B-Tagging in CMS
HCP conference 17-11-2011
Cristina Ferro, on behalf of CMS collaboration

Institut Pluridisciplinaire Hubert CURIEN STRASBOURG
UNIVERSITÉ DE STRASBOURG
CNRS IN2P3 Les deux infinis

*The identification of b-jets is crucial to characterize a variety of Standard Model (SM) and discovery channels like the measurement of bottom or top quark production, the search for Higgs boson, and many New Physics scenarios.
*The b-tag algorithms in CMS rely on the long life time, high mass and large momentum fraction of b hadrons produced in b-quark jets, as well as on the presence of soft leptons from semi-leptonic b-decays.

B-TAGGING OBSERVABLES

Effect of Pile-Up

*Due to the high instantaneous luminosity during the 2011 data taking, the number of collision taking place in the same bunch crossing (pile-up events) is of the order of 5 to 11 on average
*The presence of pile-up strongly increases the track multiplicity in the events (**Fig. 1: Left**). A specific offline selection on the tracks is applied to remove the tracks originating from pile-up [1] (**Fig.1: Right**).
*Variation in the pile-up condition over the running period is taken in to account by weighting the number of pile-up interaction in the MC simulation.

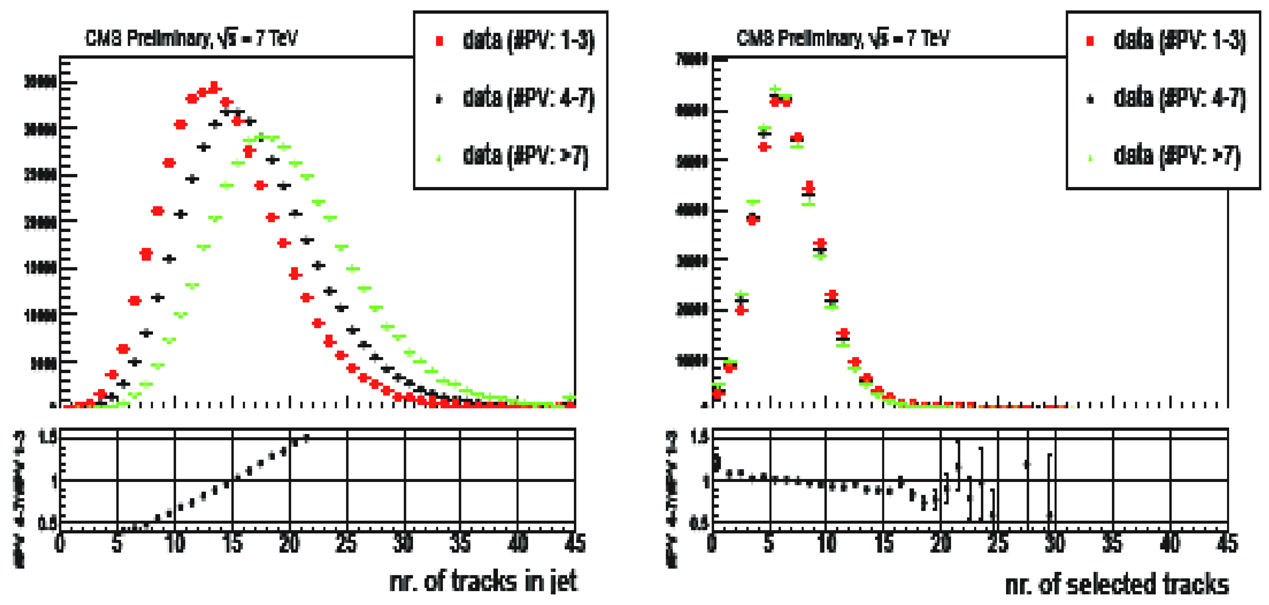


Fig.1: (Left) number of tracks associated to a jets without any selection cut. **(Right)** Number of tracks associated to a jet passig the selection cut

1) Impact Parameter

*The impact parameter (IP) is the distance between the track and the primary interaction vertex (PV) at the point of closest approach (**Fig.2**).
*The IP is calculated in 3 dimensions thanks to the good x-y-z resolution provided by the pixel detector.
*The IP is positive (negative) if the track is produced downstream (upstream) with respect to the PV (**Fig.2**).
*The impact parameter significance IP/ σ(IP) is used in order to take into account resolution effects.

