

RECHERCHE DES DÉSINTÉGRATIONS $B \rightarrow K^{(*)} \nu \bar{\nu}$ À BELLE II

HABILITATION À DIRIGER DES RECHERCHES

École doctorale de physique et chimie-physique de l'Université de Strasbourg

GIULIO DUJANY

Institut Pluridisciplinaire Hurien Curiel
IN2P3 – CNRS

10 Septembre 2026

TABLE DES MATIÈRES

Introduction	9
0.1 Choix de productions scientifiques	10
1 Contexte scientifique	13
1.1 L'expérience Belle II	13
1.2 Les transitions $b \rightarrow s\nu\bar{\nu}$	14
2 Recherche de $B^+ \rightarrow K^+\nu\bar{\nu}$	19
2.1 Sélection	20
2.2 Corrections de la simulation, incertitudes systématiques et ajustement	21
2.3 Résultats	26
3 Le GraFEI	29
4 Recherche de $B^+ \rightarrow K^+$ axion	37
5 Perspectives pour $B \rightarrow K^{(*)}\nu\bar{\nu}$	41
5.1 Recherche de $B \rightarrow K^{(*)}\nu\bar{\nu}$ avec l'approche STA	41
5.2 Perspectives à plus long terme	44
Conclusions	47
Références	49

ABSTRACT

Nom de l'université : Université de Strasbourg

Nom du candidat : Giulio Dujany

Titre du diplôme : Habilitation à diriger des recherches

Garant : DR. Isabelle Ripp-Baudot

Titre du document : Recherche des désintégrations $B \rightarrow K^{(*)} \nu \bar{\nu}$ à Belle II

Date de submission : 10 Septembre 2026

La recherche de désintégrations rares des mésons B, en particulier les transitions $b \rightarrow s \nu \bar{\nu}$, constituent un moyen puissant pour explorer la physique au-delà du Modèle Standard de la physique des particules. Ce document présente les avancées récentes dans la recherche de la désintégration $B^+ \rightarrow K^+ \nu \bar{\nu}$ au sein de l'expérience Belle II, située auprès du collisionneur SuperKEKB. Grâce à une approche combinant des méthodes inclusives et exclusives, une première évidence de ce processus a été obtenue avec une significativité de 3,5 écarts type, révélant une tension de 2,7 écarts type par rapport aux prédictions du Modèle Standard. De plus, une réinterprétation des résultats a permis de poser des limites contraignantes sur l'existence de l'axion de QCD, améliorant d'un ordre de grandeur les contraintes précédentes. Le développement du GraFEI, un réseau de neurones de type graphe, est également détaillé. Cet outil permet une reconstruction générique et efficace des désintégrations des mésons B. Enfin, les perspectives futures incluent l'utilisation de l'approche semi-leptonique (STA) pour affiner la mesure de la fraction d'embranchement de $B^+ \rightarrow K^+ \nu \bar{\nu}$, ainsi que l'exploration de la désintégration $B^0 \rightarrow K^{*0} \nu \bar{\nu}$, qui pourrait offrir des contraintes complémentaires sur les modèles de nouvelle physique.

On fait la science avec des faits comme une maison avec des pierres mais une accumulation de faits n'est pas plus une science qu'un tas de pierres n'est une maison.

— Henri Poincaré

Nombreuses sont les personnes à qui
je suis redevable et qui ont rendu cette thèse possible, ma
famille, en particulier ma femme Marie-Lys et mes enfants Paul et
Martin, mes amis, les autres membres de Belle II spécialement ceux du groupe
Belle II à l'IPHC. Je souhaite remercier particulièrement Isabelle et Jérôme pour
m'avoir accueilli fraternellement dans le groupe, pour les échanges humains et
scientifiques et pour m'avoir confié leurs étudiants à superviser. Je remercie égale-
ment les étudiants et postdocs que j'ai supervisé pour produire le contenu de ce do-
cument : Léonard, Lucas, Ilias, Arthur, Jacopo, Corentin et Merna. Je remercie aussi
les collègues impliqués dans l'analyse hadronique de $B^+ \rightarrow K^+ \nu \bar{\nu}$ Elisa, Stefano,
Claudia, Roberta et les collègues de l'analyse inclusive Sasha, Sally, Filippo, Cy-
rille, Yubo, Yulan et Eldar. Je remercie par ailleurs James, Oskar, Lea, Tobias
et Markus pour le travail ensemble sur les réseaux de neurones profonds.
Je remercie pareillement Diego, MÉRIL, Claudio et Axel pour le travail
phénoménologique ensemble. Un remerciement aussi à l'IN2P3
qui me paye pour faire ce que j'aime et l'ANR qui a financé
une partie des recherches documentées ici. Un re-
merciement enfin à Yasmine, Olcyr, Anne-
Catherine et Karim d'avoir accepté
de faire partie de mon
jury de HDR.



INTRODUCTION

La recherche des transitions $b \rightarrow s\nu\bar{\nu}$ est mon sujet de recherche principal depuis que j'ai rejoint le groupe Belle II de l'IPHC. Avec le soutien d'Isabelle Ripp-Baudot, j'ai initié cette activité dans le groupe et j'ai supervisé plusieurs étudiants et un postdoc sur ce sujet.

Ces transitions, qui dans le Modèle Standard de la physique des particules procèdent via des courants neutres qui changent la saveur, sont parmi les plus prometteuses pour trouver des signes de la nouvelle physique nécessaire pour compléter le Modèle Standard. De plus la paire de neutrinos de l'état final $K^{(*)}\nu\bar{\nu}$ n'est pas reconstruite explicitement, mais seulement déduite par la présence de particules manquantes. Cette recherche peut donc permettre de découvrir aussi des nouvelles particules invisibles comme celles qui pourraient constituer la matière noire.

Belle II est la seule expérience au monde permettant d'observer ce type de désintégrations. Des paires de mésons B sont produites dans des collisions entre électrons et positrons par l'accélérateur le plus lumineux au monde. Comme les neutrinos ne sont pas détectables, il est nécessaire de reconstruire complètement l'événement pour contraindre la cinématique et identifier correctement le signal.

Ce document est organisé de la manière suivante.

[Le premier chapitre](#) offre le contexte scientifique de mes recherches et une description succincte de l'expérience Belle II.

[Le deuxième chapitre](#) décrit la première évidence de $B^+ \rightarrow K^+\nu\bar{\nu}$ obtenue à Belle II et qui présente une tension par rapport à la prédiction du Modèle Standard.

[Le troisième chapitre](#) décrit le développement d'un réseau de neurones profond pour reconstruire une désintégration générique d'un méson B. Ce réseau peut être utilisé pour la reconstruction des événements avec particules non détectées dans l'état final comme $B \rightarrow K^{(*)}\nu\bar{\nu}$.

[Le quatrième chapitre](#) décrit la réinterprétation de l'évidence de $B^+ \rightarrow K^+\nu\bar{\nu}$ décrite dans le premier chapitre en terme de $B^+ \rightarrow K^+a$ où a est un axion ou une particule de type axion.

[Le dernier chapitre](#) Décrit la recherche en cours de $B^+ \rightarrow K^+\nu\bar{\nu}$ et $B^0 \rightarrow K^{*0}\nu\bar{\nu}$ en utilisant la méthode du tag semi-leptonique et offre quelques perspectives futures.

Pour tous les travaux décrits dans ce document, j'ai été un des analystes principaux et j'ai supervisé des étudiants ou postdocs. Dans ce document, j'adopte la convention $c = 1$ et la conjugaison de charge est implicite.

0.1 CHOIX DE PRODUCTIONS SCIENTIFIQUES

Mes publications récentes en lien avec ce document sont les suivantes :

1. Réf. [1] décrit la première évidence de la désintégration $B^+ \rightarrow K^+ \nu \bar{\nu}$. Cette désintégration est recherchée avec deux approches : une approche inclusive qui permet une grande efficacité et une approche exclusive moins efficace, mais plus pure où on reconstruit les désintégrations hadroniques de l'autre méson B de l'évènement. J'ai été, avec le doctorant Lucas Martel et le postdoc Jacopo Cerasoli que je supervisais, un des analystes principaux en me focalisant principalement sur l'approche hadronique. La désintégration $B^+ \rightarrow K^+ \nu \bar{\nu}$ est sensible à plusieurs types de possibles physiques au-delà du Modèle Standard. En outre la fraction d'embranchement obtenue est en tension de 2,7 écarts type avec le Modèle Standard. Ce résultat a donc été accueilli avec grand intérêt par la communauté. Ce document est accessible au lien <https://arxiv.org/pdf/2311.14647> et est décrit au chapitre 2.
2. Réf. [2] présente une méthode basée sur l'apprentissage automatique pour reconstruire des arbres enracinés (représentant des interactions hiérarchiques) à partir uniquement des feuilles, en utilisant une nouvelle représentation appelée matrice des Générations du Plus Bas Ancêtre Commun (LCAG). Cette méthode est appliquée à la reconstruction de la désintégration générique d'une particule dans un modèle jouet et présente des résultats encourageants avec plus de 90% de désintégrations correctement reconstruites pour des désintégrations ayant jusqu'à six particules dans l'état final. Ce travail a jeté les bases pour le développement du GraFEI et est né d'une collaboration entre mes stagiaires Ilias Tsaklidis et Arthur Thaller avec d'autres collègues de Karlsruhe et Bonn. Ce document est accessible au lien <https://arxiv.org/pdf/2208.14924> et est décrit au chapitre 3.
3. Réf. [3] présente le GraFEI qui est un modèle d'apprentissage automatique basé sur des réseaux neuronaux de type graphe permettant de reconstruire de manière générique la désintégration d'un méson B dans l'expérience Belle II. Cet algorithme peut être utilisé pour reconstruire des désintégrations avec de l'énergie manquante comme $B^+ \rightarrow K^+ \nu \bar{\nu}$. J'ai initié le développement de cet algorithme et supervisé les étudiants et le postdoc qui l'ont développé, optimisé et intégré dans le code de Belle II. Ce document est accessible au lien <https://arxiv.org/pdf/2503.09401> et est décrit au chapitre 3.
4. Réf. [4] présente la réinterprétation de l'analyse en Réf. [1] pour rechercher l'axion de QCD. En effet, l'axion, ou une autre particule similaire, pourrait ne pas laisser de trace dans le détecteur Belle II comme les neutrinos de $B^+ \rightarrow K^+ \nu \bar{\nu}$. Cette réinterprétation permet de poser la

limite la plus contraignante au monde sur l'échelle de Peccei-Quinn divisée par le couplage de l'axion avec les quarks s et b . Ce travail a été effectué en collaboration avec les théoriciens Diego Guadagnoli, Axel Iohner Claudio Toni et mon étudiante Merna Abumusabh. Ce document est accessible au lien <https://arxiv.org/pdf/2510.18953> et est décrit au chapitre 4.

Le Modèle Standard de la physique des particules a été couronné de succès depuis sa formulation dans la deuxième moitié du XX^e siècle jusqu'à la découverte de sa dernière pièce manquante, le boson de Higgs en 2012. Cet événement a toutefois laissé la communauté de la physique des particules dans une situation contradictoire. D'une part le Modèle Standard a prédit de nombreuses particules qui ont été ensuite découvertes, il a atteint un niveau d'accord inédit entre prédictions théoriques et mesures expérimentales et a permis d'augmenter considérablement notre connaissance de l'univers à son échelle la plus petite. D'autre part, malgré ses succès, le Modèle Standard est une théorie effective qui nécessite d'être complétée par de la "nouvelle physique" à une échelle d'énergie supérieure. Il ne peut pas expliquer, par exemple, l'asymétrie observée entre matière et antimatière ou encore la nature de la matière obscure et de l'énergie obscure qui constituent plus de 95% de notre univers.

Deux approches complémentaires sont aujourd'hui utilisées dans la recherche de la nouvelle physique. Dans la première approche, des nouvelles particules sont recherchées parmi les produits des collisions à haute énergie. Cette approche est surtout adoptée au LHC. La deuxième approche consiste dans la mesure de précision de phénomènes desquels on connaît avec grande précision la prédiction théorique. La nouvelle physique pourrait amener des déviations significatives entre mesure et prédiction théorique. Cette deuxième approche peut tester des échelles d'énergie beaucoup plus hautes que celle du collisionneur utilisé, elle pourrait donc permettre de découvrir de la nouvelle physique aussi dans le cas où la première approche échouerait. De plus, dans le passé, les mesures de précision ont porté à la découverte de nombreuses nouvelles particules, par exemple trois des six quarks.

1.1 L'EXPÉRIENCE BELLE II

L'expérience Belle II [5] est située à Tsukuba au Japon auprès du collisionneur électron-positron SuperKEKB. Elle est une actrice de premier plan dans la recherche de la nouvelle physique via des mesures de précision. Dans cette expérience, des paires de mésons B intriqués quantiquement sont produits dans des collisions entre électrons et positrons via la désintégration

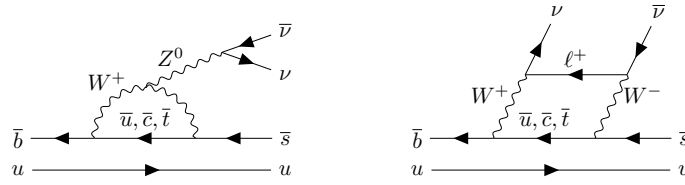


FIGURE 1.1 – Diagrammes de Feynman principaux contribuant à la désintégration $B^+ \rightarrow K^+ \nu \bar{\nu}$.

de la résonance $\Upsilon(4S)$. Belle II est la mise à niveau de l'expérience Belle qui fut responsable de plusieurs découvertes majeures comme celle de la violation de CP dans les mésons B ou la première observation des désintégrations $b \rightarrow sll$. Belle II a plusieurs atouts majeurs. Avant tout, les collisions électron-positron produisent des événements avec beaucoup moins de particules finales qu'au LHC (environ 11 contre plusieurs dizaines, centaines ou milliers selon la luminosité instantanée) et l'énergie de la collision est connue. De plus, le détecteur est hermétique ce qui permet de reconstruire la totalité d'un événement. Il est possible en outre de discriminer parmi les différents types de hadrons. Par ailleurs la résolution sur la position des vertex de désintégration et sur l'impulsion et l'énergie des particules de l'état finale est excellente. Enfin le collisionneur SuperKEKB détient le record mondial de la plus haute luminosité instantanée, $5.32 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, et dans les prochaines années, il devrait atteindre une luminosité instantanée encore plus grande. Cela va permettre à Belle II de récolter, pendant la prochaine décennie, un très large échantillon de données qui va permettre des mesures de haute précision. Cet échantillon correspondra à plusieurs dizaines de ab^{-1} , soit plusieurs dizaines de fois celui collecté par Belle. Durant le run 1 de 2019 au 2023 un échantillon d'environ 427 fb^{-1} a été collecté : 365 fb^{-1} au centre de masse du $\Upsilon(4S)$, 42 fb^{-1} avant le seuil de production de deux mésons B, pour offrir un échantillon de contrôle pour le bruit de fond continuum¹ et 20 fb^{-1} pour un scan en énergie au-dessus du $\Upsilon(4S)$. Le run 2 qui a débuté à l'automne 2024 est en cours. Au moment de la soumission de ce document 914 fb^{-1} ont été collectées en tout dont 808 fb^{-1} au centre de masse du $\Upsilon(4S)$.

