



SAPIENZA
UNIVERSITÀ DI ROMA

Tesi di Laurea in Fisica

Ricostruzione di tracce in un telescopio Čerenkov sottomarino per neutrini astrofisici di alta energia

**Facoltà di
Scienze Matematiche, Fisiche e Naturali**

Candidato
Dario Benvenuti
matr. 690486

**Relatore
Antonio Capone**

**Correlatore
Fabrizio Lucarelli**

A.A. 2006/2007

Ringraziamenti

Desidero ringraziare il Professor Antonio Capone per la disponibilità, i preziosi consigli e gli stimolanti incoraggiamenti; Fabrizio Lucarelli e Tommaso Chiarusi per l'impegno profuso nell'indispensabile aiuto offertomi per portare a termine questo lavoro; la collaborazione NEMO, in particolare il gruppo di Roma.

Ringrazio anche mia moglie Olivia, per la pazienza mostrata nei weekend e nelle serate sacrificati per lo studio, e i miei genitori che da sempre mi sostengono nello studio e nella vita.

Un ringraziamento è dovuto anche alla azienda dove lavoro, Elettronica S.p.A., che mi ha agevolato nel portare a termine i miei studi.

Indice

Introduzione.....	4
Cap. 1 Raggi cosmici	6
1.1 La scoperta dei raggi cosmici.	6
1.2 Spettro dei raggi cosmici	7
1.3 Composizione dei raggi cosmici	10
1.4 Generazione dei raggi cosmici	12
1.5 Meccanismi di accelerazione.....	14
1.6 Sciami atmosferici	16
1.7 Sorgenti astrofisiche di neutrini.....	18
Cap. 2 I telescopi a neutrini	24
2.1 Schema di principio	24
2.2 Interazione dei neutrini con la materia	26
2.3 Propagazione dei muoni	29
2.4 Effetto Čerenkov	32
2.5 Fotomoltiplicatori	34
2.6 Efficienza di rivelazione.....	36
2.7 Telescopi a neutrini nel Mediterraneo	37
2.7.1 ANTARES.....	37
2.7.2 NEMO	39
2.8 Altri Esperimenti	42
2.9 Eventi attesi e fondo	44
Cap. 3 Ricostruzione delle tracce di muoni con algoritmi standard	48
3.1 Generazione delle tracce: OPNEMO.....	48
3.2 Ricostruzione geometrica degli eventi	51
3.3 Filtri e trigger.....	60
3.4 Algoritmi di minimizzazione e pre-fit.....	62
3.5 Programma di Ricostruzione	63
3.6 Risultati della ricostruzione OPNEMO	66
Cap. 4 Ricostruzione delle tracce con reti neurali	79
4.1 Introduzione alle Reti Neurali	79
4.1.1 Neuroni	79
4.1.2 Elaborazione neurale	80
4.1.3 Funzioni di attivazione	82
4.1.4 Architetture di reti neurali	83
4.1.5 Metodologie di apprendimento.....	85
4.1.6 Applicazioni	87
4.2 Ricostruzione delle tracce di muoni con reti neurali	87
4.3 Risultati della ricostruzione con reti neurali.....	89
Cap. 5 Ricostruzione delle tracce con algoritmi genetici.....	94
5.1 Introduzione agli Algoritmi Genetici	94
5.1.1 Teoria: perché funziona? Iperpiani negli spazi di ricerca.....	98
5.2 Applicazione per la ricostruzione delle tracce.....	105
5.3 Implementazione della ricostruzione con Algoritmi Genetici.....	107
Cap. 6 Conclusioni.....	119
Appendice.....	121
Bibliografia.....	139

Introduzione

L'astronomia e l'astrofisica sono nate con l'osservazione della radiazione elettromagnetica, inizialmente per ovvie ragioni nella banda del visibile, perché i fotoni essendo neutri viaggiano in linea retta¹ conservando l'informazione della direzione della loro sorgente. Negli anni recenti esperimenti come CGRO e INTEGRAL hanno disegnato mappe del cielo dettagliate delle emissioni di raggi γ , specialmente per le energie più alte. Ma il forte assorbimento dei fotoni dalla materia interstellare, specialmente nelle immediate vicinanze di oggetti come i Nuclei Galattici Attivi (AGN), e l'interazione dei fotoni di alta energia con il fondo intergalattico di fotoni infrarossi, limitano l'orizzonte di visibilità dei raggi γ in energia e in distanza (cfr Figura 1).

Esiste la possibilità di estendere l'osservazione degli eventi di alta energia attraverso la rivelazione dei neutrini. A differenza delle particelle cariche, deviate dai campi magnetici, i neutrini non perdono l'informazione della loro direzione di provenienza e non vengono assorbiti come i fotoni γ .

L'astrofisica con i neutrini può fornire risposte a problemi aperti dell'astrofisica delle alte energie; innanzi tutto può confermare i modelli e i siti di accelerazione dei raggi cosmici, che non possono essere identificati direttamente proprio a causa del rimescolamento delle direzioni ad opera dei campi magnetici galattici ed intergalattici; può gettare luce sulla fisica degli eventi più estremi dell'universo, come gli AGN e i Lampi di Raggi Gamma (GRB), può risolvere il problema della materia oscura se si riveleranno i decadimenti delle Particelle Massive Debolmente Interagenti (WIMPs).

¹ A parte le deformazioni spaziotemporali nelle vicinanze di corpi massivi che hanno dimostrato la teoria di Einstein con l'osservazione della deviazione della luce di una stella in direzione del sole, durante l'eclissi nel 1919, e che causano gli effetti lente nei dintorni di grandi accentramenti di materia (buchi neri) nello spazio

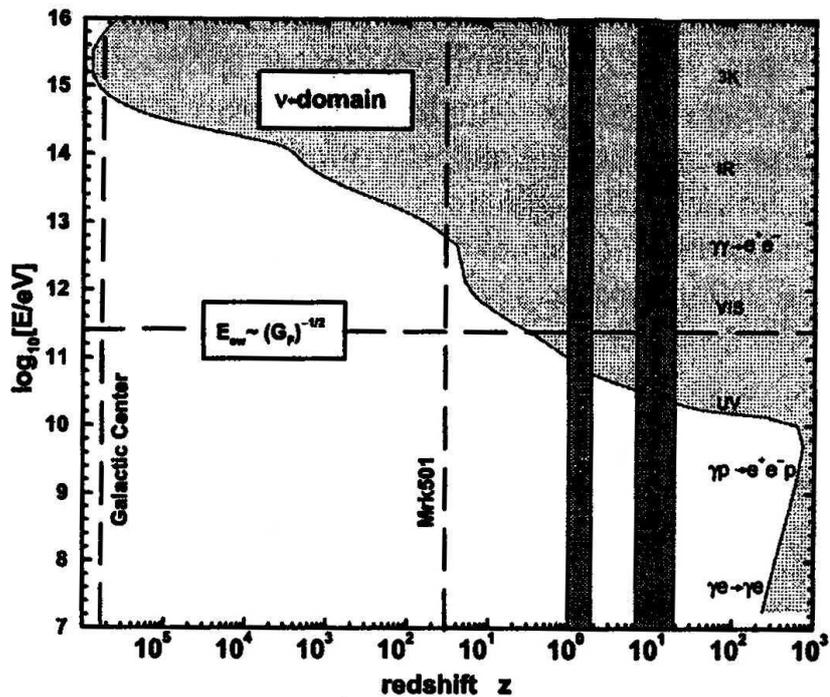


Figura 1 – Orizzonte di visibilità dei raggi γ e dei neutrini in funzione di energia e distanza (redshift) [2].

Un metodo di rivelazione dei neutrini e della loro direzione di provenienza è stato proposto da Markov nel 1961 [1]. I leptoni carichi (muoni) generati per interazione debole di corrente carica da neutrini emettono luce per effetto Čerenkov. Un sistema di fotomoltiplicatori viene utilizzato per misurare il numero di fotoni ed il loro tempo di arrivo, e permette di ricostruire la traccia del muone, e del neutrino.

Per far questo occorre posizionare l'apparato di rivelazione nelle profondità del mare (o dei ghiacci) dove il buio consenta di individuare i pochi fotoni di interesse, e in modo che gli altri raggi cosmici vengano filtrati. Nonostante questi accorgimenti i muoni che sicuramente provengono da un neutrino sono quelli provenienti dal basso: la Terra scherma efficacemente i raggi cosmici, a parte i neutrini, che hanno una cross section estremamente bassa.

La bassa probabilità di interazione fa sì che per rivelare un numero ragionevole di neutrini astrofisici sia necessario creare dei rivelatori di dimensioni enormi, come si vedrà dell'ordine del km^3 .

In questa tesi sono presentati nuovi metodi di ricostruzione delle tracce di muoni, basati sugli algoritmi genetici e sulle reti neurali, e i risultati sono confrontati con gli algoritmi già utilizzati nella collaborazione NEMO.

Cap. 1 Raggi cosmici

1.1 La scoperta dei raggi cosmici.

I raggi cosmici sono l'insieme di particelle cariche e neutre che colpiscono la terra da tutte le direzioni. Essi furono scoperti ai primi del '900, durante lo studio del fenomeno della scarica di elettroscopi isolati.

Questo fenomeno veniva inizialmente attribuito alla radioattività, che era stata scoperta da poco grazie ai lavori di Becquerel, dei coniugi Curie e in seguito di Rutherford, Soddy e collaboratori. Ma la quantità di radioattività necessaria a spiegare il tasso di scarica degli elettroscopi sembrava superiore a quella naturale; inoltre sembrava che la scarica degli elettroscopi fosse più veloce ad alta quota rispetto agli elettroscopi posti sulla superficie terrestre.

Fu Victor Hess nel 1912 (Figura 2) che, con l'aiuto di palloni sonda e di una mongolfiera dimostrò che questa radiazione ionizzante aumentava con la quota e che aveva quindi origine extraterrestre, scoprendo di fatto i raggi cosmici (RC) [3].

Fino allo sviluppo di acceleratori di energia sufficientemente elevata negli anni '50, i RC furono le uniche sorgenti disponibili di particelle di alta energia. Fu proprio durante lo studio dei raggi cosmici che vennero scoperti l'antimateria, con l'osservazione delle prime tracce di positroni da parte di Anderson, Blackett e Occhialini, il muone (da parte di Anderson e Neddermeyer), i mesoni π (predetti da Yukawa nel 1936) e vennero osservati per la prima volta anche i decadimenti di particelle dotate di *stranezza* (K, Λ , Σ e Ξ).



Figura 2 - Esperimento di Victor Hess con la mongolfiera (1912).

1.2 Spettro dei raggi cosmici

I RC primari, quelli cioè che sono distribuiti uniformemente nello spazio e che incidono sugli strati più esterni dell'atmosfera, sono costituiti principalmente da protoni (86-89%), nuclei di He (9%) e per il resto da nuclei più pesanti, elettroni, neutrini e gamma di alta energia [6].

In Figura 3 è riportato il flusso differenziale dei RC, cioè il numero di particelle incidenti per unità di energia, di tempo, di area e di angolo solido. L'andamento del flusso in funzione dell'energia può essere espresso tramite una legge di potenza del tipo [7]

$$\frac{dN}{dE} \propto E^{-\alpha}$$

dove α , l'indice spettrale è pari a

$$\alpha = \begin{cases} 2.7 & \text{per } E < 3 \cdot 10^{15} \text{ eV} \\ 3 & \text{per } 3 \cdot 10^{15} < E < 10^{18} \text{ eV} \\ 2.7 & \text{per } E > 3 \cdot 10^{18} \text{ eV} \end{cases}$$

Le variazioni della pendenza dello spettro dei RC sono più visibili moltiplicando il flusso differenziale per $E^{2.7}$ come nella parte bassa della Figura 3.

Il primo punto di variazione di pendenza (a $E = 3 \cdot 10^{15}$ eV) è detto “ginocchio” (knee) e il secondo (a $E = 10^{18}$ eV) è detto “caviglia” (ankle). L'andamento osservato può essere spiegato ipotizzando che i RC siano accelerati attraverso i meccanismi di accelerazione di Fermi (vedi §1.5). Tale meccanismo prevede per lo spettro “alla sorgente” dei raggi cosmici un andamento proporzionale ad E^{-2} . L'effetto di confinamento dovuto al campo magnetico galattico, più efficiente alle basse energie (leaky box model), distorce lo spettro delle particelle accelerate facendo sì che quelle da noi osservate seguano una legge di potenza del tipo $E^{-\gamma}$ con $\gamma \approx 2.6 \div 2.7$.

Le particelle cariche in un campo magnetico, subiscono la forza di Lorentz

$$F_{Lorentz} = q \mathbf{B} \times \mathbf{v}$$

ortogonale al loro vettore velocità, e compiono quindi traiettorie circolari o a forma di spirale.

Il raggio della traiettoria è chiamato raggio di Larmor o giroradio:

$$R_{Larmor} = \frac{E}{Z|e|B\beta c}$$

che per un protone assume la forma:

$$R[m] = \frac{E[GeV]}{0.3 B[T]}.$$

Per $B = 3 \mu\text{G}$ (campo magnetico medio della galassia), $E = 10^{16}$ eV si ha un giroradio pari a circa 4 pc ($1 \text{ pc} = 3.086 \cdot 10^{16} \text{ m}$).

Basandosi sulle considerazioni precedenti, si può assumere che, almeno fino al ginocchio (cioè fino a 10^{15} - 10^{16} eV), i RC siano prevalentemente di origine galattica: infatti dopo il punto di ginocchio una frazione sempre maggiore dei RC sfugge al campo magnetico

galattico, fino ad arrivare alla caviglia dove il giroradio è maggiore dello spessore del disco galattico²; oltre tale energia bisogna ipotizzare evidentemente un'origine extragalattica.

Il massimo dello spettro si ha intorno a qualche GeV, al di sotto del quale la componente dei RC è sommersa dalle particelle cariche di bassa energia provenienti dal Sole (vento solare).

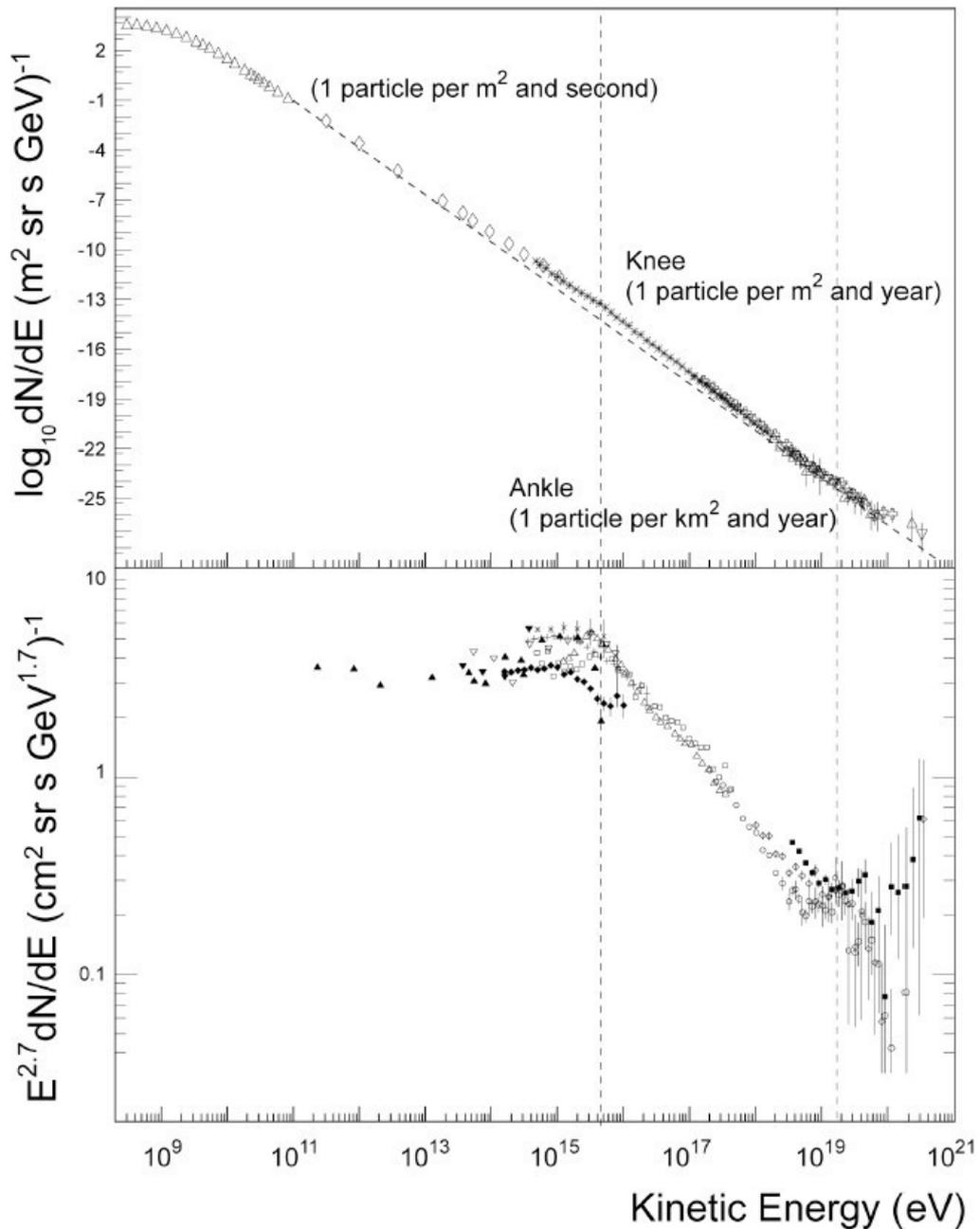


Figura 3 - Spettro di energia dei Raggi Cosmici e andamento moltiplicato per $E^{2.7}$.

² il disco galattico ha un raggio di circa 10 kpc e uno spessore pari a circa 500 pc

Oltre i 10^{20} eV quasi tutti gli esperimenti (a parte AGASA) mostrano una rapida diminuzione del flusso dei RC, recentemente confermata dall'esperimento AUGER [8]. Questo fenomeno di soglia è stato previsto negli anni '60 da Greisen, Zatsepin e Kuzmin ed è noto come effetto GZK dalle loro iniziali. La soglia (GZK cutoff) è dovuta al fatto che i RC di queste energie possono interagire con i fotoni del fondo a microonde (a 2.7 K) per fotoprodurre la risonanza Δ^+ :



La soglia di energia per questa reazione è assai elevata; la conservazione del quadrato del quadrimomento dà infatti

$$E_{Th} = \frac{c^4 (m_{\Delta}^2 - m_p^2)}{4E_{\gamma}} \approx 4 \cdot 10^{19} \text{ eV}$$

nel caso in cui venga prodotta la Δ^+ a riposo.

Il flusso di RC con energie superiori a questa soglia è quindi attenuato durante il percorso, e questo limita la distanza osservabile delle sorgenti più potenti. Nel contempo si osserva però un flusso maggiore di RC per energie immediatamente inferiori alla soglia, e questo è dovuto ai prodotti della reazione (1). Si osservi che tra i prodotti secondari vi sono anche neutrini di altissima energia.

1.3 Composizione dei raggi cosmici

Attraverso esperimenti condotti su satelliti si sono potute indagare le caratteristiche dei RC primari, in particolare la loro composizione.

Dal confronto tra le abbondanze relative dei nuclei pesanti nei RC rispetto alle abbondanze relative nel sistema solare (Figura 4), si possono dedurre importanti indicazioni.

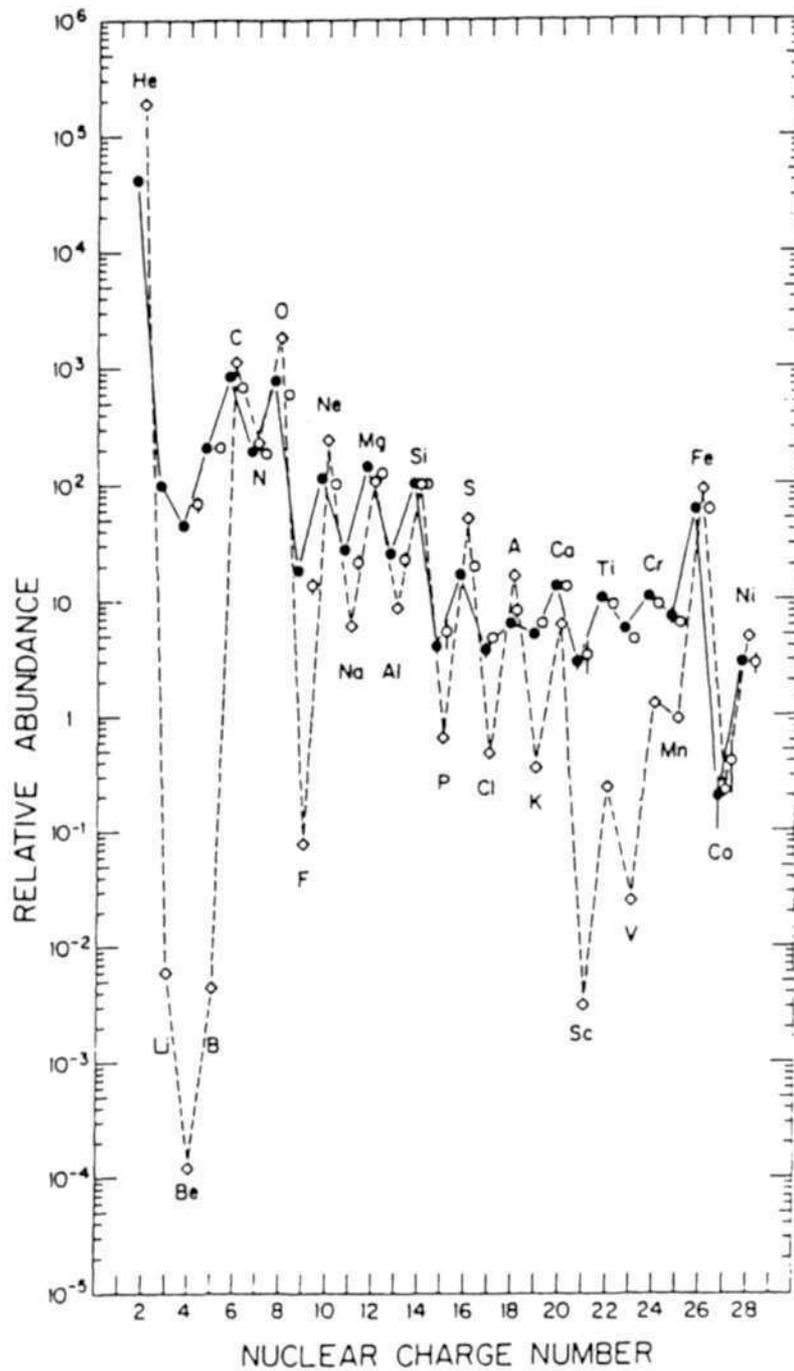


Figura 4 - Abbonanze relative dei nuclei He-Ni nei RC (linea continua) e nel Sistema solare (linea tratteggiata) [6].

Nella figura si vede che in entrambi i casi si ha il tipico effetto pari-dispari dovuto al fatto che i nuclei con Z pari hanno un legame più forte e quindi sono più abbondanti. Invece gli elementi Li, Be, B, e Sc, Ti, V, Cr, Mn, sono vari ordini di grandezza più abbondanti nei RC che nel sistema solare (per il Be si sfiora 10^6). Questi elementi sono infatti praticamente

assenti nei prodotti dei processi di nucleo-sintesi stellare, mentre nei RC sono creati dalla spallazione dei più abbondanti nuclei di C e O (Li, Be, B), e di Fe (Sc, Ti, V, Cr, Mn) presenti nel mezzo interstellare (ISM). Questa spiegazione ha anche il pregio di fornire una indicazione sulla propagazione dei RC e sul loro confinamento nella nostra galassia. Infatti conoscendo la sezione d'urto del processo di spallazione si può risalire allo spessore (X [g/cm²]) di ISM attraversato tra la sorgente e l'osservazione dei RC. Per la maggior parte dei RC lo spessore attraversato risulta compreso tra 5 e 10 g/cm². La densità media del disco galattico (ρ) è dell'ordine di un protone al cm³, quindi i RC hanno viaggiato per una distanza almeno pari a

$$l = \frac{X}{m_p \rho} = 3 \cdot 10^{24} \text{ cm} \approx 1000 \text{ kpc}$$

dove $m_p = 1.67 \cdot 10^{-27}$ kg è la massa del protone. Dato che lo spessore del disco galattico è pari a circa 0.5 kpc e poiché la maggior parte dei RC che osserviamo è di origine galattica, esiste una sorta di confinamento, ad opera del campo magnetico galattico (in media 3 μ G), che costringe i RC a propagarsi per un tempo molto lungo ($\sim 10^7$ anni) all'interno della galassia prima di essere assorbiti o sfuggire nello spazio intergalattico.

1.4 Generazione dei raggi cosmici

La vera origine dei RC rimane ancora incerta, soprattutto a causa del fatto che questi sono deviati dai campi magnetici e ciò impedisce di ricollegare la direzione di arrivo del primario alla sorgente acceleratrice. Nella nostra galassia le sorgenti più probabili sono: le esplosioni di supernovae; eventi violenti sulle superfici di stelle (come le tempeste solari che danno origine a grandi emissioni di plasma); i dischi di accrescimento intorno ad oggetti compatti (buchi neri o di stelle di neutroni) in sistemi binari; le microquasar. Possibili sorgenti extragalattiche sono i nuclei galattici attivi (AGN) e le sorgenti di Lampi di Raggi Gamma (Gamma Ray Bursts GRB) [9].

La possibilità che i RC di origine galattica siano dovuti alle esplosioni di supernovae era già stata avanzata negli anni '30 da Zwicky e Baade, e negli anni '60 è stata corroborata da Ginzburg e Syrovatskii [10]: la densità di energia necessaria per accelerare i RC alle energie osservate è circa $\rho_{\text{CR}} \cong 1 \text{ eV/cm}^3$ ($1.6 \cdot 10^{-13} \text{ J/m}^3$) distribuita su un volume galattico pari a circa $V_{\text{GAL}} = \pi R^2 D \cong \pi (15 \text{ kpc})^2 \cdot 0.2 \text{ kpc} = 1.3 \cdot 10^{60} \text{ m}^3$, per un tempo di permanenza

medio di $3 \cdot 10^6$ anni. Partendo da questi valori si può calcolare la potenza necessaria per accelerare i RC:

$$W_{RC} = \rho_{CR} V_{GAL} / \tau_{CR} \approx 10^{41} \text{ J/yr}$$

D'altra parte, una supernova di tipo II espelle nello spazio circostante una quantità di materia dell'ordine di 10 masse solari ($2 \cdot 10^{31}$ kg) ad una velocità di circa 10^7 m/s, con una frequenza media approssimativamente di una per secolo; questo corrisponde alla potenza

$$W_{SN} \approx 5 \cdot 10^{42} \text{ J/yr}$$

superiore a quella necessaria [7]. Questo è un forte indizio al fatto che una buona parte dei RC di origine galattica siano generati ed accelerati dalle onde d'urto delle supernove (§1.5).

La generazione da parte delle stelle ci è nota a causa del vento solare. Il sole emette una gran quantità di particelle (il vento solare) con bassa energia, con un andamento oscillante legato al ciclo solare di periodo 11 anni. Sulla terra si ha una correlazione dell'abbondanza di RC a bassa energia con il ciclo solare, mentre si ha anticorrelazione degli altri RC. Questo perché il vento solare, essendo formato da plasma in movimento, genera un campo magnetico più forte nei periodi di massima attività, e questo campo magnetico scherma in maniera più efficace i RC provenienti dall'esterno del sistema solare.

Lo stesso campo magnetico terrestre funge da filtro sui RC con energia minore di qualche decina di GeV [7]. La traiettoria effettiva di una particella in avvicinamento alla terra è una spirale, ma l'ordine di grandezza della energia minima per colpire l'atmosfera terrestre può essere calcolata con un modello semplice: consideriamo una particella di carica $Z|e|$, e momento $p = mv$, che viaggia in un cerchio equatoriale di raggio r , intorno ad un dipolo magnetico di momento M . Uguagliando la forza magnetica e quella centrifuga

$$Z|e| |\mathbf{B} \times \mathbf{v}| = \frac{mv^2}{r}$$

e considerando che il campo dovuto ad un dipolo magnetico al suo equatore vale

$$B = \frac{\mu_0 M}{4\pi r^3}$$

si ottiene il raggio di Størmer $r_s = \left(\frac{4\pi B}{\mu_0 M} \right)^{1/3}$. Il valore di soglia del momento della particella

può essere calcolato uguagliando il raggio di Størmer al raggio terrestre, r_E ;

$$\frac{pc}{Z} = \frac{\mu_0 M c |e|}{4\pi r_E^2} = 59.6 \text{ GeV}$$

1.5 Meccanismi di accelerazione

Nel 1949 Fermi pubblicò un articolo nel quale ipotizzava l'accelerazione dei RC attraverso la diffusione delle particelle in nubi di plasma in moto nella galassia.

Questo processo è ora noto come meccanismo di Fermi del secondo ordine, visto che l'incremento medio relativo di energia di ogni nucleone è proporzionale al β^2 della nube di plasma.

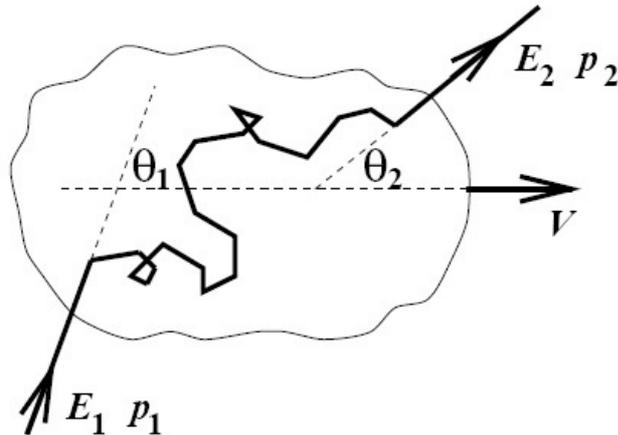


Figura 5 - Accelerazione di una particella in una nube di plasma in moto con velocità V : E_1 (E_2), p_1 (p_2) e θ_1 (θ_2) energia, momento e angolo di incidenza iniziali (finali).

Indicando con E_1 l'energia iniziale della particella e con E_2 l'energia della particella in uscita dalla nube, si ha infatti [6]:

$$\frac{\langle E_2 \rangle}{E_1} = \gamma^2 (1 - \beta \langle \cos \theta_1 \rangle) = \frac{1}{1 - \beta^2} \left(1 + \frac{\beta^2}{3} \right) \cong (1 + \beta^2) \left(1 + \frac{\beta^2}{3} \right) \cong 1 + \frac{4}{3} \beta^2$$

dove $\gamma = \frac{1}{\sqrt{1 - \beta^2}}$, $\beta = \frac{V}{c}$ dove V è la velocità della nube.

Poiché le nubi di plasma hanno velocità piuttosto basse ($\beta \approx 10^{-4}$) questo meccanismo non è in grado di spiegare le alte energie in gioco nei RC, in particolare darebbe luogo ad uno spettro in energia con un indice spettrale troppo elevato [6].

Applicando l'idea dell'accelerazione di Fermi alle onde d'urto si ottiene un guadagno medio di energia proporzionale a β , (meccanismo di Fermi del primo ordine). Un'onda d'urto (shock wave) è una perturbazione che si propaga ad una velocità maggiore di quella del suono nel mezzo considerato. Durante l'esplosione di una supernova ad esempio, una grande quantità di materia (dell'ordine delle 10 masse solari) viene espulsa ad una velocità di circa 10^4 km/s,

molto maggiore della velocità del suono nel mezzo interstellare, creando quindi un'onda d'urto che si propaga radialmente nello spazio. La velocità del fronte d'onda (V_S) dipende dalla velocità della materia espulsa ($V_P \cong 10^4$ km/s) e dal rapporto tra i calori specifici, (trattandosi di gas ionizzato, quindi mono-atomico $\gamma = c_P/c_V = 5/3$); risulta [11]

$$\frac{V_S}{V_P} \cong \frac{4}{3}$$

Le particelle che vengono diffuse dalle irregolarità dei campi magnetici delle nubi di plasma, attraversando il fronte d'onda di shock guadagnano energia, in entrambi i sensi di attraversamento dello shock, poiché la velocità relativa del plasma prima e dopo lo shock è sempre V_P (Figura 6).

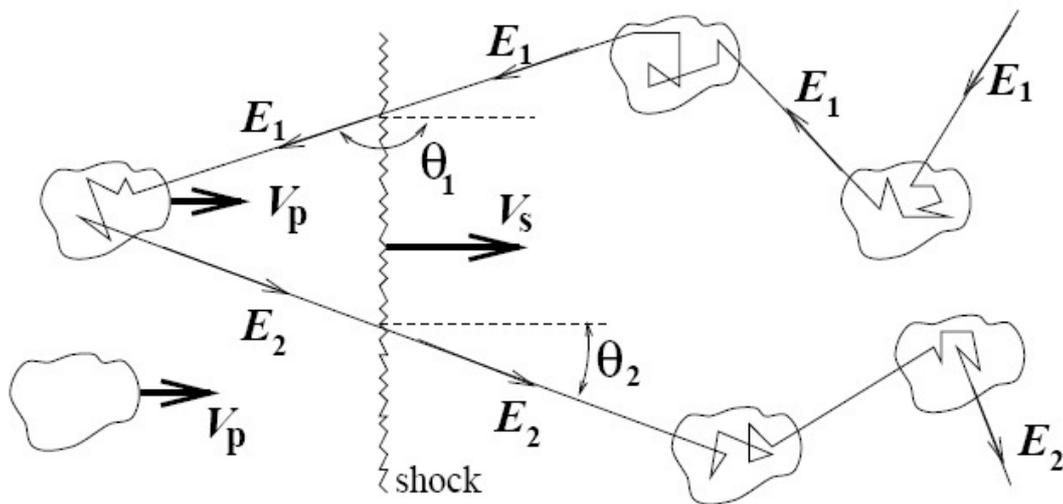


Figura 6 - Schema di interazione di una particella con un'onda di shock. E_1 (E_2) energia iniziale (finale) della particella, V_S velocità dello shock, V_P velocità della materia espulsa [11].

Questo è il motivo per cui nell'onda d'urto si ha un'accelerazione più efficiente. Il guadagno medio di energia per ogni passaggio avanti e indietro è [11]

$$\frac{\langle \Delta E \rangle}{E} \approx \frac{4}{3} \beta, \text{ con } \beta = V_P/c.$$

Mediando sull'angolo di incidenza delle particelle si può calcolare il tasso di attraversamento dello shock e l'indice spettrale dello spettro in energia della popolazione di particelle accelerate. Lo spettro calcolato ha l'andamento descrivibile tramite una legge di potenza

$$\frac{dN}{dE} \propto E^{-2}$$

consistente con quanto osservato nel §1.2.

Durante la propagazione nella galassia le particelle più energetiche hanno maggiore probabilità di sfuggire al campo magnetico galattico. Questo effetto è stato modellizzato matematicamente (Leaky Box Model) e se ne è potuta calcolare la variazione di indice spettrale pari a $0.6 \div 0.7$, che rende l'indice spettrale in accordo con quello osservato, pari a ~ 2.7 per $E < 10^{16}$ eV.

1.6 Sciame atmosferici

Una volta entrato nell'atmosfera terrestre, un raggio cosmico primario percorre al massimo una distanza³ dell'ordine di una decina di km, prima di interagire con un nucleone dell'atmosfera. L'interazione dà inizio ad uno sciame di particelle secondarie (Figura 7), formato da π^\pm , K^\pm , etc. (componente hard) che poi decadono in μ e ν , e da pioni neutri che, decadendo, danno inizio ad uno sciame puramente elettromagnetico (componente soft).

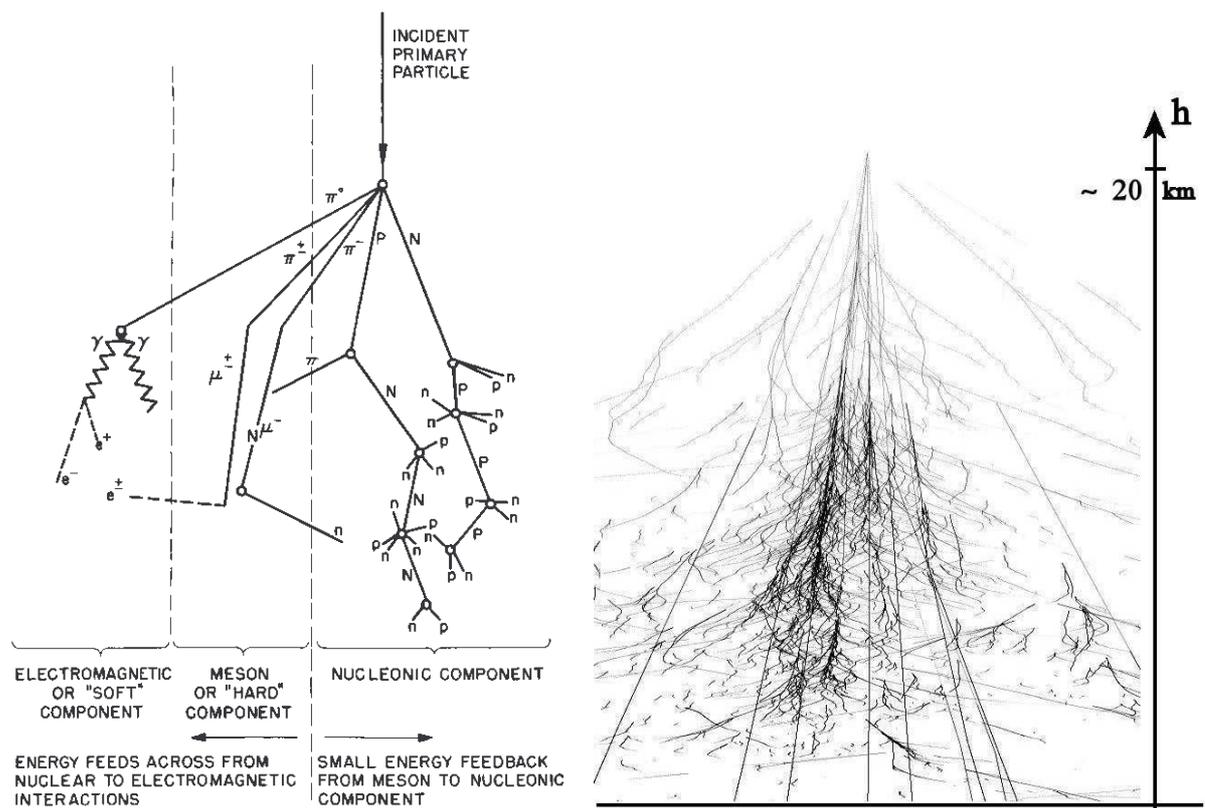


Figura 7 - Sviluppo di uno sciame atmosferico prodotto da raggi cosmici [4] (a sinistra), e simulazione di uno sciame indotto da un protone di energia 1 TeV (a destra).

³ La lunghezza di interazione dipende dall'energia e dalla massa (numero atomico) del primario

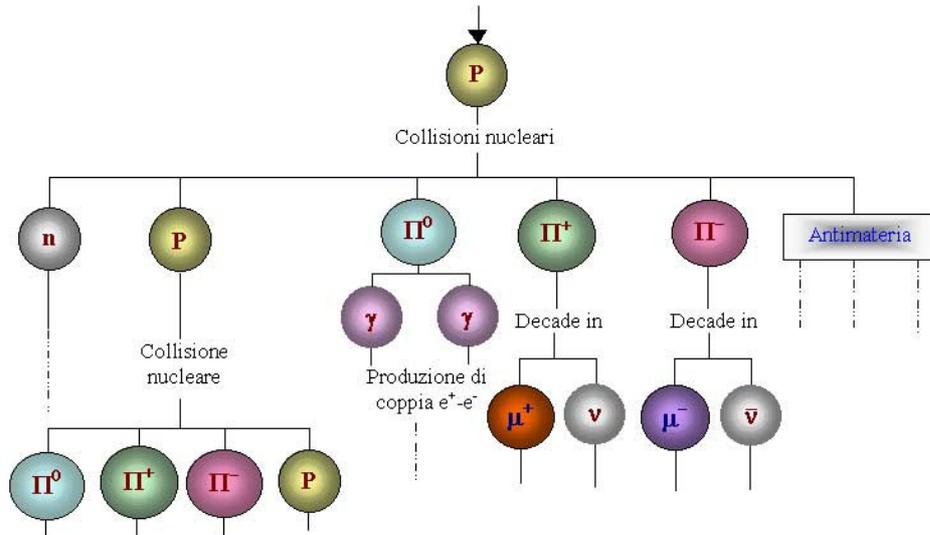


Figura 8 - Particelle secondarie, e loro decadimenti, prodotte dai raggi cosmici primari in una collisione nucleare [5].

I neutrini prodotti in questi sciami sono detti neutrini atmosferici. La alta densità dell'atmosfera (rispetto agli spazi interstellari dove probabilmente hanno origine i neutrini di origine astrofisica) fa sì che una parte dei mesoni prodotti interagisca prima di decadere, facendo sì che la pendenza dello spettro dei neutrini atmosferici sia maggiore di quella che ci aspettiamo nel caso di neutrini astrofisici $\left(\frac{dN}{dE} \propto E_\nu^{-2}\right)$; questo è un bene, perché ad energie sufficientemente elevate ci si aspetta che il flusso di neutrini astrofisici diventi superiore a quello dei neutrini atmosferici (vedi §1.7).

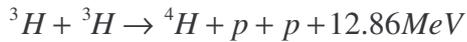
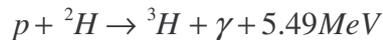
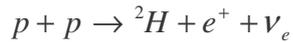
Inoltre questo fenomeno dà origine ad una forma caratteristica del flusso dei neutrini atmosferici rispetto all'angolo zenitale: i mesoni generati da RC provenienti da angoli più radenti l'orizzonte percorrono un tragitto più lungo in zone di bassa densità rispetto a quelli più verticali, avendo quindi una maggiore probabilità di decadere. Il flusso di neutrini avrà quindi un massimo in corrispondenza di angoli zenitali dell'ordine di $80 \div 90^\circ$.

1.7 Sorgenti astrofisiche di neutrini

Fino ad oggi neutrini di origine extraterrestre sono stati osservati solo dal Sole e dalla supernova SN1987 [12].

- Neutrini solari

Il sole produce neutrini a bassa energia dalle reazioni di fusione, in particolare nella catena protone-protone:



dove i neutrini hanno energia massima pari a 0.42 MeV.

Esistono altri rami di produzione dell'elio (ciclo C-N-O, $\text{Be}^7 \rightarrow \text{Li}^7$, etc.), meno frequenti, che producono neutrini ad energie leggermente più alte, ma comunque minori di 15 MeV.

Quindi i neutrini prodotti dal sole non causano problemi di background per le energie di interesse dell'astronomia di neutrini ($E < 100$ GeV). Sono stati invece molto importanti per le misure delle oscillazioni di neutrino.

Dal sole provengono tuttavia anche neutrini ad alta energia dovuti all'interazione di RC di alta energia sull'atmosfera solare: l'atmosfera solare è più densa di quella terrestre, quindi lo spettro di questi neutrini è più ripido.

- Neutrini da Supernovae

Le supernovae sono sorgenti dirette di neutrini ad energie relativamente basse, qualche decina di MeV. Dalla SN1987A sono stati osservati neutrini [12].

- Neutrini di alta energia: il “beam dump” astrofisico

I neutrini di origine extraterrestre possono essere prodotti tramite interazioni di protoni (accelerati tramite meccanismi di Fermi) di alta energia su bersagli del mezzo interstellare o del gas che circonda la sorgente. Questo meccanismo viene chiamato Astrophysical beam dump (schema in Figura 9).

FASCI di NEUTRINI: LABORATORI e CIELO

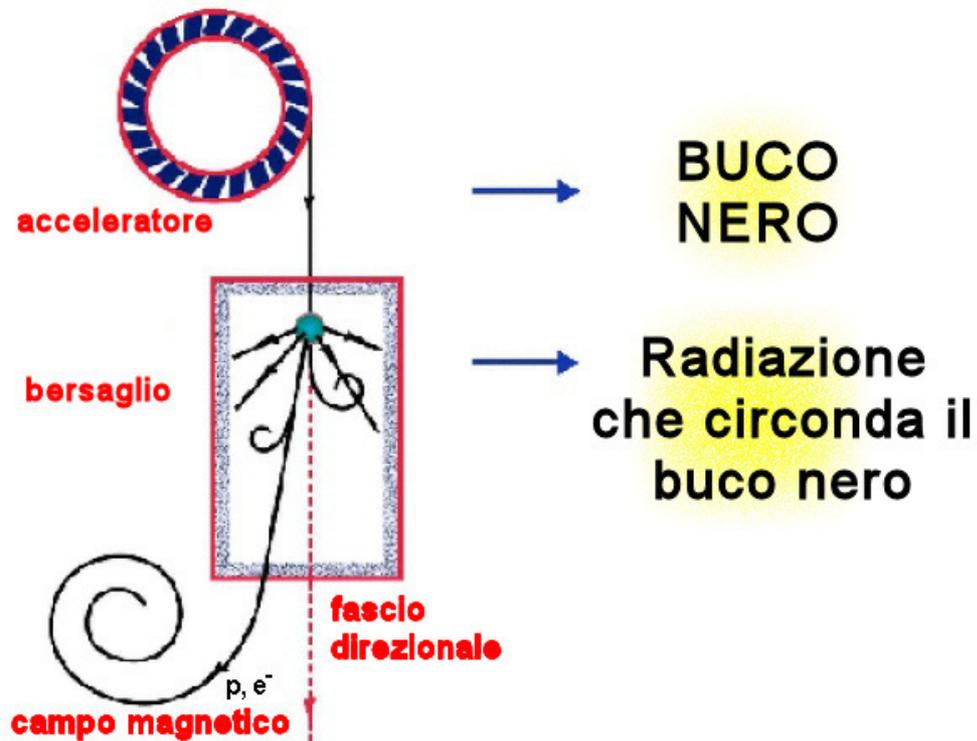
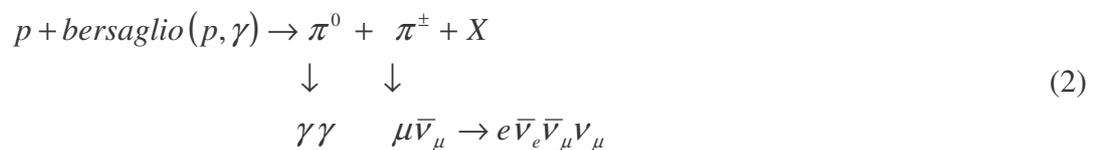


Figura 9 – Meccanismo di produzione dei neutrini: analogia tra acceleratori di laboratorio con target per la produzione di fasci di neutrini ed acceleratori cosmici (buchi neri) [15].

Nelle vicinanze dei grandi acceleratori cosmici i protoni possono interagire con altri protoni o con fotoni ambiente o che sono emessi per radiazione di sincrotrone da elettroni accelerati insieme ai protoni. Nell'urto vengono generati mesoni π e K carichi e neutri, e dal loro decadimento sono emessi neutrini e fotoni di alta energia, secondo la relazione:



Possibili sorgenti di neutrini extraterrestri, possono essere suddivise in galattiche:

- neutrini dal piano galattico
- neutrini da resti di supernova
- neutrini da stelle binarie

ed extragalattiche:

neutrini da nuclei galattici attivi

neutrini da gamma ray burst

neutrini cosmologici (da WIMPs e difetti topologici)

- Neutrini dal piano galattico

La polvere interstellare funge da bersaglio per i RC come l'atmosfera quindi è atteso un flusso di neutrini dalle regioni dello spazio più ricche di materia: un esempio vicino a noi è dato dal piano galattico. Ovviamente qui la densità di materia è bassissima e quindi lo spettro atteso dei neutrini segue quello dei RC incidenti perché i mesoni non vengono assorbiti.

- Neutrini da stelle binarie

Nei sistemi binari in cui uno dei due astri è un buco nero o una stella di neutroni e l'altro è una stella gigante, si forma un disco di accrescimento intorno all'oggetto più compatto.

I RC prodotti e accelerati (con il meccanismo descritto in 1.5) possono interagire con la materia del disco di accrescimento e produrre così un flusso di neutrini sempre secondo la relazione (2).

- Neutrini da resti di Supernovae

I resti di supernova sono sorgenti di raggi γ di alta energia (RX1713, Vela Jr. [16], [17]) e possono essere anche delle sorgenti di neutrini di alta energia. Infatti, come abbiamo visto nel §1.5 le particelle vengono accelerate tramite il meccanismo di Fermi del primo ordine attraversando l'onda d'urto del resto di supernova. Nelle SNR i protoni possono essere accelerati fino a $E \sim 10^{16} \div 10^{17}$ eV.

- Neutrini da nuclei galattici attivi

I nuclei galattici attivi (AGN) sono probabilmente gli oggetti più luminosi dell'universo osservabile, emettendo da 10^{42} a 10^{48} erg/s [14]. Gli AGN sono stati osservati per la prima volta nel 1949 e da allora ne sono stati classificati diversi tipi, ad esempio, in base allo spettro di emissione o alla luminosità. La maggior parte di questi oggetti sono situati ai confini dell'universo osservabile ($z \cong 2$). Ora sembra esserci accordo su un modello unificato in cui un AGN è una galassia molto giovane, con al centro un buco nero molto massiccio (da 10^4 a 10^{10} M_{\odot}) attorno al quale si è formato un grande disco di accrescimento. La materia (decine di M_{\odot} all'anno) in caduta sul buco nero si riscalda ed emette raggi X. Una parte della materia in caduta, a causa del forte campo magnetico, viene deviata verso i poli del buco nero, e respinta in due jets simmetrici perpendicolari al piano contenente il disco di accrescimento.

La produzione di neutrini negli AGN è spiegata con due modelli: uno basato sulla formazione di onde di shock nella regione di accrescimento, che danno luogo all'accelerazione dei protoni con il meccanismo di Fermi del primo ordine e susseguente “beam dump”; l'altro modello è basato sui jets relativistici, in cui i protoni accelerati interagiscono nel jet con i fotoni che circondano il buco nero o che sono emessi all'interno del jet da elettroni (interazione $p \gamma$), generando mesoni carichi che danno origine poi ai neutrini di alta energia [18].

- Neutrini dalle sorgenti di lampi di raggi gamma

Nel 1969 i satelliti americani lanciarono un allarme a seguito della rivelazione di lampi di raggi gamma (GRB) nello spazio che si pensava fossero dovuti ad esperimenti nucleari russi. Ad allarme rientrato, si tentò di inquadrare il fenomeno in un modello, ma non era nemmeno chiaro se le sorgenti fossero galattiche o extragalattiche. Solo nel 1997 con il satellite BeppoSAX si riuscì a determinare una distanza approssimativa dei GRB che corroborò la loro origine extragalattica e che mise in luce il fatto che la sorgente doveva essere molto luminosa. Le ipotesi attuali più accreditate considerano i GRB come eventi estremamente violenti (battezzati ipernove) come la caduta di una stella di neutroni su un buco nero o la collisione tra due buchi neri.

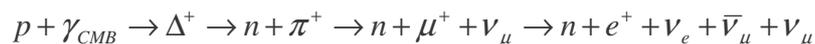
L'accelerazione delle particelle e l'interazione con la materia della regione di accrescimento danno origine a neutrini di alta energia.

- Neutrini da sorgenti esotiche

Altre possibili sorgenti di neutrini di altissima energia possono essere i decadimenti di particelle massive debolmente interagenti (WIMPs) che si depositerebbero nel centro delle stelle e dei pianeti o difetti topologici.

- Effetto GZK

Una sorgente sicura di neutrini ad altissima energia è data dall'effetto GZK (§1.4). I neutrini di $E > 10^{19}$ eV sono il prodotto secondario del decadimento della risonanza Δ^+ creata nell'interazione di protoni con la radiazione di fondo cosmica:



Nelle figure seguenti sono riportati i flussi attesi di neutrini dalle sorgenti menzionate: in Figura 10 sono riportati i flussi provenienti da sorgenti estese, mentre in Figura 11 i flussi da sorgenti puntiformi.

