

Les activités autour du Particle Flow et du tau ID à l'IPHC

Jean-Laurent Agram, Anne-Fleur Barfuss, Auguste Besson,
David Bodin, Eric Conte, Mehdi Karim, Anne-Catherine Lebihan

Réunion CMS France – 27 & 28 mai 2009

Contenu

- Les interactions nucléaires (ACL, EC, AFB)
- Le HF Cal dans le PFlow (AB)
- L'identification des taus hadroniques (DB)
- L'étude des clusters (JLA, MK)
- Le Pflow et les CRAFT (JLA, EC)

Le Particle Flow (PF)

- x *Rappel des objectifs du Pflow : description globale des événements*
 - Fournir (aux analystes) une *liste de particules* reconstruites et identifiées (~ MC truth !)
 - *Définir* les jets, déterminer la MET, l'isolation,... de même que la reco
 - *Améliorations* wrt reco : JES (réduction des systématiques), résolution en énergie des jets <10% (100 GeV), tau-ID, MET scale et résolution,...
- x *Structure hits/clusters → blocs → candidats*
 - À partir de dépôts d'énergie et de traces (rechHits) → associés en *blocs*
 - Les blocs sont interprétés en tant que particules : *PFCandidates*

Le Particle Flow (PF)

x Inefficacité de la reconstruction due à :

- *Interactions annexes* (radiations d'électrons, conversions de photons, interactions nucléaires) : quantité de matière élevée devant le calorimètre
- *Performances des sous-détecteurs* (granularité/résolution en énergie insuffisantes, Hcal non compensant et non-linéaire)

⇒ perte d'énergie, double comptage

Interactions nucléaires

x *Personnes impliquées* : Anne-Catherine Lebihan, Eric Conte, AFB

x *Contexte* :

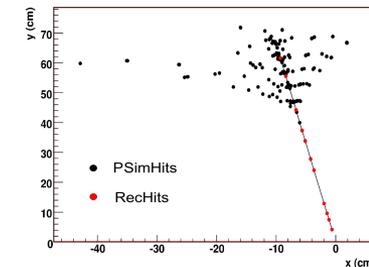
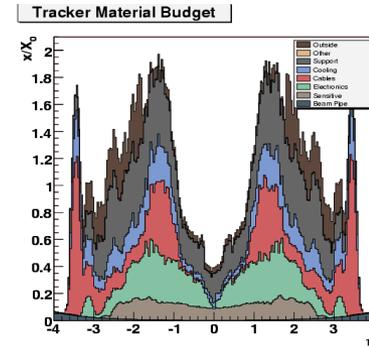
- Le budget matière est suffisamment important pour qu'entre 5 et 15% des pions interagissent avec le trajectographe
- Implications sur le *tracking* et la *reconstruction* dans le PF

x *Dans le PF* :

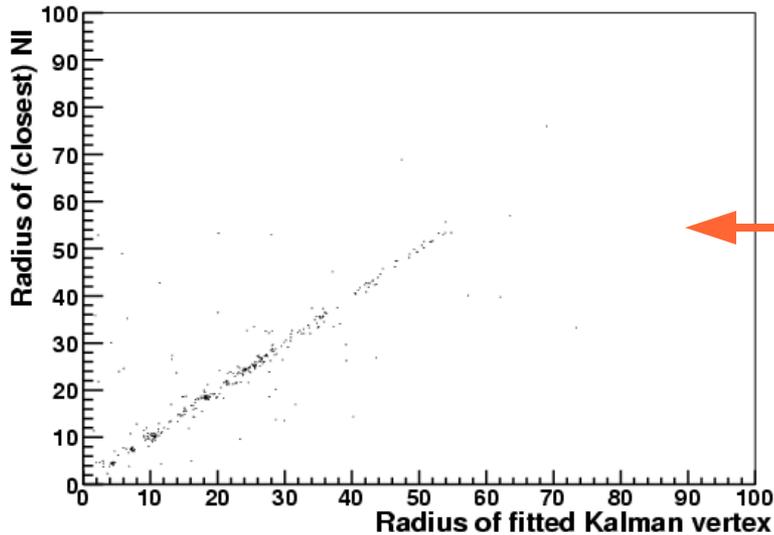
- Reconstruire les traces venant de NI avec une efficacité décente est possible, mais les inclure directement dans le PF induirait un *double comptage de l'énergie* (reco traces 1^{res} et 2^{res})
⇒ *vertexing et refitting des traces secondaires*
⇒ *choix de reconstruction des 1^{res} ou des 2^{res} à implémenter*

x *Travail en cours* :

- *Simulation* : single et di- pions 10 GeV
- *Sélection* des traces secondaires ajustées
- *Comparaison* de différents ajustements

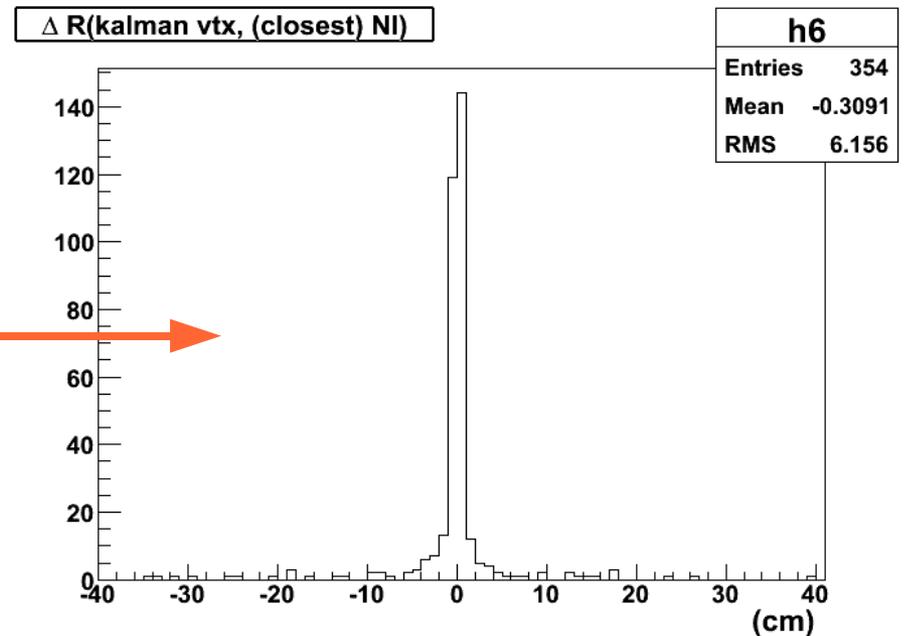


Interactions nucléaires (2)



Corrélation entre rayon du vertex ajusté et rayon de l'interaction nucléaire la + proche

Distance ΔR entre le vertex ajusté et l'interaction nucléaire la + proche



Le HF Cal dans le PF

x *Personne impliquée : Auguste Besson*

x *Buts, stratégie :*

- *Buts* : exploiter l'info. du HF pour le PF (et MET), utiliser la *profondeur* des fibres courtes/longues
- Avant cette étude, pas de clustering dans le HF
1 rec hit \Rightarrow 1 cluster \Rightarrow 1 HF PFCandidate
- *Clustering sur 2 "couches" en séparant EM et HAD*
rec hits EM/HAD \Rightarrow 1 cluster EM/HAD \Rightarrow PFCandidates

x *Tests sur les paramètres du clustering*

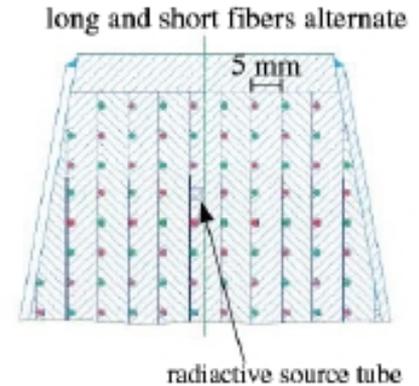
- Simulation (K_L^0 et gamma), nombre de voisins (4, 8 ?), seuils

x *Calibration*

- K_L^0 et gamma de 1 GeV à 2 TeV, $2.8 > \eta > 5.2$

x *État d'avancement*

- Calibration : $E_{\text{rec}} = a E_{\text{em}} + b E_{\text{had}}$, sélective (gamma, h, h\em), bins en eta
- *Intégré à CMSSW 3_1_0_pre6*



