Sensitivity studies of colour reconnection in top UE measurements

Arturo Rodriguez Rodriguez

Instituto Superior de Ciencias y Tecnologias Aplicadas, Cuba

Supervisors: Paolo Gunnellini & Hannes Jung

September 10, 2015

Abstract

Sensitivities studies of color reconnection (CR) effects in $t\bar{t}$ underlying events were performed for the fully leptonic and fully hadronic final states (FLFS and FHFS respectively) events. Comparisons between predictions obtained with PYTHIA 8 and the data showed disagreements around 20%, mainly due to detector effects, but a good description of the data shape was achieved. Effects of CR were studied. Differences between predictions with and without CR were observed of $\sim 8-15\%$ for two of the investigated observables, as well as for predictions considering a variation of the fragmentation parameters. For different color reconnections models, effects around 5% were observedoptions studied were effects around 5%. No differences in the CR switched off between FLFS and FHFS were found for all the observables. This study shows the sensitivity of the UE observables to CR effects and may help to lower the uncertainties due to the UE simulation in top mass measurements.

Contents

1 Introduction

1.1 Motivation

In 1973, Makoto Kobayashi and Toshihide Maskawa predicted the existence of a third generation of quarks to explain observed CP violations in kaon decay. And two new quarks, the top and botton, were introduced. It did not take too long until a fifth quark, the bottom, was discovered by the E288 experiment team, at Fermilab in 1977. This strongly suggested that there must also be a sixth quark, the top, to complete the pair. It was known that this quark would be heavier than the bottom, requiring more energy to be created in particle collisions. It took another 18 years before the existence of the top was confirmed.

The top quark was discovered by the Tevatron experiments *CDF* and DØ in 1995, about 20 years after it had been predicted. The top quark has a lifetime of about 10^{-25} s, which is shorter than the average time of hadronization. Due to this reason it always decays into other fundamental particles, a W boson and a b quark (BR$\sim$ 99%)before forming bound states; thanks to its properties it is considered an important fundamental particle for precision measurements of the Standard Model (SM) as well as for searches of new physics beyond the SM. In addition it is the heaviest elementary particle, with a mass $m_t = 172.38 \pm 0.10\,(stat.) \pm 0.65\,(syst.)$ GeV [1] which is of the same order of magnitude of the mass of the gold atom. Due to its huge mass, it is a excellent candidate for studies of the fermionic coupling to the Higgs boson.

Due to the importance of accurate measurements of top mass at LHC, these ones are becoming more and more precise, with the most recent ones achieving a precision of less than 1 GeV [1].

At the LHC, the experimental measurement of the top mass, as well as other properties, relies on the interpretation of the observed final state in terms of the parton-level kinematics. However this strategy has some issues, related to one of the decay products, the b quark, which carries the color flow after the decay of the top quark.The fragmentation of the b quark is expected to occur in a B hadron, plus other hadronic particles which form the so-called b-jet. This feature may compromise the reconstruction of the initial kinematics for to different reasons. First, the fact same particles originating from th b-quark evalotuion may tend to escape the clustered jet, affecting the reconstructed final state; secondly the possibility of interaction and interference between the top decay products during the hadronization has also an impact in the reconstruction of the initial kinematics. This occurrence is known as color reconnection (CR), and it is responsible for a decrease in the precision that can be achieved in the top mass measurements and constitutes 20 to 40% of the uncertainty. Because of that, modeling of top events constitutes a well grounded motivation to look for a better understanding of the "underlying events" (UE).

In this report will be shown results of the sensitivity studies of color reconnection in top UE measurements, as a part of the DESY *summer student program* 2015, based on the

CMS PAS TOP-13-007 [2] analysis and the main goal was to investigate the role of the CR in $t\bar{t}$ events.

