



**Università del Salento**

---

FACOLTÀ DI SCIENZE MATEMATICHE FISICHE E NATURALI  
Corso di Laurea Triennale in Fisica

TESI DI LAUREA

**Studio di scenari oltre il Modello Standard  
all'esperimento ATLAS a LHC  
con tecniche di analisi multivariata**

Candidato:  
**Marta Dell'Atti**

Relatore:  
**Dott.ssa Margherita Primavera**

---

Anno Accademico 2010–2011



# Indice

<b>Introduzione</b>	<b>vii</b>
<b>1 Il Modello Standard e la Supersimmetria</b>	<b>1</b>
1.1 Il Modello Standard . . . . .	2
1.1.1 I costituenti della materia . . . . .	2
1.1.2 Le interazioni fondamentali . . . . .	4
1.1.3 Il meccanismo di Higgs . . . . .	5
1.1.4 I limiti del Modello Standard . . . . .	5
1.2 La Supersimmetria . . . . .	5
1.2.1 Il Modello Standard Supersimmetrico Minimale . . . . .	7
<b>2 L'esperienza ATLAS al collider LHC del CERN</b>	<b>9</b>
2.1 Il Large Hadron Collider . . . . .	9
2.2 L'esperienza ATLAS ad LHC . . . . .	12
<b>3 L'analisi multivariata</b>	<b>19</b>
3.1 Tecniche di analisi multivariata . . . . .	19
3.1.1 I classificatori . . . . .	19
3.1.2 Metodi di analisi multivariata . . . . .	21
3.2 TMVA . . . . .	21
3.2.1 Metodi per l'analisi multivariata inglobati in TMVA . . . . .	22
<b>4 Applicazione dell'analisi multivariata ad un processo di nuova fisica ad ATLAS</b>	<b>31</b>
4.1 Gli eventi di segnale e di fondo . . . . .	31
4.2 Le variabili di input . . . . .	32
4.3 Il Training . . . . .	33
4.4 L'Application . . . . .	35
4.5 Risultati ottenuti . . . . .	38
<b>Conclusioni</b>	<b>44</b>
<b>Riferimenti bibliografici</b>	<b>49</b>



*Al Campus*



# Introduzione

In Fisica delle Particelle, per riuscire a dare una risposta coerente ad alcuni dei quesiti irrisolti del Modello Standard si introduce la Supersimmetria, che nasce come un modello teorico, basato sulla simmetria tra bosoni e fermioni, che permette di estendere il Modello Standard. Seppure la Supersimmetria preveda eleganti modelli per descrivere la realtà è necessario disporre di evidenze sperimentali per avere la conferma della sua validità: anche a questo scopo è nato il Large Hadron Collider del CERN e gli esperimenti che vi prendono dati, tra cui ATLAS. Nell'esperimento ATLAS si utilizzano diverse tecniche di analisi dei dati tra le quali, recentemente, riscuote un certo successo l'analisi multivariata, volta a fronteggiare il problema della discriminazione del segnale e del fondo in condizioni estremamente critiche di eventi di segnale rarissimi su un fondo largamente predominante.

Questo lavoro di tesi si articola in: un primo capitolo in cui sono brevemente descritti il Modello Standard e la Supersimmetria; un secondo capitolo che concerne la descrizione l'esperimento ATLAS al Large Hadron Collider; un terzo capitolo in cui vengono presentate le tecniche di analisi multivariata e, infine, un quarto capitolo in cui è effettuato uno studio di fattibilità per l'applicazione dell'analisi multivariata al problema concreto di isolare il processo di nuova fisica di produzione diretta del partner supersimmetrico del quark top, identificato nei suoi decadimenti in top e neutralino, da un enorme fondo di processi noti descritti dal Modello Standard.



# Capitolo 1

## Il Modello Standard e la Supersimmetria

A partire dalla fine dell'800 la Fisica delle Particelle ha subito una notevole evoluzione, basata sulla ricerca volta a determinare i costituenti fondamentali della materia. In principio le sole particelle considerate erano l'elettrone e il protone, a cui si sono aggiunte moltissime altre particelle scoperte nel corso degli anni. All'inizio degli anni Sessanta il gran numero di particelle fino ad allora supposte elementari creò non poca confusione tra i fisici, portando inevitabilmente a chiedersi se queste fossero effettivamente tali. In questo contesto fu proposto il modello a quark, i costituenti degli adroni, ossia delle particelle che interagiscono fortemente, a cui seguì l'elaborazione del Modello Standard: un quadro completo che descrive la struttura della materia e le forze che la governano, in cui la simmetria assume un ruolo preponderante. Inizialmente il Modello Standard era formulato con i soli leptoni, mentre i quark furono introdotti successivamente, negli anni Settanta. Un gran numero di verifiche sperimentali confermano con grande precisione il Modello Standard. Tuttavia la teoria presenta lacune, tra le quali ci sono l'asimmetria fra materia e antimateria nell'universo, l'impossibilità di giustificare l'esistenza della materia oscura, la mancata unificazione delle interazioni fondamentali. Un'ipotesi è che il Modello Standard possa costituire un'approssimazione alle basse energie di una teoria più ampia, e che questa sia in grado di fornire una risposta esaustiva ai quesiti che rimangono senza risposta.

Un promettente candidato a estendere il Modello Standard è la Supersimmetria, che potrebbe fornire un'interpretazione della materia oscura, potrebbe risolvere il cosiddetto problema della gerarchia di massa, e potrebbe collocarsi all'interno della teoria di grande unificazione delle quattro interazioni fondamentali (GUT). La Supersimmetria descrive uno scenario che prevede l'introduzione di nuove particelle rispetto a quelle considerate nel Modello Standard, la cui ricerca viene condotta negli esperimenti agli acceleratori ad energie sufficientemente elevate attraverso lo studio degli eventi generati nelle collisioni. In particolare, l'esperimento ATLAS al Large Hadron Collider (LHC) del CERN di Ginevra ha raccolto dati pari ad una luminosità integrata di  $\approx 5 \text{ fb}^{-1}$  da collisioni pp a un'energia nel centro di massa di  $\sqrt{s} = 7 \text{ TeV}$ . L'analisi dei dati raccolti e l'eventuale verifica dell'esistenza delle particelle previste dalla Supersimmetria, permetterà

di compiere un notevole passo in avanti per la comprensione della Natura.

## 1.1 Il Modello Standard

Il Modello Standard (MS) è una teoria di campo quantistica che descrive esaurientemente i costituenti primi della materia e le loro interazioni, avvalendosi della QCD (*Quantum Chromo-Dynamics*) e della QED (*Quantum Electro-Dynamics*). Il MS è coerente con la meccanica quantistica e con la teoria della relatività speciale e stabilisce che ciascuna interazione tra campi di materia è regolata da una simmetria locale gauge.

La necessità di introdurre una teoria quantistica di campo risiede nell'inadeguatezza della meccanica quantistica ordinaria nel descrivere particelle relativistiche, considerando le particelle come indistruttibili. Infatti l'equivalenza einsteiniana fra massa e energia comporta che particelle dotate di energie sufficientemente elevate possono combinarsi in particelle massicce o decadere in particelle più leggere, in processi detti di *creazione* e *annichilazione*. Per descrivere questi fenomeni nella teoria quantistica si introduce la cosiddetta seconda quantizzazione. Un sistema di particelle identiche è descritto da una funzione d'onda simmetrica rispetto allo scambio di due particelle qualsiasi se le particelle sono *bosoni*, o antisimmetrica se sono *fermioni*. Il campo quantistico è un'entità presente in ogni punto dello spazio, il cui stato cambia a seconda della presenza o meno di particelle, che interagiscono con il campo scambiando quanti di campo. Al campo quantistico è associato un Hamiltoniano di campo, che obbedisce all'equazione di Schrödinger e a cui sono associati autostati interpretati come la presenza o l'assenza di una particella a una data energia. La funzione d'onda di un sistema di particelle diviene così la composizione di autostati dell'Hamiltoniano di campo, a cui si possono applicare gli operatori di *creazione* e *annichilazione* di particelle.

Nel MS esistono tre tipi di particelle elementari: i *leptoni*, i *quark* e i *bosoni vettoriali*, e sono descritte tre delle quattro interazioni fondamentali: l'interazione forte, l'interazione debole e l'interazione elettromagnetica e le ultime due sono unificate nell'interazione elettrodebole. L'interazione gravitazionale non viene descritta dalla teoria e i suoi effetti, alle energie considerate, sono del tutto trascurabili.

### 1.1.1 I costituenti della materia

I quark e i leptoni sono i costituenti della materia ordinaria e sono *fermioni*. I fermioni sono particelle caratterizzate dall'aver spin semi-intero, per le quali un sistema da esse composto è descritto da una funzione d'onda completamente antisimmetrica per lo scambio di due particelle e che obbediscono al principio di Pauli. In virtù di questo principio due fermioni non possono occupare lo stesso stato quantistico, descritto dagli stessi numeri quantici, e sistemi con un gran numero di fermioni sono descritti dalla statistica di *Fermi-Dirac*.

Esistono 6 diversi tipi di quark, che si distinguono per *sapore* (o *flavour*): up (u), down (d), charm (c), strange (s), top (t) e bottom (b). I quark sono raggruppati, secondo le loro masse, in tre *famiglie* o *generazioni*. I quark u e d costituiscono la prima famiglia, i quark c e s la seconda e i quark t e b la terza. I quark u, c, t sono dotati di carica elettrica pari a  $+2/3$  la carica dell'elettrone  $e$ , e i quark d, s, b hanno

**Tabella 1.1:** I costituenti della materia: quark

	Famiglia			Carica elettrica	Carica colore			Spin	Numero barionico
	1	2	3		<i>r</i>	<i>g</i>	<i>b</i>		
Quark	u	c	t	+2/3				+1/2	+1/3
	d	s	b	-1/3				+1/2	

**Tabella 1.2:** I costituenti della materia: leptoni

	Famiglia			Carica elettrica	Carica colore	Spin	Numero leptonic
	1	2	3				
Leptoni	e	$\mu$	$\tau$	-1	-	+1/2	+1
	$\nu_e$	$\nu_\mu$	$\nu_\tau$	0		+1/2	

carica elettrica  $-1/3$  di  $e$ . Oltre ad interagire attraverso l'interazione elettrodebole i quark subiscono l'interazione forte, e per questo sono caratterizzati anche da una carica di *colore*, di cui esistono tre tipi: *rosso*, *verde*, *blu* ( $r$ ,  $g$ ,  $b$ ). Le particelle libere in natura hanno carica di colore *neutra*, e l'impossibilità di osservare quark liberi a causa dell'interazione forte prende il nome di *confinamento*. I quark  $u$  e  $d$  sono caratterizzati dal numero quantico di *Isospin* forte, che si comporta come un momento angolare e la cui terza componente  $I_3$  vale  $+1/2$  per  $u$  e  $-1/2$  per  $d$ . Gli altri quark hanno valore di isospin nullo, ma si introducono altri numeri quantici da associare a tali particelle: il  $c$  ha *fascino*  $C=1$ , lo  $s$  ha *stranezza*  $S=1$ , il  $t$  ha *verità*  $T=1$  e il  $b$  ha *bellezza*  $B'=1$ . Il MS prevede che a ogni particella sia associata un'antiparticella, con stessa massa e stessa vita media di questa, ma che si distingue da essa per l'aver numeri quantici opposti (carica elettrica, di colore, ecc.). I quark si combinano per formare gli adroni, che si distinguono in due classi: i *barioni* e i *mesoni*. I barioni sono adroni costituiti da una qualsiasi combinazione di tre quark, sono dotati di spin semi-intero e sono caratterizzati da un nuovo numero quantico, il numero barionico  $B$ , che assume valore  $+1$  per i barioni,  $-1$  per gli antibarioni e  $0$  per tutte le altre particelle, e che dà luogo a una nuova legge di conservazione che governa i processi del Modello Standard. I mesoni sono costituiti da un quark e un antiquark e hanno spin intero[1]. Nella tabella 1.1 sono riassunte le caratteristiche dei diversi quark.

Fra i leptoni si annoverano sei particelle: l'elettrone ( $e$ ), il muone ( $\mu$ ), il tau ( $\tau$ ) e i corrispondenti neutrini  $\nu_e$ ,  $\nu_\mu$ ,  $\nu_\tau$ . Si individuano, come per i quark, tre famiglie, come descritto nella tabella 1.2 ognuna delle quali è costituita da un leptone della prima riga e dal neutrino corrispondente. I leptoni sono fermioni e non essendo soggetti all'interazione forte, hanno carica di colore nulla e possono esistere come particelle libere. I leptoni della prima riga hanno carica elettrica  $-e$ , mentre i neutrini hanno carica elettrica nulla, dunque interagiscono molto debolmente con la materia e sono identificabili solo in reazioni in cui interviene l'interazione debole. Ad ogni famiglia leptonica viene associato un nuovo numero quantico: il *numero leptonic*. Si ha che  $L_e = L_\mu = L_\tau = 1$  per i leptoni e  $-1$  per gli antileptoni[1].

Tabella 1.3: Interazioni considerate nel MS

Interazione	si accoppia a	Bosone vettore	Massa (GeV/c <sup>2</sup> )	Range (fm)	J <sup>P</sup>
Elettromagnetica	carica elettrica	$\gamma$	0	$\infty$	1 <sup>-</sup>
Forte	carica di colore	8 g	0	$\approx 1$	1 <sup>-</sup>
Debole	carica debole	W <sup>±</sup> , Z <sup>0</sup>	$\approx 10^2$	$\approx 10^{-3}$	1

### 1.1.2 Le interazioni fondamentali

I bosoni vettori assumono un ruolo di particolare importanza nella teoria del MS, perché costituiscono le particelle mediatrici delle interazioni. Sono particelle caratterizzate da valori di spin intero. Un sistema di bosoni è descritto da una funzione d'onda completamente simmetrica rispetto allo scambio di due particelle e il suo comportamento è descritto dalla statistica di *Bose-Einstein*.

I bosoni vettori sono: il fotone ( $\gamma$ ), gli 8 gluoni (g), W<sup>±</sup> e, rispetto allo scambio di due particelle qualsiasi allo scambio di due particelle qualsiasi Z<sup>0</sup>. I fotoni sono i mediatori dell'interazione elettromagnetica, mentre i gluoni sono responsabili dell'interazione forte e sono dotati di carica di colore, quindi a differenza dei fotoni interagiscono fra loro. Sia i fotoni che i gluoni hanno massa nulla. I bosoni W<sup>±</sup> e Z<sup>0</sup> sono i mediatori dell'interazione debole e a differenza delle altre due tipologie di mediatori, sono particelle massicce, con masse dell'ordine del centinaio di GeV. Analogamente a quanto avviene per i gluoni che hanno carica forte, essi posseggono carica debole e possono dunque interagire tra loro.

Il range d'azione dell'interazione considerata dipende dalla massa del bosone mediatore, così come teorizzato da Yukawa. Il principio di indeterminazione di Heisenberg, nella formulazione rispetto al tempo e all'energia<sup>1</sup>, permette di asserire che un bosone vettore dotato di massa  $m$  esiste per un intervallo di tempo  $\Delta t \leq \hbar/mc^2$ . Quindi, rispetto al range dell'interazione si ha che  $\Delta R \leq \hbar/mc$ . L'interazione elettromagnetica, essendo mediata da una particella a massa nulla, il fotone, ha raggio d'azione infinito. L'interazione forte è mediata dai gluoni i quali, seppure abbiano massa nulla, hanno carica di colore e interagiscono, rendendola un'interazione a corto range. L'interazione debole, essendo mediata da particelle dotate di masse non nulle, ha raggio d'azione limitato, di circa  $10^{-3}$  fm. L'interazione debole è responsabile della violazione della parità e dei cambi di sapore dei quark[1].

Nella tabella 1.3 sono descritte alcune delle proprietà delle interazioni descritte dal Modello Standard.

A ciascuna interazione è associata una costante di accoppiamento, che varia rispetto alle energie cui ci si riferisce: la più intensa è quella dell'interazione forte, seguita da quella elettromagnetica e quindi dalla debole.

<sup>1</sup>Il principio di Heisenberg formulato rispetto al tempo e all'energia è:

$$\Delta t \Delta E \leq \hbar$$

dove in questo caso  $E = mc^2$ .

### 1.1.3 Il meccanismo di Higgs

Nel MS la presenza di  $W^\pm$  e  $Z^0$  massicci porta a un'incoerenza della teoria, che prevede, affinché sia preservata l'invarianza di gauge su cui è basata, la presenza di bosoni intermedi privi di massa. L'interazione elettromagnetica e l'interazione debole, infatti, vengono unificate nell'unica interazione elettrodebole e sono mediate da bosoni vettori a massa nulla. Il meccanismo mediante il quale i  $W^\pm$  e  $Z^0$  acquisiscono massa nel Modello Standard è il *meccanismo di Higgs*. Si introduce nel modello un campo quantistico aggiuntivo, chiamato *campo di Higgs*, che consiste in un doppietto di campi scalari complessi:

$$\varphi = \begin{pmatrix} \varphi^+ \\ \varphi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \varphi_1 + i\varphi_2 \\ \varphi_3 + i\varphi_4 \end{pmatrix}$$

I  $\varphi_i$  sono quattro campi scalari, che, dopo la rottura spontanea della simmetria, si accoppiano allo  $Z^0$  e ai  $W^\pm$ , generando nel modello i termini di massa di queste particelle. Il fotone rimane, invece, privo di massa. L'accoppiamento del campo di Higgs con i fermioni spiega invece, senza predirle quantitativamente, le loro masse. Conseguenza del fatto che sono stati introdotti quattro campi scalari per dare massa a tre bosoni nel modello è l'esistenza di un bosone di Higgs libero, la cui ricerca è in corso da anni e che ha fortemente motivato la costruzione dell'LHC e di alcuni degli esperimenti a questo acceleratore.

La prova dell'esistenza del bosone di Higgs è fondamentale per la validità del Modello Standard.

### 1.1.4 I limiti del Modello Standard

Nel corso degli anni sono state numerose le conferme sperimentali delle predizioni del MS. Tuttavia, nell'ambito del modello, ci sono problemi rimasti insoluti e alcuni aspetti che lo rendono non completamente soddisfacente[2]:

- nel MS non viene considerata l'interazione gravitazionale e le uniche interazioni unificate sono quella elettromagnetica e quella debole;
- il MS riesce a giustificare solo una piccola parte della *materia oscura*, composta di neutrini e che prende il nome di *materia oscura calda*, mentre non prevede l'esistenza di una particella massiccia e stabile in grado di interagire molto debolmente con la materia ordinaria che costituisca la *materia oscura fredda*;
- non permette di formulare una spiegazione per l'asimmetria tra materia e antimateria che si osserva nell'universo;
- il MS non spiega perché esistano proprio tre famiglie di quark e di leptoni.

## 1.2 La Supersimmetria

La Supersimmetria (SUSY), elaborata negli anni Settanta, costituisce un valido candidato ad estendere il MS. Diversi aspetti rendono la teoria promettente ad

Tabella 1.4: Partner supersimmetrici delle particelle del MS

Particelle	Simbolo	Spin	Sparticelle	Simbolo	Spin
elettrone, muone, tau	$e, \mu, \tau$	1/2	selettrone, smuone, stau	$\tilde{e}, \tilde{\mu}, \tilde{\tau}$	0
quark	$q$	1/2	squark	$\tilde{q}$	0
neutrino	$\nu$	1/2	sneutrino	$\tilde{\nu}$	0
fotone	$\gamma$	1	fotino	$\tilde{\gamma}$	1/2
gluone	$g$	1	gluino	$\tilde{g}$	1/2
bosone $W^\pm$	$W^\pm$	1	Wino	$\tilde{W}^\pm$	1/2
bosone $Z^0$	$Z^0$	1	Zino	$\tilde{Z}^0$	1/2

essere un superamento del MS, in quanto, tra i vari aspetti che la caratterizzano, prevede l'unificazione delle interazioni debole, forte e elettromagnetica, introduce l'interazione gravitazionale, teorizza l'esistenza di un valido candidato a costituire la materia oscura fredda.

