Measuring b-tag performance with data : mistags

Daniel Bloch, Jérémy Andrea

- Based on Tevatron experience, set a strategy to evaluate b-tagging and light (udsg)-tagging efficiencies
  using, as much as possible, the “real” Data themselves.

- Of course in CMS, it can be seen as an academic exercise, but one can look in the Monte-Carlo, with the present b-taggers, if such a strategy is reliable or not.

- Main motivation at Tevatron :
  the MC overestimates the tagging efficiencies by ~20-30%.

- One can hope that it will not be the case in CMS after a few years of running, but at least at the beginning, the strategy may be as follows : one can distinguish 3 steps

Step 1 : "Taggability"

- Taggability = efficiency for a jet to have enough tracks pointing to the primary vertex
- following Phys.TDR1-12.2.1.1, one can require at least 2 tracks with pt > 1 GeV, $|IP_{R\varphi}| < 2$ mm, ≥ 8 hits, etc.... (should be the same for all lifetime taggers !)

- In real Data : the taggability can be computed using multi-jet triggers, or electromagnetic triggers, or using high pt lepton triggers and relaxing the lepton-id requirements.

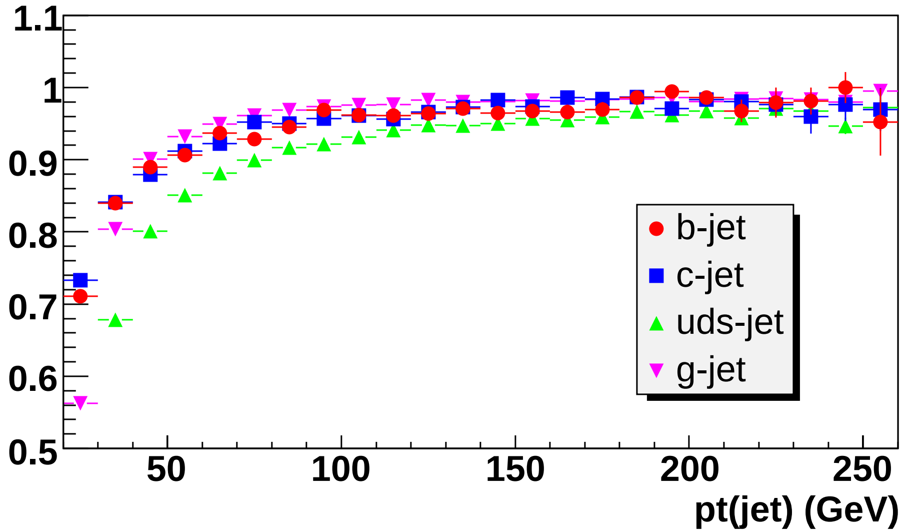
- Taggability is parametrized versus jet p_T and η (with factorization)
- It should not depend (too much) on the jet flavour

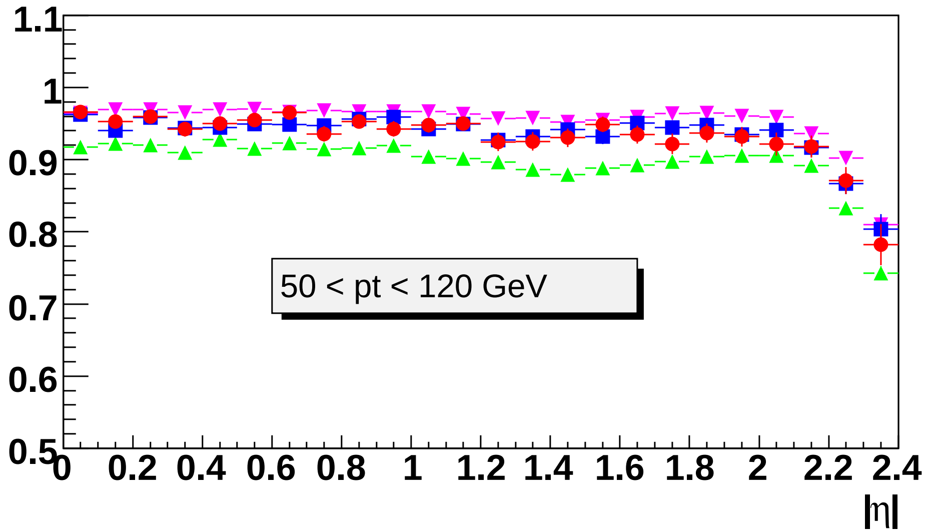
- Systematics are estimated by comparing different data samples, varying some cuts, …: ususally they are around ±2-3%

- Note that in DØ, the ratio between Data and MC is found to be taggability(Data) / taggability(MC) ≈ 0.85 … (which justifies this step 1)

Taggability in MC

- RunTrackCounting in BReco/BtagPTDR/test from ORCA_8_13_3
- Use qcd MC (ORCA_8_7_1) 250k events with pt in range 50-80 + 80-120 + 120-170 + 170-230
- Significant difference between uds and (c/b or g) jets ? (but same behaviour with “physics” or “algorithmic” definitions, see next slide)
- Is the MC too optimistic anyway ?





How to assign a flavour to jets ?

(at least) **2 ways** (cf. Phys.TDR1-12.2.1.2) **:**

- "physics" definition :
  consider all *initial* partons associated to jets within $\Delta R<0.3$,
  set flavour = 0 if no initial parton found (may be a radiative gluon)
  → used for lifetime taggers (but large amount of flavour = 0…)

