
Extraction de Δm_d et de la fraction de mauvais étiquetage avec le canal $B_d^0 \rightarrow J/\psi K^{*0}$

LA MESURE de $\sin(2\beta)$ présentée dans cette thèse nécessite la connaissance de la fraction de mauvais étiquetage ω ainsi que de la fréquence du mélange $B_d^0 - \bar{B}_d^0$, Δm_d . La méthode suivie pour extraire ces deux paramètres est décrite dans ce chapitre. Elle utilise le canal $B_d^0 \rightarrow J/\psi K^{*0}$ qui présente l'avantage d'être spécifique de saveur, la saveur du méson B_d^0 à l'instant de sa désintégration est déterminée à partir de la charge des hadrons observés dans l'état final.

Afin de pouvoir transposer directement la fraction de mauvais étiquetage déterminée ici à l'étude du canal $B_s^0 \rightarrow J/\psi \phi$ pour la mesure de $\Phi_{J/\psi \phi}$, l'analyse est limitée à l'étiquetage du côté opposé au signal.

Dans un premier temps, la modélisation du signal et du bruit de fond est discutée. Ensuite les paramètres du modèle sont ajustés sur des données de la simulation détaillée du détecteur, afin d'extraire les valeurs utilisées pour déterminer la sensibilité à la fraction de mauvais étiquetage et Δm_d pour une année nominale de données. Ces valeurs de sensibilités sont estimées à l'aide de simulations simplifiées.

5.1 Modélisation

La fraction de mauvais étiquetage se mesure en étudiant le mélange $B_d^0 - \bar{B}_d^0$ en fonction du temps propre dans le canal $B_d^0 \rightarrow J/\psi K^{*0}$. Les quantités observables sont donc t , le temps propre du méson B_d^0 , et q la différence entre la saveur du méson

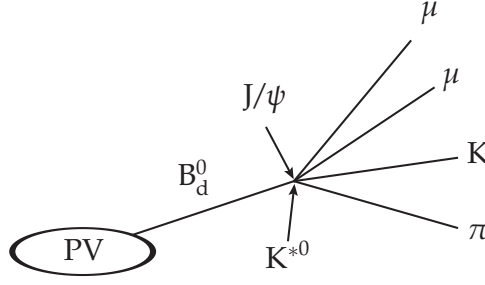


FIG. 5.1: Topologie d'une désintégration d'un méson B_d^0 en $J/\psi K^{*0}$: les quatre traces sont issues d'un vertex secondaire.

B_d^0 à sa création et sa saveur lors de sa désintégration. La convention suivante est adoptée : $q = +1$ si les deux saveurs sont identiques, et $q = -1$ dans le cas contraire. De plus, la masse invariante $\mu\mu K\pi$ est utilisée pour discriminer signal et bruits de fond.

5.1.1 Signal

La topologie d'une désintégration $B_d^0 \rightarrow J/\psi K^{*0}$ est illustrée par la Figure 5.1. Du point de vue temps propre, la désintégration est décrite par les relations 1.21 en prenant en compte l'approximation $\Delta\Gamma = 0$ pour les mésons B_d^0 . La probabilité qu'un méson B_d^0 ou $\bar{B}_d^0$ avec une certaine saveur initiale se désintègre après un temps t avec une saveur identique est

$$\text{PDF}_{\text{non mélangé}}(t; \tau_{B_d^0}, \Delta m_d) = \frac{1}{2\tau_{B_d^0}} e^{-\frac{t}{\tau_{B_d^0}}} [1 + \cos(\Delta m_d t)].$$

De même, la probabilité qu'un B_d^0 ou $\bar{B}_d^0$ se désintègre après un temps t avec une saveur opposée est donnée par

$$\text{PDF}_{\text{mélangé}}(t; \tau_{B_d^0}, \Delta m_d) = \frac{1}{2\tau_{B_d^0}} e^{-\frac{t}{\tau_{B_d^0}}} [1 - \cos(\Delta m_d t)].$$

Nous pouvons alors écrire une probabilité dépendante du temps t et de q :

$$\text{PDF}_S(t, q; \tau_{B_d^0}, \Delta m_d) = \frac{1}{2\tau_{B_d^0}} e^{-\frac{t}{\tau_{B_d^0}}} [1 + q \cos(\Delta m_d t)].$$

En prenant en compte la fraction de mauvais étiquetage ω_S , cette densité de probabilité devient

$$\text{PDF}_S(t, q; \tau_{B_d^0}, \omega_S, \Delta m_d) = \frac{1}{2\tau_{B_d^0}} e^{-\frac{t}{\tau_{B_d^0}}} [1 + q(1 - 2\omega_S) \cos(\Delta m_d t)].$$

De manière à prendre en compte la résolution temporelle, il est nécessaire de convoluer cette densité de probabilité par R , la somme de deux gaussiennes de largeur σ_{t1} et σ_{t2} , centrées sur la même moyenne μ_t pour simplifier :

$$R(t; \mu_t, \sigma_{t1}, \sigma_{t2}, f_{\text{res G1}}) = f_{\text{res G1}} e^{-\frac{1}{2} \left(\frac{t - \mu_t}{\sigma_{t1}} \right)^2} + (1 - f_{\text{res G1}}) e^{-\frac{1}{2} \left(\frac{t - \mu_t}{\sigma_{t2}} \right)^2}, \quad (5.1)$$

où $f_{\text{res G1}}$ est le poids attribué à la première gaussienne. Comme la fréquence des oscillations Δm_d est faible ($\Delta m_d = 0.502 \text{ ps}^{-1}$) par rapport à la résolution moyenne ($\sigma_t \sim 40 \text{ fs}$), l'utilisation d'une résolution dont la largeur est indépendante du temps est adaptée pour l'ajustement de l'amplitude de l'asymétrie de mélange.

La fonction de densité de probabilité précédente devient :

$$\begin{aligned} \text{PDF}_S(t, q; \tau_{B_d^0}, \omega_S, \Delta m_d, R) &= \left(\frac{1}{2\tau_{B_d^0}} e^{-\frac{t}{\tau_{B_d^0}}} [1 + q(1 - 2\omega_S) \cos(\Delta m_d t)] \right) \\ &\otimes R(t; \mu_t, \sigma_{t1}, \sigma_{t2}, f_{\text{res G1}}). \end{aligned}$$

Par ailleurs, la distribution en masse invariante du signal est décrite par une simple gaussienne pour prendre en compte la résolution du détecteur :

$$\text{PDF}_S(m; \mu_{G1}, \sigma_{G1}) = e^{-\frac{1}{2} \left(\frac{m - \mu_{G1}}{\sigma_{G1}} \right)^2}.$$

La fonction de densité de probabilité totale du signal est le produit des deux densités précédentes :

$$S = \text{PDF}_S(t, q; \tau_{B_d^0}, \omega_S, \Delta m_d, R) \times \text{PDF}_S(m; \mu_{G1}, \sigma_{G1}). \quad (5.2)$$

5.1.2 Bruits de fond

L'étude de l'origine des particules reconstruites, dans des données Monte-Carlo, fait apparaître trois grandes classes de bruit de fond avec des propriétés cinématiques très différentes. La modélisation de ces trois bruits de fond est présentée ci-dessous.

