# B-Tagging in CMS


**LHCC meeting 23/03/2011**

**Cristina Ferro, on behalf of CMS collaboration**


- 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

### 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.1).
- The IP is calculated in 3 dimensions thanks to the good z resolution provided by the pixel detector.
- The IP is positive (negative) if the track is produced downstream (upstream) with respect to the PV. The impact parameter significance *IP/ $\sigma$(IP)* is used in order to take into account resolution effects.

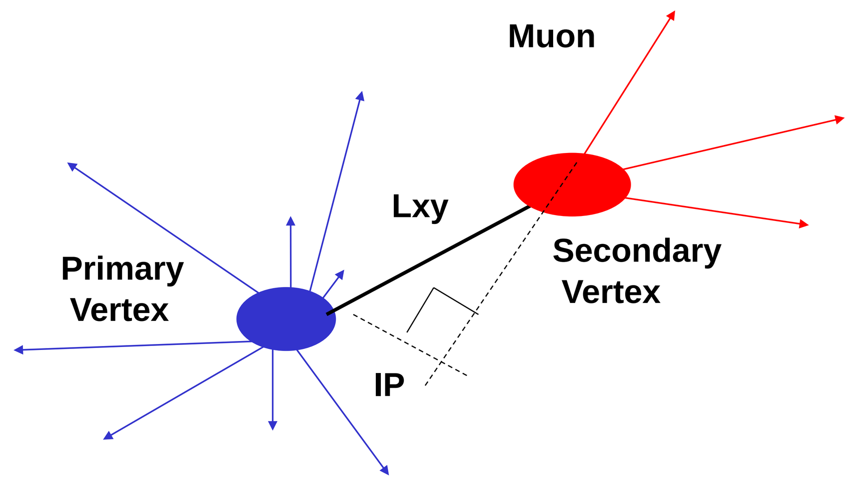



**Fig.1:** Geometric definition of the 3D Impact Parameter of the tracks., and, picture of SV

**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.
- From light quarks (u, d, s) or gluons (g), the signed IP/σ(IP) is expected to be symmetric; for b-hadrons decaying weakly, the IP/σ(IP) is mostly positive [1] (Fig.2: 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-jets.

