

UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia “Galileo Galilei”

Corso di Laurea in Fisica

Tesi di Laurea

Ricerca del decadimento raro $B \rightarrow \eta' K$ a Belle II

Relatore

Prof. Stefano Lacaprara

Laureando

Filippo Conforto

Anno Accademico 2019/2020

Sommario

Il decadimento $B \rightarrow \eta' K$ è un decadimento raro con uno stato finale a simmetria CP definita, che permette una misura dell'angolo $\phi_1(\beta)$ del triangolo di unitarietà in presenza di loop adronici. I dati raccolti all'esperimento Belle II in Giappone durante il 2019 e 2020 permettono una riscoperta di questo canale di decadimento. Utilizzando dati simulati ed eventi di collisione raccolti da Belle II è possibile individuare una serie di modalità di selezione che permettono di migliorare la significanza del segnale rispetto al fondo. Verranno discusse in dettaglio le tecniche di ottimizzazione del segnale e i risultati da esse ottenute.

Indice

1	Introduzione	1
1.1	Belle II	1
1.2	Il decadimento raro $B \rightarrow \eta' K$	2
1.3	<i>Dataset e dataframe</i>	2
2	Metodo e strumenti di analisi	4
2.1	Distribuzione delle variabili di selezione	4
2.1.1	Insieme di variabili per il canale $B^0 \rightarrow \eta'(\rightarrow \eta(\rightarrow \gamma\gamma)\pi^+\pi^-)K_S^0$	4
2.1.2	Insieme di variabili per il canale $B^0 \rightarrow \eta'(\rightarrow \rho\gamma)K_S^0$	5
2.1.3	Insieme di variabili per i canali carichi	6
2.1.4	Variabili di <i>Continuum Suppression</i>	6
2.2	Algoritmo di selezione	7
3	Applicazione al canale $B^0 \rightarrow \eta'(\rightarrow \eta(\rightarrow \gamma\gamma)\pi^+\pi^-)K_S^0$	10
3.1	Risultati per il fondo <i>Continuum + Peaking</i>	11
3.2	Risultati per il fondo <i>Continuum</i> con ottimizzazione sulle variabili di <i>Continuum Suppression</i>	13
3.3	Risultati per il fondo <i>Continuum + Peaking</i> in SR con preselezione sulle variabili di <i>Continuum Suppression</i>	15
4	Applicazione agli altri canali	17
4.1	Risultati per il canale $B^\pm \rightarrow \eta'(\rightarrow \eta(\rightarrow \gamma\gamma)\pi^+\pi^-)K^\pm$	17
4.2	Risultati per canale $B^0 \rightarrow \eta'(\rightarrow \rho\gamma)K_S^0$	18
4.3	Risultati per il canale $B^\pm \rightarrow \eta'(\rightarrow \rho\gamma)K^\pm$	19
	Conclusioni	21

Capitolo 1

Introduzione

1.1 Belle II

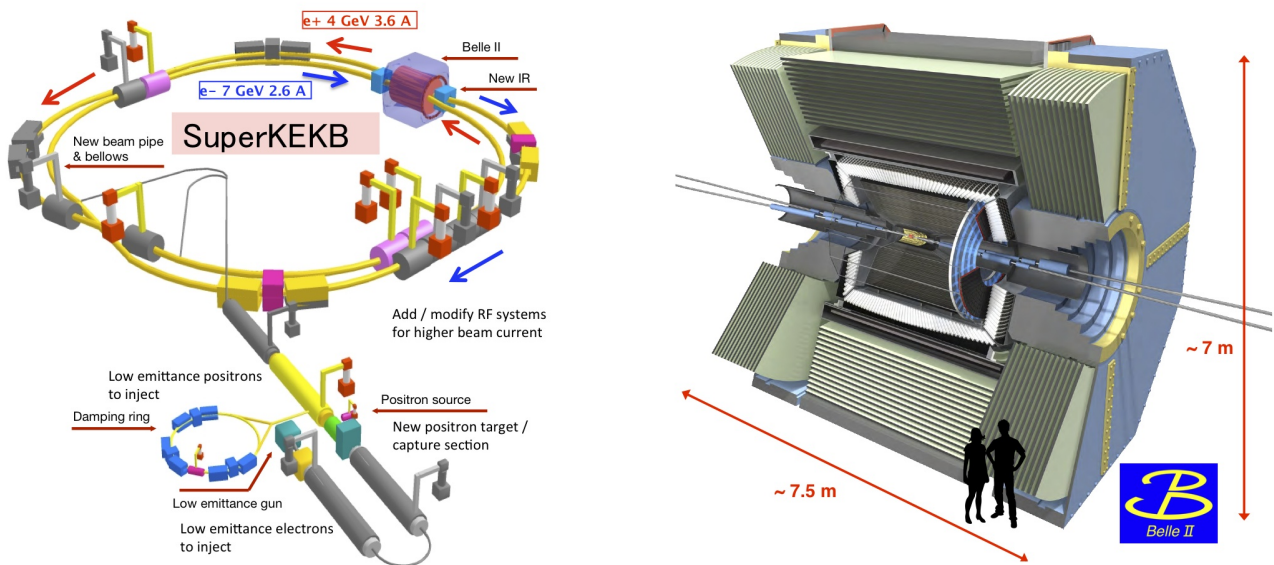


Figura 1.1: L'acceleratore di particelle superKEKB e la struttura del rivelatore Belle 2.

Belle II è un rivelatore adibito allo studio dei mesoni B che raccoglie dati sulle collisioni elettrone-positrone prodotte all'interno dell'acceleratore superKEKB; uno schema dell'acceleratore è presentato in figura 1.1. Belle II costituisce il successore di Belle, che, durante il suo periodo di attività (1999 - 2009), ha raccolto dati per una luminosità integrata di 1000 fb^{-1} [1] sui processi in atto nell'acceleratore di particelle KEKB, un anello con 3 km di circonferenza situato a KEK, in Giappone. L'acceleratore KEKB è stato poi aggiornato in superKEKB [2] che, attraverso il miglioramento delle sorgenti di positroni ed elettroni e a un nuovo schema di collisione, permetterà di raggiungere una luminosità istantanea 40 volte maggiore e circa 50 ab^{-1} di luminosità integrata. All'interno dell'anello vengono fatti collidere fasci di elettroni e positroni prodotti in modo asimmetrico, accelerando gli elettroni a una energia di 7 GeV e i positroni a una energia di 4 GeV [1], in modo tale che il centro di massa si trovi nella regione di risonanza per la particella $\Upsilon(4S)$ e che abbia un *boost* rispetto al laboratorio. Le collisioni così prodotte costituiscono l'origine di una serie di processi, tra cui proprio la formazione della particella $\Upsilon(4S)$ con sezione d'urto $\sigma \approx 1.1 \text{ nb}$, la produzione di coppie quark-antiquark, la cui adronizzazione sarà origine della maggior parte del fondo per l'esperimento, e di coppie di leptoni (e^+e^- , $\tau^+\tau^-$, $\mu^+\mu^-$). La particella $\Upsilon(4S)$, se prodotta, decade con $\text{BR}(\text{Branching Ratio}) > 96\%$ [3] in coppie $B\bar{B}$ in stato *entangled*, garantendone una copiosa produzione. I dati raccolti da Belle hanno permesso di ottenere molti risultati significativi, tra i quali la conferma sperimentale dei meccanismi

di Kobayashi-Maskawa per la violazione CP (simmetria Carica-Parità) nei sistemi $B\bar{B}$ [1], mentre la costruzione di Belle II permetterà in futuro misure di grande precisione per la fisica dei mesoni B e la ricerca di possibili fenomeni di nuova fisica non previsti dal modello standard [4]. Il rivelatore è costituito da una serie di strati concentrici, mostrati in figura 1.1, focalizzati sulla zona di collisione delle coppie e^+e^- . A partire dal centro si hanno:

- Due strati di sensori a pixel di silicio (PXD - *PiXel Detector*) e quattro di sensori a *strip* di silicio (SVD - *Silicon Vertex Detector*), che si occupano di misurare le posizioni dei vertici di decadimento delle coppie di mesoni B.
- Una camera a *drift* centrale (CDC-*Central Drift Chamber*), che si occupa di rivelare le tracce di particelle cariche, calcolarne il momento e identificarle tramite la misura della perdita di energia.
- Una struttura cilindrica di contatori TOP (*Time-Of-Propagation*), presente nella parte centrale del rivelatore, che permette d'identificare le particelle (soprattutto kaoni e pioni) attraverso la luce Cherenkov che emettono attraversando alcuni radiatori di quarzo. Un rivelatore simile (ARICH - *Aerogel Ring-Imaging Cherenkov*), sempre basato sulla luce Cherenkov, è posizionato nella parte più avanzata di Belle II e lavora invece con radiatori di aerogel.
- Un calorimetro elettromagnetico (ECL - *Electromagnetic Calorimeter*), costituito da cristalli scintillatori.
- Un solenoide superconduttore, che produce un campo magnetico da 1.5 T.
- Il detector KLM (K_L *Muon*), adibito a rivelare kaoni e muoni che arrivano fino a questo strato più esterno.

Il rivelatore produce in questo modo una serie d'informazioni che sono utilizzate da un algoritmo di ricostruzione della catena di decadimento. [5]

1.2 Il decadimento raro $B \rightarrow \eta' K$

Tra i possibili canali di decadimento per i mesoni B, il decadimento $B \rightarrow \eta' K$ risulta molto interessante per lo studio di fenomeni rari. Lo studio di questi processi è utile per diversi motivi, tra i quali la possibilità di studiare l'asimmetria CP dipendente dal tempo [6] [7] o i processi di decadimento a pinguino; i diagrammi di Feynman di questi eventi contengono infatti *loop* che permettono la creazione di particelle di massa arbitraria, fondamentali per la ricerca di nuova fisica [8]. In questo lavoro sono stati trattati 2 canali di decadimento (carico e neutro), di cui sono stati studiati due stati finali per i decadimenti della particella η' . Inoltre, per il canale neutro, sono stati trattati solo gli stati per i quali avviene il decadimento $K_S \rightarrow \pi^+\pi^-$ (BR 69.2%).

Decadimento principale	BR	Stato finale	BR
$B^0 \rightarrow \eta' K^0$	$6.6 \cdot 10^{-5}$	$\eta' \rightarrow \eta(\rightarrow \gamma\gamma)\pi^+\pi^-$	$3.83 \cdot 10^{-6}$
		$\eta' \rightarrow \rho(\rightarrow \pi^+\pi^-)\gamma$	$6.60 \cdot 10^{-6}$
$B^\pm \rightarrow \eta' K^\pm$	$7.04 \cdot 10^{-5}$	$\eta' \rightarrow \eta(\rightarrow \gamma\gamma)\pi^+\pi^-$	$1.19 \cdot 10^{-5}$
		$\eta' \rightarrow \rho(\rightarrow \pi^+\pi^-)\gamma$	$2.04 \cdot 10^{-5}$

Tabella 1.1: *Branching ratios* per i decadimenti in studio e i loro relativi stati finali.

1.3 Dataset e dataframe

Il *dataset* fornito dal rivelatore contiene tutti i dati raccolti sulle collisioni in atto all'interno dell'acceleratore di particelle. Da questo *dataset* è possibile tentare di ricostruire il decadimento in studio, ottenendo un *dataframe* [9] contenente gli eventi e un insieme di variabili a essi associati. La procedura di ricostruzione produce però spesso eventi di segnale che non corrispondono effettivamente al decadimento in studio; questo avviene poiché l'impossibilità sperimentale d'identificare sempre correttamente le particelle attraversanti il rivelatore produce errate assegnazioni d'identità. L'algoritmo può quindi

considerare erroneamente un processo estraneo come se fosse di segnale, producendo così tre tipi di fondo:

- *Continuum* : Segnale proveniente dalla produzione di coppie quark-antiquark ($u\bar{d}sc$), con relativa adronizzazione, e di coppie $\tau\bar{\tau}$. Si tratta del contributo più ampio per il fondo, dovuto al largo numero di particelle cariche prodotte in questi processi, che possono essere identificate erroneamente in fase di ricostruzione.
- *Peaking*: Segnale proveniente dalla produzione di coppie $B\bar{B}$ che decadono in canali alternativi al decadimento in studio. Il fatto che il processo sia molto raro consente infatti ai mesoni B di decadere in altri canali e generare particelle che la ricostruzione può interpretare in maniera sbagliata, producendo un contributo non trascurabile al fondo. Il fondo *Peaking* è ulteriormente classificabile in due contributi: associato a particelle figlie provenienti dalle coppie B^+B^- (*charged*) o invece derivato dalle coppie $B^0\bar{B}^0$ (*mixed*).
- *SxF (Signal cross Feed)*: Segnale proveniente dalla produzione di decadimenti $B \rightarrow \eta'K$ non correttamente ricostruiti, ovvero per i quali una delle particelle figlie è stata male assegnata.