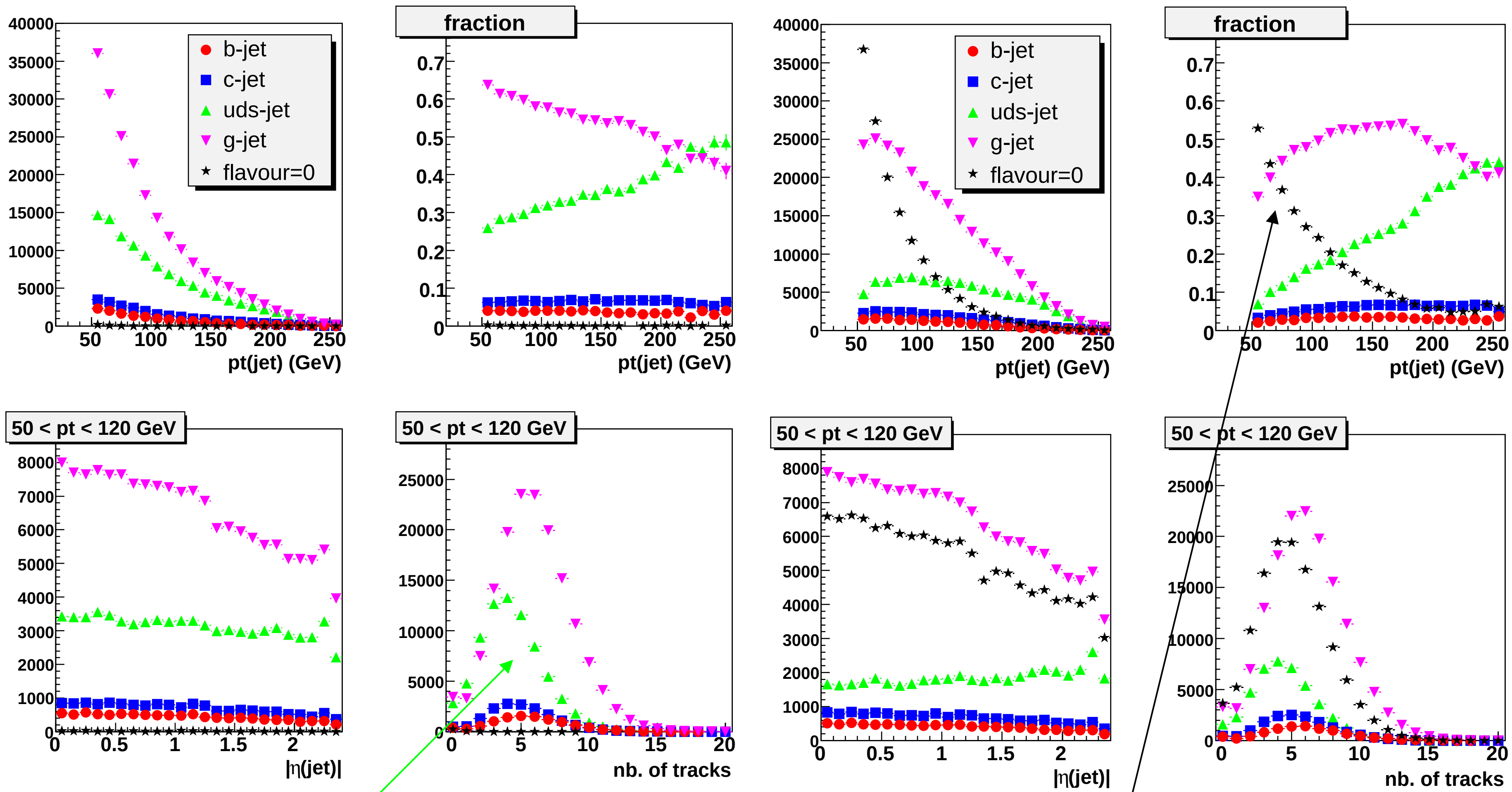
- "algorithmic" definition :
  set "b" if *final* b-quark found within $\Delta R<0.3$, otherwise set "c" if c-quark is found, otherwise set "uds" or "g"
  → used here because all jets are used in real life…

- Some flavour = 0 may remain if jet from pile-up event or tau-jet or noise…

- Note that here : jets with g→bb splitting are called "b"
  and jets with g→cc splitting are called "c" (if not already "b")

Jet flavour in qcd MC (cont.)

algorithmic **physics**



track multiplicity smaller in uds than in g-jets

large fraction of flavour=0

Step 2 : parameterizations

- For each jet of a given flavour, the tagging efficiency is then defined as :

  ε_{tag} = number of tagged jets / number of taggable jets

- As the MC may not describe well the Data, the tagging efficiencies are
  1. Computed for a limited set of **working points** (wp) : for instance corresponding to some average mistag rates of 0.1%, 0.3%, 0.5%, 1%, 2% and 4%
  2. Mainly estimated **from the Data** (see below for the mistags)
  3. Then parameterized : **Tag Rate Function**

     $$\text{TRF} = \varepsilon_{tag}(\text{pt}, \eta, \text{flavour,wp})$$

     Some physics analyses may also prefer to use **Scale Factors**

     $$\text{SF}(\text{pt}, \eta, \text{flavour,wp}) = \varepsilon_{tag}(\text{Data}) / \varepsilon_{tag}(\text{MC})$$

     (with pt / η factorization)

Step 3 : analysis

- **For the real Data :** apply the tagger to each jet

- **For the MC :** weight each jet of a given flavour by the product of its trigger efficiency(pt, η)
  taggability(pt, η) and
  TRF(pt, η, flavour,wp) (or SF(pt, η, flavour,wp))

Mistag estimate

- Use a sample of real Data with a "small" contamination of b/c jets : consider multi-jet trigger Data for instance

- Then apply the tagger by only using tracks (or secondary vertices) with negative significance (tracks with $IP/\sigma < 0$ or SV with $\Delta L/\sigma < 0$)

- The qcd MC is used to apply a correction SF_l :

  $\varepsilon_{tag}^{Data}(udsg) = \varepsilon_{neg}^{Data}(total) \times SF_l$

  with $SF_l = \varepsilon_{tag}^{MCqcd}(udsg) / \varepsilon_{neg}^{MCqcd}(total = udsg+c+b)$

