Quark and gluon jet fractions in jet sample selected in semi-leptonic $t\bar{t}$-events with backgrounds at 8 and 14 TeV pp-collisions

JINR Summer Student Program 2015

Egor Shulga, BSUIR, Minsk, Belarus

Supervisor: Dr.Sergei Shmatov
Head of Sector of CMS physics department, LHEP, JINR

1. Introduction

According to modern science conceptions parton interactions on the LHC occur in several stages. On the first stage, according to the empirical function PDF (parton distribution function) a parton (quark or a gluon) is selected from each proton. Than by the laws of QCD (quantum chromodynamics, which is more probable process) and by the theory of electro-weak interaction (which is less probable process) the partons are collided. The laws aggregate the central part of the so called Standard Model.

Usually the "2->2" process occurs. The process is accompanied by radiation in initial and final states and by the process Underline Event (UE). In terms of experiment final particles of the hadron-hadron collision are divided in two groups: jets and particles outside the jets. Particles outside the jets come from UE; its physical origin is interaction of the remnants of the initial hadrons. Remnants are partons without two active participants of the "2->2" hard process.

Most of the events do not have jets. That means that a hard partons collision did not occur. If there are jets in the events it is for hard process occurred. Jet is a collimated particles set that can be distinguished from the uniformly distributed UE.

A jet is a macroscopic object; its characteristics depend on the parameters used at the jet particles matching. Consequently for one analysis physical object we have several objects in accordance to the used jets identification algorithm (JetFinder).

At the same time a jet is an object that stands for the elementary microscopic process of the partons branching and hadronization. Particles that are the part of the jet bear information of color correlations that occurs on each stage from parton birth to its hadronization. The information does not depend on the JetFinder algorithm.

Jets characteristics depend on the parton-ancestor: whether it is a jet of the light quark (u, d, s), a jet of the heavy quark (c, b) or a gluon jet.

The most simple multiplicity distribution moments, such as mean multiplicity or dispersion of distribution, can evidently depend on the type of the JetFinder. E.g. mean multiplicity grows with the radius of the cone that is used for jet matching in the simple cone algorithm. Mean multiplicities in jets that were measured at the CDF [1] (jet matching cone 0.47 rad) and at the e^+e^- experiment [2] (unbiased jets – jets that are defined as objects in different semi spheres at the 2-jet events) differ almost by 3 times. It is necessary to use JetFinder at the pp experiment, because additional jets are born from radiation in the initial and final states and also because there is uncorrelated UE-contribution inside and outside the jets.

It is known from the experiments [1] [2] that jets characteristics depend on its flavor: gluon jet, light or heavy quark jet. It was predicted theoretically and confirmed experimentally that mean multiplicity in the gluon jets is about 1.5 larger than in light quark jets. That is why it is necessary to separate jets by their flavor when studying them.

2. Motivation

In the e^+e^--experiments birth of the jets in the final state and its structure corresponds to the theoretical concepts of parton branching and parton system hadronization. Respectively exclusive events with 2 jets and 3 jets are selected and jets in the separate semi-spheres are studied. This jets definition is called "unbiased" [3]. In the hadron collisions (hh-experiment: pp at the LHC, $p\bar{p}$ at the Tevatron) exclusive 2-jet events are suppressed because hard process is accompanied with the jets from the ISF, FSR. Under the conditions "biased" jet definition is applied i.e. using JetFinder algorithms.

Besides there is an UE-tracks contribution to the jets in the hh-experiment. Its density could be estimated in the zones outside the jets and then its mean value could be subtracted from the jet particles multiplicity. The presence of uncorrelated UE-tracks in the jet is a source of the systematic uncertainty of measurement.

Also in the LHC Run-I events there is a contribution of the Pile Up tracks concerned to the several pp-vertices tracks recording (on the average 15 vertices). In the Run-I events it is necessary to determine the signal vertex that would be matched to by signal tracks. In the case in the mean multiplicity of signal tracks in the jet there could be a residual Pile Up dependence that is another source of measurements uncertainty.