Per lo studio dei decadimenti sono stati utilizzati parte dei dati raccolti Belle II (34.6 fb^{-1} [10] su 74.1 fb^{-1} totali). Oltre ai dati sono stati generati eventi simulati, tramite tecniche Montecarlo, al fine di studiare separatamente il comportamento di segnale e fondo; in particolare, la simulazione include la generazione dell'intera collisione e^+e^- e la risposta del rivelatore alle particelle prodotte. I *dataset* così generati comprendono eventi dal *Continuum* ($u\bar{u}, d\bar{d}, s\bar{s}, c\bar{c}, \tau\bar{\tau}$), dal *Peaking* (*mixed* e *charged*) e segnale proveniente dai quattro possibili stati finali considerati in questo studio. Dopo la generazione sono quindi stati sottoposti all'algoritmo di ricostruzione, al fine di produrre un *dataframe* di eventi simulati.

L'analisi è stata condotta su questi *dataframe*, costituiti da un alto numero di eventi al fine migliorare la rilevanza statistica delle ottimizzazioni. In particolare, per ogni canale sono state usate simulazioni con le seguenti luminosità integrate:

- *Continuum*: 100 fb^{-1}
- *Peaking*: 1000 fb^{-1}
- Segnale + *SxF*: Per ogni canale sono stati simulati 200.000 eventi, la cui luminosità integrata associata varia tra 9 e 72 ab^{-1} a seconda del canale scelto

In fase di analisi, per poter confrontare *dataframe* con diversa luminosità integrata, si è deciso di scalare il numero di eventi di ogni gruppo, in modo tale che corrispondessero a quelli previsti per il *dataframe* dei dati raccolti, corrispondenti a una luminosità integrata di 34.6 fb^{-1} . In questo modo è stato inoltre possibile confrontare i dati simulati con quelli realmente prodotti dal rivelatore, al fine di controllare se l'attesa fosse più o meno verificata.

Capitolo 2

Metodo e strumenti di analisi

Il segnale raccolto dal rivelatore si trova in quantità molto ridotta rispetto al fondo, questo avviene a causa della rarità del decadimento e della grande quantità di processi che avvengono nel rivelatore. Al fine di aumentare la sua rilevanza statistica è quindi necessario selezionare gli eventi con alcune modalità, tra le quali l'imposizione di limiti al possibile valore delle variabili del *dataframe* in studio.

2.1 Distribuzione delle variabili di selezione

La ricostruzione della catena di decadimento permette di definire la distribuzione degli eventi per una lunga serie di variabili associate al processo di decadimento raro. Alcune di esse, per motivi legati alla cinematica, hanno distribuzioni particolari associate alla presenza o meno del decadimento raro stesso. Dato che i canali studiati risultano molto differenti fra loro per quanto riguarda il comportamento di queste variabili, vi saranno diversi insiemi a seconda del canale scelto. Per illustrare il significato delle selezioni discuterò ora le variabili utilizzate e le loro distribuzioni.

2.1.1 Insieme di variabili per il canale $B^0 \rightarrow \eta'(\rightarrow \eta(\rightarrow \gamma\gamma)\pi^+\pi^-)K_S^0$

Le variabili usate per il primo canale sono presentate in figura 2.1, e sono caratterizzate dalle seguenti distribuzioni di eventi:

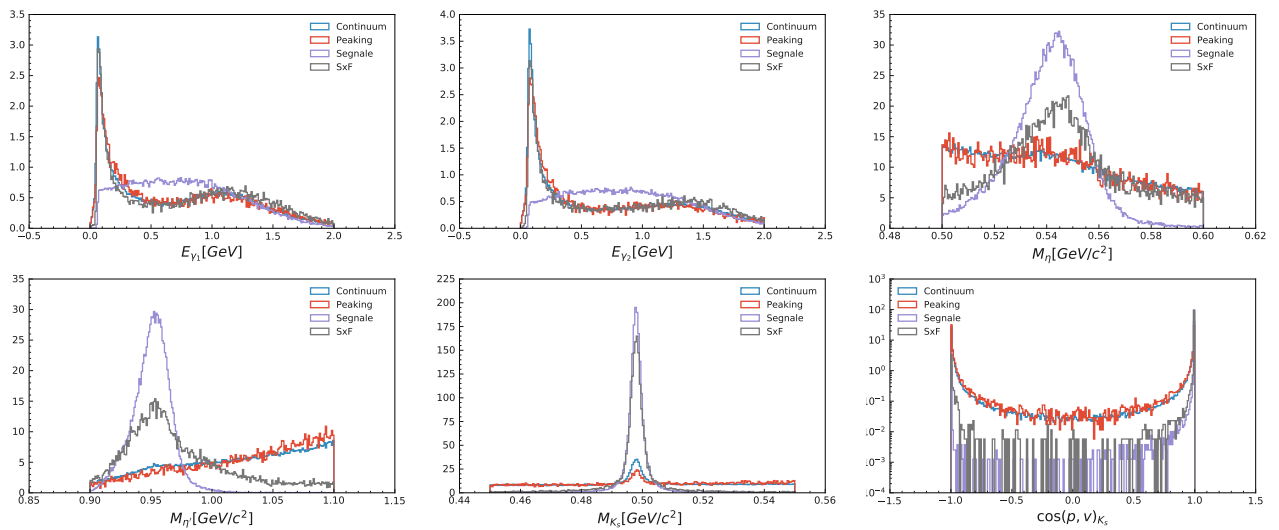


Figura 2.1: Distribuzioni normalizzate alla stessa area degli eventi del canale $B^0 \rightarrow \eta'(\rightarrow \eta(\rightarrow \gamma\gamma)\pi^+\pi^-)K_S^0$ per le variabili a esso associate, come descritto nel testo.

- E_{γ} : Le due variabili associate all'energia dei fotoni emessi risultano avere una distribuzione abbastanza uniforme per il segnale, con un andamento decrescente per valori di energia molto alti; il valore dell'energia del fotone è infatti superiormente limitato dall'energia del mesone B

prodotto. Gli eventi di fondo sono invece concentrati a energie molto basse, dove si trovano anche molti eventi di SxF ; la parziale ricostruzione è infatti spesso tale per l'assegnazione errata dei fotoni. Per quanto riguarda i contributi di *Continuum* e *Peaking*, questi si trovano anch'essi concentrati a energie basse, effetto probabilmente dovuto ai molti fotoni di fondo emessi da processi secondari, quali decadimenti di π^0 o il fondo di rivelatore.

- M_η : Essendo una massa invariante, gli eventi si trovano concentrati attorno al valore medio e nella regione di dispersione adiacente. Quest'ultima, come per la maggior parte delle masse invarianti considerate, è dovuta alla risoluzione sperimentale del rivelatore che, essendo prevalente rispetto alla larghezza di risonanza, risulta produrre un'ampia gaussiana. Il contributo di SxF , nonostante la ricostruzione errata, restituisce comunque una distribuzione abbastanza simile, mentre per i contributi di *Continuum* e di *Peaking*, l'assenza della particella nello stato finale genera una distribuzione abbastanza uniforme.
- $M_{\eta'}$: Si tratta nuovamente di una massa invariante, gli eventi di segnale sono quindi prevalentemente concentrati attorno al valore medio e nella regione di dispersione a esso adiacente. Gli eventi di SxF sono invece più dispersi, segno che l'errore sull'assegnazione di una particella influenza molto negativamente il valore di questa variabile. L'assenza di questa particella nello stato finale dei contributi di *Continuum* e *Peaking* genera una distribuzione abbastanza uniforme, che permette di distinguere questi ultimi eventi da quelli di segnale.
- M_{K_S} : Dato che la variabile rappresenta la massa invariante della coppia $\pi^+\pi^-$ associata al decadimento K_S , gli eventi di segnale e SxF sono prevedibilmente concentrati attorno al valore medio della massa ($\approx 500 \text{ MeV}/c^2$). Si nota inoltre che la dispersione degli eventi assume una forma molto più simile a una gaussiana, a causa della piccola larghezza di risonanza per il decadimento $K_S \rightarrow \pi^+\pi^-$. I contributi di fondo hanno un leggero picco centrale, probabilmente dovuto a K_S provenienti da altri decadimenti, mentre la maggior parte degli eventi è distribuita esternamente, a causa della mancanza di una vera K_S .
- $\cos(p, v)_{K_S}$: La variabile rappresenta il coseno dell'angolo tra il momento della particella K_S e il vettore congiungente i vertici di decadimento del mesone B e della K_S stessa. Se la particella ricostruita esiste, questi due vettori avranno la stessa direzione e perciò, per gli eventi di segnale, la variabile si concentrerà dunque attorno al valore 1. Il fondo è facilmente riconoscibile, in quanto l'assenza di una particella K_S permette alla variabile di assumere valori molto diversi. Si nota inoltre un picco per $\cos(p, v)_{K_S} = -1$, probabilmente dovuto al fatto che l'errore sperimentale induce l'algoritmo di ricostruzione a classificare kaoni prodotti molto vicini al punto di collisione come se fossero prodotti all'indietro.

In particolare, dato che le distribuzioni di E_{γ_1} e di E_{γ_2} sono sostanzialmente intercambiabili, in quanto rappresentano i due fotoni con i quali viene ricostruita la particella η , la selezione su di esse è stata ottimizzata imponendo gli stessi limiti contemporaneamente su entrambe le variabili.

2.1.2 Insieme di variabili per il canale $B^0 \rightarrow \eta'(\rightarrow \rho\gamma)K_S^0$

Al secondo canale neutro viene associato l'insieme di variabili presentato in figura 2.2, costituito da una serie di variabili in comune con il primo canale neutro e un gruppo di variabili caratteristiche, quali:

- E_γ : La distribuzione in figura 2.2, anche se riportata in scala logaritmica, conserva una distribuzione simile a quella già vista per il precedente canale. Si nota nuovamente l'accumulo di fondo in corrispondenza dell'origine, mentre il segnale si trova principalmente nella zona con $E_\gamma < 1 \text{ GeV}$.
- $\cos\theta_\gamma$: Si tratta dell'angolo tra la direzione del fotone e l'asse del rivelatore. La produzione di fotoni per gli eventi di segnale è più centrale, mentre gli eventi di fondo e di SxF sono caratterizzati da un accumulo nella regione con $\cos\theta_\gamma \lesssim -0.6$, in corrispondenza con la fine della parte centrale dell'ECL, e nella regione con $\cos\theta_\gamma \gtrsim 0.9$.

- M_ρ : In questo caso l'effetto della larga risonanza produce una distribuzione di tipo Breit-Wigner che amplia molto la zona d'individuazione di segnale. Sia il contributo di *Continuum* che quello di *Peaking* presentano una distribuzione non molto diversa dal segnale, mentre il contributo di *SxF* ha un picco spostato di 20 MeV rispetto a quello di segnale. Si può notare inoltre la presenza di un picco di eventi dovuto alla presenza di K_S , a circa 500 MeV/c².

Per le variabili $M_{\eta'}$, $\cos(p, v)_{K_S}$ e M_{K_S} valgono le considerazioni già fatte in precedenza.

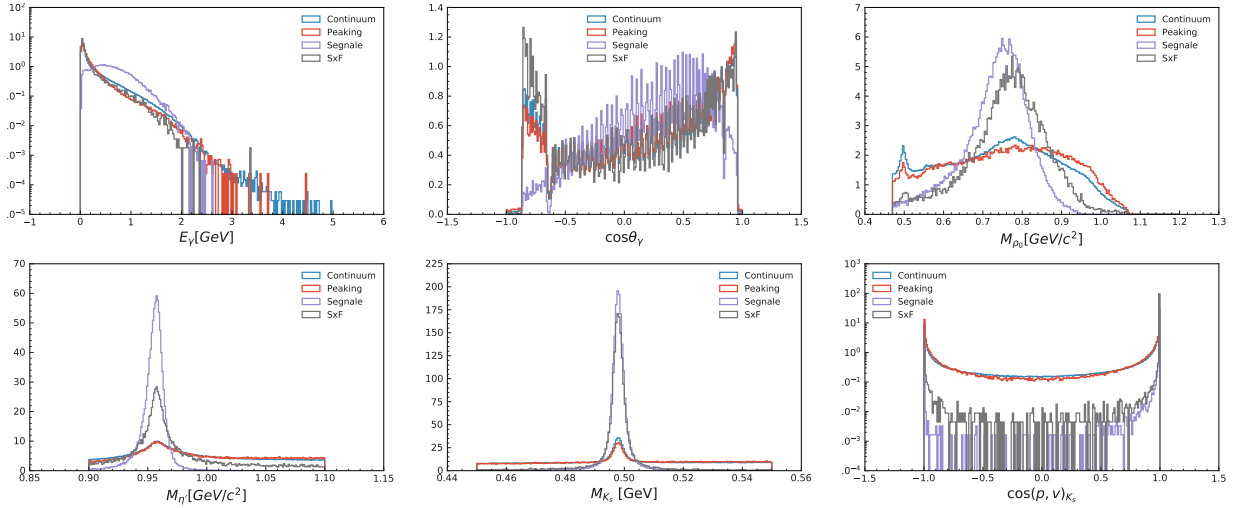


Figura 2.2: Distribuzioni normalizzate alla stessa area degli eventi del canale $B^0 \rightarrow \eta'(\rightarrow \rho\gamma)K_S^0$ per le variabili a esso associate, come descritto nel testo.

