

Etiquetage des jets issus de quarks b avec ATLAS au LHC



Cécile Lapoire
Dirigée par A. Rozanov et L. Vacavant

Séminaire doctorants CPPM 3^e année
16/11/2009



Plan

- Motivations
- Dispositif expérimental
- Etiquetage des jets b :
 - Avec les premières données : **JetProb**
 - Mesure des **performances** dans les **données**
- Conclusion et perspectives

Motivations

- **Mesure de $\sigma(tt)$:**

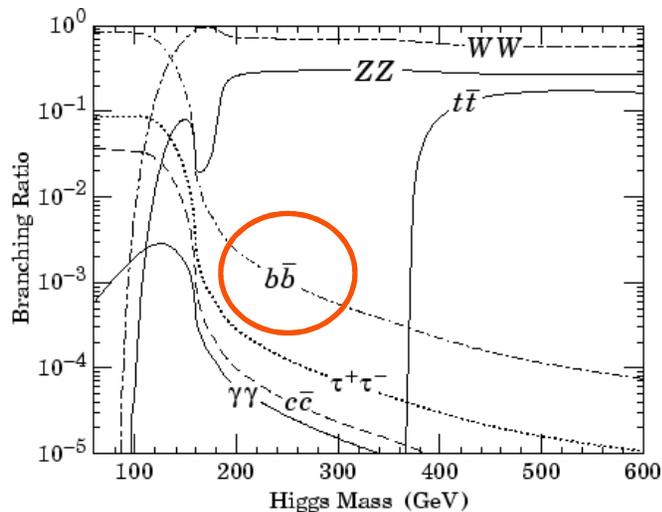
- $BR(t \rightarrow Wb) \sim 100\%$
- Canal semileptonique : $tt \rightarrow bbjjlv_l$
 - **2 jets b** à identifier



+ de détails dans le slide suivant

- **b-tagging également nécessaire pour :**

- $\sigma(bb)$
- Nouvelle physique
- Higgs ($H \rightarrow bb$ dominant à basse masse)
- etc.



Motivations : $\sigma(tt)$ dans les 1^{ères} données

- Energie des faisceaux revue à la baisse (14TeV→10TeV→7TeV→...) → **b-tagging encore + important**

- S/B de $\sigma(tt)$ à souffert : 10TeV→7TeV :

- $\sigma(tt)$ réduit d'un facteur 2.5
- $\sigma(W+jets)$ (bdf principal) de ~1.6

- Effet du **b-tagging** :

- permet de retrouver un rapport S/B satisfaisant :



20pb ⁻¹ @7TeV	No b-tag	1 b-tag : JP60
ttbar	51	~35
W+jets	28	<1
S/B	1.3	2.9

evb jjb	200pb ⁻¹ @10TeV	20pb ⁻¹ @7TeV
ttbar	1286	51
W+jets	448	28
Single top	81	4
Other bkgd	67	7
S/B	2.1	1.3
S/ $\sqrt{S+B}$	30	5.4

#evts passant la sélection
(sans utiliser le b-tagging)

- Etudes menées dans ATLAS :

- supposaient énergie transverse manquante (MET) comprise rapidement et correctement
- s'affranchissaient du b-tagging

- **MAIS : Calibration MET non triviale**

- **mon objectif** : privilégier le b-tagging et se concentrer sur un étiquetage satisfaisant avec les premières données (en relâchant éventuellement la coupure sur MET)

Dispositif expérimental

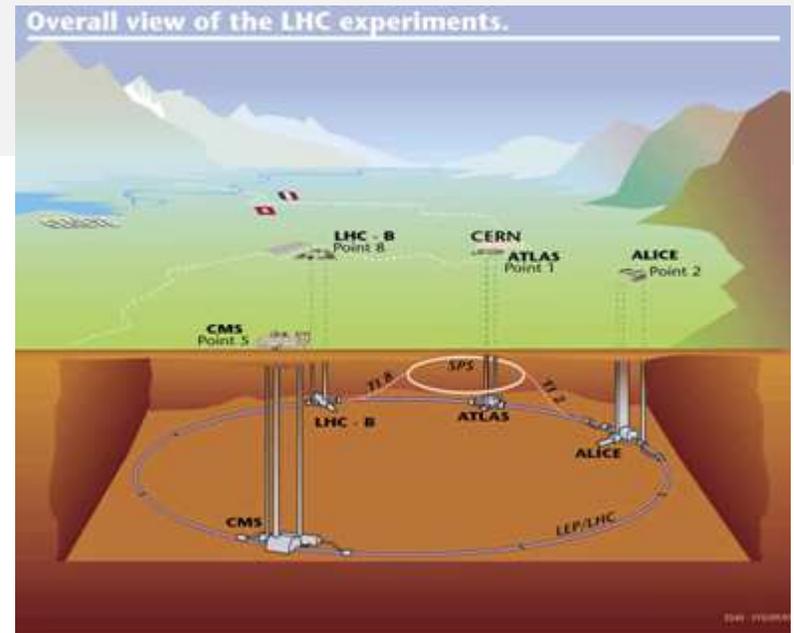
LHC

- Collisionneur p⁺-p⁺
- 27 km de circonférence



- Paramètres nominaux :
 - 14 TeV au centre de masse
 - Luminosité : 10⁻³⁴cm⁻²s⁻¹
 - 40MHz
 - ~100fb⁻¹ par an

- 1^{ere} mise en service en sept. 2008 : accident sérieux
- **Remise en service prévue dans les prochains jours**
 - 900GeV jusqu'à l'arrêt hivernal
 - ~ 1nb⁻¹
 - 2.2 à 7 TeV en 2010
 - ~200pb⁻¹

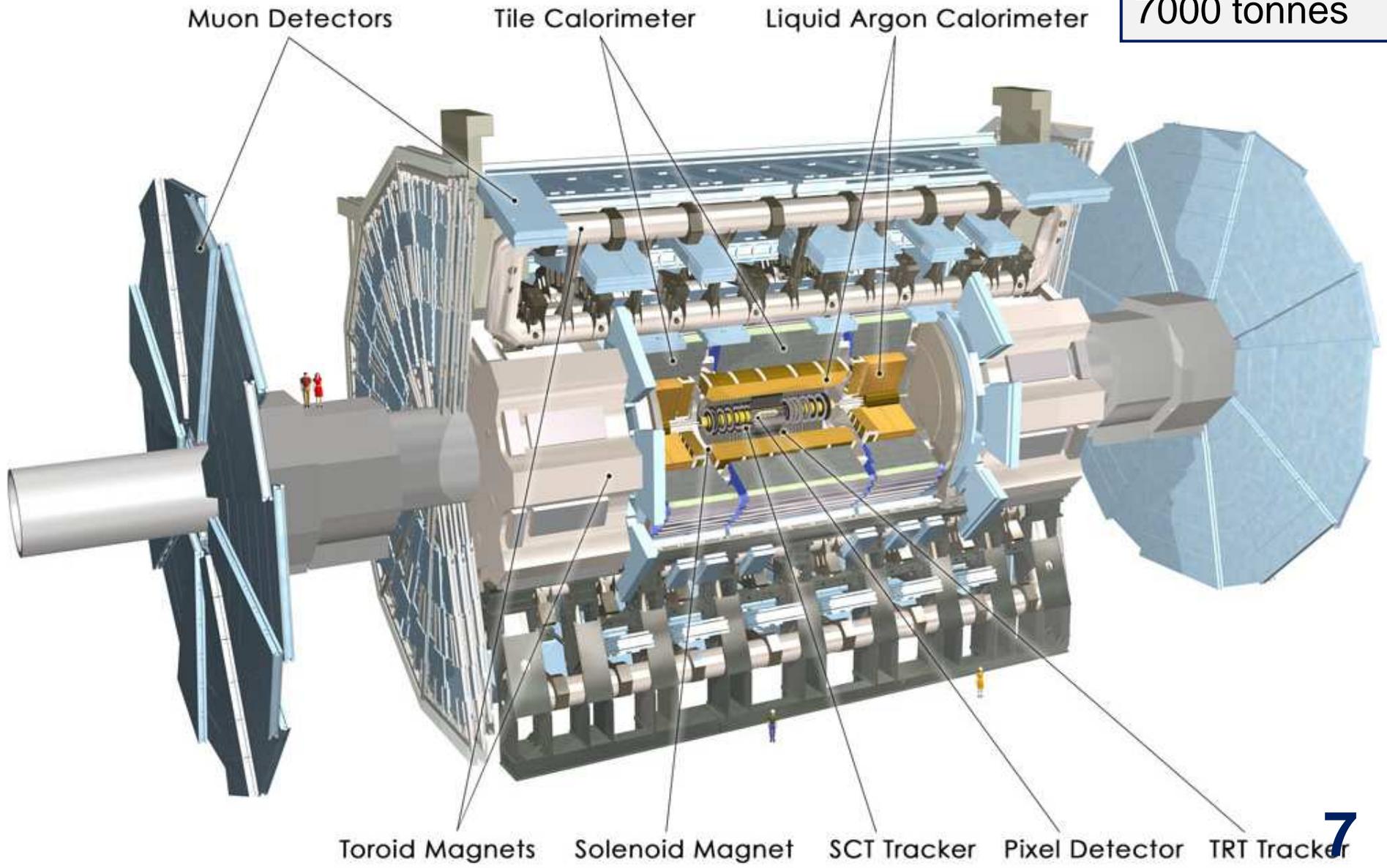


Quadrupole-dipole
interconnection (27R3)



ATLAS

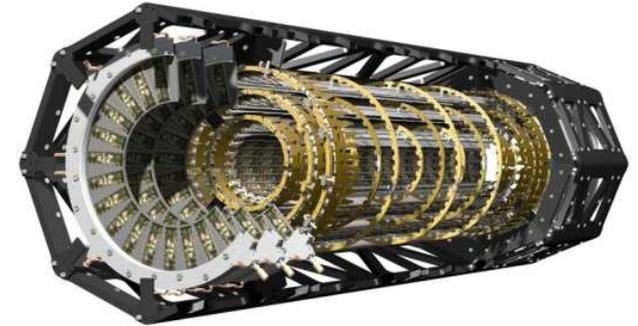
42 m de long
20 m de haut
7000 tonnes





Détecteur à pixels

Essentiel pour le
b-tagging



- Contraintes :
 - Résistant aux radiations
 - Transparent aux particules qui le traversent
- **3 couches** (à 5.05, 8.85 et 12.25 cm de l'axe du faisceau) + **3 disques** de chaque côté
 - 3 hits dans les pixels jusqu'à $|\eta| < 2.5$
- **80 363 520 pixels*** de dimensions 50 ($R\phi$) x 400 (z) μm
- **Essentiel pour l'étiquetage des jets b :**
 - **Traces des particules chargées reconstruites très précisément** grâce à :
 - **Très faible occupation**
 - **Très grande efficacité de détection**
 - **Très grande granularité :**
 - Très bonne résolution du **paramètre d'impact transverse** d_0**
 - Résolution intrinsèque : $\sim 10\mu\text{m}$ en $R\phi$, $\sim 115\mu\text{m}$ en z

