
L'étiquetage de la saveur des mésons B dans LHCb

 LA DÉTERMINATION DE LA SAVEUR des mésons B neutres est fondamentale pour les études de la violation de la symétrie CP. En effet, pour mesurer l'asymétrie CP dépendante du temps, par exemple dans les désintégrations $B_d^0 \rightarrow J/\psi K_S^0$:

$$\mathcal{A}_{CP}(t) = \frac{\Gamma(B_d^0(t) \rightarrow J/\psi K_S^0) - \Gamma(\bar{B}_d^0(t) \rightarrow J/\psi K_S^0)}{\Gamma(B_d^0(t) \rightarrow J/\psi K_S^0) + \Gamma(\bar{B}_d^0(t) \rightarrow J/\psi K_S^0)} \approx \sin(2\beta) \sin(\Delta m_d t), \quad (3.1)$$

nous devons déterminer la saveur du méson beau à l'instant de sa création. Un algorithme d'étiquetage a été conçu [43] pour déterminer cette saveur le plus efficacement possible et le plus proprement possible. On verra dans la première section comment ces deux aspects ont un impact sur la sensibilité de $\sin(2\beta)$. La deuxième section décrira l'algorithme en lui-même avec ses limitations. Enfin dans la troisième section, une étude d'optimisation de l'algorithme sera présentée.

3.1 Définition

Un lot de désintégrations de mésons beaux se divise en trois catégories : non étiqueté (N), étiqueté de manière correcte (C) ou incorrecte (I). Le nombre total d'événements est $N_{\text{tot}} = N_C + N_I + N_N$.

L'efficacité d'étiquetage est définie comme le nombre d'événements étiquetés par rapport au nombre total d'événements :

$$\epsilon_{\text{tag}} = \frac{N_C + N_I}{N_{\text{tot}}},$$

et la fraction de mauvais étiquetage :

$$\omega = \frac{N_I}{N_C + N_I}.$$

Les quantités ε_{tag} et ω interviennent dans l'erreur statistique de l'asymétrie CP. En effet, le taux de désintégration mesuré dépendant du temps $\Gamma^m(t)$ d'un mode quelconque $B \rightarrow f$ est modifié par rapport au taux de désintégration théorique $\Gamma(t)$:

$$\begin{aligned}\Gamma^m(t) &= (1 - \omega)\Gamma(t) + \omega\bar{\Gamma}(t) \\ \bar{\Gamma}^m(t) &= \omega\Gamma(t) + (1 - \omega)\bar{\Gamma}(t)\end{aligned}$$

$\bar{\Gamma}^m(t)$ et $\bar{\Gamma}(t)$ sont les taux de désintégration CP conjugués.

L'asymétrie dépendante du temps devient

$$\mathcal{A}^m(t) = \frac{\bar{\Gamma}^m(t) - \Gamma^m(t)}{\bar{\Gamma}^m(t) + \Gamma^m(t)} = (1 - 2\omega)\mathcal{A}(t) = D\mathcal{A}(t), \quad (3.2)$$

où D est appelé dilution. Un étiquetage parfait donne une dilution $D = 1 - 2\omega = 1$ alors qu'un étiquetage aléatoire ($\omega = 0.5$) donne une dilution nulle.

Dans le cas de $\sin(2\beta)$, en combinant les équations 3.1 et 3.2, nous obtenons l'asymétrie expérimentale dépendante du temps :

$$\mathcal{A}_{B_d^0 \rightarrow J/\psi K_S^0}^m(t) = (1 - 2\omega) \sin(2\beta) \sin(\Delta m_d t).$$

L'erreur statistique sur l'asymétrie est donnée par

$$\sigma(\mathcal{A}) = \frac{\sigma(\mathcal{A}^m)}{1 - 2\omega}$$

De plus, l'erreur sur l'asymétrie mesurée est :

$$\sigma(\mathcal{A}^m)^2 = \frac{1 - (\mathcal{A}^m)^2}{(\Gamma^m + \bar{\Gamma}^m)} = \frac{1 - (\mathcal{A}^m)^2}{N^m} = \frac{1 - (\mathcal{A}^m)^2}{\varepsilon_{\text{tag}} N}$$

où N^m est le nombre d'événement dont on connaît la saveur initiale.

Alors, l'erreur statistique sur la mesure de $\sin(2\beta)$ est alors inversement proportionnelle à la dilution :

$$\sigma(\sin(2\beta)) \propto \frac{\sqrt{1 - (\mathcal{A}^m)^2}}{(1 - 2\omega)\sqrt{\varepsilon_{\text{tag}} N_{\text{tot}}}} \propto \frac{1}{\sqrt{\varepsilon_{\text{eff}} N_{\text{tot}}}}. \quad (3.3)$$

La quantité $\varepsilon_{\text{eff}} = \varepsilon_{\text{tag}}(1 - 2\omega)^2$ est appelé *puissance de l'étiquetage* ou *efficacité d'étiquetage effective*. C'est la quantité à maximaliser pour minimiser l'erreur statistique due à l'étiquetage sur $\sin(2\beta)$.

La section suivante décrit l'algorithme utilisé par LHCb pour l'étiquetage de la saveur initiale. Ensuite, la procédure d'optimisation choisie pour celui-ci sera présentée.

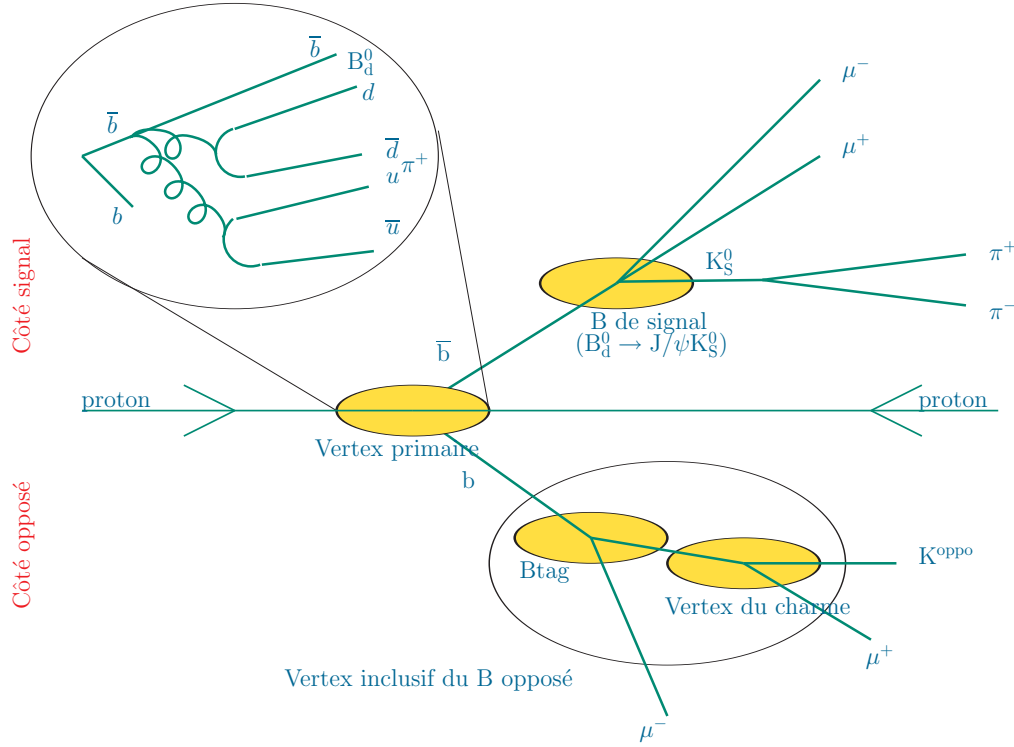


FIG. 3.1: Schéma représentant l'ensemble des sources d'information utilisées pour la détermination de la saveur initiale du méson beau de signal, ici un B_d^0 .

3.2 Algorithme et performances

L'algorithme choisi par LHCb utilise toute l'information disponible dans l'événement. En particulier il utilise la corrélation de saveur entre le quark $\bar{b}$ contenu dans le méson B_d^0 de signal et l'autre quark b utilisé pour former l'autre hadron beau, car les deux quarks sont produit par l'interaction forte, qui conserve la saveur. Il utilise aussi les corrélations de saveur entre le quark spectateur du méson B_d^0 de signal et ceux issus de la fragmentation.

Dans la suite, nous définissons le côté dit de signal, qui fait référence au méson beau étudié et aux particules issues de sa chaîne de fragmentation. Le côté opposé est lié au deuxième hadron beau. La Figure 3.1 montre l'ensemble des sources d'information utilisées par l'algorithme.

