# Simulation of measurement of jet properties for jets selected in semi-leptonic $t\bar{t}$-events at 8 TeV pp-collisions

**JINR Summer Student Program 2017**

**Dmitry Budkouski, BSU, Minsk, Belarus**

**Supervisor: Dr. S. Shulga**
**Leading research scientist, LHEP, JINR**

## Abstract

The work was performed within the scope of the problem of measurement of charged-particle multiplicity (CPM) in jets, which is performed in a JINR CMS physics group. We report about the simulation of measurements of mean chargedparticle multiplicity in jets for two groups of jets collected in events with topology of semi-leptonic $t\bar{t}$ events at 8 TeV pp-collisions: (1) pair of jets with invariant mass in 20 GeV interval around W-meson mass and (2) additional jets ($5^{th}$-, $6^{th}$- etc.). The properties under study are: mean jet CPM's, flavour jet fractions and mean flavour jet CPM's in two jet samples. The simulation is performed in framework of Pythia8 model. Pythia8 results are compared with known Pythia6 results that were obtained earlier.

# 1. Introduction

The form of the distribution of jets on the number of particles in jet reflects the dynamics of the jet formation process. According to current views, based on QCD, process of multiple production, initiated by energetic parton is a cascade of branching processes (emission of gluons and $q\bar{q} -$ pairs), with a gradual decrease of parton energy. This cascade process terminates for values of parton energies in a few hundred MeV, when the nonperturbative hadronization processes start to dominate.

Multiplicity distribution and mean multiplicity in jet have different energy dependences for quark and gluon jets but the dependence on flavour decreases with increasing the jet energy. In ratio $r_{theor} = < n^{(g)} > / < n^{(q)} >$ the main energy dependences are cancelled and it remains a weak dependence associated with the running coupling constant $\alpha_s$.

The value $r_{theor}$ is sensitive to higher QCD corrections. In the lowest order $r_{theor}$= 9/4, which is significantly higher than the experimental results. Corrections up to the third order in $\alpha_s$ reduce the value to $r_{theor} = 1{,}94$ , which is about 20%

higher than the experimental results in the jet energy range 10-100 GeV The value $r_{theor}$ in theory is calculated for partons emitted in inclusive gg (for gluon jets) and $q\bar{q}$ (for quark jets) systems produced from a color-singlet source. To compare $r_{theor}$ with the experimental value it is necessary to take into account the hadronization stage. In case of measurements for particular final-state jet topology and with using specific algorithm of jet identification the experimental conditions have not matched those of the theoretical calculation. To consider all of these effects (hadronization, jet topology, jet finder algorithm) in conjunction we must to use particular Monte Carlo (MC) model. Using the MC model, particular channel and jet selection conditions we can compare the predicted value $r_{theor}$ with experimental one. To provide the quantitative test of the QCD results the perturbative part of MC program should contain terms beyond the leading order.

In the asymptotic region of jet energy the hypothesis about the local partonhadron duality (LPHD) is assumed . According to LPHD hypothesis (originated from the idea of soft preconfinement) hadronization begins in soft energy region and controlled by soft processes, which leads to the fact that the jet MD at the parton level and at the level of particles are similar in shape $z = r_{hadron}/r_{theor} \approx 1$. At the intermediate jet energies the ratio *z* may differ from unity, and the measured multiplicity ratio of quark and gluon jets differs from the $r_{theor}$ by a factor *z*. Measurements of the ratio of the mean multiplicities in quark and gluon jets at intermediate energies shed light on the quantity *z*, which accumulates information about the hadronization stage of jet in given environment. The *z* value can be obtained by taking a certain hadronization model and performing MC calculation. For particular channel with several jets the *z* values for quark and gluon jets may be different. It brings additional uncertainty for the interpretation of measurement results.

In hadron-hadron collisions jets are identified in event by special algorithm (jet finder). "Good jet finder" satisfies the requirements of "collinear and infrared safety". The so-called "$anti - k_T$" algorithm satisfies these requirements and is used in the LHC experiments as the main jet finder algorithm.

Due to low efficiency to reconstruct neutral particles in jets we will measure only charged-particle multiplicity (CPM) in jets. To measure mean CPM's in quark and gluon jets we collect two disjoint groups of jets in events with jet topology corresponding to the semi-leptonic channel with $q\bar{q}$ quarks in the intermediate state ( Fig.1).