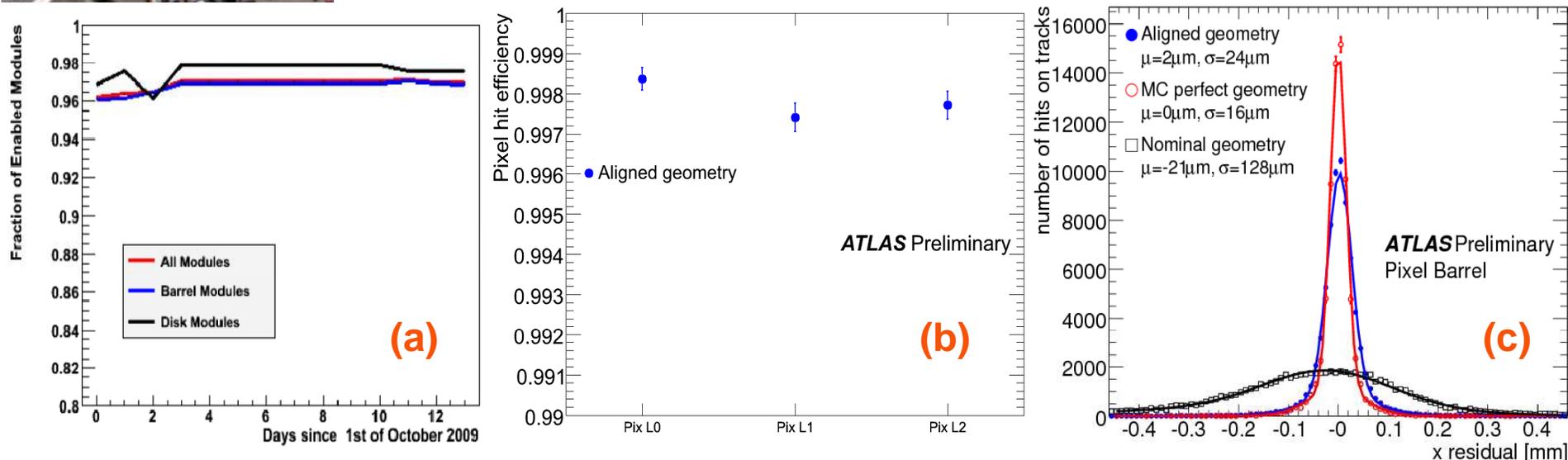
*1 pixel = diode de détection (jonction PN) + électronique complète de traitement associée reliées par une soudure à bille

** distance minimale d'approche de la trace au vertex primaire



Détecteur à pixels

Performances actuelles



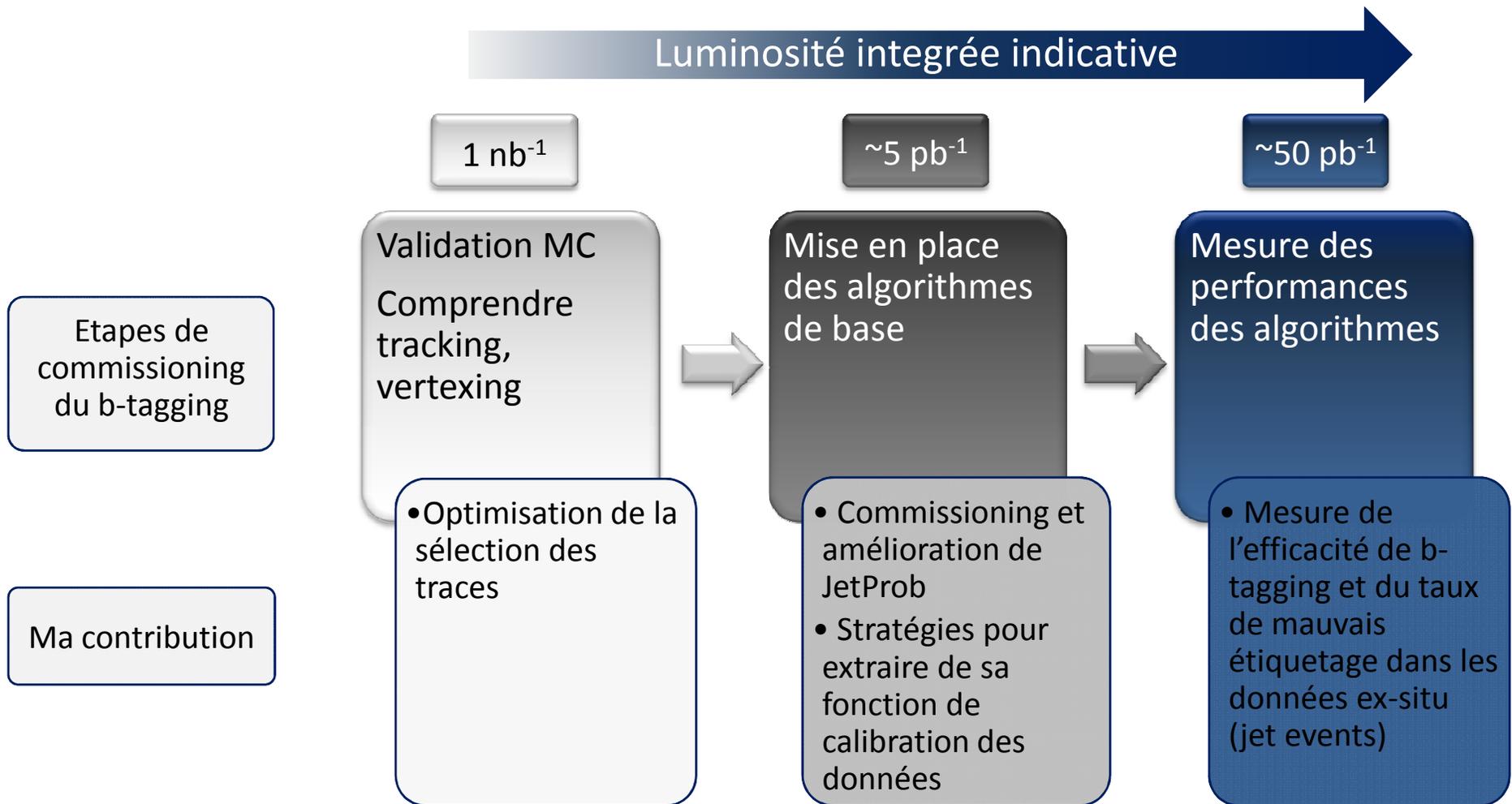
- Fraction de **modules désactivés** (simulation : 5%) :
 - sur tout le détecteur : **~2.1% (a)**, **b-layer*** : **1.7%**
- **Très faible occupation** : 10^{-10} (après avoir masqué les pixels bruyants)
- **Très grande efficacité de détection** : **>99.8% (b)**
- **Résolution des hits (en $R\phi$)** :
 - Avec un premier alignement de la géométrie (muons cosmiques) : **24 μm** (idéal 16 μm) **(c)**
 - Meilleur alignement attendu rapidement avec les premières données (dès 900GeV)

*crucial pour le b-tagging qui requiert des traces avec au moins 1 hit dans la b-layer

Grâce à la bonne condition du détecteur à pixels, le b-tagging devrait être très vite opérationnel pour la physique dans les premières données

b-tagging

Perspectives du b-tagging

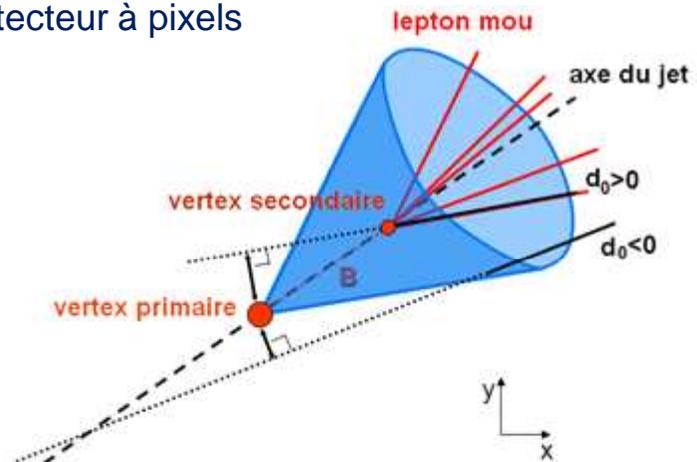


- Statistique attendue @900 GeV avec 1nb⁻¹:
 - 30k de jets légers
 - <300 jets b

b-tagging : un aperçu

- Propriétés des **hadrons B** :

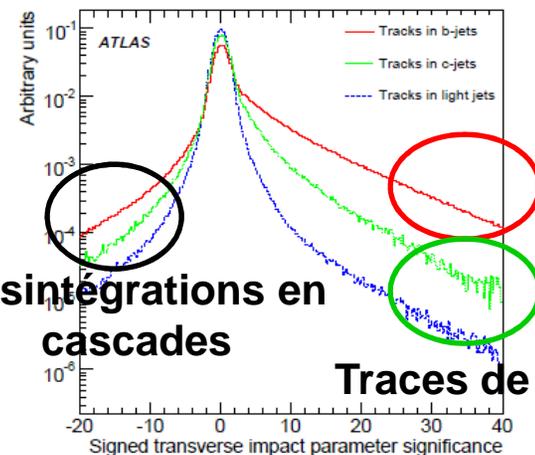
- Emportent 70% de l'impulsion du quarks b initial (fragmentation dure)
- Grande masse (5GeV) → produits de désintégration avec un grand p_T
- **Long temps de vie** (~ 1.5 ps, $c\tau=450\mu\text{m}$)
 - Typiquement : un hadron B dans un jet ($p_T=50\text{GeV}$) vole $\sim 5\text{mm}$ dans le plan transverse avant de se désintégrer ($d_0\sim 500\mu\text{m}$) : mesurable avec le détecteur à pixels
- **Désintégration semi leptonique**
 - $\text{Br}(b\rightarrow lvX)+\text{Br}(b\rightarrow c\rightarrow lvX) = 11\% + 10\%$ ($l=e, \mu$)



- **Expérimentalement :**

- Présence d'un **vertex secondaire (SV)**
- Plus grand **paramètre d'impact transverse d_0** pour les traces dans les jets venant du SV
 - **Signé** par rapport à l'**axe du jet**
 - Significance S_{d_0} :
- Présence de **leptons mous**

$$S_{d_0} = \frac{d_0}{\sigma_{d_0}}$$



Désintégrations en cascades

Traces de **B**, **C**

Nécessite un très bon tracking (slide suivant)

Sélection des traces pour le b-tagging

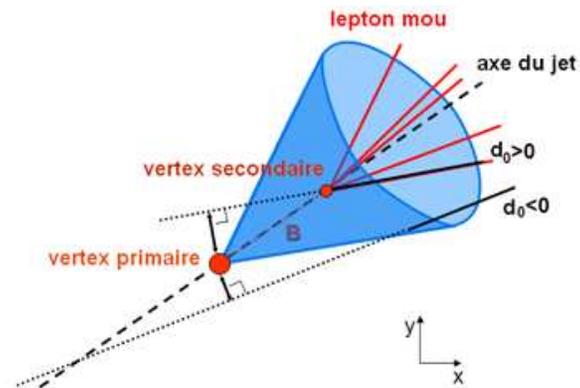
Le b-tagging requiert des traces reconstruites dans l'ID

- Sélection optimisée pour rejeter :
 - “fake tracks”
 - traces venant de la désintégration de particules à long temps de vie (Kaons, Lambda, etc.)
 - traces issues de l'interaction avec la matière (conversion de photons et interactions hadroniques)
 - **Coupures de sélection** utilisées pour le b-tagging :
 - $|p_T| > 1 \text{ GeV}$
 - $|\eta| < 2.5$
 - $|d_0| < 1 \text{ mm}$
 - $|z_0| < 1.5 \text{ mm}$
 - # hits b-layer : $N_{bla} \geq 1$
 - # hits in pixels : $N_{pix} \geq 2$
 - # hits in silicon (Pixel+SCT) ≥ 7
- Diagram illustrating the selection criteria and their purposes:
- Acceptance du détecteur: $|p_T| > 1 \text{ GeV}$ and $|\eta| < 2.5$
 - Rejet des particules à long temps de vie (V^0) et des interactions secondaires: $|d_0| < 1 \text{ mm}$ and $|z_0| < 1.5 \text{ mm}$
 - Bon σ_{d0} : # hits b-layer : $N_{bla} \geq 1$
 - Rejet des fakes: # hits in pixels : $N_{pix} \geq 2$ and # hits in silicon (Pixel+SCT) ≥ 7