Another properties measuring difficulty that is related to the multiplicity in jets in hh-experiments consists of impossibility to obtain a sufficiently pure sample of the quark and gluon jets. The reason for this is the presence of additional jets (ISR, FSR), that at the selections cause combinatory uncertainty. In the hh-experiments we can only get a mixture of jets. Mixture purity is a parameter that is beyond the modeling. The parameter contributes to the measurements the largest theoretical uncertainty. At the samples quark fraction calculations the issue of background and signal processes contribution calculations emerges. Channel list is usually well defined. These calculations are performed with the use of collisions generators (Pythia6, Pythuia8, Herwig, Herwig ++ etc.).

Thus in hadron-hadron collisions we have a number of uncertainties that e^+e^--experiments do not have. However hh-experiments are important for the measurements of structure and correlations in jets, because they provide knowledge about the universality of the jets, about their properties in a suppressed and unsuppressed additional radiation. Besides, the energy spectrum of the jets in the hh-collision is bounded above by the energy of the colliding beams, while in e^+e^--experiments the energy of the 2 jets is usually fixed (e.g. a jet in Z-boson mass [2]).

In the e^+e^--experiment OPAL [2] "unbiased" jets in the exclusive 2-jets and 3-jets events were used. In the quark and gluon jets samples following purity was achieved:

$$\alpha^{(q)} = 0.932 \pm 0.002,\ \alpha^{(g)} = 0.830 \pm 0.017.$$

Samples jets energy: $E^{jet} = 41.8\ GeV$. For ratio of mean multiplicities of quark and gluon jets following value was obtained: $r_n = 1.471 \pm 0.049$ [2]. Another OPAL data processing presented in [4] with the following results: $r_n = 1.422 \pm 0.051$.

The first and for now the only measurement of the value r_n in the hh-experiment was presented in 2005 in the work of CBA collaboration [1]. In the measurements two jet samples (q-g mixtures) were used: jet sample of the events dijet and $\gamma + jet$ events in $p\bar{p}$ collisions. Eqation to find quark and gluon jet multiplicities from q-g-mixtures:

$$\langle n^{(jj)} \rangle = \boldsymbol{\alpha_q}^{(jj)} \langle n^{(q)} \rangle + (1 - \boldsymbol{\alpha_q}^{(jj)}) \langle n^{(g)} \rangle$$
$$\langle n^{(\gamma j)} \rangle = \boldsymbol{\alpha_q}^{(\gamma j)} \langle n^{(q)} \rangle + (1 - \boldsymbol{\alpha_q}^{(\gamma j)}) \langle n^{(g)} \rangle$$

Jets energy (in the center of mass system) in the samples assumes two values: $E_{jet} = 41$ and 53 GeV in dijet/$(\gamma + jet)$ center-of-mass frame. In the work [1] simple cone algorithm to collect jets was used. The results for jet cone 0.47:

$$r_N = 1.64 \pm 0.09 \pm 0.14\ (E_{jet} = 41\ GeV)$$
$$r_N = 1.66 \pm 0.13 \pm 0.18\ (E_{jet} = 53\ GeV)$$

If we have jets mixture than probability to have an exclusive jet with the number n of charged tracks is following: $P_n = \alpha^{(q)} P_n^{(q)} + [1 - \alpha^{(q)}] P_n^{(g)}$. To find mean charged-particle multiplicities $\langle n^{(q)} \rangle$ and $\langle n^{(g)} \rangle$ for q- and g-jets separately we need two q/g-mixtures:

$$\langle n_1 \rangle = \alpha_1^{(q)} \langle n^{(q)} \rangle + (1 - \alpha_1^{(q)}) \langle n^{(g)} \rangle$$
$$\langle n_2 \rangle = \alpha_2^{(q)} \langle n^{(q)} \rangle + (1 - \alpha_2^{(q)}) \langle n^{(g)} \rangle$$

Two mixture fractions $\alpha_{1,2}^{(q)}$ are defined by MC simulation.