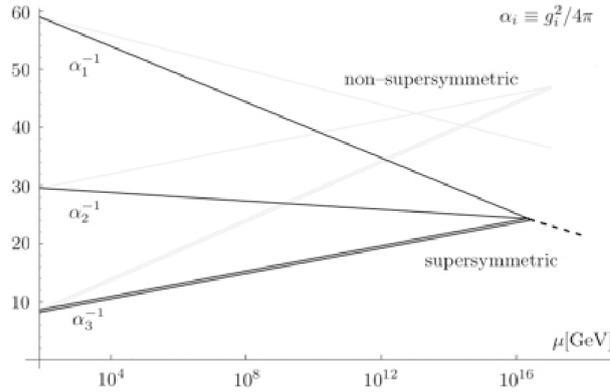
In SUSY viene postulata l'esistenza di una simmetria tra bosoni e fermioni. Una trasformazione supersimmetrica muta uno stato bosonico in uno fermionico e viceversa, dunque è necessario introdurre un partner supersimmetrico per ognuna delle particelle presenti nel MS. Alcune delle particelle previste in supersimmetria sono elencate nella tabella 1.4. I partner supersimmetrici delle particelle differiscono dalle particelle per il valore dello spin a essi associato. Se  $s$  è lo spin di una particella lo spin della *sparticella* corrispondente è  $s - 1/2$ [3].

Nel Modello Standard rimane irrisolto il problema della gerarchia, che si riscontra quando i parametri fondamentali, in tal caso le costanti di accoppiamento delle interazioni o le masse, sono molto diversi dai parametri misurati da un esperimento[4]. È questo il caso della massa dell'Higgs, per la quale gli esperimenti che hanno preceduto l'LHC sembrano indicare valori un po' più alti della scala elettrodebole, mentre nel Modello Standard essa dovrebbe essere innaturalmente elevata, dovendo tener conto di grandi correzioni radiative nei loop. Nella Supersimmetria il problema è naturalmente risolto, poiché i contributi delle particelle e dei loro partner supersimmetrici si elidono nelle correzioni della massa dell'Higgs. In realtà la Supersimmetria è necessariamente rotta, pertanto il contributo delle particelle virtuali e dei loro superpartner non è completamente compensato. Questo permette di fare delle previsioni riguardo le masse dei superpartner delle particelle: il contributo quantistico alla massa della particella di Higgs proveniente da una particella e dal suo superpartner virtuali, seppure diverso da zero, non supererà mai la differenza tra le loro masse, quindi tale differenza dovrà essere al più pari alla massa della scala debole ( $\sim 250$  GeV)[5].

La differenza tra le interazioni si manifesta nei valori estremamente diversi delle costanti di accoppiamento associate alle quattro interazioni fondamentali, come evidenziato nella tabella 1.5. In realtà per le interazioni debole forte e elettromagnetica le costanti di accoppiamento assumono valori differenti per processi che avvengono a differenti energie. Nel Modello Standard le costanti di accoppiamento, anche ad altissime energie, non convergono, mentre SUSY prevede la convergenza di tali costanti a valori elevati di energia. In particolare nel Modello Standard Supersimmetrico Minimale (MSSM) è prevista l'unificazione delle in-

**Tabella 1.5:** Costanti di accoppiamento delle interazioni fondamentali

Interazione	Simbolo	Valore costanti
forte	$\alpha_s$	1
elettromagnetica	$\alpha$	1/137,04
debole	$\alpha_w$	$10^{-6}$
gravitazionale	$\alpha_g$	$10^{-32}$



**Figura 1.1:** Confronto tra i reciproci delle costanti di accoppiamento nel MS e nel MSSM in funzione della scala di energia. Nel MS si ha solo l'unificazione delle interazioni elettromagnetica e debole, mentre nel MSSM si ha l'unificazione dell'interazione elettromagnetica, debole e forte. Con  $\alpha_3$  si indica la costante di accoppiamento per l'interazione forte, mentre  $\alpha_1$  e  $\alpha_2$  sono le costanti di accoppiamento associate all'interazione elettrodebole[6].

terazioni debole, forte ed elettromagnetica a un'energia dell'ordine di  $10^{16}$  GeV. Nella figura 1.1 è rappresentata l'interpolazione delle costanti di accoppiamento in funzione della scala di energia per il MS e il MSSM.

In SUSY si definisce la R-parità [7] è un numero quantico definito come:

$$R_p = (-1)^{2J+3B+L}$$

dove  $J$  è lo spin totale,  $B$  e  $L$  sono rispettivamente il numero barionico e leptonico. Per le particelle del MS  $R_p = 1$ , per i loro partner supersimmetrici  $R_p = -1$ . Se la R-parità è conservata, non possono esistere reazioni in cui le particelle supersimmetriche decadono in sole particelle del Modello Standard e deve esistere una particella SUSY più leggera (LSP, Lightest SUSY Particle) stabile e che sfugge alla rivelazione. La LSP è un ottimo candidato a spiegare l'origine della materia oscura fredda.

### 1.2.1 Il Modello Standard Supersimmetrico Minimale

La teoria supersimmetrica più semplice è il Modello Standard Supersimmetrico Minimale(MSSM). Nel MSSM si annovera il doppio delle particelle del MS, più due doppietti di Higgs[6].

Poiché i fermioni del MS sono caratterizzati da due gradi di libertà per lo spin, nel MSSM esisteranno due particelle supersimmetriche bosoniche corrispondenti. I neutrini costituiscono un'eccezione, perché, avendo un solo grado di libertà nel MS per lo spin hanno un corrispondente supersimmetrico, lo *sneutrino*. Per i quark e i leptoni esisteranno due partner bosonici per ogni specie, che costituiranno gli *squark* e gli *sleptoni*. Nel MSSM compaiono, dunque, 12 stati distinti di squark e 9 di sleptoni (considerando i 6 dei leptoni e i 3 dei neutrini).

I fermioni del MSSM sono i partner supersimmetrici dei bosoni vettori, i *gau-gini*, e i partner supersimmetrici dei bosoni scalari di Higgs, gli *higgsini*.

In particolare i gau-gini elettrodeboli e gli higgsini possono combinarsi per formare i chargini e i neutralini[3]. Il chargino è una miscela di *wini* carichi ( $\widetilde{W}^\pm$ ) e higgsini carichi. Nel MSSM esistono due stati di chargino:  $\widetilde{\chi}_1^\pm$  e  $\widetilde{\chi}_2^\pm$ , con masse  $M_{\widetilde{\chi}_1^\pm} \leq M_{\widetilde{\chi}_2^\pm}$ .

Il neutralino è una miscela di *zino*, *foto*, *wino* neutro e dei due higgsini neutri. Nel MSSM il neutralino esiste in 4 stati:  $\widetilde{\chi}_1^0$ ,  $\widetilde{\chi}_2^0$ ,  $\widetilde{\chi}_3^0$  e  $\widetilde{\chi}_4^0$ , con masse  $M_{\widetilde{\chi}_1^0} \leq M_{\widetilde{\chi}_2^0} \leq M_{\widetilde{\chi}_3^0} \leq M_{\widetilde{\chi}_4^0}$ .

Infine i partner dei gluoni sono i *gluini*, particelle con spin 1/2 che formano un ottetto di colore.

Tra i vari modelli proposti per giustificare la rottura spontanea della supersimmetria uno dei più significativi considera la rottura come dovuta all'interazione con la gravità, costituendo le fondamenta per i modelli di supergravità (SUGRA). Una teoria di supergravità è una teoria di campo che sintetizza la supersimmetria e la relatività generale. Nella supergravità si considera un campo il cui quanto è il gravitone, particella di spin 2. La supersimmetria richiede che al gravitone sia associato un partner supersimmetrico, il *gravitino*, particella di spin 3/2.

Il più semplice tra i modelli SUGRA è il modello di supergravità minimale (mSUGRA), in cui si associa a tutti gli squark e gli sleptoni massa  $m_0$  (bosoni) e a tutti i gau-gini massa  $m_{1/2}$  (fermioni) alla scala di energia della Grand Unification Theory (GUT),  $\sim 10^{16}$  GeV.

La rottura spontanea della simmetria implica che i partner supersimmetrici abbiano masse diverse e più grandi delle corrispondenti particelle. Le sparticelle sono instabili e decadono rapidamente, fino alla produzione della particella più leggera LSP e della materia ordinaria.

D'altronde se le particelle supersimmetriche fossero meno pesanti dei loro partner avrebbero dovuto essere già state osservate nei range di energia finora esplorati dai vari esperimenti condotti agli acceleratori.

La LSP è una particella che interagisce molto debolmente con la materia ordinaria e questo è il motivo per cui gli esperimenti di alte energie non possono rivelarla, se non considerando, così come si procede per individuare la presenza di neutrini, l'energia mancante nell'evento, che si ottiene imponendo la conservazione dell'energia. Nei modelli supersimmetrici validi candidati a ricoprire questo ruolo sono il neutralino e lo sneutrino.

## Capitolo 2

# L'esperimento ATLAS al collider LHC del CERN

Il Large Hadron Collider (LHC)[8] ha sede presso l'Organizzazione Europea per la Ricerca Nucleare (CERN) a Ginevra e l'esperimento ATLAS occupa uno dei quattro siti di interazione sull'anello. L'LHC è sorto proprio per cercare di dare una risposta ai quesiti ancora irrisolti nel Modello Standard, alcuni dei quali sono stati discussi nel capitolo 1. Per quanto concerne la Supersimmetria, l'LHC costituisce un valido "strumento" per riuscire ad individuare le particelle SUSY, se esse sono alla scala del TeV. In questo capitolo saranno descritti brevemente sia l'LHC che l'esperimento ATLAS, frutto del lungo lavoro e della collaborazione di migliaia di fisici (e non solo) di tutto il mondo.

All'interno dell'acceleratore avvengono collisioni protone-protone ad altissima energia e gli esperimenti acquisiscono i dati che vengono in seguito elaborati per ricostruire gli eventi di fisica interessanti.

### 2.1 Il Large Hadron Collider

Il Large Hadron Collider costituisce il più grande e complesso acceleratore finora concepito. L'LHC è installato in un tunnel circolare del diametro di circa 27 km e si estende dai 50 ai 175 m sottoterra. L'acceleratore si trova nei pressi di Ginevra, al confine tra Francia e Svizzera[9] (figura 2.1).

Nell'LHC si fanno collidere fasci di protoni contro protoni. I vantaggi di usare collisioni adroniche per studi di nuova fisica sono molteplici:

- permettono investigare un ampio intervallo di energia, in quanto l'energia nel centro di massa disponibile per la creazione di particelle è legata alle distribuzioni dei momenti dei partoni nei protoni collidenti;
- consentono di avere una grande varietà di stati iniziali (per la presenza di  $u$ ,  $d$ ,  $\bar{u}$ ,  $\bar{d}$ , gluoni e quark del mare);



Figura 2.1: Ubicazione geografica dell'acceleratore

- l'energia massima raggiungibile è molto maggiore rispetto a quella generata dagli acceleratori che lavorano con  $e^- e^+$ , dove c'è il problema della perdita di energia per radiazione di sincrotrone <sup>1</sup>.

Alcuni svantaggi dell'usare collisori di adroni sono, invece:

- le elevate sezioni d'urto degli eventi non interessanti per lo studio che si vuole condurre;
- le collisioni multiple tra partoni nello stesso incrocio di fasci, che danno come risultato stati finali piuttosto complessi;
- il sistema di riferimento del centro di massa di una collisione non è in quiete rispetto al sistema di riferimento del laboratorio.

All'LHC sono presenti 1232 dipoli superconduttori che operano alla temperatura di 1.9 K e deflettono i fasci di protoni sulla traiettoria circolare. A 7 TeV questi magneti producono un campo magnetico di 8.33 Tesla, con una corrente di 11700 A[9]. I magneti hanno due cavità, una per ognuno dei due fasci, che procedono in versi opposti nel tunnel. Ogni dipolo è lungo circa 14 metri. Le bobine contengono il cavo superconduttore in niobio e titanio e sono circondate da guaine di acciaio austenitico, un materiale che combina le proprietà richieste di buona risposta termica e permeabilità magnetica. Le guaine costringono le bobine a rimanere nella loro posizione, intervenendo contro la notevole forza magnetica generata all'interno dell'acceleratore. I dipoli dell'LHC comprendono circa 7600 km di cavo superconduttore, che pesa 1200 tonnellate. Ogni cavo è costituito da 36 filamenti di materiale superconduttore, ognuno dei quali "ospita" 6300 filamenti di NbTi. Un complesso di acceleratori pre-accelera i fasci di protoni, fino a raggiungere un'energia di circa 450 GeV, quindi circa 1500 pacchetti sono iniettati nell'LHC e sono accelerati fino a 7 TeV di energia del centro di massa (14 TeV è l'energia di

<sup>1</sup>La radiazione di sincrotrone è una radiazione elettromagnetica, prodotta per bremsstrahlung da particelle cariche, solitamente  $e^\pm$ , che viaggiano a velocità prossime alla velocità della luce e sono costrette a muoversi da un campo magnetico su traiettorie curve.

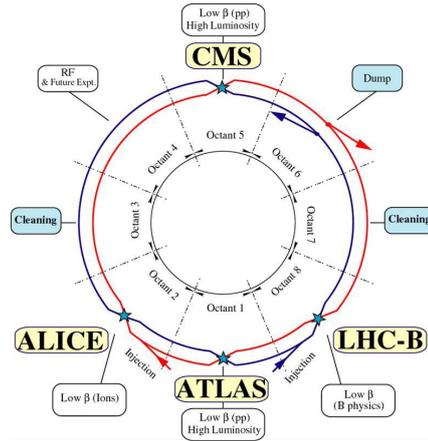


Figura 2.2: I punti d'interazione sull'acceleratore e esperimenti principali che vi operano

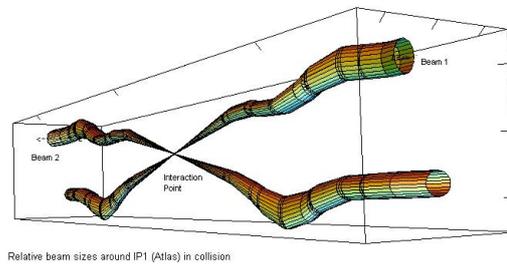


Figura 2.3: Rappresentazione delle dimensioni dei fasci che collidono nel punto d'interazione

progetto, ma verrà raggiunta nel corso dei prossimi anni). I due fasci sono collimati e fatti collidere in 4 punti d'interazione dove sono collocati gli esperimenti principali, mostrati in figura 2.2.

In figura 2.3 sono illustrati i due fasci nel punto d'interazione, in cui avviene la collisione. Sebbene ci siano i 100 miliardi di protoni per fascio nelle specifiche finali del progetto, molti non collidono e procedono senza dare origine a fenomeni di scattering, dunque lo stesso fascio è in grado di circolare nell'acceleratore per ore.

La *luminosità istantanea* di un collisore è data da:

$$L = f \frac{N n_1 n_2}{4\pi \sigma_x \sigma_y} F(\Phi, \sigma_{xy}, \sigma_s)$$

dove  $f$  rappresenta la frequenza di rotazione dei fasci che collidono, Con  $N$  si indica il numero dei pacchetti collidenti; con  $n_1$  e  $n_2$  si indica il numero dei protoni contenuti nei pacchetti. I valori  $\sigma_x$  e  $\sigma_y$  corrispondono alle dimensioni delle sezioni trasverse dei fasci che collidono, mentre  $F$  è un fattore di riduzione geometrico che dipende da  $\Phi$  (l'angolo formato tra le direzioni dei due fasci) da  $\sigma_{xy}$  (lo scarto quadratico medio nel piano trasverso) e  $\sigma_s$  (la lunghezza dei fasci). La luminosità istantanea di progetto all'LHC è di  $\approx 10^{34} \text{cm}^{-2} \text{s}^{-1}$ , mentre la luminosità istantanea di picco raggiunta sino ad oggi è di  $3.65 \cdot 10^{33} \text{cm}^{-2} \text{s}^{-1}$ .

Il rate d'interazione è dato da:

$$R = L \cdot \sigma$$

dove con  $\sigma$  si indica la sezione d'urto completamente anelastica, che ha un valore di circa 60 millibarn (1 barn =  $10^{-24}\text{cm}^2$ ). Poiché il tempo che intercorre tra due collisioni è di 25 ns e in ogni collisione avvengono più di 20 interazioni, il rate di interazioni è  $\approx 1$  GHz.

Le interazioni tipiche dei protoni infatti hanno un elevato parametro d'impatto e producono stati finali caratterizzati da valori dei momenti relativamente bassi per le particelle. Collisioni del genere sono chiamate "minimum bias collisions" e costituiscono un fondo enorme per gli eventi di fisica interessanti.

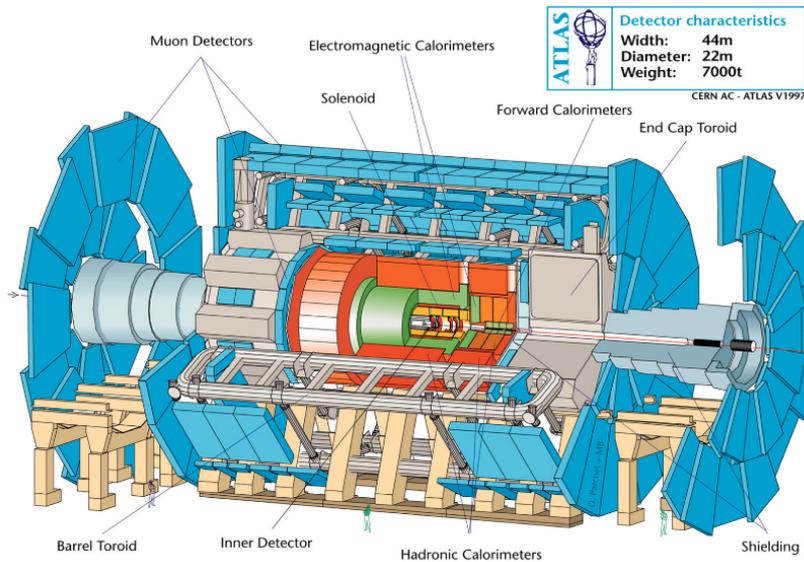
È necessario dunque avere a disposizione un sistema di trigger robusto e veloce, che effettui una selezione dei dati grezzi a partire dai segnali analogici provenienti dai detector in tempo reale. Il trigger filtra i dati che provengono dai detector e decide velocemente se questi sono sufficientemente interessanti perché possano passare a una successiva fase di selezione. Sono presenti tre diversi livelli di trigger, che permettono di ridurre il rate di eventi da 1 GHz, fino a  $\sim 10^2$  Hz degli eventi di interesse per le analisi, i quali vengono registrati permanentemente per poi essere analizzati offline. Un trigger ottimale è un trigger che dia una risposta veloce e che sia riprogrammabile, in modo da essere adattato alle esigenze dell'acquisizione dei dati. Ciononostante la quantità di dati prodotti all'LHC è smisurata e senza precedenti. Questo ha comportato la necessità di utilizzare sofisticate architetture di computing, per poter processare l'innumerabile quantità di dati acquisiti, su una griglia computazionale diffusa in tutti i paesi del mondo partecipanti agli esperimenti.

