



**Università degli Studi di Napoli
"Federico II"**

TESI DI DOTTORATO DI RICERCA IN
*Tecnologie Innovative per Materiali Sensori e
Imaging*

ciclo XIX

**Progettazione e realizzazione di un
dispositivo microelettronico ASIC per
il controllo della trasmissione dati dai
sensori di un telescopio sottomarino
per neutrini cosmici**

Coordinatore:
Prof. Ruggero Vaglio
Tutor:
Prof. Giancarlo Barbarino

Candidato:
Stefano Russo

Anno Accademico 2005 - 2006

Indice

Introduzione	5
1 L'astrofisica delle particelle	7
1.1 Introduzione	7
1.2 Composizione e spettro energetico dei raggi cosmici	9
1.3 Limiti dell'astronomia con raggi cosmici	13
1.3.1 L'effetto GZK	15
1.4 L'astronomia con gamma e neutrini	17
1.4.1 Astrophysical beam dumps	18
1.4.2 Propagazione γ e ν nell'universo	19
2 La neutrino-astrofisica	21
2.1 Introduzione	21
2.2 Produzione di neutrini in ambienti astrofisici	23
2.3 Sorgenti galattiche di neutrini	24
2.3.1 Il piano galattico e il Sole	24
2.3.2 Resti di Supernova	24
2.3.3 Microquasar e i sistemi binari X	25
2.4 Sorgenti extragalattiche di neutrini	26
2.4.1 Nuclei Galattici Attivi (AGN)	26
2.4.2 Gamma Ray Burst (GRB)	27
2.5 Altre sorgenti	28
2.5.1 Materia Oscura e WIMP	28
2.5.2 Difetti Topologici	29
2.6 Rivelazione dei neutrini di alta energia	29

2.6.1	Interazione dei neutrini	33
2.6.2	Possibilità di osservazione di differenti <i>flavour</i> di neutrini	36
2.7	Propagazione ed interazione dei muoni	37
2.7.1	Perdite di energia dei muoni nella materia	37
2.7.2	Effetto Čerenkov	38
2.7.3	Il tracciamento dei muoni	40
2.8	La situazione internazionale dei telescopi Čerenkov per neutrini	41
2.8.1	DUMAND [17]	41
2.8.2	BAIKAL	42
2.8.3	AMANDA	42
2.8.4	IceCube	44
2.8.5	NESTOR	44
2.8.6	ANTARES	45
2.8.7	Il futuro dei telescopi per neutrini: Km ³ Net	45
3	Il telescopio per neutrini da 1 Km³ NEMO	49
3.1	Introduzione	49
3.2	La scelta del sito di installazione	50
3.2.1	Proprietà ottiche dell'acqua	51
3.2.2	Le correnti marine	53
3.2.3	Il fondo ottico inorganico	54
3.2.4	Il fondo ottico da bioluminescenza	55
3.3	La struttura del rivelatore	57
3.3.1	La geometria del rivelatore	58
3.3.2	La torre ed i piani di NEMO	59
3.4	L'elettronica del telescopio	61
3.4.1	La topologia dei collegamenti	62
3.4.2	L'elettronica di piano	65
3.4.3	Il modulo ottico	67
3.5	Il sistema di trasmissione dati	69
3.5.1	Lo standard SDH	69
3.5.2	Lo standard DWDM	70

4	Codici per il controllo e la correzione di errori nella trasmissione dati	73
4.1	Introduzione	73
4.2	Trasmissione dati su canali rumorosi	74
4.2.1	La soluzione fisica	77
4.2.2	La soluzione di sistema	78
4.3	Codici a controllo e correzione d'errore	79
4.3.1	Codici per la rivelazione di errori	79
4.3.2	Il controllo di parità	80
4.3.3	Codici per la rivelazione e la correzione di errori	82
4.3.4	Codici a ripetizione d'informazione	82
4.4	Come si valuta un codice	86
4.4.1	Il codice perfetto	89
4.5	Il codice di Hamming	90
4.5.1	Codice di Hamming per la correzione di un singolo errore e rivelazione di un doppio errore	94
4.6	Il codice a correzione d'errore per l'esperimento NEMO	95
5	Realizzazione del chip ASIC per il controllo della trasmissione dati in NEMO	97
5.1	Introduzione	97
5.2	Le specifiche di progetto	98
5.3	Confronto tra ASIC e FPGA	100
5.4	La tecnologia ASIC digitale	102
5.4.1	La scelta del processo tecnologico	103
5.4.2	Flusso di progetto a celle standard per un ASIC digitale	104
5.5	L'architettura del chip	109
5.5.1	Funzionamento in trasmissione	110
5.5.2	Funzionamento in ricezione	111
5.5.3	Funzioni aggiuntive	112
5.6	Lo schema a blocchi del chip	113

6	Risultati sperimentali	117
6.1	Introduzione	117
6.2	Simulazioni del progetto	117
6.3	Piazzamento e produzione del chip	122
6.4	Collaudo del dispositivo	125
6.5	Considerazioni finali	128
	Conclusioni	135

Introduzione

Il presente lavoro di tesi si inquadra nell'ambito dell'esperimento NEMO (NEutrino Mediterranean Observatory) finanziato dall' INFN (Istituto Nazionale di Fisica Nucleare) e volto all'osservazione dei neutrini provenienti da sorgenti astrofisiche. Lo studio di tale particelle risulta molto importante per la comprensione dei meccanismi di formazione e di evoluzione stellare e permette di ampliare le conoscenze attuali in merito alle formazione e all'evoluzione dell'universo. Le sonde astrofisiche normalmente utilizzate sono fotoni e raggi cosmici. I fotoni rappresentano la sonda più antica mai utilizzata. Di essa ci si è serviti fin dall'antichità quando si osservava il cosmo mediante strumenti ottici. Tale impiego tuttavia comporta una serie di limiti poiché con i fotoni è possibile osservare solo sorgenti luminose che si trovano a breve distanza dalla Terra. Queste particelle, infatti, vengono facilmente assorbite dalle polveri interstellari e quindi non percorrono distanze superiori alle frazioni di Megaparsec. Anche le particelle più energetiche, come i raggi gamma di alta energia, soffrono di problemi simili e quindi la loro propagazione non supera i Megaparsec. L'evoluzione scientifica e tecnologica ha portato ad utilizzare come sonde astrofisiche anche i raggi cosmici ossia quelle particelle cariche che vengono emesse dai corpi celesti durante la loro evoluzione. Con tali particelle è possibile spingere l'osservazione a distanze più elevate (dell'ordine della decina di Mpc). Tuttavia esse, durante la propagazione verso la Terra, subiscono una serie di deflessioni dovute all'interazione con i campi magnetici intergalattici perdendo ogni informazione riguardo la posizione del corpo celeste che le ha generate. I neutrini, invece, avendo una sezione d'urto bassissima (circa 10^{-40} cm²) si propagano praticamente indisturbati per distanze enormi senza subire alcuna deviazione. E'

anche importante considerare che tali particelle vengono emesse da una vasta gamma di corpi celesti e quindi sono considerate delle ottime sonde astrofisiche. Attraverso lo studio del flusso e dell'energia dei neutrini che arrivano dal cosmo è possibile quindi ottenere molte informazioni preziose riguardo la sorgente che le ha generate. La loro propagazione rettilinea consente, inoltre, di "puntare" le sorgenti permettendo di costruire in maniera sufficientemente dettagliata una mappa del cielo. Sfortunatamente la bassissima probabilità d'interazione comporta grandi difficoltà nella rivelazione dei neutrini che raggiungono la Terra. Infatti, la piccola sezione d'urto unita alla carica neutra fanno dei neutrini una delle particelle più difficili da rivelare.

La tecnica attualmente più diffusa per lo studio dei neutrini prevede di rivelare i muoni prodotti dalla loro interazione con la materia. Ovviamente, avendo una ridottissima sezione d'urto è necessario che i neutrini attraversino grandi quantità di materia prima che possano interagire producendo particelle rivelabili. Per questo motivo, l'esperimento NEMO prevede di usare come bersaglio d'interazione la Terra. Ponendo, infatti, un rivelatore di muoni nella grandi profondità marine si ha una probabilità accettabile che i neutrini, attraversando tutto il globo, interagiscano producendo muoni che si propagano dal fondo del mare verso la superficie. Ricostruendo e studiando le tracce dei muoni è poi possibile ottenere una precisa indicazione dei neutrini che le hanno prodotte. L'acqua del mare nel quale il rivelatore è immerso ha una duplice funzione. Essa è utilizzata innanzitutto come rivelatore. I muoni che l'attraversano con un'energia sufficiente provocano, per effetto Čerenkov, una scia luminosa che può essere osservata mediante l'impiego di fotomoltiplicatori. L'acqua poi ha anche una funzione di schermo per i muoni atmosferici che in questo tipo di misure rappresentano rumore. L'esperimento NEMO prevede di realizzare un rivelatore per neutrini delle dimensioni di 1 Km^3 posto in mare a circa 100 Km da Capo Passero (Sicilia) ad una profondità di circa 3000 metri. Il principio di funzionamento di tale apparato prevede che i segnali provenienti dai fotomoltiplicatori posti in acqua vengano campionati, digitalizzati ed inviati alla stazione di raccolta dati a terra dove saranno analizzati. In questo modo si limita il più possibile l'elettronica di gestione e controllo da posizionare in mare contribuendo all'aumento dell'affidabilità.

Chiaramente gran parte delle probabilità di riuscita di un esperimento con questo tipo di architettura risiedono nell'affidabilità delle comunicazioni. Il presente lavoro è quindi volto allo studio ed alla realizzazione di sistemi che permettano di ridurre drasticamente la perdita di informazioni sulle linee di comunicazione.

Quando si affronta il problema legato alla trasmissione dati su canali rumorosi ci si trova di fronte a due possibili soluzioni. Si può tentare di ridurre la perdita d'informazioni adottando una soluzione che può essere definita fisica oppure una soluzione di sistema. La soluzione fisica prevede di abbassare la rumorosità del canale migliorando i componenti elettronici coinvolti nella trasmissione e utilizzando protocolli di comunicazione estremamente affidabili. Spesso però una soluzione di questo tipo comporta dei costi elevati e comunque non consente di eliminare del tutto la probabilità di perdita di dati. La teoria dell'informazione e dei codici offre un approccio alternativo e per certi versi più furbo per la soluzione di questo problema. Esso consiste non tanto nel migliorare il canale fisico tentando di aumentarne l'affidabilità ma utilizzando delle tecniche di sistema che permettano la correzione di eventuali errori. In definitiva con questo approccio si accetta la rumorosità del canale e si aggiungono sistemi in grado di rivelare gli errori indotti dal canale. Una volta che l'informazione è arrivata a destinazione tali errori vengono individuati e corretti. Queste operazioni vengono effettuate mediante degli algoritmi matematici detti codici per il controllo e la correzione degli errori. Lo studio di questo tipo di codici, iniziato a partire dagli anni 40, ha subito una forte innovazione con l'avvento dei computer. Attualmente esistono moltissimi codici, alcuni notevolmente complessi, che permettono di aumentare l'affidabilità dei canali di trasmissione fino ad annullare quasi del tutto la possibilità di perdita dei dati. Tuttavia il codice perfetto non esiste. La scelta del codice migliore, infatti, va effettuata non solo tenendo presente la rumorosità del canale ed il livello di affidabilità che si vuole raggiungere ma anche considerando la complessità di calcolo richiesta dall'algoritmo. Essa può comportare l'impiego di risorse computazionali notevoli e tempi di elaborazione elevati che non possono essere tollerati dal sistema. Nel caso dell'esperimento NEMO, infatti, l'affidabilità della trasmissione, già elevata

per l'utilizzo di componenti elettronici e protocolli sicuri, deve essere migliorata senza aggiungere alcun tempo di latenza nel flusso dei dati, utilizzando componenti estremamente affidabili e con consumi ridottissimi. A partire da queste considerazioni il mio lavoro di tesi è consistito dapprima nell'individuare il codice più adatto all'aumento dell'affidabilità della trasmissione dati dai sensori del telescopio NEMO e successivamente nel realizzare un dispositivo elettronico ASIC (*Application Specific Integrated Circuit*) in grado di realizzare l'algoritmo selezionato.

Nella prima parte del mio lavoro vengono presentati le sonde astrofisiche classiche quali i fotoni e i raggi cosmici mettendo in evidenza soprattutto i limiti che si incontrano quando si effettua lo studio dell'universo per mezzo di tali particelle. Successivamente è discussa la possibilità di effettuare lo studio dei corpi celesti utilizzando i neutrini. Vengono poi esposte le sorgenti presenti nel cosmo in grado di emettere direttamente o indirettamente i neutrini dimostrando come esse siano tali da poter effettuare uno studio approfondito di molti fenomeni astrofisici diversi e per certi versi sconosciuti. Si passa successivamente all'analisi della propagazione dei neutrini astrofisici verso il pianeta Terra sottolineando come non subiscano praticamente alcun tipo di perturbazione del loro moto. Tale caratteristica risulta estremamente importante poichè permette a queste particelle di conservare una serie di informazioni preziosissime della sorgente che le ha generate. La caratteristica che rende il neutrino particolarmente interessante come sonda astrofisica crea grandi difficoltà di rivelazione una volta che esso giunge sulla Terra; per questo motivo vengono illustrate le tecniche più moderne per la rivelazione dei neutrini ponendo particolare attenzione ai rivelatori che sfruttano l'effetto Čerenkov. La sezione si chiude con una breve rassegna dei principali esperimenti che sfruttano tale tecnica di rivelazione.

Nel terzo capitolo viene presentato in dettaglio l'esperimento NEMO. Sono esposte brevemente le motivazioni che hanno portato alla selezione del sito per l'installazione del rivelatore. Vengono qui esposti i risultati principali della campagna di studio effettuata per determinare le caratteristiche ambientali del tratto di mare preso in esame. Si passa poi alla descrizione delle strutture necessarie alla realizzazione del telescopio e all'architettura

del rivelatore. Un'ampia sezione è dedicata all'elettronica di acquisizione dati descrivendo la sezione di front end, il sistema di comunicazione e gli apparati che permettono il monitoraggio del sistema. In ultimo si passa alla descrizione del sistema di comunicazione tra il rivelatore e la stazione di raccolta dati posta a terra descrivendo il sistema di cavi e connessioni necessario per il trasporto dei dati.

Il quarto capitolo è dedicato alla teoria dei codici a correzione di errore. Dopo una breve introduzione alla materia in oggetto vengono presentati alcuni esempi di algoritmi per il rilevamento e la correzione di errori. In questa sezione vengono forniti gli elementi base per valutare un codice ed i parametri salienti che permettono di selezionare un algoritmo a partire dall'analisi del canale a cui esso deve essere applicato. Viene poi presentato il risultato della prima parte del mio lavoro che è consistita nel determinare l'algoritmo che meglio si adatta alla trasmissione dei dati in NEMO. Si passa, infine, alla descrizione dell'algoritmo di Hamming selezionato e si espongono le sue caratteristiche principali.

Il quinto capitolo è relativo alla descrizione della fase di progettazione del dispositivo elettronico digitale ASIC da me realizzato che riproduce in hardware le funzioni dell'algoritmo di Hamming. Dapprima viene giustificata la scelta dell'impiego di questa tecnologia partendo dal confronto con le principali piattaforme presenti sul mercato. Vengono poi descritti tutti i passi da me compiuti per la realizzazione del progetto. Si parte dalla determinazione delle specifiche di che ho personalmente messo a punto per raggiungere un duplice scopo. Il chip, infatti, deve da un lato essere perfettamente integrabile con l'elettronica del telescopio e dall'altro deve essere di uso del tutto generale. Una volta determinate le specifiche viene esposto in dettaglio il flusso di progetto che ho seguito per la realizzazione del dispositivo presentandone tutte le diverse fasi. Si passa poi all'esposizione delle diverse modalità di funzionamento da me previste per il dispositivo.

Nell'ultimo capitolo sono presentati i risultati sperimentali ottenuti durante la fase di progettazione e durante la campagna di misure effettuata sul chip una volta realizzato. In questo modo è possibile constatare la perfetta riuscita del progetto e le ottime prestazioni ottenute.

Capitolo 1

L'astrofisica delle particelle

1.1 Introduzione

Fin dall'antichità l'astronomia, e più in particolare l'astrofisica, si è basata sull'osservazione dei fenomeni celesti prima ad occhio nudo e poi con telescopi ottici sempre più sofisticati. Questo tipo di osservazione, però, comporta lo studio solo di una piccolissima parte dei fenomeni astrofisici cioè quelli che sono in grado di far arrivare sulla terra radiazione elettromagnetica di lunghezza d'onda visibile compresa tra 380 nm e 780 nm. Alla luce delle conoscenze fisiche acquisite nel corso degli anni, oggi risulta chiaro che questo tipo di osservazione è molto limitata sia perchè non tutti i fenomeni che avvengono nell'universo sono associati all'emissione di radiazione elettromagnetica visibile sia perchè non tutta la luce emessa riesce ad arrivare fino alla Terra. Lo studio dell'astrofisica ha quindi richiesto la necessità di individuare delle sonde che permettessero di osservare un numero più elevato di fenomeni. Con il termine sonda, infatti, si definisce un qualunque tipo di oggetto, sia esso particella o radiazione elettromagnetica in grado di portare con se informazioni relative al fenomeno che l'ha generato. In questo ambito durante il secolo scorso la fisica ha compiuto molti passi in avanti effettuando una lunga serie di scoperte. Nei primi anni del novecento vennero eseguiti studi sull'elettroscopio a foglie in grado di rivelare eventuali cariche indotte su di esso dall'esterno. Tali cariche, depositandosi sulle lamine d'oro presenti sullo strumento tendono a farle divaricare per effetto della repulsione elettro-



Figura 1.1: Victor Hess dopo una delle ascensioni che lo portarono a scoprire l'aumento della velocità di scarica degli elettroscopi con l'altitudine.

statica. L'effetto svanisce se le cariche vengono neutralizzate. Gli scienziati si accorsero che il fenomeno di scarica avveniva in maniera apparentemente spontanea anche quando lo strumento si trovava al buio e lontano da possibili sorgenti radioattive. In un primo momento questo fenomeno fu spigato imputando la ionizzazione delle foglioline agli elementi radioattivi presenti nella crosta terrestre. Tale teoria fu confutata dai fisici Hess e Kolhörster [1] che compirono una serie di osservazioni con elettroscopi montati su palloni aerostatici. Essi raggiunsero un'altezza di 5 Km nel 1912 e di 9 km nel 1914 e notarono che la ionizzazione media, osservata tramite l'elettroscopio, diminuiva fino ad un'altezza di 1,5 km, e poi cresceva proporzionalmente con l'altezza dal suolo. Era un indizio chiaro che la sorgente ionizzante si trovava fuori dall'atmosfera terrestre. Da qui ebbe inizio la fisica della radiazione cosmica o dei *raggi cosmici* come furono chiamati da Millikan per la prima volta nel 1925 [2].

