

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

Dipartimento di Fisica e Astronomia “Augusto Righi”
Corso di Laurea in Fisica

Ricostruzione del barione charmato Λ_c^+ con tecniche di Machine Learning e Reti Neurali Grafiche

Relatore:
Prof. Andrea Alici

Presentata da:
Sofia Aleotti

Anno Accademico 2024/2025

Sommario

In questo elaborato è presentata l'implementazione di un modello di Rete Neurale Grafica per la ricostruzione di particelle nell'esperimento ALICE a LHC. Lo scopo di questo esperimento è lo studio del Quark-Gluon Plasma ottenuto tramite collisioni tra ioni pesanti alle alte energie. Il Quark-Gluon Plasma è uno stato deconfinato della materia che si forma in condizioni di elevata temperatura e densità. Attraverso lo studio dei rapporti di produzione dei barioni contenenti quark pesanti in questo sistema si possono studiare i meccanismi di adronizzazione. In particolare è stato considerato il barione Λ_c^+ (*udc*) e il suo canale di decadimento $\Lambda_c^+ \rightarrow p K_S^0$. Poiché la sua vita media è molto breve, non è possibile ricostruirlo attraverso i sistemi di localizzazione di vertici di ALICE. Inoltre, il rapporto tra segnale e fondo è molto basso. Perciò vengono impiegate tecniche di Machine Learning per la ricostruzione del barione. Il programma è stato realizzato in Python, usando la API Keras e la libreria open source Spektral. I dati impiegati per l'allenamento e la valutazione che compongono il segnale sono stati generati dal generatore Monte Carlo PYTHIA8 e tramite GEANT3 sono stati simulati gli effetti dell'interazione delle particelle con i rivelatori di ALICE. I dati del fondo, invece, sono dati reali raccolti durante la RUN 2 di ALICE. La rete è in grado di distinguere i dati appartenenti alle due classi con buone accuratezza, precisione e sensibilità.

Indice

Introduzione	1
1 La fisica di ALICE	3
1.1 Il Modello Standard	3
1.2 La Cromodinamica Quantistica (QCD)	4
1.2.1 QCD su reticolo	6
1.2.2 QCD in regime perturbativo	6
1.3 Quark-Gluon Plasma (QGP)	7
1.3.1 Temperatura critica del Quark-Gluon Plasma	7
1.4 Collisioni tra ioni pesanti	8
1.4.1 Evidenze sperimentali del QGP	9
1.5 Jet quenching	11
1.6 Evoluzione del QGP	12
1.7 Adronizzazione di sapori pesanti	14
1.7.1 Adronizzazione di sapori pesanti in collisioni pp	15
1.7.2 Adronizzazione di sapori pesanti in collisioni AA	17
1.8 Risultati sperimentali	17
2 L'esperimento ALICE	22
2.1 Large Hadron Collider	22
2.2 ALICE: A Large Ion Collider Experiment	23
2.2.1 Inner Tracking System (ITS)	25
2.2.2 Time Projection Chamber (TPC)	26
2.2.3 Time-of-Flight (TOF)	27
2.2.4 High-Momentum Particle Identification Detector (HMPID)	29
2.2.5 Calorimetri elettromagnetici	31
2.2.6 Forward Muon Spectrometer (FMS)	32
2.2.7 Altri rivelatori	33

INDICE

3 Ricostruzione del barione Λ_c^+	35
3.1 Introduzione	35
3.2 Reti Neurali Grafiche	36
3.3 Keras e Spektral	38
3.3.1 Grafi in Spektral	39
3.3.2 Loaders	40
3.4 Dati e variabili fisiche	40
3.5 Definizione del modello	43
3.5.1 Ranking delle variabili	44
3.6 Valutazione	46
3.6.1 Metriche	46
3.6.2 Confusion Matrix	48
3.6.3 Curva di ROC	48
Conclusioni	50
Bibliografia	52

Introduzione

ALICE (A Large Ion Collider Experiment) è uno degli esperimenti attualmente installati a LHC, dedicato allo studio delle collisioni tra ioni pesanti ad alte energie. Lo scopo dell'esperimento è lo studio del Quark-Gluon Plasma (QGP), uno stato della materia con breve vita media che si forma in condizioni di temperatura e densità estremamente elevate. Si ritiene che fosse lo stato dell'Universo pochi microsecondi dopo il Big Bang: studiarne l'evoluzione ci potrebbe fornire importanti informazioni sui meccanismi che hanno portato alla formazione della materia che osserviamo oggi. In seguito ad una collisione tra ioni pesanti ultrarelativistici il sistema raggiunge temperature e densità sufficienti alla formazione del QGP. Oltre alle collisioni tra nuclei (A-A), nell'esperimento ALICE vengono studiate anche collisioni tra fasci di protoni (pp) e tra protoni e ioni (p -A) in modo da poter confrontare i sistemi che si ottengono nei tre diversi casi.

Lo studio del QGP avviene tramite misure indirette, come le particelle che produce. Di particolare importanza è lo studio degli adroni pesanti, ovvero quelli contenenti almeno un quark charm o bottom. Questi quark, infatti, vengono prodotti nei primi istanti successivi alla collisione, prima ancora che si formi il QGP e ne attraversano tutte le fasi dell'evoluzione. Inoltre, lo studio dei rapporti di produzione barione/mesone degli adroni pesanti permette di testare ipotesi sui meccanismi di adronizzazione. Di particolare interesse è lo studio del rapporto Λ_c^+/D^0 , rispettivamente il barione e il mesone contenenti un quark charm più leggeri. Infatti, questi adroni vengono prodotti in un intervallo molto ampio di impulsi e, poiché le loro masse sono sufficientemente elevate, il loro rapporto di produzione può essere calcolato perturbativamente, permettendo la verifica sperimentale di modelli teorici. I risultati sperimentali suggeriscono l'esistenza di due meccanismi di adronizzazione: la frammentazione, dominante ad impulsi elevati, e la coalescenza, dominante a basso impulso.

Il barione Λ_c^+ (*udc*) ha una vita media molto breve e può decadere attraverso molti canali. In questa tesi è stato considerato il canale di decadimento $\Lambda_c^+ \rightarrow p K_S^0$. A causa della vita media estremamente breve di Λ_c^+ il vertice di interazione in cui viene prodotto il barione e quello in cui decade sono così vicini da non poter essere distinti dai sistemi di localizzazione dei vertici di ALICE, come ITS e TPC. Perciò è necessario utilizzare altri metodi per ricostruire questo barione.

Nella fisica delle alte energie già da diversi anni sono impiegate tecniche di Machine

Introduzione

Learning sia per la generazione di eventi che per l'identificazione e la ricostruzione di particelle. In particolare sono state impiegate Reti Neurali di diversi tipi. In questa tesi è stato implementato in Python un modello di Rete Neurale Grafica, cioè un modello di Rete Neurale che agisce sui grafi senza modificarne la connettività, per la ricostruzione del barione Λ_c^+ . Il modello è stato definito, allenato e valutato usando la API Keras e la libreria Spektral. La qualità dell'allenamento del modello è stata valutata attraverso diverse metriche, come l'accuratezza, la precisione, la sensibilità, il fattore F1 e la curva ROC.

L'elaborato è strutturato in tre capitoli. Nel Capitolo 1 è presentata un'introduzione teorica dei fenomeni studiati nell'esperimento ALICE: a partire da una breve introduzione del Modello Standard e della Cromodinamica Quantistica (QCD) si introduce il Quark-Gluon Plasma. Nel Capitolo 2 viene descritto l'esperimento ALICE a LHC. Vengono descritti i principali rivelatori e le rispettive funzionalità. Nel capitolo 3 vengono introdotte le Reti Neurali Grafiche, il loro principio di funzionamento e viene presentata l'implementazione in Python realizzata. Infine sono presentati i risultati ottenuti dall'allenamento della Rete Neurale Grafica.

Capitolo 1

La fisica di ALICE

1.1 Il Modello Standard

La fenomenologia della fisica delle particelle è spiegata dall'insieme di teorie che chiamiamo Modello Standard (SM). Questo modello incorpora la relatività ristretta, necessaria poiché le particelle spesso si muovono con velocità vicine a quella della luce, e la meccanica quantistica, date le piccole dimensioni delle particelle. L'equazione che descrive sistemi quantistici relativistici è l'equazione di Dirac. Da questa equazione si deduce la quantizzazione dei campi, idea alla base delle teorie di campo quantizzato (QFT) che compongono il Modello Standard. Nel Modello Standard l'entità fisica fondamentale è il campo e le particelle sono quanti di tale campo: è una teoria di campi interagenti.

Le interazioni fondamentali sono quattro: gravitazionale, elettromagnetica, forte e debole. Il Modello Standard riesce a spiegare solamente le ultime tre. Il problema dell'interazione gravitazionale è tuttavia marginale nello studio della fisica delle particelle, poiché in questa scala la forza gravitazionale è trascurabile.

Secondo il Modello Standard, l'Universo è formato da fermioni elementari che interagiscono tra loro tramite campi, di cui sono sorgente. Anche ai campi sono associate delle particelle, ovvero i mediatori dell'interazione, che sono bosoni.

I fermioni fondamentali sono 12 (e le rispettive antiparticelle), divise in due tipi: leptoni e quark. I 6 leptoni sono soggetti all'interazione elettromagnetica, se sono dotati di carica elettrica, e debole. I 6 quark, invece, sono soggetti anche all'interazione forte. L'interazione elettromagnetica è mediata dai fotoni, bosoni privi di massa e di carica elettrica. Poiché i fotoni non sono massivi quest'interazione è a lungo raggio, e poiché non sono dotati di carica elettrica non è possibile l'interazione tra fotoni. L'interazione forte è mediata dai gluoni, con massa nulla e dotati di carica di colore. Infine, l'interazione debole è mediata dai bosoni W^\pm e Z , dotati di massa: è un'interazione a corto raggio. Le particelle fondamentali del Modello Standard sono schematizzate in Fig. 1.1.

Nel Modello Standard sono fondamentali le simmetrie, formalmente descritte dal-

1.2. La Cromodinamica Quantistica (QCD)

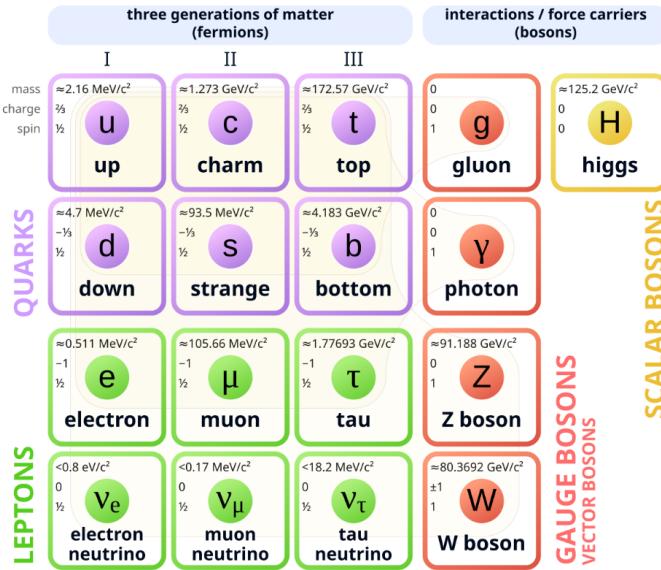


Figura 1.1: Schema delle particelle fondamentali del Modello Standard.

la teoria dei gruppi. Dal teorema di Noether sappiamo che ad ogni simmetria (ovvero invarianza dell'hamiltoniana del sistema in seguito ad una trasformazione unitaria) corrisponde una legge di conservazione. Nei processi quantistici le leggi di conservazione sono espresse attraverso la conservazione di numeri quantici. Le cariche delle tre interazioni fondamentali sono conservative, ciò significa che le hamiltoniane dei sistemi soggetti a queste interazioni sono invarianti per le trasformazioni appartenenti a gruppi unitari che le descrivono. Questi gruppi prendono il nome di gruppi di gauge e l'invarianza viene detta invarianza di gauge. In particolare, il gruppo di gauge dell'interazione elettromagnetica è $U(1)$, quello dell'interazione forte è $SU(3)$ e quello dell'interazione debole è $SU(2)$.

1.2 La Cromodinamica Quantistica (QCD)

La teoria dell'interazione forte è la cromodinamica quantistica (QCD), costruita in analogia alla teoria dell'interazione elettromagnetica, la QED. È una delle teorie che compongono il Modello Standard.

Le particelle fondamentali del Modello Standard soggette all'interazione forte sono i quark. I sistemi di quark prendono il nome di adroni. All'interno di un adrone si possono distinguere due tipi di quark: i quark *di valenza*, quelli che effettivamente determinano i numeri quantici dell'adrone, e i quark *del mare*, i quark e gluoni che sono creati e distrutti nei processi virtuali che avvengono all'interno dell'adrone secondo la QCD. In generale, si definiscono *partoni* i quark e i gluoni che formano un adrone.

1.2. La Cromodinamica Quantistica (QCD)

La sorgente dell'interazione è la carica di colore. Per descrivere la fenomenologia degli adroni sono necessarie tre cariche di colore. Infatti, i barioni sono composti da tre quark, che sono fermioni. In base al principio di Pauli in un sistema non ci possono essere due fermioni nello stesso stato: i quark che formano il barione devono quindi differire tra loro nella carica di colore. Convenzionalmente si scelgono i colori rosso (r), blu (b) e verde (g).

Il gruppo di gauge dell'interazione forte è $SU(3)$, che ha otto generatori, rappresentati dalle matrici di Gell-Mann, perciò esistono otto campi mediatori: i gluoni. I gluoni che mediano l'interazione sono dotati di una carica di colore e una di anticolore. L'interazione forte può modificare la carica di colore dei quark e dei gluoni coinvolti, ma non può modificarne il sapore. Inoltre, i gluoni, poiché sono dotati di carica di colore, possono interagire tra loro: sono possibili vertici a tre o quattro gluoni, caratteristica distintiva di questo tipo di interazione. Proprio questa caratteristica è alla base della libertà asintotica e, di conseguenza, del confinamento dei quark. Infatti, non sono mai stati osservati quark o gluoni isolati e si osserva che, nonostante i gluoni abbiano massa nulla, l'interazione forte è a corto raggio.

Consideriamo lo scattering di un quark e un antiquark. A partire da questo è possibile calcolare la costante di accoppiamento $\alpha_s(Q^2)$ dell'interazione forte, con Q^2 impulso trasferito nel sistema di riferimento del centro di massa

$$\alpha_s(Q^2) = \frac{g^2/4\pi}{1 + \frac{g^2}{16\pi^2} \left(11 - \frac{2}{3}n_f\right) \ln\left(\frac{Q^2}{\lambda^2}\right)} \quad (1.1)$$

indicando con n_f il numero di sapori cinematicamente accessibili, cioè con massa minore di Q , e con λ un parametro con le dimensioni di un'energia. Questa espressione si ottiene tenendo conto, oltre all'interazione tra le due particelle, degli effetti quantistici, come i loop di gluoni e i loop quark-antiquark. I primi hanno un effetto antischermante, mentre i secondi hanno un effetto schermante, analogamente ai loop e^+e^- . Poiché i sapori sono 6, il fattore $11 - \frac{2}{3}n_f$ è positivo. Ciò significa che

$$\alpha_s(Q^2) \rightarrow 0 \text{ per } Q^2 \rightarrow \infty \quad (1.2)$$

ovvero, all'aumentare dell'impulso trasferito o al diminuire della distanza l'intensità dell'interazione forte diminuisce, al contrario dell'effetto della polarizzazione del vuoto nel caso elettromagnetico. Questo effetto viene chiamato libertà asintotica, scoperta nel 1973 da Gross, Wilczek e Politzer. Possiamo quindi vedere la forza forte come una forza elastica, proporzionale alla distanza. Al contrario, all'aumentare della distanza o al diminuire dell'impulso trasferito, l'intensità dell'interazione aumenta: vediamo quindi come l'energia necessaria per separare i quark in un adrone sia estremamente elevata.

Possiamo definire un parametro Λ_{QCD} con le dimensioni di un'energia per riscrivere α_s :

$$\alpha_s = \frac{4\pi}{\left(11 - \frac{2}{3}n_f\right) \ln\left(\frac{Q^2}{\Lambda_{QCD}^2}\right)} \quad (1.3)$$

1.2. La Cromodinamica Quantistica (QCD)

in questa forma risulta evidente che per $Q^2 \sim \Lambda_{QCD}^2$, con $\Lambda_{QCD} \sim 300$ MeV, il fattore di accoppiamento diverge, portando la QCD nel regime non perturbativo. Il fattore Λ_{QCD}^{-1} viene detto distanza di confinamento.

A causa degli effetti antischermanti, il vuoto in prossimità di una carica di colore isolata è instabile: può avvenire il fenomeno dell'adronizzazione dei quark.

1.2.1 QCD su reticolo

La QCD su reticolo, o LQCD (Lattice QCD), nasce dalla necessità di svolgere calcoli accurati riguardanti la QCD in modo da ottenere previsioni a partire dai principi primi della teoria. L'idea alla base è quella di sostituire lo spazio-tempo 4-dimensionale infinito e continuo con un reticolo ipercubico finito. In questo modo, nota la densità della grandezza fisica di interesse, l'integrazione su uno spazio infinito diventa una somma finita sulle celle del reticolo. Si ottiene quindi un problema che si può risolvere utilizzando metodi numerici.

Nel reticolo i quark occupano i vertici delle celle, mentre i gluoni, che sono matrici 3×3 , sono rappresentati dai collegamenti tra i quark. Indichiamo con $V = L^4$ il volume dell'ipercubo e con a la separazione tra i vertici.

Una volta ottenuto il risultato sul reticolo, si deve passare al limite $V \rightarrow \infty$ e $a \rightarrow 0$. Per raggiungere il limite del continuo si procede ripetendo il calcolo per valori decrescenti di a , con una conseguente crescita del numero di passaggi e quindi dell'impegno computazionale. Più passaggi svolgiamo, migliore sarà l'accuratezza del risultato. Per confrontare i risultati della LQCD con i quelli della QCD possiamo usare le previsioni sulle masse degli adroni: per $a \sim 0.125$ fm si ottengono risultati con una discrepanza di circa l'1% dai risultati ottenuti dal calcolo sul continuo. Per quanto riguarda il limite all'infinito, per $L \geq 6.3$ fm si ottengono risultati quasi coincidenti a quelli della QCD [1].

1.2.2 QCD in regime perturbativo

In alcuni casi è possibile usare un approccio perturbativo per svolgere calcoli di QCD. Infatti, è molto difficile fare calcoli precisi in QCD, a causa delle infinite possibili interazioni con diagrammi non topologicamente equivalenti che è necessario considerare. Se la costante di accoppiamento α_s è sufficientemente piccola, però, il numero infinito di diagrammi può essere accuratamente approssimato da un numero finito. Questo approccio prende il nome di QCD perturbativa (pQCD). Si può applicare nel caso di interazioni ad alta energia o a breve distanza, cioè ad elevato Q^2 , come si vede dall'eq. 1.1. In particolare, dall'eq. 1.3 otteniamo che la scala di energia in cui è possibile applicare la pQCD è $Q^2 \gg \Lambda_{QCD}$.

