[65, 125)$	2.372±0.003	$[-1.6, -1.0)$	2.308±0.005
$[125, 200)$	2.324±0.003	$[-1.0, -0.5)$	2.075±0.004
$[200, 290)$	2.0346±0.005	$[-0.5, 0.0)$	2.038±0.004
$[290, 400)$	1.737±0.007	$[0.0, 0.5)$	2.036±0.003
		$[0.5, 1.0)$	2.078±0.004
		$[1.0, 1.6)$	2.296±0.005
		$[1.6, 2.5)$	4.018±0.014

Table 5: The particle to parton-level correction factors with the corresponding statistical errors for $p_{\mathrm{T}}^{t\bar{t}}$, $y_{t\bar{t}}$, and $m_{t\bar{t}}$ bins in the dilepton channel.

$p_{\mathrm{T}}^{t\bar{t}}$ (GeV)	$C_i(p_T^{t\bar{t}})$	$y_{t\bar{t}}$	$C_i(y_{t\bar{t}})$	$m_{t\bar{t}}$ (GeV)	$C_i(m_{t\bar{t}})$
$[0,30)$	2.600±0.004	$[-2.5,-1.5)$	5.792±0.039	$[340,380)$	2.007±0.005
$[30,80)$	2.027±0.003	$[-1.5,-1.0)$	2.636±0.009	$[380,470)$	2.546±0.004
$[80,170)$	2.139±0.005	$[-1.0,-0.5)$	2.104±0.005	$[470,620)$	2.478±0.005
$[170,300)$	2.230±0.010	$[-0.5,0.0)$	1.967±0.004	$[620,820)$	2.336±0.008
		$[0.0,0.5)$	1.962±0.004	$[820,1100)$	2.194±0.014
		$[0.5,1.0)$	2.097±0.005	$[1100,1600)$	2.077±0.028
		$[1.0,1.5)$	2.641±0.009		
		$[1.5,2.5)$	5.781±0.039		

4 Data vs Theory Predictions

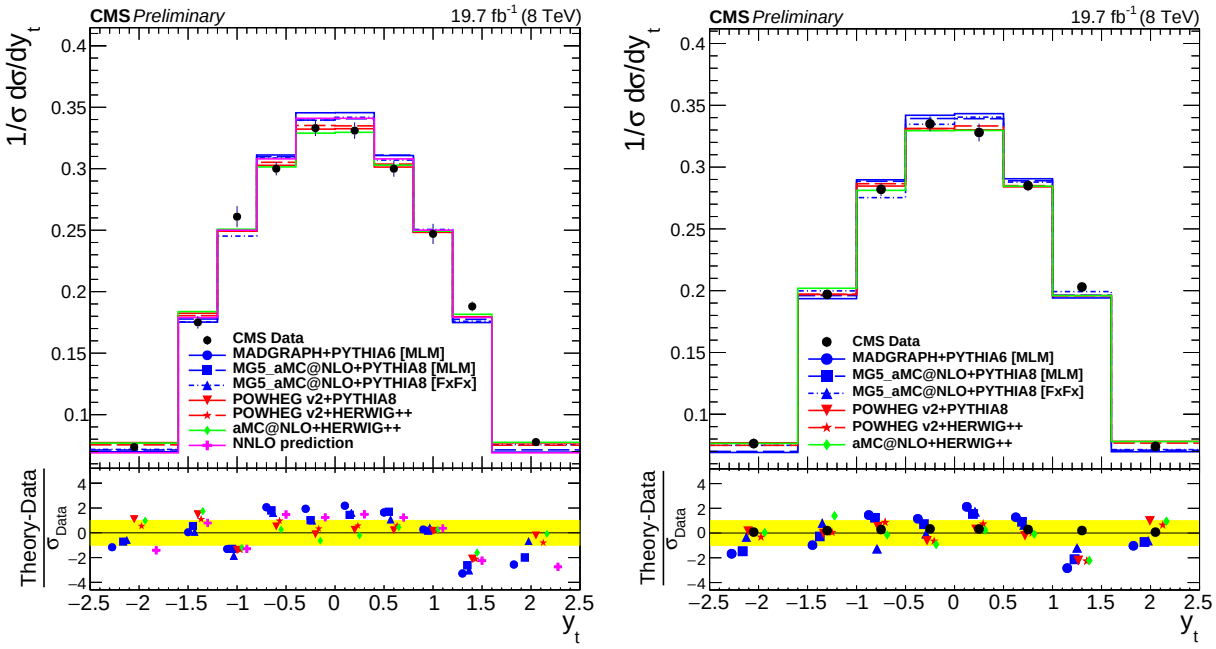
Figures 1-5 display comparisons of the unfolded data in the lepton+jets and the dilepton channels for the five kinematic variables at the parton-level: p_{T}^{t}, y_t, $p_{\mathrm{T}}^{t\bar{t}}$, $y_{t\bar{t}}$, and $m_{t\bar{t}}$. Each distribution is compared to the five different generator configurations described in Table 1 plus the reference MADGRAPH +PYTHIA 6 distribution used for the unfolding. The transverse momentum and rapidity distributions of the top quarks are also compared to the NNLO predictions described in [29]. All simulation configurations considered provide, in general, a good description of the data. POWHEG v2+HERWIG++ describes the top-quark p_{T} well, for both channels, while in the other MC setups discrepancies are observed in both the dilepton and the lepton+jets channels (Figure 1), with the exception of the FXFX configuration which shows good agreement in the dilepton channel.The complete NNLO QCD calculations [29] indicate that these deviations could be explained by missing higher order effects given that the level of agreement with the data is improved using this prediction. All MC configurations and the NNLO prediction provide a good description of the top-quark rapidity distributions (Figures 2). The rapidity of the $\mathrm{t\bar{t}}$ system (Figure 4) is also described well by all MC configurations. The transverse momentum of the $\mathrm{t\bar{t}}$ system, $p_{\mathrm{T}}^{t\bar{t}}$, is described well, however, deviations up to $\sim$15% are observed for MG5_aMC@NLO+PYTHIA 8 (MLM) at $p_{\mathrm{T}}^{t\bar{t}}$ $\sim$150 GeV in the lepton+jets channel, for MG5_aMC@NLO+PYTHIA 8 (FxFx) for all $p_{\mathrm{T}}^{t\bar{t}}$ bins and POWHEG v2+HERWIG++ for $p_{\mathrm{T}}^{t\bar{t}}$ >80 GeV in the dilepton channel (Figure 3). The invariant mass of the top-quark pair system is described well up to $\sim$ 800 GeV in the lepton+jets channel for all MC configurations. The only exceptions are POWHEG v2+HERWIG++ which shows discrepancies up to $\sim$ 15% and aMC@NLO +HERWIG++ that does not describe the data well for $m_{t\bar{t}} > 1.1$ TeV (Figure 5). In the dilepton channel, MG5_aMC@NLO+PYTHIA 8 (MLM) and POWHEG v2+PYTHIA 8 show a discrepancy of about 10% with respect to data for $m_{t\bar{t}} < 375$ GeV, MG5_aMC@NLO+PYTHIA 8 (FxFx) underestimates the data by $\sim$ 20% at $m_{t\bar{t}} \sim 900$ GeV.



