

# Recherche de nouvelle Physique grâce à l'identification des jets de saveurs $b$ au sein d'ATLAS



Nicolas Bousson  
Centre de Physique des Particules de Marseille

Directeur de these: Laurent Vacavant  
Co-directeur: Alexandre Rozanov



# Plan

## L'identification des jets $b$

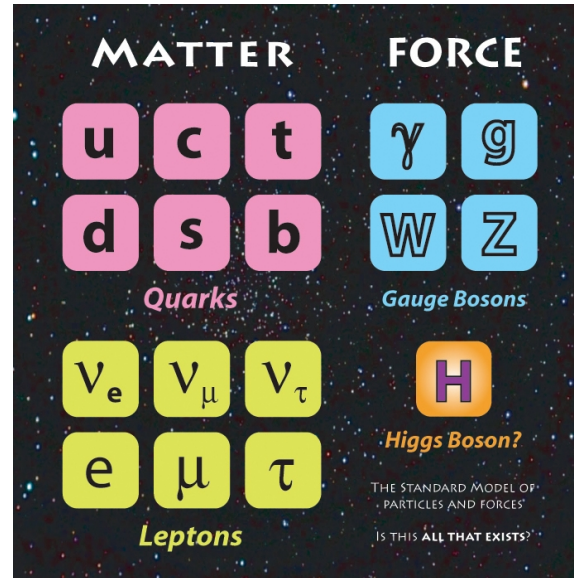
**Etiquetage dans les premières données**

**Les algorithmes les plus évolués**

**Vers les hautes luminosités du LHC**

# Introduction

Le 'Modele standard' de la Physique des Particules est une theorie couronnee de succes

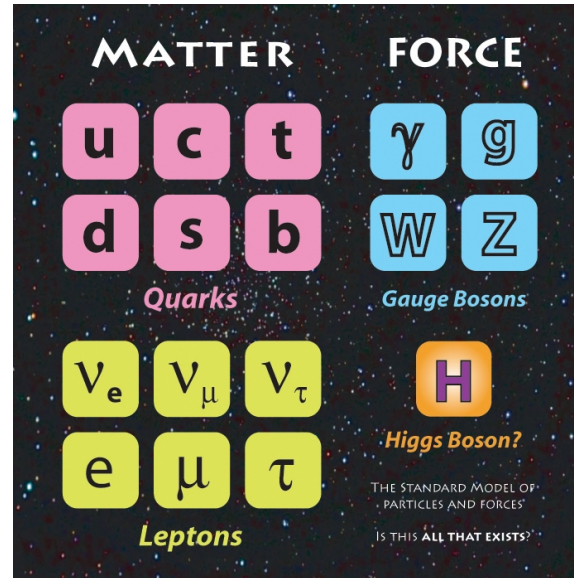


**SU(3)xSU(2)xU(1)**

Meme s'il n'est realistement qu'une approximation a basse energie d'une theorie plus 'globale'

# Introduction

**Le 'Modele standard' de la Physique des Particules est une theorie couronnee de succes**



**Mais comment expliquer:**

**Le probleme de hierarchie des masses**

**L'incompatibilite de la mecanique quantique avec la RG**

**Les observations cosmologiques**

**Etc etc ?**

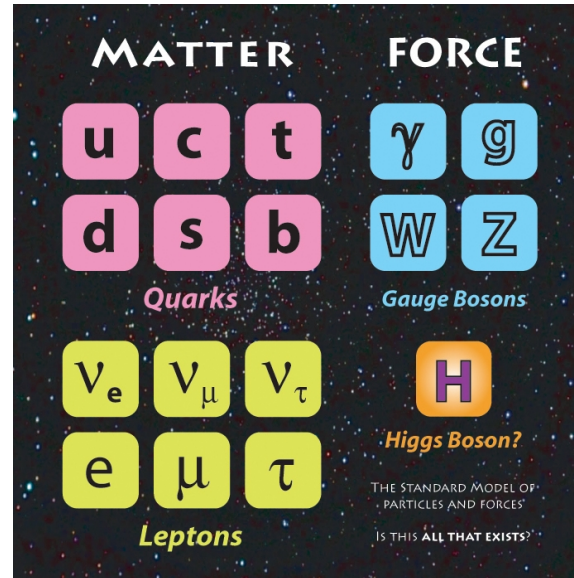
# Introduction

**Le 'Modele standard' de la Physique des Particules est une theorie couronnee de succes**

**Particules - cordes ?**

**Quantification de l'espace ?**

**Technicolor ?**



**Symetrie entre bosons et fermions (SUSY) ?**

**4<sup>eme</sup> generation ?**

**Extra-dimensions ?**

**→ Theories predisant l'existence de nombreuses particules a des energies encore non-atteintes !**

**Comment les detecter ?**

# Introduction

Le 'Large Hadron Collider'

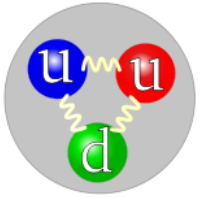
Le plus grand accélérateur de  
particules jamais construit

Collisions proton-proton à une  
énergie de 7 TeV

(14 TeV nominale)

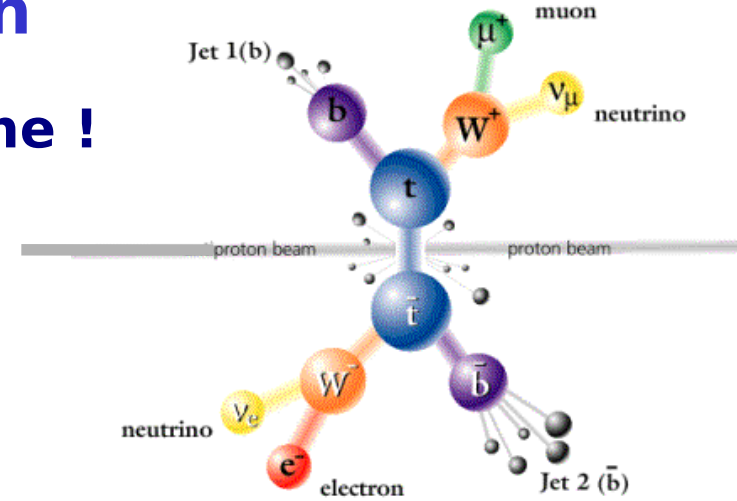
Tevatron ~ 1.96 TeV

Un croisement de faisceaux toutes  
les 25 ns

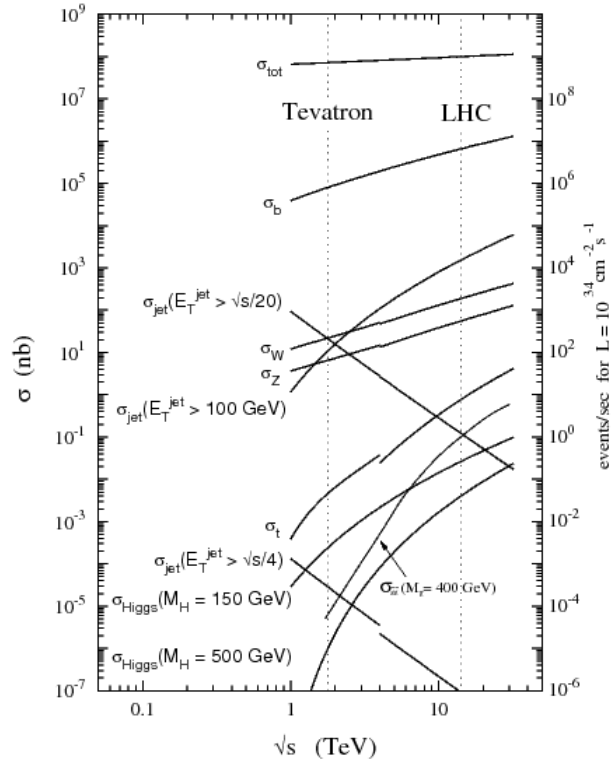


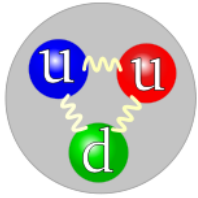
# Introduction

Phénoménologie très riche !



proton - (anti)proton cross sections

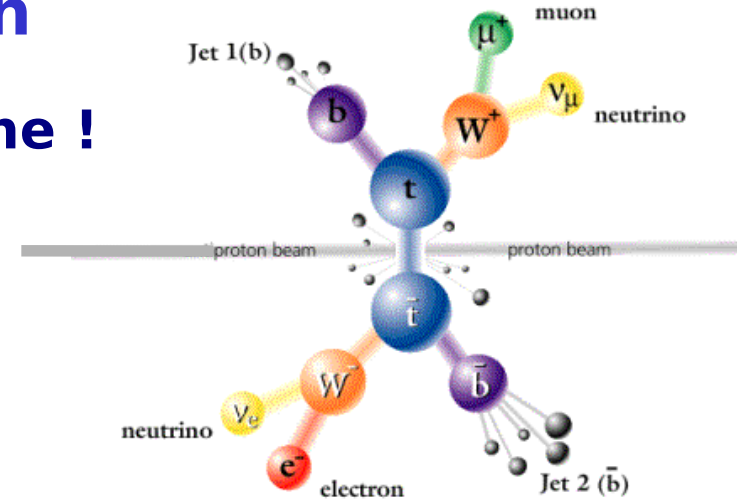




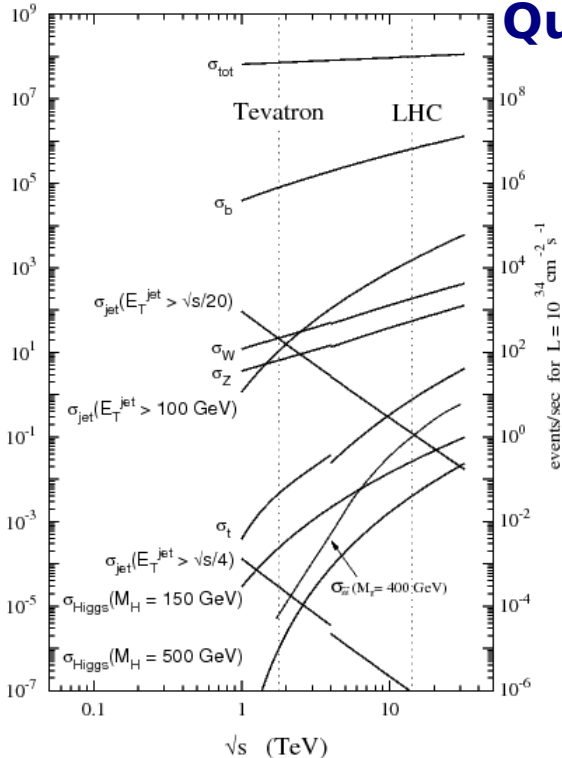
# Introduction

Phénoménologie très riche !

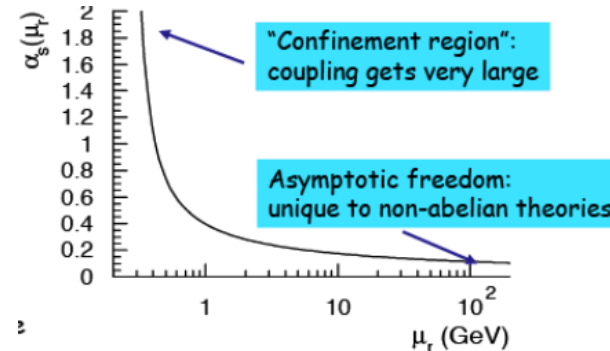
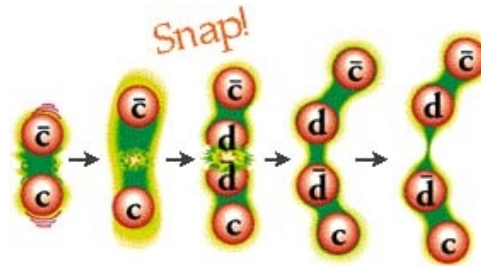
Mais fortement 'polluée' par certains processus d'interaction *forte* (QCD) entre les constituants des protons (quarks et gluons)



proton - (anti)proton cross sections



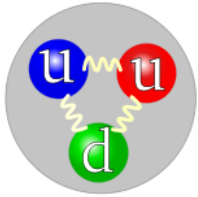
Quand 2 objets *couleurs* s'éloignent l'un de l'autre, l'intensité de l'interaction augmente !



Un quark créé lors d'une collision va:

- *fragmenter* en créant des paires q/q
- *S'hadroniser* (les quarks résultant se combinent en hadrons)

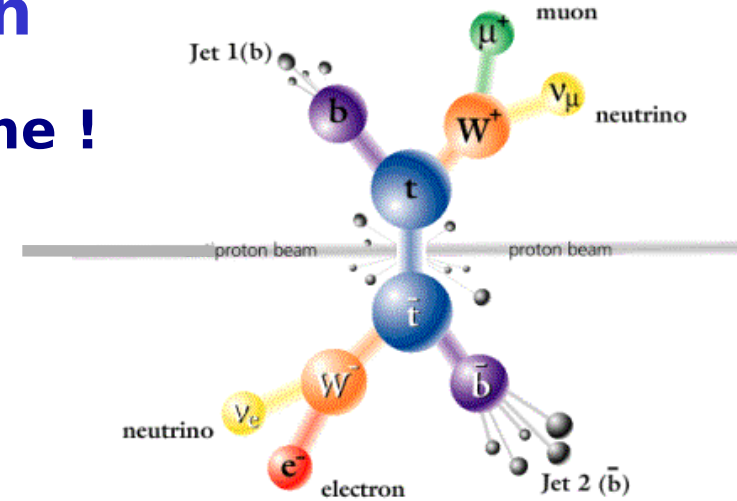




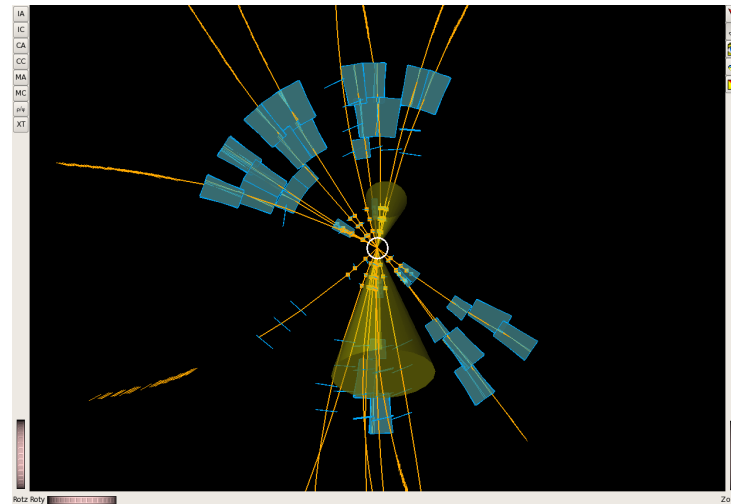
# Introduction

Phénoménologie très riche !