2 Theory

2.1 ttbar production and decay products

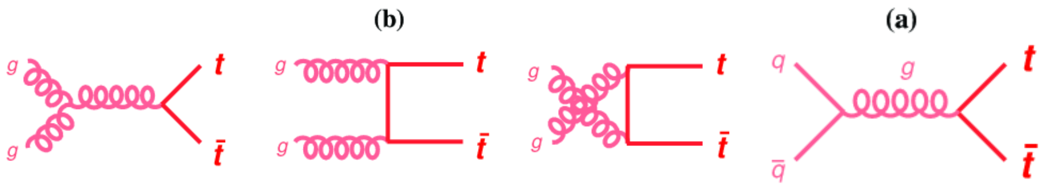
$t\bar{t}$ quarks pair are mainly produced through a QCD process. This pair can be produced from either a pair of gluons or of quarks in the initial state. Figure 1a, Figure 1b,. At the LHC $t\bar{t}$ quarks pairs are produced after the scattering of two protons being the first process in Figure1a dominant. Top-quark is produced at a scale of the order of its mass, this means that its production is a hard process that can be pertubatively described. Top quarks are colored particles, the decay is produced before the hadronization and fragmentation owing to its large mass and width. Due to this, all its properties are transferred to its decays products, which are expected to be $> 99\%$ of the cases a W boson and a b quark. Depending on the decay of the W boson, there are three possible final states.

- *Fully hadronic channel:* Both W bosons decay into a pair of quark and anti-quark of different flavor. As a result, 6 quarks are produced in total. This channel has the largest branching fraction, about 44%, but also has a very high background coming from QCD processes (UE).
- *Semileptonic channel:* One W boson decays into a pair of quark and anti-quark and the other one to a charged lepton and a neutrino of the same flavor. As a result, 4 quarks, 1 charged lepton and 1 neutrino are produced. The branching ratio of this channel is about 30% and it has a moderate background mainly coming from W production with additional jets (referred to as W+jets). Given the very short lifetime of the tau lepton, only electrons and muons are considered as possible lepton candidates.
- *Dilepton channel:* Both W bosons decay into 1 charged lepton and 1 neutrino. As a result, 2 quarks, 2 charged leptons and 2 neutrinos are produced. This channel has the lowest branching ratio, about 5% but also has a very low background coming mainly from Z/γ^* boson production with additional jets (referred to as Z+jets) and QCD processes.

The leading order Feynman decay diagram is shown in Figure 2a as well as in Figure 2b the top pair branching ratios.

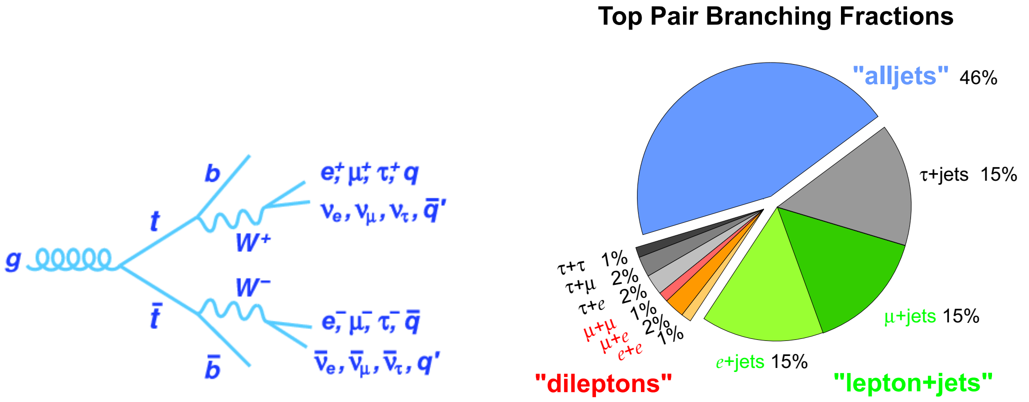
2.2 Underlying events (UE)

In this report we defined as UE any hadronic activity which can not be attributed to the particles arising from the hard scattering, i.e. the decay products of the $t\bar{t}$ system, as it



