

# UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia “Galileo Galilei”  
Corso di Laurea in Fisica

TESI DI LAUREA TRIENNALE

**Caratterizzazione del processo di decadimento del  
bosone di Higgs in due muoni e studio preliminare dei  
dati del Run3 raccolti dall’esperimento CMS**

**Relatrice**

Dott.ssa Mia Tosi

**Correlatore**

Prof. Pierluigi Bortignon

**Laureando**

Michelangelo Boscolo

Anno Accademico 2023/2024



# Abstract

Il bosone di Higgs è stato scoperto dalle collaborazioni ATLAS e CMS nel 2012, e grazie ai dati raccolti dal Run1 ad oggi se ne stanno studiando le proprietà e caratteristiche. L'osservazione statisticamente significativa del decadimento del bosone di Higgs in una coppia di muoni non è ad oggi ancora avvenuta, pertanto l'estrazione del segnale di produzione di Higgs che decade in due muoni con i circa  $250 \text{ fb}^{-1}$  di dati disponibili nel Run3 è uno tra gli obiettivi scientifici più importanti per le collaborazioni ATLAS e CMS nei prossimi anni. In questo progetto, con l'impiego di simulazioni Monte Carlo dedicate, si è caratterizzata la cinematica del processo di decadimento del bosone di Higgs in una coppia di muoni e del processo di Drell-Yan in due muoni, in quanto fondo dominante. Si è definita una strategia di analisi volta ad ottimizzare l'efficienza di segnale e la risoluzione in massa invariante dei due muoni.



# Indice

<b>1</b>	<b>Introduzione</b>	<b>1</b>
<b>2</b>	<b>Apparato Sperimentale</b>	<b>3</b>
2.1	LHC: Large Hadron Collider . . . . .	3
2.2	CMS: Compact Muon Solenoid . . . . .	4
2.3	Sistema di coordinate a CMS . . . . .	6
<b>3</b>	<b>Analisi delle simulazioni</b>	<b>9</b>
3.1	Descrizione del Dataset . . . . .	9
3.2	Modi di produzione del bosone di Higgs . . . . .	10
3.3	Analisi del fondo (Drell-Yan) . . . . .	11
3.3.1	Muoni da $Z$ a Livello di Generatore . . . . .	11
3.3.2	Muoni da $\tau$ a Gen Level . . . . .	15
3.3.3	Muoni da $Z$ a Reco Level . . . . .	16
3.3.4	Selezioni . . . . .	18
3.4	Analisi del segnale (ggF e VBF) . . . . .	23
3.4.1	Selezioni . . . . .	23
3.4.2	Cinematica del segnale . . . . .	25
3.4.3	Confronto Fondo Segnale . . . . .	27
<b>4</b>	<b>Conclusioni</b>	<b>29</b>
	<b>Bibliografia</b>	<b>31</b>
	<b>Ringraziamenti</b>	<b>33</b>



# Capitolo 1

## Introduzione

Il bosone di Higgs (indicato nel seguito con il simbolo  $h$ ) è stato osservato per la prima volta a LHC attraverso collisioni protone-protone ( $p$ - $p$  collisions) dagli esperimenti ATLAS e CMS ([2, 11, 12]) attraverso i suoi decadimenti  $h \rightarrow \gamma\gamma$  e  $h \rightarrow ZZ^* \rightarrow \ell^+ \ell^- \ell'^+ \ell'^-$ , dove le due coppie leptone-antileptone possono essere o un elettrone ed un positrone oppure una coppia di muoni di segno opposto. A partire dal 2012, anno della sua scoperta, si sono iniziate a studiare le sue proprietà: in particolare si sono cercate evidenze di accoppiamenti del bosone di Higgs sia ai bosoni vettori, mediatori delle interazioni, che ai fermioni elementari, costituenti della materia. La misura di questi accoppiamenti permette di testare le previsioni del Modello Standard delle particelle elementari, ed in particolare di confermare come i fermioni e i bosoni massivi acquisiscano massa attraverso il meccanismo di Higgs.

Finora sono stati osservati i decadimenti del bosone di Higgs solamente nelle particelle più pesanti del Modello Standard: per quanto concerne i bosoni di gauge si sono osservati gli accoppiamenti di  $h$  con i mediatori dell'interazione debole, ossia i bosoni  $W$  e  $Z$ <sup>1</sup>. Per ciò che riguarda invece le particelle di materia sono stati osservati gli accoppiamenti con fermioni appartenenti alla terza generazione (terza colonna da sinistra in Fig 1.2), ossia quark beauty e leptone  $\tau$  (tramite decadimenti) e quark top (tramite produzione associata  $t\bar{t}h$ ). Ciò è facilmente spiegato dal fatto che gli accoppiamenti del bosone di Higgs ai fermioni (rispettivamente bosoni) sono proporzionali alla loro massa (al quadrato) e di conseguenza i decadimenti in particelle più pesanti sono favoriti [3].

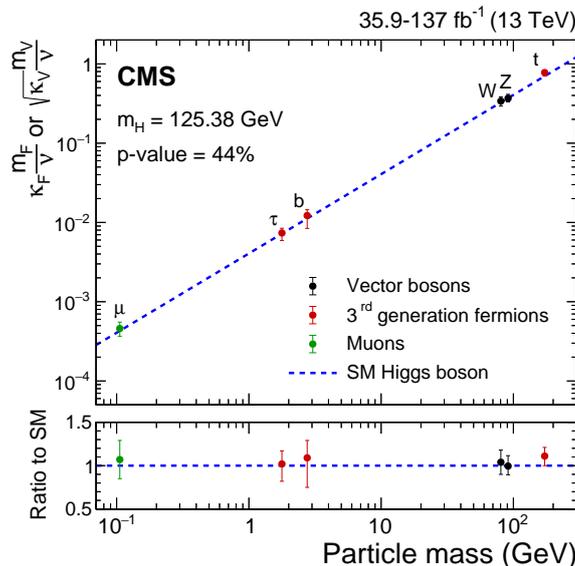


Figura 1.1: Confronto tra previsioni teoriche e risultati sperimentali per gli accoppiamenti del bosone di Higgs con fermioni e bosoni. Immagine da [1].

<sup>1</sup>Il decadimento  $h \rightarrow \gamma\gamma$  è in realtà del tipo  $h \rightarrow W^+ W^- \rightarrow \gamma\gamma$ , dove  $h$  non si accoppia direttamente a  $\gamma$ , bensì ad una coppia di  $W$  (o anche ad una coppia  $t\bar{t}$ ) che in seguito si annichila in due fotoni.

Non si sono ancora osservati tuttavia i decadimenti del bosone di Higgs in leptoni della prima e seconda generazione, e proprio su questi ultimi si sta concentrando la ricerca scientifica. In particolare si sta cercando il segnale del processo di decadimento del bosone di Higgs in una coppia di muoni:  $h \rightarrow \mu^+ \mu^-$ . L'osservazione di questo canale di decadimento è particolarmente difficile, poiché la probabilità attesa è molto bassa:  $\mathcal{B}(h \rightarrow \mu^+ \mu^-) = 2.18 \times 10^{-4}$ . Di recente l'esperimento CMS ha riportato le prime evidenze di osservazione di questo canale di decadimento (Fig. 1.1) [1].

L'obiettivo di questo lavoro di tesi è la ricerca di strategie che permettano di migliorare la significanza del segnale, migliorando la risoluzione della massa invariante nell'analisi di questo canale di decadimento. Ciò verrà fatto studiando le simulazioni del processo  $h \rightarrow \mu^+ \mu^-$ . Tramite un metodo Monte Carlo si generano gli eventi di interesse, a partire dalla collisione iniziale tra i protoni fino al processo che si vuole studiare; in seguito viene riprodotta la presa dati attuata dal rivelatore e infine viene simulata la ricostruzione del processo.

Il seguente lavoro di tesi si sviluppa nel seguente modo: nel Capitolo 2 si dà una breve descrizione del rivelatore CMS, che sta attualmente raccogliendo i dati che in seguito verranno usati per studiare questo decadimento, e del sistema di coordinate in cui i dati raccolti sono descritti. Nel Capitolo 3 si procederà all'analisi delle simulazioni, studiando sia il processo di fondo principale che il segnale, ossia l'effettivo decadimento  $h \rightarrow \mu^+ \mu^-$ . Verranno implementate delle scelte per migliorare il più possibile la risoluzione in massa invariante e infine si farà un confronto tra i risultati trovati per il fondo e quelli trovati per il segnale.

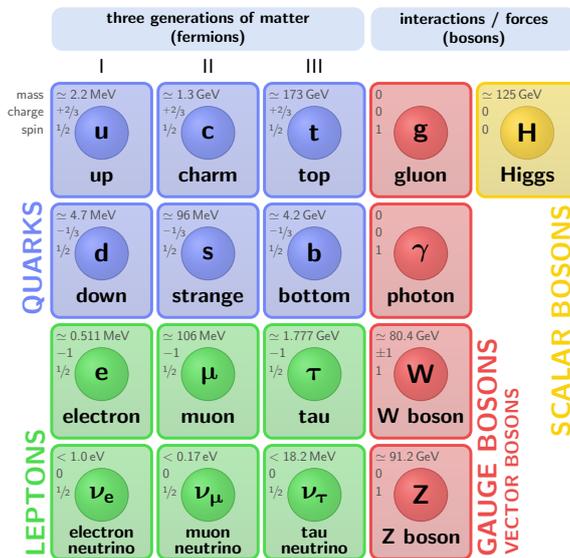


Figura 1.2: Particelle del Modello Standard.

## Capitolo 2

# Apparato Sperimentale

In questo capitolo si dà una breve descrizione delle proprietà dell'acceleratore LHC, del rivelatore CMS e del sistema di coordinate in cui sono descritti i dati raccolti.

### 2.1 LHC: Large Hadron Collider

Il Large Hadron Collider (LHC) è un acceleratore circolare di particelle, deputato alle collisioni fra adroni, in particolare protoni e ioni pesanti. Esso è posto sottoterra nei pressi del CERN a Ginevra ed è costituito da una circonferenza di lunghezza pari a 27 km, all'interno della quale fasci di protoni (o ioni), mantenuti in rotazione da campi magnetici molto intensi, vengono accelerati fino a energie nel sistema del centro di massa pari a  $\sqrt{s} = 13.6$  TeV (dal 2022). I fasci circolano in due anelli separati, l'uno in senso orario e l'altro in senso antiorario, e possono collidere nei quattro punti di intersezione dei due anelli. In corrispondenza di questi si trovano quattro rivelatori, tra cui CMS, come si vede in Figura 2.1.

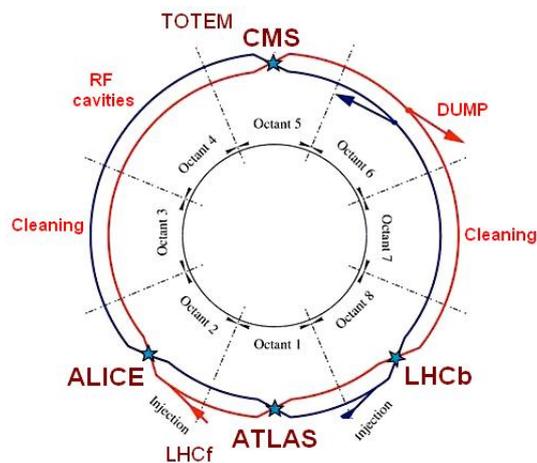


Figura 2.1: L'anello di LHC con i quattro punti di intersezione. Immagine da [https://www.lhc-closer.es/taking\\_a\\_closer\\_look\\_at\\_lhc/0.lhc\\_layout](https://www.lhc-closer.es/taking_a_closer_look_at_lhc/0.lhc_layout).

I fasci di protoni non sono continui ma sono composti da *bunch*, ossia pacchetti di protoni, distanziati l'uno dall'altro da 7.5 m o, equivalentemente, distanziati da un intervallo temporale pari a 25 ns. Questo implica che ogni 25 ns si verifica una collisione. Ogni fascio di protoni contiene fino a un massimo di 2808 bunch, e ogni bunch contiene circa  $10^{11}$  protoni. Ciò si traduce in una frequenza di intersezioni di 40 MHz e in una frequenza di collisioni pari a circa 30 MHz.

Si ha un *evento* quando l'intersezione tra due bunch dà luogo ad almeno una collisione. In realtà poiché all'interno di un bunch la densità di protoni è molto elevata si verificano molte collisioni nella stessa intersezione tra due bunch, gran parte delle quali indesiderate perché non di interesse per il processo che si vuole studiare. Queste collisioni aggiuntive vengono dette *Pile up* (nel 2022 il numero medio di pile up registrato da CMS è stato di 46 per una singola intersezione, come si vede in Fig 2.2).

LHC è stato progettato per assicurare ad ATLAS e CMS (i cosiddetti *high luminosity experiments*) un'elevata luminosità istantanea. In particolare nel 2022 le condizioni di presa dati hanno visto  $\mathcal{L} = 2.0 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ .

I periodi di produzione dei fasci non sono continui, ma hanno una durata dell'ordine delle 10 ore, e sono inframmezzati da intervalli in cui non vengono accelerate particelle. Questi periodi di produzione sono detti *Fill*. La luminosità non rimane costante durante un Fill, ma tende a diminuire esponenzialmente nel tempo a causa dello svuotamento dei fasci [4].

I periodi di produzione dei fasci vengono separati in *Runs*. Quello in corso al momento è il Run3: è iniziato nel 2022 e terminerà nel 2025, e ad esso seguirà un periodo di shutdown per incrementare la luminosità fornita dall'acceleratore. I dati del Run3 saranno utilizzati proprio per cercare il decadimento  $h \rightarrow \mu^+ \mu^-$ . Durante il Run3 LHC collide protoni ad un'energia nel centro di massa pari a  $\sqrt{s} = 13.6 \text{ TeV}$ . Per quanto riguarda la luminosità, nel 2022 (che è il periodo a cui fanno riferimento le simulazioni analizzate in seguito) LHC ha fornito una luminosità integrata  $L$  (definita come  $L = \int dt \mathcal{L}$ ) pari a  $L = 41.47 \text{ fb}^{-1}$ . Di questa, nello stesso periodo, CMS ne ha registrato una frazione pari a  $L = 38.01 \text{ fb}^{-1}$ . Non tutti gli eventi registrati vengono però utilizzati: può capitare che in alcuni casi l'acquisizione dei dati non avvenga perfettamente (ad esempio a causa di un rivelatore spento). Questi eventi *spuri* vengono scartati, riducendo ulteriormente la luminosità integrata. Sempre nel 2022 agli eventi ripuliti da contaminazioni spurie (detti *validated*) corrisponde una luminosità di  $L = 37.19 \text{ fb}^{-1}$  [5].

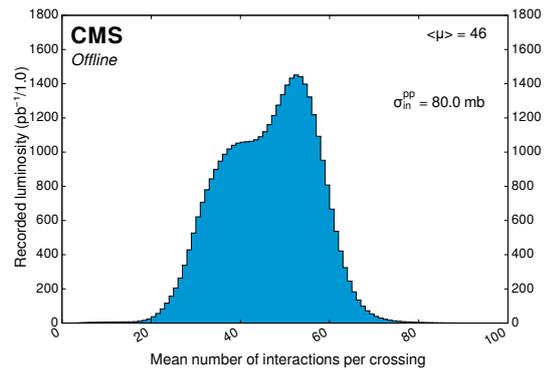


Figura 2.2: La distribuzione degli eventi di pile up durante il 2022 a LHC registrati da CMS. Fonte: [5].

## 2.2 CMS: Compact Muon Solenoid

L'esperimento CMS (Compact Muon Solenoid) è un esperimento del tipo *general purpose* ed ha, tra gli altri obiettivi, lo studio della fisica del Modello Standard. Il detector ha uno sviluppo cilindrico attorno all'asse di collisione dei fasci e pesa 12500 tonnellate, ha una lunghezza di 21,6 m e un diametro di 14,6 m. Esso è costituito da diversi sotto-rivelatori, disposti a distanze crescenti dall'asse dei fasci ed è diviso in due macro regioni: la parte centrale (*Barrel*), ossia la zona cilindrica attorno all'asse di interazione, e le zone periferiche (*Endcaps*), che chiudono il rivelatore assicurando un aumento dell'accettanza e quindi una migliore ermeticità. In Fig. 2.3 se ne può vedere una sua sezione.

Di seguito si riporta una breve descrizione dei vari sotto-rivelatori ([6]).

### Tracker

Il primo livello è dedicato al *tracking*. Il tracker è costituito da due differenti zone e si trova immediatamente attorno al punto di interazione, ad una distanza di circa qualche cm dall'asse dei fasci e si estende fino a un raggio di 1 m da quest'ultimo. Lo strato più interno è costituito da pixels, quello più esterno da strips, entrambi in Silicio. I pixels hanno una migliore risoluzione e forniscono una lettura tridimensionale della traccia, mentre gli strips solo bidimensionale. Viene utilizzato per identificare le tracce delle particelle cariche e abbinarle ai vertici da cui hanno avuto origine. Inoltre il campo magnetico presente in questa regione permette di ricavare una stima dell'impulso delle particelle. È importante sottolineare l'accettanza del tracker, individuata da dei valori di pseudorapidità di  $|\eta| < 2.5$  (per una definizione della pseudorapidità si veda la Sezione 2.3).

### Calorimetro Elettromagnetico

Il secondo livello è costituito da un calorimetro elettromagnetico (ECAL). In particolare è un calorimetro compatto, formato da cristalli assorbitori di Tungstano di piombo, il cui ruolo è assorbire completamente fotoni ed elettroni. Viene utilizzato per ricostruire la loro energia.