Mais fortement 'polluée' par certains processus d'interaction *forte* (QCD) entre les constituants des protons (quarks et gluons)

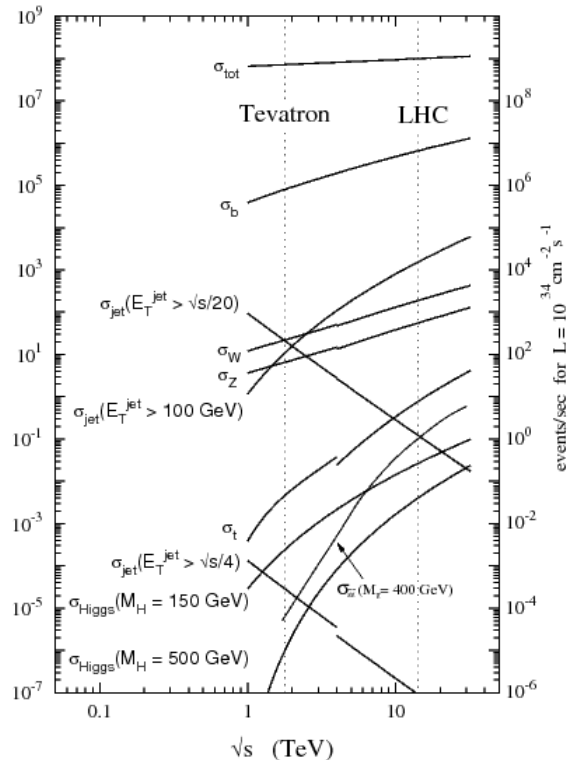


## Experimentalement:



**l'hadronisation de ces quarks crée des 'jets' de particules très localisées**

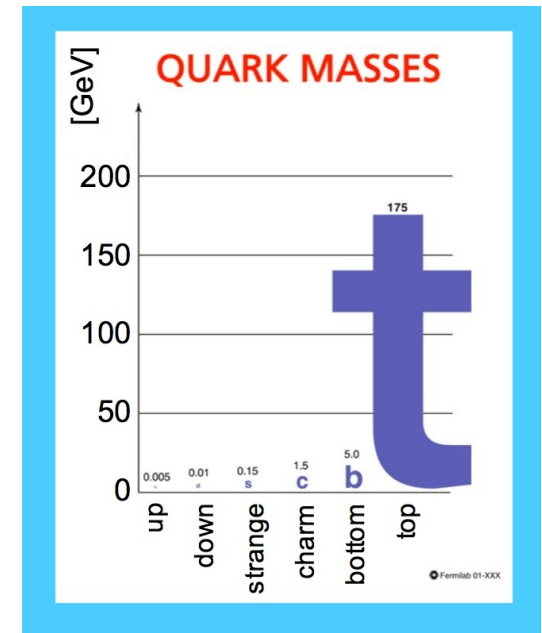
proton - (anti)proton cross sections



# Introduction

**Pourquoi chercher a identifier les *jets* issus originellement d'un quark b ?**

**→ car beaucoup d'hypothetiques particules tres massives peuvent se desintegrer en quark b ou t  
(B.R.  $t \rightarrow Wb \sim 100\%$ )**



**Identifier *le mieux possible* les jets b permettra donc:**

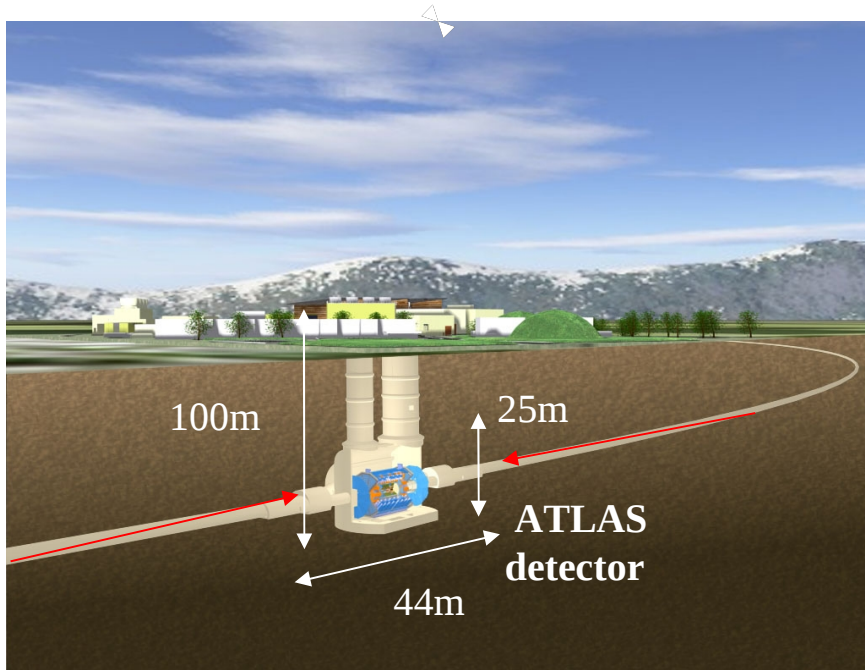
- **D'ameliorer les analyses de Physique (top,  $\sigma(bb)$ , boson de Higgs, SUSY, 4<sup>eme</sup> generation...)  
ex: Avec identification d'un jet b, ~5 fois moins de bruit de fond pour re-decouvrir le quark top**
- **De faire des decouvertes majeures plus rapidement !**

# Comment identifier les jets b ?

Le plus simplement, dans les 1<sup>ère</sup> données du LHC



# Le detecteur ATLAS

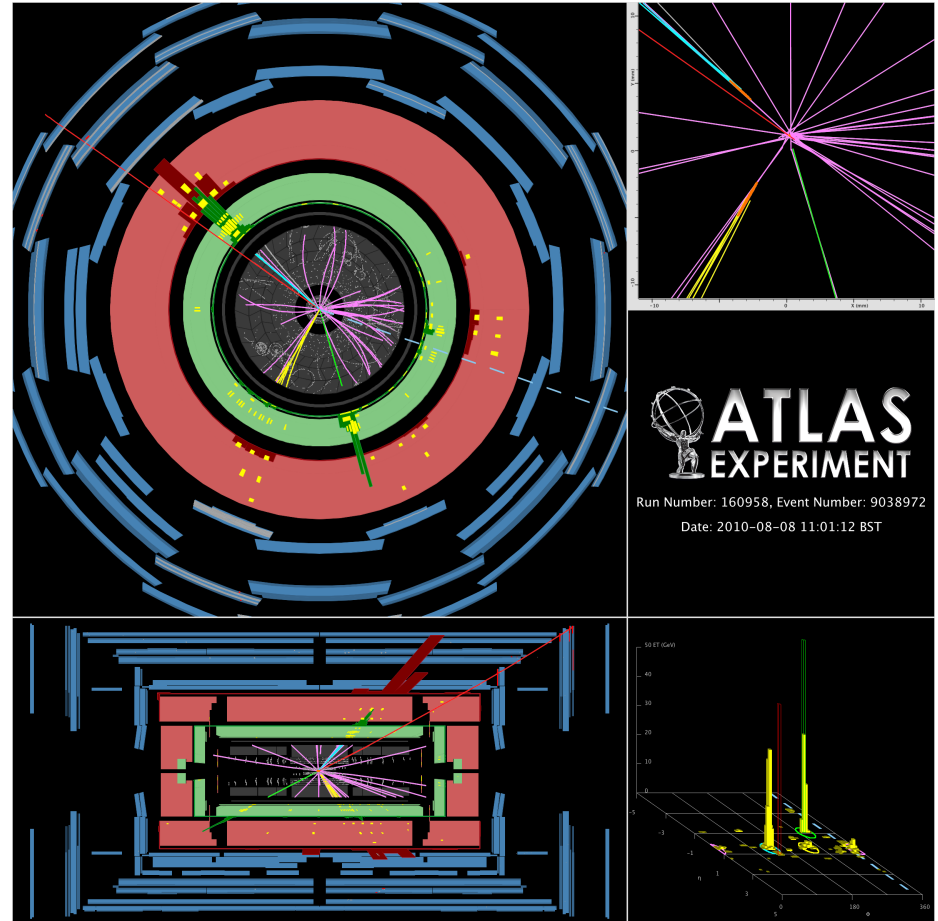


**Le plus grand detecteur au monde enregistrant les collisions du LHC**

**Il permet de reconstruire les traces des particules produites**

**De reconstituer l'energie des objets physiques**

**De reconnaître specifiquement les muons/electrons/photons...**



# Le detecteur a 'pixels' d'ATLAS

Situe au plus pres des collisions

80 millions de pixels

Detecte le passage des particules chargees

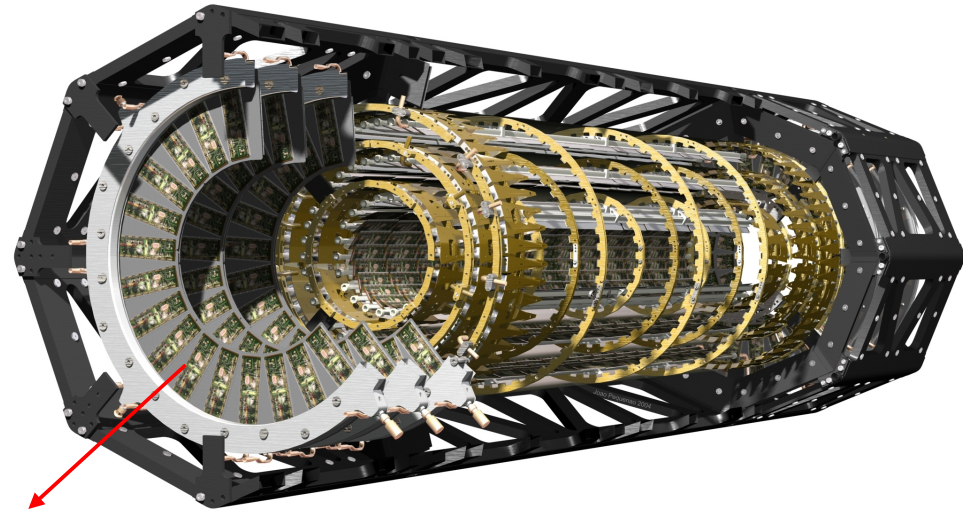
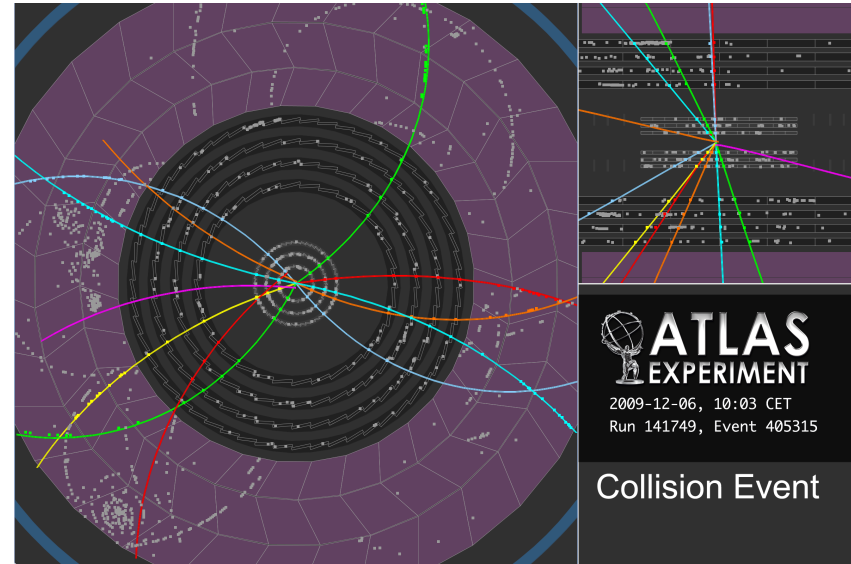
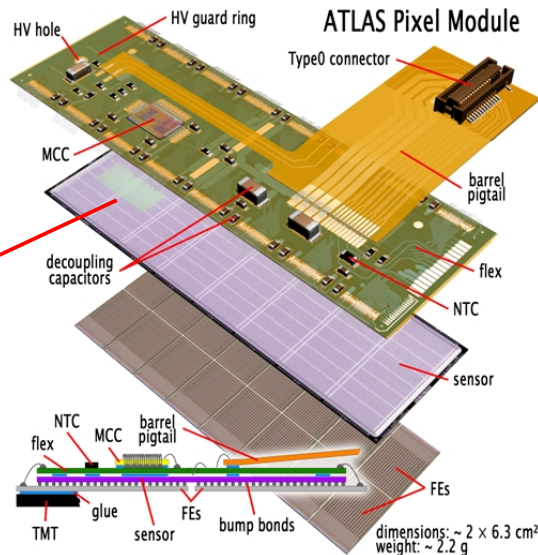
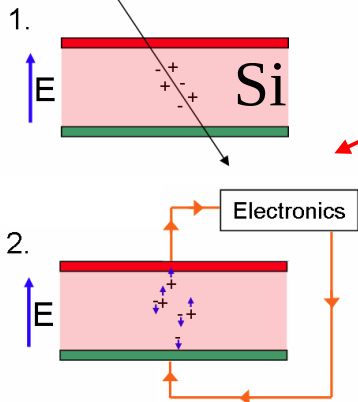


