

Numero XIX
Anno 2022


DIPARTIMENTO DI MATEMATICA E FISICA
"Ennio De Giorgi"


UNIVERSITÀ
DEL SALENTO

Viaggio nella Scienza

Ithaca



Raggi Cosmici

Ithaca: Viaggio nella Scienza

Una pubblicazione del Dipartimento di Matematica e Fisica “Ennio De Giorgi” dell’Università del Salento.

Registrazione presso il Tribunale di Lecce n. 6 del 30 Aprile 2013.
e-ISSN: 2282-8079

Direttore Responsabile
Luigi Spedicato.

Ideatore
Giampaolo Co’.

Comitato di Redazione
Adriano Barra,
Rocco Chirivì,
Paolo Ciafaloni,
Maria Luisa De Giorgi,
Vincenzo Flaminio,
Luigi Martina,
Giuseppe Maruccio,
Marco Mazzeo,
Francesco Paparella,
Carlo Sempi.

Segreteria di Redazione
Daniela Dell’Anna.

© 2013-2023 Dipartimento di Matematica e Fisica “*Ennio de Giorgi*”.

© 2023 per i singoli articoli dei rispettivi autori.

Il materiale di questa pubblicazione può essere riprodotto nei limiti stabiliti dalla licenza
“*Creative Commons Attribuzione – Condividi allo stesso modo 3.0 Italia*” (CC BY-SA 3.0 IT).

Per il testo della licenza: <http://creativecommons.org/licenses/by-sa/3.0/it/deed.it>

Ithaca: Viaggio nella Scienza
è disponibile sul sito:

<http://ithaca.unisalento.it/>

Scriveteci all’indirizzo:

ithaca@unisalento.it

Ithaca

Viaggio nella Scienza

XIX 2022

- 3** In questo numero
- 5** I Raggi Cosmici e l'universo delle alte energie
Paolo Lipari
- 35** Raggi cosmici, ieri, oggi, domani
Alessandro De Angelis
- 51** Caratteristiche principali degli esperimenti per la
misura dei raggi cosmici da terra
Andrea Chiavassa
- 63** Osservazione dei raggi cosmici dallo spazio
Mirko Boezio, Riccardo Munini
- 79** UHECR: raggi cosmici di energia ultra elevata
Sergio Petrerá, Francesco Salamida
- 93** Come sono accelerati i raggi cosmici?
Stefano Gabici

103 L'importanza dei neutrini cosmici di altissima energia

Carla Distefano, Francesco Vissani

115 Astroparticelle di altissima energia

Antonella Castellina

131 Sul tetto del mondo per studiare la radiazione cosmica

Paolo Bernardini

Fisica, Matematica e Medicina

137 Machine Learning

Giorgio De Nunzio

Sistemi complessi

169 Il valore dell'informazione

Luca Peliti

In questo numero

Lo studio dei raggi cosmici è l'argomento principale di questo numero di **Ithaca**.

Nel suo articolo Paolo Lipari presenta un'ampia panoramica della fisica legata allo studio dei raggi cosmici: rivelazione, composizione, meccanismi di accelerazione, loro uso per lo studio delle sorgenti astrofisiche che li producono, rilevanza del loro studio nell'ambito della fisica delle alte energie, nucleare, astrofisica e cosmologia. Si tratta di un compendio degli argomenti che saranno discussi in maggiore dettaglio negli articoli successivi, inquadrati nel contesto più ampio delle ricerche di base.

L'articolo di Alessandro De Angelis presenta lo studio dei raggi cosmici in una visione storica che parte dal passato, descrive gli esperimenti moderni ed offre una visione sui progetti futuri.

I vari tipi di esperimenti che permettono l'identificazione dei raggi cosmici sono presentati nell'articolo di Andrea Chiavassa e in quello di Mirko Boezio e Riccardo Munini. Nel primo articolo si descrivono le caratteristiche degli esperimenti che si effettuano sulla superficie terrestre. Nel secondo articolo, invece, si parla degli esperimenti che si svolgono nello spazio.

Sergio Petrera e Francesco Salamida presentano gli studi riguardanti la composizione dei raggi cosmici, e si concentrano, soprattutto, sulle caratteristiche di quei raggi cosmici che hanno l'energia più elevata che siamo riusciti a misurare.

Come possono essere accelerati raggi cosmici di questa energia? È la domanda che affronta Stefano Gabici nel suo articolo in cui discute dei possibili meccanismi proposti per spiegare le altissime energie che raggiungono alcuni raggi cosmici.

L'importanza, ed il ruolo, dei neutrini di al-

tissima energia contenuti nei raggi cosmici sono discussi nell'articolo di Carla Distefano e Francesco Vissani.

Antonella Castellina descrive in dettaglio l'esperimento Auger, l'esperimento con la maggiore estensione superficiale mai costruito. Un enorme sforzo tecnologico, logistico e organizzativo che ha permesso di identificare eventi ad altissime energie, irraggiungibili nei laboratori moderni, ma anche in quelli che potrebbero essere costruiti in futuro con le attuali tecnologie.

Un esempio di cosa implichi, sul campo, lo studio e l'osservazione dei raggi cosmici è fornito dall'articolo di Paolo Bernardini che presenta, usando anche dell'aneddotica, l'esperimento ARGO-YBJ che ha terminato la presa dati nel 2013.

Gli ultimi due articoli di questo numero XIX sono legati ad argomenti presentati in altri numeri di **Ithaca**.

Giorgio De Nunzio fa un'ampia presentazione di come si possano utilizzare tecniche legate all'intelligenza artificiale per l'analisi di dati diagnostici nella medicina. Un articolo in linea con i temi trattati nel numero XVIII. Nell'articolo di Luca Peliti si parla del concetto statistico di informazione e del suo ruolo nella termodinamica e nella fisica dei sistemi complessi, e utilizza come sistema semplice da investigare quello del gioco d'azzardo. Questo tema è ben legato a quelli trattati nel numero XVI.

Buona lettura,
il Comitato di Redazione

I Raggi Cosmici e l'universo delle alte energie

Paolo Lipari *Istituto Nazionale di Fisica Nucleare, Sezione Roma*

L'esistenza dei raggi cosmici, è stata scoperta all'inizio del ventesimo secolo come una causa extraterrestre di ionizzazione. Dopo due decenni di studi i fisici sono riusciti a stabilire che i raggi cosmici sono formati da particelle cariche relativistiche, con uno spettro che si estende fino ad altissime energie. Questo risultato richiede l'esistenza di sorgenti astrofisiche la cui natura è però rimasta a lungo sconosciuta, perché le traiettorie delle particelle sono deviate dai campi magnetici interstellari. Negli ultimi anni sono stati fatti enormi progressi nell'identificazione di queste sorgenti grazie ad osservazioni che utilizzano raggi gamma e neutrini. Queste particelle (elettricamente neutre, e che dunque viaggiano in linea retta) sono generate come secondarie dalle interazioni dei raggi cosmici nelle sorgenti, e le loro direzioni d'arrivo puntano verso di esse. L'universo delle alte energie si è mostrato molto ricco, con diverse classi di sorgenti (Supernovae, Pulsar, Nuclei Galattici Attivi, Gamma Ray Bursts, ...). Queste sorgenti sono anche spesso associate ad eventi astrofisi-

ci dove grandi masse subiscono forti accelerazioni, e dunque emettono onde gravitazionali. Futuri studi che combinano questi quattro messaggeri: raggi cosmici, raggi gamma, neutrini ed onde gravitazionali potranno chiarire molte delle questioni ancora aperte, ed esplorare più profondamente un universo pieno di sorprese.

Cosa sono i raggi cosmici?

I **raggi cosmici** sono una pioggia di particelle di altissima energia e velocità molto vicina a quella della luce che arriva continuamente sulla Terra da tutte le direzioni. La maggior parte di queste particelle sono protoni e nuclei atomici completamente ionizzati, una piccola parte è formata da elettroni, ed esiste anche un minuscolo ma importante contributo di particelle di anti-materia: anti-protoni ed anti-elettroni.

L'esistenza dei raggi cosmici è nota ai fisici da oltre un secolo, e dopo alcuni decenni di studio siamo riusciti a comprendere la loro natura. Ma da dove vengono? Quali sorgenti li emettono? Come ottengono la loro elevatissima energia? Molte fondamentali domande sull'origine dei raggi cosmici attendono ancora risposte certe. Gli ultimi anni hanno però visto grandi progressi

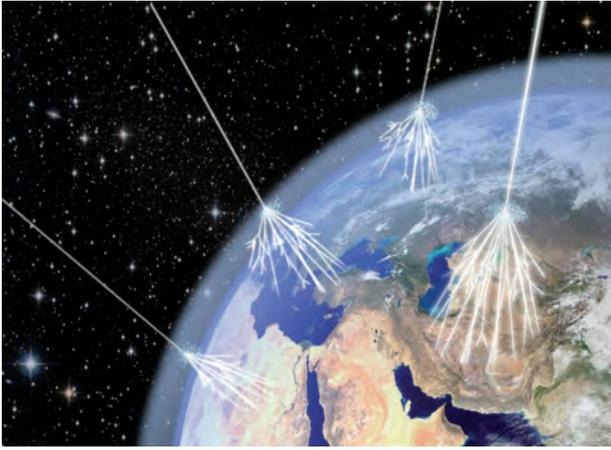


Figura 1: I raggi cosmici primari arrivano continuamente sulla Terra da tutte le direzioni, e generano sciame di secondari nell'atmosfera.

nello studio dei raggi cosmici che ci stanno facendo avvicinare alla soluzione di questi problemi, e allo stesso tempo stanno rivelando l'esistenza di nuovi affascinanti oggetti nel cielo.

La prima e più sorprendente caratteristica dei raggi cosmici è la loro grandissima energia. Per dare un'idea di quanto sia grande possiamo confrontarla con quella di particelle più familiari. L'energia può essere misurata utilizzando molte unità differenti, Joules, ergs, kilowattora, barili di petrolio e così via. I fisici che studiano i raggi cosmici utilizzano un'unità chiamata elettron-Volt (eV), adatta a descrivere i fenomeni del mondo microscopico. Le molecole dell'aria ad una temperatura normale, di circa 20 gradi centigradi, si muovono caoticamente viaggiando ad una velocità media di circa 2000 km/sec e posseggono un'energia cinetica media di circa 0.04 eV. Alla superficie del Sole la temperatura è molto più elevata, circa 6000 gradi, e le particelle del plasma solare trasportano in media un'energia venti volte maggiore, ovvero circa 1 eV. La luce del Sole è formata da un insieme di granuli discreti, o fotoni, emessi come scintille quando le particelle del plasma si scontrano, e trasportano la stessa energia media di circa 1 eV. Le particelle emesse dalle sostanze radioattive, i raggi alfa, beta e gamma emergono dal profondo dei nuclei atomici con un'energia elevatissima che si misura in milioni di elettron-Volt (10^6 eV) (e per questo possono essere molto distruttivi per gli organismi biologici).

Ora l'energia dei raggi cosmici ha uno spettro che comincia a circa un miliardo di eV (10^9 eV)

e si estende ad energie veramente spaventose. Le particelle di energia più elevata possono infatti trasportare cento miliardi di miliardi di eV (10^{20} eV). Per chiarezza, bisogna dire che queste energie sono spaventose perchè sono trasportate da un'unica particella. Le energie massime osservate sono più o meno uguali a quella di una palla da tennis nel servizio di un buon giocatore, che certo non è poco, ma non sembra nulla di straordinario, se non ci si ferma a pensare che una palla da tennis contiene circa 3×10^{25} tra protoni e neutroni, mentre nei raggi cosmici tutta questa energia può essere trasportata da un singolo protone.

Le energie di cui stiamo parlando sono energie cinetiche ovvero associate alla velocità ed alla massa di un corpo (o di una particella). Nella nostra esperienza quotidiana, quando consideriamo palle da tennis, proiettili o automobili in movimento, ovvero cose che si muovono a velocità piccole rispetto a quella della luce (di trecentomila chilometri al secondo) la relazione tra energia, massa e velocità è data dalla semplice equazione $E = \frac{1}{2} m v^2$, dove m è la massa del corpo, e v la sua velocità. Questa equazione dice che l'energia è proporzionale alla massa di una particella e cresce con il quadrato della sua velocità, ma l'equazione non è più valida quando la velocità di un corpo si avvicina a quella della luce. Questa infatti è una velocità limite non raggiungibile da un corpo di massa finita, mentre particelle di massa nulla come i fotoni devono muoversi sempre esattamente a questa velocità, senza mai potere rallentare. Una particella può comunque trasportare un'energia arbitrariamente elevata¹, e quanto più grande è questa energia, tanto più la sua velocità si avvicina a quella della luce, senza però mai riuscire a raggiungerla. Il risultato è che le particelle dei raggi cosmici hanno tutte una velocità molto vicine a quella della luce.

¹L'equazione che descrive l'energia di una particella per una velocità arbitraria è stata formulata da Albert Einstein, ed è: $E = mc^2 / \sqrt{1 - v^2/c^2}$ dove c indica la velocità della luce. È facile vedere che quando v si avvicina a c , il denominatore si annulla, e l'energia diverge all'infinito. Quando la velocità v è uguale a zero, l'energia di una particella in quiete non si annulla ma prende il valore $E = mc^2$.

La scoperta dei raggi cosmici

È interessante ripercorrere brevemente la storia di come si sia prima scoperta l'esistenza dei raggi cosmici e poi compresa la loro natura. (Si veda anche l'articolo di Alessandro De Angelis in questo numero di Ithaca.) La scoperta è avvenuta all'inizio del ventesimo secolo. A quell'epoca i fisici avevano cominciato a comprendere la struttura della materia, e scoperto che gli atomi non sono indivisibili come suggerito dal loro nome, ma sono formati da componenti più elementari e dotati di carica elettrica. La prima particella, l'elettrone venne infatti scoperta nel 1897 dal fisico inglese J.J. Thomson. Si comprese anche che gli atomi possono ionizzarsi, e cioè dividersi in elettroni e ioni positivi.

Lo studio dei fenomeni elettrici dimostrò che l'aria, e più in generale tutti i gas, sono sempre, anche se debolmente, ionizzati, cioè contengono una piccola percentuale di atomi scissi in elettroni liberi e ioni positivi. Poiché elettroni e ioni positivi tendono a ricombinarsi riformando atomi neutri, qualcosa come una radiazione ionizzante, deve dunque agire sulla materia per estrarre continuamente gli elettroni dagli atomi.

Un'importante sorgente di radiazione ionizzante, la radioattività, era stata scoperta negli ultimi anni del XIX secolo quando il fisico francese Henri Becquerel nel 1896 si rese conto che l'Uranio emette continuamente una radiazione capace di annerire le emulsioni fotografiche e appunto di ionizzare la materia. Nuovi studi sulla radioattività, in particolare quelli di Pierre e Marie Curie rivelarono presto l'esistenza di altre sostanze radioattive come Torio, Polonio e Radio, e gli esperimenti di Ernest Rutherford dimostrarono l'esistenza di tre tipi di radioattività: i raggi alfa, beta e gamma, dovuti all'emissione di nuclei di elio, elettroni e fotoni, e che queste particelle trasportavano ciascuna una grandissima energia di milioni di elettron-Volt. L'osservazione della radioattività apriva una finestra sul mondo ancora sconosciuto dei nuclei atomici.

La capacità delle sostanze radioattive di ionizzare la materia offriva un attraente e semplice metodo per studiare sperimentalmente la presenza di tracce di elementi radioattivi in diversi materiali, perché l'intensità della causa ionizzante può essere facilmente misurata osservando

la velocità con cui si scarica un elettroscopio². Questa velocità è infatti proporzionale al tasso di creazione di ioni nel gas contenuto nello strumento. Gli studi dei fisici, in particolare quelli di Julius Elster ed Hans Geitel, dimostrarono che anche nell'apparente assenza di sostanze radioattive, un tasso di ionizzazione risulta sempre misurabile. Già nel 1901 Charles Wilson suggerì che forse questo non era dovuto alla presenza di piccole tracce di materiali radioattivi, ma era invece l'indicazione dell'esistenza di una sorgente extraterrestre di radiazione ionizzante.

Per oltre un decennio fisici di diversi paesi in Europa cercarono di chiarire questo problema, portando elettroscopi in diversi luoghi per misurare come il tasso di creazione di ioni cambiava in diverse situazioni. Ad esempio nel 1910 Theodor Wulff (un fisico tedesco che era anche un sacerdote) portò un elettroscopio in cima alla Tour Eiffel a Parigi, per vedere come l'azione ionizzante cambiava sollevandosi a trecento metri di altezza, ma i suoi risultati furono inconclusivi.



Figura 2: Domenico Pacini (1878–1934) ha studiato la radiazione ionizzante “alla superficie ed in seno alle acque” ed è arrivato vicinissimo a dimostrare l'esistenza di una componente extraterrestre.

Un importante pioniere in queste ricerche è stato il fisico italiano Domenico Pacini, che ebbe l'idea di misurare l'azione ionizzante sott'acqua. Pacini osservò che se la radiazione ionizzante

²L'elettroscopio è un semplice strumento messo a punto da Alessandro Volta, che permette di stabilire se un corpo possiede una carica elettrica. Lo strumento consiste in due lamine di metallo racchiuse in un recipiente di vetro. Le lamine si separano quando posseggono una carica elettrica per la repulsione elettrostatica. Quando nel gas del recipiente si formano ioni, questi si trasferiscono alle lamine ed annullano la carica. La velocità con cui le lamine (o foglioline) ritornano alla posizione normale (quella che si ha in assenza di carica) è proporzionale all'intensità della radiazione ionizzante.

è extraterrestre, ed arriva dal cielo, i suoi effetti dovrebbero diminuire con la profondità. Nel suo lavoro più importante pubblicato nel 1911 sul Nuovo Cimento con titolo “La radiazione penetrante alla superficie ed in seno alle acque”, Pacini presentò osservazioni del tasso di ionizzazione misurato alla superficie ed alla profondità di circa tre metri fatte nel mare a Livorno e nel lago di Bracciano. Pacini trovò che la causa ionizzante diminuiva con la profondità, ed era “indipendente dall’azione diretta delle sostanze radioattive del terreno”. Per molti questi risultati possono essere considerati la prima dimostrazione dell’esistenza di una radiazione ionizzante extraterrestre.

Altri scienziati di diversi paesi europei contribuirono a dimostrare l’esistenza di questa radiazione extraterrestre. I risultati più convincenti vennero ottenuti dal fisico austriaco Victor Hess, che riuscì a misurare il tasso di ionizzazione ad alta quota, portando i suoi elettroscopi in volo su una mongolfiera. Il volo più lungo e di maggiore successo si svolse il 7 agosto del 1912, quando una mongolfiera con a bordo Hess, un pilota ed un meteorologo, volò per sei ore da Aussig in Boemia fino a Berlino superando i cinquemila metri di altezza. Le misure dimostrarono chiaramente che il tasso di ionizzazione aumentava con la quota, e Victor Hess arrivò alla conclusione che: “I risultati delle mie osservazioni si spiegano meglio assumendo che una radiazione di alto potere ionizzante entri dall’alto nella nostra atmosfera”. Grazie a questo risultato, nel 1936 Victor Hess ricevette il premio Nobel per la fisica con la motivazione “per la sua scoperta della radiazione cosmica”.

Negli anni seguenti il tasso di ionizzazione venne estesamente studiato in funzione della quota e della posizione geografica, senza ottenere grandi progressi nella comprensione della sua natura. La teoria comunemente accettata era che la radiazione extraterrestre era formata da raggi gamma, la più penetrante delle tre forme di emissione delle sostanze radioattive. La parola “raggi cosmici” venne introdotta da Robert Millikan nel 1925, che formulò un’ambiziosa (e completamente sbagliata) teoria secondo la quale questi raggi gamma venivano generati nello spazio extragalattico durante la continua formazione di nuclei atomici, ed erano dunque i “vagiti



Figura 3: Una dimostrazione convincente che una forma di radiazione ionizzante è di natura extraterrestre venne ottenuta con misure su pallone. Walter Kolhörster (nell’immagine con un assistente) estese nel 1913 le misure di Victor Hess.

di nascita” degli elementi.

La natura extragalattica dei raggi cosmici era suggerita dal fatto che la radiazione è in ottima approssimazione isotropa, ed arriva in modo quasi perfettamente uniforme da tutte le direzioni. L’isotropia dei raggi cosmici in realtà è generata dal fatto che le particelle primarie, quelle che arrivano sulla Terra dallo spazio, hanno carica elettrica e sono quindi deviate dai campi magnetici che permeano lo spazio. Questo effetto rimescola le direzioni delle particelle dei raggi cosmici che non puntano più alle loro sorgenti.

Anche la Terra (come fanno tutti coloro che hanno utilizzato una bussola) possiede un campo magnetico, e gli effetti del campo geomagnetico lasciano la loro impronta sulle distribuzioni dei raggi cosmici, se questi hanno una carica elettrica non nulla. L’effetto più importante è che il flusso di particelle che arriva alla superficie della Terra dipende dalla latitudine magnetica³, ed è massimo ai poli e minimo all’equatore. Questo effetto di latitudine venne osservato per la

³Il campo magnetico della Terra è in buona approssimazione dipolare, ma l’asse di questo dipolo non coincide con l’asse di rotazione terrestre, e perciò i poli magnetici non coincidono con quelli geografici.

prima volta da Jacob Clay nel 1927, ed in seguito in modo più convincente da Arthur Compton nel 1933, grazie ad un dettagliato studio della ionizzazione esteso a tutto il mondo. Queste osservazioni dimostrarono che la maggior parte dei raggi cosmici hanno carica elettrica ma non erano sufficienti per determinare il segno di questa carica. Un secondo effetto del campo magnetico terrestre venne però osservato dopo poco tempo, l'effetto est-ovest, vale a dire un piccolo eccesso di particelle che viaggiano da ovest verso est, rispetto a particelle che viaggiano in direzione opposta (da est verso ovest). Questo effetto dimostra che la maggior parte delle particelle dei raggi cosmici hanno carica elettrica positiva. Quando questo risultato fu ottenuto per la prima volta, fu una grande sorpresa, perché la maggior parte dei fisici si aspettava che se i raggi cosmici avevano carica elettrica, dovevano essere elettroni.

La composizione (ed anche lo spettro di energia) dei raggi cosmici primari venne misurata in seguito grazie a telescopi posizionati ad alta quota su palloni sonda o satelliti, ed in grado di osservare le particelle primarie prima della loro interazione con l'atmosfera. Le misure confermarono che la maggior parte delle particelle primarie sono protoni.

Particelle elementari

Lo studio dei raggi cosmici è stato intimamente legato alla nascita ed allo sviluppo della fisica delle particelle elementari. Alla fine degli anni '20 del ventesimo secolo, lo sviluppo di nuove tecniche sperimentali che permettevano di misurare e visualizzare le particelle della radiazione cosmica diede inizio a una nuova fase di rapido progresso. Un primo esperimento chiave venne svolto nel 1928-1929 da due fisici tedeschi Walter Bothe and Werner Kolhörster, utilizzando dei nuovi rivelatori conosciuti come i tubi Geiger-Müller (dal nome dei fisici che li avevano inventati). Questi rivelatori segnalavano con un clic, ovvero un impulso elettrico, lo spezzarsi di un atomo con la creazione di un elettrone libero ed uno ione positivo. La nuova, ed importantissima idea di Bothe e Kolhörster fu di mettere in coincidenza due rivelatori Geiger-Müller uno sopra l'altro ad una distanza di alcuni centime-

tri. I due fisici trovarono che molto spesso i due rivelatori davano segnali allo stesso tempo, indicando che erano stati attraversati da una stessa particella, capace di ionizzare atomi in tutti e due gli strumenti. Alcune di queste particelle dovevano anche essere straordinariamente penetranti, perché inserendo un blocco di oro di alcuni centimetri di spessore tra i due contatori, la frequenza delle coincidenze diminuiva solo parzialmente. Questi risultati dimostravano che i raggi cosmici non erano raggi gamma, ma avevano una natura corpuscolare, con proprietà nuove e sorprendenti.

Bruno Rossi, un giovane fisico italiano (all'epoca aveva solo 23 anni) che avrebbe dato un grande contributo allo studio dei raggi cosmici scrive così della sua reazione nel leggere il lavoro dei fisici tedeschi:

“L'articolo di Bothe and Kolhörster arrivò come un lampo di luce, che rivelava l'esistenza di un mondo insospettato, pieno di misteri, e che nessuno aveva ancora cominciato ad esplorare”.

Gli anni successivi videro i fisici lavorare intensamente ad esplorare quel mondo insospettato. Seguire passo dopo passo i progressi dei fisici nella comprensione della natura dei raggi cosmici sarebbe molto interessante, ma è troppo lungo per questo articolo (un bellissimo racconto di quegli anni è contenuto nel classico libro di Bruno Rossi “Raggi Cosmici” che lo descrive come “una delle più straordinarie avventure nella storia della scienza”) [1].

Per semplificare la nostra discussione possiamo partire dalla soluzione. Abbiamo già detto che i raggi cosmici primari, quelli che arrivano alla Terra dallo spazio, sono particelle di altissima energia, che viaggiano nell'aria lasciando dietro di loro una scia di atomi ionizzati. Generare questa ionizzazione fa però perdere alle particelle solo una piccola parte della loro energia, ma ogni tanto un raggio cosmico passa sufficientemente vicino ad un nucleo atomico ed avviene un'interazione nucleare. Queste collisioni nucleari sono rare (le particelle primarie possono attraversare in media circa un decimo dell'atmosfera prima di interagire) perché i nuclei atomici hanno dimensioni lineari circa centomila volte più piccole di quelle dell'involucro di elettroni

che li circonda. In queste interazioni con i nuclei la particella proiettile viene assorbita, e dalla collisione emergono particelle di tipo diverso, in un numero che in media (ma con grandi fluttuazioni) cresce con l'energia del proiettile, arrivando anche a centinaia di secondari per interazioni ad alta energia. Le particelle nello stato finale di una collisione nucleare non sono solo i frammenti delle particelle interagenti, ma per la maggior parte sono create dal vuoto nel momento dell'interazione, in modo analogo a come lo scontro frontale tra due macchine produce non solo rottami, ma crea anche scintille che certo prima non facevano parte delle macchine stesse. Per comprendere la natura di queste scintille bisogna introdurre due concetti fondamentali, il primo è la celebre relazione di equivalenza tra massa ed energia scoperta da Einstein: $E = mc^2$. Questa equazione dice che è possibile materializzare l'energia cinetica delle particelle in collisione creandone di nuove, se l'energia è sufficientemente grande. Il secondo concetto importante è che esistono delle leggi di conservazione che devono essere rispettate quando nuove particelle vengono create dal vuoto. Ad esempio, la legge di conservazione della carica elettrica dice che la somma delle cariche di tutte le particelle di un sistema fisico rimane sempre costante. Un'altra importantissima legge fondamentale che gioca un ruolo cruciale nei processi di creazione di nuove particelle è il principio che per ogni particella di materia, ne esiste un'altra di anti-materia che ha massa identica, lo stesso spin, e tutte le cariche (o numeri quantici interni) opposti. Ad esempio, oltre all'elettrone, che ha carica elettrica negativa, esiste l'anti-elettrone (chiamato anche positrone) che ha la stessa massa ma carica elettrica uguale ed opposta (e dunque positiva). Oltre al protone, esiste l'anti-protone con carica elettrica opposta (e dunque negativa), ed anche carica nucleare opposta⁴ La legge sulla sim-

⁴Il protone, oltre alla forza elettromagnetica, è sensibile anche alla forza nucleare conosciuta anche come l'interazione forte) che è quella che tiene protoni e neutroni legati nei nuclei atomici. Le particelle sensibili a questa forza posseggono una carica nucleare, analoga alla carica elettrica, che i fisici chiamano anche colore), e le loro anti-particelle hanno carica opposta. Il colore è definito non da uno, ma da due numeri, e può essere visualizzato come un punto su un piano. I quark esistono in tre colori differenti, che su questo piano sono rappresentati dai vertici di un triangolo equilatero, men-

metria materia/anti-materia è stata compresa per primo dal grande fisico inglese Paul Dirac, ed emerge necessariamente quando si mettono insieme la teoria della Relatività di Einstein con i principi della Meccanica Quantistica. La creazione di una coppia particella/anti-particella rispetta automaticamente tutte le leggi di conservazione e dunque è possibile quando l'energia disponibile è sufficiente.

Le particelle ed anti-particelle create nelle interazioni dei raggi cosmici primari con i nuclei dell'aria si propagano nell'atmosfera, in alcuni casi si disintegrano spontaneamente, in altri possono interagire nuovamente, generando così altre particelle (ed antiparticelle). Sono questi raggi cosmici secondari le particelle che sono osservate vicino alla superficie della Terra (attraverso la ionizzazione dei gas, oppure in rivelatori di diverso tipo). L'intricato problema affrontato dai fisici nelle decadi degli anni '30 e '40 fu di capire quali particelle sono prodotte nelle interazioni di alta energia, quali sono le loro proprietà, come si propagano nell'aria, per poi risalire alla natura e proprietà della radiazione primaria. Lo studio dei raggi cosmici coincise dunque con la nascita e lo sviluppo della fisica delle particelle elementari, dedicata allo studio delle componenti fondamentali della materia.

Possiamo ricordare alcuni degli sviluppi più profondi e rivoluzionari. Il primo avvenne nel 1932, e fu l'osservazione dell'anti-particella dell'elettrone da parte di Carl Anderson utilizzando un rivelatore, chiamato camera a nebbia che permette di visualizzare la traiettoria di una particella carica relativistica come una sequenza di goccioline che si formano intorno agli atomi ionizzati dal suo passaggio. Nell'esperimento di Anderson, la camera a nebbia era posta all'interno di un magnete, in modo che le traiettorie delle particelle cariche che la attraversavano venissero incurvate dal campo magnetico. Il verso della curvatura indicava il segno della carica elettrica. Con grande sorpresa Anderson trovò che alcune delle particelle osservate avevano la massa di un elettrone, ma carica positiva. Paul Dirac aveva da poco predetto l'esistenza dell'anti-elettrone, e

tre gli anti-quark hanno colore opposto. Neutrone ed anti-neutrone sono formati da quarks ed anti-quarks e sono particelle differenti.

dunque la scoperta fu una spettacolare conferma di questa predizione. Ci si rese presto conto che elettroni e anti-elettroni (chiamati anche comunemente positroni) sono in buona approssimazione ugualmente abbondanti nella radiazione cosmica secondaria, e ne formano una gran parte. La spiegazione di questo fatto venne presto trovata sviluppando la teoria di Paul Dirac, che aveva previsto che fotoni di energia sufficientemente elevata (vale a dire con energia $E > 2m_e c^2$ dove m_e è la massa dell'elettrone) possono (nel campo elettrico di un nucleo dell'aria) trasformarsi in una coppia elettrone/anti-elettrone (e^+e^-).

Gli esperimenti mostrarono anche che gli elettroni e positroni dei raggi cosmici secondari sono accompagnati da un flusso di fotoni di alta energia (o raggi gamma). Queste tre particelle, elettroni, positroni e fotoni formano insieme delle cascate elettromagnetiche. Il meccanismo si può illustrare in questo modo: un primo fotone si converte in una coppia e^-e^+ , queste particelle possono poi emettere altri fotoni cedendo loro parte della loro energia (in processi del tipo $e^\mp \rightarrow e^\mp + \gamma$) che possono a loro volta generare nuove coppie elettrone/positrone e così via. La presenza di elettroni/positroni/fotoni nei raggi cosmici si poteva dunque spiegare assumendo che le particelle primarie dei raggi cosmici nelle loro interazioni potessero generare una gran quantità di raggi gamma. Un'ipotesi che sarebbe stata confermata qualche anno dopo.

Una seconda scoperta di profonda importanza ottenuta già negli anni '30, è stata quella dell'esistenza dei leptoni mu o muoni (indicati con la lettera greca μ^- e μ^+). La prima indiretta indicazione dell'esistenza della nuova particella (e della sua anti-particella) era emersa dall'osservazione che i raggi cosmici secondari erano molto più penetranti delle predizioni teoriche per elettroni ed anti-elettroni, un fenomeno già trovato da Bothe e Kohlörster nei primi esperimenti di coincidenza. Questo risultato era spiegabile solo con una profonda revisione della teoria, oppure con l'esistenza di una nuova particella che costituiva una seconda componente (accanto a quella elettromagnetica formata da elettroni, positroni e fotoni) dei raggi cosmici secondari. Il problema venne chiarito alla fine degli anni '30, quando nelle camere a nebbia che osservavano i raggi cosmici secondari venne identificata una nuova

particella con la stessa carica elettrica dell'elettrone, ma una massa circa 205 volte maggiore. Le proprietà del muone vennero comprese poco a poco, ed un ruolo fondamentale in questo fu svolto dai tre giovani fisici italiani Marcello Conversi, Oreste Pancini ed Ettore Piccioni che grazie a delle ingegnose misure, fatte a Roma in condizioni straordinariamente difficili tra il 1943 ed il 1946, riuscirono a dimostrare che il muone non sente la forza nucleare (quella che tiene insieme protoni e neutroni nei nuclei) e si comporta dunque come un "elettrone pesante"⁵.

La sorgente dei muoni nei raggi cosmici secondari venne compresa grazie alle osservazioni ottenute con una nuova tecnica di rivelazione: le emulsioni nucleari, cioè un tipo di emulsione fotografica, come quella che veniva utilizzata per ottenere immagini quando le tecniche digitali ancora non esistevano. Nella gelatina che è la base dell'emulsione sono sospesi dei cristalli che possono essere impressionati dalla luce, oppure dal passaggio di una particella carica di alta energia. Dopo un processo di sviluppo, i cristalli impressionati si anneriscono e la traiettoria di una particella può essere visualizzata come una sequenza di punti, visibili al microscopio.

Questa tecnica venne inizialmente sviluppata da due ricercatrici di Vienna, Marietta Blau e la sua studentessa Hertha Wambacher, che nel 1937 esposero delle emulsioni preparate a questo scopo in un laboratorio costruito ad alta quota (2300 metri) nelle alpi austriache da Victor Hess. Nelle emulsioni erano visibili delle immagini che chiamarono stelle di disintegrazione ("Zertrümmerungsterne") che interpretarono come le tracce disegnate dai frammenti di un nucleo atomico disintegrato dalla collisione con una particella dei raggi cosmici. Si trattava delle prime immagini di interazioni nucleari. Purtroppo Marietta Blau non ebbe modo di continuare questa promettente linea di studi. Nel 1938 ci fu l'Anschluss, e l'Austria venne assorbita dalla

⁵L'esistenza del muone fu una grande sorpresa per i fisici. È rimasto celebre il commento alla scoperta fatto da Isidor Isaac Rabi: "E questo chi lo ha ordinato?". Anche oggi questa domanda rimane senza risposta. Oggi sappiamo che non solo l'elettrone ma tutte le particelle di spin 1/2 sono non solo duplicate, ma triplicate. Esistono cioè (per ogni tipo) tre diverse particelle con identiche proprietà ma massa differente. L'origine di queste tre generazioni di particelle rimane un mistero che intriga ed affascina i fisici ormai da molti anni.

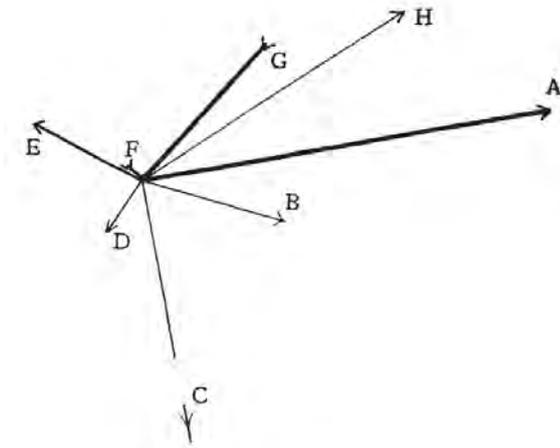


Figura 4: Schema della prima immagine di una stella di disintegrazione (una interazione di un raggio cosmico) ottenuta in emulsione fotografica da Marietta Blau ed Hertha Wambacher.

Germania nazista, e lei, ebrea, fu costretta all'esilio, prima in Messico e poi negli Stati Uniti, e la sua attività di ricerca venne irrimediabilmente interrotta⁶.

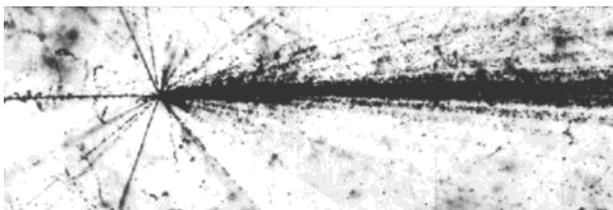


Figura 5: Interazione in emulsione fotografica con la produzione di molte particelle.

La tecnica delle emulsioni nucleari venne presto perfezionata da altri fisici, in particolare da Cecil Powell, Giuseppe Occhialini e Cesar Lattes, che riuscirono a migliorare la sensibilità delle emulsioni aumentando la concentrazione dei cristalli di bromuro d'argento, in modo da permettere la visualizzazione non solo dei frammenti del nucleo (che sono fortemente ionizzanti perché lenti e con grande carica elettrica) ma anche di particelle cariche relativistiche. Powell, Occhialini e Lattes esposero le loro emulsioni in laboratori di alta quota, prima nelle Alpi svizzere, e poi ad altezza ancora maggiore (più di

⁶Un aspetto paradossale di questa storia è che Hertha Wambacher era invece stata iscritta al partito nazista (NSDAP) già dal 1934, quando era ancora illegale in Austria. Dopo la guerra, Marietta Blau venne proposta per due volte al comitato del Nobel da Erwin Schrödinger, ma questa proposta non venne accettata.

5000 metri) nelle Ande boliviane. Le immagini registrate permisero delle scoperte importanti. Nella sua autobiografia Powell scrive:

“[...] un intero nuovo mondo si rivelava [...] Era come se fossimo penetrati all'interno di un giardino recintato, nella protezione del quale molti alberi erano fioriti, e frutta esotica di tutti i tipi era maturata in grande abbondanza”.

I frutti esotici rivelati dalle immagini delle emulsioni erano nuovi tipi di particelle. La prima e più importante scoperta fu quella dei pioni carichi (in simbolo π^+ e π^- , uno l'antiparticella dell'altro) che venivano prodotti in gran numero nelle interazioni nucleari. I pioni sono particelle instabili, che si disintegrano dopo breve tempo, in un muone ed una particella invisibile: un neutrino. Anche il muone è instabile e a sua volta si disintegra in un elettrone (o un positrone a seconda della sua carica) ed altre (due come si comprese in seguito) particelle invisibili⁷. La bellezza della tecnica delle emulsioni nucleari era che in alcuni casi l'intera traiettoria di una particella, dalla nascita al decadimento, avveniva dentro l'emulsione e poteva essere studiata in dettaglio. I risultati ottenuti da Powell, Occhialini e Lattes possono essere riassunti scrivendo le reazioni come:

$$\begin{cases} \pi^\pm \rightarrow \mu^\pm + (\text{una particella invisibile}) \\ \mu^\pm \rightarrow e^\pm + (2 \text{ particelle invisibili}) \end{cases}$$

Le particelle emesse negli stati finali di questi decadimenti non sono veramente invisibili, ma interagiscono solo molto debolmente, e si tratta di neutrini, una componente di importanza fondamentale nel mondo delle particelle elementari.

L'esistenza del neutrino era già stata compresa grazie allo studio dei decadimenti radioattivi beta, dove un nucleo si trasmuta in un altro emettendo un elettrone (il raggio beta). Nello stato finale di un decadimento deve essere presente

⁷Utilizzando le leggi di conservazione di energia e momento, dalla misura della particella visibile nello stato finale è possibile determinare se l'energia e momento invisibile è trasportata da un'unica particella, oppure da diverse particelle. Oggi sappiamo che nel decadimento dei μ^\pm vengono emessi due neutrini.

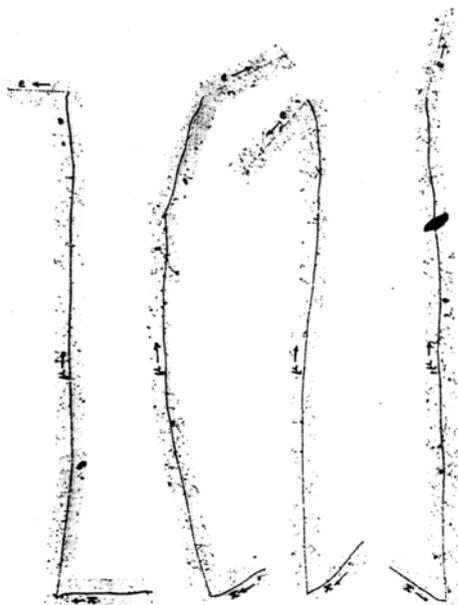


Figura 6: Nella figura si vedono quattro eventi di decadimento di pioni carichi (ottenuti da Powell, Occhialini e Lattes in Bolivia). Il pione (π^\pm) si arresta nell'emulsione, ed emette un muone (μ^\pm) ed un (invisibile) neutrino. Anche il muone si arresta e decade in un elettrone.

anche un'altra particella invisibile (cioè con interazioni debolissime) e di piccola massa, per trasportare l'energia ed il momento mancante, come suggerito per primo da Wolfgang Pauli nel 1930. La prova sperimentale dell'esistenza del neutrino venne ottenuta solo un quarto di secolo dopo nel 1956 da Clyde Cowan e Frederick Reines che riuscirono ad osservare le rare interazioni dei neutrini emessi da un reattore nucleare (dove i frammenti delle fissioni nucleari generano un grande numero di decadimenti beta).

Una questione aperta era se i neutrini emessi nei decadimenti dei pioni carichi (e quelli emessi nel decadimento dei muoni) fossero o no lo stesso tipo di particella emessa nei decadimenti beta. Questa domanda trovò una risposta quindici anni dopo, quando un esperimento nell'acceleratore di Brookhaven negli Stati Uniti riuscì a dimostrare che si trattava di particelle differenti.

I neutrini infatti formano un doppietto con una particella carica sorella (l'elettrone oppure il muone nei casi presi in considerazione) e questa associazione definisce il sapore del neutrino. Ad esempio, in un decadimento beta un neutrone si trasmuta in un protone, emettendo una cop-

pia ($e^- \bar{\nu}_e$) formata da un elettrone ed un anti-neutrino di tipo sapore elettronico. Il neutrino si ricorda del suo sapore, e quando interagisce si trasforma in un anti-elettrone. Le reazioni⁸ si possono scrivere come:

$$\begin{cases} n \rightarrow p + e^- + \bar{\nu}_e \\ \bar{\nu}_e + p \rightarrow e^+ + n \end{cases}$$

Nei decadimenti dei pioni invece i neutrini sono emessi insieme ad un μ^\pm ed hanno sapore muonico, e quando interagiscono generano nuovamente un muone. Le reazioni si possono scrivere come:

$$\begin{cases} \pi^- \rightarrow \mu^- + \bar{\nu}_\mu \\ \bar{\nu}_\mu + p \rightarrow \mu^+ + n \end{cases}, \quad \begin{cases} \pi^+ \rightarrow \mu^+ + \nu_\mu \\ \nu_\mu + n \rightarrow \mu^- + p \end{cases}$$

È ora abbastanza facile ricostruire il sapore dei due neutrini che sono generati nel decadimento dei muoni. Nella notazione usata oggi queste disintegrazioni si possono scrivere come:

$$\mu^- \rightarrow \nu_\mu + e^- + \bar{\nu}_e, \quad \mu^+ \rightarrow \bar{\nu}_\mu + e^+ + \nu_e.$$

Questi decadimenti possono essere descritti nel seguente modo: il muone si trasmuta nel neutrino (o anti-neutrino) del suo stesso sapore, ed emette una coppia ($e^- \bar{\nu}_e$) oppure ($e^+ \nu_e$) di un elettrone ed un anti-neutrino elettrone (o lo stato simmetrico).

Per completare questa discussione va osservato che studi recenti hanno dimostrato che in realtà quello che abbiamo descritto è il sapore dei neutrini subito dopo la loro creazione, ma questo sapore cambia periodicamente nel tempo. Un fenomeno che è stato predetto dal fisico italiano Bruno Pontecorvo, e dimostrato sperimentalmente solo negli ultimi anni. Per la scoperta di questo importante fenomeno, conosciuto come le oscillazioni di sapore dei neutrini, il premio Nobel per la fisica del 2015 è stato assegnato a Takaaki Kajita ed Arthur McDonald. È per studiare queste oscillazioni (o cambi periodici) del sapore che fasci di neutrini creati in laboratorio vengono fatti viaggiare per lunghe

⁸Guardando la forma di queste reazioni non è difficile capire perché si è scelto di chiamare anti-neutrino (e usare la notazione con la barra sopra il simbolo) la particella quasi invisibile emessa insieme all'elettrone.

distanze, come ad esempio dal laboratorio del CERN a Ginevra al laboratorio sotterraneo del Gran Sasso.

La scoperta dei pioni carichi (che sono la sorgente dei muoni) suggerì anche quale fosse la sorgente della componente elettromagnetica (elettroni, positroni e fotoni) nei raggi cosmici secondari. Nel 1948 alcuni fisici (Lewis, Oppenheimer e Wouthuysen) formularono l'ipotesi che i pioni carichi avevano un compagno neutro, il π^0 che veniva creato con abbondanza nelle interazioni nucleari, e che questa particella (che essendo priva di carica elettrica non lasciava tracce nelle emulsioni nucleari) si disintegrava in due fotoni:

$$\pi^0 \rightarrow \gamma + \gamma$$

e che questi fotoni fossero la sorgente di cascate elettromagnetiche. L'esistenza del π^0 venne confermata due anni dopo, in uno dei primi esperimenti che studiavano le interazioni nucleari utilizzando non più i raggi cosmici, ma fasci di particelle accelerate artificialmente.

I tre pioni (π^+ , π^- e π^0) sono le particelle create con maggior abbondanza nelle interazioni nucleari, ma non sono le sole. I fisici cominciarono a scoprirne anche altre (i mesoni K , il barione Λ , ... un vero zoo). Una spiegazione di questi fenomeni, venne ottenuta negli anni successivi grazie alla formulazione della teoria dei **quark**. Secondo questa teoria il protone ed il neutrone sono in realtà particelle composte, formate dall'unione di tre oggetti più fondamentali, e dotati di carica elettrica frazionaria: i quark. Il protone è formato da due quark di tipo up, con carica elettrica $+2/3$, ed un quark di tipo down, con carica elettrica $-1/3$ (in questo modo la carica complessiva è $+1$). Il neutrone invece è formato da due quark di tipo down ed un quark di tipo up (con carica elettrica totale nulla). Accanto ai quarks up e down che formano la materia ordinaria, esistono altri sapori di quark di massa maggiore come il quark strano e quello con charm. I quark hanno la sorprendente proprietà di non potere essere separati, e si manifestano solo combinandosi in tre per particella (come nei protoni e neutroni) oppure in combinazioni di un quark ed un anti-quark⁹.

⁹Questo fenomeno del confinamento dei quarks è sorprendente, e può sembrare impossibile, ma è verificato

I pioni sono le particelle create più facilmente nelle interazioni nucleari perché sono le combinazioni più stabili di quark/anti-quark formate dai due quark più leggeri up e down (il π^+ è la particella $u\bar{d}$, il π^- quella $d\bar{u}$, mentre il π^0 è una miscela di $u\bar{u}$ e $d\bar{d}$), e vengono creati facilmente, come le scintille delle interazioni nucleari.

I raggi cosmici sono stati la prima sorgente di particelle di alta energia, e per molto tempo, oltre due decenni, gli studi sperimentali sulla fisica delle particelle si sono basati sulla loro osservazione. A partire dagli anni '50 però i fisici impararono a costruire degli acceleratori, delle macchine capaci di produrre fasci di particelle di alta energia, e divenne possibile ed anzi preferibile svolgere gli esperimenti utilizzando questi fasci in condizioni più favorevoli e controllate, e la fisica sperimentale delle particelle si trasferì dai palloni e dagli osservatori di alta quota a laboratori come quelli di Frascati o del CERN. Le interazioni tra fisica dei raggi cosmici e fisica delle particelle continuano però ad essere intense e molteplici. Da una parte gli esperimenti agli acceleratori permettono di studiare in dettaglio, ed utilizzando condizioni iniziali esattamente conosciute, le proprietà delle interazioni delle particelle, e questi risultati sono preziosi per interpretare le osservazioni dei raggi cosmici, dall'altra va anche notato che lo spettro di energia dei raggi cosmici si estende fino ad energie molto più elevate di quelle finora ottenute negli acceleratori¹⁰.

sperimentalmente in modo molto solido. Immaginate di prendere da una parte due quark di un protone e dall'altra il terzo quark, e di tirare il più possibile per separarli. Ad un certo punto, l'energia utilizzata per separare le due parti si converte in una coppia quark/anti-quark, e il sistema si ricombina a formare un barione con tre quarks, ed un mesone composto da un quark ed un anti-quark. La separazione è fallita.

¹⁰Al CERN di Ginevra le energie più elevate ottenute sono di 6.5 TeV (1 TeV = 10^{12} eV), molti ordini di grandezza più piccola di quelle dei raggi cosmici. La variabile rilevante per le interazioni è però l'energia nel centro di massa delle collisioni. Al CERN, dove due fasci di protoni vengono fatti scontrare uno contro l'altro questa energia arriva a 13 TeV (il doppio dell'energia di un singolo fascio) mentre per i raggi cosmici le interazioni possono arrivare (anche se solo in rari casi) ad oltre 400 TeV.

Telescopi per Raggi Cosmici

Palloni e Satelliti

Le particelle dei raggi cosmici primari interagiscono nell'atmosfera ad una quota media di 15–20 km sul livello del mare, e dunque per osservarle direttamente è necessario portare dei rivelatori a grande altezza. Questo è diventato possibile a partire dagli anni '50 del ventesimo secolo con lo sviluppo di palloni capaci di raggiungere la stratosfera e rimanere per molte ore ad una quota di circa 30 chilometri. La tecnica dei palloni ha continuato a svilupparsi e oggi sono possibili voli anche molto più lunghi di oltre cento giorni. Con l'avvento dell'era spaziale è poi diventato possibile mettere degli strumenti in orbita intorno alla Terra come satelliti indipendenti, o più recentemente a bordo della International Space Station.



Figura 7: Telescopio per raggi cosmici PAMELA.

Uno sviluppo molto importante è stato quello di riuscire a mettere in orbita dei rivelatori con campo magnetico, perché questo permette di misurare il segno della carica delle particelle. Questa separazione è particolarmente importante per identificare e studiare la presenza di antimateria (anti-elettroni, anti-protoni ed anti-nuclei) nei raggi cosmici. Il primo di questi rivelatori magnetici è stato il satellite PAMELA (progettato e costruito da una collaborazione internazionale con un ruolo fondamentale dell'Italia) lanciato nel 2006 e che ha operato per dieci anni [2]. Qualche tempo dopo, nel 2011 un altro rivelatore magnetico l'Alpha Magnetic Spectrometer (AMS) [3] è stato installato sulla ISS (Internatio-

nal Space Station), e sta ancora raccogliendo dati.



Figura 8: Telescopio per raggi cosmici AMS02, durante la sua installazione sulla International Space Station (ISS).

I rivelatori su pallone e su satellite osservano direttamente la particella primaria, e possono identificare la sua natura (protoni, nuclei, elettroni, ...) e misurare la sua energia e direzione. Le misure hanno mostrato che la maggior parte dei raggi cosmici primari sono protoni, ma che circa il 10% delle particelle sono nuclei completamente ionizzati di diverso tipo. Tutti gli elementi sono presenti (elio, carbonio, ossigeno, ferro, ...) con diverse abbondanze che danno informazione sulle proprietà delle sorgenti.

Uno studio di importanza fondamentale è quello di ricostruire la distribuzione (o spettro) di energia delle particelle primarie. Questo spettro ha una forma caratteristica di legge di potenza, vale a dire che il flusso dipende dall'energia in modo proporzionale ad E^{-p} , dove l'esponente p (in un ampio intervallo di E) ha un valore di circa 2.7–2.8. Questo vuole dire che il flusso diminuisce molto rapidamente quando l'energia aumenta (di un fattore circa 500 per ogni decade). Capire quali meccanismi generano questa forma spettrale rimane uno dei problemi fondamentali per l'astrofisica dei raggi cosmici.

Gli elettroni rappresentano solo circa due centesimi del flusso delle particelle primarie, ed hanno uno spettro che decresce più rapidamente di quello dei protoni e dei nuclei, di nuovo con la forma di una legge di potenza, ma con un esponente più grande ($p \approx 3.0-3.1$). Questo è un

altro importante indizio per decifrare l'origine dei raggi cosmici.

Le osservazioni di PAMELA ed AMS che hanno generato più discussioni sono state quelle ottenute, con un dettaglio senza precedenti, dei flussi di particelle di anti-materia (positroni ed ed anti-protoni) nei raggi cosmici. Si tratta di contributi relativamente piccoli, meno di due particelle su mille sono positroni, e meno di una su diecimila è un antiprotone, ma è certamente di grande interesse capire dove e come queste anti-particelle sono prodotte¹¹. Una possibile sorgente di particelle di antimateria è la loro creazione nelle interazioni dei raggi cosmici più comuni con il gas interstellare, ma esistono anche altre possibilità.

La teoria più affascinante è che le anti-particelle nei raggi cosmici siano state generate dalla misteriosa materia oscura di cui conosciamo l'esistenza solo grazie ai suoi effetti gravitazionali. Alcune teorie, discusse intensamente da diversi anni, prevedono che la materia oscura presente nella nostra Galassia sia formata da un nuovo tipo di particelle con una massa centinaia o forse migliaia di volte quella del protone, che rimangono quasi invisibili perché hanno solo una piccola probabilità di interazione con la materia ordinaria. Queste particelle possono però annihilarsi tra di loro, oppure (in alcune varianti della teoria) disintegrarsi spontaneamente con una vita media molto lunga. Lo stato finale di queste reazioni è formato da uguali quantità di materia ed anti-materia, nella forma di particelle di alta energia, e molti lavori teorici suggeriscono la possibilità che una parte del flusso dei positroni (e forse anche degli anti-protoni) sia generato da questo meccanismo.

Un'altra possibilità interessante, anche questa oggetto di molti studi teorici, è che le anti-particelle siano create ed accelerate in sorgenti astrofisiche. Questo è molto più facile nel caso degli anti-elettroni, perché hanno una massa molto più piccola di quella degli anti-protoni, e quindi il costo in energia per creare una coppia (e^-e^+) è molto minore. Molti lavori hanno discusso la possibilità che nelle Pulsar (di cui

¹¹Al momento non è ancora stata osservata la presenza di anti-nuclei nei raggi cosmici, a parte la presenza di una manciata di possibili candidati da parte di AMS per i nuclei più leggeri (anti-deuterio ed anti-elio) che sono ancora da confermare.

parleremo brevemente nel seguito) si possano creare ed accelerare ad alta energia dei positroni.

L'origine della componente di antimateria nei raggi cosmici rimane un problema aperto, e oggetto di un'intensa discussione. Solo nuove osservazioni permetteranno di arrivare a delle conclusioni convincenti.

Le cascate atmosferiche estese

Le dimensioni geometriche dei telescopi per raggi cosmici che osservano direttamente le particelle primarie devono essere posizionati su un pallone sonda o su un satellite, e quindi non possono superare (almeno oggi) un'area di qualche metro quadrato. Questa limitazione non permette di osservare il flusso dei raggi cosmici di energia molto elevata. Questo perché il flusso decresce rapidamente con l'aumentare dell'energia, e diventa troppo piccolo. Ad esempio, il flusso per $E \gtrsim 3 \times 10^{15}$ eV è di circa 20 particelle per metro quadrato per anno, ed i più grandi rivelatori su satellite non riescono ad osservare che pochi eventi in tutta la loro vita. Per energie ancora maggiori il flusso diventa minuscolo, per $E \simeq 10^{19}$ eV si scende al livello di una particella per chilometro quadrato per anno.

Lo studio dei raggi cosmici ad alta energia rimane però possibile grazie all'osservazione degli sciame o cascate generate dalle particelle primarie in atmosfera. Le prime osservazioni di questi sciame vennero ottenute già negli anni '30, quando venne inventata la tecnica delle osservazioni in coincidenza. I primi esperimenti che utilizzavano questo metodo avevano posizionato i rivelatori di particelle uno sopra l'altro per osservare il passaggio della stessa particella attraverso due o più strumenti, ma è chiaramente possibile anche posizionare i rivelatori uno accanto all'altro, a distanza variabile, per vedere se le particelle arrivano isolate, oppure in gruppi. Esperimenti svolti alla fine degli anni '30 dal fisico francese Pierre Auger e dai suoi collaboratori ottennero dei risultati sorprendenti. Il tasso di coincidenza rimaneva elevato anche per rivelatori posizionati a centinaia di metri l'uno dall'altro. Questo è possibile se una singola particella primaria dei raggi cosmici è in grado di generare un grande numero (molti milioni in alcuni casi) di particelle secondarie che viaggiano insieme,

su traiettorie quasi parallele. Queste particelle, al livello del suolo, riempiono (con una distribuzione non uniforme, che decresce con la distanza dal centro) un disco con un raggio di dimensione variabile ma che può arrivare a essere di centinaia di metri (come osservato da Auger), o anche di diversi chilometri (come osservato oggi). Pierre Auger ed i suoi collaboratori, dimostrarono anche che la maggior parte delle particelle rivelate in coincidenza sono elettroni e positroni (accompagnati da un flusso di fotoni) e muoni (identificabili perché più penetranti, e capaci di attraversare lamine di piombo), e cercarono (con risultati in buona approssimazione corretti) di stimare l'energia della particella primaria, utilizzando ciò che era noto sulle interazioni, ed arrivarono alla conclusione che lo spettro dei raggi cosmici si doveva estendere fino ad energie veramente straordinarie, almeno fino a milioni di miliardi di elettron-Volt (10^{15} eV).

Questo insieme di particelle (tutte generate da un singolo primario) che si propagano insieme nell'atmosfera è paragonabile ad uno sciame di api che volano tutte nella stessa direzione, o alle gocce di uno scroscio d'acqua o di una cascata, ed i fisici lo chiamano sciame atmosferico esteso o cascata atmosferica estesa.

Il grande numero di particelle in uno sciame è dovuto al fatto che il numero cresce esponenzialmente quando inizia a svilupparsi, perché le particelle generate nella prima interazione (quella della particella primaria) possono interagire a loro volta e così di seguito, e penetrando nell'atmosfera lo sciame è formato da particelle sempre più numerose e di più bassa energia. Questo processo di moltiplicazione si arresta solo quando le particelle hanno un'energia sufficientemente bassa e vengono assorbite dall'aria depositando la loro energia come ionizzazione, e così il numero di particelle in uno sciame raggiunge un massimo e poi inizia a decrescere.

L'esistenza degli sciami atmosferici offre un attraente metodo per misurare le proprietà dei raggi cosmici di altissima energia con rivelatori al livello del suolo capaci di osservare questi sciami. Queste osservazioni hanno mostrato che lo spettro dei raggi cosmici si estende fino ad energie anche molto più grandi di quelle osservate da Auger. A queste energie altissime, che arrivano fino ad oltre 10^{20} eV il disco delle particelle

secondarie che arriva al suolo ha un raggio di alcuni chilometri. Questo permette di osservare uno sciame distribuendo dei rivelatori a grande distanza l'uno dall'altro, permettendo così di costruire un grande rivelatore con dei costi accettabili.

Il telescopio per raggi cosmici di energia più grande oggi in operazione si trova in Argentina, ed è chiamato Osservatorio Pierre Auger in onore del pioniere di cui abbiamo parlato, e vi lavorano i fisici di molti paesi di tutto il mondo, compresa l'Italia, che ha anzi avuto un ruolo di primo piano nel suo disegno e costruzione. (Una descrizione più approfondita dell'Osservatorio Pierre Auger è presentata dall'articolo di Sergio Petrer e Francesco Salamida in questo numero di Ithaca). L'osservatorio si estende su un'area di più di 3000 chilometri quadrati di pampa, dove sono distribuiti dei rivelatori di particelle ad una distanza di 1.5 chilometri l'uno dall'altro¹². Quando arriva uno sciame sufficientemente grande (e quindi generato da una particella primaria di energia sufficientemente elevata) un certo numero dei rivelatori registra un segnale proporzionale al numero di particelle che lo attraversano. I segnali generati dalle particelle di un singolo sciame nei diversi rivelatori sono quasi, ma non esattamente simultanei, questo perché le particelle secondarie arrivano al suolo come un disco dello spessore di qualche metro che si muove alla velocità della luce, così che le differenze nei tempi di arrivo in diversi punti permette di ricostruire l'orientamento del disco, e quindi la direzione del raggio cosmico primario. I segnali di tutti i rivelatori colpiti dallo sciame vengono inviati ad un computer di presa dati che li combina, per ricostruire direzione, energia, e (sia pure con qualche incertezza) il tipo di particella del primario che ha generato lo sciame.

L'Osservatorio Pierre Auger utilizza anche un altro ingegnoso metodo per misurare gli sciami dei raggi cosmici di altissima energia, quello della luce di fluorescenza. Quando particelle cariche relativistiche si propagano nell'atmosfera, queste possono non solo ionizzare gli atomi del mezzo, ma anche depositare energia nelle molecole dell'aria eccitando i loro involu-

¹²In un'area più ristretta i rivelatori sono più vicini, per permettere lo studio anche del flusso di particelle ad energia minore che generano sciami più piccoli

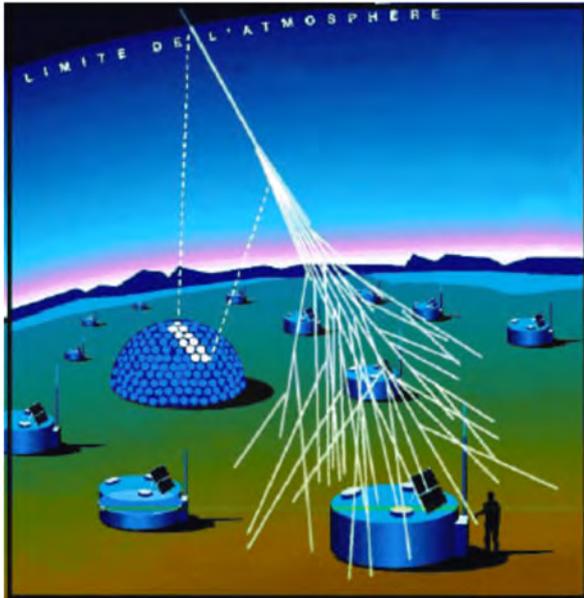


Figura 9: Schema del Pierre Auger Observatory in Argentina.

cri di elettroni. Le molecole eccitate ritornano quasi istantaneamente nello loro stato normale (o fondamentale) emettendo quella che viene chiamata luce di fluorescenza. Il passaggio nell'atmosfera di uno sciame di alta energia visto da lontano appare dunque come un sorgente di luce che si muove di moto rettilineo nel cielo alla velocità della luce. La sorgente prima cresce in brillantezza, mentre il numero di particelle aumenta esponenzialmente, raggiunge un massimo ed inizia poi ad affievolirsi. Un macchina da presa cinematografica che guarda verso il cielo (anzi meglio due o più di due per avere una visione stereoscopica) può filmare questi lampi di luce che attraversano il cielo, e lo studio delle immagini permette di nuovo di ricostruire direzione, energia e identità delle particelle primarie. Questa tecnica ingegnosa è possibile solo nelle notti senza luna (e senza nuvole che nasconderebbero i lampi generati dagli sciami) quando lo sfondo del cielo è sufficientemente scuro, ma nonostante questa limitazione è di grandissima utilità, ed ha permesso di ottenere misure molto importanti. L'Osservatorio Pierre Auger è considerato un telescopio per raggi cosmici ibrido, perché può utilizzare simultaneamente le due tecniche (dei rivelatori di superficie e della luce di fluorescenza) e questo è molto utile per limitare errori sistematici e migliorare la qualità delle misure.

La ricostruzione dell'energia, ed ancor più dell'identità delle particelle primarie a partire dalle osservazioni indirette degli sciami atmosferici richiede di avere una descrizione completa e dettagliata dello sviluppo di questi sciami, che è determinato dalle proprietà delle interazioni di alta energia. L'intima relazione tra l'astrofisica dei raggi cosmici e la fisica delle particelle elementari, continua quindi ad essere molto importante. Le conoscenze sono fluite e continuano a fluire in tutte e due le direzioni. Per alcune decadi la fisica sperimentale delle interazioni ad alta energia è stata possibile solo grazie ai raggi cosmici. Poi la costruzione di acceleratori di particelle ha permesso di studiare in laboratorio le interazioni di alta energia con grande precisione, e quindi di sviluppare modelli più accurati per lo sviluppo degli sciami atmosferici poi utilizzati per interpretare i dati dei telescopi per raggi cosmici. Le particelle di energia più elevata mai osservate sono però ancora quelle generate dalla Natura che ci arrivano dal cielo.

L'origine dei raggi cosmici

Se i fisici sono riusciti a fare rapidi progressi nel comprendere la natura dei raggi cosmici, e a ricostruire in buona approssimazione il loro spettro di energia, il problema della loro origine è rimasto a lungo senza soluzione, ed ancora oggi molte domande non hanno una risposta certa.

Il problema dell'origine dei raggi cosmici è così difficile perché le loro traiettorie durante la propagazione dalla sorgente alla Terra sono incurvate dai campi magnetici. La direzione con cui una particella è osservata non dice nulla sul punto del cielo dove si trova la sorgente che la ha emessa.

La situazione è opposta a quella della luce (o più in generale con i fotoni di tutte le lunghezze d'onda). Quando guardiamo il cielo questo ci appare puntellato di stelle, galassie ed altri oggetti astrofisici. Questo è perché la luce (o meglio i fotoni, che hanno carica elettrica nulla, e dunque non sono sensibili ai campi magnetici) si muove in linea retta. La distribuzione angolare dei fotoni forma dunque un'immagine con la posizione delle sorgenti sulla sfera celeste. L'immagine del cielo ottenuta con i raggi cosmici è invece quasi perfettamente uniforme, ed il numero di particel-

le che arriva alla Terra da ogni punto della sfera celeste è quasi esattamente uguale¹³.

Le prime teorie sui raggi cosmici, quando ancora si pensava che la radiazione primaria fosse formata da particelle neutre, interpretavano l'isotropia dei raggi cosmici come evidenza del fatto che fossero prodotti uniformemente in tutto lo spazio, e dunque di natura extragalattica o cosmologica, e queste idee sono riflesse nel nome. Ma una volta dimostrato che le particelle primarie sono cariche, l'isotropia può essere spiegata come la conseguenza dell'effetto dei campi magnetici che rimescolano le direzioni delle particelle, e diverse interpretazioni diventano possibili.

Alcuni fisici, ad esempio Walter Baade e Fritz Zwicky continuarono a pensare che i raggi cosmici avessero un'origine extragalattica e riempissero uniformemente tutto l'universo. Questa ipotesi richiede l'esistenza di una sorgente straordinariamente potente per riuscire a generare una tale gigantesca quantità di energia sotto forma di particelle relativistiche. Baade e Zwicky fecero allora una proposta ardita.

Gli astronomi conoscevano da diversi secoli il raro fenomeno dell'apparizione nel cielo di una nuova stella (in latino *Stella Nova* o più in breve di una *Nova*), che rimaneva visibile per qualche giorno o qualche settimana, per poi spegnersi lentamente. I due fisici, nel 1929, si resero conto che esistevano due classi di Novae di tipo completamente diverso, e in una delle classi le esplosioni, a cui diedero il nome di *Supernovae*, erano enormemente più luminose¹⁴. Baade

¹³Alle energie più basse ci sono anisotropie generate dal campo magnetico terrestre, discusse in precedenza e che hanno rivelato il fatto che i raggi cosmici hanno carica elettrica positiva. Gli effetti del campo geomagnetico (che sono comunque calcolabili) diminuiscono velocemente con la rigidità magnetica delle particelle, e le anisotropie rimangono dell'ordine di una parte per mille, fino ad $E \simeq 10^{18}$ eV, e anche alle energie più elevate rimangono sotto il 10%.

¹⁴Oggi il termine *Nova* è usato dagli astronomi per indicare solo la classe dei transienti meno luminosi. Si tratta di esplosioni che avvengono sulla superficie di stelle del tipo nana bianca che sono in un sistema binario "stretto" con una seconda stella di tipo gigante rossa relativamente vicina. In questi sistemi, la nana bianca, grazie alla forza di gravità, sottrae materia alla stella compagna. Il materiale si accumula sulla sua superficie, e quando raggiunge una massa critica si può innescare una reazione di fusione nucleare che libera rapidamente energia e genera il lampo osservabile dalla Terra. Queste esplosioni possono ripetersi molte volte, e si conoscono infatti

e Zwicky fecero la speculazione (che si è poi rivelata corretta) che le Supernovae fossero dovute al collasso gravitazionale di una stella ed alla formazione di una a stella di neutroni. Al centro delle stelle sufficientemente massicce, alla fine della loro vita, si forma un nocciolo di ferro che cresce grazie alle reazioni di fusione nucleare. Quando il nocciolo supera una massa critica (del valore di circa 1.4 masse solari, nota come massa di Chandrasekhar, il nome del fisico indiano che ha compreso per primo la sua esistenza) inizia a collassare sotto il suo stesso peso. Quando la materia del nocciolo raggiunge e supera la densità nucleare (cioè quando il suo raggio si è contratto da 5000 chilometri a circa 10), la sua compressione genera un gigantesco rimbalzo, un'onda d'urto che si propaga verso l'esterno, lanciando nello spazio alla velocità di migliaia di chilometri al secondo gli strati esterni della stella. Questi eventi straordinari hanno avuto per noi un'importanza letteralmente vitale, perché la materia della stella lanciata nello spazio interstellare dall'esplosione è formata non solo da idrogeno ed elio (che costituiscono la quasi totalità della materia primigenia che emerge dal Big Bang) ma contiene anche tutti gli altri elementi (carbonio, ossigeno, ferro . . .) che sono stati sintetizzati dalle reazioni di fusione nucleare all'interno della stella, e che poi formano la nostra Terra, ed i nostri stessi corpi¹⁵. I frammenti della stella esplosa si disperdono nello spazio ma dopo essersi mescolati con il resto del gas interstellare possono poi ricondensarsi per formare nuovi sistemi stellari, simili a quello solare con i suoi pianeti e la nostra Terra. La conclusione è che la materia che forma la Terra ed i nostri corpi si deve essere trasformata nel centro di una (o forse più di una) stella, che è poi violentemente esplosa alla fine della sua vita. Siamo letteralmente "polvere di stelle".

diverse *Novae ricorrenti*. Al contrario, le *Supernovae* sono esplosioni molto più potenti che distruggono completamente un'intera stella, e che quindi non si possono ripetere.

¹⁵Per maggiore completezza, nei noccioli interni delle stelle vengono formati gli elementi fino al ferro, in reazioni di fusione nucleare che rilasciano energia. Gli elementi più pesanti sono formati in processi di nucleosintesi esplosiva quando i nuclei (già formati nel centro delle stelle) sono bombardati da flussi di neutroni, che avvengono in parte nelle Supernovae ed in parte nelle coalescenze di stelle compatte che verranno discusse in seguito.

Baade e Zwicky osservarono che l'energia rilasciata durante la formazione di una stella di neutroni durante il collasso gravitazionale, secondo i loro calcoli, era molto più grande di quella visibile nella forma di energia cinetica dei frammenti lanciati nello spazio, ed ipotizzarono che l'energia mancante venisse emessa (con un meccanismo sconosciuto) nella forma di raggi cosmici. Le Supernovae sono molto rare (un paio per Galassia per secolo), ma sono così energetiche che sommando i contributi di tutte le esplosioni durante la storia dell'Universo a partire dal Big Bang, diventa possibile riempire uniformemente di particelle relativistiche l'intero Universo.

Oggi sappiamo che le considerazioni di Baade e Zwicky sull'origine delle *Supernovae* da collasso gravitazionale sono corrette, ma anche che la maggior parte dell'energia rilasciata in un collasso gravitazionale viene emessa non come raggi cosmici ma sotto forma di neutrini in un lampo della durata di una decina di secondi. Questa non è solo un'ipotesi teorica, perché ha ricevuto una spettacolare conferma sperimentale con l'osservazione dell'emissione di neutrini dalla supernova SN1987A esplosa il 23 febbraio del 1987 nella Grande Nube di Magellano (una galassia satellite della Via Lattea). La massima energia disponibile per generare raggi cosmici dalle *Supernovae* è quindi migliaia di volte più piccola di quella ipotizzata da Baade e Zwicky, ma come discuteremo tra poco, oggi sappiamo che i raggi cosmici non riempiono uniformemente l'Universo, e dunque richiedono una sorgente meno potente. Perciò è possibile, ed è anzi l'ipotesi considerata oggi più probabile, che le *Supernovae* siano in effetti la principale sorgente di raggi cosmici.

Nel 1949 tre fisici Hannes Alfvén, Robert Richtmyer ed Edward Teller formularono un'ipotesi completamente diversa, quasi opposta a quella di Baade e Zwicky, e cioè che i raggi cosmici hanno un'origine locale, sono generati dal Sole e confinati ad una regione intorno al sistema solare. Questa idea nasceva dagli studi di Alfvén che aveva capito che lo spazio interplanetario non è vuoto, ma contiene un gas rarefatto in parte in forma di plasma, dove possono esistere correnti elettriche che generano un campo magnetico caotico, sufficiente a rimescolare ed isotropizzare le traiettorie di particelle cariche.

Nello stesso anno uscì un importantissimo e

oggi celebrato articolo di Enrico Fermi, in cui veniva fatta l'ipotesi in un certo senso intermedia, e cioè che i raggi cosmici osservati alla Terra fanno parte di una popolazione di particelle che riempie la nostra Galassia, mentre lo spazio intergalattico ne conterrebbe una densità molto più piccola. Il confinamento delle particelle cariche nella Galassia sarebbe dovuto ai campi magnetici generati dalle correnti elettriche che attraversano il sottile plasma che riempie il mezzo interstellare. Questa ipotesi, come vedremo in seguito, ha avuto poi una chiara conferma sperimentale. Il confinamento dei raggi cosmici però non può essere permanente, in parte perché la bottiglia magnetica non è perfettamente ermetica, e poi perché ad ogni modo le particelle possono interagire con il gas interstellare, e se non fuggono nello spazio extragalattico verrebbero comunque assorbite da queste interazioni dopo avere viaggiato per un tempo dell'ordine di 10–100 milioni di anni. Questo tempo può sembrare lungo ma è invece corto rispetto alla vita della Galassia, e dunque i raggi cosmici devono essere continuamente rinnovati da sorgenti nella nostra Galassia. Fermi si pose quindi il problema di quale potesse essere il meccanismo capace di accelerare delle particelle ad altissime energie, e nel suo lavoro propose un metodo noto oggi come accelerazione di Fermi, che, anche se in una forma modificata, è ancora oggi una base fondamentale per le teorie sull'origine dei raggi cosmici.

Abbiamo visto come un elemento fondamentale nelle idee di Fermi (introdotto già da Hannes Alfvén) è che lo spazio interstellare contiene un gas rarefatto dove delle correnti elettriche generano un campo magnetico disordinato che curva e rimescola le traiettorie delle particelle cariche. Ora Fermi ipotizzò anche che il gas interstellare non è in quiete, ma è formato da nuvole che si muovono nello spazio in diverse direzioni. Ora, per un osservatore che vede una nuvola magnetizzata muoversi con velocità non nulla, il campo magnetico appare nel suo sistema di riferimento come un campo elettrico, che è dunque capace di accelerare (o anche decelerare) particelle dotate di carica elettrica. Un semplice calcolo mostra che in media l'incontro di una particella carica con una nuvola aumenta la sua velocità¹⁶, e che

¹⁶Come analogia, si può immaginare di lanciare una palla

ripetuti incontri con queste nuvole magnetizzate possono fornire ad una particella energie anche elevatissime, perché la sua energia cresce esponenzialmente, ed un oggetto macroscopico come una nuvola di gas possiede una energia enorme (praticamente infinita) rispetto ad una singola particella.

Fermi ipotizzò dunque che i raggi cosmici sono accelerati continuamente e gradualmente durante la propagazione nella Galassia da questo meccanismo associato ai movimenti del plasma del mezzo interstellare. L'accelerazione è anche bilanciata da perdite (Fermi ipotizzò come meccanismo dominante l'assorbimento per interazioni con il gas interstellare, ma più in generale si deve considerare anche la fuga dal confinamento). Dall'equilibrio tra accelerazione e perdite emerge uno spettro dei raggi cosmici in buona approssimazione costante (o più precisamente stazionario) della forma di legge di potenza osservata sperimentalmente.

Oggi sappiamo che molte cose nella proposta di Fermi non funzionano, ma alcune delle sue idee rimangono di importanza fondamentale. La principale differenza tra la proposta originale del 1949 di Fermi e le teorie discusse oggi, è che si pensa che i raggi cosmici siano accelerati non da un meccanismo diffuso in tutta la Galassia, ma in sorgenti astrofisiche ben definite nello spazio ed anche nel tempo. Il meccanismo di accelerazione è uno sviluppo di quello proposto da Fermi, nel senso che le particelle ricevono sempre la loro energia interagendo con un plasma magnetizzato in movimento, ma questo avviene all'interno di certi sistemi (o meglio eventi) astrofisici.

Nelle *Supernovae* l'esplosione della stella è generata da un'onda d'urto sferica che parte dal nocciolo della stella progenitrice, compresso a densità nucleare, ne attraversa gli strati più esterni e si propaga nello spazio circumstellare con una velocità iniziale di oltre diecimila chilometri al secondo, trascinandosi dietro la materia della stella. Secondo la teoria (che è considerata come ben motivata, ma che ancora non ha una conferma osservativa indiscutibile) particelle con carica elettrica possono essere accelerate

nella pista di autoscontro di un luna park. Ripetuti scontri con le automobili in moto sulla pista accelerano gradualmente la palla a velocità sempre più alte, fino a che la palla non vola fuori dalla pista, o raggiunge la stessa energia cinetica media delle automobili.

dai campi elettromagnetici trasportati dal plasma, ed aumentare gradualmente la loro energia ricevendo una spinta ogni volta che attraversano l'onda d'urto in espansione. Questo meccanismo è conosciuto in inglese come *Fermi diffusive shock acceleration* che possiamo tradurre come accelerazione diffusiva di Fermi in presenza di onde d'urto.

Due argomenti teorici danno supporto all'idea che la maggior parte dei raggi cosmici sia generata nelle esplosioni di *Supernova*. Il primo è basato su considerazioni di bilancio energetico. Sappiamo (sia pure con un'incertezza non trascurabile) che i raggi cosmici possono rimanere confinati nella Galassia per un tempo dell'ordine di alcune decine di milioni di anni. Usando questa stima, la densità osservata dei raggi cosmici alla Terra (che estrapoliamo anche ad altri punti della nostra Galassia) ed il volume della Via Lattea, è possibile stimare quale è la luminosità dell'insieme delle sorgenti necessaria a rigenerare le particelle dei raggi cosmici. Il risultato, espresso in luminosità solari e dell'ordine di circa dieci milioni di soli (che però emettono energia non come luce visibile, ma come particelle cariche relativistiche). Non è facile trovare (o immaginare) una sorgente astrofisica così potente, ma le esplosioni di *Supernova* sono una possibilità. L'energia cinetica dei frammenti di stella lanciati nello spazio è infatti grandissima (dieci masse solari lanciate alla velocità di alcune migliaia di chilometri al secondo corrispondono a circa 10^{51} erg). Gli eventi di *Supernova* nella nostra Galassia sono molto rari, circa una o due per secolo, ma se una frazione dell'ordine del 10% dell'energia cinetica dell'esplosione viene trasformata in raggi cosmici, il totale corrisponde proprio (in approssimazione ragionevolmente buona) a quello richiesto.

Il secondo argomento teorico è invece di dinamica dell'accelerazione, ed è basato sul fatto che calcoli teorici sull'accelerazione di raggi cosmici nelle *Supernovae* basati sul meccanismo di Fermi predicono una distribuzione di energia che ha la forma di una legge di potenza, con un esponente che, dopo avere considerato gli effetti di propagazione nella Galassia, è consistente con quello osservato.

È chiaramente possibile immaginare l'esistenza di altre sorgenti di raggi cosmici che non siano

le *Supernovae*, e le idee generali dell'accelerazione diffusiva di Fermi possono essere applicate anche ad altri eventi astrofisici, in particolare a quelli esplosivi dove grandi masse sono violentemente messe in movimento. Diverse recenti osservazioni hanno in effetti dimostrato che l'accelerazione di particelle cariche ad energie relativistiche è un fenomeno che avviene in diverse classi di oggetti astrofisici.

L'Astrofisica "multi-messaggeri"

Come è possibile identificare una sorgente di raggi cosmici visto che le particelle che osserviamo qui sulla Terra arrivano con una direzione che non dice nulla sulla loro origine? Determinare la presenza di una popolazione di particelle relativistiche in una regione lontana dello spazio è possibile grazie al fatto che una frazione forse piccola, ma comunque non nulla, di queste particelle interagirà con la materia o con i campi di radiazione presenti nello spazio, e da queste interazioni emergono secondari neutri: fotoni e neutrini che si propagano in linea retta, e possono essere osservati in telescopi sulla Terra, formando un'immagine della sorgente.

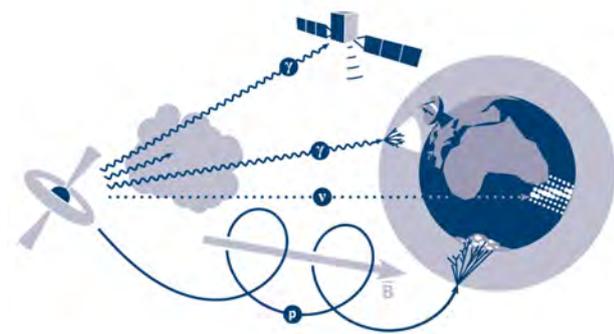


Figura 10: Il concetto dell'astronomia multimessaggera. Una sorgente emette raggi cosmici, raggi gamma e neutrini.

È facile descrivere queste emissioni di fotoni e neutrini perché i processi sono essenzialmente gli stessi che generano la radiazione secondaria a partire dalle interazioni dei raggi cosmici con l'atmosfera della Terra. Le interazioni di protoni e neutroni generano un grande numero di pioni (le particelle π^+ , π^- e π^0 formate dalle combinazioni $q\bar{q}$ dei due quark più leggeri up e down). I pioni si disintegrano spontaneamente in un

tempo molto breve. Il decadimento di un pione neutro crea due fotoni di alta energia, mentre il decadimento di un pione carico, crea tre neutrini, uno creato direttamente insieme ad un muone, e due dal decadimento a catena di questo muone¹⁷. Il risultato è che una popolazione di protoni e nuclei di alta energia presenti in un acceleratore astrofisico, genera flussi di fotoni e neutrini, approssimativamente di uguale intensità, proporzionali sia al numero di particelle presenti che alla densità del mezzo in cui si propagano.

Se nella sorgente sono presenti elettroni (o positroni) relativistici, esistono anche altri meccanismi di emissione di fotoni (ma non di neutrini). Un meccanismo è quello chiamato radiazione di sincrotrone dovuto al fatto che le particelle con carica elettrica sono deviate dalla presenza di campi magnetici, e l'accelerazione di una carica è sempre accompagnato dalla radiazione elettromagnetica¹⁸. Un secondo meccanismo di emissione è quello dell'effetto Compton inverso (dal nome del fisico americano Arthur Compton) cioè la collisione elastica tra elettroni e fotoni. In queste collisioni elettroni relativistici possono cedere la loro energia a fotoni di bassa energia, come quelli ottici o della radiazione di fondo cosmico di 2.7 Kelvin (il residuo del Big Bang) e generare raggi gamma¹⁹.

L'osservazione di raggi gamma (e più in generale di fotoni in diverse regioni dello spettro) e di neutrini permette di determinare in quali regioni dello spazio sono presenti particelle cariche di alta energia (cioè raggi cosmici), e se si conoscono sufficientemente bene le proprietà del sito di emissione (che determinano il bersaglio delle interazioni) e dello spazio tra la sorgente e la Terra

¹⁷Una buona parte dei μ^\pm creati nell'atmosfera della Terra arriva fino al suolo (formando una parte importante della radiazione cosmica secondaria), e le particelle decadono da ferme dopo avere perso tutta la loro energia cinetica in ionizzazione del suolo. In un ambiente astrofisico, la densità è in genere molto piccola, ed i muoni decadono prima di avere perso una quantità apprezzabile di energia.

¹⁸La potenza della radiazione emessa è proporzionale al quadrato dell'accelerazione. Per la stessa forza dunque l'accelerazione è molto più grande per elettroni che per protoni, perché la loro massa è circa duemila volte più piccola. La radiazione di sincrotrone è dunque importante solo per elettroni e positroni.

¹⁹Anche in questo caso la probabilità di interazione dipende dalla massa della particella carica, ed è significativa solo per particelle leggere come elettroni e positroni.

(dove è possibile che il flusso di fotoni –ma non di neutrini – sia parzialmente assorbito) diventa possibile ricostruire l'entità e la distribuzione di energia di queste popolazioni di raggi cosmici. Per questo motivo sono stati disegnati e costruiti speciali telescopi sia per raggi gamma che per neutrini.

Da queste considerazioni emerge che per conoscere le proprietà delle sorgenti di alta energia è utile, anzi necessario, combinare osservazioni che utilizzano tre diversi tipi di particelle, tre diversi *messaggeri*: raggi cosmici, fotoni e neutrini.

Esiste anche un quarto tipo di messaggero che può essere di fondamentale importanza nello studio delle sorgenti di alta energia: le onde gravitazionali. Queste onde sono una delle più importanti predizioni della teoria della Relatività Generale di Albert Einstein, uno dei pilastri fondamentali della fisica moderna. Questa teoria è stata riassunta da John Wheeler in due brevi frasi:

“La materia dice allo spazio-tempo come curvarsi. Lo spazio-tempo dice alla materia come muoversi”.

Come scoperto da Galileo osservando la caduta di oggetti dalla cima della torre di Pisa, corpi diversi si muovono sotto l'effetto della gravità con identica accelerazione. Nella teoria di Einstein questo si spiega con il fatto che la materia si muove in un campo gravitazionale lungo traiettorie, le geodesiche, determinate unicamente dalla curvatura dello spazio-tempo²⁰. La curvatura dello spazio-tempo è determinata non semplicemente dalla massa (come nella teoria della gravità di Newton) ma dalla massa e dal suo movimento (o più tecnicamente dal tensore energia-impulso).

Le onde gravitazionali emergono perché quando una massa accelera, genera una increspatura nello spazio tempo che si propaga alla velocità della luce, un modo analogo alle onde circolari

²⁰La geodesica è la generalizzazione in uno spazio curvo del concetto di linea retta. Ad esempio le geodesiche su una sfera sono gli archi di cerchio massimo. È interessante notare come la separazione tra due punti che si muovono a velocità costante lungo due differenti archi non varia nel tempo in modo costante, ma accelera. La curvatura dunque può apparire come una forza attrattiva.

sulla superficie dell'acqua che partono dal punto dove abbiamo gettato un sasso. L'effetto è straordinariamente piccolo, ed è impossibile generare onde gravitazionali osservabili in esperimenti di laboratorio. Esistono però eventi astrofisici (ne discuteremo alcuni tra breve) dove grandissime masse hanno elevatissime accelerazioni, arrivando rapidamente a velocità vicine a quella della luce. Questi eventi, che anche se avvengono a distanze enormi dalla Terra, possono generare effetti, che sebbene minutissimi sono osservabili in speciali antenne costruite a questo scopo. L'osservazione delle onde gravitazionali è stata finalmente ottenuta da pochi anni, e per questa scoperta il premio Nobel per la fisica del 2017 è stato assegnato a Rainer Weiss, Barry Barish e Kip Thorne.

La relazione tra le onde gravitazionali e le emissioni di alta energia (nella forma di raggi cosmici, fotoni e neutrini) può essere compresa in modo semplice. L'accelerazione di particelle cariche ad energie relativistiche richiede la presenza di campi elettrici sufficientemente elevati. Questi campi elettrici si formano quando della materia si muove ad alta velocità. Abbiamo discusso del caso delle esplosioni di *Supernova*, dove l'involucro esterno della stella viene violentemente accelerato dall'onda d'urto di rimbalzo nella formazione di una stella di neutroni, ma esistono anche altri casi²¹. Una condizione qualitativa ma generale delle condizioni per le quali si generano particelle di alta energia si può riconoscere nell'aggettivo *violento*: grandi masse acquistano violentemente grandi velocità, e quindi subiscono grandi accelerazioni. Ma grandi masse e grandi accelerazioni sono esattamente le condizioni che portano all'emissione di onde gravitazionali.

Sulla base di queste considerazioni ci si può dunque attendere di osservare delle coincidenze, ovvero che ad una parte degli eventi che generano onde gravitazionali sia associata l'emissione di particelle (raggi cosmici, fotoni e neutrini) di alta energia, e viceversa che almeno alcune

²¹È da notare che stiamo parlando di due accelerazioni di tipo diverso. Nel caso delle *Supernovae* l'intera massa degli strati esterni della stella è accelerata fino ad una velocità di migliaia di chilometri al secondo. Il plasma in movimento può in seguito accelerare una (relativamente piccola) parte delle particelle fino ad velocità vicine a quella della luce.

delle sorgenti di alta energia siano associate ad eventi nei quali sono emesse onde gravitazionali. Questa è la base della **Astronomia a multi-messaggeri**, che studia quello che possiamo chiamare l'Universo delle Alte Energie. Questa nuova astronomia ha cominciato a fare i suoi primi passi con le prime osservazioni che mettono insieme neutrini e raggi gamma e, in uno spettacolare evento del 17 agosto del 2017, onde gravitazionali e raggi gamma.

Telescopi per raggi gamma

I fotoni di alta energia (o raggi gamma) sono assorbiti nell'aria dopo avere attraversato solo un piccolo spessore (circa il tre per cento dell'atmosfera), e dunque come per i raggi cosmici, la osservazione diretta richiede di posizionare il telescopio al di sopra dell'atmosfera. I primi telescopi per raggi gamma sono stati messi in orbita negli anni '60. Tra il 1991 ed il 2000 il satellite CGRO (Compton Gamma Ray Observatory) [4] con a bordo quattro differenti strumenti ha ottenuto risultati di grande interesse. Dal 2008 è in orbita il Fermi Gamma Ray Telescope [5] chiamato in questo modo per onorare Enrico Fermi ed i suoi studi sull'accelerazione delle particelle. Questo telescopio ha rivelato più di seimila differenti sorgenti di alta energia e fatto fare grandi progressi alla nostra comprensione dell'Universo delle Alte Energie.



Figura 11: Il telescopio per raggi gamma FERMI, lanciato in orbita l'11 giugno 2008.

I telescopi su satellite sono troppo piccoli per osservare i flussi di raggi gamma ad energie più grandi di circa 10^{12} eV, ma come per i raggi cosmici, per energie sufficientemente elevate diventa possibile misurare indirettamente i raggi gamma osservando gli sciami che generano nell'atmosfera. Risultati importanti sono stati ottenuti grazie

ai telescopi che utilizzano la tecnica della luce Čerenkov. Questa luce, chiamata con il nome del fisico (e premio Nobel) sovietico Pavel Čerenkov che l'ha scoperta nel 1934, è emessa in aria in un cono di angolo circa 1.4 gradi rispetto alla direzione della velocità da particelle relativistiche con velocità $v > 0.9997 c$. Gli sciami di raggi gamma in un ampio intervallo di energia sono assorbiti prima di arrivare a terra, ma al suolo arriva la luce Čerenkov emessa dalle particelle dello sciame sotto forma di un lampo che è osservabile (durante le notti senza luna e senza nuvole) da telescopi disegnati a questo scopo. Dalla forma, orientamento e luminosità dei lampi di luce Čerenkov è possibile selezionare sciami generati da raggi gamma (riducendo il fondo di sciami generati dalle particelle cariche dei raggi cosmici) e determinare direzione ed energia della particella primaria. La tecnica di osservazione dei telescopi Čerenkov si è raffinata nelle ultime decadi, costruendo telescopi con specchi più grandi al fine raccogliere più luce, fotocamere con maggiore risoluzione, e sviluppando il metodo stereoscopico dove uno stesso sciame è osservato da diversi telescopi, per ricostruire meglio le proprietà dei raggi gamma che li generano. Telescopi Čerenkov come HESS, MAGIC e VERITAS [6] hanno identificato alcune centinaia di sorgenti di raggi gamma.



Figura 12: Il telescopio Čerenkov MAGIC.

Ad energie ancora più elevate i raggi gamma possono essere osservati da rivelatori che osservano gli sciami estesi, come quello di LHAASO in Cina [7].



Figura 13: Il telescopio per raggi cosmici e raggi gamma LHAASO si trova in Cina alla quota di 4300 metri.

Telescopi per Neutrini

I neutrini interagiscono molto debolmente, e questo rende la loro osservazione molto difficile; d'altra parte questa loro proprietà li rende molto attraenti come messaggeri astrofisici perché possono emergere da luoghi che sono opachi all'emissione di fotoni. Tre esempi spettacolari sono:

1. I neutrini solari, generati dalle reazioni di fusione nucleare nel Sole.
2. I neutrini generati dalla Supernova SN1987A (del 23 febbraio 1987), che sono stati emessi in un lampo incredibilmente potente e della durata di solo una decina di secondi dall'interno della stella di neutroni appena formato dal collasso gravitazionale della stella progenitrice.
3. I neutrini geofisici emessi dalle sostanze radioattive all'interno della Terra.

Le tre classi di neutrini hanno energie dell'ordine di 1–10 MeV, e sono state osservate con rivelatori disegnati a questo scopo (e due premi Nobel sono stati assegnati per i neutrini solari e quelli della Supernova del 1987). L'osservazione dei neutrini associati all'Universo delle alte energie richiede rivelatori di diverso disegno, e di massa grandissima. Per fare questo è nata l'idea di allestire con rivelatori un chilometro cubo di acqua e di ghiaccio.