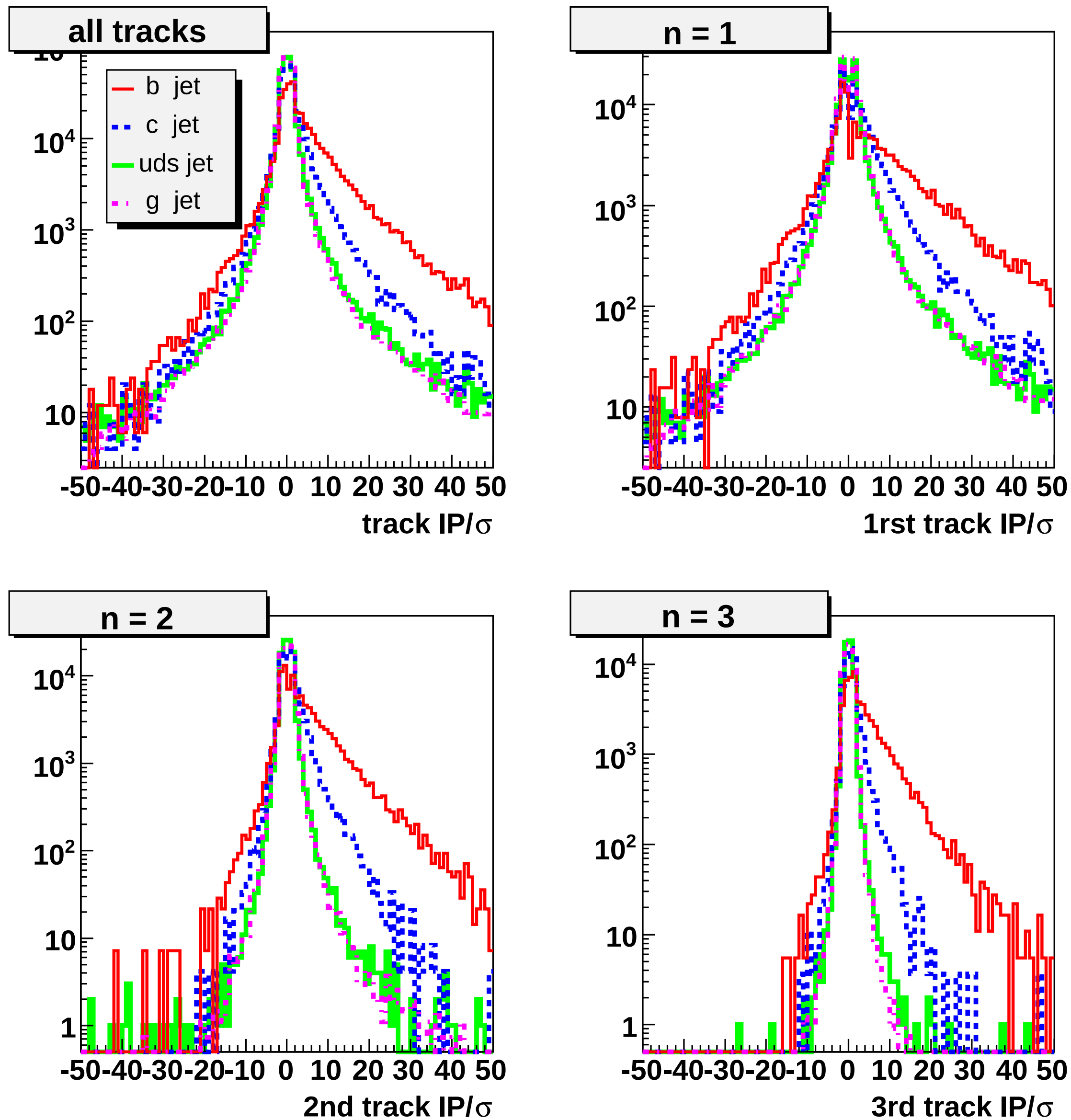
- Systematics are of various sources :
  - observed difference in negative tag rates between different real Data samples : ~ ±2-3% (at least in DØ) but neglected here
  - fraction of remaining b/c jets after negative tags in the qcd MC : see below, and vary these fractions by ±20%
  - difference between uds and g-jets ? vary gluon fraction by ±10%

IP significance

- In the present BReco, negative significances are only available with the track counting method

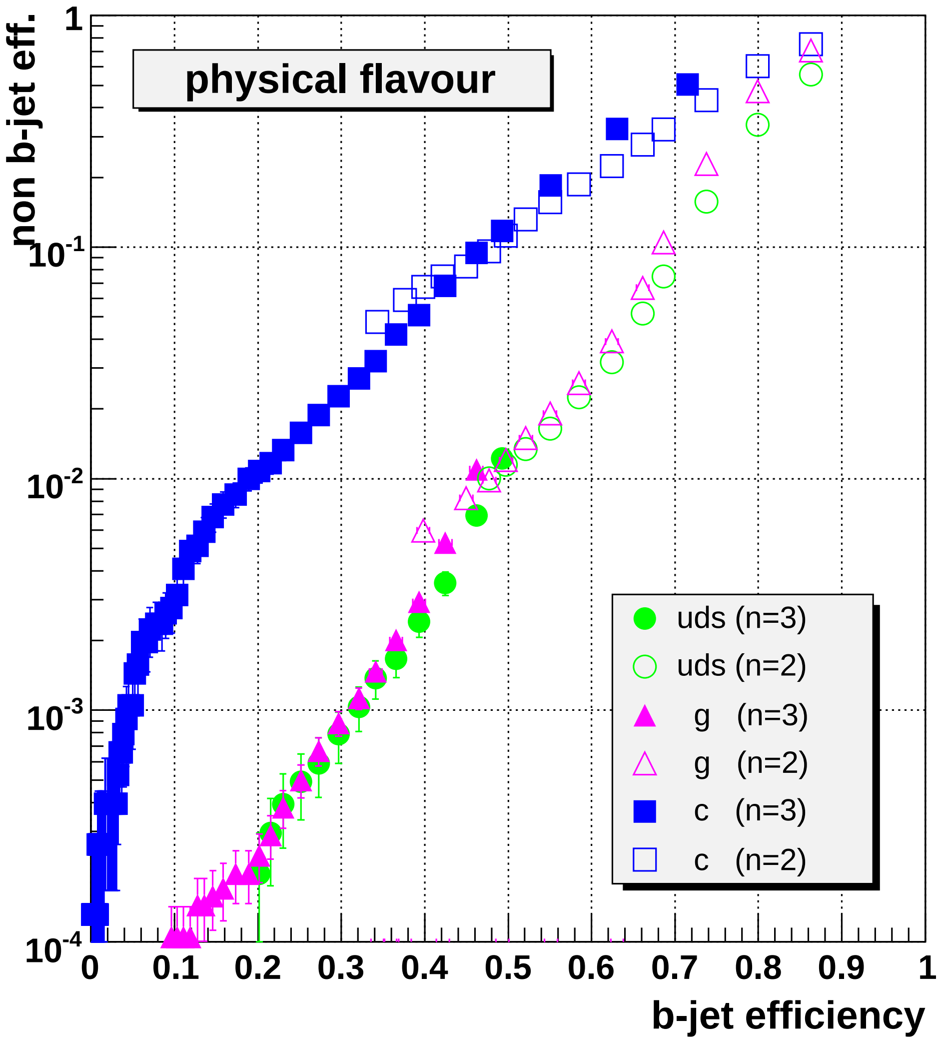
→ require 2 or 3 tracks with IP/σ > cut value