1.3 Quark-Gluon Plasma (QGP)

Appena dopo il Big Bang, si stima intorno ai 10^{-6} s [2], l’Universo era molto denso e la temperatura era estremamente elevata ($T > 100$ GeV). In queste condizioni, a causa della libertà asintotica quark e gluoni non erano confinati: si trovavano in uno stato chiamato Quark-Gluon Plasma (QGP), simile ad un liquido con viscosità molto piccola. In seguito, con il raffreddamento e l’espansione dell’Universo, quark e gluoni si sono raggruppati in adroni, che formano la materia che osserviamo oggi. Lo studio del QGP è quindi di fondamentale importanza per la comprensione dell’origine della materia ordinaria, fatta di nuclei composti da protoni e neutroni.

1.3.1 Temperatura critica del Quark-Gluon Plasma

Per temperature $T > 100$ GeV possiamo considerare tutte le particelle come ultrarelativistiche, trascurandone quindi la massa. In queste condizioni, la materia può essere trattata come un gas di particelle del Modello Standard usando metodi statistici.

Consideriamo quindi un gas di bosoni e fermioni. Le particelle saranno descritte dalla statistica di Bose-Einstein e di Fermi-Dirac, rispettivamente. La densità di energia della specie i -esima:

$$\epsilon_i = \int \frac{d^3 p_i}{(2\pi)^3} \frac{E_i}{e^{\beta E_i} \pm 1} = \begin{cases} \frac{\pi^2}{30 T^4} & \text{bosoni} \\ \frac{7\pi^2}{8 T^4 30} & \text{fermioni} \end{cases} \quad (1.4)$$

Per determinare l’energia totale del sistema è necessario conoscere anche la molteplicità g_i delle singole specie

$$\epsilon = \sum g_i \epsilon_i \quad (1.5)$$

Dall’eq.1.4 otteniamo:

$$\epsilon = g^* \frac{\pi^2 T^4}{30} \quad (1.6)$$

con $g^* = g_b + \frac{7}{8} g_f$, dove g_b è la molteplicità dei bosoni e g_f dei fermioni.

Con questa trattazione possiamo determinare la temperatura critica T_C , ovvero la temperatura alla quale si ha la transizione da QGP a gas di adroni, una transizione di fase del primo ordine. Otteniamo questa temperatura confrontando la pressione del QGP come gas di particelle elementari con la pressione di un gas di adroni. In particolare, consideriamo gli adroni più leggeri, i pioni, poiché saranno i primi a formarsi. Per farlo, adottiamo come modello per gli adroni il MIT bag model.

Secondo questo modello, gli adroni possono essere pensati come “sacche” (*bags*) di quark, antiquark e gluoni. All’interno della sacca i componenti dell’adrone sono liberi di muoversi, ma sono confinati al suo interno. In questo modello è necessario introdurre una costante $B \sim 200$ MeV, la *bag constant*, che rappresenta l’energia relativa alla formazione

1.4. Collisioni tra ioni pesanti

dell'adrone: l'energia necessaria per creare un volume all'interno del quale i quark siano liberi di muoversi.

La pressione di un gas è data da:

$$p = \frac{\epsilon}{3} \quad (1.7)$$

Un pion è dotato di tre gradi di libertà. Tenendo conto della bag constant, la pressione di un gas di pioni

$$P_\pi = B + \frac{3\pi^2 T^4}{90} \quad (1.8)$$

La pressione del QGP

$$P_{QGP} = \frac{37\pi^2 T^4}{90} \quad (1.9)$$

equagliando la 1.8 alla 1.9 otteniamo

$$T_C = \left(\frac{45B}{17\pi^2} \right)^{\frac{1}{4}} \sim 180 \text{ MeV} \quad (1.10)$$

Tuttavia, per ottenere risultati più precisi è necessario il calcolo numerico, utilizzando la LQCD.

1.4 Collisioni tra ioni pesanti

In seguito alla collisione tra ioni pesanti ad alte energie si ottiene un sistema con densità di nucleoni molto più elevata rispetto a quella del nucleo. Se nella collisione viene raggiunta una temperatura $T > T_C$ avviene il deconfinamento dei quark: si riesce quindi a ottenere il Quark-Gluon Plasma. Considerando le collisioni centrali, o *head-on*, cioè quelle con la massima sovrapposizione tra gli ioni durante la collisone, si osserva una densità di energia nella regione centrale di almeno 12 GeV/fm^3 all'energia di $\sqrt{s} = 2, 76 \text{ TeV}$. Questa densità di energia è molto maggiore rispetto a quella all'interno di un adrone ($\sim 500 \text{ MeV/fm}^3$) [3]. In questo tipo di collisioni abbiamo il maggior numero di nucleoni coinvolti, e quindi la massima probabilità di ottenere il QGP.

Per definire una collisione centrale si utilizza il parametro d'impatto b , cioè la distanza tra i centri dei nuclei in collisione. Se è piccolo, la collisione è centrale, altrimenti è periferica. Sperimentalmente, però, non è possibile misurarla, perciò si utilizzano altri parametri, come la molteplicità, la rapidità e la pseudorapidità. La molteplicità è il numero di particelle prodotte dalla collisione. La rapidità di una particella prodotta è definita come

$$y = \frac{1}{2} \ln \frac{E + p_L}{E - p_L} \quad (1.11)$$

1.4. Collisioni tra ioni pesanti

dove E è l'energia della particella e p_L l'impulso nella direzione del fascio. La pseudorapidezza si definisce, per ogni particella prodotta, come

$$\eta = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right] \quad (1.12)$$

indicando con θ l'angolo tra la direzione della particella prodotta e il fascio.

Nelle collisioni tra ioni pesanti, rappresentate schematicamente in Fig. 1.2, modellizziamo i nuclei incidenti come dei dischi, a causa della contrazione di Lorentz, che contengono nucleoni, a loro volta composti da partoni. Definiamo spettatori i nucleoni che non partecipano alla collisione, ovvero quelli che non incrociano nessun nucleone appartenente ad altri nuclei. I nucleoni partecipanti, invece, sono i nucleoni che prendono parte alla collisione. Indichiamo il numero di nucleoni dei due tipi con N_{spec} e N_{part} , rispettivamente. Indichiamo con N_{coll} il numero totale di collisioni tra nucleoni.

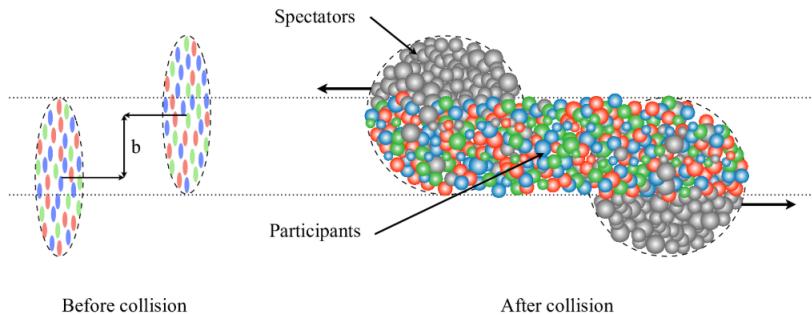


Figura 1.2: Rappresentazione schematica di una collisione tra ioni pesanti. È indicato il parametro d'impatto b , i nucleoni partecipanti e spettatori.[4]

In seguito alla collisione la maggior parte dell'energia cinetica dei nuclei si concentra nella regione centrale: si forma una fireball ad elevata densità di energia, che dopo circa 1 fm/ c si trasforma in QGP. Questo liquido, che si può studiare attraverso l'idrodinamica quantistica, si espande e si raffredda, portando all'adronizzazione. Questo processo ha inizio quando la temperatura scende al di sotto di quella critica e la densità del sistema diminuisce. L'adronizzazione si verifica grazie ai fenomeni di scattering inelastico che avvengono all'interno del QGP, mentre l'espansione è dovuta allo scattering elastico tra adroni. Questi fenomeni si verificano fino al raggiungimento delle temperature di freeze-out chimico e poi termico, rispettivamente. Accade circa a 10-15 fm/ c dalla collisione. A questo punto, gli adroni prodotti si disaccoppiano dal sistema e possono essere rivelati.

1.4.1 Evidenze sperimentali del QGP

Il confronto tra le collisioni A-A e le collisioni pp ci fornisce informazioni sulla natura del sistema prodotto in seguito alla collisione.

1.4. Collisioni tra ioni pesanti

Le particelle cariche che vengono prodotte dalle collisioni, che per motivi pratici sono le uniche su cui abbiamo informazioni, sono proporzionali a N_{part} . Inoltre, si osserva che il numero di particelle cariche prodotte per partecipante è circa uguale a quello di una collisione pp [3]. Ciò sorprende poiché l'energia persa dai nuclei in collisione è maggiore rispetto a quella persa dai protoni in collisione: nel primo caso circa l'85% del totale e circa il 50% nel secondo.

Inoltre, il numero di particelle prodotte aumenta con l'energia. La distribuzione della rapidità è molto simile a una gaussiana con larghezza $\propto \ln(s_{NN})$, indicando con $\sqrt{s_{NN}}$ l'energia della collisione nel sistema di riferimento del centro di massa, e aumenta di poco tra collisioni centrali e periferiche. Di conseguenza, la densità di particelle $\frac{dN}{dy}$ ha la forma di Fig. 1.3. Questo ci suggerisce che la maggior parte della produzione di entropia, proporzionale al numero di particelle prodotte, avvenga nei primi istanti successivi alla collisione. È consistente con la formazione di un liquido a piccola viscosità, che produce poca entropia nel suo sviluppo idrodinamico.[3]

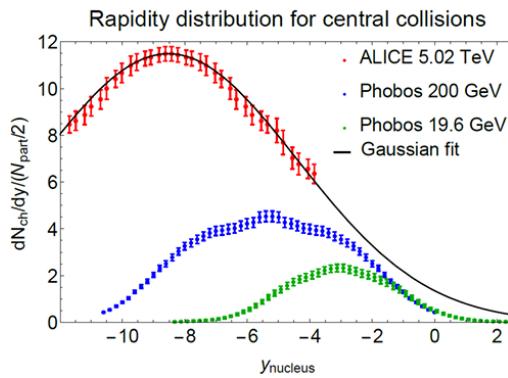


Figura 1.3: *Distribuzione della rapidità y degli adroni carichi prodotti nella collisione. Dati riguardanti collisioni AuAu a 19.6 e 200 GeV e PbPb a 5.02 TeV.[3]*

Un'ulteriore conferma del fatto che la maggior parte delle particelle viene prodotta subito dopo la collisione è il fatto che si osservino forti correlazioni tra particelle con rapidità molto diverse, in base al principio di causalità.

Infine, possiamo fare alcune considerazioni sulla densità di energia del sistema. I nuclei coinvolti nella collisione, come tutti i nuclei, hanno una densità di energia “grumosa”. Perciò, anche la densità di energia della materia prodotta appena dopo la collisione dovrà avere la stessa struttura. Questa materia può quindi essere simile ad un gas, formato da particelle poco interagenti tra loro, oppure un liquido, in cui le interazioni sono più intense. Se il QGP fosse un gas, alla scomparsa della grumosità iniziale, le particelle che lo compongono si allontanerebbero in modo isotropo. In un liquido, invece, la grumosità iniziale porterebbe alla formazione di gradienti di pressione: si formerebbero nel liquido

1.5. Jet quenching

dei flussi anisotropi. Questo secondo caso è proprio quello che si osserva in seguito alle collisioni tra ioni pesanti.

1.5 Jet quenching

Quando avvengono collisioni tra nucleoni si verificano fenomeni di scattering ad alto impulso trasferito tra partoni in cui si formano quark o gluoni con elevato impulso trasverso p_T . I partoni così prodotti emettono gluoni e coppie quark-antiquark, in base a quanto previsto dalla QCD, che si ricombinano in adroni. Si forma quindi un getto (*jet*) di adroni, che possono essere rivelati. Si definisce *jet quenching* l'attenuazione o la scomparsa di un jet di adroni a causa della perdita di energia dei partoni nel mezzo, come schematizzato in Fig. 1.4.

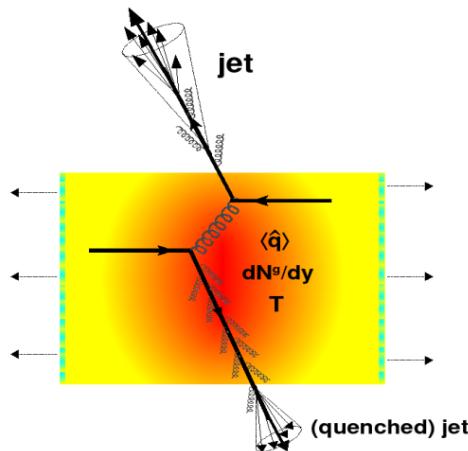


Figura 1.4: Schema del jet quenching in una collisione nucleo-nucleo. Sono rappresentati due quark che subiscono un fenomeno di scattering. Uno dei due (in alto) si allontana velocemente dal plasma, emettendo qualche gluone e portando alla formazione di un jet di adroni. L'altro (in basso) attraversa il plasma, perdendo energia a causa dell'emissione di gluoni indotta dal mezzo: dal plasma esce un jet smorzato di adroni.

La produzione di jet avviene sia nelle collisioni pp che in quelle AA , ma il loro comportamento nei due casi è diverso. Per quantificare la differenza tra i due tipi di collisioni, si definisce il nuclear modification factor

$$R_{AA}(p_T) = \frac{dN_{AA}/p_T}{N_{coll} dN_{pp}/dp_T} \quad (1.13)$$

dove dN_{AA}/dp_T è il numero di particelle di un tipo prodotte in una collisione AA e analogamente dN_{pp}/dp_T . Se i nucleoni all'interno dei nuclei avessero lo stesso comportamento

1.6. Evoluzione del QGP

di quelli liberi avremmo $R_{AA} = 1$, cioè se le collisioni tra nuclei fossero solo una sovrapposizione di collisioni pp . Si misura però $R_{AA} < 1$, cioè una soppressione, nel caso dei jet. In particolare, si osserva una perdita di energia e di impulso nella direzione originale del jet proporzionale alla distanza percorsa all'interno del QGP. I jet prodotti nelle collisioni centrali sono soggetti allo smorzamento maggiore, poiché devono percorrere una distanza maggiore nel mezzo.

È interessante studiare la variazione di energia dei jet nel QGP perché ci fornisce informazioni sulla struttura interna del mezzo: l'obiettivo è quello di usarli come sonde. Infatti, non è possibile usare sonde esterne per studiare il QGP a causa della sua vita estremamente breve, circa 10^{-23} s [5].

La variazione totale di energia di una particella in un mezzo è data da due termini: lo scattering elastico con i componenti del mezzo, detto termine di collisione e prevalente quando la particella ha basso impulso, e lo scattering inelastico, detto termine radiativo e prevalente nelle particelle ad elevato impulso. Indichiamo la variazione di energia nei due casi con ΔE_{coll} e ΔE_{rad} , rispettivamente. Nel caso dell'interazione forte, a causa dell'elevato valore di α_s , il contributo principale è quello radiativo, dovuto all'emissione di gluoni da parte dei partoni. Pertanto, l'informazione principale che otteniamo dallo studio della variazione di energia dei jet è la prova che il QGP interagisce fortemente. Un'ulteriore conferma è data dal fatto che per i fotoni e per il bosone Z , che non sono soggetti all'interazione forte, si misura $R_{AA} = 1$.

Il limite dello studio del QGP attraverso il jet quenching è il fatto che possiamo misurare solamente i getti di adroni finali: non è possibile confrontare i jet prima e dopo il passaggio nel mezzo.

1.6 Evoluzione del QGP

Dalla sezione 1.4.1 abbiamo visto come il QGP prodotto dalla collisione tra nuclei pesanti si comporta come un liquido a bassa viscosità, perciò può essere trattato usando l'idrodinamica quantistica. Possiamo dividere l'evoluzione del QGP in diverse fasi, schematizzate in Fig. 1.5.

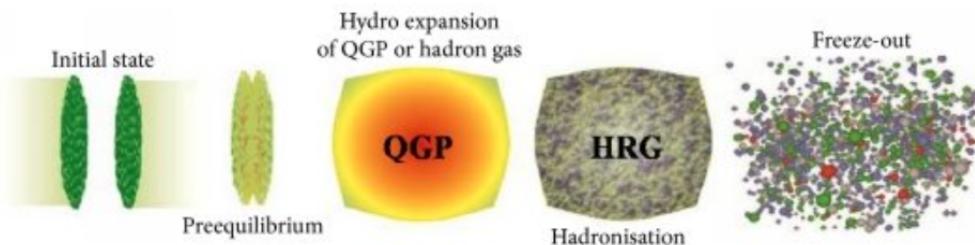


Figura 1.5: Fasi dell'evoluzione del QGP.

Fase di pre-equilibrio

La fase di pre-equilibrio si osserva per $t < 1 \text{ fm}/c$. In questa fase avvengono i principali fenomeni di scattering, che portano alla formazione di un gran numero di quark e gluoni deconfinati. Poiché sono processi ad alto impulso trasferito, possono essere descritti dalla pQCD.

Termalizzazione

A causa delle interazioni elastiche e inelastiche tra i partoni, che possono causare anche una variazione del sapore delle particelle, si verifica l'espansione del QGP. In questa fase il QGP si comporta come un fluido di partoni in espansione idrodinamica. Ciò si verifica per $1 \text{ fm}/c < t < 10 \text{ fm}/c$.

Freeze-out chimico

Con l'espansione, il QGP si raffredda. Al di sotto della temperatura critica quark e gluoni condensano in adroni. L'adronizzazione avviene a $t \sim 20 \text{ fm}/c$.

Un modello che descrive con un buon accordo con i dati sperimentali la produzione di adroni nel QGP è il modello termico. Con questo modello si può descrivere la fase del freeze-out chimico utilizzando la meccanica statistica. Il sistema viene descritto attraverso un ensemble grancanonico. I parametri del modello sono molto pochi: la temperatura T , il potenziale chimico barionico μ_B e il volume V .

Il modello mette in relazione l'adronizzazione e il diagramma di fase del QGP. Si osserva, infatti, che per piccoli valori di μ_B la temperatura di freeze-out chimico coincide con la temperatura critica T_C .

A partire dal modello termico si ottengono buone previsioni sulla produzione delle diverse specie di adroni, come si vede in Fig. 1.6 con il caso dei pioni e dei kaoni. In particolare si osserva una produzione prevalentemente di pioni rispetto ad adroni che contengono anche il quark strange. Ciò è dovuto al fatto che la fireball si forma in seguito alla collisione tra nucleoni, composti dai quark up e down: la maggior parte degli adroni viene quindi prodotta da quark già presenti nel sistema, intrappolati nel QGP. Ad energie più elevate, come ci si aspetta, questa differenza è meno marcata.