Il più grande rivelatore di neutrini esistente oggi al mondo è stato costruito al Polo Sud e si chiama IceCube. Il nome cubo di ghiaccio descrive perfettamente questo telescopio. Al polo

Sud si incontrano condizioni uniche. La superficie dell'Antartide è qui coperta da uno strato di ghiaccio dello spessore di due chilometri e mezzo. I fisici che hanno progettato e costruito IceCube [8] hanno distribuito nel ghiaccio circa cinquemila rivelatori di luce, alle profondità tra 1500 e 2500 metri²². Costruire un rivelatore di queste dimensioni al Polo Sud ha rappresentato uno sforzo straordinario, che però ha dato dei risultati affascinanti. Le osservazioni di IceCube hanno rivelato che esistono sorgenti extragalattiche che emettono neutrini con energie altissime (tra 10^{14} e 10^{17} eV).

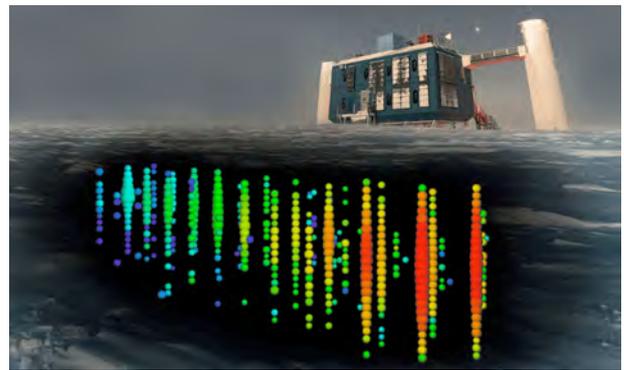


Figura 14: Il telescopio per neutrini IceCube al Polo Sud.

L'interesse per nuovi telescopi di neutrini è molto forte, ed esistono progetti sia per l'ampiamiento di IceCube, sia per la costruzione di altri telescopi per neutrini, basati sulla stessa concezione, ma anche su altri principi. Un progetto sul quale fisici italiani stanno lavorando intensamente è la costruzione di un rivelatore (KM3NeT) di struttura simile ad IceCube ma nel profondo del mar Mediterraneo, vicino alle coste della Sicilia [9].

²²I rivelatori sono dentro sfere del diametro di circa 30 cm, attaccati lungo un cavo, come le perle di una collana, ad una distanza di circa quindici metri l'uno dall'altro. Per posizzarli nel ghiaccio, i fisici fanno un buco di due chilometri e mezzo, scavato utilizzando un getto di acqua calda. Il cavo (che serve a trasmettere i dati dei rivelatori) con le sue perle viene poi calato nel buco. Il ghiaccio in breve si riforma, ed i rivelatori sono per sempre imprigionati nel ghiaccio, ma i segnali che ricevono possono essere trasmessi attraverso il cavo ai computer di acquisizione dati, e rivelare le interazioni dei neutrini.

L'Universo delle Alte Energie

L'Astronomia multimessaggera è un campo in rapido sviluppo, e gli ultimi anni hanno visto molti nuovi risultati e molte sorprese. Una discussione dettagliata di questi sviluppi richiederebbe più spazio di quello che abbiamo a disposizione qui, e tra l'altro rischierebbe di essere presto in parte obsoleta, perché è molto probabile che nuove osservazioni costringeranno i fisici a rivedere le loro teorie (come è avvenuto molte volte nel passato in questo campo). Nel seguito ci limiteremo per questo solo ad alcune osservazioni generali.

I telescopi per raggi gamma hanno ottenuto risultati di un interesse che senza esitazioni si può chiamare straordinario. I raggi gamma sono generati dalle interazioni dei raggi cosmici e viaggiano in linea retta, quindi la loro distribuzione angolare forma una mappa che è la proiezione bidimensionale sulla sfera celeste dei siti dove sono presenti particelle di alta energia²³. I telescopi su satellite (AGILE e FERMI) hanno ottenuto delle mappe dell'intero cielo gamma. Queste mappe mostrano che il cielo visto in fotoni di alta energia ($E \gtrsim 10^8$ eV) è formato da diverse componenti. La prima è un flusso diffuso che segue la forma del disco della Galassia, con un massimo verso il suo centro. Questo flusso è generato dai raggi cosmici che interagiscono nello spazio interstellare della Galassia, e le osservazioni confermano l'idea che i raggi cosmici riempiono la Via Lattea, mentre la loro densità nello spazio extragalattico è molto inferiore. Una seconda componente è un insieme di sorgenti puntiformi o quasi-puntiformi, in parte Galattiche ed in parte extragalattiche che appartengono a diverse classi. Sullo sfondo esiste un flusso isotropo di raggi gamma, uguale da tutte le direzioni. Questo flusso è di natura extragalattica, e la sua isotropia riflette il fatto che l'Universo su scala abbastanza grande è omogeneo. Si pensa che questo flusso isotropo sia dovuto ad un insieme di sorgenti extragalattiche deboli e lontane e quindi non risolte e forse anche in parte dell'emissione dallo spazio extragalattico.

²³L'emissione di raggi gamma dipende anche dalle proprietà del mezzo dove si propagano i raggi cosmici (come densità del gas e campi di radiazione). E questo genera una incertezza nel ricostruire lo spettro e composizione dei raggi cosmici a partire dai flussi di gamma.

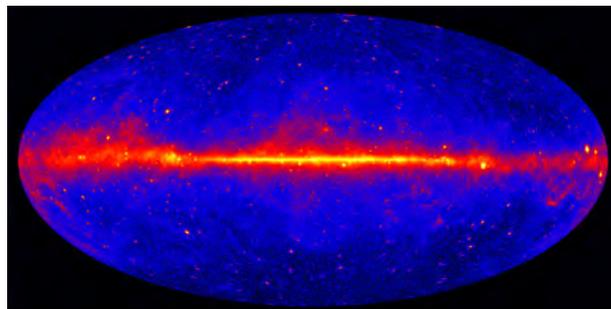


Figura 15: La sfera celeste è raffigurata sul piano (in una proiezione che conserva l'area). Il centro Galattico è al centro, ed il piano Galattico lungo la linea orizzontale che passa per il centro. Chiaramente visibile è il contributo dei raggi gamma emessi dalle interazioni dei raggi cosmici che interagiscono con il gas interstellare (che si concentra sul piano Galattico). Si possono anche intravedere molte sorgenti puntiformi.

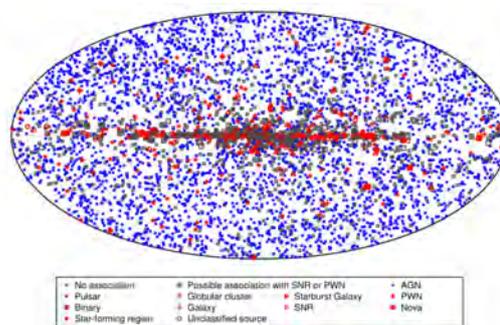


Figura 16: Sorgenti di raggi gamma rivelati dal telescopio FERMI. È evidente che esistono sia sorgenti Galattiche (presenti nel piano Galattico), e sorgenti extragalattiche (distribuite uniformemente sulla sfera celeste). Le sorgenti appartengono a diverse classi.

Una caratteristica di molte delle sorgenti gamma è la loro grande *variabilità*. Il flusso di molte sorgenti non è costante, ma varia nel tempo, con una struttura temporale in molti casi caotica, così che periodi di quiete si alternano a periodi dove la sorgente è luminosa (con intensità differenti, e per lunghezze di tempo differenti). In alcuni casi le variazioni sono molto rapide, e la luminosità può cambiare di molto in pochi minuti. Questa situazione è esattamente il contrario di come appare il cielo visto in luce ordinaria, dove le variazioni sono molto rare.

Alcune delle sorgenti più importanti ed affascinanti del cielo gamma, sono transienti molto rapidi, bagliori luminosissimi che però brillano nel cielo per un tempo brevissimo, di pochi secondi,

o anche solo di frazioni di secondo. Questi lampi di raggi gamma sono conosciuti in inglese come *Gamma Ray Bursts*.

“Supernova Remnants” e “Pulsars”

Molte delle sorgenti di raggi gamma nella nostra Galassia sono associate ai *Supernova Remnants* (Resti di Supernova), quelle sfere di gas in espansione che circondano il punto dove in un passato non troppo lontano è avvenuta la gigantesca esplosione di una stella.

Abbiamo discusso come gli astrofisici teorici hanno ipotizzato che la gran parte dei raggi cosmici siano accelerati dalle onde d’urto generate in queste esplosioni. Questo porta alla predizione che all’interno di Resti di Supernova giovani (con meno di qualche migliaio di anni) si debba trovare una gran quantità di raggi cosmici accelerati da poco (o che sono ancora in fase di accelerazione). La domanda è dunque: “questa predizione è effettivamente verificata oppure no?” La risposta si deve dare in due parti. La prima è che le osservazioni dimostrano che i Resti di Supernova giovani emettono raggi gamma, e dunque contengono raggi cosmici che sono stati accelerati dall’esplosione. Rimane però da dimostrare se le *Supernovae* sono veramente la principale sorgente dei raggi cosmici che osserviamo sulla Terra (e nello spazio interstellare). La risposta finale a questa domanda è ancora dibattuta.

Le esplosioni di *Supernova* sono in realtà generate da due tipi di meccanismo completamente differenti, ma che secondo le teorie (e oggi anche le osservazioni) possono entrambi accelerare raggi cosmici. Un meccanismo è quello dell’esplosione nucleare di una stella nana bianca in un sistema binario, evento che la distrugge completamente. Il secondo (che abbiamo già introdotto in precedenza) è quello del collasso gravitazionale del nocciolo interno di ferro di una stella alla fine della sua vita, il nocciolo si trasforma in una stella di neutroni, oppure (in casi più rari quando la stella progenitrice ha massa sufficientemente grande) in un buco nero. Nelle stelle di neutroni la materia è compressa alla densità dei nuclei atomici in una sfera del raggio di circa dieci chilometri, mille volte più piccola del nocciolo

iniziale²⁴. Il rimpicciolimento della stella implica che la sua rotazione diventa più veloce, di circa un fattore un milione, per lo stesso motivo (la conservazione del momento angolare) per cui una pattinatrice su ghiaccio accelera la sua rotazione quando raccoglie le braccia vicino al corpo. In questo modo un periodo di rotazione di circa un mese (come quello del Sole) si accorcerebbe ad un paio di secondi. Per lo stesso motivo, il campo magnetico alla superficie della stella viene amplificato di circa un milione di volte durante il collasso. La rapida rotazione ed i forti campi magnetici generano complicati (e ancora solo parzialmente compresi) fenomeni elettromagnetici intorno alla stella di neutroni con importanti conseguenze. Il primo fenomeno ad essere stato scoperto è l’emissione di pulsazioni in onde radio con un periodo uguale a quello delle rotazioni, come i segnali luminosi di un faro. La stella di neutroni diventa così una Pulsar.

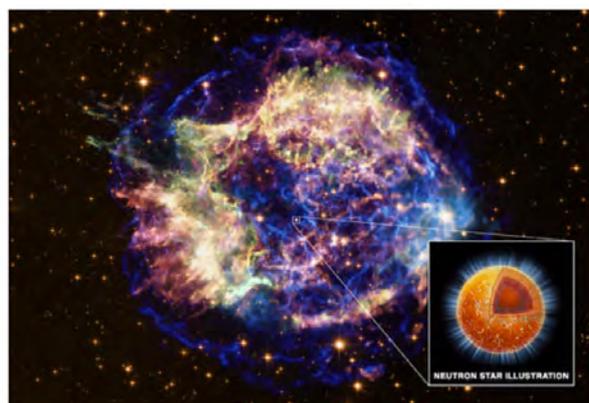


Figura 17: *Supernova Cassiopeia A* (l’immagine combina ottico e raggi X). Questa *Supernova* molto probabilmente è l’ultima osservata nella nostra Galassia, registrata il 16 luglio del 1680 dall’astronomo reale John Flamsteed. Si trova a 11mila anni luce dalla Terra, ed ha un diametro di 10 anni luce. Al centro si trova una sorgente puntiforme di raggi X: una stella di neutroni che si sta lentamente raffreddando.

La scoperta della prima Pulsar è stata fatta nel 1967 da Jocelyn Bell (utilizzando un sistema di telescopi radio disegnati da Martin Ryle

²⁴Nella materia a questa densità, è energeticamente favorevole per i protoni ed elettroni fondersi, emettendo neutrini (che poi fuggono dalla “proto-stella di neutroni”) con la reazione $e^- + p \rightarrow n + \nu_e$. È per questo che si parla di stelle di neutroni. Una piccola quantità di protoni ed elettroni rimane comunque dentro stella, e nella crosta di superficie.



Figura 18: *Crab Nebula* (l'immagine combina radio, ottico, e raggi X). La *Crab Nebula* è una Supernova esplosa nel 1054, alla distanza di 6500 anni luce dalla Terra. Al centro si trova una *Pulsar* con un periodo di 33 millisecondi. L'intricato aspetto della nebula è dovuto all'interazione di un vento di particelle emesse dalla *Pulsar* che interagisce con il materiale espulso dalla stella progenitrice nell'esplosione. Questa è la più luminosa delle *Pulsar Wind Nebulae*.

ed Anthony Hewish), che con sua grande sorpresa si rese conto che da un punto del cielo arrivava un ticchettio radio con un periodo di 1.336 secondi, di una brevità mai vista nella storia dell'astrofisica fino ad allora. Il fisico Thomas Gold suggerì subito che il segnale proveniva da una stella di neutroni. L'idea ebbe una spettacolare conferma quando circa un anno dopo una *Pulsar* con un periodo di 3.3 centesimi di secondo venne scoperto al centro di uno dei più famosi (e spettacolari) Resti di Supernova: la Nebulosa del Granchio, che sappiamo essere stata formata da un'esplosione osservata quasi mille anni fa, nell'anno 1054 e registrata come una nuova stella da osservatori in diverse parti del mondo (particolarmente dettagliata è stata la descrizione dell'evento da parte degli astronomi/astrologi dell'imperatore cinese).

Le osservazioni ad alta energia hanno mostrato che le *Pulsar* sono potenti sorgenti di raggi gamma, e potenti acceleratori di particelle. Le emissioni generate dalle *Pulsar* in alcuni casi si mescolano a quelle generate dalle particelle

accelerate dall'esplosione, ma hanno una diversa origine e sono alimentate da un meccanismo che estrae energia dalla rotazione della stella di neutroni, che rallenta gradualmente (il periodo della *Crab Pulsar* aumenta di 13 microsecondi all'anno).

In molti casi (oltre duecento) parte dell'emissione osservata è pulsata (con il periodo della rotazione), ma si osservano anche raggi gamma da una regione più grande intorno al corpo centrale, chiamata *Pulsar Wind Nebula*, alimentata da un vento di particelle che emerge dalla superficie della stella di neutroni. La più luminosa delle *Pulsar Wind Nebulae* è la *Crab Nebula*, che ha uno spettro che si estende fino al PeV (10^{15} eV o un milione di miliardi di elettron-Volt).

Nuclei Galattici Attivi

La maggior parte delle sorgenti di raggi gamma sono extra-galattiche, e sono distribuite uniformemente sulla sfera celeste. Queste sorgenti extra-galattiche appartengono nella quasi totalità alla classe di oggetti astrofisici conosciuti come *Active Galactic Nuclei* (AGN) o Nuclei Galattici Attivi. Questa è una delle scoperte più significative degli studi recenti sull'universo delle alte energie.

Oggi sappiamo che praticamente tutte le galassie hanno al loro centro una grande massa concentrata in un volume molto piccolo, che quasi certamente è nella forma di un buco nero supermassivo, dove super vuole dire qualcosa che va da milioni a molti miliardi di masse solari. Il premio Nobel per la fisica del 2020 è stato assegnato per metà a Roger Penrose "Per la scoperta che i buchi neri sono una robusta predizione della Teoria Generale della Relatività", e per la seconda metà a Reinhard Genzel ed Andrea Ghez "per la scoperta di un oggetto compatto supermassivo al centro della nostra Galassia". Il buco nero della nostra Galassia ha una massa di circa 4.1 milioni di masse solari, ma è un peso leggero a confronto di altri. Un buco nero supermassivo, recentemente è diventato notissimo in tutto il mondo, è quello al centro della galassia Messier 87 (M87), a circa 55 milioni di anni luce dalla Terra, che ha una massa di circa (6.5 ± 0.7) miliardi di masse solari, e di cui si è riusciti, combinando le osservazioni di telescopi radio distribuiti

in tutto il mondo, ad ottenere un'immagine con una risoluzione senza precedenti [10].

I buchi neri supermassivi sono (paradossalmente se si pensa al loro nome) le sorgenti più luminose dell'universo. La sorgente dell'energia che emettono è la gravità, che agisce accelerando la massa che vi cade dentro. Il meccanismo è fondamentalmente semplice. Immaginiamo un oggetto di massa m che cade verso la superficie solida di un corpo celeste. L'oggetto acquista velocità ed energia nella sua caduta. Quando colpisce la superficie del corpo, la sua velocità si annulla rapidamente, e l'energia cinetica acquistata nella caduta si converte in altre forme, come l'emissione di radiazione. Sappiamo che un buco nero non ha una superficie solida, ma lo si può immaginare come un profondissimo imbuto. Quando una stella si avvicina a sufficienza al buco nero, viene distrutta dalle forze di marea (dovute al fatto che le forze che agiscono su diversi punti della stella sono differenti) e la sua materia, disintegrata in gas, precipita verso il centro, come in un profondo imbuto che si restringe sempre di più. La strozzatura interrompe (momentaneamente) la caduta e l'energia cinetica del gas si può trasformare in altre forme²⁵.

Questo meccanismo di ottenere energia dalla forza di gravità può avere un'efficienza molto alta, che cresce con la profondità dell'imbuto. L'efficienza è massima quando il corpo che accresce massa è un buco nero²⁶. Un tasso di accrescimento di una massa solare all'anno corrisponde ad una luminosità di circa 10^{46} erg/sec (ovvero tremila miliardi di luminosità del Sole).

L'energia dell'emissione da un buco nero è dunque proporzionale al tasso di accrescimen-

to di massa. Per questo alcuni buchi neri (come quello al centro della nostra Galassia) non sono luminosissimi, mentre altri (quelli attivi) lo sono molto di più, e per questo la luminosità di un nucleo galattico attivo può avere grandi variazioni nel tempo che riflettono un tasso di accrescimento non costante.

Le considerazioni appena fatte permettono di ottenere una stima della quantità di energia che si può ottenere da un buco nero supermassivo che accresce massa, ma non dicono nulla della forma con cui questa energia viene rilasciata. Le emissioni dagli AGN prendono forme complesse ed affascinanti che sono ancora un oggetto di studio intenso. Un punto fondamentale è che le osservazioni hanno dimostrato che il rilascio di energia è associato all'accelerazione di raggi cosmici, ed alla emissione di raggi gamma e neutrini anche di altissima energia.

I nuclei galattici attivi si presentano come uno zoo di oggetti molto diversi: Seyferth galaxies, Radio galaxies, Quasars, Blazars, ... che hanno in comune il fatto che il centro della Galassia (una piccola regione spazialmente non risolta) è una potente sorgente di energia. Dalle osservazioni emerge il fatto che la grande variabilità nell'apparire di un AGN è in buona parte dovuta all'orientamento dell'asse di rotazione del buco nero al centro rispetto alla linea di vista dalla Terra, perché le emissioni sono fortemente anisotrope, e sono molto più forti lungo l'asse di rotazione.

Quello che emerge dalle osservazioni è che negli AGN vengono emessi due *getti* di plasma in direzioni opposte, presumibilmente lungo l'asse di rotazione del buco nero. Il getto non è uniforme, ma è formato da una sequenza di *blob*. Questi elementi non hanno una posizione costante nel cielo, ma misure ad alta risoluzione mostrano che si allontanano dal centro dell'AGN. Dalla velocità del movimento è possibile ricostruire la velocità con cui la materia che forma il getto viene emessa. Il risultato è sconcertante, perché si trova che la materia è emessa ad una velocità molto vicina a quella della luce. Il getto di M87 per esempio ha una velocità tra il 98 ed il 99% di quella della luce. Questo è particolarmente sorprendente perché questa non è la velocità di una particella, ma è la velocità di una materia (quasi certamente in forma di plasma) composta

²⁵Il gas possiede il momento angolare delle stelle disintegrate, e quando si avvicina al buco nero forma un disco di accrescimento che si riscalda ad altissima temperatura mentre il suo bordo interno scompare gradualmente oltre l'orizzonte.

²⁶L'efficienza è definita come il rapporto tra la massa che si accresce e l'energia che si può estrarre. L'energia cinetica che si può ottenere dalla massa m che cade da grande distanza fino alla distanza r_{min} dal centro di una massa M è $E = GMm/r_{min}$. Nel caso di un buco nero si può prendere per r_{min} una lunghezza dell'ordine di qualche volta l'orizzonte del buco nero (dato dal raggio di Schwarzschild $R_S = 2GM/c^2$). L'efficienza dunque è data da $m c^2/E \simeq 1/(2r_{min}/R_S)$, e può diventare molto grande. Calcoli più dettagliati che tengono conto della rotazione del buco nero ottengono efficienze che vanno da 0.06 a 0.42.



Figura 19: La “Nebula” Messier 87 (M87). Una galassia attiva alla distanza di 55 milioni di anni luce. Contiene il primo “jet” mai osservato. Venne notato da Heber Curtis nel (1918): “. . . un curioso raggio diritto appare connesso al nucleo da una sottile linea di materia”.

da moltissime particelle che si muovono insieme.

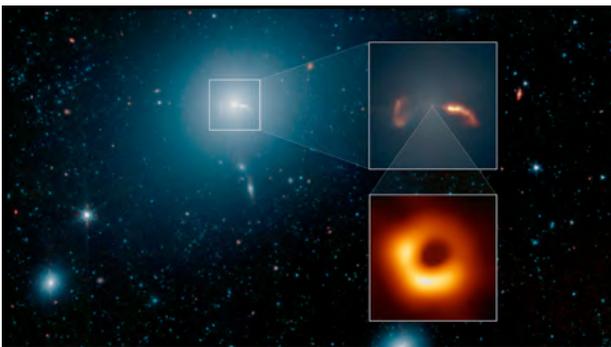


Figura 20: La galassia ellittica M87 contiene un buco nero supermassivo, di cui l’Event Horizon Telescope ha ottenuto un’immagine. Dal centro della galassia emergono due getti simmetrici, e quello emesso verso di noi appare molto più luminoso. L’accrescimento di materia sul buco nero è la sorgente dell’energia che alimenta i getti relativistici.

Le osservazioni dei telescopi per raggi gamma, sia da satellite che da terra, hanno mostrato che i nuclei galattici attivi in cui uno dei due getti viene emesso verso la Terra con un angolo minore di circa una decina di gradi (e che vengono chiamati *Blazars*) sono delle sorgenti molto luminose di fotoni nell’ intervallo 10^8 – 10^{13} eV (GeV–TeV). Questo dimostra che i getti sono siti dove particelle cariche sono accelerate almeno fino a queste energie.

Diversi studi teorici hanno anche suggerito che

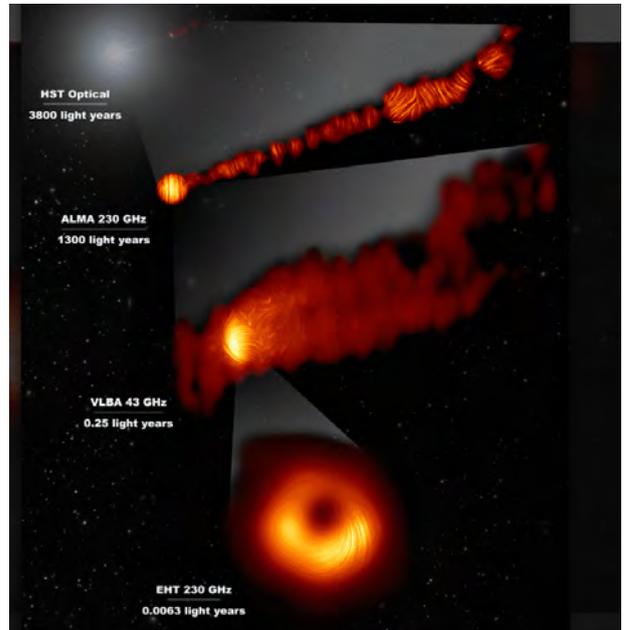


Figura 21: M87 Immagine in onde radio del buco nero al centro di M87, ed il getto emesso in una direzione a circa dieci gradi dalla linea di vista dalla Terra. Il contro getto emesso nella direzione opposta non è visibile.

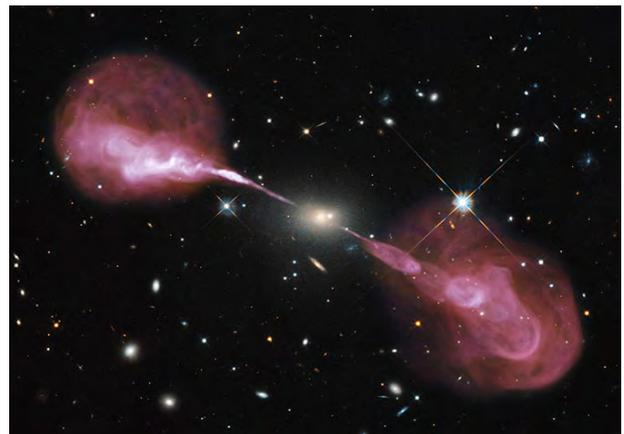


Figura 22: Nucleo Galattico Attivo (AGN) Hercules A, distante 2.1 miliardi di anni luce. L’immagine combina lo spettro di radiazione ottico (Hubble telescope) con quello radio (Jansky VLA telescope). L’immagine radio si estende molte volte oltre le dimensioni della galassia in ottico, ed ha la forma di due getti che terminano in due giganteschi blob radio. Questa è la struttura tipica degli AGN.

gli AGN sono in grado di accelerare particelle fino alle massime energie osservate per i raggi cosmici (che arrivano fino a 10^{20} eV). Abbiamo discusso in precedenza come la maggior parte dei raggi cosmici osservati alla Terra siamo di origine Galattica (cioè generati nella Via Lattea). Raggi cosmici da sorgenti extragalattiche posso-

no però anche penetrare nella Galassia e raggiungere la Terra, e ci si attende che questo contributo diventi dominante ad energie sufficientemente grandi²⁷. I nuclei galattici attivi sono un candidato naturale per essere l'origine della componente extragalattica dei raggi cosmici, e questa ipotesi è oggetto di studi sperimentali da parte dei rivelatori di sciame estesi come l'osservatorio Pierre Auger. [11]. La deviazione subita dalle particelle cariche a causa della presenza dei campi magnetici diminuisce con il crescere della loro energia (o più precisamente rigidità magnetica), e ci si attende che queste deviazioni diventino sufficientemente piccole anche per distanze extragalattiche, in modo da potere "correlare" le direzioni di arrivo dei raggi cosmici di energia più grande con le posizioni nel cielo delle loro sorgenti lontane, permettendo così la loro identificazione. Questa "astronomia" con particelle cariche richiede osservazioni prolungate con telescopi di grande accettazione, perché i flussi sono molto deboli, ed i risultati raggiunti finora non permettono ancora di raggiungere conclusioni certe [12], ma studi futuri saranno di grande interesse.

Evidenza sperimentale che i nuclei galattici attivi sono in grado di accelerare protoni e nuclei fino ad energie altissime è stata ottenuta grazie alle osservazioni dei neutrini di alta energia (10^{14} – 10^{16} eV) da parte di IceCube. Questo telescopio per neutrini ha rivelato un flusso extragalattico di neutrini, e dimostrato che una parte importante di questo flusso è generato nei *Blazars*. In particolare un AGN (il Blazar conosciuto con il nome di TXS 0506+56, distante circa 5.7 miliardi di anni luce dalla Terra, e studiato con fotoni in un ampio intervallo di lunghezze d'onda) è stato chiaramente identificato come sorgente di neutrini [13]. Le osservazioni hanno anche dimostrato che l'emissione non è costante nel tempo, ma è formata da una sequenza di "fiotti" di neutrini di diversa lunghezza temporale, che coincidono con i periodi in cui la sorgente è più luminosa anche in fotoni.

²⁷Questo avviene in parte perché nella Galassia scarseggiano o sono assenti sorgenti capaci di raggiungere le energie più elevate, ed in parte perché i campi magnetici Galattici non sono sufficientemente forti da confinare particelle di altissima rigidità magnetica.

Gamma Ray Bursts

La scoperta dei *Gamma Ray Bursts* (Lampi di Raggi Gamma o GRB) è uno dei capitoli più affascinanti nella storia dell'astrofisica delle alte energie. La scoperta avvenne per caso nel 1967 grazie ai satelliti Vela messi in orbita dagli USA per identificare test di esplosioni nucleari da parte dell'Unione Sovietica. I satelliti registrarono bagliori di raggi gamma, di una durata variabile che andava dalla frazione di secondo a molte decine di secondo. Fu subito chiaro che se gli eventi erano extragalattici, come suggerito dalla loro isotropia, il rilascio di energia doveva essere enorme.

Per tre decenni questi eventi restarono misteriosi fino a quando nel 1997 venne messo in orbita il satellite Beppo-SAX [15], che riuscì per la prima volta a misurare anche la coda di emissione (in inglese *afterglow*) formata da raggi X, più debole, ed in continua diminuzione che segue ai lampi. Questa coda permette di determinare con precisione molto maggiore la posizione dell'evento nel cielo, e quindi di associarlo ad osservazioni con altre lunghezze d'onda. Questo ha permesso di stabilire che i GRB sono effettivamente eventi esplosivi (e che non si ripetono) extragalattici.

L'astrofisica dei GRB è un argomento complesso ed ancora controverso, ma alcune cose sono state solidamente stabilite. Esistono due diversi tipi di Gamma Ray Bursts. Il primo tipo, i GRB lunghi, con una durata mediana di circa trenta secondi, è associato alle *Supernovae*. Una frazione delle *Supernovae* del tipo generato da collasso gravitazionale, produce non solamente una onda d'urto quasi sferica che disintegra la stella, ma anche due getti che espellono plasma a velocità relativistiche lungo l'asse di rotazione della stella progenitrice, ed il lampo gamma è visibile quando la linea di vista dalla Terra è allineata con questo asse.

Il secondo tipo, i GRB brevi, con una durata mediana di circa un secondo, è invece associato alla coalescenza di oggetti compatti, ed in particolare di due stelle di neutroni. Questa idea è stata a lungo un'ipotesi teorica, ma avuto uno spettacolare conferma il 17 agosto del 2017, quando è stata osservata una coincidenza tra un segnale di onde gravitazionali (indicato dalla sigla

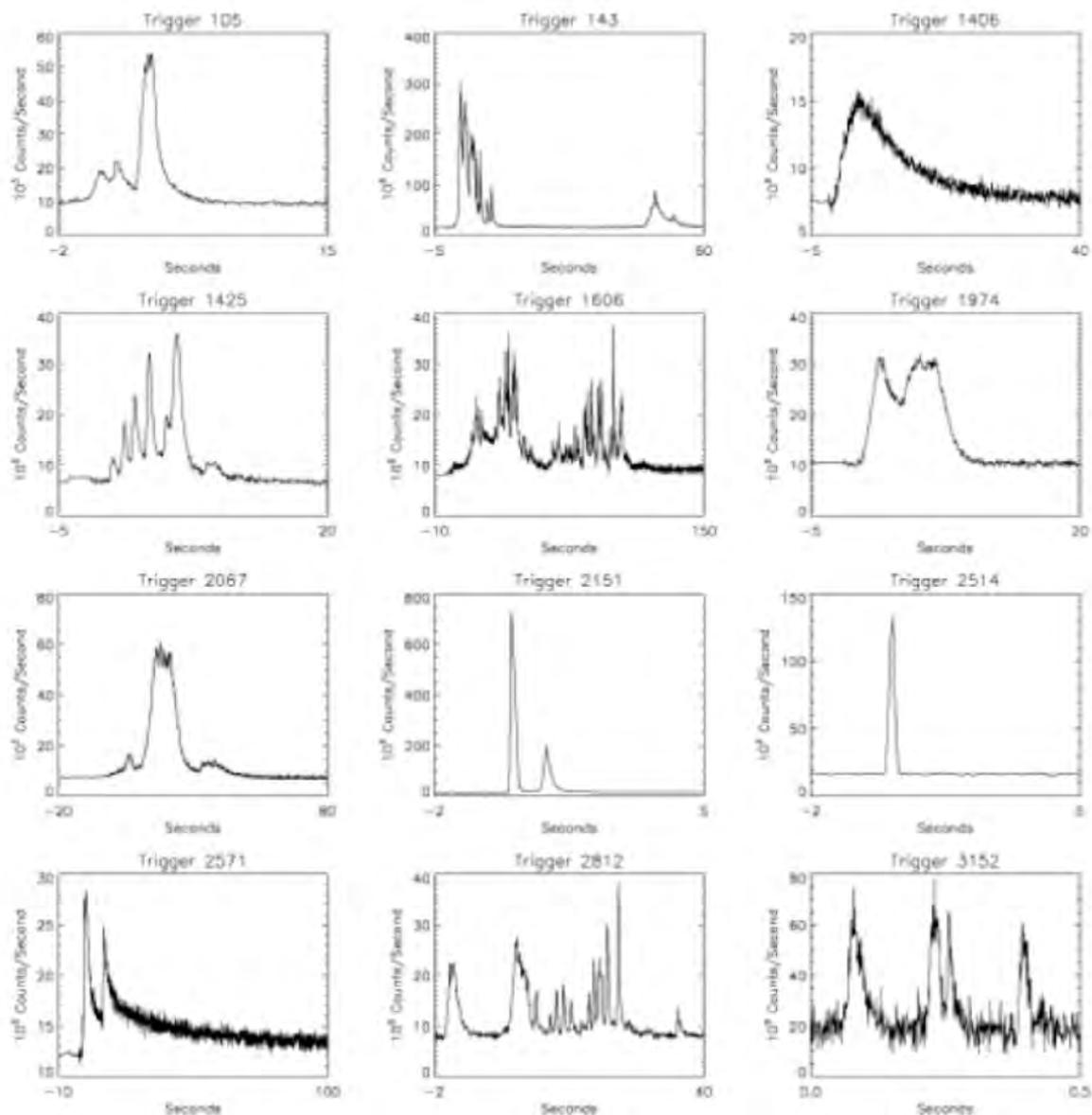


Figura 23: Esempi di curve di luce dei Gamma Ray Bursts [14].

GW170817 che contiene la sua data) ed un GRB breve (GRB 170817A). Moltissimi telescopi hanno puntato nella direzione dell'evento, ed osservato che in una galassia a 140 milioni di anni luce da noi era avvenuta una esplosione di *supernova* qualcosa di nuovo che è stato studiato con grande dettaglio, una sorta di *Supernova*, ma differente da tutte le altre, chiamata ora *kilonova*²⁸.

²⁸La parola *kilonova* è entrata nel vocabolario scientifico solo da una decina di anni (dal 2010) per indicare il transiente ottico di una coalescenza di stelle di neutroni. Il termine ha la sua origine nel fatto che le predizioni teoriche calcolavano un transiente circa mille volte più luminoso di una Nova ordinaria. L'emissione del transiente è generata dai decadimenti dei nuclei radioattivi formati nell'esplosione che segue ad una coalescenza. In queste coalescenze una parte (dell'ordine di un deci-

Il futuro

La scoperta dei raggi cosmici all'inizio del ventesimo secolo ha rappresentato il primo sguardo su quello che oggi è comunemente chiamato l'universo delle alte energie. L'osservazione di particelle relativistiche di altissima energia

mo) delle stelle di neutroni viene "lanciata" nello spazio interstellare. I nuclei (che formano la "crosta" esterna delle stelle) sono investiti da un intenso flusso di neutroni, li assorbono, ed in questo modo, dopo successivi decadimenti beta si formano nuclei più "pesanti". I calcoli teorici suggeriscono che questo è il meccanismo per la formazione degli elementi con grande numero atomico Z (come ad esempio l'oro). Le osservazioni della *kilonova* associata all'evento GW170817 sono in ottimo accordo con le previsioni.



Figura 24: La *kilonova* associata alla coalescenza di due stelle di neutroni osservata il 17 agosto 2017 come segnale di onde gravitazionali (GW170817) e come GRB breve (GRB 170817A).

ovviamente indicava che dovevano esistere sorgenti e meccanismi capaci di generarle, ma la natura di queste sorgenti e di questi meccanismi era completamente sconosciuta, e sarebbe rimasta misteriosa per molte decadi

Gli ultimi anni hanno visto grandissimi progressi nel sollevare i veli che hanno tenuto nascosto l'universo delle alte energie. La chiave di questi successi è stata la comprensione che per studiare i fenomeni di alta energia è necessario combinare le misure dei raggi cosmici con le osservazioni dei raggi gamma (e più in generale dei fotoni in tutte le lunghezze d'onda) e dei neutrini. Lo sviluppo di telescopi più potenti ha finalmente permesso di identificare non solo molte sorgenti di alta energia, ma molte classi di sorgenti: *Supernovae*, Pulsar, Nuclei Galattici Attivi, *Star Burst galaxies*, *Gamma Ray Bursts*, ...

Alcune di queste sorgenti corrispondono anche agli eventi più violenti dell'universo dopo il Big Bang iniziale, perché alla base dei meccanismi che generano le particelle di alta energia, c'è spesso la fortissima accelerazione di grandi masse. Questi sono esattamente gli eventi capaci di increspature lo spazio-tempo con l'emissione di onde gravitazionali, e dunque possono essere osservabili dalle presenti e future antenne.

La chiave per progredire nella comprensione dei fenomeni di alta energia è dunque di combinare le osservazioni di questi quattro messaggeri: raggi cosmici, fotoni, neutrini ed onde gravitazionali. I primi risultati di questa nuova astronomia multimessaggera sono già di grandissimo interesse, ma le prospettive per il futuro sono ancora più promettenti. La qualità delle osservazioni nel prossimo futuro migliorerà in modo significativo, perché esistono progetti di nuovi telescopi e nuovi strumenti per tutti e quattro i messaggeri. Le nuove osservazioni promettono di comprendere meglio le proprietà delle sorgenti note, e di scoprirne di nuove.

La storia dell'astronomia fino ad oggi può essere vista come una graduale espansione dell'intervallo di lunghezza d'onda (o di energia) dei fotoni utilizzati per le osservazioni a partire dalla ristretta banda dei fotoni ottici (visibili per nostri occhi). Ogni passo in questa direzione, l'apertura di ogni nuova finestra (luce infrarossa, luce ultravioletta, onde radio, raggi X, raggi gamma, ...) non solo ha permesso una migliore comprensione di ciò che era già conosciuto, ma ha portato a sorprese, alla scoperta di nuovi oggetti astrofisici e nuovi fenomeni. Siamo oggi testimoni di un'espansione molto più profonda, con l'utilizzo di *nuovi messaggeri* che permettono di vedere l'universo intorno a noi in modo completamente nuovo, con occhi diversi.

La comprensione del cosmo che ci circonda è un progetto che ha impegnato l'umanità da tempi remoti, e dunque conoscere l'universo delle alte energie e gli strani oggetti/eventi da una parte è qualcosa che di per se è profondamente affascinante. D'altra parte gli oggetti/eventi dell'universo delle alte energie sono anche i luoghi dove si incontrano le condizioni più estreme: le densità di materia più grandi, i campi magnetici più elevati, le interazioni alle energie più elevate, ..., e quindi possono anche essere considerati come dei laboratori per testare le leggi fondamentali della fisica in condizioni che non possono essere create in un laboratorio sulla Terra. Astronomia/astrofisica e fisica fondamentale sono sempre state profondamente intrecciate nella storia della scienza (a partire da Newton e la scoperta delle equazioni della dinamica, la comprensione della sorgente di energia del Sole, delle stelle e la fusione nucleare, il collasso gravitazio-

nale e le leggi della statistica quantistica, ...) e continuano ad essere legate da mille fili.



- [1] Bruno Rossi: *I raggi cosmici*, Einaudi, Torino (1971).
- [2] Pamela Collaboration: *The PAMELA Mission: Heralding a new era in precision cosmic ray physics*, Phys. Rep., 544 (2014) 323.
- [3] AMS Collaboration: *The Alpha Magnetic Spectrometer (AMS) on the international space station: Part II — Results from the first seven years*, Phys. Rep., 894 (2021) 1.
- [4] EGRET Collaboration: *The Third EGRET catalog of high-energy gamma-ray sources*, Astrophys. J. Suppl., 123 (1999) 79.
BATSE Collaboration: *Identification of two classes of gamma-ray bursts*, Astrophys. J. Lett., 413 (1993) L101.
- [5] Fermi–LAT Collaboration: *Fermi-LAT Observations of the Diffuse Gamma-Ray Emission: Implications for Cosmic Rays and the Interstellar Medium*, Astrophys. J., 750 (2012) 3.
Fermi Large Area Telescope Fourth Source Catalog, Astrophys. J. Suppl., 247 (2020) 33.
- [6] Cherenkov Telescope Array Collaboration: *Science with the Cherenkov Telescope Array*, [arXiv:1709.07997 [astro-ph.IM]].
- [7] LHAASO Collaboration: *The Large High Altitude Air Shower Observatory (LHAASO) Science Book*, Chin. Phys. C, 46 (2022) 035001.
- [8] F. Halzen and A. Kheirandish: *IceCube and High-Energy Cosmic Neutrinos*, [arXiv:2202.00694 [astro-ph.HE]].
IceCube Collaboration: *Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector*, Science, 342 (2013) 1242856.
- [9] KM3Net Collaboration: *Letter of intent for KM3NeT 2.0*, J. Phys. G, 43 (2016) 084001.
- [10] Event Horizon Telescope Collaboration, K. Akiyama et al. [Event Horizon Telescope]: *First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole*, Astrophys. J. Lett., 875 (2019) L1.
- [11] Pierre Auger Observatory Collaboration: *Features of the Energy Spectrum of Cosmic Rays above 2.5×10^{18} eV Using the Pierre Auger Observatory*, Phys. Rev. Lett., 125 (2020) 121106.
- [12] Pierre Auger Observatory Collaboration: *An Indication of anisotropy in arrival directions of ultra-high-energy cosmic rays through comparison to the flux pattern of extragalactic gamma-ray sources*, Astrophys. J. Lett., 853 (2018) L29.
- [13] IceCube Collaboration: *Multimessenger observations of a flaring blazar coincident with high-energy neutrino IceCube-170922A*, Science, 361 (2018) eaat1378.
Neutrino emission from the direction of the blazar TXS 0506+056 prior to the IceCube-170922A alert, Science, 361 (2018) 147
- [14] BATSE Collaboration: *BATSE observations of gamma-ray burst spectra. 1. Spectral diversity*, Astrophys. J., 413 (1993) 281.
- [15] E. Costa, et al.: *Discovery of an X-ray afterglow associated with the gamma-ray burst of 28 February 1997*, Nature, 387 (1997) 783.



Paolo Lipari: è Dirigente di Ricerca presso la Sezione di Roma Sapienza dell'Istituto Nazionale di Fisica Nucleare. Da molti anni il suo principale interesse scientifico è lo studio dell'universo delle alte energie.

Raggi cosmici, ieri, oggi, domani

Alessandro De Angelis

*Dipartimento di Fisica e Astronomia "Galilei" dell'Università di Padova, INFN e INAF
Instituto Superior Técnico dell'Università di Lisbona*

Scoperti all'inizio del XX secolo, i raggi cosmici sono particelle di energie molto elevate (fino a un miliardo di TeV) che raggiungono la Terra in provenienza da acceleratori extraterrestri il cui meccanismo di funzionamento non è ancora pienamente compreso. Sono un potente strumento per indagare la fisica fondamentale, e in modo complementare possono spiegarci interessantissimi fenomeni astrofisici, in particolare quelli che avvengono in prossimità di oggetti compatti. Possono far progredire la nostra conoscenza dell'Universo su larga scala, dalla cosmologia alla ricerca sulle origini della vita. La rinnovata attenzione degli scienziati per i raggi cosmici ha portato negli ultimi anni al fiorire di attività sperimentali e teoriche nei campi chiamati fisica astroparticellare e astronomia multimessaggero.

Che cosa sono i raggi cosmici?

Ogni secondo, particelle ad alta energia (sopra ~ 1 GeV) di origine extraterrestre raggiungono la Terra. Portano messaggi da regioni dove avvengono fenomeni altamente energetici che non possiamo esplorare direttamente. Sono chiamate raggi cosmici, ed è naturale cercare di utilizzarle per studiare la fisica di questi fenomeni [1, 2]. I raggi cosmici sono prevalentemente protoni, con

circa il 10% di nuclei di elio, e una piccola, anche se molto importante, frazione di elettroni, positroni, raggi gamma e neutrini (chiamo in questo articolo raggi cosmici tutte le particelle di origine extraterrestre – alcuni invece definiscono come raggi cosmici solo i nuclei, o solo protoni e ioni, cioè separano raggi gamma e neutrini dai raggi cosmici).

Lo spettro di energia dei raggi cosmici (Fig. 1) è approssimativamente descritto da una legge di potenza E^{-p} , con p intorno a 3 in media. Prima di andare più in dettaglio, una breve introduzione sulle unità di misura usate nella fisica delle particelle.

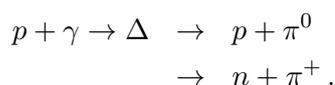
Un elettron-volt (eV) è la quantità di energia acquistata da una particella con carica elettrica unitaria, per esempio un elettrone o un protone, quando viene accelerata tra due punti fra cui ci sia una differenza di potenziale di 1 volt. L'elettron-volt è circa l'energia della luce visibile, ed è una scala naturale per i fenomeni chimico-fisici in cui la luce viene generata o assorbita (transizioni di elettroni nei livelli esterni degli atomi o nei cristalli, ad esempio nella fotosintesi delle piante), ma è un'unità piuttosto piccola rispetto ai fenomeni macroscopici: vale infatti circa 1.5×10^{-19} joule, dove il joule è l'unità di misura del sistema internazionale.

Le energie caratteristiche degli elettroni più interni di un atomo sono invece dell'ordine dei keV (1 keV = 1000 eV), e sono caratteristiche dei cosiddetti raggi X. Le energie che legano i protoni

nei nuclei sono ancora maggiori, dell'ordine dei MeV ($1 \text{ MeV} = 10^6 \text{ eV} = 1000000 \text{ eV}$); qui cominciano i cosiddetti raggi gamma. Arrivati a 1 GeV, ossia un miliardo (10^9) di eV, abbiamo un'energia sufficiente a creare un protone compatibilmente con la famosa relazione di Einstein $E = mc^2$. Una zanzara ha un'energia cinetica di circa 1000 GeV, o 1 TeV, e l'energia dei fasci dell'acceleratore LHC al CERN di Ginevra, il più potente acceleratore artificiale, è di 7 TeV. 1000 TeV fanno un PeV, e 100 PeV fanno un EeV, cioè 10^{18} eV (e qui ci avviciniamo a un joule).

Ora andiamo nel dettaglio della dipendenza del flusso dei raggi cosmici dall'energia. Dopo una regione intorno al GeV dominata dai raggi cosmici prodotti o moderati dal Sole, lo spettro diventa più ripido ($p \sim 2.7$) per valori di energia inferiori a 1000 TeV, cioè 1 PeV (circa 150 volte l'energia massima dei fasci del collisore LHC al CERN); sappiamo che questa regione di energia è dominata dai raggi cosmici prodotti da sorgenti astrofisiche nella nostra Galassia. Per energie più elevate si verifica un ulteriore irrigidimento ($p \sim 3.3$); la regione in cui avviene questo cambiamento di pendenza è chiamata **ginocchio**. I raggi cosmici sopra il ginocchio sono prodotti principalmente da sorgenti extragalattiche, per lo più buchi neri supermassicci che crescono al centro di altre galassie. Per energie ancora più alte (oltre $\sim 10^{18} \text{ eV}$, cioè 1 EeV) lo spettro dei raggi cosmici diventa meno ripido, determinando un altro cambiamento di pendenza, chiamato **caviglia**. Infine, alle energie più alte nella figura, si vede una drastica riduzione del flusso.

Un limite superiore all'energia dei raggi cosmici da sorgenti lontane è stato calcolato nel 1966 da Greisen, Kuzmin e Zatsepin, ed è chiamato oggi cutoff GZK [3, 4]. I protoni con energie superiori alla soglia di circa 10^{20} eV subiscono un'interazione con i fotoni del fondo cosmico di microonde (CMB) e degradano la loro energia producendo pioni, in particolare attraverso le reazioni



Questa degradazione continua fino a quando l'energia dei nucleoni scende sotto la soglia di produzione della Δ (1232 MeV). A causa del percorso medio associato all'interazione, i raggi cosmici

extragalattici provenienti da distanze maggiori di $\sim 50 \text{ Mpc}$ dalla Terra e con energie maggiori di circa 10^{20} eV sono fortemente soppressi, e non ci sono sorgenti conosciute a distanze minori che possano produrli. Un effetto simile (fotodisintegrazione nucleare) limita il cammino libero medio dei nuclei più pesanti.

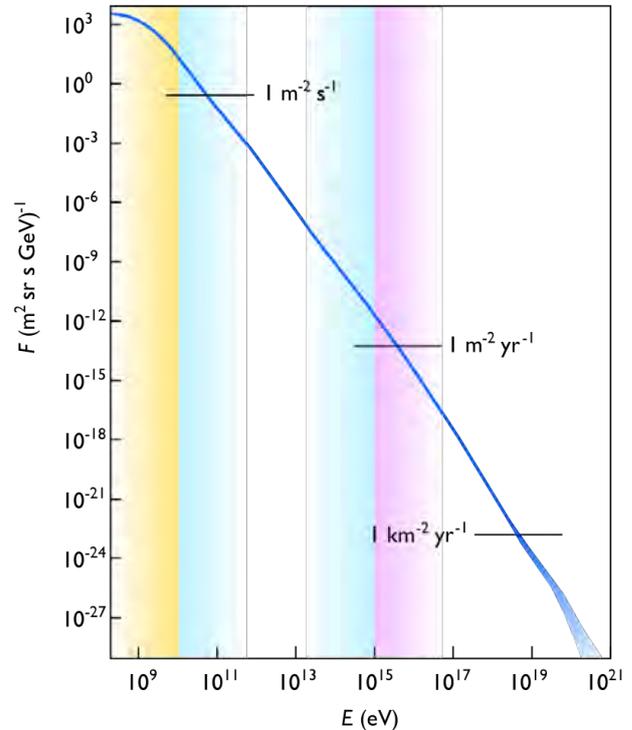


Figura 1: Spettro energetico (numero di particelle incidenti per unità di energia, per secondo, per unità di area, e per unità di angolo solido) dei raggi cosmici primari. La banda verticale a sinistra indica la regione energetica in cui si suppone che l'emissione dal Sole sia dominante; la banda centrale la regione in cui la maggior parte dell'emissione è presumibilmente di origine galattica; la banda a destra la regione di origine extragalattica. Wikimedia Commons.

La composizione dei raggi cosmici, come prevedibile, varia molto con l'energia. Il numero di neutrini ad alta energia è stimato essere paragonabile a quello dei fotoni ad alta energia, ma è molto più alto ad energie più piccole, fino alle energie del MeV, a causa dei processi nucleari che avvengono nel Sole, che comportano una grande produzione di neutrini.

I messaggeri cosmici neutri e stabili (raggi gamma, neutrini) sono molto preziosi poiché non sono deflessi dai campi magnetici extragalattici (dell'ordine di $10^{-9} - 10^{-15} \text{ G}$) o galattici (dell'ordine di 10^{-6} G) e puntano direttamente alle

loro sorgenti, consentendone la localizzazione. Infatti la curvatura indotta dai campi magnetici

$$\frac{R_L}{1\text{kpc}} \simeq \frac{E/1\text{EeV}}{B/1\mu\text{G}} \quad (1)$$

è tale che è impossibile localizzare sorgenti vicino al Centro Galattico, a circa 9 kpc (28mila anni-luce) da noi, con protoni meno energetici di ~ 10 EeV, un'energia molto vicina al cutoff GZK. Abbiamo rivelato invece un grande flusso di raggi gamma provenienti da diversi siti di produzione cosmica. Infine, l'esistenza di neutrini astrofisici è stata provata solo nel 2013 [5].

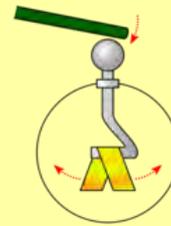
I raggi cosmici che colpiscono l'atmosfera ad energie altissime (ad esempio fotoni sopra i 100 GeV e nuclei sopra circa 1 TeV) possono produrre particelle secondarie che raggiungono la superficie terrestre attraverso il fenomeno degli sciami moltiplicativi.

Circa una volta al minuto, una singola particella subatomica entra nell'atmosfera terrestre con un'energia maggiore di 10 J ($1\text{ J} \sim 6\text{ EeV}$): da qualche parte nell'Universo esistono acceleratori che impartiscono ai singoli protoni energie milioni di volte più grandi delle energie raggiunte dal Collisore LHC al CERN. Come può accadere? Pensiamo che il motore ultimo dell'accelerazione dei raggi cosmici sia la gravità. In collassi gravitazionali giganteschi, come quelli che avvengono nelle supernove e nell'accrescimento di buchi neri supermassicci (equivalenti a milioni o miliardi di masse solari) a spese della materia circostante nei centri delle galassie (i cosiddetti nuclei galattici attivi, AGN), parte dell'energia potenziale gravitazionale si trasforma in energia cinetica delle particelle. La ragione per cui la massima energia raggiunta dagli acceleratori costruiti dall'uomo con le tecnologie di accelerazione attualmente conosciute non può competere con gli acceleratori cosmici è semplice. Il modo più efficiente per accelerare le particelle cariche richiede il loro confinamento entro un raggio R da un campo magnetico B , e l'energia finale è proporzionale al prodotto RZB , dove Z è la carica della particella in unità di e :

$$\begin{aligned} \frac{E}{1\text{PeV}} &\simeq \frac{ZB}{1\mu\text{G}} \times \frac{R}{1\text{pc}} \simeq 5 \frac{ZB}{1\text{G}} \times \frac{R}{1\text{AU}} \\ &\simeq 0.0003 \frac{ZB}{1\text{T}} \times \frac{R}{1\text{km}}. \end{aligned} \quad (2)$$

Elettroscopio

L'elettroscopio è costituito da un pomello conduttore collegato a due sottili lamine metalliche chiamate foglioline (solitamente d'oro) che sono racchiuse in un recipiente di vetro per isolarle.



Avvicinando al pomello metallico un corpo caricato elettricamente, ad esempio una bacchetta di plastica strofinata con un panno di lana, si vedranno le due lamine separarsi. La vicinanza del corpo elettricamente carico trasmette le cariche dello stesso segno alle due foglioline che si respingono.

Sulla Terra, è difficile immaginare raggi di confinamento più grandi di qualche km, e campi magnetici più forti di qualche T. Questa combinazione può fornire energie di qualche TeV, come quelle dell'LHC. Esistono acceleratori astrofisici con raggi molto più grandi, come i resti di supernova (anni luce) e i nuclei galattici attivi (decine di migliaia di anni luce). Naturalmente gli acceleratori creati dall'uomo hanno importanti vantaggi, come la luminosità e la conoscenza delle condizioni iniziali.

Breve storia della ricerca sui raggi cosmici

La storia di come abbiamo imparato tutto questo è affascinante. Nel 1785, Coulomb scoprì che gli elettroscopi possono scaricarsi spontaneamente, e non semplicemente a causa di un isolamento difettoso. Nel 1879 il britannico Crookes osservò che la velocità di scarica diminuiva quando la pressione dell'aria all'interno degli elettroscopi veniva ridotta: la scarica era quindi probabilmente dovuta alla ionizzazione dell'aria. Ma qual è la causa della ionizzazione atmosferica? La radioattività fu osservata nel 1896 da Becquerel, e Marie e Pierre Curie scoprirono che alcuni elementi subiscono trasmutazioni spontanee che

generano radioattività. Un elettroscopio carico si scarica velocemente in presenza di materiali radioattivi. Si concluse che la scarica era dovuta all'emissione di particelle cariche, che inducono la formazione di ioni che neutralizzano le cariche delle foglioline. La velocità di scarica degli elettroscopi fu quindi usata per misurare il livello di radioattività.

Durante la prima decade del ventesimo secolo, diversi ricercatori in Europa e nel Nuovo Mondo progredirono nello studio della ionizzazione. Per quanto riguarda l'origine delle radiazioni che la inducono, l'ipotesi più semplice era che fossero legate a materiale radioattivo nel suolo. L'origine terrestre di parte della radioattività era certa, ma non si era sicuri che non ci fosse qualcos'altro.

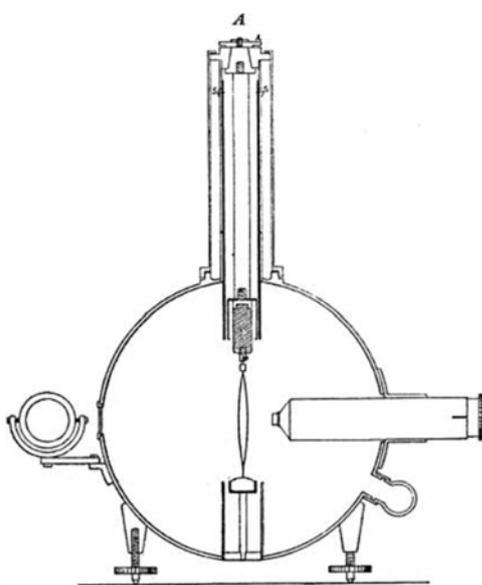


Figura 2: Schema dell'elettroscopio di Wulf (disegnato da Wulf stesso [6]). Il cilindro principale era fatto di zinco, e il volume totale era di circa un litro. La distanza tra i due fili di vetro al silicio (al centro) era misurata usando il microscopio a destra. I fili erano illuminati usando lo specchio a sinistra. Secondo Wulf, la sensibilità dello strumento era di 1 V, come misurato dalla diminuzione della distanza tra i fili.)

Padre Theodor Wulf, uno scienziato tedesco e sacerdote gesuita che insegnava in Olanda dopo alcune esperienze al Collegio Romano, misurò la variazione della ionizzazione con l'altezza per verificare se l'origine della radioattività fosse, in parte, extraterrestre. Nel 1909, utilizzando un elettroscopio trasportabile (Fig. 2) in cui le due foglie erano state sostituite da fili di vetro al silicio rivestiti di metallo, misurò il tasso di ioniz-

zazione in cima alla Torre Eiffel, a circa 300 m di altezza. Nell'ipotesi che la maggior parte della radiazione fosse di origine terrestre, si aspettava che il tasso di ionizzazione fosse significativamente inferiore rispetto al valore al suolo. La diminuzione misurata era troppo piccola per confermare l'ipotesi. Tuttavia, il risultato era afflitto da due problemi: la torre Eiffel era una fonte non trascurabile di radioattività, e l'interpretazione dei risultati era stata fatta nell'ipotesi che i raggi cosmici fossero raggi gamma, e quindi il valore assunto per la lunghezza di attenuazione era basso rispetto a quello che oggi sappiamo essere tipico della radiazione cosmica [6].



Figura 3: Pacini al lavoro nel 1910. Per gentile concessione della famiglia Pacini (Wikimedia Commons).

La conclusione che la ionizzazione atmosferica fosse principalmente dovuta alla radioattività della crosta terrestre fu messa in discussione dal geofisico Domenico Pacini (Fig. 3). Pacini sviluppò una tecnica per misure subacquee e condusse esperimenti nel mare del Golfo di Genova (usando il cacciatorpediniere Fulmine della Marina Militare Italiana presso l'Accademia Navale di Livorno) e nel lago di Bracciano [7]. Trovò una diminuzione significativa (circa il 20%, con una significatività pari a 4.3 deviazioni standard) della velocità di scarica spontanea di elettroscopi posti a tre metri sott'acqua. Scrisse:

"Le osservazioni effettuate nel mare nel 1910 mi conducevano a concludere che una parte non trascurabile della radiazione penetrante che si riscontra nell'aria avesse origine indipendente dall'azione diretta delle sostanze attive con-

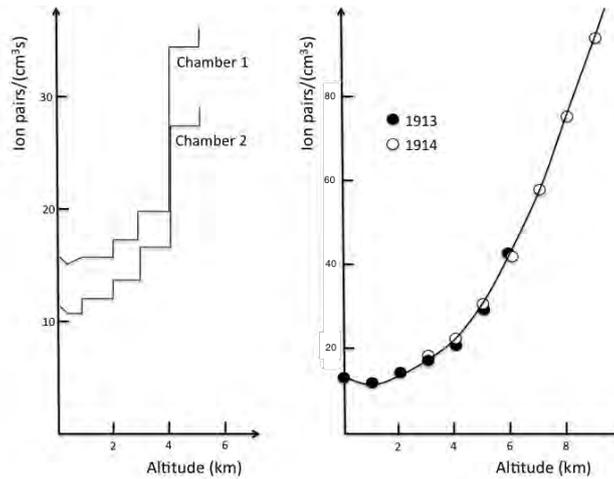


Figura 4: A sinistra: Hess durante il volo in mongolfiera nell'agosto 1912 (Wikimedia, pubblico dominio). A destra: La ionizzazione misurata nel volo ad alta quota di Hess del 1912 e nel volo di Kollhörster nel 1914 (media dei suoi due rivelatori) [1].

tenute negli strati superiori della crosta terrestre [...] [Per provare questa conclusione] l'apparecchio [...] venne racchiuso in una scatola di rame per poterlo immergere in seno alle acque. [...] Si fecero delle osservazioni coll'apparecchio alla superficie, e coll'apparecchio immerso nelle acque, a 3 m di profondità."

Pacini concluse che la diminuzione era dovuta a radiazione non proveniente dalla crosta terrestre.

Dopo le osservazioni di Wulf e i risultati di Pacini, la necessità di esperimenti con palloni aerostatici (ampiamente utilizzati per studi sull'elettricità atmosferica dal 1885, e prima ancora da Gay-Lussac per testare la sua legge dei gas) divenne chiara.

Il primo pallone ad alta quota con lo scopo di studiare la radiazione penetrante fu fatto volare in Svizzera nel 1909 da Albert Gockel, professore all'Università di Friburgo, che salì fino a 4500 m s.l.m., ma fece misure solo fino a 3000 m (lo strumento poi si ruppe). Trovò che [8] il tasso di ionizzazione non diminuiva con l'altitudine come ci si aspettava nell'ipotesi di origine terrestre, e concluse che

"una parte non trascurabile della radiazione penetrante è indipendente dall'azione diretta delle sostanze radioattive negli strati più alti della Terra".

Le misure sulla scarica spontanea avevano dimostrato che la radiazione di fondo era presente anche in ambienti isolati e poteva penetrare schermi metallici. Era difficile spiegarla in termini di radiazione α (nuclei di elio) o β (elettroni); si suppose quindi che fosse una radiazione γ , cioè fatta di fotoni, che era la più penetrante tra i tre tipi di radiazione conosciuti all'epoca (il potere penetrante delle particelle cariche relativistiche non era ancora noto).

L'esperimento decisivo fu eseguito nel 1912 dall'austriaco Victor Hess (Fig. 4, a sinistra), dopo un'accurata preparazione e calibrazione degli strumenti. Hess volò sette volte con tre elettroscopi, schermati da metalli di diversa spessore per distinguere le radiazioni α , β e γ . Nell'ultimo volo, il 7 agosto, raggiunse un'altezza di 5200 m [9]. La ionizzazione, dopo essere passata attraverso un minimo, aumentò considerevolmente con l'altezza (Fig. 4, a destra).

- "(i) Immediatamente sopra il suolo la radiazione totale diminuisce un po'.
- (ii) Ad altitudini da 1000 a 2000 m si verifica di nuovo una crescita notevole della radiazione penetrante.
- (iii) L'aumento raggiunge, ad altitudini da 3000 a 4000 m, il 50% della radiazione totale osservata al suolo.
- (iv) A 4000 a 5200 m la radiazione è più forte [più del 100%] che al suolo".

Hess concluse che l'aumento della ionizzazione con l'altezza è dovuto alla radiazione proveniente dall'alto, e che questa radiazione è di origine extra-terrestre. Escluse il Sole come fonte diretta di questa ipotetica radiazione penetrante a causa dell'assenza di variazioni giorno-notte.

Questi risultati furono poi confermati dal suo allievo Kolhörster in un certo numero di voli fino a 9200 m: fu misurata una ionizzazione fino a dieci volte più grande di quella a livello del mare (Fig. 4 a destra). Kolhörster misurò il coefficiente di assorbimento della radiazione, e trovò che era otto volte più piccolo del coefficiente di assorbimento dei raggi gamma; tuttavia non indagò ulteriormente la discrepanza. L'ultimo volo di Kolhörster fu eseguito il 28 giugno del 1914, giorno dell'assassinio dell'arciduca Francesco Ferdinando, erede dell'impero austro-ungarico, a Sarajevo – un giorno che ha cambiato la storia.

Durante la guerra 1914-1918 e gli anni che seguirono, furono eseguite pochissime indagini sulla radiazione penetrante; il centro della ricerca si spostò fino al 1930 negli Stati Uniti, e molte informazioni sui risultati in Europa andarono perse. L'idea dei raggi cosmici, nonostante le impressionanti prove sperimentali, non fu immediatamente accettata.

Nel 1926 Millikan eseguì misure di assorbimento della radiazione a diverse profondità in laghi ad alta quota. Egli concluse, erroneamente, che la radiazione era costituita da raggi γ ad alta energia "[che viaggiano] attraverso lo spazio allo stesso modo in tutte le direzioni" e li chiamò **raggi cosmici**. Millikan aveva azzardato l'ipotesi che questi raggi fossero prodotti quando protoni ed elettroni formavano nuclei di elio nello spazio interstellare, e che fossero "grida di atomi neonati".

Un esperimento chiave sulla natura dei raggi cosmici fu la misurazione della variazione di intensità con la latitudine geomagnetica. Durante due viaggi in nave tra Java e Genova nel 1927 e 1928, il fisico olandese Clay trovò che la ionizzazione aumentava con la latitudine; questo dimostrò che i raggi cosmici venivano deviati dal campo geomagnetico: quindi erano soprattutto particelle cariche.

Nel 1928 fu introdotto il tubo contatore Geiger-Müller - secondo Giuseppe Occhialini, uno dei pionieri dell'esplorazione della fisica fonda-

mentale con i raggi cosmici, era come il revolver Colt nel Far West: uno strumento economico utilizzabile da tutti nel proprio cammino attraverso una dura frontiera. Alla fine degli anni '20, i tedeschi Bothe (premio Nobel per la fisica nel 1954 "per il metodo della coincidenza e le sue scoperte fatte con esso") e Kolhörster introdussero la tecnica della coincidenza per studiare i raggi cosmici con il contatore Geiger, permettendo così la misura della direzione di arrivo. Un circuito di coincidenza attiva l'acquisizione di dati solo quando i segnali da rivelatori predefiniti vengono ricevuti entro una determinata finestra temporale.

Nel 1933, tre esperimenti indipendenti di Alvarez e Compton [10], Johnson [11], Rossi [12] scoprirono che vicino all'equatore c'erano più raggi cosmici provenienti da ovest che da est. Questo effetto, dovuto all'interazione con il campo geomagnetico, ha dimostrato che i raggi cosmici sono per lo più carichi positivamente, quindi molto probabilmente protoni.

Con lo sviluppo della fisica dei raggi cosmici, gli scienziati sapevano che sorgenti astrofisiche fornivano particelle di alta energia che entravano nell'atmosfera. L'ovvio passo successivo era quello di indagare la natura di tali particelle, e di usarle per sondare la materia in dettaglio, proprio come nell'esperimento di Rutherford. La fisica delle particelle iniziò così con i raggi cosmici, e molte delle scoperte fondamentali furono fatte grazie ai raggi cosmici. In parallelo, la comprensione teorica della fisica progrediva rapidamente: alla fine degli anni 1920, gli scienziati cercarono di mettere insieme la relatività e la meccanica quantistica, e le scoperte successive a questi tentativi cambiarono completamente la nostra visione della natura. La camera a nebbia (una scatola che contiene un vapore soprassaturo in modo tale che piccole goccioline si formino intorno alla traiettoria di una particella carica che vi passa attraverso) e le emulsioni nucleari (lastre fotografiche che registrano le tracce del passaggio di particelle cariche) furono i primi strumenti a rendere possibili queste scoperte.

Nel 1933 Anderson, uno studente di Millikan, scoprì nei raggi cosmici l'antimateria sotto forma di una particella positiva di massa uguale alla massa dell'elettrone, poi chiamata positrone [13]. Una previsione dell'equazione di Dirac fu confermata; questo fu un risultato chiave per

la fisica dei raggi cosmici. Anderson condivise con Hess il premio Nobel per la fisica nel 1936; furono nominati da Compton, con la seguente motivazione:

"è arrivato il momento [...] in cui possiamo dire che i cosiddetti raggi cosmici hanno sicuramente la loro origine a distanze così remote dalla Terra che possono essere propriamente chiamati cosmici e che l'uso di questi raggi ha ormai portato a risultati di tale importanza che possono essere considerati una scoperta di prima grandezza. [...] Hess è stato il primo a stabilire l'aumento della ionizzazione osservata negli elettroscopi con l'aumentare dell'altitudine; e certamente il primo ad attribuire con sicurezza questo aumento di ionizzazione alla radiazione proveniente dall'esterno della Terra".

Perché un riconoscimento così tardivo alla scoperta dei raggi cosmici? Compton scrive:

"Prima che fosse opportuno assegnare il premio Nobel per la scoperta di questi raggi, era necessario attendere prove più certe riguardo alle loro caratteristiche uniche e alla loro importanza in vari campi della fisica" [14].

Dopo la fondamentale scoperta dell'antimateria, nuovi risultati sperimentali nella fisica delle particelle elementari con i raggi cosmici furono guidati e accompagnati dal miglioramento degli strumenti di rivelazione, in particolare dal miglioramento delle camere a nebbia e dei circuiti elettronici che utilizzavano tubi Geiger. Nel 1934 Bruno Rossi, che aveva migliorato la tecnica delle coincidenze, riportò l'osservazione di segnali quasi simultanei da coppie di contatori Geiger a grandi distanze durante una prova della sua attrezzatura. Scrisse:

"[Sembra che] ogni tanto l'apparecchio di registrazione sia colpito da sciame di particelle molto estesi, che provocano coincidenze tra i contatori, anche posti a grande distanza l'uno dall'altro" [15].

Nel 1937 Pierre Auger, probabilmente all'oscuro dell'articolo di Rossi, fece un'osservazione simile e studiò il fenomeno in dettaglio. Concluse

che gli sciame estesi hanno origine quando i raggi cosmici primari ad alta energia interagiscono con i nuclei nell'alta atmosfera, portando ad una serie di interazioni che alla fine producono uno sciame atmosferico di particelle che raggiungono il suolo. Questa era la spiegazione della scarica spontanea di elettroscopi dovuta ai raggi cosmici [16].

La fisica delle particelle elementari è stata per diversi anni essenzialmente un prodotto della ricerca sui raggi cosmici. Molte nuove particelle furono scoperte negli anni successivi nei raggi cosmici: muoni, mesoni, particelle strane. Con gli anni 1950, tuttavia, gli acceleratori cominciarono a prendere il dominio, e nella Conferenza Internazionale sui Raggi Cosmici a Bagnères de Bigorre del 1953, Leprince-Ringuet disse nelle sue osservazioni conclusive:

"In futuro dovremo usare gli acceleratori di particelle [...] anche se [...] abbiamo l'esclusiva di alcuni fenomeni [...] per i quali le energie sono molto più grandi" [17].

Anche dopo il 1950, coerentemente con la previsione di Leprince, sebbene il centro dell'attenzione della fisica delle particelle si sia spostato sulla fisica degli acceleratori (almeno fino al 2012), molti risultati di prima grandezza nella fisica fondamentale sono dovuti ai raggi cosmici, in particolare:

- La scoperta del quark charm. La produzione associata di particelle charmate è stata osservata per la prima volta in Giappone nel 1971 negli studi sulle emulsioni di raggi cosmici [18], tre anni prima della scoperta del mesone J/ψ negli acceleratori (1974).
- La scoperta che i neutrini hanno una massa diversa da zero. Questa è finora l'unica prova che dimostra che deve esistere una fisica oltre il Modello Standard: una particella (il neutrino destrorso) è sterile nel modello, essendo la sua interazione puramente gravitazionale (ossia la sua esistenza fisica è legata a un tipo di interazione non compresa nel modello).

Come e dove vengono prodotti i raggi cosmici?

Nel 1933, lo svizzero Zwicky e il tedesco Baade avanzarono una congettura rivoluzionaria: le stelle massicce terminano la loro vita in collassi seguiti da esplosioni; tali esplosioni producono raggi cosmici e si lasciano dietro una stella fatta di neutroni densamente impacchettati. Ma come può un resto di supernova (o qualsiasi altro residuo di un collasso gravitazionale) accelerare le particelle? Fu necessario aspettare fino al 1949 quando Enrico Fermi concepì un modello che spiegava questa congettura.

Fermi propose un meccanismo in cui le particelle possono essere accelerate in collisioni stocastiche rimbalzando avanti e indietro tra le pareti dei plasmi associati alle onde d'urto che partono dal resto di un collasso gravitazionale - un collasso stellare, ma anche, come sappiamo oggi, l'intorno di un buco nero che si accresce nel centro di una galassia. Il modello di Fermi spiega approssimativamente la dipendenza dalla legge di potenza osservata nel flusso dei raggi cosmici.

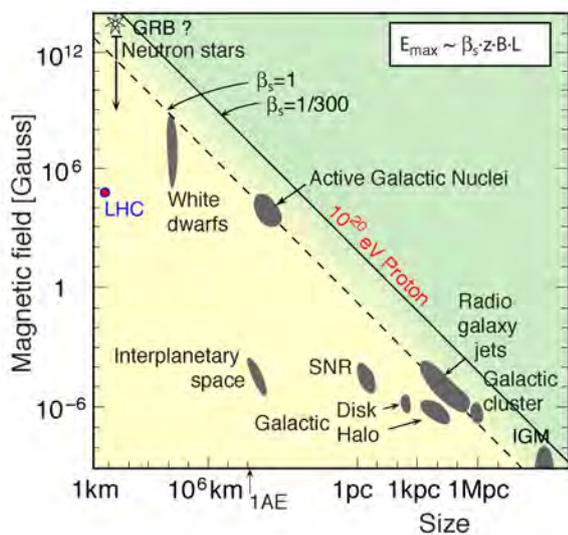


Figura 5: Plot di Hillas: sull'asse delle ordinate il campo magnetico, sull'asse delle ascisse la dimensione delle potenziali sorgenti di raggi cosmici.

Gli ingredienti principali del meccanismo di Fermi sono un oggetto compatto e le onde d'urto che si espandono da esso, ma le particelle non devono scappare subito dalla regione di accelerazione: ci vogliono un raggio e un campo ma-

gnético abbastanza grandi. Quali sono i possibili oggetti cosmici in cui queste condizioni possono essere realizzate? L'Eq. (2) implica la cosiddetta relazione di Hillas, che è illustrata nella Fig. 5. Le energie nel plot di Hillas sono le energie massime raggiungibili: l'efficienza del processo di accelerazione sono incorporate nel termine $\beta_s < 1$, che esprime l'efficienza dell'acceleratore descritto dall'Eq. (2). Si può vedere che:

- Le sorgenti capaci di accelerare particelle nella Galassia sono prevalentemente resti di supernova, detti brevemente SNR. In questo articolo intendiamo con questo termine i punti finali dell'evoluzione di grandi stelle: ciò che resta di queste stelle (oggetti compatti derivanti dal collasso e onde d'urto del materiale eiettato) dopo che hanno esaurito il loro combustibile e sono esplose. Le onde d'urto prodotte dalle esplosioni di supernova, che sopravvivono nelle SNR fino a qualche migliaio di anni dopo l'evento cataclismatico, sono capaci di accelerare particelle cariche ad energie fino al ginocchio. Abbiamo potuto recentemente misurare che alcune giovani (500 anni o meno) SNR si espandono a velocità $> c/100$ (Fig. 6 a sinistra), abbastanza per accelerare particelle ad energie relativistiche.
- Oltre il ginocchio, le possibili sorgenti possono essere AGN o collassi di grandi supernove o di sistemi binari, che hanno frequentemente una corrispondenza con l'emissione di un transiente veloce di raggi X e gamma, il cosiddetto lampo di raggi gamma o GRB (*gamma-ray burst*). Tuttavia, se i GRB fossero le fonti dominanti di raggi cosmici di altissima energia ci si aspetterebbe una corrispondente emissione di neutrini dovuta ai decadimenti dei pioni carichi prodotti dall'interazione degli adroni con il mezzo circostante, fenomeno che non si è osservato. Gli AGN devono quindi essere la fonte dominante di raggi cosmici di altissima energia.

Come localizzare gli acceleratori cosmici di particelle? L'Eq. (1) mostra che è molto difficile puntare le sorgenti usando particelle cariche: si ha bisogno di particelle di energia estremamente



Figura 6: A sinistra: il resto della supernova di Tycho, una delle sette supernove viste a occhio nudo nella storia. Apparve nel 1572 e fu indipendentemente osservata e studiata da molti scienziati, in particolare da Tycho Brahe. È relativamente giovane e si sta ancora espandendo ad una velocità di 3000 km/s, come misurato dall'osservazione diretta. A destra: L'AGN Centaurus A: un buco nero supermassiccio che si accresce ingoiando le stelle e le nubi molecolari vicine, ed emette getti di raggi gamma e particelle cariche (NASA).

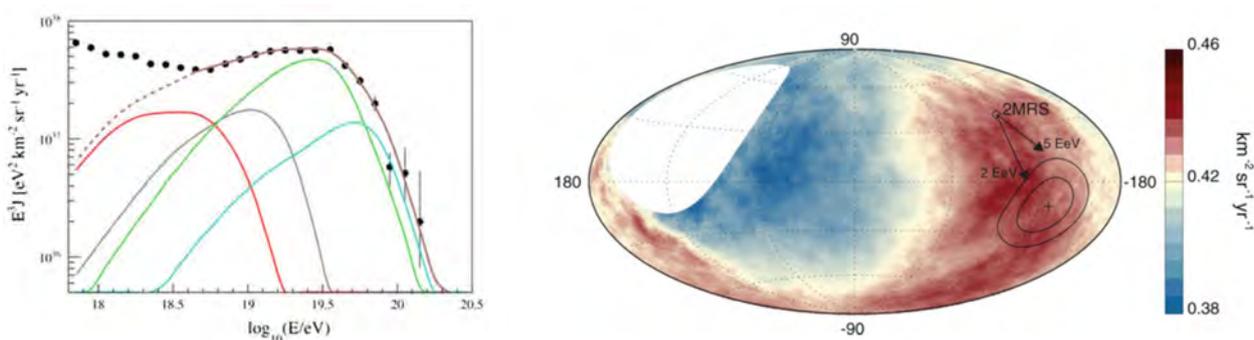


Figura 7: A sinistra: spettro ad altissima energia dei raggi cosmici misurati da Auger (pallini chiusi); lo spettro è stato moltiplicato per E^3 . Sovrapposto un fit dei dati ottenuto sommando diverse componenti: da sinistra a destra protoni, nuclei di He, una componente indistinta di nuclei di C, N, O, e nuclei di Fe. A destra: Mappa del cielo in coordinate galattiche che mostra il rapporto tra il numero di eventi osservati e quelli attesi per una distribuzione isotropa delle direzioni di arrivo, per $E > 32 \text{ EeV}$.

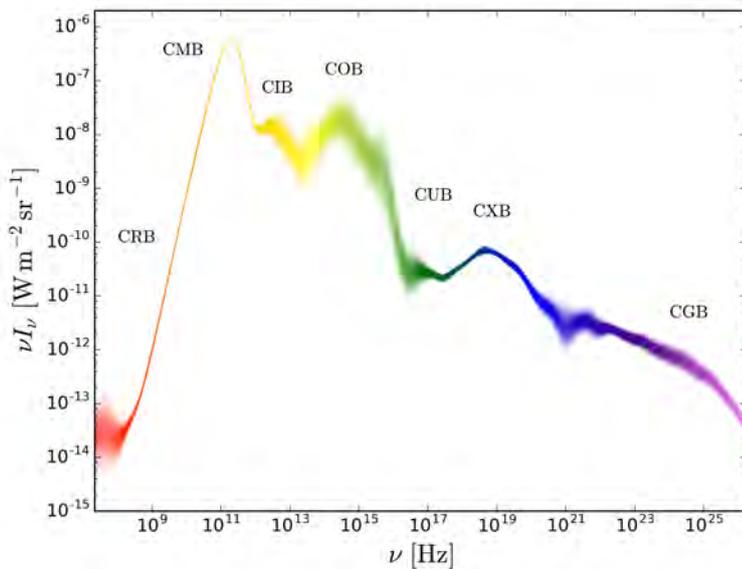


Figura 8: Distribuzione energetica della radiazione diffusa extragalattica di fondo [19]. La sigla CRB indica la radiazione di fondo (background) nella regione delle onde radio; CMB indica il cosmic microwave background, la radiazione fossile del Big Bang; CIB indica il background infrarosso, proveniente dal rescattering della luce visibile delle stelle sulle nuvole di plasma che popolano l'Universo; COB indica la radiazione ottica di fondo, proveniente principalmente dalle stelle; CUB indica la radiazione di fondo ultravioletta; CXB indica la radiazione di fondo nella regione dei raggi X; CGB indica il gamma-ray background.

alta, il che significa rivelatori molto grandi dato il rapido calo del flusso con l'energia.

L'Osservatorio Pierre Auger in Argentina è il più grande rivelatore del mondo, circa 3000 km² di superficie (1% della superficie totale dell'Italia), e rivela le particelle secondarie in sciame prodotti dalla interazione dei raggi cosmici primari con i nuclei d'aria nell'atmosfera (questo tipo di rivelatori sono chiamati rivelatori Extensive Air Shower, o EAS). Auger è un rivelatore ibrido, che impiega 1660 serbatoi d'acqua, distanziati di 1.5 km l'uno dall'altro, che rivelano le particelle cariche degli sciame attraverso la luce Čerenkov prodotta nella loro interazione con l'acqua stessa, e 4 telescopi che osservano lo sviluppo di sciame in aria mediante la luce ultravioletta di fluorescenza emessa nell'atmosfera. Il Telescope Array nello Utah funziona secondo lo stesso principio; i rivelatori Čerenkov ad acqua sono sostituiti da scintillatori e la superficie totale è circa 1/4 di quella di Auger.

L'Osservatorio Auger ha trovato due risultati particolarmente importanti: l'esistenza del *cut-off* GZK, con un probabile arricchimento della composizione in termini di nuclei più pesanti alle energie più alte (Fig. 7, a sinistra) e un'anisotropia dipolare che indica una correlazione posizionale dei raggi cosmici di più alta energia con AGN (Fig. 7 a destra). Anche se la correlazione con il campione di AGN mostra una significatività maggiore di 5 deviazioni standard, le correlazioni individuali non sono significative.

La più probabile (più di 3 deviazioni standard di significatività) è con l'AGN Centaurus A (Fig. 6 a destra).

La maggior parte delle localizzazioni di sorgenti di raggi cosmici di alta energia sono state possibili grazie ai raggi gamma. Nello spettro di fotoni dell'Universo (vedi Fig. 8) vediamo le tracce di molti sistemi termalizzati o quasi termalizzati: per cominciare, il CMB, che ha la più alta densità di fotoni (circa 400 per cm³). Salendo in energia, osserviamo altri spettri termici: quelli delle stelle, che raccontano la storia dell'evoluzione dell'Universo, e ad energie più alte gli spettri relativi alle regioni di accrescimento dei buchi neri, che corrispondono a temperature fino a milioni di kelvin - alcuni keV. Oltre queste energie si sondano meccanismi non termici: i fotoni prodotti sono il risultato di interazioni di altissima energia tra particelle nucleari. Questo chiarisce l'importanza dell'astrofisica dei raggi X e gamma.

Possiamo dividere in due grandi categorie gli strumenti per i raggi gamma: i rivelatori spaziali e quelli a terra. Queste due tipologie sono complementari. Lo spettro sperimentale dei raggi gamma abbraccia infatti più di 20 decadi (70 ottave) in energia e circa 14 in flusso, diminuendo rapidamente alle alte energie. È quindi chiaro che più grande è l'energia, più grande deve essere l'area efficace, definita come il prodotto dell'area per l'efficienza del rivelatore. A causa del costo della tecnologia spaziale, la superficie dei



Figura 9: A sinistra i due telescopi da 17 metri MAGIC nell'isola canaria di La Palma; a destra, il primo telescopio di grandi dimensioni CTA (23 m), inaugurato nell'ottobre 2018. Fotografia di Koji Noda.

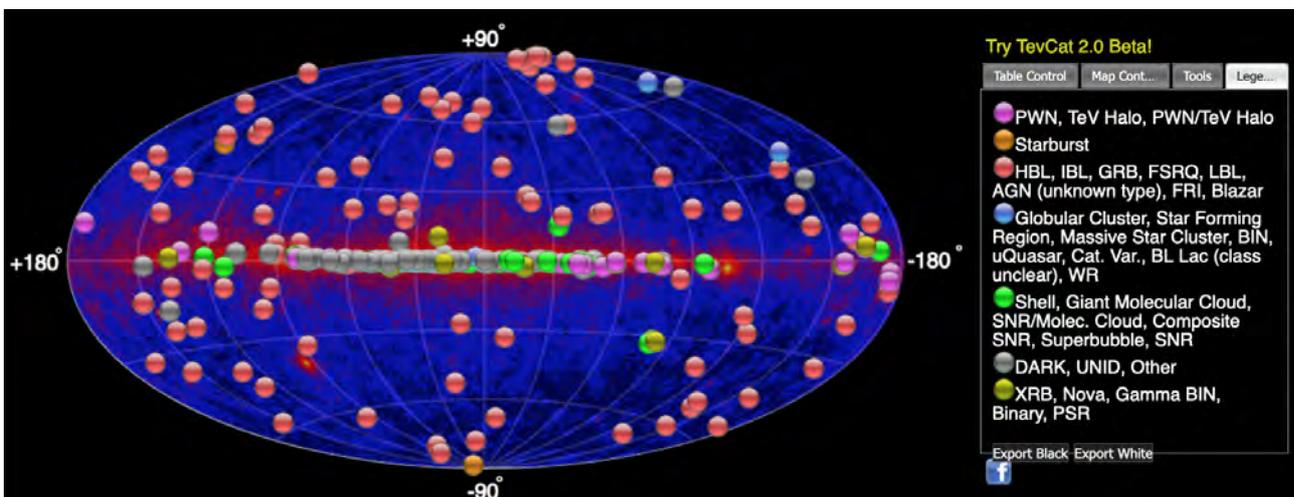


Figura 10: Mappa degli emettitori di fotoni sopra i 100 GeV nell'Universo, in coordinate galattiche (dal catalogo TeVCAT, marzo 2022). Le sorgenti sono indicate come cerchi; i colori rappresentano diversi tipi di emettitori (vedi la legenda). La maggior parte degli emettitori galattici è associata alle SNR, mentre gli emettitori extragalattici sono compatibili con le galassie attive – gli strumenti non hanno la risoluzione necessaria per studiare la morfologia delle galassie al di fuori del gruppo locale.

rivelatori non può comunque superare è circa 1 m^2 . Questo aspetto rende i rivelatori spaziali più adatti a misurare i raggi gamma nell'intervallo di energia dal MeV a qualche GeV. Per passare a energie più alte, sono necessarie grandi aree di rivelazione che possono essere ottenute solo a terra, sfruttando il fatto che, per energie superiori a circa 30 GeV, gli sciame elettromagnetici in aria iniziano a diventare rivelabili (mentre se l'energia è troppo bassa, lo sciame non può svilupparsi adeguatamente). Quando un raggio gamma entra nell'atmosfera, genera una cascata di particelle secondarie: il fotone si converte ad alta quota in coppie e^+e^- e ogni elettrone/positrone ad alta energia irradia raggi gamma secondari

soprattutto attraverso il meccanismo di *bremsstrahlung*; questi si convertono ulteriormente in coppie e^+e^- di energie inferiori, e così via.

Le cascate elettromagnetiche nell'atmosfera possono essere rivelate:

- misurando la radiazione Čerenkov di particelle cariche nell'aria (tecnica Čerenkov). La maggior parte delle particelle cariche secondarie hanno una velocità maggiore di quella della luce nell'aria ed emettono quindi radiazione Čerenkov, che può essere rivelata da telescopi speciali (Fig. 9). L'area effettiva di rivelamento è paragonabile a quella della proiezione della radiazione Čeren-

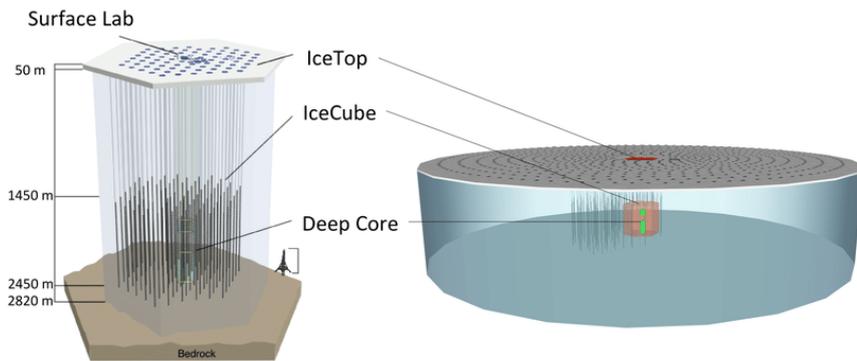


Figura 11: A sinistra: IceCube, uno strumento di un chilometro cubo di volume incorporato nel ghiaccio antartico. La sala di controllo è posta nella base antartica al di sopra della superficie del ghiaccio. In basso l'immagine della torre Eiffel dà un'idea della scala. A destra un'immagine di IceCube confrontato con la prevista estensione IceCube Gen2.

kov dello sciame a terra, cioè qualche decina di chilometri quadrati.

- rivelando direttamente le particelle cariche che raggiungono il suolo (tecnica EAS).

In Fig. 10 mostriamo una mappa degli emettitori di fotoni di altissima energia in coordinate galattiche; si possono identificare sia emettitori galattici (nel piano equatoriale) sia extragalattici. Molti degli emettitori di raggi gamma di altissima energia corrispondono posizionalmente a SNR, confermando così la congettura di Zwicky: la maggior parte degli acceleratori di raggi cosmici nella nostra Galassia sono SNR.

Un meccanismo per generare fotoni ad alta energia è simile a quello che produce i neutrini: ioni ad alta energia (protoni, nuclei di He, ...) accelerati attraverso il meccanismo di Fermi interagiscono con il mezzo circostante (raggi gamma, nubi di idrogeno) e producono adroni; circa un terzo degli adroni sono pioni neutri che decadono in circa 10^{-16} s in coppie di raggi gamma. Diversamente dal caso dei neutrini, tuttavia, meccanismi concorrenti generano fotoni senza la necessità di adroni accelerati, in particolare meccanismi di radiazione e interazioni Compton.

Una trentina di emettitori di raggi γ galattici sono stati associati a emettitori di raggi cosmici carichi: o attraverso l'identificazione di un picco nel flusso corrispondente al decadimento di π^0 da interazioni adroniche degli ioni accelerati con il mezzo circostante, o dall'*imaging* della sorgente, in cui si possono distinguere i raggi gamma irradiati da particelle accelerate nell'acceleratore primario da quelli provenienti da nubi molecolari (e quindi dalle collisioni di raggi cosmici che hanno interagito con i nuclei delle nubi molecolari producendo sciame adronici), o ancora, più

recentemente, rivelando fotoni dell'energia del PeV, per i quali la produzione adronica è quasi sicura (il meccanismo di produzione leptonica diventa inefficiente oltre i 100 GeV; le sorgenti sono dette in tal caso PeVatroni, e sono gli emettitori più energetici della Galassia). Solo dal 2021 è entrato in funzione a pieno regime uno strumento adatto a rivelare fotoni di queste energie: è il rivelatore EAS chiamato LHAASO (Large High Altitude Air Shower Observatory) [23].

LHAASO è un osservatorio costruito nel Sichuan, in Cina; si trova ad un'altitudine di 4400 metri sul livello del mare e copre un'area di circa 145 ettari. Consta di tre piscine di osservazione sotterranee, ciascuna contenente circa 100 000 tonnellate di acqua che viene utilizzata per generare luce Čerenkov; circa 6000 rivelatori di particella posti sulla superficie, e 18 telescopi Čerenkov ad ampio campo visivo.

L'energia dei protoni genitori è circa 20 volte più grande dell'energia dei raggi gamma (lo stesso rapporto vale anche per i neutrini). La localizzazione (molto recente) di una sorgente extragalattica ha richiesto però una nuova evoluzione dell'astrofisica delle alte energie, chiamata astronomia (o astrofisica) multimessaggero.