Prompt

La classe dite *prompt* décrit les événements où les quatre traces utilisées pour former la désintégration $B_d^0 \rightarrow J/\psi(\mu\mu)K^{*0}(K\pi)$ proviennent en réalité du vertex primaire. Un tel événement est schématisé dans la Figure 5.2. Le temps propre moyen de ces événements est donc en principe nul. La résolution temporelle a les mêmes caractéristiques que pour le signal. Par ailleurs, la saveur déterminée par la charge des hadrons dans l'état final est complètement décorrélée du résultat de

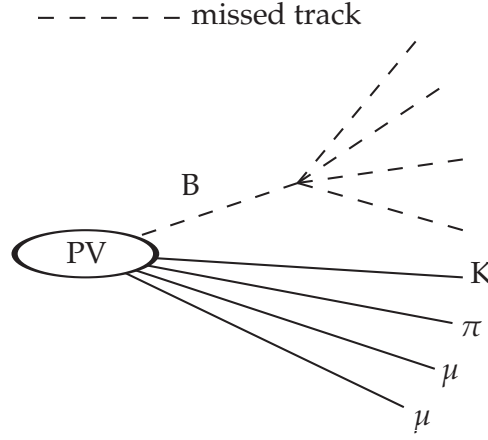


FIG. 5.2: Topologie d'un évènement de type *prompt* : les quatre traces utilisées pour la reconstruction du méson B_d^0 sont issues du vertex primaire. C'est le bruit de fond dominant avec la sélection utilisée.

l'algorithme d'étiquetage. Aucun terme n'impliquant l'étiquetage n'apparaît donc dans le modèle de ce bruit de fond.

La distribution du temps propre est décrite par la somme de deux gaussiennes dont les paramètres sont identiques à ceux utilisés pour la résolution temporelle du modèle du signal :

$$\text{PDF}_{\text{prompt}}(t; \mu_t, \sigma_{t1}, \sigma_{t2}, f_{\text{res G1}}) = R(t; \mu_t, \sigma_{t1}, \sigma_{t2}, f_{\text{res G1}}).$$

Comme la distribution de masse invariante de ces événements ne doit pas présenter de pics, elle est modélisée par une fonction lisse. Lorsque la statistique est élevée, une distribution exponentielle correspond généralement aux données :

$$\text{PDF}_{\text{prompt}}(m, \text{slope}_p) = e^{m \times \text{slope}_p}.$$

La fonction de densité de probabilité totale pour le bruit de fond *prompt* est alors :

$$\begin{aligned} \text{PDF}_{\text{prompt}}(t, m; \mu_t, \sigma_{t1}, \sigma_{t2}, f_{\text{res G1}}, \text{slope}_p) &= \text{PDF}_{\text{prompt}}(t; \mu_t, \sigma_{t1}, \sigma_{t2}, f_{\text{res G1}}) \\ &\times \text{PDF}_{\text{prompt}}(m, \text{slope}_p). \end{aligned}$$

Deux traces erronées

Si deux traces sur quatre ne sont pas issues du méson de signal (Figure 5.3), la longueur de vol n'est plus nulle, et la distribution du temps propre reconstruit devient approximativement exponentielle.

Lors de l'étiquetage de ces événements, trois cas se présentent :

1. les deux traces utilisées pour former le K^{*0} sont issues du vertex primaire : le méson de signal reconstruit a alors une saveur aléatoire ;

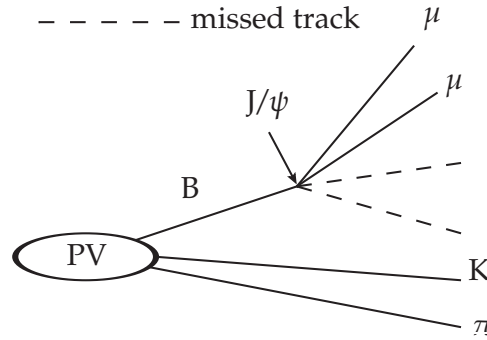


FIG. 5.3: Topologie d'une désintégration du type B2 : les deux traces utilisées pour la reconstruction du K^{*0} ne proviennent pas du méson B_d^0 , mais du vertex primaire.

2. ce sont les traces du J/ψ qui sont issues du vertex primaire : dans ce cas, la saveur reconstruite est correcte. Dans l'étude présentée ici, ce cas de figure n'est pas représenté parce qu'il n'y a pas de tel événements dans les lots de bruit de fond disponibles pour l'analyser.
3. enfin, une trace utilisée pour former le K^{*0} et une trace pour former le J/ψ sont issues du vertex primaire. Ce cas n'est pas considéré ici car très minoritaire.

Comme ce bruit de fond est dominé par le premier cas, il n'y a pas de terme oscillant, et la fraction de mauvais étiquetage attendue est proche de 0,5, soit un étiquetage aléatoire. La fonction de densité de probabilité choisie pour le temps propre est alors :

$$\text{PDF}_{B2}(t, q; \tau_{B2}, \omega_{B2}, R) = \left(\frac{1}{2\tau_{B2}} e^{-\frac{t}{\tau_{B2}}} [1 + q(1 - 2\omega_{B2})] \right) \otimes R(t; \mu_t, \sigma_{t1}, \sigma_{t2}, f_{\text{res G1}}),$$

où τ_{B2} et ω_{B2} sont respectivement le temps propre et la fraction de mauvais étiquetage de ce bruit de fond. La distribution temporelle est convoluée avec la même fonction de résolution que le signal. La résolution n'est théoriquement pas identique à celle du signal car deux traces ne viennent pas du même point de l'espace, et donc la résolution devrait être plus mauvaise. Cependant, parce que Δm_d est faible, et pour simplifier le modèle, nous considérons la même fonction de résolution que pour le signal. Dans les données réelles, lorsque la statistique sera suffisante, il faudra prendre en compte cette différence.

Pour la dépendance en masse, comme deux traces sur quatre sont décorréliées du méson de signal, la masse invariante reconstruite de ce dernier ne doit pas montrer de pic. Elle est par conséquent modélisée par une exponentielle décroissante :

$$\text{PDF}_{B2}(m; \text{slope}_{B2}) = e^{m \times \text{slope}_{B2}}.$$

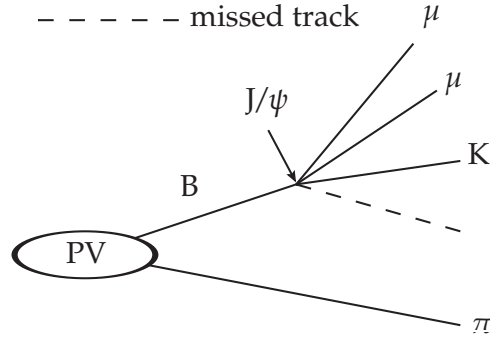


FIG. 5.4: Topologie d'une désintégration du type B3 : une des deux traces utilisées pour la reconstruction du K^{*0} ne provient pas du méson B, mais du vertex primaire.

Finalement, la fonction de densité de probabilité totale de ce bruit de fond est :

$$\text{PDF}_{B2}(t, q, m; \tau_{B2}, \omega_{B2}, R, \text{slope}_{B2}) = \text{PDF}_{B2}(t, q; \tau_{B2}, \omega_{B2}, R) \times \text{PDF}_{B2}(m; \text{slope}_{B2})$$

Une trace erronée

Le troisième bruit considéré est le cas où une seule trace ne provient pas du méson de signal, illustré dans la Figure 5.4. La position du vertex secondaire de désintégration est très alors proche de celle de la vraie position, puisque trois traces sur quatre portent la bonne information. Pour les propriétés d'étiquetage, deux cas de figure se présentent :

1. Soit le méson B reconstruit est bien un B_d^0 , et ce bruit de fond a la même fraction de mauvais étiquetage que le signal
2. Soit le méson B reconstruit n'est pas un B_d^0 , mais par exemple un B^+ se désintégrant en $J/\psi K^{*+}$. Le rapport d'embranchement de cette désintégration est similaire à celui du signal $B_d^0 \rightarrow J/\psi K^{*0}$. Dans ce cas, la saveur reconstruite de l'état final est toujours la même que celle à la production, puisque le B^+ n'oscille pas. La fraction de mauvais étiquetage de ces événements n'a pas de raison d'être identique à celle du signal.