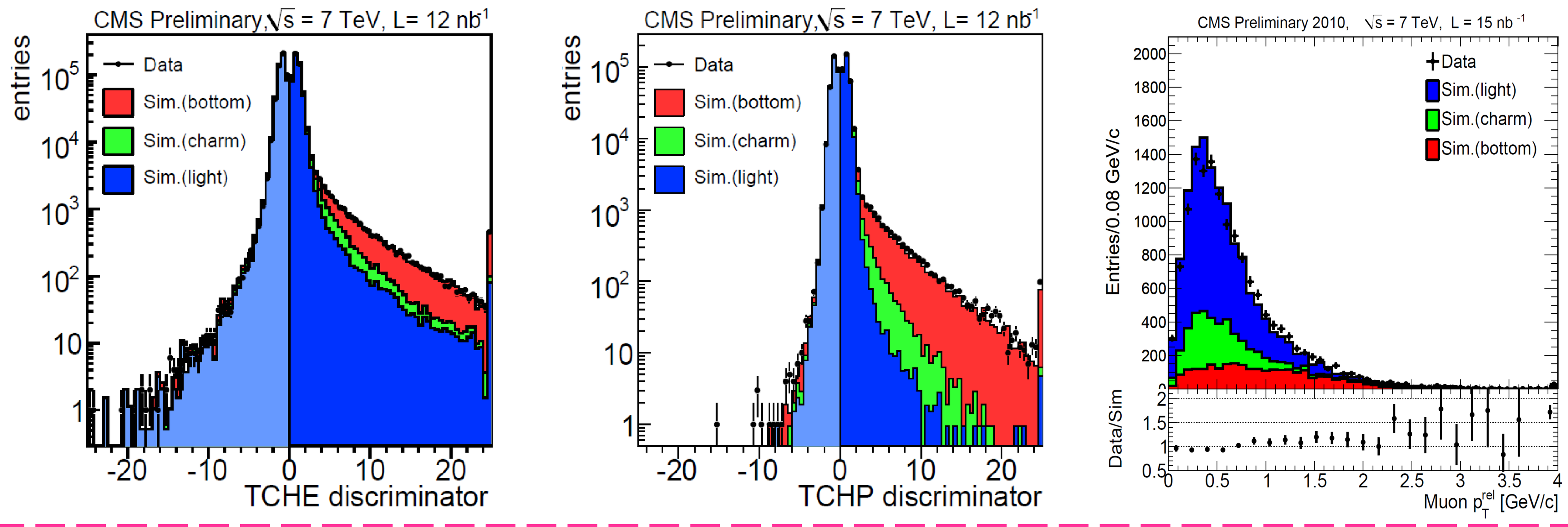
### 2) Secondary Vertex

- 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.1)

### 3) Lepton

- We can also use muon from b-hadron decay to tag the b-jets. In Fig.2(right) we show the $p_T$ of the muon relative to the jet direction [1].

**Fig.2:** **(Left, center)** IP/σ(IP) distribution. **(Right)** pT 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 exceeding 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.2)