(a) lepton+jets

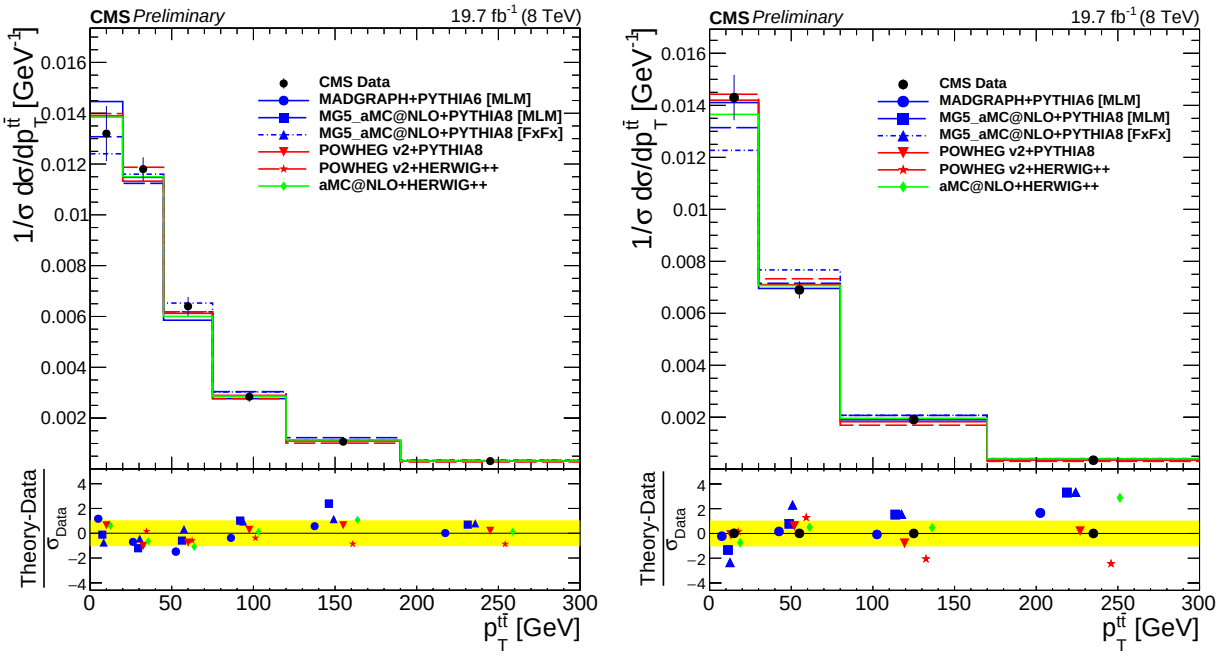
(b) dilepton

Figure 1: Normalized $t\bar{t}$ cross section in bins of p_T^t in data and MC at the parton-level. The yellow band indicates the 1σ difference of data from theory predictions. In the lepton+jets channel, the NNLO predictions are also shown.



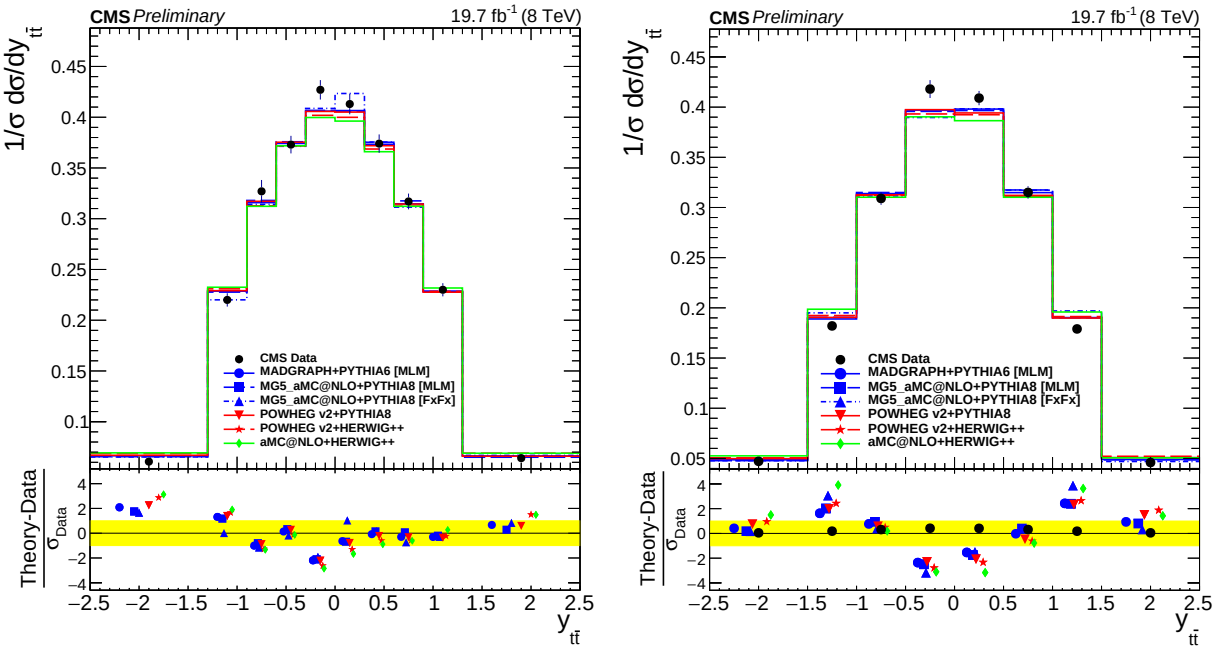
(a) lepton+jets (b) dilepton

Figure 2: Normalized $t\bar{t}$ cross section in bins of y_t in data and MC at the parton-level. The yellow band indicates the 1σ difference of data from theory predictions. In the lepton+jets channel, the NNLO predictions are also shown.



