From Particle Flow to Colour Flow

Author:
Henrik Löfberg
Supervisors:
Jürg Eugster, Pedro Silva, Benjamin Stieger, Martijn Mulders

Abstract

A deeper understanding of the underlying event in $t\bar{t}$ pair production is expected to improve the current uncertainty on the measurements of the top quark mass. By selecting events with an electron, a muon and two b-tagged jets a high purity signal is obtained. The main properties of the underlying event are isolated and compared between data and different Pythia Monte Carlo Tunes. Discrepancies between the total number of charged particles for different models is observed. Furthermore a contribution of colour reconnection to the modeling of the average transverse momentum of the charged particles is identified.

Introduction

The top quark is the heaviest of the elementary particles. In contrast to all other quarks the top quark does not hardronize. This yields a unique opportunity to study its properties, for example the mass [1].

In a proton-proton collision there are contributions from initial and final state radiation, multi parton interactions and beam remnants which is called the underlying event. Mass measurements of the top quark are affected by different uncertainties. One of them is the modeling of colour reconnection from the b-jets to the underlying event. This means that the uncertainty of the mass measurements of the top quark could be improved by better understanding the modeling of the underlying event. Furthermore the models describing the underlying event should be universal, that is independent of the energy scale of the hard process. Comparing the models against data should provide insight where the modeling can be improved.

Event Selection

In this study data from proton-proton collisions at $\sqrt{s} = 8$ TeV corresponding to an integrated luminosity of 19.7 fb^{-1} are analyzed. The particle flow algorithm is used for particle reconstruction [2]. Jets are reconstructed using

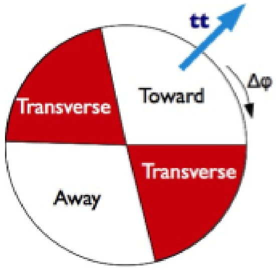

Figure 1: The regions defined in the transverse plane relatively to the $t\bar{t}$ direction.

the anti $-$ k$_T$ jet clustering algorithm with a distance parameter R $= 0.5$ and are required ot have a p_T greater than 20 GeV. [3].

Events are selected in data applying the following requirements:

1. Two hight transverse momenta isolated leptons $|\eta| < 2.5$, $p_T > 20$ GeV
2. Two b-tagged jets $|\eta| < 2.5$, $p_T > 30$ GeV [4]
3. One electron and one muon with invariant mass $m_{\ell\ell} > 12$ GeV

This is expected to yield a high purity $t\bar{t}$ candidate sample where $S/(S+B) \approx 96\%$ [5]. All the particles associated with the hard processes are removed, that is the leptons and all particles associated with the b-jets. A cut as mentioned above is applied on the charged particles and b-jets. The remaining particles are analyzed with the constraint that they are required to have $d_Z < 3$ cm, $d_0 < 1$ cm and a distance significance < 10 with respect to the primary vertex in the longitudinal and transverse directions. This cut is expected to efficiently remove charged particles originating from pileup collisions.

In order to define a reference axis in each event the $t\bar{t}$ momentum is reconstructed by adding the momenta of the leptons, the b-jets as well as the missing energy to compensate for the neutrinos created in the event.

$$\overrightarrow{p}_T^{t\bar{t}} = \overrightarrow{p}_T^{e} + \overrightarrow{p}_T^{\mu} + \overrightarrow{p}_T^{b_1} + \overrightarrow{p}_T^{b_2} + \overrightarrow{p}_T^{missing}$$

Regions are defined in the transverse plane with the help of the reconstructed transverse momentum direction of the $t\bar{t}$, as seen in fig 1.