### Calorimetro Adronico

Il terzo livello è formato da un calorimetro adronico (HCAL). Esso nello specifico è un calorimetro segmentato destinato a raccogliere l'energia degli adroni che lo attraversano. È composto da strati di metallo alternati a strati di scintillatori plastici. Per questo layer l'accettanza è data da  $|\eta| < 5$ .

### Solenoido

Vi è poi un solenoide, costituito da materiale superconduttore, che tramite una corrente molto intensa permette di generare un campo magnetico di 3.8 T nella regione dei calorimetri e del tracciatore. Ciò fa curvare le particelle cariche e permette una stima del loro impulso dalla misura del raggio di rotazione.

### Camere a muoni

All'esterno del solenoide vi sono le camere a muoni, necessarie a rilevare il passaggio di queste particelle. Queste sono delle camere a deriva e sono contenute in una struttura di ferro magnetizzato che chiude le linee del campo magnetico, e che all'interno delle camere genera un campo magnetico con verso opposto e modulo pari a 2 T. Le camere a muoni hanno un'accettanza di  $|\eta| < 2.4$ .

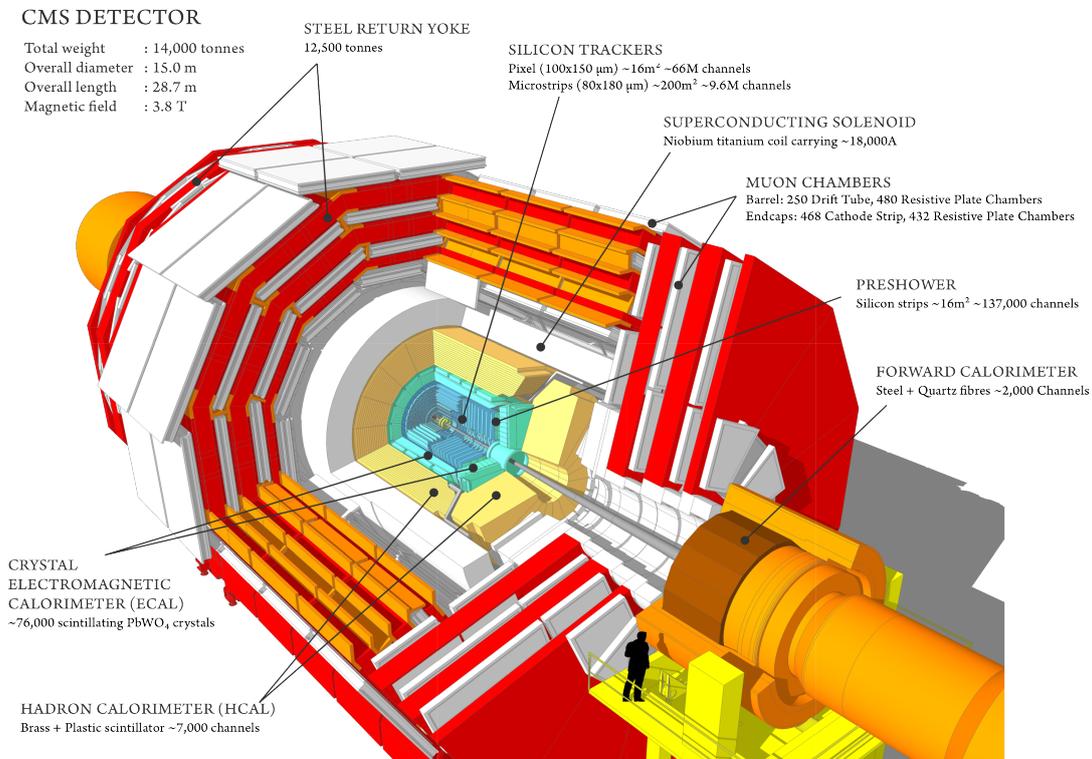


Figura 2.3: Lo schema del rivelatore CMS con le varie sottosezioni. Figura presa da [https://www.roma1.infn.it/exp/cms/introduzione\\_ita\\_eng.html](https://www.roma1.infn.it/exp/cms/introduzione_ita_eng.html).

### Trigger

In seguito alla rivelazione delle particelle è però necessario salvare le informazioni trovate. La quantità di dati da immagazzinare per ogni collisione è all'incirca 1 MB; la frequenza di intersezione dei fasci è pari a 40 MHz, e ciò richiederebbe quindi la necessità di salvare circa 40 TB/s. Poiché ciò non è implementabile è necessario attuare una drastica riduzione delle informazioni salvate. Ciò viene fatto tramite un trigger a più livelli ([9]). I due livelli sono, nello specifico, il *Level-1 trigger* (abbreviato in L1) e l'*High Level Trigger* (HLT).

**Level 1 Trigger:** questo è implementato a livello hardware, e riduce la frequenza - in un tempo dell'ordine dei  $3.8 \mu\text{s}$  - ad un valore di circa 100 kHz, basandosi su alcune proprietà di interesse come jet di alta energia, un alto valore di momento trasverso, etc.

**High Level Trigger:** a seguito della selezione attuata dal trigger L1 i dati vengono analizzati da un secondo trigger implementato a livello software - l'HLT appunto. La minor quantità di eventi da processare permette una selezione più accurata e la frequenza finale è dell'ordine del kHz.

Questa è a tutti gli effetti la prima selezione del segnale, e gli oggetti identificati a questo livello si denotano tramite la dicitura *online*. Gli oggetti ricostruiti invece nella successiva analisi e selezione dei dati si dicono *offline*. La ricostruzione dei diversi oggetti di fisica viene fatta tramite l'algoritmo *Particle Flow (PF)*, che usa informazione da tutti i sotto-rivelatori [8].

La risoluzione relativa in  $p_T$  dei muoni che attraversano il Barrel (quindi  $|\eta| < 0.83$ ) varia dall'1.5 al 2%. Quelli che invece attraversano l'Endcap ( $|\eta| > 1.2$ ) hanno una risoluzione che va dal 2 al 4%. La risoluzione relativa sull'energia  $\mathcal{E}$  è data invece dalla seguente formula

$$\frac{\sigma(\mathcal{E})}{\mathcal{E}} = \frac{2.8\%}{\sqrt{\mathcal{E}}} \oplus \frac{12\%}{\mathcal{E}} \oplus 0.3\% \quad (2.1)$$

Per lo studio del processo  $h \rightarrow \mu^+ \mu^-$  è ovviamente necessario ricostruire i muoni: questi vengono ricostruiti usando sia informazioni di tracker che informazioni dalle camere a muoni. Inoltre si usano anche informazione del calorimetro elettromagnetico ed adronico per stimarne l'isolamento. È importante sottolineare l'accettanza, che per questo processo è determinata dalle camere a muoni:  $|\eta| < 2.4$ .

## 2.3 Sistema di coordinate a CMS

Per descrivere gli eventi del processo che si vuole studiare è necessario individuare, tra le altre cose, le componenti dell'impulso  $\vec{p}$  di ogni particella. Spesso quest'ultimo viene parametrizzato in coordinate cartesiane come  $(p_x, p_y, p_z)$ , o in coordinate sferiche come  $(|p|, \theta, \phi)$ , con  $\theta$  angolo azimutale e  $\phi$  angolo polare (vedi Figura 2.4a). Poiché però il rivelatore CMS ha uno sviluppo cilindrico risulta più favorevole usare un sistema di coordinate cilindrico (che può essere visto sempre in Figura 2.4a).

Si inizia identificando l'asse polare  $z$  con l'asse dei fasci; allora è univocamente individuato il piano  $xy$  come quello perpendicolare a  $z$ . Si usa inoltre la convenzione di rivolgere l'asse  $x$  verso l'interno di LHC (vedi Figura 2.5). Si definisce poi  $\phi \in [-\pi, \pi]$  come angolo polare,  $\theta \in [0, \pi]$  come angolo formato con l'asse  $z$  e  $\vec{p}_T$  come l'impulso nel piano  $xy$ , dato da  $\vec{p}_T = \vec{p} \sin \theta$ , spesso indicato solamente con  $p_T$ .

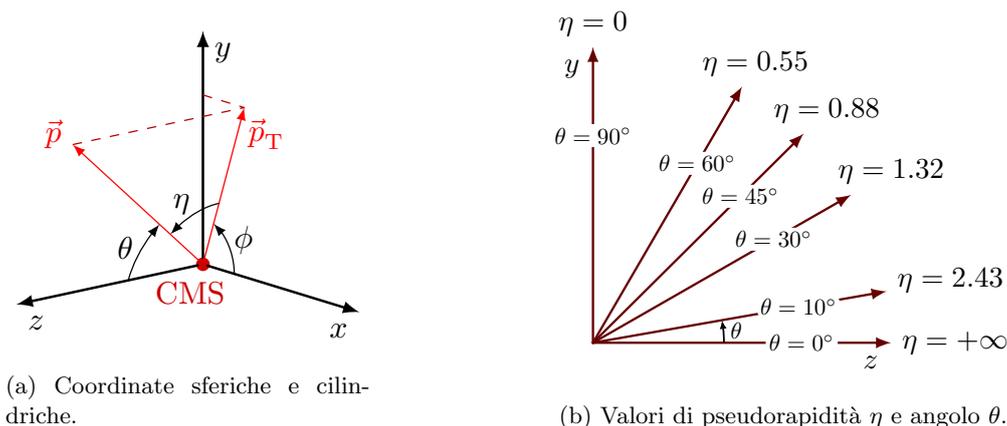


Figura 2.4: Sistema di coordinate a CMS.

Risulta però più utile sostituire  $\theta$  con una nuova variabile adimensionale, la pseudorapidità  $\eta$ , definita come

$$\eta = -\ln \left[ \tan \left( \frac{\theta}{2} \right) \right] \quad (2.2)$$

Come si vede dalla Fig. 2.4b, una particella che si muove nel piano  $xy$  ha pseudorapidità nulla. Questa aumenta in valore assoluto al diminuire dell'angolo con l'asse  $z$  e per valori di  $\eta > 3$  si parla di *produzione in avanti*. Si vede che  $\eta$  varia nel range  $(-\infty, +\infty)$ .

Ci sono diverse motivazioni di origine pratica per preferire questo sistema di coordinate. Una di queste è legata alla difficoltà di misurare l'impulso nella direzione  $z$ . Ciò che viene misurato infatti è solamente l'impulso trasverso  $p_T$ . Tuttavia, conoscendo  $\eta$  è possibile ricavare il valore della componente  $p_z$  tramite la relazione  $p_z = p_T \sinh \eta$ . Inoltre, al momento dell'accelerazione, i fasci vengono collimati lungo l'asse  $z$  con un momento trasverso totale il più vicino possibile a zero e, poiché  $p_T$  deve conservarsi, alla fine del processo che si sta studiando ci si aspetta di trovare lo stesso valore per il  $p_T$  totale (dato da  $\sum_i p_T^i$ ). Se ciò non succede, la somma vettoriale dei singoli  $p_T$  indica la presenza di una o più particelle non rilevate, che devono avere un  $p_T$  totale tale da annullare quello delle particelle rilevate. Questa componente di momento trasverso mancante si indica con il nome di *Energia trasversa*, ed è data da  $\mathcal{E}_T = \sum_i p_T^{i, \text{miss}}$ .

Ricordando infine che l'accettanza delle camere a muoni è  $|\eta| < 2.4$  si ricava che tutti i muoni in un range  $10^\circ < \theta < 170^\circ$  possono venir rilevati.

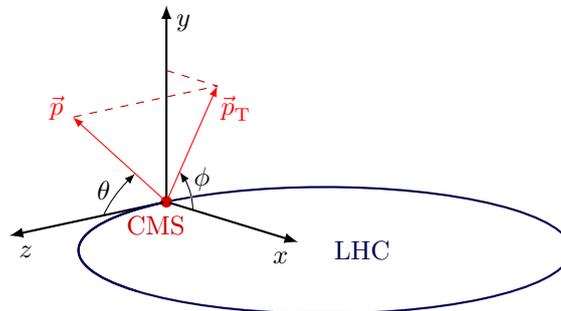


Figura 2.5: Le coordinate usate a CMS nel frame di LHC. Figura presa da [https://tikz.net/axis3d\\_cms/](https://tikz.net/axis3d_cms/).



# Capitolo 3

## Analisi delle simulazioni

L'obiettivo di questo lavoro è caratterizzare il decadimento del bosone di Higgs in una coppia di muoni, e per far ciò si inizierà uno studio dettagliato di simulazioni del processo  $h \rightarrow \mu^+ \mu^-$ . In particolare si vuole studiare la cinematica dei muoni e trovare degli algoritmi per distinguere i muoni provenienti dal decadimento del bosone di Higgs da muoni provenienti da altri fenomeni, costituenti il *fondo*.

### 3.1 Descrizione del Dataset

Gli eventi vengono inizialmente generati tramite un metodo Monte Carlo<sup>1</sup>, da quello che in gergo è detto il *Generatore*. In alcuni casi - come nelle simulazioni analizzate di seguito - i generatori possono essere più di uno.

Nello specifico i generatori usati per i campioni analizzati in seguito sono:

- AMC@NLO: usato per simulare il processo di fondo: nello specifico esso simula la produzione del bosone  $Z$  a partire dai quark (vedi Sezione 3.2).
- POWHEG: usato per simulare il segnale, quindi la produzione del bosone di Higgs a partire da fusione di gluoni o di bosoni vettori (vedi Sez. 3.2).
- PYTHIA: tramite esso si simula la frammentazione e l'adronizzazione dei quark o dei gluoni che porta alla formazione dei jet.

Nello specifico gli eventi sono stati generati con un'accuratezza al *Next-to-leading order* (NLO) in QCD perturbativa. Le corrispondenti sezioni d'urto sono state calcolate con un'accuratezza al NNLO in QCD e NLO in teoria elettrodebole (EW) per il fondo (DY, vedi 3.2), al N3LO in QCD e NLO in EW per il ggF e al NNLO in QCD e NLO in EW per il VBF.

Gli oggetti fisici salvati nei file - con le rispettive proprietà - a questo stage si denotano con la dicitura *Generator Level* (o *Gen Level*). Ad ogni particella a Gen level sono associate alcune proprietà che la caratterizzano, come le variabili cinematiche ( $\eta, \phi, p_T$ ), la massa, il *pdg id* (che identifica la natura della particella, [10]), etc. Si fa notare infine che non tutte le particelle generate vengono salvate nel file definitivo, ma alcune vengono eliminate, attuando quello che in gergo si chiama *pruning*.

In seguito questi eventi vengono processati attraverso un simulatore del rivelatore CMS, basato sul software GEANT4. Questo passaggio riproduce la rilevazione delle particelle, inclusa una simulazione dello stage di trigger.

Infine gli eventi vengono processati dal software di ricostruzione del CMS (il Particle Flow citato in Sezione 2.2). Gli oggetti salvati in questo stage si dicono anche *Reconstructed level object* (o *Reco level*). Anche qui ad ogni particella è associato un certo numero di proprietà che la caratterizzano.

Il tutto viene salvato in un formato di file del tipo NanoAOD. Questo è una versione ridotta del formato AOD ed è costituito da un formato simile a delle Ntuple, leggibile tramite ROOT e contenente le informazioni per i singoli eventi necessarie nella maggior parte delle analisi generiche. La dimensione di un singolo evento in un file NanoAOD è dell'ordine di 1 Kb.

L'analisi dei campioni di dati di simulazione è stata fatta in C++, con l'ausilio della libreria ROOT.

---

<sup>1</sup>I metodi Monte Carlo sono una classe di metodi computazionali basati sul campionamento casuale.

## 3.2 Modi di produzione del bosone di Higgs

Vi sono diverse strategie per produrre il bosone di Higgs (i diversi canali sono detti anche *modi di produzione*). Quelli considerati in questa analisi sono il *gluon gluon Fusion* (**ggF**) e il *vector boson fusion* (**VBF**).

- **ggF**: come si evince dal nome, in questo caso il bosone di Higgs viene generato tramite fusione di gluoni, dove i gluoni fanno parte dei partoni contenuti nei protoni che collidono. Ciò è dovuto al fatto che, a energie di  $\sqrt{s} = 13.6$  TeV, la probabilità di avere collisioni tra gluoni invece che tra quark è maggiore<sup>2</sup>. Questo è il metodo di produzione con sezione d'urto maggiore a LHC, e quindi quello più probabile. Questo perché il bosone di Higgs viene prodotto da un loop di quark top e il quark  $t$  è quello con coefficiente di accoppiamento ad  $h$  maggiore. Si noti che il bosone di Higgs viene prodotto centralmente.
- **VBF**: Questo è il modo di produzione più probabile a LHC dopo il ggF, e ciò è dovuto al fatto che in questo caso il bosone  $h$  si accoppia ai bosoni vettori  $W$  o  $Z$  (indicati genericamente con  $V$ ). In questo modo di produzione due quark emettono ciascuno un bosone  $V$  e questi, fondendosi, creano un bosone di Higgs, anche in questo caso prodotto centralmente. Oltre ad  $h$  nello stato finale si trovano anche due quark, troppo distanziati tra loro per interagire, e che quindi tendono ad adronizzare, dando origine a due jet prodotti in avanti.