Nel 1929 si scoprì che la radiazione cosmica è composta da particelle cariche. Negli anni '30, lo studio dei raggi cosmici ha condotto alla scoperta del *positrone* e nel 1936 fu scoperto il *muone*. In quegli stessi anni Pauli teorizzò l'esistenza del *neutrino*.

La scoperta di queste nuove sonde astrofisiche ha permesso di ampliare l'orizzonte dell'universo conosciuto permettendo di progredire nella comprensione della genesi dell'universo e di indagare le possibili sorgenti di radiazione solare, galattiche ed extragalattiche. Lo studio di queste particelle provenienti dallo spazio è anche utile per determinare le caratteristiche della materia interstellare attraversata nel loro percorso verso terra.

1.2 Composizione e spettro energetico dei raggi cosmici

Con il termine raggi cosmici si indica l'insieme di particelle cariche, incidenti sulla sommità dell'atmosfera terrestre, cioè ad un'unità astronomica dal Sole¹.

I raggi cosmici che raggiungono la terra sono in maggioranza di origine galattica. I meccanismi con cui essi vengono generati ed i punti della galassia da cui provengono sono tuttora sconosciuti, tuttavia sono stati elaborati dei modelli per tentare di penetrare questi misteri. Uno di essi ipotizza che le particelle che costituiscono i raggi cosmici galattici o GCR sarebbero frutto della nucleosintesi stellare, successivamente accelerate da onde d'urto provocate da esplosioni di supernovae. Altre teorie affermano che i GCR vengono prodotti da corpi celesti collassati, quali stelle di neutroni e buchi neri, i cui campi elettromagnetici sarebbero in grado di accelerare le suddette particelle. Un dato comune a tutte le teorie è che i raggi cosmici vengono prodotti molto lontano dal sistema solare per cui è estremamente difficile determinarne l'origine. Essi, infatti, durante il lungo cammino che compiono verso la Terra incontrano campi di forze di diversa natura (gravitazionale, elettromagnetica, ecc.) in grado di deviarne la traiettoria facendo così perdere ogni

¹L'unità astronomica (UA) è la distanza media Terra-Sole

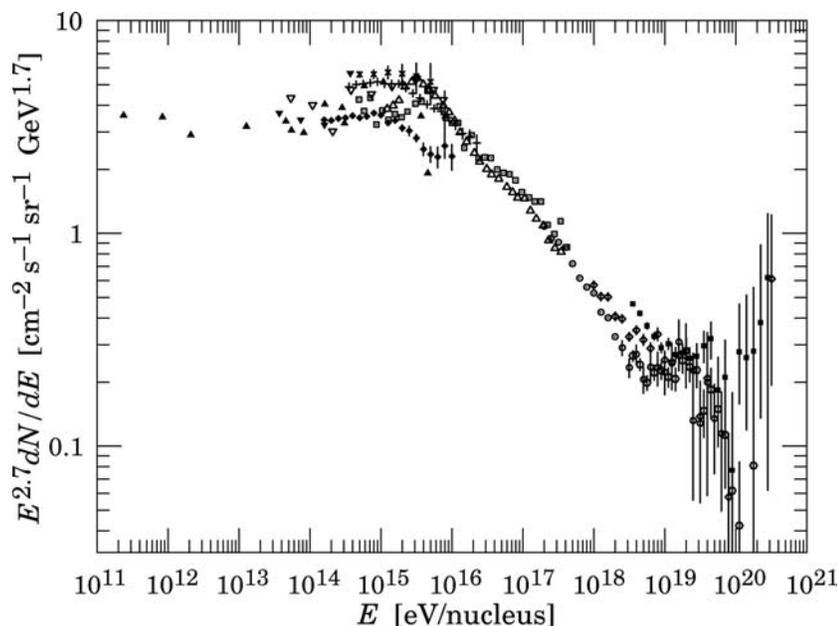


Figura 1.2: Spettro dei raggi cosmici d'alta energia. Si distinguono le cosiddette regioni del ginocchio, tra 10^{15} e 10^{16} eV, e della caviglia tra 10^{18} e 10^{19} eV [4].

informazione riguardo al luogo di generazione e facendo sì che verso la terra giunga un flusso di particelle pressochè isotropo.

Lo spettro energetico dei GCR (figura 1.2) si estende per molti ordini di grandezza da circa 10 GeV fino ad energie dell'ordine dei 100 TeV. Tale spettro può essere descritto dalla seguente legge di potenza:

$$N(E)dE = kE^{-\gamma}dE$$

dove $N(E)$ rappresenta il flusso di particelle che arrivano dall'unità di angolo solido per unità di energia, mentre γ individua l'indice spettrale differenziale. Osservando lo spettro è possibile isolare quattro differenti zone caratterizzate da un diverso valore di γ . La prima, comprendente valori di energie fino a 10^{13} eV, è caratterizzata da valori di γ compresi tra 2.6 e 2.7, la seconda racchiude valori di energia tra i 10^{14} e i 10^{15} eV e rappresenta il ginocchio dello spettro con valori di γ di circa 2.5 [3]. Nella terza parte è possibile notare una brusca variazione di pendenza con valore di γ circa 3 con energie che si estendono fino a 10^{18} eV mentre la quarta zona, detta

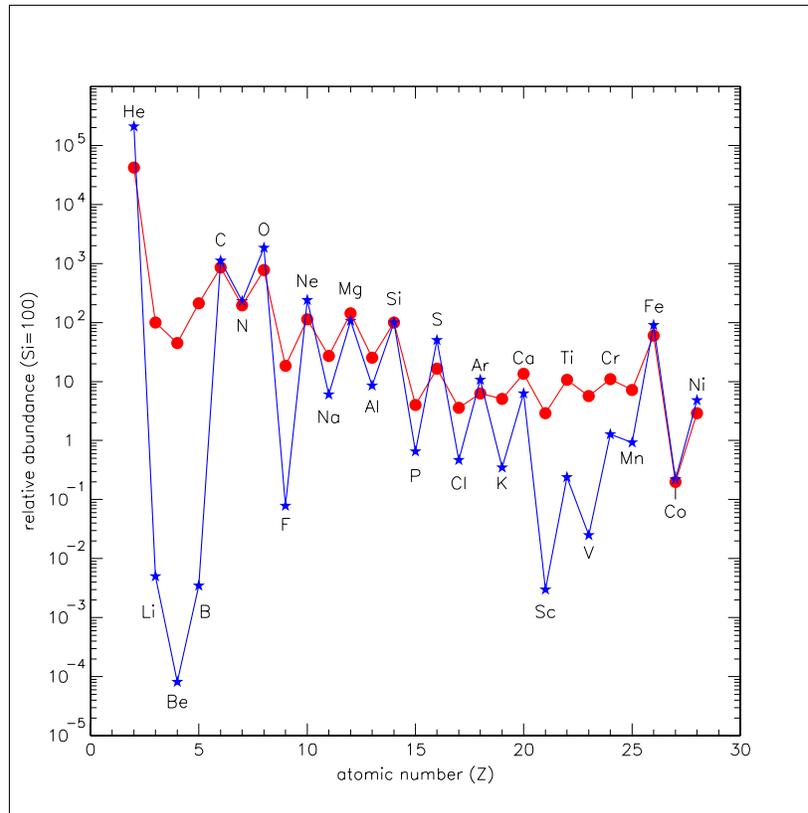


Figura 1.3: Abbondanze relative (misurate a 1 AU) degli elementi con $Z \leq 28$ nei raggi cosmici (pallini) rispetto alle quantità medie del Sistema Solare (stelline). I valori sono relativi al silicio ($\text{Si} = 100$) [7].

anche caviglia dello spettro, ha un indice γ di nuovo pari a 2.5 e racchiude le energie più alte finora misurate (10^{20} eV). Le varie zone che costituiscono lo spettro ed i corrispondenti valori di γ descrivono i diversi meccanismi di accelerazione delle particelle e sono tutti tipici di processi non termici.

I raggi cosmici sono costituiti per il 90% da protoni, per l'1% da elettroni, e per la restante parte da nuclei di 400 specie isotopiche differenti. L'abbondanza chimica nei raggi cosmici fornisce importanti informazioni sull'origine delle particelle che li costituiscono e sui meccanismi di propagazione dalla sorgente fino al pianeta Terra. Dal confronto tra le abbondanze relative nei raggi cosmici con quelle presenti nel sistema solare (figura 1.3) è possibile notare delle sostanziali differenze.

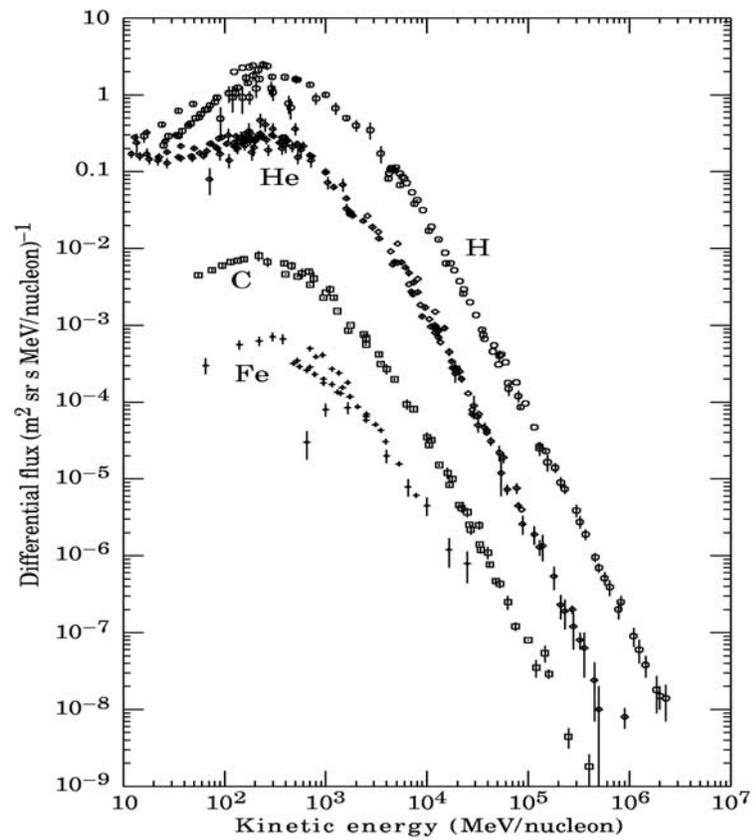


Figura 1.4: Flusso differenziale dei raggi cosmici tra 10 e 10⁷ MeV/nucleone misurato a 1 unità astronomica dal Sole [6].

Per spiegare questa apparente anomalia conviene suddividere le particelle che costituiscono i raggi cosmici in due categorie:

- particelle primarie,
- particelle secondarie.

Le particelle primarie provengono direttamente dalle sorgenti cosmiche mentre le secondarie sono il risultato dell'interazione delle primarie con il mezzo interstellare o con l'atmosfera terrestre. Tale interazione comporta la presenza di specie isotopiche quali ^2H , ^3He , Li, Be, B e altre che non sono prodotte della nucleosintesi stellare ma dall'effetto della frammentazione o *spallation* di nuclei più pesanti quali C, N e O (per la produzione di Li, Be, B) o Fe [5]. La figura 1.4 rappresenta il flusso differenziale di alcuni elementi in funzione della perdita di energia cinetica per nucleone per valori fino a 10^4 GeV/nucleone (prima della zona del ginocchio). Dal grafico è possibile notare che nell'intervallo $10^2 \div 10^3$ MeV/nucleone il flusso differenziale presenta un massimo. La regione dello spettro che si trova ad energie più basse è particolarmente sensibile all'attività del Sole che nei periodi di massima attività modifica il campo magnetico interplanetario responsabile della deviazione delle particelle galattiche e della diminuzione della loro intensità. Il fenomeno è conosciuto come *modulazione solare* dei raggi cosmici.

Lo studio dello spettro energetico e della composizione isotopica, come detto, è molto utile per l'individuazione delle sorgenti delle particelle che costituiscono i GCR; pertanto la misura accurata di tali grandezze risulta utilissima per rendere più precisi i modelli matematici della creazione e propagazione dei raggi cosmici.

1.3 Limiti dell'astronomia con raggi cosmici

Anche se, come detto, l'uso dei raggi cosmici come sonda astrofisica ha consentito e consente di studiare e comprendere molti fenomeni che avvengono nell'universo essi presentano anche una serie di limiti. Un primo inconveniente è che i raggi cosmici che arrivano sulla terra interagiscono con l'atmosfera.

Per studiare raggi cosmici primari è necessario quindi porre rivelatori fuori dall'atmosfera terrestre. Per questo motivo tale studio si può effettuare solo con rivelatori posti su palloni aerostatici e satelliti. Con questi rivelatori è possibile osservare particelle con energia dell'ordine di 10 GeV 100 GeV poichè il flusso di tali particelle è abbastanza elevato. Gli esperimenti su pallone hanno però un limite dovuto al tempo ristretto in cui possono permanere a grandi altezze; ciò comporta un tempo di presa dati limitato e quindi una bassa statistica delle misure. Gli esperimenti su satellite sono molto più efficaci sia perchè sono operativi a quote più elevate sia perchè hanno un tempo di presa dati maggiore. Tuttavia essi hanno problemi dovuti alla realizzazione degli apparati e legati alla necessità di rendere contenute le dimensioni dei rivelatori. I problemi descritti comportano che lo studio dei raggi cosmici primari sia limitato ad energie relativamente basse. Infatti, poichè la probabilità di rivelare particelle con energia superiore a 10^{12} eV (cioè nella zona del ginocchio dello spettro) è dell'ordine di una particella per m^2 per anno per avere misure statisticamente sensate a tali energie si ha bisogno di apparati di grandissime dimensioni. Perciò esperimenti di questo tipo possono essere realizzati solo a terra. In tali esperimenti è necessario tenere conto che le particelle rivelabili sono solo di tipo secondario. Come detto, infatti, i raggi cosmici primari interagendo con i diversi strati dell'atmosfera provocano sciame di particelle adronici ed elettromagnetici studiando i quali è possibile risalire alle particelle che li hanno generati.

I problemi fin qui descritti possono essere considerati di natura tecnologica ma nell'astronomia effettuata con i raggi cosmici si incontrano anche problemi legati alla natura stessa delle particelle. Per quanto riguarda le particelle di bassa energia il problema principale deriva dal fatto che i raggi cosmici sono composti da particelle cariche e quindi la loro traiettoria viene deviata dai campi magnetici che incontrano sul loro percorso. Anche se il campo magnetico galattico ha un'intensità di $\sim 3 \mu G$ ed i campi caotici generati dal moto dei gas ionizzati all'interno di nubi sono anch'essi piuttosto deboli, essi sono estesi per migliaia di anni luce. In queste condizioni le traiettorie dei raggi cosmici sono irregolari e subiscono deflessioni dovute alla forza di Lorentz. Si può facilmente dimostrare che per una particella che ha

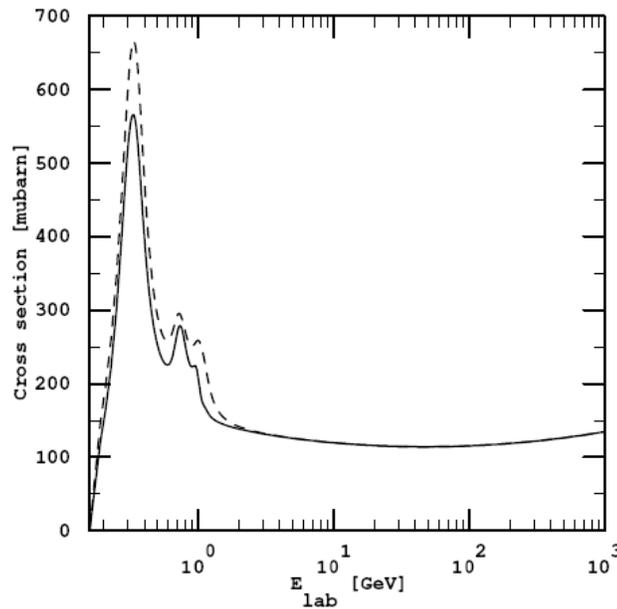


Figura 1.5: Sezione d'urto per interazione $p\gamma$ nel sistema di riferimento del protone (linea continua), la linea tratteggiata indica la sezione d'urto per $n\gamma$.

attraversato anche solo una piccola frazione di galassia non è possibile ricostruirne la direzione di provenienza. Questo effetto comporta che esse non possono essere usate per individuarne le sorgenti. Le particelle che hanno energia elevata ($\sim 10^{18}$ eV zona della caviglia) non subiscono molto l'effetto dei campi magnetici ma il loro flusso è estremamente basso. Tale problema deriva dal cosiddetto effetto GZK.

1.3.1 L'effetto GZK

All'epoca del Big Bang l'Universo era permeato da radiazione γ . Con l'espandersi della materia la distribuzione del fondo dei fotoni è cambiata ed attualmente corrisponde ad uno spettro di corpo nero a temperatura pari a 2.7° K pari cioè ad un'energia media di 10^{-3} eV. Tale radiazione prende il nome di radiazione cosmica di fondo (*Cosmic Microwave Background CMB*).

I protoni di energia maggiore di 50 EeV interagiscono con essa in maniera anelastica producendo pioni (figure 1.5 e 1.6) secondo i processi

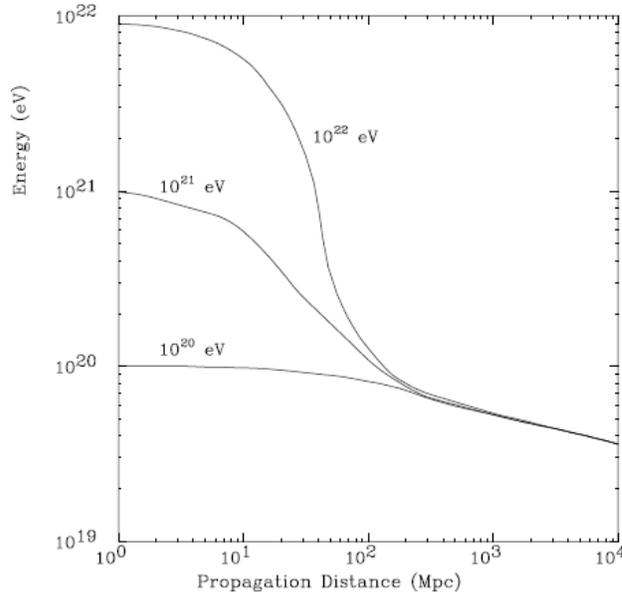
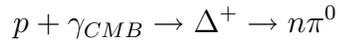


Figura 1.6: Energia di un protone in funzione della distanza di propagazione per diversi valore dell'energia.