Schéma d'un des 1744 modules:

46080 pixels / module



<http://atlas.web.cern.ch/Atlas/public/EVTDISPLAY/events.html>

Taille d'un pixel:  $\sim 50\mu\text{m} \times 400\mu\text{m} \times 270\mu\text{m}$

→ Permet la reconstruction precise des traces et vertex

# 'Shifts' pixels

~20% de mon temps

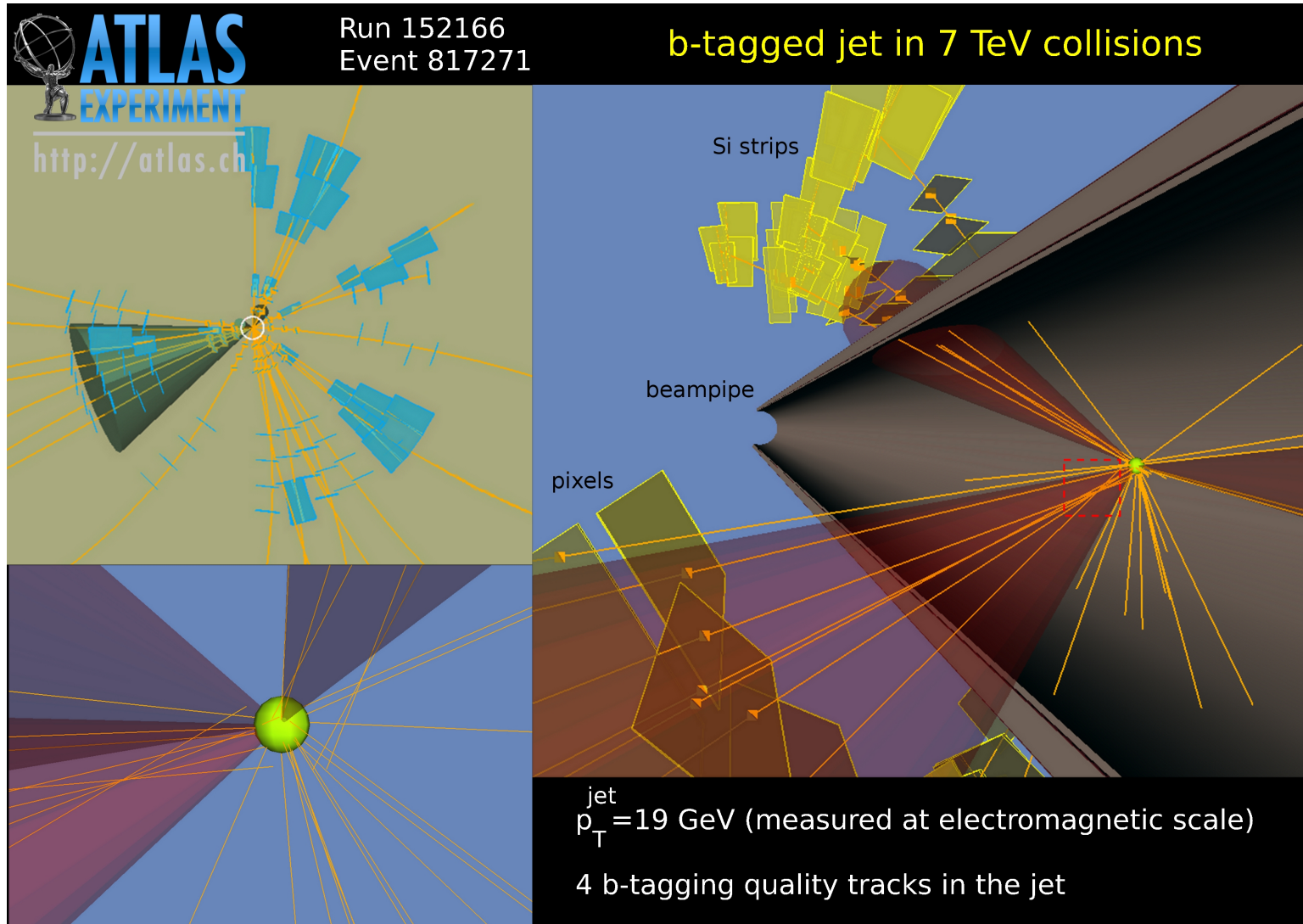
29 novembre 2009 : 1ere collisions a 900GeV

30 mars 2010 : 1ere collisions a 7TeV



Durant collisions: surveillance pendant 8h  
Sans faisceaux: **calibrations du detecteur** (configuration des modules / communication avec les modules / identification des pixels « bruyants »)

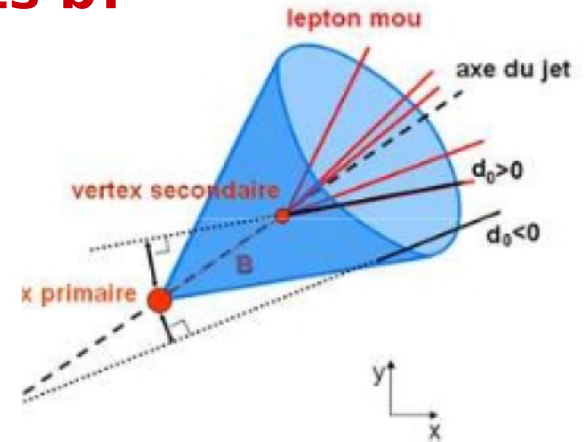
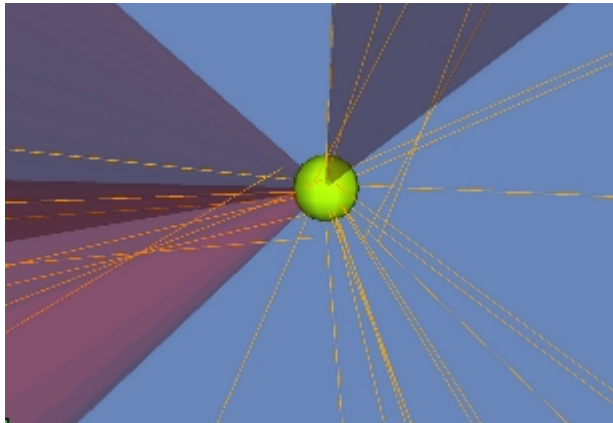
# Pourquoi le detecteur a pixels est crucial pour l'identification des jets b?



# Pourquoi le detecteur a pixels est crucial pour l'identification des jets b?

Car un hadron B conserve  $\sim 70\%$  de l'impulsion du quark b initial (5GeV):

- les traces issus de sa desintegration ont une *forte* impulsion transverse
- il 'vole' dans le detecteur du fait de sa *longue* duree de vie
- il se desintegre  $\sim 20\%$  du temps en lepton



Distance PV-SV  $\sim 5\text{mm}$   
pour un jet de  $p_T = 50\text{GeV}$

La resolution des pixels ( $\sim 15\mu\text{m}$ ) permet:

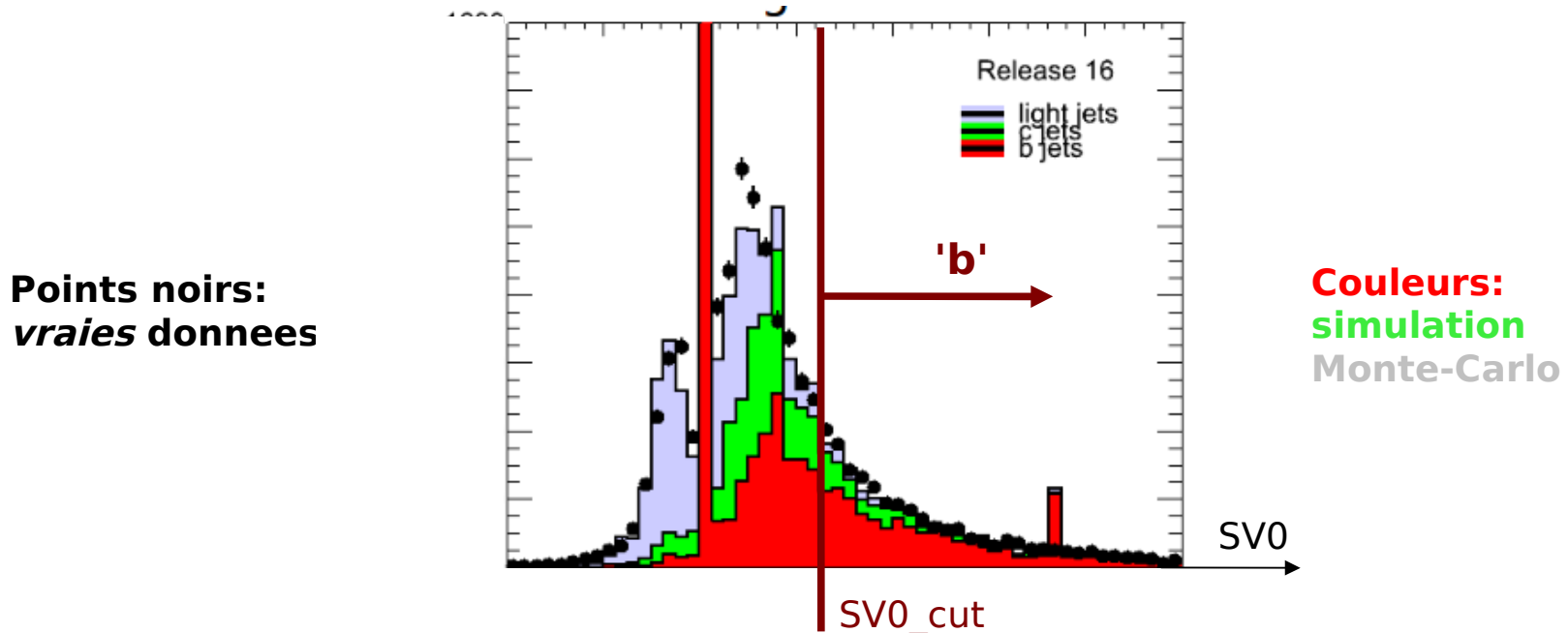
- La detection d'un vertex secondaire (SV) eventuel
- Une mesure precise du parametre d'impact ( $d_0$ ) des traces



# Etiquetage dans les 1<sup>eres</sup> donnees

## Comment construire une methode d'etiquetage simple?

**Ex1:** Regarder la distance entre le PV et le SV  
divisee par l'erreur de mesure (SV0):



**Si on decide d'etiqueter un jet comme b si son  $SV0 > SV0\_cut$  :**

**→ on se trompe un certain nombre de fois**

**→ mais on a raison tres souvent !**

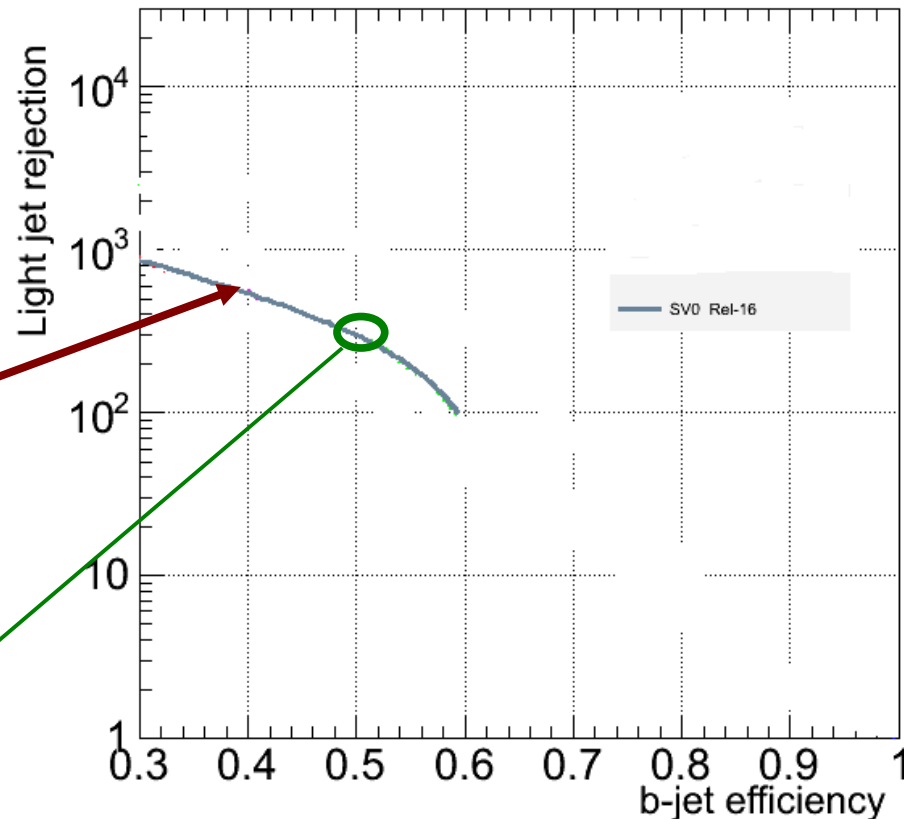
**Probleme statistique**

# Etiquetage dans les 1<sup>eres</sup> donnees

Inverse de la fraction des jets identifies incorrectement

On se deplace sur cette courbe en faisant varier la valeur de SV0\_cut

50% des jets b identifies correctement ; 1 jet mal identifie sur 300 !