Des estimateurs spécifiques sont définis :

- pour le côté opposé (OS) : muons, électrons et kaons venant de la chaîne de désintégration $b \rightarrow c \rightarrow s$ du hadron beau opposé au signal ;
- une charge de vertex est construite à partir des traces associées au vertex du hadron beau opposé ;
- pour le côté signal (SS) : pions ou kaons venant de la chaîne de fragmentation

Coupure	valeur
p	$> 2 \text{ GeV}/c$ et $< 200 \text{ GeV}/c$
p_T	$< 10 \text{ GeV}/c$
θ	$> 12 \text{ mrad}$
IP/σ_{IP}	> 3
Type de trace	<i>long</i> ou <i>upstream</i>

TAB. 3.1: Coupure de présélection des particules utilisées pour l'étiquetage de la saveur des mésons B. Elles sont appliquées séquentiellement. θ est l'angle polaire entre la particule et le faisceau, et IP/σ_{IP} est la signification du paramètre d'impact entre la particule et les vertex primaires de pile-up. Cette coupure élimine les particules venant d'interactions secondaires.

du méson de signal. Les pions issus des désintégrations de mésons excités B^{**} sont aussi recherchés et utilisés.

Il existe un facteur limitant dans l'étude du côté opposé, du fait que le hadron beau opposé au signal peut être un méson neutre, B_d^0 ou B_s^0 . Il peut alors changer de saveur, dans $\chi_d = 18\%$ des cas pour un méson B_d^0 et dans $\chi_s = 50\%$ des cas pour un méson B_s^0 . La saveur donnée par l'algorithme d'étiquetage est donc fautive dans les mêmes proportions. C'est une fraction de mauvais étiquetage irréductible. Comme les quarks b s'hadronisent dans $f_{B_d^0} = 40\%$ des cas en B_d^0 et dans $f_{B_s^0} = 10\%$ des cas en B_s^0 , la fraction irréductible de mauvais étiquetage est égale à :

$$\bar{\chi} = f_{B_d^0} \chi_d + f_{B_s^0} \chi_s = 0.4 \times 0.18 + 0.1 \times 0.5 = 12.1\%$$

Bien entendu cette fraction de mauvais étiquetage irréductible ne concerne que le côté opposé. Il n'y a pas de telle restriction pour le côté signal puisque la charge portée par le kaon ou le pion est directement liée à la saveur du quark b produit.

Les traces utilisées par l'algorithme d'étiquetage sont pré-sélectionnées par un ensemble de coupures, résumées dans la Table 3.1, pour augmenter leur pouvoir discriminant. Seules les traces qui ne sont pas issues du méson beau de signal sont utilisées. En effet, dans la plupart des cas, les filles issues de la désintégration d'un hadron beau ne portent pas d'information sur sa saveur initiale. La coupure en impulsion minimale augmente la probabilité que les particules soient issues d'un hadron beau. Les coupures sur les valeurs maximales de l'impulsion et de l'impulsion transverse éliminent les particules dont l'identification est problématique. La coupure sur l'angle polaire des traces par rapport à l'axe du faisceau permet la réjection des traces mal reconstruites. La coupure en signification sur le paramètre d'impact évite d'utiliser des traces venant d'un second vertex primaire. La signifi-

cance est calculée par rapport à tous les vertex primaires différents de celui choisi comme origine du méson de signal.

L'optimisation des sélections des estimateurs individuels suit la même procédure pour le côté opposé et pour le côté signal. Elle se base sur les corrélations cinématiques entre les hadrons beaux et les estimateurs. Un paramètre est choisi puis l'optimisation se fait en variant la valeur de la coupure sur ce paramètre. En enregistrant le nombre de fois où l'information portée par les estimateurs passant la coupure sont en accord (N_C) et en désaccord (N_I) avec la saveur vraie du méson de signal, déterminée en utilisant la vérité Monte-Carlo ou la saveur reconstruite d'un méson beau chargé, l'efficacité d'étiquetage, la fraction de mauvais étiquetage, ainsi que l'efficacité effective d'étiquetage sont calculées. La valeur de coupure choisie est celle qui maximalise l'efficacité effective d'étiquetage.

Pour la charge de vertex inclusive, la procédure d'optimisation sera expliquée ultérieurement.

3.2.1 Étiquetage utilisant les produits de désintégration du hadron beau du côté opposé

Cette sous-section présente les sélections des estimateurs et la reconstruction du vertex associées au hadron beau côté opposé. Seuls les électrons, muons et kaons sont considérés parce que ce sont les particules qui sont le plus souvent produites dans les désintégrations d'un méson B. En effet, la probabilité qu'un méson beau donne dans l'état final un lepton est de l'ordre de 20% [13]. Parmi ces 20%, environ la moitié mettent en jeu un méson charmé. Celui-ci se désintègre en donnant un kaon chargé dans environ 30% des cas. Le diagramme de Feynman à l'ordre de l'arbre pour une telle chaîne de désintégration est montré sur la Figure 3.2. La charge du lepton issu du quark b porte une charge opposée à celui-ci, alors que celui associé au quark c porte la même charge que le quark b. L'utilisation d'un tel lepton dégrade les performances d'étiquetage.

L'ensemble des coupures utilisées pour chacun des estimateurs du côté opposé est résumé dans les Tables 3.2 pour les muons, 3.3 pour les électrons et 3.4 pour les kaons. Les coupures sont appliquées séquentiellement, dans l'ordre dans lequel elles sont données dans les tables.

La Figure 3.3 illustre l'optimisation de la coupure sur l'impulsion transverse du muon. Les performances pour chacun des estimateurs du côté opposé sont données dans la section 3.2.4.

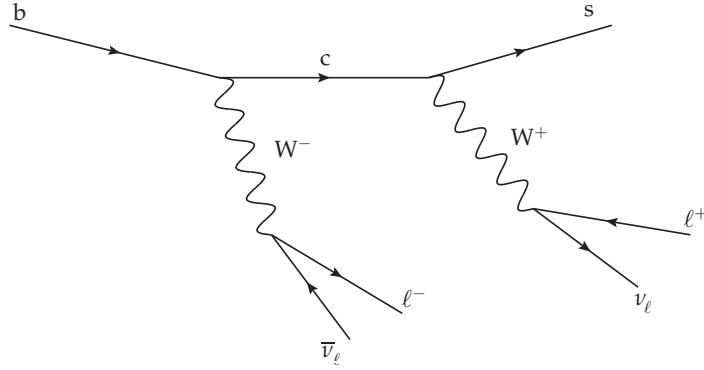


FIG. 3.2: Chaîne de désintégration semi-leptonique du quark b opposé au signal.

coupure	valeur
$\Delta \ln \mathcal{L}_{\mu\pi}$	> 0
p_T	$> 1,1 \text{ GeV}/c$
$\chi^2_{\text{trace}}/\text{nDoF}$	< 2
Partage de hits avec une autre trace	non

TAB. 3.2: Coupures appliquées dans la sélection des muons du côté opposé. La coupure sur $\Delta \ln \mathcal{L}_{\mu\pi}$ permet l'identification du muon. La coupure en impulsion transverse augmente la probabilité que le muon sélectionné vienne d'un hadron beau. La qualité de la trace $\chi^2_{\text{trace}}/\text{nDoF}$ permet d'éliminer les fausses traces (*ghost*). La dernière coupure, le partage de hits avec un autre trace, est un critère qui élimine les traces clonées.

coupure	valeur
$\Delta \ln \mathcal{L}_{e\pi}$	> 5
p_T	$> 1,1 \text{ GeV}/c$
p	$> 4 \text{ GeV}/c$
$\chi^2_{\text{trace}}/\text{nDoF}$	< 3
Vraisemblance de la trace	> -25
Charge d'ionisation	$< 1,7$
E/p	$> 0,9$

TAB. 3.3: Coupures appliquées pour la sélection de l'électron du côté opposé. La vraisemblance de la trace permet d'éliminer les fausses traces, la charge d'ionisation permet de rejeter les photo-productions de paires d'électrons-positrons dans le détecteur de vertex, et E/p permet un rejet des hadrons. En effet, la faible masse de l'électron implique que ce rapport soit proche de 1.

coupure	valeur
$\Delta \ln \mathcal{L}_{K\pi}$	> 8
$\Delta \ln \mathcal{L}_{K\pi} - \Delta \ln \mathcal{L}_{p\pi}$	> -1
p_T	$> 0,6 \text{ GeV}/c$
p	$> 4 \text{ GeV}/c$
IP/σ_{IP}	$> 3,5$
IP	$> 2 \text{ mm}$
Vraisemblance de la trace	> -25
traces <i>long</i> $\chi^2_{\text{trace}}/\text{nDoF}$	$< 2,5$
traces <i>upstream</i> $\chi^2_{\text{trace}}/\text{nDoF}$	< 2

TAB. 3.4: Coupures appliquées lors de la sélection du kaon opposé. Les coupures en signification de paramètre d'impact se font par rapport au vertex primaire. Cette coupure demande que le kaon soit loin du vertex. Pour les dernières coupures, comme deux types de traces sont utilisés, il faut prendre en compte leur différence de qualité. Pour les autres estimateurs, un seul type de trace est utilisé (*long*), d'où l'absence d'une telle différence.