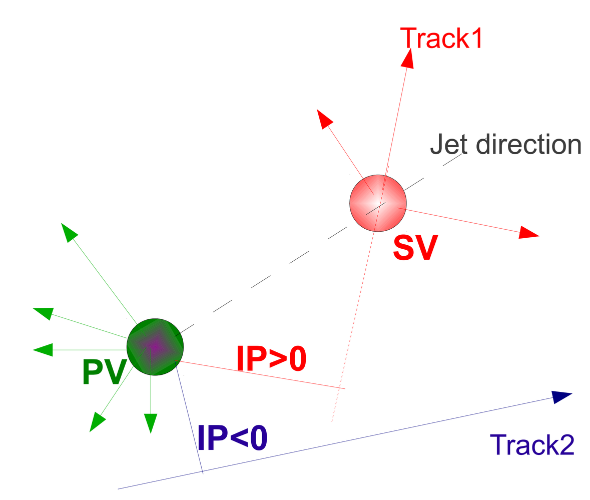


Fig.2: Geometric definition of the 3D Impact Parameter of the tracks.

Effects of the finite b lifetime on the IP /σ(IP)

*IP is Lorentz invariant and due to the b-hadron lifetime the typical IP scale is set by cτ ~480 μm.
*For light quarks (u, d, s) or gluons (g), the signed IP/σ(IP) is expected to be symmetric;
for weakly decaying b-hadrons, the IP/σ(IP) is mostly positive [2] (**Fig.3: left and center**).
*One can use the negative tail of the IP/σ(IP) distribution to extract the probability density function for tracks not coming from b/c-jets.

2) Secondary Vertex

*Thanks to the high resolution of the CMS traking system, one can reconstruct the Secondary Vertex (SV), the point where the b-hadron decays (**Fig.2**).

3) Lepton

*We can also use muon from b-hadron decay to tag the b-jets. In (**Fig.3: Right**) we show the p_T of the muon relative to the jet direction [1].

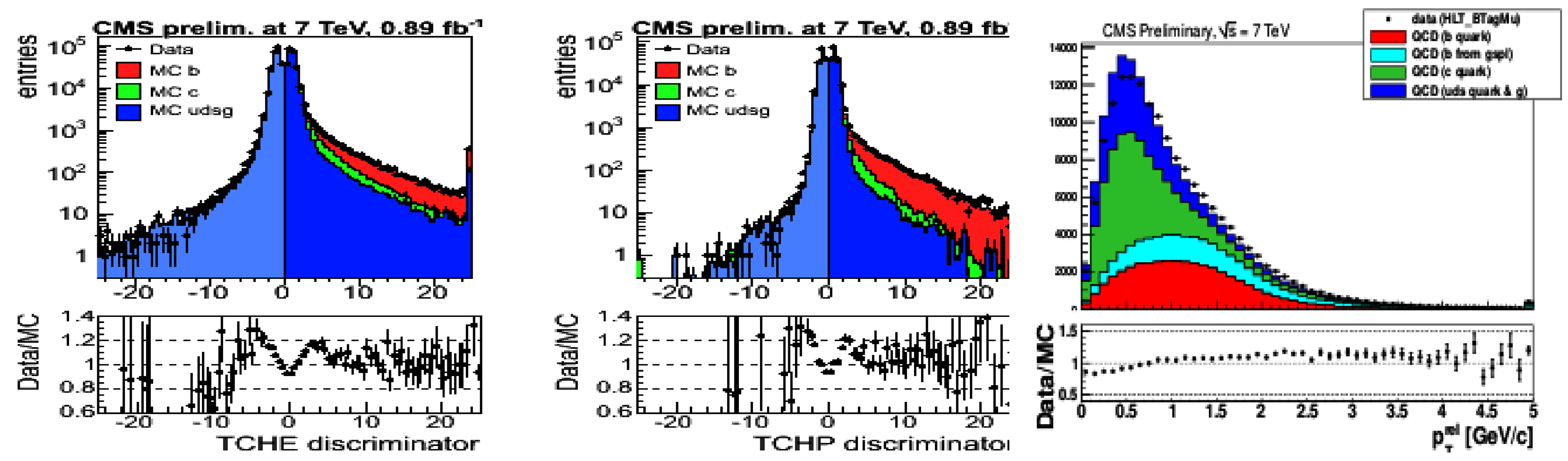


Fig.3: (Left , center)IP/σ(IP) distribution. **(Right)** p_T of the muon relative to the jet axis.