Di seguito sono riportati i digrammi di Feynman dei processi appena descritti:

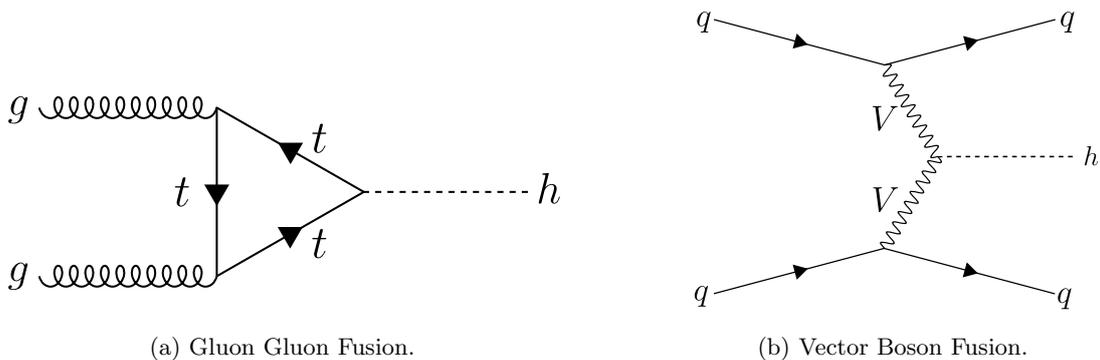


Figura 3.1: Diagrammi di Feynman per la produzione di  $h$ .

Questi tuttavia non sono gli unici processi incontrati nell'analisi. Poiché gli oggetti finali che si vogliono studiare sono i muoni c'è un ampio spettro di processi di fondo che contaminano la produzione del bosone di Higgs e che rendono difficile l'individuazione e l'isolamento del processo considerato. Il processo di fondo preponderante in questo tipo di eventi è il Drell-Yan (indicato nel seguito anche con la dicitura **DY**) in cui un quark ed un anti-quark collidono, producendo un bosone  $Z$  oppure un fotone virtuale, che decade in seguito in una coppia di leptoni carichi di segno opposto. Si noti che il quark incidente è un quark di valenza del protone, mentre l'anti-quark incidente fa parte dei quark del mare del protone restante. Dalla teoria sappiamo inoltre che per il bosone  $Z$  sono possibili anche i decadimenti in coppie di leptoni neutri, ossi neutrini:  $Z \rightarrow \nu_\ell \bar{\nu}_\ell$ , con  $\ell = e, \mu, \tau$ . Tuttavia si esclude questa casistica poiché non interferisce con la produzione di muoni da  $h$ .

È importante sottolineare che - al fine di ottimizzare la produzione del campione di simulazione per questo processo - nei file analizzati di seguito si considera solo il caso in cui il propagatore è un bosone  $Z$ , e si esclude il caso di un fotone, sin dall'inizio (quindi a livello del generatore). Ciò viene fatto in particolare escludendo tutti gli eventi con massa invariante dei leptoni minore di  $50 \text{ GeV}/c^2$ . Infatti, la probabilità di avere un  $\gamma$  invece che una  $Z$  è molto maggiore, e includendo anche il caso del fotone si dovrebbero generare un numero molto maggiore di eventi per avere la stessa sensibilità per  $Z$ . Inoltre per lo studio di  $h$  gli eventi con  $m_{\ell\ell} < 50 \text{ GeV}/c^2$  spesso sono poco significativi.

Infine nel campione di Drell-Yan viene simulata anche la presenza di due jet, in questo caso originati da gluoni emessi dai quark nello stato iniziale (*Initial State Radiation*).

<sup>2</sup>Queste probabilità sono date dalle cosiddette *Parton Distribution Functions* (PDF), che danno la probabilità di trovare un partone con un certo impulso all'interno di un protone di data energia.

Il processo si può quindi schematizzare come  $q\bar{q} \rightarrow Z \rightarrow \ell^+ \ell^-$ , con  $\ell = e, \mu, \tau$ . Di questo processo è riportato il diagramma di Feynman nella sottostante Figura 3.2.

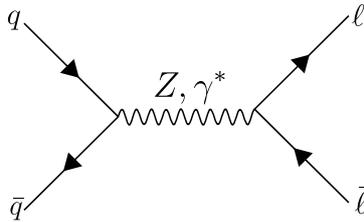


Figura 3.2: Diagramma di Feynman per il processo Drell-Yan.

Di seguito una tabella che riassume - per i tre processi menzionati sopra - la sezione d'urto, il numero totale di eventi analizzati e la luminosità equivalente dei dataset.

Processo	Sezione d'urto(fb)	Eventi	Luminosità ( $\text{fb}^{-1}$ )
DY	$(6.283^{+0.077}_{-0.094}) \times 10^6$	$9.85 \times 10^7$	$15.69^{+0.19}_{-0.23}$
ggF	$11.42 \pm 0.57$	$2.00 \times 10^5$	$(1.751 \pm 0.087) \times 10^4$
VBF	$0.889^{+0.019}_{-0.018}$	$3.99 \times 10^5$	$(4.492^{+0.099}_{-0.094}) \times 10^5$

Tabella 3.1: Gli eventi si riferiscono a quelli totali, a monte di tutte le selezioni descritte nella successiva Sezione. Con Luminosità si intende Luminosità integrata equivalente. Le sezioni d'urto sono prese da [13, 14].

La sezione d'urto del Drell-Yan è stata calcolata a partire dalla sezione d'urto per una coppia elettrone-positrone  $\sigma(pp \rightarrow e^+ e^-) = 2.0942 \times 10^6$  fb (valore preso da [13]) usando l'universalità leptonica, ossia come  $\sigma_{\text{DY}} = 3\sigma(pp \rightarrow e^+ e^-)$ .  $\sigma(pp \rightarrow e^+ e^-)$  è stata calcolata numericamente tramite l'apposito software MATRIX, con un'accuratezza al NNLO in QCD e NLO in teoria EW, e queste correzioni sono state unite additivamente.

Le sezioni d'urto per il segnale sono state prese invece da [14]. Si fa notare che i valori delle sezioni d'urto qui riportate - calcolati come interpolazione tra  $\sigma(13\text{TeV})$  e  $\sigma(14\text{TeV})$  - sono provvisori, in attesa della pubblicazione di risultati più precisi. I valori riportati nella tabella per il ggF e il VBF sono le sezioni d'urto inclusive; per trovare poi le sezioni d'urto relative al decadimento  $h \rightarrow \mu^+ \mu^-$  si sono moltiplicate queste ultime per il *Branching Ratio* di questo decadimento:  $\mathcal{B}(h \rightarrow \mu^+ \mu^-) = 2.18 \times 10^{-4}$ . Queste ultime sono state usate in seguito per calcolare la luminosità equivalente. Questa è stata calcolata a partire dalla formula  $N = L\sigma$ , con  $N$  il numero di eventi riportato nella tabella.

### 3.3 Analisi del fondo (Drell-Yan)

Si procede ora con uno studio dettagliato del processo di fondo. Poiché si sta cercando il decadimento  $h \rightarrow \mu^+ \mu^-$  l'obiettivo è studiare i muoni presenti in stato finale. Verrà quindi imposto un filtro per selezionare a Gen Level solo gli eventi in cui è presente una coppia di muoni prodotta dal decadimento della  $Z$ , e di questi si studierà la cinematica. In seguito si studieranno i muoni ricostruiti (quindi a Reco Level) e si selezionerà solo una frazione di eventi che rispettano determinate condizioni (le cosiddette *selezioni*). L'obiettivo di ciò è migliorare il più possibile la risoluzione in massa invariante della coppia di muoni.

#### 3.3.1 Muoni da $Z$ a Livello di Generatore

Come detto in precedenza in questa sezione si sono isolati gli eventi  $Z \rightarrow \mu^+ \mu^-$ , che corrispondono a circa 1/3 degli eventi totali  $Z \rightarrow \ell^+ \ell^-$ , come si può vedere anche in Fig. 3.3 (rispettivamente, terzo e primo bin). Dal grafico si può vedere anche che se si selezionano solo gli eventi con almeno un muone nello stato finale (senza chiedere che sia prodotto da  $Z$ ) si ha circa la metà degli eventi iniziali (secondo bin del grafico), e questo è dovuto agli eventi in cui  $Z \rightarrow \tau^+ \tau^-$  e in cui successivamente uno dei due o

entrambi i  $\tau$  decadono come  $\tau \rightarrow \mu \bar{\nu}_\mu \nu_\tau$  (in realtà questi  $\mu$  possono venire anche da decadimenti leptonici di adroni all'interno dei jet, ma questi in genere vengono esclusi chiedendo che i muoni siano isolati - vedi Sez. 3.3.4). Questi muoni avranno ovviamente una cinematica diversa da quelli prodotti da  $Z$ , e verranno studiati brevemente nella sezione 3.3.2.

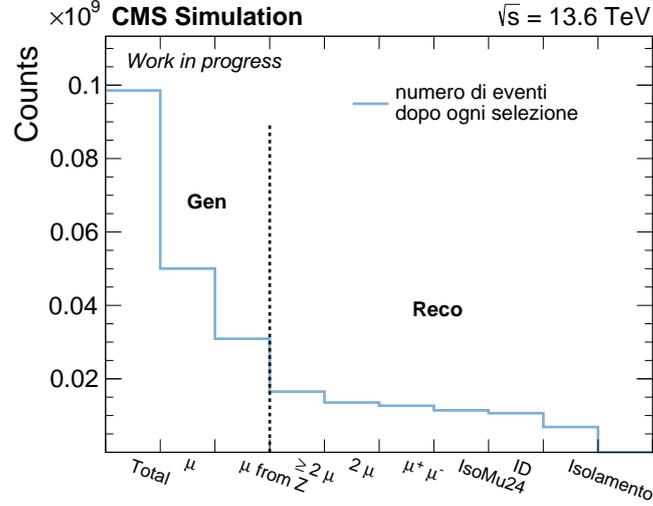


Figura 3.3: Riassunto di tutte le selezioni applicate. Ad ogni bin corrisponde una selezione. A sinistra della linea tratteggiata quelle a Gen Level, a destra quelle a Reco Level. In ordinata c'è il numero di eventi per ogni selezione.

Negli eventi selezionati (quindi  $Z \rightarrow \mu^+ \mu^-$ ) si è poi studiata la cinematica della coppia di muoni. Una prima classificazione che è stata fatta è separare i due muoni in base al loro valore di  $p_T$ : nello specifico si è definito il **Leading  $p_T$  muon** (spesso abbreviato in **Lead  $\mu$** ) come quello con valore di  $p_T$  maggiore tra i due, e analogamente il **Sub-Leading  $p_T$  muon** (Sub-Lead  $\mu$ ) come quello con  $p_T$  minore. Una volta fatto ciò si è proceduto a graficare i valori  $\eta$ , di  $\phi$  e di  $p_T$  (Fig. 3.4 e 3.5) separatamente per il Leading e per il Sub-Leading muon.

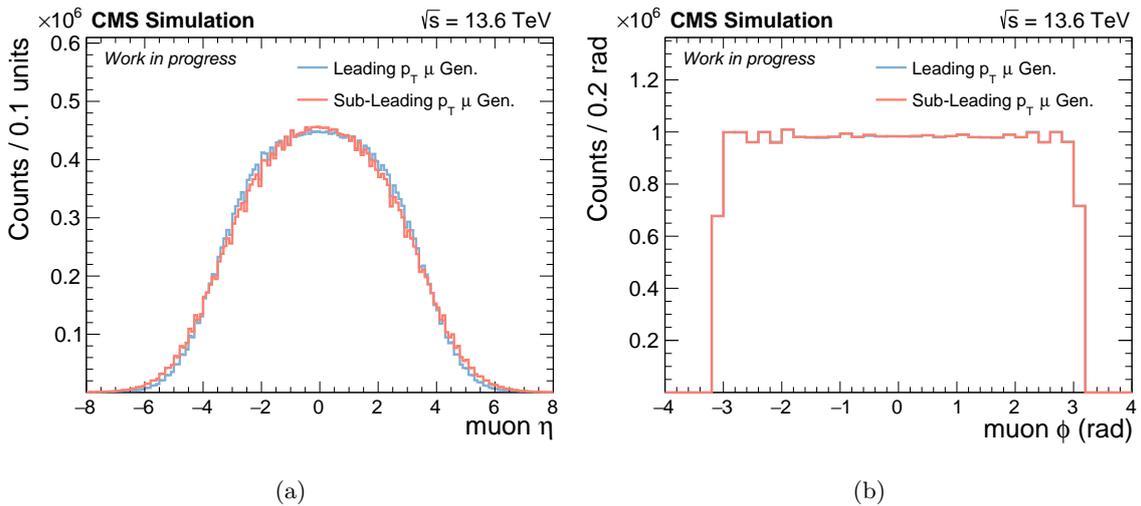


Figura 3.4: Variabili cinematiche angolari per i muoni da  $Z$  a Generator Level.

In  $\eta$  si può vedere come la distribuzione vada ben oltre l'accettanza di  $|\eta| < 2.4$  e si notano anche lievi differenze tra Leading e Sub-Leading muon, come ad esempio il fatto che questi ultimi, più dei primi, tendano ad essere prodotti ad  $\eta = 0$ . La percentuale di muoni che passano attraverso la regione del Barrel ( $|\eta| < 0.83$ ) è di circa il 25.9%, mentre i muoni inclusi nell'accettanza del rivelatore (quindi con  $|\eta| < 2.4$ ) sono circa il 66.8%. Solo questa frazione di muoni può quindi venir rilevata negli esperimenti.

Queste differenze sono invece molto più evidenti nel grafico dell'impulso trasverso 3.5: in entrambe le distribuzioni si può però individuare un picco intorno a 45 GeV/c, ossia a circa  $m_Z/2$ . Questo implica che nella maggior parte dei casi la  $Z$  è prodotta a riposo nel piano trasverso (come si vede dalla Figura 3.9a). Infatti dalla teoria dei decadimenti si ha (in unità naturali)

$$\begin{cases} \mathcal{E}_Z = \mathcal{E}_{\mu,1} + \mathcal{E}_{\mu,2} \\ \vec{p}_Z = \vec{p}_{\mu,1} + \vec{p}_{\mu,2} \end{cases} \quad (3.1)$$

che per  $\mathcal{E}_Z = m_Z$  (produzione a riposo) dà  $p_{\mu,1} = p_{\mu,2} \equiv p_\mu$  e di conseguenza  $\mathcal{E}_{\mu,1} = \mathcal{E}_{\mu,2} \equiv \mathcal{E}_\mu \simeq 45$  GeV/c<sup>2</sup>. Infine, poiché  $m_\mu \ll m_Z/2$ , si ha  $p_\mu = \mathcal{E}_\mu$ . Questo non garantisce che la  $Z$  sia però globalmente a riposo e in effetti, come si mostrerà successivamente, ciò non succede.

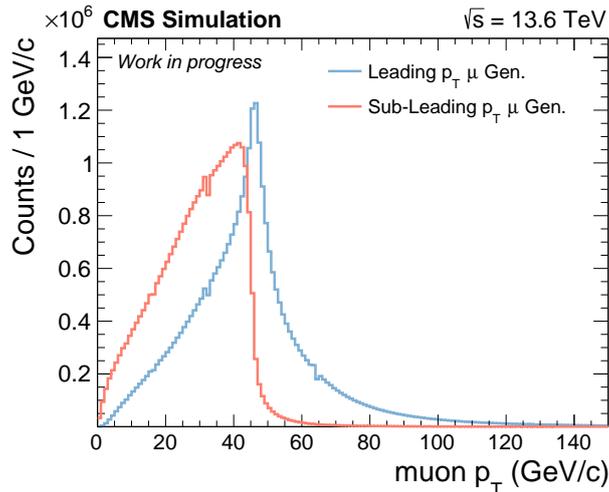


Figura 3.5: Distribuzione di  $p_T$  per muoni da  $Z$  a Gen Level.

Nonostante questo però le due distribuzioni sono abbastanza diverse, e come si vede i Sub-Leading muons hanno una distribuzione più spostata verso valori minori di  $p_T$  rispetto ai Leading muons. È possibile studiare anche la separazione angolare tra i due muoni, calcolando le quantità  $\Delta\eta = \eta_{\mu,1} - \eta_{\mu,2}$  e  $\Delta\phi = \phi_{\mu,1} - \phi_{\mu,2}$ , riportate in Fig 3.6.

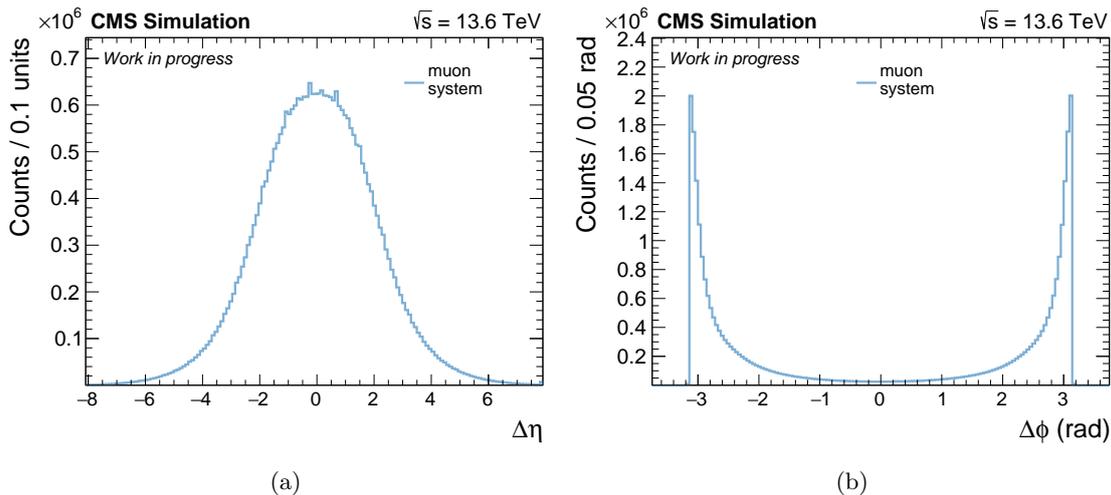


Figura 3.6:  $\Delta\eta$  e  $\Delta\phi$  per i muoni da  $Z$  a Generator Level.