Le previsioni rimangono buone anche per quanto riguarda la produzione di adroni che contengono quark pesanti (charm e bottom), aggiungendo un parametro: la sezione d'urto di produzione dei quark pesanti σ_{HQ} . Proprio questo modello ipotizza come meccanismo principale di produzione del charmonio J/ψ ($c\bar{c}$) la generazione statistica, in accordo con le osservazioni sperimentali. Infatti si osserva una produzione in media di circa 30 mesoni J/ψ nelle collisioni tra ioni pesanti, circa lo stesso delle collisioni pp [3]. Dai dati provenienti dalle collisioni pp si era ipotizzato che questi mesoni avessero origine dai quark charm e anticharm prodotti in fenomeni di hard scattering nelle fasi iniziali della collisione. Nel QGP, però, questo non sarebbe possibile a causa dello screening

1.7. Adronizzazione di savori pesanti

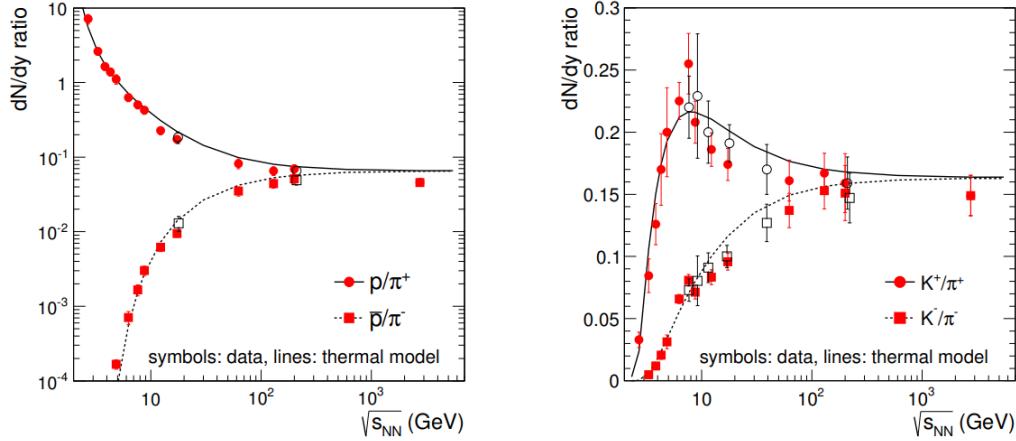


Figura 1.6: Rapporto di produzione di pioni e antipioni (a destra) e di kaoni e antikaoni (a sinistra) in funzione dell’energia $\sqrt{s_{NN}}$. I simboli rappresentano i dati sperimentali, le linee le previsioni del modello termico [6].

da parte del mezzo: separerebbe i mesoni formatisi inizialmente. Secondo il modello statistico, invece, nelle collisioni tra ioni pesanti i mesoni J/ψ si formano in seguito ad un processo di ricombinazione: i c e \bar{c} prodotti si muovono indipendentemente tra loro nel QGP, fino a trovarsi durante l’adronizzazione. La validità del modello sarebbe quindi un’ulteriore conferma del deconfinamento all’interno del QGP.

Freeze-out termico

Dopo il freeze-out chimico nel sistema si verificano solamente urti elastici tra adroni, responsabili dell’espansione del sistema. Quando la distanza media tra gli adroni diventa maggiore del raggio d’azione dell’interazione forte, ad una temperatura di circa 120 MeV, anche gli urti elastici si interrompono. Questa temperatura viene detta di freeze-out termico. L’impulso delle particelle in espansione, quindi, non viene più modificato: avviene il congelamento dello spettro in impulso delle particelle prodotte dal sistema iniziale.

1.7 Adronizzazione di savori pesanti

Lo studio dei quark charm e bottom ci permette di ottenere informazioni su tutta l’evoluzione del QGP, poiché la loro vita media è più lunga di quella del QGP e il loro tempo di formazione è più breve del tempo di termalizzazione. Infatti, questi quark si formano nelle prime fasi della collisione A-A da collisioni ad alto impulso trasferito tra partoni. Otteniamo informazioni sul QGP confrontando la produzione di adroni che

1.7. Adronizzazione di svari pesanti

contengono svari pesanti nelle collisioni A-A, in cui si forma, con le collisioni pp e pA , in cui invece non è presente.

1.7.1 Adronizzazione di svari pesanti in collisioni pp

L'adronizzazione di svari pesanti in collisioni pp avviene attraverso il processo di frammentazione: un singolo partone, attraverso processi di hard scattering, dà origine a diversi adroni. Possiamo descrivere matematicamente questo meccanismo con il teorema di fattorizzazione [7]. Il processo complessivo è

$$p + p \rightarrow h + X \quad (1.14)$$

dove h è l'adrone di interesse. h è il prodotto del decadimento del partone c prodotto dallo scattering

$$a + b \rightarrow c + d \quad (1.15)$$

con a, b partoni del protone. Attraverso il teorema di fattorizzazione possiamo separare il contributo perturbativo da quello non perturbativo. Il primo è dovuto alla produzione del *leading parton*, processo ad alta energia, mentre il secondo è dato dalla sua successiva conversione allo stato adronico, a bassa energia. Quindi è possibile calcolare perturbativamente la sezione d'urto del processo 1.15. Possiamo ora esprimere la sezione d'urto invariante della produzione dell'adrone nel medio range di rapidità

$$\frac{d\sigma_{pp}^h}{dy d^2 p_T} = K \sum_{abcd} \int dx_a dx_b f_a(x_a, Q^2) f_b(x_b, Q^2) \frac{d\sigma}{dt}(ab \rightarrow cd) \frac{D_{h/c}^0}{\pi z_c} \quad (1.16)$$

dove con $f_i(x_i, Q^2)$ si indicano le Funzioni di Distribuzione dei Partoni (PDF) dei partoni del protone, con $\frac{d\sigma}{dt}$ la sezione d'urto elementare del processo $a + b \rightarrow c + d$ e con $D_{h/c}^0$ la Funzione di Frammentazione (FF).

Funzioni di Distribuzione dei Partoni (PDF)

Le Funzioni di Distribuzione dei Partoni rappresentano la densità di probabilità di trovare un partone in un adrone, data la scala di energia Q^2 , in funzione della frazione dell'impulso dell'adrone trasportata dal partone x . Sono fondamentali per lo studio dei processi ad alta energia che coinvolgono uno o più adroni nello stato iniziale.

Le PDF devono essere coerenti con il numero barionico dell'adrone e rispettare la momentum sum rule:

$$\sum_i \int dx x q_i(x) = 1 \quad (1.17)$$

1.7. Adronizzazione di sapori pesanti

dove abbiamo indicato con $q_i(x)$ la PDF del quark i -esimo. Possiamo considerare come esempio il protone. I suoi quark di valenza sono uud , quindi:

$$\int_0^1 dx u^v(x) = 2 \quad \int_0^1 dx d^v(x) = 1 \quad (1.18)$$

Mentre per i quark del mare:

$$\int_0^1 dx [q^s(x) - \bar{q}^s(x)] = 0 \quad (1.19)$$

poiché sono sempre prodotti in coppie quark-antiquark e non contribuiscono al numero barionico dell'adrone. La condizione 1.19 deve essere verificata per tutti i sapori di quark. La momentum sum rule in questo caso diventa

$$\int_0^1 dx x [u^v(x) + d^v(x) + \sum_q (q^s(x) + \bar{q}^s(x))] = 1 \quad (1.20)$$

Funzioni di Frammentazione (FF)

Le Funzioni di Frammentazione $D_{h/c}^0$ rappresentano la probabilità che un partone c adronizzi emettendo gluoni dell'adrone h , trasportando una frazione del momento del partone iniziale. Storicamente sono state introdotte per spiegare le ampie asimmetrie azimutali osservate nelle collisioni pp .

Consideriamo il processo di annichilazione elettrone-positrone

$$e^- e^+ \rightarrow q\bar{q} \quad (1.21)$$

Supponiamo che l'energia della collisione sia Q . L'energia del fascio $E_f = \frac{Q}{2}$, quindi ciascun quark ha energia $E_q = E_f$. Nello stato finale otterremo un adrone h con energia E_h . La frazione di energia iniziale portata dall'adrone è

$$z = \frac{E_h}{E_q} = \frac{2E_h}{Q} \quad (1.22)$$

Definiamo la Funzione di Frammentazione $D_q^h(z)$ come la probabilità che l'adrone finale h porti una frazione z dell'energia dello stato iniziale. La sezione d'urto differenziale per la produzione di adroni in funzione di z , applicando il teorema di fattorizzazione:

$$\frac{d\sigma(e^- e^+ \rightarrow hX)}{dz} = \sum_q \sigma(e^- e^+ \rightarrow q\bar{q}) [D_q^h(z) + D_{\bar{q}}^h(z)] \quad (1.23)$$

Osserviamo che non sono necessarie PDF poiché gli elettroni sono particelle elementari. Per come è definita la FF deve valere la condizione

$$\sum_h \int_0^1 dz z D_q^h(z) = \sum_h \int_0^1 dz z D_{\bar{q}}^h(z) = 1 \quad (1.24)$$

1.8. Risultati sperimentali

La molteplicità di un adrone h è data dalla somma delle probabilità di produrre h da tutti i possibili quark e antiquark:

$$n_h = \sum_q \int_{2m_h/Q}^1 dz [D_q^h(z) + D_{\bar{q}}^h(z)] \quad (1.25)$$

integrando quindi su tutte le energie che permettono la formazione di tale adrone, indicando con m_h la sua massa.

Le FF non possono essere calcolate perturbativamente, si ricavano da fit sui dati sperimentali ottenuti dalle collisioni e^-e^+ . Si pensa che i parametri di queste funzioni siano universali, perciò applicabili anche agli altri tipi di collisioni.

1.7.2 Adronizzazione di svari pesanti in collisioni AA

Nelle collisioni AA l'adronizzazione avviene attraverso meccanismi diversi rispetto al caso pp e e^-e^+ . Ciò è suggerito dalla soppressione del rapporto protoni/pioni nel secondo caso. Infatti, quando la frammentazione è il meccanismo principale si ha

$$\frac{D_q^p(z)}{D_q^\pi(z)} < 0.25 \quad (1.26)$$

mentre nelle collisioni AA si osserva un rapporto protoni/pioni anche superiore all'unità [8]. Questa soppressione della produzione di pioni non può essere dovuta alla variazione di z a causa dell'interazione forte all'interno del QGP, che porta ad una perdita di energia: se fosse così tutte le particelle dovrebbero esserne soggette in modo pressoché uguale e il rapporto rimarrebbe invariato. Lo stesso si osserva anche per il rapporto Λ/K .

Il meccanismo che agisce in questi casi prende il nome di *ricombinazione* o *coalescenza*: quando due partoni sono vicini nello spazio delle fasi si possono ricombinare in barioni o mesoni, con processi $3 \rightarrow 1$ e $2 \rightarrow 1$ rispettivamente. Con questo meccanismo gli adroni prodotti hanno impulso $p_h = np_T$, con $n = 3$ per i barioni e $n = 2$ per i mesoni, mentre con la frammentazione del singolo partone $p_h = zp_T$, come rappresentato in Fig. 1.7. Infatti, nel caso della frammentazione, l'impulso iniziale è diviso tra gli adroni ottenuti. A basso p_T domina la ricombinazione, mentre ad alto p_T la frammentazione. Ad alto impulso, infatti, è meno probabile che due partoni si trovino vicini nello spazio delle fasi.

1.8 Risultati sperimentali

Lo studio sperimentale dei meccanismi di adronizzazione si basa sulle misure di adroni che contengono i quark pesanti, charm e bottom, come visto in Sez. 1.7. Le loro masse, inoltre, sono molto maggiori della scala di energia Λ_{QCD} : la loro produzione può essere

1.8. Risultati sperimentali

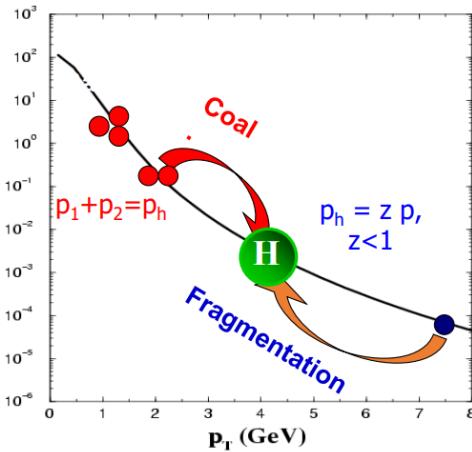


Figura 1.7: Adronizzazione tramite ricombinazione (o coalescenza), in rosso, e frammentazione, in blu. Attraverso la ricombinazione si ottiene un adrone con impulso dato dalla somma di quello dei partoni che lo formano, con la frammentazione lo stato finale ha impulso minore del partone iniziale.

calcolata perturbativamente. Sono quindi adatte ad una verifica sperimentale di modelli teorici. In particolare, sono interessanti le misure riguardanti gli adroni Λ_c^+ (*udc*) e D^0 ($c\bar{u}$) e il loro rapporto di produzione barione/mesone Λ_c^+/D^0 : sono, rispettivamente, il barione e il mesone più leggeri contenenti il quark charm e possono essere osservati in un ampio intervallo di impulsi. Il rapporto di produzione tra questi adroni è sensibile solamente alle variazioni delle Funzioni di Frammentazione del quark charm, poiché le PDF sono comuni ad entrambi. In collisioni pp con $\sqrt{s} = 5.02$ TeV, si è osservato che Λ_c^+/D^0 è dipendente da p_T e il suo andamento non è ben descritto da modelli che utilizzano le FF parametrizzate a partire da collisioni e^+e^- [9]. Alcuni esempi sono:

- PYTHIA 8.243 Monash 2013 [10]: un generatore Monte Carlo usato per la generazione di eventi di collisione nella fisica delle alte energie. Si basa sul modello a stringhe di Lund: quark e gluoni sono connessi tra loro da stringhe di colore che naturalmente frammentano in adroni. I partoni possono interagire tra loro attraverso le stringhe di colore. Questo modello assume un rapporto Λ_c^+/D^0 costante con p_T e sottostima le misure sperimentali;
- HERWIG 7 [11]: generatore di eventi basato sul modello a cluster dell'adronizzazione, secondo cui nella regione non perturbativa, coppie di colore-anticolore si trovano vicine nello spazio delle fasi, andando a formare stati adronici molto eccitati: i cluster. Dal decadimento dei cluster si formano gli adroni dello stato finale.

Da altri modelli si ottengono previsioni più accurate:

1.8. Risultati sperimentali

- PYTHIA 8 CR-BLC [12]: una versione di PYTHIA che implementa la Color Reconnection Beyond the Leading Colour (CR-BLC). In questo modo si considera un contributo maggiore dovuto alla frammentazione delle “giunzioni” in barioni: è prevista una maggiore produzione di barioni;
- Catania [13]: assume la formazione del QGP anche in collisioni pp ad alta energia, e perciò che la produzione di adroni avvenga attraverso il meccanismo della frammentazione e della coalescenza;
- Statistical Hadronization Model (SHM) [14]: calcola le frazioni di quark charmatis basandosi sulle densità termiche, che dipendono dalla massa dello stato e dal fattore di degenerazione di spin. Per spiegare l'elevato valore del rapporto Λ_c^+/D^0 misurato in collisioni pp in ALICE fa uso di ulteriori stati barionici eccitati ancora mai misurati, ma che si assume esistano secondo il modello relativistico dei quark (RQM). Descrive accuratamente i dati sperimentali a tutti i valori di p_T .

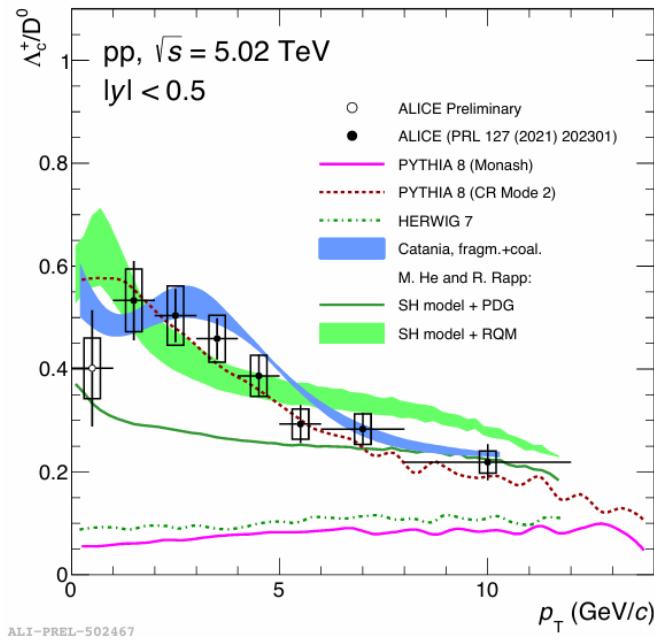


Figura 1.8: Rapporto Λ_c^+/D^0 in funzione di p_T nelle collisioni pp a $\sqrt{s} = 5.02$ TeV misurato a confronto con i modelli PYTHIA 8 (Monash), PYTHIA 8 (CR-BLC), HERWIG 7, Catania e SHM [9].

Le misure a confronto con i modelli sono in Fig. 1.8. Ciò che si può concludere da queste misure è che non esiste un unico modello di adronizzazione comune a tutti i sistemi e che esiste un limite all'universalità delle FF con parametrizzazioni ottenute da collisioni

1.8. Risultati sperimentali

e^+e^- . Anche per i rapporti $\Xi_c^{0,+}/D^0$, Ω_c^0/D^0 e $\Sigma_c^{0,++}/D^0$ si osserva una dipendenza da p_T , ma non per i rapporti D^+/D^0 e D_s^+/D^0 , che sono coerenti con le previsioni della pQCD [9].

Dai dati ottenuti in collisioni PbPb a $\sqrt{s} = 2.76$ TeV possiamo ottenere informazioni anche sugli spettri dei singoli adroni, riportati in Fig. 1.9. Per quanto riguarda D^0 i dati sperimentali sono in buon accordo con un modello di adronizzazione tramite frammentazione e coalescenza, con il contributo principale proveniente dalla frammentazione. Per il bario Λ_c^+ alle energie di LHC la coalescenza è dominante per $p_T < 5$ GeV/c. Infine, si osserva che in questo sistema, la dipendenza dall'impulso trasverso degli adroni pesanti è minore rispetto a quella degli adroni leggeri. Il rapporto Λ_c^+/D^0 presenta comunque un picco a circa 1.5 per $p_T \sim 3$ GeV/c, come si vede in Fig. 1.10. In generale, si osserva un'incremento del rapporto solamente a basso impulso trasverso, dove domina la coalescenza. Per questo Λ_c^+/D^0 può essere usato come indicatore dei meccanismi di adronizzazione presenti nel sistema. Dalla Fig. 1.11 vediamo come tale rapporto non abbia variazioni significative in funzione della molteplicità, del sistema di collisione o dell'energia.

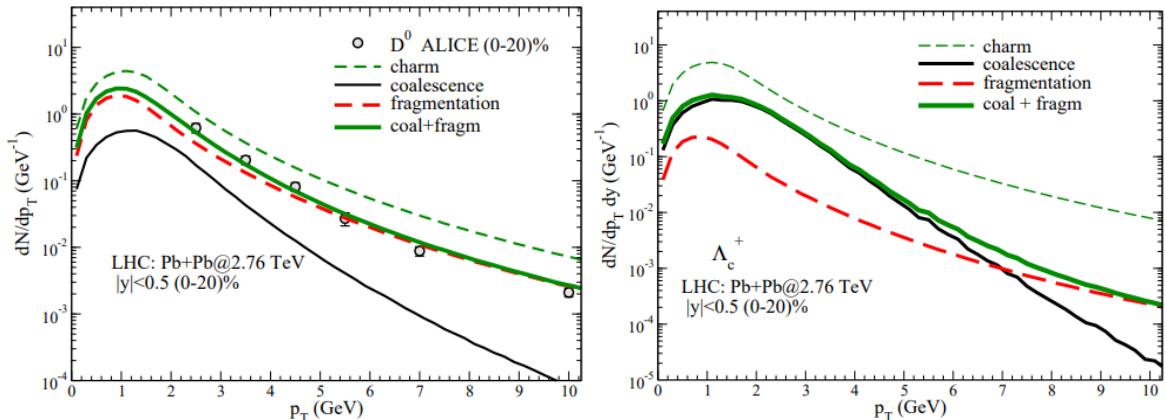


Figura 1.9: Spettro di barioni D^0 (a sinistra) e di Λ_c^+ in funzione di p_T in collisioni PbPb a $\sqrt{s} = 2.76$ TeV. La linea tratteggiata verde e quella continua nera si riferiscono allo spettro del charm [13].