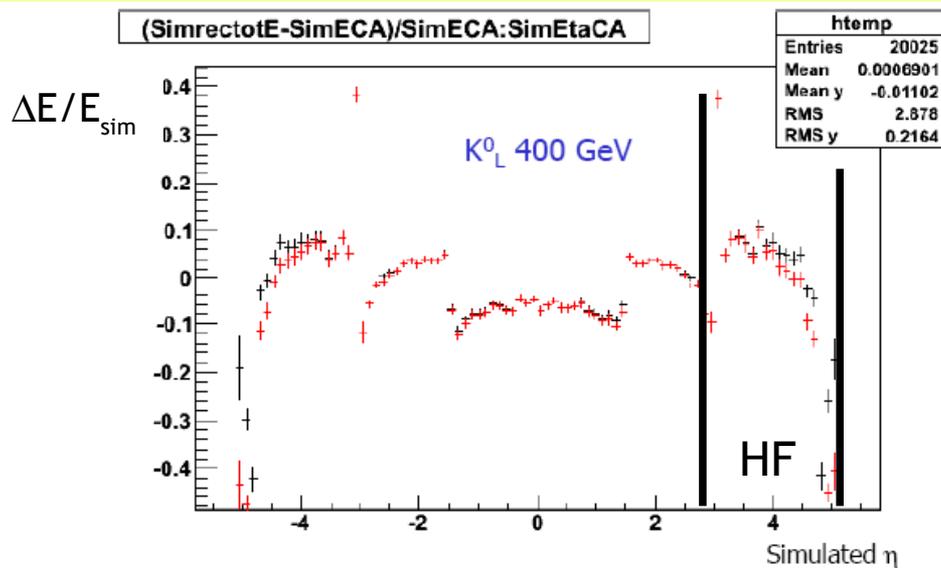
$$E = E_L + E_S$$

$$E_{\text{em}} \sim E_L - E_S$$

$$E_{\text{had}} \sim E - E_{\text{em}} \sim 2 E_S$$

Le HF Cal dans le PF (2)

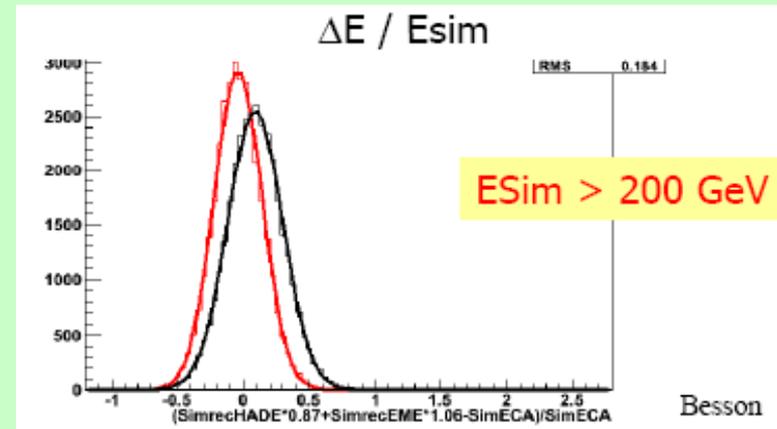
Test du clustering



Noir = E totale des clusters dans un cône
 Rouge = somme de l'E des 2 clusters les plus énergétiques (1EM+1HAD)

Calibration

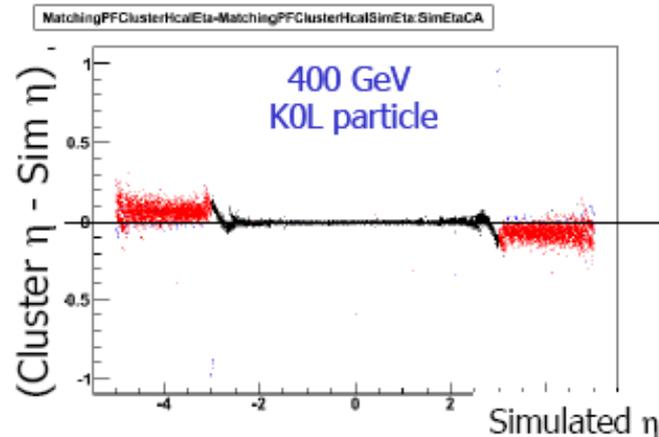
Black = Uncalibrated
 Red = calibrated



$\langle \rangle = 0.07$ $\langle \rangle = -0.03$
 RMS = 0.26 RMS = 0.23

x Perspective

Corriger la non-projectivité en eta



L'identification des taus had

x *Personnes impliquées : David Bodin*

x *Buts, stratégie :*

- *But* : Étude et optimisation de l'identification des jets de *taus hadroniques*, utilisation du Particle Flow dans l'environnement de PAT (*PF2PAT*).

Étude selon le type d'événement ($Z \rightarrow \tau\tau$, QCD, SUSY)

- *Actuellement* : Identification à l'aide de cônes signal et isolation centrés sur le constituant de plus haute énergie transverse

⇒ Utilisation d'un cône signal réductible (shrinking cone) pour les h^\pm

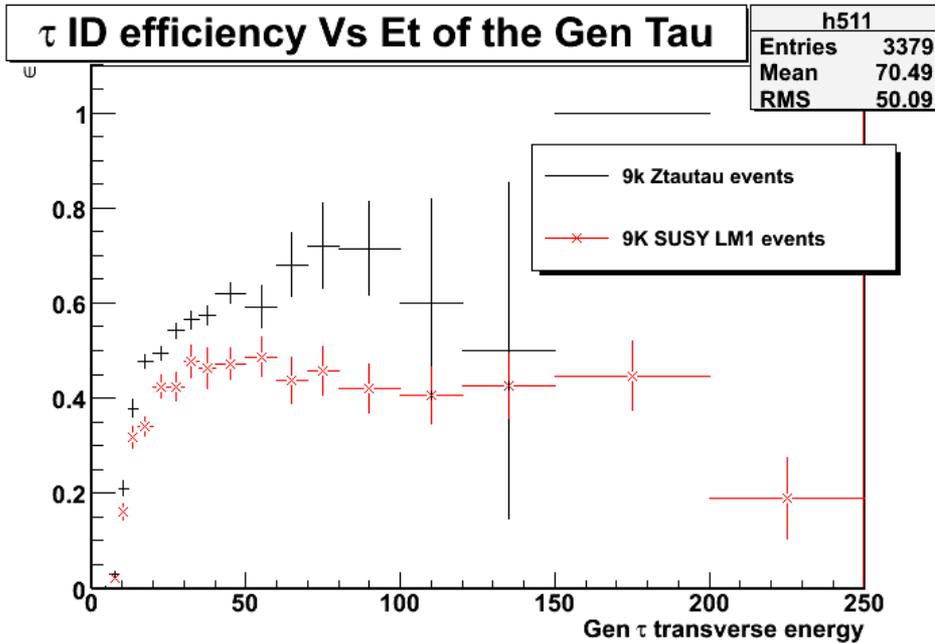
⇒ Utilisation d'un cône signal fixe pour les photons venant de $\pi\pi^0$