Qui si vede che i muoni vengono prodotti preferibilmente con  $\Delta\eta = 0$ , ossia con lo stesso angolo  $\theta$  rispetto all'asse  $z$ . Viceversa, il grafico di  $\Delta\phi$  mostra che la maggior parte dei muoni sono prodotti *back-to-back*, ossia con  $|\Delta\phi| = \pi$ . Questo accade quando  $p_T^Z \simeq 0$ , come si può vedere anche dal grafico di correlazione riportato di seguito (Fig. 3.7b). Queste informazioni possono poi essere combinate per

definire una separazione *globale* tra i due muoni, data dalla variabile  $\Delta R$ : questa è definita come  $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ , e il suo andamento è riportato in Fig. 3.7a. Si può vedere il picco a  $\Delta R = \pi$  ereditato da  $\Delta\phi$ . Dal grafico di correlazione si vede invece come gli eventi in cui la  $Z$  è prodotta con momento trasverso nullo (o quasi) sono correlati agli eventi in cui la separazione in  $\eta$  tra i due muoni è circa zero.

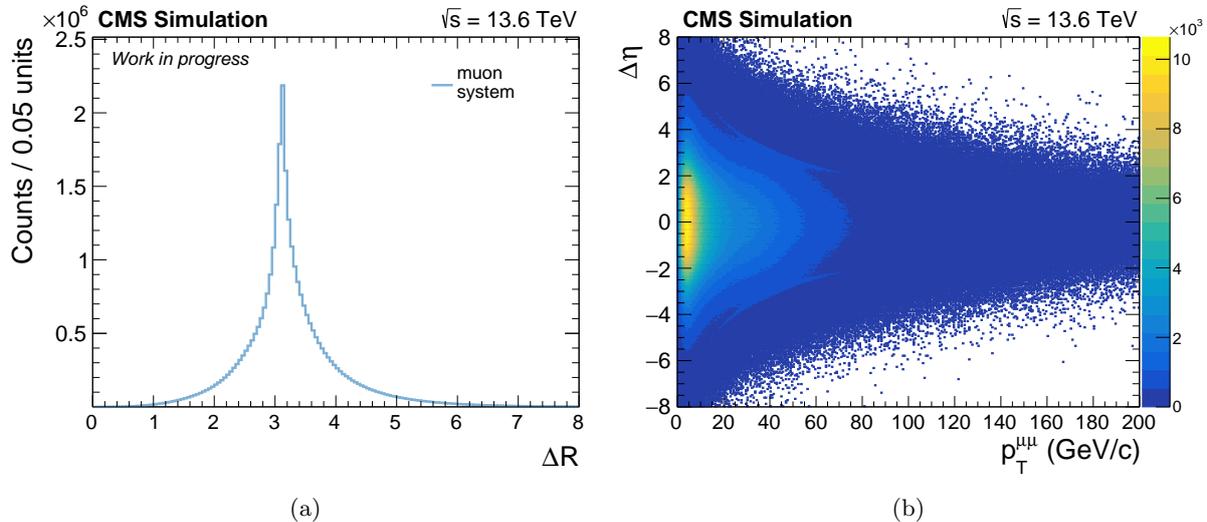


Figura 3.7: A sinistra  $\Delta R$  tra muoni, a destra correlazione  $\Delta\eta$  tra muoni -  $p_T^Z$ , entrambi a Gen Level. Per un grafico di  $p_T^Z$  vedere la Fig. 3.9a

Dai muoni in stato finale si possono poi ricostruire le proprietà del bosone  $Z$  che li ha prodotti: dati i 4-vettori dei muoni in stato finale  $\mathbf{p}_{\mu,1}, \mathbf{p}_{\mu,2}$ , si può ricostruire il vettore  $\mathbf{p}_Z = \mathbf{p}_{\mu,1} + \mathbf{p}_{\mu,2}$  e studiarne la cinematica. Le componenti di  $\mathbf{p}_Z$  vengono indicate anche con  $m_{\mu\mu}, \eta_{\mu\mu}$ , etc. Nello specifico la massa invariante  $m_{\mu\mu}$  è data dalla seguente relazione

$$m_{\mu\mu} = \sqrt{\left(\sum_i \mathcal{E}_i\right)^2 - \left(\sum_i \vec{p}_i\right)^2} \quad (3.2)$$

dove in questo caso  $\mathcal{E}_i$  e  $\vec{p}_i$  rappresentano energia e impulso dei due muoni. Di seguito è riportata la cinematica del bosone  $Z$  ricostruito (Figure 3.8 e 3.9).

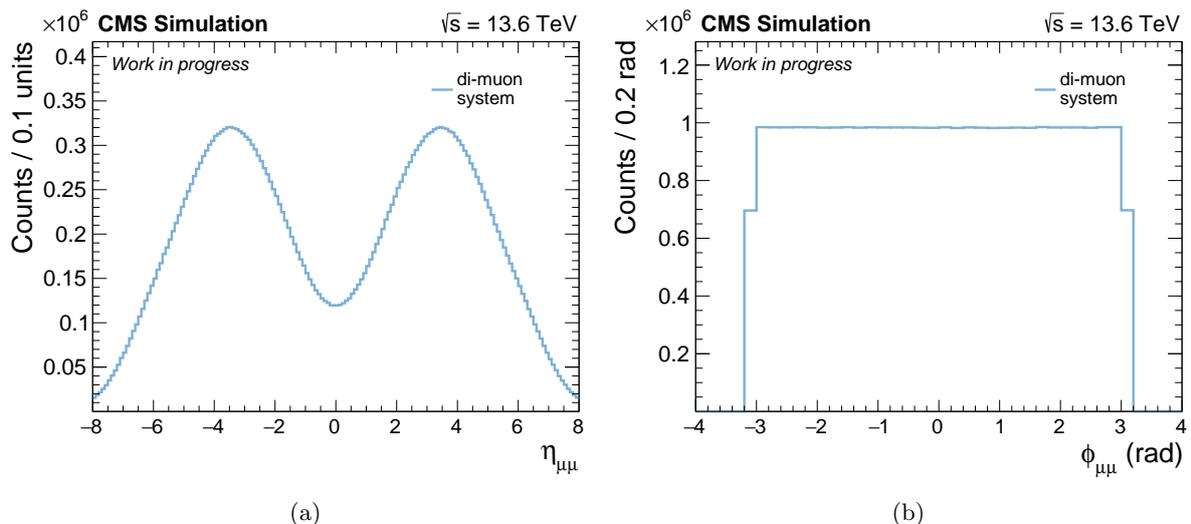


Figura 3.8:  $Z$  ricostruita a Generator Level:  $\eta$  e  $\phi$ .

Dal secondo grafico (3.8b) si vede sempre una distribuzione uniforme in  $\phi$ , mentre nel primo grafico (3.8a) si vede un andamento abbastanza diverso da quanto visto per i muoni. Questo è dovuto al fatto che

la  $Z$  ha una massa molto grande. Inoltre  $Z$  tende a essere prodotta in avanti: gli eventi in cui  $|\eta_{\mu\mu}| > 3.5$  ( $\theta < 3.46^\circ$ ) sono circa il 48.4%, ossia in quasi metà dei casi la  $Z$  è prodotta in avanti. Dal grafico di  $p_T$  (3.9a) si vede che  $Z$  è prodotta principalmente a riposo (o quasi) nel piano trasverso.

Per la massa invariante  $m_{\mu\mu}$  si trova l'usuale andamento dato da una Breit-Wigner (3.9b), piccato intorno a  $m_Z$ . Questa ha, oltre al taglio a  $m_{\mu\mu} < 50 \text{ GeV}/c^2$  descritto in Sez. 3.2, anche un taglio superiore, per  $m_{\mu\mu} > 130 \text{ GeV}/c^2$ , che si può vedere in Fig. 3.13c.

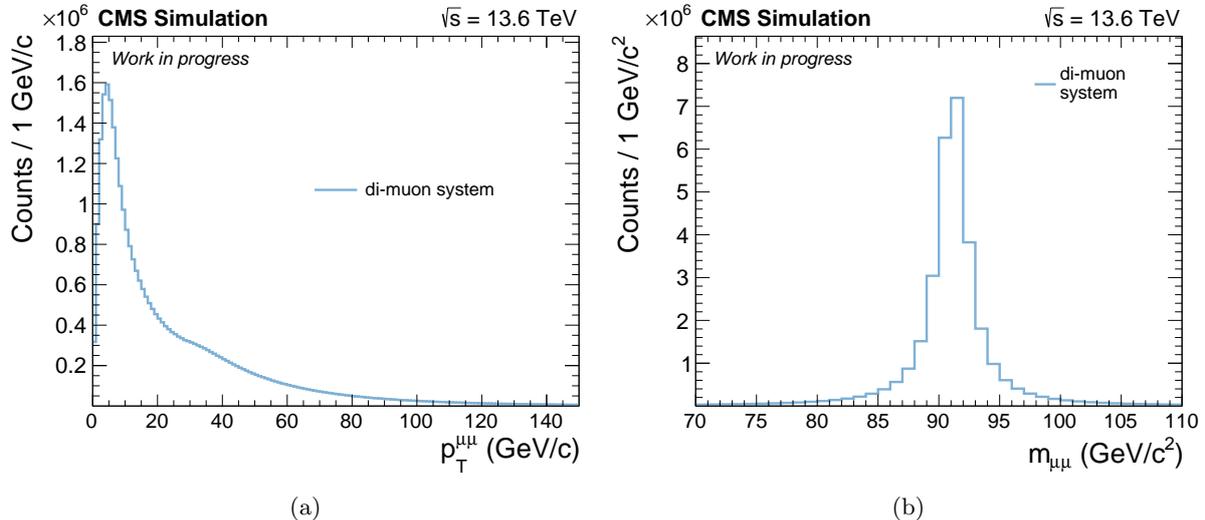


Figura 3.9:  $Z$  ricostruita a Generator Level:  $p_T$  e massa.

### 3.3.2 Muoni da $\tau$ a Gen Level

In un terzo degli eventi avviene il decadimento in leptone  $\tau$ :  $Z \rightarrow \tau^+ \tau^-$ . Questi hanno una cinematica identica (a Gen Level) a quella dei muoni descritta nella sezione precedente poiché, fissata la cinematica del bosone  $Z$ , l'energia a disposizione della coppia di  $\tau$  è comunque molto grande: infatti  $m_\tau \simeq 1.8 \text{ GeV}/c^2 \ll m_Z$ . In particolare i leptoni  $\tau$  hanno la stessa distribuzione in  $p_T$  (che si può vedere in Fig. 3.10b), e di conseguenza il loro successivo decadimento come  $\tau \rightarrow \mu \bar{\nu}_\mu \nu_\tau$  (o anche  $\tau \rightarrow e \bar{\nu}_e \nu_\tau$  o  $\tau \rightarrow had$ , ininfluenti per questa analisi) sarà in volo.

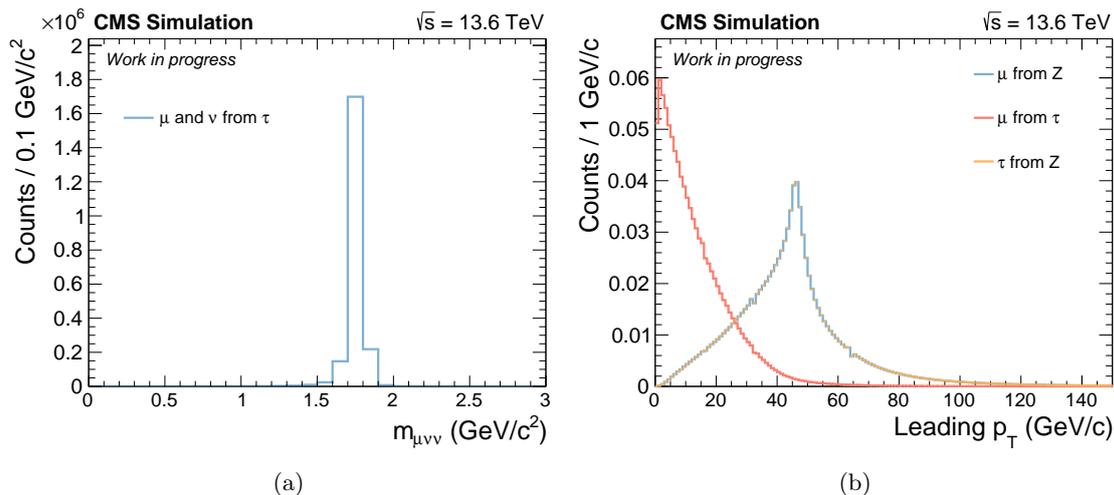


Figura 3.10: A sinistra massa invariante del sistema di  $\mu$  e  $\nu$  da  $\tau$ . A destra sovrapposizione del Leading  $p_T$  per  $\mu$  da  $Z$ ,  $\mu$  da  $\tau$  e  $\tau$  da  $Z$ , dove le distribuzioni sono state normalizzate.

Si può studiare allora a Gen Level il sistema formato dal muone più la coppia di neutrini provenienti dal decadimento del  $\tau$  - farlo con gli oggetti ricostruiti è più difficile poiché i neutrini difficilmente vengono

rilevati - e si ritrova la cinematica dei  $\tau$  discussa precedentemente, e in particolare nella massa invariante si ritrova il picco per valori prossimi a  $m_\tau$  (Figura 3.10a).

La cinematica dei muoni è abbastanza simile in  $\eta$ , con il solito andamento a campana centrato in  $\eta = 0$  ma è sostanzialmente diversa in  $p_T$  (vedi sempre 3.10b), dove nel 97,3% degli eventi il momento trasverso del muone è inferiore a 40 GeV/c, al contrario del caso  $Z \rightarrow \mu^+ \mu^-$  dove gli eventi con  $p_T < 40$  GeV/c sono il 41,9% per Lead e l'82,0% per Sub-Lead. Nell'analisi dei dati sarà allora possibile rimuovere i muoni provenienti dal  $\tau$  imponendo delle selezioni cinematiche sul  $p_T$  e sulla massa invariante.

### 3.3.3 Muoni da $Z$ a Reco Level

Si passa ora a descrivere la cinematica degli oggetti ricostruiti. Si ricorda che gli eventi di partenza sono quelli a Gen Level in cui una  $Z$  va in due muoni (terzo bin del grafico 3.3). In questo caso non essendo più possibile discriminare i muoni da  $Z$  rispetto a quelli provenienti da altri processi (in particolare dai 2 jet) il risultato sarà quello di avere due ma anche più muoni nello stato finale, come si vede dal grafico 3.11. Ovviamente di questi muoni solo due sono quelli “giusti” (cioè provenienti da  $Z$ ), gli altri sono muoni derivanti da altri decadimenti o anche da errori di ricostruzione.

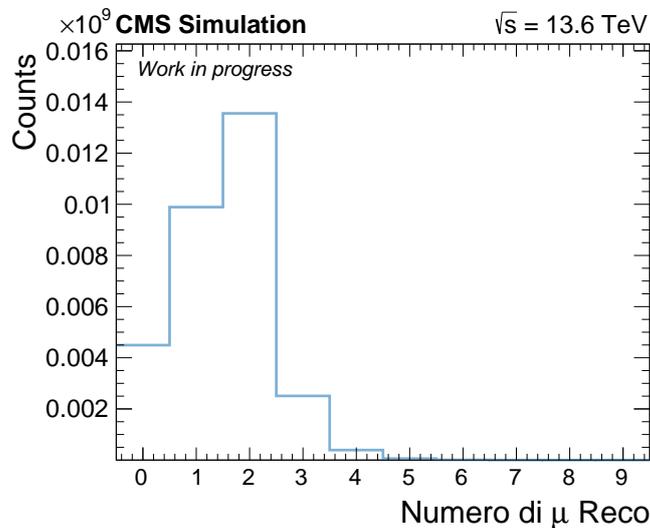


Figura 3.11: Eventi divisi in base al numero di muoni ricostruiti. Si arrivano ad avere fino a 8 muoni.

Nel grafico si vede anche che vi è una percentuale di eventi in cui il numero di muoni è nullo: ciò può succedere o perché i muoni non vengono ricostruiti oppure perché non rientrano nell'accettanza del rivelatore, e quindi non vengono proprio rilevati.

Una prima selezione attuata è quella di considerare (per semplicità) gli eventi con solamente due muoni: dopo averli ordinati in base a  $p_T$  decrescente si sono identificati il Leading e il Sub-Leading  $p_T$  muon come descritto nella sezione precedente. Indicando con  $N_G$  gli eventi a Gen Level in cui  $Z \rightarrow \mu^+ \mu^-$  e con  $N_{\geq 2}$  (rispettivamente  $N_{=2}$ ) gli eventi a Reco con almeno (risp. solo) 2 muoni si può definire l'efficienza di ricostruzione per gli eventi con almeno due muoni come  $\varepsilon_{\geq 2} = \frac{N_{\geq 2}}{N_G}$ , e analogamente per  $\varepsilon_{=2}$ . I valori trovati sono  $\varepsilon_{\geq 2} = 53.5\%$  e  $\varepsilon_{=2} = 43.9\%$ . Di questi si sono selezionati in seguito solo gli eventi in cui i due muoni hanno carica opposta; questo ulteriore taglio ha ridotto l'efficienza a  $\varepsilon_{\mu^+\mu^-} = 41.0\%$ . Si fa notare che in questi valori di efficienza è in realtà racchiusa anche l'accettanza geometrica del rivelatore.