Greisen, Zatsepin e Kuzmin [10] [11] predissero che l'interazione dei raggi cosmici con la radiazione cosmica di fondo dovrebbe produrre un taglio (*cut-off*) nello spettro di energia. Un protone con energia pari a $\sim 10^{20}$ eV ha infatti un cammino libero medio nella radiazione di fondo di circa 10 Mpc, una distanza piccola se confrontata con le dimensioni dell'universo. Detto ciò risulta chiaro che il GZK cut-off non è un vero e proprio taglio dello spettro di energia ma una soppressione del flusso di raggi cosmici di altissima energia. Tale effetto fa sì che particelle con energia superiore a 10^{20} eV possono essere rivelate sulla terra solo se sono prodotte o accelerate in un raggio di circa 10 Mpc. Poiché attualmente non sono state osservate, a queste distanze, strutture in grado di produrre accelerazioni di questa intensità è possibile supporre che le particelle osservate con energia superiore al GZK cut-off

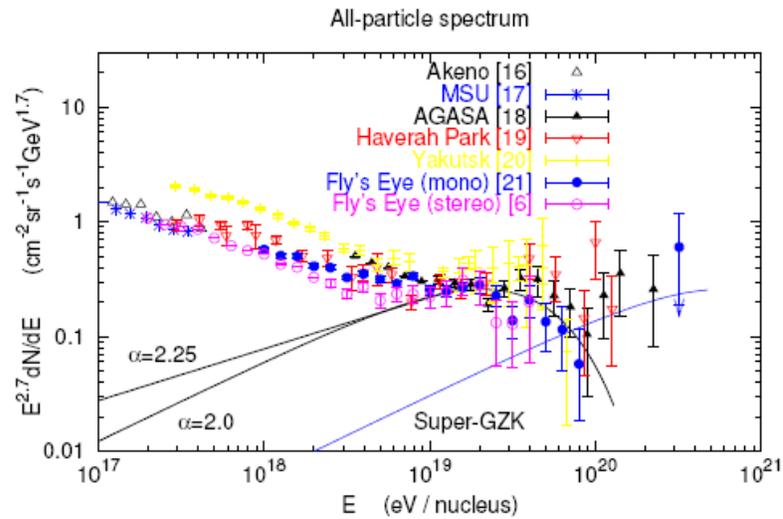


Figura 1.7: Particolare dello spettro dei raggi cosmici nella zona della caviglia.

siano riconducibili al decadimento di particelle supermassive o ad eventi di energia estremamente più grande (figura 1.7).

1.4 L'astronomia con gamma e neutrini

Per quanto fin qui detto, risulta chiaro che l'astrofisica basata sull'osservazione di fotoni di bassa energia e raggi cosmici presenta una serie di limiti dovuti essenzialmente all'impossibilità di studiare eventi di altissima energia che avvengono lontano dalla Terra e all'impossibilità di determinare le sorgenti che producono particelle e radiazione nel cosmo. Queste considerazioni inducono ad individuare delle sonde astrofisiche diverse in grado di superare tali limiti. Analizzando le motivazioni delle difficoltà incontrate viene fuori la necessità di utilizzare come sonde delle particelle che siano poco sensibili ai campi magnetici e allo stesso tempo debolmente interagenti con la materia. In tal modo sarebbe possibile studiare le sorgenti che le producono individuandone la posizione anche se molto distante dal nostro pianeta. Le particelle più utili allo scopo risultano sicuramente i raggi γ e i neutrini (ν). Come vedremo più avanti, infatti, entrambe sono prodotte da sorgenti sia

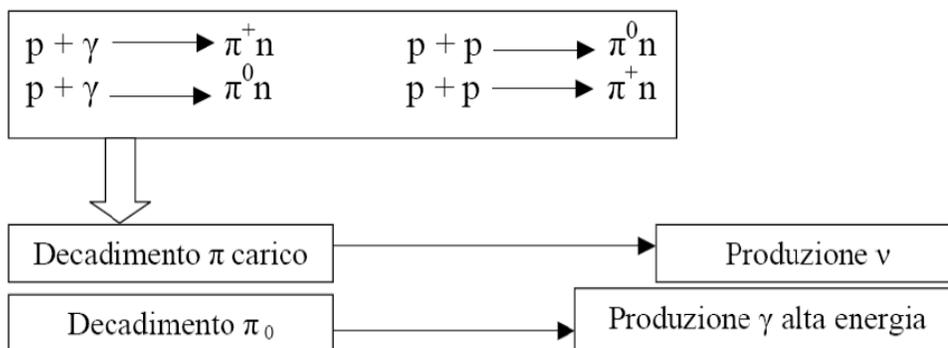


Figura 1.8: Rappresentazione schematica dell'astrophysical beam dump.

galattiche che extra galattiche ed hanno processi di interazioni più deboli con tutto ciò che attraversano durante il percorso verso la Terra.

1.4.1 Astrophysical beam dumps

L'interesse verso lo studio dei raggi γ e dei neutrini è divenuto rilevante quando studi teorici hanno descritto il meccanismo di produzione di tali particelle in prossimità dei siti di produzione dei raggi cosmici. Questo meccanismo, noto sotto il nome di *Astrophysical beam dump*, è simile a quello che si verifica in prossimità dei fasci di particelle studiati negli esperimenti con acceleratori. Il fenomeno descrive la produzione di particelle secondarie prodotte dall'interazione delle primarie con la materia circostante. In ambito astrofisico l'attenzione viene posta in particolare su alcune sorgenti quali, ad esempio, i nuclei galattici attivi e i resti di supernova. Qui la produzione di secondari è favorita sia dalla grande energia a cui vengono accelerati i raggi cosmici primari sia dall'elevata densità della materia circostante il sito di produzione. Le particelle, interagendo con la materia, generano una serie di prodotti secondari tra cui fotoni di altissima energia e neutrini. I meccanismi di produzione più importanti sono l'interazione protone protone (p - p) e protone gamma (p - γ) che portano alla produzione di pioni in tutti gli stati di carica: positivo negativo e neutro. Dal decadimento di tali particelle si ottiene la produzione di neutrini e γ di altissima energia. Lo schema riportato in

figura 1.8 riassume quanto detto fin qui. Tale modello di produzione, precedentemente supportato solo da studi teorici, ha di recente trovato conferma grazie all'esperimento HESS (*High Energy Stereoscopic System*) che studiando la radiazione γ di altissima energia è riuscito ad effettuare una prima caratterizzazione dei siti di produzione confermando di fatto l'esistenza dei meccanismi descritti in precedenza. Questo studio risulta tuttavia preliminare e sarebbe molto interessante approfondire la correlazione esistente tra i flussi di neutrini e di γ .

1.4.2 Propagazione γ e ν nell'universo

Per quanto detto fin ora appare chiaro che i raggi γ e i neutrini risultano essere ottimi candidati come nuove sonde astrofisiche poichè essi vengono prodotti da sorgenti molto interessanti dal punto di vista astrofisico. Per poter essere utilizzate per lo studio di tali sorgenti occorre tuttavia valutare come queste particelle si muovono all'interno dell'universo nel loro tragitto verso la Terra. Come visto precedentemente è fondamentale che esse si propaghino per quanto più possibile liberamente su distanze intergalattiche e che subiscano il meno possibile effetti dovuti alla forza di Lorentz e ad altre possibili interazioni.

Una caratteristica che accomuna i fotoni γ ed i neutrini è che entrambe le particelle hanno carica neutra. Ciò comporta che esse non interagiscono con i campi magnetici che incontrano sul loro cammino. L'osservazione di tali particelle rende possibile determinarne la direzione di provenienza ed individuare la sorgente di produzione.

Anche se dal punto di vista delle carica elettrica i γ ed i neutrini sono del tutto identici bisogna considerare un aspetto che mette in luce i limiti dell'astronomia γ a discapito di quella basata sullo studio dei neutrini. I fotoni γ , infatti, interagiscono con la radiazione cosmica di fondo e quindi non riescono a compiere grandi distanze. Come detto, in tutto l'universo è presente una radiazione assimilabile a quella di corpo nero a circa 3° K che costituisce la principale causa di assorbimento dei raggi cosmici che attraversano l'universo. Anche i fotoni di altissima energia interagiscono con tale

radiazione secondo la relazione

$$\gamma + \gamma_{(3^{\circ}K)} = e^+e^-$$

La reazione appena descritta è un noto processo di interazione chiamato produzione di coppie o pair production. Tale processo permette di determinare la distanza massima di propagazione di un γ di alta energia nell'universo. Un fotone con energia di un PeV (10^{15} eV) riesce a percorrere una distanza media di circa 10 Kpc. Tale distanza è circa uguale alla dimensione media della nostra galassia. Questa considerazione suggerisce che l'astronomia γ è utile allo studio solo di fenomeni galattici ma è sostanzialmente inutilizzabile per lo studio di quelli extragalattici.

Come vedremo in dettaglio più avanti i neutrini, da questo punto di vista, risultano estremamente favoriti. Essi infatti interagiscono solo debolmente con la materia ed hanno una sezione d'urto estremamente bassa. Ciò comporta che essi possano viaggiare praticamente indisturbati nell'universo percorrendo grandissime distanze attraversando campi magnetici anche di notevole intensità e zone molto dense di materia senza di fatto subire interazioni o venire deviati. Queste considerazioni rendono i neutrini una sonda astrofisica ideale e giustificano i grandi sforzi necessari a costruire telescopi in grado di rivelarli.

Capitolo 2

La neutrino-astronomia

2.1 Introduzione

Nel capitolo precedente sono stati introdotti i concetti fondamentali dell'astronomia basata sui raggi cosmici e sono stati evidenziati i vantaggi che lo studio di tali particelle ha rispetto all'osservazione dei fotoni. L'astronomia dei raggi cosmici e delle particelle in generale, si basa sulla possibilità di ricostruire, a partire dalla misura della radiazione, le caratteristiche e la posizione dei corpi celesti in modo da produrre come risultato una mappa del cielo. In questo tipo di astronomia due aspetti risultano fondamentali per ottenere risultati: il **puntamento direzionale** e l'**orizzonte di osservabilità**. Il puntamento direzionale è la proprietà per cui la misura della direzione della radiazione cosmica incidente permette di individuare direttamente la posizione della sorgente. E' fondamentale quindi, per questo tipo di osservazione, che la particella in analisi non subisca significative deflessioni nella propagazione attraverso il mezzo interstellare. L'orizzonte di osservabilità definisce la massima distanza da cui è possibile trarre informazioni. Tale distanza è determinata da tutti i meccanismi di assorbimento che interessano la particella utilizzata come sonda astrofisica durante la sua propagazione. I fotoni contenuti nella componente primaria dei raggi cosmici soddisfano bene le condizioni di puntamento direzionale poichè tali particelle, essendo prive di carica elettrica non vengono deviate dai campi magnetici. L'orizzonte di osservabilità risulta invece fortemente limitato da una serie di interazioni che

i fotoni subiscono. In particolare, per quanto riguarda la radiazione γ di alta energia, l'interazione con la radiazione cosmica di fondo, e la conseguente formazione di coppie, limita il cammino di un fotone a distanze dell'ordine di 10 Mpc (effetto GZK).

La componente carica dei raggi cosmici subisce deflessioni per azione dei campi magnetici galattici e questo ne limita notevolmente le proprietà di puntamento direzionale della sorgente. L'effetto GZK riduce inoltre le distanze medie di propagazione limitandone l'impiego allo studio di fenomeni galattici.

Da questo tipo di analisi risulta chiaro che la particella ideale da utilizzare come sonda astrofisica è quella in grado di trasportare, a qualunque energia, l'informazione della sorgente che l'ha emessa. Per fare ciò deve avere essenzialmente due caratteristiche:

- deve essere neutra e con massa piccolissima o nulla, in modo da non subire deviazioni da parte dei campi gravitazionali e magnetici soprattutto a basse energie;
- deve essere debolmente interagente altrimenti può dare luogo, durante la propagazione, a processi che comportano il suo assorbimento.

La particella che possiede entrambe le caratteristiche è il neutrino che grazie alle sue proprietà può essere considerato un'ottima sonda astrofisica. Chiaramente è anche fondamentale che la particella sia prodotta da la gran parte di sistemi astrofisici poichè tale caratteristica definisce i tipi di sistemi che possono essere osservati per mezzo di essa. Risulta inoltre molto importante conoscere i meccanismi con i quali i suddetti sistemi astrofisici siano in grado di produrre neutrini. Per realizzare un telescopio basato su questo tipo di particelle è essenziale poi conoscere approfonditamente le caratteristiche della particella e quindi essere al corrente dei meccanismi di interazione con la materia. Tutte queste informazioni serviranno sia a capire le tecniche da impiegare per la rivelazione delle particelle sia a risalire alle sorgenti che le hanno generate. Il problema della rivelazione dei neutrini è particolarmente ostico poichè le caratteristiche che le permettono di

viaggiare praticamente indisturbata per grandissime distanze nell'universo comportano numerose difficoltà di rivelazione e, conseguentemente, rendono tecnologicamente complessi i telescopi atti alla sua osservazione.

In questo capitolo sarà innanzitutto effettuata una breve rassegna delle principali sorgenti astrofisiche che emettono neutrini esplorando anche i meccanismi di produzione degli stessi. Si procederà poi all'esposizione delle principali tecniche di rivelazione dei neutrini a terra passando poi in rassegna i telescopi attualmente realizzati od in via di realizzazione per lo studio di tali particelle.

2.2 Produzione di neutrini in ambienti astrofisici

I neutrini sono generalmente particelle secondarie, vengono cioè prodotte dall'interazione delle particelle emesse da una sorgente astrofisica. I principali meccanismi di produzione derivano da particelle primarie quali il protone (p) e il neutrone (n) che, emesse e accelerate dalla sorgente astrofisica, interagiscono con la materia circostante producendo neutrini. Il principale canale d'interazione per un generico nucleone N (protone o neutrone) con un fotone è quello che porta alla produzione di un singolo pione:

$$\gamma + N \rightarrow N + \pi$$

a seconda della carica del nucleone si hanno i seguenti processi:

$$p + \gamma \rightarrow n + \pi^+; \quad n + \gamma \rightarrow p + \pi^-$$

$$p + \gamma \rightarrow p + \pi^0; \quad n + \gamma \rightarrow n + \pi^0$$

Dal decadimento dei pioni vengono successivamente prodotti neutrini secondo i seguenti processi:

$$\pi^+ \rightarrow \mu^+ + \nu_\mu; \quad \mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu$$

$$\pi^- \rightarrow \mu^- + \bar{\nu}_\mu; \quad \mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu$$

L'energia persa da un protone in una singola collisione è $\Delta E_p/E_p \sim 0.2$. Quest'energia è statisticamente suddivisa tra i prodotti finali delle catene di decadimento ed in particolare ogni neutrino ha il 5% dell'energia del protone interagente.

2.3 Sorgenti galattiche di neutrini

2.3.1 Il piano galattico e il Sole

Il Sole è una ben nota sorgente di neutrini di bassa energia[14] ma si pensa che da esso possano provenire anche neutrini di alta energia. Un possibile meccanismo di produzione di tali particelle prevede l'interazione di un raggio cosmico di altissima energia (*Ultra High Energy Cosmic Ray* UHECR) con le nuvole di gas che formano l'atmosfera solare [15]. Un'altra fonte di interazione dei UHECR che porta alla produzione di neutrini di alta energia è rappresentata dalla materia che forma il piano galattico e più in particolare le dense nuvole di gas presenti nei bracci a spirale della nostra galassia. Le sorgenti descritte potrebbero essere considerate come puntiformi e quindi i neutrini prodotti in questo modo dovrebbero essere facilmente distinguibili dal fondo di neutrini atmosferici (quelle particelle prodotte dall'interazione dei raggi cosmici con la nostra atmosfera).

2.3.2 Resti di Supernova

La stabilità di una stella è data dall'equilibrio tra la forza gravitazionale e la pressione di radiazione prodotta dall'energia rilasciata dalle reazioni nucleari che avvengono all'interno della stella stessa. Per le stelle con massa maggiore di $5 M_{\odot}$ giunte al ciclo del ferro, la pressione di radiazione e quella di degenerazione degli elettroni non riescono a sostenere la struttura della stella che collassa su se stessa. Il collasso del nucleo si arresta quando, raggiunta la densità critica, l'interazione nucleare forte genera un'onda d'urto che si propaga verso gli strati esterni della stella producendo l'esplosione di Supernova. L'onda d'urto si propaga per molto tempo nello spazio interstellare producendo una struttura nota come resto di Supernova (*Supernova*

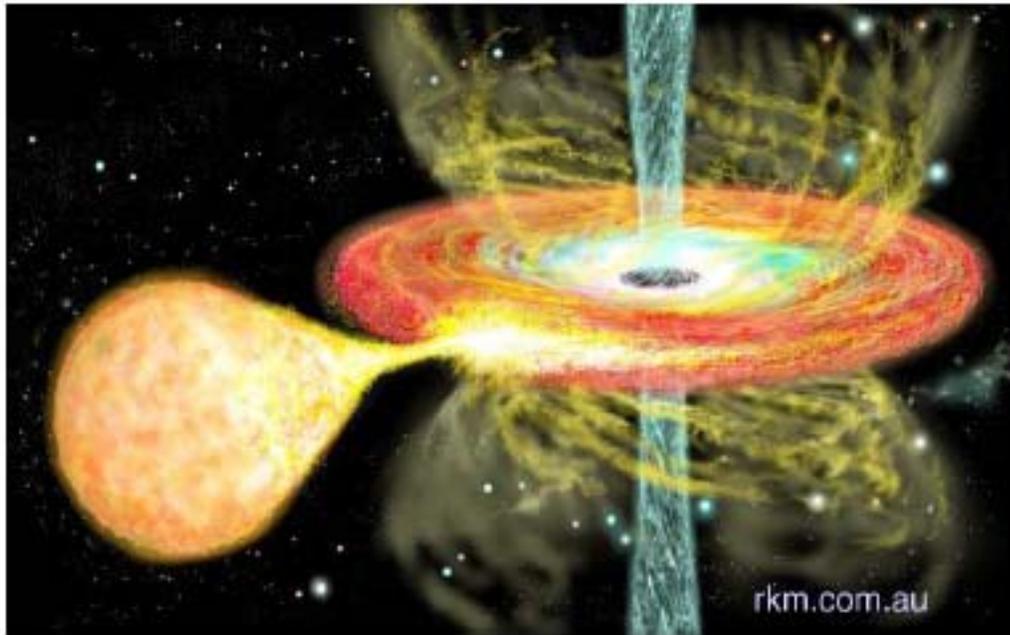


Figura 2.1: Meccanismo di formazione di una Microquasar.

Remnants SNR). Quest'onda d'urto, in combinazione con i campi magnetici, può favorire l'accelerazione di adroni che successivamente interagiscono con la materia circostante. In tali interazioni, di tipo pp o $p\gamma$, possono essere generati neutrini di alta energia.