(a) (b)

Figure 1: Leading order Feynman diagranm for strong production of top quark pairs from (a) quarks and (b) gluons. *Source*: DØ Collaboration



(a) (b)

Figure 2: (a) Leading order Feynman diagranm for $t\bar{t}$ decay (a) quarks and (b) Top pair branching ratios. *Source*: DØ Collaboration

was defined in [2]. The hadronization of initial- and final-state radiation (ISR/FSR) is also considered part of the UE as long as the particles are not clustered within the two reconstructed b-jet candidates. In the interaction process other partonic constituents within the initial colliding hadrons can scatter. This can lead multiparton interactions (MPI) which produce particles that contribute to the UE. In addition, particles from the hadronization of beam-beam remnants (BBR) are considered part of the UE.

2.3 Color reconnection (CR)

Color reconnection is an ad hoc mechanism mainly used to describe the interactions that can occur between colored fields during the hadronization process [3]. At the LHC, due to the high number of colored partons, from a combination of MPI and parton showers as well as from BBR, this process is expected to occur at a significant rate. CR is an important ingredient of the UE contributions and since there is no first-principles models that could give a unique answer, the best that one can do is to study a range of realistic models and evaluate the spread of effects.

2.4 Simulation with PYTHIA 8

The simulation process was performed with the PYTHIA 8 [4] event generator , CUETP8M1 tune, at the center of mass energy $\sqrt{s}$ =8 TeV. To achieve a good statistic a generation of 1 million events was set for each simulated sample. PYTHIA 8.1 currently contains only one model (sometimes referred to as the MPI-based one); with two possibilities for resonance system, denoted as "default" and "default Early Resonance Decays (ERD)". Our simulation was performed using only the default possibility, which evaluated the probability of CR between the hard scattering and the UE according to the following equation:

$$P_{rec}\left(p_T\right) = \frac{\left(R_{rec} p_{T0}\right)^2}{\left(R_{rec} p_{T0}\right)^2 + p_T^2}, \tag{1}$$

where $0 \leq R_{rec} \leq 10$ (for $R = 10$, saturation effects take place) is a phenomenological parameter and p_{T0}

Starting from the lowest-p_T interaction in a set of parton interactions, a reconnection probability for an interaction with hardness scale p_T is an energy dependent parameter used to damp the low-p_T divergence of the $2 \rightarrow 2$ QCD cross section. The reconnection probability is chosen to be higher for soft systems, reflecting the fact that soft-p_T system are described by more extended wave functions, thus having a higher probability to overlap and interact with other systems. If an interaction does not reconnect with the next highest one in p_T, then consecutively higher ones are tried, so that the total reconnection probability for an interaction is $1 - \left(1 - P_{rec}\right)^n$ if there are n interactions at higher p_T scales.

2.5 Event selection

Event selection was performed with Rivet 2.2.0 [5], based on the CMS-PAS-TOP-13-007 [2] analysis. For the top selection two b-jets were required, selected with $p_T > 30GeV$ in a pseudorapidity range of $|\eta| \leq 2.5$. Two leptons were also required with $p_T > 20GeV$ in $|\eta| \leq 2.5$. The charged particle selection was made by applying a cut for $|\eta| \leq 2.1$ and $p_T > 0.5GeV$. In our study three basic quantities were considered to describe the UE in $t\bar{t}$: the charged particle multiplicity resulting from the simple count of selected charged candidates (N_{ch}); the charged flux in the transverse plane resulting from the scalar sum of the transverse momentum of the selected candidates ($\sum p_T$); the average flux per charged particle, computed from the ratio of the two previous quantities $\overline{p}_T = \sum p_T / N_{ch}$. These three quantities were studied with respect to an axis defined event by event, after computing the reconstructed momentum of the $t\bar{t}$ system, as:

$$\vec{p}_T\left(t\bar{t}\right) = \vec{p}_T\left(b_1\right) + \vec{p}_T\left(b_2\right) + \vec{p}_T\left(l\right) + \vec{p}_T\left(l'\right) + \vec{p}_T^{miss} \tag{2}$$

where $\vec{p}_T\,(b)$ and $\vec{p}_T\,(l)$ are the transverse momentum of the b-tagged jets and the charged leptons and $\vec{p}_T^{miss}$ is the imbalance in the transverse momentum of the event computed from the negative of the sum of the momenta of all reconstructed particles. It is expected that this axis be correlated with the direction of the $t\bar{t}$ system. The regions constructed in that way are shown in Figure 3 according to the distance in azimuthal angle with respect to the direction of the $t\bar{t}$ system.

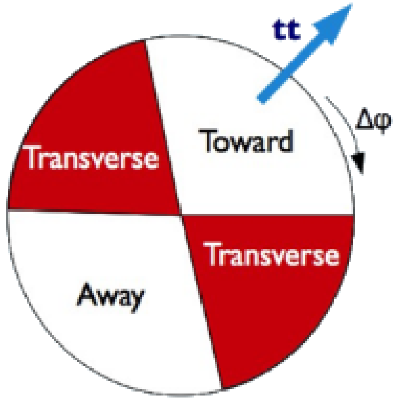


Figure 3: Constructed regions with respect to the direction of the $t\bar{t}$ system

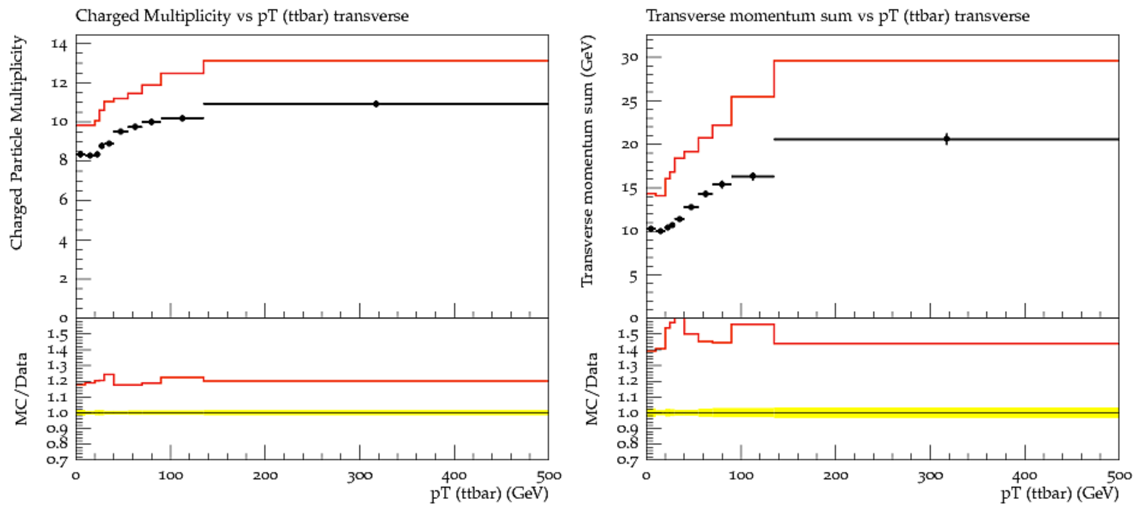
3 Study results

3.1 Predictions and data

A first comparison with the default predictions of PYTHIA 8 CUETP8M1 tune and the data was performed, as observed in the Figure a),b). Differences around 20% were found, which are related to detector effects, since the data are not yet unfolded to stable-particle level. In spite of that, the default predictions describe properly the general trend of the measurements points of the data. This is the reason why it was decided to use P8 for our purpose.