x *Travail en cours*

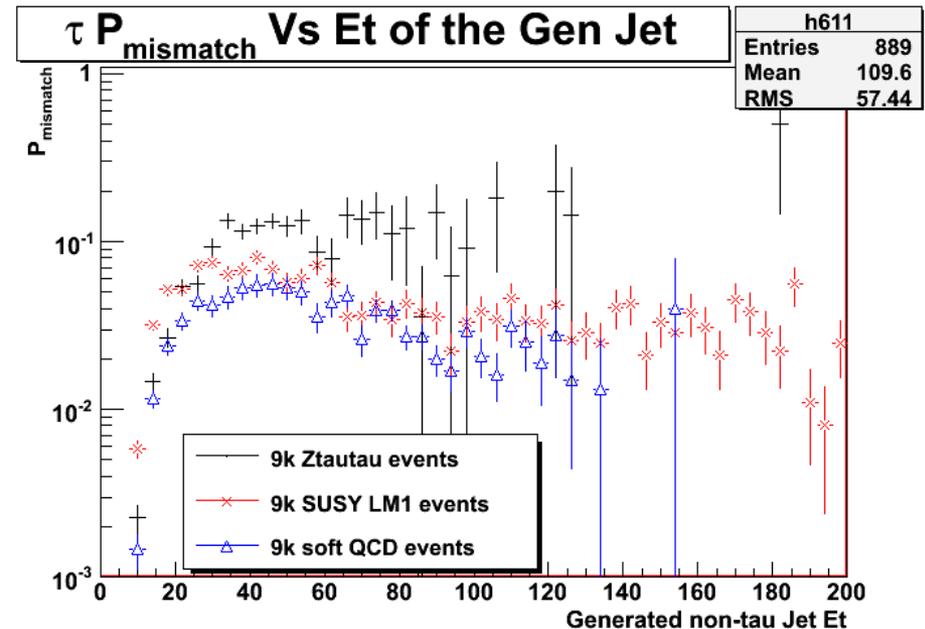
- Test de jeux de paramètres pour trouver l'identification optimale

- Essai du cône évolutif sur les photons

L'identification des taus had (2)



Efficacité de tau-ID =
 $E_T(\text{reco tau}) / E_T(\text{gen tau})$



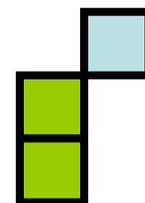
Mismatch =
 $E_T(\text{reco tau}) / E_T(\text{gen jet} \neq \text{taujet})$

Étude des clusters

x *Personnes impliquées* : Jean-Laurent Agram, Mehdi Karim (stage M2)

x *Contexte* :

- Cruzet : dépôts de faible énergie (μ) \Rightarrow toute l'E pas agglomérée
- Peut être amélioré en ajoutant *les cellules connectées par les angles*
- Événements majoritaires : particules de basse énergie
 \Rightarrow *l'amélioration pourrait être sensible.*



x *Dans le cadre d'un stage de M2* :

- Étude des particules générées dans les *événements QCD*
- Étude des *dépôts d'énergie* (Ecal, Hcal) reconstruits dans le Pflow pour chaque type de particules
- Étude de l'effet de *modifications dans l'algorithme de clustering*, dont l'ajout des cellules connectées par l'angle

x *Occasion* de lire en détail le code du clustering, de corriger des bugs

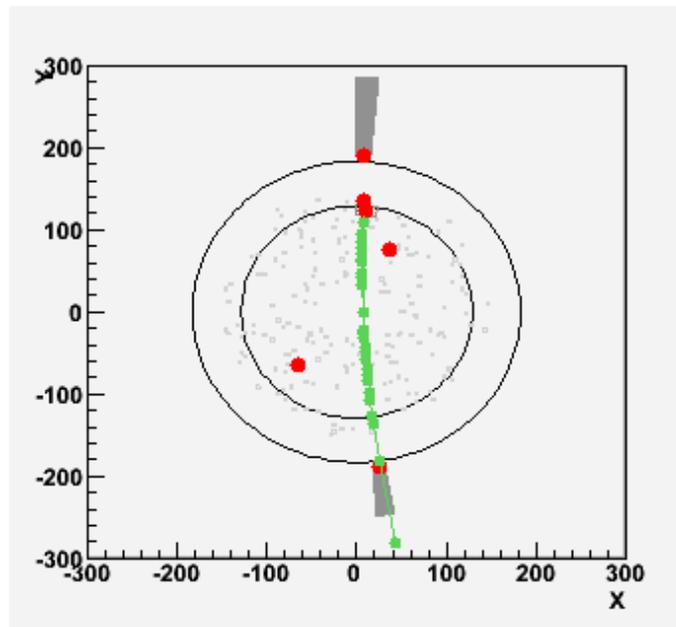
Le PF et les CRAFT

- x *Personnes impliquées : Jean-Laurent Agram, Eric Conte*
- x *Contexte : suite d'une étude des données de Cruzet en 2008, centrée sur l'ajustement des seuils (cellules, seeds) pour le clustering*
- x *Motivations :*
 - *Tester les algorithmes du PF avec des données et prévoir les changements de design de l'algorithme pour les premières collisions*
 - *Test et tuning des paramètres à chaque niveau de l'algorithme (éléments, blocs et candidats)*
 - *Comparaison avec les MC pour mettre en évidence d'éventuelles différences*
 - *Mesure d'efficacité de reconstruction*

Le PF et les CRAFT (2)

x Tâches en cours :

- *Adaptation de l'algorithme de PFlow aux CRAFT* : software CMSSW_2_2_9
- *Comparaison avec les MC* de la reconstruction de traces & clusters et de leurs associations (plots à venir)
- *Compréhension et tuning préliminaire* de l'algorithme de production des candidats



Traces & clusters au
sens du PFlow

Gris : RechHits
Vert : PFTrack
Rouge : PFCluster

Conclusion

- x Activités du groupe réparties sur différentes améliorations de l'algorithme et orientées vers leur validation sur des données
- x Une partie des contributions du groupe est déjà intégrée à CMSSW ou en voie de l'être

Backup slides

Some details on the HF

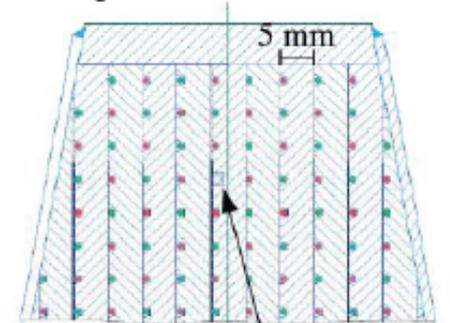
steel absorbers + quartz fibers + phototubes \Rightarrow Cherenkov light

Table 5.1: Sizes of the HCAL readout towers in η and ϕ as well as the segmentation in depth. The HF has a non-pointing geometry, and therefore the tower η ranges provided here correspond to $|z| = 11.2$ m. *The first two depth segments of tower 28 have a finer η segmentation, divided at $|\eta| = 2.688$.

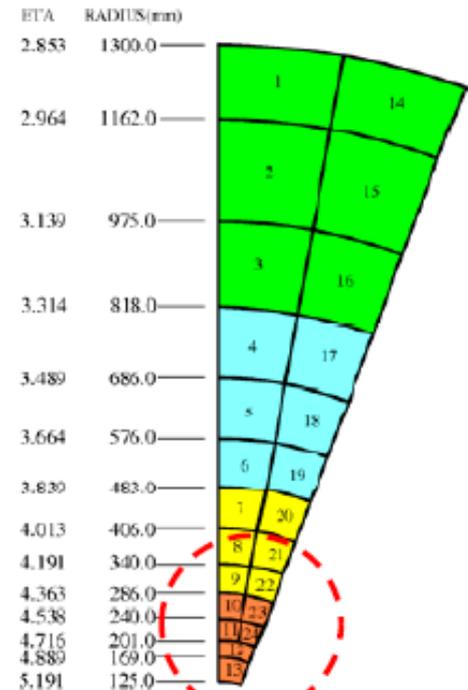
Tower index	η range		Detector	Size		Depth segments
	Low	High		η	ϕ	
27	2.500	2.650	HE	0.150	10°	HE=3
*28	2.650	3.000	HE	0.350	10°	HE=3
29	2.853	2.964	H ²	0.111	10°	HF=2
30	2.964	3.139	H ²	0.175	10°	HF=2
31	3.139	3.314	H ²	0.175	10°	HF=2
32	3.314	3.489	H ²	0.175	10°	HF=2
33	3.489	3.664	H ²	0.175	10°	HF=2
34	3.664	3.839	H ²	0.175	10°	HF=2
35	3.839	4.013	H ²	0.174	10°	HF=2
36	4.013	4.191	H ²	0.178	10°	HF=2
37	4.191	4.363	H ²	0.172	10°	HF=2
38	4.363	4.538	H ²	0.175	10°	HF=2
39	4.538	4.716	H ²	0.178	10°	HF=2
40	4.716	4.889	H ²	0.173	20°	HF=2
41	4.889	5.191	H ²	0.302	20°	HF=2

2.8 < η < 5.2
 \Rightarrow no tracks !