2.3.3 Microquasar e i sistemi binari X

I sistemi binari X sono tra gli oggetti più luminosi mai osservati nella nostra galassia. Essi sono formati da un sistema di due corpi: un Buco Nero (o una Stella di Neutroni) e una seconda stella che trovandosi molto vicina alla prima fornisce ad essa materia (figura 2.1). Il Buco Nero risucchiando la materia della stella forma un disco di accrescimento molto esteso. In esso gli elevati campi magnetici ed il plasma denso creano un ambiente dal quale si sprigionano intense onde d'urto che accelerano protoni e portano alla produzione di γ e neutrini. Un caso particolare di sistemi binari X sono le Microquasar (μ QSO). Esse emettono jet di plasma visibili nella banda radio

dello spettro elettromagnetico. Secondo il modello previsto da Levinson e Waxman [10] all'interno del jet si creano delle *shell* in cui il plasma produce una serie di onde d'urto. Grazie a queste onde di shock i protoni vengono accelerati secondo il meccanismo di Fermi sino ad energie di 10^{16} eV ed interagiscono con i fotoni di sincrotrone emessi dagli elettroni accelerati. Questa interazione, come visto in precedenza, produce pioni che decadendo producono a loro volta neutrini con energia compresa tra 1 e 100 TeV. I flussi di neutrini attesi per questa sorgente su un Km^2 a terra vanno da uno a qualche decina di eventi l'anno.

2.4 Sorgenti extragalattiche di neutrini

2.4.1 Nuclei Galattici Attivi (AGN)

I Nuclei Galattici Attivi sono stimati essere la sorgenti continue più luminose presenti nell'universo. In realtà con la sigla AGN sono descritti una serie di oggetti celesti che hanno in comune uno spettro piatto di emissione radio, un'emissione continua non termica, ed una forte dipendenza dal tempo dell'emissione di energia. Diversi modelli teorici prevedono che il nucleo di un AGN sia costituito da un Buco Nero supermassivo ($10^6 \div 10^9 M_{\odot}$) in grado di attrarre una grande quantità di polveri e materia interstellare per formare un grande disco di accrescimento. L'equilibrio dinamico tra la pressione di radiazione uscente dal sistema e la materia entrante in essa permette all'AGN di non collassare. la grande quantità di materia risucchiata dal Buco Nero produce un'energia tale da accelerare le particelle fino a permetterle di vincere la grande forza di attrazione. Successivamente esse vengono espulse in direzione perpendicolare rispetto al disco di accrescimento poichè in questa zona la densità di materia risulta minore.

Vengono così prodotti due jet di materia e radiazione (figura 2.2) in grado di accelerare particelle fino ad energie pari a 10^{20} eV e produrre neutrini di alta energia. Il flusso atteso da questo tipo di sorgenti dovrebbe aggirarsi attorno a qualche neutrino per anno per Km^2 .

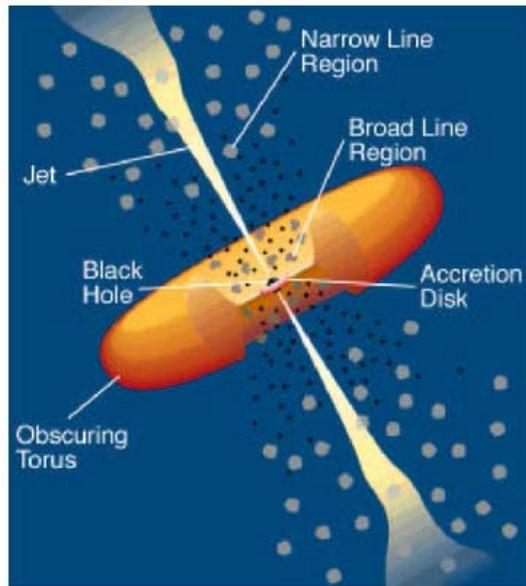


Figura 2.2: Schema rappresentativo di un AGN.

2.4.2 Gamma Ray Burst (GRB)

I Gamma Ray Burst (o lampi di raggi gamma) furono osservati per la prima volta dal satellite VELA e, con un'energia rilasciata pari a 10^{52} erg s^{-1} , tuttora rimangono il fenomeno più violento mai osservato nell'universo. Lo spettro di emissione di questi eventi si estende dai gamma di altissima energia fino al radio ed una durata massima che varia da un secondo fino ad una decina di secondi. Si ritiene che questi eventi, piuttosto rari, siano causati dalla collisione di un Buco Nero compatto e da una Stella di Neutroni oppure da un sistema binario di due Buchi Neri compatti. La collisione formerebbe una sorgente di grande massa molto calda chiamata *fireball*. Secondo i modelli teorici nei primi istanti dopo la formazione la *fireball* è otticamente opaca ed al suo interno si formano una serie di onde d'urto in grado di accelerare protoni secondo il meccanismo di Fermi. Con il passare del tempo la *fireball* si estende e diviene improvvisamente trasparente ai gamma. In questo istante viene rilasciato un'enorme flusso di radiazione che investe il mezzo interstellare freddo. Il riscaldamento improvviso del mezzo genera onde di

shock termico che causano il cosiddetto *afterglow*. I protoni accelerati in questo modo interagiscono con i fotoni contenuti nel *fireball* e producono pioni dal cui decadimento vengono generati neutrini con energie comprese tra 10^{14} e 10^{16} eV. Il numero di eventi atteso dai GRB dovrebbe aggirarsi attorno ai 25 eventi l'anno per Km^2 . L'osservazione di questi eventi potrebbe, inoltre, essere favorita dall'osservazione combinata e dalla correlazione spaziale e temporale di dati raccolti con telescopi per neutrini e telescopi gamma.

2.5 Altre sorgenti

2.5.1 Materia Oscura e WIMP

La grande discrepanza tra la velocità calcolata e quella misurata dei cluster stellari può essere spiegata ipotizzando che nell'universo ci sia una enorme quantità di materia che non riesce ad essere rivelata poichè da essa non proviene luce. Questo tipo di materia è detta materia oscura. Modelli teorici tendono a suddividere la materia oscura in due categorie: la materia oscura calda (*Hot Dark Matter* HDM), caratterizzata da piccola massa e grande energia cinetica, e la materia oscura fredda (*Cold Dark Matter* CDM) avente grande massa e bassa energia cinetica. Recenti studi hanno messo in evidenza un'anisotropia della radiazione cosmica di fondo che potrebbe essere spiegata con un'abbondanza di materia oscura fredda. Uno dei costituenti della CDM potrebbero essere le *WIMPs* (*Weakly Interactive Massive Particles*) che, essendo particelle stabili potrebbero costituire la materia mancante dell'universo.

Una delle particelle candidate a far parte della materia oscura fredda è il *neutralino* χ che, secondo la teoria della supersimmetria, sarebbe la particella supersimmetrica più leggera con una massa superiore a 50 GeV.

I neutralini potrebbero essere gravitazionalmente accumulati nell'alone galattico, al centro della Terra o del Sole, oppure al centro della galassia. Tali particelle potrebbero dar luogo a fenomeni di annichilazione di tipo $\chi\chi$ che porterebbero ad una catena di decadimenti fino alla produzione di un flusso di neutrini di alta energia. Gli effetti descritti comporterebbero un eccesso

di neutrini proveniente da questi grandi corpi celesti che potrebbero essere osservati con telescopi con area sufficientemente grande e buona risoluzione angolare.

2.5.2 Difetti Topologici

All'epoca della rottura della simmetria prevista nella teoria di Grande Unificazione (*Grand Unification Theory* GUT) potrebbero essere stati generati difetti topologici (*Topological Defects* TD) quali stringhe e monopoli magnetici con masse dell'ordine della scala della teoria di Grande Unificazione. Particelle estremamente massive prodotte in quell'epoca avrebbero potuto giocare un ruolo fondamentale come punti di accumulazione per la formazione di strutture a larga scala. I modelli attuali prevedono che dall'interazione tra difetti topologici (come ad esempio la sovrapposizione di stringhe) possano essere prodotte particelle X con masse prossime all'energia della scala della GUT. Queste particelle decadendo a loro volta, porterebbero alla generazione di protoni, fotoni e poi neutrini di altissima energia.

2.6 Rivelazione dei neutrini di alta energia

L'interesse per l'osservazione e lo studio dei neutrini astrofisici risale agli anni '60 quando, per la prima volta, fu presentato alla comunità scientifica internazionale un progetto che prevedeva la realizzazione di apparati per la rivelazione di queste particelle [11]. Tali apparati, poichè avrebbero consentito l'identificazione delle sorgenti astrofisiche di neutrini furono chiamati *telescopi per neutrini*. La proposta prevedeva di utilizzare una serie di sensori ottici ed acustici al fine di ricostruire la direzione di un neutrino grazie al tracciamento dei leptoni carichi prodotti dall'interazione del neutrino stesso con la materia.

Attualmente la tecnica di rivelazione più utilizzata e più efficace è quella di realizzare un rivelatore di radiazione Čerenkov. Tale tecnica, come vedremo in seguito, consiste nell'attrezzare con una serie di sensori ottici grandi volumi di acqua o ghiaccio utili a rivelare la luce Čerenkov prodotta dai muoni generati dai neutrini emessi nel mezzo stesso. Date le sue caratteristiche un

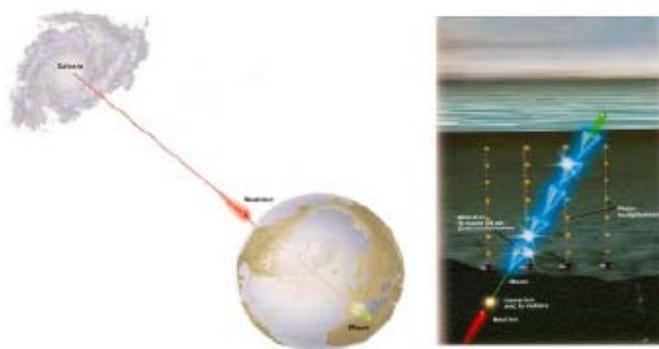


Figura 2.3: I rivelatori di luce Čerenkov come telescopi sottomarini per neutrini.

rivelatore di questo tipo viene normalmente installato nelle grandi profondità marine o sotto i ghiacci dei poli. La ricostruzione del fronte d'onda di luce Čerenkov permette di determinare la direzione di propagazione del muone da cui è possibile ricostruire la direzione del neutrino incidente. Per poter rivelare efficacemente il muone, il mezzo in cui si propaga la traccia deve avere delle ben definite caratteristiche ottiche:

- la *lunghezza di assorbimento* (L_a) deve essere sufficientemente grande in modo da accrescere la distanza che intercorre tra due moduli ottici vicini. In definitiva, a parità di sensori ottici, è possibile avere un rivelatore con volume tanto più grande quanto più grande è la lunghezza di assorbimento del mezzo utilizzato;
- la *lunghezza di scattering* deve essere il più possibile elevata in modo da lasciare invariata la direzione di propagazione della luce Čerenkov e quindi ottenere una buona ricostruzione della traccia rivelata.

Proprietà ottiche significativamente diverse caratterizzano l'acqua ed il ghiaccio. In particolare, le caratteristiche di questi due mezzi risultano in qualche maniera complementari. Il ghiaccio, infatti, possiede una maggiore lunghezza di assorbimento rispetto all'acqua ma ha una minore lunghezza di scattering. Pertanto, a parità di sensori ottici, un rivelatore che utilizzi il

ghiaccio avrà una maggiore area efficace rispetto ad uno immerso in acqua ma presenterà una minore risoluzione angolare.

I rivelatori di luce Čerenkov possono essere utilizzati sia come telescopi traccianti, essendo in grado di determinare la direzione di provenienza delle particelle, sia come calorimetri quando sono in grado di misurarne l'energia. Per poter misurare in maniera diretta l'energia delle particelle è necessario che esse vengano generate e fermate all'interno del rivelatore. I telescopi per neutrini attualmente in costruzione possono effettuare una misura diretta dell'energia di particelle dell'ordine di 300 GeV. Ad energie più elevate, la particella fuoriesce da rivelatore e quindi non rilascia al suo interno tutta l'energia. Ad ogni modo è possibile effettuare una stima indiretta dell'energia misurando la quantità di luce Čerenkov emessa per unità di lunghezza. Stime di questo tipo sono comunque affette da errori rilevanti. Risulta comunque sempre possibile ottenere informazioni sulla direzione di propagazione delle particelle con una precisione che aumenta all'aumentare dell'energia della particella in analisi.

Come descritto nei precedenti paragrafi, i flussi di neutrini attesi indicano che i telescopi devono avere un'area efficace dell'ordine di 1 Km² affinché gli studi effettuati siano statisticamente rilevanti. Tali telescopi devono, inoltre, avere un volume sufficientemente grande da tracciare il muone che lo attraversa. Per ricostruire un evento in maniera qualitativamente accettabile è necessario partire da una traccia della lunghezza di almeno 1 Km. Dalle considerazioni fin qui fatte risulta chiaro che un telescopio per neutrini astrofisici deve avere un volume non inferiore ad 1 Km³.

Un telescopio di questo tipo dovrebbe essere il più possibile isotropo in modo da essere sensibile ad eventi provenienti da direzioni differenti. Tuttavia è necessario ricordare che le tracce interessanti sono quelle prodotte dai muoni generati dai neutrini astrofisici e che la probabilità che un neutrino abbia un'interazione che porti alla produzione di un muone è tanto più elevata quanto maggiore è lo strato e la densità di materia che esso attraversa. Come vedremo più approfonditamente in seguito, il neutrino ha una sezione d'urto molto bassa e quindi, per fare in modo che esso interagisca, è necessario che attraversi grandi strati di roccia. Per questo motivo è molto

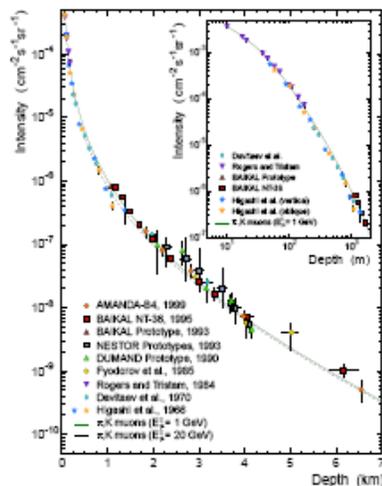


Figura 2.4: Flusso di muoni atmosferici in funzione della profondità marina.

più probabile risalire ad una traccia prodotta da un neutrino astrofisico studiando un muone che provenga dal fondo del mare e si sviluppi verso l'alto. In questo modo si è ragionevolmente certi che il neutrino abbia prodotto il muone osservato attraversando strati di roccia dell'ordine del diametro terrestre. Da queste considerazioni risulta chiaro che il telescopio debba essere più efficiente nell'osservare eventi che si propaghino dal basso verso l'alto. Questo tipo di architettura, inoltre, contribuisce ad eliminare le tracce prodotte da muoni generati dall'interazione dei raggi cosmici con l'atmosfera e quindi migliorare il rapporto segnale rumore. Il flusso di tali particelle, che costituiscono rumore per il telescopio, è circa 10^{11} volte più intenso di quello atteso per i neutrini astrofisici ed è quindi necessario trovare un metodo efficace per escludere questo tipo di eventi. Per schermare tale contributo si può partire dalla considerazione che il flusso di muoni atmosferici diminuisce sensibilmente all'aumentare della quantità di acqua attraversata [18].

Come si evince dal grafico 2.4 a profondità di circa 4000 m il detector viene raggiunto da un flusso di particelle solo 6,5 volte più intenso di quello atteso per i muoni generati da interazione di neutrino; ponendo quindi il telescopio a queste profondità e orientando i sensori verso il basso si stima di aumentare l'efficienza del rivelatore.

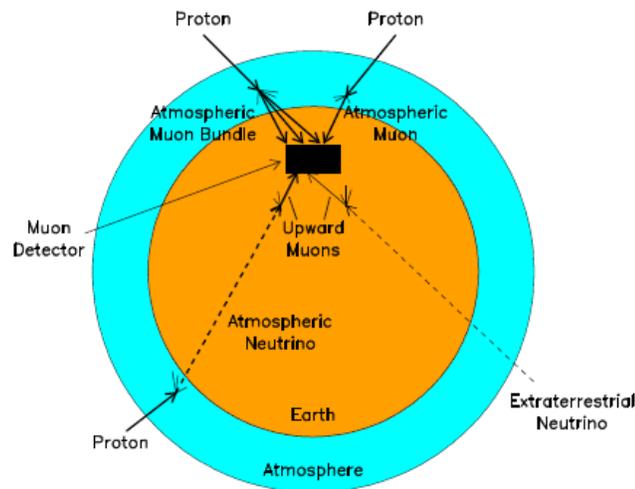


Figura 2.5: Schema di principio che mostra la formazione dei muoni atmosferici e l'interazione dei neutrini con la Terra.

In definitiva lo strato di acqua o ghiaccio che sovrasta il rivelatore svolge una triplice funzione:

- costituisce uno schermo per il flusso di muoni che provengono dall'alto facendo aumentare il rapporto segnale rumore del telescopio;
- serve da bersaglio per i neutrini che interagendo producono muoni;
- costituisce un radiatore trasparente dove la luce Čerenkov può essere prodotta e propagata.

In figura 2.5 viene rappresentato, con una semplice illustrazione, il principio di funzionamento del telescopio e di come i neutrini, a differenza dei muoni, riescono ad attraversare la Terra giungendo fino al rivelatore. In figura 2.5 viene anche rappresentato il meccanismo di produzione dei neutrini atmosferici a partire dai raggi cosmici.

2.6.1 Interazione dei neutrini

Come detto, i neutrini sono particelle prive di carica e dotate di massa trascurabile. Potendo interagire solo debolmente, essi sono in grado di percorrere

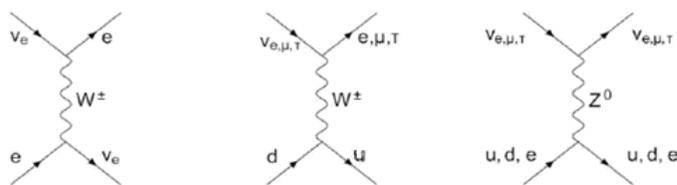


Figura 2.6: Diagrammi ad albero che descrivono schematicamente l'interazione di corrente neutra e corrente carica del neutrino.

grandi distanze senza essere assorbiti. Quando però raggiungono un oggetto denso come la Terra la probabilità d'interazione aumenta. Le interazioni possibili sono di due tipi:

- interazione di corrente carica CC: $\nu_l + N \rightarrow l + X$;
- interazione di corrente neutra NC: $\nu_l + N \rightarrow \nu_l + X$.

Nelle formule l indica uno dei tre leptoni (e, μ, τ), N un nucleone e X è una qualsiasi particella si possa originare. Il processo di interesse per un telescopio Čerenkov è l'interazione di corrente carica su nucleone poichè da essa viene generato un leptone della stessa famiglia che, propagandosi in un mezzo trasparente emette luce Čerenkov rivelabile.

Per energie del neutrino dell'ordine di qualche decina di GeV l'interazione di corrente carica dominante è il *Deep Inelastic Scattering* su nucleone che ha una sezione d'urto $\sigma \sim 10^{-36} \text{cm}^2$ per E_ν dell'ordine del TeV. Ad energie minori vi sono anche altri processi che contribuiscono quali lo scattering quasi-elastico e le interazioni risonanti. Tali interazioni possono essere trascurate per un telescopio di un Km^3 la cui soglia di rivelazione è dell'ordine del TeV.