1.8. Risultati sperimentali

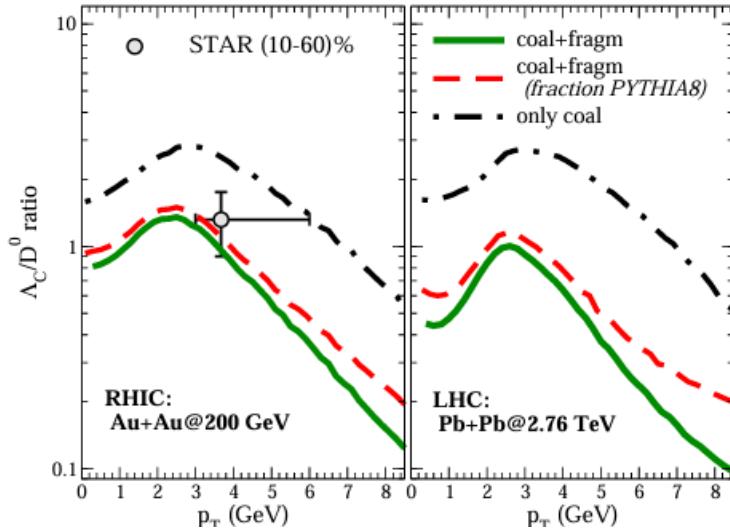


Figura 1.10: Rapporto Λ_c^+/D^0 in funzione di p_T a media rapidità in collisioni AuAu a $\sqrt{s} = 200$ GeV e PbPb a $\sqrt{s} = 2.76$ TeV. Le linee nere tratteggiate si sono ottenute assumendo solo la coalescenza, le linee verdi continue assumendo coalescenza e frammentazione. Le linee rosse tratteggiate sono ottenute con PYTHIA 8 [13].

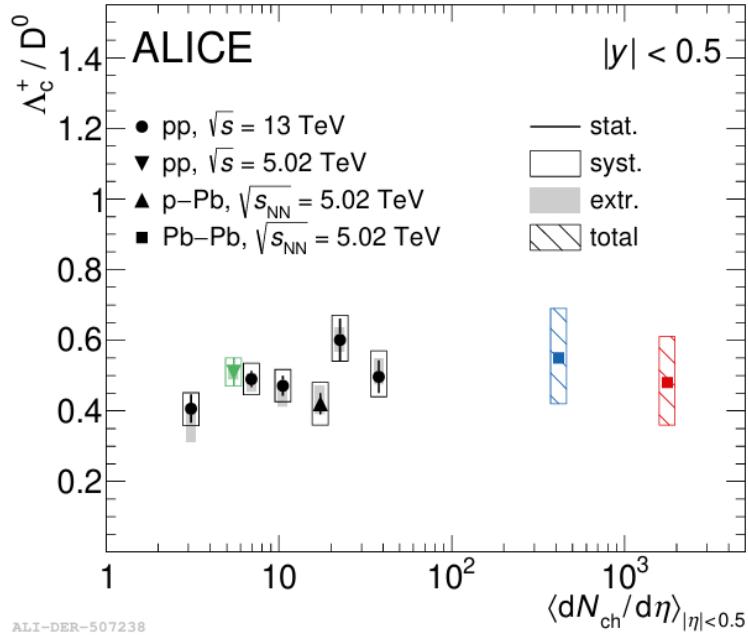


Figura 1.11: Rapporto di produzione Λ_c^+/D^0 integrato su p_T in funzione della molteplicità di particelle cariche in collisioni pp, p-Pb e PbPb a diverse energie di collisione, indicate in figura [9].

Capitolo 2

L'esperimento ALICE

2.1 Large Hadron Collider

Il Large Hadron Collider (LHC) è il più grande e potente acceleratore di protoni e ioni al mondo. Fa parte del complesso di acceleratori del CERN, a Ginevra, schematizzato in Fig. 2.1. Si trova a 100 m di profondità, con una circonferenza di circa 27 km. A LHC avvengono collisioni tra fasci di protoni o ioni che viaggiano in direzione opposta a velocità molto prossime a quella della luce. I due fasci viaggiano all'interno di due tubi separati, i *beam pipes*, guidati da un campo magnetico generato da magneti superconduttori. Per ottenere la supercondutività è necessario mantenere il sistema ad una temperatura di -271.3°C . Per mantenere il sistema a queste temperature si usa un sistema di distribuzione di elio liquido [15].

I magneti impiegati sono di diverso tipo: 1232 dipolari si usano per piegare il fascio, 392 quadripolari si usano per collimarli e altri magneti si usano per restringerli, aumentando la probabilità di collisione. Inoltre, sono disposti in modo tale che le collisioni avvengano in corrispondenza dei quattro rivelatori:

- ATLAS (A Toroidal LHC Apparatus): uno dei rivelatori *general purpose* ad LHC. Studia un'ampia gamma di campi, dal Modello Standard e il bosone di Higgs alla materia oscura. È formato da 6 sottosistemi di rivelazione disposti a strati intorno al punto della collisione. È il più grande rivelatore di particelle al mondo in termini di volume;
- CMS (Compact Muon Solenoid): come ATLAS è un rivelatore *general purpose*, ma i due rivelatori usano metodi e tecnologie di rivelazione diverse. CMS è costruito intorno ad un solenoide realizzato con una spira cilindrica di materiale superconduttore, che produce un campo magnetico di 4 T;
- LHCb (LHC beauty): è impiegato nello studio delle differenze tra materia e antimateria attraverso lo studio del beauty quark. A differenza degli altri rivelatori,

2.2. ALICE: A Large Ion Collider Experiment

non circonda il punto di collisione ma è composto da una serie di rivelatori posti uno dopo l'altro per una distanza di 20 m nella direzione *forward* della collisione;

- ALICE (A Large Ion Collider Experiment): rivelatore dedicato allo studio della fisica delle collisioni tra ioni pesanti e del QGP.

Attualmente ci sono 9 esperimenti installati a LHC, tra cui ALICE.

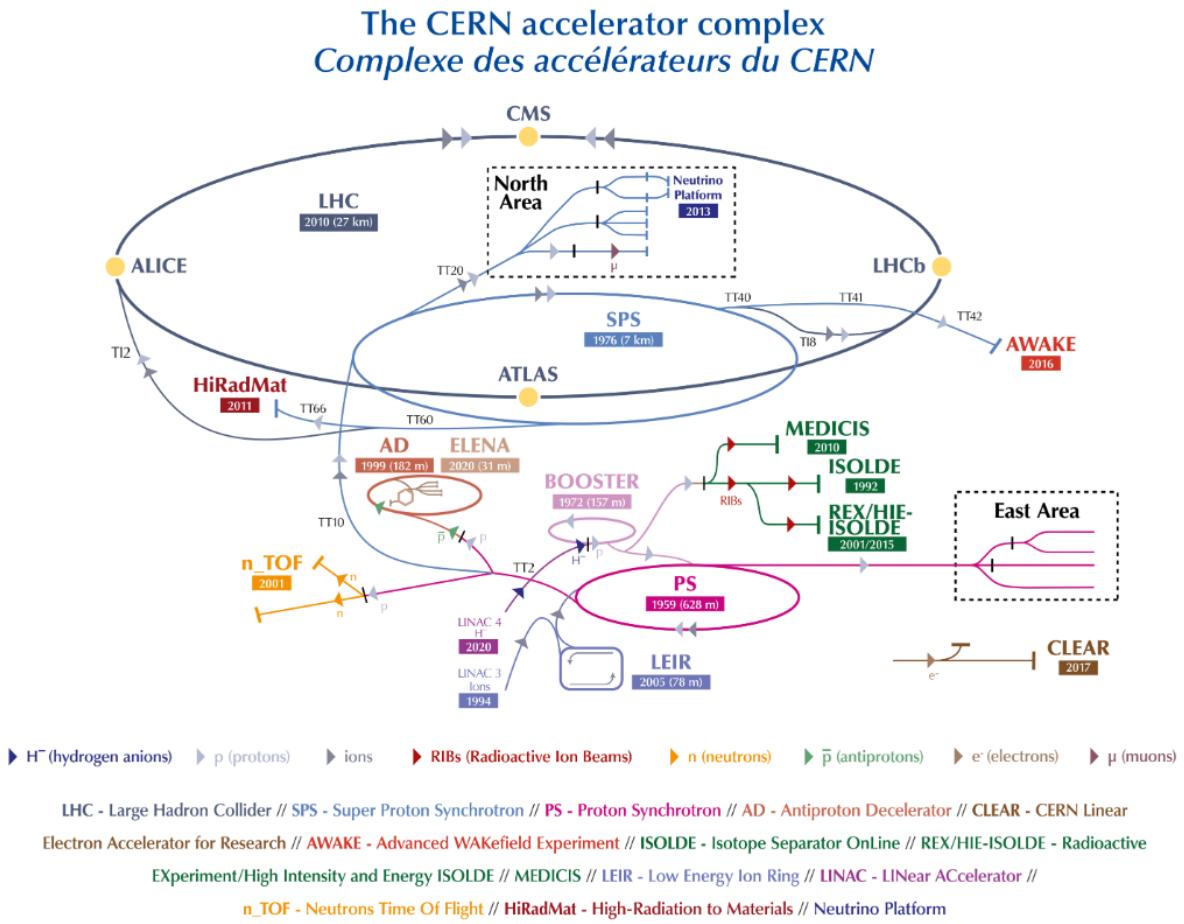


Figura 2.1: Il complesso di acceleratori del CERN [16].

2.2 ALICE: A Large Ion Collider Experiment

Il rivelatore ALICE è dedicato allo studio della QCD, in particolare attraverso lo studio della materia fortemente interagente e del Quark-Gluon Plasma, prodotto in collisioni tra ioni pesanti.

2.2. ALICE: A Large Ion Collider Experiment

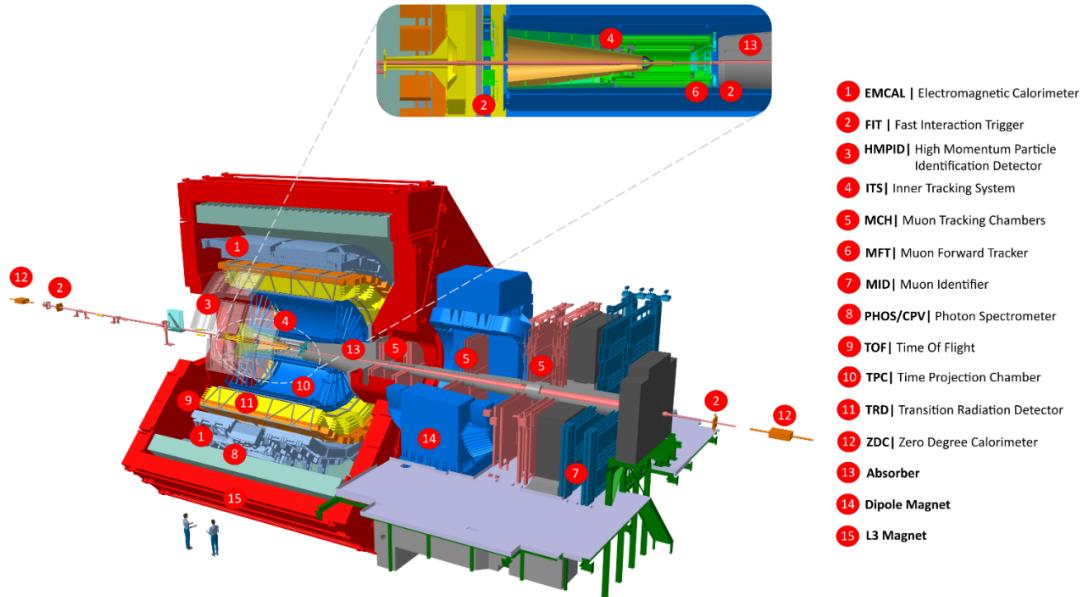


Figura 2.2: Schema del sistema di rivelatori di ALICE [17].

ALICE è formato da una parte centrale che misura adroni, elettroni e fotoni e da uno spettrometro di muoni, che si trova nella regione *forward* della collisione. La parte centrale, integrata in un solenoide riutilizzato dall'esperimento L3 a LEP, copre gli angoli polari compresi tra 45° e 135° . Contiene diversi rivelatori:

- Inner Tracking System (ITS): un sistema di tracciamento a sei livelli;
- Time Projection Chamber (TPC): il principale dispositivo di tracciamento di ALICE;
- Time-of-Flight (TOF): un rivelatore dedicato all'identificazione di particelle;
- High-Momentum Particle Identification Detector (HMPID);
- PHOS e EMCAL: due calorimetri elettromagnetici.

Tutti i rivelatori, tranne HMPID, PHOS e EMCAL coprono l'intero angolo azimutale. Inoltre, sono presenti diversi rivelatori più piccoli, come ZDC, PMD, FMD, T0, V0 per la caratterizzazione globale dell'evento e per il trigger, posizionati ad angoli più piccoli. La disposizione dei rivelatori è rappresentata in Fig. 2.2

ALICE misura caratteristiche globali degli eventi, come la molteplicità, che permettono di definire la geometria della collisione, ovvero il parametro d'impatto e i nucleoni in collisione. Viene misurata, inoltre, la produzione di svari pesanti e la frammentazione

2.2. ALICE: A Large Ion Collider Experiment

dei jet per studiare la cinematica e la perdita di energia dei partoni nel QGP. Le misure di interferenza permettono di indagare lo sviluppo spazio-temporale della collisione: fluttuazioni non statistiche segnalano la vicinanza di una transizione di fase del primo ordine, ovvero del processo di adronizzazione.

In ALICE si studiano collisioni A-A, ma anche collisioni pp e p -A, in modo da avere un confronto tra i diversi sistemi.

2.2.1 Inner Tracking System (ITS)

L'Inner Tracking System (ITS) è uno dei principali sistemi di tracciamento di ALICE, utilizzato in tutti gli ambiti dell'esperimento. I suoi compiti principali sono:

- la localizzazione dei vertici primari con risoluzione migliore di $100 \mu\text{m}$;
- la ricostruzione dei vertici secondari di decadimento di svari pesanti, mesoni D e B e iperoni, cioè barioni contenenti uno o più quark strange;
- l'identificazione e tracciamento delle particelle a basso impulso p_T ($< 200 \text{ MeV}/c$);
- il miglioramento dei parametri d'impatto e della risoluzione di impulso e angoli rispetto a TPC;
- la ricostruzione delle particelle che attraversano le *dead regions* di TPC.

ITS è formato da 6 strati cilindrici di rivelatori, di raggio compreso tra 4 e 43 cm. Il raggio esterno è scelto in modo da avere corrispondenza con le traiettorie di TPC, mentre quello interno è dovuto al raggio del *beam pipe*.

A queste energie la densità di particelle attesa è molto elevata: per gli strati più interni è $\sim 50 \text{ cm}^{-2}$, per quelli intermedi 7 cm^{-2} , mentre per quelli più esterni è circa 1 cm^{-2} . Perciò, per ottenere la risoluzione desiderata è necessario realizzare i due strati più interni con Silicon Pixel Detectors (SPD), i successivi due in Silicon Drift Detectors (SDD) e i due più esterni in Silicon micro-Strip Detectors (SSD), come rappresentato in Fig. 2.3.

I quattro strati più esterni hanno un readout analogo, perciò possono essere impiegati per l'identificazione di particelle nella regione non relativistica attraverso la misura della variazione di energia dE/dx . In questo modo ITS può essere impiegato in modo indipendente come spettrometro di particelle a basso p_T .

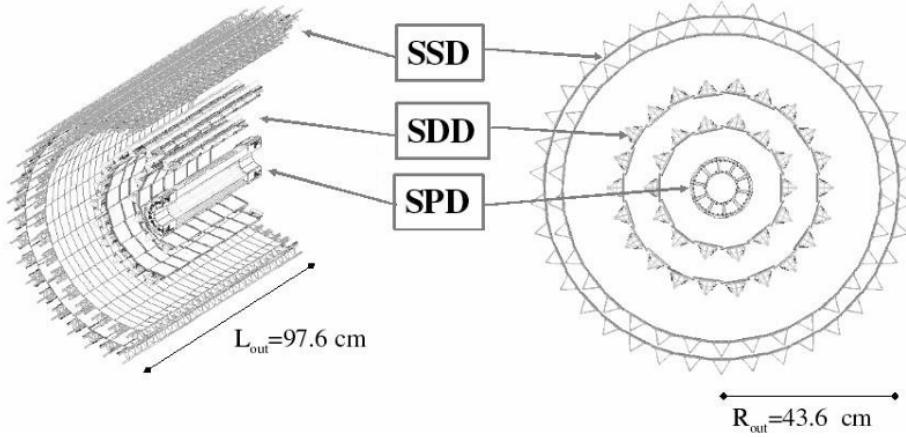


Figura 2.3: Rappresentazione di ITS. Sono riportati i materiali che compongono i diversi strati del rivelatore[18].

2.2.2 Time Projection Chamber (TPC)

La Time Projection Chamber (TPC) è il principale rivelatore di tracciamento della regione centrale. I suoi scopi sono la misura dell’impulso di particelle cariche, l’identificazione di particelle e la determinazione di vertici. La sua accettanza copre tutto l’angolo azimutale e l’intervallo di pseudorapidità $|\eta| < 0.9$. TPC ha un range di impulso molto ampio, da 0.1 GeV/c a 100 GeV/c.

TPC è un rivelatore cilindrico con l’asse coincidente con il *beam pipe*, di raggio interno 85 cm e raggio esterno 250 cm, con una lunghezza di 500 cm. Il suo volume, 90 cm^3 , è riempito di una miscela gassosa di Ne-CO₂-N₂ (90-10-5). La miscela è stata scelta in base alla velocità di deriva e alla diffusione degli elettroni in tale gas, oltre che alla sua stabilità. È presente un sistema di purificazione del gas attraverso minime immissioni di gas “nuovo”. Nel gas è presente un campo elettrico altamente uniforme, generato da una *field cage*. La *field cage* è formata da un elettrodo centrale ad elevato voltaggio e da due divisorì di potenziale assiali. A causa della scelta della miscela è necessario mantenere un campo di 400 V/cm. La *field cage* è isolata attraverso un volume di CO₂ gassoso. Sia la *field cage* che il volume isolante sono cilindri coassiali al TPC, saldati a entrambe le piastre finali. Lo schema di TPC è in Fig. 2.4.

