

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

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

Studio di tecniche di machine learning per
la ricostruzione del barione charmato
 Λ_c in ALICE ad LHC

Relatore:
Dott. Andrea Alici

Presentata da:
Sofia Strazzi

Anno Accademico 2018/2019

Indice

| | |
|--|-----------|
| Indice | 1 |
| Introduzione | 3 |
| 1 La fisica di ALICE | 4 |
| 1.1 Introduzione al Modello Standard | 4 |
| 1.2 Cromodinamica Quantistica | 5 |
| 1.2.1 QCD su reticolo | 7 |
| 1.3 Quark-Gluon Plasma (QGP) | 7 |
| 1.3.1 Dalla materia adronica al QGP | 7 |
| 1.4 Collisioni tra ioni pesanti | 9 |
| 1.4.1 Evoluzione del QGP | 10 |
| 1.5 Evidenze sperimentali del QGP | 11 |
| 1.5.1 Soft probes | 11 |
| 1.5.2 Hard probes | 14 |
| 1.5.3 Prove elettromagnetiche: fotoni diretti | 19 |
| 2 Il rivelatore ALICE a LHC | 21 |
| 2.1 Large Hadron Collider (LHC) | 21 |
| 2.1.1 Il sistema di magneti | 21 |
| 2.2 L'esperimento ALICE | 23 |
| 2.2.1 Inner Tracking System | 25 |
| 2.2.2 Time Projection Chamber | 25 |
| 2.2.3 Transition Radiation Detector | 26 |
| 2.2.4 High-Momentum Particle Identification Detector | 27 |
| 2.2.5 Calorimetri e altri rivelatori | 28 |
| 2.3 Il sistema a tempo di volo (TOF) | 29 |
| 2.3.1 Multigap Resistive Plate Chamber | 30 |
| 2.3.2 Identificazione di particelle con TOF | 32 |

| | | |
|----------|--|-----------|
| 3 | Studio di tecniche di machine learning per la ricostruzione di Λ_c | 34 |
| 3.1 | Introduzione | 34 |
| 3.2 | TMVA | 35 |
| 3.3 | I tre metodi a confronto | 35 |
| | 3.3.1 Tagli rettangolari | 35 |
| | 3.3.2 Neural Networks di tipo Multilayer Perceptron (MLP) | 35 |
| | 3.3.3 Boosted Decision Tree (BDT) | 36 |
| 3.4 | Training del TMVA | 36 |
| | 3.4.1 Simulazione Monte Carlo | 37 |
| | 3.4.2 Variabili di input | 37 |
| 3.5 | Configurazione dei tre metodi | 42 |
| | 3.5.1 Tagli rettangolari | 43 |
| | 3.5.2 Reti neurali | 44 |
| | 3.5.3 Boosted Decision Tree | 45 |
| 3.6 | Ranking delle variabili | 45 |
| 3.7 | Plot di controllo | 47 |
| 3.8 | Confronto tra i metodi utilizzati | 56 |
| | Conclusione | 59 |
| | Bibliografia | 60 |

Introduzione

L'esperimento ALICE (A Large Ion Collider Experiment), situato al Large Hadron Collider (LHC) del CERN, è dedicato allo studio delle collisioni tra ioni pesanti ultra-relativistici, in particolare di ioni di piombo, fino ad energie di 5.5 TeV per coppia di nucleoni nel centro di massa. L'obiettivo principale dell'esperimento è indagare lo stato della materia che si forma in condizioni di alta temperatura e densità di energia, noto come Quark-Gluon Plasma (QGP). Il QGP risulta essere difficilmente osservabile e avendo una vita media estremamente breve, le prove della sua esistenza sono tutte dovute a misure indirette. Uno degli strumenti fondamentali per investigare le sue proprietà è lo studio dei quark pesanti. Essi infatti, grazie alle loro masse elevate, vengono prodotti nelle primissime fasi della collisione e, propagandosi all'interno del mezzo, interagiscono con i suoi costituenti durante tutta l'evoluzione del sistema, fornendo misure sulle sue proprietà. In particolare, in questa tesi è stato preso in considerazione il barione charmato Λ_c^+ nel suo canale di decadimento $\Lambda_c^+ \rightarrow p + K_s^0$. A causa della breve vita media e del basso rapporto tra segnale e fondo, la ricostruzione di tale particella risulta complicata e si rende necessario l'uso di tecniche di analisi multivariata, che permettono di considerare contemporaneamente e indipendentemente più variabili per evento, e di sfruttare tutte le informazioni disponibili attraverso metodi di machine learning. Nella presente tesi sono state analizzate e confrontate tre diverse tecniche per l'estrazione del segnale: i tagli rettangolari, le reti neurali e i Boosted Decision Tree (BDT). Viene in particolare riportata una descrizione dei tre metodi, compresi i principi di funzionamento e i parametri di configurazione adottati, e un confronto tra le loro performance.

Capitolo 1

La fisica di ALICE

1.1 Introduzione al Modello Standard

Il Modello Standard (MS) [1] è la teoria che descrive le particelle elementari ad oggi conosciute e tre delle quattro interazioni fondamentali: debole, elettromagnetica e forte. Esso si basa sulle teorie quantistiche dei campi, in grado di conciliare i principi della meccanica quantistica e della relatività speciale: ad ogni particella corrisponde un campo che può assumere stati quantizzati di energia; le particelle sono i quanti dei rispettivi campi quando essi si trovano in uno stato diverso da quello fondamentale. In particolare, come si osserva in figura 1.1, il Modello Standard comprende 12 fermioni (particelle con spin semintero che obbediscono alla statistica di Fermi-Dirac), suddivisi in 6 quark e 6 leptoni (e le relative antiparticelle), che rappresentano i campi materiali, 4 bosoni (particelle con spin intero che obbediscono alla statistica di Bose-Einstein), mediatori delle interazioni e il bosone di Higgs con spin zero, necessario per il processo di rottura spontanea della simmetria (Spontaneous Symmetry Breaking, SSB) attraverso cui si dà massa ai bosoni W^+ , W^- , Z^0 e ai fermioni. I bosoni Z^0 e W^\pm sono i mediatori delle interazioni deboli, il fotone γ è il mediatore delle interazioni elettromagnetiche e i gluoni quelli delle interazioni forti.

Le interazioni sono descritte utilizzando le teorie di gauge, basate sull'idea che alcune simmetrie siano possibili non solo globalmente, ma anche localmente: l'interazione forte è descritta dal gruppo $SU(3)$, che caratterizza la carica di colore, quella elettromagnetica dal gruppo $U(1)$ e l'interazione debole dal gruppo $SU(2)_s$ e agisce sulle particelle sinistre. Se quest'ultima simmetria fosse esatta, la teoria non sarebbe rinormalizzabile e i bosoni non dovrebbero essere massivi, come emerge invece dalle evidenze sperimentali. Risulta quindi necessario introdurre il meccanismo di Higgs, secondo il quale i bosoni e i fermioni acquisiscono una massa mediante l'interazione con il campo scalare di Higgs. Le evidenze sperimentali confermano la validità del Modello Standard; tale teoria, però, non si può ancora considerare completa per la mancanza di una spiegazione dell'interazione

Standard Model of Elementary Particles

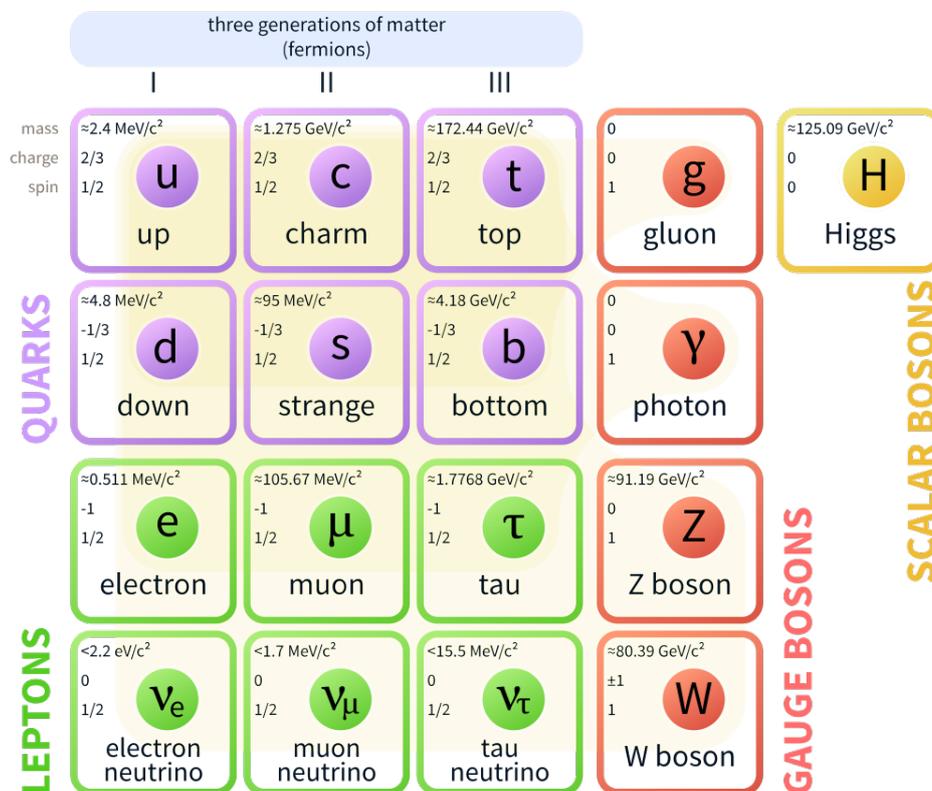


Figura 1.1: Modello Standard delle particelle elementari: quark e leptoni, costituenti della materia ordinaria, bosoni di gauge, mediatori delle interazioni e bosone di Higgs.

gravitazionale e per l'incompatibilità con i principi della relatività generale.

1.2 Cromodinamica Quantistica

La Cromodinamica Quantistica (QCD) è la teoria quantistica di campo delle interazioni forti che si fonda sull'assunzione che ogni sapore di quark esista in tre diversi colori, o cariche di gauge, convenzionalmente denominate rosso, blu e verde, ed è una teoria di gauge non abeliana, basata sul gruppo di simmetria $SU(3)$, caratterizzata da 8 campi di gauge, detti campi gluonici. I generatori infinitesimi dell'algebra di $SU(3)$ sono le matrici di Gell-Mann nella rappresentazione fondamentale. Questi 8 campi sono portatori essi stessi di carica di colore [1]. Ad alte energie o a piccole distanze, l'intensità dell'interazione forte diminuisce molto velocemente (libertà asintotica). Questo significa che nei

processi di scattering ad alta energia (Deep Inelastic Scattering, DIS) i quark si comportano come particelle libere o debolmente legate. Al contrario, a basse energie o a grandi distanze, l'intensità dell'interazione aumenta enormemente, provocando l'impossibilità di osservare quark liberi in natura (confinamento). Per comprendere questo comportamento è necessario analizzare l'andamento della costante di accoppiamento forte α_s , definita come la grandezza adimensionale che determina l'intensità dell'interazione forte ad una data scala energetica, in funzione dell'energia (figura 1.2).

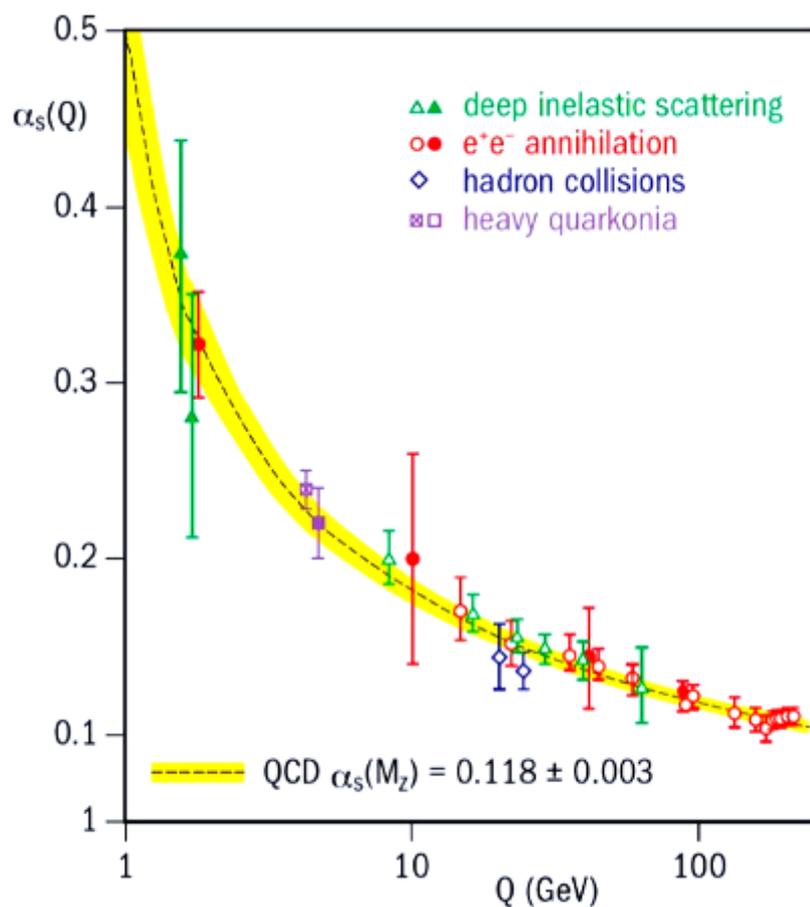


Figura 1.2: Andamento della costante di accoppiamento α_s in funzione del momento scambiato Q . I simboli vuoti e pieni si riferiscono rispettivamente a misure basate su calcoli al NLO (Next-To-Leading-Order) QCD che corrispondono a correzioni di ordine $O(\alpha)$ e di NNLO (Next-to-Next-to-Leading-Order) QCD, corrispondenti a correzioni di ordine $O(\alpha^2)$.

Essa vale, al primo ordine di approssimazione,

$$\alpha_s(Q) = \frac{12\pi}{(33 - 2N_f)\ln(Q^2/\Lambda_{QCD}^2)}$$

dove Q^2 è il quadri-impulso relativistico [2], che definisce la scala di energia, N_f rappresenta il numero di sapori di quark e Λ_{QCD} può essere interpretata come la scala di energia per cui l'interazione forte diventa importante. Nel regime asintotico la bassa intensità delle interazioni fa sì che $\alpha_s \ll 1$, rendendo quindi possibili calcoli perturbativi; mentre in regime di confinamento l'intensità delle interazioni rende il calcolo perturbativo impossibile e si rende necessario l'uso della tecnica del QCD su reticolo.

1.2.1 QCD su reticolo

Per $\alpha_s \sim 1$, l'interazione quark-quark è così intensa da non rendere possibile la separazione dei due fermioni. In queste condizioni si rende necessario quindi utilizzare la procedura di QCD su reticolo [3] che consiste in una discretizzazione delle coordinate dello spazio-tempo, mediante la costruzione di un reticolo di lato L e di passo a (cut off), che definisce il limite inferiore della distanza possibile che può essere studiata, impedendo che possano esserci integrali che divergono. Facendo il limite per $a \rightarrow 0$ e $L \rightarrow \infty$ è possibile tornare al campo continuo. La QCD su reticolo è una delle tecniche principali per lo studio della transizione di fase della materia ordinaria ad un plasma di quark e gluoni.

1.3 Quark-Gluon Plasma (QGP)

Il concetto di libertà asintotica precedentemente introdotto risulta centrale per la comprensione dello stato fisico che si ottiene in condizioni di alte temperature e alta densità di energia, chiamato Quark-Gluon Plasma (QGP).

Il QGP è un particolare sistema fisico presente nella prima fase evolutiva dell'universo, tra 10^{-12} e 10^{-6} secondi dopo il Big Bang, caratterizzato dal deconfinamento di quark e gluoni, liberi di muoversi all'interno del plasma e all'acquisizione da parte dei costituenti di una carica di colore non nulla, proprietà opposte a quelle della materia adronica alle basse temperature. È ragionevole denotare questi due stati come fasi diverse dello stesso sistema fisico (figura 1.3), il cui passaggio avviene in corrispondenza di una determinata temperatura T , detta temperatura critica T_c .

1.3.1 Dalla materia adronica al QGP

Nel passaggio dalla materia adronica al QGP sono due i parametri da tenere in considerazione: la temperatura e il potenziale bariochimico. Quest'ultimo è importante in quanto il sistema che si viene a formare risulta essere gran canonico, cioè scambia energia e particelle con l'ambiente esterno. Esso è definito come $\mu_b = \frac{\partial E}{\partial N_b}$, con N_b numero di

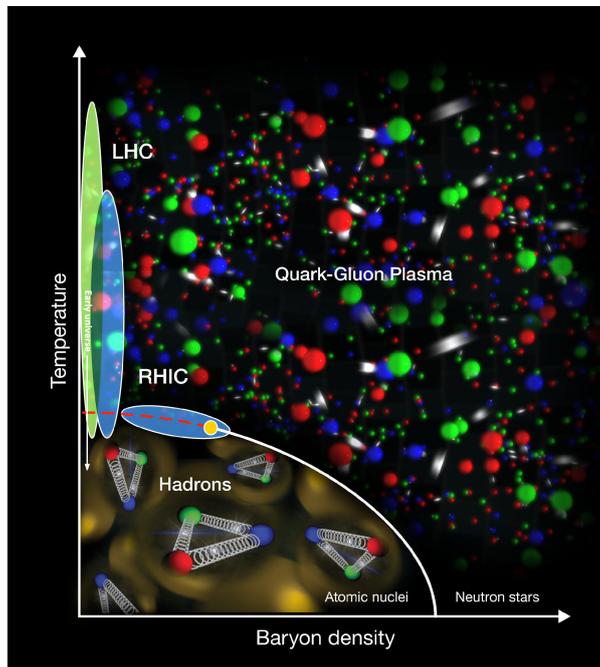


Figura 1.3: Grafico della transizione di fase per il QCD, che riporta la temperatura in funzione della densità barionica [4].

barioni, che riguarda le particelle dotate di carica barionica. Il numero barionico infatti si conserva nella maggior parte delle interazioni del Modello Standard, per cui è facile intuire che μ_b stima la differenza tra la materia e l'antimateria, le quali si trovano in equilibrio solo nel caso in cui μ_b sia uguale a zero; come si può osservare in figura 1.3, se μ_b aumenta, la temperatura critica diminuisce. Negli esperimenti ad alta energia riprodotti in laboratorio, come anche nei primi istanti dell'universo [5], si ha una situazione in cui particelle e antiparticelle possono essere create e annichilate e le loro densità sono molto più elevate della loro differenza; in queste condizioni $\mu_b \rightarrow 0$ e la temperatura critica vale circa 173 MeV. Questo valore si ricava eguagliando la pressione del gas di materia adronica con quella del QGP. Se consideriamo, ad esempio, un gas ideale di pioni senza massa, la pressione è data dalla legge di Stefan-Boltzmann:

$$P_\pi = \frac{\pi^2}{30} T^4$$

a cui va aggiunto il termine $B \simeq 0.5 \text{ GeV}/\text{fm}^3$ di "bag pressure", ottenuto da calcoli su reticolo, derivante dall'energia dello stato di vuoto non perturbativo della QCD, nel modello bag, in cui i quark sono immaginati come in un vuoto perturbativo confinati in una "sacca", l'adrone, immersa nel vuoto non perturbativo della QCD.

Bisogna quindi considerare il contributo energetico necessario a creare nel vuoto fisico non perturbativo, una cavità di vuoto perturbativo; ciò corrisponde ad esercitare

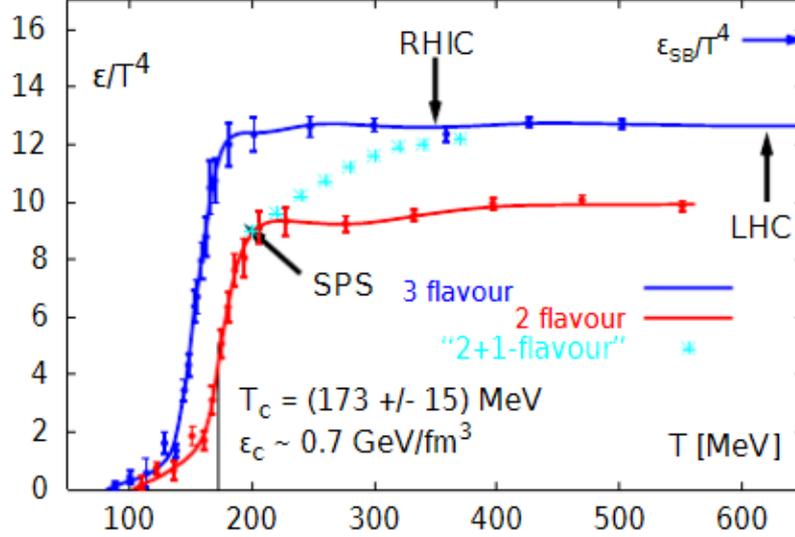


Figura 1.4: Densità di energia ϵ in unità di T^4 in funzione della temperatura T per $\mu_b \rightarrow 0$, ottenuta mediante la tecnica di QCD su reticolo. La linea rossa si riferisce a calcoli effettuati per due quark leggeri, la linea blu corrisponde al calcolo per tre quark leggeri e quella azzurra a due quark leggeri e uno pesante. La transizione avviene ad una temperatura critica pari a $\simeq 173$ MeV.

una pressione del vuoto non perturbativo sul campo di colore confinando le particelle all'interno della cavità. Eguagliando quindi questa pressione con quella del QGP non soggetto a confinamento, pari a

$$P_{QGP} = \frac{37\pi^2}{90} T^4,$$

si ottiene una temperatura critica [3]

$$T_c = \left(\frac{45B}{17\pi^2} \right)^{1/4} \simeq 173 \text{ MeV}$$

come si osserva nel grafico di figura 1.4 per $\mu_b \rightarrow 0$. Dal grafico si osserva anche che l'andamento della densità energetica è proporzionale alla quarta potenza della temperatura.

1.4 Collisioni tra ioni pesanti

Quando avviene una collisione tra ioni pesanti ultrarelativistici si vengono a creare le condizioni di temperatura ed energia necessarie alla produzione del QGP. Data la velocità ultrarelativistica delle particelle coinvolte nell'urto, tenendo quindi conto della

contrazione di Lorentz nella direzione di propagazione del fascio, si possono trattare i due nuclei che collidono come dischi sottili di raggio $2R_A \simeq 2A^{\frac{1}{2}}$ fm, dove A è il numero di nucleoni.

Maggiore è la centralità della collisione e maggiore è la probabilità che si creino le condizioni adatte alla formazione del QGP. Quando la collisione non è centrale solo una parte dei nucleoni viene coinvolta, mentre gli altri, chiamati ‘spettatori’, continuano il loro moto lungo la direzione del fascio praticamente imperturbati. La descrizione delle traiettorie delle particelle coinvolte può essere fatta in termini di rapidità, definita come [6]

$$y = \frac{1}{2} \ln \frac{E + p_L}{E - p_L}$$

dove E è l’energia e p_L la componente del momento lungo la direzione del fascio. Questa si traduce, nel limite ultrarelativistico, in un’altra quantità chiamata pseudorapidità

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

dove θ è l’angolo tra l’impulso e la direzione del fascio.