**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.

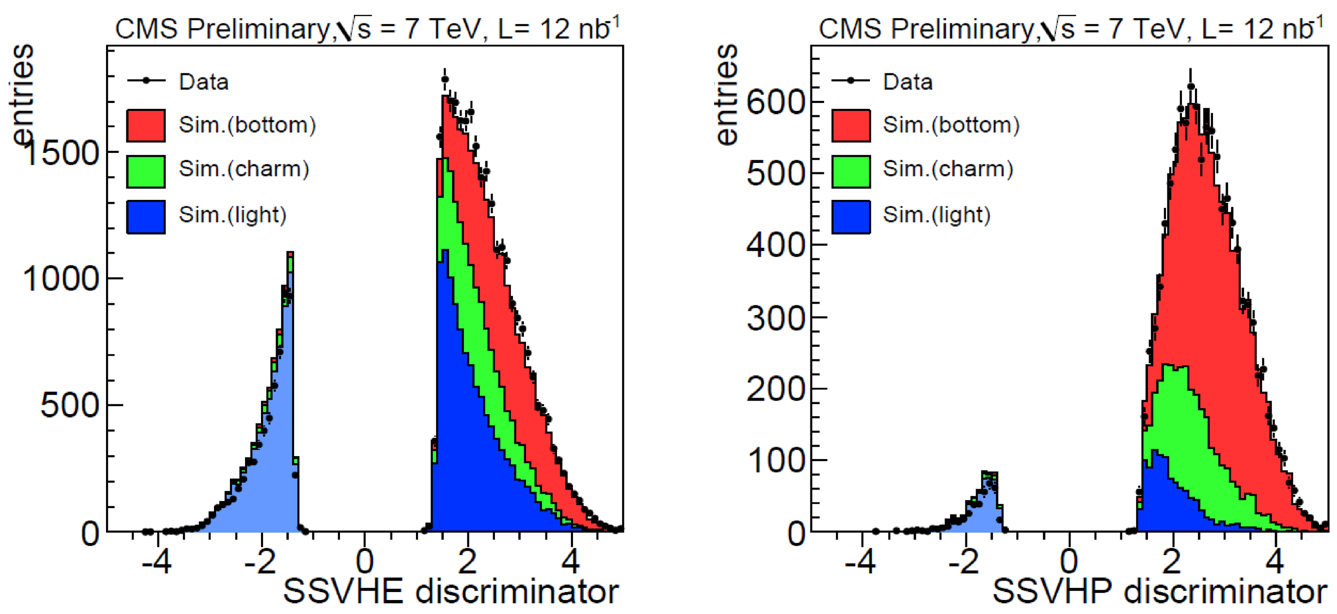



**Fig.3: Secondary Vertex Discriminator**

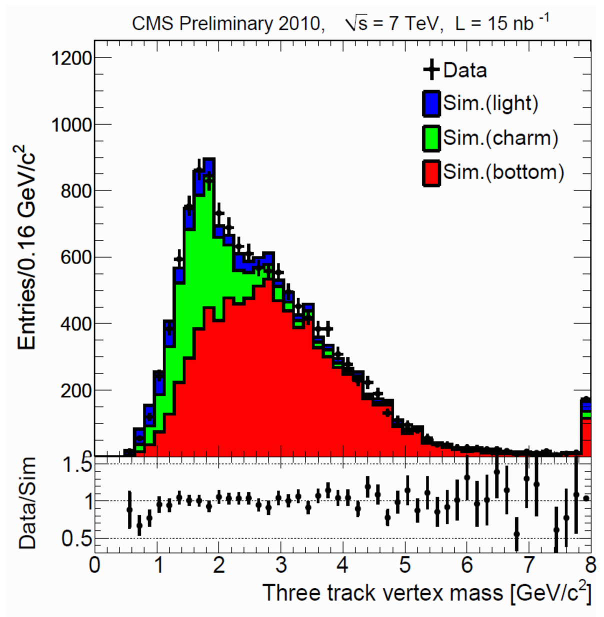


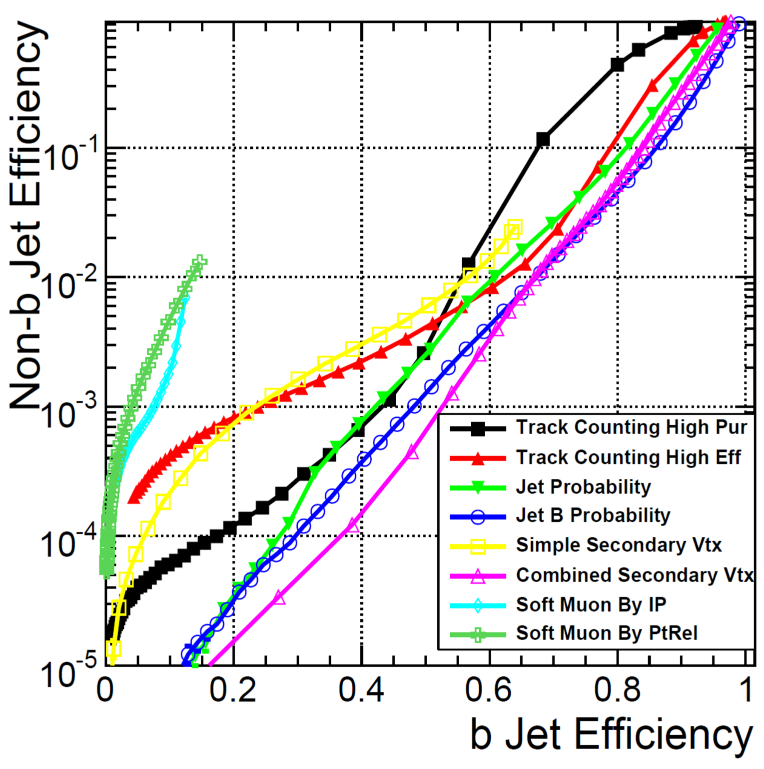

**Fig.4: SSVHP mass distribution**

**3) Soft-Lepton tagging algorithms** rely on the properties of muons or electrons from semileptonic b-decay. (Fig.2 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), and $N_{tr} \geq 3$ for “higt purity” (SSVHP) (Fig.3).
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 [3, 4, 5]. (Fig.4)

## PERFORMANCE OF THE TAGGERS

Varying the cuts on the discriminator we obtain a different performance of the tagger, so 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 are shown the results for different taggers. In Fig.6 the dependence of the b-efficiency as a function of the $p_T$ and rapidity are presented.



