

UNIVERSITÀ DEGLI STUDI DI TRIESTE  
FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI

Tesi di Laurea  
in  
Fisica

**Simulazione di eventi SIDIS in Grid  
per l'esperimento COMPASS**

Candidato:  
Giulio Sbrizzai

Relatrice:  
prof.ssa Anna Martin

Correlatore:  
dott. Venicio Duic

Anno Accademico 2006-2007



# Indice

<b>Introduzione</b>	<b>1</b>
<b>1 I processi semi-inclusivi nel DIS</b>	<b>5</b>
1.1 Interazione leptone-nucleone . . . . .	5
1.2 Urto profondamente inelastico . . . . .	6
1.3 Il modello a partoni . . . . .	10
1.4 SIDIS non polarizzato . . . . .	15
1.4.1 Momento trasverso intrinseco . . . . .	17
1.4.2 Espressione completa della sezione d'urto SIDIS . . . . .	20
<b>2 Apparato sperimentale di COMPASS</b>	<b>25</b>
2.1 Principio di funzionamento dello spettrometro . . . . .	26
2.2 I sottosistemi . . . . .	30
2.3 Dati raccolti . . . . .	46
<b>3 La ricostruzione degli eventi e l'analisi dei dati</b>	<b>47</b>
3.1 La selezione degli eventi . . . . .	50
3.1.1 La selezione degli eventi DIS . . . . .	51
3.2 Misura dell'asimmetria . . . . .	52
3.3 Gli eventi simulati . . . . .	57
3.3.1 Programmi per la simulazione . . . . .	57
3.4 Test preliminari . . . . .	59
3.4.1 Confronto con i dati . . . . .	59
3.4.2 Situazione finale . . . . .	66

<b>4</b>	<b>Sistema per la simulazione distribuita di COMPASS</b>	<b>71</b>
4.1	Introduzione al sistema distribuito . . . . .	71
4.2	Grid . . . . .	72
4.2.1	L'Architettura Grid . . . . .	73
4.2.2	EGEE ed il middleware gLite . . . . .	76
4.3	Introduzione del sistema COMPASS . . . . .	79
4.3.1	Architettura del sistema . . . . .	82
4.4	Valutazione tecnologie Grid . . . . .	84
4.4.1	AMGA . . . . .	84
4.4.2	DIANE . . . . .	85
4.4.3	GANGA . . . . .	87
4.5	Implementazione . . . . .	88
4.5.1	Test di performance . . . . .	96
<b>5</b>	<b>Risultati</b>	<b>105</b>
5.1	Funzionamento del sistema distribuito per la simulazione . . . . .	105
5.1.1	Il sistema utilizzato . . . . .	105
5.1.2	Produzione finale . . . . .	108
5.2	Distribuzioni azimutali . . . . .	108
5.2.1	L'analisi dei dati . . . . .	108
5.2.2	La stima dell'accettanza . . . . .	112
5.2.3	La distribuzione azimutale . . . . .	113
5.3	Estrazione delle asimmetrie azimutali . . . . .	123
5.3.1	Le asimmetrie misurate . . . . .	123
5.3.2	Stabilità dei risultati . . . . .	128
5.4	Confronto con misure precedenti e previsioni teoriche . . . . .	134
	<b>Conclusioni</b>	<b>139</b>
	<b>Bibliografia</b>	<b>141</b>
	<b>Ringraziamenti</b>	<b>147</b>

# Inroduzione

La maggior parte delle informazioni sulla struttura dei costituenti dei nuclei, i nucleoni, è stata ottenuta studiando l'urto di leptoni carichi su di essi. In particolare, l'urto profondamente anelastico (DIS), misurato per la prima volta a SLAC, ha permesso di sondare la struttura del nucleone, confermando nel 1972 l'ipotesi dell'esistenza di costituenti, detti partoni, subito identificati nei quark, ipotizzati da Gell-Mann e da Zweig otto anni prima per spiegare la spettroscopia adronica. Questo successo diede un grande impulso alle speculazioni di fisica teorica, nel cui contesto si svilupparono modelli che fornirono alcune previsioni (tra cui la regola di somma di Bjørken [1] e quella di Ellis-Jaffe [2]) . Molti esperimenti, a SLAC e al CERN di Ginevra, cercarono di verificare sperimentalmente queste previsioni. I risultati ottenuti furono molto spesso sorprendenti, e rivoluzionarono la visione del nucleone, portando nuovi impulsi allo sviluppo di teorie che dessero una spiegazione completa e consistente con i risultati sperimentali.

COMPASS<sup>1</sup> è un esperimento delle particelle approvato nel 1997 ed in atto al Super Proton Synchrotron (SPS) al CERN(Ginevra) che si prefigge numerose misure nel campo della struttura e della spettroscopia adronica; tra cui:

- studi sulla struttura dello spin del nucleone, tra cui:
  - misura della polarizzazione del gluone,  $\frac{\Delta G}{G}$
  - misura delle funzioni di struttura di spin trasverso,  $\Delta_T q$
- argomenti di spettroscopia adronica, in particolare:
  - ricerca di *glueballs*
  - studio di decadimenti leptonici e semileptonici di particelle charmate

---

<sup>1</sup>Common Muon Proton Apparatus for Structure and Spectroscopy

– ricerca di particelle doppiamente charmate

Nell'ambito del programma muonico di COMPASS è stata studiata l'interazione di un fascio polarizzato longitudinalmente di muoni positivi (con un momento medio di 160 GeV/c) in un bersaglio solido polarizzato di Litio deuterato ( ${}^6\text{LiD}$ ) e sono stati acquisiti dati dal 2002 al 2006, mentre nella presa dati in corso il bersaglio è costituito da ammoniaca,  $\text{NH}_3$ . La maggior parte (circa l'80%) del tempo dell'acquisizione è stato dedicato alla misura della polarizzazione dei gluoni mediante la stima dell'asimmetria longitudinale di spin in DIS con produzione semi inclusiva di adroni charmati.

La parte restante dell'acquisizione è stata dedicata alla misura di effetti di spin trasverso, che originariamente prevedeva la sola misura della funzione di Collins alla quale si sono aggiunte nel tempo altri effetti (effetto Sivers, funzioni TMD, misurabili tramite dipendenze azimutali).

Lo studio di asimmetrie azimutali nelle distribuzioni degli adroni prodotti nel processo di DIS seminclusivo su bersaglio polarizzato trasversalmente è il principale campo di ricerca in cui è impegnato il gruppo COMPASS-Trieste.

In aggiunta, negli ultimi anni un crescente interesse è stato mostrato per gli effetti di asimmetrie azimutali presenti nella produzione di adroni DIS non polarizzato legate alla dipendenza delle funzioni di distribuzione e di frammentazione dal momento trasverso intrinseco del quark e dell'adrone ed alla correlazione tra questo ed il vettore di spin trasverso. La possibilità di analizzare i dati presi da COMPASS, sia in configurazione longitudinale che trasversa, per studiare anche le asimmetrie azimutali per processi di SIDIS non polarizzato, è l'obiettivo principale di questa tesi. La misura di tali asimmetrie richiede una procedura di analisi diversa da quella utilizzata per lo studio delle asimmetrie in processi SIDIS polarizzati e richiede una buona simulazione dell'apparato con programmi di Monte Carlo. In particolare è richiesta una massiccia produzione di eventi simulati per stimare l'accettanza dell'apparato che da l'effetto dominante nella distribuzione dell'angolo azimutale dell'adrone prodotto. Per la simulazione è stato ideato, costruito e testato un sistema distribuito che avesse accesso sia alle risorse locali sia a quelle disponibili tramite Grid.

Nel primo capitolo della tesi è presentata una breve introduzione alla teoria dei

processi DIS e l'espressione completa della sezione d'urto SIDIS non polarizzata con l'esplicita dipendenza dall'angolo azimutale dell'adrone.

Nel secondo capitolo è descritto lo spettrometro di COMPASS ed i principi del suo funzionamento, mentre nel terzo capitolo è descritta la processatura e l'analisi dei dati presi dall'esperimento e di quelli ottenuti con la simulazione. Nello stesso capitolo sono inoltre spiegate: la procedura per pesare i dati, presi su bersaglio polarizzato, in modo da ottenere la distribuzione finale dell'angolo azimutale dell'adrone; gli studi fatti per verificare e migliorare la simulazione dell'accettazione dell'apparato.

Nel quarto capitolo viene descritto il sistema che è stato sviluppato, nella tesi, per la simulazione Monte Carlo di COMPASS. Tale sistema, per consentire l'uso trasparente di risorse computazionali in Grid, fa uso di alcuni interessanti prodotti del middleware del progetto LCG (DIANE e AMGA) ed ha rappresentato anche lo spunto per suggerimenti, e fonte di debug, per i loro sviluppatori alla sezione IT del CERN.

Infine, nel quinto capitolo vengono presentati i risultati ottenuti per lo studio delle asimmetrie azimutali e sono confrontati con le recenti previsioni teoriche ed alcune misure fatte da esperimenti precedenti.



# Capitolo 1

## I processi semi-inclusivi nel DIS

### 1.1 Interazione leptone-nucleone

È possibile misurare la distribuzione di carica di un bersaglio dotato di struttura interna studiando la diffusione elastica di leptoni carichi su di esso. Se ipotizziamo che il bersaglio abbia spin nullo, l'unica interazione tra esso e il leptone carico è quella elettrica; in ambito non relativistico, la distribuzione angolare dei leptoni diffusi è legata a quella relativa ad un bersaglio puntiforme dalla relazione[5]:

$$\left(\frac{d\sigma}{d\Omega}\right) = \left(\frac{d\sigma}{d\Omega}\right)_{point} |F(q^2)|^2, \quad (1.1)$$

dove  $q$  è il modulo del trimomento trasferito dal leptone al bersaglio, e  $F(q^2)$  è il fattore di forma del bersaglio, che per basse energie è direttamente riconducibile alla distribuzione di carica tramite una trasformazione di Fourier. I primi esperimenti che hanno sondato la struttura interna del nucleone si sono avvalsi di questo approccio.

Per studiare il nucleone, che ha spin  $1/2$ , è necessario considerare anche l'interazione magnetica tra esso e il leptone incidente: avremo allora una sezione d'urto puntiforme che moltiplica un termine contenente due fattori di forma,  $G_E$  e  $G_M$ , che descrivono rispettivamente la distribuzione elettrica e magnetica del nucleone. La sezione d'urto risultante, calcolata nel sistema di riferimento del laboratorio, è quindi data dalla formula di Rosenbluth [6]:

$$\left(\frac{d\sigma}{d\Omega}\right)_{Ros} = \left(\frac{d\sigma}{d\Omega}\right)_{point} \cdot \left[ \frac{G_E^2(Q^2) + \tau G_M^2(Q^2)}{1 + \tau} + 2\tau G_M^2(Q^2) \tan^2 \frac{\theta}{2} \right], \quad (1.2)$$

dove  $\tau = \frac{Q^2}{4M^2c^2}$ , con  $Q$  il quadrimomento trasferito ( $Q^2 = -q^2$ ) e

$$\left. \frac{d\sigma}{d\Omega} \right|_{point} = \frac{\alpha^2}{4E^2 \sin^4 \frac{\theta}{2}} \cdot \frac{\cos^2 \frac{\theta}{2}}{1 + \frac{2E}{M} \sin^2 \frac{\theta}{2}}. \quad (1.3)$$

dove:  $E$  è l'energia del leptone prima dell'urto,  $M$  è la massa del nucleone (in questa formula viene trascurata la massa del leptone),  $\theta$  è l'angolo di diffusione del leptone e  $\alpha$  è la costante di interazione elettromagnetica.

Nel caso di collisione leptone-nucleone ad alte energie, non è possibile usare il limite non relativistico, e cade l'interpretazione che i fattori di forma siano le trasformate di Fourier della carica e del momento magnetico.

Nel caso limite  $Q^2 \rightarrow 0$ , i fattori di forma che appaiono nella sezione d'urto di Rosenbluth descrivono le proprietà statiche dei nucleoni:  $G_E$  è la carica elettrica del nucleone, normalizzata alla carica elementare,  $G_M$  è il momento magnetico normalizzato al magnetone nucleare:

$$\begin{aligned} G_E^p(Q^2 = 0) &= 1 & G_E^n(Q^2 = 0) &= 0 \\ G_M^p(Q^2 = 0) &= 2.79 & G_M^n(Q^2 = 0) &= -1.91 \end{aligned}$$

L'andamento dei fattori di forma è stato determinato sperimentalmente, misurando la sezione d'urto a vari angoli, per valori di  $Q^2$  fissato: se si esclude il fattore di forma elettrico per il neutrone, che è vicino a zero, gli altri tre fattori mostrano lo stesso andamento dipolare con  $Q^2$ :

$$G(Q^2) = \left( 1 + \frac{Q^2}{0.71(\text{GeV}/c)^2} \right)^{-2}. \quad (1.4)$$

A causa di questo andamento, il contributo della sezione d'urto elastica a grandi  $Q^2$  viene fortemente soppresso.

## 1.2 Urto profondamente inelastico

Per alte energie trasferite dal leptone incidente è possibile studiare la struttura interna del nucleone in modo più dettagliato.

In reazioni ad alto  $Q^2$  il canale principale di interazione diventa quello inelastico. In particolare la grande energia a disposizione fa sì che vengano creati mesoni e lo

stato finale è costituito da uno stato adronico più complesso di quello iniziale, anche se ovviamente il numero barionico è conservato.

Di solito, le reazioni inelastiche vengono studiate con approccio inclusivo e si parla di DIS (Deep Inelastic Scattering), come schematizzato in figura 1.1: si prescinde dallo stato adronico generato, e si studia la reazione tramite l'energia iniziale e finale del leptone e il suo angolo di diffusione. Da queste variabili sono state ottenute le prime informazioni sulla struttura del nucleone iniziale.

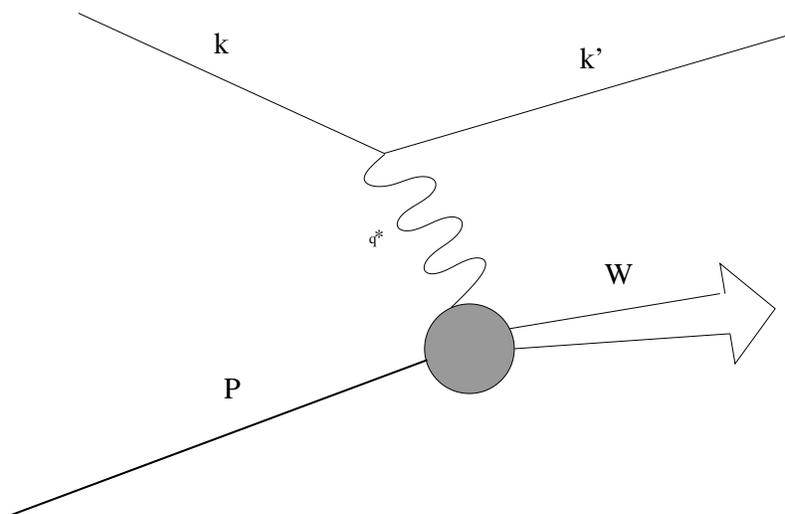


Figura 1.1: Rappresentazione schematica di una diffusione profondamente inelastica (DIS):  $k$  e  $k'$  sono i quadrivettori del leptone incidente e diffuso,  $P$  il quadrivettore del nucleone iniziale e  $q$  quello del fotone scambiato,  $q^*$ ;  $W$  è la massa invariante dello stato adronico finale.

La massa invariante dello stato adronico  $W$ , si ricava dai quadrimomenti del fotone scambiato  $q$ , e del nucleone iniziale  $P$  come mostrato nella seguente uguaglianza:

$$W^2 = (P + q)^2 \stackrel{lab}{=} M^2 + 2M\nu - Q^2, \quad (1.5)$$

dove  $M$  è la massa del nucleone iniziale, e  $\nu$  è un invariante relativistico definito come:

$$\nu = \frac{P \cdot q}{M}; \quad (1.6)$$

nel laboratorio,  $\nu$  coincide con l'energia trasferita dal leptone al nucleone:  $\nu = E - E'$ .

Nel caso elastico ad una energia incidente  $E$  fissata, c'è solo un parametro libero: la sezione d'urto di Rosenbluth e i fattori di forma che vi compaiono dipendono solo da  $Q^2$ . Invece nel caso inelastico l'energia di eccitazione del nucleone aggiunge un ulteriore grado di libertà e la sezione d'urto è funzione di due parametri indipendenti, ad esempio  $\nu$  e  $Q^2$ . Descrivendo la dinamica del processo inelastico in termini di fattori di forma, questi dipendono pertanto da due variabili e la generalizzazione della sezione d'urto di Rosenbluth (calcolata nel laboratorio) diventa [6]:

$$\left(\frac{d^2\sigma}{d\Omega dE'}\right)_{lab} = \frac{\alpha^2}{4E^2 \sin^4 \frac{\theta}{2}} \cdot \left[ W_2(Q^2, \nu) \cos^2 \frac{\theta}{2} + 2W_1(Q^2, \nu) \sin^2 \frac{\theta}{2} \right]. \quad (1.7)$$

Introduciamo a questo punto una variabile che misura l'inelasticità dell'urto, la *variabile di scala di Bjørken*:

$$x = \frac{Q^2}{2P \cdot q} = \frac{Q^2}{2M\nu}. \quad (1.8)$$

La variabile di Bjørken è compresa tra zero e uno:  $0 < x < 1$ . Nel caso elastico è  $x = 1$ ; in questo caso infatti è  $W = M$  e dall'equazione 1.5 si ottiene:

$$2M\nu - Q^2 = 0. \quad (1.9)$$

Nel caso inelastico invece  $W > M$ , per cui:

$$2M\nu - Q^2 > 0. \quad (1.10)$$

Si definisce urto profondamente inelastico (DIS) un'interazione con  $Q^2 \rightarrow \infty$ ,  $\nu \rightarrow \infty$  e valori di  $x$  finiti e le quantità che vengono valutate sono i fattori di forma adimensionali (chiamati anche funzioni di struttura):

$$\begin{aligned} F_1(x, Q^2) &= Mc^2 W_1(Q^2, \nu) \\ F_2(x, Q^2) &= \nu W_2(Q^2, \nu). \end{aligned} \quad (1.11)$$

Dai dati sperimentali, come mostrato in figura 1.2, si vede che per valori fissi di  $x$  e alti valori di  $Q^2$ , i fattori di forma non dipendono (o dipendono solo debolmente) da  $Q^2$ . Questa caratteristica è detta *invarianza di scala*, formalmente indicata in questo modo:

$$\begin{aligned} F_1(x, Q^2) &\longrightarrow F_1(x) \\ F_2(x, Q^2) &\longrightarrow F_2(x). \end{aligned} \quad (1.12)$$

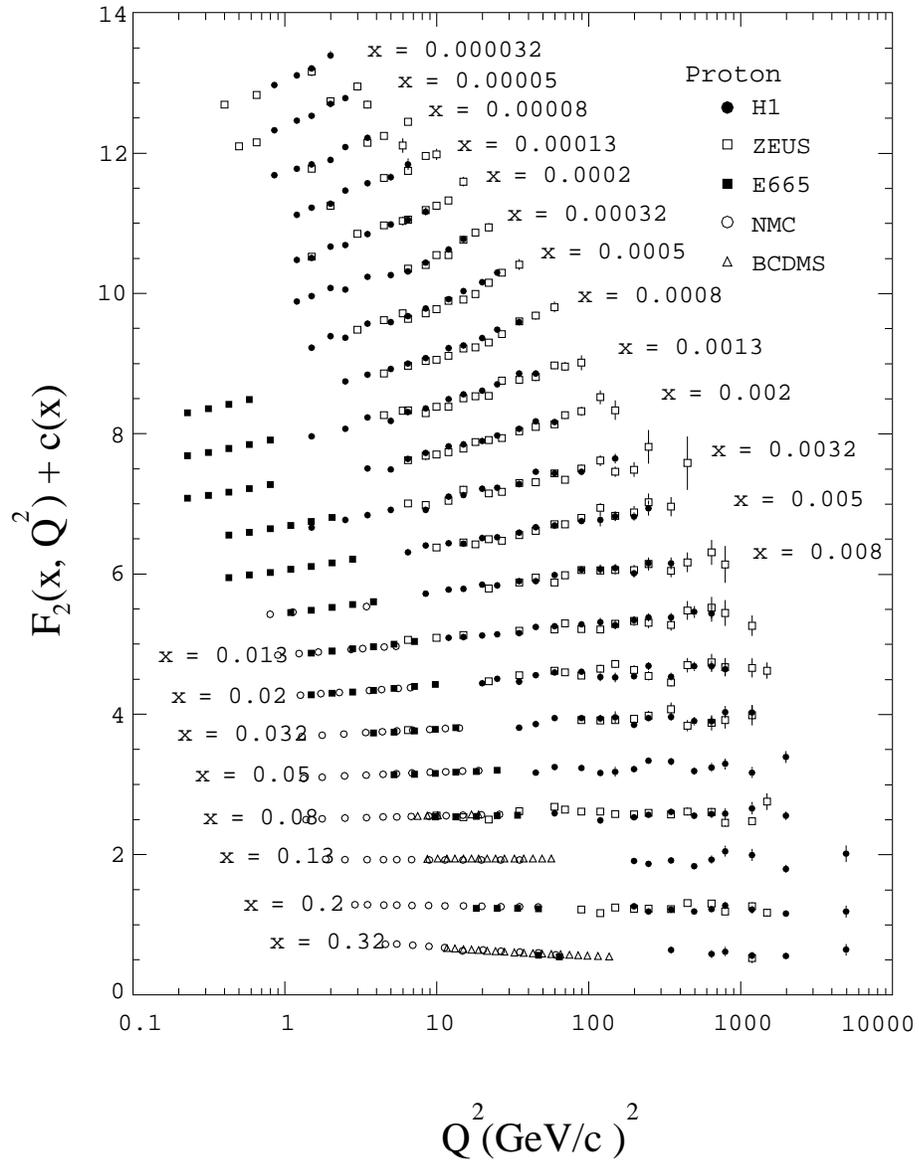


Figura 1.2: La funzione di struttura  $F_2$  in funzione di  $Q^2$  in intervalli di  $x$ . I risultati provengono da diversi esperimenti eseguiti a SLAC, al CERN, a DESY.

L'indipendenza dei fattori di forma dalla variabile  $Q^2$  è tipica di urti su sorgenti puntiformi (basti pensare, nel caso elastico, che la trasformata di Fourier di un fattore di forma costante corrisponde ad una distribuzione di carica a delta di Dirac); ne segue che i nucleoni sono costituiti da componenti puntiformi. Inoltre, dalla relazione di Callan-Gross (anch'essa verificata sperimentalmente come mostrato in figura 1.3):

$$2xF_1(x) = F_2(x), \quad (1.13)$$

segue che la sezione d'urto DIS presenta la tipica dipendenza angolare della sezione d'urto che descrive l'interazione tra un leptone e una particella carica puntiforme di spin 1/2. Tale analogia implica che anche i costituenti puntiformi del nucleone abbiano spin 1/2 e, scegliendo  $x$  ed  $y$  come le due variabili indipendenti che descrivono il processo di DIS, dove

$$y = \frac{P \cdot q}{P \cdot k} \stackrel{lab}{=} \frac{\nu}{E}; \quad (1.14)$$

la sezione d'urto scritta nell'equazione 1.7 diventa:

$$\left( \frac{d^2\sigma}{dx dy} \right) = \frac{2\pi\alpha^2}{Q^4} s [1 + (1 - y)^2] F_2(x). \quad (1.15)$$

Il modello a partoni, descritto nel paragrafo successivo, fornisce un'interpretazione dei risultati sperimentali appena descritti.

### 1.3 Il modello a partoni

Nel modello a partoni [6] il nucleone è costituito da un certo numero di particelle senza struttura, i partoni, ciascuna delle quali porta una frazione  $x$  del momento e dell'energia del nucleone:

	<i>nucleone</i>	<i>partone</i>	
<i>energia</i>	$E$	$xE$	
<i>momento</i>	$P$	$xP$	
<i>massa</i>	$M$	$xM$	(1.16)

e l'urto tra leptone e nucleone è rappresentato come in figura 1.4.

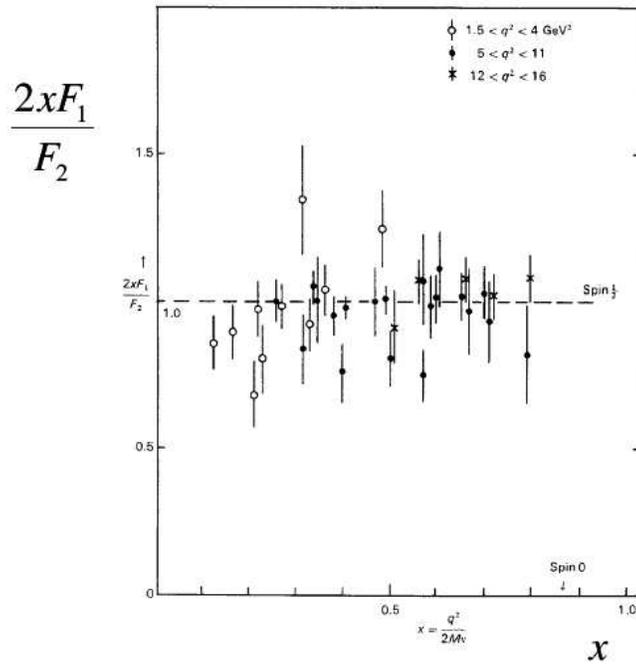


Figura 1.3: Conferma sperimentale della relazione di Callan Gross per diversi valori di  $Q^2$ .

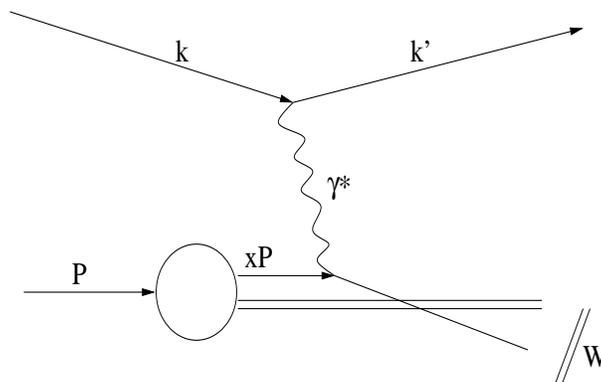


Figura 1.4: Cinematica di una reazione DIS descritta con il modello a partoni.

Se si descrive il processo DIS nel sistema di riferimento in cui il bersaglio ha un momento che tende ad infinito e si possono quindi trascurare sia le componenti trasverse del momento sia le masse a riposo dei partoni e del nucleone, la sua interpretazione tramite il modello a partoni diventa molto intuitiva:

$$(xP + q)^2 = (xP)^2 - Q^2 + 2xPq = x^2 M^2 = 0, \quad (1.17)$$

da cui, trascurando il termine  $(xP)^2$ , si ottiene:

$$x = \frac{Q^2}{2P \cdot q}. \quad (1.18)$$

Poiché la frazione del quadrimomento del nucleone portato dal partone viene così a coincidere con la variabile di Björken precedentemente definita,  $xP$  è interpretato come il quadrimomento del partone che ha assorbito il fotone scambiato nel processo. Inoltre, se tale assorbimento avviene in un intervallo di tempo tale per cui si possa considerare libero il partone colpito (non interagente con gli altri costituenti del nucleone), l'interazione leptone-nucleone diventa la somma incoerente delle interazioni tra leptone e singoli partoni.

La sezione d'urto può essere scritta come:

$$\left( \frac{d^2\sigma}{dx dy} \right)_{\ell N \rightarrow \ell' X} = \sum_i f_i(x) \cdot \left( \frac{d^2\sigma}{dx dy} \right)_{\ell p_i \rightarrow \ell' p'_i} \quad (1.19)$$

dove  $\ell$  e  $N$  indicano rispettivamente il leptone ed il nucleone iniziali,  $i$  indica il partone  $p_i$  che interagisce con il leptone  $\ell$ . In seno a questo modello la funzione  $f_i(x)$  (Parton Distribution Function - PDF) è interpretata come il numero medio di partoni  $p_i$  con quadrimomento  $xP$ , dove  $P$  è il quadrimomento del nucleone nello stato iniziale.

La sezione d'urto dell'interazione tra leptone e partone vale:

$$\left( \frac{d^2\sigma}{dx dy} \right)_{\ell p_i \rightarrow \ell' p'_i} = \frac{2\pi\alpha^2}{Q^4} e_i^2 \hat{s} [1 + (1 - y)^2], \quad (1.20)$$

dove  $\hat{s} = (xP + k)^2 \simeq 2xP \cdot k$  è la variabile di Mandelstam relativa all'urto leptone-partone. Sostituendo 1.20 nella sezione d'urto scritta in 1.19 si ottiene:

$$\left( \frac{d^2\sigma}{dx dy} \right)_{\ell p \rightarrow \ell' X} = \frac{2\pi\alpha^2}{Q^4} s [1 + (1 - y)^2] \sum_i e_i^2 x f_i(x), \quad (1.21)$$

dove  $s = (P + k)^2 \simeq 2P \cdot k$  è la variabile di Mandelstam relativa, invece, all'urto leptone-nucleone.

Alla luce della descrizione del modello non relativistico a quark il nucleone consiste di tre quark di valenza, che gli conferiscono le proprietà statiche (spin e carica); nel modello quindi i partoni vengono identificati con i tre quark di valenza. La funzione  $f_i(x)$  è interpretata come il numero medio di quark con frazione  $x$  di momento all'interno del nucleone e, chiamando  $e_i$  la frazione di carica portata da ciascun quark di sapore  $i$ , la funzione di struttura  $F_2$  che compare nella sezione d'urto 1.15, può essere scritta come:

$$F_2(x) = \sum_i e_i^2 x f_i(x). \quad (1.22)$$

I tre quark di valenza che compongono il nucleone dovrebbero spartirsi equamente il suo impulso, cioè la distribuzione  $F_2(x)$  misurata dovrebbe essere piccata attorno al valore  $1/3$ , come mostrato nel primo grafico in alto di figura 1.5. Invece, nell'interazione leptone nucleone, è stato osservato che  $F_2(x)$  ha un andamento simile a quello mostrato nel grafico in basso in figura 1.5, dove l'allargamento della distribuzione è causato dall'interazione tra i quark ed il comportamento a piccoli  $x$  è dovuto alla presenza di altre componenti identificate come coppie quark-antiquark del mare continuamente create nell'interazione.

Inoltre nel modello a partoni la somma degli impulsi dei singoli quark di valenza e dei quark (antiquark) del mare dovrebbe essere uguale ad 1 se questi fossero i soli componenti del nucleone, invece è stato osservato che:

$$\sum_i \int_0^1 dx x f_i(x) \simeq 0.5 \quad (1.23)$$

dove la somma è sulle distribuzioni di momento di tutti i quark ed antiquark che risultano presenti nel nucleone la cui struttura viene indagata studiandone l'urto con un leptone carico. L'interazione dominante tra leptone e partone in questo caso (corrispondente alle energie sperimentali raggiunte), è quella elettromagnetica che diventa trascurabile se il partone è neutro. È stato quindi dedotto che  $\sim$  il 50% del momento del nucleone è portato da particelle elementari neutre identificate poi con i gluoni che, secondo la teoria dell'interazione forte (Quantum Chromo Dynamics - QCD), sono le particelle scambiate nell'interazione forte e responsabili della coesione del nucleone.

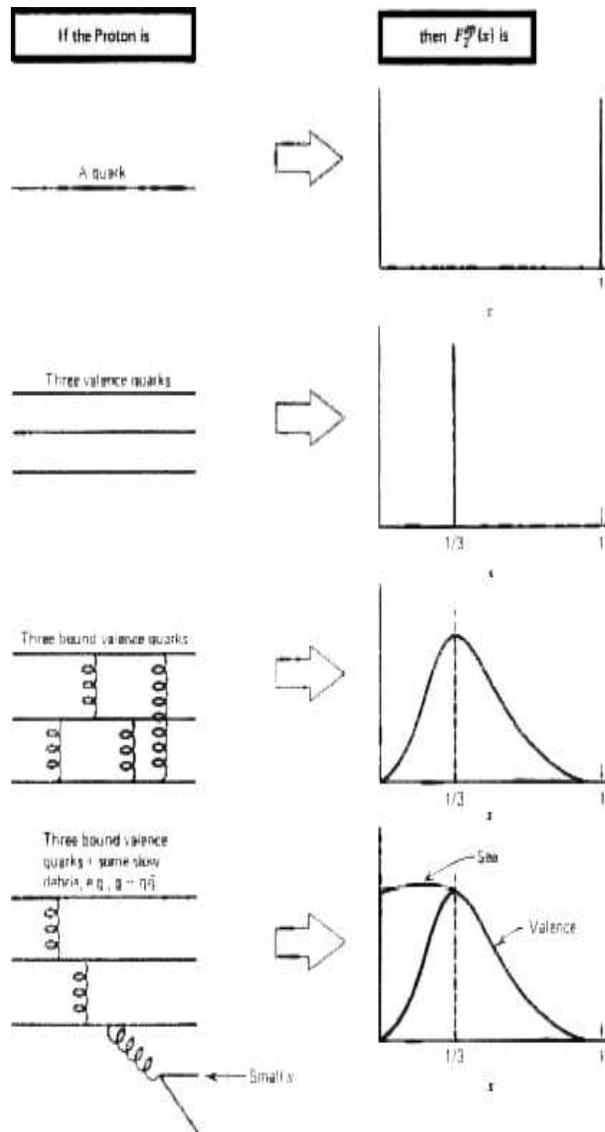


Figura 1.5: Funzione di struttura  $F_2$  nei casi in cui il nucleone sia costituito da un solo quark, prima riga in alto, da tre quark di valenza liberi, seconda riga, od interagenti, terza riga, ed infine l'ultima riga, nel caso in cui nel nucleone siano presenti anche le coppie quark-antiquark del mare.

## 1.4 SIDIS non polarizzato

Se, oltre al leptone diffuso, viene rilevato anche uno degli adroni prodotti nel processo di DIS, l'urto viene definito semi inclusivo e si parla di SIDIS (Semi Inclusive Deep Inelastic Scattering). Il processo studiato è

$$\ell(l) + N(P) \rightarrow \ell'(l') + h(P_h) + X, \quad (1.24)$$

dove  $l$  ed  $l'$  sono i quadrimomenti del leptone prima e dopo l'urto,  $P$  è il quadrimomento del nucleone  $N$  iniziale mentre  $P_h$  è quello dell'adrone prodotto.

La sezione d'urto viene scritta come:

$$d\sigma^{\ell p \rightarrow \ell' h X} \sim \sum_i f_i \cdot d\sigma^{\ell q_i \rightarrow \ell' q'_i} \cdot D_i^h. \quad (1.25)$$

Essa viene quindi espressa in termini di prodotti tra le distribuzioni partoniche, la sezione d'urto "dura" relativa all'interazione che avviene tra il fotone virtuale ed il quark, e le funzioni di frammentazione (FF)  $D_i^h$  che rappresentano la probabilità di ottenere un adrone ben definito a partire da un quark di un certo sapore. Le FF dipendono da una nuova variabile cinematica associata all'adrone rilevato,  $z = \frac{P \cdot P_h}{P \cdot q}$ , che è la frazione dell'energia del quark presa dall'adrone, calcolata nel laboratorio. Si definiscono FF favorite quelle per cui l'adrone prodotto contiene un quark di valenza dello stesso sapore di quello che ha frammentato e quindi si ha una probabilità relativamente grande di ottenere l'adrone, mentre si dicono sfavorite quelle FF per cui il sapore del quark che frammenta è diverso da quello dei quark di valenza che compongono l'adrone prodotto. Per la simmetria di isospin e l'invarianza per coniugazione di carica, considerando tre sapori del quark (u,d,s) e due possibili stati finali carichi ( $\pi^+$  e  $\pi^-$ ), in genere si assume:

$$D_1 = D_u^{\pi^+} = D_d^{\pi^+} = D_d^{\pi^-} = D_u^{\pi^-}, \quad (1.26)$$

$$D_2 = D_d^{\pi^+} = D_u^{\pi^+} = D_u^{\pi^-} = D_d^{\pi^-}, \quad (1.27)$$

$$D_3 = D_u^{\pi^+} = D_u^{\pi^-} = D_u^{\pi^-} = D_d^{\pi^-}, \quad (1.28)$$

e la loro dipendenza da  $z$  è mostrata nella figura 1.6. Nella figura 1.7, invece, sono mostrate le parametrizzazioni delle PDF relative ai diversi partoni presenti nel nucleone: quark u, d, s,  $\bar{u}$ ,  $\bar{d}$  e gluone. La parametrizzazione delle PDF e delle FF sono ottenute da fit globali di dati di processi DIS ed  $e^+e^-$ .

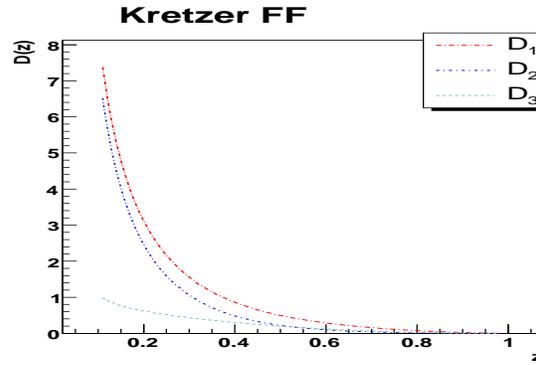


Figura 1.6: Dipendenza da  $z$  delle funzioni di frammentazione  $D_1$ ,  $D_2$  e  $D_3$  [7].

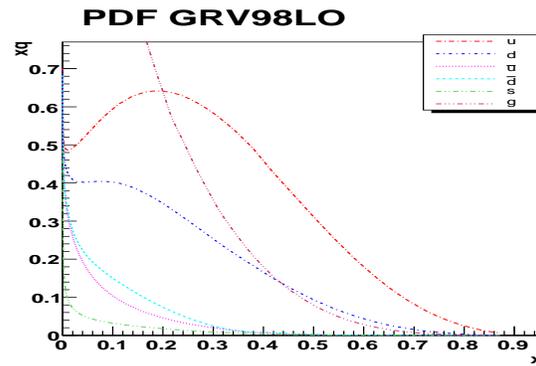


Figura 1.7: Funzioni di distribuzione partoniche (PDF) moltiplicate per  $x$ , ottenute dai fit globali, per i quark  $u$ ,  $d$ ,  $\bar{u}$ ,  $\bar{d}$ ,  $s$  e per il gluone [8].

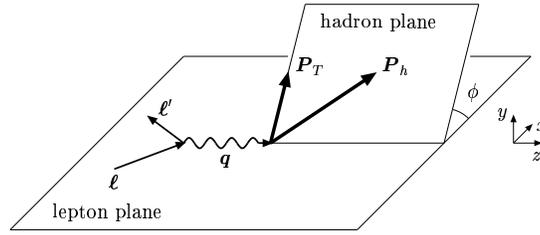


Figura 1.8: Piani dell’adrone e del leptone nell’urto SIDIS e alcune delle principali quantità considerate nella relativa sezione d’urto ( $\vec{P}_h$  è il momento dell’adrone prodotto e  $\vec{P}_T$  ne è la componente nel piano ortogonale alla direzione del fotone scambiato;  $\phi$  è l’angolo azimutale dell’adrone formato con il piano di diffusione). In piccolo è illustrata l’orientazione degli assi X, Y, Z nel Gamma Nucleon System (GNS) in cui l’asse Z coincide con la direzione del fotone scambiato e del nucleone bersaglio e X-Z è il piano contenente il muone incidente e diffuso.

### 1.4.1 Momento trasverso intrinseco

Quanto detto fino ad ora vale nel caso in cui il quark può essere effettivamente considerato collineare con il nucleone bersaglio e con l’adrone prodotto nella frammentazione, ma in realtà non si possono trascurare né il momento trasverso intrinseco del quark nel nucleone  $\vec{k}_T$ , né quello degli adroni rispetto alla direzione del quark che frammenta  $\vec{p}_T$ .

Negli ultimi anni è sempre più chiara l’importanza ed il ruolo di  $\vec{k}_T$ , che spesso non può essere ignorato se si vogliono spiegare sezioni d’urto non polarizzate in processi “duri” tipo  $pp \rightarrow hX$  [9] e la dipendenza dall’angolo azimutale della distribuzione degli adroni prodotti nel SIDIS, misurate ad esempio dall’esperienza EMC al CERN nel 1983 [10].

Misurare la dipendenza da  $\vec{k}_T$  delle PDF è quindi di fondamentale importanza per capire la struttura dei nucleoni, e le misure delle asimmetrie nelle distribuzioni azimutali degli adroni prodotti nel SIDIS non polarizzato è un passo importante in questa direzione.

Fu mostrato da Cahn [12] [13] che, nell’ambito del modello a partoni, introducendo una componente del momento trasverso del quark  $\vec{k}_T$  diversa da zero nella cinematica dell’urto, si ottengono asimmetrie nella distribuzione dell’angolo azimutale  $\phi$  tra il piano di produzione dell’adrone ed il piano di diffusione (figura 1.8).

L'effetto così generato risulta dominante, rispetto a quello previsto includendo la radiazione gluonica nella dinamica del processo [11].

Per calcolare l'asimmetria azimutale dovuta a  $\vec{k}_T$  (effetto Cahn) si consideri la sezione d'urto dell'interazione tra leptone incidente e quark, scritta nell'equazione 1.20, in funzione delle variabili di Mandelstam dell'urto elementare  $\hat{s} = (k + l)^2$ ,  $\hat{u} = (k - l')^2$  e  $\hat{t} = (l - l')^2 = Q^2$ , dove  $k$  è il quadrimomento portato dal quark e che nel caso collineare è  $k = xP$ . La sezione d'urto risulta essere:

$$d\sigma^{\ell q_i \rightarrow \ell' q'_i} \propto (\hat{s}^2 + \hat{u}^2), \quad (1.29)$$

Il momento trasverso intrinseco del quark nel GNS è  $k_T = (0, \vec{k}_T)$ , con  $\vec{k}_T = (k_T \cdot \cos(\phi_q), k_T \cdot \sin(\phi_q), 0)$ , ed è ortogonale alla direzione del fotone scambiato. Trascurando le masse del quark, del leptone e del nucleone, le variabili di Mandelstam valgono:

$$\hat{s} = 2l \cdot (xP + k_T) = x2l \cdot P - 2\vec{l}_T \vec{k}_T = xs - 2l_T k_T \cos \phi_q, \quad (1.30)$$

$$\hat{u} = -2l' \cdot (xP + k_T) = -\left(x2l' \cdot P - 2\vec{l}'_T \vec{k}_T\right) = xu + 2l'_T k_T \cos \phi_q, \quad (1.31)$$

$\phi_q$  è l'angolo azimutale del quark diffuso rispetto al piano di scattering ed  $s$  ed  $u$  sono le variabili di Mandelstam relative all'urto leptone-nucleone.

Poiché le traiettorie del leptone iniziale e finale giacciono sul piano di diffusione e per la conservazione del momento si ha:

$$\vec{l}_T = \vec{l}'_T = (l_T, 0, 0), \quad (1.32)$$

dove  $l_{T_x}$  è la componente lungo l'asse X del GNS.

Usando la relazione  $u = -(1 - y)s$  si ha:

$$\hat{s}^2 = (xs)^2 \left(1 - \frac{2l_T k_T}{xs} \cdot \cos(\phi_q)\right)^2, \quad (1.33)$$

$$\hat{u}^2 = (xs)^2 \left(1 - y - \frac{2l_T k_T}{xs} \cdot \cos(\phi_q)\right)^2. \quad (1.34)$$

Sia  $\alpha$  l'angolo formato dalla direzione del leptone incidente con quella del fotone scambiato nell'urto, si ha:

$$l_T = |\vec{l}| \cdot \sin(\alpha) \Big|_{Lab} = \frac{E E' \sin(\theta)}{\sqrt{Q^2 + \nu^2}} \Big|_{Lab}, \quad (1.35)$$

dove  $E$  ed  $E'$  sono le energie del leptone incidente e diffuso, rispettivamente, calcolate nel sistema del laboratorio e  $\theta$  è l'angolo formato dalla direzione del leptone incidente con quella del leptone diffuso. Esprimendo  $l_T$  in funzione di variabili Lorentz-invarianti si ha:

$$l_T = \frac{Q}{y} \cdot \sqrt{1-y}, \quad (1.36)$$

poiché:

$$E \cdot E' = \frac{\nu}{y} \cdot \frac{\nu}{y} (1-y), \quad (1.37)$$

$$\sin\theta = \frac{yQ}{\nu\sqrt{1-y}}, \quad (1.38)$$

dove si è considerato  $\nu^2 \gg Q^2$ .