1.2 LES TRANSITIONS $b \rightarrow s \nu \bar{\nu}$

Les modes de désintégration $B \rightarrow K^{(*)} \nu \bar{\nu}$ sont parmi les plus prometteurs pour trouver des signes de la nouvelle physique nécessaire pour compléter le Modèle Standard.

Prédiction théorique

Dans le Modèle Standard, ces désintégrations procèdent via des courants

1. Ce bruit de fond est dû à $e^+e^- \rightarrow q\bar{q}$ où q est un quark u, d, s ou c et à $e^+e^- \rightarrow \tau^+\tau^-$.

neutres qui changent de saveur et les diagrammes d'ordres inférieurs contiennent des boucles quantiques comme montré en figure 1.1. Ces transitions peuvent être décrites par une théorie effective où les champs plus lourds que le quark b sont intégrés et tous les opérateurs formés par les champs restants et qui respectent les symétries du Modèle Standard sont considérés :

$$\mathcal{L}_{\text{eff}}^{b \rightarrow s \nu \bar{\nu}} = \frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \sum_a \mathcal{C}_a \mathcal{O}_a + \text{h.c.}$$

où G_F est la constante de Fermi et V_{ij} sont les éléments de la matrice CKM. Seulement deux opérateurs sont présents à dimension 6 pour chaque couple de neutrinos :

$$\begin{aligned} \mathcal{O}_L &\propto (\bar{s}_L \gamma_\mu b_L)(\bar{\nu} \gamma^\mu (1 - \gamma_5) \nu), \\ \mathcal{O}_R &\propto (\bar{s}_R \gamma_\mu b_R)(\bar{\nu} \gamma^\mu (1 - \gamma_5) \nu). \end{aligned}$$

Les rapports d'embranchements sont alors [6, 7] :

$$\begin{aligned} \mathcal{B}(B^+ \rightarrow K^+ \nu \bar{\nu}) &\propto |f_+^K(q^2)|^2 |C_L + C_R|^2, \\ \mathcal{B}(B^+ \rightarrow K^{*+} \nu \bar{\nu}) &\propto |V(q^2)|^2 |C_L + C_R|^2 + (|A_1(q^2)|^2 + |A_{12}(q^2)|^2) |C_L - C_R|^2, \end{aligned}$$

où f_+^K , V , A_1 et A_{12} sont des facteurs de forme. Enfin

$$\mathcal{B}(B^0 \rightarrow K^{(*)0} \nu \bar{\nu}) = \frac{\tau_{B^+}}{\tau_{B^0}} \mathcal{B}(B^+ \rightarrow K^{(*)+} \nu \bar{\nu}) \quad (1.1)$$

où τ_{B^+}/τ_{B^0} est le rapport entre les temps de vie des mésons B^+ et B^0 . Il est clair de ces équations que les désintégrations avec le méson scalaire K et celles avec le méson vecteur K^* fournissent des contraintes complémentaires sur C_L et C_R . Cela est évident en figure 1.2, où la relation entre $B \rightarrow K \nu \bar{\nu}$ et $B \rightarrow K^* \nu \bar{\nu}$ est montrée en faisant varier δC_L ou δC_R séparément². Dans le Modèle Standard $C_L = -6,32 \pm 0,07$ et $C_R = 0$. La prédiction du Modèle Standard pour les rapports d'embranchement est alors [8, 9] :

$$\begin{aligned} \mathcal{B}(B^+ \rightarrow K^+ \nu \bar{\nu}) &= (4,44 \pm 0,30) \times 10^{-6}, \\ \mathcal{B}(B^+ \rightarrow K^{*+} \nu \bar{\nu}) &= (9,8 \pm 1,4) \times 10^{-6}. \end{aligned}$$

Pour $B^+ \rightarrow K^{(*)+} \nu \bar{\nu}$, il y a aussi une contribution non perturbative due à la désintégration $B^+ \rightarrow \tau^+ (\rightarrow K^{(*)+} \nu) \nu$ montrée en figure 1.3 et qui corres-

Contribution non perturbative

2. Pour δC_L ou δC_R on entend la partie du coefficient de Wilson due à de la nouvelle physique.

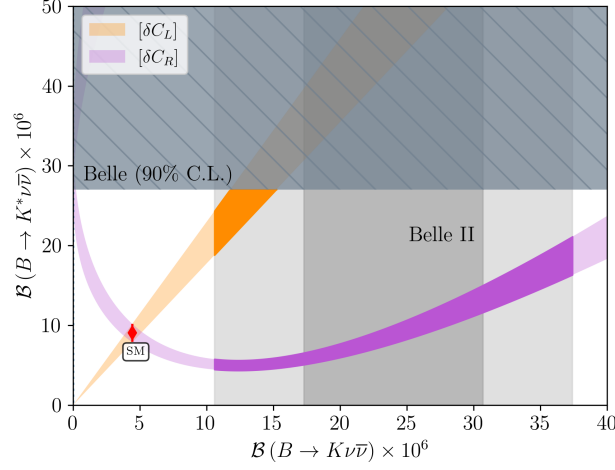


FIGURE 1.2 – Corrélation entre les rapports d’embranchement de $B \rightarrow K\nu\bar{\nu}$ et $B \rightarrow K^*\nu\bar{\nu}$ en fonction de la variation de δC_L ou δC_R . La zone grisée correspond aux intervalles 1σ et 2σ de réf. [1]. Le point rouge correspond aux prédictions du Modèle Standard pour ces observables. Et la région hachurée grise est exclue à 90% CL par réf. [10]. Reproduit de réf. [9].

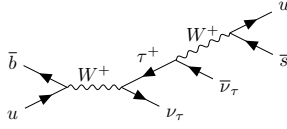


FIGURE 1.3 – Diagramme de Feynman de la contribution non perturbative aux désintégrations $B^+ \rightarrow K^{(*)+}\nu\bar{\nu}$.

pond à un peu plus que 10% de la partie perturbative et qui, par la suite, est considérée comme un bruit de fond irréductible [11] :

$$\mathcal{B}(B^+ \rightarrow K^+\nu\bar{\nu})_{\text{NP}} = (6,28 \pm 0,06) \times 10^{-7},$$

$$\mathcal{B}(B^+ \rightarrow K^{*+}\nu\bar{\nu})_{\text{NP}} = (1,07 \pm 0,10) \times 10^{-6}.$$

*Nouvelles particules
invisibles*

La recherche des transitions $b \rightarrow s\nu\bar{\nu}$ n’est pas sensible uniquement à des nouvelles particules lourdes qui pourraient entrer dans les boucles quantiques et modifier la valeur de C_L et C_R par rapport au modèle standard. Vu que les neutrinos ne sont pas directement reconstruits dans les expériences, cette recherche est sensible aussi à des possibles autres particules légères non détectées, par exemple des candidats de matière noire [12, 13]. Dans ce cas, $B \rightarrow K\nu\bar{\nu}$ et $B \rightarrow K^*\nu\bar{\nu}$ contraignent aussi différents modèles de nouvelle physique comme montrée en figure 1.4.

Les désintégrations $B \rightarrow K^{(*)}\nu\bar{\nu}$ n’ont pas encore été observées, mais si leurs taux de désintégration étaient proches de ceux prévus par le Modèle Standard, une première observation avec une significativité à cinq écarts type devrait être possible prochainement avec environ 5 ab^{-1} de données collectées par Belle II. Cette dernière est la seule expérience au monde per-

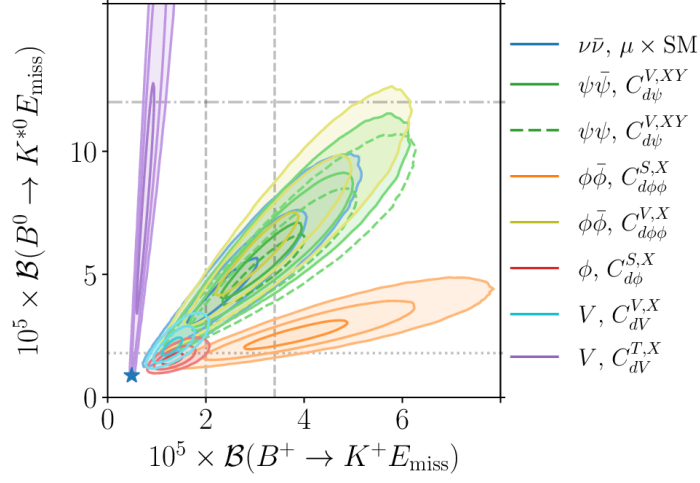


FIGURE 1.4 – Corrélation entre les rapports d'embranchement de $B \rightarrow K E_{\text{miss}}$ et $B \rightarrow K^* E_{\text{miss}}$ dans le Modèle Standard et des modèles de nouvelle physique où une ou deux particules légères invisibles sont ajoutés. Le spin de ϕ , ψ , et V est respectivement 0, 1/2 et 1. Reproduit de réf. [13].

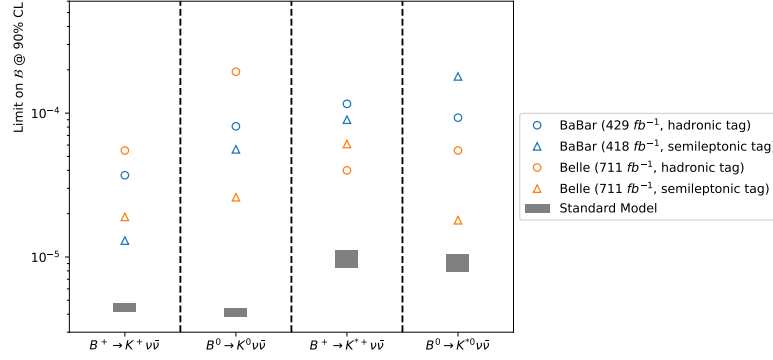


FIGURE 1.5 – Prédictions théoriques [9] et limites supérieures des rapports d'embranchement des désintégrations $B \rightarrow K^{(*)} \nu \bar{\nu}$ avant Belle II [10, 14-17].

mettant d'observer ce type de désintégrations. La figure 1.5 montre les limites supérieures obtenues par les expériences précédentes [10, 14-17].

Le défi principal de la recherche de la désintégration $B^+ \rightarrow K^+ \nu \bar{\nu}$ est la présence de deux neutrinos dans l'état final. Ces particules ne sont pas détectées directement, mais leur présence et leurs propriétés doivent être déduites du reste de l'évènement, c'est-à-dire du kaon et du B_{tag} , l'autre méson B de l'évènement. Il y a deux stratégies pour reconstruire le B_{tag} :

- l'approche tagged où le B_{tag} est reconstruit complètement via ses désintégrations hadroniques (HTA) ou semi-leptoniques (STA) ;
- l'approche inclusive (ITA) où le signal est reconstruit en premier et on assume que le reste de l'évènement constitue le B_{tag} .

L'approche tagged présuppose de reconstruire entièrement le B_{tag} qui peut se désintégrer de milliers de façons différentes. L'algorithme actuellement utilisé pour cela s'appelle Full Event Interpretation (FEI) [18]. Il s'agit d'une approche hiérarchique où une série de classificateurs peu profonds combinent les particules de l'état final pour former des particules intermédiaires et finalement le méson B. Il y a deux types de tag possibles selon les modes de désintégrations considérés pour la désintégration du méson B. Avec le HTA le méson B se désintègre dans un mode hadronique où toutes les particules de l'état final sont visibles en permettant de contraindre parfaitement la cinématique du couple de neutrinos. Ce tag a toutefois une efficacité inférieure au pourcent. Avec le tag STA, le méson B se désintègre en $D^{(*)} l \nu$. Ce tag permet une efficacité de quelques pourcents, mais la présence d'un neutrino aussi dans le tag ne permet pas de contraindre parfaitement la cinématique de l'évènement. Dans ces deux approches, on reconstruit d'abord le B_{tag} et, ensuite, on vérifie que le reste de l'évènement contient seulement le $K^{(*)}$ provenant du signal. Dans l'ITA au contraire, on reconstruit d'abord $K^{(*)}$ et on assume que le reste de l'évènement forme le B_{tag} ; cette approche permet donc une efficacité majeure au prix d'une pureté inférieure.

Au début de la prise de données de Belle II, j'ai développé et optimisé avec l'étudiant de M2, Léonard Polat, la sélection des désintégrations $B \rightarrow K^{(*)} \nu \bar{\nu}$ sur des données simulées en utilisant un tag hadronique [19]. J'ai été ensuite un des analystes principaux de réf. [1] avec le doctorant Lucas Martel et le postdoc Jacopo Cerasoli que je supervisais. Cette publication est synthétisée dans le reste du chapitre.

2.1 SÉLECTION

Sélection ITA

Dans réf. [1] la désintégration $B^+ \rightarrow K^+ \nu \bar{\nu}$ est recherchée avec deux des approches décrites avant : ITA et HTA en utilisant l'échantillon complet de run 1. Dans ITA, des sélections basiques pour rejeter les contributions provenant du bruit de fond des faisceaux sont appliqués aux traces utilisées pour reconstruire les particules chargées et aux clusters du calorimètre électromagnétique. Le vecteur de l'impulsion manquante doit être dans l'acceptance du détecteur et l'évènement doit contenir au moins quatre traces pour réduire la contribution de fonds électromagnétiques. Le signal est identifié avec la trace qui minimise q_{rec}^2 entre celles définies comme étant un kaon grâce à une coupure serrée sur la variable d'identification de particules (PID).

$$q_{\text{rec}}^2 = s/(4c^4) + m_K^2 - \sqrt{s}E_K^*/c^4 \quad (2.1)$$

où m_K est la masse du kaon, $\sqrt{s}$ est l'énergie du centre de masse et E_K^* est l'énergie du kaon dans le centre de masse du $\Upsilon(4S)$. Cette variable est une approximation de q^2 , la masse au carré du couple du neutrino, sous l'hypothèse que le méson B soit à repos dans le centre de masse du $\Upsilon(4S)$ ¹. Cette sélection permet de choisir le bon candidat avec une efficacité de 96% en cas de candidats multiple et ne cause pas un biais de sélection. Les traces et clusters restants sont identifiés comme provenant de l'autre méson B de l'évènement.