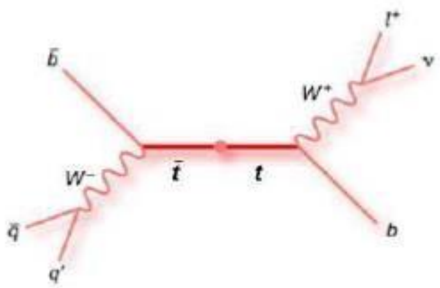

Figure 1. Semi-leptonic $tt$-channel

Using two different MC mixtures of quark and gluon jets increases the number of model-dependent parameters (parameters of fractions of q/g-jets and heavy flavor jets in two jet samples).

Interpretation of the measurement becomes more model-dependent. This increases the theoretical uncertainty, but the uncertainties are under control and are expressed quantitatively through the uncertainties of these parameters. On the other hand, the use the same events to select the two groups of jets creates more definite MC model for the theoretical interpretation of the measurements.

The differences between the properties of the jets in two groups of jets are provided by the differences in the rules of physically motivated jet selections: in one group of jets we collect jets mainly from decays of W-mesons (W-jets ), and the second group (denoted in this work as $5^{th}$-jets) contains jets from additional activity (initial-state radiation, final-state radiation, underline event activity) in the event which has exactly two b-jets and two W-jets. W-jets and $5^{th}$-jets we denote as signal jets.

Final-state of the semi-leptonic event with intermediate $t\bar{t}$ pair has one isolated lepton with high transverse momentum, significant missing transverse energy, and four or more jets. Since there is well-defined lepton primary vertex (PV), we can reliably separate tracks from pile-up vertices, tracks from secondary vertices (SV) and signal tracks produced in muon PV. We do not recluster jets after PU-tracks subtraction. It is possible because the fraction of PU jets is small in region of studied $P_T^{jet}$ > 30 GeV (at the level of uncertainty, named below as PU jet uncertainty, which is at the level of other systematical uncertainties). In our measurements of jet CPM we treat the admixtures of PU jets and heavy flavor jets as irreducible admixtures. We mentioned heavy flavor jets as irreducible admixture, because we have no means to separate the c-jets and light quark jets. The developed discriminators allow us to separate quark jets (including u, d, s and c quarks) and gluon jets. All we can do is to determine the fractions of these jets in signal jet sample and use it as modeldependent parameters or source of systematical uncertainty (depending on the sizes of these fractions) to interpret the measurements.

# 2. Task and motivation

The work was performed within the scope of the problem of measurement of charged-particle multiplicity in jets, which is performed in a JINR CMS physics group. The task is following: we need to simulate the measurements of mean charged-particle multiplicity in jets for two groups of jets collected in events with topology of semi-leptonic $t\bar{t}$ events at 8 TeV pp-collisions. For simulation of the measurements Pythia6 generator is used before. The using of the other generators in the analysis is important because it allows to determine the systemic uncertainty associated with variations of process models. It is also important to have an independent analysis to control the simulation results.

# 3. Packages and tools

Jets sample used in the study was obtained with the help of the Pythia8 generator with default settings. Protons are initial particles; their total energy in the c.m.s. is $\sqrt{S}$ = 8 TeV, that corresponds to the CMS detecting conditions at Run-I.

Jets are found using JetFinder Pytha8::SlowJet (R=0.5) (anti-kT with the angular distance $\Delta R = 0.5$ in ($\eta$, $\varphi$)-space ( $R = \sqrt{\Delta\eta^2 + \Delta\varphi^2}$ , $\Delta\eta$ and $\Delta\varphi$ – pseudorapidity and azimuth angle of the parton angle with respect to the jet axis).These algorithms is lite versions of the "fastjet3" algorithms. Jets are divided into several bins by their transverse momentum $P_T^{jet}$ ([30-60], [60-90], [90-120], [120-150], [150-180], [180-210], [210-300], [300-400].

Treatment the obtained results were carried out with Root - a modular scientific software framework. It provides all the functionalities needed to deal with big data processing, statistical analysis, visualization and storage.

# 4. Event and jets selection

The selection of the required event consist of several steps. At the first step we select the events with leptons having $P_t > 26\ GeV, |\eta| < 2.1$. After that we cut all events with additional leptons. Then we select the events with two b – jets with $P_t^{jet} > 15\ GeV, |\eta| < 2.4$. In the last step we select the events with two jets with invariant mass in 20 GeV interval around W-meson mass.

A flavor of jets is defined with simple algorithm: the most energetic parton in the $\Delta R = 0.5$ cone define flavor of jets. There are two different methods of flavor definition: PJF (Physics Jet Flavor definition) and AJF (Algorithmic Jet Flavor definition). In the first method partons are selected at the state immediately after the hard process. This method of identification gives a more stable identity and is a more priority in the CMS. The second method use partons are selected at the state just before hadronization after parton branching processes finishing.