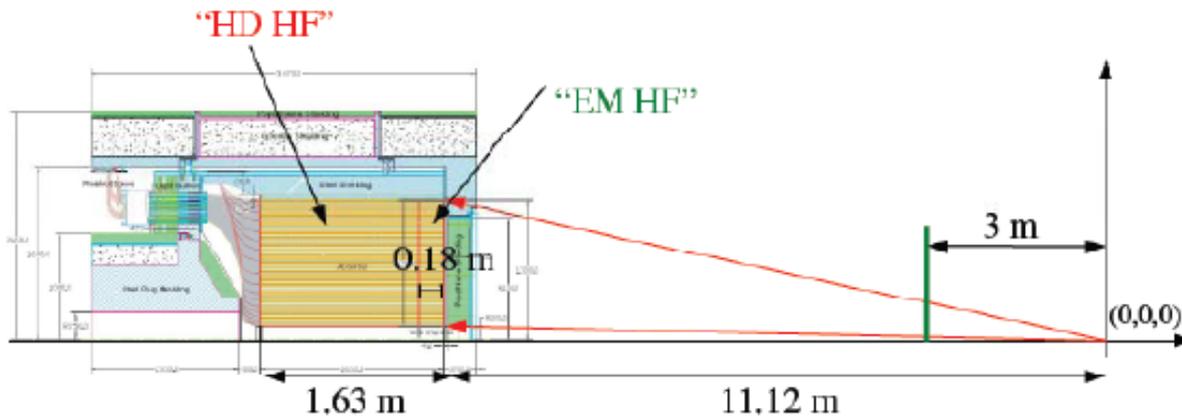
long and short fibers alternate



radiative source tube

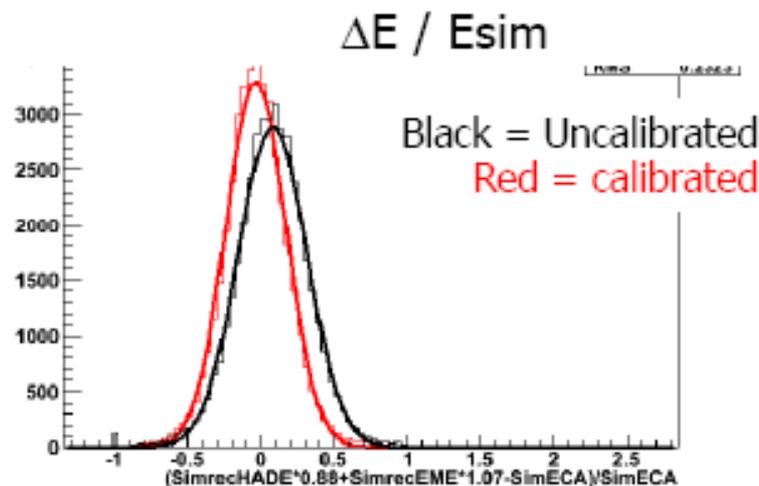


r- ϕ view @ z = 11.2 m



First minimization in the « flat eta region »

- K_L^0 1-1000 GeV in the 3.2-4.2 eta region



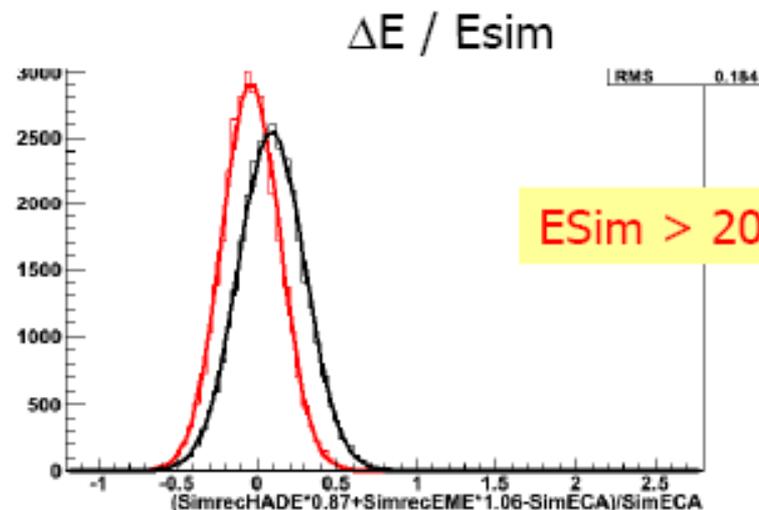
$\langle \rangle = -0.03$

RMS = 0.23

$\langle \rangle = 0.07$

RMS = 0.26

NAME	VALUE	ERROR
		calib
1 Constant	3.28 e+03	2.16 e+01
2 Mean	-3.56 e-02	1.02 e-03
3 Sigma	<u>2.06 e-01</u>	8.87 e-04
		uncalib
1 Constant	2.89 e+03	1.90 e+01
2 Mean	7.91 e-02	1.16 e-03
3 Sigma	<u>2.34 e-01</u>	1.01 e-03



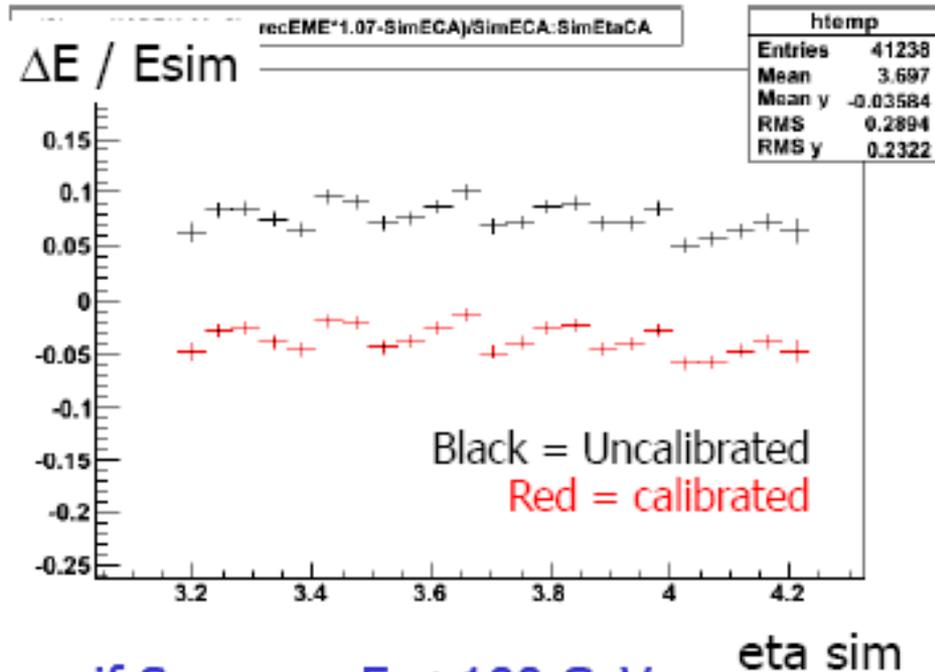
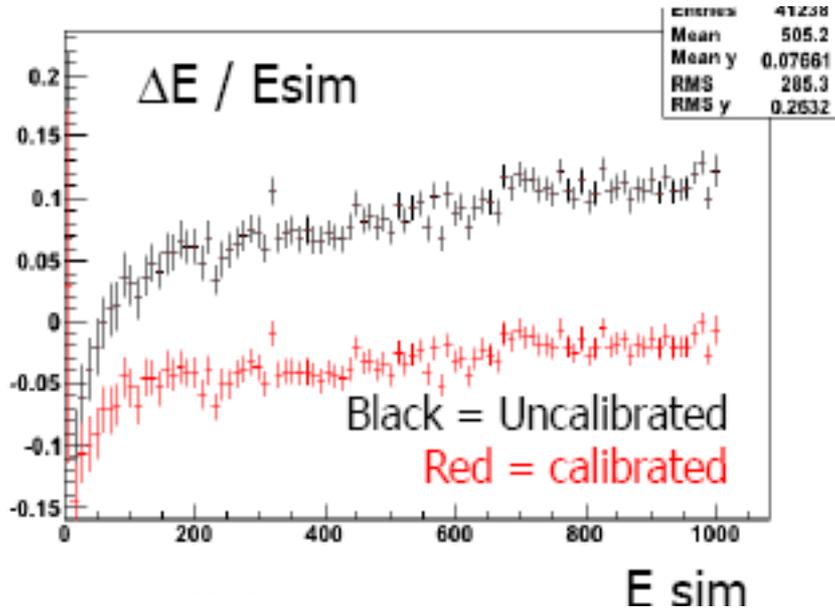
ESim > 200 GeV