Pour réduire ultérieurement le bruit de fond, deux arbres de décision consécutifs sont utilisés. Le premier utilise 12 variables, dont la plus importante est ΔE , la différence entre l'énergie du B_{tag} et $\sqrt{s}/2$. Le deuxième arbre de décision est entraîné sur les candidats où la sortie du premier arbre de décisions est plus grand que 0,9² et utilise 35 variables, dont les plus importantes sont les variables sur la forme de l'évènement et celles obtenues associant le kaon du signal avec les pions de l'évènement pour rejeter les mésons D. La région de signal est définie par une coupure sur la sortie de ce deuxième arbre de décision. L'efficacité totale de sélection est de 8% avec une pureté de 0,8%.

Sélection HTA

Dans HTA, le B_{tag} est reconstruit en premier en utilisant le FEI hadronique. Le signal est une trace identifiée comme un kaon, par une coupure serrée sur la variable d'identification de particules. L'évènement est rejeté s'il contient des candidats K_S^0 , π^0 ou Λ , ou des traces proches du point de collision et qui laissent au moins 20 hits dans le trajectographe gazeux. Un arbre de décision qui utilise 12 variables est ensuite utilisé pour réduire le bruit de fond. Les variables principales sont l'énergie extra (énergie dans le calorimètre non associé au B_{tag} ou au signal), la somme de l'énergie et

1. Sous cette hypothèse $m_B c^2 = E_B^* = \sqrt{s}/2$ et $p_B^* = 0$ quand en réalité $p_B^* \sim 0.3$ GeV.

2. Cela correspond à une efficacité de 35% en gardant 1.5% du fond.

l'impulsion manquante calculée en utilisant l'information de l'énergie de collision et du B_{tag} et le nombre de traces extra (traces non associé au B_{tag} ou au signal qui ne passent pas les critères pour faire rejeter l'évènement.).

2.2 CORRECTIONS DE LA SIMULATION, INCERTITUDES SYSTÉMATIQUES ET AJUSTEMENT

Le rapport d'embranchement du signal est obtenu par un ajustement par modèle de la sortie du BDT pour l'HTA et de la sortie du BDT et de la masse invariante du couple de neutrinos pour l'ITA. Les histogrammes avec la forme du signal et des différentes composantes du bruit de fond sont obtenus avec des échantillons simulés. Les incertitudes systématiques sont directement prises en compte dans l'ajustement en utilisant des paramètres auxiliaires qui sont contraints avec des gaussiennes. Ces systématiques peuvent changer les histogrammes de trois manières différentes : en changer la normalisation, en changer la forme, mais avec une corrélation établie entre les bins ou varier les contenus des bins indépendamment. Les incertitudes systématiques peuvent intéresser une seule composante de l'ajustement ou être corrélées entre plusieurs composantes. Dans l'ajustement de ITA huit composantes sont considérés (signal, B^+B^- , $B^0\bar{B}^0$, $c\bar{c}$, $s\bar{s}$, $d\bar{d}$, $u\bar{u}$ et $\tau^+\tau^-$) tandis que l'ajustement de HTA considère quatre composantes (signal, B^+B^- et $B^0\bar{B}^0$ ensemble, $s\bar{s}$, $d\bar{d}$, $u\bar{u}$ et $\tau^+\tau^-$ ensemble et $c\bar{c}$).

Ajustement

La simulation est centrale dans cette analyse : elle est utilisée pour estimer l'efficacité du signal et les templates des différentes composantes du fond utilisées dans l'ajustement. Il est donc important d'avoir une simulation la plus fidèle possible à la réalité et de couvrir les possibles différences entre simulation et données par des incertitudes systématiques adaptées. Plusieurs échantillons de contrôles sont utilisés pour estimer les corrections à apporter à la simulation et les incertitudes systématiques résiduelles. Les échantillons principaux sont :

Échantillons de contrôles

offres Un échantillon sous-résonance, collecté à environ 60 MeV en dessous de la masse du $\Upsilon(4S)$ permet de contraindre le bruit de fond continuum qui provient de $u\bar{u}$, $d\bar{d}$, $s\bar{s}$, $c\bar{c}$ et $\tau^+\tau^-$.

pion-enriched Des échantillons dans lesquels, au lieu de sélectionner un kaon, un pion est sélectionné.

wrong-charge Pour HTA un échantillon dans lequel le B_{tag} et le kaon de signal ont la même charge.

embedding Pour évaluer la correction sur l'efficacité du signal des événements contenant $B^+ \rightarrow J/\psi K^+$ sont reconstruits dans les données et

la simulation. La désintégration $B^+ \rightarrow J/\psi K^+$ est ensuite remplacé par $B^+ \rightarrow K^+ \nu \bar{\nu}$ simulé.

Corrections et
incertitudes
systématiques

Plusieurs corrections de la simulation sont donc apportées et pour chacune d'elle une incertitude systématique est associée. La liste des incertitudes systématiques et des corrections est reportée en tableau 2.1 pour ITA et tableau 2.2 pour HTA. Les corrections et incertitudes systématiques principales sont les suivantes.

- Dans ITA l'énergie des faux photons³ est augmentée de 10%. Cette valeur est obtenue en utilisant un échantillon de contrôle de $B^+ \rightarrow J/\psi K^+$ comme montré en figure 2.1. Une incertitude systématique équivalente à 100% de la correction est appliquée. Dans HTA une autre sélection des photons est utilisée et l'énergie des photons extra⁴ présente un bon accord entre données et simulation. Leur nombre toutefois présente un désaccord et est corrigé à partir de l'échantillon de contrôle **wrong-charge**. Cette correction est validée dans l'échantillon de contrôle **pion-enriched** comme montré en figure 2.2. Le désaccord résiduel dans cet échantillon est assigné comme incertitude systématique.
- La distribution en q^2 du signal est corrigée pour tenir en compte de la prédiction théorique la plus récente et l'incertitude sur les facteurs de forme est propagée à l'ajustement.
- L'efficacité de la sélection et reconstruction du signal est corrigée grâce aux deux échantillons de **embedding**. Le rapport entre le nombre d'événements dans l'échantillon obtenu à partir des données et celui à partir de la simulation permet d'estimer le rapport d'efficacité du signal entre données et simulation dû au B_{tag} . Ce rapport est utilisé comme facteur de correction et son incertitude statistique est utilisée comme incertitude systématique. L'efficacité liée au signal dépend de l'efficacité de reconstruction des traces et de la sélection du kaon basée sur la PID. Cette dernière est corrigée à l'aide des facteurs de correction obtenus à partir d'un échantillon de $D^{*+} \rightarrow \pi^+ D^0 (\rightarrow K^- \pi^+)$ et l'incertitude correspondante est utilisée dans l'ajustement. L'efficacité de reconstruction des traces est évaluée dans un échantillon de $e^+ e^- \rightarrow \tau^+ (\rightarrow \pi \nu) \tau^- (\rightarrow 3 \pi \nu)$. Elle est compatible avec un et l'incertitude de 0.3% est propagée dans l'ajustement.
- L'accord entre données et simulation est vérifié pour le continuum à l'aide de l'échantillon **offres**. Une BDT est entraînée pour séparer les données de la simulation et un poids de $p/(1-p)$, ou p est la sortie de la BDT, est appliqué aux événements simulés. Pour couvrir le possible désaccord résiduel, une incertitude systématique est associée où la correction peut varier entre 0 et 200%.

3. Candidats photons qui ne correspondent pas à des photons simulés mais qui proviennent erronément des gerbes électromagnétiques associées aux traces.

4. Les photons non associés au B_{tag} ou au signal

FIGURE 2.1 – Distribution de la somme de l'énergie des photons obtenus dans les données (points avec barres d'erreur), simulation non corrigée (histogramme ouvert) et simulation corrigée. Reproduit de réf. [1].

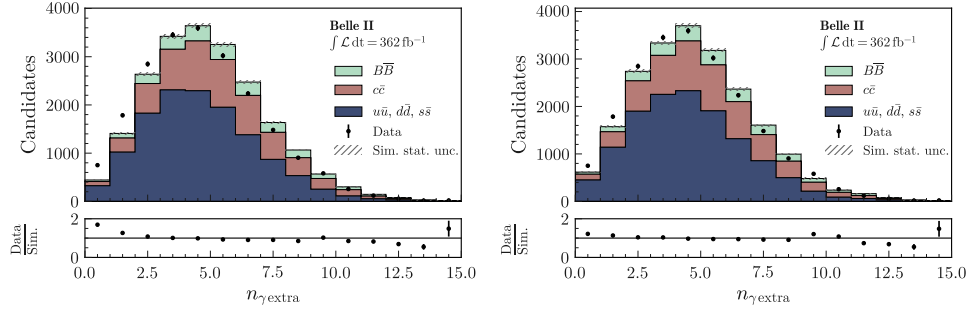
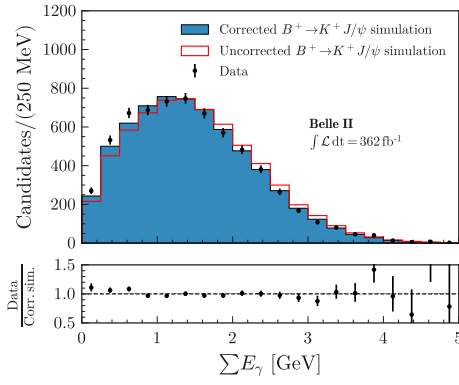


FIGURE 2.2 – Distribution du nombre de photons dans le calorimètre non associés au B_{tag} ou au signal dans le canal de contrôle (*gauche*) avant et (*droite*) après la correction. Reproduit de réf. [1].

- La normalisation des différentes composantes du bruit de fond est laissée libre de varier de 50% dans ITA car un désaccord de $(40 \pm 5)\%$ dans la normalisation du continuum entre simulation et échantillon **offres** est observé. Dans HTA, après l'application des autres corrections, un bon accord dans la normalisation du bruit de fond est observé dans les échantillons de contrôle **offres**, **wrong-charge** et **pion-enriched**. Les incertitudes statistiques de ces contrôles sont assignées comme incertitudes systématiques sur les normalisations des différents fonds : 50% pour le continuum et 30% pour le $B\bar{B}$.
- Les incertitudes sur les rapports d'embranchements des désintégrations principales du méson B présentes dans le bruit de fond sont propagées dans l'ajustement.
- La distribution cinématique des désintégrations $B^+ \rightarrow K^+ K_L^0 K_L^0$ et $B^+ \rightarrow K^+ n \bar{n}$ est corrigée dans la simulation en se basant sur $B^+ \rightarrow K^+ K_S^0 K_S^0$ et $B^+ \rightarrow K^+ p \bar{p}$.
- La simulation des clusters formés par des K_L^0 dans le calorimètre est validée grâce à un échantillon de $\phi \rightarrow K_S^0 K_L^0$. La sélection des clusters du calorimètre est différente entre ITA et HTA. Le nombre d'événements de bruit de fond contenant un K_L^0 est ainsi augmenté de 17% avec une incertitude de 8% dans ITA tandis que seulement une incertitude de 17% est assignée dans HTA.

TABLE 2.1 – Sources d’incertitude systématique dans l’ITA, facteurs de correction correspondants (le cas échéant), leur traitement dans l’ajustement, leur amplitude et leur impact sur l’incertitude de la force du signal μ . Le type d’incertitude peut être “Global”, correspondant à un facteur de normalisation global commun à toutes les catégories de régions de signal (SR), ou “Shape”, correspondant à une incertitude dépendant de la catégorie. Chaque source est décrite par un ou plusieurs paramètres de nuisance. L’impact sur l’incertitude de la force du signal σ_μ est estimé en excluant la source de la minimisation et en soustrayant, en quadrature, l’incertitude résultante de l’incertitude de l’ajustement nominal. Reproduit de réf. [1].

Source	Correction	Uncertainty type, parameters	Uncertainty size	Impact on σ_μ
Normalization of $B\bar{B}$ background		Global, 2	50%	0.90
Normalization of continuum background		Global, 5	50%	0.10
Leading B-decay branching fractions		Shape, 6	O(1%)	0.22
Branching fraction for $B^+ \rightarrow K^+ K_L^0 K_L^0$	q^2 dependent O(100%)	Shape, 1	20%	0.49
p-wave component for $B^+ \rightarrow K^+ K_S^0 K_L^0$	q^2 dependent O(100%)	Shape, 1	30%	0.02
Branching fraction for $B \rightarrow D^{**}$		Shape, 1	50%	0.42
Branching fraction for $B^+ \rightarrow K^+ n \bar{n}$	q^2 dependent O(100%)	Shape, 1	100%	0.20
Branching fraction for $D \rightarrow K_L^0 X$	+30%	Shape, 1	10%	0.14
Continuum-background modeling, BDT _c	Multivariate O(10%)	Shape, 1	100% of correction	0.01
Integrated luminosity		Global, 1	1%	< 0.01
Number of $B\bar{B}$		Global, 1	1.5%	0.02
Off-resonance sample normalization		Global, 1	5%	0.05
Track-finding efficiency		Shape, 1	0.3%	0.20
Signal-kaon PID	p, ϑ dependent O(10–100%)	Shape, 7	O(1%)	0.07
Photon energy		Shape, 1	0.5%	0.08
Hadronic energy	−10%	Shape, 1	10%	0.37
K_L^0 efficiency in ECL	−17%	Shape, 1	8.5%	0.22
Signal SM form factors	q^2 dependent O(1%)	Shape, 3	O(1%)	0.02
Global signal efficiency		Global, 1	3%	0.03
Simulated-sample size		Shape, 156	O(1%)	0.52

TABLE 2.2 – Sources d’incertitude systématique dans l’HTA (voir la légende de Tableau 2.1 pour plus de détails). Reproduit de réf. [1].