3.2 Color reconnection studies in fully leptonic final state (FLFS)

A first comparison between the default predictions and the predictions with CR switched off was performed for the three observables in the three regions of the transverse plane as is shown in Figure 5. Significant differences, around 10%, can be observed, between it default and CR off predictions for all the regions. It can also be observed that the most sensitive observables are N_{ch} and $\overline{p}_T$. For the three regions it is possible to observe an increase in the charged particle multiplicity for the CR off predictions, which is due



(a) (b)

Figure 4: Comparison between default predictions PYTHIA 8 with data a) charged particle multiplicity and b) sum of transverse momentum in the transverse region

to length of the color strings which are larger for CR off predictions. As a consequence these strings are more energetic, and produce more particles. The sum of the transverse momentum does not show noticeable changes. In this case this behavior is related to the interplay of two processes, usually the length of the strings and the different momentum transferred from the hard scattering to the MPI products due to CR. This may explain the absence of significant difference between these two predictions.

The difference observed among both predictions indicate the possibility of using these two observables to constrain the CR in top UE, which would bring to a better constrain of the UE contribution and to an improvement in the top mass uncertainties.

3.2.1 Color reconnection range

Predictions with different CR range values (R_{rec}) in the CR PYTHIA model, (see 1), were also studied, the result of these can be observed in Figure 6. As can be appreciated, there is no significant difference between the default predictions (range=1.8) and the other R_{rec} choices. This can be related to the fact that for values above unity the reconnection rates tend to be saturated, since then most systems are already connected with each other [4].

3.3 Fragmentation studies in FLFS

PYTHIA 8 uses the Lund fragmentation model to simulate the fragmentation process. In our studies we change the a and b parameters of the Lund symmetric fragmentation