2.1.3 Insieme di variabili per i canali carichi

Per il canale $B^\pm \rightarrow \eta'(\rightarrow \eta(\rightarrow \gamma\gamma)\pi^+\pi^-)K^\pm$ è stato utilizzato lo stesso insieme di variabili del rispettivo canale neutro, rimuovendo M_{K_S} e $\cos(p, v)_{K_S}$, e inserendo la variabile $\cos\theta_K$. Analogamente, per il canale $B^\pm \rightarrow \eta'(\rightarrow \rho\gamma)K^\pm$ sono state rimosse dall'insieme le variabili $\cos(p, v)_{K_S}$ e M_{K_S} , ed è stata inserita la variabile $\cos\theta_K$.

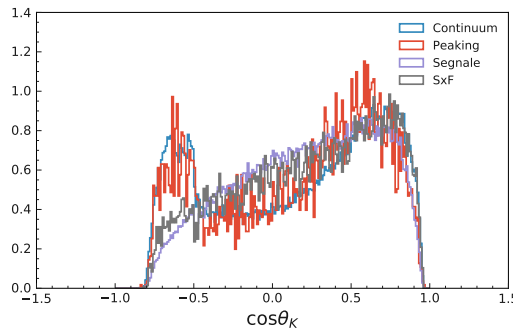


Figura 2.3: Distribuzioni normalizzate alla stessa area degli eventi del canale $B^\pm \rightarrow \eta'(\rightarrow \eta(\rightarrow \gamma\gamma)\pi^+\pi^-)K^\pm$ per la variabile $\cos\theta_K$.

La variabile $\cos\theta_K$, la cui distribuzione è visibile in figura 2.3, rappresenta l'angolo tra la direzione del kaone e l'asse del rivelatore, e presenta una distribuzione molto simile in entrambi i canali carichi. Il segnale è ben distribuito, mentre i contributi di *Continuum* e *Peaking* risultano avere un picco per bassi valori di coseno, probabilmente dovuti a particelle cariche rilevate in direzione opposta al fascio. Si tratta di un effetto dovuto alla difficoltà nel distinguere pioni e kaoni per $\cos\theta_K \lesssim -0.5$, in corrispondenza con la fine del rivelatore TOP.

2.1.4 Variabili di *Continuum Suppression*

Oltre alle variabili specifiche per i canali, è possibile utilizzare informazioni più generali per distinguere gli eventi $B\bar{B}$ dagli eventi di *Continuum*, grazie alla loro topologia; le coppie $B\bar{B}$ sono infatti prodotte

a partire da una risonanza ($\Upsilon(4S)$), decadendo a riposo nel sistema di riferimento del centro di massa. In questo modo i mesoni sono emessi in maniera pressoché isotropa, mentre gli eventi di *Continuum* sono caratterizzati dalla presenza di *jet* adronici prodotti in direzione opposta, che conferiscono così una forma prolata all'evento [11]. La variabile $\cos(\theta_B - \theta_O)$ distingue gli eventi di segnale da quelli di fondo utilizzando questo principio; θ_B è infatti l'angolo tra la direzione in cui viene emesso il mesone B ricostruito e l'asse dei fasci e θ_O l'angolo tra la direzione di emissione del resto dell'evento (per il decadimento in studio un altro mesone B) e l'asse dei fasci [10]. Nel caso di un evento planare, come quelli del contributo di *Continuum*, si avrà appunto una distribuzione concentrata sul valore 1. Una seconda variabile, R_2 , calcolata come rapporto tra i momenti di Fox-Wolfram di grado 2 e grado 0 [12], è anche in grado di quantificare la forma del processo. R_2 assume valori molto piccoli per gli eventi di tipo $B\bar{B}$, mentre per gli eventi di *Continuum* la distribuzione è concentrata attorno al valore 0.5.

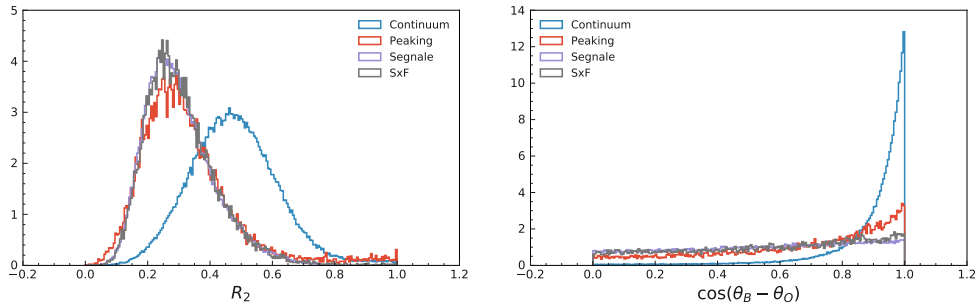


Figura 2.4: Distribuzioni normalizzate alla stessa area degli eventi del canale $B^0 \rightarrow \eta'(\rightarrow \eta(\rightarrow \gamma\gamma)\pi^+\pi^-)K_S^0$ per le variabili R_2 e $\cos(\theta_B - \theta_O)$.

Come si nota dalle distribuzioni in figura 2.4, il *Continuum* è molto localizzato, ed è quindi possibile introdurre una selezione sui domini delle due variabili che permetta di rimuovere la maggior parte del fondo.

2.2 Algoritmo di selezione

Il lavoro di ottimizzazione prevede la massimizzazione di un parametro che è fondamentale per stabilire la significanza del segnale: la figura di merito. La figura di merito (FoM - *Figure of Merit*) viene definita per un *dataframe* di eventi come:

$$\frac{S}{\sqrt{S+B}}$$

dove S e B rappresentano rispettivamente il conteggio del numero di eventi di segnale e di fondo per l'intero *dataframe*. Al fine di considerare solo l'insieme di eventi più importante per lo studio del segnale, sarà in seguito introdotta una regione di segnale, sui cui eventi verrà fatto il calcolo della figura di merito. La FoM, oltre a esprimere quanto il segnale risulti significativo rispetto l'errore statistico, di tipo poissoniano, sul numero totale di eventi, dipende direttamente dalla proporzione di segnale e fondo; attraverso delle selezioni sugli eventi del *dataframe* si è quindi in grado d'intervenire direttamente sul valore della figura di merito.

Applicando una selezione "rettangolare" sulle variabili, cioè limitandone superiormente e inferiormente i possibili valori, è possibile aumentare o diminuire la figura di merito e di conseguenza la significanza del segnale. L'approccio seguito è stato dunque quello di "scansionare" il dominio delle variabili, cioè calcolare la figura di merito al variare dei parametri di selezione, in modo da individuare quale fosse la soluzione più adatta per il problema. In particolare, è stato studiato il comportamento della figura di merito, come mostrato in figura 2.5, per tre tipi di selezione sul dominio delle variabili:

1. Selezione degli eventi appartenenti a un dominio inferiormente limitato.
2. Selezione degli eventi appartenenti a un dominio superiormente e inferiormente limitato.
3. Selezione degli eventi appartenenti a un dominio superiormente limitato.

Gli estremi della selezione sono scelti utilizzando come riferimento i quantili del *dataset* di segnale, in modo tale da poter escludere progressivamente un numero di eventi di segnale sempre costante.

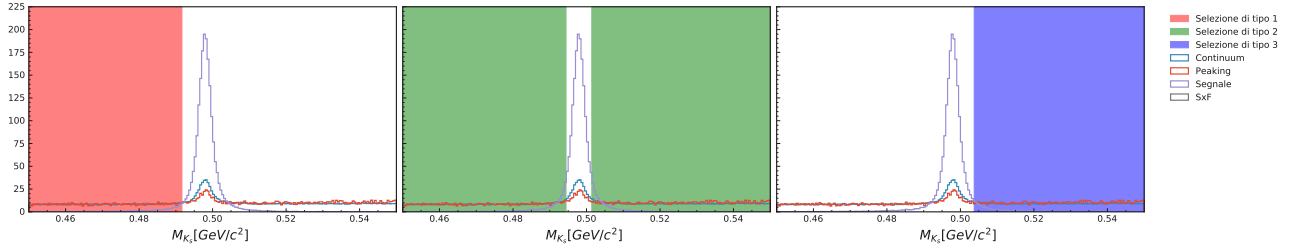


Figura 2.5: Esempio di eventi esclusi dal *dataframe* per le tre diverse selezioni sul dominio di M_{K_S} .

Il quantile di segnale scelto durante la "scansione" corrisponde al limite inferiore o superiore, a seconda del tipo di selezione, mentre l'imposizione di entrambi i limiti contemporaneamente prevede che i due estremi del dominio corrispondano al quantile scelto e al suo reciproco. Per poter avere un'espressione della figura di merito approssimativamente continua e con una densità sufficiente per avere buoni risultati, la scansione è stata eseguita su 200 quantili e, una volta ottenuta l'espressione della figura di merito per la variabile in studio, per ognuno dei tipi di selezione sono stati individuati i limiti che massimizzassero la FoM.

Per rendere ancora più efficiente l'analisi, sono state provate selezioni asimmetriche, ottenute bloccando un estremo della selezione e scansionando su 400 quantili il dominio, alla ricerca di un massimo della figura di merito per la selezione complementare. Se si era già imposto un limite inferiore (superiore) si è dunque cercato un limite superiore (inferiore) a esso complementare, mentre se erano già presenti due estremi di selezione si è imposto solo uno dei due alla volta e si è cercato un limite a esso complementare, come mostrato in figura 2.6.

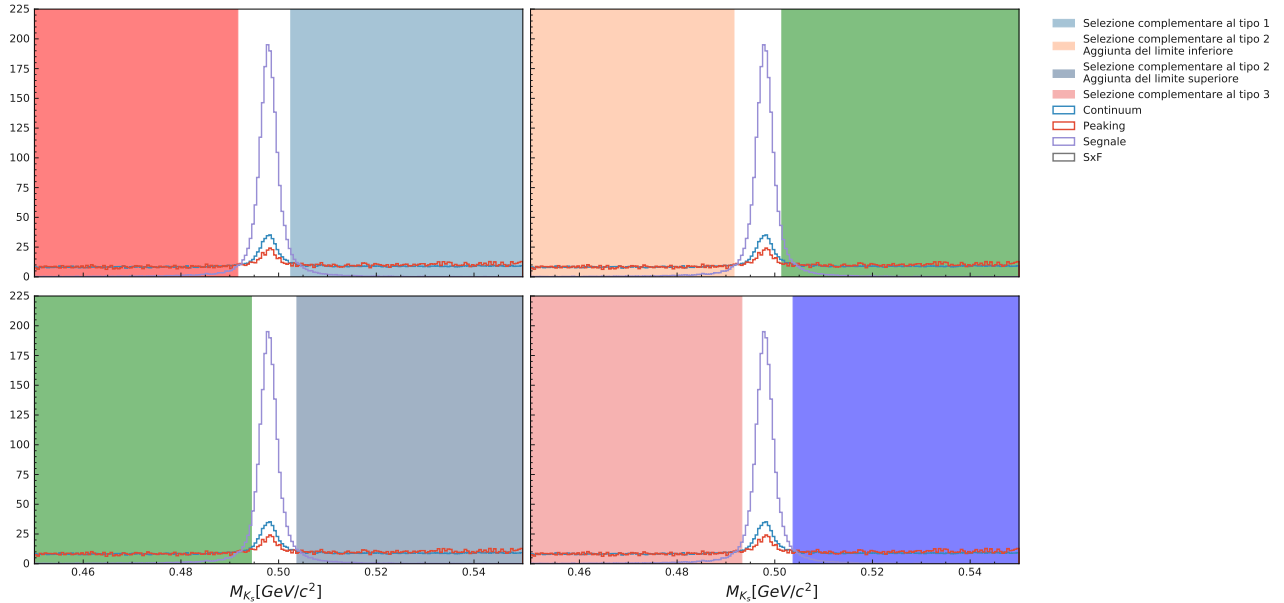


Figura 2.6: Esempio di eventi esclusi dal *dataframe* per le quattro possibili selezioni complementari sul dominio di M_{K_S} .

In questo modo è stato possibile individuare 4 possibili coppie di limiti, tra le quali è stata scelta quella per la quale la selezione di eventi avesse figura di merito massima.

Utilizzando questo principio di base vengono eseguiti diversi tentativi di ottimizzazione. Il risultato migliore in termini di conservazione del segnale, ma allo stesso tempo di rimozione del fondo, è restituito dalla seguente serie di operazioni:

1. Si definisce innanzitutto un insieme di selezioni proveniente da un primo approccio al problema, che permetta di rimuovere le zone che a occhio nudo contengono più fondo:

- $E_\gamma > 150$ MeV
- $0.5 < M_\eta < 0.57$ GeV/ c^2

- $0.92 < M_{\eta'} < 1.0 \text{ GeV}/c^2$
- $0.51 < M_{\rho} < 1 \text{ GeV}/c^2$
- $\cos \theta_{\gamma} > -0.64$
- $\cos \theta_K > -0.5$
- $\cos(p, v)_{K_s} > 0.99$
- $0.49 < M_{K_s^0} < 0.51 \text{ GeV}/c^2$
- $R_2 > 0.5$
- $\cos(\theta_B - \theta_O) > 0.7$ [10]

Si tratta di un insieme di selezioni preliminari abbastanza ragionevoli, ma non ottimizzate.