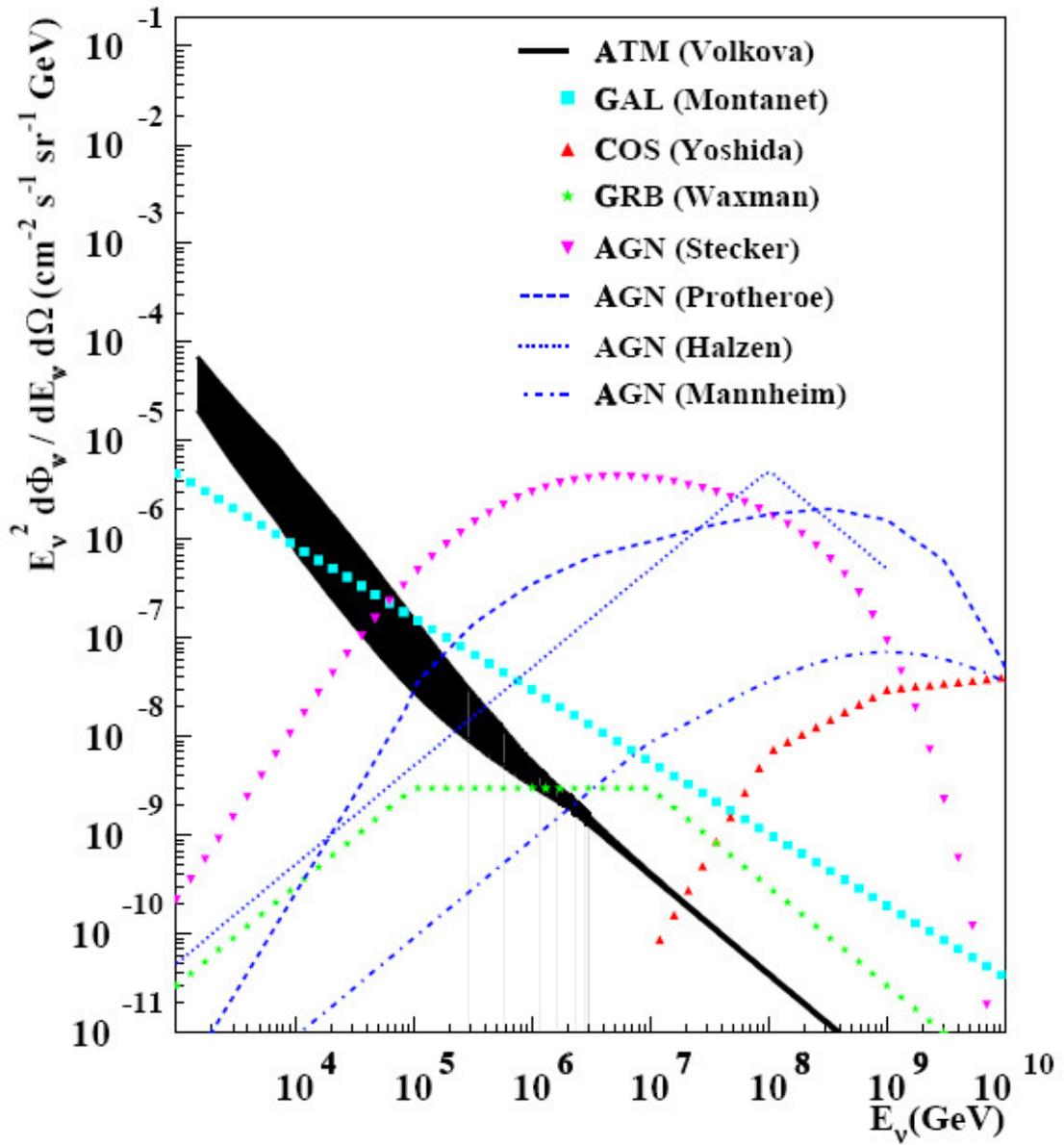


Figura 10 – Flussi attesi di neutrini da sorgenti estese in funzione dell'energia dei neutrini emessi: (GAL) neutrini dal piano galattico, (COS) neutrini cosmologici da effetto GZK, (GRB) neutrini provenienti da Gamma Ray Bursts, (AGN) neutrini provenienti da Nuclei Galattici Attivi, secondo vari modelli [9]. (ATM) indica il flusso diffuso di neutrini atmosferici calcolato in base al modello di Volkova et al. [19].

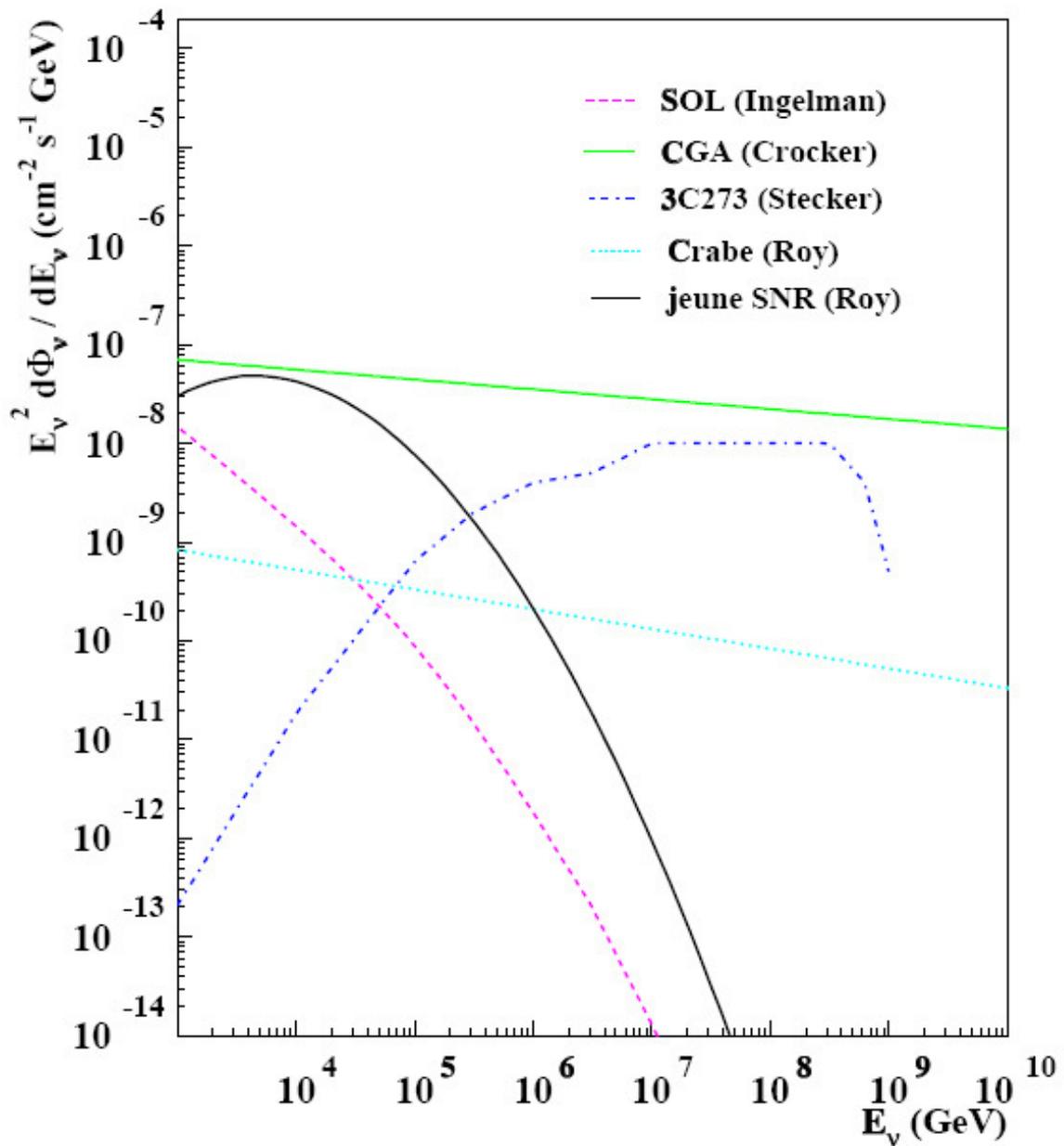


Figura 11 - Flussi attesi di neutrini da sorgenti puntiformi: (CGA) neutrini provenienti dal centro della nostra galassia, (3C273) dalla galassia attiva 3C273, (Crabe) dalla nebulosa del Granchio, (jeune SNR) dai resti di una giovane supernova. (SOL) in linea rossa tratteggiata indica il flusso differenziale di neutrini misurato dal Sole [9].

Cap. 2 I telescopi a neutrini

2.1 Schema di principio

I neutrini interagiscono con la materia solo attraverso le interazioni deboli, per questo la loro lunghezza di attenuazione è enorme, e per questo i rivelatori di neutrini devono avere volumi molto grandi o densità di materia elevate. Le interazioni dei neutrini con la materia possono essere di corrente carica (CC) (attraverso lo scambio di un bosone W^\pm) o di corrente neutra (CN) (attraverso lo Z):

$$\begin{aligned} \nu_\ell (\bar{\nu}_\ell) + N &\rightarrow \ell (\bar{\ell}) + X && (CC) \\ \nu_\ell (\bar{\nu}_\ell) + N &\rightarrow \nu_\ell (\bar{\nu}_\ell) + X && (CN) \end{aligned} \quad (3)$$

dove ℓ indica il sapore del neutrino interagente ($\ell = e, \mu, \tau$) e il leptone corrispondente generato nella interazione di CC.

Un rivelatore di neutrini di origine astrofisica, che chiameremo d'ora in avanti telescopio per neutrini, sfrutta principalmente l'interazione di CC del neutrino muonico. Il muone prodotto nell'interazione $\nu_\mu, N(\bar{\nu}_\mu, N)$, infatti può viaggiare per lunghe distanze (fino a qualche km) e provocare l'emissione di radiazione elettromagnetica per effetto Čerenkov nel mezzo attraversato. Questo permette di:

- aumentare molto il volume efficace (cfr §2.6) del rivelatore
- consente di ricostruire accuratamente la traccia del muone e la sua direzione di arrivo

Dato che l'angolo tra il neutrino e il muone prodotto è molto piccolo e diminuisce all'aumentare di E_ν ($\theta \propto 1/\sqrt{E}$) ciò permette di risalire anche alla direzione di arrivo del neutrino, cioè di fare neutrino-astronomia.

Un ottimo mezzo per avere effetto Čerenkov con volumi sufficienti allo scopo e a buon mercato è l'acqua di mare. Un telescopio a neutrini sarà allora costituito da un array di rivelatori di fotoni (fotomoltiplicatori) posizionati sul fondo marino a diversi km dalla superficie che rivelano i fotoni emessi per effetto Čerenkov al passaggio del muone. In fondo al mare, oltre ad avere un ambiente oscuro necessario per la rivelazione dei pochi fotoni Čerenkov, si ha anche una certa attenuazione delle particelle secondarie (principalmente muoni) generate negli sciame atmosferici. Questa parziale attenuazione non sarebbe comunque affatto sufficiente ad osservare tracce di muoni dovute all'interazione di un neutrino di origine extraterrestre proveniente dall'alto: infatti, il numero di muoni atmosferici che filtrano fino alle profondità accessibili per questi esperimenti è comunque alcuni ordini di grandezza superiore a quello dei muoni dovuti a neutrini astrofisici. L'unico modo per avere un "segnale pulito" è quello di cercare le tracce di muoni provenienti dal basso (Figura 12). Questi sono necessariamente dovuti a neutrini, perché nessun muone può attraversare l'intero pianeta, mentre per i neutrini, fino ad energie di ~ 100 TeV la terra è quasi completamente trasparente⁴.

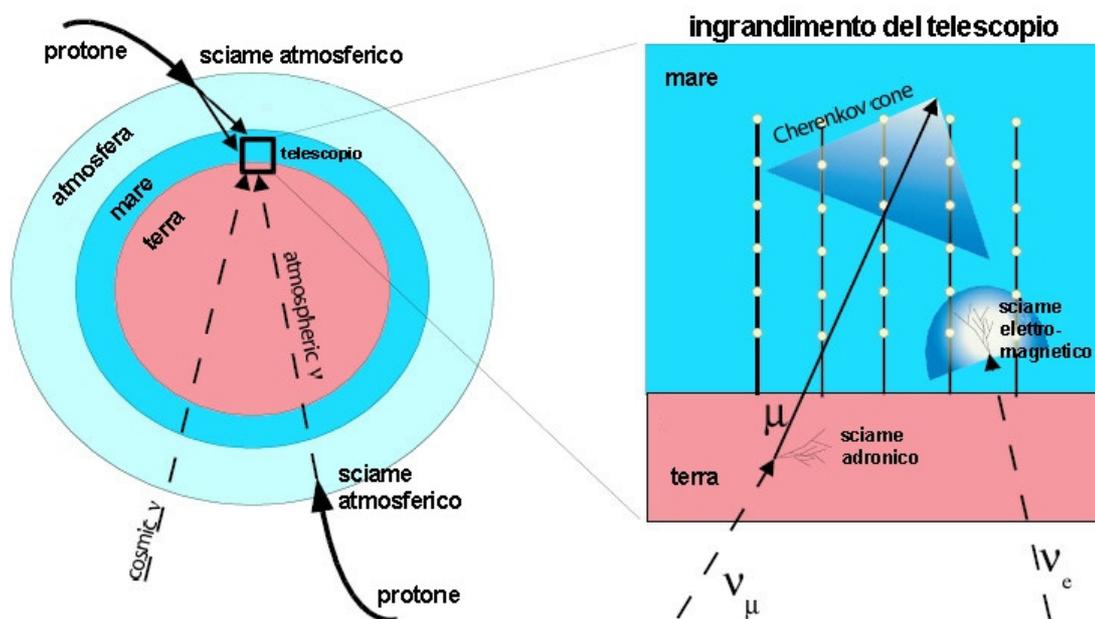


Figura 12 - Schema di principio di un telescopio a neutrini [20].

⁴ A energie più elevate, poiché la cross section νN cresce con l'energia, la terra diventa opaca, e il telescopio a neutrini può rivelare solo quelli provenienti dall'alto o da direzioni prossime all'orizzonte.

Nella Figura 12 a sinistra si vedono due protoni incidenti sull'atmosfera terrestre che danno inizio a sciame atmosferici estesi, uno proveniente dall'alto che può dare luogo ad un muone che sarà filtrato una volta ricostruita la traiettoria, uno proveniente dal basso che genera un neutrino atmosferico che attraversa la Terra e giunge nei pressi del rivelatore. I muoni prodotti dai neutrini atmosferici provenienti dal basso sono indistinguibili dai muoni indotti dai neutrini astrofisici e quindi costituiscono un fondo irriducibile per il rivelatore. Solo su base statistica, con anni di osservazione e di raccolta dati, si potranno identificare le eventuali sorgenti puntiformi di neutrini cosmici. I neutrini atmosferici possono essere distinti da quelli astrofisici anche considerando la loro energia, perché il loro spettro decade molto più rapidamente ($\sim E^{-2.7}$) rispetto a quello aspettato per i neutrini astrofisici ($\sim E^{-2}$) come si osserva in Figura 10. A destra di Figura 12 è rappresentato un ingrandimento del telescopio in cui è evidenziata la traccia di un μ , con il cono di luce Čerenkov, e un neutrino elettronico. L'elettrone prodotto da questo neutrino ha una traccia molto breve, causando uno sciame elettromagnetico completamente contenuto nell'apparato, da cui si può difficilmente ricostruire la traccia del ν_e incidente. Per questo, per fare astronomia con i neutrini, ci si limita a considerare le tracce di muoni.

2.2 Interazione dei neutrini con la materia

La cross section $\sigma_{\nu N}(E_\nu)$ dell'urto profondamente anelastico di un neutrino muonico su un nucleone in una interazione CC, è descritta dalla relazione [21]:

$$\frac{d^2\sigma}{dx dy} = \frac{2G_F^2 m_N E_\nu}{\pi} \left(\frac{M_W^2}{Q^2 + M_W^2} \right)^2 [xq(x, Q^2) + x\bar{q}(x, Q^2)(1 - y^2)]$$

dove E_ν è l'energia del neutrino interagente, $x = Q^2 / [2m_N(E_\nu - E_\mu)]$ e $y = (E_\nu - E_\mu) / E_\nu$ sono le variabili di scala di Bjorken, m_N è la massa del nucleone, M_W è la massa del bosone W, G_F è la costante di Fermi, Q^2 è l'impulso scambiato e $q(x, Q^2)$, $\bar{q}(x, Q^2)$ sono le funzioni di distribuzione dei quark. In Figura 13 è riportato l'andamento della sezione d'urto di neutrini e antineutrini muonici al variare dell'energia del neutrino interagente.

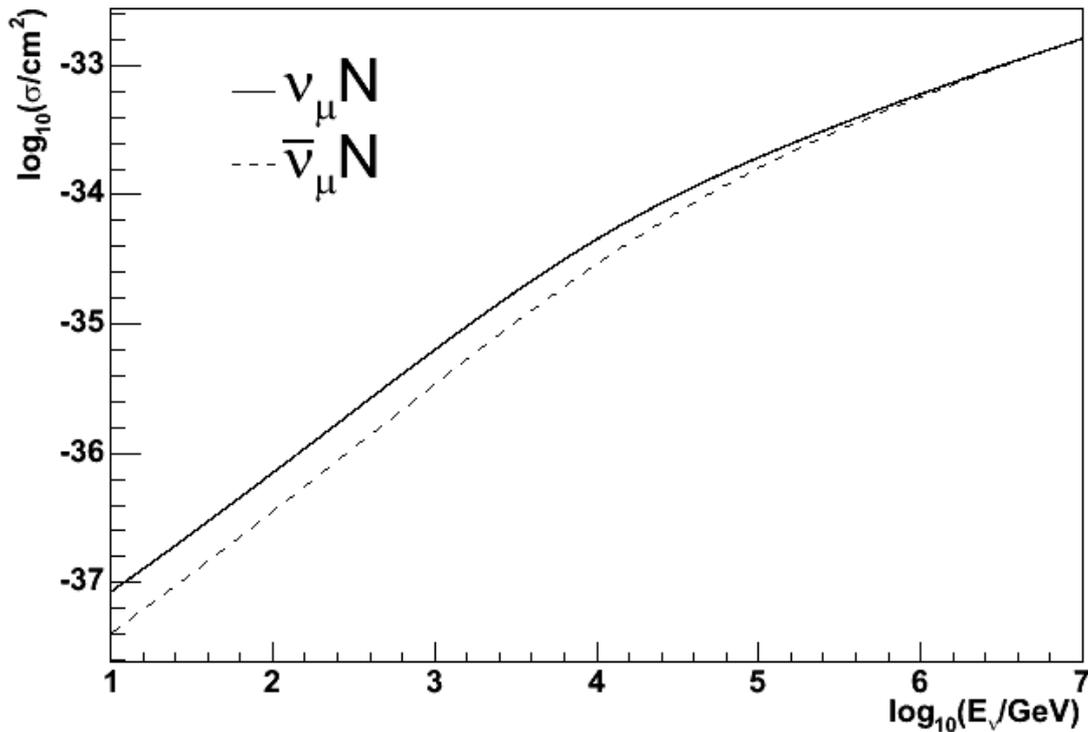


Figura 13 – Sezione d’urto di neutrini e antineutrini muonici su nucleoni in funzione dell’energia del neutrino incidente.

La cross section dell’interazione dei neutrini, estremamente bassa a basse energie (10^{-37} cm^2 a 10 GeV) cresce all’aumentare dell’energia (10^{-33} cm^2 a 10 PeV). La lunghezza di interazione ν - N , cioè la distanza oltre la quale il flusso si riduce di un fattore e , diminuisce di conseguenza:

$$L_{int}(E_\nu) = \frac{1}{\sigma_{\nu N}(E_\nu)} \frac{\rho A}{N_A}$$

dove N_A è il numero di Avogadro, ρ è la densità della materia attraversata ed A è il numero di massa, e la lunghezza di attenuazione è espressa in g/cm^2 .

La densità della terra varia con la profondità, ed è riportata in Figura 14, secondo un modello ben consolidato; un neutrino proveniente dal basso, con una direzione di arrivo formante un angolo θ rispetto al nadir del punto di osservazione, incontra una colonna di materia con densità equivalente $z(\theta)$, che dà luogo al coefficiente di assorbimento che è funzione della direzione di provenienza θ e di E_ν :

$$S(E_\nu, \theta) = \exp\left[\frac{-z(\theta)}{\mathcal{L}_{\text{int}}(E_\nu)}\right]$$

In Figura 14 è riportato anche il coefficiente di assorbimento in funzione di θ ed E_ν .

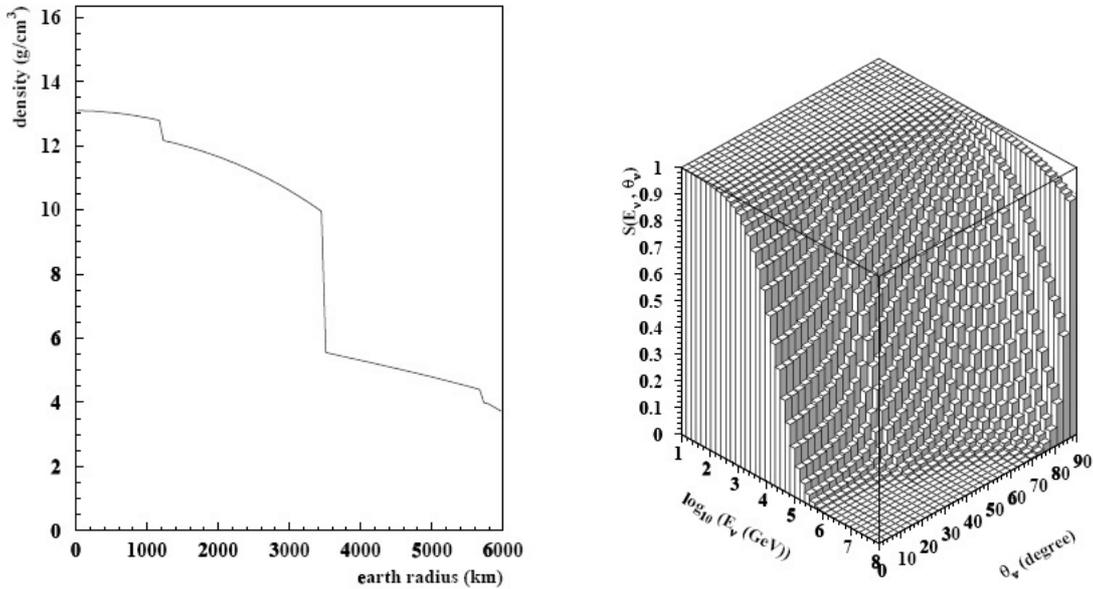


Figura 14 – Densità della terra in funzione della profondità (a sinistra) e coefficiente di trasmissione in funzione di θ ed E_ν (a destra) [9].

I neutrini orizzontali ($\theta = 90^\circ$) che attraversano una piccola quantità di materia, non subiscono apprezzabile assorbimento neanche ad altissime energie, mentre quelli provenienti dal basso ($\theta = 0^\circ$) a 100 TeV sono quasi del tutto assorbiti.

I neutrini muonici che interagiscono per CC (3), generano muoni la cui direzione di propagazione alle alte energie forma un angolo abbastanza piccolo rispetto alla direzione d'arrivo dei neutrini. Questo permette quindi di risalire alla direzione della sorgente dei neutrini. Il valore quadratico medio dell'angolo tra la traccia del neutrino e quella del muone è

$$\langle \theta_{\nu\mu} \rangle \approx \frac{1^\circ}{\sqrt{E_\nu [\text{TeV}]}}$$

Nella Figura 15 è riportato il coseno dell'angolo tra muone e neutrino in funzione dell'energia, ottenuto tramite una simulazione dell'interazione $\nu_\mu - N$ effettuata con il programma Genhen [22].

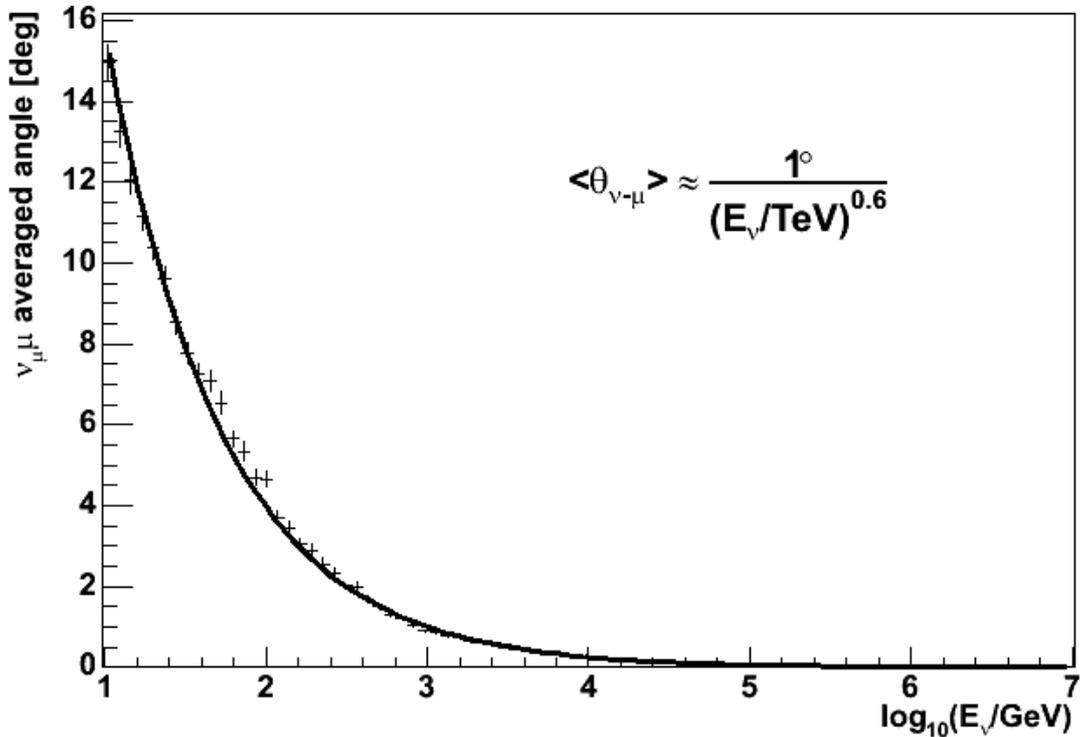


Figura 15 – Angolo medio tra muone e neutrino in funzione dell'energia.

2.3 Propagazione dei muoni

I muoni prodotti dai neutrini perdono energia durante il loro tragitto nella roccia e nell'acqua, e subiscono una diffusione coulombiana multipla.

Le perdite di energia sono dovute alla ionizzazione degli atomi della materia attraversata, all'irradiazione per brehmsstrahlung, alla produzione di coppie elettrone-positrone, e all'interazione con i nuclei; quest'ultima è la sola responsabile di apprezzabile deviazione angolare.

La perdita di energia per ionizzazione, dominante ad energie del muone $E_\mu < 1$ TeV (cfr. Figura 16), è descritta dall'equazione di Bethe-Bloch [22]:

$$\frac{dE}{dx} = \frac{4\pi z^2 e^4}{m_e v^2} \frac{N_A Z}{A} \left[\ln \left(\frac{2m_e \gamma^2 v^2 E_{MAX}}{I^2} \right) - \beta^2 \right] \quad (4)$$

dove z è la carica della particella incidente e v è la sua velocità, ($\beta = v/c$), e è la carica dell'elettrone, m_e è la sua massa, Z ed A sono il numero atomico e il numero di massa del materiale attraversato, e N_A è il numero di Avogadro. Ci sono delle correzioni a questa

formula, in particolare ad alte energie bisogna tenere conto dell'effetto densità, dovuto al fenomeno di schermo ad opera del campo elettrico della stessa particella incidente, che diminuisce il valore stimato dall'eq. (4). Comunque le perdite per ionizzazione dipendono in modo logaritmico dall'energia: $\alpha_l \approx \log(E)$.

Le perdite di energia per bremsstrahlung, produzione di coppie e interazioni nucleari, che diventano dominanti per $E_\mu > 1$ TeV, sono invece lineari in E_μ :

$$\frac{dE_\mu}{dx} = -\frac{E_\mu}{X_0}$$

dove X_0 è la lunghezza di attenuazione, ma si tratta di fenomeni stocastici che vanno modellizzati tramite la simulazione Monte Carlo.

In totale si ha:

$$-\left\langle \frac{dE(E_\mu)}{dx} \right\rangle = \alpha + \beta E_\mu$$

In Figura 16 è riportato l'andamento delle perdite di energia in funzione dell'energia del muone, simulate nell'ambito dell'esperimento ANTARES [24], utilizzando il pacchetto software MUSIC [25].

In Figura 17 sono riportate le distribuzioni di probabilità del range del muone $dP/d(\log_{10}(R_\mu))$, in acqua e roccia, espresso in km equivalenti di acqua (kmwe)⁵ per diverse energie.

In Figura 18 è invece riportata la deviazione angolare tra la traccia del muone e la direzione del ν incidente, dovuta alla diffusione colombiana multipla e all'interazione neutrino-muone; si vede che la deviazione angolare rispetto alla direzione iniziale del muone dovuta allo scattering sui nuclei è molto minore dell'angolo tra muone e neutrino.

⁵ Un km equivalente di acqua è una misura di lunghezza di attenuazione, espressa per densità di materiale unitaria; in un mezzo con densità ρ corrisponde ad una lunghezza di attenuazione pari a $1 \text{ km}/\rho$

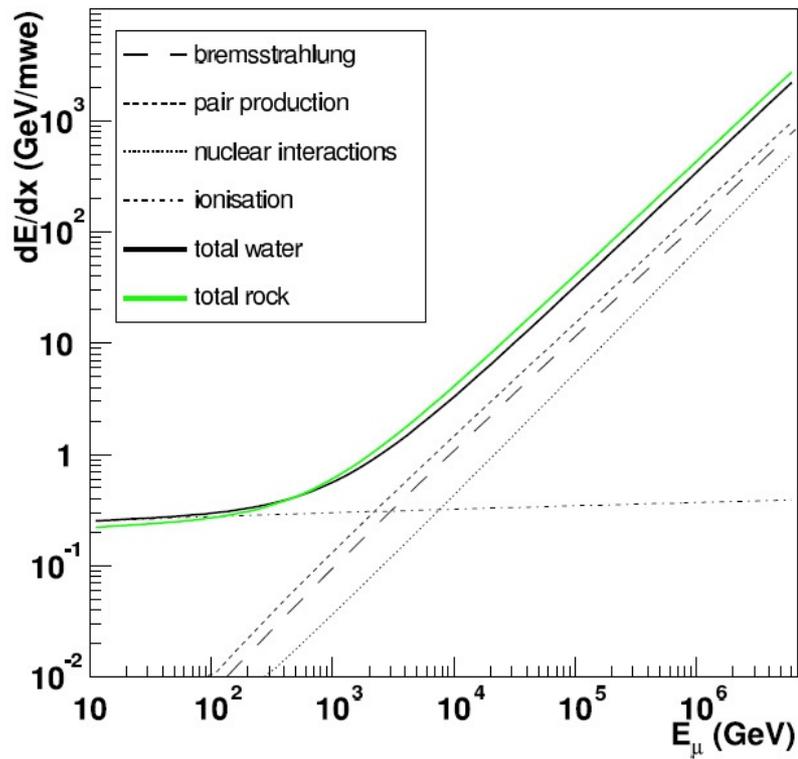


Figura 16 – Perdite di energia nella propagazione del muone in acqua e roccia, in funzione dell'energia del muone.

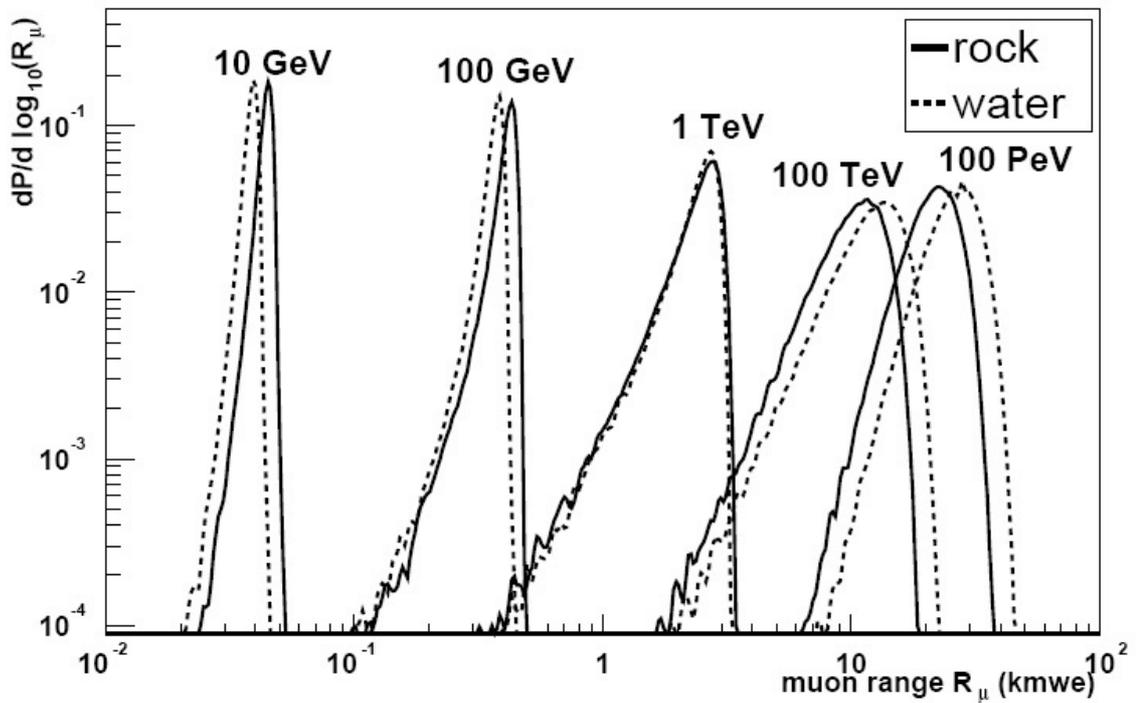


Figura 17 – Range del muone in acqua e roccia per diverse energie [24].

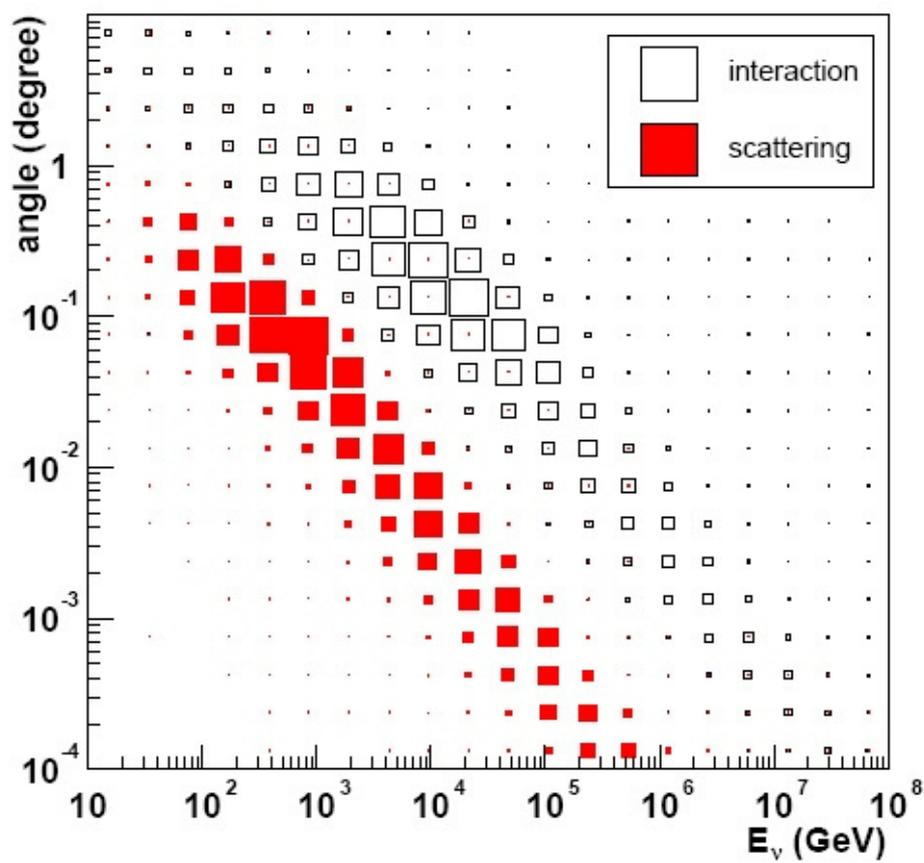


Figura 18 – Deviazione angolare tra la traccia del muone e la direzione del neutrino incidente, dovuta allo scattering multiplo e all'interazione da correnti cariche [24].

Il fatto che il range dei muoni di alta energia in acqua e roccia sia dell'ordine di qualche km di fatto aumenta la capacità del telescopio di rivelare i neutrini, cioè il Volume Efficace. Infatti non è necessario che l'interazione di CC (3) avvenga nel volume fisico del detector, ma anche qualche km prima (cfr.§2.5).

2.4 Effetto Čerenkov

I muoni di alta energia che attraversano l'acqua di mare hanno una velocità superiore a quella della luce nel mezzo. Infatti l'acqua di mare ha un indice di rifrazione n tipicamente pari a $1.350 \div 1.355$, e la velocità della luce si riduce a $v_g = c/n$.

Le particelle cariche che attraversano un mezzo lo polarizzano, e le molecole del mezzo, tornando in condizione di riposo emettono fotoni isotropicamente (cfr. Figura 19). Quando però la velocità della particella è maggiore della velocità della luce nel mezzo c'è una

direzione in cui i fotoni emessi si sommano in fase (principio di Huygens) e producono un'onda di shock con fronte d'onda che forma un angolo ben definito, caratteristico del mezzo, con la traiettoria della particella (cfr. Figura 20) [3].

Dalla Figura 20 si ricava che l'angolo caratteristico dell'emissione Čerenkov è tale che

$$\cos \theta_c = \frac{1}{n\beta}.$$

Il numero di fotoni emessi per unità di percorso e per unità di frequenza è:

$$\frac{d^2 N}{dx d\omega} = \frac{z^2 \alpha}{c} \sin^2 \theta_c$$

che dà un numero pari a circa 300 fotoni al cm nella banda del visibile (tra 300 e 600 nm), dove l'acqua è più trasparente e i fotomoltiplicatori presentano la massima sensibilità.

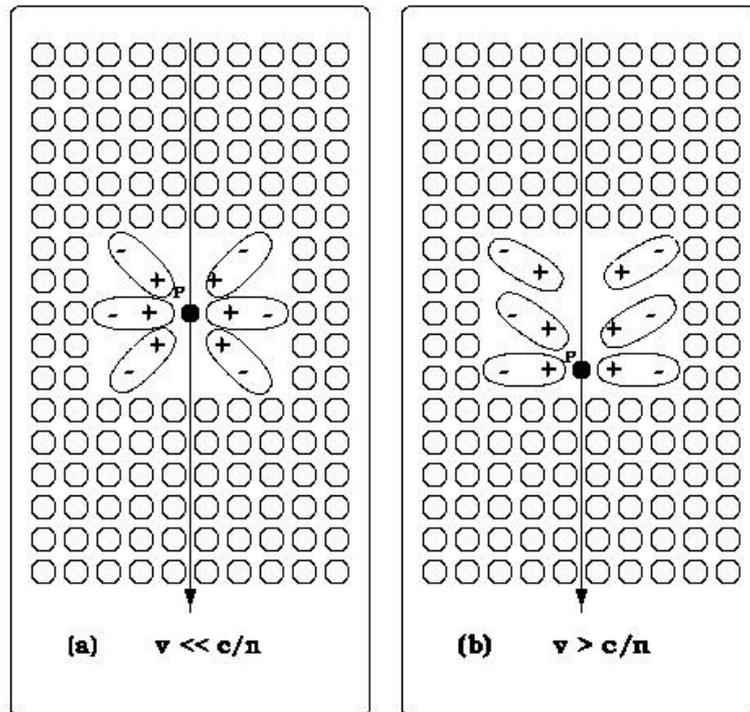


Figura 19 - Polarizzazione delle molecole del mezzo attraversato, in condizioni di velocità molto minori (a), e maggiori (b) della velocità della luce nel mezzo.

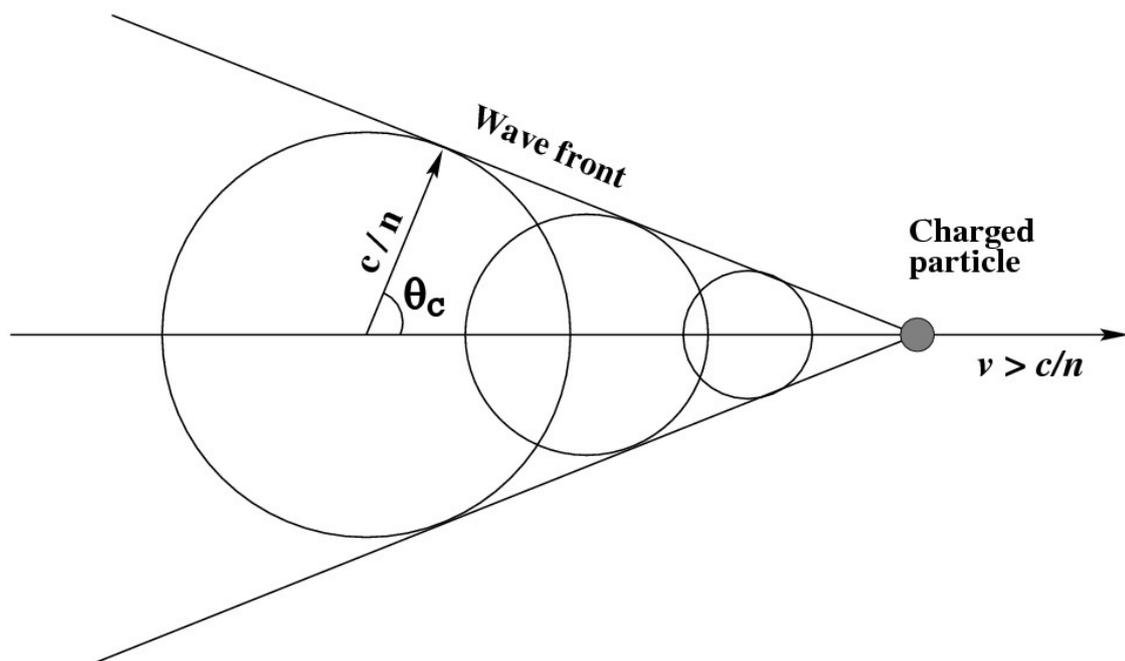


Figura 20 - Somma in fase dei fotoni emessi dalle molecole del mezzo (in direzione θ_c) attraversato da una particella carica con velocità ultra-relativistica.

Anche i fotoni emessi per effetto Čerenkov, durante la propagazione in acqua subiscono fenomeni di assorbimento e diffusione. Entrambi i processi sono modellizzabili tramite un andamento esponenziale, con L_a parametro di lunghezza di attenuazione:

$$I(x) = I_0 e^{-\frac{x}{L_a}}$$

Il fenomeno dello scattering introduce un ritardo aleatorio nel tempo di transito di un fotone dal punto di emissione al rivelatore [26].

2.5 Fotomoltiplicatori

I fotoni emessi per effetto Čerenkov al passaggio del muone devono essere rivelati e, per ricostruire la direzione del muone, deve essere misurato il loro tempo di arrivo con un'accuratezza di pochi ns. Un detector in grado di rivelare singoli fotoni con un'accuratezza temporale dell'ordine dei ns è il fotomoltiplicatore (PMT). Questo tipo di rivelatore è costituito da un tubo a vuoto in cui ad un'estremità c'è un fotocatodo, cioè una lamina di materiale metallico che colpito da un fotone libera un elettrone per effetto fotoelettrico con una certa probabilità, definita dall'efficienza quantica del PMT. L'elettrone emesso

(fotoelettrone) viene accelerato da un forte campo elettrico di polarizzazione e viene ad urtare una serie di elettrodi interni (dinodi), ognuno ad un potenziale maggiore, liberando ad ogni urto altri elettroni, e producendo quindi una moltiplicazione a valanga. Alla fine, sull'elettrodo di raccolta si ha un impulso di corrente di intensità sufficiente ad essere misurata: infatti, se un PMT ha 14 dinodi, ed ognuno quando urtato da un elettrone ne riemette (con legge poissoniana) in media 3, si hanno all'uscita $3^{14} \cong 5 \cdot 10^6$ elettroni, cioè una carica pari a circa $8 \cdot 10^{-13}$ C, che in un tempo pari alla durata media dell'impulso, pari a ~ 10 ns, dà una corrente di circa $80 \mu\text{A}$, e quindi un segnale di 4 mV su una resistenza di 50Ω .

In Figura 21 è rappresentato un PMT, con uno schermo magnetico di mu-metal; infatti il percorso degli elettroni nel tubo è disturbato dal campo magnetico terrestre, e specialmente per i primi dinodi, in cui gli elettroni hanno una bassa energia, potrebbe causare la perdita degli elettroni.



Figura 21 - PMT con schermo magnetico.

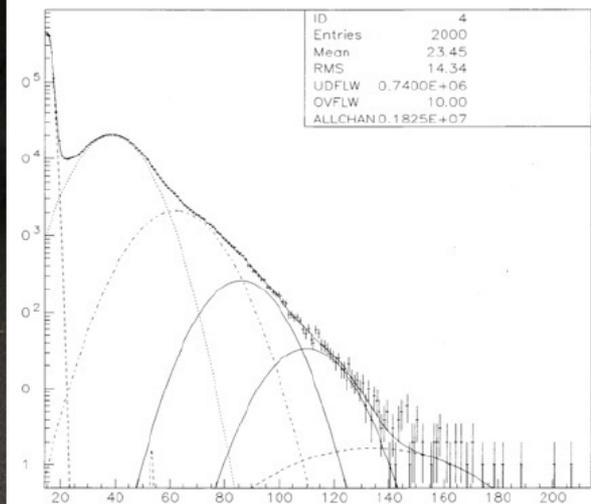


Figura 22 - Spettro in carica di un PMT.

In Figura 22 è invece graficato lo spettro in carica di un PMT; si noti che anche in assenza di fotoelettroni emessi dal fotocatodo si può avere un impulso in uscita, dovuto all'emissione spontanea di elettroni dai dinodi, (piedistallo). La distribuzione del numero di fotoelettroni (e quindi della carica raccolta sull'anodo del PMT) è poissoniana con media 0.25.

2.6 Efficienza di rivelazione

Si è detto che il fatto che i muoni abbiano un range molto lungo (qualche km) aumenta decisamente la capacità di rivelazione del telescopio. La capacità del rivelatore di identificare una traccia di un muone viene indicata con i parametri Area Efficace e Volume Efficace, intendendo con questi l'area attorno ad una sezione dell'apparato all'interno della quale i muoni che passano vengono rivelati, e il volume ottenuto proiettando l'Area Efficace per il range dei muoni (cfr §2.3).

Si ha pertanto:

$$V_{eff} = A_{eff} R_{\mu}$$

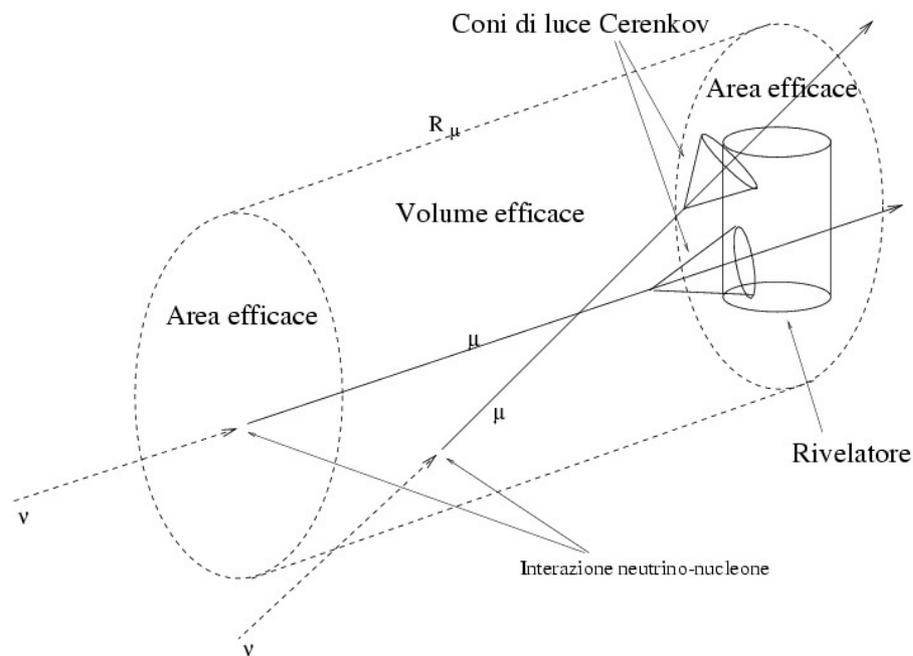


Figura 23 - Area efficace e volume efficace di un rivelatore per neutrini [31].

L'area efficace può essere calcolata conoscendo le efficienze di rivelazione, oppure stimata tramite simulazioni Monte Carlo. Scelta un'area di riferimento su cui calcolare il flusso di muoni che la attraversa, vengono conteggiate le tracce meglio ricostruite. La frazione di eventi ricostruiti moltiplicata per l'area geometrica di generazione dà l'area efficace.

L'area geometrica di riferimento deve essere più grande dell'area che contiene il rivelatore, ad esempio con un margine di 5 lunghezze di attenuazione dei fotoni.

L'area efficace dipende, oltre che dall'energia del muone, anche dalla geometria del rivelatore, per esempio dalla distanza tra i PMT rispetto alla lunghezza di attenuazione dei fotoni in acqua e dagli algoritmi di filtraggio e ricostruzione utilizzati nell'analisi degli eventi.

2.7 Telescopi a neutrini nel Mediterraneo

2.7.1 ANTARES

La collaborazione ANTARES (Astronomy with a Neutrino Telescope and Abyss environmental RESearch) [32] ha lo scopo di costruire un telescopio Čerenkov sottomarino nel mar Mediterraneo, al largo della costa Francese nei pressi di Tolone (cfr. Figura 24).



Figura 24 – Posizione geografica del telescopio ANTARES.

Il telescopio avrà un'area efficace di circa 0.1 km^2 , e lo scopo è quello di iniziare ad osservare i neutrini di origine cosmica per preparare accuratamente lo sviluppo di un telescopio da 1 km^3 (NEMO).

Il sito si trova ad una profondità di 2400 m dalla superficie e presenta una lunghezza di assorbimento dei fotoni in acqua pari a 55 m alla lunghezza d'onda del blu.

L'apparato è formato da 12 stringhe flessibili contenenti 25 piani separati di 15 m l'uno dall'altro. Ogni piano ha tre PMT separati di 120° l'uno dall'altro e orientati a 45° verso il basso, in modo da ottimizzare la rivelazione dei muoni provenienti dal basso e minimizzare l'effetto del biofouling [12]. In Figura 25 è rappresentata la pianta di posizionamento delle stringhe e in Figura 26 è riportato lo schema di tutto il telescopio.

I segnali elettrici in uscita dai PMT sono campionati, digitalizzati e trasmessi via fibra ottica alla Junction Box, connessa con un unico cavo elettro-ottico alla stazione di terra a Seyne sur le mer. Ci si aspetta una precisione temporale di circa 1 ns nella determinazione del tempo di arrivo dei fotoni Čerenkov sui PMT.

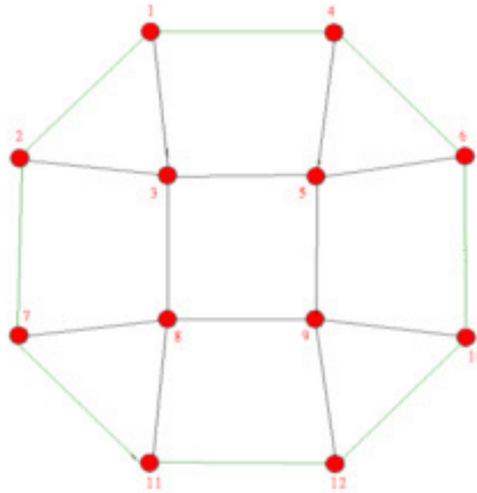


Figura 25 - Pianta di posizionamento delle stringhe di ANTARES.