1. Forward $0^\circ < |\Delta\phi| < 60^\circ$

2. Transverse $60^\circ < |\Delta\phi| < 120^\circ$

3. Away $120^\circ < |\Delta\phi|$

The variables of interest to study in these region are: the number of charged particles, the scalar sum of transverse momenta of the charged particles and the average transverse momentum per particle. Besides measuring the inclusive distribution the variables are profiled against the transverse momentum of $t\bar{t}$ and the angle $\Delta\phi$ between the particles and the $t\bar{t}$ axis. This is done in the three separate regions: towards, transverse and away as well as inclusively. The data is compared to different Pythia Monte Carlo tunings: Z2*, P11, P11-noCR, P11-TeV [6] [7]. Where P11-noCR is a variation of P11 where the colour reconnection is switched of and P11-TeV is a variation which is tuned to data from the Tevatron experiments.
The agreement between the data and the simulation is evaluated by computing the following χ^2 which is quoted on some of the figures: $\chi^2 = \sum_{i=1}^{n} \frac{(N_i - N_i^{exp})^2}{N_i + (\sigma_i^{exp})^2}$ where N_i (N_i^{exp}) is the number of observed (expected) events and σ_i^{exp} is the uncertainty associated to the expectations.

Results and Discussion

Figure 2 shows the distribution of the number of charged particles. One can observe that the simulation, Z2*, predicts a larger multiplicity than the data. Figure 3 (Left) shows the the mean number of charged particles per event. The simulation, P11, predicts a larger multiplicity than the data in all regions and for the whole spectrum of p_T of $t\bar{t}$, as observed in fig 3 (Left). Furthermore it is observed that all P11 tune variations, except P11-TeV overshoot for high transverse momentum as seen in fig 3 (Right), where the data to simulation of number of charged particles has been plotted. In fig 4 (Left) the nominal cut at 0.5 GeV is applied and in (Right) a cut at 1.0 Gev is applied. It is observed that there is a better agreement between data and simulation for a higher p_T cut of 1.0 GeV on the charged particles. In this report however effects from softer interactions are studied and therefore the p_T cut at 0.5 GeV is kept throughout this studie.

Figure 5 shows the distribution of the scalar sum of p_T of the particles which we denote p_T flux. It is observed that the agreement for the p_T flux between the Z2* simulation and data is good. In fig. 6 (Left) it is seen that the P11 tune variations tend to overshoot the data for systems close to rest.

The average p_T is observed to be described in slight deficit with respect to data by the Z2* tune in Fig. 8. This is a natural consequence of the dis-

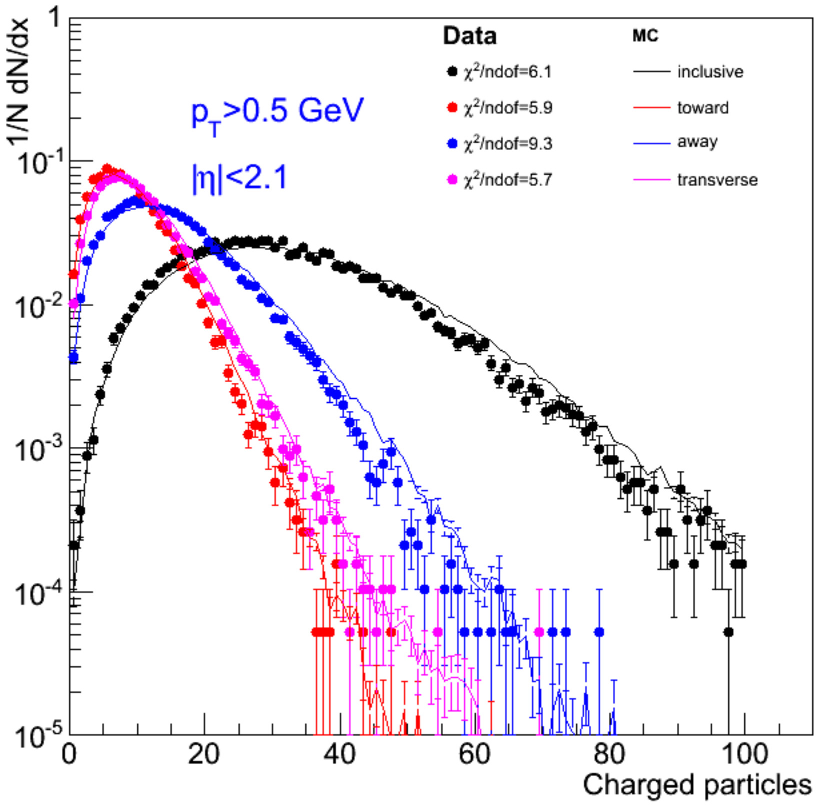

Figure 2: The distribution of number of charged particles for all three regions and inclusively. Using the Z2* tune.