This work is devoted to the value $\alpha^{(q)}$ modelling anв calculation based on the signal pp events with $t\bar{t}$ pair in an intermediate state. In the initial sample we select semi-leptonic events with two b-quarks, single muon and two light quark jets (fig. 1).

3. Events and jets modelling and selection

We choose process with a birth of $t\bar{t}$ pair. From the events we extract events with semi-leptonic process presented on the figure 1.

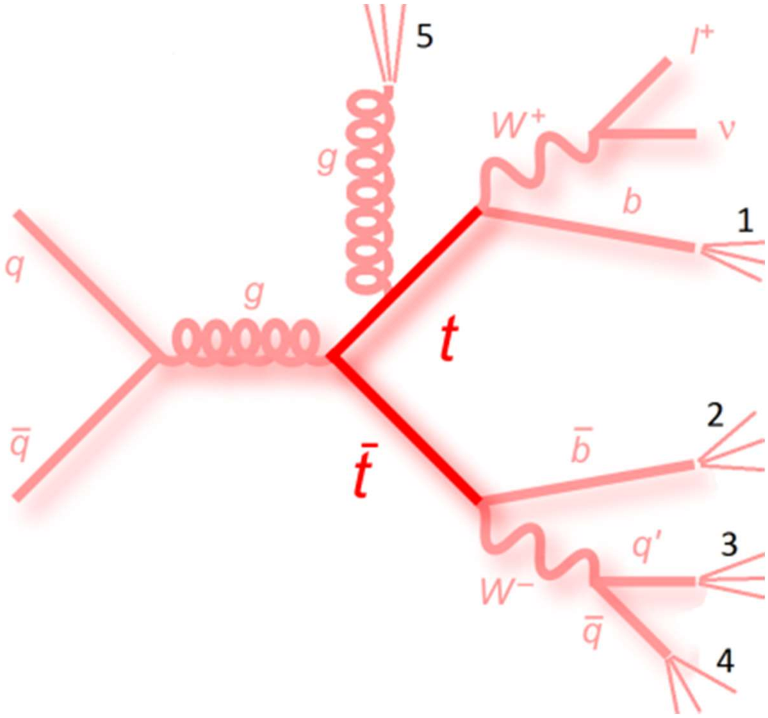

Figure 1: Semi-leptonic event: $pp \rightarrow t\bar{t} \rightarrow b\bar{b}WW \rightarrow b\bar{b}q\bar{q}\mu\nu_{\mu}$

A list of the all channels (signal and background) is presented in the table 1.

Table 1: List of channels and their cross-sections used for modelling

Channel	σ, pb; $\sqrt{s} = 8\ TeV$
$t\bar{t}$	225.2
W+Jets	36257.2
DY+Jets	3503.7
t (tW-channel)	11.1
$\bar{t}$ (t-channel)	30.7
$\bar{t}$ (tW-channel)	11.1
QCD	135000
t (t-channel)	56.4
WW	33.6
WZ	12.6
ZZ	5.2

In the study we explore first 6 background channels from the table including QCD background.

Selection steps for the semi-leptonic events :

Cut 1: Isolated muon with ${P_T}^{muon} > 26\ GeV$, $|\eta| < 2.1$;

Cut 2: Veto for additional tight isolated muons/electrons;

Cut 3: Two b-jets with ${P_T}^{b-jet} > 15\ GeV$, $|\eta| < 2.4$;

Cut 4: Two additional jets with ${P_T}^{jet} > 15\ GeV$, $|\eta| < 2.4$,
$|\,M_{jj} - M_W| < 10\ GeV$, $M_{jj} = \sqrt{(P_{jet_3} + P_{jet_4})^2}$.

3.1. Pythia6

Jets sample used in the study has been selected from events samples (signal and background electroweak channels) that had been obtained from the full CMS simulation modelling. “TuneZ2Star” are the generator settings used; the settings are the official settings used for CMS modelling at Run-I (2012). Protons are initial particles, their total energy in the c.m.s. is $\sqrt{s}$ = 8 TeV, that corresponds to the conditions of the events detecting at the CMS at Run-I.