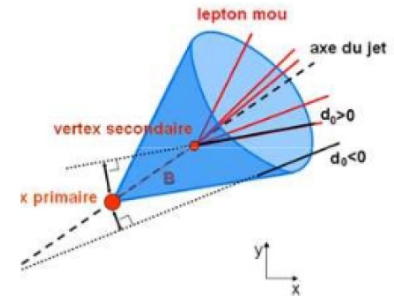
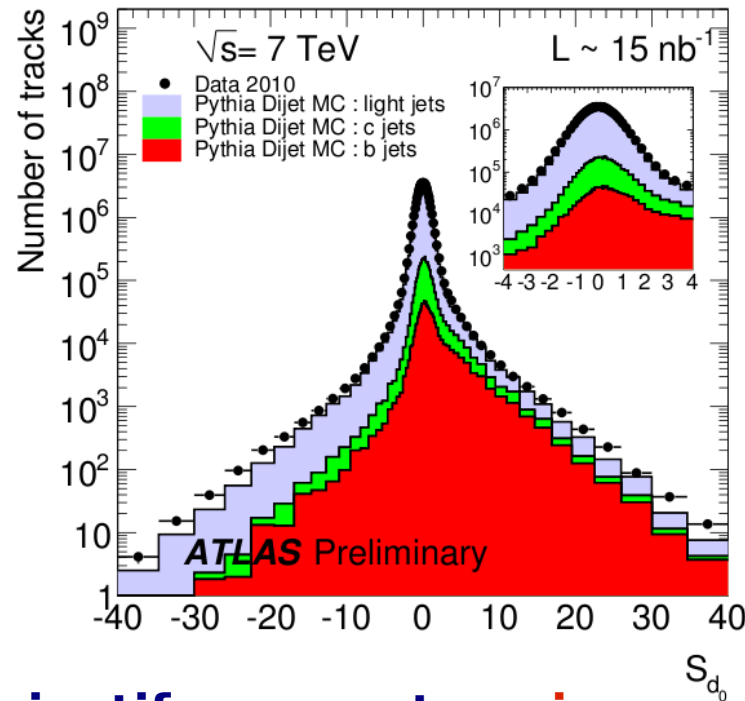


Fraction des jets b identifies correctement

# Etiquetage dans les 1<sup>eres</sup> donnees

**Comment construire une methode d'etiquetage simple?**

**Ex2: Utiliser le parametre d'impact signé des traces associees au jet (et diviser par l'erreur de mesure):**

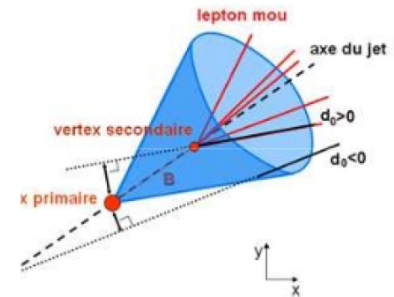
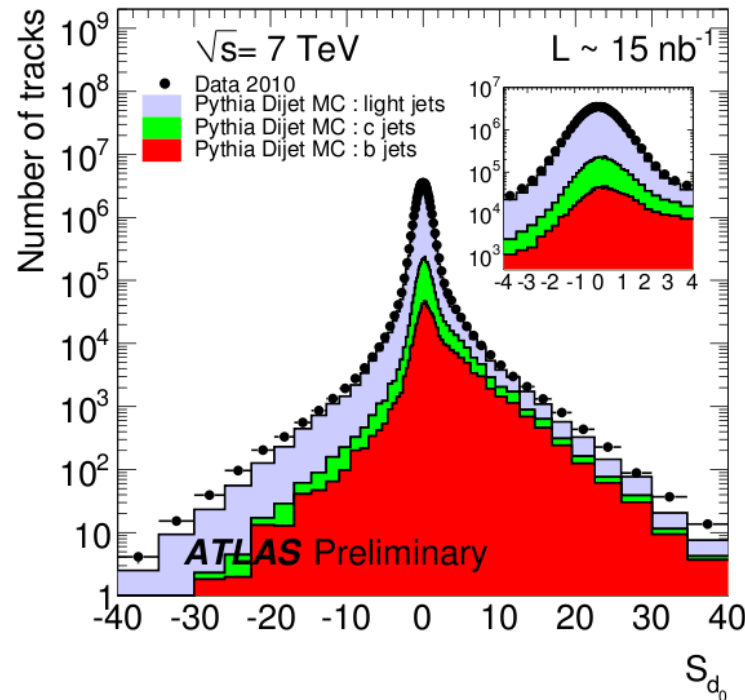


**Pouvoir discriminatif apparent, mais comment construire un poids pour le jet sans se servir du MC ?**

# Etiquetage dans les 1<sup>eres</sup> donnees

## Comment construire une methode d'etiquetage simple?

**Ex2:** Utiliser le parametre d'impact signe des traces associees au jet (et diviser par l'erreur de mesure):



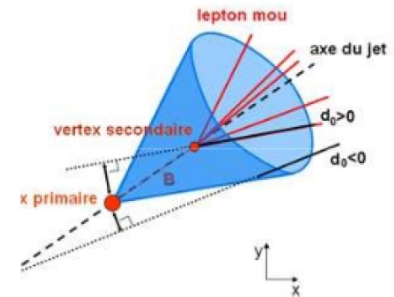
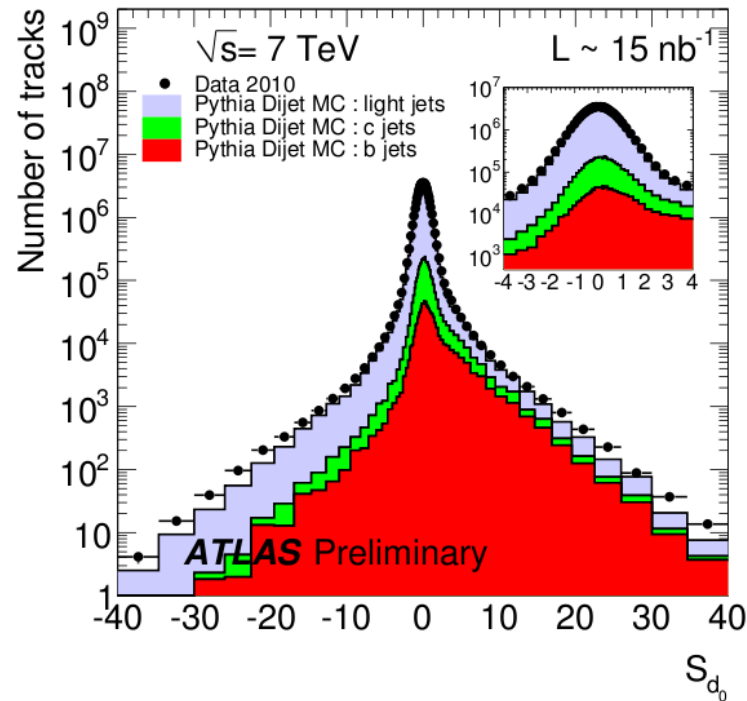
**Idee d'ALEPH:** tester la compatibilite des traces avec l'hypothese qu'elles proviennent du vertex primaire

**A PRIORI vrai pour les jets legers**

# Etiquetage dans les 1<sup>eres</sup> donnees

## Comment construire une methode d'etiquetage simple?

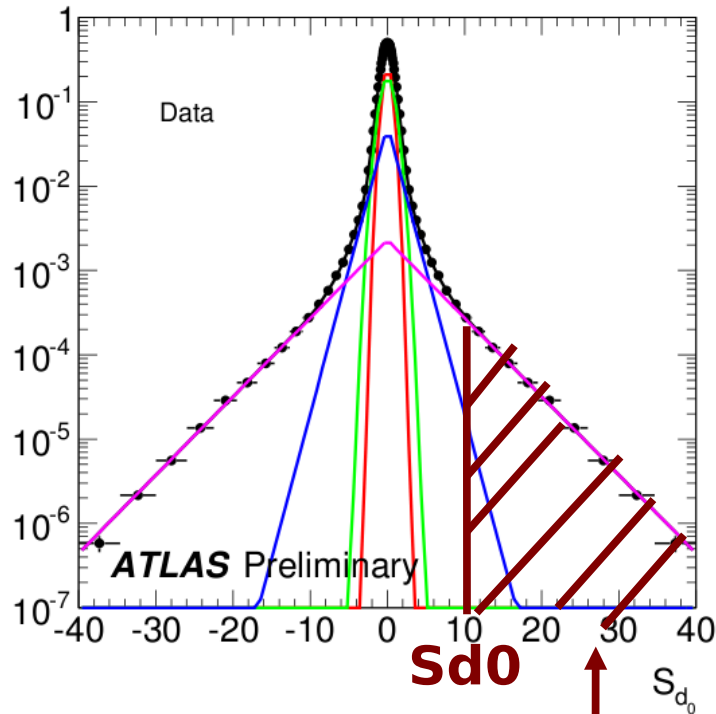
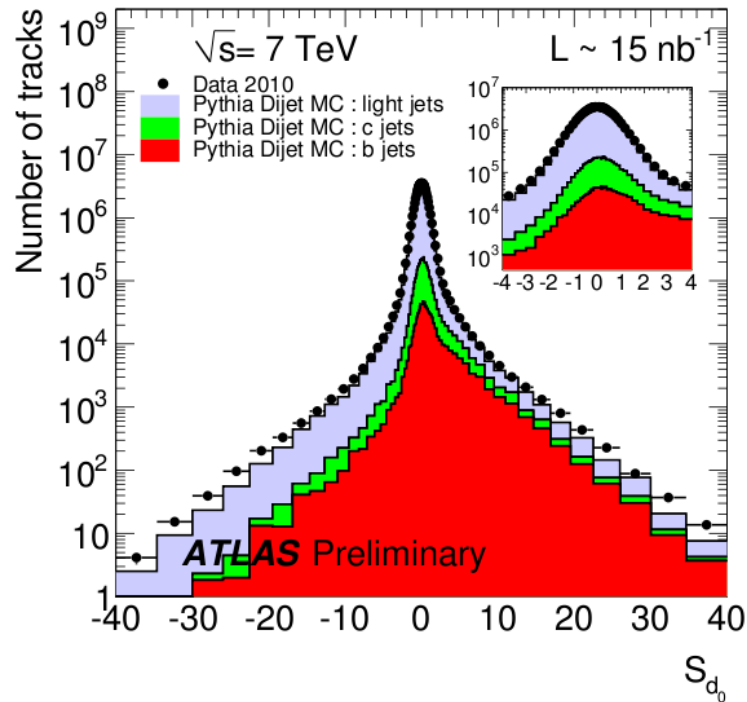
**Ex2: Utiliser le parametre d'impact signe des traces associees au jet (et diviser par l'erreur de mesure):**



→ **En pratique: Construire une 'fonction de resolution' a partir de la partie negative de la distribution des donnees**

# Etiquetage dans les 1<sup>eres</sup> donnees

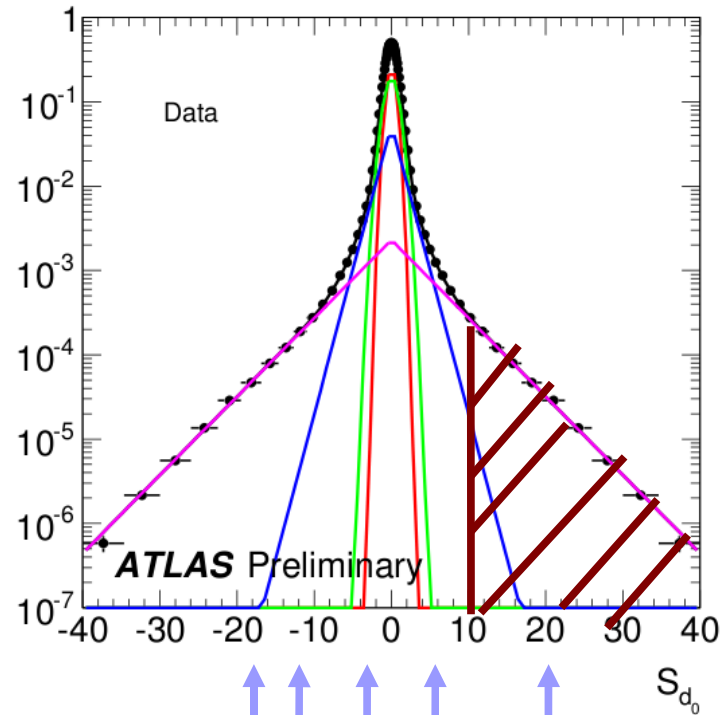
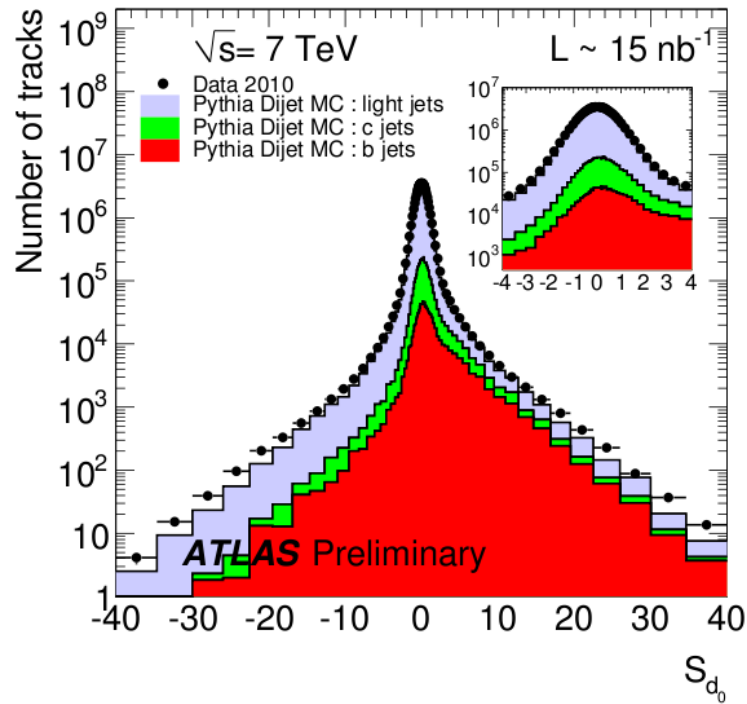
Comment construire une methode d'etiquetage simple?