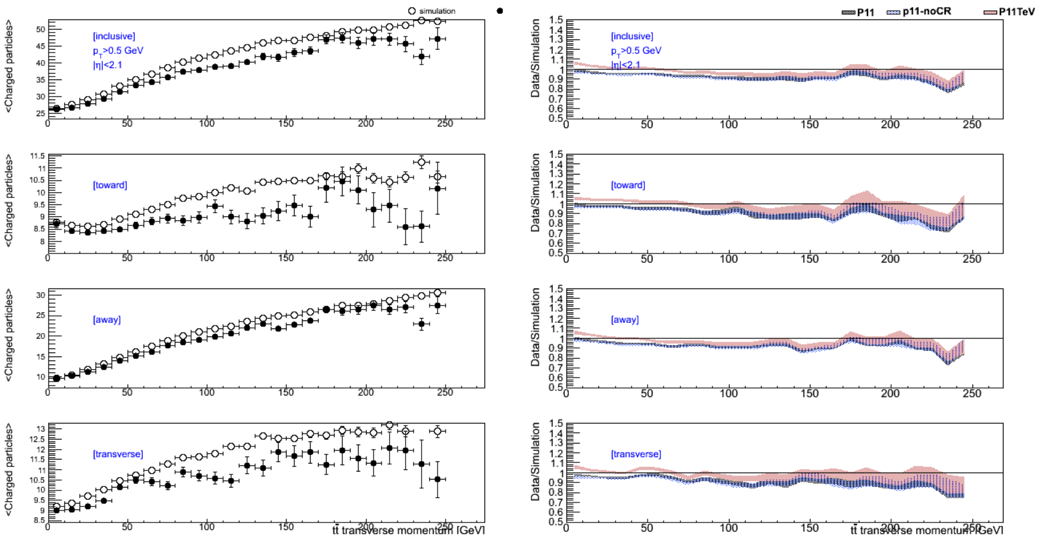

Figure 3: (Left) Mean number of charged particles profiled versus p_T of $t\bar{t}$. Using the Z2* tune.
(Right) Data to simulation ratio of the mean number of charged particles per event profiled versus p_T of $t\bar{t}$ for different P11 tune variations.

crepancy observed in the number of charged particles but a good agreement for the total p_T flux.

The ratio of data to simulation for the average transverse momentum per event shows that the P11-noCR undershoots the data significantly as seen in fig. 9 (Left). Also the undershoot seems to be independent of the boost in the transverse region. The undershoot, of P11-noCR, occurs in all regions but is observed to be less pronounced in the away region and strongest in the toward region, as shown in fig. 9 (Right)

Summary and Future Work

The Z2* Monte Carlo is observed to have better agreement for a cut at 1.0 GeV instead of 0.5 GeV. For the number of charged particles the simulation overestimates the data in all three regions and for high p_T of $t\bar{t}$ all P11 tune variations overshoots the data.
For p_T flux of the charged particles the agreement is observed to be good between Z2* and data. The P11 variations tend to overshoot the data for systems close to rest.
The P11-noCR model is observed to underestimate the data for the whole spectrum when considering the average p_T of the charged particles. The Z2*

