

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

Scuola di Scienze
Dipartimento di Fisica e Astronomia
Corso di Laurea in Fisica

**Studio di tecniche di estrazione del segnale
per barioni Λ_c^+ ricostruiti nell'esperimento
ALICE**

Relatore:
Prof. Andrea Alici

Presentata da:
Matteo Fusconi

Anno Accademico 2021/2022

Abstract

L'adronizzazione è il processo, descritto dalla QCD, attraverso il quale quark e gluoni liberi si combinano per formare stati con carica netta di colore neutra: gli adroni. Tale processo può avvenire per frammentazione, se il quark o il gluone originario si legano con altri partoni generati per la rottura di una stringa QCD nel vuoto, o per coalescenza, se quark e gluoni liberi si trovano in condizioni sufficientemente simili nello spazio delle fasi da aggregarsi in un adrone. Quest'ultimo processo è fortemente favorito in condizioni di Quark-Gluon Plasma (QGP), l'unica fase della materia in cui si trovano quark e gluoni liberi. L'esperimento ALICE a LHC ha tra i suoi obiettivi principali proprio lo studio della materia fortemente interagente e le proprietà del QGP attraverso collisioni di particelle ultrarelativistiche di vario tipo: protone-protone, protone-ione e ione-ione. Le recenti analisi dei dati raccolti ad ALICE dimostrano che la nostra comprensione dei fenomeni di adronizzazione dei sapori pesanti è ancora incompleta, perché le misure effettuate su collisioni pp, p-Pb e Pb-Pb non sono riproducibili da modelli teorici basati su altre tipologie di collisione come e^+e^- . In particolare, i risultati sembrano indicare che il principio di universalità, che assume che le funzioni di frammentazione di quark e gluoni siano indipendenti dal tipo di sistema interagente, non sia valido. Per questo motivo sono stati sviluppati nuovi modelli teorici e fenomenologici, capaci di riprodurre in modo più o meno accurato i dati sperimentali. Questi modelli differiscono tra di loro soprattutto a bassi valori di impulso trasverso p_T . L'analisi dati a basso p_T si rivela dunque di fondamentale importanza, in quanto permette di discriminare, tra i vari modelli, quelli che sono realmente in grado di riprodurre i dati sperimentali e quelli che non lo sono. Inoltre può fornire una conferma sperimentale dei fenomeni fisici su cui tale modello si basa.

In questa tesi è stato estratto il numero di barioni Λ_c^+ (*yield*) prodotto in collisioni pp a $\sqrt{s} = 13 TeV$, nel range di impulso trasverso $0 < p_T(\Lambda_c^+) < 1 GeV/c$. È stato fatto uso di una tecnica di machine learning che sfrutta un algoritmo di tipo Boosted Decision Trees (BDT) implementato dal pacchetto TMVA, al fine di identificare ed eliminare una grossa parte del fondo statistico e semplificare notevolmente l'analisi vera e propria. Il grado di attendibilità della misura è stata verificata eseguendo l'estrazione dello *yield* con due approcci diversi: il primo, modellando il fondo combinatoriale con una funzione analitica; successivamente con la creazione di un template statistico creato ad hoc con la tecnica delle *track rotations*.

Indice

Indice	4
1 L'adronizzazione dei quark pesanti	6
1.1 Il Modello Standard	6
1.2 La Cromodinamica Quantistica	8
1.3 Quark-Gluon Plasma	9
1.4 Collisioni tra ioni pesanti	10
1.5 Evoluzione del QGP	11
1.6 Adronizzazione di sapori pesanti in collisioni pp	13
1.7 Adronizzazione di sapori pesanti in collisioni A-A	17
1.8 Risultati sperimentali	18
1.8.1 Adroni charmati in collisioni Pb-Pb a $\sqrt{s_{NN}} = 5.02 TeV$	18
1.8.2 Adroni charmati in collisioni pp a $\sqrt{s} = 5.02 TeV$ e a $\sqrt{s} = 13 TeV$	19
2 L'esperimento ALICE	22
2.1 Il Large Hadron Collider	22
2.2 ALICE: <i>A Large Ion Collider Experiment</i>	24
2.2.1 Inner Tracking System (ITS)	26
2.2.2 Time-Projection Chamber (TPC)	27
2.2.3 Transition Radiation Detector (TRD)	28
2.2.4 Time-Of-Flight (TOF)	28
2.2.5 Rilevatori esterni	29
2.2.6 Forward Muon Spectrometer (FMS)	30
2.2.7 Rilevatori ad alta pseudorapidità	30
2.2.8 ALICE COsmic Ray DEtector (ACORDE)	32
3 Ricostruzione del barione Λ_c^+	34
3.1 Introduzione	34
3.2 TMVA	36
3.3 Boosted Decision Trees	37
3.4 Dati e Variabili in input	38

3.5	Training del metodo BDT	41
3.6	Testing e Grafici di Controllo	42
3.6.1	Test dell'overtraining	42
3.6.2	Curva di ROC	44
3.6.3	Ranking delle variabili	45
3.7	Metodo di estrazione del segnale con polinomiale di terzo grado	46
3.8	Metodo alternativo con template rotazionale	49
3.9	Risultati	53
	Conclusioni	54
	Bibliografia	56
	Ringraziamenti	58

Capitolo 1

L'adronizzazione dei quark pesanti

1.1 Il Modello Standard

Il Modello Standard (SM) [1] è l'insieme delle più recenti teorie che descrivono le interazioni fondamentali tra particelle elementari (esclusa l'interazione gravitazionale) e le particelle stesse, il cui schema è rappresentato in Figura 1.1.

Si trovano i sei diversi sapori di quark: up e down, charm e strange, top e bottom, organizzati in doppietti (generazioni) in cui il primo elemento è generalmente il più massivo ed ha una carica elettrica positiva di modulo uguale a due terzi di quella dell'elettrone, mentre il secondo ha una carica elettrica negativa di modulo uguale a un terzo di quella dell'elettrone. In seguito si trovano i sei leptoni: elettrone, muone, tauone, con i relativi neutrini. I primi tre hanno carica elettrica negativa ed unitaria, mentre i neutrini hanno carica elettrica nulla, ed una massa di diversi ordini di grandezza più piccola rispetto a quella del relativo leptone. L'origine di questa discrepanza è tuttora un tema di ricerca molto importante. Sono inoltre presenti nel modello anche le rispettive antiparticelle di ciascun quark e di ciascun leptone: esse sono particelle identiche alle originali in ogni caratteristica fisica ad eccezione delle cariche elettriche che risultano invertite. Sia quark che leptoni sono fermioni di spin $1/2$, ed in quanto tali sottostanno al principio di esclusione di Pauli e seguono la statistica di Fermi-Dirac.

In seguito si trovano le particelle mediatrici delle interazioni fondamentali: 8 gluoni, mediatori dell'interazione forte, il fotone, mediatore dell'interazione elettromagnetica, e i bosoni Z^0 e W^\pm , mediatori dell'interazione debole. Questi ultimi sono bosoni di spin 1 e seguono la statistica di Bose-Einstein.

É importante sottolineare che le precedenti non sono particelle nel senso classico, ma si fa sempre riferimento a campi quantizzati. Le teorie che compongono il Modello Standard sono teorie di campo quantizzato: la QCD (Cromodinamica Quantistica), che spiega l'interazione tra quark attraverso gli scambi di gluoni, e la teoria elettrodebole, la quale generalizza la QED (Elettrodinamica Quantistica), che spiegava la forza elettromagneti-

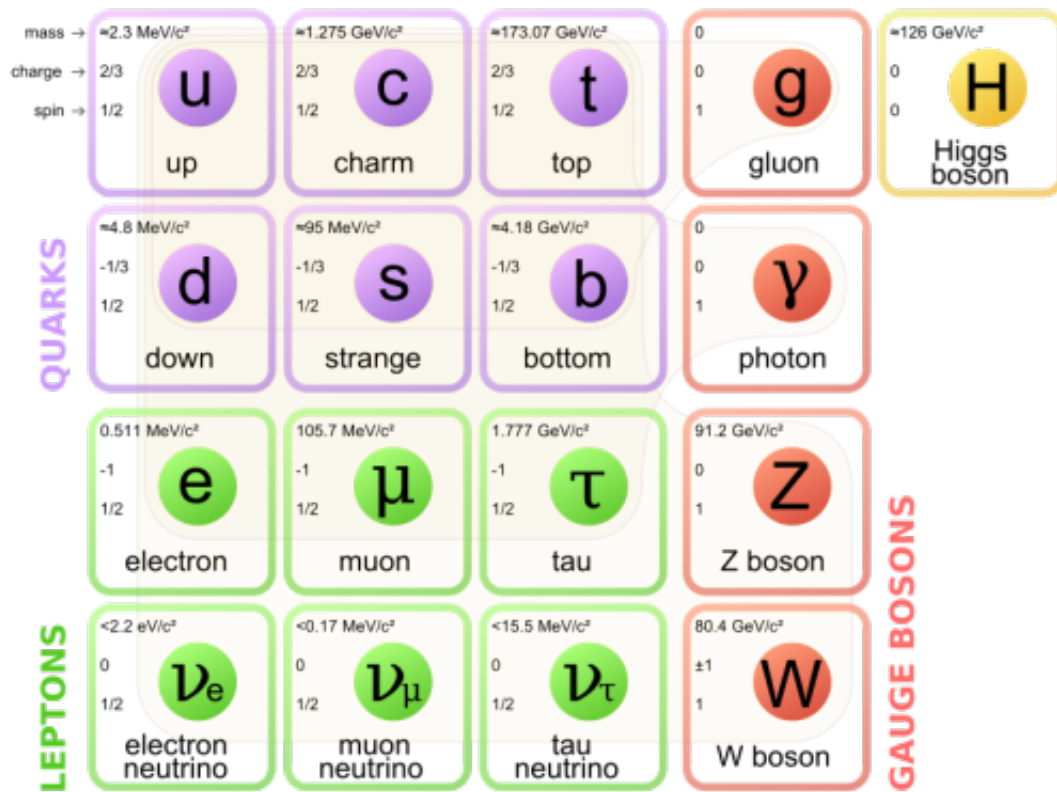


Figura 1.1: Lista delle particelle elementari, con relativa massa, carica e spin.

ca, includendo anche i fenomeni di interazione debole.

Infine il Modello Standard include il bosone di Higgs, un bosone scalare di spin nullo, che, attraverso il cosiddetto meccanismo di Higgs, è responsabile tra l'altro della massa dei mediatori dell'interazione debole. La sua scoperta è molto recente, ed è avvenuta nel 2012 simultaneamente negli esperimenti ATLAS e CMS.

1.2 La Cromodinamica Quantistica

La Cromodinamica Quantistica (QCD) [2] è la teoria di campi quantizzati che descrive le interazioni forti tra i quark attraverso scambi di gluoni. Essa è una teoria di gauge non abeliana con gruppo di simmetria $SU(3)$. Ciò significa che l'interazione forte avviene tra particelle con carica di colore, analoga ma diversa dalla carica elettrica della QED, che è una teoria di gauge con gruppo di simmetria $U(1)$, in quanto esistono tre colori (e anti-colori), solitamente indicati con *red* (r), *green* (g) e *blue* (b). Le interazioni tra quark sono mediate dai gluoni, che sono bosoni di massa nulla con spin 1. Esistono 8 possibili gluoni, ed ognuno possiede una carica di colore e una di anti-colore. A differenza di quelli dell'interazione elettromagnetica, i fotoni, questi mediatori possono anche interagire direttamente tra di loro. Sperimentalmente non sono mai stati osservati quark liberi, ma avviene il cosiddetto *confinamento di colore*: i quark si legano in doppietti o tripletti che devono necessariamente essere *bianchi*, ossia la carica di colore deve essere complessivamente nulla. La QCD garantisce che le uniche combinazioni possibili sono i *mesoni* (coppie quark-antiquark) e i *barioni* (tripletti di quark o di antiquark). Se si provasse a fornire energia per separare completamente un quark da un antiquark in un mesone, il sistema avrebbe abbastanza energia per creare una nuova coppia quark-antiquark, detta *jet adronico*, rendendo impossibile l'ottenimento di un quark libero. Il processo è presentato schematicamente in Figura 1.2.

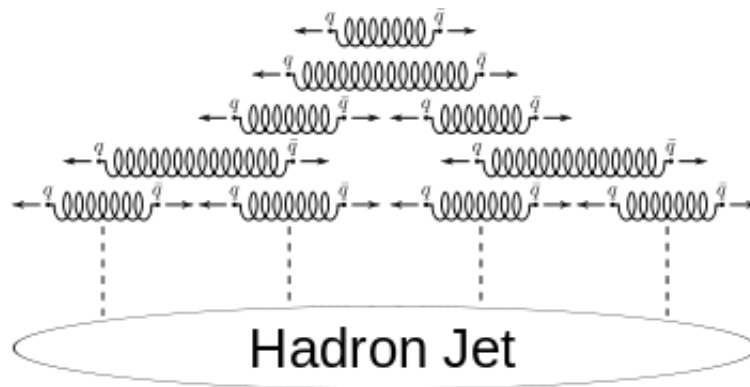


Figura 1.2: Rappresentazione della rottura di stringa QCD nel vuoto: il confinamento di colore fa sì che si crei un nuovo mesone tutte le volte che si fornisce energia al sistema, impedendo la formazione di un quark libero.

Un'altra proprietà importante della QCD è la *libertà asintotica*, per cui a basse scale di energia, o equivalentemente ad alte scale di distanza, l'intensità dell'interazione forte è estremamente alta, per cui è favorito lo stato strettamente legato degli adroni; invece, ad alte energie o a piccole distanze, l'interazione è meno intensa e permette l'utilizzo di approcci di calcolo perturbativi.

1.3 Quark-Gluon Plasma

Un modello euristico per descrivere i quark confinati negli adroni è il *MIT bag model* [3]. Secondo questa teoria i quark sono particelle aventi massa nulla all'interno di una scatola di dimensioni finite, e infinitamente massivi all'esterno. Così il confinamento non è altro che il risultato del bilancio tra la pressione esterna e quella interna data dall'energia cinetica dei quark stessi. Anche i gluoni scambiati tra i quark sono confinati nella scatola, e la carica di colore totale di quest'ultima deve essere bianca. Questo modello fornisce una conoscenza intuitiva del perché ci si dovrebbe aspettare di trovare nuove fasi della materia formata da quark, oltre alla materia adronica: se la pressione dei quark interni fosse incrementata oltre al valore della pressione esterna, si accederebbe ad uno stato della materia nuovo in cui i quark e i gluoni non sono più legati, chiamato *Quark-Gluon Plasma* (QGP). Una rappresentazione approssimativa del diagramma di fase della materia fortemente interagente è riportata in Figura 1.3. Quando il sistema raggiunge la temperatura critica di $T_C \simeq 150 - 200 \text{ MeV}$, il sistema presenta una transizione di fase del primo ordine tra materia adronica e QGP. La densità di energia nei pressi di questa transizione presenta una discontinuità, detta *calore latente di deconfinamento*. La regione con basse temperature e alte densità è detta regione di *diquark matter*, o regione di *superconduttività di colore*. Qui avviene la formazione di coppie di quark non neutre di colore, analoghe alle coppie di Cooper dei superconduttori.

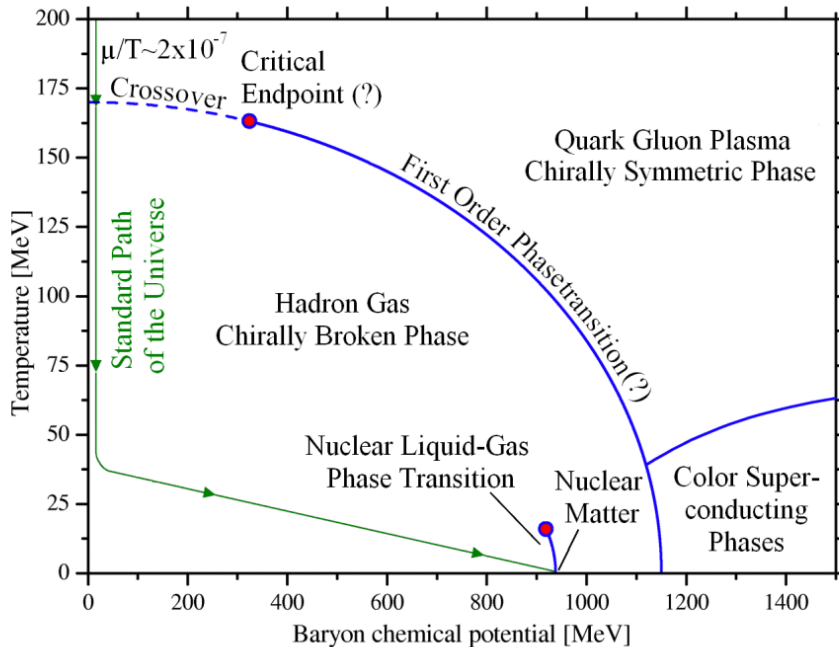


Figura 1.3: Diagramma di fase qualitativo della materia fortemente interagente.

1.4 Collisioni tra ioni pesanti

L'unico metodo conosciuto per creare le condizioni tipiche di temperatura e densità energetica elevate e produrre artificialmente il QGP sono le collisioni tra ioni pesanti ultrarelativistici ($\beta \approx 1$).

I nuclei degli ioni sono oggetti la cui estensione risulta molto grande se confrontata con tutte le scale di interesse della fisica delle alte energie. Per questa ragione, la geometria delle collisioni gioca un ruolo fondamentale nell'analisi e nell'interpretazione dei risultati sperimentali.

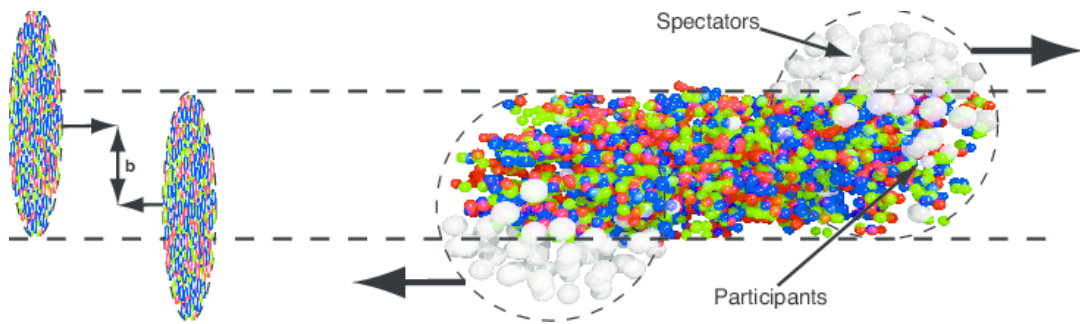


Figura 1.4: Schematizzazione di un urto tra due ioni pesanti.