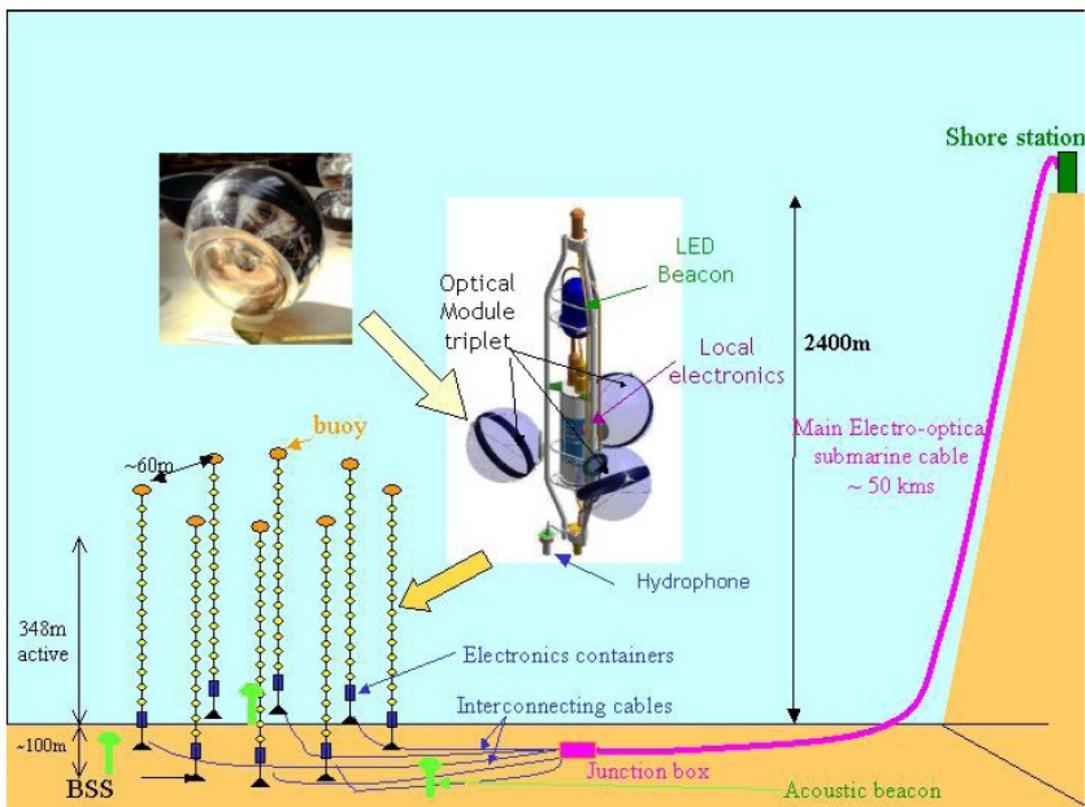


Figura 26 - Schema dell'intero telescopio ANTARES.

2.7.2 NEMO

La collaborazione NEMO (NEutrino Mediterranean Observatory) [33] ha lo scopo di sviluppare un telescopio a neutrini di scala 1 km^3 . Sono stati considerati vari siti sottomarini dove installare l'apparato (riportati in Figura 27) ed è stato scelto il sito al largo delle coste orientali della Sicilia, a sud di Capo Passero, ad una profondità di 3350 m, con ottime proprietà di trasparenza dell'acqua ($L_a = 68 \text{ m}$ alla lunghezza d'onda del blu).

L'apparato consiste di un reticolo a base quadrata di 81 torri, ciascuna formata da una sequenza di 18 piani ruotati di 90° l'uno rispetto al superiore, mantenuti in posizione attraverso due coppie di cavi che uniscono il piano inferiore e superiore, come rappresentato in Figura 28. Questa soluzione, rispetto alle stringhe di ANTARES, offre una maggiore stabilità della struttura, specialmente rispetto alla rotazione.

La distanza tra le torri è 140 m, corrispondenti a circa due lunghezze di attenuazione dei fotoni alla lunghezza d'onda del blu. La distanza tra i piani è di 40 m. Agli estremi dei piani sono posizionati i due moduli ottici, per un totale di 4 PMT per piano, contenenti i PMT e l'elettronica di conversione analogico-digitale. Due PMT sono orientati in senso orizzontale, gli altri due sono orientati verso il basso.

Il controllo del posizionamento delle torri avviene mediante un sistema acustico formato da una serie di idrofoni collocati sulle torri e da un sistema di emettitori (beacon) poggiati sul fondale marino.

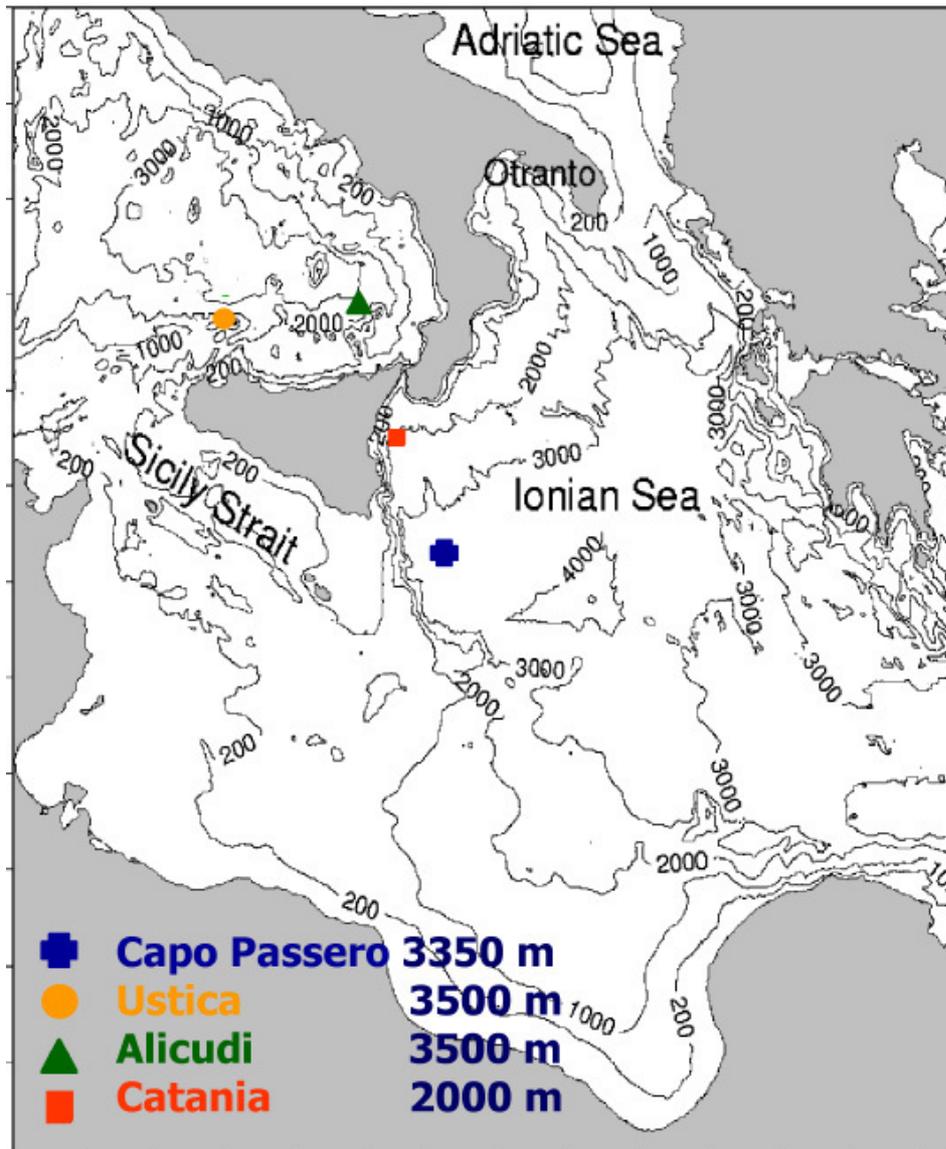


Figura 27 – Posizione di alcuni siti analizzati per la scelta dell'istallazione di NEMO.

I 5832 PMT dell'apparato danno un contributo al segnale (hit) se la carica misurata supera una soglia prefissata, e i tempi di hit consentono la ricostruzione della traccia del muone indotto da un neutrino, come spiegato dettagliatamente nel §3.2. In Figura 29 è rappresentata la simulazione di una traccia di un muone che attraversa l'apparato. I punti neri sono le posizioni di tutti i PMT, quelli gialli sono i PMT che hanno dato origine ad un hit, e quelli rossi sono hit filtrati, cioè non utilizzati per la ricostruzione.

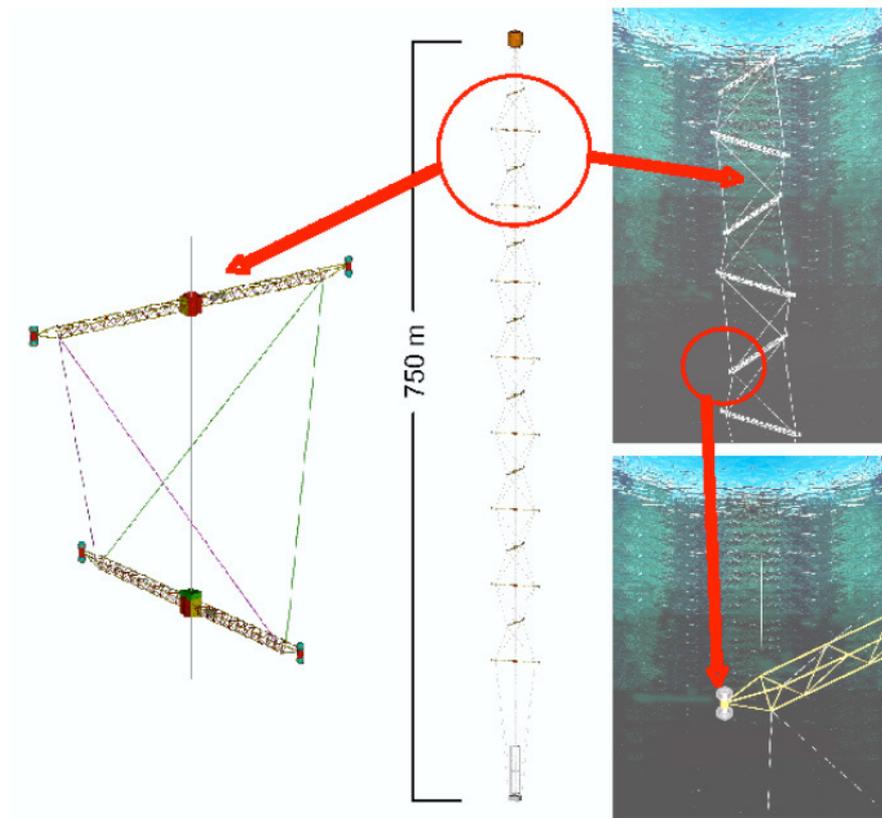


Figura 28 – Struttura delle torri di NEMO.

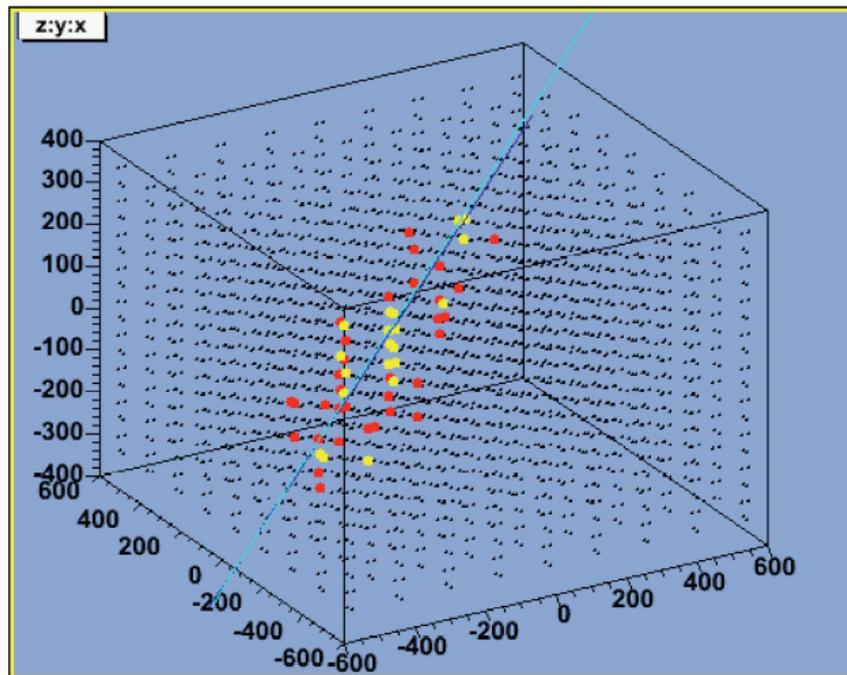


Figura 29 – Evento ricostruito: in nero PMT spenti, in rosso hit filtrati, in giallo hit per la ricostruzione.

2.8 Altri Esperimenti

NEMO ed ANTARES raccolgono l'esperienza di alcuni esperimenti pionieristici come DUMAND, presso le isole Hawaii, e gli attuali Baikal, nell'omonimo lago in Russia, e AMANDA in Antartide.

L'esperimento Baikal, ad una profondità di circa 1000 m soffre di una bassa lunghezza di assorbimento, che richiede una densità di PMT maggiore di quella degli esperimenti marini, ma ha il vantaggio di avere per questo motivo anche una soglia di energia più bassa. D'altra parte la limitata area efficace ($1000 \div 5000 \text{ m}^2$) esclude la possibilità di osservare neutrini cosmici, e l'obiettivo è stato quello di ricostruire neutrini atmosferici provenienti dal basso [35]. In Figura 30 è riportato il coseno dell'angolo zenitale degli eventi ricostruiti in 70 giorni di presa dati nel 1997.

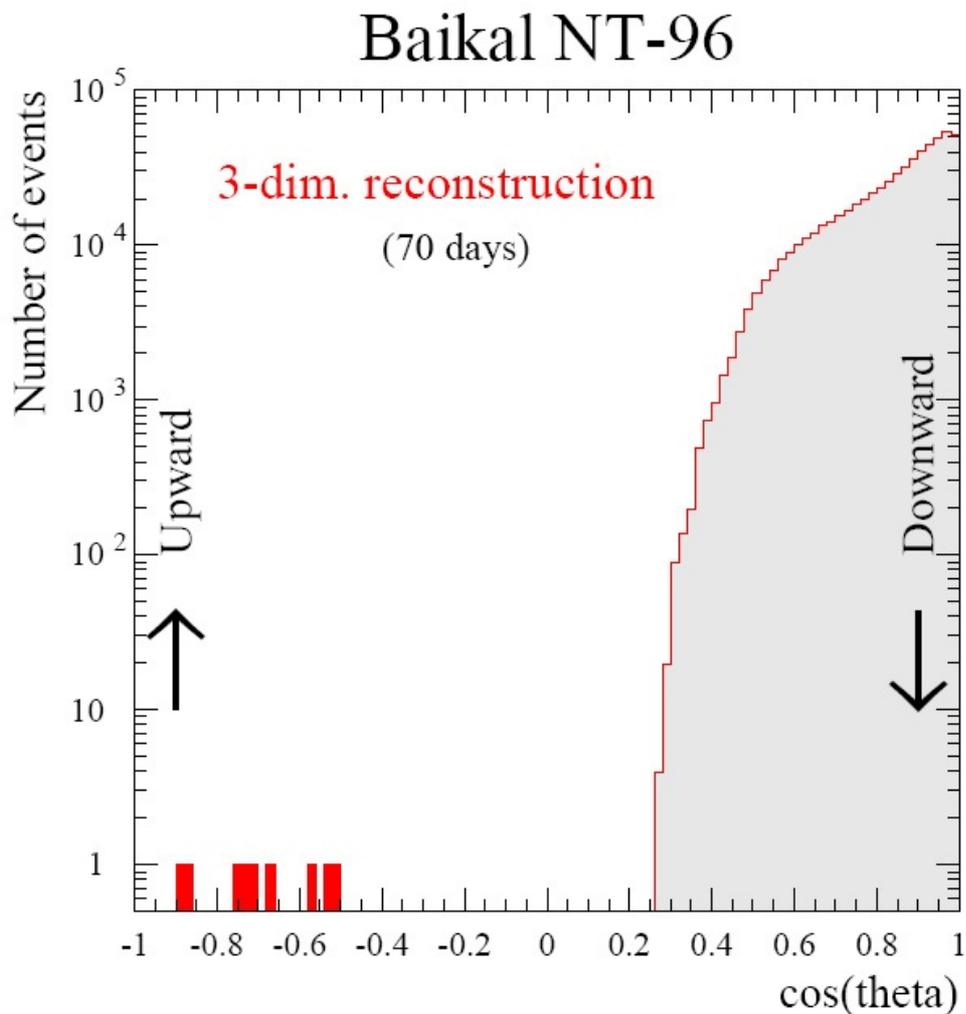


Figura 30 – Ricostruzione di eventi con l'esperimento Baikal [35].

Il telescopio AMANDA prima generazione è stato completato nell'estate australe 1996-97, ed è installato nel ghiaccio del polo sud, ad una profondità di $1500 \div 2000$ m. Consiste di un array di 302 PMT, esteso poi, con la seconda generazione a 677 PMT disposti in 19 stringhe verticali.

Il ghiaccio ha proprietà ottiche complementari a quelle dell'acqua, avendo una lunghezza di assorbimento molto grande, da 85 m nel blu fino ad oltre 200 m nell'UV. Però la lunghezza di diffusione è minore di quella in acqua⁶, e si hanno quindi maggiori ritardi nei tempi di hit, che risultano in una minore accuratezza di ricostruzione, che si deve fronteggiare con elaborati algoritmi di filtraggio.

Anche AMANDA ha raccolto dati compatibili con il flusso atteso di neutrini atmosferici, come rappresentato in Figura 31, senza trovare evidenza di eventi cosmici [37].

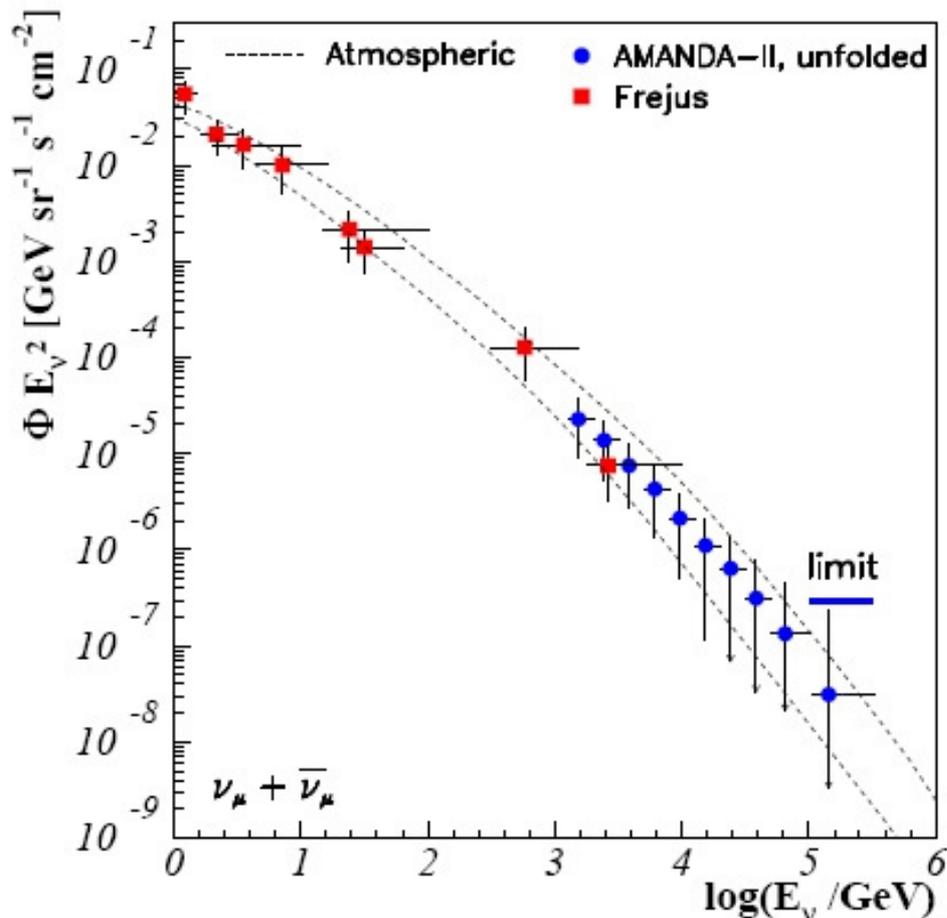


Figura 31 – Flusso di eventi misurati da AMANDA [37].

⁶ La lunghezza di assorbimento tiene conto dell'attenuazione dovuta all'assorbimento dei fotoni da parte delle molecole, mentre la lunghezza di diffusione tiene conto dello scattering.

L'estensione futura di AMANDA, un telescopio di scala 1 km^3 chiamato Ice-cube sarà messo in funzione verso il 2010 sempre nei pressi del polo sud (Figura 32).

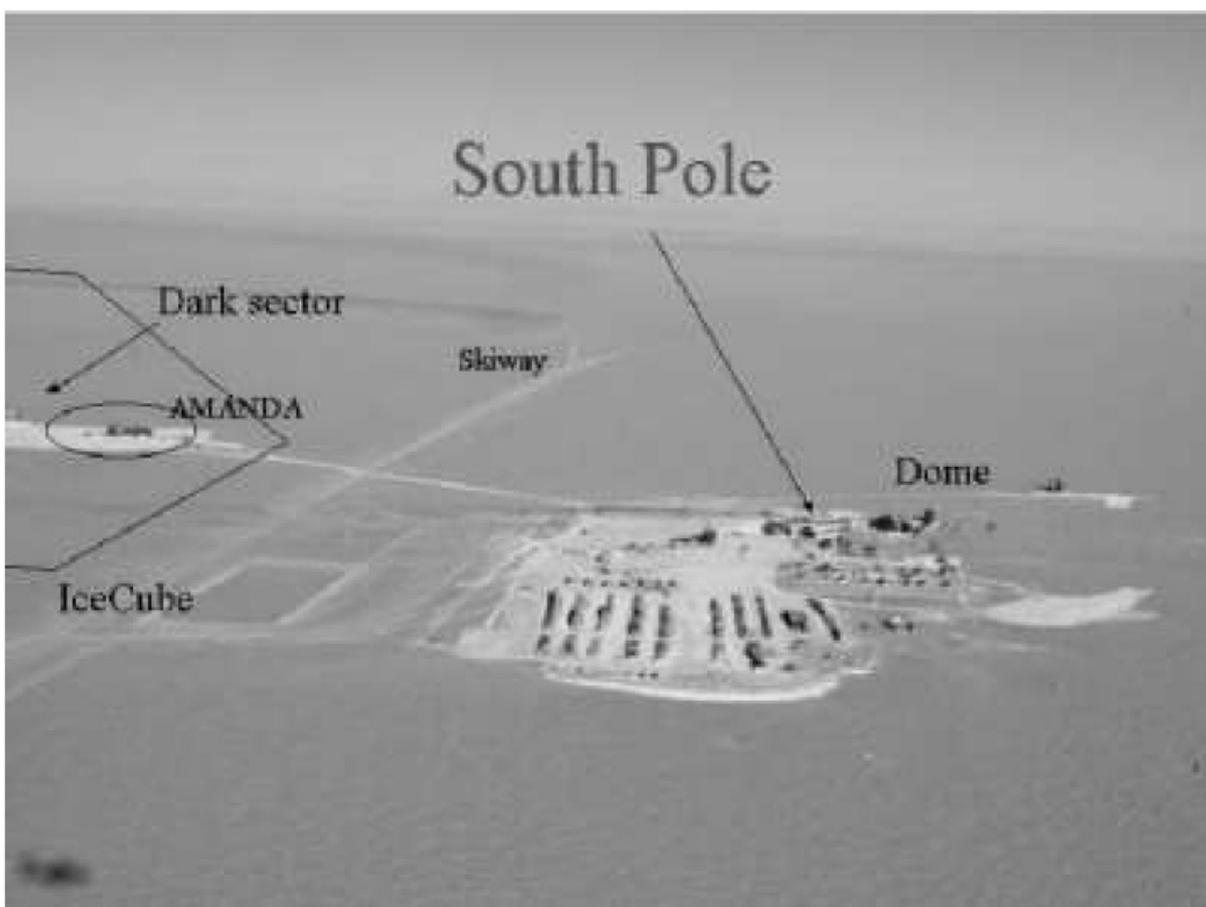


Figura 32 – Posizione di AMANDA e di Ice-cube, presso i laboratori dell'Antartide.

Sarà formato da 4800 moduli ottici disposti in circa 80 stringhe disposte in un'area di circa 1 km^2 , ad una profondità tra 1400 e 2400 m.

2.9 Eventi attesi e fondo

I neutrini interessanti dal punto di vista astrofisico sono immersi in una grande quantità di neutrini e muoni atmosferici. Mentre i muoni generati negli sciami atmosferici estesi provengono solo dall'alto, i neutrini atmosferici provenienti dal basso sono indistinguibili da quelli di origine astrofisica. La rivelazione di sorgenti puntiformi avverrà quindi solo su base statistica, osservando un'abbondanza di eventi rispetto al numero di eventi aspettato solo a causa del fondo, da certe direzioni. La direzione di arrivo potrà essere stimata con un fit gaussiano, mediando così le incertezze sulla ricostruzione della direzione dei singoli eventi

(cfr. §2.2) ottenendo accuratzeze migliori al crescere del numero di eventi osservati:

$$\sigma_{\hat{\theta}} \approx \sigma_{ev} \frac{1}{\sqrt{N}}.$$

I flussi di neutrini aspettati in base alle stime del §1.7, sono riassunti in Figura 33 [27]. Sopra al TeV i neutrini astrofisici (curva g in Figura 33) hanno un flusso pari o superiore a quello dei neutrini atmosferici (curve f ed f').

In Tabella 1 e Tabella 2 sono riportate le stime degli eventi attesi in un telescopio a neutrini di area efficace 0.1 km² e 1 km² rispettivamente [14], [29], [30], per differenti potenziali sorgenti di neutrini (sia estese che puntiformi).

Atmosferici (eventi/sr)	>1GeV >1TeV	8000 120÷170	
Astrofisici (eventi/sr)	>1GeV >1TeV	Piano Galattico 12÷20 1.5÷3	AGN (flusso diffuso) 80÷200 40÷200
Atmosferici (in un cerchio di 1°)	>1GeV >1TeV	cos θ = 0.05 12.6 0.21	cos θ = 0.95 5.6 0.05
Astrofisici (sorgenti puntiformi)	>1GeV >1TeV	2.6 Galattici (SNR, microquasar...) 0.1÷10 Extragalattici (AGN, GRB...)	

Tabella 1 - Eventi attesi all'anno per un telescopio a neutrini di area efficace 0.1 km².

Atmosferici (eventi/sr)	5 10⁵
Centro galattico (eventi/sr)	300÷1000
Sorgenti di GRB (eventi/sr)	200
AGN jets (eventi/sr)	500
AGN cores (eventi/sr)	3 10⁴

Tabella 2 - Eventi attesi all'anno per un telescopio a neutrini di area efficace 1 km².

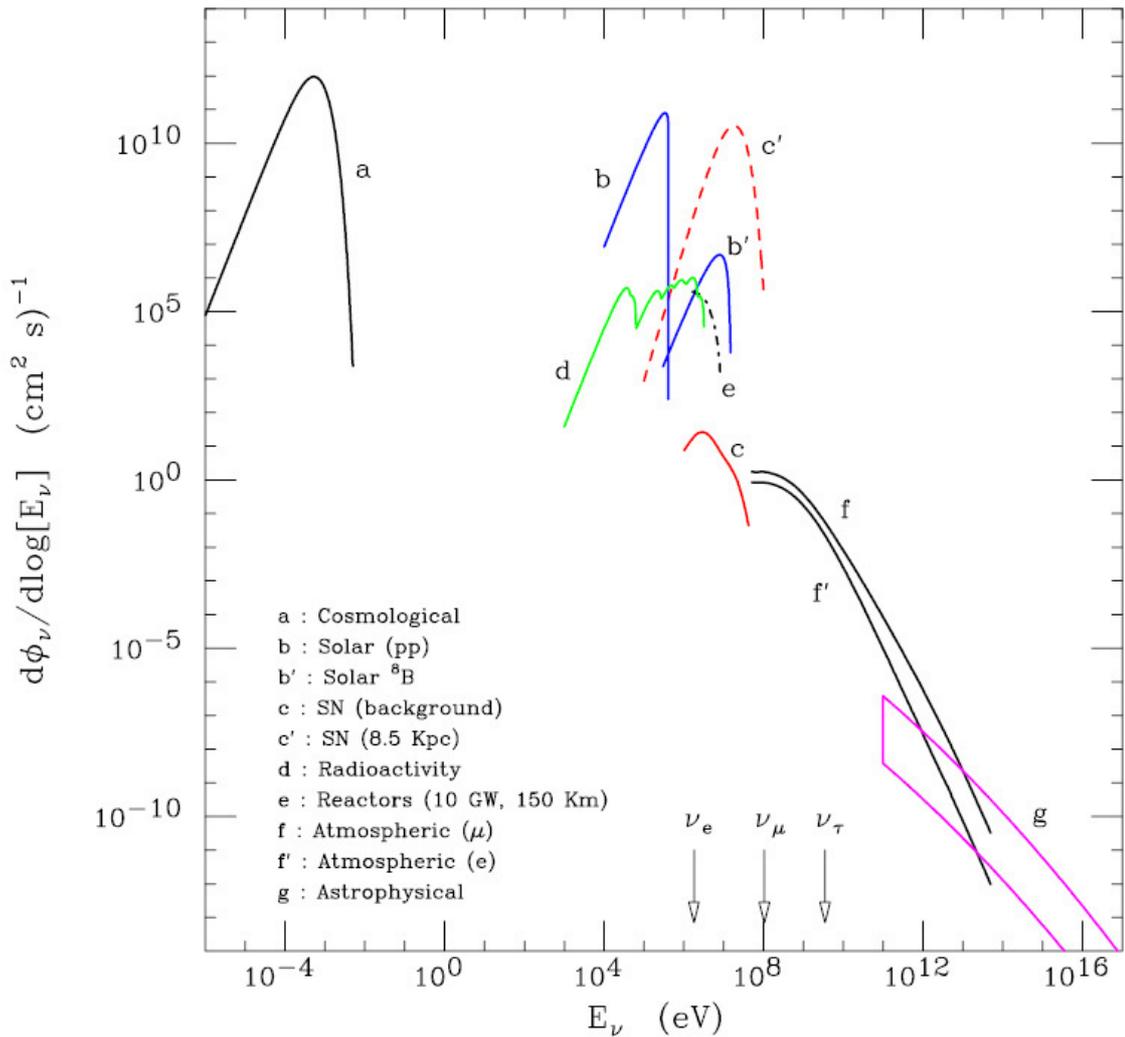
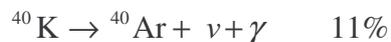
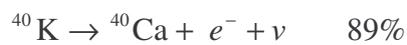


Figura 33 - Flussi attesi di neutrini da varie sorgenti [27].

Come si vede dalle tabelle gli eventi da sorgenti puntiformi sono solo qualche unità all'anno per il telescopio da 0.1 km^2 (cioè la taglia di ANTARES).

Oltre al fondo costituito da neutrini e muoni atmosferici, un altro contributo importante di disturbo nella rivelazione di neutrini astrofisici è dato dall'emissione di fotoni da parte degli elementi radioattivi contenuti nell'acqua di mare. Tra le sorgenti più importanti ci sono il ${}^{40}\text{K}$, il ${}^{238}\text{U}$ e il ${}^{87}\text{Rb}$. Ma la quasi totalità del fondo è dato dal decadimento del ${}^{40}\text{K}$



L'attività, cioè il numero di decadimenti per unità di massa e di tempo, può essere calcolata mediante [23]

$$N = \frac{\ln 2}{\tau} S \varepsilon_K \eta_{^{40}\text{K}} \frac{N_A}{A}$$

dove $\tau = 1.28 \cdot 10^9$ anni è la vita media del ^{40}K , $S = 38.4$ g/l è la salinità dell'acqua del mare Mediterraneo, ε_K è la percentuale di potassio rispetto agli altri sali, $\eta_{^{40}\text{K}}$ è l'abbondanza relativa dell'isotopo ^{40}K , N_A è il numero di Avogadro e A è il peso atomico.

Risulta un'attività pari a 13600 decadimenti al m^3 al s, che dà una frequenza media di segnale sul singolo PMT di circa 35 kHz. Si capisce quindi come sia importante utilizzare opportune logiche di trigger per registrare solo dati interessanti e logiche di filtraggio per ricostruire la traccia dell'evento in modo accurato.

Un'altra causa di fondo di fotoni nella banda del visibile è data dai microorganismi bioluminescenti, che producono impulsi di luce di durata variabile dai ms al secondo e temporalmente correlati. I PMT eventualmente interessati da questo tipo di fondo sono completamente accecati da questi “burst” di bioluminescenza. Fortunatamente la massima densità di questi organismi si trova a profondità minori di quelle dei telescopi NEMO ed ANTARES, come si vede dalla Figura 34.

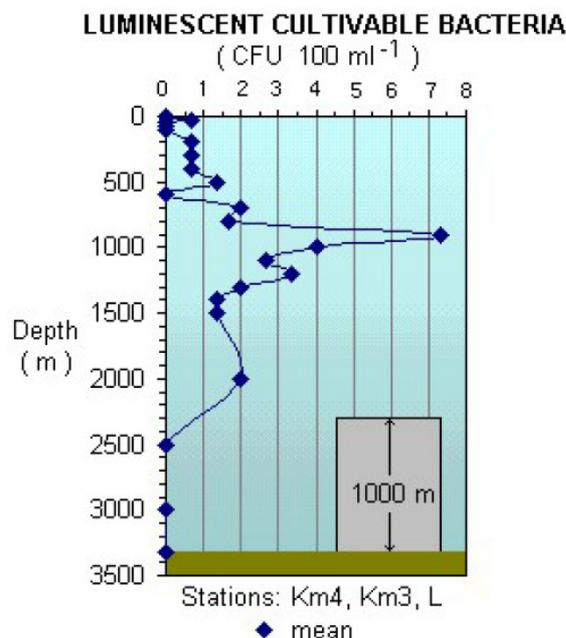


Figura 34 - Densità di organismi luminescenti in funzione della profondità nel sito di Capo Passero (CFU: Colony Forming Units) [28].

Cap. 3 Ricostruzione delle tracce di muoni con algoritmi standard

3.1 Generazione delle tracce: OPNEMO

La collaborazione NEMO ha già sviluppato un programma di simulazione di tracce di muoni indotte da neutrini, in linguaggio fortran, chiamato opnemo. È strutturato, come indicato nel diagramma di flusso di Figura 35, con le tre subroutine principali

- ✓ KINEMA
- ✓ GENERA
- ✓ PHODET

che si occupano, rispettivamente, della generazione dei parametri geometrici che individuano le tracce di muoni (KINEMA), della generazione dei fotoni emessi sia per effetto Čerenkov che per sciami elettromagnetici lungo la traccia (GENERA), e della simulazione dell'effetto del fotomoltiplicatore (PHODET).

La routine KINEMA genera una traccia di un muone con energia random uniforme in $\log(E)$; la distribuzione in energia dei raggi cosmici sappiamo essere più ripida (cfr §1.2), quindi lo spettro generato non è del tutto realistico, ma per stimolare il detector a tutte le energie di interesse, mantenendo una statistica significativa di eventi senza allungare eccessivamente il tempo di simulazione, risulta più comoda la distribuzione piatta.

La distribuzione angolare in θ (angolo rispetto allo zenit) può essere scelta attraverso un parametro di controllo: uniforme in $\cos(\theta)$ (muoni atmosferici) o uniforme in θ da $-\pi/2$ a $\pi/2$ (muoni da neutrini), mentre l'azimuth ϕ è sempre generato in modo uniforme tra 0 e 2π . Il vertice della traccia viene generato in maniera casuale sulla superficie di un cubo che contiene completamente l'apparato. La dimensione del lato di questo cubo si ottiene aggiungendo alle

dimensioni reali dell'apparato 5 lunghezze di attenuazione dei fotoni in acqua. Nel caso di OPNEMO è stato scelto pari a 1890 m, per una superficie della sezione (A_G area geometrica) pari a 3.57 km^2 . La superficie totale di generazione è pari a 21.4 km^2 . Il volume geometrico della scatola è 6.75 km^3 .

Il centro di questa scatola coincide con il centro dell'apparato.

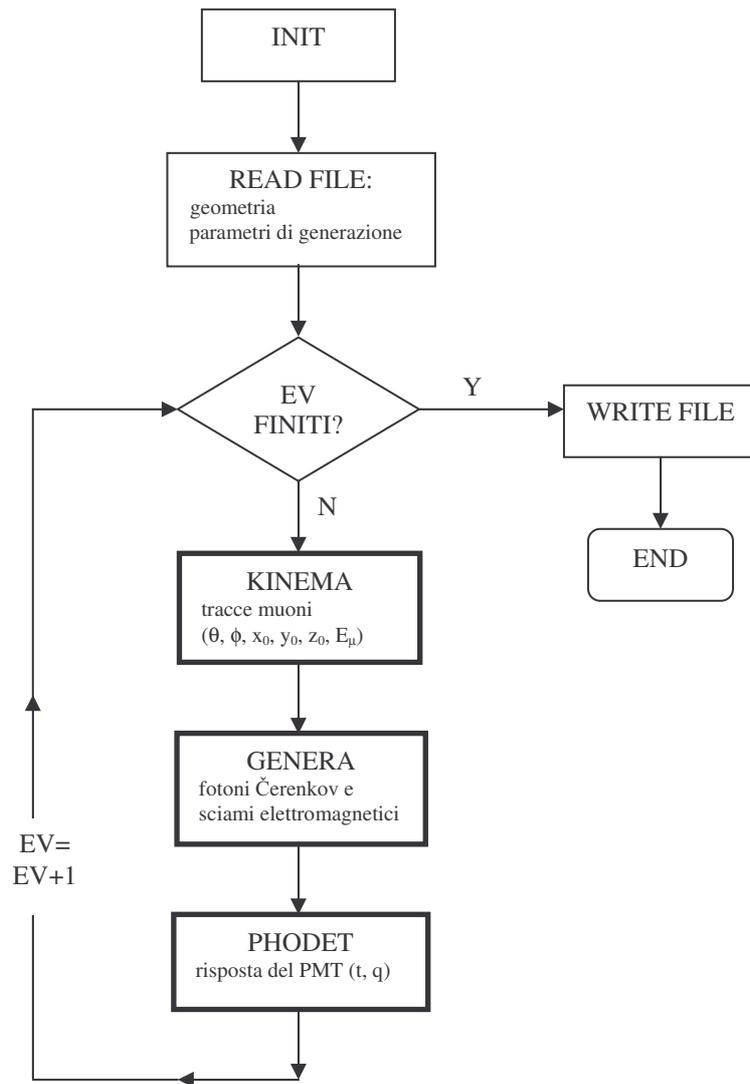


Figura 35 – Diagramma di flusso del programma di simulazione di eventi OPNEMO.

La routine GENERA calcola il numero effettivo di fotoelettroni generati su ogni PMT. Per prima cosa, attraverso la chiamata alla subroutine MU_BURST, data l'energia del muone, viene stimata la lunghezza della traccia dovuta solo alle perdite per ionizzazione (pressoché costanti); la traccia viene quindi suddivisa in un numero sufficientemente elevato di step (attualmente 3000) e per ogni step vengono estratte le probabilità di emissione per

bremstrahlung, creazione di coppie e interazioni nucleari. Se i vari processi devono essere generati vengono calcolate le densità di fotoni emessi e viene diminuita l'energia del muone. La subroutine PHOMIZ esegue il calcolo della densità di fotoni che raggiungono ciascun PMT, e del loro tempo di arrivo. Infine la routine OMSIM, partendo dalla densità di fotoni, considerando l'area del fotomoltiplicatore effettivamente esposta al loro flusso (proiezione lungo la direzione di arrivo), e tenendo presente l'efficienza dell'accoppiamento ottico e l'efficienza quantica, determina il numero di fotoelettroni generati in ciascun PMT.

La routine PHODET non è ancora definitiva, per adesso simula in modo approssimato il comportamento del PMT, generando una carica pari a quella nominale di singolo fotoelettrone (p.e.) moltiplicata per il numero di fotoelettroni incidenti più una variabile aleatoria gaussiana che tiene conto dello spreading nello spettro. Il processo di moltiplicazione in un PMT dà luogo in realtà ad una distribuzione delle cariche in uscita con code più lunghe di una gaussiana, e può essere approssimata da una distribuzione Polya. Si avrà un hit nel PMT quando la carica integrata all'uscita supera una soglia fissata a $\frac{1}{4}$ p.e.

Alla fine, per ogni evento, vengono calcolati i tempi e le cariche di ogni hit che vengono poi memorizzati in un file di uscita in formato ASCII (Tabella 3), insieme alle variabili cinematiche dell'evento stesso. Il file di eventi sarà poi letto ed elaborato dai programmi di ricostruzione.

GEOMETRIA										
1	-560.	-570.	-340.	0.	0.	-1.	0.050999999	2.5	0.25	1
2	-560.	-570.	-340.	0.	0.	1.	0.050999999	2.5	0.25	1
...										
5832	570.	560.	340.	0.	0.	1.	0.050999999	2.5	0.25	81
TORRI										
1	-560.	-560.								
2	-420.	-560.								
...										
81	560.	560.								
N_EV										
ievent,nhit,tfirst,p1,p2,p3,v1,v2,v3,thvero,phvero,energia										
index _{PMTi} i=1..nhit										
PMT _i , charge, flag, time 1 st p.e., num p.e.										
time j-th p.e. j=1..num p.e.										
charge j-th p.e. j=1..num p.e.										
...										
PMT _{nhit} , charge, flag, time 1 st p.e., num p.e.										
time j-th p.e. j=1..num p.e.										
charge j-th p.e. j=1..num p.e.										
ievent,nhit,tfirst,p1,p2,p3,v1,v2,v3,thvero,phvero,energia										
index _{PMTi} i=1..nhit										
PMT _i , charge, flag, time 1 st p.e., num p.e.										
time j-th p.e. j=1..num p.e.										
charge j-th p.e. j=1..num p.e.										
...										

Tabella 3 – Formato del file ASCII di eventi in uscita di OPNEMO.

3.2 Ricostruzione geometrica degli eventi

Per quanto detto nel §2.3, la traccia di un muone con energia superiore a 50 GeV può essere considerata rettilinea, e la velocità del muone pari a quella della luce nel vuoto. La traccia è quindi completamente identificata dalla posizione $\vec{p} = [x_0, y_0, z_0]$ del muone al tempo di riferimento t_0 , e dai coseni direttori:

$$\vec{d} = [\cos\phi \sin\theta \quad \sin\phi \sin\theta \quad \cos\theta]$$

dove ϕ e θ in un sistema di riferimento cartesiano (x, y, z) sono l'azimuth e l'angolo rispetto allo zenith rispettivamente, come rappresentato in Figura 36 a destra.

Tenendo presente la Figura 36, il tempo t_{PMTi} di arrivo del generico fotone Čerenkov (emesso con un angolo θ_C rispetto alla traccia) sull' i -esimo PMT del rivelatore, in posizione $\vec{q} = (x_i, y_i, z_i)$, può essere calcolato sommando il tempo t_1 (t_{OA}) che il muone impiega per arrivare al punto di emissione del fotone, al tempo t_2 (t_{AB}) di propagazione del fotone nel mezzo con indice di rifrazione n (linea tratteggiata) [49]:

$$t_{PMTi} = t_0 + t_1 + t_2$$

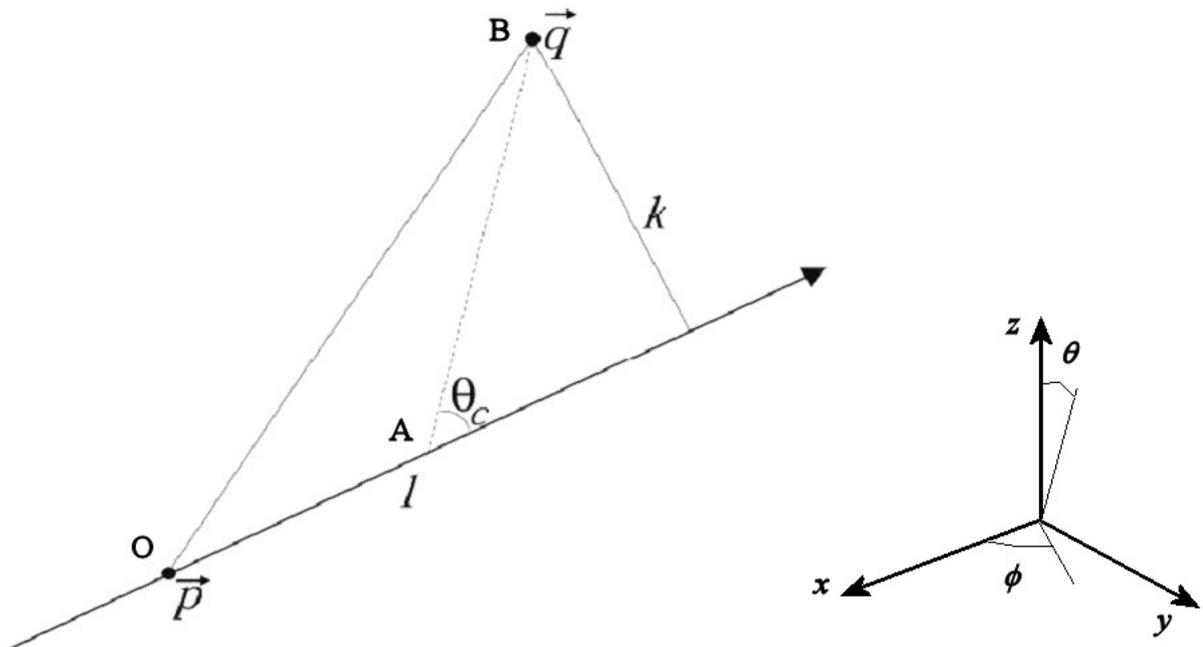


Figura 36 - Traiettoria del muone e del generico fotone Čerenkov, e a destra sistema di riferimento.

Definendo

$$\vec{v} = \vec{q} - \vec{p}$$

si ha

$$l = \vec{v} \cdot \vec{d}$$

e

$$k = \sqrt{|\vec{v}|^2 - l^2}$$

che è la minima distanza della traccia dal PMT_i.

Il tempo di arrivo su \vec{q} (t_i) è allora

$$t_i = t_0 + \frac{1}{c} \left(l - \frac{k}{\tan \theta_C} \right) + \frac{1}{v_g} \frac{k}{\sin \theta_C} \quad (5)$$

dove v_g è la velocità di gruppo della luce nell'acqua $v_g = \frac{c}{n}$ (n è l'indice di rifrazione dell'acqua).

Si noti che se t_0 è maggiore di t_i , l è negativo, mentre k è sempre positivo, e la formula (5) continua a valere.

La (5) può essere posta nella forma

$$t_i = t_0 + \frac{1}{c} (l + \kappa k) \quad (6)$$

con

$$\kappa = \frac{n}{\sin \theta_C} - \frac{1}{\tan \theta_C} = \tan \theta_C$$

dove abbiamo usato l'espressione (cfr. §2.4)

$$n = \frac{1}{\cos \theta_C}$$

Si può dimostrare che qualunque altro fotone emesso lungo la traiettoria del muone, per bremsstrahlung o a causa di sciami elettromagnetici, arriva al generico PMT dopo il fotone Čerenkov. Cioè il tempo nella traiettoria dei fotoni Čerenkov è il minimo rispetto a tutte le possibili traiettorie per giungere sull' i -esimo PMT (analogo del teorema di Fermat del minimo cammino ottico).

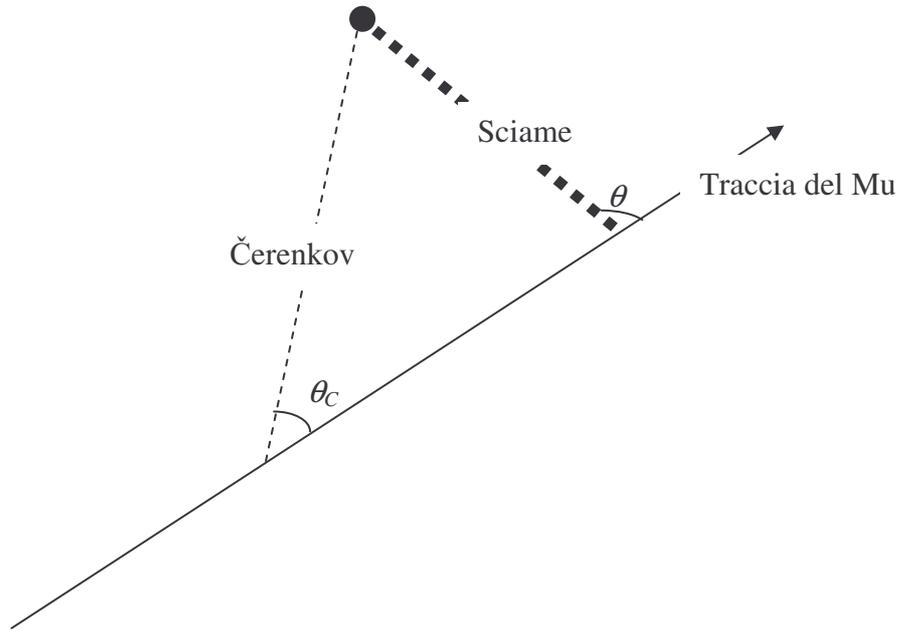


Figura 37 – Traiettoria dei fotoni Čerenkov e di uno sciame.

L'espressione (5) per il calcolo del tempo di arrivo di un fotone continua a valere anche per i fotoni provenienti da sciame aventi un generico angolo di emissione θ rispetto alla traccia (Figura 37):

$$t(\theta) = t_0 + \frac{l}{c} + \left(\frac{n}{\sin \theta} - \frac{1}{\tan \theta} \right) \frac{k}{c} \quad (7)$$

Derivando la (7) rispetto all'angolo di emissione θ si ha:

$$\frac{dt}{d\theta} = -\frac{nk}{c} \frac{\cos \theta}{\sin^2 \theta} + \frac{k}{c} \frac{1}{\sin^2 \theta} = \frac{k}{c} \frac{1}{\sin^2 \theta} (1 - n \cos \theta)$$

che si annulla per $\cos \theta = 1/n \equiv \cos \theta_c$ che è proprio l'angolo di emissione dei fotoni Čerenkov.

In Figura 38 è riportato il tempo di arrivo dei fotoni emessi lungo la traccia di un muone che attraversa l'acqua di mare ($n = 1.355$, $\theta_c = 42.5^\circ$) in funzione dell'angolo di emissione θ rispetto alla traccia. Per il calcolo di t è stata considerata una traccia diretta lungo l'asse x e un PMT di coordinate (10, 1, 0).

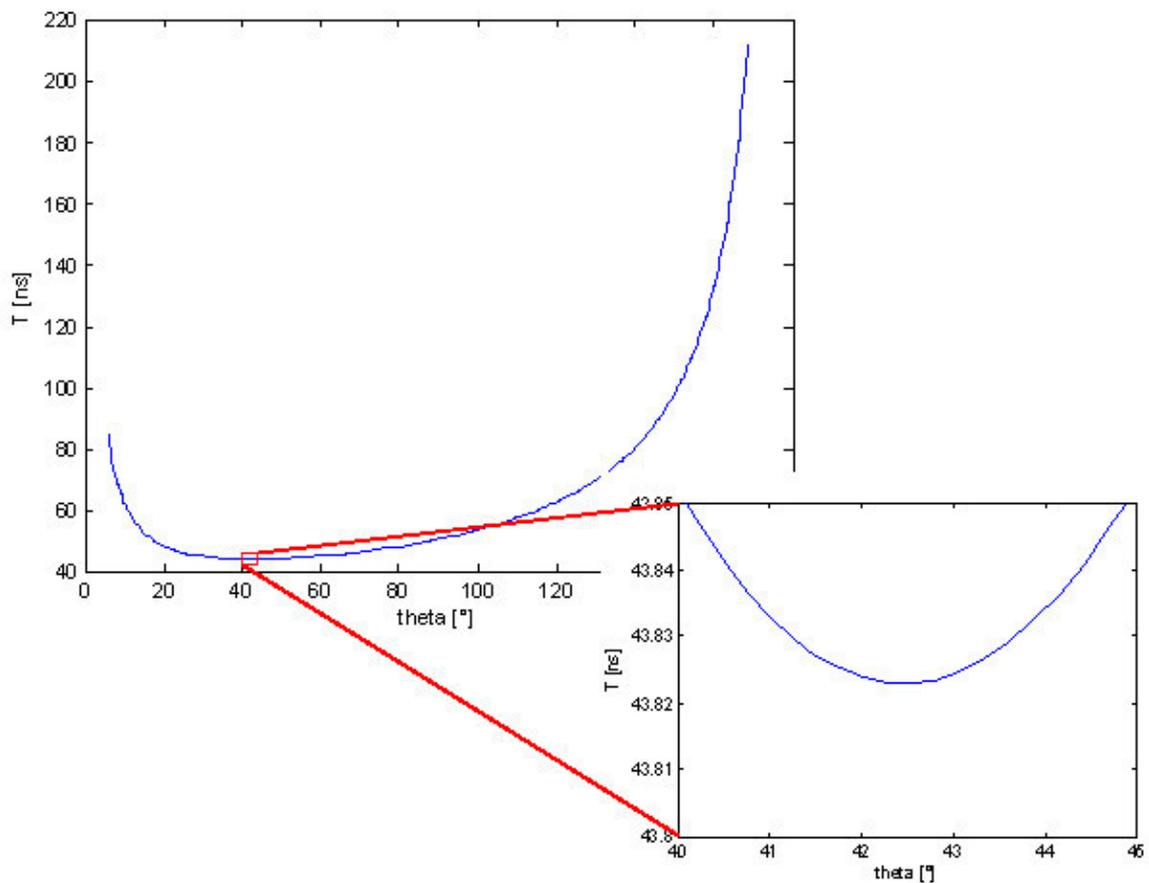


Figura 38 – Tempo di arrivo di un fotone al variare dell'angolo di emissione dalla traccia del muone.