1.4.1 Evoluzione del QGP

Il QGP si forma solo se si vengono a creare le condizioni di temperatura e densità di energia critiche. In caso contrario il sistema inizia ad evolversi in maniera idrodinamica, passando attraverso una fase preadronica, nella quale non avviene il deconfinamento [7][8].

Nei primi istanti della collisione si ha quella che viene chiamata fase di pre-equilibrio, in cui si verifica la produzione di particelle con impulso trasverso elevato ($p_T \gg 1$ GeV/c) e l’emissione di fotoni diretti, reali o virtuali (questi ultimi si manifestano in coppie leptone-antileptone e^+e^- , oppure $\mu^+\mu^-$). A differenza di quanto avviene nelle collisioni protone-protone, in quelle tra ioni pesanti i prodotti non possono uscire dalla zona in cui è avvenuto l’urto, continuando a scontrarsi e creando una nuvola di particelle interagenti che, termalizzando con una densità di energia sufficientemente elevata, dà origine al QGP. Il sistema che si viene a creare è in equilibrio termico, ed esercita quindi una pressione sul vuoto circostante iniziando ad espandersi e conseguentemente a raffreddarsi. Quando il sistema torna nuovamente alla densità critica, inizia un processo di adronizzazione dei quark e dei gluoni. La formazione di adroni può avvenire per frammentazione di un partone ad alto impulso o coalescenza, ovvero fusione di adroni a basso impulso. Questi processi continuano fino a che l’energia non diventa troppo bassa da non permettere più collisioni inelastiche, raggiungendo uno stato in cui le abbondanze delle specie chimiche sono fissate, chiamato freeze-out chimico. L’espansione, ora dovuta solo ad

urti elastici, si arresta quando le distanze tra gli adroni diventano maggiori del raggio delle interazioni forti dando origine al freeze-out termico [9]. L'intero processo appena descritto è rappresentato in figura 1.5.

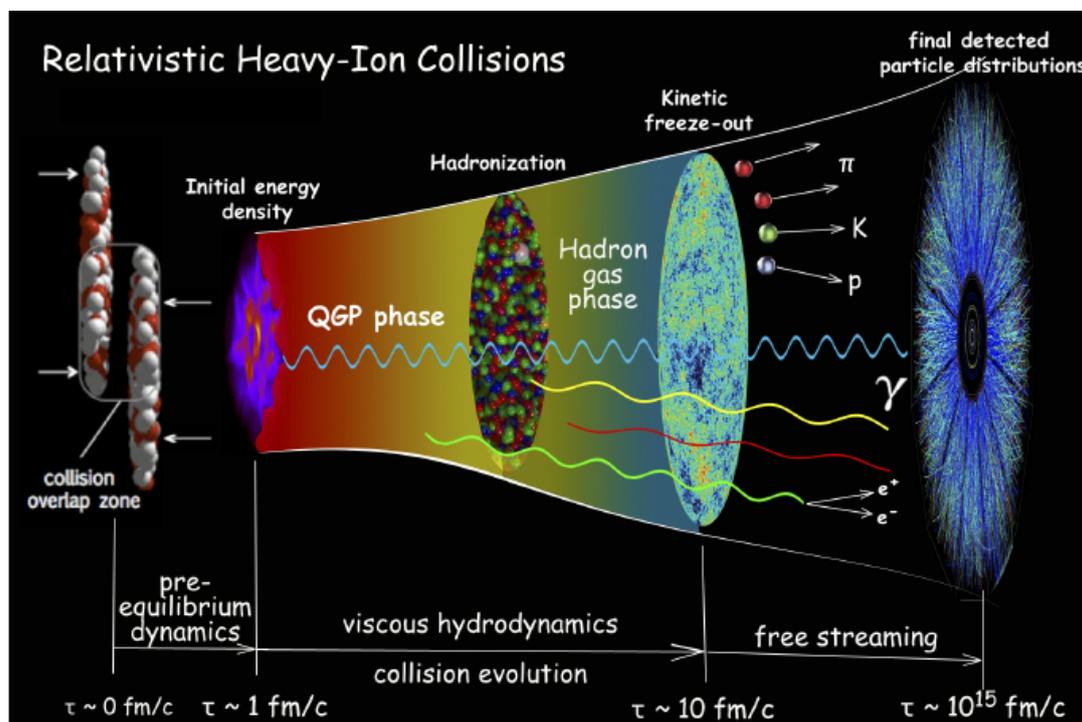


Figura 1.5: Rappresentazione di una collisione tra ioni pesanti ultrarelativistici e del processo di formazione ed evoluzione del QGP.

1.5 Evidenze sperimentali del QGP

Il QGP è uno stato della materia difficilmente osservabile e con vita media estremamente breve che non ne permette misure dirette. Le prove sperimentali della sua esistenza sono tutte dovute a misure indirette e si suddividono in “soft probes”, nel regime non perturbativo della QCD, dovute all’interazione di partoni con basso momento trasferito e “hard probes”, nel regime perturbativo, dovute all’interazione di partoni con momento trasferito elevato.

1.5.1 Soft probes

Aumento della stranezza

Una delle prime caratteristiche riscontrabili nella formazione del QGP è l’aumento della

stranezza nei prodotti di collisioni tra ioni pesanti [10].

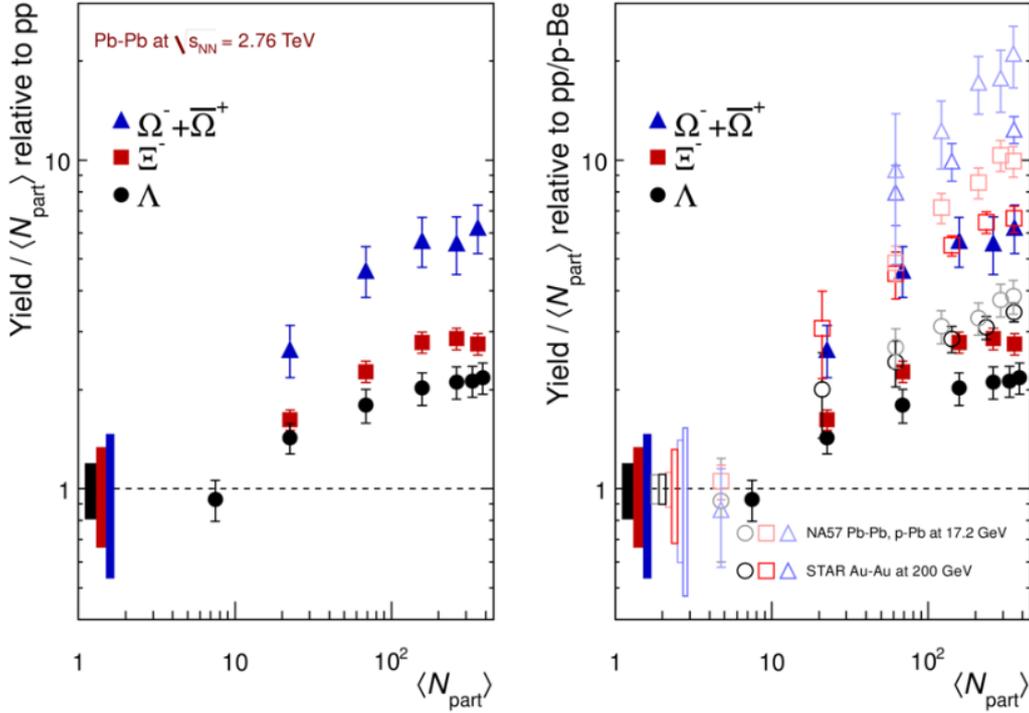


Figura 1.6: Nel grafico di sinistra è riportato l’aumento della produzione di iperoni misurato in ALICE in collisioni Pb-Pb a $\sqrt{s_{NN}} = 2.76$ TeV, che rappresentano l’aumento di stranezza in funzione del numero medio di partecipanti. Nel grafico di destra è mostrato il confronto con i dati presi a SPS e RHIC ad energie inferiori. La linea tratteggiata indica il riferimento pp [5].

I quark strange e i corrispettivi antiquark non sono presenti negli ioni di piombo iniziali e risultano quindi essere dei nuovi prodotti del processo. Nei normali processi di urti tra adroni, la creazione di stati contenenti quark strange viene soppressa a causa della minore massa dei quark up e down; tale soppressione aumenta con l’aumentare del contenuto in stranezza. In una fase di QCD invece la produzione di quark s è abbondante poichè, in seguito al parziale ripristino della simmetria chirale, la loro massa vale $m_s \sim 150$ MeV/ c^2 (massa “nuda”, ovvero senza effetti dinamici). Inoltre il principio di esclusione di Pauli limita il numero di quark u e d presenti nel plasma per la saturazione degli stati accessibili. L’elevatissima densità di gluoni all’interno del Quark-Gluon Plasma facilita la produzione di coppie $s\bar{s}$ attraverso il processo di fusione gluonica. Solo il 10 – 15% risulta dovuto all’annichilazione tra quark e antiquark, la restante parte è attribuita ad interazioni tra i gluoni che si trovano liberi a causa del deconfinamento del QGP. Quest’ultimo processo prende il nome di thermal production, in quanto assimila-

bile ad un gas di partoni in equilibrio termico. In questo stato deconfinato l'aumento di stranezza dà origine ad una maggiore probabilità, durante il freeze-out chimico, di formazione di adroni contenenti uno o più quark strange, come $\Omega^+(\bar{s}\bar{s}\bar{s})$, $\Lambda^0(uds)$, $\Xi^0(uss)$, $\Xi^-(dss)$, $\Omega^-(sss)$. Come si osserva in figura 1.6 [10], il rapporto aumenta al crescere del numero di nucleoni partecipanti alla collisione, ovvero al crescere della centralità.

Flusso ellittico

La distribuzione azimutale delle particelle prodotte in collisioni di ioni pesanti ultrarelativistici nel piano perpendicolare alla direzione dei fasci è un importante strumento per lo studio delle proprietà del mezzo creato in tali collisioni; in particolare esso è in grado di mettere in evidenza la presenza o meno di moti collettivi.

Un flusso collettivo è una correlazione tra la velocità di un elemento di volume (o di una particella) e la sua posizione nello spazio-tempo, dovuta all'esistenza di un gradiente di pressione, ed è una caratteristica dei sistemi idrodinamici.

Se consideriamo un urto centrale (parametro di impatto $b = 0$) ci aspettiamo una distribuzione azimutale isotropica. Se al contrario $b \neq 0$, l'asimmetria geometrica iniziale provoca un gradiente di pressione anisotropo nel piano trasverso alla collisione. La distribuzione azimutale finale delle particelle emesse può essere scritta utilizzando un'espansione in serie di Fourier come

$$E \frac{d^3N}{dp^3} = \frac{1}{2\pi} \frac{d^2N}{p_T dp_T dy} \left(1 + 2 \sum_{n=1}^{\infty} \nu_n \cos n(\phi - \Phi_R) \right)$$

dove Φ_R è l'angolo di inclinazione del parametro d'impatto nel piano trasverso.

Il secondo termine armonico ν_2 , chiamato coefficiente di flusso ellittico, è quello che maggiormente caratterizza la geometria iniziale del sistema.

La figura 1.7 mostra la misura del coefficiente ν_2 in ALICE per diverse particelle prodotte in collisioni Pb-Pb a $\sqrt{s_{NN}} = 5.02$ TeV in funzione dell'impulso trasverso, suddivisa in intervalli di centralità. Come si può vedere, per bassi valori di p_T ($p_T < 2-3$ GeV/c) l'andamento di ν_2 è dominato dalla massa delle particelle: particelle più leggere possiedono un coefficiente ν_2 maggiore a parità di p_T . Questo comportamento è in accordo con una descrizione idrodinamica di un fluido a bassissima viscosità in espansione. Per p_T maggiori invece le particelle sono raggruppate in base al numero di quark costituenti (2 per i mesoni e 3 per i barioni), in accordo quindi con l'assunzione che tali particelle siano state prodotte per coalescenza, o ricombinazione, di quark all'interno di uno stato di QGP in espansione.

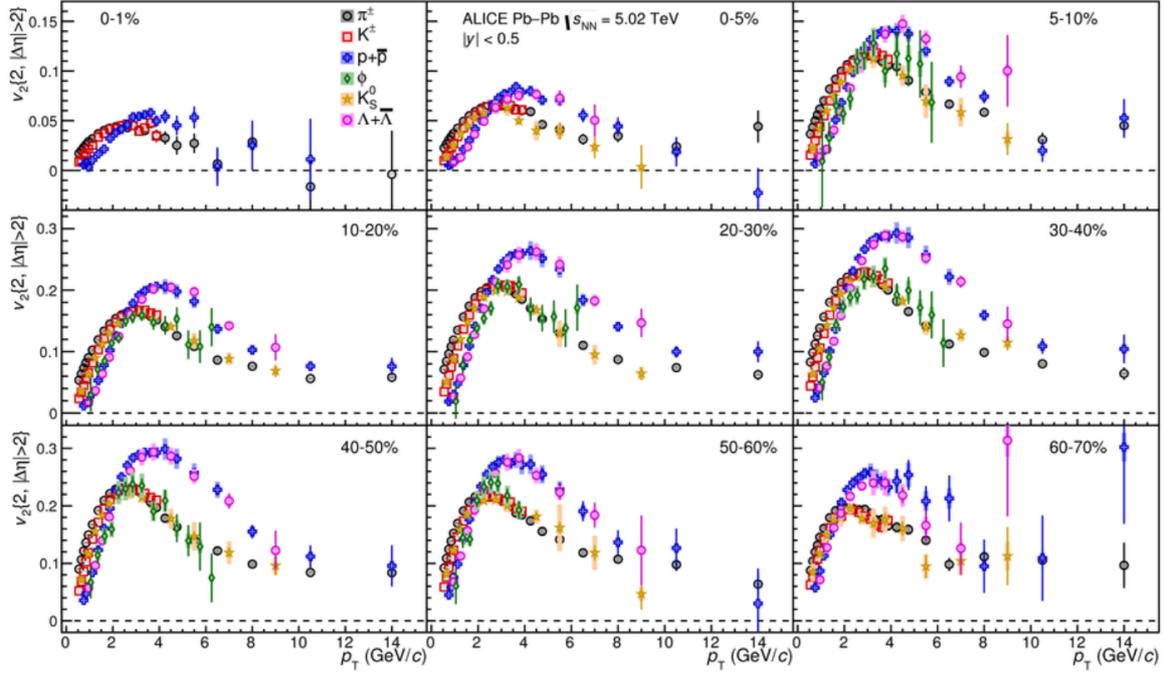


Figura 1.7: Misura del coefficiente ν_2 in ALICE per diverse particelle prodotte in collisioni Pb-Pb a $\sqrt{s_{NN}} = 5.02$ TeV in funzione dell'impulso trasverso, suddivisa in intervalli di centralità.

1.5.2 Hard probes

Jet quenching

Un jet è un aggregato di particelle con proprietà cinematiche strettamente correlate, derivanti dall'adronizzazione di uno o più partoni con elevato impulso trasverso. Con jet quenching ci si riferisce alla perdita di energia di tali jet, dovuta all'interazione forte che avviene con il QGP in cui essi si propagano. Questa perdita di energia avviene principalmente per irraggiamento di gluoni emessi per Bremsstrahlung¹. Come si osserva in figura 1.8, i partoni ad alta energia vengono prodotti in coppie con un angolo di 180° nel sistema del centro di massa. Questo porta ad una correlazione tra le particelle contenute nei 2 jet con un picco a $\Delta = \pi$ in urti pp e d-Au; il secondo picco a 180° (visibile in figura 1.9) è più largo solo per effetti di trigger. Questa correlazione sopravvive per collisioni ione-ione periferiche, mentre scompare se si considerano urti centrali [11] [12]. Tale effetto può essere spiegato assumendo che tutti gli adroni ad alto p_T su cui si effettua il trigger siano stati prodotti vicino alla superficie del mezzo caldo in espansione. In

¹Il Bremsstrahlung è la radiazione che viene emessa quando particelle cariche, in questo caso con carica di colore, vengono accelerate o decelerate

questo modo si riesce a vedere solo il gett che punta verso l'esterno, mentre l'altro viene assorbito perdendo gradualmente energia all'interno del QGP.

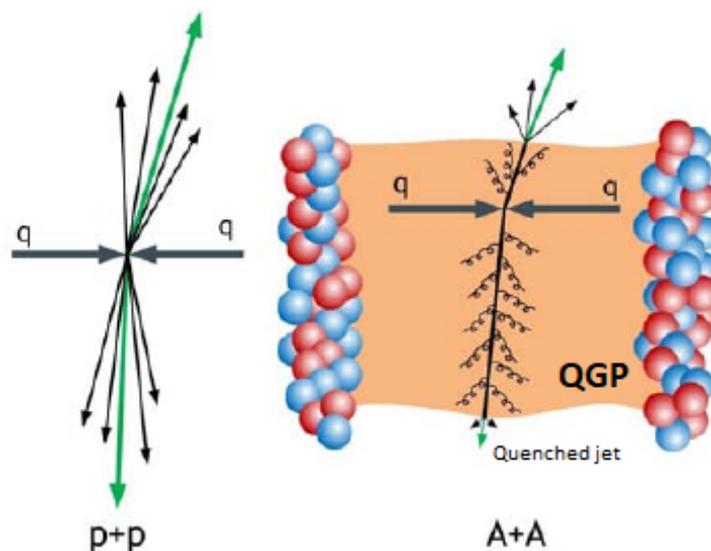


Figura 1.8: In figura è riportato il fenomeno del jet quenching in una collisione centrale tra due ioni pesanti, in confronto ad una collisione protone-protone. Come si osserva, attraversando il QGP, il partone perde energia mediante Bremsstrahlung, dando origine ad una riduzione del jet che risulta avere una dipendenza quadratica dalla distanza percorsa all'interno del plasma [12].

Il jet che emerge dal QGP ha un'energia minore rispetto a quella che avrebbe in sua assenza. In particolare l'effetto che viene riscontrato è una perdita di energia trasversale, mentre gli spettri delle collisioni periferiche rimangono praticamente inalterati. La soppressione di impulso trasversale è molto evidente nelle collisioni centrali quando comparato con quello osservato nelle collisioni pp. La centralità viene misurata sfruttando il percentile definito come

$$c = \frac{1}{\sigma_{AA}} \int_0^b \frac{d\sigma}{db'} db'$$

dove σ è la sezione d'urto e b è chiamato parametro di impatto tra i due nuclei A e A che collidono e rappresenta la distanza tra i centri dei due nuclei, sul piano perpendicolare al fascio.

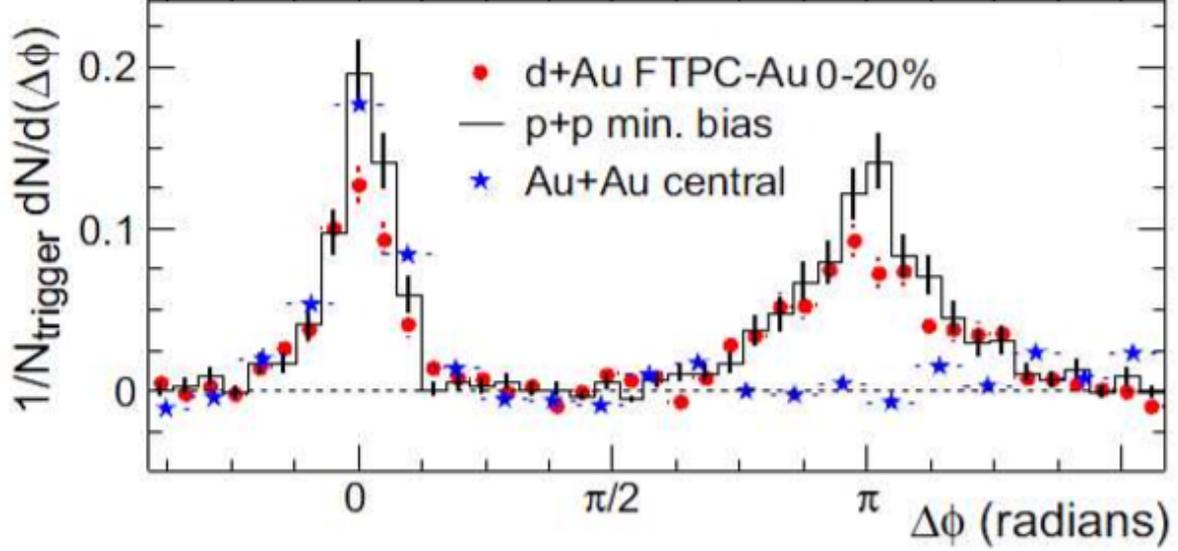


Figura 1.9: Correlazione angolare tra coppie di particelle cariche, in riferimento ad una particella “trigger”. Come si osserva, le collisioni protone-protone e quelle deutero-oro danno origine ad una coppia di jet rappresentati dai due picchi ben visibili ad angoli di 0 e π radianti, mentre i dati della collisione Au-Au danno origine al picco sullo zero, ma non a quello a π , indice che il secondo jet ha subito una riduzione dovuta al QGP fino ad annullarsi [13].

L’entità della soppressione che si osserva è quantificata dal rapporto R_{AA} , detto fattore di modificazione nucleare e definito dalla seguente formula basata sul modello di Glauber²:

$$R_{AA} = \frac{d^2 N^{AA}/dp_T d\eta}{\langle T_{AA} \rangle d^2 \sigma_{inel}^{pp}/dp_T d\eta}$$

dove $\langle T_{AA} \rangle = \langle N_{coll} \rangle / \sigma_{inel}^{NN}$ è la funzione di sovrapposizione, proporzionale al numero medio di collisioni binarie nucleone-nucleone e alla loro sezione d’urto, mentre $d^2 N^{AA}/dp_T d\eta$ rappresenta il numero di particelle prodotte in una collisione A-A per unità di impulso trasverso (p_T) e di pseudorapidità e $\sigma_{inel}^{pp}/dp_T d\eta$ è la sezione d’urto inelastica della collisione pp in un dato intervallo di impulso trasverso e pseudorapidità.

Se $R_{AA} = 1$, significa che il numero di particelle prodotte negli urti A-A scala con il numero di collisioni binarie nucleone-nucleone, mentre $R_{AA} \neq 1$ indica la presenza di effetti dovuti alla formazione di un mezzo deconfinato (QGP).

²Modello di diffusione multipla il quale stabilisce che un adrone che attraversa un nucleo subisce più interazioni, ciascuna delle quali produce nuovi adroni che interagiscono a loro volta con il nucleo stesso dando il via ad una cascata nucleare.