B-TAGGING ALGORITHMS

The output of each b-tagging algorithm in CMS is a "discriminator" value on which the user can cut on to select different regions in the efficiency versus purity phase space. We use mainly four different algorithms[1, 2]:

1)Track counting algorithm: it identifies a b-jet if there are at least N tracks each with a significance of the impact parameter above a given threshold. The tracks are ordered in decreasing IP/σ(IP) and the discriminator is the impact parameter significance of the Nth track .
To get an high b-jet efficiency we can use the IP/σ(IP) of the second track (TCHE),
to select b-jets with high purity the third track is the better choise (TCHP). (**Fig.3: Left, Center**)

2) Jet Probability algorithm The jet probability algorithm combines information from all the selected tracks in the jet to compute a "probability" for tracks to originate from the PV. (**Fig.4 : Left**). [1]
The JetBprobability is then defined in a similar way but giving more weight to the four most displaced tracks. (**Fig.4 Right**)

3) Soft-Lepton tagging algorithms rely on the properties of muons or electrons from semileptonic B-decay. (**Fig.3: Right**)

4) Secondary Vertex tagging algorithms rely on the reconstruction of at least one secondary vertex. The significance of the 3D flight distance is used as a discriminating variable. Two variants based on the number of tracks at SV are considered: $N_{tr} \geq 2$ for "high efficiency" (SSVHE), (**Fig.4: Bottom left**) and $N_{tr} \geq 3$ for "higt purity" (SSVHP) [2] (**Fig.4: Bottom center**). The **"combined secondary vertex"** algorithm provides discrimination even when no secondary vertices are found. The "Mass of reconstructed charged particles the Secondary Vertex" is used to measure the b-tagged sample purity [1] (**Fig.4: Bottom right**)