I raggi cosmici nell'era dell'astrofisica multimessaggero

Nel 1610 Galilei pubblica il *Sidereus Nuncius*, la prima opera scientifica basata sulle osservazioni fatte con il telescopio. In quel libro commenta e interpreta l'osservazione delle montagne della Luna, di centinaia di stelle mai viste prima e

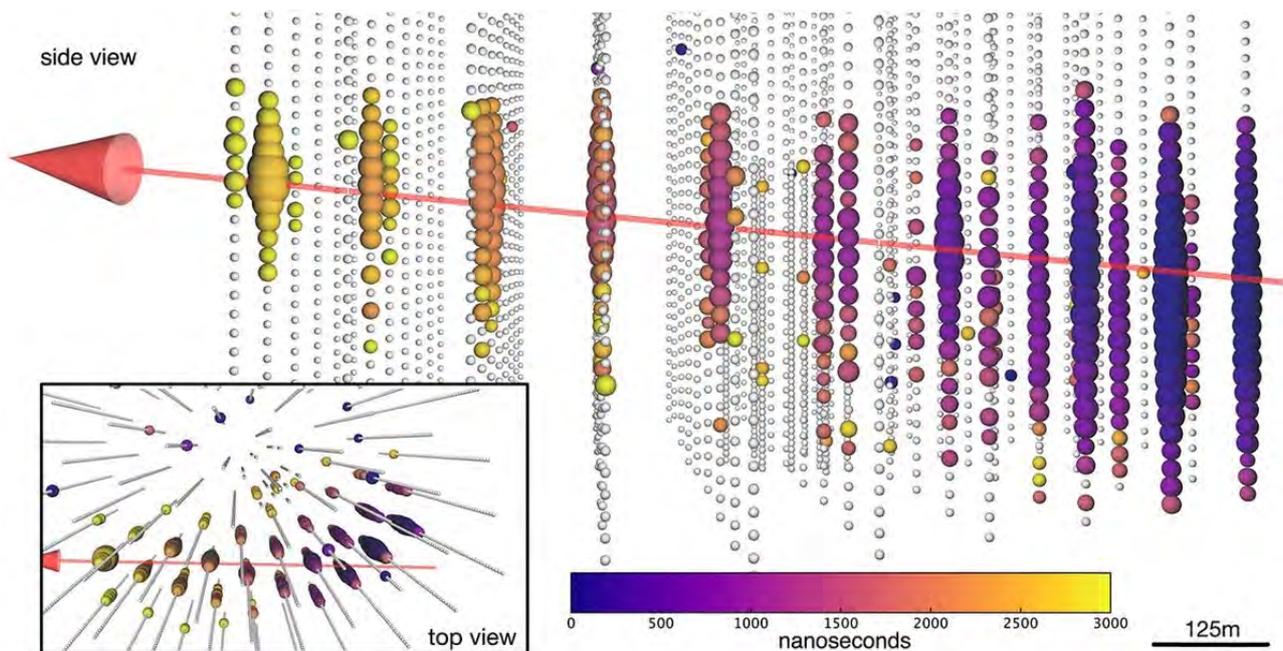


Figura 12: Il segnale del muone originato da un neutrino di 300 TeV proveniente dal blazar TXS 0506 +056 rivelato da IceCube (NSF). La sala di controllo è posta nella base antartica al di sopra della superficie del ghiaccio.

dei satelliti di Giove. Tutto questo era stato possibile grazie alla luce emanata dai corpi celesti: secondo Galilei, la luce è il *nuncius* (messaggero) delle stelle. La storia dei messaggeri cosmici, quindi, inizia con i fotoni, e il fotone è anche oggi il messaggero supremo, poiché è relativamente abbondante e punta alle sorgenti.

Gli esseri umani possono vedere solo una piccola parte dello spettro dei fotoni, corrispondente alle energie tipiche che gli elettroni atomici rilasciano quando si spostano a un livello energetico inferiore; la luce visibile si estende su un'ottava di frequenza. A partire dagli anni 1930, lo spettro di onde che possiamo osservare nell'Universo ha iniziato ad espandersi. Abbiamo potuto vedere nuovi colori invisibili a occhio nudo: le onde radio prodotte dal Sole e dalla galassia, miliardi e milioni di volte meno energetiche della luce visibile; poi le microonde, e infine i raggi X e i raggi gamma, da milioni a miliardi di volte più energetici della luce visibile. La storia dell'astronomia tra il 1930 e il 2015 può essere interpretata come un lungo viaggio alla scoperta dei colori dell'Universo invisibili all'occhio umano. In particolare per i raggi gamma, che sono meno abbondanti della luce visibile, questo ha richiesto lo sviluppo di strumenti innovativi: satelliti come il Fermi LAT che rivelano la conversione dei fotoni in coppie elettrone-positrone, telesco-

pi a terra che rivelano la luce Čerenkov emessa dalle particelle cariche negli sciami elettromagnetici generati dai raggi gamma, come MAGIC, H.E.S.S., Veritas, e recentemente grandi rivelatori a terra come LHAASO. Ora possiamo vedere 70 ottave di frequenza: una tastiera di pianoforte lunga dodici metri.

L'astronomia ha subito un'accelerazione nell'ultimo decennio. Nel febbraio 2016 è stata annunciata la prima rivelazione di un'onda gravitazionale. Le onde gravitazionali, prodotte nell'accelerazione di masse con asimmetria sferica, aprono un modo completamente nuovo di osservare gli eventi più violenti dell'Universo: viaggiano senza ostacoli alla velocità della luce e forniscono informazioni uniche sulle collisioni cataclismiche [20].

Nell'ottobre 2017, un altro annuncio rivoluzionario: per la prima volta sono state rivelate onde gravitazionali insieme a raggi gamma nella fusione di due stelle di neutroni. Questo processo può produrre elementi più pesanti del ferro ed è quindi centrale per l'evoluzione della vita come la conosciamo. Il segnale copre gran parte dello spettro elettromagnetico, dai raggi gamma ai raggi X, alla luce visibile e all'infrarosso, con un velocissimo gamma-ray burst. L'osservazione simultanea dell'onda gravitazionale GW170817 da parte della collaborazione LIGO/Virgo ha con-

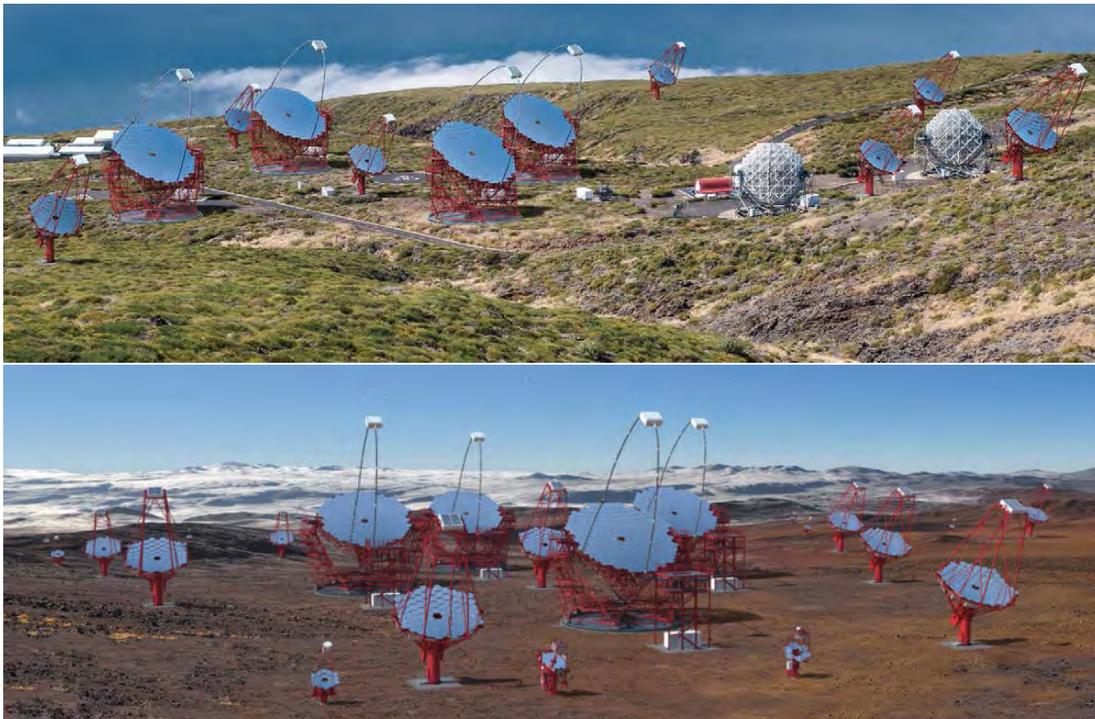


Figura 13: Rendering dell'array multi-telescopio CTA, in costruzione: sito settentrionale a La Palma (in alto) e sito meridionale a Paranal (in basso). Fonte: CTA.

giunto i modelli di collasso stellare dalla fisica teorica e le osservazioni del mondo reale [21].

Luglio 2018: un altro *big bang* nella scienza. Le collaborazioni IceCube (Fig. 12), Fermi e MAGIC (Fig. 9) hanno annunciato la rivelazione simultanea di un segnale di raggi gamma e di un neutrino dal blazar TXS 0506 +056, un AGN che si accresce a spese della massa circostante, con getti che puntano verso la Terra. Ancora una volta, abbiamo potuto risolvere un mistero: il confronto delle energie dei neutrini e dei raggi gamma ha mostrato che, nelle vicinanze del buco nero supermassiccio, la materia (protoni) è accelerata a energie decine di migliaia di volte superiori a quelle di LHC, e ha svelato il meccanismo di produzione adronica di raggi gamma e neutrini: collisioni subnucleari tra protoni e un fondo di fotoni ultravioletti. Questo processo è molto importante poiché indica che abbiamo localizzato per la prima volta un acceleratore di protoni oltre il ginocchio, a una quarantina di PeV [22].

Le onde gravitazionali e i neutrini hanno aperto il campo dell'astrofisica multi-messaggero, integrando ulteriormente la fisica delle particelle e l'astrofisica. Oggi possiamo iniziare a rispon-

dere ad alcune domande fondamentali che sembravano essere fuori dalla nostra portata. Dopo aver costruito strumenti capaci di osservare nuovi colori, stiamo sviluppando nuovi sensi e cominciamo a conoscerli. Come il tatto, l'olfatto, l'udito e il gusto ci danno informazioni sulla realtà che ci circonda, completando ciò che appare attraverso la vista, ora cominciamo a raccogliere e analizzare nuove informazioni da regioni remote dell'Universo trasmesse non dalla luce, ma da diversi messaggeri.

Cosa abbiamo imparato nei primi anni di questa nuova astronomia? Abbiamo spiegato il meccanismo che genera *burst* di raggi gamma brevi, della durata di meno di due secondi (coalescenza di stelle di neutroni) e lunghi (collasso di supernove di grande massa), fino a pochi anni fa un *puzzle* astrofisico. Abbiamo rivelato una galassia che emette simultaneamente neutrini e raggi gamma, scoprendo al suo centro un acceleratore da decine di milioni di GeV. Abbiamo localizzato una trentina di acceleratori di alta energia nella nostra Galassia. Abbiamo visto come si creano gli atomi più pesanti del ferro, e questo è importante per capire la vita. Abbiamo scoperto che l'Universo è sorprendentemente pieno

di piccoli buchi neri di poche decine di masse solari, e questo può avere una conseguenza sulla nostra comprensione del costituente dominante e ancora sconosciuto dell'Universo: la materia oscura.

Il futuro

Cosa prevediamo di fare e speriamo di capire nei prossimi 15 anni? Nuovi satelliti sensibili ai fotoni nella regione energetica tra MeV e GeV (ASTROGAM) potrebbero esplorare l'ancora sconosciuta regione del MeV, la regione dell'astrofisica nucleare. Nuovi strumenti sulla Terra sensibili ai fotoni nella regione energetica da centinaia di GeV a circa 100 TeV (come il telescopio CTA, Fig. 13, un ordine di grandezza più sensibile degli attuali strumenti per i raggi gamma) e oltre (i rivelatori a grande campo di vista LHAASO all'emisfero nord e in futuro il suo analogo SWGO all'emisfero sud) apriranno la strada alla comprensione delle più alte energie di raggi gamma nell'Universo e probabilmente porteranno alla scoperta di nuovi fenomeni fisici, in particolare in sinergia con i rivelatori di neutrini e onde gravitazionali. IceCube, che ha rivelato i primi segnali astrofisici di neutrini, passerà da 1 a 10 chilometri cubi, e un rilevatore complementare, Km³NeT, opererà nel Mar Mediterraneo. Il panorama dei rivelatori di onde gravitazionali si sta espandendo a una rete globale. I colori dei fotoni si arricchiranno di nuovi sapori nel contesto multimessaggero, e la fisica dei raggi cosmici rimarrà all'avanguardia della scienza.

Ringraziamento

Questo articolo è in gran parte una versione tradotta in italiano e aggiornata dell'articolo *Cosmic Rays* di A. De Angelis pubblicato sulla rivista *Nuovo Saggiatore*, vol. 35 (2019), p. 9.

Ringrazio Vincenzo Flaminio per l'accurata revisione.



[1] Si veda per esempio A. De Angelis, *L'Enigma dei Raggi Cosmici*, con prefazione di Margherita Hack, Springer, Milano (2010); seconda edizione in preparazione

- [2] Si veda per esempio A. De Angelis e M. Pimenta, *Introduction to Particle and Astroparticle Physics: Multimessenger Astronomy and its Particle Physics Foundations*, con prefazione di Francis Halzen, Springer, Heidelberg (2018)
- [3] K. Greisen, *End to the cosmic-ray spectrum?*, Phys. Rev. Lett. 16 (1966) 748
- [4] G.T. Zatsepin and V.A. Kuzmin, *Upper limit of spectrum of cosmic rays*, J. Exp. Theor. Phys. Lett. 4 (1966) 78
- [5] IceCube Collaboration, *Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector*, Science 342 (2013) 1242856
- [6] T. Wulf, *On the radiation of high penetrating power that exists in the atmosphere*, Phys. Zeit. 1 (1909) 152
- [7] Domenico Pacini, *La radiazione penetrante alla superficie ed in seno alle acque*, Nuovo Cim. VI/3 (1912) 93
- [8] Albert Gockel, *Measurement of penetrating radiation during balloon flights*, Phys. Zeit. 12 (1911) 595
- [9] Victor F. Hess, *Observations in low level radiation during seven free balloon flights*, Phys. Zeit. 13 (1912) 1084; trad. inglese in arXiv:1808.02927
- [10] L.W. Alvarez, Arthur H. Compton, *A positively charged component of cosmic rays*, Phys. Rev. 43 (1933) 835
- [11] T.H. Johnson, *Coincidence counter studies of the corpuscular component of the cosmic radiation*, Phys. Rev. 45 (1934) 569
- [12] B. Rossi, *Directional measurements on the cosmic rays near the geomagnetic equator*, Phys. Rev. 45 (1934) 212
- [13] Carl D. Anderson, *The positive electron*, Phys. Rev. 43 (1933) 491
- [14] I rapporti e le altre informazioni relative al Premio Nobel per la Fisica del 1936 a cui si fa riferimento nel testo sono materiale non pubblicato, ma disponibile per la ricerca storica presso il Centro di Storia della Scienza dell'Accademia Reale Svedese delle Scienze di Stoccolma, Svezia, e sono stati ottenuti grazie al professor Per Carlson del KTH di Stoccolma
- [15] B. Rossi, *Misure della distribuzione angolare di intensità della radiazione penetrante all'Asmara*, Suppl. la Ricerca Scientifica 1 (1934) 579
- [16] P. Auger, P. Ehrenfest, R. Maze, J. Daudin, R. A. Fréon, *Extensive Cosmic-Ray Showers*, Rev. Mod. Phys. 11 (1939) 288
- [17] J.W. Cronin, *The 1953 Cosmic Ray Conference at Bagnères de Bigorre*, Eur. Phys. J. H 36 (2011) 183
- [18] K. Niu, E. Mikumo, Y. Maeda, *Possible Decay in Flight of a New Type Particle*, Prog. Theor. Phys. 46 (1971) 1644
- [19] R. Hill, K.W. Masui, D. Scott, *The Spectrum of the Universe*, Applied Spectroscopy 72 (2018) 663
- [20] B.P. Abbott et al., *Observation of gravitational waves from a binary black hole merger*, Phys. Rev. Lett. 116 (2016) 061102
- [21] B.P. Abbott et al., *Multi-messenger Observations of a Binary Neutron Star Merger*, Astrophysical Journal Lett. 848 (2017) L12

- [22] The IceCube Collaboration et al., *Multimessenger observations of a flaring blazar coincident with high-energy neutrino IceCube-170922A*, *Science* 361 (2018) 6398
- [23] Z. Cao et al., *Ultrahigh-energy photons up to 1.4 PeV from 12 γ -ray Galactic sources*, *Nature* 594 (2021) 33



Alessandro De Angelis: è un fisico delle alte energie e un astrofisico, professore a Padova e a Lisbona. Il suo principale interesse di ricerca è la fisica fondamentale (in particolare l'astrofisica e la fisica delle particelle elementari); recentemente studia soprattutto la propagazione cosmologica dei fotoni, ed è il Principal Investigator del progetto spaziale ASTROGAM. È stato *staff member* al CERN fino al 2000, e successivamente tra i membri fondatori del telescopio Fermi della NASA e per molti anni Project Scientist dei telescopi MAGIC a La Palma. Affianca alla passione per la ricerca la passione per la storia della scienza. È stato *visiting professor* all'ICRR di Tokyo, all'Istituto Max Planck di Monaco, a Parigi VI.

Caratteristiche principali degli esperimenti per la misura dei raggi cosmici da terra

Andrea Chiavassa

Dipartimento di Fisica - Università degli Studi di Torino

Introduzione

Quando una particella della radiazione cosmica impatta con l'atmosfera terrestre, interagisce con un nucleo della stessa, viene distrutta e non raggiunge la superficie terrestre. Questo ne impedisce la misura diretta con strumentazione posta a terra. La misura diretta dei raggi cosmici è solo possibile per mezzo di esperimenti operanti al di fuori dell'atmosfera e operanti quindi su satelliti. Questo limita il range energetico che può essere studiato con questa tecnologia; infatti, dovendo essere posti in orbita, le dimensioni di questi apparati sono vincolate sia in estensione che in massa, ponendo due grandi limitazioni. La prima è dovuta alla statistica. Il numero di raggi cosmici diminuisce rapidamente con l'energia (circa $E^{-2.7}$, vedi figura 1), quindi, con apparati che hanno una superficie dell'ordine della decina di metri quadrati, non si riesce ad accumulare un numero di eventi sufficiente nel tempo di operazione di un esperimento (tipicamente dieci/quindici anni). La seconda limitazione è dovuta alla misura dell'energia delle particelle primarie che è solitamente effettuata con calori-

metri al cui interno viene assorbita e rivelata tutta l'energia della particella. Al crescere dell'energia è necessaria una massa sempre maggiore per poter contenere tutta l'energia della particella. Il limite sulla massa degli strumenti che si possono portare in orbita implica che la precisione con cui si misura l'energia del primario diminuisce al suo crescere (diminuendo la frazione di energia del primario contenuta nello strumento).

Tutto questo comporta che oltre una certa energia (pari a circa 10^{14} eV) la misura dei raggi cosmici primari può solo essere effettuata con esperimenti che operano a terra dove si può ovviare ai problemi di statistica realizzando esperimenti di grandi superfici (da 10^4m^2 a 10^3km^2), perdendo al contempo la possibilità di misurare direttamente le particelle primarie.

In seguito all'interazione dei raggi cosmici con i nuclei dell'atmosfera si genera una cascata di particelle (solitamente indicata con l'acronimo EAS, *Extensive Air Shower*). Il numero di particelle comprese nello sciame cresce all'aumentare dell'energia della particella primaria. Man mano che lo sciame procede nel suo sviluppo diminuisce l'energia delle particelle che lo compongono.

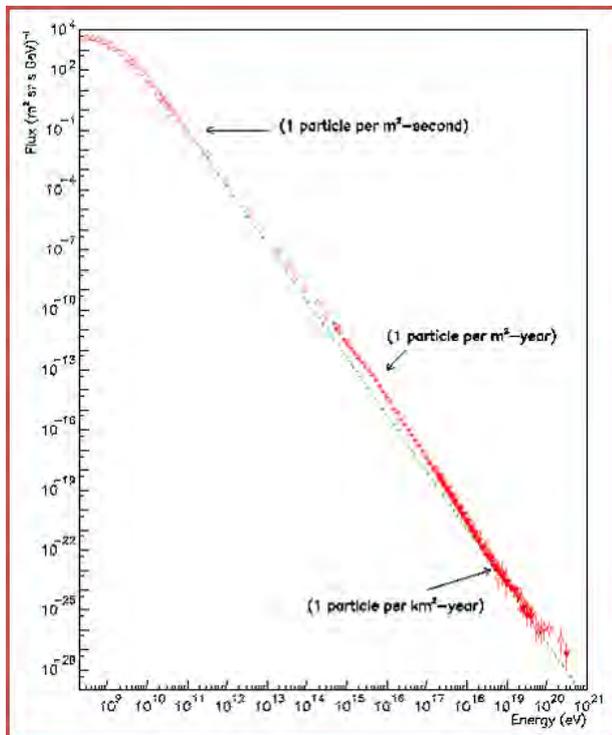


Figura 1: Spettro differenziale dei raggi cosmici

Questo processo prosegue fino a che l'energia delle particelle nella cascata è tale da permettere la loro moltiplicazione, quando questa non è più sufficiente alla loro moltiplicazione le stesse depositano la loro energia in atmosfera, venendo assorbite. Il numero di particelle comprese nella cascata a questo punto diminuisce e la stessa viene rapidamente estinta. È molto importante evidenziare che le cascate che si sviluppano in atmosfera comprendono diversi tipi di particelle: le più abbondanti sono fotoni ed elettroni, seguite da muoni e da una piccola componente di adroni (principalmente pioni).

Senza entrare nei dettagli dello sviluppo degli EAS possiamo dire che: l'energia del primario (E_0) è proporzionale al numero di particelle (N_e numero di elettroni $N_e \propto E_0$, N_μ numero di muoni $N_\mu \propto E_0^\beta A^{1-\beta}$, dove A è il numero di massa del primario, e $\beta < 1$) alla quota di osservazione. La direzione di arrivo è ricavabile dal tempo di arrivo delle particelle dello sciame e la natura della particella primaria (tra cui A) si può inferire da diversi parametri quali il rapporto tra le diverse particelle comprese nello sciame, la sua morfologia e la quota in atmosfera ($X_{max} \propto (\ln E_0 - \ln A)$) in cui lo sciame raggiunge la massima estensione (cioè il massimo numero di particelle). Quindi, con esperimenti

che misurano uno sciame alla sola quota di osservazione, la conoscenza del numero di elettroni, quello di muoni e del tempo di transito è sufficiente per risalire alle caratteristiche del primario. Se invece siamo in grado di misurare tutto lo sviluppo longitudinale dello sciame possiamo ottenere una misura quasi calorimetrica dell'energia della particella primaria in quanto utilizziamo l'atmosfera come un calorimetro al cui interno viene depositata tutta l'energia della particella primaria. Questa misura è solo possibile ad altissime energie, dell'ordine di $\sim 10^{17}$ eV, con la tecnica della misura della luce di fluorescenza (sulla quale torneremo più avanti). È anche importante sottolineare come durante lo sviluppo di un EAS le particelle che lo compongono si distribuiscono attorno all'asse dello sciame per centinaia di metri e su distanze sempre maggiori al crescere dell'energia del primario. Nella figura 2 è mostrato un esempio di simulazione dello sviluppo longitudinale di uno sciame atmosferico esteso generato da un protone di energia 10^{15} eV; mentre nella figura 3 è riportata la simulazione dello sviluppo laterale dello stesso sciame.

Avendo delineato quali sono le caratteristiche principali e comuni a tutti gli esperimenti che operano rivelando degli EAS dobbiamo differenziarli a seconda del *range* energetico cui siamo interessati e al tipo di particelle della radiazione cosmica che vogliamo studiare.

La differenziazione dovuta al tipo di particella primaria cui siamo interessati ci porta a dividere gli esperimenti in due famiglie: quelli interessati alla misura di EAS generati da fotoni primari e quelli che studiano la radiazione cosmica elettricamente carica. Nel primo caso, il punto essenziale è selezionare tra tutti gli eventi misurati quelli, meno dell'1%, che sono originati da fotoni primari. Il principale interesse di questi esperimenti è dovuto al fatto che, essendo i fotoni elettricamente neutri, non sono deviati dai campi magnetici galattico ed extra-galattico, questo permette di risalire agli oggetti celesti che li hanno originati, e di fare dell'astrofisica delle altissime energie.

Nel caso di esperimenti dedicati allo studio della radiazione cosmica carica non è possibile (come vedremo) determinare il tipo di particella che ha originato l'EAS. Le particelle primarie, essendo deviate dai campi magnetici, quando im-

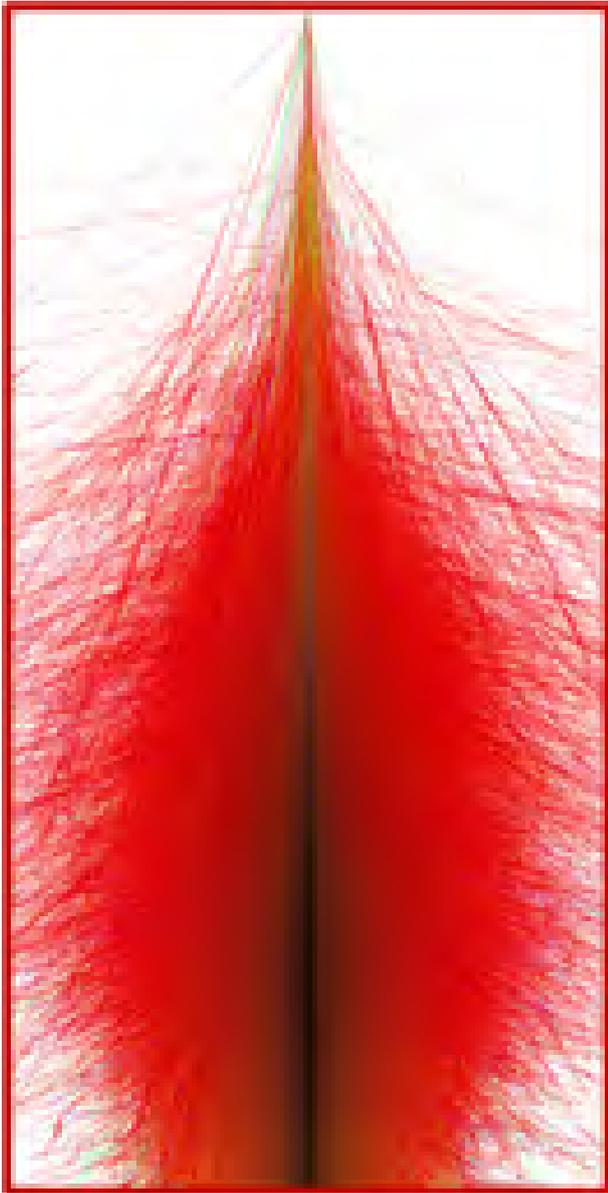


Figura 2: Simulazione dello sviluppo longitudinale di un EAS generato da un protone di energia 10^{15} eV. Nel grafico sono rappresentate le tracce delle singole particelle originate nello sviluppo dello sciame. Le ordinate rappresentano l'altezza sul livello del mare (o la profondità atmosferica attraversata). Si può osservare come il numero di particelle aumenti man mano che lo sciame si evolve in atmosfera, raggiunge un valore massimo e poi diminuisce.

pattano con l'atmosfera terrestre non provengono dalla direzione della loro sorgente. Gli studi di queste particelle sono comunque molto importanti perchè la quantità di energia distribuita nella nostra galassia sotto forma di particelle relativistiche ed elettricamente cariche è confrontabile con le altre principali forme di energia diffusa, ad esempio quella del campo magnetico galat-

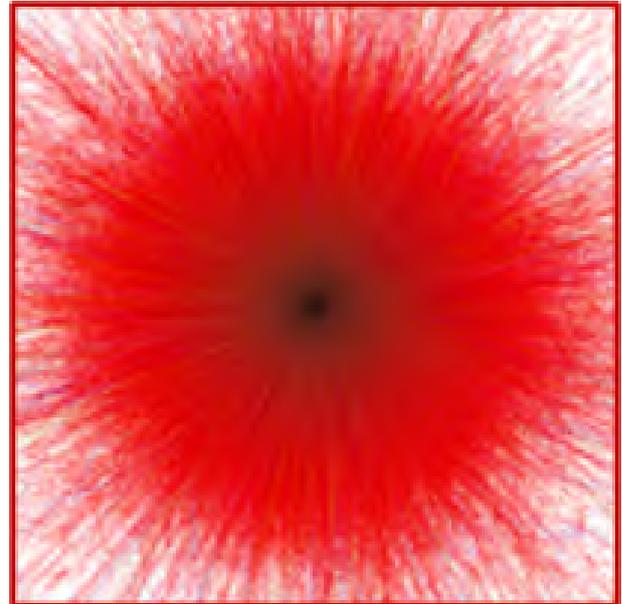


Figura 3: Simulazione dello sviluppo laterale di un EAS generato da un protone di energia 10^{15} eV. È rappresentato l'allargamento delle particelle dello sciame, nel piano orizzontale, attorno al suo asse (retta centrale che rappresenta la proiezione della direzione di arrivo del raggio cosmico che ha originato lo sciame). Le dimensioni di questa figura sono dell'ordine delle centinaia di metri.

tico, e non può quindi essere ignorata. Inoltre, questi eventi permettono uno studio dei fenomeni più energetici che avvengono nell'universo e, non ultimo, in essi si raggiungono energie che non sarà mai possibile produrre agli acceleratori di particelle costruiti dall'uomo.

Questo articolo è quindi organizzato con una sezione in cui descriverò le caratteristiche generali di questi esperimenti: tecniche di misura utilizzate, dimensioni degli esperimenti, quota sul livello del mare cui realizzare gli apparati di misura, problematiche relative alla loro calibrazione. Nella terza sezione descriverò brevemente alcuni dei più importanti (a mio giudizio) esperimenti recentemente realizzati; concluderò questa breve discussione facendo il punto sullo stato attuale di questo tipo di studi e cercando di evidenziare possibili sviluppi futuri.

Generalità degli esperimenti per la misura degli EAS

Seppur schematicamente abbiamo visto come un EAS sia composto da vari tipi di particelle.

Queste sono, in ordine di numerosità e senza differenziare tra particelle e le corrispondenti antiparticelle: fotoni, elettroni, muoni e adroni in generale (principalmente pioni). Quindi, con gli apparati per la misura degli EAS, dobbiamo misurare direttamente queste particelle oppure rivelare un segnale da loro generato.

Riassumo brevemente alcuni degli strumenti che è possibile utilizzare per la misura delle particelle degli EAS.

Scintillatori

Uno degli strumenti più utilizzati per la misura delle particelle degli EAS sono gli scintillatori, in cui le particelle cariche comprese nello sciame depositano parte della loro energia eccitando gli atomi che emettono dei fotoni. Con dei fotomoltiplicatori (o altri sensori di luce come i SiPM, *silicon photomultipliers*) vengono raccolti i fotoni emessi, generando un segnale di ampiezza proporzionale all'energia depositata nello scintillatore. Questi strumenti sono molto utilizzati per contare il numero di particelle che attraversano il rivelatore e misurare il tempo di transito delle stesse (con una precisione dell'ordine dei ns). Questi strumenti possono anche essere posti sotto un opportuno strato di materiale che assorbe gli elettroni permettendo il transito della sola componente muonica dello sciame, che può quindi essere misurata.

Rivelatori Čerenkov ad acqua

Un altro modo per misurare il numero di particelle nello sciame e misurarne il tempo di transito è quello dei rivelatori a Čerenkov ad acqua. Le particelle cariche dello sciame attraversando l'acqua del rivelatore emettono luce Čerenkov che viene poi rivelata da fotomoltiplicatori o SiPM. Anche in questo caso la quantità di luce rivelata è proporzionale al numero di particelle che attraversano il rivelatore e si ottiene una misura del tempo di transito delle particelle. Si utilizzano questi rivelatori invece degli scintillatori per due motivi, la prima è principalmente economica essendo il costo minore di quello degli scintillatori. Il secondo è legato al fatto che, realizzando rivelatori con almeno un metro di acqua (in altezza) si ha una grande probabilità che i fotoni dello sciame interagiscano producendo una

coppia elettrone-positrone che a sua volta genera della luce Čerenkov che può essere misurata. Questi rivelatori sono quindi più sensibili degli scintillatori ai fotoni.

Rivelatori di posizione

In alcune applicazioni, e soprattutto per la realizzazione di rivelatori di dimensioni relativamente contenute posti nella parte centrale dell'esperimento, si possono utilizzare camere a filo che permettono di tracciare la singola particella oltre a misurare l'energia che questa deposita in questi strumenti.

Rivelatori di luce Čerenkov

Le particelle cariche degli EAS, muovendosi in atmosfera con velocità maggiore di quella della luce nel mezzo, emettono luce Čerenkov; questa può essere raccolta o utilizzando degli specchi con dei fotomoltiplicatori posti nel loro fuoco o con fotomoltiplicatori direttamente esposti all'aria aperta. La luce Čerenkov emessa durante lo sviluppo degli EAS è molto poca (circa 10^{-4} volte della luce notturna) ma essendo concentrata in una decina di nanosecondi è possibile rivelarla. Da questo si comprende come questi strumenti possano lavorare solo di notte ed in assenza di luna piena (duty cycle circa del 10-15%). In compenso permettono una misura quasi calorimetrica dell'energia del primario in quanto tutti i fotoni Čerenkov emessi durante lo sviluppo dell'EAS non vengono successivamente assorbiti e possono essere misurati.

Rivelatori di luce di fluorescenza

Un altro tipo di strumentazione con cui si misura della luce emessa durante lo sviluppo degli EAS sono i cosiddetti rivelatori di luce di fluorescenza. Il segnale che si misura con questi strumenti è la luce emessa in atmosfera a causa dell'eccitazione e del successivo ritorno allo stato fondamentale degli atomi dell'atmosfera: in sostanza si utilizza l'atmosfera come un enorme scintillatore. Questi strumenti permettono di misurare la luce emessa nei vari stadi di sviluppo in atmosfera degli EAS e permettono quindi una misura del loro sviluppo longitudinale. Si ottiene quindi una misura calorimetrica dell'energia del primario

e si misura il punto di massimo sviluppo dello sciame. Anche questi rivelatori (come quelli di luce Čherenkov) operano solo di notte ed in assenza di luna, quindi hanno un basso duty cycle, inoltre, essendo l'aria un pessimo scintillatore, la luce emessa è poca e quindi questi strumenti funzionano solo ad energie superiori a 10^{17} eV. In compenso, come accennato in precedenza, forniscono una misura calorimetrica dell'energia del primario e una misura diretta del punto di massimo sviluppo dell'EAS (misure al momento possibili solo con questa tecnica). Dobbiamo infine notare che mentre la luce Čherenkov degli EAS è emessa in un cono di angolo molto piccolo rispetto alla direzione di propagazione dello sciame, la luce di fluorescenza è emessa isotropicamente. Quindi, con un rivelatore che osserva frontalmente lo sciame il segnale è dominato dalla luce Čherenkov, mentre con uno posto lateralmente domina quella di fluorescenza.

Da questa veloce carrellata delle tecniche di misure utilizzabili per rivelare gli EAS vediamo che possiamo misurare per ogni evento il numero totale di particelle nello sciame alla quota di osservazione, che abbiamo la possibilità di separare il conteggio della componente penetrante (muoni) dello sciame, e che possiamo determinare il tempo di transito delle particelle.

Dimensioni degli apparati per la misura di EAS e geometria dei rivelatori

Come mostrato in figura 1 il flusso di raggi cosmici varia rapidamente con l'energia, passando da una particella al secondo per metro quadrato a 10^{15} eV fino a una particella al secolo per chilometro quadrato alle energie estreme ($> 10^{19}$ eV). Questo ci dice subito come, per poter ottenere un numero di eventi statisticamente significativo, le dimensioni degli apparati debbano essere molto diverse a seconda dell'intervallo energetico che vogliamo studiare: per la ricerca di sorgenti di fotoni ($E < 1$ PeV) e per misure relative ai raggi cosmici galattici ($E < 10^{17}$ eV) sono sufficienti apparati di superficie compresa tra 10^4 e 10^6 m²; mentre per studiare i raggi cosmici di energia estrema (di origine extra-galattica, $E > 10^{18}$ eV) la superficie necessaria è di migliaia di km².

Dato che l'osservabile sperimentale che si può correlare all'energia del primario è il numero di

particelle alla quota di osservazione, idealmente si potrebbe coprire tutta la superficie dell'esperimento di strumentazione sensibile al passaggio delle particelle e semplicemente contare quante particelle sono contenute in ciascuno sciame. Chiamiamo *fill factor* il rapporto tra la superficie attiva e quella totale dell'esperimento. D'altra parte, è immediatamente evidente come non sia possibile coprire superfici molto grandi di rivelatori, questo ci dice come il *fill factor* tipico degli esperimenti vari a seconda dell'intervallo di energia delle particelle primarie che si vogliono studiare con i diversi esperimenti. Di conseguenza varierà anche molto la disposizione geometrica dei rivelatori. Possiamo quindi schematizzare questi due aspetti dandone una descrizione a seconda dell'intervallo energetico e del tipo di primari che vogliamo investigare con un esperimento.

Per lo studio delle energie appena superiori a quelle investigabili con misure dirette, quindi tra le centinaia di GeV e circa dieci TeV, è necessario rivelare EAS formati da poche decine di particelle. In particolare, con gli esperimenti dedicati alla ricerca di sorgenti di fotoni di queste energie la misura della distribuzione spaziale di particelle dello sciame è il parametro che meglio permette di selezionare eventi generati da fotoni tra quelli generati da adroni. Si devono quindi realizzare apparati con *fill factor* molto alto ($>70\%$) per poter misurare la maggior parte delle particelle dello sciame e per ricostruirne l'immagine a terra con buona precisione. Esempi di apparati di questo genere sono MILAGRO [1], LHAASO-WCDA [2] in cui sono state realizzate delle piscine d'acqua (*fill factor* circa del 100%); HAWC [3] composto da *tank* cilindriche di acqua affiancate tra di loro; ARGO-YBJ [4] costituito da un tappeto di rivelatori RPC (*Resistive Plate Chambers*), di forma rettangolare, disposti l'uno a fianco dell'altro. La dimensione tipica di questi apparati è dell'ordine dei 10^4 m², la dimensione della cella elementare (che definisce la granularità con cui posso ricostruire il profilo laterale dello sciame) va dalle centinaia di cm² (ARGO-YBJ) alla decina di m² degli altri esperimenti.

Salendo con l'energia, avendo quindi come scopo la ricerca di sorgenti di fotoni di energia maggiore di cento TeV, o lo studio dei raggi cosmici carichi di origine galattica ($10^{14} < E < 10^{17}$

eV, la superficie degli esperimenti arriva fino ai 10^6 m^2 e il *fill factor* che si può raggiungere è molto più basso, tipicamente inferiore al 5%. Infine, arrivando alle energie estreme dello spettro gli apparati coprono migliaia di chilometri quadrati e il *fill factor* deve scendere ulteriormente arrivando a valori pari a circa $10^{-4} - 10^{-5}$. In questi esperimenti i rivelatori vengono disposti su una superficie, possibilmente piana, con geometria regolare (sempre compatibilmente con la logistica del sito sperimentale). Solitamente i rivelatori vengono disposti ai vertici o di quadrati (KASCADE [5] o di triangoli equilateri (KASCADE-Grande [6], TUNKA [7], LHAASO-KM2A [2]), la distanza tra i vari elementi dell'apparato varia a seconda delle dimensioni dello stesso e del *fill factor* che si vuole (può) raggiungere. Con questi apparati si campiona il numero di particelle in diversi punti del piano, da queste misure si ricostruisce il profilo laterale dello sciame il cui fattore di normalizzazione è un'osservabile sperimentale da cui si può stimare l'energia del primario (vedi in seguito la discussione sulla calibrazione di questi apparati), la distribuzione laterale di un evento è mostrata come esempio in figura 4. Salendo ulteriormente in energia e

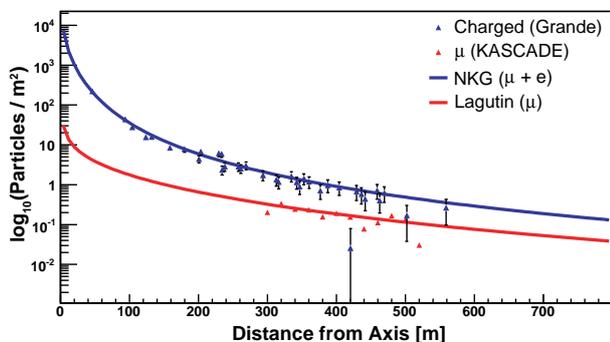


Figura 4: Distribuzione laterale misurata per un evento di raggi cosmici (esperimento KASCADE-Grande [6]).

arrivando quindi allo studio delle energie estreme dello spettro primario ($E > 10^{18}$ eV) dove la radiazione cosmica è prevalentemente di origine extra-galattica, i flussi diventano così bassi che per poter accumulare una statistica sufficiente gli apparati devono avere estensioni enormi. I due maggiori esperimenti che operano a queste energie sono il Pierre Auger Observatory [8] e Telescope Array [9], che si estendono su superfici di 3000 km^2 e di 700 km^2 rispettivamente. In que-

sto caso il *fill factor* è dell'ordine di $10^{-4} - 10^{-5}$, la disposizione dei rivelatori è su una griglia triangolare di lato 1500 m nel caso di Auger (la superficie del singolo elemento è 10 m^2) e su una griglia quadrata di lato 1200 m per Telescope Array (area del singolo elemento 3 m^2). A queste energie è anche possibile la rivelazione della luce di fluorescenza degli EAS ed in entrambe gli esperimenti sono stati posti ai margini della superficie coperta dagli apparati dei telescopi di luce di fluorescenza.

Quota sul livello del mare

Uno dei parametri fondamentali da stabilire nell'identificare il luogo in cui costruire un apparato è la quota sul livello del mare. Osservando lo sviluppo longitudinale di uno sciame (vedi figura 2) possiamo notare come il numero di particelle aumenta fino a raggiungere un punto di massimo sviluppo, dopodiché le particelle vengono assorbite e il numero di particelle nello sciame si riduce. Già da questo è chiaro che, per avere il massimo numero di particelle disponibili per la rivelazione, la localizzazione ideale di un apparato è alla quota cui lo sciame raggiunge il massimo sviluppo. Questo è tanto più importante quanto è più bassa l'energia dei primari che si vuole studiare. In questo senso, gli apparati per studiare i primari di centinaia di GeV devono essere costruiti sopra i 4000 m sul livello del mare s.l.m. (4100 m HAWC, 4300 m ARGO-YBJ, 4410 m LHAASO).

Nel punto di massimo sviluppo di uno sciame il numero di elettroni non dipende dalla natura del primario (il che rappresenta un vantaggio per la calibrazione energetica, vedi sezione seguente); a questa profondità atmosferica solo la componente muonica degli sciami è sensibile alla natura del primario. Quando invece l'esperimento osserva gli EAS nella fase di assorbimento, dopo aver raggiunto il massimo sviluppo, sia il numero di elettroni che quello di muoni alla quota di osservazione dipendono dalla natura del primario (oltre che dalla sua energia). Esperimenti posti a queste quote hanno maggiori incertezze nella calibrazione energetica ma hanno una maggiore sensibilità alla natura del primario. Gli apparati per lo studio della radiazione cosmica di origine galattica sono stati realizzati a

quote molto diverse tra di loro dai 100 m s.l.m. di KASCADE, ai 600 m s.l.m. di TUNKA, ai 2850 m s.l.m. di ICE-TOP.

Esperimenti che misurano lo sviluppo longitudinale dello sciame (ad esempio i rivelatori di luce di fluorescenza), per poter eseguire una misura quasi calorimetrica dell'energia, devono poter osservare tutto lo sviluppo dello sciame e quindi devono essere realizzati ben al di sotto della quota di massimo sviluppo dello sciame. È importante osservare come tutto questo è riferito a sciami che arrivano in direzione verticale (angolo zenitale $\theta = 0^\circ$), quando un evento è rivelato ad un angolo zenitale θ , rispetto alla verticale, lo spessore atmosferico attraversato è aumentato di un fattore $1/\cos\theta$, di conseguenza è possibile osservare lo sviluppo completo di uno sciame anche essendo a quote atmosferiche relativamente elevate (per esempio l'esperimento Auger è realizzato a circa 1500 m s.l.m.).

Gli esperimenti che rivelano luce Čerenkov o di fluorescenza emessa in atmosfera è indispensabile che siano costruiti in luoghi remoti lontani da sorgenti artificiali di luce.

Solitamente, a parte rare eccezioni, gli esperimenti dedicati alla misura di EAS non sono realizzati all'interno di centri di ricerca attrezzati, rendendo gli aspetti logistici (dalle vie di accesso, alla vicinanza di strutture ricettive, alla disponibilità di manodopera locale, alla disponibilità di potenza elettrica sufficiente per operare la strumentazione e di una rete internet per la trasmissione di dati) un aspetto molto rilevante nella scelta del luogo in cui costruire gli apparati.

Calibrazione degli apparati per la misura di EAS

Non esistendo delle sorgenti calibrate di raggi cosmici il problema della calibrazione degli apparati per la misura di EAS è molto rilevante e rappresenta la maggior sorgente di errori sistematici nella determinazione delle caratteristiche delle particelle primarie.

In realtà per quanto riguarda la calibrazione assoluta della misura della direzione di arrivo dei raggi cosmici è vero che non esiste una sorgente puntiforme, ma essendo la luna un corpo celeste di dimensioni non trascurabili, vicino alla terra, ci aspettiamo una diminuzione del flusso

di raggi cosmici da questa direzione. Essendo le dimensioni della luna dell'ordine di 0.5° se la risoluzione angolare di un apparato sciami è di questo ordine di grandezza un esperimento deve misurare un deficit di raggi cosmici provenienti dalla direzione della luna. Questa tecnica è ampiamente sfruttata da tutti gli esperimenti dedicati alla misura dei raggi cosmici carichi. Gli esperimenti dedicati alla ricerca di sorgenti di fotoni devono invece dimostrare di osservare la sorgente di fotoni di alta energia più intensa conosciuta, la CRAB Nebula (se questa si trova nel loro campo di vista). Questo permette anche di valutare la capacità di reiezione del fondo adronico da parte dell'esperimento.

Per quanto riguarda la calibrazione energetica questi esperimenti non possono in alcun modo utilizzare una sorgente calibrata e devono quindi ricorrere a una procedura indiretta che sfrutta una simulazione completa dello sviluppo di un EAS. In questo momento il codice più utilizzato e di riferimento nella comunità si chiama CORSIKA, è stato sviluppato presso il Karlsruhe Institute for Technology ed è un software ad accesso libero. Una simulazione degli EAS deve partire dalla prima interazione tra la particella primaria e un nucleo dell'atmosfera e procedere poi attraverso tutte le successive interazioni di energia inferiore ed i decadimenti delle particelle. Le interazioni tra particelle adroniche sono studiate agli acceleratori di particelle, le energie più alte sono state raggiunte al collisore protone-protone LHC in funzione al CERN di Ginevra. Con questo acceleratore si è raggiunta un'energia nel centro di massa che corrisponde a quella dell'interazione di un raggio cosmico di energia dell'ordine dei 10^{16} eV. Le simulazioni di EAS generati da particelle di energia superiore ricorrono a modelli di interazione adronica che non possono essere testati agli acceleratori di particelle, è chiaro che questo introduce una incertezza nei risultati di queste simulazioni. Con queste simulazioni si possono riprodurre EAS di energia e primario di natura note e ricavare il numero di elettroni e di muoni alla quota di osservazione, nonché la quota cui lo sciame ha raggiunto il massimo sviluppo.

Bisogna poi considerare altri fattori che fanno variare il numero di particelle in uno sciame alla quota di osservazione, *in primis* la natura

del primario. Come accennato in precedenza, a parità di energia, sciame generati da particelle pesanti raggiungono il massimo sviluppo più in alto in atmosfera e quindi quando sono osservati durante la fase del loro assorbimento sono più lontani dal punto di massimo sviluppo. Questo fa sì che sotto il punto di massimo sviluppo il numero di elettroni in un EAS dipenda, oltre che dall'energia del primario, dalla sua natura: più è alta la massa del primario (A) minore è il numero di elettroni generati. Inoltre, il numero di elettroni in uno sciame dipende (a parità di altezza sul livello del mare) dall'angolo zenitale θ cui si osserva lo sciame. Uno sciame osservato a θ maggiore ha attraversato una quantità di atmosfera maggiore e quindi si trova più lontano dal punto di massimo sviluppo, in una fase di assorbimento più avanzata: a parità di energia il numero di elettroni è minore all'aumentare di θ .

Riassumendo, possiamo dire che il numero di elettroni osservato da un esperimento è una funzione dell'energia del primario E_0 , della natura del primario (il suo numero atomico Z) e dell'angolo zenitale cui viene osservato lo sciame θ . Per avere una calibrazione di N_e in funzione di E_0 è necessario rimuovere le dipendenze da Z e da θ . La dipendenza dall'angolo zenitale viene trattata utilizzando una procedura di analisi dati (nota come *constant intensity cut*) basata sui dati sperimentali e che ha come unica ipotesi la non dipendenza dall'angolo azimutale del flusso dei raggi cosmici, ipotesi verificata a meno di un fattore $10^{-3} - 10^{-4}$. La dipendenza dalla massa del primario non può essere eliminata con i dati sperimentali (non possiamo determinare per ogni evento la Z del primario) e dobbiamo quindi ricorrere ad un'ipotesi basata su modelli, questa è una seconda fonte di errore sistematico nella calibrazione energetica di questi apparati.

Queste considerazioni sugli errori sistematici valgono per gli esperimenti che campionano uno sciame a un determinato punto del suo sviluppo, il discorso è molto diverso per gli apparati che misurano la luce di fluorescenza.

Breve descrizione di alcuni esperimenti

LHAASO

Per esemplificare le caratteristiche principali di questi apparati scelgo di descrivere in un certo dettaglio un esperimento appena completato in Cina, LHAASO, che racchiude al suo interno le caratteristiche principali sia di un rivelatore dedicato alla ricerca di sorgenti di fotoni di alta energia che allo studio della radiazione carica di origine galattica (cioè di energia inferiore a 10^{18} eV). Al termine elencherò brevemente gli altri principali apparati in presa dati al momento.

LHAASO (Large High Altitude Air Shower Observatory) il cui *layout* è mostrato in figura 5, è un esperimento che racchiude tutte le principali caratteristiche descritte finora, è infatti stato concepito per la ricerca di sorgenti di fotoni da un'energia di qualche centinaio di GeV fino a qualche PeV, almeno, quindi deve essere in grado di selezionare gli eventi generati da protoni primari scartando gli sciame adronici. D'altra parte, questi sciame adronici sono eventi con i quali si possono fare studi legati ai raggi cosmici carichi. LHAASO è un esperimento realizzato in Cina, nella provincia del Sichuan a una quota di 4410 m s.l.m., ed è composto da tre apparati: LHAASO-KM2A un rivelatore distribuito su 10^6 m² composto da scintillatori non schermati (per la misura della componente elettromagnetica dello sciame: elettroni e fotoni); LHAASO-WCDA un sistema di tre piscine per un totale di 78000 m² posto al centro dell'esperimento per la misura degli sciame generati da primari di bassa energia; LHAASO-WFCTA composto da 18 telescopi per la misura della luce chernkov e/o di quella di fluorescenza. KM2A è l'apparato di grande superficie realizzato per coprire le energie più alte sia per la misura dei fotoni primari che per quella dei raggi cosmici carichi, ed è composto da due tipi di rivelatori ED (Electromagnetic Detector) e MD (Muon Detector). Il rivelatore ED è distribuito uniformemente su una superficie di 1.3 km² ed è composto da una parte centrale che comprende 4911 rivelatori dislocati in un area circolare di raggio 575 m su una griglia triangolare di lato 15 m e da un anello esterno composto da 305 rivelatori posti in un anello centrale che

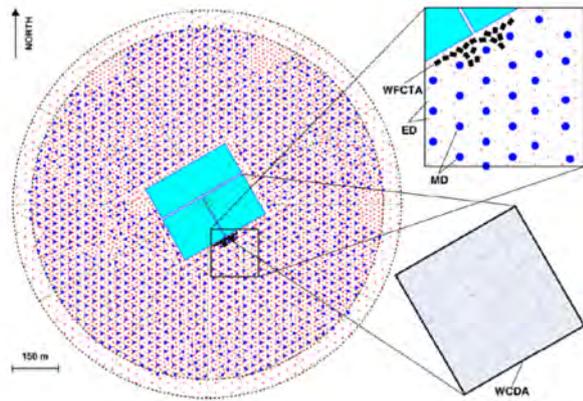


Figura 5: Layout dell'esperimento LHAASO

circonda quello centrale di raggio esterno di 625 m, in questa regione più esterna i rivelatori sono dislocati su una griglia triangolare di lato 30 m. Il singolo rivelatore di ED è uno scintillatore plastico realizzato da quattro elementi di $100 \times 25 \text{ cm}^2$, di spessore 1 cm; in ogni elemento sono poste 12 fibre ottiche che raccolgono la luce emessa nello scintillatore e la convogliano verso un fotomoltiplicatore. Per aumentare la sensibilità ai fotoni compresi negli EAS e per migliorare la risoluzione temporale del rivelatore ogni scintillatore di ED è coperto da uno strato di 5 mm di piombo. Per aumentare il *range* dinamico del rivelatore il segnale è estratto sia dall'anodo che dal sesto dinodo del fotomoltiplicatore.

Il rivelatore MD consiste di 1196 rivelatori disposti nella regione interna, di raggio 575 m, di ED e sono localizzati su una griglia triangolare di lato 30 m. Il singolo rivelatore di MD è una vasca, con pareti riflettenti, di acqua purificata al cui interno le particelle cariche producono luce Čerenkov che viene misurata da un fotomoltiplicatore da otto pollici posto al centro dello strato superiore del rivelatore. Le vasche di acqua sono coperte da uno strato di terra che può raggiungere lo spessore di 2.5 m per assorbire le particelle della componente elettromagnetica degli sciami permettendo il passaggio dei soli muoni.

Al centro dell'esperimento è costruito il rivelatore denominato WCDA che è composto da tre piscine per una superficie totale di 78000 m^2 , la profondità dell'acqua è di 4.4 m. L'acqua delle piscine è otticamente divisa da teli di plastica nera in celle elementari di $5 \times 5 \text{ m}^2$, al centro di ciascuna di queste sono posti due fotomoltiplicatori:

uno da 8 pollici e uno da 1.5 pollici per aumentare il *range* dinamico del rivelatore. Questo rivelatore è dedicato alle più basse energie cui è sensibile LHAASO e in particolare studia la radiazione cosmica tra 100 GeV e 30 TeV.

Per poter misurare il profilo longitudinale degli sciami (i.e. la luce di fluorescenza) e per rivelare la luce Čerenkov emessa negli sciami sono stati realizzati 18 telescopi a grande campo di vista (WFCTA). Ognuno di questi ha una superficie riflettente di 5 m^2 , la luce riflessa è poi raccolta da un sistema di Silicon Photomultipliers (SiPM). Ogni telescopio è posto in un container che può essere spostato per poter realizzare diverse geometrie di misura a seconda dei diversi scopi osservativi.

Altri esperimenti

Per concludere faccio una breve rassegna dei principali apparati per la misura degli EAS attualmente in presa dati, senza trattare gli esperimenti che lavorano a energie superiori a 10^{18} eV che sono discussi in un altro lavoro di questa pubblicazione.

Partendo dalle basse energie voglio ricordare i cosiddetti IACT (Imaging Atmospheric Čerenkov Telescopes) che non ho discusso in questo lavoro ma che hanno portato importanti risultati nella ricerca di sorgenti di fotoni per energie tra circa 100 GeV e 10 TeV. Attualmente sono in funzione tre esperimenti, il primo HESS [10] è localizzato in Namibia, nell'emisfero sud (con la possibilità quindi di osservare il centro della nostra galassia), composto da quattro telescopi di 12 m di diametro ciascuno al cui centro è poi stato recentemente aggiunto un quinto telescopio di 2.8 m di diametro. Gli altri due esperimenti sono costruiti nell'emisfero Nord: MAGIC [11] alle isole Canarie, due telescopi di 17 m di diametro e VERITAS [12] negli USA composto da 4 telescopi di 12 m di diametro.

Nell'emisfero Nord è attualmente in funzione, in Messico a una quota di 4100 m *s.l.m.*, l'esperimento HAWC [3] composto da 300 rivelatori di luce Čerenkov in acqua. Ciascun elemento è una tank di acqua purificata di diametro 7.3 m e di altezza 4m al cui interno sono posti quattro fotomoltiplicatori. Il tutto copre un'area di circa

22000 m² e i rivelatori hanno un *fill factor* del 57% circa.

Nel sito sperimentale della valle di TUNKA (Siberia) è stato realizzato un esperimento TAIGA [13] in cui operano sia rivelatori di luce Čerenkov dedicati alla misura dei primari carichi che IACT dedicati alla ricerca di sorgenti di fotoni di altissima energia.

Per i raggi cosmici carichi di origine galattica tra gli esperimenti attualmente in presa dati è importante menzionare ICE-TOP [14] che è stato realizzato al polo Sud al di sopra del rivelatore di neutrini astrofisici ICE-CUBE. ICE-TOP è composto da 80 elementi, ciascuno dei quali è realizzato con due rivelatori di luce Čerenkov nel ghiaccio, posti a una distanza di 125 m l'uno dall'altro. L'esperimento copre una superficie complessiva di circa 1 km².

Sviluppi futuri e conclusioni

Con gli esperimenti per la misura dei raggi cosmici da terra che ho brevemente descritto in questo lavoro si è aumentata la nostra conoscenza sulla radiazione cosmica di origine galattica.

Per la ricerca delle sorgenti di fotoni di energia superiore a circa 100 GeV sono state sviluppate due tecniche di misura tra loro complementari: gli IACT sono strumenti in grado di selezionare con grande accuratezza gli eventi generati da fotoni primari rispetto al fondo di eventi dovuti a primari adronici e con questi telescopi si ottengono misure con un'alta precisione angolare e una buona risoluzione energetica (circa 15-20%). D'altra parte questi osservatori hanno un piccolo campo di vista e un basso *duty cycle*, quindi sono molto indicati per osservare oggetti celesti già noti e studiarne nel dettaglio le caratteristiche. Con questi strumenti sono state osservate molte sorgenti sia galattiche (principalmente *Supernovae remnant* e *pulsar*) che extragalattiche (essenzialmente AGN e tra questi *blazar*) e sono recentemente stati osservati un paio di eventi sporadici, in particolare dei *gamma ray burst* che sono stati segnalati da esperimenti operanti su satelliti.

A fianco agli IACT anche con gli apparati a sciami si sono ottenuti dei notevoli risultati soprattutto per quanto riguarda l'estensione della misura delle sorgenti di fotoni oltre i 100 TeV e fino al PeV. Questi risultati sono stati prima otte-

nuti dagli esperimenti HAWC e TIBET-AS γ e in seguito dall'esperimento cinese LHAASO. In particolare, quest'ultimo ha localizzato 13 sorgenti oltre i 100 TeV, tutte si trovano sul piano galattico (anche se non si è ancora osservato il centro della galassia). Questi risultati hanno aperto la strada all'astronomia ad energie superiori ai 100 TeV e ci aspettiamo grandi sviluppi in questo campo. A questo proposito segnalo il progetto SWGO [15] che si propone di realizzare un osservatorio a grande campo di vista, basato su rivelatori Čerenkov ad acqua, da costruire nell'emisfero sud a una quota superiore a 4500 m s.l.m. che possa osservare queste sorgenti tra 100 GeV e i PeV. Questi apparati a grande campo di vista, rispetto agli IACT, hanno *performance* meno buone (sia sulla selezione gamma/adroni che sulla precisione angolare ed energetica) ma grazie al loro grande *duty cycle* e campo di vista sono molto più indicati per la ricerca di sorgenti inattese e per la rivelazione di eventi sporadici come, per esempio, i *gamma ray burst*. Inoltre, questi esperimenti hanno una sensibilità anche migliore di quella degli IACT alle altissime energie ($E > 100$ TeV).

Per quanto riguarda la misura dei raggi cosmici carichi di origine galattica, gli studi degli ultimi anni si sono concentrati sulla regione attorno al cambio di pendenza dello spettro primario dei raggi cosmici, il cosiddetto ginocchio osservato a circa $3 - 4 \times 10^{15}$ eV, la cui origine è stata associata agli elementi più leggeri (H, He). Il cambio di pendenza della componente pesante (Fe) della radiazione cosmica è stato osservato a un'energia superiore (circa 10^{17} eV) da parte dell'esperimento KASCADAE [5]. Bisogna però sottolineare come questi risultati sono stati ottenuti con esperimenti operanti molto al sotto del punto di massimo sviluppo degli sciami, mentre con i dati raccolti ad alta quota dall'esperimento ARGO-YBJ si è osservato un cambio di pendenza simile per la componente leggera dei primari a energia inferiore (circa 400 TeV).

In questo senso è fondamentale che gli apparati di prossima generazione (a partire da LHASSO) siano in grado di separare, con una selezione applicata ai singoli eventi, almeno due, ma possibilmente più, gruppi di massa. In questo modo si potranno misurare gli spettri della componente leggera e pesante dei raggi cosmici indagando la

questione aperta con i recenti dati citati in precedenza. Naturalmente se si riuscirà a separare anche un terzo (o magari anche un quarto) gruppo di massa la nostra conoscenza della radiazione cosmica sarebbe notevolmente migliorata. Se si riuscisse a separare questi gruppi di massa su una statistica molto alta di eventi sarebbe anche molto interessante investigare l'anisotropia della direzione di arrivo della radiazione cosmica per diversi gruppi di massa (risultato cui non si è ancora pervenuti).

Indico qui sotto alcuni testi per un approfondimento degli argomenti trattati.

- M. Spurio: *Particles and Astrophysics*, Springer, Berlino (2015).
- A. De Angelis, M. Pimenta: *Introduction to Particle and Astroparticle Physics*, Springer (Berlino) (2015).
- A. Matthews: *A Heitler model of extensive air showers*, *Astroparticle Physics*, 22 (2005) 387.
- J. A. Hinton, W. Hofmann: *Teraelectronvolt Astronomy*, arXiv:1006.5210v2, (2010)
- A. Haungs, H. Rebel, M. Roth: *Energy spectrum and mass composition of high-energy cosmic rays*, *Rep. Prog. Phys.*, 66 (2003) 1145.



- [1] R. Atkins et al.: *Milagro, a TeV air shower array*, *Nucl. Inst. Meth. A*, 449 (2000) 478.
- [2] LHAASO collaboration: *The Large High Altitude Air Shower Observatory (LHAASO) Science Book (2021 Edition)*, *Chinese Physics C*, 46 (2022) 035001.
- [3] A.U. Abeysekara et al.: *Data Acquisition Architecture and Online Processing System for the HAWC gamma-ray observatory*, *Nucl. Inst. Meth. A*, 888 (2018) 138.
- [4] B. Bartoli et al.: *Calibration of the RPC charge readout in the ARGO-YBJ experiment with the iso-gradient method*, *Nucl. Inst. Meth. A*, 783 (2015) 68.
- [5] T. Antoni et al.: *The Cosmic ray experiment KASCADE*, *Nucl. Inst. Meth. A*, 513 (2003) 490.
- [6] D. Apel et al.: *The KASCADE-Grande experiment*, *Nucl. Inst. Meth. A*, 620 (2010) 202.
- [7] R.D. Monkhoev et al.: *The Tunka-Grande experiment*, *Jour. Inst.*, 12 (2017) C06019.

- [8] The Pierre Auger Collaboration: *The Pierre Auger Cosmic Ray Observatory*, *Nucl. Inst. Meth. A*, 798 (2015) 172.
- [9] T. Abu-Zayyad et al.: *The surface detector array of the Telescope Array experiment*, *Nucl. Inst. Meth. A*, 689 (2012) 87.
- [10] J.A. Hinton: *Status of the HESS project*, *New. Astron. Rev.*, 48 (2004) 331.
- [11] The MAGIC Collaboration: *The MAGIC telescope*, *Nucl. Inst. Meth. B*, 114 (2003) 247.
- [12] J. Holder et al.: *VERITAS: Status and Highlights*, arXiv:1111.1225 (2011).
- [13] N. M. Budnev et al.: *TAIGA - A hybrid array for high-energy gamma astronomy and cosmic-ray physics*, *Nucl. Inst. Meth. A*, 936 (2019) 254.
- [14] R. Abbasi et al.: *IceTop: the surface component of IceCube*, *Nucl. Inst. Meth. A*, 700 (2013) 188.
- [15] www.swgo.org



Andrea Chiavassa: è Professore Ordinario presso l'Università di Torino. Si interessa da anni allo studio dei raggi cosmici partecipando ad esperimenti condotti con apparati operanti a terra (e quindi nella misura degli EAS), in particolare ha contribuito agli esperimenti EAS-TOP e KASCADE-Grande.

Osservazione dei raggi cosmici dallo spazio

Mirko Boezio

Istituto Nazionale di Fisica Nucleare, Sez. di Trieste, I-34149 Trieste, Italy

Riccardo Munini

Institute for Fundamental Physics of the Universe, I-34014 Trieste, Italy

Istituto Nazionale di Fisica Nucleare, Sez. di Trieste, I-34149 Trieste, Italy

Institute for Fundamental Physics of the Universe, I-34014 Trieste, Italy

I raggi cosmici sono particelle energetiche con origine sia nel nostro sistema solare e galassia ma anche fuori della galassia. I raggi cosmici sono in gran parte protoni e nuclei di elio ma sono presenti tutti i nuclei della tavola periodica oltre ad elettroni ed antiparticelle. La loro origine, i processi di accelerazione ad alte energie come pure la propagazione verso la terra sono argomenti che stanno impegnando gli scienziati dalla loro scoperta da parte di Victor Hess nel 1912. Per studiarli adeguatamente è necessario misurarne con precisione lo spettro energetico e la composizione. Queste ricerche sono state, ed è, affrontate sperimentalmente sia con strumentazione a terra che con apparati al sommo dell'atmosfera installati su palloni stratosferici o in satelliti o stazioni spaziali. Quest'ultimo campo ha visto una significativa attività negli ultimi decenni con molti apparati equipaggiati con strumentazione all'avanguardia che hanno prodotto importanti risultati rivoluzionando la comprensione della fisica dei raggi cosmici.

Cosa sono i raggi cosmici

La scoperta dei raggi cosmici avvenne nel 1912 grazie ad un volo in mongolfiera di Victor Hess, successivamente premiato con il premio Nobel nel 1936, durante il quale egli misurò il livello di ionizzazione "della radiazione ionizzante" a diverse altitudini. Hess notò che questo livello di ionizzazione, dopo aver raggiunto un minimo poco sopra il livello del terreno, aumentava considerevolmente man mano che la mongolfiera si allontanava dal terreno, raggiungendo, ad una altezza di 5200 metri, un valore del 100% maggiore rispetto a quello osservato a terra.

I meticolosi esperimenti di Hess erano stati preceduti da quelli condotti sulla cima della torre Eiffel da Theodor Wulf, a bordo di palloni aerostatici da Albert Gockel e sott'acqua dall'italiano Domenico Pacini. Tutti questi dati portavano alla conclusione che una parte non trascurabile della radiazione misurata non era correlata con la naturale radioattività proveniente dalle sostanze radioattive presenti negli strati superficiali della crosta terrestre ma doveva venire dallo spazio.

Oggi sappiamo che i raggi cosmici sono particelle energetiche che includono i nuclei di tutti gli elementi, compresi i loro isotopi, oltre ad elettroni, positroni e antiprotoni. Queste particelle si propagano attraverso la galassia e sono accelerate dai sistemi più energetici dell'universo come esplosione di supernove, pulsars, getti

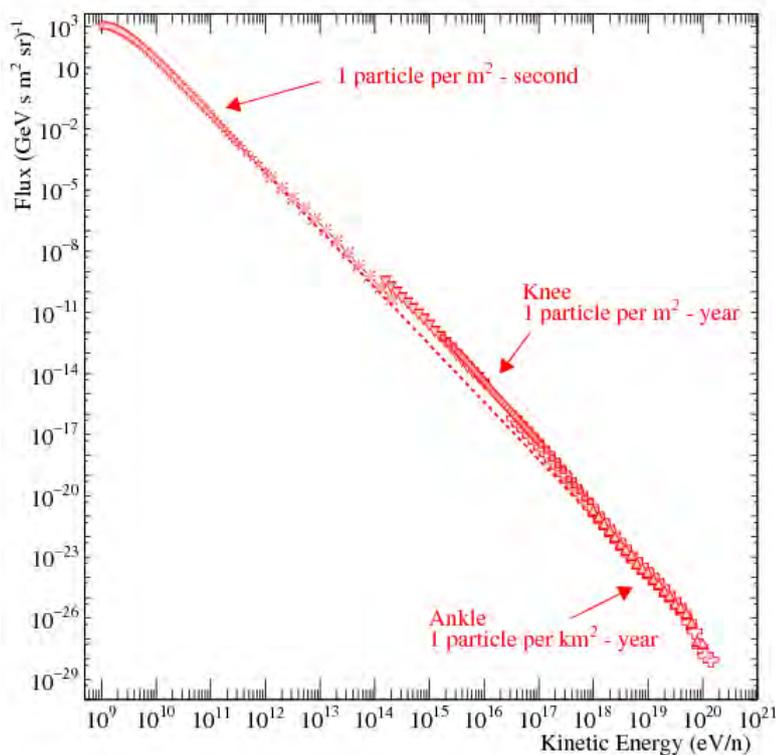


Figura 1: Lo spettro in energia dei raggi cosmici: fino a circa 10^{14} eV si tratta di flussi di protoni da misure dirette. Ad energie superiori viene mostrato lo spettro di tutte le componenti misurato da osservatori a terra. La linea tratteggiata rappresenta una legge di potenza di indice spettrale -2.7 .

relativistici e nuclei galattici attivi.

Per una discussione più approfondita sulla storia dello studio dei raggi cosmici si veda l'articolo di Alessandro De Angelis in questo numero di **Ithaca** e la referenza [1].

Storicamente, la scoperta dei raggi cosmici gettò le basi per la nascita della fisica delle particelle che si sviluppò per molti anni studiando i raggi cosmici nell'atmosfera terrestre. I positroni, i muoni, i pioni e le cosiddette particelle strane furono rilevate per la prima volta grazie ai raggi cosmici ed ai prodotti delle loro interazioni. Ci vollero quaranta anni di innovazione tecnologica prima di riprodurre parzialmente con gli acceleratori in laboratorio le energie dei raggi cosmici.

Lo studio della radiazione cosmica dallo spazio o nell'alta atmosfera è iniziata negli anni sessanta con la serie di pionieristici satelliti Proton di Grigorov e con la rivelazione di elettroni di Earl, Meyer e Vogt tramite apparati su palloni stratosferici. Gli anni ottanta e novanta vedono una forte accelerazione dello studio della radiazione cosmica con campagne di voli in pallone o piccoli satelliti. Questo contribuì a completare la transizione dalle iniziali osservazioni dei rag-

gi cosmici come laboratorio per la produzione e la rivelazione di nuove particelle allo studio dei meccanismi di produzione, accelerazione e propagazione dei raggi cosmici nella galassia e del loro ruolo nel cosmo.

Nella figura 1 sono mostrati i flussi dei raggi cosmici in funzione della loro energia; il cosiddetto spettro energetico. Esso rappresenta uno dei più impressionanti risultati della ricerca sperimentale astroparticellare, estendendosi per più di 32 ordini di grandezza sulla scala dell'intensità del flusso e per oltre 10 ordini di grandezza sulla scala delle energie. In questo spettro energetico sono nascoste le risposte alle principali domande della ricerca sui raggi cosmici. Da dove vengono queste particelle? Come sono accelerate? Come si propagano attraverso il mezzo interstellare e quali sono le loro interazioni? A che energia il contributo extra galattico diventa dominante? I raggi cosmici possono darci indizi sull'esistenza di particelle esotiche, residui di un universo primordiale, come l'antimateria o la materia oscura? Nel corso dei prossimi paragrafi si discuterà come si sono e si stanno affrontando sperimentalmente questi quesiti.

Equazione di diffusione

La diffusione nello spazio e nel tempo di una specie i di raggi cosmici viene descritta considerando le variazioni della sua densità $\psi_i \equiv \psi_i(\mathbf{x}, p, t)$ definita al tempo t nella posizione \mathbf{x} e con modulo dell'impulso p . L'equazione che descrive queste variazioni è

$$\frac{\partial \psi_i}{\partial t} = q_i(\mathbf{x}, p, t) + \nabla \cdot (D \nabla \psi_i - \mathbf{u} \psi_i) + \frac{\partial}{\partial p} p^2 D_p \frac{\partial}{\partial p} \frac{1}{p^2} \psi_i - \left[\frac{\partial}{\partial p} \dot{p} \psi_i - \frac{p}{3} (\nabla \cdot \mathbf{u}) \psi_i \right] - \frac{\psi_i}{(\tau_f)_i} - \frac{\psi_i}{(\tau_r)_i},$$

dove $q(\mathbf{x}, p, t)$ indica il contributo delle sorgenti sia della componente primaria che secondaria da interazione o decadimento, D è il coefficiente di diffusione spaziale, \mathbf{u} la velocità di convezione, D_p è il coefficiente di diffusione nello spazio degli impulsi, $p = |\mathbf{p}|$, $\dot{p} = dp/dt$ è la frequenza dei guadagni o perdite di impulso, τ_f rappresenta la scala dei tempi per la frammentazione e τ_r per il decadimento radioattivo.

Il paradigma dei raggi cosmici

Il paradigma dei raggi cosmici afferma che i raggi cosmici fino alle energie del ginocchio, $10^{15} - 10^{17}$ eV ("Knee" in figura 1), sono primariamente accelerati nei resti di supernova (SNR) che si trovano nel disco galattico. Le SNR hanno l'energetica giusta ed attraverso meccanismi di accelerazione diffusiva (*diffusive shock acceleration*) accelerano e iniettano particelle nel mezzo interstellare con uno spettro energetico che segue approssimativamente una legge di potenza, come si vede in figura 1, e simile per tutte le specie di particelle con indice spettrale di circa -2.7 (linea tratteggiata in figura).

Dopo aver lasciato i siti di accelerazione, queste particelle, diventate ora raggi cosmici, si diffondono nella galassia interagendo col mezzo interstellare e coi campi magnetici galattici. La propagazione di ogni componente è solitamente descritta da un'equazione di diffusione che ne descrive il trasporto nel campo magnetico galattico. Da questa diffusione ne consegue una isotropizzazione della distribuzione spaziale degli stessi. Per raggi cosmici primari come protoni e nuclei di elio i guadagni e le perdite di energia si possono trascurare ad alta energia e l'equazione di diffusione in condizione stazionaria si semplifica:

$$0 = q_i(\mathbf{x}, p, t) + \nabla \cdot (D \nabla \psi_i), \quad (1)$$

assumendo uno spettro di iniezione dato da $q \propto p^{-\gamma}$, la soluzione diventa

$$\psi_i \sim p^{-\gamma-\delta}.$$

Ovvero, secondo il paradigma dei raggi cosmici il processo di propagazione produce un indebolimento degli spettri energetici osservati nella galassia. Questo spiega come mai lo spettro dei raggi cosmici di figura 1 è più ripido di quanto atteso per la sola accelerazione (per una discussione più approfondita si può vedere [2]).

Nello spettro di figura 1 c'è un ulteriore effetto particolarmente significativo a bassa energia (sotto alcune decine di GeV) legato al fatto che la quasi totalità delle osservazioni sono state condotte nella vicinanza della terra. I raggi cosmici galattici per raggiungere la Terra devono propagarsi nell'eliosfera dove interagiscono col campo magnetico congelato nel vento solare. Da questa interazione risulta la così detta modulazione solare dei raggi cosmici che ha la stessa ciclicità undecennale dell'attività solare.

Il vento solare è un flusso continuo di gas magnetizzato emesso radialmente dal Sole a velocità comprese tra i 300 e gli 800 km/s. Attraverso un meccanismo noto come congelamento del flusso, il vento solare trasporta il campo magnetico del Sole nel sistema solare. A distanze dal Sole tra 80 e 90 AU (1 AU: distanza tra terra e sole) la pressione del vento solare eguaglia la pressione del mezzo interstellare ed il vento solare rallenta

a velocità subsoniche formando un fronte d'urto chiamato *termination shock*. La regione contenuta all'interno del *termination shock* si chiama eliosfera e definisce la regione di spazio dominata dal vento solare.

Da secoli è noto che l'attività solare ha una ciclicità di periodo 11 anni in cui passa da una fase di minima attività ad una di massima. In aggiunta a questo effetto, ogni 11 anni la polarità del campo magnetico solare cambia introducendo una seconda ciclicità di 22 anni. Il campo magnetico solare come pure la velocità del vento solare sono significativamente minori durante un periodo di minima attività solare, influenzando conseguentemente di meno i raggi cosmici. Inoltre cicli a polarità opposta impattano i raggi cosmici in maniera diversa a seconda del segno della carica dei raggi cosmici.

Da queste brevi considerazioni si può capire che le misure precise degli spettri in energia e della loro variabilità nel tempo e della composizione della radiazione cosmica forniscono informazioni fondamentali per spiegare la vita dei raggi cosmici.

Misure di raggi cosmici

Come si può intuire dall'andamento dello spettro in energia dei raggi cosmici mostrato in figura 1 un singolo approccio sperimentale non è una soluzione praticabile. Quindi nel corso degli anni sono stati adottati molti approcci diversi per misurare la radiazione cosmica. La distinzione più significativa è tra misure dirette ed indirette. La presenza dell'atmosfera terrestre impone che le misure dirette siano condotte al sommo dell'atmosfera o tramite pallone stratosferici, con residui di frazioni del percento di atmosfera, o nello spazio. La realizzazione di apparati per tali misure è, però, temperata sia dai costi di messa in opera, particolarmente importanti soprattutto per missioni spaziali, sia dall'andamento a legge di potenza negativa dello spettro dei raggi cosmici che comporta che più alte siano l'energia delle particelle da misurare maggiori diventino l'accettanza sperimentale ed il tempo di acquisizione richiesti. Per esempio nella regione del ginocchio la frequenza attesa di particelle è di 1 particella per m² per anno. Appare ovvio che a queste ed ad energie più elevate misure dirette

non sono praticabili. Nel corso degli anni varie tecniche sono state utilizzate per misurare la radiazione cosmica alle più alte energie, tutte basate sulla misura dei prodotti dell'interazione dei raggi cosmici con l'atmosfera terrestre, da cui la definizione di misure indirette. La quasi totalità di queste misure sono state condotte da terra. Comunque, recentemente si sono cominciate a sperimentare tecniche di osservazione dal sommo dell'atmosfera tramite palloni stratosferici e dallo spazio.

Nel seguito vengono brevemente presentati gli esperimenti spaziali più recenti nel campo delle misure dirette dei raggi cosmici assieme ai risultati più rilevanti che sono stati ottenuti.

Esperimenti dallo spazio

Come discusso in precedenza per comprendere l'origine dei raggi cosmici è necessario misurarne con molto dettaglio lo spettro in energia e la composizione per poter capire se ci sono strutture negli andamenti spettrali e se queste dipendono dal tipo di particelle. Quindi gli esperimenti devono essere equipaggiati con apparati in grado di identificare le particelle e misurarne l'energia. Nel corso dei decenni molti approcci diversi sono stati utilizzati per queste misure di rivelazione diretta.

Per quanto riguarda le misure di energie le due tecniche principalmente usate sono stati spettrometri magnetici con cui ricavare la rigidità¹ e, quindi, l'impulso e l'energia delle particelle e calorimetri con cui misurare direttamente l'energia delle particelle. Dai primi spettrometri magnetici equipaggiati con magneti superconduttori e lanciati con voli in pallone negli anni '70 del secolo scorso, si è passati in anni più recenti a magneti permanenti installati su satelliti come l'apparato PAMELA [3] o, sulla stazione spaziale ISS, come l'apparato AMS-02 [4]. La figura 2 mostra gli apparati PAMELA ed AMS-02. PAMELA, lanciato il 15 giugno del 2006 dalla base di Baikonur in Kazakistan, ha preso dati a circa 600 km di altitudine fino a gennaio 2016. L'apparato PAMELA (figura 2 sinistra), presenta, attorno al magnete permanente per la misura della rigidità

¹La rigidità R è definita come $R = pc/(Ze)$, dove c è la velocità della luce, p l'impulso e Ze la carica della particella.

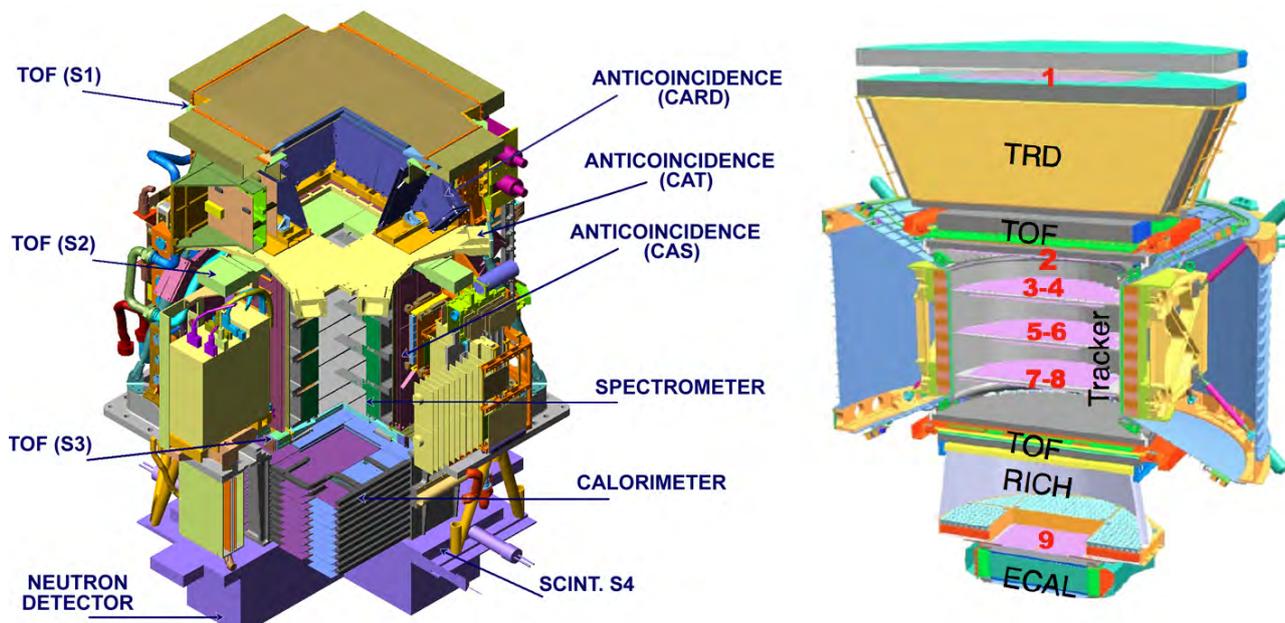


Figura 2: A sinistra: l'apparato PAMELA. A destra: l'apparato AMS-02.

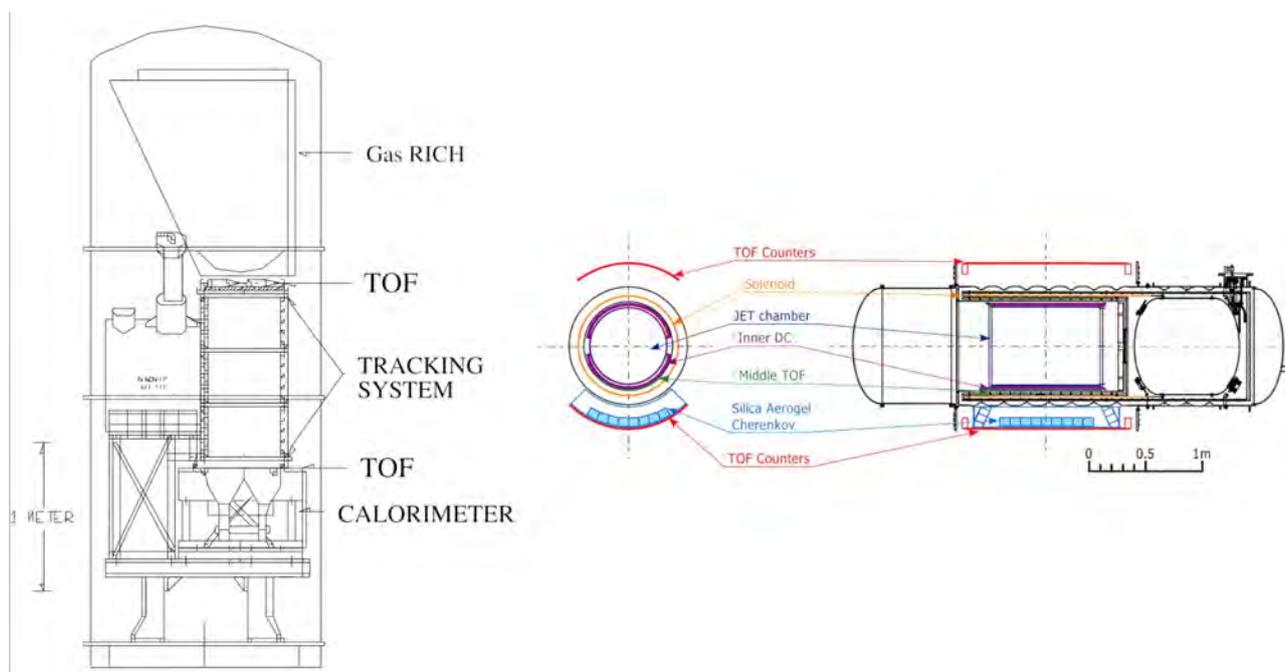


Figura 3: Sinistra: l'apparato CAPRICE98. Destra: l'apparato BESS-Polar.

delle particelle, un sistema di tempo di volo per il *trigger*, misura di velocità e perdite di energia per ionizzazione, un calorimetro elettromagnetico ed un rivelatore a neutroni per la separazione tra particelle elettromagnetiche e particelle adroniche.

L'apparato AMS-02 (figura 2 destra), installato sulla ISS il 19 maggio 2011, consiste, oltre che ad un magnete permanente per la misura della rigidità delle particelle, di un sistema di tempo

di volo per il trigger, misura di velocità e perdite di energia per ionizzazione, di un calorimetro elettromagnetico per la separazione tra particelle elettromagnetiche e particelle adroniche e di un rivelatore di radiazione di transizione ed un *Ring Imaging Čerenkov* (RICH) per ulteriore potere di separazione tra particelle veloci, come elettroni e positroni, e lente, come protoni e nuclei più pesanti.

Questo approccio segue la strada indicata dai precedenti esperimenti su palloni stratosferici

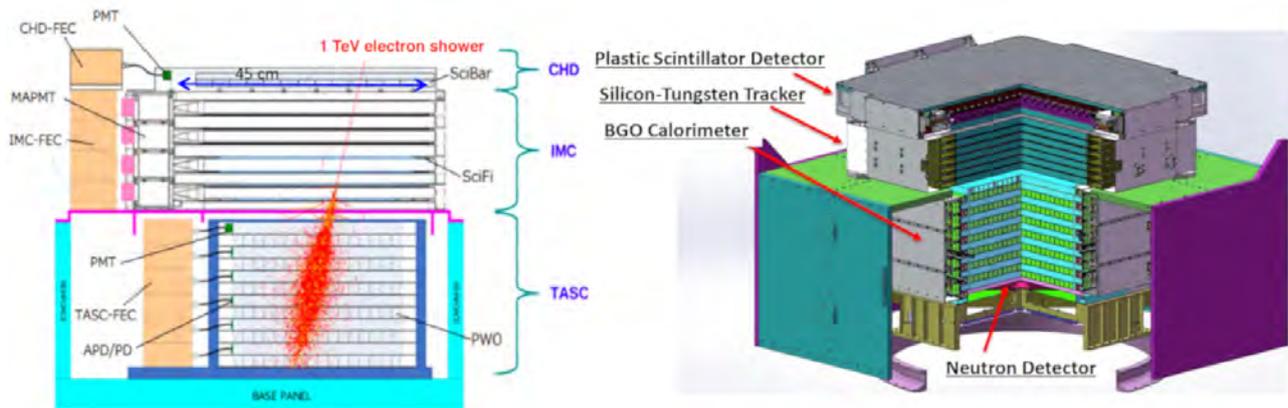


Figura 4: Sinistra: l'apparato CALET. Destra: l'apparato DAMPE.

come l'apparato equipaggiato con un magnete superconduttore con cui il compianto Robert Golden misurò nel 1979 per la prima volta antiprotoni nei raggi cosmici galattici [5]. O come gli apparati che negli anni 1990 e 2000 hanno affiancato a magneti superconduttori avanzati rivelatori per l'identificazione delle particelle sviluppati per esperimenti ad acceleratori. La figura 3 mostra due di questi apparati: l'esperimento CAPRICE98 [6], a sinistra, dotato di calorimetro elettromagnetico simile a quello di PAMELA e di un RICH, l'esperimento BESS-Polar [7], a destra, dotato di rivelatore Čerenkov e sistema di tempo di volo ad alta risoluzione.

Accanto all'utilizzo di magneti, permanenti o superconduttori, l'altra tecnica ampiamente usata per misurare l'energia dei raggi cosmici è stata quella di utilizzare dei calorimetri. Questo è evidenziato in figura 4 che mostra due recenti missioni spaziali: CALET, a sinistra, e DAMPE a destra. CALET [8] è stato lanciato il 19 agosto 2015 e poi installato a bordo della stazione spaziale internazionale ISS da dove sta raccogliendo ininterrottamente dati. Oltre ad un calorimetro omogeneo di tustanato di piombo per misurare l'energia delle particelle e separare particelle elettromagnetiche da adroniche, l'apparato include degli scintillatori plastici per la misurazione della carica delle particelle e un calorimetro ad immagine che funziona sia come tracciante che come *preshower*. DAMPE [9] è stato installato a bordo di un satellite cinese e lanciato il 17 dicembre 2015. In maniera analoga a CALET, è dotato di un calorimetro di barre di BGO per la misura di energia e per l'identificazione delle particelle, di un misuratore di carica, di un tracciante e

di un rivelatore di neutroni per fornire ulteriore capacità di identificazione. Il vantaggio dell'approccio calorimetro è la compattezza degli apparati che permette di ottenere accettanze più grandi a parità di peso e dimensioni, ovvero di costi per missioni spaziali. Un immediato svantaggio è l'assenza di strumentazione in grado di misurare il segno della carica e, quindi, di studiare, per esempio, la componente antiparticellare della radiazione cosmica che ha un grande interesse scientifico per le sue ricadute nella ricerca di nuova fisica, come materia oscura.

La componente primaria

I raggi cosmici sono normalmente definiti essere primari, cioè accelerati alle sorgenti (di qualunque tipo queste possano essere), o secondari, cioè prodotti dall'interazione dei raggi cosmici col mezzo interstellare. Protoni, nuclei di elio, carbonio ed ossigeno sono un esempio di componente primaria mentre i nuclei di litio, berillio e boro sono tra le più abbondanti componenti secondarie. Vale la pena rimarcare che, con l'eccezione di una piccola componente di raggi cosmici di bassa energia (da poche decine a qualche centinaio di MeV) originati nell'eliosfera, i raggi cosmici misurati nelle vicinanze della terra sono nuclei in quanto gli elettroni atomici sono stati strappati durante la fase di accelerazione.

Protoni e nuclei di elio sono la componente più abbondante dei raggi cosmici galattici rappresentando oltre il 95% del flusso totale di raggi cosmici. Non deve quindi sorprendere che sono anche le componenti più studiate e meglio misurate. La figura 5 mostra un insieme di misure

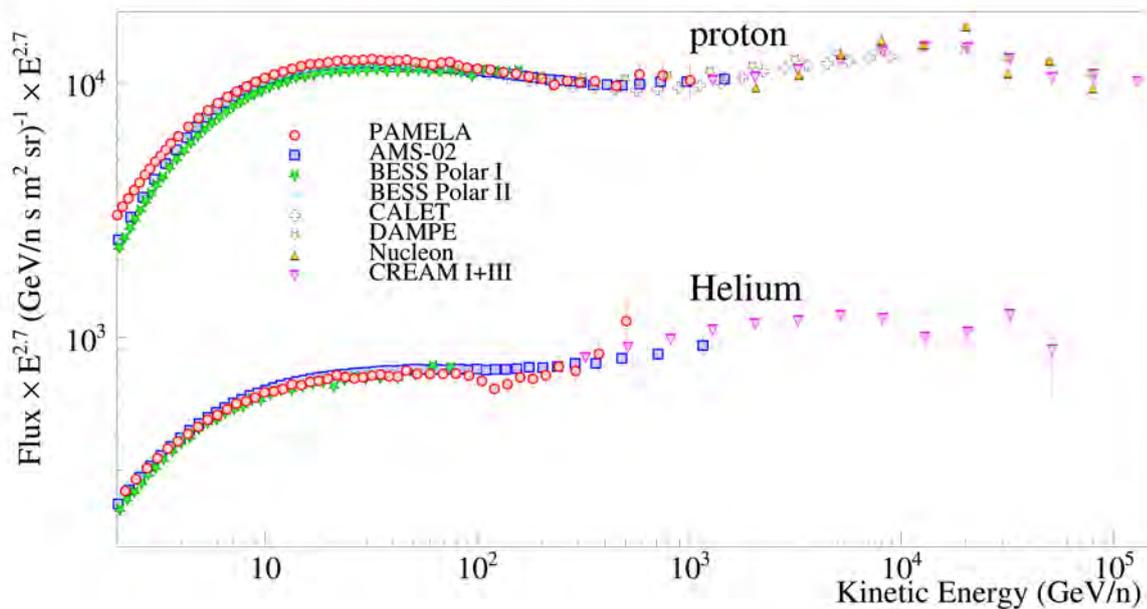


Figura 5: Misure dirette recenti dei flussi di protoni e nuclei di elio. I flussi sono moltiplicati per $E^{2.7}$, con E energia cinetica in GeV per nucleone.

recenti dei flussi di protoni e di nuclei² di elio ottenute con esperimenti su palloni stratosferici o nello spazio. In figura i flussi sono moltiplicati per una potenza dell'energia ($E^{2.7}$), questo viene comunemente fatto per visualizzare meglio le forme di spettri che seguono delle leggi di potenza.