Sostituendo la relazione 1.36 nella 1.34 e ponendo  $xs = \frac{Q^2}{y}$  si ottiene che la sezione d'urto elementare è:

$$\begin{aligned} d\sigma^{\ell q_i \rightarrow \ell' q'_i} &\propto (\hat{s}^2 + \hat{u}^2) \\ (\hat{s}^2 + \hat{u}^2) &= x^2 s^2 \left( \left( 1 - 2\sqrt{1-y} \cdot \frac{k_T}{Q} \cdot \cos(\phi_q) \right)^2 + (1-y)^2 \left( 1 - \frac{2k_T}{Q\sqrt{1-y}} \cdot \cos(\phi_q) \right)^2 \right) = \\ &= x^2 s^2 \left( 1 + (1-y)^2 - \frac{4k_T}{Q} (2-y) \sqrt{1-y} \cdot \cos(\phi_q) + \right. \\ &\quad \left. + \left( \frac{2k_T}{Q} \right)^2 (1-y) \cdot \cos(2 \cdot \phi_q) \right), \end{aligned} \quad (1.39)$$

si è usata la relazione trigonometrica  $\cos^2(\phi_q) = \frac{\cos(2 \cdot \phi_q) + 1}{2}$  e si è trascurato il termine  $\left( \frac{2k_T}{Q} \right)^2 \cdot (1-y)$  perché piccolo rispetto al termine  $1 + (1-y)^2$ , che è il termine consueto della sezione d'urto dell'interazione elettromagnetica tra due fermioni carichi puntiformi ed è l'unico termine che rimane se  $k_T = 0$ . Considerando quindi la dipendenza della funzione di distribuzione da  $\vec{k}_T$  si introducono due asimmetrie azimutali, con andamento  $\cos(\phi_q)$  e  $\cos(2\phi_q)$ , ma con diverse dipendenze da  $k_T/Q$ . In realtà si misura la direzione dell'adrone prodotto e quindi occorre tener conto anche del momento trasverso  $\vec{p}_T$  dell'adrone rispetto alla direzione di frammentazione del quark colpito, il che comporta una diluizione dell'asimmetria azimutale risultante. Si introducono quindi le funzioni di distribuzione e di frammentazione dipendenti dal momento trasverso  $f_i(\vec{k}_T, x)$  e  $D_i^h(\vec{p}_T, z)$  (*Transverse Momentum Dependent-TMD*).

### 1.4.2 Espressione completa della sezione d'urto SIDIS

L'espressione completa per la sezione d'urto SIDIS può essere scritta in modo indipendente dai modelli teorici [15] considerando lo scambio di un fotone singolo. Nell'equazione seguente è scritta la sezione d'urto calcolata in [15] per bersagli non polarizzati, al "leading order" di QCD e trascurando i termini proporzionali ad  $(\frac{M}{Q})^2$ , dove  $M$  è la massa del nucleone bersaglio:

$$\begin{aligned} \frac{d^5\sigma}{dx dy dz d\phi dP_T^2} &= \frac{2\pi\alpha^2}{xyQ^2} \cdot \\ &\left\{ \frac{1 + (1-y)^2}{2} F_{UU} + \right. \\ &+ (2-y) \sqrt{1-y} \cos\phi F_{UU}^{\cos\phi} + \\ &+ (1-y) \cos 2\phi F_{UU}^{\cos 2\phi} + \\ &\left. + \lambda_l y \sqrt{1-y} \sin\phi F_{LU}^{\sin\phi} \right\} \end{aligned} \quad (1.40)$$

dove  $\lambda_l$  è la polarizzazione del leptone incidente e, riguardo la notazione delle funzioni di struttura  $F$ , il primo indice definisce la polarizzazione del leptone incidente ed il secondo quella del nucleone bersaglio (U sta per non polarizzato, L sta per polarizzato longitudinalmente), il suffisso in apice indica la forma della modulazione azimutale relativa.

Usando la notazione compatta per la convoluzione sui momenti trasversi:

$$\mathcal{C}[w f D] = x \sum_i e_i^2 \int d^2p_T d^2k_T \delta^{(2)}(z\vec{k}_T + \vec{p}_T - \vec{P}_T) w(\vec{p}_T, \vec{k}_T) f_i(x, \vec{k}_T) D_i(z, \vec{p}_T). \quad (1.41)$$

dove  $w(\vec{k}_T, \vec{p}_T)$  indica una funzione dei momenti trasversi,  $f$  la generica TMD PDF e  $D$  la generica TMD FF e dove la somma è effettuata su tutti i quark ed antiquark

presenti nel nucleone, la forma esplicita delle funzioni di struttura è:

$$F_{UU} = \mathcal{C}[f D], \quad (1.42)$$

$$F_{UU}^{\cos \phi} = \frac{2M}{Q} \mathcal{C} \left[ -\frac{\hat{h} \cdot \vec{p}_T}{M_h} \cdot \left( \frac{k_T^2}{M^2} \cdot h^\perp H^\perp \right) - \frac{\hat{h} \cdot \vec{k}_T}{M} f D \right], \quad (1.43)$$

$$F_{UU}^{\cos 2\phi} = \mathcal{C} \left[ -\frac{2(\hat{h} \cdot \vec{p}_T)(\hat{h} \cdot \vec{k}_T) - \vec{k}_T \cdot \vec{p}_T}{MM_h} h^\perp H^\perp + \frac{2(\vec{k}_T \cdot \hat{h})^2 - \vec{k}_T^2}{Q^2} \cdot f D \right] \quad (1.44)$$

$$F_{LU}^{\sin \phi} = \frac{2M}{Q} \mathcal{C} \left[ -\frac{\hat{h} \cdot \vec{p}_T}{M_h} \frac{m}{M} f H^\perp + \frac{\hat{h} \cdot \vec{k}_T}{M} \frac{m}{M} h^\perp D \right], \quad (1.45)$$

$$(1.46)$$

con  $\hat{h} = \vec{P}_T/|P_T|$ ,  $\vec{P}_T$  il momento trasverso dell'adrone nel GNS,  $M$  ed  $M_h$  le masse del nucleone bersaglio e dell'adrone prodotto rispettivamente ed  $m$  è quella del quark.

- Il termine  $F_{UU}$  esprime l'interazione tra il leptone ed il nucleone nel caso in cui siano entrambi non polarizzati; esso è l'unico termine che non dipende dall'angolo azimutale. Integrando la 1.41 su  $\vec{P}_T$  e su  $z$  e sommando su tutti gli adroni  $F_{UU}$  diventa  $2xF_1$ .  $F_{UU}$  è proporzionale alla convoluzione tra le TMD PDF  $f(x, \vec{k}_T)$  e TMD FF  $D(z, \vec{k}_T)$  e, assumendo valida la fattorizzazione (espressa nelle equazioni 1.51), ai prodotti di  $f(x)$  ed  $D(z)$  che sono le funzioni attualmente meglio conosciute.
- il termine  $F_{UU}^{\cos \phi}$ , al primo ordine in  $\frac{k_T}{Q}$ , è dovuto all'effetto Cahn descritto nel paragrafo precedente e ancora contiene  $f(x, \vec{k}_T)$  e  $D(z, \vec{k}_T)$ . Un altro contributo è dovuto a effetti di QCD perturbativa, ma è molto minore [11];
- il termine  $F_{UU}^{\cos 2\phi}$  contiene la distribuzione di Boer-Mulders  $h^\perp$  che, accoppiata alla funzione di frammentazione di Collins  $H^\perp$ , genera un'asimmetria azimutale il cui effetto è visibile già all'ordine  $\frac{k_T}{Q}$  e domina sul contributo derivante dall'effetto Cahn, calcolato nel paragrafo precedente e dell'ordine di  $(\frac{k_T}{Q})^2$ . La funzione di Boer-Mulders è stata introdotta di recente e descrive la correlazione tra il momento intrinseco del quark ed il suo vettore di spin trasverso all'interno di un nucleone non polarizzato, mentre la funzione di Collins descrive la correlazione tra lo spin trasverso del quark e la direzione di emissione dell'adrone prodotto. Per esempio un quark che si muove orizzontalmente lungo la direzione positiva di Z e polarizzato lungo Y verso l'alto, emette il mesone

più energetico prodotto nella propria frammentazione, preferibilmente lungo la direzione positiva dell'asse X [16] [17]. Un altro contributo è dato da calcoli di QCD perturbativa, ma può essere trascurato [65].

- infine il termine  $F_{LU}^{\sin\phi}$  descrive l'interazione tra il leptone polarizzato longitudinalmente ed il nucleone non polarizzato.

La distribuzione dell'angolo azimutale dell'adrone, prevista, è quindi:

$$d\sigma \propto 1 + \varepsilon_1 A_1 \cos(\phi) + \varepsilon_2 A_2 \cos(2\phi) + \lambda_l \varepsilon_3 A_3 \sin(\phi), \quad (1.47)$$

dove i termini  $\varepsilon_i$  sono le funzioni di  $y$ :

$$\begin{aligned} \varepsilon_1 &= (2-y) \cdot \sqrt{1-y} \cdot \frac{2}{1+(1-y)^2}, \\ \varepsilon_2 &= (1-y) \cdot \sqrt{1-y} \cdot \frac{2}{1+(1-y)^2}, \\ \varepsilon_3 &= y \sqrt{1-y} \cdot \frac{2}{1+(1-y)^2}, \end{aligned} \quad (1.48)$$

e

$$\begin{aligned} A_1 &= \frac{F_{UU}^{\cos\phi}}{F_{UU}}, \\ A_2 &= \frac{F_{UU}^{\cos 2\phi}}{F_{UU}}, \\ A_3 &= \frac{F_{LU}^{\sin\phi}}{F_{UU}}. \end{aligned} \quad (1.49)$$

In generale, per calcolare esplicitamente  $A_i$  (le asimmetrie misurabili), si assume una dipendenza gaussiana dal momento trasverso intrinseco, si veda per esempio [13], e quindi si ha:

$$\begin{aligned} f(x, k_T) &= f(x) \cdot \frac{1}{\pi \langle k_T^2 \rangle} \cdot e^{\frac{-k_T^2}{\langle k_T^2 \rangle}}, \\ D(z, p_T) &= D(z) \cdot \frac{1}{\pi \langle p_T^2 \rangle} \cdot e^{\frac{-p_T^2}{\langle p_T^2 \rangle}}, \end{aligned} \quad (1.50)$$

così che  $\int d^2\vec{k}_T f(x, k_T) = f(x)$  e  $\int d^2\vec{p}_T D(z, p_T) = D(z)$ . In questo modo nelle asimmetrie misurabili  $A_i$  compare una funzione di  $\langle k_T^2 \rangle$  e  $\langle p_T^2 \rangle$  e del prodotto

delle PDF e FF integrate sul momento trasverso. Nel caso di  $A_1$  le PDF e FF sono quelle usuali non polarizzate, e quindi dalla sua misura si possono estrarre informazioni su  $\langle k_T^2 \rangle$  e  $\langle p_T^2 \rangle$ , utilizzabili poi per estrarre altre TMD PDF e FF dalle rimanenti due asimmetrie.

Le prime misure dell'asimmetria in  $\cos(\phi)$  furono pubblicate nel 1983 dall'esperimento EMC al CERN [18] e nel 1993 da E665 [19]. Misure più recenti per  $\cos(\phi)$  e  $\cos(2\phi)$  sono state fatte a ZEUS [20] [21] e prime misure di  $\sin(\phi)$  sono state fatte a CLAS [22] ed HERMES [23].

Data l'importanza delle asimmetrie azimutali presenti nell'urto SIDIS non polarizzato COMPASS, che già studia la struttura dello spin del nucleone, sta intraprendendo, parallelamente, lo studio della distribuzione angolare dell'adrone prodotto nell'urto SIDIS non polarizzato. Un primo lavoro preliminare è stato compiuto su parte dei dati disponibili in [24].

È in questo contesto che si colloca il lavoro di tesi, il cui scopo è lo studio della fattibilità di queste misure, più difficili delle misure di asimmetrie di spin in quanto richiedono una buona conoscenza e simulazione dell'apparato sperimentale, come verrà spiegato in seguito.



## Capitolo 2

# Apparato sperimentale di COMPASS

L'esperimento COMPASS è un esperimento a bersaglio fisso installato al CERN, nella hall 888, sulla linea di fascio M2 dell'SPS (Super Proton Synchrotron). La realizzazione del programma di fisica attuato da COMPASS impone specifici requisiti alle caratteristiche dello spettrometro: grande accettazione angolare ed in momento; capacità di ricostruire tracce lasciate da particelle diffuse ad angoli molto piccoli; precisa ricostruzione della cinematica di ogni evento; efficiente identificazione di particelle e buona risoluzione nella determinazione della loro massa. Le condizioni di alta luminosità, indispensabili per le misure di COMPASS, richiedono la capacità di avere un fascio ad alta intensità, di sostenere una elevata frequenza (rate) di conteggi e trigger e di poter sostenere un enorme flusso di dati. In figura 2.1 viene illustrato lo spettrometro nella sua configurazione per la presa dati del 2004. In essa si possono distinguere tre parti principali. La prima è costituita dai rivelatori che sono posizionati prima del bersaglio (upstream) e che servono a misurare le particelle del fascio prima dell'urto. I restanti rivelatori sono stati posti dopo il bersaglio (downstream), per una lunghezza totale di circa 50 m, e costituiscono lo spettrometro a grandi angoli (LAS) e quello a piccoli angoli (SAS) costruiti, ciascuno, attorno ad un magnete analizzatore, e completati con un calorimetro adronico ed un filtro muonico per l'identificazione di muoni ad alte energie. Un rivelatore RICH, per l'identificazione di adroni, è posto nella prima parte dello spettrometro, mentre nella parte a piccolo angolo è presente un calorimetro elettromagnetico.

L'uso di due spettrometri per rilevare le particelle diffuse è dettato dalla necessità di disporre di grande accettazione, angolare ed in momento. Il LAS (Large Angle Spectrometer), è stato progettato per assicurare una accettazione polare di 180 mrad.

È costruito attorno al magnete SM1, posto a circa 4 m downstream, lungo 110 cm, con un foro (gap) largo 229 cm ed alto 152 cm, la cui grandezza corrisponde all'accettanza angolare richiesta di  $\pm 180$  mrad. La componente principale del campo di dipolo prodotto va dall'alto verso il basso ed il suo potere deflettore è di 1 Tm, corrispondente ad una deflessione di 300 mrad per particelle con momento 1 GeV/c. Per questo, i rivelatori appartenenti a questa parte dello spettrometro devono avere un'accettanza angolare di  $\pm 250$  mrad nel piano orizzontale. Il SAS (Small Angle Spectrometer) rivela particelle a piccoli angoli ( $\pm 30$  mrad) ed alti momenti (5 GeV/c o più). L'elemento centrale è il magnete SM2, posto a 18 m downstream ed il cui gap è grande  $2 \times 1$  m<sup>2</sup>. Il potere deflettore è di 4.4 Tm e la principale componente del campo di dipolo è verticale, come per SM1.

In questo capitolo viene descritto il funzionamento dell'apparato sperimentale durante la presa dati del 2004 [26], per coerenza con gli studi fatti nella tesi.

## 2.1 Principio di funzionamento dello spettrometro

Prima di interagire con il bersaglio, il fascio di muoni viene rivelato da alcune stazioni di rivelatori di fibre scintillanti e due rivelatori a microstrip di Silicio, che hanno la funzione di misurarne la traccia entrante nel bersaglio. È presente, inoltre, un odoscopio di scintillatori che funge da "veto": esso copre un'ampia regione attorno al fascio e intercetta i muoni del fascio disperso (alone) che non interagiranno con il bersaglio ma attraverseranno lo spettrometro, interferendo con la ricostruzione delle tracce appartenenti ad adroni, e, soprattutto, generando segnali di trigger spuri.

Dopo questi rivelatori, si trova il bersaglio, circondato dal sistema criogenico e dal solenoide. I muoni del fascio interagiscono con i nucleoni contenuti nel <sup>6</sup>LiD, e a seguito di questa reazione, dal bersaglio escono particelle con diverso angolo polare, la cui rivelazione si avvale dei due diversi spettrometri dedicati.

La collocazione del LAS è subito a valle del bersaglio, in quanto le tracce delle particelle emesse a grande angolo si allontanerebbero troppo dalla linea di fascio se rivelate ad una maggiore distanza dal bersaglio, mentre il SAS è posto immediatamente a valle del LAS, quando le tracce delle particelle emesse a piccolo angolo si sono allontanate a sufficienza dall'asse del fascio.

Il costituente principale del LAS, illustrato in figura 2.2, è il magnete analiz-

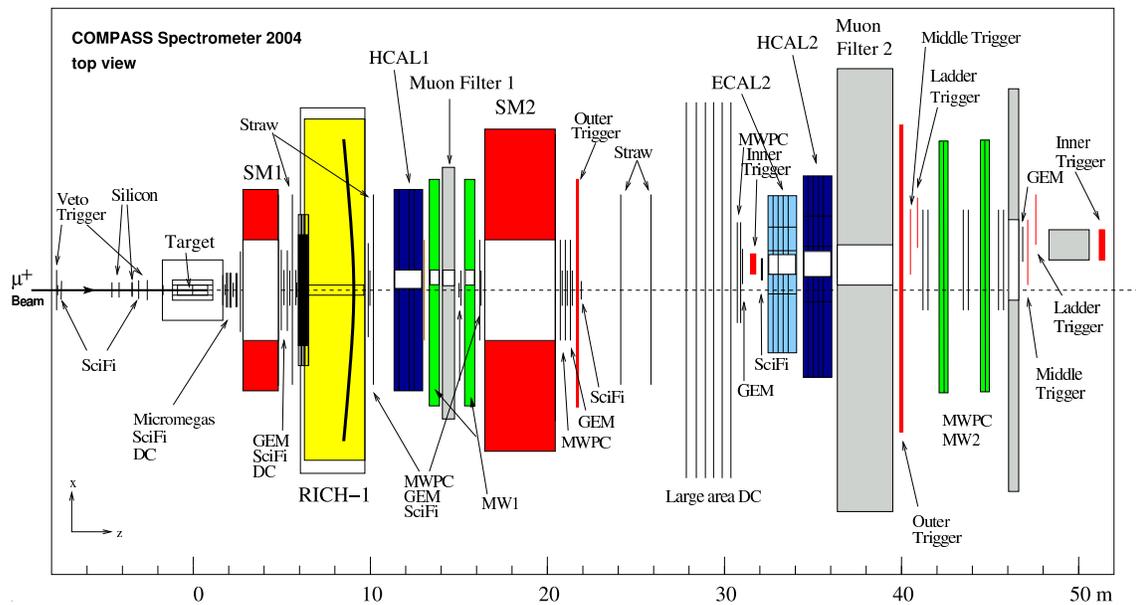
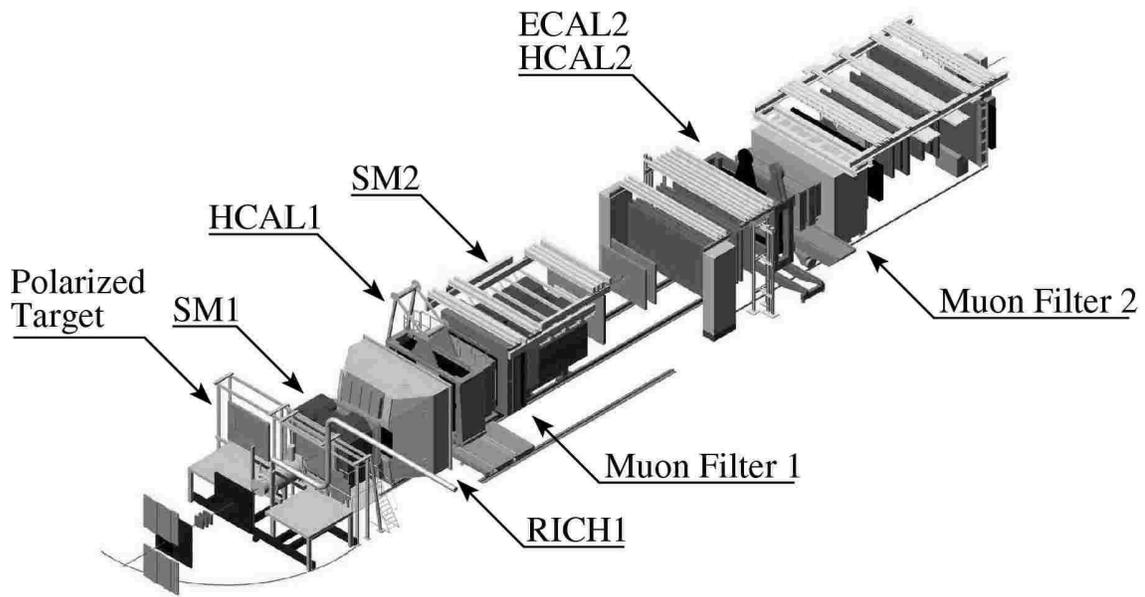


Figura 2.1: Lo spettrometro COMPASS durante la presa dati 2004; vista dall'alto. Viene anche indicato il sistema di riferimento la cui origine è posta al centro della cella downstream: l'asse Z (parallelo alla linea nominale di fascio e crescente nel verso che va dal bersaglio alla fine dello spettrometro), l'asse X (orizzontale, i cui valori sono crescenti andando dal lato Salève a quello Jura della sala sperimentale) e l'asse Y (verticale con valori crescenti verso l'alto) formano una terna ortogonale destrorsa.

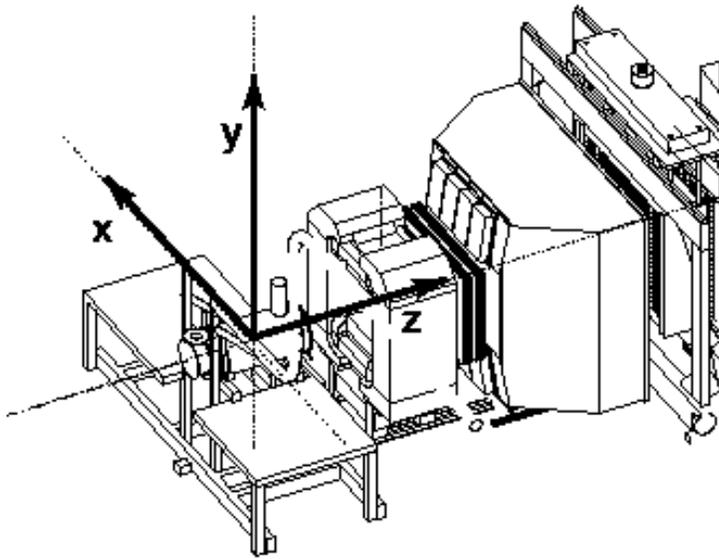


Figura 2.2: Visione assonometrica della zona del bersaglio e di parte del LAS, e definizione del sistema di riferimento usato in COMPASS. Il sistema di riferimento usato da COMPASS definisce l'asse Z lungo la linea ideale del fascio e con origine a metà della cella downstream, mentre l'asse X (coordinata orizzontale) ed Y (coordinata verticale con direzione positiva verso l'alto) sono definiti in modo da ottenere una terna destrorsa.

zatore, SM1, preceduto e seguito da numerose stazioni di rivelatori di diverso tipo (GEM, MICROMEGAS, SDC, *straw tube*). I tracciatori a monte e a valle di SM1 permettono la misura della deviazione angolare di una traccia, da cui si ricostruisce il momento della particella. La misura, quanto più possibile accurata, del momento, è essenziale per determinare la cinematica dell'evento, ed è indispensabile per l'identificazione della particella: usando la sola variabile misurata dal RICH, la velocità della particella, non sarebbe possibile infatti alcuna discriminazione in massa. L'identificazione dipende quindi anche dalla precisione con cui il momento è misurato, che a sua volta dipende dal potere deflettore del magnete e dalla risoluzione dei tracciatori.

Subito dopo i tracciatori a valle del magnete, si trova il RICH-1, il cui ruolo è misurare la velocità delle particelle per permetterne l'identificazione. Il RICH adopera un metodo non distruttivo di rivelazione, al contrario dei calorimetri, in quanto le particelle lo attraversano quasi indisturbate (a parte effetti di diffusione multipla, per ridurre i quali si è cercato di diminuire quanto più possibile la densità e lo spessore del materiale attraversato), e continuano il loro percorso verso le ultime componenti dello spettrometro. A valle del RICH-1 si trovano ancora tracciatori, (MWPC) e poi il calorimetro adronico, HCAL1.

Anche il calorimetro adronico contribuisce all'identificazione delle particelle, in quanto dalla misura dell'energia, di nuovo accoppiata alla misura del momento, si può ricavare la massa dell'adrone considerato. L'identificazione fatta tramite RICH o HCAL è diversa: il RICH è in grado di rivelare adroni e leptoni carichi, mentre HCAL è sensibile anche ad adroni neutri. Lo scopo principale di HCAL è infatti la rivelazione di tutti gli adroni (per es. neutroni provenienti da decadimenti di barioni charmati); inoltre è utilizzato per il sistema di trigger.

Infine, l'ultimo elemento che serve all'identificazione di particelle nel LAS è il filtro di muoni, che si basa sull'alta penetrazione dei muoni rispetto alle altre particelle: a valle dell'assorbitore, rimangono solo i muoni, che vengono rivelati da tracciatori posti a valle del filtro, e le cui tracce sono associate a quelle misurate a monte dell'assorbitore.

Il secondo spettrometro inizia subito dopo il primo, a  $\sim 15$  m di distanza dal bersaglio, e si estende per circa 30 m, fino alla fine dell'apparato. Le componenti di questo spettrometro sono le stesse, si ripete la sequenza di elementi del primo

spettrometro: le differenze sostanziali sono la mancanza del RICH-2 e la presenza di ECAL2 posto davanti ad HCAL2 ed atto a misurare l'energia rilasciata dagli sciame elettromagnetici.

## 2.2 I sottosistemi

### I magneti

I magneti dei due spettrometri sono dotati di poteri deflettori diversi, adatti alle caratteristiche delle particelle che devono analizzare: il momento medio delle particelle a grande angolo, analizzate da SM1, è minore. Il potere deflettore di SM1 è 1 Tm, mentre quello di SM2 è 4.4 Tm; questo implica una maggiore precisione nella ricostruzione del momento misurato nel SAS rispetto a quello ottenuto nel LAS.

Le *figure di merito* dei due spettrometri nel caso in cui una traccia sia misurata a monte e a valle dei rispettivi magneti si possono calcolare dalla formula:

$$\frac{\sigma_p}{p^2} = \frac{\sigma_\theta}{0.3(BL)} (\text{GeV}/c)^{-1}, \quad (2.1)$$

dove il parametro (BL) è il potere deflettore del magnete, e  $\sigma_\theta$  è l'errore sulla deflessione della traccia. L'errore sulla direzione della traccia corrisponde a  $\sim 0.4$  mrad per tracce rivelate solo nel LAS, e a  $\sim 0.08$  mrad per quelle rivelate anche nel SAS (questi valori sono stati ottenuti [28] confrontando i parametri delle tracce generate con MC con quelli delle tracce ricostruite con l'algoritmo usato nell'analisi), quindi si trova:

$$\left(\frac{\sigma_p}{p^2}\right)_{LAS} \sim 2 \cdot 10^{-3} (\text{GeV}/c)^{-1}, \quad (2.2)$$

$$\left(\frac{\sigma_p}{p^2}\right)_{SAS} \sim 10^{-4} (\text{GeV}/c)^{-1}. \quad (2.3)$$

### Il fascio di $\mu$

La linea di fascio M2 permette il trasporto di fasci di  $\mu^+$  di alta intensità e momento fino a 190 GeV/c. Il fascio primario usato per produrre i  $\mu^+$  è estratto dal SPS, ed è costituito da protoni aventi momento 400 GeV/c che incidono su un bersaglio solido di Berillio spesso 50 cm. Nell'interazione vengono prodotti soprattutto pioni (con 3.6% di contaminazione da Kaoni). I pioni di momento 225 GeV/c  $\pm 10\%$

vengono selezionati ad opera di quadrupoli e dipoli e quindi trasportati lungo una linea di 600 m nella quale in parte decadono in  $\mu\nu$ . La selezione dei muoni è fatta tramite un assorbitore di Berillio che ferma le altre particelle, meno penetranti. La polarizzazione longitudinale del fascio, indispensabile per le misure di asimmetria di doppio spin, è ottenuta naturalmente: l'elicità di un muone positivo nel sistema di riferimento del pione che lo ha prodotto è -1, dunque la polarizzazione dei muoni nel sistema del centro di massa è del 100%. Nel sistema del laboratorio, il valore della polarizzazione dipende dall'energia iniziale del pione e dall'angolo di decadimento del muone: per il fascio di COMPASS la polarizzazione media longitudinale è dell'ordine di 80%.

Prima di entrare nella sala dell'esperimento, il fascio percorre ancora  $\sim 400$  m: lungo questo percorso, una serie di magneti deflettori seleziona la carica e il momento, e alcuni magneti focalizzatori ne diminuiscono l'estensione spaziale. Nella zona sperimentale le dimensioni del fascio sono di  $8 \times 8 \text{ mm}^2$  ( $\sigma_x \times \sigma_y$ ) e la sua divergenza è di  $0.4 \times 0.8 \text{ mrad}^2$  ( $\sigma_x \times \sigma_y$ ). Il momento ha una RMS del 5%. Il fascio risultante è caratterizzato da un'alta intensità, corrispondente a  $2 \cdot 10^8 \mu/\text{spill}$ ; il ciclo utile è di 4.8 s/16.8 s, quindi il numero di muoni al secondo è  $\sim 4 \cdot 10^7 \mu/\text{s}$ .

Un'altra caratteristica del fascio di muoni così ottenuto è il cosiddetto *alone* dovuto ai muoni non sufficientemente deflessi od assorbiti. Nella zona del bersaglio, l'*alone vicino*, fino a circa 15 cm dalla linea di fascio, ha una intensità per unità di superficie tre ordini di grandezza inferiore rispetto a quella massima, mentre l'*alone lontano* ha una intensità per unità di superficie cinque ordini di grandezza inferiore e si estende fino a circa 3 m dalla linea di fascio.

Prima che il fascio entri nella sala dell'esperimento, il momento di ogni muone è misurato da una stazione dedicata, la Beam Momentum Station (BMS), illustrata in figura 2.3; essa è formata da una serie di magneti deflettori (indicati come B6 nella figura 2.3), che separa le particelle in base al loro momento, e da quattro odoscopi di contatori a scintillazione caratterizzati da ottima risoluzione temporale ( $\sim 300$  ps), in modo da fornire una misura del momento del fascio evento per evento, (la risoluzione in momento è  $\lesssim 1\%$  e l'efficienza di ricostruzione  $\sim 93\%$ ).

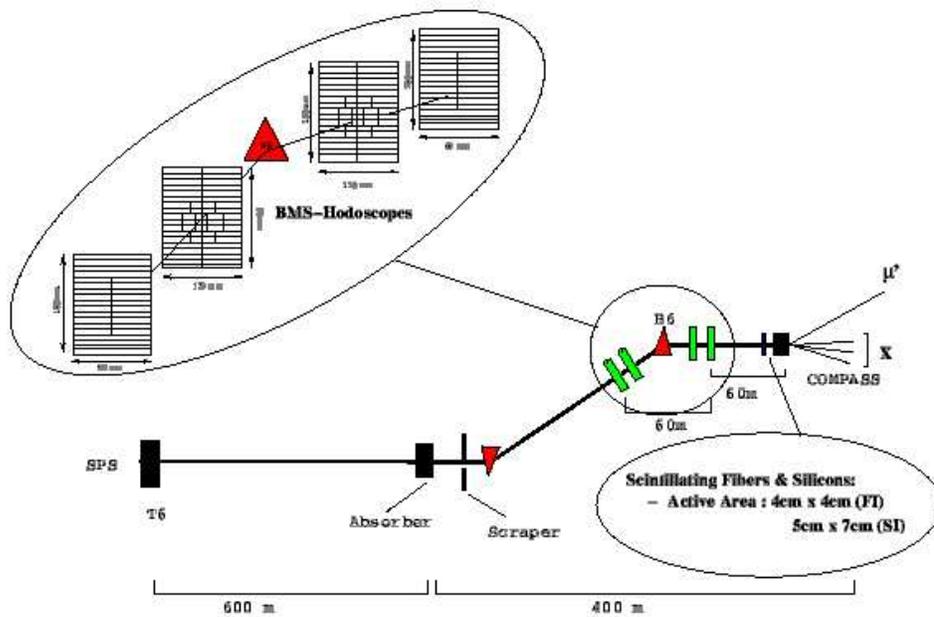


Figura 2.3: Percorso del fascio prima di entrare nella hall dell'esperimento; in evidenza la collocazione della stazione BMS.

### Bersaglio polarizzato

La scelta di un bersaglio polarizzato a stato solido è stata imposta dalla ingente statistica richiesta per la valutazione delle asimmetrie misurate e quindi dall'alta luminosità richiesta. La descrizione che segue e la figura 2.4 si riferiscono agli anni di presa dati 2002 e 2004.

Il materiale che costituisce il bersaglio, il  ${}^6\text{LiD}$ , è contenuto in due celle, raffigurate in figura 2.4, lunghe 60 cm e di diametro 3 cm, separate 10 cm l'una dall'altra. Il materiale scelto è caratterizzato da un alto fattore di diluizione, ovvero da una grande percentuale di nucleoni polarizzabili contenuti in esso. L'alto fattore di diluizione è un parametro essenziale per le misure di asimmetria che dipendono dall'orientazione degli spin dei nucleoni contenuti nel bersaglio rispetto alla direzione del fascio poiché gli eventi in cui i muoni interagiscono con i nucleoni non polarizzati costituiscono fondo per le analisi. Considerando il  ${}^6\text{Li}$  come uno stato legato [27] tra una particella  $\alpha$ , un protone e un neutrone, i nucleoni polarizzabili del  ${}^6\text{LiD}$  sono quattro (il neutrone, il protone e i due nucleoni del deuterio), in quanto la particella  $\alpha$  ha spin nullo; il fattore di diluizione del  ${}^6\text{Li}$  è dunque  $f = \frac{4}{8} = 0.5$ .

Per la polarizzazione del bersaglio viene usata la tecnica DNP (Dynamic Nuclear Polarization) che permette di trasferire la polarizzazione acquistata dagli elettroni immersi in un campo magnetico, ai nucleoni usando un campo di microonde. Questo processo richiede una temperatura inferiore ad un grado Kelvin e campi magnetici intensi ed omogenei.

Il sistema del bersaglio polarizzato di COMPASS include un solenoide superconduttore che produce un campo magnetico di 2.5 T lungo la direzione del fascio, con una omogeneità migliore di 20 ppm su una lunghezza di 1.5 m. Un dipolo permette inoltre di creare un campo magnetico di 0.42 T ortogonale alla direzione del fascio e orientato verso l'alto, che viene combinato con quello del solenoide nel ruotare gli spin del bersaglio per la presa dati per la misura dello spin trasverso.

Le basse temperature necessarie sono ottenute grazie ad un refrigeratore, indicato in figura 2.4, a diluizione di  ${}^3\text{He}/{}^4\text{He}$ . L'elio liquido riempie sia la zona del bersaglio sia lo spazio tra i cristalli di  ${}^6\text{LiD}$  di cui è costituito il bersaglio stesso, riducendo il fattore di diluizione a circa 0.40. La figura 2.5 mostra la polarizzazione misurata nelle due celle del bersaglio durante la presa dati dell'anno 2004.

Il fatto di raccogliere contemporaneamente dati con due celle con polarizzazione

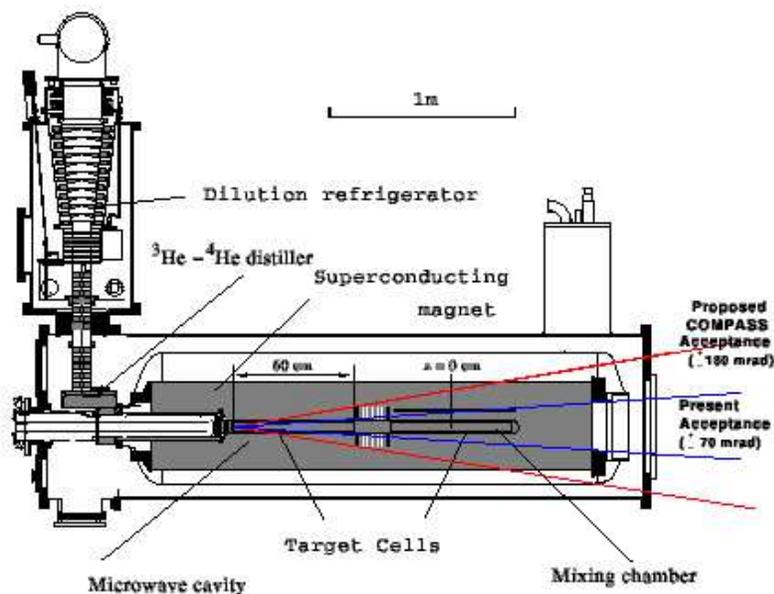


Figura 2.4: Il sistema criogenico e il supporto del bersaglio di COMPASS.

opposta, per le misure di asimmetrie con polarizzazione del bersaglio sia parallela che ortogonale (“spin trasverso”) alla direzione del fascio, permette di ridurre gli effetti sistematici dovuti a variazioni nel tempo dell’efficienza dello spettrometro. L’ulteriore possibilità di invertire tale polarizzazione in entrambe le celle (operazione eseguita ogni otto ore con polarizzazione longitudinale ed ogni quattro giorni con polarizzazione trasversa) permette poi di ridurre gli errori sistematici dovuti alla diversa accettazione delle due celle, nell’angolo polare, che varia da circa  $\pm 70$  mrad a  $\pm 170$  mrad. Il solenoide costruito per COMPASS che ha l’accettazione dello spettrometro è stato installato solo nel 2006.

Le misure considerate in questa tesi sono relative alle asimmetrie non dipendenti dallo spin. Per eliminare gli effetti della polarizzazione del bersaglio, i dati relativi alle due celle ed alle due orientazioni sono stati combinati in modo diverso rispetto all’analisi “standard” (come verrà spiegato nel capitolo 3).

## I rivelatori di traccia

Le varie stazioni di rivelatori per la ricostruzione delle tracce rilasciate dalle particelle cariche sono distribuite lungo tutta la lunghezza dello spettrometro COMPASS.

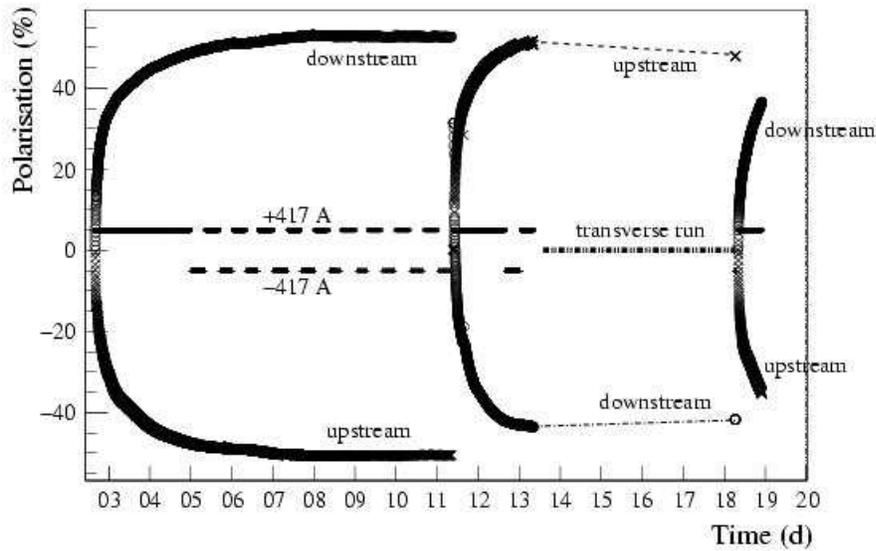


Figura 2.5: Tipico andamento della polarizzazione media nelle celle upstream e downstream (mostrato per 20 giorni del 2004); dopo l'undicesimo giorno la polarizzazione è invertita cambiando la frequenza delle microonde.

Ciascuna stazione consiste in un gruppo di rivelatori dello stesso tipo, situati molto vicini (rispetto alla coordinata  $Z$ ) lungo la linea del fascio, posti in modo da misurare diverse proiezioni della posizione della traccia. Differenti tecnologie di rivelazione sono state adoperate per far fronte alle necessità dell'esperimento.

Il flusso di particelle che investe la regione coperta dall'intera accettazione dello spettrometro varia di 5 ordini di grandezza. Nella zona lungo il fascio e vicino al bersaglio, dove il flusso è maggiore, i rivelatori devono essere caratterizzati da un'eccellente risoluzione spaziale e capaci di sostenere un'alta frequenza di particelle. Inoltre la quantità di materiale presente in questa regione deve essere minima per ridurre gli effetti dovuti alla diffusione multipla ed alle interazioni secondarie. In particolare, nella zona che precede il magnete SM1, dove il flusso incidente si arricchisce di particelle a bassa energia generate dalle interazioni secondarie che avvengono nella zona vicina al bersaglio, i limiti imposti da tali richieste diventano più stringenti. Nella regione più lontana dalla linea di fascio la risoluzione spaziale richiesta ai rivelatori è minore ma aumenta l'area che essi devono ricoprire.

Le tecniche di tracciamento utilizzate dai rivelatori appartenenti allo spettrometro

e posizionati a differenti distanze dalla linea di fascio sono scelte a seconda delle condizioni imposte dalla frequenza di particelle incidenti, dalla risoluzione spaziale e temporale, e dall'estensione della superficie che deve essere ricoperta dai materiali sensibili e dalla strumentazione elettronica.

Una prima classificazione dei rivelatori si può basare sulla grandezza della zona che ricoprono, e la distanza di questa dall'asse del fascio; in questo senso distinguiamo tra VSAT (*Very small Area tracker*), SAT (*Small Area Tracker*) e LAT (*Large Area Tracker*).

- **Tracciatori a piccolissimo angolo** (Very Small Area Trackers, VSAT).

Questi rivelatori ricoprono un'area attorno alla linea di fascio di  $\sim 5 \times 5 \text{ cm}^2$ ; la loro regione attiva è attraversata dal fascio stesso e sono destinati quindi a misurare la traiettoria del muone incidente e quella del muone diffuso. I VSAT sono caratterizzati dalla capacità di operare a flussi intensi, da un'alta risoluzione temporale e da un'ottima risoluzione spaziale per poter misurare piccoli angoli di diffusione.

Nell'apparato si trovano due tipi di VSAT, gli odoscopi di fibre scintillanti (SciFi) e i rivelatori a microstrip di Silicio:

- **Fibre scintillanti** di diametro di 0.5 o 1 mm sono sistemate in diversi strati, sovrapposti per ridurre le zone morte, a formare i piani di rivelazione come mostrato in figura 2.6; diverse coordinate si misurano utilizzando piani di fibre scintillanti disposti con differente orientazione. Lo spessore di materiale scintillante che le particelle attraversano è di  $\sim 3\text{-}3.5 \text{ mm}$ . La risoluzione temporale varia da 350 a 550 ps, quella spaziale 130 o 250  $\mu\text{m}$ ;
- i rivelatori a **microstrip di Silicio**, anch'essi collocati prima e dopo il bersaglio, hanno la miglior risoluzione spaziale dello spettrometro, 14  $\mu\text{m}$ ; la risoluzione temporale è dell'ordine di 3 ns; si misurano quattro coordinate corrispondenti a quattro diverse orientazioni delle strip.