En cours : optimisation des coupures de sélection pour les traces utilisées par JetProb
(slides suivants)

JetProb

Algorithme d'étiquetage basé sur le paramètre d'impact transverse d_0



Présentation de JetProb

- **Algorithme d'étiquetage à la ALEPH**

- Mesure la **compatibilité des traces avec le PV**

- Calcul du poids :

- Pour chaque trace : $P_i = \int_{-\infty}^{\infty} \mathcal{R}(x) dx$

- Poids du jet : $P_{jet} = P_0 \sum_{j=0}^{N_{trk}-1} \frac{(-\ln P_0)^j}{j!}$ where : $P_0 = \prod_{i=1}^{N_{trk}} P'_i$

and

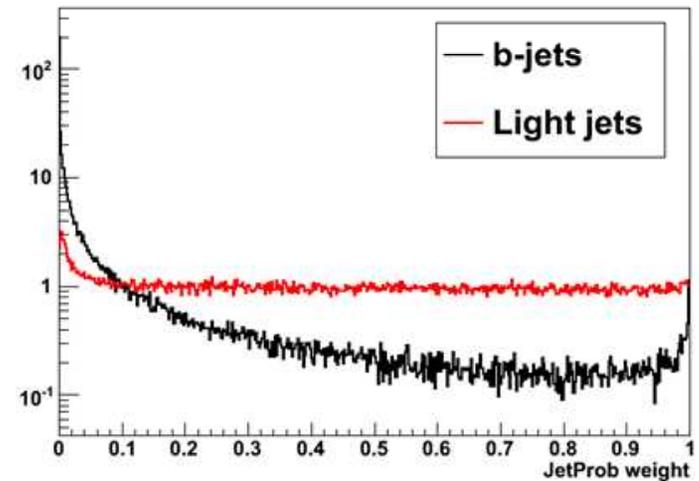
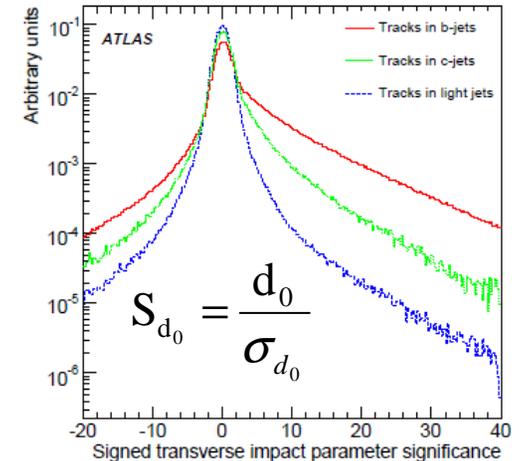
$$\begin{cases} P'_i = \frac{P_i}{2} & \text{if } d_0^i > 0 \\ P'_i = (1 - \frac{P_i}{2}) & \text{if } d_0^i < 0 \end{cases}$$



- **Calibration : Fonction de résolution R**

- Requier la distribution de S_{d_0} des traces venant du PV
- Peut être assez facilement extraite des données

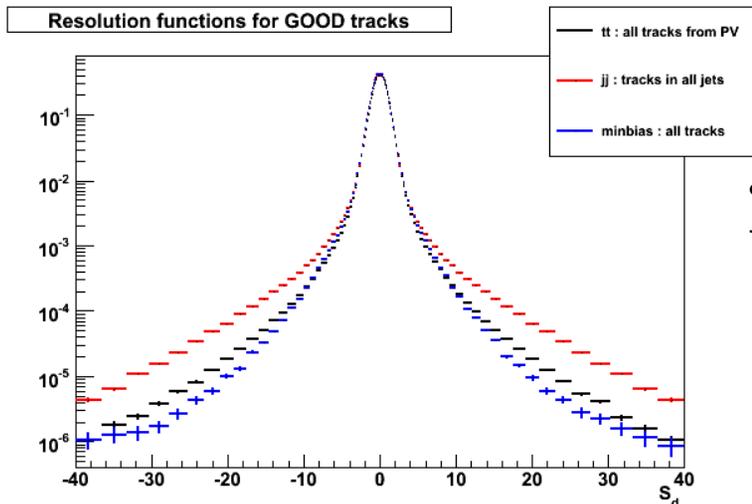
- En utilisant un lot avec un contenu en saveur lourde le plus petit possible
 - dijets @900 GeV ou même événements de biais minimum
- En symétrisant la distribution obtenue en n'utilisant que les traces de $d_0 < 0$ (contamination en saveurs lourde reduite)



Slide suivant

Stratégies pour l'extraction de la fonction de résolution

- **Cas idéal : traces venant du PV dans tt (nécessite de connaître la vérité)**
- Stratégies étudiées :
 - **Événements de biais minimum : toutes les traces**
 - Avantage : extraction de R avec les toutes premières données, avant la bonne compréhension des jets (mais signe de d_0 pas défini)
 - **Événements dijets : toutes les traces dans les jets** (essentiellement des légers) :



• Distributions des traces Good :
tt et minbias compatibles

Fraction de jets b, c et légers ($p_T > 15$ GeV, $|\eta| < 2.5$):

Lot	b-jet	c-jet	l-jet
tt	37%	8%	55%
jj	3%	6%	91%
min bias	2%	5%	93%

- **Performances :**
- Lot test : tt
- Avec différentes fonctions de résolution

Rejet des jets légers pour différentes efficacités de b-tagging

	tt		minbias		jj	
	50 %	60 %	50 %	60 %	50 %	60 %
	53.0 ± 0.2	22.86 ± 0.05	53.8 ± 0.2	23.58 ± 0.05	48.7 ± 0.2	21.04 ± 0.05

Fonction de résolution extractible des événements de biais minimum

Extraction avec données @900GeV : étude en cours dans le MC

Nouvelles méthodes (en cours de développement)

- Utilisation des traces de $d_0 > 0$ uniquement pour le calcul du poids de JetProb
- Pour tenir compte de la **plus grande multiplicité des traces dans les jets b** par rapport aux jets légers :
 - Introduction d'un **facteur $\alpha^{N_{trk}}$** (perf. optimales avec $\alpha=0.5$)
 - Utilisation d'une **distribution de la multiplicité des traces dans les jets légers ($R \otimes h$)**
- "JetProb with track pruning" (**élaguage**) : **poids recalculé en enlevant 1, voire 2 traces**
→ décision du poids selon # traces dans jet

Méthodes
détaillées en
back-up

Stratégie	50 %	60 %
Toutes les traces	56.8 ± 0.9	23.3 ± 0.2
Seulement $d_0 > 0$	76.7 ± 1.4	27.4 ± 0.3
Toutes les traces + $0.5^{N_{trk}}$	104.7 ± 2.2	36.5 ± 0.4
Toutes les traces + $R \otimes h$	108.4 ± 0.5	35.3 ± 0.1
Toutes les traces + JP with track pruning	120.1 ± 0.6	35.2 ± 0.1

**Performances
améliorées d'un
facteur 2**

Compromis à trouver entre **bonnes performances** et **simplicité**
(JetProb devant rester un algorithme d' étiquetage simple)

- MC à valider avec les données

Mesure des performances dans les données

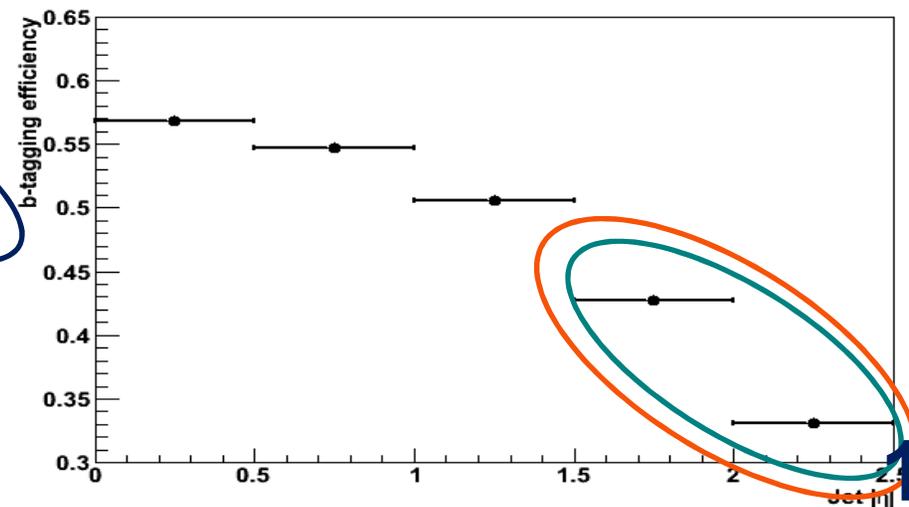
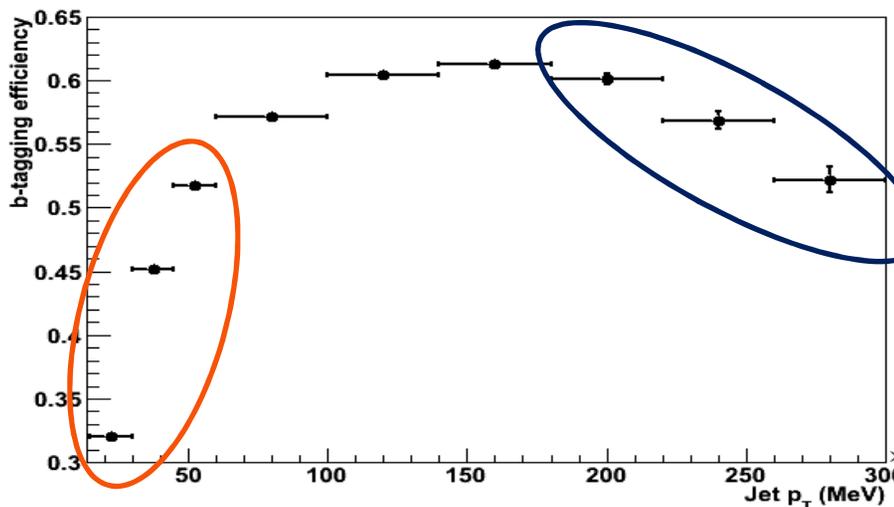
- Mesure de **l'efficacité de b-tagging** :
 - En utilisant des événements dijet (jets contenant un muon) : méthode p_T^{rel}
 - **En prenant en compte les dépendances à la cinématique pour pouvoir appliquer le b-tagging à n'importe quel canal (slide suivant)**

Dépendance cinématique

- Plus de **diffusion multiple et d'interactions secondaires** à bas p_T et à grand $|\eta|$
- **Mauvaise résolution sur z_0** à grand $|\eta|$
- Problèmes à grand p_T :
 - **jets denses** (traces plus collimées) :
 - reconnaissance des schémas des traces plus difficile
 - signe de d_0 aléatoire
 - **désintégration des B tardive** : e.g. $p_T \sim 200$ GeV :
 - $\sim 8\%$ des B se désintègrent après la b-layer
 - sélection des traces ($|d_0| < 1$ mm, $N_{\text{BLayer}} > 1$)
 - trop sévère pour les grands boosts