- Note that IP/σ is the same for uds and g jets (g→bb or cc are assigned as b or c-jet)



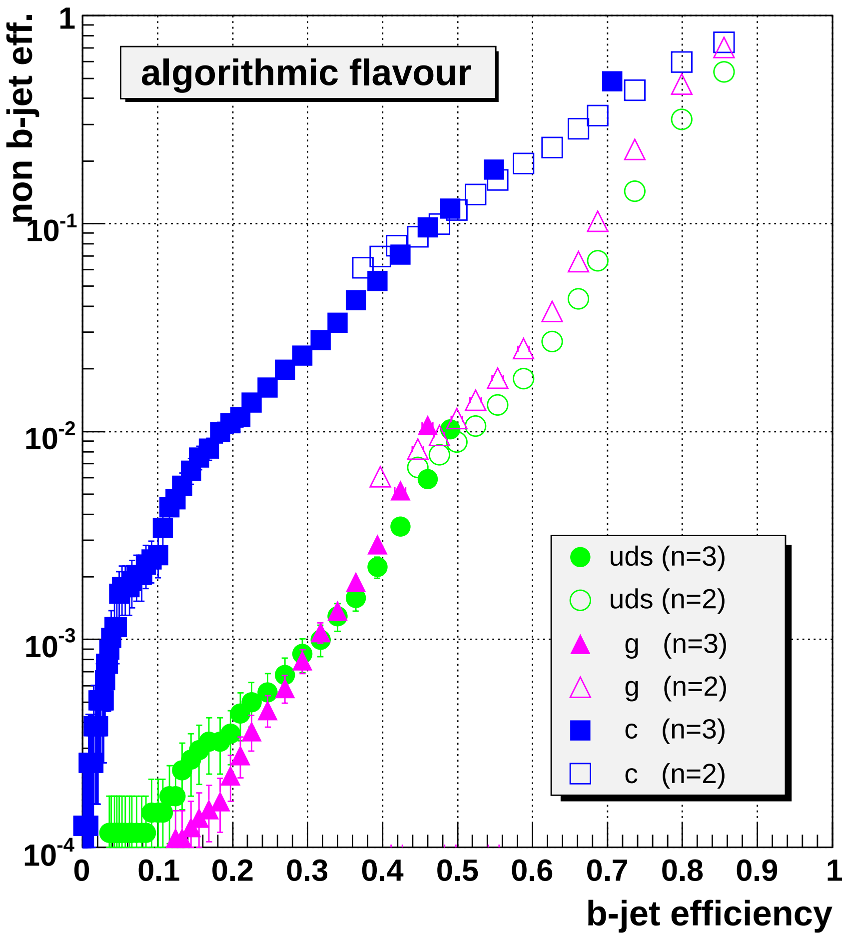
Track Counting performance

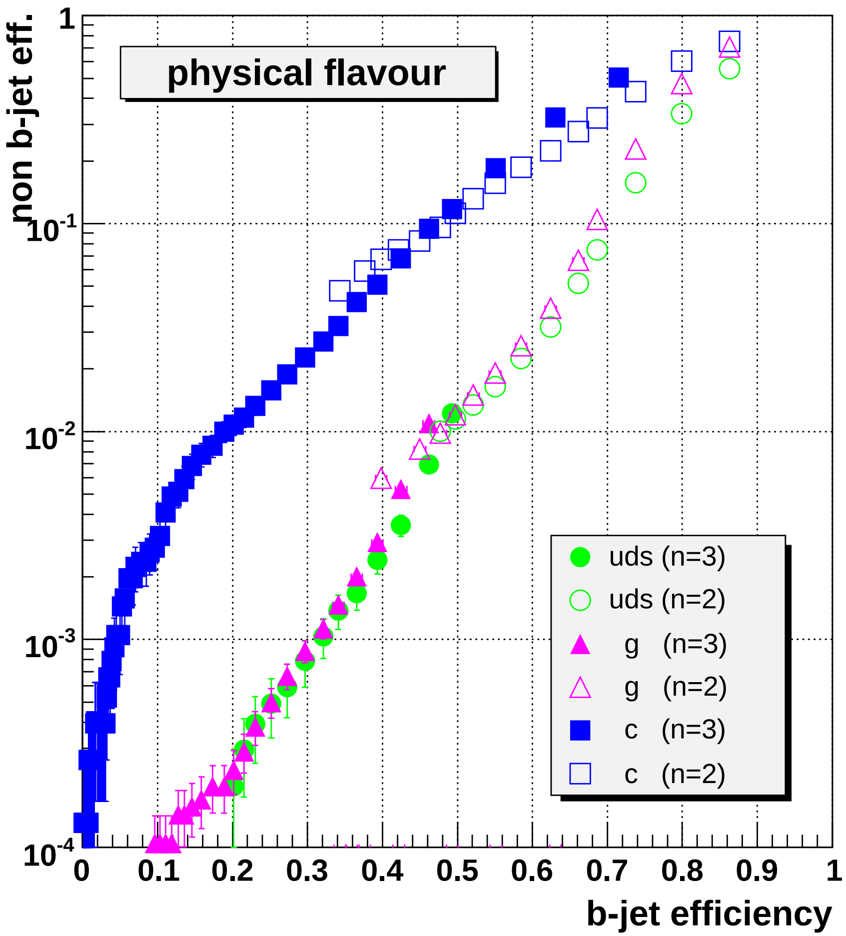
- Use qcd MC with here 50<pt<120 GeV , |η|<1.4
- One can require ≥2 or ≥3 tracks with IP/σ > cut
- Get similar performance than in Note 2006/019 (Fig.8)
- (here only, the efficiency is computed w.r.t all jets, not only the taggable ones)



Performance (cont.)