Source	Correction	Uncertainty type, parameters	Uncertainty size	Impact on σ_μ
Normalization of $B\bar{B}$ background		Global, 1	30%	0.91
Normalization of continuum background		Global, 2	50%	0.58
Leading B-decay branching fractions		Shape, 3	O(1%)	0.10
Branching fraction for $B^+ \rightarrow K^+ K_L^0 K_L^0$	q^2 dependent O(100%)	Shape, 1	20%	0.20
Branching fraction for $B \rightarrow D^{**}$		Shape, 1	50%	< 0.01
Branching fraction for $B^+ \rightarrow K^+ n \bar{n}$	q^2 dependent O(100%)	Shape, 1	100%	0.05
Branching fraction for $D \rightarrow K_L^0 X$	+30%	Shape, 1	10%	0.03
Continuum-background modeling, BDT _c	Multivariate O(10%)	Shape, 1	100% of correction	0.29
Number of $B\bar{B}$		Global, 1	1.5%	0.07
Track finding efficiency		Global, 1	0.3%	0.01
Signal-kaon PID	p, ϑ dependent O(10–100%)	Shape, 3	O(1%)	< 0.01
Extra-photon multiplicity	$n_{\gamma\text{extra}}$ dependent O(20%)	Shape, 1	O(20%)	0.61
K_L^0 efficiency		Shape, 1	17%	0.31
Signal SM form factors	q^2 dependent O(1%)	Shape, 3	O(1%)	0.06
Signal efficiency		Shape, 6	16%	0.42
Simulated-sample size		Shape, 18	O(1%)	0.60

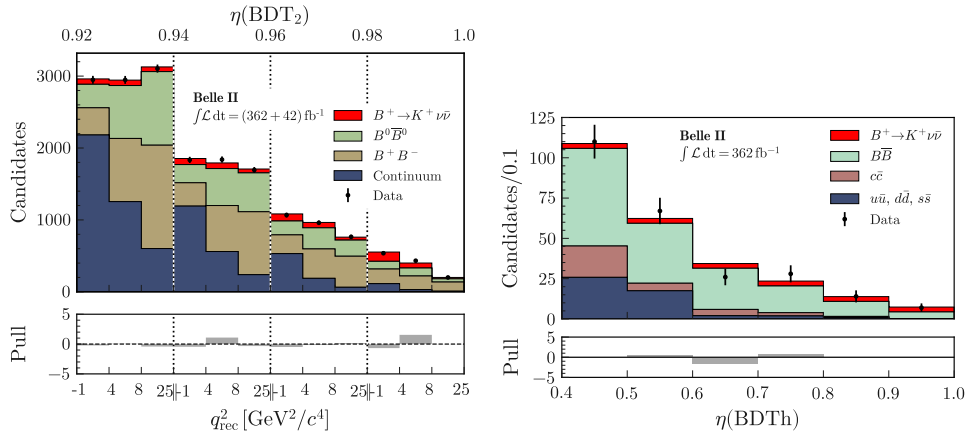


FIGURE 2.3 – Candidats observés et résultat de l’ajustement pour (gauche) ITA en bins de la sortie du BDT et de la masse invariante du couple de neutrinos (droite) HTA en bins de la sortie du BDT. Reproduit de réf. [1].

2.3 RÉSULTATS

La fraction d’embranchement obtenue est $[2,7 \pm 0,5(\text{stat}) \pm 0,5(\text{syst})] \times 10^{-5}$ pour l’ITA et $[1,1^{+0,9}_{-0,8}(\text{stat})^{+0,8}_{-0,5}(\text{syst})] \times 10^{-5}$ pour l’HTA. En combinant les deux approches, on obtient $[2,3 \pm 0,5(\text{stat})^{+0,5}_{-0,4}(\text{syst})] \times 10^{-5}$. Cela constitue la première évidence pour cette désintégration à 3,5 écarts type et présente une tension de 2,7 écarts type avec la prédiction du Modèle Standard. La figure 2.3 montre le résultat de l’ajustement et la figure 2.5 gauche montre le rapport de log-vraisemblance du profil négatif. la figure 2.4 montre la distribution en q^2 après l’ajustement. La figure 2.5 droite compare le résultat obtenu avec ceux des expériences passées et la prédiction du modèle standard. Tableau 2.3 compare les résolutions obtenues par les différentes analyses ainsi que ces résolutions reportées à la même luminosité⁵. Il est intéressant de noter que l’approche HTA à Belle II a permis une meilleure précision que la même stratégie utilisée dans les expériences précédentes malgré le plus petit échantillon de données utilisé. De plus, le résultat de ITA est à ce jour le plus précis si reporté à la luminosité utilisée.

Modèles de nouvelle
physique

Cette mesure en tension avec le Modèle Standard a suscité beaucoup d’intérêts dans la communauté et nombreux modèles de nouvelle physique ont été proposés pour l’expliquer. Figure 2.6 montre un schéma de plusieurs d’entre eux regroupés par catégories. La nouvelle physique peut soit être à haute énergie et modifier directement le rapport d’embranchement de $B^+ \rightarrow K^+ \nu \bar{\nu}$ (paramétrée en termes de coefficients de Wilson ou décrite explicitement dans un modèle complet), soit être à basse énergie et contenir des nouvelles particules invisibles qui remplacent le couple de neutrinos.

5. Cela est une simplification qui suppose que les incertitudes statistiques et systématiques évoluent avec la racine de la luminosité.

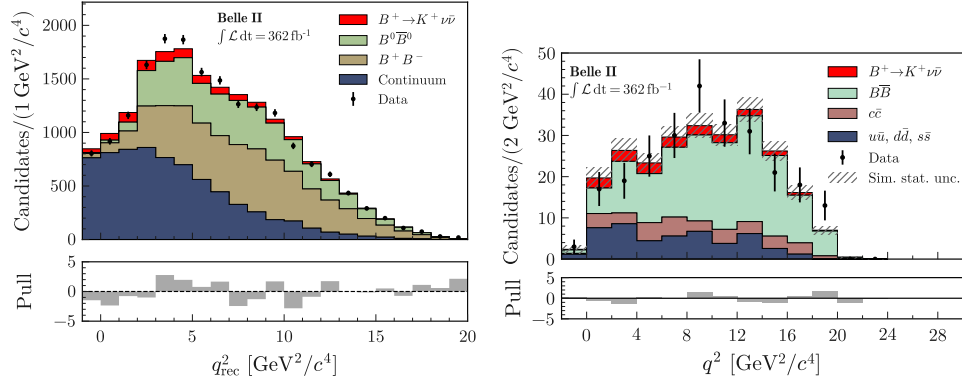


FIGURE 2.4 – Distribution en q^2 des candidats dans la région du signal pour (gauche) ITA en utilisant eq. 2.1 et (droite) HTA en utilisant l'information du B_{tag} . Reproduit de réf. [1].

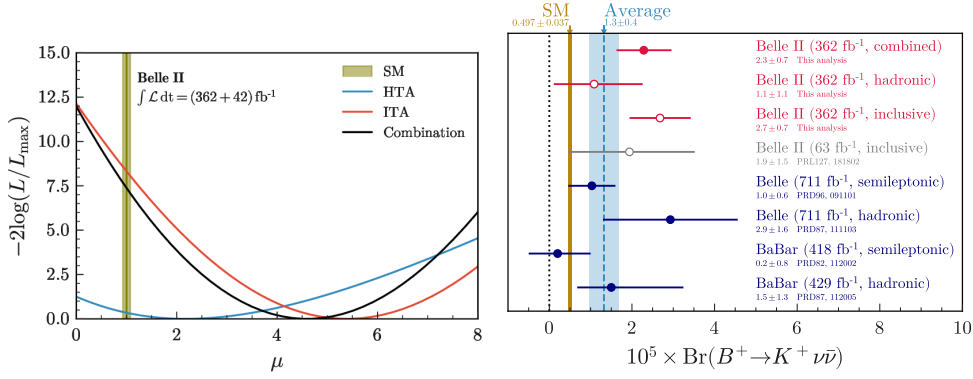
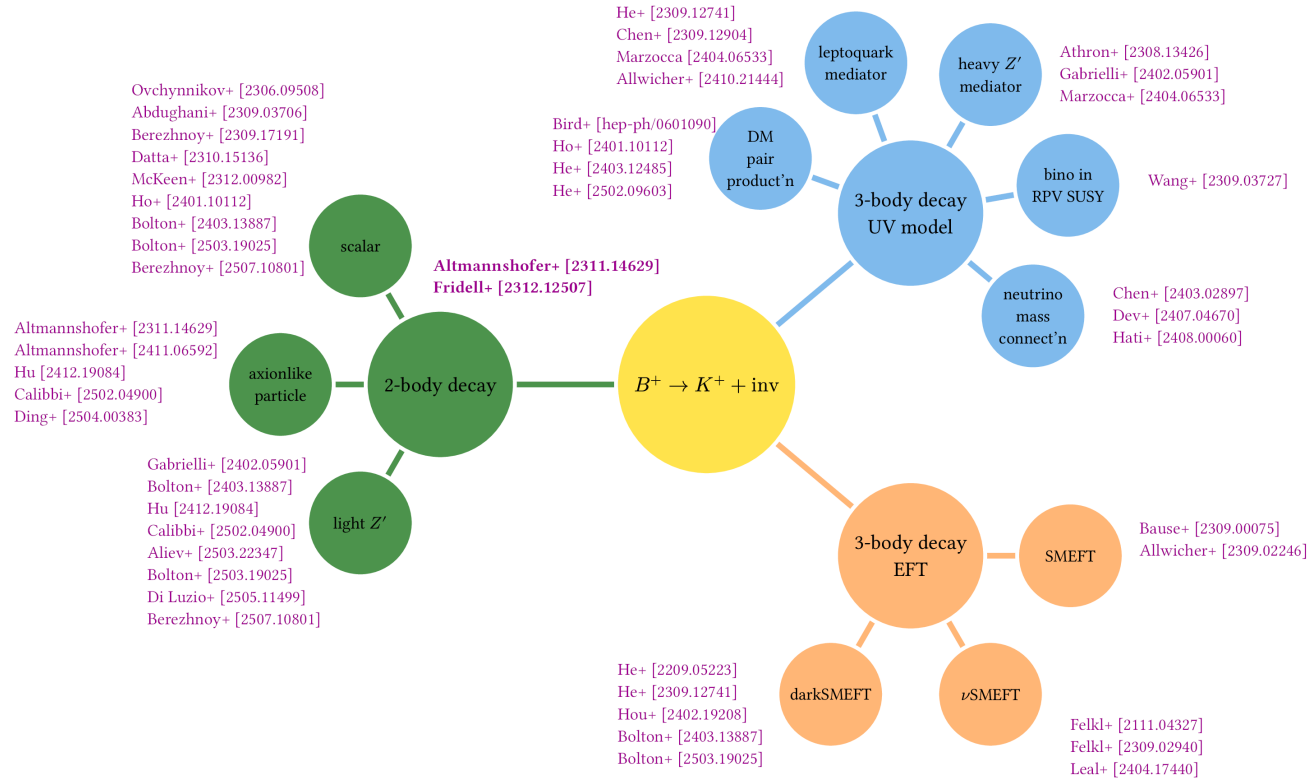


FIGURE 2.5 – (Gauche) deux fois le rapport de log-vraisemblance du profil négatif en fonction de l'intensité du signal μ de la désintégration $B^+ \rightarrow K^+ \nu \bar{\nu}$ pour ITA, HTA et le résultat combiné. (Droite) Valeur de la fraction d'embranchement de $B^+ \rightarrow K^+ \nu \bar{\nu}$ mesurée par différentes expériences et comparés à la prédiction théorique. Reproduit de réf. [1].

TABLE 2.3 – Comparaison des incertitudes totales des différentes recherches de $B^+ \rightarrow K^+ \nu \bar{\nu}$ et incertitudes rescalées naïvement par la luminosité à 365 fb^{-1}

Analyse $B^+ \rightarrow K^+ \nu \bar{\nu}$	Incrtitude [10^{-6}]	Incrtitude pour 365 fb^{-1} [10^{-6}]
BaBar HTA [16]	12.9	14.0
BaBar STA [15]	7.5	8.0
Belle HTA [17]	16.8	23.5
Belle STA [10]	5.8	8.1
Belle II HTA [1]	10.7	10.7
Belle II ITA [1]	7.1	7.1

Proposed new physics explanations



Michael Schmidt - BSM for $B \rightarrow K + \text{invisible}$

2

FIGURE 2.6 – Schéma des modèles de nouvelle physique proposés jusqu'en octobre 2025 pour expliquer la tension avec le Modèle Standard de la mesure de Belle II. Reproduit de réf. [20].

La reconstruction de la désintégration d'un méson B avec des particules non détectées dans l'état final nécessite de l'information sur tout l'évènement. À Belle II, où les mésons B sont produits en paire dans la désintégration du méson $\Upsilon(4S)$ sans d'autres particules, cela se traduit par la reconstruction du B_{tag} , l'autre B de l'évènement. La reconstruction d'une désintégration générique d'un méson B est un problème complexe, car il peut se désintégrer dans des milliers de modes possibles.

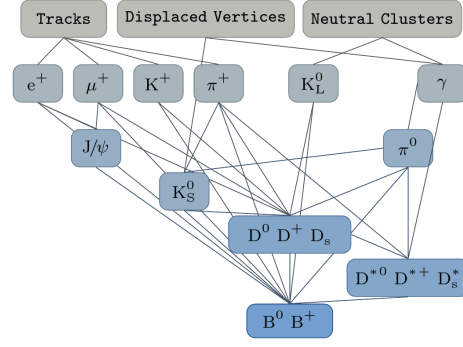
Actuellement dans Belle II l'algorithme *Full Event Interpretation* (FEI) [18] *Le FEI* est utilisé pour cette tâche. Il est basé sur six niveaux de BDTs qui reconstruisent le méson B à partir des particules de l'état final comme schématisé en figure 3.1. Cet algorithme reconstruit environ 10 000 chaînes complètes de désintégration et permet une efficacité de quelques pourcents. La plus grande partie de son efficacité est toutefois due aux modes qui procèdent via une transition $b \rightarrow cW$ semi-leptoniques $B \rightarrow D^{(*)}l\nu$ ou hadroniques $B \rightarrow D^{(*)}n\pi$.

Les désavantages de cette méthode sont que les six couches sont déconnectées et que les désintégrations possibles sont établies à priori. De cette manière 43% de désintégrations du méson B sont couvertes par le FEI, mais, si on considère toute la chaîne de désintégration jusqu'aux particules de l'état final, seulement 15% de chaînes de désintégrations sont couvertes.