NAME	VALUE	ERROR
		calib
1 Constant	2.91 e+03	1.99 e+01
2 Mean	-4.06 e-02	1.00 e-03
3 Sigma	<u>1.82 e-01</u>	7.55 e-04
		uncalib
1 Constant	2.54 e+03	1.74 e+01
2 Mean	8.99 e-02	1.15 e-03
3 Sigma	<u>2.08 e-01</u>	8.53 e-04

Besson

20

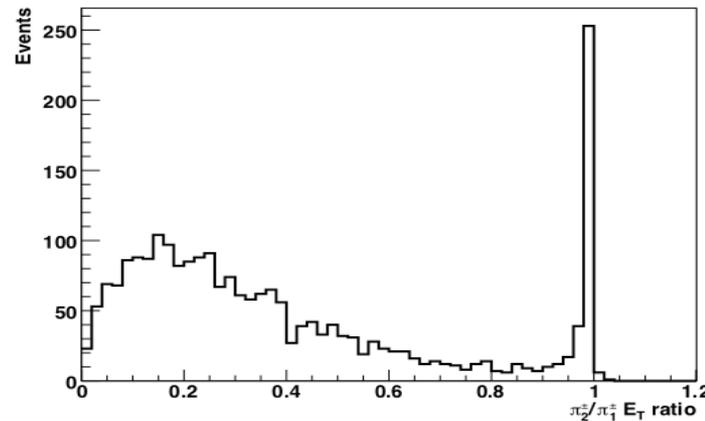
Calibration



- if Suppress $E < 100$ GeV
 - minimization found:
 $E_{rec} = 1.06 E_{em} + 0.87 E_{had}$
- $E = 100$ GeV; $\eta = 4 \Rightarrow p_T \sim 4$ GeV

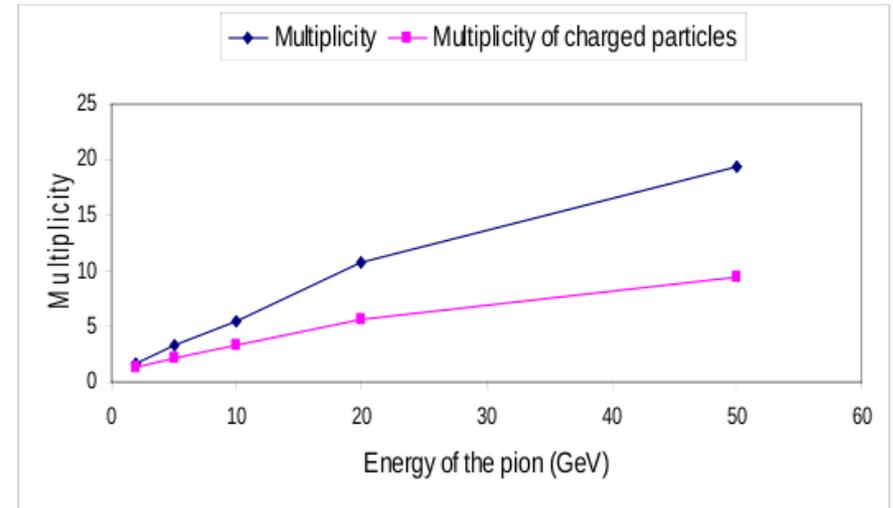
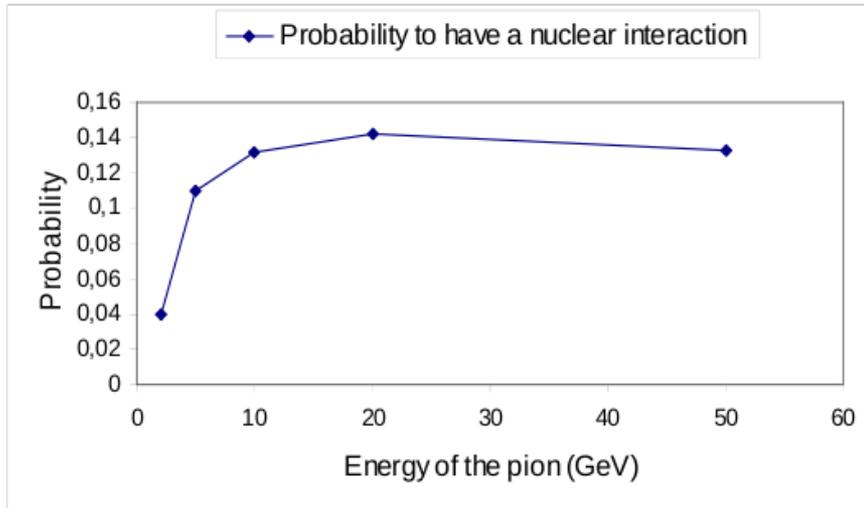
Elastic & inelastic nuclear interactions

- x ~15 % of nuclear interactions are inelastic ones, defined as ones with $E_T(\pi^\pm(2)/\pi^\pm(1)) < 0.9$, where $\pi^\pm(1)$ and $\pi^\pm(2)$ are respectively the primary and secondary pions*



- x The secondary particles consist of :*
 - 83 % - pions
 - 9.4% - protons
 - 6.1% - K
 - 1.5% - n, Σ ,...