Calibrations : BTagCalib-02-01
Lot de test : ttbar @10TeV (500 000 evts)
Algorithme de reconstruction des jets :
standard actuel d'ATLAS : cône de $\Delta R = 0.4$

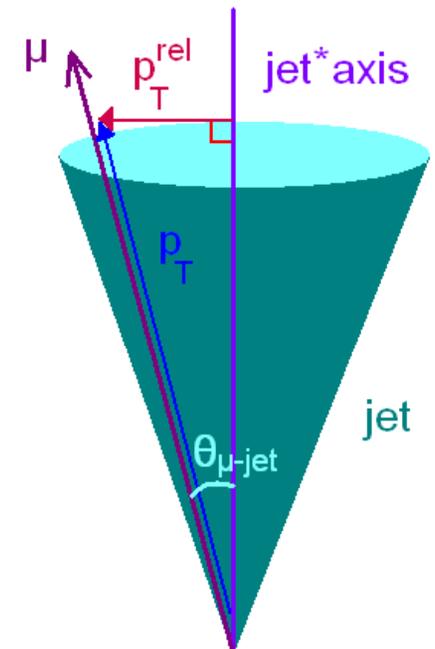
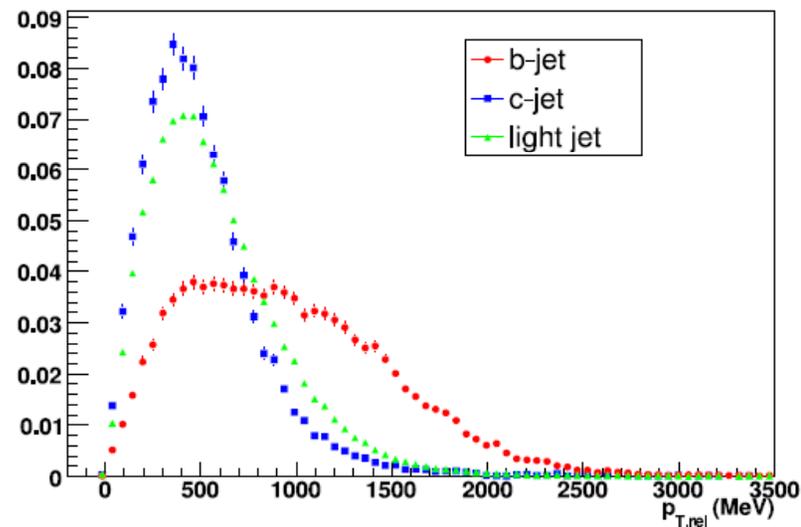
Efficacité de b-tagging en utilisant JetProb pour une coupure fixe ($w > 0.005$) VS p_T , $|\eta|$:



Méthode p_T^{rel} : introduction

- Permet de **mesurer l'efficacité d'étiquetage de jets b semileptoniques** dans les données
- Basée sur la propriété **des hadrons B de se désintégrer semileptoniquement**
- N'utilise que les jets dans lesquels un muon a été reconstruit
 - Variable : p_T^{rel} ** du muon : $p_T^{\text{rel}} = p \times \sin(\theta_{\mu - \{jet + \mu\}})$
 - Discrimine jets **b** et **jets de saveur moins lourde** :

**impulsion du muon associé au jet dans le plan perpendiculaire à l'axe du système $\{\mu + jet\}$



*en prenant en compte le muon pour déterminer l'axe du jet

- **Méthode déjà utilisée au TeVatron**
- Etude menée dans **ATLAS @ 14 TeV** (CSC book)
- **Mon étude a été faite pour une simulation @ 10TeV**
- L'extrapolation à **tous les jets** nécessite le MC et fait l'objet d'une autre étude

Méthode p_T^{rel} : démarche

Idée générale : fit des fractions de b, c et l avant et après tagging grâce aux distributions p_T^{rel}

- Méthode pour obtenir l'efficacité de b-tagging semileptonique :
 - Lot de dijets contenant un muon (\rightarrow enrichi en saveurs lourdes) obtenu grâce à un trigger dédié
 - Extractions de templates p_T^{rel} pour les jets b, c (MC) et légers
 - Extraction des distributions de p_T^{rel} de tous les jets et des jets taggés
 - Utilisation des templates p_T^{rel} pour les jets b, c (MC) et légers
 - Mesure du contenu en jets b en fittant les templates des b avant et après tagging (TFractionFitter)

$$\epsilon_b^{\text{data},i} = \frac{N_{\mu\text{-jet}}^{\text{tag},i} F_b^{\text{tag},i}}{N_{\mu\text{-jet}}^i F_b^i}$$

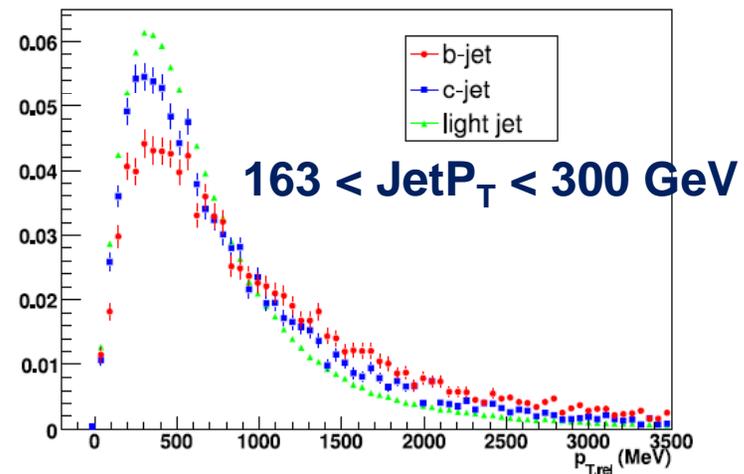
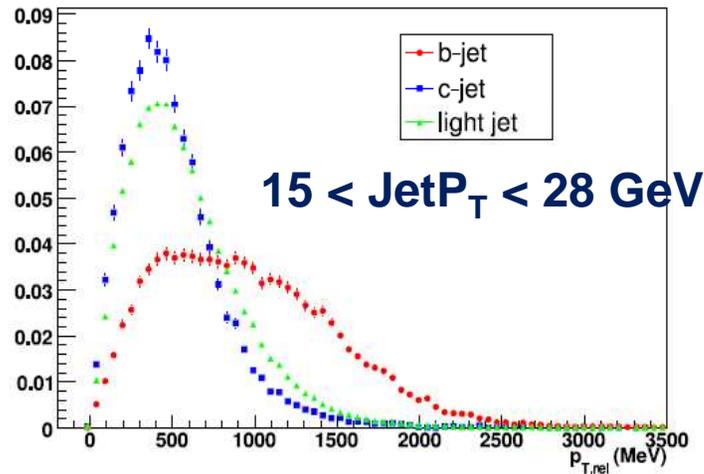
Avec :

- $N_{\mu\text{-jet}}^{\text{tag},i}$ • Nombre de jets taggés contenant un muon
- $F_b^{\text{tag},i}$ • Fraction de jets taggés b parmi eux
- $N_{\mu\text{-jet}}^i$ • Nombre de jets contenant un muon
- F_b^i • Fraction de jets b parmi eux
- $F_b^{\text{tag},i}$ and F_b^i • Sont les 2 inconnues obtenues en fittant les templates aux données

En bins de p_T et $|\eta|$ étant donnée la forte dépendance du b-tagging à ces variables

Méthode p_T^{rel} : problèmes

- Templates p_T^{rel} c et légers~ identiques → **problèmes de fit**
- Extraction templates b et c nécessite **MC**
- **Méthode ne fonctionne que pour des jets de $p_T < 80$ GeV**
 - car traces + collimées dans jets à grand p_T



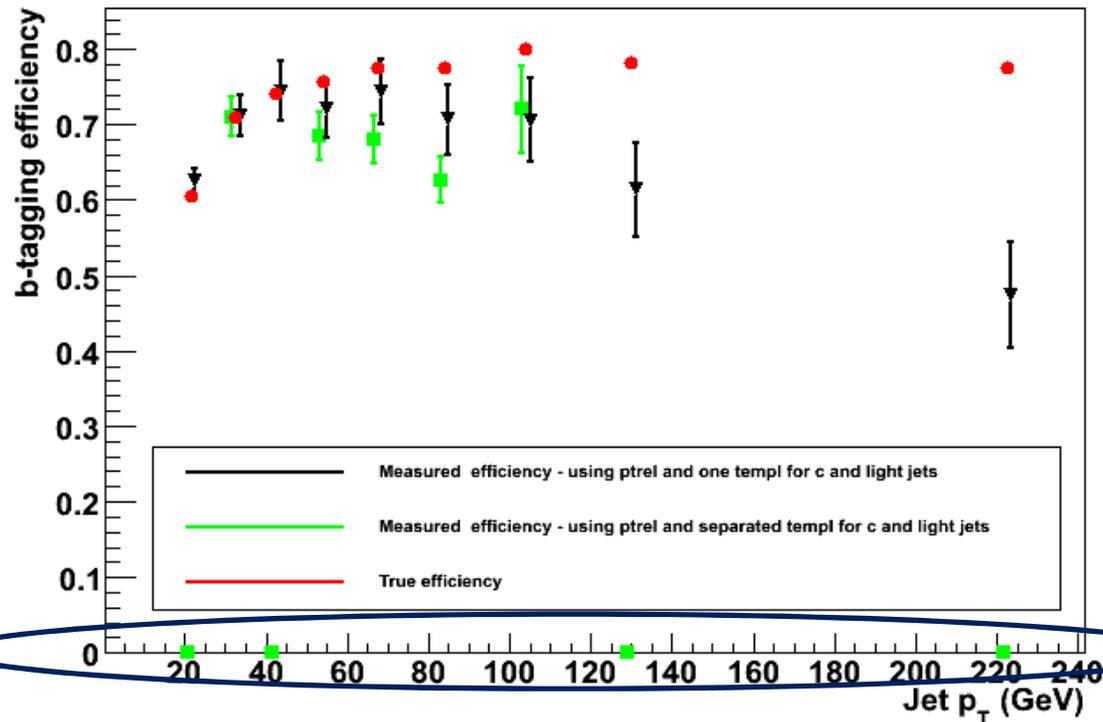
• Améliorations :

- **1 seul template** pour jets c et légers : fit avec plus grand succès
- Utilisation de **fonctions analytiques** (en cours d'étude)

Slides
suivants

Méthode p_T^{rel} : résultats

- Résultats avec **2 templates (default)** et **1 template** pour c et l
 - Comparaison à **l'efficacité** calculée à partir de la **vérité**



- En utilisant 1 seul template :
 - fit fonctionne
 - **bon accord MC/méthode jusqu'à 80 GeV**



Fit a echoué

- **Détermination des erreurs** : étude pas encore effectuée
- Résultats obtenus dans l'étude ATLAS :
 - Erreurs @14 TeV pour 100pb^{-1} :
 - **Statistiques bien inférieures à 1%**
 - **Systématiques contrôlables à 6% (abs.)** - dues à :
 - Utilisation du MC
 - Contamination en saveurs lourdes dans le template de jets légers