Charge inclusive du vertex opposé

Un estimateur supplémentaire est la charge inclusive du vertex de désintégration du hadron beau opposé. La procédure de reconstruction du vertex et son optimisation étant très différentes de celles des traces individuelles, elles nécessitent un traitement particulier. De plus, comme l'optimisation de cette charge de vertex avec les données du *Data Challenge 2006* a fait partie des travaux de cette thèse, cette partie sera plus amplement détaillée.

L'objectif de l'algorithme de reconstruction du vertex est de construire le vertex de désintégration inclusif de la chaîne de désintégration du quark *b* opposé. Une fois ce vertex obtenu, il faut extraire la saveur du hadron beau opposé.

Les traces utilisées pour la reconstruction de ce vertex doivent respecter les critères suivants : avoir une impulsion supérieure à $2 \text{ GeV}/c$, et être de type *long* ou *upstream*. Selon le type de trace, une coupure différente en qualité de trace $\chi^2_{\text{trace}}/\text{nDoF}$ est appliquée. Elle doit être plus faible que 2,5 dans le cas *long* et inférieure à 5 dans le cas *upstream*.

La reconstruction du vertex commence par la sélection d'une paire de traces qui sera utilisée comme graine pour faire un vertex. Ces deux traces doivent avoir une signification de paramètre d'impact par rapport au vertex primaire compris entre 2 et 100. L'erreur sur ce paramètre d'impact ne doit pas dépasser 1 mm.

Une fois le vertex créé, des coupures sur sa qualité $\chi^2_{\text{vtx}}/\text{nDoF}$ ainsi que sur sa distance par rapport au vertex primaire suivant l'axe *z* sont appliquées. Cette distance doit être supérieure à 1 mm.

Dans le cas où la masse invariante des deux traces est compatible avec celle

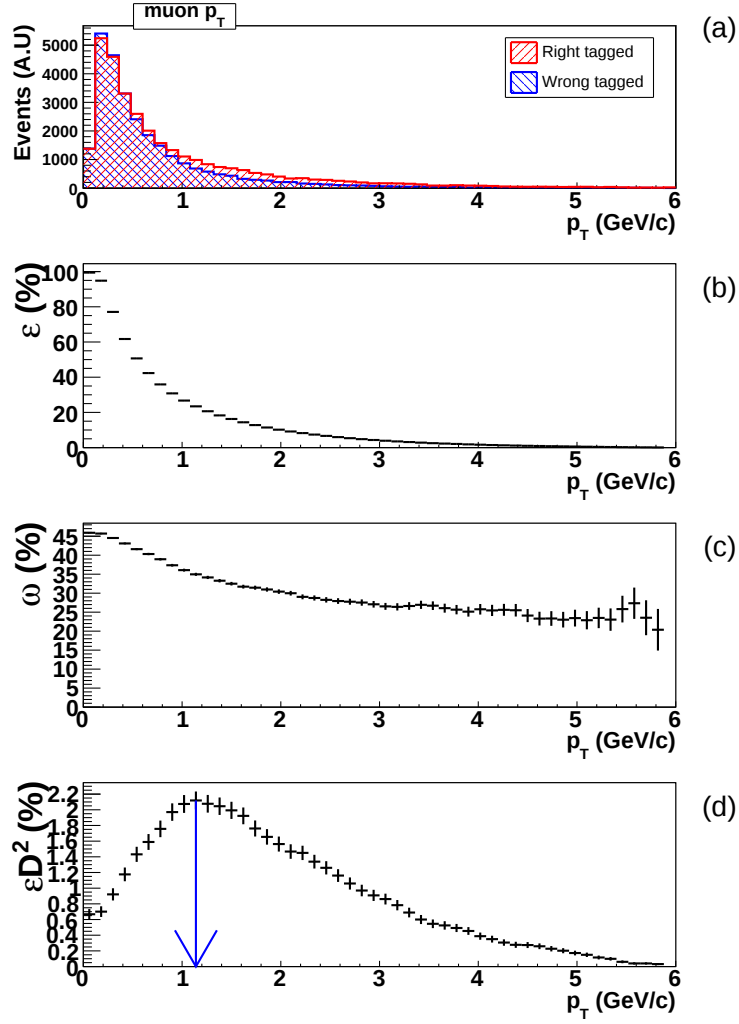


FIG. 3.3: Optimisation de la coupure sur l'impulsion transverse des candidats muons dans un lot d'événements $B^+ \rightarrow J/\psi K^+$. Ces distributions sont obtenues après coupure sur $\Delta \ln \mathcal{L}_{\mu\pi}$. (a) distribution de l'impulsion transverse des candidats muons. En rouge sont donnés les événements correctement étiquetés, en bleu les événement incorrectement étiquetés. (b) efficacité d'étiquetage en fonction de la coupure ; (c) fraction de mauvais étiquetage en fonction de la coupure ; (d) efficacité effective de l'étiquetage. Une coupure à 1.1 GeV/c correspond au maximum de l'efficacité effective.

d'un K_S^0 ou celle d'un Λ^0 ¹, alors ces deux traces sont exclues des traces candidates à la création du vertex. En effet, les traces issues de ces deux particules ne portent pas d'information sur la charge du hadron beau, et de plus le temps de vol de ces particules étant grand, le vertex inclusif serait artificiellement déplacé loin de sa position nominale.

Quand une paire de traces passe cet ensemble de coupures, les propriétés de celle-ci sont utilisées pour calculer une vraisemblance qui teste l'hypothèse que cette paire de trace soit issue de la désintégration d'un quark b. Les paramètres utilisés sont les paramètres d'impact, les impulsions transverses, et l'angle entre les deux traces. Les densités de probabilités de ces trois quantités observables ont été obtenues en recherchant dans la vérité Monte-Carlo l'origine vraie des traces. La vraisemblance du vertex choisi doit être supérieure à 0,2.

Si la vraisemblance est plus grande que 0,32, alors d'autres traces peuvent être ajoutées au vertex existant. Celles-ci doivent avoir une signification du paramètre d'impact par rapport au vertex primaire comprise entre 1,8 et 100, et ne pas être issue d'un K_S^0 ou d'un Λ^0 . La trace est finalement sélectionnée en utilisant la même fonction de probabilité que pour le calcul de vraisemblance. Le vertex résultant doit avoir une qualité $\chi_{\text{vtx}}^2/\text{nDoF}$ plus faible que 5 et être à plus de 1 mm du vertex primaire suivant l'axe z. La plus mauvaise trace utilisée pour construire le vertex peut être retirée si la signification de son paramètre d'impact par rapport au vertex reconstruit est plus grande que 3. Bien entendu, il doit toujours rester au moins deux traces issues du vertex.

Une fois le vertex inclusif de désintégration du hadron beau opposé reconstruit, une charge inclusive pondérée de vertex est calculée :

$$Q_{\text{vtx}} = \frac{\sum_{i=1}^n Q_i p_T^\kappa(i)}{\sum_{i=1}^n p_T^\kappa(i)},$$

où Q_i est la charge de la trace i et $p_T(i)$ son impulsion transverse. La charge Q_{vtx} est représentée dans la Figure 3.4. Le facteur de pondération $\kappa = 0,35$ est utilisé pour prendre en compte la corrélation entre l'origine d'une trace et son impulsion transverse. Lorsque cette charge est peu discriminante, c'est-à-dire lorsqu'elle est comprise entre $-0,15$ et $0,15$, elle est fixée à 0, et l'algorithme ne va pas la considérer dans la combinaison des décisions. La valeur de κ ainsi que la largeur de la coupure centrale sont optimisées pour maximaliser l'efficacité effective d'étiquetage.

Les performances pour le côté opposé seront données dans la section 3.2.4.