L'apprentissage profond a révolutionné plusieurs domaines dans la dernière décennie grâce à sa capacité à définir de manière autonome les variables de haut niveau les mieux adaptées au problème. Par exemple, pour la reconnaissance d'image des réseaux profonds qui travaillent avec des informations de bas-niveau, ont pris la place des réseaux peu profonds qui travaillaient avec des données de haut niveaux prédéfinies par des experts. D'autres problèmes complexes, caractérisés par une représentation hiérarchique des données comme le traitement du langage naturel et les recherches Web ont aussi vu une évolution similaire. Il me semblait donc naturel d'essayer d'améliorer la reconstruction générique d'un méson B, autre problème complexe caractérisé par une représentation hiérarchique des données, à l'aide d'un réseau de neurones profond. Ce réseau aurait reconstruit une désintégration générique de méson B sans devoir explicitement définir toutes les voies possibles de désintégration. J'ai poursuivi ces recherches d'abord dans le cadre d'un projet Seed Money du programme of Eucor The European Campus et ensuite grâce à un projet ANR JCJC. J'ai supervisé sur ce sujet

Apprentissage profond

FIGURE 3.1 – Aperçu schématique du FEI. L’algorithme opère sur des objets identifiés par la reconstruction de Belle II : traces des particules chargées, clusters du calorimètre et vertex déplacés. En six distincts étapes, ces objets de base sont interprétés comme un état final particules (e^+ , μ^+ , K^+ , π^+ , K_L^0 , γ) combinés pour former des particules intermédiaires (J/ψ , π^0 , K_S^0 , D , D^*) et enfin former le méson B. Reproduit de réf. [18].



le travail du postdoc Jacopo Cerasoli et des étudiants de M2 Ilias Tsaklidis, Arthur Taller, Corentin Santos et Merna Abumusabh [21-23].

Matrice LCAG

Avec Ilias, Arthur et des collègues physiciens et informaticiens de Karlsruhe et de Bonn, nous avons jeté les bases théoriques en définissant la *matrice de la Génération du Plus Bas Ancêtre Commun* qui permet de représenter la structure hiérarchique d’une désintégration directement à partir des particules de l’état final sans hypothèses préalables sur le nombre de particules intermédiaires [2]. Une désintégration peut, en effet, être représentée par un arbre hiérarchique comme montré en figure 3.2 gauche. Toutefois, cela nécessite de connaître déjà les particules intermédiaires impliquées dans la désintégration et ne peut donc pas être utilisé dans l’entraînement d’un réseau de neurones qui se base seulement sur les particules de l’état final. La matrice du plus bas ancêtre commun (LCA) associe à chaque pair de particules de l’état final l’ancêtre commun le plus loin de la particule d’origine dans la chaîne de désintégration comme montré en figure 3.2 droite haut. Si on remplace la particule elle-même par sa génération¹ on obtient la matrice de la génération du plus bas ancêtre commun (LCAG) montrée en figure 3.2 droite bas. Le premier résultat de réf [2] est de prouver qu’on peut décrire sans ambiguïté un arbre de désintégration avec la LCAG. Cette matrice est donc une bonne cible pour un réseau de neurones.

L’ordre des particules de l’état final n’a pas une signification au contraire de l’ordre des mots dans une phrase ou des pixels dans une image. Le réseau devra donc être invariant par permutation des particules considérées. En ce sens les réseaux de type graphe (GNN) ou les transformers sont particulièrement adaptés. Dans réf [2] nous avons entraîné un GNN [24] et un transformer [25] sur un jeu de donnée synthétique où une particule de masse 100 (en unités arbitraires) peut se désintégrer dans des particules intermédiaires avec masses entre 90 et 10 et particules finales avec masses entre 21 et 1.

1. combien de niveaux, il faut descendre pour arriver aux particules de l’état final

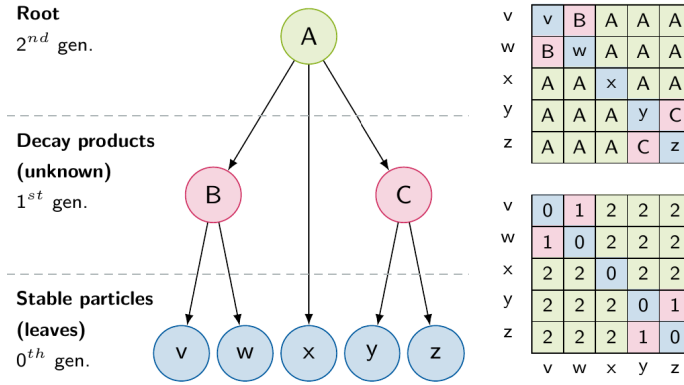
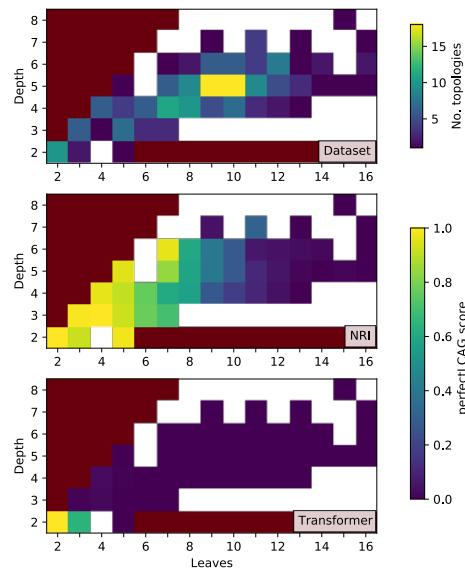


FIGURE 3.2 – (Gauche) Exemple d’arbre hiérarchique pouvant représenter une désintégration, (droite haut) matrice LCA et (droite bas) matrice LCAG. Reproduit de réf. [2].

FIGURE 3.3 – (Haut) Couverture de l’ensemble de données et (centre et bas) fraction des désintégrations où la matrice LCAG a été parfaitement prédite pour le réseau de type graphe et le transformer. Les régions en rouge sombre indiquent les régions pas autorisées dans l’échantillon et les régions en blanc celles autorisées, mais non couvertes par l’échantillon considéré. Reproduit de réf. [2].



Chaque particule peut se désintégrer en deux à cinq particules suivant un modèle d’espace de phases. De cette manière, nous considérons 200 topologies différentes et nous entraînons sur un échantillon de 2000 désintégrations par topologie². Figure 3.3 *haut* montre la répartition de ces topologies par profondeur de l’arbre de désintégration et nombre de particules de l’état final. Figure 3.3 *centre* et *bas* montre les performances du réseau de type graphe et du transformer en termes de fraction des désintégrations où la LCAG a été parfaitement prédite.

On peut observer que le transformer arrive à apprendre seulement les topologies les plus simples, peut-être à cause d’un échantillon d’entraînement trop petit. Le GNN par contre arrive à prédire correctement les désintégrations avec au plus huit niveaux dans 92,5% des cas s’il y a jusqu’à six particules dans l’état final et dans 59,7% des cas s’il y en a jusqu’à dix.

Utiliser la matrice LCAG comme cible d’entraînement présente toutefois

Matrice LCAS

2. Et le même nombre pour les échantillons de validation et de test.

deux désavantages majeurs : la profondeur de l'arbre de la désintégration n'a pas de limite de profondeur ce qui complexifie la tâche. En outre, dans une désintégration, le même état final peut être atteint par des résonances différentes qui peuvent interférer quantiquement entre elles. De ce fait, la matrice LCAG n'est pas bien définie en général. Par la suite, on a donc utilisé comme cible d'entraînement la matrice du niveau du plus bas ancêtre commun (LCAS) où sept classes d'ancêtres possibles sont considérées³. On a en même temps prédit aussi l'identité des particules de l'état final avec un deuxième système de classes⁴.

Le GraFEI J'ai supervisé Arthur, Jacopo et Corentin qui ont développé et optimisé le GraFEI, un GNN plus général que celui utilisé dans la publication précédente et basé sur les *blocs de réseau de graphes* [26]. Ce modèle travaille directement sur des graphes : un ensemble de nœuds connectés par des arêtes. Des caractéristiques peuvent être associées à chaque nœud, arête ou au graphe dans son ensemble. Pour notre cas, on considère en entrée un graphe totalement connecté où les nœuds sont les particules de l'état final. Les variables considérées dans notre cas sont les suivantes.

- **Les caractéristiques des nœuds** encodent des informations sur les particules de l'état final. Le modèle utilise les caractéristiques de nœuds suivantes : charge de la particule, p_T , p_z , distance en longitudinale et transversale du point de collision, variables d'identification des particules, ainsi que certaines informations sur les clusters. Dans le cas où une valeur n'est pas disponible (par exemple, les variables des clusters calculées pour les particules chargées), la variable est définie à 0.
- **Les caractéristiques des arêtes** sont des quantités qui dépendent des paires de particules. Le modèle utilise les caractéristiques d'arêtes suivantes : le cosinus de l'angle azimutal entre les moments des particules ($\cos \vartheta$) et la distance d'approche la plus proche entre deux traces (DOCA). Afin de calculer le DOCA pour les photons, on suppose qu'ils passent par l'origine du système de coordonnées.
- **Les caractéristiques globales** peuvent être des informations sur le graphe global (par exemple, le nombre de particules de l'état final, la quantité de mouvement totale, etc.). Le modèle n'utilise actuellement aucune caractéristique globale en entrée.

Un bloc de réseau de graphes prend en entrée un graphe et retourne en sortie le même graphe, mais avec des variables différentes comme montré en figure 3.4. Le GraFEI est composé d'un bloc d'entrée, un de sortie et un ou plusieurs blocs intermédiaires. Chaque bloc de réseau de graphes est à son tour composé de trois sous-blocs : les sous-blocs d'arêtes, de nœuds et globaux :

3. 6 pour $\Upsilon(4S)$, 5 pour B, 4 pour D^* , 3 pour D, 2 pour K_s^0 , 1 pour π^0 ou J/ψ et 0 pour les particules en dehors de l'arbre de désintégration.

4. 1 pour e, 2 pour μ , 3 pour π , 4 pour K, 5 pour p, 6 pour γ et 0 pour tout le reste.

- **Le sous-bloc des arêtes** est utilisé pour mettre à jour les caractéristiques des arêtes. Le cœur du bloc est un Perceptron Multi-Couches (MLP) ϕ^e appelé *fonction de mise à jour* du sous-bloc. Pour chaque arête e , le vecteur d'entrée de ϕ^e contient les caractéristiques des arêtes, les caractéristiques des deux nœuds liés par l'arête et les caractéristiques globales. La sortie du sous-bloc d'arêtes est également utilisée comme entrée pour les sous-blocs de nœuds et globaux, via l'utilisation de *fonctions d'agrégation* $\rho^{e \rightarrow v}$ et $\rho^{e \rightarrow u}$. Pour chaque caractéristique de sortie, ces fonctions calculent la moyenne arithmétique pour les arêtes connectées au même nœud et pour toutes les arêtes du même graphe. La sortie du sous-bloc d'arêtes dans le dernier bloc de réseau de graphes est utilisée pour prédire la matrice LCAS du graphe. Cela est réalisé en évaluant les probabilités pour chaque arête d'appartenir aux sept classes définies dans la matrice LCAS.
- **Le sous-bloc des nœuds** est utilisé pour mettre à jour les caractéristiques des nœuds. Le cœur du bloc est un MLP ϕ^v . Pour chaque nœud v , le vecteur d'entrée de ϕ^v contient les caractéristiques du nœud, la sortie de $\rho^{e \rightarrow v}$, et les caractéristiques globales. La sortie du sous-bloc de nœuds est également utilisée comme entrée pour le sous-bloc global, via l'utilisation de la fonction d'agrégation $\rho^{v \rightarrow u}$, qui, pour chaque caractéristique de sortie, calcule la moyenne arithmétique pour tous les nœuds du même graphe. La sortie du sous-bloc de nœuds dans le dernier bloc de réseau de graphes est utilisée pour prédire les hypothèses de masse des particules de l'état final. Cela est réalisé en évaluant les probabilités pour chaque nœud d'appartenir aux 7 classes définies pour l'identité des particules de l'état final.
- **Le sous-bloc global** n'est pas utilisé pour prédire des quantités physiques pour l'instant, mais il est utilisé pour propager l'information à travers le réseau. Le cœur du bloc est un MLP ϕ^u . Le vecteur d'entrée de ϕ^u contient les caractéristiques globales et les sorties de $\rho^{e \rightarrow u}$ et $\rho^{v \rightarrow u}$.

Le nombre de blocs de réseau de graphes intermédiaires ainsi que le nombre de couches des MLPs et la dimension de sortie des MLPs sont des hyperparamètres du modèle. Leurs valeurs optimales sont 1, 1 et 64 respectivement. Le modèle a donc environ 4 millions de paramètres. Le GraFEI est entraîné sur plusieurs millions d'événements de Belle II, simulés en utilisant comme cible les éléments de la matrice LCAS et l'identité des particules de l'état final en utilisant l'entropie croisée comme fonction de coût. Pour chaque particule de l'état final et élément de la matrice LCAS le modèle associe une probabilité à sa prédiction. La moyenne géométrique de ces probabilités peut être utilisée comme variable pour sélectionner les désintégrations qui ont une plus haute probabilité d'être bien reconstruites.

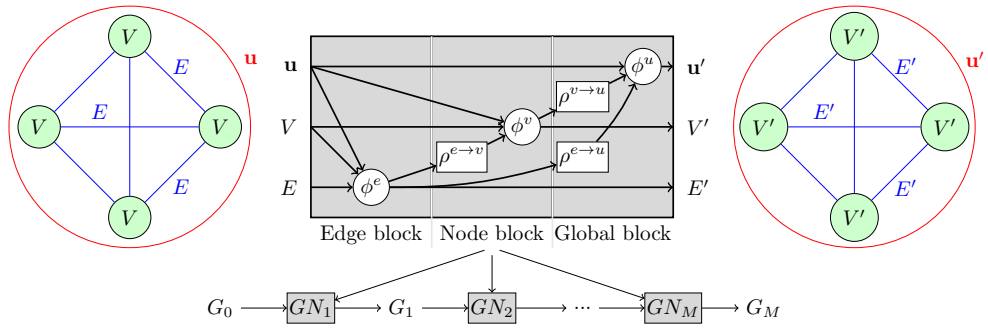


FIGURE 3.4 – Schéma du modèle GraFEI : le graphe d'entrée G_0 est passé à travers une série de blocs de réseau de graphes (GN_i), chacun composé de trois sous-blocs. Le graphe de sortie G_M a la même structure que le graphe d'entrée, mais avec des caractéristiques mises à jour, et est utilisé pour faire des prédictions sur les quantités d'intérêt. Reproduit de réf. [3].