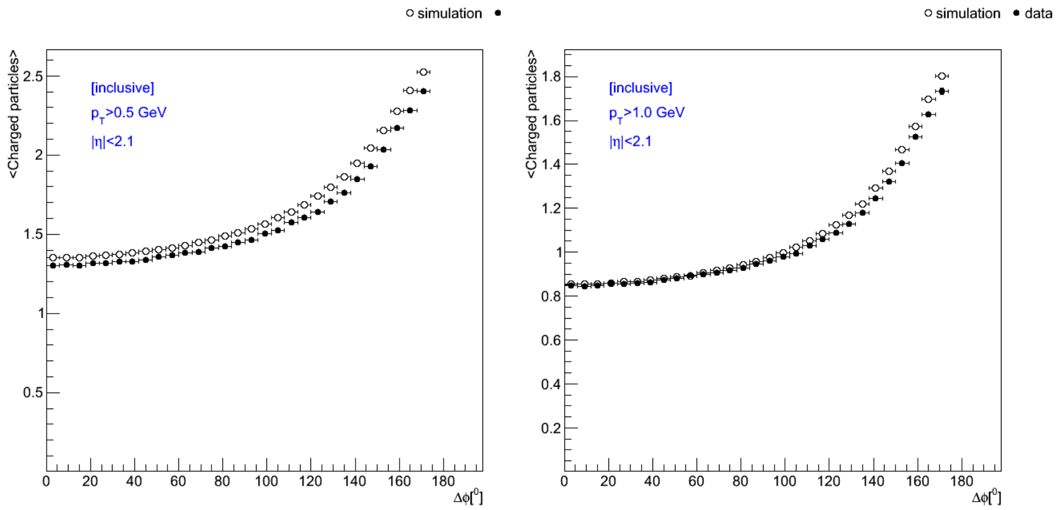

Figure 4: (Left) Mean number of charged particles profiled versus $\Delta\phi$ with a $p_T > 0.5$ GeV cut. Using the Z2* tune.
(Right) Mean number of charged particles profiled versus $\Delta\phi$ with a $p_T > 1.0$ GeV cut. Using the Z2* tune.

slightly underestimates the data.
In the future more variables of interest should be investigated for example the pull of the particles with respect to the jet and the beam. Also a deeper analysis should be made regarding the P11-noCR undershoot and how this could be used to calibrate our models.

References

[1] J. Beringer et al. Review of Particle Physics (RPP). *Phys.Rev.*, D86:010001, 2012.

[2] Commissioning of the particle-flow reconstruction in minimum-bias and jet events from pp collisions at 7 tev. Technical Report CMS-PAS-PFT-10-002, CERN, Geneva, 2010.

[3] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Anti-k(t) jet clustering algorithm. *JHEP*, 0804:063, 2008.

[4] Serguei Chatrchyan et al. Identification of b-quark jets with the CMS experiment. *JINST*, 8:P04013, 2013.

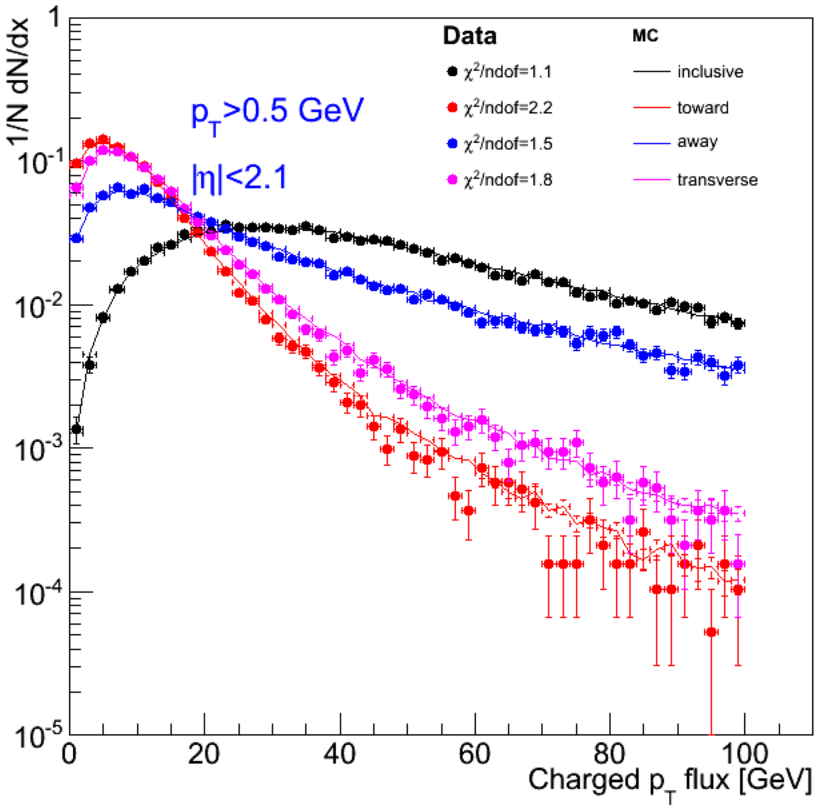

Figure 5: The distribuiton of the transverse momentum flux of the charged particles for all three regions and inclusively. Using the Z2* tune.