La figura 2.7 mostra l'andamento della sezione d'urto in funzione dell'energia nell'intervallo $10^{10} \div 10^{21}$ eV per il processo d'interazione neutrino-nucleone. Nel grafico vengono rappresentati con curve distinte i processi di corrente neutra (NC) e quelli di corrente carica (CC). Poichè la sezione d'urto dipende dalla massa del bersaglio, le interazioni tra neutrini e quark del

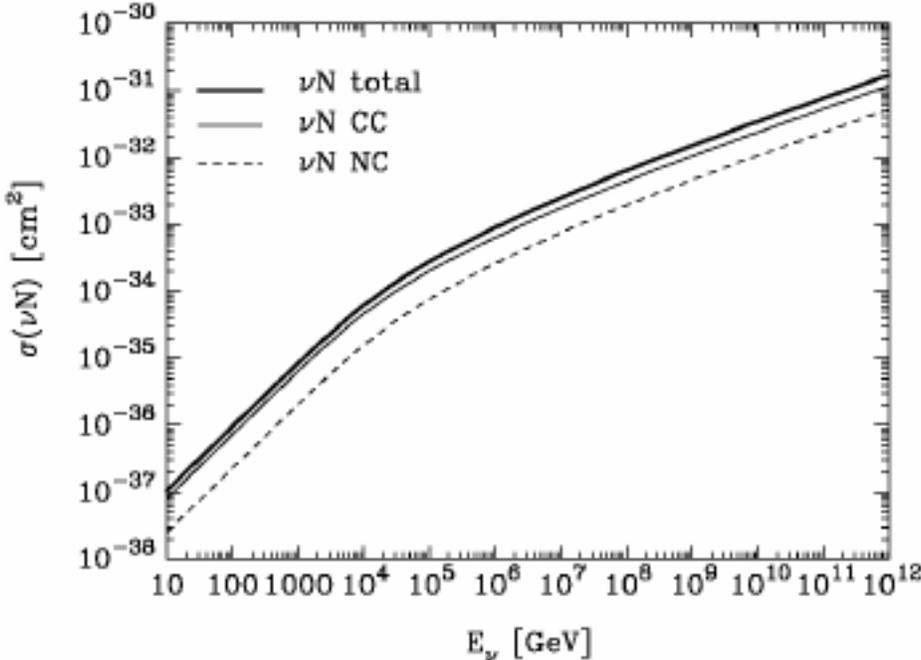


Figura 2.7: Diagramma della sezione d’urto dei neutrini in funzione dell’energia.

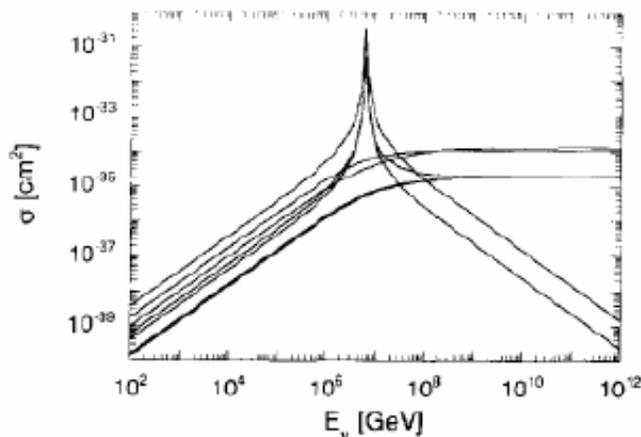


Figura 2.8: Sezione d'urto di interazione neutrino-elettrone in funzione dell'energia del neutrino. Le diverse curve distinguono le diverse interazioni con l'elettrone.

mezzo attraversato, sono estremamente più probabili rispetto a quelle tra neutrini ed elettroni; l'interazione neutrino-elettrone risulta pertanto un processo più raro. L'eccezione a quanto detto è rappresentata dalla formazione risonante del bosone intermedio W nei processi d'interazione anti-neutrino elettronico su elettrone con un'energia di soglia $E^{th} \sim 6.3$ PeV. A queste energie la sezione d'urto di risonanza risulta maggiore della sezione d'urto d'interazione neutrino-nucleone (figura 2.8).

2.6.2 Possibilità di osservazione di differenti *flavour* di neutrini

I telescopi sottomarini per neutrini rivolgono la loro attenzione principalmente alla ricerca del neutrino muonico ν_μ , dal momento che i muoni sono i leptoni più facilmente individuabili tra i prodotti d'interazione descritti in precedenza. Le interazioni di corrente carica dei ν_μ producono μ^\pm che ad energie elevate lasciano lunghe tracce nell'apparato e producono cascate adroniche localizzate. I neutrini elettronici ν_e danno origine a sciame adronici localizzati e a sciame elettromagnetici con dimensioni longitudinali di pochi metri, essendo la lunghezza di radiazione in acqua per gli elettroni circa 36

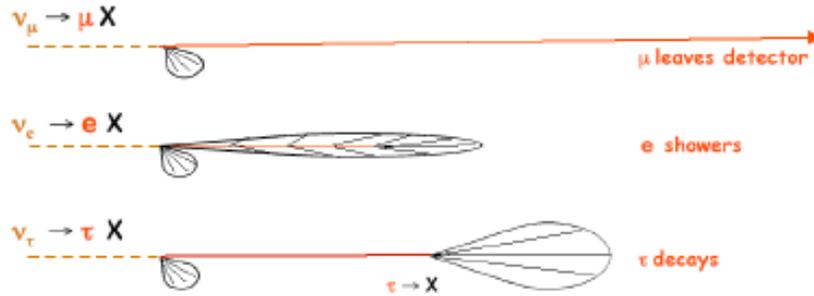


Figura 2.9: Interazione dei tre flavour di neutrini.

cm. Questi eventi sono quindi visti dal telescopio come tracce puntiformi.

I telescopi sottomarini possono anche studiare eventuali fenomeni di oscillazione di sapore di neutrini attraverso l'osservazione del neutrino tauonico ν_τ . Se si verificasse oscillazione di sapore del neutrino circa la metà dei ν_μ ultra-energetici emessi da sorgenti astrofisiche si dovrebbe trasformare in ν_τ durante il percorso verso la Terra. Attraversandola, i ν_τ producono leptoni τ attraverso interazioni di corrente carica. I leptoni τ sono particelle con vita media molto breve ($t_\tau = 2.9 \times 10^{-13} s$) e decadendo producono ν_τ (35%), ν_μ e ν_e dando luogo ad un processo di rigenerazione di neutrini di tutti i *flavours*. Dunque il contributo dei ν_τ potrebbe manifestarsi come un eccesso di particelle che riescono ad attraversare la Terra ad energie per cui questa dovrebbe essere opaca. I diversi tipi di interazione sono illustrati nella figura 2.9.

2.7 Propagazione ed interazione dei muoni

2.7.1 Perdite di energia dei muoni nella materia

La perdita di energia dei muoni nella materia può essere espressa come somma di contributi differenti: ionizzazione, bremsstrahlung, produzione di coppie e^+e^- , interazione fotonucleare (con conseguente produzione di adroni) e altri. Solo una minima parte di energia viene persa per emissione di radiazione Čerenkov.

Per muoni con energia inferiore a qualche centinaio di GeV le perdite radiative sono talmente piccole da poter essere trascurate e la perdita di energia per unità di cammino (*stopping power*) è espressa dalla formula di Bethe-Block. Per energie superiori, nel computo, è necessario considerare anche i processi radiativi che generano, lungo la traccia della particella, sciami elettromagnetici ed adronici che a loro volta producono luce Čerenkov. Fino ad energie intorno a 100 TeV è possibile esprimere lo *stopping power* dei muoni come [12]:

$$-\frac{dE}{dx} = a(E) + b(E)E$$

dove $a(E)$ rappresenta la perdita di energia dovuta alla sola ionizzazione mentre $b(E)$ rappresenta il contributo dovuto alle perdite radiative che può essere espresso nel seguente modo:

$$b(E) = b(E)_{bre} + b(E)_{pair} + b(E)_{nucl}$$

Per i muoni in acqua si può assumere il contributo $a(E) \sim 2 \text{ MeV/cm}$ mentre $b(E) \sim 3.4 \times 10^{-6} \text{ cm}^{-1}$.

Il range del muone che si ottiene integrando le espressioni precedenti risulta circa 200 m per energie di 60 GeV. La lunghezza ottenuta è già sufficiente ad effettuare il tracciamento della particella nel telescopio.

2.7.2 Effetto Čerenkov

Quando una particella carica si propaga all'interno di un materiale dielettrico con una velocità superiore a quella della luce nel mezzo ($v > c/n$ dove n è l'indice di rifrazione del mezzo) viene emessa radiazione Čerenkov. Tale processo radiativo è quindi a soglia. Per un muone l'energia di soglia per l'emissione di luce Čerenkov in acqua (il cui indice di rifrazione $n=1.33$) è $E_\mu = 54.6 \text{ MeV}$.

Tale fenomeno radiativo scaturisce dagli effetti di polarizzazione del mezzo indotti dalla presenza del campo elettrico generato dalla particella carica in movimento. La direzione della radiazione, come mostrato in figura 2.10, è individuata da un angolo θ_c tale che le onde emesse lungo la traccia della

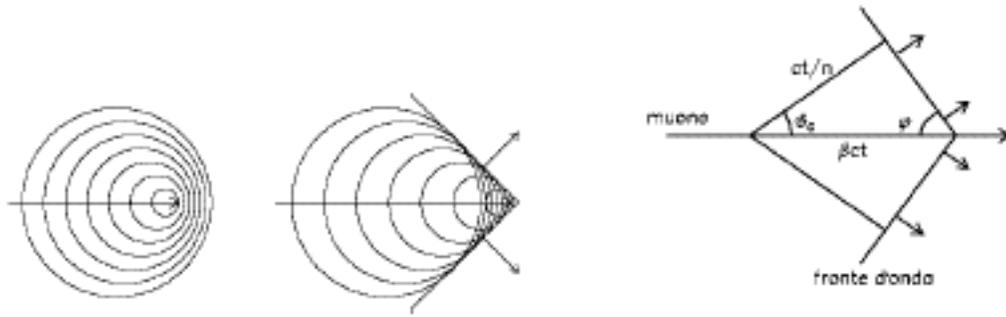


Figura 2.10: Fronte d’onda prodotto dall’involuzione del campo di polarizzazione di una particella relativistica.

particella interferiscano costruttivamente generando un fronte d’onda piano. L’angolo di apertura del cono di luce generato dalla particella è funzione dell’indice di rifrazione del mezzo e della velocità della particella. Utilizzando la costruzione di Huygens del fronte d’onda piano riportata in figura 2.10 si ottiene:

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

Il numero di fotoni emessi in un cammino infinitesimo dx ad una lunghezza d’onda infinitesima $d\lambda$ da una particella di carica unitaria è:

$$\frac{d^2 N}{d\lambda dx} = 2\pi\alpha \sin^2 \frac{\theta_c}{\lambda^2}$$

Poichè tale numero è funzione di λ^{-2} l’emissione di fotoni, nella regione del visibile, risulta maggiore per lunghezze d’onda corrispondenti al blu.

Come visto in precedenza, con un’interazione di corrente carica un neutrino può generare un muone. Questo mediamente conserva un’energia pari al 50-60% del neutrino che l’ha generato [13]. Ciò significa che, per un neutrino da 1 TeV, l’energia dei muoni da esso generati risulta circa 500-600 GeV. Tali muoni sono relativistici e quindi hanno energia sufficiente per generare luce Čerenkov in acqua. Attraverso una rete di sensori ottici è quindi possibile identificare la traccia del muone. Per ricostruire la direzione del neutrino

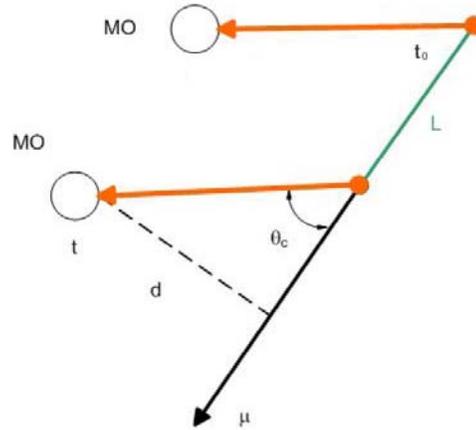


Figura 2.11: Tempo di arrivo del fotone Čerenkov sul modulo ottico (MO) rispetto al tempo di generazione di un fotone.

che lo ha generato è necessario conoscere o stimare l'angolo di emissione del muone rispetto alla direzione di propagazione del neutrino. Sperimentalmente si trova che l'angolo medio tra neutrino e muone diminuisce all'aumentare dell'energia del neutrino che lo ha generato secondo la relazione:

$$\langle \theta_{\nu\mu} \rangle = \frac{1.5^\circ}{\sqrt{E_\nu [TeV]}}$$

2.7.3 Il tracciamento dei muoni

Come visto, il principio di funzionamento di un telescopio sottomarino per neutrini è completamente basato sull'identificazione e la ricostruzione delle tracce generate dai muoni che attraversano il volume sensibile dell'apparato. Per un corretto tracciamento è però necessario combinare l'informazione spaziale (quali sensori sono stati investiti dalla luce Čerenkov) con quella temporale (il tempo di arrivo dei fotoni sui sensori). Il tempo di arrivo di un fotone sul singolo sensore ottico posto a distanza d dalla traccia è espresso dalla relazione:

$$t = t_0 + \left[\frac{L}{c} + \frac{2d}{c \sin 2\theta_c} \right]$$

dove t_0 è un tempo di riferimento t è l'istante di arrivo del singolo fotone sul sensore ottico considerato e θ_c è l'angolo di emissione della luce Čerenkov in acqua che, dalla relazione precedente, risulta essere $\theta_c \sim 42^\circ$. L rappresenta invece la proiezione sulla traccia della distanza tra due moduli ottici come indicato nella figura 2.11.

Ovviamente la ricostruzione della traccia effettuata in questo modo è affetta da un'incertezza dovuta essenzialmente ai seguenti fattori:

- la risoluzione temporale finita dei sensori ottici e dell'elettronica;
- l'indeterminazione sulla posizione relativa dei fotomoltiplicatori in acqua;
- la diffusione dei fotoni in acqua che fa perdere direzionalità alla luce;
- la presenza di segnali spuri dovuti al fondo ottico.

2.8 La situazione internazionale dei telescopi Čerenkov per neutrini

In questa ultima parte del capitolo verrà illustrata la situazione internazionale riguardo i telescopi sottomarini per neutrini passando brevemente in rassegna i principali esperimenti che hanno come obiettivo la costruzione e l'impiego di questo tipo di rivelatore. In tale panoramica, come vedremo, saranno presentati soprattutto una serie di dimostratori tecnologici che rappresentano un necessario studio di fattibilità prima della realizzazione di un apparato avente un'area strumentata di 1 Km^3 .

2.8.1 DUMAND [17]

Dumand è il primo progetto di telescopio per neutrini passato alla fase operativa. Negli anni '80 una collaborazione, costituita da enti di ricerca americani, tentò di installare 200 fotomoltiplicatori, disposti su 9 stringhe verticali e parallele, ad una profondità di 4800 m nell'oceano pacifico a largo delle Hawaii. La superficie di rivelazione per muoni con energia superiore a 1 TeV

era stimata pari a 2000 m². La collaborazione, pur avendo superato una lunga serie di difficoltà logistiche non riuscì a portare a termine la missione a causa della mancanza di un supporto tecnologico adeguato.

2.8.2 BAIKAL

BAIKAL è stato il primo rivelatore per neutrini posto sott'acqua ed attualmente, dopo circa dieci anni di attività, è ancora l'unico telescopio presente nell'emisfero nord della Terra. Situato a circa 1 Km di profondità nel lago Baikal (in Siberia, Russia) è costituito da stringhe verticali di fotomoltiplicatori orientati sia verso l'alto che verso il basso. Le operazioni di installazione e manutenzione dell'apparato avvengono durante l'inverno quando uno strato di ghiaccio ricopre il lago rendendo tali operazioni molto più semplici. BAIKAL è un telescopio ad alta granularità con una soglia di energia di rivelazione $E_\mu \sim 10$ GeV e un'area di 10^5 m².

Il telescopio tuttavia mostra una serie di inconvenienti. La profondità a cui è installato l'apparato non consente un perfetto schermaggio tanto che il fondo di muoni atmosferici è di parecchi ordini di grandezza maggiore del segnale. Il rivelatore, inoltre, può funzionare solo durante i mesi invernali e quindi il tempo di presa dati risulta molto limitato. Infine, le scarse proprietà ottiche dell'acqua del lago che ha una lunghezza di assorbimento di circa 27.9 m limitano le performance di questo apparato. Nonostante ciò con esso si è comunque riusciti a misurare il flusso di neutrini atmosferici e a porre un limite superiore ai flussi diffusi di neutrini di origine astrofisica pari a $E_\nu^2 \Phi_\nu < 4 \times 10^{-7} \text{ GeV cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}$

2.8.3 AMANDA

L'uso del ghiaccio come radiatore Čerenkov venne proposto dalla collaborazione AMANDA negli anni '90. Allo stato attuale, l'apparato, denominato AMANDA II, è il più grande fino ora realizzato e si trova nei ghiacci del Polo Sud. Esso consta di 677 fotomoltiplicatori, ciascuno da 8", tutti orientati verso il basso e montati su 19 stringhe verticali situate ad una profondità compresa tra i 1300 ed i 2400 m. La spaziatura verticale tra i moduli ottici

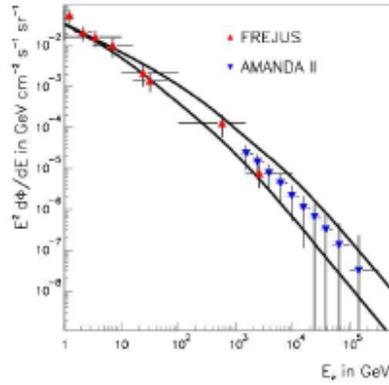


Figura 2.12: Spettro dei neutrini atmosferici misurato da AMANDA, confrontato con quello misurato da FREJUS a basse energie. Le linee piene indicano le previsioni dei modelli per un flusso verticale (linea superiore) e per uno orizzontale (linea inferiore).

va dai 10 ai 20 m e quella orizzontale tra le stringhe va da 30 a 50 m. I risultati degli studi fatti sulle caratteristiche del ghiaccio in cui è stato installato il rivelatore hanno mostrato una lunghezza di assorbimento di circa 100 m ed una lunghezza di diffusione di 20 m. Queste caratteristiche rendono AMANDA un buon calorimetro per eventi astrofisici mentre penalizzano il rivelatore per il puntamento. La massima risoluzione angolare ottenibile è, infatti, compresa tra 1.5° e 3.5° per i muoni. Il principale vantaggio del sito in cui è costruito AMANDA risiede nell'assenza di rumore ottico dovuto a sorgenti quali il decadimento radioattivo del potassio o la bioluminescenza.