¹C'est-à-dire que la masse invariante est comprise entre $490 \text{ MeV}/c^2$ et $505 \text{ MeV}/c^2$ ou entre $1110 \text{ MeV}/c^2$ et $1120 \text{ MeV}/c^2$

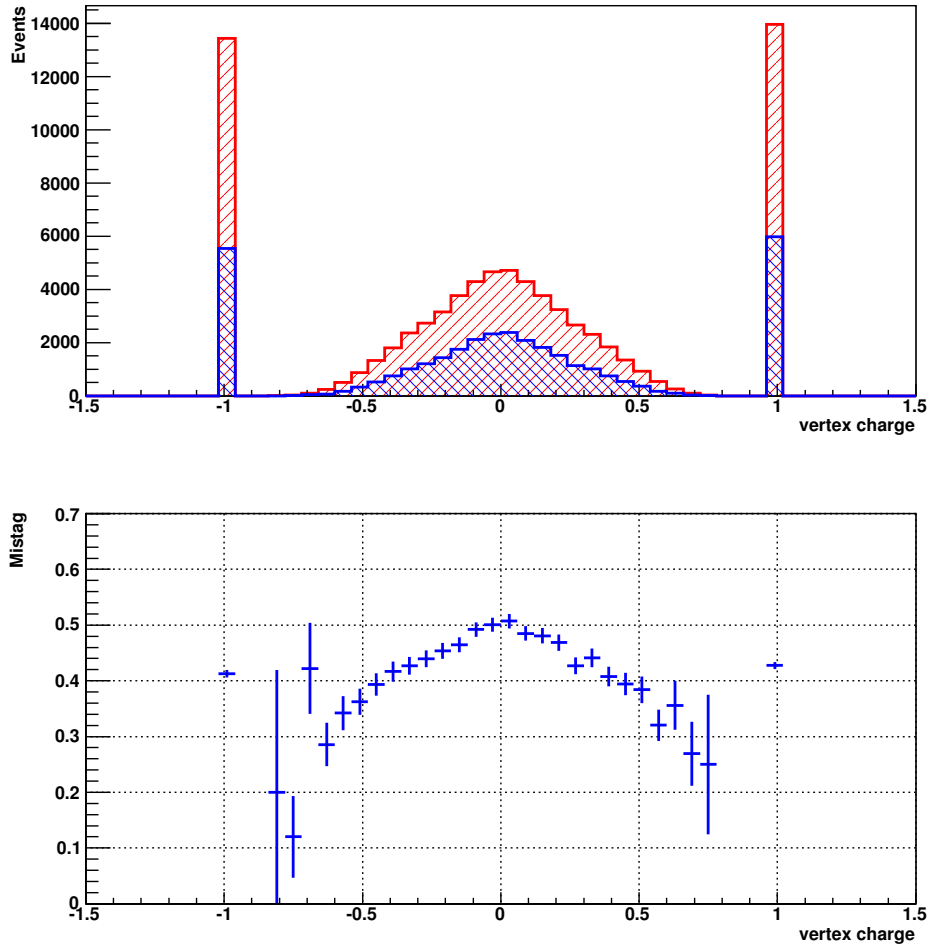


FIG. 3.4: Distribution du haut : charge de vertex inclusive dans un lot d'événements $B^+ \rightarrow J/\psi K^+$. En rouge sont donnés les événements dont la charge de vertex est en accord avec la saveur vraie du méson B de signal, et en bleu ceux où elle est en désaccord. Les pics à ± 1 apparaissent dans le cas où toutes les traces utilisées portent la même charge. La région comprise entre -0.15 et 0.15 est retirée, c'est-à-dire que la charge est mise à 0 du fait du trop faible pouvoir séparateur. La distribution du bas donne la fraction de mauvais étiquetage en fonction de la charge de vertex.

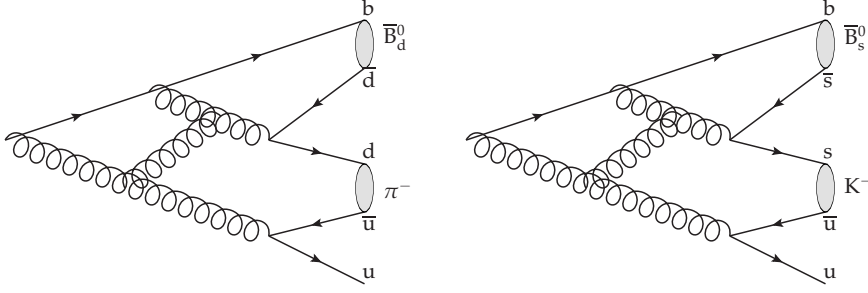


FIG. 3.5: Diagrammes de Feynman de l'hadronisation en méson $\bar{B}_d^0$ (gauche) et $\bar{B}_s^0$ (droite). La création d'un méson π ou K n'est pas systématique.

3.2.2 Étiquetage utilisant les traces issues de la fragmentation du méson beau de signal

Lors de l'hadronisation des quarks beaux, un certain nombre de particules sont créées, comme le montrent les diagrammes de Feynman à l'ordre de l'arbre montrés sur la Figure 3.5. Dans le cas où le quark b s'hadronise en B_d^0 ou en B^+ , un méson π^+ sera très probablement créé. Une contribution de pion provient des désintégrations de B^{**} . Dans le cas d'un B_s^0 , c'est un méson étrange qui sera certainement créé. Celui-ci, en se désintégrant donnera souvent un kaon chargé. Ces particules sont donc des estimateurs de la saveur initiale du méson beau de signal. Cette section s'intéresse aux critères utilisés pour sélectionner ces estimateurs.

Dans le cas où le méson étudié est un B_d^0 ou un B^+ , le pion recherché doit respecter les critères de la Table 3.5. Dans le cas où le méson de signal est un B_s^0 , le kaon utilisé doit passer les coupures données dans la Table 3.6.

Avant de donner les performances de ces estimateurs, il faut s'intéresser au cas où plusieurs estimateurs sont disponibles pour un événement. Dans ce cas, une procédure de combinaison des réponses est mise en place, détaillée dans la section suivante.

3.2.3 Combinaison des décisions

Lorsque plusieurs sources d'information sont disponibles, il faut combiner les réponses pour obtenir la meilleure décision possible. Pour cela, il faut déterminer une fraction de mauvais étiquetage par estimateur, qui est utilisée pour déterminer la fraction de mauvais étiquetage événement par événement, essentielle pour l'extraction de paramètres de violation de CP.

Coupure	valeur
p_T	$> 200 \text{ MeV}/c$
p	$> 2 \text{ GeV}/c$
Signature dans le calo. électromag.	non
$\Delta \ln \mathcal{L}_{e\pi}$	> 5
$\Delta \ln \mathcal{L}_{K\pi}$	< 8
$\Delta \ln \mathcal{L}_{Kp}$	< -1
IP/σ_{IP} (Vertex Primaire)	< 3
$M_{B\pi} - M_B$	$< 1,5 \text{ GeV}/c^2$
$\chi^2_{\text{trace}}/\text{nDoF}$	< 3
Vraisemblance de la trace	> -25

TAB. 3.5: Coupures utilisées pour la sélection du pion du côté signal. Il n'est recherché que si le méson de signal est un B^+ ou B_d^0 . La coupure $M_{B\pi} - M_B$ est la différence de masse invariante entre celle du système $B - \pi$ et celle du méson B de signal. Les autres coupures sont expliquées dans les tables précédentes.

Coupure	valeur
$\Delta \ln \mathcal{L}_{K\pi}$	> 8
$\Delta \ln \mathcal{L}_{Kp}$	> -4
p_T	$> 400 \text{ MeV}/c$
p	$> 4 \text{ GeV}/c$
IP/σ_{IP} (Vertex Primaire)	< 3
$\Delta\eta = \eta_K - \eta_B$	$< 1 \text{ rad}$
$\Delta\phi = \phi_K - \phi_B$	$< 1,1 \text{ rad}$
$M_{BK} - M_B$	$< 1,4 \text{ GeV}/c^2$
$\chi^2_{\text{trace}}/\text{nDoF}$	< 4

TAB. 3.6: Coupures utilisées pour la sélection du kaon du côté signal. Celui-ci n'est recherché que dans le cas où le méson de signal est un B_s^0 .

La technique utilisée par LHCb tire profit des propriétés cinématiques des particules et des propriétés générales de l'événement. Elle repose sur l'utilisation d'un réseau de neurone (Neural Net). Celui-ci permet de quantifier la qualité de l'information apportée par un estimateur donné, et donc de donner une fraction de mauvais étiquetage par estimateur et par événement. Les ingrédients fournis aux réseaux de neurones sont les suivants :

- pour les estimateurs du côté opposé : muons, électrons, et kaon ; le nombre de traces disponibles pour l'algorithme d'étiquetage, l'impulsion transverse du hadron beau de signal, l'impulsion et l'impulsion transverse de l'estimateur, et enfin la signification de son paramètre d'impact par rapport au vertex primaire.
- pour les estimateurs du côté signal, kaon ou pion, les mêmes informations