Méthode p_T^{rel} : utilisation de fits

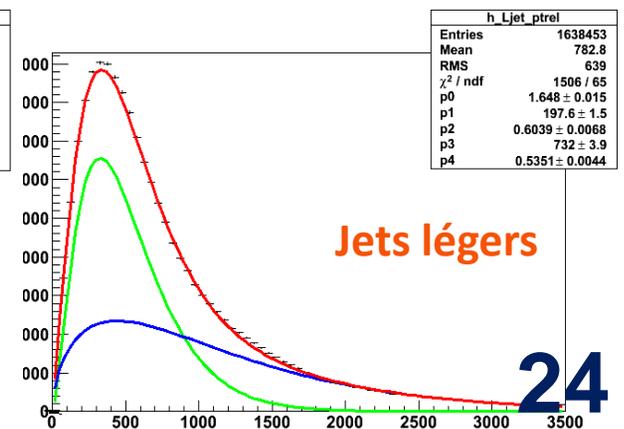
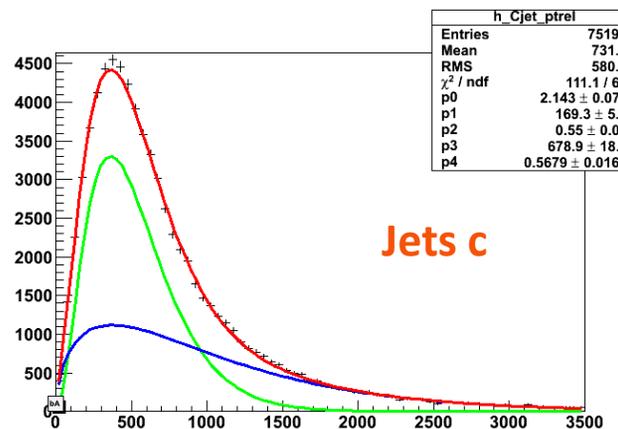
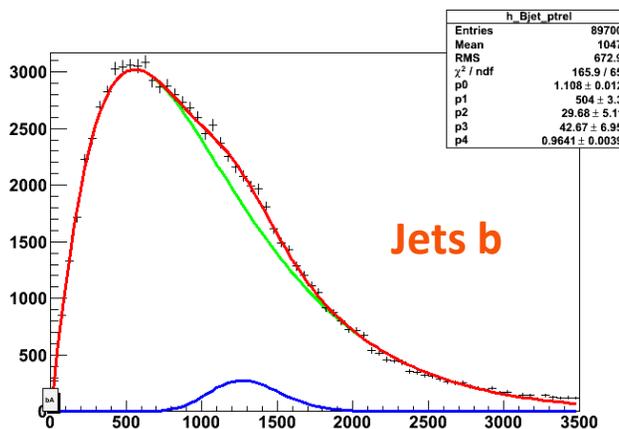
- Etude commencée très récemment
- Déjà appliquée avec succès dans **CDF**
- **Fonctions analytiques** utilisées : sommes de **2 fonctions gamma incomplètes** :

$$f(p_T^{rel}) = Nf_1 \underbrace{\left(\frac{(p_T^{rel})^{\alpha_1+1} e^{-\frac{p_T^{rel}}{\beta_1}}}{\beta_1^{\alpha_1+1} \Gamma(\alpha_1+1) \int_0^{3500} x^{\alpha_1+1} e^{-x} dx \int_0^\infty x^{\alpha_1+1} e^{-x} dx} \right)}_{G1} + N(1-f_1) \underbrace{\left(\frac{(p_T^{rel})^{\alpha_2+1} e^{-\frac{p_T^{rel}}{\beta_2}}}{\beta_2^{\alpha_2+1} \Gamma(\alpha_2+1) \int_0^{3500} x^{\alpha_2+1} e^{-x} dx \int_0^\infty x^{\alpha_2+1} e^{-x} dx} \right)}_{G2}$$

- **5 paramètres** : $f_1, \alpha_1, \beta_1, \alpha_2, \beta_2$
- N=facteur de normalisation

- **Décrivent très bien les templates** (points noirs) : ✓

Rouge : fit global
Vert et bleu : contributions de G1, G2



Méthode p_T^{rel} : utilisation de fits

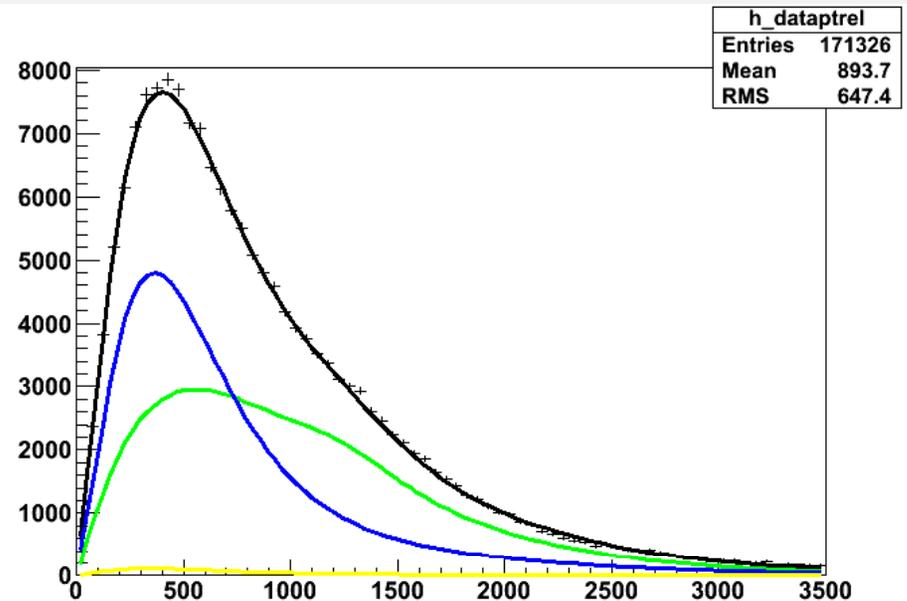
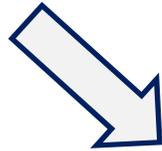
Points noirs : distribution p_T^{rel} “data like”

Vert : fraction de b en utilisant la fonction analytique

Bleu : fraction de c en utilisant la fonction analytique

Jaune : fraction de légers en utilisant la fonction analytique

Fit noir : somme des fractions de b, c et légers



Fit ok même avec des distributions de c et légers séparées

$\epsilon_b(\%)$	Bins de p_T (GeV) (Globale = tous p_T)						
	Globale	15-28	28-40	40-50	50-60	60-85	85-115
Vraie	72.9	60.5	71.6	74.5	75.4	77.4	79.4
Mesurée avec templates	X	X	72.1 ± 4.0	76.5 ± 9.9	69.9 ± 7.1	65.6 ± 3.0	X
Mesurée avec fct fittée	71.0	62.0	71.4	71.1	70.3	72.9	66.6

*Calcul des erreurs en cours



Conclusion et Perspectives

- **Optimisation de JetProb**

- **Stratégies** pour l'extraction de la fonction de résolution dans les données
- **Nouvelles méthodes** : performances doublées



- **Mesure de l'efficacité de b-tagging dans les données**

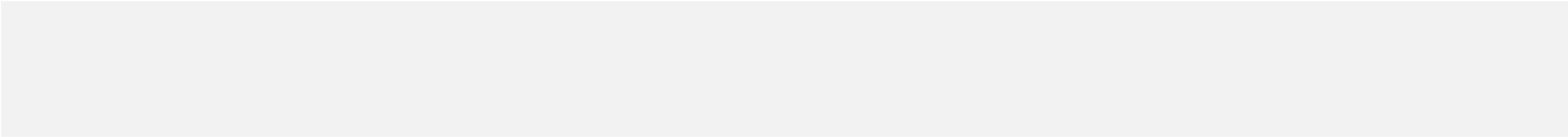
- **Méthode p_T^{rel}** : dès $\sim 20\text{pb}^{-1}$
 - Efficacité bien mesurée pour des jets de $p_T < 80\text{ GeV}$
 - au-delà : méthode utilisant des lots $t\bar{t}$ bien reconstruits (nécessite 100 à 200 pb^{-1})



- **Perspectives :**

- **Collisions à 900 GeV :**
 - Tracking
 - Commissioning de JetProb :
 - extraction de la fonction de résolution dans les données
- Préparation à la mesure du **taux de mauvais étiquetage** dans les données





Back-up

INTRODUCTION : what is pTrel ?

pTrel is defined as :

$$p_T^{rel} = p \sin(\theta_{\mu-jet})$$

pTrel is a discriminating variable of b-jets against lighter flavoured jets

Why ?

Let's suppose the muon takes the whole momentum :

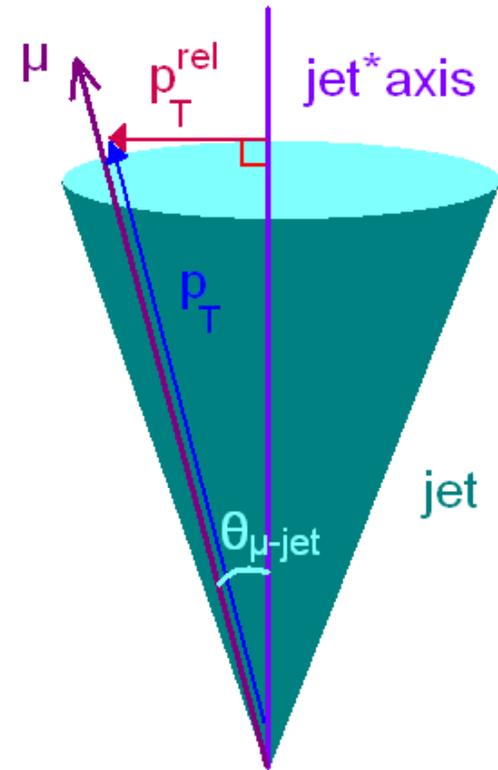
$$p_T^{rel} Max \approx M_{Hadron} - M_{ProductsDecay}$$

Thus, b quarks being heavier than lighter flavoured ones, their maximum pTrel is bigger

Typically :

b jets : max pTrel = 2.5 GeV

c and light jets : max pTrel = 1 GeV **BUT...**

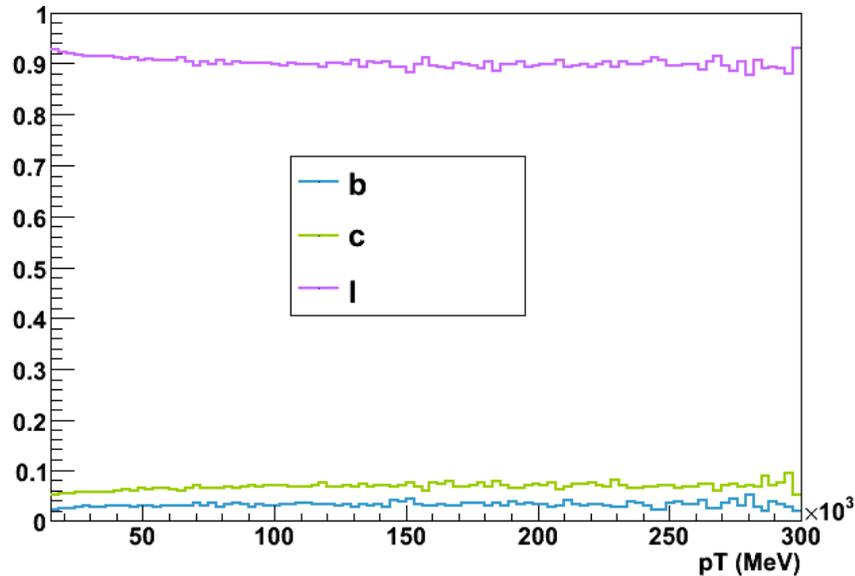


*muon taken into account to define the jet axis

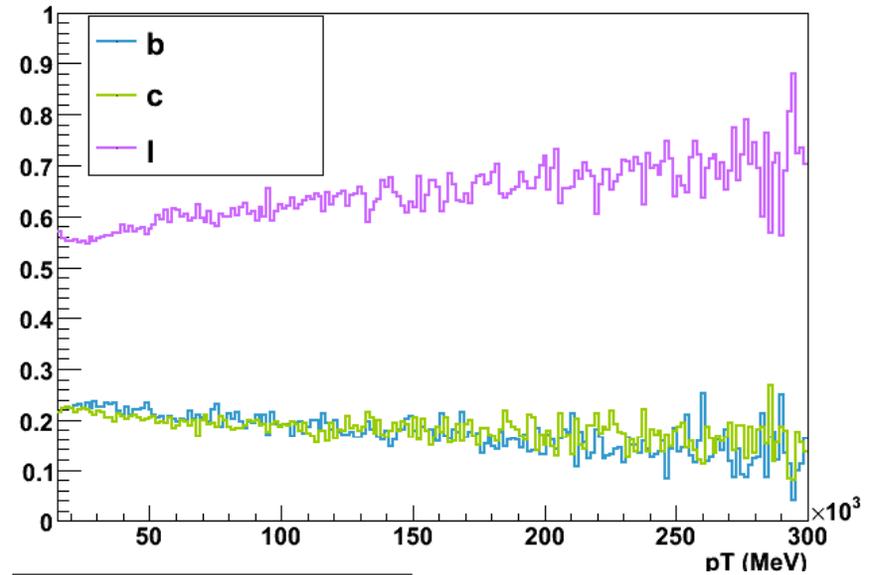
INTRODUCTION : dijets events study

Fraction of b, c and light jets before (left) and after (right) requiring* à muon in dijets events :