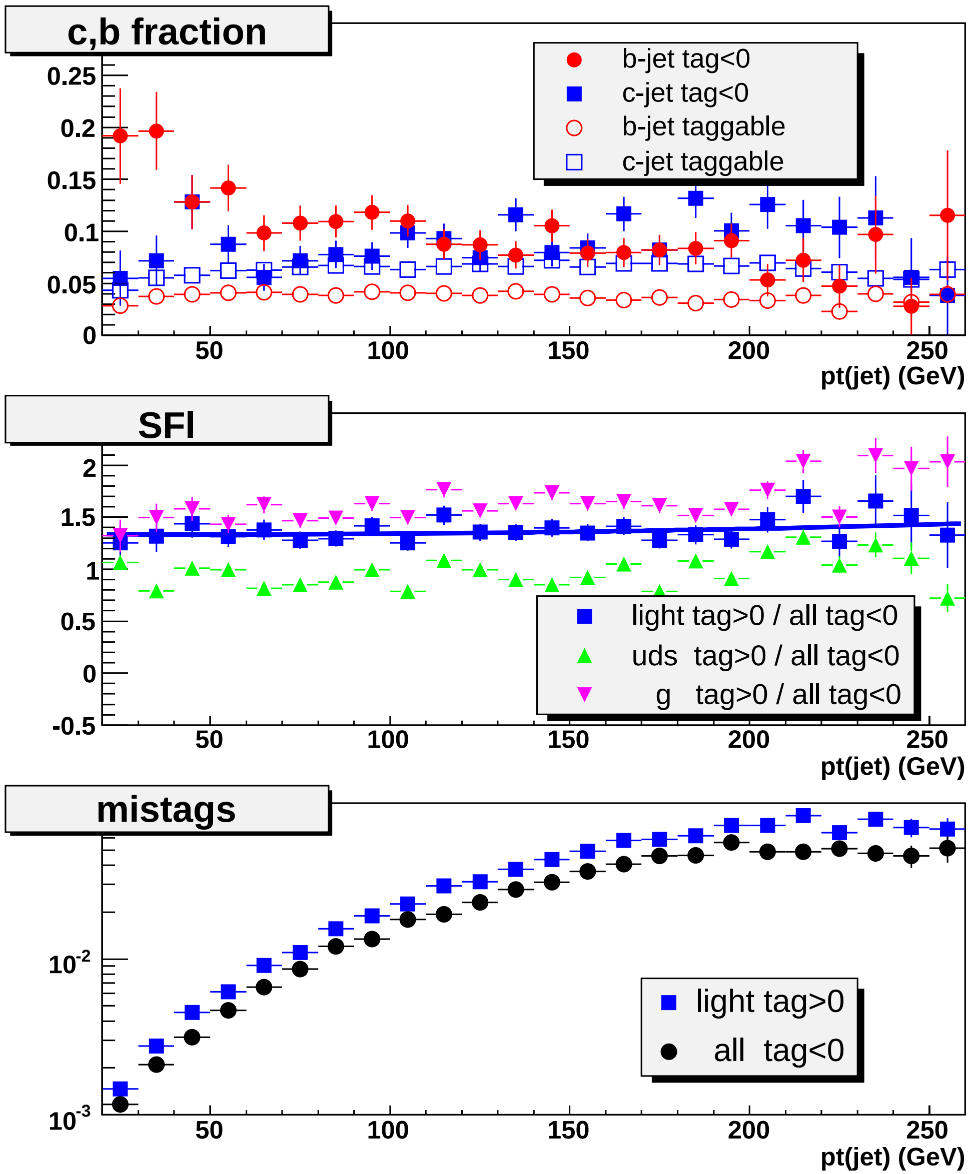
- In fact the b-tag performance doesn't seem to depend too much on the flavour definition





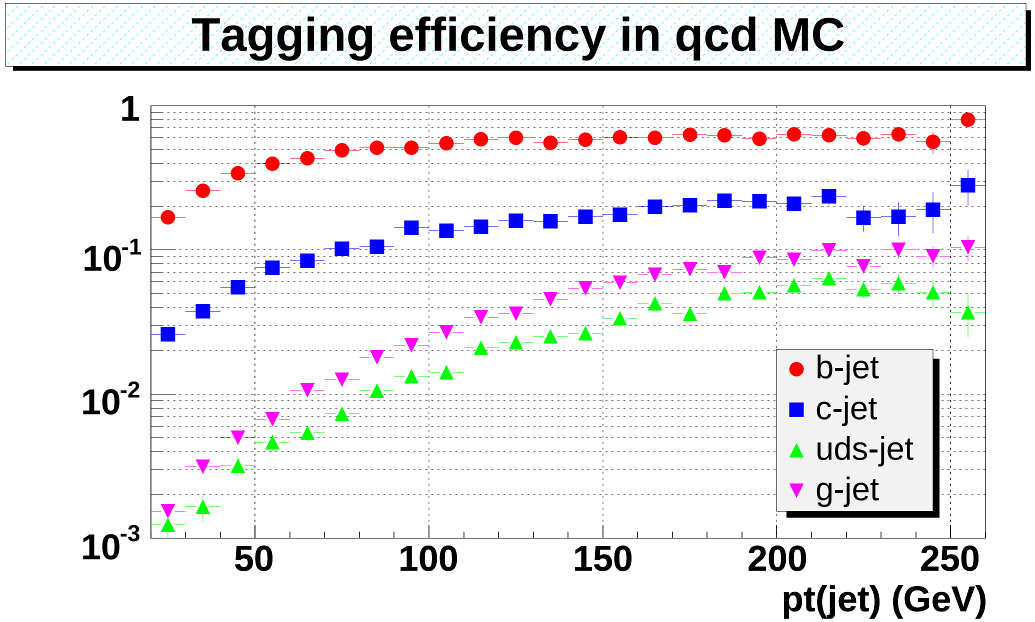
Mistags (cont.)