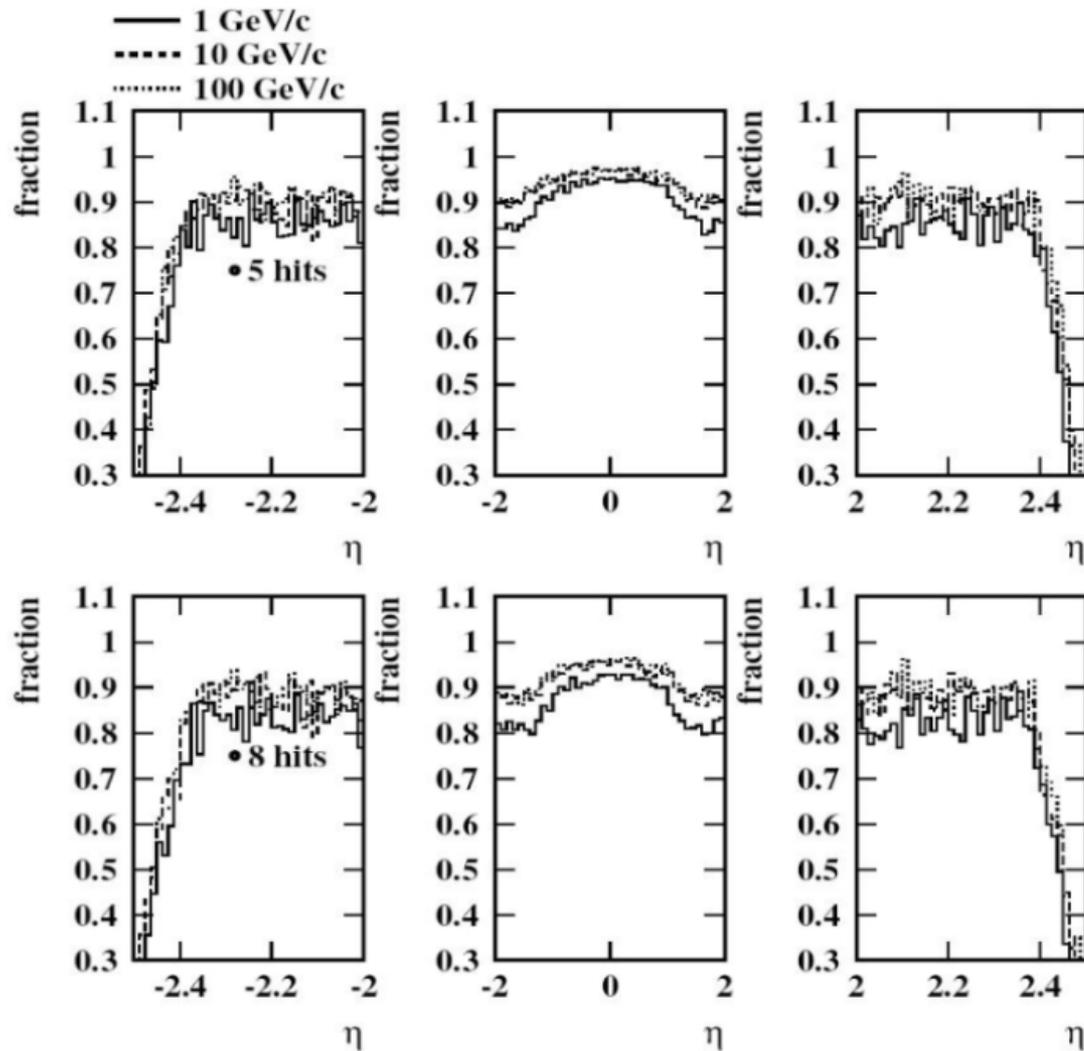
How often does a nuclear int. occur ?



→ the multiplicity of secondaries increases logarithmically with the energy of the pion

Effect of nuclear interactions on tracking

The pion track reconstruction efficiency varies between 80 to 95 %...

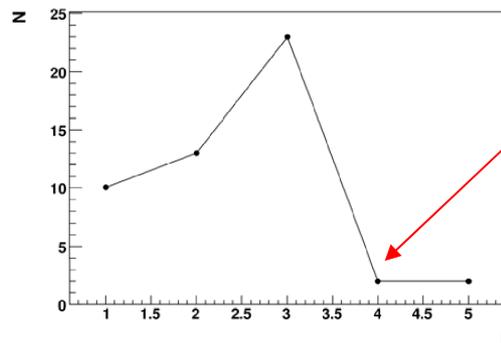


What has been developed so far...

Developed by Vincent Roberfroid

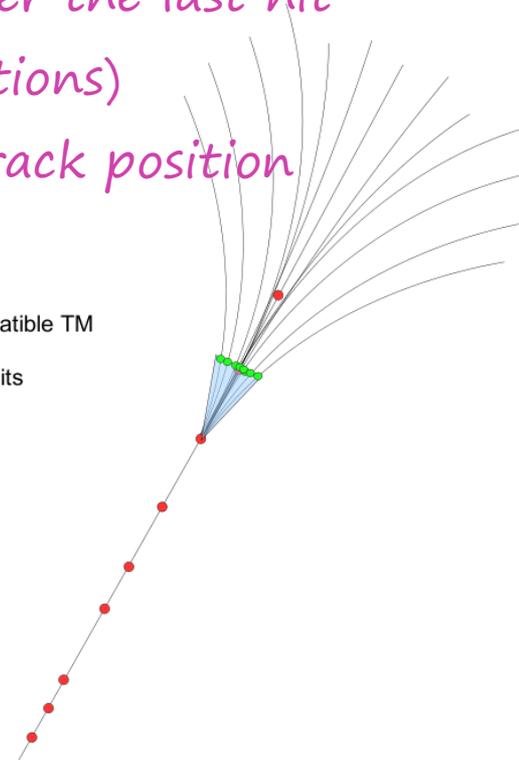
- First step : find the tracks compatible with a nuclear interaction in the tracker

- 1) All tracks with no hits in the last tracker layers
- 2) If $p > 5$ GeV check if the hit density increases after the last hit (in a 2 GeV cone to account for inelastic interactions)
- 3) Find the NI position, by finding the first inner track position with minimum hit density



Minimum of hit density

● compatible TM
● RecHits



What has been developed so far...

■ Second step : reconstruct secondary tracks

1) Build seeds starting from the last hit of the primary track

2) Keep those with :

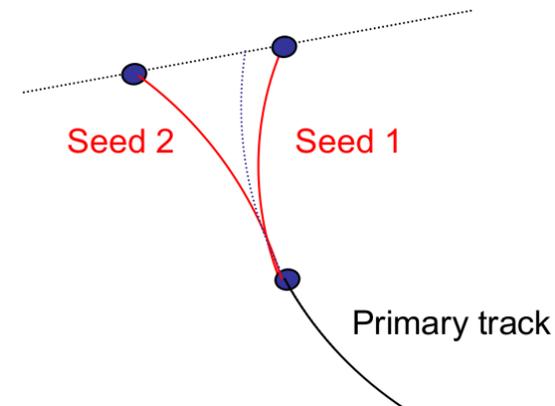
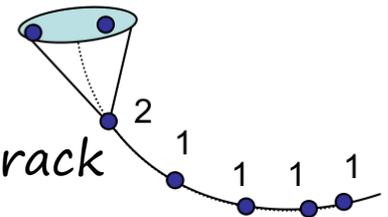
- 3 hits
- within a cone of $2 \text{ GeV}/p_1$ around the primary track
- $p > 0.3 \text{ GeV}$

3) Apply pattern recognition and keep :

- $\chi^2 < 30$
- $\Delta p/p < 0.3$
- $p_2 - 2\Delta(p_2) < p_1 + 2\Delta(p_1)$
- $d_{12} < 10 \text{ cm}$

4) Fit all tracks to a common vertex :

- Rank tracks from the best to the worst (number of hits and χ^2)
- If $\Sigma(p_2) < p_1$ remove the worst track, refit the vertex until $\Sigma(p_2) < p_1$



What has been developed so far...

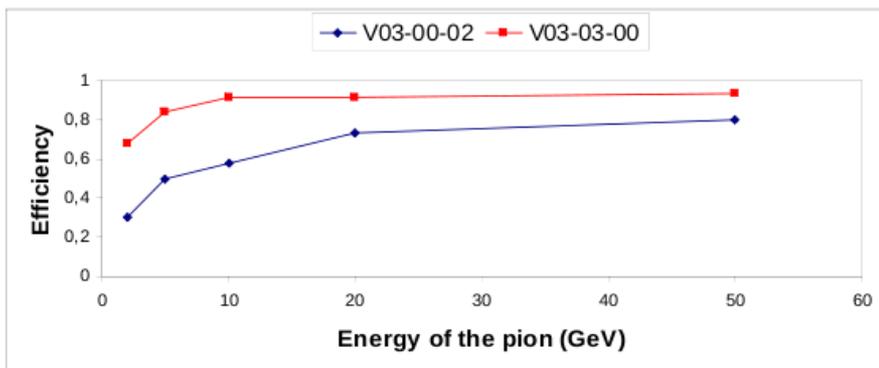
- Third step : implementation in PFlow

- 1) Propagate the secondary tracks to the calorimeter
- 2) Create links between secondary tracks and clusters
- 3) The primary gives rise to a PFCandidate
The clusters linked to the secondaries are removed

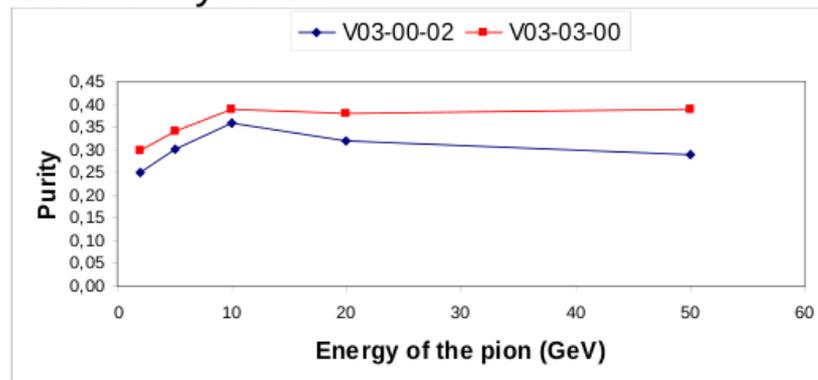
All code implemented in CMSSW and in good shape !