Le particelle cariche prodotte dalla collisione che attraversano la miscela ionizzano il gas, liberando elettroni. Questi elettroni, accelerati dal campo elettrico, sono trasportati fino ad una delle due piastre finali, dove avviene il readout del segnale. Ciascuna è divisa in 18 sezioni trapezoidali. È necessario, inoltre, un sistema di amplificazione del segnale dato dagli elettroni. Questo compito è svolto dalle camere GEM (Gas Electron

2.2. ALICE: A Large Ion Collider Experiment

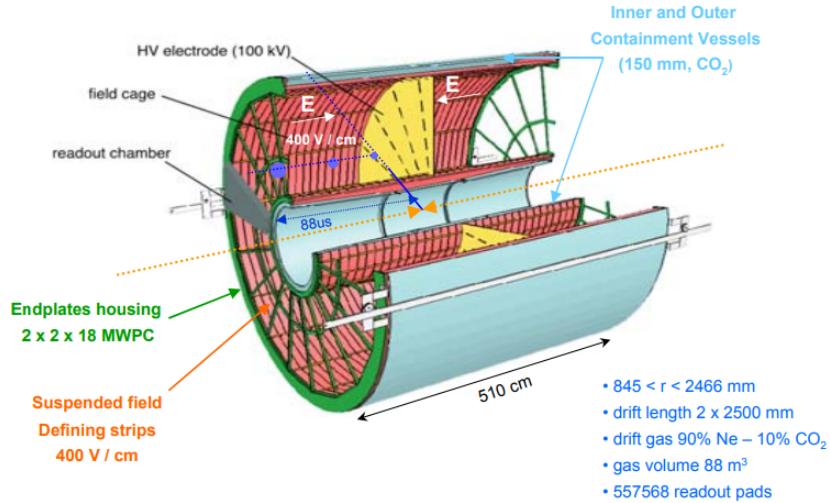


Figura 2.4: Schema della Time Projection Chamber.

Multipliers), situate sulle piastre finali. Il sistema di readout è composto da un numero molto elevato di blocchi, ciascuno dei quali costituisce l'anodo di una camera GEM. Ogni blocco di readout è connesso a un canale elettronico di front-end, dove i segnali ottenuti vengono digitalizzati. TPC ha una velocità di trasmissione dei dati estremamente elevata, fino a 3.28 TB/s [19].

2.2.3 Time-of-Flight (TOF)

Il rivelatore a tempo di volo (Time of Flight detector, TOF) misura con elevata precisione il tempo che una particella impiega per viaggiare dal vertice al rivelatore. Copre la regione centrale di pseudorapidità $|\eta| < 0.9$ e identifica le particelle in un range medio di impulso, tra $0.5 \text{ GeV}/c$ e $2.5 \text{ GeV}/c$. Le sue misure, accoppiate a quelle di ITS e TPC sulla lunghezza di traccia L e l'impulso p , permettono di calcolare la massa delle particelle, e quindi la loro identificazione. Infatti, dalla relazione relativistica energia-impulso

$$E^2 = m^2 c^4 + p^2 c^2 \quad (2.1)$$

e considerando la definizione di energia come componente temporale del quadrimpulso $E = \gamma mc^2$, dove γ è il fattore di Lorentz, otteniamo la massa della particella:

$$m = \frac{p}{c} \sqrt{\frac{c^2 t^2}{L^2} - 1} \quad (2.2)$$

TOF, rappresentato in Fig. 2.5, ha forma cilindrica, coassiale al beam pipe, posto a circa 3.8 m dal centro della collisione. Copre l'angolo polare tra i 4° e i 135° e tutto

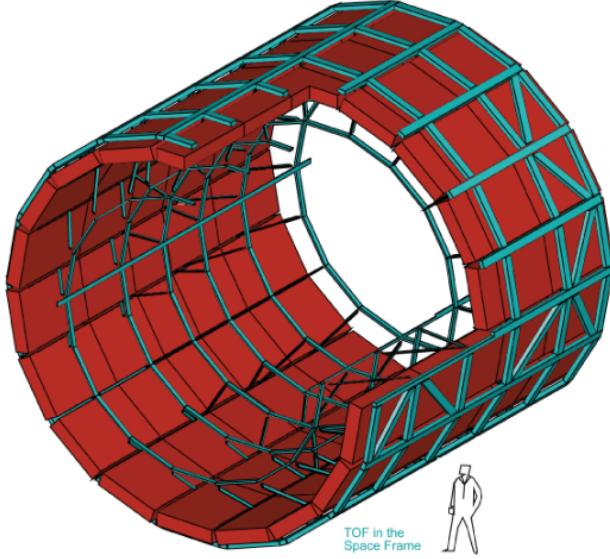


Figura 2.5: Rappresentazione schematica del rivelatore TOF [21].

l’angolo azimutale. Le unità fondamentali del rivelatore sono le Multigap Resistive Plate Chamber (MRPC): due pile di 6 sottili piani resistivi di vetro separati da spazi riempiti di gas, di spessore $\sim 250 \mu\text{m}$. Alle estremità delle pile sono posti due catodi, un anodo è posto nel centro: viene prodotto un campo elettrico uniforme e intenso. Quando una particella attraversa il gas, lo ionizza, analogamente a quanto accade nella TPC. Per la presenza del campo, gli elettroni prodotti acquisiscono energia e provocano un’ulteriore ionizzazione, generando una valanga elettronica. Il segnale così prodotto viene raccolto da sensori di rame, i *pickup pads*, posti al di fuori delle lastre esterne. Ciascuna di queste celle ha una risoluzione temporale di 50 ps [20].

Possiamo esprimere la capacità di TOF di distinguere due particelle differenti, di massa m_1 e m_2 , caratterizzate dalla stessa lunghezza di traccia e stesso impulso, ma con diverso tempo di volo in funzione della risoluzione temporale totale. Infatti, dall’eq. 2.2 si ricava:

$$t = \frac{L}{c} \sqrt{\frac{m^2 c^2}{p^2} + 1} \quad (2.3)$$

Nel limite di impulsi elevati ($m^2 c^2 / p^2 \rightarrow 0$), svolgendo uno sviluppo di Taylor:

$$t = \frac{L}{c} \left(\frac{m^2 c^2}{2 p^2} + 1 \right) \quad (2.4)$$

Perciò, la differenza tra i tempi di volo delle due particelle è

$$\Delta t = \frac{L c}{2} \cdot \frac{m_1^2 - m_2^2}{p^2} \quad (2.5)$$

2.2. ALICE: A Large Ion Collider Experiment

Otteniamo quindi la capacità di risolvere le due particelle in numero di deviazioni standard:

$$n_\sigma = \frac{\Delta t}{\delta t} \quad (2.6)$$

dove δt è la risoluzione temporale totale di TOF. In un esperimento di collisioni Pb-Pb a $\sqrt{s_{NN}} = 2.76$ TeV con una risoluzione temporale di ~ 80 ps si ottiene una separazione maggiore di 3σ per pioni e kaoni fino a 2.5 GeV/c e per kaoni e protoni fino a 4 GeV/c [22]. In Fig. 2.6 sono riportate le misure delle velocità $\beta = v/c$ in funzione dell'impulso di tale collisione.

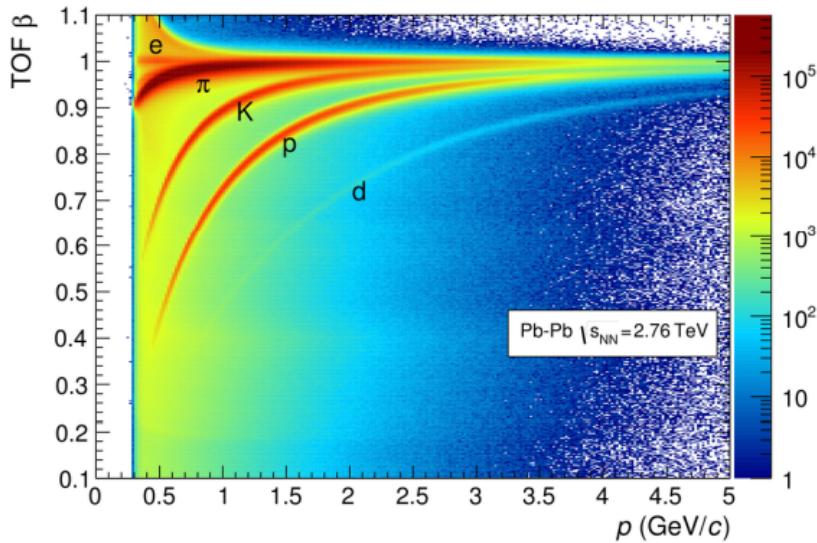


Figura 2.6: Velocità β calcolata a partire dalle misure dei tempi di volo di TOF in funzione dell'impulso in una collisione Pb-Pb con $\sqrt{s_{NN}} = 2.76$ TeV [23].

2.2.4 High-Momentum Particle Identification Detector (HMPID)

L'High-Momentum Particle Identification Detector (HMPID) è un rivelatore specializzato nella rivelazione di adroni ad elevato impulso trasverso ($p_T > 1$ GeV/c), permette di riuscire ad identificare anche quelle particelle non coperte dalle misure di ITS, TPC e TOF. È posto a 5 m dal centro della collisione, ed è formato da 7 rivelatori RICH (Ring-Imaging Cherenkov detector) $1.4\text{ m} \times 1.3\text{ m}$ ciascuno, rappresentati in Fig. 2.7.

I rivelatori RICH sfruttano l'effetto Cherenkov: una particella carica che attraversa un mezzo denso con velocità superiore a quella della luce in quel mezzo provoca l'emissione di radiazione elettromagnetica, detta radiazione di Cherenkov, che si propaga con un fronte

2.2. ALICE: A Large Ion Collider Experiment

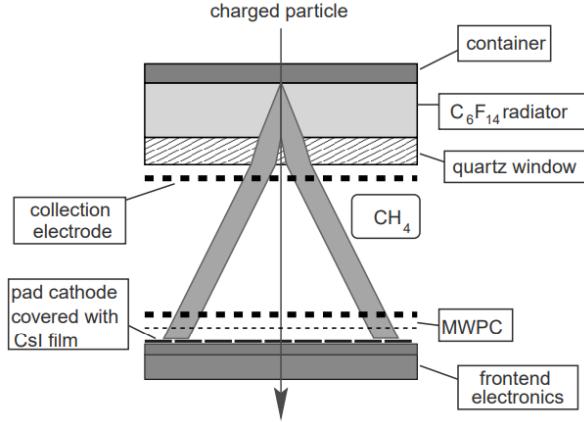


Figura 2.7: *Rappresentazione di un'unità fondamentale di HMPID, un rivelatore RICH [18].*

d'onda conico. L'emissione avviene poiché la particella carica polarizza gli atomi del mezzo lungo la sua traccia, trasformandoli in dipoli. La variazione nel tempo della carica di dipolo porta all'emissione della radiazione. Attraverso considerazioni geometriche si può ottenere la relazione tra $\beta = v/c$, con v velocità della particella, e l'angolo di emissione del fotone θ_C . Nell'intervallo di tempo t la particella percorre la distanza $AB = t\beta c$, e il fotone la distanza $AC = tc/n$, n indice di rifrazione del mezzo. Poiché, come si vede in Fig. 2.8, $AC = AB \cos \theta_C$:

$$\cos \theta_C = \frac{1}{n\beta} \quad (2.7)$$

da cui si ottiene che l'effetto Cherenkov è un effetto a soglia. Infatti, l'emissione avviene se e solo se

$$\beta \geq \frac{1}{n} \quad (2.8)$$

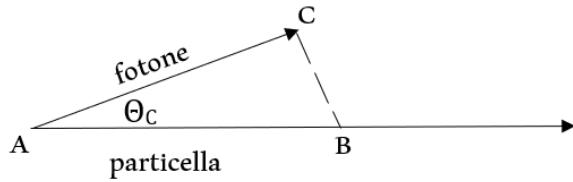


Figura 2.8: *Emissione di un fotone con angolo θ_C rispetto alla direzione di volo della particella per effetto Cherenkov.*

I RICH sono formati da due componenti: un radiatore contenente materiale dielettrico (C_6F_{12}), in cui viene emessa la radiazione, e un contatore di fotoni. Il contatore di fotoni è una camera proporzionale multifili (Multiwire Proportional Chamber, MWPC), il cui catodo è ricoperto da un sottile strato di CsI [24]. Dai contatori di fotoni vengono prodotte immagini di corone circolari i cui raggi interni ed esterni sono proporzionali a θ_C : misurandoli riusciamo a risalire alla velocità delle particelle tramite eq. 2.7. Noti velocità e impulso di una particella possiamo risalire alla sua massa e quindi alla sua identità, come visto nella sezione precedente.

2.2.5 Calorimetri elettromagnetici

Un calorimetro converte l'energia di una particella in un segnale, che può essere di diversi tipi, mantenendo la proporzionalità tra l'energia e il segnale. I calorimetri sono vantaggiosi poiché, a differenza degli spettrometri, la loro precisione aumenta all'aumentare dell'energia. Nel set-up di ALICE sono presenti due calorimetri elettromagnetici: PHOS (PHOton Spectrometer) e EMCal.

PHOS è un calorimetro ad alta risoluzione posizionato a 460 cm dal centro della collisione, che copre la regione centrale di pseudorapidità $|\eta| < 1.2$ [18] e un angolo azimutale di 70° . PHOS misura la radiazione elettromagnetica del QGP e ricostruisce lo spettro degli adroni tramite i loro decadimenti radiativi. Gli scopi di queste misure sono l'indagine delle prime fasi del sistema in seguito alla collisione e lo studio del jet quenching, attraverso le misure delle correlazioni tra π^0 e getti di fotoni. Inoltre, PHOS funziona come sistema di trigger: invia segnali al sistema di trigger centrale di ALICE in modo da selezionare solamente gli eventi in cui sono presenti fotoni altamente energetici. PHOS è formato da 12544 canali a base di tungstato di piombo $PbWO_4$, rappresentati in Fig. 2.9 [25]. Quando questo materiale è attraversato da una particella, produce luce nel visibile, che viene raccolta e trasformata in segnale da un fotodiode a valanga. Ogni canale è accoppiato a un CPV (Charged Particle Veto detector). In questo modo si ottiene un'efficienza di rivelazione di particelle cariche superiore al 99%.

EMCal (Electro Magnetic Calorimeter) è impiegato nello studio del jet quenching, riuscendo a misurare anche le componenti neutre dei jet. Copre la regione di pseudorapidità $|\eta| < 0.7$ e un angolo azimutale di ampiezza $\Delta\phi = 107^\circ$ [26]. Inoltre, è disposto in modo diametralmente opposto a PHOS. È un calorimetro a campionamento, ovvero costruito alternando strati di materiale assorbitore, in questo caso il piombo, con strati di materiale scintillatore. È formato da 4416 moduli raggruppati in 20 Super Modules, rappresentati in Fig. 2.9. Ciascun modulo è formato da 4 torri otticamente isolate, in grado di trasferire segnali tra due circuiti isolati usando la luce. Attraverso un fotodiode a valanga si ottiene il readout ottico [27]. Poiché EMCal deve essere integrato nel set-up di ALICE, i Super Modules sono realizzati in diverse scale: intera, 2/3 e 1/3. Dieci Super Modules full-size e due 1/3 formano EMCal.

2.2. ALICE: A Large Ion Collider Experiment

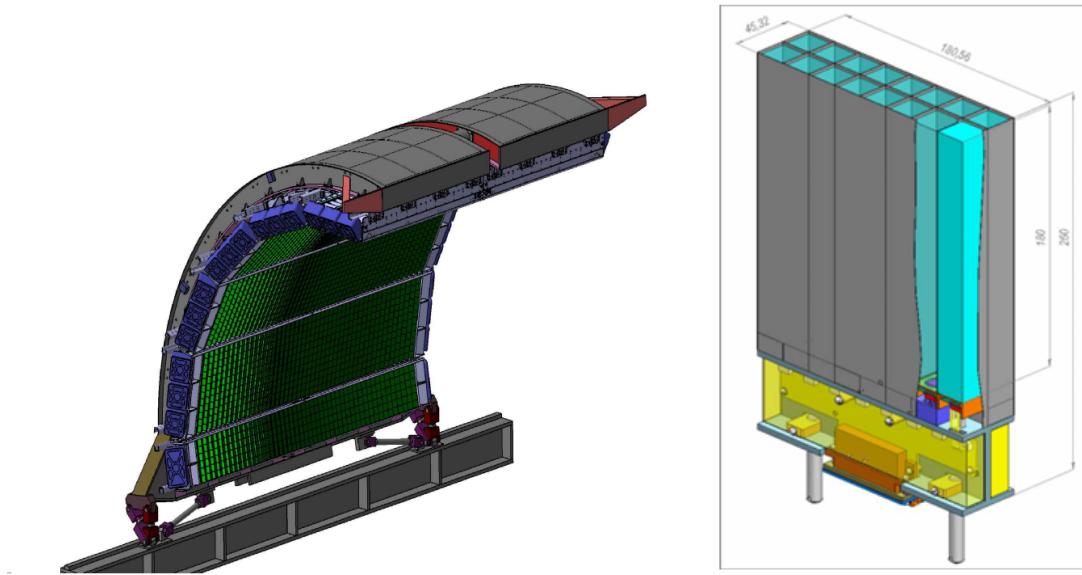


Figura 2.9: *A sinistra:* i Super Modules di EMCal installati nelle loro strutture di supporto [26]. *A destra:* un sistema di 8×2 unità di cristalli rivelatori di PbWO_4 che compongono PHOS [18].

2.2.6 Forward Muon Spectrometer (FMS)

Il Forward Muon Spectrometer (FMS) è posizionato nella regione *forward* di ALICE. Misura lo spettro di mesoni composti da quark pesanti come J/ψ , ψ' , Υ , Υ' , Υ'' attraverso il canale di decadimento $\mu^+ \mu^-$. La misura di tutte queste specie svolta con lo stesso apparato permette il confronto diretto dei loro rapporti di produzione, fornendo prove dell'esistenza del QGP, come visto in Sez. 1.6. FMS copre l'intervallo di pseudorapidità $-4 < \eta < -2.5$ e un angolo polare compreso tra 171° e 178° .

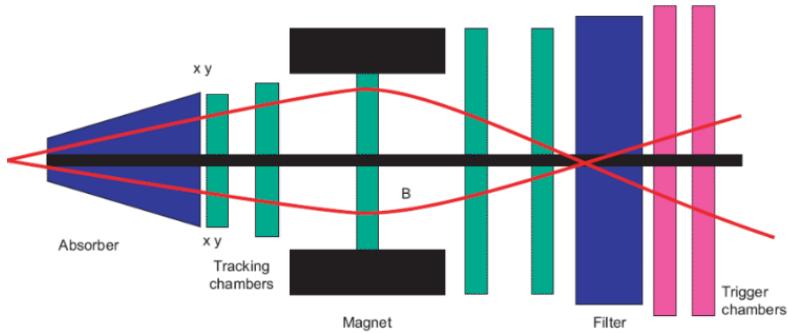


Figura 2.10: *Rappresentazione schematica della sezione longitudinale di FMS* [28].

Come rappresentato in Fig. 2.10, è presente un assorbitore, realizzato tramite uno strato di carbonio e calcestruzzo che sopprime tutte le particelle ad eccezione dei muoni del vertice d'interazione. Infatti, negli eventi ad elevata rapidità vengono prodotte particelle primarie e secondarie che compongono il background delle misure di FMS. Oltre all'assorbitore, per prevenire che queste particelle arrivino ai rivelatori, è presente uno strato schermante di tungsteno, piombo e acciaio inossidabile intorno al *beam pipe*. Dopo l'assorbitore sono presenti 10 rivelatori planari, alcuni dei quali posti all'interno di un magnete a dipolo. Infine, è presente un filtro passivo di muoni e quattro camere planari di trigger, che selezionano i decadimenti di quark pesanti in base all'impulso trasverso dei muoni [18].