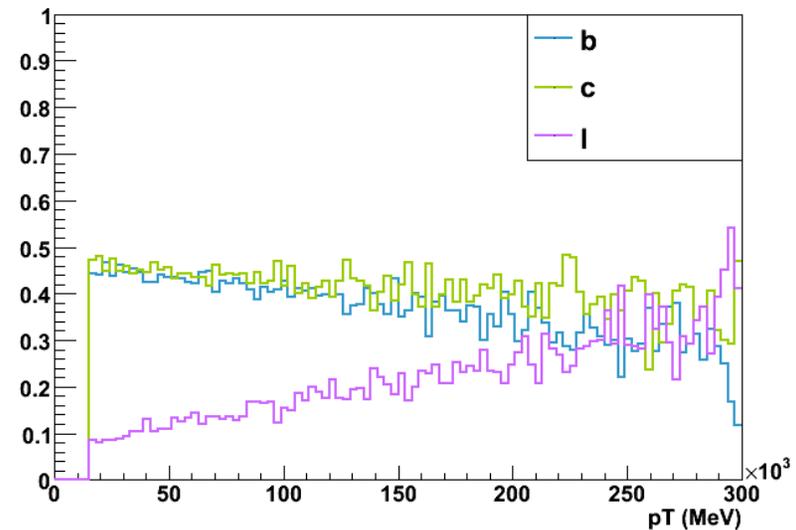
Flavour rates without muon filter



Flavour rates with muon filter



Flavour rates with muon filter



Fraction of b, c and light jets before (left) and after (right) requiring* à muon in jets :
(most muons associated to light jets come from the contamination of heavier flavoured jets)

**the muon requirement is made at the generator level (pythia)*

In data, à trigger will be used to select events with at least one jet containing à muon.

INTRODUCTION : the pTrel method

This method was used in the CSC note (BT10), and my analysis is based on it

-The semi-leptonic b-tagging efficiency is obtained by measuring the b-jets content before and after tagging à sample of dijets containing à muon.

-Use à fraction fitter

-Binned in pt and η

-The b-tagging efficiency in the i'th jet pT and η bin is given by :

$$\mathcal{E}_b^{data,i} = \frac{N_{\mu-jet}^{tag,i} F_b^{tag,i}}{N_{\mu-jet}^i F_b^i}$$

$N_{\mu-jet}^{tag,i}$ Where :

$F_b^{tag,i}$ is the number of tagged jets containing muons in the i'th bin

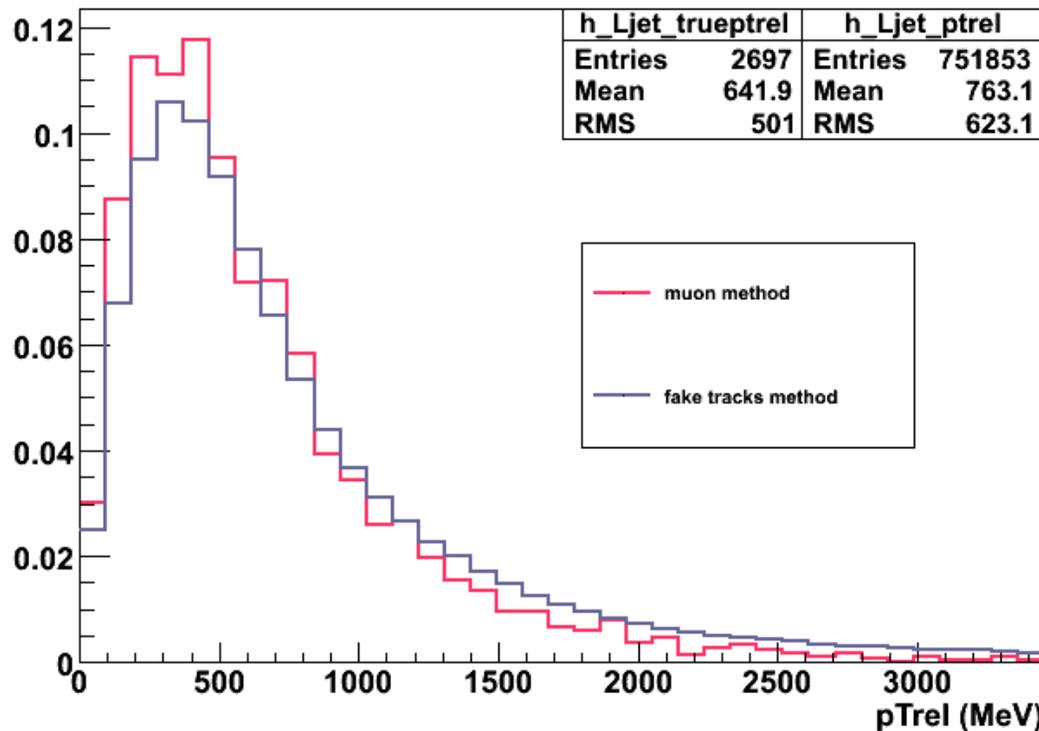
$N_{\mu-jet}^i$ is the fraction of tagged b jets among them

F_b^i is the number of jets containing muons in the i'th bin

$F_b^{tag,i}$ and F_b^i is the fraction of b jets among them

are the two unknowns obtained by fitting templates to the data (before and after tagging) using à discriminating variable :

LIGHT JETS TEMPLATE



- This method uses all tracks able to fake à muon is used because of à lack of statistics
 - On the plot above, the distributions of the p_{Trel} of all the tracks able to fake à muon in purified light jets and of the muons associated to purified light jets
 - The distribution are about the same
 - Method valid, but for data :
 - b and c jets templates will be taken from MC
 - templates for light jets extracted from data
- an equivalent to the purification has to be found

Detecteur à pixels

↻ Résolution en $R\phi$ sur d_0 :

$$\sigma(d_0) = A \oplus \frac{B}{P_T}$$

↻ A : résolution intrinsèque $A = \frac{50}{\sqrt{12}} = 14,5 \mu m$
:
pour un système binaire
MAIS partage de la charge
entre plusieurs pixels $\rightarrow A = 11,5 \mu m$

↻ B/P_T : diffusion multiple à bas P_T . Ex : impulsion typique d'une trace dans un jet b : $5 \text{ GeV} \rightarrow$
résolution totale : $\sigma(d_0) = 40 \mu m$

Perspectives du b-tagging

Etapes de commissioning du b-tagging :

-comprendre le tracking – vertexing

1nb-1

base

-mise en place des algorithmes de

~5pb-1

-eventuellement : Calibration avec les données (extraction fonctions ref)

taggers

-mesure des performances des early

~50pb-1

Mon travail :

traces

-Optimisation de la selection des

-defintion des categories de traces

-Améliorations de JetProb

-stratégies pour l'extraction de R dans les données

-mesure de l'efficacité de b-tagging dans les données (ptrel)

Ingredients clés pour les performances de b-tagging

Les performances de b-tagging sont sensibles à la cinématique des jets, au tracking, à l'alignement et à la matière dans l'ID

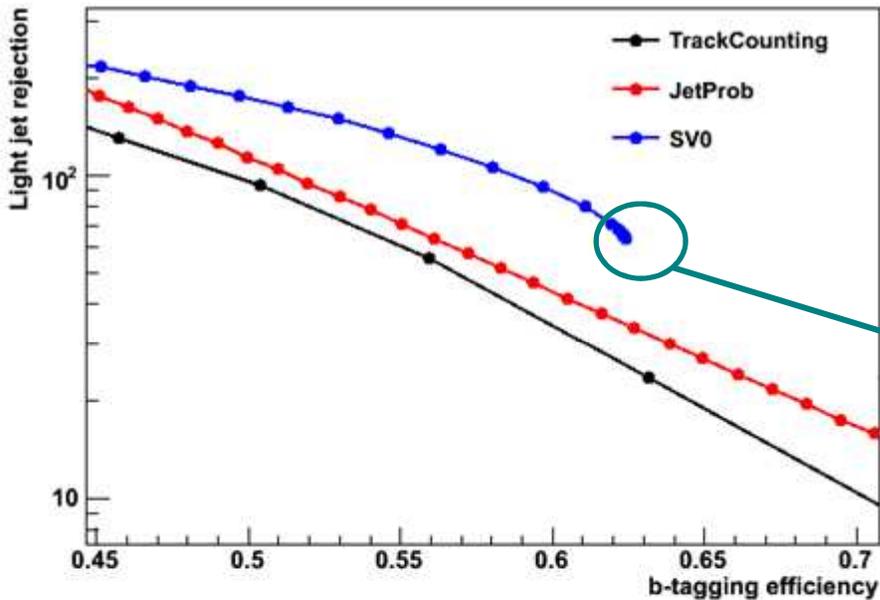
- Effet de certains de ces éléments sur le b-tagging

Early taggers performance

efficacité de b-tagging

Rejet des jets légers pour différentes efficacités de b-tagging

	50%	60%
JetProb	114±1	43.7±0.3



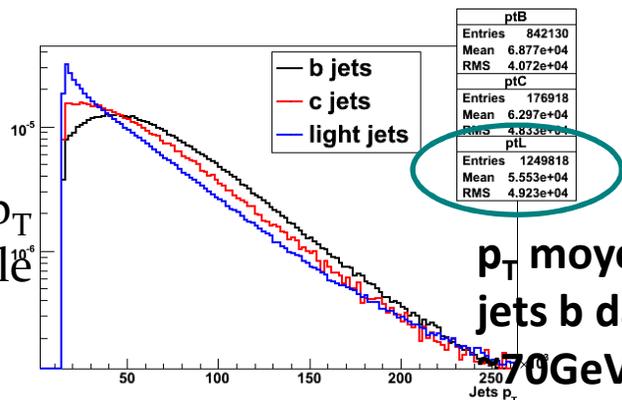
Calibrations : BTagCalib-02-01

Lot de test : ttbar @10TeV (500 000 evts)

