performance of track and vertex reconstruction and b-tagging studies with CMS

Alexander Schmidt

Universität Zürich

Physik-Institut

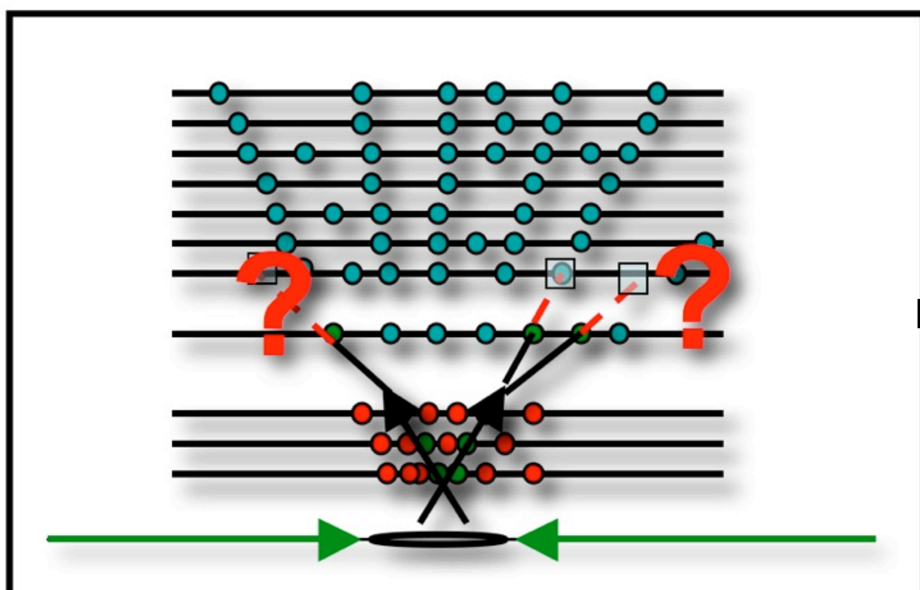
on behalf of the CMS collaboration

track reconstruction

- details on CMS silicon tracker given by Ettore earlier today
- this talk: tracker performance:
  - track reconstruction
  - vertex reconstruction
  - b-jet tagging

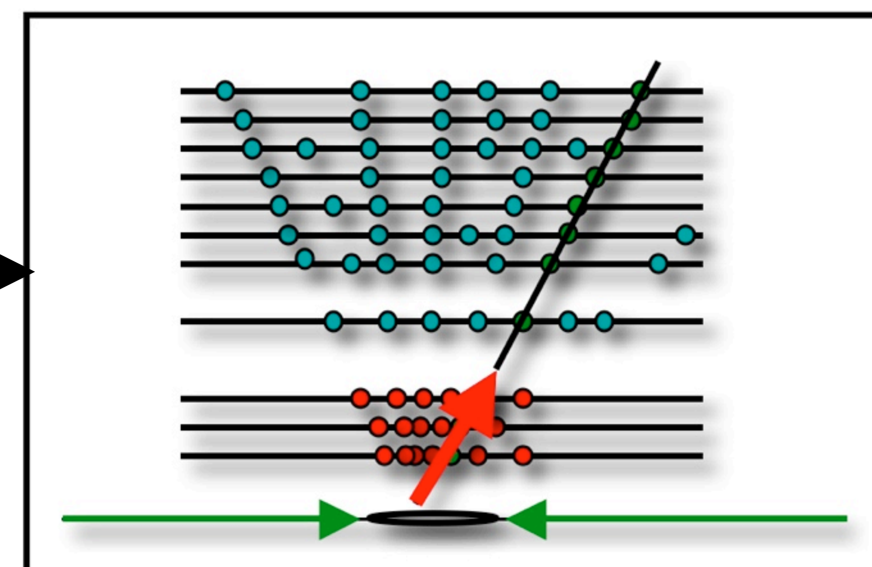
- basics of track reconstruction in CMS:

seeding from pixel hits



iterative tracking
(details next slide)

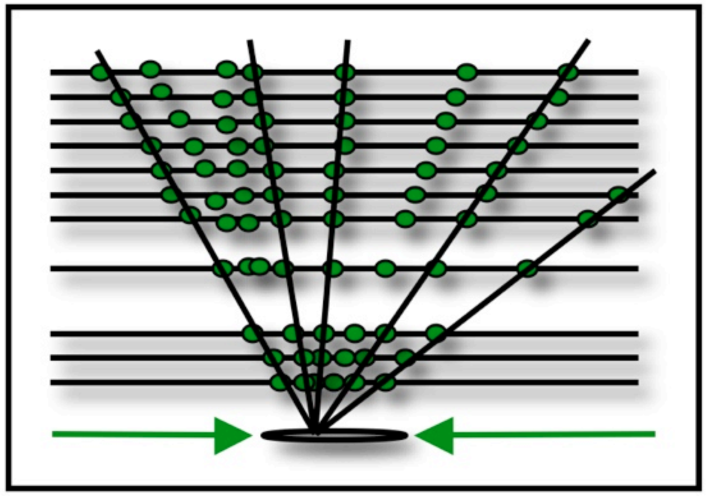
final fit using Kalman Filter/Smoother

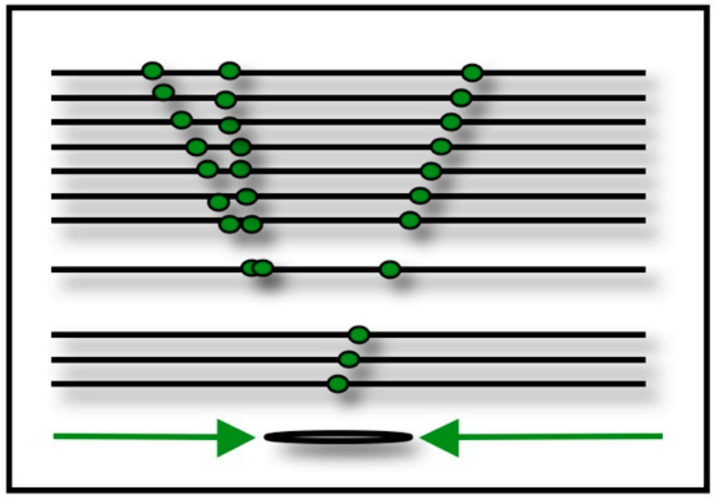


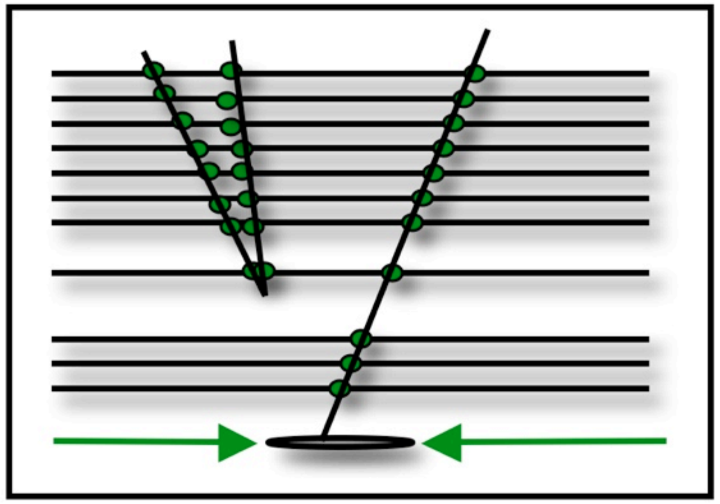
iterative tracking

six iterations:

- propagate seed outwards and search for new hits
- unambiguously assigned hits are removed from the list
- filter track collection to remove fakes or bad tracks
- repeat with remaining hits





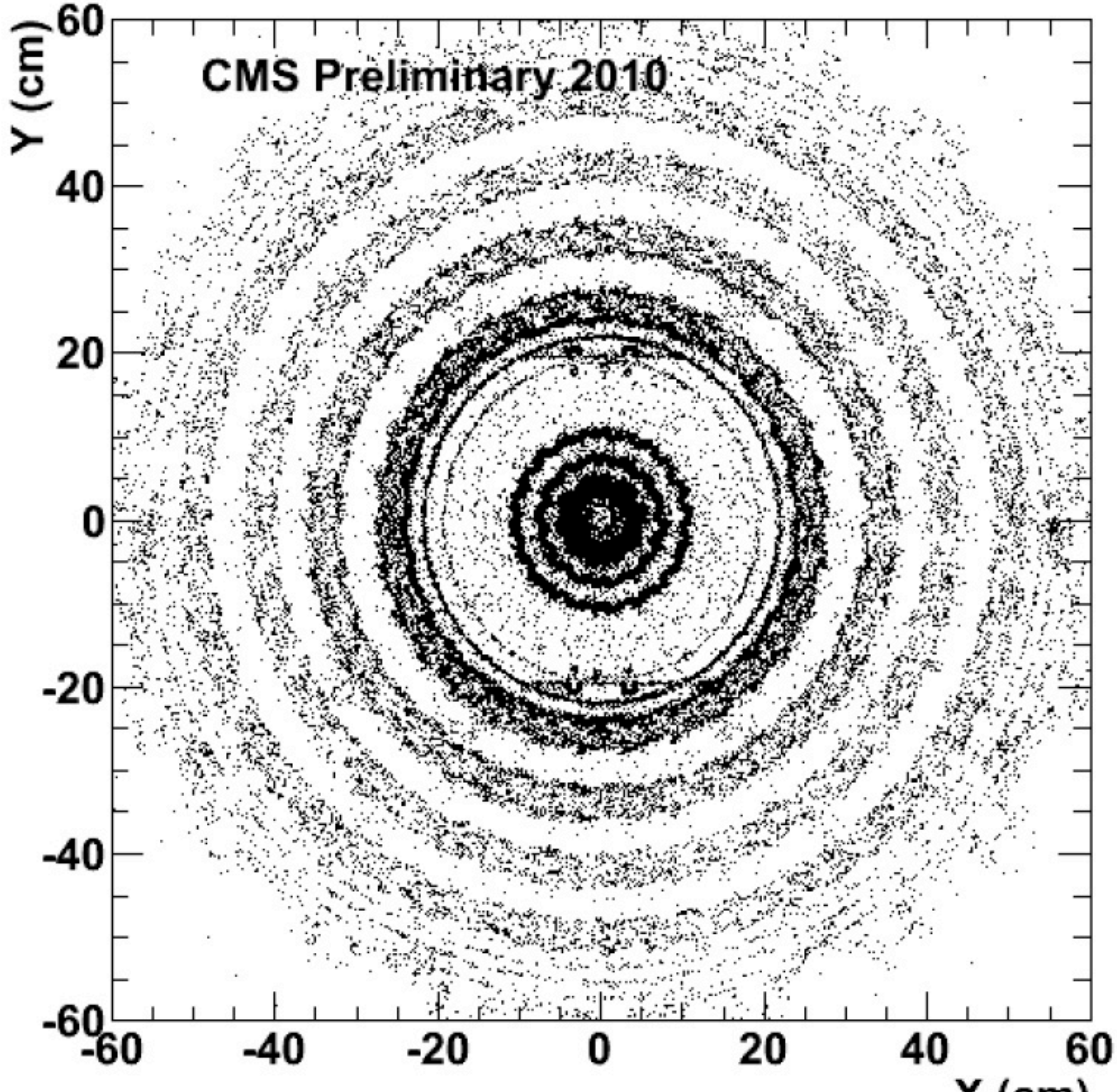


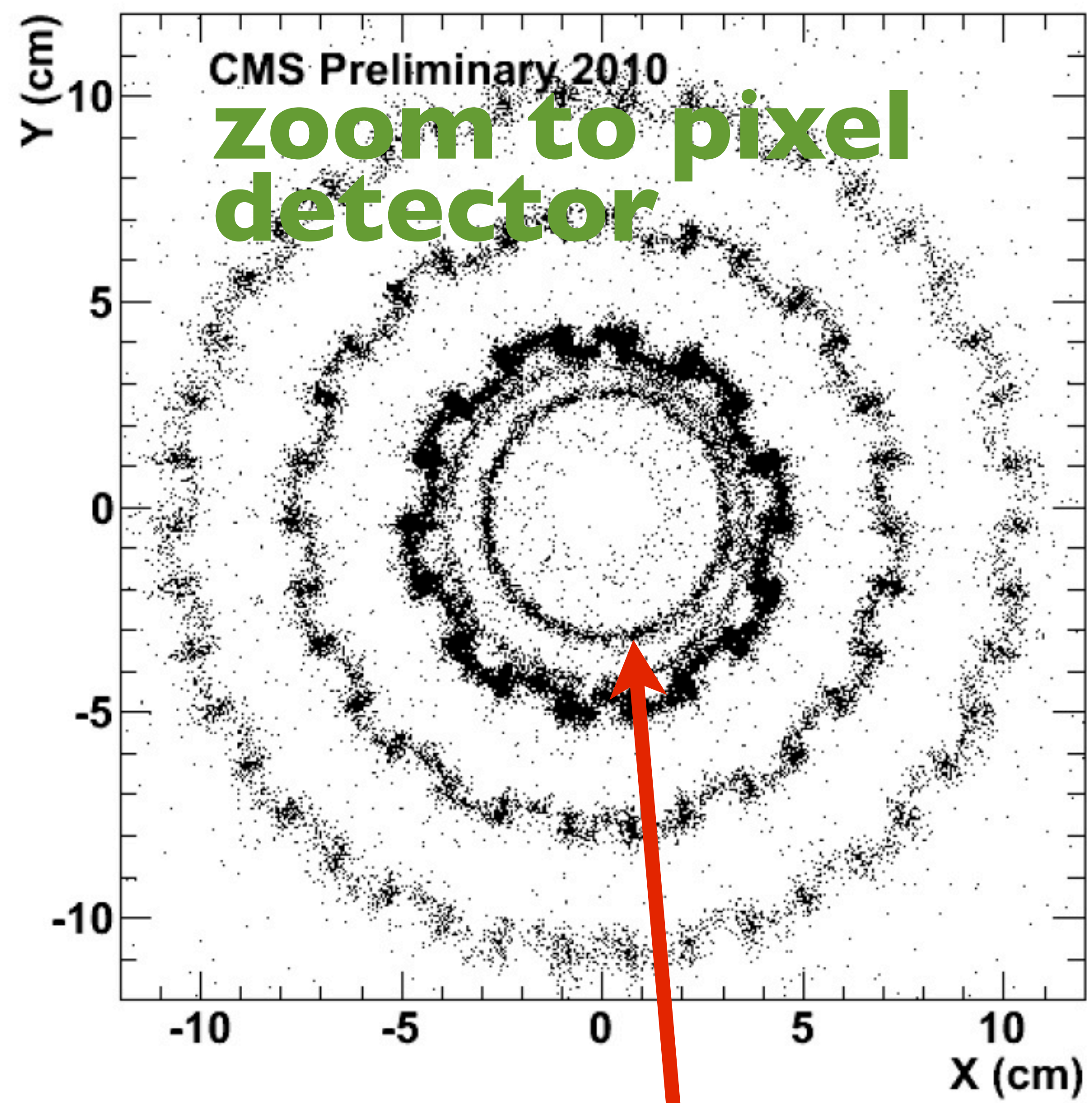
differences in seeding:

- first two iterations: pixel pairs or pixel triplets, p_t>0.9GeV
- third iteration: pixel triplets, low momentum tracks
- fourth iteration: pixel + strip layers as seeds (find displaced tracks)
- fifth, sixth iterations: strip pairs (for tracks lacking pixel hits)

photon conversions

radiography from reconstructed $\gamma \rightarrow e^+e^-$ vertices

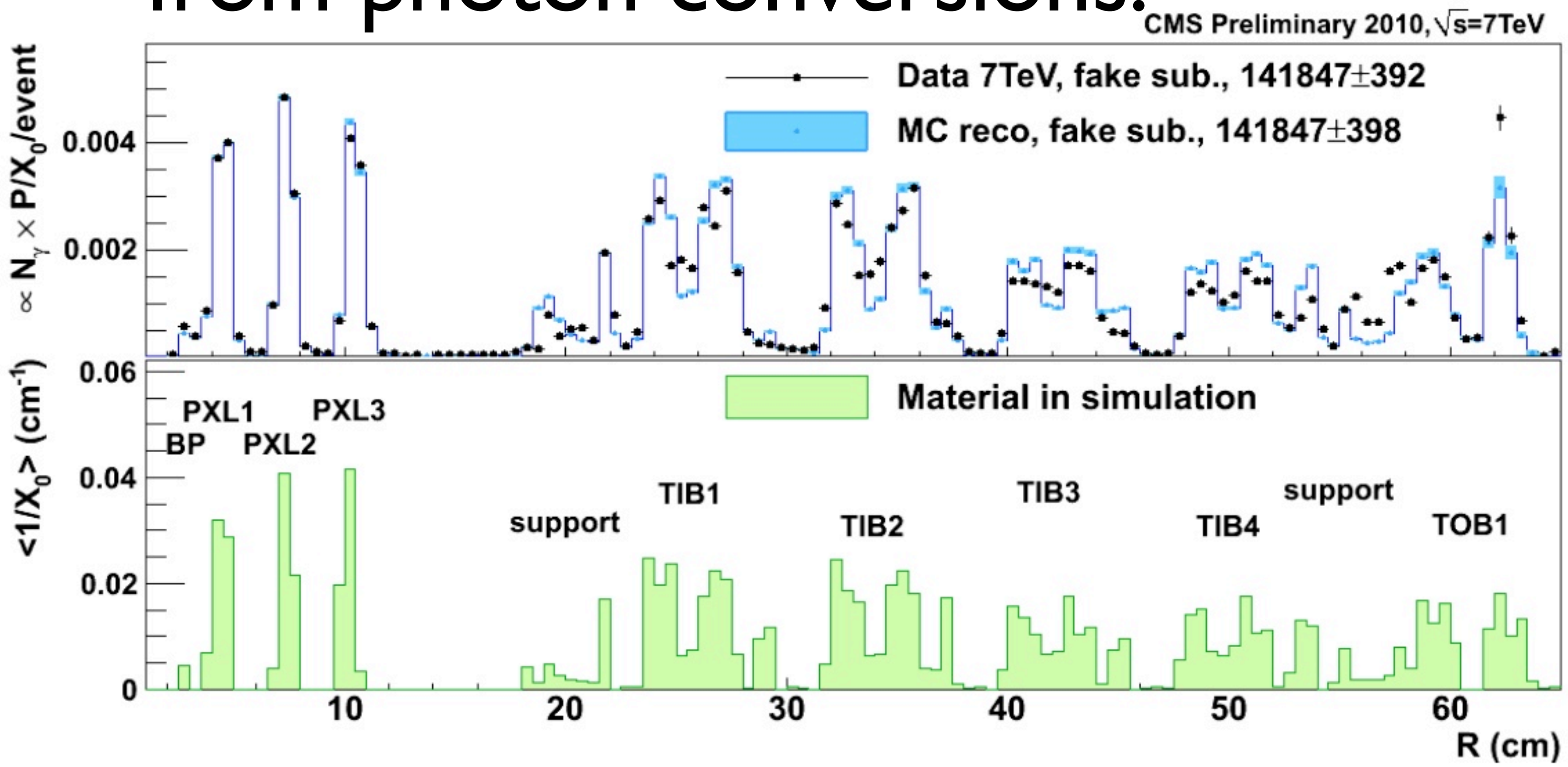




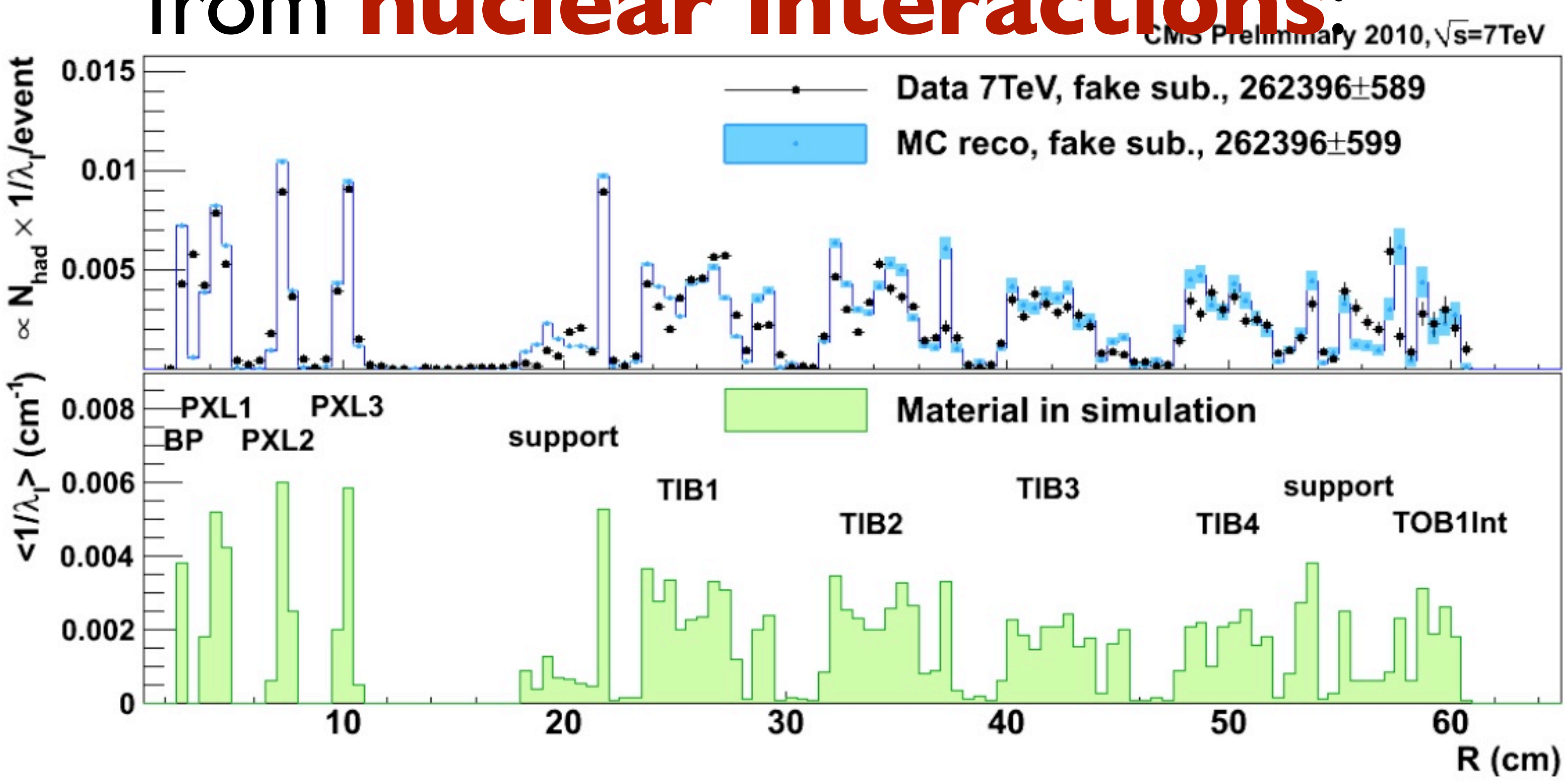
pixel barrel displaced wrt. beam pipe

CMS PAS TRK-10-003

from photon conversions:



from **nuclear interactions**:



plots corrected for:

- reco. efficiency
- photon (or hadron) flux
- fake subtraction

shown in fiducial volumes:

- R < 20cm : |z| < 26cm
- 20 < R < 65cm: |z| < 66cm

includes Pixel Barrel and Tracker Inner Barrel

agreement in data and simulation within **10%**

CMS PAS TRK-10-003

tracking efficiency

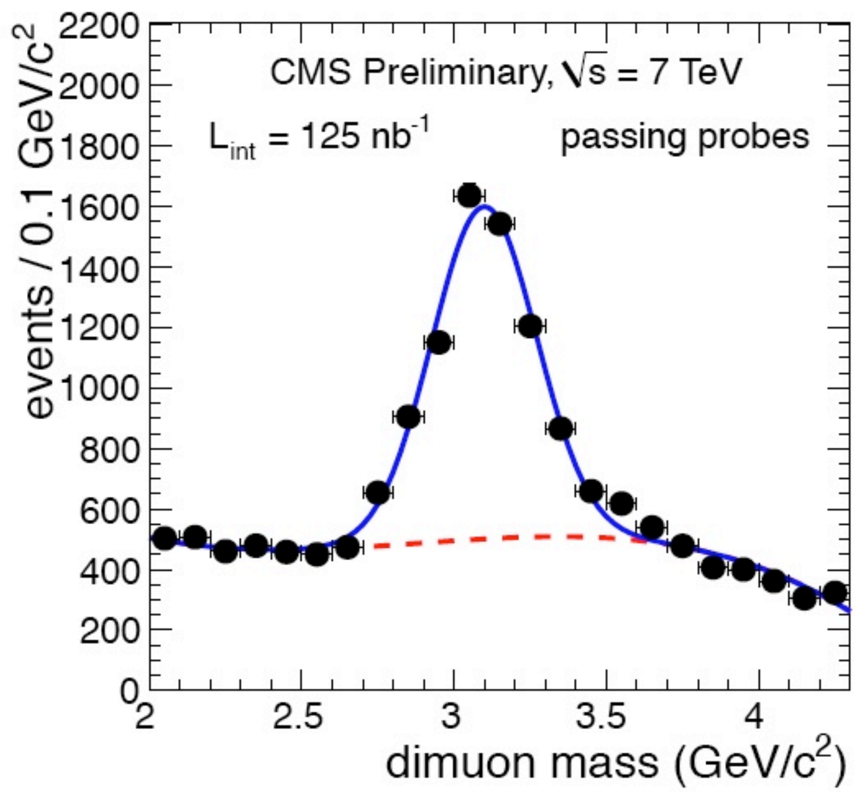
data driven eff. measurement for muons:

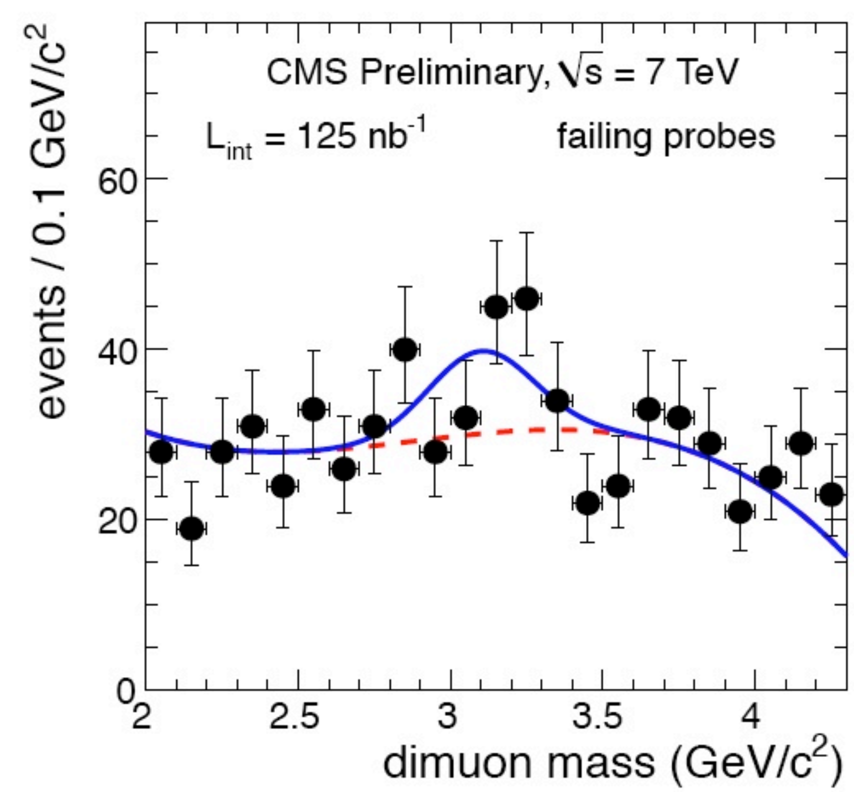
- using $J/\psi \rightarrow \mu\mu$ events
- **tag** one muon
- the second one is the **probe**
- probe reconstructed from muon system
- efficiency ϵ to match the probe to a track

$$\epsilon_T \epsilon_M = \frac{\epsilon - \epsilon_F}{1 - \epsilon_F}.$$

fake

CMS PAS TRK-10-002





Region	Data Eff. (%)	Sim Eff. (%)	Data/Sim
$0.0 \leq \lvert\eta\rvert < 1.1$	$100.0^{+0.0}_{-0.3}$	$100.0^{+0.0}_{-0.1}$	$1.000^{+0.001}_{-0.003}$
$1.1 \leq \lvert\eta\rvert < 1.6$	$99.2^{+0.8}_{-1.0}$	$99.8^{+0.1}_{-0.1}$	$0.994^{+0.009}_{-0.010}$
$1.6 \leq \lvert\eta\rvert < 2.1$	$97.6^{+0.9}_{-1.0}$	$99.3^{+0.1}_{-0.1}$	$0.983^{+0.009}_{-0.010}$
$2.1 \leq \lvert\eta\rvert < 2.4$	$98.5^{+1.5}_{-1.6}$	$97.6^{+0.2}_{-0.2}$	$1.010^{+0.015}_{-0.016}$
Combined	$98.8^{+0.5}_{-0.5}$	$99.2^{+0.1}_{-0.1}$	$0.996^{+0.005}_{-0.005}$

eff. measurement for pions: using $D^0 \rightarrow K\pi$ decays (backup slides)

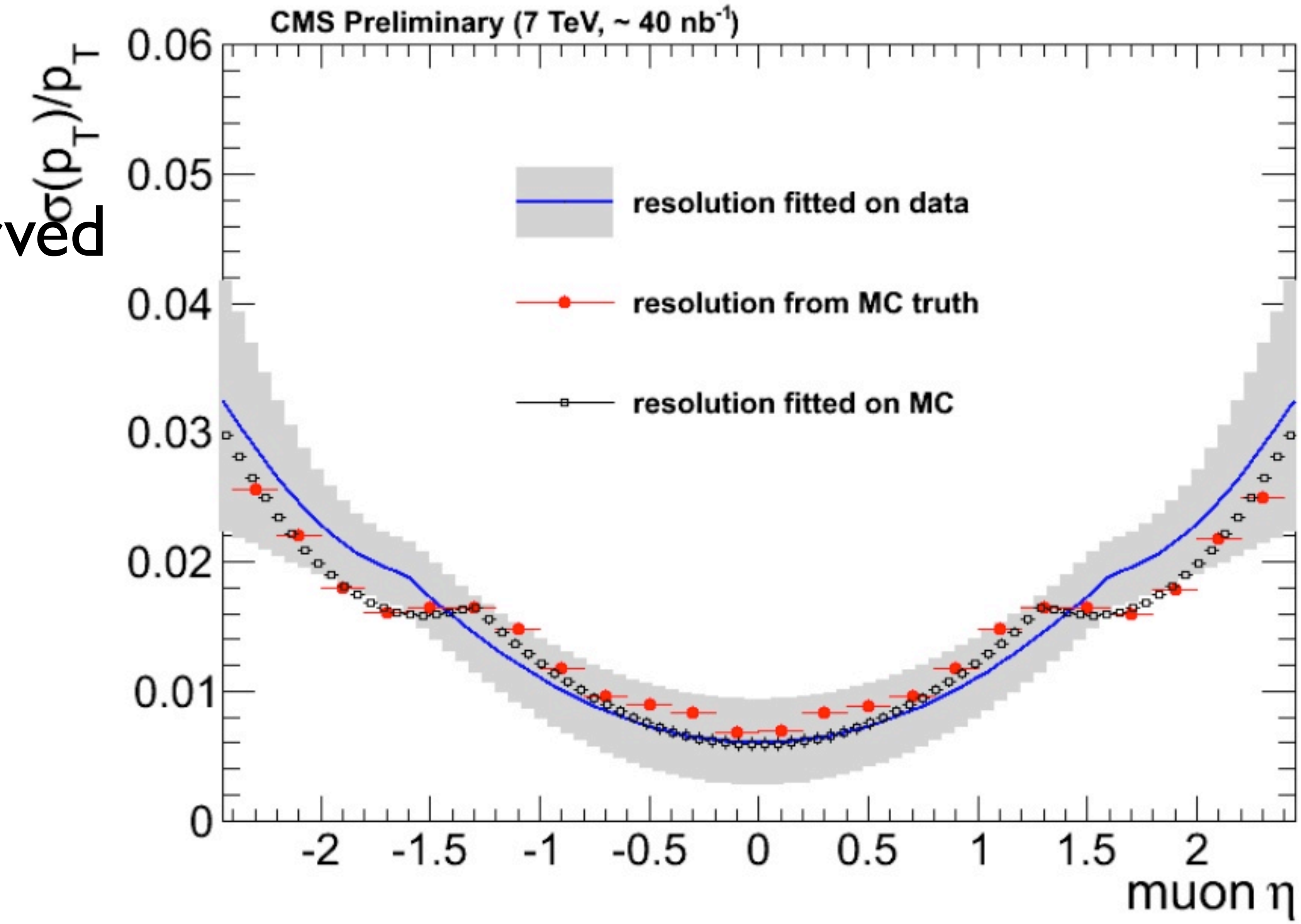
momentum scale and resolution

CMS PAS TRK-10-004

using J/ψ mass lineshape:

express functional form, depending on two-track kinematics, gaussian resolution and momentum scale

- intrinsic width 90keV
- unbinned maximum likelihood fit to observed mass distribution



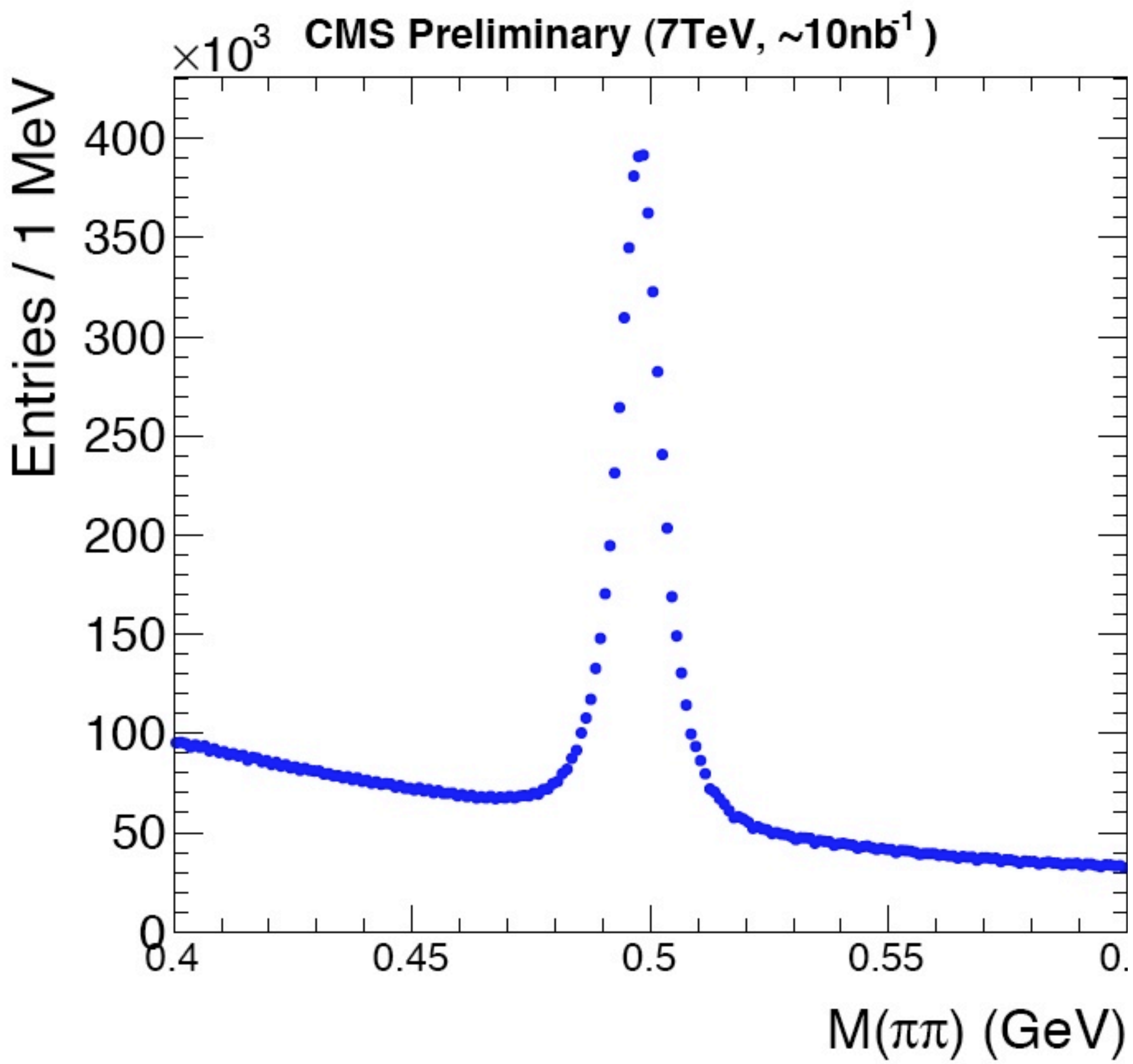
relevant effects:

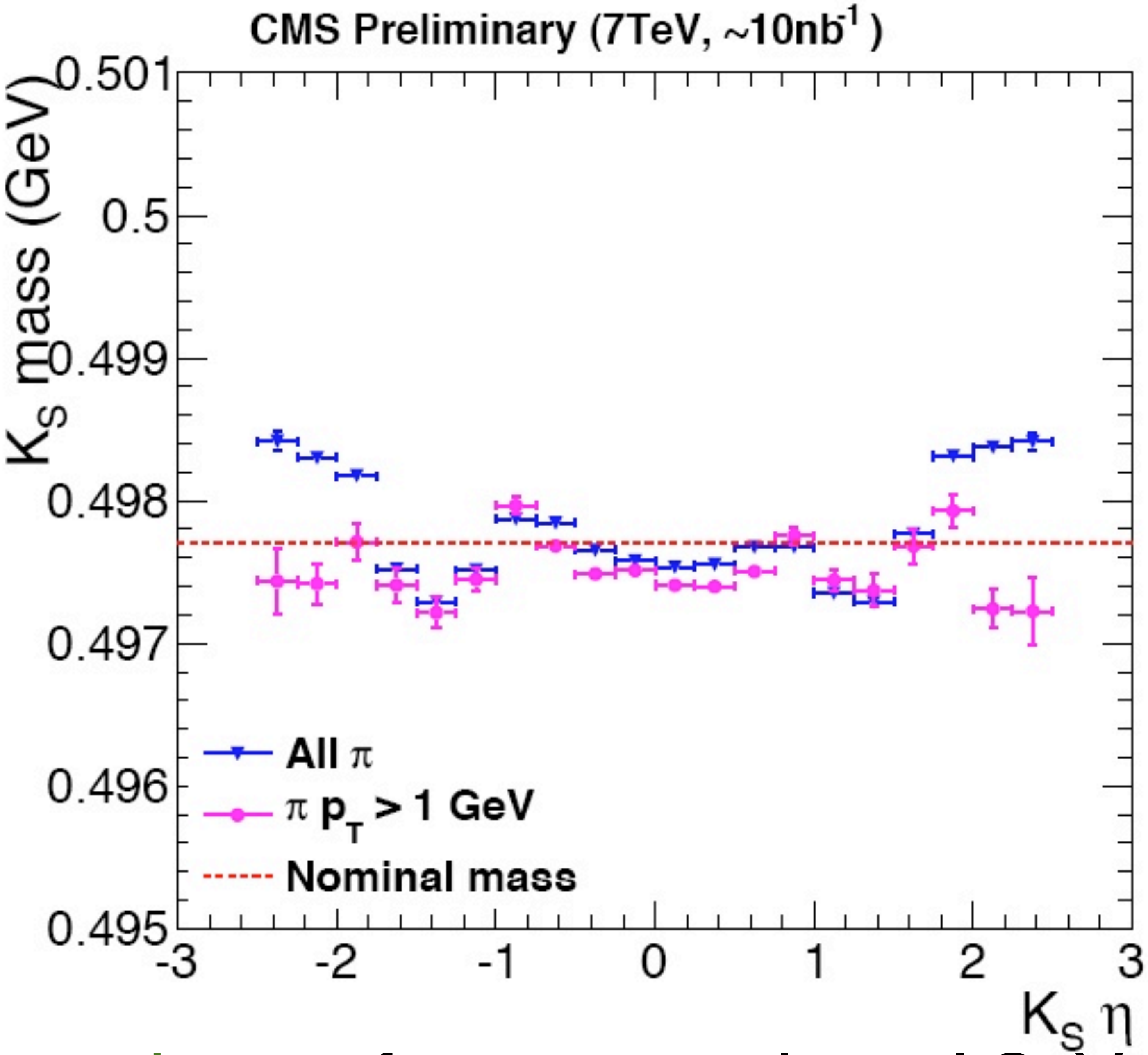
- detector material
- misalignment
- magnetic field

momentum scale and resolution

CMS PAS TRK-10-004

looking at $K_s \rightarrow \pi^+\pi^-$ resonance (low pt tracks)