Come si vede dalla figura il minimo di $t(\theta)$ si ha intorno ai 42.5° che corrispondono proprio all'angolo di emissione della luce Čerenkov in acqua.

Per ogni quintupla di parametri $(x_0, y_0, z_0, \theta, \phi)$, che identifica univocamente la traccia del muone, la relazione (6) permette di calcolare il tempo di arrivo dei fotoni Čerenkov su ogni PMT dell'apparato. Quindi, dati i tempi di arrivo effettivamente misurati (hit), per ricostruire la traccia del muone si può costruire una funzione dei cinque parametri $(x_0, y_0, z_0, \theta, \phi)$ che rappresenta l'errore quadratico medio tra i tempi di hit teorici e quelli misurati, e minimizzare tale funzione per trovare i valori geometrici della traccia nello spazio. Questo equivale a fare un fit di una funzione non lineare, in cui la funzione di errore, se opportunamente normalizzata all'incertezza sulla misura dei tempi di hit, rappresenta il χ^2 .

Poiché i parametri di una traccia sono cinque, occorrono almeno cinque hit indipendenti (fototubi distinti e non allineati) per avere una soluzione univoca del problema. Gli eventi con meno di cinque hit saranno quindi automaticamente scartati a priori nel

processo di ricostruzione, e la frazione di tali eventi pesa sul calcolo dell'area efficace dell'apparato. Nel file di eventi generato per le simulazioni si ha una percentuale di eventi con meno di 5 hit pari al 22.5%.

Se tutti i tempi fossero misurati con accuratezza infinita la funzione di errore si annullerebbe in corrispondenza dei parametri veri della traccia. La presenza della fluttuazione del tempo di transito dei fotoelettroni (TTS) nei PMT, e soprattutto il ritardo dei fotoni dovuto allo scattering e agli sciami elettromagnetici, e la possibile presenza di segnali spuri dovuti al fondo del ^{40}K e della bioluminescenza, fanno sì che la funzione di errore non si annulli. In ogni caso minimizzandola rispetto ai 5 parametri si può ottenere la migliore stima della traccia che meglio approssima la distribuzione dei tempi teorici.

Dato che la dimensione del fotocatodo dei PMT è dell'ordine di ~ 20 cm, un fotone incidente su un PMT che dà luogo ad un hit, può essere emesso in un intervallo di lunghezza di poco superiore lungo la traccia del muone; l'incertezza dovuta all'istante di emissione del fotone Čerenkov in questo intervallo di lunghezza risulta minore di 1ns ed è pertanto trascurabile rispetto alla fluttuazione dei tempi di transito nei PMT ($2 \div 3$ ns).

L'espressione della funzione da minimizzare è:

$$\left(\hat{x}_0, \hat{y}_0, \hat{z}_0, \hat{\theta}, \hat{\phi}\right) = \min \left\{ \sum_{k \in \text{hit}} \left(T_k^{(MIS)} - T_k^{(TEO)}(x_0, y_0, z_0, \theta, \phi) \right)^2 \right\} \quad (8)$$

dove $T_k^{(MIS)}$ è il tempo di hit misurato sul k-esimo PMT dell'apparato, e $T_k^{(TEO)}$ è il tempo di arrivo teorico sullo stesso PMT, calcolato con la (6) in base ai parametri della traccia simulata. Normalizzando la funzione di errore con l'incertezza sui tempi si ottiene il χ^2 :

$$\chi^2 = \sum_{k \in \text{hit}} \frac{\left(T_k^{(MIS)} - T_k^{(TEO)}(x_0, y_0, z_0, \theta, \phi) \right)^2}{\sigma^2}$$

In Figura 39 un evento simulato con il programma OPNEMO (cfr. §3.1), con le seguenti caratteristiche:

energiadel muone ~ 8 TeV

numero di hit 89

ϕ vero 250.48°

θ vero 29.62°

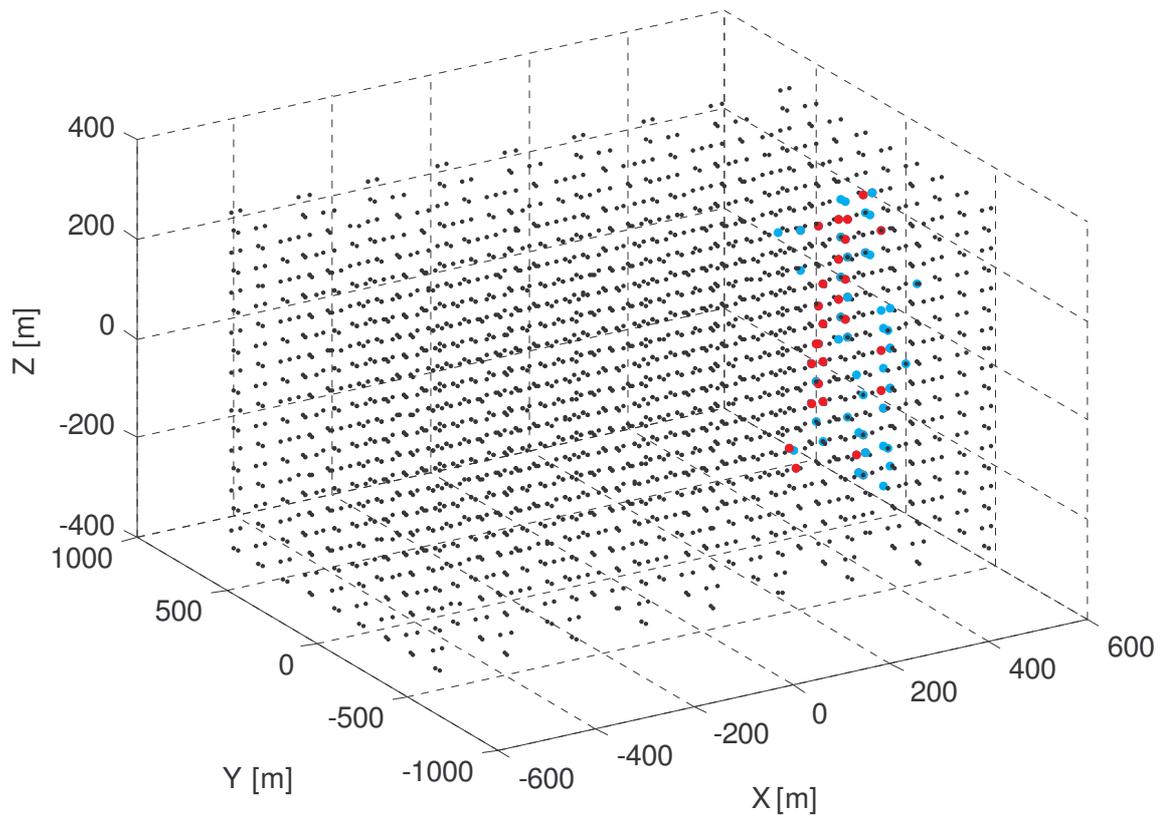


Figura 39 - Evento simulato con OPNEMO, $E_\mu = 8$ TeV, $N_{\text{HIT}} = 89$, $\theta = 29.62^\circ$, $\phi = 250.48^\circ$.

In Figura 40 è rappresentato il χ^2 ottenuto con il programma di ricostruzione geometrica dell'evento di Figura 39 avendo imposto per le coordinate (x, y, z) del punto di impatto del muone quelle "vere" x_0, y_0, z_0 , del vertice della traccia.

L'andamento del χ^2 è parabolico nell'intorno del minimo. In figura il minimo del χ^2 è indicato dall'asterisco nero. L'asterisco rosso rappresenta invece la direzione "vera" della traccia simulata.

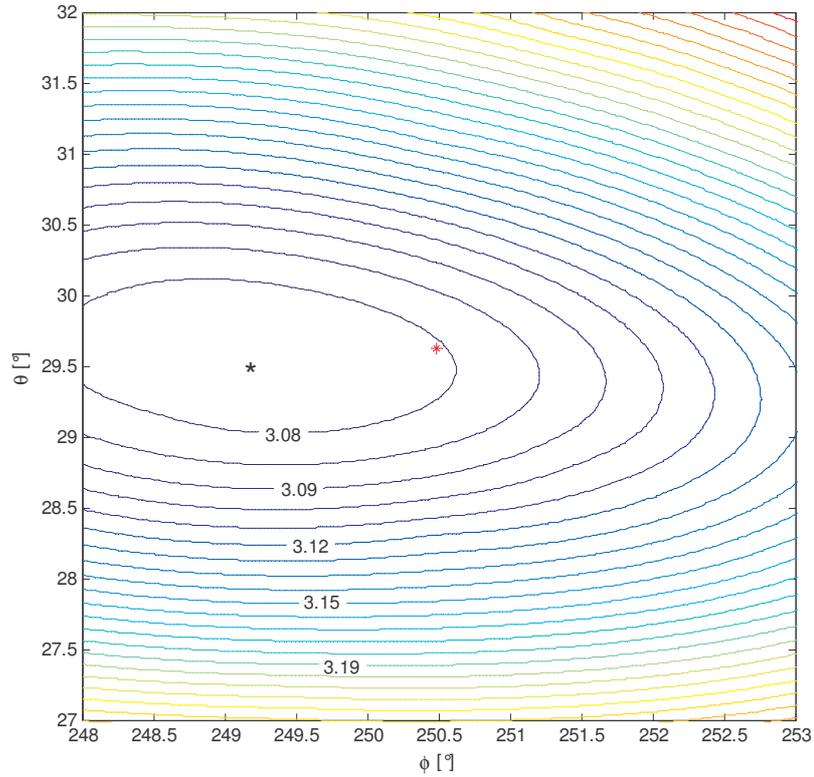


Figura 40 - Curve di livello della funzione di errore di un evento simulato con il programma OPNEMO. La posizione iniziale è fissata al suo valore vero, l'asterisco rosso indica la direzione vera, l'asterisco nero indica la coppia di valori (θ, ϕ) per i quali la (8) assume il suo valore minimo.

La discrepanza tra il minimo della funzione di errore e la direzione vera della traccia, pari a circa 0.7° , è dovuta ai tempi di arrivo dei fotoni diffusi ed a quelli emessi negli sciami elettromagnetici. Questi come si è detto arrivano in ritardo rispetto ai fotoni emessi per effetto Čerenkov; in Figura 41 sono riportati gli scarti dei tempi degli hit dell'evento calcolati come $\Delta T = T_{\text{Misurato}} - T_{\text{Teorico}}$.

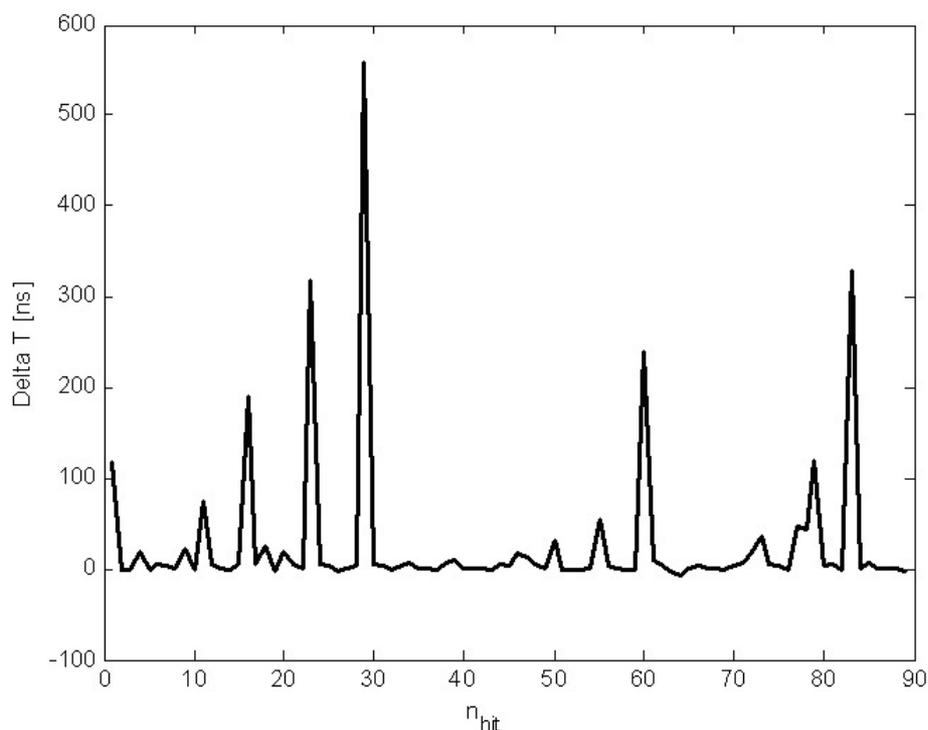


Figura 41 - Differenza ΔT tra i tempi misurati e quelli aspettati per i fotoni Čerenkov per ogni hit dell'evento della Figura 39.

Gli hit con i tempi misurati che si discostano maggiormente da quelli aspettati “sporcano” la ricostruzione della traccia e possono essere eliminati con l’applicazione dei filtri software. Possiamo qui anticipare brevemente alcuni risultati che possono essere ottenuti applicando delle condizioni di causalità spazio-temporali che verranno descritte in seguito (cfr. §3.3). In Figura 42 è riportata la funzione di errore, in funzione di θ e ϕ , dopo il filtraggio e in Figura 43 gli scarti dei tempi di hit che sopravvivono al filtro (soglia posta a 200 ns). Si può osservare che dopo il filtraggio il minimo della funzione si avvicina al valore vero della direzione della traccia, e si fa più pronunciato con un errore angolare che è ora pari a circa 0.2° . Si osservi che alcuni hit con discrepanze temporali significativamente maggiori dell'incertezza (TTS) riescono a passare il filtro, mostrando errori di picco nei tempi di arrivo pari a 50 ns.

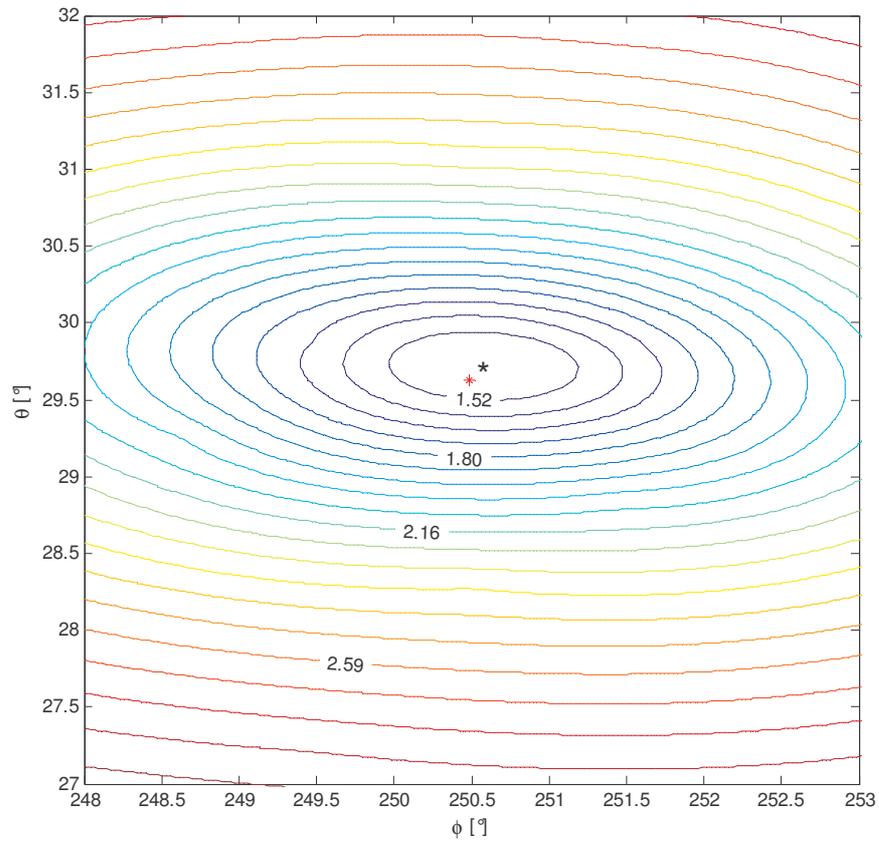


Figura 42 - Funzione di errore dell'evento della Figura 39 dopo i filtri.

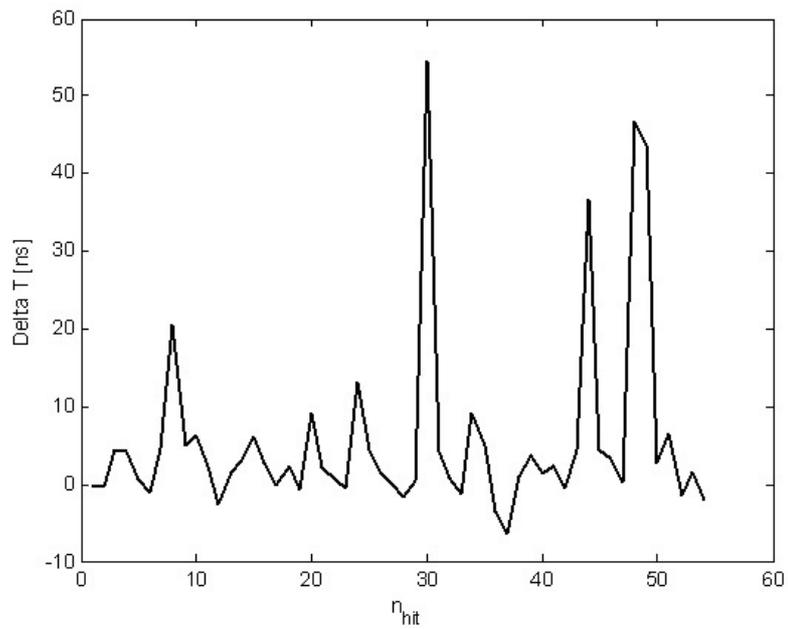


Figura 43 - Errori nei tempi di hit dell'evento della Figura 39 dopo i filtri.

Si osservi che ora si notano anche degli scarti negativi, dovuti alla statistica gaussiana dello spreading dei tempi di transito nei PMT (TTS).

3.3 Filtri e trigger

L'impatto del fondo del ^{40}K sul data rate atteso da un esperimento come NEMO è enorme (una frequenza tipica di hit da ^{40}K su singolo PMT è di circa 35 kHz, che moltiplicata per 5832 PMT dà una frequenza di hit di 200 MHz), ed è importante quindi poter riconoscere velocemente un evento significativo per trasmettere a terra solo i dati relativi ai momenti in cui è presente (con una buona probabilità) una traccia, cioè il cosiddetto trigger di livello 0.

La bontà di un trigger si misura con due caratteristiche contrastanti: la selettività nei confronti del fondo e la efficienza nei confronti dei segnali veri. Si tratta di un tipico test di ipotesi, cioè dato un set di misure appartenenti al segnale o al rumore un test statistico può commettere due tipi di errore:

- ✓ può scartare un segnale buono interpretandolo come rumore (errore di tipo I, mancata rivelazione)
- ✓ può considerare come buono un segnale composto da solo rumore (errore di tipo II, falso allarme)

Gli errori di tipo I influiscono sulla efficienza, mentre gli errori di tipo II determinano la selettività.

Nelle collaborazioni NEMO ed ANTARES sono stati proposti e studiati vari tipi di trigger [52][53][54]. In generale si può imporre una condizione sulla coincidenza temporale di più hit, oppure sulla carica misurata da uno o più PMT, o combinazioni delle due tecniche.

Il trigger basato sulla coincidenza temporale (entro un intervallo prefissato) di tre hit su PMT di una stessa torre o di un piano ha un'efficienza accettabile, maggiore dell'80% il primo, > 60% il secondo, per $E_\mu > 1 \text{ TeV}$. Qui l'efficienza è definita rispetto al numero di eventi ricostruibili in linea di principio, cioè quelli con almeno 5 hit. In Figura 44 è riportata l'efficienza di trigger al variare dell'energia dell'evento per tre tipi di trigger di selettività crescente: 3 hit in una torre, 3 hit in un piano, 3 hit in un piano per almeno due torri. Alle energie più basse (range minore, meno hit su ogni torre) è più probabile perdere eventi. Tuttavia la selettività di questi trigger non è ottimale [53], la frequenza del fondo è abbastanza elevata da rendere le coincidenze casuali di tre hit non troppo rare e c'è il rischio di sovraccaricare la linea di comunicazione tra l'apparato e terra: questo accade quando prima

che sia terminata la trasmissione dei dati di una finestra intorno ad un trigger, si ha già un nuovo segnale di trigger, evidentemente questo verrà scartato.

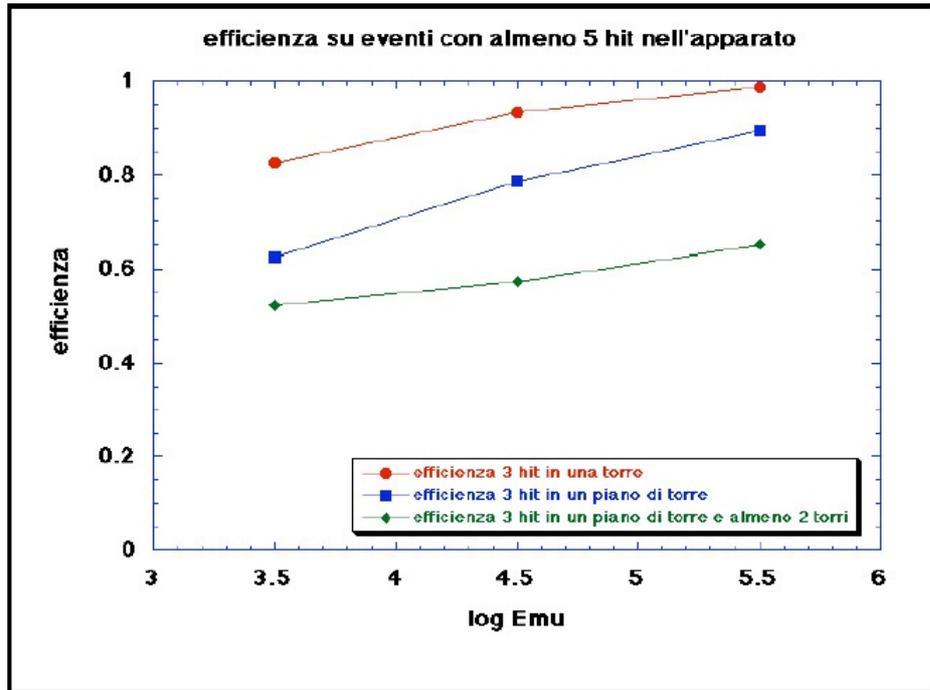


Figura 44 - Efficienza di diversi tipi di trigger al variare dell'energia.

Una volta che il trigger è scattato vengono inviati a terra i dati relativi a tutti i PMT che hanno ricevuto un segnale in una finestra temporale ΔT_{ev} centrata intorno all'istante di trigger, abbastanza ampia da contenere tutti i segnali (causalità). Per NEMO, in cui le distanze massime tra i PMT sono dell'ordine del km, una finestra temporale in grado di contenere tutti i dati è dell'ordine di ± 5500 ns.

Gli hit dovuti al segnale sono completamente sommersi dagli hit casuali del decadimento del ^{40}K : per una tipica frequenza f_K di singolo hit da ^{40}K di 35 kHz si ha un numero medio di hit in un evento pari a

$$N_K \approx f_K \Delta T_{ev} N_{PMT} = 2250 \quad \text{con } N_{PMT} = 5832$$

a fronte di eventi che possono teoricamente avere un numero di hit superiore a 5.

Le tecniche di filtraggio sono quindi molto importanti per avere dati affidabili su cui operare i fit di ricostruzione traccia.

Prima di tutto si impone la condizione di causalità su tutto l'apparato rispetto al trigger: due segnali di hit relativi a PMT ad una certa distanza d non possono essere separati di un tempo maggiore di d/c . Naturalmente questa condizione, sebbene sia un buon filtraggio

iniziale che riduce di molto il numero di hit casuali, diventa via via più debole man mano che si prendono in esame PMT più distanti dal trigger.

Si utilizza allora il filtro a cluster: si cercano gruppi di hit correlati in tempo e con carica superiore a una data soglia (ad es. 2 fotoelettroni: p.e.), entro distanze contenute (superando così il problema descritto sopra della diminuzione della selettività del filtro causale a grandi distanze).

L'applicazione di questi filtri riesce ad evidenziare soltanto gli hit della traccia, pur con una certa perdita di hit e di eventi a bassa energia. In Figura 45 sono riportati gli hit di un evento con e senza fondo, in rosso gli hit scartati dal filtro, in giallo gli hit usati per la ricostruzione.

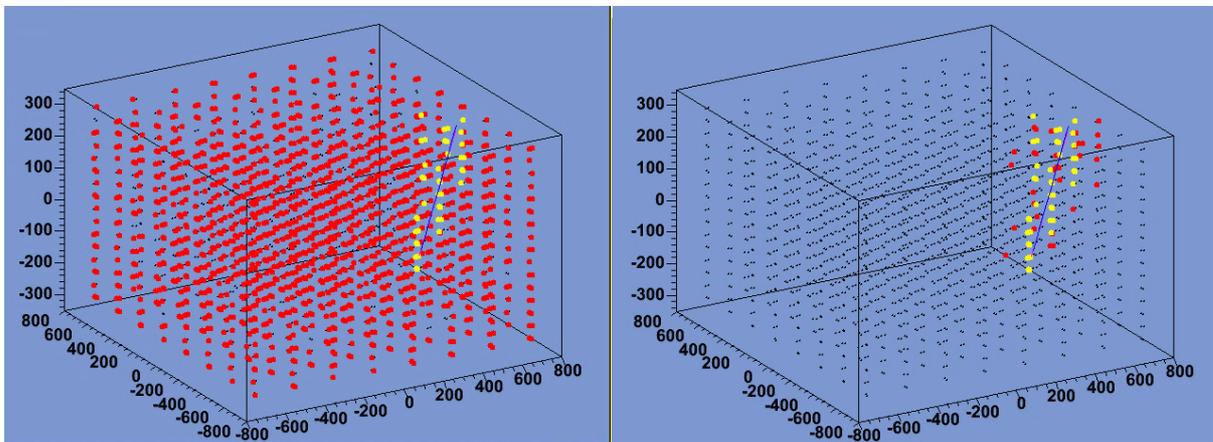


Figura 45 - Evento con e senza fondo. In rosso gli hit scartati dal filtro, in giallo gli hit usati per la ricostruzione geometrica della traccia.

3.4 Algoritmi di minimizzazione e pre-fit

La funzione di errore è piuttosto complessa per essere minimizzata in forma analitica; si utilizzano metodi numerici che hanno ormai raggiunto un'elevata maturità. A parte la strategia cosiddetta della "forza bruta" che cerca il minimo di una data funzione in modo esaustivo, calcolando la funzione per tutte le possibili configurazioni dei parametri, esistono metodi di minimizzazione basati sul gradiente, con o senza l'aiuto dell'espressione delle derivate prime, e metodi puramente geometrici come il metodo del Simplex. Presso il CERN è stata sviluppata la libreria MINUIT (parte delle CERNLIB), che consente di attuare varie strategie di minimizzazione, come quella del gradiente [55].

Per avere un buon comportamento della routine di minimizzazione e per contenere i tempi di elaborazione occorre però far partire la procedura di minimizzazione da un valore

iniziale dei parametri di traccia che non sia del tutto casuale. I valori iniziali possono essere stimati con una procedura di pre-fit, ad esempio ottenuto tramite un fit lineare sulle posizioni degli hit di un evento, imponendo la condizione di passaggio della traccia per il baricentro delle posizioni dei PMT che hanno dato hit, oppure imponendo il passaggio per i due cluster più lontani [56]. In questo modo non vengono considerati i tempi di arrivo e si ha una buona stima solo per quegli eventi con molti hit. Comunque l'accuratezza angolare è sufficiente per dare un valore iniziale alla procedura di fit, (cfr. Figura 46).

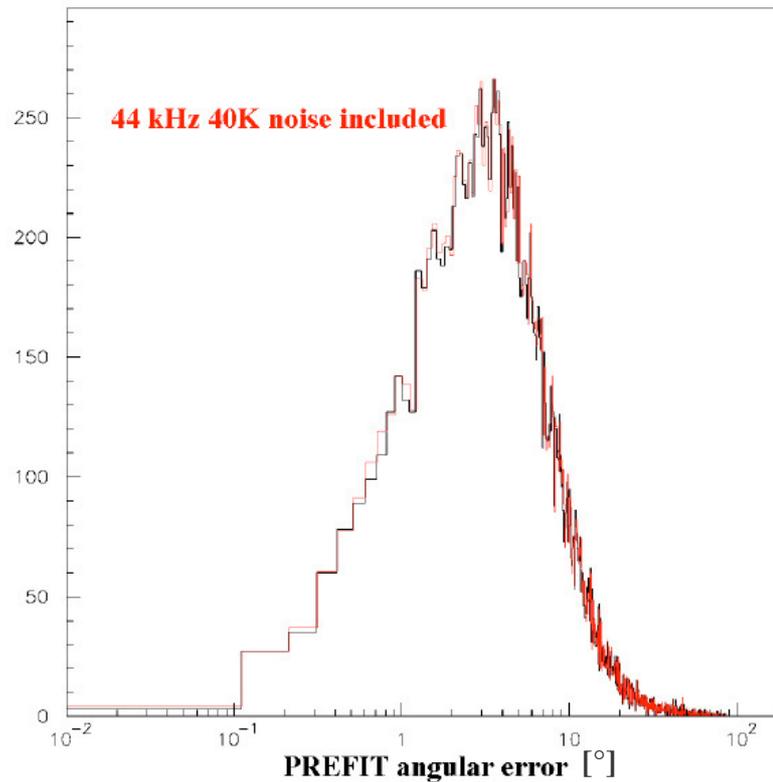


Figura 46 – Distribuzione della differenza ($\theta_{vero}^{spazio} - \theta_{Pr\text{fit}}^{spazio}$) [56].

3.5 Programma di Ricostruzione

La collaborazione NEMO ha sviluppato anche il programma di analisi e ricostruzione degli eventi. Legge in ingresso il file di uscita di OPNEMO, se richiesto aggiunge il fondo dovuto al ^{40}K , applica le logiche di trigger e i filtri descritti nel §3.3, e ricostruisce gli eventi (cfr. §3.2), generando in uscita un file di eventi ricostruiti e tutti i grafici di interesse. In Figura 47 è riportato il diagramma di flusso del programma di analisi.

La serie di routine è suddivisa in 4 fasi, indicate attraverso un parametro globale:

- ✓ inizializzazione
- ✓ elaborazione
- ✓ operazioni sugli istogrammi
- ✓ memorizzazione dei risultati

Le routine principali sono:

- ✓ ADD40K
- ✓ READEV
- ✓ ANALISI
- ✓ TRKFIT

Dopo aver letto i file di ingresso, se si vuole aggiungere il fondo viene chiamata la routine add40K, che genera un numero di hit random con frequenza nominale selezionabile a priori e li integra nel file di eventi.

La routine Analisi implementa il trigger di carica e di coincidenza, il filtro di causalità e il filtro a cluster, alzando un'opportuna flag sulla struttura dati degli hit. Sugli hit che passano viene poi fatto il prefit che andrà ad inizializzare la funzione di minimizzazione.

La routine trkfit implementa la ricostruzione, utilizzando la libreria del CERN "MINUIT" [55]. La routine FCNTRK, in particolare implementa la funzione (8), chiamando la routine tpmcal che esegue il calcolo (6).

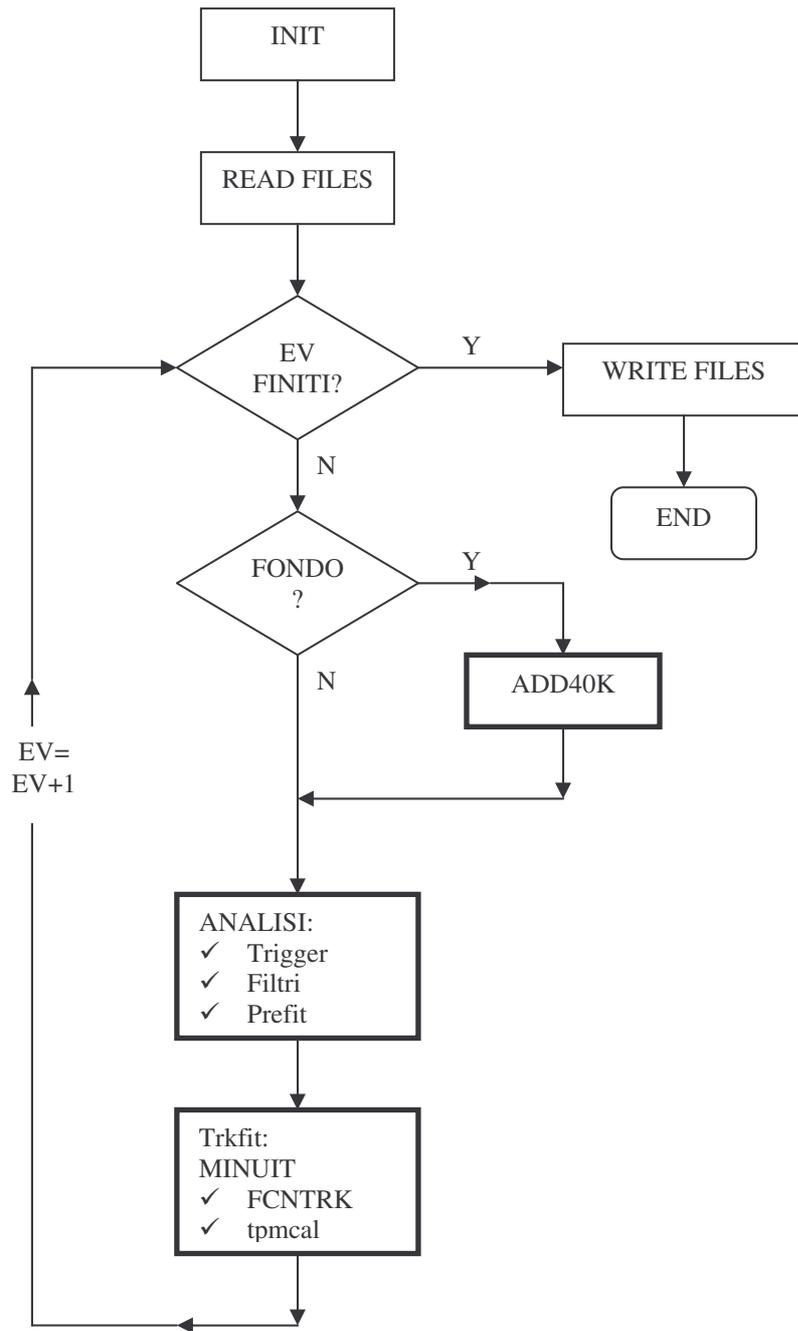


Figura 47 – Diagramma di flusso del programma di ricostruzione.

3.6 Risultati della ricostruzione OPNEMO

Mediante il programma OPNEMO è stato generato un file di 1000 eventi di energia compresa tra 1 e 1000 TeV, e questo è stato elaborato tramite il programma di analisi. Sarebbe meglio avere una statistica superiore, soprattutto per valutare le prestazioni della ricostruzione in funzione degli angoli della traccia; nei bin degli istogrammi in questo caso ci sarebbero troppo pochi eventi per avere una buona stima delle prestazioni: con 10 bin si hanno 100 eventi in media per bin, con una incertezza relativa nel bin pari al 10%.

Gli eventi generati hanno $-1 < \cos(\theta) < 1$ e $0 < \phi < 2\pi$. In Figura 48 sono riportate le distribuzioni in $\cos(\theta)$ e ϕ degli eventi generati. In Figura 49 sono riportate le coordinate y_0 e z_0 in funzione di x_0 . In Figura 50 è riportata l'energia degli eventi.

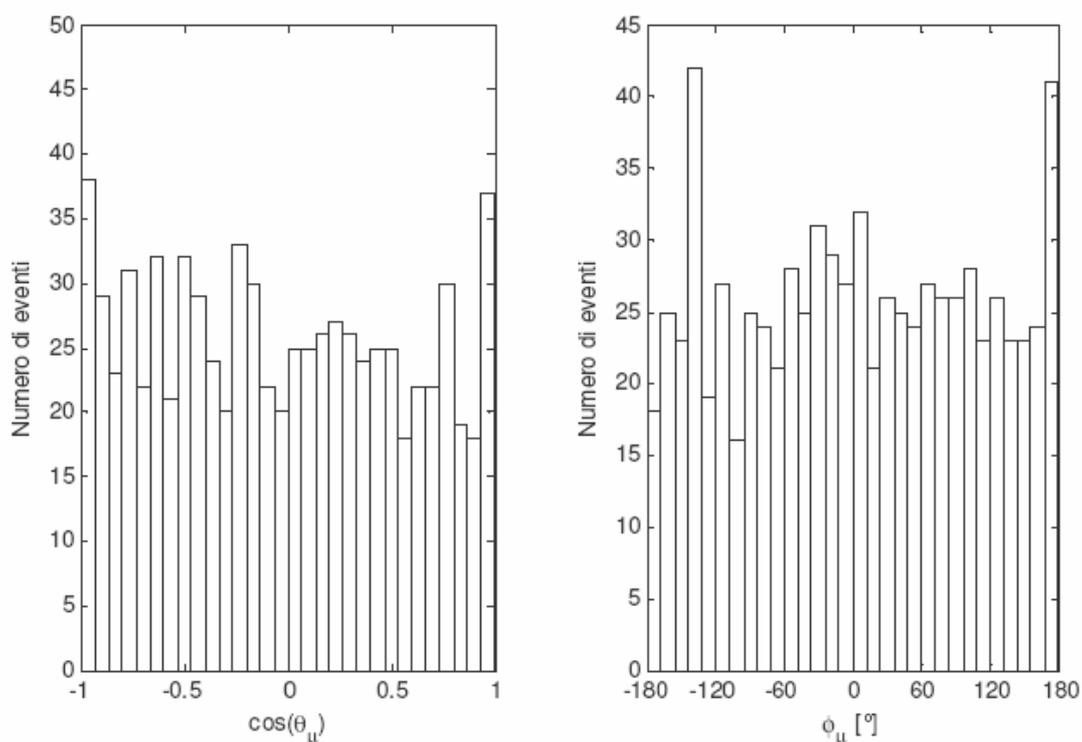


Figura 48 - Istogrammi di $\cos(\theta)$ e ϕ degli eventi simulati.

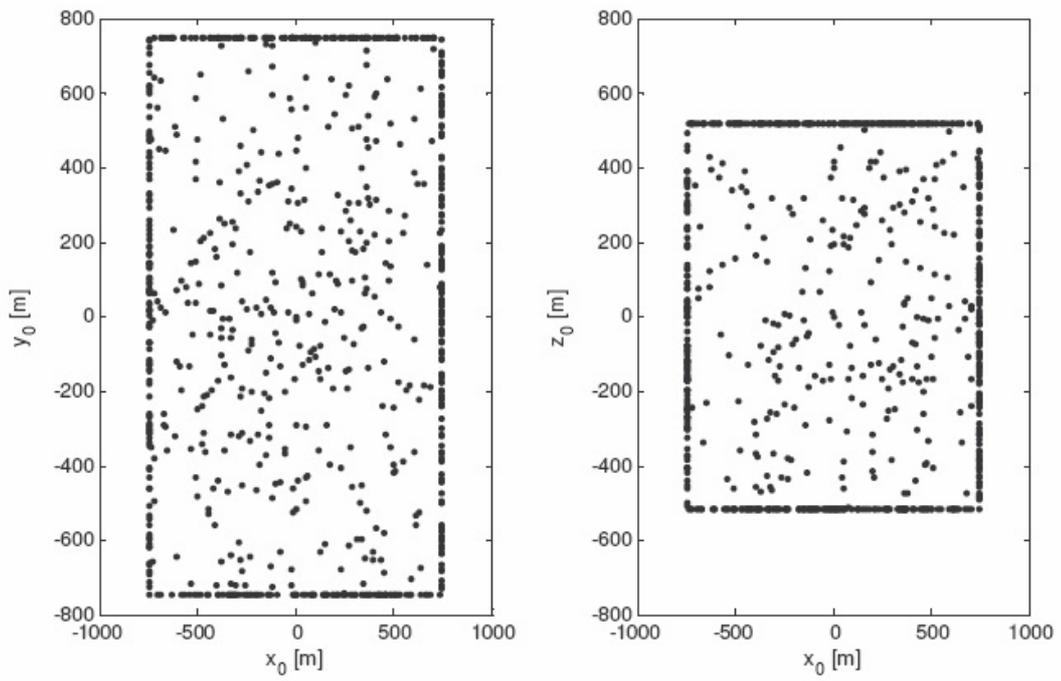


Figura 49 - Coordinate y_0 vs x_0 e z_0 vs x_0 degli eventi simulati.

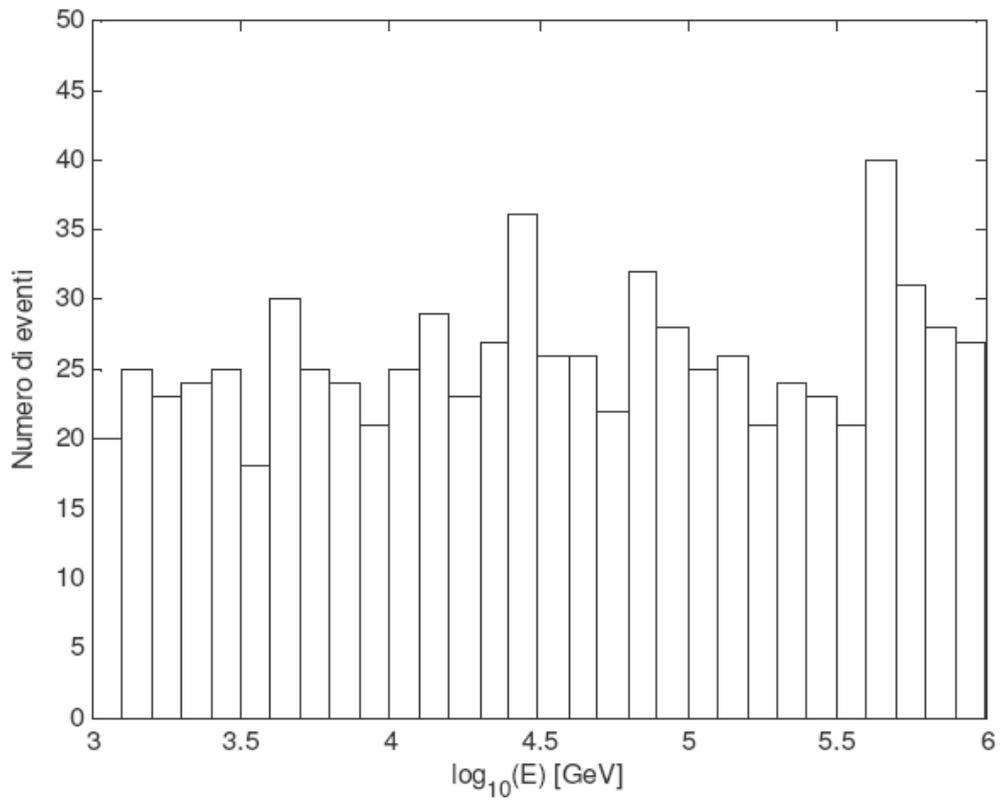


Figura 50 - Istogramma dell'energia degli eventi simulati.

Risultati globali degli errori angolari.

In Figura 51 sono riportati gli istogrammi delle differenze angolari tra gli angoli θ e ϕ "veri" della traccia e gli angoli ricostruiti mentre in Figura 52 l'istogramma dell'angolo nello spazio tra la traccia vera e quella ricostruita, calcolato come $\arccos(\hat{d}_{VERA} \cdot \hat{d}_{RIC})$

Il numero di eventi ricostruibili, cioè quelli con almeno 5 hit, in questo caso è di 775 su 1000.

La deviazione standard non è la grandezza statistica più adatta a descrivere l'accuratezza angolare di ricostruzione, perché la distribuzione delle differenze angolari ha delle lunghe code che danno un grande contributo all'errore quadratico medio. È utile allora anche la mediana dell'errore angolare nello spazio. Nel caso della ricostruzione OPNEMO la mediana globale dell'errore angolare è $0.5^{\circ} \begin{smallmatrix} +0.1 \\ -0.07 \end{smallmatrix}$ (vedi Figura 52).

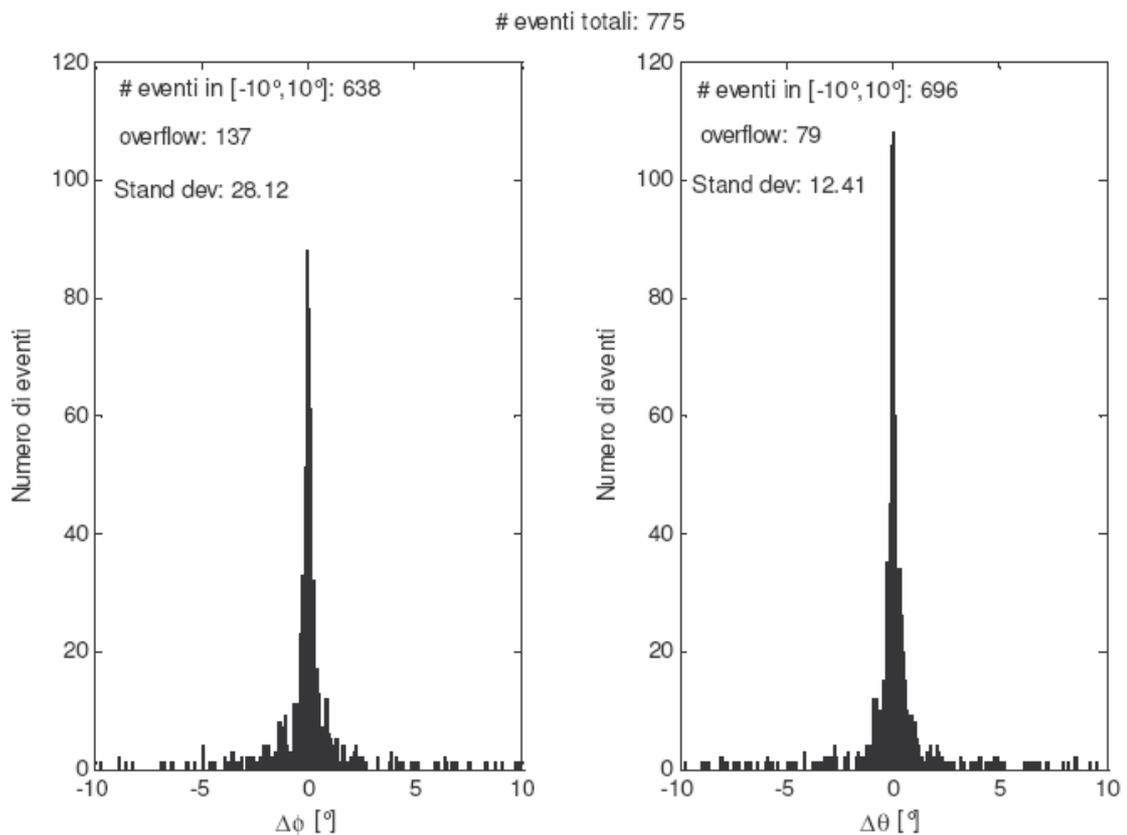


Figura 51 – Iistogrammi delle differenze tra gli angoli ϕ e θ della traccia "vera" e gli angoli ricostruiti con il programma di analisi.

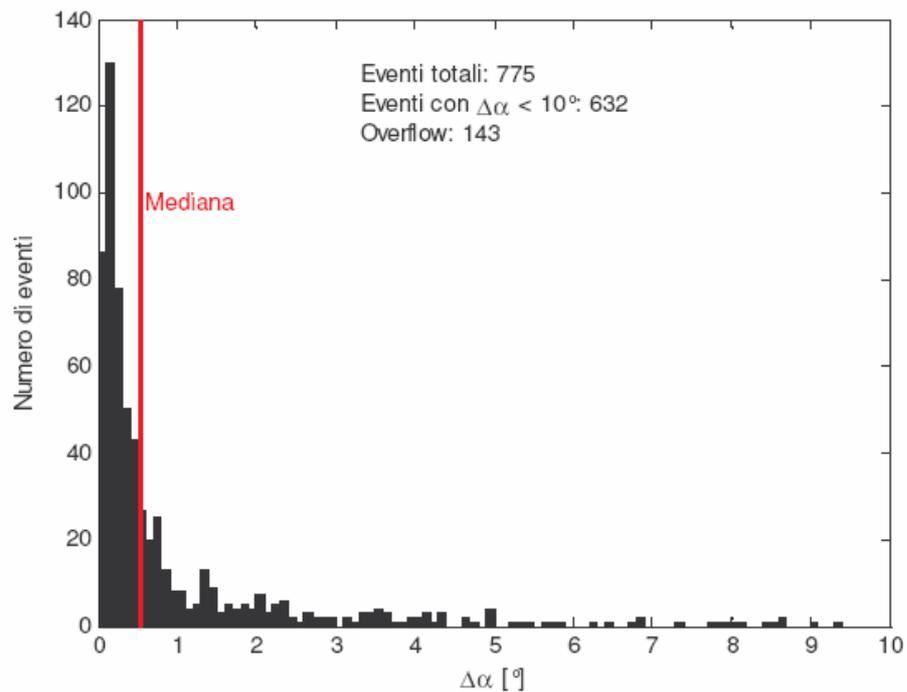


Figura 52 – Istogramma dell'angolo tra la traccia "vera" e quella ricostruita. Il 50% delle tracce (la mediana della distribuzione) è ricostruito entro 0.5° dalla direzione della traccia vera.

Errore angolare in funzione delle grandezze geometriche

In Figura 53-54 sono riportati gli errori angolari in funzione delle coordinate del vertice e in Figura 55-56 sono in funzione di θ e ϕ .

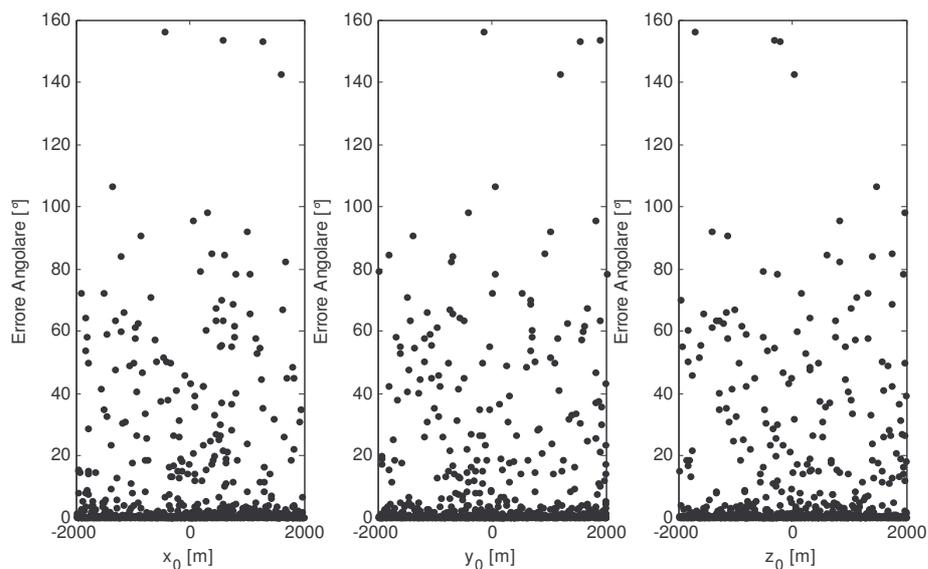


Figura 53 – Scatter plot dell'errore angolare in funzione delle coordinate del vertice.