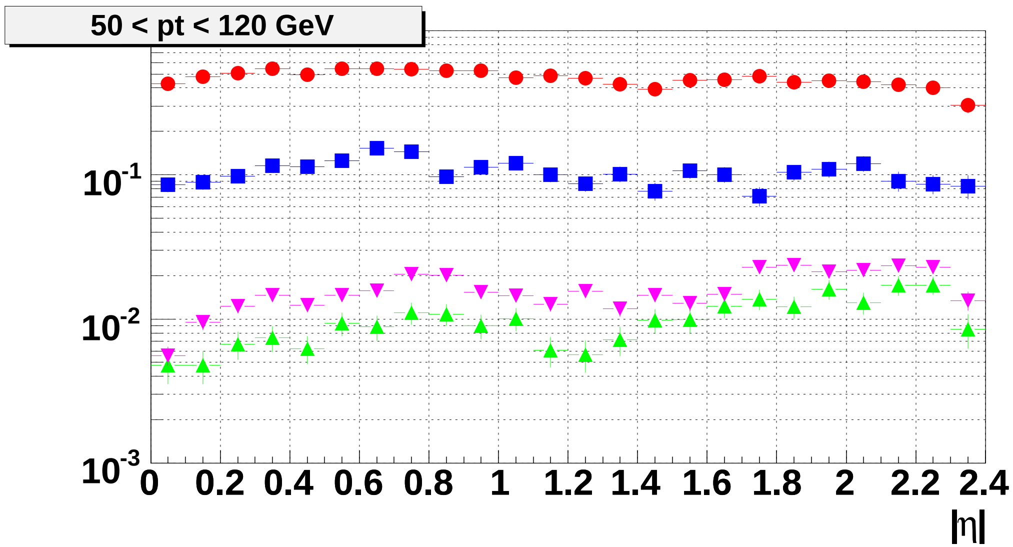
- Choose a working point : n=2 with IP/σ>2.5 for instance
- qcd MC : fraction of c and b jets before and after negative tagging
- Scale Factor SF_l (l = udsg) : close to 1.3
- different in uds-jet and g-jet
- udsg positive tag efficiency and overall negative tag rate



Tagging efficiency

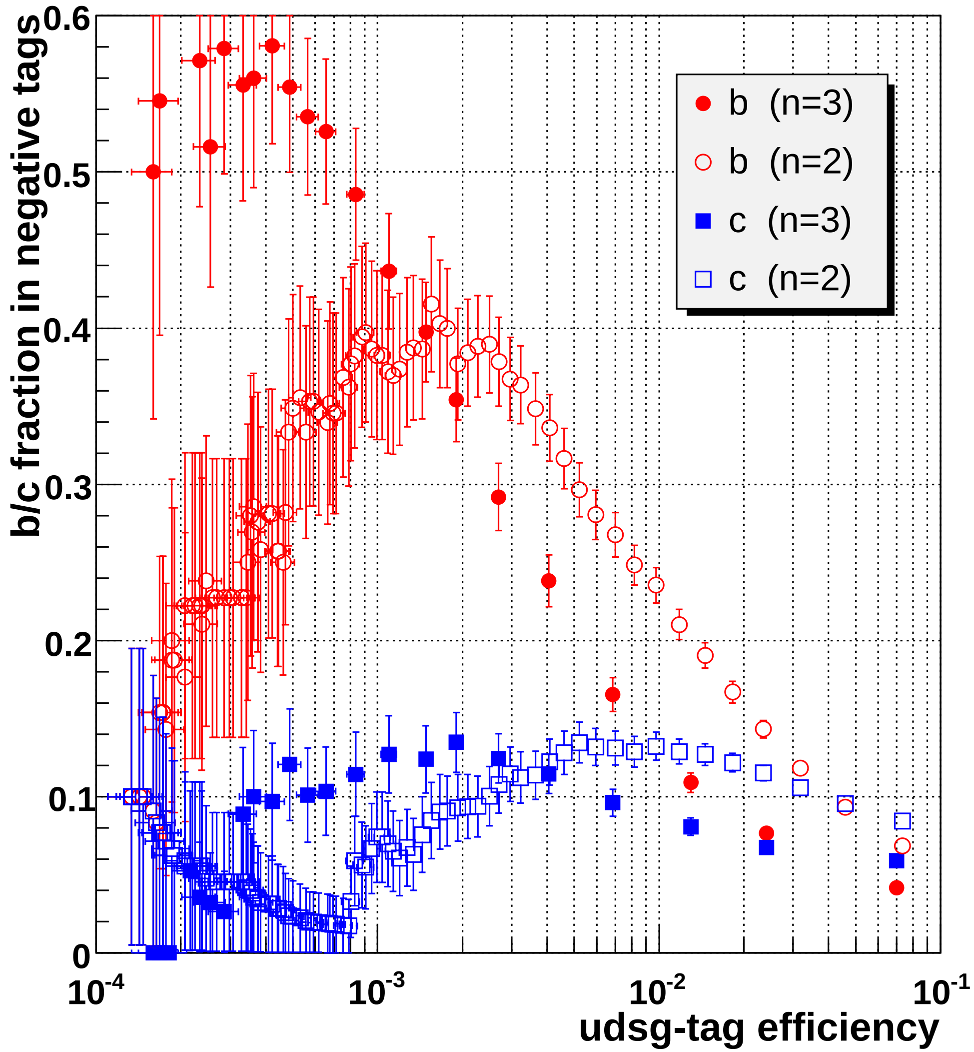
- For this working point
  ≥ 2 tracks with $IP/\sigma > 2.5$
- g-jet tagging efficiency slightly larger than for uds-jets (due to the larger track multiplicity I guess)