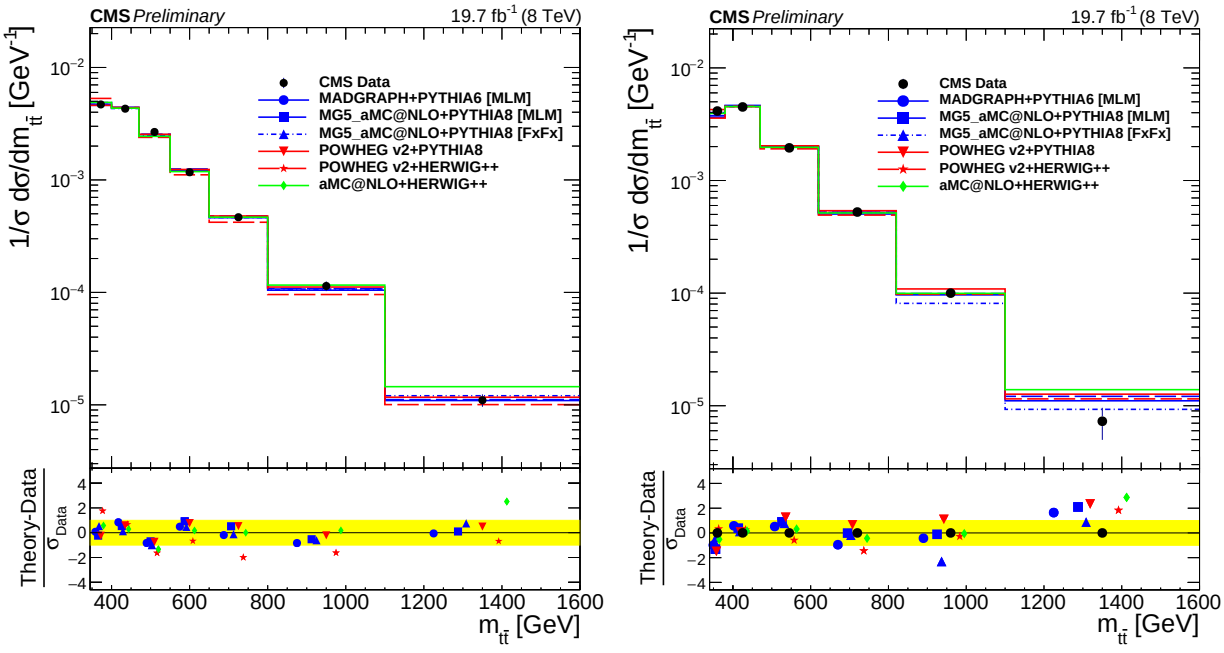
(a) lepton+jets (b) dilepton

Figure 3: Normalized $t\bar{t}$ cross section in bins of $p_T^{t\bar{t}}$ in data and MC at the parton-level. The yellow band indicates the 1σ difference of data from theory predictions.



(a) lepton+jets (b) dilepton

Figure 4: Normalized $t\bar{t}$ cross section in bins of $y_{t\bar{t}}$ in data and MC at the parton-level. The yellow band indicates the 1σ difference of data from theory predictions.



(a) lepton+jets (b) dilepton

Figure 5: Normalized $t\bar{t}$ cross section in bins of $m_{t\bar{t}}$ in data and MC at the parton-level. The yellow band indicates the 1σ difference of data from theory predictions.

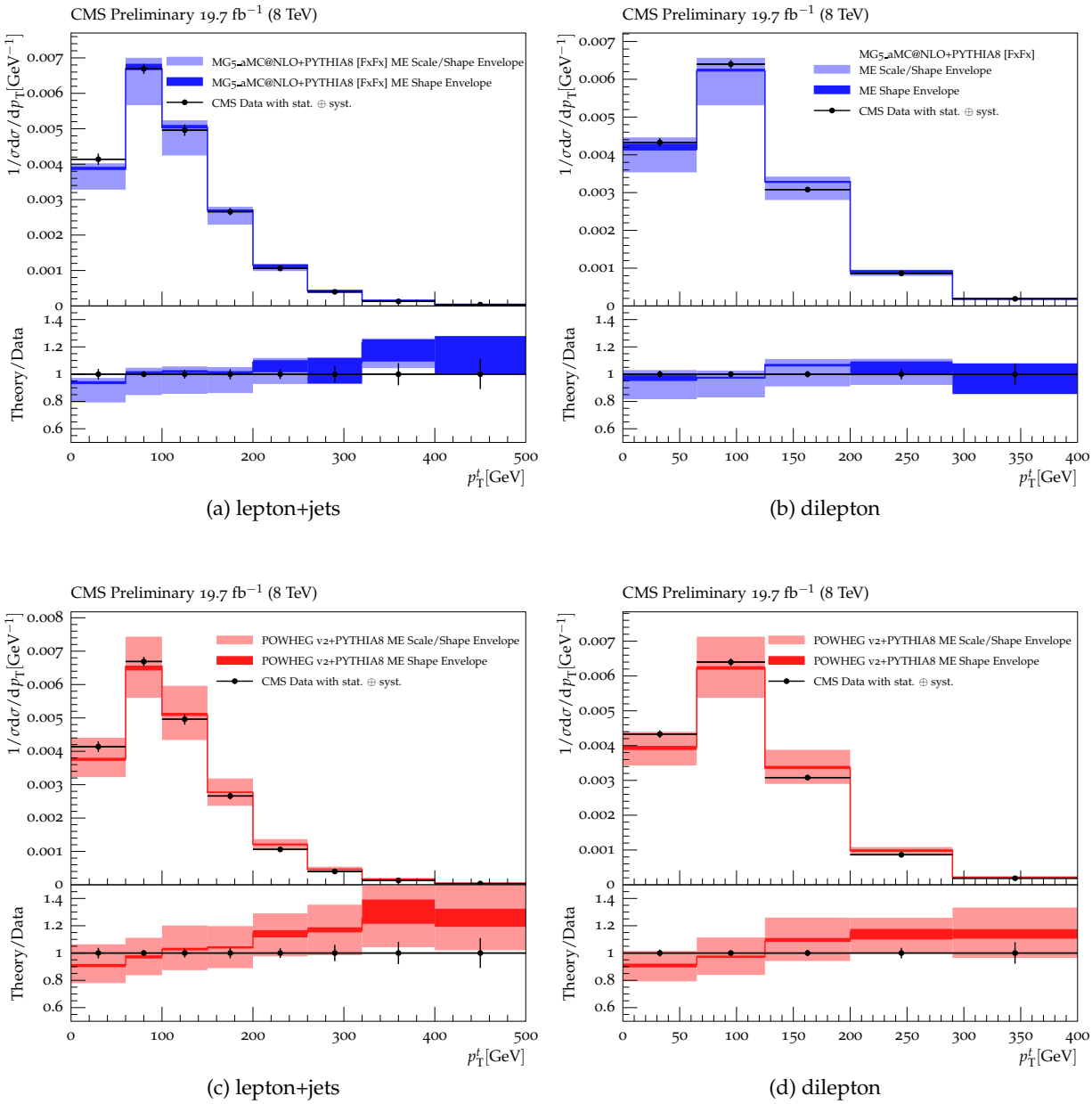
5 Renormalization and Factorization Scale Variations

Uncertainty in the choice of the renormalization and factorization scale leads to one of the significant theoretical uncertainties in the prediction of $t\bar{t}$ production. To investigate and quantify the effects on the predictions, the following matrix element scale variations have been considered around the central scale for each NLO simulation matched with PYTHIA 8:

- μ_R^{down} = 0.5: lower bound for renormalization scale variations.
- μ_R^{up} = 2.0: upper bound for renormalization scale variations .
- μ_F^{down} = 0.5 : lower bound for factorization scale variations.
- μ_F^{up} = 2.0 : upper bound for factorization scale variations.

All seven possible combinations are considered, excluding the most extreme cases of $\mu_{R,F} = \mu(2, 0.5)$ and $\mu_{R,F} = \mu(0.5, 2)$ which could potentially result in uncontrolled large logarithms in resummation (see e.g. [30, 31]). The scale varied differential cross sections obtained using the ME weights in MG5_aMC@NLO+PYTHIA 8 (with FxFx merging) and POWHEG v2+PYTHIA 8 are shown in Figures 6-10. The bands shown in the distributions represent the envelope of maximum and minimum bin-by-bin deviations from the nominal scenario, $\mu_{R,F} = \mu = 1$, either considering only the shape variation or both normalization scale factor and shape (respectively shown as a darker and a lighter band), while the error bars represent the total uncertainty in the data.