For tagging the b-jets different discriminators are used in CMS. The best discriminator is Combined Secondary Vertex algorithm. We chose "medium" work point for this discriminator. At this point b–jets efficiency equals about 30%, misidentification probability for c–jets equals about 20% and for udsg – jets equals about 1%.

# 5. Results of simulation

Obtained with the help of the Pythia8 generator results was compared with previously obtained Pythia6 generator results. From the table one can see number of events in Pythia8 and Pythia6 at each step of cuts.

Table 1. Selection of events

| N | Cuts | $t\bar{t}$ evts (Pythia6+full sim.) | $t\bar{t}$ evts (Pythya8) |
|---|---|---|---|
| 1 | Events with isolating leptons ($p_T{}^{\mu} > 26$ GeV, $\lvert\eta\rvert < 2.1$ ) | 1469713(100% ) | 33432 ( 100%) |
| 2 | Events with additional isolating leptons | 1136229 ( 77% ) | 30203( 90±2% ) |
| 3 | Events with two b – jets with && $P_T{}^{jet} > 15$ GeV | 158639 ( 11% ) | 7361 ( 22 ±1,5% ) |
| 4 | Events with W – jets with , $P_T{}^{jet} > 15$ GeV | 22224 ( 1,5 % ) | 676 ( 2,0±0,5%) |

After the last step of the selection percentage of events in Pythia8 and Pythia6 are coincide.

Results of modelling of measurements of jet CPM's are presented in Fig.2, Fig.3 and Fig.4.

Fig.2 demonstrate the mean jet CPM's for W – jets and $5^{th}$ – jets (left pictures) and the difference between mean CPM's of W – jets and $5^{th}$ – jets (right pictures).

Difference between mean CPM's in Pythia8 less than in Pythia6, but the absolute values of mean CPM's are equal within uncertainties.

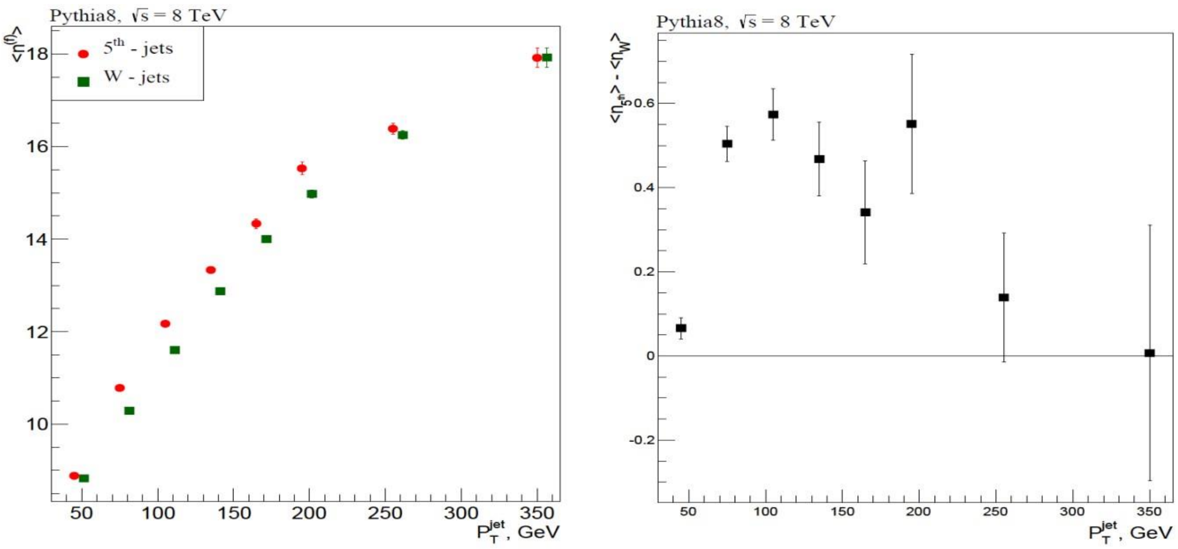


Fig.2 Mean charged-particle multiplicity and difference between mean multiplicities of two groups in Pythia8.