- **Tracciatori a piccolo angolo** (Small Area Trackers, SAT).

La regione che coprono è di  $\sim 40 \times 40 \text{ cm}^2$  attorno alla direzione del fascio, con una zona morta di  $\sim 5 \times 5 \text{ cm}^2$ . Nell'apparato, oltre ai contatori a scintillazione

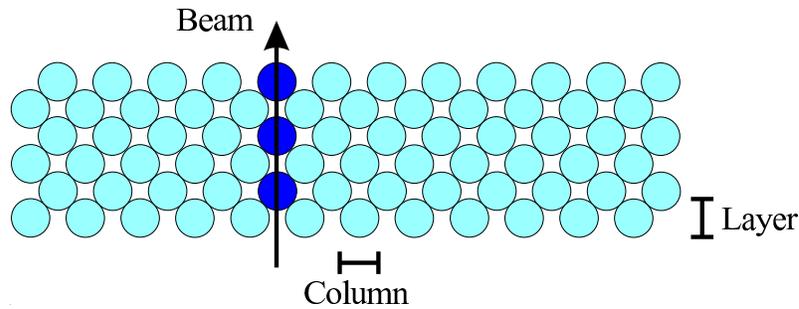


Figura 2.6: Configurazione di un piano composto dalle fibre scintillanti (il numero di strati di fibre per ogni piano è 8, 12 o 14 a seconda della stazione a cui questo appartiene).

delle quattro stazioni che formano la Beam Momentum Station, si trovano due tipi di SAT di nuova concezione, le MICROMEGAS (MICRO MESH Gaseous Structure) e le GEM (Gas Electron Multiplier): entrambi sono rivelatori a gas, in cui la zona di conversione e quella di moltiplicazione sono separate, caratterizzati da un'efficienza pari a  $\sim 97\%$ , da una buona risoluzione spaziale e temporale.

- Nelle **MICROMEGAS**, il cui principio di funzionamento è illustrato in figura 2.7, la risoluzione spaziale è di  $70 \mu\text{m}$ , quella temporale di  $10 \text{ ns}$ .
- Nelle **GEM**, il cui principio di funzionamento è illustrato in figura 2.8, la risoluzione spaziale è di  $50 \mu\text{m}$ , quella temporale di  $12 \text{ ns}$ .

- **Tracciatori a grande angolo** (Large Area Trackers, LAT).

Le zone sensibili di questi rivelatori sono dell'ordine di qualche  $\text{m}^2$ . Essi si dividono in camere a deriva (DC, Drift Chamber) e *straw tube* nel primo spettrometro, e in camere proporzionali multifilo (MWPC, Multi Wire Proportional Chamber) e altre due camere a deriva di ancora maggiori dimensioni (W4-5) nel secondo spettrometro.

- Nell'apparato si trovano tre stazioni di **DC** poste prima e subito dopo SM1, ognuna in grado di misurare otto coordinate; il tempo di deriva massimo è di  $70 \text{ ns}$ , e la risoluzione spaziale di circa  $170 \mu\text{m}$ . Ciascuno dei

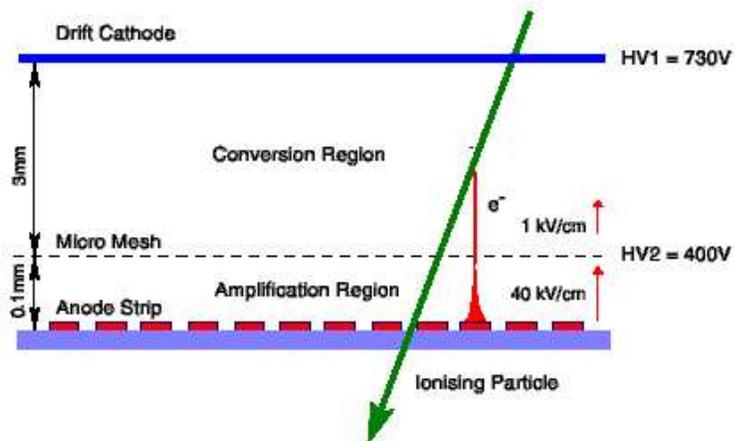


Figura 2.7: Schema del principio di funzionamento dei rivelatori MICROMEGA.

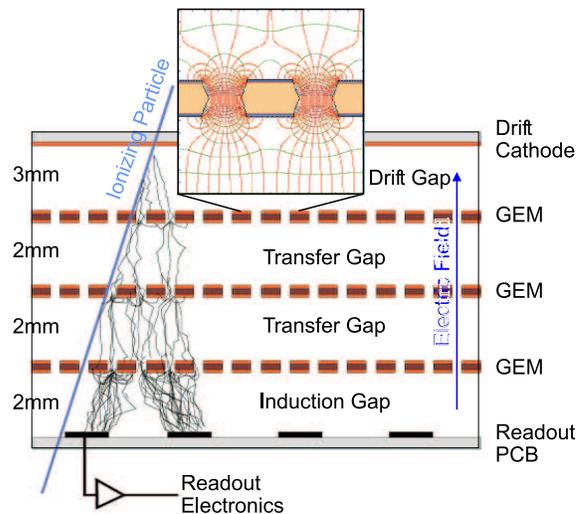


Figura 2.8: Schema del principio di funzionamento dei rivelatori GEM.

piani che le compongono ha un'area di  $180 \times 127 \text{ cm}^2$  ed è caratterizzato da un'efficienza del 95%.

- Le stazioni di *straw tube* consistono in sei insiemi di doppi strati di tubi di kapton alluminati disposti con tre orientazioni diverse, con diametro di 6 mm nella zona più interna, e di 10 mm in quella più esterna; la superficie coperta è di  $3.2(3.6) \times 4.6(4.1) \text{ m}^2$ . La risoluzione spaziale varia da 200 a 300  $\mu\text{m}$ .
- Le MWPC sono caratterizzate da una distanza tra i fili anodici di 2 mm e da una risoluzione spaziale di 500  $\mu\text{m}$ .
- Le altre due camere a deriva dello spettrometro, **W4-5**, si trovano dopo SM2 e coprono un'area di  $5 \times 2.5 \text{ m}^2$ ; la loro risoluzione spaziale è di  $\sim 500 \mu\text{m}$ .

## II RICH

Il rivelatore RICH (Ring Imaging Cherenkov) è installato alla fine del LAS ed è in grado di identificare pioni, kaoni e protoni carichi in un ampio intervallo di momento, da sopra soglia Cherenkov (corrispondente a  $\sim 2.5 \frac{\text{GeV}}{c}$  per i pioni,  $\sim 9 \frac{\text{GeV}}{c}$  per i kaoni e  $17 \frac{\text{GeV}}{c}$  per i protoni) fino a  $\sim 50 > \frac{\text{GeV}}{c}$ . I fotoni prodotti nel gas radiatore ( $\text{C}_4\text{F}_{10}$ ) per effetto Cherenkov, vengono focalizzati da specchi sferici con raggio di curvatura di 6.6 m sulle superfici di due fotorivelatori posti al di fuori dell'accettanza dello spettrometro per non perturbare le traiettorie delle particelle. I fotorivelatori sono MWPC, equipaggiate con un catodo segmentato ricoperto con uno strato di Ioduro di Cesio che funge da materiale fotoconvertitore.

## I calorimetri

I calorimetri adronici presenti nell'apparato hanno una struttura "a sandwich": in HCAL1 strati di Fe con spessore di 25 mm si alternano a strati di 5 mm di scintillatore plastico, mentre in HCAL2 si trovano strati di 16 mm di Pb e di 4 mm di scintillatore plastico. Essi misurano l'energia degli adroni prodotti in bersaglio e partecipano al trigger.

Entrambi i calorimetri vengono calibrati all'inizio del periodo di presa dati analizzando lo spettro di energia prodotto dai muoni dell'alone e l'andamento lineare

dell'energia misurata dal calorimetro rispetto al momento misurato dai rivelatori traccianti è mostrato in figura 2.9, sia per HCAL1 che per HCAL2. L'efficienza dei calorimetri comprende anche la fase di ricerca del “cluster” e dipende dall'energia della particella.

Il calorimetro elettromagnetico (ECAL2) è composto da  $64 \times 48$  moduli costituiti, ciascuno, da un parallelepipedo di vetro al piombo di dimensioni  $38 \times 38 \times 450 \text{ mm}^3$ . Esso misura l'energia degli sciami elettromagnetici; le informazioni provenienti da ECAL2 non vengono utilizzate nell'analisi standard.

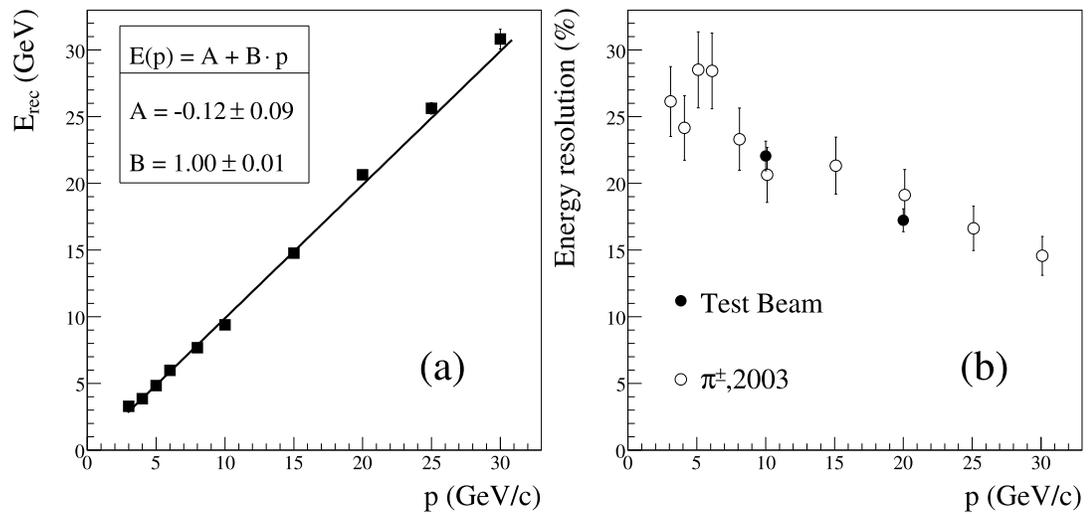


Figura 2.9: Nella figura a sinistra (a) è mostrata l'andamento dell'energia depositata in HCAL1 rispetto al momento di una particella ricostruita ed identificata come adrone (entrambe le quantità sono ricostruite da CORAL). Nella figura a destra (b) è mostrato invece l'andamento della risoluzione in energia sempre in funzione del momento. Le energie sono misurate in GeV ed i momenti in GeV/c.

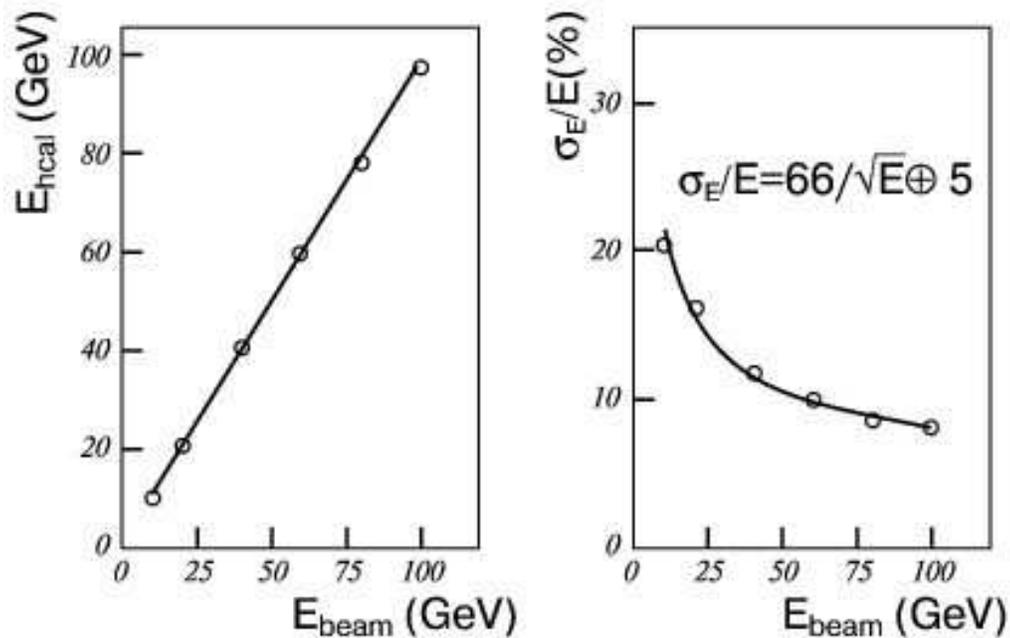


Figura 2.10: Nella figura a sinistra è mostrato l'andamento dell'energia depositata in HCAL2 rispetto al momento del fascio usato per la calibrazione. Nella figura a destra è mostrato invece l'andamento della risoluzione in energia. Le energie sono misurate in GeV

## I rivelatori di muoni

I rivelatori di muoni MW1 e MW2 si trovano alla fine del LAS e del SAS rispettivamente. Sono composti da due stazioni di tracciatori a monte e a valle di un blocco di materiale assorbitore: 60 cm di Fe per MW1 e 2.4 m di cemento per MW2. Le stazioni di rivelatori ricoprono una grande superficie, di area  $\sim 4 \times 2 \text{ m}^2$ . Essi permettono di distinguere i muoni dalle altre particelle che costituiscono il rumore (principalmente adroni provenienti dai calorimetri) e che vengono fermate dal materiale assorbitore.

## Il sistema del Trigger

Dato l'alto flusso di particelle incidenti sullo spettrometro, il sistema di trigger deve essere in grado di selezionare gli eventi di diffusione dal fondo, con alta efficienza ed un tempo di decisione inferiore a 500 ns. Gli elementi principali dai quali è costituito sono odoscopi di contatori a scintillazione posizionati prevalentemente nel SAS, i calorimetri adronici ed il sistema del Veto; una rappresentazione schematica del Trigger è illustrata in figura 2.11.

Gli eventi utili sono caratterizzati da  $Q^2$  che va da  $\approx 0$ , per la misura di  $\Delta G/G$ , fino al massimo permesso dalla cinematica, (gli eventi DIS hanno un  $Q^2 > 1 \text{ (GeV/c)}^2$ ). La logica del Trigger varia a seconda dei diversi intervalli di  $Q^2$  in cui opera. Per gli eventi con  $Q^2 > 0.5 \text{ (GeV/c)}^2$  il Trigger si basa principalmente sulla rivelazione del muone diffuso che avviene in due odoscopi di scintillatori orizzontali, per verificare che sia compatibile con una traccia proveniente dal bersaglio.

Per sopprimere la principale fonte del rumore, dovuto ai muoni dell'alone, è utilizzato anche un sistema di Veto costituito da due contatori a scintillazione posizionati a monte del bersaglio, forati nella parte centrale dove passa il fascio. Se uno dei rivelatori appartenenti al Veto dà un segnale l'evento non viene acquisito, come illustrato in figura 2.12.

Gli odoscopi del Trigger sono stati divisi in quattro sottosistemi costituiti da due stazioni di odoscopi ciascuno: IT (H4I, H5I); LT (H4L, H5L); MT (H4M, H5M); OT (H3O, H4O). Esse permettono di coprire diverse regioni in  $Q^2$ , come da figura 2.13.

Per i Trigger IT (Inner trigger), LT (Ladder Trigger), MT (Middle Trigger), in aggiunta alle informazioni date dagli odoscopi viene richiesto un deposito di ener-

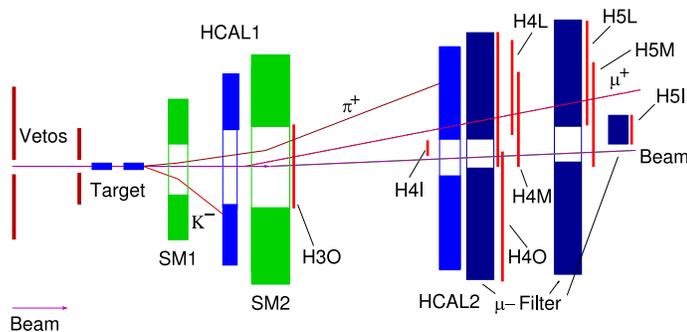


Figura 2.11: Schematica panoramica sulle principali componenti del trigger.

gia, in uno dei calorimetri adronici, che sia al di sopra del rilascio energetico medio da parte di particelle che possono, altrimenti, causare un segnale di trigger spurio. Infatti a causa dell’alta frequenza ( $\sim 2 \cdot 10^7/s$ ) di muoni dell’alone che attraversano un calorimetro il tempo medio tra due successivi muoni è inferiore alla finestra temporale in cui il calorimetro raccoglie il segnale. Quindi si ha il passaggio “contemporaneo” di due (o più) particelle che, pur essendo minimo ionizzanti, rilasciano un’energia totale comparabile con quello causato da un adrone. Viene quindi impostato un valore di soglia, ottenuto come compromesso tra la purezza del trigger e la sua efficienza, per i diversi tipi di trigger: è più basso nel caso dei trigger in cui il segnale proveniente dal calorimetro è in coincidenza con quello proveniente dagli odoscopi (e quindi viene rilevato anche il muone diffuso) rispetto al trigger in cui si richiede il segnale dal solo calorimetro (CT).

Infine, è stato utilizzato anche un “trigger calorimetrico” (CT) che richiede solo il segnale proveniente da uno dei due calorimetri adronici, ma la soglia energetica è maggiore rispetto a quella impostata per i trigger “misti”. Come si può vedere nella figura 2.13 la maggiore estensione “trasversale” dei calorimetri, in particolare HCAL1, permette di raggiungere una regione di  $Q^2$  proibita agli odoscopi.

### Elettronica di front-end e acquisizione dati

L’esperimento COMPASS ha un sistema di acquisizione dati ottimizzato che permette di registrare una quantità di dati che raggiunge i 600 TByte all’anno raccolti da 250K canali di lettura dei rivelatori con una rate di acquisizione di  $\sim 100K$  eventi al secondo.

Le informazioni prodotte dai rivelatori vengono digitalizzate dall’elettronica di

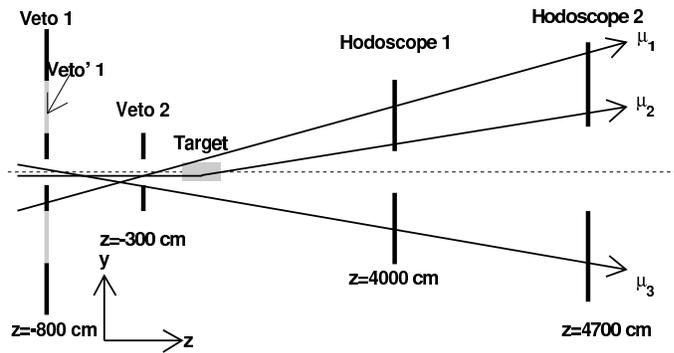


Figura 2.12: Funzionamento schematico del Veto: solo la traccia  $\mu_2$  soddisfa le condizione dei trigger inclusivi.

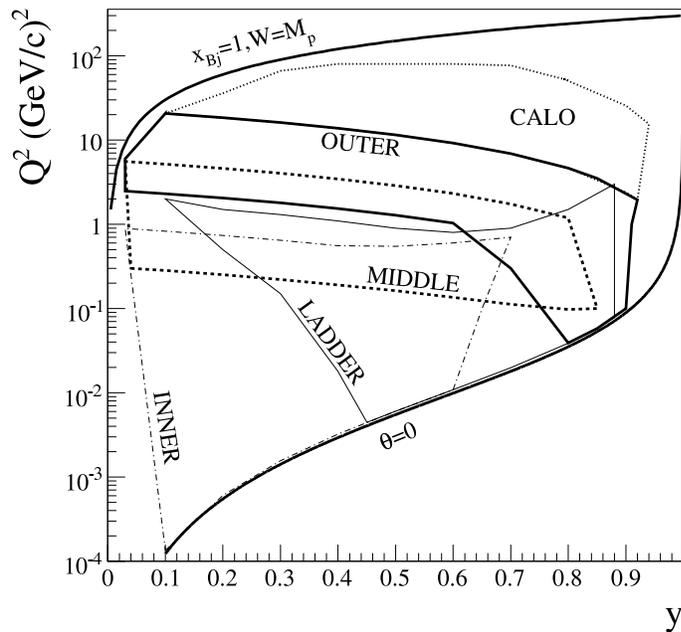


Figura 2.13: La copertura cinematica (in  $y$  e  $Q^2$ ) dei quattro sottosistemi costituiti dai soli odoscopi: Inner Trigger (IT), Ladder Trigger (LT), Middle Trigger (MT), Outer Trigger (OT) e dal Calorimetric Trigger (CT).

front-end e trasportate tramite fibre ottiche a PC dedicati chiamati *read-out buffer* [31] (ROB), dove avviene un temporaneo immagazzinamento dei dati. In uscita dai ROB si ha un flusso continuo di dati che vengono trasferiti ad altri PC, detti *event builder*, che hanno il compito di combinare le informazioni relative ad uno stesso evento, ovvero di “costruire l’evento” che viene eventualmente filtrato da un algoritmo “on-line” basato sulla possibilità di ricostruire il muone incidente.

A livello di event builder ha luogo anche l’estrazione dei metadati con i quali gli eventi sono registrati in una base di dati ORACLE, che consente di disporre di un successivo accesso avanzato ai dati. Infine, gli eventi vengono trasportati attraverso la rete locale verso il centro di calcolo del CERN dove vengono memorizzati temporaneamente su disco e migrati su nastri magnetici sotto il controllo di CASTOR (CERN Advanced STORage manager); in una tipica giornata di presa dati si raccolgono  $\sim 3$  TB di dati, corrispondenti a 25 cassette da 120 GB l’una.

### Controllo dei rivelatori e monitoraggio

Per garantire la qualità dei dati acquisiti diversi aspetti dell’esperimento vengono monitorati in tempo reale: l’elettronica di *front end* e la catena di lettura dati, la stabilità del fascio ed il conteggio del trigger.

Il DCS (Detector Control System) monitora i principali parametri che regolano il funzionamento dei vari rivelatori quali l’alimentazione, la pressione e le miscele dei gas, le temperature ed i campi magnetici e segnala la presenza di valori diversi da quelli impostati tramite un allarme sonoro. È inoltre possibile effettuare eventuali modifiche di alcuni di questi parametri tramite una interfaccia grafica *user friendly*, riducendo al minimo le situazioni in cui è necessario accedere all’interno dell’area dello spettrometro.

Un continuo ed automatico monitoraggio del fascio e del trigger viene effettuato via software ed i valori dei parametri letti vengono immagazzinati in un *DataBase* locale.

Oltre al software che mostra lo stato della *readout chain* e rileva eventuali errori generati dall’inconerenza del formato della struttura dei dati il pacchetto software COOL (COMPASS Object Oriented Online) [30] permette di analizzare i flussi di dati *raw* in tempo reale. Con l’interfaccia grafica fornita da questo programma è possibile ottenere istogrammi di quantità, come numero di canali che hanno dato

Periodo	Sottoperiodo	Pol. cella U	Pol. cella D
1	1 (W33)	+51%	-44%
	2 (W34)	-42%	+43%
2	1 (W35)	-46%	+47%
	2 (W36)	+50%	-43%

Tabella 2.1: Valori della polarizzazione misurati nelle celle upstream e downstream e relativi ai diversi sottoperiodi di presa dati del 2004 con spin trasverso. Accanto ai sottoperiodi sono indicate le settimane a cui corrispondono (per esempio W33 significa che è la settimana numero 33 a partire dall'inizio del 2004)

segnale, ampiezze dei segnali e tempi, riferite ai piani dei diversi rivelatori. La distribuzione campione è sovrapposta ad ogni istogramma, in modo da avere una veloce diagnostica del comportamento del rivelatore in questione.

## 2.3 Dati raccolti

Negli anni 2002, 2003, 2004 sono stati raccolti dati con bersaglio ( ${}^6\text{LiD}$ ) polarizzato longitudinalmente e trasversalmente (nel 20% del tempo) e con fascio composto da muoni positivi di 160 GeV/c. Per questo lavoro di tesi sono stati usati i dati raccolti con spin trasverso nel 2004 poiché, oltre ad avere una maggiore statistica rispetto agli anni precedenti, sono stati scrutinati nelle analisi per le misure delle asimmetrie di Collins e Sivers dalle quali sono emerse la buona qualità e la stabilità dei dati. I dati sperimentali (raw) sono stati raccolti in due periodi distinti, entrambi composti da due sottoperiodi (che corrispondono ad una settimana di presa dati) caratterizzati da polarizzazioni opposte nelle due celle del bersaglio con valori mostrati in tabella 2.1. Altri dati con fascio di muoni e bersaglio di  ${}^6\text{LiD}$  sono stati raccolti nel 2006, mentre nel 2007 si stanno raccogliendo dati con bersaglio di  $\text{NH}_3$  polarizzato longitudinalmente e trasversalmente.

## Capitolo 3

# La ricostruzione degli eventi e l'analisi dei dati

Per la loro mole e complessità i dati raw non vengono direttamente utilizzati per l'analisi fisica, ma sono sottoposti ad una processatura per la ricostruzione degli eventi, che avviene al centro di calcolo del CERN nella *COMPASS Computing Farm*, utilizzando una potenza di calcolo pari a 600K SPECint2000<sup>1</sup>.

La ricostruzione degli eventi viene realizzata da CORAL [32] (COmpass Reconstruction and AnaLysis), un software scritto in C++ con programmazione orientata agli oggetti, sviluppato all'interno della collaborazione COMPASS, la cui architettura è costituita da un framework centrale che fornisce le funzionalità di base come l'accesso alle necessarie librerie esterne e da un insieme di moduli indipendenti spesso intercambiabili, che hanno la funzione d'interfacciarsi con i dati o con specifiche classi per la ricostruzione degli eventi di fisica stesse. Tramite basi di dati vengono gestite le informazioni specifiche dell'apparato sperimentale come la geometria (dimensione, posizione e caratteristiche dei rivelatori) e le calibrazioni dei rivelatori (condition database), oltre, naturalmente, al reperimento dei dati stessi. L'architettura modulare permette d'integrare e/o adattare in modo agevole nel framework principale eventuali classi che offrano nuove funzionalità alla ricostruzione o accesso a diverse fonti di dati. Il flusso di ricostruzione dei dati (reali o simulati) di input è rappre-

---

<sup>1</sup>La grandezza SPECint2000 (d'ora in poi indicata come SI2K) è un'unità di misura arbitraria delle prestazioni ("benchmark") che consente di confrontare le prestazioni di diversi sistemi sulla base di un punteggio ottenuto nella computazione di operazioni con interi. Periodicamente viene aggiornato (per meglio adattarlo all'evoluzione dei sistemi sul mercato) e SPECint2000 è stato recentemente sostituito da SPECint2007. [<http://www.spec.org/>]

sentato in figura 3.1, dove i rettangoli identificano le varie classi responsabili della decodifica dei dati, della ricostruzione delle tracce e dei vertici e dell'identificazione delle particelle che avvengono durante la processatura. Le quantità ricostruite per ogni evento costituiscono l'output di CORAL e vengono scritte in file chiamati Data Summary Tapes (DST), in formato binario proprietario, ed in file mini DST (mDST) in formato ROOT [29]. In quest'ultimo caso, prima della scrittura, viene applicata una prima selezione agli eventi. La stessa procedura, usata per i dati reali, viene seguita per la ricostruzione dei dati generati dalla simulazione Monte Carlo, descritta nel paragrafo 3.3. Nella iniziale fase di "decoding" viene estratta l'informazione (tradotta in coordinate "fisiche") riguardo allo specifico "canale" del rivelatore che ha dato segnale e dopo la fase di "clustering", dove le informazioni vengono raggruppate da un algoritmo e viene calcolato il punto del rivelatore in cui la particella è passata (hit), vengono ricostruite le informazioni relative alle particelle cariche e neutre che vengono infine identificate. Le informazioni che invece provengono dai tracciatori dello spettrometro vengono utilizzate per ricostruire le traiettorie delle particelle cariche mentre, a partire da quelle provenienti dal calorimetro elettromagnetico sono ricostruite le particelle neutre. La separazione tra muoni ed adroni è basata sui "cluster" relativi ai calorimetri adronici e l'identificazione degli adroni viene svolta dal software (RICH-1) responsabile della ricostruzione delle informazioni ottenute dal RICH.

Fra le diverse quantità fisiche ricostruite (a partire dalle informazioni raccolte dalle risposte dei rivelatori) in CORAL e scritte su mDST ci sono informazioni relative alla direzione ed al momento del muone incidente, di quello diffuso e degli adroni prodotti nell'interazione, la cui posizione viene pure registrata. Nel prossimo paragrafo vengono descritte le tre principali selezioni effettuate durante la processatura affinché un evento venga ricostruito e scritto sui mDST: l'identificazione del muone incidente, l'identificazione del muone diffuso e la definizione di vertice primario (quello in cui avviene l'interazione tra il muone incidente ed il bersaglio). L'analisi fisica è di norma effettuata a partire dai mDST replicati nei vari istituti utilizzando PHAST [33] (PHysics Analysis Software Tools), il programma di analisi di COMPASS (anch'esso scritto in C++).

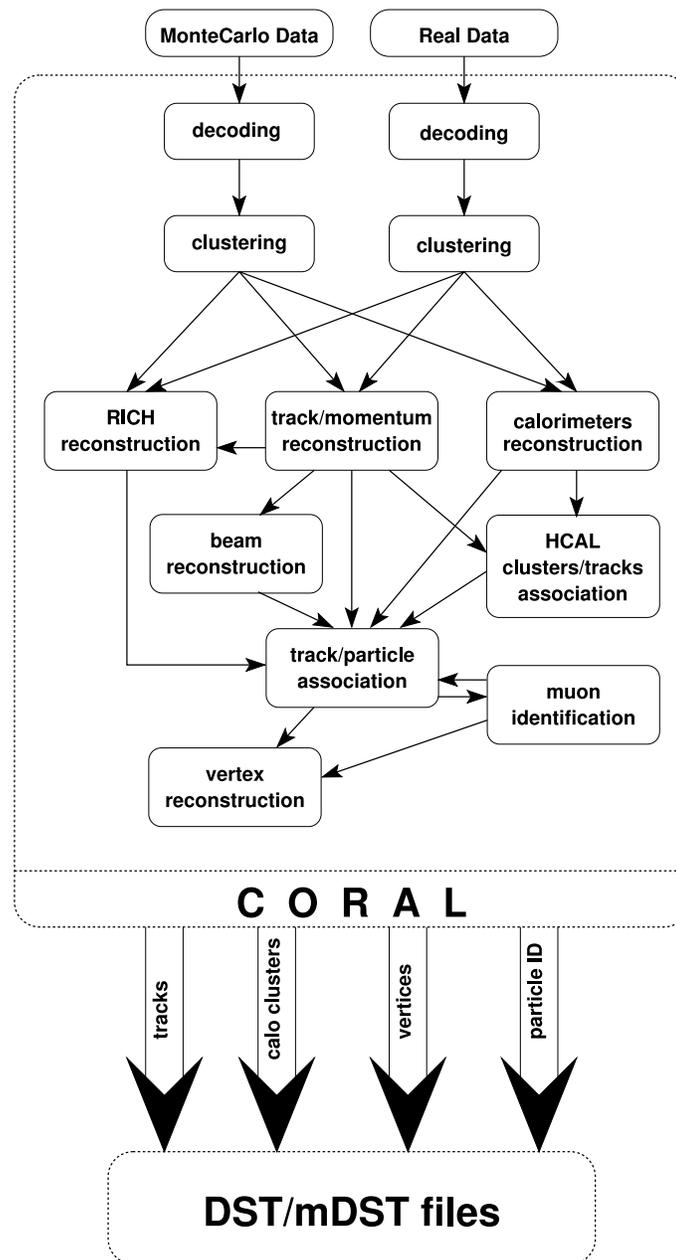


Figura 3.1: Rappresentazione schematica della ricostruzione degli eventi in CORAL, il software di ricostruzione di COMPASS. I dati (reali o simulati) ne costituiscono l'input e le informazioni prodotte relative ad ogni evento vengono scritte su file DST o mDST, (i rettangoli rappresentano le classi che effettuano le diverse funzioni che costituiscono la processatura: decodifica; “clustering”, ricostruzione delle tracce; l'identificazione delle particelle e la ricostruzione dei vertici di interazione).

### 3.1 La selezione degli eventi

Gli eventi utilizzati per l'analisi compiuta nel lavoro di tesi (contenuti in mDST) sono gli stessi che sono stati selezionati per la misura degli effetti Collins e Sivers pubblicata in Nuclear Physics [34]. Essi provengono da una selezione dei dati (del 2004 in configurazione di spin trasverso) per cui c'era stabilità dei rivelatori presenti nello spettrometro (e quindi era stabile l'efficienza di ricostruzione [34]) e sono relativi a quattro settimane di presa dati in cui il bersaglio era polarizzato trasversalmente. Ulteriori selezioni, relative al fisica del DIS, sono state applicate ai mDST ed i file infine analizzati occupano uno spazio di memoria pari a  $\sim 50GB$  (mentre i file contenenti i raw data del trasverso e ottenuti come output di CORAL dopo la processatura avvenuta al CERN, occupavano  $\sim 70TB$ ).

Nel seguito viene descritta l'identificazione del muone (incidente e diffuso) e degli adroni, e la definizione di vertice primario.

**Identificazione del muone incidente.** È considerata appartenente al fascio una traccia ricostruita nei rivelatori al silicio ed in quelli a fibre a monte del bersaglio, a cui sia associato un momento misurato nella BMS.

**Identificazione del muone diffuso.** Candidate sono tutte le tracce cariche positivamente, ricostruite dopo SM1 con un momento maggiore di 1 GeV/c, e la cui estrapolazione, sia all'entrata che all'uscita del bersaglio, disti meno di 5 cm dall'asse del bersaglio. Inoltre se l'evento è caratterizzato dai trigger basati sulle informazioni degli odoscopi, la traccia ricostruita deve intersecare un numero di scintillatori, appartenenti alle stazioni di odoscopi, sufficiente a dare il segnale di trigger. Qualora l'evento sia caratterizzato dal solo trigger calorimetrico, è richiesto che la traccia candidata sia associata ad un numero minimo di hits nei filtri muonici e che, quindi, la particella corrispondente abbia attraversato un ammontare di materiale sufficiente a minimizzare la contaminazione dovuta agli adroni. Ogni altra traccia ricostruita viene associata ad un adrone.

**Definizione di vertice primario.** Il punto di interazione è definito come il punto di minima distanza tra la traccia ricostruita come dovuta al  $\mu$  incidente e quelle associate al  $\mu$  diffuso ed agli adroni, se la distanza tra loro è minore di

una soglia di  $\sim 2-3$  mm (corrispondente a 10 deviazioni standard). Nel caso in cui, per un evento, più vertici sono ricostruiti, si definisce “vertice migliore” come quello con il maggior numero di tracce uscenti ed, in caso di parità, col minore valore di probabilità di  $\chi^2$  associato.

### 3.1.1 La selezione degli eventi DIS

I mDST adoperati contengono solo gli eventi a  $Q^2 > 1$  (Gev/c)<sup>2</sup> con almeno un vertice di interazione che abbia almeno una ulteriore traccia associata (ottenendo una riduzione di circa un fattore 100 sulla grandezza dei file iniziali che contenevano i “raw data”). Per migliorare la qualità degli eventi esaminati sono stati applicati ulteriori tagli in fase di analisi, descritti nel seguito.

Per assicurare lo stesso flusso del fascio incidente, su entrambe le celle è stato adottato un taglio che escluda gli eventi per cui la traccia del muone incidente non percorra entrambe le celle per tutta la loro dimensione longitudinale. L’ammontare di materiale attraversato dal muone diffuso deve essere superiore a 30 lunghezze di radiazione; ciò riduce ulteriormente la contaminazione di adroni (che, alle energie tipiche dell’esperimento sono meno penetranti), che quindi risulta trascurabile. Ogni traccia uscente dal vertice primario è associata ad un adrone solo se:

- almeno uno degli hit, ad essa associati, appartiene ad un rivelatore posizionato dopo il primo magnete, affinché venga utilizzato tutto il campo del magnete per calcolarne il momento;
- la quantità di materiale attraversato è inferiore a 10 lunghezze di radiazione;
- il momento trasverso  $P_T$ , rispetto alla direzione del fotone virtuale nel Gamma Nucleon System (GNS), è maggiore di  $0.1\text{GeV}/c$ , in modo da assicurare una buona risoluzione nell’angolo azimutale misurato;
- infine, la frazione  $z$  di energia portata dall’adrone, rispetto all’energia del fotone virtuale, calcolata nel laboratorio, è compresa tra 0.2 e 1.

Sono stati applicati due tagli geometrici alla posizione del vertice primario: uno sulla distanza tra il punto di interazione e l’asse nominale del fascio, per selezionare una zona del bersaglio uniforme, ed uno sulla coordinata  $Z_{vtx}$  lungo l’asse nominale

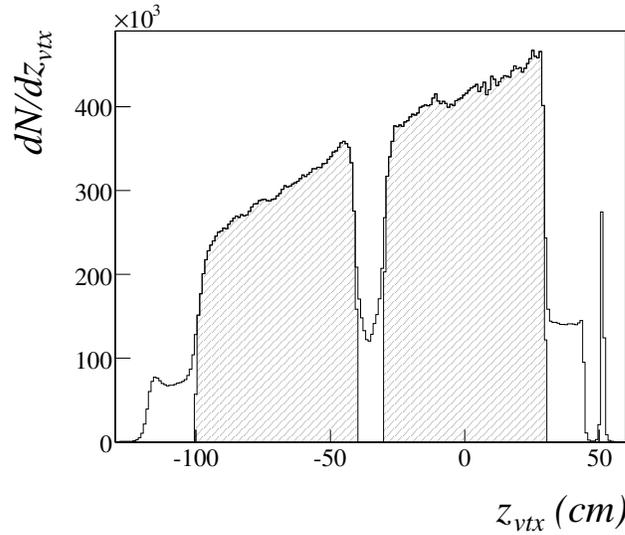


Figura 3.2: Distribuzione del numero di eventi in funzione della coordinata  $Z$  del vertice primario. È tratteggiata la zona corrispondente al taglio applicato per associare ogni evento ad una cella in modo non ambiguo.

del fascio, come illustrato in figura 3.2, per poter assegnare l'evento ad una cella in modo non ambiguo.

Infine sono stati adottati i tipici tagli cinematici associati al DIS:

- il taglio  $W > 5 \text{ GeV}/c^2$  esclude la regione delle risonanze nucleari;
- $y > 0.1$  per escludere la regione di diffusione elastica e  $y < 0.9$  per garantire che il muone diffuso vada nell'accettanza dello spattrometro.

A valle di tutti i tagli gli eventi rimasti sono circa il 20% di quelli presenti nei mDST.

Per questi eventi in figura 3.3 vengono mostrate la distribuzione del numero di adroni in funzione di  $Q^2$  ed in  $x$ , e la loro correlazione, e di  $y$ ,  $W$ ,  $z$  e  $P_T$ .

## 3.2 Misura dell'asimmetria

Il numero di eventi in funzione dell'angolo azimutale  $\phi$  dell'adrone calcolato nel GNS per le due celle e le due polarizzazioni può essere scritto come (considerando un solo periodo):

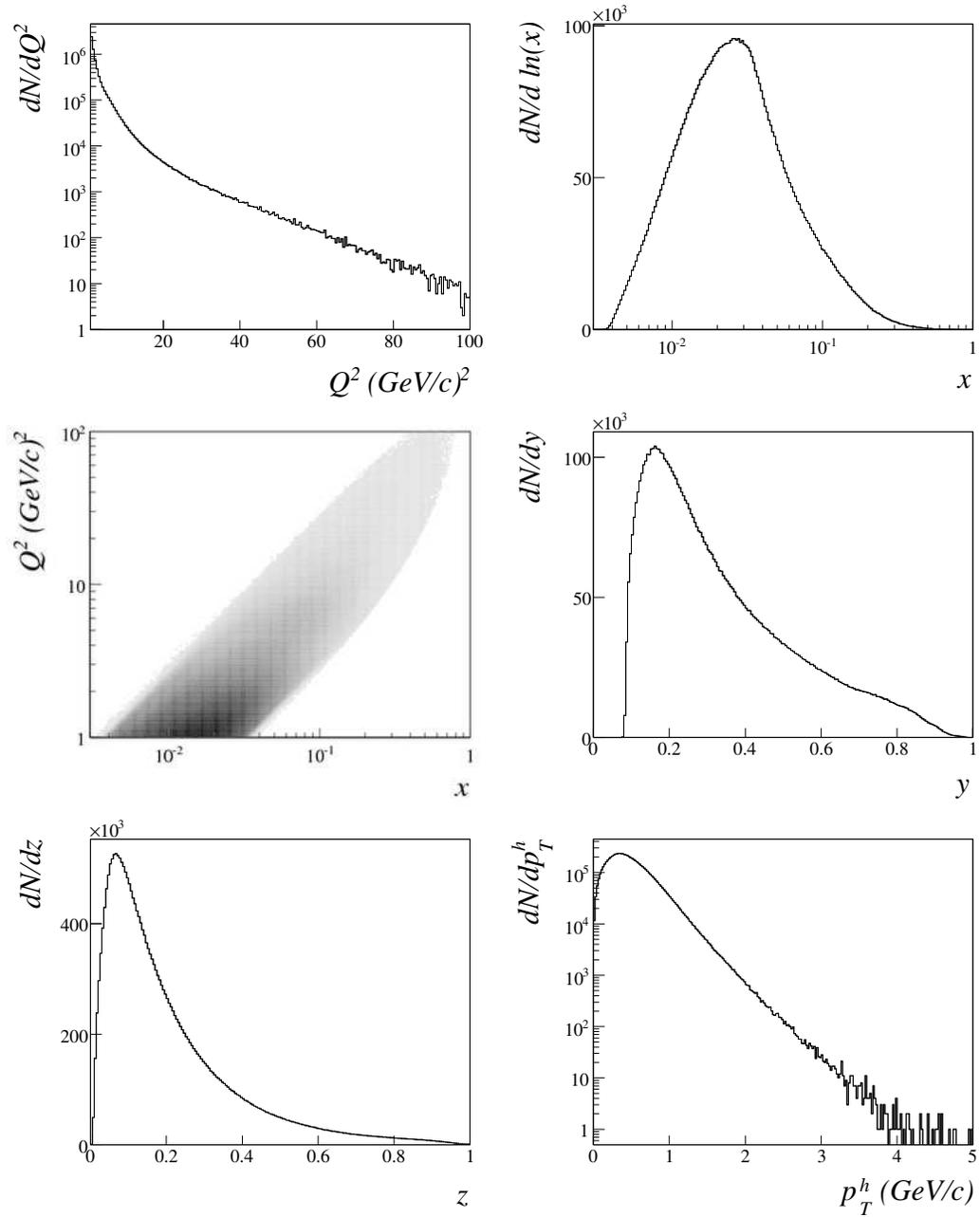


Figura 3.3: Distribuzioni delle principali quantità cinematiche ( $Q^2$ ,  $x$ ,  $y$ ,  $z$  e  $P_T$ ) ottenute dai dati dopo tutti i tagli standard dell'analisi.

$$N_{u,d}^{\pm} = B_{u,d}^{\pm} \cdot (\rho \cdot l) \cdot a_{u,d}(\phi) \cdot \sigma(\phi) \quad (3.1)$$

$$\sigma(\phi) = \sigma_0 \cdot (1 + X(\phi) \pm P_{u,d}^{\pm} \cdot Y(\phi)) \quad (3.2)$$

Dove:

- $B_{u,d}^{\pm}$  è il numero di muoni incidenti che incidono sul bersaglio ad una data polarizzazione, per esempio la cella upstream ( $u$ ) con polarizzazione positiva (+) e grazie al taglio sulle tracce dei muoni incidenti, è  $B_u^+ = B_d^-$  e  $B_u^- = B_d^+$ ;
- il prodotto  $\rho \cdot l$  è la densità di centri diffusori per unità di superficie, presenti nel bersaglio;
- $\sigma(\phi)$  è la sezione d'urto dipendente da  $\phi$ ,  $X$  è la parte indipendente dallo spin del bersaglio, mentre  $Y$  ne dipende e  $\sigma_0$  è la sezione d'urto totale non polarizzata;
- $a_{u,d}$  rappresenta l'accettanza dello spettrometro, costante rispetto al tempo, quindi la stessa per tutti i sottoperiodi presi in considerazione e include anche l'efficienza dell'apparato e della ricostruzione;
- infine,  $P_{u,d}^{\pm}$  è il valore assoluto della polarizzazione misurata nelle celle  $u$  e  $d$  durante un sottoperiodo.