agreement with nominal mass within 0.3MeV

- better for pions with p_t>1GeV
- more difficult in endcap region

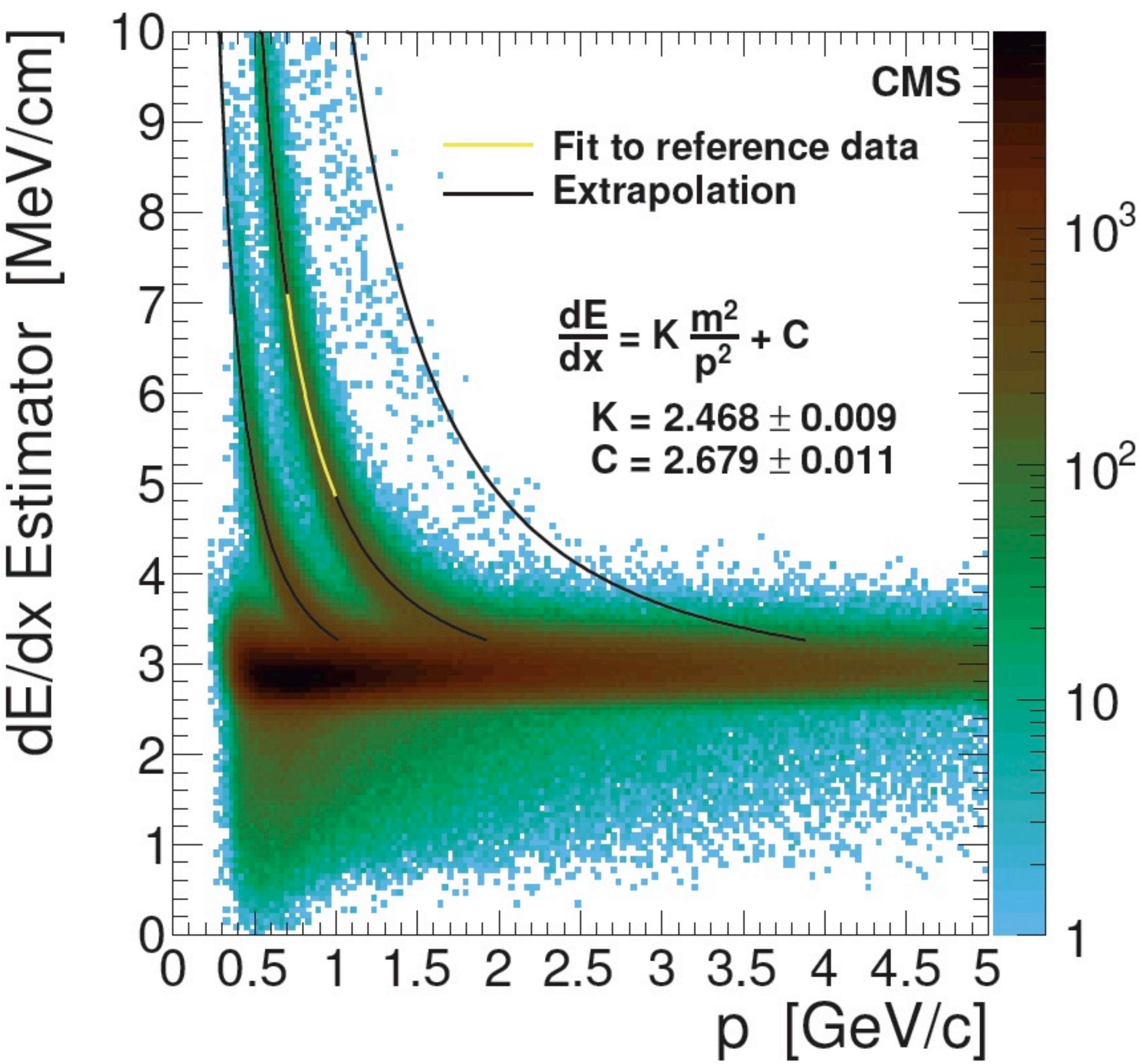
energy loss

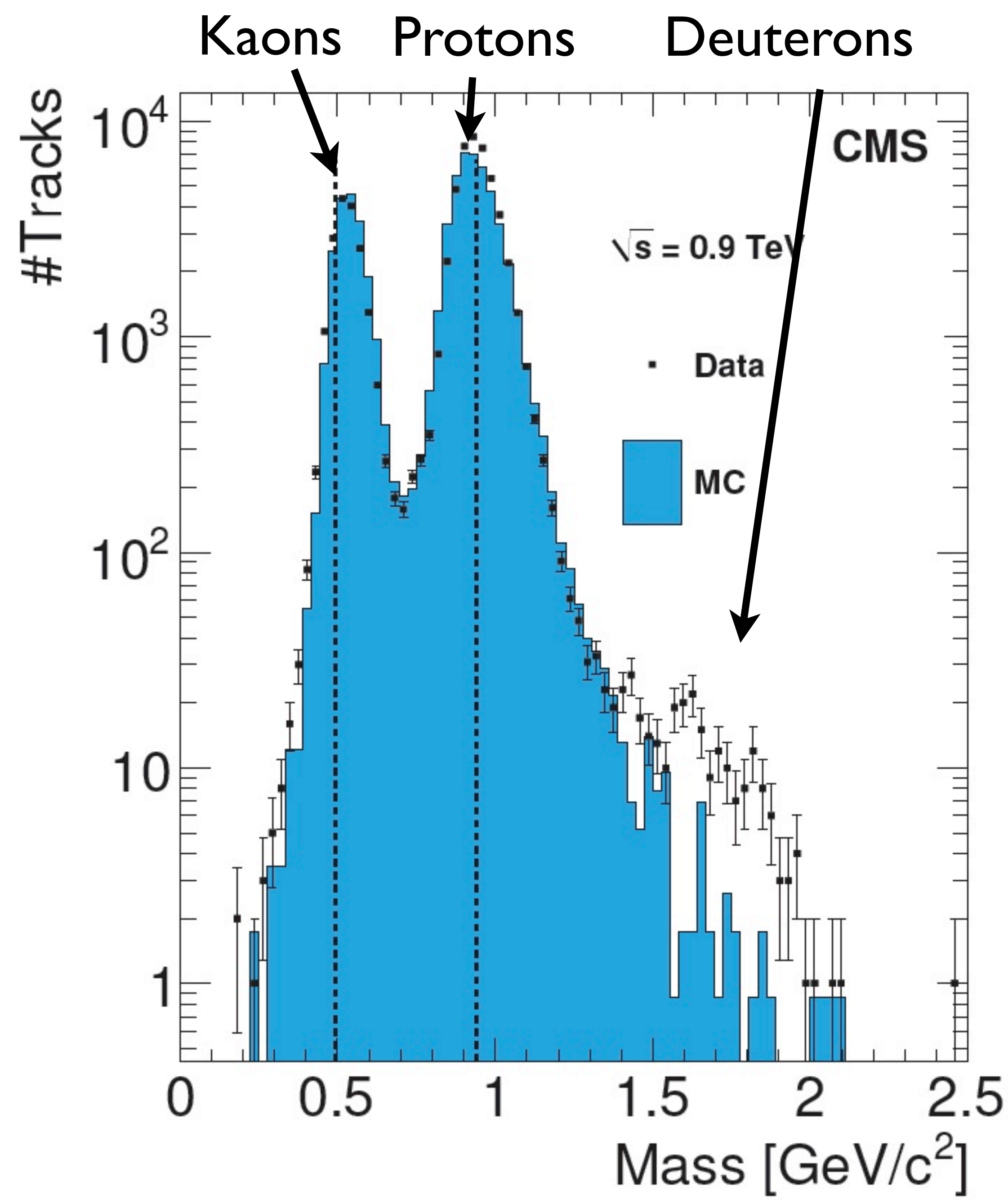
CMS PAS TRK-10-001

hit cluster charge proportional to energy deposit:

→can calculate energy loss **dE/dx** along the trajectory from all hits

→particle identification possible up to 1GeV



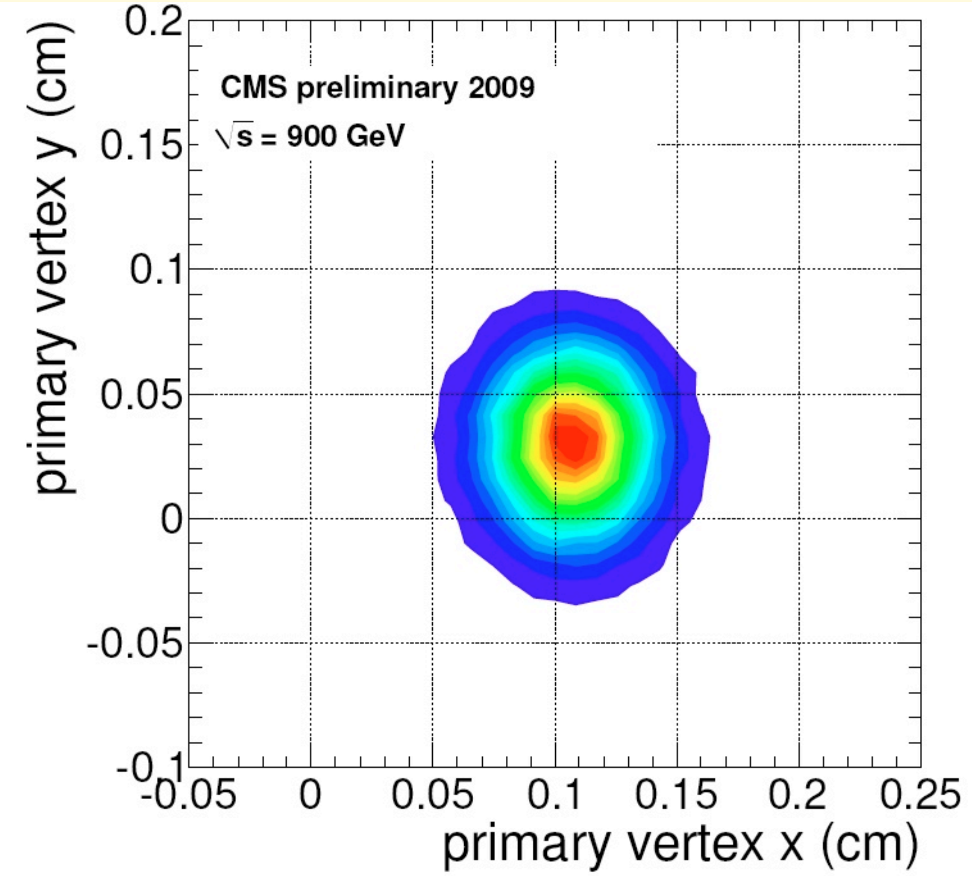


primary vertex

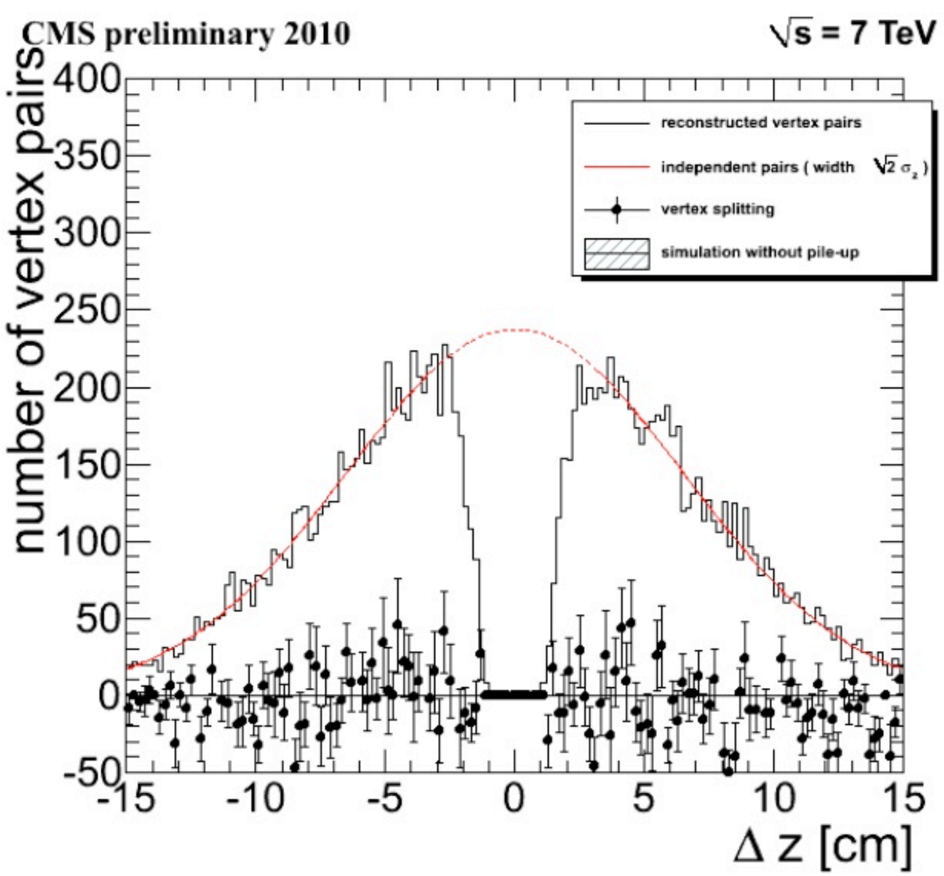
CMS PAS TRK-10-005

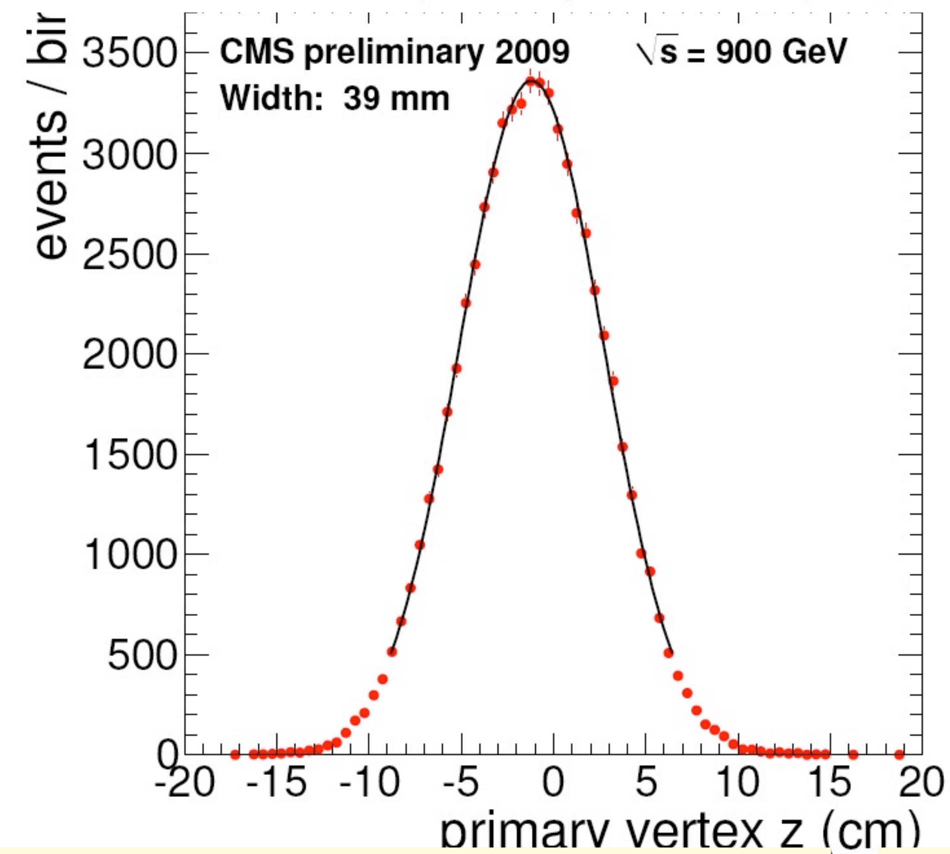
PV reconstruction:

- select prompt tracks (wrt. Beamspot)
- “Adaptive Vertex Fitter”: weight tracks according to compatibility with common vertex (outliers downweighted)
- z-separation>1cm (depends on pileup)



Δz for multi-vertex events:



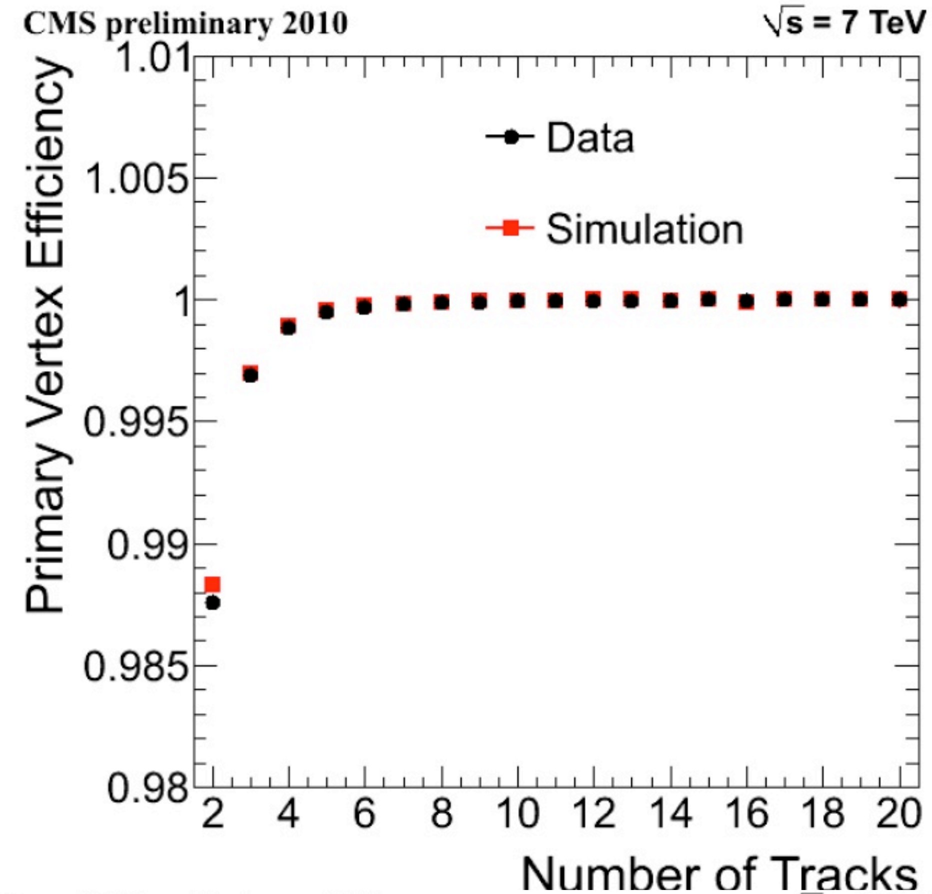


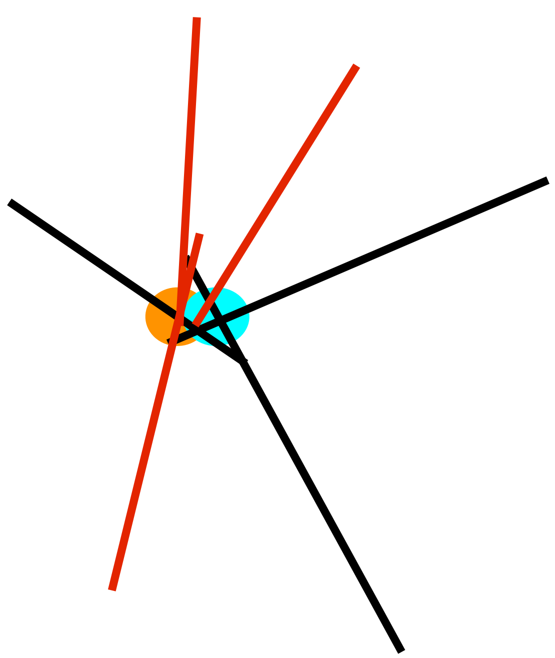
primary vertex

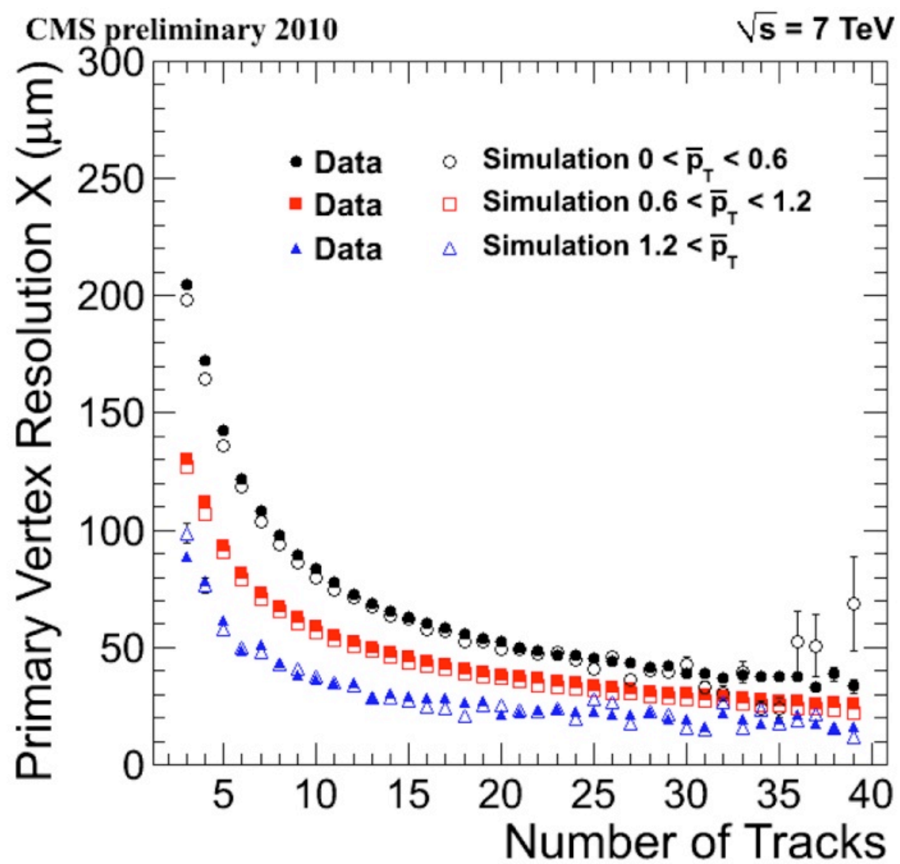
CMS PAS TRK-10-005

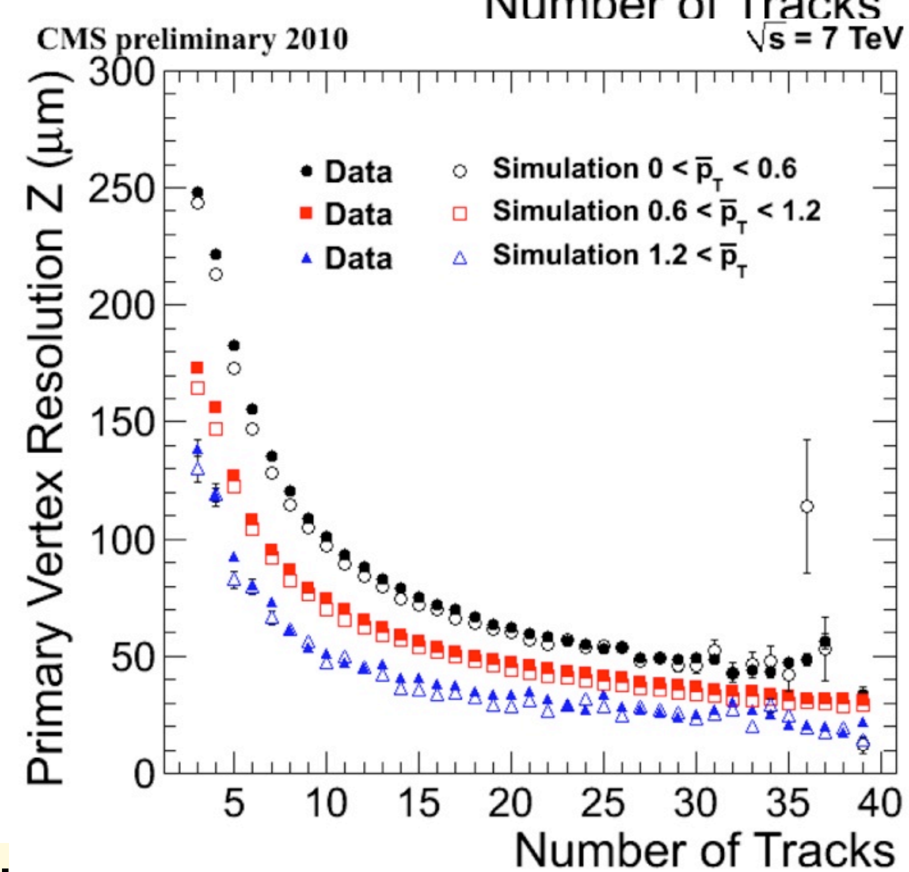
PV resolution and efficiency from ***split-method***:

- tracks of a reconstructed PV randomly split into two sets
- these two sets are fit independently
- vtx difference gives resolution
- depends on number of tracks



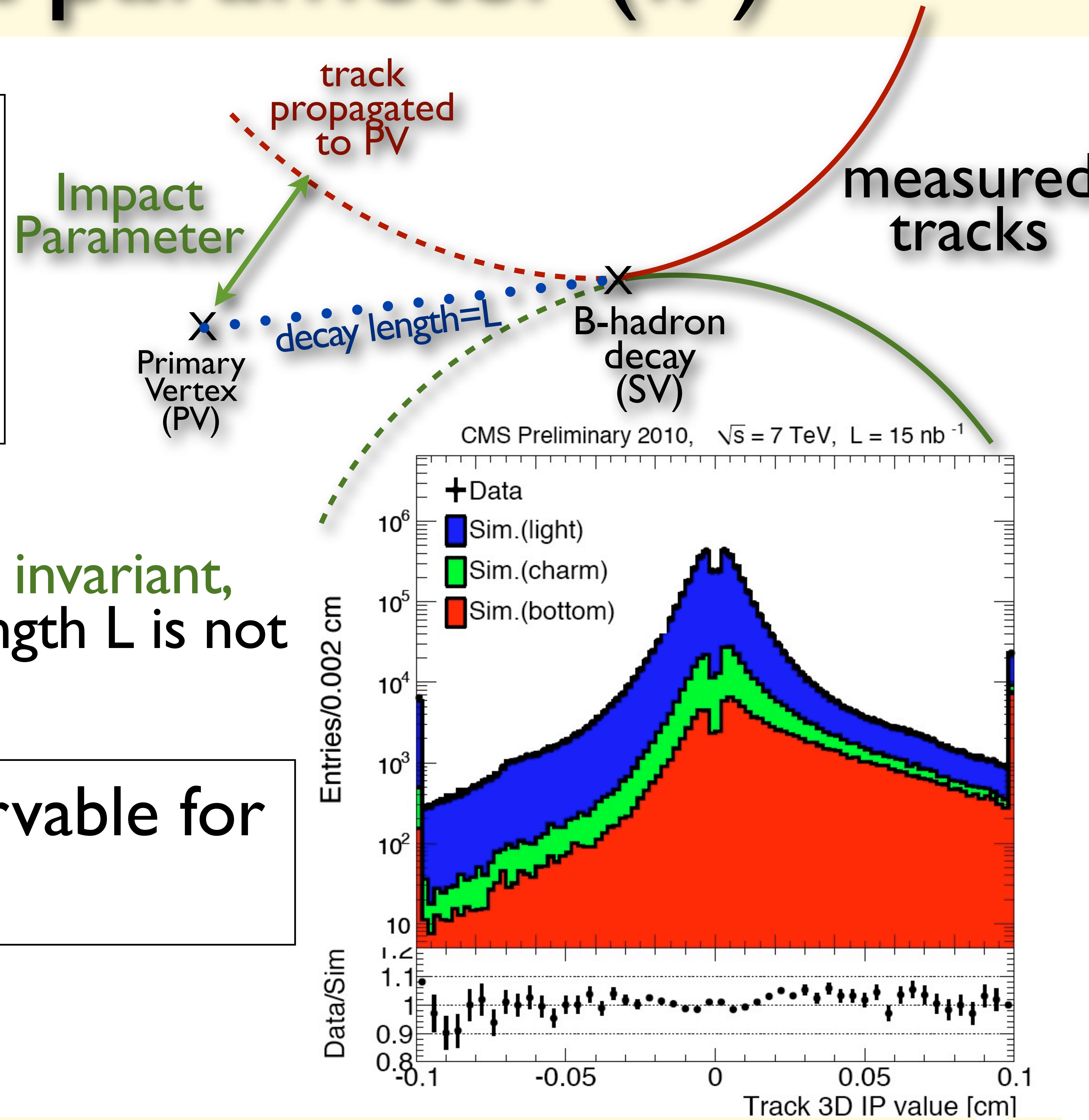






impact parameter (IP)

Impact Parameter (IP) defined as distance between track and PV at their point of closest approach



note: in the UR limit,
the IP is ~ Lorentz invariant,
while the decay length L is not

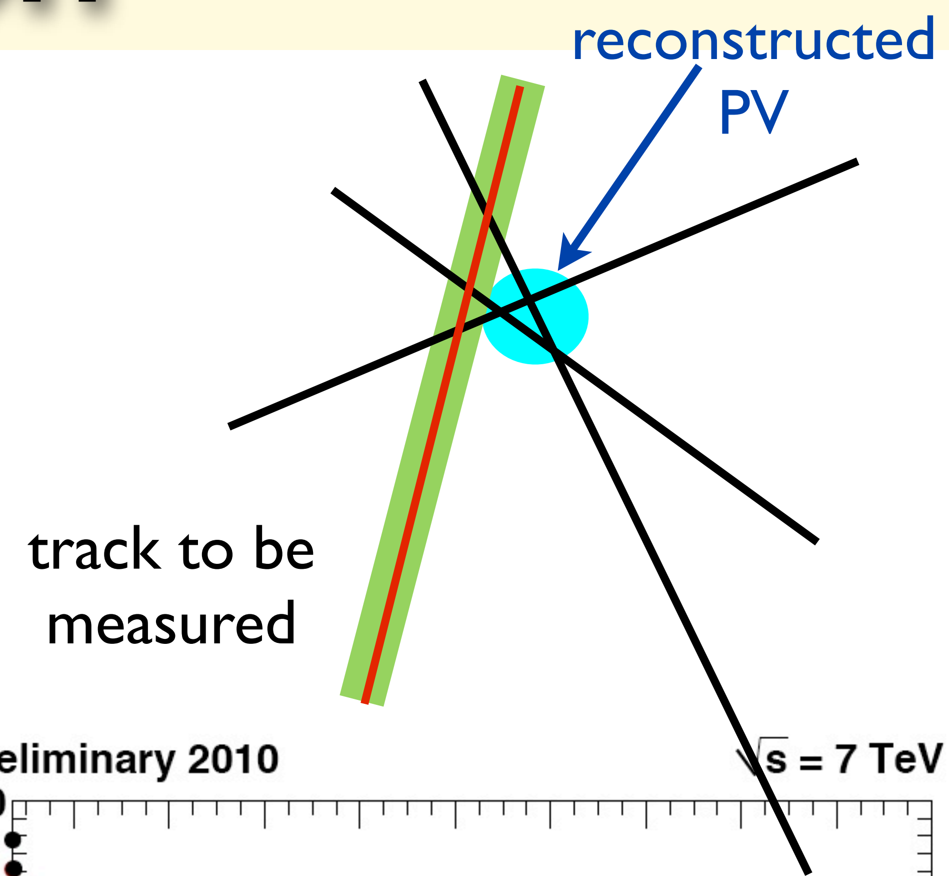
IP is important observable for b-jet tagging!

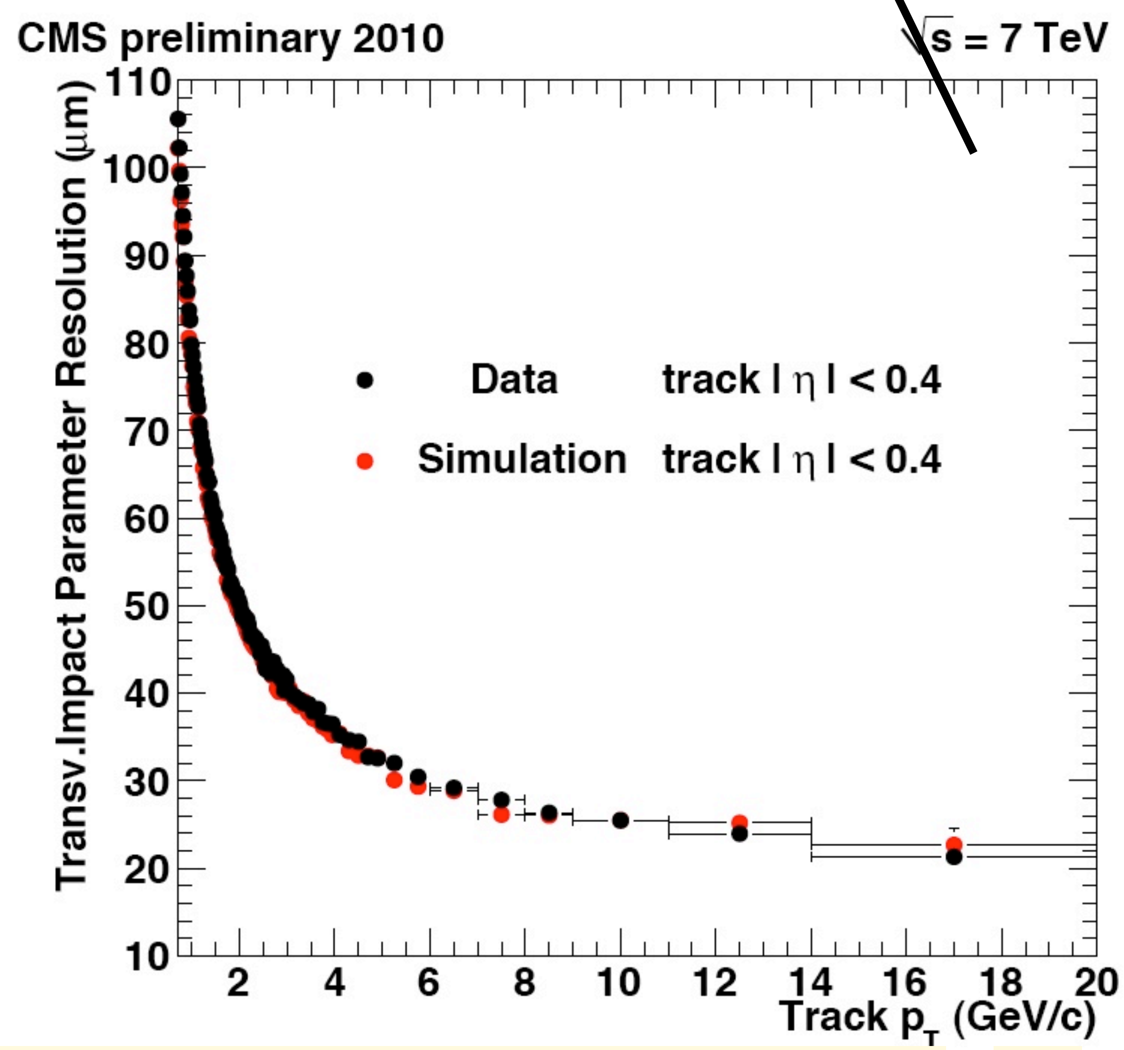
CMS PAS BTV-10-001

IP resolution

data driven method for tracks from primary vertex:

- reconstruct PV from all other tracks
- IP centered around zero
- width of IP is convolution of
  - IP resolution
  - PV resolution (known, see previous slides)
  - displaced particles (small contribution)
- extract IP resolution by fit of convoluted gaussians



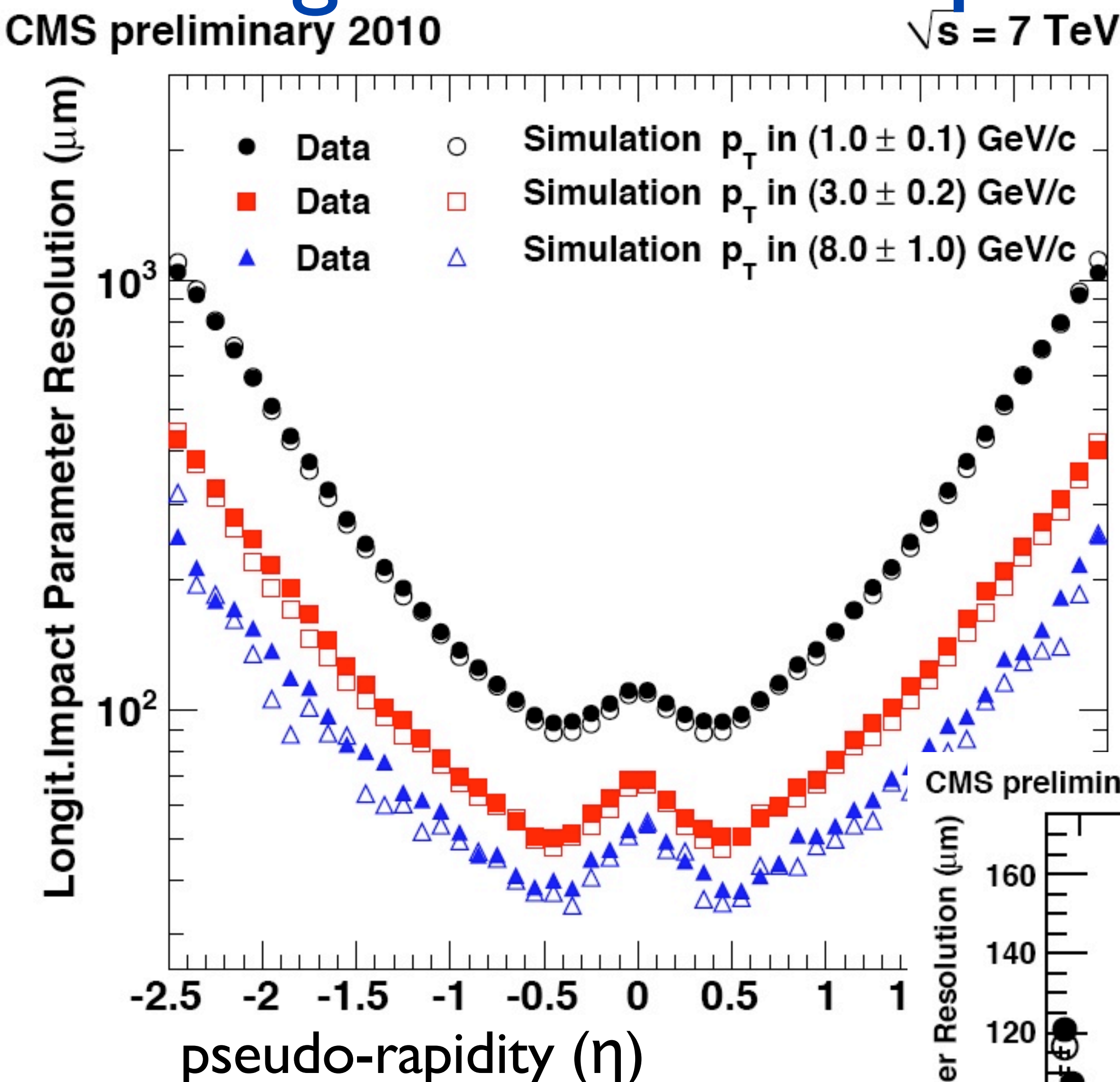


CMS PAS TRK-10-005

IP resolution

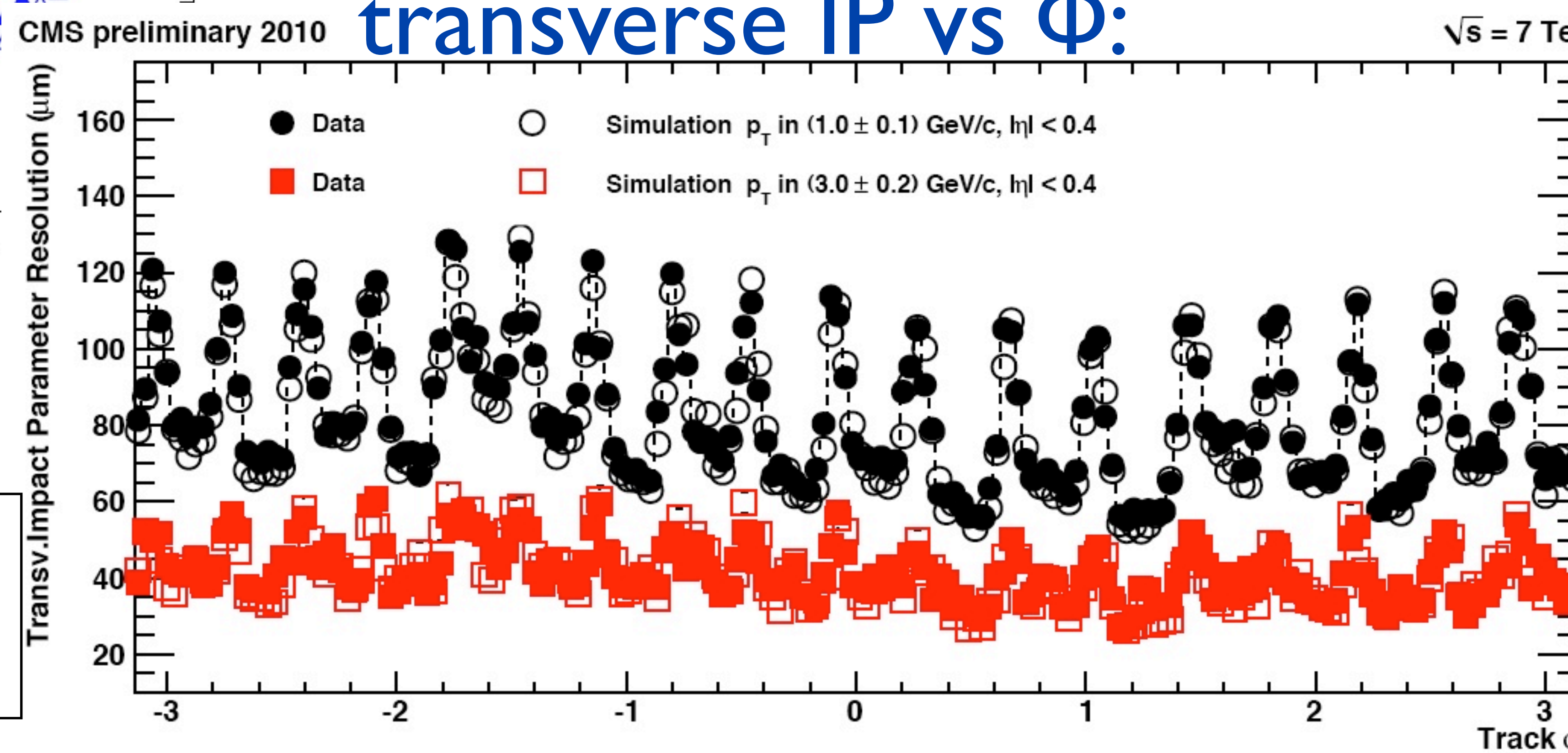
CMS PAS TRK-10-005

longitudinal IP vs η:



- IP resolution worse for large η (more material to be traversed)
- improvement at |η|=0.5 due to pixel charge sharing
- modulation in transv. IP due to pixel detector displacement

transverse IP vs Φ:

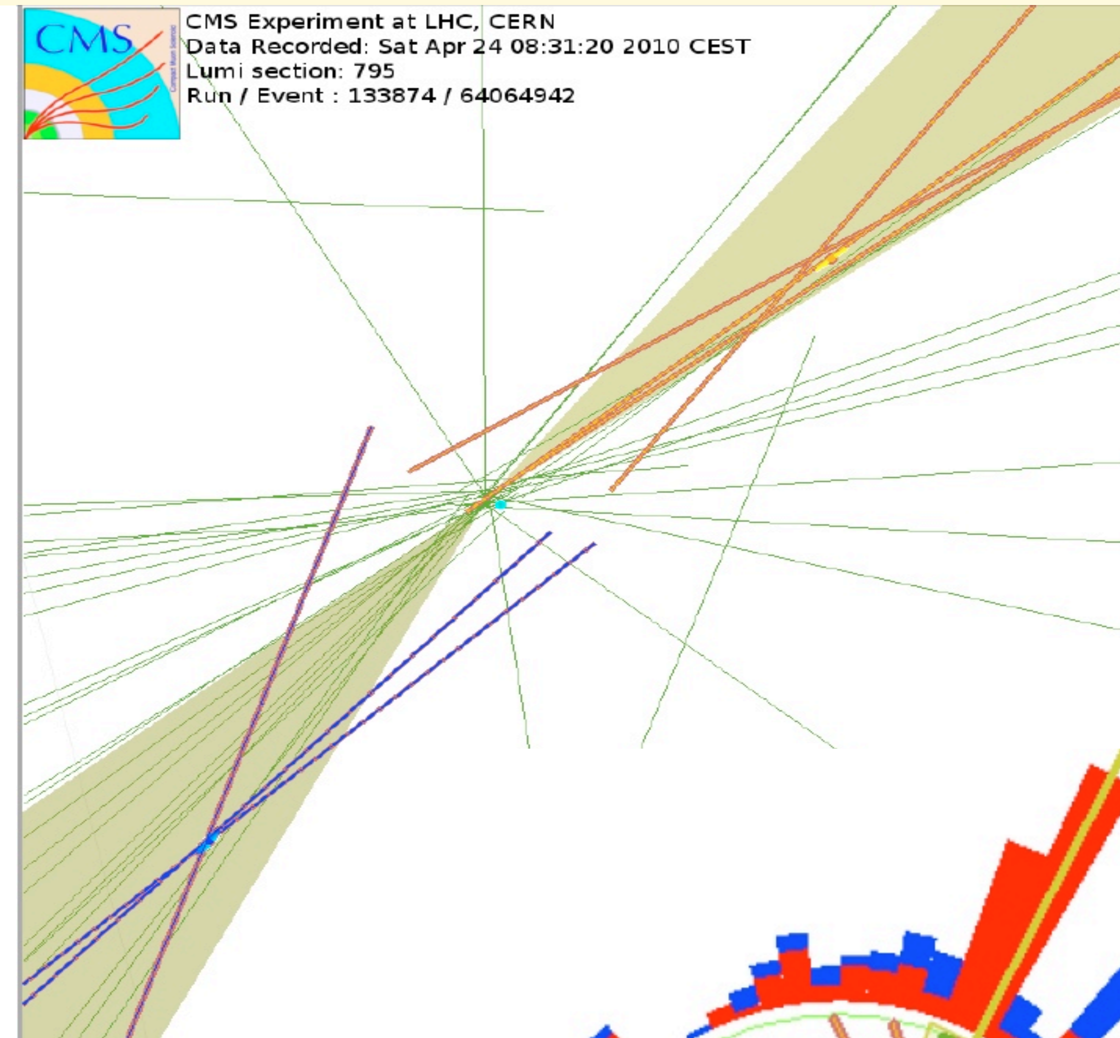


pixel layer 1 cooling pipes visible!