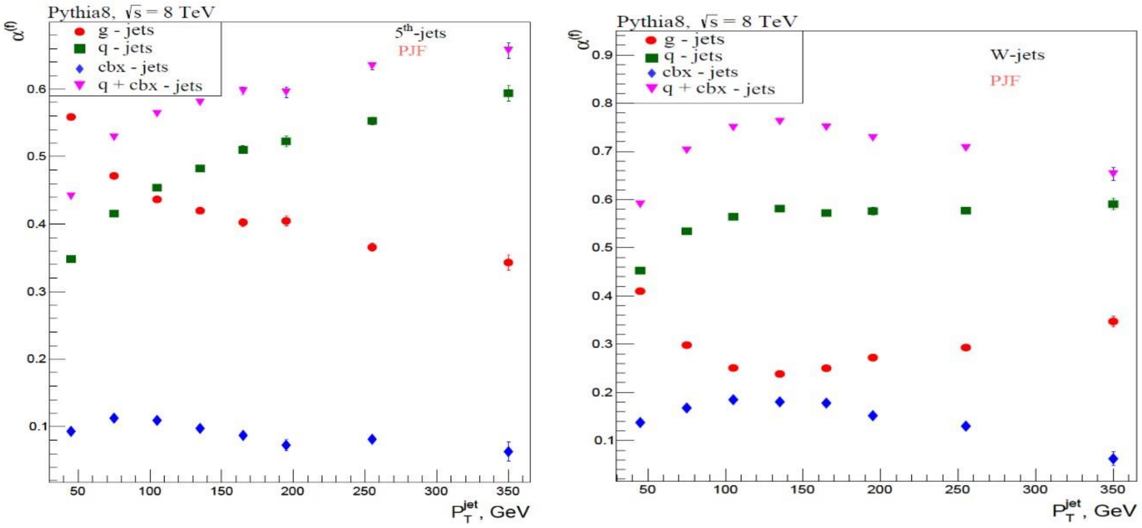


Fig. 3 The fraction q/g, cbx and q+cbx in Pythia8 for 5th-jets (left) and W-jets (right).

At the figure 3 the fractions of q/g, cbx and q+cbx obtained in Pythia8 generator are shown. Jet flavor is defined by PJF algorithm. From the Fig.3 we can see that fraction of q – jets ( and q+cbx – jets, that we assume as q - jets) in W–jets more than g – gets fraction, that match with the expectations, but in 5th – jets g - jets fraction suppressed, that contradicts to our expectations and the reasons for this are not known yet and will be a subject of our further study. Mean CPM's for g-jets, qjets, cbx-jets and q+cbx-jets in Pythia6 and Pythia8 are equal within errors.

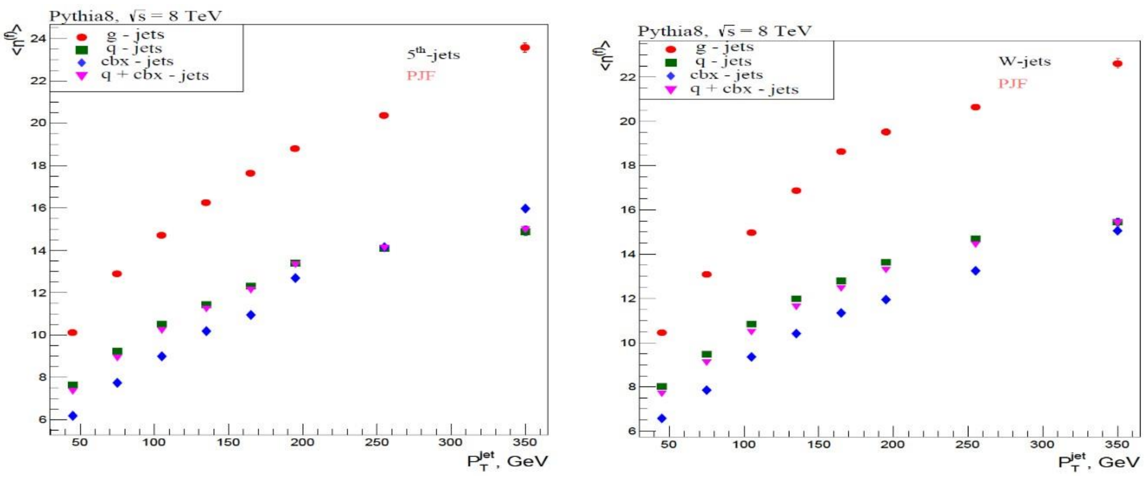


Fig.4 Mean CPM's for g-jets, q-jets, cbx-jets and q+cbx-jets in Pythia8 for 5th-jets (left) and W-jets (right).

At the figure 4 the mean CPM's for g-jets, q-jets, cbx-jets and q+cbx-jets are shown. Jet flavor is defined by PJF algorithm.