La prima cosa che si può notare è l'eccellente accordo tra le varie misure tenendo conto che, come discusso precedentemente, le più basse energie (< 30 GeV) sono affette dalla modulazione solare che era diversa per esperimenti che si sono svolti in periodi diversi. Poi si può notare che sopra alla regione di influenza della modulazione solare, accanto all'andamento approssimativamente a legge di potenza previsto dallo scenario standard dei raggi cosmici, l'eccellente precisione di queste recenti misure indica chiaramente un inatteso indurimento dello spettro³ ad alcune centinaia di GeV. Questo effetto fu chiaramente evidenziato per la prima volta dall'esperimento PAMELA e successivamente confermato, con un significativo miglioramento della precisione statistica, dall'esperimento AMS-02, poi ulteriormente verificato, con sistematiche diverse, dagli esperimenti DAMPE e CALET.

²I nuclei di elio sono per oltre l'80% ^4He .

³In inglese si usa il termine *hardening*, ovvero l'indice spettrale γ della legge di potenza $E^{-\gamma}$ diminuisce.

Molte idee sono state presentate per spiegare quest'indurimento dello spettro, idee che vanno dal contributo di altre sorgenti astrofisiche oltre alle SNR, tipo stelle nova, che ad effetti dovuti alla propagazione, spiegazione che sta diventando la più accettata (per una rassegna di queste discussioni si veda [10]). Si può anche notare che ci sono indicazioni di un ulteriore cambio di pendenza dello spettro a $\approx 10 - 20$ TeV che, però, necessita di maggiori conferme sperimentali. Come predetto dal modello standard dei raggi cosmici i cambi spettrali osservati nei protoni si dovrebbero osservare anche nelle altre componenti nucleari, ed effettivamente questo è quanto si osserva nei nuclei di elio, figura 5. Però si può anche notare che protoni e nuclei di elio hanno forme spettrali diverse (chiaramente evidenti facendo il rapporto dei due flussi). Al momento non c'è ancora una spiegazione condivisa per questa differenza. Sono state fatte varie ipotesi (per una rassegna di queste si veda [11]): effetti dovuti ai processi di interazione tra raggi cosmici di specie diverse ed il mezzo interstellare; sorgenti diverse; e, forse la spiegazione più intrigante, differenti spettri di iniezione alla sorgente causati dai diversi rapporti di carica su massa.

La figura 6 mostra una raccolta di spettri in rigidità di nuclei più pesanti dell'elio. Come si può notare soprattutto nelle recenti misure di

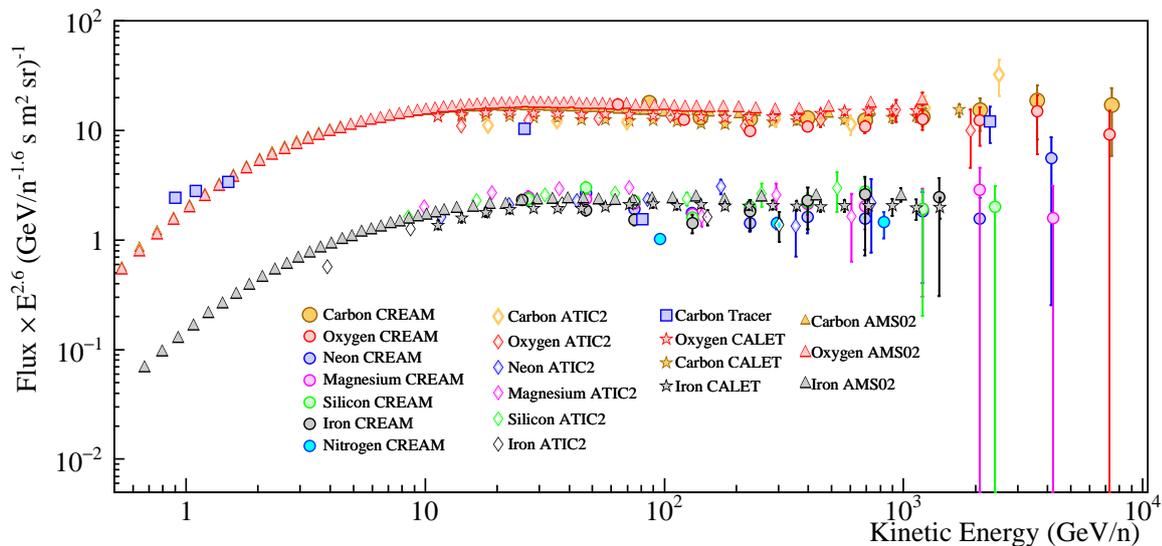


Figura 6: Misure dirette recenti degli spettri in energia di nuclei primari più pesanti dell'elio. I flussi sono moltiplicati per $E^{2.6}$, con E energia cinetica in GeV/n .

AMS-02 e CALET l'indurimento nello spettro è chiaramente visibile.

Molto interessanti sono anche state le misure di nuclei pesanti ad energie tra alcune decine di MeV/n ad alcune centinaia di MeV/n o qualche GeV/n a seconda del nucleo e dell'esperimento. Infatti, grazie soprattutto alle misure degli esperimenti su pallone TIGER e SUPER-TIGER ed all'esperimento CRIS della missione spaziale ACE, si è osservata un'abbondanza maggiore di metalli con alta temperatura di condensazione, rispetto a quella di elementi chimici con bassa temperatura di condensazione come idrogeno, non metalli e gas nobili. Questa differenza è stata spiegata col fatto che la gran parte dell'accelerazione di raggi cosmici avvenga nelle cosiddette associazioni di stelle del tipo OB supermassive e che il materiale accelerato sia preferibilmente polvere interstellare. Un altro interessante risultato ottenuto dall'esperimento ACE-CRIS è stata la scoperta che la maggior parte dei raggi cosmici origina da materia che è rimasta a lungo nel mezzo interstellare prima di essere accelerata dal fronte d'urto di esplosioni di supernova. Questo risultato è stato ottenuto misurando l'abbondanza di isotopi di nickel e cobalto. Dato che uno di questi isotopi, il ^{59}Ni , è radioattivo e decade soltanto per cattura elettronica con tempo di dimezzamento di 0.76×10^5 s, e visto che il processo di accelerazione strappa tutti gli elettroni dagli atomi, quindi impedendo il processo di de-

cadimento, la misurata assenza di questo isotopo nella radiazione cosmica implica che il processo di accelerazione è avvenuto ben oltre 10^5 anni dopo la nucleo-sintesi da parte della stella degli isotopi di nickel.

La componente secondaria

Come si diceva elementi secondari come, per esempio, i nuclei di boro sono prodotti dall'interazione dei raggi cosmici col mezzo interstellare. Conseguentemente, misurarne lo spettro e la composizione fornisce importanti informazioni sulla propagazione dei raggi cosmici nella galassia e sui processi fisici che li condizionano, come pure sul mezzo interstellare.⁴

Per questi studi è particolarmente interessante osservare il rapporto dei flussi tra secondari e primari in quanto è strettamente legato ai processi di diffusione ed alla quantità di materiale attraversato dai raggi cosmici. Dei vari rapporti che si possono studiare, quello che ha avuto la maggior attenzione è stato il rapporto tra il flusso del boro e quello del carbonio, essendo il boro la componente secondaria più abbondante nella radiazione cosmica. La figura 7 mostra questo rapporto misurato da recenti esperimenti su

⁴Il boro non è prodotto nei processi di combustione stellare, e quindi è relativamente meno abbondante rispetto, ad esempio, al ^{12}C che, invece, si produce nelle stelle con un processo chiamato 3α .

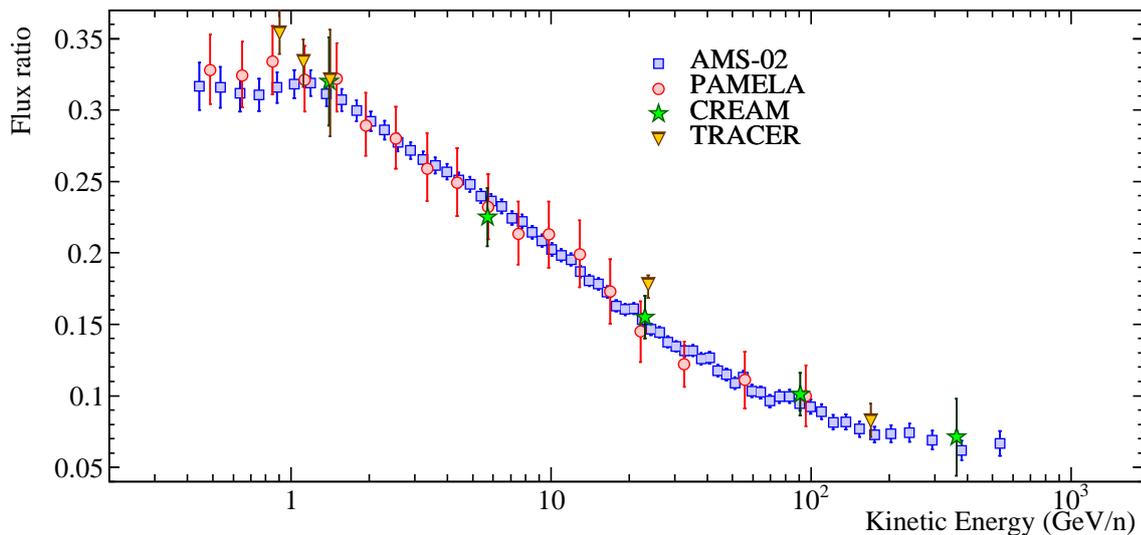


Figura 7: Il rapporto tra i flussi di boro e carbonio misurato da recenti esperimenti su palloni stratosferici e spaziali.

palloni stratosferici e spaziali. L'andamento del rapporto in funzione dell'energia è sostanzialmente consistente con quanto ci si aspetterebbe teoricamente ed il suo valore permette di stimare il coefficiente di diffusione nella galassia (il termine D dell'equazione (1)). Inoltre, mostra che anche per i nuclei di boro vi è un indurimento dello spettro consistente con l'ipotesi che siano effetti nella diffusione dei raggi cosmici che causano i cambi spettrali nei flussi dei nuclei di raggi cosmici.

Anche per i secondari è molto interessante studiare le componenti isotopiche dei nuclei, in particolare quelle instabili come ^{10}Be , ^{26}Al e ^{36}Cl . Infatti, siccome questi isotopi hanno vite medie di milioni di anni, studiarne la frazione rispetto agli isotopi stabili nella radiazione cosmica dà informazioni sul tempo di permanenza, ora stimato ~ 15 milioni anni, dei raggi cosmici nella galassia e sulla loro propagazione. La maggior parte delle misure disponibili sono limitate ad energie inferiori ad alcune centinaia di MeV per nucleone. È in previsione un esperimento, HELIX, su un volo in pallone di lunga durata dall'Antartide che dovrebbe estendere queste misure sopra il GeV/n.

Componenti rare

Accanto agli elementi più abbondanti come protoni ed elio, nella radiazione cosmica è presente anche una componente di elettroni ed una, mol-

to più rara, di antiparticelle. Per motivi diversi lo studio di queste componenti, benché rare, è particolarmente interessante e foriero di risultati potenzialmente molto importanti.

Elettroni

Gli elettroni sono la componente di carica negativa più abbondante della radiazione cosmica ma rappresentano solo l'1% del flusso totale dei raggi cosmici. Tuttavia, a causa della loro piccola massa risentono di processi di perdita di energia diversi dagli elementi di massa maggiore fornendo, quindi, informazioni sulla propagazione nella galassia diversi da quelli ottenibili con la componente adronica. Infatti gli elettroni perdono frazioni consistenti della loro energia per radiazione di sincrotrone interagendo col campo magnetico galattico, e, mediante un processo Compton inverso, interagendo coi fotoni ambientali. Le perdite di energia sono talmente significative che gli elettroni di alta energia ($\gg 100$ GeV) possono raggiungere il sistema solare solo se le loro sorgenti sono relativamente vicine su scala galattica. Gli elettroni sono principalmente di origine primaria accelerati da sorgenti astrofisiche come gli altri raggi cosmici ma vi è anche una importante, dell'ordine del 10%, componente secondaria risultante dalle interazioni dei raggi cosmici col mezzo interstellare.

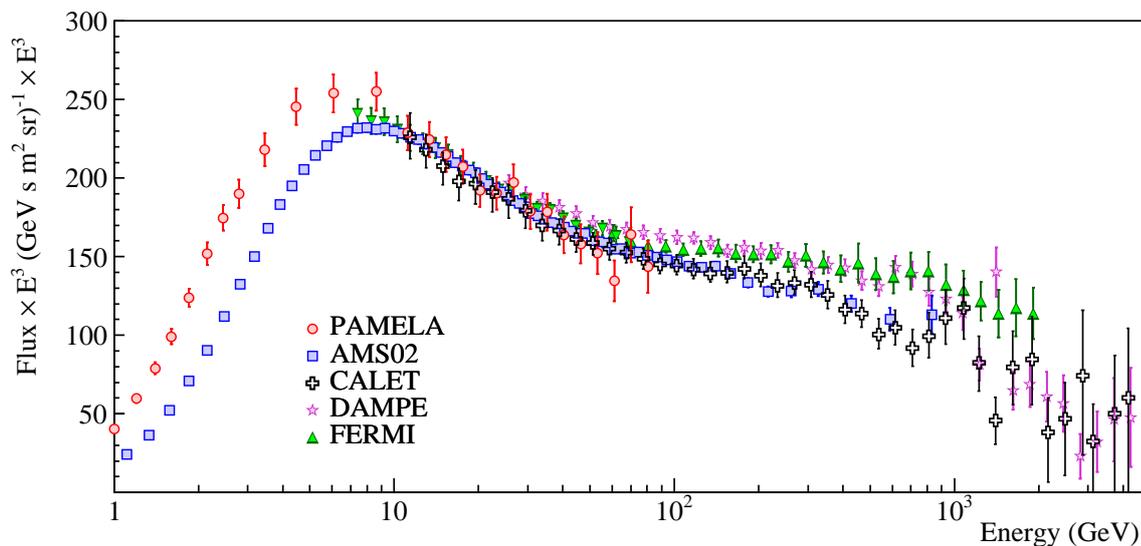


Figura 8: Lo spettro in energia di tutti gli elettroni misurato da recenti esperimenti spaziali. I flussi sono moltiplicati per E^3 , con E l'energia cinetica in GeV.

La figura 8 mostra le più recenti misure dirette di elettroni ottenute da esperimenti spaziali. È da notare che questi risultati riguardano la somma del flusso di elettroni con quello dei positroni (l'antiparticella degli elettroni rappresenta circa il 10% di questo flusso) in quanto diversi dei risultati presentati in figura sono stati ottenuti da apparati privi di magneti, quindi non in grado di separare le particelle in base al segno della loro carica.

Alcune considerazioni interessanti che si possono trarre da questi dati sono le seguenti:

- l'andamento dello spettro cambia significativamente sopra ad 1 TeV. Questo potrebbe essere dovuto al fatto che elettroni di alte energie siano primariamente prodotti da una singola sorgente, probabilmente localizzata a poche centinaia di parsec dal sistema solare. Oppure potrebbe essere il risultato della combinazione degli effetti di accelerazione presso sorgenti distribuite su scala galattica e delle perdite di energia dovute alla propagazione nella galassia;
- ad energie intermedie, inferiori al TeV ma superiori a quelle per cui la modulazione solare ha un'influenza significativa, lo spettro è una legge di potenza con indice spettrale $\sim -3.1 \div -3.2$, consistente con quanto atteso dai modelli standard di propagazione dei raggi cosmici;

- ci sono delle evidenti differenze tra le varie misure, non ancora capite ma probabilmente imputabili ad effetti sistematici degli apparati non ancora identificati.

Elettroni e positroni

Le misure di figura 8 si riferiscono alla somma di elettroni e positroni, che il modello standard dei raggi cosmici assume avere origini parzialmente diverse. Ovvero, la maggior parte degli elettroni sono di origine primaria accelerati in SNR con una piccola frazione di origine secondaria comparabile alla componente di positroni, totalmente di origine secondaria. Si capisce, quindi, che la misura separata delle due componenti assume una notevole valenza nello studio dei raggi cosmici. La figura 9 a sinistra mostra le misure più recenti dei flussi separati di elettroni e positroni ottenute dagli esperimenti spaziali Fermi, PAMELA ed AMS-02. È degno di nota il fatto che l'esperimento Fermi, benché non dotato di magneti, sia riuscito ad ottenere dati sulle due componenti utilizzando il campo magnetico terrestre osservando la diversa direzione di arrivo che questo induceva nelle particelle elettromagnetiche identificate dall'apparato. Escludendo le energie più basse affette dal vento solare, un eccellente accordo si osserva tra i tre gruppi di misura con l'unica eccezione lo spettro di positroni misurato da Fermi che è sistematicamente più

alto di quello misurato da AMS-02 e PAMELA. Questo potrebbe essere dovuto ad un residuo di contaminazione di protoni nei dati dell'esperimento Fermi, considerando che l'apparato, progettato per misurare fotoni gamma, non ha la stessa capacità identificativa degli altri due.

La figura 9, a destra, mostra la frazione di positroni, ovvero il rapporto tra i flussi (ϕ) di positroni e la somma dei flussi di elettroni e positroni: $\phi(e^+) / (\phi(e^+) + \phi(e^-))$. Storicamente è questa la quantità che è stata maggiormente misurata essendo decisamente più semplice da ottenere sperimentalmente non richiedendo la completa conoscenza di tutti gli effetti strumentali che entrano nel calcolo dei flussi, in quanto questi possono essere assunti simili per elettroni e positroni e, quindi, cancellarsi nel rapporto. Nella figura, accanto a misure recenti, viene mostrata anche una predizione teorica di questa frazione assumendo che tutti i positroni siano prodotti dall'interazione dei raggi cosmici col mezzo interstellare.

Diversi punti balzano all'occhio osservando queste figure:

- i dati sperimentali sotto i 5 GeV sono abbastanza consistenti o più bassi della predizione per la produzione secondaria;
- tuttavia i dati sperimentali per i diversi esperimenti differiscono significativamente a queste basse energie;
- al contrario sono in sostanzialmente in accordo ad energie più elevate, sopra i 10 GeV;
- dove, però, sono significativamente più alti nell'ipotesi di produzione secondaria.

In realtà le differenze tra i dati sperimentali a basse energie (< 10 GeV) non sono preoccupanti in quanto sono molto probabilmente dovute ad effetti di modulazione solare che condizionano differentemente raggi cosmici di segno di carica diversa (si ricordi che l'interazione col vento solare è praticamente un'interazione con campi magnetici). Infatti queste misure sono state ottenute durante fasi diverse e polarità diverse del ciclo solare e se si confrontano misure effettuate nello stesso periodo tra PAMELA ed AMS-02 si trova un ottimo accordo tra le stesse.

L'aumento della frazione di positroni ad alta energia, per la prima volta chiaramente identificato dall'esperimento PAMELA, è, come si vede dalla figura 9 a sinistra, dovuto ad uno spettro di e^+ più duro di quello degli e^- . Questo aumento è difficile da spiegare col modello standard dei raggi cosmici, rappresentato dalla linea continua in figura 9 a destra, che, invece, predice una diminuzione della frazione ad alte energie. Ciò implica che o vi è una componente primaria di e^+ (e, probabilmente, in egual misura un'ulteriore componente primaria di e^-) oppure il paradigma dei raggi cosmici richiede delle modifiche.

Negli anni successivi alle misure di PAMELA ed AMS-02 innumerevoli ipotesi sono state avanzate per spiegare questi risultati, con l'esistenza di sorgenti di positroni quella che ha attirato il maggior interesse. Tra le sorgenti proposte vi sono sorgenti astrofisiche come le pulsar ma anche la possibilità che positroni vengano prodotti ed accelerati durante la fase di accelerazione dei raggi cosmici nelle SNR. Per ovvi motivi, molti studi si sono concentrati sulla possibilità che questi positroni in eccesso possano risultare dall'annichilazione o dal decadimento della materia oscura. Le prove dell'esistenza della materia oscura sono molteplici ma la sua esatta natura è ancora ignota. Tra i possibili candidati per la materia oscura che sono stati proposti vi sono le cosiddette WIMP ("Weakly Interactive Massive Particle"), particelle risultanti da un'estensione del modello standard della fisica fondamentale con masse possibili dal GeV a molti TeV. Queste particelle possono annichilare producendo particelle del modello standard tra cui positroni ed elettroni. Questa spiegazione però incontra due problemi: l'assenza di un segnale equivalente negli antiprotoni, come verrà discusso a breve, e nella radiazione gamma da sorgenti dove ci si attende una forte presenza di materia oscura, come le galassie sferiche nane. Per una discussione più approfondita sulla materia oscura si può vedere [12] e sulla ricerca della materia oscura con antiparticelle nei raggi cosmici si può vedere [13].

Antinuclei

Come si diceva, segnali di materia oscura possono essere cercati anche in altre componenti rare

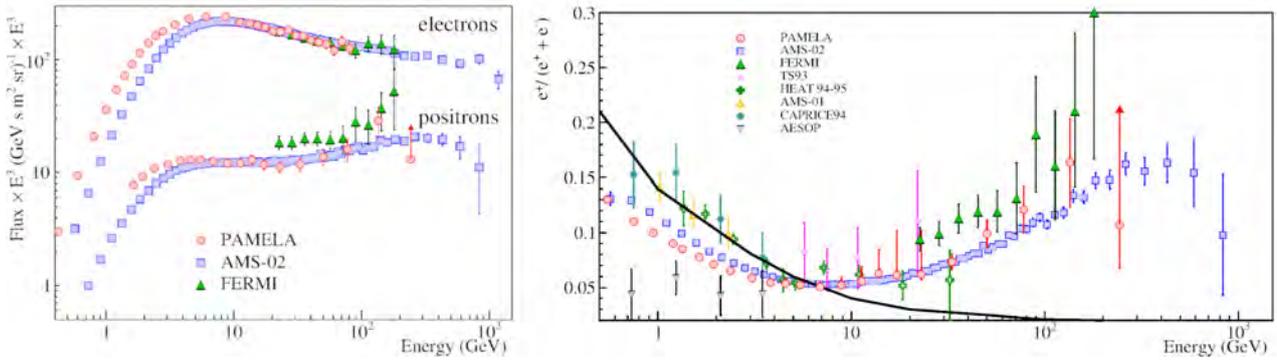


Figura 9: A sinistra: gli spettri in energia degli elettroni e dei positroni misurati degli esperimenti spaziali PAMELA, Fermi ed AMS-02. I flussi sono moltiplicati per E^3 , con E l'energia cinetica in GeV. A destra: una raccolta di alcuni dei risultati più recenti sulla frazione di positroni. La linea continua è la predizione assumendo che tutti i positroni siano di origine secondaria.

come gli antiprotoni e gli antinuclei. Sin dalle prime osservazioni degli antiprotoni nella radiazione cosmica negli anni 70 del secolo scorso, si capì subito che gli antiprotoni, come i positroni, permettono di studiare importanti aspetti della radiazione cosmica. Come i positroni, gli antiprotoni sono prodotti durante la propagazione dei raggi cosmici nella galassia. Accanto a questa produzione secondaria, una componente primaria potrebbe risultare dall'annichilazione o decadimento della materia oscura. Negli ultimi anni, nuovi esperimenti, come PAMELA ed AMS-02 nello spazio o come BESS-Polar su palloni stratosferici con voli di lunga durata sopra l'Antartide, hanno significativamente cambiato lo scenario sperimentale con un aumento di ordini di grandezza della statistica esistente sulle antiparticelle.

La figura 10 mostra i risultati più recenti sul rapporto dei flussi di antiprotoni su protoni assieme ad una predizione teorica per antiprotoni di origine secondaria (stesso modello della predizione per positroni in figura 9, destra). I dati sperimentali sono in buon accordo tra di loro e con la predizione secondaria anche se vi è un appiattimento ad alte energie non previsto dal modello. Differenza, però, che si riduce significativamente se il modello viene aggiornato coi dati più recenti sulle sezioni d'urto di produzione di antiprotoni da esperimenti ad acceleratori.

Il buon accordo tra dati sperimentali di antiprotoni e predizioni secondarie mette limiti importanti sulla presenza di componenti primarie da annichilazione e decadimento di materia oscura, escludendo anche molte interpretazioni dei

risultati sui positroni come segnali indiretti di materia oscura. Purtroppo, un lieve eccesso rispetto alla predizione secondaria nei dati di PAMELA e, soprattutto, AMS-02 tra alcuni GeV e 20 GeV ha spinto molti autori a proporre la presenza di una componente primaria da annichilazione di una particella di materia oscura con massa di alcune decine di GeV.

Purtroppo queste misure sono pesantemente condizionate dalle incertezze sulle sezioni d'urto di produzione e sulla propagazione dei raggi cosmici nella galassia. Questo perché l'eventuale componente primaria è una frazione della componente secondaria, rendendo le incertezze teoriche sulla propagazione dei raggi cosmici importanti quanto l'eventuale segnale stesso. Altri antinuclei, però, come l'antideuterio e l'antielio non soffrono di questo problema, in quanto la componente secondaria, soprattutto sotto qualche GeV, è pesantemente soppressa dall'alta soglia energetica richiesta per la produzione in collisioni di questi antinuclei. Esistono e permangono come plausibili molti modelli di annichilazione o decadimento di WIMP che predicono flussi di antideuterio ed antielio ordini di grandezza superiori alla componente secondaria. Per il momento non vi sono misure validate di antideuterio o di antielio, però sia l'attualmente operativo esperimento AMS-02 che il futuro esperimento GAPS [14], previsto volare su pallone stratosferico sopra l'Antartide nel prossimo futuro, promettono risultati su questi antinuclei che potrebbero cambiare fondamentalmente la ricerca sulla natura della materia oscura.

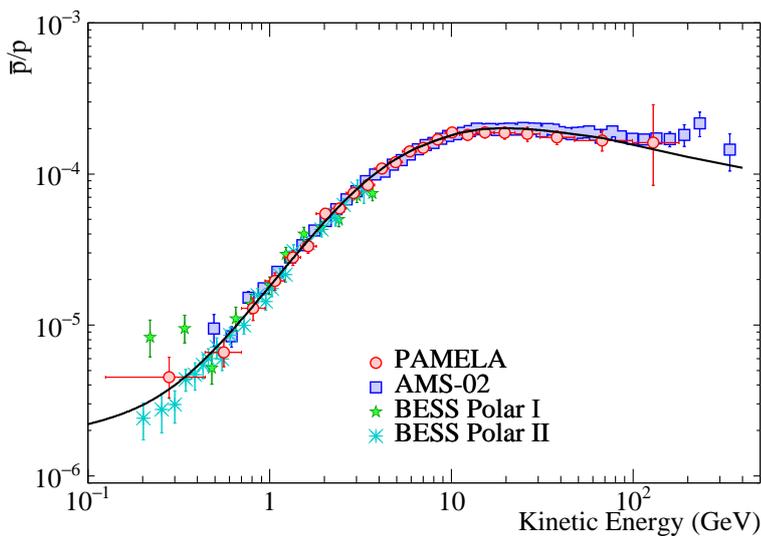


Figura 10: Il rapporto tra i flussi di antiprotoni e di protoni misurato da recenti esperimenti su palloni stratosferici e spaziali. La linea continua è la predizione assumendo che tutti gli antiprotoni siano di origine secondaria.

Future missioni spaziali

In questi ultimi anni nuovi esperimenti, soprattutto nello spazio, equipaggiati con tecnologie all'avanguardia hanno permesso di far progredire significativamente la conoscenza della fisica dei raggi cosmici spingendo a rivedere i modelli di accelerazione e propagazione degli stessi. Con questa grande quantità di nuovi dati, però, sono sorte anche nuove domande che richiedono nuovi esperimenti, più grandi e più sofisticati. Alcuni sono in fase di sviluppo, altri ancora in una fase di progettazione.

I nuovi esperimenti dovranno espandere la regione di energia esplorata misurando direttamente lo spettro e composizione dei raggi cosmici fino alla zona del ginocchio per permettere di capire se vi sono altre strutture negli spettri e come varia la composizione quando ci si avvicina al ginocchio, dove il paradigma dei raggi cosmici prevede che siano assenti, ed in maniera diversa, in base al loro numero di carica, i raggi cosmici accelerati da SNR. Inoltre, dovranno cercare di misurare con più precisione le componenti rare, estendendone le misure ad energie superiori al TeV.

L'esperimento in fase di sviluppo più avanzata è l'esperimento HERD [15] proposto per essere installato sulla futura stazione spaziale cinese nella seconda metà di questo decennio. L'apparato, mostrato nella figura 11, sarà composto da un calorimetro a cubetti (CALO), circondato da un sistema di tracciamento (STK) di microstrip di silicio a suo volta circondato da scintillatori

plastici (PSD) e un rivelatore di radiazione di transizione (TRD) piazzato in un lato. L'esperimento è stato disegnato per misurare lo spettro e la composizione dei raggi cosmici fino al ginocchio (≈ 1 PeV) e per misurare la componente di elettroni (e^- più e^+ non avendo campo magnetico) fino ad una decina di TeV. Non avendo un campo magnetico, HERD non potrà chiarire completamente l'origine della componente di antimateria.

La misura di questa componente fino ad alcuni TeV è, invece, l'obiettivo scientifico di due missioni spaziali attualmente in fase di studio: ALADInO (Antimatter Large Acceptance Detector In Orbit) [16] ed AMS-100 [17], un'estensione dell'attuale esperimento AMS-02. Entrambi questi esperimenti sono pensati aventi rivelatori per l'identificazione di particelle assieme ad un magnete superconduttore che permetterebbe di estendere di almeno un ordine di grandezza l'accettanza e la rigidità misurabile con precisione di apparati esistenti come AMS-02.

Benché l'osservazione di raggi cosmici di altissima energia, dal ginocchio fino a sopra i 10^{19} eV, sia condizionata dall'estremamente basso flusso di particelle, circa 1 particella per km^2 al secolo per energie dell'ordine di 5×10^{19} eV, che rende l'osservazione diretta di queste particelle non praticabile, esperimenti dallo spazio sono stati proposti per questo campo della fisica dei raggi cosmici. Come descritto nell'articolo di A. Chiavassa in questo numero di **Ithaca**, tutte le misure di raggi cosmici di altissima energia sono state

ottenute con strumentazione a terra che misura i prodotti dell'interazione dei raggi cosmici con l'atmosfera, derivando da questi informazioni sulle energie e tipo di raggio cosmico. In questo modo si sono potuti costruire esperimenti che campionano enormi superfici (l'osservatorio Pierre Auger in Mendoza, Argentina ha 1600 water rivelatori Čerenkov detectors distribuiti su una superficie di 3,000 km²) compensando conseguentemente il bassissimo flusso della radiazione cosmica. Pur tuttavia, i dati disponibili sull'energie estreme (UHECR, energie $\geq 10^{19}$ eV) sono ancora limitati e la loro origine e natura rimangono un'incognita. Per studiare queste energie estreme, già molti anni fa fu proposto di andare nello spazio da cui osservare la luce di fluorescenza prodotta dagli sciami atmosferici estesi originati dai raggi cosmici che interagiscono con l'atmosfera terrestre. Osservando l'atmosfera da centinaia di km di altezza è possibile aumentare l'accettanza osservativa di oltre 2 ordini di grandezza rispetto ad apparati terrestri come l'osservatorio AUGER, migliorando di più di un fattore 10 (vi è una riduzione dovuta al ridotto tempo osservativo, dato che questi apparati possono operare solo in condizioni di buio) la statistica esistente. Attualmente c'è una mis-

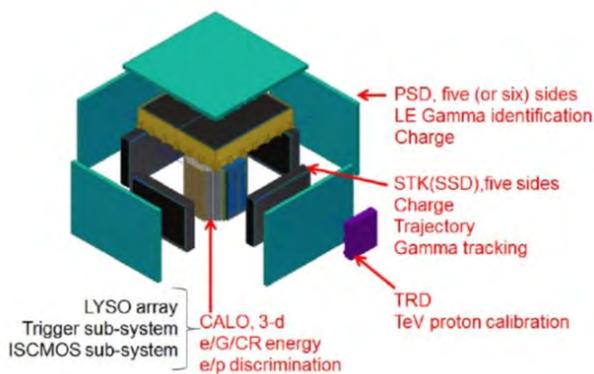


Figura 11: L'apparato HERD.

sione spaziale, POEMMA [18], dedicata ai UHECR in fase di studio da parte della NASA per un'eventuale missione nel prossimo decennio. POEMMA è l'evoluzione di proposte precedenti come OWL ed JEM-EUSO e di studi di fattibilità condotti con palloni stratosferici: EUSO-SPB1 e 2. La missione verrebbe posta in un'orbita circolare a 525 km di altitudine da cui osservare sia la luce di fluorescenza che la radiazione Čerenkov degli sciami atmosferici. Se e quando POEMMA

entrerà in operazione migliorerà di un fattore 10 la statistica sugli UHECR permettendo di dare risposta alle domande ancora aperte sulla fine dello spettro dei raggi cosmici.

Per una discussione più approfondita su questi argomenti si veda [19].



- [1] P. Carlson and A. De Angelis: *Nationalism and internationalism in science: the case of the discovery of cosmic rays*, Eur. Phys. J. H, 35 (2010) 309.
- [2] V. L. Ginzbur, S. I. Syrovatskii: *The Origin of Cosmic Rays*, Macmillan, New York (1964).
- [3] P. Picozza et al.: *PAMELA - A Payload for Antimatter Matter Exploration and Light-nuclei Astrophysics*, Astropart. Phys., 27 (2007) 296.
- [4] R. Battiston: *The antimatter spectrometer (AMS-02): A particle physics detector in space*, Nucl. Instrum. Methods Phys. Res., Sect. A, 588 (2008) 227.
- [5] R. L. Golden et al.: *Evidence for the Existence of Cosmic-Ray Antiprotons*, Phys. Rev. Lett., 43 (1979) 1196.
- [6] M. L. Ambriola et al.: *CAPRICE98: A balloon borne magnetic spectrometer to study cosmic ray antimatter and composition at different atmospheric depths*, Nucl. Phys. B, 78 (1999) 32.
- [7] A. Yamamoto: *BESS-Polar long duration balloon flights at Antarctica to search for primordial antiparticle*, Prog. Theor. Phys. Suppl., 151 (2003) 240.
- [8] S. Torii et al.: *The CALorimetric Electron Telescope (CALET): High Energy Astroparticle Physics Observatory on the International Space Station*, Proc. of Science, ICRC2015 (2015) 581.
- [9] J. Chang et al.: *The DARK Matter Particle Explorer mission*, Astropart. Phys., 95 (2017) 6.
- [10] P. Blasi: *Recent results in cosmic ray physics and their interpretation*, Braz. J. Phys., 44 (2014) 426.
- [11] P. Serpico: *Possible physics scenarios behind cosmic-ray "anomalies"*, Proc. of Science, ICR2015 (2016) 09.
- [12] M. Cirelli: *La Materia Oscura*, Ithaca, XII (2018) 117.
- [13] M. Boezio et al.: *PAMELA and Indirect Dark Matter Searches*, New J. Phys., 11 (2009) 105023.
- [14] C. J. Hailey: *An indirect search for dark matter using antideuterons: the GAPS experiment*, New J. Phys., 11 (2009) 105022.
- [15] S. N. Zhang et al.: *Introduction to the High Energy cosmic-Radiation Detection (HERD) Facility onboard China's Future Space Station*, Proc. of Science, 2017 (2018) 1077.
- [16] R. Battiston et al.: *High precision particle astrophysics as a new window on the universe with an Antimatter Large Acceptance Detector In Orbit (ALADInO)*, Exper. Astron., 51 (2021) 1331., (erratum) Exper. Astron. 51 (2021) 1299.

- [17] S. Schael et al.: *AMS-100: The Next Generation Magnetic Spectrometer in Space - An International Science Platform for Physics and Astrophysics at Lagrange Point 2*, Nucl. Instrum. Meth. Phys. Res. A, 944 (2019) 162561.
- [18] A. Olinto et al.: *The POEMMA (Probe of Extreme Multi-Messenger Astrophysics) mission*, Proc. of Science, ICRC2019 (2020) 378.
- [19] M. Boezio, R. Munini, P. Picozza: *Cosmic ray detection in space*, Prog. Part. Nucl. Phys., 112 (2020) 103765.



Mirko Boezio: è dirigente di ricerca dell'Istituto Nazionale di Fisica Nucleare presso la sezione di Trieste. Lavora da trent'anni nel campo della fisica sperimentale astroparticellare e dei raggi cosmici. È stato il responsabile nazionale dell'esperimento PAMELA ed attualmente è il responsabile nazionale dell'esperimento GAPS.

Riccardo Munini: è ricercatore dell'Istituto Nazionale di Fisica Nucleare presso la sezione di Trieste. Lavora da una decina di anni nel campo della fisica sperimentale dei raggi cosmici partecipando all'analisi dei dati dell'esperimento PAMELA ed alla realizzazione dell'esperimento GAPS.

UHECR: raggi cosmici di energia ultra elevata

Non est ad astra mollis e terris via.

Seneca, Hercules Furens

Sergio Petrerà

Gran Sasso Science Institute, L'Aquila, Italy

Francesco Salamida

Dipartimento di Scienze Fisiche e Chimiche - Università dell'Aquila, Italy

Lo studio dei raggi cosmici di energia ultra elevata (UHE, $E > 10^{18}$ eV) permette di aprire uno spiraglio nella comprensione della loro origine. A queste energie, infatti, l'Universo non è trasparente ai raggi cosmici, nel loro percorso dalle sorgenti fino a Terra, a causa del fondo cosmico di fotoni che lo pervade. Mostreremo, con argomenti di base, come è possibile trattare la propagazione di questi raggi cosmici, qual è la fisica particellare e nucleare coinvolta e, infine, come lo studio combinato delle osservabili a Terra (in particolare, spettro di energia e composizione) permette di inferire le caratteristiche delle sorgenti di raggi cosmici.

UHECR: particolarità di scienza e tecnica di rivelazione

Dopo oltre 100 anni dalla scoperta dei raggi cosmici da parte di Victor Hess, la loro origine è in gran parte sconosciuta. Il motivo principale è che, indipendentemente dalla risoluzione angolare con cui vengono rivelati, la loro direzione è

in gran parte scorrelata con quella delle sorgenti. Il metodo intuitivo di associazione alla sorgente attraverso la direzione, come avviene per la luce (o più in generale, per i fotoni alle varie frequenze), è infatti reso impossibile dai campi magnetici presenti nella nostra Galassia (dell'ordine di 10^{-6} G) e dello spazio extragalattico (dell'ordine di 10^{-9} G).

Dagli studi sulle abbondanze degli elementi presenti nei raggi cosmici, relative a quelle solari [1], si deduce che i raggi cosmici di origine galattica con energie inferiori a $\sim Z \cdot 10^{18}$ eV (Z , carica del nucleo) sono confinati dal campo magnetico per milioni di anni. Questo confinamento preclude ogni possibilità di associazione con le loro sorgenti. È invece possibile l'associazione a classi di sorgenti astrofisiche: argomenti di energetica stellare hanno portato Ginzburg and Syrovatskii [2] a proporre i *Supernova Remnants* (SNR) come possibili sorgenti. Il meccanismo di accelerazione in questi siti astrofisici sarebbe la *shock acceleration* [1], anche nota come *first-order Fermi acceleration*.

Detto questo, quale finestra di scienza aprono i raggi cosmici di energia superiore a 10^{18} eV (UHECR, **Ultra-High Energy Cosmic Rays**)? Essi sono prevalentemente extra-galattici e, inol-

tre, il paradigma della *SNR shock acceleration* risulta inadatto ad accelerare questi UHECR fino alle energie osservate (oltre 10^{20} eV). Le quantità osservabili disponibili sono le stesse che in altri *range* di energia (spettro, composizione, anisotropia), ma dal punto di vista interpretativo interviene un nuovo elemento: la propagazione nel fondo cosmico, cioè le interazioni dei raggi cosmici con i campi di fotoni che pervadono l'Universo.

Per quanto il tempo di residenza dei raggi cosmici nella nostra Galassia sia di milioni di anni, la probabilità che essi interagiscano con il fondo cosmico è praticamente nulla. Questo rende le osservabili praticamente scorrelate fra loro per energie al di sotto di 10^{18} eV. Quindi protoni/nuclei componenti i raggi cosmici sono rivelati a Terra senza alcun cambiamento¹ (energia e massa) rispetto a come sono prodotti dalle sorgenti.

Il carattere extra-galattico degli UHECR ha delle conseguenze che incidono su ognuna delle osservabili:

- poichè le distanze coinvolte (Sorgente-Terra) possono essere su scala cosmica, occorre tenere conto degli effetti cosmologici: perdite di energia adiabatiche (dovute al raffreddamento delle energie per l'espansione dell'Universo) ed evoluzione cosmologica delle sorgenti;
- i raggi cosmici emessi dalle sorgenti attraversano il fondo cosmico di fotoni, interagendo con esso. Alcuni processi sono, dal punto di vista del nucleo, di tipo elastico e comportano solo la perdita di energia cinetica, altri, inelastici, anche il cambiamento della massa del nucleo.

È proprio attraverso queste interazioni, le cui caratteristiche fondamentali sono note, che nel caso degli UHECR le osservabili risultano differenti da quelle alla sorgente e correlate tra loro, proprio a seguito della propagazione. Quindi, attraverso la modellizzazione del processo di propagazione, è possibile risalire alle proprietà delle

¹A parte ovviamente la produzione di nuclei prodotti di spallazione nell'interazione con la materia intergalattica, nuclei che costituiscono la principale evidenza del confinamento galattico.

sorgenti. Questa è una peculiarità di questa regione energetica nel vasto scenario della Fisica dei Raggi Cosmici.

Va anche aggiunto che per gli UHECR l'associazione delle direzioni alle sorgenti è resa possibile alle energie più elevate. Anche se le incertezze sull'intensità e la coerenza dei campi magnetici extra-galattici sono ancora molto grandi, la deflessione angolare attesa per i raggi cosmici di energia vicino a 10^{20} eV è di qualche grado $\times Z$ [3] e quindi eventuali protoni in questo range di energia possono essere associati alle loro sorgenti.

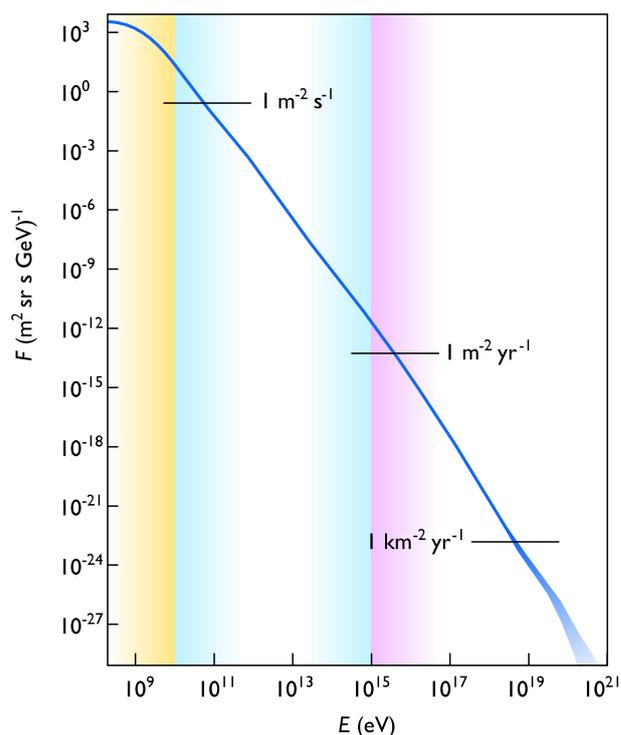


Figura 1: Flusso dei raggi cosmici da 10^9 fino a 10^{21} eV. Con un andamento a legge di potenza il flusso si estende su 30 ordini di grandezza. La zona attorno ai 10^{15} eV è chiamata *ginocchio knee*, quella attorno ai 10^{18} eV *caviglia ankle*.

Esiste, infine, un'altra peculiarità. Questa è legata alla loro rivelazione. Gli UHECR incidono sull'atmosfera terrestre con un'intensità così bassa da richiedere superfici di osservazione di centinaia o migliaia di km^2 . Una griglia di rivelatori sparsi su tale area (**SD, Surface Detector Array**) è la tecnica comunemente usata, con sostanziale continuità da energie di 10^{14} eV in su. Ma proprio nel *range* UHE si rende disponibile una tecnica di rivelazione aggiuntiva, la rivelazione della fluorescenza atmosferica, **FD, Fluore-**

science Detection, che risulta invece inefficiente a energie inferiori. Questa si basa sull'emissione ultravioletta (UV) prodotta dalla diseccitazione (isotropa) delle molecole di azoto dell'atmosfera e indotta dagli elettroni di ionizzazione, numerosissimi negli sciami atmosferici UHE. L'accoppiamento delle due tecniche è stato il principale *breakthrough* degli ultimi 20 anni [4], iniziato dal Pierre Auger Observatory e in seguito adottato anche dall'esperimento Telescope Array [5]. Tale accoppiamento permette la calibrazione dell'energia degli sciami atmosferici negli eventi ibridi (cioè rivelati da entrambi i sistemi di rivelazione, SD e FD) e ha sgomberato il campo dalle innumerevoli incertezze insite nella calibrazione basata sulle simulazioni Monte Carlo a cui si affidavano gli esperimenti precedenti.

L'Osservatorio Pierre Auger

Il flusso dei raggi cosmici si estende su 30 ordini di grandezza (Fig. 1) in un intervallo di energia che va dal GeV fino a quelle estreme di oltre 10^{20} eV, dove il flusso si abbassa fino a 1 particella per km^2 per secolo. La tecnica di rivelazione utilizzata per catturare queste particelle dipende dunque fortemente dall'energia alla quale si vogliono misurare. In particolare, se si vuole avere adeguata statistica di eventi ad energie sopra i 10^{19} eV (dove, peraltro, la probabilità che i raggi cosmici puntino direttamente alle sorgenti è più alta), è fondamentale costruire rivelatori sempre più grandi. Con questa idea nel 1992 James Cronin e Alan Watson proposero l'idea di costruire un rivelatore di superficie di dimensioni colossali e tentare così di rispondere alle domande fondamentali sull'origine dei raggi cosmici di altissima energia. La proposta, ampiamente supportata dalla comunità scientifica ed elaborata durante un workshop di 6 mesi presso il Fermi National Accelerator Laboratory, portò alla preparazione di un progetto denominato Giant Array Project che ha poi dato vita all'Osservatorio Pierre Auger. In una fase iniziale, si era pensato ad un Osservatorio da 3000 km^2 in ciascuno degli emisferi, in modo da ottenere una copertura completa del cielo, al fine di identificare le sorgenti, cosa poi non realizzatasi. Nella ricerca di un sito adatto ad ospitare l'Osservatorio sono stati considera-

ti diversi fattori, fondamentali al raggiungimento degli obiettivi scientifici prefissi. Ad esempio la quota avrebbe dovuta essere tra i 500 e i 1500 metri sul livello del mare, per ottimizzare la rivelazione di sciami originati da primari con energia di 10^{19} eV. Inoltre, i requisiti per la comunicazione tra le diverse stazioni e per il posizionamento dei rivelatori richiedevano un sito relativamente pianeggiante, con poca vegetazione e scarsa popolazione. Nel 1998 dopo una accurata ricerca in Australia, Sud Africa e Sud America, la Collaborazione internazionale Auger ha scelto di costruire l'Osservatorio nella Pampa Amarilla in Argentina e precisamente nel Dipartimento di Malargüe, Provincia di Mendoza situato tra i 1340 e i 1610 metri sul livello del mare, corrispondenti a circa 875 g/cm^2 di profondità atmosferica verticale. A parte le dimensioni, la caratteristica innovativa dell'Osservatorio Pierre Auger è il concetto di *hybrid detection*, ossia la combinazione delle due diverse tecniche di misura SD e FD, usate, in precedenza, in modo separato da altri esperimenti, come AGASA[6] e Fly's Eye[7].

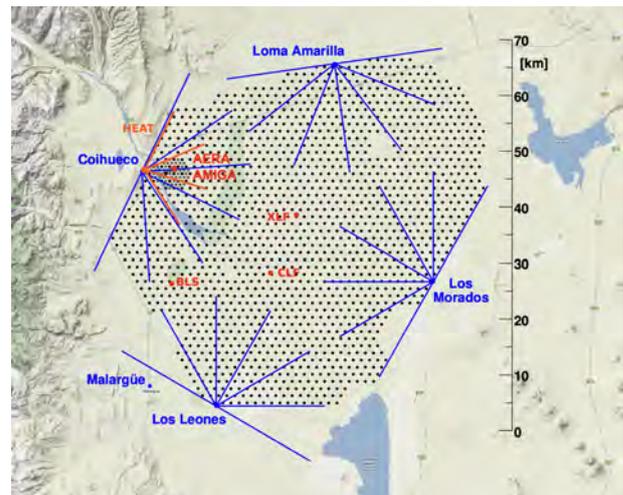


Figura 2: Schema dell'Osservatorio Pierre Auger. I punti neri rappresentano le posizioni delle circa 1600 stazioni SD, mentre le linee azzurre rappresentano il campo di vista azimutale dei 24 telescopi FD, distribuiti in 4 siti (Los Leones, Los Morados, Loma Amarilla, Coihueco).

Il design di base dell'Osservatorio Pierre Auger [8], mostrato in Fig. 2, è costituito da un array di SD formato da 1600 rivelatori di superficie, disposti su di una griglia triangolare e spazati di 1500 m. Il rivelatore di fluorescenza (FD) è invece costituito da 24(+3) telescopi di fluore-

scenza, suddivisi in 4(+1) edifici, posti sul bordo dell'array in modo da catturare la luce di fluorescenza prodotta dagli sciame in atmosfera. In aggiunta un'area di 23.5 km² è coperta da rivelatori di superficie posti ad una distanza di 750 m in modo da rivelare gli eventi di energia più bassa (intorno a 3×10^{17} eV) con alta efficienza.

Array di Superficie (SD)

L'array SD misura la densità di particelle secondarie che, prodotte nello sciame, raggiungono la superficie dando informazioni su energia, direzione e tempo di arrivo del primario. Ogni rivelatore dell'SD è costituito da una cisterna cilindrica del diametro di 3.6 m e altezza di 1.2 m. La cisterna di polietilene è rivestita internamente da un materiale riflettente ed è riempita con 1200 litri di acqua ultra-pura. Le particelle cariche relativistiche, attraversando la cisterna, generano luce Čerenkov che viene poi rivelata da 3 fotomoltiplicatori (PMT) da 9 pollici, posti sulla superficie della cisterna ad una distanza di 1.2 m dal centro dalla stessa. Tutta la strumentazione, che include i PMT, un ricevitore GPS ed un ricevitore/trasmittitore radio per le comunicazioni tra il rivelatore e le stazioni di controllo, è alimentata da batterie connesse a due pannelli solari. Uno schema viene mostrato in Fig. 3.

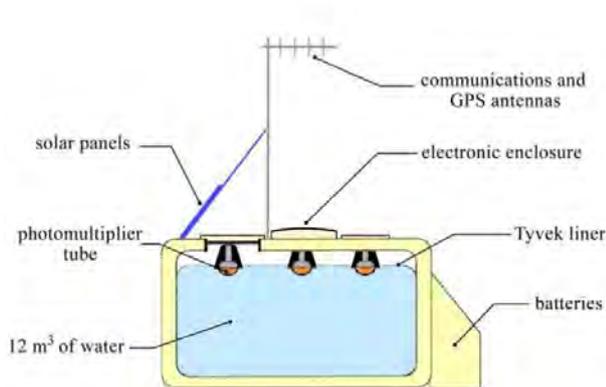


Figura 3: Vista schematica di un rivelatore water Čerenkov posto nella Pampa Amarilla.

Rivelatore di Fluorescenza (FD)

Ogni telescopio di fluorescenza ha un campo di vista di $30^\circ \times 30^\circ$, che a gruppi di sei realizzano una copertura angolare di 180° in azimuth per

ogni sito FD. L'elevazione minima di ogni telescopio è 1.5° sopra l'orizzonte e, dunque, ognuno copre angoli di elevazione tra 1.5° e 31.5° rispetto all'orizzontale. Una vista schematica di un telescopio di fluorescenza è riportata in Fig. 4. Le molecole di azoto nell'atmosfera eccitate dalle particelle cariche dello sciame, emettono, diseccitandosi, luce UV nell'intervallo tra 300 nm e 400 nm. Questa luce, detta di fluorescenza, raggiunge i telescopi entrando attraverso un diaframma circolare di raggio 1.1 m. Un anello correttore, di raggio interno da 0.85 m e composto da 24 settori di materiale trasparente alla luce UV, è posto intorno al diaframma al fine di correggere le aberrazioni sferiche e di coma. Di seguito un specchio sferico segmentato da $3.5 \text{ m} \times 3.5 \text{ m}$ convoglia la luce su di una camera composta da 440 fotomoltiplicatori (*pixels*) a sezione esagonale, disposti in una matrice da 22 righe e 20 colonne. Due porte scorrevoli dette *shutters* proteggono il telescopio durante i periodi di luce chiudendosi automaticamente durante la presa dati in caso di vento o pioggia.

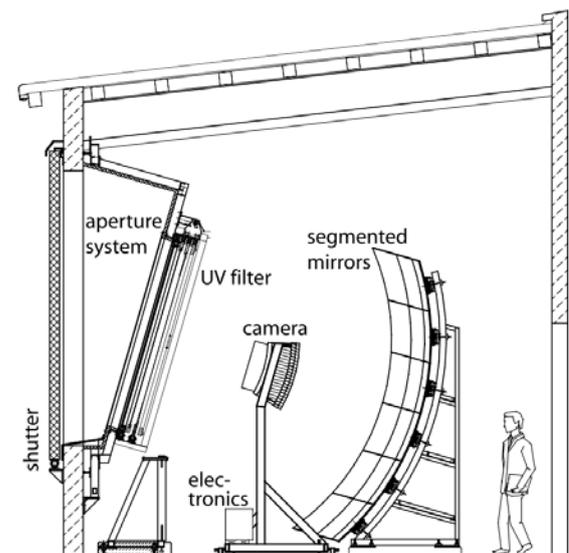


Figura 4: Vista schematica di un telescopio di fluorescenza e delle sue componenti principali.

I risultati dell'esperimento Auger

In più di 20 anni di attività l'Osservatorio Pierre Auger ha prodotto risultati di notevole rilevanza scientifica, che comprendono, tra gli altri, la misura del flusso dei raggi cosmici UHE con

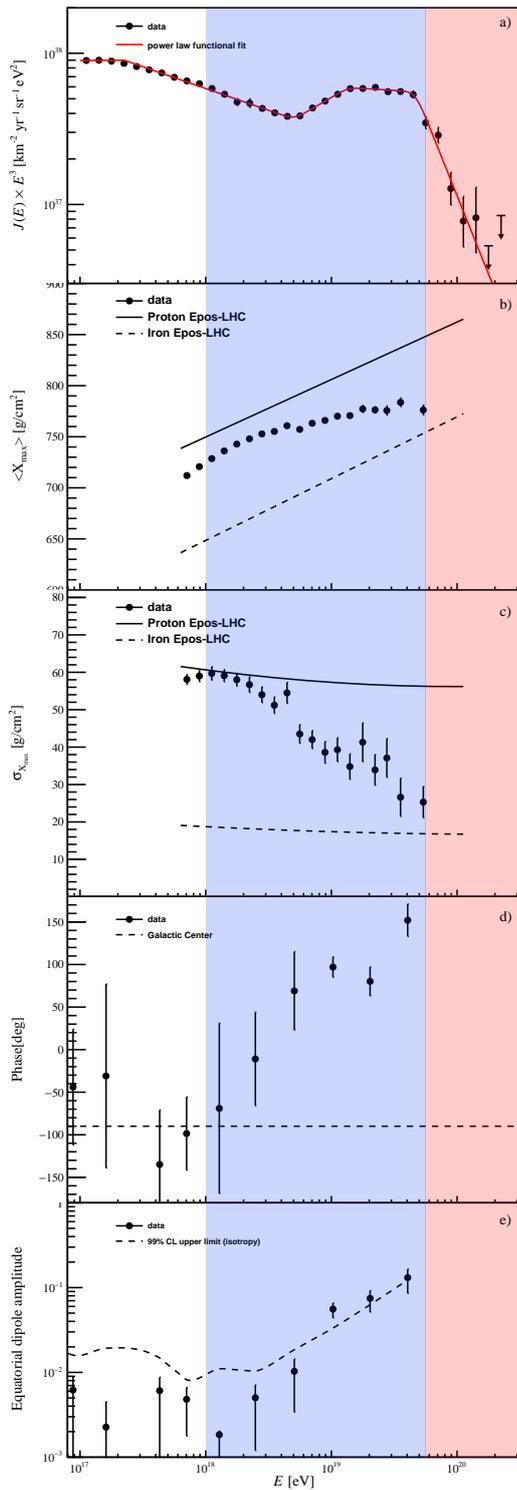


Figura 5: Sommario dei principali risultati dell'Osservatorio Pierre Auger. La banda blu corrisponde alla zona dell'*ankle* dove tutte le osservabili sono disponibili: spettro (a), composizione di massa (b)(c) e anisotropia su larga scala (d)(e). La banda rossa corrisponde alla zona della soppressione.

massa dei primari con l'energia e la distribuzione delle direzioni di provenienza degli stessi. In Fig. 5 è mostrato un sommario dei risultati di Auger. Diverse osservabili sono illustrate nella stessa figura in modo da mostrare la coerenza della loro evoluzione in energia. La banda blu corrisponde alla regione, intorno all'*ankle*, dove sono disponibili le diverse osservabili: lo spettro, la composizione di massa e l'anisotropia su larga scala. La banda rossa corrisponde alla zona della soppressione, dove mancano le misure relative alla massa dei primari, a causa del limitato tempo vivo² del rivelatore FD. Nella stessa regione anche le misure di anisotropia sono assenti poiché la statistica è insufficiente.

In Figura 5(a) è mostrato il flusso dei Raggi Cosmici al di sopra di 10^{17} eV, misurato dall'Osservatorio Pierre Auger utilizzando principalmente il rivelatore SD. Gli eventi, misurati simultaneamente dall'FD e l'SD e disponibili solo per il 10% del tempo di misura, vengono utilizzati per calibrare in modo calorimetrico l'estimatore dell'energia del primario fornito dall'SD. La distribuzione degli eventi rivelati viene poi corretta dagli effetti di rivelatore attraverso il calcolo dell'esposizione. Nella figura è mostrato il risultato del *fit* del flusso con un modello che consiste in una serie di 4 leggi di potenza, collegate tra loro da delle transizioni continue. Questo permette di individuare le diverse posizioni dell'*ankle*, a circa 5×10^{18} eV, e della soppressione a circa 4.6×10^{18} eV. Inoltre la statistica aggiornata al 2018 ha reso possibile individuare una nuova caratteristica spettrale, denominata *instep*, intorno all'energia di 14×10^{18} eV.

In Figura 5(b) e 5(c) rispettivamente, sono mostrati il valor medio $\langle X_{\max} \rangle$ e la dispersione $\sigma_{X_{\max}}$ della distribuzione del massimo dello sviluppo longitudinale degli sciame, misurato dall'FD, in funzione dell'energia. La profondità atmosferica alla quale gli sciame raggiungono il loro massimo sviluppo longitudinale è infatti un indicatore della massa del primario che li ha iniziati. Massimi più (meno) profondi e distribuzioni più (meno) larghe indicano primari più leggeri (pesanti). Il quadro è però complicato dalla possibile presenza di primari differenti al-

un dettaglio senza precedenti, l'evoluzione della

²Il sistema FD può operare solo in condizioni di sufficiente oscurità nel campo visivo. Questo corrisponde a un tempo vivo (o *duty cycle*) di circa il 10%.

la stessa energia, per cui la dispersione di X_{\max} dipende, in parte, anche dalla dispersione delle masse dei primari. Quello che si osserva nelle misure effettuate da Auger è un cambio di pendenza in corrispondenza dell'*ankle*, per $\langle X_{\max} \rangle$ e una decrescita della $\sigma_{X_{\max}}$. Questo fornisce una chiara indicazione di una progressione da masse leggere a masse intermedie, con limitata sovrapposizione di masse. Come già accennato, lo studio della composizione primaria, attraverso la misura diretta del massimo della profondità dello sviluppo dello sciame, non è possibile nella regione della soppressione a causa del bassissimo flusso e del tempo di presa dati limitato del rivelatore di fluorescenza. In particolare non è possibile determinare, in modo univoco, se esista una componente di protoni alle altissime energie. L'individuazione di una tale componente (anche intorno al 10%) sarebbe di particolare importanza, in quanto i protoni di alta energia permetterebbero di individuare i siti di produzione ed accelerazione dei raggi cosmici di altissima energia.

La distribuzione su larga scala delle direzioni di arrivo degli UHECR insieme allo spettro e alla composizione in massa rappresenta una osservabile molto importante per studiare la loro natura e origine. La regione che precede l'*ankle* è dove viene attesa la transizione da raggi cosmici galattici ad extra-galattici. A tal proposito sono di particolare rilevanza i risultati ottenuti dall'Osservatorio Pierre Auger per quanto concerne le anisotropie su larga scala. Osservando eventi fino a 80° in zenith è stato possibile determinare, con alta significatività statistica, una anisotropia nelle direzioni di arrivo degli UHECR per energie sopra $8 \cdot 10^{18}$ eV. Questo studio viene fatto attraverso una analisi di Rayleigh degli eventi SD, ossia effettuando uno sviluppo in serie di Fourier delle direzioni di arrivo di tali eventi, in coordinate equatoriali, e determinando così l'ampiezza e la fase della prima armonica. L'analisi armonica mostra che tale anisotropia ha le caratteristiche di un dipolo con una ampiezza di circa il 7% e direzione che punta 115° lontano dal Centro Galattico (Fig. 5(d)). Questo fatto insieme all'incremento dell'ampiezza del dipolo con l'energia (Fig. 5(e)) è una chiara indicazione dell'origine extra-galattica dei raggi cosmici di altissima energia.

Il ruolo della propagazione degli UHECR

La propagazione delle particelle e nuclei dalle sorgenti fino alla loro rivelazione a Terra avviene attraversando il fondo di fotoni cosmici. Le componenti rilevanti per gli UHECR sono il *Cosmic Microwave Background* (CMB) e l'*Extragalactic Background Light* (EBL) (Fig. 6).

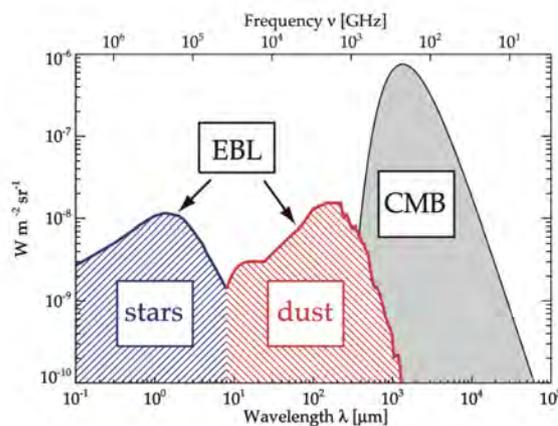
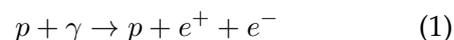


Figura 6: Schema dello spettro energetico dei fotoni di fondo dell'Universo [9]. Mostra i diversi range di lunghezza d'onda e intensità associate all'EBL e al CMB.

I processi più importanti che intervengono nella propagazione degli UHECR sono schematicamente:

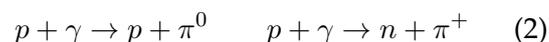
- per i protoni: la foto-produzione di coppie elettrone-positrone e di pioni;
- per i nuclei: la foto-disintegrazione.

La produzione di coppie



costituisce per il protone una causa di perdita di energia che può ragionevolmente essere trattata come continua ed è presente a partire da circa 10^{18} eV per interazioni con i fotoni del CMB.

La produzione di pioni



(in larga parte mediata da stati risonanti) determina invece una brusca attenuazione del flusso di protoni con conseguente soppressione, per energie intorno a $5 \cdot 10^{19}$ eV, del flusso di protoni

da sorgenti distanti. La soppressione del flusso di protoni prende il nome di *GZK-cutoff* da Greisen, Zatsepin e Kuzmin [10]. Il processo ha luogo anche per i nuclei, ma ad una energia di soglia A volte superiore e quindi non influenza lo spettro fino a 10^{20} eV. Per entrambi i processi precedenti il fondo rilevante è il CMB.

Berezinsky e collaboratori [11] hanno calcolato lo spettro UHE dei protoni per flussi di iniezione a legge di potenza $E^{-\gamma_g}$, come mostrato in Fig. 7. Il modello così ottenuto, noto come *dip model*,

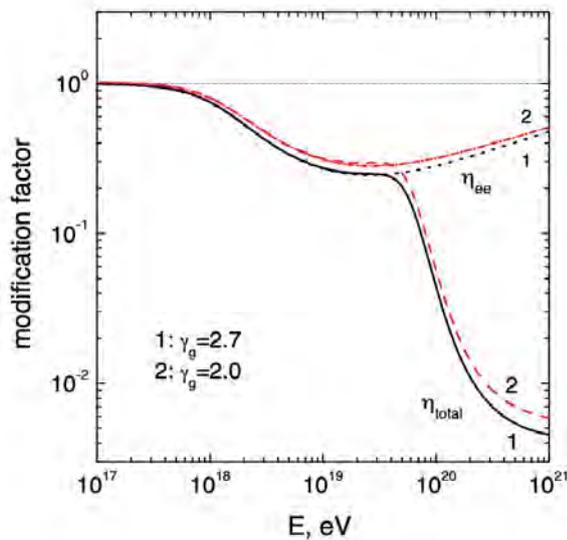
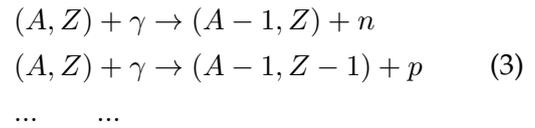


Figura 7: Rapporto tra lo spettro dei protoni, con le perdite di energia tenute in conto e senza perdite, calcolato da Berezinsky et al. [11] per iniezioni con indice spettrale da 2 a 2.7. Le curve identificate con η_{ee} e η_{total} si riferiscono, rispettivamente, alle perdite per produzione di coppie e totale.

mostra, oltre al *GZK-cutoff*, una variazione dell'indice spettrale tra 10^{18} e 10^{20} eV, che permette di interpretare l'*ankle* osservato nello spettro UHE come il *dip*, prodotto dalla produzione di coppie. Va osservato che il *dip model* è l'unico modello capace di interpretare lo spettro UHE con un'unica componente di protoni nell'intero range UHE. Però, essendo i processi coinvolti conosciuti molto bene, l'adattamento ai dati non permette molti gradi di libertà, oltre ad essere automaticamente escluso nel caso di osservazione di composizione mista.

Per quanto riguarda i nuclei dei raggi cosmici, il processo di foto-disintegrazione



risulta quello più importante e dominante rispetto ai processi già illustrati per i protoni. Questo processo causa la produzione di uno o più nucleoni e la disintegrazione in nuclei di massa inferiore. Nel sistema di riferimento a riposo del nucleo, il processo avviene nel range da 10 a 150 MeV ed è dominato alle energie più basse dalla risonanza gigante di dipolo elettrico. La Figura 8 mostra l'evoluzione con l'energia delle lunghezze di perdita di energia per foto-disintegrazione (in Mpc), per tutti i processi e per l'effetto combinato. Da questa si evidenzia che il processo di foto-disintegrazione su fotoni EBL (CMB) domina alle energie più basse (alte).

È importante notare che i tre processi (1,2,3) sono studiati attraverso gli esperimenti ad acceleratori con fasci di fotoni contro protoni/nuclei fermi. Quindi si tratta di fisica nota e modellizzabile.

In una visione semplificata (Fig. 9) si potrebbe dire che la propagazione degli UHECR nel mezzo cosmico ha lo stesso ruolo che, nella fisica agli acceleratori, svolge la simulazione della risposta dei rivelatori. In quest'ultimo caso, il tracciamento dei prodotti dell'interazione viene effettuato con programmi di simulazione dedicati (ad es. GEANT4 [12]), che descrivono il trasporto di tutte le particelle, primarie e secondarie, nell'apparato sperimentale, considerando tutte le possibili interazioni delle particelle con i materiali del rivelatore.

Nel caso degli UHECR, sono stati sviluppati codici di propagazione, che utilizzano la simulazione Monte Carlo, come ad es. CRPropa e SimProp³. Questi sono in grado di connettere gli scenari astrofisici (caratterizzati dai flussi emessi dalle sorgenti, loro distribuzione, etc.) alle osservabili alla sommità dell'atmosfera terrestre. In essi, il mezzo è costituito dal fondo di radiazione: il CMB è noto in tutti i suoi aspetti, inclusa la sua evoluzione cosmologica. Il fondo EBL è invece

³Per un confronto tra i due codici di propagazione si veda la referenza [13]

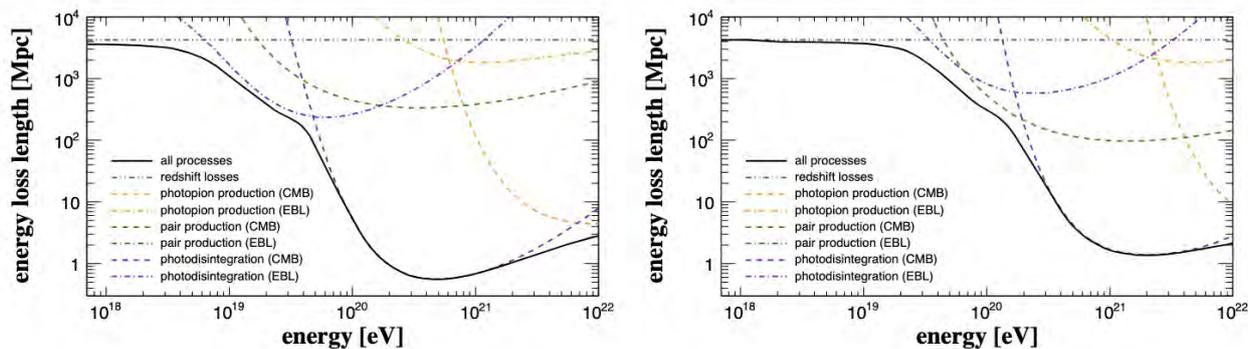
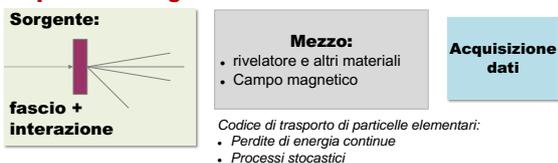


Figura 8: Lunghezze di perdita di energia ($= E/(dE/dx)$) per ^{14}N (sinistra) e ^{56}Fe (destra). I contributi dei vari processi sui fondi di fotoni è mostrato con diversi tratteggi. La linea continua rappresenta l'effetto combinato.

Esperimento agli acceleratori



Esperimento UHECR



Figura 9: Confronto tra gli schemi di simulazione in esperimenti ad acceleratori e con UHECR.

meno noto ed è descritto attraverso parametrizzazioni disponibili nella letteratura⁴. Alle perdite di energia conseguenti dai processi (1,2,3), che costituiscono la parte più importante, vanno aggiunte le così dette perdite adiabatiche, dovute all'espansione dell'Universo. Queste si sanno trattare con esattezza. Nel trattamento della propagazione degli UHECR, ci sono sicuramente altre incognite rilevanti, come ad esempio i campi magnetici extra-galattici, poco noti sia per il loro valore sia per come sono distribuiti. Per quanto una propagazione tridimensionale già trovi alcune applicazioni, solitamente si usa una propagazione unidimensionale con eventuali correzioni per effetti magnetici.

⁴I fotoni EBL hanno energie maggiori rispetto al CMB, ma densità nettamente inferiori. Essi sono comunque rilevanti nel processo di foto-disintegrazione a bassa energia (Fig. 8)

Un altro passo rilevante della simulazione riguarda lo sviluppo delle cascate (*Extensive Air Showers*), conseguenti all'interazione degli UHECR con l'atmosfera terrestre. In questo caso ci si avvale di codici di cascata (CORSIKA [17], CONEX [18], ...), a loro volta basati su modelli di interazione adronica, al fine di consentire l'estrapolazione dalla fenomenologia nota agli acceleratori fino alle energie più elevate, che si presentano specialmente nelle fasi iniziali degli sciame. L'adattamento di alcuni di questi modelli adronici (Sibyll, EPOS, QGSJet-II) ai dati ottenuti al Large Hadron Collider (LHC) ha sensibilmente ridotto le loro incertezze, anche se l'estrapolazione alle energie degli UHECR è di oltre un ordine di grandezza, nel sistema del centro di massa, rispetto alle misure di LHC.

È evidente che questo è solo uno schema semplificato, perché ci sono anche sostanziali differenze tra la simulazione degli esperimenti agli acceleratori e quella per gli esperimenti con i raggi cosmici. Prima di tutto, c'è differenza tra gli scopi degli esperimenti. Nella Fig. 9 questa differenza è indicata dal punto interrogativo nel primo riquadro in basso. Infatti lo scopo ultimo degli esperimenti con UHECR è la ricerca delle sorgenti dei raggi cosmici. In altre parole, rispetto agli acceleratori, viene sostituita una sorgente delle interazioni ben controllata (proiettile e bersaglio entrambi definiti), con una distribuita e con proprietà note solo dalle osservabili a Terra.

Inoltre è anche molto diverso il livello di conoscenza dei processi che avvengono nei mezzi. Infatti i processi relativi alla propagazione delle particelle nei rivelatori, nel caso degli accelerato-

La soppressione dello spettro degli UHECR

La soppressione (*cutoff*) osservata nello spettro energetico dei raggi cosmici tra 10^{19} e 10^{20} eV è sicuramente la caratteristica più significativa di tutto lo spettro. Tuttavia la sua possibile interpretazione non è unica. In particolare essa può essere causata sia da effetti di propagazione sia dall'energia massima associata al processo di accelerazione dei raggi cosmici alla sorgente.

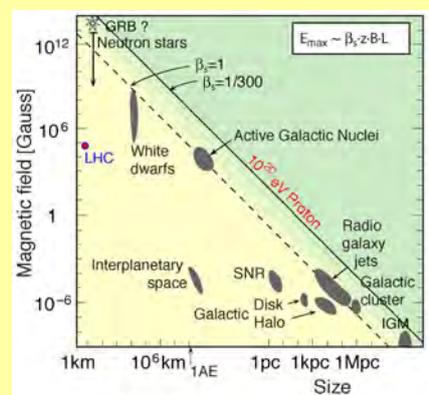
1. Propagazione. I protoni, interagendo con i fotoni del CMB, perdono parte della loro energia nella produzione di pioni, per energie superiori a

$$E_{th}(\theta) = \frac{m_{\pi}^2 + 2m_p m_{\pi}}{2\epsilon_{\gamma}(1 - \cos\theta)}.$$

Per $\epsilon_{\gamma} \approx 1$ meV (energia media del CMB) e collisione frontale (*head-on*), si ha l'energia di soglia minima $\approx 7 \cdot 10^{19}$ eV. Greisen e, indipendentemente, Zatsepin e Kuzmin [10] ipotizzarono l'esistenza di questo *cutoff* nel flusso dei raggi cosmici, subito dopo la scoperta, nel 1965, del fondo di radiazione cosmica da Penzias e Wilson.

Per i nuclei, l'interazione col fondo di fotoni avviene prevalentemente attraverso la foto-disintegrazione. Il processo ha luogo per energie di soglia dipendenti dal nucleo ($E_{th} \approx M\Delta B/(2\epsilon_{\gamma})$, con M massa del nucleo e ΔB differenza di energia di legame tra nuclei padre-figlio), attraverso l'eccitazione della Risonanza Gigante di Dipolo. Il fondo EBL è quello che determina l'attenuazione del flusso negli UHECR (vedi fig. 8), mentre le interazioni su CMB sono efficaci a energia più alte. Gerasimova e Rozental [15] per primi, nel 1961, hanno considerato l'effetto della foto-disintegrazione sui raggi cosmici, pur riferendosi a raggi cosmici galattici e alla componente ottica del fondo EBL. Nei già citati lavori di Greisen e Zatsepin, Kuzmin viene accennato al caso dei nuclei e all'eventuale loro attenuazione per foto-disintegrazione. Per questo motivo nella letteratura attuale, il nome *GZK-cutoff* viene usato, in senso esteso, come soppressione dello spettro per effetti di propagazione.

2. Energia massima di accelerazione. Nelle SNR, il meccanismo di *shock-acceleration* [1] raggiunge un'energia massima $E_{max} \approx \beta_s ZeBR$, dove B e R sono il campo magnetico e la dimensione del sito di accelerazione, e β_s è la velocità dello shock. Hillas ha proposto un simile meccanismo, di origine elettromagnetica, per tutte le sorgenti astrofisiche [16]. Nell'*Hillas plot* qui accanto sono rappresentati i campi magnetici e le dimensioni tipiche di tutte le sorgenti. Dalla figura risulta evidente che nessuna sorgente è capace di accelerare i raggi cosmici oltre circa 10^{20} eV.



ri sono per lo più elettromagnetici (e quindi ben noti). Nel mezzo cosmico invece, mentre la foto-produzione di coppie e di pioni è nota con buon dettaglio, la foto-disintegrazione (3) si manifesta nella propagazione con un'ampia varietà di modi e i casi per cui esistono misure sperimentali sono relativamente pochi. I nuclei potenzialmente coinvolti nella propagazione degli UHECR sono quelli della carta dei nuclei dal ferro all'idrogeno. La situazione sperimentale sulle sezioni d'urto

di foto-disintegrazione è mostrata in Fig. 10 [14]. Da questa figura si evince che le lunghezze di interazione associate ai vari nuclei e le probabilità di decadimento dei possibili modi di disintegrazione sono misurati per una ristretta minoranza di isotopi, lungo la linea di stabilità. Per il resto, tutto è affidato alla validità di modelli teorici.

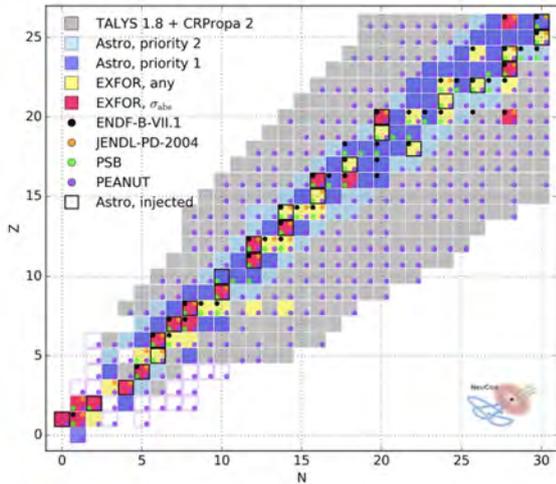


Figura 10: I nuclei di interesse per la propagazione degli UHECR sono indicati in grigio. I nuclei per i quali sono disponibili dati sperimentali (dal database EXFOR) sono in rosso, quando la sezione d'urto totale di fotoassorbimento è nota, oppure in giallo quando è stata misurata una specifica modalità di frammentazione, ad esempio la fotoemissione di un neutrone. Per maggiori dettagli si veda la Ref. [14].

Scenari astrofisici

Facendo uso dei codici di propagazione è possibile calcolare le osservabili a Terra, una volta scelto uno scenario astrofisico. Questi scenari, nel caso più semplice, sono basati sull'ipotesi di sorgenti distribuite uniformemente (*comoving volume*) su larga scala. Le due classi principali sono lo scenario dominato dai protoni e quello con composizione mista.

Protoni

Nello scenario in cui le sorgenti emettono principalmente protoni, il risultato degli effetti di propagazione porta al così detto *dip model* [11]. Poiché i processi dominanti sono quelli indicati da (1) e (2), i protoni emessi dalle sorgenti mantengono la loro identità, ma il loro spettro è modificato. Partendo, ad esempio, da uno spettro di iniezione $\propto E^{-2.7}$, lo spettro rivelato è mostrato in Fig. 11.

La produzione di coppie genera il *dip*, cioè l'ampia depressione centrata intorno a circa $10^{18.7}$ eV, mentre la produzione di pioni genera il *GZK-cut-off*. Come si vede dalla Fig. 11, lo scenario di protoni dà una descrizione di entram-

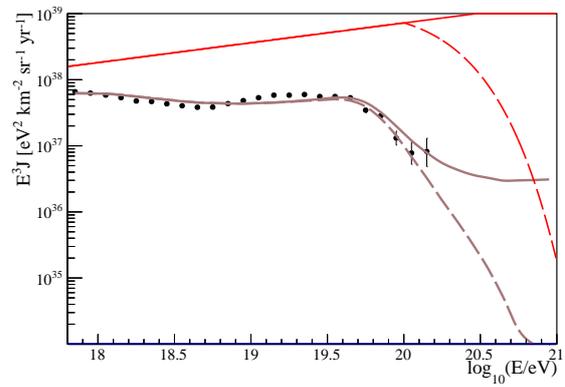


Figura 11: Spettro energetico osservato (linea marrone), e quello per sorgenti di protoni $\propto E^{-2.7}$ (linea rossa) uniformemente distribuite. Le linee tratteggiate si riferiscono al caso in cui l'energia massima dei protoni è di 10^{20} eV. I punti sperimentali si riferiscono allo spettro di Auger e sono mostrati solo per confronto.

be le caratteristiche dello spettro UHE. La forma dello spettro osservato può in parte cambiare, se l'energia massima delle sorgenti è dello stesso ordine di grandezza del *GZK-cut-off* (linee tratteggiate) oppure se si considerano evoluzioni cosmologiche delle sorgenti diverse da quella uniforme.

Confrontando con i dati sperimentali, questo accordo si realizza meglio per i dati di Telescope Array, ma meno per quelli di Auger, principalmente per la maggiore precisione sperimentale di quest'ultimo e la conseguente migliore determinazione delle caratteristiche spettrali. Tuttavia l'incompatibilità con l'esperimento Auger viene prevalentemente dalla composizione di massa, in quanto, a parte le energie più basse, la deviazione di $\langle X_{\max} \rangle$ da una composizione puramente protonica risulta piuttosto evidente, indipendentemente dal modello adronico considerato.

Composizione mista

Questo scenario ha interesse per l'interpretazione dei dati di Auger, a causa della composizione di massa dai dati di X_{\max} .

Infatti una composizione che include diverse masse, come nella Fig. 5 per i dati di Auger, richiama necessariamente ad uno scenario in cui le sorgenti emettono protoni e nuclei secondo una composizione mista.

Come abbiamo visto, nel caso dei nuclei, il processo più importante è la foto-disintegrazione (3). Nel regime ultra-relativistico in cui si trovano i nuclei degli UHECR, le loro energie di legame si possono trascurare e la conservazione del quadrimpulso nella reazione (3) si traduce quindi nella conservazione del fattore di Lorentz, $\Gamma = E/A$. Questo implica che, ad ogni disintegrazione di un nucleo di massa A e energia E , il nucleo uscente ha energia $E(A - 1)/A$ (per disintegrazione singola) ed è accompagnato dall'emissione di (almeno) un nucleone⁵ di energia E/A . Pertanto, nella loro propagazione, i nuclei degradano in massa ed energia e, inoltre, producono un flusso secondario di protoni ad energia $1/A$ rispetto al primario. Questo meccanismo è tanto più efficace quanto maggiore è la distanza della sorgente, per cui, per le massime distanze, il nucleo si disintegra in A protoni di energia circa E/A .

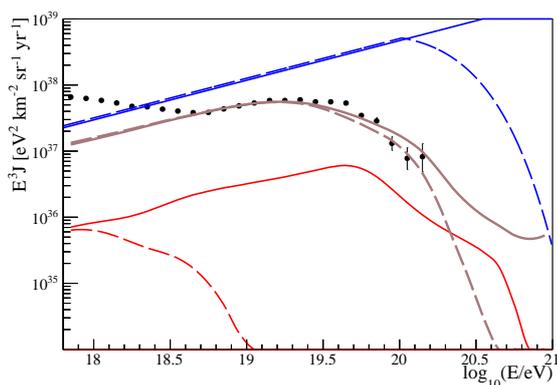


Figura 12: Spettro *all-particle* osservato (linea marrone), per sorgenti di nuclei Fe (linea blu) uniformemente distribuite con spettro d'iniezione $\propto E^{-2.4}$. Le linee rosse mostrano gli spettri di protoni secondari generati dalla foto-disintegrazione. Le linee tratteggiate si riferiscono al caso di energia massima dei nuclei di 10^{20} eV.

Per mostrare le conseguenze di queste considerazioni, esaminiamo, per semplicità, il caso in cui le sorgenti emettano un solo nucleo, il ferro. Nella Figura 12, ottenuta con il codice *SimProp* [19], vengono mostrati, per uno spettro di iniezione $\propto E^{-2.4}$ (blue), lo spettro *all-particle* osservato (marrone) e la sola componente dei protoni se-

condari (rosso)⁶. In pratica, nonostante siano emessi solo nuclei Fe, si osserva che:

1. lo spettro *all-particle* osservato è modificato;
2. lo spettro viene soppresso ad iniziare da energie $> 10^{19.3}$ eV circa
3. lo spettro riproduce ragionevolmente i dati sopra l'*ankle*, ma lo sottostima ad energie inferiori;
4. la composizione è mista: in particolare, sono presenti protoni secondari.

Il secondo punto (soppressione dello spettro) evidenzia che anche la foto-disintegrazione produce un *cutoff* a energie prossime al *GZK-cutoff* dei protoni.

La presenza di protoni secondari è la caratteristica più rilevante della propagazione dei nuclei, ma, nel caso in cui non esista un'energia massima di accelerazione (linee continue), la relazione della loro energia con quella primaria è, in parte, nascosta dagli effetti di soppressione (dei protoni e dei nuclei). Invece, se esiste un'energia massima E_{\max} , per i nuclei di massa A , dalla conservazione del fattore di Lorentz consegue che i protoni secondari non potranno raggiungere energie superiori a E_{\max}/A . Questo fatto risulta evidente dalle linee tratteggiate in Fig. 12, corrispondenti al caso in cui l'iniezione dei nuclei di Fe ha una soppressione esponenziale a partire da 10^{20} eV.

La composizione di Auger, come rappresentata in Fig. 5, mostra per $E \gtrsim 10^{18.3}$ un aumento della massa ($\langle X_{\max} \rangle$) e una diminuzione delle fluttuazioni ($\sigma_{X_{\max}}$). Poichè ad una composizione mista corrisponde una dispersione maggiore di quelle relative alle singole masse, il quadro complessivo che emerge dalla composizione di Auger è quello di masse crescenti con mescolamento limitato. Tra tutti i nuclei, i protoni, avendo le fluttuazioni maggiori, debbono necessariamente essere soppressi al crescere dell'energia.

Questa descrizione, sicuramente semplificata, porta alla necessità di un meccanismo di

⁵Il nucleone, non importa se si tratti di protone o neutrone, verrà rivelato come protone. Infatti, su questa scala, decade pressoché istantaneamente in protone.

⁶In questa descrizione, ci limitiamo a considerare tra i secondari solo i protoni, ma anche altri nuclei partecipano alla composizione di massa

soppressione dei protoni secondari di foto-disintegrazione. Questo si realizza, ad esempio, con un meccanismo di accelerazione limitato da una rigidità magnetica massima (vedi il riquadro).

Pertanto gli scenari a composizione mista, cioè con emissione dalle sorgenti di nuclei accelerati fino a una rigidità massima $R_{\max} = E_{\max}/(Ze)$, sono quelli che meglio si adattano ai dati di Auger nel loro complesso.

Per trovare qual è lo scenario che meglio si adatta ai dati di Auger, si ricorre ad un programma di *fit* combinato dello spettro e della composizione, in cui si ipotizza che lo spettro di iniezione segua una legge di potenza con rigidità massima e sia composto da cinque masse elementari (H, He, N, Si e Fe), ciascuna con un peso relativo. I valori spettrali e i pesi relativi costituiscono i parametri del *fit*. Per ogni insieme dei valori dei parametri, gli spettri delle singole componenti vengono propagate dal codice e le osservabili a Terra (spettro di energia e distribuzioni di X_{\max}) sono confrontate con quelle sperimentali, con l'uso di una funzione di verosimiglianza, *Likelihood*. In Fig. 13 sono mostrati i risultati del *best fit*, pubblicati in [20], limitatamente alla regione di energia sopra l'*ankle*.

Direzioni future

Come abbiamo visto precedentemente, lo studio combinato dello spettro e della composizione degli UHECR permette di individuare le caratteristiche delle sorgenti da cui originano. Però, diversamente dal caso dello scenario dominato dai protoni, l'adattamento ai dati degli scenari a composizione mista è soddisfacente soltanto al di sopra dell'*ankle*. Questo appare evidente in Fig. 13, se si osservano le linee tratteggiate, escluse dal *fit*. Infatti, la propagazione dei nuclei non riesce a produrre un *dip*, come avviene per i protoni, perché questa caratteristica, generata dalla produzione di coppie, risulta sub-dominante per ogni specie nucleare.

Pertanto, l'estensione del *fit* a energie inferiori richiede che l'*ankle* sia generato da una componente aggiuntiva. Una componente di origine Galattica è poco supportata dai limiti di accelerazione della nostra Galassia (paradigma SNR) ed inoltre produrrebbe un livello di anisotropia

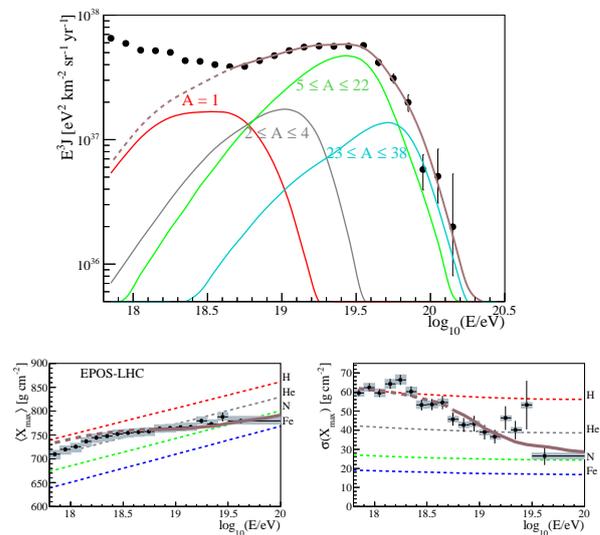


Figura 13: Alto: Spettro di energia ottenuto dal fit combinato. Gli spettri parziali e quello totale sono mostrati con colori differenti. Basso: Valor medio e deviazione standard di X_{\max} per i dati e per la predizione del fit (marrone). Il tratteggio si riferisce alla regione di energia non inclusa nel fit. Per tutte le notazioni e riferimenti vedi [20].

su larga scala che non si osserva nei dati. Quindi una soluzione è quella di ipotizzare un'altra popolazione di sorgenti extra-galattiche, il cui flusso, sovrapponendosi a quello ottenuto dal *fit* in Fig. 13, produca l'*ankle* e il conseguente cambio dell'indice spettrale osservato nei dati.

In alternativa a questa ipotesi, sono stati proposti in alcuni lavori dei meccanismi in cui l'*ankle* è prodotto dalle interazioni dei raggi cosmici con l'ambiente che circonda le sorgenti. Infatti, gran parte delle sorgenti astrofisiche ipotizzate come possibili candidate a sorgenti di UHECR, sono circondate da intensi campi magnetici, fotoni e polveri. In questo ambiente sono possibili sia processi di foto-disintegrazione che di spallazione dei nuclei. Tra i prodotti di questi processi i neutroni, non essendo intrappolati dai campi magnetici, possono uscire liberamente, causando un'emissione di protoni con diverso indice spettrale, che, per opportuni valori dei parametri ambientali, può generare lo spettro e la composizione osservata intorno all'*ankle*.



- [1] T. K. Gaisser, R. Engel, E. Resconi: *Cosmic Rays and Particle Physics: 2nd Edition*, Cambridge University Press, Cambridge (2016).
- [2] V. L. Ginzburg, S. I. Syrovatskii: *The Origin of Cosmic Rays*, Pergamon Press, Oxford (1964).
- [3] J. Abraham *et al.* [Pierre Auger]: *Correlation of the highest-energy cosmic rays with the positions of nearby active galactic nuclei*, *Astropart. Phys.*, 29 (2008) 188. [erratum: *Astropart. Phys.* 30 (2008), 45]
- [4] R. Aloisio, P. Blasi, I. De Mitri and S. Petrer; *Selected Topics in Cosmic Ray Physics* doi:10.1007/978-3-319-65425-6_1 [arXiv:1707.06147 [astro-ph.HE]].
- [5] T. Abu-Zayyad *et al.* [Telescope Array]: *The surface detector array of the Telescope Array experiment*, *Nucl. Instrum. Meth. A*, 689 (2013) 87.
- [6] M. Teshima: *The highest energy cosmic rays observed by AGASA*, *AIP Conference Proceedings*, 516 (2000) 365.
- [7] P. Sokolsky for the HiRes Collaboration: *Final Results from the High resolution Fly's Eye (HiRes) Experiment*, *Nuclear Physics B - Proceedings Supplements*, 212-213 (2011) 74.
- [8] The Pierre Auger Collaboration: *The Pierre Auger Cosmic Ray Observatory*, *Nuclear Instruments and Methods in Physics Research A*, 798 (2015) 172.
- [9] D. Mazin: *Constraints on Extragalactic Background Light from Čerenkov telescopes: status and perspectives for the next 5 years*, *AIP Conf. Proc.*, 1112 (2009) 111..
- [10] K. Greisen: *End to the Cosmic-Ray Spectrum?*, *Phys. Rev. Lett.* 16 (1966) 748.
G. T. Zatsepin, V.A. Kuz'min: *Upper Limit of the Spectrum of Cosmic Rays*, *Sov. Phys. JETP Lett. (Engl. Transl.)*, 4 (1966) 78.
- [11] V. Berezhinsky, A. Z. Gazizov and S. I. Grigorieva: *On astrophysical solution to ultrahigh-energy cosmic rays*, *Phys. Rev. D*, 74 (2006) 043005.. doi:10.1103/PhysRevD.74.043005.
- [12] J. Allison *et al.*: *Geant4 developments and applications*, *IEEE Transactions on Nuclear Science*, 53 (2006) 270.
- [13] R. Alves Batista, A. Dundovic, M. Erdmann, K. H. Kampert, D. Kuempel, G. Müller, G. Sigl, A. van Vliet, D. Walz and T. Winchen: *CRPropa 3 - a Public Astrophysical Simulation Framework for Propagating Extraterrestrial Ultra-High Energy Particles*, *JCAP*, 5 (2016) 38.
- [14] D. Boncioli, A. Fedynitch and W. Winter: *Nuclear Physics Meets the Sources of the Ultra-High Energy Cosmic Rays*, *Sci. Rep.*, 7 (2017) 4882.
- [15] N. M. Gerasimova and I. L. Rozental: *Influence of the Nuclear Photoeffect on the Cosmic Ray Spectrum*, *Sov. Phys. JETP (Engl. Transl.)*, 14 (1962) 350.
- [16] A. M. Hillas: *The Origin of Ultrahigh-Energy Cosmic Rays*, *Ann. Rev. Astron. Astrophys.*, 22 (1984) 425.
- [17] D. Heck, J. Knapp, J. N. Capdevielle, G. Schatz and T. Thouw: *CORSIKA: A Monte Carlo code to simulate extensive air showers*, FZKA-6019.
- [18] T. Pierog, M. Alekseeva, T. Bergmann, V. Chernatkin, R. Engel *et al.*: *First results of fast one-dimensional hybrid simulation of EAS using CONEX*, *Nucl. Phys. Proc. Suppl.*, 151 (2006) 159.
- [19] R. Aloisio, D. Boncioli, A. Di Matteo, A. F. Grillo, S. Petrer, F. Salamida: *SimProp v2r4: Monte Carlo simulation code for UHECR propagation*, *JCAP*, 11 (2017) 9.
- [20] A. Aab *et al.* [Pierre Auger]: *Combined fit of spectrum and composition data as measured by the Pierre Auger Observatory*, *JCAP*, 4 (2017) 38. [erratum: *JCAP* 03 201, E02]

Sergio Petrer: già Professore Ordinario di Fisica Generale presso l'Università di Lecce e poi di Fisica Nucleare e Subnucleare presso l'Università dell'Aquila, dove ha ricoperto diverse cariche. Attualmente è Senior Professor al Gran Sasso Science Institute a L'Aquila. Fino al 1985 ha svolto principalmente esperimenti di fisica delle particelle ad acceleratori, al CERN e al Fermilab, con ricerche di quark pesanti prodotti da adroni e neutrini. Successivamente ha indirizzato i suoi interessi alla fisica astroparticellare con l'esperimento MACRO ai Laboratori Nazionali del Gran Sasso, indirizzato alla ricerca di monopoli magnetici, oscillazioni di neutrini atmosferici e bursts di neutrini da collapsi stellari. A partire dal 2000 ha indirizzato le sue ricerche allo studio dei raggi cosmici di energia ultra-elevata presso l'Osservatorio Pierre Auger in Argentina.