2. Una volta scelto il canale da studiare, si procede quindi ad applicare tutte le selezioni preliminari per l'insieme di variabili a esso associato, tranne una. Per quest'ultima variabile si scansiona quindi il dominio, cercando la coppia di estremi che massimizzi la figura di merito per l'insieme di eventi che passano la selezione.
3. Eseguendo questa ottimizzazione su tutte le variabili associate a un canale, è possibile ottenere un gruppo di selezioni che ha la caratteristica di minimizzare la correlazione reciproca, permettendo in questo modo di aumentare la significatività del segnale.

Questo tipo di analisi è stata eseguita utilizzando come fondo di riferimento per il calcolo della figura di merito diversi tipi di contributi. Sono stati trattati principalmente tre tipi di fondo:

- Fondo costituito dalla combinazione di *Continuum* e *Peaking*.
- Fondo costituito dal solo *Continuum*, ma applicando l'ottimizzazione anche alle variabili di *Continuum Suppression*.
- Fondo costituito dalla combinazione di *Continuum* e *Peaking*, considerato solo nella regione di segnale (SR - *Signal Region*) e con l'applicazione delle selezioni preliminari sulle variabili di *Continuum Suppression*.

Per regione di segnale si intende una regione di particolare interesse, definita come l'insieme di eventi che rispettano la selezione $M_{bc} > 5.2 \text{ GeV}/c^2$ e $-0.07 < \Delta E < 0.05 \text{ GeV}$ [10], dove M_{bc} è ottenuto come $\sqrt{E_{beam}^2 - p^{*2}}$ (con E_{beam} energia dei fasci e p^* momento del mesone B ricostruito), mentre ΔE è ottenuto come differenza tra l'energia della coppia di B e l'energia complessiva dei fasci. La regione a essa complementare è detta *SideBand* (SB) e permette di studiare la distribuzione del fondo per l'esperimento.

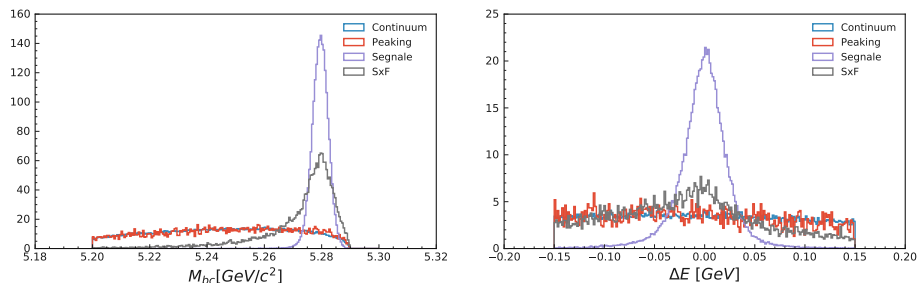


Figura 2.7: Distribuzioni normalizzate alla stessa area degli eventi del canale $B^0 \rightarrow \eta'(\rightarrow \eta(\rightarrow \gamma\gamma)\pi^+\pi^-)K_S^0$ per le variabili M_{bc} e ΔE .

Per le due variabili, M_{bc} e ΔE , le distribuzioni in figura 2.7 presentano un segnale con una forma molto precisa e ben localizzata. Sono infatti variabili utilizzate per misurare la quantità di segnale (e di fondo) nei dati, tramite un *fit*, e per questo motivo non sono state oggetto di ottimizzazione.

Capitolo 3

Applicazione al canale

$$B^0 \rightarrow \eta'(\rightarrow \eta(\rightarrow \gamma\gamma)\pi^+\pi^-)K_S^0$$

Al fine d'illustrare l'efficacia dell'algorithm, saranno ora presentati i risultati dell'ottimizzazione sul primo canale, per le classi di fondo elencate. Le distribuzioni, mostrate in figura 3.1, di dati simulati e dati sperimentali saranno utilizzate per confrontare i risultati delle selezioni sui *dataframe*.

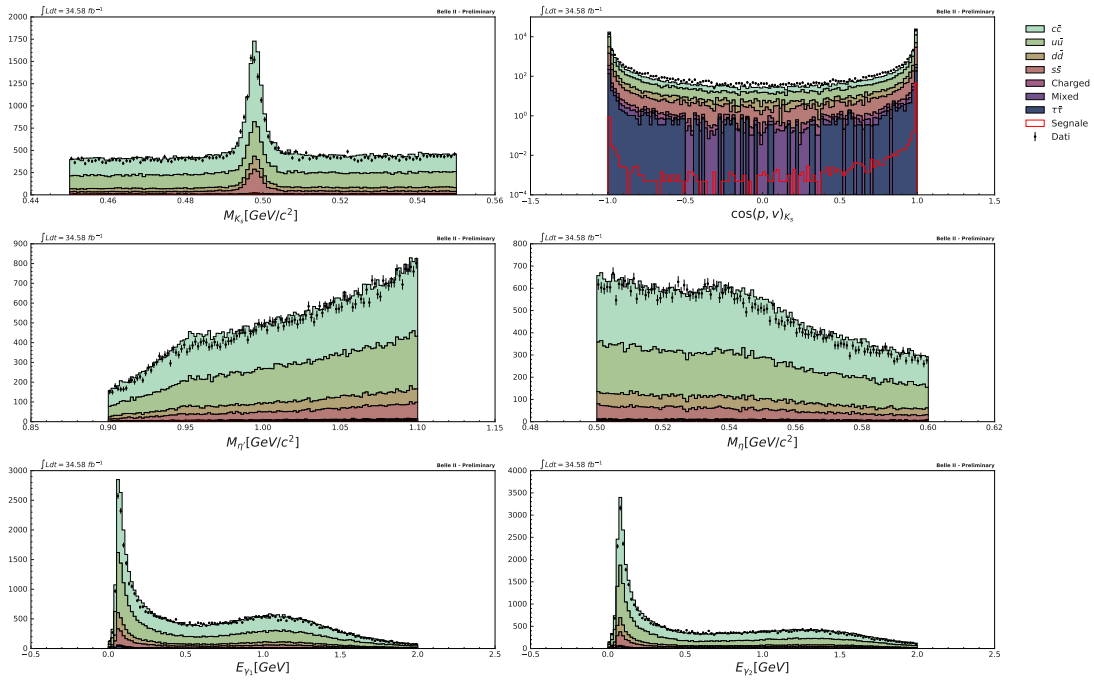


Figura 3.1: Distribuzione normalizzata a 34.6 fb^{-1} degli eventi per le variabili associate al canale, prima di qualsiasi selezione.

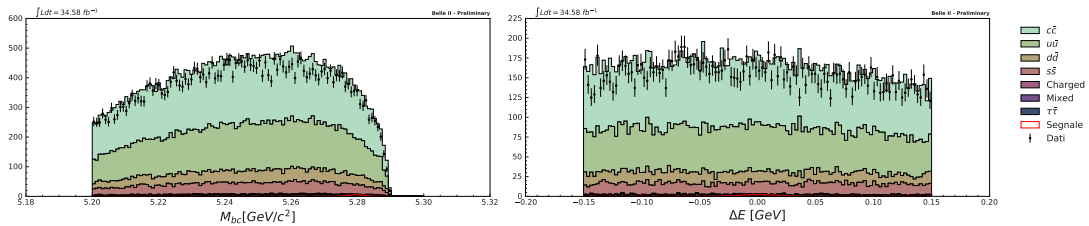


Figura 3.2: Distribuzione normalizzata a 34.6 fb^{-1} degli eventi per le variabili di estrazione del segnale, prima di qualsiasi selezione.

Prima delle selezioni il segnale è difatti sostanzialmente invisibile, mentre il contributo di *Continuum* risulta essere la componente principale del *dataframe*. Si nota inoltre che le distribuzioni per gli eventi

simulati sono molto simili a quelle per gli eventi sperimentali, sia come forma che come normalizzazione, e possono essere quindi considerate un buon modello per ottimizzare le selezioni

Un discorso analogo vale per le variabili di riferimento per l'estrazione del numero di eventi di segnale, in figura 3.2, che presentano un contributo di segnale troppo ridotto per qualsiasi analisi.

3.1 Risultati per il fondo *Continuum* + *Peaking*

Analizzando il fondo principale, composto sia dal contributo di *Peaking*, che da quello di *Continuum*, è possibile rimuovere la maggior parte degli eventi estranei a quelli d'interesse. In questo modo si può minimizzare il fondo sia nella regione di segnale, sia nella *sideband*.

Come si nota nelle selezioni in tabella 3.1, gli estremi risultano molto simili a quelli visti in sezione 2.2; in alcuni casi risultano più stringenti, in altri meno. È possibile studiare il loro effetto verificandone l'efficienza per segnale e fondo.

	Valore minimo	Valore massimo	
M_{K_s}	0.4927	0.5050	GeV/c ²
$\cos(p, v)_{K_s}$	0.9999	1.000	
$M_{\eta'}$	0.9236	0.9720	GeV/c ²
M_{η}	0.5206	0.5642	GeV/c ²
E_{γ}	0.2212	2.027	GeV

Tabella 3.1: Miglior insieme di selezioni ottenibile avendo come fondo la combinazione di *Continuum* e *Peaking* per gli eventi associati al canale $B^0 \rightarrow \eta'(\rightarrow \eta(\rightarrow \gamma\gamma)\pi^+\pi^-)K_S^0$.

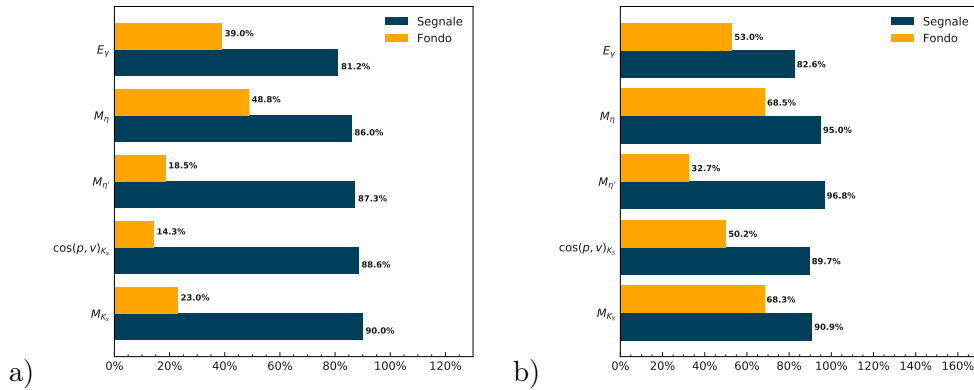


Figura 3.3: a) Efficienze delle selezioni presentate se applicate sulle singole variabili, in modo tale da essere le uniche selezioni eseguite; b) efficienze delle selezioni presentate se applicate sulla variabile in grafico dopo aver applicato tutte le selezioni sulle altre variabili. In entrambi i casi l'efficienza è calcolata dividendo separatamente il numero di eventi di segnale e fondo totale dopo la selezione per il numero di eventi di segnale e fondo totale presenti prima della selezione.

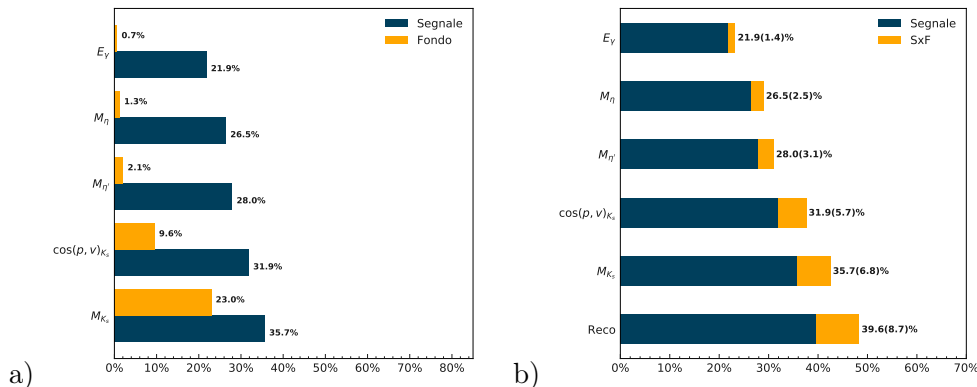


Figura 3.4: a) Efficienza delle selezioni presentate se applicate in maniera sequenziale su tutte le variabili, in modo tale che tutte le selezioni siano eseguite; b) efficienza delle selezioni presentate su segnale e *SxF*, se applicate in maniera sequenziale su tutte le variabili, in modo tale che tutte le selezioni siano eseguite. Le efficienze di a) sono state ottenute dividendo rispettivamente il numero di eventi di fondo totale dopo la selezione per quelli che erano presenti senza selezioni, e il numero di eventi di segnale dopo la selezione per il numero totale di decadimenti $B^0 \rightarrow \eta'(\rightarrow \eta(\rightarrow \gamma\gamma)\pi^+\pi^-)K_S^0$ simulati. In modo analogo sono state calcolate anche per b) tenendo conto solamente del numero totale di decadimenti $B^0 \rightarrow \eta'(\rightarrow \eta(\rightarrow \gamma\gamma)\pi^+\pi^-)K_S^0$ simulati. Per Reco si intende inoltre lo stato che si trova prima dell'applicazione di qualsiasi selezione.