From the Fig.4 can see that mean CPM's for q – jets and q+cbx-jets are close one to other. This equality allows to use only mean CPM's of q+cbx-jets as independent parameter and to reduce the number of independent parameters in equations.

At the Table 2 we see the values of mean CPM's and difference between mean CPM's in q/g – jets for W – jets and $5^{th}$ – jets. The negative values in $4^{th}$ and $7^{th}$ columns show so called "non-universality" of g/q–jets. With the same production conditions the properties of g/q – jets should be equal in any group of jets. "Nonuniversality" of g/q – jets is followed by the different kinematics of the jets and possible dependence of underlying event content on the pseudorapidity region.

Table2. Mean g/q – jets CPM's for two jet samples.

| $P_T^{Jet}$ - bin, GeV | $\langle n^{(q+cbx)} \rangle_W$ | $\langle n^{(q+cbx)} \rangle_{5th}$ | $\Delta n^{(q+cbx)}$ | $\langle n^{(g)} \rangle_W$ | $\langle n^{(g)} \rangle_{5th}$ | $\Delta n^{(g)}$ |
|---|---|---|---|---|---|---|
| [30,60] | 7,684±0,015 | 7,331±0,016 | -0,354±0,021 | 10.449±0.018 | 10.108±0.019 | -0.3418±0.026 |
| [60,90] | 9,087±0,021 | 8,913±0,028 | -0,174±0,034 | 13.084±0.025 | 12.879±0.036 | -0.2046±0.045 |
| [90,120] | 10,207±0,027 | 10,207±0,044 | -0,262±0,052 | 14.971±0.035 | 14.704±0.056 | -0.2663±0.062 |
| [120,150] | 11,612±0,037 | 11,225±0,063 | -0,387±0,069 | 16.871±0.049 | 16.245±0.083 | -0.6256±0.094 |
| [150,180] | 12,225±0,054 | 12,112±0,087 | -0,333±0,092 | 18.636±0.068 | 17.632±0.111 | -1.0038±0.128 |
| [180,210] | 13,278±0,077 | 13,314±0,121 | 0,036±0,136 | 19.521±0.095 | 18.787±0.149 | -0.7347±0.163 |
| [210,300] | 14,425±0,082 | 14,104±0,101 | -0,321±0,137 | 20.644±0.106 | 20.344±0.133 | -0.2991±0.152 |
| [300,400] | 15,421±0,186 | 15,021±0,173 | -0,401±0,244 | 22.612±0.235 | 23.552±0.221 | 0.9403±0.345 |

# 6. Conclusion

We have developed a package based on the generator Pythia8, that is designed to simulate the measurement of mean CPM's in the samples of jets, the determination of the fractions in jets samples and the definition of mean CPM's of jets. We study two samples of jets selected in semi-leptonic channel with a $t\bar{t}$ pair at 8 TeV pp collision. The results are compared with the results obtained earlier by generator Pythia6.

Our main results are following:

1. Parts of events selected in $t\bar{t}$ channel with two b-jets in Pythia8 and Pythia6 are different in 2 times.
2. Faction g-jets in the sample of jets in Pythia8 much more than in Pythia6;
3. The mean CPM's of jets with fixed flavors in Pythia8 and Pythia6 are approximately the same;
4. The difference of CPM's between W-jets and 5-jets in Pythia8 is 2 times less than in Pythia6. The mechanism that produce this difference is not associated with a difference of flavor fractions.
5. Confirmed previously obtained in Pythia6 observation of
5. "Non-universality" of q(g)-jets (dependence on jet sample) is studied in Pithia8. This "non-universality" is much less than the difference of CPM's in q - and g-jets.
6. It is shown that the contribution of the secondary vertex tracks in jet CPM's is about 30%. This means that you need to use quite strict rules to match the tracks to the vertices, to avoid the counting of secondary vertex tracks.

# References


1. Y. Azimov, Y. Dokshitzer, V. Khoze, and T. S., “Similarity of Parton and Hadron Spectra in QCD Jets”, Z. Phys. **C27** (1985) 65
2. I. Dremin and J. Gary, “Hadron Multiplicities”, Phys. Rep. **349** (2001) 301
3. I. Dremin, “Multiparticle production and quantum chromodynamics”, Physics-Uspekhi **45** (2002) 507
4. I. Dremin, “Quantum chromodynamics and multiplicity distributions”, Physics-Uspekhi **37** (1994) 715
5. D. Amati and G. Veneziano, “Preconfinement as a property of perturbative QCD”, Phys. Lett. 83 (1979) 87