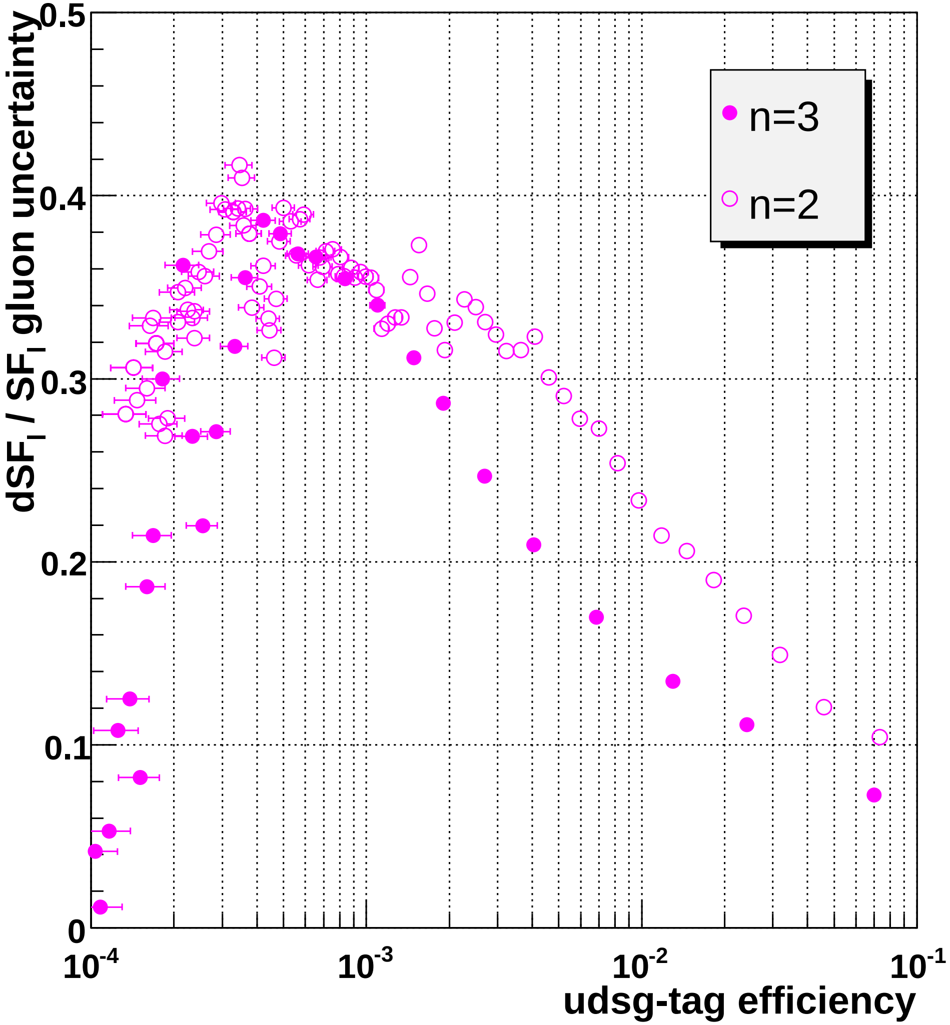
Fraction of b/c jets

- In qcd MC, here if 50<pt<120 GeV
- displayed versus the light=udsg tag efficiency (for any working point) :
  one has to use (see slide 10)
  n=2 if $\varepsilon^{l}_{tag} > 10^{-2}$
  n=3 if $\varepsilon^{l}_{tag} < 10^{-2}$
- for n=3, $\varepsilon^{l}_{tag} \approx 10^{-2}$:
  ~12% b and ~8% c in negative tag jets



Gluon fraction dependence

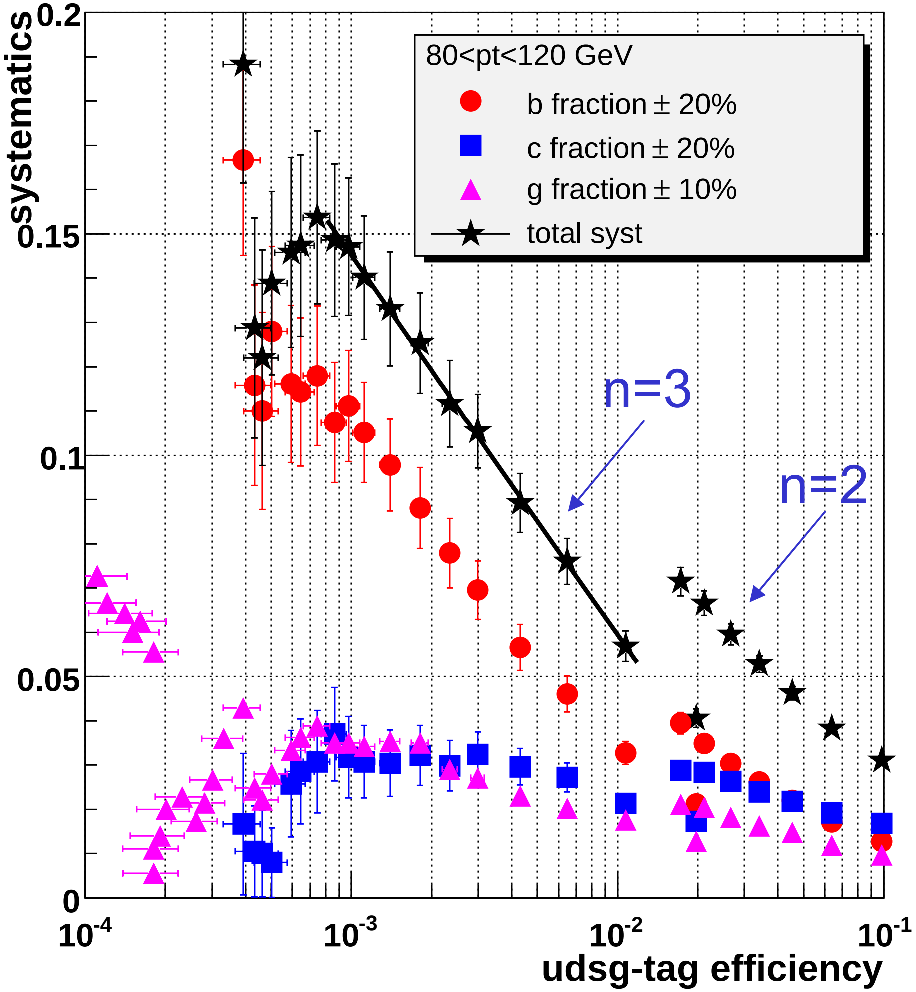
- The gluon fraction also affects $SF_l = \varepsilon_{tag}(udsg) / \varepsilon_{neg}(udsg+c+b)$
- Here $[dSF_l / SF_l] / [df_g / f_g]$ is shown , where df_g / f_g is the uncertainty on the gluon fraction,
  i.e it has to be scaled by 0.1 if $df_g / f_g = \pm 10\%$