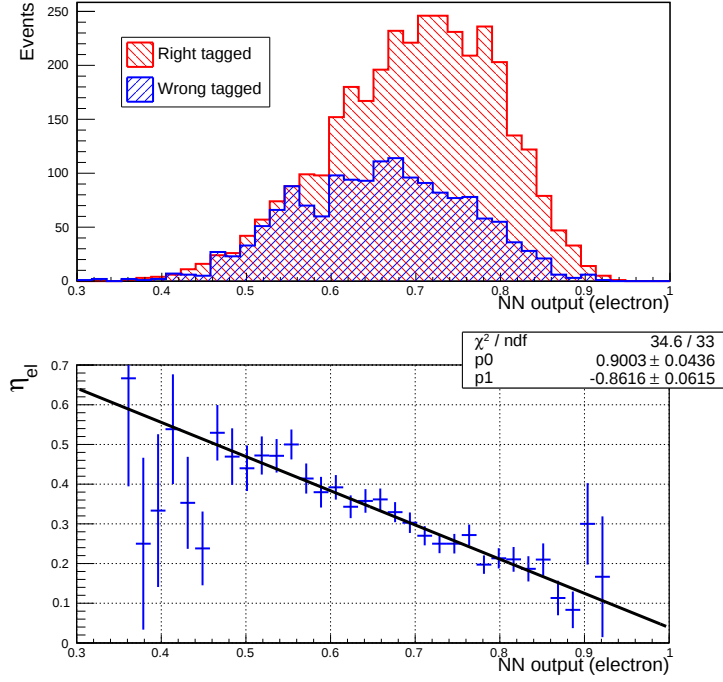


FIG. 3.6: Distribution de la sortie du réseau de neurone pour l'estimateur électron dans un lot d'événements $B^+ \rightarrow J/\psi K^+$. Distribution du haut : distribution pour les événements correctement étiquetés (rouge) et incorrectement étiquetés (bleu). Distribution du bas : fraction de mauvais étiquetage η_{el} en fonction de la sortie du réseau de neurone.

que pour les estimateurs du côté opposé, mais complétées par les distances polaires par rapport au méson beau de signal $\Delta\theta$ et $\Delta\phi$. De plus, la différence de masse invariante entre le système B – K et B est utilisée.

Le critère pour l'entraînement des réseaux est la saveur vraie du méson beau de signal qui est connue dans les données Monte-Carlo utilisées.

La sortie des réseaux de neurones est alors histogrammée pour les événements correctement étiquetés et pour les événements incorrectement étiquetés. Dès lors, il est possible d'extraire la fraction de mauvaise étiquetage η_i en fonction de la sortie des réseaux de neurones, pour chaque estimateur. En pratique cette fonction $\eta_i(\text{NNet})$ est, avec une bonne approximation, un polynôme du premier degré comme le montre la Figure 3.6.

En ce qui concerne la charge inclusive du vertex opposé, il est possible d'extraire directement la fraction de mauvais étiquetage $\eta_{Q_{\text{vtx}}}$ en fonction de la valeur de la charge, montrée dans la Figure 3.7. Lorsque toutes les traces utilisées dans la reconstruction ont la même charge, alors la valeur de $\eta_{Q_{\text{vtx}}}$ est fixée à une valeur moyenne. Cette moyenne dépend du nombre de traces utilisées : s'il n'y en a que

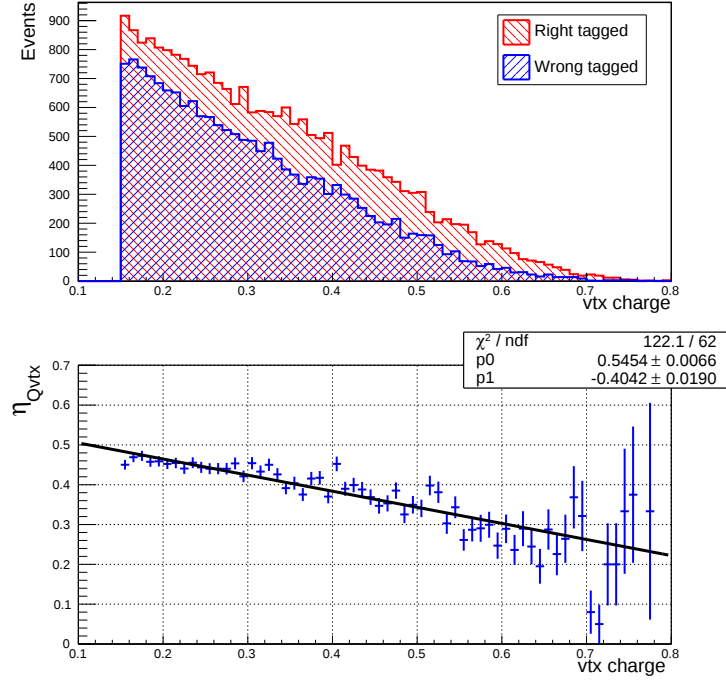


FIG. 3.7: Distribution de la fraction de mauvais étiquetage η_{Qvtx} en fonction de la charge inclusive de vertex opposé dans un lot d'événements $B^+ \rightarrow J/\psi K^+$. Les cas où toutes les traces utilisées ont la même charge sont traités en particulier, et n'apparaissent pas dans cette distribution.

2 alors $\eta_{Qvtx} = 0.42$, et s'il y en a plus que 2 alors $\eta_{Qvtx} = 0.35$. Ces valeurs sont extraites de l'optimisation.

Les distributions des η_i pour les estimateurs muon, électron, kaon du côté opposé, pion du côté signal, et charge inclusive de vertex sont données dans la Figure 3.8. Les pics dans la distribution de η_{Qvtx} sont le résultat du fait que toutes les traces utilisées dans la reconstruction du vertex ont la même charge.

L'apprentissage des réseaux de neurones se fera en utilisant les données issues de la simulation, car il est nécessaire de connaître la saveur vraie des mésons beaux. Alors, puisque les données réelles et les données Monte-Carlo ne peuvent être en accord parfait, il est nécessaire d'étalonner les η_i en utilisant les données réelles pour corriger ces différences :

$$\omega_i(\eta_i) = p0_i + p1_i(\eta_i - \bar{\eta}_i)$$

où $\bar{\eta}_i$ est la moyenne des η_i . Cette transformation permet de limiter les corrélations linéaires entre les couples $p0_i$ et $p1_i$ [44]. L'extraction des $p0_i$ et $p1_i$ sera expliquée

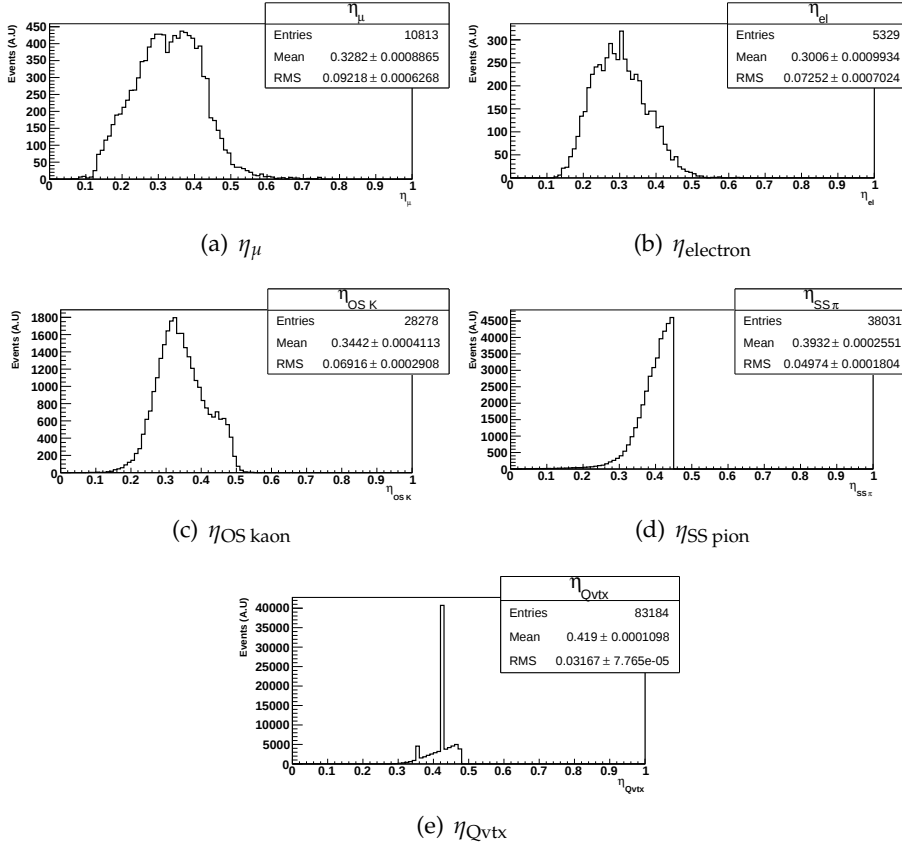


FIG. 3.8: Distribution des η pour chacun des estimateurs dans un lot d'événements $B^+ \rightarrow J/\psi K^+$. Ces distributions seront utilisées dans la section 3.2.5.

dans la sous-section 3.2.5.