Per isolare il termine  $Y(\phi)$  dipendente dallo spin e valutare quindi le asimmetrie di Collins e Sivers [34], viene utilizzata la quantità  $F = \frac{N_u^+ \cdot N_d^+}{N_u^- \cdot N_d^-}$  che non dipende dal fascio incidente e, in ipotesi del tutto ragionevoli, dall'accettanza  $a(\phi)$ . Invece, per la misura dell'asimmetria azimutale indipendente dallo spin, bisogna combinare queste quantità in modo da eliminare tale termine ed ottenere una grandezza proporzionale al solo termine  $(1 + X(\phi))$ . Ciò si ottiene con la quantità:

$$n = C \cdot \left( b \cdot P_u^- \cdot (P_d^+ + P_d^-) \cdot N_u^+ + P_u^+ \cdot (P_d^+ + P_d^-) \cdot N_u^- + \right. \\ \left. + P_d^- \cdot (P_u^+ + P_u^-) \cdot N_d^+ + b \cdot P_d^+ \cdot (P_u^+ + P_u^-) \cdot N_d^- \right) \quad (3.3)$$

dove  $b = \frac{B_u^-}{B_u^+} = \frac{B_d^-}{B_d^+}$ , stimato usando le misure del flusso del fascio relative a ciascun periodo di presa dati e  $C$  è una costante, ininfluenza dal momento che vengono valutate delle asimmetrie.

Si vede facilmente che  $n$  risulta essere proporzionale alla somma delle accettanze per le due celle ed al termine non dipendente dallo spin del bersaglio, cioè:

$$n(\phi) = C \cdot (a_u(\phi) + a_d(\phi)) \cdot (1 + X(\phi)), \quad (3.4)$$

dove  $C$  è una quantità indipendente da  $\phi$ . Per risalire, analizzando i dati dell'esperimento, alla distribuzione dell'angolo azimutale dell'adrone e quindi ottenere le asimmetrie azimutali che compaiono nell'equazione 1.50 del capitolo 1, è necessario stimare la quantità  $(a_u(\phi) + a_d(\phi))$  ovvero l'accettanza dell'apparato che può essere ottenuta tramite simulazione Monte Carlo come descritto nel paragrafo 3.3.

L'idea di base è di generare eventi che differiscano il meno possibile da quelli reali. Nella simulazione vengono inclusi tutti gli effetti fisici, dall'interazione del  $\mu$  nelle due celle del bersaglio alla rivelazione delle particelle prodotte, e la ricostruzione viene compiuta con la stessa catena di analisi (CORAL+PHAST) utilizzata per i dati reali. Viene calcolato il rapporto delle distribuzioni di adroni ricostruiti e generati:

$$\frac{N_R^{mc}(\phi)}{N_G^{mc}(\phi)} = \frac{N_0 \cdot f(\phi) \cdot \alpha(\phi)}{N_0 \cdot f(\phi)}. \quad (3.5)$$

Dove:

- l'indice  $mc$  indica che entrambe le distribuzioni sono state ottenute dalla simulazione Monte Carlo;
- $N_0$  è il numero di adroni (carichi) generati nell'urto nella regione cinematica corrispondente al DIS ( $Q^2, W, y$ ) e  $f(\phi)$  ne rappresenta l'eventuale dipendenza dall'angolo azimutale introdotta nella generazione dell'evento DIS;
- $\alpha(\phi)$  dipende dall'efficienza dei rivelatori, dall'efficienza del trigger e dall'accettanza angolare e fornisce la stima di  $a_u(\phi) + a_d(\phi)$  necessaria per estrarre le asimmetrie azimutali.

Infine,  $N = \frac{n(\phi)}{\alpha(\phi)}$  (dove  $n$  è definito nell'equazione 3.3) è la distribuzione, funzione dell'angolo azimutale dell'adrone, dal cui fit si ottengono le asimmetrie azimutali fisiche come descritto nel capitolo 5.

L'errore statistico su  $N$  e quindi sulle asimmetrie finali, dipende sia dall'errore statistico su  $n$ ,  $\sigma_n$  sia da quello su  $\alpha$ ,  $\sigma_\alpha$  secondo la formula espressa nell'equazione 3.6:

$$\sigma_N = N \cdot \sqrt{\left(\frac{\sigma_n}{n}\right)^2 + \left(\frac{\sigma_\alpha}{\alpha}\right)^2}, \quad (3.6)$$

ovviamente si vuole avere una produzione Monte Carlo che permetta di avere un'errore sulla stima dell'accettanza trascurabile rispetto a quello associato ad  $n$ . Assumendo, ragionevolmente, che i valori assoluti delle polarizzazioni  $p$  siano uguali e con errori trascurabili, che il numero di adroni ricostruiti sia lo stesso, cioè  $n_p = N_u^+ = N_u^- = N_d^+ = N_d^-$ , e che sia  $B_u^- = B_u^+ = N_\mu$ , si ottiene:

$$\begin{aligned} n &\simeq 4 \cdot p^2 \cdot n_p \cdot (b + 1), \\ \frac{\sigma_n}{n} &= \sqrt{\left(\frac{\sigma_{n_p}}{n_p}\right)^2 + \left(\frac{\sigma_b}{b+1}\right)^2} = \sqrt{\frac{1}{n_p} \cdot \left(1 + \frac{2 \cdot n_p}{N_\mu}\right)} \end{aligned} \quad (3.7)$$

Trascurando l'errore su  $b$  poiché il rapporto  $\frac{n_p}{N_\mu}$  è dell'ordine di  $10^{-6}$  si ha:

$$\frac{\sigma_n}{n} \simeq \sqrt{\frac{1}{n_p}}. \quad (3.8)$$

L'accettanza calcolata  $\alpha = \frac{N_R^{mc}}{N_G^{mc}}$  ha un errore di tipo binomiale in quanto il numeratore è un sottoinsieme del denominatore che viene considerato esatto ed  $\alpha$  può essere considerata la probabilità che un adrone vada nell'accettanza dell'apparato e venga ricostruito, mentre  $1 - \alpha$  la probabilità che l'adrone non venga ricostruito. In prima approssimazione il termine  $1 - \alpha$  può essere preso uguale ad uno e si ha:

$$\frac{\sigma_\alpha}{\alpha} = \frac{\sqrt{N_R^{mc} \cdot (1 - \alpha)}}{N_G^{mc} \cdot \alpha} \simeq \sqrt{\frac{1}{N_R^{mc}}}. \quad (3.9)$$

Se si vuole che il contributo all'errore statistico espresso nell'equazione 3.6 portato dal termine  $\frac{\sigma_\alpha}{\alpha}$  sia dell'ordine di quello dovuto al termine  $\frac{\sigma_n}{n}$  o minore, è necessario generare con il Monte Carlo un numero di eventi sufficiente ad avere un numero di adroni ricostruiti pari a:

$$N_R^{mc} = k^2 \cdot n_p, \quad (3.10)$$

dove  $k$  è il fattore che definisce il rapporto tra i due errori:  $\frac{\sigma_\alpha}{\alpha} = \frac{1}{k} \cdot \frac{\sigma_n}{n}$ . Una prima stima sul rapporto tra numero di adroni ricostruiti nel Monte Carlo ed il numero di eventi generato viene data alla fine di questo capitolo.

### 3.3 Gli eventi simulati

La tecnica Monte Carlo (MC) è definita come una simulazione statistica in cui si ha l'uso sistematico del campionamento con numeri casuali per la stima di una distribuzione incognita [35]. I metodi basati su tale tecnica vengono applicati nei casi in cui la dimensionalità o la complessità del problema rendono impraticabile o impossibile l'applicazione diretta di metodi analitici, è questo il caso di molte misure nel campo delle alte energie e dell'asimmetria qui studiata in quanto non è conveniente cercare di determinare analiticamente l'accettanza dell'apparato. Benché il processo fondamentale che descrive il processo di diffusione tra quark e leptone si basi su formule di sezioni d'urto ben definite nell'ambito del modello standard elettrodebole, l'inclusione delle PDF e delle correzioni di QCD perturbativa, rendono i calcoli analitici applicabili solo in casi semplificati. Inoltre la natura statistica delle interazioni tra particelle e materia e la complessità dell'apparato sperimentale rendono l'uso delle simulazioni MC indispensabile per ottenere la probabilità di ricostruzione degli adroni generati nel DIS.

#### 3.3.1 Programmi per la simulazione

In COMPASS la simulazione degli eventi fisici, a partire dall'urto elementare fino alla ricostruzione dell'evento all'interno dell'apparato, è costituita da una "catena" di programmi, specifici per COMPASS, eseguiti in serie ed il cui output viene letto come input dal programma successivo.

La generazione degli eventi DIS e la propagazione delle particelle generate attraverso l'apparato vengono eseguite in modo indipendente. Il programma scelto per l'implementazione della prima parte varia a seconda dell'effetto fisico che si vuole analizzare. In questo lavoro di tesi gli eventi da simulare sono eventi di DIS ed è quindi stato usato LEPTO<sup>2</sup>, un generatore di tipo Monte Carlo, che simula in modo completo l'urto profondamente inelastico tra leptone e nucleone basandosi sulla sezione d'urto descritta nel modello standard delle interazioni elettrodeboli, corretta con i processi perturbativi al leading order della QCD e con le parametrizzazioni delle PDF. Il processo di adronizzazione viene implementato adottando il modello

---

<sup>2</sup>Per simulare l'urto tra leptone e nucleone con  $Q^2 < 1 \left(\frac{GeV}{c}\right)^2$  viene usato PYTHIA, per la fisica dello charm AROMA.

fenomenologico a stringhe di LUND, dove è descritto come l'interazione forte generi stati finali adronici quando un costituente (quark o gluone) del nucleone si allontana da esso. Tale programma è molto usato nell'ambito della fisica delle alte energie e di fatto si è dimostrato essere il generatore di DIS migliore a disposizione per affidabilità ed accordo con i dati sperimentali. È inoltre possibile impostare la specie di leptone e la sua energia media, il bersaglio, i range cinematici ( $x$ ,  $Q^2$ , ecc.) e molti altri parametri, di cui, qui, sono stati usati i valori di default.

L'output generato da questo programma contiene, per ogni evento, tutte le particelle finali, con i relativi momenti ed energie, e viene poi letto da COMGEANT [37]: il programma responsabile della simulazione della propagazione delle particelle prodotte in DIS, all'interno dell'apparato di COMPASS. Esso è basato sul pacchetto di simulazioni GEANT 3.2.1 [38] e provvede sia al tracciamento di ogni particella generata simulando i principali processi fisici (per esempio per gli adroni vengono simulati: il decadimento, la diffusione multipla, la ionizzazione e la produzione di raggi  $\delta$ , interazioni adroniche, ecc.) che riguardano l'interazione tra la particella e le diverse parti dell'apparato attive (rivelatori) e no, sia alla simulazione della risposta dei rivelatori. In COMGEANT la coordinata Z del punto di interazione associato all'evento DIS viene generata uniformemente all'interno delle celle del bersaglio (dove non è implementato, però, il materiale liquido refrigerante,  $^3\text{He}/^4\text{He}$ ). Esso si avvale di una descrizione estremamente dettagliata dell'apparato (per esempio rivelatori, dimensioni, magneti, trigger, ecc.) con tutto il materiale (attivo e passivo) definito al meglio delle conoscenze e che viene letta in input da file (in formato *ffr*) aggiornati ogni anno per le variazioni di geometria dell'apparato. La risposta dei rivelatori viene simulata essenzialmente in base alla loro tipologia, risoluzione ed efficienza, queste ultime ottenute analizzando il comportamento dell'apparato reale. In input viene letto anche il file contenente la descrizione delle particelle appartenenti al fascio e per la quale vengono svolte periodicamente apposite misure. L'output principale di COMGEANT consiste in un file formato "zebra" (formato binario proprietario) che contiene sia il risultato della propagazione dell'evento nell'apparato sia le informazioni provenienti da LEPTO ed in un file contenente la descrizione (in forma semplificata) della geometria dei rivelatori dell'apparato. Entrambi i file vengono letti da CORAL che provvede prima alla digitalizzazione dei segnali delle particelle rilasciate nei rivelatori e poi alla ricostruzione degli eventi, in analogia con quanto

accade per i dati reali.

## 3.4 Test preliminari

Alcuni test preliminari sono stati eseguiti sia per verificare l'accordo tra i dati simulati e quelli reali, sia per stimare il numero di eventi simulati necessari e le corrispondenti risorse (cpu e storage) necessarie per la produzione finale.

Affinché la stima dell'accettanza sia attendibile è ovviamente necessario riprodurre in modo soddisfacente il funzionamento dello spettrometro. Quindi, prima di iniziare il lavoro sulle asimmetrie azimutali, è stato compiuto un confronto tra la simulazione, in cui sono stati generati 500 K urti tra muone positivo e deuterio, ed i dati relativi al sottoperiodo W35 (vedi paragrafo 2.3 del capitolo 2). In base ai tagli applicati nell'analisi degli eventi reali sono stati generati eventi con  $Q^2 > 0.8$   $(\text{GeV}/c)^2$ ,  $0.003 < x < 0.99$ ,  $0.05 < y < 0.95$  e  $W > 4 \text{ GeV}/c^2$ . Per processare i dati MC è stata usata la stessa versione di CORAL con cui sono stati processati i dati reali ed entrambi i campioni sono stati analizzati con lo stesso programma, scritto utilizzando le classi fornite da PHAST, dove sono stati implementati tutti i tagli presentati nel capitolo precedente. I dati così selezionati sono stati scritti in file formato root per rendere più rapida l'analisi interattiva e produrre gli istogrammi relativi qui presentati. Le risorse computazionali utilizzate per generare gli eventi (LEPTO+COMGEANT), processarli (CORAL) ed infine analizzarli (PHAST), appartengono alla farm di calcolo di COMPASS locale ("ACID") gestita, attualmente, dal Servizio Calcolo e Reti della sezione di Trieste dell'INFN.

### 3.4.1 Confronto con i dati

Per comparare le distribuzioni provenienti dai dati veri e da quelli simulati sono state adottate due normalizzazioni: per le quantità cinematiche e muoniche ( $Q^2$ ,  $x$ ,  $y$ ; angoli e impulsi associati al muone diffuso o del fascio) si è normalizzato allo stesso numero di eventi  $e$ , per le quantità adroniche ( $P_T$ ,  $z$ , impulsi ed angoli), allo stesso numero di adroni. La normalizzazione viene calcolata a valle dei tagli dell'analisi e viene mantenuta inalterata per ogni successiva selezione come, per esempio, nella scelta di eventi che hanno innescato uno specifico trigger. In questo caso tale

metodo ha il vantaggio di fornire una stima dell'efficacia di ciascun trigger rapportato agli altri.

Il primo confronto ha presentato una situazione non del tutto soddisfacente. In figura 3.4 viene mostrata la sovrapposizione delle distribuzioni di  $x$  e  $y$  relative ai dati (linea) ed alla simulazione (punti), mentre nella figura 3.5 sono mostrati i rispettivi rapporti.

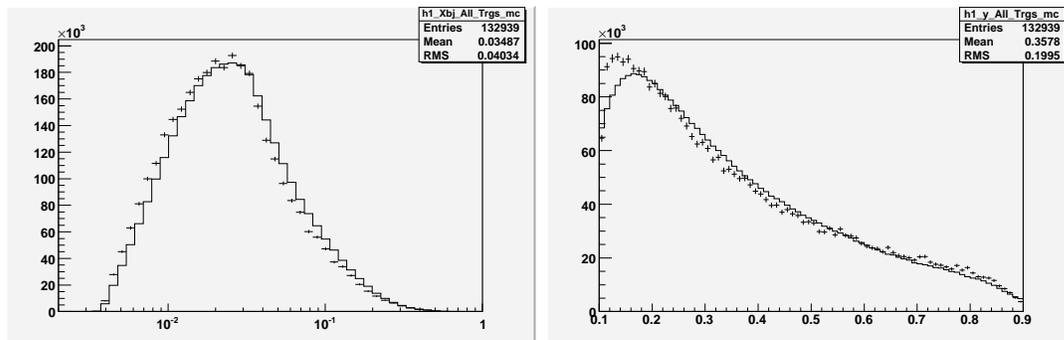


Figura 3.4: Confronto tra dati e MC per la distribuzione del numero di adroni in funzione di  $x$ , a sinistra, e di  $y$ , a destra. La linea indica la distribuzione relativa ai dati reali (2004 W35), i punti quella relativa ai dati simulati.

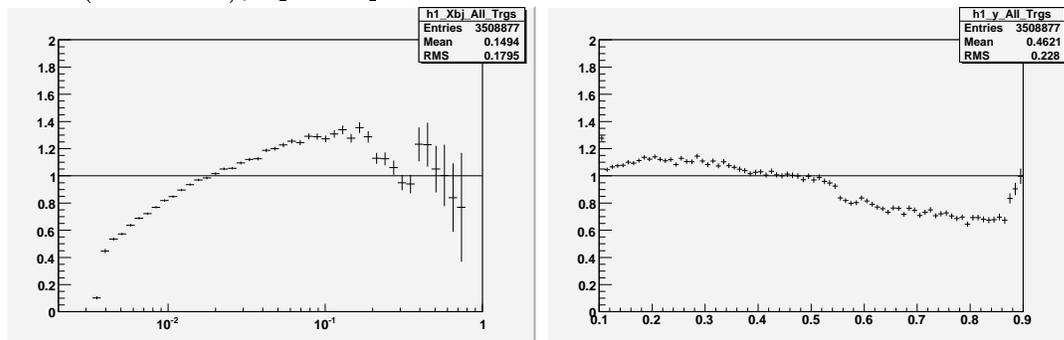


Figura 3.5: Rapporto (tra dati reali e dati simulati) della distribuzione del numero di adroni in funzione di  $x$ , a sinistra, e di  $y$ .

Ulteriori studi sono stati fatti per cercare la cause delle discrepanze. Degli eventi selezionati, relativi ai diversi trigger, sono state studiate le distribuzioni degli angoli polare ed azimutale misurati nel sistema del laboratorio di ciascuna traccia (sia muonica che adronica) e la distribuzione dell'energia degli adroni. Sono emersi due problemi, descritti nei paragrafi successivi.

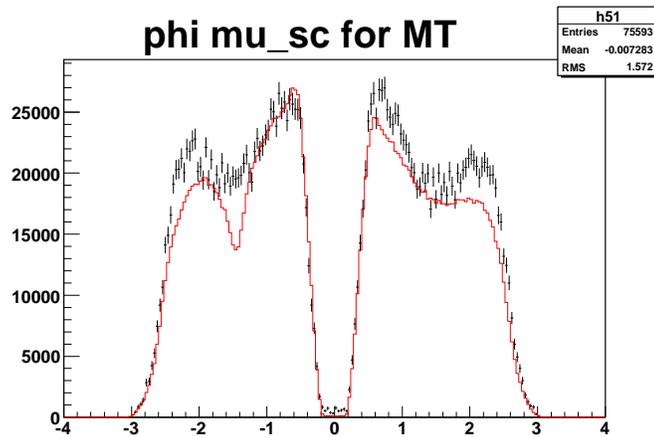


Figura 3.6: Sovrapposizione delle distribuzioni del numero di muoni diffusi in funzione dell'angolo azimutale calcolato nel laboratorio per gli eventi caratterizzati dal MT (i punti rappresentano il Monte Carlo). Si notano tre zone dove i dati sembrano contare meno eventi rispetto al MC (la zona intorno a  $-\frac{\pi}{2}$  le due zone a  $\pm\frac{3}{4}\pi$ )

### Efficienza degli odoscopi

Il primo problema riguarda il confronto per la distribuzione dell'angolo azimutale calcolato nel laboratorio, in particolare per gli eventi caratterizzati dal *Middle Trigger* (MT). In figura 3.6 viene mostrata la sovrapposizione di tale distribuzione per i dati reali (linea) e quelli simulati (punti); si notano tre zone (la zona intorno a  $-\frac{\pi}{2}$  le due zone a  $\pm\frac{3}{4}\pi$ ) dove ci sono meno eventi reali che MC. Una possibile spiegazione è che alcuni contatori a scintillazione degli odoscopi appartenenti al MT siano stati inefficienti, per verificare ciò è stata presa in considerazione la proiezione della traccia del muone diffuso sul piano dell'ultimo odoscopio e ne sono state studiate le distribuzioni in X e Y nel caso di eventi caratterizzati dal MT.

In figura 3.7 viene mostrata la sovrapposizione delle distribuzioni in X, per i due piani di odoscopi verticali (Y positivi e negativi), estrapolata in corrispondenza della coordinata Z degli odoscopi appartenenti al MT. Dal momento che la larghezza dei bin è stata presa uguale alla larghezza degli scintillatori si vedono chiaramente cinque contatori inefficienti con un'efficienza relativa che va dal 40% (per il contatore appartenente al piano inferiore e con  $X \sim 35$  cm) a circa il 70-90% per gli altri. Questa inefficienza viene simulata selezionando ulteriormente gli eventi Monte Carlo con MT all'interno del codice usato per l'analisi: per ogni traccia muonica us-

cente dal vertice primario che interseca il piano scelto nelle zone corrispondenti agli odoscopi inefficienti, viene estratto un numero casuale tra 0 e 1 (con distribuzione uniforme); se questo è maggiore della soglia di efficienza impostata, la traccia, e quindi anche l'evento, vengono rigettati. Il risultante confronto dati/Monte Carlo della distribuzione nell'angolo azimutale dell'adrone è decisamente migliorato, come si può vedere in figura 3.8.

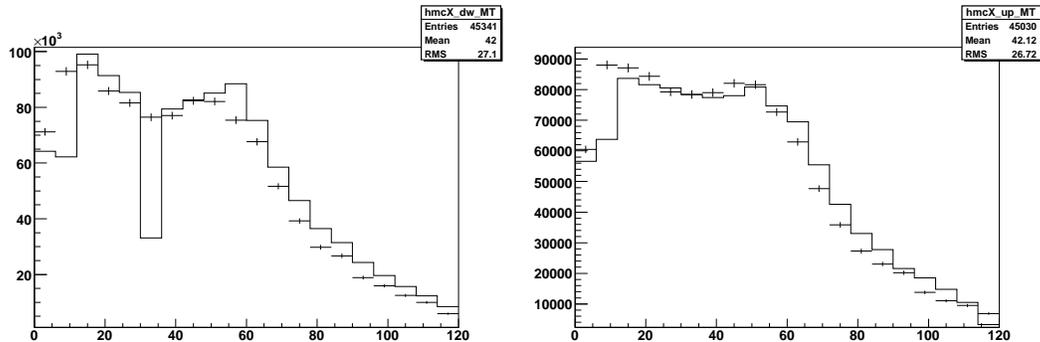


Figura 3.7: A sinistra la distribuzione del numero di muoni diffusi, in funzione dell'angolo azimutale calcolato nel laboratorio, la cui traiettoria ha intersecato il piano inferiore (Y negativi) corrispondente agli odoscopi; a destra la distribuzione analoga per il piano superiore (Y positivi). In entrambi i grafici i punti rappresentano il Monte Carlo e la larghezza del bin corrisponde alla larghezza di un contatore a scintillazione. Oltre ad evidenziare le zone di “inefficienza” di alcuni scintillatori le distribuzioni indicano un leggero spostamento della posizione (in X) delle stazioni di odoscopi (nella descrizione del MC sembrano essere spostate verso valori negativi di X); anche questo effetto è praticamente scomparso con la simulazione dell'efficienza.

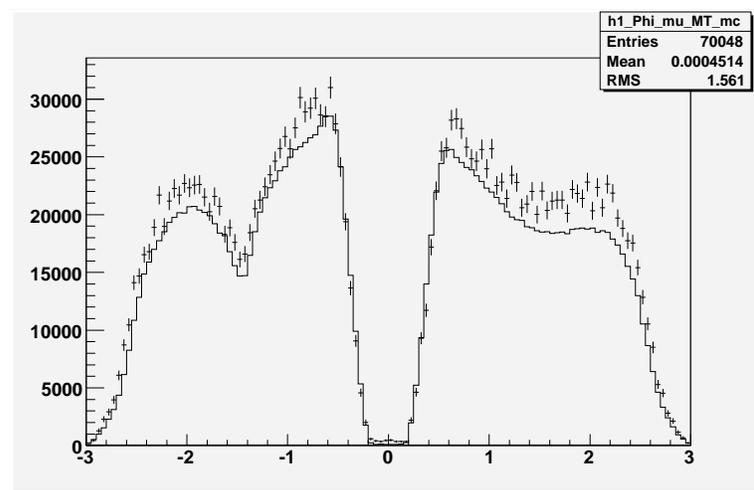


Figura 3.8: Sovrapposizione delle distribuzioni del numero di muoni diffusi in funzione dell'angolo azimutale calcolato nel laboratorio (i punti rappresentano il Monte Carlo), dopo una prima simulazione dell'efficienza nella fase di analisi.

### Soglie del calorimetro

Come si vede dal grafico a sinistra in figura 3.9 la distribuzione dell'energia dell'adrone per eventi caratterizzati dal trigger Calorimetrico non mostra un accordo soddisfacente tra dati reali e simulati, in particolare la distribuzione relativa al Monte Carlo sembra essere spostata a valori più alti rispetto a quella reale.

Tale spostamento è imputabile all'impostazione sbagliata della soglia associata al CT nella simulazione e quindi è stata rifatta una produzione Monte Carlo con tale valore posto a 9 GeV/c invece di 18 GeV/c ottenendo un notevole miglioramento come mostrato in figura 3.10.

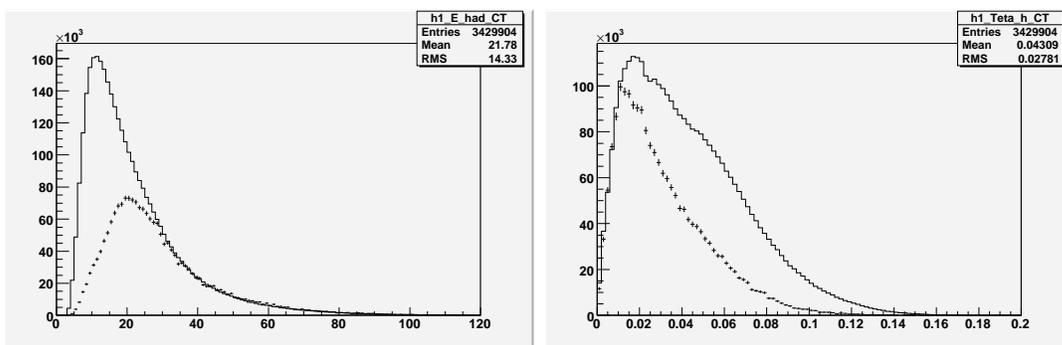


Figura 3.9: A sinistra la sovrapposizione delle distribuzioni dell'energia dell'adrone; a destra le distribuzioni in funzione dell'angolo polare nel laboratorio. In entrambi i casi i punti indicano la distribuzione per il Monte Carlo.

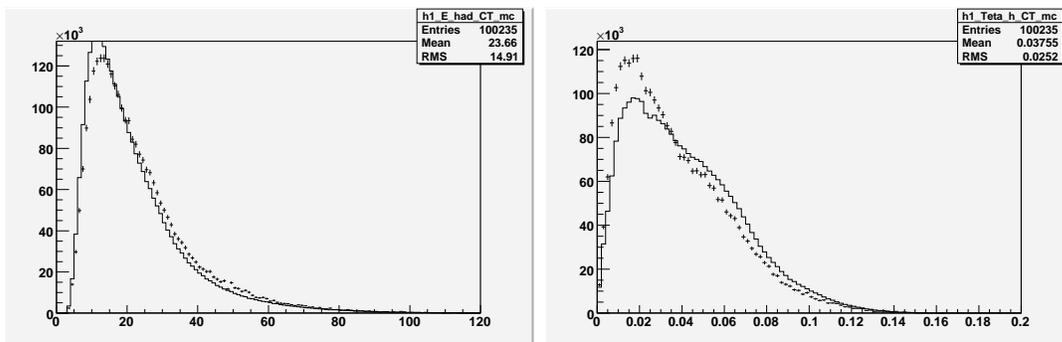


Figura 3.10: Produzione con le soglie per i segnali del calorimetro abbassate. A sinistra la sovrapposizione delle distribuzioni dell'energia dell'adrone; a destra le distribuzioni dell'angolo polare, dell'adrone, nel laboratorio. In entrambi i casi i punti indicano la distribuzione per il Monte Carlo.

Dalla figura 3.10 si vede anche che rimane una leggera discrepanza in quanto nella simulazione ci sono più adroni a ad alte energie e piccoli angoli polari. Ulteriori test sull'efficienza dei rivelatori la cui superficie attiva copre un'area piccola intorno alla linea di fascio come GEM e  $\mu$ M sono stati effettuati per migliorare l'accordo, ma senza risultati positivi. È stato quindi deciso di aggiungere un ulteriore taglio (considerare solo adroni la cui traiettoria, nel laboratorio, ha un angolo polare maggiore di 20 mrad) nella selezione di eventi applicata per ottenere la stima dell'accettanza. La distribuzione risultante dell'energia ed angolo polare dell'adrone è mostrata in figura 3.11 dove si vede un netto miglioramento nel confronto relativo alla distribuzione dell'energia.

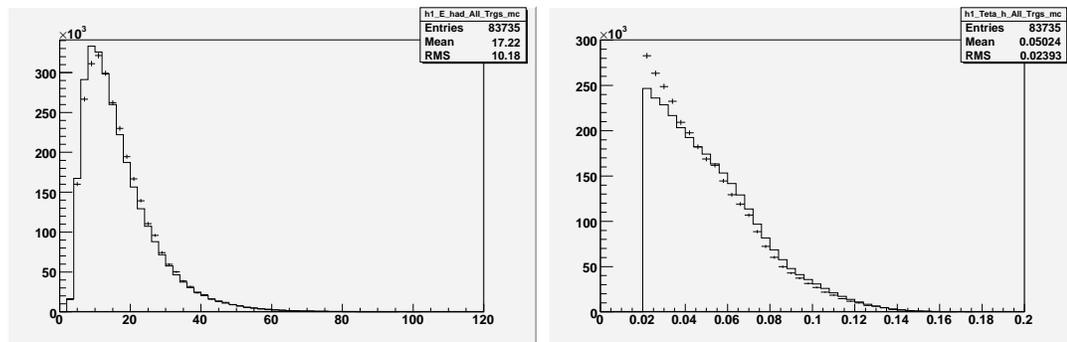


Figura 3.11: Produzione con soglie abbassate, taglio in angolo polare dell'adrone e simulazione dell'efficienza (per il MC). A sinistra la sovrapposizione delle distribuzioni dell'energia dell'adrone; a destra le distribuzioni dell'angolo polare nel laboratorio. In entrambi i casi i punti indicano la distribuzione per il Monte Carlo.

Per completare il confronto, sono state esaminate le distribuzioni di  $z$  e  $P_T$  dell'adrone i cui rapporti sono mostrati in figura 3.12. L'accordo tra dati reali e simulati non è soddisfacente per valori di  $z$  superiori a  $\sim 0.85$  e per  $P_T$  superiore a  $\sim 1.5\text{GeV}/c$ . Considerando che tali zone corrispondono ad una frazione limitata della statistica a disposizione è stato deciso di valutare l'accettazione utilizzando solo gli adroni con  $z < 0.85$  e  $P_T < 1.5\text{GeV}/c$ .

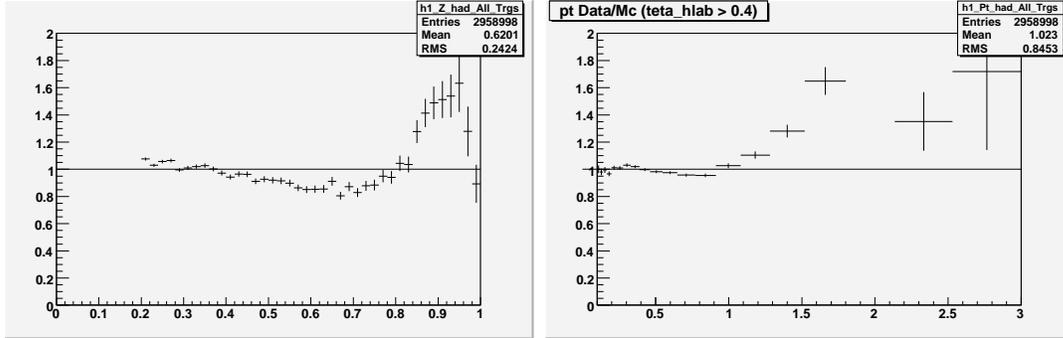


Figura 3.12: A sinistra è mostrato il rapporto tra la distribuzione di  $z$  dell'adrone ottenuta sul campione di dati reali con quella ottenuta con i dati simulati. A destra è mostrato il rapporto tra le distribuzioni di  $P_T$ .

### 3.4.2 Situazione finale

Dopo l'applicazione delle correzioni nella simulazione ed i tagli introdotti, l'accordo tra dati reali e Monte Carlo è decisamente migliore. In figura 3.13 viene mostrato il rapporto delle distribuzioni di  $x$  e  $y$ , da confrontare con i rapporti iniziali mostrati in figura 3.5; in figura 3.14 è mostrata la sovrapposizione delle distribuzioni dell'energia e dell'angolo polare dell'adrone calcolati nel laboratorio; infine in figura 3.15 sono riportate le distribuzioni dell'angolo azimutale del muone diffuso calcolato nel laboratorio e nell'angolo azimutale dell'adrone calcolato nel GNS.

Dal confronto risulta una soddisfacente descrizione dell'apparato sperimentale da parte della simulazione Monte Carlo ai fini della stima dell'accettazione azimutale  $\alpha(\phi)$ .

A questo punto si può fornire una stima dell'efficienza globale  $\varepsilon$  definita come:

$$\varepsilon = \frac{\frac{h_{dati}}{B_{dati}}}{\frac{h_{mc}}{B_{mc}}}, B = \frac{n_{ev}^{mc}}{\rho \cdot l \cdot \sigma_{DIS}} \quad (3.11)$$

$$B_{mc} = \frac{n_{ev}^{mc}}{\rho \cdot l \cdot \sigma_{DIS}}, \quad (3.12)$$

dove  $h_{mc} = 87 \cdot 10^3$  è il numero di adroni ricostruiti dal Monte Carlo a partire da  $n_{ev}^{mc} = 5 \cdot 10^5$  eventi generati, a cui corrisponde il fascio utile  $B_{mc} = 5.7 \cdot 10^{10}$  che è il numero di muoni che hanno attraversato il bersaglio ed è stato calcolato, come scritto nell'equazione 3.12, a partire dal numero di centri diffusori presenti nel bersaglio  $\rho \cdot l$  e dalla sezione d'urto per eventi DIS  $\sigma_{DIS}$  calcolata da LEPTO;  $h_{dati} = 8.8 \cdot 10^6$  è invece il numero di adroni ricostruiti dall'analisi fatta sui dati di entrambi i periodi analizzati ed a cui corrisponde il flusso  $B_{dati} = 1.36 \cdot 10^{13}$ .  $\varepsilon$  è un'indice dell'efficienza globale dell'apparato sperimentale confrontata con il modello implementato nella simulazione e quindi dovrebbe essere 1 nel caso di un apparato perfettamente simulato. Nel calcolo riportato  $\varepsilon = 0.42$ . Il motivo per cui tale valore sia così diverso da uno richiede indagini ulteriori.

Dalla stima sul numero di adroni ricostruiti dalla simulazione, espressa nell'equazione 3.10 del paragrafo 3.2 e conoscendo il rapporto  $\frac{h_{mc}}{n_{ev}^{mc}}$  si può risalire al numero di eventi che bisogna generare per il calcolo dall'accettanza  $\alpha$  perché il suo contributo all'errore della distribuzione finale sia dello stesso ordine di quello apportato dai dati:

$$N_G^{mc} = \frac{n_{ev}^{mc}}{h_{mc}} \cdot n_p \simeq 10^7, \quad (3.13)$$

dove  $n_p \simeq 2 \cdot 10^6$  è il numero di adroni ricostruiti nei dati per un dato sottoperiodo e provenienti da un urto avvenuto in una delle due celle che compongono il bersaglio. Tuttavia, per avere che l'errore portato da  $\alpha$  sia trascurabile, per esempio di un fattore 10 più piccolo di quello portato da  $n$ , il numero di adroni ed il tempo  $T$  impiegato a produrli aumenterebbero di un fattore 100. Inoltre lo studio di simmetrie azimutali può essere fatto usando i dati relativi sia alla configurazione longitudinale che trasversa, che COMPASS ha preso sia nel 2003 che nel 2002 (oltre ovviamente a quelli relativi al 2006 e 2007 non disponibili nel periodo di lavoro di tesi) e ciò comporterebbe un ulteriore incremento del numero di eventi da generare.

Nella prospettiva quindi di simulazioni Monte Carlo future il cui numero di eventi prodotto sia tale da soddisfare questi requisiti, e per portare a termine le quali il tempo impiegato diventerebbe dell'ordine degli anni se si usassero le sole risorse computazionali a disposizione della farm di Trieste, si è optato per l'implementazione di un nuovo sistema, tramite il quale effettuare la simulazione richiesta, che abbia

accesso sia alle risorse computazionali locali che a quelle, molto più numerose, fornite da Grid; tale sistema è descritto nel capitolo 4 ed è stato implementato e testato per effettuare la simulazione che ha prodotto i dati con cui è stata calcolata l'accettanza  $\alpha(\phi)$  dell'apparato.

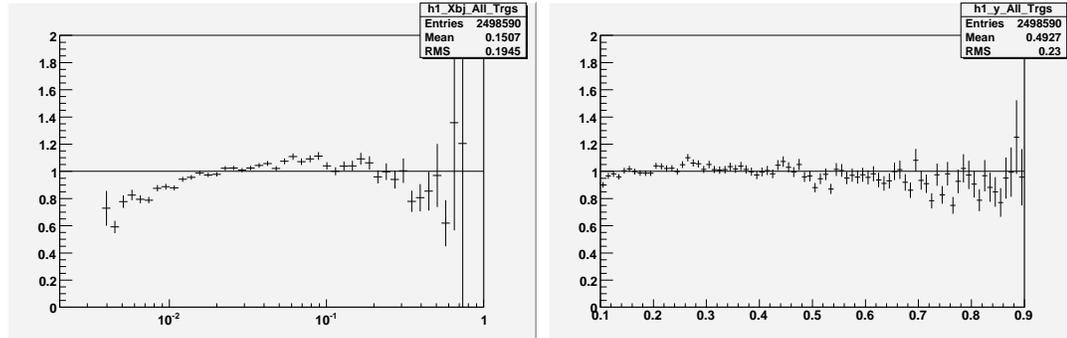


Figura 3.13: Rapporto tra dati e MC per; a sinistra per la distribuzione in  $x$ , a destra quello per  $y$ . In entrambi i casi le selezioni applicate comprendono, oltre a quelle descritte nel paragrafo 3.1 del capitolo 3, quelle descritte in questo capitolo ed i grafici vanno confrontati con i quelli mostrati in figura 3.4.

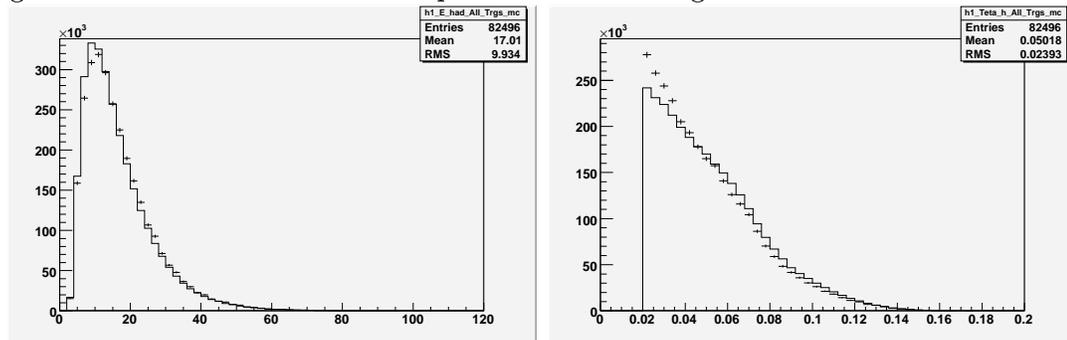


Figura 3.14: Rapporto (ottenuto in modo analogo a quello mostrato nei due grafici di figura 3.13) per l'energia (a sinistra) e per l'angolo polare dell'adrone (a destra), entrambi calcolati nel laboratorio.

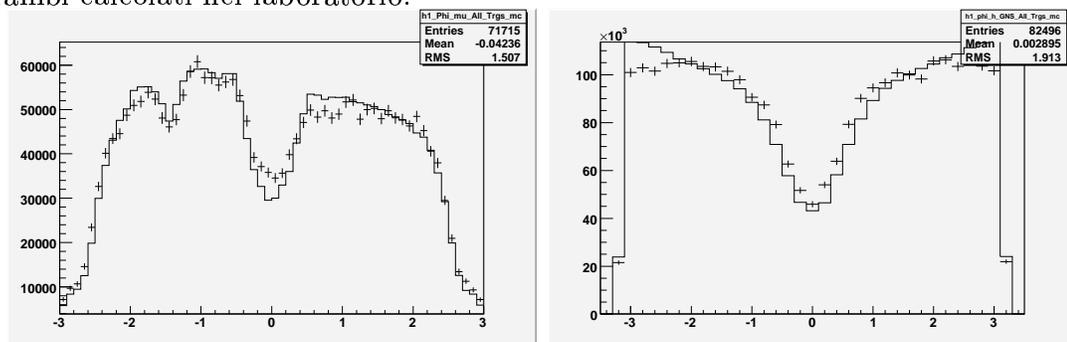


Figura 3.15: Rapporto (ottenuto in modo analogo a quello mostrato nei due grafici di figura 3.13) per l'angolo azimutale del muone diffuso calcolato nel laboratorio (a sinistra) e l'angolo azimutale dell'adrone calcolato nel GNS (a destra).



# Capitolo 4

## Sistema per la simulazione distribuita di COMPASS

### 4.1 Introduzione al sistema distribuito

Un sistema distribuito è un insieme di computer indipendenti che appaiono all'utente come un singolo sistema coerente [41]. L'interesse è naturalmente quello di presentare il sistema come coerente e coeso, nascondendo i dettagli quali diversi sistemi di autenticazione, di accesso dati e di accesso alle risorse di calcolo. Al tempo stesso, perché il sistema possa attirare l'interesse ed aggregare risorse in maniera efficace, l'integrazione deve essere possibile senza essere intrusiva. Questo significa che un centro di calcolo che serva già delle comunità di utenti tramite meccanismi convenzionali (tipicamente tramite un sistema di code *batch*), possa fornire servizi agli utenti del sistema distribuito senza modificare in maniera radicale il suo sistema di operazioni e senza danneggiare le comunità che già utilizzano il sistema.

Un sistema distribuito permette di accedere ed utilizzare risorse anche remote, che possono essere differenti computer, stampanti, pagine Web, dalla stessa macchina, indipendentemente dal fatto che altri utenti stiano utilizzando lo stesso sistema distribuito. Tale condivisione di risorse avviene in modo controllato ed uniforme. Inoltre, l'utente non deve preoccuparsi di dove venga eseguita una operazione da lui richiesta oppure se alcune componenti siano temporaneamente guaste o, ancora, se alcune di esse siano state sostituite con altre. In altre parole la proprietà principale di un sistema distribuito è quella di essere *trasparente*.