Come si nota dalla figura 3.4, l'effetto di selezione è ottimo, infatti il fondo risulta essere lo 0.7% di quello iniziale, mentre il segnale invece passa da essere dato dal 39% degli eventi di decadimento in studio a circa la metà; si tratta di un risultato accettabile, dato che comunque la riduzione del fondo è molto buona. L'applicazione delle singole selezioni, in figura 3.3, risulta equilibrata sia se sono imposte come prime, che come ultime, si nota però che le selezioni sui valori delle masse risultano essere le più efficaci. L'altra selezione importante è quella sulla variabile $\cos(p, v)_{K_s}$, che permette di rimuovere il fondo accumulato per valori inferiori a 1. Per quanto riguarda il contributo di SxF , si nota un andamento molto simile a quello del fondo globale.

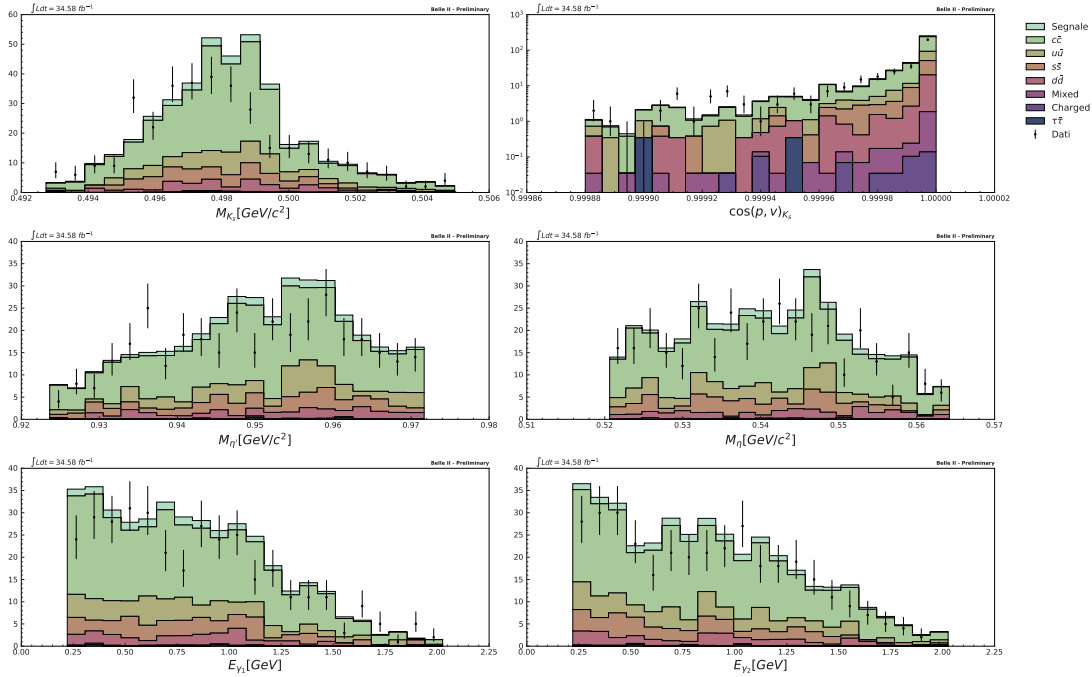


Figura 3.5: Distribuzione normalizzata a $34.6fb^{-1}$ degli eventi per le variabili associate al canale, dopo le selezioni.

Dalla distribuzione delle variabili, in figura 3.5, si può notare come il segnale abbia acquisito una rilevanza maggiore rispetto allo stato iniziale. Si può inoltre verificare come il risultato delle selezioni sulle due distribuzioni (simulata e non) sia molto simile.

Zona	Segnale	Continuum	Peaking	SxF	Totale	Eventi rilevati
SR	20.4 ± 0.1	11 ± 2	2.1 ± 0.3	0.82 ± 0.02	35 ± 2	40 ± 6
SB	0.40 ± 0.01	368 ± 11	1.0 ± 0.2	0.47 ± 0.01	370 ± 11	306 ± 17
Totale	20.8 ± 0.1	380 ± 11	3.1 ± 0.3	1.29 ± 0.02	405 ± 11	346 ± 18

Tabella 3.2: Numero di eventi attesi e rilevati per la selezione in discussione con una luminosità integrata di $34.6 fb^{-1}$.

Come si nota dalla tabella 3.2, la selezione restituisce un buon risultato, considerando l'assenza di selezioni sulle variabili di *Continuum Suppression*. Nella regione di segnale rimane un numero di eventi di *Continuum* ancora molto alto e i contributi di *Peaking* e di SxF risultano anch'essi non trascurabili. Nella *sideband* invece è ancora molto presente il contributo di *Continuum*, mentre gli altri due risultano limitati.

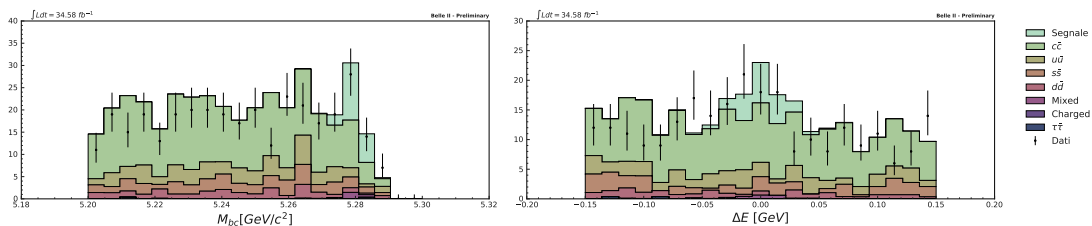


Figura 3.6: Distribuzione normalizzata a $34.6fb^{-1}$ degli eventi per le variabili di estrazione del segnale, dopo le selezioni.

Dalla figura 3.6 si vede come il picco di eventi di segnale risulti molto più evidente sul fondo, pur essendo un risultato migliorabile attraverso selezioni sulle variabili di *Continuum Suppression*.

3.2 Risultati per il fondo *Continuum* con ottimizzazione sulle variabili di *Continuum Suppression*

L'ottimizzazione è avvenuta massimizzando la figura di merito rispetto alla sola componente di *Continuum*, che risulta rappresentare il contributo principale da ridurre. L'aggiunta delle due variabili di *Continuum Suppression* permette di ridurre ulteriormente il peso del fondo, con la possibilità di ottimizzare la selezione a seconda delle correlazioni con le altre variabili. Le selezioni così trovate sono presentate in tabella 3.3.

	Valore minimo	Valore massimo	
R_2	0.01593	0.4119	
$\cos(\theta_B - \theta_O)$	$1.000 \cdot 10^{-5}$	0.5691	
M_{K_S}	0.4918	0.5057	GeV/c ²
$\cos(p, v)_{K_S}$	0.9999	1.000	
$M_{\eta'}$	0.9244	0.9732	GeV/c ²
M_{η}	0.5206	0.5661	GeV/c ²
E_{γ}	0.2192	1.913	GeV

Tabella 3.3: Miglior insieme di selezioni ottenibile avendo come fondo il solo *Continuum* per gli eventi associati al canale $B^0 \rightarrow \eta'(\rightarrow \eta(\rightarrow \gamma\gamma)\pi^+\pi^-)K_S^0$.

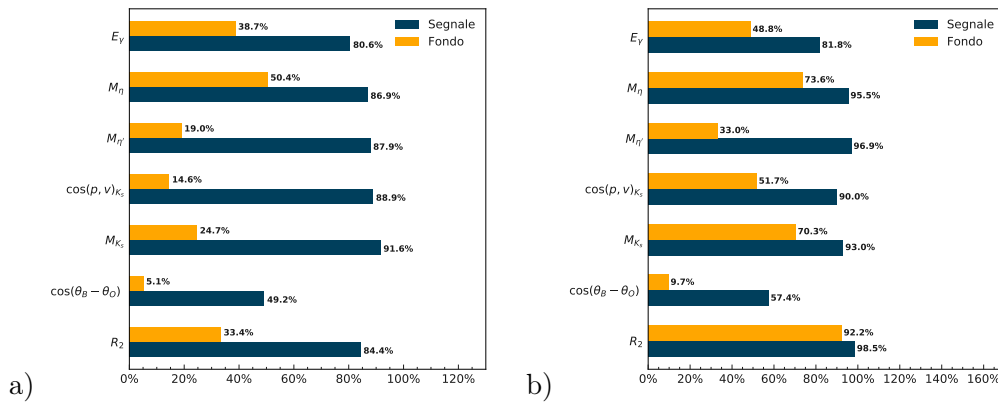


Figura 3.7: a) Efficienze delle selezioni presentate se applicate sulle singole variabili, in modo tale da essere le uniche selezioni eseguite; b) efficienze delle selezioni presentate se applicate sulla variabile in grafico dopo aver applicato tutte le selezioni sulle altre variabili. In entrambi i casi l'efficienza è calcolata dividendo separatamente il numero di eventi di segnale e fondo totale dopo la selezione per il numero di eventi di segnale e fondo totale presenti prima della selezione.

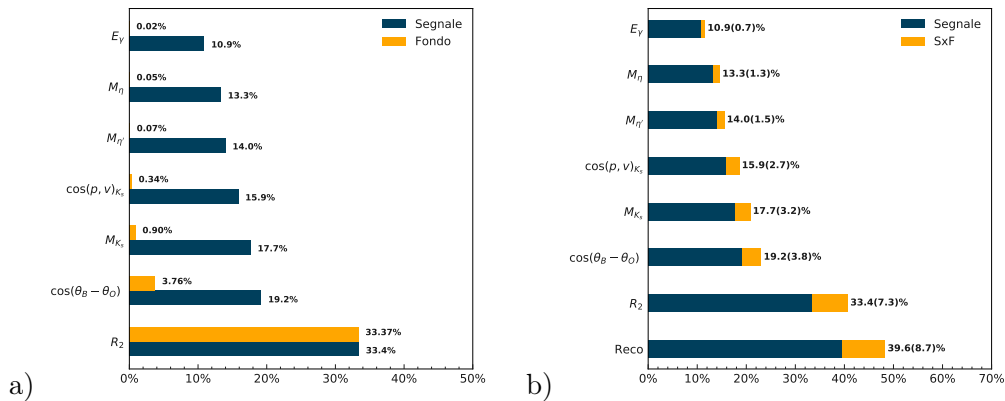


Figura 3.8: a) Efficienza delle selezioni presentate se applicate in maniera sequenziale su tutte le variabili, in modo tale che tutte le selezioni siano eseguite; b) efficienza delle selezioni presentate su segnale e SxF , se applicate in maniera sequenziale su tutte le variabili, in modo tale che tutte le selezioni siano eseguite. Le efficienze di a) sono state ottenute dividendo rispettivamente il numero di eventi di fondo totale dopo la selezione per quelli che erano presenti senza selezioni, e il numero di eventi di segnale dopo la selezione per il numero totale di decadimenti $B^0 \rightarrow \eta'(\rightarrow \eta(\rightarrow \gamma\gamma)\pi^+\pi^-)K_S^0$ simulati. In modo analogo sono state calcolate anche per b) tenendo conto solamente del numero totale di decadimenti $B^0 \rightarrow \eta'(\rightarrow \eta(\rightarrow \gamma\gamma)\pi^+\pi^-)K_S^0$ simulati. Per Reco si intende inoltre lo stato che si trova prima dell'applicazione di qualsiasi selezione.

In figura 3.8 si può notare come l'effetto della selezione sulle variabili di *Continuum Suppression* sia estremamente utile al fine di ridurre il fondo. Dalla figura 3.7 si nota inoltre come la selezione sulla variabile $\cos(\theta_B - \theta_O)$ risulti essere quella con impatto maggiore, mentre quella sulla variabile R_2 sembra essere invece correlata con le altre; la sua applicazione infatti ha un effetto molto ridotto sul fondo e inoltre, se applicata dopo le altre selezioni, risulta praticamente trascurabile. Il SxF è anch'esso estremamente ridotto, mentre il segnale passa dal 39.6% degli eventi al 10.9%, a causa della selezione molto restrittiva sui domini.