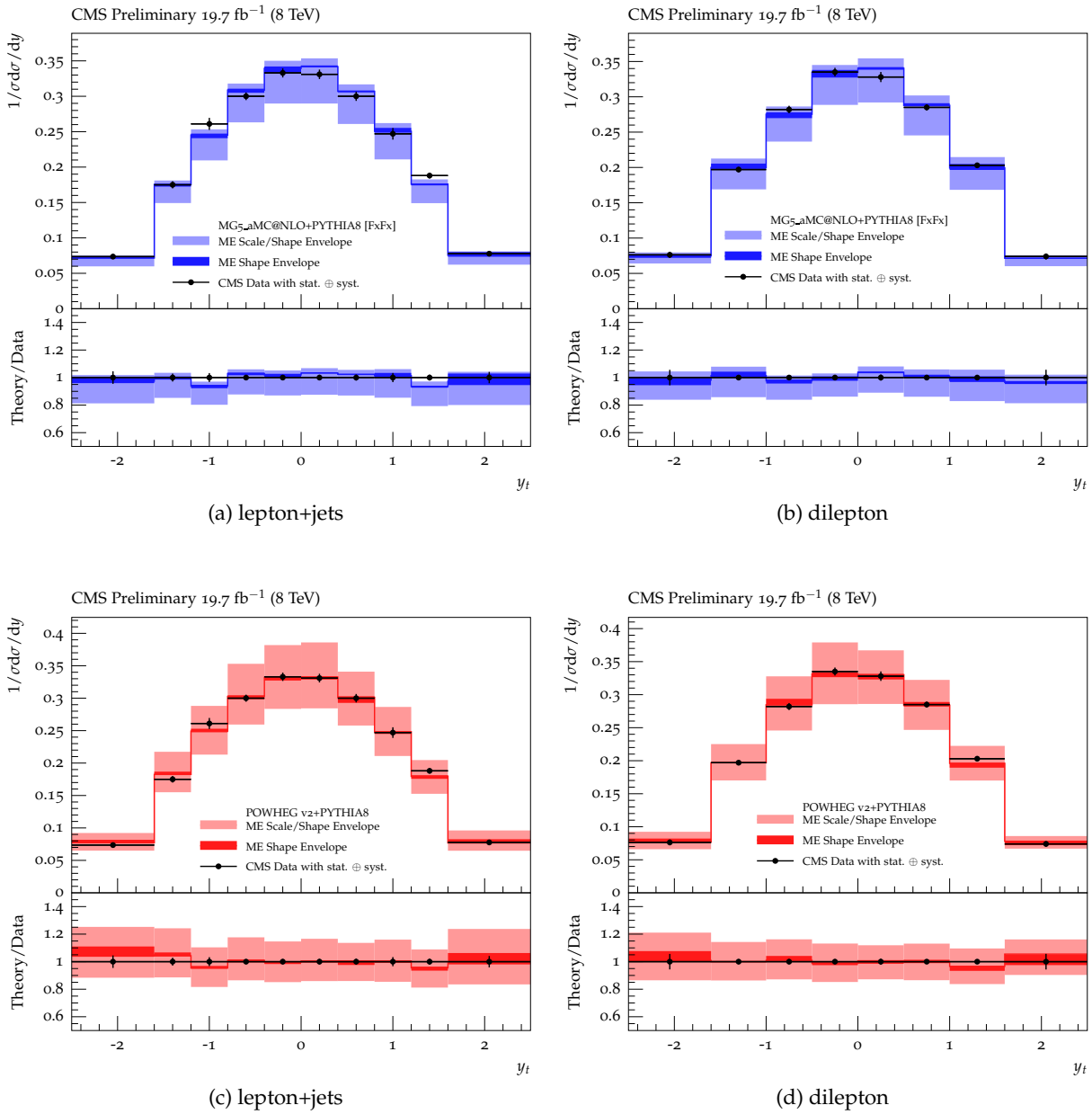
For MG5_aMC@NLO+PYTHIA 8 (with FxFx merging), the scale variations are asymmetric. The theory predictions with the ME scale uncertainties agree well with data except $m_{t\bar{t}} \sim$900 GeV in the dilepton channel. For POWHEG v2+PYTHIA 8, the scale variations are symmetric in both directions and in general larger than for the MG5_aMC@NLO+PYTHIA 8 (with FxFx merging) sample. The theory predictions agree with data within the ME scale uncertainties. In its full NLO calculation, POWHEG v2+PYTHIA 8 includes first real and virtual corrections, while, MG5_aMC@NLO+PYTHIA 8 calculates up to 2 real parton radiations merged at NLO including the first two real and virtual corrections. Therefore, for MG5_aMC@NLO+PYTHIA 8, the scale variations are expected to be smaller as is observed in Figures 6-10. Moreover, it can be seen that this uncertainty includes changes in both shape and normalization (scale). Renormalizing away the change in scale for all the considered variations, distributions end up with a quite small shape only uncertainty. For a more complete assessment of uncertainties due to the scale choices, one should also consider parton shower scale variations [32], however, this is outside the scope of the current study.



(a) lepton+jets (b) dilepton

(c) lepton+jets (d) dilepton

Figure 6: Normalized $t\bar{t}$ cross section in bins of p_T^t in data and MG5_aMC@NLO sample with FxFx merging (top) and POWHEG v2+PYTHIA 8 sample (bottom) at the parton-level. The data points are shown with total error bars and the envelope of different factorization and renormalization assumptions in the matrix elements with a band for both scale+shape and only shape variations.