Nel sistema del centro di massa, grazie alla contrazione di Lorentz nella direzione longitudinale (di propagazione del fascio), i due nuclei possono essere visti come dischi sottili di raggio (dimensione trasversale) $2R_A \simeq 2A^{\frac{1}{3}}$, dove A è il numero di nucleoni. Visibili in Figura 1.4, alcune delle quantità rilevanti sono [4]:

- Il parametro di impatto b , ossia la distanza tra i centri dei nuclei in procinto di collidere. Il parametro di impatto caratterizza la centralità della collisione; quest'ultima si dirà centrale se b è molto piccolo e lo scontro è pressoché frontale, si dirà invece periferica se b è grande rispetto alle dimensioni delle particelle.
- Il numero di nucleoni coinvolti (*participants*) N_{part} all'interno dei nuclei collidenti, ossia il numero di neutroni e protoni dei due ioni che prendono parte alla collisione; i restanti vengono chiamati spettatori (*spectators*) e perseverano nella loro traiettoria quasi imperturbati.
- Il numero totale di collisioni nucleone-nucleone incoerenti N_{coll} .

Numerosi modelli teorici sono stati sviluppati per descrivere le dinamiche di collisione a partire da queste quantità. Inoltre, è consueto in fisica subnucleare descrivere le traiettorie delle particelle in termini della variabile *rapidità* y , oppure della variabile

pseudorapidità η , definite rispettivamente come segue:

$$y = \frac{1}{2} \ln \left(\frac{E + p_L}{E - p_L} \right)$$

dove E è l'energia e p_L è la componente longitudinale del momento della particella, rispetto all'asse del fascio;

$$\eta = \frac{1}{2} \ln \left(\frac{p + p_L}{p - p_L} \right) = - \ln \left(\tan \frac{\theta}{2} \right)$$

dove θ è l'angolo tra impulso della particella e asse del fascio. Per particelle di massa nulla come il fotone, la rapidità e la pseudorapidità coincidono, ma per quelle massive questa corrispondenza vale solo nel limite ultrarelativistico.

1.5 Evoluzione del QGP

A collisione avvenuta, la formazione di un plasma di quark e gluoni avviene solamente nel caso in cui si raggiungano le condizioni critiche di temperatura e densità energetica [5]. Se ciò non accade, il sistema entra in *evoluzione idrodinamica* (Figura 1.5 a sinistra). Questo è il caso tipico delle collisioni tra protoni, o di collisioni tra ioni pesanti non sufficientemente energetiche e centrali. Subito dopo la collisione vi è la fase pre-adronica in cui, sebbene avvenga un sostanziale incremento della pressione e della temperatura, non si manifesta alcun deconfinamento dei partoni. Se, invece, l'energia dell'urto tra ioni pesanti è sufficiente da soddisfare le condizioni di creazione del QGP, il processo è più complesso (Figura 1.5, a destra):

1. *Pre-Equilibrium Phase* ($t < 1 \text{ fm}/c$): in questa fase i partoni diffondono l'uno sull'altro dando luogo ad una abbondante produzione di quark e gluoni deconfinati. Inoltre vengono prodotte molte particelle ad elevato impulso trasverso ($p_T \gg 1 \text{ GeV}/c$), una grande quantità di fotoni, sia reali che virtuali che decadono in coppie leptone-antileptone.
2. *Termalizzazione* ($t \approx 1 - 10 \text{ fm}/c$): questa fase è provocata dalle interazioni elastiche ed inelastiche tra i partoni nel QGP. Le interazioni inelastiche possono anche cambiare la composizione di sapore delle particelle. A causa della pressione interna, il sistema all'equilibrio termico inizia ad espandersi rapidamente, convertendosi in un gas adronico (fase mista).

3. *Adronizzazione* ($t \approx 20 fm/c$): durante l'espansione il sistema si raffredda e, dopo aver raggiunto il valore critico di densità energetica, inizia il processo di adronizzazione e i quark ed i gluoni del QGP si condensano in nuovi adroni. L'interazione tra gli adroni continua finché il relativo tasso è in grado di sostenere l'espansione del QGP; dopodiché, la composizione di sapore del QGP si fissa raggiungendo il congelamento chimico (*chemical freeze-out*).
4. *Congelamento termico* (*thermal freeze-out*): quando la distanza media tra gli adroni diventa maggiore del raggio d'interazione forte (a $T \approx 120 MeV$), le diffusioni elastiche tra adroni cessano e resta fisso anche lo spettro cinematico della materia risultante.

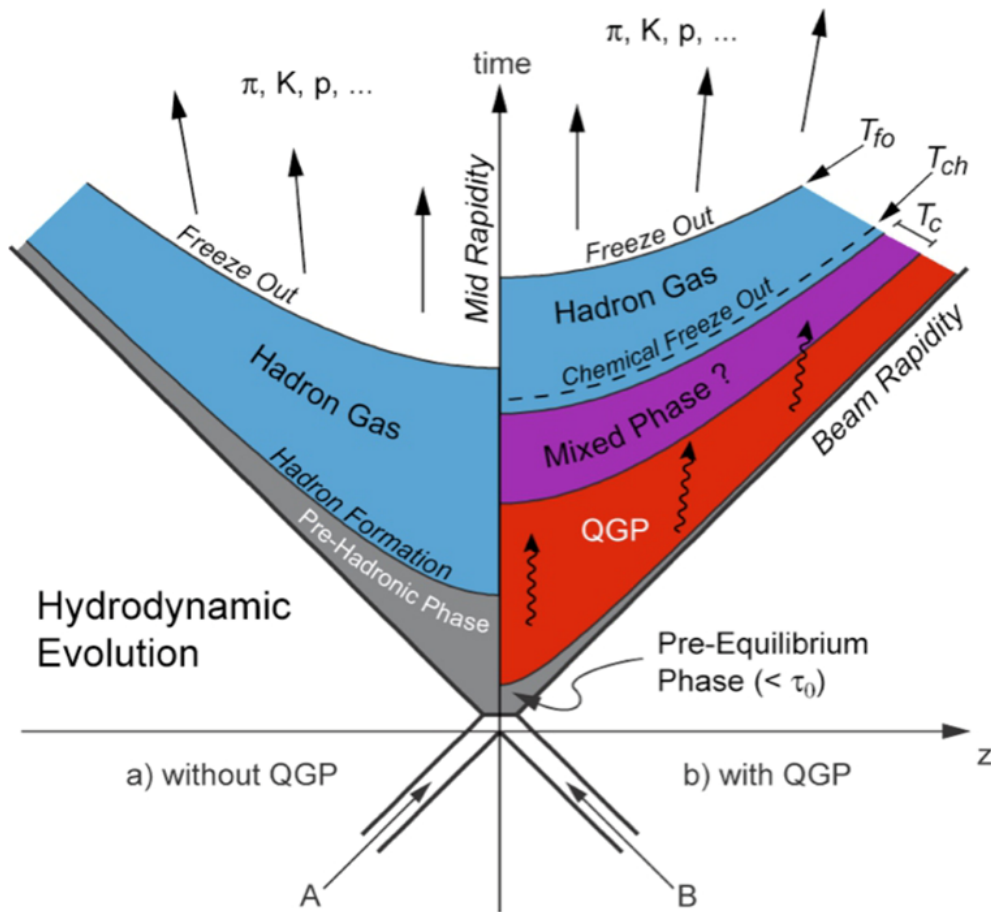


Figura 1.5: Fasi dell'evoluzione di un sistema in cui è avvenuta una collisione tra particelle. A sinistra il caso in cui il sistema entra in evoluzione idrodinamica, a destra il caso in cui le condizioni consentono la formazione del Quark-Gluon Plasma.

1.6 Adronizzazione di sapori pesanti in collisioni pp

Nello studio delle proprietà del QGP, i quark pesanti (charm e bottom) rivestono un ruolo fondamentale perché, in virtù della loro massa elevata, vengono prodotti in collisioni hard (ossia ad alto momento Q^2 trasferito) tra i partoni dei nucleoni nelle primissime fasi della collisione nucleo-nucleo, prima ancora che il sistema termalizzi e si formi lo stato di QGP. Essi quindi si propagano attraverso il sistema ultra-denso, interagendo con i suoi costituenti e fornendo una misura diretta delle sue proprietà. Per poter comprendere appieno una misura effettuata in collisioni A-A è necessario confrontarla con la stessa misura effettuata in collisioni pp e p-A.

L'adronizzazione di sapori pesanti in collisioni pp attraverso il processo di frammentazione viene descritta matematicamente utilizzando il *teorema di fattorizzazione* [6]. Data la scala del momento Q^2 trasferito nel processo di collisione, esso consiste nel separare il contributo perturbativo ad alta energia della produzione del *leading parton* dalla successiva conversione nello stato adronico a bassa energia (non perturbativo). Il processo complessivo è

$$p + p \rightarrow h + X$$

dove l'adrone di riferimento h è dato dal decadimento del partone c proveniente dallo scattering $a + b$ dei partoni del protone $a + b \rightarrow c + d$.

Si può ora esprimere la sezione d'urto invariante della produzione dell'adrone nel medio range di rapidità per collisioni pp come

$$\frac{d\sigma_{pp}^h}{dyd^2p_T} = K \sum_{abcd} \int dx_a dx_b f_a(x_a, Q^2) f_b(x_b, Q^2) \frac{d\sigma}{d\hat{t}}(ab \rightarrow cd) \frac{D_{h/c}^0}{\pi z_c}$$

dove:

- $f_i(x_i, Q^2)$ sono le *Funzioni di Distribuzione dei Partoni* (PDF) riferite ai partoni del protone;
- $\frac{d\sigma}{d\hat{t}}(ab \rightarrow cd)$ è la sezione d'urto elementare perturbativa QCD della produzione della particella c a partire dallo scattering dei partoni $a + b$;
- $D_{h/c}^0$ è la *Funzione di Frammentazione* (FF), un oggetto adimensionale che fornisce la probabilità che il partone c adronizzerà emettendo gluoni nell'adrone finale h , trasportando una frazione del momento del partone iniziale. Le Funzioni di Frammentazione non sono calcolabili perturbativamente e devono essere quindi misurate sperimentalmente.

Funzioni di Distribuzione dei Partoni

L'evidenza sperimentale dimostra che gli adroni non sono particelle elementari puntiformi, ma sono composte da *partoni*: quark e gluoni. Successivamente vi è la distinzione tra i *quark di valenza*, ossia i quark che effettivamente determinano i numeri quantici dell'adrone, come gli *uud* per il protone, mentre si definiscono *partoni del mare* tutti i restanti partoni, includendo sia i gluoni che i quark creati e distrutti nei processi virtuali che avvengono all'interno dell'adrone secondo la QCD.

Consideriamo l'esempio del protone. Denoteremo con $q^v(x)$ la densità di probabilità di un quark di valenza, con $q^s(x)$ quella di un quark del mare e con $g(x)$ quella di un gluone, dove x è la frazione del momento totale trasportato da q o g . Sapendo che i quark di valenza del protone sono *uud*, otteniamo la condizione:

$$\int_0^1 dx u^v(x) = 2 \quad \int_0^1 dx d^v(x) = 1$$

I quark del mare sono sempre prodotti in coppie $q\bar{q}$, e danno un contributo nullo al numero barionico

$$\int_0^1 dx [u^s(x) - \bar{u}^s(x)] = 0 \quad \int_0^1 dx [d^s(x) - \bar{d}^s(x)] = 0$$

La stessa condizione è valida per gli altri quark del mare, s^s , c^s , b^s e t^s .

Il momento totale portato da tutti i partoni deve contribuire al momento totale, perciò si ha la condizione:

$$\int_0^1 dx x [u^v(x) + d^v(x) + \sum_q (q^s(x) + \bar{q}^s(x))] = 1$$

I quark pesanti sono inclusi in questa trattazione, ma sono attivi solamente se la scala di energia Q del sistema è superiore alla massa m_q del quark pesante stesso.

In Figura 1.6 sono mostrate le funzioni di distribuzione del momento dei partoni del protone a $Q^2 = 10 \text{ GeV}^2$. È interessante notare che il termine dei gluoni rappresenta circa metà del momento totale.

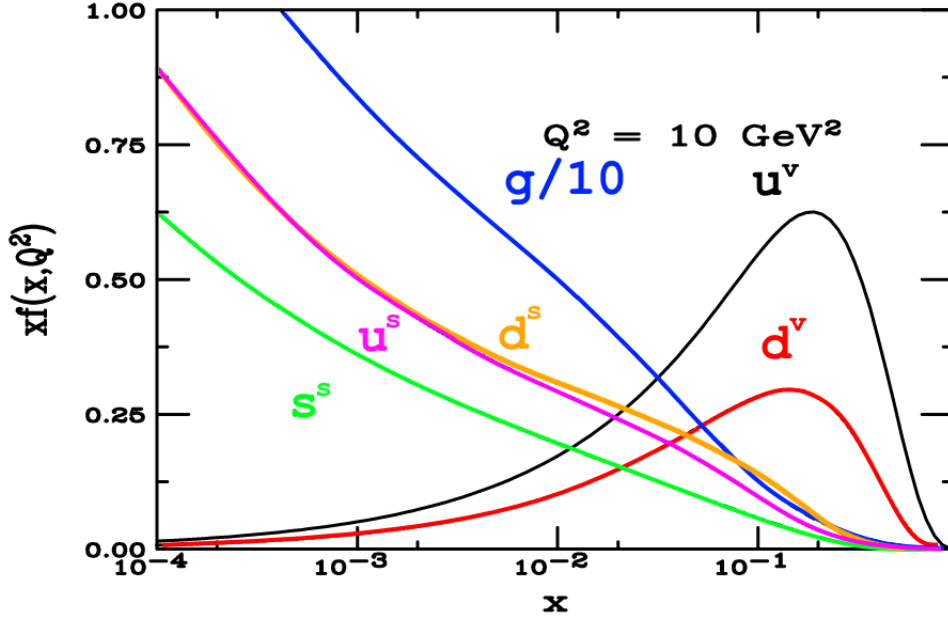


Figura 1.6: Funzioni di distribuzione del momento dei partoni del protone $xf(x)$, secondo la parametrizzazione CTEQ6M dei partoni a $Q^2 = 10 \text{ GeV}^2$. La distribuzione dei gluoni è divisa per 10 per migliorarne la visualizzazione.

Funzioni di Frammentazione

Per comprendere meglio le funzioni di frammentazione, consideriamo il processo di annichilazione di un sistema elettrone-positrone per produrre una coppia quark anti-quark [7]

$$e^+e^- \rightarrow q\bar{q}$$

Se l'energia della collisione è Q , allora l'energia del fascio è $E_f = Q/2$, e i quark prodotti hanno energia E_q uguale a E_f . Dunque se l'adrone h dello stato finale ha energia E_h , porterà una frazione di energia data da

$$z = \frac{E_h}{E_q} = \frac{2E_h}{Q}$$

La sezione d'urto differenziale per la produzione di adroni come funzione di z è:

$$\frac{d\sigma(e^+e^- \rightarrow hX)}{dz} = \sum_q \sigma(e^+e^- \rightarrow q\bar{q})[D_q^h(z) + D_{\bar{q}}^h(z)]$$

Questa formula è data dall'applicazione del teorema di fattorizzazione senza le PDF, giustificata dal fatto che gli elettroni sono particelle elementari. La Funzione di Frammentazione $D_q^h(z)$ rappresenta la probabilità che l'adrone h dello stato finale trasporti

una frazione z del momento iniziale del quark. La FF descrive la transizione partone-adrone nello stesso modo in cui la PDF descrive la struttura partonica di un adrone. Dunque la somma delle energie di tutti gli adroni prodotti deve formare l'energia del quark iniziale,

$$\sum_h \int_0^1 dz z D_q^h(z) = \sum_h \int_0^1 dz z D_{\bar{q}}^h(z) = 1$$

La molteplicità di h è data dalla somma delle probabilità per produrre h da tutti i possibili quark e anti-quark:

$$n_h = \sum_q \int_{z_{min}}^1 dz [D_q^h(z) + D_{\bar{q}}^h(z)]$$

dove z_{min} è l'energia di soglia per produrre un adrone di massa m_h : $z_{min} = 2m_h/Q$. Le FF possono avere diverse parametrizzazioni. Spesso è usata la parametrizzazione

$$D_q^h(z) = N \frac{(1-z)^n}{z}$$

dove N e n sono costanti specifiche per un dato adrone h . I parametri sono ottenuti sperimentalmente dal fitting della grande mole di dati disponibile per le collisioni e^+e^- . Si pensa che le Funzioni di Frammentazione siano universali, pertanto una volta calcolati i parametri per le collisioni e^+e^- , dovrebbero essere applicabili per altri casi, come le collisioni ep , pp e $p\bar{p}$.

1.7 Adronizzazione di sapori pesanti in collisioni A-A

Fin dalle prime osservazioni di produzione di adroni in collisioni di ioni pesanti fu evidente che il processo di adronizzazione è diverso rispetto al caso della pura frammentazione nel vuoto. I modelli che provano a spiegare questa differenza considerano che avvenga in concomitanza anche un secondo meccanismo chiamato *ricombinazione* o *coalescenza*. La differenza tra i due processi è che nella frammentazione, il momento iniziale è distribuito tra i frammenti, mentre nella ricombinazione, due o tre partoni vicini nello spazio delle fasi (posizione e momento) producono un adrone con momento trasverso pari alla somma dei momenti dei partoni iniziali, come si vede in Figura 1.7. Il calcolo degli effetti di ricombinazione nelle collisioni di ioni pesanti è particolarmente complesso perché non è possibile scrivere la funzione d'onda di tutti i partoni che compongono il QGP. La probabilità di trovare due o tre partoni vicini nello spazio delle fasi diminuisce all'aumentare del momento trasverso dell'adrone dello stato finale, perciò la coalescenza contribuisce meno ad alti impulsi trasversi p_T , e la frammentazione risulta il fenomeno dominante. Inoltre, l'effetto della ricombinazione è più significativo in collisioni centrali, poiché queste ultime favoriscono maggiormente la transizione a QGP, mentre la frammentazione è tipica delle collisioni periferiche.