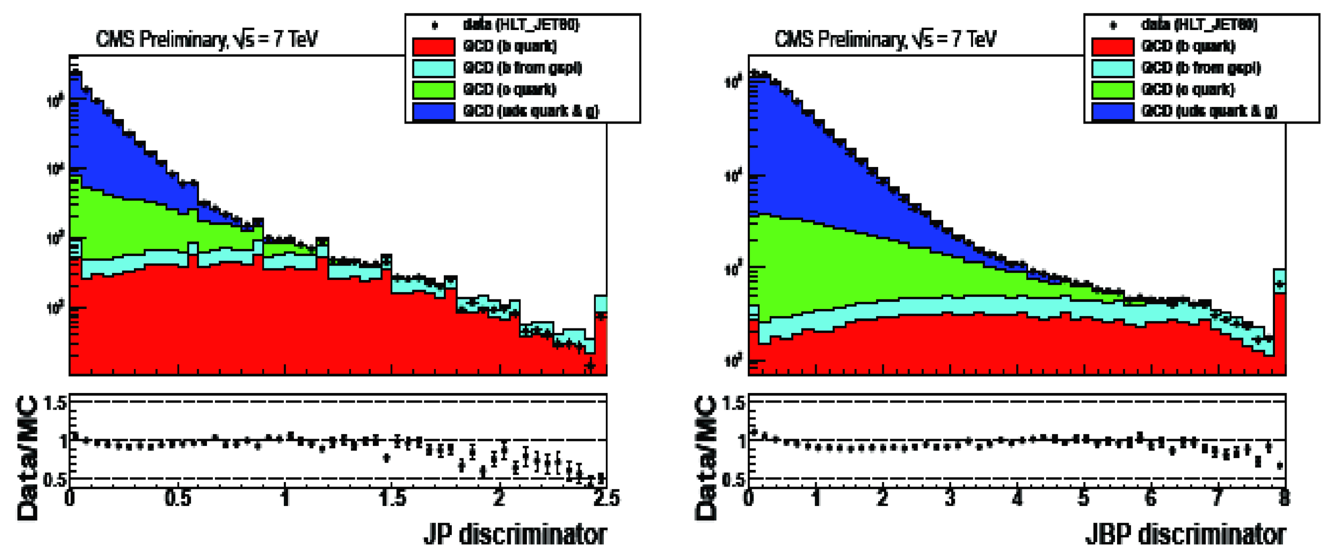
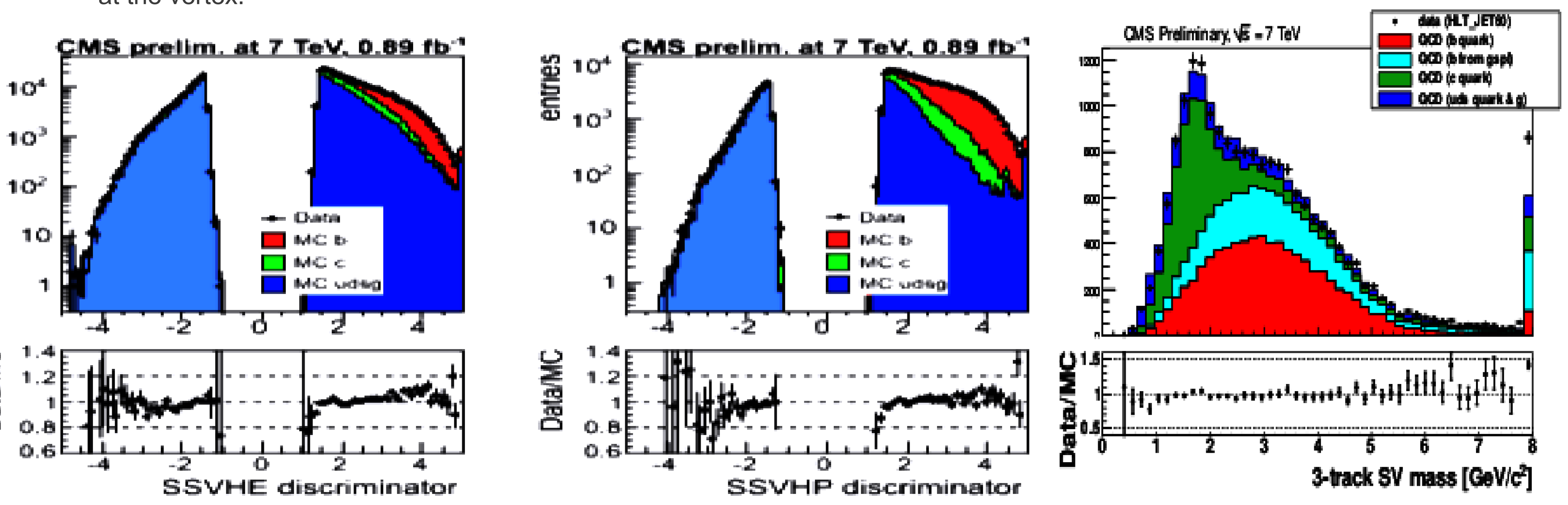


Fig. 4: (Left) Jet Probability discriminator. **(Right)** JetBProbability discriminator. **(Bottom left)** Simple Secondary Vertex High Efficiency Discriminator. **(Bottom center)** Simple Secondary Vertex Hihg Purity. **(Bottom right)** vertex mass with three or more tracks at the vertex.



PERFORMANCE OF THE TAGGERS

Varying the cuts on the discriminator we obtain a different performance of the tagger.
We establish standard operating points as, ``loose'' (L), ``medium'' (M), and ``tight'' (T), being the value at which the tagging of light jets is estimated from MC to be 10%, 1%, or 0.1%, respectively.
In (**Fig.5:Left**) the performance for different taggers are shown.
In (**Fig.6:Right**) the effects of the pile-up on the performance of the TCHE tagger is presented [2].