## 2.2 L'esperimento ATLAS ad LHC

ATLAS (A Toroidal LHC Apparatus)[10] è uno dei quattro rivelatori in funzione all'LHC. Il progetto ATLAS ha avuto origine dalla collaborazione dei gruppi di ricercatori che lavoravano ai rivelatori EAGLE (Experiment for Accurate Gamma, Lepton and Energy Measurements) e ASCOT (Apparatus with Super COnducting Toroids), che hanno fatto convergere il loro lavoro verso la costruzione di un unico rivelatore *general purpose*, ossia in grado di studiare molti aspetti delle collisioni dei protoni alle alte energie, senza focalizzarsi su un particolare processo fisico.

ATLAS è lungo 44 m, ha un diametro di 25 m e pesa 7000 tonnellate. Nella figura 2.4 è rappresentata un'immagine dell'esperimento. I rivelatori sono installati su due tipi di strutture; una cilindrica, coassiale alla direzione del fascio (*barrel*) e due "tappi" (*end-cap*) posti alle estremità di ATLAS e ortogonali al piano dell'orbita. Nell'apparato si individuano quattro componenti principali: il sistema dei magneti, l'inner detector, i calorimetri elettromagnetici e adronici e lo spettrometro a muoni.

Il sistema di riferimento è fatto di assi cartesiani ortogonali, con l'origine coincidente con il punto d'interazione tra i due fasci. L'asse  $Ox$  è la congiungente origine-centro dell'anello; l'asse  $Oy$  è diretto verso l'alto; l'asse  $Oz$  è diretto parallelamente al flusso di protoni e ha un verso tale che la terna appena definita sia



**Figura 2.4:** Il rivelatore di ATLAS: è alto ~25 m e lungo ~44 m. Sono indicati le diverse parti che lo compongono: il tracciatore interno, i calorimetri, il rivelatore di muoni, i magneti

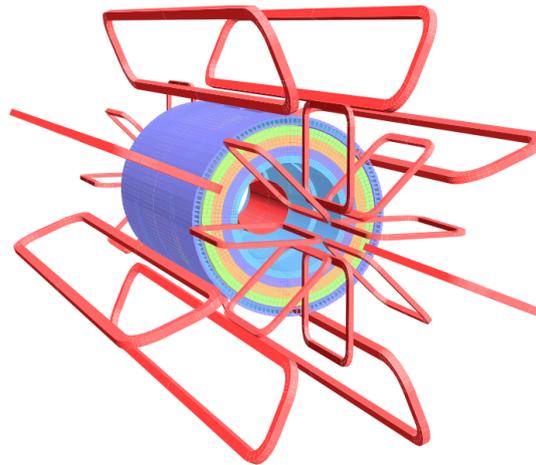
destrorsa. In coordinate sferiche, invece, si definisce  $\varphi$  l'angolo *azimutale* formato nel piano  $xy$  e  $\vartheta$  l'angolo *polare* formato con l'asse  $z$ . Per le misure dell'angolo polare si preferisce riferirsi a una nuova coordinata, la *pseudorapidità*:

$$\eta = -\log\left(\tan \frac{\vartheta}{2}\right)$$

### Il sistema dei magneti

Per misurare la carica e il momento delle particelle cariche ATLAS utilizza due diversi sistemi magnetici:

- un solenoide superconduttore che avvolge il tracciatore interno (inner detector) e genera un campo magnetico di 2 T, diretto parallelamente all'asse del fascio. Il solenoide è mantenuto a una temperatura di 4.5 K mediante un flusso di elio liquido;
- tre grandi toroidi in aria (air-core), uno nel barrel (BT) e due negli end-cap (ECT), posti al di fuori dai calorimetri, per lo spettrometro a muoni. Il campo non è uniforme, ha una struttura cilindrica con l'asse del cilindro coincidente con quello dei fasci ed è di 4 T in prossimità delle bobine del barrel, mentre assume un valore massimo di circa 5 T negli end-cap. Il campo magnetico prodotto dai tre toroidi copre complessivamente un intervallo  $0 < |\eta| \leq 3$ . Nella figura 2.5 è illustrata la struttura del sistema magnetico di ATLAS.



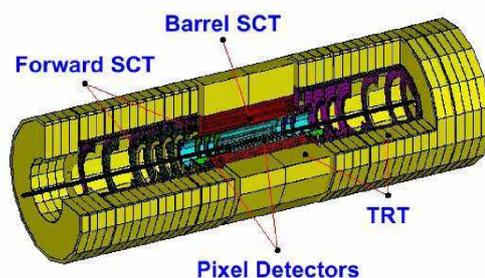
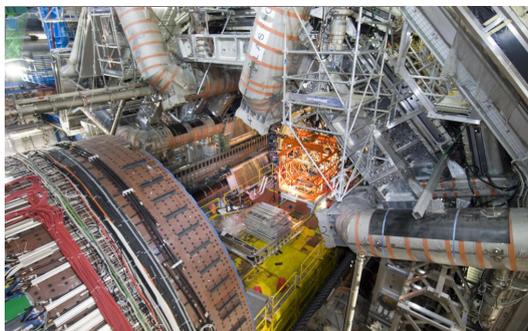
**Figura 2.5:** Schema dei magneti di ATLAS: sono evidenziati il solenoide, le 8 bobine del toroide nel barrel e le bobine dei due toroidi negli end-cap

### L'inner detector

Il tracciatore interno (inner detector) è ubicato in prossimità del tubo a vuoto dei fasci e costituisce il rivelatore più interno dell'apparato. Trovandosi nei pressi del luogo in cui avvengono le collisioni, è necessario che sia resistente all'elevata dose di radiazioni e che sia caratterizzato da un'alta risoluzione, in modo da ricostruire le traiettorie delle particelle cariche e i vertici primari e secondari. L'apparato è complessivamente una struttura cilindrica, lunga  $\sim 6$  m e con il diametro di  $\sim 2$  m ed è immerso in un campo magnetico solenoidale di 2 T, che determina la deflessione delle particelle cariche nel piano trasverso ai fasci, in modo che sia possibile misurarne il momento delle particelle.

L'inner detector si compone di tre strati, disposti come cilindri concentrici nel barrel e come dischi ortogonali al piano dell'orbita negli end-cap (figura 2.6).

Poiché è necessario avere un'alta risoluzione nella regione più interna sono qui installati i *pixel detector*, che permettono di ottenere un'elevata granularità. I pixel detector sono rivelatori al silicio rettangolari con dimensioni che vanno da 50 a 300  $\mu\text{m}^2$  e sono adibiti a identificare la posizione della particelle che li attraversa e



**Figura 2.6:** Installazione dell'inner detector di ATLAS all'interno del calorimetro (sinistra) e schema dell'inner detector e sue componenti (destra).

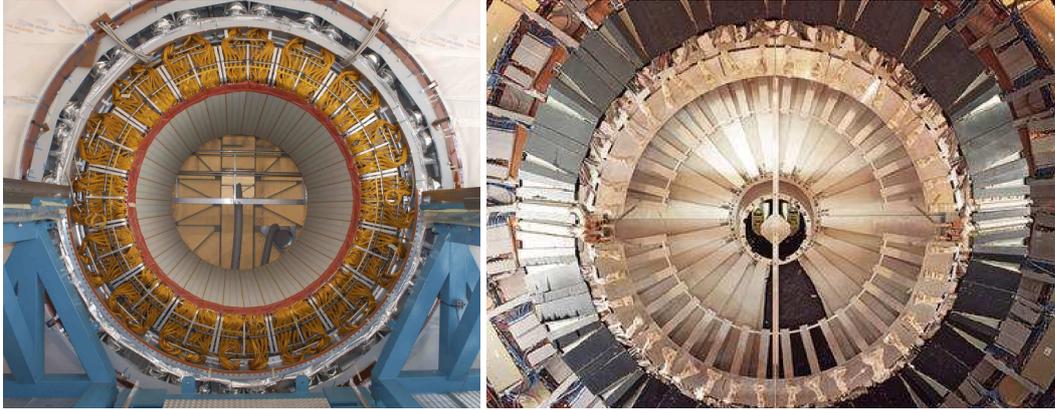


Figura 2.7: Calorimetri elettromagnetico (sinistra) e adronico (destra) installati ad ATLAS

a determinare i vertici secondari di decadimento.

Quando una particella attraversa un pixel viene generato un segnale elettrico che viene collegato a quelli prodotti dagli altri pixel per ricostruire la traiettoria della particella sfruttando algoritmi software di tracciamento. Il sistema è costituito di tre cilindri di raggio  $\approx 5$  cm,  $\approx 9$  cm e  $\approx 12$  cm nel barrel e di cinque dischi sull'end-cap di raggio compreso tra gli 11 e i 20 cm. Nella regione immediatamente successiva ai pixel detector sono installate le *Silicon Microstrips* (SCT), strisce di rivelatori al silicio molto sottili, lunghe qualche cm e larghe  $\sim 80$   $\mu$ m, disposte a piccoli angoli l'una rispetto all'altra, per misurare in maniera stereoscopica le coordinate  $\varphi$  e  $\eta$ . Il sistema è progettato per produrre 8 misure di precisione per traccia, per le misure di momento, parametri d'impatto e posizione dei vertici.

L'ultimo strato di rivelatori del tracciatore consiste di un *Transition Radiation Tracker* (TRT), fatto di 36 strati di sottili tubi di materiale plastico (di diametro  $\sim 4$  mm), con un filo sensibile centrale immerso in una miscela gassosa di Xenon, in grado sia di effettuare il tracciamento delle particelle cariche che di rivelare la radiazione di transizione prodotta in un radiatore posto tra i tubi in seguito al passaggio di una particella carica. Questa radiazione è prodotta solo da elettroni e da adroni di altissima energia e viene pertanto utilizzata per effettuare la discriminazione tra elettroni e adroni carichi.

## I calorimetri

I calorimetri hanno la funzione di assorbire e misurare l'energia di elettroni, fotoni e adroni.

I calorimetri elettromagnetici (figura 2.7) effettuano misure di energia, posizione e direzione di fotoni ed elettroni[9]. I calorimetri adronici (figura 2.7) sono ottimizzati per la misura dell'energia dei jet adronici e, poiché l'energia trasversa mancante viene misurata bilanciando nel piano trasverso l'energia totale misurata dai calorimetri, sia elettromagnetico che adronico, sono fondamentali per riconoscere la presenza di particelle che interagiscono molto poco con il rivelatore. Infine i calorimetri assumono un'importanza cruciale nella fase di trigger, che è fondamentale per la selezione o la scrittura degli eventi. Il sistema è posto al di fuori del campo magnetico solenoidale e si compone di:

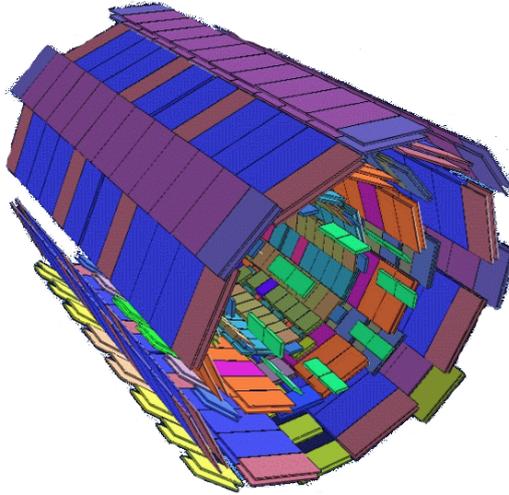
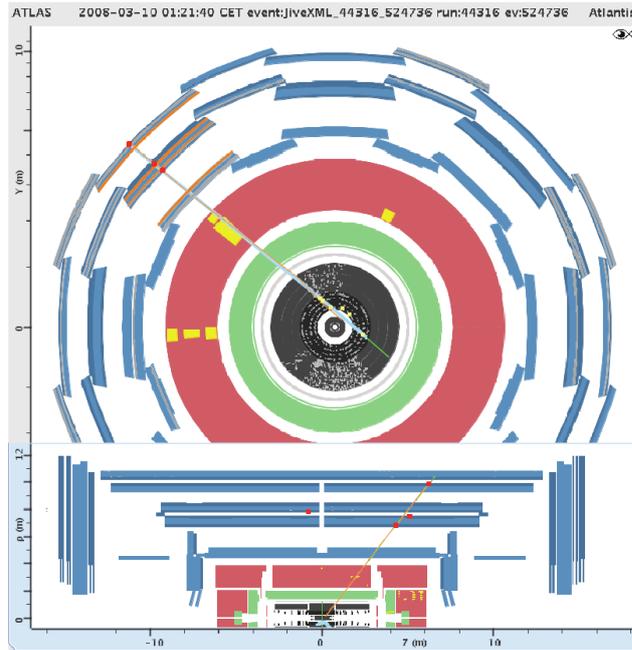


Figura 2.8: Lo spettrometro a muoni di ATLAS

- un calorimetro elettromagnetico che copre la regione di pseudorapidità  $|\eta| < 3.2$  ed è suddiviso in una parte collocata nel barrel ( $|\eta| < 1.475$ ) e due negli end-cap ( $1.475 < |\eta| < 3.2$ ). Per valori di pseudorapidità inferiori a 1.8 è installato un pre-campionatore per correggere le eventuali perdite di energia prima del calorimetro elettromagnetico. Nel barrel il calorimetro è costituito da due semi-cilindri identici separati da un gap di 6 mm lungo il piano  $z = 0$ , mentre negli end-cap è diviso in due ruote coassiali, una più interna ( $1.375 < |\eta| < 2.5$ ) e l'altra più esterna ( $2.5 < |\eta| < 3.2$ ). Utilizza piombo e argon liquido come mezzo attivo, mantenuto a una temperatura di 89 K, ed è realizzato secondo una geometria ad "accordion" (fisarmonica), per garantire un completo ricoprimento in funzione dell'angolo  $\varphi$ . La tecnologia ad argon liquido garantisce una buona resistenza alla radiazione e ottime risoluzioni in energia e posizione;
- un calorimetro adronico, che copre la regione  $|\eta| < 4.9$  e, nella regione  $|\eta| < 1.7$ , è basato su una tecnica di campionamento che usa un fitto insieme di mattonelle di materiale plastico (*scintillatori*) e assorbitori di ferro. Nelle regioni end-cap si è scelto di usare argon liquido per la sua proprietà intrinseca di tolleranza di radiazione. I calorimetri end-cap sono infatti costituiti da rame e argon liquido.
- un calorimetro forward, ossia nelle regioni a grandi valori di  $|\eta|$ , che utilizzano argon liquido ad alta densità e consiste di tre sezioni: la prima realizzata con elettrodi di rame e le altre due con elettrodi a barrette in una matrice di tungsteno.

### Lo spettrometro a muoni

I sistemi per rivelare i muoni occupano tipicamente vaste aree e sono costituiti da rivelatori a gas. In particolare, in ATLAS, diverse migliaia di metri quadri sono occupati da *camere* essenzialmente divise in due tipologie: quelle adibite alla



**Figura 2.9:** Evento di muone cosmico registrato ad ATLAS che attraversa i veri detector. Il tracker (silicon strip e straw tracker), i calorimetri adronici e i muon detector effettuano il read out, mentre il calorimetro ad argon liquido e il pixel system non effettuano il read out.

misurazione precisa della traccia del muone e quelle adibite al trigger dei muoni (figura 2.8). Il sistema dei toroidi genera un campo magnetico che produce la deflessione delle particelle nel piano  $R - \eta$ . Gli effetti di scattering multiplo, essendo una struttura aperta (“air-core”), sono minimizzati e si ha un’eccellente risoluzione del momento dei muoni, tracciati con tre stazioni di camere ad alta precisione, in una regione di pseudorapidità fino a  $|\eta| = 2.7$ . Le camere traccianti sono realizzate con strati di MDT (Monitored Drift Tubes), tubi di deriva di 3 cm di diametro orientati secondo le linee del campo magnetico solenoidale e riempiti da gas ad alta pressione (3 bar).

A grandi valori della pseudorapidità ( $2 < |\eta| < 2.7$ ) e nei pressi del punto d’interazione, in cui sia il flusso di particelle che il fondo di radiazione sono elevati, si utilizzano le CSC (Cathod Strip Chamber), camere a fili con strip catodiche con alta granularità. Il sistema di trigger dei muoni, caratterizzato da una risposta estremamente rapida, comprende RPC (Resistive Plate Chamber) nella regione del barrel ( $|\eta| < 1$ ) e TGC (Thin Gap Chamber) nelle regioni end-cap. Oltre ad effettuare il trigger di primo livello dei muoni queste camere effettuano la misura della “seconda coordinata”, ossia della coordinata nel piano  $R - \phi$ , perpendicolare a quello di curvatura. Le camere sono disposte in tre stazioni di trigger, due delle quali usate per identificare i muoni di basso valore del momento trasverso, e la terza, affiancata alle altre due, per identificare i muoni con alto valore del momento trasverso[9].

## Il trigger

Come già accennato nel paragrafo 2.2 alla luminosità istantanea dell'LHC vengono prodotte circa un miliardo di interazioni protone-protone al secondo, di cui solo pochissime sono realmente interessanti per la fisica che si vuole studiare e che, comunque, non potrebbero mai essere tutte registrate, ricostruite e analizzate in tempi ragionevoli per il sistema di acquisizione e di processamento. Solo una piccola frazione del rate iniziale ( $\sim 300 \div 400$  Hz) viene infine ritenuto.

Gli esperimenti necessitano dunque di un sistema di trigger che permetta di selezionare solo gli eventi interessanti, e che raggiunga un fattore di reiezione  $\approx 10^7$ .

Il trigger di ATLAS è basato su tre livelli: il primo livello riceve i dati dal rivelatore, mentre gli altri due (HLT, High Level Trigger) elaborano i dati dell'evento e le informazioni raccolte al livello precedente. Il primo livello è implementato in hardware, si basa sull'utilizzo di processori *custom* e usa una limitata quantità di informazioni derivanti dai detector per prendere decisioni in un tempo caratteristico di  $2.5 \mu\text{s}$  per selezionare gli eventi di interesse, riducendo il rate dei dati fino a raggiungere i  $\sim 75\text{kHz}$ .

Il secondo e il terzo livello di trigger sono realizzati in software e usano diverse migliaia di processori commerciali per processare algoritmi che consentono la riduzione finale dei dati fino al livello richiesto. La sincronizzazione degli elementi dei detector deve essere accurata per permettere che l'acquisizione dei dati sia coerente, per "assemblare" tutti i dati provenienti dalle varie parti del detector in una determinata collisione di fasci e ricostruire l'evento completo (event-building).

La dimensione media di un evento ATLAS è di  $\sim 1.5$  MB.

La decisione del trigger è basata sulla presenza di "oggetti" nei calorimetri e nello spettrometro a muoni, come fotoni, elettroni, muoni e jet, oltre che di energia trasversa e di energia trasversa mancante.

## Capitolo 3

# L'analisi multivariata

Dopo aver descritto il contesto teorico e sperimentale in cui si inserisce questo lavoro di tesi, in questo capitolo ci si soffermerà sulla descrizione delle tecniche impiegate per l'analisi preliminare che si è sviluppata per la ricerca del partner supersimmetrico del top nei suoi decadimenti in un top e un neutralino. Queste tecniche utilizzano l'analisi multivariata. L'analisi multivariata, come suggerisce la parola stessa, consiste in un'analisi statistica in cui vengono impiegate più variabili. Questo tipo di analisi si utilizza per l'identificazione di eventi rari in presenza di una forte componente di fondo.

Lo scopo finale è quello di ottimizzare il rapporto segnale-fondo per massimizzare la sensibilità a una possibile scoperta o i limiti di esclusione delle sezioni d'urto di produzione diretta dello stop.

### 3.1 Tecniche di analisi multivariata

L'analisi convenzionale per la ricerca di nuove particelle consiste nel ridurre il contributo del fondo applicando un insieme di tagli su un campione sulle osservabili che discriminano il fondo dal segnale che si vuole cercare. Il valore dei tagli è generalmente determinato dal confronto tra le distribuzioni del segnale e quelle dei fondi generati con metodi di simulazione Monte Carlo. Questa procedura è corretta nel momento in cui i fondi sono conosciuti e il Monte Carlo è stato accuratamente calibrato e riproduce i dati.