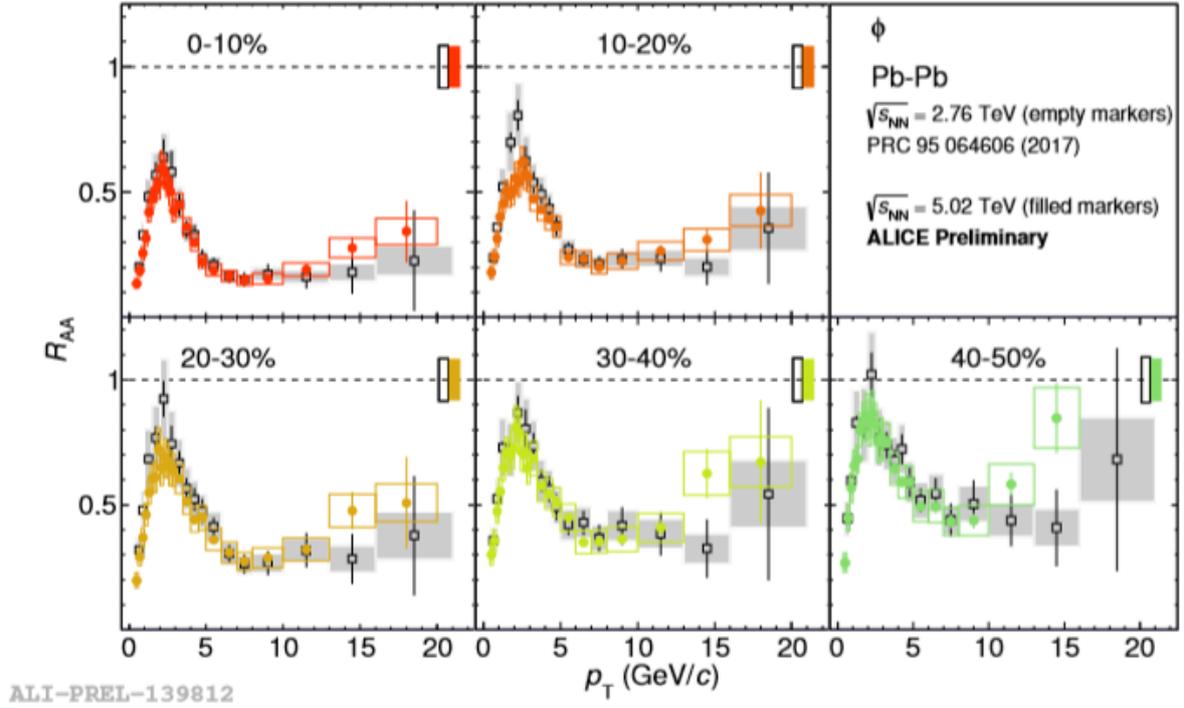


Figura 1.10: Fattore di modificazione nucleare R_{AA} in funzione dell'impulso trasverso in vari intervalli di centralità. I dati corrispondono ad una collisione Pb-Pb in ALICE a $\sqrt{s_{NN}} = 2.76$ TeV [14].

In particolare:

- $R_{AA} < 1$ indica soppressione
- $R_{AA} > 1$ indica amplificazione

La misura del fattore R_{AA} effettuata in ALICE in collisioni piombo-piombo a $\sqrt{s_{NN}} = 5.02$ TeV indica la soppressione di particelle cariche ad alto impulso trasverso.

Come si osserva in figura 1.10, la distribuzione in funzione dell'impulso trasverso presenta un minimo marcato per $p_T = 6-7$ GeV/c rispetto a collisioni pp. I risultati ottenuti in ALICE e a RHIC arrivano alle medesime conclusioni, con una diversa magnitudine, dovuta ad un volume maggiore del QGP ottenuta ad LHC, a causa della maggiore energia del centro di massa nella collisione.

Produzione di J/ψ

L'insieme degli stati legati di quark pesanti come il charmonium ($c\bar{c}$) e bottonium ($b\bar{b}$) è chiamato Quarkonia.

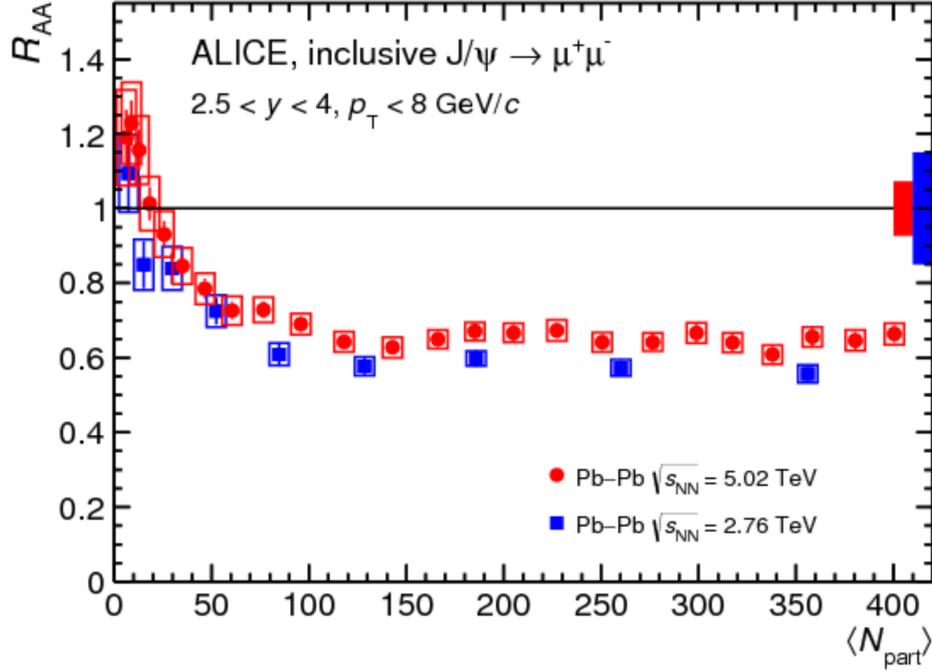


Figura 1.11: Fattore di modificazione R_{AA} per la produzione di J/ψ , in funzione dei nucleoni partecipanti collisioni Pb-Pb a $\sqrt{s_{NN}} = 5.02$ TeV e a $\sqrt{s_{NN}} = 2.76$ TeV misurato in ALICE [15].

In un mezzo deconfinato ad alte temperature come il Quark-Gluon Plasma ci si aspetta che la produzione di stati legati appartenenti a Quarkonia sia soppressa rispetto a collisioni pp, come si osserva in figura 1.11. La causa di tale soppressione è la schermatura di colore della forza che tiene legati tali quark. Stati fortemente legati come J/ψ e Υ si dissolvono a temperature più alte rispetto a stati più debolmente legati. L'effetto di schermaggio della carica di colore che avviene all'interno della materia nucleare densa e calda, divide le J/ψ nei suoi costituenti c e \bar{c} ; finché il raggio di schermatura è maggiore del raggio adronico, i quark vengono tenuti assieme dalla forza forte, nel momento in cui le dimensioni dello stato legato diventano maggiori del raggio di schermatura, avviene il deconfinamento: i costituenti si separano e nel momento dell'adronizzazione si ricombinano con i quark leggeri per formare mesoni D. Il grado di soppressione della J/ψ è proporzionale alla temperatura e alla densità di carica di colore.

1.5.3 Prove elettromagnetiche: fotoni diretti

I fotoni diretti, definiti come fotoni non prodotti dal decadimento di adroni, rappresentano un importante strumento per lo studio dell'evoluzione del QGP.

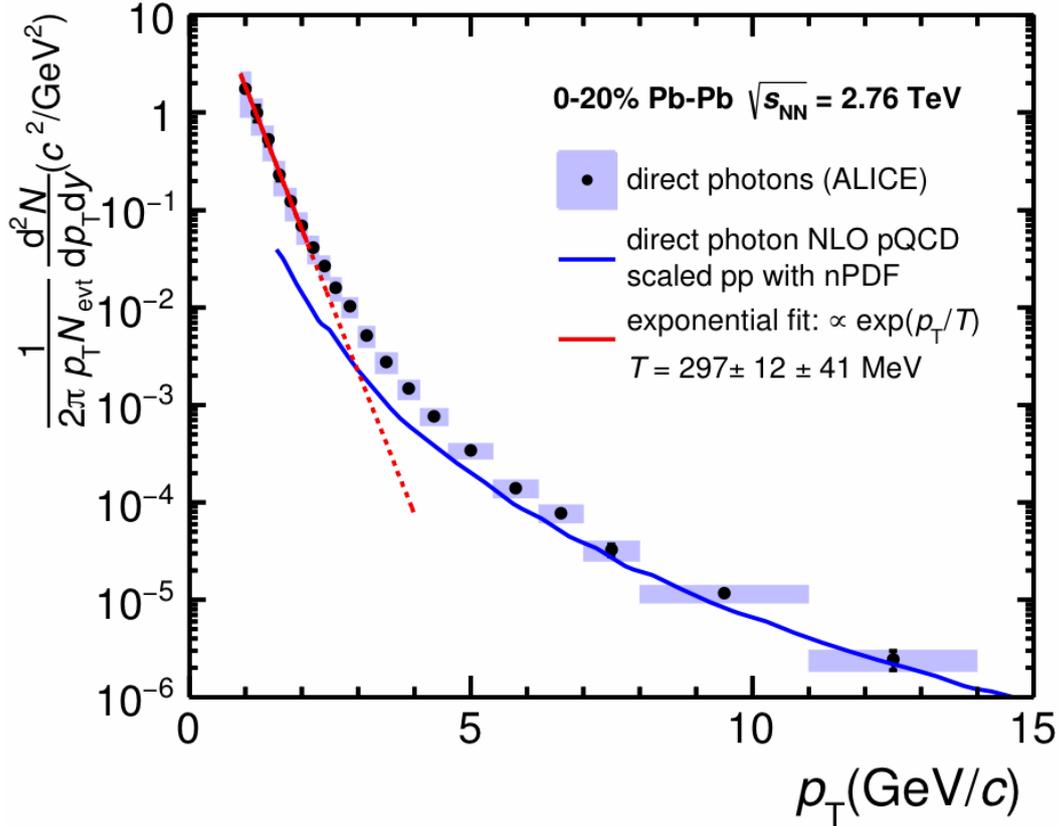


Figura 1.12: Spettro di fotoni diretti misurato in ALICE in collisioni centrali Pb-Pb a $\sqrt{s_{NN}} = 2.76$ TeV

A differenza degli adroni, i fotoni diretti vengono prodotti durante tutte le fasi della collisione, ma, non avendo carica di colore forte, non interagiscono con i costituenti del plasma, fornendo informazioni dirette sulle condizioni presenti nell'istante in cui sono stati creati, come temperatura, moti collettivi ed evoluzione spazio-temporale del sistema.

Diversi intervalli di p_T corrispondono a fotoni emessi in momenti diversi; ad alti p_T ($\gtrsim 5$ GeV/c) lo spettro è dominato da fotoni prompt, creati in urti ad alto momento trasferito (hard probes), mentre per impulsi minori lo spettro è dominato da fotoni termici prodotti nella fase partonica e adronica.

Questo tipo di misura risulta complicato nella pratica a causa del grande fondo, dovuto soprattutto al decadimento di mesoni π^0 e η , e della separazione delle diverse componenti nello spettro inclusivo. In figura 1.12 è riportato lo spettro di fotoni diretti misurato in ALICE in collisioni centrali Pb-Pb a $\sqrt{s_{NN}} = 2.76$ TeV. Confrontando gli spettri misurati in collisioni protone-protone con quelli in collisioni tra ioni pesanti, si osserva per questi ultimi un maggior numero di fotoni prodotti a bassi impulsi trasversi. Dal fit esponenziale dello spettro nell'intervallo di centralità 0-20% si ottiene una temperatura di circa 297 MeV, che corrisponde alla temperatura della sorgente di fotoni mediata lungo tutta la sua evoluzione temporale. Tale valore indica una temperatura iniziale molto al di sopra della temperatura critica alla quale si forma il QGP, pari a circa 150–160 MeV.

Capitolo 2

Il rivelatore ALICE a LHC

2.1 Large Hadron Collider (LHC)

Il Large Hadron Collider (LHC) è il più grande e potente acceleratore al mondo, situato al CERN di Ginevra e attivo dal 2008. Si tratta di un acceleratore adronico circolare che si estende all'interno di un tunnel di 27 chilometri di circonferenza, nel quale vengono fatti circolare due fasci di particelle fino ad una velocità prossima a quella della luce (99.99999%) e fatti collidere in più punti, in corrispondenza dei quattro esperimenti: ALICE, ATLAS, CMS, LHCb, come mostrato in figura 2.1. I fasci viaggiano in direzioni opposte all'interno dei tubi a vuoto (beam pipe) in condizione di ultralight vacuum¹ e vengono guidati da un sistema di magneti superconduttori che lavorano ad una temperatura di 1.9K, inferiore a quella dello spazio aperto.

2.1.1 Il sistema di magneti

Le particelle cariche accelerate (protoni e ioni pesanti) effettuano circa undicimila giri al secondo all'interno di LHC. Prima di essere iniettate nell'anello principale, esse vengono accelerate passando all'interno di una serie di acceleratori circolari e rettilinei, venendo inserite nel successivo ogni volta che raggiungono la velocità massima alla quale l'acceleratore in cui si trovano può portarle. Il fascio viene mantenuto sulla traiettoria circolare attraverso un vasto reticolo, composto da più di 50 tipi di magneti differenti. I più potenti sono i 1232 Main Dipoles, ciascuno di 15 metri di lunghezza e pesante 35 tonnellate, che producono un campo magnetico di circa 8.3 T. Prima di entrare nelle aree sperimentali dei vari esperimenti, i fasci di particelle vengono deflessi dai magneti di inserimento in modo da collidere frontalmente nei punti di interazione. Oltre a deviare il fascio per mantenerlo sulla traiettoria, è infatti necessario focalizzarlo, in quanto, per

¹Vuoto dell'ordine di $10^{-10} - 10^{-11}$ mbar, ottenuto mantenendo i tubi ad una a temperature molto basse e pompando criogenicamente 9000 metri cubi di gas.

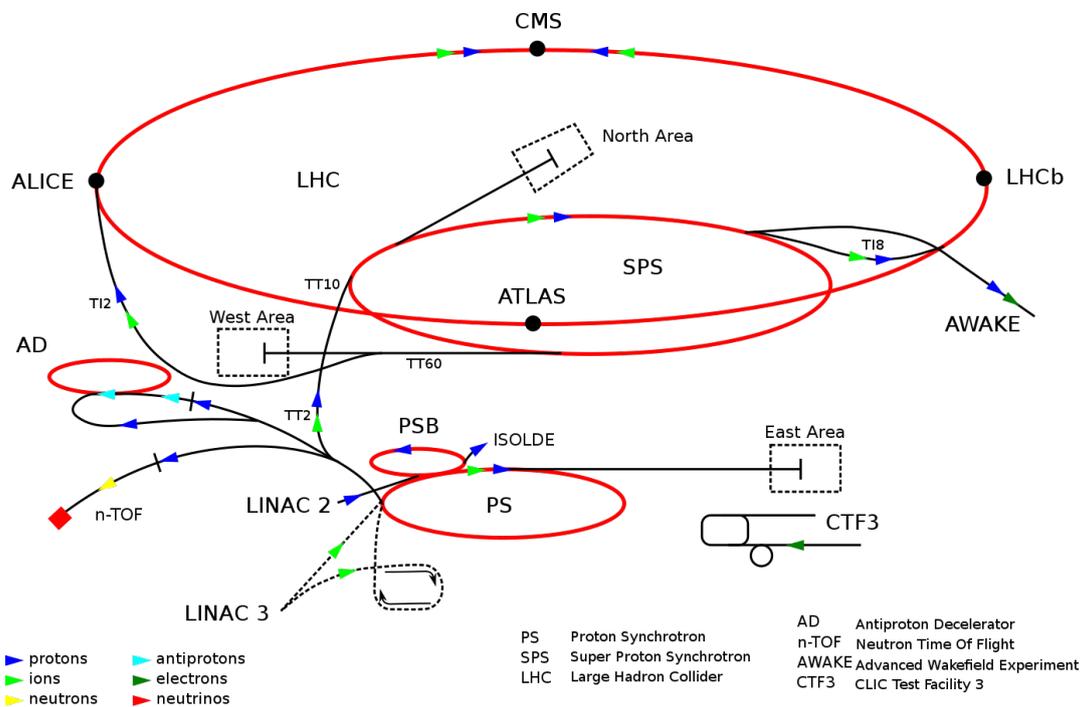


Figura 2.1: Rappresentazione del complesso di acceleratori al CERN, e degli esperimenti dislocati sull'acceleratore principale LHC.

repulsione coulombiana, le particelle all'interno del fascio tenderebbero a disperdersi. Questo compito è svolto da 392 quadripoli magnetici disposti in modo da agire alternativamente sulla verticale del fascio (quadripoli D) e sul piano orizzontale (quadripoli F), come mostrato in figura 2.2.

In diversi punti si trova una disposizione particolare di multipoli chiamata FODO, nella

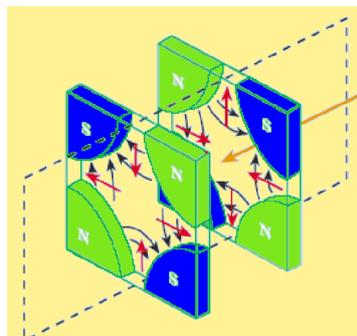


Figura 2.2: Rappresentazione dei quadripoli magnetici che agiscono in modo alternato su due piani perpendicolari per mantenere il fascio focalizzato.

quale si osserva un dipolo O alternato a un dipolo F e successivamente ad uno D, il primo crea la deflessione del fascio causandone, però, nel frattempo, una delocalizzazione, che viene corretta dagli altri due [16].

2.2 L'esperimento ALICE

ALICE (A Large Ion Collider Experiment) è uno dei quattro grandi esperimenti a LHC, l'unico espressamente progettato per lo studio della fisica degli ioni pesanti ultrarelativistici e del QGP. In particolare, l'esperimento è stato ottimizzato per collisioni di nuclei di piombo ad energie fino a 5.5 TeV per coppia di nucleoni. L'apparato sperimentale, mostrato in figura 2.3, ha una lunghezza di circa 20 metri, con una larghezza e un'altezza di 16 metri, ed è costituito da due parti principali: la parte centrale, centrata attorno al punto di interazione, racchiusa all'interno di un magnete solenoidale che crea un campo magnetico costante di 0.5 T e formata da rivelatori dedicati allo studio dei segnali adronici ed elettromagnetici, e lo spettrometro di muoni, dedicato allo studio delle proprietà degli stati Quarkonia.

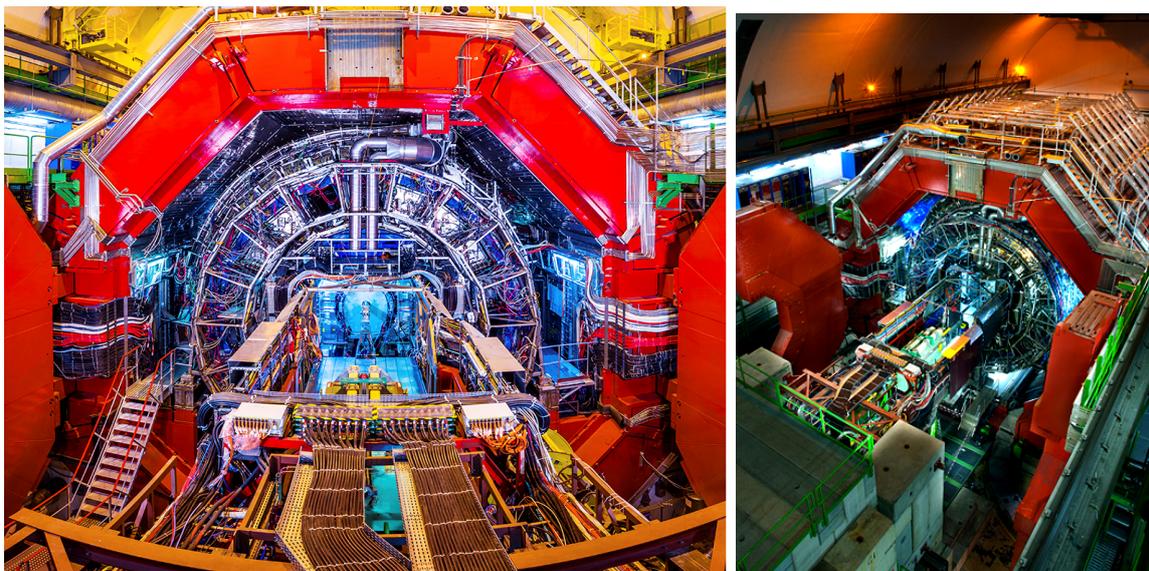


Figura 2.3: Foto dell'esperimento ALICE.

L'apparato sperimentale è composto da 18 rivelatori, che possono essere suddivisi in base alla loro funzione [17]:

- I *rivelatori di tracciamento* (tracking detectors) hanno il compito di ricostruire i vertici di interazione e di decadimento e le traiettorie delle particelle cariche nel campo magnetico e permettono di ottenere misure accurate di posizione e impulso.

THE ALICE DETECTOR

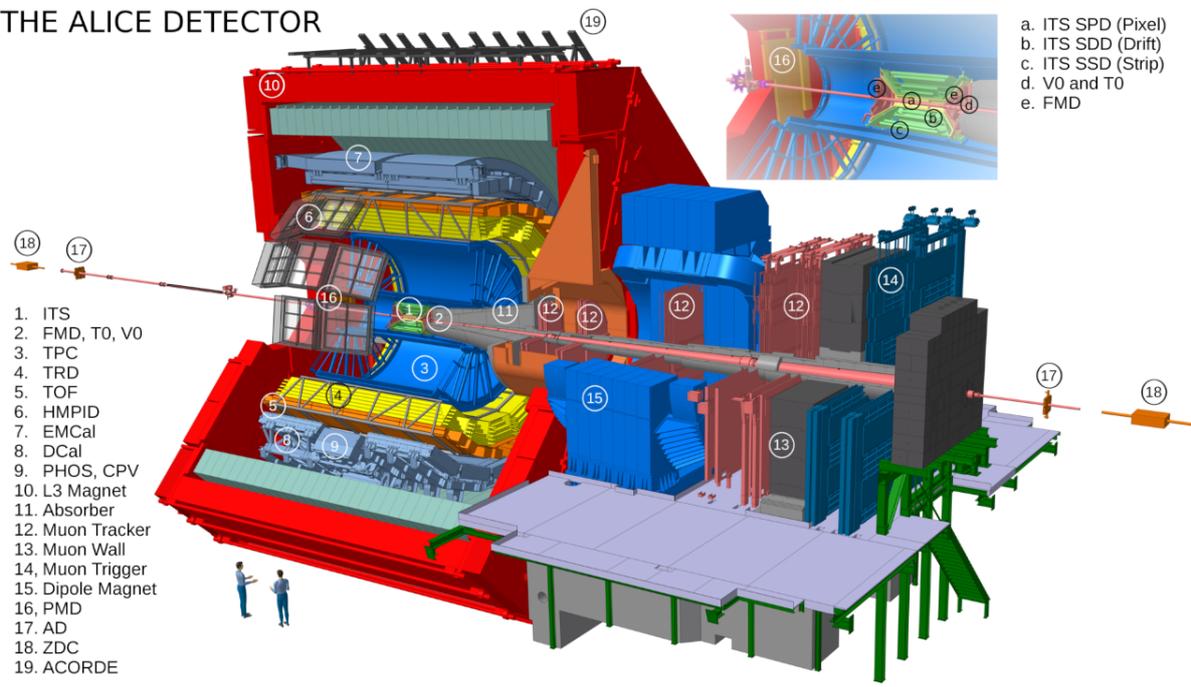


Figura 2.4: Rappresentazione schematica della struttura del rivelatore ALICE. In alto a destra si osserva un ingrandimento degli strati più vicini al punto di collisione.