Une fois les $\omega_i(\eta_i)$ obtenus, il est possible de combiner les différentes décisions. On définit $p_i = 1 - \omega_i$ comme la probabilité que l'estimateur i donne la bonne décision, notée $q_i = \pm 1$. Alors la probabilité normalisée que le méson beau de signal contiennent un quark b est :

$$\mathcal{P}(b) = \frac{p(b)}{p(b) + p(\bar{b})},$$

dans laquelle

$$p(b) = \prod_i \left(\frac{1 - q_i}{2} + q_i p_i \right), \quad \text{et} \quad p(\bar{b}) = \prod_i \left(\frac{1 + q_i}{2} - q_i p_i \right),$$

où $p(b)$ est la probabilité non normalisée que l'ensemble des estimateurs i soient compatibles avec la présence d'un quark b , et $p(\bar{b})$ celle qu'ils soient compatibles avec un quark $\bar{b}$. La probabilité normalisée qu'il contienne un quark $\bar{b}$ est :

$$\mathcal{P}(\bar{b}) = 1 - \mathcal{P}(b).$$

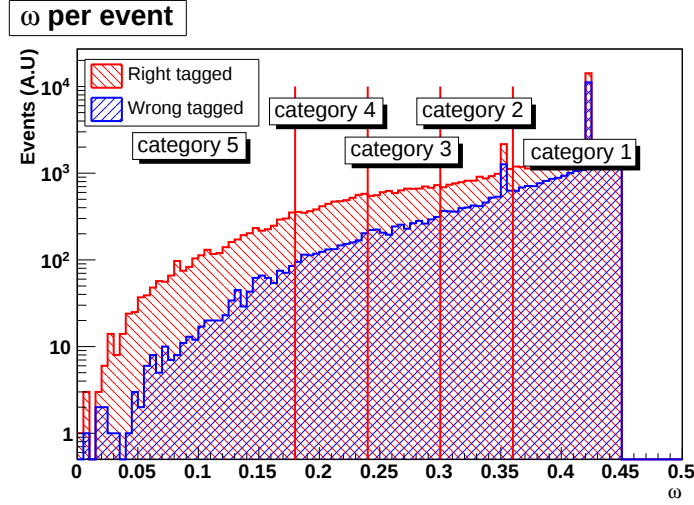


FIG. 3.9: Distribution de la fraction de mauvais étiquetage par événement après combinaison des décisions individuelles de chaque estimateurs dans un lot d'événements $B^+ \rightarrow J/\psi K^+$. Cette distribution est similaire entre différents lots. Les deux pics à 0,35 et 0,42 sont dûs aux valeurs discrètes des probabilités de la charge inclusive de vertex opposé dans le cas où toutes ses traces portent la même charge. Les différentes catégories d'étiquetage sont indiquées.

Si $\mathcal{P}(b)$ est supérieur à $\mathcal{P}(\bar{b})$, alors la décision est que le quark b du méson de signal est un b . Sinon c'est un $\bar{b}$. Si la probabilité normalisée $\mathcal{P}$ est inférieure à 0.55, alors l'algorithme fixe la décision à 0, et la probabilité $\mathcal{P}$ à 0,5. Le méson beau de signal n'est alors pas étiqueté. Dans le cas contraire, la fraction de mauvais étiquetage pour cet événement est $\omega = 1 - \mathcal{P}$. La distribution des ω par événement est donnée dans la Figure 3.9.

Dans cette combinaison, les estimateurs sont considérés comme indépendants. Or la charge inclusive de vertex est en corrélation avec les autres estimateurs du côté opposé. Cette corrélation peut diminuer le pouvoir discriminant de l'algorithme d'étiquetage, mais ne devrait pas biaiser la détermination de la fraction de mauvais étiquetage utilisée [43] dans la mesure où celle-ci sera mesurée directement dans les données avec des canaux de contrôle, voir par exemple le chapitre 5.

Si les événements sont utilisés tels quels pour évaluer des asymétries CP sans prendre en compte leur répartition dans la distribution Figure 3.9, alors de bons événements à petit ω sont mélangés à de mauvais événements de grand ω . L'efficacité effective moyenne est plus faible que si les événements sont triés par valeurs de fraction de mauvais étiquetage similaires. La séparation en 5 catégories de fractions de mauvais étiquetage permet d'obtenir une amélioration notable des performances de l'étiquetage [43]. Les différentes limites de catégories sont représentées

Catégorie	$\varepsilon_{\text{eff}}(\%)$	$\varepsilon_{\text{tag}}(\%)$	$\omega(\%)$
1	0.70 ± 0.15	27.98 ± 0.40	42.1 ± 0.8
2	1.10 ± 0.18	10.54 ± 0.28	33.9 ± 1.3
3	1.64 ± 0.21	8.09 ± 0.24	27.5 ± 1.4
4	1.45 ± 0.20	6.09 ± 0.21	25.6 ± 1.6
5	3.53 ± 0.27	7.12 ± 0.23	14.8 ± 1.2

(a) Performances par catégories d'étiquetage

Catégorie	$\varepsilon_{\text{eff}}(\%)$	$\varepsilon_{\text{tag}}(\%)$	$\omega(\%)$
OS muons	1.33 ± 0.20	9.64 ± 0.26	31.4 ± 1.3
OS elect	0.41 ± 0.11	3.06 ± 0.15	31.8 ± 2.4
OS kaons	1.81 ± 0.23	16.44 ± 0.33	33.4 ± 1.0
SS kaon	3.38 ± 0.31	29.89 ± 0.41	33.2 ± 0.8
Q_{vtx}	1.39 ± 0.21	45.88 ± 0.45	41.3 ± 0.7

(b) Performances par estimateurs

Efficacité d'étiquetage = $59.82 \pm 0.44\%$ Fraction de mauvais étiquetage = $31.25 \pm 0.55\%$ Efficacité effective combinée = $8.41 \pm 0.46\%$
--

(c) Performances de la combinaison

TAB. 3.7: Performances de l'algorithme d'étiquetage sur un échantillon de 12 000 événements $B_s^0 \rightarrow D_s \pi$ [45].

dans la Figure 3.9. Lors de l'extraction d'un paramètre de violation de CP tel que $\sin(2\beta)$, chaque catégorie est utilisée dans un ajustement simultané, chacune ayant sa propre fraction de mauvais étiquetage.

3.2.4 Performances

Les performances de l'algorithme d'étiquetage sont données dans les Tables 3.7 pour les désintégrations $B_s^0 \rightarrow D_s \pi$ [45], en utilisant la vérité Monte-Carlo pour déterminer la saveur vraie du méson beau de signal.

La Table 3.7(a) présente les performances des différentes catégories, telles que définie dans la section précédente. Elles ont des efficacités d'étiquetage et des fractions de mauvais étiquetage décroissantes. Cependant, l'efficacité effective est croissante, de sorte que la catégorie 5 possède la meilleure efficacité effective et la catégorie 1 la plus mauvaise.

La Table 3.7(b) donne les performances de l'étiquetage par estimateur. Le muon du côté opposé est bien meilleur que l'électron. Cela provient de la supériorité de l'efficacité de reconstruction et d'identification des muons par rapport à celle des électrons, ce qui a été brièvement expliqué dans le chapitre 2.

La charge inclusive de vertex possède la même puissance d'étiquetage que les

Catégorie	$\epsilon_{\text{eff}}(\%)$	$\epsilon_{\text{tag}}(\%)$	$\omega(\%)$
SS pion	0.66 ± 0.08	18.09 ± 0.20	40.5 ± 0.6

TAB. 3.8: Performances de l'étiquetage de l'estimateur pion côté signal. Elles sont extraites d'un échantillon de 38 000 événements $B_d^0 \rightarrow J/\psi K_S^0$

muons, grâce à sa très grande efficacité d'étiquetage. En effet, sa fraction de mauvais étiquetage reste très élevée par rapport aux autres estimateurs. Cette différence s'explique par la faible pureté du vertex reconstruit. Cette dernière est définie comme le rapport entre le nombre de traces venant réellement (dans la vérité Monte-Carlo) du vertex de désintégration du hadron beau opposé et le nombre total de traces utilisées pour sa reconstruction. Il vaut en moyenne 28,4%.

Le kaon du côté opposé reste le meilleur estimateur du côté opposé. Cette performance se justifie par les critères très stricts de sa sélection. En effet, les coupures appliquées permettent une pureté après sélection finale de l'ordre de 80%.

Dans l'exemple ci-dessus, l'estimateur du côté signal est un kaon puisque le méson de signal est un B_S^0 . Dans le cas où c'est un méson B^+ ou B_d^0 , l'estimateur est un pion. Dans ce cas, ses performances sont données dans la Table 3.8. Elles sont moins bonnes que celles du kaon du côté signal dans la mesure où la multiplicité des pions est très supérieure à celle des kaons, et donc la probabilité d'utiliser un mauvais pion est plus grande.