Performances du GraFEI

Le GraFEI est implémenté dans le software de Belle II et peut être utilisé dans les recherches de particules manquantes de deux différentes façons : si le signal est reconstruit en premier, comme dans l'ITA, l'autre B peut être reconstitué à partir du reste de l'événement. Alternativement, l'événement entier peut être reconstruit. La figure 3.5 gauche compare les performances du GraFEI pour la reconstruction d'un méson B avec celles du FEI. Le GraFEI présente une efficacité maximale plus grande et permet d'obtenir une efficacité double par rapport au FEI avec le même niveau de bruit de fond. Il a toutefois des performances similaires dans la région à haute pureté. Merna a appliqué le GraFEI à la recherche de $B^+ \rightarrow K^+ \nu \bar{\nu}$ avec une analyse de test similaire à celle décrite dans le chapitre 2, mais où elle a considéré seulement les incertitudes statistiques. L'erreur statistique sur la fraction d'embranchement est améliorée de près d'un facteur deux par rapport à l'HTA, alors qu'il est d'environ 20% pire que l'ITA comme on peut voir en figure 3.5 droite. On s'attend à ce que ces résultats puissent être améliorés ultérieurement avec une analyse plus affinée.

La plus grande efficacité du GraFEI par rapport au FEI est due à deux facteurs : le nombre plus important de chaînes de désintégrations reconstruites, mais aussi le fait que le GraFEI reconstruit aussi des chaînes incomplètes où des particules de l'état final n'ont pas été reconstruites. De ce fait, il se rapproche plus de l'ITA que du HTA. Un autre point d'attention est que ses performances dépendent de la qualité de la simulation utilisée pendant l'entraînement. Sa calibration est possible en utilisant l'embedding décrit dans la section 2.2.

GraFEI à LHCb

L'utilisation de GNNs pour la reconstruction générique d'un méson B a aussi été exploré à LHCb [27, 28] à la suite de la définition de la matrice LCAG dans ref. [2]. Le but est toutefois différent que dans Belle II : dans les différentes jouvences de LHCb l'énorme volume de données et le pileup fait en sorte qu'il ne soit pas possible de sauvegarder l'événement entier, mais

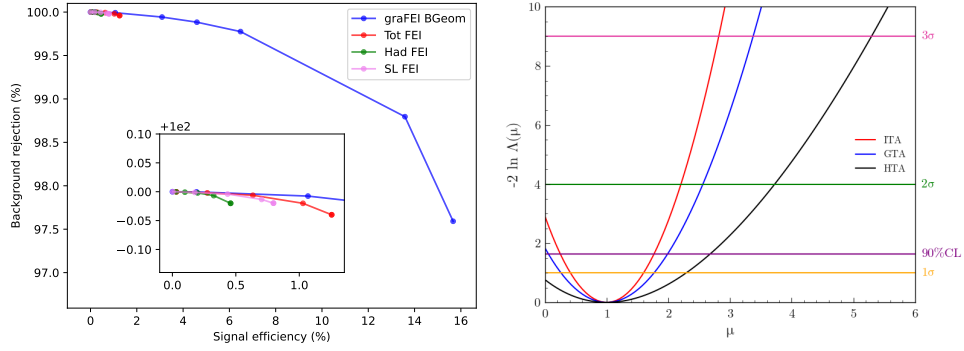


FIGURE 3.5 – (*Gauche*) Comparaison des courbes ROC du FEI et du GraFEI. (*Droite*) deux fois le rapport de log-vraisemblance du profil négatif en fonction de μ le rapport entre le rapport d'embranchement mesuré et théorique de la désintégration $B^+ \rightarrow K^+ \nu \bar{\nu}$ pour ITA, HTA et analyse utilisant le GraFEI (GTA). Seulement les incertitudes statistiques sont considérées dans l'ajustement. Reproduit de réf. [3].

seulement la partie qui contient la désintégration d'intérêt. Pouvoir identifier dans un évènement toutes les particules qui proviennent de la même chaîne de désintégration est utile pour deux aspects. Premièrement, il permet une sélection générique ce qui est intéressant pour des recherches inclusives. Deuxièmement, même quand la désintégration d'intérêt est simple à identifier, il permet l'identification des autres particules provenant de la même collision proton-proton ce qui est utile par exemple pour savoir si un méson B neutre est un B^0 ou un $\bar{B}^0$.

Les axions et les particules de type axion (ALPs), dénotés dans la suite par a , sont des états hypothétiques légers, pseudoscalaires, activement recherchés par le biais d'expériences en laboratoire et d'observations astrophysiques. L'axion de QCD [29] est une solution au problème de l'absence de violation de CP observée dans l'interaction forte. L'axion et les ALPs sont également des candidats viables pour la matière noire. Contrairement aux ALPs génériques, la masse m_a et son couplage de l'axion QCD sont liées : un axion QCD plus léger se couple nécessairement plus faiblement, ses interactions évoluant comme $1/f_a$ où f_a est l'échelle de Peccei-Quinn.

Axions et ALPs

La recherche de $B^+ \rightarrow K^+ \nu \bar{\nu}$ peut être réinterprétée comme $B^+ \rightarrow K^+ a$ car l'axion ne serait pas détecté et aurait donc les mêmes caractéristiques expérimentales que le couple de neutrinos. La seule différence est que $B^+ \rightarrow K^+ a$ est une désintégration à deux corps et pas à trois comme $B^+ \rightarrow K^+ \nu \bar{\nu}$. On s'attend donc à un pic monochromatique à m_a^2 dans la distribution du q^2 . La fraction d'embranchement de $B^+ \rightarrow K^+ a$ attendue est [12, 30]

$$\mathcal{B}(B^+ \rightarrow K^+ a) \propto |f_+^K(m_a^2)|^2 |(F_V)_{sb}|^2$$

où $(F_V)_{sb} = 2f_a/(k_V)_{sb}$ est l'échelle de Peccei-Quinn redimensionné par le couplage $(k_V)_{sb}$ de l'axion avec les quarks s et b . Le facteur de forme f_+^K est le même que pour $B^+ \rightarrow K^+ \nu \bar{\nu}$ introduit dans la section 1.2.

Avec des collègues théoriciens d'Annecy et ma doctorante Merna Abumusabh, j'ai réinterprété comme recherche de $B^+ \rightarrow K^+ a$ les résultats de réf [1], décrits aussi dans le chapitre 2. Le point de départ est figure 2.4 qui montre la distribution en q^2 dans HTA et ITA. On a effectué un ajustement par modèle à trois composantes : $B^+ \rightarrow K^+ a$, $B^+ \rightarrow K^+ \nu \bar{\nu}$ et le bruit de fond. La forme du bruit de fond est prise directement des figures, tandis que pour les deux autres composantes nous connaissons la distribution théorique en q^2 . Dans le cas de HTA, où q^2 est reconstruit à partir de l'information du B_{tag} , cette distribution théorique est convoluée avec une gaussienne¹ pour inclure les effets de résolution. Dans le cas de ITA, où le B_{tag} n'est pas reconstruit explicitement, q^2 est approximé par q_{rec}^2 comme montré dans eq. 2.1. On dérive alors la correspondance entre les distributions en ces deux variables. Dans l'ajustement la fraction d'embranchement de $B^+ \rightarrow K^+ a$, $\mathcal{B}_a$ est le paramètre libre, la normalisation du bruit de fond est contrainte

Analyse

1. De largeur $(450 \text{ MeV})^2$, comparable à la résolution en masse contrainte par l'énergie du faisceau du B_{tag} [18].

avec une gaussienne de largeur 1% pour ITA e 10% pour HTA ². Nous considérons deux cas distincts pour la normalisation de $B^+ \rightarrow K^+ \nu \bar{\nu}$: soit elle est fixée à la valeur prédite par le Modèle Standard ($\mu_{\nu \bar{\nu}} = 1$) soit elle est libre de varier ($\mu_{\nu \bar{\nu}} \neq 1$). L'ajustement est considéré séparément par ITA, HTA et leur combinaison. Une limite supérieure peut être ainsi obtenu sur $\mathcal{B}_a$ ainsi qu'une limite inférieure correspondante sur $(F_V)_{sb}$.

Résultats

La masse de l'axion de QCD est essentiellement zéro à l'échelle des collisionneurs. En considérant ITA et HTA ensemble, nous avons ainsi obtenu pour $m_a = 0$, sous l'hypothèse ($\mu_{\nu \bar{\nu}} = 1$), $\mathcal{B}(B^+ \rightarrow K^+ a) < 1,4 \times 10^{-6}$ et $(F_V)_{sb} > 7,6 \times 10^8$ GeV. Avec $\mu_{\nu \bar{\nu}}$ libre de varier, nous obtenons alors $\mathcal{B}(B^+ \rightarrow K^+ a) < 1,2 \times 10^{-6}$ et $(F_V)_{sb} > 8,0 \times 10^8$ GeV. Cette limite est la plus contraignante au monde et améliore de presque un ordre de grandeur la meilleure limite précédente [31]. Les résultats des autres fits sont résumés en tableau 4.1 Figure 4.1 montre l'évolution des limites sur $\mathcal{B}(B^+ \rightarrow K^+ a)$ et $(F_V)_{sb}$ en fonction de la masse de l'ALP. On voit que pour ITA et l'ajustement combiné, les limites sont pires pour $m_a \sim 2$ GeV où il y a une discrédance entre données et l'ajustement en figure 2.4.

Il est aussi intéressant de remarquer que les deux courbes, pour $\mu_{\nu \bar{\nu}} = 1$ et pour $\mu_{\nu \bar{\nu}} \neq 1$, sont proches. Cela est dû au fait que les distributions en q^2 de $B^+ \rightarrow K^+ a$ et $B^+ \rightarrow K^+ \nu \bar{\nu}$ sont très différentes et donc l'ajustement peut partager entre les deux composantes. En cas de présence de nouvelle physique dans ce canal, il serait donc possible de savoir si elle est due à nouvelles particules à hautes énergies et donc une modification des coefficients de Wilson pour $B^+ \rightarrow K^+ \nu \bar{\nu}$, ou à une nouvelle particule légère comme c'est le cas pour $B^+ \rightarrow K^+ a$.

2. Cette valeur permet de reproduire la distribution de deux fois le rapport de log-vraisemblance du profil négatif de figure 2.5 gauche dans l'absence de $B^+ \rightarrow K^+ a$. Il a été vérifié qu'augmenter significativement cette incertitude n'a pas d'effet notable sur le résultat final.

TABLE 4.1 – Limites supérieures sur le rapport de branchement de $B^+ \rightarrow K^+ a$ et limites inférieures sur l'échelle de Peccei-Quinn redimensionnée par le couplage à 90% de niveau de confiance pour les différents cas considérés. Toutes les limites supposent $m_a = 0$. Reproduit de réf. [4].

	ITA	HTA	Combined
$\mu_{\nu\bar{\nu}} = 1$	$n_{\text{dof}} = 2$ $\mathcal{B}_a < 1.1 \times 10^{-6}$ $ (F_V)_{sb} > 8.4 \times 10^8 \text{ GeV}$	$n_{\text{dof}} = 2$ $\mathcal{B}_a < 5 \times 10^{-6}$ $ (F_V)_{sb} > 4 \times 10^8 \text{ GeV}$	$n_{\text{dof}} = 3$ $\mathcal{B}_a < 1.4 \times 10^{-6}$ $ (F_V)_{sb} > 7.6 \times 10^8 \text{ GeV}$
$\mu_{\nu\bar{\nu}} \neq 1$	$n_{\text{dof}} = 3$ $\mathcal{B}_a < 1.0 \times 10^{-6}$ $ (F_V)_{sb} > 8.9 \times 10^8 \text{ GeV}$	$n_{\text{dof}} = 3$ $\mathcal{B}_a < 6 \times 10^{-6}$ $ (F_V)_{sb} > 4 \times 10^8 \text{ GeV}$	$n_{\text{dof}} = 4$ $\mathcal{B}_a < 1.2 \times 10^{-6}$ $ (F_V)_{sb} > 8.0 \times 10^8 \text{ GeV}$

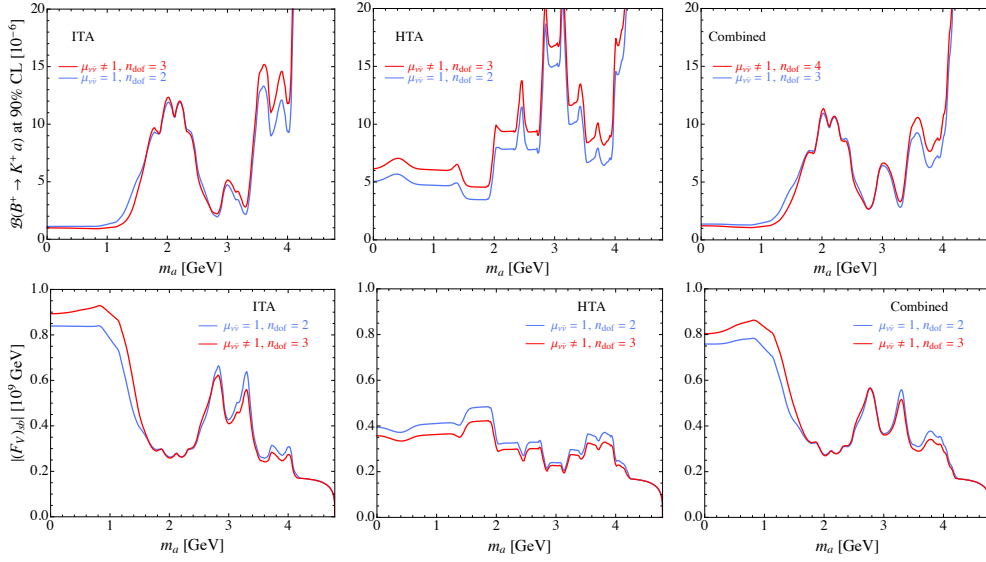


FIGURE 4.1 – (*Haut*) Limite supérieure sur $\mathcal{B}_a$ et (*bas*) limite inférieure sur l'échelle de Peccei-Quinn redimensionnée par le couplage $(F_V)_{sb}$ à 90% de niveau de confiance en fonction de la masse m_a , avec $\mu_{\nu\bar{\nu}}$ fixé à un (ligne bleue) ou laissé libre (ligne rouge). Reproduit de réf. [4].

Comme décrit dans le chapitre 2, la recherche de $B^+ \rightarrow K^+ \nu \bar{\nu}$ avec les méthodes ITA et HTA par Belle II a porté à une évidence et une tension de 2,7 écarts types par rapport à la prédiction du Modèle Standard, ce qui a suscité beaucoup d'intérêt dans la communauté. Belle II est la seule expérience actuelle à pouvoir effectuer cette mesure, il n'y a donc pas de confirmation totalement indépendante possible. Nous pouvons toutefois effectuer une analyse avec l'approche tagged semi-leptonique (STA), où l'autre B de l'événement est reconstruit en $B \rightarrow D^{(*)} l \nu$, avec l un électron ou un muon. La comparaison avec les analyses de Belle et BaBar, au tableau 2.3, nous montre que cette approche devrait permettre une précision similaire à celle de ITA. L'approche STA, où le B_{tag} est reconstruit en seulement quatre modes de désintégrations, est très différente de l'approche inclusive de ITA. On s'attend donc qu'il soit assez complémentaire en terme d'effets systématiques. En figure 2.5 droite, les analyses STA sont les plus proches de la valeur prédite par le Modèle Standard, il est donc intéressant d'effectuer ce type de mesure aussi à Belle II pour comprendre si cela est seulement une fluctuation statistique ou le résultat d'une différence systématique. Comme détaillé dans le chapitre 1, $B \rightarrow K^* \nu \bar{\nu}$ offre des contraintes complémentaires à $B \rightarrow K \nu \bar{\nu}$ sur les modèles de nouvelle physique possible. Il est donc important de rechercher au moins un de ces modes.