Some efficiency plots

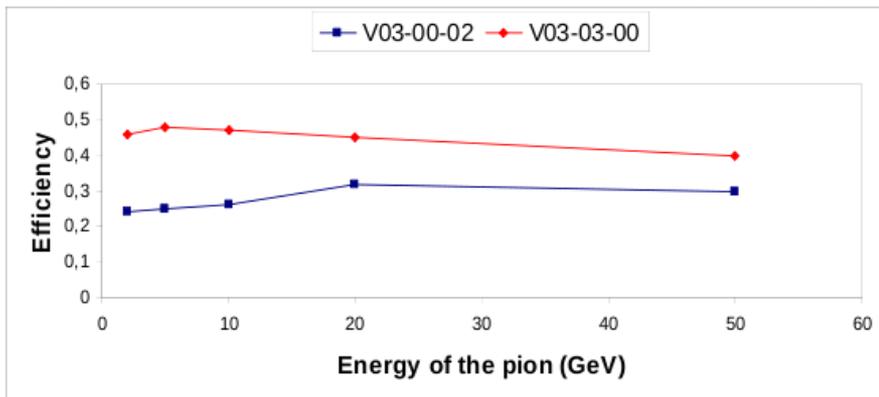
Efficiency to identify a nuclear interaction



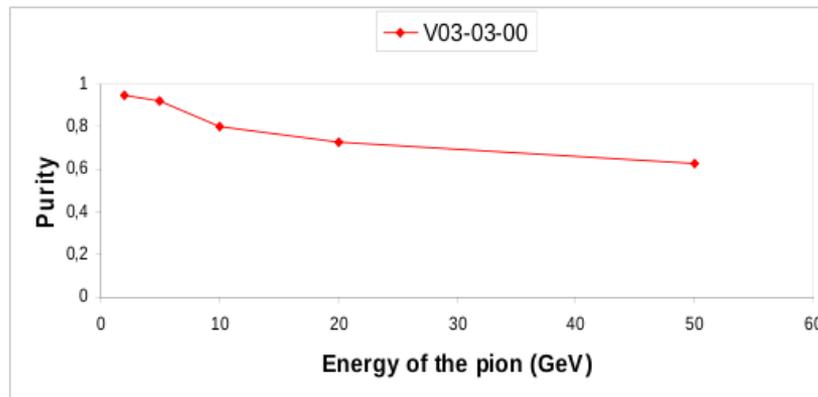
Percentage of seeds leading to a secondary track



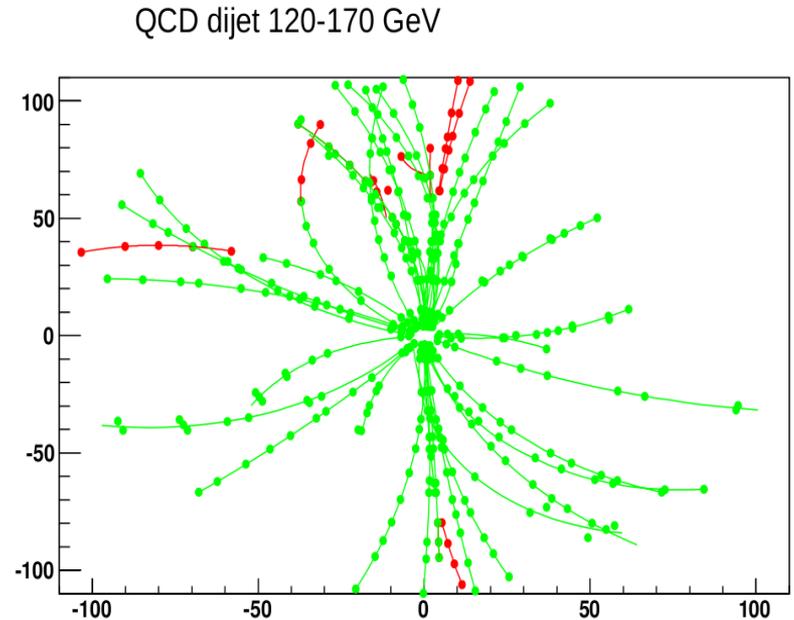
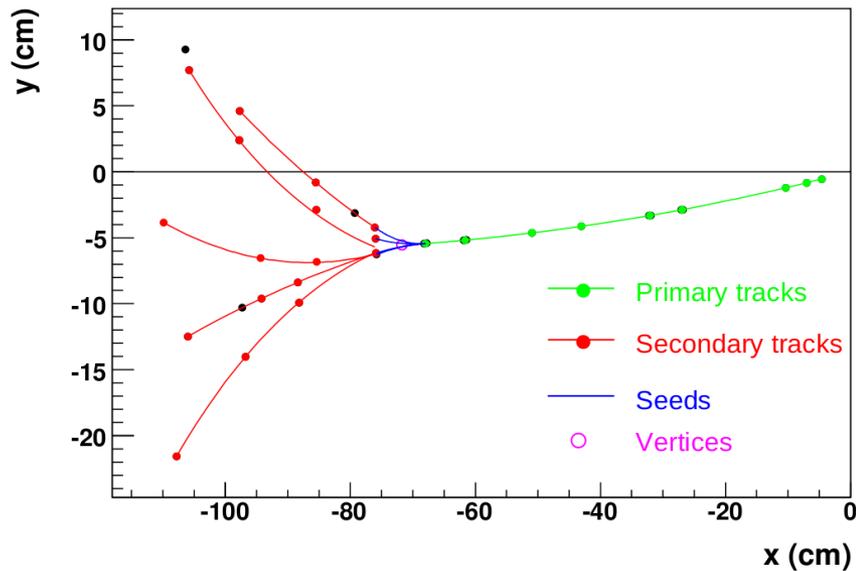
Efficiency to reconstruct secondary tracks (with $P_t > 1$ GeV and attached to a primary)



Percentage of secondary tracks which match a simTrack

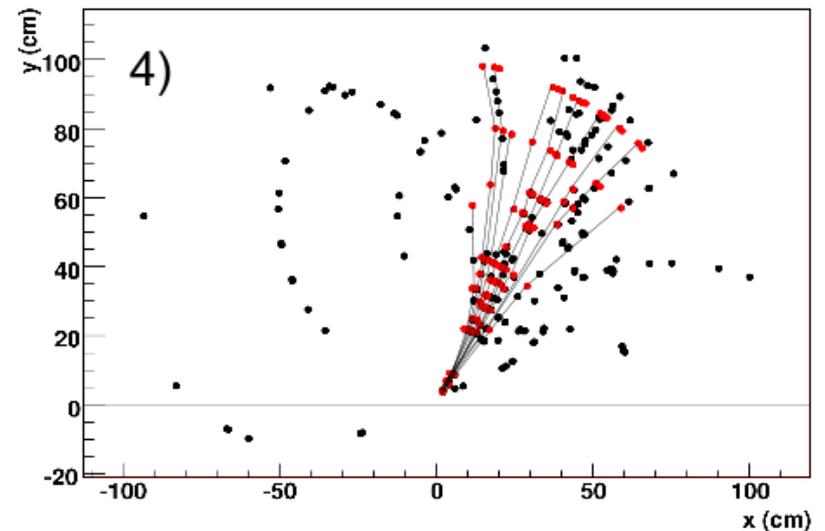
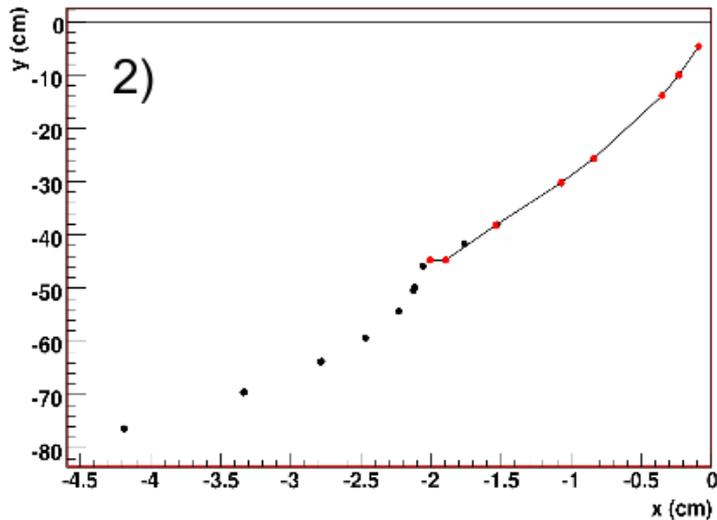