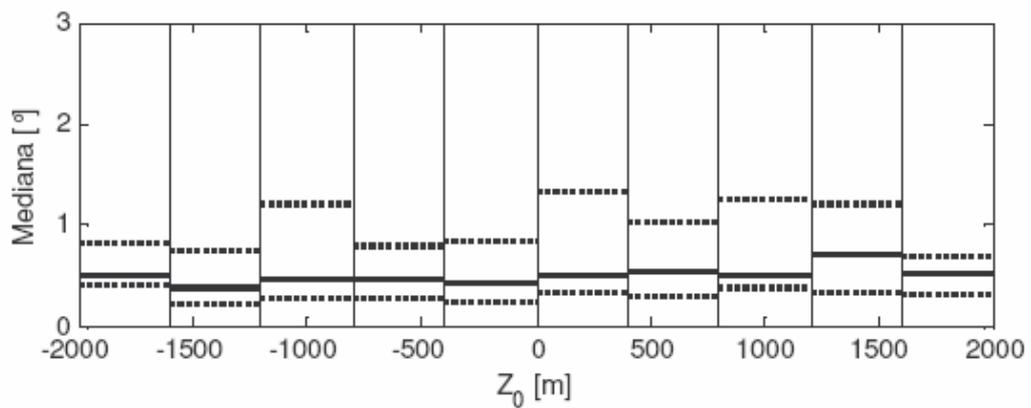
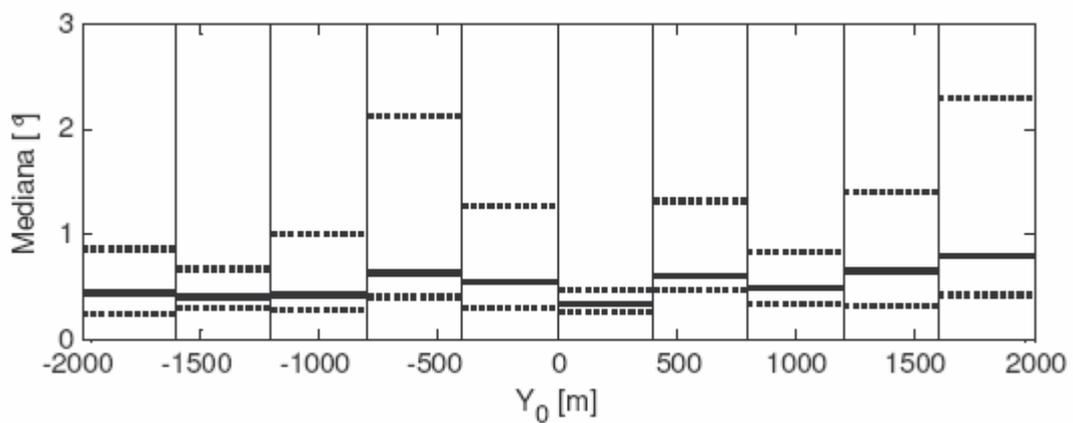
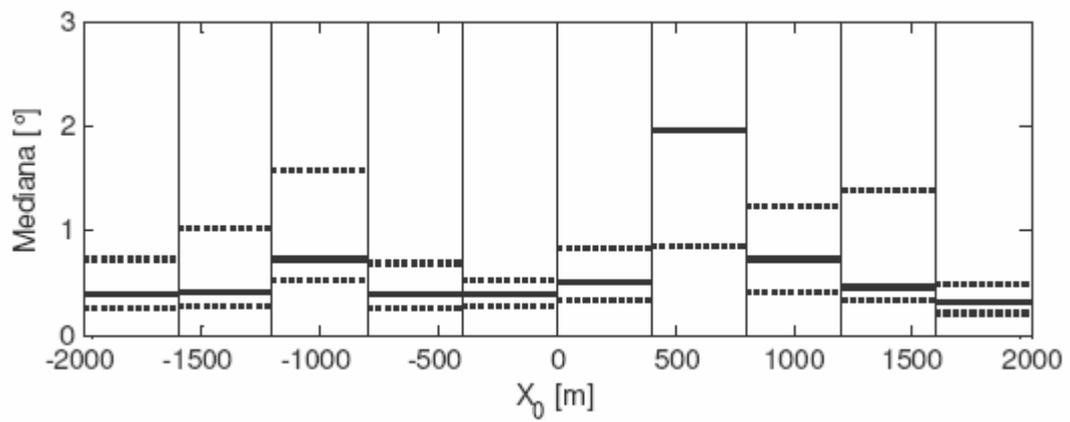


Figura 54 – Istogrammi della mediana dell'errore angolare rispetto alle coordinate x_0, y_0, z_0 del vertice, in alto la deviazione standard, in basso la mediana, della distribuzione dell'errore angolare (in linea tratteggiata l'errore a 1σ).

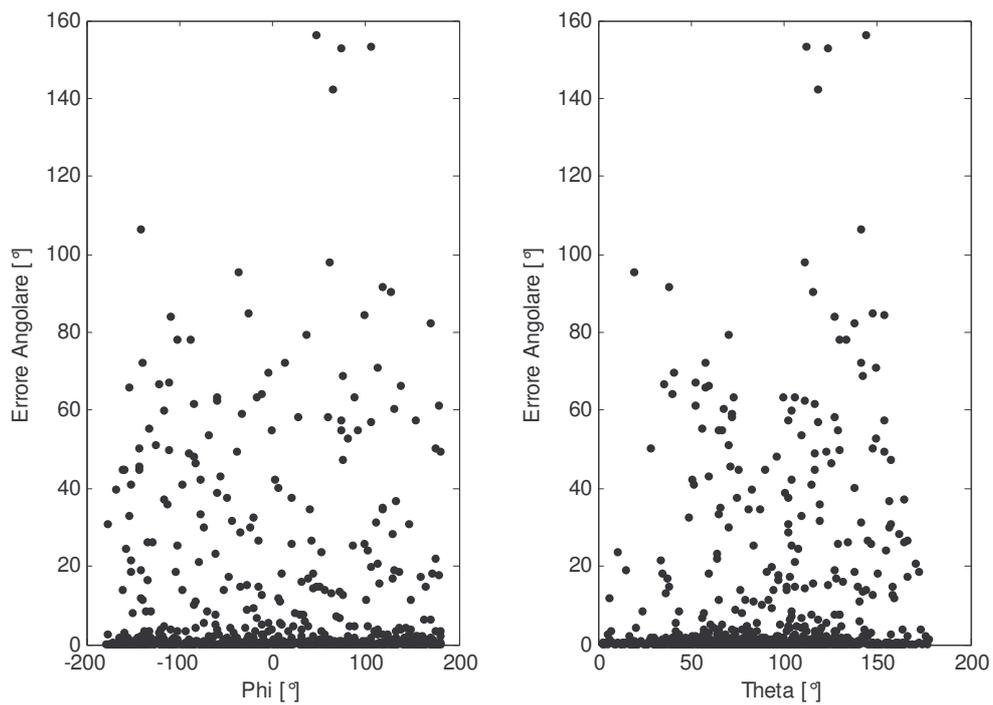


Figura 55 – Scatter plot dell'errore angolare rispetto a θ e ϕ .

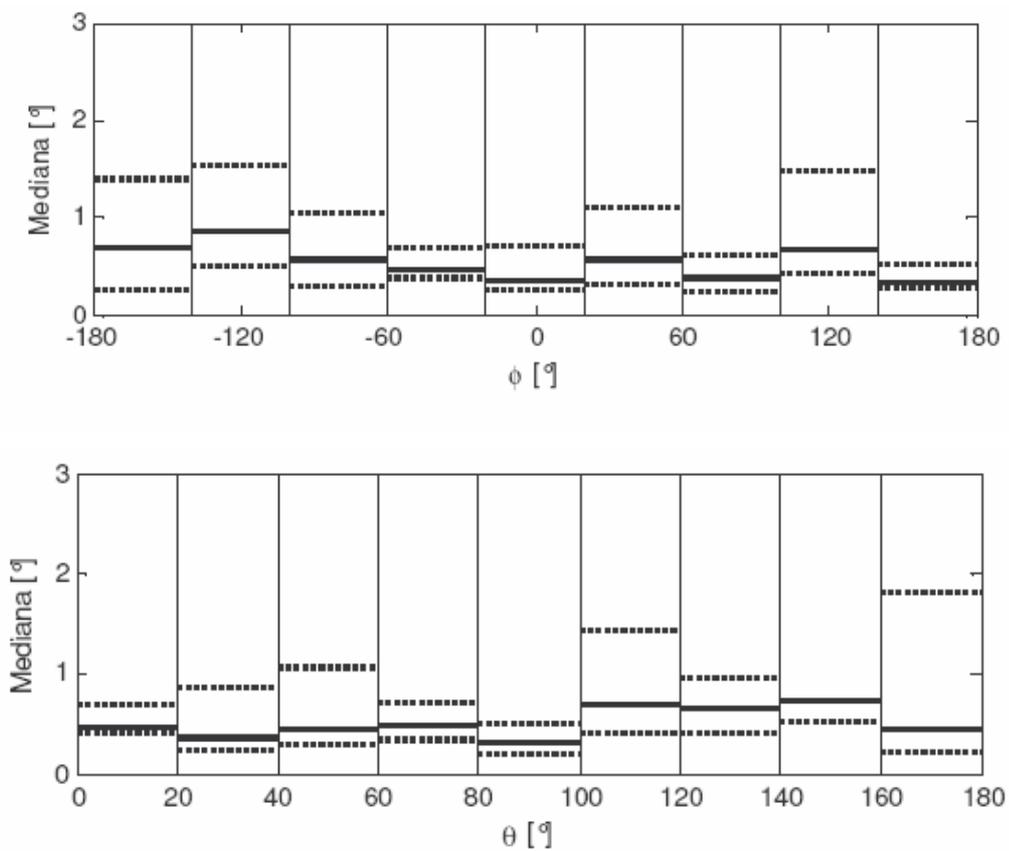


Figura 56 - Istogrammi della mediana dell'errore angolare rispetto a ϕ e θ .

Le variazioni dell'errore angolare rispetto alle coordinate del vertice e rispetto alla direzione vera della traccia, osservate nelle figure precedenti sono da attribuirsi alle fluttuazioni statistiche; dagli scatter plot di Figura 53 e Figura 55 non si notano dipendenze evidenti.

Errore angolare in funzione dell'energia

Di seguito sono tracciati gli andamenti dell'errore angolare in funzione dell'energia, in Figura 57 lo scatter plot e in Figura 58 le grandezze statistiche in forma di istogramma.

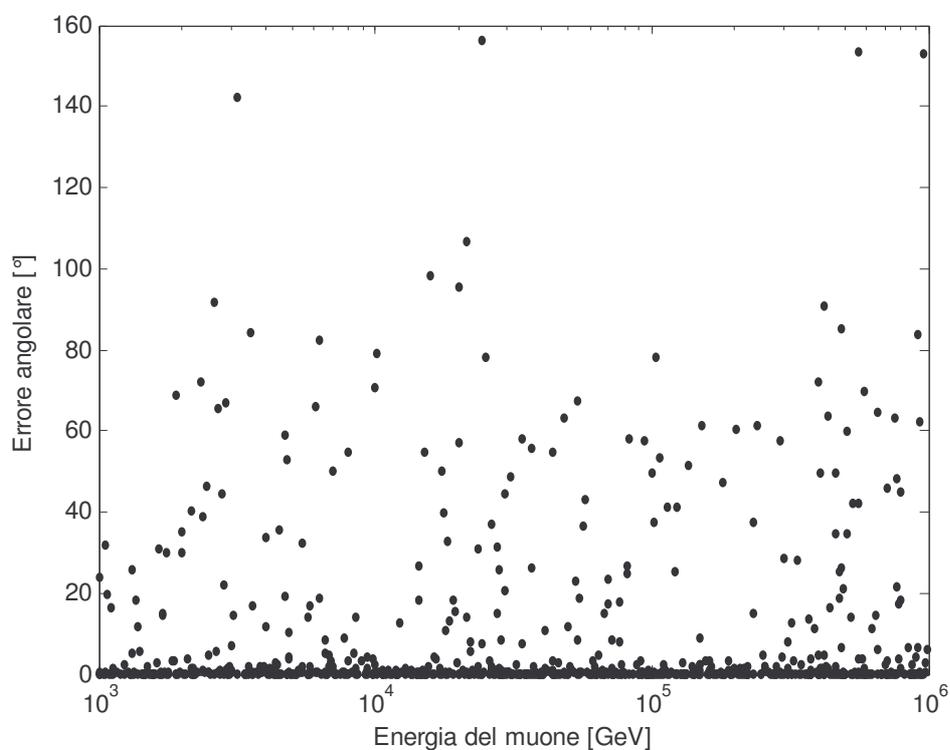


Figura 57 – Scatter plot dell'errore angolare in funzione dell'energia del muone.

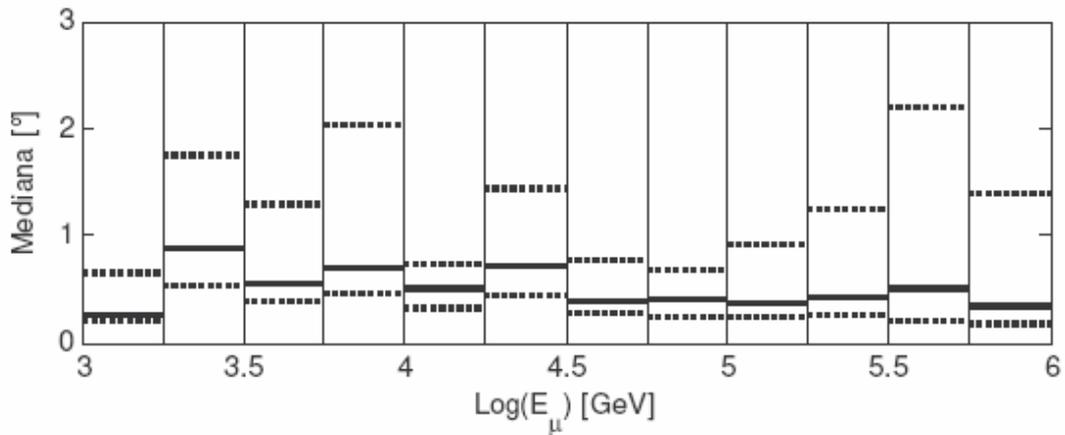


Figura 58 – Istogrammi dell’errore angolare (standard deviation e mediana) in funzione dell’energia del muone.

Si può vedere che la dipendenza dell’errore dall’energia è debole, anche se sopra i 100 TeV la mediana è sotto il $\frac{1}{2}$ grado. Ben più significativa è la dipendenza dell’errore dal numero di hit nell’evento, come si può vedere dalla Figura 59 a sinistra. A destra di Figura 59 è riportato il numero di hit in funzione dell’energia; si vede che il numero di hit cresce con l’energia, ma cresce anche la probabilità di avere sciami elettromagnetici durante la traccia, e quindi aumenta l’errore del tempo di arrivo dei relativi fotoni (cfr. Figura 60)

Questa dipendenza può essere sfruttata per operare una selezione a priori sugli eventi, scartando quelli con un numero di hit inferiore ad una data soglia, accettando quindi una perdita di efficienza, ma guadagnando in accuratezza di ricostruzione.

Selezione degli eventi

In Figura 61 è riportata la mediana dell’errore istogrammata in funzione del numero di hit, e in rosso è riportata la mediana cumulativa di tutti gli eventi con N_{hit} maggiori del valore corrente. Tenendo presente la distribuzione del numero di hit (Figura 62) si può scegliere, con un buon compromesso (perdendo circa 200 eventi), la soglia pari a 50 hit.

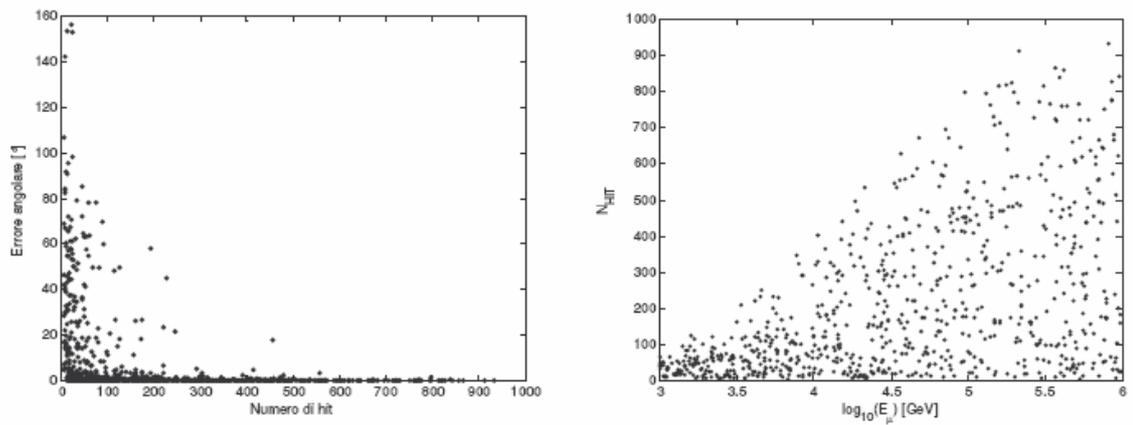


Figura 59 – Scatter plot dell'errore angolare in funzione del numero di hit nell'evento.

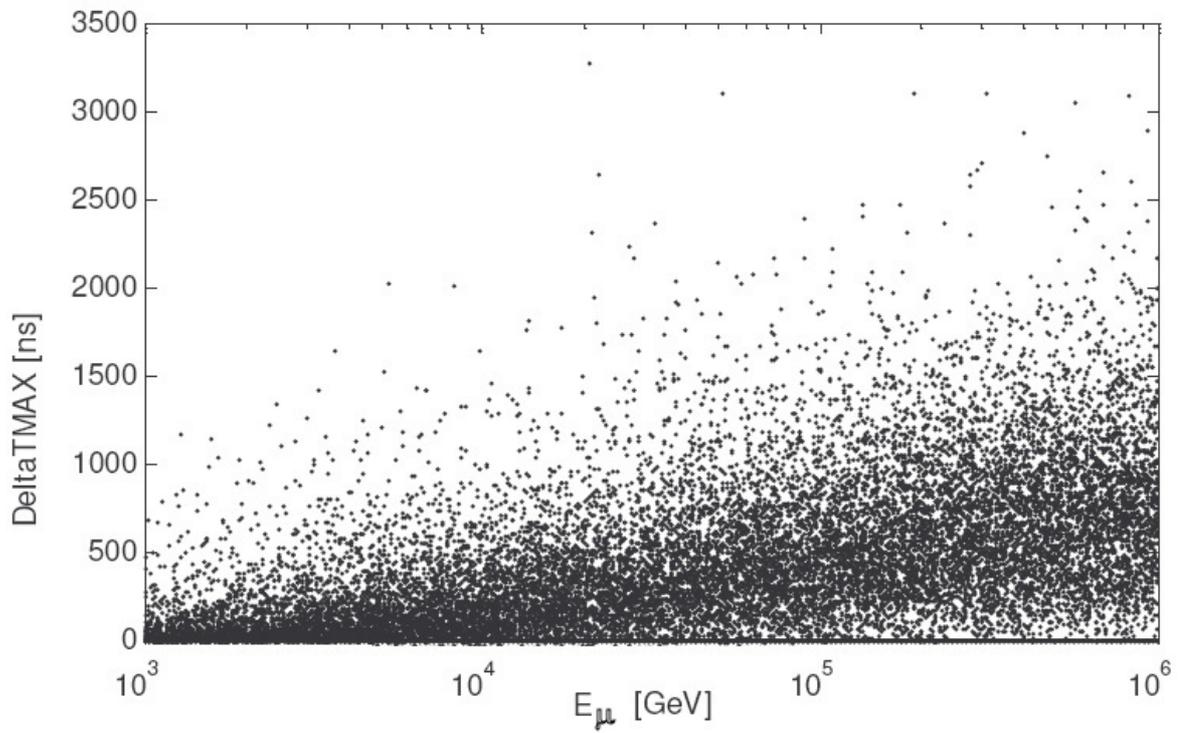


Figura 60 - Differenza massima dei tempi di arrivo dei fotoni (rispetto al tempo di arrivo teorico) in funzione dell'energia.

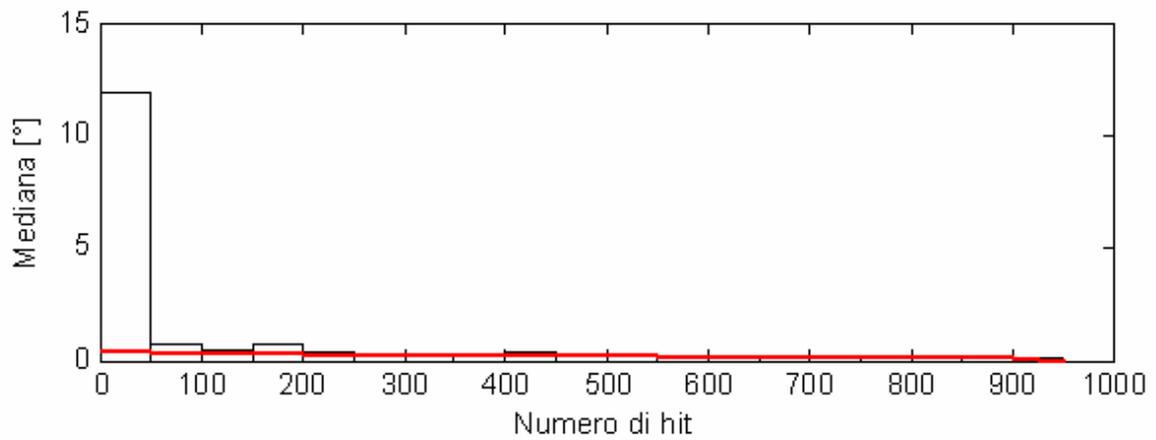


Figura 61 – Istogrammi dell’errore angolare (deviazione standard e mediana) in funzione del numero di hit; in rosso è riportata la mediana globale degli eventi con N_{hit} maggiore del valore corrente.

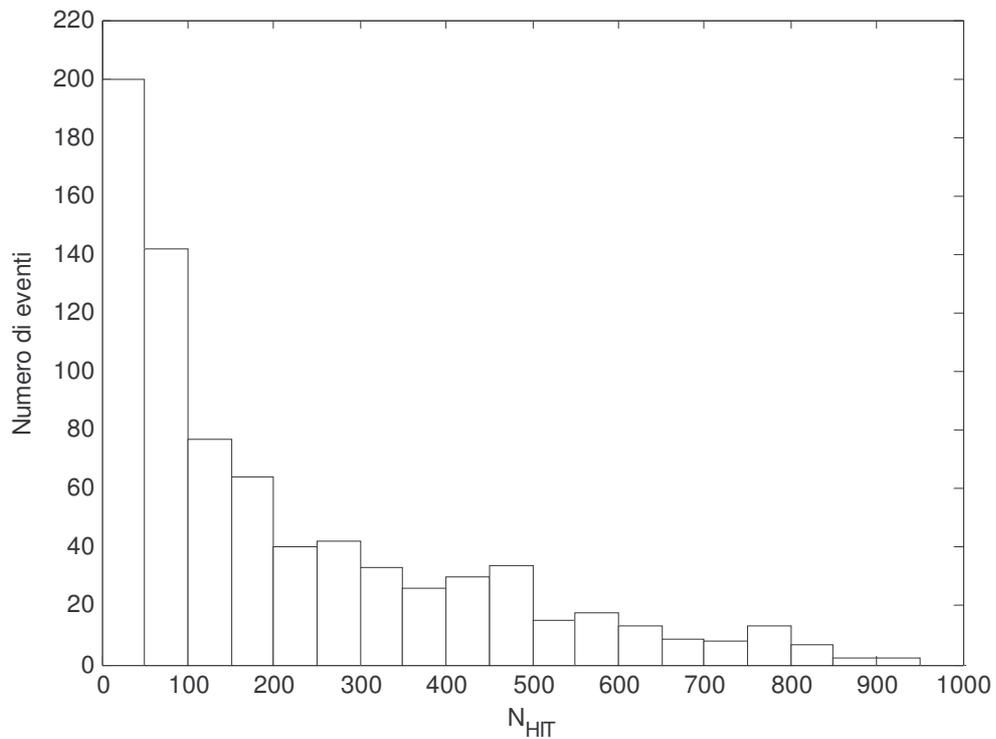


Figura 62 – Istogramma del numero di hit (N_{HIT}) nell’evento.

Un’altra selezione può essere operata sulla base del χ^2 ridotto, cioè dell’errore quadratico medio dei tempi di hit, normalizzato all’incertezza nominale di misura dei tempi nei PMT, pari a 2.5 ns. Il logaritmo del χ^2 ottenuto alla fine del processo di minimizzazione per ciascun evento è riportato in Figura 63

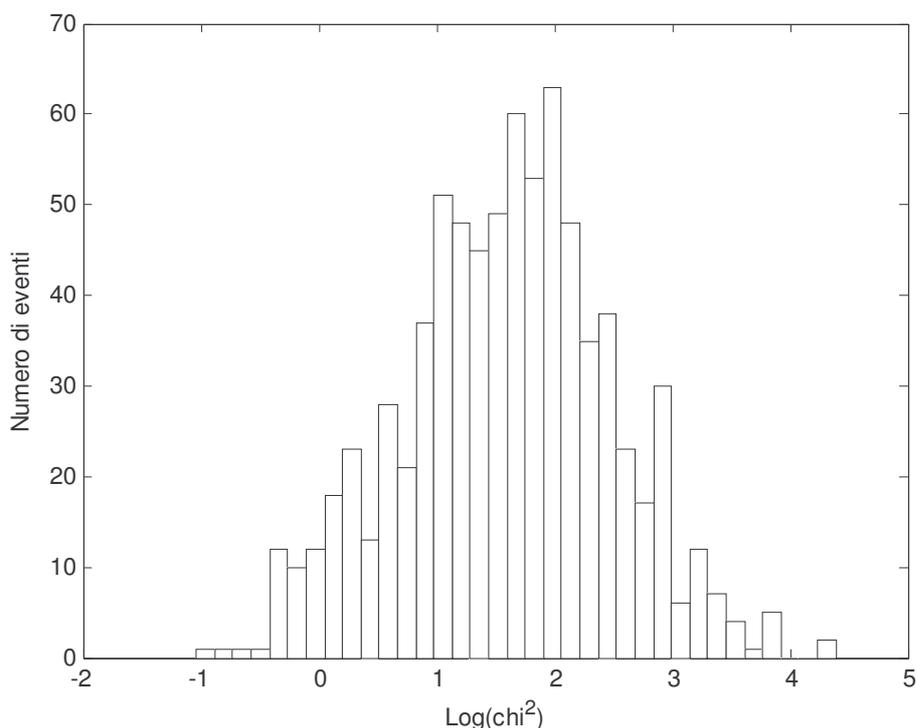


Figura 63 – Istogramma del logaritmo del valore del χ^2 ridotto ottenuto alla fine del processo di ricostruzione della traccia.

La definizione del χ^2 data precedentemente non tiene conto degli errori sui tempi degli hit dovuti ad altri fenomeni diversi dal puro spreading del tempo di transito nei fototubi; invece altre fonti di errore sono dovute alle code nelle distribuzioni dei tempi date dallo scattering dei fotoni, peraltro non ancora modellizzato in OPNEMO, e dagli sciame elettromagnetici. Questi ritardi (cfr.§3.2) sono eliminati completamente se il PMT corrispondente ha rivelato un fotone Čerenkov, perché questo è il primo ad arrivare, e basta quindi considerare come tempo di hit il tempo di arrivo del primo fotone; ma se il fotone Čerenkov non viene rivelato (per attenuazione in acqua, o per efficienza di rivelazione) rimangono i fotoni ritardatari a dare un contributo di errore ben maggiore del TTS. Molti di questi hit errati sono filtrati dagli algoritmi descritti in §3.3, ma alcuni passano il filtro ed il loro risultato si vede in termini di errore di ricostruzione e di valori del χ^2 ridotto molto maggiore di 1. Ad esempio, l'evento di Figura 43 dopo il filtro ha degli errori temporali di 50 ns, che con un TTS di 2.5 ns danno un contributo al χ^2 ridotto molto maggiore di 1.

In Figura 64, il χ^2 è messo in relazione con l'errore angolare, a sinistra con uno scatter plot, a destra è istogrammata la mediana dell'errore angolare in funzione del χ^2 .

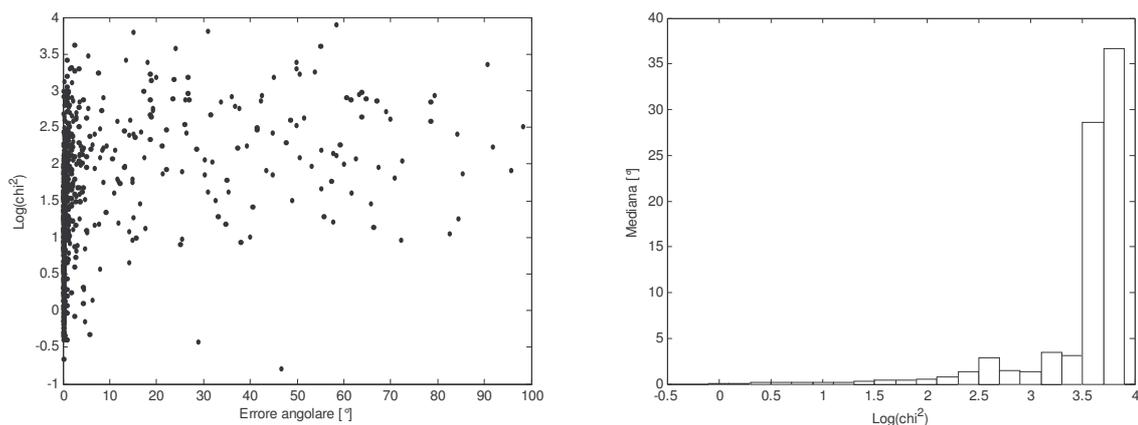


Figura 64 – Errore angolare in funzione del χ^2 , a sinistra scatter plot, a destra istogramma della mediana.

Nonostante non vi sia una grande correlazione tra il χ^2 e l'errore angolare visto che molti eventi ben ricostruiti mostrano un χ^2 molto più alto del normale, tuttavia gli eventi ricostruiti con i valori più elevati del χ^2 (sopra a 1000) hanno una mediana dell'errore molto più alta, quindi si può operare un'ulteriore selezione, filtrando gli eventi ricostruiti con $\chi^2 > 1000$ (Figura 65).

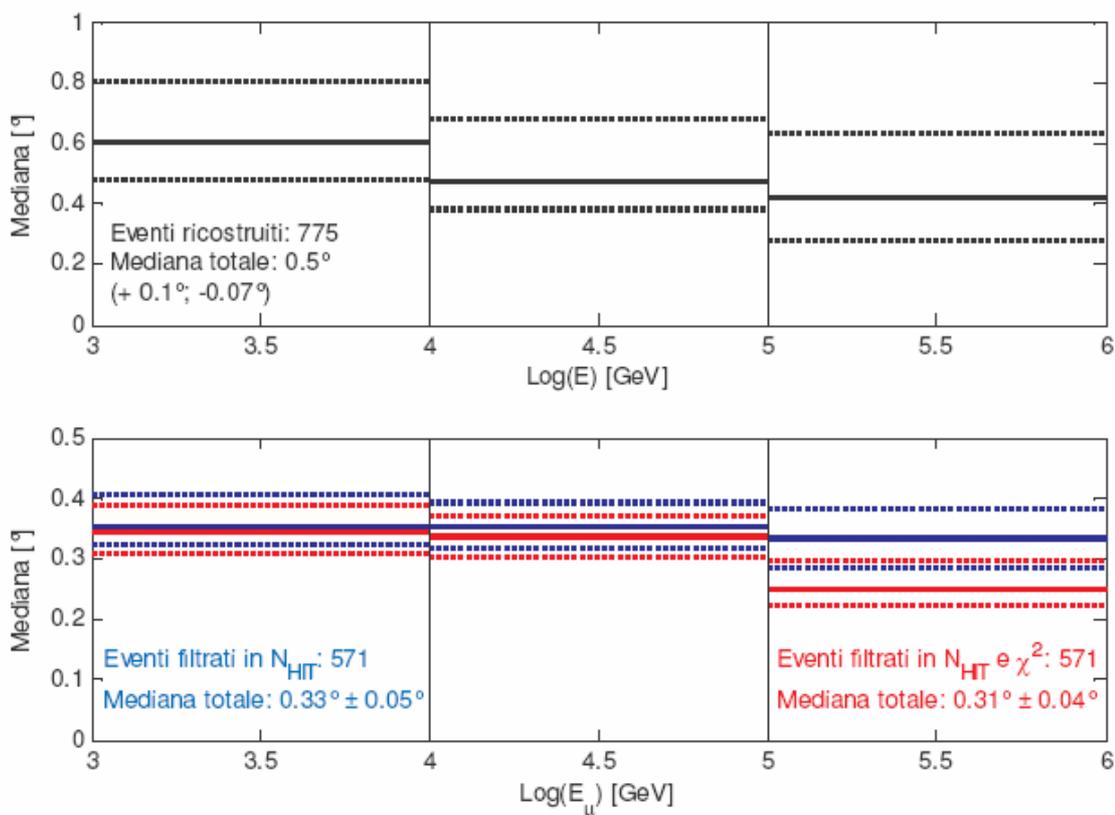


Figura 65 – Confronto tra la mediana degli eventi ricostruiti e la mediana degli eventi dopo la selezione.

La mediana totale dell'errore per tutti gli eventi ricostruiti ($N_{hit}>5$) è quindi $0.5^{+0.1}_{-0.07}$.

Con la selezione sul numero di hit la mediana dell'errore totale è $0.33^\circ \pm 0.05$; con la selezione sul numero di hit e sul χ^2 la mediana dell'errore è $0.31^\circ \pm 0.04$.

Efficienza di ricostruzione

L'area geometrica di generazione degli eventi, come già spiegato in §2.6 e in §3.1, è l'area che contiene l'apparato più 5 lunghezze di attenuazione dei fotoni in acqua, ed è pari a $A_G = 3.57 \text{ km}^2$. L'area efficace prima della selezione è pari a $A_{Eff} = A_G \cdot 775/1000 = 2.77 \text{ km}^2$; dopo la selezione l'area efficace risulta $A_{Eff} = A_G \cdot 550/1000 = 1.96 \text{ km}^2$, con una mediana dell'errore pari a $0.31^\circ \pm 0.04^\circ$.

In Figura 66 è riportata l'area efficace (rapportata all'area geometrica) al variare dell'energia.

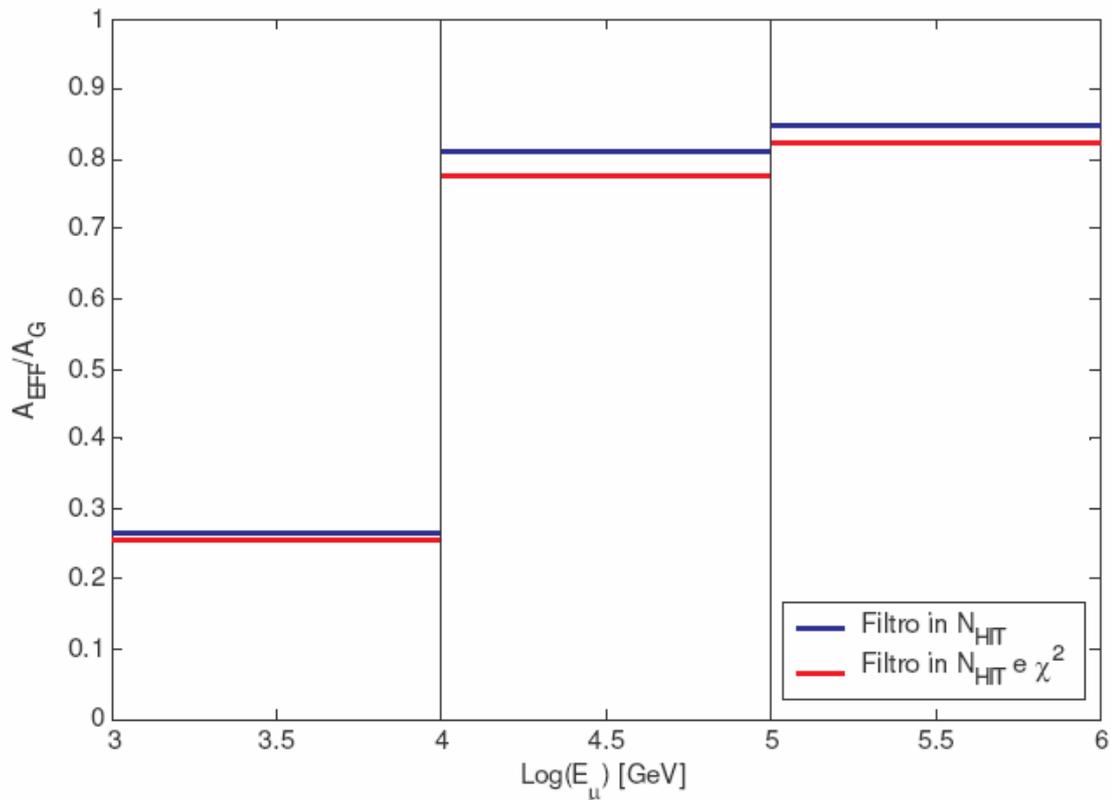


Figura 66 - Area efficace normalizzata all'area geometrica in funzione dell'energia.

Cap. 4 Ricostruzione delle tracce con reti neurali

4.1 Introduzione alle Reti Neurali

4.1.1 Neuroni

Le reti neurali si ispirano al funzionamento dei neuroni biologici, le cellule alla base dell'elaborazione nel cervello animale.

Perfino il cervello degli insetti, molto più piccolo e più semplice di quello umano, contiene milioni di queste cellule elementari, mentre quello umano ne contiene centinaia di miliardi. Il cervello più piccolo in natura è probabilmente quello del verme nematode che ha circa 300 neuroni.

La caratteristica fondamentale del cervello è l'alto grado di connessione tra i neuroni: ognuno ha migliaia di "ingressi" e migliaia di "uscite"; l'immagazzinamento dell'informazione e la capacità elaborativa sono legate alle intensità di queste connessioni, cioè come si vedrà a dei pesi che modulano i segnali in ingresso al neurone. In Figura 67 è riportata una rappresentazione schematica di un neurone: gli ingressi sono in corrispondenza dell'albero dendritico, mentre le ramificazioni dell'assone sono le uscite; i neuroni sono quindi caratterizzati dall'unidirezionalità del flusso di informazione.

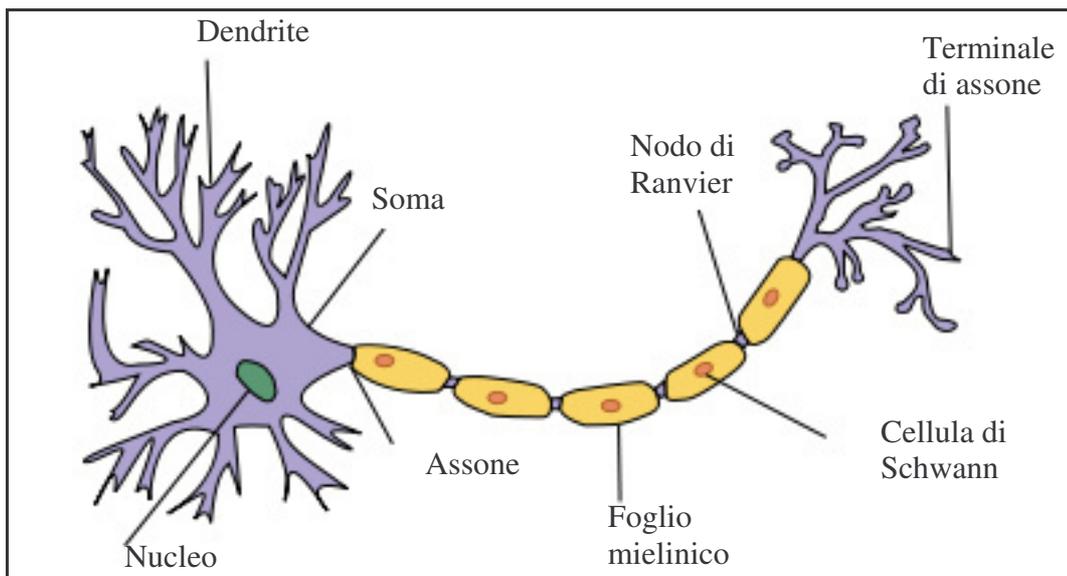


Figura 67 - Schema di un neurone.

Le uscite di un neurone vanno ad alimentare ingressi di altri neuroni, attraverso interfacce dette sinapsi, costituendo quindi una vera e propria rete. Ci sono neuroni afferenti, cioè legati a organi di senso, neuroni intermedi, che hanno per ingresso le uscite di altri neuroni e alimentano altri neuroni ancora, e ci sono neuroni efferenti, che trasmettono segnali alle cellule di organi motori.

La trasmissione dell'informazione all'interno di un neurone avviene tramite impulsi elettrici, condotti attraverso la guaina mielinica. La trasmissione sinaptica invece è di tipo chimico; se il segnale elettrico che giunge all'assone supera il potenziale di azione vengono emesse in corrispondenza delle sinapsi molecole di neurotrasmettitori che attivano i dendriti dei neuroni successivi.

4.1.2 Elaborazione neurale

Un semplice modello di neurone, che ricalca quanto descritto precedentemente, è riportato in Figura 68

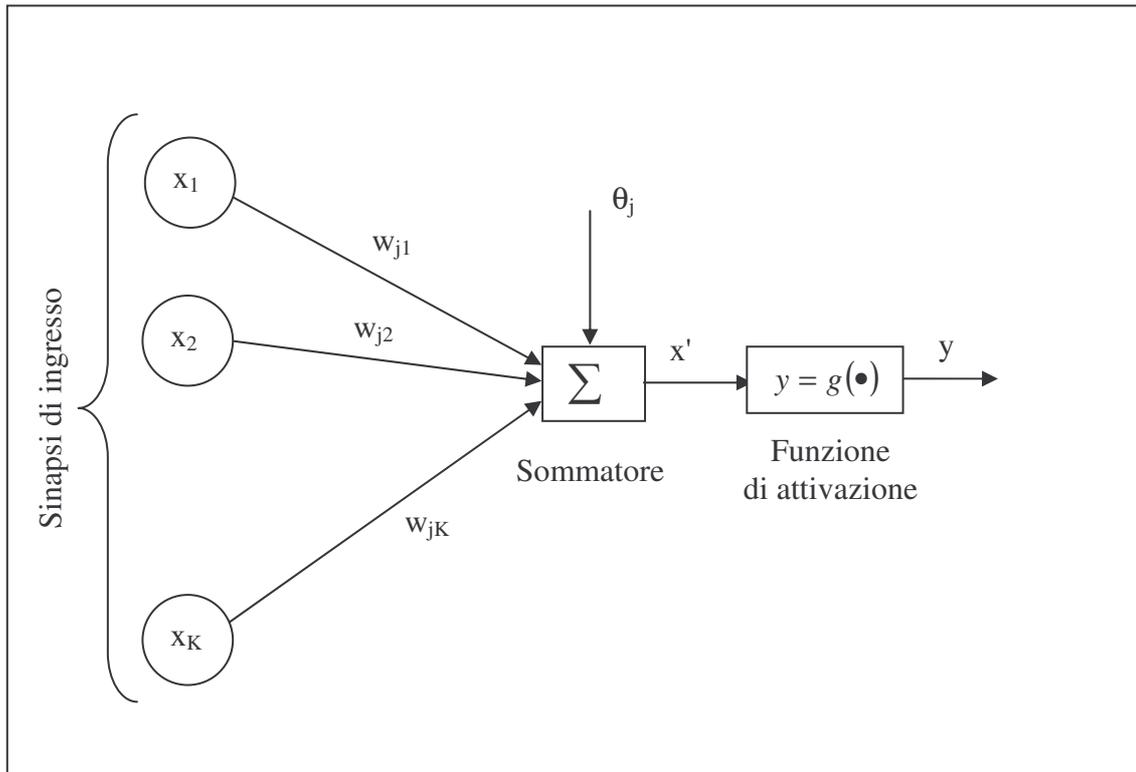


Figura 68 - Modello di neurone.

Questo è composto da:

- un insieme di sinapsi di ingresso x_1, x_2, \dots, x_K , ognuna caratterizzata da un'ampiezza o forza $w_{j1}, w_{j2}, \dots, w_{jK}$;
- un sommatore di tutte le sinapsi di ingresso pesate con la propria ampiezza, con un eventuale offset (soglia)
- una funzione di attivazione (di solito non-lineare)

Nella figura l'uscita di un neurone k (x_k) è collegata all'ingresso del neurone rappresentato, e l'ampiezza della sinapsi è indicata dal peso w_{jk} .

Matematicamente l'uscita y_j del neurone j può essere espressa come:

$$y_j = g\left(\sum_{k=1}^K w_{jk} \cdot x_k - \theta_j\right)$$

In pratica ogni singolo neurone di una rete funziona come un interruttore che si accende se la somma pesata di tutti gli ingressi supera la soglia di attivazione [38].

4.1.3 Funzioni di attivazione

Le funzioni di attivazione utilizzate nelle implementazioni di reti neurali sono di vari tipi, ma tutte hanno due caratteristiche in comune:

- saturano ad un valore minimo e un valore massimo
- sono non-decrescenti

Alcune funzioni tipicamente usate in letteratura sono:

1. Funzione soglia, definita come

$$g(x) = \begin{cases} 1 & \text{se } x \geq 0 \\ 0 & \text{se } x < 0 \end{cases}$$

che definisce quindi un interruttore ideale (Figura 69 a)

2. Funzione lineare a tratti, dove le due saturazioni sono raccordate da un tratto rettilineo (Figura 69 b):

$$g(x) = \begin{cases} 1 & \text{se } x \geq 1/2 \\ x + 1/2 & \text{se } -1/2 < x < 1/2 \\ 0 & \text{se } x \leq -1/2 \end{cases}$$

3. Funzione sigmoide, una delle funzioni di attivazione più usate, anche perché la sua caratteristica di essere derivabile ovunque facilita la convergenza degli algoritmi di apprendimento

$$g(x) = \frac{1}{1 + e^{-ax}}$$

L'effetto del parametro a è evidente nella Figura 69 (c).

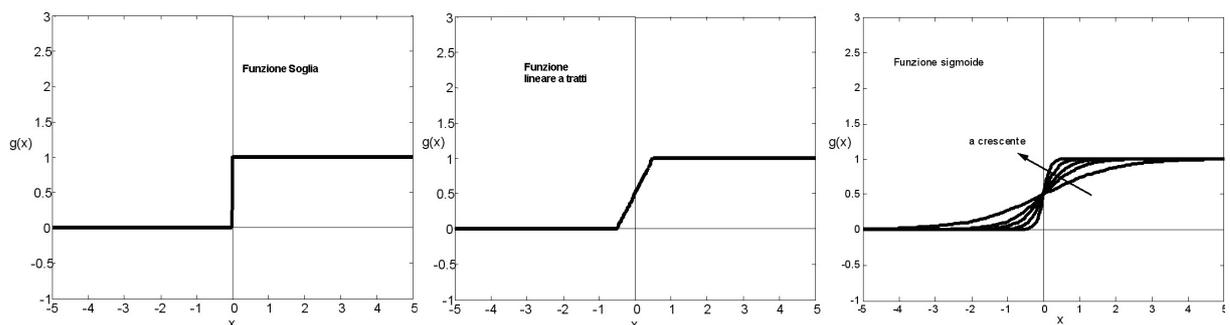


Figura 69 - Funzioni di attivazione

Questi tre esempi di funzione di attivazione sono definiti tra 0 ed 1; per alcuni scopi è comodo definire la funzione tra -1 e 1, e questo può essere fatto mantenendone la forma.

4.1.4 Architetture di reti neurali

Single layer Feed-forward

L'architettura più semplice per una rete neurale è composta da uno strato di N neuroni ad ognuno dei quali sono collegati K ingressi attraverso i pesi w_{nk} come rappresentato in Figura 70

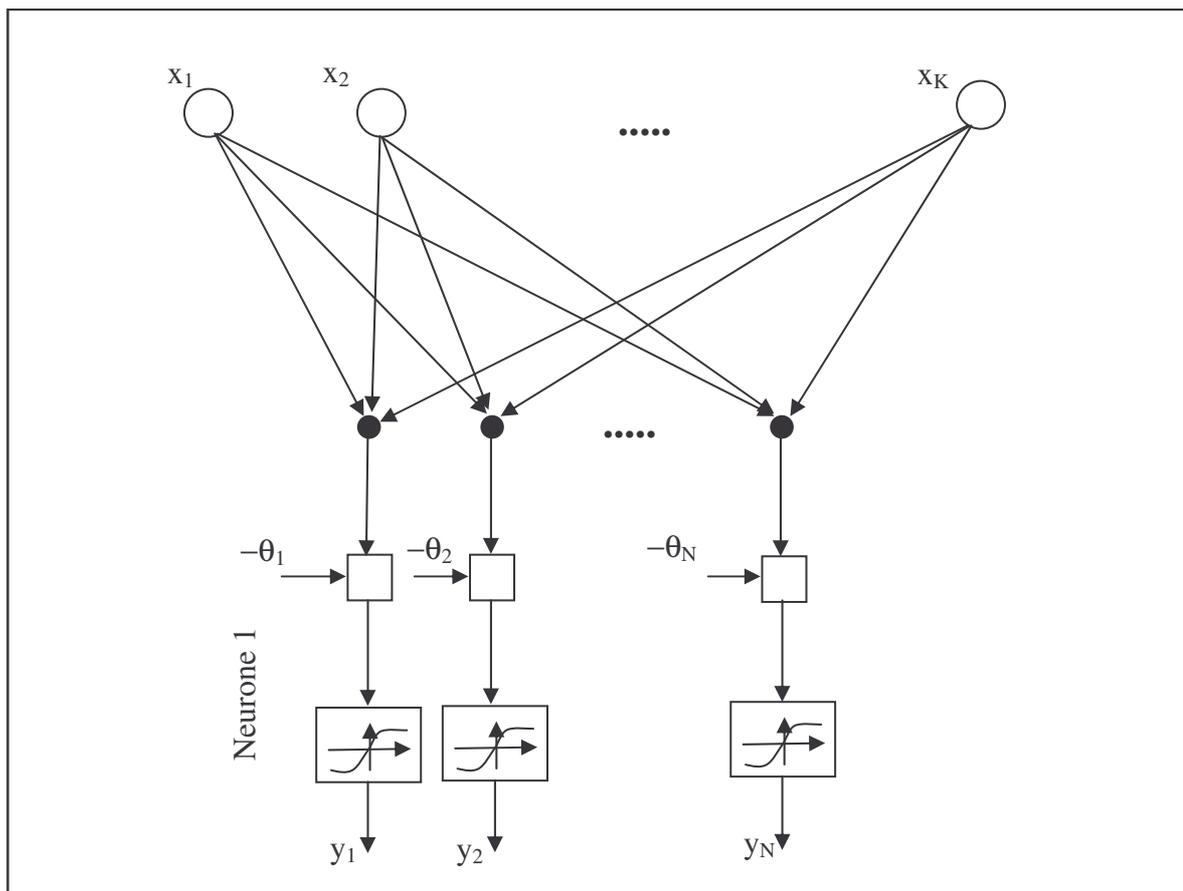


Figura 70 - Schema di rete single layer feed-forward.

Questa struttura è molto semplice e permette di rappresentare delle mappe cosiddette a classi linearmente separabili, ma fallisce su problemi non-linearmente separabili, di cui l'esempio più classico è la funzione XOR [40].

Multilayer feed-forward

Un'architettura più generale, in grado di superare le limitazioni viste per il Single Layer è naturalmente una struttura a più livelli, in cui oltre allo strato di ingresso e allo strato

di uscita, vi siano uno o più livelli intermedi (nascosti o hidden). Questo tipo di architettura è detta multilayer perceptron.

In effetti è stato dimostrato che un'architettura con due livelli, (uno hidden) e funzioni di attivazione non-lineari è in grado di implementare qualsiasi funzione di interesse pratico [41], e comunque aggiungere livelli interni (a parità di gradi di libertà) non aggiunge capacità di elaborazione alla rete.