Ne fanno parte l'ITS (Inner Tracking System), la TPC (Time Projection Chamber) e il TRD (Transition Radiation Detector).

- I *rivelatori per l'identificazione di particelle* (Particle Identification) permettono di associare una identità alle traiettorie ricostruite dai rivelatori traccianti. Ne fanno parte il TOF (Time-Of-Flight) e HMPID (High Momentum Particle Identification).
- I *calorimetri* elettromagnetici PHOS (Photon Spectrometer), EMCAL (Electro-Magnetic Calorimeter) e DCAL (Di-jet Calorimeter) permettono di misurare l'energia di particelle cariche e neutre.
- Lo *spettrometro muonico* composto da MCH (Muon Chamber) e MTR (Muon Trigger), che copre una regione di rapidità forward ($2.5 \leq \eta \leq 4$) ed è dedicato allo studio delle risonanze dei quark pesanti.
- I *rivelatori di forward e trigger* di cui fanno parte l'FMD (Forward Multiplicity Detector), il PMD (Photon Multiplicity Detector), AD (ALICE Diffractive), lo ZDC (Zero Degree Calorimeter), V0 e T0.
- Il *rilevatore di trigger di raggi cosmici* ACORDE (A COsmic Ray Detector).

2.2.1 Inner Tracking System

Il primo rivelatore che incontrano le particelle dopo la collisione è l'ITS, che costituisce lo strato più interno del cilindro di ALICE. Esso copre un range di pseudorapidità $|\eta| < 0.9$ su tutto l'angolo azimutale e si estende da un raggio di 3.9 cm dal centro ad uno di 43 cm. È costituito da sei strati concentrici di rivelatori al silicio, come mostrato in figura 2.4: 2 Silicon Pixel Detector (SPD), ideali per far fronte alla grande densità di particelle cariche prodotte in collisione ione-ione, mentre più esternamente, dove la densità di particelle è minore sono presenti 4 rivelatori 2-dimensionali, di cui 2 Silicon Drift Detector (SDD) e 2 Silicon Strip Detector (SSD). L'ITS ha come principali obiettivi la localizzazione dei vertici primari, con una risoluzione inferiore a $100 \mu\text{m}$ e la ricostruzione dei vertici secondari a partire dai decadimenti dei mesoni B e D e di iperoni². Esso ha inoltre il compito di raffinare la misura dell'impulso e la risoluzione angolare delle particelle ricostruite dalla TPC, tracciando anche quelle cariche con $p_T < 200 \text{ MeV}/c$ non riescono ad arrivare alla Time Projection Chamber.

2.2.2 Time Projection Chamber

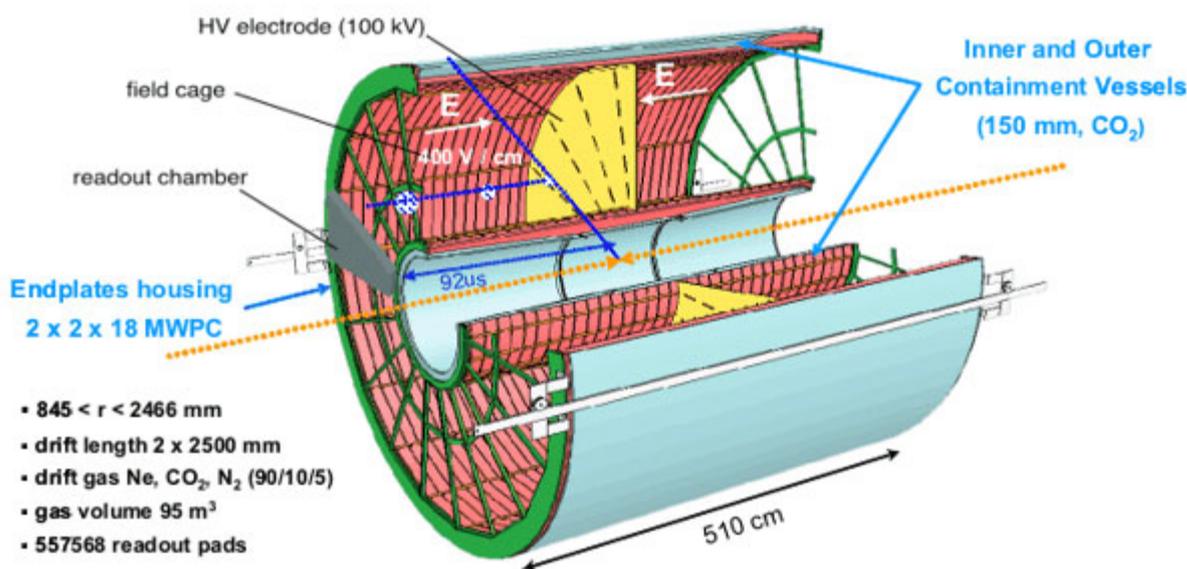


Figura 2.5: Rappresentazione schematica della TPC.

²I mesoni B sono formati da un antiquark bottom unito ad un quark leggero o ad un charm, mentre i mesoni D contengono un quark charm e un antiquark leggero. Gli iperoni sono dei barioni contenenti uno o più quark strange.

La TPC è il principale rivelatore di tracciamento di ALICE, ottimizzato per identificare particelle in un range molto ampio di impulsi, fino ad un valore di 10 GeV/c, con pseudorapidità $|\eta| < 0.9$, su tutto l'angolo azimutale. L'identificazione delle particelle a basso momento avviene attraverso la misura della perdita di energia di ionizzazione $-\frac{dE}{dx}$, descritta dalla formula relativistica di Bethe-Bloch, valida per le particelle cariche che viaggiano a velocità elevate [18]. La TPC ha inoltre il compito di ricostruire i vertici primari di particelle cariche partendo dalle regioni più esterne, dove la densità di tracce è minore, fino a quelle più interne nelle quali esse sono così fitte da avere numerose sovrapposizioni. Il rivelatore ha una forma cilindrica con raggio interno pari a 85 cm e raggio esterno di 250 cm ed una lunghezza di circa 5 metri, come mostrato in figura 2.5, coprendo un volume di 88 m³, riempiti da una miscela di gas composta per il 90% di Neon e per il 10% di CO₂ che riduce gli effetti di scattering multiplo [19].

2.2.3 Transition Radiation Detector

Esternamente alla TPC si trova il TRD, un rivelatore che si estende da un raggio di 2.9 m a uno di 3.7 m e che copre una lunghezza di 7 m. Esso è formato da un materiale inhomogeneo chiamato radiatore e da 522 camere di readout disposte su 6 livelli, come mostrato in figura 2.6, ciascuna riempita con una miscela di gas formata da 85% di Xe e da 15% di CO₂ e suddivisa da una griglia catodica in 3 regioni di drift e una di amplificazione, nella quale è presente un anodo.

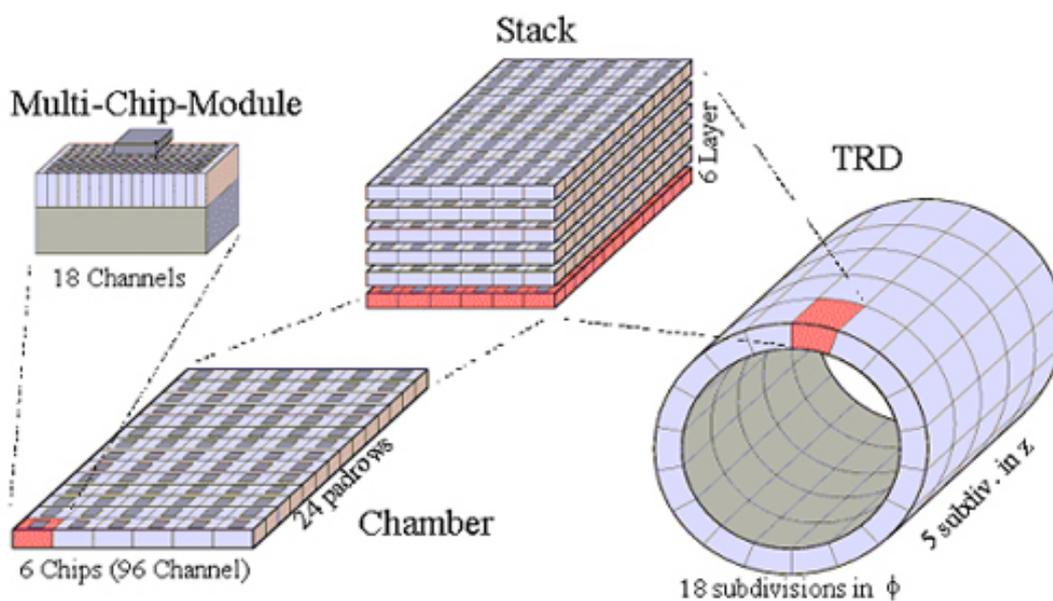


Figura 2.6: In figura è riportata una schematizzazione del TRD.

Il compito principale di questo rivelatore è l'identificazione di elettroni e positroni con impulso al di sopra di 1 GeV/c. L'identificazione avviene sfruttando la radiazione di transizione³ che le particelle cariche relativistiche e in particolare gli elettroni emettono quando passano da un materiale ad un altro con diversa costante dielettrica.

In figura 2.7 è riportata la forma del segnale elettrico indotto da un pione e da un elettrone in una camera di readout; la presenza di un picco a tempi lunghi dovuto al fotone emesso per radiazione di transizione permette di identificare gli elettroni.

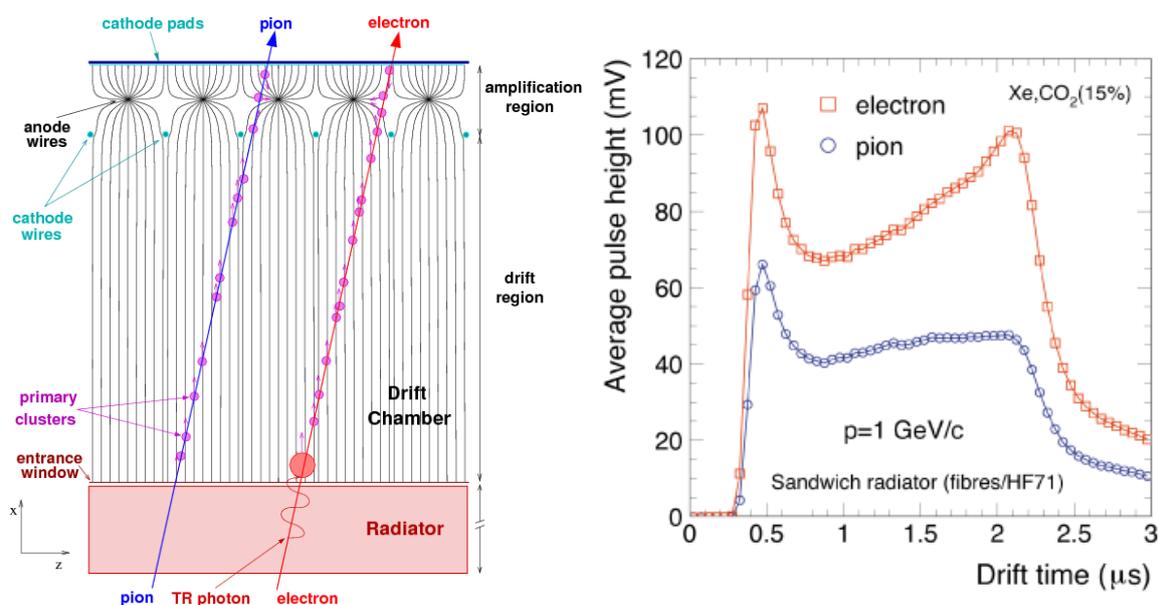


Figura 2.7: A sinistra è rappresentata la schematizzazione di una camera del TRD e a destra è riportato il segnale indotto da un pione e da un elettrone in funzione del tempo di deriva; il secondo picco a tempi lunghi sulla curva rossa, dovuto al fotone emesso, denota che tale curva appartiene ad un elettrone.

2.2.4 High-Momentum Particle Identification Detector

Il rivelatore HMPID è dedicato alla rilevazione degli adroni con alto p_T . In particolare è ottimizzato per discriminare π e K fino a 3 GeV/c e K e p fino a 5 GeV/c. Si tratta di un dispositivo di 10 m² situato ad una distanza di 4.9 m dal centro di collisione.

È un rivelatore di tipo RICH (Ring-Imaging Cherenkov detector), cioè si basa sullo studio delle caratteristiche della radiazione Cherenkov che viene emessa quando una particella passa in un radiatore contenente un materiale dielettrico (C_6F_{12}) con una velocità

³Essa è proporzionale al fattore relativistico γ che ad alte energie risulta $\simeq \frac{p}{m}$; per cui un pione carico deve avere un impulso molto elevato per produrre la stessa radiazione dell'elettrone.

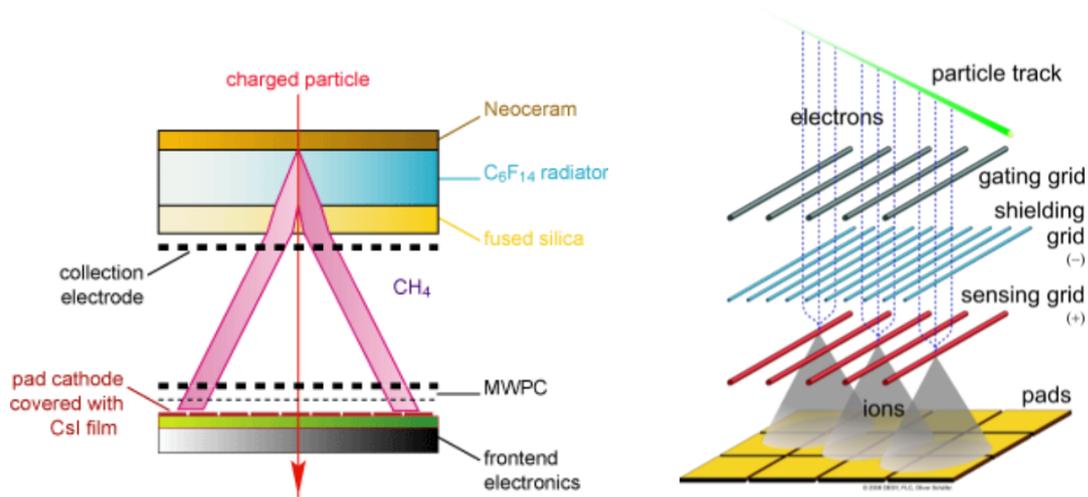


Figura 2.8: A sinistra è riportata una rappresentazione di HMPID, mentre a destra è descritto lo schema di funzionamento di un rivelatore MWPC.

maggiore di quella che avrebbe la luce al suo interno. L'interferenza dei fronti d'onda sferici produce un'onda elettromagnetica in direzione θ_c legata alla velocità secondo la formula $\cos(\theta_c) = \frac{1}{n\beta}$ dove n è l'indice di rifrazione del mezzo [20]. I fotoni Cherenkov vengono quindi rivelati da contatori di fotoni, realizzati ricoprendo il catodo di una camera proporzionale multifili (Multiwire Proportional Chamber, MWPC) con un sottile strato di CsI; vengono quindi prodotte immagini di corone circolari in cui raggi interni ed esterni sono proporzionali a θ_c . Misurando il raggio si riesce quindi a risalire alla velocità delle particelle e, conoscendone l'impulso, anche alla loro massa e di conseguenza alla loro identità. In figura 2.8 sono riportate la schematizzazione dell'HMPID e del funzionamento di una MWPC con fotocatodo al CsI [21].

2.2.5 Calorimetri e altri rivelatori

I calorimetri misurano l'energia di una particella assorbendola completamente e arrestandone quindi la corsa. Essi si dividono in calorimetri adronici ed elettromagnetici; i primi si occupano degli adroni, mentre i secondi sono dedicati alle misure di fotoni, elettroni e positroni.

Il PHOS è un calorimetro elettromagnetico ad alta risoluzione formato da 17920 celle costituite da metalli di piombo tungstanato ($PbWO_4$), che effettua misure sullo spettro dei fotoni ad alta energia per avere informazioni sulle proprietà termiche e dinamiche delle fasi iniziali della collisione. L'analisi dei fotoni diretti è importante perché un aumento di tali particelle con impulso $p_T > 2 \text{ GeV}/c$ potrebbe indicare la presenza di QGP. In aggiunta la misura dei fotoni prodotti in alcuni decadimenti come $\pi_0 \rightarrow \gamma\gamma$, permette

anche di ottenere alcune prime informazioni riguardanti il jet quenching. Lo studio della fisica di questo fenomeno viene però attribuita principalmente all'EMCal e al DCal capaci di lavorare su un ampio range di impulsi. Si tratta di calorimetri a campionamento⁴ scintillatore-piombo a energia cilindrica, ottimizzati per l'analisi degli hard jet, permettendone una ricostruzione dettagliata in tutti i tipi di collisioni, attraverso anche la misura delle componenti neutre [22].

Sono poi presenti diversi altri rivelatori, come il PMD, adibito alla misura della molteplicità e della distribuzione dei fotoni, lo Spettrometro a Muoni, che attraverso la rivelazione dei muoni permette di ricostruire lo spettro dei Quarkonia pesanti e del mesone ϕ e l'FMD che misura le particelle cariche negli intervalli di rapidità $-3.4 < \eta < -1.7$ e $1.7 < \eta < 5.0$ [23]. Infine si trovano calorimetri adronici, come lo ZDC, che effettua misure sui neutroni e i protoni spettatori per ottenere informazioni sulla centralità degli eventi [24] e i rivelatori V0 e T0, il primo con lo scopo di stimare la centralità delle collisioni e rigettare il fondo dovuto alle interazioni con il gas residuo contenuto nella beam pipe, mentre il secondo fornisce il tempo reale in cui è avvenuto l'urto.

2.3 Il sistema a tempo di volo (TOF)

Il TOF è un rivelatore a tempo di volo, dedicato all'identificazione di particelle cariche ad impulsi intermedi in grado di separare a 3σ π/K fino ad impulsi di $2.5 \text{ GeV}/c$ e p/K fino a $4 \text{ GeV}/c$.

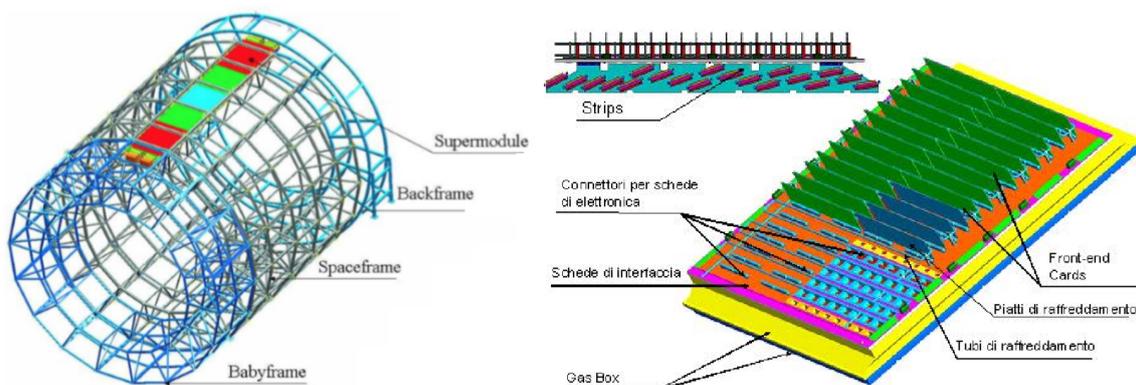


Figura 2.9: In figura si osserva la struttura modulare del TOF e la schematizzazione di uno dei moduli che la compongono.

Esso è collocato a 3.7 m di distanza dall'asse centrale ed ha una struttura cilindrica che copre la regione centrale di pseudorapidità $|\eta| < 0.9$ sull'intero angolo azimutale Φ ,

⁴Un calorimetro a campionamento è formato da un'alternanza di rivelatori e strati di assorbimento, a differenza del calorimetro omogeneo, nel quale lo stesso materiale risponde ad entrambe le funzionalità.

per un'area complessiva di 140 m². L'apparato ha una struttura modulare, formata da 18 settori in Φ e 5 segmenti lungo la coordinata assiale. Ciascuno dei 90 moduli contiene un numero variabile di Multigap Resistive Plate chamber strips (MRPC), di cui si tratterà nel paragrafo seguente, per un totale di 1638 ed è diviso in due regioni, una interna contenente una miscela gassosa di tetrafluoroetano ($C_2H_2F_4$), isobutano (C_4H_{10}) ed esafluoro di zolfo (SF_6) con rapporto 90:5:5, in cui sono collocate le MRPC e una esterna in cui è presente l'elettronica di front-end collegata attraverso fili di varia lunghezza ai moduli di readout. Una rappresentazione di un modulo è riportata in figura 2.9 [25].