2.2.7 Altri rivelatori

Sono presenti anche altri piccoli rivelatori specializzati, posti al di fuori della regione centrale. Svolgono misure ad angoli molto piccoli, quindi ad elevata pseudorapidità.

- ZDC (Zero Degree Calorimeter): fornisce una misura della centralità della collisione attraverso la misura dell'energia dei nucleoni spettatori. Infatti, a partire da tale misura e nota l'energia per nucleone specifica del tipo di collisione si può ottenere il numero di spettatori e, per sottrazione, il numero dei partecipanti. È formato da due sistemi identici di calorimetri, ciascuno formato da un calorimetro specializzato nella rivelazione di neutroni (ZN) e uno nella rivelazione di protoni (ZP). I due sistemi sono posti ad una distanza di 116 m dal punto di collisione, ai lati opposti [18].
- PMD (Photon Multiplicity Detector): misura, evento per evento, la molteplicità e la distribuzione spaziale dei fotoni con pseudorapidità $2.3 < \eta < 3.9$ [18].
- FMD (Forward Multiplicity Detector): effettua misure di molteplicità delle particelle cariche con pseudorapidità $-3.4 < \eta < -1.7$ e $1.7 < \eta < 5$ [18].
- T0: la sua funzione principale è fornire il tempo iniziale per la misura di tempo di volo di TOF, ovvero determinare l'istante della collisione. Inoltre effettua la misura della posizione del vertice di interazione, in modo da distinguere le interazioni tra particelle del fascio da quelle tra particelle del fascio e del gas, fornendo quindi un segnale di trigger. È formato da due array di contatori Cherenkov, T0A e T0C, ciascuno contenente 12 contatori. Gli array sono posti, rispettivamente, a una distanza di 375 cm e 72.7 cm dal punto di interazione, nelle direzioni opposte. Coprono l'intervallo di pseudorapidità $4.61 < \eta < 4.92$ e $-3.28 < \eta < -2.97$.
- V0: il rivelatore che fornisce il trigger per gli altri rivelatori della regione centrale, in base alla centralità della collisione. Due scintillatori di diametro 90 cm e 64 cm

2.2. ALICE: A Large Ion Collider Experiment

e spessore 2.5 cm e 2 cm, posti a distanza 340 cm e 90 cm dal punto di collisione, V0A e V0C, in Fig.2.11, compongono il rivelatore. In questo modo riesce a coprire un'ampia regione di pseudorapidità [18].

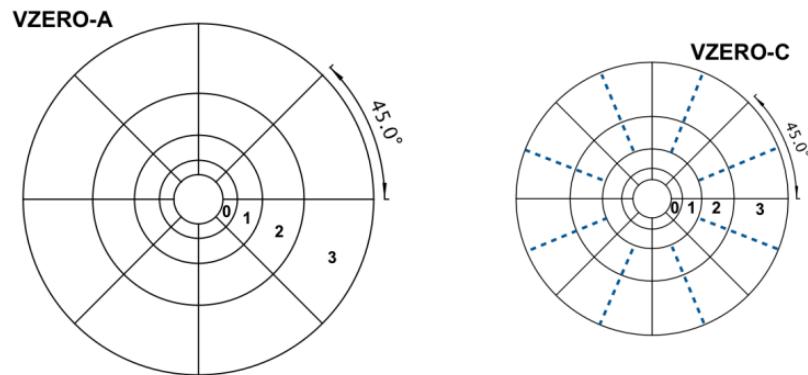


Figura 2.11: *Schema degli scintillatori V0A e V0C in cui è mostrata la loro segmentazione* [29].

Capitolo 3

Ricostruzione del bario Λ_c^+

3.1 Introduzione

Il bario Λ_c^+ è il bario contenente un quark charm più leggero, e si può osservare in un intervallo molto ampio di impulsi. Come visto in Sez. 1.8, lo studio dei rapporti di produzione di questo bario, in particolare attraverso lo studio del rapporto di produzione bario/mesone Λ_c^+/D^0 , è utile per la comprensione dei meccanismi di adronizzazione.

Λ_c^+ è uno dei barioni lambda, ovvero quei barioni che contengono un quark up, un quark down e un terzo quark che può essere uno strange (Λ^0), un charm (Λ_c^+) o un bottom (Λ_b^0). Ha una massa a riposo di (2286.46 ± 0.14) MeV/ c^2 , un isospin nullo e uno spin $J^P = 1/2$. La sua vita media è $\tau = (2.00 \pm 0.06) \cdot 10^{-13}$ s [30], molto breve. Sono possibili diversi canali di decadimento, ma nell'esperimento ALICE ne vengono studiati solo tre:

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

$$\Lambda_c^+ \rightarrow p K_S^0 \quad (3.2)$$

$$\Lambda_c^+ \rightarrow \Lambda e^+ \nu_e \quad (3.3)$$

i primi due adronici, il terzo semileptonico. I Branching Ratio di questi canali, ovvero la frazione di Λ_c^+ che decade in ciascun canale sono: $(6.28 \pm 0.32)\%$ per 3.1, $(1.59 \pm 0.08)\%$ per 3.2 e $(3.6 \pm 0.4)\%$ per 3.3 [31]. Considereremo, d'ora in poi, solo il canale 3.2.

Il bario Λ_c^+ è prodotto dall'urto tra due protoni nel punto che chiamiamo vertice primario. Il suo decadimento avviene, invece, nel vertice secondario. Anche il mesone K_S^0 prodotto nel decadimento ha vita media molto breve e decade per decadimento debole in una coppia $\pi^+ \pi^-$. I pioni carichi prodotti in questo decadimento sono le particelle che effettivamente vengono rivelate da ALICE. Nel sistema di riferimento proprio, la Λ_c^+ decade dopo aver percorso mediamente la distanza $c\tau = 60 \mu\text{m}$. I due vertici sono quindi molto vicini: per impulsi fino a 4 GeV/ c la distanza tra vertice primario e secondario è

3.2. Reti Neurali Grafiche

inferiore a $100 \mu\text{m}$. I localizzatori di vertici di ALICE, come ITS, hanno una precisione di circa $100 \mu\text{m}$, come visto in Sez. 2.2.1. Non è possibile quindi distinguere il vertice primario dal secondario, perciò è necessario usare altri metodi per distinguere il segnale, ovvero le Λ_c^+ , dal background, tutte quelle particelle con caratteristiche simili al segnale ma non provenienti dal vertice secondario. Inoltre, il rapporto segnale fondo è molto basso. Risulta quindi conveniente usare metodi di Machine Learning per distinguere le particelle del segnale da quelle del background. In questa tesi è stato impiegato il modello delle Reti Neurali Grafiche (Graph Neural Network, GNN).

3.2 Reti Neurali Grafiche

Un grafo è un modo per rappresentare un insieme di oggetti e le relazioni che esistono tra loro. Gli oggetti prendono il nome di nodi o vertici, le relazioni tra loro vengono dette *edges*. Per una migliore rappresentazione, è possibile aggiungere informazioni che ne descrivono le proprietà, i *features*, a ciascun elemento del grafo e anche al grafo complessivamente. Inoltre, è possibile distinguere i grafi in direzionali e non direzionali: nel primo caso è associata una direzionalità agli *edges*, nel secondo no.

I grafi e le Reti Neurali Grafiche sono impiegati per la rappresentazione e lo studio di reti sociali, molecole, immagini e testi. In ambito fisico sono state impiegate per lo studio di sistemi complessi di particelle interagenti [32].

3

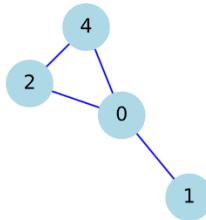


Figura 3.1: *Rappresentazione grafica di un grafo non direzionale con cinque nodi e quattro edges.*

Esistono diversi modi per rappresentare la connettività di un grafo. Tutti devono essere invarianti per permutazione: ordinando i nodi in modo diverso, la connettività non

3.2. Reti Neurali Grafiche

deve variare. Uno di questi modi è la matrice di adiacenza: una matrice $N_{nodi} \times N_{nodi}$ il cui elemento ij è non nullo se e solo se è presente un *edge* tra il nodo i -esimo e il nodo j -esimo. Considerando il grafo in Fig. 3.1, la sua matrice di adiacenza è in Tab. 3.1.

	0	1	2	3	4
0	0	1	1	0	1
1	1	0	0	0	0
2	1	0	0	0	1
3	0	0	0	0	0
4	1	0	1	0	0

Tabella 3.1: *Matrice di adiacenza del grafo in Fig. 3.1. La quarta riga e la quarta colonna sono nulle poiché il nodo 3 è isolato. La diagonale è nulla dato che nessun nodo è in relazione con se stesso.*

Le Graph Neural Network (GNN), o Reti Neurali Grafiche, sono reti neurali che agiscono su grafi, acquisendo informazioni sulle relazioni del grafo tramite l'aggregazione delle informazioni di nodi vicini, senza modificarne la connettività. Per sfruttare la connettività del grafo per ottenere previsioni, le GNN usano layer *message passing* [33]. Questo tipo di layer lavora in tre passaggi. Per prima cosa, per ogni nodo del grafo vengono raccolte tutte le informazioni dei nodi vicini. Il numero dei nodi vicini dipende dal grafo. Le informazioni, o messaggi, raccolte vengono aggregate tramite una funzione di aggregazione, come la somma. Infine, i messaggi aggregati diventano l'input di una funzione di update, che aggiorna i parametri della GNN. I *message passing* layers possono essere impilati l'uno sull'altro: aumentando i layer ogni nodo ha accesso alle informazioni di nodi sempre più lontani, eventualmente dell'intero grafo. La trasmissione di messaggi può avvenire sia tra nodi che tra *edges*.

Con le GNN possono essere risolti tre diversi tipi di problemi:

- Graph level: lo scopo è prevedere una proprietà dell'intero grafo. Un problema di questo tipo è l'identificazione di molecole o di una loro proprietà.
- Edge level: la GNN prevede le proprietà delle relazioni tra i nodi. Questo tipo di problema è tipico delle applicazioni di computer vision, nella comprensione della scena rappresentata da un'immagine.
- Node level: la GNN svolge previsioni sulle proprietà di un singolo nodo, come la sua identità. L'identificazione dei barioni Λ_c^+ dal background è un problema di questo tipo.

Le Reti Neurali sono formate da diversi neuroni, ognuno dei quali applica al dato in input una specifica funzione che dipende da alcuni parametri, numeri reali che prendono

il nome di pesi e bias. Lo scopo dell’allenamento di una Rete Neurale è determinare l’insieme dei pesi e bias che permettono al modello di effettuare le migliori previsioni possibili. Quando il modello viene inizializzato, ai pesi e al bias sono assegnati valori casuali piccoli [34]. L’allenamento delle GNN è supervisionato, ovvero avviene su dati di cui conosciamo l’identità `y_true`. La funzione che guida l’allenamento del modello è la *loss function*: è una funzione che calcola la distanza tra le previsioni del modello `y_pred` e la vera identità dei dati di allenamento `y_true`. Tutte le funzioni che agiscono nelle GNN, come le funzioni di aggregazione e di update, sono differenziabili. In particolare lo è la loss function. Perciò possiamo determinare il minimo della loss function, e quindi il set di parametri ideale del modello, trovando i punti in cui il suo gradiente è nullo. Tuttavia, a causa dell’elevato numero di parametri, non è possibile risolvere questo problema analiticamente. Perciò, i dati utilizzati per l’allenamento vengono suddivisi in *batch*: dopo che il modello viene eseguito su un *batch* si calcola il gradiente della loss function, in base al quale si aggiornano i pesi del modello. Il modello viene iterato sui dati di allenamento diverse volte. Ogni iterazione viene detta epoca. Un algoritmo di questo tipo si dice Stochastic Gradient Descent (SGD) [35]. Esistono diverse versioni di questo algoritmo. L’ottimizzatore è la funzione che definisce quale di queste è implementata nello specifico modello, poiché è la funzione che determina come vengono aggiornati pesi e bias. La velocità con cui i parametri vengono modificati è detta *learning rate*.

Esistono diverse loss function, si deve scegliere quella più adatta in base al tipo di problema. In questo caso il problema è di classificazione binaria: gli elementi del segnale appartengono alla classe 1, quelli del background alla classe 0. Perciò la loss function più adatta è la *binary crossentropy*, che misura la distanza tra due distribuzioni di probabilità. È definita come:

$$BCE = -\frac{1}{N} \sum_{i=1}^N [y_{\text{true}}_i \log(p_i) + (1 - y_{\text{true}}_i) \log(1 - p_i)] \quad (3.4)$$

dove N è il numero di dati su cui è stato eseguito il modello e p_i è la probabilità prevista dal modello che l’ i -esimo dato sia di classe 1.

3.3 Keras e Spektral

Keras è una API (Application Programming Interface) per il deep learning in Python. Fornisce strumenti di alto livello per la definizione e l’allenamento di molti tipi di modelli di deep learning. Keras, inizialmente, era un’interfaccia indipendente, con Tensorflow, Microsoft Cognitive Toolkit e Theano come librerie di back-end. Con il rilascio di Tensorflow 2.0, Keras è stato incorporato in Tensorflow come `tf.keras`. Tensorflow è una libreria open-source sviluppata dal Google Brain Team impiegata principalmente per il Machine Learning, in particolare per le Reti Neurali. Può essere usata in diversi linguaggi di programmazione: Python, JavaScript, Java e C++.

3.3. Keras e Spektral

L’idea di Keras è di fornire strumenti semplici da usare, che permettano di non concentrarsi sui dettagli matematici dei tensori e delle loro forme, ma anche flessibili, con funzioni che possono essere estese dagli utenti esperti. Sono presenti moduli per l’allenamento, la definizione e la valutazione delle prestazioni dei modelli.

Uno dei modelli di Rete Neurale più usati di Keras è **Sequential**. Costruisce un modello con una struttura lineare di livelli impilati in cui l’output di un layer è l’input del successivo. Tutti i layer devono necessariamente avere sia come input che come output un solo tensore. Trattando GNN non è possibile usare **Sequential**, poiché sono necessari due o più input per ogni livello: i nodi e la matrice di adiacenza, ed eventualmente gli *edges*. Diventa necessario usare ulteriori strumenti per adattare il tipo di dato alle funzioni di Keras, come Spektral. Spektral [36] è una libreria di Python basata sulla API Keras e su Tensorflow 2.0. Fornisce strumenti per lavorare in modo semplice e flessibile con le Reti Neurali Grafiche. Può essere usata insieme a Keras senza generare problemi, in modo da poter continuare ad usare gli strumenti di Keras fondamentali per il deep-learning come il training loop e gli strumenti di valutazione del training.

3.3.1 Grafi in Spektral

In Spektral i grafi possono essere rappresentati come istanze della classe `spektral.data.Graph`. Gli oggetti di questa classe hanno quattro attributi principali: la matrice di adiacenza `a`, il tensore delle proprietà dei nodi (*node features*) `x`, il tensore delle proprietà degli *edges* (*edge features*) `e` e il tensore dei labels `y`. Ogni istanza della classe può avere tutti, nessuno o solo alcuni di questi attributi.

La matrice di adiacenza può essere un array denso di Numpy, ma questo comporta un elevato costo in termini di memoria, dato che è una matrice sparsa: molti dei suoi elementi sono nulli. Per migliorare la gestione della memoria, `a` può essere una matrice sparsa di Scipy. In questo modo vengono inseriti in memoria solamente gli elementi non nulli della matrice e i rispettivi indici. Questo formato prende il nome di COOrdinate format, lo stesso usato da TensorFlow per gli sparse tensor. In questo formato una matrice è rappresentata da una `list` di tuple `[r, c, v]`: `r` indica l’indice di riga, `c` l’indice di colonna e `v` il valore non nullo dell’elemento, in questo caso 1. La maggior parte dei modelli di Spektral richiede che la matrice `a` sia di questo tipo. Anche il tensore degli *edge*, di forma $N_{nodi} \times N_{nodi} \times N_{edge-features}$, con $N_{edge-features}$ il numero di proprietà di ciascun *edge*, è una matrice sparsa e si rappresenta in questo formato.

Il tensore delle proprietà dei nodi `x` è una matrice $N_{nodi} \times N_{features}$, con $N_{features}$ il numero delle proprietà dei nodi. In questo caso non sono presenti zeri inutili, quindi `x` è un array di Numpy.

I labels possono appartenere sia all’intero grafo che al singolo nodo. Nel primo caso possono essere scalari o array di Numpy unidimensionali di lunghezza N_{labels} , con N_{labels} il numero di labels. Nel secondo caso possono essere array di Numpy unidimensionali di lunghezza N_{nodi} , nel caso di un label scalare per ciascun nodo, o di forma $N_{nodi} \times N_{labels}$.

3.4. Dati e variabili fisiche

È utile definire il dataset, un container di oggetti di tipo `Graph`. Per rappresentarlo possiamo usare istanze della classe `spektral.data.dataset.Dataset`. È possibile personalizzare i dataset definendo classi che ereditano da `Dataset`. Per farlo è necessario ridefinire almeno il metodo `read()`, che restituisce una `list` di oggetti di tipo `Graph`: permette quindi di leggere gli elementi del dataset.

3.3.2 Loaders

Uno dei problemi principali delle GNN è la definizione dei *batch*. Infatti, non potendo modificare la connettività del grafo, la suddivisione del campione in sottoinsiemi non è banale. In Spektral sono definiti degli oggetti, i `Loader`, che iterano su un dataset e producono *batch* che possono essere forniti ad un modello di Keras. Il metodo fondamentale di questa classe è `load()`: restituisce un oggetto che può essere passato come argomento alle funzioni `fit()`, `predict()` e `evaluate()` di un modello di Keras.

A seconda del tipo di grafo (*data mode*), è necessario usare un diverso `Loader`. In Spektral ne esistono quattro, che si distinguono per il numero di grafi nel dataset e per la forma della matrice di adiacenza. Nel problema trattato, il dataset è composto da un singolo grafo, in cui ogni nodo è una particella misurata. Questo tipo di dataset è del tipo *single mode* e il `Loader` adatto è `SingleLoader`. Per inizializzare un oggetto di questo tipo sono necessari quattro argomenti:

- `dataset`: un oggetto di tipo `Dataset` o di una sua classe figlia contenente un solo grafo;
- `epochs`: un intero, il numero di iterazioni che il `Loader` svolge sul datatset. Di default è `None`, il numero di iterazioni è indefinito;
- `shuffle`: un oggetto di tipo booleano, definisce se i dati verranno rimescolati all'inizio di ogni epoca oppure no;
- `sample_weights`: un array con lunghezza pari al numero di nodi. Sono i pesi da applicare al calcolo della loss function per ciascun dato. `None` di default.

L'oggetto inizializzato dal costruttore è la tupla `(inputs, labels)` o `(inputs, labels, sample_weights)`, nel caso in cui l'ultimo argomento non sia `None`. Con `inputs` è indicata la tupla contenente gli attributi `graph.x`, `graph.a`, `graph.e` del grafo e con `labels` l'attributo `graph.y`.