b-reconstruction

at hadron colliders (LHC)
b-hadrons are produced
inside of ***jets***:

their ***lifetime*** (1.5ps) and the Lorentz boost lead to displaced decay vertices

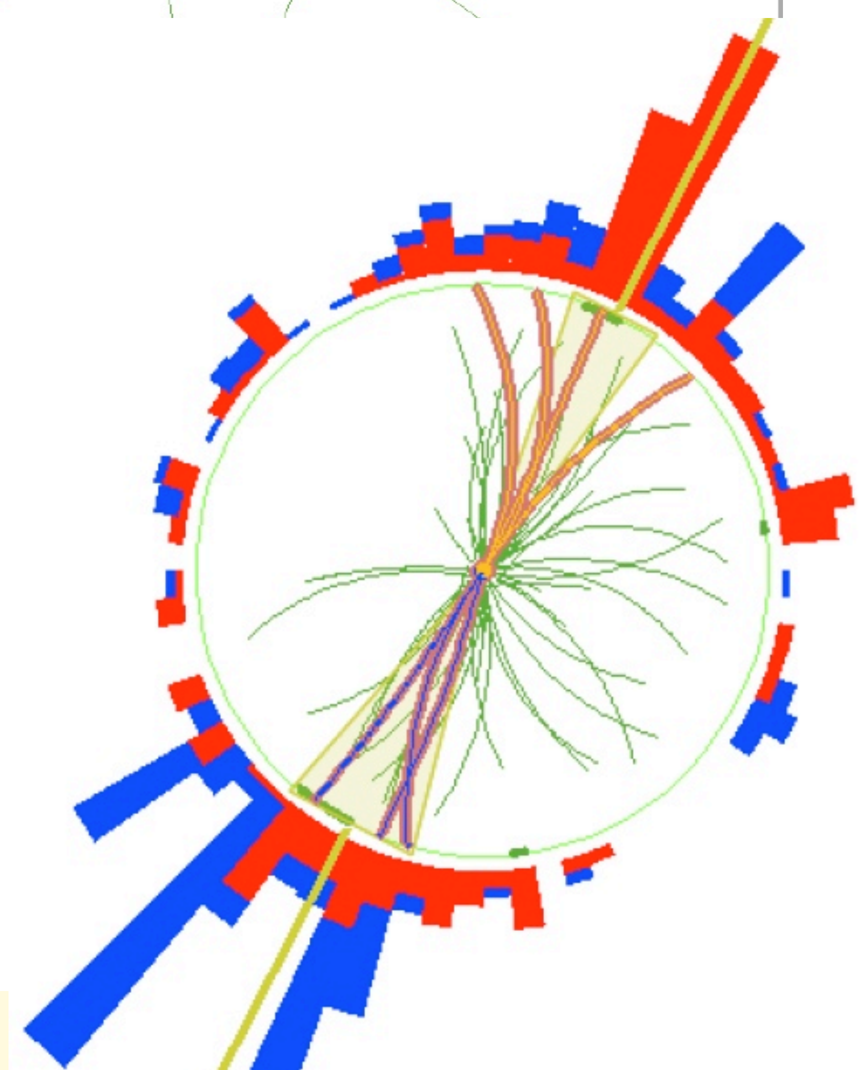


inclusive:

look for displaced tracks and vertices within jets *(b-jet tagging)*

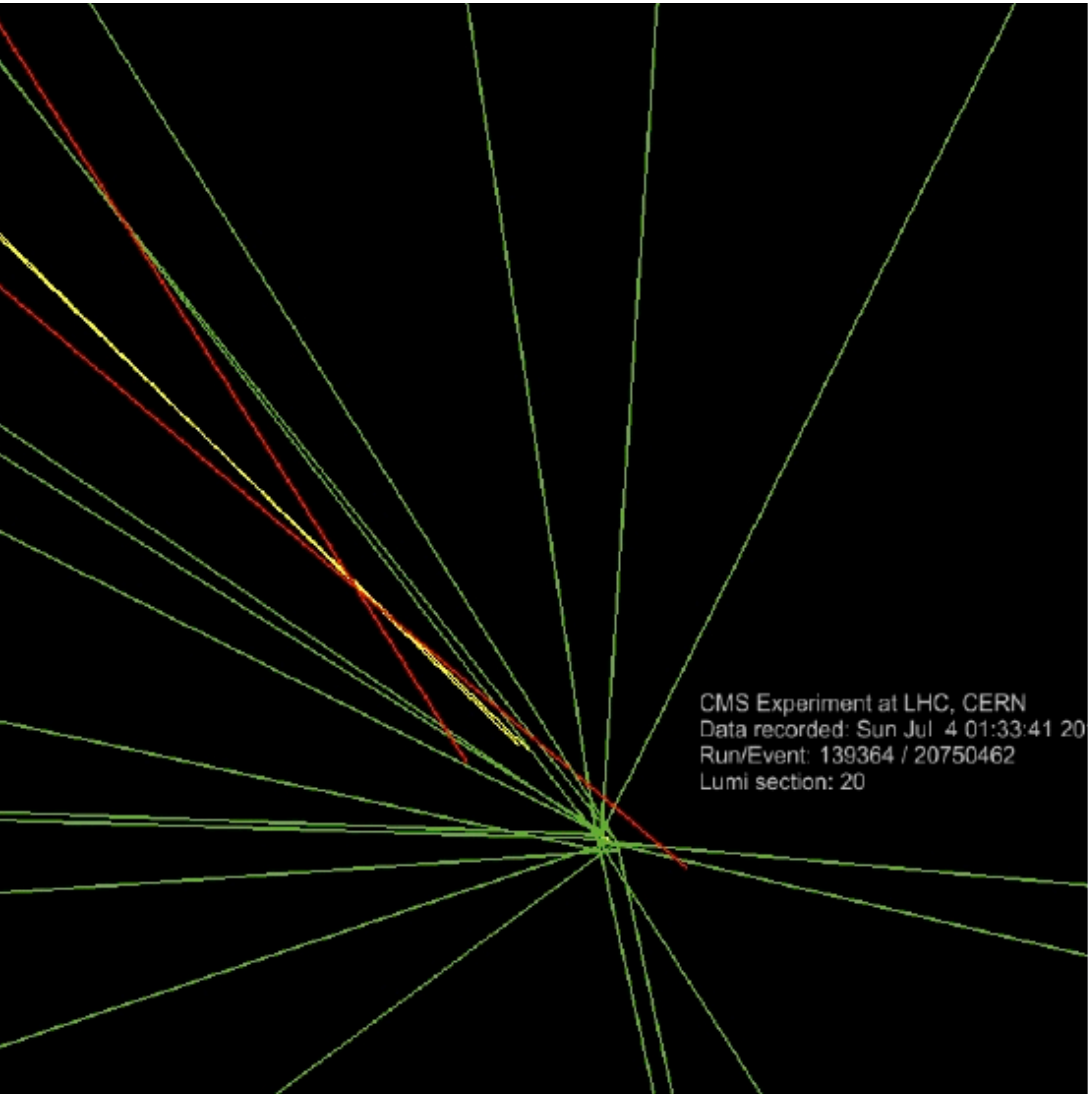
exclusive:

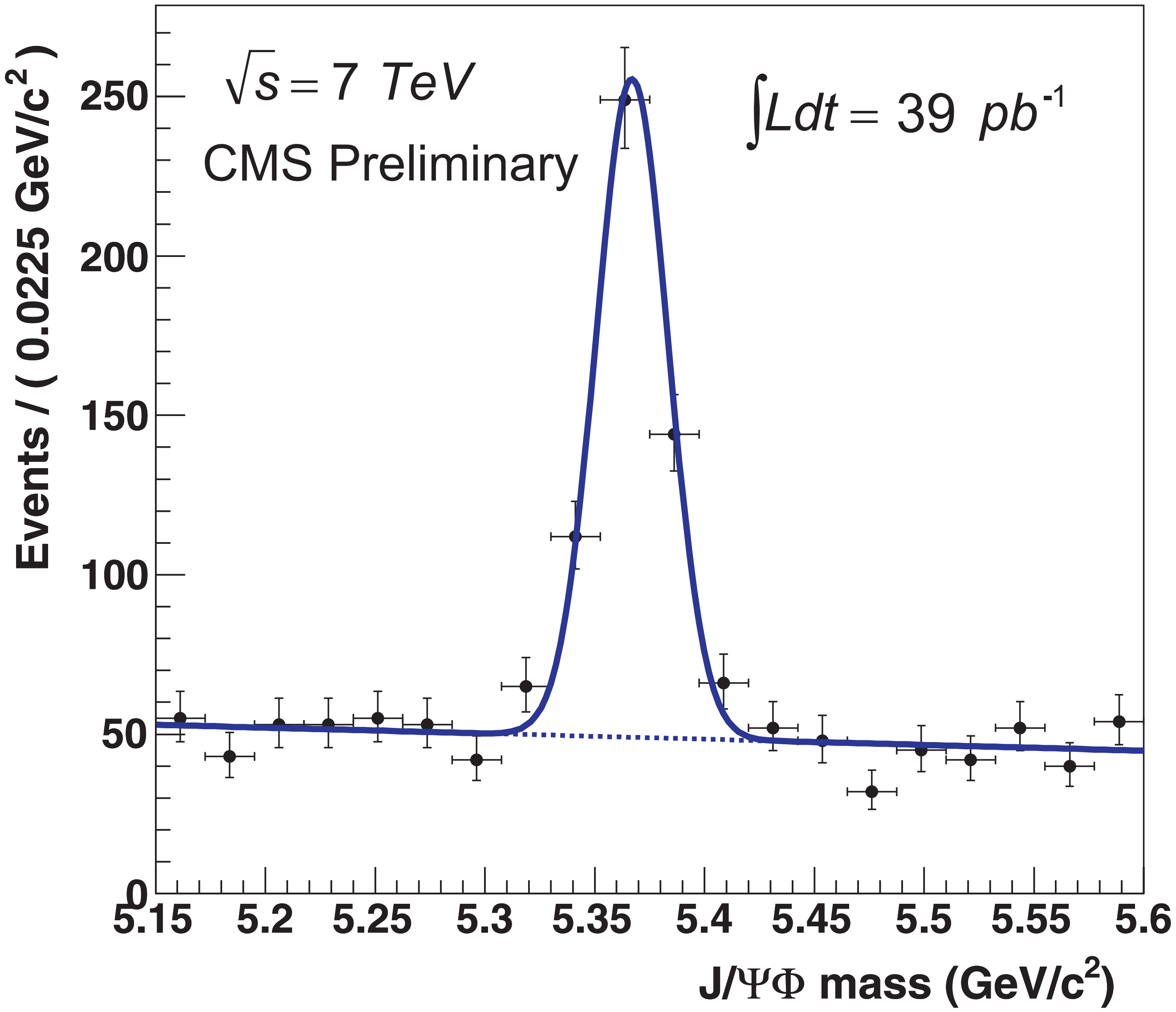
full reconstruction of the B meson four-vector using decays with $B \rightarrow J/\psi\ X$ with $J/\psi \rightarrow \mu^+\mu^-$



exclusive example: $B_s \rightarrow J/\psi\ \Phi \rightarrow \mu^+\mu^-K^+K^-$

candidate event display:



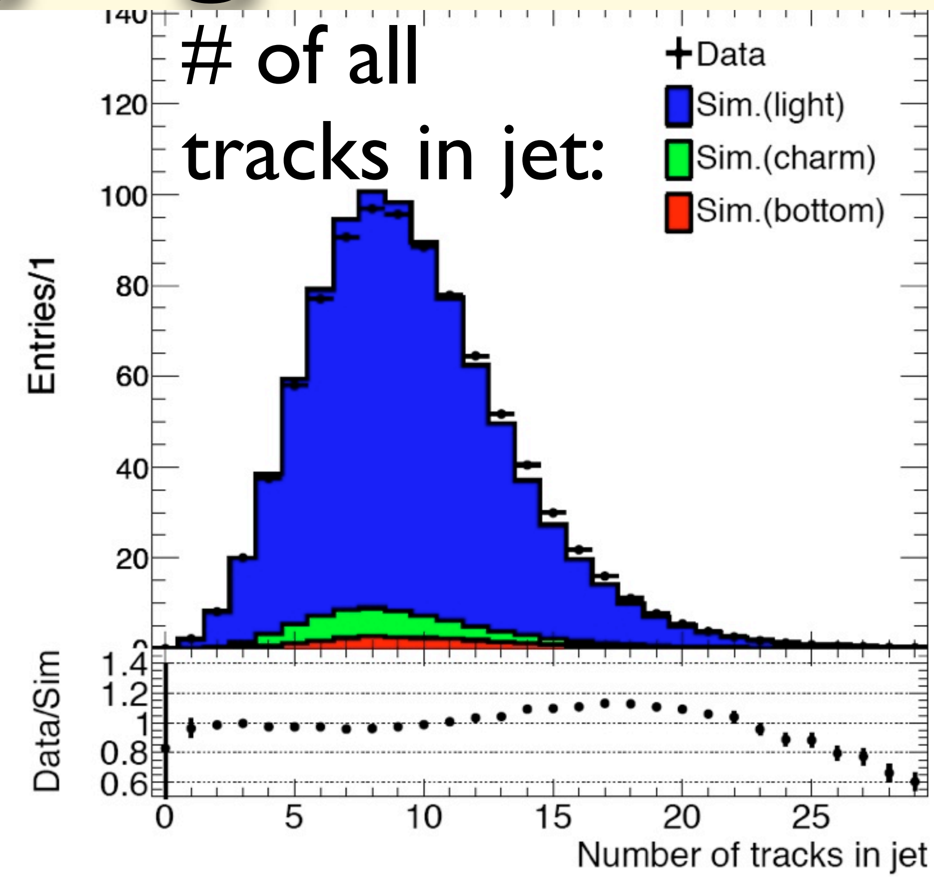


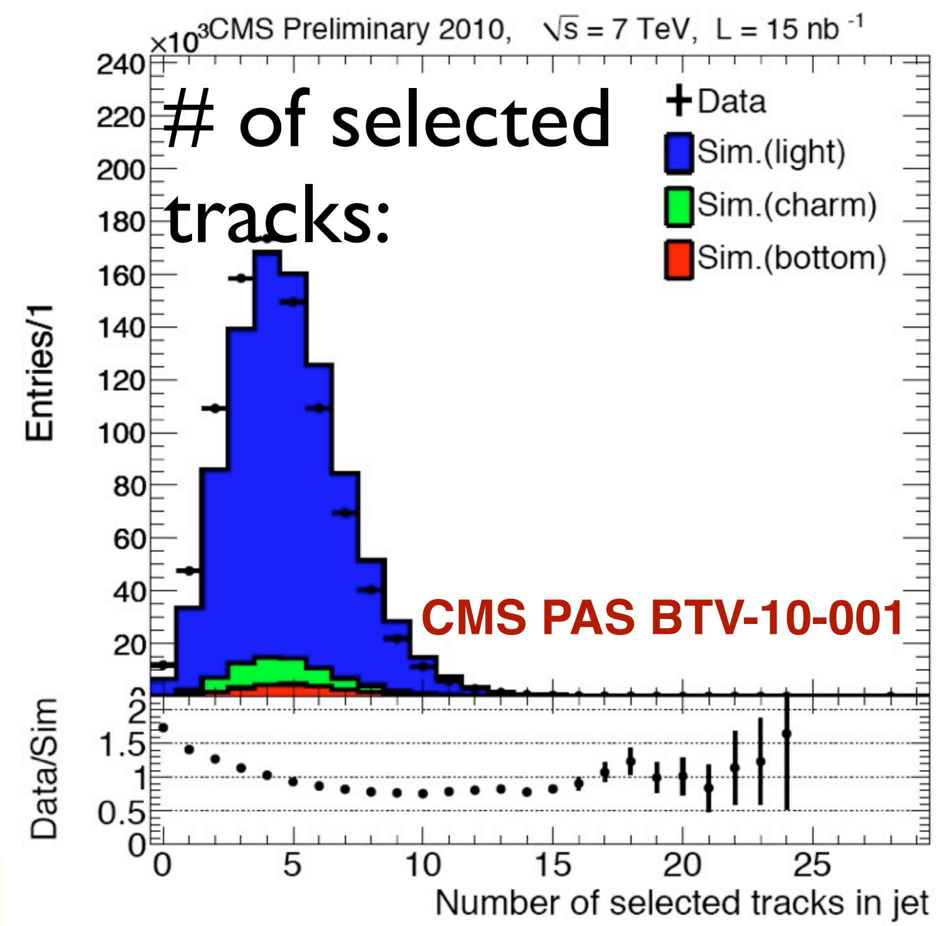
goal: measure CP violation

inclusive b-jet-tagging algorithms

b-tagging methods:

- Impact Parameter based b-tagging:
  - IP of one track in the jet ("track counting")
  - or combination of all tracks in the jet ("jet probability")
- Secondary Vertex algorithms
- combined methods using multivariate techniques
- (soft lepton tagging, mostly used to calibrate the other methods)