5.1 RECHERCHE DE $B \rightarrow K^{(*)} \nu \bar{\nu}$ AVEC L'APPROCHE STA

Pour ces raisons, je travaille actuellement avec les doctorants Corentin Santos et Merna Abumusabh à la recherche de $B^+ \rightarrow K^+ \nu \bar{\nu}$ et $B^0 \rightarrow K^{*0} \nu \bar{\nu}$ avec l'approche STA en utilisant 488 fb^{-1} de données au pic du $\Upsilon(4S)$ ¹. La stratégie d'analyse est similaire à celle suivie par HTA.

Le B_{tag} est reconstruit en premier en utilisant le FEI semi-leptonique. Le signal est identifié avec un kaon ou une paire kaon-pion chargés avec une masse compatible avec celle du K^{*0} . L'événement ne doit pas contenir d'autres traces, K_S^0 , π^0 ou Λ que ceux provenant du B_{tag} ou du signal. Comme un neutrino est présent dans l'état final du B_{tag} , il n'est donc pas possible d'utiliser l'impulsion reconstruite du B_{tag} pour calculer q^2 et

Sélection

¹. Cet échantillon correspond aux données collectées par Belle II pendant le run 1 et une partie run 2 jusqu'à la fin de 2024.

l'équation 2.1 est utilisée comme pour ITA. Un arbre de décision différent pour chaque canal avec une vingtaine de variables chacun est ensuite utilisé pour réduire ultérieurement le bruit de fond. Les variables principales sont l'énergie extra, la somme de l'énergie et l'impulsion manquante, $\cos \vartheta_{BY}^2$, q^2 et, pour $B^0 \rightarrow K^{*0} \nu \bar{\nu}$, la masse du K^{*0} . Les rapports d'embranchements sont obtenus d'un ajustement par modèle simultané des sorties des BDTs des deux canaux.

Échantillons de contrôles

Dans cette analyse aussi la simulation joue un rôle central. Pour cette analyse nous utilisons une simulation améliorée par rapport à celle utilisée pour HTA et ITA : le bruit de fond des faisceaux est pris directement des données et l'état du détecteur et les caractéristiques du point de collision sont variés avec le temps pour reproduire celles observées pendant la prise de données [32]. Il est toutefois encore nécessaire de corriger la simulation et estimer plusieurs incertitudes systématiques. Plusieurs échantillons de contrôles sont utilisés à ce propos :

offres Un échantillon sous-résonance pour de contraindre le bruit de fond continuum.

wrong-charge Un échantillon dans lequel le B_{tag} et le kaon de signal ont la même charge pour $B^+ \rightarrow K^+ \nu \bar{\nu}$ et la même saveur pour $B^0 \rightarrow K^{*0} \nu \bar{\nu}$. Dans ce deuxième cas une contamination du signal est attendue à cause du mixing entre B^0 et $\bar{B}^0$, mais elle est négligeable pour les fins de l'analyse.

extra-tracks Pour construire cet échantillon les critères de qualité utilisés pour sélectionner les traces qui ne proviennent ni du B_{tag} ni du signal sont plus stricts. De cette manière, on rejette moins d'événements qu'avec les critères utilisés dans la sélection de base en demandant qu'il n'y ait pas de ces traces en plus. Les événements sélectionnés en plus par rapport à la sélection de base forment l'échantillon extra track.

embedding Pour évaluer la correction sur l'efficacité du signal des événements contenant $B^+ \rightarrow J/\Psi K^{(*)+}$ et $B^0 \rightarrow J/\Psi K^{*0}$ sont reconstruits dans les données et la simulation. La désintégration $B^+ \rightarrow J/\Psi K^{(*)+}$ est ensuite remplacée par un candidat $B^+ \rightarrow K^+ \nu \bar{\nu}$ simulé, et $B^0 \rightarrow J/\Psi K^{*0}$ est remplacée par $B^0 \rightarrow K^{*0} \nu \bar{\nu}$.

Corrections et incertitudes systématiques

La plupart des corrections et incertitudes systématiques sont similaires à celles décrites dans la section 2.2, les différences principales sont les suivantes.

- Les rapports d'embranchement de toutes les désintégrations présentes dans le bruit de fond de la région de signal sont corrigés à la valeur la plus à jour et leurs incertitudes propagées dans l'ajustement. Envi-

2. L'angle entre B_{tag} et $Y = D^{(*)}l$ de la désintégration $B_{tag} \rightarrow D^{(*)}l \nu$ dans le référentiel du $\Upsilon(4S)$ calculé comme $\cos \vartheta_{BY} = \frac{2E_B^* E_Y^* - m_B^2 - m_Y^2}{2|p_B^*||p_Y^*|}$.

ron 40% de la fraction d’embranchement des mésons B est composé de désintégrations inconnues. Dans la simulation ces désintégrations sont obtenues par hadronisation des quarks constituant le méson B avec le logiciel Pythia [33]. Un nombre très limité de désintégrations produites par Pythia est présent dans la région de signal. Nous appliquons la même correction globale à ces désintégrations afin de reporter à un la fraction d’embranchement totale du méson B. Elles sont ensuite laissées libre de varier individuellement dans l’ajustement.

- L’accord entre simulation et données de l’énergie extra est fortement amélioré. Nous n’appliquons plus donc une correction spécifique. Nous appliquons en revanche la correction sur le bruit de fond continuum obtenu de la BDT entraînée à l’aide de l’échantillon **offres**. Nous entraînons ensuite un deuxième BDT à l’aide de l’échantillon **extra-tracks**, qui nous permet de corriger les différences résiduelles entre simulation et données, due surtout au bruit de fond provenant des mésons B. Pour couvrir le possible désaccord résiduel, une incertitude systématique est associée à chacune des deux corrections. Chaque correction peut ainsi varier entre 0 et 200%.
- Les quatre canaux $B^+ \rightarrow K^+ \nu \bar{\nu}$, $B^0 \rightarrow K_S^0 \nu \bar{\nu}$, $B^+ \rightarrow K^{*+} \nu \bar{\nu}$, $B^0 \rightarrow K^{*0} \nu \bar{\nu}$ sont considérés comme des composantes séparées dans l’ajustement, car ils peuvent être présents soit comme signal, soit comme bruit de fond “cross-feed”. Nous considérons liés les rapports d’embranchement $\mathcal{B}(B^0 \rightarrow K^{(*)0} \nu \bar{\nu})$ et $\mathcal{B}(B^+ \rightarrow K^{(*)+} \nu \bar{\nu})$ selon équation 1.1.
- On considère dans l’ajustement aussi les désintégrations $B \rightarrow X_s \nu \bar{\nu}$ où X_s est une résonance contenant un quark s de masse plus grande que 1.1 GeV. La distribution de masse de X_s est inspirée par l’analyse en amplitude de $B^0 \rightarrow K_S^0 \pi^- \pi^+ \gamma$ [34] et des incertitudes systématiques sont assignées en faisant varier la proportion relative des différentes composantes. $B \rightarrow X_s \nu \bar{\nu}$ est libre de varier dans l’ajustement, mais contraint par la limite supérieure obtenue dans réf [35].
- Les échantillons d’**embedding** sont utilisés pour contraindre l’efficacité de sélection du signal, mais aussi la normalisation du bruit de fond $B\bar{B}$ où le B_{tag} est bien reconstruit. La normalisation du bruit de fond $B\bar{B}$ où le B_{tag} n’est pas bien reconstruit par le FEI est corrigée avec l’échantillon **extra-tracks**.

L’optimisation de cette analyse n’est pas encore terminée et, pour le moment, l’échantillon des données n’a pas encore été inspecté. Nous avons, toutefois, déjà une estimation des efficacités et puretés attendues dans la région de signal, présentes dans le tableau 5.1 avec les efficacités et puretés des analyses précédentes. Figure 5.1 montre le nombre d’événements attendus dans la région du signal pour le signal et les différentes composantes du bruit de fond. L’incertitude totale attendue sur $\mathcal{B}(B^+ \rightarrow K^+ \nu \bar{\nu})$

Résultats attendus

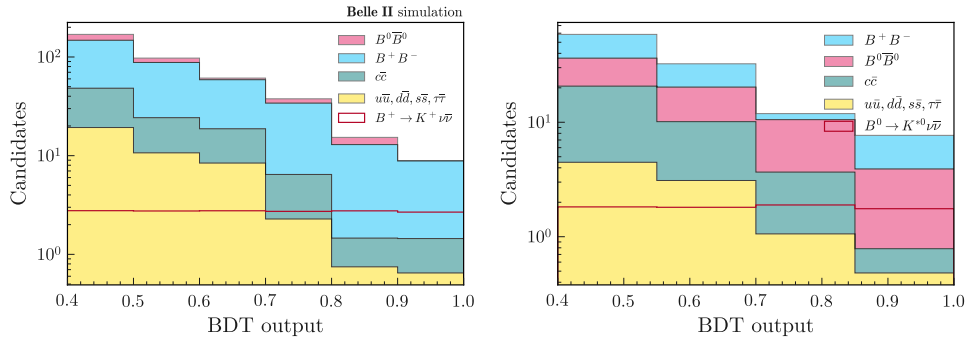


FIGURE 5.1 – Distributions attendues de la sortie de la BDT dans la région de signal pour le signal et les différentes composantes du bruit de fond pour (gauche) $B^+ \rightarrow K^+ \nu \bar{\nu}$ et (droite) $B^0 \rightarrow K^{*0} \nu \bar{\nu}$.

est $\mathcal{O}(5 \times 10^{-6})$, donc en accord avec celle attendue avec l'approche STA³ et celle sur $\mathcal{B}(B^0 \rightarrow K^{*0} \nu \bar{\nu})$ est $\mathcal{O}(2 \times 10^{-5})$. Figure 5.2 compare cette incertitude attendue avec les incertitudes obtenues par les analyses précédentes. Cette mesure sera limitée statistiquement et les incertitudes systématiques principales seront dues à la dimension de l'échantillon simulé et à la normalisation des fonds.

TABLE 5.1 – Comparaison des efficacités et puretés des différentes recherches de $B^+ \rightarrow K^+ \nu \bar{\nu}$

Analyse $B^+ \rightarrow K^+ \nu \bar{\nu}$	Efficacité	Pureté
BaBar HTA [16]	0.04%	17%
BaBar STA [15]	0.16%	35%
Belle HTA [17]	0.05%	5%
Belle STA [10]	0.2%	5%
Belle II HTA [1]	0.4%	3.5%
Belle II ITA [1]	8%	0.8%
Belle II STA	0.7%	5%
Belle II STA $B^0 \rightarrow K^{*0} \nu \bar{\nu}$	0.2%	5%

5.2 PERSPECTIVES À PLUS LONG TERME

Analyse combiné STA,
HTA, ITA

La prochaine étape logique après l'analyse décrite dans la section précédente est une analyse plus complète qui étudie aussi les canaux $B^0 \rightarrow K_S^0 \nu \bar{\nu}$ et $B^+ \rightarrow K^{*+} \nu \bar{\nu}$. Comme décrit dans la section 1.1 Belle II a maintenant collecté un échantillon de 808 fb^{-1} au pic du $\Upsilon(4S)$, ce qui permet une amélioration d'environ 30% sur l'incertitude statistique. Tableau 5.1 montre que les différentes approches, STA, HTA, ITA, ont des points de travail différents en terme d'efficacité et pureté, avec STA qui offre la plus haute pureté sui-

3. Cette incertitude rescalée par la luminosité à 365 fb^{-1} est 8×10^{-6} comme pour les analyses STA de Belle et BaBar montrées dans la tableau 2.3.

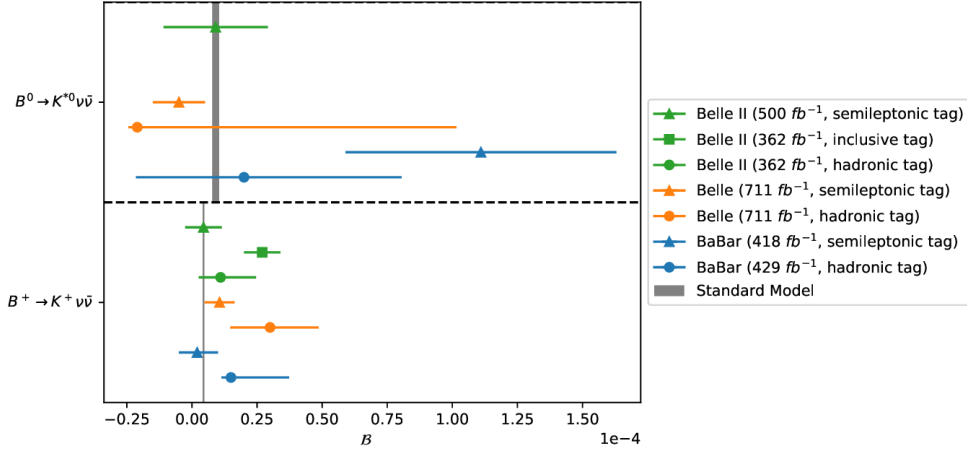


FIGURE 5.2 – Valeur de la fraction d’embranchement de $B^+ \rightarrow K^+ \nu \bar{\nu}$ et $B^0 \rightarrow K^{*0} \nu \bar{\nu}$ obtenue par différentes expériences [10, 14-17] comparée aux prédictions théoriques [9] et à la valeur attendue dans l’analyse STA de Belle II.