Pour prendre en compte ces deux contributions en une seule pour simplifier le modèle, nous considérons une composante oscillante, avec la même fréquence de mélange que le signal Δm_d , mais avec une fraction de mauvais étiquetage plus élevée pour rendre compte du fait que nous donnons des propriétés oscillantes à une partie des événements alors que celle-ci n'en a pas.

La densité de probabilité correspondante pour le temps propre reconstruit est donc :

$$\text{PDF}_{B3}(t, q; \tau_{B3}, \omega_{B3}, \Delta m_d, R) = \left(\frac{1}{2\tau_{B3}} e^{-\frac{t}{\tau_{B3}}} [1 + q(1 - 2\omega_{B3}) \cos(\Delta m_d t)] \right) \otimes R(t; \mu_t, \sigma_{t1}, \sigma_{t2}, f_{\text{res } G1}),$$

où τ_{B3} est le temps propre de ces événements et ω_{B3} leur fraction de mauvais étiquetage. La résolution temporelle R est considérée identique à celle du signal, pour les mêmes raisons que pour le bruit de fond précédent.

En ce qui concerne la masse invariante, comme une trace porte une information complètement décorrélée de celle du méson de signal, la distribution de masse doit être une gaussienne de grande largeur comparée au signal. Pour simplifier l'analyse et parce que ce bruit de fond est supposé faible devant les autres, nous choisissons de le modéliser par une densité de probabilité exponentielle :

$$\text{PDF}_{B3}(m; \text{slope}_{B3}) = e^{m \times \text{slope}_{B3}}.$$

La densité de probabilité totale pour ce bruit de fond est le produit des deux précédentes :

$$\text{PDF}_{B3}(t, q, m; \tau_{B3}, \omega_{B3}, \Delta m_d, R, \text{slope}_{B3}) = \text{PDF}_{B3}(t, q; \tau_{B3}, \omega_{B3}, \Delta m_d, R) \times \text{PDF}_{B3}(m; \text{slope}_{B3}).$$

Autres types de bruits de fond

Le cas où trois traces sur quatre sont issues du vertex primaire n'est pas traité ici, parce qu'il est considéré comme très marginal. De même, les cas où la désintégration est $B_d^0 \rightarrow J/\psi K^{*0} X$, comme par exemple $B_d^0 \rightarrow J/\psi K^{*0} \pi^+ \pi^-$, sont ignorés ici parce que leur rapports d'embranchement sont très faibles par rapport à celui du signal.

5.1.3 Combinaison

Le modèle complet est la somme des contributions de signal et celles des bruits de fond :

$$\text{PDF} = f_S \text{PDF}_S + (1 - f_S) \text{PDF}_B, \quad (5.3)$$

où f_S est la fraction de signal $B_d^0 \rightarrow J/\psi K^{*0}$, PDF_S la densité de probabilité du signal (eq. 5.2), et PDF_B celle du bruit de fond. Elle est définie comme :

$$\text{PDF}_B = f_{\text{prompt}} \text{PDF}_{\text{prompt}} + (1 - f_{\text{prompt}}) \text{PDF}_{\text{LL}},$$

où f_{prompt} est la fraction du bruit de fond *prompt* par rapport à l'ensemble du bruit. La densité de probabilité PDF_{LL} est la somme des bruits de fond B2 et B3, LL signifiant *Long Living* :

$$\text{PDF}_{\text{LL}} = f_{\text{B2}}\text{PDF}_{\text{B2}} + (1 - f_{\text{B2}})\text{PDF}_{\text{B3}},$$

où f_{B2} est la fraction du bruit de fond B2 par rapport au bruit de fond à grand temps propre.

5.2 Ajustement sur les données de la simulation complète

Un ajustement de ce modèle sur les données issues de la simulation complète est effectué pour le valider et déterminer les valeurs des paramètres à utiliser dans les simulations simplifiées. Dans un premier temps, nous déterminerons les paramètres du signal et des bruits de fond à grand temps propre B2 et B3 avec un ajustement sur un lot d'événements $B_{u,d} \rightarrow J/\psi X$, puis nous déterminerons les propriétés des événements du bruit de fond *prompt* en étudiant un lot d'événements J/ψ inclusif.

5.2.1 Le lot $B_{u,d} \rightarrow J/\psi X$

Le lot utilisé se compose de 2,41 millions d'événements $B^+ \rightarrow J/\psi X$ et 2,35 millions d'événements $B_d^0 \rightarrow J/\psi X$, après coupure au niveau générateur et *stripping*. Cela assure qu'il y ait autant de B^+ que de B_d^0 avant les coupures, comme les proportions naturelles l'imposent. D'après la Table 4.1, la statistique générée correspond à une luminosité intégrée de $\sim 0,2 \text{ fb}^{-1}$. Après la sélection finale, le niveau 0 de déclenchement, et l'étiquetage par le côté opposé, il reste 65 751 événements, parmi lesquels 45 948 événements sont des B_d^0 et 19 803 événements sont des B^+ .

L'ajustement des données se fait en utilisant la technique de maximum de vraisemblance, et la fonction à maximiser est :

$$\mathcal{L}(m, t, q; \theta) = \prod_{i=0}^N \text{PDF}_i(m, t, q; \theta) \quad (5.4)$$

où i désigne un événement, PDF est défini par la relation 5.3, et θ est l'ensemble des paramètres rappelés dans la table 5.1.

Dans cet ajustement, comme le bruit de fond *prompt* est absent, f_{prompt} est fixé à 0. Seuls les paramètres du signal et ceux des bruits de fond à grand temps propre sont extraits. Trois ajustements utilisant des configurations d'étiquetage différentes sont effectués. Nous discuterons plus loin les mérites comparés de chacun.

Le premier utilise tous les événements sans prendre en compte la catégorisation donnée par l'algorithme d'étiquetage, c'est-à-dire dans un seul lot. La fraction

Paramètre	Valeur MC	Résultat	corrélacion globale
f_S	–	$0,3966 \pm 0,0026$	0,361
$\mu_{G1} [\text{GeV}/c^2]$	5,2778	$5,2775 \pm 1,34 \times 10^{-4}$	0,087
$\sigma_{G1} [\text{GeV}/c^2]$	0,01611	$0,01664 \pm 1,2 \times 10^{-4}$	0,345
$\tau_{B_d^0} [\text{ps}]$	1,536	$1,528 \pm 0,011$	0,227
$\omega_{\text{sig}}^{\text{OS}}$	0,3896	$0,3936 \pm 0,0048$	0,283
$\Delta m_d [\text{ps}^{-1}]$	0,502	$0,475 \pm 0,020$	0,234
$\mu_t [\text{ps}]$	–	$(-5,6323 \pm 7,43) \times 10^{-4}$	0,727
$f_{\text{res } G1}$	–	$0,98584 \pm 0,00153$	0,399
$\sigma_{t1} [\text{fs}]$	–	$34,177 \pm 0,667$	0,660
$\sigma_{t2} [\text{ps}]$	–	$0,44438 \pm 0,0882$	0,456
f_{B2}	–	$0,729 \pm 0,008$	0,869
$\text{slope}_{B2} [(\text{GeV}/c^2)^{-1}]$	–	$-0,669 \pm 0,066$	0,531
$\tau_{B2} [\text{ps}]$	–	$0,177 \pm 0,003$	0,835
ω_{B2}	0,5	$0,498 \pm 0,004$	0,495
$\text{slope}_{B3} [(\text{GeV}/c^2)^{-1}]$	–	$-5,222 \pm 0,15$	0,546
$\tau_{B3} [\text{ps}]$	–	$0,854 \pm 0,018$	0,762
ω_{B3}	–	$0,40 \pm 0,01$	0,522

TAB. 5.1: Résultat de l'ajustement de 65751 événements de données $B_{u,d} \rightarrow J/\psi X$ utilisant la technique du maximum de vraisemblance. La première colonne définit les paramètres. La seconde leur valeurs dans la vérité Monte-Carlo, lorsqu'elle est connue. La troisième donne le résultat de l'ajustement. L'erreur est statistique uniquement. La quatrième donne la corrélation globale de chaque paramètre par rapport aux autres.