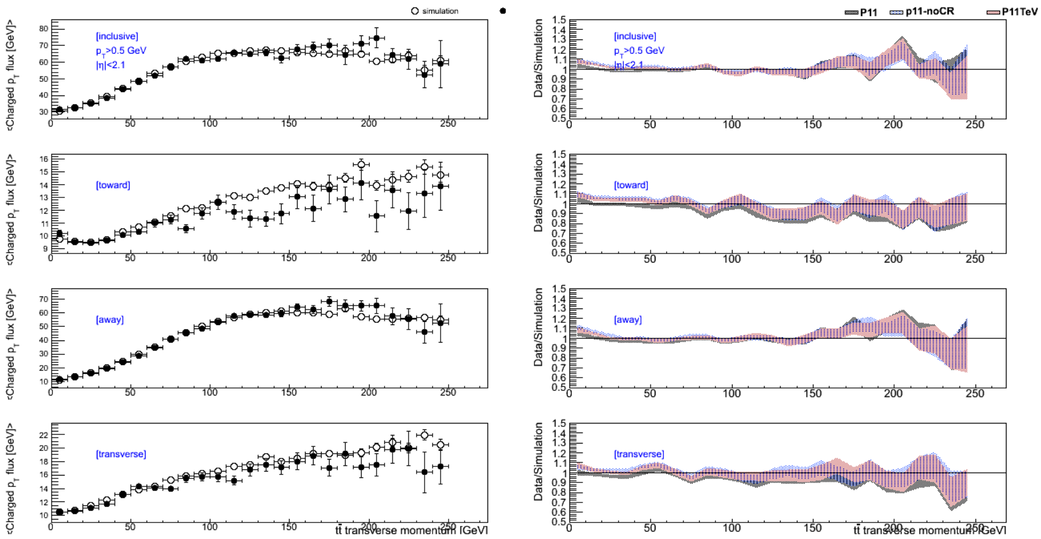

Figure 6: (Left) Scalar sum of transverse momentum profiled versus p_T of $t\bar{t}$. Using the Z2* tune.
(Right) Data to simulation ratio of the scalar sum of transverse momentum profiled versus p_T of $t\bar{t}$ for different P11 tune variations.

[5] Study of the underlying event and fragmentation and hadronization properties of b quarks in tt events. 2013.

[6] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 0605:026, 2006.

[7] Peter Zeiler Skands. Tuning Monte Carlo Generators: The Perugia Tunes. *Phys.Rev.*, D82:074018, 2010.