Francesco Salamida: ha conseguito la Laurea in Fisica nel 2003 presso l'Università degli studi di Lecce e il Dottorato in Fisica nel 2007 presso l'Università dell'Aquila. Dopo Post-Doc all'Aquila, al Karlsruhe Institute of Technology, all'Institut de Physique Nucléaire d'Orsay e all'INFN di Milano Bicocca, dal 2019 è Professore Associato presso il Dipartimento di Scienze Fisiche e Chimiche dell'Università dell'Aquila dove tiene vari corsi tra i quali quello di Fisica Nucleare e Subnucleare nel Corso di Laurea Magistrale in Fisica. La sua ricerca si concentra sulla fisica sperimentale dei raggi cosmici ad altissima energia nell'ambito dell'Osservatorio Pierre Auger in Argentina e sulla ricerca del doppio decadimento beta senza neutrini come membro dell'esperimento GERDA/LEGEND presso i Laboratori Nazionali del Gran Sasso.

Come sono accelerati i raggi cosmici?

Those who became interested in cosmic rays tended to be rugged individualists, to be iconoclastic, and to march to the drummer in their own heads rather than some distant one.

V. Fitch

Stefano Gabici *Université de Paris, CNRS, Astroparticule et Cosmologie, F-75013 Paris, France*

L'atmosfera terrestre è costantemente bombardata da un flusso di particelle di alta energia (in stragrande maggioranza nuclei atomici) provenienti dal cosmo: i raggi cosmici. L'energia media di queste particelle è diverse migliaia di volte inferiore a quelle riproducibili negli acceleratori del CERN, mentre quella massima è circa cento miliardi di volte quella media. Questo implica che nell'Universo debbano operare dei meccanismi di accelerazione di particelle estremi, la cui natura è tuttora dibattuta.

Tipi bizzarri?

In un bell'articolo apparso nel 1999 su *Reviews of Modern Physics* [1], il premio Nobel Val Logsdon Fitch ripercorre gli anni durante i quali la fisica delle particelle maturò fino a diventare una disciplina autonoma. Fino agli anni cinquanta, infatti, l'unico strumento a disposizione dei fisici per studiare il mondo delle particelle elementari era l'osservazione delle interazioni dei raggi cosmici nell'atmosfera terrestre. Quando gli acceleratori di particelle riuscirono a eguagliare l'e-

nergia (media) delle particelle che compongono la radiazione cosmica, divenne possibile creare in laboratorio collisioni tra particelle di alta energia. L'avvento degli acceleratori fu accompagnato da un grande (e comprensibile) entusiasmo all'interno della comunità scientifica, e la maggior parte dei fisici spostarono la loro attenzione dai raggi cosmici ai laboratori, dove gli scontri tra particelle, col passare degli anni, erano caratterizzati da energie sempre più elevate.

La minoranza che continuò, testardamente, a studiare i raggi cosmici era formata, secondo Fitch, da "burberi individualisti iconoclasti". Gente che remava controcorrente. Tipi bizzarri?

John Linsley: fino a che energie si estende lo spettro dei raggi cosmici?

Uno degli ultimi fisici ad impersonificare piuttosto fedelmente la definizione di Fitch fu lo statunitense John Linsley che, a cavallo tra gli anni cinquanta e sessanta, costruì insieme al collega italiano Livio Scarsi, e poi operò in quasi totale solitudine, il più grande rivelatore di raggi cosmici dell'epoca. A Volcano Ranch, nel deserto

del New Mexico, Linsley e Scarsi installarono 19 rivelatori di particelle sparsi su una superficie di circa 8 km^2 .

Il principio del funzionamento di questo tipo di rivelatori è basato sul fatto che quando un raggio cosmico impatta su una molecola dell'atmosfera terrestre, una parte dell'energia della collisione può essere convertita in particelle di più bassa energia. Queste particelle, dette **secondarie**, possono a loro volta interagire nell'atmosfera generando altre particelle, in un processo a catena che porta alla formazione di uno sciame atmosferico formato da un numero progressivamente crescente di particelle. Quando l'energia delle particelle dello sciame scende sotto una certa soglia, il processo di moltiplicazione si arresta (l'energia delle collisioni non è più sufficiente a creare nuove particelle). Le particelle iniziano quindi a perdere gradualmente la loro energia negli urti con le molecole dell'atmosfera e lo sciame, gradualmente, si estingue.

Se il raggio cosmico primario è di energia sufficientemente elevata, lo sciame che genera può raggiungere il suolo prima di estinguersi nell'atmosfera, investendo un'ampia superficie di terreno. Quanto più grande è l'energia del raggio cosmico primario, tanto più numerose sono le particelle nello sciame, e tanto maggiore è l'area di terreno investita. Disseminando un gran numero di rivelatori di particelle nel deserto del New Mexico, Linsley voleva spingere le osservazioni dei raggi cosmici a energie mai esplorate prima (Fig. 1). L'impresa non era semplice, perché era già noto all'epoca che il numero di raggi cosmici che piovono sulla terra diminuisce rapidamente al crescere dell'energia.

Il 22 febbraio 1962, la perseveranza e la tenacia di Linsley vennero premiate. I rivelatori di Volcano Ranch misurarono uno sciame contenente circa 30 miliardi di particelle! Niente del genere era mai stato osservato, e l'energia stimata per il raggio cosmico primario era davvero sensazionale: l'equivalente dell'energia di una palla da baseball lanciata a 50 km/h ($\approx 10^{20} \text{ eV}$, per usare un'unità di misura in voga nella comunità scientifica dei raggi cosmici). Questa energia è decine di miliardi di volte più alta dell'energia del raggio cosmico medio che bombarda l'atmosfera terrestre. Gli studi di Linsley e di altri studiosi che hanno continuato per la sua

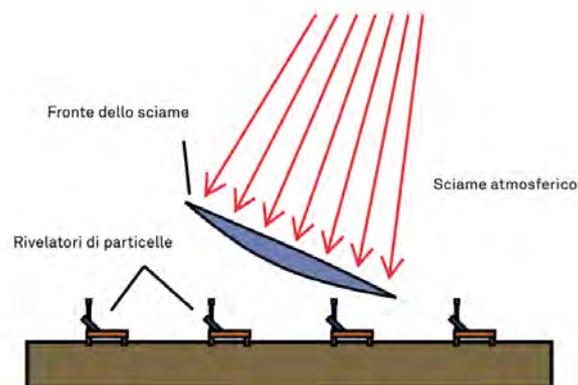


Figura 1: Una schiera di rivelatori di particelle permette di registrare l'arrivo al suolo di uno sciame atmosferico (Wikimedia Commons).

strada nei decenni seguenti mostrano che, mentre i raggi cosmici di energia media bombardano l'atmosfera terrestre a un ritmo di migliaia di particelle per metro quadro per secondo, i raggi cosmici di energia massima sono rarissimi: uno per kilometro quadro per secolo (Fig. 2). La loro esistenza dimostra che nell'Universo debbano esistere dei veri e propri mostri capaci di accelerare minuscoli nuclei atomici fino ad energie esorbitanti.

Chi volesse approfondire le tecniche di rivelazione dei raggi cosmici può far riferimento al classico testo di Bruno Rossi [2]. Resoconti sugli sviluppi degli studi sui raggi cosmici fino a tempi recenti possono essere trovati nelle referenze [3, 4], mentre Linsley ha raccontato la sua versione dei fatti nella referenza [5].

Michael Hillas: chi accelera i raggi cosmici?

Chi accelera, quindi, le particelle più energetiche mai osservate nell'Universo? Come spesso accade nella scienza, le risposte migliori a questioni estremamente complesse si ottengono grazie a brutali semplificazioni. Le semplificazioni, va da sé, devono essere non solo brutali, ma anche intelligenti. Spetta al fiuto del buon fisico individuare il minor numero possibile di elementi che sono essenziali per la risoluzione del problema. La questione è delicata, perché se si semplifica troppo poco il problema resta spesso troppo complesso, mentre se si semplifica troppo la risposta

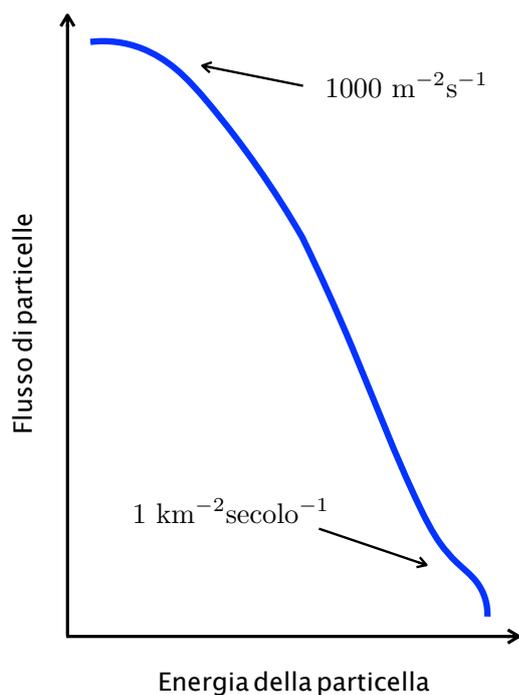


Figura 2: Il flusso di particelle cosmiche che impattano l'atmosfera terrestre decresce rapidamente all'aumentare dell'energia delle particelle.

che si ottiene è quasi sempre sbagliata. Fu il fisico inglese Michael Hillas a compiere un piccolo capolavoro, e a proporre alla comunità scientifica un criterio semplicissimo ma anche efficacissimo per decidere se un dato oggetto astrofisico possa, potenzialmente, accelerare i raggi cosmici di una certa energia. Questa procedura, ben presto battezzata con il nome di "criterio di Hillas", fu presentata in un famosissimo articolo del 1984 ed è tuttora utilizzata da chiunque voglia studiare l'origine dei raggi cosmici [6].

Hillas decise di non affrontare di petto il problema dell'accelerazione dei raggi cosmici. In un primo momento, stabilì di non interessarsi a come i raggi cosmici vengano accelerati, ma a chi li acceleri. Il suo ragionamento partì dalla considerazione che i campi magnetici sono onnipresenti nell'Universo. Inoltre, i raggi cosmici sono particelle elettricamente cariche (nuclei atomici), e saranno quindi deflessi da questi campi magnetici. Qualsiasi tentativo di comprendere le sorgenti dei raggi cosmici deve quindi tenere in conto questo fatto. Supponiamo quindi di avere una potenziale sorgente di raggi cosmici, all'in-

terno della quale vi sia un campo magnetico di una certa intensità. Supponiamo inoltre, che la sorgente abbia accelerato un raggio cosmico fino a una certa energia. Per semplificare il problema, supponiamo che il raggio cosmico si muova in una direzione ortogonale (a 90 gradi) a quella delle linee di forza del campo magnetico. In questo caso la traiettoria della particella sarà perfettamente circolare. Il raggio di questa traiettoria sarà più grande per particelle di energia maggiore, e più piccolo per campi magnetici più intensi. L'intuizione di Hillas fu quella di confrontare il raggio di questa traiettoria con la dimensione della sorgente astrofisica. Se il raggio della traiettoria è molto maggiore della dimensione della sorgente, la particella potrà facilmente fuggire dalla sorgente prima di essere deflessa in maniera significativa. Se invece il raggio della traiettoria è di molto inferiore alla dimensione della sorgente, la particella resterà intrappolata al suo interno. È difficile pensare che un oggetto astrofisico possa accelerare una particella se questa può sfuggirgli troppo rapidamente, e questo portò alla formulazione del **criterio di Hillas**: una sorgente astrofisica caratterizzata da una data dimensione e da un dato campo magnetico può potenzialmente accelerare una particella fino a una data energia solamente se la dimensione della sorgente è superiore al raggio di girazione della particella nel campo magnetico.

Per una discussione più quantitativa, va ricordato che il **raggio di girazione** di un nucleo atomico è definito come

$$r_g = E(\text{eV})/300 Z B(\text{G}) \text{ cm} ,$$

dove E è l'energia della particella in eV, Z il numero atomico e B l'intensità del campo magnetico in Gauss. La condizione per accelerare una particella di energia simile a quella rivelata da Linsley in un acceleratore cosmico di raggio L è quindi $r_g < L$, ovvero:

$$L > 0.1 Z^{-1} \left(\frac{E}{10^{20} \text{ eV}} \right) \left(\frac{B}{\text{G}} \right)^{-1} \text{ pc} \quad (1)$$

dove abbiamo introdotto il **parsec** che è un'unità di misura spesso usata in astrofisica ($1 \text{ pc} = 3 \times 10^{18} \text{ cm}$).

Hillas aveva ridotto un problema molto complesso alla determinazione di due proprietà fi-

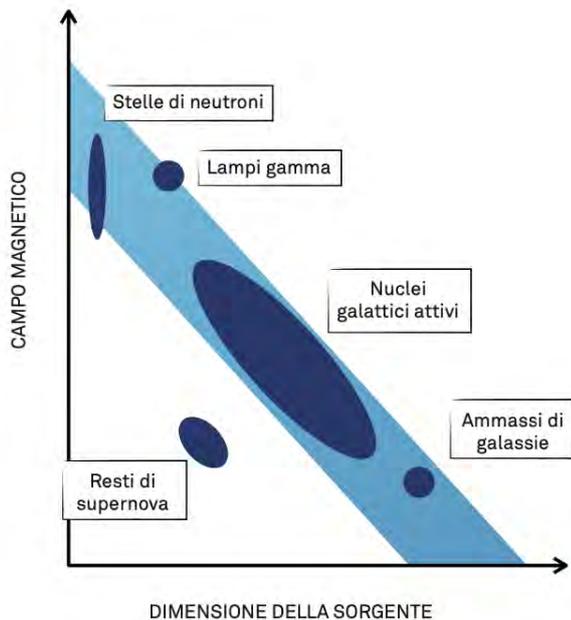


Figura 3: Il diagramma di Hillas.

siche degli oggetti astrofisici: dimensione della sorgente stessa e intensità del campo magnetico in essa contenuto. È possibile rappresentare visivamente questo concetto con un grafico noto come **diagramma di Hillas**. Vediamo come. Nel diagramma cartesiano mostrato in Fig. 3, l'asse orizzontale rappresenta il raggio della sorgente astrofisica L , mentre l'asse verticale l'intensità del campo magnetico B . Una data sorgente astrofisica occuperà una certa regione in questo diagramma. Il criterio di Hillas impone che solamente le sorgenti che si trovino al di sopra (Eq. 1) della banda diagonale visibile nel diagramma possano potenzialmente accelerare raggi cosmici di energia estrema. Il fatto che la banda colorata sia inclinata significa che i raggi cosmici più energetici possono essere accelerati sia da oggetti piccolissimi, ma dotati di campi magnetici di intensità elevatissima (ovvero quelli che occupano la regione in alto a sinistra del grafico), sia da oggetti enormi dotati di campi magnetici di intensità molto bassa (regione in basso a destra nel diagramma). Solo una manciata di oggetti astrofisici noti soddisfa questo criterio. Come descritto nel riquadro "Acceleratori cosmici", questi sono tra gli oggetti più estremi conosciuti nell'Universo, ma nonostante ciò si collocano nel diagramma a cavallo, non al di sopra, della barra colorata. Questo significa che accelerare i raggi cosmici più energetici è dif-

ficile anche per gli oggetti più estremi presenti nell'Universo! Comprendere quali tra questi oggetti siano i responsabili dell'accelerazione dei raggi cosmici che osserviamo sulla Terra resta una sfida ancora aperta per gli scienziati.

L'idea di Fermi

L'ultima importante questione da affrontare riguarda come i raggi cosmici vengano accelerati. Quale meccanismo astrofisico può imprimere a minuscoli corpuscoli energie così elevate da farli muovere a una velocità vicinissima a quella della luce? Fu Enrico Fermi, a cavallo tra gli anni Quaranta e Cinquanta, ad avere l'intuizione che permise di comprendere come i raggi cosmici potessero acquistare la loro energia [7]. Questa idea semplice ma fondamentale è la base sulla quale poggia la grande maggioranza degli studi teorici sull'accelerazione dei raggi cosmici effettuati fino a oggi. Il meccanismo di accelerazione di Fermi fu pensato per spiegare i raggi cosmici galattici, ma in realtà ha avuto e ha tuttora un grande impatto anche per quanto riguarda l'accelerazione dei raggi cosmici extragalattici.

Fermi notò che il disco della nostra galassia è pervaso da un gas di bassa densità detto mezzo interstellare, e anche da un campo magnetico. Sono due le proprietà fondamentali di questo gas magnetizzato. Per prima cosa, la sua densità e l'intensità del campo magnetico non sono uguali ovunque nel disco galattico, ma vi sono regioni in cui queste grandezze fisiche sono più elevate che altrove. In altre parole, gas e campo magnetico interstellari sono disomogenei. La seconda importante caratteristica è che gas e campo magnetico non sono entità statiche, bensì dinamiche. Il gas interstellare si muove in modo vorticoso, e trascina con sé il campo magnetico. Questa proprietà è chiamata turbolenza.

Ancora una volta, l'intuizione geniale fu quella di semplificare enormemente il problema. Per prima cosa, per tener conto della disomogeneità nella distribuzione di gas e campo magnetico nel disco galattico, Fermi fece la seguente ipotesi: gas e campo magnetico sono concentrati in nuvolette, al di fuori delle quali non c'è nulla. Un raggio cosmico, quindi, viaggerà indisturbato, muovendosi lungo una linea retta, fino a che non incontrerà una nuvoletta. La collisione del

Acceleratori cosmici

Scivolando dalla regione in alto a sinistra del diagramma di Hillas (Fig. 3) fino a quella in basso a destra incontriamo gli oggetti più estremi dell'Universo.

Stelle di neutroni – Nascono dalle esplosioni di supernova, e sono uno dei possibili stadi finali dell'evoluzione di stelle molto più massicce del Sole. Hanno una massa confrontabile a quella del Sole, ma un raggio di appena una decina di chilometri, il che significa che sono densissime ($\sim 10^8$ tonnellate/cm³). Si formano a causa del collasso gravitazionale della parti più interne della stella originaria: durante il collasso il campo magnetico resta intrappolato nel gas, e viene compresso fino a raggiungere valori elevatissimi.

Lampi gamma – Sono intensissimi lampi di raggi gamma di durata variabile dai millisecondi ai minuti. Appaiono nel cielo al ritmo di circa uno al giorno. All'origine di questi violentissimi eventi c'è il collasso gravitazionale di una stella supermassiva, o la fusione di due stelle di neutroni orbitanti una attorno all'altra.

Nuclei galattici attivi – Una galassia su cento emette dal suo nucleo quantitativi di radiazione molto superiori alla norma. Ciò è dovuto alla presenza di un buco nero supermassiccio (fino a miliardi di masse solari) che inghiotte materia dal mezzo circostante e la emette a una velocità prossima a quella della luce.

Ammassi di galassie – Sono enormi strutture contenenti fino a migliaia di galassie. Benché il loro campo magnetico sia piuttosto debole, le loro enormi dimensioni (più di cento volte quelle della nostra galassia) li rende dei buoni candidati per l'accelerazione di raggi cosmici.

Resti di supernova – La parte al di sotto della banda colorata nel diagramma di Hillas è popolata da moltissimi oggetti. Segnaliamo qui i resti di supernova perché gli scienziati ritengono che siano questi oggetti a produrre la stragrande maggioranza dei raggi cosmici (quelli di energia media). I resti di supernova non sono altro che l'onda d'urto che si propaga nel gas interstellare a seguito dell'esplosione.

raggio cosmico con la nuvoletta magnetizzata provocherà una deflessione della traiettoria del raggio cosmico. Se le nuvolette sono ferme, l'energia del raggio cosmico resterà invariata dopo la collisione. Possiamo servirci di un'analogia per comprendere quest'ultima affermazione. Sostituiamo il raggio cosmico con una pallina da tennis, e la nuvoletta con una racchetta. Se la racchetta è immobile (e se trascuriamo gli effetti dell'attrito e altri dettagli) la pallina rimbalzerà sul piatto delle corde e dopo la collisione si allontanerà dalla racchetta alla stessa velocità con la quale vi si avvicinava prima dell'impatto. Stessa velocità significa anche stessa energia.

Fermi a questo punto doveva tener conto anche dell'aspetto dinamico e turbolento del problema. E lo fece assegnando a ogni nuvoletta una velocità. Per semplificare ulteriormente il problema, si può ipotizzare che le nuvolette si muovano tutte alla stessa velocità, ma in direzioni casuali. A questo punto, si possono dividere le collisioni tra raggio cosmico e nuvolette in due categorie:

le collisioni frontali, nelle quali il raggio cosmico e la nuvoletta si corrono incontro, e le collisioni da tergo, nelle quali il raggio cosmico insegue la nuvoletta. Ripensando all'analogia con la pallina e la racchetta da tennis sarà facile intuire che dopo una collisione frontale il raggio cosmico avrà aumentato la propria energia, mentre dopo una collisione da tergo l'energia sarà minore di quella iniziale.

A prima vista si potrebbe pensare che, dato che le nuvolette si muovono lungo direzioni casuali, un raggio cosmico subirà in media lo stesso numero di collisioni frontali e da tergo. In questo caso, il guadagno netto di energia sarebbe nullo. Una riflessione più attenta, però, rivela che questo ragionamento non è corretto. Il motivo è che il tempo che un raggio cosmico impiega per compiere una collisione frontale è minore di quello necessario per compiere una collisione da tergo. Questo perché nel primo tipo di collisione raggio cosmico e nuvoletta si corrono incontro, mentre nel secondo il raggio cosmico impiega un

tempo maggiore per raggiungere una nuvoletta che si allontana. In media, quindi, un raggio cosmico subisce più collisioni frontali che da tergo, e quindi la sua energia cresce gradualmente con il passare del tempo. Questa intuizione di Fermi resta tutt'oggi un pilastro della Fisica dell'accelerazione dei raggi cosmici. I campi magnetici, onnipresenti nell'Universo, giocano quindi un ruolo fondamentale anche nei meccanismi di accelerazione dei raggi cosmici.

Meccanismi di accelerazione al primo e secondo ordine

Lo stesso Fermi si rese ben presto conto che la sua idea aveva un punto debole. L'incremento relativo medio di energia di un raggio cosmico è dell'ordine di $\langle \Delta E/E \rangle \approx (V/c)^2$ per collisione, dove V è la velocità tipica delle nuvolette magnetizzate, e c la velocità del raggio cosmico, dell'ordine della velocità della luce (per una derivazione di questo risultato si veda il riquadro "Il meccanismo di Fermi per l'accelerazione dei raggi cosmici"). Nell'espressione compare il quadrato del rapporto V/c , e per questo motivo si parla spesso di meccanismo di Fermi al secondo ordine. Il fatto che le velocità tipiche dei moti turbolenti del gas interstellare siano al più dell'ordine di ~ 100 km/s implica che un raggio cosmico incrementi la sua energia di un fattore minuscolo dopo ogni collisione, $\langle \Delta E/E \rangle \approx 10^{-5}(V/100 \text{ km/s})^2$, ovvero, l'energia cambia in maniera apprezzabile dopo centinaia di migliaia di collisioni! Questo rende il meccanismo al secondo ordine poco efficace, perché il tempo di fuga delle particelle dalla regione magnetizzata (in questo caso il disco della nostra Galassia) è molto più breve del tempo necessario ad accelerare sufficientemente le particelle cosmiche. Per spiegare le alte energie delle particelle che compongono la radiazione cosmica bisogna quindi ipotizzare l'esistenza di un meccanismo di accelerazione molto più efficace di quello proposto da Fermi.

La soluzione a questo problema venne proposta indipendentemente da quattro diversi gruppi di ricercatori alla fine degli anni '70 (si veda [8] per un articolo di rassegna). L'idea proposta è, in buona sostanza, una variazione del meccanismo di Fermi. Per rendere questo meccanismo

più efficace, basta ipotizzare che i raggi cosmici subiscano solamente collisioni frontali. Tornando all'analogia proposta nel capitolo precedente, questo equivarrebbe a considerare una pallina da tennis intrappolata tra due racchette in moto l'una verso l'altra. In questo caso, solo collisioni frontali sono possibili e la pallina guadagna energia ogni volta che rimbalza su una racchetta.

Considerare solo collisioni frontali significa introdurre un meccanismo di Fermi al primo ordine, molto più efficace, in quanto l'incremento di energia per collisione frontale è dominato dal termine V/c , e non da $(V/c)^2$. Questo risultato può essere facilmente derivato nel caso di particelle non-relativistiche (si veda il riquadro), imponendo $v \gg V$ e $v \sim c$:

$$\frac{\Delta E_+}{E} \sim \frac{V}{c^2}(V+c) \approx \left(\frac{V}{c}\right) + \left(\frac{V}{c}\right)^2 \quad (2)$$

dove abbiamo trascurato i piccoli fattori numerici moltiplicativi. In questo caso, il numero di collisioni necessarie per cambiare significativamente l'energia di una particella è dell'ordine di $\approx (V/c)^{-1} \sim 300 (V/100 \text{ km/s})^{-1}$

Esistono oggetti astrofisici nei quali il meccanismo di Fermi al primo ordine possa operare? Fortunatamente, la risposta è sì: le onde d'urto che si formano in plasmi astrofisici magnetizzati sono infatti acceleratori di particelle molto efficienti.

L'accelerazione diffusiva in onde d'urto

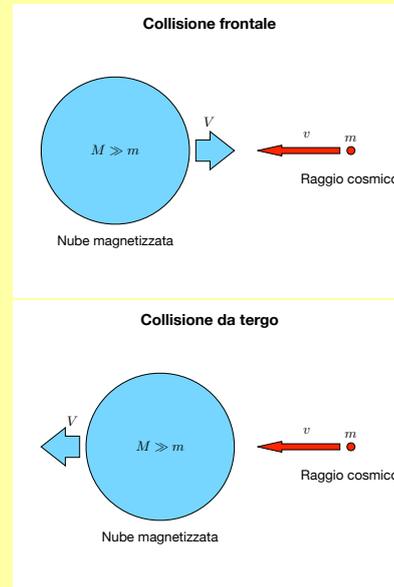
La velocità del suono c_s definisce la velocità massima alla quale una informazione può essere comunicata da un punto all'altro di un mezzo per via meccanica (per esempio in un plasma astrofisico). Le onde d'urto si formano quando una perturbazione si muove in un mezzo a una velocità supersonica.

Quando questo avviene, una discontinuità appare nel plasma, a dividere il volume occupato dal mezzo imperturbato (a monte dell'onda d'urto, o *upstream*) da quello perturbato (a valle dell'onda d'urto, o *downstream*).

Attraversando la discontinuità si registrano forti variazioni della pressione, temperatura, densità e velocità del mezzo.

Il meccanismo di Fermi per l'accelerazione dei raggi cosmici

Il 4 dicembre 1948, Enrico Fermi riempì di calcoli quattro pagine del suo quaderno, gettando le basi della teoria per l'accelerazione dei raggi cosmici. Questa teoria, nota oggi come **meccanismo di Fermi**, è tuttora il punto di partenza di numerosissimi studi sull'accelerazione dei raggi cosmici. Fermi considerò inizialmente particelle non-relativistiche, caratterizzate cioè da velocità v molto inferiori a quella della luce. Inoltre, distinse due tipi di collisioni elastiche tra particelle di massa m e nubi magnetizzate di massa $M \gg m$ e velocità V : le collisioni frontali, nelle quali la particella e la nube si muovono in direzioni opposte prima della collisione, e quelle da tergo, nelle quali la particella insegue la nube.



Nel sistema di riferimento nel quale la nube è a riposo, la velocità della particella sarà $v \pm V$ (il segno + si riferisce a collisioni frontali, il - a collisioni da tergo). In quel sistema di riferimento, a seguito della collisione la particella inverte semplicemente la direzione di moto. Ne segue che, nel sistema di riferimento dell'osservatore, l'energia cinetica della particella cambierà da $E_1 = mv^2/2$ a $E_2 = m(v \pm 2V)^2/2$. La variazione relativa di energia per collisione è quindi:

$$\frac{\Delta E_{\pm}}{E} = \frac{E_2 - E_1}{E_1} = \frac{4V}{v^2}(V \pm v)$$

Dal momento che $v \gg V$ (i raggi cosmici si muovono più velocemente delle nubi interstellari), le particelle guadagnano energia a seguito di collisioni frontali ($\Delta E_+ > 0$), mentre ne perdono in collisioni da tergo ($\Delta E_- < 0$). Bisogna considerare, però, che le collisioni frontali avvengono a una frequenza $\nu_+ = (v + V)/L$ maggiore di quella alla quale avvengono le collisioni da tergo, $\nu_- = (v - V)/L$, dove L è la distanza caratteristica tra le nubi. La probabilità di avere collisioni frontali o da tergo è quindi $P_{\pm} = (v \pm V)/2v$ e la variazione media di energia per collisione è:

$$\left\langle \frac{\Delta E}{E} \right\rangle = \frac{\Delta E_+}{E} P_+ + \frac{\Delta E_-}{E} P_- = 8 \left(\frac{V}{v} \right)^2 > 0$$

Le particelle, quindi, guadagnano energia! [Ponendo $v \sim c$ (i raggi cosmici sono relativistici) e mediando su tutti i possibili angoli tra v e V si ottiene il risultato esatto: $\langle \Delta E/E \rangle = (4/3)(V/c)^2$.]

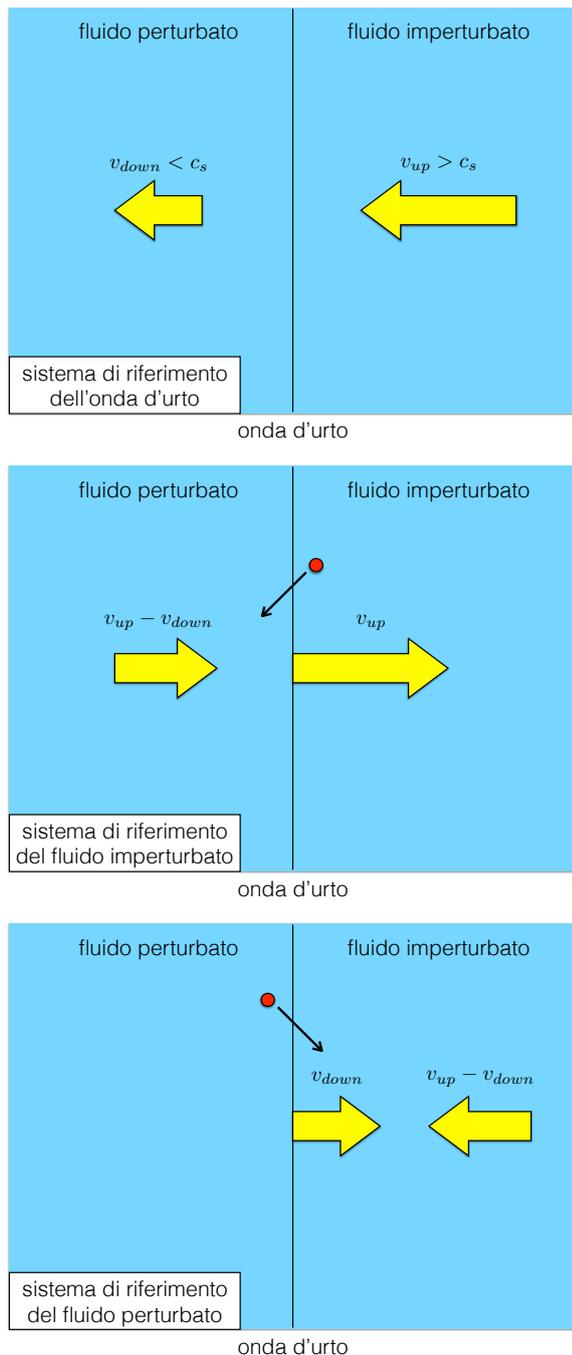


Figura 4: *Onda d'urto piana vista dal sistema di riferimento in cui la discontinuità (pannello superiore), il plasma a monte (pannello centrale) e a valle dell'onda d'urto (pannello inferiore) sono a riposo. Il punto rosso rappresenta un raggio cosmico che attraversa l'onda d'urto.*

Per semplicità, consideriamo una discontinuità piana che si muova in un plasma omogeneo a una velocità costante V . Spesso, è conveniente studiare le onde d'urto nel sistema di riferimento nel quale la discontinuità è a riposo (si veda il pannello superiore della Fig. 4). In quel sistema di riferimento il gas a monte della discontinuità si

avvicinerà ad essa con una velocità supersonica $v_{up} = V$. Il plasma, attraversando la superficie di discontinuità viene compresso e riscaldato, e si allontana a valle dell'onda d'urto a una velocità subsonica $v_{down} < v_{up}$ (si veda [9] per una trattazione dettagliata delle onde d'urto).

Immaginiamo ora che il plasma nel quale propaga l'onda d'urto sia strutturato in nuvolette magnetizzate, proprio come nello scenario descritto da Fermi nel suo importante lavoro sull'accelerazione dei raggi cosmici. Assumiamo inoltre che queste nuvolette magnetizzate si muovano in maniera solidale con il plasma, ovvero che tutte le nuvolette a monte della discontinuità si muovano verso di essa a una velocità v_{up} , e tutte quelle a valle della discontinuità si allontanino da essa a una velocità v_{down} . A questo punto, è utile considerare il problema nel sistema di riferimento nel quale il plasma a monte della discontinuità è a riposo (pannello centrale in Fig. 4). In questo sistema di riferimento le nuvolette magnetizzate che si trovano a monte della discontinuità saranno anch'esse a riposo, mentre la discontinuità sarà in moto verso destra con velocità v_{up} . In queste circostanze, le collisioni tra queste nuvolette (che sono a riposo) e un raggio cosmico hanno come unico effetto quello di cambiare la direzione di moto del raggio cosmico, lasciando invariata la propria energia.

Consideriamo ora un raggio cosmico che, dopo un certo numero di collisioni con nuvolette magnetizzate, attraversi la discontinuità e passi a valle dell'onda d'urto (pannello inferiore in Fig. 4). Il raggio cosmico vedrà le nuvolette a valle dello shock venirgli incontro a una velocità $v_{up} - v_{down}$.

Passiamo ora al sistema di riferimento dove il plasma a valle della discontinuità è a riposo (pannello inferiore in Fig. 4). In questo sistema di riferimento, le nuvolette magnetizzate a valle della discontinuità saranno a riposo, e la discontinuità si muoverà verso destra a una velocità v_{down} . Inoltre, le collisioni tra un raggio cosmico e le nuvolette (a riposo) lasceranno inalterata l'energia della particella. Un raggio cosmico che, dopo diverse collisioni, riattraversi la discontinuità per ritornare a monte dell'onda d'urto vedrà le nuvolette a valle dello shock venirgli incontro a una velocità $v_{up} - v_{down}$, esattamente come nel passaggio da monte a valle dell'onda

d'urto descritto sopra.

Riassumendo: un raggio cosmico che attraversa la discontinuità (poco importa se da monte a valle o viceversa) vedrà invariabilmente una nuvoletta magnetizzata corrergli incontro a una velocità $v_{up} - v_{down}$. Questo significa che ad ogni ciclo attorno all'onda d'urto (monte \rightarrow valle \rightarrow monte, o valle \rightarrow monte \rightarrow valle) il nostro raggio cosmico subirà una collisione frontale e guadagnerà energia! Ne segue che una particella che esegua innumerevoli cicli attorno alla discontinuità sarà soggetta a un meccanismo di accelerazione di Fermi al primo ordine, e aumenterà rapidamente la propria energia.



Figura 5: Il resto di una supernova esplosa nell'anno 1006, come è stata osservata nelle onde radio (rosso), nei raggi X (blu) e nella radiazione visibile (giallo). Immagine da *Astronomy Picture of the Day*.

Il meccanismo di accelerazione diffusiva ad onde d'urto è stato verificato in varie osservazioni, e ci sono quindi pochi dubbi sul fatto che sia operativo in un gran numero di sorgenti astrofisiche. Strumenti a bordo di sonde spaziali ci hanno permesso di rivelare *in situ* la presenza di particelle accelerate in onde d'urto interplanetarie, ed osservazioni dell'emissione prodotta da particelle energetiche proveniente da resti di supernove dimostrano che queste onde d'urto sferiche in espansione sono in grado di accelerare raggi cosmici (si veda ad esempio Fig. 5). Come abbiamo già avuto modo di discutere (si veda la discussione relativa alla Fig. 3), la combinazione

di campo magnetico e dimensione dei resti di supernova li rende incapaci di accelerare particelle fino alla massima energia osservata nei raggi cosmici. Una delle possibilità è che l'accelerazione di queste particelle di energia estrema possa essere prodotta da onde d'urto relativistiche, ovvero in moto a velocità vicinissime a quella della luce. Queste onde d'urto si formano nelle esplosioni che originano i lampi gamma, o a seguito di eventi eruttivi che avvengono frequentemente nei nuclei di galassie attive.

La spiegazione dell'origine dei raggi cosmici di energia altissima resta un mistero. Ma la soluzione di questo mistero potrebbe essere non troppo lontana.



- [1] V. L. Fitch, *Elementary particle physics: The origins*, *Reviews of Modern Physics*, 71 (1999) S25.
- [2] B. Rossi, *I raggi cosmici*, Einaudi, Torino (1971).
- [3] S. Gabici, *Raggi Cosmici*, RCS MediaGroup (2019).
- [4] A. de Angelis, *L'enigma dei raggi cosmici*, Springer, Milano (2011).
- [5] J. Linsley, *Search for the end of the cosmic ray energy spectrum*, *AIP Conference Proceedings*, 433 (1998) 1.
- [6] A. M. Hillas, *The Origin of Ultra-High-Energy Cosmic Rays*, *Annual Review of Astronomy and Astrophysics*, 22 (1984) 425.
- [7] E. Fermi, *On the Origin of the Cosmic Radiation*, *Physical Review*, 75 (1949) 1169.
- [8] L. O'C. Drury, *An introduction to the theory of diffusive shock acceleration of energetic particles in tenuous plasmas*, *Reports on Progress in Physics*, 46 (1983) 973.
- [9] F. H. Shu, *The physics of astrophysics. Volume II: Gas dynamics*, University Science Books, Sausalito (1992)
- [10] B. Achterberg, *Cosmic Accelerators*, Lecture notes for the IAC2008, http://www.rci.uct.ac.za/sites/default/files/image_tool/images/395/Course_Resources/Pieter_Meintjies/ba-lecture.pdf



Stefano Gabici: è ricercatore al laboratorio AstroParticelle e Cosmologia (APC) di Parigi. Si occupa di raggi cosmici e di astronomia nei raggi gamma.

L'importanza dei neutrini cosmici di altissima energia

Carla Distefano

INFN - Laboratori Nazionali del Sud (LNS), Catania

Francesco Vissani

INFN - Laboratori Nazionali del Gran Sasso (LNGS), Assergi, L'Aquila

Presentiamo una breve ed elementare introduzione ai neutrini cosmici di altissima energia, esaminando le ragioni del loro interesse, le connessioni con i raggi cosmici e con i raggi gamma, e alcune plausibili direzioni di sviluppo del campo.

Origine dei raggi cosmici e neutrini

Una delle principali motivazioni per la ricerca e lo studio dei neutrini cosmici di altissima energia¹ è la speranza di imparare qualcosa sulle sorgenti dei raggi cosmici. Richiamiamo l'argomento a supporto di questo programma di ricerca, chiarendo entro quali condizioni ci aspettiamo un collegamento e quali sinergie ci potrebbero essere.

¹Con questa locuzione, ci riferiremo in particolare a neutrini di energia compresa tra le frazioni del TeV e le decine di PeV (ovvero $10^{11} - 10^{16}$ eV) dove $1 \text{ eV} = 1.6 \times 10^{-12}$ erg. Inoltre, parleremo come regola di neutrini senza distinguerli - a meno di indicazione esplicita - dagli antineutrini, in quanto, fatta eccezione per certi casi speciali, non li possiamo identificare per tali nei rivelatori.

Cosa sappiamo sui raggi cosmici

Nonostante si conosca da più di un secolo l'esistenza della continua pioggia o bombardamento di nuclei provenienti dallo spazio, che chiamiamo raggi cosmici, ne ignoriamo ancora l'origine. Le ragioni sono piuttosto semplici da capire. Lo spazio tra le stelle è permeato da campi magnetici; per esempio, nella nostra galassia, il valore tipico è di qualche micro-Gauss, mentre nello spazio intergalattico è 1000 volte più piccolo. Pertanto, le particelle cariche cambiano direzione in tempi relativamente brevi. Per convincercene, notiamo che il raggio di curvatura è dato dalla formula di Larmor

$$R = \beta \frac{E}{qB},$$

dove E è l'energia della particella, q è la carica elettrica, B il campo magnetico e $\beta = v_T/c$ è il rapporto tra la velocità della particella v_T in direzione ortogonale al campo magnetico e la velocità della luce nel vuoto. Come riferimento consideriamo un protone di energia pari a 1 PeV ($1 \text{ PeV} = 10^{15} \text{ eV}$), avente a questa energia una velocità $\beta \sim 1$. Poiché un protone ha una carica

elettrica pari a $q = 4.8 \times 10^{-10}$ esu, otteniamo

$$R \sim 1 \text{ parsec} \times \frac{E}{\text{PeV}} \times \frac{\mu G}{B}$$

dove un parsec (pc) è circa tre anni luce. Per confronto, la distanza dal centro galattico è 8.5 kpc e lo spessore del disco galattico è 100 – 300 pc. Pertanto, non ci aspettiamo che la direzione di arrivo dei raggi cosmici ci possa dare indizi importanti sulla direzione originaria di provenienza, se non forse alle più alte energie osservate.

L'informazione principale di cui possiamo disporre è la distribuzione di energia dei raggi cosmici, detta anche spettro.² Esso è stato misurato fino ad energie dell'ordine di 10^{20} eV, e si estende dunque da energie pari a $m_p c^2 \sim 1$ GeV (m_p = massa del protone) per più di 10 ordini di grandezza, riducendosi in intensità di circa 30 decadi (vedi fig. 1). I raggi cosmici hanno energie molto più elevate di quelle fino ad oggi raggiunte nei nostri laboratori di fisica. L'acceleratore LHC del CERN ha permesso di produrre protoni fino con energie appena superiori a 10 TeV (10^{12} eV), ben otto ordini di grandezza inferiori alle energie più estreme dello spettro dei raggi cosmici. spiegare l'esistenza di questi ultimi è quindi necessario ipotizzare l'esistenza di acceleratori cosmici molto più potenti di quelli costruiti dall'uomo.

Le osservazioni mostrano uno spettro continuo che può essere descritto da leggi di potenza, $I \sim E^{-\gamma}$ in tratti successivi, dove γ è detta la pendenza (in inglese, *slope*). Un marcato cambiamento avviene per i protoni ad un'energia

$$E_p \sim 4 \text{ PeV}$$

e viene detto il **ginocchio**; si tratta di un incremento della pendenza, da $\gamma \approx 2.7$ fino a 3 (diventando circa 3.3 poco prima di 1000 PeV). La sua origine non è nota con sufficiente certezza. Una possibilità è che questa energia sia legata ad una popolazione di raggi cosmici tipici della nostra galassia, che hanno raggiunto il massimo valore dell'accelerazione; un'altra è che, a partire da quella energia in poi, i raggi cosmici riescano

²Una affermazione un po' maliziosa, attribuita a Bruno Pontecorvo, è che "la pretesa di capire i raggi cosmici dal loro spettro è tanto ambiziosa quanto quella di capire cosa si mangia in una certa città, dopo avere annusato gli scarichi delle fogne".

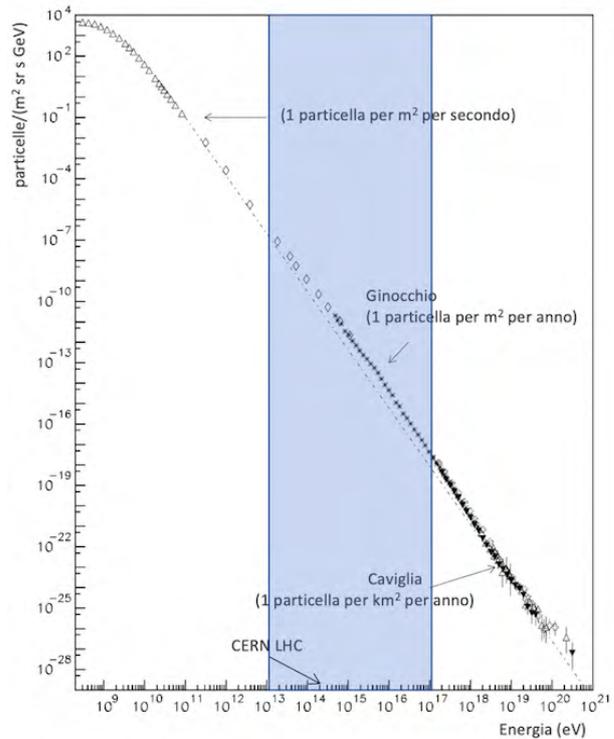


Figura 1: Lo spettro di energia dei raggi cosmici ottenuto da Swordy [1] come compilazione di dati di diversi esperimenti. La regione in azzurro mette in evidenza la parte di spettro connessa al flusso di neutrini discusso in questo articolo.

con sempre maggior successo a sfuggire, avendo un raggio di Larmor maggiore delle dimensioni della nostra galassia. In questo quadro interpretativo, ci si aspetta una serie di ginocchi in posizioni corrispondenti alle cariche elettriche dei vari nuclei, fino a circa $E_p \times Z_{Fe} \sim 100$ PeV e ne abbiamo qualche evidenza. Ma, soprattutto, si deve concludere che da qualche parte debba subentrare una nuova popolazione di raggi cosmici, plausibilmente di origine extra-galattica. Si noti che intorno a 5 EeV si osserva un ulteriore cambiamento di pendenza dello spettro dei raggi cosmici, detto **caviglia**, dove γ ritorna simile al valore 2.7; ma una prima diminuzione della pendenza avviene già intorno a 20 PeV.

La connessione con i neutrini di altissima energia

Siamo pronti a presentare l'idea di massima della connessione che ci interessa: ci si aspetta che la densità dei raggi cosmici sia particolarmente alta all'interno delle loro sorgenti e che, a causa della presenza di qualche tipo di bersaglio sul quale

i raggi cosmici impattano, si possano produrre particelle secondarie tra le quali neutrini e raggi gamma di altissima energia. Questi, essendo particelle prive di carica elettrica, non vengono deflessi dai campi magnetici come nel caso dei raggi cosmici. Dall'osservazione delle loro direzioni di arrivo e dalle informazioni note alla astronomia, si spera quindi di identificare quali siano le sorgenti dei raggi cosmici.

Prima di procedere, sottolineiamo che l'ipotesi relativa ai bersagli è cruciale; a seconda della loro composizione e consistenza, si produrranno più o meno particelle secondarie. In altre parole, non sembra possibile imparare molto sui raggi cosmici, senza capire abbastanza bene quali siano i bersagli.

La produzione di queste particelle secondarie (gamma e neutrini) passa attraverso certe particelle instabili, le più importanti delle quali sono i pioni, che vengono prodotte nelle collisioni dei raggi cosmici: vedi fig. 2. Ci sono pioni di due tipi: quelli neutri che danno origine ai raggi gamma, quelli carichi che danno origine ai neutrini. Molto rozzamente, l'energia dei raggi gamma è 10 volte più piccola di quella dei raggi cosmici primari, quella dei neutrini circa la metà di quella dei gamma.³ Quindi una regione estremamente interessante, in considerazione delle stime sopra, è quella nell'intervallo di energia dei neutrini

$$E_\nu = [0.01 , 50] \text{ PeV}$$

del quale, la parte inferiore a qualche PeV è stato con successo esplorato da IceCube come ricordiamo sotto. Bisogna però notare che i raggi gamma possono essere assorbiti molto più facilmente dei neutrini e, a differenza di questi, possono essere anche prodotti tramite interazioni elettromagnetiche; in altre parole, non è automaticamente garantito che i flussi di neutrini e raggi gamma di alte energie si corrispondano esattamente. Infine, ci potremmo aspettare che qualche cosa di

³Questa stima deriva dall'assumere che il pione di energia principale porti via all'incirca il 20% dell'energia iniziale, che poi viene ripartita in modo simile tra le particelle in cui esso decade. Ma ripetiamo che va presa con molta cautela, siccome deviazioni anche molto importanti da queste aspettative sono possibili in vari casi concreti.

interessante accada intorno a

$$E_\nu \sim 0.2 \text{ PeV}$$

che corrisponde grosso modo al ginocchio dei raggi cosmici galattici. In effetti un nucleo che porta una carica ed una energia Z volte maggiore di quella di un protone, ripartisce la sua energia tra gli A nucleoni, e dunque - per quanto attiene alla produzione di neutrini - conta come un insieme di nucleoni con energia $E_p \times Z/A$, che è inferiore a quella del protone, E_p . Sulla base di queste considerazioni la regione del ginocchio dovrebbe essere ben marcata nel caso dei neutrini, anche se ci aspettiamo l'allargamento tipico delle distribuzioni secondarie per ogni nucleo componente. Sarebbe interessante osservare una caratteristica del genere nello spettro galattico; un ostacolo a farlo è la presenza concomitante di altri segnali di neutrini (plausibilmente non galattici) che sembrerebbero fare davvero parte dello spettro osservato come discusso in seguito.

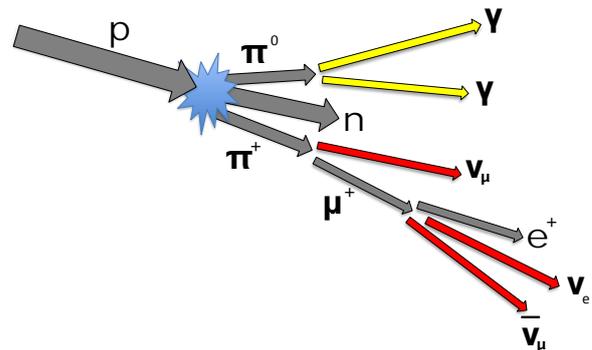


Figura 2: Le collisioni dei raggi cosmici, presso il sito d'origine, nell'atmosfera terrestre, o altrove producono delle particelle instabili, il cui decadimento dà luogo a radiazione secondaria potenzialmente osservabile. Nella figura i fotoni sono mostrati in giallo, ed indicati con il simbolo γ ; i neutrini e gli antineutrini in rosso. Si noti che vengono prodotti due tipi di neutrini muonici ed uno di tipo elettronico.

Ipotesi sulle sorgenti di neutrini

Per essere certi di procedere su un terreno sicuro, bisognerebbe avere informazioni credibili su quali siano le sorgenti, quali siano lo spettro e

l'intensità iniziali dei raggi cosmici, e quali siano i bersagli sui quali avviene la conversione in particelle secondarie. Purtroppo i dati di cui disponiamo sono ancora lacunosi e si deve continuare a procedere avanzando ipotesi, nelle quali abbiamo solo una relativa confidenza.

Sorgenti galattiche

Nel caso dei raggi cosmici galattici, c'è una ipotesi che viene considerata con speciale attenzione: quella che essi provengano dai resti di supernova. Ricordiamo un importante argomento formulato da Ginzburg e Syrovatskii [2] più di mezzo secolo fa. Immaginiamo che i raggi cosmici siano confinati in un volume $V \sim 4000 \text{ kpc}^3$ (p.e., in un cilindro di altezza 6 kpc e raggio 15 kpc) per un tempo T intorno a 50 milioni di anni; siccome la loro densità vale circa $u \sim 1 \text{ eV/cm}^3$, abbiamo che la potenza irradiata dalla nostra galassia vale

$$P_{\text{out}} \sim u \times V/T \sim 4 \times 10^{48} \text{ erg / anno.}$$

Ora, sappiamo che ogni supernova espelle una gran massa di gas, detta resto di supernova che immette nella galassia una quantità di energia cinetica tipicamente intorno a $\mathcal{E}_k \sim 10^{51} \text{ erg}$. Considerando che nella nostra galassia abbiamo una supernova ogni $\tau = 50$ anni circa, possiamo compensare la potenza irradiata se una frazione $f \sim 20\%$ di questa energia cinetica viene convertita in raggi cosmici ed avere quindi

$$P_{\text{in}} = f \mathcal{E}_k / \tau \sim P_{\text{out}}.$$

Se questa congettura vale, i guadagni compensano le perdite ed il sistema galassia (che vive miliardi di anni) è all'equilibrio su questi tempi scala molto più brevi. Alla luce della teoria e delle informazioni note, una frazione del genere è del tutto ragionevole.

Passiamo adesso a discutere la produzione di particelle secondarie. Nel primo migliaio di anni dopo l'esplosione, non ci si aspetta che i raggi cosmici siano fuggiti in maniera sostanziale dalla regione intorno alla supernova, dove vengono accelerati. Dunque, sembra ragionevole pensare alla possibilità di una significativa emissione di neutrini, prodotti dalle collisioni tra i raggi cosmici e lo stesso gas in espansione. Questo tipo di ipotesi ha due interessanti conseguenze:

- * le collisioni tra adroni ed adroni danno luogo ad uno spettro di neutrini che segue abbastanza fedelmente la stessa legge di potenza che caratterizza i raggi cosmici primari;
- * la radiazione gamma non è assorbita in misura importante. Detto altrimenti, ci aspettiamo una corrispondenza tra i neutrini ed (una parte della) radiazione gamma di altissima energia che viene emessa.

Queste osservazioni consentono di calcolare il massimo dell'emissione neutrinica attesa, utilizzando i dati osservativi sui gamma di altissima energia, che ipotizziamo essere interamente della natura appena descritta⁴, e sfruttando le relazioni note tra produzione di pioni neutri e pioni carichi e la cinematica dei decadimenti coinvolti. Queste considerazioni selezionano naturalmente, come obiettivi per l'osservazione dei neutrini, i più vicini tra i resti di supernova noti, e tra essi, risultano particolarmente interessanti RXJ 1713-3947 e Vela Jr, essendo questi anche tra le più intense sorgenti del cielo gamma.

Esistono anche altre possibilità promettenti per l'osservazione di neutrini di origine galattica: le zone particolarmente dense di gas, come la cosiddetta regione del Cigno che ammonta a ben 100 mila masse solari; i conglomerati di esplosioni di supernove, dette superbolle (in inglese *superbubbles*); la regione intorno al centro galattico; ed infine, il disco galattico stesso. In tutti questi casi, la regione cruciale di energia, alle quali si potrebbero osservare i neutrini, è relativamente bassa: intorno ai 10 TeV (=0.01 PeV). I rivelatori di raggi gamma in funzione come LHAASO [3] o quelli in progetto come CTA [4] dovrebbero fornirci informazioni particolarmente utili in proposito. Ricordiamo che LHAASO ha recentemente ottenuto evidenze di sorgenti galattiche con energie del PeV, che hanno sorpreso la comunità scientifica, e che riguardano in particolare l'interessantissima nebulosa del Granchio [5].

⁴Ovvero trascurando l'altro meccanismo di produzione: quello elettromagnetico. Per migliorare queste stime, al momento sembra inevitabile ricorrere a dei modelli delle supernove.

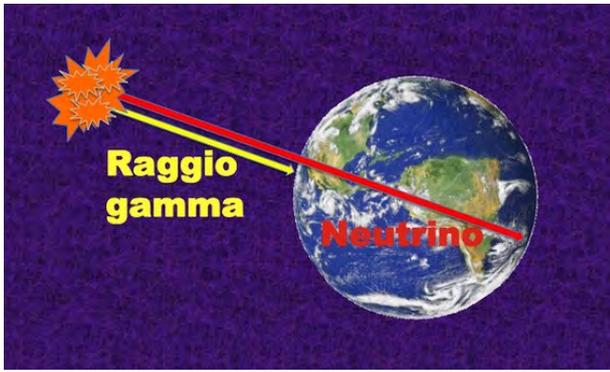


Figura 3: Una buffa caratteristica dei telescopi di neutrino è che essi vedono meglio la regione sotto l'orizzonte, in quanto la Terra fornisce il modo di convertire i neutrini muonici in segnale e funge da schermo ai muoni di origine atmosferica. Pertanto, sono in massima sinergia con i rivelatori di gamma di altissima energia posizionati in emisferi diversi.

Sorgenti extragalattiche

Disponiamo di osservazioni di onde elettromagnetiche provenienti dal cielo che si estendono di 20 ordini di grandezza in frequenza, dalle onde radio ai raggi gamma,⁵ che testimoniano la presenza di fenomeni di altissima energia nell'universo. Ma quando ci interessiamo di gamma di origine extra-galattica, bisogna ricordare che quelli con energia superiore ai 100 GeV vanno incontro alla conversione in coppie elettrone-positrone, per effetto dell'interazione con i fotoni del fondo diffuso di radiazione. Il flusso di gamma di altissima energia, corrispondente ai neutrini a cui siamo interessati, viene pertanto fortemente assorbito prima di raggiungere i nostri telescopi. Dunque, nel caso di sorgenti extra-galattiche, dobbiamo stare sull'avviso: la connessione attesa tra neutrini e gamma è solo indiretta e dipende dai dettagli dei modelli con cui la descriviamo.

D'altro canto, possiamo misurare le energie emesse da varie classi di oggetti extra-galattici, ed è possibile sincerarci che ci sono diversi siti astrofisici che, sulla carta, sarebbero in grado di compensare la perdita di energia dei raggi cosmici di più alta energia, causata tra le altre cose dall'espansione dell'universo⁶. Più nello specifico,

⁵I radiotelescopi osservano lunghezze d'onda fino a $\ell \sim 1$ metro circa, che corrisponde ad energie del μeV , mentre i moderni telescopi Cherenkov arrivano fino al PeV.

⁶Durante l'espansione dell'universo, l'energia delle par-

possiamo ipotizzare che essi riescano, in un modo o nell'altro, ad iniettare in nuovi raggi cosmici una potenza simile a quella che osserviamo per mezzo della radiazione elettromagnetica; in altre parole, pur non conoscendo i dettagli dei meccanismi di produzione, siamo ragionevolmente certi dell'esistenza di sistemi fisici che immagazzinano tutta l'energia necessaria per continuare a contribuire alla popolazione di raggi cosmici di origine extra-galattica.

È possibile proporre un argomento simile, basandosi direttamente sui raggi cosmici. In particolare, nel cosiddetto limite di Waxmann e Bahcall [4], il flusso dei neutrini viene postulato essere una legge di potenza del tipo $I \propto E_\nu^{-2}$ come suggerito dalla teoria, e la normalizzazione viene fissata proprio con considerazioni simili a quelle descritte appena sopra, ovvero supponendo che una frazione significativa della energia trasportata dai raggi cosmici di più alta energia osservati sopra una certa soglia⁷ venga convertita in particelle secondarie. Anche se certi fatti, che sono stati acquisiti più recentemente, portano a ritenere che le ipotesi alla base di questo limite non siano del tutto affidabili, esso resta un utile riferimento e un valido modello per procedere a definire meglio le aspettative teoriche.

Sono stati suggeriti vari possibili siti per l'origine dei neutrini di altissima energia. Originariamente si pensava ad oggetti stellari estremi, come i lampi di raggi gamma (in inglese, *gamma ray burst*) che crediamo siano prodotti da una ristretta classe di supernove con comportamenti estremi. Oggi si ritengono più plausibili le regioni intorno ai buchi neri che sono al centro di ogni galassia, nelle quali la conversione di energia gravitazionale, causata dalla materia che accresce la massa del buco nero, dà luogo a fenomeni di emissione. Si parla di nuclei galattici attivi (in inglese, *active galactic nuclei*) e tra questi, sembrano di particolare interesse quelli che danno luogo ad un getto di radiazione e materia pun-

ticelle diminuisce. Infatti, le lunghezze d'onda ℓ delle particelle libere presenti nell'universo crescono proporzionalmente al fattore di scala; pertanto le frequenze $f = c/\ell$ diminuiscono, e assieme ad esse diminuiscono le energie, che sono proporzionali alle frequenze secondo la relazione $E = hf$.

⁷Precisiamo che si tratta dei raggi cosmici che misuriamo a terra e che postuliamo essere rappresentativi della distribuzione media nel cosmo.

tato in direzione della Terra, come i cosiddetti Blazars e FSRQ⁸, che sono i più notevoli oggetti del cielo extra-galattico nella finestra osservativa gamma. In tutti questi casi, il bersaglio per la conversione dei raggi cosmici sarebbe molto probabilmente costituito da fotoni, piuttosto che da altri nucleoni.

Oltre a questi oggetti, sembrano particolarmente interessanti gli ammassi di galassie e forse ancora di più, le cosiddette galassie soggette ad intensa formazione stellare (dette in inglese *starforming* o in certi casi addirittura *starburst*). Queste ultime, proprio come la nostra galassia, sono certo dotate di una popolazione di raggi cosmici e - forse a differenza della nostra galassia - dispongono anche di moltissima materia in forma di gas, che consente la formazione stellare. Se, per qualche ragione, il massimo valore dei raggi cosmici, intrappolati in quei siti, fosse di uno due ordini di grandezza superiore a quello della nostra galassia, esse potrebbero essere siti ideali per la produzione di neutrini di altissima energia. In questi casi, naturalmente, ci aspetteremmo che i bersagli per la conversione in neutrini siano costituiti da nucleoni.

Un esempio di nucleo galattico attivo è NGC 5128, ovvero la galassia Centaurus A, uno di galassia *starburst* è M82, detta anche la galassia sigaro; il primo è a 4.6 Mpc, il secondo è a 3.5 Mpc dalla Terra.

I risultati di IceCube

L'osservatorio IceCube, in Antartide, monitora un volume di circa un chilometro cubo di ghiaccio, ad una profondità che va da 1450 a 2450 metri, alla ricerca di segnali di neutrini cosmici. Nel 2013 ha ottenuto i primi risultati: due eventi di energia intorno al PeV non compatibili con il flusso di neutrini prodotti dall'interazione dei raggi cosmici con l'atmosfera terrestre e candidati quindi ad essere dovuti a neutrini di origine cosmica.

⁸Essi corrispondono agli acronimi inglesi *blazing quasi-stellar object* - quasar fiammeggiante - e *flat spectrum radio quasar* - quasar con spettro radio piatto. Come è ben noto, quasar indica un *quasi stellar radio object* - oggetto radio quasi stellare - una buffa locuzione che conserva traccia della sorpresa per la loro scoperta negli anni 60.

Le varie classi di eventi

Nel corso di questi anni, l'osservatorio ha scoperto varie classi di eventi candidati ad essere descritti come neutrini cosmici. Ecco la lista delle principali classi di eventi identificati.

1. Degli eventi simili a lunghe tracce, prodotti ancor prima di entrare nel rivelatore e provenienti da sotto l'orizzonte del rivelatore stesso. Questi sono dovuti a neutrini muonici che interagiscono con dei nucleoni scambiando un bosone W ⁹. Dall'interazione vengono prodotti muoni, che generano le lunghe tracce osservate.¹⁰
2. Degli eventi di tipo simile ma che si originano all'interno del rivelatore. Questi vengono prodotti dallo stesso tipo di neutrini (muonici) nel caso in cui il punto di interazione con i nucleoni risulta osservabile.
3. Degli eventi che si originano nel rivelatore sviluppandosi in una regione di spazio sferoidale all'interno della quale si deposita una grande quantità di energia. Questi eventi sono dovuti a neutrini di qualsiasi tipo che interagiscono tramite un W , o anche attraverso interazioni con scambio di particelle Z^0 .¹¹
4. Tre eventi costituiti da due segnali in rapida successione temporale. Questi sono interpretati come neutrini di tipo tau che interagiscono con un nucleone, producendo il primo rilascio di energia visibile ed un leptone

⁹Le particelle W e Z , che caratterizzano il modello standard delle particelle e delle interazioni, sono quelle responsabili delle interazioni deboli, proprio come i fotoni sono responsabili delle interazioni elettromagnetiche. Sono state identificate per la prima volta dall'esperimento UA1 del CERN guidato da Carlo Rubbia.

¹⁰Questo è il primo tipo di segnale di neutrino che è stato preso in esame: venne discusso sin dalla fine degli anni '50 del secolo scorso e venne osservato per la prima volta un decennio dopo.

¹¹Questo tipo di segnale è stato evidenziato per la prima volta proprio da IceCube, ed include i primi 2 eventi rivelati. Vengono detti in gergo HESE e saranno discussi più approfonditamente in seguito. Sottolineiamo che questa classe di eventi può essere ulteriormente ripulita dai neutrini di tipo muonico, quando si riesce ad identificarne le tracce; in questo modo, si migliora la corrispondenza con il tipo di neutrini e si diminuisce il rischio di una contaminazione da segnali spuri.

Oscillazione dei neutrini (approfondimento)

Esistono tre tipi di neutrini, elettronici, muonici e tau, e abbiamo imparato che essi si trasformano gli uni negli altri non appena percorrono distanze pari a

$$L = \frac{E}{\text{TeV}} \times \begin{cases} 0.5 \text{ Gm} \\ 15 \text{ Gm} \end{cases}$$

(Gm=gigametri=milioni di km) corrispondenti ai due valori noti delle differenze di massa quadra dei neutrini. Alle energie $E > \text{TeV}$ che ci riguardano, questo significa che i neutrini di origine atmosferica non oscillano, mentre quelli di origine cosmica oscillano completamente. Dalle misure dei coefficienti di mescolamento dei neutrini concludiamo che quelli cosmici si ripartiscono approssimativamente in uguali proporzioni tra tutti e tre i tipi.

carico τ . Quest'ultimo si allontana dal vertice di produzione e, decadendo, produce il secondo rilascio di energia.

5. Un evento speciale con energia intorno a 6.3 PeV, interpretabile come un antineutrino elettronico che interagisce con un elettrone producendo un W reale. Il processo, chiamato risonanza di Glashow, è stato previsto da Sheldon Glashow nel lontano 1960, prima ancora che il modello standard delle interazioni elementari prendesse del tutto forma.

Le prime tre classi comprendono molte decine di eventi, anche se alcuni non sono dovuti al segnale di neutrini cosmici ma al fondo, e le ultime due pochissimi eventi, ma di grande interesse. Nella schema di fig. 3 ci riferiamo ai primi due tipi di segnali, che sono quelli che ci danno le migliori informazioni sulla direzione di provenienza degli eventi.

Eventi spuri (fondo sperimentale) Come abbiamo già accennato, oltre al segnale dei neutrini

cosmici, esiste una classe di eventi di origine per così dire locale, che è dovuta alla interazione tra raggi cosmici ed atmosfera terrestre. Essa tende a mascherare i neutrini cosmici, specialmente alle basse energie, tanto che l'analisi del segnale sotto il TeV diventa sempre più difficile se non proibitiva.

Si tratta sia di muoni (che tendono a contaminare gli eventi provenienti da sopra l'orizzonte) ma anche di neutrini, che arrivano invece da ogni direzione. Tali neutrini, detti atmosferici, sono composti soprattutto da neutrini muonici, che hanno uno spettro ancora più ripido di quello dei raggi cosmici ($\gamma \approx 3.7$); essi includono inoltre, in misura circa 30 volte inferiore, anche dei neutrini elettronici.

Inoltre, sappiamo dalla teoria che c'è anche una ulteriore componente minoritaria, pure prodotta in atmosfera, ma con una pendenza $\gamma \approx 2.7$ e ugualmente composta da neutrini muonici ed elettronici, dovuta alla produzione di quark di tipo charm. Anche se la teoria non ne prevede con precisione l'intensità attesa, e non ne è stata ancora misurata con sicurezza la consistenza, esistono interessanti indizi che non siamo lontani dall'osservarli per la prima volta.

Cosa abbiamo imparato sui neutrini cosmici

La scoperta Non ci sono praticamente dubbi che sia stata scoperta una nuova popolazione di neutrini, sebbene non confermata con sufficiente sicurezza da altri rivelatori come Antares e Bajkal a causa delle dimensioni molto più piccole di quella di IceCube, e a dispetto della presenza di eventi spuri, che sicuramente contaminano i campioni osservati.

Le oscillazioni Come sappiamo per certo ormai da una ventina di anni, i tre tipi di neutrini noti (elettronico, muonico e tau) sono combinazioni di particelle di massa definita, e quindi non corrispondono a stati stazionari: vedi lo schema di fig. 4. A causa di questo essi si trasformano gli uni negli altri durante la loro propagazione nello spazio, e lo stesso succede anche ai corrispondenti antineutrini (nella terminologia corrente si parla di oscillazione di sapore o più sinteticamente di oscillazione). Questa circostanza

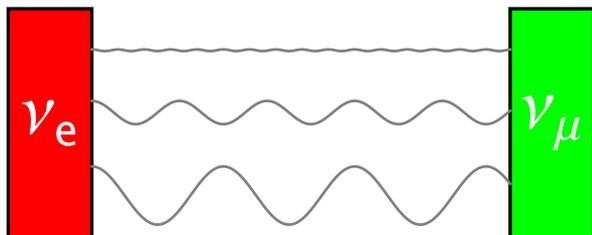


Figura 4: Schema delle oscillazioni: i neutrini inizialmente prodotti (elettronici nell'esempio) sono combinazioni lineari di stati di massa differente. Durante la propagazione dal sito di produzione a quello di rivelazione si accumulano degli sfasamenti, che fanno sì che i neutrini rivelati possano risultare differenti da quelli iniziali (muonici nell'esempio).

comporta importanti conseguenze per il caso in esame, a partire dalla seguente: i neutrini di origine astrofisica, a differenza di quelli atmosferici, hanno oscillato prima di giungere nei nostri rivelatori. I neutrini prodotti sono quasi esclusivamente di tipo elettronico o muonico. Quelli prodotti nell'atmosfera non hanno a disposizione abbastanza tempo per potersi trasformare in neutrini tau, prima di essere rivelati, al contrario di quelli di origine cosmica, che, come discusso nella scheda di approfondimento della pagina precedente, di tempo ne hanno a iosa. Per questo, deve esistere una popolazione di neutrini tau, che è assente nei neutrini atmosferici; e le evidenze di questo tipo (che per adesso sono limitate dalla statistica) devono essere considerate una dimostrazione dell'origine cosmica dei nuovi neutrini osservati.

La distribuzione angolare complessiva Un altro punto di importanza capitale è la distribuzione angolare degli eventi, che testimonia come la gran parte di quelli non-atmosferici non correli fortemente con il piano galattico, suggerendo invece che siano, in prima approssimazione e per quanto si riesce a vedere, uniformemente distribuiti. Questo porta a credere che il segnale scoperto da IceCube sia di origine extra-galattica.¹²

¹²Da un punto di vista formale, non possiamo escludere in modo assoluto che parte di questi eventi sia originata nell'alone galattico, e dunque abbiano natura diffusa. D'altro canto, queste interpretazioni alternative sono state proposte dopo la scoperta dei neutrini cosmici e non prima. In effetti, somigliano a quelle che vennero avanzate dopo la scoperta di nuove popolazioni di

La statistica raccolta e le caratteristiche di un telescopio di neutrini non consentono affermazioni molto precise, nonostante ciò i dati di IceCube danno adito a qualche dubbio sulla perfetta isotropia. Essi suggeriscono una qualche asimmetria alto-basso, di cui si ignora la causa. Allo stato dell'arte, le ipotesi più plausibili sono tre: un effetto fisico, come una piccola componente galattica (che dovrebbe venire soprattutto dal cielo nord); la presenza di una qualche contaminazione residua, dovuta ai neutrini atmosferici; una fluttuazione statistica dei dati. Purtroppo, tali ipotesi non sono esclusive tra di loro. Non ci sono dubbi però che lo studio della distribuzione angolare meriti indagini più approfondite.

La distribuzione in energia Allo stato dell'arte, gli spettri di energia delle classi di eventi statisticamente più significative, vengono di regola descritti con leggi di potenza. Un problema che però emerge abbastanza evidentemente è che i vari risultati si accordano tra loro solo parzialmente. In particolare, lo spettro che viene dal campione inclusivo di eventi che hanno origine nel rivelatore, detto in gergo HESE (dall'acronimo inglese *high-energy starting events*) indica uno spettro decisamente ripido, non troppo lontano da $\gamma = 2.7$; invece i neutrini muonici di energia più alta non sono così lontani dal valore $\gamma = 2$; e altri campioni di dati suggeriscono valori intermedi della pendenza. Le oscillazioni di neutrino implicano che lo spettro dei vari tipi di neutrino sia lo stesso, si pone quindi il problema di capire se esista un unico spettro in grado di descrivere allo stesso tempo i vari sottoinsiemi di dati. Inoltre, questi dati sollevano alcune specifiche domande, ad esempio le seguenti.

- * Se lo spettro è davvero ripido come suggerito dagli eventi HESE, come evitare di produrre una popolazione di gamma associati ai neutrini, che si manifesterebbe alle energie più basse, che in effetti non vediamo?
- * Riteniamo che i neutrini associati alla produzione di quark charmati abbiano una di-

oggetti astrofisici, come i quasar o i lampi di raggi gamma. Sono utili inviti alla cautela, ma per ora non meritano tanta attenzione quanto quelle descritte nel testo.

istribuzione con spettro $E^{-2.7}$. IceCube, tramite lo studio degli eventi di tipo muoni che provengono dal basso, ha fornito dei limiti stringenti su di essi. Ma allora, perché mai lo stesso limite non si dovrebbe applicare agli eventi di tipo HESE?¹³

Queste domande al momento non hanno risposte certe. Sono state proposte alcune vie d'uscita, per esempio nello studio di Ref. [6] che si basa sull'ipotesi di collisioni tra raggi cosmici ed adroni e prevede degli spettri non troppo diversi dal caso $\gamma \sim 2$. Sarebbe desiderabile usare campioni di dati più puliti possibile, e ben caratterizzati dal punto di vista del tipo di neutrino; inoltre, se non si desidera deviare troppo da una ipotesi minimale per interpretare i dati, converrebbe procedere ad analisi globali dei dati disponibili, anziché esaminarli uno per uno, separatamente.

Verso l'astronomia con i neutrini di altissima energia

C'è stato un evento che IceCube ha visto nella direzione di un particolare blazar, TXS 0506+056, che è a circa 1.75 Gpc dalla Terra, e che è stato contemporaneamente registrato da un certo numero di osservatori gamma: i risultati sono descritti nella referenza [7]. Ci sono risultati contrastanti sulla correlazione con delle popolazioni, per esempio, Troitsky e collaboratori [5] hanno recentemente avanzato il caso di una interessante correlazione, con significatività di 3σ , tra le posizioni degli eventi osservati da IceCube ed i dati del catalogo radio RFC <http://astrogeo.org/rfc>. In precedenza, simili analisi non avevano trovato niente di altrettanto interessante.

Discussione e prospettive di sviluppo

Le osservazioni di IceCube hanno dato un enorme impulso al campo dei neutrini di altissima energia, che oggi è uno di quelli più vivaci della moderna fisica astroparticellare. Sono state

¹³Detto in termini più formali: se vale l'isotropia e se si applicano gli effetti delle oscillazioni di neutrino, il limite sulla componente di neutrini prodotti da quark charmati, ottenuto da IceCube studiando gli eventi di tipo traccia sotto 200 TeV, porterebbe ad escludere l'ipotesi che i dati HESE siano dovuti a neutrini cosmici con spettro del tipo $E^{-2.7}$.

raccolte molte informazioni interessanti e molti indizi, che non sembrano ancora formare un quadro chiaro. Se si volesse assumere un atteggiamento distaccato se non addirittura critico, si potrebbe arrivare a sostenere che l'informazione principale che abbiamo ottenuto riguarda lo spettro dei neutrini, che per adesso ci aiuta solo fino ad un certo punto ad investigare il mistero dell'origine dei raggi cosmici. Ma dall'altro lato, ci sono pochissimi dubbi che si sia fatta una grossa scoperta, e ad essere obiettivi, il decennio appena trascorso ha registrato enormi passi avanti.

Tutto porta a credere che il quadro interpretativo verrà chiarito nei prossimi anni, ed è plausibile che le successive osservazioni dei neutrini di altissima energia (in sinergia con quelle che riguardano i raggi gamma) consentiranno ulteriori significativi avanzamenti nella conoscenza dei raggi cosmici.

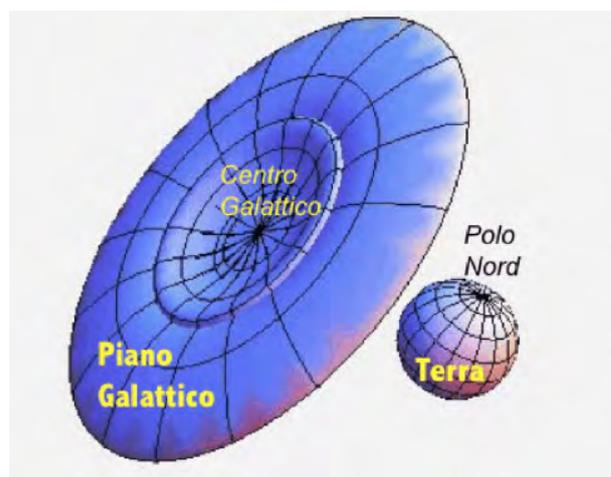


Figura 5: L'orientamento relativo della Terra e della via Lattea nello spazio fa sì che i telescopi di neutrino dell'emisfero nord siano favoriti nell'indagine della nostra galassia per mezzo di eventi di tipo traccia (muoni di altissima energia).

Il futuro sembra ricco di opportunità - o almeno, il lavoro ancora da svolgere non manca. È certo che IceCube abbia ancora molto da dare al campo, e sono ammirabili i risultati ottenuti dai rivelatori subacquei Antares e Baikal nell'emisfero nord, che pure sono molto più piccoli di IceCube. Diventa sempre meno procrastinabile l'esigenza di disporre di un rivelatore di classe analoga ad IceCube nel nostro emisfero.

Un tale rivelatore consentirà per prima cosa di effettuare verifiche cruciali riguardo ai risultati

ottenuti. Poi, come evidenziato nelle figure 3 e 5, saremo in posizione assai più vantaggiosa per indagare l'emissione dalla nostra galassia: il buco nero centrale, noto come Sgr A*, ha $\delta = -29^\circ$, ovvero ha una declinazione di 29 gradi sud; similmente, RX J117-3946 ha $\delta = -39^\circ$ e Vela Jr ha $\delta = -46^\circ$. Inoltre, è ragionevole aspettarsi che gli eventi interessanti siano correlati al piano galattico, con conseguente riduzione del numero di eventi spuri, dovuti a fondo o a segnali concomitanti.

Per le ragioni discusse precedentemente, la regione di energie più interessante non sembrerebbe quella delle energie più estreme, ma piuttosto quella sotto qualche centinaio di TeV, che oggi non capiamo così bene.

Inoltre, e soprattutto, i rivelatori che invece del ghiaccio useranno l'acqua, come KM3NeT nel Mediterraneo [8] GVD nel lago Bajkal [9], o il rivelatore P-ONE nel Pacifico recentemente proposto [10], raggiungeranno risoluzioni angolari migliori di quella di IceCube di quasi un ordine di grandezza – e dunque potranno contare su una maggiore capacità di reiezione del segnale di fondo e di identificazione delle sorgenti astrofisiche. Questo sarà di importanza dirimente per identificare le singole sorgenti.

Concludiamo ricordando che esistono anche progetti concreti per una estensione significativa di IceCube stesso, e anche rivelatori di diversa concezione, che mirano soprattutto all'esplorazione delle energie più alte, nella regione 10-100 PeV ed oltre ancora: una regione della quale non sappiamo quasi nulla al momento e dalla quale sembra legittimo aspettarsi ulteriori sorprese.

Una lista di domande

Da ultimo, riportiamo una lista di domande presentate da F. Vissani, in qualità di presidente della sessione di discussione sui neutrini di altissima energia, al convegno (tenuto a distanza durante la prima emergenza Covid) "Cosmic Rays and Neutrinos in Paris" nel gennaio del 2021, nella speranza che possano stimolare il lettore a ragionarci sopra, impostandole la discussione, e magari a proporre qualche buona risposta o a arricchire ulteriormente la lista.

- Siamo tutti d'accordo che la scoperta di una popolazione inattesa di neutrini sia già stata fatta?
- Ai fini di effettuare i controlli definitivi, ulteriori verifiche dei risultati sperimentali di IceCube sono necessarie o sono semplicemente auspicabili?
- Nel caso siano sentite come necessarie, quali sono le quantità cruciali che dovremmo verificare con la massima priorità?
- Quanto è essenziale migliorare la risoluzione angolare?
- Quanto sono convincenti i casi di sorgenti adroniche desunti dai dati dei raggi gamma?
- Le aspettative di oscillazioni a 3 sapori sono del tutto credibili ed affidabili per i neutrini di origine cosmica?
- I vari campioni di dati di IceCube rientrano in un'interpretazione coerente?
- Quanto è importante la ricerca di neutrini da quark charmati?
- È possibile verificare la regione del ginocchio dei raggi cosmici?
- L'origine astrofisica della nuova componente è un'ipotesi definitiva o almeno convincente?
- Quali sono le sorgenti più promettenti di neutrini? Quali sono i prossimi passi che suggeriscono di intraprendere?
- Quanto sono grandi le incertezze teoriche?
- Cosa sappiamo con certezza della nuova componente vista da IceCube?
- Abbiamo una qualche evidenza di uno spettro di potenza in qualche intervallo di energia?
- Ci sono caratteristiche fisiche speciali a basse/alte energie?
- La distribuzione angolare è completamente isotropa?
- Quanto è cruciale disporre di un telescopio di neutrini nell'emisfero nord?
- Come confrontare tra di loro segnali indipendenti e dipendenti dal tempo?
- L'approccio multi-messaggero è buono di per sé o è necessaria una teoria/modello per porre le domande migliori e inquadrare le risposte?
- Cosa sarebbe più utile per corredare l'astronomia di neutrini: poter contare su ulteriori studi di raggi cosmici? Una ricerca a tappeto di fotoni nella regione 0.01-1 PeV? Altre informazioni dall'astronomia X, IR, UV o magari proprio da quella ordinaria?

nome	tipo	luogo	stato	riferimento
ICECUBE	neutrini	Polo Sud	operativo	https://icecube.wisc.edu/
ANTARES	neutrini	Mediterraneo	concluso	https://antares.in2p3.fr/
KM3NET	neutrini	Mediterraneo	in costruzione	https://www.km3net.org/
BAIKAL-GVD	neutrini	Lago Bajkal	in costruzione	https://baikalgvd.jinr.ru/
P-ONE	neutrini	Bacino Cascadia	futuro	https://www.pacific-neutrino.org/
LHAASO	gamma	Sichuan	operativo	http://english.ihep.cas.cn/lhaaso/
CTA	gamma	La Palma; Atacama	in costruzione	https://www.cta-observatory.org/

Tabella 1: I rivelatori di neutrini e di raggi gamma menzionati nel presente lavoro e alcune delle loro principali caratteristiche.

Lecture consigliate

Inseriamo qui una lista, incompleta, di testi per ulteriori approfondimenti.

- Un'introduzione generale all'astronomia di neutrino, di una dozzina di anni fa ma ancora valida: G. Pagliaroli, F. L. Villante, F. Vissani, *Neutrini dallo spazio (discutendo gli obiettivi di una nuova astronomia)*, Nuovo Saggiatore 25 (2009), 5. Per qualche aggiornamento, si vedano A. Gallo Rosso, C. Mascaretti, A. Palladino and F. Vissani, *Introduction to neutrino astronomy*, Eur. Phys. J. Plus 133 (2018) 267, arXiv:1806.06339 e la monografia pubblicata dalla Springer (2018) *Multiple Messengers and Challenges in Astroparticle Physics*, che discute di onde gravitazionali, raggi cosmici, fotoni, neutrini e contiene in particolare il capitolo *Neutrino Astronomy* di A. Capone, P. Lipari and F. Vissani. Un lavoro di rassegna più recente per approfondire è: A. Palladino, M. Spurio and F. Vissani, *Neutrino Telescopes and High-Energy Cosmic Neutrinos*, Universe 6 (2020) no.2, 30, arXiv:2009.01919.
- Alcuni fra i risultati più significativi di IceCube: *Evidence for High-Energy Extraterrestrial Neutrinos with IceCube*, Science 342 (2013) 1242856 in cui viene riportata la prima evidenza di neutrini di origine cosmica; la referenza [7] in cui viene presentata per la prima volta la possibile emissione di neutrini da una sorgente astrofisica; *Detection of a particle shower at the Glashow resonance with IceCube*, Nature 591 (2021) 220 in cui viene discussa la rivelazione del primo evento consistente con la risonanza di Glashow.

- La descrizione degli obiettivi dell'esperimento KM3NeT può essere reperita presso la referenza [8].

RINGRAZIAMENTI

F.V. ringrazia il MIUR per la borsa 017W4HA7S "NAT-NET: Neutrino and Astroparticle Theory Network", PRIN 2017.



- [1] S. P. Swordy: *Space Science Reviews*, The Energy Spectra and Anisotropies of Cosmic Rays, 99 (2001) 85.
- [2] V.L. Ginzburg, S. I. Syrovatskii: *The Origin of Cosmic Rays*, Macmillan, New York (1964).
- [3] Z. Cao, F. Aharonian, Q. An, et al.: *Ultrahigh-energy photons up to 1.4 petaelectronvolts from 12 γ -ray Galactic sources*, Nature, 594 (2021) 33.
- [4] E. Waxman, J. N. Bahcall: *High-energy neutrinos from astrophysical sources: An Upper bound*, Phys. Rev. D, 59 (1999) 023002.
- [5] A. V. Plavin, Y. Y. Kovalev, Yu. A. Kovalev, S. V. Troitsky: *Directional association of TeV to PeV astrophysical neutrinos with radio blazars*, Astroph. J., 908 (2021) 151.
- [6] Yu-Chen Ding et al.: *Prospects of detecting dark matter through cosmic-ray antihelium with the antiproton constraints*, Journal of Cosmology and Astroparticle Physics, 08 (2019) 004.
- [7] ICECUBE collaboration: *Neutrino emission from the direction of the blazar TXS 0506+056 prior to the IceCube-170922A alert*, Science, 361 (2018) 147.
- [8] S. Adrián-Martínez et al.: *Letter of intent for KM3NeT 2.0*, J. Phys. G: Nucl. Part. Phys., 43 (2016) 084001.
- [9] A. D. Avrorin et al.: *High-Energy Neutrino Astronomy and the Baikal-GVD Neutrino Telescope*, Phys. Atom. Nuclei, 84 (2021) 513.
- [10] M. Agostini et al.: *The Pacific Ocean Neutrino Experiment*, Nat. Astron., 4 (2020) 913.

Carla Distefano: laureatasi in fisica all'Università degli Studi di Catania nel 2000, consegue successivamente il titolo di dottore di ricerca nello stesso ateneo. Dal 2007 è ricercatrice presso i Laboratori Nazionali del Sud dell'INFN presso i quali si occupa di fisica del neutrino e neutrino astronomia. Svolge principalmente la sua attività di ricerca nell'ambito del progetto KM3NeT di cui coordina le attività di simulazione Monte Carlo. Appassionata di divulgazione scientifica, dal 2016 coordina il gruppo di Public Engagement dei Laboratori Nazionali del Sud.

Francesco Vissani: ha ricevuto il PhD alla SISSA di Trieste nel 1994 e si interessa da molti anni dei neutrini in fisica ed astrofisica. Negli ultimi 20 anni circa ha lavorato sugli aspetti teorici e fenomenologici relativi ai neutrini di altissima energia, assieme a molti colleghi, inclusi Felix Aharonian, Antonio Capone, Arman Esmaili, Paolo Lipari, Chiara Righi, Esteban Roulet, Maurizio Spurio, Fabrizio Tavecchio e Francesco Villante e seguendo le tesi di dottorato sull'argomento di Silvia Celli, Maria Laura Costantini, Carlo Mascaretti, Andrea Palladino e Narek Sahakyan. Trova che l'argomento stia diventando sempre più stimolante ed è convinto che non siamo lontani da un nuovo punto di svolta. È un forte fautore dell'importanza della divulgazione scientifica nel proprio paese e per questo ha fondato e coordina il premio ASIMOV che cura assieme a moltissime scuole ed amici.

Astroparticelle di altissima energia

La notte fu esplorata più del giorno perché era tanto più vasta. Il pensiero ha forzato i segreti, scippato conoscenze per allargare il campo della poca vita. Sbirciare l'infinito fa aumentare lo spazio, il respiro, la testa, di chi lo sta ad osservare.

Erri de Luca

Antonella Castellina

Osservatorio Astrofisico di Torino (INAF-OATO), Torino, Italy
INFN, Sezione di Torino, Torino, Italy

La fisica astro-particellare è un campo interdisciplinare giovane, che studia la radiazione e le particelle del cosmo utilizzando tecniche tipiche degli esperimenti di alta energia. Alcune di queste particelle cosmiche raggiungono energie incredibilmente elevate, ed il loro studio è uno degli argomenti più affascinanti dell'astrofisica moderna. Questi messaggeri ultra-energetici dallo spazio profondo possono raggiungere la Terra ed essere osservati con apparati sperimentali dedicati e ci forniscono informazioni ineguagliabili sulle loro sorgenti e sulla struttura dell'Universo.

Introduzione

Ogni giorno, in ogni istante, dal cielo piovono su di noi argento, oro, idrogeno, carbonio, ferro e altri elementi. Bombardano la Terra costantemente, da tutte le direzioni; possono provenire dal Sole, dalla nostra Galassia, dallo spazio extragalattico. Chiamati "radiazione dall'alto" dal loro scopritore Viktor Hess nel 1912, vennero

ribattezzati "raggi cosmici" da Robert Millikan, ma la loro natura è corpuscolare: sono in realtà particelle subatomiche, dai protoni ai nuclei più pesanti, e sarebbe più corretto definirle dunque astroparticelle, benchè il termine raggi cosmici sia ancora largamente utilizzato. È facile comprendere che maggiore è l'energia che devono ricevere al momento della produzione, maggiore sarà la difficoltà di produrli: in effetti, il flusso di raggi cosmici che arriva sulla Terra diminuisce molto rapidamente al crescere della loro energia. Su un m² di atmosfera, arrivano circa 10000 particelle/secondo con energia pari a circa 10⁹ eV¹, ma solo 10 particelle/ora se la loro energia è di 10¹² eV, fino a numeri incredibilmente bassi come 1 particella/km² ogni secolo alle massime energie misurate. Misure dirette dei raggi cosmici, ad esempio con apparati installati su palloni o su satelliti, sono possibili solo fino ad energie dell'ordine dei 10¹³ – 10¹⁴ eV; poi, a causa del loro bassissimo

¹l'elettronvolt, eV, è l'energia guadagnata (o persa) da un elettrone che si muove nel vuoto attraversando una differenza di potenziale di 1 Volt. Nel seguito, indicheremo spesso le energie in unità di exaelettronvolt: 1 EeV = 10¹⁸ eV.

flusso, le particelle devono essere registrate con installazioni di grandissima area, dunque poste a Terra, e studiate in modo indiretto attraverso i secondari che esse generano in atmosfera.

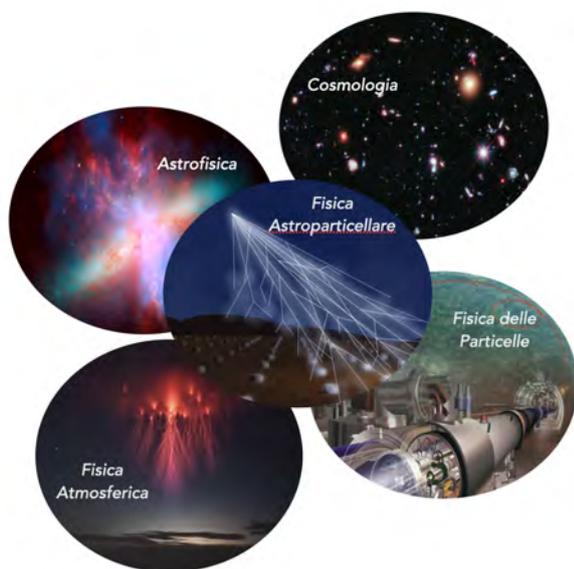


Figura 1: Le diverse connessioni della fisica astroparticellare con altri campi di ricerca.