Some examples...



Expected inefficiencies

- Nuclear interaction in the outermost layer of the tracker
- Very elastic interactions
- Nuclear interaction close to the primary vertex
- Mixing of clusters

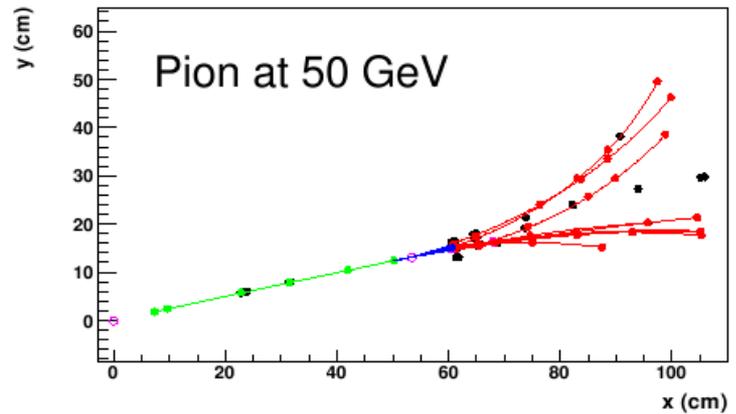
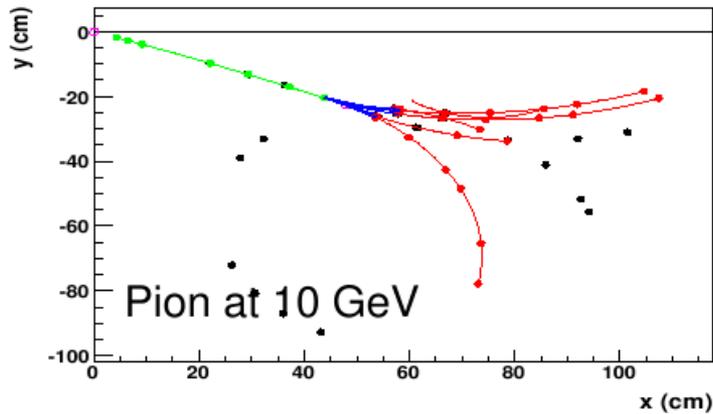
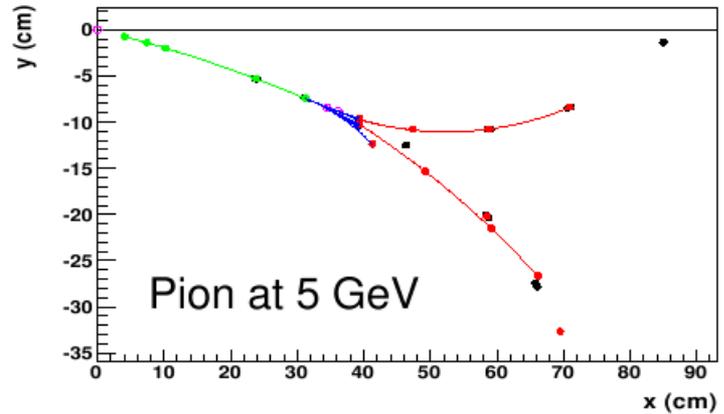
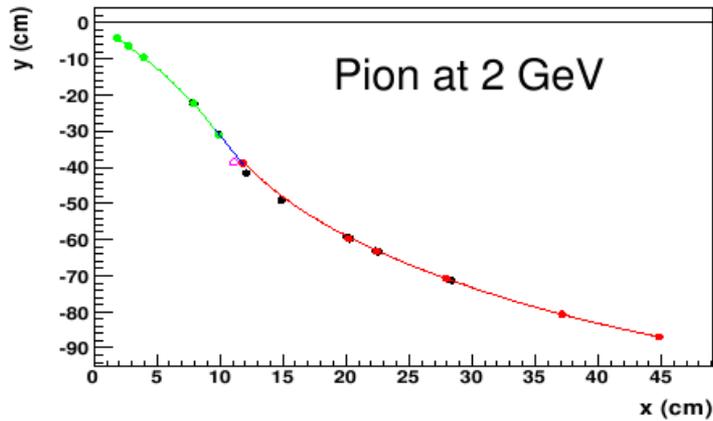


Prospects

(as defined during the PFlow workshop)

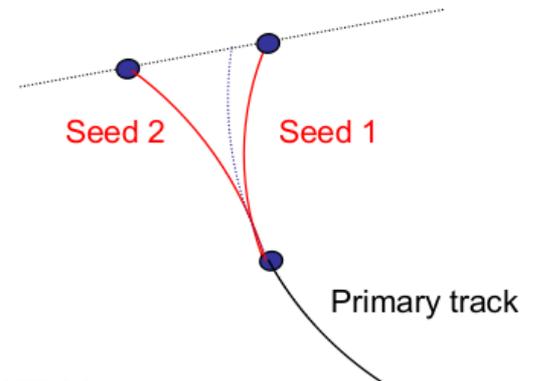
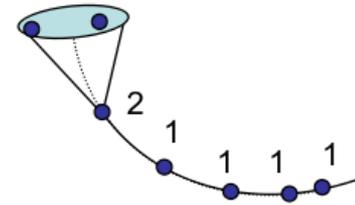
- Evaluate performances with latest tracking improvements / releases
- Refit the primary track (removing all hits found after the nuclear interaction)
- Consider tracks with out-in seeding, recover additional secondary tracks not tangent to the primary track...
- Linking between clusters and secondary tracks: use χ^2 of vertex and d_{12} as distance instead of default linking
- Look at di-jet events, hadronic taus, effect on resolutions ...

Some examples...



Brief description of the algorithm

1. The **NuclearSeedGenerator** module build for each Trajectory a vector of the number of compatible hits.
2. Using this vector, the algorithm identifies which trajectories are stopped due to a nuclear interaction (NI) and calculates roughly the position of this NI.
3. For each compatible hits, “nuclear seeds” are constructed assuming that the secondary tracks are emitted tangent to the primary track.
4. The standard CKF-CTF sequence is then used to reconstruct secondary tracks from these nuclear seeds.
5. The **NuclearInteractionEDProducer** produces NuclearInteractions (secondary vertices) from these secondary tracks using an Adaptive Vertex Fitter
6. Additional tracks (such as out-in seeding) can be added as input to the NuclearInteractionEDProducer. The compatibilities of each of those tracks with the NuclearInteractions in the event are checked.
7. The possible track candidates are merged with the tracks from the nuclear seeds using a method similar as the SimpleTrackListMerger from S. Wagner (+ speed up from P. Janot).
8. The NuclearInteractionProducer filters all the input tracks and recalculate the vertices.



NEW

Results : Efficiencies VS Pt

- E=50 GeV
- E=20 GeV
- E=10 GeV
- E=5 GeV
- E=2 GeV