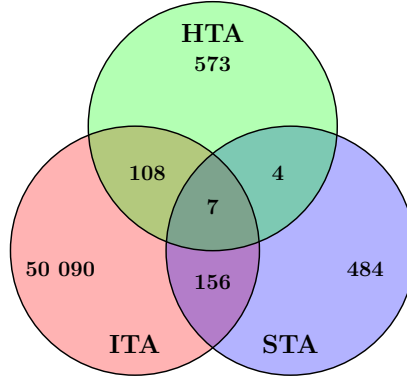


FIGURE 5.3 – Diagramme de Venn avec le nombre d’évènements extrapolés à 1 ab^{-1} dans la région de signal des analyses $B^+ \rightarrow K^+ \nu \bar{\nu}$ ITA et HTA de réf. [1] et de l’analyse STA en cours.

vie de HTA et enfin de ITA qui a une efficacité beaucoup plus grande. On envisage donc une analyse combinée, où les évènements sélectionnés par plusieurs approches seraient attribués à l’approche la plus pure. Figure 5.3 montre la proportion d’évènements qui seraient sélectionnés par plusieurs approches. Dans une analyse combinée, HTA perdrait donc seulement 1.6% des évènements sélectionnés et ITA 0.5%. Les analyses actuelles STA, HTA et ITA sont optimisées pour la recherche du signal à trois corps $B \rightarrow K^{(*)} \nu \bar{\nu}$. Elles sont donc plus efficaces à bas q^2 et ne seraient donc pas adaptées à la recherche de $B \rightarrow K^{(*)} X_{\text{dark}}$ si la nouvelle particule invisible avait une masse importante. À cette fin, une analyse alternative pourrait être envisagée, où la sortie du BDT soit indépendant du q^2 en utilisant, par exemple, les techniques de réf [36].

Une dernière remarque sur l’efficacité et pureté relative de STA et HTA. Dans HTA le B_{tag} est reconstruit complètement et peut être utilisé pour

Pureté relative de HTA et STA

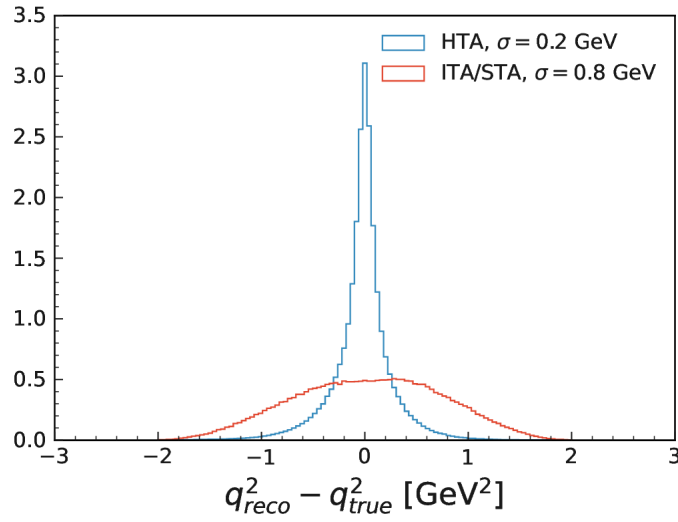


FIGURE 5.4 – Comparaison de la résolution en q^2 entre HTA où l'information cinématique du B_{tag} est utilisée et ITA ou STA où seulement le B_{sgn} est utilisé avec l'approximation qu'il soit à repos dans le référentiel du $\Upsilon(4S)$.

contraindre la cinématique du signal, ce qui n'est pas possible dans STA à cause du neutrino dans l'état final du B_{tag} . L'efficacité du FEI semi-leptonique est plus grande que celle du FEI hadronique. En outre, la reconstruction de $B_{tag} \rightarrow D^{(*)}l\nu$ est aussi généralement plus pure que celle du B_{tag} dans un état hadronique, qui contient généralement plus de particules dont des π^0 . La différence en pureté entre les deux approches dépend toutefois du signal recherché. L'approche HTA sera plus pure si l'information cinématique du B_{tag} peut être utilisée efficacement pour discriminer le signal du bruit de fond, comme c'est le cas d'états finaux avec un seul neutrino. Dans notre cas, au contraire, l'information cinématique n'offre pas d'avantage significatif et STA est alors plus pur comme montré dans le Tableau 5.1. L'approche HTA, cependant, permet d'utiliser l'information du B_{tag} pour estimer le q^2 sans faire recours à l'approximation d'équation 2.1. Cela permet un gain en résolution d'approximativement un facteur quatre comme montré en figure 5.4. Cela est particulièrement intéressant pour la recherche de nouvelles particules invisibles.

CONCLUSIONS

Ce document retrace mes contributions à la recherche des désintégrations $B \rightarrow K^{(*)}\nu\bar{\nu}$. Le résultat principal est l'évidence de $B^+ \rightarrow K^+\nu\bar{\nu}$, illustrée dans le chapitre 2, et qui présente une déviation à 2,7 écarts types par rapport à la prédiction du Modèle Standard. Cette mesure a été réinterprétée comme $B^+ \rightarrow K^+\alpha$ dans le chapitre 4, ce qui a permis d'obtenir la limite la plus contraignante au monde sur l'échelle de Peccei-Quinn divisée par le couplage de l'axion avec les quarks s et b. Une approche alternative pour la recherche de $B^+ \rightarrow K^+\nu\bar{\nu}$ en utilisant un réseau profond de type graphe est discutée dans le chapitre 3. Enfin une nouvelle recherche de $B^+ \rightarrow K^+\nu\bar{\nu}$ et $B^0 \rightarrow K^{*0}\nu\bar{\nu}$ est en cours comme décrit dans le chapitre 5, qui contient aussi quelques perspectives plus à long terme.

RÉFÉRENCES

- [1] Belle II collaboration, I. ADACHI et al., *Evidence for $B^+ \rightarrow K^+ \nu \bar{\nu}$ decays*, *Phys. Rev. D* **109**, 112006 (2024), [arXiv :2311.14647](#).
- [2] J. KAHN, I. TSAKLIDIS, O. TAUBERT, L. REUTER, G. DUJANY, T. BOECKH, A. THALLER, P. GOLDENZWEIG, F. BERNLOCHNER, A. STREIT et M. GÖTZ, *Learning tree structures from leaves for particle decay reconstruction*, *Machine Learning: Science and Technology* **3**, 035012 (2022), [arXiv :2208.14924](#).
- [3] M. ABUMUSABH, J. CERASOLI, G. DUJANY et C. SANTOS, *Graph-based Full Event Interpretation : a graph neural network for event reconstruction in Belle II*, in 27th International Conference on Computing in High Energy and Nuclear Physics (mars 2025), [arXiv :2503.09401](#).
- [4] M. ABUMUSABH, G. DUJANY, D. GUADAGNOLI, A. IOHNER et C. TONI, *Using $B^+ \rightarrow K^+ \nu \bar{\nu}$ to search for the QCD axion*, *Phys. Rev. Res.* **8**, 023129 (2026), [arXiv :2510.18953](#).
- [5] Belle II collaboration, T. ABE et al., *Belle II Technical Design Report*, (2010), [arXiv :1011.0352](#).
- [6] W. ALTMANNSHOFER, A. J. BURAS, D. M. STRAUB et M. WICK, *New strategies for New Physics search in $B \rightarrow K^* \nu \bar{\nu}$, $B \rightarrow K \nu \bar{\nu}$ and $B \rightarrow X_s \nu \bar{\nu}$ decays*, *JHEP* **04**, 022 (2009), [arXiv :0902.0160](#).
- [7] A. J. BURAS, J. GIRRBACH-NOE, C. NIEHOFF et D. M. STRAUB, *$B \rightarrow K^{(*)} \nu \bar{\nu}$ decays in the Standard Model and beyond*, *JHEP* **02**, 184 (2015), [arXiv :1409.4557](#).
- [8] D. BEČIREVIĆ, G. PIAZZA et O. SUMENSARI, *Revisiting $B \rightarrow K^{(*)} \nu \bar{\nu}$ decays in the Standard Model and beyond*, *Eur. Phys. J. C* **83**, 252 (2023), [arXiv :2301.06990](#).
- [9] L. ALLWICHER, D. BECIREVIC, G. PIAZZA, S. ROSAURO-ALCARAZ et O. SUMENSARI, *Understanding the first measurement of $B(B \rightarrow K \nu \bar{\nu})$* , *Phys. Lett. B* **848**, 138411 (2024), [arXiv :2309.02246](#).
- [10] Belle collaboration, J. GRYGIER et al., *Search for $B \rightarrow h \nu \bar{\nu}$ decays with semileptonic tagging at Belle*, *Phys. Rev. D* **96**, [Addendum: *Phys. Rev. D* **97**, 099902 (2018)], 091101 (2017), [arXiv :1702.03224](#).
- [11] J. F. KAMENIK et C. SMITH, *Tree-level contributions to the rare decays $B^+ \rightarrow \pi^+ \nu \bar{\nu}$, $B^+ \rightarrow K^+ \nu \bar{\nu}$ and $B^+ \rightarrow K^{*+} \nu \bar{\nu}$ in the Standard Model*, *Phys. Lett. B* **680**, 471-475 (2009), [arXiv :0908.1174](#).

- [12] P. D. BOLTON, S. FAJFER, J. F. KAMENIK et M. NOVOA-BRUNET, *Signatures of light new particles in $B \rightarrow K^{(*)} \text{Emiss}$* , *Phys. Rev. D* **110**, [Erratum: *Phys.Rev.D* **111**, 039903 (2025)], 055001 (2024), arXiv :2403.13887.
- [13] P. D. BOLTON, S. FAJFER, J. F. KAMENIK et M. NOVOA-BRUNET, *Impact of new invisible particles on $B \rightarrow K^{(*)} \text{Emiss}$ observables*, *Phys. Rev. D* **112**, 035010 (2025), arXiv :2503.19025.
- [14] BaBar collaboration, B. AUBERT et al., *Search for $B \rightarrow K^* \nu \bar{\nu}$ decays*, *Phys. Rev. D* **78**, 072007 (2008), arXiv :0808.1338.
- [15] BaBar collaboration, P. del AMO SANCHEZ et al., *Search for the Rare Decay $B \rightarrow K \nu \bar{\nu}$* , *Phys. Rev. D* **82**, 112002 (2010), arXiv :1009.1529.
- [16] BaBar collaboration, J. P. LEES et al., *Search for $B \rightarrow K^{(*)} \nu \bar{\nu}$ and invisible quarkonium decays*, *Phys. Rev. D* **87**, 112005 (2013), arXiv :1303.7465.
- [17] Belle collaboration, O. LUTZ et al., *Search for $B \rightarrow h^{(*)} \nu \bar{\nu}$ with the full Belle $\Upsilon(4S)$ data sample*, *Phys. Rev. D* **87**, 111103 (2013), arXiv :1303.3719.
- [18] T. KECK et al., *The Full Event Interpretation*, *Comput. Softw. Big Sci.* **3**, 6 (2019), arXiv :1807.08680.
- [19] G. POLAT, *Search for $B \rightarrow K^{(*)} \nu \bar{\nu}$ decays in the Belle II experiment*, BELLE2-MTHESIS-2021-064, Université de Strasbourg (2019).
- [20] M. SCHMIDT, *BSM for $B \rightarrow K$ + invisible and connections to other physics*, in 2025 Belle II Physics Week (October 2025).
- [21] I. TSAKLIDIS, *Demonstrating learned particle decay reconstruction using Graph Neural Networks at BelleII*, BELLE2-MTHESIS-2020-006, Université de Strasbourg (2020).
- [22] C. SANTOS, *Optimisation of a deep graph neural network to reconstruct generic B decays at Belle II*, BELLE2-MTHESIS-2024-002, Université de Strasbourg (2023).
- [23] M. ABUMUSABH, *Search for $B^+ \rightarrow K^+ \nu \bar{\nu}$ at Belle II using Graph Neural Networks*, BELLE2-MTHESIS-2024-021, Université de Strasbourg (2024).
- [24] J. GILMER et al., *Neural Message Passing for Quantum Chemistry*, in *Proceedings of the 34th International Conference on Machine Learning (ICML '17)*, t. 70, Proceedings of Machine Learning Research (2017), p. 1263-1272.
- [25] A. VASWANI et al., *Attention Is All You Need*, in *Advances in Neural Information Processing Systems*, t. 30 (2017).
- [26] P. W. BATTAGLIA et al., *Relational inductive biases, deep learning, and graph networks*, (2018), arXiv :1806.01261.

- [27] J. GARCÍA PARDINAS, M. CALVI, J. ESCHLE, A. MAURI, S. MELONI, M. MOZZANICA et N. SERRA, *GNN for Deep Full Event Interpretation and Hierarchical Reconstruction of Heavy-Hadron Decays in Proton–Proton Collisions*, *Comput. Softw. Big Sci.* **7**, 12 (2023), [arXiv :2304.08610](#).
- [28] W. SUTCLIFFE, M. CALVI, S. CAPELLI, J. ESCHLE, J. GARCÍA PARDIÑAS, A. MATHAD, A. UZUKI et N. SERRA, *Scalable multi-task learning for particle collision event reconstruction with heterogeneous graph neural networks*, *Mach. Learn. Sci. Tech.* **6**, 045060 (2025), [arXiv :2504.21844](#).
- [29] R. D. PECCEI et H. R. QUINN, *CP Conservation in the Presence of Instantons*, *Phys. Rev. Lett.* **38**, 1440-1443 (1977).
- [30] Particle Data Group, S. NAVAS et al., *Review of particle physics*, *Phys. Rev. D* **110**, 030001 (2024).
- [31] J. MARTIN CAMALICH, M. POSPELOV, P. N. H. VUONG, R. ZIEGLER et J. ZUPAN, *Quark Flavor Phenomenology of the QCD Axion*, *Phys. Rev. D* **102**, 015023 (2020), [arXiv :2002.04623](#).
- [32] G. GAUDINO, *Run Dependent Monte Carlo at Belle II*, *PoS EPS-HEP2025*, 559 (2026).
- [33] C. BIERLICH et al., *A comprehensive guide to the physics and usage of PYTHIA 8.3*, *SciPost Phys. Codeb.* **2022**, 8 (2022), [arXiv :2203.11601](#).
- [34] BaBar, P. del AMO SANCHEZ et al., *Time-dependent analysis of $B^0 \rightarrow K_S^0 \pi^- \pi^+ \gamma$ decays and studies of the $K^+ \pi^- \pi^+$ system in $B^+ \rightarrow K^+ \pi^- \pi^+ \gamma$ decays*, *Phys. Rev. D* **93**, 052013 (2016), [arXiv :1512.03579](#).
- [35] Belle-II, M. ABUMUSABH et al., *First Search for $B \rightarrow X s \nu \bar{\nu}$ Decays*, *Phys. Rev. Lett.* **136**, 231801 (2026), [arXiv :2511.10980](#).
- [36] A. ROGOZHNIKOV, A. BUKVA, V. V. GLIGOROV, A. USTYUZHANIN et M. WILLIAMS, *New approaches for boosting to uniformity*, *JINST* **10**, T03002 (2015), [arXiv :1410.4140](#).