Puis pour chaque trace, calculer cette integrale (Ptrack) a partir de son  $S_{d_0}$

# Etiquetage dans les 1<sup>eres</sup> donnees

Comment construire une methode d'etiquetage simple?

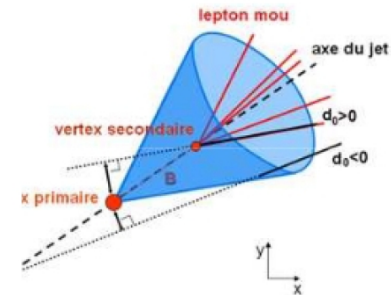
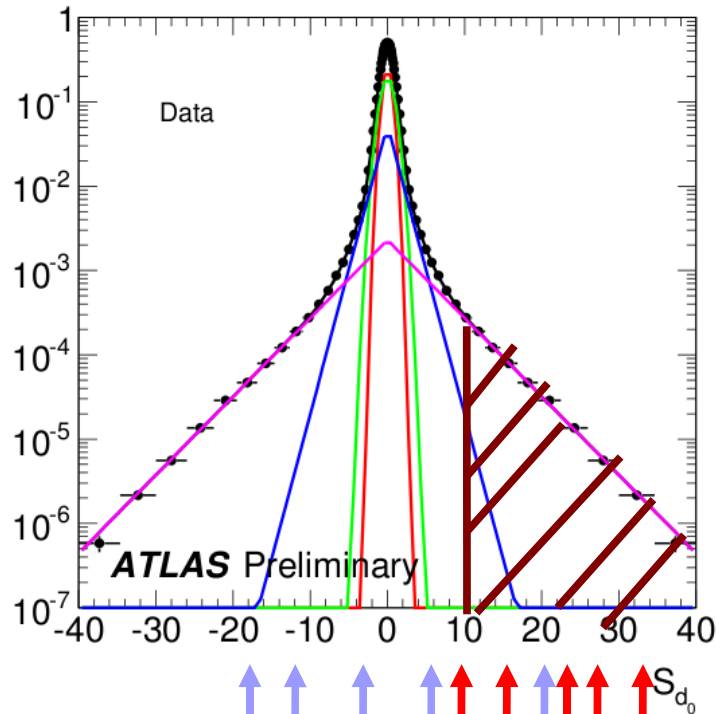
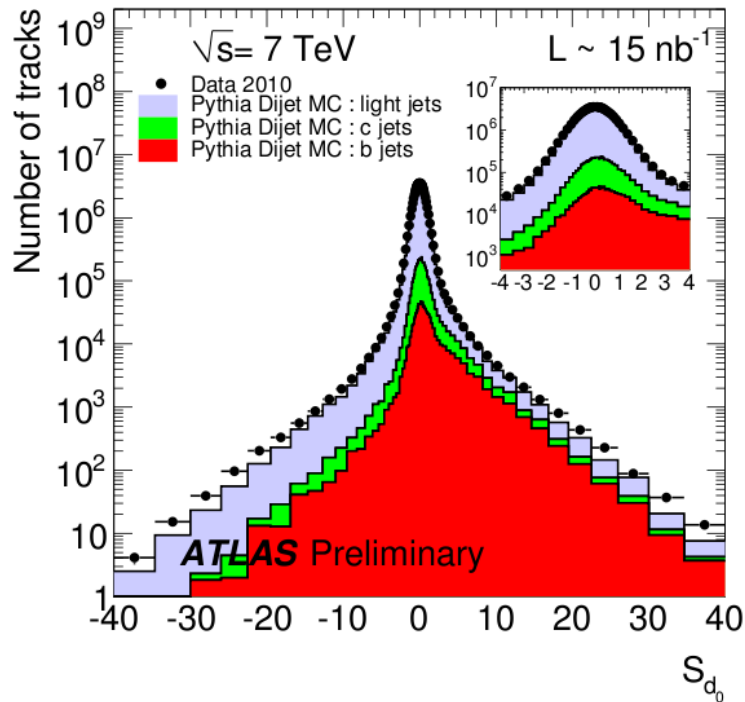


**Intuitivement:**

→ la distribution de Ptrack sera plutot symetrique p/r a 0.5 pour les jets legers

# Etiquetage dans les 1<sup>eres</sup> donnees

Comment construire une methode d'etiquetage simple?



**Intuitivement:**

→ la distribution de  $P_{\text{track}}$  sera plutot symetrique p/r a 0.5 pour les jets legers

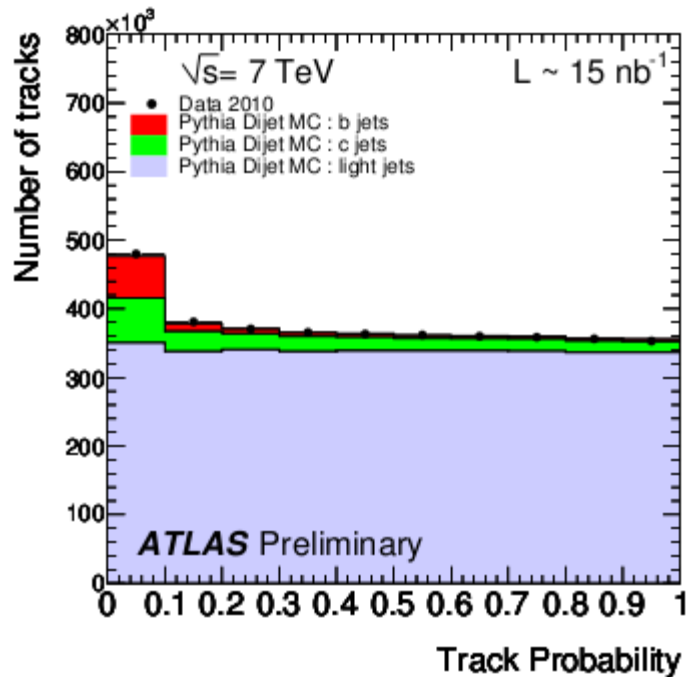
→ mais piquee vers 0 pour les traces des jets b



# Etiquetage dans les 1<sup>eres</sup> donnees

Comment construire une methode d'etiquetage simple?

Track probability



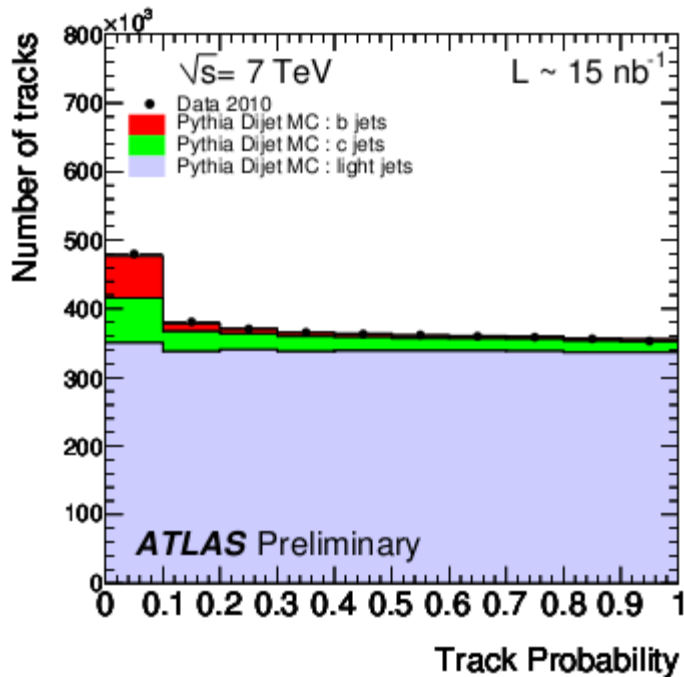
$$\mathcal{P}_0 = \prod_{i=1}^N \mathcal{P}_{\text{trk}i}$$

**N = # de traces ds le jet**

# Etiquetage dans les 1<sup>eres</sup> donnees

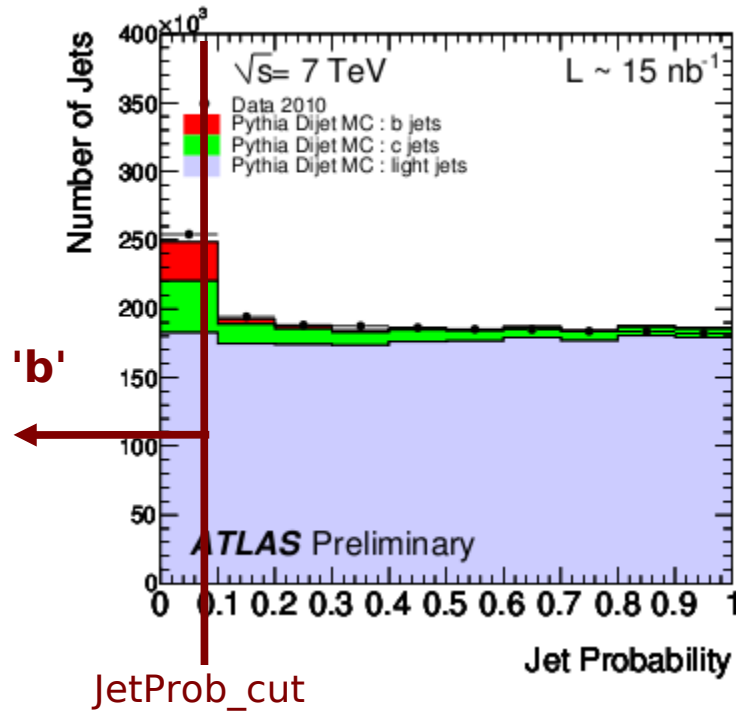
Comment construire une methode d'etiquetage simple?

Track probability



$$\mathcal{P}_0 = \prod_{i=1}^N \mathcal{P}_{\text{trki}}$$

Jet probability



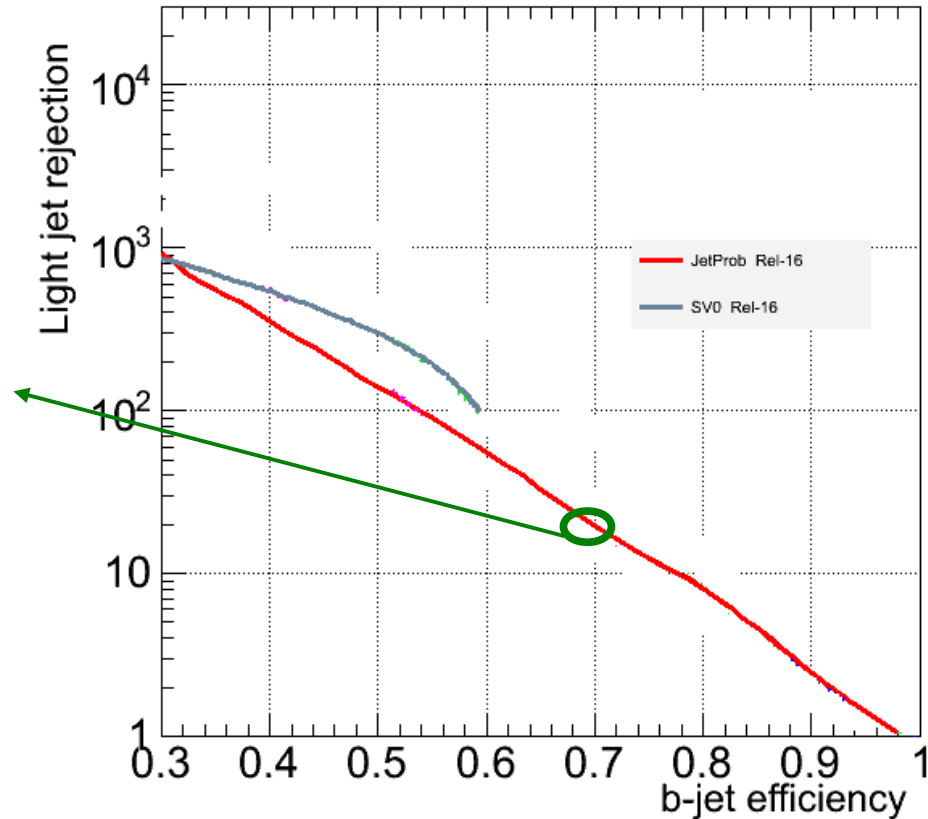
$$\mathcal{P}_{\text{jet}} = \mathcal{P}_0 \sum_{k=0}^{N-1} \frac{(-\ln \mathcal{P}_0)^k}{k!}$$

Jet etiquete comme b si son  $\mathcal{P}_{\text{jet}} < \text{JetProb\_cut}$

# Etiquetage dans les 1<sup>eres</sup> donnees

**JetProb permet  
d'atteindre des plus  
hautes efficacites**

**70% des jets b identifiés  
correctement ; 1 jet mal  
identifié sur 20**



# Les algorithmes les plus évolués



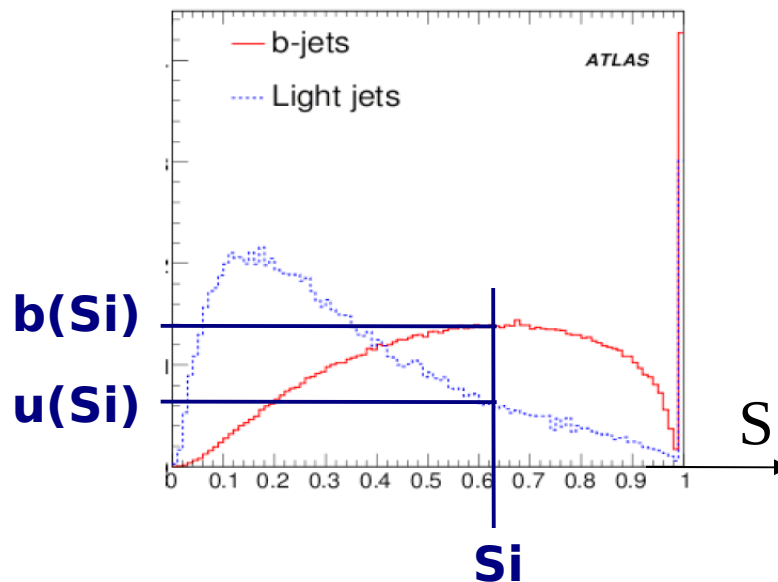
# Les algorithmes les plus evolues

Comment *bien mieux* identifier les jets b ?