Le tecniche di analisi multivariata costituiscono un approccio in principio più efficiente dell'analisi convenzionale. Il vantaggio dell'usare tali tecniche risiede nella capacità intrinseca dell'analisi di utilizzare un gran numero di variabili combinandone il potere discriminante, ma tenendo conto delle correlazioni esistenti tra le variabili utilizzate, in modo da ridimensionare la complessità del problema.

#### 3.1.1 I classificatori

I classificatori nell'analisi multivariata sono entità che permettono di identificare la classe di appartenenza di un campione statistico. Si consideri un'osservazione sperimentale, dalla quale si ottiene un campione di valori che costituiscono il vettore dell'evento, che chiamiamo  $x$ : gli elementi del vettore sono i valori delle osservabili che caratterizzano il processo fisico in esame. Il vettore  $x$  costituisce

l'oggetto di una regola di decisione, che permette di determinare a quale classe tale vettore appartiene. È importante notare che se il vettore è costituito da variabili casuali la cui distribuzione di probabilità condizionata dipenda dalle categorie a cui l'evento può appartenere e se tali distribuzioni sono note l'analisi multivariata si riduce al problema dei test di ipotesi[3].

Si consideri ad esempio il seguente problema di assegnazione di un evento a una data classe di appartenenza, la classe di segnale ( $s$ ) e la classe di fondo ( $b$ ). Una regola di decisione può essere determinata a partire dal confronto delle probabilità condizionate o a posteriori, come segue:

$$\begin{cases} P(s|\mathbf{x}) > P(b|\mathbf{x}) \rightarrow \mathbf{x} \in S \\ P(s|\mathbf{x}) \leq P(b|\mathbf{x}) \rightarrow \mathbf{x} \in B \end{cases} \quad (3.1)$$

dove le probabilità  $P(s|\mathbf{x})$  e  $P(b|\mathbf{x})$  sono le probabilità a posteriori che un dato evento appartenga rispettivamente alla classe di segnale o alla classe di fondo, condizionatamente al fatto che tale evento è il vettore  $\mathbf{x}$ .

Sfruttando in questo caso il teorema di Bayes è possibile determinare i valori delle probabilità a posteriori a partire dalle probabilità a priori  $P(s)$  e  $P(b)$  e dalle funzioni di densità di probabilità condizionate di un evento nelle classi di segnale  $P(\mathbf{x}|s)$  e di fondo  $P(\mathbf{x}|b)$ :

$$P(b[s]|\mathbf{x}) = \frac{P(\mathbf{x}|b[s])P(b[s])}{P(\mathbf{x})} \quad (3.2)$$

Si sostituisce l'equazione 3.2 nella 3.1 e poiché compare la  $P(\mathbf{x})$  al denominatore in entrambe le disequazioni si ottiene la seguente regola di decisione per l'appartenenza del vettore  $\mathbf{x}$  alle suddette classi:

$$R(\mathbf{x}) = \frac{P(\mathbf{x}|s)}{P(\mathbf{x}|b)} \begin{cases} \geq \frac{P(b)}{P(s)} = \beta \rightarrow \mathbf{x} \in s \\ < \frac{P(b)}{P(s)} = \beta \rightarrow \mathbf{x} \in b \end{cases} \quad (3.3)$$

dove con  $\beta$  si indica il valore di soglia per la decisione e  $R(\mathbf{x})$  è il rapporto di verosimiglianza di Bayes.

Scelto  $\beta$ , valore di soglia, moltiplicandolo opportunamente per un fattore, la regola che ne risulta è efficiente, nel senso che minimizza la possibilità di assegnare l'evento alla classe di appartenenza errata.

La regola (3.3) può essere riscritta introducendo un discriminante tra le due classi di appartenenza, definito a partire dalle funzioni di distribuzione di densità di probabilità condizionata  $P(\mathbf{x}|s)$  e  $P(\mathbf{x}|b)$  che indichiamo rispettivamente con  $f_s(\mathbf{x})$  e  $f_b(\mathbf{x})$ :

$$D(\mathbf{x}) = \frac{f_s(\mathbf{x})}{f_s(\mathbf{x}) + f_b(\mathbf{x})}$$

La regola (3.3) allora sarà:

$$R'(\mathbf{x}) = \begin{cases} D(\mathbf{x}) \geq \alpha \rightarrow \mathbf{x} \in s \\ D(\mathbf{x}) < \alpha \rightarrow \mathbf{x} \in b \end{cases} \quad (3.4)$$

dove  $\alpha$  assume valori nell'intervallo  $[0, 1]$ .

In realtà non si dispone delle funzioni densità di probabilità suddette, ma il discriminante è considerato a partire dalle stime delle distribuzioni di probabilità. Quindi il discriminante da considerare sarà:

$$\tilde{D}(\mathbf{x}) = \frac{\tilde{f}_s(\mathbf{x})}{\tilde{f}_s(\mathbf{x}) + \tilde{f}_b(\mathbf{x})}$$

I metodi di analisi multivariata forniscono una stima delle densità di probabilità reali e dei discriminanti ad esse correlati.

### 3.1.2 Metodi di analisi multivariata

I metodi di analisi multivariata consistono nella costruzione di classificatori che permettano di ottenere la massima separazione tra segnale e fondo, una volta dati in input il campione di segnale che si vuole studiare e il fondo, entrambi con dimensione finita. L'identificazione è resa possibile per la capacità dei classificatori di selezionare, a partire da un insieme di variabili che descrivono il processo, il sottoinsieme di variabili migliore per discriminare tra segnale e fondo, ovviamente rispetto alla statistica a disposizione.

L'utilizzo dei metodi di analisi multivariata consiste di due fasi. Nella prima fase, detta *training*, si considerano due campioni di eventi Monte Carlo di dimensione finita, uno per il segnale e uno per il fondo. Rispetto a sottoinsiemi di tali campioni si implementa l'algoritmo che permette al metodo di costruire il discriminante. Per valutare la prestazione del metodo scelto, si utilizzano valori di efficienza e reiezione su sottoinsiemi del campione di training scorrelati da quelli su cui è costruito il classificatore.

Nella seconda fase, detta *application*, il discriminante viene utilizzato per decidere sulla classificazione di un campione di eventi dato in input al metodo. La classificazione può essere effettuata in due o più classi, a seconda del metodo. In linea di principio il campione di application, di cui è ignota la classificazione, è completamente scorrelato da quelli di training. In realtà, nel nostro caso, questi campioni non sono del tutto scorrelati, perché il nostro studio di fattibilità mira a valutare la sensibilità raggiunta dall'analisi su campioni Monte Carlo perché possa, in seguito, essere applicata sui dati. Parte degli stessi campioni Monte Carlo è stata pertanto utilizzata sia per il training che per l'application.

## 3.2 TMVA

Per implementare l'analisi multivariata è stato utilizzato il programma ROOT[11] (versione 5.27) e il suo ambiente integrato per l'analisi multivariata Toolkit for Multivariate Analysis (TMVA versione 4.1.2). Nel pacchetto TMVA-v4.1.2 sono presenti diverse macro, modificabili rispetto alle esigenze dell'utente[12]. Come esposto nel paragrafo precedente l'analisi multivariata consiste di due fasi: una di apprendimento o training e una fase successiva di esecuzione o application.

Per la fase di training viene utilizzata la macro `TMVAClassification.C` e per la fase di esecuzione viene usata la macro `TMVAClassificationApplication.C`, entrambe fornite dal pacchetto TMVA.

Tutte le tecniche di analisi multivariata di TMVA appartengono alla famiglia di algoritmi di “apprendimento supervisionato”. Tali tecniche utilizzano eventi di training, per i quali è noto l'output, per determinare la funzione di mapping che descrive la regola di decisione (*classification*). La funzione di mapping può contenere diversi gradi di approssimazione e può essere una singola funzione globale o un insieme di modelli locali.

TMVA consente di confrontare diversi metodi di analisi multivariata sugli stessi dati, per riuscire a valutare quello che meglio si adatta a discriminare tra il segnale e il fondo per il caso specifico in analisi.

L'utente interagisce con il pacchetto mediante le classi definite *Factory* e *Reader*. Le *Factory* gestiscono le diverse fasi dell'analisi TMVA. Effettuano un'analisi e un processamento preliminari dei dati di training per determinare le proprietà basilari delle variabili discriminanti usate come input per i classificatori. Dalla prima fase sono calcolati i valori dei coefficienti di correlazione lineare delle variabili di input e i cosiddetti “pesi” che descrivono la funzione di mapping che verrà utilizzata nella seconda fase di esecuzione.

Per confrontare le prestazioni dei metodi si considera l'efficienza del segnale e la reiezione del fondo, come calcolate nella fase di training.

Dal training viene prodotto un file root accessibile grazie alla macro *TMVAGui.C* che permette di visualizzare graficamente i risultati ottenuti e semplificarne l'interpretazione: sono rappresentate graficamente diverse informazioni riguardo le variabili di input, le correlazioni tra le variabili, l'efficienza del taglio del classificatore, la reiezione del fondo in funzione dell'efficienza del segnale. Viene creato inoltre un file xml in cui sono contenuti i pesi. Proprio alla lettura e all'interpretazione dei file xml è adibita la classe *Reader*.

Per effettuare il confronto dei metodi utilizzati ci si riferisce al valore della significanza statistica, definita in TMVA come:

$$\eta = \frac{S(y)}{\sqrt{S(y) + B(y)}} \quad (3.5)$$

dove con  $S(y)$  e  $B(y)$  si indicano rispettivamente il segnale e il fondo, opportunamente pesati per tener conto occorrenze relative, rispetto alla variabile  $y$ , che rappresenta il classificatore. In TMVA viene calcolato e restituito all'utente il valore del taglio che massimizza la (3.5).

Dopo la fase di training con la macro *TMVAClassificationApplication.C* si processano i segnali. La macro prende come input: il training scelto che, fissate le variabili di input è funzione del segnale, del fondo e del metodo; un campione di eventi di segnale, generato nel nostro caso variando le masse del neutralino e dello stop su una griglia di punti ben definita, frammisto a varie sorgenti di fondo.

### 3.2.1 Metodi per l'analisi multivariata inglobati in TMVA

Il pacchetto TMVA fornisce diversi metodi di analisi multivariata, alcuni dei quali vengono brevemente descritti nel seguito e sono quelli che sono stati utilizzati per il processo di training e confrontati per scegliere quello che meglio si adattasse alla natura del problema in esame e alle variabili di input scelte.

Prima di passarle al metodo, per poter trattare le variabili nei processi di classificazione, TMVA può operare quattro operazioni di pre-processamento:

- normalizzazione delle variabili;
- decorrelazione lineare mediante la radice quadrata della matrice di covarianza;
- decorrelazione mediante la Principal Components Analysis (PCA);
- trasformazione delle variabili in distribuzioni gaussiane (“gaussianizzazione”), utile prima di effettuare le procedure di decorrelazione (che assumono correlazioni lineari e variabili gaussiane).

Ogni metodo risulta essere particolarmente efficiente per un determinato tipo di problema di classificazione. In particolare, tra i metodi discussi nel seguito, si individuano:

- metodi che non hanno generalmente una buona prestazione rispetto a grandi insiemi di training (k-NN, H-Matrix);
- metodi ottimali nella trattazione di dati con variabili aventi distribuzioni gaussiane e correlazioni lineari (Fisher, LD);
- metodi la cui prestazione dipende dalla funzione discriminante inserita dall’utente (FDA);
- metodi ottimali per trattare grandi insiemi di dati sia per la fase di training che per la fase di application (PDE, MLP, BDT).

Nel nostro caso, per esempio, i metodi con prestazione migliore sono stati i PDE, MLP e BDT, tra i quali si è preferito usare il BDT (e BDTG in particolare descritto nel seguito), che risulta essere particolarmente utile nel caso in cui le variabili di input scelte abbiano un modesto potere discriminante.

### Funzione discriminante di Fisher

Il metodo si basa sulla selezione degli eventi effettuata in uno spazio delle variabili trasformate con zero correlazioni lineari, distinguendo i valori medi delle distribuzioni del segnale e del fondo. L’analisi con discriminante lineare porta a delineare un asse nell’iperspazio delle variabili di input, in modo che, quando vengono proiettate su tale asse le classi di output (segnale e fondo) tali classi sono quanto più separate possibile, mentre gli eventi appartenenti a una stessa classe sono quanto più “vicini” possibile l’uno all’altro.

La classificazione degli eventi si basa su:

- i valori medi dell’intero campione  $\bar{x}_k$  per ogni variabile di input  $k = 1, \dots, n_{\text{var}}$ ;
- i valori medi dei campioni delle classi specifiche  $\bar{x}_{S(B),k}$ ;
- la matrice totale di covarianza  $C$  del campione.

La matrice di covarianza può essere suddivisa in due matrici distinte: la matrice  $W$  e la matrice  $B$ . La matrice  $W$  è la cosiddetta *within-class matrix* e descrive la dispersione degli eventi dalla media all'interno della classe specifica. La matrice  $B$  è la *between-class matrix* e descrive la dispersione degli eventi relativamente all'intero campione.<sup>1</sup>

I coefficienti di Fisher  $F_k$  sono dati da:

$$F_k = \frac{\sqrt{N_S N_B}}{N_S + N_B} \sum_{j=1}^{n_{\text{var}}} W_{kj}^{-1} (\bar{x}_{S,j} - \bar{x}_{B,j})$$

in cui  $N_{S(B)}$  è il numero di eventi di segnale (fondo) nel campione di training.

Il discriminante di Fisher per l' $i$ -esimo evento è:

$$y_F(i) = F_0 + \sum_{k=1}^{n_{\text{var}}} F_k x_k(i)$$

dove il termine  $F_0$  trasla la media  $\bar{y}_F$  del campione della totalità degli eventi ( $N_S + N_B$ ) a zero.

I coefficienti di Fisher possono dunque essere determinati a partire dalla matrice della covarianza come segue:

$$F_k = \frac{\sqrt{N_S N_B}}{N_S + N_B} \sum_{j=1}^{n_{\text{var}}} C_{kj}^{-1} (\bar{x}_{S,j} - \bar{x}_{B,j})$$

dove  $C_{kj} = W_{kj} + B_{kj}$ .

### Projective likelihood estimator (PDE)

Il metodo del massimo della likelihood consiste nella costruzione di un modello delle funzioni densità di probabilità che riproducono le variabili di input per il segnale e il fondo. Per un dato evento la likelihood che questo evento appartenga alla classe segnale è ottenuta moltiplicando la funzione densità di probabilità del segnale per tutte le variabili di input, considerate indipendenti, e normalizzando tale quantità con la somma delle probabilità di segnale e fondo. Poiché vengono ignorate le correlazioni tra le variabili di input questo metodo PDE viene talvolta chiamato "stimatore naive di Bayes".

<sup>1</sup>La matrice within-class è costituita da elementi:

$$W_{kj} = \sum_{U=S,B} \langle x_{U,k} - \bar{x}_{U,k} \rangle \langle x_{U,j} - \bar{x}_{U,j} \rangle = C_{S,kj} + C_{B,kj}$$

con  $C_{S(B)}$  matrice di covarianza del campione del segnale (fondo).

La matrice between-class ha elementi dati da:

$$B_{kj} = \frac{1}{2} \sum_{U=S,B} (\bar{x}_{U,k} - \bar{x}_k)(\bar{x}_{U,j} - \bar{x}_j)$$

dove con  $\bar{x}_{S(B)}$  si indica il valor medio della  $k$ -esima variabile per il campione del segnale (fondo) e con  $\bar{x}_k$  si indica il valor medio della suddetta variabile valutata sull'intero campione.

Si definisce il *likelihood ratio* usato dal PDE per l' $i$ -esimo evento:

$$y_L(i) = \frac{L_S(i)}{L_S(i) + L_B(i)} \quad (3.6)$$

in cui si ha:

$$L_S(i) = \prod_{k=1}^{n_{\text{var}}} p_{S,k}(x_k(i)) \quad L_B(i) = \prod_{k=1}^{n_{\text{var}}} p_{B,k}(x_k(i))$$

con  $p_{S(B),k}$  densità di probabilità di segnale (fondo) per la  $k$ -esima variabile  $x_k$ . Le funzioni densità di probabilità sono normalizzate, dunque vale la seguente relazione:

$$\int_{-\infty}^{+\infty} p_{S(B),k}(x_k(i)) dx_k = 1 \quad \forall k$$

Si può mostrare che in assenza di correlazioni tra le variabili il rapporto (3.6) porta a una separazione ottimale tra il segnale e il fondo per un dato insieme di variabili di input.

### Discriminante H-Matrix

L'approccio proposto si delinea nel contesto dei classificatori gaussiani. La discriminazione è svolta tra due classi, una di segnale e una di fondo. Gli elementi correlati di un vettore sono assunti avere una distribuzione gaussiana e l'inverso della matrice di covarianza è chiamata *H-Matrix*. È costruito uno stimatore  $\chi^2$  multivariato che sfrutta le differenze dei valori medi degli elementi dei vettori tra le due classi con lo scopo di effettuare la discriminazione. Per l'evento  $i$  sono costruiti gli stimatori  $\chi^2$  per la classe di segnale e di fondo,  $\chi_S^2$  e  $\chi_B^2$ , sfruttando le stime per le medie dei campioni  $\bar{x}_{S(B),k}$  e le matrici di covarianza  $C_{S(B)}$  determinate a partire dai dati di training:

$$\chi_{S(B)}^2(i) = \sum_{j,k=1}^{n_{\text{var}}} (x_j(i) - \bar{x}_{S(B),j}) C_{S(B),jk}^{-1} (x_k(i) - \bar{x}_{S(B),k})$$

dalla suddetta espressione si determina il discriminante:

$$y_H(i) = \frac{\chi_B^2(i) - \chi_S^2(i)}{\chi_B^2(i) + \chi_S^2(i)}$$

utilizzato per separare le classi di segnale e fondo.

### Classificatore k-Nearest Neighbour (kNN)

Il metodo confronta un evento osservato (test) con degli eventi di riferimento derivanti da un training. Risulta particolarmente efficiente nel caso in cui il "confine" tra il segnale e il fondo è irregolare.

Il metodo appronta la ricerca degli eventi più “vicini” all’evento di test. La “vicinanza” è determinata a partire da una metrica. La più semplice metrica è la distanza euclidea:

$$d = \left( \sum_{i=1}^{n_{\text{var}}} |x_i - y_i|^2 \right)^{\frac{1}{2}}$$

dove  $n_{\text{var}}$  è il numero di variabili di input,  $x_i$  sono le coordinate di un evento di training e  $y_i$  sono le coordinate dell’evento di test. I  $k$  eventi con il più piccolo valore di  $d$  sono i  $k$ -nearest neighbours. Il valore di  $k$  determina la dimensione dell’intorno per cui viene valutata la funzione densità di probabilità. Grandi valori di  $k$  non permettono di descrivere in maniera ottimale il comportamento locale della funzione densità di probabilità, mentre valori troppo piccoli di  $k$  causano notevoli fluttuazioni per la funzione densità di probabilità. Applicazioni del metodo su dati reali hanno individuato il range ottimale per il valore di  $k$  nell’intervallo  $[10, 100]$ , con centinaia di migliaia di eventi e un basso valore di  $n_{\text{var}}$ .