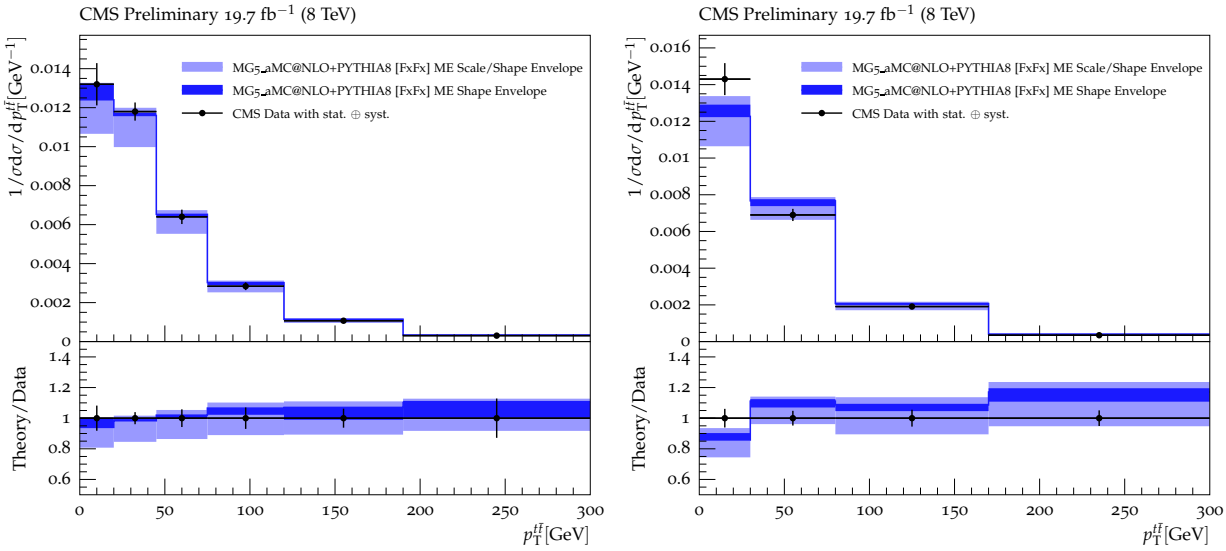
(a) lepton+jets

(b) dilepton

(c) lepton+jets

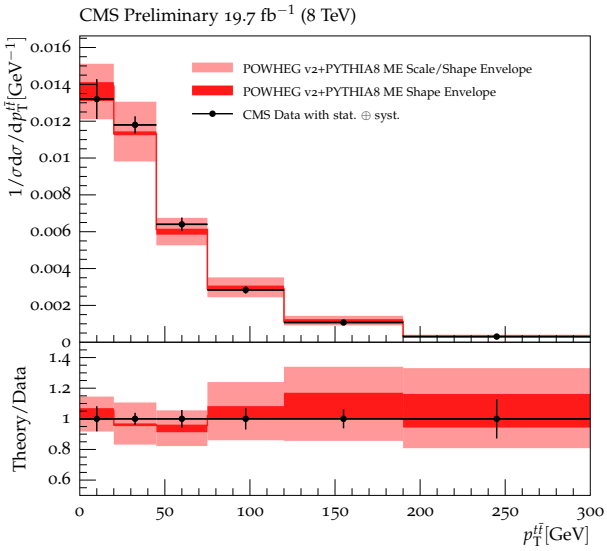
(d) dilepton

Figure 7: Normalized $\mathrm{t\bar{t}}$ cross section in bins of y_t in data and MG5_aMC@NLO sample with FxFx merging (top) and POWHEG v2+PYTHIA 8 sample (bottom) at the parton-level. The data points are shown with total error bars and the envelope of different factorization and renormalization assumptions in the matrix elements with a band for both scale+shape and only shape variations.

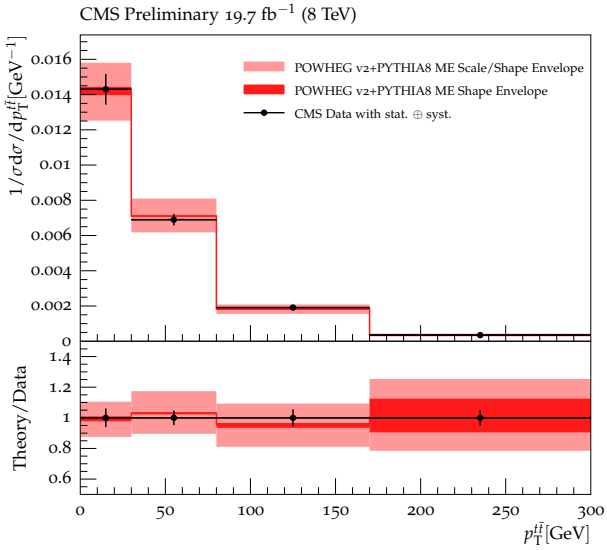


(a) lepton+jets

(b) dilepton

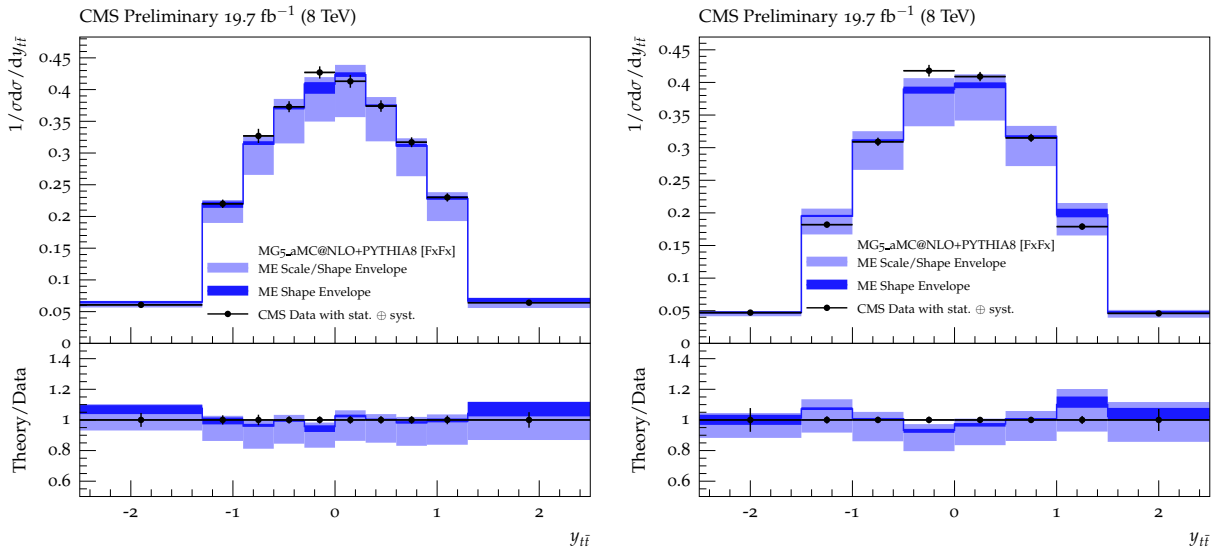


(c) lepton+jets



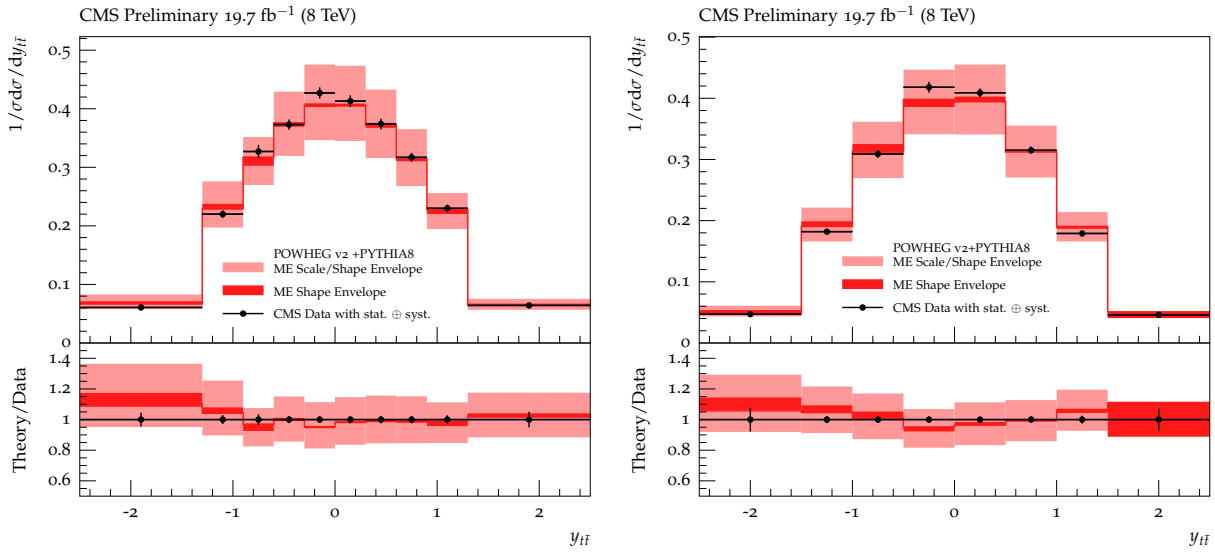
(d) dilepton

Figure 8: Normalized $t\bar{t}$ cross section in bins of $p_T^{t\bar{t}}$ in data and MG5_aMC@NLO sample with FxFx merging (top) and POWHEG v2+PYTHIA 8 sample (bottom) at the parton-level. The data points are shown with total error bars and the envelope of different factorization and renormalization assumptions in the matrix elements shown with a band for both scale+shape and only shape variations.