→ En utilisant la simulation Monte-Carlo !

D'un point de vue statistique, on peut montrer que le classificateur le plus puissant est le ratio:

$\frac{b(S_i)}{u(S_i)}$   $b(u)$  etant la p.d.f. de la variable  $S$  dans l'hypothese que le jet provient d'un jet b (leger)



# Les algorithmes les plus evolues

Comment *bien mieux* identifier les jets b ?

→ En utilisant la simulation Monte-Carlo !

D'un point de vue statistique, on peut montrer que le classificateur le plus puissant est le ratio:

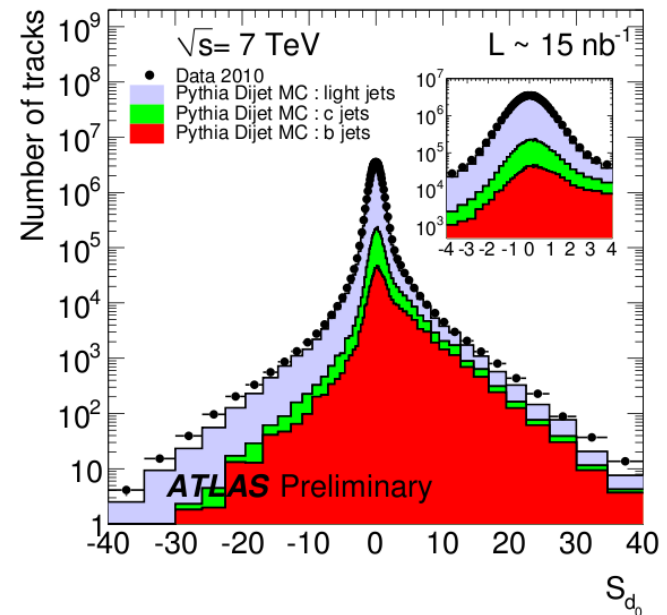
$$\frac{b(S_i)}{u(S_i)}$$

$b(u)$  etant la p.d.f. de la variable  $S$  dans l'hypothese que le jet provient d'un jet b (leger)

Exemple avec la distribution  $S_{d0}$ :

$$W_{Jet} = \sum_{i=1}^{N_T} \ln W_i = \sum_{i=1}^{N_T} \ln \frac{b(S_i)}{u(S_i)}$$

$N_T = \#$  de traces dans le jet



# Les algorithmes les plus evolues

Comment *bien mieux* identifier les jets b ?

→ En utilisant la simulation Monte-Carlo !

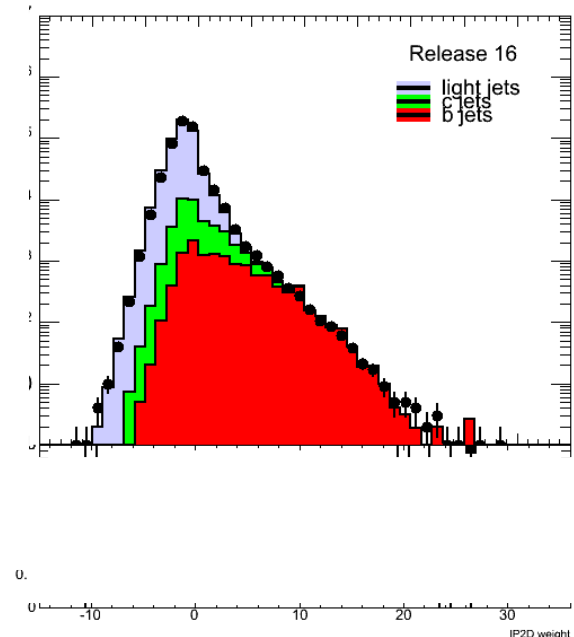
D'un point de vue statistique, on peut montrer que le classificateur le plus puissant est le ratio:

$\frac{b(S_i)}{u(S_i)}$   $b(u)$  etant la p.d.f. de la variable S dans l'hypothese que le jet provient d'un jet b (leger)

Exemple avec la distribution Sd0:

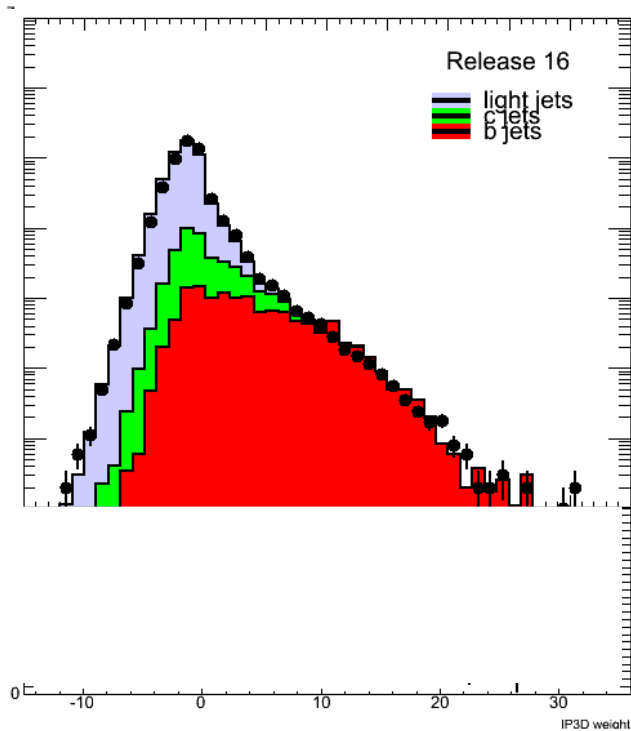
$$W_{Jet} = \sum_{i=1}^{N_T} \ln W_i = \sum_{i=1}^{N_T} \ln \frac{b(S_i)}{u(S_i)}$$

Distribution des 'poids' des jets pour ce tagger (IP2D) :

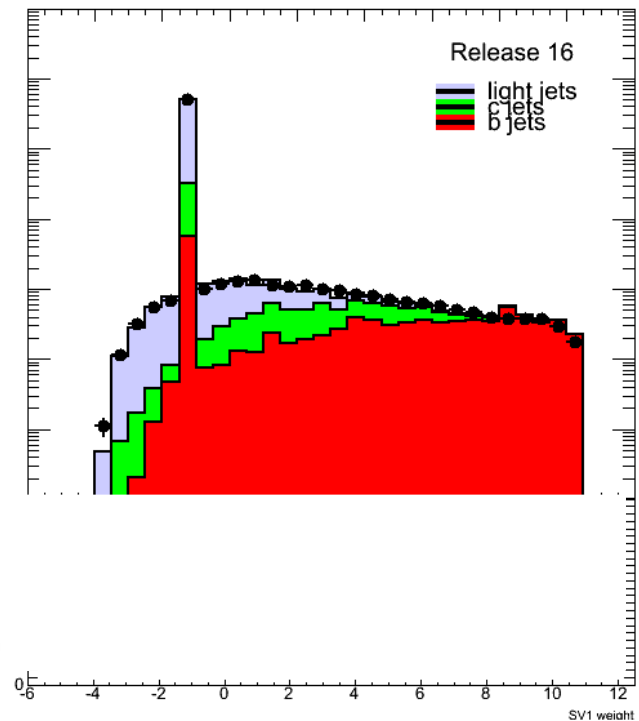


# Les algorithmes les plus evolues

On peut faire de meme avec la  
significance 3D du parametre  
d'impact des traces (IP3D)



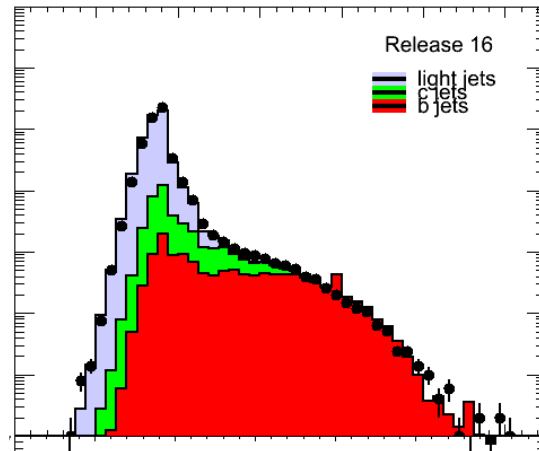
Ou encore a partir des  
proprietes du vertex  
secondaire (SV1)





# Les algorithmes les plus evolues

On peut aussi additionner les poids de ces differents taggers pour obtenir un tagger plus performant (Ex: IP3D+SV1)

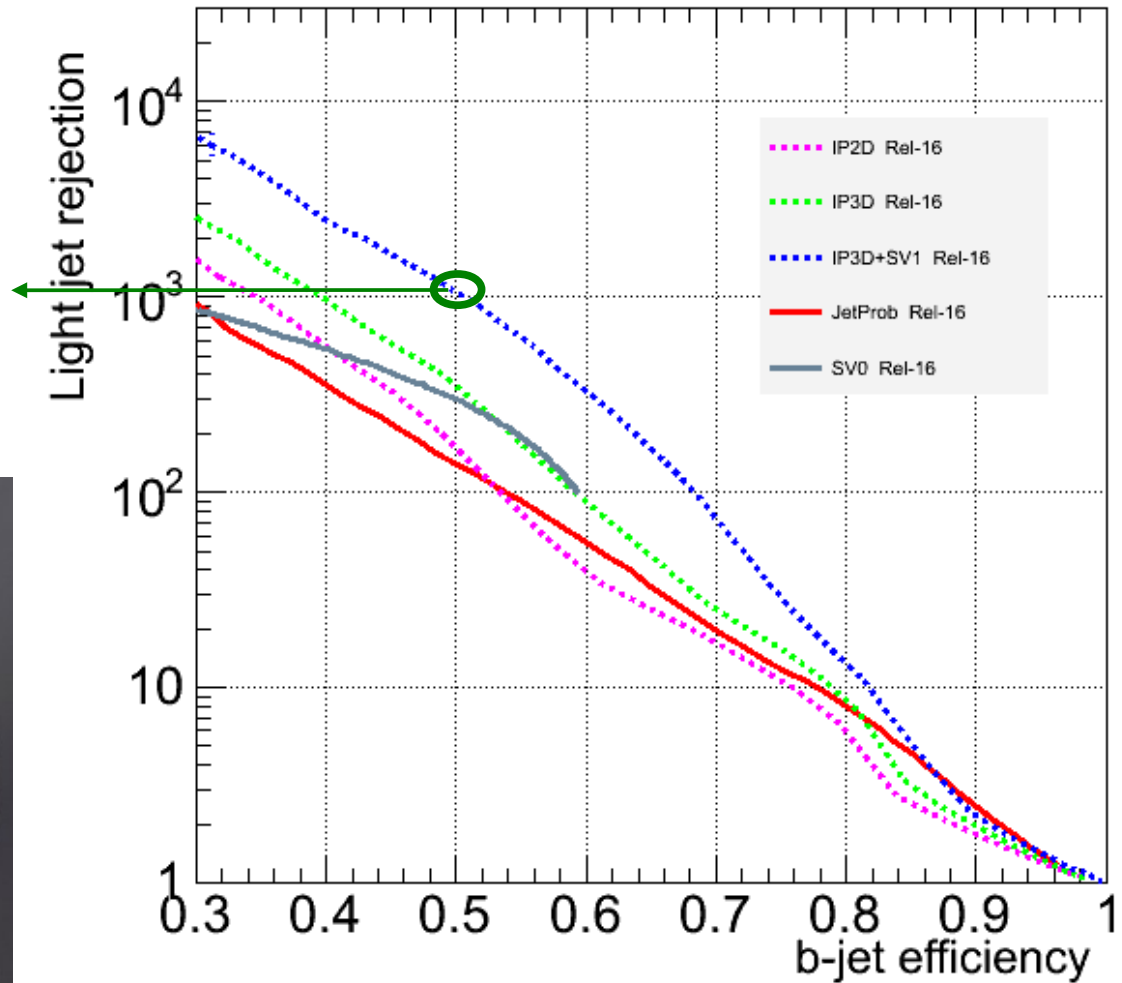


**Non-optimal si les variables d'entree sont correlees entre elles:**

→ **Motivation pour developper des methodes d'analyse 'multivariee' plus puissante (Arbres de decision, reseaux neuronaux...)**

# Les algorithmes les plus evolues

**50% des jets b identifiés  
correctement ; 1 jet mal  
identifié sur 1000 !**



# Contributions personnelles

**Etudes preliminaires sur la reconstruction des traces  
utilisees pour le b-tagging**

**Participation a la mise en oeuvre de JetProb, et des  
taggers les plus evolues**

**Production des histogrammes de 'reference' -  
Interfacage avec la base de donnees ATLAS**

**'Shifts' pixels online**

**Conference Notes:** Tracking studies for b-tagging with 900 GeV collision data - Bousson, N et al - ATL-COM-PHYS-2010-087

First look at the JetProb b-tagging algorithm in 900 GeV collision data - Lapoire, C et al - ATL-COM-PHYS-2010-086

Tracking studies for b-tagging with 7 TeV collision data with the ATLAS detector - Bousson, N et al - ATL-COM-PHYS-2010-273

The background is a deep blue, almost black, space scene. It features a dense field of small, distant stars and a prominent, glowing nebula or star-forming region. The nebula has a complex, filamentary structure with various shades of blue and purple, suggesting the presence of interstellar dust and gas. The overall effect is that of a vast, cosmic environment.