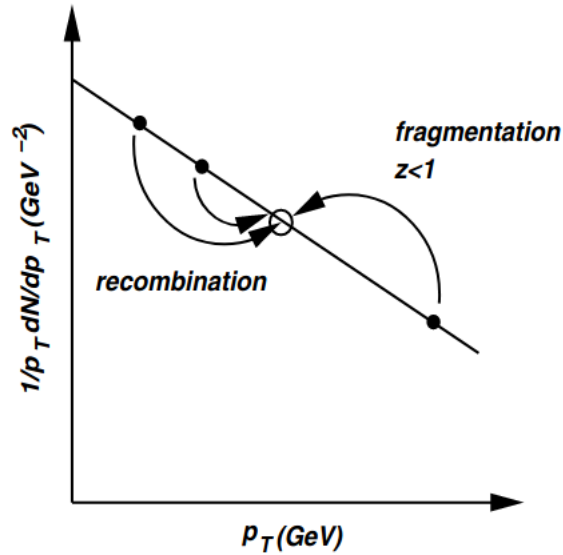


Figura 1.7: I meccanismi di ricombinazione e frammentazione che competono per creare lo stesso adrone dello stato finale.

1.8 Risultati sperimentali

1.8.1 Adroni charmati in collisioni Pb-Pb a $\sqrt{s_{NN}} = 5.02 \text{ TeV}$

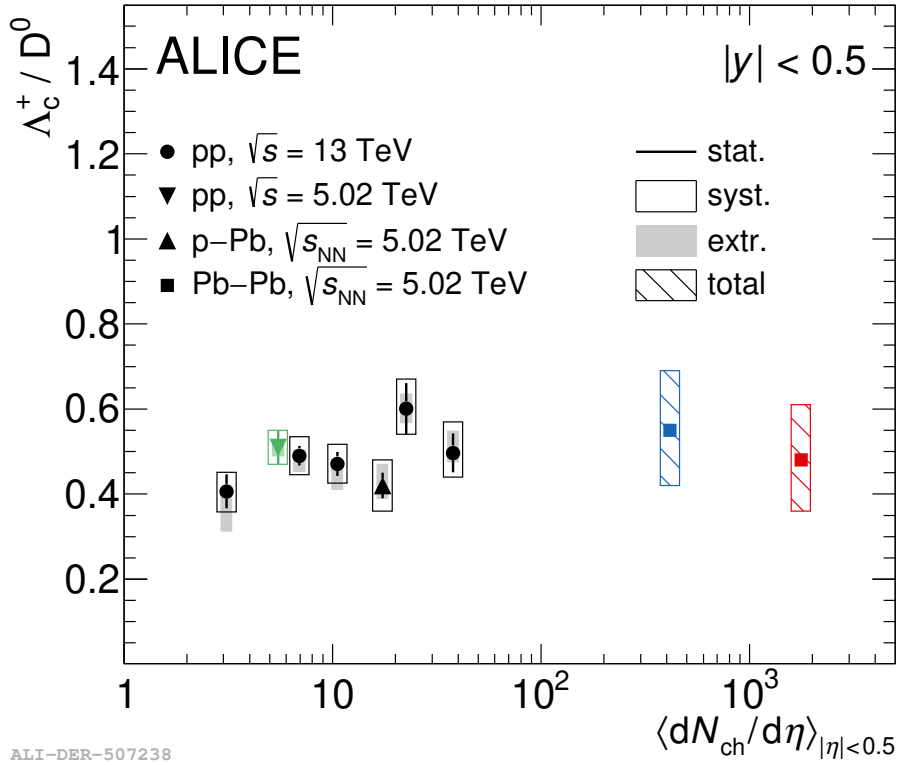
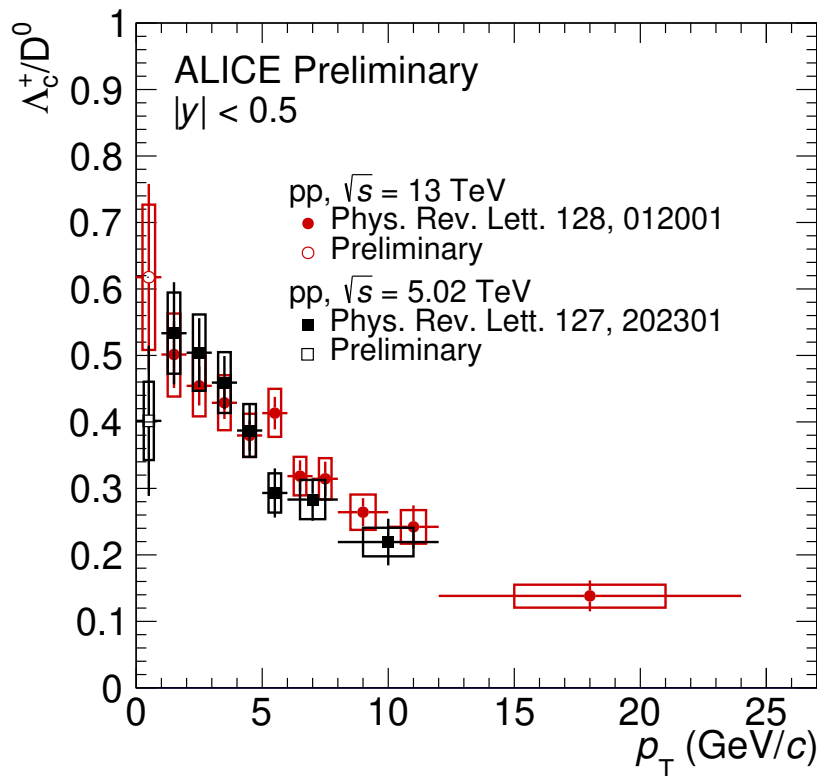


Figura 1.8: Rapporto Λ_c^+ / D^0 misurato in ALICE per varie tipologie di sistema collidente.

In Figura 1.8 è riportato il valore del rapporto barione-mesone Λ_c^+ / D^0 misurato dall'esperimento ALICE a LHC e integrato su tutto lo spettro di impulsi trasversi del barione Λ_c^+ (utilizzando valori estrapolati dove non sono presenti misure sperimentali), per diversi sistemi collidenti: protone-protone, protone-nucleo e nucleo-nucleo, sia centrali che periferiche. Contrariamente a quanto atteso, i valori misurati del rapporto non differiscono in maniera significativa tra di loro, e in particolare non si osserva il consistente aumento della produzione di barioni charmati, riferito alla produzione di mesoni charmati, previsto dall'insorgere di meccanismi di ricombinazione in collisioni A-A. Infatti, nei processi di ricombinazione o coalescenza, la formazione di un barione è molto meno sfavorita rispetto alla formazione di un mesone rispetto a processi di frammentazione, e questo porta a prevedere un valore del rapporto barione-mesone maggiore in collisioni A-A. Una possibile spiegazione è che tali o simili meccanismi siano già presenti e importanti, soprattutto a basso p_T , anche in collisioni pp e p-A alle energie di LHC. L'andamento del

rapporto in funzione della molteplicità dimostra che, se presenti, tali meccanismi sono all'opera anche a basse molteplicità. Inoltre, i valori del rapporto Λ_c^+/D^0 in collisioni pp risultano essere maggiori rispetto ai valori misurati in collisioni e^+e^- e e^-p , e alle previsioni dei modelli teorici basati su tali misurazioni. Questi modelli prevedono un valore intorno a 0.1, significativamente inferiore al valore compreso tra 0.4 e 0.6 misurato in ALICE [8] [9] [10] [11]. Questa discrepanza può essere interpretata come una indicazione che le probabilità con cui un quark charm adronizza in uno specifico adrone charmato (le Funzioni di Frammentazione) non sono universali, come ritenuto fino ad oggi, ma dipendono dalle caratteristiche del sistema collidente.

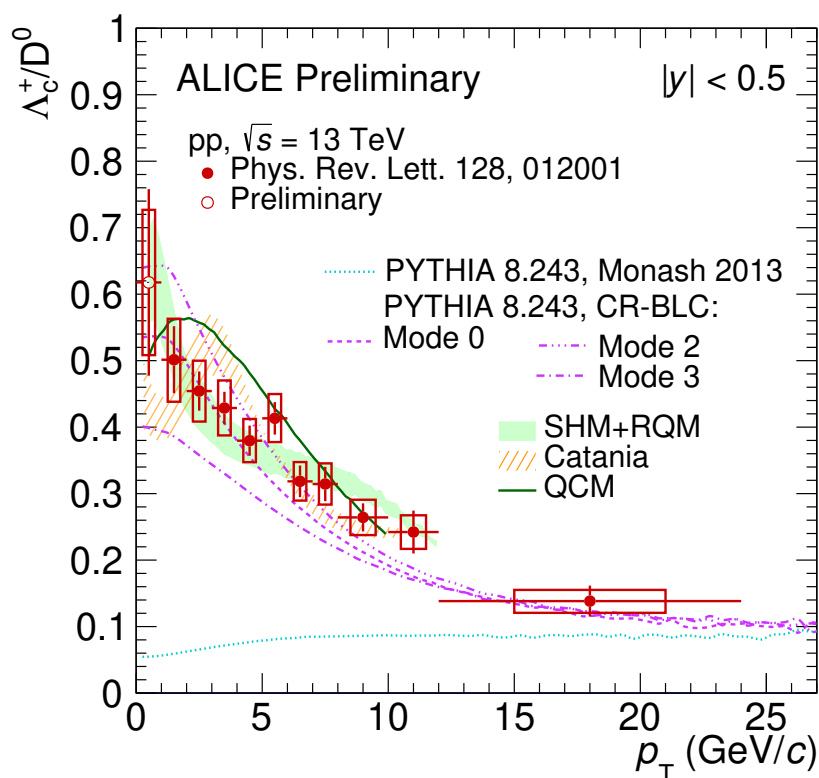
1.8.2 Adroni charmati in collisioni pp a $\sqrt{s} = 5.02$ TeV e a $\sqrt{s} = 13$ TeV



ALI-PREL-502456

Figura 1.9: Rapporto di produzione di adroni charmati Λ_c^+ e D^0 in funzione di p_T misurato in collisioni pp a diversa energia.

La Figura 1.9 mostra il rapporto Λ_c^+/D^0 in funzione di p_T in collisioni pp a $\sqrt{s} = 5.02 TeV$ e a $\sqrt{s} = 13 TeV$ misurati con il rivelatore ALICE a LHC. Entrambi i rapporti mostrano un certo accordo per valori di impulso fuori dal bin $p_T = [0, 1] GeV/c$, con un andamento decrescente, mentre i valori al suo interno, calcolati unicamente attraverso il canale di decadimento $\Lambda_c^+ \rightarrow pK\pi$, sono molto diversi. In effetti l'analisi dei decadimenti a basso impulso trasverso è particolarmente delicata in questo range di impulso.



ALI-PREL-502450

Figura 1.10: Rapporto di produzione di adroni charmati Λ_c^+ e D^0 in funzione di p_T in collisioni pp a $\sqrt{s} = 13 TeV$, confrontato con diversi modelli teorici.

La Figura 1.10 mostra il rapporto Λ_c^+/D^0 in funzione di p_T in collisioni pp a $\sqrt{s} = 13 TeV$, posto a confronto con diverse predizioni teoriche.

- PYTHIA 8.243 Monash 2013 [12] è un generatore Monte Carlo (MC) che implementa processi di frammentazione con produzione di adroni charmati, misurati dalle collisioni e^+e^- . Esso predice un rapporto Λ_c^+/D^0 di circa 0.1 con una debole dipendenza da p_T , e costituisce una grossa sottostima dei dati sperimentali,

soprattutto a bassi range di impulso trasverso. Questo fatto supporta l'ipotesi che la frammentazione è un meccanismo che domina ad alti p_T .

- PYTHIA 8.243 CR-BLC (Color Reconnection - Beyond Leading Color) [13] è un generatore MC che implementa la *riconnesione di colore*. Questo modello di adronizzazione è basato sul modello a stringhe di Lund. Sono presenti tre possibili modalità, che introducono vincoli più o meno restrittivi sulla generazione: Mode 0, senza vincoli, Mode 2, con vincoli stretti, Mode 3, con vincoli più larghi. Questo modello predice piuttosto bene l'andamento del rapporto Λ_c^+/D^0 , in particolare nella Mode 0.
- Il SHM+RQM (Statistical Hadronization Model - Relativistic Quark Model) [14] calcola le frazioni di adroni charmati basandosi su densità termiche, dunque dipendenti dalla massa dello stato e dal fattore di degenerazione di spin. Fa uso di ulteriori stati barionici eccitati, che non sono stati ancora misurati ma si assume che esistano secondo il modello relativistico dei quark (RQM). Le previsioni di questo modello sono buone per tutti i range di p_T
- Il modello Catania [15] assume che anche in collisioni pp si possa creare uno stato di QGP, e che dunque l'adronizzazione avvenga sia per frammentazione che per coalescenza.

Questi modelli, escluso PYTHIA 8.243 Monash 2013, forniscono previsioni simili in quasi tutto il range di impulso trasverso, ma presentano valori piuttosto diversi nel bin $[0, 1]$ di p_T . L'analisi dati nel bin $[0, 1]$ di p_T è dunque molto importante, perché permette di distinguere i modelli teorici più affidabili da quelli che lo sono meno. È tuttavia un'analisi molto difficile, a causa del bassissimo rapporto segnale su fondo. Inoltre, come si può vedere da Figura 1.10, l'errore statistico e quello sistematico sono significativi, il che rende la misura sperimentale meno attendibile e di difficile interpretazione.

Capitolo 2

L'esperimento ALICE

2.1 Il Large Hadron Collider

Il Large Hadron Collider (LHC) [16] è il più grande e potente acceleratore di particelle al mondo, e rappresenta l'ultimo elemento del complesso sistema di acceleratori del CERN presso Ginevra, in Figura 2.1. Terminato di costruire nel 2008 con la collaborazione di Italia, Francia e Svizzera, è costituito da un anello di 27 km di circonferenza, formato da magneti superconduttori e strutture adibite per l'accelerazione delle particelle. I magneti sono mantenuti nello stato di superconduttività grazie ad un sistema di raffreddamento a elio liquido, che mantiene i cavi ad una temperatura di $-273.1\text{ }^\circ\text{C}$. Dentro l'acceleratore, due fasci di particelle confinati in due *beam pipes* separate sono accelerati in direzioni opposte fino a velocità prossime a quella della luce, e vengono fatti collidere presso i siti dove si trovano i quattro rivelatori: ATLAS, CMS, ALICE e LHCb [17].

- ATLAS e CMS sono i due rivelatori più grandi, e sono *general-purpose*. Non hanno dunque uno scopo specifico ma investigano numerosi aspetti della fisica, dalla ricerca del bosone di Higgs, a quella di evidenza di altre dimensioni, oppure di particelle che compongono la materia oscura.

I due esperimenti, anche se si prefiggono gli stessi obiettivi, lavorano in modo completamente indipendente, e i rivelatori sono realizzati con soluzioni tecnologiche differenti. Per questo, misure e risultati ottenuti da ATLAS e CMS sono considerate indipendenti.

- LHCb e ALICE, sono invece esperimenti specializzati nello studio di particolari fenomeni. L'esperimento LHCb si occupa di indagare le sottili differenze tra materia e antimateria attraverso lo studio del quark bottom. ALICE è dedicato allo studio degli urti di ioni pesanti ed indaga le caratteristiche e le proprietà del Quark-Guon Plasma.

CERN's Accelerator Complex

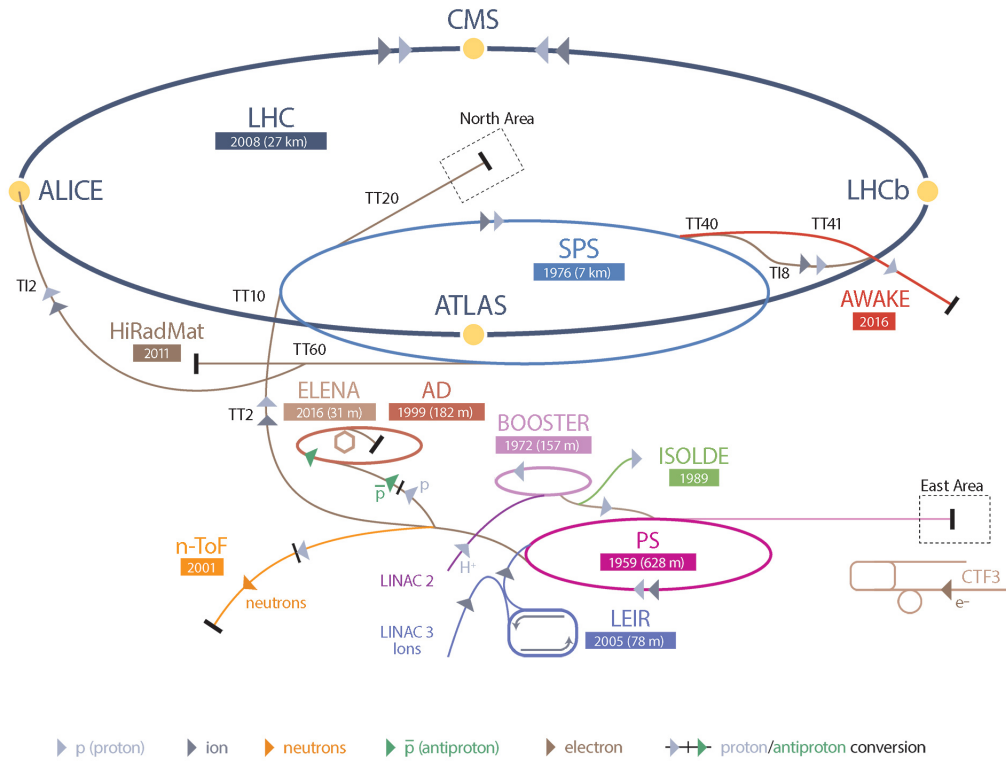


Figura 2.1: Schema del complesso di acceleratori del CERN a Ginevra.

2.2 ALICE: A Large Ion Collider Experiment

ALICE [18] è il rivelatore espressamente dedicato alla fisica degli ioni pesanti ad LHC, allo scopo di studiare la fisica della materia fortemente interagente ad elevate densità e temperature (oltre 100000 volte quella del nucleo del sole), ossia nella fase del Quark-Gluon Plasma. La Figura 2.2 rappresenta lo schema generale del rivelatore, e la Figura 2.3 la sua sezione trasversale. ALICE è formato da una parte centrale, il *central barrel*, dove vengono effettuate le misure per adroni, elettroni e fotoni, ed una parte *forward* con lo spettrometro per muoni. Il *central barrel* copre un angolo polare da 45° a 135° ed i sensori che contiene sono immersi in un campo magnetico di intensità $B = 0.5 T$ generato da un grande magnete solenoidale. Dall'interno verso l'esterno si trovano i seguenti rivelatori:

- Inner Tracking System (ITS);
- Time-Projection Chamber (TPC);
- Transition Radiation Detector (TRD);
- Time-Of-Flight (TOF);
- High-Momentum Particle Identification Detector (HMPID);
- Due calorimetri elettromagnetici: PHOS e EMCal.