de mauvais étiquetage moyenne est obtenue de cette manière. Les résultats de l'ajustement sont donnés dans les Tables 5.1 et 5.2(a). La distribution de la masse invariante $\mu\mu K\pi$ est montrée dans la Figure 5.5, et la distribution du temps propre est donnée dans la Figure 5.6. La largeur moyenne de la résolution temporelle est $\langle\sigma\rangle = \sqrt{f_{\text{res } G1}\sigma_{t1}^2 + (1 - f_{\text{res } G1})\sigma_{t2}^2} \approx 62,8 \text{ ps}$. Pour effectuer une vérification supplémentaire, nous construisons l'asymétrie de mélange, définie par l'équation 1.22, à partir de la densité de probabilité totale 5.3 ajustée :

$$\mathcal{A}^{\text{mix}}(t) = \frac{\text{PDF}(t, q = 1) - \text{PDF}(t, q = -1)}{\text{PDF}(t, q = 1) + \text{PDF}(t, q = -1)}.$$

Celle-ci est superposée à l'asymétrie construite à partir des données dans la Figure 5.7. Le χ^2/nDoF entre la courbe et les données est proche de 1, signe que l'ajustement est cohérent. Le modèle détaillé précédemment correspond donc bien aux données.

Ensuite un ajustement simultané des événements étiquetés par chacun des cinq estimateurs de l'étiquetage est effectué. Les lots ne sont pas indépendants car deux

(a) Ajustement en utilisant tous les événements dans une catégorie

	Valeur MC	Résultat	déviations (σ)
Moyenne	0,3896	$0,3936 \pm 0,005$	0,85

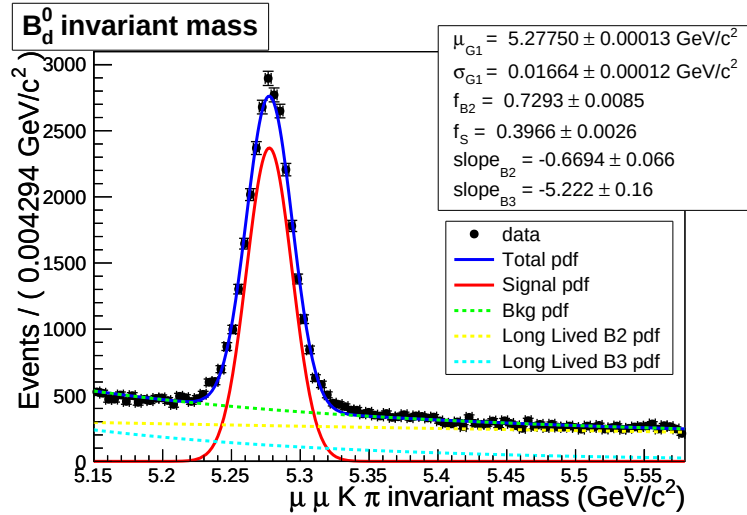
(b) Ajustement par estimateurs

	Valeur MC	Résultat	déviations (σ)
muon	0,3194	$0,324 \pm 0,013$	0,32
électron	0,3201	$0,330 \pm 0,018$	0,57
OS Kaon	0,3272	$0,3391 \pm 0,0082$	1,45
Q_{vtx}	0,4221	$0,4232 \pm 0,0049$	0,22
SS pion	0,3944	$0,4002 \pm 0,0065$	0,89

(c) Ajustement par catégorie d'étiquetage

	Valeur MC	Résultat	déviations (σ)
NNet 1	0,447	$0,448 \pm 0,006$	0,20
NNet 2	0,3525	$0,358 \pm 0,013$	0,45
NNet 3	0,2793	$0,285 \pm 0,013$	0,41
NNet 4	0,2388	$0,245 \pm 0,016$	0,37
NNet 5	0,2025	$0,211 \pm 0,020$	0,40

TAB. 5.2: Valeur Monte-Carlo et résultat de chaque ajustement des fractions de mauvais étiquetage. Les erreurs sont statistiques uniquement.

FIG. 5.5: Masse invariante $\mu\mu K\pi$ des événements du lot $B_{u,d} \rightarrow J/\psi X$, ajustée par le modèle décrit dans ce chapitre.

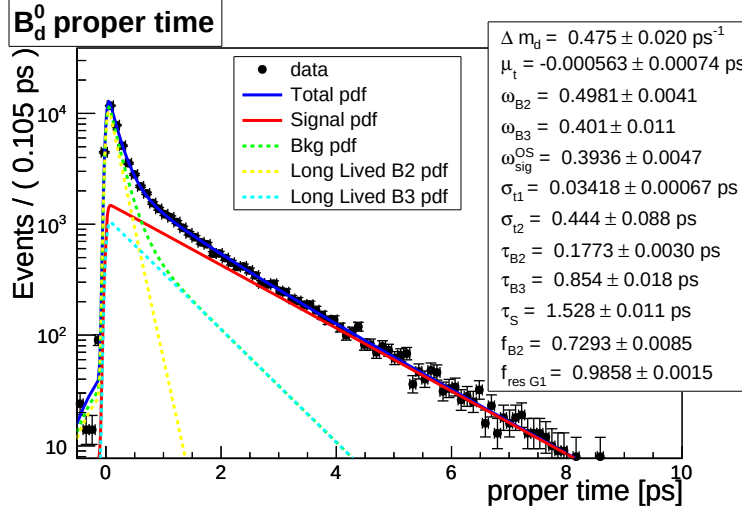


FIG. 5.6: Temps propre des événements du lot $B_{u,d} \rightarrow J/\psi X$, ajusté par le modèle décrit dans ce chapitre.

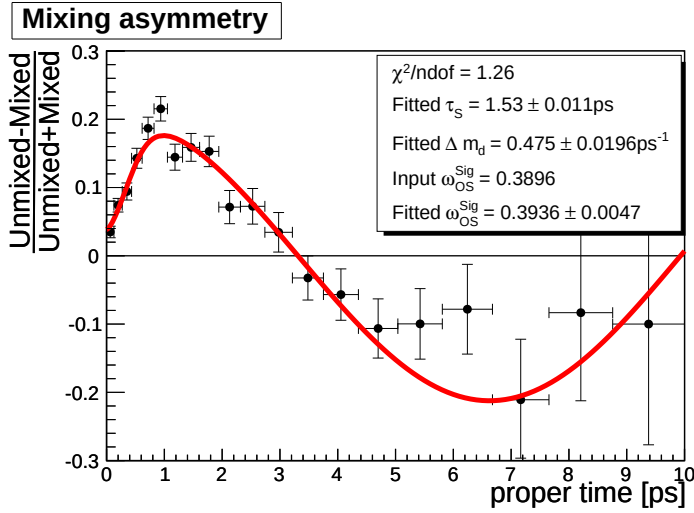


FIG. 5.7: Asymétrie de mélange des événements du lot $B_{u,d} \rightarrow J/\psi X$, ajustée par le modèle décrit dans ce chapitre. Le χ^2/nDoF correspond à l'écart entre la courbe de la PDF et les données. La valeur $\text{Input}\omega_{OS}^{\text{sig}}$ est celle de la fraction de mauvais étiquetage du côté opposé du signal mesuré dans la vérité Monte-Carlo.

estimateurs peuvent servir à l'étiquetage d'un événement. Cette méthode permet cependant de mesurer les fractions de mauvais étiquetage de chaque estimateur. Les résultats sont donnés dans la Table 5.2(b). Toutes les fractions de mauvais étiquetage sont en accord avec les valeurs issues de la vérité Monte-Carlo. C'est un test de la stabilité du modèle.