L’algoritmo trova i  $k$  eventi di training più vicini al punto richiesto, con  $k = k_S + k_B$ , dove  $k_{S(B)}$  è il numero di eventi di segnale(fondo) del campione di training. La probabilità relativa che l’evento di test appartenga alla classe di segnale è:

$$P_S = \frac{k_S}{k_S + k_B} = \frac{k_S}{k}$$

### Analisi con discriminante lineare (LD)

L’analisi effettua la classificazione dei dati utilizzando un modello lineare, nel senso che la funzione discriminante  $y(\mathbf{x})$  è lineare nei parametri  $\beta$  e ha la forma:

$$y(\mathbf{x}) = \mathbf{x}^T \beta + \beta_0$$

dove  $\beta_0$  è determinato in modo tale che si abbia  $y(\mathbf{x}) \geq 0$  per il segnale e  $y(\mathbf{x}) < 0$  per il fondo. Tale metodo risulta essere equivalente al discriminante di Fisher, in cui i dati sono proiettati su un sottospazio lineare.

Si considerino  $m + 1$  parametri  $\beta_0, \dots, \beta_m$  che devono essere stimati su un insieme di training di  $n$  eventi. Si può scrivere l’equazione in  $\beta$ , vettore dei parametri, in forma matriciale, introducendo il vettore  $Y$  e la matrice  $X$  come segue:

$$Y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \quad X = \begin{pmatrix} 1 & x_{11} & \dots & x_{1m} \\ 1 & x_{21} & \dots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \dots & x_{nm} \end{pmatrix}$$

La relazione che lega  $X$  e  $Y$  sarà allora:

$$Y = X\beta$$

In particolare il vettore  $Y$  è costituito da elementi tali che  $y_i = 1$  se l’evento appartiene alla classe del segnale,  $y_i = 0$  se l’evento appartiene alla classe del fondo. La suddetta equazione si risolve con il metodo dei minimi quadrati e si ha:

$$X^T X \beta = X^T Y \iff \beta = (X^T X)^{-1} X^T Y$$

dove il vettore  $\beta$  è quello che minimizza i minimi quadrati e la matrice  $(X^T X)^{-1} X^T$  prende il nome di matrice *pseudo-inversa di Moore-Penrose*, che può essere considerata una generalizzazione della matrice inversa per matrici non quadrate. Si può determinare la pseudo-inversa di  $X$  solo se  $X$  ha rango completo, cioè se i vettori colonna che formano la matrice sono tutti linearmente indipendenti.

### Analisi con funzione discriminante (FDA)

L'analisi permette di risolvere problemi relativamente semplici o parzialmente non-lineari. L'utente deve specificare la funzione discriminante da utilizzare e i suoi parametri. La FDA permette di elaborare un fit di tali parametri, in modo che il valore della funzione sia quanto più vicino possibile al valore reale, che è pari a 0 per il fondo e 1 per il segnale.

Rispetto alla complessità della funzione discriminante scelta TMVA permette di utilizzare un solo algoritmo o una convergenza di algoritmi per definire l'estimatore migliore. l'estimatore pesato utilizzato nei problemi di classificazione è il seguente:

$$\varepsilon = \frac{1}{W_S} \sum_{i=1}^{N_S} (F(\mathbf{x}_i) - 1)^2 w_i + \frac{1}{W_B} \sum_{i=1}^{N_B} (F(\mathbf{x}_i))^2 w_i$$

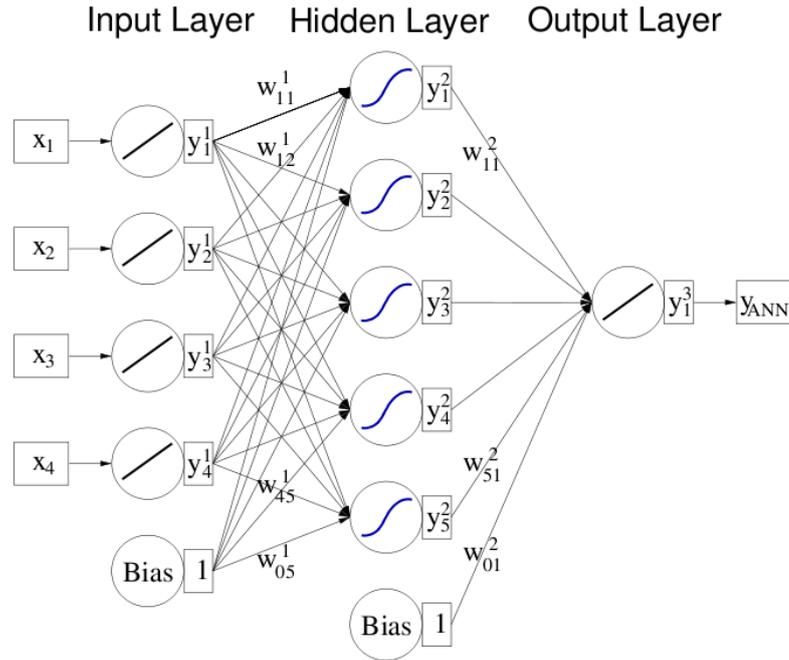
dove le somme sono sugli eventi di segnale e fondo del training,  $F(\mathbf{x}_i)$  è la funzione discriminante calcolata sull' $i$ -esimo evento rappresentato dal vettore  $\mathbf{x}_i$  i cui elementi sono i valori delle  $n_{\text{var}}$  variabili di input per quell'evento,  $w_i$  è il peso dell' $i$ -esimo evento e  $W_{S(B)}$  è la somma dei pesi degli eventi di segnale (fondo) per il campione di training.

### Multilayer Perceptron (MLP)

L'analisi rientra nella più vasta classe di analisi con discriminatori non-lineari, l'*Artificial Neural Network* (ANN). Si configura una rete neurale, in cui ogni neurone dà un responso rispetto al dato insieme di dati di segnale di input. I dati di segnale "attivano" uno o più neuroni (input) della rete, che formano una configurazione o stato della rete. In seguito alla risposta di tali neuroni (output) si determina la configurazione della rete. Questo processo può essere considerato un mapping di uno spazio delle variabili di input  $x_1, \dots, x_{n_{\text{var}}}$  in uno spazio di variabili unidimensionale, come nel caso di discriminazione segnale-rumore, o multidimensionale di variabili di output  $y_1, \dots, y_{m_{\text{var}}}$ .

Il comportamento di una rete neurale artificiale è dato dalla configurazione dei neuroni, dalle connessioni tra i neuroni e dalla risposta dei neuroni alle variabili di input, determinata dalla *funzione di risposta neurale*  $\rho$ .

Una rete neurale costituita da  $n$  neuroni comprende  $n^2$  connessioni e il problema risulta complesso. Per ovviare a tale complessità si procede a disporre i neuroni su strati (*layer*), in modo da ottenere unicamente connessioni dirette tra uno strato e lo strato successivo: si parla allora di *Multilayer perceptron*. Tutte le implementazioni di reti neurali in TMVA sono di questo tipo. Il primo strato è lo strato di input, l'ultimo lo strato di output e gli strati intermedi sono gli strati *hidden*. Per un problema di classificazione con  $n_{\text{var}}$  variabili di input lo strato iniziale assegna



**Figura 3.1:** Schema del funzionamento delle connessioni tra neuroni di una rete neurale artificiale a molti layer con uno strato *hidden*

i valori delle variabili di input a  $n_{\text{var}}$  neuroni, e lo strato di output comprende un neurone che costituisce l'estimatore della rete neurale,  $y_{\text{ANN}}$ . Nella figura 3.1 è mostrato uno schema del funzionamento della rete.

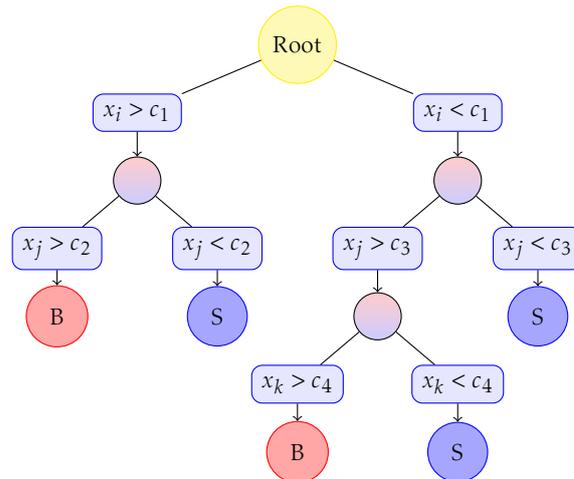
La funzione di risposta neurale  $\rho$  connette  $n$  variabili di input  $i_1, \dots, i_n$  al neurone di output. Spesso la funzione  $\rho$  può essere considerata una funzione composta:

$$\rho = \alpha \circ \kappa$$

con  $\alpha: \mathbb{R} \rightarrow \mathbb{R}$  funzione di *attivazione neurale*,  $\kappa: \mathbb{R}^n \rightarrow \mathbb{R}$  funzione di *sinapsi*, dove le funzioni possono avere la seguente forma:

$$\kappa: (y_1^{(l)}, \dots, y_n^{(l)} | w_{0j}^{(l)}, \dots, w_{nj}^{(l)}) \rightarrow \begin{cases} w_{0j}^{(l)} + \sum_{i=1}^n y_i^{(l)} w_{ij}^{(l)} \\ w_{0j}^{(l)} + \sum_{i=1}^n (y_i^{(l)} w_{ij}^{(l)})^2 \\ w_{0j}^{(l)} + \sum_{i=1}^n |y_i^{(l)} w_{ij}^{(l)}| \end{cases}$$

$$\alpha: x \rightarrow \begin{cases} x \\ \frac{1}{1 + e^{-kx}} \\ \frac{e^x - e^{-x}}{e^x + e^{-x}} \\ e^{-x^2/2} \end{cases}$$



**Figura 3.2:** Decision tree. A partire dal nodo root si dividono i dati in una sequenza di suddivisioni binarie, utilizzando le variabili di discriminazione  $x_j$ . Ogni suddivisione usa la variabile che in quel preciso nodo garantisce la miglior separazione tra il segnale e il fondo rispetto a quanto si taglia su di essa. Allora la stessa variabile potrebbe essere usata per discriminare in diversi nodi, altre potrebbero non essere usate affatto. Al termine del tree si trovano le *leaf* (foglie) contrassegnate da S per il segnale e B per il fondo, a seconda della maggioranza degli eventi che finiscono in quel nodo.

### Boosted Decision trees (BDT)

Questo metodo, e in particolare il BDTG che sarà trattato di seguito, è stato estensivamente usato nell'analisi descritta poiché, come vedremo nel capitolo seguente, si è rivelato il migliore per la classificazione dei dati nell'analisi.

Un *decision tree* è un classificatore binario di cui si propone uno schema nella figura 3.2. Su una singola variabile sono prese decisioni del tipo yes/no fino all'ultima decisione presa rispetto a un predefinito criterio di stop. Quindi l'intero spazio viene suddiviso in regioni che saranno classificate come di segnale o di fondo, rispetto alla maggioranza degli eventi che finiscono nei nodi finali, che prendono il nome di *leaf*. Il *boosting decision tree* estende lo schema descritto da uno a più tree, formando una *forest* (foresta). I tree sono costruiti a partire dallo stesso insieme di dati di training ripesando gli eventi e convogliano verso un singolo classificatore che è dato dalla media pesata dei decision tree.

I decision tree sono classificatori che permettono una semplice interpretazione, poiché sono visualizzabili come una struttura di albero bidimensionale. Per un decision tree il percorso dal nodo root a ogni nodo leaf costituisce una sequenza individuale di tagli che permette di selezionare il segnale o il fondo rispetto al tipo di nodo leaf.

Una limitazione dei decision tree risiede nell'instabilità rispetto alle fluttuazioni statistiche nel campione di training da cui la struttura è derivata. Per esempio, se compaiono due variabili caratterizzate da una simile capacità di discriminazione, una fluttuazione sul campione di training potrebbe portare l'algoritmo che costruisce il tree alla decisione di dividere gli eventi rispetto a una variabile, mentre sarebbe selezionata l'altra senza la fluttuazione. Questo causerebbe un'alterazione

dell'intero percorso, con conseguente instabilità nella risposta del classificatore. Il boosting sopracitato è introdotto per ovviare a questo problema, per cui si costruisce una forest di decision tree e si classifica un evento sul valore della maggioranza delle classificazioni effettuate da ogni tree della forest. Il boosting migliora la performance del metodo di discriminazione rispetto a un singolo tree, anche se complica notevolmente l'interpretazione di come questo lavora.

**BDTG** Questo metodo si basa sul BDT, con il miglioramento dato dal *gradient boosting*. Consideriamo la funzione di discriminazione  $F(\mathbf{x})$  come la somma pesata di funzioni di base parametriche  $f(x; a_m)$ . Ogni funzione di base costituisce un decision tree in questo sviluppo:

$$F(\mathbf{x}, P) = \sum_{m=0}^M \beta_m f(x; a_m) \quad \text{con } P \in [\beta_m; a_m] \quad m = 0, \dots, M$$

Il boosting interviene per adattare  $P$  in modo che sia minimizzata la deviazione tra il valore di  $F(\mathbf{x})$  e il valore vero  $y$  ottenuto dal campione di training. La deviazione è determinata mediante la funzione  $L(F, y)$  detta *loss-function*, che può essere di vario tipo e determina completamente la procedura di boosting. La più semplice loss-function è:

$$L(F, y) = (F(\mathbf{x}) - y)^2$$

Nel gradient boosting la loss-function ha la seguente espressione:

$$L(F, y) = \ln\left(1 + e^{-2F(\mathbf{x})y}\right)$$

Poiché l'algoritmo scelto è complicato si procede per passi. Si calcola il gradiente della loss-function e quindi si costruisce un tree in cui i valori dei nodi leaf sono adattati in modo da convergere verso il valore medio del gradiente in ogni regione definita dalla struttura del tree. L'iterazione di questa procedura conduce all'insieme di tree desiderato, che minimizza la loss-function. Il BDTG rende più robusto il BDT perché diventano meno influenti i dati specifici utilizzati nel training.

## Capitolo 4

# Applicazione dell'analisi mutivariata ad un processo di nuova fisica ad ATLAS

È stato effettuato uno studio di fattibilità di un'analisi che utilizza le tecniche descritte nel capitolo 3 per identificare un processo di nuova fisica in ATLAS: la produzione diretta di coppie di stop e il loro decadimento, ciascuno in un top e un neutralino. Gli eventi di segnale e di fondo sono stati generati con opportuni generatori Monte Carlo, processati attraverso la simulazione dell'esperimento e prefiltrati con un'analisi standard (cut-based) in cui si sono richiesti due soli leptoni di segno opposto (*OS, opposite sign*) nell'evento e di ottima qualità (ossia rispondenti a richieste particolarmente esigenti per definire elettroni e muoni in ATLAS), con momenti trasversi  $> (25, 20)$  GeV per gli elettroni e  $> (20, 10)$  GeV per i muoni, masse invarianti  $> 20$  GeV. I due leptoni possono essere sia dello stesso flavour (*SF, same flavour*) che di flavour diversi (*OF, opposite flavour*). Sono stati considerati solo eventi triggerati con elettroni o con muoni.

In particolare, in questo lavoro di tesi lo studio si è concentrato sul canale muonico, dando particolare rilevanza alla scelta del punto di segnale che ottimizza la significanza dell'analisi e il numero di eventi di segnale ottenibili.

### 4.1 Gli eventi di segnale e di fondo

Sono considerati processi di produzione diretta di coppie di stop via interazioni forti all'LHC, in cui li stop decadono in:

$$\tilde{t}\tilde{t} \rightarrow \tilde{\chi}_1^0 t + \tilde{\chi}_1^0 \bar{t} \rightarrow \tilde{\chi}_1^0 l^+ \nu_l + \tilde{\chi}_1^0 l^- \bar{\nu}_l \quad (4.1)$$

per cui risulta evidente che si debba avere  $m_{\tilde{t}} - m_{\tilde{\chi}_1^0} > m_t$ .

Il decadimento (4.1) è rappresentato nella figura 4.1, nel caso che il processo di produzione avvenga attraverso un vertice a tre gluoni.

Come già accennato, in questo lavoro di tesi si sono considerati i decadimenti in cui gli stati finali leptonici sono stati triggerati come muoni, accompagnati dai neutrini corrispondenti (eventi  $\mu^+ \mu^-$  e  $\mu^\pm e^\mp$  che non hanno soddisfatto il trigger di elettroni; questi ultimi costituiscono in realtà una piccola percentuale,  $\sim 1\%$  dei

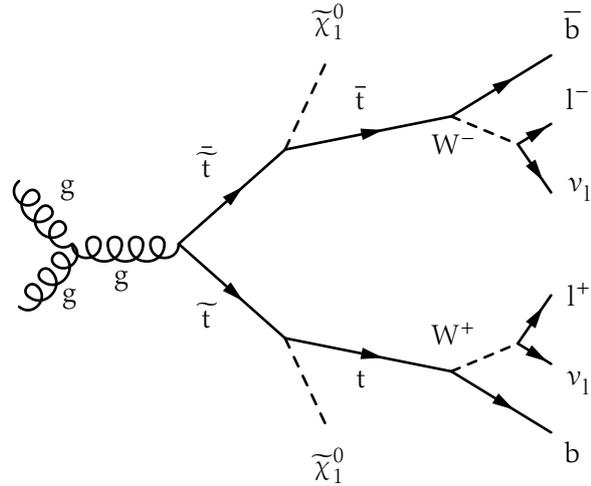


Figura 4.1: Rappresentazione del decadimento considerato

primi). Gli eventi di  $t\bar{t}$ , top singolo, coppie di bosoni, bosoni accompagnati da jet, processi Drell-Yan costituiscono un enorme fondo alla produzione di stop.

Di seguito sono elencati i processi e i corrispondenti programmi di generazione impiegati per ottenere i campioni Monte Carlo[13]:

- eventi di segnale (Herwig), per una griglia di valori nel piano  $m_{\tilde{t}} - m_{\tilde{\chi}_1^0}$  con masse dello stop fino a 800 GeV e masse del neutralino fino a 550 GeV, con passi di 50 GeV;
- $t\bar{t}$  con almeno un leptone carico nello stato finale (MC@NLO);
- di-bosoni Z Z, W Z, W W (Alpgen + Herwig);
- $Z \rightarrow l^+ l^-$  e processo Drell - Yan (DY) con uno o più jet nello stato finale e  $l = e, \mu, \tau$  (Alpgen + Jimmy);
- $W^\pm \rightarrow l^\pm \nu_l$  con uno o più jet nello stato finale e  $l = e, \mu, \tau$  (Alpgen + Herwig);
- top singolo (MC@NLO).

## 4.2 Le variabili di input

Come approccio preliminare all'analisi multivariata sono state selezionate le variabili di input tra quelle normalmente utilizzate per l'analisi standard dei processi di supersimmetria[13].

$E_T^{\text{miss}}$

L'energia trasversa mancante, anche indicata con *met*, è una variabile che identifica la presenza di particelle che non interagiscono con i rivelatori, come i neutrini, nel caso delle SUSY dei neutralini.

$M_{\text{eff}}$

Consiste nella massa effettiva dell'evento, data da:

$$M_{\text{eff}} = \sum_{i=1}^2 |p_T(\text{jet})_i| + E_T^{\text{miss}} + p_T^{\text{Lep1}} + p_T^{\text{Lep2}}$$

e permette di valutare la scala di massa dei processi in esame. La nuova fisica si suppone essere caratterizzata da elevati valori di  $M_{\text{eff}}$ .

$p_T^{\text{Lep2}}$

È il momento trasverso del secondo leptone nell'evento, ovvero del meno energetico tra i due.

$p_T^{\text{Jet2}}$

È il momento trasverso del secondo jet nell'evento, supponendo che i jet siano ordinati per valori decrescenti del momento trasverso. Nel caso esso non esista la variabile assume valore zero.