3.2. Pythia8

Jets sample used in the study was obtained with the help of the Pythia8 generator [5] with default settings. Protons are initial particles; their total energy in the c.m.s. is $\sqrt{s}$ = 8 TeV and 14 TeV, that corresponds to the CMS detecting conditions at Run-I and Run-II.

Jets are found using the built-in JetFinder (SlowJet class). These algorithms are lite versions of the “fastjet3” [6] 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]).

A flavor-based jets dividing was performed using event partons information. Partons are selected at the state just before hadronization after parton branching processes finishing. The first method to classify quark/gluon jets puts a jet in accordance with a parton with a maximum transverse momentum P_T, which was selected from the number of the partons in the $\Delta R = \sqrt{\Delta\eta^2 + \Delta\varphi^2} < 0.3$ cone ($\Delta\eta$ and $\Delta\varphi$ – pseudorapidity and azimuth angle of the parton about the axis). The second method puts in accordance a jet and the nearest by the ΔR parton in the $\Delta R < 0.3$ cone. In the both methods there is a limitation of the P_T deviation between a parton and a jet: $\frac{\Delta p_T}{p_T} < 3$ [6]. Let $\Delta n = \langle n_q \rangle - \langle n_g \rangle$ be a difference between mean multiplicities in a jet of quark and gluon jet samples. Jets that were selected by the first method have Δn by 10-15% larger than in the second method. In the study the first method is used because it allows distinguishing quark and gluon jets characteristics more evidently.

3.3. Event and jet selections

Results of modelling and jets and events selection for the signal channel ($t\bar{t}$) and 6 background channels (listed in the table 1) are presented in the table 2.

Table 2. Channels list and their cross-sections

	$t\bar{t}$ [1]	Wjets [1]	DY [1]	t (tW-channel) [1]	$\bar{t}$ (t-channel) [1]	$\bar{t}$ (tW-channel) [1]	$t\bar{t}$ [2]	QCD [2]	$t\bar{t}$ [3]	QCD [3]
N_0^{event}	$6{,}74\cdot 10^6$	$18{,}4\cdot 10^6$	$4{,}16\cdot 10^6$	$1{,}94\cdot 10^6$	$0{,}498\cdot 10^6$	$0{,}493\cdot 10^6$	10^7	$908{,}7\cdot 10^6$	10^7	10^8
$\tilde{N}_0^{event}$ (cut 3)	95 468	124	633	2 198	4 533	2 161	$0{,}614\cdot 10^6$	8 404	$0{,}483\cdot 10^6$	724
Number of Jets, $P_T^{jet} \in [30, 400]$ GeV										
$\tilde{N}_{q\ selected}$ (cut 4)	12 437	4	7	254	30	227	65 800	588	55 099	42
$\tilde{N}_{g\ selected}$	5 319	2	12	74	42	96	54 000	888	46 485	74
Number of Jets per initial 10^7 event ($N_0^{event} = 10^7$), $P_T^{jet} \in [30, 400]$ GeV										
$N_{q\ selected}$	18 453	2,17	16,8	1 309	602	4 604	65 800	6,47	55 099	4,2
$N_{g\ selected}$	7 892	1,09	28,8	381	843	1 947	54 000	9,77	46 485	7,4
N_{q+g}	26 345	3,26	45,6	1690	1443	6551	119 800	16,24	101 584	11,6
σ	168,97	36257,2	3503,7	11,1	30,7	11,1	168,97	135000	731,22	559049
k_σ	1	214,6	20,7	0,0657	0,182	0,0657	1	798,96	1	764,54

In the first row there are presented numbers of generated events. In the second row there are numbers of the events after selection of events with at least two b-quarks (it corresponds to the "cut 3" selection conditions). Other table rows are the results of the "cut 4" selection, which includes following conditions: selection of the events with jets ($P_T{}^{jet} > 15$ GeV, $|\eta| < 2.1$) in the 20 GeV range of the two jets invariant mass equal to the W-boson mass (80 ± 10 GeV). We name the jets as "W-jets" further.