(a) lepton+jets

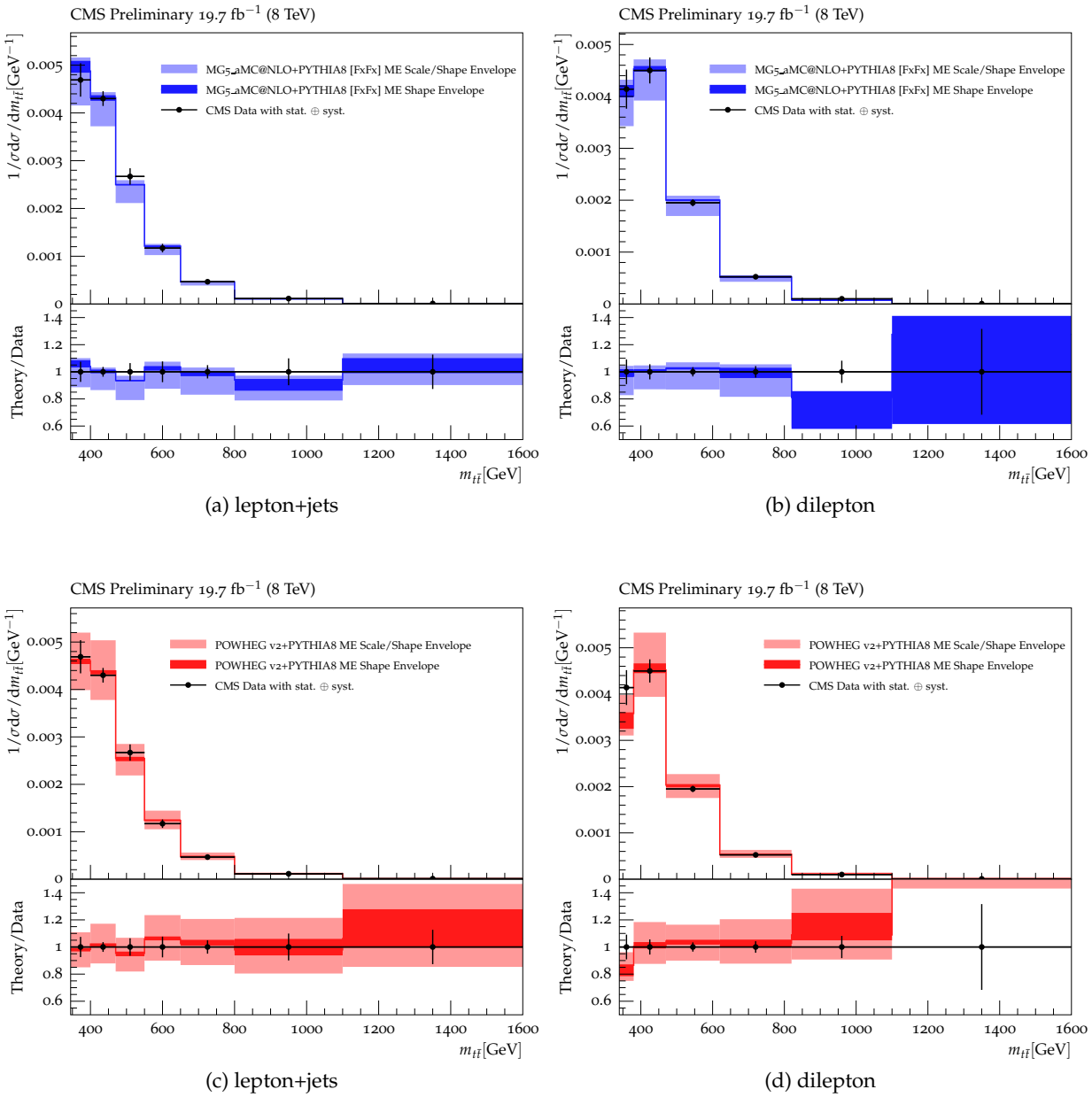
(b) dilepton



(c) lepton+jets

(d) dilepton

Figure 9: Normalized $t\bar{t}$ cross section in bins of $y_{t\bar{t}}$ in data and MG5_aMC@NLO sample with FxFx merging (top) and POWHEG v2+PYTHIA 8 sample (bottom) at the parton-level. The data points are shown with total error bars and the envelope of different factorization and renormalization assumptions in the matrix elements with a band for both scale+shape and only shape variations.



(a) lepton+jets (b) dilepton

(c) lepton+jets (d) dilepton

Figure 10: Normalized $t\bar{t}$ cross section in bins of $m_{t\bar{t}}$ in data and MG5_aMC@NLO sample with FxFx merging (top) and POWHEG v2+PYTHIA 8 sample (bottom) at the parton-level. The data points are shown with total error bars and the envelope of different factorization and renormalization assumptions in the matrix elements with a band for both scale+shape and only shape variations.

6 Summary and Conclusions

Studies are presented comparing $\mathrm{t\bar{t}}$ differential distribution data, obtained with the CMS detector at $\sqrt{s} = 8$ TeV in the lepton+jets and dilepton channels, to state-of-the-art matrix element Monte Carlo event generators at next-to-leading order in QCD, used in conjunction with the parton showers PYTHIA 8 and HERWIG++. The predictions of these improved simulation tools are compared to unfolded data distributions. A RIVET module, providing all the results discussed in this manuscript, is made publicly available. The NLO theory predictions do not always improve the agreement with data with respect to the LO ones, but from this study we can conclude that any deviation observed with respect to the central value of the NLO theory predictions is covered by the uncertainty assigned to the choice of the renormalization and factorization scales at the matrix element level.

References

[1] J. Alwall, M. Herquet, and F. Maltoni, "MadGraph 5: Going Beyond", *JHEP* **6** (2011) 128, doi:10.1007/JHEP06(2011)128.

[2] T. Sjostrand, S. Mrenna, and P. Skands, "PYTHIA 6.4 physics and manual", *Journal of High Energy Physics* **2006** (2006), no. 05, 026.

[3] CMS Collaboration, "Underlying Event Tunes and Double Parton Scattering", Technical Report CMS-PAS-GEN-14-001, CERN, Geneva, 2014.

[4] R. Field, "Early LHC Underlying Event Data - Findings and Surprises", in *Hadron collider physics. Proceedings, 22nd Conference, HCP 2010, Toronto, Canada, August 23-27, 2010*. 2010. arXiv:1010.3558.

[5] J. Pumplin et al., "New Generation of Parton Distributions with Uncertainties from Global QCD Analysis", *Journal of High Energy Physics* **2002** (2002), no. 07, 012.

[6] CMS Collaboration, "Measurement of the differential cross section for top quark pair production in pp collisions at $\sqrt{s}$ = 8 TeV", (2015). arXiv:1505.04480.

[7] J. Alwall, R. Frederix, S. Frixione et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.

[8] P. Nason, "A New method for combining NLO QCD with shower Monte Carlo algorithms", *JHEP* **11** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:0409146.

[9] 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.

[10] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX ", *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.