Di questi eventi selezionati si è studiata la cinematica, riportata in Fig. 3.12. Non si riporta  $\phi$  per motivi di concisione. D'altronde, a meno che non vi siano regione del rivelatore spente, ci si aspetta di trovare la stessa distribuzione uniforme vista a Gen Level, poiché il rivelatore copre tutta la regione angolare  $-\pi < \phi < \pi$ . Nei grafici seguenti si riportano i risultati ricostruiti (Reco) sovrapposti a quelli generati (Gen), così da poter vedere l'effetto dell'efficienza di rivelazione nella riduzione del numero di eventi.

In  $p_T$  si vede come l'andamento delle distribuzioni venga solo lievemente alterato (in particolare è presente un picco a bassi  $p_T$  per il Sub-Leading muon a Reco Level, dovuto probabilmente a errori di ricostruzione e che sparirà nelle successive selezioni). Dal grafico in basso a destra si vede come il rapporto

sia pressoché costante, tranne a bassi  $p_T$  (indicativamente  $p_T < 40 \text{ GeV}/c$ ) dove vi è più distorsione. Le fluttuazioni del rapporto in viola ad alti  $p_T$  sono dovute a valori molti piccoli che i bin assumono negli istogrammi a Reco e Gen Level del Sub-Leading muon.

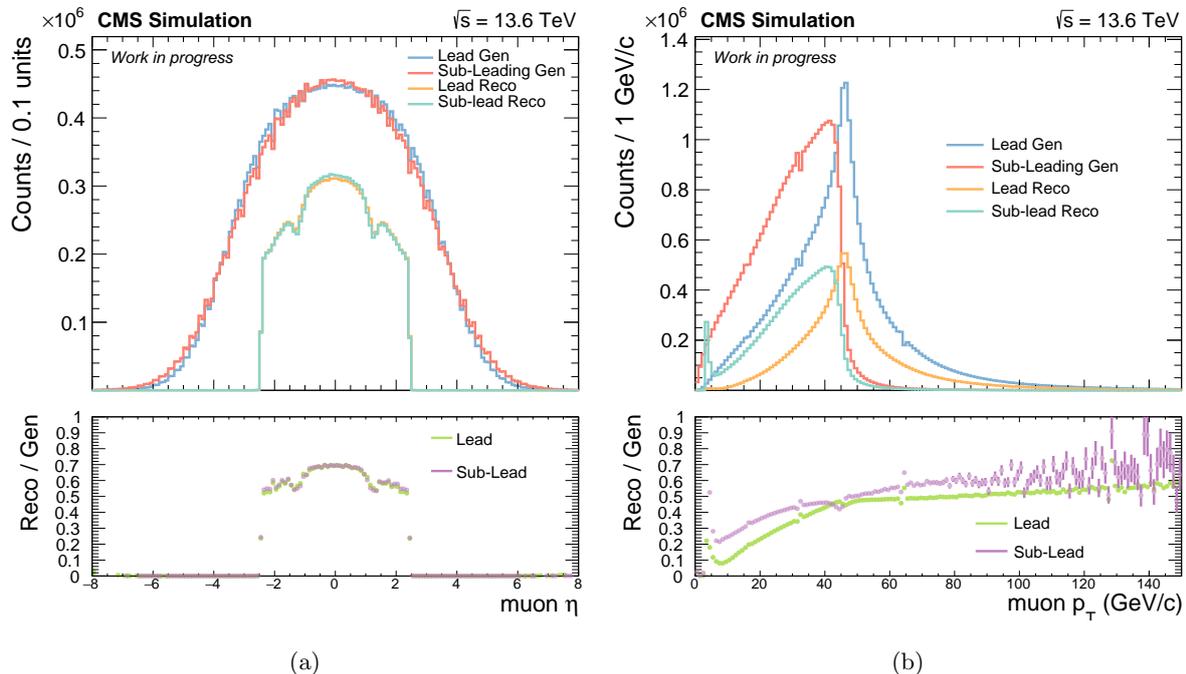


Figura 3.12:  $p_T$  e  $\eta$  per i muoni a Reco Level sovrapposti alle distribuzioni a livello di generatore. In basso si possono vedere i rapporti tra gli istogrammi Reco e Gen Level sia per i Leading che Sub-Leading muons.

Nelle distribuzioni in  $\eta$  risaltano invece di più le differenze: la più ovvia è l'effetto dell'accettazione geometrica del rivelatore, che restringe i grafici a Reco Level ad un intervallo  $-2.4 < \eta < 2.4$ . Si vede poi che nella zona centrale la ricostruzione è abbastanza fedele (il Ratio nel grafico in basso è circa costante) con le distribuzioni Lead e Sub-Lead che seguono lo stesso andamento visto in Fig. 3.4, ma ciò non vale più per  $|\eta| > 1$ , dove il rapporto inizia a decrescere. Le lievi depressioni nella zona individuata da  $0.8 < \eta < 1.2$  sono dovute al rivelatore: in questo intervallo spaziale infatti vi è la connessione tra Barrel e Endcap, e ciò si ripercuote in una minore ermeticità e quindi un numero minore di muoni rilevati.

Dopodiché si è proceduto a ricostruire il bosone  $Z$  come fatto precedentemente, usando i muoni selezionati sopra (Fig. 3.13). La distribuzione del momento trasverso è piuttosto fedele a quanto trovato a livello di generatore, e si vede sostanzialmente solo un riscaldamento dell'istogramma senza distorsioni apprezzabili. Nella distribuzione di  $\eta$  invece si vede una notevole deformazione, dovuta sempre all'accettazione, che ha come conseguenza una riduzione maggiore di eventi nei picchi a  $|\eta| \simeq 3.5$  rispetto alla zona centrale: infatti gli eventi con  $|\eta_{\mu\mu}| > 3.5$  sono ora solo il 17.6%.

Infine nel grafico della massa invariante (questa volta riportato con asse  $y$  logaritmico) si possono vedere i tagli del generatore (per il grafico in blu), una diminuzione di eventi intorno al picco e (meno evidente graficamente) un aumento della risoluzione di  $m_{\mu\mu}$  (dovuto a effetti di rivelazione e ricostruzione) per il grafico in rosso. Per quest'ultimo si vedono infine anche delle code per  $m_{\mu\mu} < 50 \text{ GeV}/c^2$  e  $m_{\mu\mu} > 130 \text{ GeV}/c^2$ . Queste possono essere dovute a errori di ricostruzione oppure a errori di accoppiamento dei muoni: infatti non è detto che i primi due muoni in  $p_T$  selezionati sopra siano quelli giusti (ossia entrambi provenienti da  $Z$ ) e un loro accoppiamento sbagliato porta quindi un valore di  $m_{\mu\mu}$  lontano dal valore della  $Z$ . Si vedrà nella successiva Sezione 3.3.4 che questi eventi spuri possono essere ridotti applicando determinati filtri sugli eventi considerati (vedi Figura 3.19).

Da ultimo si fa notare che selezionando solo i muoni che attraversano il Barrel la risoluzione di  $m_{\mu\mu}$  migliora: infatti se si riempiono tre istogrammi separati di  $m_{\mu\mu}$  in base alla richiesta che entrambi, uno solo o nessuno dei due muoni abbiano  $|\eta| < 0.83$  si trova una risoluzione migliore nel primo caso e peggiore nel terzo. In Figura 3.13d sono riportati i tre istogrammi in questione, mentre in Tabella 3.2 sono riportati i valori della risoluzione trovati dal fit gaussiano, insieme al corrispondente valore di massa invariante

(trovato dallo stesso fit) e dalla percentuale di eventi che rispettano la condizione presa in considerazione.

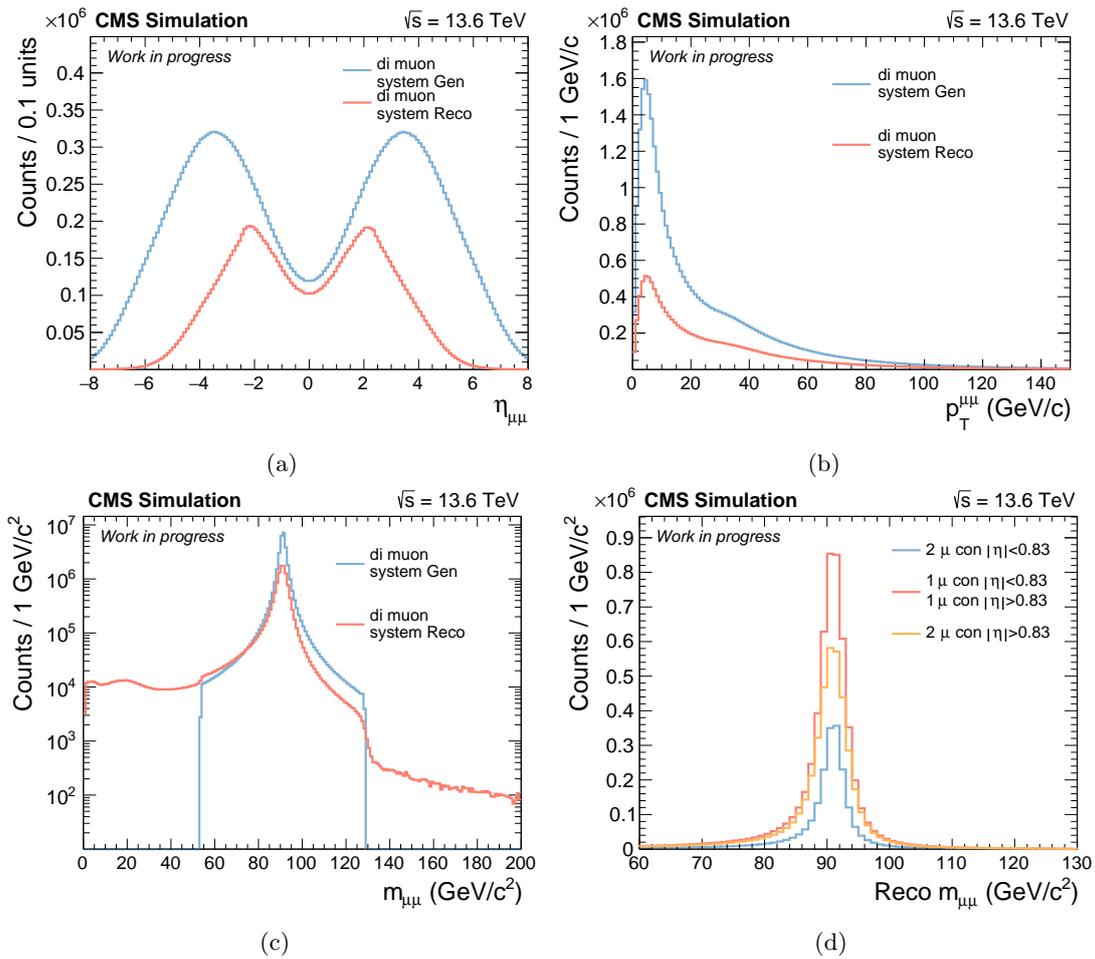


Figura 3.13: Cinematica della  $Z$  ricostruita a partire dai  $\mu$  a Reco Level;  $p_T^{\mu\mu}$ ,  $\eta_{\mu\mu}$  e  $m_{\mu\mu}$ .

Selezione	$m_{\mu\mu}$ (GeV/ $c^2$ )	Risoluzione (GeV/ $c^2$ )	Eventi (%)
entrambi i $\mu$ con $ \eta  < 0.83$	$90.939 \pm 0.002$	$1.879 \pm 0.002$	17.1
1 solo $\mu$ con $ \eta  < 0.83$	$90.907 \pm 0.001$	$2.065 \pm 0.001$	46.3
entrambi i $\mu$ con $ \eta  > 0.83$	$90.865 \pm 0.002$	$2.310 \pm 0.002$	36.6
nessuna richiesta	$90.9001 \pm 0.0008$	$2.1074 \pm 0.0009$	100

Tabella 3.2: Risultati dei fit per diversi range di pseudorapidità. Fit gaussiano nell'intervallo  $[87,95]$  GeV/ $c^2$ .

### 3.3.4 Selezioni

Di seguito si espongono le varie selezioni applicate. I valori ottimali per i tagli applicati sono stati scelti guardando anche i risultati del segnale e non solo del fondo.

#### Constrain alla Beamspot

Poiché la vita media del bosone  $Z$  è molto piccola, esso vola pochissimo prima di produrre i muoni e quindi ci si può aspettare che i muoni prodotti dal suo decadimento vengano dalla collisione primaria, e quindi dal vertice primario, a differenza dei muoni prodotti dal decadimento del  $\tau$  o da decadimenti leptonici

degli adroni. Il vertice primario di interazione è contenuto nella *Beamspot*: questa è un ellissoide (prolato lungo  $z$ ) che rappresenta l'intersezione tra due bunch, e quindi contiene tutti i punti di interazione tra protoni. La larghezza lungo gli assi è data in termini di  $\sigma$ , e vale  $\sigma_z \simeq 4$  cm e  $\sigma_x \simeq \sigma_y \simeq 10$   $\mu$ m.

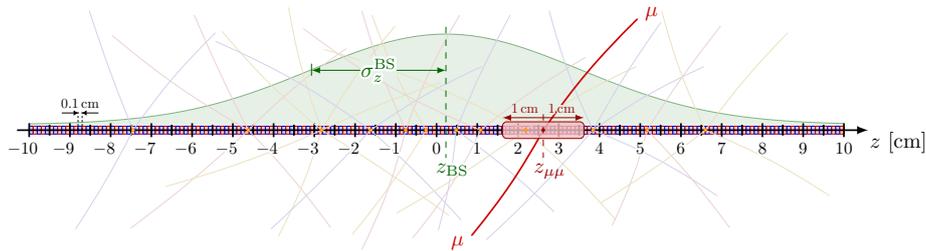


Figura 3.14: Beamspot con  $\sigma_z$  e traccia dei  $\mu$  (fonte: <https://tikz.net/beamspot/>).

Chiedere che la traccia di un oggetto venga dalla Beamspot equivale a chiedere che quello stesso oggetto sia prodotto nel vertice primario. Si parla in questo caso di *Constrain alla Beamspot*. Questo ha l'effetto di allungare la traccia di qualche cm, ossia della distanza tra la Beamspot e il tracker più interno. Ciò si traduce in una migliore risoluzione del momento trasverso  $p_T$  (vedi Grafico 3.15b) e poiché da questa dipende la risoluzione della massa invariante si avrà un miglioramento della risoluzione su  $m_{\mu\mu}$  usando il  $p_T$  costretto alla Beamspot. Quest'ultimo è il momento trasverso calcolato usando non la traccia usuale, bensì la traccia su cui si è imposto il Constrain. A questo Constrain è associato un test  $\chi^2$  che ne verifica la bontà, ossia che gli oggetti in considerazione (in questo caso i muoni) vengano effettivamente dalla Beamspot. Un grafico dei valori del  $\chi^2$  è riportato di seguito (Figura 3.15).

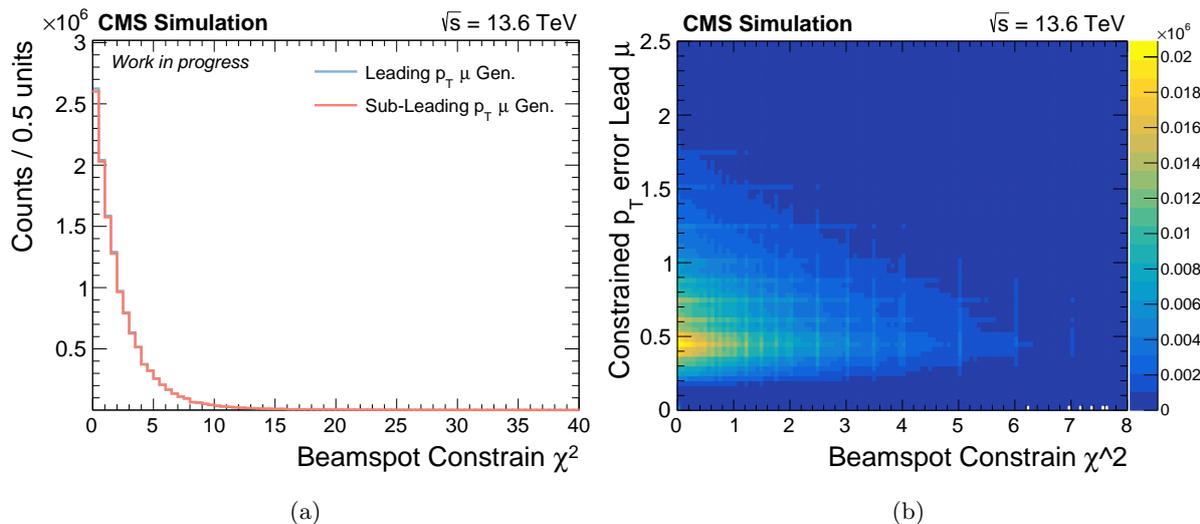


Figura 3.15:  $\chi^2$  del Constrain alla Beamspot a sinistra; le distribuzioni per Lead  $\mu$  e Sub-Lead  $\mu$  sono molto simili e quindi si sovrappongono. A destra correlazione tra errore sul  $p_T$  costretto (del solo Leading  $\mu$  per semplicità) alla Beamspot e  $\chi^2$  del Constrain alla Beamspot.