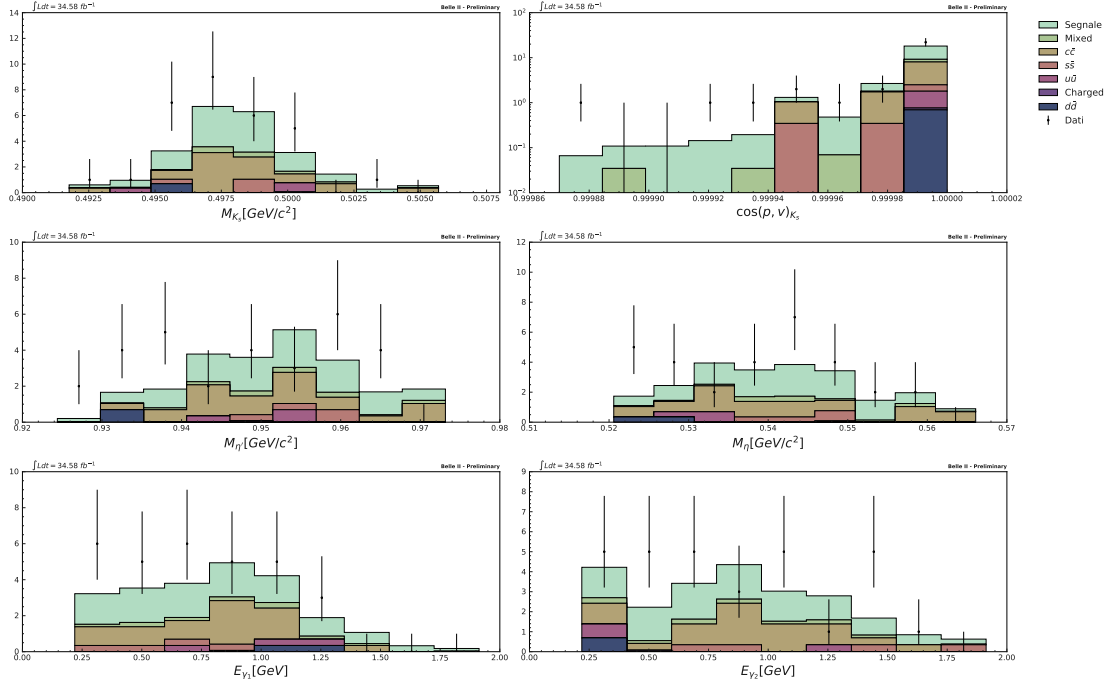


Figura 3.9: Distribuzione normalizzata a $34.6fb^{-1}$ degli eventi per le variabili associate al canale, dopo le selezioni.

Dalla figura 3.9 si nota come il segnale sia stavolta molto più evidente del fondo, con una buona corrispondenza tra il numero di eventi selezionati presente in simulazione e quelli del *dataframe* proveniente dai dati sperimentali.

Zona	Segnale	<i>Continuum</i>	<i>Peaking</i>	<i>SxF</i>	Totale	Eventi rilevati
SR	10.20 ± 0.07	0.3 ± 0.3	1.1 ± 0.2	0.42 ± 0.01	12.1 ± 0.4	14 ± 4
SB	0.20 ± 0.01	10 ± 2	0.3 ± 0.1	0.20 ± 0.01	11 ± 2	16 ± 4
Totale	10.40 ± 0.07	11 ± 2	1.5 ± 0.2	0.63 ± 0.02	23 ± 2	30 ± 5

Tabella 3.4: Numero di eventi attesi e rilevati per la selezione in discussione con una luminosità integrata di $34.6fb^{-1}$.

Dalla tabella 3.4 si vede come nella SR il segnale sia prevalente, mentre nella *sideband* il fondo risulti molto ridotto, e globalmente confrontabile con il numero di eventi di segnale totali. I dati raccolti da Belle II presentano un numero di eventi stavolta maggiore di quelli previsti dalla simulazione, ma comunque compatibile entro gli errori sperimentali. La concordanza con i valori della distribuzione Montecarlo non è essenziale, dato che il fit finale avverrà sui valori sperimentali, ma risulta essere un'ottima conferma del lavoro fatto.

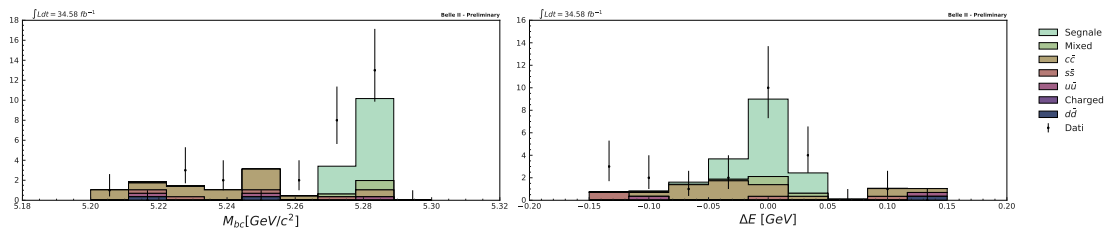


Figura 3.10: Distribuzione normalizzata a $34.6fb^{-1}$ degli eventi per le variabili di estrazione del segnale, dopo le selezioni.

Il picco di segnale in figura 3.10 è evidente e risulta molto più importante della componente di fondo, ma il numero di eventi offre una statistica limitata.

3.3 Risultati per il fondo *Continuum* + *Peaking* in SR con preselezione sulle variabili di *Continuum Suppression*

Ottimizzando il segnale contro il fondo nella regione di segnale è possibile massimizzare il segnale estraibile e nel frattempo conservare una componente di fondo nella *sideband*. L'applicazione della selezione preliminare sulle variabili di *Continuum Suppression* permette inoltre di lavorare con un fondo già notevolmente ridotto. Le selezioni in tabella 3.5 risultano comunque simili a quelle presentate nella precedente sezione.

	Valore minimo	Valore massimo	
M_{K_s}	0.4914	0.5058	GeV/c^2
$\cos(p, v)_{K_s}$	0.9999	1.000	
$M_{\eta'}$	0.9244	0.9732	GeV/c^2
M_{η}	0.5206	0.5642	GeV/c^2
E_{γ}	0.2192	1.913	GeV

Tabella 3.5: Miglior insieme di selezioni ottenibile avendo come fondo la combinazione di *Continuum* e *Peaking* nella regione di segnale per gli eventi associati al canale $B^0 \rightarrow \eta'(\rightarrow \eta(\rightarrow \gamma\gamma)\pi^+\pi^-)K_S^0$.

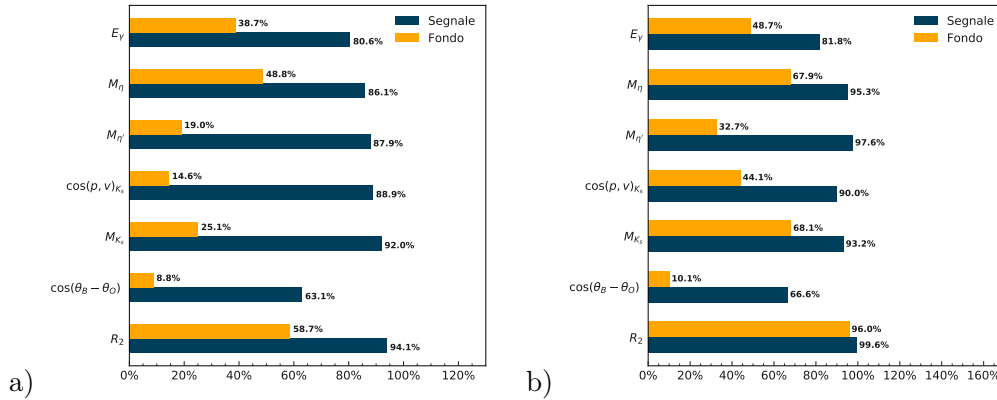


Figura 3.11: a) Efficienze delle selezioni presentate se applicate sulle singole variabili, in modo tale da essere le uniche selezioni eseguite; b) efficienze delle selezioni presentate se applicate sulla variabile in grafico dopo aver applicato tutte le selezioni sulle altre variabili. In entrambi i casi l'efficienza è calcolata dividendo separatamente il numero di eventi di segnale e fondo totale dopo la selezione per il numero di eventi di segnale e fondo totale presenti prima della selezione.

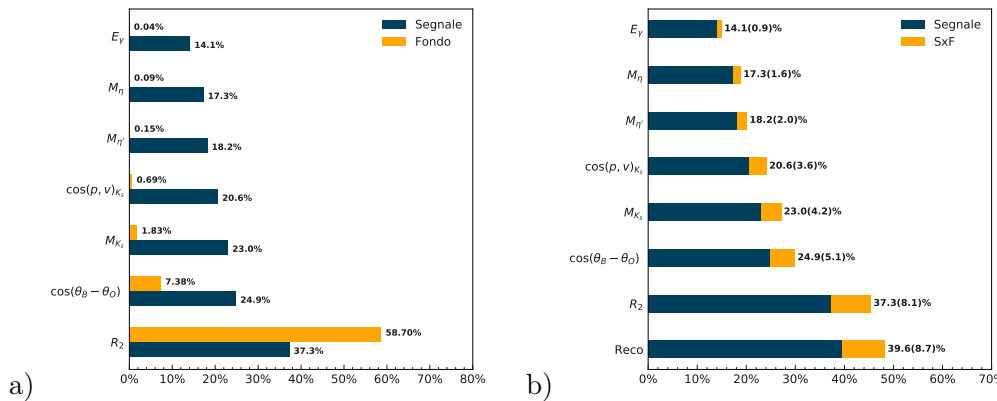


Figura 3.12: a) Efficienza delle selezioni presentate se applicate in maniera sequenziale su tutte le variabili, in modo tale che tutte le selezioni siano eseguite; b) efficienza delle selezioni presentate su segnale e *SxF*, se applicate in maniera sequenziale su tutte le variabili, in modo tale che tutte le selezioni siano eseguite. Le efficienze di a) sono state ottenute dividendo rispettivamente il numero di eventi di fondo totale dopo la selezione per quelli che erano presenti senza selezioni, e il numero di eventi di segnale dopo la selezione per il numero totale di decadimenti $B^0 \rightarrow \eta'(\rightarrow \eta(\rightarrow \gamma\gamma)\pi^+\pi^-)K_S^0$ simulati. In modo analogo sono state calcolate anche per b) tenendo conto solamente del numero totale di decadimenti $B^0 \rightarrow \eta'(\rightarrow \eta(\rightarrow \gamma\gamma)\pi^+\pi^-)K_S^0$ simulati. Per Reco si intende inoltre lo stato che si trova prima dell'applicazione di qualsiasi selezione.

Nonostante la selezione fosse ottimizzata solo per una regione specifica, i risultati sono ottimi, con un fondo ridottissimo e un numero di eventi rimanenti più alto del caso precedente. Anche qui risultano

fondamentali le selezioni sulle variabili di *Continuum Suppression*, per le quali si nota nuovamente, in figura 3.11, la sostanziale correlazione della selezione sulla variabile R_2 con le altre e l'efficacia della selezione sulla variabile $\cos(\theta_B - \theta_O)$. Per quanto riguarda il SxF , si nota in figura 3.12, come il risultato sia molto buono, con un contributo finale estremamente ridotto.

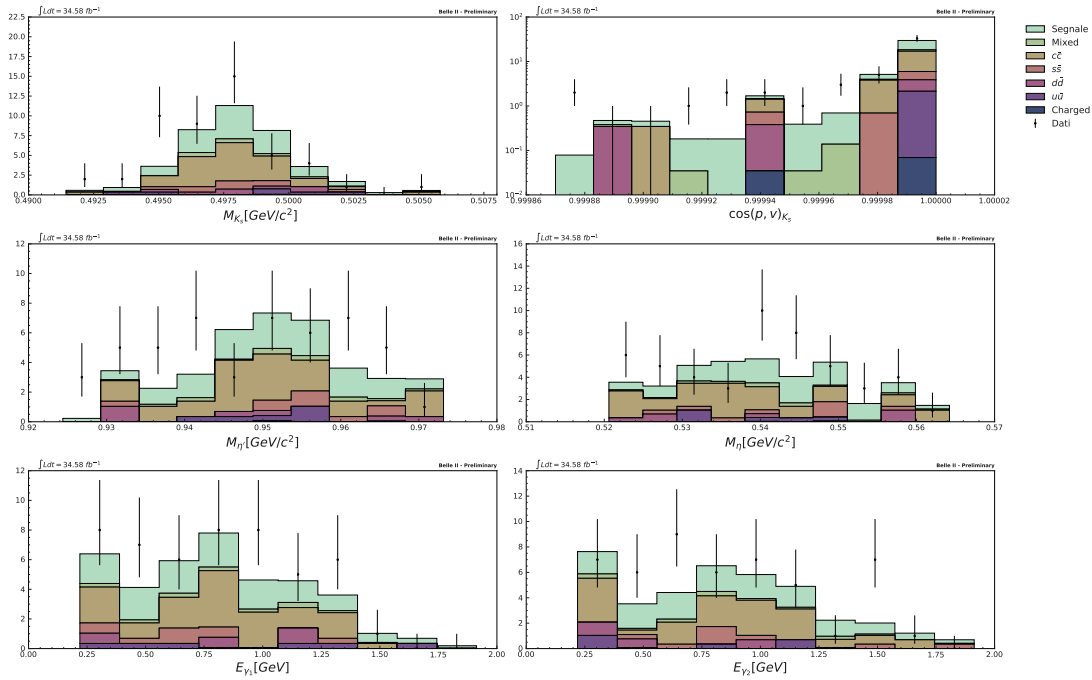


Figura 3.13: Distribuzione normalizzata a 34.6fb^{-1} degli eventi per le variabili associate al canale, dopo le selezioni.