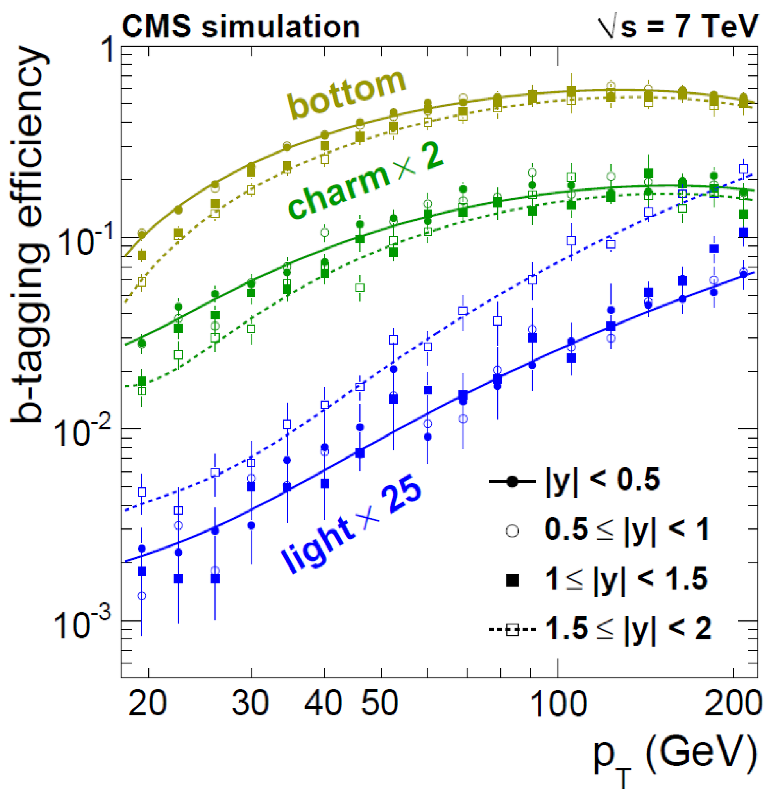




**Fig.5: (Left)** Performance of all b-taggers obtained on the simulated QCD events. The performance are shown in the form of uds jets versus b-jets [2].
**Fig.6: (Right)** The b-tag efficiency and light, charm mistag rates from MC truth for SSVHP tagger, versus $p_T$ for different rapidity intervals [3].

## MISTAG ESTIMATION FROM DATA

We can estimate the mistag rate using a negative discriminator (using tracks with negative IP or using SV with negative decay length). The measured mistag rates in data and data/MC scale factors are presented in Fig. 7 as a function of the jet $p_T$ for the TCHE algorithm and working point (L) [1].

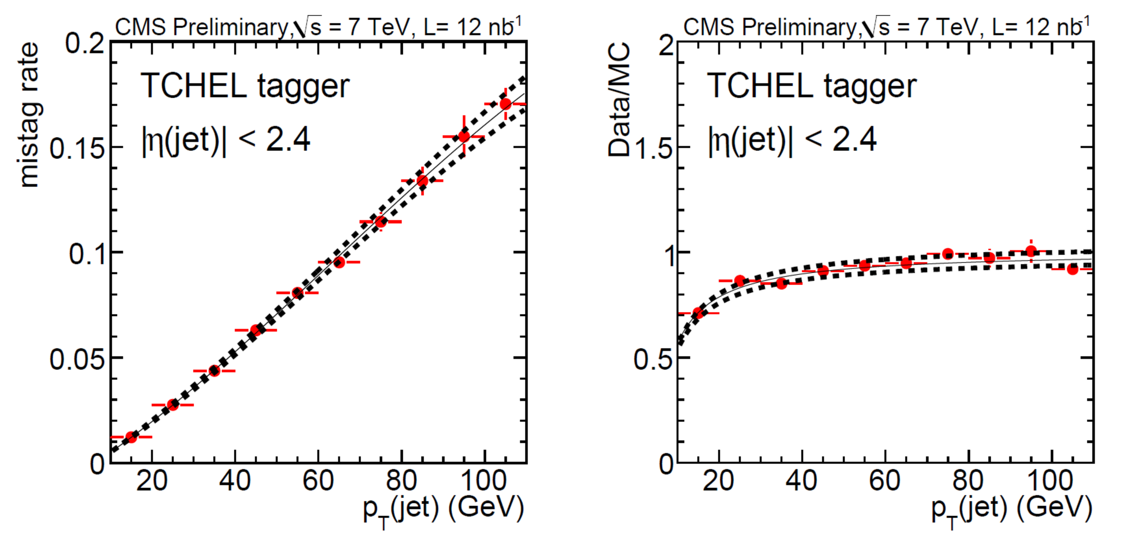



**Fig.7:** Mistag rate and data/MC scale factor, as a function of the jet $p_T$.

## B-TAG EFFICIENCY FROM DATA

One can estimates the b-tag efficiency from data using jets with a muon from semi-leptonic b-decay [1], or using top quark pair event candidates [6].