Per rendere possibile la comunicazione tra il sistema distribuito e altri servizi es-

terni ad esso si ricorre all'implementazione di interfacce che rispondono ai requisiti di un unico insieme di regole che descrivono la sintassi e la semantica di tali servizi. Per esempio, in una rete di computer, un unico protocollo formalizza le regole che definiscono il formato, il contenuto ed il significato dei messaggi scambiati tra macchine o sistemi diversi. Tali interfacce permettono il dialogo tra diverse applicazioni indipendentemente dalla loro implementazione purché siano definiti in modo univoco la sintassi ed i parametri che si scambiano. È importante che un sistema distribuito sia flessibile, ovvero che sia organizzato in un insieme di parti relativamente piccole e facilmente rimpiazzabili o adattabili ad eventuali nuove esigenze. In questo modo è possibile sviluppare o sostituire un singolo componente, quando necessario, senza alterare le parti restanti. Un sistema distribuito che soddisfa queste caratteristiche viene definito sistema aperto.

Un sistema distribuito si dice scalabile se soddisfa un certo numero di condizioni.

- Ad un aumento di risorse deve essere possibile far seguire un aumento corrispondente in qualità del servizio. Per esempio, raddoppiando i nodi computazionali, il numero di *job* processati in un giorno deve pure raddoppiare.
- In generale la scalabilità ha intervalli di validità al di fuori dei quali una o più componenti iniziano a rallentare il sistema: in un sistema scalabile, questi problemi devono essere risolvibili aggiungendo risorse senza invalidare il punto precedente in maniera sostanziale.
- Le prestazioni del sistema (a parità di risorse) non devono dipendere dal numero di utenti e di richieste. Per esempio, il numero di *job* processati in un giorno non diminuisce all'aumentare di utenti distinti o di richieste (in generale, raggiunto un valore massimo, le prestazioni integrate restano costanti).

## 4.2 Grid

Il termine Grid (griglia computazionale o semplicemente “griglia”) fu coniato da Ian Foster e Carl Kesselman nel 1998 [42] in analogia con la rete di energia elettrica, in inglese “Grid”, che consente ai suoi utenti di connettere uno strumento elettrico ad una presa, senza che si debbano preoccupare di come questa venga prodotta o

di chi la fornisca. Nel caso ideale, un utente potrebbe ottenere risorse di calcolo, di accesso ai dati ed altro, in modo analogo, dalla griglia.

La condivisione delle risorse su larga scala, non soggette a controllo centralizzato, è la base del progetto sviluppato da Grid. Dalla realizzazione pratica di questa idea discendono gli altri fondamenti necessari per lo sviluppo dell'intero progetto: la sicurezza, la gestione delle risorse computazionali, l'annullamento delle distanze e l'uso di uno standard comune. L'obiettivo finale è di fornire: all'utente, i vantaggi di avere accesso alle risorse di un sistema distribuito su scala mondiale, minimizzando gli svantaggi che un tale sistema potenzialmente comporta, quali la comunicazione tra computer fisicamente molto distanti ed amministrati secondo regole diverse; a chi mette a disposizione le proprie risorse, la garanzia che l'accesso sia controllato da efficaci politiche di sicurezza. La creazione di un sistema integrato ha naturalmente la potenzialità di essere più efficiente della somma delle parti: gli utenti infatti possono approfittare dell'eccesso di risorse generato dall'uso non costante delle varie comunità ed accedere ad un sistema che continua a fornire servizi anche quando una componente (un centro di calcolo) risultasse non disponibile, per esempio a causa di manutenzioni o eventi imprevisti.

Le componenti software che permettono la condivisione delle risorse sono conosciute come *middleware* e costituiscono uno *strato* intermedio tra le risorse fisiche e le applicazioni dell'utente. L'accesso alle risorse avviene tramite l'appartenenza ad una o più *Organizzazioni Virtuali*, ciascuna di esse è costituita da un insieme di individui e/o istituti che condividono l'accesso a computer, dati, ed altre risorse [43], controllato da determinate *policy* d'accesso.

### 4.2.1 L'Architettura Grid

Grid ha una architettura a strati sovrapposti, ognuno dei quali fornisce una funzionalità aggiuntive, come suggerito in figura 4.1. Nell'impossibilità pratica di approfondirne in questo lavoro la descrizione, faremo soltanto cenno alla suddivisione principale in tre strati che si può individuare in tale architettura.

Lo strato più basso comprende le risorse fisiche, quali computer, elementi per immagazzinare dati, insieme alla rete ed ai vari elementi *hardware* che assicurano la connettività tra le varie parti. Attualmente Grid usa *internet*, la stessa rete uti-

lizzata dal *World Wide Web* per la maggior parte del traffico di controllo e scambio dati.

Lo strato intermedio comprende il *software* che rende possibile l'utilizzo di Grid come un unico ambiente, a partire dalla gestione della sicurezza nell'accesso alle risorse fino alla gestione ed al monitoraggio dei processi la cui esecuzione è stata richiesta dall'utente; è quell'insieme di applicativi che sono stati indicati come *middleware*. In questo livello troviamo i servizi che generalizzano (virtualizzano) l'accesso alle risorse permettendo agli utenti di formulare le richieste in maniera generale (per esempio: accedere ad un file dati senza preoccuparsi dei protocolli in uso in un dato centro di calcolo).

Infine, lo strato superiore, è lo strato "visibile" dall'utente, quello cioè costituito dalle applicazioni che vengono eseguite nella griglia. Spesso questo è uno strato estremamente sottile, con l'utente che sottomette programmi e recupera risultati in modo sostanzialmente analogo a quello che avviene in un sistema *batch*. Altre volte, ed è il caso dei grandi esperimenti, in questo livello troviamo servizi che svolgono funzioni ausiliarie, come fornire informazioni sui dati disponibili (cataloghi di metadati).

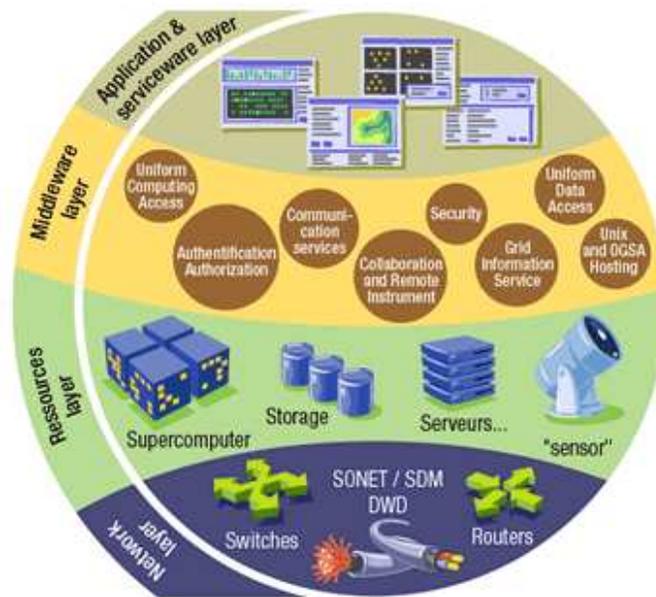


Figura 4.1: Struttura a strati dell'architettura Grid. Lo strato inferiore, "Fabric", comprende le risorse fisiche, quello superiore, "User Application", comprende le applicazioni disponibili agli utenti Grid, infine i due strati intermedi costituiscono il "middleware" con cui viene chiamato il software che rende uniforme e coerente l'utilizzo delle risorse appartenenti a Grid.

### 4.2.2 EGEE ed il middleware gLite

Il progetto EGEE (*Enabling Grids for E-science*) è nato con lo scopo di sviluppare e gestire un'infrastruttura Grid europea multidisciplinare, promuovere interazioni con le realtà extraeuropee, facilitare il trasferimento tecnologico a gruppi scientifici ed industriali attuando nuove applicazioni. Coinvolge scienziati provenienti da più di 250 istituti in 45 paesi e fornisce una infrastruttura Grid attiva 24 ore su 24 che dispone attualmente di oltre 36.000 CPU, di circa 5 PB di dischi in aggiunta allo spazio disponibile su nastro magnetico e, in media, mantiene 30.000 job<sup>1</sup> attivi concorrenti.

Il progetto EGEE si basa sul middleware basato sul lavoro del suo predecessore, il progetto European DataGrid (EDG) [45], modificato in seguito nel middleware LCG (LHC Computing Grid) [46], ed usato nella prima infrastruttura di EGEE. Questo middleware ha come fondamento il middleware mantenuto dal progetto VDT (Virtual Data Toolkit), che contiene contributi di vari progetti, come Condor [47] e Globus Toolkit [48].

Parallelamente, EGEE sviluppa un nuovo insieme di servizi, chiamato gLite. Distribuito sotto una licenza commerciale open source, gLite integra sia componenti provenienti da VDT, sia componenti sviluppati per il progetto LCG. Il risultato è un middleware che è sufficiente a costruire ed operare l'infrastruttura EGEE. L'uso di fondamenti comuni con i maggiori progetti esistenti, rende il sistema aperto a collaborazioni (ed interoperazioni con altri progetti, in particolare negli Stati Uniti). EGEE è a tutt'oggi la più grande infrastruttura grid disponibile per uso scientifico.

I servizi Grid di gLite seguono l'Architettura Orientata ai Servizi, ciò significa che diventa molto semplice collegare il software ad un altro servizio Grid, ed inoltre ciò facilita la compatibilità con gli standard Grid di nuova generazione; per esempio il Web Service Resource Framework [50] di OASIS [51] e la Open Grid Service Architecture del Global Grid Forum (recentemente rinominato Open Grid Forum) [52]. La struttura di gLite è concepita come un sistema modulare, consentendo agli utenti di sviluppare servizi differenti idonei alle loro esigenze, piuttosto che costringerli ad usare l'intero sistema. Basandosi sull'esperienza dello sviluppo del middleware EDG e LCG, gLite mira a rendere facile ed efficiente l'accesso e l'utilizzo delle risorse Grid.

---

<sup>1</sup>Nella terminologia Grid, il termine *Job* si riferisce ad un compito assegnato che richiede computazione per essere portato a termine.

I principali servizi forniti da EGEE (attraverso gLite) sono:

- Servizi informativi e di monitor (essenziali per controllare le risorse disponibili)
- Servizi dati: in particolare accesso (Storage Element), funzioni di trasferimento e funzioni di catalogazione
- Servizi computazionali: questi forniscono tra l'altro accesso ai sistemi *batch* remoti (Computing Element) e i meccanismi di selezione delle risorse (Workload Management System).

La coerenza tra questi servizi è fornita da meccanismi approvati per garantire la sicurezza (basati su standard condivisi forniti da Globus e gLite stesso) e per far funzionare il sistema in maniera controllata (per esempio, il controllo delle risorse utilizzate via il sistema di accounting).

### **I servizi di gestione dei *Job***

Il *Workload Management System* (WMS) è l'insieme dei servizi gLite responsabili della distribuzione e della gestione dei job all'interno di Grid. Esso fornisce l'efficiente sottomissione dei job compatibilmente con le informazioni sulle risorse disponibili. L'idea di base è che il WMS riesca a stimare il tempo di transito totale di un *job* al momento della sottomissione (Estimated Transit Time: ETT). L'ETT è la somma del tempo di attesa (in una coda *batch*, quindi dipende dall'utilizzazione attuale delle risorse) e dal tempo di esecuzione (quindi dal tipo di computer disponibili in una data coda).

Il *Computing Element* (CE) è il servizio che rappresenta la risorsa computazionale localizzata in un sito, tipicamente dotato di un sistema di esecuzione *batch*. Il suo principale compito è di ricevere i *job* da un utente (o da un servizio come il WMS) in modalità sicura e di farli eseguire dalla coda *batch* sottostante. Inoltre, il CE è responsabile di comunicare al sistema informativo il suo stato (numero di calcolatori totali, lunghezza delle code, organizzazioni virtuali accettate ecc...). Quindi il CE è la porta d'accesso (gateway) alle risorse di calcolo di un sito della griglia. Esso deve operare efficacemente nonostante le caratteristiche specifiche di ogni sito siano diverse (diversi sistemi di *batch* preesistenti al sito) e l'utilizzo dei computer possa essere condiviso con altri utenti in locale.

## I servizi per la gestione dei dati

Nell'architettura di EGEE i servizi per la gestione dei dati sono divisi in 3 categorie: *storage*, trasferimento e cataloghi. La granularità fondamentale con cui vengono effettuate le operazioni sui dati è in genere costituita dal file.

Il componente fondamentale dell'infrastruttura Grid è lo *Storage Element* (SE), il servizio che fornisce a Grid un'interfaccia con le differenti risorse di storage locali. Il compito dello SE è di nascondere all'utente Grid la complessità derivante sia dalle diverse tipologie di hardware, su cui sono fisicamente memorizzati i dati, sia dai diversi protocolli di accesso disponibili al sito. Per questo, il middleware EGEE richiede che ogni SE implementi una interfaccia di SRM (*Storage Resource Management* [55]), il cui scopo è mascherare le differenze dei diversi sistemi di gestione della memorizzazione dei dati virtualizzati da Grid.

Lo SE deve inoltre supportare un insieme di protocolli per il trasferimento dei file tra cui gridFTP [56], originariamente sviluppato da Globus. Per permettere un funzionamento ottimale del sistema, ci sono servizi specifici per migliorare il trasferimento dati. EGEE utilizza FTS (File Transfer Service) che controlla il traffico di richieste sulla rete in modo da ottimizzare il traffico e garantendo minimi di utilizzo per le differenti organizzazioni virtuali. Inoltre FTS è in grado di risottomettere un trasferimento dati nel caso di fallimento per tenere conto di instabilità temporanee della rete o di uno SE.

Il servizio di catalogo provvede a mantenere le informazioni sui dati immagazzinati. Esso gestisce: la collocazione ed il *namespace* dei file, la scrittura ed il recupero dei metadati e le informazioni sui permessi di accesso ad essi associati.

L'identificazione di un file registrato in Grid avviene con un *Logical File Name* (LFN), generalmente seguendo un'organizzazione della struttura dei nomi di tipo gerarchico, facilmente leggibile ed interpretabile e regolata da una logica dipendente dalla VO di appartenenza. Il LFN deve essere unico all'interno di una VO e conflitti tra differenti VO sono evitati in quanto ogni VO ha uno spazio dei nomi separato. La semantica del LFN è praticamente uguale a quella usata da un file system UNIX, dove differenti VO corrispondono a differenti directory. Tra il LFN ed i file non c'è una corrispondenza biunivoca, ad un file possono essere associati diversi LFN. Per identificare in modo univoco un file si usa il *Global Unique Identifier* (GUID), che è immutabile ed è una stringa di 36 byte generata con il meccanismo di UUID [62] che

garantisce l'unicità del nome. Il GUID viene generato automaticamente ogni volta che un file viene registrato in una risorsa Grid. Mentre ad ogni file è sempre associato un GUID diverso, non necessariamente gli viene associato anche un LFN. Infine lo *Storage URL* (SURL), detto anche *Physical File Name* (PFN), identifica sia il luogo fisico dove i dati sono scritti sia il protocollo utile per il proprio trasferimento. In gLite il catalogo contenente la mappatura tra GUID LFN e SURL dei file è implementato dal *LCG File Catalogue* (LFC).

### 4.3 Introduzione del sistema COMPASS

Il modello di simulazione adottato dal gruppo di Trieste sfrutta la farm locale (denominata ACID<sup>2</sup>) come risorsa computazionale principale per l'esecuzione e la conservazione dei dati prodotti e prevede l'uso di risorse esterne (in particolare Grid) solo in caso di grosse produzioni di dati di simulazione. Per esempio per gli studi sistematici dell'accettazione dell'apparato necessari per la valutazione di asimmetrie azimutali, le risorse di calcolo locali non sono sufficienti. È stato quindi studiato, progettato ed implementato un nuovo sistema che integra, in modo coordinato ed efficiente, tali risorse con quelle appartenenti a Grid. Un'altra idea portante è quella di potere offrire, in prospettiva, simili servizi anche ai collaboratori di COMPASS non basati a Trieste (sia dal punto di vista di accedere alla griglia per generare nuovi insiemi di dati, sia per condividere dati generati in precedenza).

Utilizziamo il termine *dataset* per identificare un insieme di dati omogeneo, in generale generato da una o più richieste di un utente (per esempio, 10.000 eventi simulati con Lepto corrispondenti ad un set di parametri di variabili e parametri per la generazione degli eventi, come il momento del muone incidente, intervalli di  $Q^2$  e  $x$  del muone diffuso, funzioni di distribuzione dei partoni ecc.). L'apparato è simulato con COMGEANT ed è anch'esso determinato da un insieme di parametri quali la versione della geometria, le soglie minime per il tracciamento delle particelle

---

<sup>2</sup>La farm di COMPASS-Trieste è caratterizzata da una potenza complessiva di circa 50 kSI2K (50.000 SpecInt 2000) fornita da 12 PC Opteron 250 e 275 bi e quadri processore e da 17 PC Pentium III. Lo spazio disco dedicato è di 16 TB. Il sistema dispone di un'unità nastro, in grado di consentire l'accesso a 10 TB di spazio on-line, che è gestito in modo trasparente tramite CASTOR (Sistema di accesso nastri sviluppato al CERN ed installato a Trieste dal 2003.)

Stato	“new” “submitted” “validated” “done” “aborted”
Numero di eventi	scelto dall’utente
richiedente	nome e indirizzo email
versione Lepto	
versione COMGEANT	
Geometria	2004,2006...
polarizzazione del bersaglio	longitudinale o trasverso

Tabella 4.1: In tabella sono riportati i più significativi parametri che caratterizzano un task di produzione presente nella base dati (eventuali modifiche sono possibili e possono essere eseguite facilmente manipolando le entrate della base dati).

secondarie nell’apparato, risoluzioni ed efficienza dei rivelatori ecc.). La lista dei principali parametri è riportata in tabella 4.1.

In tale sistema, la produzione dei dati e la loro aggregazione sono effettuati mantenendo le informazioni sul corrente stato di produzione di ciascun dataset disponibile o semplicemente richiesto. Tutte queste informazioni, che includono la collocazione dei dataset e sul loro stato di produzione sono mantenute in un’apposita base di dati che viene utilizzata dal sistema per accettare nuove richieste, aggiornare lo stato di avanzamento dei task in corso di elaborazione e reperire la collocazione dei dati e dei file di log.

Questa base di dati è il cuore del sistema, nel senso che svolge sia la funzione di archivio dei dati sia quella di controllo della processatura. L’interazione tra le richieste di nuovi dataset e l’effettiva processatura è permessa dalla base di dati. Le richieste sono semplicemente depositate dagli utenti. La processatura vera e propria viene attivata in maniera asincrona a partire dalle richieste trovate nella base di dati. Alla fine della processatura, i dati vengono pubblicati attraverso la base di dati.

In questa prima fase il sistema si appoggia ai servizi Grid di base nell’ambito del progetto EGEE (attualmente alcuni servizi sono resi disponibili dal CERN e dagli altri istituti della collaborazione) ma integra, in maniera trasparente anche le risorse locali della sezione di Trieste (farm ACID).

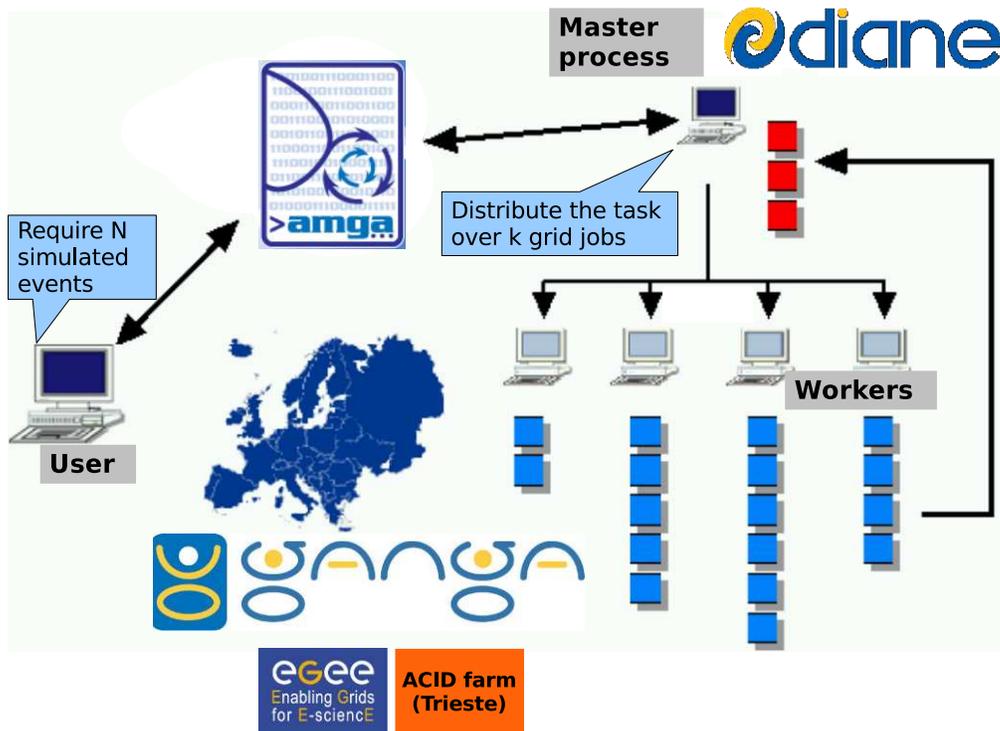


Figura 4.2: Nella figura è riportato uno schema di massima del sistema.

### 4.3.1 Architettura del sistema

Le componenti del sistema sono illustrate in figura 4.3. Ciascuno modulo è indipendente e svolge un compito ben preciso. L'interazione tra le componenti principali è asincrona, in modo da rendere più semplice e robusto il funzionamento del sistema. In questo modo è possibile alterare le diverse parti riducendo al minimo necessario il numero di cambiamenti sugli altri programmi che fanno parte dell'intera struttura (manutenzione software). Inoltre un fallimento di un particolare modulo può essere recuperato intervenendo sulla parte interessata e non pregiudica l'esecuzione dell'intero sistema che può riprendere non appena il modulo è di nuovo funzionante.

DIANE è lo scheduler adottato perché in grado di consentire la sottomissione di job di simulazione indifferentemente su risorse computazionali locali o Grid, in modo equivalente (si usa il termine trasparente in quanto le differenze dei vari sistemi utilizzati congiuntamente, per esempio la farm e Grid, vengono nascoste all'utente finale). DIANE è stato integrato in un framework di script coordinati incentrato su una base dati informativa cui il sistema si interfaccia attraverso AMGA (il catalogo per metadati sviluppato nel progetto EGEE che permette di accedere a basi di dati attraverso la griglia). L'importanza della base di dati è di permettere un semplice meccanismo di sincronizzazione e di isolare la parte utente (richieste di dati esistenti o di nuovi) dalla processatura. Inoltre AMGA ha la possibilità di avere politiche di accesso integrate con le strutture già implementate da gLite (per esempio l'autenticazione tramite certificato proxy). In ultima analisi, AMGA armonizza l'accesso alle basi di dati con i protocolli Grid, fornendo dei client leggeri ed efficienti.

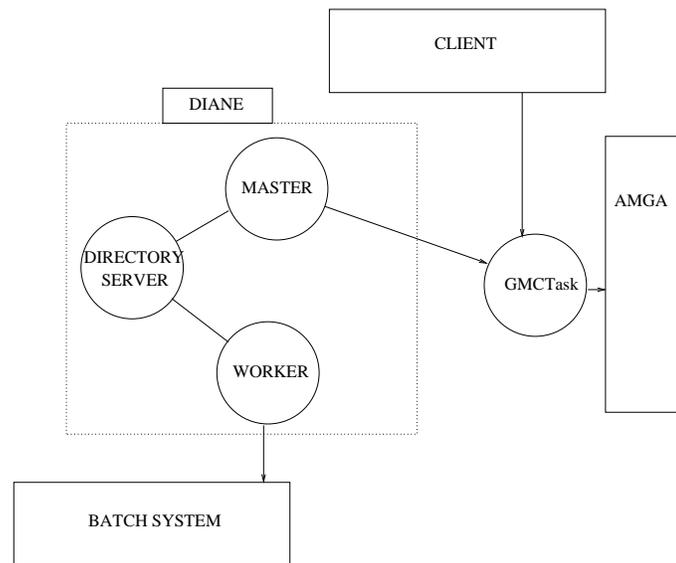


Figura 4.3: Funzionamento del sistema. GMCTask è un modulo (python) che implementa le classi tramite le quali il resto del sistema dialoga con AMGA. La parte client fa delle richieste che vengono registrate nella base dati mentre DIANE interroga la base dati e si occupa della sottomissione delle simulazioni richieste.

## 4.4 Valutazione tecnologie Grid

### 4.4.1 AMGA

In AMGA (ARDA Metadata Grid Application) i metadati sono definiti in termini di *schema*, *entry* e *attribute*. Uno *schema* è l'insieme di *entry* gerarchicamente organizzate, analogo alla struttura a directory di un filesystem Linux. Le *entry* secondo questa analogia corrispondono ai file, e vengono identificate dalla loro posizione nella gerarchia. Allo *schema* viene assegnata una lista di *attribute* che contiene il nome, unico per ogni *schema*, ed il tipo di variabile. Ogni *entry* assegna un valore a ciascun attributo dello schema a cui appartiene. Nella base dati usata come back-end, gli *schema* corrispondono a tabelle in cui le colonne sono gli *attribute* e le righe sono le *entry*.

AMGA è un servizio che fornisce una interfaccia ad un sottostante sistema di base di dati relazionale dove vengono registrati i metadati. Attualmente sono supportati Oracle, PostgreSQL, MySQL, e SQLite. Le richieste ricevute dalle applicazioni client, vengono ricevute, accettate, convalidate ed infine redirette alla base di dati che contiene i dati veri e propri. AMGA si incarica anche di raccogliere i risultati delle richieste degli utenti e di inviarle alle applicazioni client, utilizzando tutta una serie di accorgimenti per massimizzare le prestazioni (in termini di flusso di dati). Inoltre AMGA implementa dei meccanismi per ottimizzare il numero di interazioni client-server per ridurre i fenomeni di latenza .

Un catalogo di metadati deve essere in grado di servire migliaia di client che accedono alla base dati contemporaneamente e che possono innescare risposte molto impegnative in termini di quantità di dati. Le prestazioni [58] di AMGA sono comparabili ad altri cataloghi adottati da Grid, ma migliori, di almeno un ordine di grandezza, quando AMGA effettua *streamed bulk operations*, in questo caso viene quasi raggiunto il limite della base dati.

L'autenticazione può essere regolata sia tramite l'impostazione delle Access Control List (ACL) associate ad ogni schema, sia con la possibilità di introdurre dei permessi per ogni entry, alla maniera del file system Linux, quest'ultima va, però, a scapito delle prestazioni. Inoltre sono supportati anche meccanismi per l'autenticazione, basati su password o certificati (anche grid-proxy e con estensione VOMS) e connessioni sicure basate su SSL. In generale AMGA è più veloce degli altri sistemi

di accesso alle basi di dati presenti in Grid anche utilizzando accessi sicuri, grazie alla capacità di gestire le connessioni che, aperte in modalità sicura, vengono utilizzate efficientemente dal server (sostanzialmente evitando che interazioni successive tra client e server inneschino ad ogni interazione una nuova procedura completa di autenticazione: queste operazioni hanno una durata incompressibile che dipende dal numero di interazioni client-server e dalla latenza determinata dalla distanza client-server - per le connessioni più lontane è dell'ordine di 100-200 ms).

#### 4.4.2 DIANE

DIANE [59] (DIstributed ANalysis Environment) è un sistema per effettuare calcolo distribuito in corso di sviluppo al CERN. Allo stato attuale ha già dimostrato di essere particolarmente duttile adattandosi ad una serie di attività diverse, tra le quali, le principali sono <sup>3</sup>:

- controllo dello sviluppo del pacchetto Geant4 attraverso il confronto di numerose configurazioni (rapida esecuzione di un gran numero di test utilizzati per validare il software prima del rilascio di una nuova versione),
- applicazioni biomediche come la ricerca di principi attivi contro la febbre aviaria (calcoli di fisica molecolare dell'interazione di elementi potenzialmente attivati su di un insieme di proteine del virus. Le coppie proteina-principio attivo che mostrino una particolare affinità sono selezionate per test di laboratorio classici)
- processatura su larga scala di programmi per stabilire la divisione di frequenze per le trasmissioni digitali in Europa, Africa, Medio Oriente e territori dell'ex Unione Sovietica [ITU radiocommunication conference, Maggio-Giugno 2006]

DIANE è una infrastruttura (implementata in Python) che permette un uso efficiente di grandi sistemi distribuiti quando il carico di lavoro possa essere diviso in un gran numero di task indipendenti. Nel caso della generazione di eventi per COMPASS è chiaro che, controllando opportunamente i generatori di numeri casuali che determinano la generazione, un job di N eventi possa essere suddiviso in n task indipendenti. Le caratteristiche principali sono:

---

<sup>3</sup><https://twiki.cern.ch/twiki/bin/view/ArdaGrid/ITUConferenceIndex>

- Uso uniforme di risorse provenienti da sistemi diversi per uno stesso insieme di task (per esempio, integra risorse della farm ACID con job eseguiti sulla griglia).
- Ogni task è distribuito a un nodo di esecuzione solo quando questo è pronto a riceverlo. Questa tecnica si chiama “late binding”. Un sistema batch tradizionale lavora in modalità “early binding” nel senso ogni job riceve uno (o più) task quando viene sottomesso.
- Il sistema è in grado di prendere semplici ma efficaci misure per aumentare la probabilità che tutti task siano eseguiti con successo.

DIANE implementa un sistema, di tipo client-server, “Master-Worker” in cui il master gestisce il flusso di informazioni e di dati con i vari worker che sono, essenzialmente, responsabili dell’esecuzione dell’applicazione richiesta dall’utente. In base ai parametri che riceve dal cliente, il master suddivide il job assegnatogli in task, che rappresentano l’unità di lavoro fondamentale e che vengono assegnati ai worker. Il numero di task assegnato a ciascun worker risulta determinato dalla velocità con cui ogni worker riesce portare a termine i task assegnatigli. Per fissare le idee, chiamiamo  $n$  il numero di task e  $k(t)$  il numero di worker attivi all’istante  $t$ . Di fatto  $k(t)$  è funzione non banale del tempo perché a seconda dello stato delle risorse il numero di worker può cambiare; con  $k$  indichiamo il numero tipico di worker attivi durante l’esecuzione.

Il primo risultato interessante è che la durata dell’intero job non è dominata dal tempo di esecuzione della macchina più lenta (sempre nell’ipotesi  $n \gg k$ ). Il meccanismo di late-binding fa sì che una macchina che processa a metà velocità delle altre riceverà circa metà task e non sarà necessariamente l’ultima a terminare l’ultimo task.

Tipicamente un utente ha accesso a  $10^2$  worker (un numero di worker accessibile ad un utente medio della griglia), quindi il sistema diventa molto competitivo rispetto alla sottomissione di un job tradizionale (batch) se il job può essere suddiviso in un numero di task dell’ordine di  $10^3$  (o più). Il parallelismo ed il meccanismo di late-binding permettono che l’accelerazione rispetto alla processatura su di un singolo nodo sia in questo caso effettivamente dell’ordine di  $10^2$ .

L'altro motivo di interesse è la possibilità già menzionata di regire automaticamente al fallimento di un task (per esempio con la sua riesecuzione). In un sistema batch, per motivi di efficienza, ogni job esegue molti task, per esempio un gran numero di eventi di simulazione, in una singola unità di esecuzione e il fallimento di uno di questi task invalida l'intero job o richiede delle procedure sofisticate per isolare l'errore. Grazie all'adozione di DIANE, se un task fallisce per cause "transienti" (per esempio dovute alla memoria non sufficiente di un certo PC, mancanza di software, problemi di accesso ai dati sul nodo di esecuzione, hardware difettoso etc...) viene rieseguito su un'altro worker. Errori persistenti (task falliti più volte, tipicamente 3 volte) non vengono recuperati, ma studiati da esperti dell'applicazione in un secondo momento: rappresentano efficienti test case per individuare errori nel codice, nelle librerie o nelle strutture dati come errori di (de)codifica, "memory leak" o altro.

L'interazione tra master e worker è opzionalmente mediata da un servizio, che prende il nome di Directory Service (DS), che è in grado di ottimizzare il riutilizzo dei worker che abbiano eseguito tutti i task assegnati da un master reindirizzandoli ad un altro master. Come illustrato in figura 4.4, il DS controlla se ci sono master attivi che richiedono l'esecuzione di task e contatta i worker inattivi assegnandoli agli eventuali master che ne hanno bisogno.

DIANE è in grado di adattarsi al tipo di lavoro richiesto attraverso l'*Application Adapter*. Esso consiste in un modulo (Python) in cui sono definiti alcune classi e metodi che vengono opportunamente gestiti dal resto del software e attraverso i quali viene regolato il flusso di messaggi e di dati che avviene tra master e worker. Modificando opportunamente tale modulo, l'utente è in grado di pianificare la suddivisione del job in task. Va evidenziato che questi vengono assegnati ai worker in base al loro arrivo (al primo worker che lo richiede viene assegnato il primo task) senza sapere a priori a quale worker sarà assegnato uno specifico task (all'interno di questo meccanismo non è banale assegnare determinati task a specifici worker, ma è anche inessenziale per gli scopi che ci si prefigge).

### 4.4.3 GANGA

Per nascondere le differenze dei sistemi batch (in questo caso LSF e Grid) a cui DIANE può appoggiarsi viene utilizzato GANGA [61], il prodotto di un progetto congiunto ATLAS e LHCb. Esso è un software in grado di gestire in modo

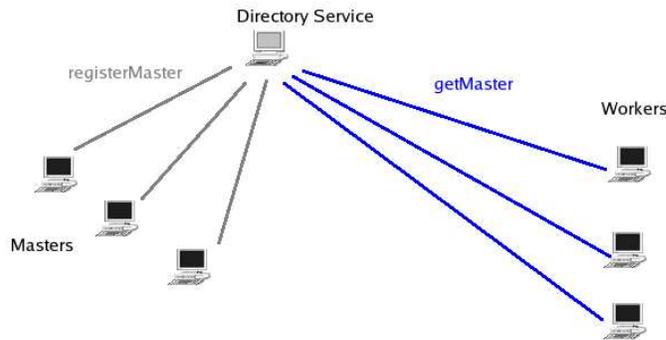


Figura 4.4: Il funzionamento del Directory Service. Esso consiste in una lista dei master: non appena un master viene attivato si registra in tale lista. In modo asincrono i worker periodicamente richiedono al DS eventuali master attivi. Il DS assegna i worker che lo contattano ai master presenti nella lista.

trasparente i diversi meccanismi di sottomissione dei job relativi ai diversi sistemi batch tradizionali o di Grid ed è supportato da DIANE, che lo utilizza nella fase di sottomissione dei worker.

Ganga è in realtà un sistema complesso che permette, per esempio, di assistere l'utente ATLAS (per esempio) a costruire job che utilizzano il software di ricostruzione dell'esperimento (chiamato ATHENA), assistendolo nella scelta delle opzioni, dei dati ecc. Nel nostro caso, la funzionalità di Ganga utilizzata è appunto quella di isolare dai dettagli del sistema di esecuzione dei job. In pratica tutti i job che veicoleranno i worker del sistema DIANE sono sottomessi tramite Ganga. Alcuni sono diretti su ACID, altri sulla griglia, ma tutti, nel momento in cui entrano in esecuzione contattano il DS e, se ci sono master attivati, li contattano e iniziano a processare tasks.

## 4.5 Implementazione

Come già accennato, il sistema è pensato come un insieme di programmi indipendenti, asincroni, che gravitano attorno al database (DB) il cui compito è mantenere le informazioni necessarie affinché l'attività del sistema si svolga in modo coerente. Ogni singola produzione è innescata da un processo che rileva (in modo asincrono, attraverso una query al DB) la richiesta di un utente al sistema e attiva un mas-

ter DIANE a cui è affidato lo scheduling del job corrispondente. Parallelamente il carico di lavoro del sistema batch di riferimento viene monitorato e, in base alle informazioni raccolte, vengono sottomessi alcuni worker pronti ad incaricarsi dell'effettiva esecuzione dei singoli task che il master assegna loro. Anche questo processo è asincrono, il numero dei worker ed il back-end (LSF o Grid) vengono scelti in modo da minimizzare il tempo di esecuzione del job richiesto (efficienza del sistema), in questo caso la simulazione, e l'utente è liberato dalla necessità di occuparsi di dettagli che non siano la scelta dei dati da generare e dalla loro utilizzazione finale.

La tattica per la scelta delle risorse di esecuzione prevede di sfruttare le risorse di calcolo locali e di ricorrere, in modo automatico, a Grid solo in caso in cui esse non siano sufficienti.

Vi sono due "scenari" di base. Il primo corrisponde al caso in cui il sistema non è utilizzato ed una richiesta viene inserita. Nel secondo, invece, un gran numero di worker è già presente ed è utilizzato da alcuni master per svolgere altri task.

**Scenario 1:** il fattore limitante è dato dal tempo per istanziare i worker. In particolare in Grid la concorrenza di altre attività può essere tale che i worker sottomessi per eseguire il task dell'utente rimangano in coda per tempi lunghi, tipicamente dell'ordine di qualche ora. Per garantire che la processatura inizi il più presto possibile (in modo da fornire primi risultati parziali con minima latenza), le risorse locali sono quindi essenziali e di fatto è possibile avere un minimo di worker latenti, che non consumano CPU, in attesa in ogni momento senza disturbare il normale uso della farm<sup>4</sup>. La velocità di esecuzione globale crescerà col tempo man mano che più worker entrano in esecuzione.

**Scenario 2:** in questo caso (ricordiamo che il sistema è saturo) la latenza iniziale può essere importante perché il sistema sta già attingendo a "tutte" le risorse disponibili. D'altra parte il numero di worker che terminano l'esecuzione di task (per altri job) nell'unità di tempo è relativamente garande e tramite il directory service questi migrano progressivamente sull'ultimo master. Nella implementazione attuale di DIANE, la migrazione dei worker tra un master

---

<sup>4</sup>Una volta che un worker viene sottomesso occupa una slot, che quindi viene preclusa ad altri job. L'uso di cpu, quando non esegue il task assegnatogli dal master, è tuttavia minimo e quindi non impegna la macchina e la cpu può essere utilizzata, praticamente per intero, da eventuali job sottomessi dagli altri utenti.

e l'altro avviene di fatto solo quando un job termina e rilascia tutti i worker, ma a breve è previsto il rilascio di un prototipo che permette migrazioni in maniera dinamica.

Per essere maggiormente versatile, il sistema trae le proprie configurazioni da un'unica classe dove sono definiti i parametri che ne caratterizzano l'ambiente.

Per motivi tecnici il prototipo con cui si è iniziato è quello in cui tutti gli script Python vengono eseguiti sulla stessa macchina, da cui si ha accesso sia a Grid che al DB, e si possono dividere in tre gruppi:

- script che vengono eseguiti periodicamente dal sistema e ne gestiscono le funzionalità;
- script che contengono le definizioni delle classi o dei metodi utilizzati per dialogare con il DB;
- script che permettono all'utente di interagire con il sistema (inserire nuovi task, ottenere informazioni sui task già presenti, cancellare task propri).

Appartenenti al primo gruppo figurano i client principali: uno si occupa della sottomissione del job (GetTask) ed uno si occupa di controllare che gli eventuali dati prodotti siano validi (ValidateTask).

Il compito di GetTask è cercare dei task appartenenti alla base dati etichettate come nuove, rilevarne la più vecchia e, dopo averlo posto in stato di lock, invocare gli opportuni metodi che preparano i file necessari per sottomettere la simulazione corrispondente a quel task ed, eventualmente, attivare un master (in locale); il suo funzionamento è illustrato in figura 4.5.

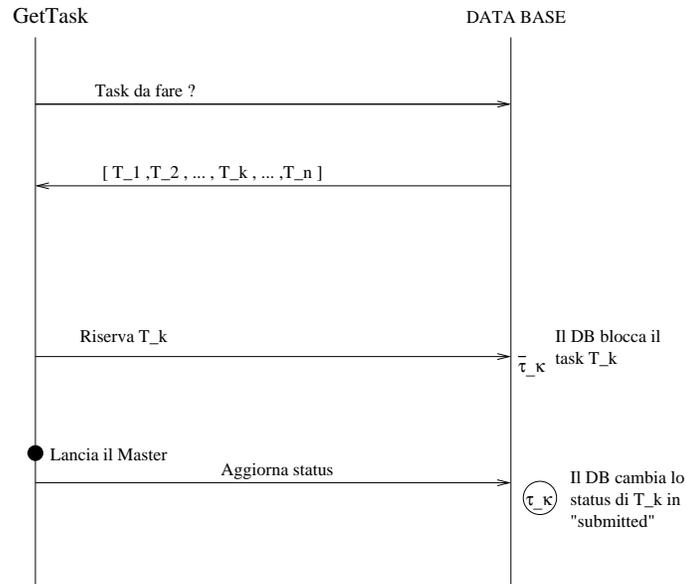


Figura 4.5: Interazione tra GetTask ed il sistema informativo. In risposta alla richiesta di nuovi task, il client riceve una lista di task (“new”) tra i quali può sceglierne uno, (attualmente in base alla data della richiesta, ma possono essere facilmente implementati altri criteri come priorità “user dependent”) e chiede al DB di “bloccarlo”: a questo punto, a meno che lo stesso task non sia stato precedentemente bloccato su richiesta di un altro client, viene attivato un master incaricato di sottomettere il job associato a tale task. Quando sia la preparazione che la sottomissione del master sono andate a buon fine, il client comunica i nuovi dati con i quali il DB aggiorna la entry corrispondente a quel task “sbloccandolo”.

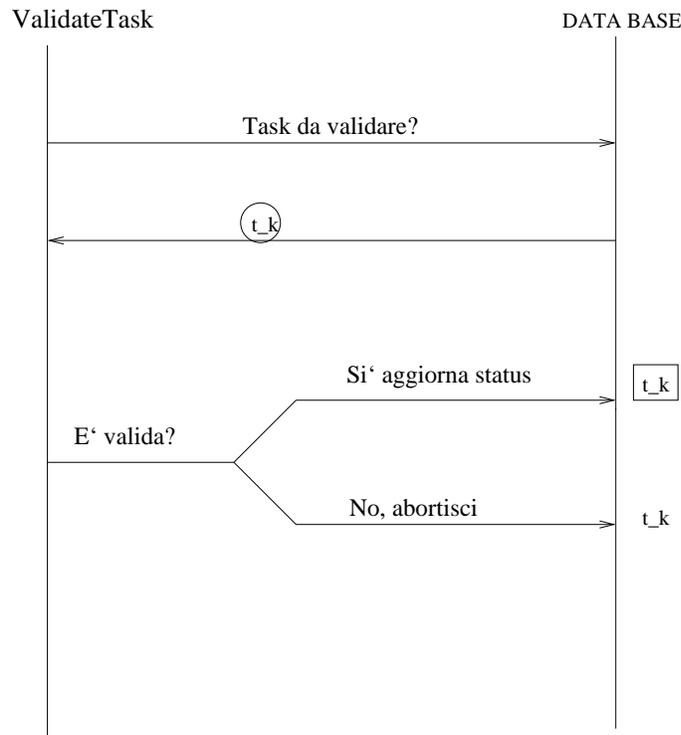


Figura 4.6: Una volta marcati come sottomessi da GetTask i task possono essere rilevati e “bloccati” da ValidateTask che ne monitora l’andamento analizzando l’output proveniente dal master. Una volta rilevata la terminazione del master, ValidateTask inizia il processo di validazione dei dati prodotti in seno allo specifico task di cui il master era responsabile. Se tale processo non ha successo il task viene marcato come “aborted”.