Quindi uno schema sufficientemente generale di rete neurale Multi Layer Feed Forward NN è quello riportato in Figura 71

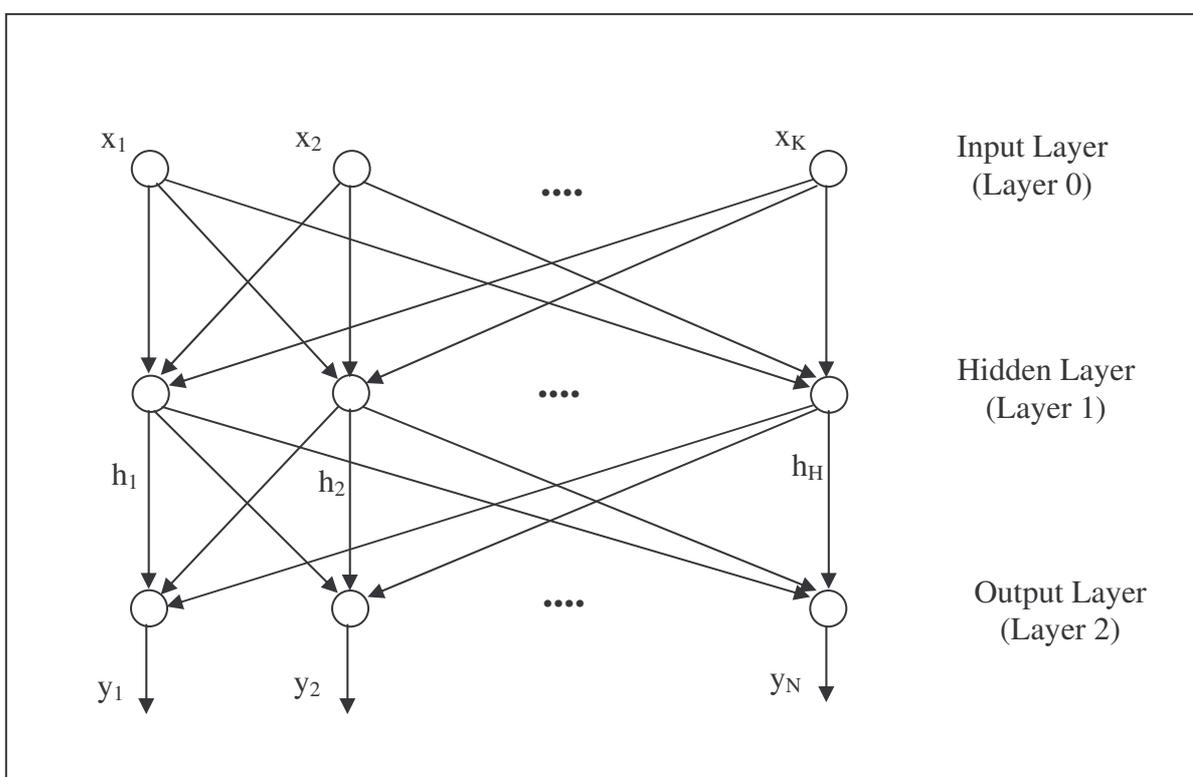


Figura 71 - Schema di rete Multi layer feed forward.

Qui è rappresentata una rete con K ingressi, H neuroni hidden e N uscite.

Questa architettura di rete neurale, sebbene fosse nota già dagli anni '60 e fosse già stato dimostrato che avrebbe superato i limiti della struttura a singolo layer, è stata utilizzata estensivamente solo dopo la scoperta di un efficace metodo per l'apprendimento: il cosiddetto back-propagation algorithm, degli anni '80 [42], di cui si parlerà in seguito.

Altri tipi di architettura

Esistono altri tipi di architetture per le reti neurali, ognuna con delle particolarità che la rendono più o meno utile a seconda dell'applicazione. In effetti la struttura classica Multi Layer Feed Forward è la più versatile, essendo buona per ogni tipo di applicazione.

Le Recurrent NN sono reti in cui è permesso il ritorno delle uscite verso gli ingressi; c'è praticamente la stessa differenza che corre tra i filtri FIR (Finite Impulse Response) e quelli IIR (Infinite Impulse Response).

Le Radial Basis Function NN sono più utili per applicazioni in cui si richiede l'approssimazione di una funzione più che la classificazione di un pattern; sono essenzialmente uguali alla multi layer, ma le funzioni di attivazione sono a simmetria radiale (per esempio una gaussiana), quindi i neuroni intervengono localmente (intorno al picco della gaussiana).

Le Adaptive Resonance Theory NN sono delle reti neurali con struttura più complessa, che permettono una certa adattabilità a nuove tipologie di ingressi, e quindi l'applicabilità a segnali non stazionari [43].

4.1.5 Metodologie di apprendimento

Le reti neurali hanno bisogno di essere addestrate per fornire le risposte corrette. L'addestramento viene condotto scegliendo un set di dati di ingresso, di cui è noto il risultato desiderato dell'elaborazione. I dati di training vengono presentati ad una rete inizializzata con valori random dei parametri, e le uscite ottenute vengono confrontate con le uscite desiderate. I pesi della rete vengono quindi corretti iterativamente per minimizzare l'errore di uscita.

Esistono tipicamente tre metodologie di apprendimento per una rete neurale: apprendimento supervisionato, apprendimento non supervisionato e apprendimento ibrido.

- Apprendimento supervisionato (supervised training)

La caratteristica di questo tipo di apprendimento è la presenza di un insegnante esterno alla rete. L'insegnante fornisce alla rete un set di dati di training di cui sono note le risposte esatte. I pesi della rete sono ottimizzati in base all'errore che questa commette in successive e reiterate prove.

- Apprendimento non supervisionato (unsupervised training)

Nell'apprendimento non supervisionato, non conoscendo le risposte ai dati di training, si deve individuare una misura della bontà della rappresentazione che la rete deve imparare. Alla fine del training la rete sarà in grado di raggruppare dati omogenei rispetto al set di parametri utilizzati per definire la rappresentazione.

- Apprendimento ibrido

In alcuni casi è conveniente utilizzare entrambi i metodi descritti o in cascata, per diminuire il set di apprendimento supervisionato, o in parallelo, per migliorare le prestazioni dell'apprendimento stesso.

Tutte queste tecniche possono essere implementate tramite l'algoritmo di retro-propagazione dell'errore, (Back-propagation algorithm [42]). Dato un insieme di ingressi di training \mathbf{x} e le relative uscite desiderate \mathbf{d} , si può calcolare la funzione d'errore

$$E(\mathbf{w}) = \sum_{i=1}^N (d_i - y_i)^2$$

dove \mathbf{w} rappresenta la matrice dei pesi e y_i le uscite.

Ora, assumendo che la funzione d'errore sia differenziabile, si possono applicare le usuali tecniche di ottimizzazione, per esempio la tecnica del *gradient descent*. Si aggiornano i pesi di una quantità proporzionale al gradiente della funzione d'errore:

$$\Delta \mathbf{w} = -\eta \nabla E(\mathbf{w})$$

per poi procedere ad una successiva iterazione.

Oltre a MINUIT, anche in MatLAB esistono già implementate molte funzioni che applicano questa tecnica, con delle piccole varianti che ottimizzano alcuni aspetti [44].

Quando la dimensione del set di training è piccola (rispetto al numero di pesi totali da ottimizzare), la convergenza del *gradient descent* è molto veloce, ma in questo caso la rete non è in genere in grado di generalizzare (a meno di un criterio di riconoscimento molto semplice).

Una buona regola empirica per addestrare correttamente una rete è quella di utilizzare un campione di training di dimensione almeno 10 volte il numero totale di pesi indipendenti; questo implica che, una volta fissato il numero di ingressi e di uscite della rete, la scelta del numero di neuroni dello strato intermedio (H):

$$H = \frac{PT}{10(I + O)}$$

dove PT è la dimensione del set di training, I ed O sono il numero di ingressi e di uscite rispettivamente.

4.1.6 Applicazioni

Le reti neurali trovano applicazioni particolarmente appropriate nel campo di pattern recognition, ad esempio nel riconoscimento vocale, o di immagini, e in tutti i campi in cui sia necessaria una classificazione veloce di segnali complessi, in cui l'applicazione di algoritmi tradizionali sarebbe molto difficile.

4.2 Ricostruzione delle tracce di muoni con reti neurali

Si vuole costruire ed addestrare una rete neurale in grado di prendere in ingresso gli hit di un evento in un telescopio Čerenkov sottomarino (tempi e carica) e restituire in uscita la direzione della traccia del muone (ϕ , θ), come rappresentato in Figura 72. Ci si limita ai parametri (ϕ , θ) con l'intento di ottenere un prefit.

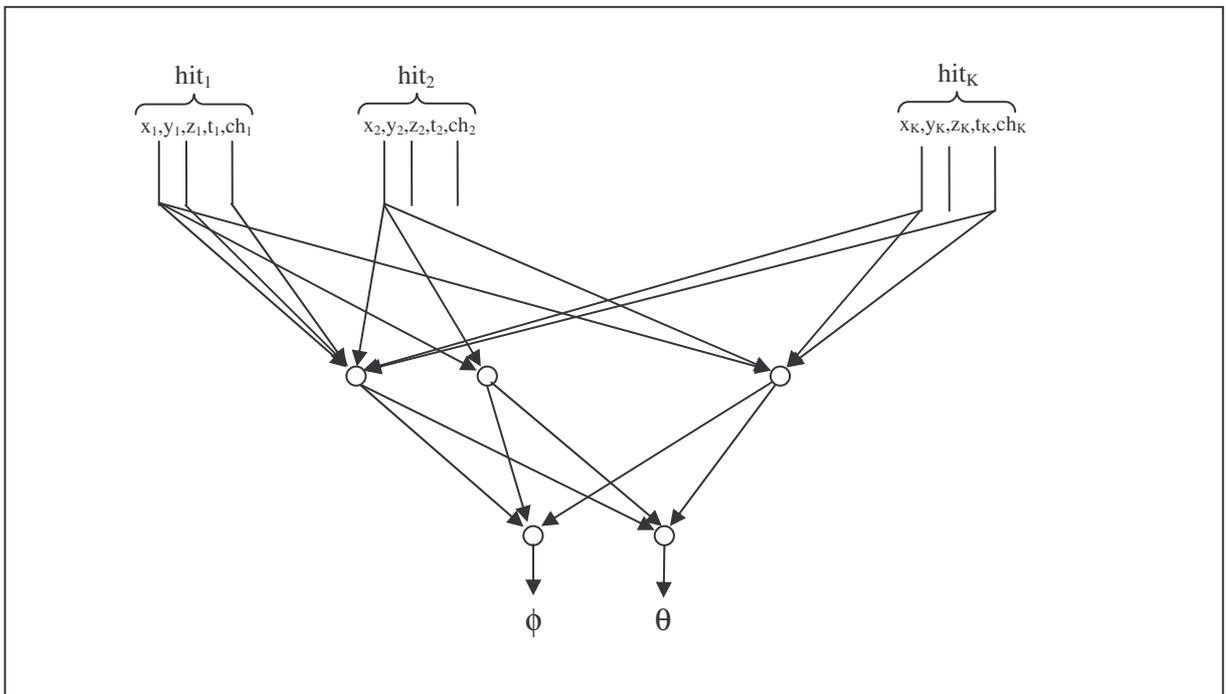


Figura 72 - Schema di principio della ricostruzione di una traccia con una rete neurale.

La scelta della topologia della rete e della modalità con cui fornire gli ingressi relativi agli eventi lascia una certa libertà. Bisogna indagare quali sono i vantaggi e gli svantaggi delle possibili scelte implementative.

Nel nostro caso un evento consiste di un certo numero di hit, da poche unità ad alcune centinaia, ogni hit è contraddistinto da una posizione nello spazio, da un tempo, e da una carica.

La rete più generale possibile probabilmente è quella che ha $2 N_{\text{PMT}}$ ingressi (N_{PMT} è il numero totale di PMT dell'apparato) a cui vanno i tempi e la carica di ogni relativo hit. Nei PMT che non hanno rivelato un segnale si può porre zero carica e un certo codice (p.es. -1) al tempo, posto che tutti gli hit abbiano tempo positivo.

In questa topologia non c'è bisogno di passare alla rete le coordinate degli hit, perché queste verrebbero apprese e memorizzate implicitamente nella rete al momento dell'addestramento; tuttavia il grande numero di detector nell'apparato NEMO ($N_{\text{PMT}} = 5832$) rende questa via poco percorribile, specialmente per la fase di addestramento.

Si possono allora fornire in ingresso alla rete solo gli hit, questa volta con la posizione e il tempo. In questo caso, poiché il numero di ingressi sarebbe variabile da evento a evento, si possono scegliere un certo numero di hit (quelli con carica maggiore) oppure si possono dare alla rete più sequenze di hit e poi mediare le uscite di sequenze appartenenti allo stesso evento come rappresentato in Figura 73. Qui le sequenze di hit sono lunghe K , e per ogni hit vengono passati alla rete le coordinate del PMT e il tempo di hit, quindi la rete avrà $4 \cdot K$ ingressi. Durante l'addestramento, per tutte le sequenze di hit dello stesso evento, il target (cioè le uscite desiderate) saranno i valori veri di ϕ e θ della traccia.

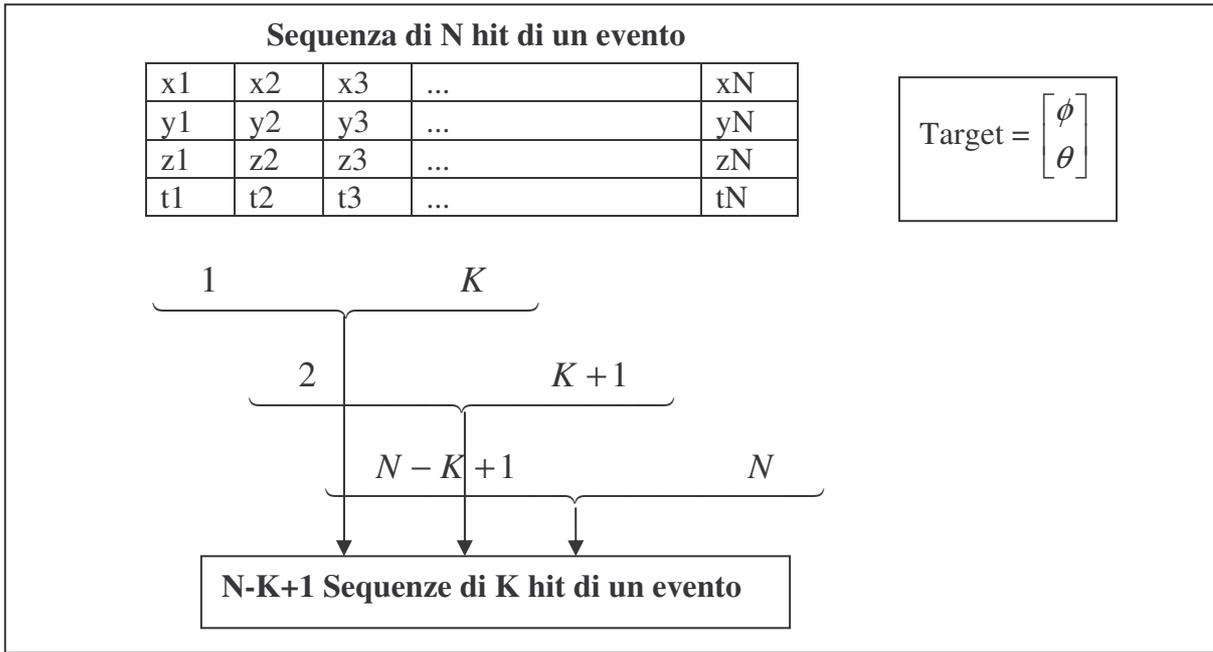


Figura 73 - Sequenze di hit in ingresso alla rete neurale.

Un altro approccio che è stato preso in considerazione è quello di dare in ingresso alla rete gli ingressi differenziali dx/dt , dy/dt , dz/dt , cioè le differenze di coordinate tra coppie di hit divise per le differenze di tempo d'arrivo. Anche in questo caso si possono selezionare gli hit con carica maggiore.

Il procedimento utilizzato per la simulazione (cfr. §4.3) è stato quello di generare un file di 1000 eventi con il programma OPNEMO, una parte del quale (da 100 a 200 eventi) è stata utilizzata per ricavare i dati di training, e la parte restante per simulare le ricostruzioni operative.

Le ricostruzioni sono state fatte al variare dei seguenti parametri:

- numero di hit di ogni sequenza in ingresso
- numero dei neuroni intermedi

e sono state generate reti per la stima di theta e phi e solo per la stima di theta (riconoscimento up/down). Il parametro di bontà utilizzato per i confronti è la mediana della distribuzione dell'angolo tra la traccia vera e la traccia ricostruita.

4.3 Risultati della ricostruzione con reti neurali

La ricostruzione mediante reti neurali, descritta in §4.2, è stata realizzata con il programma Matlab. Inizialmente è stata implementata la ricostruzione con in ingresso le

coordinate e i tempi di arrivo di gruppi di M hit: (x_i, y_i, z_i, t_i) . Valori possibili per M sono tra 10 e 30; infatti con pochi hit la ricostruzione è sempre più critica, mentre al crescere del numero di ingressi la memoria e il tempo di elaborazione crescono eccessivamente: cresce molto anche la dimensione del set di training necessario per un buon addestramento, che come detto (§4.1.5) è dell'ordine di $10 \cdot N_w$, dove N_w è il numero totale di parametri liberi della rete.

In Figura 74 è riportato l'istogramma dell'errore angolare per $M = 20$; la mediana è molto alta anche rispetto al prefit.

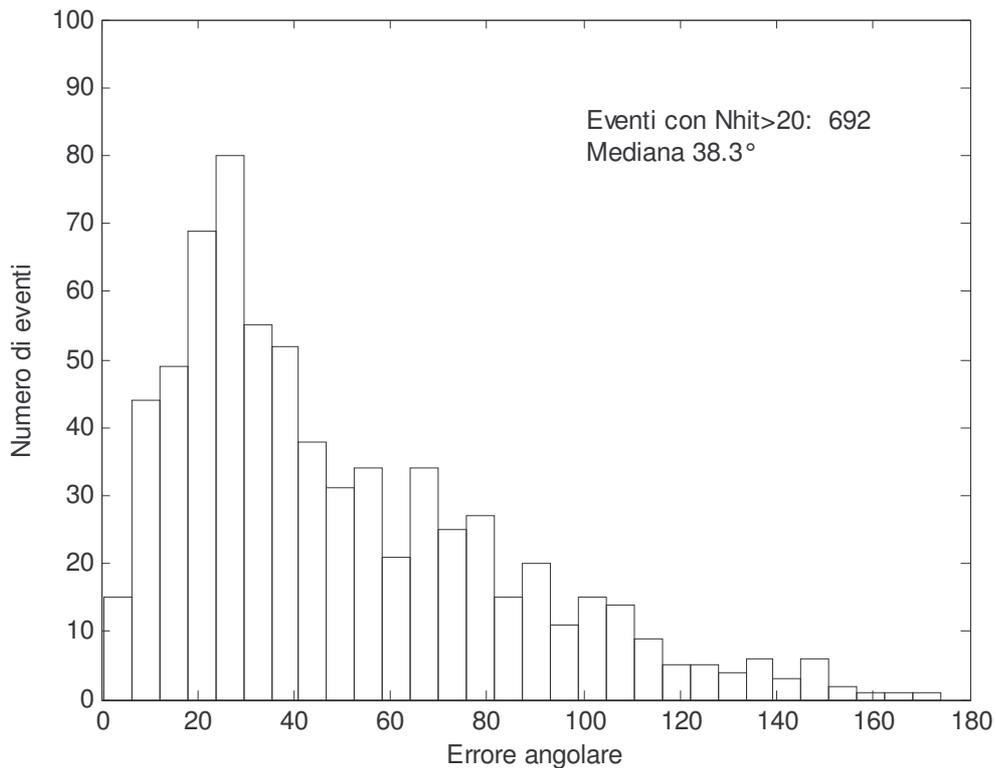


Figura 74 – Angolo nello spazio tra la direzione della traccia simulata e quella ricostruita ($M=20$).

Aumentando il numero di ingressi (in Figura 75 l'errore angolare per $M=30$) le performance non migliorano (mediana $\cong 50^\circ$), probabilmente perché il set di training non è sufficientemente ampio.

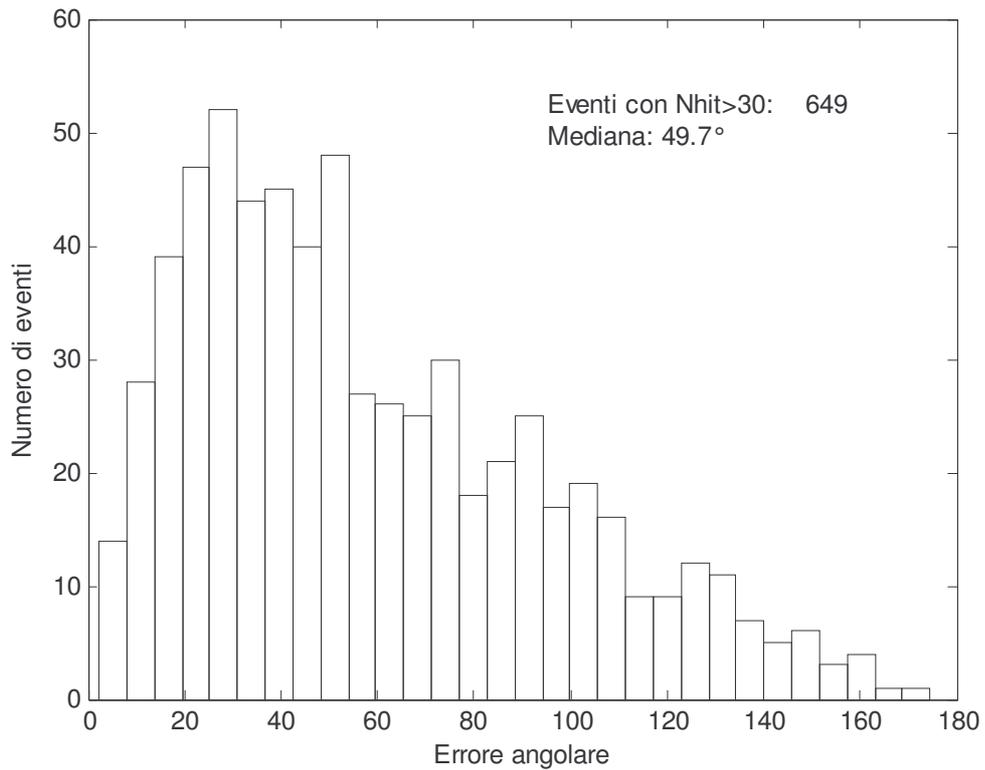


Figura 75 - Angolo nello spazio tra la direzione della traccia simulata e quella ricostruita (M=30).

Per alleggerire il compito della rete è stata tentata la ricostruzione della sola variabile angolare θ , cercando quindi di individuare solamente la direzione Up/Down. Le performance per M=10 e M=20 sono riportate in Figura 76 e in Figura 77 rispettivamente. I risultati sono migliorati, ma ancora sono peggiori del prefit (cfr. Figura 46).

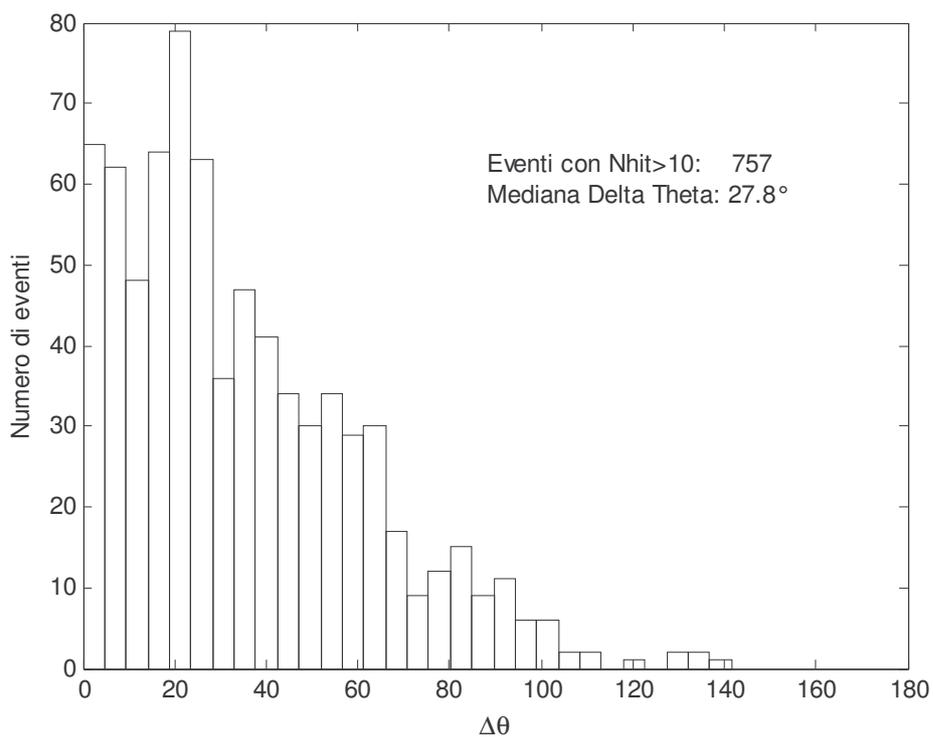


Figura 76 - Ricostruzione dell'angolo θ con rete neurale con $M=10$.

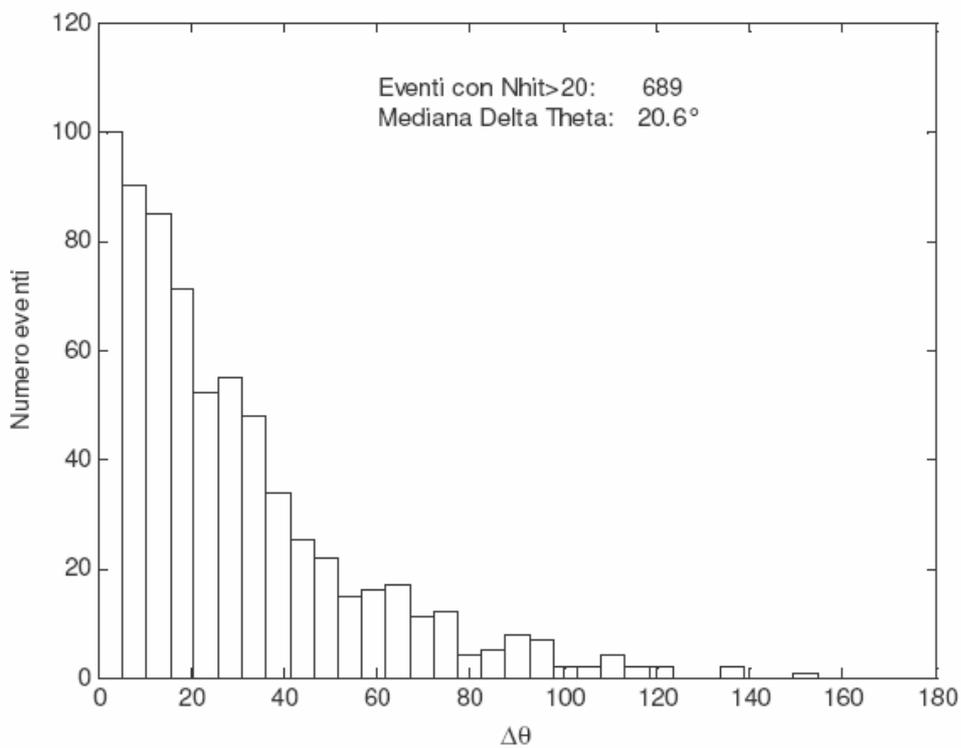


Figura 77 - Ricostruzione dell'angolo θ con rete neurale con $M=20$.

Un miglioramento ulteriore, che porta la mediana dell'errore in θ a 11.5° , è stato quello di considerare come ingressi alla rete le variabili dx_{ij}/dt_{ij} , dy_{ij}/dt_{ij} , dz_{ij}/dt_{ij} , dove i, j sono gli indici di alcune coppie di hit, opportunamente scelti. In questo caso sono stati scelti gli hit con carica misurata maggiore, e sono state fatte le differenze con tutti gli altri hit. L'istogramma dell'errore in θ è riportato in Figura 78, si ha una accuratezza confrontabile con un buon prefit.

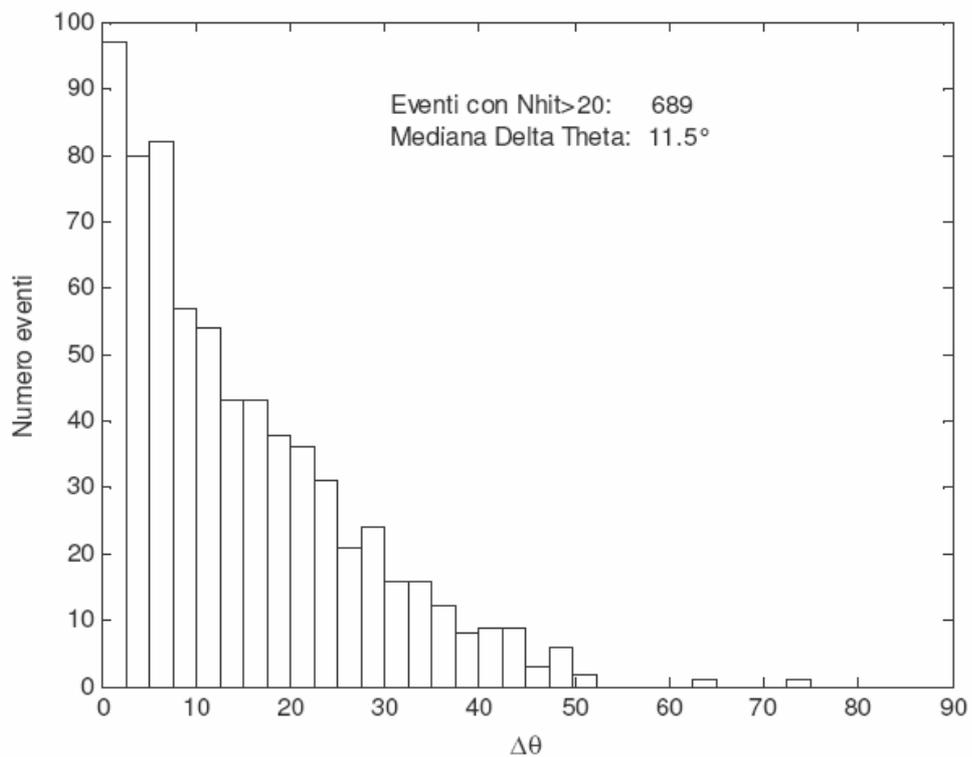


Figura 78 - Errore in θ della ricostruzione con rete neurale con ingressi differenziali e selezione degli hit tramite la carica.

Cap. 5 Ricostruzione delle tracce con algoritmi genetici

5.1 Introduzione agli Algoritmi Genetici

"Gli organismi viventi sono esperti solutori di problemi"; questa osservazione di John Henry Holland, l'inventore degli Algoritmi Genetici, è all'origine dei tentativi di emulare la versatilità e la potenza dell'evoluzione naturale per scopi informatici; i primi lavori risalenti agli anni '50 e '60 erano incentrati sulla simulazione dell'evoluzione naturale (Barricelli nel 1954 simulava l'evoluzione di automi in grado di misurarsi con un semplice gioco di carte); in quel periodo maturò l'idea che l'evoluzione "naturale" poteva essere un buon metodo di ottimizzazione. Fu però con Holland, e con i suoi lavori dei primi anni '70 che fu coniato il termine algoritmi genetici, che salirono agli onori della cronaca nel 1975 con il suo libro "Adaptation in Natural and Artificial Systems" [46].

Gli organismi viventi evolvono attraverso la selezione naturale e la riproduzione (nella maggior parte dei casi sessuata), a cui si aggiunge una possibile mutazione casuale (dai possibili tristi effetti nella nostra specie), che però ha il pregevole compito di salvaguardare intere specie dal pericoloso effetto di ristagno dato dall'involutione. In natura l'individuo che trasmette alle generazioni successive il proprio patrimonio genetico è quello che riesce a sopravvivere più a lungo e che riesce a vincere le competizioni per aggiudicarsi il diritto ad accoppiarsi, in una parola il più "adatto". I figli hanno un'alta probabilità di ereditare le buone caratteristiche dei genitori, anche dopo il parziale rimescolamento del materiale genetico tra madre e padre, anzi c'è una probabilità non trascurabile che uno dei figli (più sono e meglio è) erediti le caratteristiche migliori di entrambi i genitori; i figli peggiori saranno destinati nel migliore dei casi ad essere passivi spettatori della storia dell'evoluzione, o altrimenti a soccombere.

Questa caratteristica dell'evoluzione verso il miglior adattamento (fitness) è stata sfruttata per implementare un algoritmo di ottimizzazione simulando una popolazione di individui, una funzione di fitness (che è la funzione da massimizzare), un cross-over tra gli individui, e una mutazione. Ogni individuo rappresenta un punto nello spazio delle possibili soluzioni e durante l'algoritmo la popolazione evolve verso punti via via migliori (con fitness più elevata).

Un primo problema da risolvere è quello della codifica dell'informazione nei "geni": "mutando" o "accoppiando" il testo di un programma FORTRAN non si ottiene un programma FORTRAN migliore o peggiore di quello iniziale, ma di solito si ottiene un errore di compilazione... Bisogna che il meccanismo del cross-over e della mutazione diano risultati accettabili, anche se in qualche caso peggiori di quelli di partenza.

Bisogna pertanto individuare le variabili significative del problema e una loro funzione da massimizzare; questo potrebbe già essere un compito arduo, tuttavia questa funzione può essere molto generica, non ci sono particolari vincoli di regolarità, anzi la funzione può anche non essere definita in modo algebrico.

I parametri rispetto a cui ottimizzare la funzione vanno codificati, opportunamente quantizzati, in stringhe binarie. In questo modo una funzione di due parametri reali diciamo compresi tra 0 e 1:

$$x, y \in [0,1]$$

quantizzati a 16 bit (da 0 a 65535) sarà ricercata in una "popolazione" di cromosomi binari di lunghezza 32. Il numero di possibili soluzioni è quindi 2^{32} , pari a circa 4 miliardi di combinazioni.

La cosa più importante negli algoritmi genetici è sicuramente la funzione di "Fitness" che è il vero legame con la realtà che si vuole simulare e ottimizzare. L'algoritmo genetico infatti ricercherà il massimo (o se si vuole il minimo, e allora invece di fitness si potrà chiamare funzione costo o errore) di questa funzione.

La caratteristica che rende gli algoritmi genetici molto interessanti è che la ricerca del minimo avviene in modo "globale", cioè non c'è il pericolo che la soluzione si "intrappoli" in un minimo locale della funzione, come invece avviene per gli usuali metodi detti "Steepest Descent" (o "Hill Climbing" a seconda del verso della funzione). Appare chiaro quindi che l'utilità sarà massima quando la funzione è molto complessa, è fortemente non lineare, ha molti massimi e minimi relativi, non è facilmente descrivibile in termini matematici.

Un'altra caratteristica è il "parallelismo intrinseco" dovuto all'esistenza di una "popolazione" di possibili soluzioni che esplorano lo spazio in modo caotico (o meglio semi-caotico); questa caratteristica è comune ad esempio agli algoritmi randomici, ma si vedrà che l'operazione di accoppiamento (cross-over) degli individui migliori (selezione) indirizza l'esplorazione dello spazio in modo ottimo (o quasi, ma comunque in modo più efficiente della ricerca randomica).

L'algoritmo comincia con una "popolazione" iniziale di elementi random. Ad ogni iterazione (generazione) viene valutata la fitness di ogni elemento.

Vengono selezionati gli elementi migliori per l'accoppiamento (cross-over); di solito si utilizza un criterio di Selezione Proporzionata alla fitness (Roulette Wheel Selection), in base al quale un individuo x_j è scelto come genitore con probabilità

$$P_{SEL} = \frac{fit(x_j)}{\sum_i fit(x_i)}$$

Per ogni coppia di genitori selezionati si applica il cross-over con una certa probabilità P_C preassegnata (ad es. 80%).

Vi sono vari tipi di cross-over; il più semplice è quello con un punto di taglio: data una coppia di genitori viene scelto in modo casuale il punto di taglio (da 1 ad $n-2$ per cromosomi di n bit), e vengono scambiate le sequenze di bit successive al taglio (vedi Figura 79)

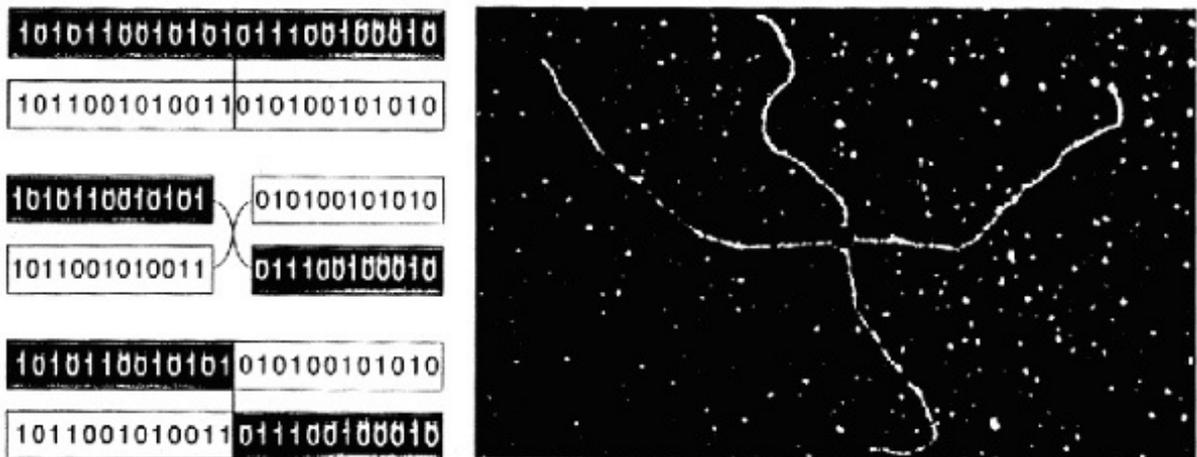


Figura 79 - Schematizzazione binaria e fotografia del momento del cross-over in cromosomi reali.

In questo modo i genitori trasferiscono ai figli parte del loro patrimonio genetico.

In seguito viene applicata la mutazione (cioè l'inversione di un bit random) con probabilità P_{Mut} di solito dell'ordine del 1%. Questo ha lo scopo di evitare in caso di convergenza evolutiva, il ristagno della popolazione.

A questo punto si riparte con la valutazione della fitness per i nuovi individui e così via, fino ad un criterio di terminazione dell'algorithm, che può essere basato sul valore del massimo della fitness per ogni generazione e/o sul numero di generazioni.

Il diagramma di flusso dell'algorithm è riportato in Figura 80.

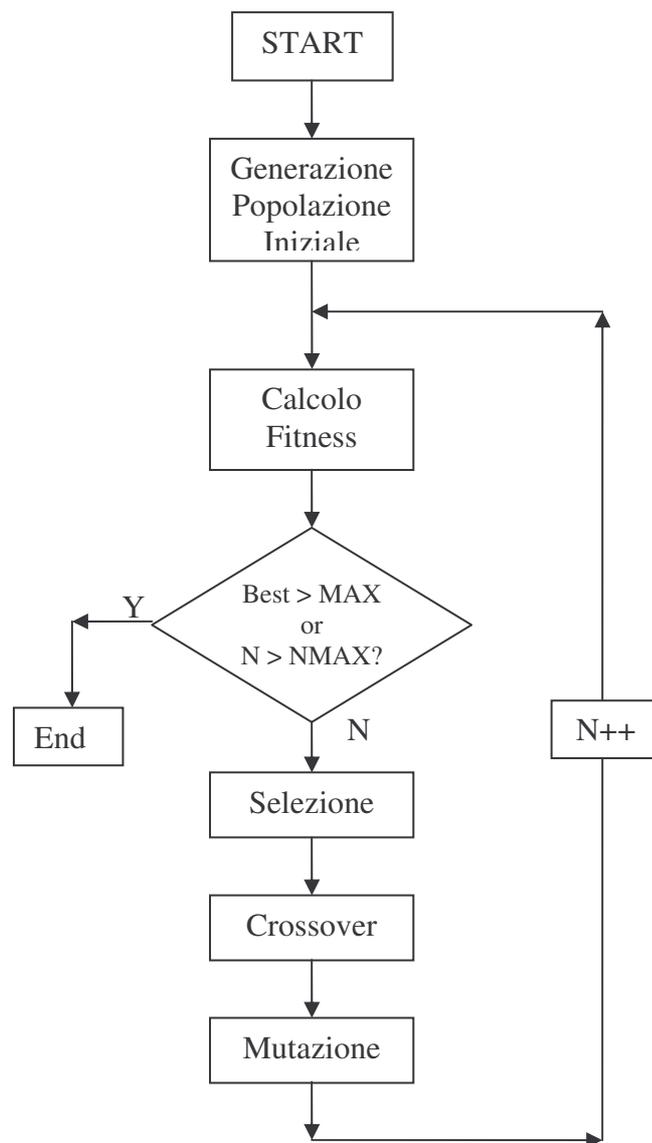


Figura 80 - Diagramma di flusso di un algoritmo genetico.

5.1.1 Teoria: perché funziona? Iperpiani negli spazi di ricerca.

La domanda che molti si pongono a questo punto è: perché un tale processo dovrebbe produrre qualcosa di utile? Perché tutto questo dovrebbe portare ad un'efficace forma di ottimizzazione?

La risposta che di solito viene data per spiegare il comportamento degli algoritmi genetici è il risultato del lavoro di Holland, che nel suo già citato libro [46] sviluppa vari argomenti per spiegare come un algoritmo genetico può fornire una ricerca complessa e robusta andando a campionare partizioni iper-planari dello spazio di ricerca.

Il miglior modo di capire come un algoritmo genetico può campionare iperpiani dello spazio di ricerca è di considerare un semplice spazio 3-dimensionale (vedi Figura 81). Assumendo di avere un problema codificato con soli 3 bit, lo spazio di ricerca può essere rappresentato da un cubo (la stringa 000 coincide con l'origine degli assi). Gli angoli del cubo sono individuati da tutte le possibili stringhe di 3 bit, e angoli adiacenti hanno stringhe che differiscono di un solo bit (vedi Figura 81). Il piano frontale del cubo contiene tutti i punti che iniziano per 0. Utilizzando il simbolo * come jolly questo piano può essere indicato dalla stringa 0**. Le stringhe contenenti * sono dette schemi⁷.

Ogni schema corrisponde ad un iperpiano nello spazio di ricerca. L'ordine di un iperpiano è il numero di bit indicati effettivamente nello schema: quindi ad esempio 0** è uno schema di ordine 1, mentre 1**1*****0* è di ordine 3.

Per rappresentare uno spazio di ricerca 4-dimensionale si può inserire un cubo nel cubo precedente, come in basso nella Figura 81. I punti possono essere indicati come per il cubo originario, sia quelli appartenenti al cubo esterno che quelli del cubo interno con le stringhe di 3 bit precedenti; a questo punto basta aggiungere all'inizio delle stringhe uno 0 per i punti del cubo esterno e un 1 per i punti del cubo interno. Il cubo interno ora è individuato dallo schema 1*** e quello esterno dallo schema 0***. Analogamente è facile vedere che lo schema *0** indica i punti appartenenti ai piani frontali di entrambi i cubi, e 10** è il piano frontale del cubo interno.

⁷ schema o al plurale schemata dal greco σχήμα, pl. σχήματα che vuol dire forma, piano

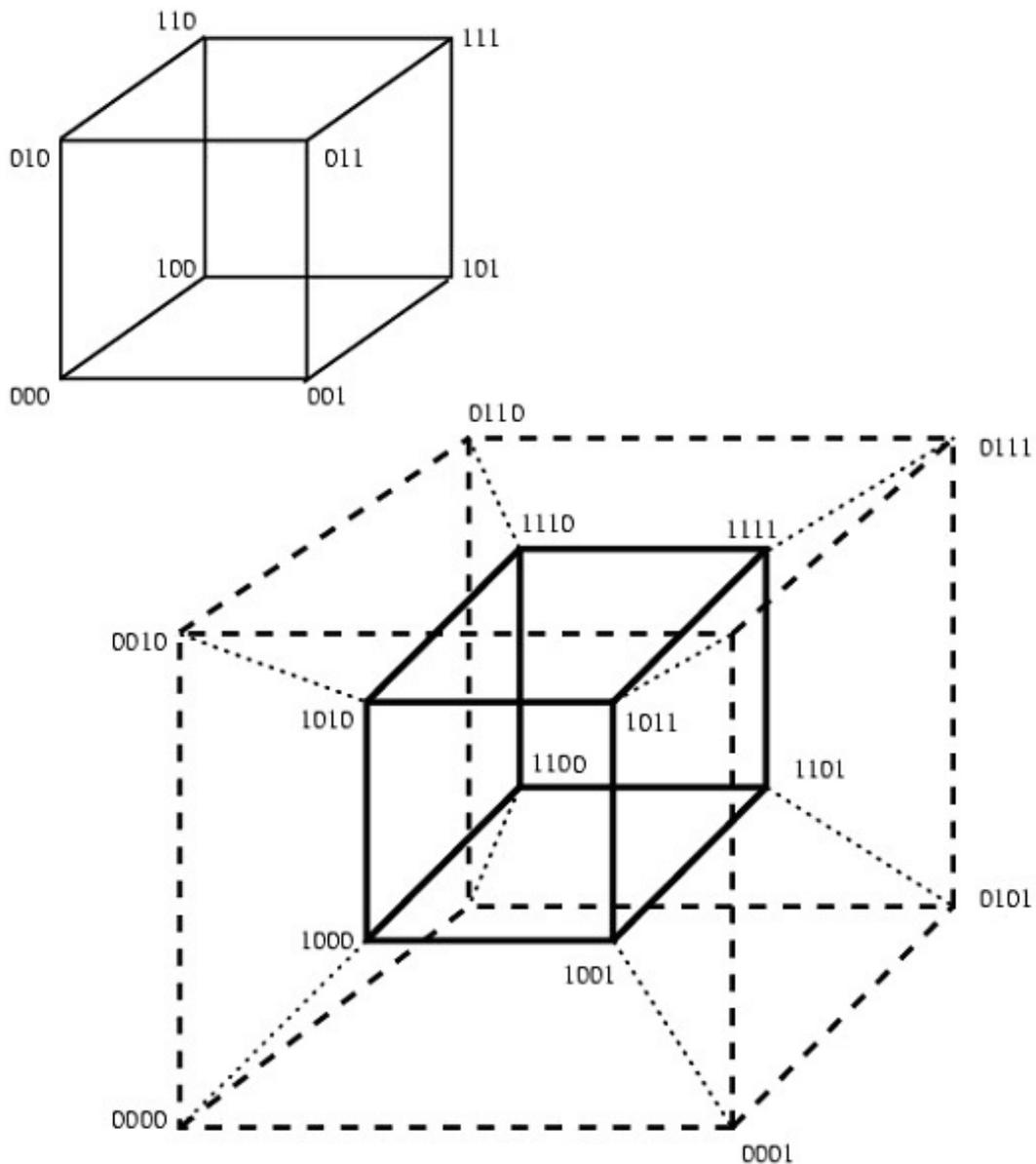


Figura 81 - Cubo e ipercubo di ordine 4 [45].

Ogni codice binario è un "cromosoma" che corrisponde ad un punto dello spazio di ricerca. Ogni punto è inoltre membro di $2^L - 1$ iperpiani: sono i modi in cui posso disporre degli * in una stringa di L bit data. Il numero totale di iperpiani nello spazio di ricerca è invece $3^L - 1$: in questo caso ogni bit può assumere il valore 0, 1 o *. L'iperpiano banale coincidente con l'intero spazio di ricerca non è considerato, mentre sono considerati come iperpiani anche i singoli punti (iperpiani di ordine L).

Il parallelismo intrinseco in un algoritmo genetico non è solamente dato dal fatto che abbiamo una popolazione di possibili soluzioni, ma anche dal fatto che ogni cromosoma in realtà campiona un gran numero di iperpiani dello spazio di ricerca. L'effetto cumulativo di N iperpiani per M individui della popolazione fornisce una informazione statistica su tutti i sottoinsiemi di iperpiani [46].

Il parallelismo intrinseco implica che molte competizioni tra iperpiani vengano risolte simultaneamente in parallelo. La teoria suggerisce che attraverso il processo di riproduzione e ricombinazione, gli schemi aumentino o diminuiscano la loro rappresentanza nella popolazione a seconda della fitness delle stringhe che giacciono in quegli iperpiani.

Un altro modo di considerare le partizioni a iperpiano è presentato in Figura 82. Viene disegnata la funzione fitness da massimizzare in funzione di una variabile di uno spazio unidimensionale. L'iperpiano $0^{***}..**$ ricopre la prima metà dello spazio, e $1^{***}...**$ ricopre la seconda metà. Poiché le stringhe dell'iperpiano $0^{***}...**$ sono in media migliori di quelle dell'iperpiano $1^{***}...**$, si vorrebbe che la ricerca fosse indirizzata preferibilmente sulla prima partizione. Nella seconda parte della figura è evidenziata la partizione $**1**...**$ e si nota anche la sovrapposizione tra $0^{***}...**$ e $**1**...**$, cioè l'intersezione $0*1^{***}...**$. Quindi nel terzo grafico si può osservare la partizione $0*10^{***}...**$.

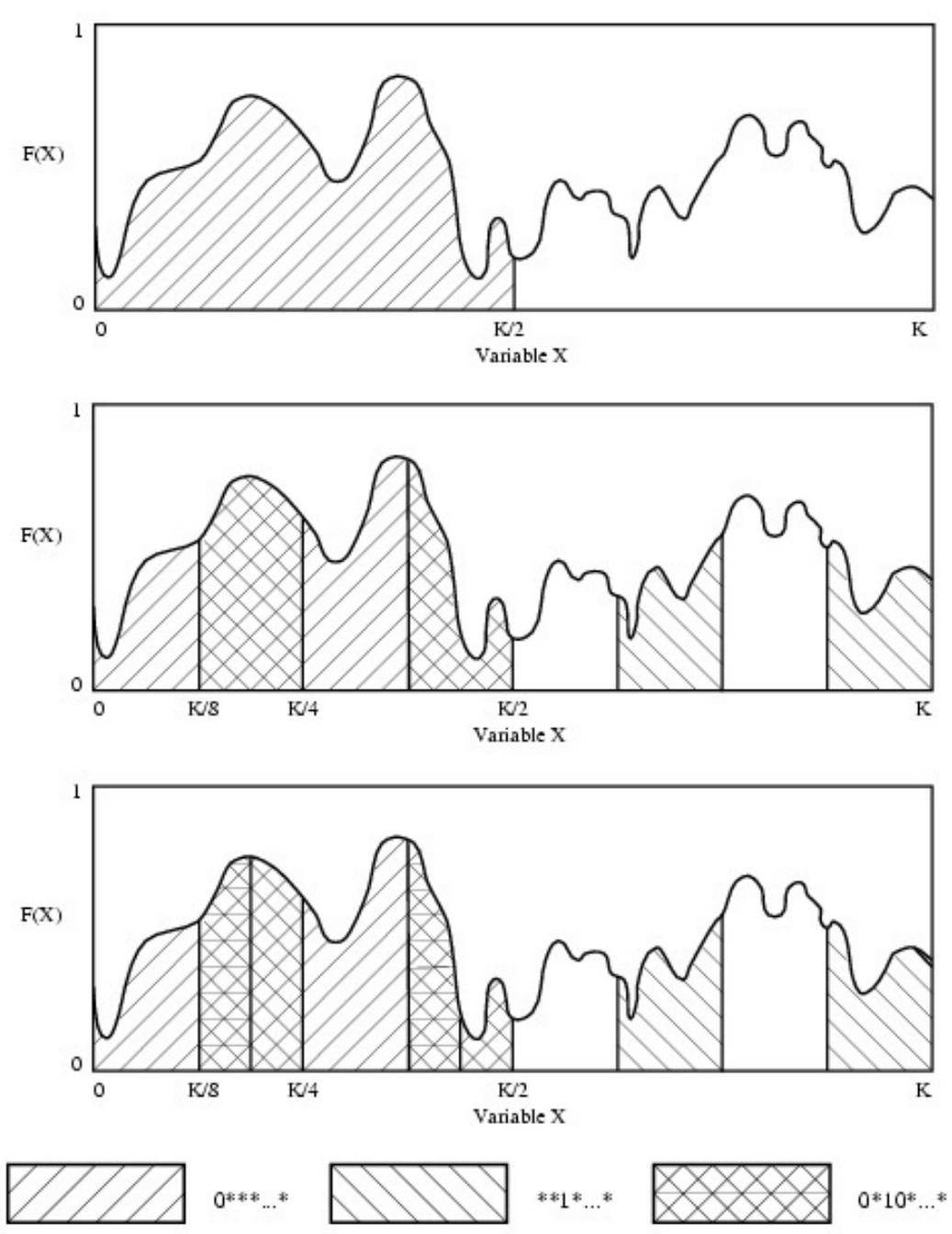


Figura 82 - Rappresentazione degli iperpiani nel dominio della funzione fitness [45].