I raggi cosmici di altissima energia (comunemente indicati con l'acronimo UHECR: *ultra-high energy cosmic rays*), definiti come particelle cosmiche di energia superiore a $\sim 10^{17}$ eV, rivestono un interesse particolare per la Fisica Astroparticellare [1].

Come evidenziato in Fig.1, questo campo è strettamente legato alla cosmologia ed all'astrofisica. Alle energie estreme in gioco, gli UHECR non possono più restare confinati nella nostra Galassia, e dunque devono avere origine extragalattica; non sappiamo ancora quali sorgenti cosmiche siano in grado di produrli ed accelerarli a valori così estremi. Lo studio degli UHECR riveste enorme importanza anche dal punto di vista della fisica fondamentale. È infatti l'unico modo di capire come interagiscono le particelle elementari ad energie altissime, non raggiungibili ai moderni acceleratori ed anche di esplorare possibili effetti fisici non previsti dal Modello Standard. Come vedremo, riuscire a registrare i secondari prodotti dagli UHECR richiede una profonda conoscenza dell'atmosfera terrestre; questo fornisce la possibilità di estendere le nostre conoscenze su effetti come i TGF (*Terre-*

trial Gamma Flashes) o gli ELVES, eventi luminosi che si manifestano in concomitanza con i fulmini.

Caccia alle sorgenti

Comprendere la natura delle sorgenti degli UHECR è uno dei problemi aperti dell'astrofisica particellare moderna. Nei modelli, detti *bottom-up*, le particelle vengono accelerate in ambienti astrofisici con meccanismi probabilmente legati a processi dovuti agli intensi campi magnetici ed elettrici presenti nelle sorgenti. Le possibili sorgenti devono garantire un certo numero di condizioni:

- la particella deve poter rimanere all'interno della sorgente mentre viene accelerata;
- la sorgente deve possedere la quantità di energia necessaria per accelerare le particelle ad energie così elevate;
- l'energia persa da una particella per radiazione nel campo accelerante oppure per interazione con le altre particelle presenti non deve superare il guadagno di energia;
- il numero totale e la potenza delle sorgenti deve essere in grado di fornire il flusso di UHECR osservato;
- i fotoni ed i neutrini generati nella sorgente candidata oppure durante la propagazione non devono superare i flussi osservati.

È stato ipotizzato, per esempio, che le particelle possano essere direttamente accelerate ad altissima energia in stelle di neutroni o pulsar, o in oggetti come le nane bianche, che sono altamente magnetizzate e in rapida rotazione. L'accelerazione sarebbe in questo caso rapidissima, ma gli UHECR perderebbero molta dell'energia acquistata prima di poter abbandonare la sorgente e cominciare il loro viaggio interstellare, a causa dell'alta densità di radiazione e di materia presente nell'ambiente della sorgente.

Un'altra possibilità è che l'energia cresca gradualmente a causa di incontri multipli con plasma magnetizzato, come ipotizzato da Enrico Fermi, oppure, più efficientemente, che le particelle siano accelerate da onde d'urto che si propagano nel mezzo interstellare. Questo tipo di accelerazione potrebbe avvenire per esempio nel disco della nostra Galassia, in radio Galassie giganti, o in nuclei galattici attivi (AGN) che emettono getti relativistici di materia nella direzione

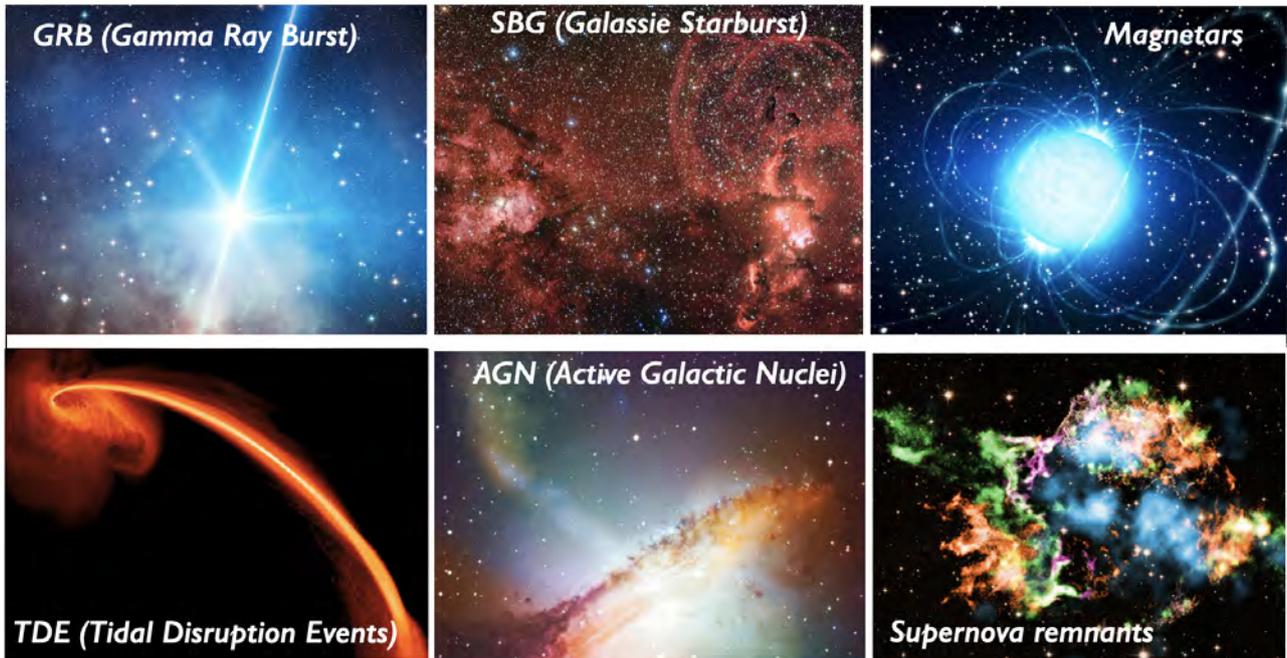


Figura 2: Possibili candidati a sorgenti di astroparticelle di altissima energia.

ne dell'osservatore. In Fig.2 sono mostrati alcuni affascinanti esempi di possibili sorgenti di UHECR.

Un'origine più esotica è quella proposta dai modelli *top-down*, secondo i quali gli UHECR possono essere prodotti nel decadimento di particelle super pesanti con masse superiori a $\sim 10^{21}$ eV. Queste particelle possono essere metastabili, prodotte direttamente nell'Universo primordiale (materia oscura super pesante), oppure possono essere emesse da difetti topologici, prodotti attraverso una transizione di fase nel primo Universo. In questo scenario, i componenti principali del flusso di particelle UHE sono fotoni e neutrini da decadimenti di pioni neutri e carichi con solo un 10% di nucleoni.

Finora, non sono state trovate sorgenti individuali di UHECR. È possibile che le sorgenti siano tante ma deboli, oppure che i campi magnetici che permeano l'Universo siano così intensi da deviare considerevolmente le direzioni d'arrivo delle particelle cariche, e quindi rendere impossibile correlarne la direzione d'arrivo con l'origine. In effetti, conosciamo abbastanza bene il campo magnetico Galattico, modellato sulla base delle osservazioni con una componente regolare e una turbolenta, dell'ordine di pochi μG . Tutte le particelle cariche di energia inferiore a circa 1 EeV sono confinate nella Galassia, poichè il loro raggio

di curvatura nel campo magnetico risulta sempre molto inferiore allo spessore tipico del disco della Galassia (~ 300 pc)². Tuttavia non abbiamo modelli basati su osservazioni sperimentali per quanto riguarda il campo magnetico extragalattico, per cui gli studi si basano su simulazioni della formazione delle strutture cosmologiche. Questi campi possono avere intensità molto variabili, a seconda che li si calcoli per vuoti cosmici, filamenti, o altre strutture extragalattiche, con valori tra 10^{-17} e 10^{-9} G.

La propagazione

L'Universo è permeato di fotoni di diverse lunghezze d'onda. La radiazione termica di fondo (CMB, *cosmic microwave background*) è una radiazione elettromagnetica proveniente dall'epoca della ricombinazione, quando dal plasma primordiale nacquero i primi atomi e l'universo divenne trasparente alla luce.

La radiazione extragalattica di fondo (EBL, *extragalactic background light*) è invece la luce emessa da tutte le galassie nel corso della storia dell'Universo, dalle lunghezze d'onda dell'infrarosso fino al visibile e ultravioletto. Durante il loro viaggio nello spazio intergalattico, gli UHECR

²1 parsec equivale a 3.26 anni luce, o 3.086×10^{13} km.

Le interazioni durante la propagazione degli UHECR

Lo spazio è permeato di fotoni di diversa origine e diverse lunghezze d'onda, come mostrato nella figura.

L'interazione tra un nucleo di energia E , massa m (e dunque con fattore di Lorentz $\Gamma = E/m$) e massa atomica A ed un fotone del fondo cosmico di energia ϵ potrà avvenire, in base alle energie in gioco, per:

(a) produzione di coppie:

$N + \gamma \rightarrow N + e^+ + e^-$, per $E/A \gtrsim 0.2$ EeV.

(b) foto-produzione di pioni da parte dei protoni:

$p + \gamma \rightarrow p + \pi^0$, oppure $p + \gamma \rightarrow n + \pi^+$.

Avviene per $E/A \gtrsim 4$ EeV oppure $E/A \gtrsim 40$ EeV se sui fotoni EBL o CMB rispettivamente

(c) foto-disintegrazione del nucleo:

${}^A Z + \gamma \rightarrow {}^{A-1} Z + n$, o ${}^A Z + \gamma \rightarrow {}^{A-1} (Z-1) + p$.

Avviene per $E/A \gtrsim 20$ EeV se sui fotoni del CMB, oppure ad un valore 10 volte inferiore se sui fotoni EBL.

In particolare già nel 1965, poco dopo la scoperta del CMB, Greisen, Zatsepin e Kuzmin predissero l'esistenza di un *cut-off* nello spettro degli UHECR, causato dall'interazione dei protoni con i fotoni del CMB. Poiché l'energia tipica dei γ_{CMB} è $\simeq 10^{-3}$ eV, si può calcolare che l'energia a cui si osserva il *cut-off*, ormai conosciuto universalmente come il **cut-off GZK**, sia attorno ai $5 \cdot 10^{19}$ eV.

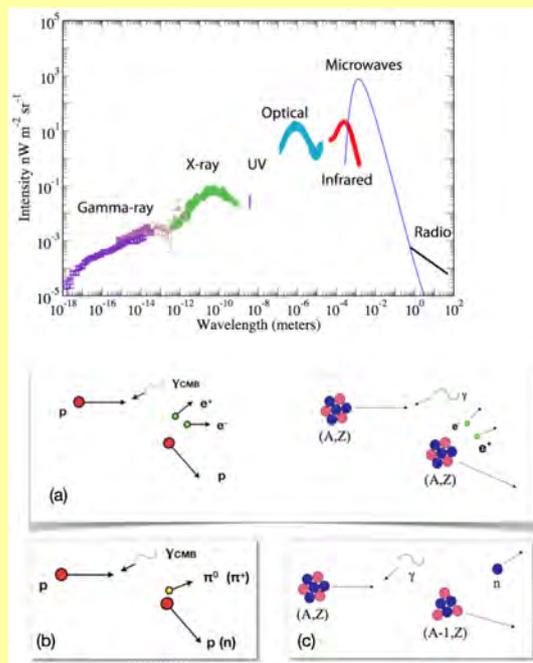


Figura 3: Il fondo di radiazione nello spazio extragalattico (in alto) e uno schema delle diverse interazioni che avvengono durante la propagazione.

possono interagire con questi fotoni, perdendo energia oppure dando origine a nuove particelle.

Ad energie inferiori a qualche EeV per i protoni e a qualche decina di EeV per i nuclei, le particelle perdono energia a causa dell'espansione dell'Universo. Al crescere dell'energia, nell'interazione possono essere prodotte coppie di elettroni e positroni. I protoni possono foto-produrre pioni carichi o neutri, mentre i nuclei subiscono processi più complessi di foto-disintegrazione. In quest'ultimo caso, una descrizione corretta della propagazione richiede un'incursione nel campo della fisica nucleare: occorre infatti conoscere le sezioni d'urto per le interazioni inelastiche e le sezioni d'urto inclusive.

La soppressione del flusso dei raggi cosmici alle energie più estreme è un dato osservativo molto solido, come vedremo, ma la sua interpre-

tazione è invece ancora oggi molto dibattuta: una delle possibilità è quella di attribuirlo al *cut-off* GZK, cioè alle perdite di energia dovute alla foto-produzione dei pioni da parte dei protoni per interazione sui fotoni del fondo cosmico di microonde. A causa dell'effetto GZK, maggiore è l'energia degli UHECR, minore sarà la distanza dalla quale possono arrivare: ad esempio, la stragrande maggioranza degli UHECR con energia superiore a ~ 60 EeV può arrivare fino a noi solo da distanze inferiori ai 150-200 Mpc. Dal punto di vista della ricerca delle sorgenti, questo è positivo: significa infatti che le sorgenti più lontane non contribuiscono al flusso osservato, rendendo più evidenti quelle vicine. Tuttavia, si potrebbe anche immaginare che esista un limite all'energia che le sorgenti possono dare agli UHECR: questo effetto potrebbe essere presen-

te insieme all'effetto GZK, oppure addirittura risultare quello dominante.

Misurare le astroparticelle

La misura dei raggi cosmici di altissima energia è possibile da Terra solo in modo indiretto: non si registrano direttamente i primari, ma soltanto i cosiddetti sciame estesi atmosferici (EAS, dall'inglese *Extensive Air Showers*) [2]. Un EAS è una cascata di particelle secondarie generate dall'interazione di un primario con un nucleo atmosferico, tipicamente tra 15 e 20 km di altezza.

La cascata si sviluppa in atmosfera fino a raggiungere un massimo numero di particelle secondarie; quando la loro energia non è più sufficiente a generare nuove particelle, lo sciame si attenua dissipando la sua energia per ionizzazione. La parte centrale dell'EAS è formata da adroni

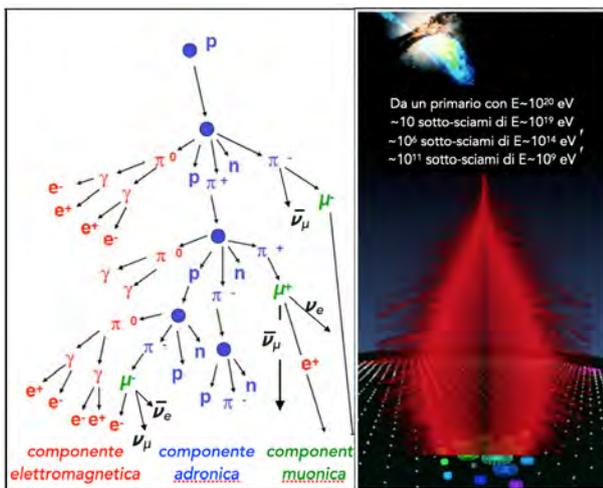


Figura 4: Schema di uno sciame atmosferico esteso. A lato, simulazione di uno sciame.

di alta energia, che attraverso la produzione ed il decadimento di pioni generano la componente elettromagnetica e la componente muonica. La prima è la più numerosa; è formata da fotoni, elettroni e positroni e trasporta circa l'85% dell'energia del primario. Dal decadimento dei pioni carichi vengono generati neutrini e muoni; questi ultimi, particelle circa 200 volte più pesanti degli elettroni, costituiscono la componente più penetrante dello sciame.

Oltre alla produzione di un enorme numero di particelle, gli sciame atmosferici generano anche diversi tipi di radiazione elettromagnetica.

I secondari carichi degli sciame eccitano le molecole di azoto nell'atmosfera che a loro volta, diseccitandosi, emettono una luce di fluorescenza, nella banda del vicino ultravioletto. L'efficienza di fluorescenza, cioè il rapporto tra l'energia emessa in fluorescenza dal gas eccitato e l'energia depositata dalle particelle cariche, è molto bassa, dell'ordine di $\sim 5 \times 10^{-5}$. Questa è la ragione per cui, solo alle energie più alte ($> 10^{17}$ eV), l'enorme numero di particelle permette di rilevare questa luce.

Si può inoltre avere emissione elettromagnetica nelle frequenze radio, attraverso due meccanismi diversi. Il più importante è quello geomagnetico, dovuto al moto di elettroni secondari e positroni degli sciame accelerati nel campo magnetico terrestre. Il secondo meccanismo, noto come effetto Askaryan, si basa sul fatto che l'aria è ionizzata dalle particelle dell'EAS e dunque si ha un eccesso di carica negativa sul fronte dello sciame; la variazione di carica nel tempo genera l'emissione.

Possiamo a questo punto definire le caratteristiche di un buon osservatorio per lo studio degli UHECR:

- a causa del loro bassissimo flusso, gli UHECR possono essere misurati con alta statistica soltanto installando un gran numero di rivelatori su vastissime aree;
 - la separazione tra i rivelatori deve essere regolata in modo da corrispondere all'estensione dello sciame al livello di osservazione (si va dalle decine di metri, per sciame di $E \sim 10^{15}$ eV, ai km, per gli array che studiano la regione di energia più estrema);
 - l'energia massima è limitata dall'area totale dell'esperimento, perché la statistica raccolta è proporzionale all'esposizione (cioè alla superficie per l'accettazione in angolo solido per il tempo di misura).
- Infine, è importante registrare quante più osservabili possibili, per massimizzare l'informazione e ridurre le incertezze sistematiche nelle misure.

L'Osservatorio Pierre Auger

Nel 1992, James Cronin e Alan Watson proposero di costruire un enorme insieme di rivelatori per studiare l'origine dei raggi cosmici delle più

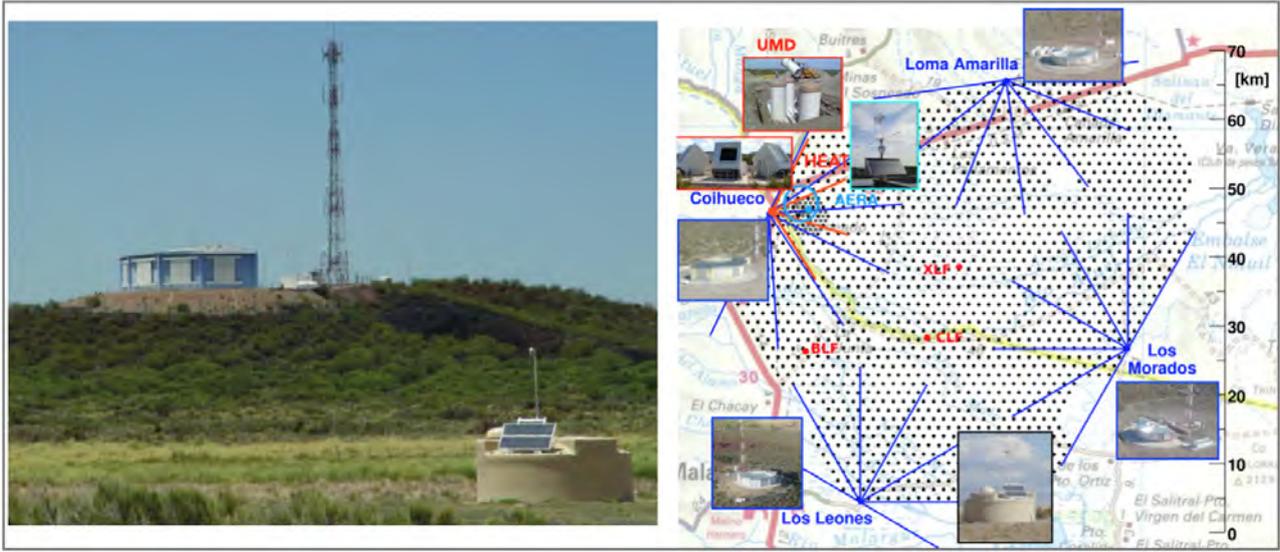


Figura 5: Un rivelatore di superficie nella Pampa; sul fondo sono visibili uno dei 4 siti con 6 telescopi di fluorescenza e la torre di comunicazione. A destra, lo schema dell'Osservatorio Pierre Auger coi suoi diversi rivelatori; ogni puntino nero corrisponde ad una stazione SD, le linee blu indicano il campo di vista dei telescopi nei 4 siti [3].

alte energie, i meccanismi della loro accelerazione, la loro energia e composizione di massa. La proposta trovò il sostegno e l'incoraggiamento della comunità scientifica e si realizzò nella costruzione dell'Osservatorio Pierre Auger in una vasta area della Pampa Argentina, presso la città di Malargüe. L'Osservatorio è ad oggi il più grande esperimento di fisica astroparticellare per le altissime energie mai realizzato [3].

Per poter ottenere il maggior numero possibile di informazioni, nell'Osservatorio si utilizza un sistema ibrido di rivelatori. Le 1600 stazioni Čerenkov ad acqua, poste a distanza mutua di 1.5 km su una griglia triangolare che copre 3000 km² di area, ne costituiscono il rivelatore di superficie (SD1500)³. Si basano sull'emissione di luce Čerenkov che le particelle relativistiche secondarie producono attraversando le 12 tonnellate di acqua contenute in ognuno dei rivelatori. Quattro siti posti attorno a quest'area, ciascuno con 6 telescopi di fluorescenza con angolo di elevazione da 1° a 30°, ne formano il rivelatore di fluorescenza (FD), per la misura della luce emessa dagli sciame attraversando l'atmosfera.

Nella parte nord-occidentale del sito sperimentale, 61 SD sono distribuiti in una griglia più densa con una distanza tra i rivelatori di 750 m (SD750), insieme a tre telescopi a fluorescenza

ad alta elevazione (HEAT), da 30° a 60°. Questa regione viene sfruttata per ampliare la regione energetica esplorabile fino a circa 10^{16.5} eV. Rivelatori a scintillazione con fibra ottica sono interati a fianco delle stazioni di SD750 per la misura diretta dei muoni degli sciame. L'Auger Engineering Radio Array (AERA) completa le misurazioni degli sciame di bassa energia rilevando la loro emissione radio grazie ad un array di più di 150 antenne. Diversi dispositivi di monitoraggio atmosferico sono inoltre impiegati per monitorare l'atmosfera. Mentre l'SD è operativo 24 ore al giorno, tutto l'anno, l'FD può funzionare solo nelle notti serene e senza luna, con un duty cycle pari a ~ 14%.

L'Osservatorio Pierre Auger è situato ad un'altitudine di 1400 m sopra il livello del mare, che corrisponde (per uno sciame verticale) a 750 g/cm². Nella Fig. 6 è schematizzata la ricostruzione di un evento ibrido in Auger cioè uno sciame che viene rilevato simultaneamente dall'FD e dall'SD. L'altitudine atmosferica viene misurata dal livello del mare ($h = 0$ km) fino ad $h \simeq 40$ km. Al contrario, la profondità atmosferica degli sciame si misura in g/cm² a partire dalla cima dell'atmosfera ($X = 0$) ed è definita come

$$X = \int \rho(h') dh' \quad (1)$$

dove $\rho(h')$ è la densità dell'aria ed h' è il cammi-

³Utilizziamo nel seguito gli acronimi inglesi: SD (Surface Detector) ed FD (Fluorescence Detector)

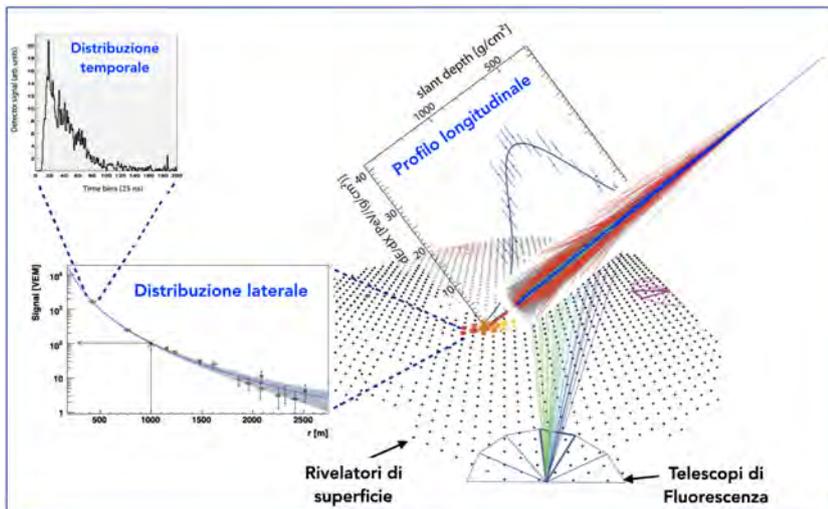


Figura 6: La misura di uno sciame esteso all'Osservatorio Pierre Auger. Con gli SD si determinano direzioni d'arrivo e punto d'impatto a terra dello sciame, e la distribuzione laterale, cioè la densità di particelle misurata in funzione della distanza dal core. In ogni stazione, si misura anche la distribuzione dei tempi d'arrivo delle particelle. Con gli FD si misura la distribuzione longitudinale, cioè l'energia depositata dallo sciame nell'attraversare l'atmosfera terrestre. Con le antenne radio si può misurare il segnale elettromagnetico nella regione di frequenza delle decine di MHz.

no percorso dall'EAS dal punto di ingresso in atmosfera lungo l'asse dello sciame (che coincide con la direzione d'arrivo del primario, ad un angolo θ rispetto coincide con la direzione d'arrivo del primario, ad un angolo θ rispetto alla verticale della superficie terrestre). Con questa definizione, risulta semplice descrivere lo sciame: in queste unità di misura le lunghezze di interazione ed attenuazione delle particelle risultano indipendenti dal mezzo in cui si propagano. Il punto d'impatto a terra dell'asse dello sciame viene comunemente chiamato il *core* dello sciame.

Poiché l'integrale della luce di fluorescenza è proporzionale all'energia del primario, usando l'FD, possiamo ottenere una misura quasi calorimetrica dell'energia elettromagnetica rilasciata in atmosfera e da questa, con una piccola correzione per tener conto dell'energia rilasciata in atmosfera in neutrini e muoni (detta energia invisibile E_{inv}), l'energia totale del primario:

$$E_{em} = \int_0^{\infty} dX \frac{dE}{dX}$$

$$E_{tot} = E_{em} + E_{inv} .$$

Utilizzando gli eventi ibridi, cioè gli sciami registrati simultaneamente dall'FD e dall'SD, possiamo correlare questa misura ad un parametro misurato con l'SD, la densità di particelle misurata a 1000 m dal *core*. In questo modo, la scala di energia dell'esperimento è ben definita per tutti gli eventi e quasi completamente indipendente dai modelli di interazione adronica, che a queste energie sono quanto mai incerti.

Le tre osservabili tipiche da ricostruire per

ogni evento registrato sono: l'energia, la direzione d'arrivo del primario, e la profondità di massimo sviluppo dello sciame in atmosfera. Nel caso degli eventi inclinati, inoltre, è possibile ricavare il numero di muoni e la profondità di massimo sviluppo della componente muonica, N_{μ} e X_{max}^{μ} . Nell'attraversare uno spessore atmosferico molto maggiore rispetto a sciami verticali, in questi eventi la componente elettromagnetica viene completamente assorbita, ed a terra arrivano soltanto i muoni.

Lo stato della ricerca

I risultati scientifici ottenuti dalla Collaborazione Pierre Auger coprono campi di ricerca diversi e complementari. La ricerca dell'origine degli UHECR si basa sulla misura dello spettro energetico e della composizione di massa dei primari, su studi di altri messaggeri neutri, i fotoni ed i neutrini, e su estese ricerche di anisotropia su scala angolare sia grande che intermedia. Oltre agli studi di tipo astrofisico, con i dati raccolti è possibile esplorare le caratteristiche delle interazioni adroniche ad energie inarrivabili agli acceleratori, e valutare l'esistenza di effetti di fisica non standard, come possibili violazioni dell'invarianza di Lorentz, o segnali di materia oscura super-pesante.

Anisotropie nel flusso degli UHECR

Il metodo più diretto per cercare di individuare le sorgenti degli UHECR è quello di studiare

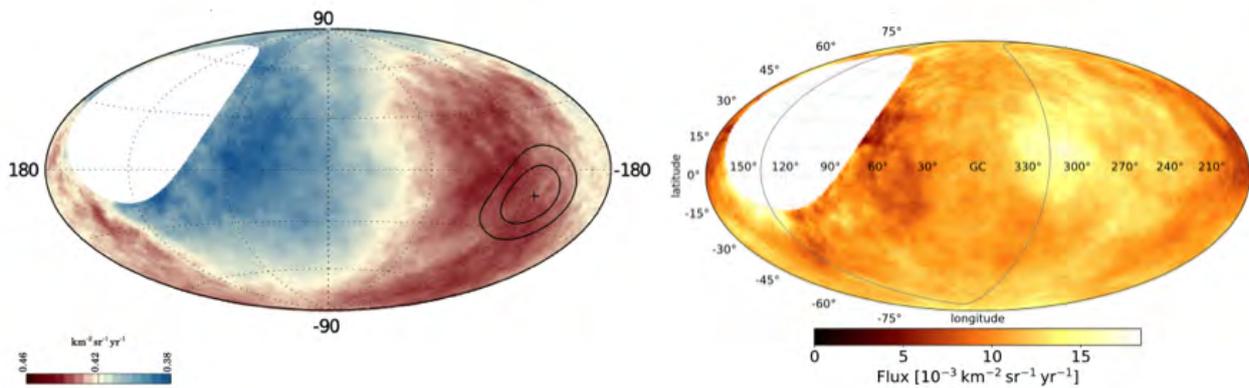


Figura 7: Il flusso degli UHECR in funzione delle coordinate Galattiche. A sinistra, per $E > 8 \text{ EeV}$; il dipolo è mostrato con una croce. A destra: per $E > 40 \text{ EeV}$; si nota l'eccesso in correlazione con NGC4945 ($305^\circ, 13^\circ$). In entrambi i casi, l'area bianca indica la parte di cielo non visibile dalla latitudine di Auger.

La misura dell'anisotropia

L'esposizione direzionale (area di raccolta degli eventi integrata nel tempo) in ascensione retta^a α è quasi uniforme per qualsiasi osservatorio a terra che operi con continuità nel tempo, grazie alla rotazione terrestre. Dati $N(\alpha)$ eventi misurati, un'analisi armonica permette di determinare le componenti di Fourier dell'armonica k : $a_k = (2/N) \sum_{i=1}^N \cos(k\alpha_i)$ e $b_k = (2/N) \sum_{i=1}^N \sin(k\alpha_i)$, e da esse l'ampiezza e la fase della modulazione della frequenza di eventi:

$$r_k = \sqrt{a_k^2 + b_k^2}; \quad \phi_k = \arctan(b_k/a_k)$$

Il caso più semplice di deviazione delle direzioni d'arrivo degli UHECR rispetto ad una distribuzione isotropa è quello di un puro dipolo ($k = 1$).

La probabilità di avere un valore di ampiezza maggiore di quello osservato solo per caso da un insieme isotropo di dati è pari a $P(> r_k) = \exp(-Nr_k^2/4)$. La significatività statistica del risultato si misura in numero di deviazioni standard σ dal valor medio atteso della distribuzione Gaussiana delle probabilità.

^aAscensione retta e declinazione (α e δ) sono le coordinate del sistema di riferimento equatoriale.

le loro direzioni d'arrivo e identificare possibili deviazioni da una loro distribuzione isotropa, eventualmente correlandole con qualche tipo di oggetto cosmico. Identificare tali anisotropie non è un compito semplice a causa dell'esistenza dei campi magnetici galattici ed extragalattici che deviano le direzioni degli UHECR, tanto più quanto minore è la rigidità delle particelle $R = E/Z$ dove E è l'energia e Z il numero atomico.

Le anisotropie possono verificarsi su una vasta gamma di scale angolari. Nella regione di energia più bassa, si esplorano scale di decine di gradi, poiché le deviazioni delle traiettorie degli UHECR sono troppo grandi per cercare sorgenti puntiformi. Ci aspettiamo ad esempio una deviazione dalla distribuzione isotropa a causa della diffusione e della fuga di particelle dalla Galassia, quando la loro energia diventa troppo alta perché il campo magnetico galattico possa mantenerle confinate all'interno. Alle energie più alte, d'altra parte, possiamo aspettarci rigidità più elevate e di conseguenza una maggiore possibilità di cercare sorgenti puntiformi o almeno anisotropie di scala intermedia.

L'enorme esposizione raggiunta all'Osservatorio Pierre Auger (circa $110,000 \text{ km}^2 \text{ sr yr}$) ha permesso di ottenere una statistica sufficiente a misurare un'anisotropia di dipolo di ampiezza pari al 5.9%, con una significatività statistica maggiore di 6σ [4]. La distribuzione del flusso degli UHECR per $E \geq 8 \text{ EeV}$ è mostrata nella Fig.7 in coordinate Galattiche. Il dipolo punta in direzione opposta al centro della nostra Galassia, dandoci una chiara indicazione del fatto

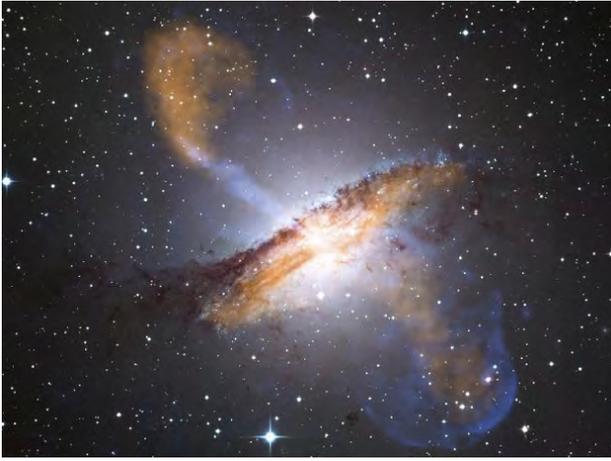


Figura 8: Centaurus A è la Galassia attiva più vicina alla Terra (~ 4.5 Mpc).



Figura 9: La Galassia starburst NGC253 (Galassia dello Scultore), a 2.7 Mpc dalla Terra.

che questi UHECR sono di origine extragalattica, un fatto previsto dalla teoria, ma solo grazie ad Auger ora confermato sperimentalmente⁴.

Ad energie molto superiori, è possibile provare a correlare le direzioni d'arrivo degli UHECR misurati a terra con le posizioni di eventuali possibili sorgenti (per esempio utilizzando i cataloghi di galassie forniti da esperimenti come Fermi-LAT, 2MRS, ecc.). Il flusso di UHECR misurato ad energie superiori a 40 EeV, utilizzando una griglia di 25° , viene mostrato nella Fig.7. L'area più interessante corrisponde alla regione del Centauro, che contiene Galassie importanti come

⁴È prassi consolidata nella fisica attribuire il significato di osservazione alle misure con significatività statistica $> 5\sigma$. Si parla invece di evidenza se la significatività è superiore a 3σ .

la radio-galassia Centaurus A, e galassie stellari come M83 e NGC4945; un ulteriore eccesso si trova vicino al Polo Sud Galattico, dove si trovano NGC1068 e NGC253.

La migliore significatività statistica, corrispondente a correlazioni con le *starburst galaxies* (SBG), è a livello di 4σ ; la correlazione con le posizioni indicate è interessante, ma ancora non sufficiente per concludere di aver identificato le sorgenti degli UHECR [5].

Uno studio completo dell'anisotropia richiede di poter osservare l'intero cielo. Questo è reso possibile dal lavoro congiunto della collaborazione Auger con Telescope Array, un esperimento volto anch'esso allo studio degli UHECR situato nell'emisfero Nord, nello Utah, di area pari a circa $1/4$ rispetto a quella di Auger [6]. La collaborazione tra i due gruppi ha consentito di confermare sia la misura del dipolo, che nel caso di una ricerca full-sky non dipende più da ipotesi di partenza sulla distribuzione angolare degli eventi, sia la correlazione con i cataloghi di diversi oggetti cosmici.

Spettro energetico e composizione

Lo spettro di energia degli UHECR è stato misurato in Auger utilizzando più di 200 mila eventi, in una regione di energia estremamente ampia, da $\sim 6 \times 10^{15}$ eV fino a $\sim 2 \times 10^{20}$ eV [7, 8]. Se si rappresenta il flusso moltiplicandolo per la terza potenza dell'energia in funzione dell'energia stessa, come in Fig.10, è possibile osservare alcuni cambi di pendenza, che sono estremamente informativi sui processi di produzione di queste astroparticelle. Le energie a cui essi sono osservati vengono generalmente indicati con nomi relativi all'anatomia della gamba: un secondo ginocchio (ne esiste un primo, attorno a $3 \cdot 10^{15}$ eV), energia a cui si presume completato il contributo dei raggi cosmici galattici; una caviglia (*ankle*, in inglese), nella regione di energia a cui si ha il passaggio da particelle di origine galattica ad extragalattica; il collo del piede (*instep*) è infine l'energia a cui comincia la soppressione del flusso.

Le caratteristiche degli sciami dipendono non solo dall'energia, ma anche dalla natura della particella primaria che li ha generati. A parità di energia iniziale la profondità atmosferica a cui si

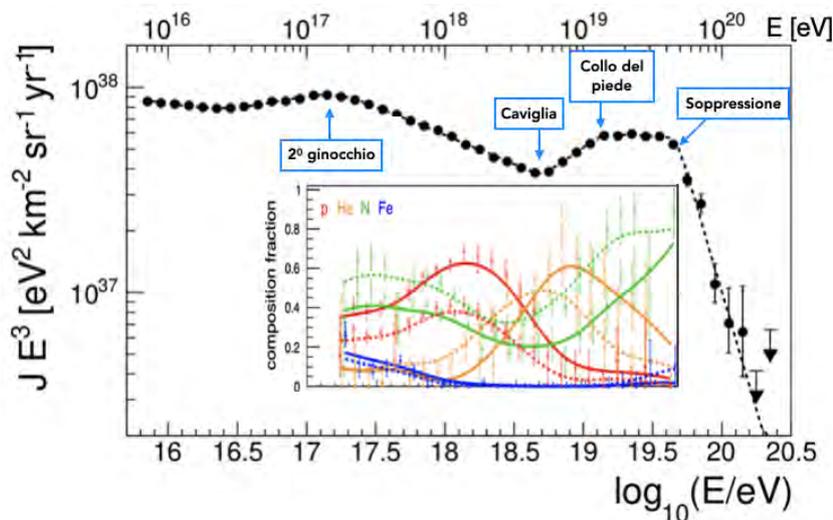


Figura 10: Il flusso degli UHECR in funzione dell'energia, moltiplicato per E^3 ad evidenziare i cambi di pendenza. Nell'inserto, si mostra la frazione di elementi di massa presenti nel fascio primario; le linee continue e tratteggiate si riferiscono a due diversi modelli di interazione adronica.

ha il massimo sviluppo dello sciame, X_{\max} , e il numero di muoni prodotti N_{μ} dipendono dalla massa della particella primaria: gli sciame originati da un protone si sviluppano più in basso nell'atmosfera, e contengono un numero minore di muoni rispetto ad uno sciame generato da un nucleo più pesante. In uno sciame generato invece da un fotone, la lunghezza di radiazione è maggiore di un fattore 2 rispetto a quella delle interazioni foto-nucleari; di conseguenza X_{\max} è molto più profondo (più vicino alla terra) rispetto a quello di uno sciame adronico alla stessa energia ed il contenuto di muoni è molto più basso.

Informazioni sulla massa dei primari possono essere ricavate dal valor medio e dalla varianza dei profili longitudinali, ($\langle X_{\max} \rangle$ e $\sigma^2(X_{\max})$; (si veda il riquadro "Misura della composizione"). Con sufficiente statistica, è possibile considerare le intere distribuzioni di X_{\max} e modellarle come sovrapposizione delle distribuzioni attese da gruppi di massa rappresentativi (di solito si usano protoni e nuclei di ^4He , ^{14}N , ^{28}Si , ^{56}Fe), con parametri che dipendono dal modello usato per descrivere le interazioni adroniche in atmosfera. Confrontando la convoluzione di queste distribuzioni con la distribuzione misurata, è quindi possibile determinare le frazioni di massa ad ogni energia che meglio descrivono i dati sperimentali; il risultato è mostrato nell'inserto di Fig.10.

Da un punto di vista astrofisico, è importante osservare come la composizione degli UHECR non sia costante, ma si evolva con l'energia; nella regione energetica fino a un po' prima del-

la caviglia dominano i nuclei più leggeri, poi appaiono nuclei sempre più pesanti [9].

Questo risultato, confermato in diverse analisi indipendenti, esclude che la regione di transizione ad un'origine extragalattica degli UHECR (attorno alla caviglia sia composta solo da primari leggeri ($p+\text{He}$), come previsto da alcuni modelli. Nella regione di energia più estrema la composizione è mista; poiché i nuclei subiscono maggiormente l'effetto delle deviazioni dovute ai campi magnetici extragalattici, diventa più complicato ricercarne le sorgenti seguendo le loro direzioni d'arrivo, a meno che la frazione dei protoni presenti non sia sufficientemente grande (non meno del 10%).

Sulla base dei risultati sperimentali su spettro e composizione, è possibile ricavare informazioni sulle caratteristiche delle possibili sorgenti. A partire da scenari astrofisici di cui si parametrizzano tipo di sorgenti e loro evoluzione cosmologica, masse delle particelle accelerate nella sorgente, forma dello spettro alla produzione, si possono infatti propagare le particelle accelerate fino a Terra, tenendo conto di tutte le possibili interazioni e perdite di energia, e ricavare spettro e composizione attesi.

Il confronto tra quantità attese e osservate permette, attraverso una procedura di *fit*, di determinare o limitare i parametri delle sorgenti. L'analisi combinata di spettro energetico e composizione di Auger suggerisce che la soppressione del flusso alle energie più estreme sia dovuta al raggiungimento della massima energia disponibile per accelerare particelle nelle sorgenti, piuttosto che ad effetti di propagazione [10]. Estendendo

La misura della composizione

Le osservabili più usate per ottenere informazioni sulla composizione sono la profondità media del massimo sviluppo degli sciami in atmosfera, $\langle X_{\max} \rangle$ e le sue fluttuazioni $\sigma^2(X_{\max})$, ed il numero di muoni N_{μ} . Per una composizione di nuclei mista all'ingresso in atmosfera, avremo che la profondità media può essere espressa come

$$\langle X_{\max}^A \rangle = \langle X_{\max}^p \rangle + f_E \langle \ln A \rangle$$

dove X_{\max}^p rappresenta la profondità media per il protone. Il parametro f_E dipende dalla frequenza di variazione di X_{\max} per decade di energia (detto *elongation rate*), e dalle caratteristiche delle interazioni adroniche. Le fluttuazioni di $\langle X_{\max}^A \rangle$ sono somma di fluttuazioni intrinseche, σ_{sh} , cioè le fluttuazioni nello sviluppo di sciami diversi generati dallo stesso tipo di nucleo, e quelle che derivano dalla dispersione dei valori di A nel fascio primario, $\sigma_{\ln A}$.

Il numero di muoni contenuto in uno sciami generato da un protone oppure da un nucleo di massa A è dato da

$$\begin{aligned} N_{\mu}^p &\sim (E/\epsilon_d^{\pi})^{0.9} \\ N_{\mu}^A &\sim N_{\mu}^p A^{0.1} \end{aligned}$$

con ϵ_d^{π} pari all'energia di decadimento dei pioni carichi.

questo lavoro fino alle energie inferiori, sotto la caviglia, i dati di Auger risultano ben descritti da una componente galattica di massa intermedia (nuclei del gruppo CNO), unita a due componenti extragalattiche, una di protoni ed una di massa mista (composta cioè da p , He e altri nuclei fino al ferro) [11].

La presenza di raggi cosmici galattici a queste energie, ben superiori a quelle a cui essi possono essere accelerati da meccanismi standard (accelerazione in onde d'urto in residui di Supernova), suggerisce che esista una seconda componente galattica, dominata da elementi del gruppo CNO, generata ad esempio in esplosioni di supernovae nei venti di stelle di tipo

Wolf-Rayet⁵.

Multi-messaggeri dallo spazio

All'Osservatorio Pierre Auger è possibile registrare anche particelle neutre, fotoni e neutrini [12].

I neutrini possono viaggiare in percorsi rettilinei dai confini dell'Universo e sono in grado di attraversare grandi profondità di materia senza assorbimento. I neutrini astrofisici sono eccellenti messaggeri da sorgenti extragalattiche, generati nel decadimento dei pioni creati nelle interazioni dei raggi cosmici con la materia e/o con la radiazione nelle loro potenziali sorgenti. I neutrini cosmologici sono prodotti nell'interazione degli UHECR con i fotoni di fondo durante la propagazione.

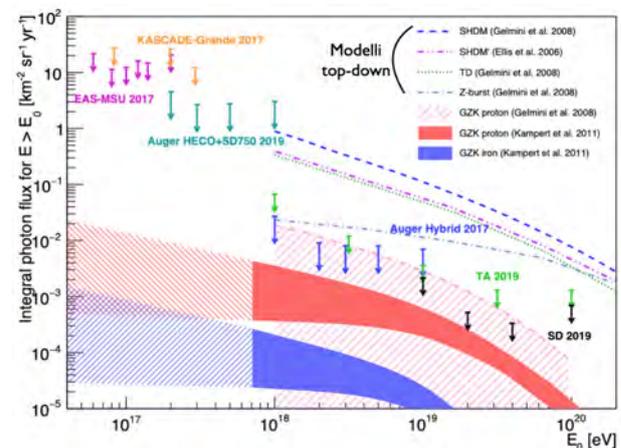


Figura 11: Limiti superiori al flusso di fotoni cosmogenici. Le linee tratteggiate mostrano le previsioni dei modelli top-down.

Identificare gli sciami generati da neutrini è relativamente facile, sulla base delle differenze dei loro profili longitudinali e laterali rispetto a quelli prodotti da sciami adronici provenienti da direzioni molto inclinate ($\theta > 60^\circ$). In quest'ultimo caso, infatti, l'atmosfera è abbastanza spessa (più di 3 atmosfere verticali!) da assorbire quasi completamente la componente elettromagnetica degli EAS quasi orizzontali, e la distribuzione dei secondari a terra risulta dunque dominata dai muoni. Al contrario, i neutrini, a causa della loro

⁵Le stelle di Wolf-Rayet sono stelle massicce, con masse iniziali superiori a 20 masse solari, in uno stadio avanzato della loro evoluzione. In base alla loro composizione, i loro spettri di emissione sono dominati o da He ed N o da He, C, O.

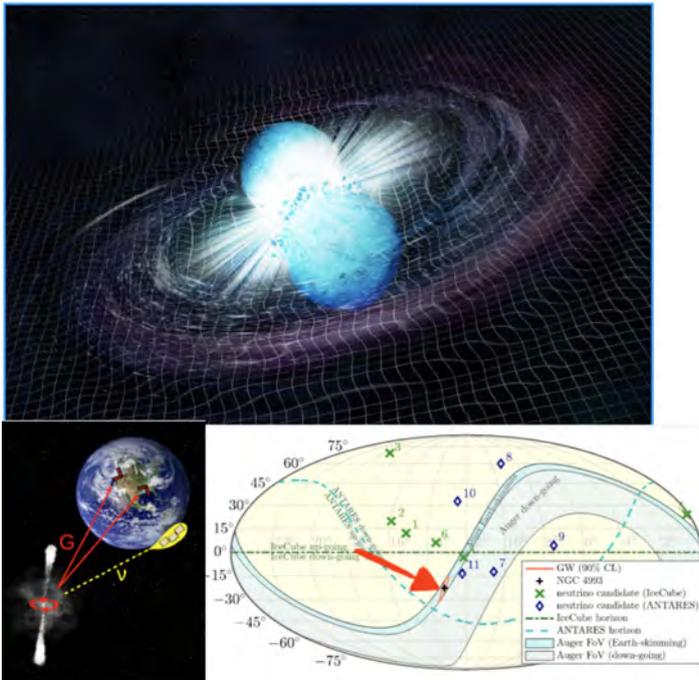


Figura 12: Schema della fusione di due stelle di neutroni che creano un buco nero, con emissione di onde gravitazionali. Oltre all'emissione di onde elettromagnetiche, questi eventi possono produrre neutrini, osservabili a Terra in apparati come IceCube (Polo Sud), ANTARES (Mar Mediterraneo), Auger (Argentina). L'evento GW170817 è stato localizzato in una regione di cielo (il cerchio rosso nel plot) corrispondente alla controparte ottica NGC4993 (croce nera). La sorgente si trovava appena sotto l'orizzonte, nel campo di vista ottimale per i neutrini nel caso di Auger, ma non degli altri Osservatori.

sezione d'urto estremamente piccola, possono interagire molto in basso in atmosfera, generando sciami ricchi di particelle elettromagnetiche vicino al rivelatore.

I fotoni di altissima energia possono derivare da una serie di processi:

- (a) fotoni cosmogenici possono essere prodotti durante la propagazione degli UHECR, nel processo GZK;
- (b) possono essere prodotti da interazioni adroniche dei raggi cosmici all'interno delle sorgenti o nel loro ambiente locale;
- (c) nei modelli *top-down* di produzione degli UHECR ci si attende un grande flusso di fotoni di altissima energia, prodotti nel decadimento di particelle super massicce come difetti topologici o particelle di materia oscura super-pesante.

A differenza dei neutrini, questi fotoni interagiscono con la radiazione di fondo extragalattica (EBL); la cascata elettromagnetica indotta si degrada fino a concludersi ad energie di $10^9 - 10^{12}$ eV. Sono quindi un mezzo per esplorare l'universo locale.

Anche gli sciami generati da fotoni hanno un basso contenuto in muoni e si sviluppano più in profondità dei raggi cosmici nell'atmosfera, e dunque sono abbastanza facilmente separabili dal fondo di sciami adronici.

Non è invece possibile separare gli sciami indotti dai neutroni da quelli generati da protoni. Tuttavia, poichè i neutroni sono direzionali, è in

linea di principio possibile identificare le loro sorgenti vicine: i neutroni possono infatti coprire distanze dell'ordine dei $9.2 \text{ kpc} \times E/\text{EeV}$. La distanza dalla Terra al Centro Galattico è di circa 8.3 kpc, e il raggio della Galassia è di circa 15 kpc. Si potrebbero quindi rivelare eventuali sorgenti nel disco galattico, compreso il centro galattico, sfruttando i neutroni di energia superiore a 1 EeV.

L'osservazione anche solo di un fotone o di un neutrino di altissima energia sarebbe un risultato davvero spettacolare, ma nessun evento di questo tipo è stato finora registrato (Fig.11).

Il fatto non è però negativo come sembra: la non-rilevazione di questi eventi è di per sé un dato sperimentale importante. Per esempio, i limiti superiori posti al flusso dei neutrini e dei fotoni cosmogenici hanno consentito di eliminare la maggior parte dei modelli di origine *top-down* degli UHECR e, nel caso dei fotoni, ha posto limiti piuttosto stretti alla massa e alla vita media di eventuali particelle di materia oscura super-pesante.

Un altro esempio è fornito dalla ricerca di neutrini in coincidenza con gli eventi di onde gravitazionali (Fig.12). La rivelazione di questi neutrini alle altissime energie, osservabili soltanto in Auger, può aiutare nella localizzazione delle sorgenti delle onde gravitazionali e consentire uno studio complementare dei loro progenitori astrofisici.

L'Osservatorio riceve e reagisce agli allarmi inviati dalla rete di rivelatori di onde gravitazionali, e fa parte dell'Astrophysical Multi-messenger Observatory Network, che riceve e distribuisce allarmi all'arrivo di eventi transienti particolarmente significativi tra un numero sempre crescente di osservatori nel mondo.

Fisica delle particelle con le astroparticelle

La capacità di dedurre la composizione nucleare dei raggi cosmici su base statistica si basa in larga misura sulla comprensione teorica delle interazioni adroniche che si verificano all'interno della cascata. Tuttavia, la produzione adronica di multiparticelle deve essere estrapolata ad energie che superano di gran lunga quelle raggiungibili negli acceleratori terrestri e in regioni dello spazio delle fasi non coperte dagli esperimenti al *collider* LHC del CERN. Le interazioni degli UHECR in atmosfera sondano infatti regioni di energia dell'ordine di $50 \lesssim \sqrt{s}/\text{TeV} \lesssim 450$, da confrontare con i 14 TeV raggiungibili a LHC negli urti protone-protone.

Misurando gli EAS è possibile verificare la validità delle estrapolazioni e dunque delle previsioni del Modello Standard delle particelle elementari. Per esempio, utilizzando le distribuzioni longitudinali degli sciami misurati con FD, in Auger è stata studiata la sezione d'urto protone-Aria fino ad energie dell'ordine dei 50 TeV, ricavandone, con l'uso di un modello fenomenologico, la sezione d'urto protone-protone (Fig.13), ben al di sopra della scala di energia di LHC [13].

Lo studio degli UHECR in Auger ha mostrato che tutti i modelli in uso descrivono bene la componente elettromagnetica degli sciami, ma ne sottostimano in modo significativo il numero di muoni. Questo deficit di muoni nei modelli inizia a vedersi attorno a 10^{17} eV, ed aumenta con il crescere dell'energia con una significatività di più di 8σ ; al contrario, le fluttuazioni nel numero di muoni sono ben riprodotte dai modelli. Le modifiche da apportare ai modelli per eliminare l'attuale deficit di muoni devono quindi essere tali da aumentare gradualmente

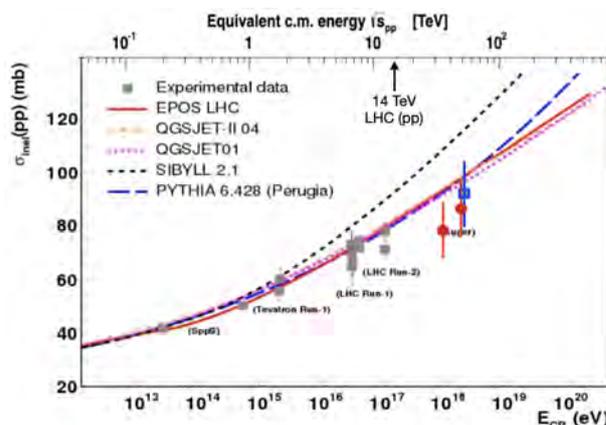


Figura 13: Sezione d'urto inelastica p - p ricavata dai dati di Auger (punti rossi e blu) confrontata con i risultati di alcuni modelli e con i dati dagli acceleratori.

la frazione adronica negli sciami senza alterare le fluttuazioni relative previste per i muoni [14].

Diversi risultati alle altissime energie sono stati ottenuti nella ricerca di effetti non previsti dal Modello Standard delle particelle.

Abbiamo visto come sia possibile porre limiti molto stringenti alla produzione di UHECR da parte di materia oscura superpesante grazie alle ricerche sui fotoni primari. La fisica delle astroparticelle è stata anche proposta come un ambiente di prova appropriato per la possibile violazione dell'invarianza di Lorentz (LIV), data la grande energia delle particelle, le enormi distanze di propagazione, l'accumulo di piccoli effetti di interazione e, recentemente, la precisione delle misure. La LIV viene descritta come modifica delle relazioni di dispersione⁶, che porterebbe ad un cambiamento nella cinematica delle interazioni causando modifiche nella propagazione extragalattica degli UHECR. I risultati di Auger in questo ambito permettono di limitare il livello di LIV possibile fino a 4 ordini di grandezza più di quanto fatto in precedenza [15].

⁶L'invarianza di Lorentz è uno dei pilastri della fisica moderna ed è stata verificata con diversi approcci sperimentali. In particolare, la massa a riposo m della particella è un'invariante di Lorentz legato all'energia E e all'impulso $p = |\mathbf{p}|$ misurati in qualsiasi sistema di riferimento dalla relazione $E^2 - p^2 c^2 = m^2 c^4$, dove c è la velocità della luce nel vuoto. Una sua violazione può essere espressa come $E^2 = p^2 c^2 + m^2 c^4 + \delta_n(E^n)$, dove δ_n esprime una correzione perturbativa di ordine n .

Cosmo-geofisica

Come visto in Fig.1, è possibile utilizzare le misure degli UHECR anche per studi di cosmo-geofisica. Sfruttando alcune delle caratteristiche uniche dei rivelatori di Auger, sono stati analizzati gli ELVES, eventi luminosi transitori che si verificano alla base della ionosfera, quando un forte impulso elettromagnetico viene emesso da un fulmine. La luce di fluorescenza generata in

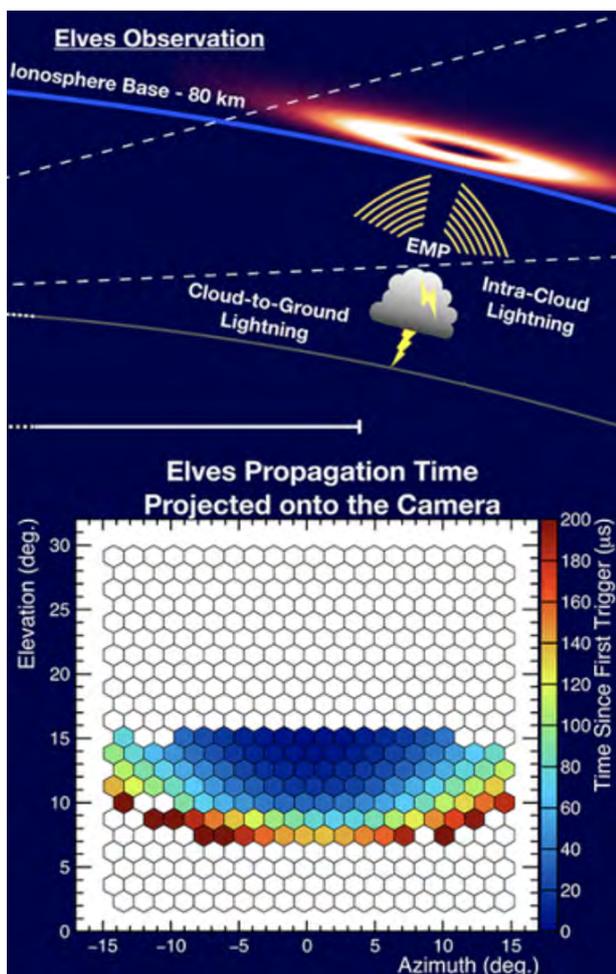


Figura 14: Schema di produzione di ELVES e immagine rilevata sulla camera di un telescopio FD.

atmosfera si propaga sotto forma di un anello in rapida espansione e può essere raccolta dai telescopi di Auger su aree fino a $\sim 10^6$ km², con una precisione senza precedenti grazie alla risoluzione temporale degli FD (100 nsec). Con questi studi si forniscono informazioni importanti sulle dinamiche degli acceleratori di plasma sul nostro pianeta, come quelli che si nascondono dietro il lampo di un fulmine [16]. L'Osservatorio è anche sensibile ai TGF (*terrestrial gamma-ray flashes*)

e alle diminuzioni nel flusso di raggi cosmici per effetto di attività solare (effetto Forbush).



Figura 15: Uno dei rivelatori di superficie di AugerPrime. Sono visibili sia il rivelatore a scintillazione (4 m²) che l'antenna radio installati sopra la stazione Čerenkov.

Il futuro

Si può certamente dire che viviamo un'epoca d'oro della fisica astroparticellare: grazie allo sviluppo tecnologico ed all'enorme numero di informazioni ottenute grazie agli osservatori attualmente in misura è nata l'astronomia multimessaggera, e si sono rafforzati i legami con la fisica delle particelle, l'astrofisica, la cosmologia.

L'enorme esposizione agli UHECR accumulata in particolare dall'Osservatorio Pierre Auger ha portato a grandi progressi nella comprensione delle loro proprietà, ma le ancora numerose incognite sulla natura e la distribuzione delle sorgenti, la composizione primaria e le interazioni adroniche sottostanti impediscono l'emergere di un quadro univocamente coerente.

I risultati attuali propongono nuove sfide:

- capire l'origine della soppressione del flusso, differenziando tra un *cut-off* dovuto agli effetti di propagazione oppure al raggiungimento dell'energia massima raggiunta nelle sorgenti. Questo ci permetterà di caratterizzare le eventuali sorgenti.
- valutare la possibile esistenza di una frazione di protoni alle più alte energie. La risposta a questa

domanda fondamentale porrà le basi per futuri esperimenti e valuterà la fattibilità dell'astronomia delle particelle cariche.

– aggiungere ulteriori informazioni sulle interazioni adroniche alle altissime energie, dove non possono arrivare le misure dei collider ed esplorare meccanismi non standard, previsti per esempio dalle teorie di gravità quantistica o nella teoria delle stringhe.

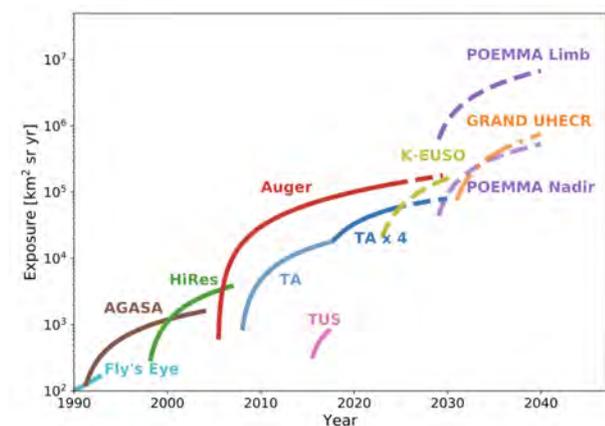


Figura 16: Confronto tra le esposizioni degli esperimenti passati (Fly's Eye, AGASA, HiRes), attuali (Auger, Telescope Array) e futuri (linee tratteggiate).

Per raggiungere questi obiettivi, è essenziale ottenere ulteriori misure di osservabili sensibili alla composizione su un campione statistico elevato di eventi ad altissima energia. Questo ha condotto la Collaborazione Pierre Auger ad effettuare un *upgrade* dell'Osservatorio, chiamato AugerPrime [17], in cui ad ogni rivelatore di superficie viene aggiunto un rivelatore a scintillazione ed un'antenna radio, vengono introdotti rivelatori posti a ~ 2.5 m sottoterra per la misura dei muoni e si ottimizzano risoluzioni temporali e range dinamico dei segnali. AugerPrime sarà dunque un Osservatorio di precisione, con lo scopo di misurare la composizione degli UHECR evento per evento, con una confidenza superiore al 50%.

È in corso attualmente anche l'*upgrade* di Telescope Array (TAx4) [18], con l'idea di avere anche nell'emisfero Nord un'esposizione agli UHECR paragonabile a quella di Auger.

Nel futuro più lontano, nuovi progetti potrebbero permettere la misura degli UHECR all'energia di soppressione ed oltre. L'osservatorio spaziale Probe of MultiMessenger

Astrophysics (POEMMA) [19] e l'osservatorio a terra multi-sito Giant Radio Array for Neutrino Detection (GRAND) [20] sono due strumenti che misureranno sia i neutrini UHE che i raggi cosmici. Il Global Cosmic Ray Observatory (GCOS) [21], in fase di studio, prevede due osservatori a terra, uno dei quali collocato con GRAND, per arrivare a 40 mila km² di area.

L'evoluzione delle esposizioni per i diversi esperimenti è mostrata in Fig.16; AugerPrime rimarrà l'esperimento *leader* nel campo della fisica degli UHECR per questo decennio. Si potranno raggiungere esposizioni 100 volte superiori con i rivelatori previsti dopo il 2030.

Lo studio dei raggi cosmici di altissima energia rimane uno dei più importanti punti chiave dell'affascinante campo della fisica astroparticellare per la comprensione e l'origine delle particelle più energetiche dell'Universo.

Un invito finale

Nel 2021, l'Osservatorio Pierre Auger ha reso disponibili pubblicamente il 10% dei suoi dati. Collegandosi al sito <https://opendata.auger.org/> è possibile per chiunque lo desideri scaricare i dati, visualizzare gli eventi, riprodurre alcune delle analisi che abbiamo descritto.

Appendice: lista delle sigle

- AERA - Auger Engineering Radio Array
- AGN - Active Galactic Nuclei
- CMB - Cosmic Microwave Background
- EAS - Extensive Air Shower
- EBL - Extragalactic Background Light
- ELVES - Emissions of Light and Very low frequency perturbations due to Electromagnetic pulse Sources
- FD - Fluorescence Detector
- GRB - Gamma Ray Burst
- GZK - Greisen, Zatsepin e Kuzmin
- LIV - Lorentz Invariance Violation
- SBG - Starburst Galaxies
- SD - Surface Detector
- TDE - Tidal Disruption Event
- TGF - Terrestrial Gamma-ray Flashes
- UHE - Ultra High Energy
- UHECR - Ultra High Energy Cosmic Rays



- [1] R. Alves Batista et al., *Open questions in cosmic-ray research at ultrahigh energies*, *Front. Astron. Space Sci.*, 6 (2019) 23.
- [2] A. Castellina, F. Donato, *Astrophysics of Galactic Charged Cosmic Rays*, *Planets, Stars and Stellar Systems: Galactic Structure and Stellar Populations*, 5 (2011) 725.
- [3] Pierre Auger Coll., *The Pierre Auger Cosmic Ray Observatory*, *Nucl. Instrum. Meth. A* 798 (2015) 172.
- [4] Pierre Auger Coll., *Observation of a large-scale anisotropy in the arrival directions of cosmic rays above 8×10^{18} eV*, *Science*, 357 (2017) 1266.
- [5] Pierre Auger Coll., *An Indication of Anisotropy in Arrival Directions of Ultra-high-energy Cosmic Rays through Comparison to the Flux Pattern of Extragalactic Gamma-Ray Sources*, *Astrop. J.* 853 (2018) L29.
- [6] S. Ogio [Telescope Array Coll.], *Telescope Array experiment*, *EPJ Web of Conf.* 208 (2019) 08002.
- [7] Pierre Auger Coll., *Measurement of the cosmic-ray energy spectrum above 2.5×10^{18} eV using the Pierre Auger Observatory*, *Phys. Rev. D* 102 (2020) 062005.
- [8] V. Novotny [Pierre Auger Coll.], *Energy spectrum of cosmic rays measured using the Pierre Auger Observatory*, *Proc.37th Int.Cosmic Ray Conf. PoS(ICRC2021)*, 395 (2021) 324.
- [9] Pierre Auger Coll., *Inferences on mass composition and tests of hadronic interactions from 0.3 to 100 EeV using the water-Cherenkov detectors of the Pierre Auger Observatory*, *Phys. Rev. D* 96 (2017) 122003.
- [10] Pierre Auger Coll., *Combined fit of spectrum and composition data as measured by the Pierre Auger Observatory* *Journ. Cosmology and Astrop. Phys. A* 798 (2015) 172.
- [11] E. Guido [Pierre Auger Coll.], *Combined fit of the energy spectrum and mass composition across the ankle with the data measured at the Pierre Auger Observatory*, *Proc.37th Int.Cosmic Ray Conf. PoS(ICRC2021)*, 395 (2021) 311.
- [12] Pierre Auger Coll., *Multi-Messenger Physics With the Pierre Auger Observatory*, *Front. Astron. Space Sci.*, 6 (2019) 24.
- [13] Pierre Auger Coll., *Measurement of the Proton-Air Cross Section at $\sqrt{s} = 57$ TeV with the Pierre Auger Observatory*, *Phys. Rev. Lett.*, 109 (2012) 062002.
- [14] Pierre Auger Coll., *Measurement of the fluctuations in the number of muons in extensive air showers with the Pierre Auger Observatory*, *Phys. Rev. Lett.*, 126 (2021) 152002.
- [15] Pierre Auger Coll., *Testing effects of Lorentz Invariance Violation in the propagation of astroparticles with the Pierre Auger Observatory*, *Journ. Cosmology and Astrop. Phys.*, 01 (2022) 023.
- [16] Pierre Auger Coll., *A 3-Year Sample of Almost 1,600 Eaves Recorded Above South America by the Pierre Auger Cosmic-Ray Observatory*, *Earth and Space Sci.*, 7 (2020) e00582.
- [17] A. Castellina [Pierre Auger Coll.], *AugerPrime: the Pierre Auger Observatory Upgrade*, *Proc. Int. Cosmic Ray Conf. on UHECRs*, 210 (2019) 06002.
- [18] Telescope Array Coll., *Surface detectors of the TAx4 experiment*, *Nucl. Instr. Meth. A* 1019 (2021) 165726.
- [19] POEMMA Coll., *POEMMA's target-of-opportunity sensitivity to cosmic neutrino transient sources*, *Journ. Cosmology and Astrop. Phys.*, 06 (2021) 66.
- [20] GRAND Coll., *The Giant Radio Array for Neutrino Detection (GRAND): Science and design*, *Sci. China Phys., Mech. and Astron.* 63 (2020) 219501.
- [21] J. Hörandel [GCOS Coll.], *GCOS, The Global Cosmic Ray Observatory*, *Proc.37th Int.Cosmic Ray Conf. PoS(ICRC2021)* 395 (2021) 027



Antonella Castellina: è co-spokesperson dell'Osservatorio Pierre Auger e Primo Ricercatore dell'Osservatorio Astrofisico (INAF) di Torino.

Sul tetto del mondo per studiare la radiazione cosmica

... eine Strahlung von sehr hoher Durchdringungskraft
von oben her in unsere Atmosphäre eindringt ...

— Viktor F. Hess, *Physik. Zeitschr.* XIII (1912) 1084

Paolo Bernardini

*Dipartimento di Matematica & Fisica "Ennio De Giorgi", Università del Salento
Istituto Nazionale di Fisica Nucleare, Sez. di Lecce, Lecce, Italia*

Vengono qui sinteticamente esposte le motivazioni che hanno spinto la collaborazione italo-cinese ARGO-YBJ ad installare un rivelatore per raggi cosmici in Tibet. Oltre agli obiettivi scientifici dell'esperimento, vengono anche presentate le difficoltà logistiche ed ambientali che i ricercatori hanno dovuto affrontare alla quota di 4300 metri.

Introduzione

I raggi cosmici furono scoperti nel 1912 da Viktor Hess e da Domenico Pacini a bordo, rispettivamente, di una mongolfiera e di un cacciatorepeditore. Ancora oggi chi progetta un esperimento sulla radiazione cosmica deve rinunciare all'idea di un comodo e ben attrezzato laboratorio nei pressi di una grande città e di un aeroporto internazionale. I rivelatori per la radiazione cosmica, nelle sue diverse componenti (raggi cosmici, raggi gamma e neutrini) sono collocati in montagna, a bordo di satelliti, in zone desertiche, al polo

Sud, in miniere abbandonate o nelle profondità marine. Collocazioni tanto esotiche hanno il loro fascino, ma evidentemente comportano difficoltà logistiche, superabili solo grazie ad un'attenta organizzazione e ad un grande entusiasmo.

L'esperimento

Questo articolo tratta dell'esperimento ARGO-YBJ [1], installato alla proibitiva quota di 4300 metri sul livello del mare. Progettato per lo studio dei raggi cosmici e per osservazioni astronomiche nella banda gamma, aveva un campo di vista molto ampio ed era in grado di rimanere operativo per mesi, senza interruzioni. Infatti, sebbene l'acronimo significhi *Astrophysical Radiation with Ground-based Observatory at YangBaJing*, il nome rimanda alla mitologia greca: Argo Panoptes ("che tutto vede") era un gigante dotato di cento occhi, capace di mantenerne aperti la metà durante il sonno e quindi era sempre vigile. YangBaJing è il nome del villaggio tibe-



Figura 1: Il capannone dell'esperimento ARGO-YBJ nel villaggio tibetano di YangBaJing. Sono visibili dei container usati per il trasporto di strumentazione dall'Italia. Sono presenti anche due abitazioni tipiche del villaggio.

tano nei pressi del quale era collocato l'apparato sperimentale. Quest'ultimo consisteva in un tappeto di RPC (*Resistive Plane Counters*) esteso all'incirca quanto un campo di calcio. Gli RPC sono rivelatori a gas, quelli usati in Tibet erano in grado di registrare il passaggio di una particella elettricamente carica con una risoluzione spaziale dell'ordine delle decine di centimetri, mentre quella temporale era del nanosecondo, cioè di un milionesimo di secondo. Gli RPC erano disposti in modo da costituire un rivelatore continuo, senza separazione tra i diversi elementi. Questa caratteristica (*full coverage*) e l'ottima risoluzione temporale hanno permesso di "fotografare" gli sciame di particelle cariche innescati dai raggi cosmici che penetrano nell'atmosfera (vedi Fig. 3).

Gli sciame hanno una loro evoluzione, inizialmente si manifesta una rapida crescita del numero di particelle, fino al raggiungimento di un massimo, seguito da una lenta riduzione della numerosità fino alla completa scomparsa dello sciame. Ovviamente tale processo è condiziona-

to dall'energia del raggio cosmico primario, proveniente dallo spazio esterno. Al crescere della sua energia lo sciame penetra sempre più profondamente nell'atmosfera, raggiungendo quote sempre più basse. In altre parole, la quota prescelta per l'installazione di un rivelatore definisce l'intervallo di energia che si vuole indagare. Realizzare un esperimento in alta montagna significa abbassare la soglia in energia dei raggi cosmici rivelati, avvicinandosi alle energie tipiche delle misure da satellite.

La collaborazione ARGO-YBJ aveva come obiettivo studiare i raggi cosmici nella banda energetica $10^{14} - 10^{16}$ eV e questo ha comportato la ricerca di un sito sufficientemente alto. Un'opzione possibile era Chacaltaya in Bolivia, a 5400 metri di quota. Ma a quelle altitudini le lunghe permanenze sono fortemente sconsigliate e così i ricercatori sarebbero stati costretti a scendere a quote inferiori per il riposo notturno, sottoponendosi ogni giorno a lunghi spostamenti. La scelta di installare l'esperimento in Tibet, nei pressi del villaggio di YangBaJing, presentava una se-

rie di vantaggi. La quota restava notevole (4300 m), ma non tale da impedire il pernottamento. I colleghi dell'Accademia Sinica e dell'Università di Lhasa erano disponibili a collaborare ed offrire supporto logistico. Inoltre il villaggio di YangBa-Jing si trova in un vasto bacino pianeggiante, in prossimità di una centrale geotermica e gode di condizioni meteo non proibitive. Infine, si prevedevano la costruzione della ferrovia Pechino-Lhasa e il rifacimento della strada che collega il villaggio al capoluogo e all'aeroporto¹.

Ad alta quota

Così alla fine degli anni '90 iniziarono i test a YangBaJing per verificare il funzionamento dei rivelatori e dell'elettronica a quelle quote [3]. In realtà i primi viaggi in Tibet servirono anche a comprendere a quali disagi e difficoltà andavano incontro tecnici e ricercatori. Il mal di montagna, dovuto essenzialmente alla ridotta ossigenazione del sangue, si può manifestare anche a 2500 metri di quota, ed è una certezza a 4300 metri. Furono quindi adottate delle contromisure. Oltre all'assunzione di diuretici e l'uso occasionale delle bombolette di ossigeno, si decise che prima si salire fino al laboratorio si rimanesse in albergo a Lhasa (solo 3600 metri di quota) per un paio di giorni in modo da abituarsi alle nuove condizioni.

Ricordo bene la prima trasferta in Tibet. Avevamo trascorso oltre 24 ore in volo o in qualche aeroporto: da Brindisi a Roma, poi a Pechino, a Chengdu e infine a Lhasa. Inizialmente non ebbi alcun malessere e pensai di poter andare in giro a fare il turista, ma bastò affrontare pochi scalini per capire che non ero nelle condizioni di fare alcunchè. Il mal di testa e l'affanno consentivano solo di starsene distesi, immobili. Anche una semplice lettura era troppo faticosa. Non fu molto diverso nei viaggi successivi anche se ormai sapevo bene come comportarmi e quali accorgimenti prendere per ridurre gli effetti dell'altitudine.

Dopo qualche giorno di acclimatamento a Lhasa, si ripartiva. Con uno sgangherato pullmino si percorrevano una novantina di chilometri, salendo di altri 700 metri. A YangBaJing l'aria era

ancora più rarefatta e quindi tornavano a manifestarsi mal di testa, affanno e disturbi del sonno. Altro problema era la secchezza dell'aria: l'Oceano Indiano è molto lontano e la catena himalayana è un ostacolo insormontabile per i monsoni, così l'umidità in Tibet è intorno al 10%. È facile immaginare cosa succede in queste condizioni alle labbra e alla pelle di chi abitualmente vive a pochi chilometri dal mare, anzi da due mari nel caso dei salentini.



Figura 2: L'autore dell'articolo in visita a un monastero tibetano. Foto del prof. Ivan De Mitri.

Ma nonostante i disagi dell'alta quota, visitare il Tibet è un'esperienza magnifica. Descrivere le montagne, le vallate, le mandrie di yak, i villaggi e i monasteri, le usanze dei tibetani richiederebbe una perizia che non ho. Per chi fosse interessato, segnalo il libro di una collega di Torino [4], che ha raccontato in maniera magistrale il Tibet e i suoi abitanti.

I risultati scientifici

Purtroppo gli impegni in laboratorio non lasciavano molto tempo per il turismo. Il montaggio dell'apparato, ultimato nel 2007, richiese un notevole impegno, mettendo a dura prova la resistenza di tecnici e ricercatori. Negli anni successivi, fino al 2013, si dovette provvedere alla manutenzione e alla gestione dell'apparato, con turni che tipicamente duravano tre settimane. Così si raccolse un'enorme mole di dati, oggetto di attente analisi.

Non è semplice sintetizzare i risultati scientifici dell'esperimento. Questi sono stati presentati in centinaia di conferenze internazionali e sono l'argomento di oltre 50 pubblicazioni su

¹Tali opere sono state poi effettivamente realizzate.

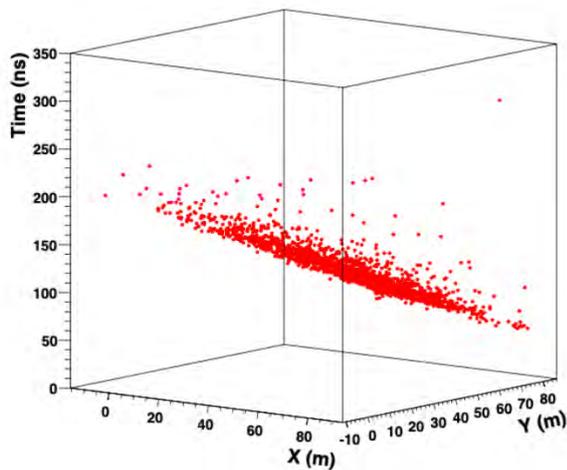


Figura 3: Immagine di uno sciame atmosferico realmente acquisito col rivelatore ARGO-YBJ. Sugli assi orizzontali è rappresentata la posizione (X, Y) delle particelle sul piano del rivelatore. Sull'asse verticale sono riportati i loro tempi di arrivo che possono essere associati alla coordinata verticale. Figura tratta da [2].

rivista. La struttura degli sciame di particelle elettricamente cariche è stata studiata con una risoluzione senza precedenti. È stato possibile ricostruirne la distribuzione laterale [5] e le deformazioni indotte dal campo geomagnetico [6]. È stato misurato il flusso della componente leggera dei raggi cosmici, cioè protoni e nuclei di elio [7, 8]. L'osservazione dell'ombra del Sole nel flusso della radiazione cosmica ha permesso di stimare l'intensità del campo magnetico interplanetario e le sue variazioni [9]. È stata studiata la correlazione tra i forti campi elettrici associati ai fenomeni temporaleschi e le variazioni nel flusso di raggi cosmici [10].

La finalità principale dell'esperimento ARGO-YBJ era l'astronomia. Il campo magnetico galattico deflette la traiettoria delle particelle cariche e quindi non è possibile determinarne la direzione iniziale. Così le osservazioni astronomiche sono possibili solo studiando la direzione d'arrivo di radiazione elettricamente neutra, nel caso dell'esperimento ARGO-YBJ fotoni gamma nella banda VHE (*Very High Energy*). Può apparire strano che un rivelatore di particelle cariche venga utilizzato per rivelare radiazione elettromagnetica, ma c'è una spiegazione. I fotoni della banda VHE sono i più energetici e quando penetrano nell'atmosfera terrestre innescano sciame di particelle, analogamente a quello che avviene

per i raggi cosmici carichi. Gli sciame innescati da fotoni (sciame elettromagnetici) hanno caratteristiche diverse rispetto agli sciame innescati da protoni o nuclei di elio (sciame adronici). Ciononostante non è facile distinguere le due tipologie e solo un eccesso di sciame provenienti da una certa direzione nel cielo permette di individuare una sorgente gamma. Con questa tecnica ARGO-YBJ ha monitorato il cielo dell'emisfero nord [11]. Senza interruzioni significative, per oltre 5 anni, ARGO-YBJ ha osservato sorgenti già note (la Nebulosa del Granchio [12], il piano galattico [13], i nuclei galattici attivi Markarian 421 e 501 [14, 15, 16]) segnalando talvolta delle brusche variazioni nel flusso gamma, i cosiddetti *flares*. Particolare attenzione è stata dedicata al *Cygnus cocoon* (bozzolo del Cigno) dove potrebbero trovarsi sorgenti capaci di accelerare i raggi cosmici fino a energie dell'ordine del PeV (10^{15} eV) [17]. Proprio in quella zona del cielo ARGO-YBJ ha scoperto una sorgente gamma [18], ribattezzata appunto ARGO J2031+4157 [19].

L'eredità scientifica

ARGO-YBJ ha dimostrato la fattibilità e l'utilità di un esperimento di alta quota dedicato alla fisica dei raggi cosmici e all'astronomia gamma, in grado di operare in maniera continuativa, con un ampio campo di vista. Infatti gli stessi colleghi cinesi che facevano parte della collaborazione ARGO-YBJ sono attualmente impegnati nel completamento del rivelatore LHAASO (*Large High Altitude Air Shower Observatory*), alla quota di 4400 m nella regione cinese del Sichuan [20]. Su un'area di 145 ettari vengono utilizzate diverse tecniche di rivelazione degli sciame, tra di loro complementari. Molti interessanti risultati sono attesi anche da questa nuova affascinante impresa scientifica.



- [1] B. D'Ettorre Piazzoli, G. Di Sciascio, *Sul tetto del mondo*, *Asimmetrie*, 10 (2010) 20 [www.asimmetrie.it].
- [2] P. Bernardini, *Gamma astronomy with the ARGO-YBJ experiment*, *PoS (Texas 2010)* 166.
- [3] C. Bacci et al. (ARGO-YBJ Collaboration), *Results from the ARGO-YBJ test experiment*, *Astroparticle Physics* 17 (2002) 151.

-
- [4] S. Vernetto: *In Tibet. Tra uomini e Dei*, Lindau, Torino (2008).
- [5] B. Bartoli et al. (ARGO-YBJ Collaboration): *EAS age determination from the study of the lateral distribution of charged particles near the shower axis with the ARGO-YBJ experiment*, *Astroparticle Physics*, 93 (2017) 46.
- [6] B. Bartoli et al. (ARGO-YBJ Collaboration): *Evidence of a geomagnetic effect on extensive air showers detected with the ARGO-YBJ experiment*, *Physical Review D*, 89 (2014) 052005.
- [7] B. Bartoli et al. (ARGO-YBJ Collaboration): *Light-component spectrum of the primary cosmic rays in the multi-TeV region measured by the ARGO-YBJ experiment*, *Physical Review D*, 85 (2012) 092005.
- [8] B. Bartoli et al. (ARGO-YBJ Collaboration): Y.X. Bai et al. (LHAASO Collaboration), *The Knee of the Cosmic Hydrogen and Helium Spectrum below 1 PeV Measured by ARGO-YBJ and a Cherenkov Telescope of LHAASO*, *Physical Review D*, 92 (2015) 092005.
- [9] G. Aielli et al. (ARGO-YBJ Collaboration): *Mean Interplanetary Magnetic Field Measurement Using the ARGO-YBJ Experiment*, *Astrophysical Journal*, 729 (2011) 113.
- [10] B. Bartoli et al. (ARGO-YBJ Collaboration): *Observation of the thunderstorm-related ground cosmic ray flux variations by ARGO-YBJ*, *Physical Review D*, 97 (2018) 04200.
- [11] B. Bartoli et al. (ARGO-YBJ Collaboration): *TeV gamma-ray survey of the northern sky using the ARGO-YBJ experiment*, *Astrophysical Journal*, 779 (2013) 27.
- [12] B. Bartoli et al. (ARGO-YBJ Collaboration): *Crab Nebula: five-year observation with ARGO-YBJ*, *Astrophysical Journal* 798 (2015) 119.
- [13] B. Bartoli et al. (ARGO-YBJ Collaboration): *Study of the diffuse gamma-ray emission from the Galactic plane with ARGO-YBJ*, *Astrophysical Journal*, 806 (2015) 20.
- [14] G. Aielli et al. (ARGO-YBJ Collaboration): *Gamma ray flares from MRK421 in 2008 observed with the ARGO-YBJ detector*, *Astrophysical Journal Letters*, 714 (2010) L208.
- [15] B. Bartoli et al. (ARGO-YBJ Collaboration): *Long-term monitoring on Mrk 501 for its VHE gamma emission and a flare in October 2011*, *Astrophysical Journal*, 758 (2012) 2.
- [16] B. Bartoli et al. (ARGO-YBJ Collaboration): *4.5 Years Multi-Wavelength Observations of MRK 421 During the ARGO-YBJ and FERMI Common Operation Time*, *Astrophysical Journal, Supplement Series*, 222 (2016) 6
- [17] B. Bartoli et al. (ARGO-YBJ Collaboration): *Observation of TeV gamma rays from the Cygnus region with the ARGO-YBJ experiment*, *Astrophysical Journal Letters*, 745 (2012) L22.
- [18] B. Bartoli et al. (ARGO-YBJ Collaboration): *Identification of the TeV Gamma-ray Source ARGO J2031+4157 with the Cygnus Cocoon*, *Astrophysical Journal*, 790 (2014) 152.
- [19] <http://tevcat.uchicago.edu/>
- [20] Huihai He on behalf of the LHAASO Collaboration, *Status and First Results of the LHAASO Experiment*, *PoS (ICRC2019) 693*.

Paolo Bernardini: è Professore Ordinario di Fisica Nucleare e Subnucleare presso l'Università del Salento. Si è principalmente occupato di fisica astroparticellare e di fisica del neutrino. Attualmente partecipa all'esperimento DAMPE su satellite ed alla progettazione degli esperimenti HERD, sempre su satellite, e DUNE negli Stati Uniti.

Machine Learning: **principi di** **funzionamento e** **applicazioni in Medicina**

Giorgio De Nunzio

Dipartimento di Matematica e Fisica "Ennio De Giorgi" - Università del Salento
Istituto Nazionale di Fisica Nucleare, Sezione di Lecce
giorgio.denunzio@unisalento.it

L'articolo espone le basi del funzionamento delle applicazioni mediche dell'Intelligenza Artificiale basate sul *Machine Learning* e illustra sinteticamente alcuni esempi dalla letteratura nei quali l'autore è o è stato coinvolto. L'approccio è essenzialmente divulgativo e pragmatico, con pochi dettagli matematici: la finalità è dare al lettore un'idea abbastanza completa della procedura tipica di realizzazione di un'applicazione di *Machine Learning* e delle difficoltà che si incontrano nel corso del lavoro.

Sistemi di Intelligenza Artificiale per la Medicina basati su *Machine Learning*

La prima parte dell'articolo descrive i sistemi di Intelligenza Artificiale per la Medicina, focalizzando in particolare l'attenzione sul *Machine*

Learning. Per il *Deep Learning*, a parte una breve introduzione, si rimanda alla vasta letteratura disponibile.

Per un elenco degli acronimi di uso comune impiegati nell'articolo, si veda la lista in coda.

Intelligenza Artificiale e *pattern Recognition*

Il termine **Intelligenza Artificiale (IA)** è molto ampio e indica l'abilità di un sistema artificiale di possedere caratteristiche tipicamente umane quali il ragionamento, l'apprendimento, la pianificazione e la creatività. L'IA ha assunto forme diverse nel corso degli ultimi cinquant'anni, passando dai sistemi esperti (*software* di supporto diagnostico-decisionale costituiti da una base di conoscenza specifica, un motore inferenziale, e un'interfaccia per l'utente) a sistemi di *Pattern Recognition*, PR, ossia di riconoscimento di schemi o modelli, basati sull'apprendimento, conven-

zionali (*Machine Learning*, ML) o profondi (*Deep Learning*, DL).

Quest'articolo tratta dei sistemi di ML, dando alcuni dettagli del funzionamento e della struttura, e accenna brevemente al DL. Utile lettura di complemento è [1].

IA in Medicina

In Medicina l'IA è finalizzata a supportare il medico nella diagnosi delle patologie, nelle scelte terapeutiche, nella prognosi, fornendo sostegno infermieristico e medico per il monitoraggio dei pazienti, assistendo durante gli interventi chirurgici per diminuire il rischio d'errore, e così via. I vantaggi di un sistema di IA sono soprattutto a favore del paziente e vanno nella direzione della Medicina personalizzata. Questa tiene conto delle differenze individuali in termini di genetica, microbioma, stile di vita, ambiente ecc. e si basa, quindi, sull'individuazione delle caratteristiche specifiche del singolo paziente, resa possibile dalla grande varietà di parametri fisiologici e patologici rilevabili grazie alle ormai disponibili tecnologie avanzate. Non sono tuttavia da trascurare le ricadute puramente economiche dell'IA sul sistema sanitario, in quanto la maggior accuratezza delle pratiche sanitarie può evitare la prescrizione di esami clinici inutili (favorendo l'appropriatezza) e riduce la probabilità di diagnosi e interventi erranei che costringano al ritorno del paziente nel percorso clinico.

Alcuni termini usuali nel campo dell'IA per la Medicina sono:

- **Radiomics** (Radiomica): termine introdotto di recente, ispirato da altre discipline omiche, ovvero genomica, proteomica o metabolomica; la radiomica è un approccio quantitativo all'*imaging* medicale (*Computed Tomography* o TAC, *Magnetic Resonance imaging* o Risonanza Magnetica, ...) che mira a incrementare e arricchire l'informazione presente nei dati mediante analisi matematiche avanzate. Attraverso il calcolo delle distribuzioni spaziali delle intensità del segnale e delle interrelazioni tra i *pixel*, la radiomica quantifica le informazioni presenti nell'immagine traendo indicazioni non rilevabili mediante la semplice osservazione visiva [2].

- **CADe** (*Computer-Assisted/Aided Detection*): sistema software progettato per individuare automaticamente lesioni (ad esempio tumorali) nei tessuti e quindi avente il fine di ridurre le sviste (in particolare il tasso di falsi negativi, FN) nell'analisi visiva e nell'interpretazione delle immagini mediche (o di altro dato o segnale di interesse diagnostico). Il tipico *output* di un CADe è la posizione di potenziali lesioni (spesso tumori). Comprende solitamente una fase automatica di contornamento dei margini (ovvero segmentazione) molto sensibile (in grado cioè di individuare la totalità delle lesioni), seguita da un algoritmo di ML avente lo scopo di ridurre il numero di falsi positivi (FP). Il *focus* è sull'individuazione di lesioni.
- **CADx** (*Computer-Assisted/Aided Diagnosis*): è progettato per diagnosticare e classificare lesioni note (quindi, ad esempio, per calcolare la probabilità che una lesione nota sia maligna). Può avere una fase di segmentazione automatica/manuale, seguita dal calcolo di *feature* radiomiche ed eventualmente da una fase di ML. Il *focus* è dunque sulla diagnosi.

Spesso l'acronimo CAD è adoperato indifferente per indicare i sistemi CADe e CADx, poiché le tecniche e gli algoritmi alla base sono comuni. Radiomica è un vocabolo di grande successo nella comunità medica e si rivolge specificamente all'analisi delle immagini mentre i CAD possono operare su dati di altro genere (come le serie temporali, ad esempio l'elettroencefalogramma o EEG). Un CAD prevede generalmente un sistema di classificazione di qualche tipo (quindi opera nel campo del ML), mentre un *software* radiomico può limitarsi al calcolo di determinati indicatori da adoperare in base a considerazioni statistiche elementari.

Pattern Recognition (PR)

Un sistema CAD è in grado di individuare *pattern*, ossia schemi comuni all'interno di dati diagnostici, e mettere in relazione tali schemi con una situazione di presenza o assenza di patologia (*detection*), o con una specifica caratterizzazione di una lesione (*diagnosis*). A ciascun *pattern*

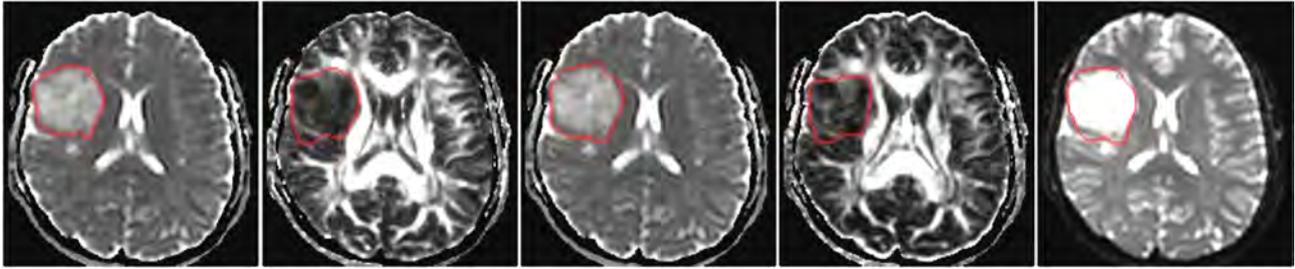


Figura 1: Segmentazione automatica di gliomi cerebrali in immagini di Risonanza Magnetica in Tensore di Diffusione (risultati dell'autore).

corrisponde una classe, ossia una popolazione di oggetti derivanti da osservazioni, aventi in comune delle caratteristiche di rilevanza clinica.