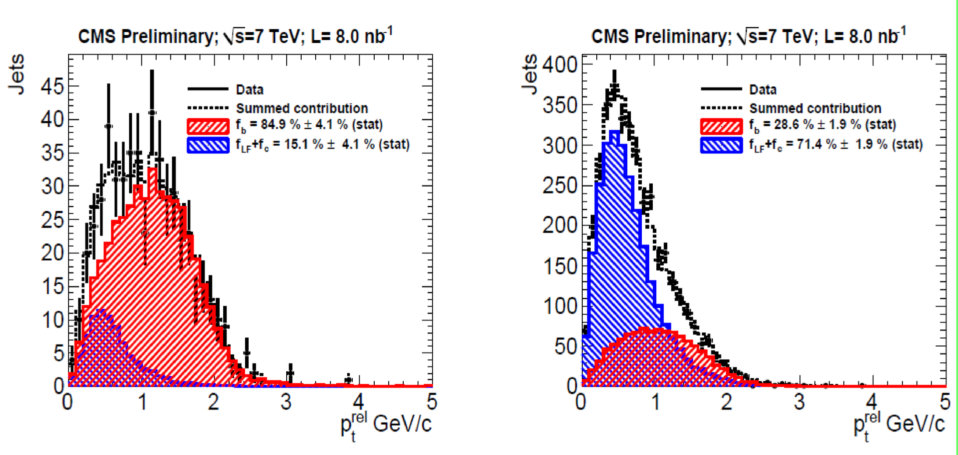



**Fig.8:** Fits of the muon $p_T$(rel) distributions to b and light flavour templates for muons that pass (left) or fail (right) the TCHPM algorithm.

Using jets with a muon, pT(rel) templates are obtained from data for the light jets, and from MC for the b and c jets. A fit is performed to the data, with or without passing b-tagging . From each fit we extract the fraction of b-jets in data. These fractions are used to estimate the b-tagging efficiency.

## 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 [3,7].
- bb angular correlation based on Secondary Vertex reconstruction [8].

**EW physics:**

- Observation of Z+bb events[4]. (Fig.9)

**Top-physics:**

- top-quark pair-production cross section in the dilepton channel [6,9], and in the lepton + jets channel [5].
- production cross section for single top t-channel [10].
- measurement of the top quark mass [11]

**New Physics:**

- search for top quark pair resonances [12]
- Search for SUSY with b-jets in the final state [13].