In modo analogo alla sezione precedente, le distribuzioni in figura 3.13 presentano un effetto visibile, con un segnale che risulta significativo. Si nota inoltre una buona corrispondenza con i dati sperimentali.

Zona	Segnale	<i>Continuum</i>	<i>Peaking</i>	<i>SxF</i>	Totale	Eventi rilevati
SR	13.21 ± 0.08	1.0 ± 0.6	1.4 ± 0.2	0.55 ± 0.02	16.2 ± 0.6	19 ± 4
SB	0.27 ± 0.01	22 ± 3	0.5 ± 0.1	0.27 ± 0.01	23 ± 3	30 ± 5
Totale	13.47 ± 0.08	23 ± 3	1.9 ± 0.3	0.81 ± 0.02	39 ± 3	49 ± 7

Tabella 3.6: Numero di eventi attesi e rilevati per la selezione in discussione con una luminosità integrata di 34.6fb^{-1} .

Il segnale in tabella 3.6 risulta più numeroso della selezione precedente, con i contributi di *Continuum* e di *Peaking* più significativi, ma comunque molto ridotti. Nella *sideband* invece si trovano ancora molti eventi di fondo, nonostante l'applicazione delle selezioni sulle variabili di *Continuum Suppression*.

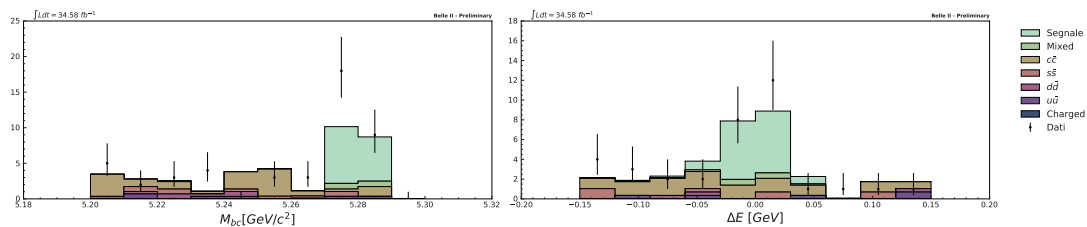


Figura 3.14: Distribuzione normalizzata a 34.6fb^{-1} degli eventi per le variabili di estrazione del segnale, dopo le selezioni.

Il picco di segnale in figura 3.14 risulta ancora una volta molto alto, ma il numero di eventi rispetto al caso precedente sembra più rilevante. La presenza di un leggero fondo permette inoltre di poter acquisire informazioni sulla sua struttura.

Capitolo 4

Applicazione agli altri canali

L'analisi effettuata nella regione di segnale con la preselezione sulle variabili di *Continuum Suppression* ha permesso di ottenere ottimi risultati a livello di conservazione del segnale e rimozione del fondo. Applicandola agli altri canali è possibile verificare ulteriormente la sua efficacia.

4.1 Risultati per il canale $B^\pm \rightarrow \eta'(\rightarrow \eta(\rightarrow \gamma\gamma)\pi^+\pi^-)K^\pm$

Per questo canale le variabili sono solo quattro, fatto che impedisce all' algoritmo di ottenere risultati buoni come i precedenti. Si introduce inoltre la variabile $\cos\theta_K$, che, come già visto, non risulta discriminare in modo ottimale le distribuzioni di fondo e di segnale. Le selezioni in tabella 4.1 risultano poco stringenti per questi motivi.

	Valore minimo	Valore massimo	
$\cos\theta_K$	-0.7287	0.9305	
$M_{\eta'}$	0.9084	0.9822	GeV/c ²
M_η	0.5092	0.5855	GeV/c ²
E_γ	0.05679	2.540	GeV

Tabella 4.1: Miglior insieme di selezioni ottenibile avendo come fondo la combinazione di *Continuum* e *Peaking* nella regione di segnale per gli eventi associati al canale $B^\pm \rightarrow \eta'(\rightarrow \eta(\rightarrow \gamma\gamma)\pi^+\pi^-)K^\pm$.

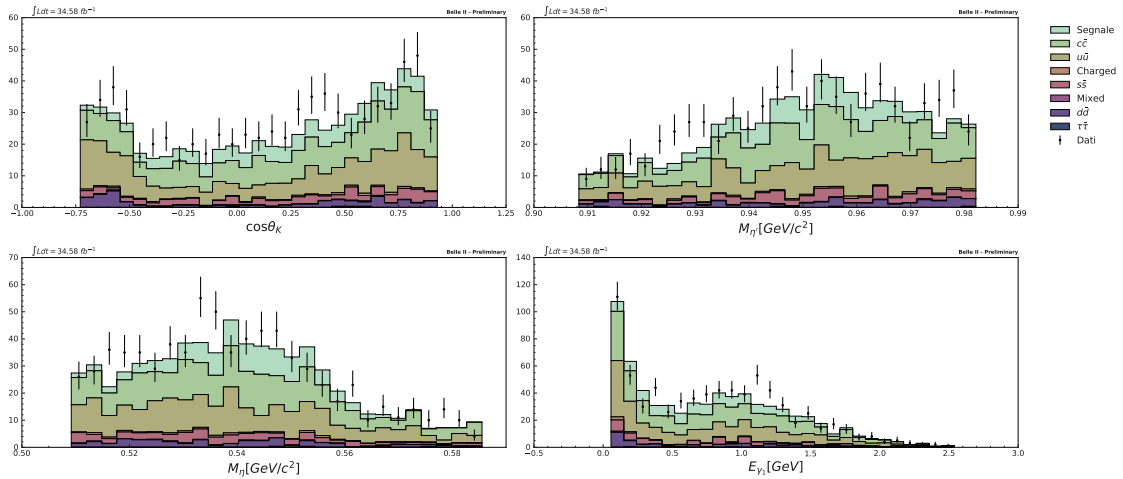


Figura 4.1: Distribuzione normalizzata a $34.6 fb^{-1}$ degli eventi per le variabili associate al canale, dopo le selezioni.

Zona	Segnale	<i>Continuum</i>	<i>Peaking</i>	<i>SxF</i>	Totale	Eventi rilevati
SR	101.8 ± 0.5	19 ± 3	1.3 ± 0.2	5.1 ± 0.1	128 ± 3	108 ± 10
SB	3.8 ± 0.1	539 ± 14	13.8 ± 0.7	6.3 ± 0.1	563 ± 14	633 ± 25
Totale	105.6 ± 0.5	558 ± 14	15.1 ± 0.7	11.4 ± 0.2	690 ± 14	741 ± 27

Tabella 4.2: Numero di eventi attesi e rilevati per la selezione in discussione con una luminosità integrata di $34.6 fb^{-1}$.

Dalla figura 4.1 e dalla tabella 4.2 si nota come nel complesso il segnale sia visibile, soprattutto nella SR, dove risulta molto più alto delle componenti di *Continuum* e *Peaking*, e con un contributo di *SxF* molto ridotto. Si nota inoltre una buona corrispondenza tra la distribuzione di eventi simulati e sperimentali.

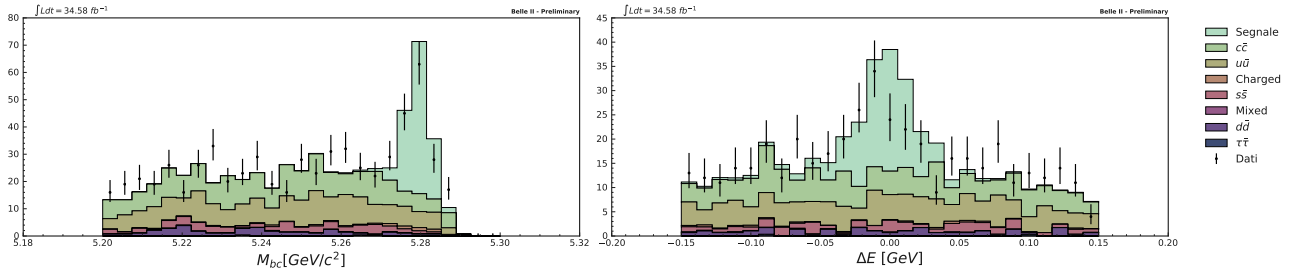


Figura 4.2: Distribuzione normalizzata a $34.6fb^{-1}$ degli eventi per le variabili di estrazione del segnale, dopo le selezioni.

Il picco in figura 4.2 è molto visibile sopra una limitata quantità di fondo per entrambe le variabili. Con un *fit* adeguato è possibile dunque estrarre il numero di eventi di segnale per il decadimento in studio.

4.2 Risultati per canale $B^0 \rightarrow \eta'(\rightarrow \rho\gamma)K_S^0$

In questo caso si hanno sei variabili di ottimizzazione, che possono essere usate per ridurre il fondo in maniera ancora più forte. Questo in realtà, come si vedrà, non avviene; le variabili $\cos\theta_K$ e M_ρ non separano bene segnale e fondo, e sono perciò poco utili allo scopo. Le selezioni in tabella 4.3 per queste due variabili saranno molto poco stringenti per questo motivo e per la sostanziale correlazione con le altre selezioni.

	Valore minimo	Valore massimo	
M_{K_S}	0.4890	0.5087	GeV/c^2
$\cos(p, v)_{K_S}$	0.9994	1.000	
E_γ	0.2418	1.580	GeV
$\cos\theta_\gamma$	-0.8143	0.9431	
$M_{\eta'}$	0.9407	0.9687	GeV/c^2
M_ρ	0.5671	0.8904	GeV/c^2

Tabella 4.3: Miglior insieme di selezioni ottenibile avendo come fondo la combinazione di *Continuum* e *Peaking* nella regione di segnale per gli eventi associati al canale $B^0 \rightarrow \eta'(\rightarrow \rho\gamma)K_S^0$.

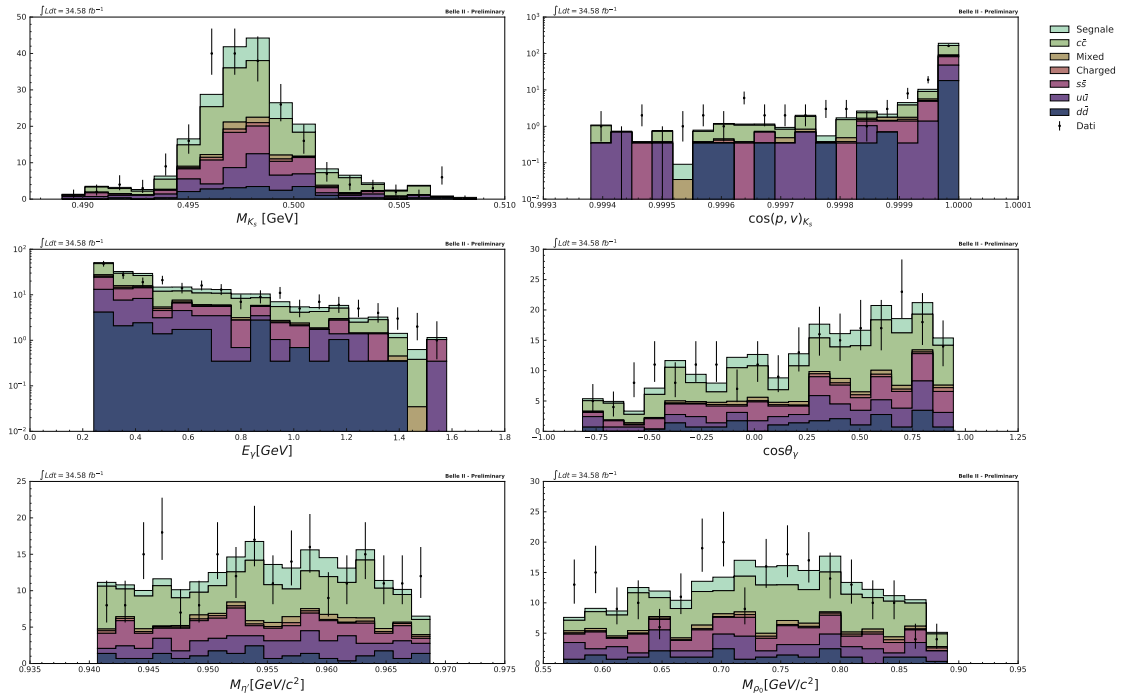
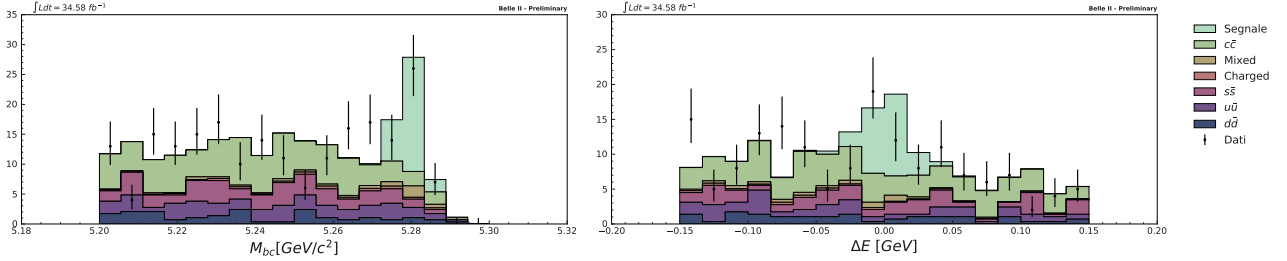


Figura 4.3: Distribuzione normalizzata a $34.6fb^{-1}$ degli eventi per le variabili associate al canale, dopo le selezioni.