Enfin, les événements ont été séparés en cinq catégories indépendantes en fonction de leurs propriétés d'étiquetage, telles que définies dans la section 3.2.3. L'ajustement est fait simultanément sur ces cinq lots pour obtenir les cinq fractions de mauvais étiquetage. Les résultats sont donnés dans la Table 5.2(c). Les valeurs ajustées sont compatibles avec les valeurs d'entrée. Les asymétries de mélange des cinq catégories sont données dans la Figure 5.8. La valeur ajustée de $\Delta m_d = (0,485 \pm 0,015) \text{ ps}^{-1}$ montre une amélioration d'environ 30% par rapport à la sensibilité donné dans la Table 5.1 ($\Delta m_d = (0,475 \pm 0,020) \text{ ps}^{-1}$). C'est en accord avec l'amélioration attendue en changeant seulement la fraction de mauvais étiquetage, de $\omega_{OS}^{\text{sig}}(\text{moyen}) = 0,39$ à $\omega_{OS}^{\text{sig}}(\text{comb.}) = 0,36$. La sensibilité à Δm_d est supposée s'améliorer car [60] :

$$\frac{1}{\sigma(\Delta m_d)} \sim \sqrt{\frac{N}{2}} f_s (1 - 2\omega) e^{-\frac{(\sigma_t \Delta m_d)^2}{2}},$$

où N est le nombre d'événements disponibles et f_s la pureté en signal du lot. De ce fait nous attendons une amélioration de la sensibilité d'un facteur :

$$\frac{1 - 2\omega_{OS}^{\text{sig}}(\text{comb})}{1 - 2\omega_{OS}^{\text{sig}}(\text{moyen})} = 1,25$$

Dans les tables de résultats 5.1 et 5.2, nous avons indiqué les valeurs des paramètres extraites de la vérité Monte-Carlo en regardant l'ancêtre des particules dans l'état final. Ceci permet de vérifier que l'ajustement du modèle sur les données donne des paramètres corrects.

Seuls les événements dont toutes les particules dans l'état final sont associées aux particules du Monte-Carlo sont utilisés dans cette phase. La probabilité qu'un candidat B_d^0 ait ses quatre filles finales reconstruites associées dans le Monte-Carlo est $89,2 \pm 0,1\%$. Parmi les événements associés, demander que les quatre filles du méson de signal soient correctement identifiées, c'est-à-dire que leur équivalent dans le Monte-Carlo soit la même particule, a une probabilité de $75,1 \pm 0,2\%$.

Les catégories considérées sont les suivantes :

- toutes les particules de l'état final ont le même ancêtre, et c'est un B_d^0 : c'est un événement de signal. $43,1 \pm 0,2\%$ du total des événements sont du signal ;
- les deux muons sont issus du même méson B_d^0 , et les deux hadrons viennent du vertex primaire, via n'importe quelle résonance : c'est le bruit de fond de type B2. La vérité Monte-Carlo contient $66,6 \pm 0,3\%$ événements de types B2 par rapport à l'ensemble du bruit ;

- les deux muons viennent d'un méson B_d^0 . Un des deux hadrons vient aussi du même B_d^0 , alors que le second vient du vertex primaire. C'est le bruit de type B3. $33,4 \pm 0,3\%$ des événements de bruit de fond sont de type B3.

Les événements qui n'ont pas été associés à la vérité Monte-Carlo peuvent être du signal et du bruit de fond, car deux effets s'ajoutent :

1. au moins une des traces utilisées est une fausse trace (*ghost*), qui n'a pas d'équivalent dans le Monte-Carlo (moins de 70% de ses hits sont associés),
2. au moins une des traces utilisées a une équivalence dans le Monte-Carlo, mais l'associateur a échoué.

Le deuxième effet implique que même de bonnes traces de signal peuvent être considérées comme fausses, ce qui explique le pic dans la distribution de masse invariante de la Figure 5.9. De ce fait, les fractions données précédemment ne reflètent pas exactement la réalité, mais donnent une estimation des proportions. L'ajustement des données produit des fractions qui sont du même ordre de grandeur, et montre donc que le modèle décrit les données.

Un dernier ajustement en séparant les mésons étiquetés B_d^0 et $\bar{B}_d^0$ a été effectué. Dans cet ajustement simultané des deux lots, les seuls paramètres différents sont les fractions des mauvais étiquetages $\omega_{B_d^0}$ et $\omega_{\bar{B}_d^0}$. Les valeurs extraites de cet ajustement sont données dans la Table 5.3, avec leur équivalent de la vérité Monte-Carlo. On peut remarquer que ces deux valeurs sont compatibles entre elles à 1,3 déviations standards, et qu'elles sont compatibles chacune avec leur valeur vraie. Dans l'étude de la mesure de $\sin(2\beta)$ présentée ici, nous n'utilisons pas ces deux valeurs différentes car ce niveau de détail n'est pas nécessaire ici. Elles seront cependant nécessaires dans l'ajustement de $\sin(2\beta)$ dans les données réelles lorsque le nombre d'événement sera suffisamment grand pour observer une différence.

L'ajustement d'un lot de données simulées contenant des événements de signal et de bruit de fond à grand temps propre, le lot $B_{u,d} \rightarrow J/\psi X$, montre que le modèle décrit bien le comportement des différentes composantes.

5.2.2 Avec le lot de J/ψ inclusifs

Un élément est absent du lot de $B_{u,d} \rightarrow J/\psi X$, les événements *prompt*, dominant le bruit de fond comme le montre la Table 4.19.

Le lot utilisé pour déterminer leurs propriétés est le lot d'événements J/ψ inclusifs décrit dans le chapitre 4 dont les événements $pp \rightarrow b \Rightarrow J/\psi X$ ont été retirés. Pour simplifier, ce lot est appelé *prompt*.

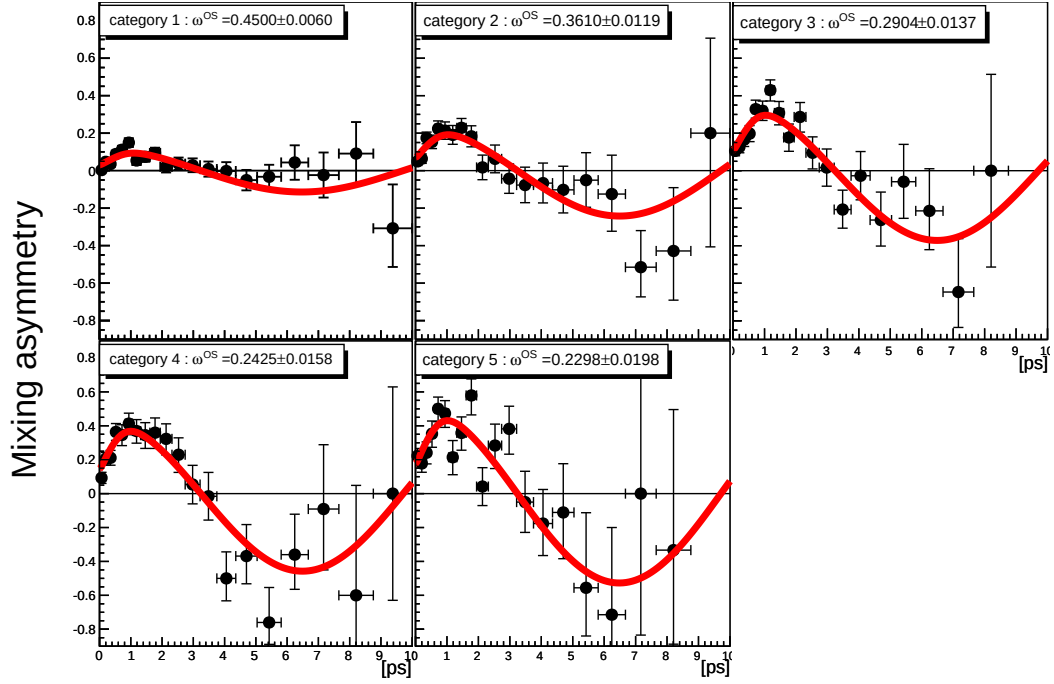


FIG. 5.8: Asymétries de mélange des événements du lot $B_{u,d} \rightarrow J/\psi X$ pour les cinq catégories d'étiquetage.