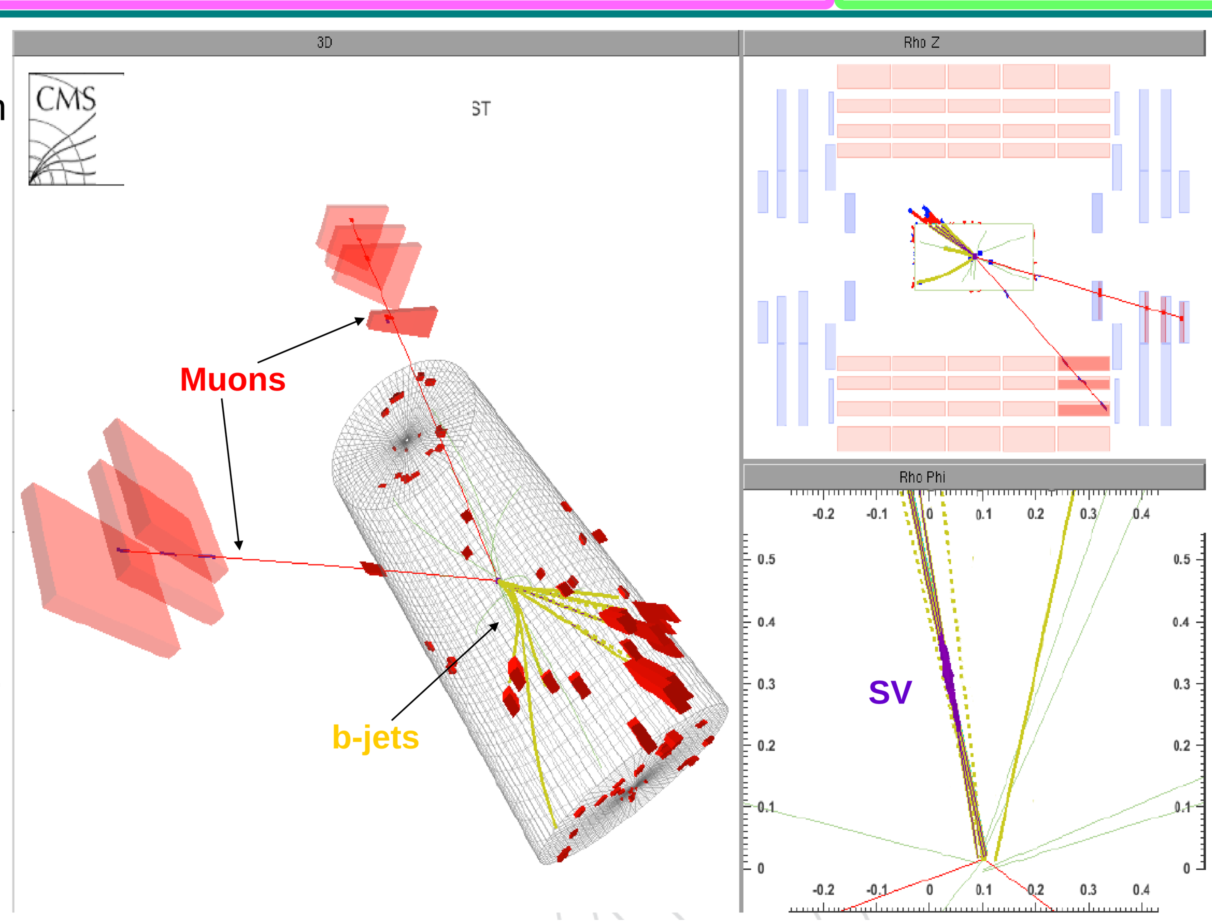



Fig.9: Event display of a Z+bb, Z->μμ candidate event in 3D (left), in ρ-z (top-right) rapresentation, and ρ-φ around the primary vertex (bottom-right).

## REFERENCES


[1] “Commissioning of b-jet dentification with pp collisions at √s = 7 TeV”, CMS PAS BTV 10 001 (2010).
[2] “Algorithms for b Jet Identification in CMS”, CMS PAS BTV 09 001 (2009).
[3] Inclusive b-jet prodaction in pp collision at √s = 7 TeV”, CMS PAS BPH-10-009, (2009)
[4] “Observation of Z+b, Z-> ee, μμ with CMS at √s = 7 TeV”, CMS PAS EWK-10-015, (2010).
[5] “Measurement of the tt Pair Production Cross Section at √s = 7TeV using b-quark Jet Identification Techniques in Lepton + Jet Events”, CMS PAS TOP-10-003, (2010)
[6] “Measurement of the top-quark pair-production cross section in the dilepton channel at √s = 7 TeV”, CMS PAS TOP-10-005, (2010)
[7] “Inclusive b-hadron production cross section with muons in pp collision at √s = 7 TeV” CERN-PH-EP-2010-088 (2011)
[8] “Measurement of BBbar Angular Colleration based on Secondary Vertex Reconstruction at √s = 7 TeV”, CERN-PH-EP-2010-093 (2011)
[9] “First measurement of the cross section for top-quark pair production in proton–proton collisions at √s = 7 TeV”, /Phys. Lett. B/ 695 (2011) 424-443
[10] “Measurement of the single top t-channel cross section in pp collision at √s = 7 TeV”, CMS PAS TOP-10-008, (2010).
[11] “First measurement of the top quark mass in pp collisions at √s = 7 TeV”,CMS PAS TOP-10-006,(2010).
[12] Search for resonances in semi-leptonic top-pair decays close to production threshold", CMS PAS TOP-10-007 (2010)
[13] “A Search for New Physics in b-tagged dijet and multi-jet events with Missing Energy in pp collisions at 7 TeV" CMS PAS SUS-10-011, (2010).