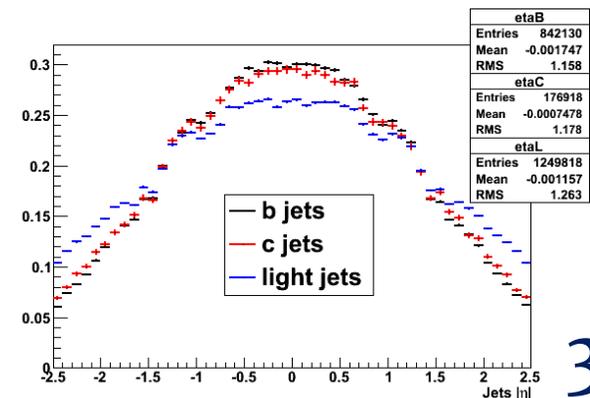
Algorithme de reconstruction des jets : standard actuel

d'ATLAS : Cone de DR =0.4

Distributions de p_T et η des jets dans le lot ttbar :



p_T moyen des jets b dans tt : 70GeV



$\alpha^{N_{trk}}$ method

Naïve approach to compute JetProb : $P_{jet} = \prod_{i=1}^{N_{trk}} P_i'$

-An ambiguity appears :

Jet with 1 track with $P_i=10^{-5}$ -> $P_{jet}=10^{-5}$ more likely to be b than
jet with 5 tracks with $P_i=10^{-1}$ -> $P_{jet}=10^{-5}$

→ Introduction of : $P_{jet} = P_0 \sum_{j=0}^{N_{trk}-1} \frac{(-\ln P_0)^j}{j!}$

But if N_{trk} large, P_{jet} tends to 1 : $\lim_{N_{trk} \rightarrow +\infty} \sum_{j=0}^{N_{trk}-1} \frac{(-\ln P_0)^j}{j!} = e^{-\ln P_0} = \frac{1}{P_0} \Rightarrow \lim_{N_{trk} \rightarrow +\infty} P_{jet} = 1$

Whereas b-jets have larger **track multiplicity distribution** than light jets :

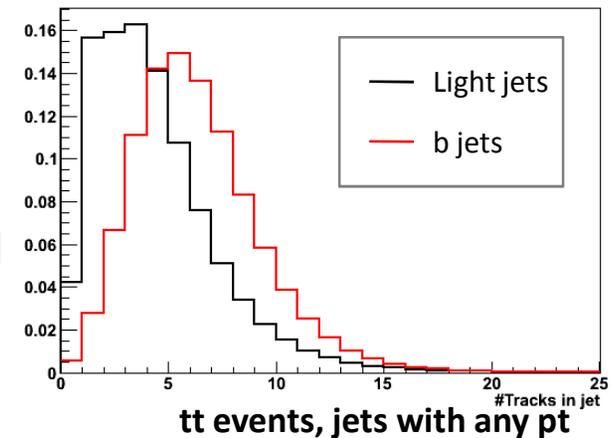
→ Introduction of $\alpha^{N_{trk}}$ ($\alpha < 1$) :

$$P_0 = \prod_{i=1}^{N_{trk}} \alpha \times P_i' = \alpha^{N_{trk}} \prod_{i=1}^{N_{trk}} P_i' \Rightarrow P_{jet} = \alpha^{N_{trk}} P_0 \sum_{j=0}^{N_{trk}-1} \frac{(-\ln(\alpha^{N_{trk}} P_0))^j}{j!}$$

Jets with à larger number of tracks have their weight reduced

Drawback : free parameter α

-optimal value for this analysis : $\alpha = 0.5$



R⊗h method

Trying to build on what was understood using $\alpha^{N_{trk}}$

Introduces a new term to compute P_0 taking into account the track multiplicity in light jets :

$$P_{jet} = P_0 \sum_{j=0}^{N_{trk}} \frac{(-\ln P_0)^j}{j!} \quad \text{where} \quad P_0 = \prod_{i=1}^{N_{trk}} P'_i \times \int_{N_{trk}}^{+\infty} h(x) dx$$

-Requires distribution of # of tracks in light jets in tt events.

For this analysis :

-use of truth (tracks in light jets)

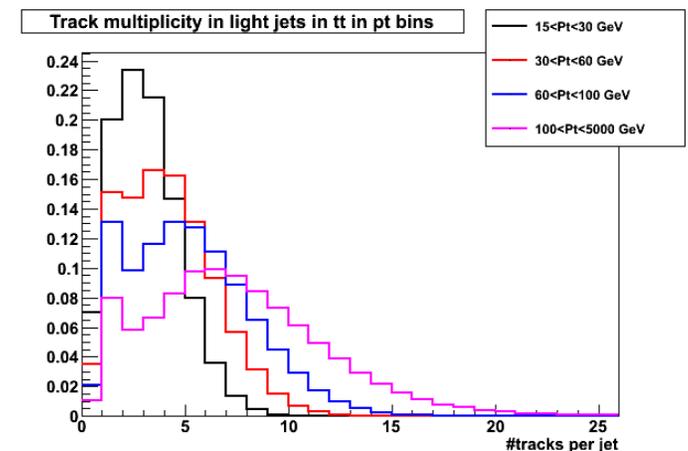
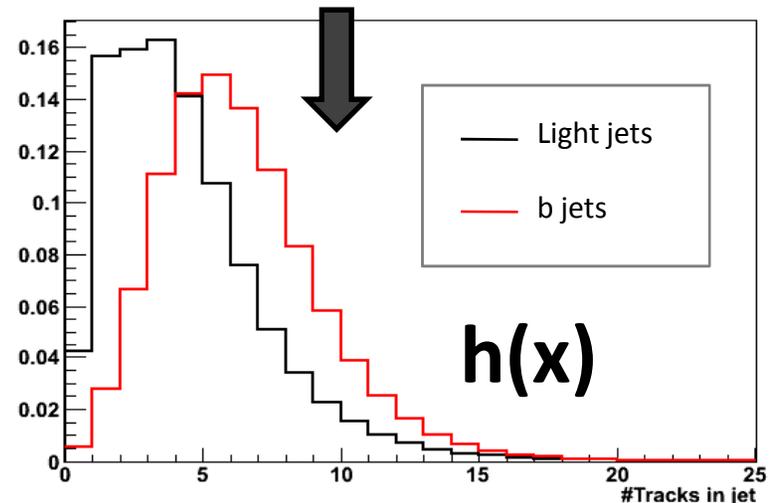
strategies for data :

-similar as extracting R :

-using jj or minimum bias events

-using tracks with $d_0 < 0$ multiplicity (times 2)

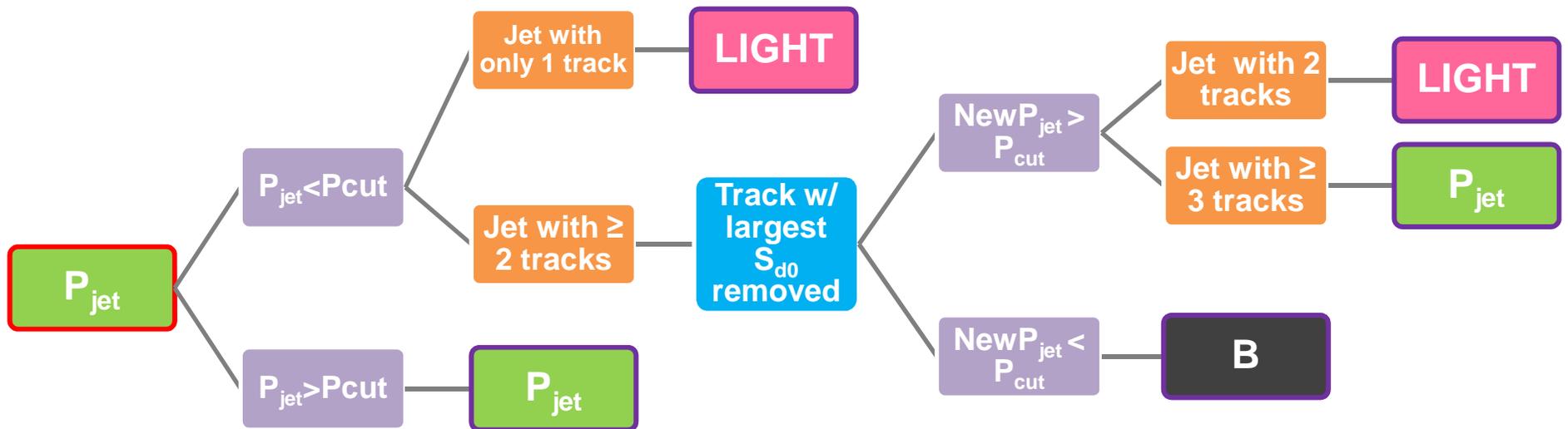
-Distributions must be binned in jet pt to isolate the track multiplicity effect



JetProb w/ track pruning

Method :

- Jets with JetProb weight $<$ defined weight P_{cut}
- Set 1-track jets as light
- Study JetProb weight after removal of track with largest S_{d0}



- LIGHT : final weight set to 1
- B : final weight set to 10^{-21}
- P_{jet} : Original weight
- Original weight
- Final weight

Drawbacks :

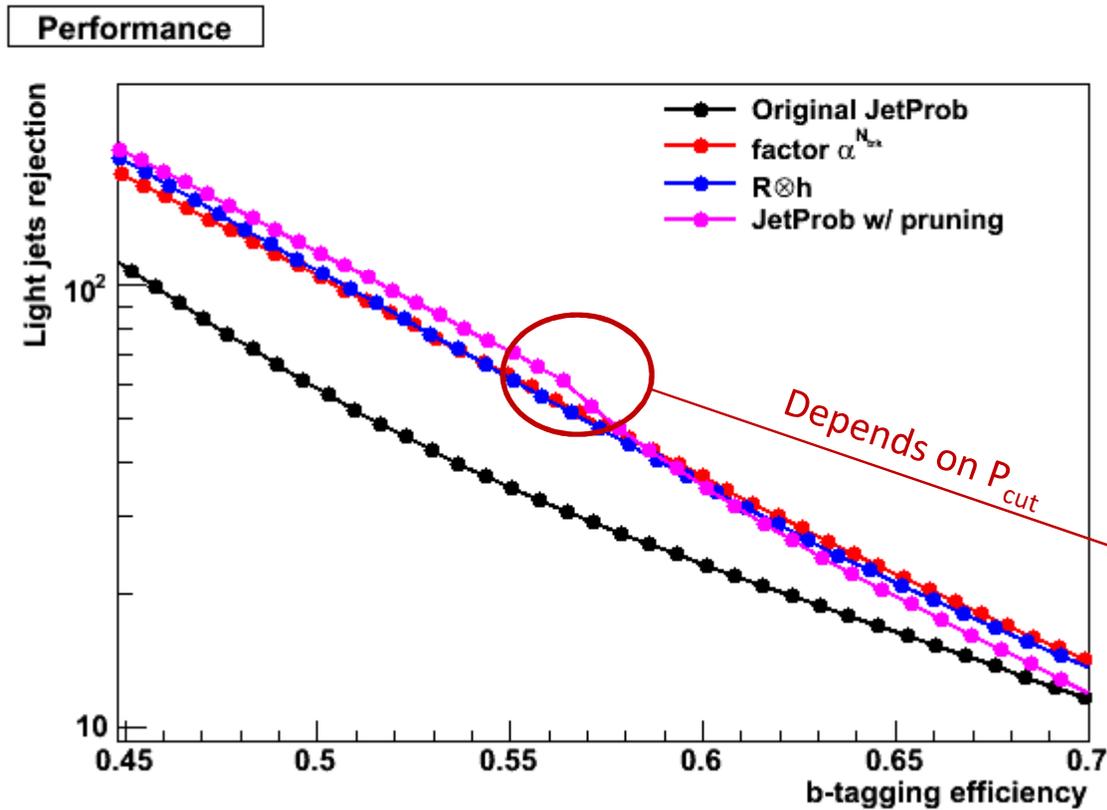
- Free parameters P_{cut} , weight for b-jets
- P_{cut} tuned for every b-tagging efficiency (see performance plot)