2.3.1 Multigap Resistive Plate Chamber

Una MRPC è un rivelatore costituito da una serie di piani paralleli alternati a spazi contenenti gas. In figura 2.10 è mostrata una rappresentazione schematica di una MRPC

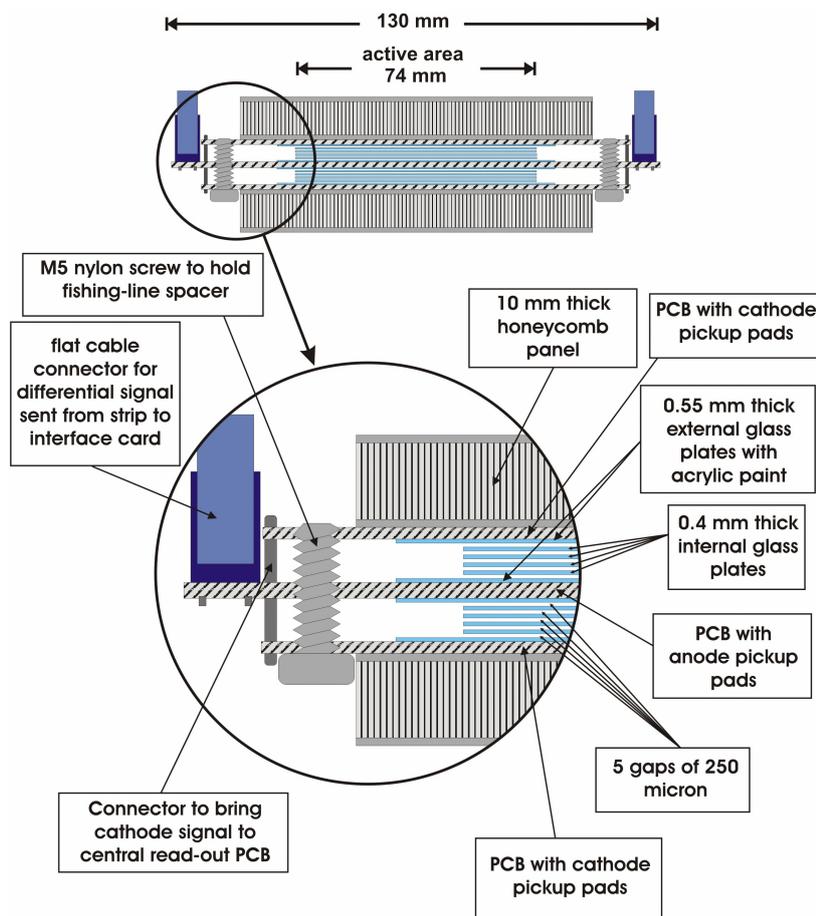


Figura 2.10: In figura è riportata la schematizzazione di un rivelatore MRPC del TOF di ALICE.

del rivelatore TOF di ALICE, in cui sono visibili i 10 gaps da $250 \mu\text{m}$ di spessore ciascuno, disposti in due serie da 5 gaps (configurazione a double stack). La MRPC rappresenta l'evoluzione della Resistive Plate Chamber (RPC) a singolo gap. Una RPC è un rivelatore dal disegno relativamente semplice, e consiste essenzialmente in due elettrodi piani paralleli, costituiti da materiale ad elevata resistività, racchiudenti al loro interno un volume di gas a pressione atmosferica. L'alta tensione viene applicata a degli elettrodi posti esternamente rispetto ai due piani resistivi; l'alto valore di resistività dei piani permette di operare il rivelatore per valori del campo elettrico estremamente elevati. L'utilizzo delle RPC come rivelatori di tempo di volo è però precluso dal fatto che è impossibile operare tali rivelatori in regime proporzionale (condizione necessaria per avere risoluzioni temporali inferiori a qualche centinaio di ps) ed avere efficienze elevate, come mostrato in figura 2.11. Questo problema è stato superato con lo sviluppo delle MRPC, che consisto-

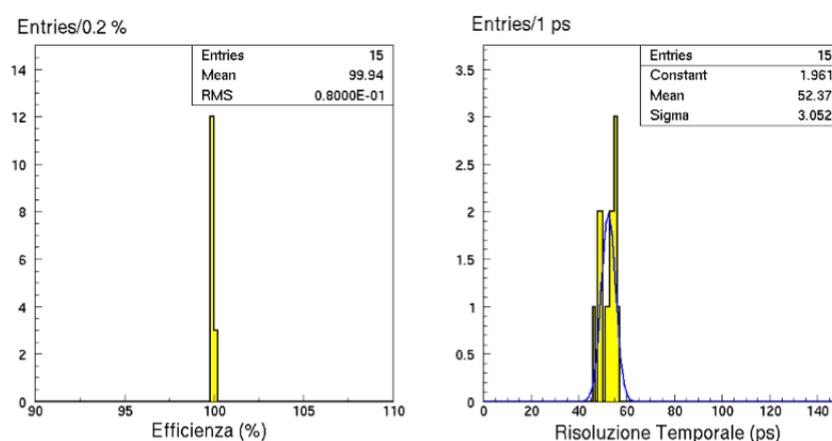


Figura 2.11: Grafici che riportano i risultati dei test effettuati su alcune MRPC del TOF, in cui si osserva un'efficienza del 99.9% e una risoluzione temporale di 52.4 ps.

no sostanzialmente in delle RPC in cui vengono inseriti dei piani resistivi equidistanziati, dello stesso materiale di quelli esterni e paralleli ad essi, in modo da suddividere la distanza tra i due elettrodi in tanti intervalli uguali ed indipendenti. Questi piani interni non sono connessi con l'alta tensione, ma sono forzati elettrostaticamente a prendere il valore del potenziale in modo da suddividere la tensione applicata ai due elettrodi esterni in maniera uguale per ogni gap. Una particella carica che attraversa il volume gassoso provoca la ionizzazione del gas e successivamente la moltiplicazione a valanga degli elettroni liberati per effetto dell'elevato campo elettrico. I piani resistivi sono trasparenti al segnale indotto dal movimento delle cariche nel rivelatore: il segnale prelevato sugli elettrodi di pick-up è quindi la somma analogica di tutti i segnali indotti dalle singole valanghe. Ogni MRPC del TOF di ALICE contiene 96 di tali pick-up pad, ognuno di area $2.5 \times 3.5 \text{ cm}^2$, per un totale di 157248. Rispetto alla configurazione asimmetrica a singolo stack, la configurazione double stack permette, a parità di numero di gap totali,

di ottenere la stessa intensità del campo elettrico all'interno del rivelatore applicando metà della tensione. Inoltre, gli elettrodi sono più vicini tra loro, per cui l'impronta della valanga su di essi risulta minore; il segnale è indotto su un'area limitata, riducendo gli effetti di bordo tra pad adiacenti [26].

2.3.2 Identificazione di particelle con TOF

L'identificazione delle particelle nel TOF avviene mediante la misura del tempo di volo. La lunghezza della traiettoria e l'impulso p delle particelle sono già noti, quest'ultimo fornito dalle misure effettuate dai tracking detectors descritti nei paragrafi precedenti e in particolare dalla TPC. Utilizzando le equazioni relativistiche $E = \gamma mc^2$ e $E^2 =$

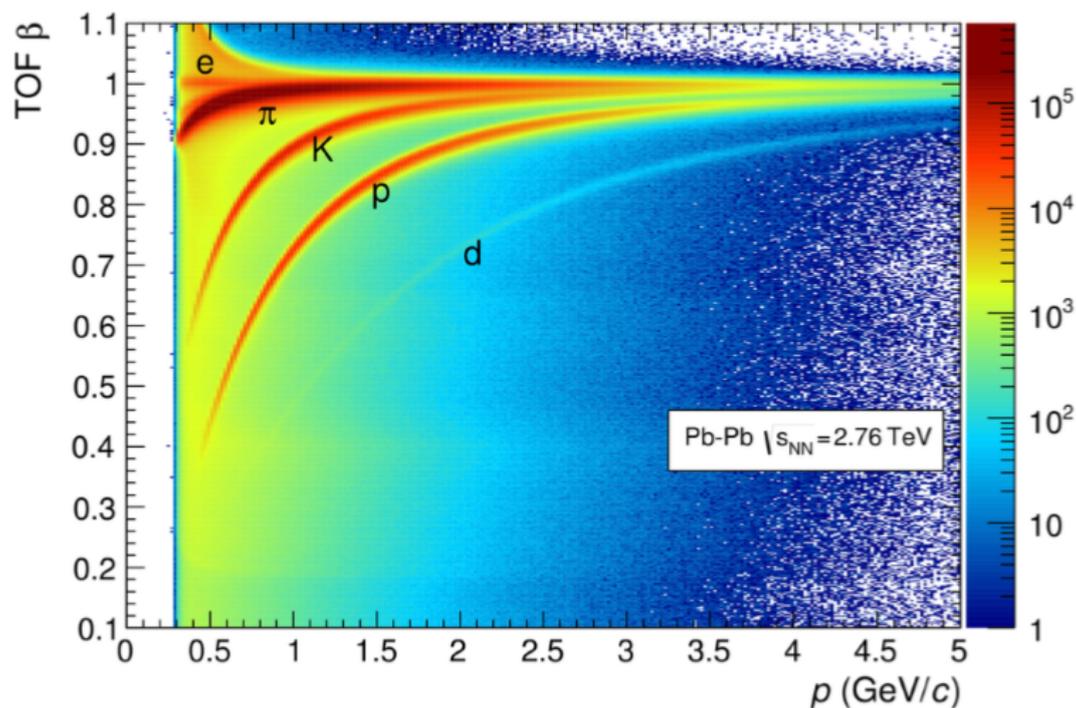


Figura 2.12: Velocità β calcolata a partire dai tempi di volo misurati dal TOF, in funzione dell'impulso per una collisione Pb-Pb con $\sqrt{s_{NN}} = 2.76$ TeV.

$m^2c^4 + p^2c^2$, dove γ è il fattore di Lorentz, è quindi possibile calcolare la massa, che insieme alla carica ci permette di conoscere l'identità della particella considerata. Si ottiene:

$$m = \frac{p}{c} \sqrt{\frac{1}{\beta^2} - 1} = \frac{p}{c} \sqrt{\frac{t^2 c^2}{L^2} - 1} \text{ con } \beta = \frac{v}{c}$$

A questa è associata un'incertezza sperimentale data dalla formula di somma degli errori in quadratura:

$$\frac{\delta m}{m} = \sqrt{\left(\frac{\delta p}{p}\right)^2 + \left(\gamma^2 \frac{\delta L}{L}\right)^2 + \left(\gamma^2 \frac{\delta t}{t}\right)^2}$$

dalla quale è evidente che l'incertezza sulla massa della particella cresce all'aumentare dell'impulso. Sapendo che il tempo di volo delle particelle può essere scritto come:

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

ed espandendo in serie di Taylor per impulsi elevati si ottiene

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

allora è possibile calcolare la differenza di tempi di volo per due particelle aventi lo stesso impulso p e la stessa lunghezza della traiettoria L , ma masse differenti m_1 ed m_2 come

$$\Delta t = t_2 - t_1 = \frac{Lc}{2} \left(\frac{m_2^2 - m_1^2}{p^2} \right)$$

da cui, mediante la risoluzione del tempo di volo δt , si ricava il numero di deviazioni standard, valore corrispondente alla capacità di TOF di separare due particelle differenti, con lo stesso impulso e la stessa lunghezza di traccia. Risulta

$$n_\sigma = \frac{\Delta t}{\delta t}$$

In figura 2.12 è riportata la misura della velocità delle particelle al TOF in funzione dell'impulso in collisioni Pb-Pb a 2.76 TeV per coppia di nucleoni. Il cut-off a sinistra (a bassi impulsi), fino a 300 MeV è dovuto al fatto che, a causa del valore del campo magnetico di ALICE, solo particelle cariche con $p_T > 300$ MeV/c riescono a raggiungere il TOF. Da questo grafico si può vedere che con una risoluzione temporale totale (data dalla risoluzione temporale intrinseca, dall'elettronica e dalle calibrazioni) di 80 ps è possibile soddisfare le richieste di design del rivelatore, ovvero la separazione a 3σ per π/K fino a 2.5 GeV/c e p/K fino a 4 GeV/c [27].

Capitolo 3

Studio di tecniche di machine learning per la ricostruzione di Λ_c

3.1 Introduzione

Lo studio dei quark pesanti rappresenta uno degli strumenti fondamentali per investigare le proprietà del Quark-Gluon Plasma creato nelle collisioni tra nuclei pesanti ultrarelativistici. Grazie alla loro massa elevata, i quark charm sono creati nelle primissime fasi della collisione e, propagandosi all'interno del mezzo, possono interagire con i suoi costituenti durante l'intera evoluzione del sistema, fornendo una misura diretta delle sue proprietà. La produzione di barioni charmati è di particolare interesse perchè sensibile ai processi di adronizzazione nel mezzo. Il barione charmato Λ_c^+ , le cui caratteristiche sono illustrate in tabella 3.1, è stato ricostruito in ALICE attraverso due canali di decadimento adronico e nel canale semileptonico; in particolare in questa tesi è stato preso in considerazione il canale di decadimento $\Lambda_c^+ \rightarrow p + K_s^0$.

| Quark che la compongono | Massa (MeV/c ²) | Tempo di vita (10 ⁻¹⁵ s) |
|-------------------------|-----------------------------|-------------------------------------|
| udc | 2286.46 ± 0.14 | 200 ± 6 |

Tabella 3.1: Caratteristiche principali del barione charmato Λ_c^+ .

Nell'analisi di questo decadimento si riscontrano due particolari difficoltà. Una è legata alla breve vita media della Λ_c^+ , dell'ordine di 2×10^{-13} s (a riposo); essa infatti, prima di decadere, percorre una distanza $c\tau$ di circa 60 μm , minore rispetto alla precisione del rivelatore di microvertice di ALICE, pari a circa 100 μm , rendendo impossibile la distinzione tra i vertici primario e secondario. La seconda è dovuta al basso rapporto che esiste tra segnale e fondo. Per riuscire a studiare questo decadimento può risultare quindi molto utile l'utilizzo di tecniche di analisi multivariata, che permettono di considerare

contemporaneamente ed indipendentemente tutte le variabili esaminate, sfruttando la maggior parte delle informazioni disponibili attraverso tecniche di machine learning. Un'implementazione software delle tecniche di analisi multivariate è già fornita all'interno del pacchetto TMVA [28] distribuito insieme a ROOT [29].

3.2 TMVA

TMVA (Toolkit for Multivariate Data Analysis) è un pacchetto integrato a ROOT che fornisce un ambiente di apprendimento automatico per l'elaborazione e la valutazione di classificazioni multivariate e include numerose tecniche apposite per lo studio della fisica delle alte energie. In particolare in questa tesi vengono analizzati e confrontati tre diversi metodi: i tagli rettangolari, le reti neurali e i Boosted Decision Tree (BDT).

3.3 I tre metodi a confronto

3.3.1 Tagli rettangolari

I tagli rettangolari rappresentano il metodo di classificazione più comune per selezionare gli eventi di segnale, distinguendoli da quelli di fondo. A differenza degli altri metodi essi restituiscono una semplice risposta binaria, 1 per il segnale e 0 per il fondo. La selezione viene effettuata applicando dei valori di massimo e di minimo in maniera indipendente a tutte le quantità (variabili) selezionate come input; esso quindi non rappresenta, a tutti gli effetti, una tecnica di analisi multivariata poichè le informazioni fornite dalle singole variabili non vengono combinate tra loro. I valori dei tagli vengono ottimizzati in automatico dall'algoritmo facendo una scansione completa di tutti i possibili valori di efficienza e massimizzando, per ognuno, la rimozione del fondo.

3.3.2 Neural Networks di tipo Multilayer Perceptron (MLP)

Le reti neurali sono un metodo di analisi basato su un'analogia con il funzionamento del cervello umano. I neuroni, che corrispondono ai nodi, vengono modellizzati come funzioni di base che si attivano o non si attivano a seconda che venga superato, oppure no, un livello di soglia. Queste funzioni vengono sommate attraverso una combinazione lineare pesata sui pesi delle variabili di input. La funzione risultante viene mandata al livello successivo mediante dei collegamenti, corrispondenti alle sinapsi, il cui spessore è proporzionale al peso a loro attribuito. La rete dà quindi come output una funzione del tipo

$$y_j = y\left(\sum_{i=0}^d w_{ji}x_i\right)$$

dove x_i sono gli input provenienti dal livello precedente, con pesi w_{ji} , e d è il numero di tali input. La rete può essere formata, oltre che dai livelli di input e output, anche da livelli intermedi nascosti: più questi sono numerosi e più l'output sarà preciso. Le MLP sono un particolare tipo di reti neurali in cui l'output di ciascun nodo di un livello viene usato solo come input di uno o più nodi del livello successivo, ma non può tornare a quelli precedenti creando un loop. Nel livello di input troviamo inoltre un particolare nodo aggiuntivo, il nodo di bias, settato sempre a 1, necessario nel caso in cui tutte le variabili di input risultino 0; in tal caso infatti il valore passato al livello successivo, calcolato come la media pesata dei loro valori risulterebbe sempre uguale a 0.

3.3.3 Boosted Decision Tree (BDT)

Un decision tree è costituito da una sequenza di decisioni di tipo binario ("yes/no") applicate a tutte le variabili di input che permettono, partendo da un campione di eventi iniziale (root node), di selezionare regioni con purezza crescente. Questo metodo può essere assimilato ai tagli rettangolari: qui però ogni taglio dipende da quello precedente. Un singolo decision tree è generalmente poco stabile (piccole variazioni del campione di training possono provocare grandi effetti sulla risposta) e soffre di overtraining, ovvero di una forte dipendenza dalle caratteristiche del campione iniziale. Per aumentare la stabilità del metodo si utilizzano tecniche di boosting. In pratica, si realizzano diversi tree, che formano quella che viene chiamata una foresta, derivati dallo stesso campione di training, assegnando un peso maggiore agli eventi che non sono stati identificati correttamente, come segnale o fondo, nel tree precedente. L'output finale è una combinazione di tutti i classificatori costruiti con la tecnica del boosting e vale

$$y = \sum_{i=0}^d w_i C_i$$

dove w_i sono i pesi e d è il numero di decision trees.

3.4 Training del TMVA

Per istruire gli algoritmi di machine learning è necessario fornire loro dei campioni statistici composti da eventi appartenenti ad una soltanto delle due categorie in cui si vogliono classificare, ovvero nel nostro caso segnale e fondo. Dal confronto tra alcune grandezze selezionate in entrambi i campioni (variabili di input), gli algoritmi apprendono le regole di classificazione che permettono poi di distinguere eventi di origine incognita in una delle due categorie.

3.4.1 Simulazione Monte Carlo

In questa tesi entrambi i campioni statistici provengono da un'apposita simulazione Monte Carlo (MC), realizzata con il framework di analisi AliRoot [33], in cui gli eventi sono generati utilizzando l'event generator HIJING [30], per descrivere le caratteristiche globali della collisione Pb-Pb, e in cui, evento per evento, è stata aggiunta una collisione protone-protone generata con PYTHIA [31], per la quale viene richiesta la presenza di almeno una particella Λ_c^+ , che decade attraverso il decadimento adronico che vogliamo studiare (pK_s^0) tra i prodotti finali. Questo accorgimento permette di massimizzare la statistica disponibile per il training dei metodi di machine learning e di minimizzare i tempi di procedimento poichè nessun evento è in teoria rigettato in fase di selezione. Le particelle così generate sono poi trasportate attraverso i rivelatori che compongono l'esperimento ALICE utilizzando il software GEANT3 [32], in cui vengono simulate anche le risposte dei singoli rivelatori. Le versioni dei software utilizzati sono riportate in tabella 3.2.

| Software | Versione utilizzata |
|----------|---------------------|
| ROOT | v5-34-30-alice10-12 |
| AliRoot | v5-09-44d-1 |
| GEANT3 | v2-5-10 |
| HIJING | v1.36-10 |
| PYTHIA | v6.425 |

Tabella 3.2: Versioni dei software utilizzati per il Monte Carlo.

3.4.2 Variabili di input

I metodi di machine learning sono stati fatti girare utilizzando 9 variabili di input (riportiamo il nome delle variabili come compaiono all'interno del codice utilizzato):

massK0: è la massa invariante della particella neutra K_s^0 , ricostruita mediante le tracce delle particelle figlie. La K_s^0 decade in π^+ e π^- , per cui quando si osservano due particelle di carica opposta provenienti da un vertice comune, si misurano i loro impulsi e tramite questi si ricava quello della particella che le ha generate; assumendo che la loro massa sia quella di un mesone π , pari a circa $138 \text{ MeV}/c^2$, si calcola l'energia come $E_i = \sqrt{m_i \cdot m_i + p_i \cdot p_i}$ e la massa invariante risulta essere $m_{inv} = \sqrt{(E_1 + E_2)^2 - (p_0 \cdot p_0)}$

tImpParBach: è il parametro di impatto del protone, definito come la distanza minima tra la traccia ricostruita del protone e la posizione del vertice primario; nel caso

di vere Λ_c^+ , ci si aspetta che esso abbia un valore piccolo, perché, dato il basso valore di $c\tau$, è come se il protone provenisse dal vertice primario.

tImpParV0: è il parametro di impatto della K_s^0 , definito come la distanza minima tra la traccia ricostruita della K_s^0 e la posizione del vertice primario.

bachelorPt: è l'impulso trasverso del protone, misurato attraverso il raggio di curvatura della sua traiettoria.

combinedProtonProb: è la probabilità che la traccia carica associata alla K_s^0 per formare una candidata Λ_c^+ sia effettivamente un protone. Il valore 1 della variabile corrisponde ad una probabilità del 100% che quella particella sia un protone, il valore 0 corrisponde ad una probabilità nulla.

ctK0S: è la vita media della K_s^0 moltiplicata per c : $c\tau$. Essa corrisponde alla distanza tra il vertice primario e quello da cui partono i due mesoni π , moltiplicata per la massa di K_s^0 e divisa per il suo impulso.

cosPaK0S: è il coseno dell'angolo di pointing, l'angolo compreso tra la direzione della traccia ricostruita della K_s^0 e la retta passante per il vertice primario e quello di decadimento della K_s^0 . Tale valore deve essere prossimo a 1.

CosThetaStar: è il coseno dell'angolo compreso tra la direzione di emissione del protone nel sistema di riferimento in cui la Λ_c^+ è a riposo e la direzione della Λ_c^+ nel sistema di riferimento del laboratorio. Per le vere Λ_c^+ ci si aspetta una prevalenza di valori negativi, mentre per quelle false ci si aspetta che tutti i valori siano equiprobabili.

signd0: è il parametro di impatto del protone che tiene conto del segno: un valore positivo corrisponde con probabilità maggiore ad un protone creato dal decadimento di una Λ_c^+ , mentre un valore negativo indica una provenienza del protone da un punto non compatibile con il vertice primario.

Si è scelto di analizzare 6 intervalli di impulso trasverso (p_T) della Λ_c^+ : [2,4], [4,6], [6,8], [8,12], [12,16], [16,25] GeV/c; i training sono stati ripetuti in modo indipendente sui diversi intervalli. La distribuzione del segnale sovrapposta a quella del fondo per le 9 variabili di input per ciascuno degli intervalli si può osservare dai grafici nelle figure 3.1-3.6.

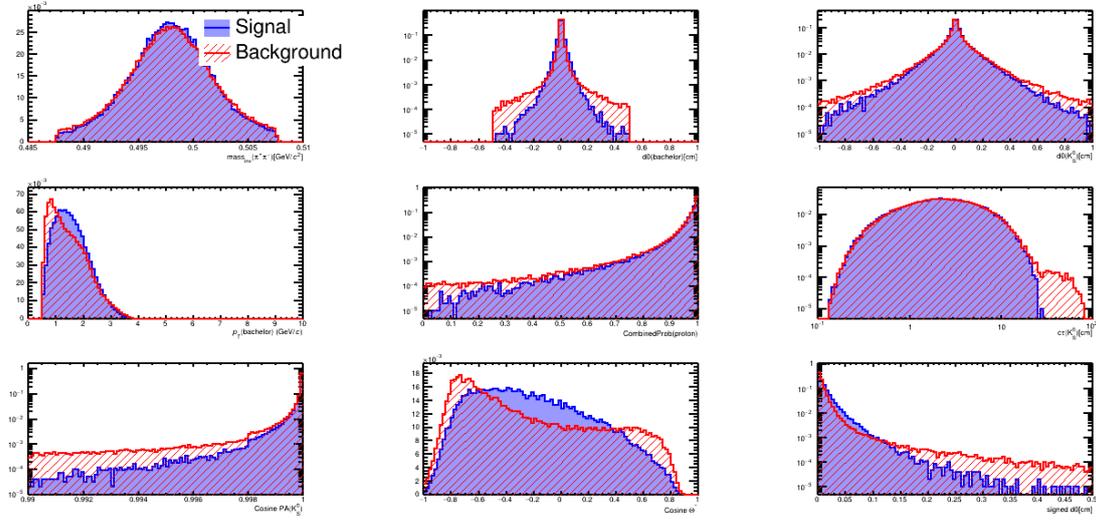


Figura 3.1: Distribuzione delle 9 variabili di input nell'intervallo di impulso $[2,4]$ GeV/c. La curva blu corrisponde all'andamento del segnale, mentre in rosso è riportato il fondo.