From the table one can see that in the selection made from the events with full CMS modelling (the first column) rate of the quark and gluon jets in the W-jets sample is about 2:1, and it has physical reasons because selections of the jets in the W-boson mass range should provide more quark jets that are produced by the W-boson decay (see figure 1). At the Pythia8 modelling (the 7^{th} and 9^{th} columns) quark jets selection purity is much worse. This mismatch requires a more detailed study of the conditions of samples selections. Besides qualitative conclusions can be made.

$k_\sigma = \frac{\sigma^{(B)}}{\sigma^{(t\bar{t})}}$ coefficient allows to calculate rates of the quark and gluon jets in the mixture of signal and background events. Its values are presented in the last row of the table 2. Background events are suppressed by the selection conditions. At the same time k_σ coefficient increases number of the background jets.

3.4. Quark jets fraction in the W-jets sample selected in semi-leptonic $t\bar{t}$-channel

Contribution of the electroweak backgrounds in the W-jets sample is presented in the table 3. Total contribution of all flavors is about 8.5%. Fraction of quark jets excluding backgrounds is 0.699. The background reduces the value by 1.3% (Pythia6). For the Pythia8 quark jets fraction is unexpectedly low: 0.55 for 8 TeV, 0.54 for 14 TeV. It is necessary to recheck these values. But it does not interfere with a shift of QCD events contribution. In the table 4 one can see QCD background does not alter α_q value.

Table 3. Fraction of EW background jets and fraction of quark jets α_q

	$\sqrt{s} = 8\,TeV$ $t\bar{t}$ & EW → Pythia6	$\sqrt{s} = 8\,TeV$ $t\bar{t}$ & EW → Pythia8	$\sqrt{s} = 14\,TeV$ $t\bar{t}$ & EW → Pythia8
Fraction of EW background jets (q+g) in jet sample			
$N_{q+g}{}^{(\mathrm{EW})}$	8,51%	-	-
Fraction of q-jets in "W-jet" sample			
$\langle \alpha_q{}^{(tt)} \rangle$	0.699±0.004	0,549	0,542
$\langle \alpha_q{}^{(tt+EW)} \rangle$	0. 689 ± 0.015	-	-

Table 4. Fraction of QCD background Jets and fraction of quark jets α_q

	$\sqrt{s} = 8\,TeV$ $t\bar{t}$, EW → Pythia6 QCD → Pythia8	$\sqrt{s} = 8\,TeV$ $t\bar{t}$ & QCD → Pythia8 EW → Pythia6	$\sqrt{s} = 14\,TeV$ $t\bar{t}$ & QCD → Pythia8 EW → Pythia6
Fraction of EW background jets (q+g) in jet sample			
$N_{q+g}{}^{(\mathrm{QCD})}$	33,00%	9,77%	8,03%
Fraction of q-jets in "W-jet" sample			
$\langle \alpha_q{}^{(tt)} \rangle$	0.699±0.004	0,549	0,542
$\langle \alpha_q{}^{(tt+\mathrm{QCD})} \rangle$	0,700	0,549	0,542
$\langle \alpha_q{}^{(tt+\mathrm{QCD+EW}}$	0,690	0,558	0,553

4. Conclusion

Preliminary estimates of the backgrounds contribution to the shift of the quark jets fraction in the W-jets samples show that electroweak background contribution is essential but not crucial for the measurements. QCD background contribution is negligible.

References

[1] "CDF," *Phys.Rev.Lett. 94, 171802,* 2005.

[2] "OPAL," *Eur.Phys.J. C1,* pp. 479-494, 1998.

[3] I.M.Dremin, "Uspehi Phys. Nauk (in russian), V.172. N 5," 2002.

[4] A. M.Kaur, *Advances in High Energy Physics V 2013,* p. 11.

[5] "Pythia," [Online]. Available: http://home.thep.lu.se/~torbjorn/Pythia.html.

[6] "FastJet," [Online]. Available: http://fastjet.fr/.