$M_{T2}$

Questa variabile è stata introdotta nello studio di eventi supersimmetrici per i decadimenti in cui vengono prodotte due particelle identiche ed entrambe decadono in una particella invisibile. Il vettore somma dei momenti trasversi delle due particelle invisibili viene misurato e costituisce la  $E_T^{\text{miss}}$  dell'evento. L'idea è di considerare tutte le possibili combinazioni di  $E_T^{\text{miss}}$  in due vettori trasversi che vengono interpretati come i momenti trasversi delle particelle invisibili nei due prodotti finali. Dai momenti trasversi sono ricostruite due masse trasverse con le particelle visibili e viene assunto come valore della variabile  $M_{T2}$  il massimo tra questi due valori. In generale si ha che  $M_{T2} < M_W$  nei processi del Modello Standard, e  $M_{T2} > M_W$  per i processi come quello in esame, per cui questa variabile risulta essere molto potente[13].

$M_{ll}$

È la massa invariante dei due leptoni.

### 4.3 Il Training

La prima fase dell'analisi multivariata consiste nell'applicare il training come descritto nel paragrafo 3.2. L'operazione è stata effettuata su diversi punti di segnale della griglia e sul fondo <sup>1</sup>, utilizzando i vari metodi di analisi multivariata descritti nel capitolo 3. Uno studio parallelo a quello proposto in questa tesi, focalizzato

<sup>1</sup>Con il termine fondo si intenderà d'ora in poi la somma di tutte le sorgenti di background descritte nel paragrafo 4.1, ognuna pesata con pesi opportuni determinati a partire dalla sezione d'urto di ogni processo e dalla luminosità integrata a cui ci si normalizza nell'analisi ( $5 \text{ fb}^{-1}$ ).

sul decadimento dello stop, ma nel canale di elettroni, ha portato a individuare nel metodo BDTG il migliore metodo di analisi per discriminare il segnale dal fondo nel problema in studio.

Sono stati selezionati, tra i diversi punti di segnale, quattro punti ottimali su cui operare il training, denominati con codici e riferiti a diversi punti sulla griglia di segnale:

- dataset 127848, che corrisponde a  $m_{\tilde{\tau}} \simeq 350$  GeV e  $m_{\tilde{\chi}_1^0} \simeq 0$  GeV;
- dataset 127849, che corrisponde a  $m_{\tilde{\tau}} \simeq 300$  GeV e  $m_{\tilde{\chi}_1^0} \simeq 50$  GeV;
- dataset 127850, che corrisponde a  $m_{\tilde{\tau}} \simeq 250$  GeV e  $m_{\tilde{\chi}_1^0} \simeq 50$  GeV;
- dataset 127851, che corrisponde a  $m_{\tilde{\tau}} \simeq 250$  GeV e  $m_{\tilde{\chi}_1^0} \simeq 0$  GeV.

Il processo di training sia sul segnale che sul fondo è stato effettuato dopo aver applicato dei tagli preliminari su alcune delle variabili di input. Questi tagli preliminari sono stati determinati ottimizzando sia i risultati del training che il tempo impiegato da TMVA per elaborare i risultati. In generale questi tagli esaltano il rapporto segnale-fondo, non introducendo distorsioni sul segnale. Per i muoni è stato inizialmente considerato il taglio denominato  $cut_5$ , che corrisponde a un taglio sulla massa efficace  $M_{eff}$  e a un taglio sull'energia trasversa mancante  $E_T^{miss}$ . Il taglio è il seguente:

$$cut_5 = M_{eff} > 300\text{GeV} \ \& \ E_T^{miss} > 50\text{GeV}$$

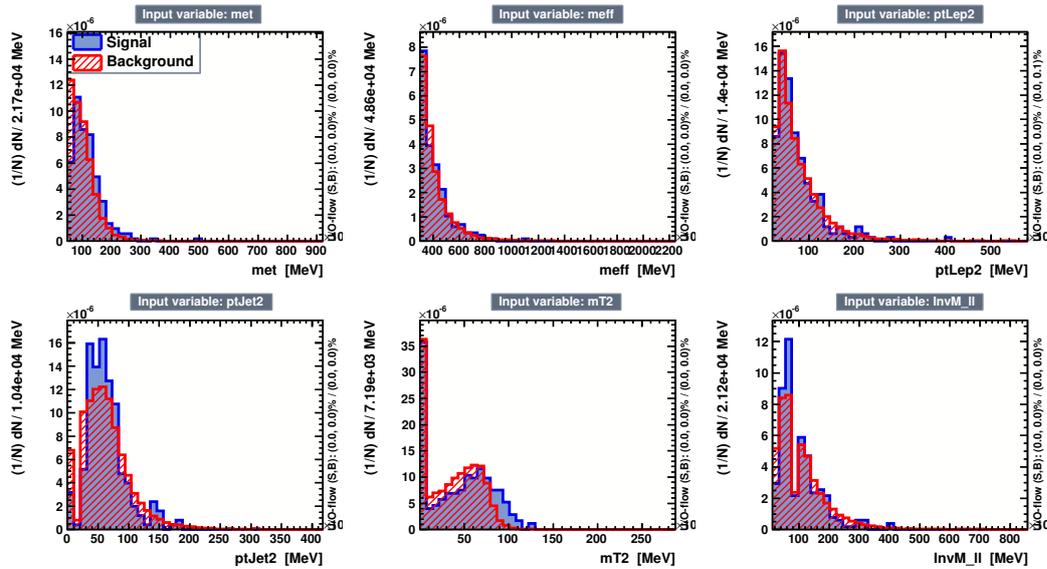
Mentre questo taglio si è rivelato particolarmente efficiente nel caso degli elettroni, i risultati ottenuti sui muoni non sono stati soddisfacenti, per cui si è operato il processo di training sugli stessi dati di segnale e fondo, sottoponendoli al taglio denominato  $cut_{10}$ , che comprende anche un taglio sulla massa invariante dei leptoni, in modo da abbattere significativamente il fondo di  $Z \rightarrow \mu^+\mu^-$  [14]:

$$cut_{10} = M_{eff} > 300\text{GeV} \ \& \ E_T^{miss} > 50\text{GeV} \ \& \ |M_{ll} - 91 \text{ GeV}| > 10\text{GeV}$$

Dal processo di training si sono ottenuti dei file in formato root e xml. Il file root permette all'utente di visualizzare i risultati ottenuti dal training per valutare il metodo migliore nei termini della reiezione del fondo e dell'efficienza sul segnale, la bontà delle variabili utilizzate, le correlazioni tra queste. Con la macro `TMVAGui.C` è possibile visualizzare i grafici ottenuti a partire dai dati processati. Nella figura 4.2 sono rappresentate le sei variabili di input, nel caso del segnale (in blu, punto 127848) e del fondo (in rosso). Le distribuzioni sono normalizzate al numero totale di eventi. Dalla figura si osserva che la variabile che meglio discrimina tra segnale e fondo è la  $M_{T2}$ .

Nella figura 4.3 sono riportati i grafici che illustrano, in funzione del taglio sull'estimatore BDTG:

- le efficienze sul segnale e sul fondo;
- la significanza, definita come nel paragrafo 3.2;
- la purezza del segnale, definita come  $S/(S + B)$ ;



**Figura 4.2:** Distribuzioni delle variabili di input per il training con il punto di segnale 127848 e taglio  $cut_{10}$ , normalizzate al numero totale di eventi. Il segnale è in blu e il fondo è in rosso.

- il prodotto di efficienza e purezza del segnale.

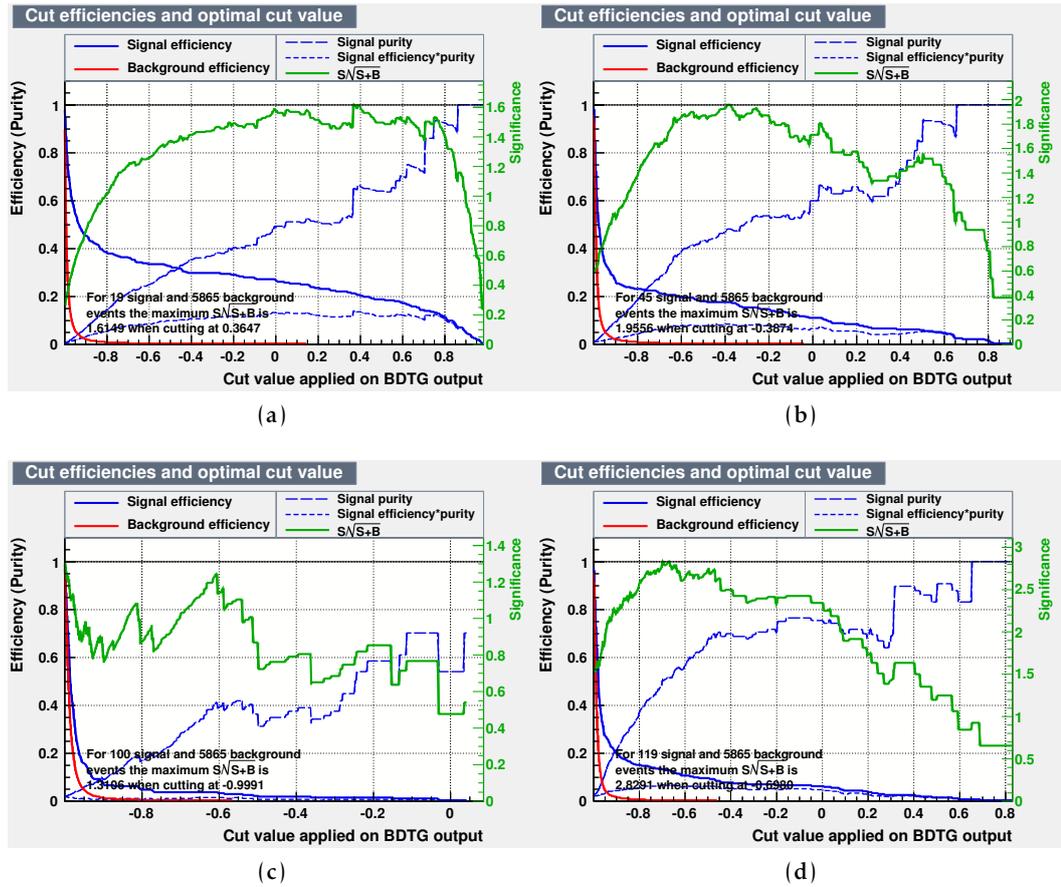
risultati dal training con il metodo BDTG per i quattro differenti punti di segnale descritti.

TMVA consente di valutare la correlazione tra le variabili di input scelte, per valutarne il potere discriminante. La macro `TMVAGui.C` permette di visualizzare una matrice in cui sono rappresentate le correlazioni lineari tra le variabili. Per i quattro punti di segnali scelti si ottengono circa le stesse matrici di correlazione lineare rispetto alle variabili sia per il segnale che per il fondo. Quelle per il training con il punto 127848 sono riportate come esempio nella figura 4.4.

Nella figura 4.5 sono riportati i grafici che illustrano la reiezione del fondo in funzione dell'efficienza sul segnale per i training effettuati con il fondo e i punti di segnale 127848, 127849, 127850, 127851. Si osserva che la migliore curva si ottiene per il punto di segnale 127848 (a).

## 4.4 L' Application

Dopo la fase di training ogni punto di segnale della griglia viene processato insieme a tutto il fondo per discriminarlo da esso e valutare la potenza del metodo e del training utilizzati. A tale scopo la macro `TMVAClassificationApplication.C` viene opportunamente modificata introducendo il metodo di analisi multivariata scelto e il percorso dei file da processare. La macro prende come input il training scelto, attraverso il relativo file xml dei pesi, e i vari punti di segnale e di fondo di cui si considera non nota a priori la classificazione. Per ogni punto del segnale e per il fondo TMVA crea un file root, che viene utilizzato per la successiva fase di analisi e che contiene gli istogrammi con il valore dell'estimatore (nel nostro caso BDTG)



**Figura 4.3:** Efficienza sul segnale e sul fondo e valore della significanza per i training effettuati con il fondo e i punti di segnale 127848 (a) 127849 (b) 127850 (c) 127581 (d) con taglio preliminare  $cut_{10}$ , in funzione del taglio applicato all'estimatore BDTG. Sui grafici è indicato il valore del taglio che massimizza la significanza.

che viene assegnato, a partire dal training scelto, a ciascun evento di segnale o di fondo processati. Per semplificare l'analisi di questi istogrammi è stata utilizzata la macro `findTcut.C`, che ha la funzione di individuare il taglio ottimale da apportare all'estimatore per massimizzare la significanza. La macro `findTcut.C` ha tre argomenti di input: il file root derivante dall'application sul punto di segnale di cui si vuole massimizzare la significanza, il file root creato dall'application sul fondo e il taglio preliminare scelto ( $cut_5$  o  $cut_{10}$  nel nostro caso). Sostanzialmente essa accede ai valori degli istogrammi degli estimatori per il segnale e per il fondo e calcola per quale valore del taglio sull'estimatore la significanza, definita come il rapporto  $S/\sqrt{B + (\delta B)^2}$ , è massimizzata, dove  $S$  e  $B$  sono rispettivamente i valori del segnale e del fondo che sopravvivono al taglio e  $\delta B$  è un termine aggiunto per tener conto dell'incertezza sistematica sul fondo, assunta essere del 60%.

L'output di `findTcut.C` consiste infine in file log contenenti i valori della significanza, il suo massimo e i corrispondenti valori dei tagli da apportare sull'estimatore. Per accelerare la procedura è utilizzato il programma `loopOnSamples.csh`, che permette di girare l'application e `findTcut.C` su tutti i dati della griglia di

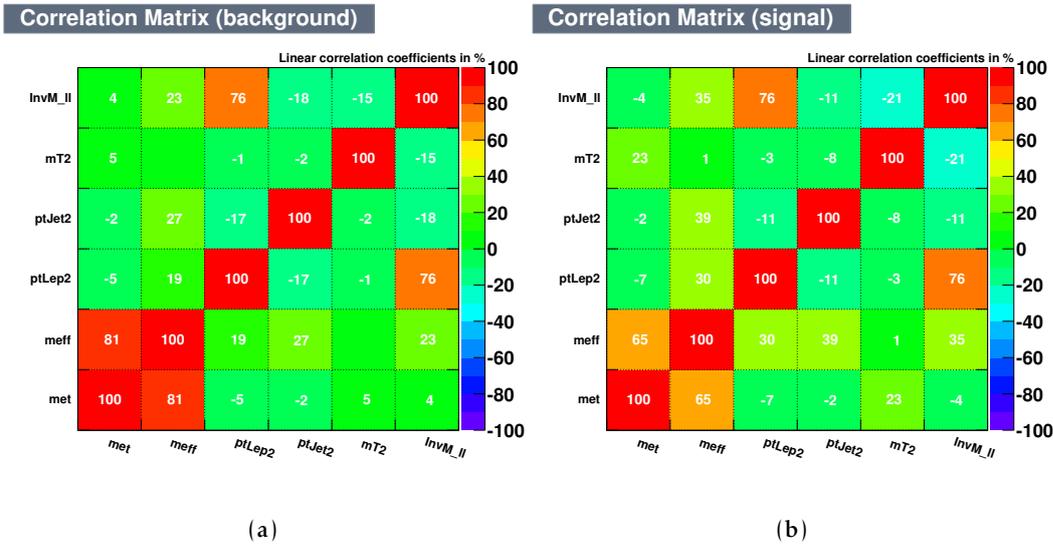


Figura 4.4: Matrici dei coefficienti di correlazione lineare tra le variabili di input per il fondo (a) e il segnale (b), ottenuti per il training con il punto 127848 con  $cut_{10}$ . Si osserva che il coefficiente di correlazione lineare tra  $M_{eff}$  e  $E_T^{miss}$  è alto, a causa della definizione di  $M_{eff}$ , che include  $E_T^{miss}$ . Lo stesso vale per il coefficiente di correlazione tra  $p_T^{Lep2}$  e  $M_{ll}$ .

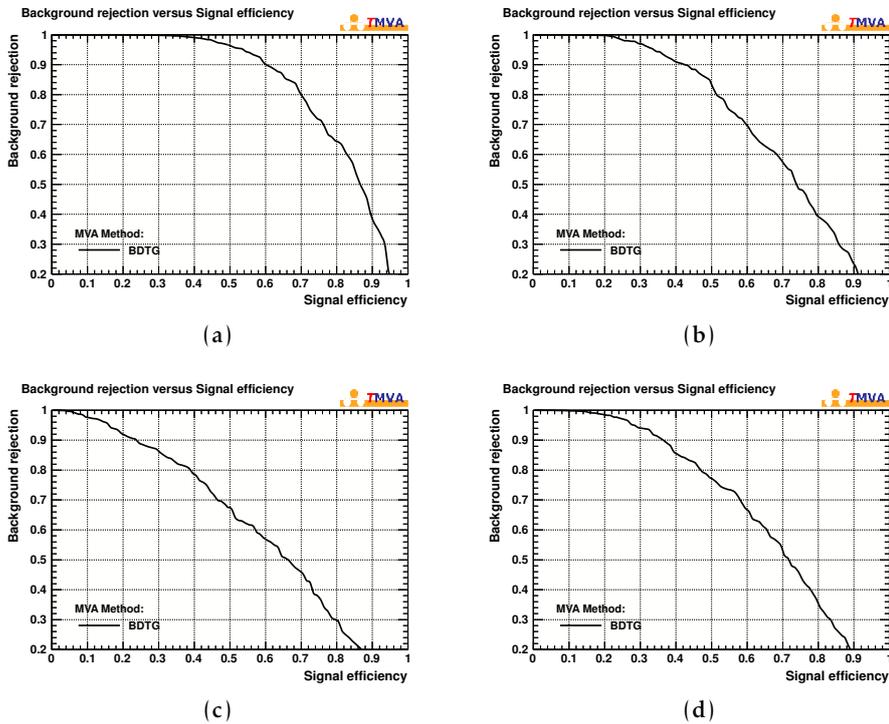


Figura 4.5: Reiezione del fondo in funzione dell'efficienza sul segnale per i training effettuati con il fondo e i punti di segnale 127848 (a) 127849 (b) 127850 (c) 127581 (d) con taglio preliminare  $cut_{10}$ .

segnale e sul fondo.

In `findTcut.C`, per il calcolo del valore del taglio ottimale, sono stati introdotti: un controllo sul valore del massimo della significanza in modo da tener conto anche dell'errore ad essa associato; un controllo sul numero degli eventi di fondo originario Monte Carlo che entrano nella definizione della significanza per stimare un valore ragionevole su questa. Questo controllo impone la richiesta di almeno due eventi di fondo, per poter considerare la stima della significanza attendibile e quindi eleggere il taglio corretto sull'estimatore.

Infine i risultati sono visualizzati con la macro `makeSgnPlot.C`, che ha come input i log file prodotti precedentemente per ogni punto della griglia di segnale e il valore del taglio che massimizza la significanza, determinato da `findTcut.C`, per il punto di segnale di cui si è utilizzato il training nell'application.