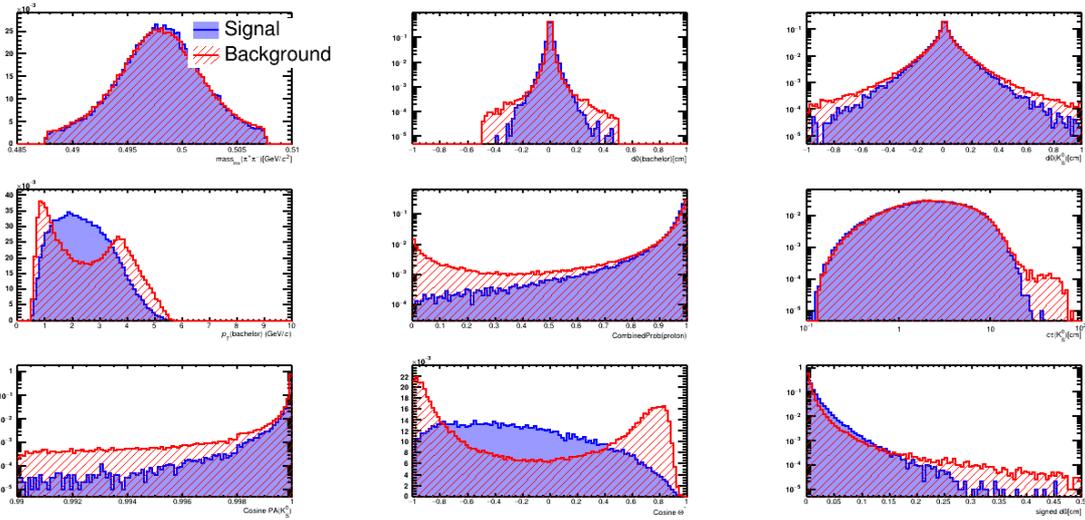


Figura 3.2: Distribuzione delle 9 variabili di input nell'intervallo di impulso $[4,6]$ GeV/c. La curva blu corrisponde all'andamento del segnale, mentre in rosso è riportato il fondo.

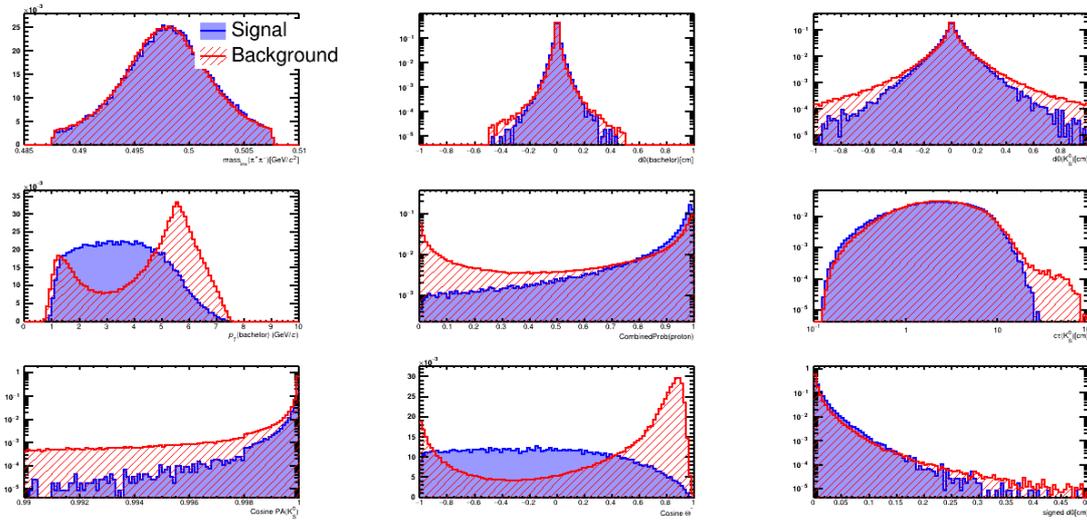


Figura 3.3: Distribuzione delle 9 variabili di input nell'intervallo di impulso [6,8] GeV/c. La curva blu corrisponde all'andamento del segnale, mentre in rosso è riportato il fondo.

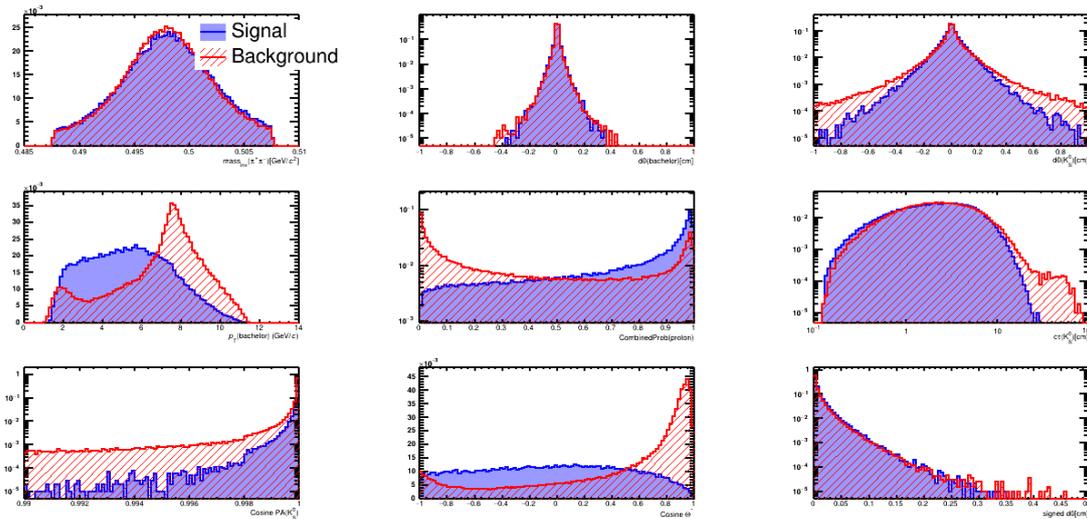


Figura 3.4: Distribuzione delle 9 variabili di input nell'intervallo di impulso [8,12] GeV/c. La curva blu corrisponde all'andamento del segnale, mentre in rosso è riportato il fondo.

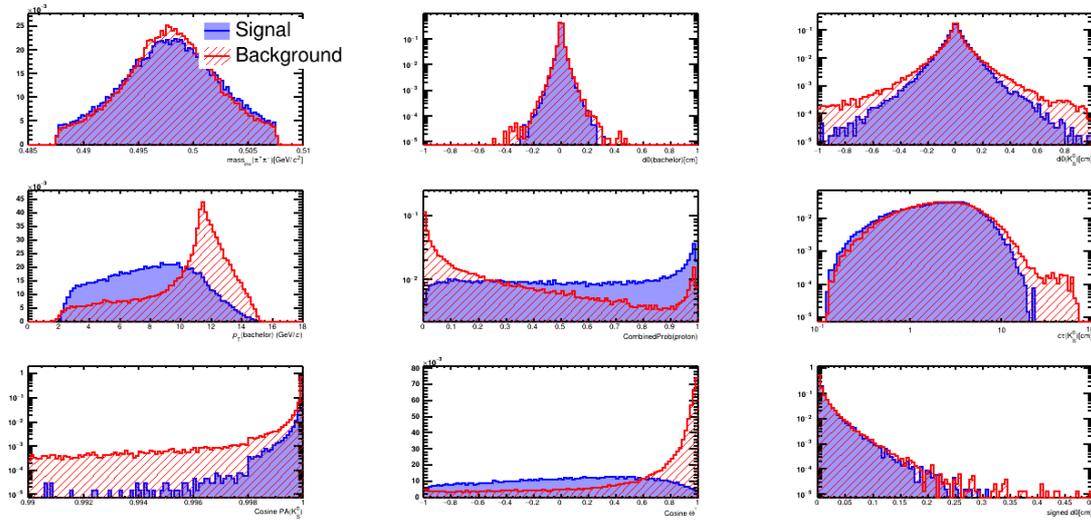


Figura 3.5: Distribuzione delle 9 variabili di input nell'intervallo di impulso [12,16] GeV/c. La curva blu corrisponde all'andamento del segnale, mentre in rosso è riportato il fondo.

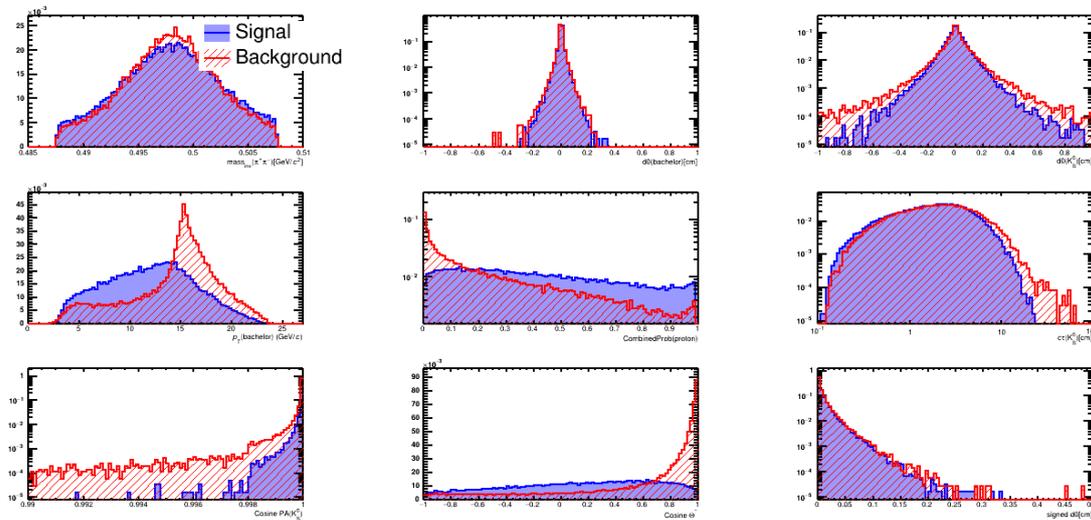


Figura 3.6: Distribuzione delle 9 variabili di input nell'intervallo di impulso [16,25] GeV/c. La curva blu corrisponde all'andamento del segnale, mentre in rosso è riportato il fondo.

Affinchè le variabili scelte costituiscano un buon set, è necessario che ci sia una bassa correlazione tra le une e le altre. Due variabili risultano correlate quando esiste una relazione che le lega. Tale relazione è quantificata da un coefficiente che vale zero nel

caso in cui le variabili non siano correlate e +1 o -1, rispettivamente se sono correlate positivamente, cioè all'aumentare della prima, aumenta anche la seconda, o negativamente, cioè all'aumentare della prima, la seconda diminuisce. La matrice di correlazione dell'intervallo [4-6] GeV/c è riportata nella figura 3.7 e risulta molto simile in tutti gli altri intervalli. Come si può osservare, sia per il segnale che per il background, è presente un'alta correlazione tra le variabili CosThetaStar e bachelorPt.

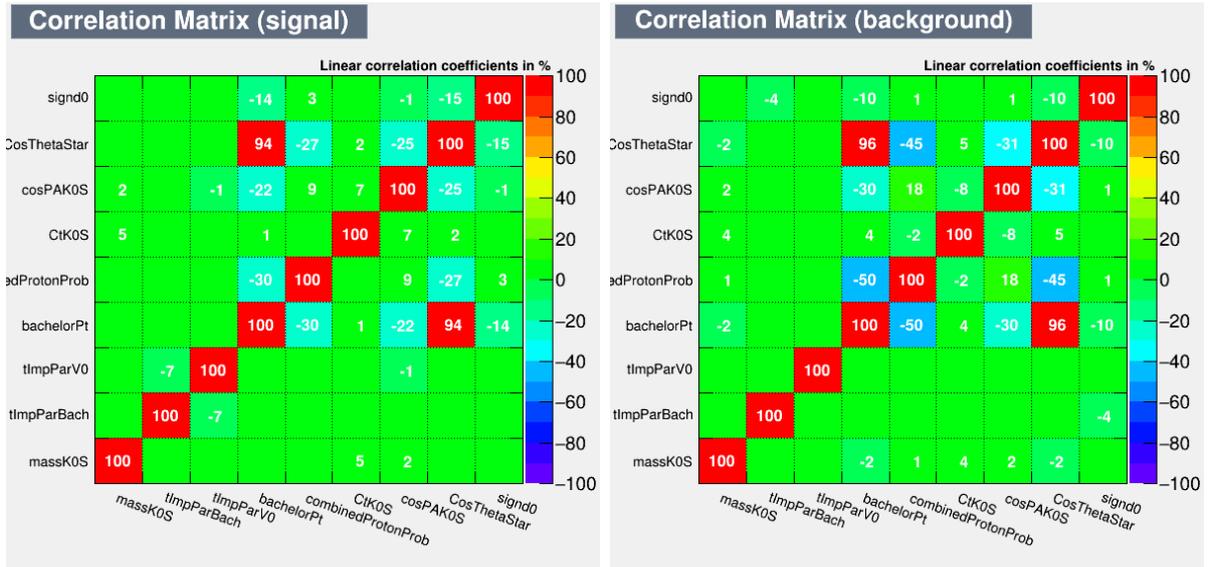


Figura 3.7: Coefficienti di correlazione lineare delle variabili di input per il segnale (a sinistra) e il fondo (a destra) per le candidate Λ_c^+ aventi un impulso trasverso $4 \leq p_T \leq 6$ GeV/c. Come si osserva, le uniche variabili con alta correlazione risultano essere CosThetaStar e bachelorPt.

3.5 Configurazione dei tre metodi

Il TMVA è stato configurato per utilizzare gli eventi sia di segnale che di fondo provenienti da una simulazione Monte Carlo, riservandone metà per fare il training dei tre metodi e metà per testarli. È stato inoltre richiesto che, per ogni intervallo, sia per il training che per il testing, venissero effettivamente utilizzati 100000 eventi per il segnale e 200000 per il fondo, quando presenti, o in caso contrario, l'intera quantità di eventi presenti nel range di impulsi dell'intervallo (tabella 3.5). Nelle tabelle 3.3 e 3.4 sono riportati il numero degli eventi totali utilizzati e il numero di eventi selezionati in ogni intervallo.

| | | | |
|---------------|---------|------------|---------------|
| | Segnale | Background | Totale eventi |
| Numero eventi | 1725004 | 20048373 | 21773376 |

Tabella 3.3: Somma degli eventi di training e di testing che corrispondono ai requisiti richiesti per il segnale e per il fondo.

| Λ_c^+ p_T bin (GeV/c) | Eventi totali segnale | Eventi totali background |
|---------------------------------|-----------------------|--------------------------|
| $2 \leq p_T \leq 4$ | 581348 | 14244738 |
| $4 \leq p_T \leq 6$ | 402615 | 3804226 |
| $6 \leq p_T \leq 8$ | 231356 | 1128900 |
| $8 \leq p_T \leq 12$ | 255373 | 642964 |
| $12 \leq p_T \leq 16$ | 133302 | 156678 |
| $16 \leq p_T \leq 25$ | 121010 | 70866 |

Tabella 3.4: Somma degli eventi di training e di testing appartenenti ad ogni intervallo di impulso trasverso, per il segnale e per il fondo.

| Λ_c^+ p_T bin (GeV/c) | Segnale-training events (=testing events) | Background-training events (=testing events) |
|---------------------------------|--|---|
| $2 \leq p_T \leq 4$ | 100000 | 200000 |
| $4 \leq p_T \leq 6$ | 100000 | 200000 |
| $6 \leq p_T \leq 8$ | 100000 | 200000 |
| $8 \leq p_T \leq 12$ | 100000 | 200000 |
| $12 \leq p_T \leq 16$ | 66651 | 78339 |
| $16 \leq p_T \leq 25$ | 60505 | 35433 |

Tabella 3.5: Numero di eventi effettivamente utilizzati per il training, uguale a quello per il testing, in ciascuno dei 6 intervalli, corrispondente al minimo tra il numero di eventi presenti nell'intervallo e 100000 per il segnale o 200000 per il fondo.

3.5.1 Tagli rettangolari

Per i tagli rettangolari è stato utilizzato il metodo di minimizzazione MC, che consiste nello scegliere in maniera casuale i valori dei tagli, per poi selezionare quelli migliori, ovvero quelli che minimizzano l'estimatore, massimizzando la background rejection per ogni data efficienza di segnale. Il metodo utilizzato per calcolare l'efficienza di un determinato taglio è il metodo EffSel, ottimale dal momento che sono presenti variabili

correlate. L'efficienza viene quindi calcolata facendo il rapporto tra il numero di eventi che vengono selezionati dal taglio e il numero totale di eventi. L'opzione FSmart, invece abilita il TMVA a scegliere automaticamente quali tagli non sono necessari e rimuoverli. La configurazione utilizzata per i tagli rettangolari è riassunta in tabella 3.6.

| Parametri | Valori | Descrizione |
|-----------|--------|---|
| FitMethod | MC | Metodo di minimizzazione utilizzato per effettuare il fit |
| EffMethod | EffSel | Metodo di selezione utilizzato per calcolare l'efficienza di un taglio su una variabile |
| VarProp | FSmart | Categorizzazione dei tagli |

Tabella 3.6: Configurazione utilizzata per i tagli rettangolari.

3.5.2 Reti neurali

La configurazione utilizzata per le reti neurali è riassunta in tabella 3.7. La funzione di attivazione utilizzata è la tangente iperbolica. Essa restituisce un peso compreso tra 1 e -1, rispettivamente quando in ingresso è presente un valore alto oppure basso. Ha la particolarità di essere una funzione continua e permette quindi di avere degli output anche intermedi ai valori di massimo e di minimo, una caratteristica fondamentale per l'algoritmo di apprendimento delle reti neurali multilivello, come quello utilizzato. In particolare l'input layer contiene $N=10$ neuroni che corrispondono alle 9 variabili di ingresso e al bias, mentre gli hidden layers sono stati configurati con un'architettura tale da avere $N+5=15$ neuroni nel primo livello, $N+5-1=14$ nel secondo e così via, fino ad arrivare ad un unico neurone di output. Per il training sono stati effettuati 600 cicli, sufficienti ad avere un output finale molto vicino a quello richiesto, come si osserva in figura 3.12.

| Parametri | Valori | Descrizione |
|--------------|--------|---|
| NeuronType | tanh | Tipo di funzione di attivazione dei neuroni |
| VarTransform | N | Trasformazione delle variabili |
| NCycles | 600 | Numero di cicli utilizzati per il training |
| HiddenLayers | $N+5$ | Specificazione dell'architettura dei livelli nascosti |
| TestRate | 5 | Test per l'overtraining eseguito ad ogni epoca |

Tabella 3.7: Configurazione utilizzata per le reti neurali.

3.5.3 Boosted Decision Tree

La configurazione per il metodo BDT è riportata in tabella 3.8. In particolare vengono creati 850 trees, con una profondità massima pari a 3. Il tipo di boosting utilizzato, è l'AdaBoost basato su una loss-function di tipo esponenziale $L(F, y) = e^{-F(x) \cdot y}$; essa misura la deviazione tra la risposta del modello $F(x)$ e il valore vero y . Il procedimento di boosting consiste in una variazione dei parametri, con lo scopo di ottenere una deviazione minima. Infine, il criterio di separazione che viene utilizzato nei nodi è il Gini Index: l'indice di Gini misura il grado di probabilità che una certa variabile sia classificata in modo errato quando presa in modo casuale. Il suo valore varia tra 0 e 1, dove 0 denota che tutti gli elementi appartengono ad una certa classe e 1 che essi sono invece distribuiti in modo casuale tra le classi. La formula per l'indice di Gini è

$$\text{Gini Index} = p(1 - p)$$

dove p è la purity del taglio. Siccome p può valere al massimo 0.5, l'indice di Gini varia tra 0 e 0.25.

| Parametri | Valori | Descrizione |
|----------------------|------------|---|
| NTrees | 850 | Numero di trees |
| MinNodeSize | 2.5% | Percentuale minima di training events richiesti in un leaf node |
| MaxDepth | 3 | Massima profondità dei trees consentita |
| BoostType | AdaBoost | Tipo di boosting utilizzato |
| AdaBoostBeta | 0.5 | Frequenza di apprendimento per l'algoritmo AdaBoost |
| BaggedSampleFraction | 0.5 | Dimensione del campione di bagged events relativa a quella del campione di dati originale |
| SeparationType | Gini Index | Criterio di separazione utilizzato per il node splitting |
| nCuts | 20 | Numero di punti dell'intervallo per avere il migliore taglio da utilizzare nel node splitting |

Tabella 3.8: Configurazione utilizzata per il BDT.

3.6 Ranking delle variabili

Fare un ranking delle variabili di input significa ordinarle con un criterio di importanza e peso che esse hanno all'interno del processo di training. Per esempio, nel metodo BDT il ranking delle variabili si ottiene contando quanto spesso una determinata variabile viene utilizzata come criterio di splitting per i nodi, pesando ogni splitting sia in base al

grado di separazione ottenuto sia in base agli eventi presenti nel nodo. Le reti neurali effettuano il ranking delle variabili utilizzando la somma del quadrato dei pesi attribuiti alle connessioni tra i neuroni del livello di input e del primo hidden layer. Per il metodo dei tegli rettangolari non è previsto un ranking delle variabili. I ranking negli intervalli a maggiore e minore statistica, per i BDT e le MLP sono riportati nelle tabelle 3.9-3.12; come si può notare essi non dipendono solo dal metodo utilizzato, ma anche dall'intervallo che si sta considerando. Oltre alla differenza netta tra i ranking effettuati dai due metodi, con variabili che si trovano negli ultimi posti per le MLP, ma nei primi per il BDT, è interessante notare che per le MLP la classifica delle variabili nell'intervallo a minore e maggiore statistica è quasi uguale, mentre per i BDT risulta essere molto diversa.

| Rank | Variable | Importance |
|------|--------------------|------------|
| 1 | signd0 | 3,67E+05 |
| 2 | cosPAK0S | 1,53E+04 |
| 3 | combinedProtonProb | 1,10E+04 |
| 4 | CosThetaStar | 3,30E+03 |
| 5 | CtK0S | 2,89E+03 |
| 6 | bachelorPt | 1,47E+03 |
| 7 | tImpParBach | 1,04E+01 |
| 8 | tImpParV0 | 2,97E+00 |
| 9 | massK0S | 2,92E+00 |

Tabella 3.9: Ranking delle variabili di input per il metodo delle reti neurali, nell'intervallo di p_T a maggiore statistica $[2,4]$ GeV/c.

| Rank | Variable | Importance |
|------|--------------------|------------|
| 1 | cosPAK0S | 1,72E+04 |
| 2 | combinedProtonProb | 1,62E+04 |
| 3 | signd0 | 1,35E+04 |
| 4 | CosThetaStar | 9,21E+03 |
| 5 | CtK0S | 4,32E+03 |
| 6 | bachelorPt | 4,07E+03 |
| 7 | tImpParBach | 2,61E+02 |
| 8 | tImpParV0 | 5,94E+01 |
| 9 | massK0S | 3,53E+00 |

Tabella 3.10: Ranking delle variabili di input per il metodo delle reti neurali, nell'intervallo di p_T a minore statistica $[16,25]$ GeV/c.

| Rank | Variable | Importance |
|------|--------------------|------------|
| 1 | bachelorPt | 1,60E+02 |
| 2 | CosThetaStar | 1,47E+02 |
| 3 | signd0 | 1,22E+02 |
| 4 | cosPAK0S | 1,20E+02 |
| 5 | massK0S | 1,18E+02 |
| 6 | combinedProtonProb | 9,15E+01 |
| 7 | tImpParV0 | 9,14E+01 |
| 8 | tImpParBach | 7,65E+01 |
| 9 | CtK0S | 7,48E+01 |

Tabella 3.11: Ranking delle variabili di input per il metodo BDT, nell'intervallo di p_T a maggiore statistica [2,4] GeV/c.

| Rank | Variable | Importance |
|------|--------------------|------------|
| 1 | combinedProtonProb | 1,85E+02 |
| 2 | CosThetaStar | 1,85E+02 |
| 3 | bachelorPt | 1,31E+02 |
| 4 | massK0S | 1,22E+02 |
| 5 | tImpParV0 | 9,28E+01 |
| 6 | tImpParBach | 8,09E+01 |
| 7 | CtK0S | 8,04E+01 |
| 8 | signd0 | 7,30E+01 |
| 9 | cosPAK0S | 4,99E+01 |

Tabella 3.12: Ranking delle variabili di input per il metodo BDT, nell'intervallo di p_T a minore statistica [16,25] GeV/c.