**Vers les hautes luminosités du LHC**

# **Le LHC fonctionne merveilleusement bien!**

**D'immenses progres realises pour augmenter la luminosite (~taux d'interaction des paquets de protons)**

**Pic de luminosite instantanee record:**

$$**L \sim 2.1 \cdot 10^{32} \text{ cm}^{-2}\text{s}^{-1}**$$

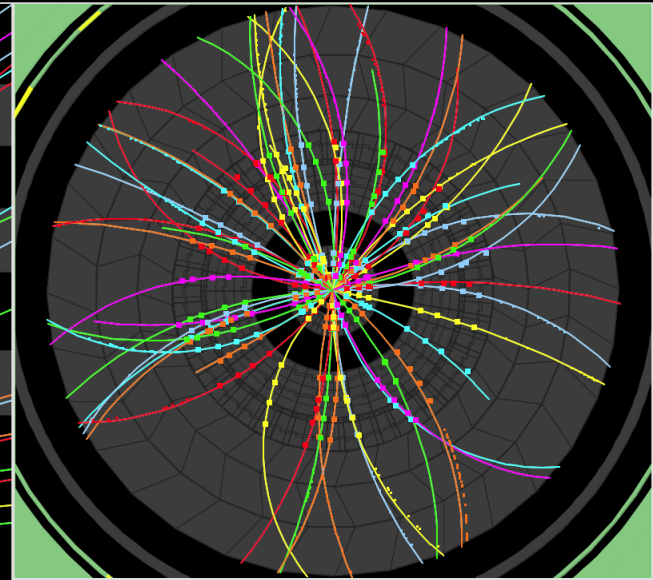
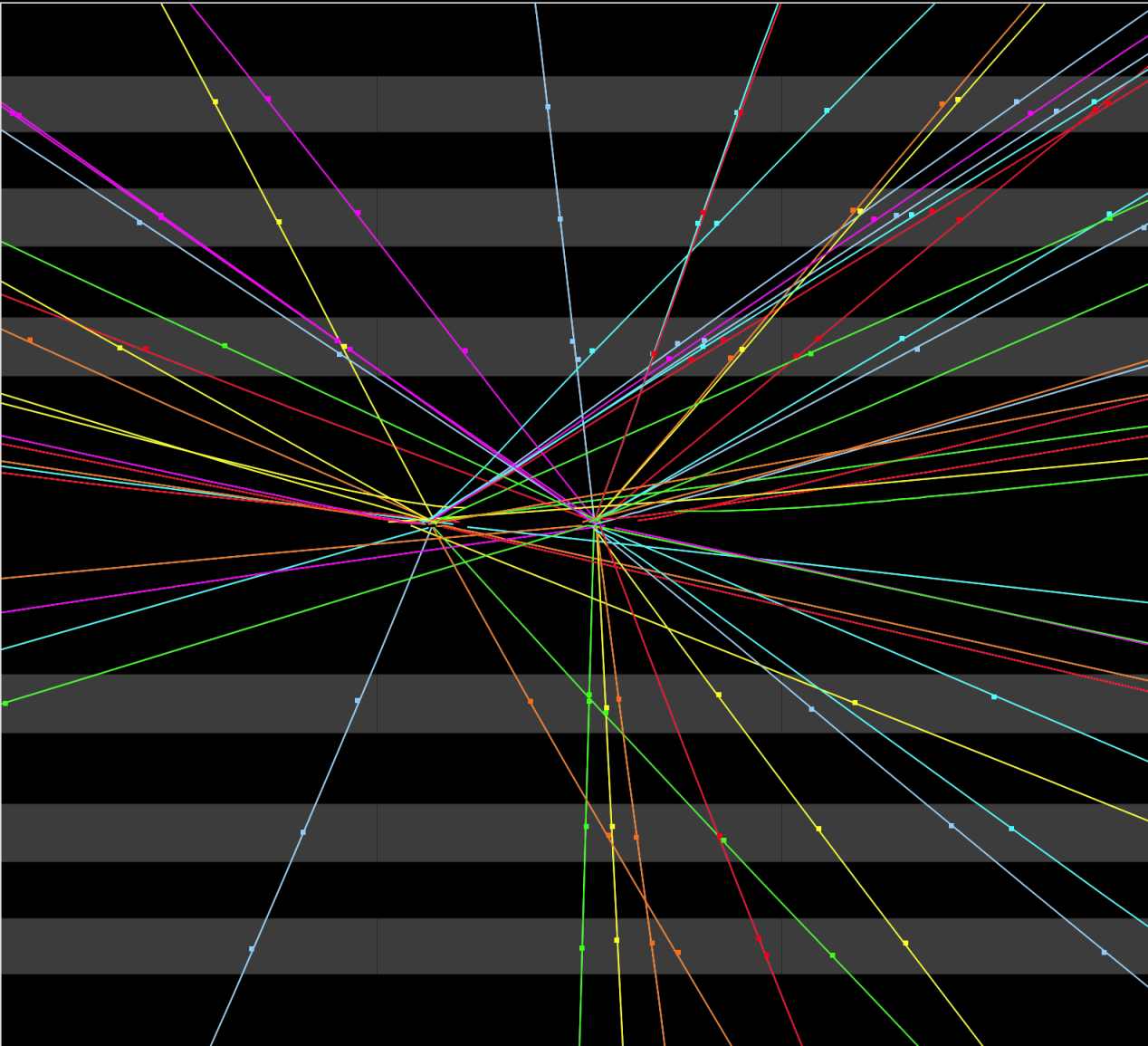
**Deja 45pb<sup>-1</sup> de luminosite integree !**

**Quelques fb<sup>-1</sup> fin 2011 !**



**MAIS...**

# Collision Event at 7 TeV with 2 Pile Up Vertices



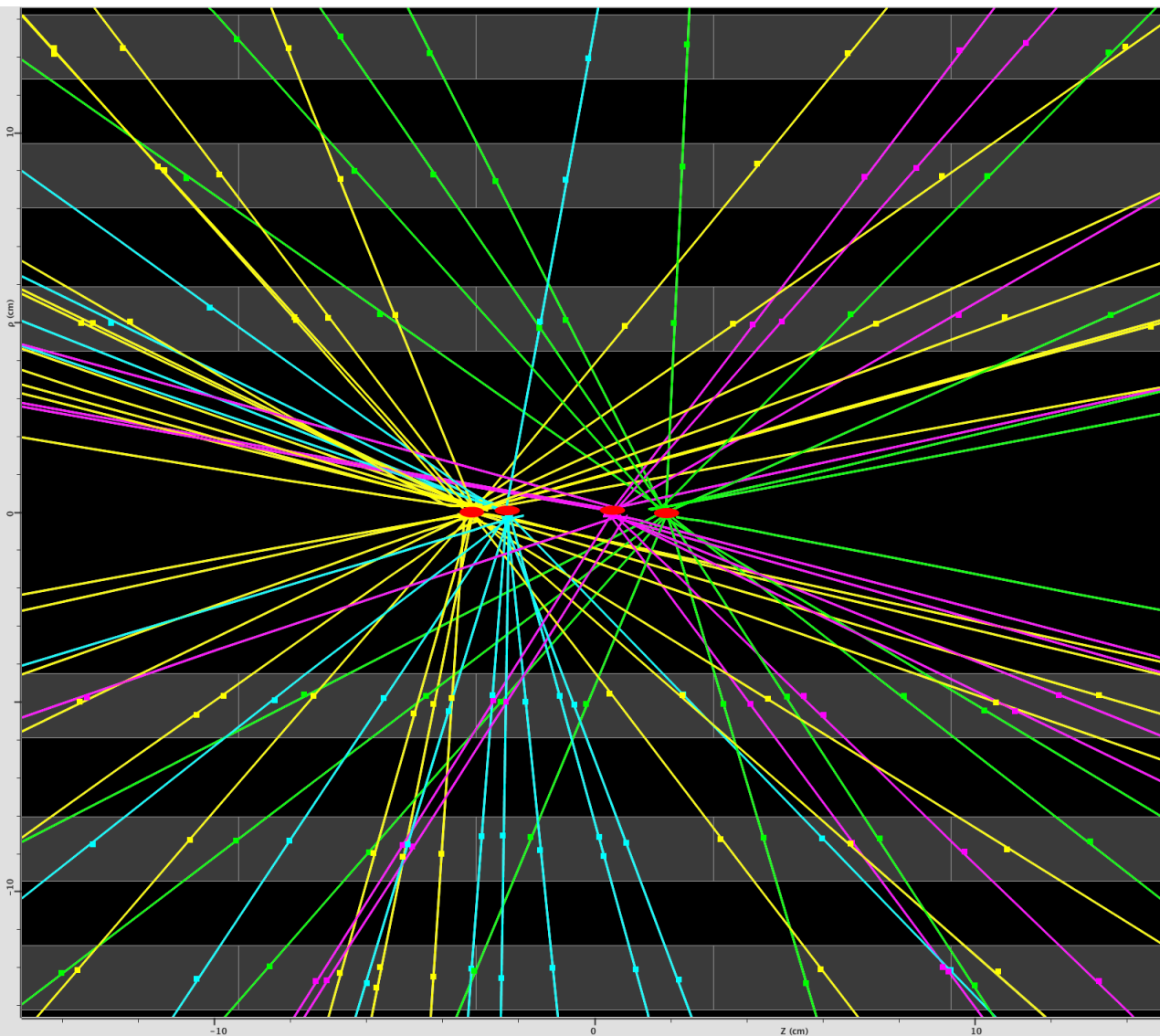
**ATLAS**  
**EXPERIMENT**

Run Number: 152166, Event Number: 467774

Date: 2010-03-30 13:31:46 CEST

<http://atlas.web.cern.ch/Atlas/public/EVTDISPLAY/events.html>



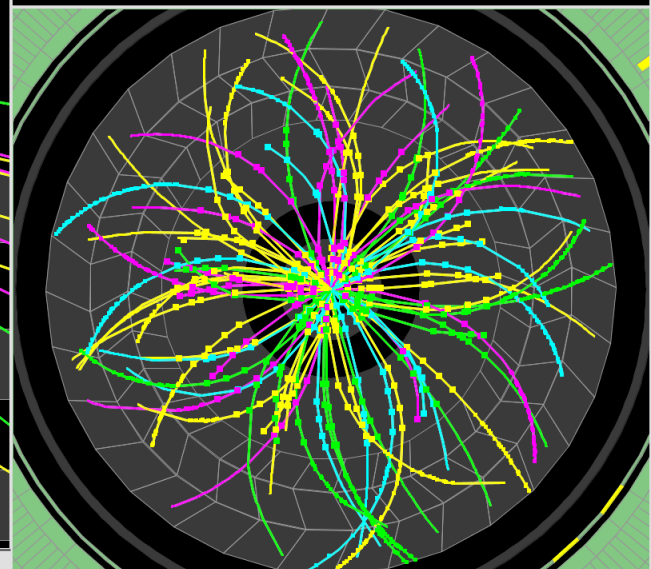


# ATLAS EXPERIMENT

Run Number: 153565, Event Number: 4487360

Date: 2010-04-24 04:18:53 CEST

**Event with 4 Pileup Vertices  
in 7 TeV Collisions**



Anakin Skywalker is shown in profile on the left, holding a glowing red lightsaber. In the background, the helmet of Darth Vader is visible. The entire scene is set against a dark red, atmospheric background.