Formalmente, il riconoscimento di un *pattern* è un processo di assegnazione di una classe C (tessuto sano o patologico, lesione di un tipo istologico o di un altro...) a un'osservazione X (un *pixel*, un organo, una lesione). Attore principale del processo è un componente *software* denominato *classificatore*, del quale si parla ampiamente nel seguito.

Le due modalità più diffuse di *pattern recognition* sono:

- **PR supervisionato:** consiste in una tecnica di apprendimento automatico finalizzata a dedurre una funzione di classificazione da un insieme di addestramento nel quale siano state preliminarmente assegnate delle etichette (classi) che rappresentano la verità, in base alle competenze di un supervisore (generalmente un medico); ad esempio, in una classificazione binaria le classi possono essere: sano/malato, o benigno/maligno; in una classificazione multiclasse esse possono riguardare i tipi istologici di una lesione tumorale; una volta addestrato, il sistema è auspicabilmente in grado di riconoscere la classe di appartenenza di nuove osservazioni (generalizzazione);
- **PR non supervisionato:** individua e raggruppa, all'interno di un *dataset*, sottoinsiemi di osservazioni basandosi su misure di similarità; non è necessario definire a priori delle classi ed etichettare le osservazioni e anzi è proprio il classificatore che deve individuare le regolarità nei dati che permettano di dedurre come partizionare in classi il *dataset*.

In Medicina la maggioranza delle applicazioni adotta lo schema di PR supervisionato.

Oltre a stabilire se un tumore sia benigno o maligno, o se un paziente abbia o no una certa patologia, o se la prognosi sia fausta o infausta, la classificazione automatica ottenuta per apprendimento (supervisionato o non) ha anche una diversa applicazione: la segmentazione automatica nelle immagini diagnostiche.

La segmentazione, o contornamento, consiste nell'individuazione dei margini dei tessuti di interesse, quali una lesione tumorale o un organo (figura 1).

Lo scopo è per esempio la misura delle dimensioni della lesione o dell'organo (eventualmente seguendone l'evoluzione nel tempo), o il successivo calcolo di variabili radiomiche all'interno degli stessi, o ancora il contornamento di tessuti da colpire o da risparmiare nel corso del trattamento radioterapico (Piano di Trattamento Radioterapico).

La segmentazione può in realtà essere realizzata mediante procedure diverse dal PR (con tecniche più o meno avanzate di *image processing*) ma i metodi più complessi richiedono usualmente qualche meccanismo di classificazione, perché i *pixel* delle immagini sono classificati come appartenenti o non appartenenti all'organo o tessuto di interesse.

Accanto alla classificazione, i sistemi di IA possono essere finalizzati alla **regressione**. Questa affronta il problema di individuare una relazione funzionale incognita tra variabili misurate su un campione di interesse medico, di cui una o più sono considerate variabili indipendenti, e una o più sono variabili dipendenti: si tratta dunque di un'estensione del caso della classificazione: in quest'ultima la funzione da individuare ha valori discreti (le classi), mentre per la regressione il

valore può essere qualunque (una n -upla di numeri reali). La regressione è un tipo di apprendimento supervisionato: si parte da un insieme di addestramento di cui si conoscono i valori assunti sia dalle variabili indipendenti che da quelle dipendenti, e si allena il *software* a dedurre la relazione funzionale, per poi applicare il sistema a dati nuovi di cui non si conosca il valore delle variabili dipendenti.

Al fine di limitare l'estensione dell'articolo e fissare le idee, il *focus* (tranne saltuari accenni) sarà sulla classificazione con apprendimento supervisionato, in special modo sul ML, e si supporrà che il problema sia di classificazione binaria (variabili *target*: sano/malato, benigno/maligno, tipo istologico A o B...).

Inoltre, si considererà essenzialmente il caso del dato visuale bi- o tri-dimensionale (in particolare, immagini quali la radiografia, la risonanza magnetica, la tomografia computerizzata, l'ecografia, la PET...) tralasciando dati con un numero maggiore di dimensioni (spazio-temporali) e serie temporali (unidimensionali) come EEG e ECG.

Le applicazioni del ML alla Medicina condividono lo schema di sviluppo di seguito illustrato (figura 2) i cui dettagli saranno chiariti nei paragrafi successivi.

1. Il quesito clinico: per sistemi di classificazione, consiste nell'individuazione delle classi di interesse (ad esempio lesione benigna e lesione maligna); per problemi di regressione, è l'individuazione delle variabili cliniche da adoperare come *target*.
2. Acquisizione dei dati (immagini o serie temporali di biosegnali o variabili cliniche scalari) ed eventuale conversione tra formati.
3. Preprocessamento del dato, consistente nella standardizzazione di dati di origine disomogenea (se necessario), e nella pulizia dal rumore e da eventuali artefatti.
4. Eventuale *data augmentation* nel caso di *dataset* sbilanciati, eseguito a livello dei dati (immagini, serie temporali...).
5. Individuazione delle regioni di interesse da considerare (ROI, *Regions of Interest*): organi o lesioni nelle immagini, eventi salienti

nelle serie temporali; eventuale approccio a finestra mobile.

6. Calcolo di variabili (dette indicatori, biomarcatori, *feature*...) nelle ROI, da scegliere tra agnostiche o semantiche, comuni o *domain specific*; eventuale *data augmentation* a livello delle *feature*.
7. Eventualmente, individuazione di una procedura di riduzione della dimensionalità dello spazio delle *feature*: selezione o estrazione.
8. Scelta di uno schema di apprendimento e validazione, con il partizionamento dei dati negli insiemi di *training*, *validation* e, possibilmente, *test*, ai fini della procedura denominata *hold-out cross validation*, o di altro schema quale il *k-fold cross validation*.
9. Scelta del modello di classificatore; ottimizzazione degli iperparametri con iterazione della procedura di allenamento e validazione, calcolando ad ogni *loop* delle appropriate figure di merito.
10. Verifica, sul *test set*, dell'avvenuta generalizzazione (ossia della capacità di fornire risposte corrette per dati non visti durante la procedura di apprendimento).
11. Trasferimento all'ambito clinico, se possibile.

La procedura è basata su un approccio iterativo che prevede di ripetere vari STEP a seconda del successo delle scelte via via operate, ottimizzando il sistema per la massima *performance*.

Lo STEP 1 è solo apparentemente semplice e banale: in realtà alle volte può non essere facile individuare una sola variabile *target* perché i medici stessi ne valutano contemporaneamente diverse, oppure usano misure fatte in momenti diversi della storia clinica del paziente, o ancora non hanno sempre la disponibilità delle variabili di interesse per l'intero campione considerato: in tal caso una discussione approfondita, allo scopo di identificare la scelta migliore, è delicata e fondamentale.

Lo STEP 2 spesso è un ostacolo non trascurabile, soprattutto nel caso delle immagini: il formato DICOM, estremamente dettagliato e personalizzabile, è uno *standard* inclusivo che permette di inserire in svariati modi, nei *file*, dati e metadati:

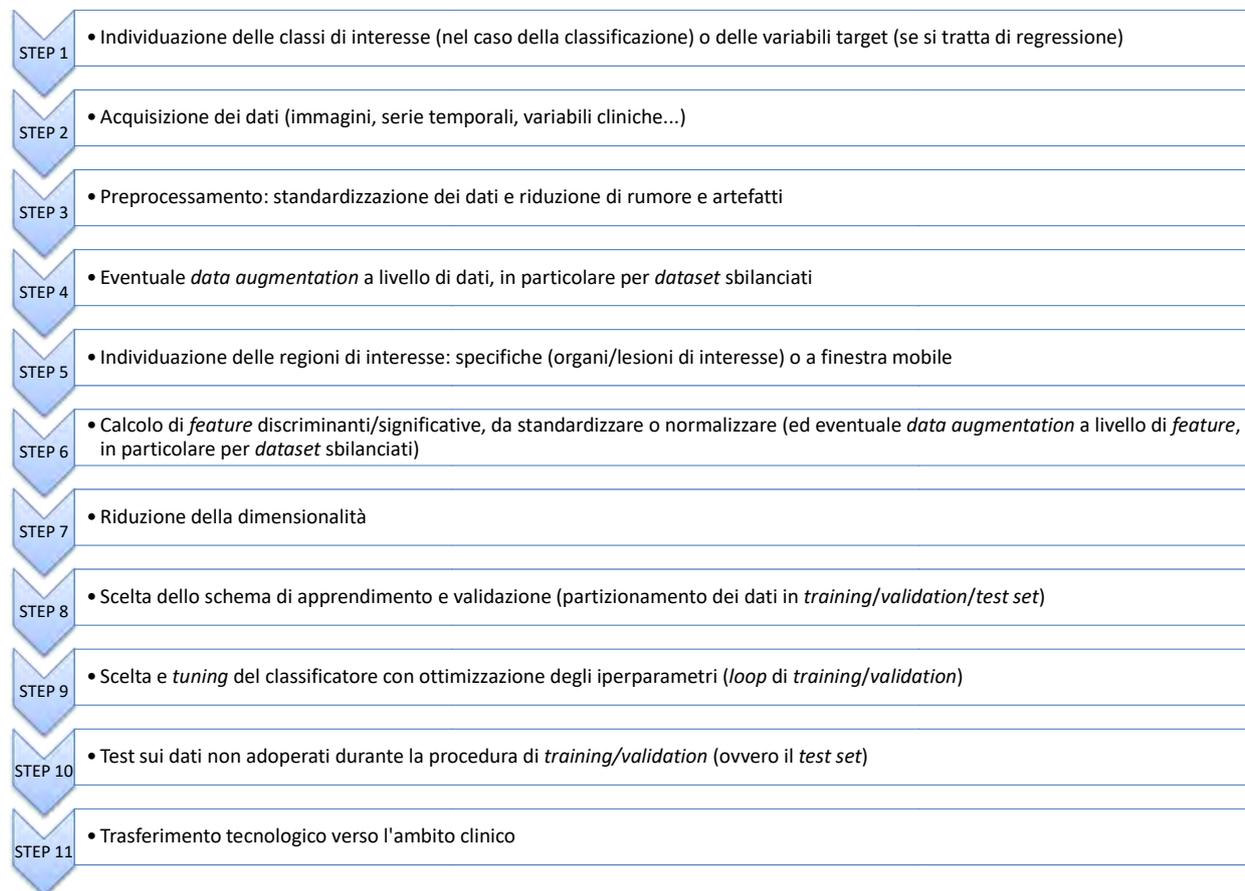


Figura 2: Schema di sviluppo di un sistema di ML per la Medicina.

i codici scritti per leggere e convertire le immagini in formati più agili devono spesso scontrarsi con questa versatilità, in particolare quando l'origine dei dati sia molteplice (studi multicentrici, che acquisiscano i dati da diversi presidi ospedalieri).

Lo STEP 3 si basa su algoritmi e tecniche di trattamento di dati e immagini ed è anch'esso molto importante soprattutto nel caso di studi multicentrici: quando, infatti, i dati provengano da diverse fonti, sebbene teoricamente tra loro compatibili essi in realtà possono presentare differenze di *range* di valori acquisiti e nelle curve di calibrazione (e quindi nel significato fisico e clinico dei valori misurati, come avviene in special modo per le immagini di risonanza magnetica); nel caso delle immagini, un ulteriore possibile problema è la disparità di *voxel size* (dimensioni del volume acquisito corrispondente a un *pixel* nell'immagine) e l'eventuale presenza di *voxel* non cubici (non isometrici o isotropici); in tutti questi casi è necessario uniformare i dati applicando riscalature opportune e interpo-

lare il dato spazialmente o temporalmente. Il preprocessamento consiste, dunque, usualmente in: normalizzazione delle intensità, omogeneizzazione delle immagini per *voxel size* con eventuale isometrizzazione del *voxel* (interpolazione per portarsi a dimensioni standard quali $1\text{ mm} \times 1\text{ mm} \times 1\text{ mm}$, figura 3), riduzione del rumore e degli artefatti.

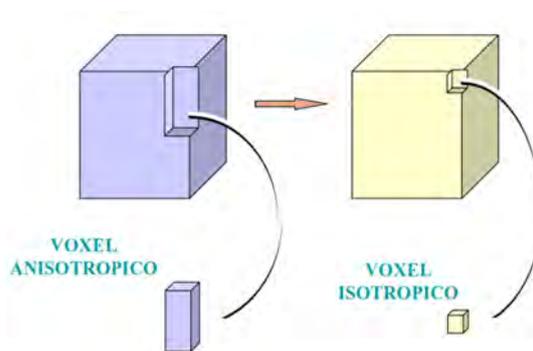


Figura 3: Differenza tra *voxel* anisotropico e *voxel* isotropico.

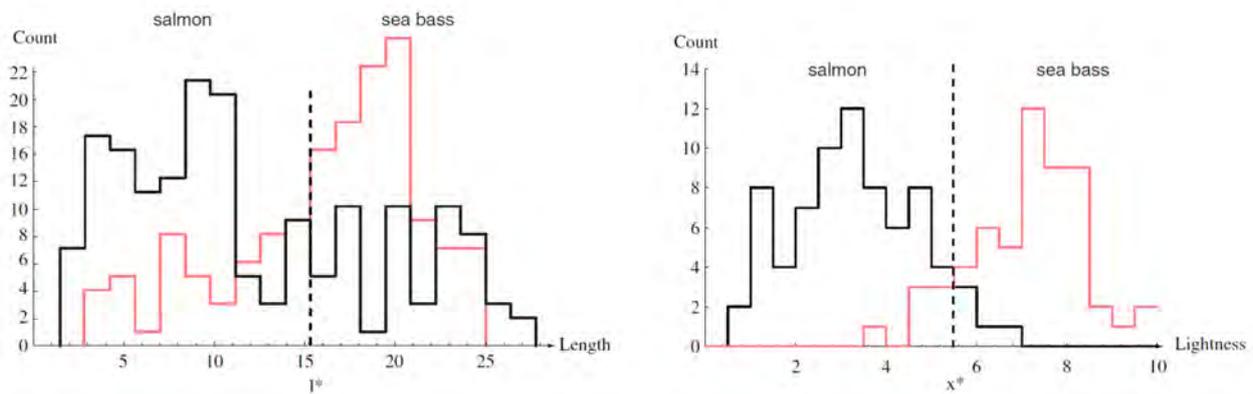


Figura 4: A sinistra, istogrammi della lunghezza dei pesci (per specie); a destra, istogrammi della lucentezza delle scaglie. Gli istogrammi sono calcolati nel training set. Da [3].

Lo STEP 10 può essere, per applicazioni mediche, arduo da eseguire in quanto spesso i *dataset* disponibili non sono tanto ampi da garantire la presenza di dati non sfruttati per *training* o *validation*, sui quali testare il sistema ormai ottimizzato. Dell'importanza di questo argomento si parlerà tuttavia nel paragrafo dedicato al punto 8.

Lo STEP 11 meriterebbe un trattamento esteso e quindi non sarà discusso.

Ove figurassero, accanto a dati complessi e strutturati come immagini e serie temporali, anche semplici variabili cliniche scalari, esse andranno direttamente ad arricchire il *set* di *feature* da adoperare nei calcoli di ML: in tal caso, naturalmente, alcuni degli STEP indicati perdono di significato, in particolare l'individuazione delle regioni di interesse (STEP 5) e il calcolo di *feature* al loro interno (STEP 6).

I prossimi paragrafi approfondiscono lo schema indicato, in particolare i punti da 4 a 9. Allo scopo, però, di dare un'idea immediata del funzionamento di un sistema di ML, conviene preliminarmente mostrare un esempio non specifico delle applicazioni alla Medicina, bensì estratto da un ben noto testo didattico [3] al quale si rimanda per approfondimenti. In tal modo sarà più facile in seguito comprendere i dettagli più tecnici.

Come funziona il *Machine Learning* supervisionato: un esempio

I gestori di un impianto di confezionamento di pesce desiderano automatizzare il processo di

smistamento in entrata in base alla specie, in particolare distinguendo il branzino dal salmone, utilizzando il rilevamento ottico e metodi matematici [3]. Possono essere esplorate diverse misure che promettono di essere discriminanti tra i due tipi di pesci, ad esempio la lunghezza o la larghezza, la lucentezza della pelle (dovuta alla riflettanza delle scaglie), misure legate alla posizione della bocca, e così via. Il fine ultimo è elaborare una regola decisionale di classificazione che riduca al minimo il costo della procedura, a sua volta dipendente anche dal numero di errori compiuti.

È possibile immaginare una procedura basata su tre fasi: (1) preprocessamento e segmentazione (i fotogrammi registrati dal rilevatore ottico, raffiguranti il nastro trasportatore e i pesci, sono elaborati in modo da isolare i pesci l'uno dall'altro e dallo sfondo), (2) estrazione di *feature* (le immagini dei singoli pesci sono inviate a un sistema di misura, che rileva/calcola le proprietà desiderate), e (3) classificazione (i valori delle *feature* sono passati a un classificatore che valuta le informazioni, deduce il legame esistente tra i valori delle *feature* e la classe alla quale ciascun pesce appartiene, e costruisce un modello per discriminare tra le specie). Si tratta evidentemente di un approccio supervisionato perché, nella fase di creazione del modello di classificazione, a ciascun pesce è associata l'informazione sulla classe alla quale esso appartiene; il modello è costruito su una parte dei dati disponibili (*training set*) e successivamente validato su un'altra parte (*validation set*).

Per la costruzione del modello, iniziamo con-

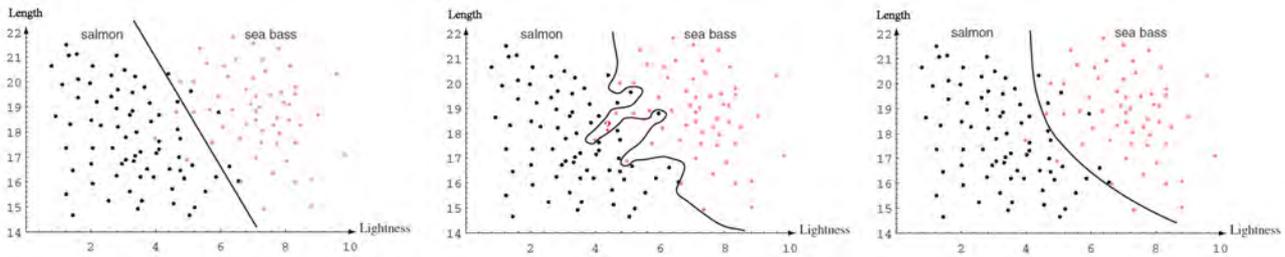


Figura 5: Scatter plot delle due feature considerate, lunghezza dei pesci e lucentezza delle loro scaglie (per specie) (da [3]). Da sinistra verso destra, il confine di decisione è una retta, una linea curva complessa, una linea curva dolce. Gli scatter plot sono calcolati nel training set.

siderando una singola *feature*, ad esempio la lunghezza del pesce. La conoscenza del dominio applicativo consente di affermare che un branzino è generalmente più lungo di un salmone (il branzino ha una lunghezza tipica, e questa è maggiore della lunghezza tipica del salmone). Dal punto di vista matematico, costruiamo gli istogrammi (per specie) della lunghezza dei pesci che abbiamo selezionato per il *training set* (figura 4, a sinistra) e osserviamo se i due istogrammi appaiono sufficientemente separati.

È facile immaginare un sistema elementare di decisione (una regola) basato sull'applicazione di una soglia: se la lunghezza del pesce è maggiore o uguale a un valore opportuno l^* , allora il pesce è un branzino, altrimenti è un salmone. Il valore di soglia va ottimizzato in maniera da minimizzare gli errori di classificazione. Esaminando gli istogrammi, però, ci rendiamo conto che il sistema non è ottimale perché, qualunque sia la soglia scelta, molti pesci saranno classificati in maniera erranea.

Si può ripetere il discorso su altre *feature*, ad esempio sulla lucentezza del manto di scaglie (figura 4, a destra). Le conclusioni saranno analoghe: tale *feature* da sola non basta. Possiamo allora pensare di considerare le due misure insieme, rappresentando ciascun pesce come una coppia (o vettore) di *feature*: $x = (x_1, x_2)$ dove x_1 e x_2 sono la lunghezza e la lucentezza. Poiché essa può essere vista come una coppia di coordinate che definiscono un punto nello spazio delle *feature*, ciascun oggetto (un pesce) sarà di fatto modellizzato in tal modo. Di conseguenza, classificare, ossia distinguere branzini da salmone, sarà equivalente a tracciare nello spazio delle *feature* un confine (una linea non necessariamente retta) che separi i punti appartenenti alle due

specie.

In base alla modellizzazione bidimensionale dei pesci, lo spazio delle *feature* può essere rappresentato come in figura 5, dove la linea di confine tra le classi (da sinistra a destra) è rispettivamente una retta, una curva che separi perfettamente le due classi, e infine una curva dolce in qualche modo intermedia tra i casi precedenti. La prima situazione darà un numero di errori non indifferente; la seconda non è realistica e mostra il cosiddetto *overfitting* ai dati di *training*, ovvero un eccessivo adattamento a questi dati durante il processo di apprendimento, come illustrato in figura 6; infine, la terza soluzione, pur dando luogo a un certo numero di errori risulta il miglior compromesso tra accuratezza sul *training set* e capacità di generalizzare a nuovi punti (appartenenti per esempio al *validation set*).

Compito del classificatore è trovare, a partire dai dati di *training*, la linea di separazione ottimale. Qualora tale linea sia una retta, il problema è detto linearmente separabile.

Generalizzando, possiamo utilizzare più di due variabili (in generale una n -pla), dimodoché un pesce sia modellizzato da un vettore (o da un punto) in uno spazio a n dimensioni e la linea di separazione tra i domini delle due classi diventi una (iper)superficie. Si osservi che alcune *feature* potrebbero essere ridondanti (a causa della correlazione tra variabili) o computazionalmente costose da ottenere, e può capitare che l'aumento eccessivo della dimensionalità possa diminuire le prestazioni. In conclusione, l'apprendimento supervisionato, tipico del ML, necessita di un set di dati di apprendimento (*learning set*) annotato (ovvero, etichettato). Il *dataset* si presenta usualmente sotto la forma di una matrice (figura 7) in cui le righe (più raramente, le colonne) sono

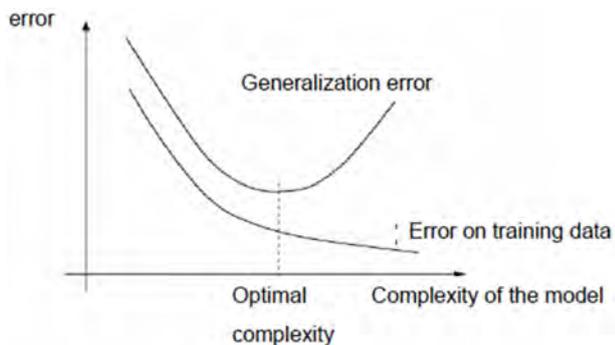


Figura 6: Un modello che si adatti eccessivamente ai dati di training può causare overfitting, a scapito della capacità di generalizzare. La figura mostra qualitativamente come l'aumento di complessità del modello porti a una diminuzione dell'errore di classificazione sul set di training, ma un peggioramento della capacità di generalizzare (aumento dell'errore su validation set e test set)

vettori di *feature*: ciascuna riga contiene i dati di un oggetto (un campione) del *dataset* e ciascuna colonna corrisponde a una *feature*; una colonna, inoltre, spesso la prima o l'ultima, contiene le etichette, ossia la classe di appartenenza di ciascun campione: in caso di classificazione binaria di un problema diagnostico medico, detta colonna potrebbe contenere ad es. 0 o 1, rispettivamente per normale e patologico. L'etichetta è assegnata da un esperto (il supervisore) e contiene la verità o comunque lo *standard* di riferimento (*ground truth*, *golden truth*, o *gold standard*). Il classificatore ha il compito di realizzare la separazione ottimale tra i campioni nello spazio delle *feature*, come giusto compromesso tra la separazione nel *training set* e la capacità di generalizzare tale separazione a dati nuovi.

Una volta comprese le basi del funzionamento di un sistema di ML, conviene entrare nei dettagli di alcuni degli STEP della procedura introdotta con la figura 2.

STEP 4: Dataset sbilanciati e Data Augmentation

I *dataset* sbilanciati, ossia quelli nei quali una delle classi (consideriamo una classificazione binaria) è considerevolmente meno rappresentata dell'altra, sono prevalenti in una moltitudine di campi e settori, come i servizi finanziari (transazioni fraudolente vs genuine) e la diagnostica medica. Ad esempio, le lesioni maligne sono me-

Label	Feature-1	Feature-2	Feature-3	Feature-4	...	Feature-n
C1	x_1^1	x_2^1	x_3^1	x_4^1	...	x_n^1
C2	x_1^2	x_2^2	x_3^2	x_4^2	...	x_n^2
C3	x_1^3	x_2^3	x_3^3	x_4^3	...	x_n^3
...
Cm	x_1^m	x_2^m	x_3^m	x_4^m	...	x_n^m

Figura 7: Esempio di matrice di feature, con i vettori disposti per righe.

no diffuse di quelle benigne e i pazienti malati sono in generale meno delle persone sane. Il ML in caso di *dataset* sbilanciati è particolarmente delicato perché gli algoritmi devono identificare casi rari in grandi insiemi di dati. A causa della disparità delle classi, l'algoritmo tende a categorizzare anche i casi rari nella classe maggioritaria, venendo meno al suo compito primario che è proprio l'individuazione di istanze della classe minoritaria. Inoltre molte misure di qualità tendono ad avere valori elevati nonostante il sistema di fatto non funzioni, dando la falsa sensazione di un modello altamente accurato.

Vi sono due diversi approcci al problema, di cui il primo è il più adatto per insiemi di dati poco numerosi.

- **Data augmentation o oversampling (aumento o sovracampionamento dei dati).** Si incrementa la numerosità della classe di minoranza aggiungendo copie leggermente modificate di dati già esistenti. Esempi: trasformazioni geometriche delle immagini come le rotazioni, aggiunta di rumore, o cancellazione casuale di parti dell'immagine. Un'alternativa è il sovracampionamento dei vettori di *feature* dopo il loro calcolo (quindi nello STEP 6) e un algoritmo noto è SMOTE (*Synthetic Minority Over-sampling Technique*). I nuovi vettori di *feature* sono da adoperarsi esclusivamente nel *training set*.
- **Undersampling (sottocampionamento).** Al contrario del precedente, ha lo scopo di ridurre la numerosità della classe maggioritaria, diminuendo quindi la quantità di dati

effettivamente adoperati, in modo che sia paragonabile a quella della classe di minoranza. È ottenuta selezionando casualmente una parte delle osservazioni della classe di maggioranza. È possibile effettuare un sottocampionamento informato, osservando la distribuzione dei dati della classe di maggioranza e selezionando opportunamente i campioni da includere. Un modo è l'uso preliminare di una tecnica di *clustering* nell'insieme maggioritario, che individui la stratificazione naturale dei dati. A questo punto, la scelta dei campioni da usare sarà fatta casualmente ma rispettando la proporzione tra i diversi cluster, mantenendo così la distribuzione originale. Anche in questo caso il bilanciamento può avvenire nello STEP 6.

In caso di *dataset* sbilanciati, ove non si intervenga con i metodi indicati, è opportuno valutare la qualità dei sistemi di classificazione in maniera specifica, ossia evitando di utilizzare la curva ROC che non è adatta a problemi affetti da squilibrio tra le classi, e adoperando piuttosto la curva Precision-Recall. Non usare, inoltre, l'accuratezza come cifra di merito (per i dettagli sui metodi di misura della qualità del sistema di ML, si veda il paragrafo corrispondente in quest'articolo).

STEP 5: Individuazione di regioni di interesse nelle immagini

Questo paragrafo è dedicato specificamente ad applicazioni del ML in immagini diagnostiche. In caso i sistemi di ML siano invece rivolti a dati come serie temporali EEG o ECG, sarà necessario adattare opportunamente il discorso, pur restando immutati i concetti di base.

Come visto nell'esempio della discriminazione tra branzini e salmoni, i calcoli delle *feature* (le variabili discriminanti in grado di differenziare tra oggetti appartenenti a classi diverse) sono condotti localmente, in regioni di interesse, o ROI. Le ROI possono essere, nelle applicazioni ad immagini diagnostiche mediche:

- lesioni o organi: la variabile *target* di classificazione potrebbe essere la malignità o non malignità di una lesione tumorale conclamata, o la presenza di una particolare patologia

in un organo; simile è il caso in cui vi siano diverse regioni sospette (candidati) di cui non si conosca e si voglia stabilire la natura: il *target* potrebbe essere allora la natura eventualmente nodulare di regioni composte da *pixel* ipo- o iper-intensi (perché le caratteristiche della patologia, e la fisica sulla quale si basa la tecnica di *imaging*, garantiscono che i noduli cercati abbiano questa caratteristica); in generale, le ROI possono essere delimitate manualmente da un medico, oppure automaticamente con tecniche di processamento di immagini;

- l'interno di una piccola finestra che scansioni sistematicamente l'immagine, eventualmente con sovrapposizione parziale da una posizione all'altra: è l'approccio a finestra mobile (o *sliding window*) in cui la classe è attribuita come etichetta al *pixel* centrale della finestra; in fase di validazione, una volta terminata l'esplorazione esaustiva del tessuto e l'etichettatura dei *pixel* in base alla classe, l'accorpamento dei *pixel* equivalenti individua le regioni appartenenti all'una o all'altra classe; questo caso è tipico della segmentazione di un'immagine.

Una volta definite le ROI, occorre scegliere le *feature* da calcolare.

STEP 6: Calcolo di *feature*

Come detto, preliminarmente all'addestramento di un classificatore è necessario effettuare, sui dati o immagini dei pazienti e all'interno delle ROI, il calcolo di *feature* significative per il problema affrontato, ovvero di proprietà misurabili (biomarcatori, attributi) che siano in grado ad esempio di differenziare tra una lesione benigna e una maligna. L'insieme delle *feature* così calcolate può essere visto come un vettore a n componenti, o una n -upla di coordinate in uno spazio a n dimensioni. In definitiva la ROI è modellizzata da un vettore di misure e l'insieme dei dati da adoperare per il processo di addestramento costituisce una matrice con tante righe quanti sono i campioni e tante colonne quante sono le *feature*, più la colonna *target*.

La scelta di *feature* informative, discriminanti e indipendenti è un compito cruciale nel ML. Esse sono solitamente numeriche, ma possono

anche essere categoriali, booleane, stringhe, ecc. Le variabili numeriche sono più facili da trattare, mentre è spesso opportuno o necessario che quelle categoriali siano convertite in numeriche. Le *feature* possono essere **semantiche** o **agnostiche**: le prime sono legate al contesto e derivano dall'uso comune (in radiologia potrebbero essere misure di vascolarizzazione, di forma, o che descrivano la spiculazione di una lesione, ovvero la presenza di margini irregolari e infiltranti); le seconde colgono la natura di una ROI mediante descrittori matematici di uso generale, come le *feature* di Haralick (descritte nel prosieguo). Il loro calcolo è anche denominato *feature extraction*.

In letteratura è stata descritta una grande varietà di *feature* potenzialmente utilizzabili per la caratterizzazione di una ROI di un'immagine. A priori, anche esaminando il problema specifico, è spesso difficile dedurre quali siano le più adatte, per cui si preferisce procedere in maniera empirica inserendo nei sistemi di classificazione ampi *set* di variabili di diversa natura e verificandone a posteriori l'utilità, ossia la capacità di discriminazione tra oggetti di classi diverse.

Poiché esistono molti modi e formule diverse per calcolare le *feature*, si raccomanda l'adesione alle linee guida dell'*Image Biomarker Standardization Initiative* (IBSI) [6], che offrono un consenso per calcoli di *feature* radiomiche standardizzate. Un esempio di libreria molto usata in linguaggio python è *pyradiomics* [7].

Vi sono *feature* basate sulla *texture* delle ROI e altre basate su caratteristiche geometriche (area e forma).

Feature basate sulla tessitura

Con il termine *texture*, o tessitura [5], si definisce una disposizione geometrica eventualmente (grosso modo) ripetitiva dei livelli di grigio o dei colori di un'immagine (figura 8: alcuni esempi di tessiture tratte dal *database* Brodatz¹). Nei prossimi paragrafi il riferimento sarà alle immagini in toni di grigio, le più diffuse in ambito medico.

L'analisi tessiturale è una tecnica che caratterizza una tessitura calcolandone le caratteristiche distintive con metodi matematici e statistici. L'analisi tessiturale valuta l'intensità di grigio di

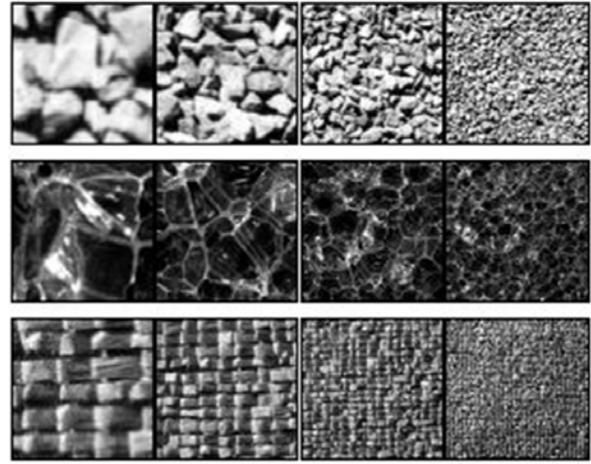


Figura 8: Esempi di texture (tratte dal database Brodatz).

ogni *pixel* nella zona in esame, in relazione ai *pixel* circostanti.

Negli ultimi decenni sono state sviluppate molte tecniche per l'estrazione di *feature* impiegate nell'analisi di *texture*. Tra queste si elencano usualmente anche quelle dette del primo ordine (media, deviazione standard o varianza, *skewness* e *kurtosis* dei livelli di grigio): si parla di *feature* del primo ordine perché esse sono calcolabili a partire dall'istogramma dei livelli di grigio, e quindi sono legate alla probabilità di osservare un particolare valore di grigio in una locazione singola scelta a caso nella ROI. Più propriamente tessiture sono invece quelle di secondo ordine (calcolate su statistiche di coppie di *pixel*, le matrici di co-occorrenza) o di ordine superiore. Come esempio, di seguito sono descritte le *feature* di Haralick (del secondo ordine) e i *Local Binary Pattern*.

Feature di Haralick. Le *feature* di Haralick sono calcolate a partire dalle matrici di co-occorrenza (COM, *Co-Occurrence Matrix*) dei livelli di grigio, o GLCM (*Gray Level Co-occurrence Matrix*) [5]. Le GLCM permettono l'estrazione di informazioni statistiche riguardanti la distribuzione di coppie di *pixel* (relazioni tra due *pixel* distinti), pertanto sono anche denominate istogrammi di secondo ordine (o bidimensionali) dei livelli di grigio. Esse sono legate alla probabilità di osservare una coppia definita di valori di grigio nei *pixel* situati alle estremità di un segmento di lunghezza data e orientamento fissato, posto nella ROI dell'immagine in una posizione casuale. Per maggiori dettagli si faccia riferimento

¹<http://sipi.usc.edu/database/database.php?volume=textures>, sito visitato in novembre 2021

all'Appendice 1.

Local Binary Pattern Il *Local Binary Pattern* (LBP) [10] è un operatore locale basato sulla tessitura, definito come segue. Si voglia calcolare il descrittore nell'intorno di un dato *pixel* di un'immagine a toni di grigio; si considerino i suoi otto primi vicini (distanza di Chebyshev pari a 1). Ai valori dei vicini è applicata un'operazione di soglia rispetto al valore del *pixel* centrale: se il livello di grigio del *pixel* è minore della soglia, ad esso è assegnato il valore 0, altrimenti 1. Si otterrà, in questo modo, un byte espresso in base 2, che verrà convertito in un valore decimale da assegnare al *pixel* dato.

La figura 9 mostra un esempio di codifica. Estendendo il concetto è possibile tenere conto di *pixel* distanti più dei primi vicini o effettuare il calcolo in 3D.

Feature basate sulla forma

Tra le *feature* basate sulla forma della ROI si citano la circolarità, l'eccentricità, i descrittori di Fourier. La figura 10, tratta da [11], ne mostra alcune.

In particolare:

Circularità. La circolarità C_{cir} è un fattore di forma utilizzato in molti campi dell'analisi di immagini (2D), in particolare in microscopia. Esso descrive la forma di particelle e oggetti indipendentemente dalla loro dimensione. Esistono diverse definizioni di circolarità, ma una molto diffusa è il quoziente isoperimetrico, dato dalla seguente formula [5]:

$$C_{cir} = \frac{4\pi A}{P^2},$$

dove A è l'area della ROI e P il suo perimetro. La circolarità di un cerchio è per definizione pari a 1 (sebbene ciò sia vero solo approssimativamente per un cerchio tracciato in un'immagine digitale) mentre, per un oggetto molto diverso da un cerchio, esso diminuisce fino a 0 (caso ideale del segmento di retta).

Eccentricità. L'eccentricità e di un'ellisse è il rapporto tra la distanza tra i fuochi e la lunghezza dell'asse maggiore. Il valore è compreso tra 0 (il caso del cerchio) e 1 (un segmento di retta). Il concetto è esteso a oggetti di forma non ellittica, ponendo e pari all'eccentricità dell'ellisse

che ha momenti del secondo ordine (calcolati sulle coordinate) uguali a quelli dell'oggetto. Eccentricità e circolarità esprimono concetti simili ma non coincidenti e sono grandezze tra loro indipendenti.

Descrittori di Fourier. Forniscono una codifica della forma di un oggetto 2D tramite la trasformata di Fourier della linea che ne definisce il bordo. Per i dettagli si rimanda al testo [5].

Standardizzazione o normalizzazione delle Feature

Prima di essere utilizzate nella procedura di *training* e *validation*, è opportuno che le *feature* siano normalizzate, tipicamente tra 0 e 1, o standardizzate, ossia convertite mediante la trasformazione

$$z_j^k = \frac{x_j^k - \mu_j}{\sigma_j},$$

dove μ_j e σ_j sono rispettivamente la media e la deviazione standard della *feature* x_j , di cui x_j^k è il valore per il k -mo campione. Lo scopo è evitare che variabili con valore medio o *range* di variazione elevati siano erroneamente considerate dal classificatore più importanti di altre di piccola entità e varianza, compromettendo il processo ottimale di addestramento. È importante che la procedura di standardizzazione o normalizzazione sia eseguita in modo da non causare *bias* statistico (vedere il paragrafo dedicato agli schemi di apprendimento e validazione). In particolare, i parametri della procedura (minimi e massimi delle *feature*, o loro medie e deviazioni standard, a seconda del metodo che si vuole applicare) devono essere rigorosamente calcolati solo nel *training set* e con questi parametri l'operazione è applicata ai *dataset* di *training*, *validation* e *test*.

A questo punto la matrice di *feature* è pronta per essere utilizzata all'interno di uno schema di addestramento e validazione, per condizionare un classificatore in modo che esso, dopo aver imparato dai dati di *training* come distinguere i vettori di *feature* riferiti a ciascuna classe, sia poi in grado di applicare queste conoscenze a nuovi dati (*set* di validazione e poi di *test*) generalizzando quanto appreso.

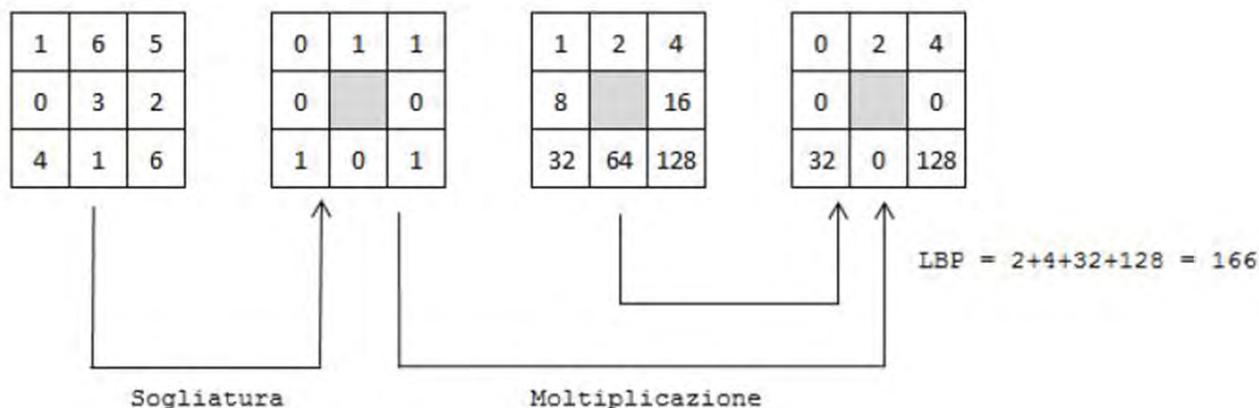


Figura 9: Codifica attraverso Local Binary pattern del pixel centrale di una regione 3×3 .

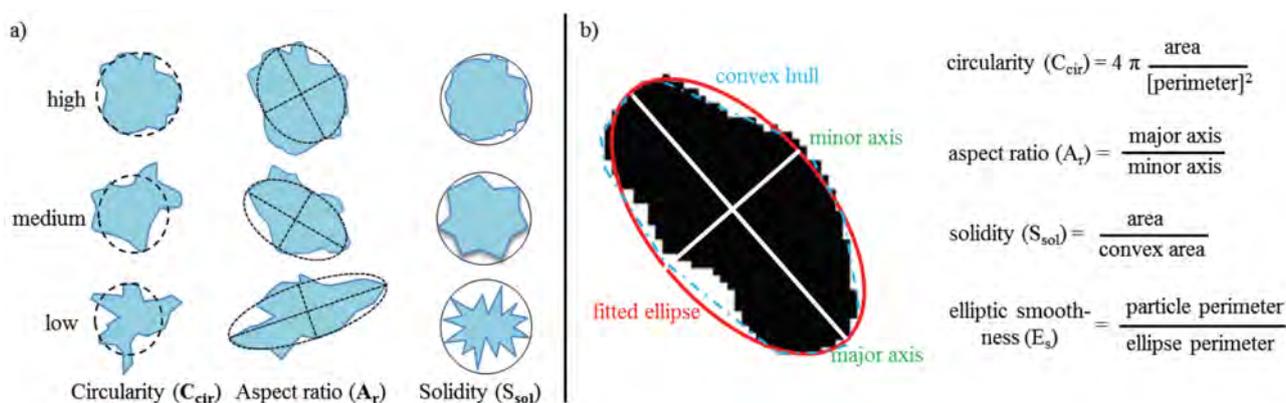


Figura 10: a) rappresentazione schematica di Circolarità (C_{cir}), Aspect ratio (A_r) e Solidità (S_{sol}) di particelle da proiezione bidimensionale; b) immagine al microscopio di una particella con le equazioni per il calcolo delle feature geometriche. Da [11].

Potrebbe tuttavia essere necessario uno *step* ulteriore detto di riduzione della dimensionalità, del quale si occupa il paragrafo seguente.

STEP 7: Riduzione della dimensionalità

L'uso di un gran numero di *feature* in un problema di ML ha il vantaggio di aumentare la ricchezza di informazioni e quindi la probabilità di inserire, nel *set* di variabili, quelle effettivamente discriminanti. Tuttavia l'eccessiva dimensionalità del problema può inficiare la qualità del risultato, sia perché un maggior numero di indicatori implica un tempo di calcolo maggiore, sia soprattutto perché si rischia di incorrere nella cosiddetta *curse of dimensionality*, la maledizione della dimensionalità. Il termine, definito per la prima volta da Richard E. Bellman in [12], indica il fatto che l'aumentare della dimensionalità dello spazio delle variabili, senza un contemporaneo

incremento nel numero di campioni disponibili, porta a una sparsità dei dati con conseguente rischio di *overfitting* e perdita di qualità.

Allo scopo di risolvere il problema è possibile ricorrere a tecniche di riduzione della dimensionalità [4], mediante le quali si individuano e si eliminano le variabili superflue conservando solo quelle effettivamente utili (*feature selection*). In alternativa si possono applicare opportune trasformazioni allo spazio delle *feature* in modo da massimizzare l'efficacia delle variabili risultanti, di cui si conservano solo le migliori eliminando quelle poco utili (*feature extraction*).

Vantaggi della riduzione della dimensionalità sono: scongiurare la maledizione della dimensionalità evitando l'*overfitting*, accorciare il tempo di addestramento del classificatore e, nel caso della *selection*, semplificare i modelli per scopi interpretativi.

Feature Selection. Seleziona un *subset* effica-

ce e ottimale dall'insieme completo delle *feature*. Allo scopo l'approccio più semplice è valutare il potere discriminante delle singole variabili, ottenibile ad esempio con il calcolo del coefficiente di Fisher (o indice di separazione tra classi, o *Fisher's Score*) [3]:

$$F = \frac{(\mu_- - \mu_+)^2}{\sigma_-^2 + \sigma_+^2},$$

dove μ_- e σ_- sono rispettivamente la media e la deviazione standard di una *feature*, calcolate sui campioni negativi (classe 0), e, analogamente, μ_+ e σ_+ sono calcolate per i campioni positivi. È intuitivo che detto *score* assuma valori elevati per variabili che separino bene le due classi, e valori bassi per variabili poco discriminanti, perché massimizza la varianza interclasse allo stesso tempo minimizzando la varianza totale intraclasse.

La tecnica di selezione mediante l'approccio a variabili indipendenti può essere utile, anche perché va nella direzione della spiegabilità dei risultati, ma non è l'opzione migliore, perché non considera le relazioni tra *feature*. Dalla figura 11 appare evidente che le due *feature* mostrate sotto forma di *scatter plot* non sono singolarmente discriminanti (immaginiamo di proiettare i numeri, che rappresentano i campioni delle classi 1 e 2, sugli assi, costruendo degli istogrammi, e ce ne renderemo facilmente conto); è tuttavia evidente che le due variabili, considerate insieme, sono perfettamente discriminanti, perché permettono il passaggio, tra le due nuvole di punti, di un confine di decisione estremamente semplice (una retta).

Appare chiaro, quindi, che occorrono tecniche di selezione più elaborate che considerino le *feature* insieme. Citiamo, senza scendere nei dettagli, la *Sequential Forward Selection* (che costruisce un insieme di variabili ridotto e ottimale partendo da quella più discriminante e via via aggiungendone altre), la *Sequential Backward Selection* (che invece considera inizialmente l'insieme completo di variabili e poi via via elimina le meno utili), la selezione basata su algoritmi genetici, tecniche LASSO (*Least Absolute Shrinkage and Selection Operator*) ecc.

Feature Extraction. Applicano opportune trasformazioni allo spazio delle variabili, con la finalità di calcolarne di più utili alla soluzio-

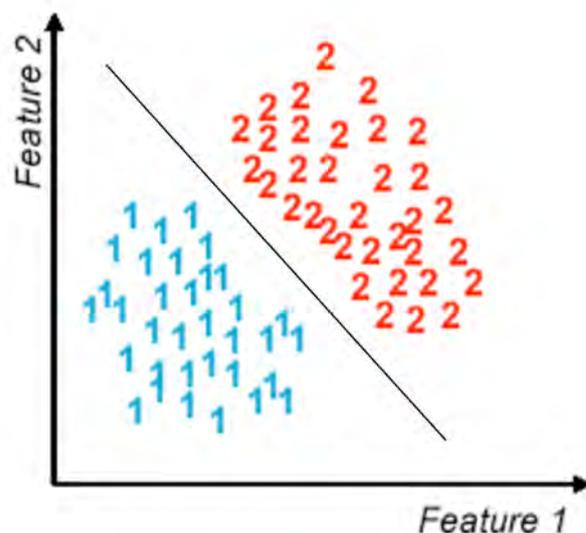


Figura 11: Due *feature* singolarmente non discriminanti, ma ottimali se considerate insieme.

ne del problema, trascurando quelle meno utili derivanti dalla trasformazione. Tra queste, la PCA (*Principal Component Analysis*), la LDA (*Linear Discriminant Analysis*), la ICA (*Independent Component Analysis*). In particolare, quella forse maggiormente adoperata per l'estrazione di *feature* orientata al ML è la LDA (o FLDA: *Fisher Linear Discriminant Analysis*). La FLDA è un metodo di riduzione della dimensionalità ma, nel caso di un problema a due classi, è di fatto un metodo di classificazione in sé, perché il risultato della riduzione è un'unica variabile ottimale. La FLDA applica una trasformazione lineare allo spazio delle *feature*, finalizzata a massimizzare il già citato coefficiente di Fisher. Essa definisce quindi nuovi assi di riferimento per i dati. Al termine della trasformazione, avviene la proiezione su un singolo asse (riduzione della dimensionalità ad una sola variabile) scegliendo come asse di proiezione quello a cui compete il massimo coefficiente di Fisher. Ciò porta ad una separazione ottimale tra i dati delle due classi (figura 12).

STEP 8: Scelta di uno schema di addestramento e validazione, classificazione, e calcolo della qualità di un classificatore

La classificazione supervisionata, caratteristica del ML (e del DL) necessita di uno schema di addestramento e validazione incrociata, ossia una

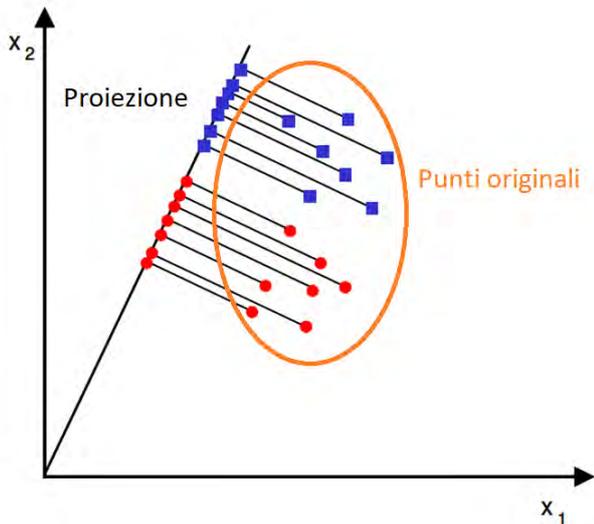


Figura 12: Esempio di applicazione della FLDA. L'asse di proiezione individuato è ottimale e massimizza la separabilità, ora ottenibile tramite semplice sogliatura.

procedura con la quale il classificatore è allenato su parte del *dataset* (*training set*) e l'apprendimento è verificato su un'altra parte di esso (*validation set*). Verificare l'apprendimento sullo stesso insieme di *training* non dà, infatti, una misura affidabile della qualità che il classificatore avrebbe su nuovi dati, che è invece la finalità ultima della procedura (capacità di *generalizzare*).

Nell'approccio più semplice, definito *hold-out cross validation* [13] i dati sono divisi in (almeno) due parti, gli insiemi di *training* e di *validation*. Il primo serve per l'ottimizzazione degli iperparametri del classificatore (per la minimizzazione dell'errore di classificazione) mentre il secondo è usato solo per la verifica (*unbiased evaluation*, ossia una valutazione non condizionata dai dati sui quali il classificatore è stato allenato) dell'apprendimento e dell'effettiva generalizzazione (capacità di classificare nuovi campioni). Tipicamente il partizionamento avviene in percentuali di 70%–30%, oppure 80%–20%. La procedura di allenamento e verifica è condotta iterativamente, ogni volta modificando leggermente gli iperparametri del modello o il tipo di classificatore sperimentato, fino a giungere al risultato migliore.

In realtà, oltre a ricavare dai dati i due insiemi di *training* e *validation*, è opportuno estrarre e mettere da parte un terzo insieme di campioni, denominati di *test*. Alla fine del processo iterati-

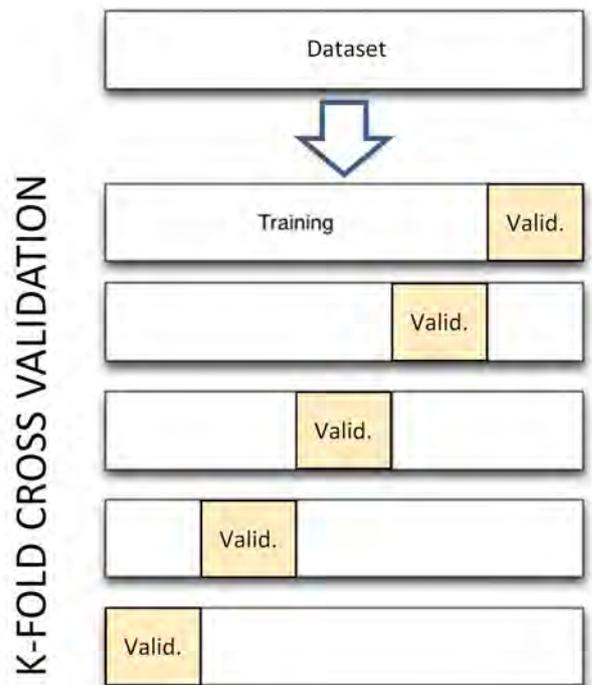


Figura 13: Schema di *k-fold cross validation* con $k = 5$.

vo di ottimizzazione degli iperparametri, il classificatore dovrebbe essere testato su questi dati per ottenere una misura affidabile della qualità. Infatti, con la ripetizione del ciclo di allenamento e verifica il modello addestrato diventa sempre più sbilanciato (*biased*) verso il *set* completo di apprendimento ($training \cup validation$) finché la verifica su un insieme terzo (*final unbiased evaluation*) diventa importante. Fatto ciò, il modello non dovrebbe più essere variato in alcun modo.

Accanto al semplice *hold-out* esistono svariati altri schemi di valutazione dell'apprendimento [14]. Tra questi, il *k-fold cross validation* o *k-fold CV*, e una procedura simile detta *leave-one-out cross validation* (LOO, LOO-CV). Nel *k-fold CV* (figura 13) partizioniamo l'insieme di apprendimento in un numero k di *fold* (ovvero in k parti). Una è riservata alla validazione mentre $k - 1$ parti sono adoperate per l'addestramento (se $k = 5$ quest'ultimo corrisponde all'80% dei dati). I pronostici sono accantonati per il calcolo delle cifre di merito (e.g. la curva ROC). Quindi l'allenamento è ripetuto prendendo una parte (diversa) per la convalida e il resto per l'allenamento. Si reitera complessivamente k volte. La curva ROC o altri dati complessivi di merito sono calcolati dalle previsioni cumulate. Il LOOCV è una procedura analoga, in cui però k coincide con

il numero di campioni. Si utilizzano anche le sigle LOPO o LOSO, riferendosi specificamente a pazienti o soggetti.

Particolare attenzione va fatta, al momento del partizionamento dei dati, nel caso in cui a ciascun campione (paziente, soggetto) corrisponda più di un vettore di *feature* (è il caso per esempio di un'applicazione medica in cui vi siano misure fatte su più campioni biotipici, come nel caso delle biopsie multiple per la diagnosi del tumore della prostata): in tal caso, per evitare *bias*, si dovrà avere l'accortezza di non suddividere mai lo stesso paziente tra insieme di allenamento e insieme di validazione.

STEP 9 Classificatori: struttura, addestramento e misura della performance

Esiste una grande varietà di classificatori supervisionati, con una tassonomia molto complessa. Un elenco non esaustivo comprende artificial neural networks, support vector machines, decision trees, metodi basati su gradient boosting e tanti altri.

Già è stato chiarito come la procedura di classificazione si basi sulla definizione di una regola di decisione che partizioni lo spazio delle *feature* in regioni disgiunte, ognuna delle quali sia attribuibile ad una delle classi note (regioni di decisione), separate da una frontiera. I classificatori si distinguono per il metodo che adoperano per individuare tale partizionamento. In quest'articolo si darà qualche informazione sul neurone artificiale (il perceptrone) e le reti neurali artificiali con esso costruite (il *multi-layer perceptron*).

Neuroni biologici e artificiali

Le reti neurali artificiali sono modelli matematici composti da semplici elementi (*Processing Elements*) operanti in parallelo, il cui funzionamento trae ispirazione dai sistemi nervosi animali. Il cervello umano contiene circa 10^{11} cellule denominate neuroni. Esse sono caratterizzate da tre parti principali (figura 14): soma o corpo cellulare, dendriti, e assone. Il soma è la parte centrale della cellula, contenente il nucleo, e da esso si irradiano prolungamenti sottili detti dendriti, e l'assone. I dendriti ricevono i segnali provenienti

dagli assoni degli altri neuroni e quindi costituiscono gli ingressi della cellula, mentre attraverso l'assone il neurone trasmette segnali ad altre cellule. La comunicazione tra assone e dendrite avviene alla giunzione, detta sinapsi. Ogni neurone è tipicamente connesso a un migliaio di altri neuroni e il numero totale di sinapsi nel cervello umano supera 10^{14} [15].

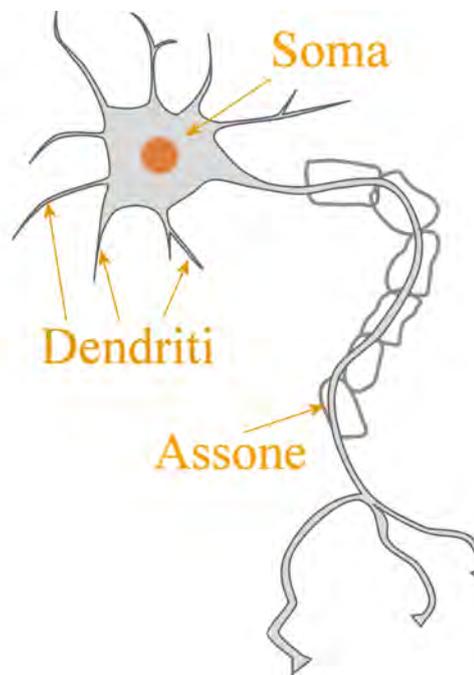


Figura 14: Struttura di un neurone biologico.

Semplificando, un neurone si può trovare essenzialmente in due stati: a riposo o attivo. Quando passa dallo stato di riposo a quello attivo, esso produce un impulso elettrico detto potenziale di azione (generalmente uno *spike* di circa 35 mV di picco), poi trasmesso dall'assone. Il segnale, raggiunte le sinapsi, provoca il rilascio di sostanze chimiche denominate neurotrasmettitori che attraversano la giunzione ed entrano, viaggiando nel dendrite, nel corpo del neurone connesso. Le sinapsi possono essere eccitatorie o inibitorie: nel primo caso i neurotrasmettitori aumentano la probabilità che il successivo neurone si attivi, nel caso contrario la diminuiscono. Ad ogni sinapsi è associato un peso che determina la natura e l'intensità dell'effetto eccitatorio o inibitorio. Se la somma pesata degli ingressi provenienti dagli altri neuroni supera, entro i tempi propri delle attivazioni neuroniche, una specifica soglia, allora il neurone si attiva.

Un neurone opera in tempi dell'ordine di 1 ms

e quindi è un sistema di elaborazione lento se confrontato con un *personal computer* i cui tempi tipici delle operazioni elementari vanno dai microsecondi dei primi modelli ai nanosecondi attuali. Tuttavia, il gran numero di neuroni e sinapsi che operano in parallelo rende il cervello un sistema potente, oltre a conferirgli alta tolleranza a informazioni affette da rumore e ad eventuali danni di estensione limitata (robustezza), e le facoltà di apprendimento e generalizzazione.

La struttura del neurone artificiale emula quella del neurone biologico [16]. Prendendo come esempio il perceptrone, introdotto da Rosenblatt nel 1958 [17] partendo dalle teorie di McCulloch e Pitts [18], vediamo che esso è costituito da (figura 15):

- alcuni ingressi x_i che ricevono informazioni dall'esterno (equivalenti ai dendriti del neurone biologico); vi sono tanti ingressi quante sono le *feature* utilizzate dal modello;
- un peso w_i su ciascun ingresso; tali pesi rappresentano l'intensità del trasferimento dell'informazione tra neuroni e modellano le sinapsi;
- un *input* ulteriore, posto sempre a 1 e denominato *bias* (naturalmente, con significato diverso rispetto al *bias* statistico di cui si è parlato più volte nell'articolo); il peso associato all'ingresso di *bias* è indicato con b ; il significato dell'*input* di *bias* sarà chiaro nel prosieguo;
- una funzione di attivazione $f()$: può assumere varie forme ma nel caso del singolo perceptrone essa è la funzione a gradino;
- un valore di *output* y , informazione trasmessa al neurone successivo, dato dal risultato dell'applicazione di f alla somma pesata degli ingressi: $y = f(\sum_i w_i x_i + b)$.

Il neurone artificiale è in grado di apprendere

In definitiva, il perceptrone riceve un insieme di valori in *input* (il vettore di *feature*) e, in base ai pesi e a una funzione di attivazione, produce un determinato *output*. Un singolo perceptrone può essere visto come un caso particolare di rete neurale artificiale *feedforward* (ovvero in cui l'informazione viaggia in un solo verso, dagli *input*

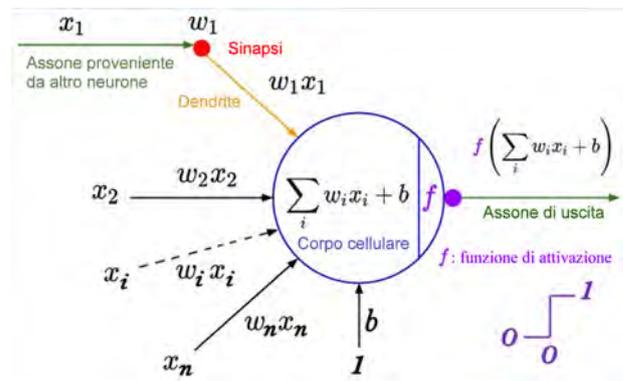


Figura 15: Struttura di un neurone artificiale. (Adattata da: <https://towardsdatascience.com/a-gentle-introduction-to-neural-networks-series-part-1-2b90b87795bc>, sito visitato in novembre 2021.)

verso l'*output*, in contrapposizione alle reti ricorsive), avente un solo strato di neuroni (composto da un solo neurone) e ad una sola uscita. Vogliamo che il neurone artificiale fornisca, come *output* dell'assone, la classe a cui appartiene il vettore (e quindi l'oggetto o il campione modellizzato da quel vettore). Senza perdere in generalità, possiamo supporre che l'*output* valga 0 o 1. È possibile dimostrare che una scelta appropriata dei pesi consente la corretta classificazione in caso di problemi linearmente separabili.

La procedura che porta a determinare il valore dei pesi è *error driven*, poiché sfrutta l'errore compiuto dal perceptrone nella classificazione, per aggiustarne l'entità, in base alla *perceptron learning rule* introdotta da Rosenblatt. I pesi sono inizialmente assegnati in maniera *random*. A seguire, l'aggiustamento è realizzato fornendo ciclicamente e più volte come *input* i vettori dell'insieme di *training* e calcolando il valore di uscita o del neurone. Ad ogni iterazione i pesi sono aggiornati secondo la formula:

$$w_{\text{new}} = w_{\text{old}} + \eta(t - o)x_i,$$

dove x_i è l' i -mo vettore di *feature*, t e o sono rispettivamente l'*output* desiderato (la classe vera alla quale appartiene il vettore di *feature* x_i) e l'*output* del neurone, mentre η è il tasso di apprendimento (o *learning rate*), relativamente piccolo per garantire la convergenza del processo iterativo.

Per una dimostrazione della capacità del perceptrone di operare come classificatore, si

veda il box “Il perceptrone come classificatore lineare”.

Il perceptrone ha capacità limitate dal fatto che il problema debba essere linearmente separabile. Per problemi più complessi si utilizzano reti di neuroni (figura 17) che prevedono uno strato (o livello) di connessioni di *input*, uno di neuroni di *output* (un solo neurone se il problema è di classificazione binaria, più neuroni in caso di classificazione multiclasse) e uno o più strati nascosti (ossia situati tra i livelli di ingresso e di uscita) popolati da neuroni (*hidden layers*). Ciascun neurone è caratterizzato da una propria funzione di attivazione che, a seconda delle applicazioni, può essere una sigmoide, una tangente iperbolica, o altre funzioni non lineari.

La regola di apprendimento in questo caso diventa più complessa. Un esempio è la regola di *backpropagation* [19], chiamata così perché l'errore si propaga all'indietro dalle unità di uscita verso lo strato di ingresso, via via modificando i coefficienti di peso, dapprima relativi al collegamento tra l'ultimo strato e il penultimo, poi a quelli tra il penultimo e il terz'ultimo, e così via. La procedura è realizzata iterativamente, iniziando i pesi con valori casuali e poi sottoponendo più volte alla rete gli esempi del *dataset* di *training*, finché non è raggiunto un errore finale complessivo ritenuto tollerabile (o sono soddisfatti altri criteri di stop). Nel *backpropagation* sono adoperate le derivate delle funzioni di attivazione tra i vari neuroni, per cui queste devono essere derivabili.

La tipologia di problema risolvibile con una rete neurale dipende dalla sua struttura o architettura, ossia dal numero di strati nascosti. Facendo riferimento alla figura 18 si osserva come il perceptrone sia in grado di risolvere esclusivamente problemi linearmente separabili, ossia in cui le classi possono essere separate da un iperpiano. Una rete con un singolo strato nascosto può risolvere problemi non linearmente separabili, in cui il *decision boundary* identifichi un dominio convesso, mentre un doppio strato nascosto permette la soluzione di problemi non linearmente separabili con confine arbitrario.

Qualità di un classificatore

Un classificatore, una volta addestrato, sarà in grado di mappare, nella maniera più corretta possibile anche compatibilmente con la qualità dei dati, un vettore di *feature* verso la sua classe di appartenenza. È naturalmente importante poter esprimere quantitativamente la qualità con cui avviene la mappatura, in modo da valutare l'efficacia dell'apprendimento e della capacità di generalizzare. Le prestazioni di un sistema di classificazione sono misurate attraverso l'introduzione di diversi parametri, tra i quali:

- numero di veri positivi, veri negativi, falsi positivi, falsi negativi, e la matrice di confusione
- sensibilità, specificità, precisione, recupero (*recall*), accuratezza, *F-measure*
- curva ROC e curva *precision-recall*.

Consideriamo nel prosieguo sistemi di classificazione binaria, fermo restando che le considerazioni esposte e le grandezze calcolate si estendono con facilità alla classificazione multiclasse. È opportuno, date le due classi *target*, stabilire quale considerare la classe dei positivi e quale quella dei negativi, ove questo non sia naturale, perché alcuni termini e alcune figure di merito fanno riferimento esplicito a positivi e negativi e quindi ciò semplifica l'argomentazione. Naturalmente la convenzione trae origine dal linguaggio medico dove i termini positivo e negativo si riferiscono rispettivamente a elementi patologici o sani, ed è esteso a un caso generico di oggetto di interesse (positivo) e oggetto non di interesse (negativo), o di distinzione tra qualità di pari importanza come ad esempio la discriminazione tra due tipi istologici di tumore.

Gran parte dei classificatori fornisce in uscita un valore continuo (ad esempio tra 0 e 1) e solo l'applicazione di una soglia discretizza l'*output* che così identifica la classe proposta (classificatore discreto binario, con i soli possibili valori di *output* 0 e 1). Vi sono figure di merito calcolate dopo l'applicazione della suddetta soglia, e altre che si applicano prima della soglia e permettono di dare un valore di qualità indipendente dalla soglia stessa.

Per meglio comprendere il significato dell'applicazione della soglia, conviene far riferimento alla figura 19 che rappresenta (qualitativamente)

Il perceptrone come classificatore lineare

Per dimostrare che il perceptrone è in grado di classificare, ove il problema sia linearmente separabile, consideriamo il suo *output*, pari a $y = f(a)$ dove $a = \sum_i w_i x_i + b$ e $f()$ è la funzione di attivazione (funzione a gradino).

Nel semplice caso di due variabili, questo significa che $f(a) = 1$ se e solo se $w_1 x_1 + w_2 x_2 \geq -b$, quindi (se $w_2 \neq 0$):

$$x_2 \geq -\frac{w_1}{w_2} x_1 - \frac{b}{w_2}.$$

Se identifichiamo x_1 e x_2 rispettivamente con le variabili dipendente e indipendente x e y , e poniamo $m = -(w_1)/(w_2)$ e $n = -b/w_2$, la disequaglianza precedente diventa $y \geq mx + n$. Ciò significa che l'*output* del neurone è 1 se e solo se la coppia di coordinate, e quindi la coppia di *feature*, (x_1, x_2) identifica un punto nel semipiano superiore delimitato dalla retta $y \geq mx + n$ (figura 16). Resta quindi dimostrato che il perceptrone possa risolvere problemi di classificazione linearmente separabili.

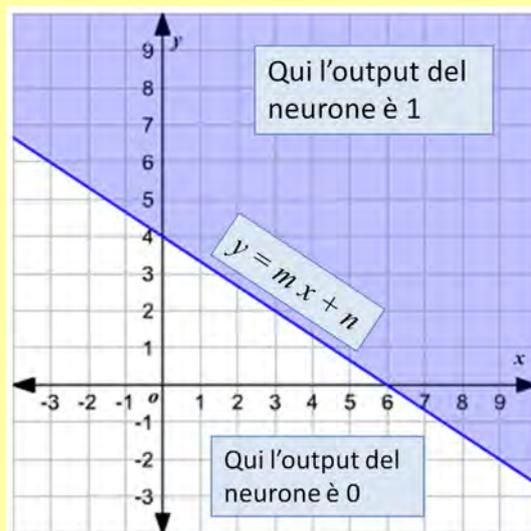


Figura 16: Output di un perceptrone.

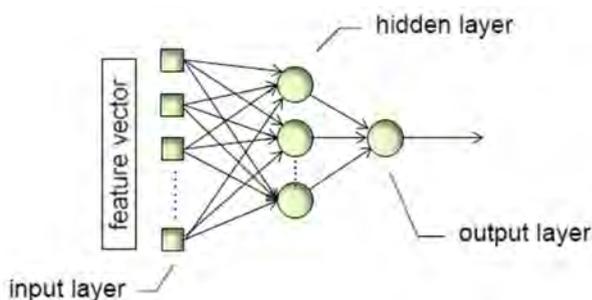


Figura 17: Esempio di rete neurale artificiale feedforward caratterizzata da uno strato di connessioni di input (di numerosità pari al numero di feature), uno strato di neuroni di output (in questo caso con singolo neurone per la classificazione binaria) e da uno strato nascosto popolato da tre neuroni.

STRUCTURE	TYPES OF DECISION REGIONS	EXCLUSIVE OR PROBLEM	CLASSES WITH MESHEDED REGIONS	MOST GENERAL REGION SHAPES
SINGLE-LAYER 	HALF PLANE BOUNDED BY HYPERPLANE			
TWO-LAYER 	CONVEX OPEN OR CLOSED REGIONS			
THREE-LAYER 	ARBITRARY (Complexity Limited By Number of Nodes)			

Figura 18: Tipologia di problemi risolvibili tramite una rete neurale artificiale feedforward, in base alla sua struttura. Dall'alto verso il basso i casi raffigurati sono: rete a singolo strato (perceptrone), rete con due strati (di cui uno nascosto) e rete con tre strati (di cui due nascosti) (da [20]).

L'*output* del classificatore, prima dell'applicazione di una soglia, mediante istogrammi per classe normalizzati. Come spesso succede nella realtà, si è scelta una rappresentazione campaniforme senza con ciò perdere di generalità. Per chiarire meglio le idee, nel caso del perceptrone la quantità rappresentata sarebbe $\sum_i w_i x_i$.

Supponiamo ora di applicare una soglia θ_0 . In figura il valore di soglia è stato scelto come pun-

to di intersezione tra le curve che approssimano gli istogrammi, ma questa è solo una delle scelte possibili. Consideriamo i campioni che contribuiscono alla parte indicata con a : essi sono casi effettivamente positivi che, essendo sopra soglia, anche il classificatore ha indicato come positivi. I campioni nella parte b sono invece negativi

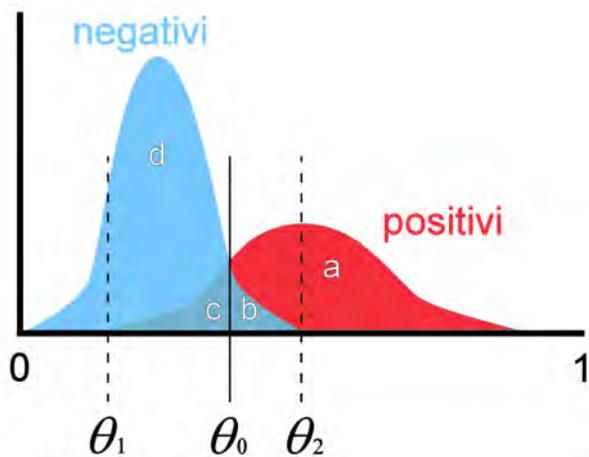


Figura 19: Istogrammi per classe, normalizzati, del valore di uscita del classificatore, prima dell'applicazione di un'operazione di sogliatura.

ma, essendo a destra della soglia, il classificatore li ha erroneamente classificati come positivi. Analogamente si considerino le parti *c* e *d*.

È possibile allora dare le seguenti definizioni:

- se il campione è positivo (patologico) ed è classificato come tale dalla rete neurale ($output \geq \theta_0$), esso costituisce un vero positivo (TP) e contribuisce alla parte etichettata come *a*; se invece è classificato come negativo dalla rete neurale ($output < \theta_0$), esso costituisce un falso negativo (FN) (caso *c*);
- se il campione è negativo (sano) ed è classificato come tale, esso è contato come vero negativo o TN (caso *d*), se è classificato come positivo, è definito falso positivo o FP (caso *b*).

Da queste definizioni di base si possono dedurre altre grandezze di interesse, ad esempio (indicando con FP il numero di falsi positivi, ecc.):

- **sensibilità** o *TP rate*, o recupero (*recall*): è la frazione di campioni positivi correttamente classificati: $TP/(TP + FN)$,
- **rapporto dei falsi positivi** (*FP rate*) o frequenza dei falsi allarmi: $FP/(TN + FP)$,
- **specificità** (è la frazione di campioni negativi correttamente classificati): $TN/(TN + FP)$, (complementare rispetto a 1 di *FP rate*)
- **accuratezza** (frazione di campioni correttamente classificati): $(TP + TN)/(TP + TN + FP + FN)$,
- **precisione**: $TP/(TP + FP)$,

- **F-measure** o *F-score* o *F1-score*:

$$2 \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

è la media armonica tra precisione e richiamo

- **matrice di confusione:**

$$\begin{pmatrix} TP & FP \\ FN & TN \end{pmatrix}$$

(definizione non univoca, poiché fonti diverse definiscono la matrice di confusione come la trasposta di quella mostrata, oppure scambiano sia le righe che le colonne).

Un sistema di IA per la Medicina deve avere:

- alta sensibilità, in modo da rilevare tutte le possibili lesioni (o tutti i pazienti affetti da una patologia);
- alta specificità, per evitare falsi allarmi e inutili e dispendiosi approfondimenti diagnostici.

Si noti, sempre dalla figura 19, che la scelta della soglia influisce direttamente sui valori di sensibilità e specificità (e sulle altre figure di merito elencate). Infatti, se la soglia è $\leq \theta_1$, ossia a sinistra del valore minimo di *output* per i campioni positivi, evidentemente tutti i positivi saranno classificati come tali e la sensibilità sarà massima e pari a 1; tuttavia, la specificità sarà molto bassa. Al contrario, se la soglia è $> \theta_2$, ossia a destra del valore massimo di *output* per i campioni negativi, tutti i negativi saranno classificati come tali e la specificità sarà massima uguale a 1, a scapito di una sensibilità ridotta. Valori intermedi della soglia daranno combinazioni diverse di sensibilità e specificità, sempre tra loro in competizione e compresi nell'intervallo (0,1).

Come accennato in precedenza, nel caso di *dataset* sbilanciati è consigliabile l'uso di *F-measure* al posto dell'accuratezza, perché quest'ultima tende a dare valori eccessivamente ottimistici anche quando il classificatore favorisce la classe maggioritaria.

È opportuno introdurre un ulteriore indicatore, detto curva ROC (*Receiver Operating Characteristics*), che fornisce una misura della capacità di

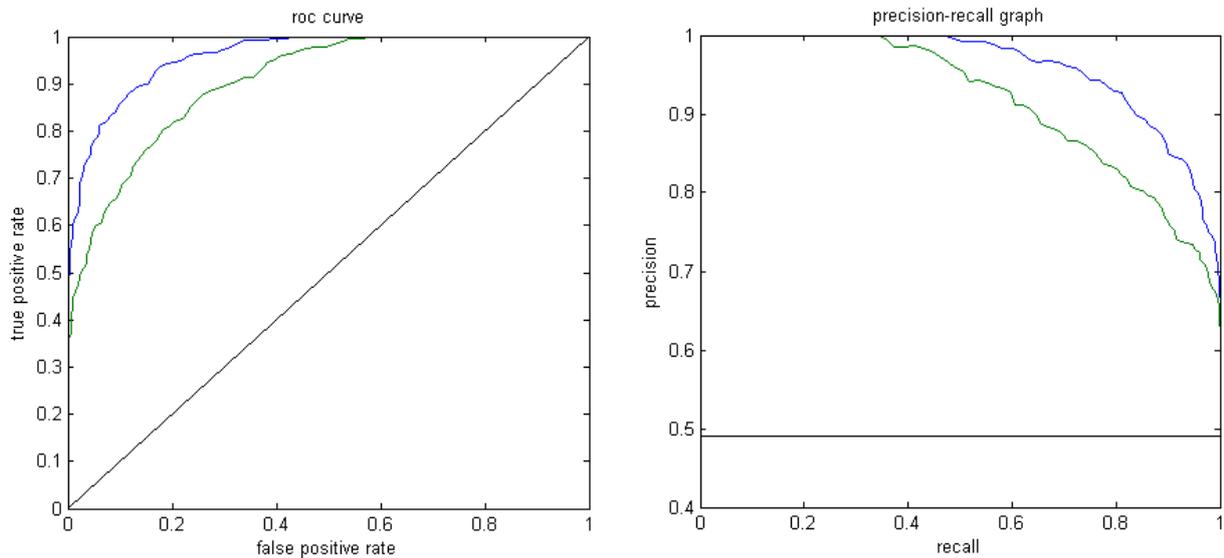


Figura 20: Sinistra: curve ROC; destra: curve P-R. Le curve in blu sottendono un'area maggiore e quindi caratterizzano un classificatore di qualità migliore.

scriminante del sistema indipendentemente dalla soglia applicata (ossia a partire dal suo *output* continuo), e in maniera grafica.

Lo spazio ROC è uno spazio bidimensionale in cui l'ascissa è rappresentata dal valore di *FP rate* e l'ordinata dalla sensibilità. All'interno di questo spazio, necessariamente delimitato dal quadrato $[0, 1] \times [0, 1]$, possono essere posti punti di coordinate $(1 - \text{specificità}, \text{sensibilità})$ che rappresentano punti di lavoro diversi del classificatore. Facendo variare la soglia di uscita del classificatore tra i valori estremi, il luogo di tali punti descriverà in generale una curva, detta curva ROC (figura 20, a sinistra), che lega la sensibilità al *FP rate*, cioè descrive la relazione fra veri e falsi allarmi. Applicando un determinato valore di soglia, la rete neurale diventa un classificatore discreto binario e lavorerà in un singolo punto definito nello spazio ROC.

Punti di lavoro diversi (definiti dalla soglia) sono scelti in base all'importanza relativa di specificità e sensibilità. Ad esempio, in un problema medico è spesso più importante riconoscere tutti i positivi, sia pure con il rischio di un numero non indifferente di FP, per cui si prediligerà un punto ad alta sensibilità. In molti casi, un punto ottimale è l'intersezione della curva con la seconda bisettrice del quadrante, dove la sensibilità è uguale alla specificità, oppure il punto della curva più vicino alle coordinate $(0,1)$, che rappresentano il punto ideale.

Prescindendo dal valore della soglia, la capacità discriminante di un classificatore è legata all'area sottesa dalla curva ROC (*Area Under the Curve*, AUC o AUROC or ROC AUC). Nel caso di un *test* perfetto, ossia di specificità e sensibilità entrambe pari a 1, il valore di AUC corrisponde all'area dell'intero quadrato delimitato dai punti di coordinate $(0,0)$, $(0,1)$, $(1,0)$, $(1,1)$, e assume valore 1 (probabilità del 100% di una corretta classificazione). Al contrario, la ROC per un *test* assolutamente privo di valore informativo è rappresentata dalla diagonale che passa per l'origine, con $AUC = 0.5$, corrispondente a una scelta casuale della classificazione.

Per l'interpretazione dei valori dell'area sottostante la curva ROC è possibile riferirsi allo schema proposto in [24]:

- $AUC = 0.5$: sistema diagnostico non informativo;
- $0.5 < AUC < 0.7$: poco accurato;
- $0.7 \leq AUC < 0.9$: moderatamente accurato;
- $0.9 \leq AUC < 1.0$: altamente accurato;
- $AUC = 1.0$: perfetto.

Nel caso della *k-fold CV* o del LOO, anziché calcolare la AUC ad ogni iterazione, spesso i risultati predittivi acquisiti nelle differenti fasi di *training* e *validation* sono accumulati e la curva ROC è calcolata alla fine della procedura.

Infine, qualora si abbia a che fare con un *data-set* sbilanciato, è opportuno adoperare, al posto

della curva ROC, la curva *Precision-Recall* (P-R) (figura 20, a destra). La curva P-R visualizza il compromesso tra precisione e richiamo al variare della soglia. Un'area elevata sotto la curva rappresenta sia un *recall* che una *precision* elevati. La P-R AUC (o AUPR) è meno comoda da trattare in quanto il suo valore di base (legato alla *baseline* della curva e corrispondente a un test non informativo) non è fisso ma dipende dall'entità dello squilibrio.

Machine Learning vs Deep Learning

Il DL è una particolare architettura basata sul ML, da cui si differenzia per alcuni aspetti importanti. Esaminare in profondità il funzionamento del DL (o, quanto meno, delle reti neurali convoluzionali, o CNN, la più nota implementazione del DL) comporterebbe diverse pagine ed è al di fuori delle finalità di quest'articolo, essenzialmente divulgativo. Saranno pertanto mostrate brevemente le caratteristiche di base e saranno evidenziate le differenze con l'approccio ML, rimandando per i dettagli alla letteratura specifica (ad esempio [21]).

Il DL è nato su ispirazione del funzionamento della corteccia visiva cerebrale e i processi di visione biologica (figura 21) ed è diventato uno strumento estremamente efficace nella visione artificiale e in generale in compiti visivi (riconoscimento di forme e oggetti in immagini, in particolare in immagini naturali ma anche in immagini di diagnostica medica).

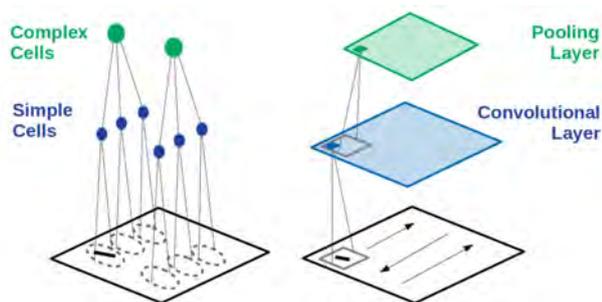


Figura 21: Relazione tra la struttura della corteccia cerebrale visiva e i componenti base di una rete neurale convoluzionale, da [21].

Hubel e Wiesel, due neurologi premi Nobel, scoprirono nel 1962 [22] che singoli neuroni da loro denominati *cellule semplici*, situati nella corteccia visiva del gatto (a sinistra in figura 21,

in blu), rispondevano intensamente a elementi visuali di forma lineare orientati specificamente, localizzati in posizioni ben definite dell'immagine. Cellule denominate *complesse*, invece, avevano un campo ricettivo più ampio (rispondendo comunque all'orientazione, ma da elementi visuali anche localizzati in regioni vicine), perché ricevevano l'*input* dall'*output* delle cellule semplici (in verde in figura).

Nel 1980, Fukushima applicò le scoperte di Hubel e Wiesel realizzando un modello informatico funzionante del sistema visivo [23], il Neocognitron, precursore delle moderne CNN. Il principio è mostrato in figura 21 a destra. Il primo strato (convoluzionale, in blu) applica un'operazione di convoluzione all'immagine da riconoscere. In particolare, l'applicazione di un piccolo filtro di convoluzione (riquadro grigio) a ogni posizione nell'immagine (finestra mobile) crea una mappa di *feature*. Uno strato (*layer*) convoluzionale ha tante mappe di *feature* quanti sono i filtri applicati (nel caso particolare, solo uno). Conservando il valore massimo di attivazione in una piccola sezione di ciascuna mappa di *feature* (riquadro grigio) di fatto si sottocampiona l'immagine e si costruisce una risposta complessa simile a quella fornita dalle cellule *complesse* del caso biologico (in verde in figura). Quest'operazione è nota come *max-pooling* e le mappe di *feature* sottocampionate così prodotte costituiscono i livelli di *pooling*.

I primi strati delle reti convoluzionali hanno quindi il compito di estrarre dalle immagini le *feature* di basso livello, che poi si combinano tra loro strato per strato. Le *feature* non sono, dunque, ingegnerizzate manualmente come nel caso del ML, ma sono individuate in maniera automatica. Gli ultimi strati delle reti hanno il compito della classificazione vera e propria. Nelle applicazioni reali, le CNN adoperate per il DL sono complesse, come si può vedere in figura 26, nella sezione dedicata agli esempi applicativi. Nel corso degli anni si è passati da reti con un numero relativamente limitato di livelli a strutture che ne impiegano oltre cento e di tipo diverso, con un sensibile incremento delle prestazioni e quindi una diminuzione dell'errore di classificazione.

La applicazioni erano inizialmente limitate al riconoscimento delle immagini naturali (impulso notevolissimo è stato dato dalla realizzazione del

database ImageNet, contenente 14 milioni di immagini) poi esteso ad altri campi, tra cui le immagini diagnostiche mediche. Una rete convoluzionale ha necessità di un grandissimo numero di immagini durante la fase di *training*, il che spesso ne limita l'applicazione a favore del tradizionale ML per situazioni nelle quali la quantità di dati non sia sufficiente. Ciò accade spesso in Medicina, sia per la relativa rarità di una patologia, sia per la difficoltà di raccogliere massicciamente i dati in maniera omogenea e utile (studi multicentrici). Vengono allora in aiuto stratagemmi quali il *transfer learning* che, anziché allenare da zero una rete, parte da un sistema preallentato su immagini di natura diversa (immagini naturali, ad esempio ImageNet) e attua una delle seguenti strategie:

- uso delle *feature* calcolate dalla rete preallentata, e classificazione mediante un classificatore in ML (ad esempio, una rete neurale tradizionale, o *shallow* in contrapposizione alle reti *deep*, oppure una SVM o *Support Vector Machine*)
- *fine tuning*, che consiste nel modificare opportunamente in base al problema gli ultimi strati (che hanno il compito della classificazione), adattando solo i pesi di questi ultimi durante una fase di *training* sul *dataset* specifico.

Ulteriore differenza sostanziale tra DL e ML sono i tempi di calcolo, molto maggiori per il DL qualora si richieda il *training* completo della rete.

Alcuni esempi di applicazioni

La letteratura registra ormai una quantità considerevole di articoli sulle applicazioni del ML o del DL in Medicina: darne un campione rappresentativo sarebbe compito improbo oltretutto improponibile per la quantità di testo necessario. È stata dunque scelta una via alternativa: la breve rassegna che segue riassume quattro articoli rappresentativi della ricerca svolta nel campo dall'autore di questo lavoro.

Preliminarmente è tuttavia utile esaminare la figura 22 che rappresenta l'evoluzione temporale

e la distribuzione geografica delle pubblicazioni internazionali referenziate, censite dal database Scopus, aventi per oggetto i sistemi di ML o DL, o riferite alla radiomica, o aventi le parole chiave CADx o CADe. La figura 23 riporta le medesime informazioni ma è riferita al solo termine *radiomica*. Si nota innanzitutto che il numero di lavori globalmente dedicati all'argomento è in crescita continua e l'Italia si pone in ottima posizione (addirittura, per la parola chiave *radiomica*, è al terzo posto dopo colossi come USA e Cina). In particolare dal 2012, anno in cui uscì il lavoro seminale sulla radiomica [2], la crescita del numero di lavori che riportano questa parola chiave è stata praticamente esponenziale (tranne negli ultimi due anni in cui un assestamento ha portato la curva vicina alla linearità). Ciò dimostra, insieme ai numeri assoluti di tutto rispetto, l'importanza che le tematiche hanno nel panorama delle letteratura scientifica odierna.

Una patologia prenatale rara: l'ernia congenita diaframmatica (CDH)

Il primo esempio [25] è riportato con la finalità di porre l'accento sull'importanza di definire un protocollo chiaro e dettagliato prima di realizzare uno studio di ML/DL in Medicina. Troppo spesso, infatti, il lavoro è organizzato a partire da dati disomogenei (come spesso capita nel caso di studi retrospettivi) e senza un'adeguata coscienza di quali siano i quesiti clinici per i quali si cerca risposta. L'articolo riporta quindi il protocollo stabilito a monte per la realizzazione di un sistema di IA a scopo prognostico per l'ernia diaframmatica congenita (CDH). Si tratta di una patologia rara del feto, consistente nel passaggio parziale dei visceri addominali nel torace, attraverso un difetto del diaframma (erniazione), con conseguente compressione e iposviluppo degli organi toracici (in particolare i polmoni), insufficienza respiratoria e ipertensione polmonare persistente, con elevata mortalità alla nascita quando il bambino dovrà utilizzare i propri polmoni per la respirazione.

Le previsioni sull'esito dei pazienti con ernia diaframmatica congenita (CDH) sono ancora limitate nella stima prenatale dell'ipertensione polmonare postnatale (PH). L'articolo propone di applicare approcci di Machine Learning (ML)

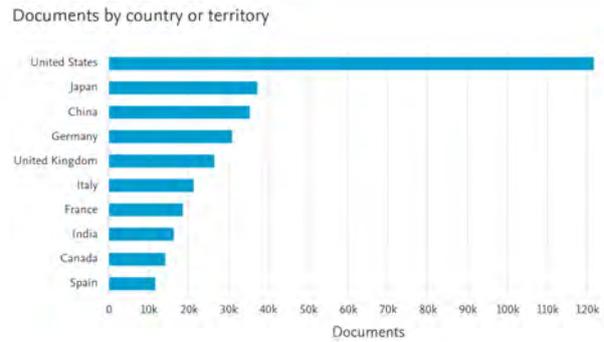
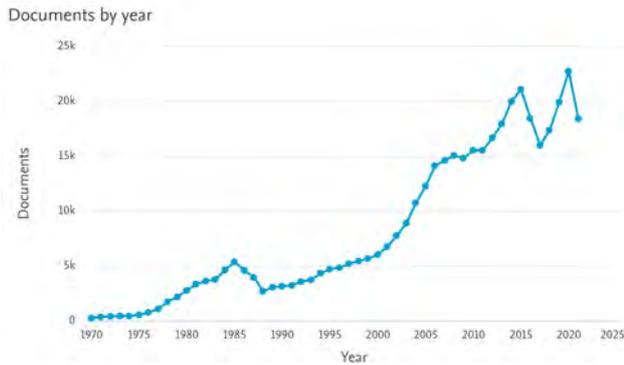


Figura 22: Evoluzione nel tempo (sinistra) e distribuzione geografica (destra) del numero di pubblicazioni aventi, come parole chiave o all'interno del titolo o dell'abstract, i termini *Machine/Deep Learning*, *radiomics*, o *CADx/CADe* (fonte: Scopus).

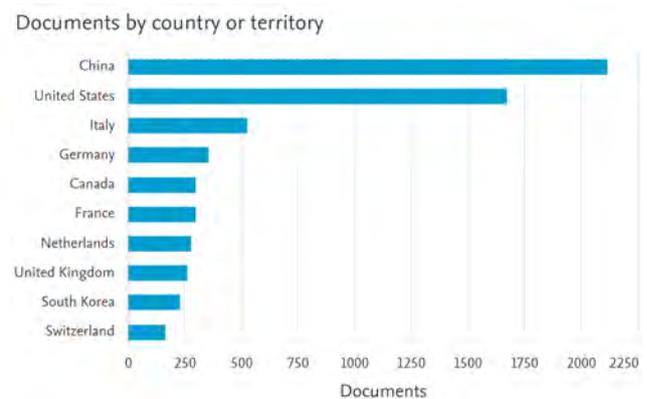
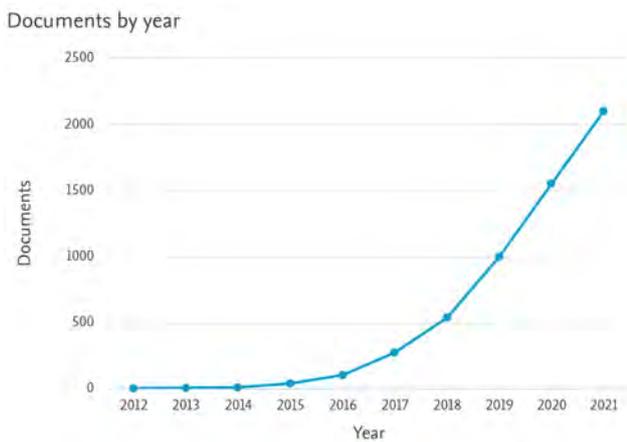


Figura 23: Evoluzione nel tempo (sinistra) e distribuzione geografica (destra) del numero di pubblicazioni aventi, come parole chiave o all'interno del titolo o dell'abstract, il termine *radiomics* (fonte: Scopus).

e Deep Learning (DL) a feti e neonati con CDH per sviluppare modelli di previsione in epoca prenatale, basati sull'analisi integrata di dati clinici, per fornire risultati prognostici quali il PH neonatale *in primis* e, possibilmente:

- risposta favorevole alla *Fetoscopic Endoluminal Tracheal Occlusion*, o FETO: si tratta di un intervento chirurgico avente la finalità di creare un'occlusione nella trachea del feto, per impedire ai fluidi prodotti dai polmoni di sfuggire attraverso di essa; è stato infatti dimostrato che il fluido, se trattenuto nelle vie aeree, aumenta la pressione nella trachea che stimola a sua volta lo sviluppo dei polmoni, altrimenti destinati a restare piccoli e insufficientemente sviluppati;
- necessità di *ExtraCorporeal Membrane Oxygenation*, o ECMO: è un trattamento post-natale che, mediante una macchina che assume le funzioni dei polmoni e del cuore,

stabilizza il neonato;

- sopravvivenza all'ECMO;
- morte post-natale.