I dati forniti da AMANDA hanno permesso, per la prima volta, di studiare il flusso di neutrini astrofisici che provengono dalla parte bassa del rivelatore nel range di energia da pochi TeV fino a 300 TeV (figura 2.12). Il telescopio è anche sensibile agli sciami generati dai leptoni derivanti dall'interazione di corrente carica dei neutrini di tipo elettronico e taunico. la sensibilità dello strumento ai tre flavor è $E_\nu^2 \Phi_{\nu_\mu + \nu_e + \nu_\tau} < 8.6 \times 10^{-7} \text{ GeV cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}$ con $50 \text{ TeV} \leq E_\nu \leq 5 \text{ PeV}$. Al disopra del PeV la Terra risulta opaca ai neutrini e quindi la ricerca di sorgenti astrofisiche è concentrata su eventi vicini all'orizzonte o al disopra di esso. Tale misura è possibile poichè a queste energie il fondo di muoni atmosferici risulta molto basso.

2.8.4 IceCube

IceCube rappresenta l'estensione di AMANDA alle dimensioni di 1 Km³. Quando sarà completo (data prevista 2011) consterà di 4800 PMT tutti orientati verso il basso e disposti su 80 stringhe. La distanza tra i fotomoltiplicatori sarà di 16 m con una spaziatura tra le stringhe di 125 m. IceCube sarà in grado di identificare le tracce dai muoni prodotti da ν_μ con energia $E_{\nu_\mu} > 100$ GeV e sciami prodotti da ν_e con $E_{\nu_e} > 10$ TeV e da ν_τ con $E_{\nu_\tau} > 1$ PeV. Le simulazioni prevedono che in tre anni di acquisizione dati l'apparato sarà in grado di raggiungere una sensibilità $E_{\nu_\mu}^2 \Phi_{\nu_\mu} = 4.2 \times 10^{-9} \text{GeV cm}^{-2} \text{s}^{-1} \text{sr}^{-1}$ per il flusso di neutrini e $E_{\nu_\mu}^2 \Phi_{\nu_\mu} = 2.4 \times 10^{-9} \text{GeV cm}^{-2} \text{s}^{-1} \text{sr}^{-1}$ per sorgenti puntiformi. Inoltre, i rivelatori installati in ghiaccio a differenza di quelli in mare sono più adatti a rivelare neutrini di bassa energia provenienti da Supernovae poichè non sono affetti da rumore ottico di natura biologica o radioattiva.

2.8.5 NESTOR

La collaborazione NESTOR è stata la prima ad aver condotto studi di fattibilità per la realizzazione di un telescopio per neutrini di 1 Km³ nel mediterraneo. La proposta formulata prevede la costruzione di un rivelatore modulare da installare a 3800 m di profondità in un sito posto a 20 Km dalla costa di Methoni nel Peloponneso (Grecia). Ogni modulo è costituito da una struttura semi rigida alta 360 m e con un diametro di 32 m equipaggiata con 170 PMT orientati sia verso l'alto che verso il basso. Nel marzo del 2003 è stato posizionato un primo piano con 12 moduli ottici avente un diametro di 12 m al fine di studiare le performances dell'apparato. Esso ha acquisito dati per circa un mese da cui sono state ricostruite 745 tracce di muoni atmosferici. Questo studio fa sperare che sia possibile misurare il flusso di muoni atmosferici in funzione dell'angolo di zenith. Nel futuro la collaborazione ha in progetto il posizionamento della prima torre completa avente una superficie di rivelazione di circa $2 \times 10^4 \text{ m}^2$ che permetterà di rivelare muoni di energia $E_\mu > 10$ TeV.

2.8.6 ANTARES

La collaborazione è impegnata nella costruzione di un apparato dimostratore con superficie di 0.1 Km^2 in un sito sottomarino vicino Tolone (Francia) ad una profondità di 2400 m. Il telescopio è un rivelatore ad alta granularità costituito da 12 stringhe verticali ciascuna equipaggiata con 75 PMT orientati verso il basso a 45° rispetto alla verticale e posti ad una distanza di 60 m. Il posizionamento dei PMT dovrebbe ridurre il più possibile l'oscuramento dei moduli ottici dovuto alla crescita biologica ed alla sedimentazione. In questo modo sarà possibile, inoltre, ridurre la soglia di rivelazione dei muoni fino ad energie di circa 10 GeV. Dovrebbe essere possibile, quindi, studiare fenomeni di bassa energia quali le oscillazioni dei neutrini atmosferici e la ricerca della materia oscura. Rispetto ad esperimenti posti in ghiaccio come AMANDA e BAIKAL, fortemente penalizzati dalla diffusione della luce nel mezzo, il telescopio ANTARES dovrebbe raggiungere una precisione di puntamento vicina a 0.1° . Dopo una prima posa in opera di due stringhe, avvenuta nella primavera del 2002, da aprile 2005 una nuova versione instrumentata con moduli ottici e sensori oceanografici sta raccogliendo dati. L'analisi effettuata fin ora mostra un inaspettato valore di rumore ottico compreso tra 60 e 100 KHz inspiegabile considerando fenomeni di decadimento radioattivo dovuto a ^{40}K . Questo rumore potrebbe essere dovuto alla bioluminescenza e potrebbe limitare molto le capacità di rivelazione di ANTARES. Per questo motivo la collaborazione sta studiando la possibilità di applicare ai dati condizioni di trigger più stringenti che permettano di ridurre drasticamente il rumore.

2.8.7 Il futuro dei telescopi per neutrini: Km^3Net

L'osservazione contemporanea di tutto il cielo è un obiettivo scientifico fondamentale. Per questo motivo la comunità scientifica è fortemente intenzionata a realizzare almeno due telescopi delle dimensioni di 1 Km^3 ciascuno posizionato in un emisfero terrestre. Un primo rivelatore potrebbe essere IceCube che, posto al polo Sud, sarebbe in grado di studiare la porzione nord del cielo (figura 2.12). L'altro potrebbe essere un telescopio per neutrini posizionato nelle acque del mediterraneo. Tale zona risulta particolarmente interessante

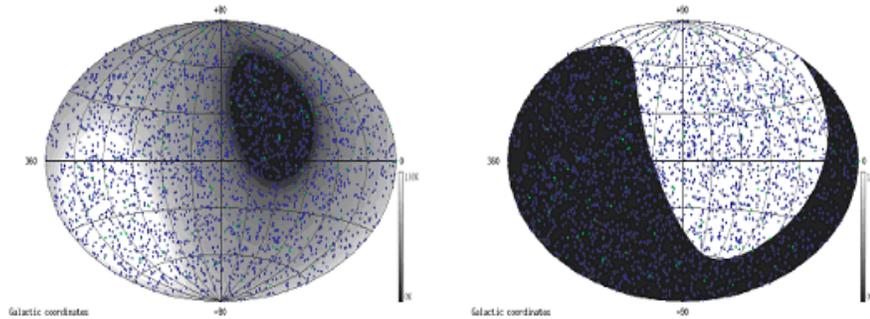


Figura 2.13: Rappresentazione dello sky map (in coordinate celesti) osservabile da un telescopio sito nell'emisfero settentrionale (figura di sinistra) e meridionale (figura di destra). La potenzialità di osservazione è deducibile dalla scale di grigi mediante la quale è stata rappresentata la figura.

poichè solo da qui sarebbe possibile osservare il centro della nostra galassia. Le condizioni climatiche e marine consentirebbero, inoltre, di acquisire dati durante tutto l'anno. I due telescopi inoltre avrebbero caratteristiche complementari: l'elevata lunghezza d'assorbimento ottenibile nei ghiacci antartici permetterebbe una buona risoluzione in energia mentre l'elevata distanza di scattering che caratterizza le profondità marine del mediterraneo permetterebbe un'eccellente risoluzione angolare.

Un altro aspetto fondamentale che caratterizza un telescopio delle dimensioni di 1 Km^3 è quello relativo alla sensibilità raggiungibile dallo strumento. Tale sensibilità viene messa in relazione al flusso differenziale in funzione dell'energia dei neutrini emessi dalle sorgenti citate (figura 2.14). In figura è riportato, inoltre, il dettaglio dei meccanismi di produzione dei neutrini da cui è possibile risalire alla sorgente fisica che li ha generati. Si nota che un telescopio delle dimensioni di 1 Km^3 offra la massima sensibilità ed è quindi in grado di osservare bene sorgenti galattiche ed extragalattiche. Attualmente le difficoltà da affrontare nella realizzazione di un telescopio di 1 Km^3 in acqua sono sostanzialmente tecnologiche. L'ambiente sottomarino risulta, infatti, particolarmente ostile. La pressione assieme alla salinità riduce di molto il tempo di vita dei materiali quindi tutte le strutture meccaniche

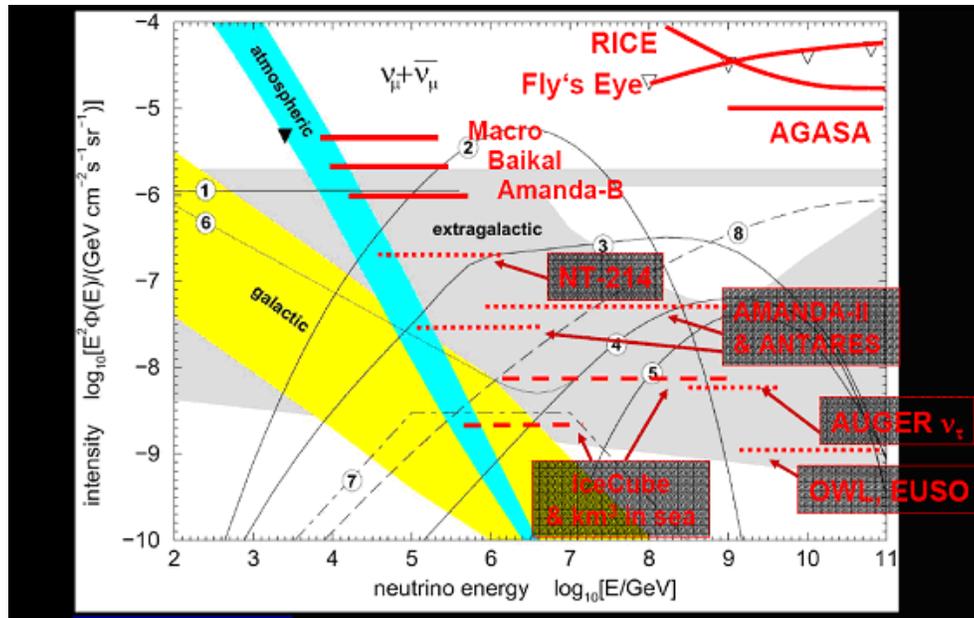


Figura 2.14: Flusso di neutrini diffuso da sorgenti astrofisiche in relazione alla sensibilità dei principali telescopi per neutrini. In figura sono rappresentati in dettaglio i meccanismi di interazione che portano alla produzione di neutrini: 1) flusso di neutrini prodotto da AGN a seguito di interazione pp; 2) neutrini prodotti da AGN tramite interazione p γ ; 3) previsione massima di produzione p γ ; 4) neutrini generati da jet e blazar; 5) GZK; 6) neutrini dalla superficie di AGN per interazione pp; 7) neutrini da Gamma Ray Burst; 8) previsione in base al modello TD signl

devono essere realizzate con grande accortezza. Anche il tipo di struttura che costituirà il rivelatore dovrà essere studiata accuratamente. Strutture molto granulari come quella di ANTARES oppure molto rigide come quella di NESTOR hanno presentato grandi difficoltà realizzative e di posa in opera. Le esperienze fin ora accumulate comunque rappresentano un'eccezionale base da cui partire per la realizzazione del telescopio nel mediterraneo. Per raccoglierle la Comunità Europea ha messo in piedi un gruppo di studio, denominato Km³Net, al quale partecipano numerosi enti di ricerca provenienti da tutta Europa. Il progetto si concluderà con uno studio di fattibilità che porterà alla progettazione ed alla realizzazione del telescopio per neutrini da 1 Km³ nel mediterraneo.

Capitolo 3

Il telescopio per neutrini da 1 Km³ NEMO

3.1 Introduzione

Nei capitoli precedenti è stato giustificato il motivo dell'impiego dei neutrini come sonda astrofisica e sono state brevemente descritte le principali sorgenti che è possibile studiare per mezzo di queste particelle. Sono state anche descritte le tecniche più efficaci per la rivelazione e lo studio di tali particelle.

Come si è discusso precedentemente le particolari caratteristiche del neutrino, unitamente al ridotto flusso di particelle atteso, rende necessaria la realizzazione di apparati di grandissime dimensioni. In base alle argomentazioni scientifiche riportate nei capitoli precedenti, se si vuole realizzare apparati con buona sensibilità, è necessario che essi abbiano un volume di rivelazione non inferiore a 1 Km³. La necessità poi di schermare, almeno parzialmente, tali apparati dai flussi di neutrini atmosferici, che in questo tipo di misura costituiscono rumore, rende essenziale la collocazione di questo tipo di strumenti in mare a profondità non inferiori a 3000 metri. Le specifiche costruttive elencate pongono enormi difficoltà tecnologiche che devono essere affrontate e risolte dal gruppo di ricercatori che intende realizzare un apparato di questo tipo.

In questo capitolo viene presentato il telescopio per neutrini più moderno in fase di realizzazione denominato NEMO. Il progetto NEMO (*NEutrino Mediterranean Observatory*) è frutto di una collaborazione di fisici italiani

provenienti da molte università ed enti di ricerca e finanziati dall'Istituto Nazionale di Fisica Nucleare (INFN). Alla realizzazione di questo progetto collaborano inoltre diversi istituti scientifici specializzati in oceanografia, biologia marina e geofisica rendendo NEMO anche un laboratorio multidisciplinare.

3.2 La scelta del sito di installazione

La scelta del sito sottomarino di installazione del rivelatore è estremamente importante. Da essa infatti dipende gran parte della riuscita dell'esperimento. Per questo motivo a partire dal 1999 la collaborazione NEMO ha svolto un programma di ricerca, durato quattro anni, che ha avuto come obiettivo lo studio e la caratterizzazione di una serie di siti sottomarini nelle vicinanze delle coste italiane. Lo studio ha portato all'individuazione di due siti nel bacino tirrenico, a largo di Ustica e di Alicudi, e due nel bacino ionico a largo di Capo Passero. Su di essi sono state effettuate varie campagne di misura volte a studiare approfonditamente le caratteristiche di queste zone. I principali parametri oggetto di studio sono stati:

- la misura delle proprietà ottiche di assorbimento e di diffusione della luce nell'acqua;
- la misura del fondo ottico dovuto a fenomeni di bioluminescenza e al decadimento radioattivo degli isotopi disciolti in acqua (principalmente il ^{40}K);
- la misura delle proprietà oceanografiche del sito quali la temperatura e la salinità dell'acqua e la quantità di particolato disciolto;
- la misura delle velocità delle correnti e delle sue variazioni.

Queste misure, come vedremo tra breve, hanno una forte incidenza sia sul progetto che sulle prestazioni del telescopio stesso. Per la scelta del sito hanno giocato un ruolo fondamentale anche la sua profondità, determinante al fine di migliorare il rapporto segnale rumore, e la vicinanza alla

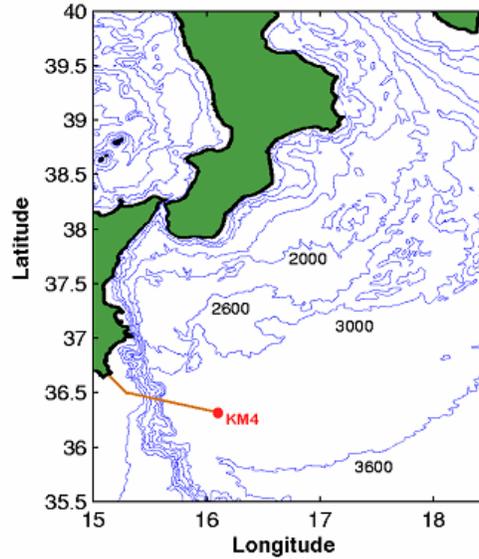


Figura 3.1: Posizione geografica del sito di installazione del telescopio per neutrini NEMO.

costa. Quest'ultima caratteristica risulta importante per aumentare l'affidabilità della trasmissione dati a terra. Se, infatti, si seleziona un sito che sia distante al massimo 100 Km dalla costa per la comunicazione è possibile utilizzare un cavo elettro-ottico su cui i dati possono viaggiare senza che siano necessari dispositivi che amplifichino i segnali. Mediante questo tipo di cavi è anche possibile fornire la potenza elettrica necessaria al funzionamento dell'apparato.

Le approfondite indagini condotte dalla collaborazione hanno portato a scegliere il sito denominato KM4 posto a largo di Capo Passero (36° 16' N, 16° 16' E). La zona selezionata per la collocazione del telescopio ha una profondità di circa 3500 metri e si trova a circa 100 Km a Sud-Est dalla punta meridionale della Sicilia (figura 3.1).