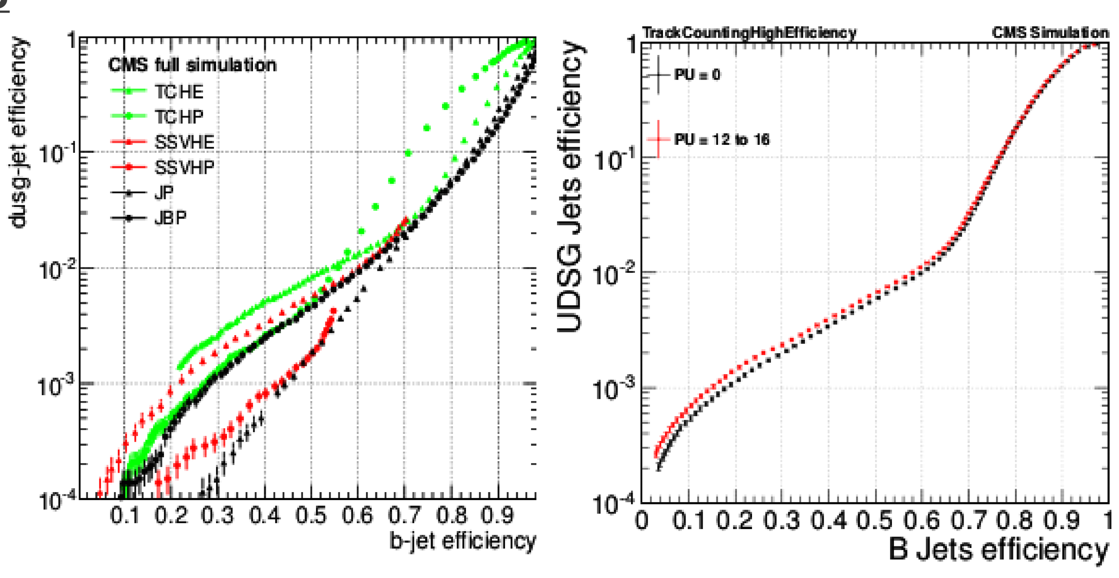


Fig.5: (Left) Performance of all b-taggers obtained on the simulated QCD events. The performance are shown in the form of udsg jets versus b-jets [2].
Fig.6: (Right) Light flavor mistag efficiency versus b-tagging efficiency for different pile-up scenario, for the TCHE tagger .

MISTAGRATE RATE AND B-TAG EFFICIENCY MEASUREMENTS FROM DATA

We can estimate the mistag rate using the negative tail of discriminators (tracks with negative IP or using SV with negative decay length). The measured mistag rates in data and data/MC scale factors are presented as a function of the jet p_T (**Fig:8**) for the TCHE algorithm and working point (M) [1].
We measure the b-tag efficiency in jets with a semi-muonic b-hadron decay, by using the information of the p_T of the muon relative to the jet axis and the request of another b-tagged jet in the same event (**Fig.9**)

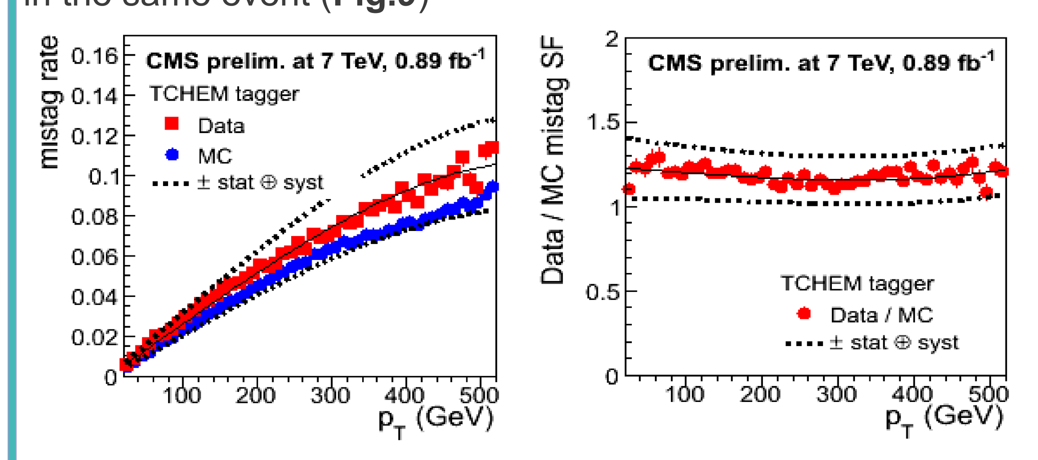
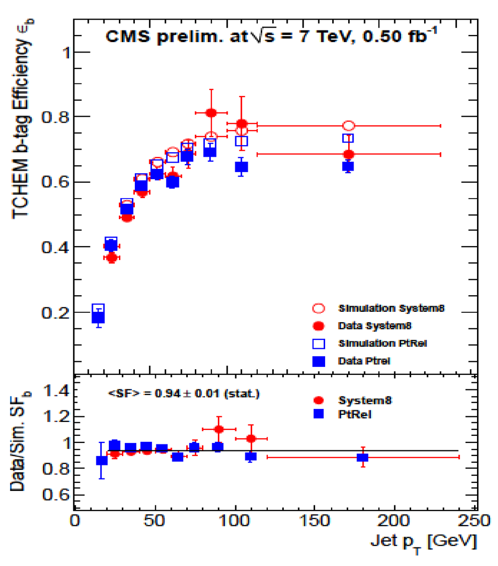