Ci aspettiamo dunque che usando il  $p_T$  costretto alla Beamspot in luogo del  $p_T$  non costretto nella costruzione dei 4-vettori associati ai muoni la risoluzione su  $m_{\mu\mu}$  migliori. Ci sono poi alcuni eventi in cui l'ipotesi nulla è falsa, ossia i muoni non vengono dalla Beamspot e in questo caso ci aspettiamo quindi un valore del  $\chi^2$  elevato. Una terza soluzione è considerare una soglia  $\chi_0^2$  per cui se  $\chi^2 < \chi_0^2$  l'ipotesi nulla è corretta ed ha effettivamente senso usare il  $p_T$  costretto, mentre se  $\chi^2 > \chi_0^2$  l'ipotesi nulla è falsa e si usa il  $p_T$  non costretto al posto di  $p_T$  costretto - quest'ultimo caso si indicherà in seguito come caso con  $p_T$  misto. Di seguito una tabella che riporta i risultati di un fit gaussiano eseguito su 3 istogrammi diversi di massa invariante  $m_{\mu\mu}$ , uno riempito considerando il  $p_T$  non costretto, uno considerando il  $p_T$  costretto e infine l'ultimo riempito considerando il caso misto. Si è scelta come soglia  $\chi_0^2 = 30$ . Gli eventi con  $\chi^2 > 30$  sono meno dello 0.7% per il Leading  $\mu$  e meno dello 0.9% per il Sub-Leading  $\mu$ .

Selezione	$m_{\mu\mu}$ (GeV/ $c^2$ )	Scarto %	Risoluzione (GeV/ $c^2$ )
$p_T$ non costretto	$90.8958 \pm 0.0009$	0.319	$2.1396 \pm 0.0009$
$p_T$ costretto	$90.9001 \pm 0.0008$	0.315	$2.1066 \pm 0.0009$
$p_T$ misto	$90.9001 \pm 0.0009$	0.315	$2.1074 \pm 0.0009$

Tabella 3.3: La risoluzione ( $\sigma$ ) e  $m_{\mu\mu}$  sono stati estratti da un fit gaussiano nell'intervallo [87,95] GeV/ $c^2$ . Lo scarto è riferito alla differenza tra il valore di  $m_{\mu\mu}$  trovato dal fit e quello riportato nel PDG [15].

Gli scarti nella seconda colonna sono stati calcolati come

$$\text{scarto} = \frac{m_{\mu\mu}(\text{PDG}) - m_{\mu\mu}(\text{FIT})}{m_{\mu\mu}(\text{PDG})} \quad (3.3)$$

con  $m_{\mu\mu}(\text{PDG}) = 91.1876 \pm 0.0021$  GeV/ $c^2$  (preso da [15]). Come si può vedere c'è un miglioramento abbastanza evidente nel secondo caso ( $p_T$  costretto) rispetto al primo caso ( $p_T$  non costretto) - la risoluzione migliora dell'1.54% - e anche nel terzo caso ( $p_T$  misto) - dove la risoluzione migliora dell'1.50%. Inoltre si vede che nel secondo caso la risoluzione è migliore che nel terzo caso, contrariamente a quanto ci si aspettava. Tuttavia questi valori sono compatibili e la differenza è esigua e probabilmente dovuta ad una fluttuazione statistica. Inoltre in un'analisi più accurata verrebbe probabilmente nascosto dall'arrotondamento dovuto all'errore, e per questo si è deciso di considerare come soluzione ottimale la terza. Si nota un lieve miglioramento anche nel valore di  $m_{\mu\mu}$ .

Si noti che nei grafici a Reco Level riportati nella precedente Sezione 3.3.3 si è già considerato il  $p_T$  misto, poiché questo non altera in modo visibile le distribuzioni in  $\eta$  o  $p_T$ .

### Simulazione del trigger

Nella presa dati gli oggetti identificati come muoni vengono selezionati a livello di Trigger (L1 e HLT) in base al loro  $p_T$ : se il momento trasverso del muone con  $p_T$  più alto è maggiore di 24 GeV/ $c$  l'evento viene salvato, altrimenti viene scartato (viene anche richiesta una condizione aggiuntiva sull'isolamento dei muoni). Per avere quindi un campione Monte Carlo che sia il più simile possibile ai dati si è implementata la simulazione di questo Trigger. Facendo un fit gaussiano di  $m_{\mu\mu}$  si trova un miglioramento della risoluzione, riportata nella seguente Tabella 3.4.

Selezione	$m_{\mu\mu}$ (GeV/ $c^2$ )	Scarto %	Risoluzione (GeV/ $c^2$ )
qualsiasi $p_T$	$90.9001 \pm 0.0009$	0.315	$2.1074 \pm 0.0009$
$p_T > 24$ GeV/ $c$	$90.9118 \pm 0.0009$	0.302	$2.0959 \pm 0.0009$

Tabella 3.4: Fit gaussiano nell'intervallo [87,95] GeV/ $c^2$ . Lo scarto è stato calcolato come nell'equazione 3.3 rispetto allo stesso valore. Si noti che la prima riga è uguale a quella della Tabella 3.3.

Si vede anche un miglioramento più evidente nella stima di  $m_{\mu\mu}$ . Poiché il taglio viene fatto sul  $p_T$  ci si aspetta che queste distribuzioni vengano modificate, e così effettivamente si può vedere in Figura 3.16b. Si vede distintamente il taglio sui Leading muon e come questo influenzi anche la distribuzione del Sub-Leading muon. Inoltre l'efficienza di questo trigger in funzione del momento trasverso non è perfetta, in particolare a valori prossimi a  $p_T = 24$  GeV/ $c$  (vedi Fig. 3.16a) e per questo in seguito si implementerà anche un taglio a  $p_T > 25$  GeV/ $c$ .

Dal Grafico 3.16a si vede inoltre come le simulazioni di Drell-Yan tendano a sovrastimare l'efficienza dei dati di uno *scale factor*, in questo caso pari a circa 0.97, di cui bisognerà tenere conto al momento del confronto delle simulazioni con i dati.

L'efficienza di questo taglio rispetto al totale - calcolata come  $\varepsilon_{\text{Trig}} = \frac{N_{\text{Trig}}}{N_G}$  - è pari a  $\varepsilon_{\text{Trig}} = 36.9\%$ , con  $N_{\text{Trig}}$  il numero di eventi che passano la simulazione del Trigger.

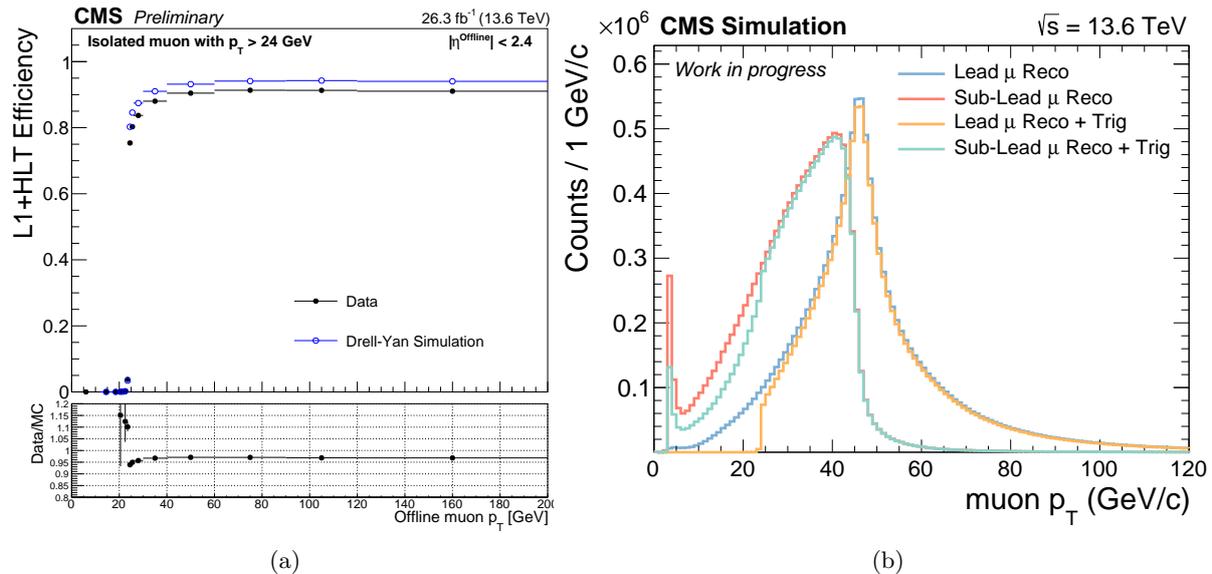


Figura 3.16: A sinistra l'efficienza di Trigger (da [16]). A destra le distribuzioni di  $p_T$  prima e dopo la simulazione del trigger.

### Purezza e ID

Non sempre i muoni che vengono identificati come tali in seguito alla ricostruzione sono dei veri muoni, ossia può capitare che per errore una particella venga ricostruita come un muone quando in realtà non lo è. Questi si dicono *fake*, e tanti più fake ci sono nel campione, tanto più la sua *purezza* (che si può immaginare come proporzionale alla percentuale di muoni "veri" all'interno del campione) diminuisce. Si usano allora degli algoritmi in fase di ricostruzione volti ad ottimizzare la purezza, ossia a ridurre il più possibile il numero di fake. Ad ogni muone viene quindi associata una probabilità che questo sia un muone vero. Questa probabilità viene data sotto forma di livelli di confidenza (o ID), e nello specifico si hanno tre livelli di confidenza: **Loose**, **Medium** e **Tight**, l'uno più stringente dell'altro. Quindi un muone identificato come **Tight** ha più probabilità rispetto ad uno identificato come **Medium** di essere un vero muone, e analogamente per un muone **Loose**. Più stringente è l'ID richiesto però e più diminuisce l'efficienza, poiché i muoni **Tight** sono in numero minore dei muoni **Medium**, che a loro volta sono meno dei muoni **Loose** (Fig. 3.17b).

Si è proceduto allora a studiare le combinazioni di ID per i Leading e Sub-Leading muon, che sono riportati nel grafico 3.17a. Come si vede il caso a efficienza massima è quello in cui entrambi i muoni hanno ID **Loose**, ma la probabilità che siano dei veri muoni sarà minore, e quindi aumenta il rischio che nel campione ci siano dei fake. Viceversa per il caso in cui hanno entrambi ID **Tight**, dove però l'efficienza è minima.

Di questi nove casi se ne sono selezionati due:

- il Leading muon è **Tight** e il Sub-Leading muon è **Medium** (abbreviato in **TM**). Corrisponde alla cella in mezzo nella colonna di destra del Grafico 3.17a.
- entrambi i muoni sono **Tight** (**TT**), e corrisponde alla cella più in alto nella colonna di destra del Grafico 3.17a.

Nei casi in cui le due ipotesi riportate sopra erano vere si sono riempiti due istogrammi separati della massa invariante, su cui in seguito si è fatto un fit gaussiano, i cui risultati sono riportati nella successiva Tab. 3.5.

Come si può vedere nel caso **Tight-Tight** la risoluzione è leggermente migliore che nel caso **Tight-Medium**, e per questo si è scelto di imporre come selezione finale che entrambi i muoni siano **Tight**.

L'efficienza di questo taglio è  $\varepsilon_{ID} = \frac{N_{ID}}{N_G} = 34.4\%$ .

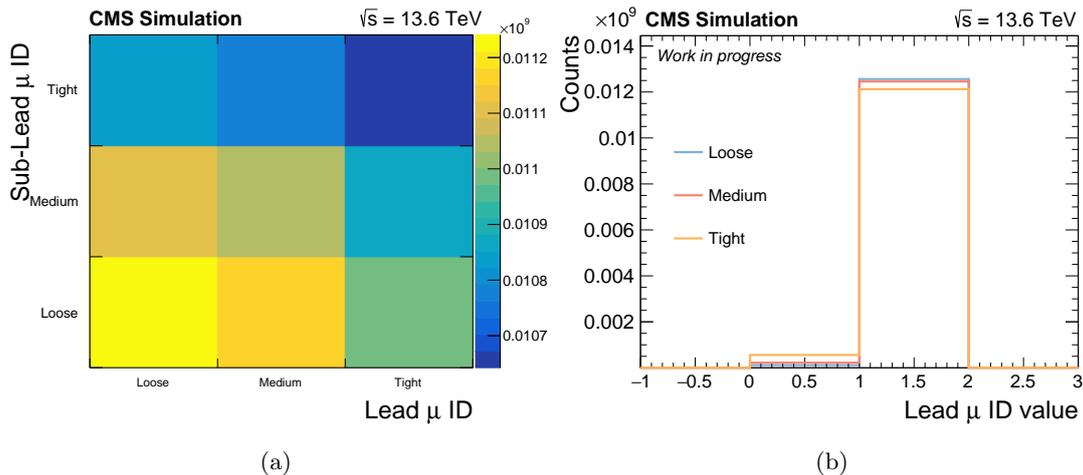


Figura 3.17: A sinistra le possibili combinazioni di ID per Leading e Sub-Leading muon. A destra la statistica degli ID (per semplicità solo del Leading  $\mu$ ).

Selezione	$m_{\mu\mu}$ ( $\text{GeV}/c^2$ )	Scarto %	Risoluzione ( $\text{GeV}/c^2$ )	Efficienza (%)
nessun ID	$90.9118 \pm 0.0009$	0.302	$2.0959 \pm 0.0009$	36.9
TM	$90.9115 \pm 0.0009$	0.303	$2.0914 \pm 0.0009$	35.2
TT	$90.9120 \pm 0.0008$	0.302	$2.0901 \pm 0.0009$	34.4

Tabella 3.5: Fit gaussiano nell'intervallo  $[87,95]$   $\text{GeV}/c^2$ . Lo scarto è stato calcolato come nell'equazione 3.3 rispetto allo stesso valore. Si noti che la prima riga è uguale a quella della Tabella 3.4.

### Isolamento

Infine si è studiato l'isolamento dei muoni; l'isolamento di una particella viene definito come segue. Data la traccia di una particella, si considera un cono attorno ad essa individuato da un certo valore di  $\Delta R$ , dove questo è lo stesso  $\Delta R$  definito in Sez. 3.3.1 (valore tipici sono ad esempio 0.4 o 0.3). Dopodiché si sommano i momenti trasversi di tutte le particelle ricostruite contenute nel cono stesso (esclusa la particella al centro del cono). Tanto più questa somma è piccola, tanto più la particella si dice *isolata* e in genere si richiama che sia minore di un certo valore di soglia. Spesso è utile anche considerare l'isolamento relativo, ossia l'impulso totale appena descritto diviso per il  $p_T$  della particella presa in considerazione:

$$\text{Iso} = \frac{\sum_i p_T^i (\Delta R(\mu) < 0.3)}{p_T(\mu)} \quad (3.4)$$

Si possono anche separare le componenti all'interno del cono: si può calcolare il  $p_T$  totale considerando solo le particelle cariche (isolamento indicato solitamente con **Charged**), indicato nel seguito con  $\text{Iso}_{\text{Chg}}$ , oppure considerare i contributi da tutte le particelle (isolamento **All**), indicato con  $\text{Iso}_{\text{All}}$ . In quest'ultimo caso si usano informazioni dai calorimetri (ECAL e HCAL) per ricostruire gli impulsi trasversi delle particelle neutre. Nei campioni analizzati è stata attuata questa separazione, con  $\Delta R = 0.3$  per la componente carica e  $\Delta R = 0.4$  per la componente All. Le rispettive distribuzioni si possono vedere in Figura 3.18a.

Per la successiva analisi sono stati scelti due diversi valori di soglia, sia per  $\text{Iso}_{\text{Chg}}$  che per  $\text{Iso}_{\text{All}}$ , pari a 0.25 e 0.02. Inoltre si sono imposti gli stessi valori sia per Leading che per il Sub-Leading Iso: ciò è giustificato dal grafico di correlazione 3.18b, dove si vede che le distribuzioni sono simmetriche rispetto a 0.02. Si trova un risultato simile per  $\text{Iso}_{\text{Chg}}$ .

Si hanno in totale quindi quattro casi distinti. In ognuno di questi si è riempito un istogramma della massa invariante se la rispettiva condizione sull'isolamento era verificata e di questi si è fatto un fit come nelle precedenti sezioni, i cui risultati sono riportati in Tabella 3.6. A titolo di esempio, nella seconda colonna la scrittura  $\text{Iso}_{\text{All}} < 0.25$  significa  $\text{Iso}_{\text{All}} < 0.25$  sia per il Leading  $\mu$  che per il Sub-Leading  $\mu$ .

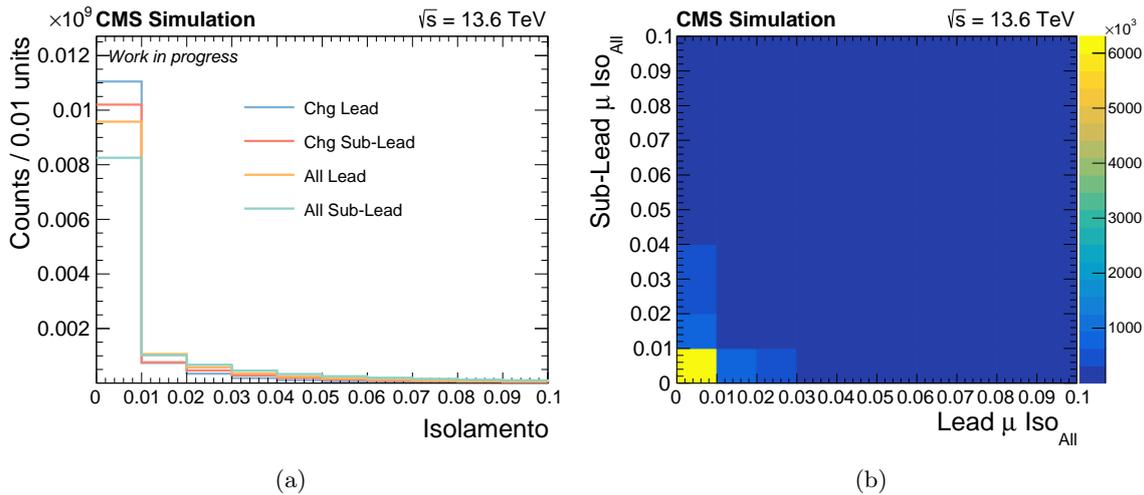


Figura 3.18: A sinistra le distribuzioni di isolamento All e Charged per Leading e Sub-Leading muon. A destra plot di correlazione per Lead e Sub-Lead  $\mu$  (per semplicità solo di Iso<sub>All</sub>).