In modo asincrono ValidateTask interroga il DB, determina quali task risultino sottomessi ed implementa il processo di validazione del task sottomesso. Esso consiste principalmente nel rintracciare i file prodotti in seno ad un dato task e verificare che ricostruiscano integralmente il dataset del task relativo ad essi. Il suo funzionamento è illustrato in figura 4.6.

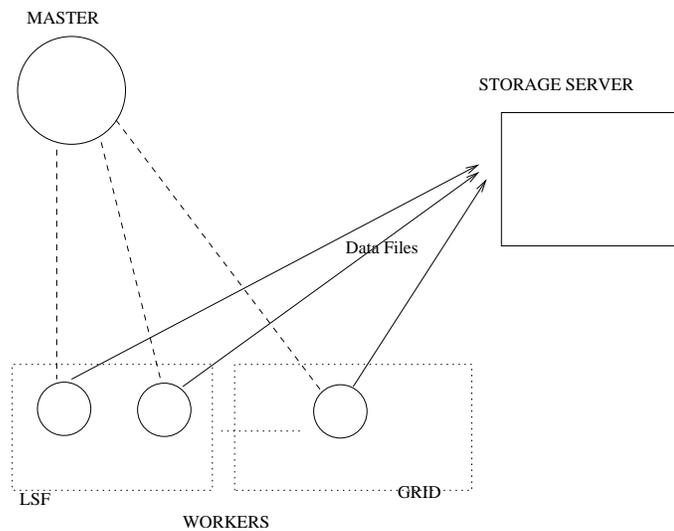


Figura 4.7: Ogni unità di lavoro assegnata al worker produce diversi file contenenti i dati ed i “log” relativi ai programmi eseguiti. Essi vengono trasferiti su uno spazio di memoria nella farm ad opera degli stessi worker, sia che questi siano stati sottomessi in locale sia in Grid, in modo da garantire un accesso uniforme ai file stessi.

Al momento attuale i singoli file contenenti i dati prodotti dalla simulazione vengono copiati, dagli stessi worker, in uno spazio di memoria predefinito (vedi figura 4.7), sia che siano stati sottomessi in Grid sia nel sistema batch locale. Tipicamente i file hanno una grandezza che si aggira intorno al GB (per circa 10K eventi processati). Successivamente essi vengono registrati nel catalogo dei file (per consentirne il reperimento trasparente anche in ambiente Grid).

Il programma responsabile della sottomissione dei worker monitora periodicamente l'attività sulla base delle informazioni fornite dal DS (sempre attivo nel sistema). Il suo compito è di mantenere, in locale, un numero minimo di worker liberi sempre attivi, pronti per essere utilizzati non appena un nuovo master venga creato, in modo che il job richiesto venga subito sottomesso. Quando il carico di lavoro diventa troppo oneroso per il sistema batch locale, cioè il numero di worker supera una certa soglia, inizia la sottomissione in Grid. Al momento la scelta sul numero di worker, da sottomettere inizialmente, non prevede di sfruttare l'informazione relativa al numero totale di eventi da produrre ed, ad ogni master, ne viene assegnato un numero prefissato.

Una volta che il DS ha messo in contatto ciascun master con i propri worker l'interazione tra loro viene regolata interamente da DIANE, come illustrato in figura 4.8.

L'unità di lavoro che viene eseguita (che consiste nella simulazione MC di eventi di DIS) consiste nella catena Monte Carlo LEPTO+COMGEANT. Per il momento le parti riguardanti la ricostruzione e l'analisi effettuate da CORAL e PHAST rispettivamente vengono processate in locale, al di fuori di questo sistema, in quanto questi programmi, scritti in C++, necessitano di numerose librerie dinamiche la cui implementazione dipende dalla versione del programma utilizzata, in continuo cambiamento (grazie anche alla relativa facilità di aggiornare parti di codice caratteristica dei programmi "modulari") per inglobare nuove informazioni ottenute riguardanti il funzionamento dello spettrometro in dato periodo di presa dati (per esempio l'efficienza dei rivelatori). Diventa molto più complicato quindi "esportare" l'intero framework necessario per la loro corretta esecuzione rispetto, per esempio, a COMGEANT il cui eseguibile comprende quasi esclusivamente librerie statiche e la descrizione dell'apparato dipende essenzialmente dall'anno di presa dati e dalla configurazione dello spin nel bersaglio (longitudinale o trasversa).

Come detto nel capitolo 3, la catena LEPTO+COMGEANT richiede la presenza di alcuni file, ognuno dei quali contiene la descrizione della geometria, della mappatura dei materiali e dei campi magnetici (in tutto sono circa una decina) e che vengono letti durante l'esecuzione. La strategia adottata prevede che gli eseguibili e tutti i file necessari vengano raggruppati in modo gerarchico in un unico file (che ha le dimensioni di quasi 20MB) che comprende anche uno script la cui esecuzione inizial-

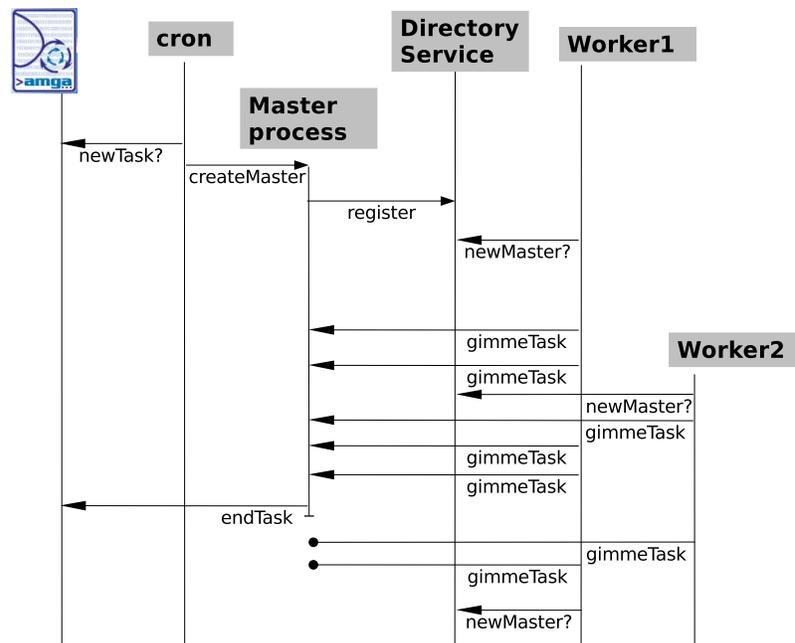


Figura 4.8: Diagramma temporale dell'esecuzione di un task da parte del sistema. Un cron script controlla periodicamente se è presente un nuovo task nella base dati, in caso affermativo attiva un master. Quando un worker viene attivato chiede al DS se c'è un master; il DS gli restituisce il riferimento (Interoperable Object Reference - IOR) al master disponibile (attualmente vengono associati solo M-W che condividono la stessa applicazione) con il quale il worker è in grado di contattare il master. Nel caso che nessun master sia attivo, il sistema decide se terminare o tenere in vita il worker in base al numero totale di worker attivi.

izza e fa partire la catena di simulazione. Questo file viene creato, insieme al master, in modo automatico, dal sistema implementato per la tesi, in seguito ad una richiesta di task presente nella base dati e viene consegnato dal master ai worker attivi che, seguendo i comandi impostati nell'application adapter, si incaricano di "spacchettarlo" e di effettuare ciascuna simulazione sulle macchine (remote) d'esecuzione restituendone infine i file di output.

In fase di test sono state provate due differenti strategie per l'implementazione della catena MC: la prima consiste nel generare con LEPTO l'intero ammontare di eventi, richiesto dal task, in locale affidando al master la suddivisione e la distribuzione (ai worker) del file così prodotto; la seconda invece prevede che l'intera catena sia eseguita dai worker ed il master si occupa di consegnare i parametri necessari (per esempio il numero di eventi per ciascuna unità di produzione, il seme per la generazione dei numeri casuali). Per la produzione finale è stata scelta la seconda strategia in quanto è stato verificato che l'overhead di comunicazione del master causato dalla suddivisione e dalla distribuzione del file LEPTO comportava una sostanziale diminuzione delle performance globali.

### 4.5.1 Test di performance

#### **Valutazione del numero di eventi da simulare in un singolo task**

Per verificare le caratteristiche del sistema sono state eseguite alcune misure di prestazioni di calcolo in parallelo. Queste hanno lo scopo di determinare il numero di eventi da simulare che caratterizza ciascuno dei task in cui il job viene suddiviso e che viene assegnato dal master a ciascun worker, garantendo l'efficienza del sistema. Si tratta in generale di una situazione di compromesso tra task grandi e piccoli. Grandi task hanno il vantaggio che l'effetto di overhead della comunicazione tra master e worker e altre inizializzazioni effettuate dal worker hanno un effetto minimo. Naturalmente task molto grandi possono limitare il livello di parallelismo e rendere poco efficiente il meccanismo di correzioni degli errori (un task con un errore viene interamente rieseguito). In aggiunta, un unico task che impieghi molto più tempo degli altri (per esempio, perché eseguito su una CPU più lenta) può sostanzialmente diminuire le prestazioni globali del sistema.

Il sistema è stato valutato nella produzione di eventi montecarlo attraverso la

eventi	T(s)	$T_i$ (s)	t/evento(s)
894	507	10	0.59
8977	4536	10	0.58

Tabella 4.2: Misure di performance sulle macchine con processore “Opteron 250” (SI2K = 1037). La quantità secondi/evento·SI2K = 601 è una costante (per i diversi tipi di PC) ed è relativa al numero di operazioni necessarie per processare un evento. T è il tempo totale di simulazione (compreso di inizializzazione e simulazione LEPTO+COMGEANT);  $T_i$  tempo per l’inizializzazione; t/evento è il tempo totale impiegato dall’eseguibile per processare tutti gli eventi diviso il numero di tali eventi.

eventi	T(s)	$T_i$ (s)	t/evento(s)
894	1620	27	1.79
1 8977	15624	26	1.76

Tabella 4.3: Misure di performance sulle macchine con processore “Pentium III” (SI2K = 317). La quantità secondi/evento·SI2K = 563 è una costante (per i diversi tipi di PC) ed è relativa al numero di operazioni necessarie per processare un evento. T è il tempo totale di simulazione (compreso di inizializzazione e simulazione LEPTO+COMGEANT);  $T_i$  tempo per l’inizializzazione; t/evento è il tempo totale impiegato dall’eseguibile per processare tutti gli eventi diviso il numero di tali eventi.

catena LEPTO+COMGEANT. Sono state misurate le performance del sistema distinguendo tempo di inizializzazione e numero di eventi al secondo. Scegliendo il numero di eventi simulati per un singolo task tra 1000 e 10000, il suo range di inefficienza varia tra il 2% e 0.2%. L’inefficienza è definita come il rapporto tra il tempo  $T_i$  impiegato per l’inizializzazione del codice contenuto nel task (cioè dall’inizio dell’esecuzione del programma e la processatura del primo evento) ed il tempo totale  $T$  impiegato per la sua esecuzione (inizializzazione compresa). I tempi misurati sono riportati nelle tabelle 4.2 e 4.3.

Le misure sono nettamente dominate dall’esecuzione di COMGEANT che richiede, mediamente, un tempo di almeno due ordini di grandezza superiore rispetto a quella di LEPTO. Visto che i dati sono distribuiti intorno alla loro media a meno di qualche punto percentuale, in tabella sono riportati i valori massimi.

### Test di Scalabilità

La misura di scalabilità del sistema, è volta a determinare la sua capacità di ottenere un incremento di performance (per esempio il numero di eventi simulati in

un dato intervallo di tempo) proporzionalmente al numero di risorse computazionali accessibili. Questa è una caratteristica chiave di un sistema distribuito e richiede che l'assegnazione di "task" alle nuove risorse venga fatto senza incorrere in "colli di bottiglia".

Nel caso del sistema master-worker il limite al numero di risorse che lavorano in parallelo è dato dalla presenza di un unico master che deve gestire tutte le comunicazioni tra le applicazioni e che quindi può costituire il bottleneck che limita la scalabilità del sistema. In questo senso un primo progresso rispetto ai tradizionali batch system (LSF) è ottenuto con l'introduzione del meccanismo di pull: sono i worker che chiedono al master nuovi task da eseguire. Viene così evitato al master l'overhead di comunicazioni che deriva dalla periodica rilevazione dello stato del task in attesa che questo sia terminato. In un certo senso, la distribuzione dei task è determinata dalle risorse esistenti e funzionanti (un worker che non funzioni correttamente o che abbia problemi di comunicazione di fatto non contatta il master e non crea ritardi per le risorse "sane").

Il test di scalabilità è stato effettuato nella Farm di Trieste usando LSF come sistema batch. Su una macchina è stato fatto partire il master ed in seguito è stato aggiunto un numero progressivo (1, 2, 5, 10, 20, 50) di worker, vedi figura 4.9.

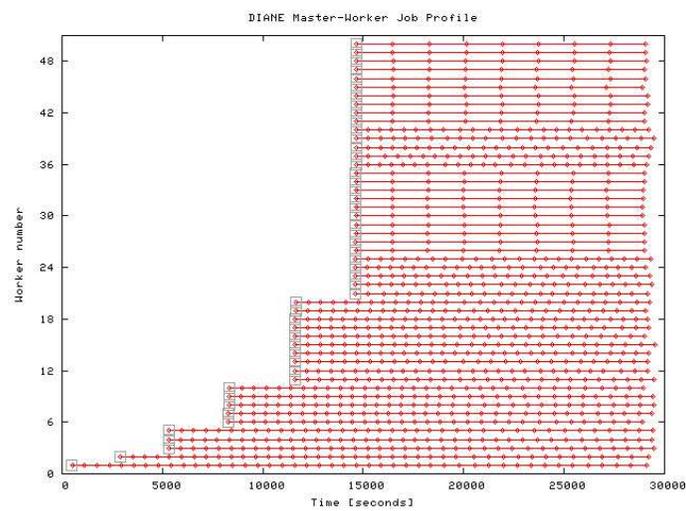


Figura 4.9: Andamento dell'attività dei worker in funzione del tempo. Ogni linea continua rappresenta un task in esecuzione e, dalla loro diversa lunghezza, si nota che le macchine della Farm non sono omogenee, vedi per esempio worker numero 26-35 e 40-49.

Sulla base dei dati osservati è stata misurata la scalabilità  $S$  secondo la formula

$$S = \frac{\sum_{i=0}^N k_i \cdot n_i \cdot t_i}{T} \quad (4.1)$$

dove:

- $N$  è il numero totale di task il cui intervallo di esecuzione è almeno parzialmente sovrapposto all'intervallo  $T$ .
- $T$  è l'intervallo scelto su cui valutare la scalabilità. Sono stati scelti intervalli in cui l'andamento dei worker fosse stabile, lontani dagli istanti in cui avveniva l'ingresso di nuovi worker (la lunghezza dell'intervallo  $T$  è maggiore della durata media di un task). Nel grafico riportato in figura 4.9 ciò corrisponde a selezionare un segmento  $T$  sull'asse delle ascisse e sommare la corrispondente attività dei worker, lungo l'ordinata.
- $k_i$  è la frazione del task contenuta in  $T$  (1 se il task è interamente contenuto)
- $n_i$  è il numero di eventi totale completato da un task,  $t_i$  è il tempo medio impiegato da COMGEANT per processare un evento. Il parametro  $t_i$  viene automaticamente calcolato dal programma COMGEANT e misura il tempo effettivo impiegato per la computazione della simulazione, scritto nel file log di output e recuperato per ogni task. Il termine  $\frac{t_i}{T}$  è un fattore adimensionale che permette di ottenere una misura di scalabilità indipendente dalle prestazioni della particolare macchina in cui il task è stato eseguito, esso riscalda automaticamente il numero di eventi prodotti dalla simulazione eseguita alla velocità di esecuzione della macchina, che dipende essenzialmente dalla cpu.

Come illustrato in figura 4.10, entro il numero di worker paralleli preso in considerazione, il sistema mostra una scalabilità molto buona cominciando a mostrare un certo degrado in corrispondenza di 50 worker concorrenti.

In generale è la frequenza di scambio di informazioni tra il master e l'insieme dei woker attivi che determina il funzionamento di DIANE. Definita tale frequenza come  $\nu = N/t$ , dove  $N$  è il numero di worker e  $t$  il tempo di esecuzione di un task, i test ufficiali mostrano che per  $\nu$  dell'ordine di  $10Hz$  DIANE permette un utilizzo stabile delle risorse mentre una frequenza  $\nu = 100Hz$  di solito mette a dura prova

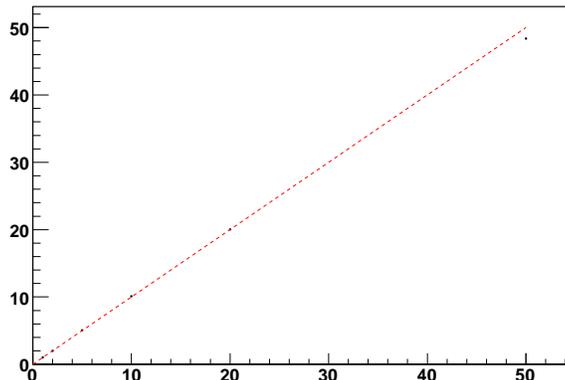


Figura 4.10: Scalabilità del sistema master-worker su LSF. In ordinata è indicata la scalabilità normalizzata alla misura fatta nel caso di un solo worker attivo; in ascissa c'è il numero di worker attivi al momento della misura. La retta tratteggiata ha coefficiente angolare uno e corrisponde all'andamento nel caso di perfetta scalabilità.

il sistema di connessioni master-worker. Nel caso delle simulazioni Monte Carlo usate per il test di scalabilità presentato la frequenza  $\nu$  è di  $\sim 0.1Hz$ , nel caso di 50 worker attivi. La deviazione dalla scalabilità osservata nell'ultimo punto del grafico 4.10 non è quindi imputabile al sistema DIANE, quanto piuttosto ad un difetto nella misura stessa. Il prodotto  $k_i \cdot n_i$  infatti fornisce una corretta stima della frazione di eventi simulati, nell'intervallo considerato, nella misura in cui la frazione di tempo impiegato dal programma per computare sia dominante su quella, per esempio, data dalle operazioni di input/output (I/O). Nel test condotto il flusso di I/O era, nell'ultima fase, essenzialmente costituito dall'output di 50 stream paralleli, dai nodi di esecuzione verso lo storage implementato su GPFS [63] all'interno della stessa sottorete, per un totale di dati non superiore a  $10MB/s$ . Le prestazioni tipiche di GPFS sono tali da giustificare l'uso di tale approssimazione.

La spiegazione del comportamento esibito da DIANE nel test può essere ragionevolmente ricondotta all'unico elemento sul quale non poteva essere esercitato un controllo: l'input/output. Non essendo lo spazio di storage a disposizione un sistema isolato, non si sono potute escludere attività di I/O concorrenti che disturbassero il test, benché si sia disposto il test in modo da escludere l'esecuzione concorrente sul sistema batch locale. Inoltre, in alcune occasioni successive al test

ed in contesti diversi, sono state osservate delle situazioni in cui il numero consistente ( $\sim 100$ ) di accessi in scrittura/lettura allo spazio disco disponibile sotto GPFS erano in grado di congestionare il sistema di storage. L'origine di tali problemi non è risultata chiara né facilmente investigabile.

Nonostante non si sia potuta dare alcuna spiegazione definitiva del fenomeno osservato nel test di scalabilità, vi sono forti indizi che l'origine sia da ascrivere ad un sovraccarico di I/O. Purtroppo, per motivi di tempo, non è stato possibile condurre ulteriori test chiarificativi.

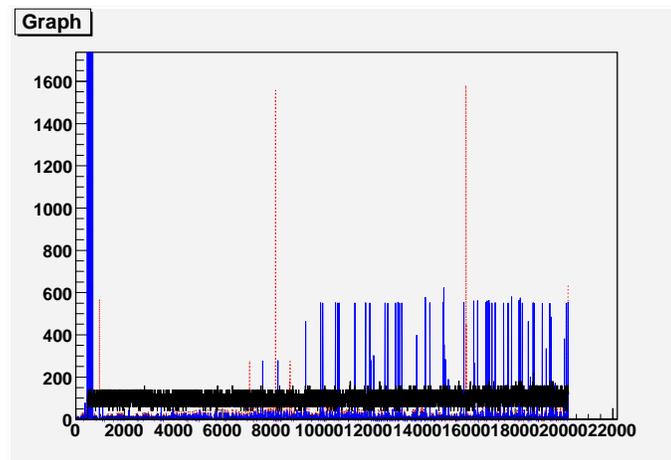


Figura 4.11: Andamento, in funzione del tempo, dell'attività del master. In ascissa è indicato il tempo in secondi (l'istante zero è lo stesso di quello scelto per il grafico in figura 4.9). La linea orizzontale mostra l'andamento dell'uso della cpu (in unità arbitrarie) pressoché costante mentre i picchi rappresentano i kB dei messaggi, in entrata ed in uscita, del master. Si nota come il flusso di dati tra master e worker è concentrato in piccoli intervalli di tempo che corrispondono sostanzialmente all'ingresso di nuovi worker (con la conseguente inizializzazione) ed al recupero (da parte del master) dei risultati.

Contemporaneamente alla misura di scalabilità è stato fatto un monitoraggio su alcune caratteristiche del master: il tempo in cui il programma usa la cpu (in modalità utente) sull'intervallo di tempo trascorso, entità dei messaggi scambiati in input ed output. Il loro andamento è mostrato in figura 4.11. L'utilizzo di cpu da parte del master è molto ridotto e costante (linea orizzontale), mentre si possono notare i picchi rappresentati dalla grandezza dei messaggi scambiati, nell'unità di tempo, che corrispondono allo scambio che avviene tra master e worker in fase di inizializzazione del worker e ad ogni completamento del task. Da tale grafico, che va sovrapposto al grafico di figura 4.9, si può notare che, nel limite del grado di concorrenzialità di worker raggiunto, né la cpu né la banda di trasmissione, a carico del master, hanno costituito un fattore limitante.



# Capitolo 5

## Risultati

### 5.1 Funzionamento del sistema distribuito per la simulazione

#### 5.1.1 Il sistema utilizzato

Già in fase di test, il sistema descritto nel capitolo precedente è stato utilizzato per la produzione Monte Carlo dei dati necessari per la stima dell'accettanza dell'apparato di COMPASS.

Nella configurazione del prototipo del sistema lo spazio di “storage” per i dati prodotti dalle simulazioni eseguite dai worker è allocato su una predefinita partizione dello spazio disco appartenente alla farm di Trieste (in grado di contenere  $\sim 6$  TB). Al momento delle prime produzioni non era ancora presente, sul gateway di accesso allo storage della farm, un nodo abilitato a supportare trasferimenti di file con protocolli di autenticazione Grid (gsiftp). Ciò ha reso impossibile la copia diretta dei file prodotti dalle macchine in cui hanno operato i worker sottomessi in Grid. L'unico modo per ovviare a questo problema è stato l'uso della macchina esterna alla farm, sulla quale erano presenti le librerie gLite e che ospita la UI usata per la sottomissione dei worker, come gateway GSI-enabled verso lo storage della farm. In tal modo i file prodotti venivano copiati dai worker su uno spazio di memoria di 20 GB che fungeva da buffer di transito verso la destinazione finale nella farm, dotata solo di autenticazione TLS/SSL.

I file in arrivo sullo storage di transito venivano trasferiti sulla farm in automatico da un daemon che ne monitorava la disponibilità alla copia, attraverso un

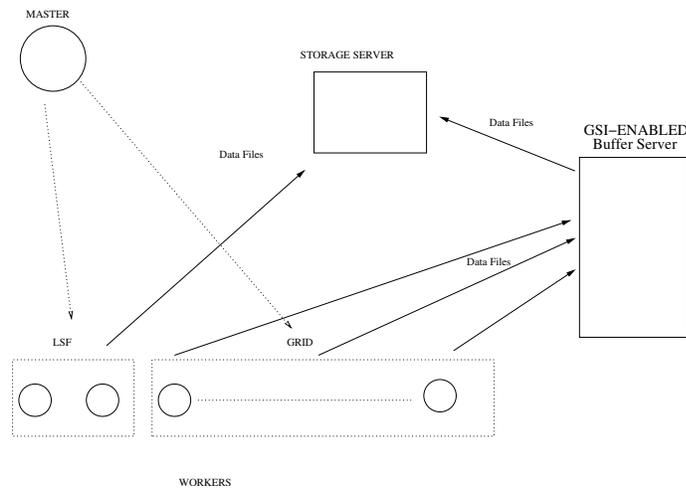


Figura 5.1: Percorso dei file dati prima di raggiungere lo spazio di “storage” apposito. Nella figura vengono evidenziati i due differenti meccanismi

meccanismo di tagging/locking implementato con la copia di un apposito file collegato a quello principale. L'andamento previsto dello spazio occupato ( $S$ ) sulla UI, in funzione del tempo, è mostrato in figura 5.2. Una prima fase avviene nell'intervallo di tempo (proporzionale al numero di eventi simulati in ogni task di DIANE) in cui i worker sottomessi in Grid copiano i rispettivi file dati prodotti contemporaneamente senza che il meccanismo implementato per il trasferimento sulla farm entri in funzione (i file che marcano i dati come trasferibili iniziano ad arrivare dopo la prima fase). In prima approssimazione si può considerare che  $S$  cresca linearmente con il tempo  $t$  (le unità nel grafico 5.2 sono arbitrarie) con un rapporto incrementale pari alla velocità di trasferimento (in Byte/secondo) con cui i file vengono copiati da remoto. Per tempi maggiori di  $t_1$  (cioè superata la prima fase) il meccanismo di copia implementato nel sistema contribuisce a ridurre tale incremento e la tangente alla retta diventa la differenza tra la velocità in uscita ed in entrata. In questo caso la differenza è positiva e quindi ad un certo istante si ha che lo spazio disponibile sulla macchina di transito (che è molto minore dello spazio occupato dalla somma dei file dati relativi ad un tipico task della base dati) si esaurisce causando la terminazione di tutti i processi di copia (in entrata) attivi. È stato quindi implementato un ulteriore meccanismo di blocco in stato di attesa di qualsiasi richiesta di copia proveniente dai worker attivi in Grid, che doveva agire ogni volta che lo spazio disco occupato avesse superato una certa soglia. È stato convenzionalmente utilizzato un

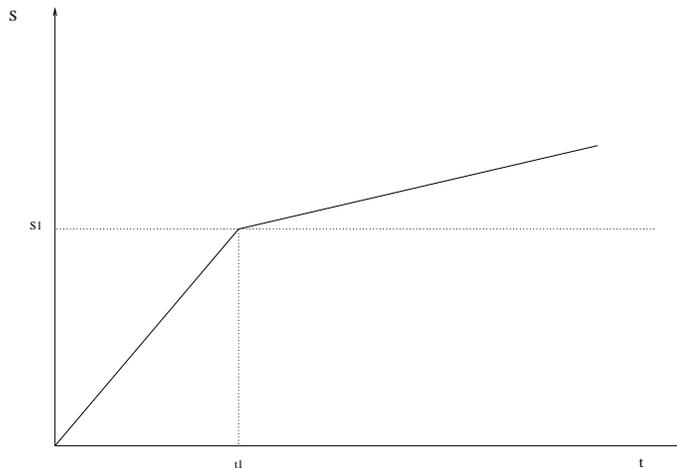


Figura 5.2: Andamento previsto per la quantità di spazio disco occupato  $S$  sulla macchina di transito, in funzione del tempo  $t$ . Si può distinguere una prima fase in cui i file arrivano, ma il meccanismo di trasferimento verso lo spazio di storage della farm non è innescato finché i primi file non sono completi, ed una seconda in cui la ricezione in input dei file dati ottenuti dai worker avviene in contemporanea con il trasferimento verso lo storage dei file già completamente trasferiti.

file “flag” che fosse rilevato dal worker prima che quest’ultimo iniziasse la richiesta di trasferimento; tale flag veniva rimosso qualora lo spazio disco disponibile si fosse sufficientemente liberato, grazie al trasferimento di file verso lo storage.

Ovviamente il numero di worker  $N_w$  da sottoporre in Grid doveva essere tale per cui lo spazio disco  $S1 \sim N_w \cdot S_t$  raggiunto sulla macchina di transito, al tempo  $t1$ , fosse ben minore di  $\sim 20GB$ ; con  $S_t$  lo spazio occupato da un file ottenuto da una singola simulazione assegnata al worker di volta in volta.

Questo complesso ed inefficiente meccanismo di recupero dei file è stato abbandonato non appena il servizio del Centro di Calcolo è stato in grado di approntare il servizio di autenticazione GSI sui gateway di accesso allo storage della farm.

Nei test di riproduzione si sono riscontrati alcuni problemi legati alla comunicazione tra Master, Worker e Directory Service implementata in DIANE, (ricordiamo che è un software ancora in fase di sviluppo), e che sono stati segnalati agli sviluppatori.

### 5.1.2 Produzione finale

Dopo i test preliminari è stata effettuata una produzione massiccia di  $2 \cdot 10^7$  eventi, con l'abilitazione al trasferimento del gateway della farm. È quindi stata valutata l'efficienza del sistema (come rapporto tra il numero di eventi recuperati e quello degli eventi richiesti nel task) e sono state monitorate in funzione del tempo: l'utilizzo di cpu del master e del directory service e l'entità dei messaggi scambiati tra master e worker, vedi figura 5.4; l'attività dei worker, sia che fossero stati sottomessi in locale, sia in Grid e mostrata in figura 5.3. Purtroppo i dati provenienti da questa produzione non sono stati utilizzati per la stima dell'accettanza per mancanza di tempo.

Nella figura 5.3, analoga a quella presente nel capitolo 4 al paragrafo 4.5.1, le linee continue rappresentano i worker, indicati in ordinata, che stanno eseguendo un task. I worker indicati in ordinata con i numeri dallo 0 al 20 sono quelli sottomessi nel sistema batch locale e, come si vede dal grafico, hanno lavorato in modo efficiente per tutta la durata della produzione; tutti gli altri sono stati sottomessi in Grid. L'attività di questi ultimi non è stata regolare per tutti, tuttavia il meccanismo grazie al quale il worker che ha problemi non chiede nessun task al master, ha reso possibile il completamento della produzione con una efficienza pari al 99.3%.

Dai grafici mostrati nella figura 5.4 si vede che il carico di cpu del master e del directory service è praticamente nullo durante tutta la durata della produzione e l'entità dei messaggi scambiati, costituita essenzialmente dai messaggi di log che il worker consegnava al master al termine dell'esecuzione di ogni task, limitata a poche decine di kB per worker attivo ogni secondo (ben al di sotto della banda trasmissiva a disposizione).

## 5.2 Distribuzioni azimutali

### 5.2.1 L'analisi dei dati

Per lo studio delle asimmetrie azimutali è stata analizzata la distribuzione  $N(\phi)$  dell'angolo azimutale  $\phi$  dell'adrone prodotto (nel GNS), ottenuta dal rapporto  $\frac{n(\phi)}{\alpha(\phi)}$  dove  $n(\phi)$  è la distribuzione ottenuta dall'analisi dei dati raccolti nei periodi menzionati alla fine del capitolo 2 opportunamente pesati con il metodo descritto nel

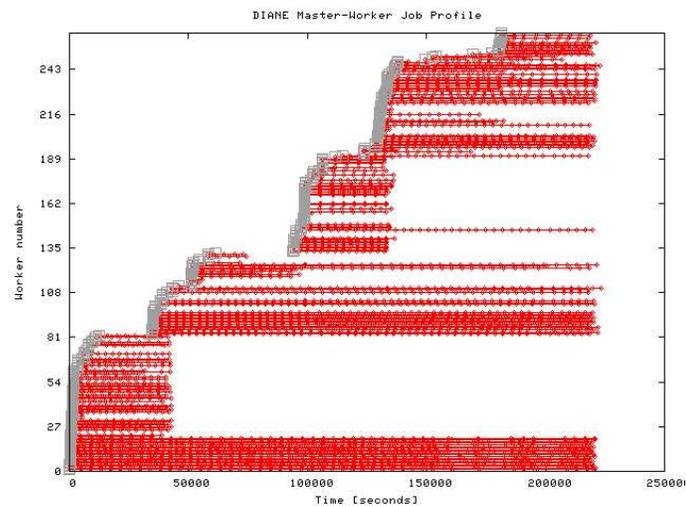


Figura 5.3: Grafico dell'attività dei worker sottomessi per la produzione finale. In ordinata è indicato il numero del worker ed in ascissa il tempo espresso in secondi. I worker tra 0 e 20 sono quelli sottomessi nel sistema batch locale, mentre i restanti worker sono stati sottomessi in Grid. L'attività dei worker locali risulta praticamente costante, il motivo per cui i worker 20-80 hanno terminato contemporaneamente è stato un errore nell'impostazione del rinnovo automatico del grid-proxy, che altrimenti dura 12 ore, risolto prima della sottomissione dei worker successivi. Le interruzioni "premature" degli altri worker in esecuzione in Grid, hanno invece origine da instabilità dei CE di destinazione.

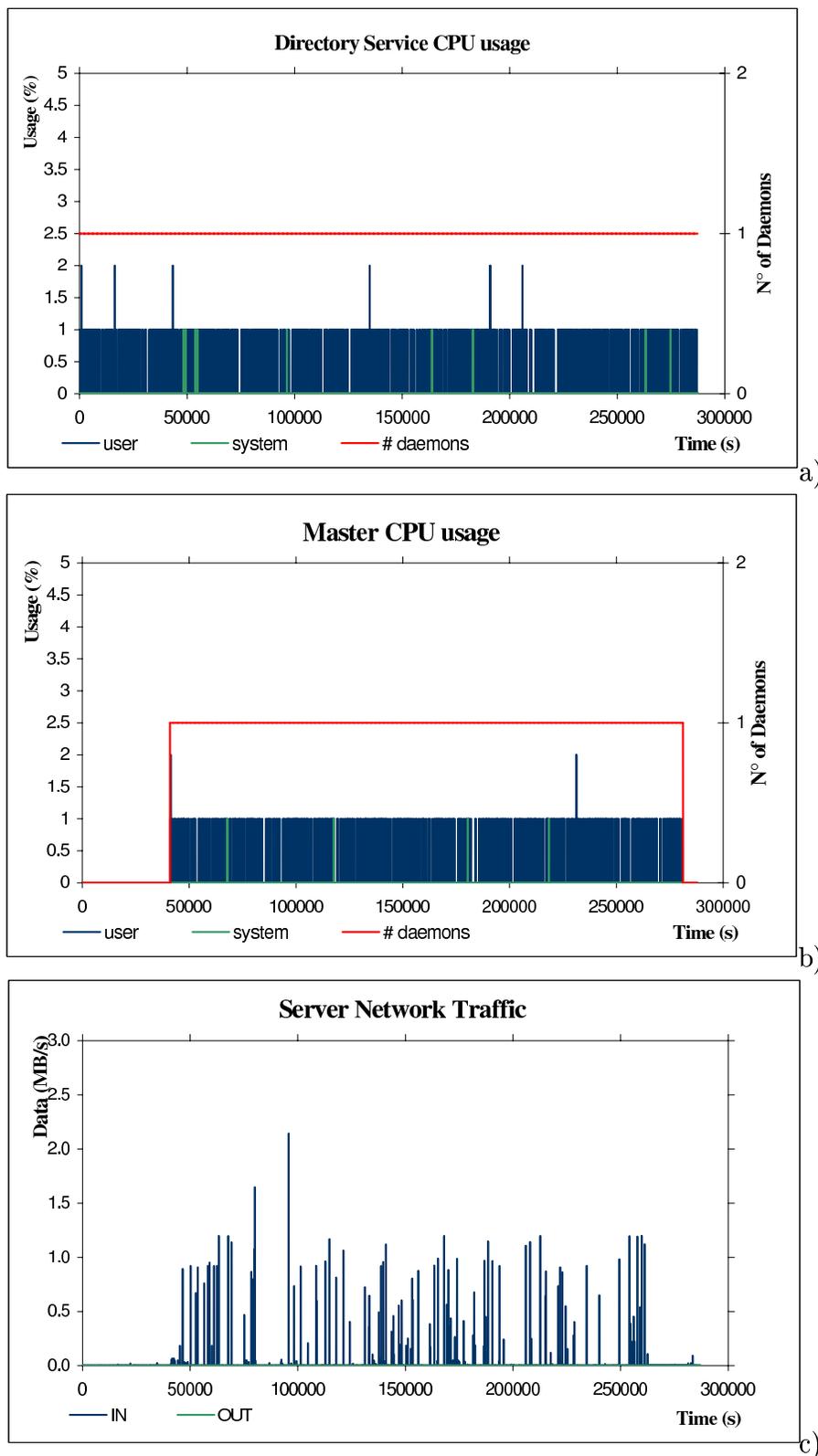


Figura 5.4: Nei tre grafici sopra riportati è mostrato l'andamento dell'utilizzo di cpu da parte del DS, grafico a), del master, grafico b), ed il traffico di rete generato dai messaggi scambiati tra master e worker. I grafici vanno raffrontati con quelli di figura 5.3

$0.1 < y < 0.9$ $Q^2 > 1(\text{GeV}/c)^2$ $W > 5\text{GeV}/c^2$
$0.2 < z < 0.85$ $0.1 < P_T < 1.5(\text{GeV}/c)$
$\theta_h^{ab} > 0.020 \text{ rad } (*)$

Tabella 5.1: Tabella riassuntiva dei tagli applicati agli eventi generati (simulati) e ricostruiti (simulati e reali). (\*) Il taglio  $\theta_h^{ab} > 0.020 \text{ rad}$  è stato applicato solo ai dati ricostruiti (reali e simulati)

capitolo 3 in modo da cancellare eventuali effetti di asimmetria dovuti alla polarizzazione (trasversa) del bersaglio. Oltre ai tagli standard dell'analisi, per ottenere  $n(\phi)$ , sono stati applicati i tagli discussi nel paragrafo 3.4 del capitolo 3 per avere un migliore accordo con la simulazione. I tagli finali sono riassunti nella tabella 5.1.  $\alpha(\phi) = \frac{N_R^{mc}(\phi)}{N_G^{mc}(\phi)}$  è l'accettanza dell'apparato sperimentale ottenuta dal rapporto tra la distribuzione dell'angolo azimutale dell'adrone ricostruito nella simulazione  $N_R^{mc}(\phi)$  e  $N_G^{mc}(\phi)$ , che è la distribuzione corrispondente per l'adrone generato.

La distribuzione finale dell'angolo azimutale dell'adrone  $N(\phi)$  da cui si ricavano le asimmetrie azimutali, è stata ottenuta dividendo l'intervallo  $(-\pi, \pi)$  in 16 parti uguali così da avere una larghezza di bin molto maggiore della risoluzione angolare dell'apparato, che da studi Monte Carlo [34] è stata calcolata in  $\sim 0.07 \text{ rad}$ .

La funzione con cui è stato effettuato il fit alla distribuzione  $N(\phi)$  è:

$$f(\phi) = A \cdot (1 + A_{\cos\phi} \cdot \cos(\phi) + A_{\cos 2\phi} \cdot \cos(2 \cdot \phi) + A_{\sin\phi} \cdot \sin(\phi)) \quad (5.1)$$

dove  $A$ ,  $A_{\cos\phi}$ ,  $A_{\cos 2\phi}$ ,  $A_{\sin\phi}$  sono i parametri liberi tra cui  $A_{\cos\phi}$ ,  $A_{\cos 2\phi}$ ,  $A_{\sin\phi}$  sono le asimmetrie azimutali che si vogliono misurare.

Sono stati analizzati separatamente adroni con carica positiva e negativa poiché l'accettanza non è la stessa e a priori le asimmetrie azimutali potrebbero essere diverse a seconda della carica dell'adrone prodotto. Per avere la dipendenza dei valori delle diverse asimmetrie in funzione di  $x$ , il campione di eventi a disposizione è stato suddiviso in 5 parti, in base al valore di  $x$ , come mostrato in tabella 5.2. Nella tabella 5.2 vengono inoltre indicati i valori medi calcolati per le variabili cinematiche  $x$ ,  $y$ ,  $Q^2$  ed il numero di adroni ricostruiti per due periodi di dati considerati a valle di tutti i tagli applicati nell'analisi descritta. Il valore medio di  $x$  di ciascuna

Carica	Bin di x	$\langle x \rangle$	$\langle y \rangle$	$\langle Q^2 \rangle$
-	$0.002 < x < 0.020$	0.012	0.493	3.475
	$0.020 < x < 0.045$	0.031	0.273	3.687
	$0.045 < x < 0.080$	0.059	0.240	4.638
	$0.080 < x < 0.125$	0.099	0.236	6.729
	$0.125 < x < 1.000$	0.202	0.240	14.519
+	$0.002 < x < 0.020$	0.012	0.500	3.475
	$0.020 < x < 0.045$	0.031	0.275	3.682
	$0.045 < x < 0.080$	0.059	0.242	4.641
	$0.080 < x < 0.125$	0.099	0.237	6.771
	$0.125 < x < 1.000$	0.199	0.238	14.295

Tabella 5.2: Suddivisione del numero di adroni nei 5 intervalli di  $x$ . Per ogni bin di  $x$  sono state valutate le asimmetrie separatamente per adroni positivi e negativi. Per ognuno dei 5 bin è riportato: il range di  $x$  ed il valore medio di  $x$ , di  $y$  e di  $Q^2$ .

variabile è calcolato come

$$\langle x \rangle = \frac{\int x f(x) dx}{\int f(x) dx}; \quad (5.2)$$

dove  $f(x)$  è la distribuzione finale di  $x$  dell'adrone, corretta per l'accettanza. Il numero totale di eventi, generati e reali, è dato in tabella 5.3.

Per ciascun intervallo di  $x$  e separatamente per adroni positivi e negativi, è stata ripetuta l'analisi completa e sono state estratte le relative asimmetrie azimutali. I risultati sono mostrati nel paragrafo 5.3.1.

## 5.2.2 La stima dell'accettanza

Per la stima dell'accettanza sono stati generati  $10^7$  eventi Monte Carlo utilizzando il sistema sviluppato per COMPASS descritto nel paragrafo precedente.

Per minimizzare il numero degli eventi da simulare, sono stati generati, con LEPTO, già con i seguenti tagli cinematici:

- $0.003 < x < 0.99$ ;
- $0.05 < y < 0.95$ ;
- $W > 4.00 \text{ GeV}/c^2$ ;
- $Q^2 > 0.90 (\text{GeV}/c)^2$ ;

Carica	Numero di adroni			
	Bin di x	ricostruiti (dati)	generati (MC)	ricostruiti (MC)
+	$0.002 < x < 0.020$	872970	1126886	379799
	$0.020 < x < 0.045$	1390735	937105	275011
	$0.045 < x < 0.080$	614446	393875	105598
	$0.080 < x < 0.125$	276525	177321	46486
	$0.125 < x < 1.000$	211161	176131	36308
-	$0.002 < x < 0.020$	1092073	1126886	379799
	$0.020 < x < 0.045$	2482808	2064091	658810
	$0.045 < x < 0.080$	3097254	2457966	764408
	$0.080 < x < 0.125$	3373779	2635287	810894
	$0.125 < x < 1.000$	3584940	2811418	847202

Tabella 5.3: Per le due cariche degli adroni e per i diversi bin di  $x$  usati, c'è il numero di adroni ricostruiti, dai dati, generati e ricostruiti, dalla simulazione.

leggermente meno restrittivi rispetto ai tagli che sono stati applicati in fase di ricostruzione, processatura ed analisi per evitare eventuali effetti di bordo, nelle distribuzioni finali, dovuti alla risoluzione con cui le varibili cinematiche vengono ricostruite. La catena di processi attraverso i quali si ottengono i campioni finali di dati simulati è illustrata in figura 5.5.

Alla fine della catena LEPTO+COMGEANT+CORAL vengono estratte ed analizzate le informazioni riguardanti gli adroni ricostruiti o generati a seconda del programma PHAST che viene usato per leggere i mDST ottenuti.

L'insieme dei tagli definitivi applicati nella fase di analisi per i dati simulati generati e ricostruiti e per i dati reali è riportato in tabella 5.1. Per i dati ricostruiti dalla simulazione, è stata implementata l'efficienza degli odoscopi con il metodo descritto nel paragrafo 3.4.1 del capitolo 3.