[11] A. Buckley, J. Butterworth, D. Grellscheid et al., "Rivet user manual", *Comput. Phys. Commun.* **184** (2013) 2803, arXiv:1003.0694.

[12] M. Dobbs and J. Hansen, "The HepMC C++ Monte Carlo event record for High Energy Physics", *Comput. Phys. Commun.* **134** (2001) 41.

[13] P. Artoisenet, R. Frederix, O. Mattelaer, and R. Rietkerk, "Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations", *Journal of High Energy Physics* **2013** (2013), no. 3, doi:10.1007/JHEP03(2013)015.

[14] Z. Was, "TAUOLA the library for tau lepton decay, and KKMC/KORALB/KORALZ/... status report", *Nucl.Phys.Proc.Suppl.* **98** (2001) 96–102, doi:10.1016/S0920-5632(01)01200-2, arXiv:hep-ph/0011305.

[15] S. Agostinelli et al., "Geant4: a simulation toolkit", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** (2003), no. 3, 250 – 303, doi:http://dx.doi.org/10.1016/S0168-9002(03)01368-8.

[16] S. Frixione, P. Nason, and G. Ridolfi, "A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction", *JHEP* **09** (2007) 126, doi:10.1088/1126-6708/2007/09/126, arXiv:0707.3088.

[17] T. Sjostrand, S. Mrenna, and P. Skands, "A Brief Introduction to PYTHIA 8.1", *Comput. Phys. Commun.* **178** (2008) 852, doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820.

[18] "Pythia8, Sample Main Programs page". http://home.thep.lu.se/torbjorn/pythia82html/SampleMainPrograms.html.

[19] M. Bahr et al., "Herwig++ Physics and Manual", *Eur.Phys.J.C* **58** (2008) 639, doi:10.1140/epjc/s10052-008-0798-9, arXiv:0803.0883.

[20] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO", *JHEP* **12** (2012) 061, doi:10.1007/JHEP12(2012)061, arXiv:1209.6215.

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

[22] J. Pumplin et al., "New Generation of Parton Distributions with Uncertainties from Global QCD Analysis", *JHEP* **07** (2002) 012, doi:10.1088/1126-6708/2002/07/012, arXiv:0201195.

[23] P. Skands, S. Carrazza, and J. Rojo, "Tuning PYTHIA 8.1: the Monash 2013 Tune", *Eur.Phys.J.C* **74** (2014) 3024, doi:10.1140/epjc/s10052-014-3024-y, arXiv:1404.5630.

[24] R. D. Ball et al., "Parton distributions with LHC data", *Nucl.Phys. B* **867** (2013) 244, doi:10.1016/j.nuclphysb.2012.10.003, arXiv:1207.1303.

[25] H. L. Lai et al., "New parton distributions for collider physics", *Phys.Rev.D* **82** (2010) 074024, doi:10.1103/PhysRevD.82.074024, arXiv:1007.2241.

[26] R. D. Ball et al., "Parton distributions for the LHC Run II", *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.

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

[28] Particle Data Group Collaboration, "Review of Particle Physics", *Chin. Phys.* **C38** (2014) 090001, doi:10.1088/1674-1137/38/9/090001.

[29] M. Czakon, D. Heymes, and A. Mitov, "High-precision differential predictions for top-quark pairs at the LHC", arXiv:1511.00549.

[30] M. Cacciari et al., "The t anti-t cross-section at 1.8-TeV and 1.96-TeV: A Study of the systematics due to parton densities and scale dependence", *JHEP* **04** (2004) 068, doi:10.1088/1126-6708/2004/04/068, arXiv:hep-ph/0303085.

[31] S. Catani, D. de Florian, M. Grazzini, and P. Nason, "Soft gluon resummation for Higgs boson production at hadron colliders", *JHEP* **07** (2003) 028, doi:10.1088/1126-6708/2003/07/028, arXiv:hep-ph/0306211.

[32] B. Cooper et al., "Importance of a consistent choice of alpha(s) in the matching of AlpGen and Pythia", *Eur. Phys. J.* **C72** (2012) 2078, doi:10.1140/epjc/s10052-012-2078-y, arXiv:1109.5295.

A Appendix: Matrix Element Configurations

A.1 Matrix Element Settings for the MG5_aMC@NLO LO [MLM] Sample

```
MadGraph5_aMC@NLO VERSION 2.2.2                 2014-11-06
set group_subprocesses Auto
set ignore_six_quark_processes False
set loop_optimized_output True
set complex_mass_scheme False
import model sm-ckm_no_b_mass
define p = g u c d s u~ c~ d~ s~
define j = g u c d s u~ c~ d~ s~
define l+ = e+ mu+
define l- = e- mu-
define vl = ve vm vt
define vl~ = ve~ vm~ vt~
define p = p b b~
define j = j b b~
generate p p > t t~ @0
add process p p > t t~ j @1
add process p p > t t~ j j @2
add process p p > t t~ j j j @3

PDF set = 263000 (lhapdf number)

Renormalization and factorization scales

Matching:
MLM (ickkw = 1)
ktscheme = Durham kT
scale factor for QCD emission vx: alpsfact = 1
cluster only according to channel diag: chcluster = false
highest quark flavor for a_s reweight: asrwgtflavor = 5
Change the way clustering information: lhe_version = 3.0
                                pass to shower.

Automatic ptj and mjj cuts if xqcut > 0: auto_ptj_mjj = true

BW cutoff (M+/-bwcutoff*Gamma): bwcutoff = 15 GeV

pt/E/eta/dr/mij cuts on decay products: cut_decays = false
(note that etmiss/ptll/ptheavy/ht/sorted cuts always apply)

Number of helicities used per event: nhel = 0
```

```
max rap for the jets: etaj = 5

min DeltaR distance between gamma and jet: draj = 0.1
min DeltaR distance between gamma and lepton: dral = 0.1

ktdurham = no cut

maximal pdg code for quark to be considered as a light jet:
Maximum jet pdg code: maxjetflavor = 5

Jet measure cuts:
minimum kt jet measure between partons: xqcut = 20 GeV
```

A.2 Matrix Element Settings for the MG5_aMC@NLO NLO [FxFx] sample

```
MadGraph5_aMC@NLO   VERSION 2.2.1                 2014-09-25
set group_subprocesses Auto
set ignore_six_quark_processes False
set loop_optimized_output True
set complex_mass_scheme False
import model loop_sm-ckm_no_b_mass
define p = g u c d s u~ c~ d~ s~
define j = g u c d s u~ c~ d~ s~
define l+ = e+ mu+
define l- = e- mu-
define vl = ve vm vt
define vl~ = ve~ vm~ vt~
define p = p b b~
define j = j b b~
generate p p > t t~ [QCD] @0
add process p p > t t~ j [QCD] @1
add process p p > t t~ j j [QCD] @2

Number of points per itegration channel (ignored for aMC@NLO runs):
  0.01   = req_acc_FO       ! Required accuracy

PDF choice: this automatically fixes also alpha_s(MZ) and its evol:
PDF set = 260000 (lhapdf number)

Include the NLO Monte Carlo subtr. terms for the following parton:
PYTHIA8   = parton_shower

Renormalization and factorization scales:
The transverse masses of all final state particles and partons.

Reweight flags to get scale dependence:
reweight to get scale dependence: reweight_scale = true
lower bound for ren scale variations: rw_Rscale_down = 0.5
upper bound for ren scale variations: rw_Rscale_up = 2.0
```

```
lower bound for fact scale variations: rw_Fscale_down = 0.5
upper bound for fact scale variations: rw_Fscale_up = 2.0

Merging:
FxFx: ickkw = 3

BW cutoff (M+/-bwcutoff*Gamma): 15

Cuts on the jets
Jet clustering is performed by FastJet.
When matching to a parton shower, these generation cuts should be
considerably softer than the analysis cuts:
FastJet jet algorithm: kT
jetradius = 1.0
Min jet transverse momentum = 10 GeV
```