3.7 Plot di controllo

Il pacchetto TMVA fornisce una serie di grafici di controllo che permettono di verificare la correttezza e la performance del processo di training. Nelle figure 3.8, 3.9 e 3.10 si osserva per ciascun metodo l'efficienza del segnale e del background rispetto ad ogni possibile valore di taglio. Nei grafici é riportata inoltre la significanza (significance); uno dei criteri di ottimizzazione consiste infatti nel massimizzare il valore della significanza $S/\sqrt{S+B}$, con S e B rispettivamente il numero di eventi del segnale e del fondo. Il punto di massimo di tale curva corrisponde al working point, ovvero al punto in cui si

ha la performance migliore.

Sono poi riportati i layout delle reti neurali in figura 3.11, dove si osservano l'input layer, composto da 9 nodi delle variabili di ingresso e dal bias node, l'output layer, in cui si trova solo il nodo di uscita, e il primo hidden layer, che è quello utilizzato per creare il ranking delle variabili; gli altri livelli nascosti, che rendono l'output sempre più preciso, non sono riportati in figura. Nel grafico sono rappresentati i collegamenti tra i nodi dei vari livelli, ciascuno con uno spessore ed un colore proporzionale al peso che ha il nodo di partenza nel determinare la funzione corrispondente al nodo del livello successivo. Come si osserva, i collegamenti più significativi sono quelli che corrispondono alle variabili di input che nei ranking hanno maggiore peso. Correlati a questi grafici ci sono quelli relativi ai test di convergenza in figura 3.12 che rappresentano l'errore sull'output trovato, rispetto a quello desiderato, al passare delle varie epoche; ogni bin rappresenta 5 epoche. Come si osserva l'errore diminuisce sempre di più all'aumentare delle epoche, arrivando quasi a stabilizzarsi già intorno alla quattrocentesima. Da questo grafico si vede anche che non è avvenuto un overtraining del metodo, che sarebbe stato rappresentato da una risalita dell'errore sulle ultime epoche.

Infine, per le reti neurali e per i BDT, sono state riportate, in figura 3.13 e 3.14, le distribuzioni sovrapposte delle funzioni di risposta per segnale e fondo, ottenute applicando i metodi ai campioni di training e di testing separatamente. Siccome i due campioni sono completamente indipendenti, il fatto che le distribuzioni normalizzate siano perfettamente compatibili dimostra che i risultati non sono affetti da overtraining, ovvero il campione statistico utilizzato per il training è sufficientemente grande, in rapporto ai parametri degli algoritmi e al numero di variabili in input, da garantire l'assenza di bias nel risultato finale.

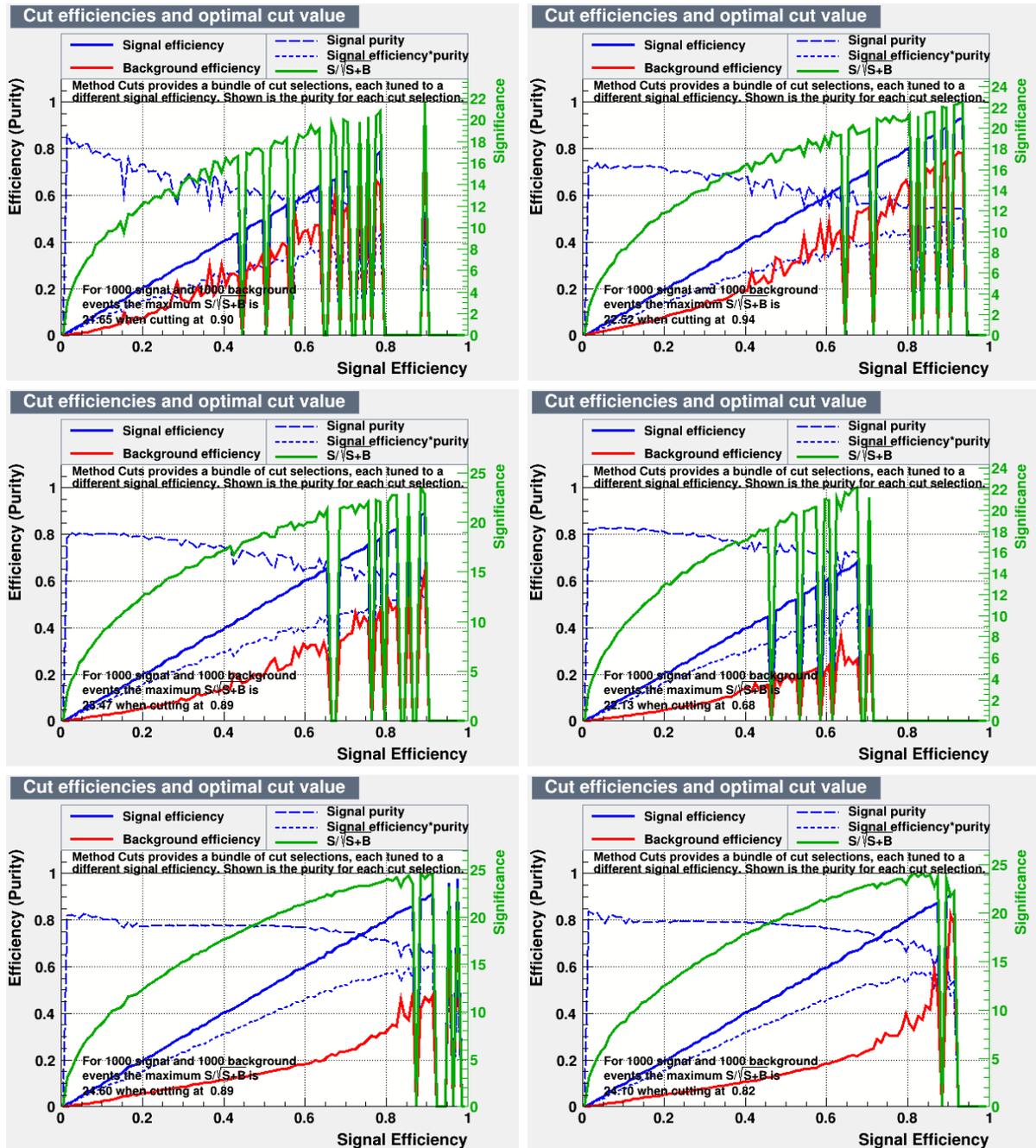


Figura 3.8: Efficienza del segnale e del fondo rispetto ad ogni possibile valore di taglio, in ciascun intervallo di impulso trasverso, per il metodo dei tagli rettangolari. In verde è rappresentata la curva della significanza.

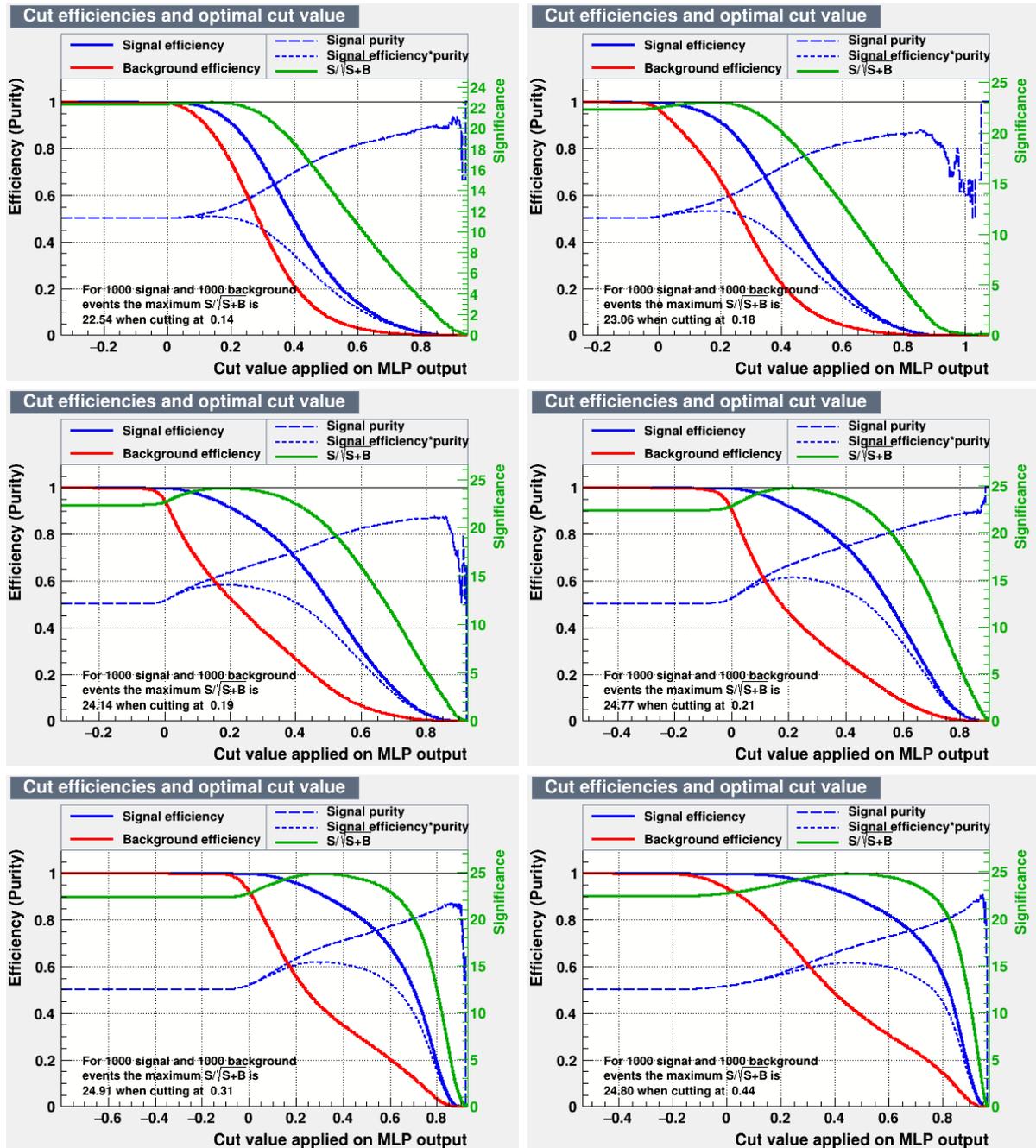


Figura 3.9: Efficienza del segnale e del fondo rispetto ad ogni possibile valore di taglio, in ciascun intervallo di impulso trasverso, per le reti neurali. In verde è rappresentata la curva della significanza, il cui massimo corrisponde al working point.

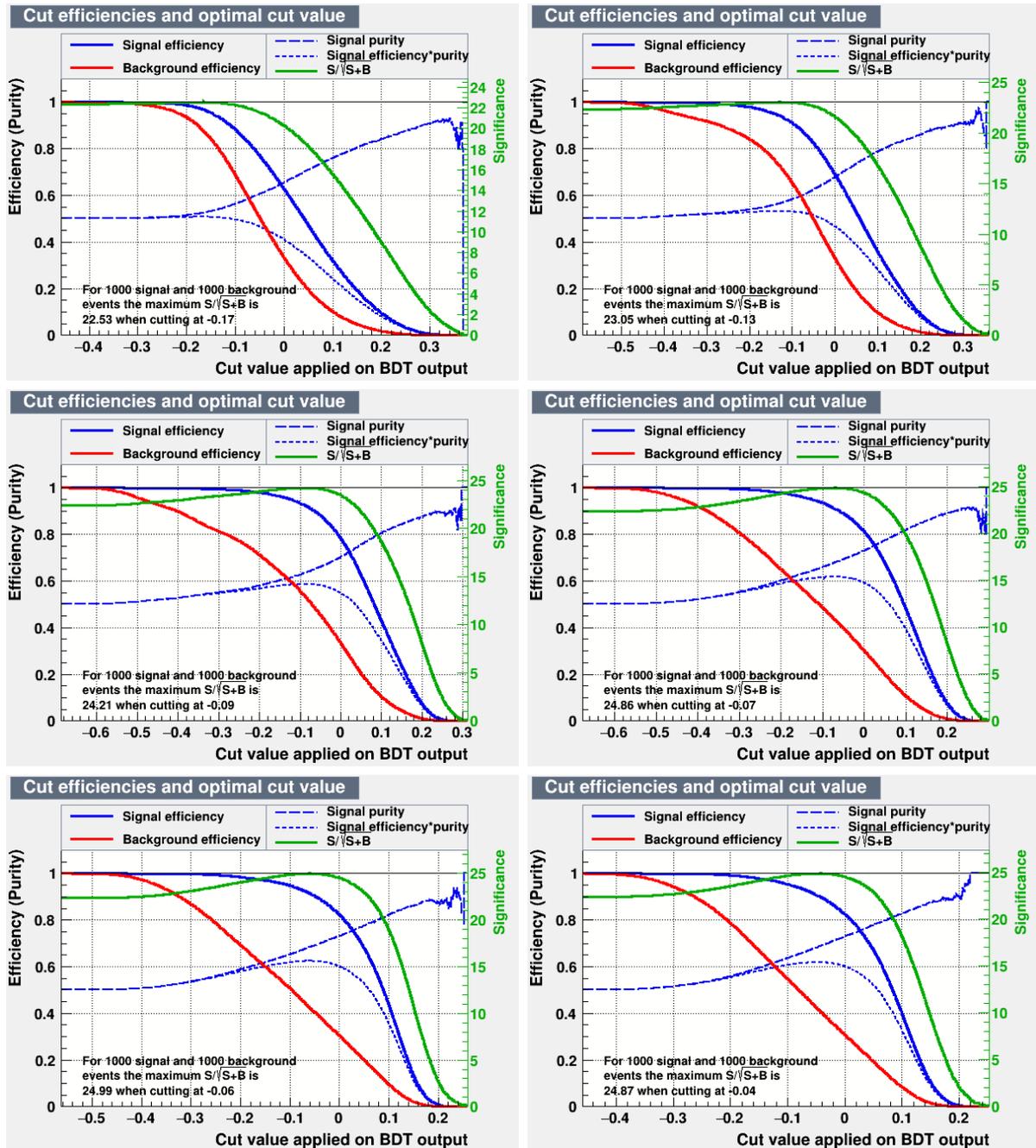


Figura 3.10: Efficienza del segnale e del fondo rispetto ad ogni possibile valore di taglio, in ciascun intervallo di impulso trasverso, per il metodo BDT. In verde è rappresentata la curva della significanza, il cui massimo corrisponde al working point.

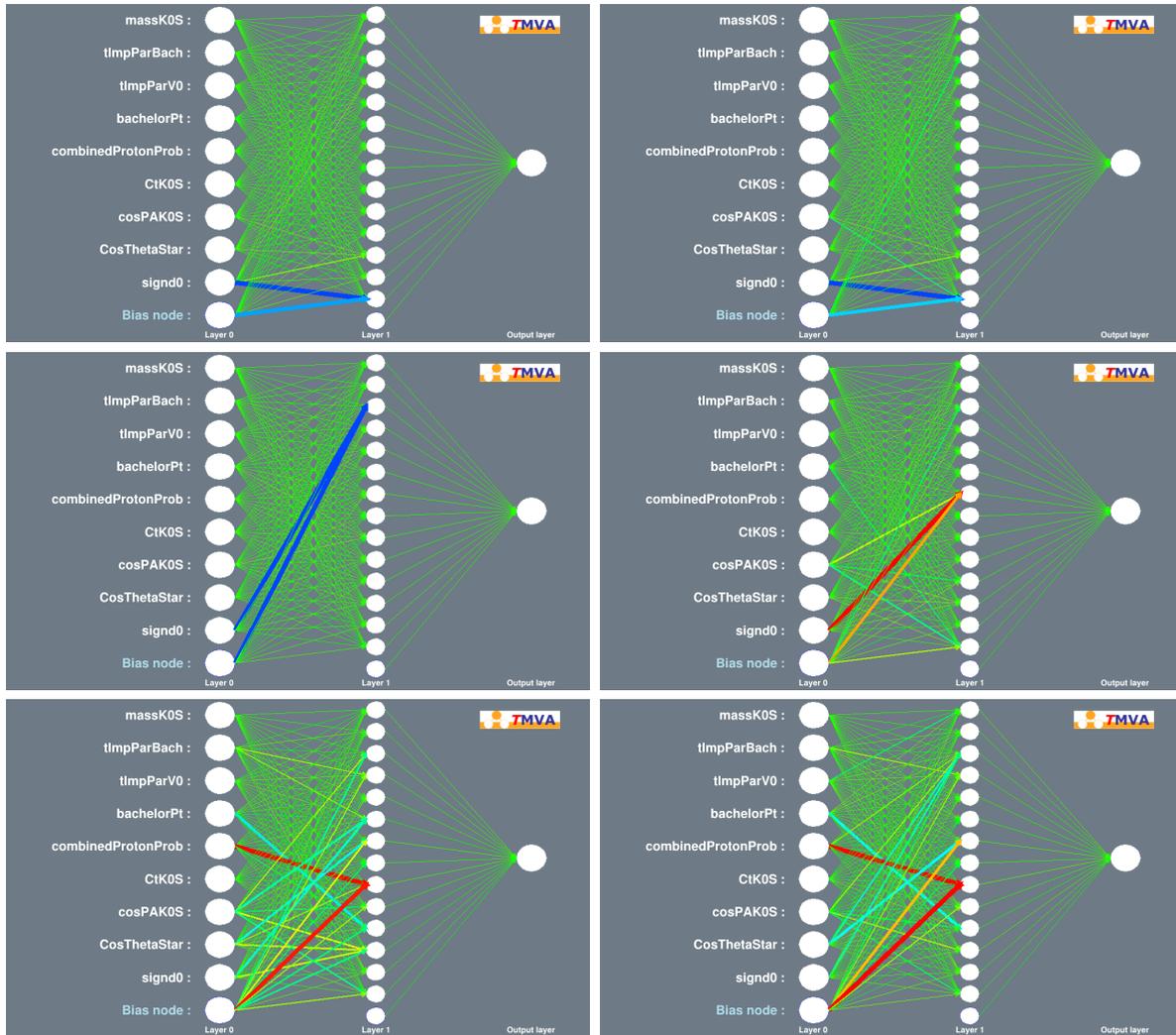


Figura 3.11: Layout delle reti neurali per i 6 intervalli di impulso trasverso. In ogni grafico si osservano l'input layer, composto da 9 nodi corrispondenti alle variabili di ingresso e dal bias node, l'output layer, con un unico nodo di uscita e il primo hidden layer. I collegamenti tra i nodi dei diversi livelli sono rappresentati con un colore e uno spessore proporzionale al peso che il nodo di partenza ha su quello di arrivo.

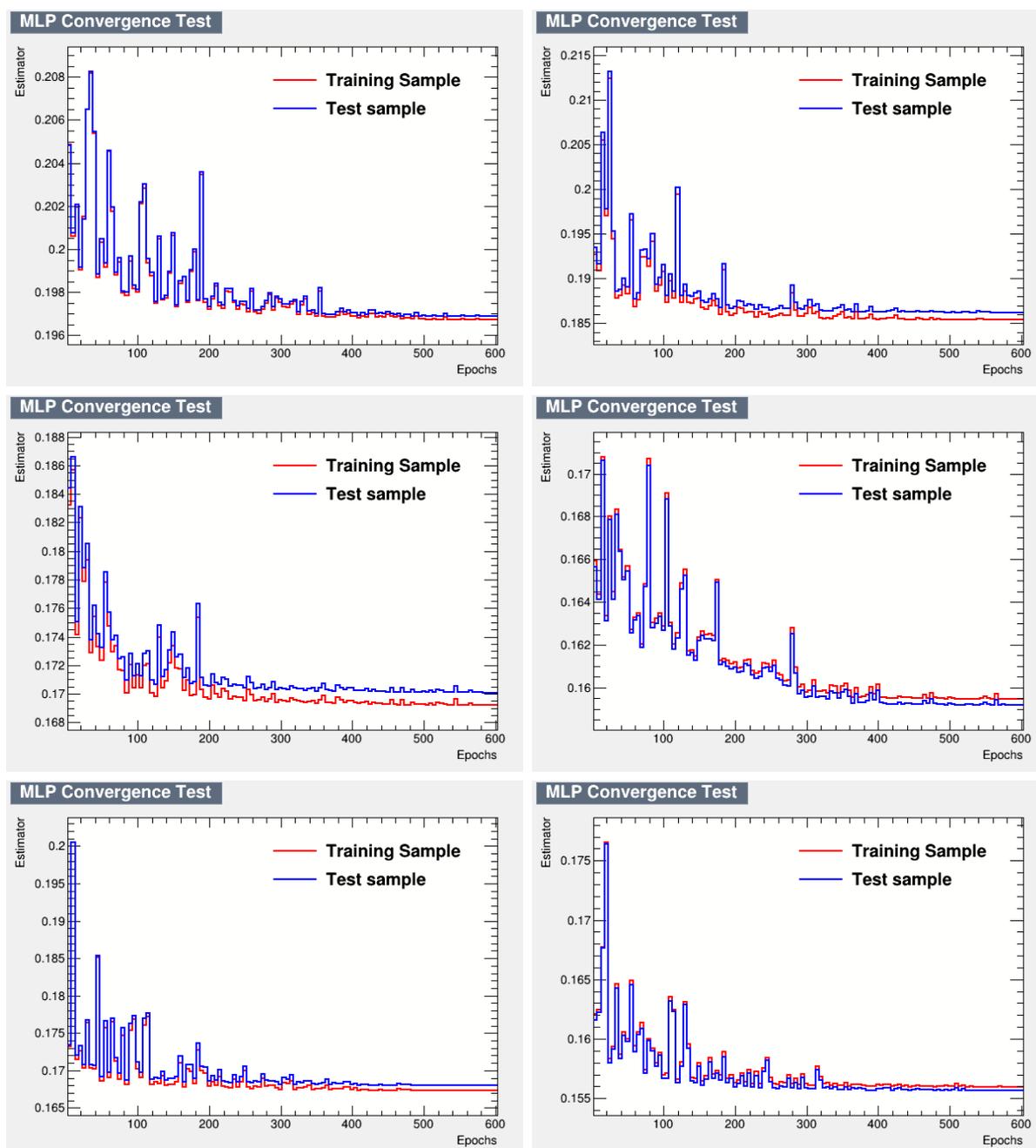


Figura 3.12: Test di convergenza per le reti neurali. Essi rappresentano, per ogni intervallo, l'errore sull'output trovato, al passare delle epoche; ogni bin corrisponde a 5 epoche.