**Et avec 50 interactions  
d'empilement ?!!**

**STAR WARS**

**EPISODE III**

**RISE OF THE EMPIRE**

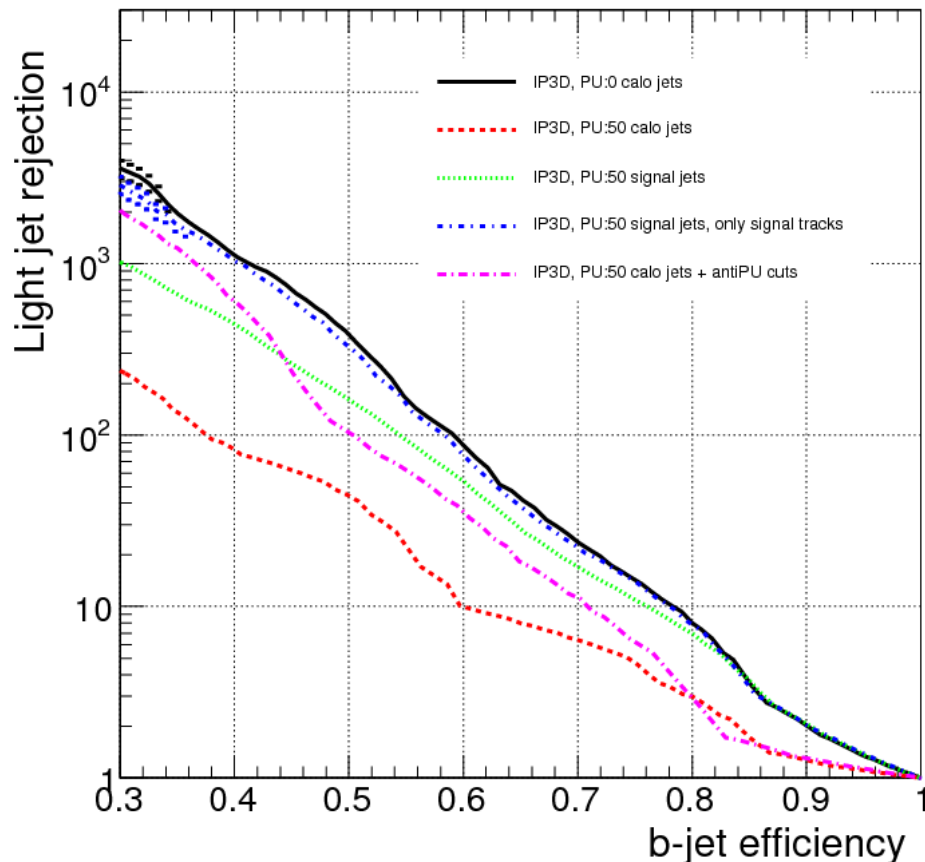


# STAR WARS

JEDI vs SITH

# Vers les hautes luminosités du LHC

**Les performances d'identification des jets b souffrent énormément (si l'on ne fait rien) !**



**Principales causes:**

**Se tromper de vertex primaire!**

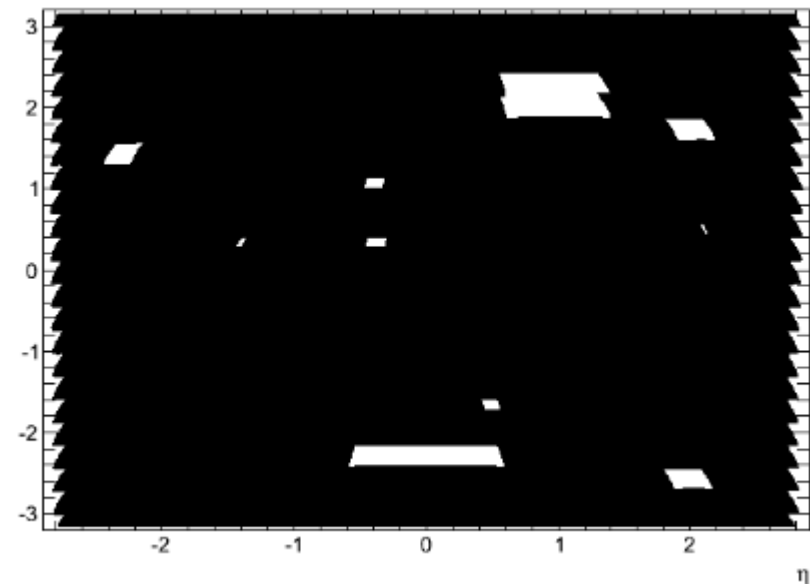
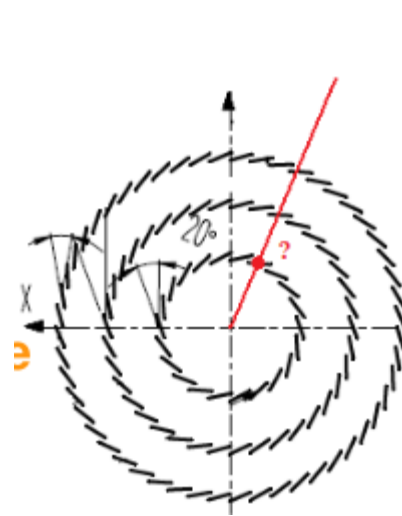
**'Extra' jets provenant d'une autre collision**

**'Extra-traces'**

# Vers les hautes luminosités du LHC

## Autre problematique:

**Le niveau de radiation a haute luminosite va progressivement endommager le detecteur a pixels (effet signifiant dès qq 100aines d' $\text{fb}^{-1}$  )**



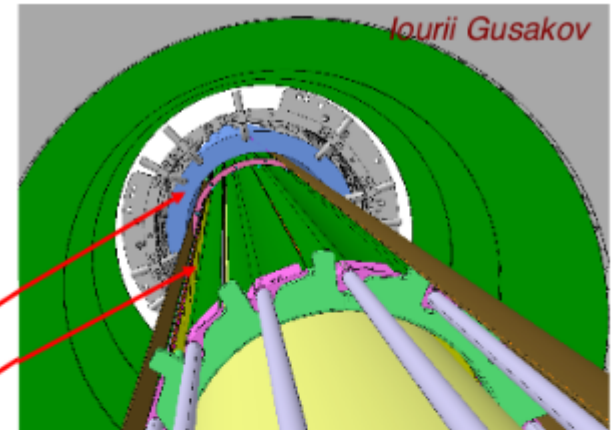
**→ Effet direct sur les performances d'identification des jets b !**

# Vers les hautes luminosités du LHC

**Il est prévu d'insérer une couche supplémentaire de détection (~2014)**

## Motivations:

**Inefficacités de détection dues aux radiations**



Existing B-layer

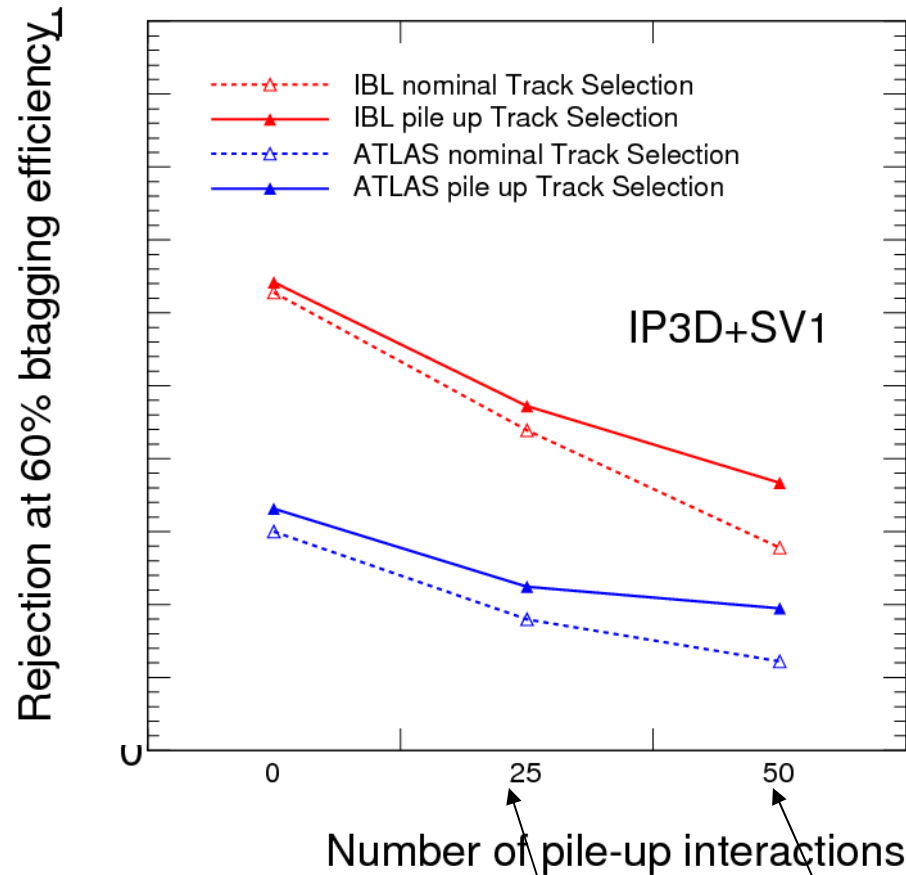
IBL (Staves)

**Assurance en cas de gros dégât sur l'actuelle b-layer**

**Il n'est pas possible de remplacer la b-layer pendant un arrêt long du LHC (~8mois)**



# Vers les hautes luminosités du LHC



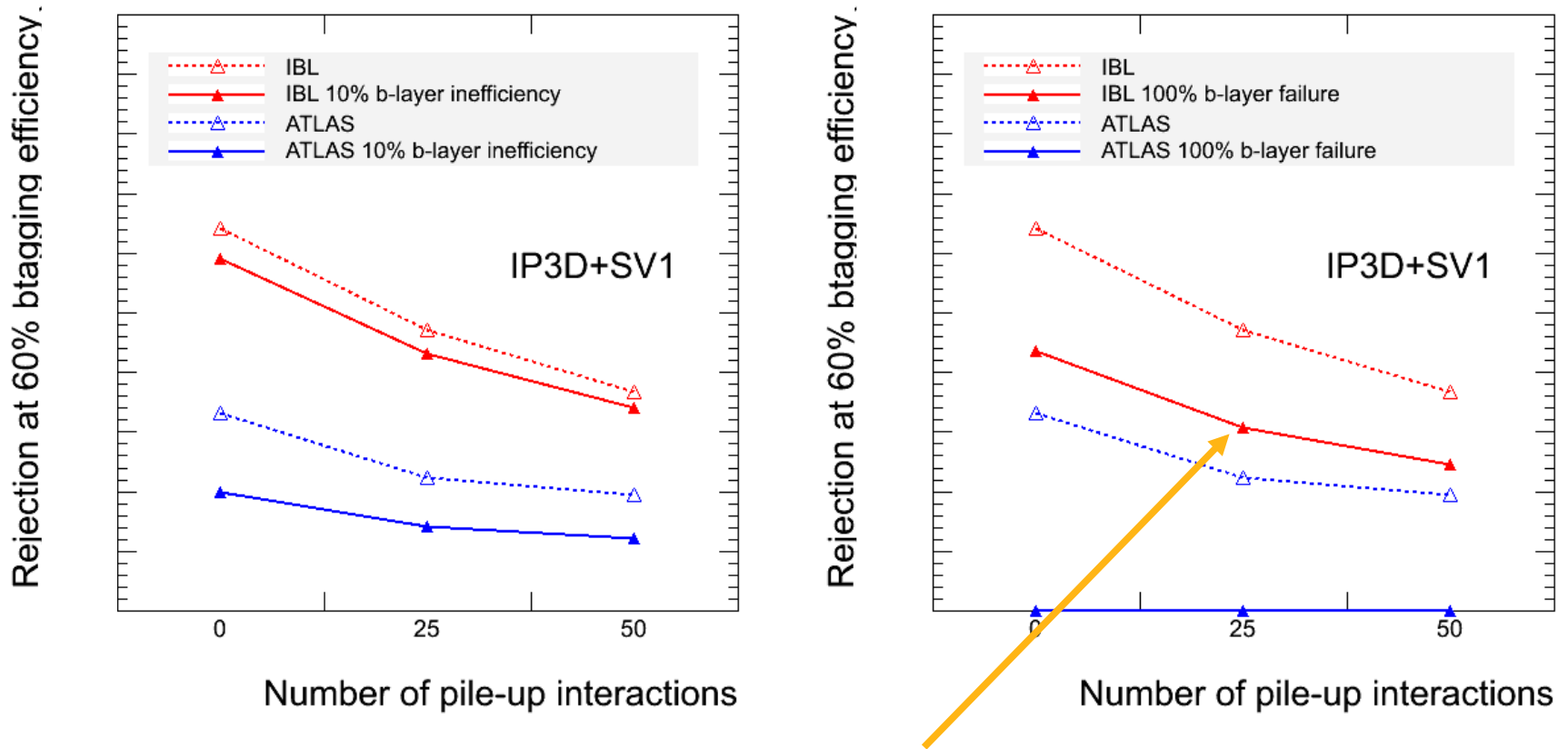
**Plot in ATLAS IBL  
Technical Design Report**

$10^{34} \text{ cm}^{-2} \cdot \text{s}^{-1}$

$2 \cdot 10^{34} \text{ cm}^{-2} \cdot \text{s}^{-1}$

# Vers les hautes luminosités du LHC

## Simulation des performances avec des dégradations du détecteur a pixels



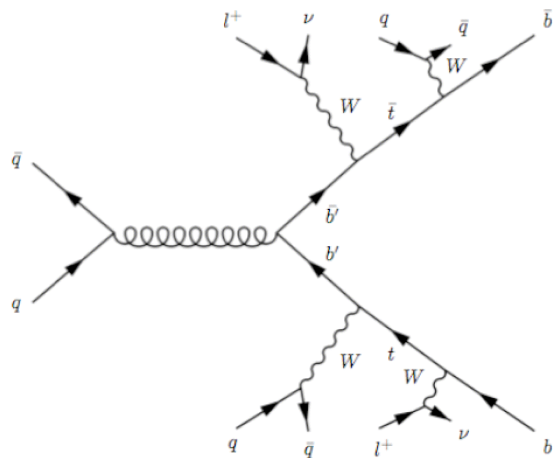
**ATLAS avec IBL sans b-layer meilleur que ATLAS actuel !**



# Conclusion et perspectives

**Les algorithmes simples d'identification des jets  $b$  fonctionnent déjà très bien, et la mise en oeuvre des taggers plus évolués est en cours**

**Mon analyse physique portera sur la recherche d'une 4<sup>ème</sup> génération de quarks**



**L'optimisation de l'identification des jets  $b$  de grande impulsion transverse sera cruciale**

**Merci pour votre attention !**

**Questions ?**