Dalla Figura 82 si capisce che il campionamento delle partizioni non è condizionato da massimi locali. Aumentando il tasso di campionamento di partizioni che hanno valori di fitness superiori alla media aumenta la probabilità di trovare soluzioni ottimali.

Indicando con $M(H,t)$ il numero di stringhe di bit che appartengono ad un certo iperpiano H alla t -esima generazione, è possibile trovare un limite inferiore a $M(H,t+1)$

(Schemata theorem). Cioè, è possibile dimostrare che il campionamento delle soluzioni si infittisce sugli iperpiani che presentano una fitness superiore alla media.

Intanto indichiamo con $M(H,t+1/2)$ il numero di stringhe appartenenti ad H dopo la selezione, ma prima del cross-over e della mutazione. Eseguendo la selezione proporzionale si avrà

$$M\left(H,t+\frac{1}{2}\right) = M(H,t) \frac{f(H,t)}{\bar{f}} \quad (9)$$

dove $f(H,t)$ è la fitness dell'iperpiano H , e \bar{f} è la fitness media su tutta la popolazione.

Vediamo come il cross-over e la mutazione intervengono a modificare la (9).

Operatori di Cross-Over e Schemi

Osserviamo innanzi tutto che gli iperpiani di ordine 1 non vengono modificati dall'operatore di cross-over, infatti l'unico bit del relativo schema di ognuno dei genitori viene ereditato da uno dei due discendenti. Invece la distribuzione dei campioni degli iperpiani di ordine superiore al primo può essere alterata dal cross-over. Inoltre iperpiani dello stesso ordine possono essere alterati con probabilità diverse.

Cross-over ad 1 punto di taglio

Questo operatore è comodo perché è abbastanza semplice quantificarne l'effetto sui vari schemi. Assumendo di lavorare su stringhe di 12 bit consideriamo i seguenti schemi: 11***** e 1*****1. La probabilità che i due bit del primo schema vengano separati dal cross-over è soltanto

$$\frac{1}{L-1}$$

perché ci sono $L-1$ punti di taglio nel cross-over a 1 punto. Al contrario la probabilità che il secondo schema venga alterato è unitaria, perché ogni punto di taglio separa i due bit dello schema. Per il cross-over ad un punto di taglio quindi la posizione dei bit dello schema è importante per determinare la probabilità che quei bit rimangano uniti dopo il cross-over, cioè che lo schema non venga alterato.

Cross-Over a 2 punti di taglio

Questo tipo di cross-over usa due punti scelti a caso; le stringhe si scambiano il segmento che cade tra i due punti. Si osservi che il cross-over a 2 tagli tratta le stringhe e gli schemi come se fossero chiusi a formare un anello, come illustrato nella Figura 83



Figura 83 - Schemi nel cross-over a 2 tagli.

dove b1 .. b12 rappresentano i bit da 1 a 12. Quando visto in questo modo, il cross-over ad un punto è un caso particolare di quello a 2 punti: dove cioè uno dei due punti capita sempre nella posizione di ricongiungimento tra il primo e l'ultimo bit. Ora il massimo disturbo per gli schemi di ordine 2 si ha quando i due bit occupano posizioni complementari sull'anello. Quindi gli schemi che hanno bit vicini tra loro hanno meno probabilità di essere scompaginati dal cross-over ad uno o a 2 tagli. Più precisamente gli iperpiani rappresentati da schemi compatti saranno campionati con frequenza più vicina a quella che si avrebbe con la sola selezione, minimizzando così la probabilità che il cross-over ne disturbi la distribuzione.

Ogni iperpiano o schema è quindi caratterizzato dalla sua "Lunghezza", cioè dalla distanza tra il primo e l'ultimo bit non jolly (0 o 1), indichiamola con $\Delta(H)$. Maggiore è la lunghezza caratteristica di un iperpiano, maggiore sarà la probabilità che esso venga distrutto durante il Cross-Over. In particolare, per il Cross-Over ad un taglio si ha che la probabilità di distruzione (p_D) dello schema che rappresenta l'iperpiano H è:

$$p_D \propto \frac{\Delta(H)}{L-1}$$

In realtà un punto di taglio interno alla lunghezza caratteristica di uno schema non è detto che causi una distruzione dello schema stesso: infatti, se due sottosequenze omologhe dei genitori sono uguali, lo schema non viene interessato dal Cross-Over, questo avviene sicuramente se entrambi i genitori appartengono allo schema. Dato un elemento appartenente allo schema H ,

la probabilità che un partner random appartenga allo stesso schema è $P(H,t)$, pari a $M(H,t)$ diviso la dimensione della popolazione;

$$p_D \cong \frac{\Delta(H)}{L-1} [1 - P(H,t)] \quad (10)$$

Il teorema degli Schemi

Abbiamo visto l'effetto della selezione sulla popolazione di stringhe appartenenti ad un dato iperpiano, ora possiamo includere l'effetto del Cross-Over, per arrivare al cosiddetto teorema degli schemi [46].

Innanzitutto bisogna considerare che il Cross-Over è applicato con una probabilità minore di 1, quindi una parte della popolazione rimane inalterata, e questo dà luogo al primo termine pesato per $(1-p_C)$. La parte di popolazione che subisce il Cross-Over verrà alterata, e ci saranno delle stringhe che usciranno dallo schema (perdite), e altre che entreranno nello schema (guadagni). Le perdite sono valutabili attraverso la "distruzione" dello schema data dall'applicazione dell'operatore di Cross-Over (10), mentre i guadagni sono dovuti alla comparsa in uno schema di una stringa proveniente da genitori fuori dallo schema stesso; un esempio è dato dallo schema 11*****, e dai genitori 10***** e 01*****, con il Cross-Over ad un taglio tra il primo e il secondo bit: uno dei figli entra nello schema dato. Dopo il Cross-Over il numero di stringhe $M(H, t+1)$ sarà:

$$M(H, t+1) = (1 - p_C) M(H, t) \frac{f(H, t)}{\bar{f}} + p_C \left[M(H, t) \frac{f(H, t)}{\bar{f}} (1 - \text{perdite}) + \text{guadagni} \right]$$

Facendo un'ipotesi semplificativa, ma conservativa, non consideriamo ulteriormente i possibili guadagni, e consideriamo le perdite uguali alla probabilità di distruzione massima data dalla (10). A parte il caso fortunato in cui entrambi i genitori appartengano allo schema, si ha quindi:

$$M(H, t+1) \geq (1 - p_C) M(H, t) \frac{f(H, t)}{\bar{f}} + p_C \left[M(H, t) \frac{f(H, t)}{\bar{f}} \left(1 - \frac{\Delta(H)}{L-1} (1 - P(H, t)) \right) \right] \quad (11)$$

Dividendo per la dimensione della popolazione e riarrangiando, si ha il teorema degli schemi, cioè il limite inferiore per l'evoluzione della rappresentatività nella popolazione al tempo $t+1$ dell'iperpiano H :

$$P(H, t+1) \geq P(H, t) \frac{f(H, t)}{\bar{f}} \left[1 - p_C \frac{\Delta(H)}{L-1} \left(1 - P(H, t) \frac{f(H, t)}{\bar{f}} \right) \right] \quad (12)$$

Dalla (9) e dalle successive correzioni, si vede che gli individui con valori di fitness più elevata hanno maggiore probabilità di sopravvivere.

Quindi gli schemi con una lunghezza caratteristica non troppo elevata e con una fitness superiore alla media avranno una rappresentatività esponenzialmente crescente. In effetti questo è vero per qualche generazione, poi il bias introdotto nella distribuzione degli individui insieme alle approssimazioni fatte per giungere alla (11) non permette di valutare esattamente l'andamento nel tempo della rappresentatività. La (12) dà pertanto un'indicazione della convergenza della popolazione in media verso soluzioni migliori. Cross-Over, Mutazione e convergenza prematura

Dal teorema degli schemi sembra che il Cross-Over crei solo svantaggi, a causa della probabilità di distruggere uno schema con una buona fitness; questo tuttavia è solo un aspetto del complesso problema dell'evoluzione di un Algoritmo Genetico; infatti, se fosse attiva solo la selezione, con poche generazioni tutti gli individui appartenerebbero allo schema migliore presente al momento dell'inizializzazione, ma si assisterebbe ad una cosiddetta convergenza evolutiva. Per evitare questo impasse si utilizza l'operatore di Mutazione, che cambia dei bit random con probabilità p_{MUT} .

Altri meccanismi euristici per favorire l'evoluzione sono legati invece alla capacità di "esplorazione" dell'operatore di Cross-Over: il Cross-Over ad un taglio, dai genitori può produrre al massimo solo $2(L-1)$ figli diversi, mentre il Cross-Over a due tagli può generare $2\binom{L}{2} = L^2 - L$ figli diversi. Considerando il cosiddetto Cross-Over uniforme in cui vengono scelti i singoli bit dall'uno o dall'altro genitore in modo casuale il numero massimo di figli diversi è in questo caso esponenziale, e pari a $2^L - 2$.

5.2 Applicazione per la ricostruzione delle tracce

Un algoritmo genetico può essere utilizzato per minimizzare l'errore quadratico medio degli scarti dei tempi di hit (8), analogamente a quanto si fa nella ricostruzione OPNEMO. Gli individui che compongono la popolazione sono formati dalle stringhe di bit che codificano i parametri della traccia x_0 , y_0 , z_0 , θ , e ϕ , come rappresentato in Figura 84

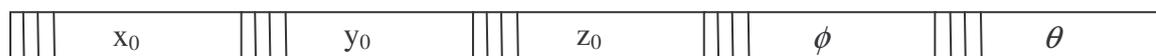


Figura 84 - Codifica dei parametri di traccia in un "cromosoma".

Posto lo schema di principio di un algoritmo genetico generico di Figura 80, i parametri liberi su cui si può agire per sperimentare ed ottimizzare la procedura sono molti. A priori è stata fissata la codifica delle variabili di traccia x_0 , y_0 , z_0 , θ , e ϕ in parole binarie con una risoluzione (Least Significant Bit: LSB) sufficientemente piccola in modo che l'errore di quantizzazione non influisca sull'accuratezza angolare. Scegliendo parole di 16 bit il LSB per θ e ϕ risulta pari a

$$LSB_{\theta} = \frac{180}{2^{16} - 1} = 0.00274662^{\circ}$$

$$LSB_{\phi} = \frac{360}{2^{16} - 1} = 0.00549325^{\circ}$$

Per quanto riguarda la posizione iniziale, poiché il volume netto occupato dai PMT ha le dimensioni (± 570 m, ± 570 m, ± 340 m) intorno al centro bisogna ricoprire 1140 m in x e y e 680 m in z, più almeno tre lunghezze di attenuazione dei fotoni. Prendendo $L_a=75$ m, cioè la lunghezza di attenuazione dell'acqua pura, si ha che le coordinate iniziali possono assumere valori in un intervallo di circa 1600 m. Una parola di 16 bit avrà quindi $LSB_{x,y,z} = 2.5$ cm, più che sufficiente per l'accuratezza sulla misura dei tempi ($\cong 0.083$ ns).

La trasformazione dal valore binario al valore in metri delle coordinate è:

$$x_0[m] = X_0 LSB_x - x_C$$

$$y_0[m] = Y_0 LSB_y - y_C$$

$$z_0[m] = Z_0 LSB_z - z_C$$

dove $x_C = y_C = z_C = 819.1875$ m sono le posizioni del centro dell'apparato rispetto al bordo.

Sono state valutate le prestazioni dell'algoritmo al variare di:

- dimensione della popolazione
- tipo di cross-over
- probabilità di mutazione.

Si è osservato a parità di altre condizioni, un miglioramento delle prestazioni per alti valori di mutazione; questo è dovuto al fenomeno della convergenza evolutiva, per cui dopo molte generazioni vi sono individui identici. Ciò causa la effettiva diminuzione della popolazione efficace. Per ovviare a questo inconveniente è stato introdotto inizialmente nell'algoritmo un controllo di esistenza di cloni, che in caso affermativo ne sostituisce uno con un nuovo individuo casuale. Tuttavia il carico computazionale di questo controllo è molto gravoso ($O(n^2)$). Lo stesso effetto benefico è stato ottenuto tramite un meccanismo di rinnovo parziale della popolazione, cioè ad ogni generazione i peggiori N_{WOR} individui vengono

sostituiti con elementi casuali, (anche N_{WOR} è un parametro da ottimizzare), e questo metodo ha dato buoni risultati con un contenuto carico computazionale.

I risultati delle simulazioni con gli algoritmi genetici sono descritti nel §5.3.

5.3 Implementazione della ricostruzione con Algoritmi Genetici

La popolazione iniziale dell'algoritmo genetico è stata generata a partire dagli angoli ottenuti dalla procedura di prefit, con una distribuzione aleatoria di deviazione standard pari a 200 m per le coordinate e 5° per gli angoli.

Sono state effettuate delle prove di ricostruzione al variare dei parametri caratteristici dell'algoritmo genetico, per trovare una configurazione ottimale. Per motivi di tempo di elaborazione sono stati testati solo dei tagli nello spazio dei parametri, e la configurazione scelta corrisponde ai valori dei parametri per cui si ha una prestazione migliore; in teoria dovrebbero essere fatti variare tutti i parametri contemporaneamente per avere un'ottimizzazione globale. D'altra parte si può ipotizzare che per qualunque valore di probabilità di cross-over (CO) o di mutazione le performance migliorino al crescere del numero di individui.

La grandezza indicativa della prestazione è la mediana dell'errore angolare, per decade di energia.

Nelle figure seguenti sono riportati i risultati della ricostruzione al variare della dimensione della popolazione, per i tre tipi di CO esaminati. In Figura 85 il CO ad un taglio, in Figura 86 il CO a 2 tagli e in Figura 87 il CO uniforme. La dimensione della popolazione è stata fatta variare da 200 a 400 individui.

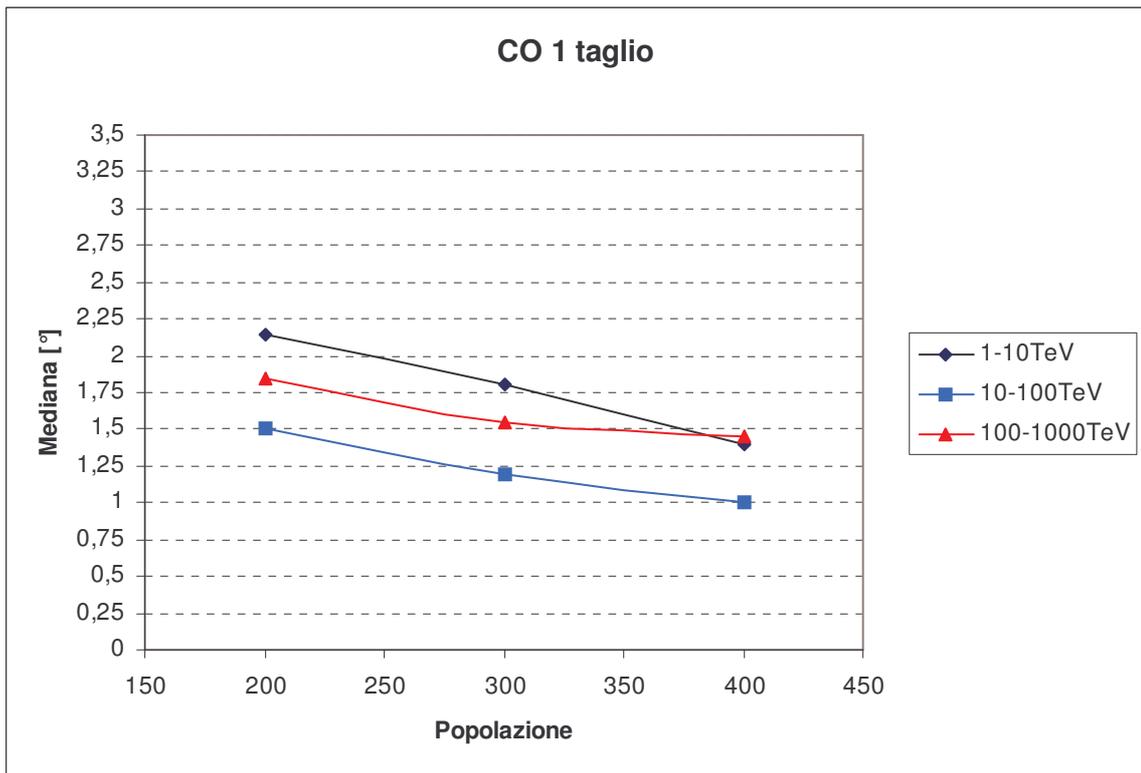


Figura 85 – Risultati della ricostruzione al variare della dimensione della popolazione, con il cross-over ad un taglio.

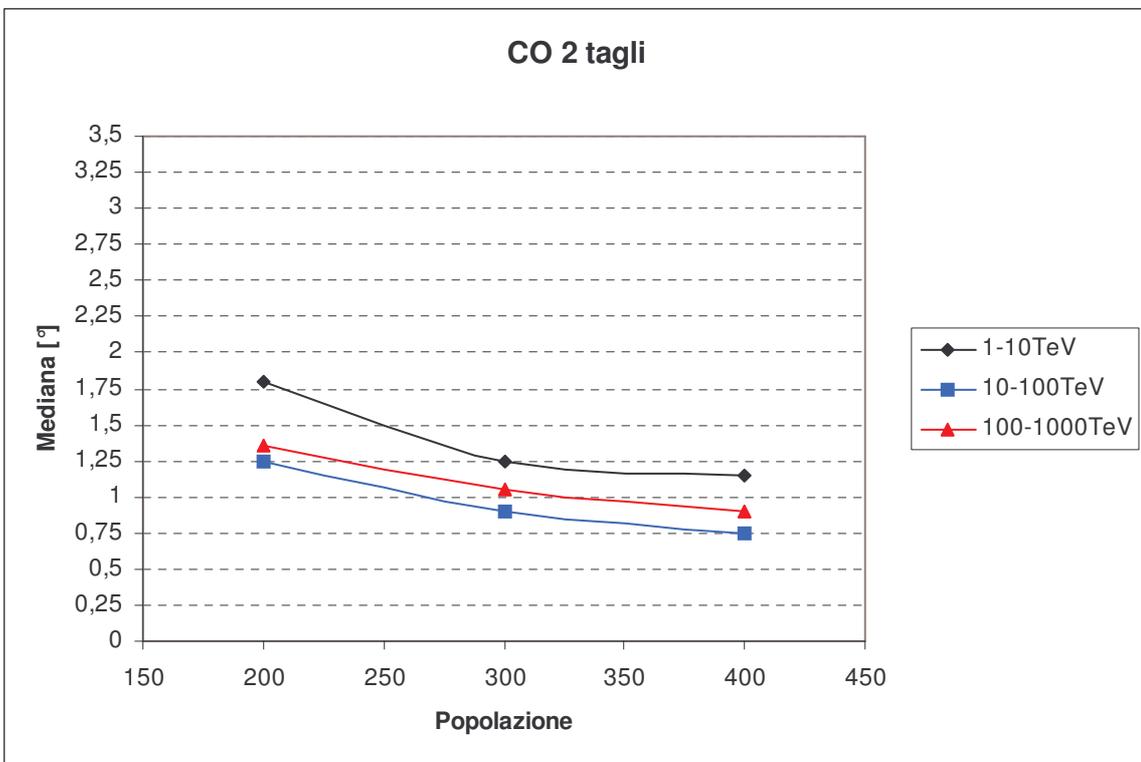


Figura 86 - Risultati della ricostruzione al variare della dimensione della popolazione, con il cross-over a due tagli.

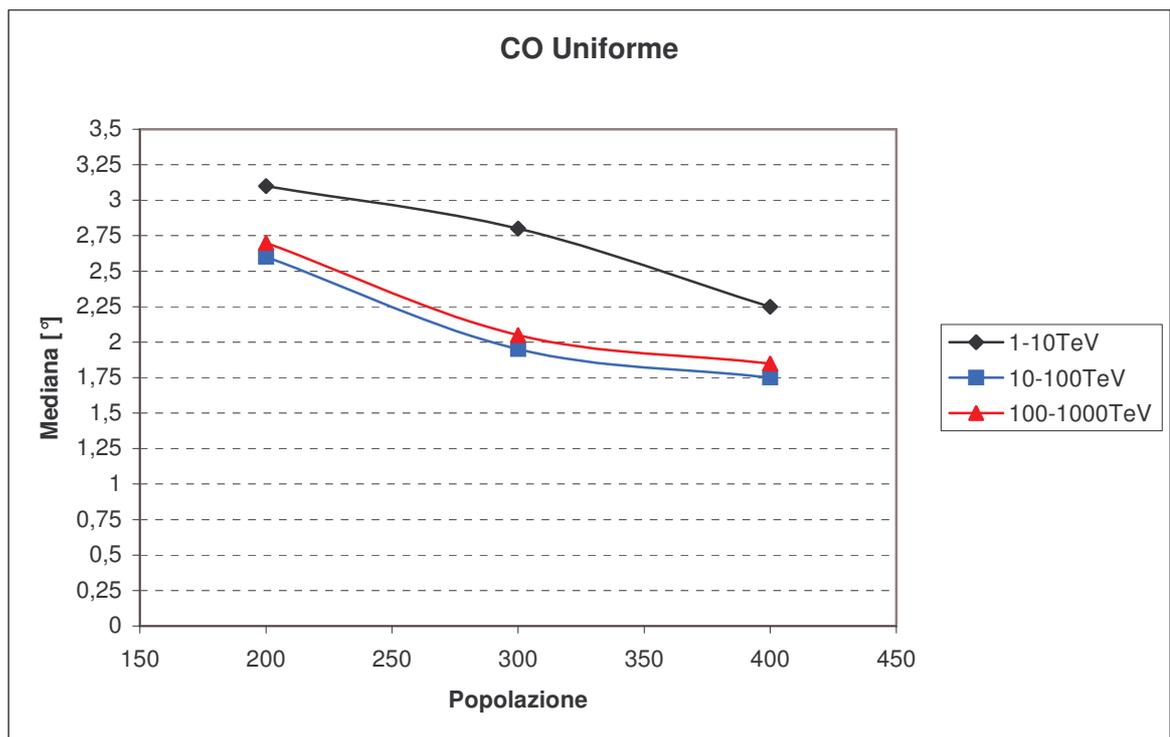


Figura 87 Risultati della ricostruzione al variare della dimensione della popolazione, con il cross-over uniforme.

Scelgo quindi di utilizzare sempre il CO a 2 tagli. Nelle prove successive, inoltre, utilizzerò un numero di individui nella popolazione pari a 200, ipotizzando che le performance migliorino comunque all'aumentare della popolazione.

Nella Figura 88 è riportata la bontà della ricostruzione al variare della probabilità di cross-over.

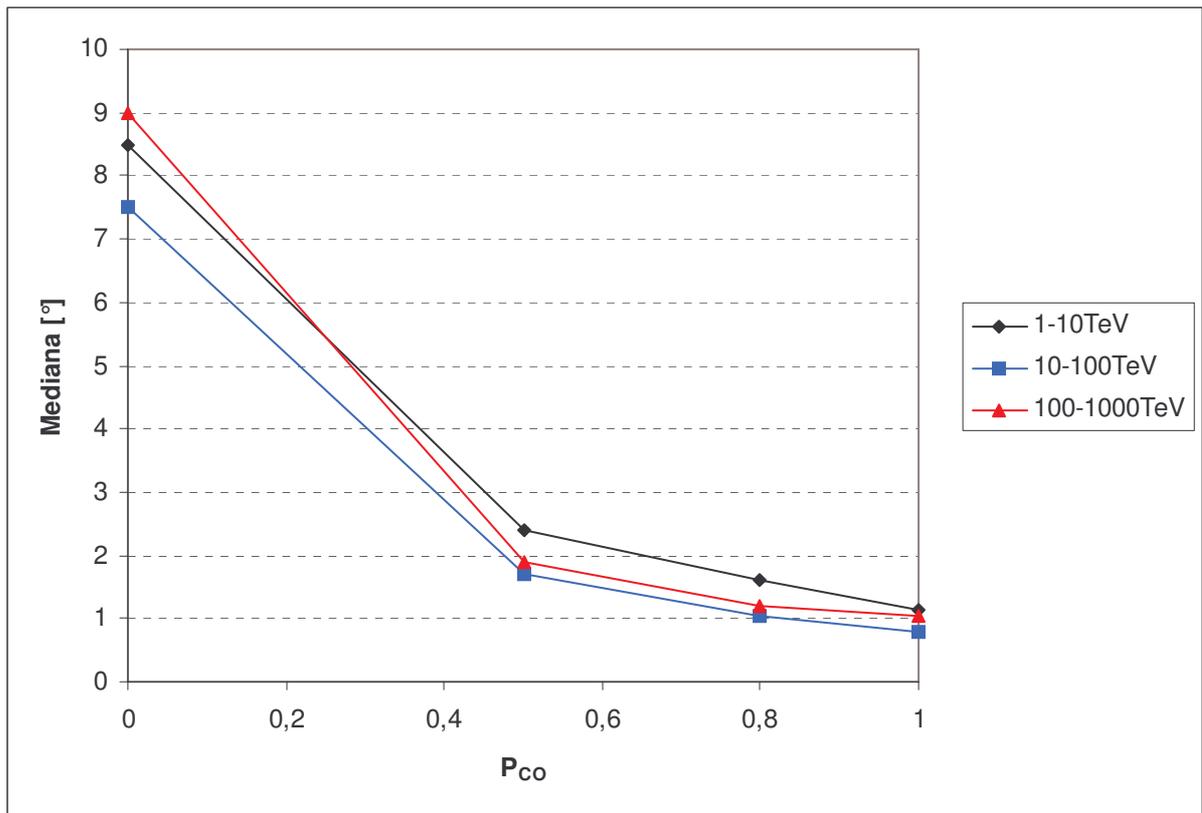


Figura 88 – Accuratezza di ricostruzione al variare della probabilità di cross-over.

Si vede che la migliore performance si ha con probabilità unitaria, d'altra parte in questa implementazione dell'algoritmo genetico il 2% dei migliori individui viene sempre trasmesso alla generazione successiva, per non perdere mai i risultati raggiunti. Il cross over viene applicato a tutti gli altri individui selezionati; è probabilmente per questo motivo che la probabilità di cross over ottima tende all'unità.

Per le successive prove è stata utilizzata una $P_{CO} = 1$.

Nella Figura 89 è riportata l'accuratezza della ricostruzione in funzione della probabilità di mutazione; si nota un miglioramento all'aumentare della mutazione. Ma per mutazioni superiori al 50% l'algoritmo genetico assomiglia sempre di più ad una ricerca randomica. Si è provato pertanto ad inserire nell'algoritmo un rinnovo della popolazione in modo casuale, sostituendo una percentuale dei peggiori individui. Il risultato di questa operazione è riportato in Figura 90 al variare della percentuale di individui rinnovati. Si vede che è stato trovato un minimo dell'errore angolare intorno al 10% di rinnovo, e poco variabile con la probabilità di mutazione.

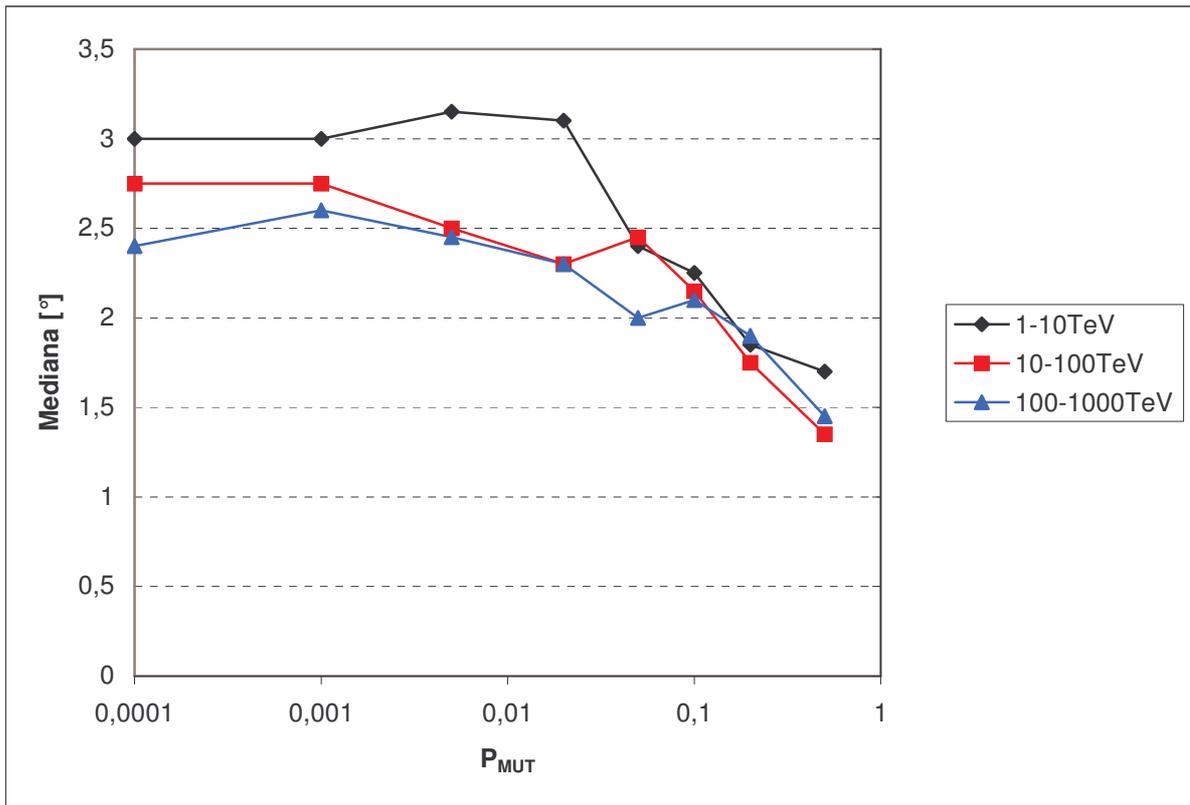


Figura 89 – Accuratezza della ricostruzione al variare della probabilità di mutazione.

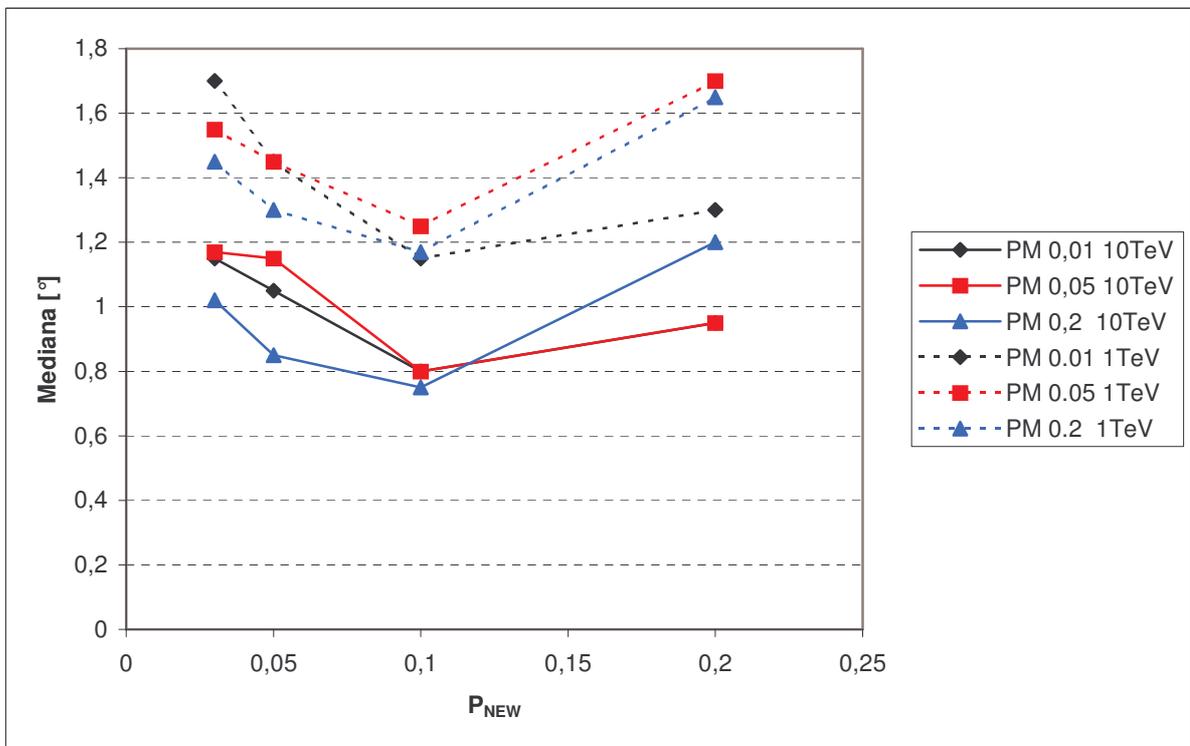


Figura 90 – Accuratezza di ricostruzione minima (linee continue, 10 - 100 TeV) e massima (linee tratteggiate, 1-10 TeV) in funzione della percentuale di rinnovo e della probabilità di mutazione.

La configurazione scelta per la ricostruzione è quella con il Cross Over a due tagli, con un rinnovo della popolazione pari al 10% ad ogni generazione, con probabilità di cross-over pari a 1, e con probabilità di mutazione del 2%.

Invece di aumentare molto il numero di individui nella popolazione, si è preferito ripetere la ricostruzione più volte per ogni evento, mantenendo la prova con più basso valore di χ^2 raggiunto.

Le prestazioni di ricostruzione ottenute utilizzando lo stesso campione di eventi del Cap. 3, sono riassunte nei grafici seguenti.

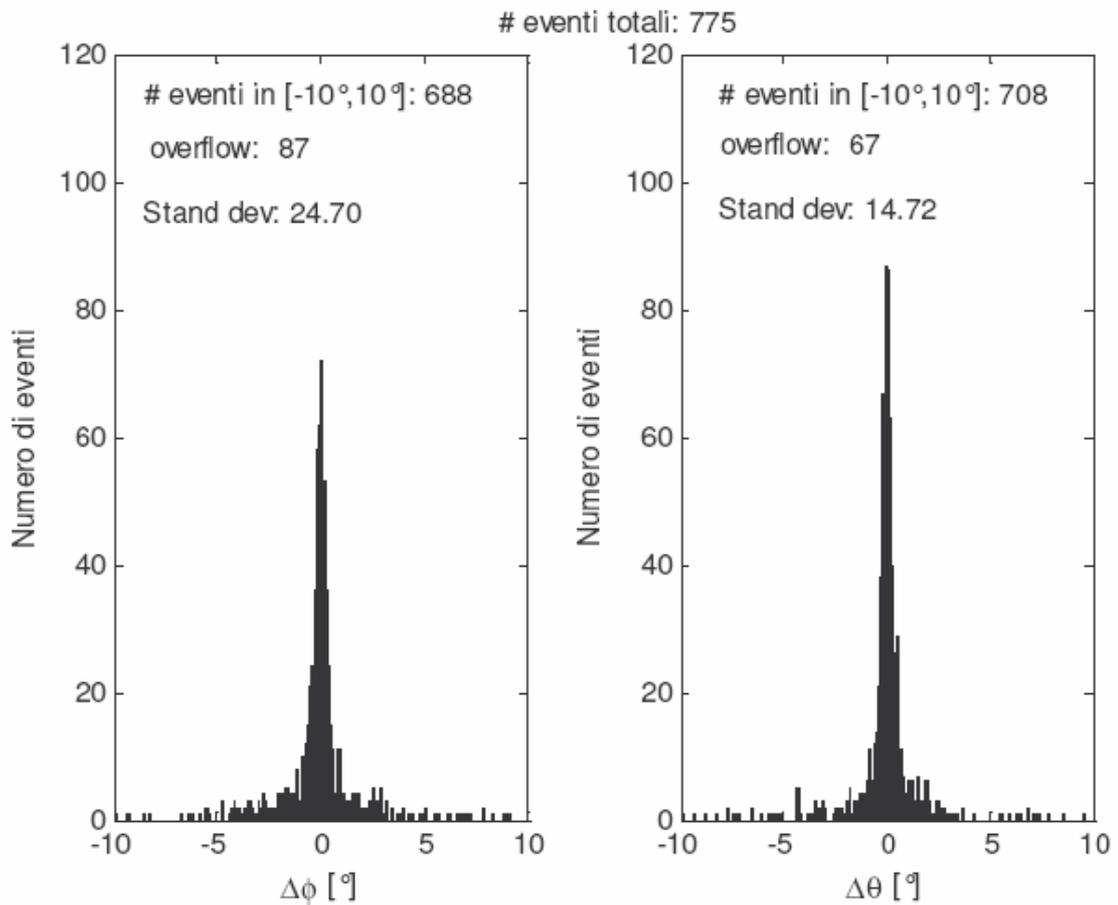


Figura 91 – Istogrammi delle differenze tra gli angoli ϕ e θ della traccia "vera" e quelli della traccia ricostruita dall'algorithm genetico.

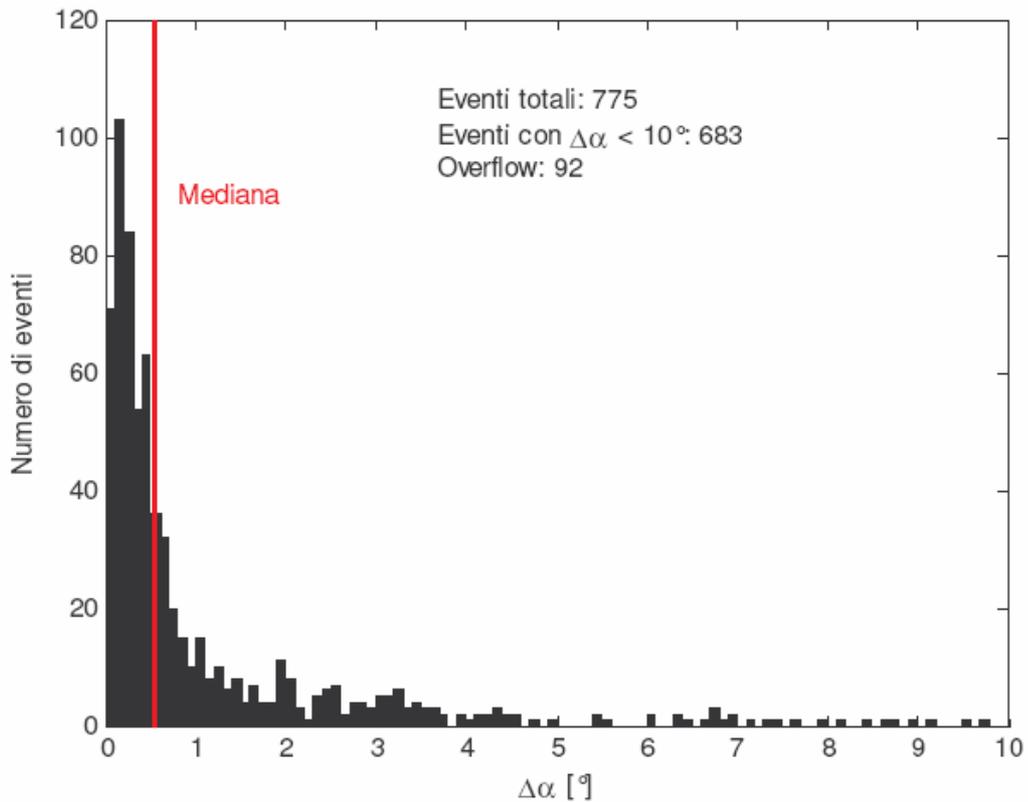


Figura 92 – Istogramma dell'angolo nello spazio tra la traccia "vera" e quella ricostruita.

Nella Figura 91 sono disegnati gli istogrammi delle differenze angolari in ϕ e θ , mentre in Figura 92 è riportato l'angolo nello spazio ($\Delta\alpha$) tra la traccia vera e quella ricostruita. Confrontandolo con quello della ricostruzione classica di Figura 52 si vede che le performance sono molto simili. La ricostruzione classica dà una mediana totale pari a $0.5^\circ_{-0.07}^{+0.1}$, mentre la ricostruzione con l'algoritmo genetico dà una mediana totale pari a $0.53^\circ_{-0.06}^{+0.08}$.

In Figura 93 è riportato lo scatter plot dell'errore angolare rispetto alle coordinate del vertice della traccia, e in Figura 94 rispetto a θ e ϕ .

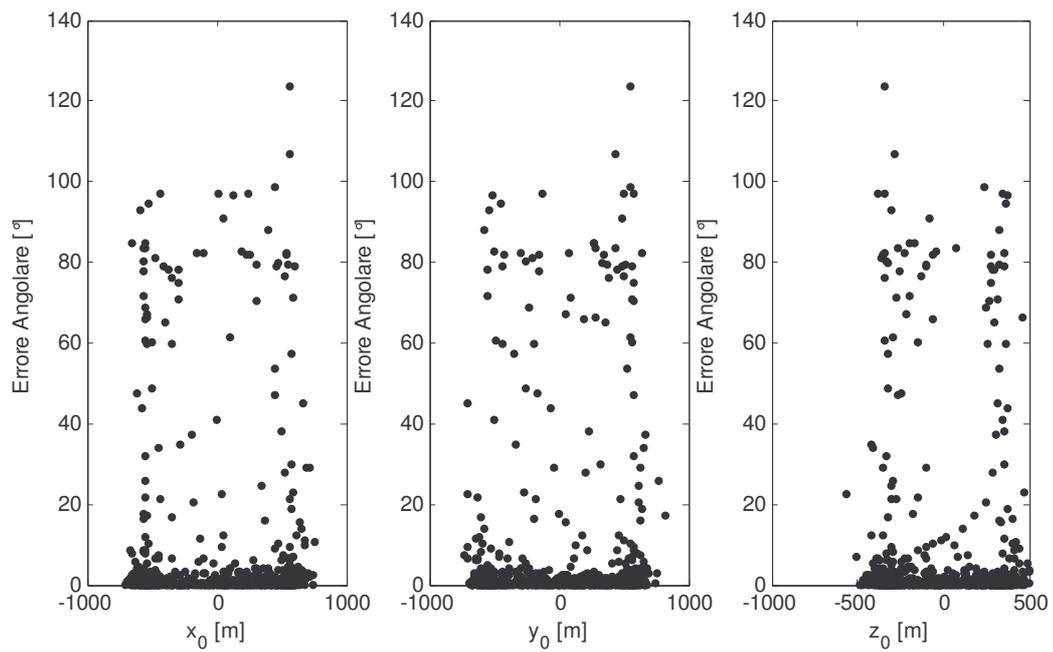


Figura 93 – Errore angolare vs coordinate dei vertici della traccia.

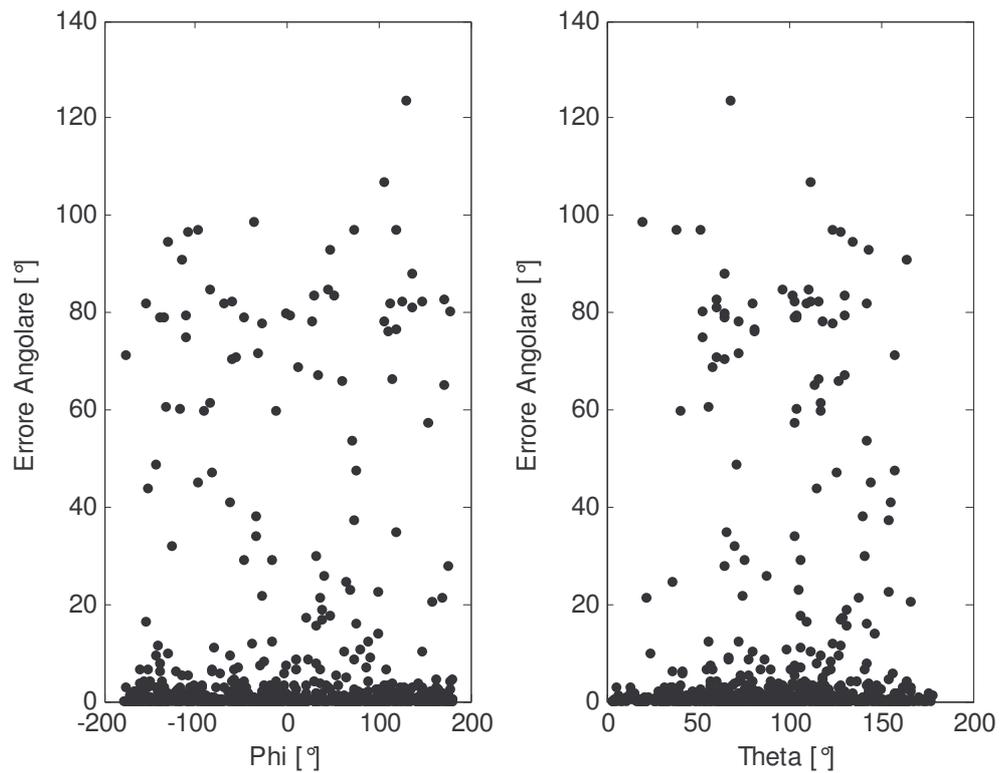


Figura 94 – Errore angolare vs θ e ϕ .

L'errore angolare rispetto all'energia è riportato in Figura 95 (scatter plot) e in Figura 96 (istogrammi delle standard deviation e delle mediane)

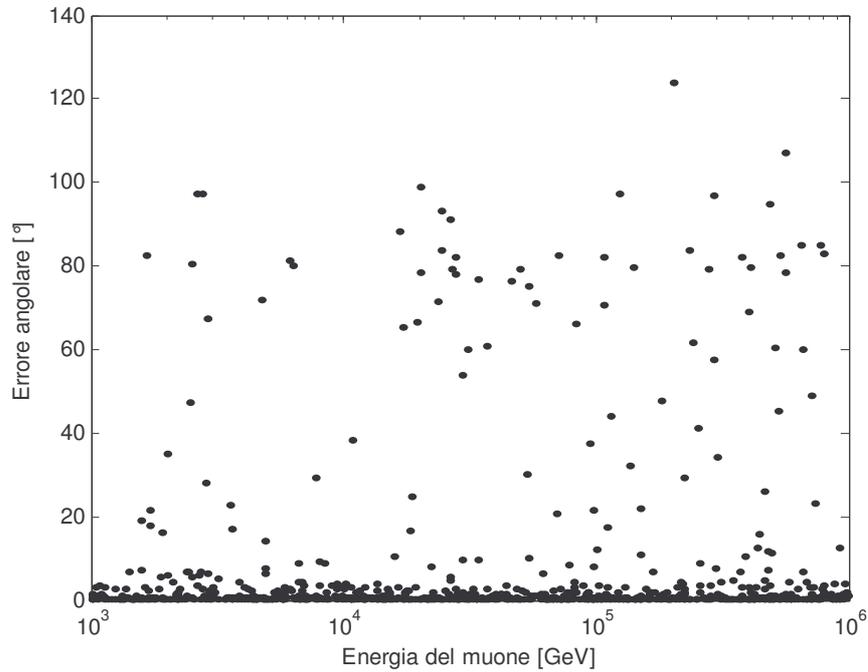


Figura 95 – Errore angolare in funzione dell'energia (scatter plot).

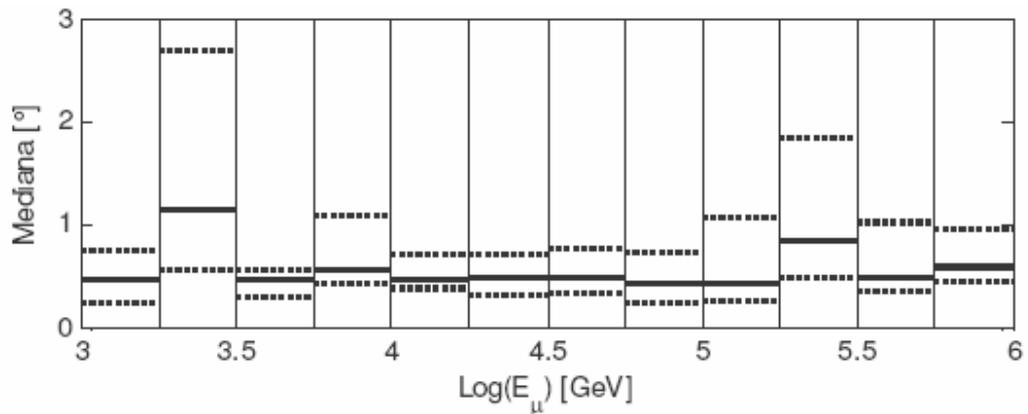


Figura 96 – Deviazione Standard e mediana della differenza angolare in funzione dell'energia (in linea tratteggiata l'errore a 1σ).

In Figura 97 è riportato l'errore angolare in funzione del numero di hit, in Figura 98 sono riportati gli istogrammi (standard deviation e mediana), insieme alla mediana totale degli eventi con N_{hit} superiore al valore corrente (linea rossa).

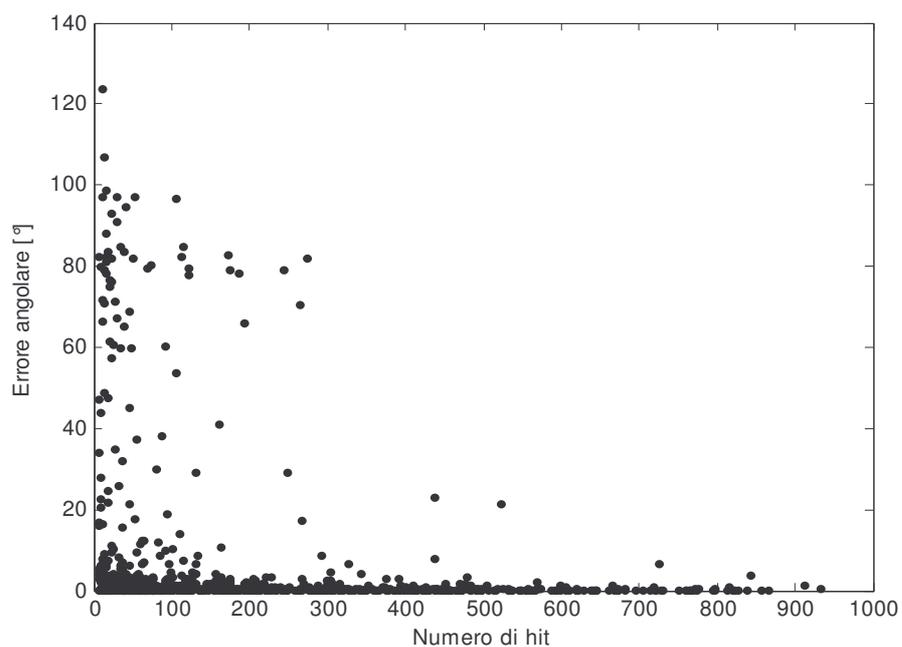


Figura 97 – Errore angolare rispetto al numero di hit.

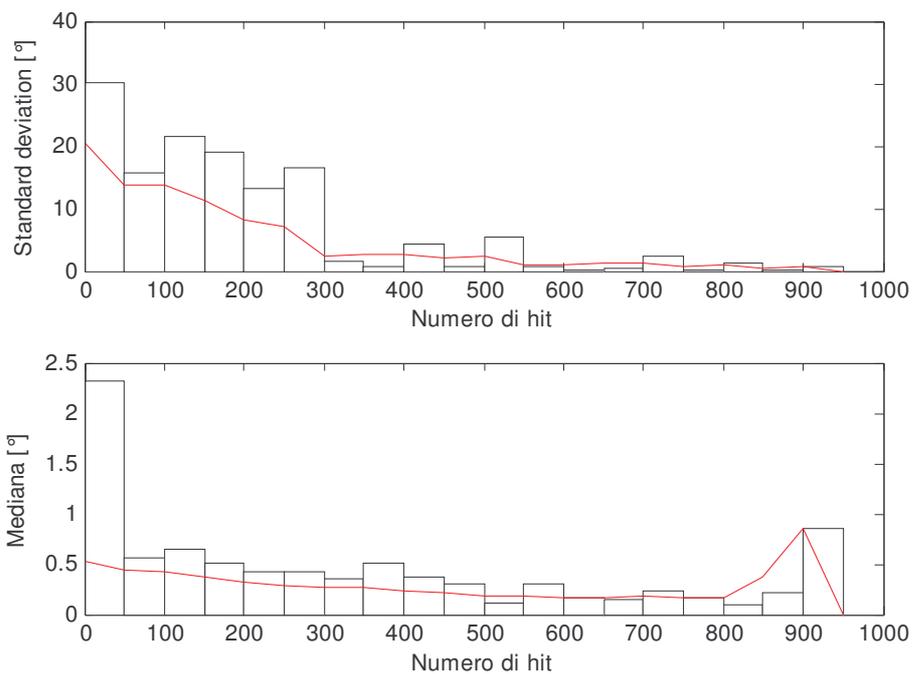


Figura 98 – Istogrammi della standard deviation e della mediana in funzione del numero di hit, e valori totali per numero di hit superiore al valore corrente (linea rossa).