Inoltre, il protocollo prevede di produrre un sistema di segmentazione polmonare (semi)automatica del feto in risonanza magnetica (MRI), utile durante l'implementazione del progetto (per il calcolo di *feature* radiomiche e per limitare l'ambito di applicazione degli algoritmi di ML e DL) ma importante anche per standardizzare le misure di volume polmonare dei feti. Il sistema sarà sviluppato sulla base dell'approccio DL denominato 3D U-NET. Essendo il fegato l'organo principalmente coinvolto nella erniazione, è prevista eventualmente anche la realizzazione di un sistema di segmentazione automatica del fegato.

Il progetto prevede di arruolare pazienti con CDH isolato da gravidanze singole, i cui controlli prenatali siano stati effettuati presso l'Unità

di Chirurgia Fetale della Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico (Milano) dalla 30ma settimana di gestazione. La ricerca sarà in retrospettiva, con la raccolta di variabili cliniche e radiologiche (MRI) di madri e neonati nati su un arco di 9 anni. I dati saranno integrati e analizzati con algoritmi previsionali in ML e DL. Essendo la patologia rara e quindi il *dataset* relativamente limitato, si prevede di adottare metodi di *data augmentation* e di riduzione della dimensionalità (selezione ed estrazione delle *feature*) allo scopo rispettivamente di aumentare la dimensione del campione ed evitare *overfitting*.

Tra le variabili considerate come *feature*, citiamo alcuni parametri radiologici (volumi polmonari, volume epatico, angolo di spostamento mediastinico o MSA e coefficiente medio di diffusione apparente, ADC, dei polmoni e del fegato del feto; altre grandezze derivano da queste, come per esempio la percentuale di fegato erniato %LH) e variabili cliniche, come quelle derivanti dalle ecografie prenatali (parametri Doppler relativi alla circolazione) e dall'ecocardiogramma.

Le figure 24 e 25, tratte dall'articolo, mostrano rispettivamente l'aspetto della segmentazione polmonare (manuale) nelle immagini di Risonanza Magnetica T_2 del feto con evidente iposviluppo polmonare, e il fegato, erniato attraverso il diaframma verso il torace. La figura 26 rappresenta lo schema di principio della rete neurale convoluzionale che verrà implementata.² In particolare si nota come nella CNN gli strati di uscita siano adattati al caso della classificazione binaria, ossia con due sole classi.

Discriminazione tra tipi diversi di tumore alla mammella

Il lavoro descritto in [26] si inserisce nel filone dell'IA come supporto alla diagnosi e riguarda il tumore della mammella. Il cancro al seno è in tutto il mondo la principale causa di morte per cancro nelle donne. Questo tumore aggressivo può essere classificato in due gruppi principali,

²Figura modificata da <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>, sito visitato in novembre 2021

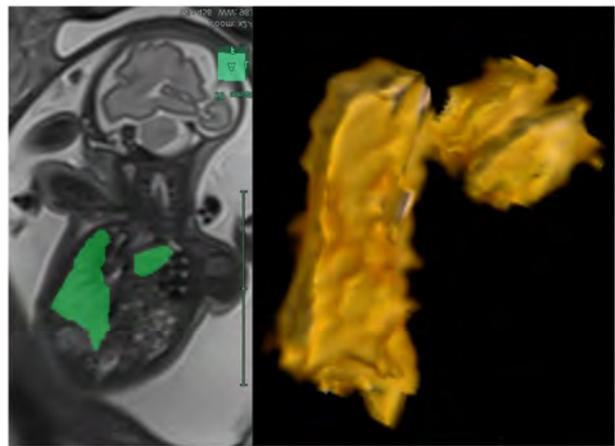


Figura 24: Segmentazione dei polmoni in immagini MRI fetali e ricostruzione 3D, da [25].

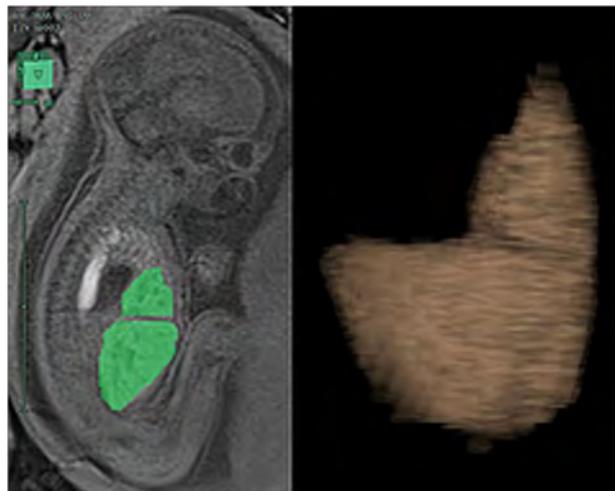


Figura 25: Segmentazione del fegato in immagini MRI fetali e ricostruzione 3D, da [25].

in situ o infiltrante, a loro volta suddivisi in sottotipi. Il tumore infiltrante è tra le lesioni maligne più comuni (il carcinoma duttale invasivo costituisce il 70% di tutti i casi maligni). L'imaging per risonanza magnetica ha dimostrato grande sensibilità nel rilevamento e nella discriminazione tra lesioni benigne e maligne, se interpretato da radiologi esperti. Conoscere precocemente il tipo di tumore è fondamentale per orientare la terapia e anche allo scopo di dare alla paziente l'opportuno supporto psicologico nei casi in cui la gravità del tumore lo suggerisca. Lo scopo di questo studio è stato lo sviluppo di un sistema software in grado di differenziare automaticamente tra tumori *in situ* e infiltranti in immagini di risonanza magnetica con mezzo di contrasto (DCE-MRI), sulla base della firma radiomica della lesione. L'importanza di un sistema siffatto

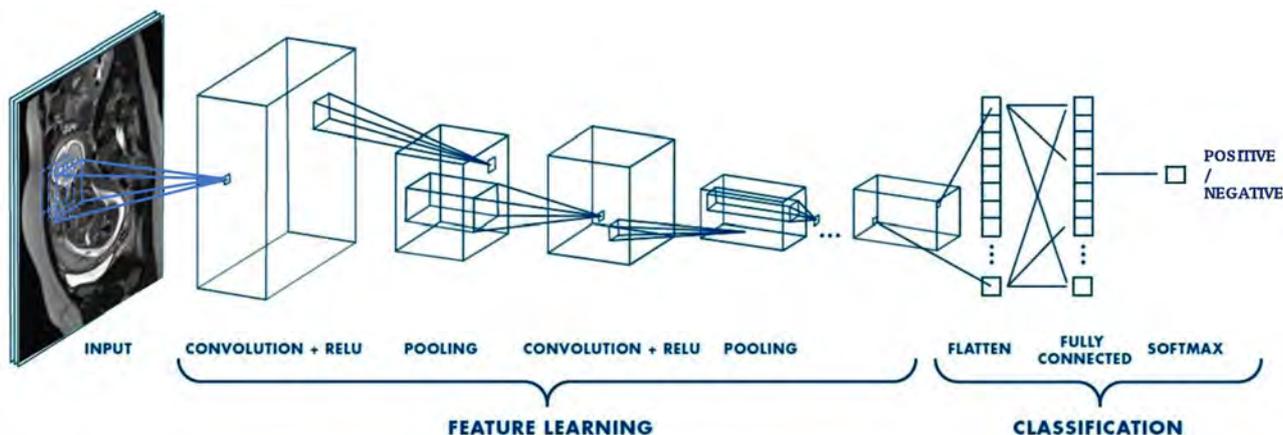


Figura 26: Schema della rete convoluzionale, da [25]. Gli strati di uscita sono adattati al caso della classificazione binaria, ossia con due sole classi, mentre le reti adoperate per il riconoscimento di oggetti, naturali o artificiali, hanno usualmente circa un migliaio di classi d'uscita.

consiste nella possibilità di riconoscere il tipo di tumore senza dover effettuare prelievi biotici e quindi in maniera totalmente non invasiva e precoce.

L'algoritmo è costituito da due livelli di elaborazione principali: (1) localizzazione di possibili regioni di interesse tumorale (ROI) attraverso una procedura iterativa basata su valori di intensità dei livelli di grigio (il cosiddetto ROI Hunter), seguita da un metodo di estrazione e classificazione di *feature* avente lo scopo di scartare i falsi positivi, e (2) caratterizzazione delle ROI selezionate e discriminazione tra tumore *in situ* e invasivo, consistente nell'estrazione di *feature* radiomiche e loro classificazione attraverso un algoritmo di ML.

Il sistema CAD è stato sviluppato e valutato utilizzando un *database* di dimensioni limitate (55 pazienti) di immagini DCE-MRI provenienti dal Presidio Ospedaliero "Di Summa - Perrino" di Brindisi e contenenti almeno una massa confermata per immagine, come diagnosticato da un radiologo esperto. Le immagini avevano *voxel size* diverso l'una dall'altra, per cui è stato necessario inserire un modulo di *preprocessing* per rendere i dati omogenei.

Il primo *step* di elaborazione, ossia l'individuazione e la segmentazione delle masse tumorali, segue lo schema usuale dell'identificazione di *candidati tumore* con la massima sensibilità possibile, seguita dalla *rejection* dei falsi positivi per aumentare la specificità. Nel caso particolare, le

feature da associare ai candidati sono state ricavate da una CNN preallentata, secondo la procedura seguente. I candidati sono stati esplorati con finestre 30×30 *pixel*, ridimensionate (per compatibilità con la CNN) a 224×224 *pixel* utilizzando interpolazione bilineare, convertite in immagini RGB attraverso la replica del bitplane dell'immagine, e infine date come *input* a una rete GoogLeNet preallentata sul *dataset* ImageNet. La rete produce 1000 *feature* che sono state ridotte mediante *recursive feature selection* a 200 e utilizzate per l'allenamento di una rete neurale artificiale multistrato *feed-forward* allenata con *backpropagation*. La procedura di *training/validation* è stata eseguita con lo schema LOPO.

Una volta individuati i tumori, è stato realizzato il secondo *step* dell'elaborazione, ossia la discriminazione tra i due tipi di lesione: infiltrante e *in situ*. Le regioni di interesse sono state modellizzate da 1820 *feature* tessiturali calcolate da un *software* scritto in linguaggio *python* e basato sulla già citata libreria *pyradiomics*. Le *feature* comprendevano variabili basate sulla forma e *feature* tessiturali del primo ordine e di ordine superiore (da matrici GLCM, GLRLM, GLDM, GLSZM, per le quali si rimanda alla documentazione di *pyradiomics*). Dette variabili sono state calcolate direttamente e previa applicazione di diversi filtri di *pre-processing* quali: *Laplacian of Gaussian* (che evidenzia i bordi), *Wavelet*, e altri, compreso il filtro *Local Binary Pattern*. È stata applicata una fase di *feature selection* per eliminare variabili ridondanti e irrilevanti.

La figura 27 riporta, sulla sinistra, lo schema

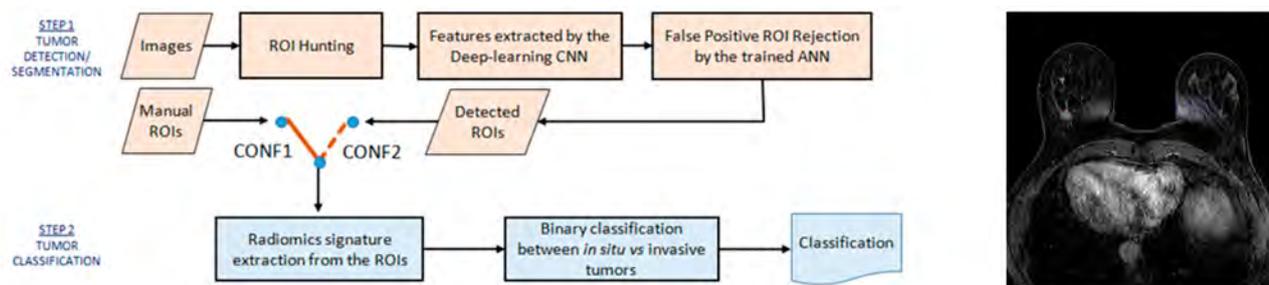


Figura 27: Sinistra: schema del software di rilevamento e classificazione del tumore della mammella; destra: esempio di detection di una massa tumorale e del riconoscimento di un falso positivo, rispettivamente in rosso e blu (da [26]).

del software.

La sensibilità dell'algoritmo al rilevamento delle masse è risultata essere del 75%, quindi migliorabile (vedere, in figura 27, sulla destra, un esempio di detection di una massa tumorale e del riconoscimento di un falso positivo, rispettivamente in rosso e blu). L'AUC della curva ROC per la discriminazione tra tumori *in situ* e infiltrativi è stata pari a 0,70: lontana da valori ottimali ma paragonabile a quanto ottenuto da altri ricercatori.

Individuazione di noduli polmonari

In [27], e in alcuni degli articoli ivi citati, è descritto un sistema CAD per l'individuazione automatica dei noduli polmonari in immagini TC a bassa dose. L'esempio è stato scelto perché, oltre a presentare molti elementi descritti nella prima parte di questo lavoro (segmentazione, calcolo di *feature* per la discriminazione, classificazione), pone anche l'accento sull'interesse dell'uso del *cloud* per rendere i sistemi facili da gestire e mantenere. Infatti, l'approccio più comune nello sviluppo dei software è l'implementazione di stazioni di lavoro *standalone*, dotate di interfacce grafiche utente (GUI) sviluppate autonomamente (con un costo non indifferente dovuto alle licenze per i software di sviluppo e alle necessità di aggiornamento). Un'alternativa è il *cloud computing*, accessibile tramite protocolli Web sicuri, e coniugato con l'approccio SaaS o *Software as a Service*.

Come motivazione del lavoro, occorre porre l'attenzione sul fatto che il rilevamento visivo, da parte del medico radiologo, dei noduli pol-

monari nelle immagini TC a bassa dose è particolarmente difficile a causa del tasso di rumore nelle immagini. L'utilizzo di sezioni TC sottili per migliorare la risoluzione produce un numero di *slice* 2D compreso tra 300 e 500, il che rende il lavoro di analisi visiva lungo e soggetto a errori. Inoltre, i noduli possono essere di dimensioni molto ridotte e collegati alla superficie della pleura o addossati a un vaso, il che rende l'identificazione ancora più difficile. La disponibilità di software che assistano durante la diagnosi può quindi essere un utile supporto a vantaggio dell'accuratezza.

Il sistema presentato, realizzato nell'ambito di un progetto finanziato dalla Commissione Scientifica Nazionale 5 dell'INFN (Istituto Nazionale di Fisica Nucleare), è composto da tre blocchi principali: WIDEN (*Web-based Image and Diagnosis Exchange Network*) gestisce il flusso di lavoro, il caricamento dell'immagine e la notifica del risultato CAD; la *batch farm* IaaS (*Infrastructure as a Service*) basata su OpenNebula, che alloca risorse di elaborazione e archiviazione virtuali; il CAD M5L che fornisce la funzionalità di rilevamento dei noduli. L'implementazione proposta gestisce in modo sicuro i dati sensibili dei pazienti, poiché le immagini vengono trasferite con il protocollo HTTPS e la *batch factory* sottostante è isolata. Inoltre è efficiente poiché scala dinamicamente in base alle richieste degli utenti grazie al *backend* fornito dal *cloud*.

Entrando nel dettaglio, WIDEN consente ai medici di scambiare studi di *imaging* e confrontare le diagnosi. Il servizio è disponibile come applicazione Web accessibile via *browser* con credenziali adeguate. I file in formato DICOM³ pos-

³Digital imaging and COmmunication in Medicine, standard

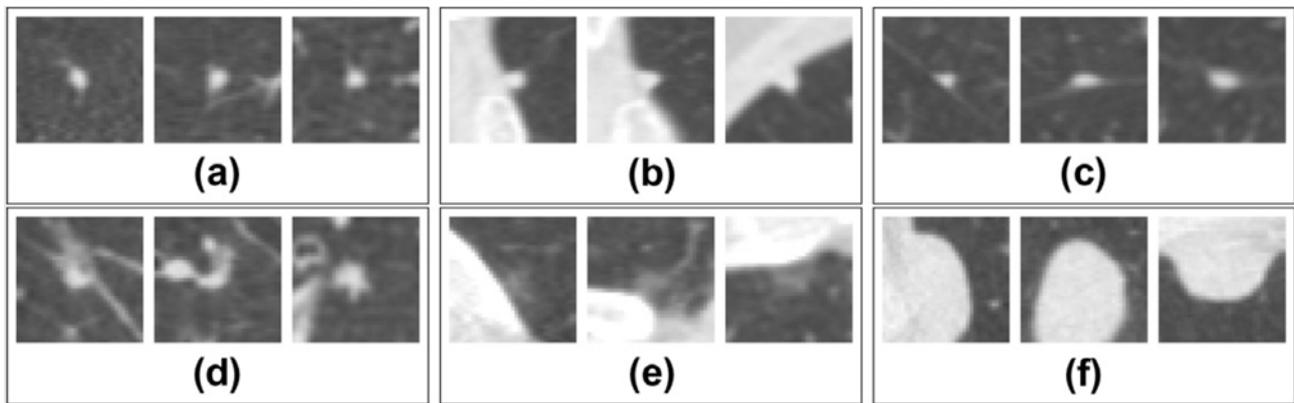


Figura 28: Esempi di noduli polmonari. In ogni riquadro è visualizzato un nodulo in vista sagittale, coronale e assiale, 35 voxel (circa 25 mm) attorno al punto centrale. La riga superiore mostra tre piccoli noduli, (a) un nodulo isolato di 4,4 mm; (b) un nodulo pleurico di 4,2 mm e (c) un nodulo peri-fissurale di 4,8 mm (la linea sottile visibile su ciascuna vista è la fessura). La seconda fila mostra tre grandi noduli, (d) un nodulo di 5,9 mm con adesione vascolare; (e) un nodulo ground-glass di 5,4 mm (relativamente raro) e (f) un grande nodulo pleurico (18,4 mm). (da [28], immagine parziale).

sono essere caricati nell'archivio remoto: i medici possono caricare la propria diagnosi e vedere quelle di altri medici. WIDEN dispone di un sistema di notifica basato su *e-mail* e SMS che informano i medici iscritti quando sono aggiunte o aggiornate informazioni al caso di studio.

Le immagini caricate sono elaborate dagli algoritmi di calcolo, a loro volta gestiti come *plugin* dal sistema di *cloud computing* OpenNebula che consente il cosiddetto *elastic computing*, ovvero l'ottimizzazione delle risorse a seconda del carico computazionale.

Il progetto MAGIC-5 dell'INFN ha dimostrato che la combinazione di analisi ottenute da algoritmi diversi e complementari può massimizzare il risultato: il sistema implementato combina quindi tre algoritmi con diversi approcci: il Channeler Ant Model (CAM), l'analisi neurale basata su voxel (Voxel-Based Neural Approach, VBNA) e il Region Growing Volume Plateau o RGVP (per i dettagli, si rimanda naturalmente agli articoli specifici). Gli algoritmi hanno uno stadio comune, la segmentazione 3D del volume parenchimale, che separa trachea, bronchi e polmoni. Il risultato della segmentazione è salvato come maschera che, sovrapposta al parenchima polmonare, limita spazialmente la ricerca dei candidati noduli.

I tre algoritmi forniscono elenchi di candida-

per la comunicazione, la visualizzazione, l'archiviazione e la stampa di informazioni di tipo biomedico come le immagini diagnostiche.

ti noduli con un certo tasso di falsi positivi, da ridurre mediante un classificatore (una rete neurale multistrato *feed-forward*). Le *feature* calcolate per i candidati sono piuttosto semplici e comprendono variabili basate sia sulla forma della ROI, sia sulle intensità dei grigi al suo interno, sia sui livelli di grigio dei *pixel* vicini; tra queste: volume della ROI, intensità massima, intensità media, deviazione standard dell'intensità, sfericità, intensità dei vicini, autovalori della matrice Hessiana, matrice dei gradienti. In particolare, gli autovalori della matrice Hessiana hanno la proprietà di descrivere bene la geometria locale e quindi di essere legati alla forma degli oggetti visibili nelle immagini.

In figura 29 sono riportate le curve FROC (*Free-Response ROC Curve*) relative ai risultati dei singoli sistemi di *detection*, e della loro combinazione. La FROC[30] è un analogo della ROC in cui l'ascissa riporta il numero di FP per immagine (ovvero per scansione TC) mentre l'ordinata misura la sensibilità. Al contrario della ROC, la FROC non è allocata in uno spazio finito. Essa è talora più utile perché serve a identificare la soglia di uscita del classificatore in base al compromesso scelto tra il contenimento del numero di FP per paziente, e la sensibilità.

Come si vede dal grafico, la combinazione dei sistemi di *detection* risulta migliore dei singoli algoritmi.

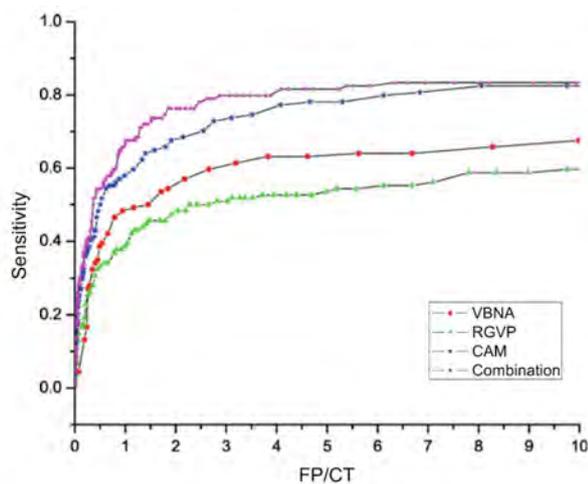


Figura 29: Le curve FROC ottenute con la combinazione di CAD confrontate con quelle ottenute con i singoli sistemi nel range di FP di interesse radiologico (0–10 FP/CT): in rosso sono riportate le curve VBNA CAD, RGVP CAD e CAM CAD, verde e blu, rispettivamente; la combinazione è mostrata in viola. Da [29].

Segmentazione di gliomi cerebrali

I gliomi sono i tumori cerebrali primari più comuni. Una diagnosi precoce e una valutazione completa dell'estensione del tumore e della relazione con le strutture anatomiche circostanti sono cruciali nel determinare la prognosi e la pianificazione del trattamento. La crescita diffusa e infiltrativa dei gliomi cerebrali è un importante determinante di prognosi infausta: le cellule tumorali invadono i tessuti circostanti preferenzialmente lungo i tratti di sostanza bianca, diffondendosi oltre l'area anormale individuabile nelle immagini di risonanza magnetica convenzionale. Pertanto, quest'ultima non sempre consente una precisa delineazione dei margini del tumore o la differenziazione del tumore dall'edema o dagli effetti della terapia. Tuttavia l'individuazione e la caratterizzazione delle infiltrazioni microscopiche in maniera non invasiva è di fondamentale importanza per il trattamento chirurgico e la pianificazione della radioterapia o per valutare la risposta alla chemioterapia. L'imaging in tensore di diffusione (*Diffusion Tensor imaging*, o DTI, o DT-MRI), al contrario della MRI convenzionale, può identificare anomalie peritumorali della sostanza bianca nei gliomi cerebrali rilevando la presenza di piccole aree con infiltrazione di cellule tumorali attorno al bordo del tumore.

Il DTI può dare informazioni sulle microstrut-

ture cerebrali poiché fornisce contrasto all'immagine in base alle differenze nell'ampiezza della diffusione delle molecole d'acqua all'interno del cervello nelle varie direzioni. Il DTI stima la diffusione delle molecole d'acqua in ciascun voxel e produce quindi mappe scalari,[31] tra cui anisotropia frazionaria (FA), la diffusività media (MD), la diffusività assiale (AD), la diffusività radiale (RD), e le mappe p e q . Questi parametri descrivono la microstruttura e l'integrità della materia bianca e grigia nel cervello in un modo molto specifico che consente di monitorare la progressione di una malattia o di studiare lo sviluppo del cervello. In particolare, è stato dimostrato che il DTI è un *marker* sensibile di danno alla sostanza bianca.

Obiettivo del lavoro descritto in [32] è stato caratterizzare il tessuto patologico e sano in DTI mediante analisi tessiturale statistica in 3D (*feature* di Haralick e altre variabili), con lo scopo di sviluppare un CAD per l'individuazione e il contornamento automatico (segmentazione) dei tumori cerebrali. L'utilità del sistema è ad esempio la possibilità di misurare automaticamente (e perciò senza incorrere nell'altrimenti inevitabile variabilità inter- e intra-operatore) un tumore durante la terapia.

Sono stati selezionati quindici pazienti con glioma (9 di basso grado, 6 di alto grado) e sei pazienti sani di controllo dell'Ospedale Vita-Salute San Raffaele (Milano). Sono state acquisite immagini in DTI a 3T. Sono state calcolate svariate mappe di diffusione utilizzando un *software* implementato in Matlab: sia mappe di isotropia, come MD e la mappa p , sia di anisotropia, come FA e la mappa q . In seguito i tumori sono stati segmentati manualmente da radiologi esperti in tutte le mappe calcolate. *Feature* tessiturali in 3D sono state calcolate nelle regioni di interesse tumorali segmentate e (al fine di identificare le variabili discriminanti) nelle regioni controlaterali, in particolare le *feature* del primo ordine (dall'istogramma dell'intensità dei livelli di grigio) e del gradiente dei grigi, delle matrici di co-occorrenza (di Haralick) e delle matrici di percorrenza (RLM, *Run Length Matrix*)[33]. Il calcolo è stato effettuato con il *software* MaZda.[34] Si è utilizzato un approccio a finestra mobile (le ROI ipsilaterali e controlaterali sono state suddivise in piccoli sotto-volumi sovrapposti in cui è stato fat-

to il calcolo, associando poi il vettore di variabili al *pixel* centrale, da classificare). La scelta delle variabili più discriminanti per ogni mappa è stata poi effettuata in base al coefficiente di Fisher, già citato nella prima parte di quest'articolo.

Per eliminare la ridondanza delle informazioni, la dimensionalità dello spazio delle *feature* è stata ulteriormente ridotta utilizzando la PCA e conservando l'insieme minimo di componenti principali alle quali competesse almeno il 97% della varianza (due o tre componenti principali). Questa procedura non è in generale necessariamente fruttuosa perché la PCA non tiene conto del potere discriminante delle variabili ma, in questo caso, si è rivelata utile.

Il classificatore, allenato mediante una procedura di *hold out*, è stato una rete *feed-forward* con *back-propagation*, implementata in MatLab tramite PRTools, una libreria di *pattern recognition* liberamente disponibile. L'architettura di rete consisteva in un livello nascosto con tre neuroni, e un neurone di uscita. La qualità della classificazione dei *pixel* dell'insieme di validazione, misurata mediante l'area sotto la curva ROC, è stata > 0.96 per tutte le mappe considerate. In figura 30 è visibile il confronto tra una segmentazione automatica e la corrispondente manuale. La figura 1 già vista nella prima parte dell'articolo mostra altre segmentazioni ricavate mediante il sistema descritto (non riportate in [32]).

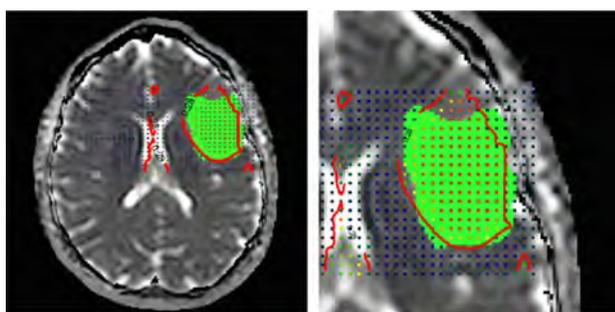


Figura 30: Segmentazione automatica e manuale di una lesione: i punti segnano le posizioni dei centri della finestra mobile durante l'esplorazione delle immagini, la linea rossa mostra la segmentazione prodotta dal sistema CAD e la regione verde è la corrispondente ROI disegnata manualmente.

Conclusioni

Le applicazioni dell'Intelligenza Artificiale alla Medicina sono ormai tantissime, alcune già in uso nella pratica clinica presso presidi ospedalieri d'avanguardia. Tutto fa presagire che il progredire delle tecniche di *Machine/Deep Learning* e la disponibilità di adeguata potenza di calcolo rendano inevitabile la transizione verso una Medicina sempre più supportata da sistemi automatici con finalità diagnostiche e previsionali. Una diagnosi precoce, non invasiva e di qualità, è di beneficio per il paziente ma anche per il Sistema Sanitario Nazionale, per la potenziale riduzione delle spese non necessarie (si pensi ad esempio al risparmio derivante dalla diminuzione del numero di biopsie e dalla mancata prescrizione di esami clinici superflui). Le applicazioni prognostiche dell'IA sono anch'esse estremamente importanti, perché permettono ai medici di orientare meglio le terapie e di essere preparati alla probabile evoluzione di una patologia, garantendo una risposta più pronta quando necessario.

Si tratta della Medicina del futuro, sempre più personalizzata, sempre più di precisione, in cui figure diverse e una volta scarsamente comunicanti (il medico, il fisico, il matematico, l'ingegnere) stabiliscono un linguaggio comune e collaborano per il bene del paziente.

Appendice 1: le *feature* di Haralick

Supponiamo che la ROI di cui vogliamo caratterizzare la *texture* sia costituita da (o strettamente contenuta in) un rettangolo I di dimensioni $N_r \times N_c$. Facciamo riferimento alla figura 31, in particolare all'immagine sulla sinistra, di dimensioni 5×5 , in cui i *pixel* sono rappresentati da caselle con fondo grigio più o meno scuro e dal valore corrispondente. I livelli di grigio che compongono l'immagine siano quantizzati in un numero pari a N (ovvero, siano contenuti nell'insieme $G = \{0, \dots, N-1\}$; nel caso particolare, $N = 4$). Ove l'immagine originale abbia un numero maggiore di livelli di grigio, è previsto che questi siano riducibili a un valore inferiore

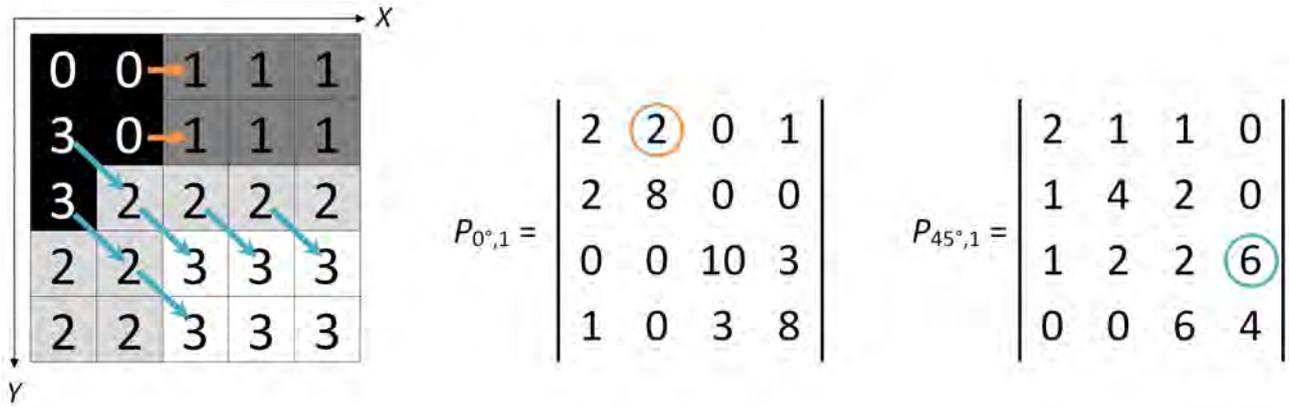


Figura 31: Costruzione delle matrici di co-occorrenza dei livelli di grigio (GLCM).

arbitrario, allo scopo di rendere le matrici meno sparse.

Ogni elemento della GLCM $P_{\alpha,d}$, ovvero $P_{\alpha,d}^{ij}$ per $(i, j) \in \{0, \dots, N-1\} \times \{0, \dots, N-1\}$, è calcolato come il numero di occorrenze di coppie di *pixel* distanti d nella direzione individuata dall'angolo α , aventi livelli di grigio (i, j) o (j, i) (quindi, indipendentemente dall'ordine). Ciò equivale (a meno di un fattore di normalizzazione) a valutare la probabilità congiunta che una coppia di punti, che soddisfano le condizioni di distanza d e fissata una direzione α , abbia livelli di grigio (i, j) o (j, i) .

Le GLCM $P_{\alpha,d}$ avranno quindi dimensione $N \times N$ e dipenderanno da:

- la distanza d (generalmente posta pari a 1) tra i *pixel* delle coppie utilizzate per il calcolo; la metrica usata è la distanza di Chebyshev: $d(p_1, p_2) = \max_k (|p_{1,k} - p_{2,k}|)$, dove p_1 e p_2 sono i *pixel* considerati e $p_{1,k}$ è la k -ma coordinata del *pixel* p_1 ;
- l'orientamento α del segmento che congiunge i due *pixel*, espresso come angolo rispetto a un asse di riferimento e usualmente pari a: $0^\circ, 45^\circ, 90^\circ$ e 135° .

Si noti che la formulazione delle GLCM adoperate per il calcolo delle feature di Haralick porta a matrici simmetriche, perché al computo dell'elemento $P_{\alpha,d}^{ij}$ contribuiscono sia le coppie (i, j) che le coppie (j, i) .

Ad esempio, nella figura 31 si può notare come l'immagine sia caratterizzata da quattro livelli di grigio, indicati come 0, 1, 2, 3. Nella matrice di co-occorrenza $P_{0^\circ,1}$, ossia con angolo 0° e $d = 1$, l'elemento $P_{0^\circ,1}^{01}$ è calcolato dal numero di occorrenze (ossia compresenze) tra *pixel* con valori di

grigio 0 e 1, a distanza 1 in direzione x (angolo pari a 0°). Osservando i valori di grigio dell'immagine, si possono contare due co-occorrenze tra *pixel* 0 e 1 in questa direzione e distanza 1 *pixel* (indicati in colore arancio). Per questo motivo $P_{0^\circ,1}^{01}$ (ma anche $P_{0^\circ,1}^{10}$) è pari a 2.

Stesso ragionamento per $P_{45^\circ,1}^{23}$ (indicato in celeste) e per tutti gli altri elementi delle due matrici. Il set delle GLCM è completato da $P_{90^\circ,1}$ e $P_{135^\circ,1}$. Nel caso in cui l'immagine di *input* sia tridimensionale, è sufficiente dichiarare un opportuno set di angoli rispetto agli assi, che corrispondano alle direzioni nello spazio.

Dalle matrici di co-occorrenza si calcolano le 19 *feature* tessiturali di Haralick, per le quali si rimanda agli articoli originali [8, 9]. Solo a titolo esemplificativo, eccone due. L'energia, o misura del grado di omogeneità della *texture*, o momento angolare di secondo ordine, è definita come:

$$ENE_{\alpha,d} = \sum_i \sum_j P_{\alpha,d}^{ij}.$$

Valori alti di E corrispondono a tessiture omogenee, ovvero in cui la maggior parte delle coppie di *pixel* (fissati α e d) hanno livelli di grigio simili. Valori bassi competono a situazioni meno omogenee.

L'entropia, definita secondo la formula seguente:

$$ENT_{\alpha,d} = - \sum_i \sum_j P_{\alpha,d}^{ij} \cdot \log P_{\alpha,d}^{ij}$$

è alta quando i valori $P_{\alpha,d}^{ij}$ sono equidistribuiti. L'entropia ha valori bassi se, per esempio, la matrice di co-occorrenza è diagonale, ossia esistono

coppie di livelli di grigio uguali dominanti per una certa direzione e distanza.

Lista degli acronimi

%LH	Percent of Liver Herniation
AD	Axial Diffusivity
AUC	Area Under the Curve
AUROC	Area Under the ROC Curve
CAD	Computer-Assisted Detection/Diagnosis
CADe	Computer-Assisted Detection
CADx	Computer-Assisted Diagnosis
CDH	Congenital Diaphragmatic Hernia
CNN	Convolutional Neural Network
COM	CoOccurrence Matrices
CT	Computed Tomography
CV	Cross Validation
DCE-MRI	Dynamic Contrast-Enhanced MRI
DICOM	Digital Imaging and COmmunications in Medicine
DL	Deep Learning
DT-MRI	Diffusion Tensor MRI
DTI	Diffusion Tensor Imaging
ECG	Elettrocardiogramma
ECMO	ExtraCorporeal Membrane Oxygenation
EEG	Elettroencefalogramma
FA	Fractional Anisotropy
FETO	Fetoscopic Endoluminal Tracheal Occlusion
FLDA	Fisher Linear Discriminant Analysis
FN	False Negative
FP	False Positive
FROC	Free-Response Operating Characteristic
GLCM	Gray-Level Cooccurrence Matrix
GLDM	Gray Level Dependence Matrix
GLRLM	Gray-Level Run-Length Matrix
GLSZM	Gray Level Size Zone Matrix
GUI	Graphical User Interface
IA	Intelligenza Artificiale
IBSI	Image Biomarker Standardisation Initiative
ICA	Independent Component Analysis
LDA	Linear Discriminant Analysis
LOO	Leave One Out
LOO-CV	Leave One Out Cross Validation
LOPO	Leave One Patient Out Cross Validation
LOSO	Leave One Subject Out Cross Validation
MD	Mean Diffusivity
ML	Machine Learning
MRI	Magnetic Resonance Imaging
PCA	Principal Component Analysis
PET	Positron-Emission Tomography
PH	Pulmonary Hypertension

P-R	Precision-Recall
PR	Pattern Recognition
RD	Radial Diffusivity
ROC	Receiver Operating Characteristic
ROI	Region of Interest
SMOTE	Synthetic Minority Oversampling TEchnique
SVM	Support Vector Machine
TAC	Tomografia Assiale Computerizzata
TC	Tomografia Computerizzata
TN	True Negative
TP	True Positive
WIDEN	Web-based Image and Diagnosis Exchange Network



- [1] G. Buttazzo, *Reti Neurali in grado di apprendere*, Ithaca, XVI (2020) 109.
- [2] P. Lambin et al., *Radiomics: Extracting more information from medical images using advanced feature analysis*, European Journal of Cancer, 48 (2012) 441.
- [3] R. O. Duda, P. E. Hart, D. G. Stork, *Pattern Classification, 2nd Edition*, Wiley, New York, (2000).
- [4] H. Liu, H. Motoda (Editors), *Feature Extraction, Construction and Selection - A Data Mining Perspective*, Springer, Berlino, (1998).
- [5] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Prentice Hall, Englewood Cliffs NJ, (2007).
- [6] A. Zwanenburg et al., *The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping*, Radiology 295 (2020) 328.
- [7] J. J. Griethuysen et al., *Computational Radiomics System to Decode the Radiographic Phenotype*, Cancer Research, 77 (2017) e104.
- [8] R. M. Haralick, K. Shanmugan, I. Dinstein, *Textural Features for Image Classification*, IEEE Transactions on Systems, Man, and Cybernetics, SMC-3 (1973) 610.
- [9] R. M. Haralick, *Statistical and structural approaches to texture*, Proc. IEEE, 67 (1979) 786.
- [10] T. Ojala, M. Pietikäinen, D. Harwood, *Performance evaluation of texture measures with classification based on Kullback discrimination of distributions*, Proc. of the 12th IAPR International Conference on pattern Recognition (ICPR 1994), (1994) 582.
- [11] M. Schmida, M. Vetterlia, K. Wegener, *Polymer powders for laser-sintering: Powder production and performance qualification*, AIP Conference Proceedings, 2065, 020008 (2019).
- [12] R. E. Bellman, *Dynamic Programming*, Princeton University Press, Princeton NJ (1957).
- [13] T. Hastie, R. Tibshirani, G. James, D. Witten, *An Introduction to Statistical Learning, with applications in R.*, Springer, Berlin (2013).

- [14] M. Stone *Cross-Validatory Choice and Assessment of Statistical Predictions*, J. R. Stat. Soc. Ser. B, 36 (1974) 111.
- [15] A. Bruce, K. Hopkin, A. Johnson, *L'Essenziale di Biologia Molecolare della cellula*, Zanichelli, Bologna, (2003).
- [16] C. Bishop, *Neural networks and their applications*, Review of Scientific Instruments, 65 (1994) 1803.
- [17] F. Rosenblatt, *The Perceptron—a perceiving and recognizing automaton*, Cornell Aeronautical Laboratory Report 85-460-1 (1957).
- [18] W.S. McCulloch, W. Pitts, *A logical calculus of the ideas immanent in nervous activity*, The bulletin of mathematical biophysics, 5 (1943) 115.
- [19] D. MacKay, *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press, Cambridge, (2005).
- [20] R. Lippmann, *An introduction to computing with neural nets*, IEEE ASSP Magazine, 4 (1987) 4.
- [21] G. W. Lindsay, *Convolutional Neural Networks as a Model of the Visual System: Past, Present, and Future*, J. Cogn. Neurosci., 33 (2021) 2017.
- [22] D. H. Hubel, T. N. Wiesel, *Receptive fields, binocular interaction and functional architecture in the cat's visual cortex*, The Journal of physiology, 160 (1962) 106.
- [23] K. Fukushima *Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position*, Biological cybernetics, 36 (1980) 193.
- [24] J. A. Swets, *Measuring the accuracy of diagnostic systems*, Science, 240 (1998) 1285.
- [25] I. Amodeo et al., *A Machine and deep Learning Approach to predict pulmonary hypertension in newborns with congenital diaphragmatic Hernia (CLANNISH): Protocol for a retrospective study*, PLoS ONE 16(11): e0259724 (2021).
- [26] L. Conte et al., *Breast Cancer Mass Detection in DCE-MRI Using Deep-Learning Features Followed by Discrimination of Infiltrative vs. In Situ Carcinoma through a Machine-Learning Approach*, Appl. Sci., 10 (2020) 6109.
- [27] D. Berzano et al., *On-demand lung CT analysis with the M5L-CAD via the WIDEN front-end web interface and an OpenNebula-based cloud back-end*, Proc. IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC) (2012) 978.
- [28] B. van Ginneken et al., *Comparing and combining algorithms for computer-aided detection of pulmonary nodules in computed tomography scans: The ANODE09 study*, Medical Image Analysis, 14 (2010) 707.
- [29] N. Camarlinghi et al., *Combination of computer-aided detection algorithms for automatic lung nodule identification*, Int J Comput Assist Radiol Surg., 7 (2012) 455.
- [30] A. I. Bandos et al., *Area under the free-response ROC curve (FROC) and a related summary index*, Biometrics, 65 (2009) 247.
- [31] A. Peña et al., *Enhanced visualization and quantification of magnetic resonance diffusion tensor imaging using the p:q tensor decomposition*, The British journal of radiology 79 (2006) 101.
- [32] G. De Nunzio et al., *A CAD system for cerebral glioma based on texture features in DT-MR images*, Nucl. Instrum. Methods Phys. Res. A, 648 Suppl. 1 (2011) S100.
- [33] G. Castellano, L. Bonilha, L. M. Li, F. Cendes, *Texture analysis of medical images*, Clin. Radiol., 59 (2004) 1061.
- [34] M. Strzelecki, et al., *A software tool for automatic classification and segmentation of 2D/3D medical images*, Nucl. Instrum. Methods Phys. Res. A, 702 (2013) 137.

Giorgio De Nunzio: laureato in Fisica presso l'Università di Lecce, ha conseguito il dottorato all'Université de Montpellier II. Professore Aggregato in Fisica Applicata (FIS/07) del Dipartimento di Matematica e Fisica "Ennio De Giorgi" dell'Università del Salento, si occupa principalmente di Fisica e Informatica per la Medicina (in particolare, elaborazione di immagini e segnali di interesse biomedicale, e intelligenza artificiale per la realizzazione di sistemi automatizzati di supporto decisionale). Interessi secondari sono la Fisica e l'Informatica applicate ai Beni Culturali.

Il valore dell'informazione

Luca Peliti

Santa Marinella Research Institute Santa Marinella (Rm), Italia

Un ben noto risultato di Kelly mostra che la mutua informazione tra un processo e un segnale relativo ad esso misura il vantaggio che può essere ottenuto nell'ambito del gioco d'azzardo (per esempio, scommesse su corse di cavalli). Questa misura del valore dell'informazione può essere ritrovata in altre forme, come in termodinamica (estrazione di lavoro da un serbatoio di calore) oppure nella dinamica delle popolazioni (adattamento fenotipico all'ambiente). In questo articolo, percorro a grandi linee il ragionamento che porta a queste conclusioni.

La misura dell'incertezza

La teoria matematica dell'informazione nasce con il lavoro di Claude Shannon "A Mathematical Theory of Communication" [1, 2], che considera un modello ideale di trasmissione di messaggi tramite un canale di comunicazione soggetto a rumore. Il problema fondamentale della comunicazione è quello di ridurre l'incertezza che il recipiente del messaggio ha rispetto al contenuto del messaggio stesso. Prima dell'emissione del messaggio, il recipiente ha una certa informazione sui possibili messaggi che possono essere trasmessi. Avendo ricevuto il messaggio, anche se è parzialmente corrotto dal rumore, la sua incertez-

za è diminuita. Al limite, se è possibile ricostruire il messaggio senza errori, la sua incertezza è pari a zero. Questo suggerisce di quantificare il contenuto d'informazione del messaggio come la differenza fra l'incertezza a priori e l'incertezza a posteriori. Dobbiamo considerare il messaggio X come una variabile aleatoria, che ha una certa probabilità a priori p_x di assumere il valore x . Allora Shannon dimostra che l'unica misura dell'incertezza su X che soddisfa degli assiomi semplici ed intuitivi è data dall'espressione

$$H(X) = -k \sum_{x=1}^N p_x \ln p_x, \quad (1)$$

dove k è una costante positiva, per altro arbitraria. Shannon scelse per k il valore $k = 1/\ln 2$. In questo modo, l'incertezza unitaria corrisponde a una semplice alternativa tra due possibili valori con uguale probabilità $p_x = \frac{1}{2}$. Seguendo un suggerimento di John W. Tukey, Shannon chiamò **bit** questa unità, come abbreviazione di *binary digit* (cifra binaria), ma corrispondente anche all'espressione inglese che traduce "un po'". In questo articolo useremo $k = 1$, a meno di avviso contrario.

L'espressione in eq. (1) è nota come **entropia di Shannon**.¹ Shannon (citato in [3]) rivelò che il nome **entropia** gli era stato suggerito da John von Neumann:

¹La dimostrazione dell'espressione (1) della misura dell'incertezza è presentata in appendice.

Dovresti chiamarla **entropia** per due ragioni: la prima perché la tua espressione è identica a quella usata da Gibbs per l'entropia termodinamica, e la seconda e più importante, perché nessuno sa sul serio che cosa sia l'entropia e quindi avresti sempre un vantaggio nelle discussioni.

In effetti, Gibbs nel suo trattato [4] aveva mostrato che l'entropia termodinamica di un sistema all'equilibrio, in cui la probabilità di trovarsi nello stato microscopico x è pari a p_x , è data dall'espressione

$$S = -k_B \sum_x p_x \ln p_x, \quad (2)$$

che corrisponde alla (2) con k uguale alla costante di Boltzmann k_B .

Shannon fece un brillante uso della sua misura dell'incertezza per impostare e risolvere il problema della codifica ottimale di messaggi lungo un canale di comunicazione soggetto a rumore. In particolare dimostrò che esisteva un valore limite quantificabile alla capacità di trasmissione di un canale. Trasmettendo messaggi a una cadenza inferiore a questa capacità è possibile codificarli in modo da recuperarli con un tasso d'errore arbitrariamente piccolo, mentre questo non è possibile se la cadenza supera questo limite.

Tuttavia la misura dell'incertezza ha trovato applicazione in campi molto diversi. In particolare vorrei descrivere in questo articolo alcune applicazioni, partendo dal classico esempio del gioco d'azzardo, alla termodinamica, e perfino a problemi di interesse biologico, come la percezione, la diversificazione fenotipica, e la possibilità di interpretare l'evoluzione darwiniana come un processo di acquisizione dell'informazione.

Corse di cavalli

Nel 1956, Kelly propose un'interpretazione della capacità di trasmissione che andava al di là del problema della codifica ottimale [5]. Il problema non è tanto quello di trasmettere dei messaggi senza errori, quanto quello di assegnare un valore a una comunicazione ricevuta con errore. L'idea fondamentale è che il destinatario dovrebbe poter approfittare della conoscenza (anche

se incerta) del messaggio trasmesso, e che quindi una misura del valore dell'informazione può essere basata sul vantaggio ottenuto.

La situazione considerata da Kelly ricorda il classico film del 1973 "La Stangata", con Paul Newman e Robert Redford [6]. Consideriamo un giocatore che può scommettere su una corsa di cavalli che si svolge in un'altra città. Un complice è in grado di comunicare al giocatore il risultato della corsa prima che essa sia nota all'allibratore, mettendolo in grado di vincere a colpo (quasi) sicuro. Il quasi è dovuto al fatto che il canale di comunicazione è soggetto a errori, e talvolta trasmette il risultato sbagliato. Supponiamo per semplicità che ci siano solo due cavalli, con uguale probabilità di vittoria a priori. Il segnale ricevuto è esatto con probabilità p e sbagliato con probabilità $q = 1 - p$. Se il giocatore punta sul cavallo corretto, vince 2 volte la posta. Qual è la migliore strategia per il giocatore?

Se il giocatore può giocare una sola volta, e $p > \frac{1}{2}$, la situazione è chiara. Il giocatore tenderà ad ottimizzare il valor medio della vincita, che è pari a due volte la posta se sceglie il cavallo giusto (con probabilità p) e a 0 se sceglie il cavallo sbagliato. Quindi il giocatore punterà tutto il suo capitale sul cavallo indicato dal complice, e ritornerà a tasche vuote con probabilità q .

Le cose cambiano se le corse si ripetono indefinitamente (sempre con le stesse probabilità) e il giocatore può reinvestire il proprio capitale. In questo caso, a ogni corsa, con probabilità p il giocatore raddoppierebbe il suo capitale, e con probabilità q lo perderebbe tutto. È vero che il valor medio della vincita viene ad essere uguale a $(2p)^N$ dopo N corse, ma il giocatore avrebbe perso tutto con probabilità $1 - q^N$, che tende rapidamente a 1 al crescere di N .

Di conseguenza conviene che il giocatore non punti tutto il suo capitale, in modo da garantirsi nel caso che le informazioni ricevute siano sbagliate. Indichiamo con b la frazione del capitale puntata sul cavallo suggerito. Quale criterio adottare per scegliere b ? Consideriamo il capitale del giocatore dopo N giocate, in cui il cavallo suggerito ha vinto N_+ volte, e ha perso N_- volte. Otteniamo

$$V_N = (1 + b)^{N_+} (1 - b)^{N_-} V_0, \quad (3)$$

dove V_0 è il capitale iniziale. Da questa espressione è chiaro che il capitale vinto tende a crescere (o decrescere) esponenzialmente. Questo suggerisce di ottimizzare non il valor medio di V_N , ma il suo logaritmo, o meglio, il logaritmo del rapporto fra V_N e V_0 . Definiamo

$$\Lambda(b) = \lim_{N \rightarrow \infty} \frac{1}{N} \ln \frac{V_N}{V_0}. \quad (4)$$

La quantità $\Lambda(b)$ è detta **tasso di crescita del capitale**. Si ha

$$\ln \frac{V_N}{V_0} = \sum_{i=1}^N \ln [\epsilon_i(1+b) + (1-\epsilon_i)(1-b)], \quad (5)$$

dove $\epsilon_i = 1$ se nella corsa i vince il cavallo suggerito, e $\epsilon_i = 0$ se esso invece perde. Questa espressione è la somma di variabili indipendenti, identicamente distribuite, e per la legge dei grandi numeri la sua media empirica, che appare in eq. (4), tende con probabilità 1 alla sua media valutata sulla distribuzione di probabilità.² Dato che $\langle \epsilon_i \rangle = p$, $\forall i$, otteniamo

$$\begin{aligned} \Lambda(b) &= \langle \ln [\epsilon(1+b) + (1-\epsilon)(1-b)] \rangle \\ &= p \ln(1+b) + q \ln(1-b). \end{aligned} \quad (6)$$

La strategia ottimale b^* si ottiene imponendo la condizione

$$\Lambda'(b^*) = \frac{p}{1+b^*} - \frac{q}{1-b^*} = 0, \quad (7)$$

che dà

$$b^* = 2p - 1. \quad (8)$$

Il tasso di crescita ottimale $\Lambda(b^*)$ è dato da

$$\Lambda(b^*) = \ln 2 + p \ln p + q \ln q, \quad (9)$$

dove abbiamo usato la relazione $q = 1 - p$. Ma questa espressione può essere scritta

$$\Lambda(b^*) = \ln 2 - H(X), \quad (10)$$

dove

$$H(X) = -(p \ln p + q \ln q) \quad (11)$$

è la misura dell'incertezza contenuta nel messag-

gio dato dal complice (con la costante $k = 1$). Possiamo interpretare questo risultato nel modo seguente: se il risultato della corsa fosse certo, il giocatore raddoppierebbe il suo capitale a ogni corsa, e si avrebbe $\Lambda = \ln 2$. A causa dell'incertezza della comunicazione, la vincita viene ridotta di una quantità esattamente uguale all'entropia del messaggio. Alla peggio, se l'incertezza del messaggio è totale, si ha $H(X) = \ln 2$ e la vincita è nulla (in media). O, detto in altro modo, la riduzione dell'incertezza sul vincitore della corsa, partendo dal valore $\ln 2$ in assenza d'informazione e raggiungendo il valore $H(X)$ in presenza del messaggio, è uguale al tasso di crescita del capitale, e costituisce una misura del valore dell'informazione contenuta nel messaggio.

Il fatto che il valore dell'informazione ricevuta stia nella riduzione dell'entropia viene confermato dall'osservazione che il tasso ottimale di crescita del capitale non dipende dai dettagli della strategia del giocatore. Supponiamo, per esempio, che il giocatore, invece di trattenere una frazione $1 - b$ del capitale, decida di utilizzarla per scommettere sull'altro cavallo, quello *non* suggerito dal messaggio. In questo caso, $\Lambda(b)$ viene espresso da

$$\begin{aligned} \Lambda(b) &= p \ln(2b) + q \ln(2(1-b)) \\ &= \ln 2 + p \ln b + q \ln(1-b). \end{aligned} \quad (12)$$

La strategia ottimale è data da

$$b^* = p. \quad (13)$$

Sostituendo, otteniamo

$$\Lambda(b^*) = \ln 2 + p \ln p + q \ln q = \ln 2 - H(X), \quad (14)$$

esattamente come nel caso precedente.

Un caso più generale

Possiamo analizzare il problema in una situazione più generale, in cui ci sono N cavalli, ciascuno con una probabilità p_x di vincere ($x \in \{1, \dots, N\}$), e in cui l'allibratore restituisce o_x volte la posta se vince il cavallo x . Questo paragrafo è necessariamente più tecnico, perché richiede di introdurre un certo numero di quantità analoghe all'entropia di Shannon, di cui faremo uso nel seguito. Supponiamo di avere a che fare

²In questo articolo, utilizzo le parentesi angolari per denotare la media. Data una funzione f_x della variabile aleatoria discreta X , si ha $\langle f \rangle = \sum_x p_x f_x$, dove p_x è la probabilità che X assuma il valore x .

con un allibratore onesto, che non trattiene per sé (in media) parte delle scommesse. In questo caso si ha

$$\sum_{x=1}^N \frac{1}{o_x} = 1. \quad (15)$$

In questa situazione $1/o_x$ può essere interpretata come la stima q_x che l'allibratore fa della probabilità che vinca il cavallo x . Indichiamo con $b = (b_x)$ la frazione del capitale che il giocatore punta sul cavallo x , con $\sum_x b_x = 1$. Si ha allora

$$\Lambda(b) = \sum_x p_x \ln \frac{b_x}{q_x}. \quad (16)$$

Ottimizzando rispetto a b (la dimostrazione si trova nell'Appendice B) e tenendo conto della condizione $\sum_x b_x = 1$ si ottiene

$$b_x^* = p_x, \quad \forall x. \quad (17)$$

Sorprendentemente, le quote $o = (o_x)$ non influenzano la strategia ottimale b^* , ma solo il tasso di crescita del capitale. Si ha infatti

$$\Lambda(b^*) = \sum_x p_x \ln(o_x b_x^*) = \sum_x p_x \ln \frac{p_x}{q_x}. \quad (18)$$

L'espressione al secondo membro è definita non negativa, e si annulla solo se $p_x = q_x, \forall x$, (queste proprietà sono mostrate nell'Appendice C). L'equazione (18) costituisce una misura di quanto le distribuzioni di probabilità $p = (p_x)$ e $q = (q_x)$ sono differenti, e viene chiamata **divergenza di Kullback-Leibler** e denotata con $D_{\text{KL}}(p||q)$. Non può essere considerata una distanza: in particolare non è simmetrica, perché si può vedere che se, p.es., $q_x = 0$ per un qualche x , mentre $p_x > 0, \forall x$, si ha $D_{\text{KL}}(p||q) = \infty$, ma $D_{\text{KL}}(q||p)$ è finito.

Anche in questa situazione possiamo considerare l'effetto di un canale di comunicazione che produce un segnale Y . Indichiamo con $p = (p_x)$ le probabilità a priori che vinca il cavallo x , e con $p_{x|y}$ la probabilità condizionata che il cavallo vincente sia x dato che il giocatore abbia ricevuto il segnale y . Indichiamo inoltre con p_y la probabilità che il segnale ricevuto sia y . Queste probabilità sono legate alla probabilità congiunta $p_{x,y}$ che vinca il cavallo x e che il segnale ricevuto

sia y :

$$p_{x,y} = p_{x|y}p_y, \quad p_x = \sum_y p_{x,y}, \quad p_y = \sum_x p_{x,y}. \quad (19)$$

Ci si riferisce a p_x e p_y , in questo contesto, come alle **marginali** della distribuzione congiunta $p_{x,y}$. Dobbiamo anche introdurre $b_{x|y}$, che è la frazione del capitale scommessa sul cavallo x quando sia stato ricevuto il segnale y . Essa soddisfa le relazioni

$$\sum_x b_{x|y} = 1, \quad \forall y. \quad (20)$$

Si ha allora

$$\Lambda_Y(b) = \sum_{x,y} p_{x,y} \ln(o_x b_{x|y}) = \sum_{x,y} p_{x|y}p_y \ln \frac{b_{x|y}}{q_x}, \quad (21)$$

dove $\Lambda_Y(b)$ ricorda che stiamo considerando il processo in presenza del segnale Y . Otteniamo così la strategia ottimale

$$b_{x|y}^* = p_{x|y}, \quad \forall x, y, \quad (22)$$

e il corrispondente tasso di crescita:

$$\Lambda_Y(b^*) = \sum_{x,y} p_{x,y} \ln \frac{p_{x|y}}{q_x} = \sum_{x,y} p_{x,y} \ln \frac{p_{x,y}}{q_x p_y}. \quad (23)$$

Possiamo riscrivere questa espressione come segue:

$$\Lambda_Y(b^*) = \sum_{x,y} p_{x,y} \ln \frac{p_{x,y}}{p_x p_y} + \sum_x p_x \ln \frac{p_x}{q_x}. \quad (24)$$

Il secondo termine è $D_{\text{KL}}(p||q)$, che corrisponde al tasso ottimale di vincita in assenza di comunicazione parallela. Il primo termine è la divergenza di Kullback-Leibler fra la probabilità congiunta $p_{x,y}$ e il prodotto delle marginali p_x e p_y . Questa quantità è non negativa e si annulla solo se X e Y sono indipendenti. Indichiamola con $I(X : Y)$. Per interpretarla, date due variabili X e Y , con distribuzione congiunta $p_{x,y}$, definiamo l'**entropia condizionata** $H(X|Y)$ mediante l'espressione

$$\begin{aligned} H(X|Y) &= - \sum_{x,y} p_{x,y} \ln p_{x|y} \\ &= - \sum_y p_y \sum_x p_{x|y} \ln p_{x|y}. \end{aligned} \quad (25)$$

L'entropia condizionata è il valor medio dell'entropia della distribuzione condizionata di x dato y , valutato sulla marginale di y . Si ha allora

$$I(X : Y) = H(X) - H(X|Y) = H(Y) - H(Y|X). \quad (26)$$

Vediamo così che $I(X : Y)$ dà una misura di quanto la conoscenza di una variabile diminuisce l'incertezza sull'altra variabile. Essa è nota come **informazione mutua**. Dalla sua espressione si vede che è simmetrica:

$$I(X : Y) = I(Y : X). \quad (27)$$

Concludendo, abbiamo ottenuto

$$\Lambda_Y(b^*) = I(X : Y) + D_{\text{KL}}(p||q), \quad (28)$$

che mostra che l'informazione mutua è pari all'incremento del tasso di crescita del capitale dovuto alla presenza del segnale Y . In particolare, nel caso di un allibratore non solo onesto ($\sum_x 1/o_x = 1$), ma anche competente ($q_x = p_x, \forall x$), il tasso di crescita ottimale del capitale è esattamente uguale all'informazione mutua fra segnale e risultato.

La macchina di Szilard

Il secondo principio della termodinamica, nella formulazione di Kelvin, stipula che non è possibile disegnare una macchina che, operando ciclicamente, estragga energia sotto forma di calore da un serbatoio a temperatura uniforme e la restituisca sotto forma di lavoro. Nel 1929, Leo Szilard (noto fra l'altro per avere scritto con Einstein la lettera a Roosevelt che lanciò il Progetto Manhattan) propose un esperimento concettuale che sembrava violare questo principio [7]. La macchina è costituita da un cilindro di volume \mathcal{V} contenente una sola molecola di gas, e in contatto con un serbatoio di calore alla temperatura T (cf. fig. 1). Un demone dispone nel cilindro una partizione mobile, dividendolo in due sezioni, non necessariamente uguali. Indichiamo con p_S la frazione del volume contenuta nella sezione di sinistra e con p_D quella della sezione destra. A questo punto il demone identifica in quale sezione si trova la molecola. Se essa, per esempio, si trova nella sezione sinistra, esso attacca una barra alla parte destra della sezione mobi-

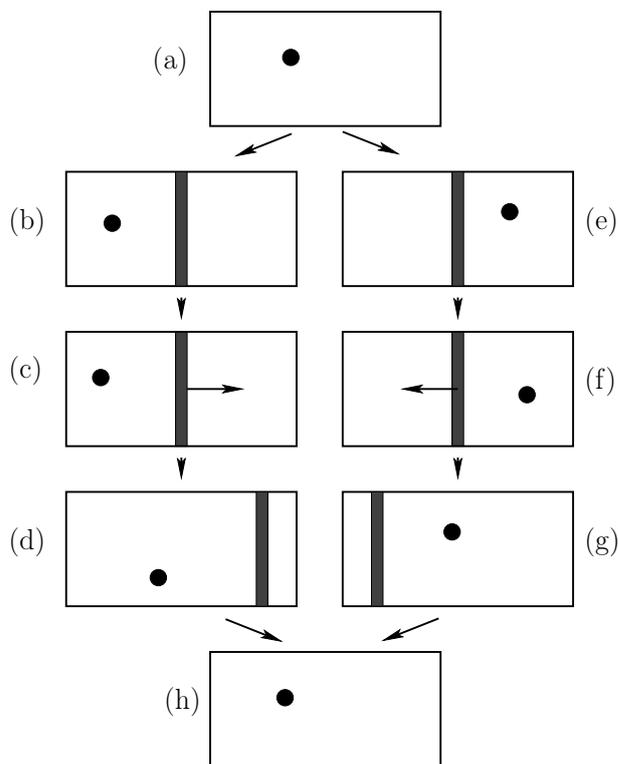


Figura 1: Il ciclo della macchina di Szilard. Si pone una partizione in un cilindro, dividendolo in due sezioni di volume $p_S\mathcal{V}$ e $p_D\mathcal{V}$ rispettivamente. Quindi si effettua una misura della posizione (S o D) della molecola. Dato il risultato $y \in \{S, D\}$ della misura, la partizione viene spostata reversibilmente dalla posizione iniziale alla posizione finale. La partizione viene rimossa e il ciclo ricomincia.

le. (Se essa si trova a destra, l'attacca alla parte sinistra.) Quindi lascia espandere reversibilmente la sezione contenente la molecola, finché la partizione tocca la parete del cilindro. In questa espansione viene compiuto lavoro sull'esterno, pari a $k_B T \ln(\mathcal{V}/\mathcal{V}_0)$, dove \mathcal{V}_0 è il volume iniziale della partizione, pari a $p_S\mathcal{V}$ se la molecola si trovava a sinistra, o a $p_D\mathcal{V}$ se si trovava a destra. A questo punto il demone rimuove la partizione e il sistema ritorna apparentemente alla condizione iniziale.

Valutiamo il valor medio W del lavoro estratto dal demone in un ciclo. La molecola si trova a sinistra con probabilità p_S . Supponendo che il demone non si sbaglia mai, esso estrarrà in questo caso una quantità di lavoro pari a $k_B T \ln(\mathcal{V}/\mathcal{V}_0) = -k_B T \ln p_S$. Ripetendo l'argomento per il caso in cui la molecola si trova a

destra, otteniamo il valor medio

$$W = -k_B T (p_S \ln p_S + p_D \ln p_D) = k_B T H(X), \quad (29)$$

dove X è una variabile aleatoria che assume i valori (S, D) rispettivamente con probabilità (p_S, p_D) . Vediamo così che la riduzione dell'incertezza sulla posizione della molecola, dovuta alla misura, può essere utilizzata per convertire calore in lavoro. Se la misura è affetta da errore, dobbiamo tener conto del fatto che, se la misura è sbagliata, comprimere il gas contro la parete (riducendo il suo volume a 0) costerebbe un lavoro infinito. Quindi la partizione non può toccare la parete del cilindro alla fine della manipolazione. Indichiamo con $b_{x|y}$ la frazione del volume \mathcal{V} in cui si trova la particella alla fine della manipolazione, data la sua posizione iniziale $x \in \{S, D\}$ e il risultato della misura $y \in \{S, D\}$. La probabilità congiunta di x e y è denotata con $p_{x,y}$, mentre $p_{x|y}$ è, come nel paragrafo precedente, la probabilità condizionata che la molecola sia in x se il risultato della misura è y . Il lavoro ottenuto dalla manipolazione è quindi dato da $k_B T \ln(b_{x|y}/p_x)$, e può essere positivo (se la misura è corretta) o negativo (se è sbagliata). Il valor medio W del lavoro estratto è dato da

$$W(b) = k_B T \sum_{x,y} p_{x,y} \ln \frac{b_{x|y}}{p_x}. \quad (30)$$

Vediamo che questa espressione coincide con il tasso di crescita del capitale dato in eq. (21), per il caso dell'allibratore competente (che sa cioè valutare correttamente le probabilità) moltiplicato per l'energia termica $k_B T$. In analogia con quel caso, vediamo che la scelta ottimale di b è data da

$$b_{x|y}^* = p_{x|y}, \quad \forall y, \quad (31)$$

e il lavoro medio W corrispondentemente estratto è dato da

$$\begin{aligned} W(b^*) &= k_B T \sum_{x,y} p_{x,y} \ln \frac{p_{x|y}}{p_x} \\ &= k_B T \sum_{x,y} p_{x,y} \ln \frac{p_{x,y}}{p_x p_y} \\ &= k_B T I(X : Y). \end{aligned} \quad (32)$$

Otteniamo così il risultato che il lavoro medio W che può essere estratto da una macchina di

Szilard soddisfa la disuguaglianza

$$W \leq k_B T I(X : Y), \quad (33)$$

dove $I(X : Y)$ è l'informazione mutua tra la misura Y e la posizione X della molecola. Questo risultato è stato esteso a situazioni più generali da Sagawa e Ueda [8].

Come si concilia l'operazione della macchina di Szilard con il secondo principio della termodinamica? Szilard [7] e altri ricercatori, fra cui Léon Brillouin [9] hanno suggerito che il processo di misura è intrinsecamente dissipativo, e che la dissipazione implicata non può essere inferiore a $k_B T \Delta H(X)$, dove $\Delta H(X)$ è la differenza fra l'incertezza su X prima e dopo la misura. Tuttavia Bennett [10] ha fatto vedere che è possibile, in linea di principio, eseguire delle misure senza dissipazione, purché l'apparato di misura sia in uno stato ben definito prima della misura stessa.

Conseguentemente, il ciclo della macchina di Szilard non riporta il sistema esattamente nello stato iniziale, perché l'apparato di misura conserva il risultato della misura stessa. Per ottenere una trasformazione perfettamente ciclica è necessario riportarlo nello stato definito iniziale. In questa operazione l'incertezza sullo stato dell'apparato di misura viene cancellata, in altri termini, l'entropia $H(Y)$ del sistema diminuisce. Questo violerebbe il secondo principio, a meno che non venga dissipata una quantità di lavoro almeno pari a $k_B T \Delta H(Y)$, dove $\Delta H(Y)$ è la differenza dell'incertezza su Y prima e dopo il ripristino dello stato iniziale dell'apparato.

Un'analisi basata sulla termodinamica stocastica mostra che in condizioni molto generali questo è esattamente ciò che avviene. Questo risultato venne suggerito da Rolf Landauer nel 1961 [11], ed è stato sperimentalmente dimostrato in diversi sistemi mesoscopici in anni recenti [12, 13]. Esso è noto come il **principio di Landauer** ed è stato enunciato da Bennett [10] come segue:

"Ogni manipolazione logicamente irreversibile dell'informazione, come la cancellazione di un bit o la confluenza di due percorsi computazionali, deve essere accompagnata da un corrispondente incremento d'entropia nei gradi di libertà non

Corse di Kelly	Macchina di Szilard
x_i – risultato della corsa i	x_i – posizione della molecola nel ciclo i
Informazione parallela	Risultato della misura
y_i – informazione sulla corsa i	y_i – misura imprecisa nel ciclo i
(p_x) – distr. di prob. del risultato	(p_x) – distr. di prob. della posizione
$p_{x y}$ – prob. cond. di x , dato y	$p_{x y}$ – prob. cond. di x data la misura
o_x – quota della scommessa su x	$1/p_x$ – inverso del volume iniziale
Piazzare scommesse	Muovere la partizione al punto finale
Logaritmo del capitale	Lavoro estratto
Tasso di crescita del capitale	Lavoro medio estratto per ciclo

Tabella 1: Analogia fra le corse di Kelly e la macchina di Szilard. Adattato dalla ref. [14].

portanti informazione dell'apparato di manipolazione dell'informazione o del suo ambiente."

Una volta tenuto conto della dissipazione conseguente al ripristino dell'apparato di misura, si vede che il valor medio W del lavoro netto ottenuto dalla macchina di Szilard non può essere positivo.

Analogia fra gioco d'azzardo ed estrazione di lavoro

L'analogia fra il funzionamento della macchina di Szilard e le corse di Kelly è stata notata da Vinkler e collaboratori [14]. Questa analogia è dettagliata nella Tabella 1. Notiamo che, a differenza dell'allibratore delle corse di Kelly, nella macchina di Szilard si ha a che fare con un allibratore sempre competente, in cui l'analogo della quota o_x è sempre uguale all'inverso della probabilità p_x della posizione iniziale. Tuttavia, a differenza delle corse di cavalli, in questo caso la posizione iniziale della partizione può essere scelta dal demone, il che corrisponde a scegliere arbitrariamente la quota o_x (sempre in maniera competente). Quindi nella macchina di Szilard è possibile scegliere la posizione della partizione iniziale in modo da ottimizzare la mutua informazione $I(X : Y)$, e disporre conseguentemente le posizioni finali $b_{x|y}$ in modo da ottimizzare W . Otteniamo così il valore massimo W_{\max} del lavoro estraibile:

$$W_{\max} = \max_{p_x} I(X : Y). \quad (34)$$

Riassumendo: l'analogia non è perfetta, perché nel caso di Kelly l'allibratore ha la possibilità di scegliere le quote o_x arbitrariamente (ma la scelta ottimale è sempre $o_x = 1/p_x$), mentre nel caso di Szilard il giocatore (cioè il demone) ha la possibilità di scegliere p_x in modo da ottimizzare il lavoro estratto.

L'analogia può essere estesa al caso in cui una parte del lavoro estratto viene dissipata irreversibilmente (per esempio a causa di attriti). Questo corrisponde al caso in cui l'allibratore detiene per sé una parte delle scommesse, e si ha così $\sum_x 1/o_x > 1$. Per questo caso, Kelly ha elaborato una strategia sofisticata, che contempla in generale la possibilità di trattenere una parte del capitale. Nel caso della macchina di Szilard, questo corrisponde alla possibilità di non compiere la manipolazione se il lavoro dissipato viene ad essere in eccesso di quello estratto.

Popolazioni e fitness

Nel caso delle popolazioni, il capitale è costituito dal numero di individui che la compongono. Si può dire che un individuo piazza una scommessa producendo prole, e che questa scommessa viene a maturazione quando la prole è a sua volta in grado di riprodursi. La **fitness** di un individuo è il numero medio dei suoi figli che giunge a potersi riprodurre. Quindi la *fitness* è analoga alla quota nelle scommesse, e possiamo in linea di principio applicare a questo caso i ragionamenti fatti nei paragrafi precedenti.

Seguendo Donaldson-Matasci e collaboratori [15], consideriamo una forma di vita esposta ad un ambiente fluttuante X , che può assumere

con probabilità p_x uno stato x fra un certo numero di stati possibili. Il fenotipo Φ dell'organismo può anch'esso assumere un certo numero di stati ϕ . Ciascun fenotipo prospera se si trova in un ambiente adatto, e altrimenti viene penalizzato. Questo significa che la *fitness* $f_{x|\phi}$ del fenotipo ϕ nell'ambiente x è grande se ϕ è adatto ad x ed è piccola altrimenti. Adottiamo la convenzione che il fenotipo adatto all'ambiente x ha il valore x . Consideriamo il caso limite in cui il fenotipo ϕ , con $\phi \neq x$, non è in grado di riprodursi.³ Quindi la taglia media della prole di un individuo con fenotipo ϕ in un ambiente x è pari a $f_{x|\phi}$ se $\phi = x$, ed a 0 altrimenti. Di conseguenza, conviene agli individui produrre prole con fenotipi diversi, allo stesso modo che conviene ai giocatori fare puntate su diversi cavalli.

Questo fenomeno è stato osservato, per esempio, in certe popolazioni batteriche, in cui una parte degli individui hanno un metabolismo più lento (*persisters*) a parità di genotipo, il che permette loro di resistere agli antibiotici (vedi, per esempio [17]).

Denotiamo $f_{x|x}$ con o_x , e con b_x la frazione di individui della prole che posseggono il fenotipo x . Allora il tasso medio di incremento della popolazione, definito da

$$\Lambda(b) = \lim_{N \rightarrow \infty} \frac{1}{N} \ln \frac{\mathcal{N}_N}{\mathcal{N}_0}, \quad (35)$$

dove \mathcal{N}_N è il numero di individui alla N -esima generazione, e \mathcal{N}_0 è il numero iniziale di individui, è dato da

$$\Lambda(b) = \sum_x p_x \ln(o_x b_x). \quad (36)$$

Questo corrisponde all'equazione (16), con la sola differenza che $\sum_x 1/o_x$ non è necessariamente uguale a 1. Anzi ci aspettiamo che in condizioni favorevoli si abbia $\sum_x 1/o_x < 1$, cioè che la popolazione tenda ad aumentare. Questo corrisponde a un allibratore generoso che distribuisce più di quanto ricava dalle scommesse. Possiamo tuttavia definire $Z = (\sum_x 1/o_x)^{-1}$ e $q_x = Z/o_x$, così che q_x è normalizzato. In questo caso il tasso di

crescita per la strategia $b = (b_x)$ è dato da

$$\Lambda(b) = \ln Z + \sum_x p_x \ln \frac{b_x}{q_x}. \quad (37)$$

Come al solito, il valore ottimale b^* di b è dato da

$$b_x^* = p_x, \quad \forall x. \quad (38)$$

Il corrispondente tasso medio di crescita è dato da

$$\begin{aligned} \Lambda(b^*) &= \sum_x p_x \ln(o_x p_x) \\ &= \ln Z + \sum_x p_x \ln \frac{p_x}{q_x} \\ &= \ln Z + D_{\text{KL}}(p||q). \end{aligned} \quad (39)$$

Vediamo che $\ln Z \geq 0$ è il minimo tasso ottimale di crescita, che si ottiene quando $p_x = q_x, \forall x$. Questo corrisponde a un processo ambientale perverso, che minimizza il benessere degli organismi, mentre in generale si possono avere tassi più elevati se le probabilità p_x dell'ambiente non coincidono con le probabilità q_x definite dalla distribuzione della *fitness*.

Consideriamo adesso le probabilità ambientali p_x fissate, e valutiamo il massimo tasso di crescita possibile, che si potrebbe ottenere se l'organismo fosse in grado di individuare con certezza l'ambiente in cui si trova, e produrre prole tutta con il fenotipo adatto. Otteniamo

$$\begin{aligned} \Lambda_{\text{max}} &= \ln Z - \sum_x p_x \ln q_x \\ &= \ln Z + \sum_x p_x \ln \frac{p_x}{q_x} - \sum_x p_x \ln p_x \\ &= \ln Z + D_{\text{KL}}(p||q) + H(X). \end{aligned} \quad (40)$$

Otteniamo così

$$\Lambda_{\text{max}} - \Lambda(b^*) = H(X), \quad (41)$$

che mostra che il tasso di crescita viene ridotto dall'entropia del processo ambientale.

Tuttavia può non essere possibile per l'organismo di aggiustare esattamente la frequenza b_x del fenotipo x alla frequenza p_x dell'ambiente x nel processo ambientale. Ci aspettiamo che in questo caso il tasso di crescita della popolazione sia più piccolo che nel caso ottimale. Un semplice

³Questa condizione può essere resa meno stringente (cf. Haccou e Iwasa [16]), ma non è il caso di discutere qui come procedere nel caso più generale.

calcolo dà infatti

$$\Lambda(b^*) - \Lambda(b) = D_{\text{KL}}(p||b). \quad (42)$$

Questo mostra che una valutazione accurata del processo ambientale permette una crescita più vigorosa.

Il valore di *fitness* della percezione

Da quanto detto nel paragrafo precedente, risulta evidente che gli organismi possono trarre vantaggio da informazioni tempestive sullo stato dell'ambiente. Supponiamo che sia disponibile un segnale Y , accessibile a tutti gli individui, tale che la probabilità che l'ambiente si trovi ad avere il valore x quando il segnale vale y è data da $p_{x|y}$. Il segnale gioca il ruolo del telegrafo nel caso delle corse di cavalli. Il tasso di crescita della popolazione è dato dalla stessa espressione (21) che per le corse. Come in quel caso, la strategia ottimale è una strategia condizionata $b_{x|y}^*$, data da

$$b_{x|y}^* = p_{x|y}, \quad \forall x, y. \quad (43)$$

Il tasso corrispondente di crescita della popolazione è dato da

$$\Lambda_Y(b^*) = \sum_{x,y} p_{x,y} \ln(o_x p_{x|y}), \quad (44)$$

dove $\Lambda_Y(b)$ ricorda che siamo in presenza del segnale Y correlato all'ambiente. Si ha allora

$$\begin{aligned} \Lambda_Y(b^*) - \Lambda(b^*) &= \sum_{x,y} p_{x,y} \ln \frac{p_{x|y}}{p_x} \\ &= \sum_{x,y} p_{x,y} \ln \frac{p_{x,y}}{p_x p_y} = I(X : Y), \end{aligned} \quad (45)$$

come nel caso delle scommesse.

A questa analisi si può obiettare che gli individui che l'informazione sull'ambiente non è comune a tutti gli individui, ma è distribuita. In altre parole, che il segnale ricevuto non è necessariamente lo stesso per tutti gli individui, ma differisce (in principio) da individuo a individuo. Indichiamo con $p_{y|x}$ la probabilità condizionata che l'individuo riceva il segnale y se l'ambiente è nello stato x . Allora la probabilità che il fenotipo

ϕ sia uguale a x è data da

$$b_x = \sum_y b_{x|y} p_{y|x}. \quad (46)$$

Il tasso di crescita della popolazione in questo caso (informazione distribuita) è dato da

$$\begin{aligned} \Lambda_Y(b) &= \ln Z + \sum_x p_x \ln \left(\sum_y \frac{b_{x|y} p_{y|x}}{q_x} \right) \\ &= \Lambda_{\max} + \sum_x p_x \ln \left(\sum_y b_{x|y} p_{y|x} \right), \end{aligned} \quad (47)$$

dove $\Lambda_{\max} = \ln Z - \sum_x p_x \ln q_x = \langle \ln o_x \rangle$. Il secondo termine è dato dalla media su x del logaritmo del valor medio su y di $b_{x|y}$, valutata con la distribuzione condizionata $p_{y|x}$. D'altra parte, quando il segnale y è lo stesso per tutti gli individui (informazione comune), si ottiene

$$\begin{aligned} \Lambda_Y(b) &= \ln Z + \sum_{x,y} p_{x,y} \ln \frac{b_{x|y}}{q_x} \\ &= \Lambda_{\max} + \sum_x p_x \sum_y p_{y|x} \ln b_{x|y}. \end{aligned} \quad (48)$$

Per ogni valore di x abbiamo nel caso dell'informazione distribuita il logaritmo della media di $b_{x|y}$ su $p_{y|x}$, e nel caso di informazione comune la media del logaritmo. Ora, data una variabile aleatoria W a valori positivi, si ha in generale

$$\ln \langle w \rangle \geq \langle \ln w \rangle. \quad (49)$$

Questa relazione è dimostrata nell'appendice C. Abbiamo ottenuto così che ci si aspetta una crescita della popolazione più grande nel caso di informazione distribuita che in quello di informazione comune. Questo risultato è dovuto a Rivoire e Leibler [18], e può essere considerato come un riflesso della "saggezza delle folle" osservata da Galton nel 1907 [19].

Discussione

Abbiamo visto che l'informazione su un processo aleatorio X contenuta in un segnale Y , e misurata dall'informazione mutua $I(X : Y)$, è una misura del valore dell'informazione stessa in ambienti tanto diversi come le scommesse, l'estrazione di lavoro, e la crescita delle popolazioni.

Un ingrediente essenziale di tutte le situazioni è di avere a che fare con un processo X stazionario, dove Y fornisce informazioni relative alle successive istanze del processo, e dove i guadagni si accumulano senza interferire (nel caso delle scommesse e della popolazione, il guadagno deve essere misurato su scala logaritmica). Il discorso può essere generalizzato al caso in cui il processo X ha memoria (e può quindi essere rappresentato da un processo di Markov). In questo caso si pone il problema della memoria del segnale, e del costo e del valore della predizione del processo. Alcuni lavori sono stati dedicati a questi aspetti, ma discuterne qui ci porterebbe troppo lontano. Per ora mi accontento di avere richiamato l'attenzione dei lettori sull'inaspettata generalità delle applicazioni di alcuni concetti base della teoria dell'informazione.

Appendici

A. Entropia di Shannon

Indichiamo con X una variabile aleatoria, con valori da 1 a N , e tale che la probabilità che essa assuma il valore x è pari a p_x , dove $0 \leq p_x$ e $\sum_x p_x = 1$. Vogliamo definire una misura $H(X)$ dell'incertezza associata a X . Questa misura non deve dipendere dalla natura di X , ma solo dalla sua distribuzione di probabilità $p = (p_1, \dots, p_N) = (p_x)$. Inoltre, chiaramente, non deve dipendere dalla scelta delle etichette $1, 2, \dots, N$ che assegniamo ai valori di X . Possiamo richiedere ancora che abbia le seguenti proprietà:

1. Se un solo valore x_0 di X è possibile, cioè se $p_x = \delta_{xx_0}$, dove δ_{xy} è la delta di Kronecker, definita da

$$\delta_{xy} = \begin{cases} 1, & \text{se } x = y; \\ 0, & \text{altrimenti;} \end{cases} \quad (50)$$

l'incertezza si deve annullare.

2. Se X può assumere uno fra N valori, tutti equiprobabili, l'incertezza $H(X)$ deve crescere al crescere di N . Indichiamo con $f(N)$ l'incertezza associata alla distribuzione $p_x =$

$1/N$ per $x \in \{1, \dots, N\}$. Si ha allora

$$f(N) > f(M), \quad \text{se } N > M. \quad (51)$$

3. Valori di X che hanno probabilità nulla non contribuiscono all'incertezza.
4. Se è possibile suddividere gli N valori di X in K gruppi, l'incertezza relativa a X diventa pari all'incertezza relativa alla variabile Y che individua il gruppo, più l'incertezza di X dato Y , mediata sulla distribuzione di probabilità di X . Indichiamo con $H_y(X)$ l'incertezza residua su X , una volta che sia noto il valore y di Y . Si ha allora

$$H(X) = H(Y) + \sum_y p_y H_y(X). \quad (52)$$

5. L'espressione $H(X)$ intesa come funzione del vettore $p = (p_x)$ delle probabilità di X , deve essere continua.

Mostriamo adesso che l'espressione (1) è l'unica espressione che soddisfa queste condizioni, a meno dell'arbitrarietà della costante k .

Consideriamo una variabile X che assuma $N = K \cdot M$ valori, tutti equiprobabili. Possiamo raggruppare i valori di X in K gruppi, ciascuno con M valori. Indicando con Y la variabile che identifica i gruppi, abbiamo $H_y(X) = f(M)$, $\forall M$. Per la (52) si ha allora

$$f(N) = f(K) + K \cdot \frac{1}{K} f(M) = f(K) + f(M). \quad (53)$$

Questo suggerisce di porre

$$f(N) = k \ln N, \quad (54)$$

con $k > 0$ per la proprietà 2. Mostrare che questa è l'unica soluzione monotona crescente dell'equazione funzionale (53) è un facile esercizio che lascio al lettore. Questa espressione soddisfa chiaramente la proprietà 1.

Consideriamo adesso una variabile X che assuma N valori, e tale che la probabilità che assuma il valore x sia pari a p_x . Data la continuità della $H(X)$ si commette un errore arbitrariamente piccolo se supponiamo che i valori di p_x siano tutti razionali. Riportiamoli tutti al comun

denominatore \mathcal{N} , ottenendo

$$p_x = \frac{n_x}{\mathcal{N}}, \quad (55)$$

con n_x intero, $\forall x$. Possiamo ora introdurre una variabile aleatoria Y che assume \mathcal{N} valori equiprobabili, divisi in N gruppi identificati da X , e tali che il gruppo x ha n_x valori equiprobabili. Si ha allora, per la (52),

$$\begin{aligned} H(Y) &= H(X) + \sum_x p_x H_x(Y) \\ &= H(X) + \sum_x p_x f(n_x) \\ &= H(X) + k \sum_x p_x \ln n_x. \end{aligned} \quad (56)$$

Dato che $H(Y) = k \ln \mathcal{N}$, otteniamo

$$\begin{aligned} H(X) &= -k \sum_x p_x \ln n_x + k \ln \mathcal{N} \\ &= -k \sum_x p_x \ln p_x. \end{aligned} \quad (57)$$

Per la proprietà 3., l'espressione $p_x \ln p_x$ deve essere posta pari a zero se $p_x = 0$.

B. Strategie ottimali

Desideriamo trovare l'estremale di $\Lambda(b)$ dato dalla Eq. (16) rispetto a b , tenendo conto della condizione di normalizzazione $\sum_x b_x = 1$. Introduciamo il moltiplicatore di Lagrange λ e ottimizziamo

$$\Psi(b) = \Lambda(b) - \lambda \sum_x b_x. \quad (58)$$

Otteniamo le equazioni

$$\frac{p_x}{b_x} - \lambda = 0, \quad \forall x, \quad (59)$$

che ammettono la soluzione

$$b_x = \frac{p_x}{\lambda}, \quad x. \quad (60)$$

Dato che p è normalizzato, si deve porre $\lambda = 1$.

Allo stesso modo, nel caso in cui è presente un'informazione parallela Y , si introduce un moltiplicatore di Lagrange λ_y per ogni valore di y , che imponga il vincolo $\sum_x b_{x|y} = 1, \forall y$. In maniera esattamente analoga si ottengono le

equazioni

$$\frac{p_{x|y}}{b_{x|y}} = \lambda_y, \quad \forall x, y, \quad (61)$$

che ammettono le soluzioni

$$b_{x|y} = p_{x|y}, \quad \forall x, y, \quad (62)$$

correttamente normalizzate.

C. Disuguaglianze

Sia W una variabile aleatoria che assume valori w positivi. Vogliamo dimostrare la disuguaglianza

$$\ln \langle w \rangle \geq \langle \ln w \rangle. \quad (63)$$

Poniamo $w = e^x$ ed esponenziamo ambo i membri. Otteniamo la disuguaglianza

$$\langle e^x \rangle \geq e^{\langle x \rangle}, \quad (64)$$

Osserviamo che la funzione $f(x) = e^x$ è tale che $f''(x) > 0, \forall x$. Allora, per lo sviluppo di Taylor si ha, dati due punti qualunque x_0 e x ,

$$\begin{aligned} f(x) &= f(x_0) + f'(x_0)(x - x_0) \\ &+ \int_{x_0}^x dx' (x - x') f''(x'). \end{aligned} \quad (65)$$

Per il teorema della media pesata, si ha

$$\int_{x_0}^x dx' (x - x') f''(x') = \frac{1}{2} f''(\bar{x})(x - x_0)^2, \quad (66)$$

dove \bar{x} è un punto compreso fra x_0 e x . Otteniamo così

$$f(x) \geq f(x_0) + f'(x_0)(x - x_0). \quad (67)$$

Scegliamo $x_0 = \langle x \rangle$ e prendiamo la media di ambo i membri. Otteniamo così la disuguaglianza (64), da cui segue l'equazione (63).

Osserviamo adesso che, date due distribuzioni $p = (p_x)$ e $q = (q_x)$ della stessa variabile X , si ha

$$D_{\text{KL}}(p||q) = - \sum_x p_x \ln \frac{q_x}{p_x} = - \left\langle \ln \frac{q}{p} \right\rangle. \quad (68)$$

Per la disuguaglianza appena dimostrata, otte-

niamo

$$D_{\text{KL}}(p||q) = - \left\langle \ln \frac{q}{p} \right\rangle \geq - \ln \left\langle \frac{q}{p} \right\rangle$$
$$= - \ln \left(\sum_x p_x \cdot \frac{q_x}{p_x} \right) = - \ln \sum_x q_x = 0,$$

data la normalizzazione di q . Quindi la divergenza di Kullback-Leibler non può essere negativa.



[1] C. E. Shannon: *A mathematical theory of communication*, Bell System Technical Journal; 27, 379 & 623 (1948).

[2] C. E. Shannon, W. Weaver: *The Mathematical Theory of Communication*, University of Illinois Press, Urbana and Chicago (1949). Tradotto in: *La teoria matematica delle comunicazioni*, ETAS Kompass, Milano (1971).

[3] M. Tribus, E. C. McIrvine; *Energy and information* Scientific American, 225 (1971) 179.

[4] J. W. Gibbs; *Elementary Principles in Statistical Mechanics*, Yale University Press, New Haven (1902).

[5] J. L. Kelly: *A new interpretation of information rate*, Bell System Technical Journal, 35 (1956) 917.

[6] G. R. Hill (regia), D. S. Ward (sceneggiatura), *The sting*, Universal Picture (USA) (1973).

[7] L. Szilard: *Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen* Zeitschrift für Physik, 53 (1929). Tradotto in inglese in: H. S. Leff, A. F. Rex (eds.), *Maxwell's Demon 2: Entropy, Classical and Quantum Information, Computing*, IoP Press, Bristol and Philadelphia (2003), p. 110.

[8] T. Sagawa, M. Ueda: *Nonequilibrium thermodynamics of feedback control*, Physical Review E, 85 (2012) 021104.

[9] L. Brillouin: *Maxwell's demon cannot operate: Information and Entropy. I*, Journal of Applied Physics, 22 (1951) 334.

[10] C. H. Bennett: *The thermodynamics of computation—A review*, International Journal of Theoretical Physics, 21, (1982) 905.

[11] R. Landauer: *Irreversibility and heat generation in the computing process*, IBM Journal of Research and Development, 3 (1961) 183.

[12] A. Bérut, A. Arakelyan, A. Petrosyan, S. Ciliberto, R. Dillenschneider, E. Lutz: *Experimental verification of Landauer's principle linking information and thermodynamics*, Nature 483 (2012) 187.

[13] Y. Jun, M. Gavrilov, J. Bechhoefer: *High precision test of Landauer's principle in a feedback trap*, Physical Review Letters, 113 (2014) 190601.

[14] D. A. Vinkler, H. H. Permuter, N. Merhav: *Analogy between gambling and measurement-based work extraction*, Journal of Statistical Mechanics: Theory and Experiment (2016) 043403.

[15] M. C. Donaldson-Matasci, C. T. Bergstrom, M. Lachmann: *The fitness value of information*, Oikos, 119 (2010) 219.

[16] P. Haccou, Y. Iwasa: *Optimal mixed strategies in stochastic environments*, Theoretical Population Biology, 47 (1995) 212.

[17] B. Gollan, G. Grabe, C. Michaux, S. Helaine: *Bacterial persists and infection: Past, present, and progressing*, Annual review of microbiology, 73 (2019) 359.

[18] O. Rivoire, S. Leibler: *The value of information for populations in varying environments*, Journal of Statistical Physics, 142 (2011) 1124.

[19] F. Galton: *Vox Populi*, Nature, 75 (1907) 450.



Luca Peliti: è stato Professore Ordinario di Meccanica Statistica presso l'Università "Federico II" di Napoli. Si è occupato di meccanica statistica, applicata in particolare a sistemi di interesse biologico, e dell'approccio statistico alla teoria dell'evoluzione. È autore, in particolare, di *Appunti di Meccanica Statistica* (Bollati Boringhieri, 2003) e di *Stochastic Thermodynamics: An Introduction*, in collaborazione con Simone Pigolotti, Princeton University Press (2021). Al momento è vice direttore del Santa Marinella Research Institute.

Numero XIX Anno 2022

Viaggio nella Scienza

Ithaca



Raggi Cosmici