Enfin, la Table 3.7(c) résume les performances globales de l'algorithme d'étiquetage qui prennent en compte la catégorisation par valeur de fraction de mauvais étiquetage. L'efficacité d'étiquetage combinée est la somme des efficacité d'étiquetage de chaque catégorie. Il en va de même pour l'efficacité effective combinée. De ces deux informations, on extrait une fraction de mauvais étiquetage combinée :

$$\omega^{comb} = \frac{\left(1 - \sqrt{\frac{\epsilon_{\text{eff}}^{comb}}{\epsilon_{\text{tag}}^{comb}}}\right)}{2}.$$

Cette valeur de mauvais étiquetage permet de quantifier la qualité de l'algorithme avec un seul nombre. De plus, elle peut être utilisée dans les simulations simplifiées, au lieu de générer explicitement les différentes catégories. Deux exemples d'utilisation seront donnés dans les chapitres 5 et 6.

Les performances varient d'un canal de désintégration à un autre. Celles-ci dépendent des sélections utilisées et du niveau de déclenchement appliqué. Le canal $B_S^0 \rightarrow D_S \pi$ est celui qui présente les meilleures performances, et le canal $B_d^0 \rightarrow J/\psi K_S^0$ les plus mauvaises, données dans la Table 4.22.

Expérience	Puissance de l'étiquetage (%)	Référence
BaBar	30.5	[46]
Belle	28.8	[47]
D0	4.5	[48]
CDF	6.6	[49]
ATLAS	4.6	[50]
LHCb	8.4	[44]

TAB. 3.9: Performances d'étiquetage comparées entre différentes expériences. Pour BaBar et Belle, il s'agit de l'étiquetage de méson B_d^0 , alors que pour les autres expériences, il s'agit de mésons B_s^0 . Même si ce sont deux types différents de mésons, les différences entre les deux types d'expériences sont telles qu'il y a un réel gain.

Pour avoir un aperçu des performances de l'algorithme développé par LHCb, il est utile de les comparer avec celles des autres expériences faisant des études de violation de CP. En particulier, les expériences installées auprès des usines à B, BaBar (SLAC, USA) et Belle (KEKB, Japon), et les expériences hadroniques du TeVatron (D0 et CDF) et du LHC (ATLAS).

La Table 3.9 compare les performances de l'étiquetage entre différentes expériences et montre la supériorité des collisionneurs e^+e^- dans ce domaine. Celle-ci s'explique par le fait que les machines leptoniques ont un bruit de fond bien plus faible que les machines hadroniques. De même le nombre de traces reconstruites est plus faible. De plus, dans ces expériences, les collisions ont lieu au seuil de production de $Y(4S)$, et les mésons B issus de sa désintégration sont corrélés. Ils oscillent donc en phase. Il n'y a alors pas la restriction présentée au début de ce chapitre sur la fraction de mauvais étiquetage du côté opposé. Par rapport aux autres expériences hadroniques, le détecteur LHCb se montre plus performant en raison de son système d'identification des particules.

Jusqu'ici, l'algorithme a été optimisé en utilisant la vérité Monte-Carlo sur des échantillons de signaux purs. Lorsque l'expérience LHCb commencera à acquérir des données, il faudra étalonner l'algorithme avec des événements réels. Il y aura du bruit de fond et la saveur vraie ne sera connue que grâce aux canaux de contrôles auto-étiquetés tels que $B^+ \rightarrow J/\psi K^+$. La section suivante détaille la procédure envisagée pour l'extraction des paramètres des polynômes qui décrivent la fraction de mauvais étiquetage ω_i en fonction des η_i . La mesure de la fraction de mauvais étiquetage est l'objet du chapitre 5.

3.2.5 Étalonnage

L'étalonnage présenté ici est détaillé dans la référence [44]. Il ne considère que la partie concernant l'extraction des $\omega_i(\eta_i)$, et ne donne pas le paramétrage des η_i en fonction des sorties des réseaux de neurones. En effet, celle-ci est extraite pour le moment de la vérité Monte-Carlo, tout comme l'optimisation de l'algorithme d'étiquetage. La mesure de la fraction de mauvais étiquetage est traitée dans le chapitre 5.

Cet étalonnage se fait sur un canal de contrôle auto-étiqueté, $B^+ \rightarrow J/\psi K^+$, en utilisant un ajustement combiné des propriétés du méson B^+ . Sa masse invariante $\mu\mu K^+$ ainsi que son temps de vie sont utilisés comme variables discriminantes du bruit de fond. Pour déterminer les fractions de mauvais étiquetage, la décision de l'algorithme d'étiquetage est comparée à la charge reconstruite du méson beau. Les distributions des η_i de chaque estimateur sont nécessaires, et sont modélisées par différentes fonctions densité de probabilités selon l'estimateur considéré.

Le modèle utilisé pour extraire la fraction de mauvais étiquetage de chaque estimateur est basé sur la construction de fonctions de densité de probabilités pour le signal et les deux bruits de fond :

- le bruit de fond appelé *prompt* : toutes les traces utilisées dans la reconstruction du méson B^+ sont issues en réalité du vertex primaire. Avec la sélection utilisée, décrite dans chapitre 4, c'est le bruit de fond dominant. Comme aucune trace n'est issue d'un méson B^+ , la charge reconstruite n'est pas corrélée avec celle retournée par l'algorithme d'étiquetage, et la fraction de mauvais étiquetage correspond à celle d'un étiquetage aléatoire : 0,5.
- le bruit de fond à grand temps de vie : certaines traces utilisées pour la reconstruction viennent du vertex primaire. Malgré le fait que les propriétés du méson de signal soient fausses puisqu'une partie de l'information utilisée pour sa reconstruction n'est pas corrélée au B^+ réel, sa saveur est corrélée avec celle retournée par l'algorithme quand le K^+ est issu d'un méson B^+ . Dans ce cas, la fraction de mauvais étiquetage ω_i extraite est supposée différente de 0,5.

Pour le signal, la densité de probabilité décrivant la fraction de mauvais étiquetage ω_i en fonction des η_i est :

$$\mathcal{S}_i^{\text{tag}} = \begin{cases} \varepsilon_i \times \{1 - [p0_i + p1_i \times (\eta_i - \bar{\eta}_i)]\} \times \mathcal{N}(\eta_i) & \text{si correctement étiqueté,} \\ \varepsilon_i \times \{p0_i + p1_i \times (\eta_i - \bar{\eta}_i)\} \times \mathcal{N}(\eta_i) & \text{si incorrectement étiqueté,} \\ \{1 - \varepsilon_i\} & \text{si non étiqueté,} \end{cases}$$

où ε_i est l'efficacité d'étiquetage de l'estimateur i , et $\mathcal{N}(\eta_i)$ la densité de probabilité des η_i , donnée dans la Figure 3.8. Pour la masse et le temps de vie, les densités de

estimateur	$p0_i$		$p1_i$	
	Résultat	Vérité MC	Résultat	Vérité MC
muon	0.312 ± 0.005	0.308 ± 0.005	0.86 ± 0.06	0.90 ± 0.05
électron	0.315 ± 0.008	0.314 ± 0.008	1.13 ± 0.10	1.13 ± 0.10
OS kaon	0.331 ± 0.003	0.331 ± 0.003	0.74 ± 0.05	0.74 ± 0.05
SS pion	0.378 ± 0.003	0.377 ± 0.003	0.88 ± 0.04	0.93 ± 0.06
Q_{vtx}	0.425 ± 0.002	0.425 ± 0.002	0.94 ± 0.06	0.96 ± 0.06

TAB. 3.10: Résumé de l'ajustement sur des données issues de la simulation complète pour l'extraction des paramètres $p0_i$ et $p1_i$. Aucun des paramètres ne montre de déviation par rapport à sa vérité Monte Carlo.

probabilité correspondantes sont plus classiques, c'est-à-dire une somme de deux gaussiennes pour la masse invariante et une exponentielle décroissante convoluée par une fonction de résolution pour le temps de vie. Les détails des composantes pour le signal et les bruits de fond sont donnés dans la référence [44].

Un ajustement sur des données issues de la simulation complète [44] donne les résultats de la Table 3.10, dans laquelle les mesures sont en accord avec les valeurs extraites de la vérité Monte Carlo. L'échantillon utilisé est un mélange d'événements de type $B_{u,d,s} \rightarrow J/\psi X$, avec des proportions naturelles entre les différents types de mésons beaux. Cette table montre que la procédure mise en place fonctionne. Pour tester la solidité du modèle, un ensemble de simulations rapides a été effectué. Le résultat est qu'aucun des paramètres mis en jeu n'a de biais.