In Figura 99 è riportata la mediana dell'errore angolare in funzione del χ^2 . Applicando lo stesso filtro sul χ^2 scelto nel §3.6, si hanno le performance di Figura 100.

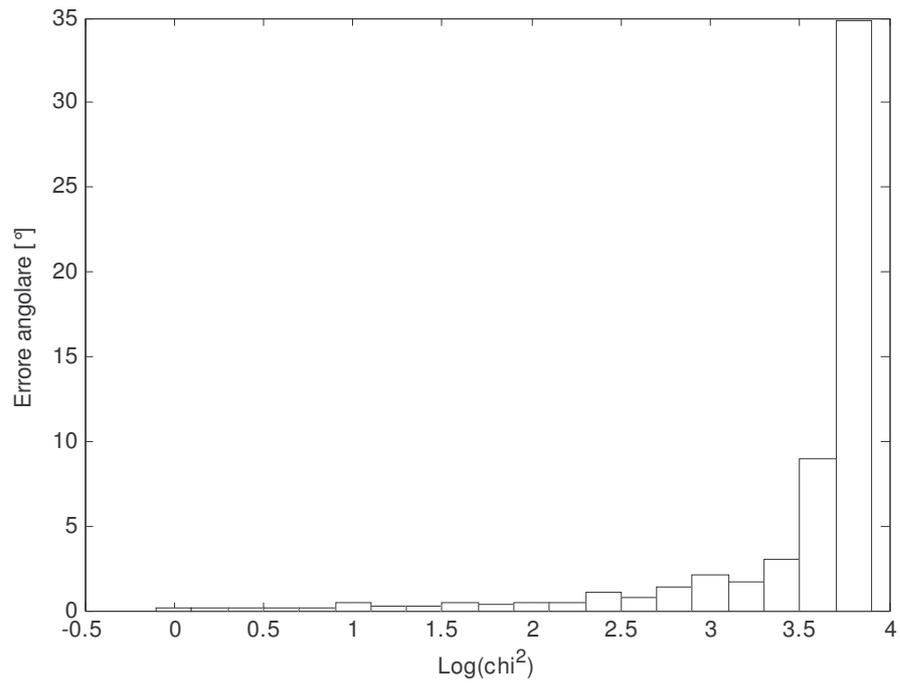


Figura 99 – Errore angolare (mediana) in funzione del χ^2 .

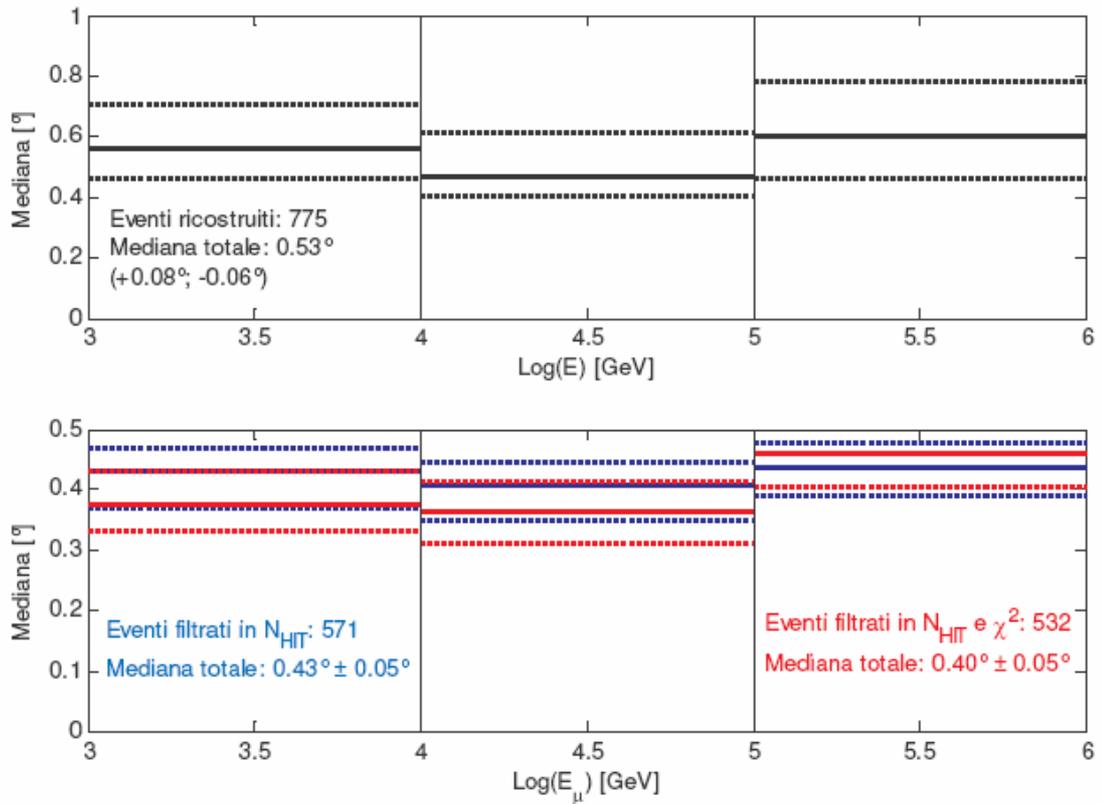


Figura 100 – Mediana dell'errore angolare in funzione dell'energia, prima e dopo la selezione.

La mediana totale dell'errore per tutti gli eventi ricostruiti ($N_{hit}>5$) è quindi $0.53^{\circ+0.08}_{-0.06}$.

Con la selezione sul numero di hit la mediana dell'errore totale è $0.43^{\circ} \pm 0.05$; con la selezione sul numero di hit e sul χ^2 la mediana dell'errore è $0.40^{\circ} \pm 0.05$.

Efficienza di ricostruzione

L'area efficace prima della selezione è pari a $A_{Eff} = A_G \cdot 775/1000 = 2.77 \text{ km}^2$; dopo la selezione l'area efficace risulta $A_{Eff} = A_G \cdot 532/1000 = 1.9 \text{ km}^2$, con una mediana dell'errore pari a $0.40^{\circ} \pm 0.05^{\circ}$.

In Figura 101 è riportata l'area efficace (rapportata all'area geometrica) al variare dell'energia.

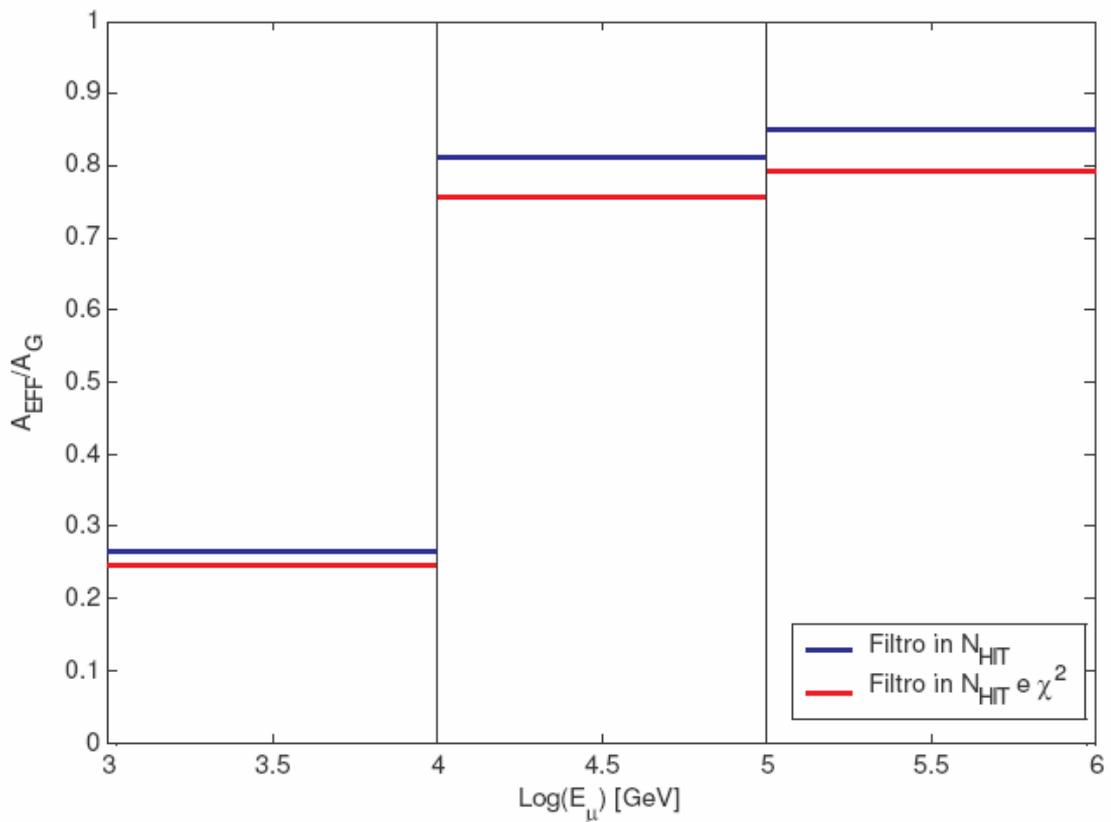


Figura 101 - Area efficace normalizzata all'area geometrica in funzione dell'energia.

Cap. 6 Conclusioni

Sono state implementate le ricostruzioni delle tracce mediante algoritmi genetici e reti neurali. Confrontando i risultati con quelli ottenuti tramite gli algoritmi di ricostruzione classici, già studiati e implementati nelle collaborazioni NEMO ed ANTARES, si osserva che le reti neurali danno prestazioni confrontabili con quelle del prefit, e possono essere utilizzate per un filtraggio veloce delle tracce up-going.

Gli algoritmi genetici permettono di raggiungere prestazioni molto simili a quelle della minimizzazione diretta dell'errore quadratico, anche se con un tempo di elaborazione più lungo di circa 50 volte. In Figura 102 e Figura 103 sono riportati i confronti della mediana di ricostruzione e dell'area efficace.

L'algoritmo genetico descritto nel Cap. 5 è più appropriato nello studio di sistemi fisici particolarmente complessi. Nel nostro caso viene usato per minimizzare una funzione (χ^2) abbastanza semplice. Il risultato ottenuto è comunque apprezzabile visto che, con un metodo totalmente indipendente dalle tecniche di correlazioni causali spazio-temporali riesce ad ottenere risoluzioni spaziali confrontabili.

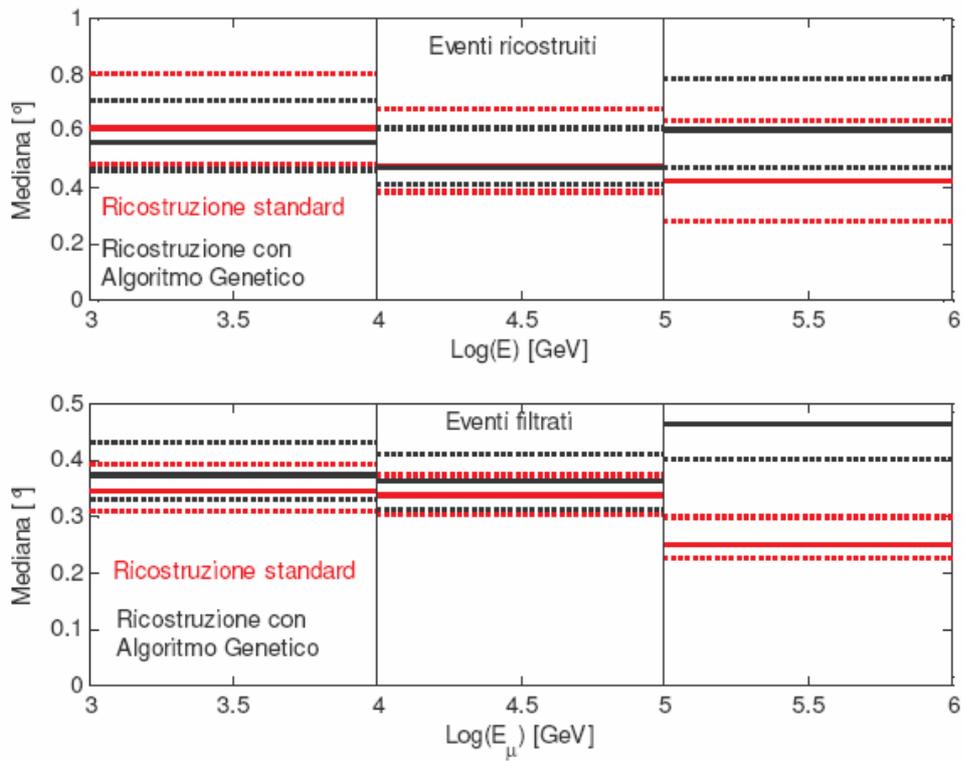


Figura 102 - Confronto tra le mediane di ricostruzione dell'algorithmo standard e dell'algorithmo genetico.

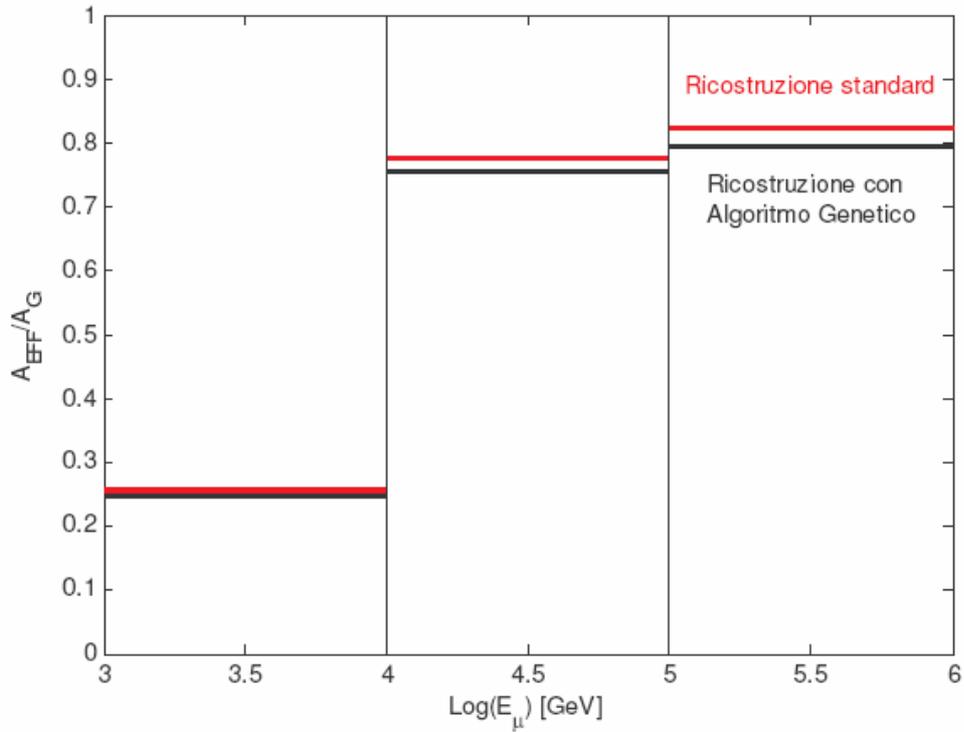


Figura 103 - Confronto tra le aree efficaci dell'algorithmo standard e dell'algorithmo genetico.

Appendice

Funzioni per l'implementazione dell'algoritmo genetico

```
subroutine trkfit(iflag)

integer N_POP, N_GEN
dimension afcn(6)

double precision tprimo,tultimo
common /forse/ IPMfirst,KPMfirst,tprimo,IPMlast,KPMlast,tultimo
*
include 'param81.inc'
include 'common_geom81.inc'
include 'common_fase181.inc'
include 'common_evread81new.inc'
include 'common_evdata81.inc'
include 'common_steer.inc'
include 'common_analysisnew.inc'

include 'paramGA.inc'

double precision fitness(N_MAX_POP)
double precision fit_temp(N_MAX_POP),best_fit(N_MAX_GEN)
double precision AAFCN(6), tempfit
double precision sigmar_init, sigmaa_init

double precision quX, quFI, quTH, DiamX, CentroX
double precision best_temp, best_mc

common/GAdata/IPOP(N_MAX_POP,5),NEW_POP(N_MAX_POP,5),
+ fitness
common/GAdata/fit_temp,best_fit
common/GAdata/igene(5),itempgen(5), AAFCN, tempfit
common/GAdata/sigmar_init, sigmaa_init, INIT_TYPE
common/GAdim/N_POP, N_GEN, quX, quFI, quTH, DiamX, CentroX
```

```

double precision rangte,rangfi,xbar,ybar,zbar
common /resultprefit/ rangte,rangfi,xbar,ybar,zbar
real RANDOM, x, y(N_MAX_POP), xi,yi,zi,fii,thi, pig

double precision X_OPT, Y_OPT, Z_OPT, FI_OPT, TH_OPT
double precision x0,y0,z0,thfit,phfit,r,cx,cy,cz
double precision argtest,arccos,tcosfit,tcosmc
double precision cxtrue,cytrue,cztrue,anglfp
double precision xx0, yy0, zz0

integer gene_mc(5)
logical CONT

```

```

*=====

```

```

N_POP=300
N_GEN=300
quX=0.025
quFI=0.0000958738
quTH=0.0000479369
DiamX=1638.4
CentroX=819.2
sigmar_init=200.0
sigmaa_init=0.1
pig = 3.141593
INIT_TYPE=1
CONT=.true.
best_mc=1000000000.0
imc=0
MC=1

```

```

if(iflag.eq.1) then

```

```

*-----1-----2-----3-----4-----5-----6-----7-*

```

```

*
*
*           IFLAG = 1
*
*

```

```

*-----1-----2-----3-----4-----5-----6-----7-*

```

```

*

```

```

elseif(iflag.eq.2) then

```

```

*-----1-----2-----3-----4-----5-----6-----7-*

```

```

*
*
*           IFLAG = 2
*
*

```

```

*-----1-----2-----3-----4-----5-----6-----7-*

```

```

*

```

```

nt=0
do i=1,npmhit
  if (flagPMT_HIT(i,3).eq.0) then
    nt=nt+1

```

```

endif
end do
if (nt.ge.5) then
do while (CONT)

if (nt.gt.200) then
MC=3
elseif (nt.gt.100) then
MC=5
elseif (nt.gt.30) then
MC=10
elseif (nt.gt.10) then
MC=15
elseif (nt.gt.5) then
MC=20
else
MC=0
endif

if (INIT_TYPE.eq.1) then
* inizializzazione pilotata dai risultati del prefit
cxpre = sin(rangte)*cos(rangfi)
cypre = sin(rangte)*sin(rangfi)
czpre = cos(rangte)
aa = 1.
bb = -2.*(cxpre*xbar + cypre*ybar + czpre*zbar)
cc = xbar**2 + ybar**2 + zbar**2 - CentroX**2
* utilizzo un raggio della sfera di 800m, più vicino al primo hit della traccia
test = bb**2 - 4.*aa*cc
if(test.lt.0.) then
solm = (-bb)/(2.*aa)
solp = (-bb)/(2.*aa)
else
solm = (-bb - sqrt(bb**2 - 4.*aa*cc))/(2.*aa)
solp = (-bb + sqrt(bb**2 - 4.*aa*cc))/(2.*aa)
endif

xx0 = xbar - cxpre * solp
yy0 = ybar - cypre * solp
zz0 = zbar - czpre * solp
*
do i=1,N_POP
* un punto random intorno al punto trovato sopra, con larghezza sigmar_init
CALL NORRAN(RANDOM)
xi=sigmar_init*RANDOM+xx0
CALL NORRAN(RANDOM)
yi=sigmar_init*RANDOM+yy0
CALL NORRAN(RANDOM)
zi=sigmar_init*RANDOM+zz0

```

```

* il punto non deve essere esterno al dominio di esistenza + o - CentroX
  xi=max(xi,-CentroX)
  yi=max(yi,-CentroX)
  zi=max(zi,-CentroX)
  xi=min(xi,CentroX)
  yi=min(yi,CentroX)
  zi=min(zi,CentroX)
* angoli random intorno agli angoli del prefit, con larghezza sigmaa_init (in radianti)
  CALL NORRAN(RANDOM)
  if (rangfi.lt.0.0) then
    fii=sigmaa_init*RANDOM+rangfi+2*pig
  else
    fii=sigmaa_init*RANDOM+rangfi
  end if
  CALL NORRAN(RANDOM)
  thi=sigmaa_init*RANDOM+rangte
* anche gli angoli devono essere limitati tra 0 e 360 fi e tra 0 e 180 teta
  fii=amax1(fii,0.0)
  fii=amin1(fii,2*pig)
  thi=amax1(thi,0.0)
  thi=amin1(thi,pig)
* POP contiene i geni codificati a 16 bit (tra 0 e 65535)
  IPOP(i,1)=32767*(xi+CentroX)/CentroX
  IPOP(i,2)=32767*(yi+CentroX)/CentroX
  IPOP(i,3)=32767*(zi+CentroX)/CentroX
  IPOP(i,4)=65535*fii/(2*pig)
  IPOP(i,5)=65535*thi/pig
  end do
  elseif (INIT_TYPE.eq.2) then
* inizializzazione random
  do i=1,N_POP
  call RANLUX(y,5)
  do j=1,5
    IPOP(i,j)=65535*y(j)
  end do
  end do
  elseif (INIT_TYPE.eq.3) then
* inizializzazione costante da un punto
  do i=1,N_POP
  IPOP(i,1)=32768
  IPOP(i,2)=32768
  IPOP(i,3)=0
  IPOP(i,4)=0
  IPOP(i,5)=0
  end do
  end if

  do igen=1,(N_GEN-1)

```

```

*-----

```

```

* ciclo sulle generazioni

    call FITNEMO
    best_temp=1000000000.0
    do ip=1,N_POP
        fitness(ip)=fit_temp(ip)
        if (fitness(ip).lt.best_temp) best_temp=fitness(ip)
    end do
    best_fit(igen)=best_temp
* calcola la fitness di tutti gli individui e trova il migliore
*-----

*-----

    call SELECTION
* trascrive gli individui migliori e seleziona quelli per il CO
*-----

*-----

    call CROSSOVER
* attua il CO
*-----

*-----

    call NEW_ENTRIES
* attua la MUTAZIONE
*-----

    do m=1,N_POP
        do k=1,5
            IPOP(m,k)=NEW_POP(m,k)
        end do
    end do

    end do
C fine ciclo sulle generazioni
C _____

*-----

* ULTIMA VALUTAZIONE e SCELTA DEL MIGLIORE
    call FITNEMO
    best_temp=1000000000.0
    do ip=1,N_POP
        fitness(ip)=fit_temp(ip)
        if (fitness(ip).lt.best_temp) then
            best_temp=fitness(ip)
            IND_MIN=ip
        endif
    end do

```

```

        best_fit(iigen+1)=best_temp
write(41,*) iev, best_temp
if (best_temp.lt.best_mc) then
    best_mc=best_temp
    do k=1,5
        gene_mc(k)=IPOP(IND_MIN,k)
    end do
endif

if (best_mc.eq.0) CONT=.false.
imc=imc+1
print *, 'imc ', imc
if (imc.ge.MC) CONT=.false.
end do

sumn=best_mc

        call DECODE_GENE(gene_mc,X_OPT,Y_OPT,Z_OPT,FI_OPT,TH_OPT)

        x0=X_OPT
        y0=Y_OPT
        z0=Z_OPT
        cx=sin(TH_OPT)*cos(FI_OPT)
        cy=sin(TH_OPT)*sin(FI_OPT)
        cz=cos(TH_OPT)
thfit=TH_OPT*180/pig
phfit=FI_OPT*180/pig
if (phfit.gt.180) then
    phfit=phfit-360.
endif
print *, 'Delta theta ', thfit, thmu, thfit-thmu
print *, 'Delta fi' , phfit, phimu, phfit-phimu

```

```

xmc0 = vxmu
ymc0 = vymu
zmc0 = vzmu
cxmc = pxmu/sqrt(pxmu**2+pymu**2+pzmu**2)
cymc = pymu/sqrt(pxmu**2+pymu**2+pzmu**2)
czmc = pzmu/sqrt(pxmu**2+pymu**2+pzmu**2)
cxtrue = dble( cxmc )
cytrue = dble( cymc )
cztrue = dble( czmc )
anglfp = 2.*pig
if((cx**2+cy**2+cz**2).ne.dble( 0.))
+ and.(cxtrue**2+cytrue**2+cztrue**2).ne.dble( 0.)) then
    if((dabs(cx*cxtrue+cy*cytrue+cz*cztrue)).lt.dble(1.01))then
        if( (cx*cxtrue+cy*cytrue+cz*cztrue) .gt.dble(1.01))
+         anglfp = dacos(dble(1.))

```

```

        if( (cx*cxtrue+cy*cytrue+cz*cztrue) .lt.dble(-1.01))
+       anglfp = dacos(dble(-1.))
    else
        write(16,*)' cannot evaluate anglfp !!!',
,         cx,cy,cz,cxtrue,cytrue,cztrue
        print *, ' attention !, cx,cy,cz,cxmc,cymc,czmc',
,         cx,cy,cz,cxmc,cymc,czmc
        print *, ' and the argument of acos :',
,         cx*cxtrue+cy*cytrue+cz*cztrue
        endif
    endif
    argtest = cx*cxtrue+cy*cytrue+cz*cztrue
    argcos = ( cx*cxtrue + cy*cytrue + cz*cztrue )/
+     (dsqrt(cx**2+cy**2+cz**2)*
+     dsqrt(cxtrue**2+cytrue**2+cztrue**2))
    tcosfit = cx**2+cy**2+cz**2
    tcosmc = cxtrue**2+cytrue**2+cztrue**2
    call hfill(3010,sngl(argtest),0.,1.)
    call hfill(3011,sngl(argcos ),0.,1.)
    call hfill(3012,sngl(tcosfit),0.,1.)
    call hfill(3013,sngl(tcosmc ),0.,1.)
    if((dabs(cx*cxtrue+cy*cytrue+cz*cztrue)).le.dble(1.))then
        anglfp = dacos( cx*cxtrue + cy*cytrue + cz*cztrue )/
+     (dsqrt(cx**2+cy**2+cz**2)*
+     dsqrt(cxtrue**2+cytrue**2+cztrue**2))
        if(iprivi.ge.3) then
            print *, ' in FCNTRK cx,cy,cz from fit and from MC',
+             cx,cy,cz,cxmc,cymc,czmc
        endif
    endif

    write(16,*) sngl(x0),sngl(y0),sngl(z0),sngl(thfit),
+ sngl(phfit),thmu,phimu,anglfp,sumn
    thepre = 180.*rangte/pig
    phipre = 180.*rangfi/pig
    write(16,2002) thepre,phipre
2002 format(' theta, phi prefit',2f10.2)
    deltang = sngl(anglfp)*180./pig
    emul = log10(emu/1000.)

* chiamate per la gestione degli istogrammi
    call hfill(3000,sngl(thfit )-thmu ,0.,1.)
* .....
* .....
* .....

    elseif(iflag.eq.3) then
*-----1-----2-----3-----4-----5-----6-----7-*

```

```

*
*           IFLAG = 3
*
*-----1-----2-----3-----4-----5-----6-----7-*
*
    elseif(iflag.eq.4) then
*-----1-----2-----3-----4-----5-----6-----7-*
*
*           IFLAG = 4
*
*-----1-----2-----3-----4-----5-----6-----7-*
*
    endif
    return
30 FORMAT(A30)
    end

*-----1-----2-----3-----4-----5-----6-----7-*
    subroutine DECODE_GENE(igene_p, XDEC, YDEC, ZDEC, FIDEC, THDEC)
* dato un gene creo i valori x, y, z, fi e theta corrispondenti
* i valori di distanza sono in metri, i valori angolari sono in radianti
    implicit double precision(a-h,o-z)
    integer igene_p(5)
    include 'paramGA.inc'
    include 'common_GAdim.inc'

    common/GAdata/IPOP(N_MAX_POP,5),NEW_POP(N_MAX_POP,5),
+ fitness(N_MAX_POP)
    common/GAdata/fit_temp(N_MAX_POP),best_fit(N_MAX_GEN)
    common/GAdata/igene(5),itempgen(5), AAFCN(6), tempfit
    common/GAdata/sigmar_init, sigmaa_init, INIT_TYPE

    XDEC=igene_p(1)*quX-CentroX
    YDEC=igene_p(2)*quX-CentroX
    ZDEC=igene_p(3)*quX-CentroX
    FIDEC=igene_p(4)*quFI
    THDEC=igene_p(5)*quTH

    end

*-----1-----2-----3-----4-----5-----6-----7-*
    subroutine FITNEMO
* calcola l'errore quadratico dei tempi delle tracce codificate nei geni della popolazione

    include 'param81.inc'
    include 'paramGA.inc'

    include 'common_geom81.inc'
    include 'common_fase181.inc'

```

```

include 'common_evread81new.inc'
include 'common_evdata81.inc'
include 'common_steer.inc'
include 'common_analysisnew.inc'

double precision fitness(N_MAX_POP)
double precision fit_temp(N_MAX_POP),best_fit(N_MAX_GEN)
double precision AAFCN(6), tempfit
double precision sigmar_init, sigmaa_init
double precision xx, yy, zz, ffi, tth
double precision quX, quFI, quTH, DiamX, CentroX
double precision t_th, s, ST, out, tme, delt

common/GAdata/IPOP(N_MAX_POP,5),NEW_POP(N_MAX_POP,5),
+ fitness
common/GAdata/fit_temp,best_fit
common/GAdata/igene(5),itempgen(5), AAFCN, tempfit
common/GAdata/sigmar_init, sigmaa_init, INIT_TYPE
common/GAdim/N_POP, N_GEN, quX, quFI, quTH, DiamX, CentroX

nt=0
do i=1,npmhith
  if (flagPMT_HIT(i,3).eq.0) nt=nt+1
end do
do i=1,N_POP
  do k=1,5
    igene(k)=IPOP(i,k)
  end do
  call DECODE_GENE(igene,xx,yy,zz,ffi,tth)
  AAFCN(1)=xx
  AAFCN(2)=yy
  AAFCN(3)=zz
  AAFCN(4)=cos(ffi)*sin(tth)
  AAFCN(5)=sin(ffi)*sin(tth)
  AAFCN(6)=cos(tth)

  out=0
  if (nt.ge.5) then

    do jj=1,npmhith

      if (flagPMT_HIT(jj,3).eq.0) then
        ip=idPMT_HIT(jj)
        call tpmcal(ip,t_th,AAFCN,s)
        ST=dbl(pmtdata(ip,8))

        tme=timePMT_HIT(jj,2)

        delt=tme-t_th
      end if
    end do
  end if
end do

```

```

        out=out+delt**2/ST**2

    endif
  end do
  fit_temp(i)=log10(out/nt)

  if (fit_temp(i).lt.0) fit_temp(i)=0
else
  fit_temp(i)=1000000000.0
endif
end do

end

*-----1-----2-----3-----4-----5-----6-----7-*
subroutine SELECTION
* copia nella NEW_POP i migliori (N_BEST) e seleziona con il metodo
* della roulette wheel quelli destinati al CrossOver
  implicit double precision(a-h,o-z)
  include 'paramGA.inc'
  include 'common_GAdim.inc'
  dimension temp_sel(N_MAX_POP)

  real RANDOM
  common/GAdata/IPOP(N_MAX_POP,5),NEW_POP(N_MAX_POP,5),
+ fitness(N_MAX_POP)
  common/GAdata/fit_temp(N_MAX_POP),best_fit(N_MAX_GEN)
  common/GAdata/igene(5),itempgen(5), AAFCN(6), tempfit
  common/GAdata/sigmar_init, sigmaa_init, INIT_TYPE
  real x, y(N_MAX_POP)
  double precision MIN_SEL
  newe=1
  N_NEW=int(N_POP/10)+1
  N_BEST=int(N_POP/30)+1
  if (newe.gt.0) then
    N_CO_S=N_POP-N_BEST-N_NEW
  else
    N_CO_S=N_POP-N_BEST
  endif

  do j=1,N_POP
    temp_sel(j)=fitness(j)
  end do

  do i=1,N_BEST
    tempfit=1000000000.0
    IND_MIN=0
    do j=1,N_POP
      if (temp_sel(j).lt.tempfit) then

```

```

        tempfit=temp_sel(j)
        IND_MIN=j
    end if
end do
do k=1,5
    NEW_POP(i,k)=IPOP(IND_MIN,k)
    temp_sel(IND_MIN)=1000000000.0
end do
end do
do i=1,N_CO_S
    call RANLUX(y,N_POP)
    do k=1,N_POP
        temp_sel(k)=fitness(k)*y(k)
    end do
    MIN_SEL=1000000000.0
    IND_MIN=0
    do k=1,N_POP
        if (temp_sel(k).lt.MIN_SEL) then
            MIN_SEL=temp_sel(k)
            IND_MIN=k
        endif
    end do
    do k=1,5
        NEW_POP(N_BEST+i,k)=IPOP(IND_MIN,k)
    end do
end do
end

```

-----1-----2-----3-----4-----5-----6-----7-

subroutine CROSSOVER

* procedura che implementa il CROSS OVER a due a due sui geni selezionati

implicit double precision(a-h,o-z)

include 'paramGA.inc'

include 'common_GAdim.inc'

common/GAdata/IPOP(N_MAX_POP,5),NEW_POP(N_MAX_POP,5),

+ fitness(N_MAX_POP)

common/GAdata/fit_temp(N_MAX_POP),best_fit(N_MAX_GEN)

common/GAdata/igene(5),itempgen(5), AAFCN(6), tempfit

common/GAdata/sigmar_init, sigmaa_init, INIT_TYPE

integer gene1(5), gene2(5), newgene1(5), newgene2(5)

integer jj, CO_TYPE, CUT_POINT, INDW_CUT, INDK_CUT

integer w11, w12, w21, w22, w(5), wn(5)

double precision cromo1, cromo2, ncrom1, ncrom2

real x,y(N_MAX_POP)

newe=1

cop=1

```

call RANLUX(x,1)
if (x.lt.cop) then

N_NEW=int(N_POP/10)+1
N_BEST=int(N_POP/30)+1
if (newe.gt.0) then
  N_CO=N_POP-N_NEW
else
  N_CO=N_POP
endif

do i=(N_BEST+1),N_CO,2
  do k=1,5
    gene1(k)=NEW_POP(i,k)
    gene2(k)=NEW_POP(i+1,k)
  end do

  CO_TYPE=2
  if (CO_TYPE.eq.1) then
* Crossover ad un taglio
  call RANLUX(y,1)
  CUT_POINT=NINT(y(1)*78+1)
* punto di taglio random da 1 a 79 (lunghezza gene 80)
  INDW_CUT=INT(CUT_POINT/16)+1
* parola selezionata per il taglio
  do jj=1,INDW_CUT-1
    newgene1(jj)=gene1(jj)
    newgene2(jj)=gene2(jj)
  end do
  do jj=INDW_CUT+1,5
    newgene1(jj)=gene2(jj)
    newgene2(jj)=gene1(jj)
  end do
* ho copiato le parole non tagliate, quelle prima del taglio in ordine,
* quelle dopo il taglio in ordine inverso
  INDK_CUT=mod(CUT_POINT,16)
* print *, 'cut ', INDW_CUT, INDK_CUT
* taglio all'interno della parola selezionata
  MASK1=0
* parte più significativa
  do kk=0,INDK_CUT
    MASK1=MASK1+2**(15-kk)
  end do
  MASK2=2**16-1-MASK1
* parte meno significativa
  newgene1(INDW_CUT)=IOR(IAND(gene1(INDW_CUT),MASK1),
+ IAND(gene2(INDW_CUT),MASK2))
  newgene2(INDW_CUT)=IOR(IAND(gene1(INDW_CUT),MASK2),
+ IAND(gene2(INDW_CUT),MASK1))
  do k=1,5

```

```

        NEW_POP(i,k)=newgene1(k)
        NEW_POP(i+1,k)=newgene2(k)
    end do
    elseif (CO_TYPE.eq.2) then
*   crossover2 è con due punti di taglio.
*   g1 e g2 sono i due geni che partecipano al cross over
*   sono due vettori riga di 5 elementi x y z fi teta
*   ogni elemento è un uint16

*   due indici random (1..62; ind1+16..79) danno i punti di taglio
    call RANLUX(y,2)
    ind1=anint(61*y(1))+1
    ind2=anint((63-ind1)*y(2))+ind1+16
    k1=int(ind1/16)+1
    k2=int(ind2/16)+1

    do jj=1,(k1-1)
        newgene1(jj)=gene1(jj)
        newgene2(jj)=gene2(jj)
    end do
    do jj=(k1+1),(k2-1)
        newgene1(jj)=gene2(jj)
        newgene2(jj)=gene1(jj)
    end do
    do jj=(k2+1),5
        newgene1(jj)=gene1(jj)
        newgene2(jj)=gene2(jj)
    end do
    indk1=mod(ind1,16)
    indk2=mod(ind2,16)
*   indk va da 0 a 15
*   -----
        MASK11=0
*   % parte più significativa
    do kk=0,indk1
        MASK11=MASK11+2**(15-kk)
    end do
    MASK21=2**16-1-MASK11
*   % parte meno significativa
    newgene1(k1)=IOR(IAND(gene1(k1),MASK11),
+   IAND(gene2(k1),MASK21))
    newgene2(k1)=IOR(IAND(gene1(k1),MASK21),
+   IAND(gene2(k1),MASK11))
    MASK12=0
*   % parte più significativa
    do kk=0,indk2
        MASK12=MASK12+2**(15-kk);
    end do
    MASK22=2**16-1-MASK12

```

```

* % parte meno significativa
    newgene1(k2)=IOR(IAND(gene1(k2),MASK22),
+   IAND(gene2(k2),MASK12))
    newgene2(k2)=IOR(IAND(gene1(k2),MASK12),
+   IAND(gene2(k2),MASK22))
    do k=1,5
        NEW_POP(i,k)=newgene1(k)
        NEW_POP(i+1,k)=newgene2(k)
    end do
    else if (CO_TYPE.eq.3) then
* % crossover_u è il crossover uniforme
* % cioè ogni bit di g1 passa a g2 con probabilità 50%
* % g1 e g2 sono i due geni che partecipano al cross over
* % sono due vettori riga di 5 elementi x y z fi teta
* % ogni elemento è un uint16

        call RANLUX(y,5)
        do k=1,5
            w(k)=int(65535*y(k))
            wn(k)=65536-w(k)
            newgene1(k)=IOR(IAND(gene1(k),wn(k)),
+   IAND(gene2(k),w(k)))
            newgene2(k)=IOR(IAND(gene1(k),w(k)),
+   IAND(gene2(k),wn(k)))
            NEW_POP(i,k)=newgene1(k)
            NEW_POP(i+1,k)=newgene2(k)
        end do
        else if (CO_TYPE.eq.4) then
*   % crossover4 è il crossover generalizzato Coli-Palazzari
*% ma agisce su due parole da 16bit
*% g1 e g2 sono i due geni che partecipano al cross over
*% sono due vettori riga di 5 elementi x y z fi teta
*% ogni elemento è un uint16

*% parametri del CO
        b=1.05;
        M=int((31*log10(2.0))/log10(b))+1

*% parole su cui si opera
        call RANLUX(y,1)
        j=aint(3*y(1))+1
*% da 1 a 4, perché prende 2 parole
*% cromo sono parole da 32 bit
        cromo1=gene1(j)*65536.0+gene1(j+1)
        cromo2=gene2(j)*65536.0+gene2(j+1)

        call RANLUX(y,1)
        k=aint((M-1)*y(1))+1
        cr=b**k

```

```

r1=dmod(cromo1,cr)
r2=dmod(cromo2,cr)
ncrom1=cromo1-r1+r2
ncrom2=cromo2-r2+r1
aa1=ncrom1/65536
aa2=ncrom2/65536

w11=anint(aa1)
w12=mod(ncrom1,65536)
w21=anint(aa2)
w22=mod(ncrom2,65536)

do jj=1,5
  newgene1(jj)=gene1(jj)
  newgene2(jj)=gene2(jj)
end do

newgene1(j)=w11
newgene1(j+1)=w12
newgene2(j)=w21
newgene2(j+1)=w22

      do jj=1,5
        NEW_POP(i,jj)=newgene1(jj)
        NEW_POP(i+1,jj)=newgene2(jj)
      end do
    end if
  end do
endif
end

*-----1-----2-----3-----4-----5-----6-----7-*
subroutine NEW_ENTRIES
* procedura che genera nuovi arrivi nella popolazione
* sostituisce la MUTAZIONE
* implementa se la flag apposita è alzata, la eliminazione dei cloni
  implicit double precision(a-h,o-z)
  include 'paramGA.inc'
  include 'common_GAdim.inc'

  common/GAdata/IPOP(N_MAX_POP,5),NEW_POP(N_MAX_POP,5),
+ fitness(N_MAX_POP)
  common/GAdata/fit_temp(N_MAX_POP),best_fit(N_MAX_GEN)
  common/GAdata/igene(5),itempgen(5), AAFCN(6), tempfit
  common/GAdata/sigmar_init, sigmaa_init, INIT_TYPE

  real x, y(5)
  logical NO_CLONE, equa
  real PMUT, pmutemp(N_MAX_POP), indtemp(2)

```

```

integer wordmut, wordgen, indmuw, indmub
newe=1
N_NEW=int(N_POP/10)+1
NO_CLONE=.false.
PMUT=0.02

if (NO_CLONE) then
  do i=1,(N_POP-1)
    do j=(i+1),N_POP
      equa=.true.
      do k=1,5
        if (NEW_POP(i,k).ne.NEW_POP(j,k)) equa=.false.
      end do
      if (equa) then
        call RANLUX(y,5)
        do k=1,5
          NEW_POP(j,k)=65535*y(k)
        end do
      end if
    end do
  end do
endif

if (newe.gt.0) then
  N_BEST=int(N_POP/30)+1
  N_CO=N_POP-N_NEW
  do i=(N_CO+1),N_POP
    call RANLUX(y,5)
    do j=1,5
      NEW_POP(i,j)=65535*y(j)
    end do
  end do
endif

if (pmut.gt.0.0)then
  call RANLUX(pmutemp,N_POP)
  do i=2,N_POP
    if (pmutemp(i).lt.PMUT) then
      call RANLUX(indtemp,2)
      indmuw=nint(4*indtemp(1))+1
      indmub=nint(15*indtemp(2))+1
      wordgen=NEW_POP(i,indmuw)
      wordmut=2*(indmub-1)
      NEW_POP(i,indmuw)=IEOR(wordgen,wordmut)
    end if
  end do
end if
end

```

Funzioni per l'implementazione delle reti neurali

Genera i file di ingresso alla rete neurale, partendo dai file degli eventi

```
load GEOM
n=1.355;
th_C=acos(1/n);
c=0.29979;

load NH_45BF
load GEOM

for k_nhh=1:1000
    if NH_45BF(k_nhh)>19
        evnum=k_nhh;
        eval(['load data/ev_' num2str(evnum)]);
        eval(['load data/T_ev_' num2str(evnum) 'BF']);
        eval(['load data/PMT_ev_' num2str(evnum) 'BF']);
        eval(['ev=ev_' num2str(evnum) ';']);
        eval(['Tempi=T_ev_' num2str(evnum) 'BF;']);
        eval(['PMT=PMT_ev_' num2str(evnum) 'BF;']);

        fiv=ev(4);
        thv=ev(5);

        ind_ev=1;
        P=zeros(80,1);
        T=zeros(2,1);
        N_PMT=length(Tempi);
        n_ev_ev=N_PMT-19;
        for jk=1:n_ev_ev
            for jkk=0:19
                P(4*jkk+1,ind_ev)=(GEOM(PMT(jk+jkk),1)+800)/1600;
                P(4*jkk+2,ind_ev)=(GEOM(PMT(jk+jkk),2)+800)/1600;
                P(4*jkk+3,ind_ev)=(GEOM(PMT(jk+jkk),3)+800)/1600;
                P(4*jkk+4,ind_ev)=(Tempi(jk+jkk)-min(Tempi))/10000;
            end
            T(:,ind_ev)=[fiv/360; thv/180];
            ind_ev=ind_ev+1;
        end
        eval(['P_' num2str(evnum) 'BF=P; save P_' num2str(evnum) 'BF P_' num2str(evnum) 'BF;']);
        eval(['T_' num2str(evnum) 'BF=T; save T_' num2str(evnum) 'BF T_' num2str(evnum) 'BF;']);

    end
end
```

Genera i file di addestramento della rete

```

load NH_45BF
load P_1BF
load T_1BF
Pt=P_1BF;
Tt=T_1BF;

for i=2:500
    if NH_45BF(i)>19
        eval(['load P_' num2str(i) 'BF; load T_' num2str(i) 'BF;'])
        eval(['newP=P_' num2str(i) 'BF; newT=T_' num2str(i) 'BF;'])
        Pt=[Pt newP];
        Tt=[Tt newT];
    end
end
end

```

Addestramento della rete

```

clear
R2=(ones(80,1)*[0 1]);
SIZ=[25 2];
net=newff(R2, SIZ, {'tansig','satlin'}, 'trainscg');

```

```

load Pt
load Tt
P=Pt;
T=Tt;

```

```

net.trainParam.epochs=50000;
[net, TR, YY, EE]=train(net,P,T);

```

Test della rete

```

clear
load NH_45BF

load net45_20
netw=net45_20;

```

```

k=1;
for n=501:1000
    if NH_45BF(n)>19
        evnum=n;
        eval(['load P_' num2str(evnum) 'BF; P=P_' num2str(evnum) 'BF;']);
        eval(['load T_' num2str(evnum) 'BF; T=T_' num2str(evnum) 'BF;']);
        Y=sim(netw, P);
        E=Y-T;
    end
end

```

Bibliografia

- [1] M.A.Markov, I.M.Zheleznykh, Nucl. Phys., **27**, p. 385, (1961)
- [2] J.G.Learned, K.Mannheim, High-energy neutrino astrophysics, Annual Review Nucl. Part. Sci., 2000
- [3] M.Longair, High Energy Astrophysics, Cambridge University Press, (1997)
- [4] http://www.ngdc.noaa.gov/stp/SOLAR/COSMIC_RAYS/cosmic.html
- [5] http://it.wikipedia.org/wiki/Radiazione_cosmica
- [6] Gaisser, Cosmic rays and particle physics, Cambridge University Press (1990)
- [7] Perkins, Particle Astrophysics, Oxford University Press, (2003)
- [8] AUGER collaboration, Proc. of Rome International Conference on Cosmic Rays and Particle Astrophysics, 2007
- [9] Anne Hoppelt, Étude de la résolution angulaire du télescope à neutrinos ANTARES, Thèse de doctorat, Université de la Méditerranée, 2001
- [10] V.L.Ginzburg, S.I.Syrovatskii, Origin of Cosmic Rays, (1964)
- [11] R.J.Protheroe, Origin and propagation of the highest energy cosmic rays
<http://it.arxiv.org/abs/astro-ph/9612212v1>
- [12] Kamiokande Collaboration, Phys. Rev. Lett. **58**, p. 1490, (1987)
- [13] Manuela Vecchi, Studio dell'effetto del biofouling sull'efficienza di rivelazione di un apparato Čerenkov per neutrini astrofisici di alta energia, Tesi di Laurea, 2006
- [14] T.K.Gaisser, F.Halzen, T. Stanev, Particle astrophysics with high energy neutrinos, Physics Reports **258**, 1995
- [15] F. Halzen, High energy neutrino astronomy, astro-ph/0402083 v1 (2004)
- [16] H.E.S.S.Collaboration, Nature, **432**, p. 75, (2004)
- [17] H.E.S.S.Collaboration, Astroph. J., **619**, p. L113, (2005)

- [18] K.Mannheim, *Astron. & Astroph.*, **69**, p. 67, (1993)
- [19] L.V.Volkova, *Sov. J. Nul. Phys.*, **31**, p.6, (1980)
- [20] ANTARES Technical Design Report
- [21] W.E.Burcham, M.Jobes, *Nuclear and particle physics*, Prentice Hall 1995
- [22] D.J.L.Bailey, Genhen v5r1, ANTARES-SOFT, 2002-004
- [23] William R.Leo, *Techniques for Nuclear and particle physics experiments*, Springer-Verlag 1993
- [24] Adrian Jacob Heijboer, *Track reconstruction and point source searches with ANTARES*, PHD thesis, 2004
- [25] P.Antonioli et al, *A three-dimensional code for muon propagation through the rock: MUSIC*, hep-ph/9705408v2, Jul 1997
- [26] Francesco Simeone, *Caratterizzazione del rumore di fondo ottico in un apparato sottomarino per la rivelazione di neutrini*, Tesi di Laurea, 2002
- [27] P.Lipari, *Introduction to neutrino physics*, CERN-CLAF School of Physics, Itacuruca, Brazil, May 2001
- [28] T.Chiarusi, *Environmental parameters and water characteristics from deep-sea surveys at the NEMO sites*, 29th International Conference on Cosmic Rays, Pune (India), 2005
- [29] R.Gandhi, C.Quigg, M.H.Reno, I.Sarcevic, *Ultrahigh-energy neutrino interactions*, *Astroparticle Physics* 5, 1996
- [30] T.K.Gaisser, *Neutrino astronomy: physics goals, detector parameters*, OECD Megascience Forum Workshop, Taormina, Maggio 1997, astro-ph/9707283v1
- [31] F.Lucarelli, *Rivelazione di neutrini di alta energia da sorgenti puntiformi extragalattiche con l'apparato Čerenkov sottomarino Nestor*, Tesi di Laurea, Università di Roma, La Sapienza, 1998
- [32] <http://antares.in2p3.fr/Collaboration/>
- [33] <http://nemo11.roma1.infn.it/AstroNeut/KM3NeT/KM3NeT.htm>
- [34] Giulia De Bonis, *Studio delle tecniche di rivelazione acustica di neutrini astrofisici di alta energia*, Tesi di laurea, 2005
- [35] F.Halzen, D.Hooper, *High energy neutrino astronomy: the cosmic ray connection*, astro-ph/0204527 v2 Jul 2002
- [36] Amanda Collaboration, *Recent results for the AMANDA-II neutrino telescope*, astro-ph/0505278 v1 May 2005

- [37] A.Groß, astro-ph/00505278v1, “40me rencontre de Moriond” La Thuile, March 2005
- [38] J.A.Freeman, D.M.Skapura; Neural Networks: algorithms, applications and programming techniques; Addison-Wesley, 1991
- [39] C.Christodoulou, M.Georgiopoulos; Application of Neural Network in Electromagnetics, 2001
- [40] M.L.Minsky, S.A.Papert; Perceptrons, 1969
- [41] C.Cybenko; Approximation by superpositions of Sigmoidal Functions, Math. Contr., 1989
- [42] Rumelhart, Hinton, Williams; "pdp book", 1986
- [43] E.Granger, M.Rubin, S.Grossberg,P.Lavoie; Classification of incomplete data using the fuzzy ARTMAP neural network, Jan 2000
- [44] Manuale MatLAB, NNET toolbox
- [45] Darrell Whitley, A Genetic Algorithm Tutorial, Statistics and Computing 4, 1994.
- [46] J.Holland, Adaptation In Natural and Artificial Systems, Un. of Michigan Press, 1975
- [47] L.D.Davis, Handbook of Genetic Algorithms. Van Nostrand Reinhold, 1991
- [48] D.Goldberg, Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley, 1989
- [49] A.J. Heijboer ANT-SOFT-2002-002 An algorithm for track reconstruction in ANTARES, ANTARES Internal Note
- [50] Y Becherini, N Cottini, T.Stolarczyk ANT SOFT 2006 003
- [51] De Jong ANT SOFT 2007 001
- [52] ANTARES-SOFT-2004-001
- [53] A.Capone, V.Valente, Possibili trigger in NEMO, Nota interna NEMO, 2001
- [54] V.Valente, Trigger, fondo e ampiezza di segnale, Nota interna NEMO, 2003
- [55] F.James, MINUIT Function Minimization and Error Analysis Reference Manual
- [56] V.Valente, Un prefit possibile, Nota interna NEMO, 2004
- [57] ANT SOFT 2004 007 Filter
- [58] Carmona, ANT SOFT 2001 006, Prefit
- [59] Carmona, Hernandez - ANT SOFT 2000 011, Ricostruzione