3.4 Dati e variabili fisiche

Il campione di dati utilizzato per gli eventi del segnale è formato da eventi simulati generati dal generatore Monte Carlo PYTHIA8 [37]. È stato imposto un vincolo sulla

3.4. Dati e variabili fisiche

generazione: deve essere presente almeno un decadimento della Λ_c^+ attraverso il canale 3.2 tra tutti gli eventi generati. Inoltre, per rendere il campione ancora più verosimile, sono stati inseriti con GEANT3 [38] gli effetti della presenza del rivelatore ALICE. In questo modo si tengono in considerazione anche le interazioni delle particelle con i materiali che compongono i rivelatori e la formazione dei segnali di risposta.

Il campione di dati utilizzato per il background, invece, è composto da misure provenienti da ALICE, ottenute nella seconda fase di presa dati (RUN 2), avvenuta dal 2016 al 2018. Tra i dati reali sono stati selezionati solo eventi con una massa invariante ricostruita non compatibile con quella di una vera Λ_c^+ .

Consideriamo il canale di decadimento 3.2. Definiamo particella bachelor il protone potenzialmente prodotto e particella V^0 la K_S^0 potenzialmente prodotta. Le variabili che selezioniamo come input per l'allenamento del modello sono:

- massK0S: la massa invariante della particella V^0 . Si ottiene dalle tracce ricostruite delle figlie. Ci si aspetta che sia compatibile con quella di K_S^0 , quindi circa 497 MeV/ c^2 .
- tImpParBach: parametro d'impatto della particella di bachelor, ovvero la distanza minima tra la traccia della particella e il vertice primario.
- tImpParV0: parametro d'impatto della particella V^0 .
- ctK0S: $c\tau$ della particella V^0 . Ci aspettiamo, assumendo che V^0 sia la particella K_S^0 , circa 2.68 cm.
- cosPAK0S: coseno dell'angolo tra la direzione della particella V^0 e la congiungente dei vertici primario e secondario. Il valore atteso è prossimo all'unità.
- nSigmapr: probabilità, in unità di deviazioni standard, che la particella bachelor sia un protone. Si ottiene dalla somma in quadratura della probabilità fornita da TOF con quella fornita da TPC. Nel caso in cui l'informazione di TOF sia assente, è data dalla probabilità di TPC.
- dcaV0: Distance of Closest Approach, distanza minima tra le tracce ricostruite delle particelle figlie della particella V0.

In Fig. 3.2 è riportata la distribuzione dei valori delle variabili tra segnale e background.

In generale, per l'allenamento di una Rete Neurale è poco conveniente usare variabili altamente correlate tra loro. Infatti, potrebbe essere compromesso l'apprendimento o aumentare significativamente il tempo necessario. Da Fig. 3.3 osserviamo che le variabili scelte non presentano correlazioni elevate tra loro, sia per i dati del segnale che per quelli del fondo.

3.4. Dati e variabili fisiche

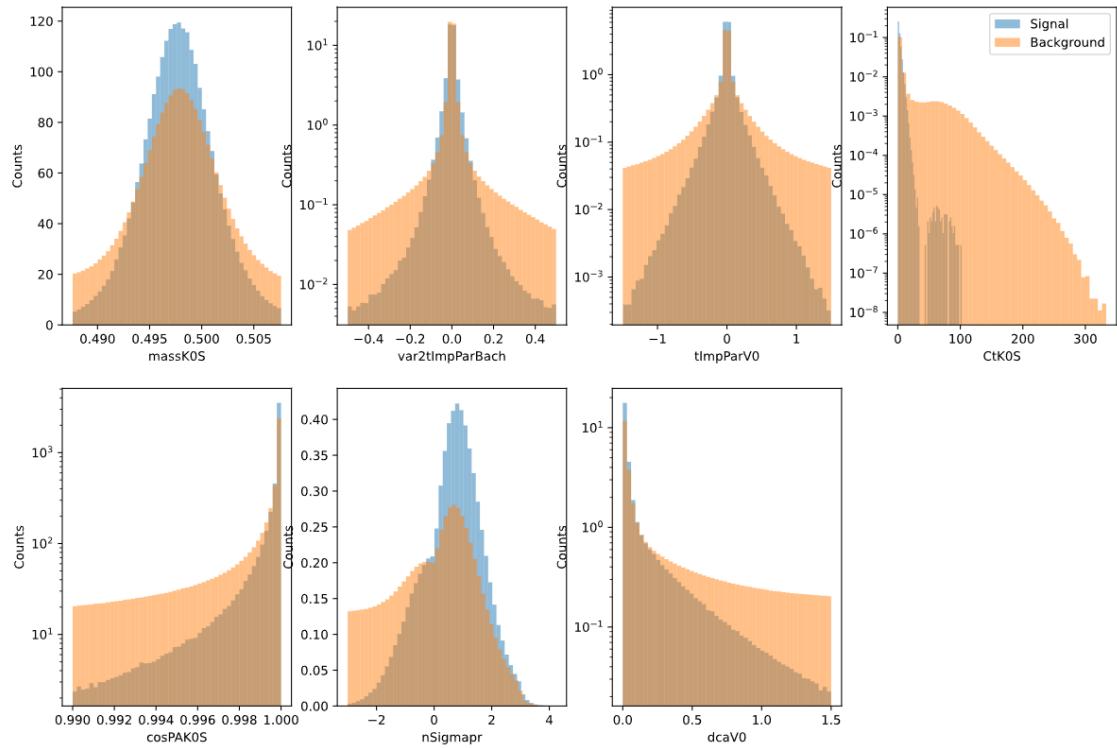


Figura 3.2: Variabili selezionate per il training della Rete Neurale Grafica. Negli istogrammi è rappresentata la differenza tra il segnale e il background.

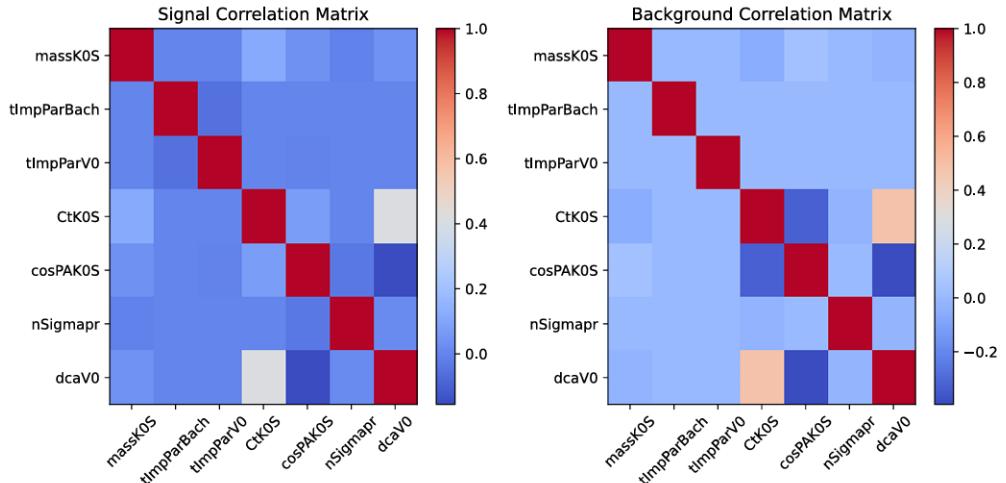


Figura 3.3: Matrici di correlazione per il segnale (a sinistra) e per il fondo (a destra). Si osserva una scarsa correlazione tra le variabili.

3.5 Definizione del modello

Il codice completo in Python impiegato in questa tesi è contenuto nella repository <https://github.com/sofialeotti/Tesi.git>.

Il campione di dati complessivo, formato sia dal segnale che dal fondo, è stato diviso in tre parti: il training set, il validation set e il testing set. I primi due sono necessari per l'allenamento del modello, l'ultimo per la valutazione. Questi tre insiemi devono avere lo stesso numero di dati, poiché l'allenamento e la valutazione di una GNN devono avvenire su grafi con lo stesso numero di nodi. Il campione è stato diviso nei tre insiemi applicando due volte la funzione `train_test_split` di scikit-learn. Successivamente è presente un controllo per verificare che abbiano tutti la stessa dimensione. Se non è così vengono rimossi gli ultimi elementi dell'insieme più numeroso. Ognuno dei tre insiemi viene usato per definire un grafo, in cui ogni nodo è in relazione solo con se stesso: la matrice di adiacenza è diagonale. Inoltre, gli *edges* non sono dotati di proprietà rilevanti, quindi non è necessario definire `e`. Infine, ogni dato è caratterizzato dalle 7 variabili definite in Sez. 3.4, dunque il tensore `x` ha forma $N_{nodes} \times 7$. Il codice che realizza la preparazione dei grafi è contenuto nel file `DataPreparation.py`.

La definizione del modello e il codice necessario per l'allenamento e la valutazione si trovano nel file `Classifier.py`, dove sono definite le classi `GNN` e `GraphNeuralNetwork`. Nella prima, una classe che eredita da `Model` di Keras, è definita la Rete Neurale Grafica formata da tre layer convoluzionali `GCNConv()` [39]. Nella seconda è contenuto il codice per l'inizializzazione, l'allenamento e la valutazione della GNN.

Inoltre, si vuole evitare l'overfitting, ovvero la situazione in cui un modello di Machine Learning impara pattern specifici dei dati su cui è stato allenato, ma irrilevanti o fuorvianti su altri insiemi di dati, peggiorando le sue prestazioni su altri campioni di dati. Perciò è stata aggiunta una funzione di dropout nella definizione della rete. Questa tecnica consiste nello spegnimento casuale di alcuni neuroni di un layer della rete neurale durante l'allenamento [35]. Il layer `Dropout` di Keras [40], in particolare, mette a zero alcuni elementi del tensore in input al layer, ciascuno con la probabilità passata in argomento, nel nostro caso 0.5. La scelta degli elementi che vengono posti a zero è indipendente tra i layer.

La definizione della classe `GNN`:

```
1 class GNN(Model):
2     def __init__(self,
3                  dataset,
4                  hidden_channels = 20,
5                  hidden_channels_1 = 16,
6                  output_channels = 1):
7         super().__init__()
8         self.mask = np.ones((dataset.n_nodes), dtype=float)
```

3.5. Definizione del modello

```
9     self.gcn0 = GCNConv(channels=hidden_channels, activation=
10    'relu')
11    self.dropout = Dropout(0.5)
12    self.gcn1 = GCNConv(channels=hidden_channels_1,
13    activation= 'relu')
14    self.gcn2 = GCNConv(channels=output_channels, activation=
15    'sigmoid')
16
17    def call(self,
18        inputs):
19        x, a = inputs
20        #layer 1
21        x = self.gcn0([x, a], mask=self.mask)
22        #layer 2
23        x = self.dropout(x)
24        x = self.gcn1([x, a], mask=self.mask)
25        #layer 3
26        x = self.dropout(x)
27        x = self.gcn2([x, a], mask=self.mask)
28
29    return x
```

Definendo una classe che eredita da `Model` è stato necessario ridefinire il metodo `call()`, in cui si specificano le operazioni che vengono eseguite sui dati in input ad ogni chiamata del modello.

Le migliori prestazioni del modello sono state ottenute con 4200 epoche e un learning rate di 0.01. In Fig. 3.4 è rappresentato l'andamento della loss function nel corso delle epoche. Osserviamo che il valore decresce in modo rapido nelle prime 1000 epoche, per poi stabilizzarsi intorno a 0.4. Il programma calcola anche il tempo di allenamento della rete, in questo caso 1820.82 s.

3.5.1 Ranking delle variabili

Un modo per comprendere cosa porta una GNN a prendere una decisione è valutare l'importanza che la rete dà ai *features* dei nodi, cioè alle diverse variabili. In Fig. 3.5 è presente tale ranking in un istogramma.

Osserviamo che la variabile più rilevante per la rete è `ctK0S`. Hanno circa metà della sua rilevanza `tImpParBach` e `tImpParV0`. Infine, la variabile meno rilevante è `nSigmaPr`.

3.5. Definizione del modello

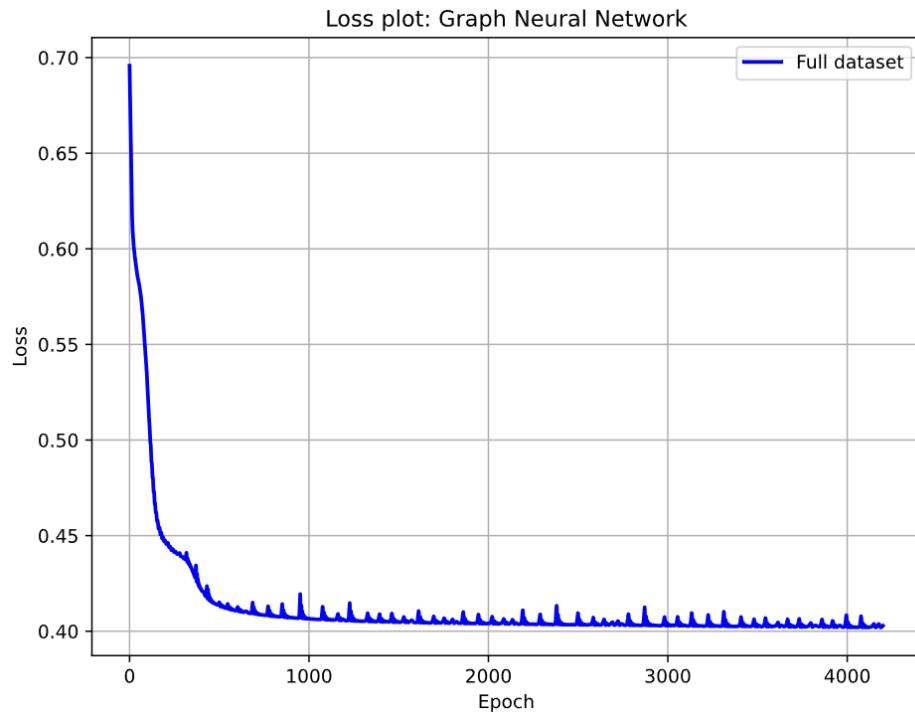


Figura 3.4: *Andamento della loss function durante il training nel corso delle epoche.*

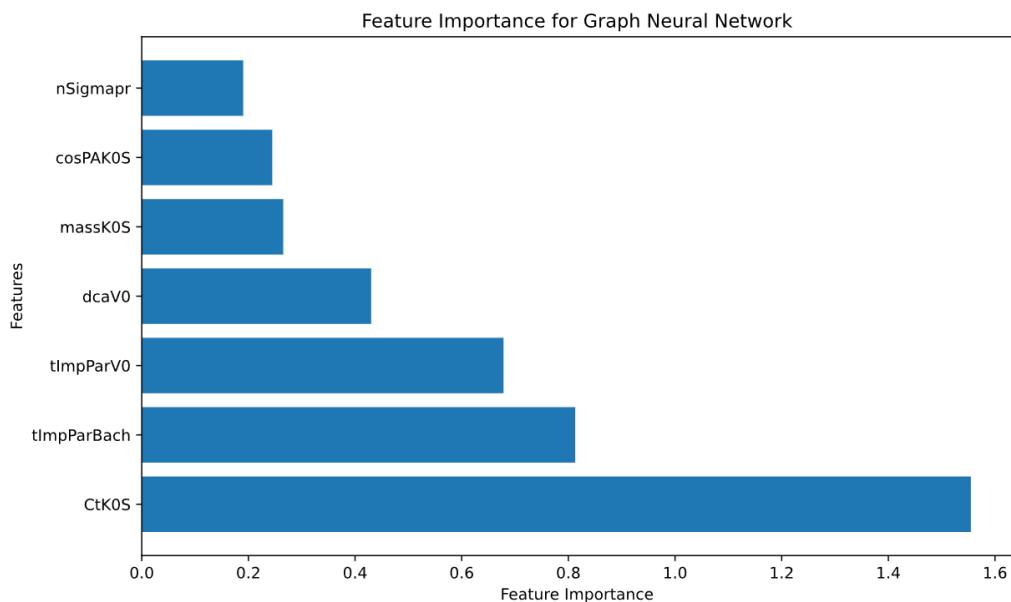


Figura 3.5: *Ranking dell'importanza delle variabili per la Rete Neurale Grafica.*

3.6 Valutazione

Per valutare la qualità dell’allenamento si fa agire la rete su un set di dati diverso da quello di allenamento e si confrontano le previsioni del modello con la reale identità dei dati. Attraverso il metodo `predict()` di Keras si ottengono tali previsioni sui dati selezionati per il testing. Esistono diversi metodi che possono essere impiegati per la valutazione.

3.6.1 Metriche

Definendo i dati del segnale come appartenenti alla classe 1 e i dati del fondo come appartenenti alla classe 0, definiamo:

- TP (True Positive): il numero di dati del segnale correttamente classificati;
- TN (True Negative): il numero di dati del background correttamente classificati;
- FP (False Positive): il numero di dati del background classificati come dati del segnale;
- FN (False Negative): il numero di dati del segnale classificati come dati del background;

L’accuratezza (*accuracy*) è definita come il rapporto tra il numero di classificazioni corrette e il numero di classificazioni totali:

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.5)$$

È una metrica utile nel caso di campioni di dati di testing bilanciati tra le due classi. In Fig. 3.6 si può vedere l’andamento dell’accuratezza nel corso delle epoche durante il training. Si osserva una crescita veloce nelle prime 1000 epoche, successivamente non ha importanti variazioni.

La precisione (*precision*) è definita come il numero di dati correttamente classificati come elementi del segnale rispetto al numero totale di dati classificati come segnale:

$$\text{precision} = \frac{TP}{FP + TP} \quad (3.6)$$

è un indice dell’affidabilità di una previsione nella classe 1.

Il recall, o sensibilità, è la frazione di dati appartenenti alla classe 1 correttamente classificati:

$$\text{recall} = \frac{TP}{TP + FN} \quad (3.7)$$

3.6. Valutazione

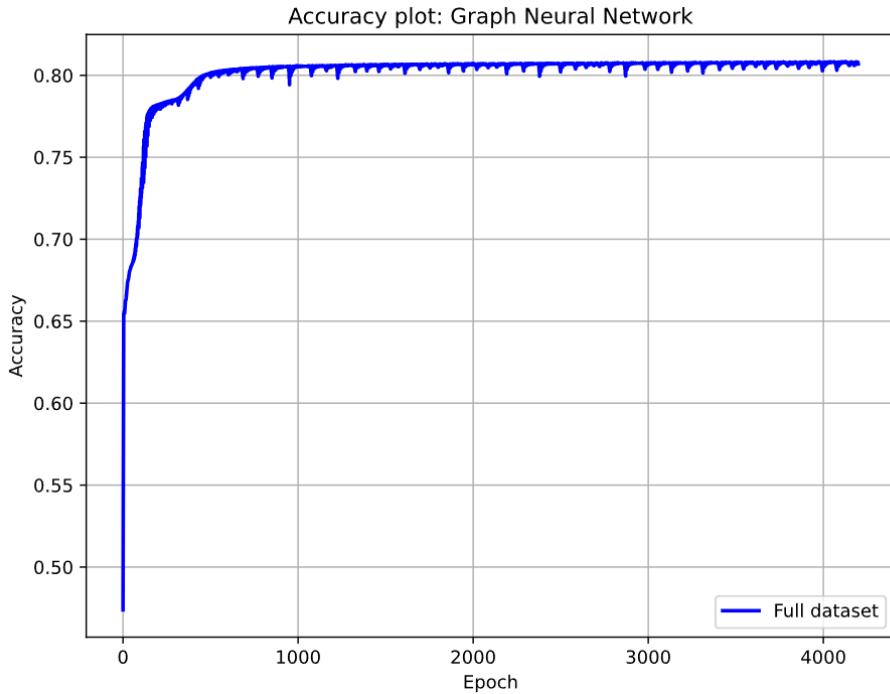


Figura 3.6: *Andamento dell'accuratezza del modello nel corso delle epoche di training.*

ovvero il True Positive Rate (TPR). È una metrica molto rilevante nei casi in cui l'obiettivo è identificare tutti i positivi.