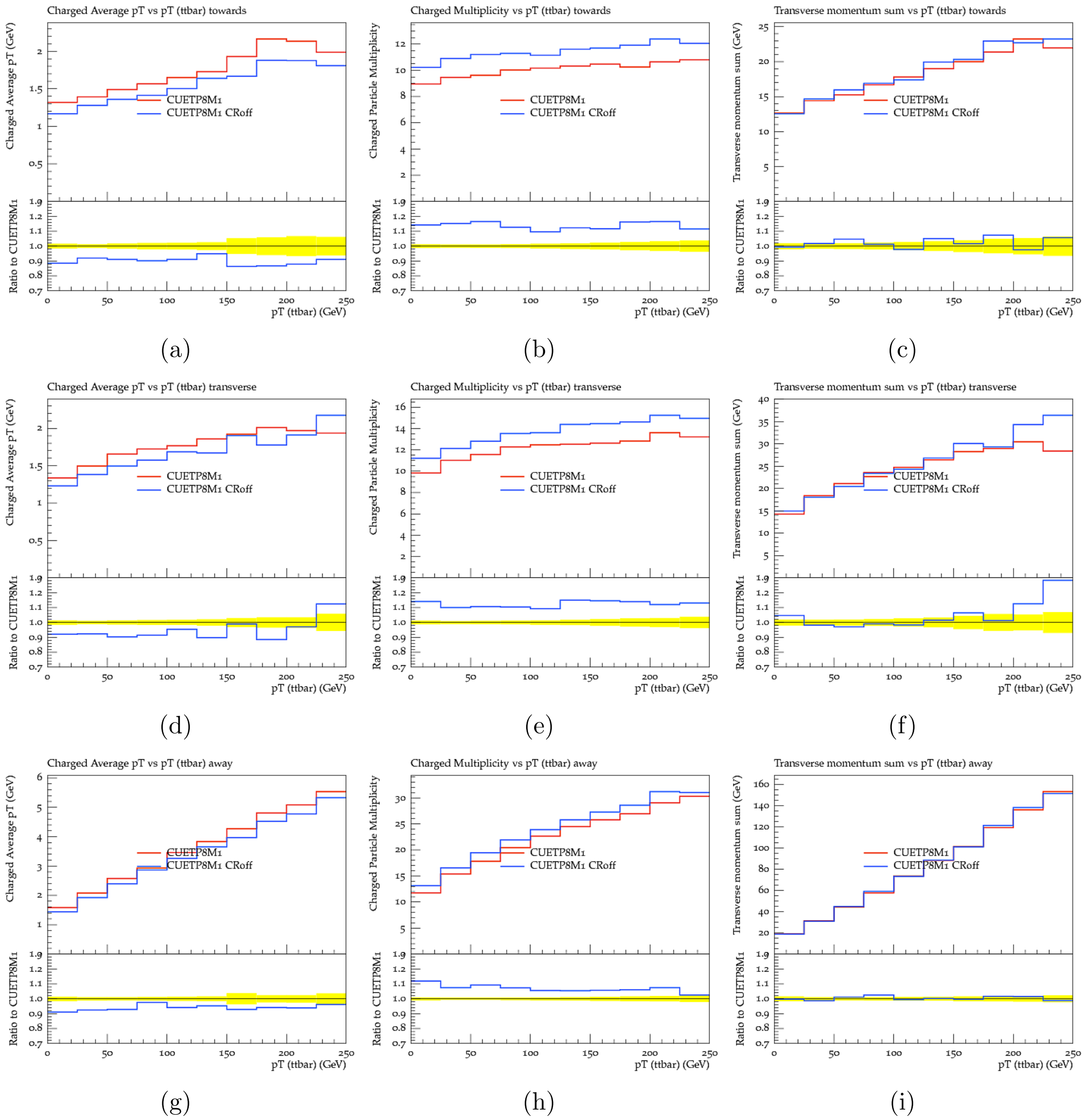
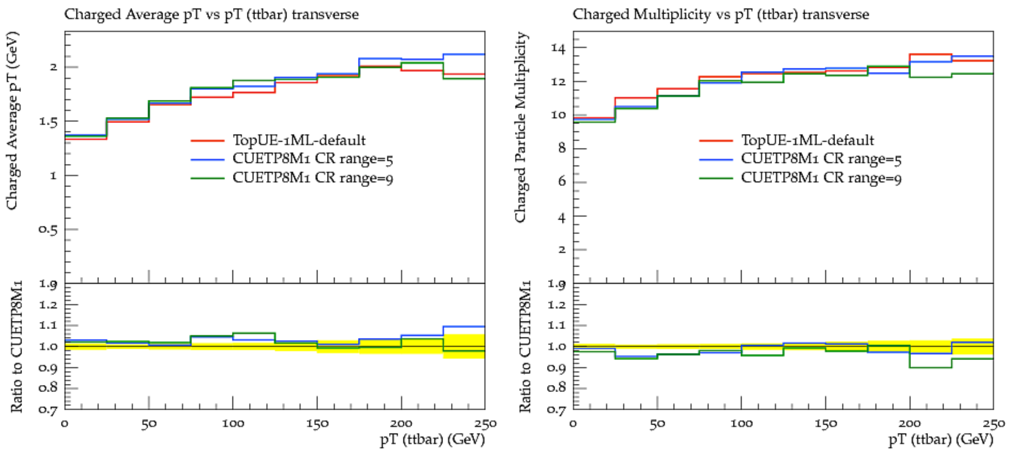


Figure 5: CR switched off effects for all regions; from a) to c) $\overline{p}_T$, N_{ch}, $\sum p_T$ in the toward region; from d) to f) $\overline{p}_T$, N_{ch}, $\sum p_T$ in the transverse region and from g) to i) $\overline{p}_T$, N_{ch}, $\sum p_T$ in the away region