## 4.5 Risultati ottenuti

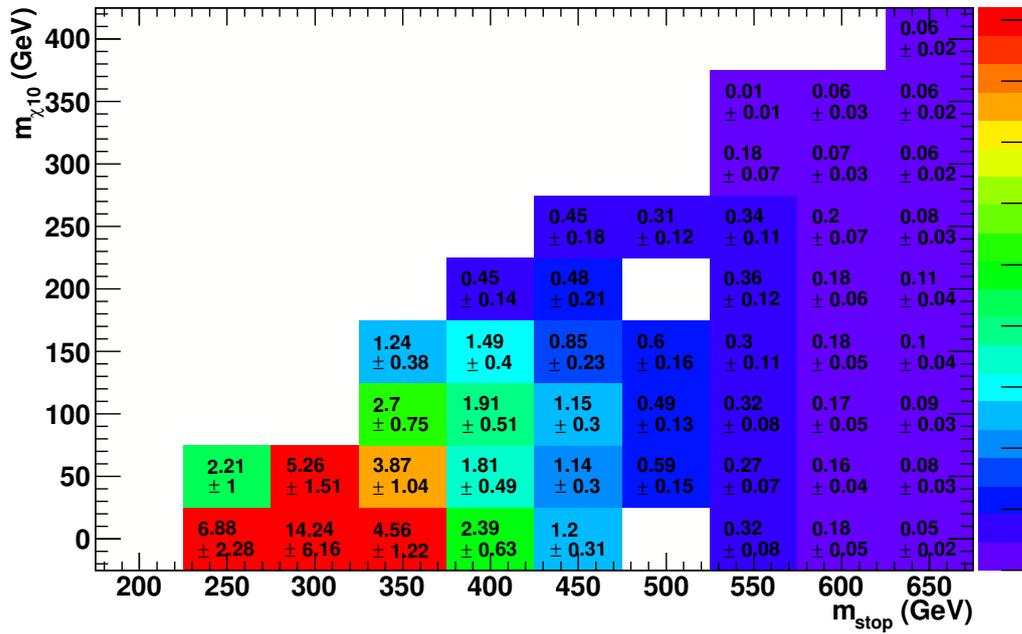
Dalla macro `makeSgnPlot.C` si ottengono due grafici, relativi alla griglia di segnale utilizzata, in cui i punti sono individuati dal valore della massa dello stop (in ascissa) e della massa del neutralino (in ordinata) prodotto nel decadimento. Nel primo nelle celle che rappresentano i diversi punti di segnale viene stampato il valore della significanza e dell'errore ad essa associato, nel secondo gli eventi di segnale da cui si è ottenuta quella significanza. Entrambi i grafici sono ottenuti con la condizione che la statistica sia normalizzata a un valore di  $5 \text{ fb}^{-1}$ , che corrisponde al valore della statistica dei dati di ATLAS raccolta nel corso del 2011.

Nelle figure dalla 4.6 alla 4.9 sono riportati i grafici della significanza e degli eventi di segnale determinati per il training eseguito sui quattro punti di segnale definiti nel paragrafo 4.3. I grafici sono facilmente interpretabili grazie alla scala di colori che li affianca e che permette di individuare i valori più alti della significanza o degli eventi di segnale.

Dal confronto dei grafici che rappresentano la significanza risulta evidente che il migliore punto di segnale su cui ottimizzare il training per il canale muonico è il 127848, in corrispondenza del quale si osservano sulla griglia di segnale vari punti con alta significanza (anche considerando l'elevata incertezza su questi numeri). Inoltre si evince che la regione di massa per cui è effettuabile questa analisi multivariata di produzione diretta dello stop con possibilità di scoperta con  $5 \text{ fb}^{-1}$  è limitata a  $m_{\tilde{t}} \leq 350 \text{ GeV}$  e  $m_{\tilde{\chi}_1^0} \leq 50 \text{ GeV}$ , in quanto per valori più alti delle masse le sezioni d'urto dei processi sono troppo piccole per renderli accessibili alla nostra sensibilità.

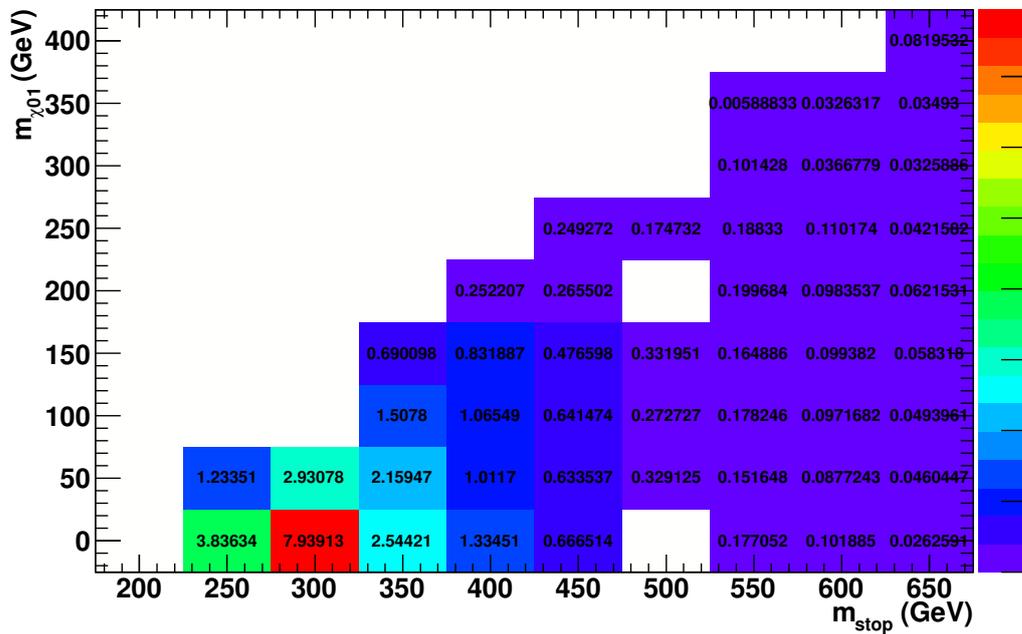
I grafici discussi finora sono stati determinati a partire dai punti di segnale e del fondo, su cui si è applicato un taglio preliminare  $\text{cut}_{10}$  e, quindi, il taglio sul valore dell'estimatore dato da `findTcut.C`. L'opportunità di utilizzare il  $\text{cut}_{10}$  rispetto al  $\text{cut}_5$  iniziale si evince dal confronto delle significanze determinate a partire dallo stesso training e con i due differenti tagli preliminari, riportato nella figura 4.10. Da questo confronto discende che la reiezione preliminare di Z con jet operata con il taglio sulla massa invariante di due leptoni  $M_{ll}$  è utile per ridurre drasticamente questo fondo, rendendo il lavoro di TMVA molto più semplice sia nella fase di training che nella fase di application.

**Significance for  $5 \text{ fb}^{-1}$  with cut at 0.456**



(a)

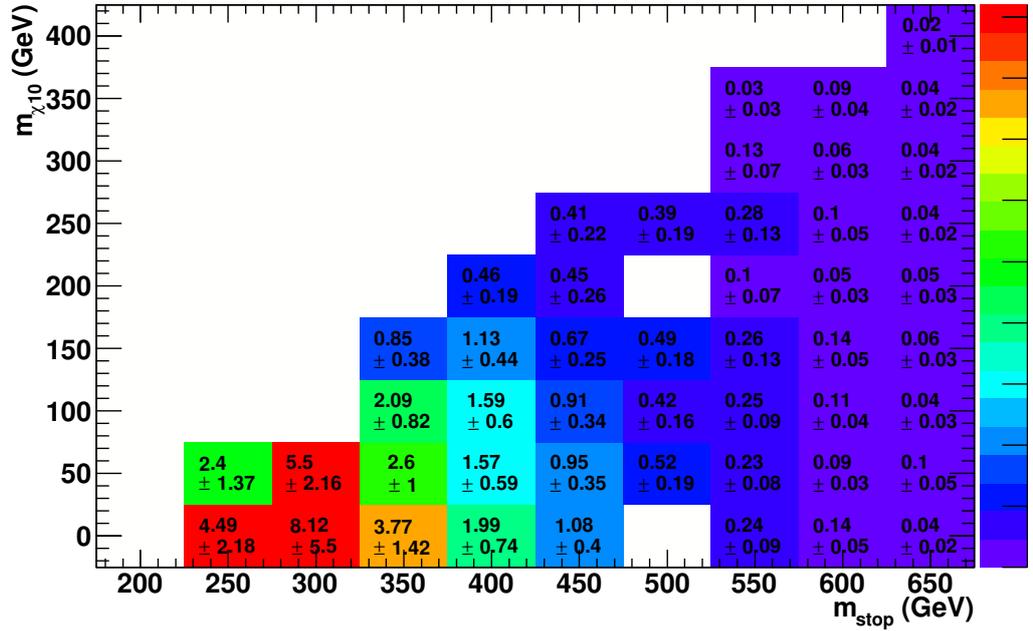
**Signal events for  $5 \text{ fb}^{-1}$  with cut at 0.456**



(b)

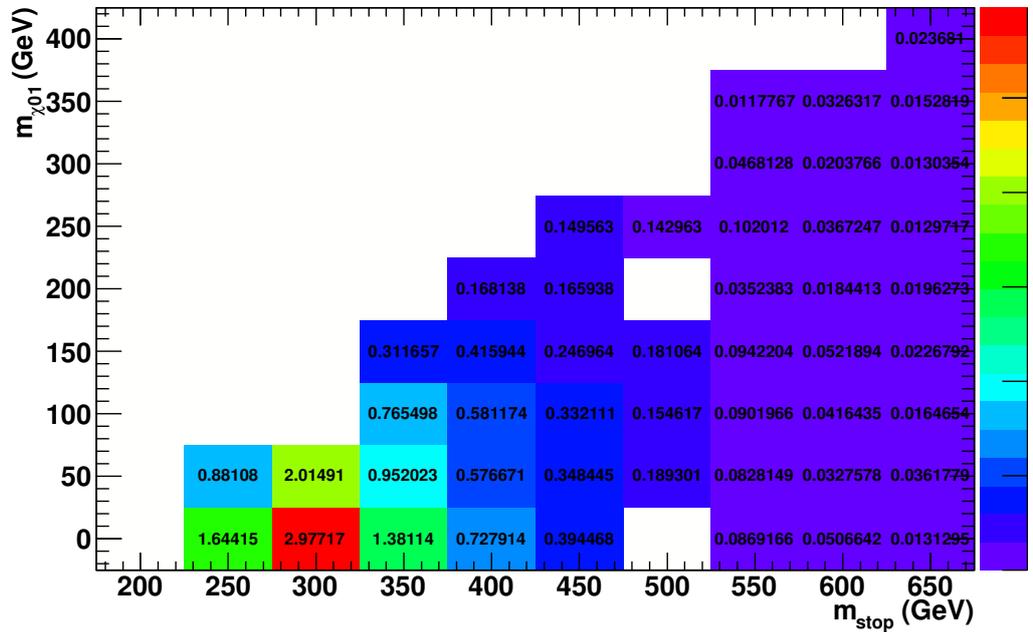
Figura 4.6: Valori della significanza (a) e numero di eventi di segnale (b) determinati con training 127848 con taglio preliminare  $\text{cut}_{10}$  e taglio sull'estimatore (BDTG) a 0.456

**Significance for  $5 \text{ fb}^{-1}$  with cut at 0.508**



(a)

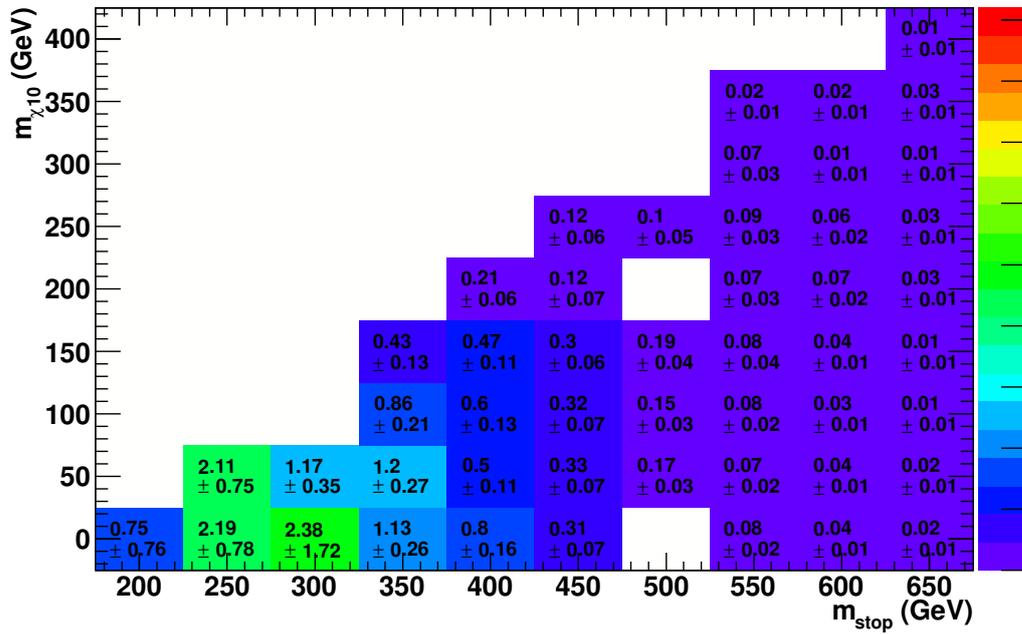
**Signal events for  $5 \text{ fb}^{-1}$  with cut at 0.508**



(b)

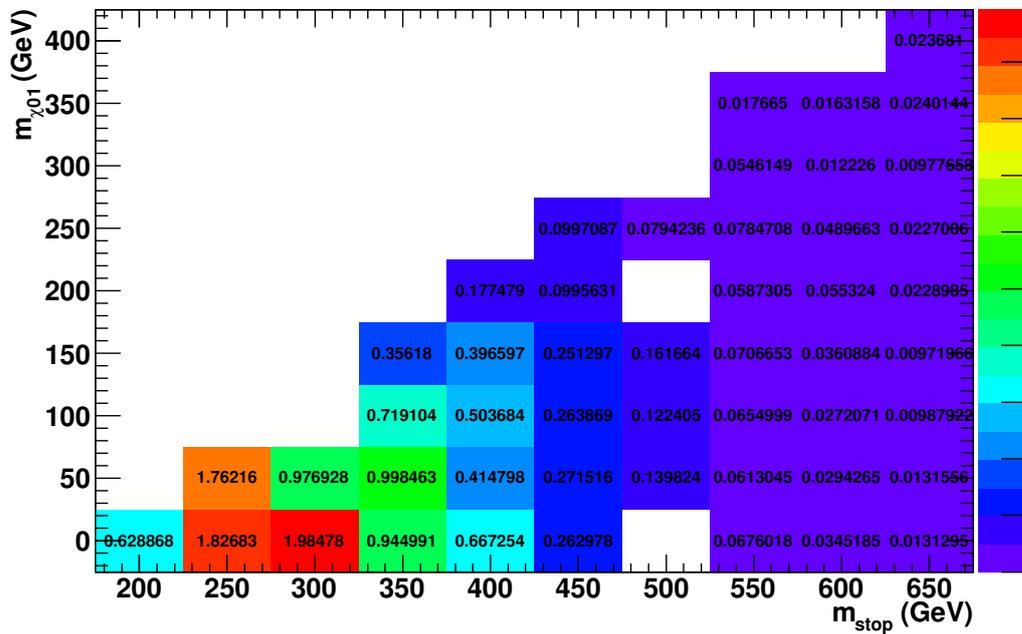
Figura 4.7: Valori della significanza (a) e numero di eventi di segnale (b) determinati con training 127849 con taglio preliminare  $\text{cut}_{10}$  e taglio sull'estimatore (BDTG) a 0.508

Significance for  $5 \text{ fb}^{-1}$  with cut at  $-0.306$



(a)

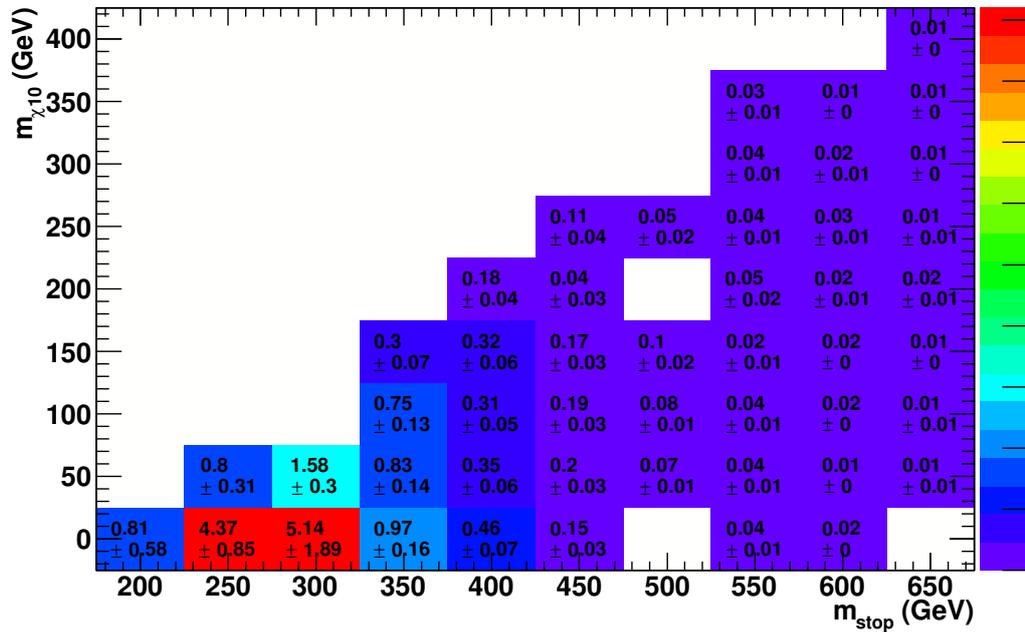
Signal events for  $5 \text{ fb}^{-1}$  with cut at  $-0.306$



(b)

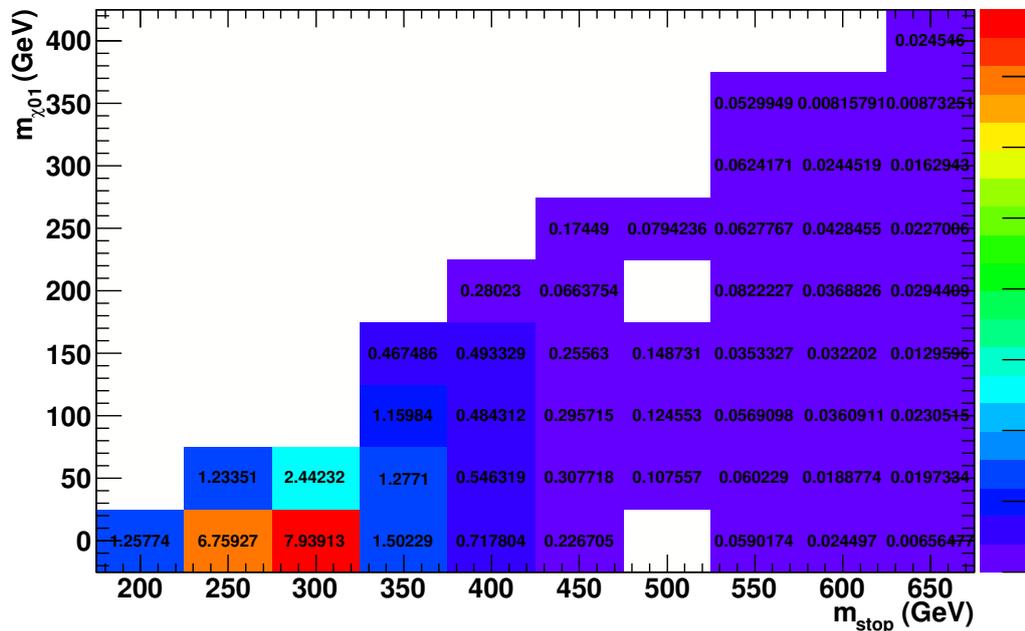
Figura 4.8: Valori della significanza (a) e numero di eventi di segnale (b) determinati con training 127850 con taglio preliminare  $\text{cut}_{10}$  e taglio sull'estimatore (BDTG) a  $-0.306$

**Significance for  $5 \text{ fb}^{-1}$  with cut at  $-0.094$**



(a)