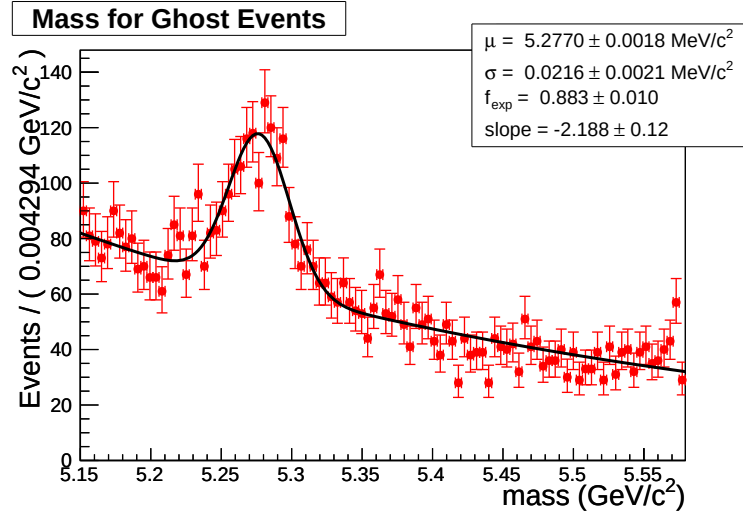


FIG. 5.9: Masse invariante $\mu\mu K\pi$ pour les événements du lot $B_{u,d} \rightarrow J/\psi X$ pour lesquels toutes les traces de l'état final n'ont pas été associées à la vérité Monte-Carlo. Le pic s'explique par l'échec de l'associateur à faire correspondre des vraies traces de signal.

	valeur MC (%)	valeur ajustée (%)
$\omega_{B_d^0}$	$38,59 \pm 0,46$	$38,77 \pm 0,66$
$\omega_{\bar{B}_d^0}$	$39,32 \pm 0,46$	$39,95 \pm 0,66$

TAB. 5.3: Fractions de mauvais étiquetage pour les mésons étiquetés B_d^0 et $\bar{B}_d^0$ extraites de l'ajustement simultané des données.

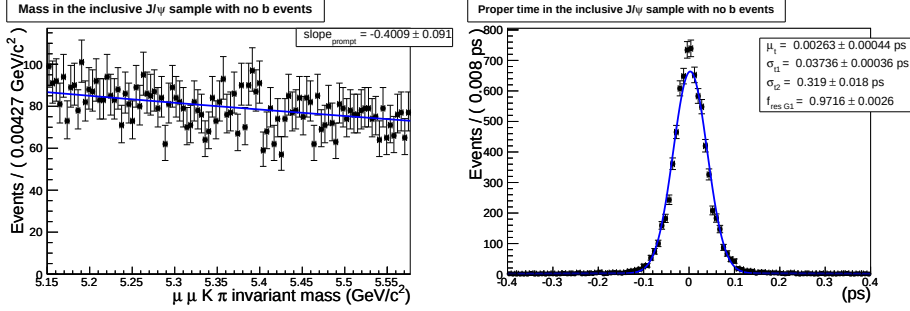


FIG. 5.10: Gauche : distribution de masse invariante des événements *prompt* passant la sélection $B_d^0 \rightarrow J/\psi K^{*0}$. La distribution est ajustée par une exponentielle décroissante. Droite : distribution du temps propre des mêmes événements. Elle est ajustée par la somme de deux gaussiennes dont la largeur moyenne est $\langle \sigma \rangle = 65$ ps.

Après la sélection, le déclenchement de niveau 0 et l'étiquetage du côté opposé, il ne reste que 1690 événements dans le lot *prompt*. Les distributions de masse invariante et de temps propre correspondantes sont données dans la Figure 5.10. La résolution en temps mesurée dans ce lot est comparable à celle trouvée dans l'ajustement précédent, $\langle \sigma \rangle = 65$ ps. La pente utilisée pour la modélisation de la masse invariante est $\text{slope}_{\text{prompt}} = -0,4 (\text{GeV}/c^2)^{-1}$, et les paramètres de la résolution en temps pour ces événements sont $\mu_t = -0,00263$ ps, $\sigma_{t1} = 0,03736$ ps, $\sigma_{t2} = 0,319$ ps, et $f_{\text{res G1}} = 0.9716$. Étant donné que ces valeurs sont utilisés comme paramètre d'entrée dans les simulation simplifiées, leurs erreurs ne sont pas prises en compte.

5.3 Simulations simplifiées

Pour déterminer les fractions de signal et des bruits de fond, ainsi que le nombre d'événements à générer dans les simulations simplifiées, nous utilisons le lot d'événements de J/ψ inclusifs dans lesquels les chaînes de désintégration $pp \rightarrow b \Rightarrow J/\psi X$ sont présentes. Après la sélection des événements $B_d^0 \rightarrow J/\psi K^{*0}$, le déclenchement de niveau 0 et l'étiquetage de la saveur utilisant le côté opposé, nous avons à disposition 2637 événements. Parmi ces événements, le Monte-Carlo nous donne une fraction de signal $f_S = 7\%$. Le bruit de fond *prompt* représente $f_{\text{prompt}} = 67\%$ du bruit de fond total. Les fractions utilisées sont résumées dans la

fraction	valeur
f_S	7%
f_{prompt}	67%
f_{B2}	68%

TAB. 5.4: Fractions de signal, *prompt*, et bruit de fond à grand temps de vie utilisées dans les simulations simplifiées. Elles sont présentées dans la section 5.1.3.

Table 5.4. D'après la Table 4.1, les événements du lot de J/ψ inclusifs correspondent à une luminosité de $0,00078 \text{ fb}^{-1}$. Pour économiser le temps de calcul, nous effectuons des simulations simplifiées de 100 000 événements, correspondantes à une luminosité de $0,03 \text{ fb}^{-1}$.

Pour tester la stabilité du modèle utilisé, nous effectuons un ensemble de simulations simplifiées et vérifions que les valeurs ajustées sont compatibles avec les valeurs d'entrées. Ces vérifications se font par l'étude des *pulls* des valeurs ajustées, définis comme :

$$pull = \frac{x_{\text{ajustée}} - x_{\text{entrée}}}{\sigma(x_{\text{ajustée}})}.$$

Les distributions des pulls doivent être des distributions gaussiennes normales centrées sur 0 pour que les paramètres puissent être considérés comme estimés correctement. La largeur des pulls doit être égale à 1 pour que l'erreur des paramètres soit correctement estimée.

Les événements générés ont une masse invariante de méson B_d^0 comprise entre $5,15 \text{ GeV}/c^2$ et $5,58 \text{ GeV}/c^2$, et un temps propre compris entre $-0,5 \text{ ps}$ et 10 ps .

Les résultats d'un exercice sont donnés dans la Table 5.5. Toutes les valeurs ajustées correspondent aux valeurs d'entrées. En particulier, la valeur de la fraction de mauvais étiquetage du signal est compatible avec la valeur d'entrée. Les distributions de masse invariante, de temps propre, et l'asymétrie de mélange sont données dans les Figures 5.11, 5.12 et 5.13 respectivement.

Un ensemble de 899 exercices similaires a été effectué en changeant la racine à la génération. Les résultats de ces exercices sont donnés dans la Table 5.6. Tous les paramètres se comportent suivant la loi Normale centrée sur 0. Seule l'erreur sur la pente de la masse invariante de la troisième composante de bruit de fond (slope_{B3}) est surestimée.