Selezione	$m_{\mu\mu}$ (GeV/ $c^2$ )	Scarto %	Risoluzione (GeV/ $c^2$ )	Efficienza (%)
Nessuna richiesta	$90.9120 \pm 0.0008$	0.302	$2.0901 \pm 0.0009$	34.4
Iso <sub>All</sub> < 0.25	$90.9130 \pm 0.0009$	0.301	$2.089 \pm 0.001$	33.4
Iso <sub>Chg</sub> < 0.25	$90.9122 \pm 0.0009$	0.302	$2.0900 \pm 0.0009$	34.2
Iso <sub>All</sub> < 0.02	$90.958 \pm 0.001$	0.252	$2.071 \pm 0.001$	22.2
Iso <sub>Chg</sub> < 0.02	$90.925 \pm 0.001$	0.288	$2.088 \pm 0.001$	28.5

Tabella 3.6: Fit gaussiano nell'intervallo [87,95] GeV/ $c^2$ . Lo scarto è stato calcolato come nell'equazione 3.3 rispetto allo stesso valore. Si noti che la prima riga è uguale a quella della Tabella 3.5.

Come si vede la condizione dove si ha il miglior risultato è quella con Iso<sub>All</sub> < 0.02, ossia richiedendo che la somma dei vari  $p_T$  all'interno del cono sia minore del 2% del  $p_T$  dei due muoni. Questa è la condizione più stringente, ed infatti è anche quella con la minore efficienza, data da  $\varepsilon_{\text{Iso}} = \frac{N_{\text{Iso}}}{N_G} = 22.2\%$ .

### 3.4 Analisi del segnale (ggF e VBF)

Si dà ora una descrizione dell'analisi svolta sul segnale, quindi sulle simulazione del decadimento  $h \rightarrow \mu^+ \mu^-$  nei due diversi canali ggF e VBF. Si descriveranno le selezioni applicate, anche qui volte a ottimizzare la risoluzione in massa invariante, e la cinematica dei muoni e del bosone di Higgs ricostruito a partire da questi. Infine verrà fatto un confronto con il processo di fondo precedentemente analizzato.

#### 3.4.1 Selezioni

Le selezioni sui muoni definite nelle sezioni precedenti derivate dall'analisi di eventi  $Z \rightarrow \mu^+ \mu^-$  sono state applicate anche al campione  $h \rightarrow \mu^+ \mu^-$ . Nel seguito si riporta una Tabella (3.7) con un riassunto di tutte di tutte le selezioni applicate a Reco Level, sia per il segnale (ossia ggF e VBF) che per il fondo.

Per il DY le efficienze sono state calcolate rispetto agli eventi  $Z \rightarrow \mu^+ \mu^-$  a livello di generatore, che si ricorda essere un terzo degli eventi totali del campione di Drell-Yan (quindi un terzo del valore riportato in Tabella 3.1). Per i due canali del segnale le efficienze sono state calcolate sempre rispetto agli eventi a Gen Level in cui  $h \rightarrow \mu^+ \mu^-$ , che però in questo caso coincidono con gli eventi totali (ossia i valori riportati sempre nella Tab. 3.1). Come si può vedere le selezioni risultano essere più efficienti per il segnale che non per il fondo. Si noti infine che non vi è una riga relativa alla selezione sul  $p_T$  costretto alla Beamsport poiché questa non riduce il numero di eventi.

Efficienze di selezione (%)			
Selezione	DY	ggF	VBF
almeno 2 $\mu$ a Reco	53.5	75.7	79.3
solo 2 $\mu$ di carica opposta	41.0	61.1	62.7
Trigger	36.9	59.2	60.9
ID	34.4	55.8	57.3
Isolamento	22.2	40.8	44.5

Tabella 3.7: Riassunto delle efficienze risultanti dalle selezioni applicate ai processi di fondo e segnale.

Di seguito si riportano i grafici di massa invariante sia per fondo che segnale con tutte le selezioni applicate, dove si può vedere l'effetto dei singoli tagli sulla distribuzione di  $m_{\mu\mu}$ . Ovviamente per il segnale ci si aspetta una distribuzione simile ma con il picco di eventi situato intorno alla massa del bosone di Higgs, ossia a circa 125  $\text{GeV}/c^2$ .

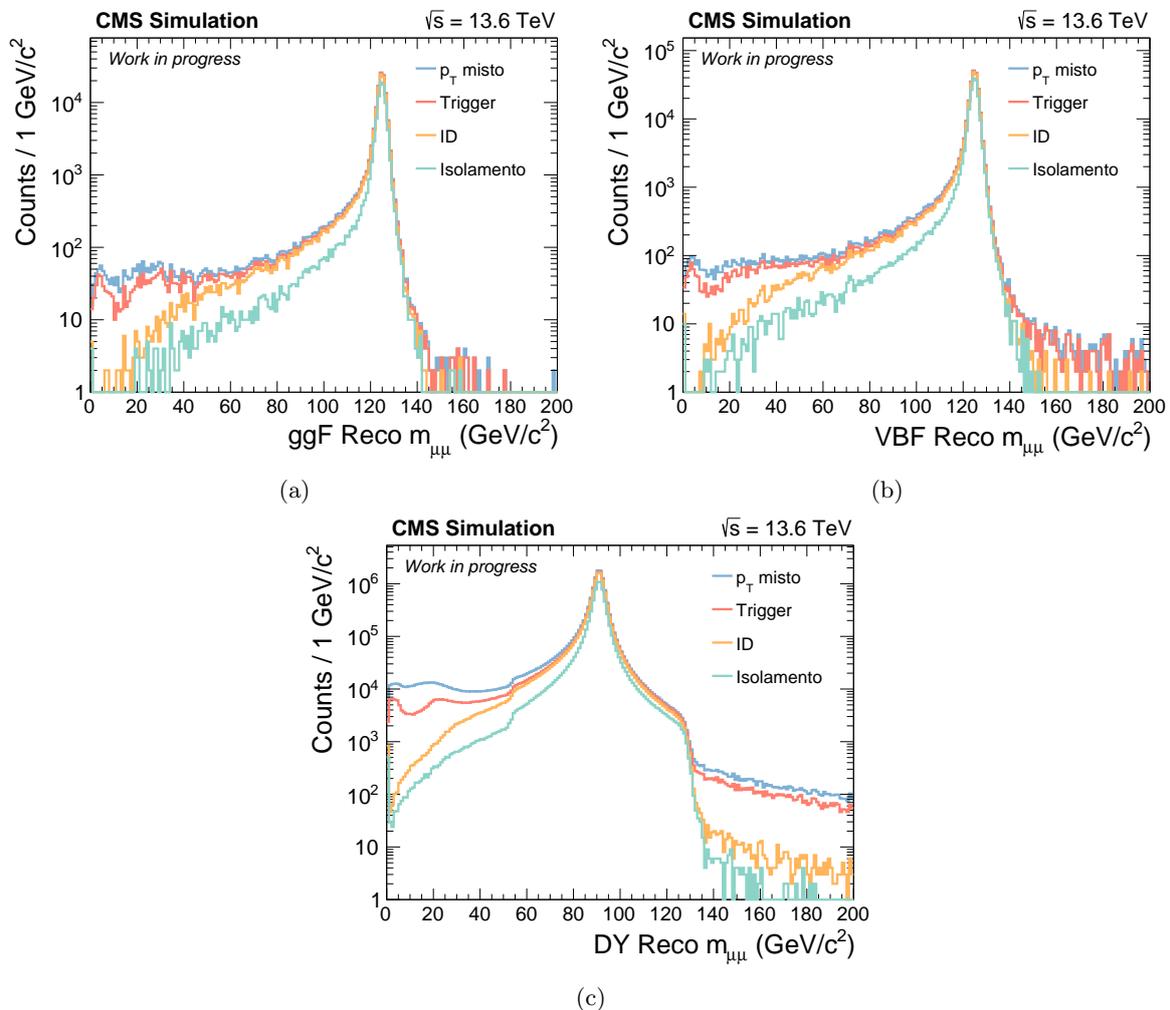


Figura 3.19: Per il DY il grafico in blu indicato come “ $p_T$  misto” è lo stesso che compare in Figura 3.13c in rosso. Con Trigger si intende la simulazione del Trigger, con ID si intende la richiesta Tight-Tight e con Isolamento la richiesta  $\text{ISO}_{\text{All}} < 0.02$ .

Dalle figure si può vedere come le diverse condizioni imposte sui campioni hanno l'effetto di ridurre gli eventi nelle code, ossia gli eventi con  $m_{\mu\mu}$  molto grande o molto piccola, cosa che si ripercuote poi nel

calcolo della risoluzione. Si vede anche come la richiesta sull'isolamento sia quella più stringente.

Poiché le efficienze riportate in Tabella 3.7 hanno dei valori abbastanza bassi si è scelto un ordine alternativo di applicazione dei tagli sopra riportati, col fine di migliorare l'efficienza senza peggiorare troppo la risoluzione su  $m_{\mu\mu}$ . Nello specifico si è proceduto nel seguente modo: invece che selezionare i primi due muoni a Reco Level come Leading e Sub-Leading e trascurare fin dall'inizio tutti gli altri muoni con  $p_T$  minore, si sono studiati tutti i muoni e si sono tenuti solo quelli che passassero le richieste di ID e Isolamento fatte prima. Per quelli che rispettavano queste condizioni si è implementata la separazione tra  $p_T$  e  $p_T$  costretto a seconda del valore del  $\chi^2$ . Dopodiché si è mantenuto l'intero evento se il muone con  $p_T$  maggiore soddisfaceva a  $p_T > 25$  GeV/c. Infine si è mantenuto l'evento se i primi due  $\mu$  (ordinati per  $p_T$  decrescente) avevano carica opposti. Se ciò era verificato, questi venivano identificati come Leading e Sub-Leading muon.

Di questi eventi poi si è proceduto a graficare la cinematica e a ricostruire la  $Z$  (rispettivamente  $h$ ) per il fondo (per il segnale) dalla coppia di muoni come descritto precedentemente. Per i grafici di massa invariante si è fatto un fit gaussiano per trovare i valori definitivi per risoluzione e  $m_{\mu\mu}$ .

Si riporta quindi nella seguente Tabella 3.8 le nuove efficienze trovate applicando il procedimento appena descritto e le risoluzioni in massa invariante.

Processo	Efficienza (%)	$m_{\mu\mu}$ (GeV/ $c^2$ )	Risoluzione (GeV/ $c^2$ )
DY	26.8	$90.957 \pm 0.001$	$2.074 \pm 0.001$
ggF	48.8	$124.791 \pm 0.006$	$1.590 \pm 0.006$
VBF	54.5	$124.774 \pm 0.004$	$1.648 \pm 0.004$

Tabella 3.8: Risultati del fit gaussiano su  $m_{\mu\mu}$  nel range [87,95] GeV/ $c^2$  (risp. [121,128] GeV/ $c^2$ ) per il DY (segnale) e efficienza del rispettivo istogramma.

Si può vedere un considerevole aumento dell'efficienza, che migliora del 20% per il DY e di circa il 19% per il segnale, a fronte di un peggioramento praticamente nullo della risoluzione (0.13% per DY, 0.63% per ggF e 0.48% per VBF).

### 3.4.2 Cinematica del segnale

Nelle seguenti Figure 3.20, 3.21 e 3.22 si riporta la cinematica del segnale per i canali ggF e VBF. Nelle figure sono riportate le distribuzioni a Gen Level e le corrispondenti distribuzioni a Reco Level a valle di tutte le selezioni. Si noti che in questo caso le distribuzioni sono meno definite essendo la statistica minore di un fattore pari a circa 400 per il ggF e 200 per il VBF.

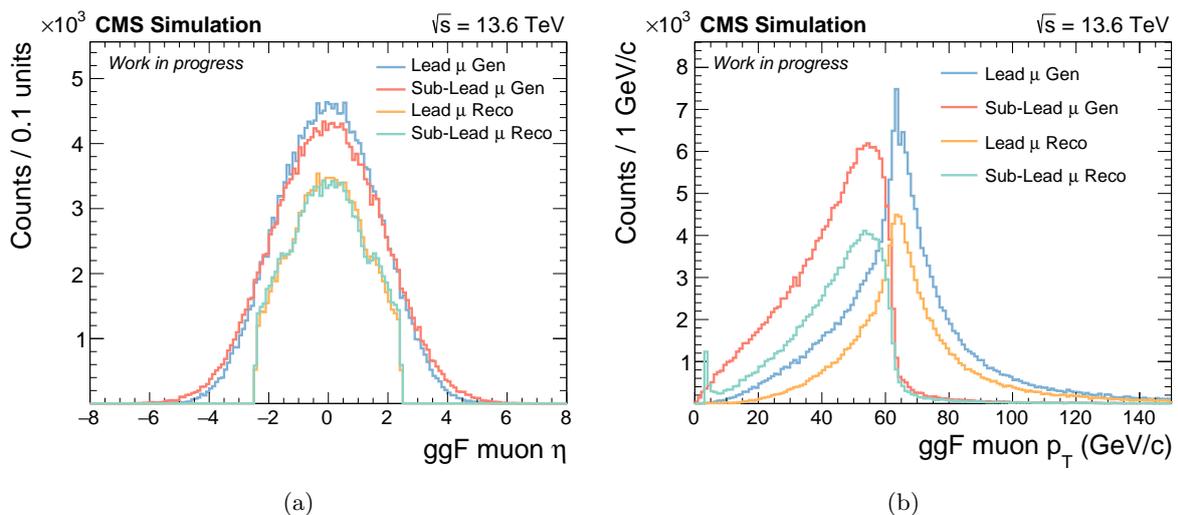
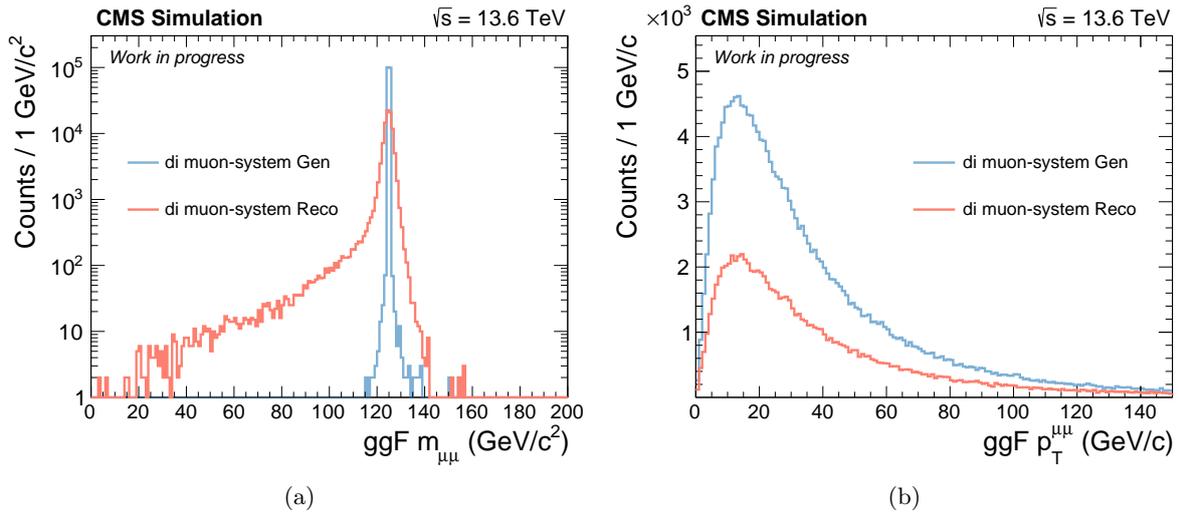
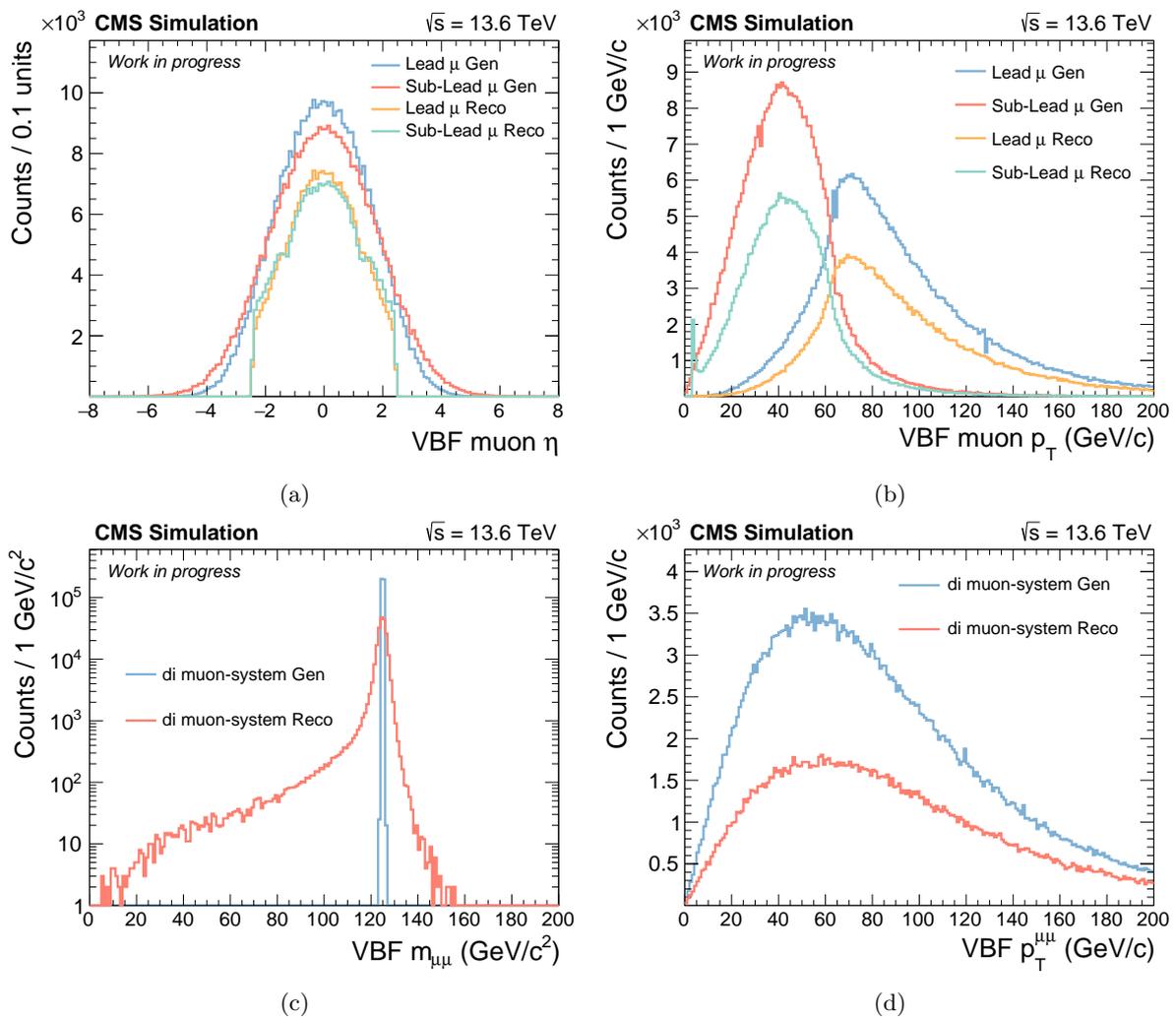


Figura 3.20: Cinematica dei  $\mu$  da  $h$  per il canale ggF.