Alla stima dell'accettanza  $\alpha = \frac{N_R^{mc}(\phi)}{N_G^{mc}(\phi)}$  è stato associato un errore di tipo binomiale, come spiegato nel paragrafo 3.2 del capitolo 3, dato dalla seguente formula:

$$\sigma_\alpha^2 = \frac{N_R^{mc}}{(N_G^{mc})^2}. \quad (5.3)$$

### 5.2.3 La distribuzione azimutale

L'errore  $\sigma_N$ , relativo a ciascun bin di  $\phi$ , della distribuzione  $N(\phi)$  finale è stato calcolato dalla somma in quadratura degli errori statistici delle distribuzioni  $n(\phi)$

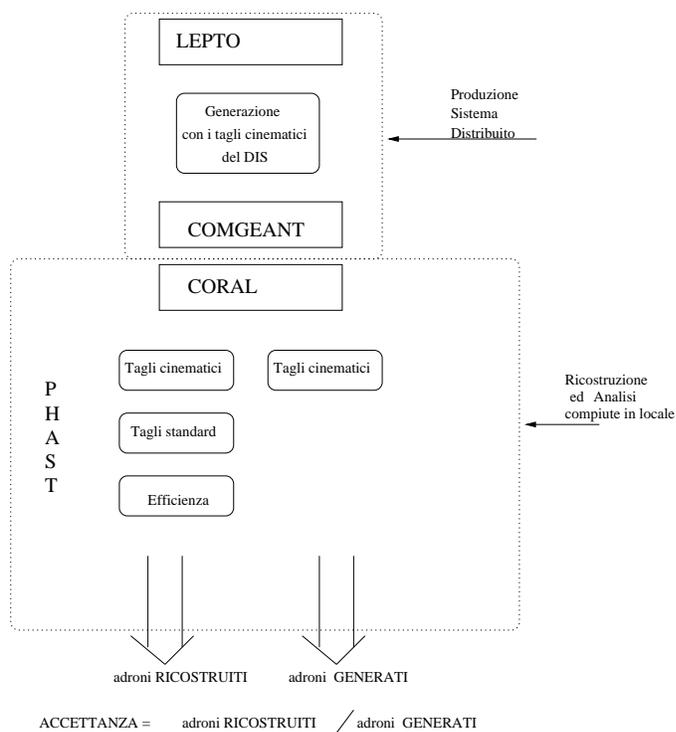


Figura 5.5: Schematica rappresentazione della catena di processi utilizzata per ottenere i file finali contenenti le informazioni necessarie per la costruzione della distribuzione  $N(\phi)$  da cui si ricavano le asimmetrie azimutali.

ed  $\alpha(\phi)$  e, nella formula seguente, sono indicati come  $\sigma_n$  e  $\sigma_\alpha$  rispettivamente:

$$\sigma_{N^j} = N^j \cdot \sqrt{\left(\frac{\sigma_{n^j}}{n^j}\right)^2 + \left(\frac{\sigma_{\alpha^j}}{\alpha^j}\right)^2}, \quad (5.4)$$

dove l'indice  $j$  si riferisce al bin di  $\phi$  preso in considerazione.

$$\sigma_{n^j} = \sqrt{(\sigma_{n1^j})^2 + (\sigma_{n2^j})^2} \quad (5.5)$$

in cui  $\sigma_{n1^j}$  e  $\sigma_{n2^j}$  sono gli errori delle distribuzioni del  $\phi$  dell'adrone relative al primo ed al secondo periodo rispettivamente, calcolate per ognuno dei sedici bin di  $\phi$ .

$\sigma_{n1^j}$  e  $\sigma_{n2^j}$  sono stati calcolati con la formula espressa nell'equazione 5.6 dove sono omessi gli indici  $j$  e gli indici 1 e 2 per non appesantirne la lettura.

$$\begin{aligned} \sigma_n^2 = & (P_u^- \cdot (P_d^+ + P_d^-) \cdot N_u^+ + P_d^+ \cdot (P_u^+ + P_u^-) \cdot N_d^-)^2 \sigma_b^2 + \\ & + (b \cdot P_u^- \cdot (P_d^+ + P_d^-))^2 \sigma_{N_u^+}^2 + \\ & + (P_u^+ \cdot (P_d^+ + P_d^-))^2 \sigma_{N_u^-}^2 + \\ & + (P_d^- \cdot (P_u^+ + P_u^-))^2 \sigma_{N_d^+}^2 + \\ & + (b \cdot P_d^+ \cdot (P_u^+ + P_u^-))^2 \sigma_{N_d^-}^2, \end{aligned} \quad (5.6)$$

dove la notazione è la stessa di quella adottata nel paragrafo 3.2 del capitolo 3. L'errore è stato ottenuto assumendo che il numero di adroni ricostruiti per un dato sottoperiodo e con vertice primario appartenente ad una cella sia indipendente dalla misura del fascio e dal numero di adroni ottenuto per un'altra cella o sottoperiodo e assumendo poissoniano l'errore associato sia al numero di adroni che a quello di muoni incidenti. Nell'equazione 5.6 i contributi dei singoli errori vengono sommati in quadratura.

Nelle figure 5.2.3 e 5.2.3 sono mostrate le distribuzioni  $n(\phi)$  dell'angolo azimutale dell'adrone (positivo e negativo rispettivamente) ricostruiti dai dati, a valle di tutti i tagli descritti in tabella 5.1, per ciascuno dei bin di  $x$  presi in considerazione. Da questi grafici emerge che la distribuzione angolare dell'adrone dipende molto fortemente da  $x$ , confrontando per esempio le distribuzioni di adroni positivi relative al primo ed all'ultimo bin di  $x$  si nota che hanno forma opposta; effetti simili si hanno anche per adroni negativi.

Si riscontra una forte dipendenza da  $x$  delle distribuzioni dell'angolo azimutale dell'adrone anche nell'accettanza  $\alpha(\phi)$  valutata con la simulazione Monte Carlo, come si vede nelle figure 5.2.3 e 5.2.3 dove sono mostrate le accettanze per i diversi bin di  $x$  per adroni positivi e negativi rispettivamente.

Le distribuzioni finali  $N(\phi) = \frac{n(\phi)}{\alpha(\phi)}$  da cui si estraggono le asimmetrie sono mostrate invece nelle figure 5.2.3 e 5.2.3 per adroni positivi e negativi rispettivamente e la curva rappresenta la funzione ottenuta dal fit. Nonostante la forte dipendenza da  $x$  sia di  $n$  che di  $\alpha$  la distribuzione mostra un andamento simile nei diversi bin di  $x$  presi in considerazione.

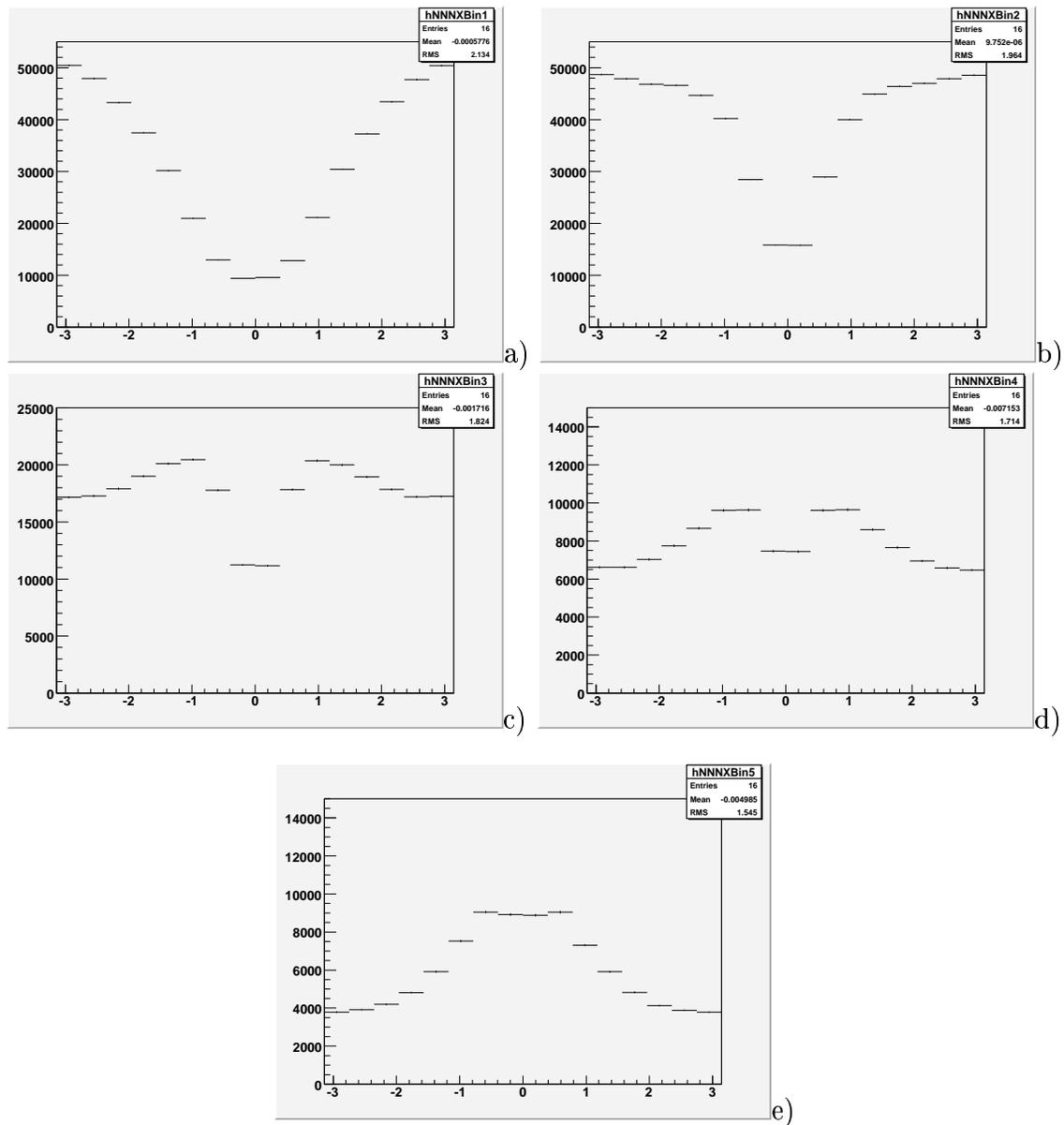


Figura 5.6: Distribuzioni  $n(\phi)$  azimutali non corrette per l'accettanza per adroni positivi ricostruiti dai dati a valle di tutti i tagli nei diversi bin di  $x$ . Nella figura a) eventi con  $x < 0.020$ ; nella figura b) eventi con  $0.020 < x < 0.045$ ; nella figura c) eventi con  $0.045 < x < 0.080$ ; nella figura d) eventi con  $0.080 < x < 0.125$ ; nella figura e) eventi con  $x > 0.125$ .

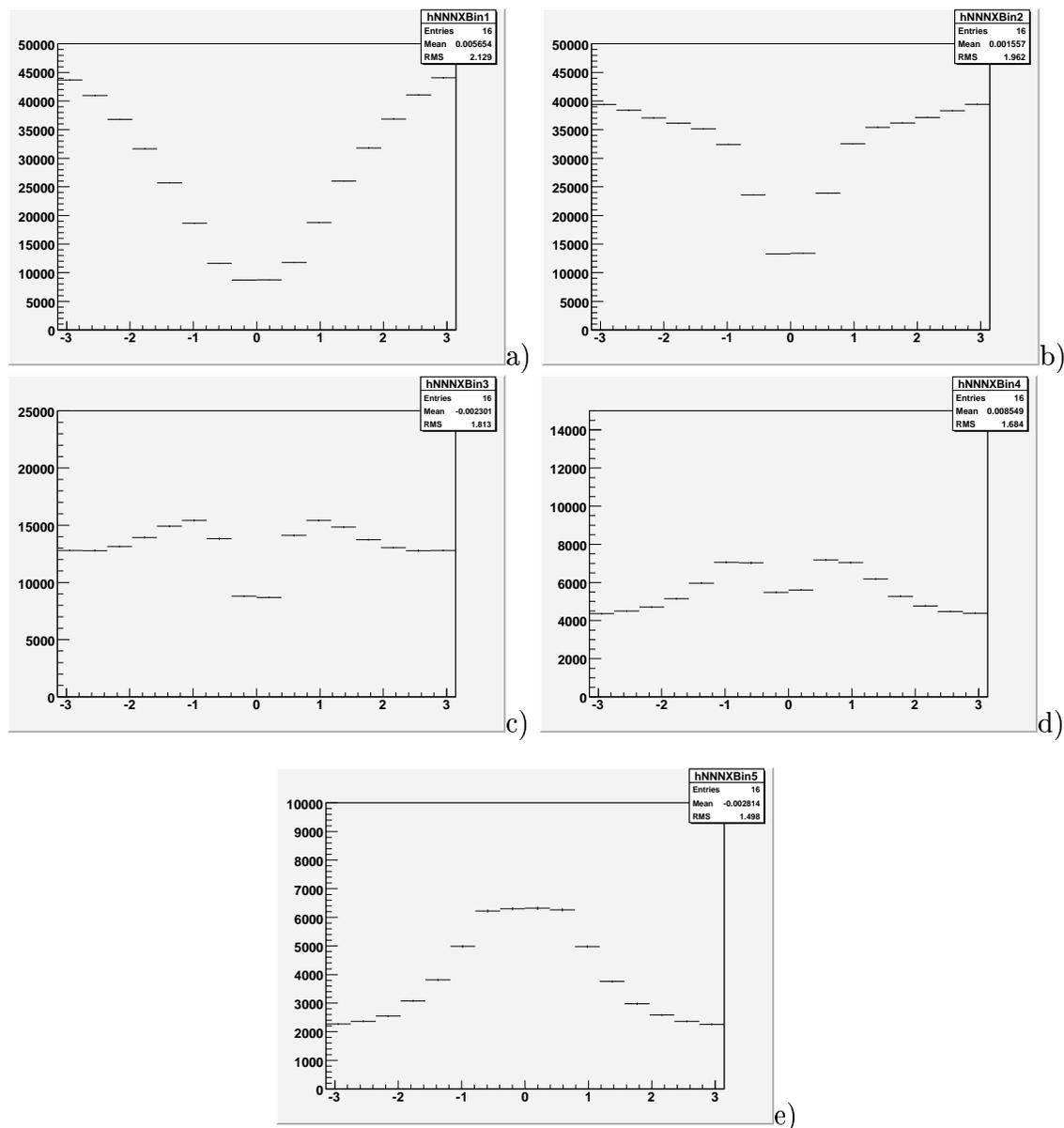


Figura 5.7: Distribuzioni  $n(\phi)$  azimutali non corrette per l'accettanza per adroni negativi ricostruiti dai dati a valle di tutti i tagli nei diversi bin di  $x$ . Nella figura a) eventi con  $x < 0.020$ ; nella figura b) eventi con  $0.020 < x < 0.045$ ; nella figura c) eventi con  $0.045 < x < 0.080$ ; nella figura d) eventi con  $0.080 < x < 0.125$ ; nella figura e) eventi con  $x > 0.125$ .

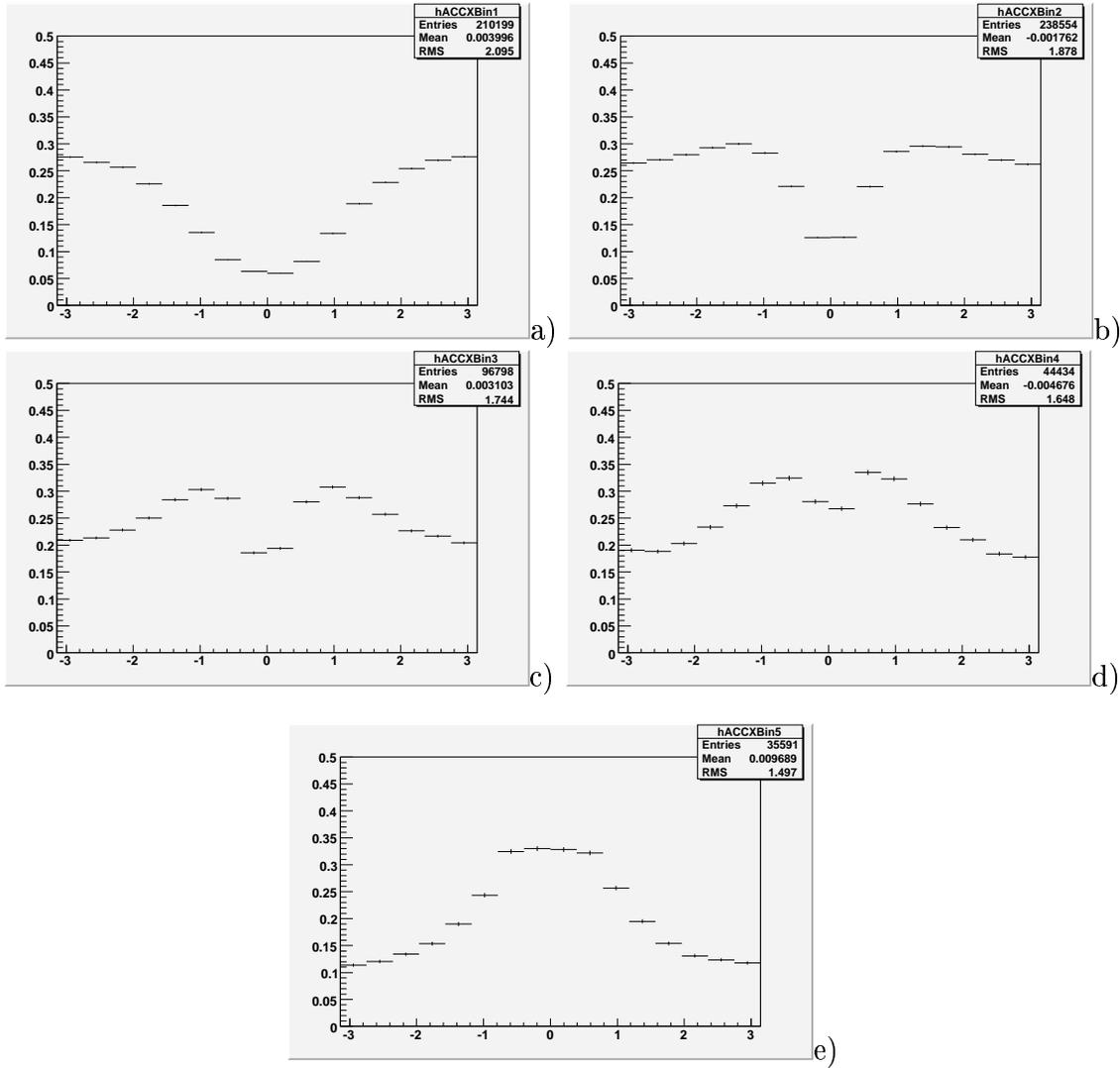


Figura 5.8: Andamento dell'accettanza  $\alpha(\phi)$  stimata con la simulazione Monte Carlo per adroni positivi ricostruiti dai dati reali a valle di tutti i tagli nei diversi bin di  $x$ . Nella figura a) eventi con  $x < 0.020$ ; nella figura b) eventi con  $0.020 < x < 0.045$ ; nella figura c) eventi con  $0.045 < x < 0.080$ ; nella figura d) eventi con  $0.080 < x < 0.125$ ; nella figura e) eventi con  $x > 0.125$ .

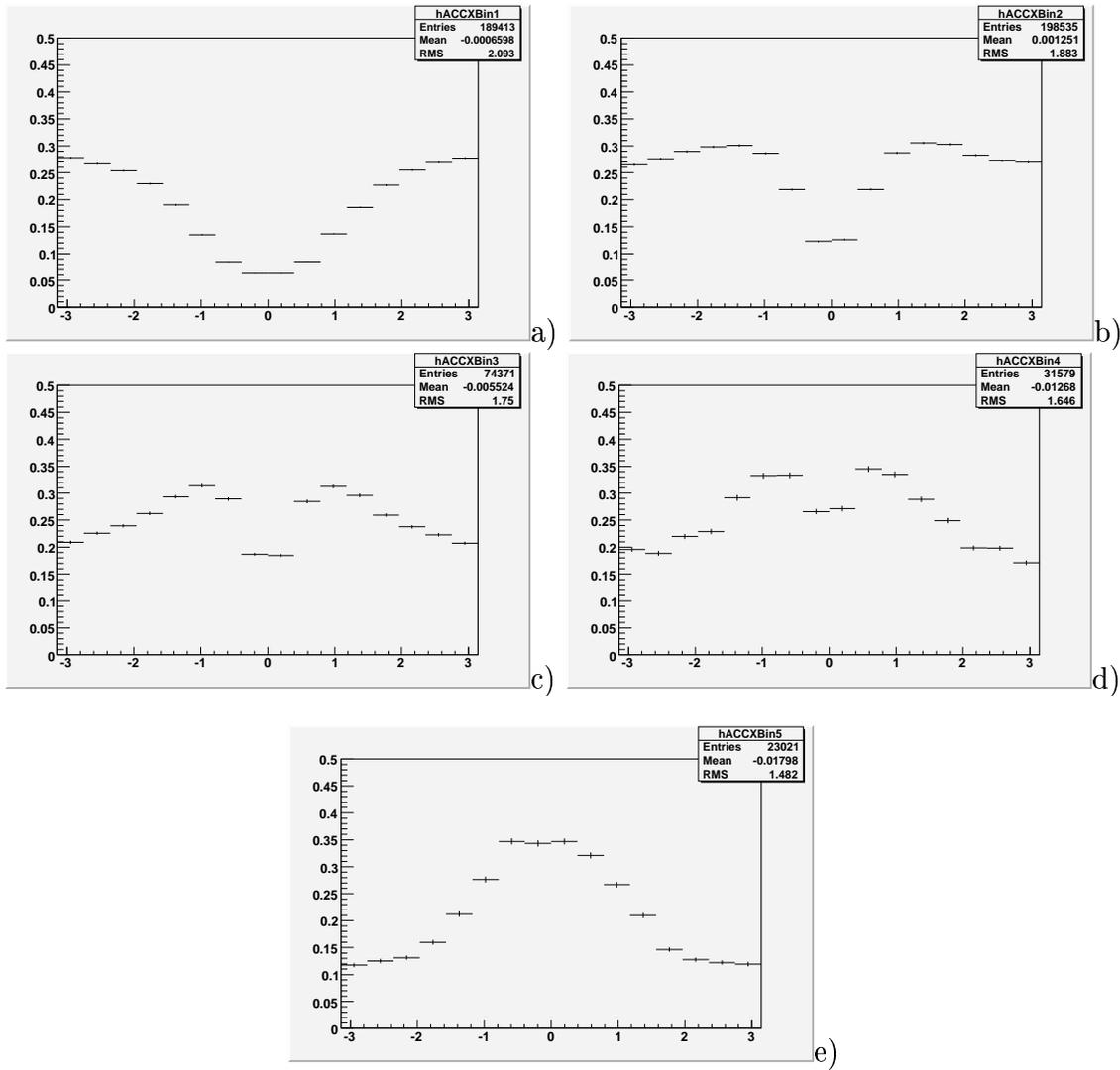


Figura 5.9: Andamento dell'accettanza  $\alpha(\phi)$  stimata con la simulazione Monte Carlo per adroni negativi ricostruiti dai dati reali a valle di tutti i tagli nei diversi bin di  $x$ . Nella figura a) eventi con  $x < 0.020$ ; nella figura b) eventi con  $0.020 < x < 0.045$ ; nella figura c) eventi con  $0.045 < x < 0.080$ ; nella figura d) eventi con  $0.080 < x < 0.125$ ; nella figura e) eventi con  $x > 0.125$ .

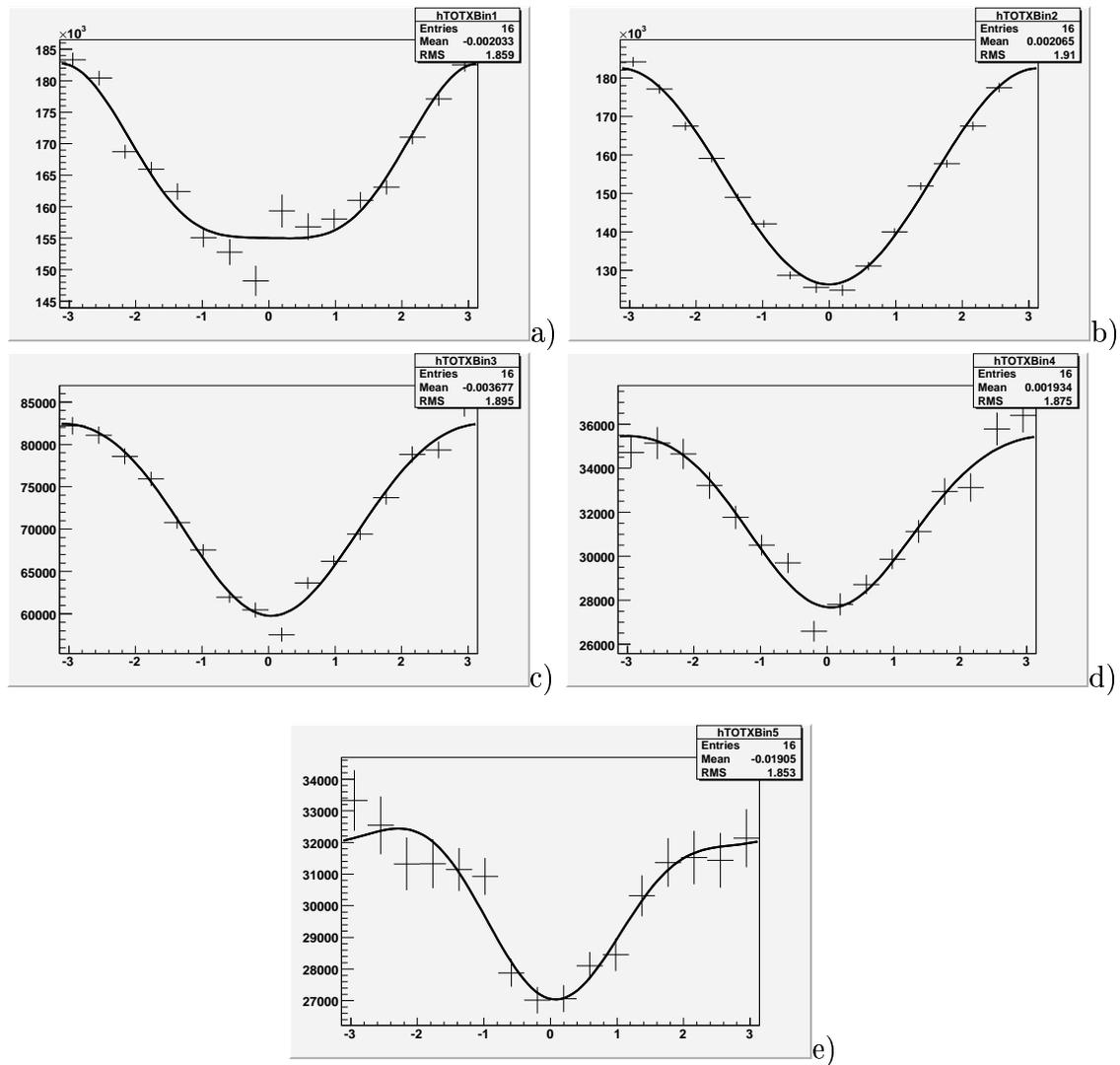


Figura 5.10: Distribuzione  $N(\phi)$  ottenuta dai dati reali e corretta per l'accettanza per adroni positivi ricostruiti a valle di tutti i tagli e per ciascun bin di  $x$ . Ad essa è sovrapposta la curva di fit. Nella figura a) eventi con  $x < 0.020$ ; nella figura b) eventi con  $0.020 < x < 0.045$ ; nella figura c) eventi con  $0.045 < x < 0.080$ ; nella figura d) eventi con  $0.080 < x < 0.125$ ; nella figura e) eventi con  $x > 0.125$ .

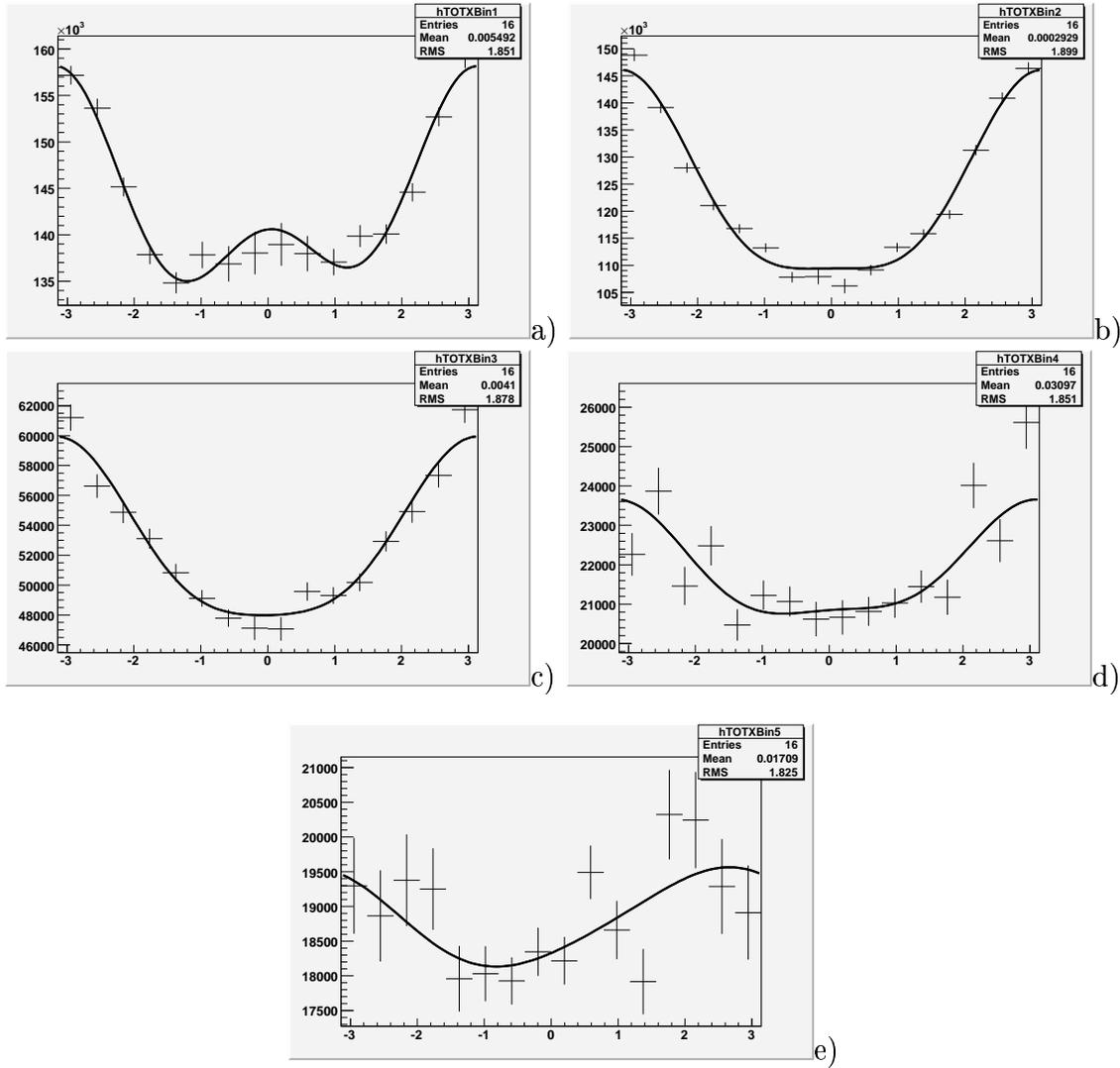


Figura 5.11: Distribuzione  $N(\phi)$  ottenuta dai dati reali e corretta per l'accettanza per adroni negativi ricostruiti a valle di tutti i tagli e per ciascun bin di  $x$ . Ad essa è sovrapposta la curva di fit. Nella figura a) eventi con  $x < 0.020$ ; nella figura b) eventi con  $0.020 < x < 0.045$ ; nella figura c) eventi con  $0.045 < x < 0.080$ ; nella figura d) eventi con  $0.080 < x < 0.125$ ; nella figura e) eventi con  $x > 0.125$ .

Carica	$\langle x \rangle$	$A_{\cos\phi}/2$	$A_{\cos 2\phi}/2$	$A_{\sin\phi}/2$	$\chi^2/\text{NDF}$
-	0.012	$-0.031 \pm 0.002$	$0.021 \pm 0.002$	$0.0028 \pm 0.0015$	1.6
	0.031	$-0.0746 \pm 0.0015$	$0.0195 \pm 0.0015$	$0.0007 \pm 0.0013$	5.7
	0.059	$-0.057 \pm 0.002$	$0.0115 \pm 0.0024$	$0.0009 \pm 0.0021$	2.0
	0.099	$-0.032 \pm 0.004$	$0.009 \pm 0.004$	$0.0029 \pm 0.0035$	3.1
	0.202	$-0.015 \pm 0.005$	$0.002 \pm 0.004$	$0.011 \pm 0.005$	1.7
+	0.012	$-0.042 \pm 0.002$	$0.0105 \pm 0.0015$	$-0.0006 \pm 0.0013$	3.2
	0.031	$-0.091 \pm 0.001$	$0 \pm 0.001$	$0.0005 \pm 0.0011$	4.6
	0.059	$-0.079 \pm 0.002$	$-0.006 \pm 0.002$	$-0.004 \pm 0.002$	2.3
	0.099	$-0.061 \pm 0.003$	$-0.007 \pm 0.003$	$-0.005 \pm 0.003$	1.6
	0.199	$-0.041 \pm 0.004$	$-0.013 \pm 0.004$	$-0.007 \pm 0.004$	1.0

Tabella 5.4: Valori misurati per le diverse asimmetrie previste per la distribuzione dell'angolo azimutale dell'adrone.

## 5.3 Estrazione delle asimmetrie azimutali

### 5.3.1 Le asimmetrie misurate

Le asimmetrie azimutali in ogni bin di  $x$  sono state ottenute eseguendo un fit delle distribuzioni finali  $N$  con la funzione

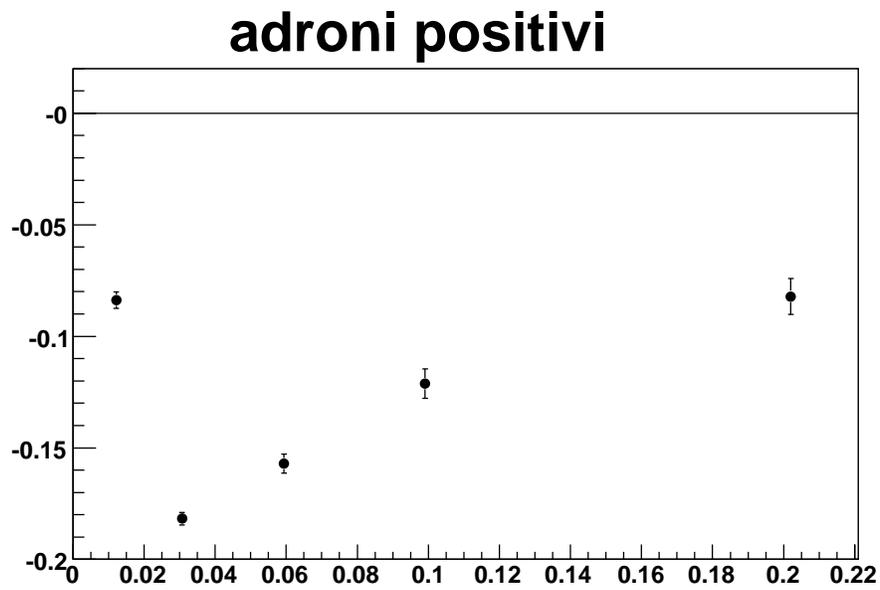
$$f(\phi) = A \cdot (1 + A_{\cos\phi} \cdot \cos(\phi) + A_{\cos 2\phi} \cdot \cos(2 \cdot \phi) + A_{\sin\phi} \cdot \sin(\phi)). \quad (5.7)$$

Nella procedura di Fit, il programma per l'analisi scritto utilizzando le classi di ROOT, utilizza MINUIT che chiama il metodo MIGRAD per trovare il minimo del  $\chi^2$  associato alla funzione calcolandone il gradiente.

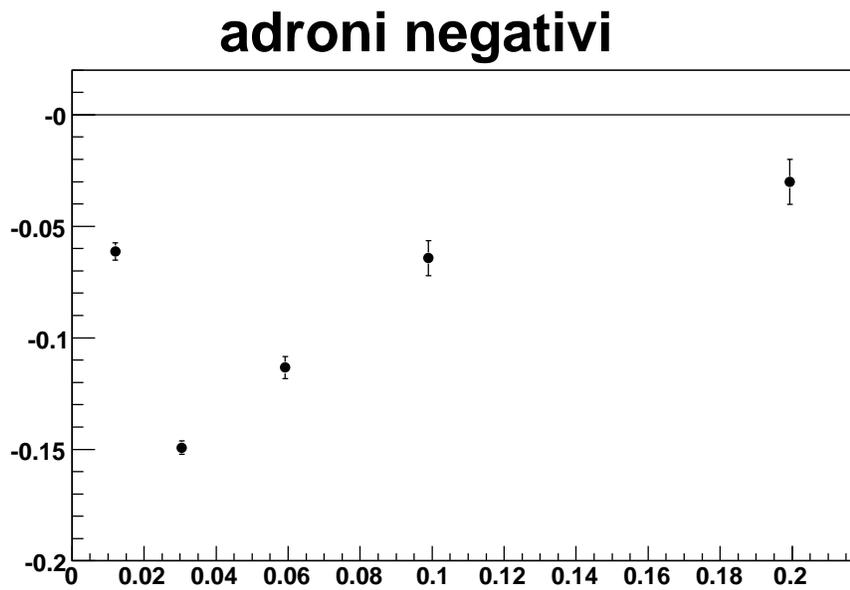
Nella tabella 5.4 vengono riportati i valori delle asimmetrie valutate per ogni bin di  $x$  con i relativi valori del  $\chi^2$  diviso per i 12 (16 bin - 4 parametri liberi) gradi di libertà del fit. Le curve nelle figure 5.2.3 e 5.2.3 sono le funzioni  $f(\phi)$  calcolate usando i valori stimati dei parametri per le asimmetrie. I grafici mostrati nelle figure 5.12, 5.13 e 5.14, mostrano i valori delle asimmetrie ottenute in funzione dei valori medi di  $x$ ; gli errori sono gli errori statistici del fit.

In figura 5.12 sono mostrati i valori ottenuti per  $A_{\cos\phi}$  in funzione di  $x$  separatamente per adroni positivi e negativi. Questa asimmetria è legata all'effetto Cahn descritto nel capitolo 1. I valori misurati raggiungono  $\sim -0.18$  e non sono compatibili con zero su tutto l'intervallo di  $x$  studiato, in accordo con gli effetti cinematici dovuti al momento trasverso intrinseco dei quark. Si nota una forte dipendenza da

$x$  e l'andamento a grande  $x$  è spiegabile con la correlazione  $x-Q^2$  e l'andamento  $1/Q$  dell'asimmetria di Cahn. Inoltre è presente una leggera differenza tra adroni positivi e negativi.



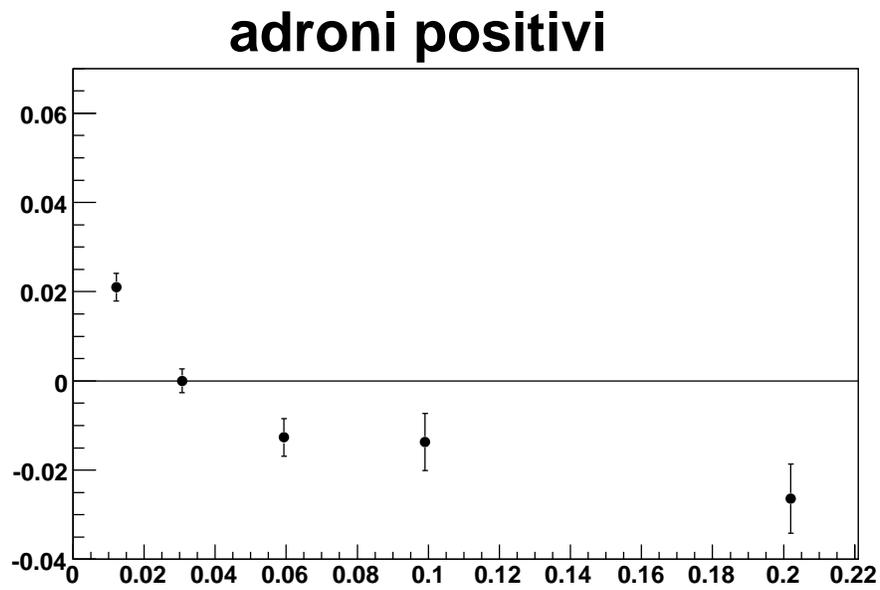
a)



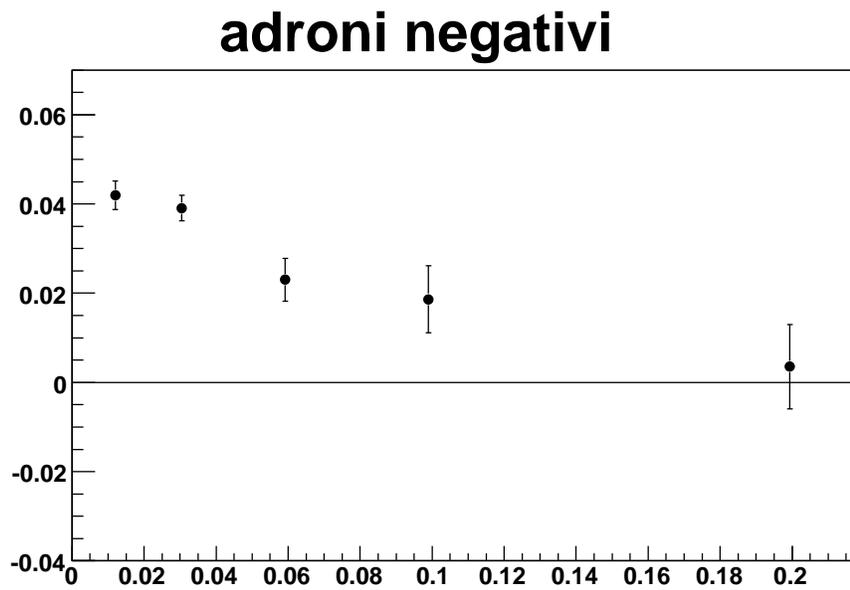
b)

Figura 5.12: Valori misurati di  $A_{\cos\phi}$  in funzione di  $x$  nel caso di adroni positivi, figura a), e di adroni negativi, figura b).

In figura 5.13 sono mostrati i valori ottenuti per  $A_{\cos 2\phi}$  in funzione di  $x$  separatamente per adroni positivi e negativi. A tale asimmetria contribuisce, all'ordine  $1/Q$ , la distribuzione partonica di Boer-Mulders accoppiata alla funzione di frammentazione di Collins. I valori assoluti di questa asimmetria (pochi punti percentuale) sono molto minori di quelli ottenuti per  $A_{\cos\phi}$ , e si osserva una forte dipendenza da  $x$ . Le differenze trovate tra adroni positivi e negativi necessita di ulteriori studi per determinare se sia da ascrivere alla fisica oppure alla sistematica.



a)



b)

Figura 5.13: Valori misurati di  $A_{\cos 2\phi}$  in funzione di  $x$  nel caso di adroni positivi, figura a), e di adroni negativi, figura b).

Il terzo parametro valutato,  $A_{\sin\phi}$ , descrive l'entità dell'asimmetria in  $\sin(\phi)$  legata alla polarizzazione del fascio incidente ed è essenzialmente compatibile con zero sia per adroni positivi che negativi, come si vede in figura 5.14.