La distribution du pull de la fraction de mauvais étiquetage du signal est montré dans la Figure 5.14. Elle est parfaitement compatible avec la valeur attendue, montrant la stabilité du modèle. L'erreur relative statistique obtenue de ces exercices est :

$$\frac{\sigma(\omega)}{\omega} = 0,010/0,390 = 2,6\%,$$

qui correspond à une sensibilité statistique de $0,3\%$ pour 2 fb^{-1} de données. En ef-

Paramètre	Valeur MC	Valeur ajustée	Corrélation globale
f_s	0,07	$0,0717 \pm 0,0011$	0,454
$\mu_{G1} [\text{GeV}/c^2]$	5,2775	$5,2774 \pm 0,0003$	0,106
$\sigma_{G1} [\text{GeV}/c^2]$	0,0166	$0,0167 \pm 0,0002$	0,362
$\omega_{\text{sig}}^{\text{OS}}$	0,3896	$0,3903 \pm 0,0097$	0,300
$\tau_{B_d^0} [\text{ps}]$	1,536	$1,501 \pm 0,021$	0,320
$\Delta m_d [\text{ps}^{-1}]$	0,502	$0,477 \pm 0,028$	0,262
f_{prompt}	0,675	$0,675 \pm 0,003$	0,698
$\text{slope}_{\text{prompt}} [(\text{GeV}/c^2)^{-1}]$	-0,4	$-0,41 \pm 0,04$	0,275
$\mu_t [\text{ps}]$	-0,00056	$-0,00065 \pm 0,00018$	0,588
$f_{\text{res } G1}$	0,986	$0,987 \pm 0,001$	0,445
$\sigma_{t1} [\text{ps}]$	0,034	$0,0339 \pm 0,0001$	0,480
$\sigma_{t2} [\text{ps}]$	0,444	$0,447 \pm 0,066$	0,630
f_{B2}	0,68	$0,69 \pm 0,01$	0,856
$\text{slope}_{B2} [(\text{GeV}/c^2)^{-1}]$	-0,67	$-0,61 \pm 0,1$	0,612
ω_{B2}	0,4999	$0,5003 \pm 0,0056$	0,511
$\tau_{B2} [\text{ps}]$	0,173	$0,176 \pm 0,004$	0,782
$\text{slope}_{B3} [(\text{GeV}/c^2)^{-1}]$	-5,22	$-5,52 \pm 0,16$	0,572
ω_{B3}	0,401	$0,411 \pm 0,011$	0,571
$\tau_{B3} [\text{ps}]$	0,85	$0,851 \pm 0,018$	0,771

TAB. 5.5: Résultat d'un exercice Monte-Carlo pour la désintégration $B_d^0 \rightarrow J/\psi K^{*0}$, modèle complet. La première colonne définit les paramètres, la seconde leur valeurs d'entrée, la troisième leur valeur ajustée, et la dernière les corrélations globales, c'est-à-dire la dépendance d'un paramètre par rapport aux autres. Aucun paramètre n'est biaisé.

fet, la sensibilité de la fraction de mauvais étiquetage varie en $1/\sqrt{N}$ où N est le nombre d'événements. Pour Δm_d , la sensibilité statistique ramenée à 2 fb^{-1} de données est $\sigma(\Delta m_d) = 0,003 \text{ ps}^{-1}$. Cette erreur, qui ne prend pas en compte les catégories d'étiquetage, présente une amélioration de 40% relative par rapport à l'erreur totale, statistique et systématique, sur la moyenne mondiale, $\Delta m_d = 0,507 \pm 0,005 \text{ ps}^{-1}$ [13]. Selon le résultat de l'étude sur les données de simulation complète, section précédente, nous attendons une amélioration de 25% de l'erreur statistique en utilisant un ajustement simultané des catégories d'étiquetage.

5.3.1 Erreurs systématiques

Comme aucun paramètre du modèle n'est fixé dans cette étude, les sources d'erreurs systématiques peuvent être les suivantes :

- tout d'abord, le modèle lui-même. En particulier, la description du bruit de fond est limitée par les données issues de la simulation complète. Lorsque les

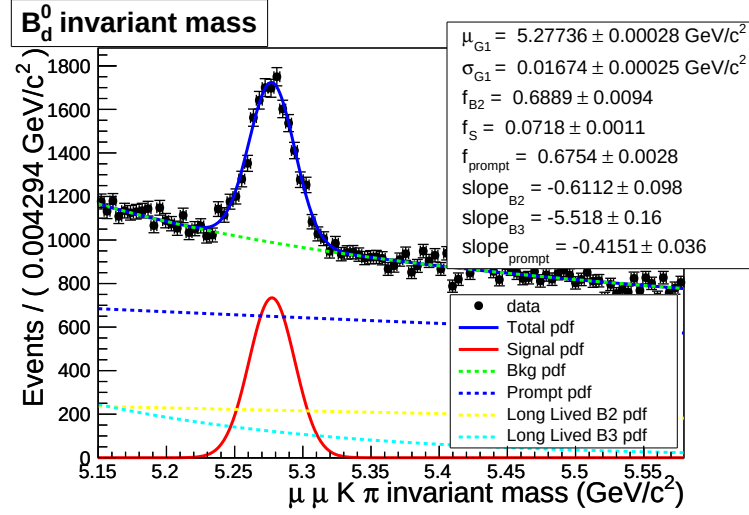


FIG. 5.11: Distribution de masse invariante $\mu\mu K\pi$ pour un exercice Monte-Carlo ajustée avec le modèle complet. Chaque contribution est représentée par une couleur différente. La Table 5.5 donne les valeurs ajustées des paramètres.

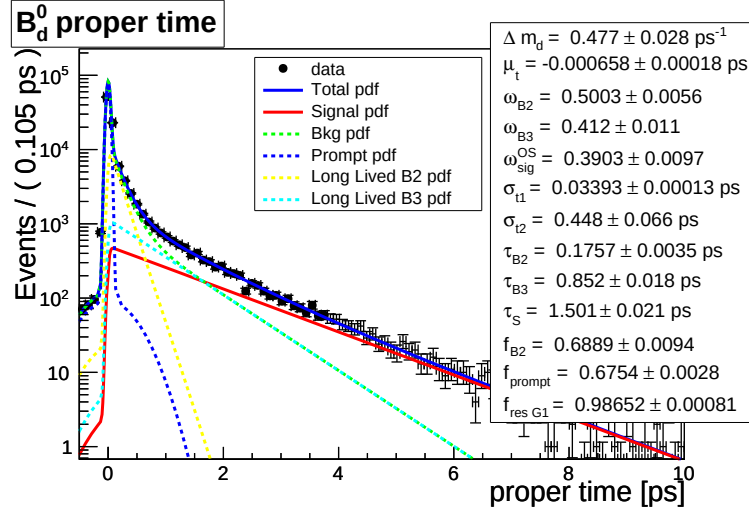


FIG. 5.12: Temps propre d'un exercice Monte-Carlo ajusté avec le modèle complet. Chaque contribution est représentée par une couleur différente. La Table 5.5 donne les valeurs ajustées des paramètres.