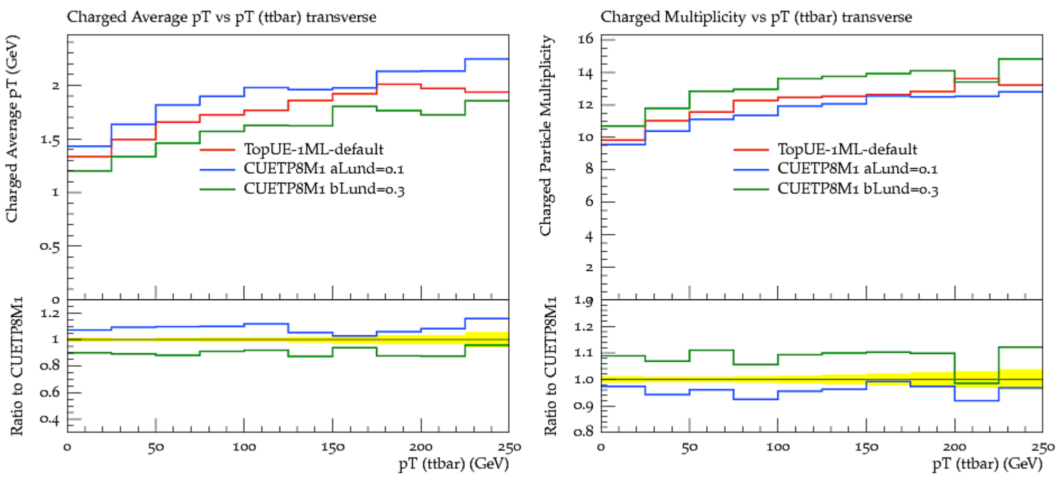
function, which give the transfer momentum fraction of the fragmentation products. In the plots of Figure 7, one can observe differences around 10% among the default simulation and the predictions with variations of the Lund fragmentation function. In Figure 7b) a decrease is observed in the charged particle multiplicity associated to the dependence $\sim (1-z)^a$ (an increase in the value of a respect to the default value $= 0.68$) of the symmetric fragmentation function which means that the momentum transfer is larger and the fragmentation products carry higher p_T. As a consequence there is less



(a) (b)

Figure 6: CR range variation, a) charged average p_T and b) charged particle multiplicity in the transverse region

energy for particle production.



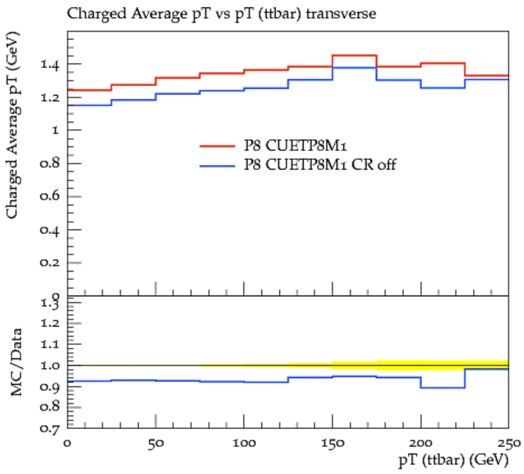
(a) (b)

Figure 7: Fragmentation variation a) charged average p_T and b) charged particle multiplicity in the transverse region

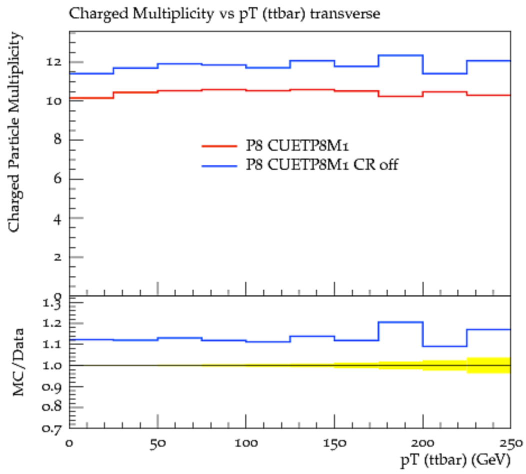
In the case of the change of the b parameter an opposite behavior is observed. A decrease in the value of b (default parameter $= 0.98$) determines an increase in the charged particle multiplicity which can be related to the same process described above, but in the opposite direction.