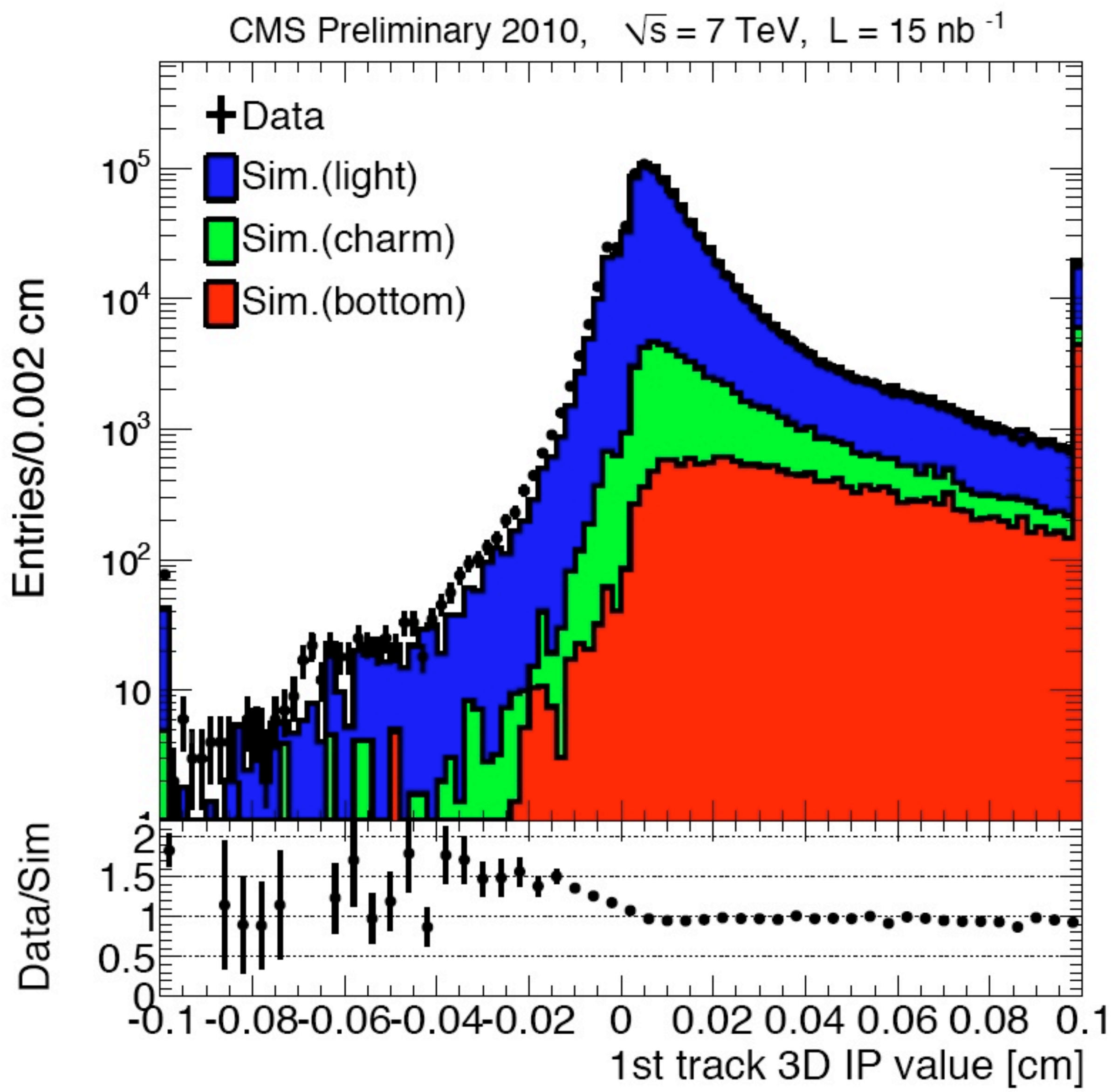




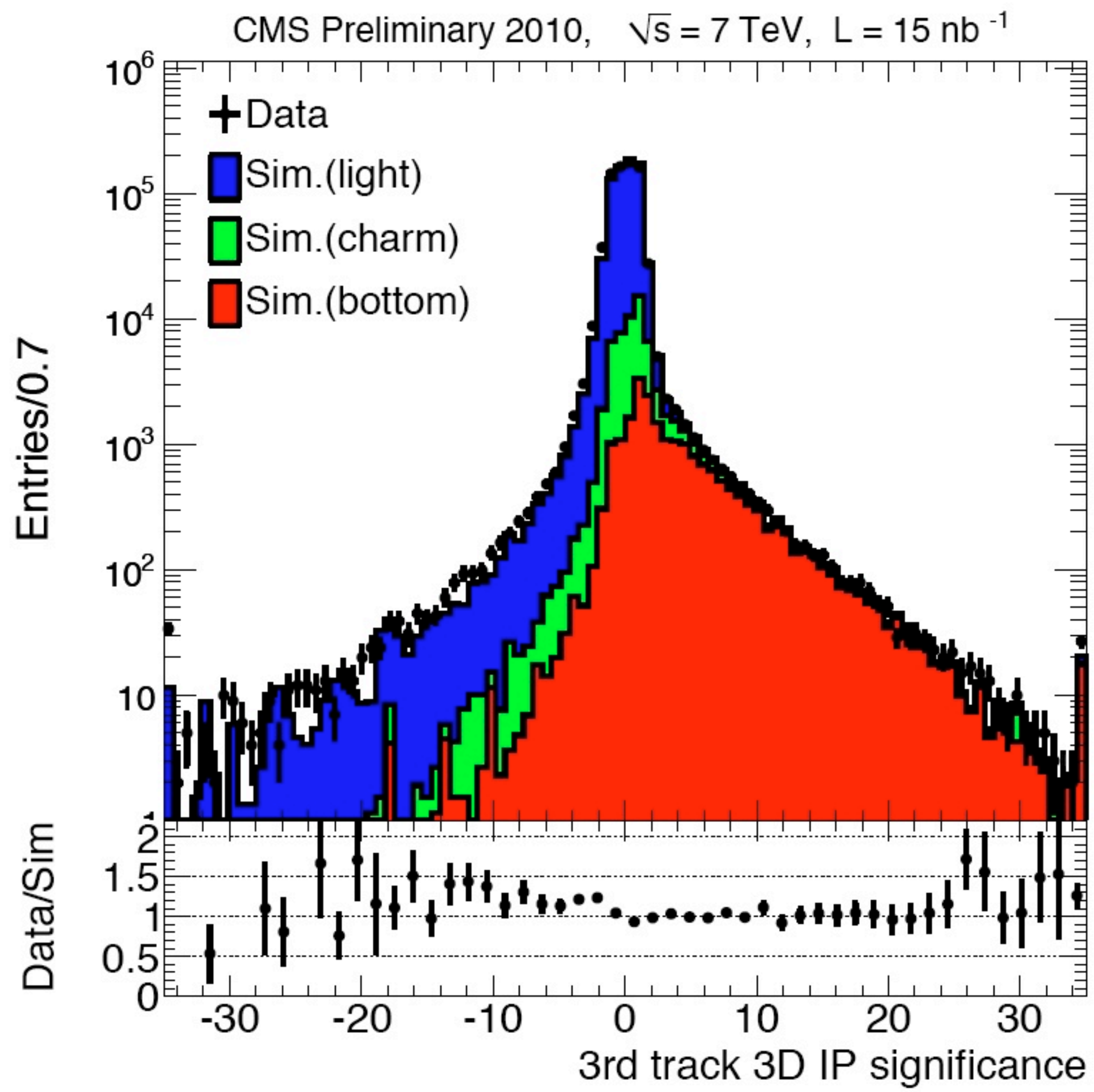
need a clean track collection to exclude fakes and badly reconstructed tracks (details in backup slides)

Impact Parameter

1st track IP value:



IP significance = value/error

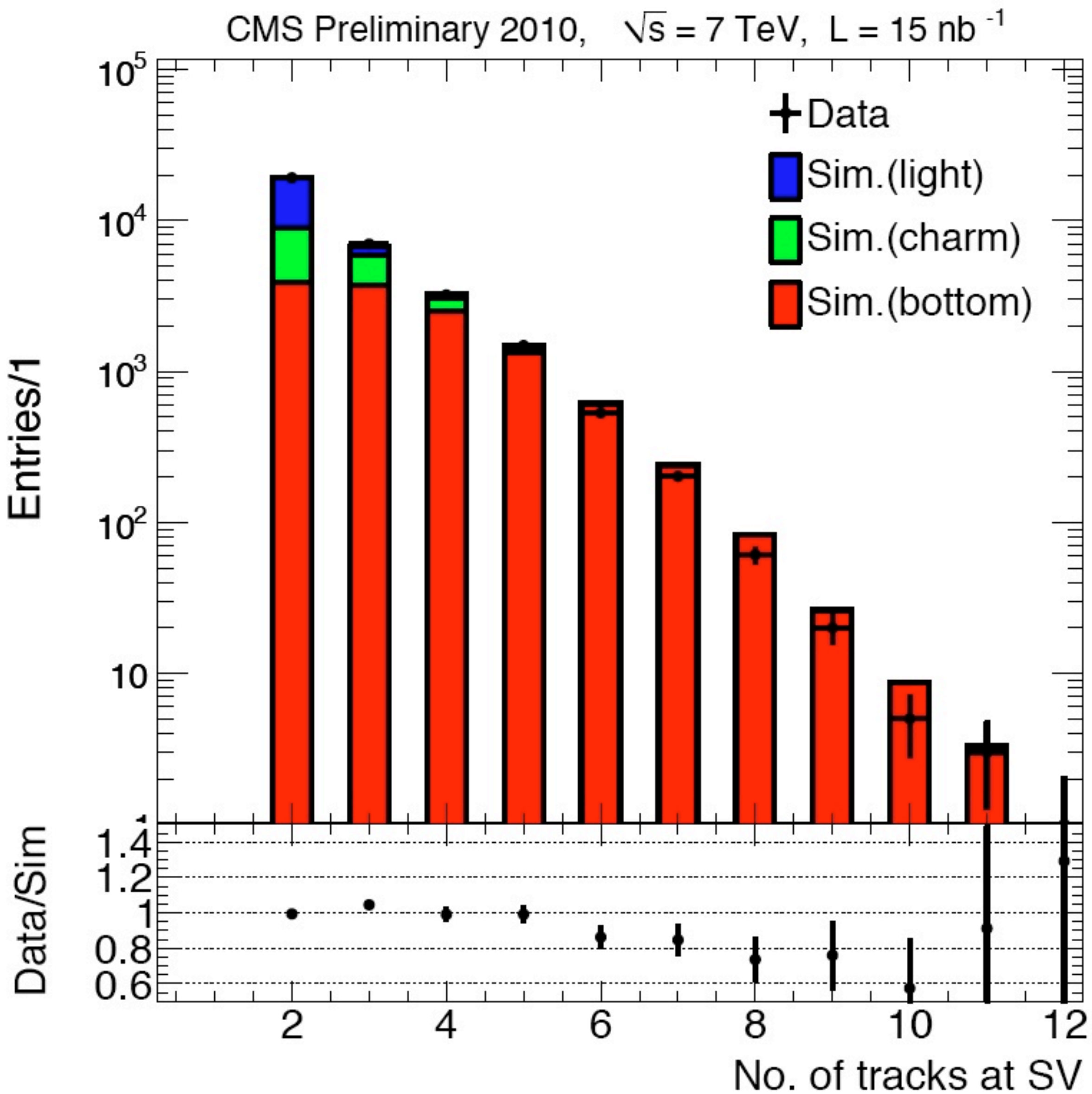


discrepancies in negative tail, when ordering tracks:

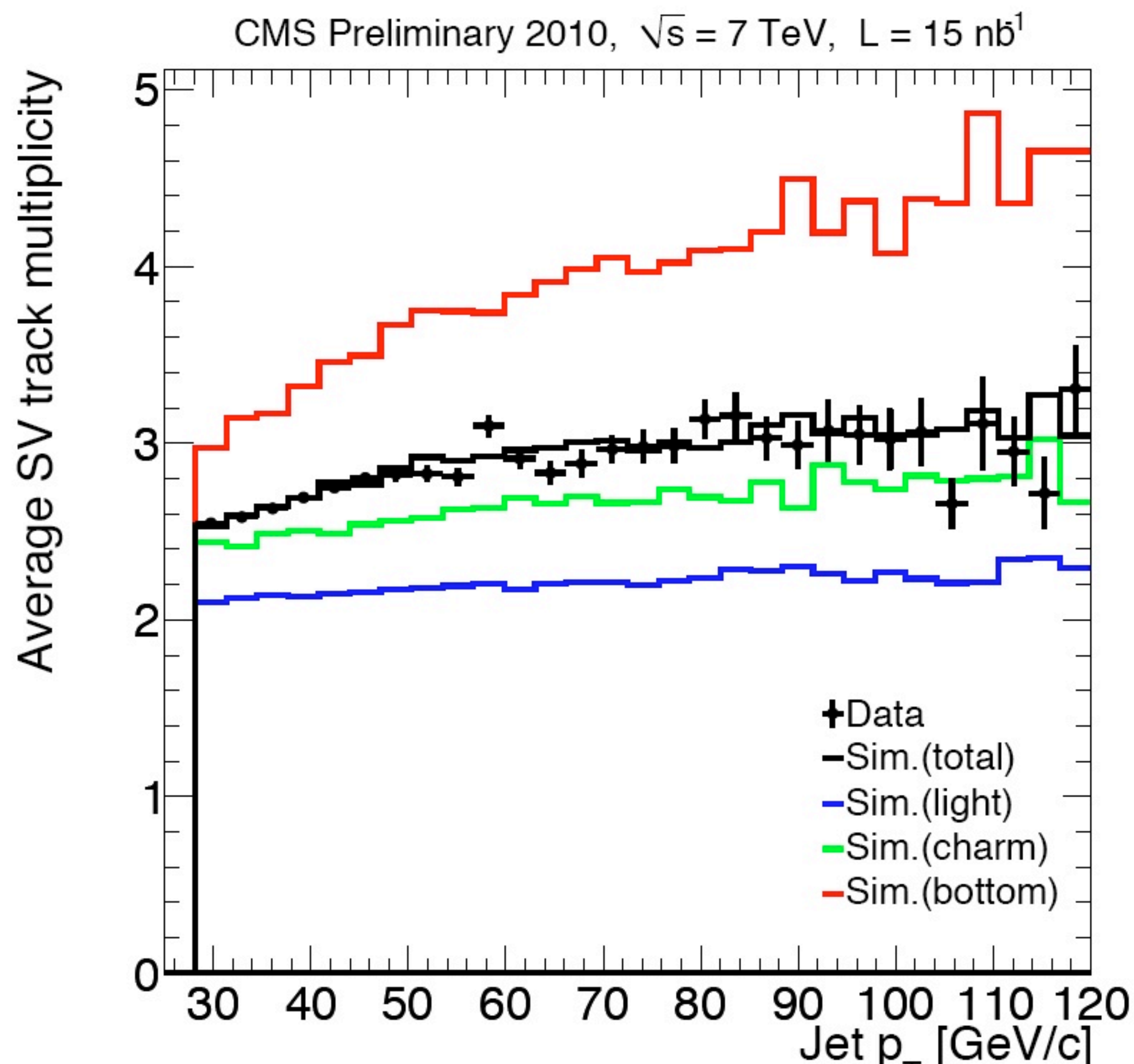
- reason is track multiplicity
- b-tagging algorithms use the positive part

CMS PAS BTV-10-001

Secondary Vertex



vertex search in a set of tracks within a jet, using the "Adaptive Vertex Fitter" (downweight outliers)

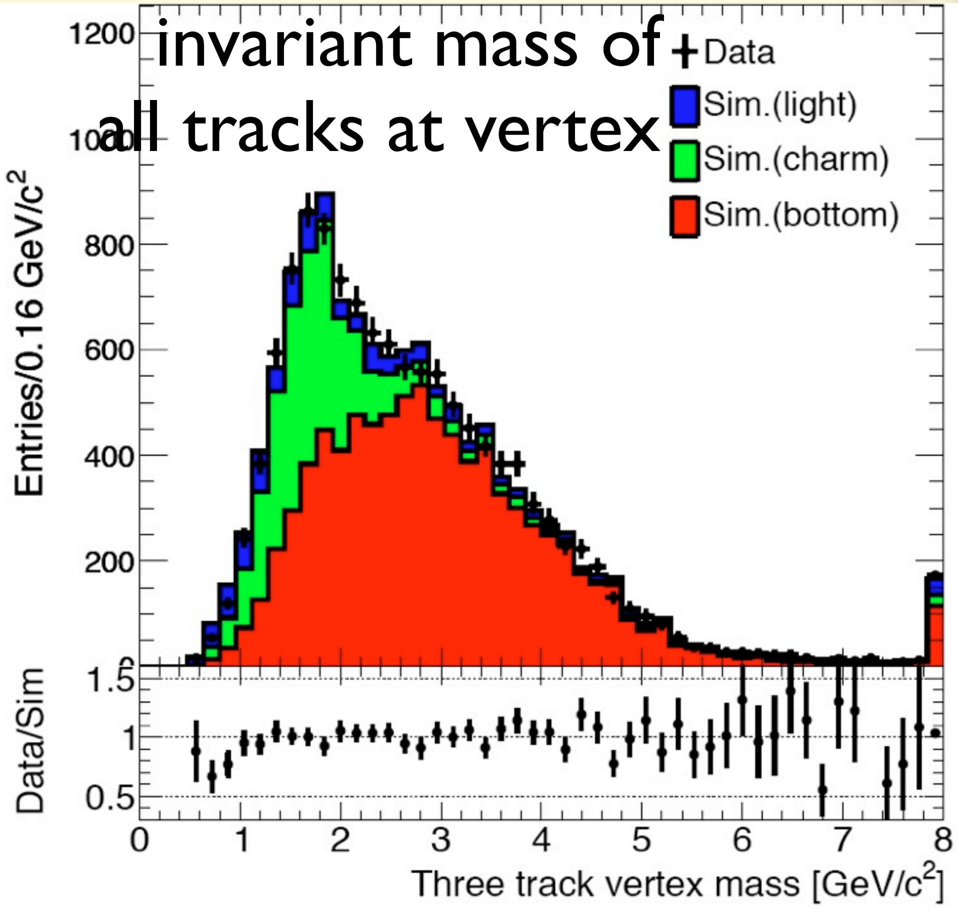


track multiplicity in agreement at 5% level

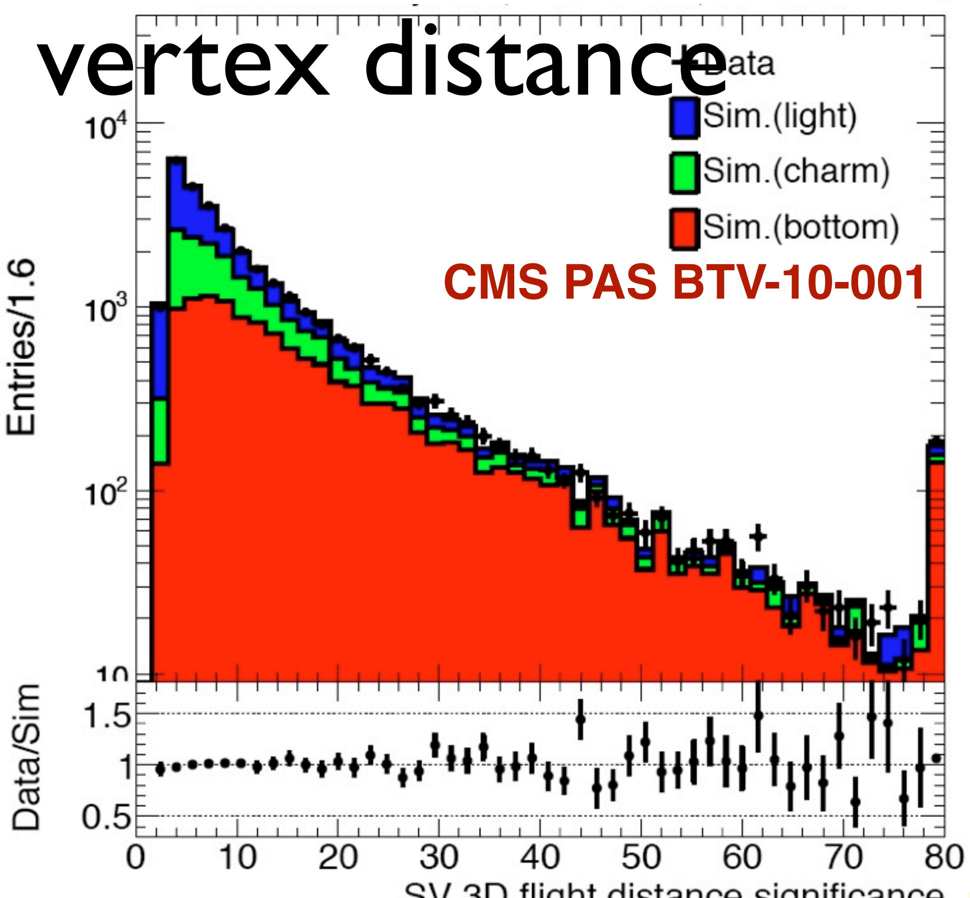
CMS PAS BTV-10-001

Secondary Vertex

invariant mass of all tracks at vertex

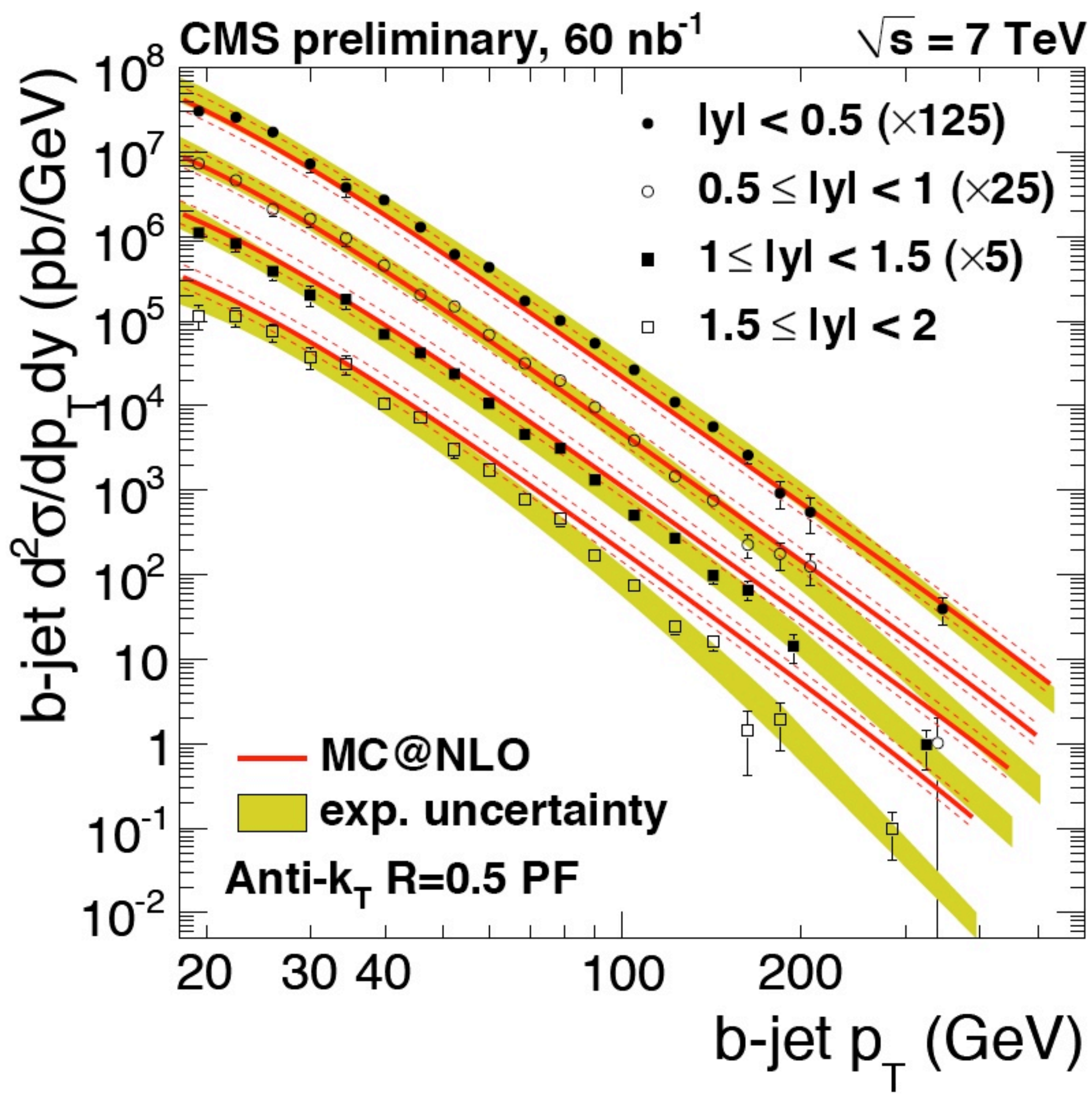


vertex distance



used in precision measurement of inclusive b-production cross-section:

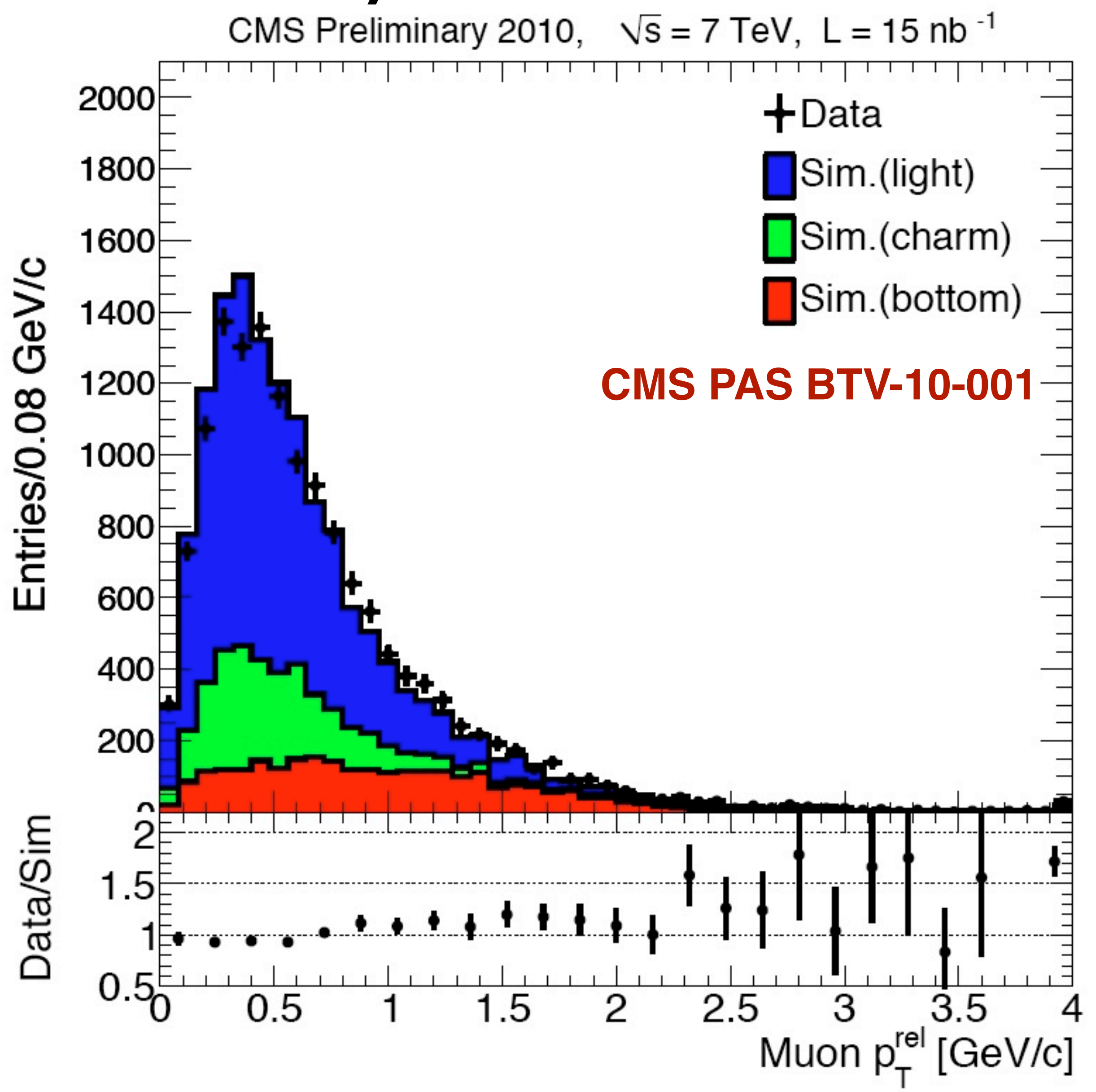
CMS PAS BPH-10-009



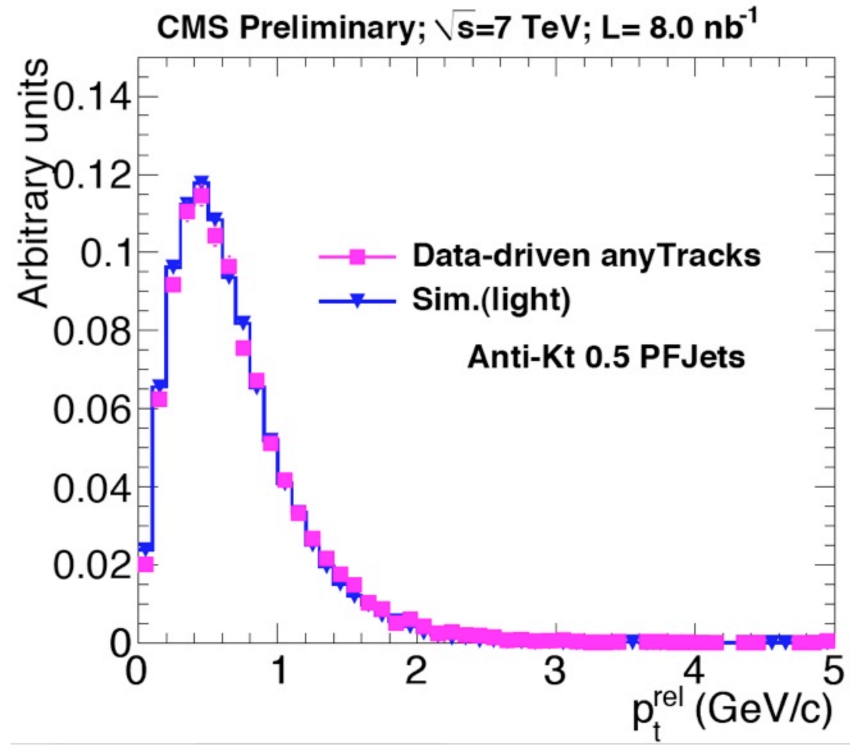
efficiency measurement: p_t^{rel}

- relative momentum of muon wrt. the jet (p_t^{rel}) is sensitive to B decays because of high B mass
- use p_t^{rel} shape to fit fractions (f_b) of b and light+c jets in tagged and anti-tagged jets
- efficiency is calculated from:

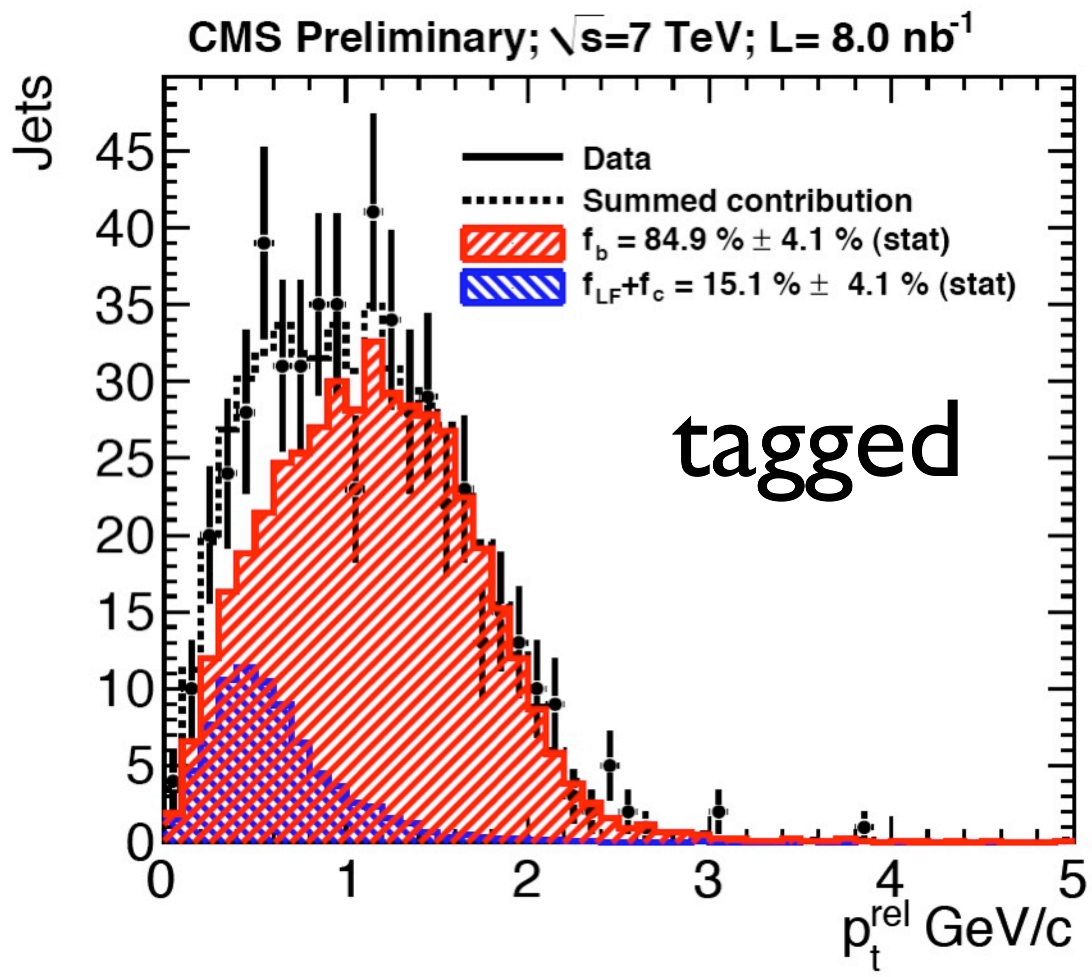
$$\epsilon_b^{data} = \frac{f_b^{tag} \cdot N_{data}^{tag}}{f_b^{tag} \cdot N_{data}^{tag} + f_b^{untag} \cdot N_{data}^{untag}}$$

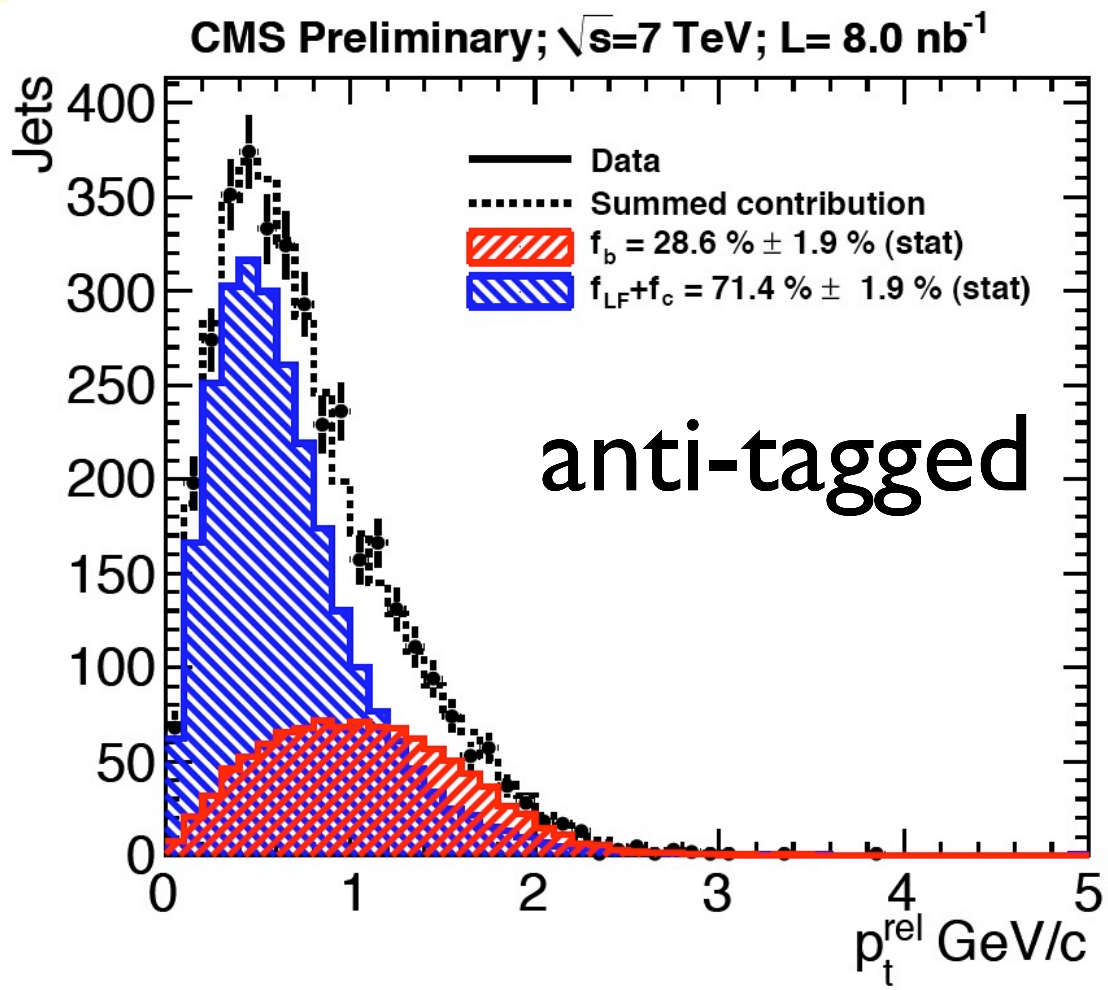


different methods to obtain data driven light flavour template:



efficiency measurement: p_t^{rel}





Tagger+Operating Point	ϵ_b^{data}	ϵ_b^{MC}	SF_b
SSVHPT	0.203 ± 0.015	0.207± 0.002	0.98± 0.08± 0.18
SSVHEM	0.405 ± 0.016	0.417± 0.003	0.97± 0.04± 0.19
SSVHET	0.127 ± 0.017	0.131± 0.002	0.97± 0.13± 0.21
TCHPL	0.404 ± 0.018	0.444± 0.003	0.91± 0.04± 0.19
TCHPM	0.303 ± 0.015	0.331± 0.003	0.92± 0.05± 0.19
TCHPT	0.233 ± 0.014	0.244± 0.002	0.95± 0.06± 0.19
TCHEL	0.562 ± 0.020	0.636± 0.003	0.88± 0.03± 0.19
TCHEM	0.455 ± 0.016	0.494± 0.003	0.92± 0.03± 0.20
TCHET	0.151 ± 0.015	0.150± 0.002	1.01± 0.10± 0.19

still large systematic uncertainties

CMS PAS BTV-10-001

efficiency measurement: System8

method: three uncorrelated identification criteria combined:

1. working point of algorithm under study
2. cut on muon P_{Trel}
3. presence of second b-tagged-jet (b-Quark pair production)

=> get system of 8 linear equations:

$$
\begin{aligned}
n &= n_b + n_{cl} \\
p &= p_b + p_{cl} \\
n^{tag} &= \varepsilon_b^{tag} n_b + \varepsilon_{cl}^{tag} n_{cl} \\
p^{tag} &= \beta \varepsilon_b^{tag} p_b + \alpha \varepsilon_{cl}^{tag} p_{cl} \\
n^{mu} &= \varepsilon_b^{mu} n_b + \varepsilon_{cl}^{mu} n_{cl} \\
p^{mu} &= \delta \varepsilon_b^{mu} p_b + \gamma \varepsilon_{cl}^{mu} p_{cl} \\
n^{tag,mu} &= \kappa_b \varepsilon_b^{tag} \varepsilon_b^{mu} n_b + \kappa_{cl} \varepsilon_{cl}^{tag} \varepsilon_{cl}^{mu} n_{cl} \\
p^{tag,mu} &= \delta \kappa_b \beta \varepsilon_b^{tag} \varepsilon_b^{mu} p_b + \gamma \kappa_{cl} \alpha \varepsilon_{cl}^{tag} \varepsilon_{cl}^{mu} p_{cl}
\end{aligned}
$$

8 unknowns

n = number of jets with muons

p = as n, but require second tagged jet in the event

n^{tag} = (lifetime-) tagged jets with muons

ε_b^{tag} = (lifetime-) btagging efficiency

n^{mu} = number of jets with muons passing PTrel cut

$\alpha, \beta, \kappa_b, \kappa_{cl}, \delta, \gamma$ are correlation factors **=> need to be determined on MC**

needs large data samples, results expected soon

mistag rate measurement

Method:

the mistag rate is evaluated as $\epsilon_{data}^{mistag} = \epsilon_{data}^{-} \cdot R_{light}$ where ϵ_{data}^{-} is the negative tag rate in data and $R_{light} = \epsilon_{MC}^{mistag} / \epsilon_{MC}^{-}$ is the ratio between the light flavour mistag rate and negative tag rate of all jets in the simulation

R_{light} is sensitive to:

- fractions of b and c in the negative tags (increase R_{light})
- other displaced processes, K_s^0 and Λ, material interactions and mismeasured tracks (increase R_{light})
- residual differences between u,d,s-quark and gluon jets
- angular resolutions of jet axis and IP (sign flips)

➔ taken as systematic errors

negative tags

CMS PAS BTV-10-001

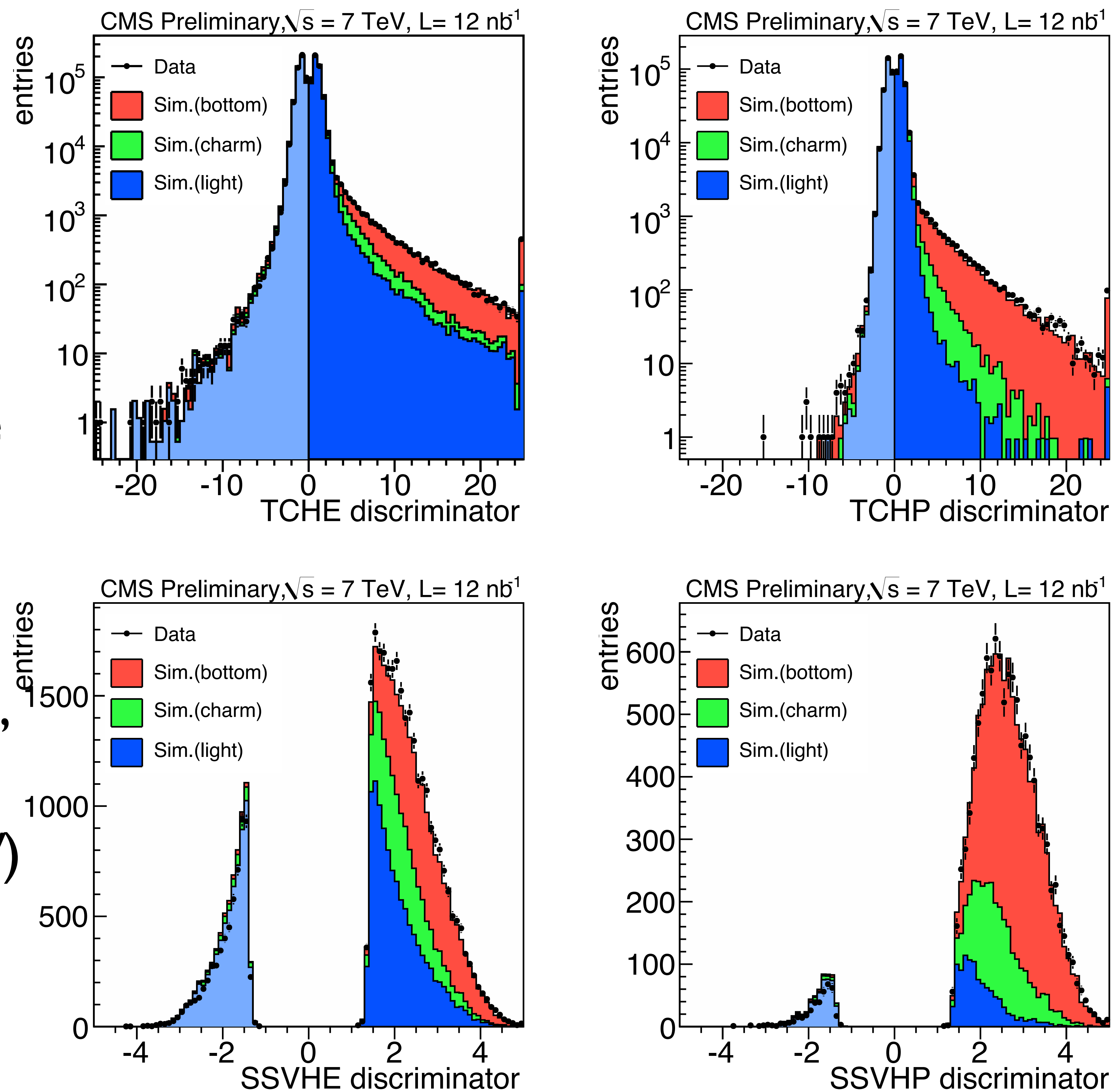
Impact Parameter:

- invert the order of negative IPs from the most negative upwards
- take the 2nd (3rd) negative IP for high eff (high purity) algorithms
- ordering on positive side remains unchanged

Secondary Vertex:

- negative sign if vertex is reconsructed "upstream" wrt. the primary vertex (on the other side of the PV)

Note: K_s^0 and Λ reweighting is applied!