Il fattore F1 è la media armonica di precisione e recall

$$F1 = 2 * \frac{precision \cdot recall}{precision + recall} \quad (3.8)$$

Idealmente, il modello deve identificare correttamente tutte le istanze, e quindi $FP = 0$, $FN = 0$, per cui il valore ideale di ciascuna di queste metriche è 1. I valori delle metriche ottenuti con il modello definito in Sez. 3.5 sono in Tab. 3.2.

<i>accuracy</i>	0.8060623080360375
<i>precision</i>	0.8212619417907132
F1	0.8015757564805696

Tabella 3.2: *Valori delle metriche del modello di Rete Neurale Grafica composto da tre livelli convoluzionali, allenato su 4200 epoche.*

3.6. Valutazione

3.6.2 Confusion Matrix

La confusion matrix, o matrice di confusione, è un metodo per la rappresentazione grafica dei risultati di un algoritmo di classificazione. Le colonne della matrice indicano i valori previsti, le righe i valori reali. Per un classificatore binario è una matrice 2×2 . Idealmente, gli elementi diagonali dovrebbero essere gli unici popolati, poiché i loro ingressi sono TP (1,1) e TN (0,0). A partire dalla confusion matrix si possono ottenere tutte le metriche definite in Sez. 3.6.1.

In Fig. 3.7 la matrice di confusione ottenuta con il modello definito in Sez. 3.5. Dal numero di ingressi in ciascun quadrante è possibile calcolare il recall: 0.782811259.

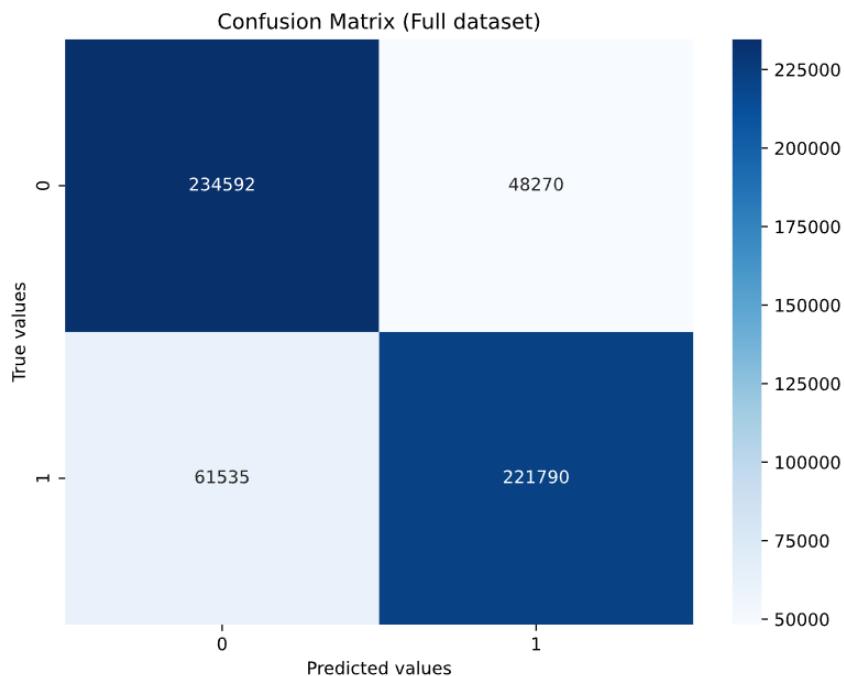


Figura 3.7: *Confusion Matrix del modello.*

3.6.3 Curva di ROC

La Receiver Operating Characteristic (ROC) mette in relazione l'efficienza del segnale, ovvero il recall, con l'eliminazione del fondo (background rejection), definita come:

$$\text{background rejection} = \frac{TN}{TN + FP} \quad (3.9)$$

3.6. Valutazione

Idealmente, entrambi questi parametri dovrebbero essere 1: tutti gli elementi del fondo vengono rigettati e tutti quelli del segnale vengono identificati. La curva ideale, quindi, si avvicina il più possibile al vertice in alto a destra.

Della curva di ROC è rilevante, oltre alla forma, l'area sottesa (Area Under the Curve, AUC): è la probabilità che il modello classifichi un'istanza positiva scelta casualmente più in alto di una negativa scelta casualmente. Se $AUC = 0.5$ il modello è un classificatore casuale, cioè non ha alcuna capacità di discriminare tra le due classi. Questo corrisponde ad una curva di ROC data da una retta che forma un angolo di 135° con l'asse TPR. Se $0 < AUC < 0.5$ il modello è un classificatore peggiore di un classificatore casuale: valori in questo range solitamente indicano problemi di costruzione del modello o dei dati, che portano a delle previsioni invertite. Questo caso corrisponde a curve di ROC che si trovano al di sotto della curva di classificatore casuale. Se $AUC > 0.5$ il modello è in grado di distinguere tra le due classi. Questi valori di AUC corrispondono a curve di ROC al di sopra della curva di classificatore casuale. Più il valore di AUC è vicino a 1, il valore ideale, migliore è il classificatore.

Per il modello costruito in questa tesi il valore di AUC è 0.896263082668049, la curva è presente in Fig. 3.8.

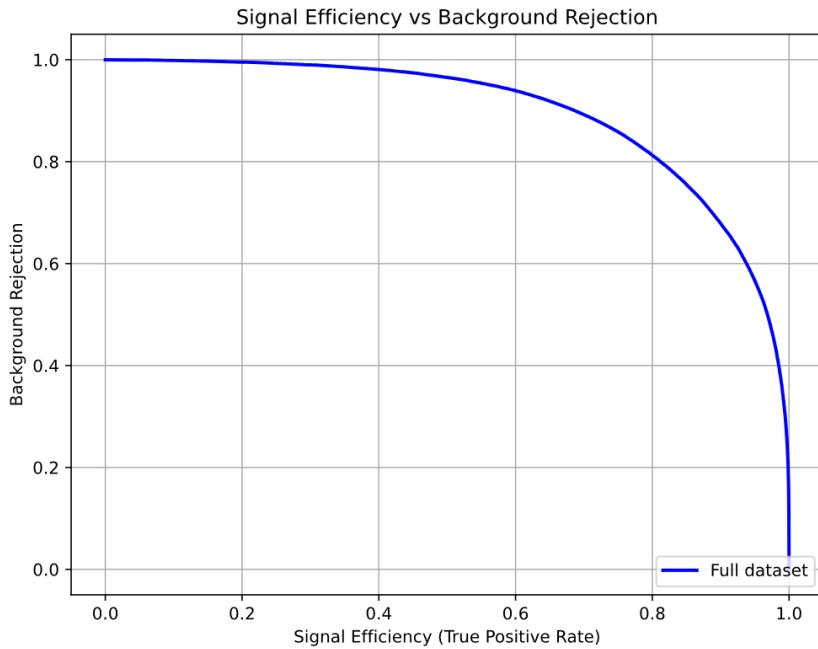


Figura 3.8: *Curva di ROC della Rete Neurale Grafica.*

Conclusioni

Nel presente elaborato è stato realizzato un programma in Python per l’allenamento e la valutazione di una Rete Neurale Grafica in grado di ricostruire il barione Λ_c^+ . Sono state impiegate le funzioni di definizione, inizializzazione, allenamento e valutazione di modelli di deep learning della API Keras. Questo tipo di Rete Neurale lavora su grafi, perciò è stato necessario utilizzare la libreria open source Spektral per adattare questo tipo di dato alle funzioni di Keras. Questo è stato necessario per la suddivisione dei set di dati di allenamento, validazione e test nei *batch*. A questo scopo sono stati usati i **Loaders**. Inoltre, per la rappresentazione dei grafi sono stati usate le cassi **Graph** e **Dataset** di Spektral. Per ottenere la convergenza del modello, il grafo è stato definito come segue: ogni elemento del campione di dati è un nodo del grafo, i valori delle variabili sono i suoi *features* e ciascun nodo è in relazione solamente con se stesso.

Il programma realizzato fornisce informazioni sull’andamento dell’allenamento tramite i grafici dell’accuratezza e della loss function nel corso delle epoche. Per quanto riguarda la valutazione, vengono fornite la confusion matrix, la curva di ROC e la corrispondente AUC e alcune metriche: accuratezza, precisione e fattore F1. A partire dalla confusion matrix è possibile ottenere la sensibilità, o True Positive Rate, del modello. Con un allenamento su 4200 epoche sono stati ottenuti i seguenti parametri:

accuratezza	0.8060623080360375
precisione	0.8212619417907132
F1	0.8015757564805696
recall	0.782811259
AUC	0.896263082668049

Osserviamo che i parametri non sono ideali ma sono prossimi all’unità: il modello è in grado di distinguere bene le due classi. In particolare osserviamo il valore di recall: il modello è in grado di identificare quasi tutti gli elementi del segnale. Il valore elevato della precisione e dell’accuratezza, inoltre, ci mostrano l’affidabilità delle previsioni.

I risultati ottenuti dimostrano quindi la capacità dell’algoritmo implementato di addestrare in maniera adeguata il modello di Machine Learning con Reti Neurali Grafiche. Tale framework, una volta completato con la parte riguardante l’applicazione del modello su dati reali, potrà fornire un ulteriore strumento potente, flessibile e personalizzabile,

Conclusioni

da utilizzare non solo nelle analisi che richiedono la ricostruzione del barione charmato Λ_c^+ ma in generale per diversi tipi di analisi nel campo della fisica delle alte energie.

Bibliografia

- [1] A. Bettini, *Introduction to Elementary Particle Physics*, 3^a ed. Cambridge University Press, 2024.
- [2] S. Braibant, G. Giacomelli e M. Spurio, *Particelle e Interazioni Fondamentali*. Springer-Verlag Italia, 2012.
- [3] W. Buza, K. Rajagopal e W. vanderSchee, “Heavy Ion Collisions: The Big Picture, and the Big Questions”, *Annual Reviews*, 2018. DOI: <https://doi.org/10.1146/TBD>.
- [4] LHCb collaboration et al., *Centrality determination in heavy-ion collisions with the LHCb detector*, 2021. arXiv: 2111.01607 [nucl-ex]. indirizzo: <https://arxiv.org/abs/2111.01607>.
- [5] R. S. Bhalerao, *Relativistic heavy-ion collisions*, 2014. arXiv: 1404.3294 [nucl-th]. indirizzo: <https://arxiv.org/abs/1404.3294>.
- [6] A. Andronic, “An overview of the experimental study of quark–gluon matter in high-energy nucleus–nucleus collisions”, *International Journal of Modern Physics A*, vol. 29, n. 22, p. 1430047, ago. 2014, ISSN: 1793-656X. DOI: 10.1142/s0217751x14300476. indirizzo: <http://dx.doi.org/10.1142/S0217751X14300476>.
- [7] J. C. Collins, D. E. Soper e G. Sterman, *Factorization of Hard Processes in QCD*, 2004. arXiv: hep-ph/0409313 [hep-ph]. indirizzo: <https://arxiv.org/abs/hep-ph/0409313>.
- [8] R. J. Fries et al., “Hadronization in Heavy-Ion Collisions: Recombination and Fragmentation of Partons”, *Physical Review Letters*, vol. 90, n. 20, mag. 2003, ISSN: 1079-7114. DOI: 10.1103/physrevlett.90.202303. indirizzo: <http://dx.doi.org/10.1103/PhysRevLett.90.202303>.
- [9] A. S. Kalteyer, *Charm production and hadronization in pp and p-Pb collisions at the LHC with ALICE*, 2022. arXiv: 2207.07590 [nucl-ex]. indirizzo: <https://arxiv.org/abs/2207.07590>.

BIBLIOGRAFIA

- [10] P. Skands, S. Carrazza e J. Rojo, “Tuning PYTHIA 8.1: the Monash 2013 tune”, *The European Physical Journal C*, vol. 74, n. 8, ago. 2014, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-014-3024-y. indirizzo: <http://dx.doi.org/10.1140/epjc/s10052-014-3024-y>.
- [11] J. Bellm, S. Gieseke e P. Kirchgaeßer, “Improving the description of multiple interactions in Herwig”, *Eur. Phys. J. C* 80, 469 (2020), 2020. DOI: <https://doi.org/10.1140/epjc/s10052-020-8002-y>.
- [12] J. R. Christiansen e P. Z. Skands, “String formation beyond leading colour”, *Journal of High Energy Physics*, vol. 2015, n. 8, ago. 2015, ISSN: 1029-8479. DOI: 10.1007/jhep08(2015)003. indirizzo: [http://dx.doi.org/10.1007/JHEP08\(2015\)003](http://dx.doi.org/10.1007/JHEP08(2015)003).
- [13] S. Plumari et al., “Charmed hadrons from coalescence plus fragmentation in relativistic nucleus-nucleus collisions at RHIC and LHC”, *The European Physical Journal C*, vol. 78, n. 4, apr. 2018, ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-018-5828-7. indirizzo: <http://dx.doi.org/10.1140/epjc/s10052-018-5828-7>.
- [14] M. He e R. Rapp, “Charm-baryon production in proton-proton collisions”, *Physics Letters B*, vol. 795, pp. 117–121, ago. 2019, ISSN: 0370-2693. DOI: 10.1016/j.physletb.2019.06.004. indirizzo: <http://dx.doi.org/10.1016/j.physletb.2019.06.004>.
- [15] “The Large Hadron Collider.” indirizzo: <https://home.cern/science/accelerators/large-hadron-collider>.
- [16] E. Lopienska, “The CERN accelerator complex, layout in 2022”,, 2022, General Photo. indirizzo: <https://cds.cern.ch/record/2800984>.
- [17] S. Acharya et al., “ALICE upgrades during the LHC Long Shutdown 2”, *Journal of Instrumentation*, vol. 19, n. 05, P05062, mag. 2024, ISSN: 1748-0221. DOI: 10.1088/1748-0221/19/05/p05062. indirizzo: <http://dx.doi.org/10.1088/1748-0221/19/05/P05062>.
- [18] The ALICE Collaboration et al., “The ALICE experiment at the CERN LHC”, *Journal of Instrumentation*, vol. 3, n. 08, S08002, ago. 2008. DOI: 10.1088/1748-0221/3/08/S08002. indirizzo: <https://doi.org/10.1088/1748-0221/3/08/S08002>.
- [19] G. Dellacasa et al., *ALICE time projection chamber* (Technical design report. ALICE). Geneva: CERN, 2000. indirizzo: <https://cds.cern.ch/record/451098>.
- [20] G. Dellacasa et al., “ALICE technical design report of the time-of-flight system (TOF)”,, feb. 2000.
- [21] R. Russo, “Mesurement of D⁺ meson production in p-Pb collisions with the ALICE detector”,, 2015. DOI: 10.48550/arXiv.1511.04380.

BIBLIOGRAFIA

- [22] R. Preghenella, “The Time-Of-Flight detector of ALICE at LHC: construction, test and commissioning with cosmic rays,” tesi di dott., alma, Maggio 2009. indirizzo: <https://amsdottorato.unibo.it/id/eprint/1373/>.
- [23] S. N. Alam, “Charge correlations using balance functions of identified particles in Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV with ALICE”, *PoS*, vol. EPS-HEP2017, p. 151, 2017, The European Physical Society Conference on High Energy Physics (EPS-HEP2017). doi: 10.22323/1.314.0151. arXiv: 1902.09730. indirizzo: <https://cds.cern.ch/record/2665084>.
- [24] F. Piuz et al., *ALICE high-momentum particle identification* (Technical design report. ALICE). Geneva: CERN, 1998. indirizzo: <https://cds.cern.ch/record/381431>.
- [25] G. Dellacasa et al., “ALICE technical design report of the photon spectrometer (PHOS)”, mar. 1999.
- [26] A. Fantoni e (. behalf ofthe ALICE collaboration), “The ALICE Electromagnetic Calorimeter: EMCAL”, *Journal of Physics: Conference Series*, vol. 293, n. 1, p. 012043, apr. 2011. doi: 10.1088/1742-6596/293/1/012043. indirizzo: <https://doi.org/10.1088/1742-6596/293/1/012043>.
- [27] “ALICE EMCAL.” indirizzo: https://alice-collaboration.web.cern.ch/menu_proj_items/emcal.
- [28] “ALICE Muon Spectrometer.” indirizzo: https://alice-collaboration.web.cern.ch/menu_proj_items/Muon-Spect.
- [29] The ALICE collaboration, “Performance of the ALICE VZERO system”, *Journal of Instrumentation*, vol. 8, n. 10, P10016–P10016, ott. 2013, ISSN: 1748-0221. doi: 10.1088/1748-0221/8/10/p10016. indirizzo: <http://dx.doi.org/10.1088/1748-0221/8/10/P10016>.
- [30] C. Amsler et al (Particle Data Group), *Particle listings Λ_c* , 2008. indirizzo: <https://pdg.lbl.gov/2008/listings/s033.pdf>.
- [31] P.A. Zyla et al. (Particle Data Group), *Prog. Theor. Exp. Phys.* 2020, 083C01 (2020). indirizzo: <https://pdg.lbl.gov/2020/tables/rpp2020-tab-baryons-Charm.pdf>.
- [32] P. W. Battaglia et al., *Interaction Networks for Learning about Objects, Relations and Physics*, 2016. arXiv: 1612.00222 [cs.AI]. indirizzo: <https://arxiv.org/abs/1612.00222>.
- [33] B. Sanchez-Lengeling et al., “A Gentle Introduction to Graph Neural Networks”, *Distill*, 2021, <https://distill.pub/2021/gnn-intro>. doi: 10.23915/distill.00033.

BIBLIOGRAFIA

- [34] N. K. Manaswi, *Deep Learning with Applications Using Python: Chatbots and Face, Object, and Speech Recognition With TensorFlow and Keras*, 1st. USA: Apress, 2018, ISBN: 1484235150.
- [35] F. Chollet, *Deep Learning with Python*, 1st. USA: Manning Publications Co., 2017, ISBN: 1617294438.
- [36] D. Grattarola e C. Alippi, *Graph Neural Networks in TensorFlow and Keras with Spektral*, 2020. arXiv: 2006.12138 [cs.LG]. indirizzo: <https://arxiv.org/abs/2006.12138>.
- [37] R. Ciesielski e K. Goulianos, *MBR Monte Carlo Simulation in PYTHIA8*, 2012. arXiv: 1205.1446 [hep-ph]. indirizzo: <https://arxiv.org/abs/1205.1446>.
- [38] R. Brun et al., *GEANT 3*. Geneva: CERN, 1987. indirizzo: <https://cds.cern.ch/record/1119728>.
- [39] T. N. Kipf e M. Welling, *Semi-Supervised Classification with Graph Convolutional Networks*, 2017. arXiv: 1609.02907 [cs.LG]. indirizzo: <https://arxiv.org/abs/1609.02907>.
- [40] “Dropout layer.” indirizzo: https://keras.io/api/layers/regularization_layers/dropout/.