Zona	Segnale	<i>Continuum</i>	<i>Peaking</i>	<i>SxF</i>	Totale	Eventi rilevati
SR	27.2 ± 0.2	5 ± 1	2.7 ± 0.3	0.45 ± 0.02	35 ± 1	184 ± 14
SB	0.33 ± 0.02	177 ± 8	9.4 ± 0.6	0.52 ± 0.03	187 ± 8	34 ± 6
Totale	27.5 ± 0.2	182 ± 8	12.1 ± 0.7	0.97 ± 0.03	223 ± 8	218 ± 15

 Tabella 4.4: Numero di eventi attesi e rilevati per la selezione in discussione con una luminosità integrata di 34.6 fb^{-1} .

Il risultato presentato in figura 4.3 e in tabella 4.4 è molto buono, con un contributo di *Continuum* e *Peaking* estremamente ridotti nella regione di segnale.

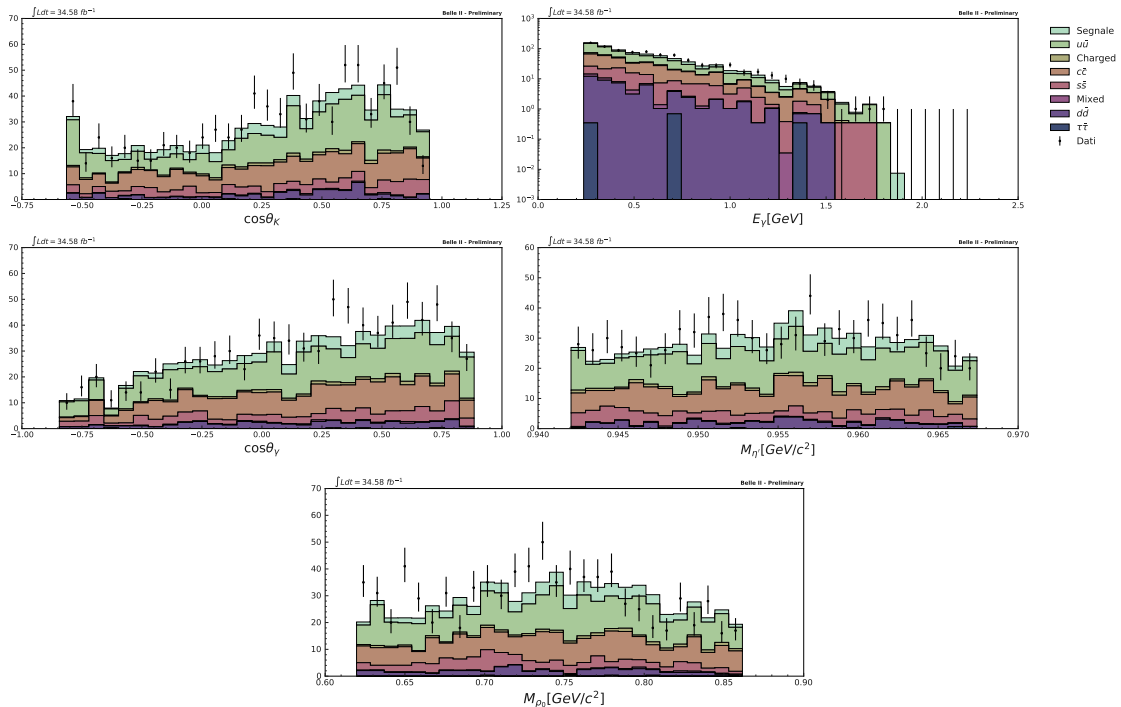

 Figura 4.4: Distribuzione normalizzata a 34.6 fb^{-1} degli eventi per le variabili di estrazione del segnale, dopo le selezioni.

Il picco in figura 4.4 risulta molto visibile sul fondo, per entrambe le variabili, con una buona rilevanza.

4.3 Risultati per il canale $B^\pm \rightarrow \eta'(\rightarrow \rho\gamma)K^\pm$

Il numero di variabili ritorna a essere cinque, ma la presenza delle variabili M_ρ , $\cos\theta_K$ e $\cos\theta_\gamma$ non permette una selezione accurata come quella del primo canale. Come nella sezione precedente, le selezioni in tabella 4.5 risultano nuovamente poco stringenti, a causa della loro correlazione con le selezioni sulle restanti variabili.

	Valore minimo	Valore massimo	
$\cos\theta_K$	-0.5642	0.9479	
E_γ	0.2373	2.271	GeV
$\cos\theta_\gamma$	-0.8440	0.8841	
$M_{\eta'}$	0.9421	0.9674	GeV/c^2
M_ρ	0.6197	0.8616	GeV/c^2

 Tabella 4.5: Miglior insieme di selezioni ottenibile avendo come fondo la combinazione di *Continuum* e *Peaking* nella regione di segnale per gli eventi associati al canale $B^\pm \rightarrow \eta'(\rightarrow \rho\gamma)K^\pm$.

 Figura 4.5: Distribuzione normalizzata a 34.6 fb^{-1} degli eventi per le variabili associate al canale, dopo le selezioni.

Zona	Segnale	<i>Continuum</i>	<i>Peaking</i>	SxF	Totale	Eventi rilevati
SR	83.6 ± 0.6	25 ± 3	2.5 ± 0.3	0.74 ± 0.05	112 ± 3	123 ± 11
SB	0.94 ± 0.06	651 ± 15	28 ± 1	1.33 ± 0.07	681 ± 15	714 ± 27
Totale	84.6 ± 0.6	676 ± 15	30 ± 1	2.07 ± 0.09	792 ± 15	837 ± 29

Tabella 4.6: Numero di eventi attesi e rilevati per la selezione in discussione con una luminosità integrata di 34.6 fb^{-1} .

Dai risultati presentati in figura 4.5 e in tabella 4.6 si nota, come previsto, che il segnale risulta essere presente, ma in quantità molto ridotta rispetto al fondo totale. Se ci si concentra invece nella regione di segnale la proporzione cambia decisamente e presenta una prevalenza del segnale. È inoltre presente una buona concordanza tra dati sperimentali e simulati.

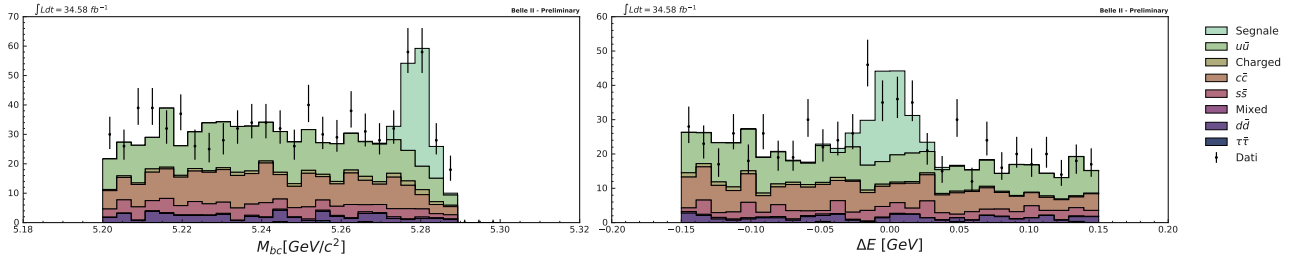


Figura 4.6: Distribuzione normalizzata a 34.6 fb^{-1} degli eventi per le variabili di estrazione del segnale, dopo le selezioni.

Come nei precedenti casi, il picco di segnale in figura 4.6 è ben visibile per entrambe le variabili, con un fondo sufficientemente ridotto alla sua base.

Conclusioni

L'algoritmo elaborato ha permesso d'individuare dei possibili insiemi di selezioni per il *dataframe* in studio. Al variare del fondo è stato possibile infatti ottenere una serie di risultati: per il fondo composto dai contributi di *Continuum* e *Peaking* si è trovato un buon insieme di selezioni che, anche senza ricorrere alle variabili di *Continuum Suppression*, è in grado di rimuovere buona parte del fondo ed evidenziare il segnale; per il fondo composto dal solo *Continuum*, con l'ottimizzazione sulle variabili di *Continuum Suppression*, si è visto come fosse possibile ridurre il fondo a poco più di una decina di eventi, conservando comunque un discreto segnale; per la selezione contro il fondo presente nella SR è stato possibile invece ottenere un insieme di selezioni che permettono di conservare una quantità maggiore di segnale, rendendolo comunque la componente principale nella regione di segnale, e contemporaneamente avere una discreta quantità di fondo nella *sideband*, garantendo la possibilità di stimare l'influenza dei processi estranei a quello d'interesse. In questo modo è possibile scegliere, a seconda di ciò che si vuole ottenere, un possibile approccio al problema.

Utilizzando il metodo più equilibrato (ottimizzando il segnale contro il fondo completo nella SR e applicando le preselezioni sulle variabili di *Continuum Suppression*) è possibile ottenere ottimi risultati anche per gli altri canali. È stato mostrato infatti che si possono ottenere picchi di segnale ben visibili, con fondi abbastanza circoscritti nella *sideband*.

L'ottimizzazione delle selezioni è necessaria per l'estrazione del numero di eventi di segnale dal *dataframe* di dati sperimentali, tramite il *fit* delle variabili M_{bc} e ΔE , e di conseguenza per la misura delle BR per i decadimenti studiati. Questo tipo di analisi è stata esclusa dal lavoro di tesi, ma tuttavia ci si aspetta, vista l'ottima concordanza tra le misure sperimentali e gli eventi simulati, un sostanziale accordo con le misure di BR presentate e utilizzate nella simulazione Montecarlo.

Queste selezioni sono solo il punto di partenza per ulteriori sviluppi in questo campo; tra le possibili opzioni, si possono citare tecniche di *Machine Learning* quali MVA (*MultiVariate data Analysis*) [13], che possono migliorare la capacità di selezione tra segnale e fondo nel *dataframe*. L'ottimizzazione del segnale renderà quindi possibile esaminare con più dettaglio questi decadimenti rari, permettendo di cercare eventuali discrepanze con il modello standard.

Bibliografia

- [1] T. Abe *et al.*, “Belle II Technical Design Report,” *arXiv e-prints*, p. arXiv:1011.0352, Nov. 2010.
- [2] K. Akai, K. Furukawa, H. Koiso, and SuperKEKB Accelerator Team, “SuperKEKB collider,” *Nuclear Instruments and Methods in Physics Research A*, vol. 907, pp. 188–199, Nov. 2018.
- [3] Particle Data Group *et al.*, “Review of Particle Physics,” *Progress of Theoretical and Experimental Physics*, vol. 2020, 08 2020. 083C01.
- [4] W. Altmannshofer *et al.*, “The Belle II Physics Book,” *PTEP*, vol. 2019, no. 12, p. 123C01, 2019. [Erratum: PTEP 2020, 029201 (2020)].
- [5] F. Tenchini and J.-F. Krohn, “Decay chain reconstruction in Belle II,” BELLE2-POSTER-CONF-2018-005, Jul 2018.
- [6] B. Oberhof and I. Belle Collaboration, “Precise measurements of CP violation in B decays at Belle II,” in *XIV International Conference on Heavy Quarks and Leptons*, p. 4, June 2018.
- [7] B. Aubert *et al.*, “Observation of CP Violation in $B \rightarrow \eta' K^0$ Decays,” vol. 98, p. 031801, Jan. 2007.
- [8] T. Aushev *et al.*, “Physics at Super B Factory,” *arXiv e-prints*, p. arXiv:1002.5012, Feb. 2010.
- [9] The pandas development team, “pandas-dev/pandas: Pandas,” Feb. 2020.
- [10] S. Lacaprara, “Rediscovery of $B \rightarrow \eta' K$ in Belle II data.” BELLE2-NOTE-PH-2020-053, unpublished.
- [11] U. Tamponi, “A very exciting lecture about continuum suppression,” Belle II Starter Kit, 2018.
- [12] C. Cecchi, G. D. Nardo, M. Merola, E. Manoni, and A. Selce, “Approved plots: Study of r^2 distribution and b counting in early phase 3 data,” Jul 2019.
- [13] A. Hoecker *et al.*, “TMVA: Toolkit for Multivariate Data Analysis,” *PoS*, vol. ACAT, p. 040, 2007.