Fig8: mistag rate as function of the jet p_T and Data/MC scale factor

Fig.9: Measured b-tagging efficiency and data/MC scale factor as a function of the jet p_T



B-TAGGING AT TRIGGER LEVEL

Data for some physics analyses in 2011 are collected using dedicated b-tag triggers [1,3].(**Fig:10**)
The HLT algorithms used to select b jets using tracking are implemented in three steps:

1) jets are selected without any tracking requirements.

2) the jets are matched to tracks reconstructed using the Pixel Tracker alone
.
3)b-quark candidates are selected if they have at least one or two tracks with a 3D impact parameter Significance above a given threshold.

The motivation for applying b-tagging in the trigger is a reduction of the trigger rates, while keeping the signal efficiency high at the same time.
The typical rate reduction is a factor of 5-10.

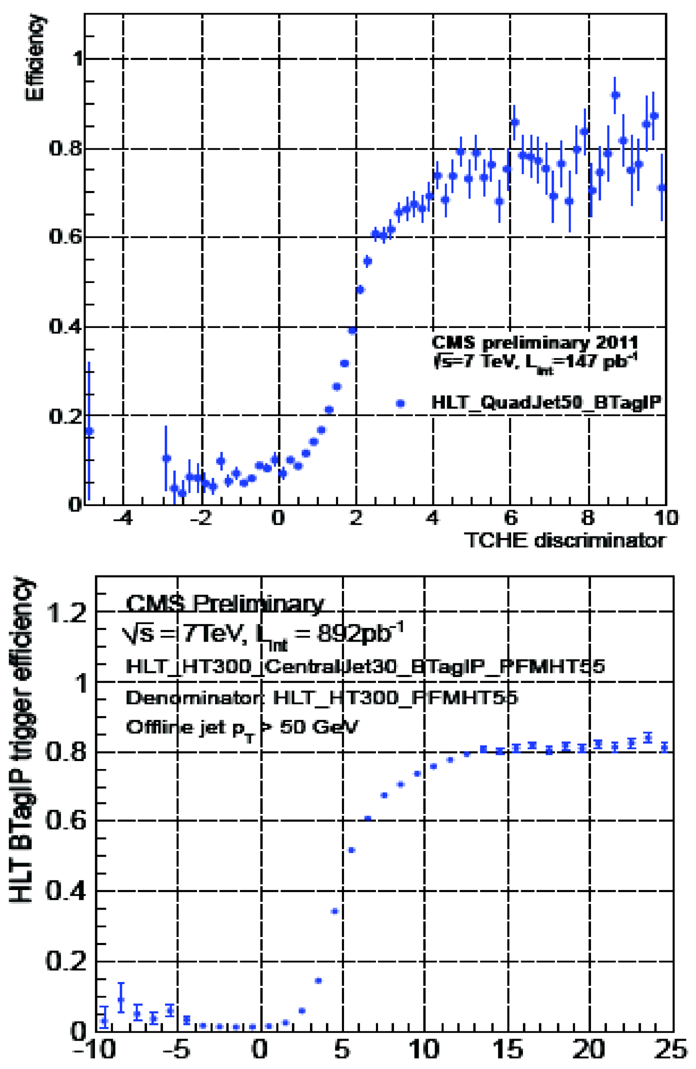


Fig 10:(top) performance of the Track Counting High Efficiency discriminator in a 4-jet trigger. **(bottom):** performance of the same discriminator in a trigger requiring missing transverse energy of clustered objects MHT > 55 GeV and HT > 300 GeV and jets with transverse energy above 30 GeV.