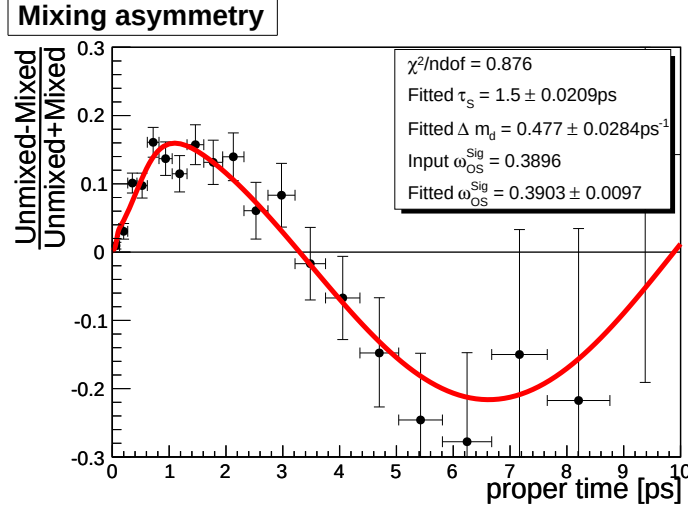


FIG. 5.13: Asymétrie de mélange d'un exercice Monte-Carlo ajustée avec le modèle complet. La valeur de la fraction de mauvais étiquetage ajustée est compatible avec la valeur d'entrée.

	Valeur centrale	moyenne du pull (μ)	$\frac{\mu-0}{\sigma(\mu)_{\text{stat}}}$	largeur du pull (σ)	$\frac{\sigma-1}{\sigma(\sigma)_{\text{stat}}}$
f_S	0.07	0.017 ± 0.035	0.47	0.989 ± 0.026	-0.40
μ_{G1}	$5.2775 \text{ GeV}/c^2$	0.020 ± 0.035	0.57	0.990 ± 0.029	-0.32
σ_{G1}	$0.0166 \text{ GeV}/c^2$	-0.025 ± 0.035	-0.70	0.985 ± 0.031	-0.46
$\omega_{\text{sig}}^{\text{OS}}$	0.3896	-0.012 ± 0.036	-0.34	0.974 ± 0.030	-0.84
$\tau_{B_d^0}$	1.536 ps	-0.061 ± 0.038	-1.60	1.037 ± 0.033	1.13
Δm_d	0.502 ps^{-1}	-0.007 ± 0.034	-0.20	0.959 ± 0.028	-1.46
f_{prompt}	0.675	0.062 ± 0.034	1.81	0.960 ± 0.025	-1.59
$\text{slope}_{\text{prompt}}$	$-0.4 (\text{GeV}/c^2)^{-1}$	-0.006 ± 0.037	-0.15	1.024 ± 0.030	0.79
μ_t	-0.00056 ps	0.002 ± 0.034	0.06	0.948 ± 0.029	-1.80
$f_{\text{res G1}}$	0.986	0.003 ± 0.032	0.10	0.989 ± 0.031	-0.35
σ_{t1}	0.034 ps	-0.026 ± 0.035	-0.74	0.995 ± 0.028	-0.19
σ_{t2}	0.444 ps	-0.056 ± 0.036	-1.54	0.928 ± 0.030	-2.37
f_{B2}	0.68	0.090 ± 0.035	2.57	0.980 ± 0.029	-0.69
slope_{B2}	$-0.67 (\text{GeV}/c^2)^{-1}$	-0.041 ± 0.033	-1.22	0.925 ± 0.030	-2.52
ω_{B2}	0.4999	0.018 ± 0.036	0.49	1.006 ± 0.032	0.20
τ_{B2}	0.173 ps	0.083 ± 0.034	2.42	0.974 ± 0.028	-0.94
slope_{B3}	$-5.22 (\text{GeV}/c^2)^{-1}$	-0.036 ± 0.033	-1.11	0.898 ± 0.026	-4.01
ω_{B3}	0.401	-0.013 ± 0.034	-0.40	0.958 ± 0.026	-1.62
τ_{B3}	0.85 ps	0.068 ± 0.033	2.06	0.937 ± 0.027	-2.31

TAB. 5.6: Résultat de 899 exercices générés et ajustés avec le modèle complet. Aucun paramètre n'est biaisé, seule la pente de l'exponentielle du troisième bruit de fond voit son erreur surestimée. Les termes $\mu - 0/\sigma(\mu)$ et $\sigma - 1/\sigma(\sigma)$ sont les pulls de la moyenne et de l'erreur.

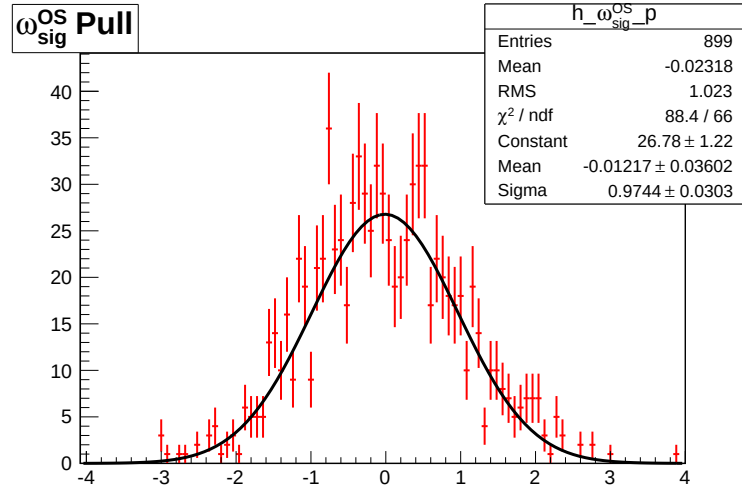


FIG. 5.14: Pull de la fraction de mauvais étiquetage du signal, extrait de 899 exercices d'ajustement.

données réelles seront disponibles, il sera alors possible d'effectuer des variations des descriptions pour étudier leur impact sur la fraction de mauvais étiquetage.

- ensuite l'algorithme d'étiquetage. Cet outil utilise de nombreuses étapes pour limiter les corrélations entre les estimateurs. Mais cette complexité le rend plus enclin à induire des erreurs systématiques. Par exemple, les polynômes utilisés pour paramétrer les η_i de la section 3.2.3, ou ceux utilisés pour les ω_i , ne sont pas absolus, chacun des paramètres possède son erreur. Il est nécessaire de comprendre quelles sont les variations induites sur la fraction de mauvais étiquetage mesurée par une modification de ces paramètres. Ces études d'erreur sont repoussées au moment où des données réelles seront disponibles.
- Enfin, le modèle de résolution en temps. Dans cette étude, le modèle utilisé prend une largeur constante pour l'ensemble des événements. Or la résolution dépend elle-même du temps propre [61]. Du fait de la faible valeur de Δm_d , on s'attend à ce que l'impact d'un tel choix soit faible, mais doit être étudié.

5.4 Conclusion

Dans ce chapitre, nous avons présenté la procédure utilisée pour mesurer la fraction de mauvais étiquetage à l'aide du canal de contrôle $B_d^0 \rightarrow J/\psi K^{*0}$. Le bruit de fond se compose de trois parties : une *prompt*, dans laquelle toutes les partic-

ules sont issues du vertex primaire, une à grand temps propre dans laquelle deux traces sur quatre sont issues du vertex primaire, et enfin une autre à grand temps propre mais dans laquelle une seule particule sur quatre est issue du vertex primaire. La modélisation de ces bruits de fond utilise leurs propriétés cinématiques. Un ajustement sur un lot de données issues de la simulation complète permet de valider une partie de ce modèle, alors qu'un ensemble d'exercices Monte-Carlo simplifiés permet d'estimer la sensibilité statistique attendue. Nous montrons qu'avec cette méthode, la sensibilité statistique à la fraction de mauvais étiquetage attendue pour 2 fb^{-1} de donnée est $\sigma(\omega)/\omega = 0,3\%$. Pour la fréquence des oscillations, nous obtenons $\sigma(\Delta m_d)_{\text{stat}} = 0,003 \text{ ps}^{-1}$. On s'attend à ce que les sources d'erreurs systématiques dans cette analyse proviennent principalement du modèle utilisé, puis de l'algorithme d'étiquetage. Ces études seront à traiter lorsque les données réelles seront disponibles.

Le prochain chapitre utilise les résultats de la mesure de la fraction de mauvais étiquetage pour étudier la sensibilité de LHCb à la mesure de $\sin(2\beta)$.