For the same reasons explained in the case of the charged particle multiplicity, the behavior of charged average p_T can be understood. For the a parameter as we have higher charged particle multiplicity with particle of low p_T a decrease is observed for this observable and the opposite case occurs for the predictions where the b parameter was changed.

This behavior confirms the possibility of using these two observables for constraining top UE.



(a)



(b)

Figure 8: CR switched off in the FHFS for a) charged average p_T and b) charged particle multiplicity in the transverse region

3.4 Color reconnection studies in hadronic final state

The routine **CMS-PAS-TOP-13-007** was modified to obtain a fully hadronic final state (FHFS) requiring the decay of the W boson in two more jets. For this final state the influence of the CR was also studied.

It is possible to note that the differences among the predictions is around 10% again, which are of the same order of the difference observed for the FLFS. Again this seems to indicate that the charged particle multiplicity and charged average p_T could be used for constraining color reconnection in top underlying events. This behavior can be understood if we assume that in PYTHIA 8 there is no implementation of the color reconnection for the decay products of the W boson, which in principle, needs to be done to describe top UE data.

4 Conclusions

A study of the top UE was performed. As a first step a comparison among the default prediction of PYTHIA and the measured data was done. Wide difference was observed (around $\sim 20\%$) between predictions and data, while a good description was achieved for the shape of the measurement. The difference seems to be related to detector effects since the data are not unfolded while the simulation was performed at particle level.

In a second step, the effects of the MPI and the hadronization level were studied. Large difference between the default predictions and the MPI off and Hadron level off was observed, showing the necessity of using these process to describe the data. Color reconnection effects in the fully leptonic and hadronic final state in $t\bar{t}$ hace an impact of around 10% for charged particle multiplicity and charged average p_T while for the sum of transverse momentum no noticeable changes were observed.

Effects of the color reconnection range and fragmentation for the fully hadronic final state were also studied. For the CR range variation no significant change was observed, which allows to say, in principle, that for the considered values. Were observed for the variation of the fragmentation parameters, differences around 10%, in the charged particle multiplicity and in the charged average p_T which confirms the idea of using these two variables for a possible tuning of the top UE measurements.

Due to all these facts this study shows that the charged particle multiplicity and the charged average p_T might be two appropriated observables for constraining color reconnection in top underlying event, and as a result would allow to lower top mass uncertainties due to color reconnection.

5 Acknowledgments

First of all, I would like to thank Hannes Jung for all his support during this period, since the beginning of the application process until now, thanks lot for giving the opportunity to come here and spend a wonderful summer. To my supervisor Paolo Gunnellini for all the time that he had to spend with me, and for all his excellent explanations and laugh moments. I want to thank Jasone Garay for all her help and concern, for being all this time my best friend, also to my officemates Ganna Dolinska and Nazar Bartosik and of course to all my new friends. A special thanks to my beloved Tante und Onkel, Edita and Werner for all your concern and support and love during this summer and to my family in Cuba for their love.

References

[1] J. Beringer *et al.* [Particle Data Group Collaboration], Phys. Rev. D **86** (2012) 010001.

[2] CMS Collaboration [CMS Collaboration], "Study of the underlying event, b-quark fragmentation and hadronization properties in ttbar events," CMS-PAS-TOP-13-007.

[3] S. Argyropoulos and T. Sjostrand, "Effects of color reconnection on $t\bar{t}$ final states at the LHC," JHEP **1411** (2014) 043.

[4] T. Sjostrand, S. Mrenna and P. Z. Skands, Comput. Phys. Commun. **178** (2008) 852 [arXiv:0710.3820 [hep-ph]].

[5] A. Buckley, J. Butterworth, L. Lonnblad, D. Grellscheid, H. Hoeth, J. Monk, H. Schulz and F. Siegert, Comput. Phys. Commun. **184** (2013) 2803 [arXiv:1003.0694 [hep-ph]].