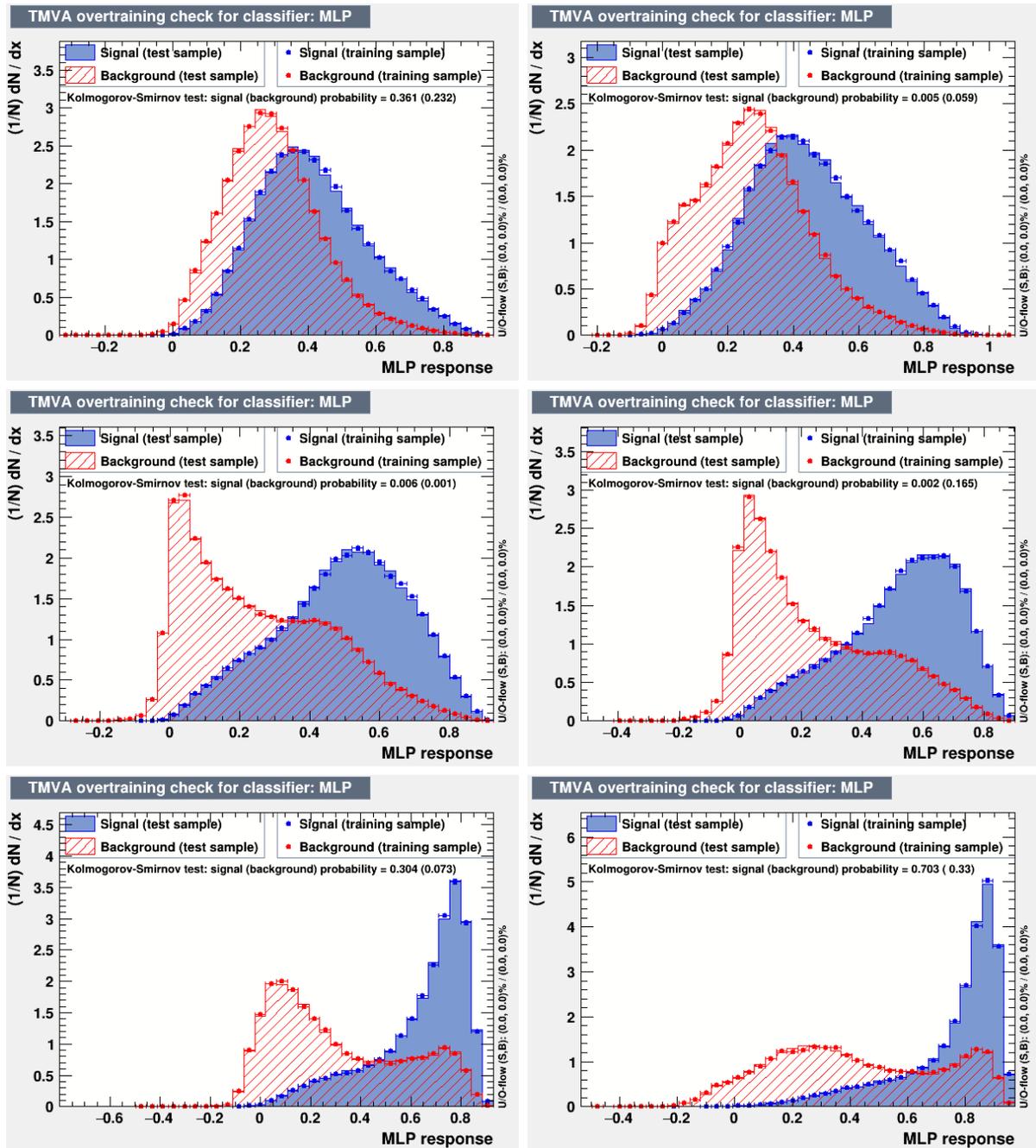


Figura 3.13: Distribuzioni sovrapposte delle funzioni di risposta per segnale e fondo, ottenute applicando il metodo MLP ai campioni di training e di testing separatamente. I risultati non risultano affetti da overtraining.

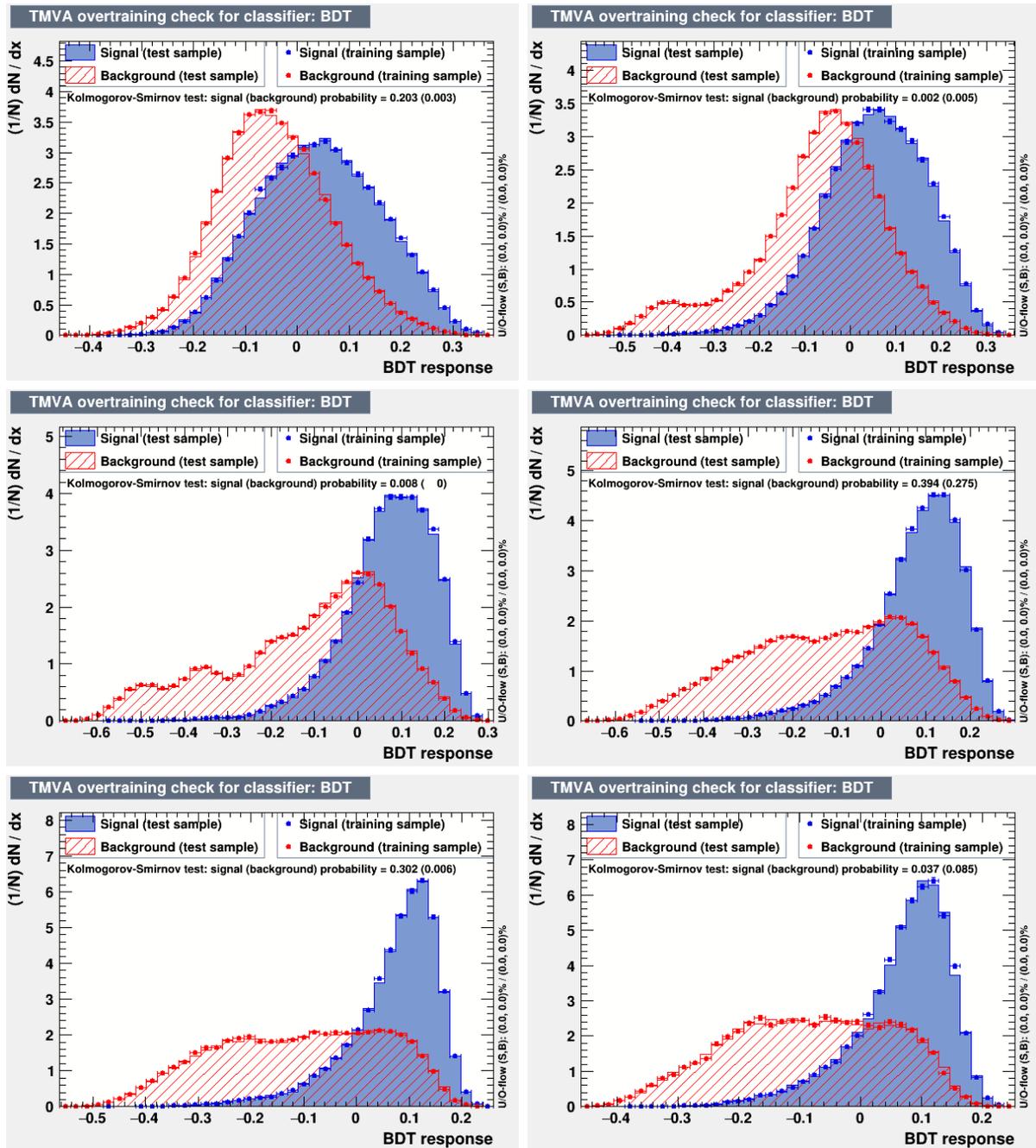


Figura 3.14: Distribuzioni sovrapposte delle funzioni di risposta per segnale e fondo, ottenute applicando il metodo BDT ai campioni di training e di testing separatamente. I risultati non risultano affetti da overtraining.

3.8 Confronto tra i metodi utilizzati

Il primo confronto tra i tre metodi studiati è stato fatto analizzando le curve di ROC (Receiver-Operating-Characteristic), visibili in figura 3.13, che rappresentano la relazione che intercorre tra l'efficienza del segnale (signal efficiency) e l'eliminazione del fondo (background rejection). Come si osserva dalle figure, maggiore è l'efficienza del segnale e minore è il fondo che viene rigettato e vice versa. La performance di un metodo risulta buona quando si riesce a rigettare molto fondo, mantenendo una buona efficienza del segnale. Il punto in cui si ha la performance migliore è chiamato working point. Come si può osservare in figura 3.15, in tutti gli intervalli le performance del BDT e delle reti neurali risultano molto simili, con curve quasi sovrapposte. È evidente che invece i tagli rettangolari hanno una prestazione molto meno performante e anche nel caso limite, in cui non viene rigettato il fondo, arrivano comunque ad un'efficienza massima del segnale pari solo a 0.8.

In tabella 3.13 sono riportati i working points di ciascun intervallo, relativi al BDT e alle reti neurali.

| Λ_c^+ p_T bin (GeV/c) | MLP | BDT |
|---------------------------------|-------|--------|
| $2 \leq p_T \leq 4$ | 0.338 | -0.006 |
| $4 \leq p_T \leq 6$ | 0.347 | 0.003 |
| $6 \leq p_T \leq 8$ | 0.387 | 0.022 |
| $8 \leq p_T \leq 12$ | 0.399 | 0.026 |
| $12 \leq p_T \leq 16$ | 0.541 | 0.028 |
| $16 \leq p_T \leq 25$ | 0.682 | 0.027 |

Tabella 3.13: Working points per ciascun intervallo, relativi al BDT e alle reti neurali.

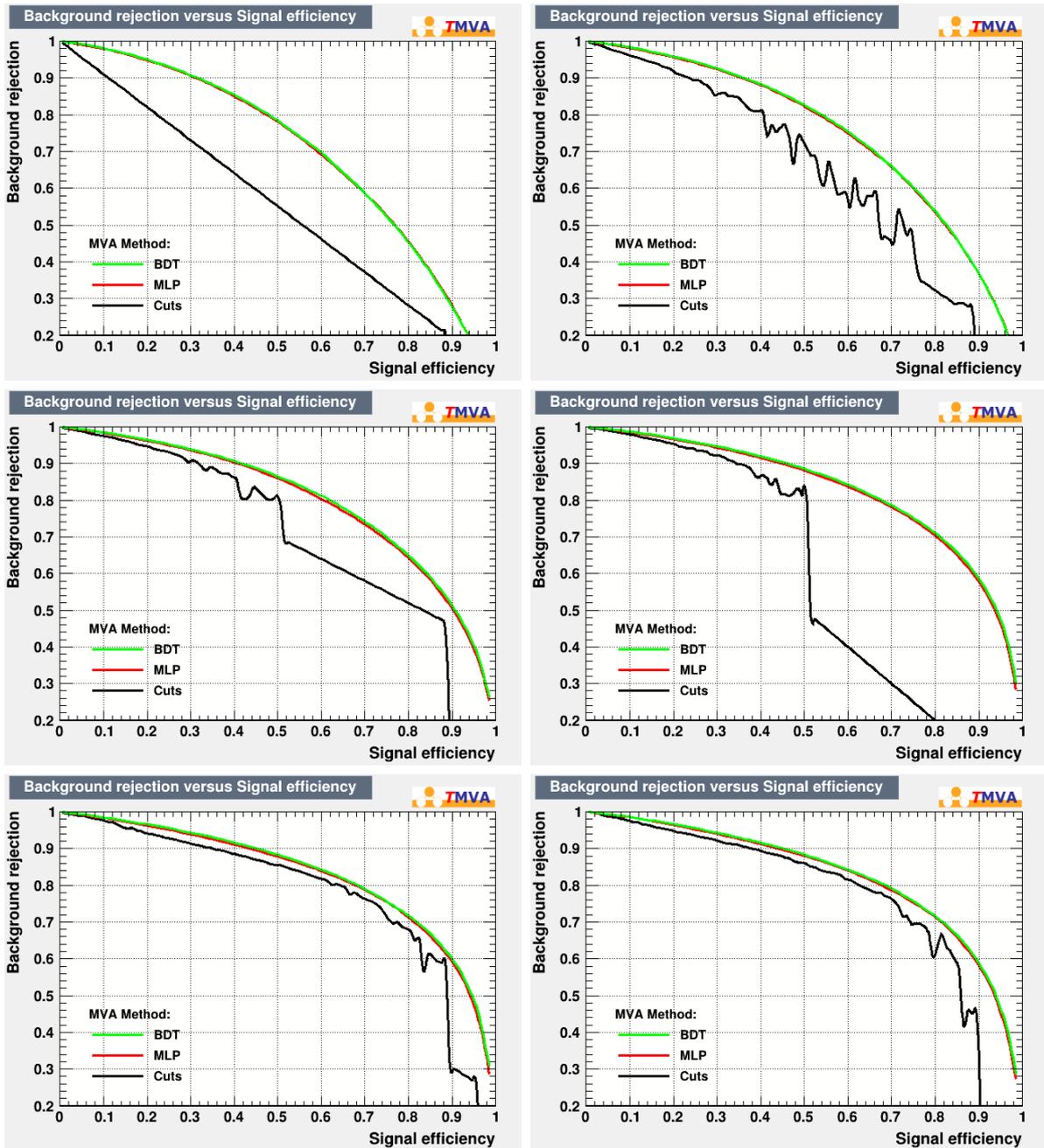


Figura 3.15: Curve di ROC per i tre metodi analizzati. Ad una maggiore efficienza del segnale corrisponde una minore quantità di fondo che viene rigettata.

Per concludere è evidente che il metodo dei tagli rettangolari risulta il meno performante, mentre i BDT e le reti neurali, per quanto osservato, danno una risposta con una

precisione molto simile. Bisogna però tenere conto anche di alcune altre caratteristiche: sebbene il training risulti essere molto più veloce per le reti neurali, in caso di presenza di molte variabili di input, il metodo dei BDT eseguirà il testing nel minor tempo; questo è dovuto al fatto che il BDT, dopo il training, utilizza meno tempo sulle variabili che ha catalogato come meno importanti, mentre le reti neurali le usano tutte, nonostante le selezioni fatte nel training. Per questo motivo, anche se le reti neurali sono performanti in una grande varietà di problemi lineari e non lineari, il loro utilizzo necessita che il numero di variabili non sia troppo elevato, richiedendo all'utente di eliminare a priori quelle meno discriminanti. Un'altra caratteristica da considerare è la difficile interpretazione degli output delle reti neurali, a differenza dei BDT molto più intuitivi.

Conclusione

Le tecniche di machine learning stanno progressivamente ampliando il loro utilizzo in fisica delle alte energie, dove i segnali che si vogliono investigare sono solitamente di difficile estrazione a causa dell'elevato fondo.

Tali tecniche permettono di considerare contemporaneamente più variabili di un singolo evento e di classificare in seguito tutti gli eventi analizzati in base a regolarità che l'algoritmo di selezione apprende direttamente dai dati.

Come si è visto in questa tesi, esistono diverse tecniche di analisi multivariata, ognuna delle quali utilizza criteri diversi per suddividere gli eventi di segnale e di fondo, e per questo presentano risultati e performance differenti.

Dall'analisi dei tre metodi presi in considerazione in questa tesi, risulta evidente che in confronto ai tagli rettangolari, che non sono propriamente una tecnica di analisi multivariata, le reti neurali e i BDT risultano molto più performanti, dando risposte con una precisione molto simile, come si osserva dal confronto delle efficienze e delle curve di ROC. Le principali differenze che rendono uno dei due metodi più conveniente rimangono quindi principalmente a livello di tempo di esecuzione e di facilità di interpretazione degli output. Per affrontare analisi complicate come quella del barione charmato Λ_c^+ , l'uso di tecniche di machine learning risulta indispensabile; è interessante notare che nonostante le reti neurali e i BDT abbiano performance molto simili, in ALICE siano stati utilizzati prevalentemente questi ultimi nelle analisi finora pubblicate. I BDT sono sicuramente la scelta più appropriata nel caso di un elevato numero di variabili in ingresso, in quanto, dando meno peso alle variabili con minore importanza, permettono di avere un grosso risparmio di tempo nella fase di testing. Diverso è il caso in cui si hanno poche variabili di ingresso, ma molti eventi da considerare: in queste condizioni le reti neurali permettono un guadagno di tempo sul training. Questa caratteristica può certamente essere presa in considerazione per le future analisi in ALICE.

Bibliografia

- [1] S. Braibant, G. Giacomelli, M. Spurio, *Particelle e Interazioni Fondamentali*, Springer-Verlag Italia 2012
- [2] S. Bethke. *Experimental Tests of Asymptotic Freedom*. Max-Planck-Institut für Physik, Munich, Germany. February 5, 2008
- [3] F. Karsch, *Lattice QCD at High Temperature and Density*, Lect.NotesPhys.583(2002), pp.209–249
- [4] *Phenomenological Review on Quark–Gluon Plasma: Concepts vs. Observations Pasechnik*, Roman et al. Universe 3 (2017) no.1, 7 arXiv:1611.01533 [hep-ph]
- [5] Rajeev S. Bhalerao. *Relativistic heavy-ion collisions* arXiv:1404.3294 [nucl-th]
- [6] *Lectures on high-energy heavy-ion collisions at the LHC*. C.A. Salgado. Departamento de Física de Partículas and IGFAE. Universidade de Santiago de Compostela, Santiago de Compostela, Spain
- [7] K.J. Eskola, K. Kajantie, P.V. Ruuskanene and K. Tuominen, *Scaling of transverse energies and multiplicities with atomic number and energy in ultrarelativistic nuclear collisions*, Nucl. Phys. B, 570, [arXiv:hep-ph/9909456], 2000
- [8] A. Andronic, *An overview of the experimental study of quark-gluon matter in highenergy nucleus-nucleus collisions*, [arXiv:nucl-ex/1407.5003], 2014
- [9] Barbara Guerzoni. *Identified primary hadron spectra with the TOF detector of the ALICE experiment at LHC* PhD thesis, Alma Mater Studiorum, Università di Bologna, 2012
- [10] J. Rafelski, B. Müller, *Strangeness production in the quark-gluon plasma*, Phys.Rev.Lett.48(1982)1066
- [11] Sarkar, Sourav, Satz, Helmut Sinha, Bikash (Ed.). (2010). *The physics of the quarkgluon plasma Introductory lectures*. Germany: Springer

- [12] ALICE Collaboration. *Anisotropic flow of charged particles in in Pb-Pb at $\sqrt{s_{NN}} = 5.02$ TeV.* arXiv:1602.01119v2 [nucl-ex]. 4 April 2016
- [13] J. Adams, et al., *Evidence from d+Au measurements for final-state suppression of high $p(T)$ hadrons in Au+Au collisions at RHIC*, Phys.Rev.Lett.91:072304,2003
- [14] Julius Gronefeld. *The nuclear modification of charged particles in Pb-Pb at $\sqrt{s_{NN}} = 5.02$ TeV measured with ALICE.* Proceedings of Science, Fourth Annual Large Hadron Collider Physics. POS(LHCP2016)126
- [15] *J/ψ suppression at forward rapidity in Pb-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV.* ALICE Collaboration. arXiv:1606.08197 [nucl-ex]
- [16] Xabier Cid Vidal Ramon Cid Manzano. url: <http://www.lhc-closer.es/taking-a-closer-look-at-lhc/0.magnetic-multipoles>
- [17] The ALICE Experiment. *A Large Ion Collider Experiment.* 16/11/2018 <http://aliceinfo.cern.ch/Public/en/Chapter2/>
- [18] J. Ashkin H. Bethe. *Expeimental Nuclear Physics.* In: a cura di E. Segré. Wiley, 1953
- [19] ALICE Collaboration. *Technical Design Report of the Time Projection Chamber.* Rapp. tecn. CERN, 1999
- [20] J. D. Jackson. *Classical Electrodynamics.* In: 3a ed. Wiley, 1999, pp. 637–640
- [21] *ALICE Technical Design Report of the High Momentum Particle Identification Detector.* CERN / LHCC 98-19, ALICE TDR 1, 14 August 1998
- [22] *ALICE Technical Design Report of the Photon Spectrometer (PHOS).* CERN / LHCC 99-4, ALICE TDR 2, 5 March 1999
- [23] P. Crochet for the ALICE collaboration. *The ALICE Forward Muon Spectrometer.* ALICE-PUB-2001-58 version 1.0 2001-12-13
- [24] *ALICE Technical Design Report of the Zero Degree Calorimeter (ZDC).* CERN / LHCC 995, ALICE TDR 3, 5 March 1999
- [25] *TOF, Time of Flight for ALICE Experiment.* INFN, The ALICE Bologna Group. 15/11/2018 <https://www.bo.infn.it/alice/tof-over/index.html>
- [26] ALICE Collaboration. *The MRPC-based ALICE Time-Of-Flight detector: status and performance.* In: Nucl. Instr. Meth. Phys. Res. A 706 (2013)

- [27] ALICE Collaboration. *Performance of the ALICE Experiment at the CERN LHC*. In: Int. J. Mod. Phys. A 29 (2014)
- [28] Höcker, Andreas (CERN); Speckmayer, Peter (CERN); Stelzer, Jörg (CERN); Tegenfeldt, Fredrik (Iowa State Univ.); Voss, Helge (Max-Planck Inst. Kernphysik, Heidelberg), *PHYSTAT-LHC Workshop on Statistical Issues for LHC Physics*, CERN, Geneva, Switzerland, 27 - 29 Jun 2007, pp.184-187 (CERN-2008-001), 10.5170/CERN-2008-001.184
- [29] <https://root.cern.ch/root/html/doc/guides/users-guide/ROOTUsersGuide.html>
- [30] X.N Wang and M. Gyulassy, *HIJING: a Monte Carlo model for multiple jet production in pp, p-A and A-A collisions*, Phys. Rev. D44 (1991) 3501
- [31] T. Sjöstrand, S. Mrenna and P.Z. Skands, *PYTHIA 6.4 physics and manual*, JHEP 05 (2006) 026 [hep-ph/0603175]
- [32] R. Brun et al., *GEANT detector description and simulation tool*, CERN-W-5013 (1994)
- [33] ALICE Collaboration, *AliRoot, ALICE Offline framework for simulation, reconstruction and analysis*, <http://alice-offline.web.cern.ch/>