PHYSICS RESULTS

Many measurements have been obtained using the b-tagging algorithms at √s = 7 TeV:

B-PHYSICS:

- Inclusive production cross section of b-jets [4].
- bb angular correlation based on Secondary Vertex reconstruction [5].
- Search for Bs -> μμ decay. [6]

EW PHYSICS:

- Measurement of the WW, WZ, ZZ cross section. [7]
- Measurement of associated charm production in W final state. [8]

Top-PHYSICS:

- Cross-section measurement of top pair production in various final states: dileptons [10,11], lepton+jets [12], all hadronic [3].
- Single top in t channel [13].
- Top mass measurement [14].

New PHYSICS:

- Search for supersymmetry in events with b-jets and missing transverse momentum. [15]
- Search for supersymmetry in all hadronic events.[16]
- Search for an Heavy Bottom-like quark. [17]
- Search for an Heavy Top-like quark.[18]
- Search for pair production of a fourth-generation t' quark in the lepton-plus-jets channel. [19]
- Inclusive search for a fourth generation of quarks. [20]

REFERENCES:

[1] "Status of b-tagging tools for 2011 data analysis". CMS PAS BTV-11-002
[2] "Performance of b-jet identification in CMS". CMS PAS BTV-11-001
[3] "Measurement of the ttbar production cross section in the fully hadronic decay channel in pp collisions at √s = 7 TeV". CMS PAS TOP-11-007
[4] "Inclusive b-hadron production cross section with muons in pp collision at √s = 7 TeV". J. High Energy Phys. 03 (2011) 090
[5] "Measurement of BBbar Angular Colleration based on Secondary Vertex Reconstruction at √s = 7 TeV ". J. High Energy Phys. 03 (2011) 136
[6] "Search for Bs⁰ -> μμ and Bs-> μμ decays in pp collisions at = 7 TeV". CERN-PH-EP-2011-120
[7] "Measurements of the WW,WZ,ZZ, cross cection at CMS". CMS PAS EWK-11-010
[8] "Measurement of associated charm production in W final state I pp collision at √s=7 TeV". CMS PAS EWK-11-013
[9] "First measurement of the cross section for top-quark pair production in p–p collisions at √s = 7 TeV". Phys. Lett. B 695 (2011) 424-443
[10] "Measurement of the ttbar production cross section in the dilepton channel in pp collisions at √s = 7 TeV with a luminosity of 1.14 fb-1". CMS PAS TOP-11-005
[11] "First measurement of the ttbar production cross section in the dilepton channel with tau leptons in the final state in pp collisions at √s=7 TeV". CMS PAS TOP-11-006
[12] "Measurement of the ttbar Pair Production Cross Section at √s = 7 TeV using b-quark Jet Identification Techniques in Lepton + Jet Events". CMS PAS TOP-11-003
[13] "Measurement of the t-channel single top quark production cross section in pp collision at √s = 7 TeV". Phys.Rev.Lett.107(2011)091802
[14] "Measurement of the ttbar production cross section and the top quark mass In the Dilepton channel in pp collision at√s = 7 TeV". J.High Energy Phys.07(2011)049.
[15] "Search for supersimmetry in events with b-jets and missing transverse momentum at LHC". J.High Energy Phys.07(2011)091802
[16] "Search for supersymmetry In all hadronic events with b-jet". CMS PAS SUS-11-006
[17] "Search for an Heavy Bottom-like qurak in 1.14 fb-1 of pp collision at √s = 7 TeV". CMS PAS EXO-11-036
[18] "Search for an Heavy Top-like quark in the Dilepton Final state in pp collision at √s = 7 TeV". CMS PAS EXO-11-050
[19] "Search for pair production of a fourth-generation t' quark in the lepton-plus-jets channel with the CMS experiment". CMS PAS EXO-11-051
[20] "Inclusive search for a fourth generation of quarks with the CMS experiment". CMS PAS EXO-11-054