Performance

Test sample : **tt** (1 930 403 evts)

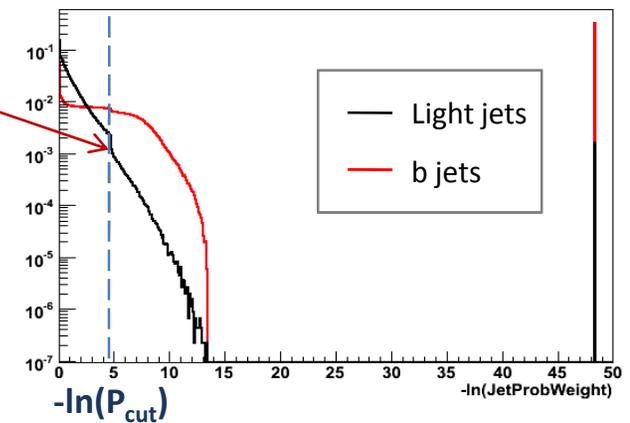
Resolution function : extracted from **tt** events using **all tracks** from **PV** (using truth)

Track categories : **good** and **shared**



Performance x2 with new methods

Light jet rejection for diff. b-tag eff	50 %	60 %
Original JetProb	58.3 ± 0.2	23.45 ± 0.05
factor $\alpha^{N_{trk}}$	105.4 ± 0.5	36.8 ± 0.1
$R \otimes h$	108.4 ± 0.5	35.3 ± 0.1
JP w/ pruning	120.1 ± 0.6	35.2 ± 0.1

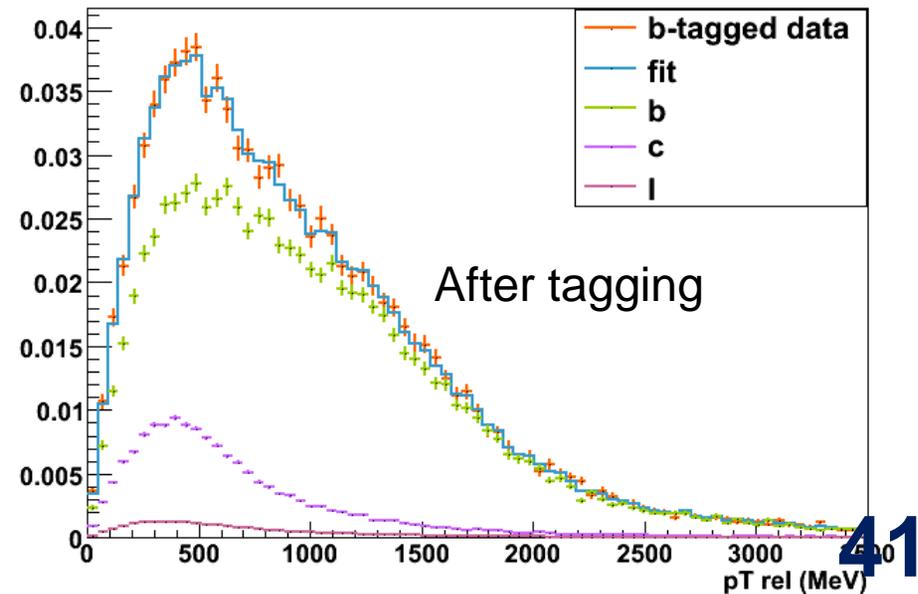
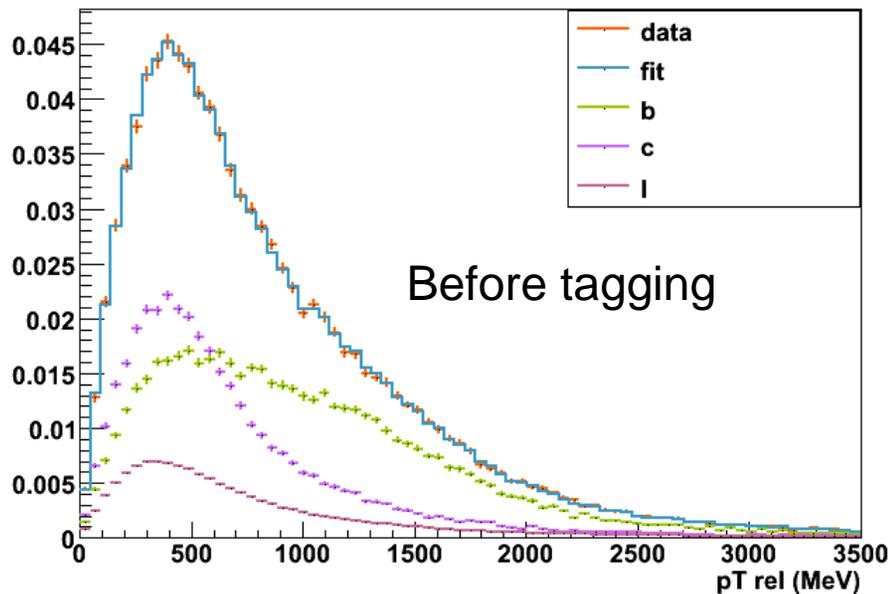


p_T^{rel} method : analyse

- Events :
 - dijet @ 10TeV
 - statistique :
- Etapes dans le MC:
 - 1ere moitié des données : Creation des templates
 - 2nde moitié : remplissage des histos dans les conditions de prise de données
 - Fit des données avec les templates en utilisant TFractionFitter
 - efficacité de b-tagging calculée et comparée à celle obtenue directement avec la vérité
- Jets used :
 - Pt>15 GeV
 - $|\eta| < 2.5$
 - label 4 for c-jets, 5 for b-jets and 0 for light-jets
 - light-jets DR=0.8 away from b, c and tau in truth (purified)
- b and c-jets templates :
 - filled with the p_T^{rel} of all muons with $p_T > 4\text{GeV}$ associated to the jet in a DR=0.4 cone around the jet axis
- light jets template :
 - p_T^{rel} of all tracks within a cone of DR=0.4 around the jet axis and with a p_T above 4 GeV
(assuming that all the tracks have a uniform probability of faking a muon)
 - For the validity of this method, see next slide

p_T^{rel} method : analyse

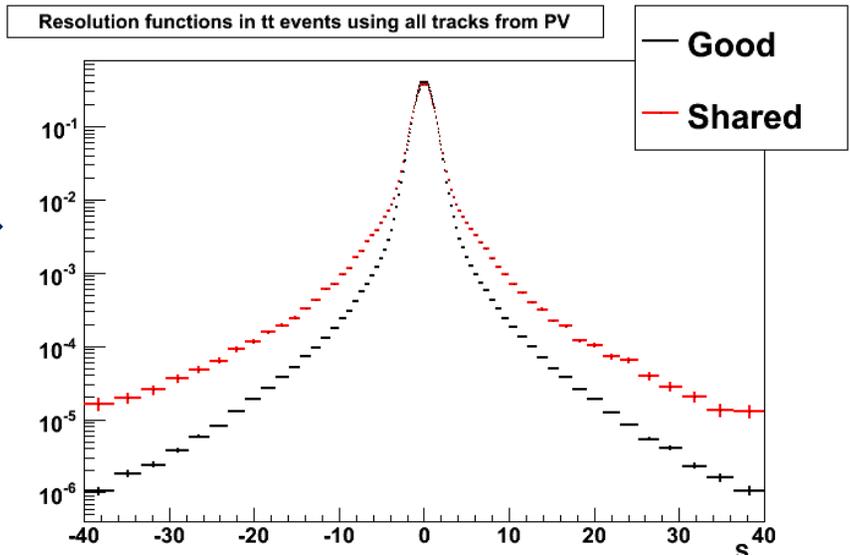
- Events :
 - dijet @ 10TeV
 - statistique :
- Etapes dans le MC:
 - 1ere moitié des données : Creation des templates
 - 2nde moitié : remplissage des histos dans les conditions de prise de données
 - Fit des données avec les templates en utilisant TFractionFitter
 - efficacité de b-tagging calculée et comparée à celle obtenue directement avec la vérité
- Exemples de fits :



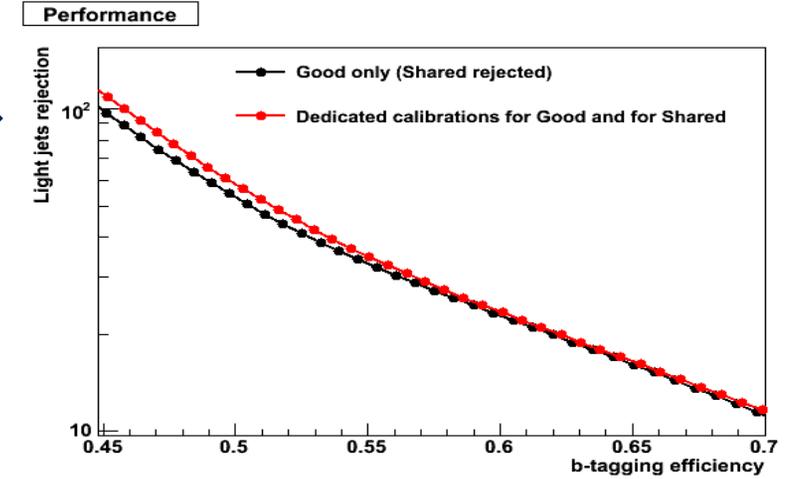
Categories de traces : Good - Shared

Defaut dans JetProb

- Distributions de S_{d0} des traces **Good** et **Shared** dans les événements tt :
 - Performances* :
- Lot test** : tt (1 930 403 evts)
Fonction de resolution : extraite de tt en utilisant **toutes les traces** issues du **PV**



Rejet des jets légers pour différentes ϵ_b	50 %	60 %
Good (Shared rejetées)	53.0 ± 0.2	22.86 ± 0.05
Calibrations dédiées pour Good et Shared	58.3 ± 0.2	23.45 ± 0.05



*La mesure des performances se fait avec :

- l'efficacité de b-tagging : rapport du nombre de jets b taggés b sur le nombre total de jets b
- le rejet des jets légers : rapport du nombre de jets légers sur le nombre de jets légers taggés b

+10% en utilisant des calibrations séparées pour les Shared



+ de categories en cours d'étude

Introduction

- **Objectif de mon travail** : préparation à la **mesure de la section efficace $t\bar{t}$** en utilisant la méthode « cut and count » et le b-tagging :

$$\sigma_{t\bar{t}} = \frac{N_{obs} - N_{bkg}}{A \varepsilon L}$$

N_{obs} : #d'événements passant les coupures
 N_{bkg} : #d'événements de bruit de fond attendu passant ces coupures
A : acceptance du détecteur
 ε : efficacité de sélection (prend en compte l'efficacité de trigger, de b-tagging, etc.)
L : luminosité intégrée

- **Pour être réalisable avec les 1eres données cette étude requiert :**

- Une bonne mesure de la luminosité intégrée
- Un bon accord MC/données pour estimer N_{bkg}
- Commissioning du b-tagging
- Mesure dans les données :
 - de l'efficacité de b-tagging
 - du taux de mauvais étiquetage

C'est essentiellement autour de ces 3 points que s'articule mon travail (voir slide suivant)