Figura 3.21: Cinematica del Bosone di Higgs  $h$  ricostruito per il canale ggF.Figura 3.22: Cinematica dei  $\mu$  da  $h$  e di  $h$  ricostruito per il canale VBF.

Come si può vedere la distribuzione in  $\eta$  dei due muoni è molto più stretta rispetto al fondo: infatti gli eventi con  $|\eta| < 2.4$  sono l'86.4% (89.4%) per il ggF (VBF) contro il 66.8% del fondo. I valori di  $p_T$

assunti dai due muoni sono mediamente più alti rispetto al DY, poiché in questo caso vi è più energia a disposizione della coppia di muoni a causa della massa di  $h$  maggiore. Inoltre è molto più evidente la separazione tra Leading e Sub-Leading muon, soprattutto nel canale VBF dove vi sono molti più muoni prodotti ad alti  $p_T$  e dove le distribuzioni presentano dei picchi meno definiti (si notino le scale diverse sull'asse  $x$ ).

Nei grafici di massa invariante si vede che la risoluzione a Gen Level è molto più stretta di quella a Reco, come ci si aspetta essendo la larghezza  $\Gamma$  del bosone di Higgs pari a circa 3 MeV. Si può vedere inoltre che la risoluzione del segnale a livello di ricostruzione è migliore nel canale ggF che non nel VBF (vedi Risoluzioni in Tabella 3.8). Infine dal grafico di  $p_T^{\mu\mu}$  si vede che, a differenza della  $Z$  nel DY,  $h$  non è prodotto a riposo nel piano trasverso; ciò è particolarmente evidente per il canale VBF. Infatti gli eventi in cui si ha  $p_T^{\mu\mu} < 50$  GeV/ $c$  sono il 74.7% per il ggF e solo il 28.5% per il VBF (nel DY questi eventi corrispondevano all'86%). Sempre in  $p_T^{\mu\mu}$  si vedono evidenti differenze anche tra ggF e VBF: queste sono probabilmente dovute al fatto che nel ggF  $h$  è prodotto da solo, mentre nel VBF è prodotto assieme al sistema dei 2 jet, su cui rincula in avanti.

### 3.4.3 Confronto Fondo Segnale

Infine si sono confrontati i risultati fin qui trovati tra fondo e segnale. Nello specifico si è confrontato il ggF con il Drell-Yan poiché quest'ultimo rappresenta il processo di fondo predominante in questa categoria di produzione dell'Higgs. Nel canale del VBF vi sono altri processi (nello specifico due jet nello stato finale con una grande massa invariante  $m_{jj}$ ) che permettono di sopprimere la contaminazione derivante dal Drell-Yan.

In Figura 3.23 si possono vedere gli istogrammi della massa invariante sovrapposti e normalizzati in modo che il loro integrale sia pari a 1. Si può notare che la larghezza del segnale è lievemente minore (come si vede d'altronde anche dai risultati del fit riportati in Tabella 3.8).

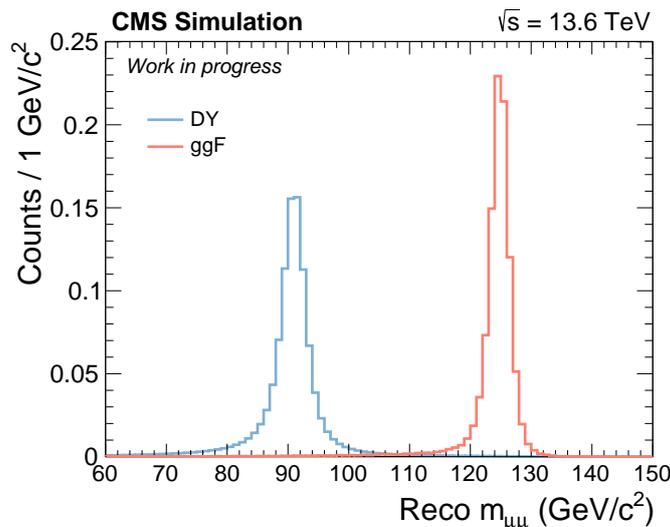


Figura 3.23: Sovrapposizione delle masse invarianti di DY e ggF a Reco Level normalizzate a 1.

Questa istogramma non è però rappresentativo dei dati: infatti il processo  $h \rightarrow \mu^+ \mu^-$  ha una probabilità molto bassa, e quindi ci si aspetta che il picco a 125 GeV/ $c^2$  sia molto più piccolo di quello a 91 GeV/ $c^2$ . Questo non si manifesterebbe facendo una semplice sovrapposizione dei due istogrammi perché i campioni Monte Carlo di fondo e segnale hanno una luminosità integrata equivalente molto diversa tra loro:  $L_{DY} = 15.6851^{+0.19}_{-0.23}$  e  $L_{ggF} = (1.751 \pm 0.087) \times 10^4$  (Tabella 3.1), ossia la luminosità del segnale è circa 1000 volte più grande del DY. Per ottenere un grafico più rappresentativo si sono quindi normalizzati gli istogrammi della massa invariante riportati sopra al valore della luminosità integrata registrata da CMS nel 2022, che ricordiamo essere pari a  $L = 37, 19 \text{ fb}^{-1}$ . Gli istogrammi così normalizzati sono stati poi sovrapposti come si vede in Figura 3.24.

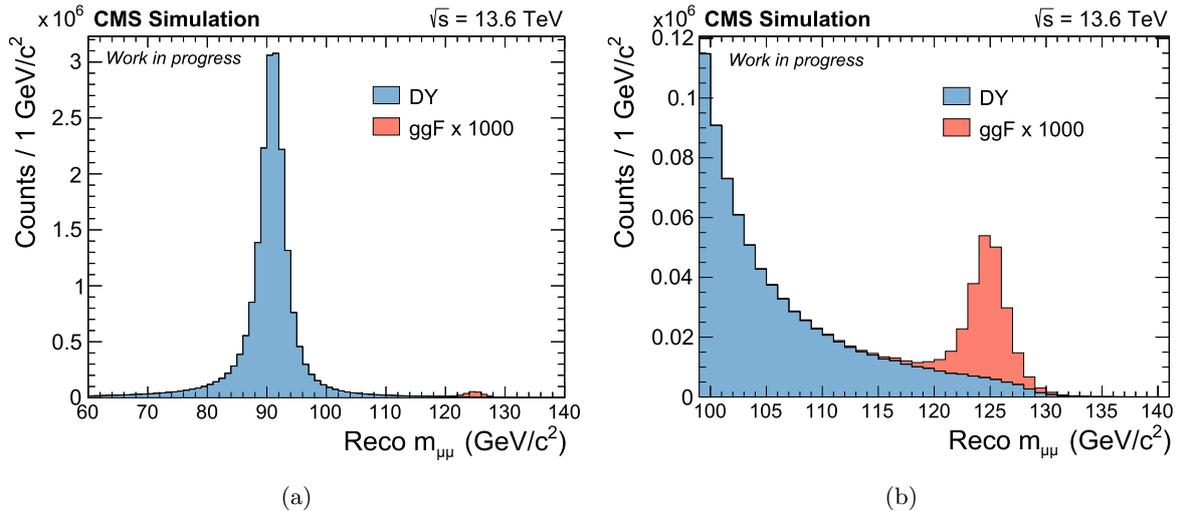


Figura 3.24: Sovrapposizione delle masse invarianti di DY e ggF a Reco Level normalizzate alla luminosità di CMS. A destra uno zoom nel range  $[100,140]$   $\text{GeV}/c^2$ .

È importante sottolineare che l'istogramma del segnale (in rosso) è stato magnificato di un fattore 1000 per renderlo visibile. Quindi nei dati ci si aspetta di trovare un segnale circa 1000 volte più piccolo di quanto riportato qui, e di conseguenza molto difficile da distinguere rispetto al processo di fondo.

Da questi istogrammi normalizzati alla luminosità si può ricavare anche una stima della significanza, ossia del rapporto tra segnale e fondo in un intorno del picco del segnale, quindi in un intorno di 125  $\text{GeV}/c^2$ . Per la significanza si è usato come stimatore  $\frac{S}{\sqrt{B}}$ , dove con  $S$  si indica l'integrale dell'istogramma del segnale in un certo range e con  $B$  si indica l'integrale dell'istogramma del fondo nello stesso range. Come range si è scelto l'intervallo  $[m_{\mu\mu} - 3\sigma, m_{\mu\mu} + 3\sigma]$ , con  $m_{\mu\mu}$  media del fit gaussiano del ggF e  $\sigma$  risoluzione dello stesso fit (quindi Risoluzione e  $m_{\mu\mu}$  della seconda riga nella Tabella 3.8). Esplicitamente:  $[120.02, 129.56]$   $\text{GeV}/c^2$ . Si è trovato quindi un valore pari a

$$\frac{S}{\sqrt{B}} = 0.752 \quad (3.5)$$

Come si vede il fattore trovato è abbastanza basso, e ciò conferma quanto trovato nell'istogramma 3.24a, ossia che la probabilità del segnale è molto più bassa rispetto a quella del fondo, e quindi molto più difficile da osservare.

# Capitolo 4

## Conclusioni

In questo lavoro si sono analizzate le simulazioni relative alla presa dati del rivelatore CMS nell'anno 2022, rivolte alla ricerca del decadimento  $h \rightarrow \mu^+ \mu^-$ . In queste simulazioni si sono studiati i due canali più probabili ad LHC per la produzione del bosone di Higgs, ossia la fusione di gluoni (ggF) e la fusione di bosoni vettori (VBF), e si è considerato come contributo principale al processo di fondo il Drell-Yan. Tramite lo studio di queste simulazioni si è definita una strategia minima di analisi con l'obiettivo di ottimizzare il più possibile la risoluzione sulla massa invariante  $m_{\mu\mu}$  dei due muoni in stato finale. I valori finali ricavati da questa analisi della risoluzione e della massa invariante (con le rispettive efficienze di selezione tra parentesi) sono:

- **DY:**  $m_{\mu\mu} = (90.957 \pm 0.001) \text{ GeV}/c^2$  e  $\sigma_{\text{DY}} = (2.074 \pm 0.001) \text{ GeV}/c^2$  (26.8%).
- **ggF:**  $m_{\mu\mu} = (124.791 \pm 0.006) \text{ GeV}/c^2$  e  $\sigma_{\text{ggF}} = (1.590 \pm 0.006) \text{ GeV}/c^2$  (48.8%).
- **VBF:**  $m_{\mu\mu} = (124.774 \pm 0.004) \text{ GeV}/c^2$  e  $\sigma_{\text{VBF}} = (1.648 \pm 0.004) \text{ GeV}/c^2$  (54.5%).

La strategia di analisi qui descritta può essere poi ulteriormente migliorata, considerando ad esempio tutte le coppie di muoni a livello di Ricostruzione e non solo i primi due muoni con  $p_T$  più grande, e imponendo una selezione sui valori di massa invariante, o implementando una selezione sul parametro d'impatto dei muoni presenti in stato finale.

L'analisi di queste simulazioni è altresì utile per la calibrazione dei dati, azione necessaria prima di poter passare all'analisi degli stessi. Un esempio è quanto accennato nella sezione 3.3.4 per quanto riguarda l'efficienza di Trigger, in cui bisogna tener conto di uno scale factor che riassume la differenza tra simulazione e dati.

Infine, come si è visto durante l'analisi, il segnale è molto tenue a causa della bassa probabilità di decadimento  $h \rightarrow \mu^+ \mu^-$  e quindi sarà necessario usare tutti i  $250 \text{ fb}^{-1}$  di dati raccolti nel corso del Run3 per avere quanta più statistica possibile così da migliorare la significanza del segnale.



# Bibliografia

- [1] CMS collaboration, *Evidence for Higgs boson decay to a pair of muons*, JHEP 01 (2021) 148
- [2] CMS collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B 716 (2012) 30
- [3] Alessandro Bettini, *Introduction to elementary particle physics*, Cambridge University Press, 2014
- [4] Lyndon Evans and Philip Bryant *LHC Machine*, Journal of Instrumentation, 2008 JINST 3 S08001
- [5] 2022 proton-proton collisions at 13.6 TeV, [https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults#2022\\_proton\\_proton\\_collisions\\_at](https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults#2022_proton_proton_collisions_at)
- [6] The CMS Collaboration et al, *The CMS experiment at the CERN LHC*, 2008 JINST 3 S08004
- [7] Alessandro De Angelis, Mário Pimenta *Introduction to Particle and Astroparticle Physics*, Springer, 2015
- [8] Milos Dordevic, *The CMS Particle Flow Algorithm*, EPJ Web Conf. 191 02016 (2018), DOI: 10.1051/epjconf/201819102016
- [9] CMS collaboration, *The CMS trigger system*, 2017 JINST 12 P01020.
- [10] F. Krauss, S. Navas, P. Richardson and T. Sjöstrand, *Monte Carlo Particle Numbering Scheme*, 2019, <https://pdg.lbl.gov/2019/reviews/rpp2019-rev-monte-carlo-numbering.pdf>
- [11] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B 716 (2012) 1, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.
- [12] CMS Collaboration, *Observation of a new boson with mass near 125 GeV in pp collisions at  $\sqrt{s} = 7$  and 8 TeV*, JHEP 06 (2013) 081, doi:10.1007/JHEP06(2013)081, arXiv:1303.4571.
- [13] Lazar Markovic, Roberto Covarelli, *Standard Model Cross Sections for CMS at 13.6 TeV*, 2023
- [14] LHC Higgs Working Group, *Ad interim recommendations for the Higgs boson production cross sections at  $\sqrt{s} = 13.6$  TeV*, 2023
- [15] R.L. Workman et al. (Particle Data Group), *Gauge and Higgs Boson Table*, Prog.Theor.Exp.Phys. 2022, 083C01 (2022) and 2023 update
- [16] CMS Collaboration, *Performance of Soft Drop Mass Jet High Level Trigger at  $\sqrt{s} = 13.6$  TeV in Run 3*, 2023



# Ringraziamenti

Anche se una tesi può sembrare il frutto del lavoro di un'unica persona, questa visione sfiora solo l'aspetto superficiale di ciò che in realtà una tesi è. Essa inevitabilmente si porta dietro i contributi di tante persone. Sento di dover dire grazie a tutte le persone che mi hanno accompagnato durante questo viaggio - iniziato ormai quasi quattro anni fa e di cui questa tesi rappresenta solo un capitolo - ispirandomi e supportandomi ogni giorno. A loro devo la mia gratitudine perché, direttamente o indirettamente, mi hanno dato la forza di andare avanti e di arrivare fin qui.

Alla mia relatrice Mia e al mio correlatore Pierluigi, che sono stati quanto di meglio potessi sperare di trovare, e che in questi mesi non hanno mai esitato ad aiutarmi durante tutti i momenti di insicurezza e difficoltà.

A mia mamma e mio papà, che hanno reso possibile tutto questo e che non mi hanno mai imposto di seguire altre vie se non quella che io ho scelto.

A Caterina, che ogni giorno è per me fonte di ispirazione.

Agli amici di una vita, in particolare Domenico, Antonio, Giuseppe, Armando, Gianvito e Mario, che nonostante la distanza e i silenzi porto sempre con me.

Ai compagni di corso, da quelli con cui ho scambiato poche parole prima di un esame a quelli con cui ho passato ore di studio sui libri, che hanno contribuito a rendere più leggero questo cammino tortuoso.

Al mio prof. Nicola, che ha acceso in me la passione per questa difficile ma affascinante materia che è la Fisica.