A.3 Matrix Element Settings for the POWHEG v2 Samples

```
pdf set for hadrons (LHA numbering): lhans1 = 10800

mass of heavy quark in GeV: qmass = 172.5 GeV

the order the maximum number of the following particles(antiparticles)
in the final state: e  mu tau up charm: topdecaymode = 22222
22222 means all decays (up to 2 units of everything)

Parameters for the generation of spin correlations in ttbar decays:
tdec/wmass = 80.398 GeV  ! W mass for top decay
tdec/wwidth = 2.141 GeV
tdec/bmass = 4.16 GeV
tdec/twidth = 1.31 GeV
tdec/elbranching = 0.108
tdec/emass = 0.00051 GeV
tdec/mumass = 0.1057 GeV
tdec/taumass = 1.777 GeV
tdec/dmass = 0.100 GeV
tdec/umass = 0.100 GeV
tdec/smass = 0.200 GeV
tdec/cmass = 1.28 GeV
tdec/sin2cabibbo = 0.051

# calls for initializing the integration grid: ncall1 = 10000
# iterations for initializing the integration grid: itmx1 = 5
# calls for computing the integral and finding upper bound: ncall2 = 100000
# iterations for computing the integral and finding upper bound: itmx2 = 5
# folds on x integration: foldcsi = 1
# folds on y integration: foldy = 1
# folds on phi integration: foldphi = 1
# bbarra calls to setup norm of upper bounding function: nubound = 100000
```

```
<= 10, normalization of upper bounding function in iunorm X iunorm square in y,
       log(m2qq): iymax = 1
<= 10, normalization of upper bounding function in iunorm X iunorm square in y,
       log(m2qq): ixmax = 1
increase upper bound for radiation generation = xupbound = 2
```

A.4 Matrix Element Settings for the aMC@NLO (+HERWIG++) Sample

```
MadGraph 5 2.0.0.beta3 2013-02-14
set group_subprocesses Auto
set ignore_six_quark_processes False
set loop_optimized_output True
set gauge unitary
set complex_mass_scheme False
import model loop_sm
define l+ = e+ mu+
define l- = e- mu-
define vl = ve vm vt
define vl~ = ve~ vm~ vt~
define p = g u c d s u~ c~ d~ s~
define l+ =  e+ mu+ ta+
define l- = e- mu- ta-
define tt = t t~
define j = g u c d s u~ c~ d~ s~
generate p p > t t~ [QCD]

PDF choice: this automatically fixes also alpha_s(MZ) and its evol.:
pdlabel = cteq6_m

Include the NLO Monte Carlo subtr. terms for the following parton shower:
parton_shower = HERWIG6

Renormalization and factorization scales:
The transverse masses of all final state particles and partons.

Reweight flags to get scale dependence:
reweight to get scale dependence: reweight_scale = true
lower bound for ren scale variations: rw_Rscale_down = 0.5
upper bound for ren scale variations: rw_Rscale_up = 2.0
lower bound for fact scale variations: rw_Fscale_down = 0.5
upper bound for fact scale variations: rw_Fscale_up = 2.0

BW cutoff (M+/-bwcutoff*Gamma): bwcutoff = 15 GeV

Cuts on the jets
Jet clustering is performed by FastJet.
When matching to a parton shower, these generation cuts should be
considerably softer than the analysis cuts:
FastJet jet algorithm: kT
jetradius = 0.7
```

```
Min jet transverse momentum = 10 GeV

maximal pdg code for quark to be considered as a light jet:
Maximum jet pdg code: maxjetflavor = 5
```

B Parton Shower Settings

B.1 PYTHIA 6 Settings for the MADGRAPH Sample

```
Tune = Z2*
ExternalDecays = Tauola with TauolaPolar, TauolaDefaultInputCards

pythiaUESettingsBlock:
MSEL=0
b quark mass: PMAS(5,1)=4.8 GeV
t quark mass: PMAS(6,1)=172.5 GeV
Fragmentation/hadronization on or off: MSTJ(1)=1
Parton showering on or off: MSTP(61)=1

jetMatching:
scheme = Madgraph
mode = auto
MEMAIN_nqmatch = 5
```

B.2 PYTHIA 8 Settings for the MG5_aMC@NLO[MLM] Sample

```
Tune = CUEP8M1

JetMatching:
setMad = off
scheme = 1
merge = on
jetAlgorithm = 2
etaJetMax = 5.
coneRadius = 1.
slowJetPower = 1
actual merging scale: qCut = 60. GeV
5-flavour scheme: nQmatch = 5
number of partons in born matrix element for highest multiplicity:
       nJetMax = 3
off for MLM matching, turn on for shower-kT matching: doShowerKt = off
```

B.3 PYTHIA 8 Settings for the MG5_aMC@NLO [FxFx] Sample

```
Tune = CUEP8M1

jetMatching:
setMad = off
scheme = 1
merge = on
```

```
jetAlgorithm = 2
etaJetMax = 999.
coneRadius = 1.
slowJetPower = 1
actual merging scale: qCut = 30. GeV
doFxFx = on
this must match the ptj cut in the lhe generation step: qCutME = 10. GeV
5-flavour scheme: nQmatch = 5
number of partons in born matrix element for highest multiplicity:
       nJetMax = 2
```

B.4 PYTHIA 8 Settings for the POWHEG v2 Sample

```
Tune = CUEP8M1

pythia8PowhegEmissionVetoSettings:
Number of final state particles(BEFORE THE DECAYS) in the LHE
       other than emitted extra parton : POWHEG:nFinal = 2
POWHEG:veto = 1
POWHEG:pTdef = 1
POWHEG:emitted = 0
POWHEG:pTemt = 0
POWHEG:pThard = 0
POWHEG:vetoCount = 100
SpaceShower:pTmaxMatch = 2
TimeShower:pTmaxMatch = 2
```

B.5 HERWIG++ Settings for the POWHEG v2 Sample

```
tune = EE5C

Shower PDF matching with the tune: hwpp_pdf_CT10
Showering LHE files from MadGraph5_aMC@NLO.
          Use the same PDF for the shower as for
          the hard subprocess afore: hwpp_LHE_Powheg
Switch off ME corrections while showering LHE files
           as recommended by Herwig++ authors: hwpp_MECorr_Off
```

B.6 HERWIG++ Settings for the aMC@NLO Sample

```
tune = EE5C

Shower PDF matching with the tune: hwpp_pdf_CTEQ6L1
Showering LHE files from MadGraph5_aMC@NLO.
          Use the same PDF for the shower as for
          the hard subprocess afore: hwpp_LHE_MadGraph
Switch off ME corrections while showering LHE files
           as recommended by Herwig++ authors: hwpp_MECorr_Off
```