I primi quattro rivelatori coprono l'intero angolo azimutale, mentre gli ultimi tre ne coprono solo una parte. Nella parte *forward* sono presenti numerosi altri rivelatori più piccoli posizionati a piccoli angoli rispetto alla *beam pipe*, per misure che caratterizzano globalmente gli eventi e per il trigger: ZDC, PMD, FMD, T0, V0. Al di sopra del *central barrel* si trova l' Alice Cosmic Ray Detector (ACORDE), che è un array di scintillatori che vengono attivati dai raggi cosmici, utilizzato per scopi di calibrazione e allineamento dei rivelatori centrali, e per studiare la fisica dei raggi cosmici. ALICE è caratterizzato da una eccellente capacità di identificazione delle particelle (Particle Identification, PID) rispetto agli altri esperimenti presenti a LHC, grazie alla combinazione di differenti tecniche di PID usate dai diversi rivelatori, ciascuno specializzato in un certo range di impulso.

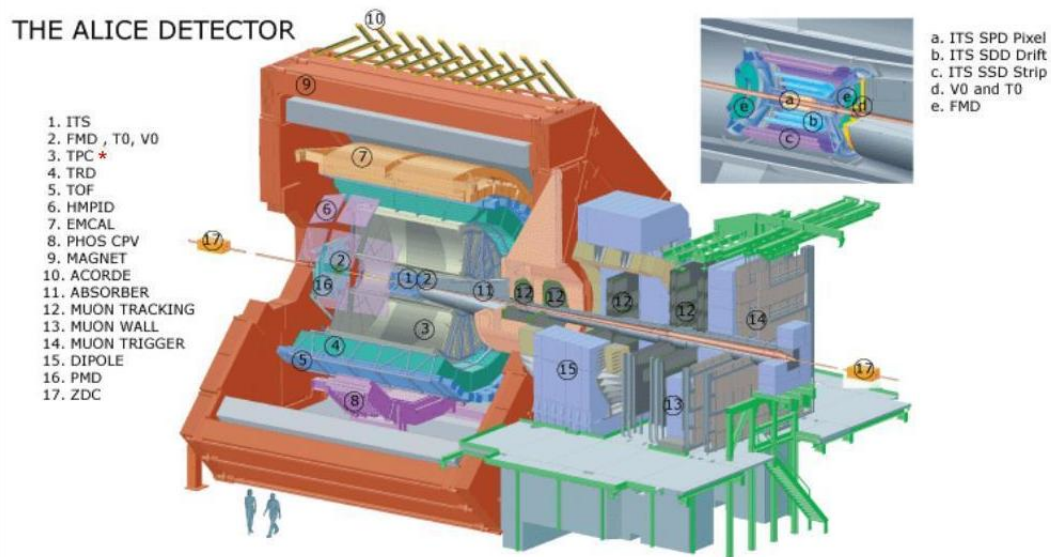


Figura 2.2: Schema del rivelatore ALICE.

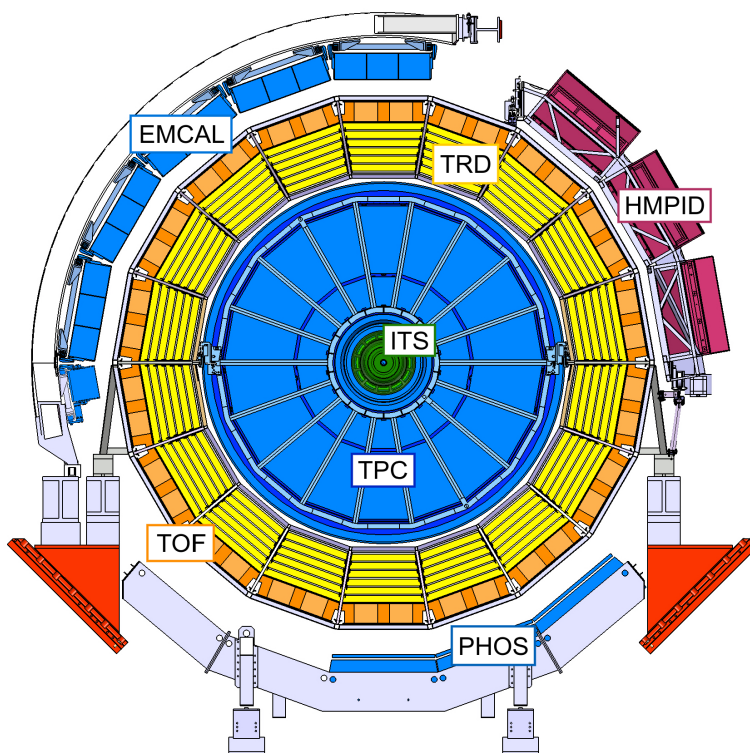


Figura 2.3: Sezione trasversale del central barrel del rivelatore ALICE.

2.2.1 Inner Tracking System (ITS)

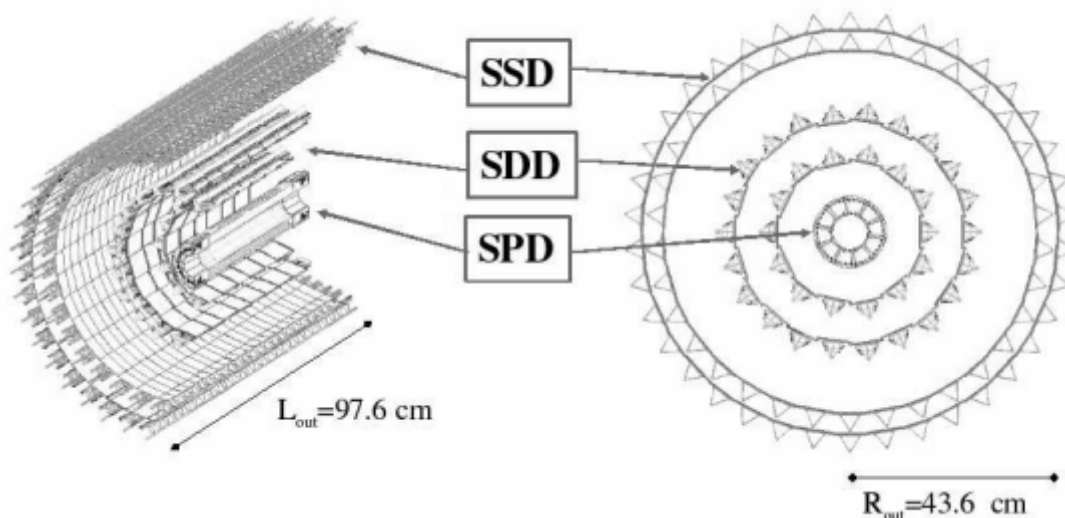


Figura 2.4: Schema longitudinale e trasversale dell'Inner Tracking System.

L'Inner Tracking System, schematizzato in Figura 2.4, circonda la *beam pipe*, e consiste di sei strati cilindrici coassiali di rilevatori al silicio, localizzati da un raggio minimo di 4 cm (imposto dalle dimensioni della *beam pipe*), a un raggio massimo di 43 cm (necessario per il matching delle traiettorie con il successivo rilevatore, la TPC). ITS copre un intervallo di pseudorapidità di $|\eta| < 0.9$, ed i suoi obiettivi sono: la localizzazione dei vertici primario e secondario, il tracciamento e l'identificazione di particelle con momento inferiore a $200 \text{ MeV}/c$, migliorare la risoluzione delle particelle ricostruite dalla TPC, e ricostruire quelle che attraverseranno le regioni morte di quest'ultima. Per acquisire un'adeguata alta risoluzione del parametro d'impatto, e data l'elevata densità di particelle aspettata nelle collisioni di ioni pesanti a LHC (circa $50 \text{ particelle}/\text{cm}^2$), sono stati scelti i Silicon Pixel Detectors (SPD) per i primi due strati a partire dall'interno, e i Silicon Drift Detectors per i successivi due; per gli ultimi due strati, dove la densità di particelle aspettata è ridotta a una particella per centimetro quadrato, sono stati scelti i Silicon micro-Strip Detectors (SSD). I quattro strati più esterni hanno un readout analogo, e possono essere usati per la Particle Identification attraverso la misura della perdita di energia dE/dx , nella regione non relativistica.

2.2.2 Time-Projection Chamber (TPC)

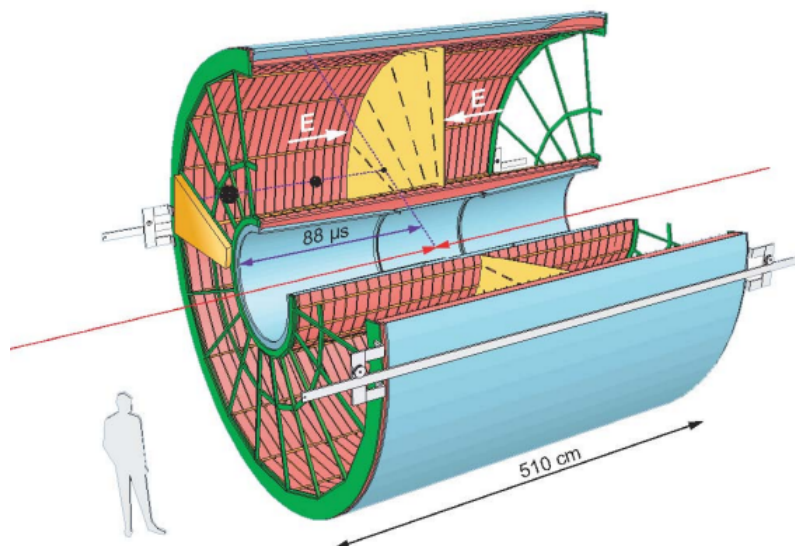


Figura 2.5: Schema della Time Projection Chamber.

La Time-Projection Chamber (raffigurata in Figura 2.5) è il principale rivelatore del *central barrel* per il tracciamento delle particelle. I suoi scopi sono fornire misure precise di impulso delle particelle cariche in un range di p_T di $0.1 - 100 \text{ GeV}/c$ e attuare l'identificazione delle particelle e la localizzazione dei vertici di decadimento. La TPC ha una simmetria cilindrica, coassiale con la direzione del fascio, con una zona attiva compresa tra un raggio di 85 cm e uno di 250 cm , ad esclusione di alcune zone morte dove non vengono rilevate particelle. L'interno del rivelatore è riempito con 90 m^3 di una miscela di $\text{Ne} - \text{CO}_2 - \text{N}_2$ in percentuali $85 - 10 - 5 \%$. Quando questo gas è attraversato da una particella carica, viene ionizzato creando una traccia formata da elettroni liberi e lacune. Questi vengono guidati tramite un campo elettrico interno alle *readout chambers*, presenti sulle due basi del cilindro e composte da rivelatori Gas Electron Multipliers (GEMs). La posizione degli hit sulle camere di readout, in aggiunta alla posizione lungo l'asse z determinata misurando il tempo di arrivo del segnale sugli endcap della TPC, permettono un tracciamento digitale in 3 dimensioni della traccia rilasciata dalla particella carica. Ricostruendo l'intera traiettoria e il suo raggio di curvatura è possibile risalire all'impulso della particella carica che l'ha generata.

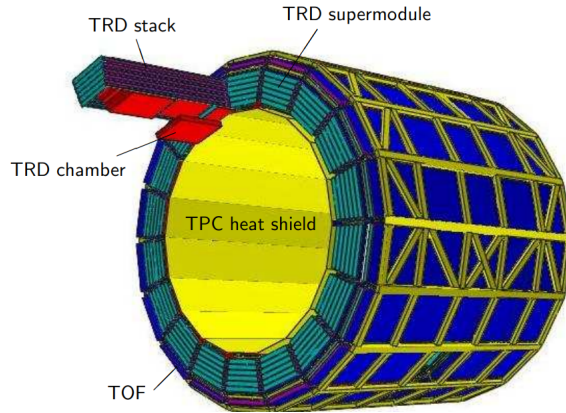


Figura 2.6: Schema del rivelatore TRD.

2.2.3 Transition Radiation Detector (TRD)

Il Transition Radiation Detector (schema in Figura 2.6) permette di identificare gli elettroni e i positroni con momento maggiore di $1 \text{ GeV}/c$ (mentre quelli con momento minore vengono rilevati dalla TPC), sfruttando la radiazione di transizione (TR) generata dal passaggio di questi attraverso una superficie di separazione tra materiali con diverso indice di rifrazione. Combinando i dati raccolti con quelli di ITS e TPC è possibile studiare la produzione di risonanze di mesoni leggeri e pesanti in collisioni pp e Pb-Pb. Il TRD è formato da 540 moduli individuali raggruppati in 18 supermoduli, ciascuno a sua volta suddiviso in cinque gruppi lungo l'asse del fascio e sei strati lungo la direzione radiale. Ogni modulo contiene un radiatore e una Multi-Wire Proportional Chamber (MWPC), un rivelatore riempito di una miscela di gas Xe/CO_2 , che misura il punto di passaggio di una particella carica.

2.2.4 Time-Of-Flight (TOF)

Il rivelatore Time-Of-Flight (schema in Figura 2.7) ha un ruolo fondamentale nella Particle Identification nel range di impulsi intermedi: è in grado di separare pioni e kaoni fino a $p_T < 2.5 \text{ GeV}/c$ e kaoni e protoni fino a $p_T < 4 \text{ GeV}/c$, nell'intervallo di pseudorapidità $|\eta| < 0.9$. Il TOF fornisce misure sul tempo di volo di ciascuna particella carica e, unendo i dati raccolti da ITS e TPC, permette di determinarne l'identità. Come i rivelatori precedentemente mostrati, il TOF ha una forma cilindrica coassiale alla *beam pipe*, ed è situato ad una distanza di 3.8 m da quest'ultima. È formato da 1638 Multi-gap Resistive Plate Chambers (MRPC), raggruppati similmente ai moduli del TRD in 18 settori azimutali, ciascuno a sua volta suddiviso in 5 moduli contenenti diversi MRPC, in base alla posizione: 15 per i moduli centrali, 19 per quelli intermedi o esterni.

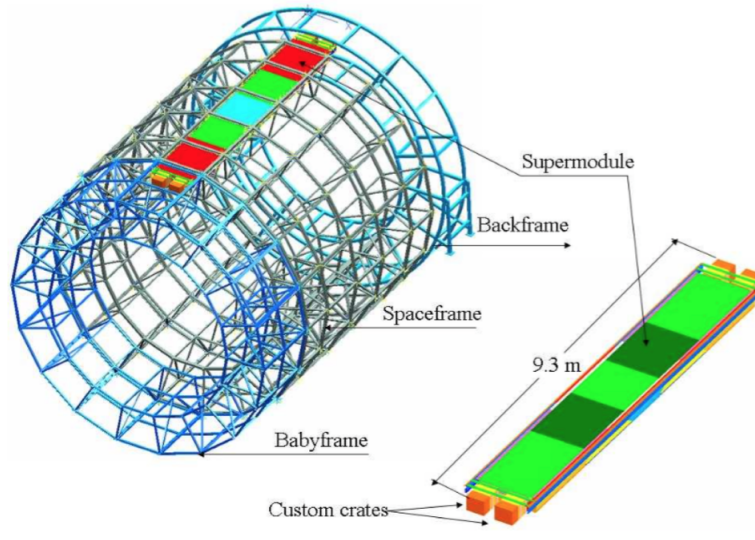


Figura 2.7: Schema del rivelatore Time-Of-Flight.

2.2.5 Rilevatori esterni

L'High-Momentum Particle Identification Detector (HMPID), il PHOton Spectrometer (PHOS) e il ElectroMagnetic Calorimeter (EMCal) sono i tre rivelatori più esterni del *central barrel*, e ognuno copre una certa porzione dell'angolo azimutale ed un diverso range di pseudorapidità.

- HMPID si trova ad una distanza radiale di 5 m , e copre una porzione di angolo polare di $60^\circ < \theta < 120^\circ$, e di angolo azimutale di $12.2^\circ < \phi < 58.8^\circ$. Il suo scopo è la Particle Identification di adroni carichi caratterizzati da un impulso trasverso troppo grande per essere identificati da ITS, TPC e TOF.
- PHOS è uno spettrometro elettromagnetico ad alta risoluzione capace di rivelare fotoni a basso p_T prodotti nelle prime fasi della collisione, per lo studio del *jet quenching*. Si trova ad una distanza radiale di 4.6 m , e copre un angolo polare nel range $83^\circ < \theta < 97^\circ$, un angolo azimutale in $220^\circ < \phi < 320^\circ$, e un intervallo di pseudorapidità di $|\eta| < 0.12$.
- EMCal è un calorimetro a campionamento con simmetria cilindrica, con lo scopo di studiare in maniera dettagliata il fenomeno del *jet quenching*. Esso è situato a circa 4.5 m dall'asse del fascio, e copre un angolo polare compreso tra $53^\circ < \theta < 127^\circ$, un angolo azimutale di ampiezza $\Delta\phi = 107^\circ$ in una posizione diametralmente opposta a PHOS, ed un intervallo di pseudorapidità di $|\eta| < 0.7$.

2.2.6 Forward Muon Spectrometer (FMS)

Nella parte *forward* del rivelatore ALICE si trova il rivelatore di muoni FMS, in un range polare di $171^\circ < \theta < 178^\circ$, corrispondente ad una pseudorapidità di $-4.0 < \eta < -2.5$. Il suo scopo è rilevare coppie di $\mu^+\mu^-$ prodotte nelle collisioni, grazie alla quali si possono ricostruire molti mesoni vettori contenenti quark pesanti, come i mesoni $\phi, \Upsilon, \Upsilon', \Upsilon'', \psi', J/\psi$. Variazioni dei conteggi di questi mesoni possono portare a provare indirettamente la formazione del QGP nelle collisioni pp. In Figura 2.8 è schematizzata la sezione longitudinale del FMS. Come si può vedere, il rivelatore è formato da un cono per l'assorbimento passivo di adroni e fotoni, seguito da 10 rivelatori planari, di cui alcuni immersi in un magnete a dipolo, un filtro passivo e quattro camere planari per il *trigger*.