Mistag systematics

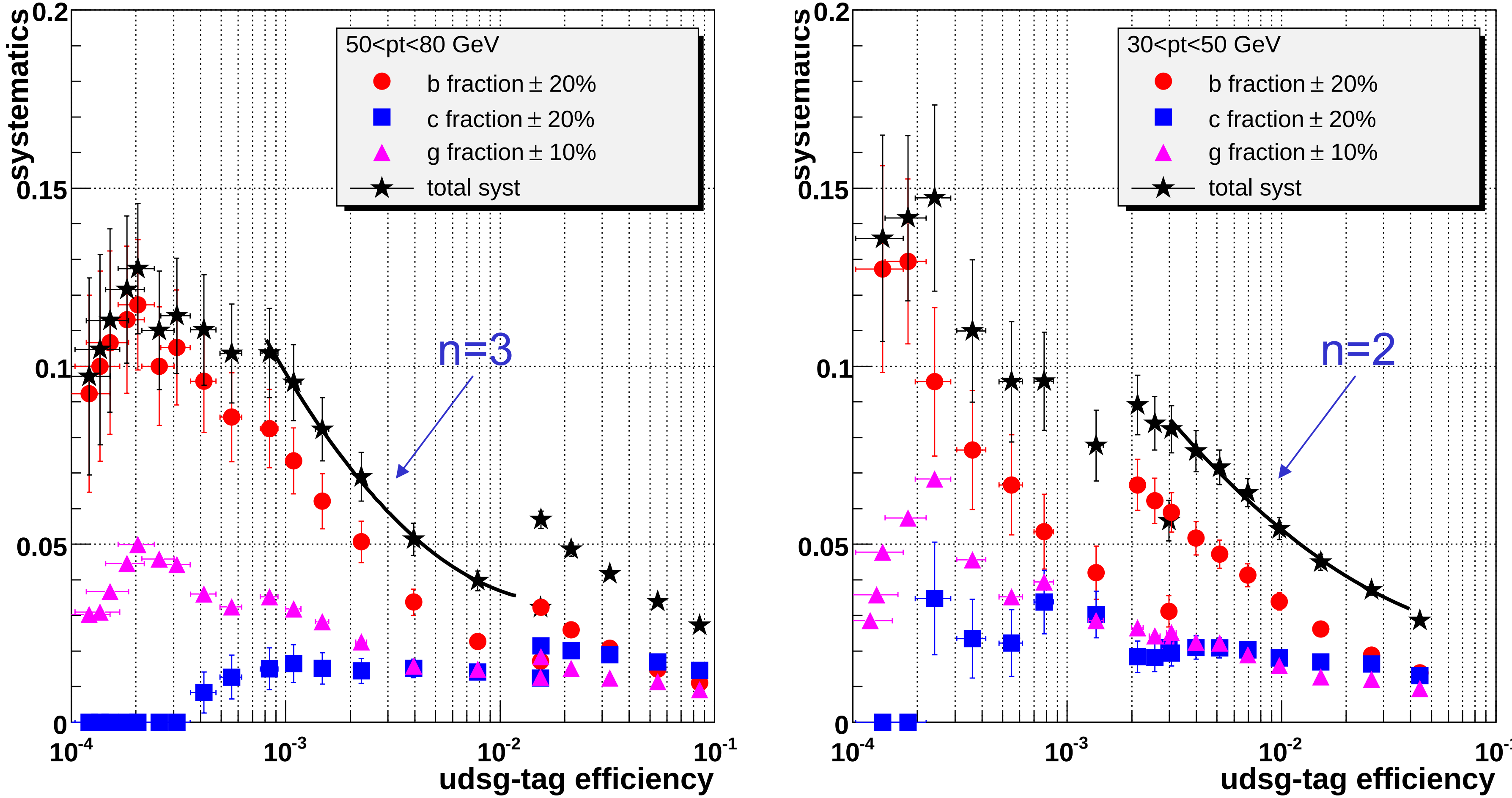
- The uncertainties on the b and c fractions are assumed to be ±20%, and fully correlated
- The uncertainty on the gluon fraction is assumed to be ±10% (and not correlated with the b and c fractions)
- With the track counting method and n=3, one gets an overall systematic of about

  $\pm\ 6\%$ if $\varepsilon^{l}_{tag} = 10^{-2}$
  $\pm 16\%$ if $\varepsilon^{l}_{tag} = 10^{-3}$
  if 80<pt<120 GeV



Mistag syst. (cont.)

- Idem in lower pt bins : systematics are somewhat smaller



Conclusion and outlook

- Based on the track counting method and qcd MC, the mistag systematics have been evaluated :

for jets within	30<pt<50 GeV	50<pt<80 GeV	80<pt<120 GeV
if $\varepsilon^{l}_{tag} = 10^{-2}$	± 5%	± 4%	± 6%
$\varepsilon^{l}_{tag} = 10^{-3}$	±10%	±11%	±16%

- Extending this mistag estimate to all lifetime taggers would require to run the track probability and the combined-tag methods with negative IP or with negative decay length only…
  (and for consistency with positive ones too !) :
  i.e 2 possibilities per algorithm :
  negative significance only or positive significance only
  which have to be implemented in BReco