De nombreux canaux de contrôles sont possibles parmi lesquels $B^+ \rightarrow \bar{D}^0 \pi^+$, $B^+ \rightarrow \bar{D}^0 \mu^+ \nu_\mu$ [51] ou encore $B_s^0 \rightarrow D_s \mu^+ \nu_\mu$ [52, 53], qui font partie de ce travail. Chacun de ces canaux de contrôle permettent d'étalonner de la même manière l'algorithme d'étiquetage pour les différentes mesures, telles que $\Phi_{J/\psi\Phi}$, γ , etc.

3.3 Optimisation de l'algorithme d'étiquetage

Les données Monte-Carlo utilisées au début du travail dataient de 2004. Un certain nombre d'améliorations dans la description du détecteur ainsi qu'une meilleure connaissance de la physique à haute énergie ont demandé une nouvelle génération de simulations Monte-Carlo. Ces nouvelles données produites en 2006 simulaient un détecteur contenant plus de matière et surtout plus bruyant. Mais elles n'en n'étaient pas moins plus réalistes. Malheureusement, ces améliorations ont eu pour conséquence une diminution des performances globales de l'étiquetage de l'ordre de 30%. Les Tables 3.11 et 3.12 mesurent respectivement les performances de l'étiquetage de 2004 et de 2006. Les performances sont moins bonnes avec les données de 2006, avec une perte d'environ 30% relatifs. Ces diminutions sont indépendantes

Estimateur	$\varepsilon_{\text{tag}}(\%)$	$\omega(\%)$	$\varepsilon_{\text{eff}}(\%)$
muon	5.6 ± 0.1	31.6 ± 0.7	0.76 ± 0.06
électron	3.3 ± 0.1	31.0 ± 0.9	0.48 ± 0.05
OS kaon	28.5 ± 0.2	38.0 ± 0.3	1.65 ± 0.09
SS kaon	26.2 ± 0.2	35.2 ± 0.3	2.31 ± 0.11
Q_{vtx}	21.2 ± 0.2	39.1 ± 0.4	1.00 ± 0.07
Combinaison	56.8 ± 0.2	33.2 ± 0.2	6.40 ± 0.24

TAB. 3.11: Performances de l'étiquetage d'un échantillon $B_s^0 \rightarrow D_s \pi$ dans les données Monte-Carlo de 2004. La combinaison est faite selon les sorties de réseaux de neurones. Cette table doit être comparée à la Table 3.12.

Estimateur	$\varepsilon_{\text{tag}}(\%)$	$\omega(\%)$	$\varepsilon_{\text{eff}}(\%)$
muon	5.6 ± 0.2	31.4 ± 1.5	0.77 ± 0.13
électron	4.1 ± 0.2	35.5 ± 1.9	0.35 ± 0.09
OS kaon	19.0 ± 0.3	38.1 ± 0.9	1.07 ± 0.16
SS kaon	20.0 ± 0.3	35.0 ± 0.8	1.80 ± 0.20
Q_{vtx}	17.7 ± 0.3	41.4 ± 0.9	0.53 ± 0.11
Combinaison	46.9 ± 0.4	34.3 ± 0.6	4.64 ± 0.44

TAB. 3.12: Performances de l'étiquetage d'un échantillon $B_s^0 \rightarrow D_s \pi$ à l'arrivée des données de 2006. La combinaison est faite selon les sorties de réseaux de neurones. Cette table doit être comparée à la Table 3.11.

de l'échantillon analysé. Un travail de collaboration a été mis en place pour comprendre l'origine des pertes et réoptimiser les estimateurs. Nous nous sommes concentrés sur l'électron.

3.3.1 Optimisation de l'estimateur électron

L'optimisation de cet estimateur a suivi la même procédure que pour les autres estimateurs. La valeur de chaque coupure a été scannée pour maximiser la puissance de l'étiquetage. Dans cette étude, pour simplifier l'analyse, nous n'avons pas inclus de nouveaux paramètres. Les données utilisées sont différentes de celles qui ont servi aux tables de référence 3.11 et 3.12. Il s'agit de données de signal $B_d^0 \rightarrow J/\psi K^*$.

Cette analyse a montré que le rapport E/p n'était pas correctement évalué dans les nouvelles données. En effet, l'étalonnage du calorimètre électromagnétique n'était pas terminé lors de cette analyse, et certaines erreurs restaient à corriger.

De plus, nous avons aussi observé que la charge d'ionisation déposée dans les détecteurs en silicium du VELO n'était pas correctement simulée par Geant 4. De ce fait, la coupure sur la valeur minimale de cette charge a été enlevée.

Coupure	avant opti.	après opti.
p_T	$> 1,2 \text{ GeV}/c$	$> 1,2 \text{ GeV}/c$
p	$> 5 \text{ GeV}/c$	$> 5 \text{ GeV}/c$
E/p	$> 0,8$	$> 0,92$
$\chi^2_{\text{trace}}/\text{nDoF}$	< 4	< 4
Velo Charge	$> 1,4 \text{ et } < 1,8$	$< 1,8$

TAB. 3.13: Coupures utilisées dans la sélection de l'estimateur électron lors de l'étude d'optimisation. Seules les coupures sur le rapport E/p et sur le dépôt d'énergie dans le détecteur VELO ont changé dans cette optimisation.

	$\varepsilon_{\text{tag}}(\%)$	$\omega(\%)$	$\varepsilon_{\text{eff}}(\%)$
Avant analyse	4.2 ± 0.1	36.2 ± 1.1	0.32 ± 0.05
Après analyse	3.72 ± 0.04	34.4 ± 0.6	0.36 ± 0.03

TAB. 3.14: Performances de l'estimateur électron avant et après l'optimisation sur deux lots d'événements $B_d^0 \rightarrow J/\psi K^*$ de taille similaire. Le gain est de l'ordre de 12%.

La Table 3.13 donne la comparaison des valeurs des coupures avant et après optimisation.

La comparaison entre les performances avant et après optimisation est donnée dans la Table 3.14. Le gain total de cette optimisation sur l'estimateur électron est de 12% relatifs.

Depuis ce travail, de nombreux points ont été améliorés. En particulier, l'identification des particules a été complètement revue courant 2008. De même la trajectographie a été optimisée. Enfin, la sélection utilisée était basée sur la vérité Monte-Carlo, et aucune simulation du système de déclenchement n'était disponible au moment de l'analyse. Les performances actuelles peuvent être trouvées dans la Table 4.21.

3.4 Conclusion

L'étiquetage de la saveur initiale des mésons beaux est un aspect fondamental des mesures de violation de CP.

Dans l'expérience LHCb, l'algorithme développé utilise l'ensemble des informations disponibles dans le détecteur. En particulier, il utilise le fait que les quarks b sont produits en paires, et qu'un des quarks s'hadronise en méson beau dit de signal, alors que l'autre s'hadronise en hadron beau appelé opposé. L'algorithme recherche les filles du hadron beau opposé, en particulier les leptons issus de sa désintégration semi-leptonique, et un kaon issu de la chaîne de désintégration $b \rightarrow c \rightarrow s$. De plus, l'algorithme tente de reconstruire le vertex inclusif de désinté-

gration du hadron beau opposé et mesure une charge de vertex inclusif. Enfin, il recherche un pion ou un kaon issu de la fragmentation du méson beau de signal.

La décision globale de l'algorithme se base sur la sortie de réseaux de neurones pour chacun des estimateurs. Une classification des événements est faite suivant la qualité du résultat, dépendante aussi de la sortie de ces réseaux de neurones.

Les performances globales de l'étiquetage donnent une efficacité effective maximale de l'ordre de 9.5%[45]. Une comparaison avec les autres expériences hadroniques montre que l'expérience LHCb a les meilleures performances. Cela s'explique essentiellement par le fait que le détecteur possède un système d'identification des particules.

Ce chapitre a aussi présenté une méthode d'étalonnage de l'algorithme. Celui-ci s'effectue en utilisant un canal de contrôle. Ce canal est auto-étiqueté, c'est-à-dire que la saveur initiale du méson beau est connue à partir de ses produits de désintégration. Typiquement, les canaux chargés tels que $B^+ \rightarrow J/\psi K^+$ ou $B^+ \rightarrow D^0 \pi^+$ sont de bons canaux de contrôle [51]. Ceux-ci servent à l'optimisation de l'algorithme et son étalonnage. D'autres canaux, tels que $B_d^0 \rightarrow J/\psi K^{*0}$ ou $B_s^0 \rightarrow D_s^- \pi^+$ sont utilisés pour mesurer les fractions de mauvais étiquetage qui seront utilisées dans les mesures de violation de CP.

Une des contraintes dans l'utilisation des canaux de contrôles est la possible apparition des erreurs systématiques. Celle-ci sont engendrées par les différences entre les fractions de mauvais étiquetage mesurées dans un canal et appliquées dans un autre. Pour limiter l'impact de ces erreurs, et simplifier l'analyse, une des solutions est de développer des sélections pour les canaux de contrôle et les canaux de signal qui n'induisent pas de différences. C'est l'objectif du chapitre suivant.