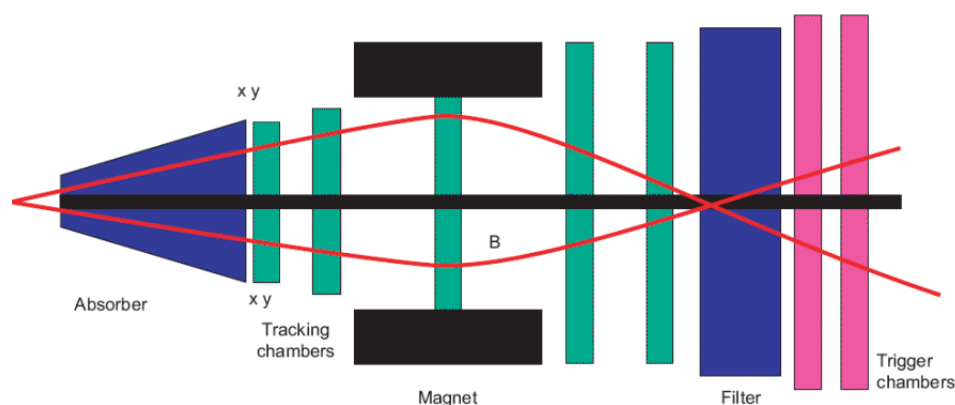


Figura 2.8: Rappresentazione della sezione longitudinale del Forward Muon Spectrometer.

2.2.7 Rilevatori ad alta pseudorapidità

Al di fuori del *central barrel* sono presenti alcuni piccoli rivelatori specializzati, con funzioni di *trigger* o per la misura di variabili globali delle collisioni. Queste misure sono effettuate per angoli molto piccoli rispetto alla direzione del fascio, ovvero hanno un'alta pseudorapidità.

- In una collisione tra due nuclei pesanti, alcuni nucleoni collidono mentre altri lasciano il punto di collisione senza aver avuto alcuna interazione. I rivelatori ZDC (Zero Degree Calorimeter) effettuano la misura dell'energia dei nucleoni spettatori, ovvero i quelli che non hanno partecipato all'urto. Da questa misura è possibile risalire al loro numero e quindi, per sottrazione, al numero di nucleoni che hanno partecipato all'urto. Questo osservabile è il più direttamente collegato alla geometria della collisione, e ne fornisce una misura della centralità. In ALICE due set

di ZDC sono posizionati a $\pm 116 m$ dal punto dove avviene la collisione e ciascuno è formato da un calorimetro specializzato nella rilevazione di neutroni spettatori (ZN) e uno in quella dei protoni spettatori (ZP).

- Il PMD (Photon Multiplicity Detector) è un grande array di contatori a gas in una struttura cellulare a nido d'ape, e misura la molteplicità e la distribuzione spaziale ($\eta - \phi$) dei fotoni con pseudorapidità $2.3 < \eta < 3.7$.
- Il FMD (Forward Multiplicity Detector) è un rivelatore progettato per misurare la molteplicità delle particelle cariche nel vasto range di pseudorapidità $-3.4 < \eta < -1.7$ e $1.7 < \eta < 5.0$. È composto da tre moduli posti in tre posizioni diverse, visibili in Figura 2.9.

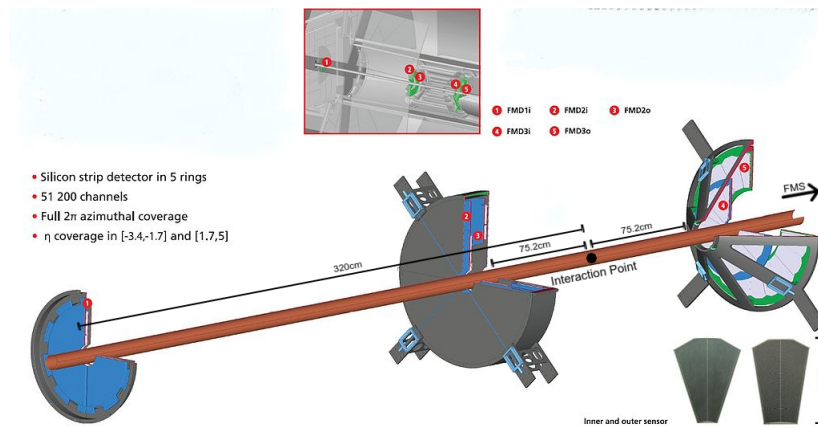


Figura 2.9: Rappresentazione grafica dei tre moduli che compongono il Forward Multiplicity Detector.

- V0 è il rivelatore che fornisce il *trigger* per gli altri rivelatori nel *central barrel*. Esso è suddiviso in due array di scintillatori, V0A e V0C, posizionati rispettivamente a $340 cm$ e a $90 cm$ in parti opposte rispetto al punto di interazione, coprendo intervalli di pseudorapidità di $2.8 < \eta < 5.1$ per V0A e $-3.7 < \eta < -1.7$ per V0C.
- Il rivelatore T0 fornisce l'istante iniziale della collisione per la misurazione del tempo di volo eseguita dal TOF e misura la posizione del vertice di interazione di ogni evento. Questo fatto può anche essere usato come *trigger*, e permette anche di distinguere gli urti tra particelle accelerate e particelle di gas, da quelli tra le sole particelle accelerate. È formato da due array di contatori Cherenkov, T0A e T0C, contenenti 12 contatori, rispettivamente posizionati a $375 cm$ e $72.7 cm$ rispetto al punto di interazione in direzioni opposte, coprendo intervalli di pseudorapidità di $4.61 < \eta < 4.92$ per T0A e $-3.28 < \eta < -2.97$ per T0C.

2.2.8 ALICE COsmic Ray DEtector (ACORDE)

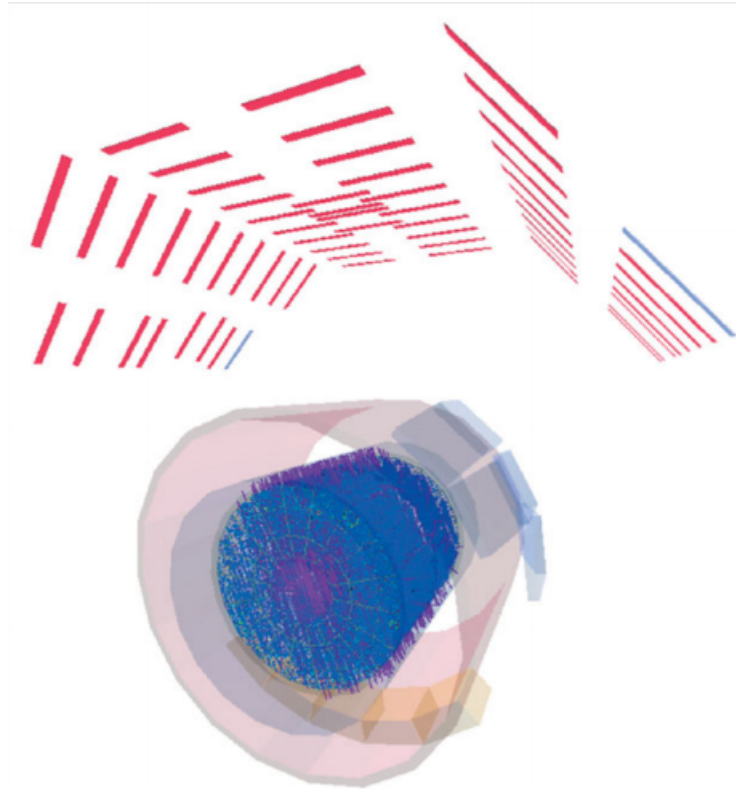


Figura 2.10: Schema del posizionamento dei moduli del rivelatore ACORDE.

Al di sopra del *central barrel* è situato ACORDE, il rivelatore di raggi cosmici di ALICE. I suoi scopi sono principalmente due. Per primo, fornire un segnale di *trigger* per le procedure di calibrazione e allineamento di altri rivelatori interni di ALICE. In secondo luogo è usato per lo studio dei raggi cosmici ad alta energia rilevando, in combinazione con TPC, TRD e TOF, muoni provenienti dall'atmosfera. È composto da 60 moduli posizionati come in Figura 2.10, con simmetria cilindrica a 8.5 m dall'asse del fascio, coprendo un angolo azimutale di 120° e un intervallo di pseudorapidità di $-1.3 < \eta < 1.3$.

Capitolo 3

Ricostruzione del barione Λ_c^+ .

3.1 Introduzione

Come si è visto nel Capitolo 1, lo studio della produzione di barioni charmati è di fondamentale importanza per investigare la formazione del QGP e le previsioni teoriche della QCD, in quanto l'adronizzazione da collisioni pp non è ancora stata compresa del tutto e i risultati sperimentali non sono riprodotti dai modelli teorici con sufficiente precisione. Il barione Λ_c^+ [19], tra gli adroni con sapori pesanti, è particolarmente interessante perché è il barione charmato con massa minore e quindi quello prodotto con più abbondanza in collisioni pp. Le caratteristiche fisiche principali della Λ_c^+ sono un contenuto di quark udc , una massa di $(2286.46 \pm 0.14)MeV/c^2$, un $I(J^P) = 0(1/2^+)$ e una vita media di $(2.024 \pm 0.031) \times 10^{-13}s$. La Λ_c^+ possiede diversi canali di decadimento, ma in ALICE se ne analizzano tre, due adronici e uno semileptonico:

$$\Lambda_c^+ \rightarrow pK^-\pi^+ \quad (3.1)$$

$$\Lambda_c^+ \rightarrow pK_S^0 \quad (3.2)$$

$$\Lambda_c^+ \rightarrow \Lambda e^+\bar{\nu}_e \quad (3.3)$$

I Branching Ratio di questi canali di decadimento (ossia la frazione di Λ_c^+ che decade in quello specifico canale) sono $(6.28 \pm 0.32)\%$ per il canale 3.1, $(1.59 \pm 0.08)\%$ per il 3.2, $(3.6 \pm 0.4)\%$ per il 3.3.

In questa tesi è stato analizzato solamente il secondo canale di decadimento, schematizzato in Figura 3.1.

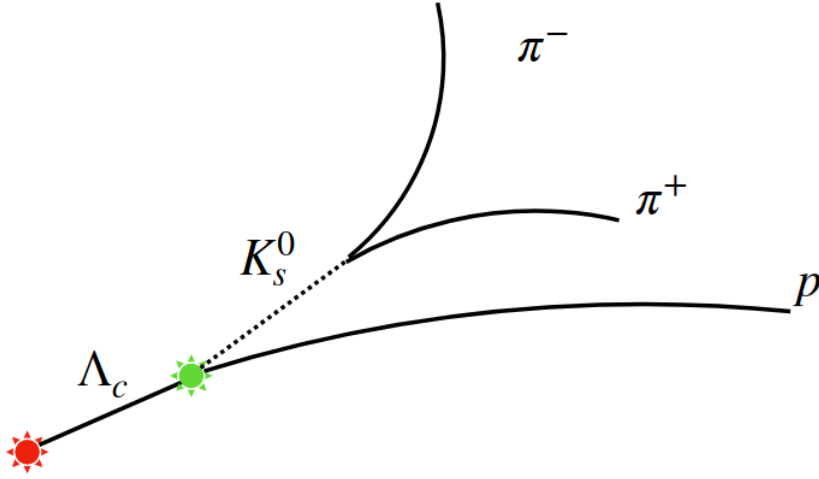


Figura 3.1: Schema del decadimento del barione Λ_c^+ secondo il canale di decadimento $\Lambda_c^+ \rightarrow pK_s^0$.

Come si vede dall'immagine, il barione Λ_c^+ è prodotto dall'urto tra protoni nel punto rosso, detto *vertice primario*, poi decade molto rapidamente per interazione debole nel punto verde, detto *vertice secondario*. Dopodiché il mesone K_s^0 , che a sua volta è instabile, decade per decadimento debole in due pioni carichi, e sono questi ultimi che vengono in effetti rilevati dall'ITS di ALICE.

La difficoltà più grande che si incontra nell'analisi di questo decadimento è che, data la sua brevissima vita media, il barione Λ_c^+ decade mediamente dopo avere percorso una distanza di $c\tau = 60 \mu m$, ben più piccola della precisione dell'ITS che è circa $100 \mu m$ per impulsi trasversi $p_T \approx 1 GeV/c$, rendendo impossibile la distinzione tra vertice primario e secondario. Questo fatto complica notevolmente l'analisi: dato che non è possibile distinguere particelle provenienti dal vertice primario e secondario, bisogna attuare dei metodi più sofisticati per separare le particelle vere di Λ_c^+ dal fondo, costituito da tutte le possibili combinazioni di particelle non provenienti dal decadimento di una Λ_c^+ , ma aventi caratteristiche simili a quelle che lo sono. A tale proposito può risultare conveniente utilizzare metodi di analisi che coinvolgono machine learning. In questo lavoro è stato utilizzato un metodo di analisi multivariata fornito dal pacchetto TMVA, descritto nella sezione successiva.

3.2 TMVA

TMVA (Toolkit for MultiVariate data Analysis) [20] è un pacchetto integrato nel framework di elaborazione dati ROOT che fornisce metodi avanzati di analisi multivariata. È nato con lo scopo di lavorare nel campo della fisica delle alte energie, ma può essere utilizzato per qualsiasi scopo che richieda l'applicazione di metodi multivariati. In particolare questi metodi sono specializzati nella previsione, a partire da una serie di variabili in input, di un output numerico continuo o di una classe di appartenenza. Nel primo caso si tratta di algoritmi di *regressione*, e nel secondo caso di algoritmi di *classificazione*. Tutte le tecniche di TMVA fanno parte della famiglia di algoritmi di *supervised learning*, ossia algoritmi che imparano da soli a classificare eventi in una fase di training in cui vengono utilizzati dati il cui output è già conosciuto. In generale si fornisce al metodo un sample di dati il cui esito è noto, il quale viene diviso in due parti; la prima viene utilizzata nella fase di training per calcolare tutti i fattori peso necessari all'apprendimento; la seconda viene usata nella fase di testing in cui l'algoritmo coi fattori peso viene applicato sulla parte restante del sample, al fine di controllare che la fase di training abbia portato al risultato corretto.

Nel caso della ricostruzione del barione Λ_c^+ , il TMVA è stato utilizzato per la classificazione di ciascun evento in segnale o fondo. Per evento si intende una combinazione di particelle, dette "candidate", aventi proprietà (le variabili in input) in un range compatibile con il decadimento di una particella Λ_c^+ . Fanno parte del segnale le candidate che vengono effettivamente riconosciute come decadimenti di Λ_c^+ , mentre fanno parte del fondo tutte quelle che casualmente si trovano nel range compatibile, ma non sono provenienti da un vero decadimento di Λ_c^+ .

3.3 Boosted Decision Trees

L'algoritmo di classificazione utilizzato da TMVA in questa tesi è di tipo BDT, dunque fa uso di *boosted decision trees*. I mattoncini fondamentali di un algoritmo BDT sono i *decision trees* (DT). Un DT è un semplice classificatore, incaricato di porre domande binarie alle variabili di ingresso \mathbf{x} . Ciò è equivalente a ripartire in varie regioni con condizioni diverse lo spazio delle fasi delle variabili di ingresso [21]. In Figura 3.2 è rappresentato schematicamente un semplice *decision tree*, mettendo in risalto il fatto che per ogni richiesta che viene fatta alle variabili di ingresso, esso si ramifica in due, e ciò corrisponde ad una suddivisione in regioni distinte dello spazio delle fasi. Un singolo DT è una struttura di per sé troppo semplice, e la sua applicazione ad un campione di dati non porta a risultati di classificazione soddisfacenti. È per questo che si usano algoritmi di *boosting*, i quali generano un grande numero di tree che insieme formano strutture più complicate, dette foreste. Il risultato finale dunque non sarà più una semplice classificazione discreta, bensì una media pesata delle risposte dei singoli tree.

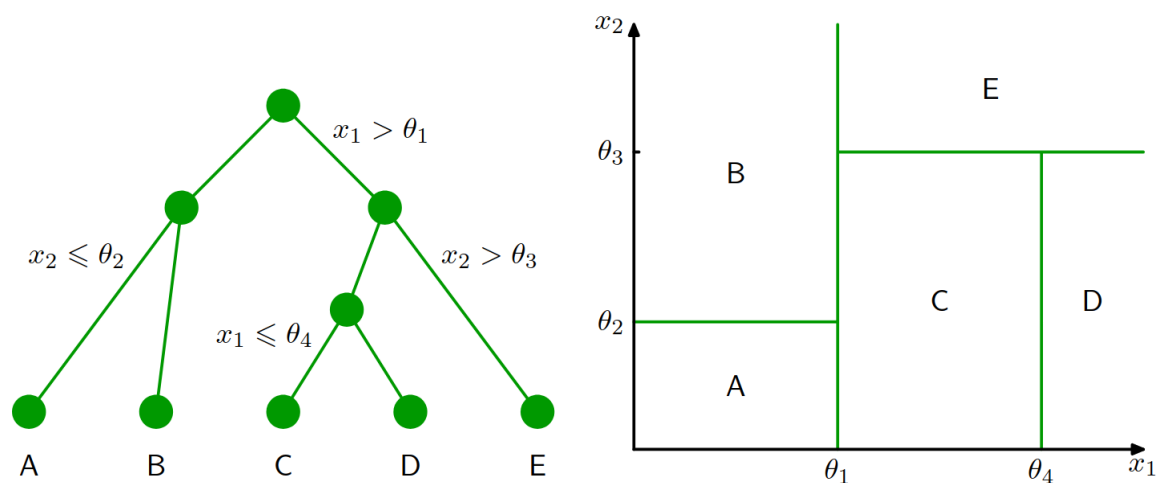


Figura 3.2: Schema di un semplice Decision Tree di un sistema con sole due variabili di ingresso. A sinistra la rappresentazione della successione di domande binarie, a destra la corrispondente partizione dello spazio delle fasi.

3.4 Dati e Variabili in input

Il campione di dati utilizzato per gli eventi di segnale è costituito da eventi simulati dal generatore Monte Carlo PYTHIA8 [22]. In queste simulazioni, per aumentare la statistica, vi è stato posto come vincolo il fatto che ci sia almeno una Λ_c^+ che decade secondo il canale di decadimento di nostro interesse, tra tutte quelle che vengono generate. Per ottenere una simulazione ancora più fedele, in queste simulazioni vengono anche inseriti i rivelatori che compongono l'esperimento ALICE, in modo da simulare nel modo più verosimile possibile le interazioni delle particelle con i materiali di cui sono costituiti tali rivelatori, nonché la formazione dei loro segnali di risposta. La simulazione del rivelatore ALICE e la propagazione delle particelle viene fatta utilizzando il software GEANT3 [23].

Il campione di dati utilizzato per il background invece è stato ottenuto dalle misure stesse di ALICE, prese nella seconda fase di presa dati (RUN 2), avvenuta negli anni 2016-2017-2018. Negli eventi dei dati reali, sono però state selezionate solo candidate con una massa invariante ricostruita non compatibile con la massa di una vera Λ_c^+ .

Il Machine Learning (ML) è lo studio di algoritmi capaci di imitare o trovare patterns nel campione di training attraverso l'esperienza. Un algoritmo di ML prende in input un insieme di variabili e applica delle funzioni per distinguerle nelle categorie conosciute. Si capisce dunque che la scelta di queste variabili di input è di grande importanza, e una buona scelta di queste permette all'algoritmo di separare le classi di eventi (in segnale e background) in una maniera più efficiente.