3.2.1 Proprietà ottiche dell'acqua

L'acqua del sito selezionato deve avere caratteristiche di basso assorbimento e bassa diffusione della luce. Questi parametri risultano determinati per

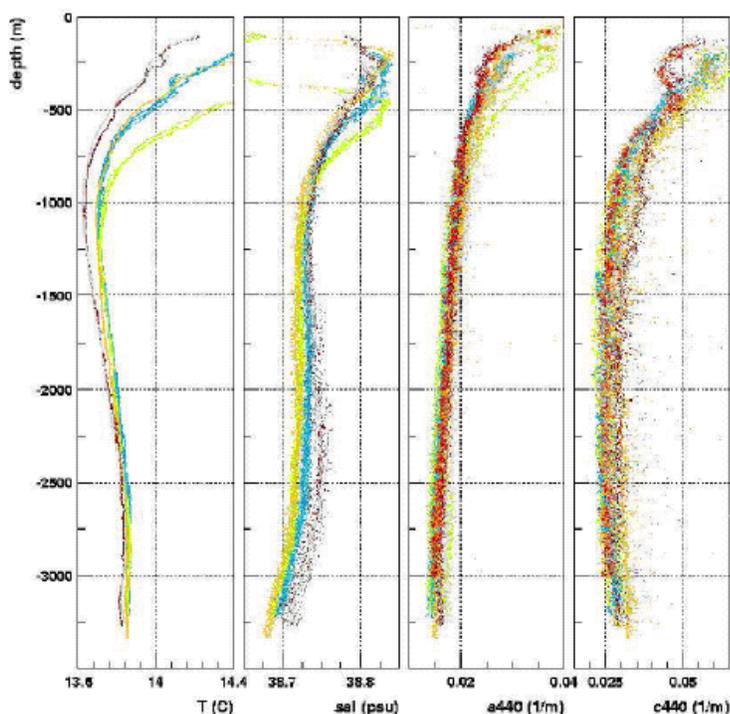


Figura 3.2: Profili di temperatura (T), salinità (sal), coefficiente di assorbimento (a) e attenuazione (c) alla lunghezza di d'onda di 440 nm misurati a Capo Passero. I dati sono riferiti a cinque diverse campagne condotte nel Dicembre 1999 (rosso), Marzo 2002 (verde), Maggio 2002 (giallo), Agosto 2002 (arancio) e Agosto 2003 (azzurro). Le variazioni a basse profondità sono di carattere stagionale.

stabilire la distanza massima a cui possono essere posti i fotomoltiplicatori preposti alla rivelazione della luce Čerenkov. In definitiva, a parità di area instrumentata, minore è la diffusione e l'assorbimento della luce e minore è il numero di fotomoltiplicatori da utilizzare. La caratterizzazione delle proprietà ottiche dell'acqua nei siti selezionati è consistita nella misura dei coefficienti di assorbimento e di attenuazione della luce per 9 diverse lunghezze d'onda comprese tra 412 nm e 715 nm. I risultati delle misure mostrano che alla profondità di 3300 metri nel sito di Capo Passero la lunghezza di assorbimento per la luce blu ($\lambda = 440$ nm) è pari a 0.014 ± 0.003 m⁻¹ ed un coefficiente di attenuazione di 0.025 ± 0.003 m⁻¹. Da questi valori è stata stimata una lunghezza di assorbimento della luce pari a circa 60 metri (com-

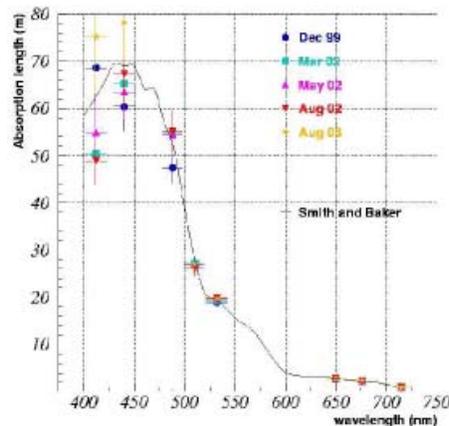


Figura 3.3: Lunghezza di assorbimento in funzione della lunghezza d'onda misurata in cinque campagne. Ogni punto rappresenta il valore medio della lunghezza di assorbimento in riferimento ad un dato valore di λ . La linea rappresenta il valore di assorbimento dell'acqua pura misurato da Smith e Baker [20].

parabile con quella dell'acqua pura) mentre la lunghezza di attenuazione risulta essere circa 40 m [19]. I dati mostrano anche che le proprietà ottiche dell'acqua sono stabili nel tempo. In figura 3.2 sono riportati i parametri dell'acqua misurati in cinque campagne condotte nel dicembre 1999 (rosso), marzo 2002 (verde), maggio 2002 (giallo), agosto 2002 (arancio), agosto 2003 (azzurro). Le figure 3.3 e 3.4 riportano la lunghezza di assorbimento e di attenuazione dell'acqua in funzione della lunghezza d'onda.

3.2.2 Le correnti marine

Un rivelatore sottomarino costituito da una serie di moduli (siano essi torri o stringhe) agganciati al fondo del mare e tenuti in tensione mediante boe risulta particolarmente sensibile alle correnti sottomarine. Esse possono creare una serie di problemi sia in fase di deposizione che in fase di utilizzo dei moduli stessi. Le correnti potrebbero infatti provocare torsioni o inclinazioni della struttura che potrebbero condurre addirittura alla collisione di moduli vicini. E' quindi fondamentale conoscere le correnti sottomarine presenti nel sito prescelto. Per questo motivo sono state effettuate una serie di mi-

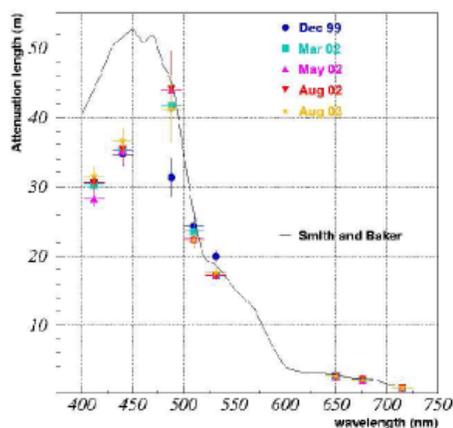


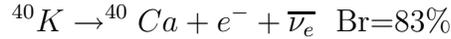
Figura 3.4: Lunghezza di attenuazione in funzione della lunghezza d'onda misurata in cinque campagne. Ogni punto rappresenta il valore medio della lunghezza di attenuazione in riferimento ad un dato valore di λ . La linea rappresenta il valore di attenuazione dell'acqua pura misurato da Smith e Baker [20].

sure che hanno mostrato, per il sito di Capo Passero, un valor medio delle correnti molto basso pari a circa 3 cm/s con picchi che non superano i 12 cm/s. Il sito risulta quindi particolarmente indicato per l'installazione di un telescopio di grandi dimensioni poichè la probabilità di collisioni dovute alle correnti risulta praticamente nulla. Oltre a problemi di natura strutturale l'inclinazione dei moduli produce anche errori di ricostruzione delle tracce se non è possibile determinare con estrema precisione la posizione dei fotomoltiplicatori. Per questo motivo è necessario dotare il rivelatore di un sistema di posizionamento estremamente preciso che consenta di conoscere la posizione dei PMT istante per istante. Ovviamente anche la realizzazione di un sistema di questo tipo risulta agevolata dalla bassa intensità delle correnti sottomarine presenti nel sito prescelto.

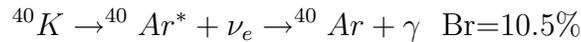
3.2.3 Il fondo ottico inorganico

Nell'acqua marina vi è una quantità di sali disciolti alcuni dei quali hanno una componente radioattiva. Il decadimento di questi elementi può produrre un rilascio di energia all'interno del volume sensibile che può creare disturbo

al rivelatore. Il contributo maggiore è fornito dal potassio ed in particolare dall'isotopo ⁴⁰K. Tale elemento decade β^- emettendo fotoni secondo la relazione:



oppure per cattura elettronica, seguita dal rilascio di un fotone come descritto dalla relazione:



il fotone successivamente produce elettroni per effetto Compton.

In entrambi i casi gli elettroni hanno abbastanza energia da produrre in acqua fotoni Čerenkov che contribuiscono al rumore rivelato dal telescopio. La statistica che regola il decadimento del ⁴⁰K è quella di Poisson e quindi tale fenomeno genera un segnale di fondo con una frequenza media di eventi costante. Nelle varie campagne di misura, effettuate nel sito di Capo Passero utilizzando fotomoltiplicatori da 2 pollici, è stata registrata una frequenza media di eventi acquisiti pari a circa 380 Hz [21]. Tale parametro fornisce informazioni preziose per la progettazione dell'elettronica di trasmissione dati. Alla luce di questi risultati è infatti noto che, anche in assenza di eventi rilevanti, ogni fotomoltiplicatore del telescopio registrerà segnali provenienti dal decadimento ora descritto. Dato che la frequenza di eventi rivelati dipende linearmente dall'area attiva del sensore ottico e poichè la collaborazione NEMO ha deciso di utilizzare PMT da 13 pollici ne consegue che la frequenza attesa dal fondo di ⁴⁰K dovrebbe essere pari a circa 16 KHz.

3.2.4 Il fondo ottico da bioluminescenza

Nelle grandi profondità marine, dove la luce del sole non riesce a penetrare, sono presenti una serie di microrganismi che si sono adattati a vivere in condizioni di totale oscurità. Questi esseri viventi hanno sviluppato organi in grado di emettere luce. Tale fenomeno prende il nome di bioluminescenza. Anche se le conoscenze in merito ai microrganismi bioluminescenti risultano piuttosto scarse, ciò che si conosce è la modalità di emissione e lo spettro

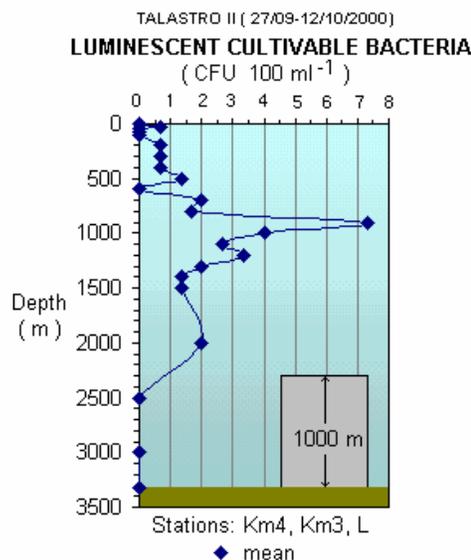


Figura 3.5: Profilo della concentrazione di batteri bioluminescenti in funzione della profondità per il sito di Capo Passero.

della luce che risulta centrato attorno alle lunghezze d'onda del verde e dell'azzurro. Ciò significa che l'utilizzo di filtri per schermare tale radiazione è totalmente da escludere poichè questa viene emessa alle stesse lunghezze d'onda della luce Čerenkov. Le raffiche di fotoni (*burst*) prodotti da questi piccoli animali possono durare fino a varie centinaia di millisecondi a frequenze che possono superare anche di vari ordini di grandezza quelle del ⁴⁰K. Le misure effettuate a Capo Passero mostrano come il sito prescelto sia caratterizzato da un'attività biotica molto limitata (figura 3.5).

Un altro elemento che disturba la raccolta dati del telescopio è rappresentato dal materiale biologico in sospensione prodotto dai biorganismi (*biofouling*). Tale materiale può depositarsi sulla parte attiva dei sensori ottici aumentando gli effetti di diffusione della luce e quindi peggiorando la risoluzione angolare nella ricostruzione delle tracce. Inoltre il sedimento che col tempo tende ad accumularsi riduce l'efficienza del rivelatore. Il problema del biofouling è stato affrontato effettuando una serie di misure. Ci si è innanzitutto concentrati nel misurare la quantità di sedimenti e si è notato come,

nel sito prescelto, essa risulta notevolmente inferiore rispetto ad altre regioni costiere. Si è poi provveduto alla misurazione dell'attenuazione della luce raccolta all'interno di una sfera in funzione dell'angolo azimutale della sfera stessa. Si è notato come in condizioni di scarsa corrente marina il sedimento tende ad accumularsi sulla parte superiore della sfera. Per questo motivo si è deciso di arginare il problema posizionando i fotomoltiplicatori in modo che la parte attiva non sia rivolta verso l'alto.

3.3 La struttura del rivelatore

Il telescopio per neutrini NEMO, oltre a ottimizzare tutte le caratteristiche necessarie alla corretta ricostruzione delle tracce, deve anche soddisfare una serie di requisiti tecnici dettati dalle particolari condizioni in cui si troverà ad operare e per un tempo non inferiore a 10 anni. Dato che il rivelatore verrà deposto in mare ad una profondità di 3500 metri, tutti gli elementi che lo compongono dovranno essere resistenti alle elevatissime pressioni alle quali saranno sottoposti. Considerando che la pressione aumenta di 1 bar ogni 10 metri di colonna d'acqua, ogni singola parte del telescopio dovrà sopportare una pressione non inferiore a 350 atmosfere. Un altro problema non trascurabile è legato alla composizione chimica dell'acqua nel quale esso risulterà immerso. I vari sali che sono in essa disciolti possono attaccare, chimicamente ed elettroliticamente, la maggior parte dei metalli corrodendoli. Tutte le parti meccaniche devono quindi essere realizzate in modo da resistere per lungo tempo a quest'azione chimica. Tale necessità riduce le materie prime utilizzabili a poche leghe di metallo. Nella progettazione non è da trascurare, infine, la fase di messa in opera dell'apparato e la possibilità di eseguire eventuali lavori di manutenzione durante il tempo di raccolta dati.

Schematicamente il telescopio NEMO è costituito da tre parti principali:

- la struttura del rivelatore costituita dalle parti meccaniche, denominate torri, che servono a reggere i fotomoltiplicatori e tutta l'elettronica necessaria alla raccolta ed alla trasmissione dati a terra;
- una stazione di terra in grado di fornire la potenza elettrica al la-

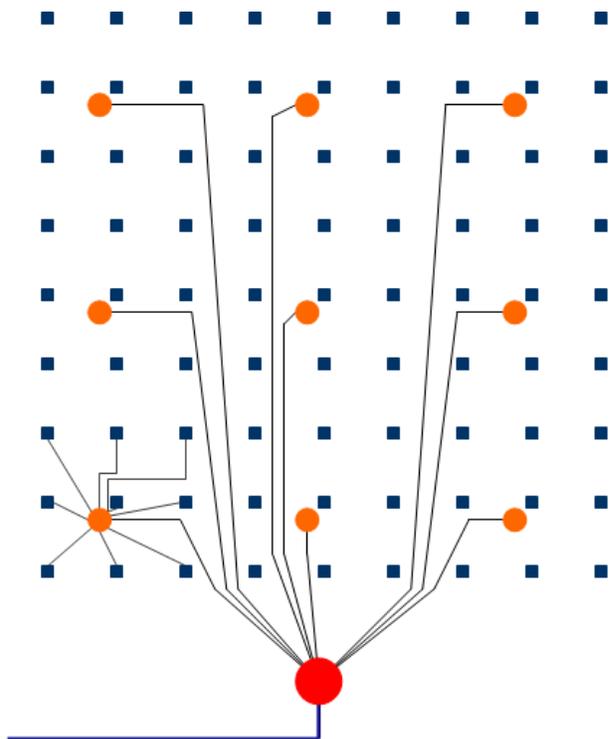


Figura 3.6: Schema di principio del rivelatore da 1 Km³ NEMO.

boratorio sottomarino, ospitare il sistema di controllo e di gestione della strumentazione nonché il sistema per la raccolta dei dati;

- una rete di cavi sottomarini in grado di connettere il telescopio con il laboratorio a terra.

3.3.1 La geometria del rivelatore

Il numero e la disposizione delle varie torri che compongono il telescopio sottomarino devono essere scelti in modo da rendere massime le potenzialità di rivelazione, rendere semplici gli algoritmi di trigger in grado di selezionare le tracce interessanti acquisite dal telescopio separandole dal fondo di rumore e rendere agevole la posa in opera e la connessione delle diverse torri con i cavi sottomarini. Tutti questi obiettivi devono essere perseguiti tentando di limitare il più possibile i costi. La geometria del telescopio, la disposizione ed

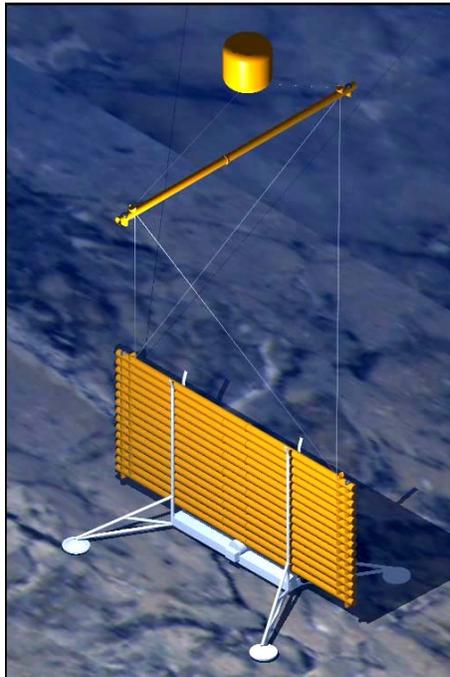


Figura 3.7: In figura è riportato un disegno dell'apertura della torre una volta in mare.

il numero dei sensori ottici sono stati a lungo studiati mediante simulazioni Monte Carlo in modo da ottenere il miglior compromesso tra le potenzialità del rivelatore ed i costi di produzione. Attualmente la struttura più accreditata prevede l'impiego di 81 torri, disposte su un reticolo 9×9 e posizionate ad una distanza di circa 140 metri l'una dall'altra come mostrato in figura 3.6.

3.3.2 La torre ed i piani di NEMO

Anche la struttura e la geometria della torre e dei piani che la compongono deve soddisfare i requisiti descritti in precedenza. Per questo motivo si è pensato di realizzare una torre costituita da 16 piani. Ogni torre, una volta messa in acqua, viene ancorata al fondo mediante un corpo morto e tenuta in tensione per mezzo di una boa. Una serie di tiranti collega i diversi piani tenendoli equidistanti ed orientati con un angolo di 90° l'uno rispetto

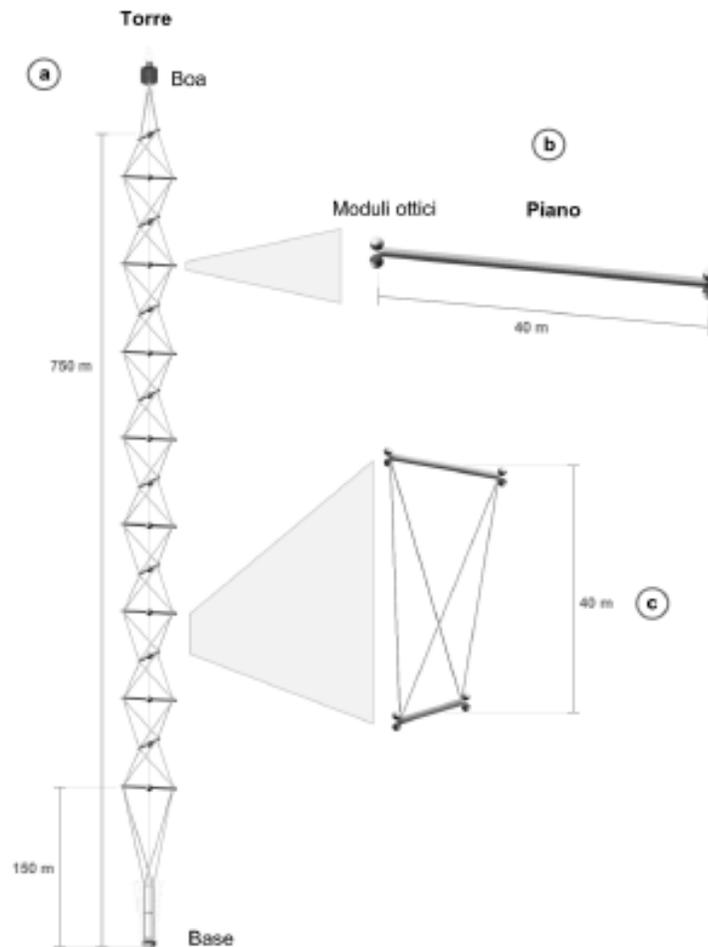


Figura 3.8: Schema di una torre di NEMO. La figura (a) mostra la struttura complessiva della torre. La figura (b) mostra un dettaglio di un piano della torre da cui è possibile notare i quattro PMT posti alle due estremità. In figura (c) viene infine mostrato il sistema di cavi che, tenuti in tensione dalla boa di galleggiamento, vincolano i piani nella loro posizione.

all'altro. Questa tecnica risulta particolarmente efficace poichè permette di compattare i piani quando sono fuori dall'acqua consentendo un agevole trasporto della strumentazione mentre, una volta messa in acqua ed ancorata al fondo è possibile aprirla mediante un robot sottomarino detto ROV (*Remote Operating Vehicles*)(figura 3.7).