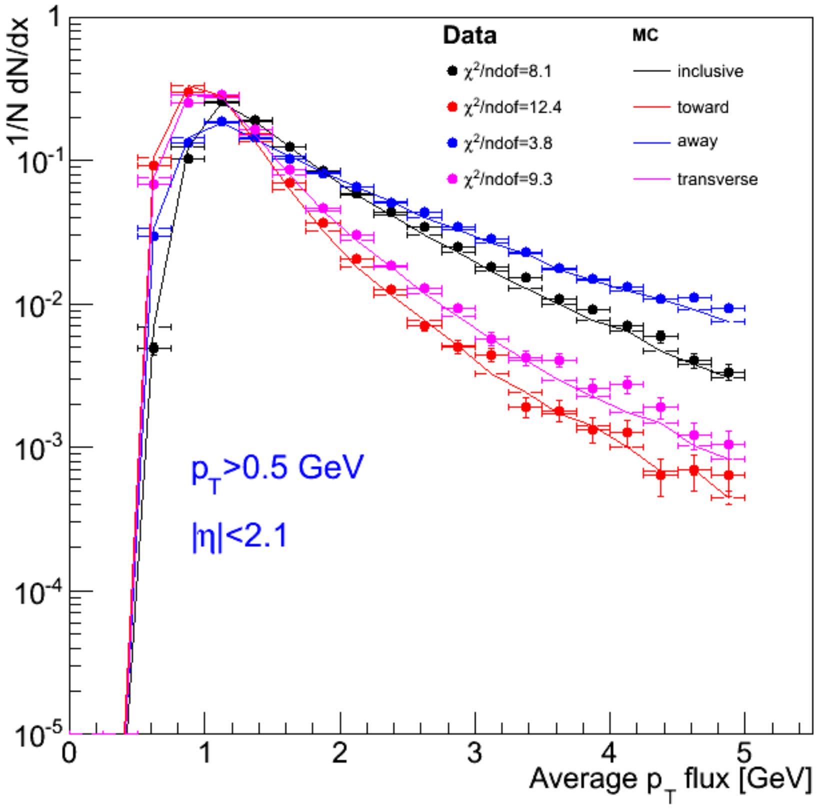

Figure 7: The distribution of the average transverse momentum of the charged particles for all three regions and inclusively. Using the Z2* tune.

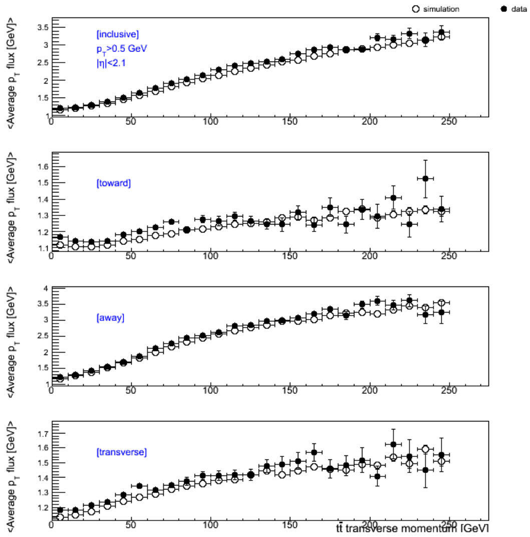

Figure 8: The average p_T per particle from data and Z2* tune profiled versus the p_T of $t\bar{t}$.

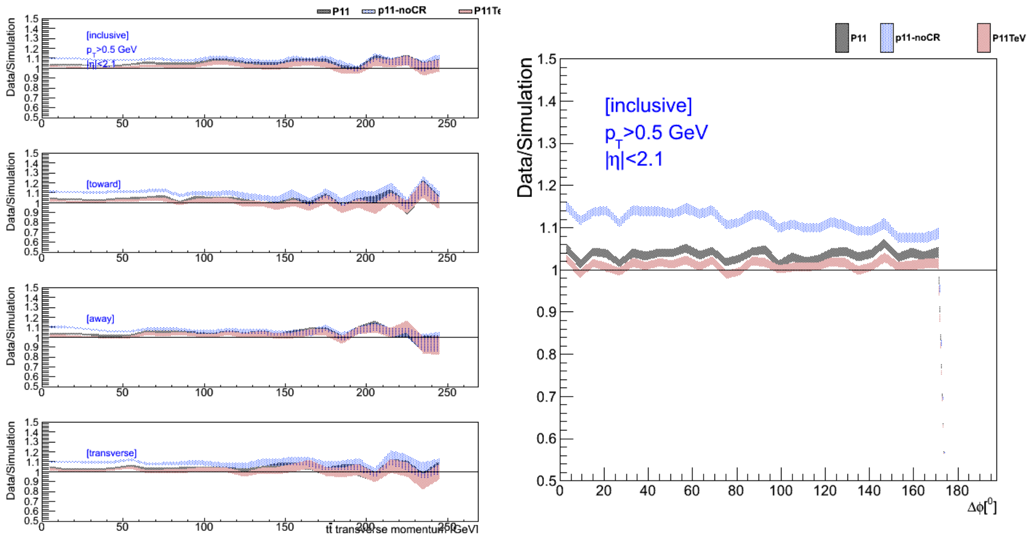

Figure 9: (Left) Data to simulation ratio of average p_T flux profiled versus p_T of $t\bar{t}$ for different P11 tune variations.
(Right) Data to simulation ratio of average p_T flux profiled versus $\Delta\phi$ for different P11 tune variations.