Prendendo come riferimento la Figura 3.1, si parlerà di potenziale protone come particella "bachelor", e di potenziale K_S^0 come particella V^0 . Allora queste sono le variabili selezionate come input:

- **massK0S**: la massa invariante della particella V^0 , ottenuta a partire dalle tracce ricostruite delle due figlie. Nel caso di decadimenti di Λ_c^+ ci si aspetta che questa massa sia compatibile con quella della particella di K_S^0 , cioè circa $497 \text{ MeV}/c^2$.
- **tImpParBach**: parametro d'impatto della particella bachelor, definito come la distanza minima tra la traccia del bachelor e il vertice primario, nel piano normale alla direzione dei fasci collidenti.
- **tImpParV0**: parametro d'impatto della particella V^0 .
- **ctK0S**: $c\tau$ della particella V^0 , aspettato per la K_S^0 a circa 2.68 cm .

- **cosPAK0S**: coseno dell'angolo tra la direzione della particella V^0 e la congiungente tra il vertice primario e il secondario. Il suo valore atteso è vicino all'unità.
- **CosThetaStar**: coseno dell'angolo tra la direzione della particella bachelor nel sistema di riferimento dove Λ_c^+ è a riposo con la direzione di Λ_c^+ in laboratorio. Il decadimento dovrebbe essere isotropo, pertanto è attesa per questa variabile una distribuzione uniforme nel dominio $[-1, 1]$ per il segnale, e valori vicini a 1 o -1 per il fondo. Questa variabile fa parte del Tree che immagazzina i dati in input, ma non è stata utilizzata nel training perché è sospetta di rendere la procedura di estrazione del segnale più difficile e instabile.
- **nSigmaTOFpr**: probabilità, in unità di deviazioni standard, che la particella bachelor sia effettivamente un protone secondo il rivelatore TOF. Il suo valore è ottenuto paragonando il tempo che impiega la particella bachelor a raggiungere TOF, rispetto a quello che impiega un vero protone.
- **nSigmaTOFpi**, **nSigmaTOFka**: similmente alla precedente variabile, le probabilità che la particella bachelor sia un pione o un kaone secondo il rivelatore TOF.
- **nSigmaTPCpr**, **nSigmaTPCpi**, **nSigmaTPCka**: probabilità, in unità di deviazioni standard, che la particella bachelor sia un protone, un pione, o un kaone secondo il rivelatore TPC. Il suo valore è ottenuto paragonando l'energia persa dE/dx della particella bachelor, rispetto a quella persa da veri protoni, pioni o kaoni.

In Figura 3.3 è possibile vedere le distribuzioni di segnale e fondo di tutte le variabili in input per candidate Λ_c^+ ricostruite con un impulso trasverso nel range $0 < p_T < 1 \text{ GeV}/c$. I casi in cui **nSigmaTOFpr**, **nSigmaTOFpi** a **nSigmaTOFka** valgono -999 corrispondono ad eventi in cui, per motivi di efficienza o di accettazione geometrica, la particella bachelor non ha rilasciato alcun segnale sul rivelatore TOF.

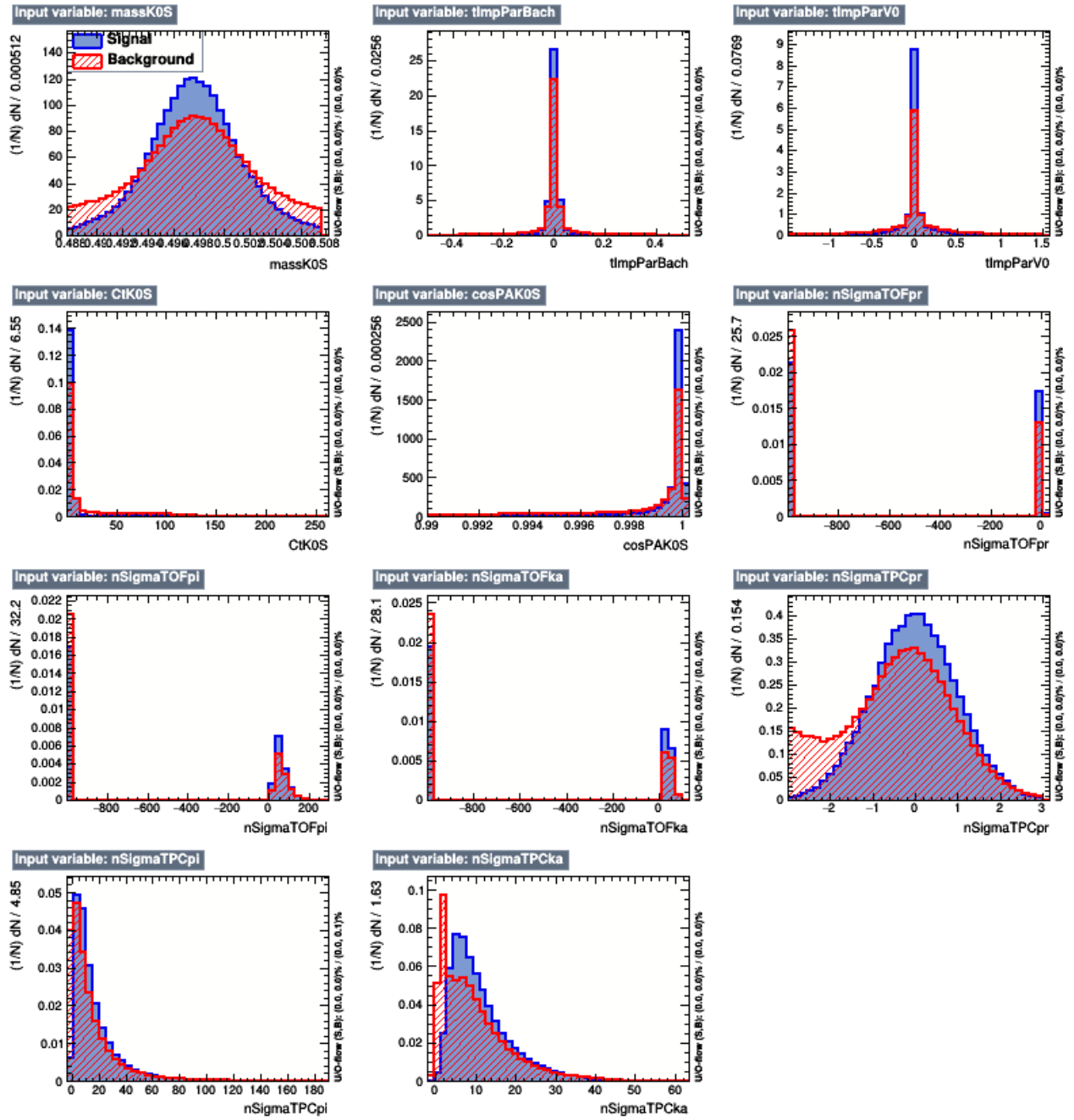


Figura 3.3: Distribuzioni delle variabili di input nel range $0 < p_T < 1 \text{ GeV}/c$ classificati come segnale (blu) o fondo (rosso).

3.5 Training del metodo BDT

Il training del metodo BDT è stato effettuato utilizzando i valori ed i settaggi riportati in Tabella 3.1 per i parametri principali di configurazione.

Parametro	Valore	Descrizione
NTrees	850	Numero di trees nella foresta
MaxDepth	3	Massima profondità del decision tree permessa
MinNodeSize	2.5%	Minima percentuale di eventi di training richiesti in un nodo
BoostType	AdaBoost	Tipo di algoritmo di boosting per generare la foresta
AdaBoostBeta	0.5	Rateo di apprendimento per l'algoritmo AdaBoost
BaggedSampleFraction	0.6	Grandezza relativa del campione di eventi bagged rispetto a quella originale del campione di dati
SeparationType	GiniIndex	Criterio di separazione per lo splitting tra i nodi
nCuts	20	Numero di punti utilizzati per trovare il taglio ottimale nella splitting mode

Tabella 3.1: Parametri di configurazione utilizzati nell'algoritmo BDT.

Prima di eseguire il training vero e proprio, il software TMVA produce le matrici di correlazione lineare delle variabili in input, sia per il segnale, sia per il fondo (Figura 3.4). Lo studio di queste matrici è importante perché anche se variabili altamente correlate non compromettono il corretto funzionamento del BDT, il loro utilizzo può comportare un aumento nel tempo di processamento del training, che dipende tra le altre cose dal numero di variabili in input. Pertanto variabili altamente correlate risultano quanto meno superflue, ed è buona norma evitarle. In questo caso, la correlazione tra le variabili di PID è aspettata, e non crea un problema.

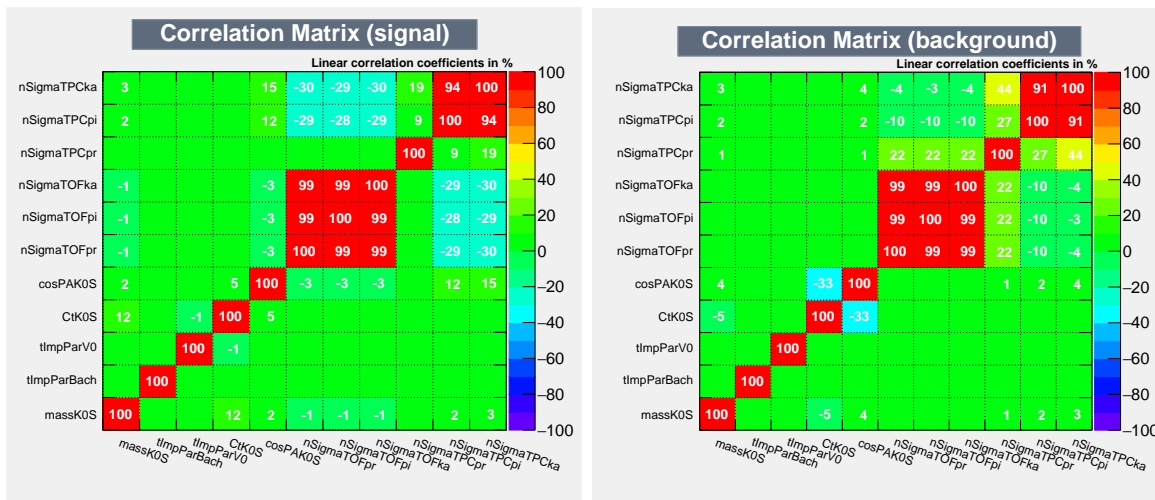


Figura 3.4: Matrici di correlazione lineare delle variabili in input a sinistra per il segnale e a destra per il fondo.

3.6 Testing e Grafici di Controllo

La seconda metà del campione di dati in ingresso, come già detto, è utilizzata nella fase di testing. Agli eventi di questa frazione di dati, TMVA applica i fattori peso calcolati nella fase di training, classificandoli in segnale o fondo, poi verifica che l'esito sia uguale alla loro classe a priori. In questa fase vengono prodotti diversi grafici di controllo, che permettono di verificare e visualizzare il buon esito del training.

3.6.1 Test dell'overtraining

Il primo grafico generato è la distribuzione di risposta per i dati di training e di testing, riportato in Figura 3.5. Confrontare tali distribuzioni è importante per controllare se l'algoritmo non ha subito il fenomeno dell'*overtraining*. Un algoritmo si dice *overtrained* quando nella fase di training interpreta fluttuazioni statistiche come patterns significativi. Questo permette all'algoritmo di classificare molto bene il campione su cui ha eseguito il training, ma di non essere altrettanto efficace quando viene applicato su altri campioni statisticamente non correlati con quello di training. La presenza o meno di questi bias può essere verificata utilizzando il campione di testing. In Figura 3.5 si vede che le distribuzioni del segnale di risposta del BDT per il segnale e per il fondo sono molto simili tra training e testing sample, per cui l'ipotesi di overtraining è esclusa. La Figura 3.5 permette anche di capire intuitivamente come funziona la selezione delle candidate attraverso il metodo BDT. L'algoritmo fornisce in uscita una singola funzione chiamata BDT output score o classificatore. Tale funzione ha un range finito $[-1,1]$, i cui estremi

corrispondono ai casi in cui il BDT ha la certezza (ossia una probabilità pari a 100%) che la candidata sia rispettivamente fondo o segnale. Fissato un certo valore del classificatore, quindi, è possibile effettuare una selezione eliminando tutte le candidate che si trovano a sinistra del valore di taglio (BDT output < cut value, ovvero candidate *background-like*) e selezionando solo quelle a destra (BDT output > cut value, ovvero candidate *signal-like*). Di norma, questa selezione rigetterà anche un certo numero di vere candidate; il numero di vere candidate selezionate sul totale presente nel training sample rappresenta l'efficienza del BDT cut scelto.

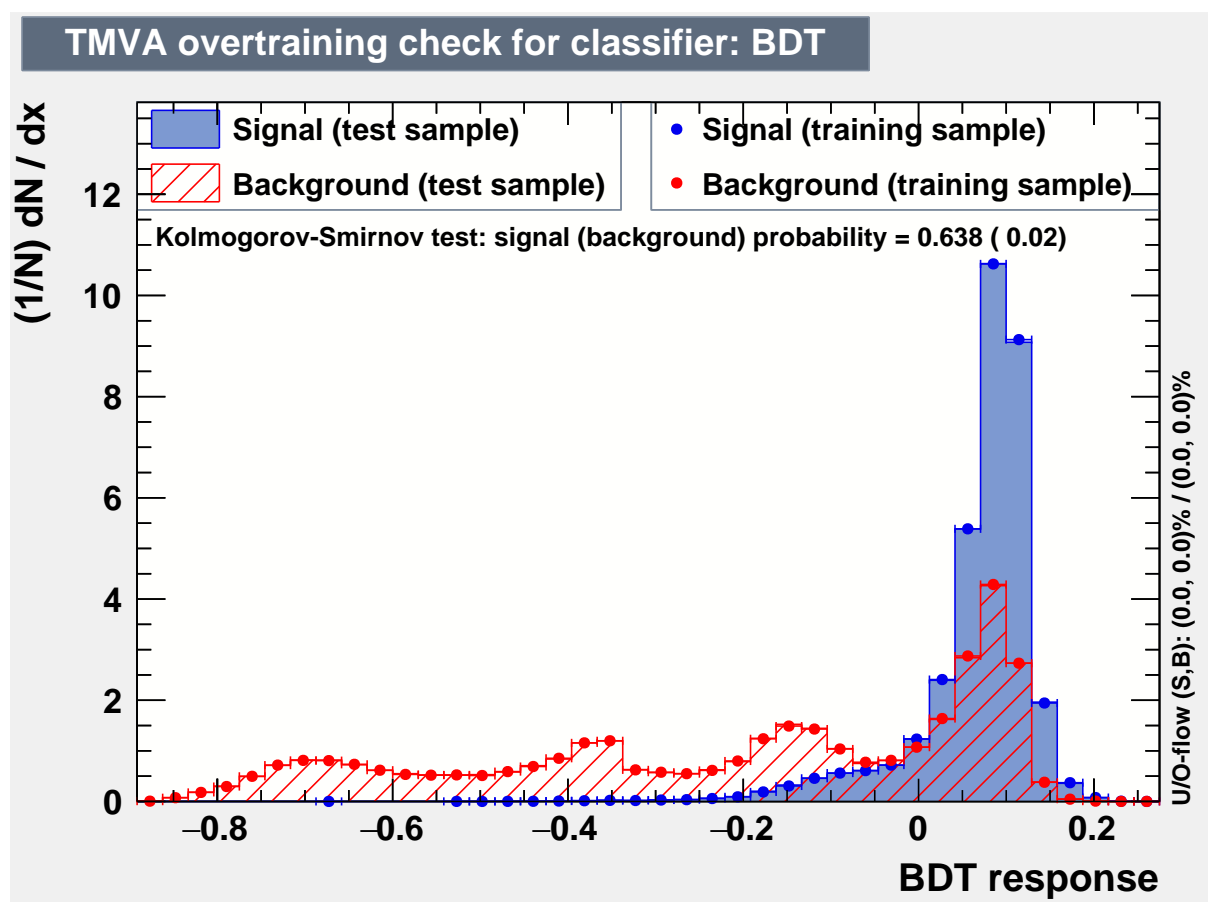


Figura 3.5: BDT output score dei dati in input per il testing e il training sample, sia segnale che fondo.

3.6.2 Curva di ROC

Le curve *receiver operating characteristic* (curve ROC), sono grafici che mostrano il rigetto del fondo in funzione dell'efficienza di segnale. La curva ROC per un classificatore ideale sarebbe rettangolare, con un rigetto del fondo del 100%, un'efficienza di segnale del 100% e un integrale pari a 1. Più la curva ROC del metodo utilizzato si avvicina alla curva ideale, migliore sarà la sua performance.

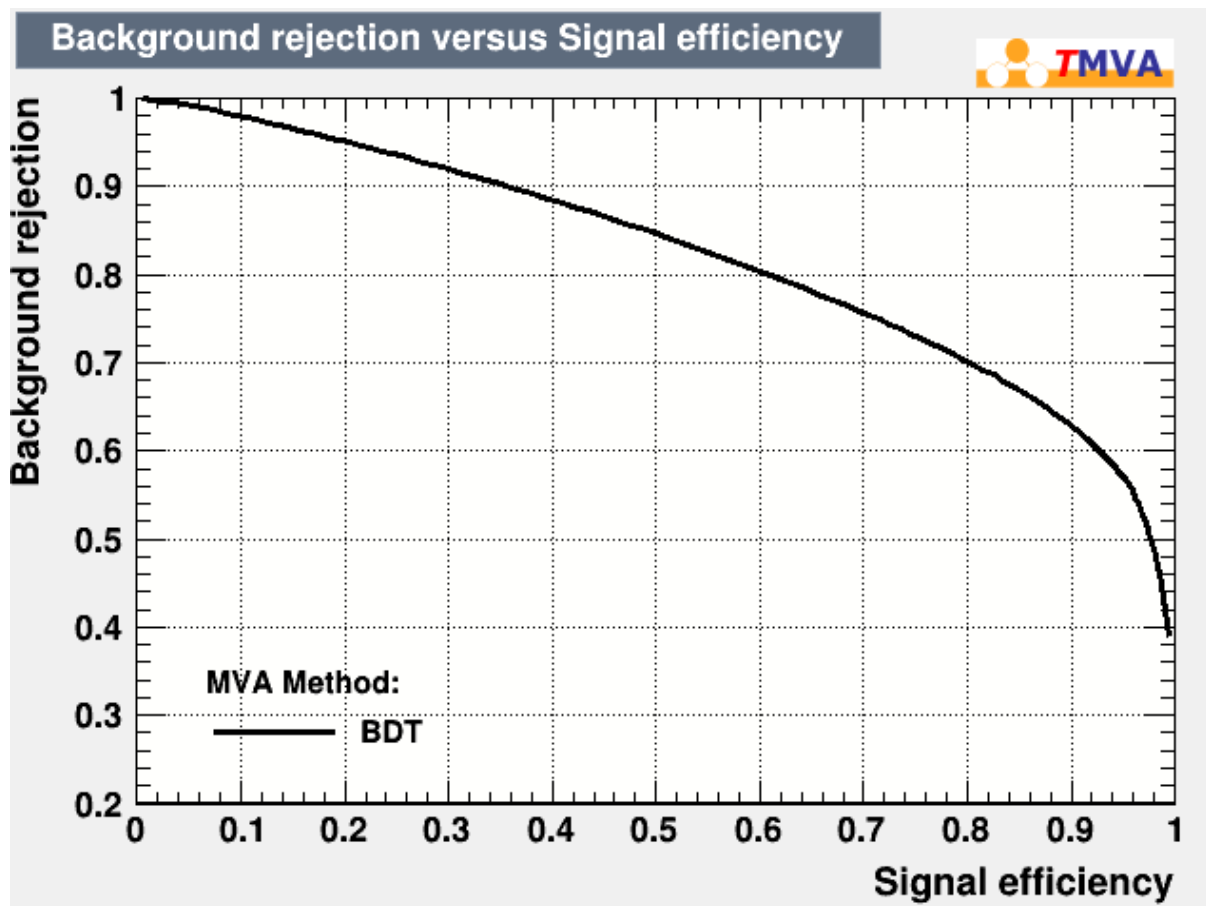


Figura 3.6: Curva di ROC nel range $0 < p_T < 1 \text{ GeV}/c$.