mistag rate: results

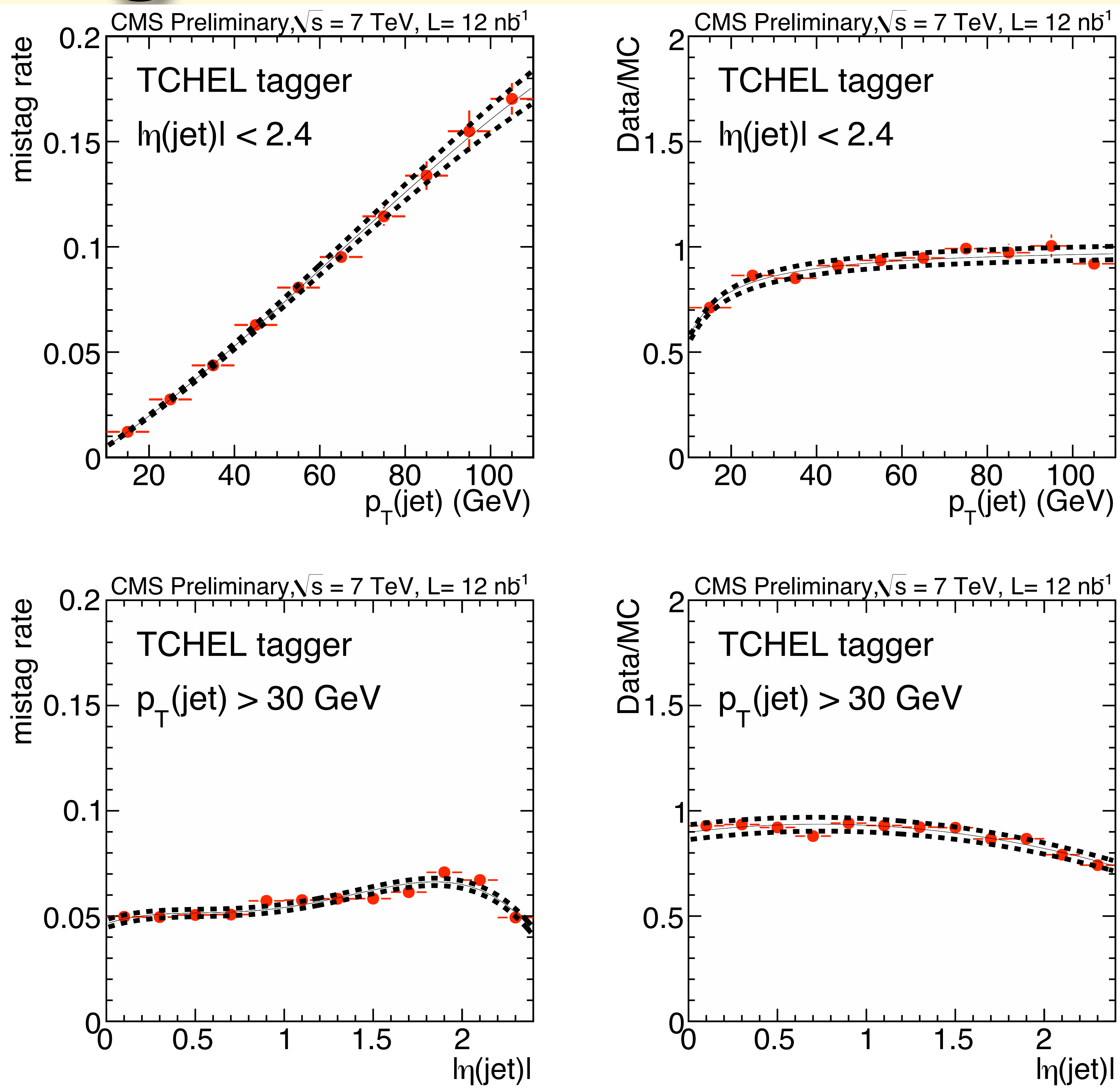
CMS PAS BTV-10-001

example: "Track counting" algorithm

mistag rate and scale factor as function of jet p_t and $|\eta|$

the dashed lines represent statistical and systematic uncertainties

results for "Secondary Vertex" algorithm in backup



conclusions

- CMS tracker operating well right from the start
- measured and validated detector geometry/material
- commissioned data driven methods to measure tracking performance (resolutions, efficiencies)
- excellent agreement with expectations from simulation
- b-tagging algorithms commissioned:
  - "Track Counting", "Secondary Vertex", "Track Probability" working well as expected
  - higher level algorithms (combined methods) still need to be commissioned
  - b-jet triggers will be important with increased luminosity
- b-tagging performance measurements show agreement of 10-15% with simulation

backup slides

tracking efficiency

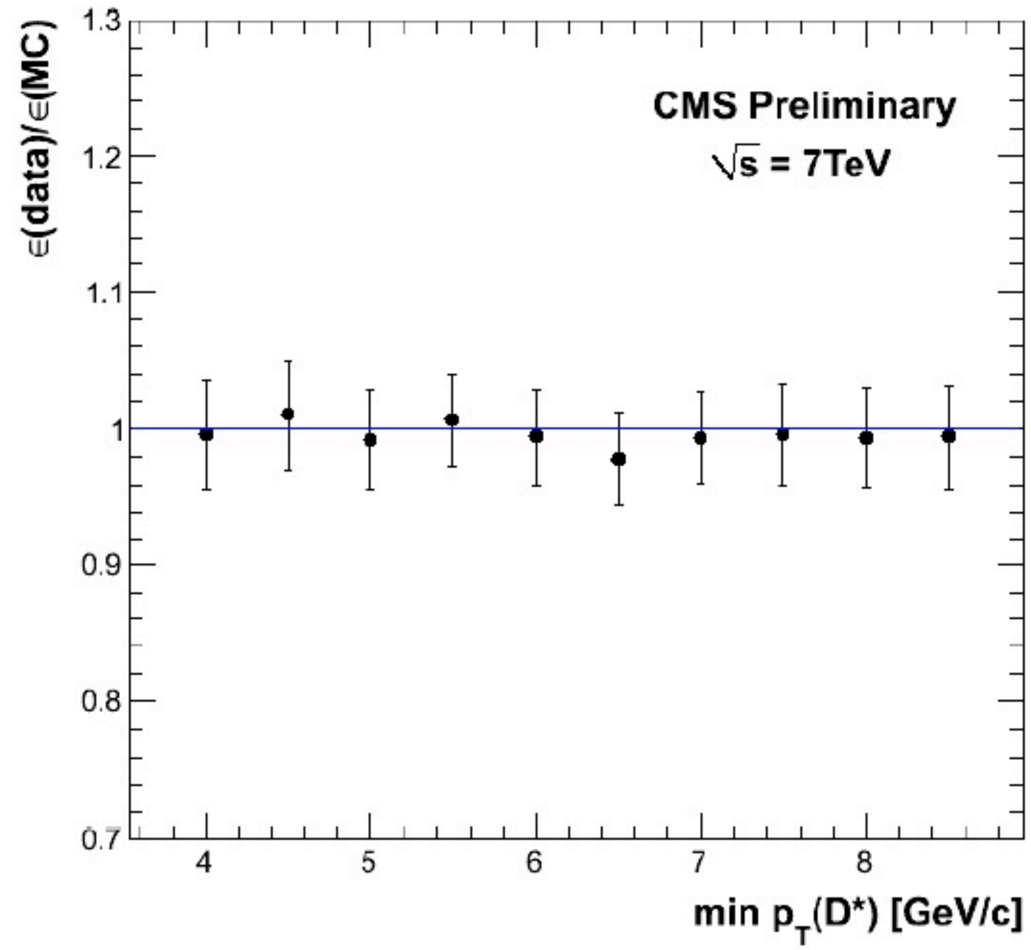
tracking efficiency for pions can be measured using ratio of $D^0 \rightarrow K^- \pi^+$ and $D^0 \rightarrow K^- \pi^+ \pi^- \pi^+$ decays

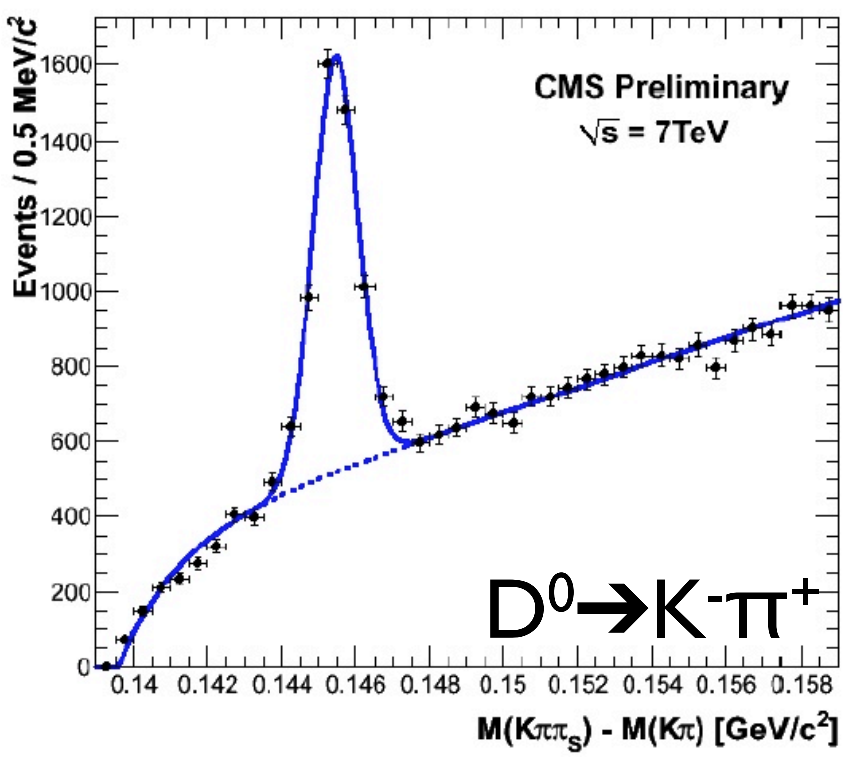
$$\mathcal{R} = \frac{N_{K3\pi}}{N_{K\pi}} \cdot \frac{\epsilon_{K\pi}}{\epsilon_{K3\pi}}$$

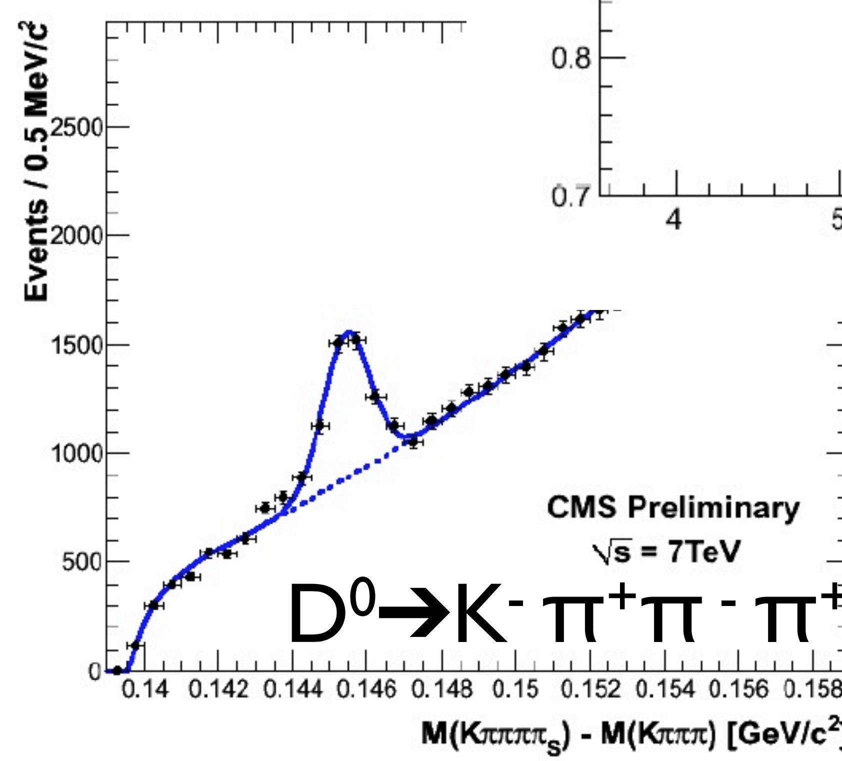
← ratio eff. corrected yields

$$\frac{\epsilon(\text{data})}{\epsilon(\text{MC})} = \sqrt{\frac{\mathcal{R}}{\mathcal{R}(\text{PDG})}}$$

← ratio of $D^0 \rightarrow K^- \pi^+$ and $D^0 \rightarrow K^- 3\pi$ branching from PDG

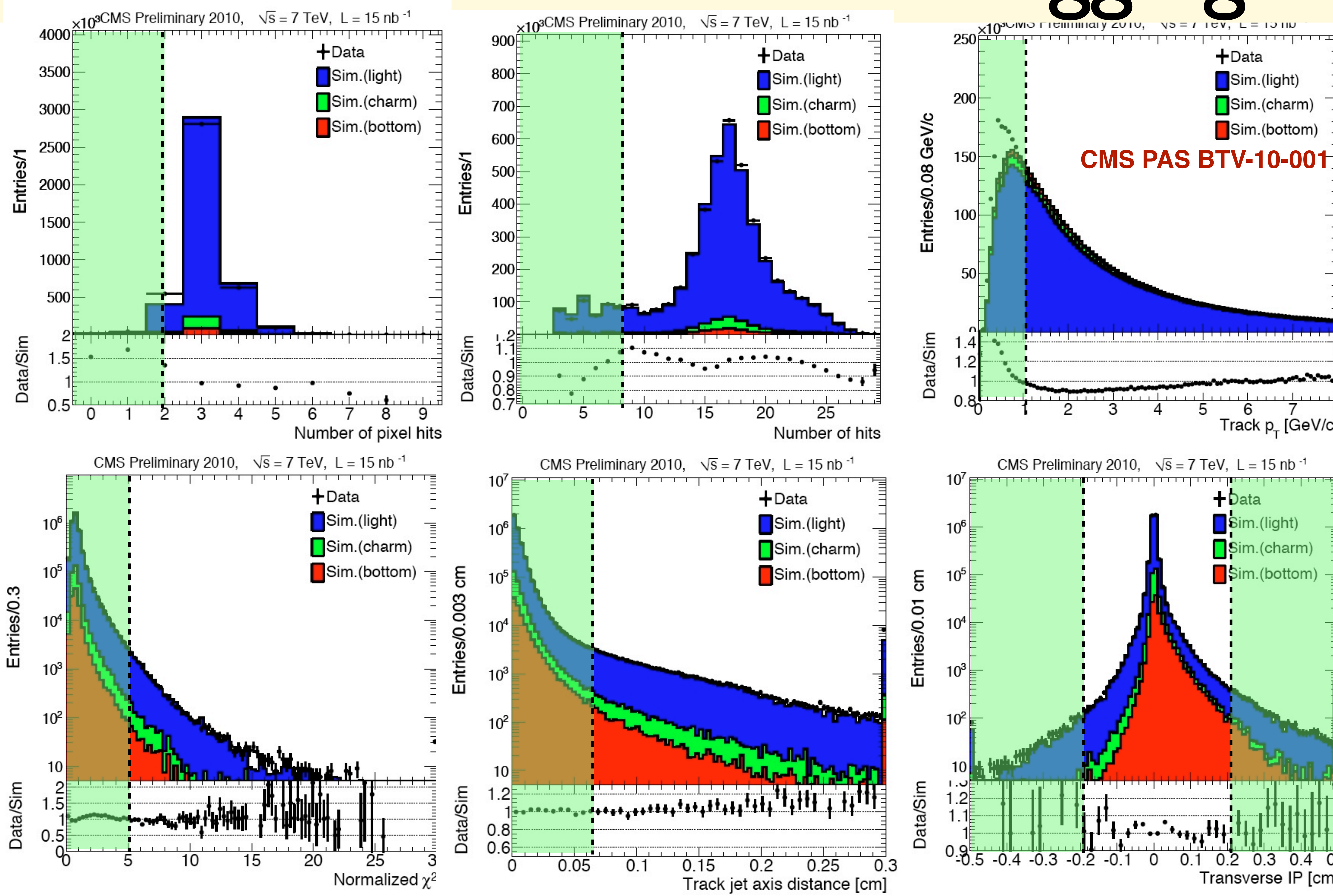






CMS PAS TRK-10-002

track selection for b-tagging

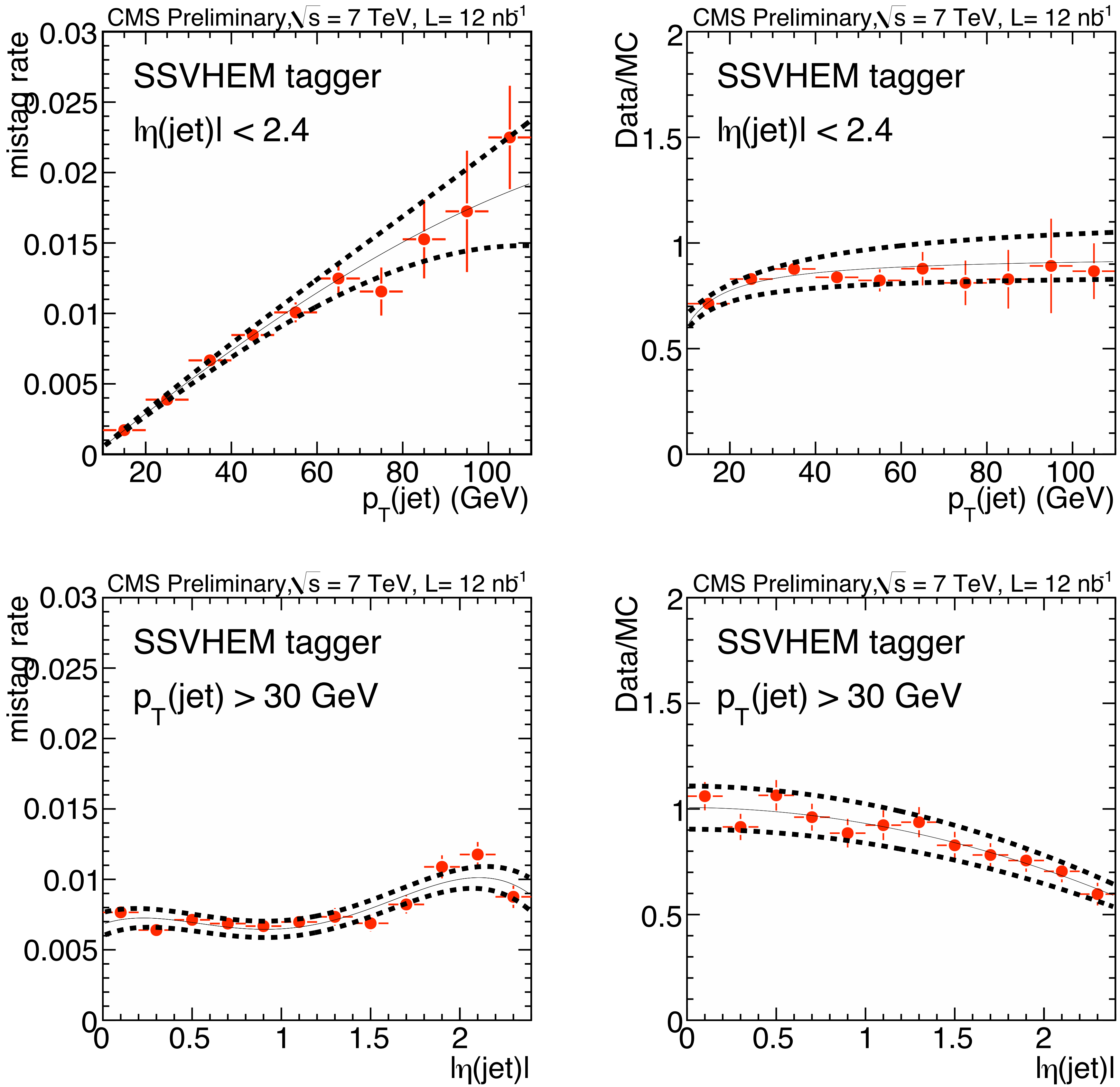


mistag rate results

example:
"Secondary Vertex"
algorithm

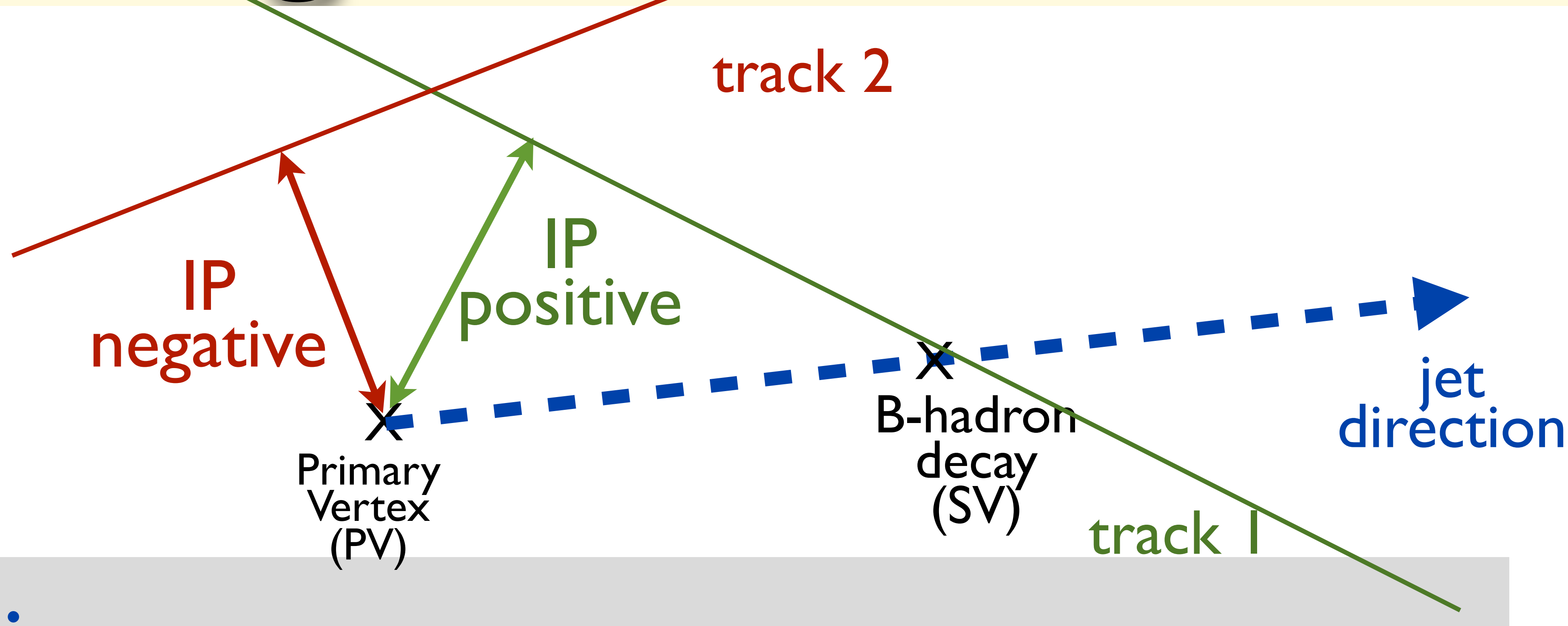
mistag rate and
scale factor as
function of jet p_t
and $|\eta|$

the dashed lines
represent statistical
and systematic
uncertainties



The Sign of the IP

the scalar product of
the IP segment
with the jet direction
determines the sign



in an ideal world:

- the IP distribution of light-flavour jets would be perfectly symmetric around 0 (and perfectly gaussian, because of various effects entering)

- the distribution would be mostly positive for b-jets

in reality, light jets are asymmetric and b-jets have negative IPs

negative tags are important for various applications, e.g. measurement of mis-tag rate (PAS BTV-07-002)