### 5.3.2 Stabilità dei risultati

#### Andamento di $A_{\cos\phi}$ a piccoli $x$

L'andamento a piccolo  $x$  di  $A_{\cos\phi}$  è stato studiato stimando le asimmetrie in intervalli di  $x$  leggermente diversi: la regione  $0.002 < x < 0.080$  è stata suddivisa in 4 intervalli anziché 3. I valori di  $A_{\cos\phi}$  così ottenuti sono riportati in figura 5.15 per adroni positivi e negativi (cerchi) dove sono confrontati con i risultati precedenti (triangoli). I nuovi valori sono coerenti con i precedenti e suggeriscono la presenza di un massimo (in valore assoluto) a  $x \simeq 0.04$  ed una diminuzione ad  $x$  minori. Non si tratta quindi di una peculiarità del primo punto misurato, ma di un andamento continuo. Prima di trarre conclusioni sulla fisica, è comunque necessario ricordare che l'accordo tra Monte Carlo e dati non è particolarmente soddisfacente proprio a piccoli  $x$ , si veda figura 3.13 al capitolo 3, e ulteriori studi sono indispensabili.

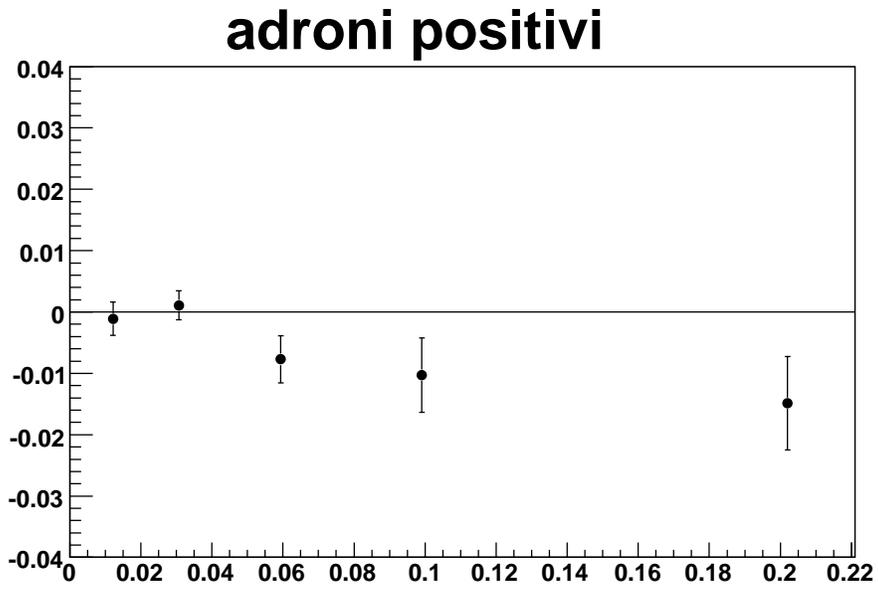
#### Effetto della selezione $\theta_h^{lab} > 0.020$ rad

Nello studio del confronto tra le distribuzioni delle variabili cinematiche ottenute con i dati e la simulazione Monte Carlo, descritto nel capitolo 3, si è cercato di individuare e risolvere eventuali disaccordi tra simulazione ed apparato per ridurre al minimo la presenza di errori sistematici dovuti alla cattiva simulazione di quest'ultimo. Nel seguito verranno presentate le distribuzioni angolari e le asimmetrie calcolate senza applicare il taglio angolare  $\theta_h^{lab} > 0.020$  rad agli adroni ricostruiti. Dal confronto con i grafici per i quali tale taglio è stato applicato, si può valutare l'entità di questo tipo di errore: nella zona  $\theta_h^{lab} < 0.020$  rad ci sono i disaccordi più evidenti per le distribuzioni dell'energia e dell'angolo dell'adrone, come mostrato nel paragrafo 3.4.1 del capitolo 3.

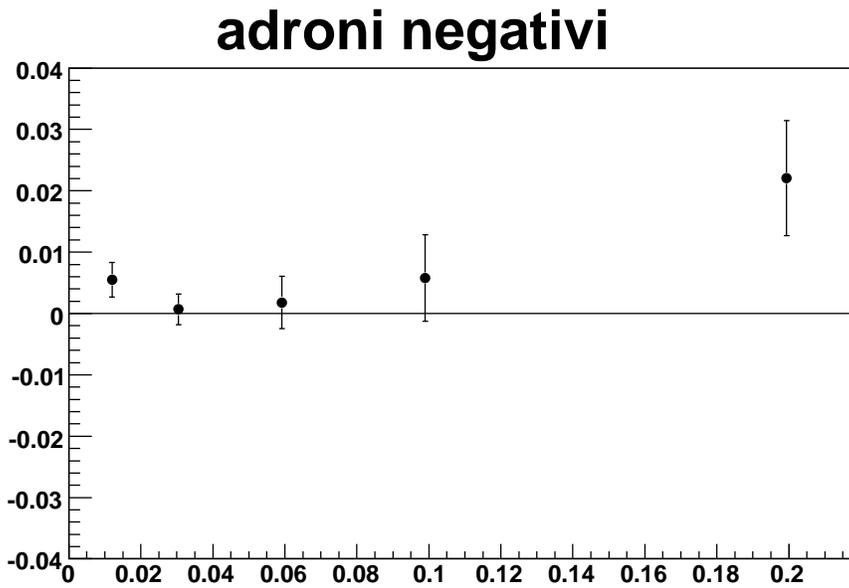
Questa selezione modifica notevolmente le distribuzioni dell'angolo azimutale degli adroni, ed è responsabile di gran parte delle variazioni dell'accettanza a valori di  $x$  minori, come si può vedere dalla figura 5.16, nella quale le distribuzioni per dati reali e simulati per adroni con  $\theta_h^{lab} > 0.020$  rad (sinistra) sono confrontate con

quelle ottenute senza il taglio (destra). In quest'ultimo caso è presente un picco nella distribuzione dell'angolo azimutale  $\phi$  dell'adrone ricostruito a partire dai dati reali, assente in quella ricostruita dai dati simulati, in corrispondenza alla regione attorno a  $\phi = 0$ . Tale picco, la cui origine non è stata ancora chiarita e necessita di ulteriori studi, non compare se si confrontano le distribuzioni azimutali per gli adroni con angolo polare  $\theta_h^{lab} > 0.020$  rad e potrebbe influire sulla misura delle asimmetrie. Per valutare le dipendenze da queste selezioni dei valori misurati delle asimmetrie, l'intera analisi è stata ripetuta utilizzando il campione di adroni senza il taglio angolare.

Nella figura 5.17 sono mostrati i risultati per le asimmetrie riportate in tabella 5.1 sovrapposti ai risultati ottenuti includendo anche gli adroni con angolo polare  $\theta_h^{lab} < 0.020$  rad nell'analisi. I valori stimati dell'asimmetria in  $\cos(\phi)$  risultano abbastanza stabili rispetto alla presenza o meno del taglio angolare, mentre nel caso dell'asimmetria in  $\cos(2\phi)$ , benché l'andamento delle asimmetrie in funzione di  $x$  sia simile in entrambi i casi, si osserva una traslazione lungo l'asse delle ordinate, sia per adroni positivi che negativi e la differenza tra le due analisi è dello stesso ordine di grandezza dell'effetto misurato. Infine, per  $A_{\sin\phi}$ , i risultati sono essenzialmente gli stessi. Il problema della stabilità dei risultati rispetto al taglio necessita di effettuare ulteriori studi.

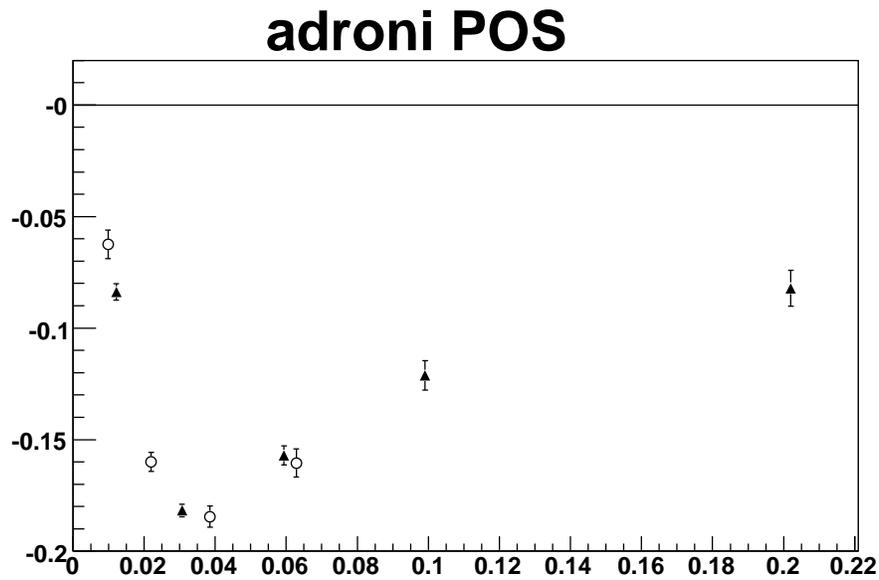


a)

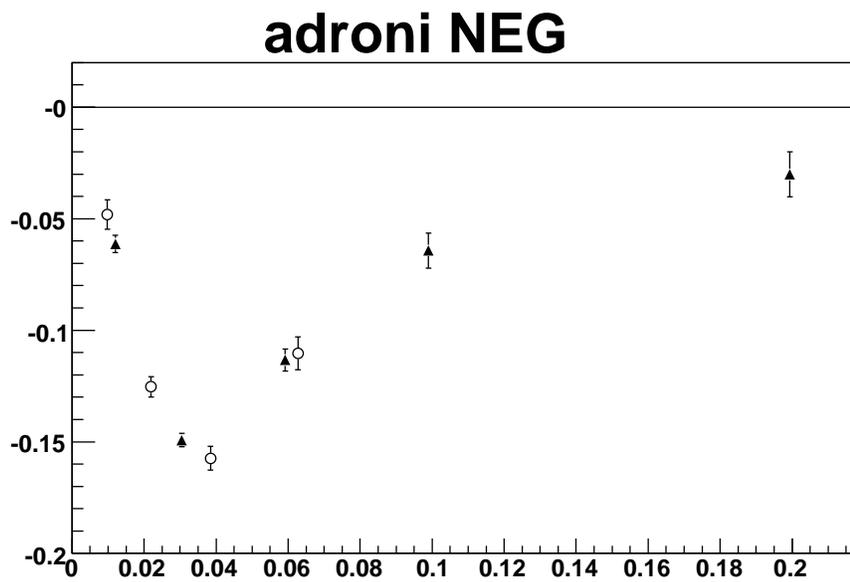


b)

Figura 5.14: Valori misurati di  $A_{\sin\phi}$  in funzione di  $x$ , per adroni positivi, figura a), e per adroni negativi, figura b).



a)



b)

Figura 5.15: Confronto a piccoli  $x$  per i valori misurati di  $A_{\cos\phi}$  in funzione di  $x$ , per adroni positivi, figura a), e per adroni negativi, figura b). I triangoli indicano i valori ottenuti dai risultati precedenti, mentre i cerchi indicano i valori ottenuti suddividendo il range  $0.002 < x < 0.080$  in 4 bin anziché 3.

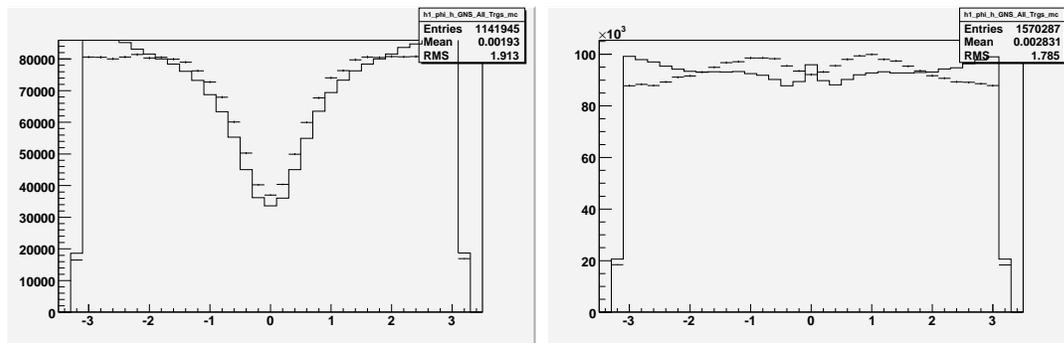


Figura 5.16: Distribuzione dell'angolo azimutale  $\phi$  dell'adrone ricostruito. A sinistra sono sovrapposte le distribuzioni ottenute dai dati del sottoperiodo W35 (linea continua) e dagli eventi simulati (punti), per adroni con angolo polare  $\theta_h^{lab} > 0.020$  rad; a destra è mostrata la sovrapposizione delle analoghe distribuzioni includendo anche gli adroni con angolo polare  $\theta_h^{lab} < 0.020$  rad, zona in cui la riproduzione dell'apparato da parte della simulazione Monte Carlo non è altrettanto buona.

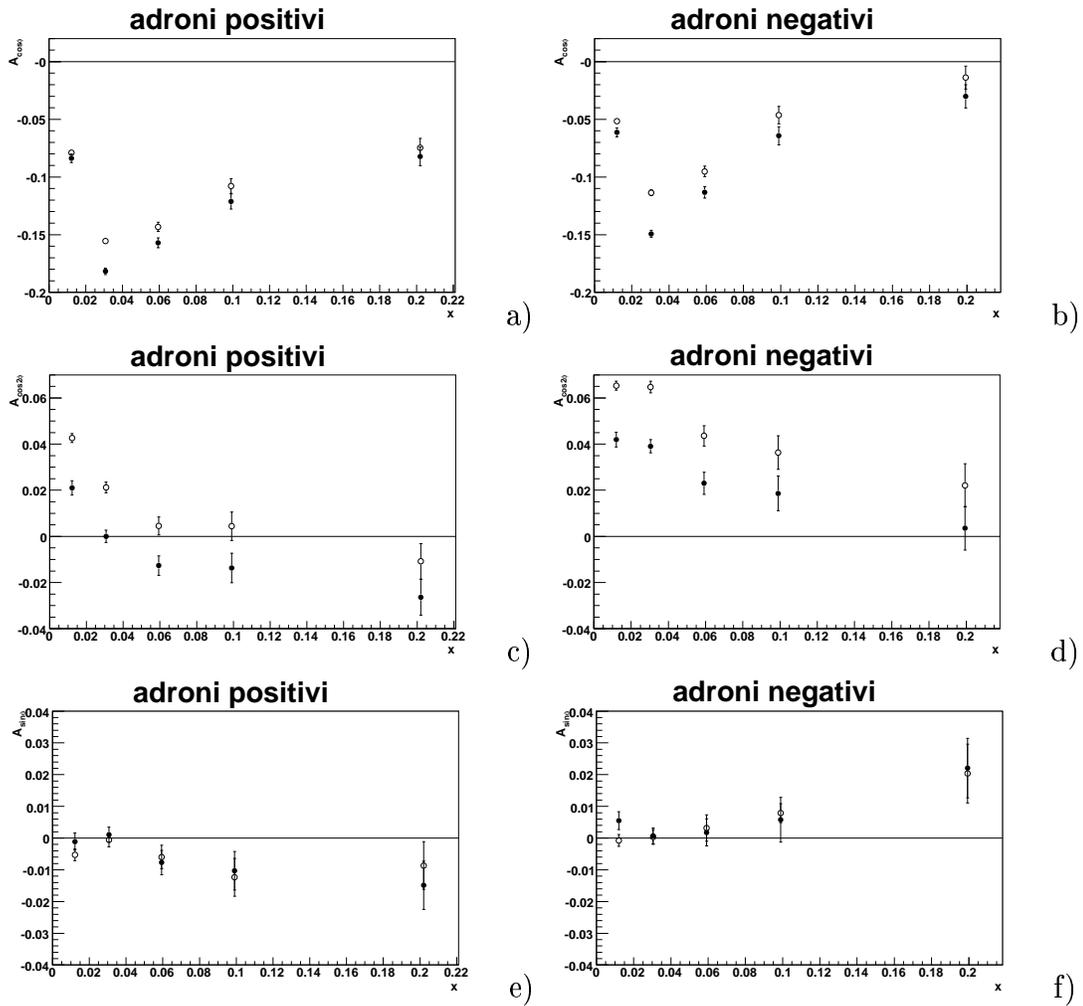


Figura 5.17: Asimmetrie azimutali in funzione di  $x$ . In alto  $A_{\cos\phi}$ , per adroni positivi a) e negativi b); al centro  $A_{\cos 2\phi}$ , per adroni positivi c) e negativi d); in basso  $A_{\sin\phi}$ , per adroni positivi e) e negativi f). Nelle figure sono sovrapposti i valori calcolati dall'analisi degli adroni con  $\theta_h^{lab} > 0.020$  rad, rappresentati con cerchi neri, a quelli calcolati inserendo nell'analisi anche gli adroni a piccolo angolo,  $\theta_h^{lab} < 0.020$  rad, rappresentati con cerchi bianchi. Si vede che i risultati non dipendono molto dal taglio effettuato, a parte per l'asimmetria in  $\cos(2\phi)$ , dove l'andamento è simile, ma i valori dell'asimmetria appaiono traslati tutti di una quantità analoga lungo l'asse delle ordinate; ciò vale sia per adroni positivi che negativi. L'origine di tale comportamento richiede ulteriori studi.

## 5.4 Confronto con misure precedenti e previsioni teoriche

Nonostante le considerazioni del paragrafo precedente, che sottolineano il carattere preliminare dei risultati presentati in questa tesi, è interessante confrontare le misure ottenute con quelle precedenti eseguite dall'esperimento EMC e con alcune curve teoriche che danno le asimmetrie previste nel dominio cinematico di COMPASS e per bersaglio costituito da deuterio.

### Confronto con l'esperimento EMC

EMC (Europeana Muon Collaboration) fu un esperimento a bersaglio fisso di prima generazione per misure di DIS al CERN, installato sulla stessa linea di fascio usata da COMPASS. Ha eseguito molte misure importanti (la presa dati fu completata nel 1982) e fu tra i primi a misurare l'asimmetria azimutale [10]. In figura 5.18 viene mostrato il confronto per le asimmetrie in  $\cos(\phi)$  ed in  $\cos(2\phi)$ , divise per i fattori cinematici  $\varepsilon_1$  e  $\varepsilon_2$ , rispettivamente, definiti nel paragrafo 1.4.2 del capitolo 1, con quelle misurate dall'esperimento EMC per  $W^2$  appartenente all'intervallo [60 – 160]  $(\text{GeV}/c^2)^2$ <sup>1</sup>, dove viene incluso anche l'errore sistematico. Le asimmetrie così ottenute da EMC sono in funzione di  $Q^2$  e quindi anche le asimmetrie ottenute per COMPASS sono state mostrate in funzione del valore di  $Q^2$ .

Dalla figura risulta evidente la grande potenzialità di COMPASS nel fornire misure statisticamente di gran lunga più precise degli esperimenti precedenti.

### Previsioni per l'asimmetria in $\cos(\phi)$

In [64] è calcolata la sezione d'urto SIDIS non polarizzata nell'ambito del modello a partoni, assumendo che la dipendenza delle PDF e delle FF non polarizzate dal momento trasverso sia uguale a quella scritta nelle equazioni 1.51 del paragrafo 1.4.2 del capitolo 1.

L'assunzione fatta per le funzioni di distribuzione e frammentazione è quindi che la dipendenza dai momenti trasversi sia fattorizzabile rispetto a quella delle altre variabili cinematiche dell'urto,  $x$  o  $z$ , e di tipo gaussiano. Inoltre i parametri  $\langle k_T^2 \rangle$  e

---

<sup>1</sup>Nei dati analizzati il valore medio di  $W^2$  è di  $\sim 104 (\text{GeV}/c^2)^2$

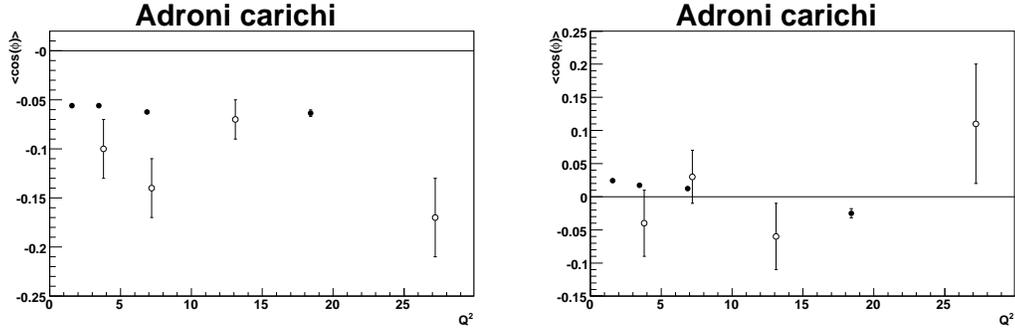


Figura 5.18: Confronto con le asimmetrie misurate da EMC (cerchi bianchi). A sinistra è mostrato il confronto per  $\frac{A_{\cos\phi}}{2 \cdot \varepsilon_1}$  in funzione di  $Q^2$ , a destra è mostrato quello per  $\frac{A_{\cos 2\phi}}{2 \cdot \varepsilon_2}$ .

$\langle p_T^2 \rangle$ , sono indipendenti sia dal sapore del quark colpito sia dalla specie dell'adrone prodotto. In particolare  $\langle p_T^2 \rangle$  è lo stesso per FF favorite e sfavorite.

Inserendo queste parametrizzazioni nella sezione d'urto, all'ordine zero di QCD ed al primo ordine in  $k_T/Q$  e assumendo che il momento trasverso dell'adrone sia  $\mathbf{P}_T = z\mathbf{k}_T + \mathbf{p}_T$ , si ha:

$$\begin{aligned} \frac{d^5\sigma^{\ell p \rightarrow \ell h X}}{dx dQ^2 dz d^2\mathbf{P}_T} \simeq \sum_q \frac{2\pi\alpha^2 e_q^2}{Q^4} f_q(x) D_q^h(z) \left[ 1 + (1-y)^2 \right. \\ \left. - 4 \frac{(2-y)\sqrt{1-y} \langle k_T^2 \rangle z P_T}{\langle P_T^2 \rangle Q} \cos \phi_h \right] \frac{1}{\pi \langle P_T^2 \rangle} e^{-P_T^2 / \langle P_T^2 \rangle}. \end{aligned} \quad (5.8)$$

Nell'articolo viene calcolato il valore medio del coseno dell'angolo azimutale ( $\langle \cos(\phi) \rangle = A_{\cos\phi}/2$ ) integrato sul dominio cinematico di COMPASS. I valori numerici di  $\langle k_T^2 \rangle = 0.25(\text{GeV}/c)^2$  e  $\langle p_T^2 \rangle = 0.20(\text{GeV}/c)^2$  usati per la predizione sono quelli calcolati in [13], ottenuti con un fit globale dei dati dell'esperimento EMC, svolto al CERN, e E665, svolto al FNAL, assumendo le stesse ipotesi sulla forma esplicita delle PDF e FF sopra descritte.

L'asimmetria prevista è mostrata in figura 5.19, dove viene sovrapposta ai valori di  $A_{\cos\phi}/2$  di tabella 5.4. L'andamento dei dati risulta essere in buon accordo con le previsioni.

La dipendenza del valore dell'asimmetria dal momento trasverso medio dei quark  $\langle k_T^2 \rangle$  può essere estratta dalla sezione d'urto scritta in 5.8 integrata in  $\mathbf{P}_T$  e nella regione cinematica di COMPASS.

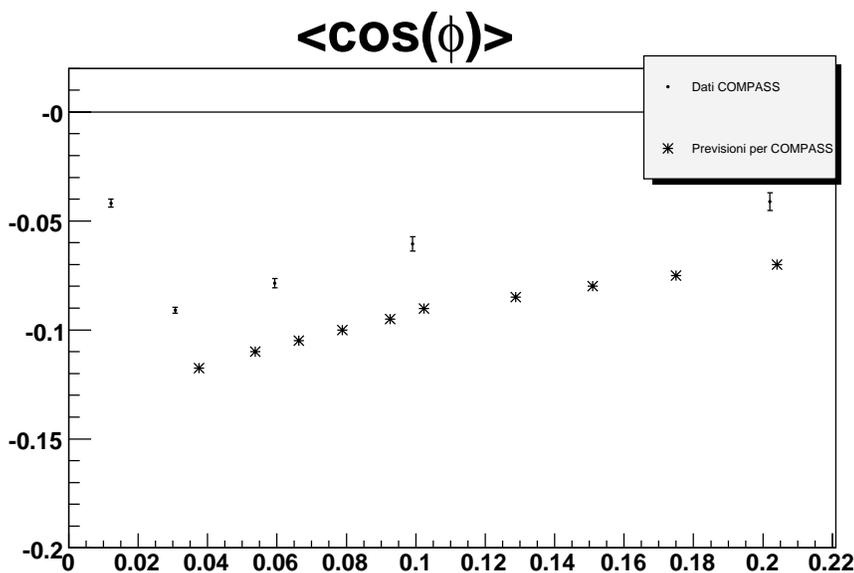


Figura 5.19: Asimmetria azimutale in  $\cos(\phi)$ : confronto tra i risultati ottenuti dal lavoro di tesi (punti) e quelli calcolati in [64] (asterischi) per la regione cinematica coperta da COMPASS.

Risulta che:

$$\langle \cos(\phi) \rangle \sim K_1 \cdot \left( \sqrt{\langle k_T^2 \rangle} + K_2 \right), \quad (5.9)$$

dove  $K_1$  è costante rispetto a  $\sqrt{\langle k_T^2 \rangle}$ , mentre  $K_2$  dipende debolmente dal rapporto  $\frac{\langle p_T^2 \rangle}{\langle k_T^2 \rangle}$ . Quindi il confronto tra l'asimmetria misurata e la predizione teorica, mostrato in figura 5.19, potrebbe indicare un valore di  $\sqrt{\langle k_T^2 \rangle}$  minore di quello usato in [64].

Va tenuto presente, comunque, che i valori ottenuti sono ancora preliminari e che ulteriori studi sistematici sono necessari prima di trarre alcuna conclusione.

**Previsioni per l'asimmetria in  $\cos(2\phi)$** 

L'interpretazione in questo caso è più delicata in quanto l'asimmetria in  $\cos(2 \cdot \phi)$  non deriva solo da un effetto cinematico, effetto Cahn, ma dipende anche dalla convoluzione tra la funzione di distribuzione di Boer-Mulders, non nota, e della funzione di frammentazione di Collins, misurata recentemente in reazioni  $e^+e^-$ . Inoltre, per alti valori del momento trasverso dell'adrone  $P_T$ , è rilevante anche il contributo calcolato nell'ambito della QCD perturbativa.

In[14] è presentata una prima previsione dell'asimmetria in  $\cos(2\phi)$ ; calcoli più dettagliati sono stati fatti per adroni positivi [65]. Le previsioni sono confrontate con i valori misurati riportati in tabella 5.4 e riportate in figura 5.20 dove sono disegnate le curve per i diversi effetti previsti: la linea punteggiata (in verde) rappresenta il contributo originato dalle funzioni di B-M e di Collins; la linea tratteggiata (in blu) rappresenta il contributo di origine cinematica dell'effetto Cahn; la linea punteggiata-tratteggiata (in nero) rappresenta il contributo derivato dai calcoli di QCD perturbativa.

Data l'instabilità nei valori misurati di  $A_{\cos 2\phi}$ , non si possono trarre conclusioni prima di ulteriori studi su effetti sistematici. È comunque interessante notare che l'ordine di grandezza dell'asimmetria misurata e prevista è, anche in questo caso, lo stesso. È inoltre evidente la potenzialità di COMPASS nella misura di questa asimmetria, di grande interesse per le informazioni che può dare sulla funzione TMD di Boer-Mulders.

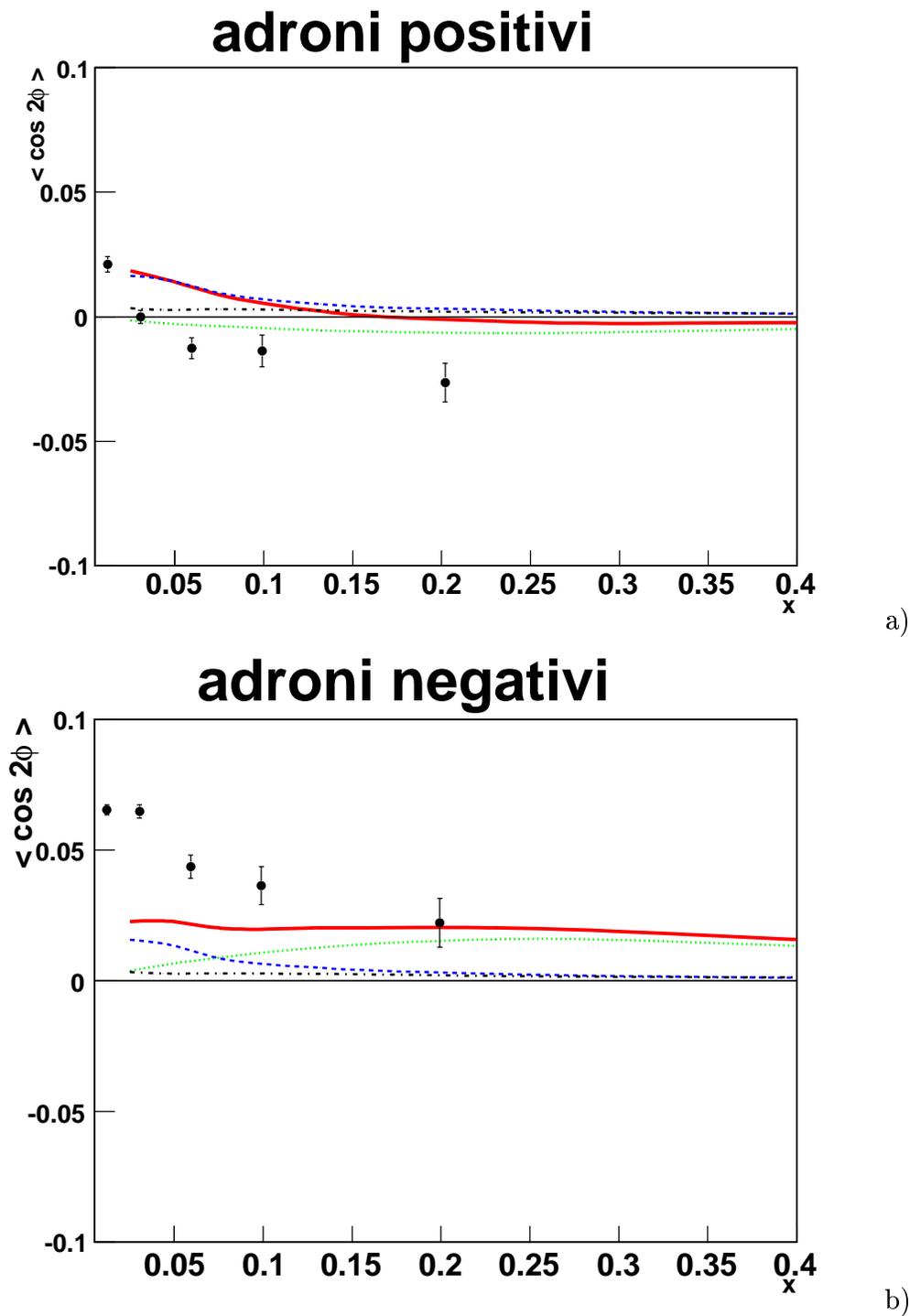


Figura 5.20: Confronto dei dati ottenuti per  $\langle \cos(2\phi) \rangle$  (punti) con la predizione calcolata da Barone e Prokudin [65] (linee colorate) dove la linea continua indica la somma dei diversi contributi previsti. Nella figura a) è riportato il confronto per adroni positivi, nella figura b) per adroni negativi.

# Conclusioni

I risultati ottenuti per le asimmetrie nella distribuzione azimutale degli adroni prodotti nel DIS non polarizzato a COMPASS sono molto interessanti e incoraggianti. Le difficoltà della misura, legate al fatto che, contrariamente al caso delle asimmetrie di spin, le distribuzioni angolari misurate debbano essere corrette per l'accettanza dell'apparato sperimentale, sono superabili. In particolare è stato superato il problema tecnico della produzione dell'ingente quantità di eventi simulati necessaria per stimare l'accettanza, disegnando, costruendo, testando e utilizzando un sistema distribuito per la simulazione, in grado di sfruttare sia le risorse di calcolo locali sia quelle accessibili tramite Grid. Ulteriori sviluppi permetteranno di automatizzarne le operazioni e aumentarne l'efficienza. I valori delle asimmetrie ottenuti in questo lavoro sono in ragionevole accordo con misure precedenti e con le più recenti previsioni teoriche, nonostante le correzioni applicate per l'accettanza e l'efficienza dell'apparato sperimentale siano tutt'altro che trascurabili. I valori ottenuti presentano però una dipendenza dai tagli eseguiti nell'analisi che è maggiore degli errori statistici, indicando quindi la presenza di possibili effetti sistematici. Per finalizzare i risultati, è quindi necessario eseguire ulteriori studi e, soprattutto, migliorare l'accordo tra le distribuzioni ottenute con eventi reali e simulati. Ciò implica ulteriori produzioni di eventi simulati, cambiando le condizioni della generazione. Inoltre, sono disponibili molti più dati di quelli analizzati in questa tesi (circa un fattore 10), raccolti in COMPASS dal 2002 al 2006, in condizioni sperimentali diverse. Sarà estremamente utile confrontare i risultati per le asimmetrie ottenute nei diversi periodi di presa dati, per meglio stimare possibili errori sistematici. Si tratta di un lavoro lungo, ma le potenzialità di COMPASS in questo campo e l'interesse crescente per le asimmetrie azimutali nel SIDIS e gli effetti legati al momento trasverso intrinseco ed allo spin dei quark lo giustificano.



# Bibliografia

- [1] J. D. Bjorken, “Applications Of The Chiral  $U(6) \times (6)$  Algebra Of Current Densities,” *Phys. Rev.* **148**, 1467 (1966).
- [2] J. R. Ellis and R. L. Jaffe, “A Sum Rule For Deep Inelastic Electroproduction From Polarized Protons,” *Phys. Rev. D* **9**, 1444 (1974) [Erratum-*ibid.* *D* **10**, 1669 (1974)].
- [3] R.L. Jaffe e A. Manohar, *Nucl. Phys.* B337,509 (1990)
- [4] K. Ackerstaff et al., *Phys. Rev. D*48, 5057 (1993)
- [5] B. Povh, K. Rith, C. Scholz, F. Zetsche, “Particles and nuclei: an introduction to the physical concepts ”, Berlin: Springer, 1995.
- [6] F. Halzen, A. D. Martin, “Quarks and leptons: an introductory course in modern particle physics”, NewYork, NY: Wiley, 1984.
- [7] S. Kretzer, “Fragmentation functions from flavour-inclusive and flavour-tagged  $e^+ e^-$  annihilations,” *Phys. Rev. D* **62**, 054001 (2000) [arXiv:hep-ph/0003177].
- [8] M. Gluck, E. Reya and A. Vogt, “Dynamical parton distributions revisited,” *Eur. Phys. J. C* **5**, 461 (1998) [arXiv:hep-ph/9806404].
- [9] U. D’Alesio and F. Murgia, “Parton intrinsic motion in inclusive particle production: Unpolarized cross sections, single spin asymmetries and the Sivers effect,” *Phys. Rev. D* **70**, 074009 (2004) [arXiv:hep-ph/0408092].
- [10] J. J. Aubert *et al.* [European Muon Collaboration], “Measurement of hadronic azimuthal distributions in deep inelastic muon proton scattering,” *Phys. Lett. B* **130**, 118 (1983).

- [11] A. Konig and P. Kroll, “A Realistic Calculation Of The Azimuthal Asymmetry In Semiinclusive Deep Inelastic Scattering,” *Z. Phys. C* **16**, 89 (1982).
- [12] R.N.Cahn, “Azimuthal Dependence In Leptoproduction: A Simple Parton Model Calculation” *Phys.Lett.*B78:269,1978
- [13] M.Anselmino, M.Boglione, U.D’Alesio, A.Kotzinian, F.Murgia, A.Prokudin “Role of Cahn and Sivers effects in deep inelastic scattering” *Phys.Rev.*D71,074006 (2005)
- [14] Vincenzo Barone, Zhun Lu, Bo-Qiang Ma, “On the  $\cos 2\phi$  asymmetry in unpolarized leptoproduction” *Phys.Lett.* B632 (2006) 277-281 [arXiv:hep-ph/0512145v1]
- [15] A.Bacchetta, M.Diehl, K.Goeke, A.Metz, P.Mulders, M.Schlegel “Semi-inclusive deep inelastic scattering at small transverse momentum” arXiv:hep-ph/0611265v1
- [16] X. Artru, “The transverse spin,” arXiv:hep-ph/0207309.
- [17] J. C. Collins, “Fragmentation of transversely polarized quarks probed in transverse momentum distributions,” *Nucl. Phys. B* **396**, 161 (1993) [arXiv:hep-ph/9208213].
- [18] **EMC** Collaboration, M. Arneodo *et al.*, “Measurement of hadron azimuthal distributions in deep inelastic muon proton scattering”, *Z. Phys.* **C34** (1987) 277.
- [19] **E665** Collaboration, M. R. Adams *et al.*, “Perturbative QCD effects observed in 490-GeV deep inelastic muon scattering”, *Phys. Rev.* **D48** (1993) 5057–5066.
- [20] **ZEUS** Collaboration, J. Breitweg *et al.*, “Measurement of azimuthal asymmetries in deep inelastic scattering”, *Phys. Lett.* **B481** (2000) 199–212, <http://xxx.lanl.gov/abs/hep-ex/0003017> hep-ex/0003017.
- [21] **ZEUS** Collaboration, S. Chekanov *et al.*, “Measurement of azimuthal asymmetries in neutral current deep inelastic scattering at HERA”, <http://xxx.lanl.gov/abs/hep-ex/0608053> hep-ex/0608053.

- [22] CLAS Collaboration, H. Avakian *et al.*, “Measurement of beam-spin asymmetries for deep inelastic  $\pi^+$  electroproduction”, *Phys. Rev.* **D69** (2004) 112004, <http://xxx.lanl.gov/abs/hep-ex/0301005> hep-ex/0301005.
- [23] A. Airapetian *et al.* [HERMES Collaboration], “Beam-spin asymmetries in the azimuthal distribution of pion electroproduction,” *Phys. Lett. B* **648**, 164 (2007) [arXiv:hep-ex/0612059].
- [24] I. Ludwig, “Azimuthal Asymmetries and the Cahn Effect at COMPASS” , Univ. Freiburg, July 2004 <http://wwwcompass.cern.ch/compass/publications>
- [25] A.Kotzinian “Including Cahn and Sivers effects into event generation” arXiv:hep-ph/0510359v1 (2005)
- [26] P. Abbon *et al.* [COMPASS Collaboration], “The COMPASS Experiment at CERN,” *Nucl. Instrum. Meth. A* **577**, 455 (2007) [arXiv:hep-ex/0703049].
- [27] N. W. Schellingerhout *et al.*, “Nucleon polarization in three-body models of polarized  ${}^6\text{Li}$ ”, *Phys. Rev. C* **48** (1993) 2714.
- [28] Y. Bedfer, S.Gerassimov, A. Korzenev, R. Windmolders, “COMPASS’s track reconstruction algorithm”, (2004); nota interna di COMPASS.
- [29] <http://root.cern.ch/>
- [30] <http://cbernet.home.cern.ch/cbernet/Cool/>
- [31] L. Schmitt *et al.*, “The DAQ of the COMPASS experiment”, *IEEE Tran. Nucl. Sci. Proceedings of IEEE Realtime 2003 Conference*.
- [32] <http://coral.web.cern.ch/coral/cvsrepository.php>
- [33] <http://ges.home.cern.ch/ges/phast/>
- [34] E.S.Ageev *et al.*, “A new measurement of the Collins and Sivers asymmetries on a transversely polarised deuteron target”, *Nuclear Physics B765* (2007) 31-70
- [35] <http://rkb.home.cern.ch/rkb/AN16pp/node177.html>
- [36] <http://www3.tsl.uu.se/thep/lepto>

- [37] <http://valexakh.web.cern.ch/valexakh/wwwcomg/index.html>
- [38] <http://wwwasd.web.cern.ch/wwwasd/geant/>
- [39] <http://www.ts.infn.it/servizi/computing.html>
- [40] <http://www.ts.infn.it/>
- [41] Andrew S. Tanenbaum e Maarten van Steen “Distributed Systems: Principles and Paradigms”, Prentice Hall, 2002
- [42] I. Foster e C. Kesselman, “The Grid: Blueprint of a New Computing Infrastructure”, Morgan Kaufman Publishers, 1998
- [43] I. Foster e C. Kesselman, “The Grid 2: Blueprint of a New Computing Infrastructure”, Morgan Kaufman Publishers, 2004
- [44] V. Welch, I. Foster, C. Kesselman, O. Mulmo, L. Pearlman, S. Tueck, J. Gawor, S. Meder e F. Siebenlist, “X.509 Proxy Certificates for Dynamic Delegation” Terzo Workshop annuale PKI R & D, Maryland, USA, Aprile 2004
- [45] <http://eu-datagrid.web.cern.ch/eu-datagrid/>
- [46] <http://lcg.web.cern.ch/LCG/>
- [47] <http://www.cs.wisc.edu/condor/>
- [48] <http://www.globus.org/>
- [49] <http://www.platform.com/Products/Platform.LSF.Family/Platform.LSF/>
- [50] <http://www.globus.org/wsrp/>
- [51] <http://www.oasis-open.org/home/index.php>
- [52] <http://www.globus.org/ogsa/>
- [53] GSI, <http://www.globus.org/security/overview.htm>
- [54] EGEE Middleware Architecture. Document identifier: EGEE-DJRA1.1-594698-v1.0

- [55] Grid Storage Managment Working Group (GSM-WG),  
<https://forge.gridforum.org/projects/gsm-wg/>
- [56] W.Allock et al. *GridFTP Protocol Specification*, Global Grid Forum  
Recomendation GFD.20, Marzo 2003
- [57] <http://cern.ch/AMGA>
- [58] B.Koblitz, N.Santos, V.Pose “The gLite AMGA metadata catalogue” CHEP06
- [59] <http://cern.ch/diane>
- [60] omniORB: Free High Performance ORB for C++ and python,  
<http://omniorb.sourceforge.net>
- [61] <http://ganga.web.cern.ch/ganga/>
- [62] P. J. Leach and R. Salz, UUIDs and GUIDs. Network Working Draft, February  
1998. <http://www.opengroup.org/dce/info/draft-leach-uuids-guids-01.txt>.
- [63] [www.ibm.com/systems/clusters/software/gpfs.html](http://www.ibm.com/systems/clusters/software/gpfs.html)
- [64] M. Anselmino, A. Efremov, A. Kotzinian and B. Parsamyan, “Transverse mo-  
mentum dependence of the quark helicity distributions and the Cahn effect in  
double-spin asymmetry  $A(LL)$  in semi inclusive DIS,” *Phys. Rev. D* **74**, 074015  
(2006) [arXiv:hep-ph/0608048].
- [65] E. Barone, Bo-Qiang Ma, A. Prokudin to be published (2007)



# Ringraziamenti

Il primo doveroso ringraziamento e' dedicato ad Anna Martin e Venicio Duic.

Grazie al professor Franco Bradamante ed a tutti gli altri membri del gruppo COMPASS-Trieste, in particolare al gruppo del trasverso.

Ringrazio tutti i colleghi dell'ufficio alte energie, in particolare Federica Sozzi ed Andrea Bressan.

Ringrazio la sezione INFN di Trieste per il supporto che mi ha dato in qualità di associato, consentendomi di seguire da vicino il run 2007.

Ringrazio Massimo Lamanna della sezione IT del CERN, per avermi dato l'opportunità di lavorare a stretto contatto con lui ed il suo team in un settore di punta della ricerca informatica. Grazie in particolare a J. Moscicki, A. Muraru, B. Koblitz.