In Figura 3.6 è riportata la curva di ROC dell'algoritmo utilizzato nel range di impulso trasverso $[0, 1] \text{ GeV}/c$. L'integrale di questa curva è 0.821.

3.6.3 Ranking delle variabili

TMVA fornisce anche due ranking delle variabili in input, il primo parametrizzato dalla variabile *separation*, e il secondo dalla variabile *importance*. La prima fornisce una stima del potere di separazione segnale da fondo (*separation power*) che ogni variabile in input ha se si effettua un taglio su di essa, presa singolarmente; la seconda testimonia l'importanza che una data variabile ha avuto nel processo di splitting dei nodi da parte del BDT nella fase di training dell'algoritmo. Nelle Tabelle 3.2 e 3.3 sono riportate rispettivamente la distribuzione di ranking secondo le variabili *separation* e *importance*.

rango	variabile	separation
1	CtK0S	1.447e-01
2	cosPAK0S	9.824e-02
3	tImpParV0	9.291e-02
4	nSigmaTPCka	8.417e-02
5	nSigmaTPCpi	7.178e-02
6	tImpParBach	5.827e-02
7	nSigmaTPCpr	5.747e-02
8	massK0S	3.147e-02
9	nSigmaTOFpi	1.589e-02
10	nSigmaTOFka	1.525e-02
11	nSigmaTOFpr	1.372e-02

Tabella 3.2: Ranking delle undici variabili in input secondo la *separation* nel range di impulso trasverso $0 < p_T < 1$ (GeV/c).

rango	variabile	importance
1	CtK0S	2.614e-01
2	cosPAK0S	1.115e-01
3	nSigmaTPCka	1.085e-01
4	nSigmaTPCpr	8.985e-02
5	nSigmaTPCpi	8.201e-02
6	tImpParBach	7.678e-02
7	tImpParV0	6.632e-02
8	nSigmaTOFpr	5.873e-02
9	massK0S	5.446e-02
10	nSigmaTOFka	4.657e-02
11	nSigmaTOFpi	4.385e-02

Tabella 3.3: Ranking delle undici variabili in input secondo la *importance* nel range di impulso trasverso $0 < p_T < 1$ (GeV/c).

3.7 Metodo di estrazione del segnale con polinomiale di terzo grado

Per calcolare la probabilità con cui un barione Λ_c^+ viene prodotto in collisioni pp ad alta energia, è necessario misurare quante di tali particelle sono presenti nei nostri dati. Questa misura è resa estremamente difficile dall'elevato fondo combinatoriale, in particolare a bassi impulsi trasversi della candidata Λ_c^+ . Il primo passo per ricostruire il numero di particelle consiste dunque nell'eliminare la quantità maggiore possibile di fondo, effettuando un taglio sulla variabile di risposta del BDT, eliminando gli eventi caratterizzati da un BDT response sotto il valore di soglia e tenendo solo quelli al di sopra. TMVA fornisce un calcolo accurato del taglio ottimale per ottimizzare la *significance*, definita come $S/\sqrt{S+B}$, riportato in Figura 3.7. In questa tesi si è scelto di usare il taglio a $BDT = 0$, molto vicino a quello consigliato dal TMVA.

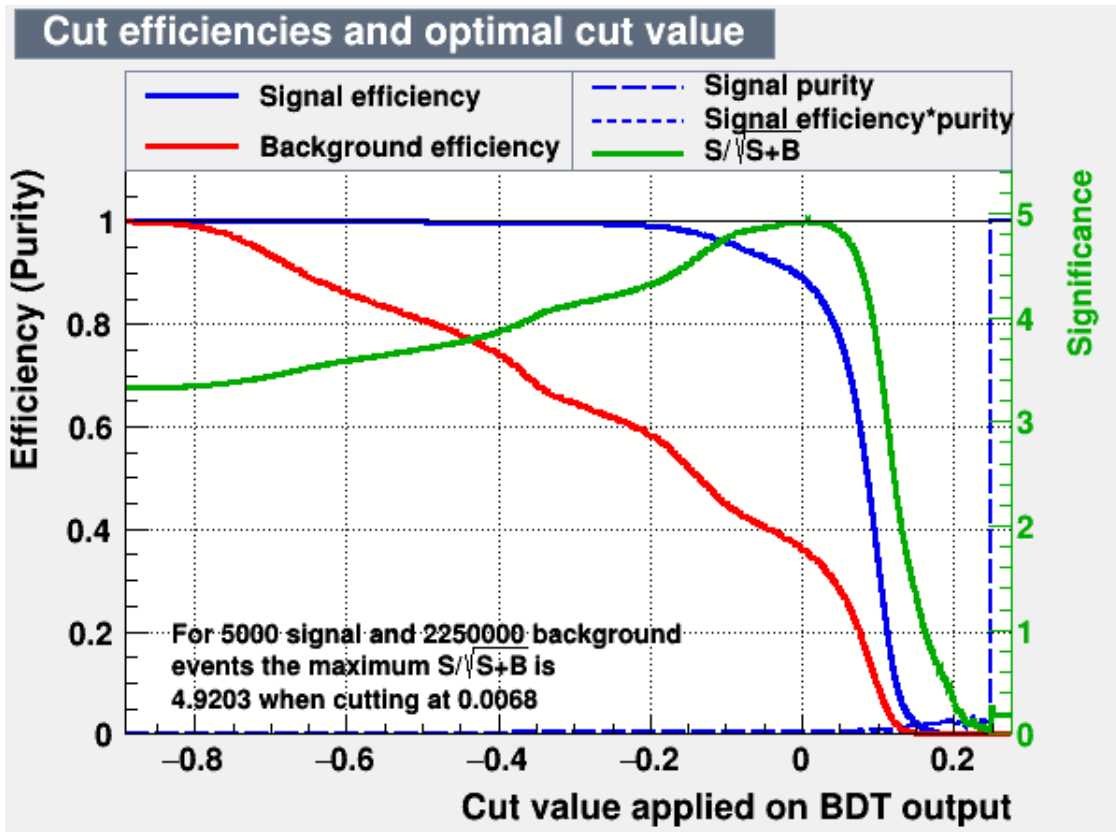


Figura 3.7: Efficienze, purezze e significance per ogni taglio possibile della variabile BDT nel range di impulsi $0 < p_T < 1 (GeV/c)$. Tutto è calcolato per una quantità di eventi di segnale dello stesso ordine di grandezza del valore aspettato dello yield e per una quantità di eventi di fondo pari al numero totale di candidate relative al range stabilito.

Ora consideriamo l'istogramma di massa invariante ottenuto dopo aver applicato il taglio sulla variabile BDT response. Per eliminare la restante porzione di fondo è stato eseguito un fit polinomiale di terzo grado dell'istogramma di Figura 3.8, esclusa la parte di dominio nel range entro $\pm 3\sigma$ dalla massa teorica della Λ_c^+ , dove σ rappresenta la larghezza del picco della Λ_c^+ , dovuta principalmente alla risoluzione del tracciamento in ALICE. Infine, per il conteggio delle Λ_c^+ ricostruite è stato effettuato un fit gaussiano dei residui, definiti come la differenza, punto per punto, tra l'istogramma di massa invariante e la funzione di fit del fondo, estrapolata sul tutto il range dell'istogramma, visibile in Figura 3.9. Infine da questa analisi è possibile estrarre lo *yield*, cioè il numero effettivo di barioni Λ_c^+ presente nei nostri dati, semplicemente calcolando l'integrale della gaussiana rappresentata dalla curva rossa in Figura 3.9. Il calcolo dello *yield* ha prodotto il seguente risultato, dove l'errore riportato è l'errore statistico della misura effettuata:

$$N(\Lambda_c^+) = (3939 \pm 839)$$

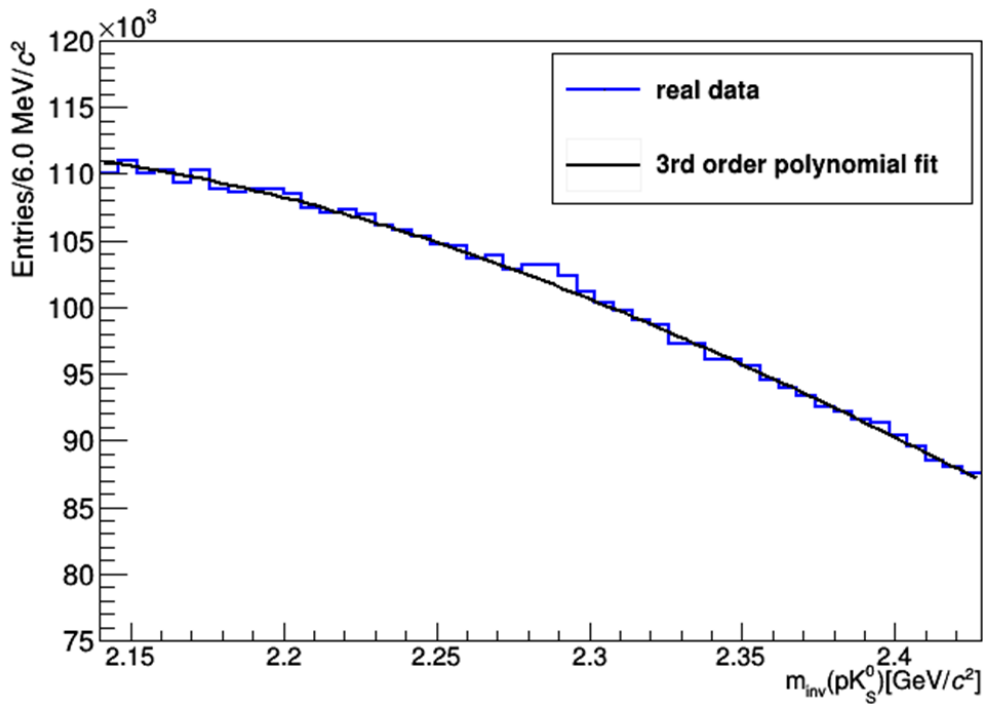


Figura 3.8: Fondo statistico con la forma di una curva polinomiale di terzo grado sovrapposto ai dati sperimentali.

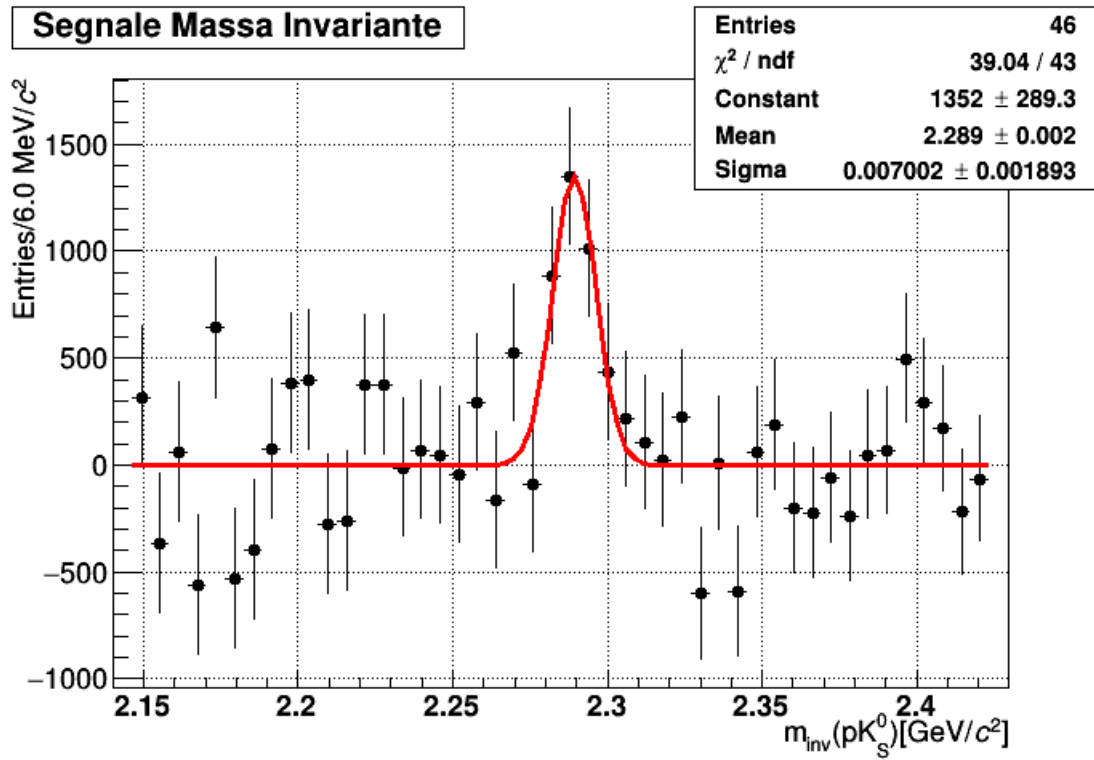


Figura 3.9: Istogramma dei residui di massa invariante ricostruita nel range di impulsi $[0, 1] \text{ GeV}/c$, con relativo fit gaussiano.

3.8 Metodo alternativo con template rotazionale

Modellare il fondo statistico con una funzione analitica è un approccio standard nell'analisi dati in fisica delle particelle; tuttavia, in presenza di un fondo combinatoriale e soprattutto quando il rapporto segnale su fondo è molto basso e gli errori statistici significativi, come nel caso della nostra analisi, questo approccio potrebbe non essere privo di bias e influenzare pesantemente il risultato finale. Per questo motivo, in questa tesi è stato sviluppato un approccio alternativo per descrivere la forma del fondo con tecniche di tipo non-analitico. In particolare, è stato utilizzato un template "rotazionale" con candidate create artificialmente, sviluppato nel seguente modo.

Per ogni candidata Λ_c^+ ricostruita, è stato cambiato l'angolo tra gli impulsi nel piano trasverso del protone e della K_S^0 ricostruiti, prendendolo da una distribuzione uniforme nel range $[0, 2\pi]$ (visualizzazione in Figura 3.10). Questa procedura è stata ripetuta dieci volte, imponendo come unica condizione che la massa invariante ricostruita fosse ancora compatibile con quella di un barione Λ_c^+ e ricalcolando per ogni ricorrenza l'impulso trasverso della particella ricostruita.

Attraverso questa procedura vi è la certezza che le tracce ruotate del protone e della K_S^0 siano totalmente scorrelate, dunque la particella ricostruita a partire da essi non sarà una vera Λ_c^+ , e l'evento in questione è da classificare come un evento di fondo.

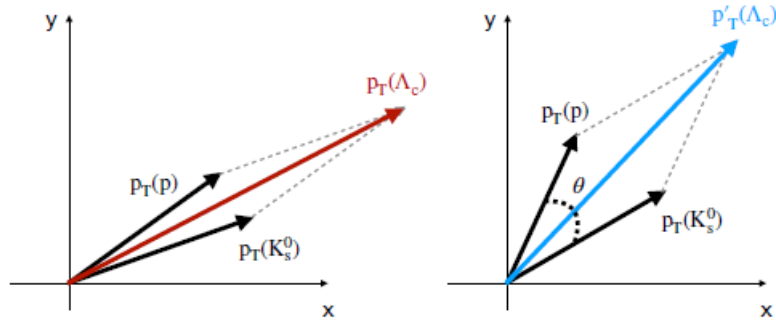


Figura 3.10: Visualizzazione della rotazione nel piano trasverso degli impulsi tra protoni ricostruiti e K_S^0 ricostruite.

A questo punto, a partire dal tree contenente i dati raccolti da ALICE, e in cui sono state salvate tutte le informazioni e le misure relative alla particella bachelor e alla V^0 , è stato costruito un nuovo tree sostituendo le misure relative alle candidate Λ_c^+ reali ricostruite con quelle generate attraverso la tecnica rotazionale. Successivamente, è stato applicato a questo nuovo tree il modello BDT ottenuto in precedenza, ed è stata effettuata la stessa selezione (taglio per valori del classificatore minori di zero) utilizzata nell'analisi con il metodo standard. Il vantaggio di utilizzare questo ulteriore step è duplice:

- da un lato, applicando lo stesso taglio sui dati e sul template, siamo sicuri di confrontare due campioni di dati "topologicamente" compatibili, in quanto il taglio sulle variabili topologiche (parametro d'impatto per protone e K_S^0 , direzione di volo del K_S^0 rispetto alla congiungente vertice primario-secondario) e di PID dovrebbe essere simile se non identico nei due casi;
- d'altra parte, se per effetto di un piccolo overtraining il modello tendesse a selezionare più candidate Λ_c^+ in un intervallo vicino al valore teorico della massa del barione, riconoscendo erroneamente alcuni eventi di fondo come vere Λ_c^+ , un eccesso di candidate in corrispondenza del picco dovrebbe manifestarsi sia dopo aver applicato il modello ai dati reali che al nostro template. Questo effetto non può essere identificato, risultando in tal modo non correggibile, con l'approccio standard. Con questo metodo alternativo invece, effettuando una sottrazione del fondo rotazionale bin per bin dai dati sperimentali, l'effetto verrebbe completamente cancellato.