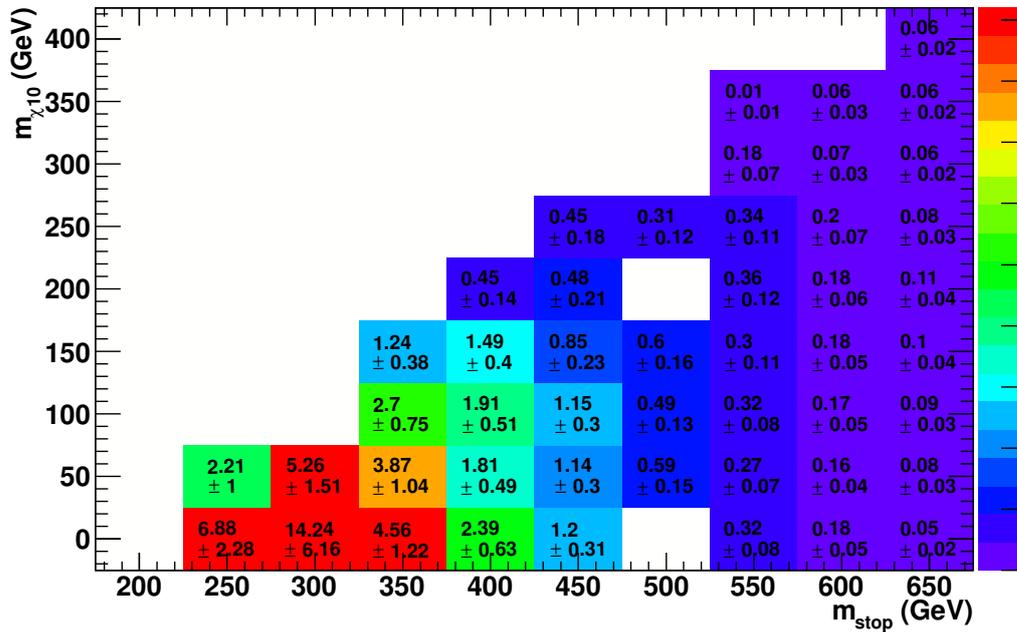
**Signal events for  $5 \text{ fb}^{-1}$  with cut at  $-0.094$**



(b)

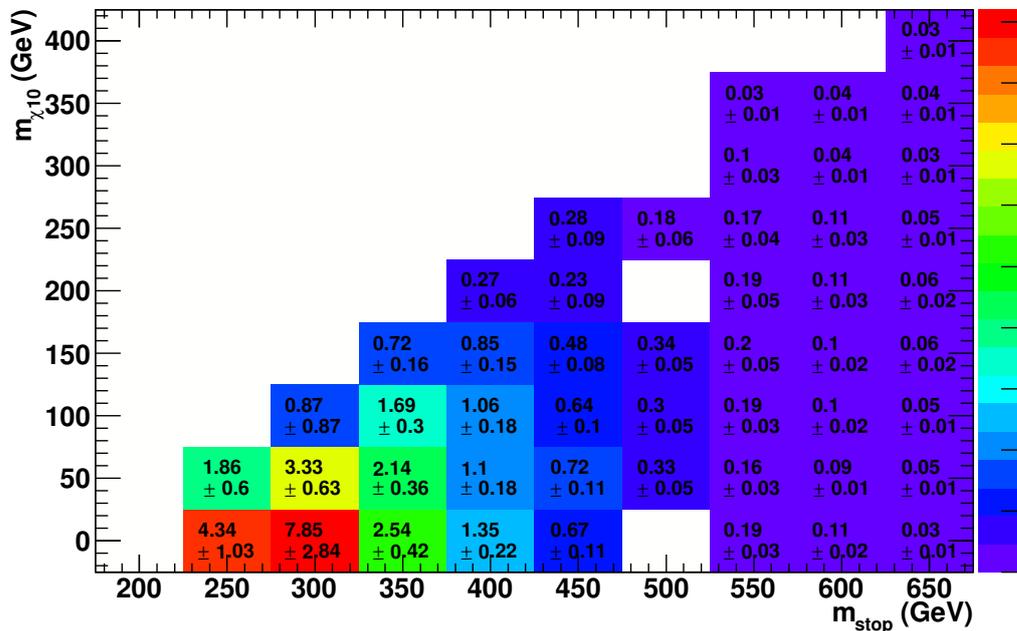
Figura 4.9: Valori della significanza (a) e numero di eventi di segnale (b) determinati con training 127850 con taglio preliminare  $\text{cut}_{10}$  e taglio sull'estimatore (BDTG) a  $-0.094$

Significance for  $5 \text{ fb}^{-1}$  with cut at 0.456



(a)

Significance for  $5 \text{ fb}^{-1}$  with cut at 0.258

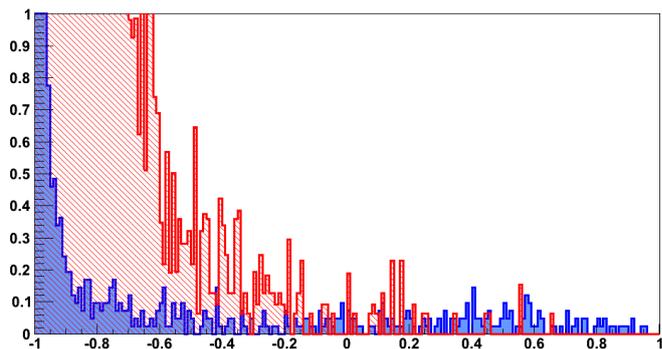


(b)

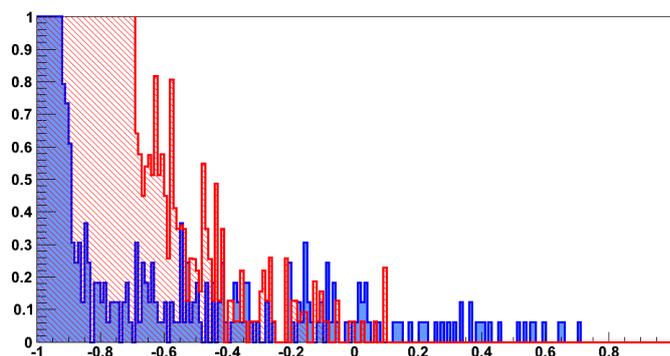
Figura 4.10: Valori della significanza ottenuti con il training 127848 con  $\text{cut}_{10}$  (a) e con  $\text{cut}_5$  (b). Dal confronto risulta evidente che il taglio preliminare migliore è il  $\text{cut}_{10}$ .

È interessante notare la differenza nel potere discriminante tra segnale e fondo rispetto al training con punti di segnale diversi, come illustrato nella figura 4.11. Nonostante il 127848 sia il migliore punto di segnale per il canale muonico non si osserva una netta distinzione tra il segnale e il fondo.

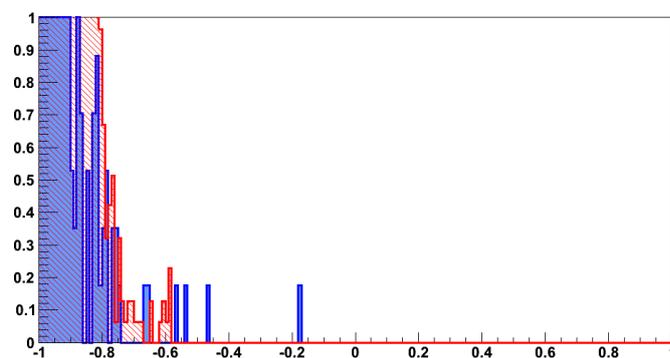
Infine lo scopo di utilizzare l'analisi multivariata è quello di migliorare la prestazione dell'analisi convenzionale. Nella figura 4.12 sono confrontate le significanze ottenute per la stessa griglia di segnale con l'analisi multivariata e con un'analisi convenzionale basata sul taglio su  $M_{T2}$  ( $M_{T2} > 120$  GeV) e che considera solo eventi OF. Dal confronto si evince che la significanza nel caso dell'analisi multivariata è migliore e che la sua sensibilità si estende ad un maggior numero di punti sulla griglia di segnale.



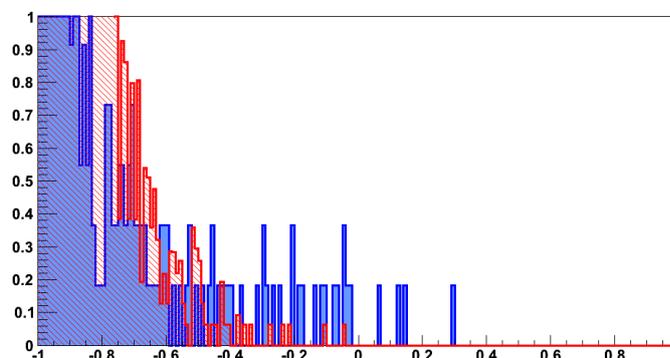
(a)



(b)

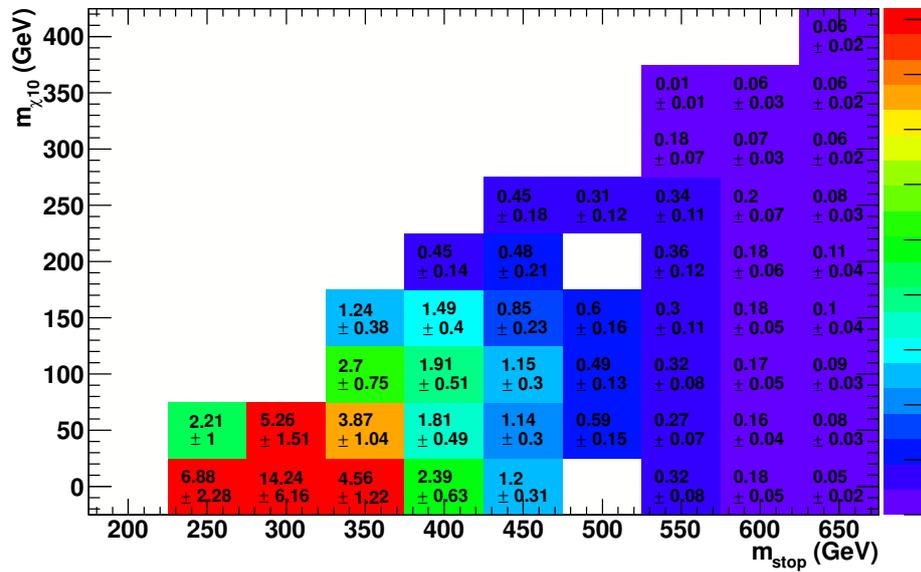


(c)

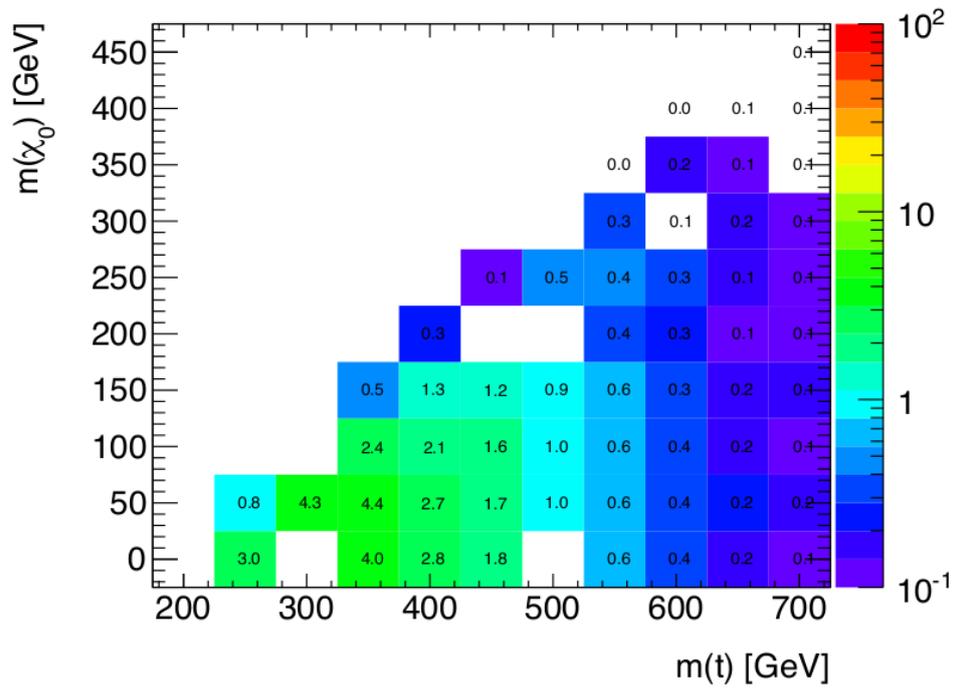


(d)

**Figura 4.11:** Distribuzioni dell'estimatore BDTG per il segnale (in blu) e il fondo (in rosso) per i punti della griglia di segnale 127848 (a), 127849 (b), 127850 (c), 127851 (d). I tagli che massimizzano la significanza sono: 0.456 per il 127848, 0.508 per il 127849, -0.306 per il 127850 e -0.094 per il 127851. Si può dunque osservare che, dopo l'applicazione del taglio ottimizzato sull'estimatore, il 127848 risulta il punto per cui si ha la migliore discriminazione tra segnale e fondo.



(a)



(b)

Figura 4.12: Numero di eventi di segnale con il training 127848 con tecniche di analisi multivariata (a) e con analisi convenzionale (b)[13]. Dal confronto risulta che l'analisi multivariata dà dei risultati migliori rispetto a quella convenzionale.

# Conclusioni

La discriminazione tra segnale e fondo è fondamentale in molti problemi, ma è vitale nel caso degli esperimenti a LHC, che cercano eventi rari di nuova fisica su un fondo che è svariati ordini di grandezza più grande del segnale.

In questo lavoro sono state studiate le prestazioni dell'analisi multivariata nell'individuare eventi rari di Supersimmetria relativi alla produzione diretta e al decadimento di stop in un top e un neutralino, con stati finali a due leptoni, con una statistica Monte Carlo normalizzata a  $5 \text{ fb}^{-1}$ , corrispondente alla statistica dei dati 2011 raccolta dall'esperimento ATLAS.

I risultati ottenuti mostrano che le tecniche di analisi multivariata costituiscono una promettente alternativa all'analisi convenzionale e che le prestazioni sono estremamente incoraggianti. Buoni risultati si ottengono, dopo l'ottimizzazione del metodo e dei training che è stato in parte oggetto di questa tesi, sia nello studio del canale con muoni nello stato finale, che per quello con elettroni, e le significanze sembrano segnalare che siamo sensibili a un'eventuale scoperta in un definito intervallo di massa. Tuttavia le incertezze sono grandi, anche a causa del fatto che la statistica Monte Carlo di fondo è limitata. L'idea è che il metodo possa essere ancora ottimizzato, introducendo nuove variabili confezionate *ad hoc* per sfruttare i vantaggi dell'analisi. Tra i vantaggi dell'usare tecniche di analisi multivariata, infatti, c'è la possibilità di considerare un gran numero di variabili discriminanti, introducendo, ad esempio, ulteriori proprietà geometriche o cinematiche dell'evento di segnale che non appartengono all'evento di fondo.

Un limite concreto e intrinseco nella trattazione e analisi dei dati risiede nella statistica a Monte Carlo a disposizione.



# Riferimenti bibliografici

- [1] B. Povh, K. Rith, C. Sholz e F. Zetsche, *Particelle e nuclei. Un'introduzione ai concetti fisici*. Bollati Boringhieri, 2006.
- [2] G. Kane, *Supersimmetria, squark, fotini, sparticelle: svelare le leggi ultime della natura*. Bollati Boringhieri, 2005.
- [3] URL: <http://cdsweb.cern.ch/record/339593/files/thesis-97-008.pdf>.
- [4] S. Martin, *A Supersymmetry Primer*. In: *Perspectives on Supersymmetry* (G. L. Kane, curatore) 1998, pagina 1. arXiv:hep-ph/9709356.
- [5] L. Randall, *Passaggi curvi*. Il Saggiatore, 2005.
- [6] URL: [http://theophys.kth.se/tepp/research\\_grand\\_unified\\_theories.html](http://theophys.kth.se/tepp/research_grand_unified_theories.html).
- [7] J. Feng, *Supersymmetry and cosmology*. Gen. 2005, *Annals of Physics*, 315, 2–51. DOI: 10.1016/j.aop.2004.09.014. arXiv:hep-ph/0405215.
- [8] Nature Insight 448, *The Large Hadron Collider*. 2007.
- [9] G. Kane e A. Pierce, *Perspective on LHC Physics*. Bollati Boringhieri, 2005.
- [10] ATLAS Collaboration, *The ATLAS experiment at CERN Large Hadron Collider*. 2005, *Journal of Instrumentation*.
- [11] *ROOT Reference Guide*. URL: <http://root.cern.ch/root>.
- [12] A. Hoecker *et al.*, *TMVA - Toolkit for Multivariate Data Analysis*. 9 Nov. 2009, *ArXiv Physics e-prints*. arXiv:physics/0703039.
- [13] M. Bianco, *Search for a heavy top partner in final states with two leptons*. 2012, *ATL-COM-PHYS-2012-022*.
- [14] M. Bianco, E. Gorini, M. Primavera e A. Ventura, *2 leptons in stop  $\rightarrow$  top + LSP with multivariate techniques*. 2011, *SUSY direct stop meeting*. URL: <http://indico.cern.ch/conferenceDisplay.py?confid=16608>.



# Ringraziamenti

Non mi ero resa conto di quanto fosse difficile scrivere due righe di ringraziamenti, ma sono così tante le persone a cui vorrei dire “grazie” che se dovessi ringraziarle come meritano tutte finirei per scrivere un’altra tesi.

Grazie a chi ha permesso che scoprissi la bellezza di studiare la Fisica, spingendomi a “cercare di capire come va il mondo”, a chi non ha mai smesso di darmi il suo marmoreo appoggio e indotto, in ogni occasione, a “stombare sobba”, a chi ha condiviso con me i piaceri della speculazione filosofica in merito alle più disparate questioni, prima fra tutte la *sciarpità*.

Grazie a chi c’è sempre stato e mi ha sostenuto incondizionatamente, a chi è stato sempre pronto ad ascoltare (anche quando la lucidità veniva a mancare, come durante la lezione sulla “minimazione”), a chi “per caso” si è trovato a passare dalla mia vita e ci è rimasto, nonostante la diversità.

Grazie ai miei (secondi) genitori, per non avermi mai lasciata sola e per avermi sempre consigliato e criticato, offrendomi nuovi punti di vista e per aver condiviso con me (quasi) tutta la mia vita.

Grazie a chi ha voluto vivere con me la quotidianità e ha sempre cercato di strapparmi un sorriso, con dolci, musica e opere d’Arte (con la A maiuscola).

Grazie a chi ha condiviso con me tutto questo, le notti insonni, i laboratori, le chiacchierate al Duomo, i *fallimenti*, le felpe, le focacce, le biciclette, i Risiko, la *folla*, gli *scami* e la *goduria* e penso si possa proprio dire che oggi:

$$\lim_{x \rightarrow \infty} G(x) = +\infty$$

Grazie a chi ha cercato di darmi equilibrio, di mettere un po’ di ordine al mio “Casino” e ha voluto condividere ogni cosa, dai maledetti grafici (figura 4.11) alla Musica.

E infine grazie alla persona che mi ha offerto sempre il suo aiuto con la sua estrema generosità, mi ha sottoposto a ramanzine indispensabili, ha apprezzato la mia “s”, mi ha accompagnato dal primo giorno in università, ha vissuto con me ogni cosa e ha sopportato i ritardi, le paranoie e le paure. Nonostante gli arcobaleni, i “tanto questo non te lo chiede” e i circuiti talvolta mal funzionanti, l’unica cosa da dire è infinitamente grazie (per l’*n*-esima volta<sup>2</sup>).

---

<sup>2</sup>Una stima attendibile di *n* è  $n \simeq 10^{10}$ .