In Figura 3.11 è mostrato il confronto tra la distribuzione in massa invariante delle candidate Λ_c^+ per dati reali, riscalati di un fattore 10, e per il template rotazionale.

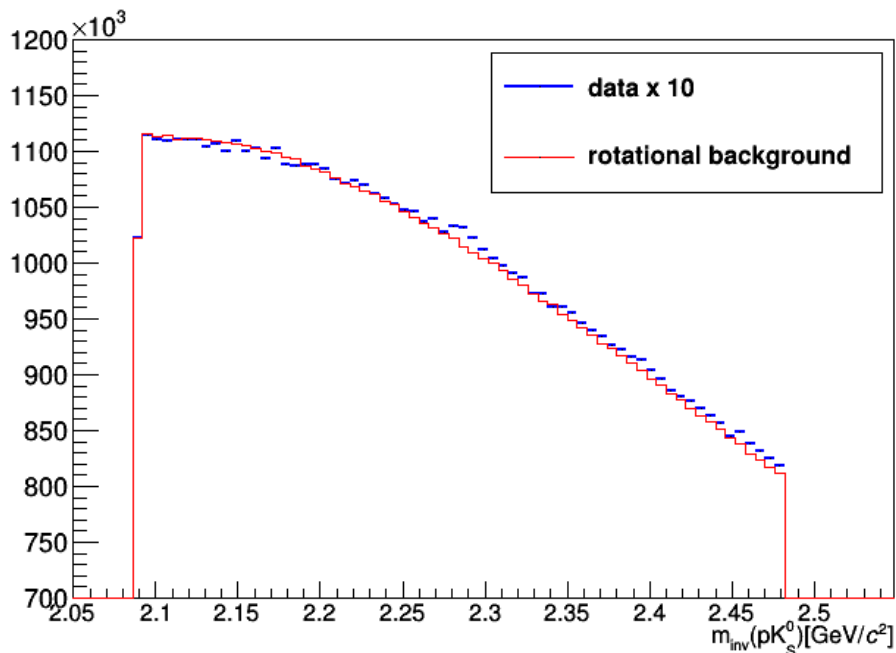


Figura 3.11: Template rotazionale sovrapposto all'istogramma della massa invariante ricostruita aumentato di un fattore $\times 10$. È possibile apprezzare come il primo sia già inizialmente piuttosto aderente al secondo, anche prima di ulteriori riscalamenti.

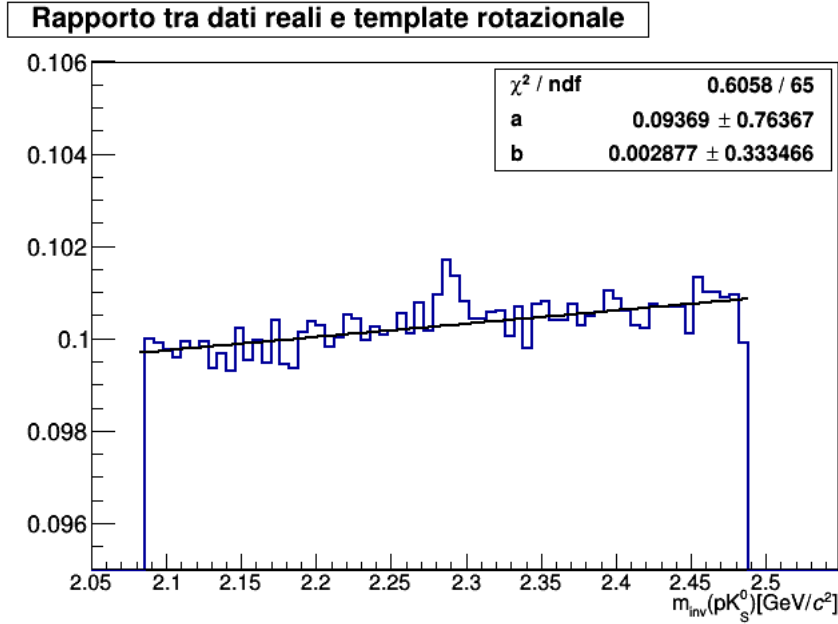


Figura 3.12: Istogramma del rapporto bin per bin tra la massa invariante ricostruita e il template rotazionale. È possibile apprezzare una lieve pendenza lineare e il fit fa riferimento alla funzione $y = a + bx$.

Come si può vedere, l'accordo tra le due distribuzioni è molto buono, anche se è possibile notare una piccola discrepanza, che risulta più visibile considerando il rapporto tra le due distribuzioni (Figura 3.12). Tale discrepanza è dovuta probabilmente al fatto che nei dati reali esiste un range di angoli tra gli impulsi trasversi del bachelor e della V^0 favorito cinematicamente rispetto agli altri valori, mentre nel nostro approccio tali angoli sono generati in maniera random e uniforme nel range $[0, 2\pi]$, e può essere parametrizzata con un trend lineare del rapporto dati/template in funzione della massa invariante ricostruita della Λ_c^+ . Il passo successivo è eseguire un fit lineare sull'istogramma del rapporto, escludendo il range $\pm 3\sigma$ attorno alla massa teorica della Λ_c^+ . Tale funzione lineare, estrapolata poi su tutto il range di massa invariante, è stata usata per correggere il template rotazionale effettuando la moltiplicazione bin per bin. Si noti che quest'ultima operazione funge anche da riscaldamento del template rotazionale, essendo esso in origine più grande di dieci volte rispetto ai dati reali, data la maniera in cui è stato costruito. Nella Figura 3.13, che mostra il fondo statistico riscaldato sovrapposto ai dati reali, è possibile apprezzare che la discrepanza è stata corretta. Infine è stato eseguito un fit gaussiano sulla distribuzione dei residui visibile in Figura 3.14, da cui è stato estratto lo yield calcolandone l'integrale:

$$N(\Lambda_c^+) = (4119 \pm 679)$$

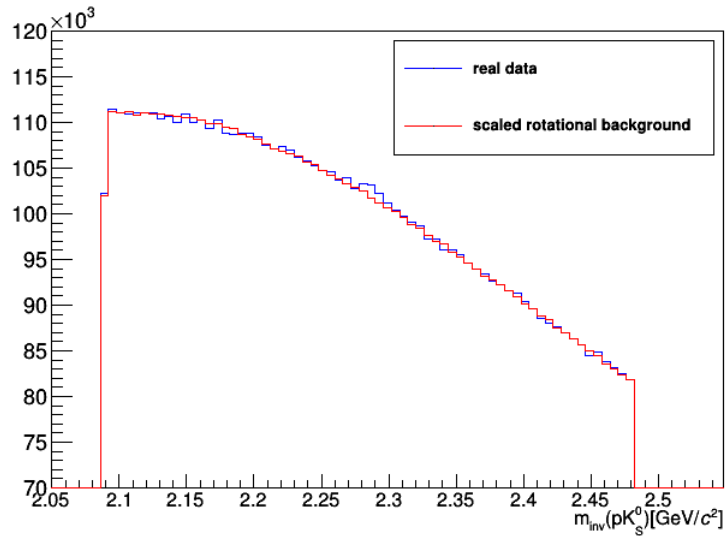


Figura 3.13: Fondo statistico riscaldato, sovrapposto ai dati reali. In questo grafico è possibile apprezzare una maggiore aderenza del fondo ai dati, rispetto alla Figura 3.11.

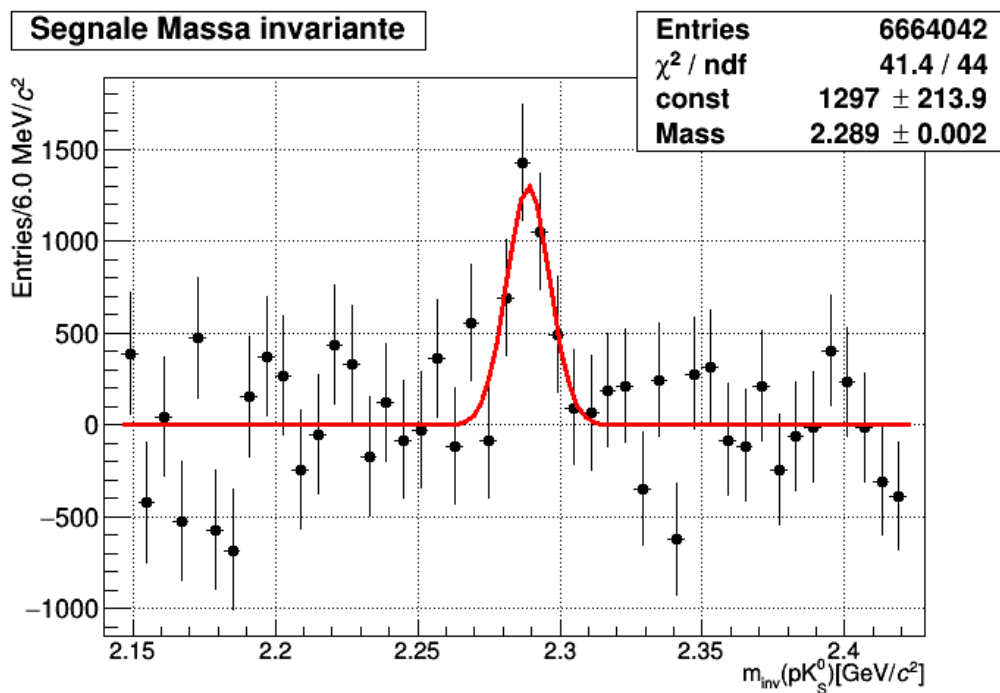


Figura 3.14: Istogramma dei residui di massa invariante ricostruita nel range di impulsi $[0, 1] \text{ GeV}/c$ attraverso l'analisi con fondo rotazionale, con relativo fit gaussiano.

3.9 Risultati

La Tabella 3.4 riassume il confronto tra gli yield ottenuti con i 2 metodi illustrati in questo lavoro di tesi.

p_T (GeV/c)	3 rd order polynomial fit	Rotational template
[0, 1]	3939 ± 839	4119 ± 679

Tabella 3.4: Segnale dello yield della Λ_c^+ per entrambi i metodi di analisi.

Come si può vedere, i risultati ottenuti sono perfettamente compatibili entro gli errori statistici. Questa compatibilità dimostra che il risultato ottenuto con l'analisi standard è attendibile e non affetto da bias dovuti ad una non ottimale sottrazione del fondo combinatoriale. Inoltre, essa rappresenta anche una validazione dell'approccio con template rotazionale e soprattutto dell'algoritmo utilizzato per la generazione del template stesso.

Conclusioni

L'esperimento ALICE ad LHC del CERN è dedicato allo studio delle proprietà della materia nucleare creata in collisioni di ioni pesanti ultrarelativistici: il Quark-Gluon Plasma. Tra le varie misure sperimentali, lo studio dei sapori di quark pesanti e dei meccanismi attraverso i quali essi si combinano con gli altri quark e anti-quark, per formare mesoni e barioni, ricopre un ruolo fondamentale. Recenti studi hanno mostrato come la produzione del barione charmato Λ_c^+ in collisioni pp a LHC differisce in modo sostanziale dalle previsioni dei modelli teorici basati sui risultati ottenuti nel passato con collisioni ep ed e^+e^- , dimostrando che la nostra comprensione del fenomeno dell'adronizzazione dei quark pesanti è attualmente fortemente incompleta. In risposta, nuovi modelli teorici e fenomenologici sono stati sviluppati con l'intento di riprodurre i dati sperimentali. Questi modelli sono per lo più compatibili tra di loro e in generale in grado di riprodurre il valore del rapporto barione-mesone Λ_c^+/D^0 per impulsi grandi e intermedi, mentre differiscono nelle loro previsioni per bassi valori di impulso trasverso. Pertanto l'analisi dati a basso p_T è di fondamentale importanza perché permette di discriminare i modelli che riproducono correttamente i dati sperimentali su tutto il dominio di p_T .

In questo lavoro è stato calcolato il numero di barioni Λ_c^+ (*yield*) prodotto in collisioni pp a $\sqrt{s} = 13 \text{ TeV}$ ricostruiti attraverso il canale di decadimento $\Lambda_c^+ \rightarrow pK_S^0$ nel range di impulsi trasversi $0 < p_T(\Lambda_c^+) < 1 \text{ GeV}/c$. L'analisi dati di questo decadimento è particolarmente difficile per vari motivi, tra cui la breve vita media di Λ_c^+ , la statistica non elevata e il basso rapporto segnale su fondo, quest'ultimo particolarmente importante a bassi p_T . Pertanto è indispensabile l'utilizzo di software di machine learning in grado di rigettare una parte considerevole del fondo combinatoriale mantenendo allo stesso tempo una buona efficienza di selezione. In questo lavoro è stato usato il software di analisi multivariata TMVA e algoritmi di machine learning di tipo BDT. L'estrazione dello yield è stata effettuata utilizzando due differenti approcci per la rimozione del fondo residuo: il primo metodo consiste nel modellare il fondo combinatoriale con una funzione analitica, in questo caso particolare con una curva polinomiale di terzo grado, mentre il secondo prevede un template combinatorio creato artificialmente attraverso la tecnica delle track rotations illustrata in dettaglio nel lavoro di tesi.

Il calcolo dello yield ha prodotto risultati compatibili nei due metodi indipendenti. L'approccio di modellare il fondo statistico con una funzione analitica è una metodologia

standard nell'analisi dati in fisica delle particelle, a patto che la forma del fondo sia ben riproducibile da tale funzione. Un approccio che fa uso di tecniche non-analitiche invece è teoricamente applicabile per qualsiasi forma del fondo, una volta individuata la procedura corretta per la creazione del template. La compatibilità tra i risultati ottenuti con i due metodi valida sia il valore dello yield ottenuto con l'approccio analitico, sia l'intero metodo di estrazione del segnale con template rotazionale sviluppato in questa tesi, che potrà essere utilizzato anche in altre analisi in ALICE nel caso in cui l'approccio analitico dovesse risultare non applicabile.

Bibliografia

- [1] W. N. Cottingham and D. A. Greenwood, *An introduction to the Standard Model of Particle Physics*, Cambridge University Press, 2007.
- [2] S. Braibant, G. Giacomelli and M. Spurio, *Particles and Fundamental Interactions: An Introduction to Particle Physics*, Springer, 2012.
- [3] Cheuk-Yin Wong. *Introduction to High-Energy Heavy-Ion Collisions*. World Scientific, 1994
- [4] *Lectures on high-energy heavy-ion collisions at the LHC*. C.A. Salgado. Departamento de Física de Partículas and IGFAE. Universidade de Santiago de Compostela, Santiago de Compostela, Spain
- [5] A. Andronic, *An overview of the experimental study of quark-gluon matter in high-energy nucleus-nucleus collisions*, [arXiv:nucl-ex/1407.5003], 2014.
- [6] John C. Collins, Davison E. Soper and George Sterman. *Factorization of Hard Processes in QCD*. 2004. arXiv: 0409313 [hep-ph].
- [7] Ramona Vogt. *Ultrarelativistic Heavy-Ion collisions*. Elsevier Science, 2007.
- [8] ALICE Collaboration, JHEP 04 (2018) 108, arXiv:1712.09581 [nucl-ex]
- [9] ALICE Collaboration, Phys. Rev. C 104 (2021) 054905, arXiv:2011.06079 [nucl-ex]
- [10] ALICE Collaboration, Phys. Rev. Lett. 127 (2021) 202301, arXiv:2011.06078 [nucl-ex]
- [11] ALICE Collaboration, Phys. Rev. Lett. 128 (2022) 012001, arXiv:2106.08278 [hep-ex]
- [12] P. Skands, S. Carrazza and J. Rojo, Eur. Phys. J. **C74**, 3024 (2014), arXiv:1404.5630 [hep-ph].
- [13] J.R. Christiansen and P. Skands, JHEP **08** 003 (2015), arXiv:1505.01681 [hep-ph].

- [14] M. He and R. Rapp, Phys. Lett. **B795**, 117 (2019), arXiv:1902.08889 [nucl-th].
- [15] S. Plumari, V. Minissale, S. K. Das, G. Coci and V. Greco, *Charmed Hadrons from Coalescence plus Fragmentation in relativistic nucleus-nucleus collisions at RHIC and LHC*, Eur. Phys. J. **C292** 78 no. 4, (2018) 348, arXiv:1712.00730 [hep-ph].
- [16] The Large Hadron Collider. <https://home.cern/science/accelerators/large-hadron-collider>
- [17] LHC experiments. <https://home.cern/science/experiments>
- [18] The ALICE Collaboration et al. The ALICE experiment at the CERN LHC". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08002-S08002. DOI: 10.1088/1748-0221/3/08/s08002. URL: <https://doi.org/10.1088/1748-0221/3/08/s08002>.
- [19] P.A. Zyla et al. (Particle Data Group), Prog. Theor. Exp. Phys. **2020**, 083C01 (2020) <https://pdg.lbl.gov/2020/tables/rpp2020-tab-baryons-Charm.pdf>
- [20] Hoecker, A., Speckmayer, P., Stelzer, J., Therhaag, J., von Toerne, E., Voss, H., ... & Zemla, A. (2007). TMVA-toolkit for multivariate data analysis. arXiv preprint physics/0703039. <https://root.cern.ch/download/doc/tmva/TMVAUsersGuide.pdf>
- [21] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.
- [22] Ciesielski, R., & Goulianos, K. (2012). *MBR Monte Carlo simulation in PYTHIA8*. arXiv preprint arXiv:1205.1446.
- [23] Brun, R., McPherson, A. C., Zandarini, P., Maire, M., & Bruyant, F. (1987). GEANT 3: *user's guide Geant 3.10, Geant 3.11* (No. CERN-DD-EE-84-01). CERN.

Ringraziamenti

Ci tengo a ringraziare in primo luogo il professor Andrea Alici che, con grande competenza, disponibilità e pazienza, mi ha permesso di avvicinarmi al mondo del machine learning e a quello dell'analisi dati in fisica delle particelle. Fin dal nostro primo incontro mi ha guidato in questa attività di tesi con gentilezza e accoglienza.

Ringrazio i miei compagni di corso, in particolare Francesco, Leonardo, Lorenzo, Niccolò e Niccolò per la condivisione di studio e momenti di vita.

Ringrazio di cuore i miei genitori Daniela e Gianluca, i miei nonni Paola, Graziella e Nino, per aver creduto in me e avermi sostenuto in ogni fase di questo percorso accademico.

Ringrazio i miei fratelli Emanuele e Francesco per aver alleggerito i momenti di stress, soprattutto con le nostre sfide a ping pong.

Infine, ma non per importanza, ringrazio la mia ragazza, Livia, per essermi sempre accanto con amore e fiducia.