Una volta aperta, la torre risulta alta 750 metri; la distanza tra ognuno dei 16 piani risulta circa 40 metri mentre la distanza del primo piano dal fondo è di circa 150 metri. Ogni piano della torre è costituito da un traliccio di tubi metallici, realizzati in alluminio marino, lungo circa 40 metri. Ad ognuna delle due estremità sono posizionati due moduli ottici (OM) costituiti da una Benthos-sfera contenente un fotomoltiplicatore e la relativa elettronica di *fornt end*. Dei quattro PMT due sono rivolti verso il basso e due in maniera orizzontale rispetto al fondo marino in direzione esterna rispetto alla torre. La struttura qui descritta è in grado di coprire un volume d'acqua di 1 Km³ ed il numero di PMT impiegati risulta pari a 5184. In figura 3.8 viene presentato un disegno della torre di NEMO.

3.4 L'elettronica del telescopio

Nella costruzione di un apparato di dimensioni così imponenti, che deve funzionare in maniera del tutto indipendente in ambiente estremamente ostile, un ruolo fondamentale è rivestito dall'elettronica di distribuzione della potenza, di controllo e di acquisizione dati. Essa, infatti, da un lato deve essere molto complessa per svolgere tutte le operazioni necessarie al corretto funzionamento dell'apparato e dall'altro deve essere estremamente affidabile per funzionare correttamente per un lunghissimo periodo. Dato il bassissimo numero di eventi attesi, il telescopio deve raccogliere dati per un tempo non inferiore a dieci anni durante i quali non sarà possibile alcun intervento di manutenzione dei componenti elettronici. D'altra parte, vista la grande quantità di apparati da alimentare, e, soprattutto, la grande distanza alla quale la potenza deve essere trasportata, è necessario progettare ogni elemento del sistema in modo tale da ridurre al massimo i consumi.

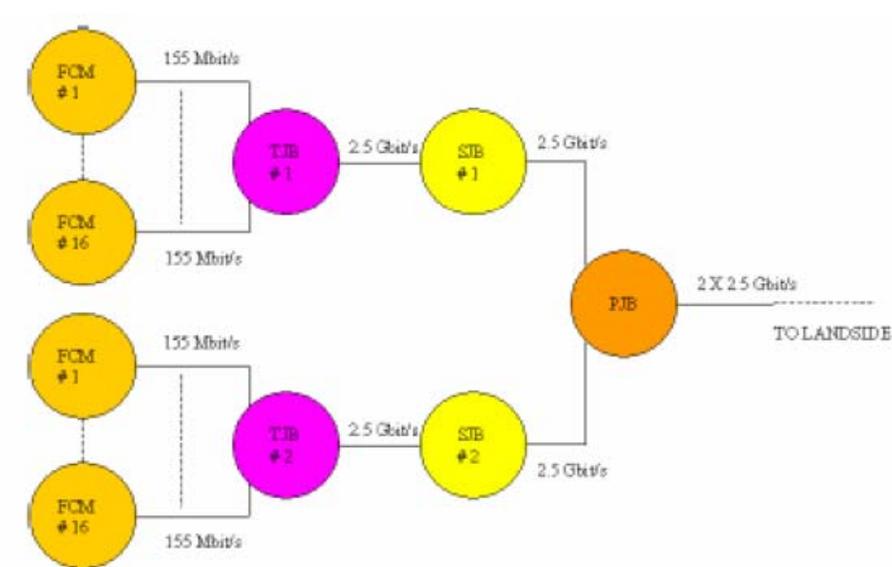


Figura 3.9: Disegno schematico delle interconnessioni tra le diverse JB di NEMO.

3.4.1 La topologia dei collegamenti

La struttura meccanica dell'apparato è stata realizzata in modo da offrire una estrema modularità al telescopio che ben si presta ad una progettazione gerarchica delle interconnessioni. Per questo motivo è stata realizzata una struttura ad albero in cui i cavi di collegamento vengono via via separati in modo da raggiungere tutti i 5184 fotomoltiplicatori del telescopio. In dettaglio, il collegamento tra la stazione di terra ed il telescopio viene effettuato mediante un cavo elettro-ottico principale lungo circa 100 Km che termina nelle vicinanze dell'apparato. La potenza elettrica viene distribuita mediante la parte in rame del suddetto cavo mentre il trasporto dei segnali è affidato alle fibre ottiche presenti in esso. Un sistema di controllo intelligente permetterà di gestire le fasi di accensione e spegnimento consentendo di attivare in modo sequenziale e controllato i vari componenti. Durante la fase operativa il sistema sarà in grado di segnalare eventuali malfunzionamenti permettendo di isolarli dall'alimentazione in modo da evitare rotture.

La figura 3.9 mostra lo schema delle interconnessioni a partire dal cavo

elettro-ottico principale fino ad arrivare alla base delle torri. Come mostrato in figura sono stati progettati una serie di dispositivi in grado di separare il cavo in tutti i suoi rami. Il cavo principale termina nella scatola di giunzione primaria (denominata *Primary Junction Box* PJB). Lì il cavo viene suddiviso in nove parti e connesso ai nove punti di derivazione secondaria denominati *Secondary Junction Box* (SBJ). In ognuna delle SBJ il cavo elettro-ottico viene ulteriormente separato ed inviato a tutte le torri alle cui basi è posto l'ultimo punto di derivazione (denominato *Tower Junction Box* TBJ). Da qui il cavo raggiunge tutti i piani della torre. Al centro di ogn'uno di essi è presente una modulo di controllo denominato *Floor Control Module* (FCM) in grado di interconnettere e gestire tutta l'elettronica di piano. In figura 3.10 viene mostrata schematicamente la posizione dei componenti elettronici della torre.

Tutti i collegamenti descritti sono bidirezionali quindi il flusso di dati proveniente da un piano della torre compie lo stesso percorso descritto in direzione inversa. Per limitare la complessità delle interconnessioni e migliorare l'affidabilità dell'intero sistema tutte le derivazioni dei segnali ottici, effettuate dai tre ordini di junction box, sono state realizzate senza utilizzare dispositivi ottici attivi. I segnali ottici vengono infatti gestiti solo da componenti passivi quali *multiplexer/demultiplexer* ottici. Il tipo di interconnessione progettato prevede che il segnale elettrico prodotto dal fotomoltiplicatore venga convertito in ottico dall'FCM e così venga trasportato fino a terra. Anche i segnali di temporizzazione vengono trasportati attraverso questo canale e ciò comporta l'eliminazione di eventuali e poco controllabili effetti di ritardo che la conversione elettro-ottica introdurrebbe. Questo aspetto riveste un ruolo cruciale per la riuscita del progetto data l'estrema importanza che l'informazione temporale riveste nella ricostruzione delle tracce. Il tipo di interconnessione ad albero fin qui descritta possiede un ulteriore vantaggio. Data la sua grande modularità permette il funzionamento dell'apparato anche se non tutti i moduli sono connessi. In questo modo da un lato sarà possibile iniziare ad acquisire dati anche prima della fine della costruzione dell'apparato e dall'altro sarà possibile continuare a far funzionare il telescopio anche in caso di rottura di una o più torri. In tal caso sarà, infatti

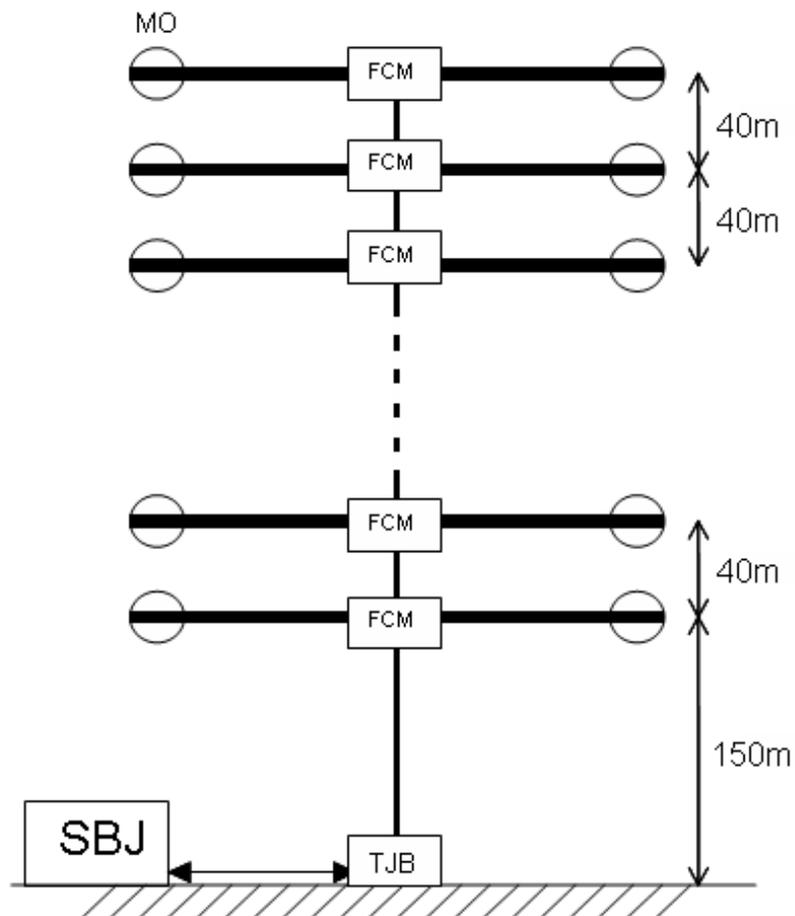


Figura 3.10: Schema che mostra i collegamenti elettro-ottici di una torre di NEMO.

possibile isolare gli elementi danneggiati senza compromettere l'acquisizione dei dati dal telescopio.

3.4.2 L'elettronica di piano

Su ogni piano sono presenti una serie di dispositivi elettronici che permettono sia di acquisire i segnali fisici provenienti dai fotomoltiplicatori sia di monitorare ed inviare a terra una serie di parametri funzionali che permettono di comprendere lo stato di funzionamento dell'intero sistema. Nel centro del piano è presente un cilindro molto spesso di alluminio marino al cui interno, protetti dalle pressione e dall'umidità sono installati due dispositivi elettronici lo *slow control module*(SCM) e il *Floor Control Module* (FCM). L'SCM permette di gestire una serie di strumenti per il monitoraggio ambientale a bassa velocità. Il piano è infatti equipaggiato con una serie di strumenti quali un correntometro in grado di misurare la corrente marina, misuratori di salinità, bussola, inclinometri e tilt-meters che assieme al sistema di posizionamento sono in grado di definire la posizione del piano, e quindi della torre, in acqua con estrema precisione. A bordo del piano sono poi presenti una serie di sensori di temperatura di tensione e di corrente elettrica. L'SCM si interfaccia con tutti questi dispositivi mediante una serie di porte *general purpose* digitali come le RS 232 ed una serie di interfacce analogiche. Lo *Slow Control Module* elabora questi dati e li invia all'FCM a cui risulta connesso.

Il *Floor Control Module*, come detto, è il dispositivo responsabile della supervisione di tutta l'elettronica di piano e della comunicazione con il laboratorio di terra. Con esso risulta connesso mediante due fibre ottiche che partendo dall'FCM arrivano al concentratore di torre TJB e da lì vengono instradate verso terra nel modo descritto in precedenza. Sul piano l'FCM, oltre che con l'SCM, è connesso con i quattro moduli ottici. Tutte le connessioni sono effettuate mediante cavi in rame dai quali l'FCM fornisce anche la corrente necessaria ai dispositivi per funzionare. Dal punto di vista strettamente operativo l'FCM può essere considerato un componente passivo nel senso che non effettua direttamente misure e quindi non produce dati fisici. Tuttavia esso svolge un compito altrettanto fondamentale per la riuscita

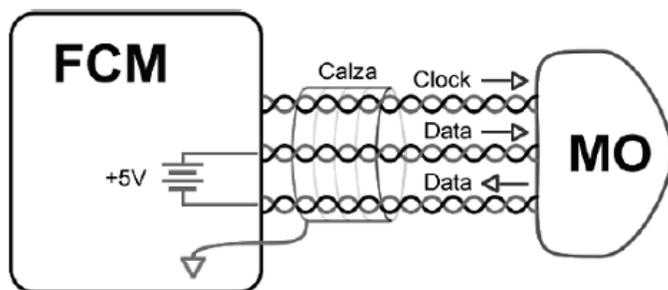


Figura 3.11: Collegamento tra FCM e MO.

dell'esperimento poichè da un lato riceve i dati provenienti dal laboratorio e li smista verso tutti i dispositivi elettronici presenti sul piano e dall'altro riceve i dati fisici delle misure effettuate dai MO e dall'SCM e li invia a terra inserendoli in una trama ottica. E' facile intuire quindi che il *Floor Control Module* è un dispositivo estremamente complesso. Esso infatti interpreta e formatta la trama di dati scambiati con gli apparati del laboratorio di terra, controlla la trasmissione correggendo eventuali errori e provvede al controllo delle alimentazioni di tutti gli apparati del piano. Mediante l'FCM è infatti possibile accendere e spegnere tutta la strumentazione ad essa connessa.

Il collegamento tra l'FCM e l'SCM è effettuato mediante un bus SPI (*Serial Peripheral Interface*), mentre quello con ognuno dei quattro moduli ottici è realizzato attraverso un cavo di rame. Esso è formato da tre coppie di cavi schermati mediante una calza metallica (figura 3.11). Sulla prima coppia viene inviato il segnale di clock mentre gli altri due servono per la comunicazione dei dati. Tutte le informazioni sono scambiate mediante lo standard elettrico differenziale LVDS (*Low Voltage Differential Signaling*) [22]. La scelta è ricaduta su questo standard poichè permette di pilotare segnali anche a grandi distanze garantendo sempre un'ottima reiezione del rumore. Il modo comune delle due coppie di cavi che trasportano dati viene anche utilizzato per la trasmissione della tensione di alimentazione dei moduli ottici.

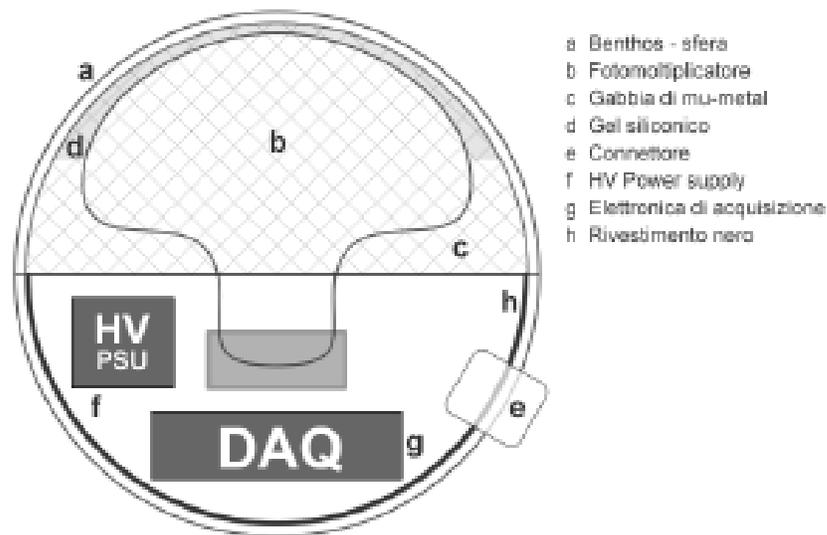


Figura 3.12: Il modulo ottico del telescopio NEMO.

3.4.3 Il modulo ottico

Ogni modulo ottico (figura 3.12) è costituito da una sfera di vetro borosilicato in grado di sopportare le enormi pressioni a cui è sottoposta nelle grandi profondità marine pur mantenendo ottime caratteristiche di trasmissione della luce. Il suo diametro esterno misura 43 cm ed il vetro ha uno spessore di 15 mm. La sfera, denominata Benthos-sfera, è formata da due emisferi che una volta chiuse sono in grado di resistere ad una pressione di 700 bar. Una delle due emisferi che costituiscono il modulo ottico è trasparente mentre l'altra è rivestita da una guaina nera che non permette il passaggio della luce. In tal modo si ha la certezza che il PMT riveli solo la luce proveniente dalla direzione in cui è orientato migliorando le caratteristiche di direzionalità del sensore. Il fotomoltiplicatore presente nel modulo ottico è il modello R7081 prodotto dalla ditta giapponese HAMATSU (figura 3.13). Esso ha un'area sensibile di 830 cm², 10 stadi di moltiplicazione ed un TTS (*Transit Time Spread*) pari a ~ 1.5 ns *full width at half maximum*. La tensione nominale di funzionamento è pari a 1900 V. La parte sensibile del PMT



Figura 3.13: Il PMT del telescopio NEMO.

è attaccata alla semisfera trasparente del modulo ottico mediante un gel a base di silicone che garantisce l'accoppiamento ottico tra il fotocatodo ed il vetro della sfera. Tale gel viene deposto in forma liquida ma a temperatura e pressione ambiente vulcanizza facendo aderire le due superfici. Per schermare il fotocatodo dagli effetti del campo magnetico terrestre il PMT è racchiuso da una gabbia di μ -metal. Nella parte opaca della Benthos-sfera trovano posto l'elettronica di *Front End* del fotomoltiplicatore, denominata DAQ (*Data AcQuisition*) board ed il modulo di alimentazione del PMT. Il compito principale della DAQ board è quello di acquisire i segnali analogici provenienti dal fotomoltiplicatore, di convertirli in digitale e di trasmetterli verso terra tramite l'FCM. La DAQ inoltre alimenta e controlla il modulo di generazione dell'alta tensione impostando quindi la tensione di lavoro del PMT. Infine la DAQ svolge una serie di funzioni di monitoraggio ambientale acquisendo i dati da sensori di temperatura e umidità presenti nel modulo ottico.

3.5 Il sistema di trasmissione dati

Il problema della trasmissione dati riveste un ruolo fondamentale nella realizzazione di un telescopio sottomarino per neutrini. Per questo motivo, in fase di progettazione sono state affrontate una serie di problematiche relative alla sincronizzazione dei dati inviati a terra dalle varie componenti dell'apparato. La sincronizzazione è un requisito fondamentale per i dati poichè da essa dipende la corretta ricostruzione delle tracce dei muoni acquisite dai fotomoltiplicatori.

In generale le caratteristiche richieste al sistema di trasmissione dati del telescopio NEMO sono le seguenti:

- perfetto sincronismo tra i dati inviati a terra da tutti i piani;
- grande ampiezza di banda nel canale monodirezionale per il trasferimento dei dati acquisiti dai PMT;
- possibilità di scambiare in modo bidirezionale dati di controllo tra il telescopio e la base di terra.

3.5.1 Lo standard SDH

Le specifiche richieste al sistema di comunicazione possono essere soddisfatte impiegando un protocollo standard denominato SDH (*Synchronous Digital Hierarchy*). Tale protocollo definisce il livello fisico del canale di trasmissione specificandone il metodo di formattazione e trasmissione dei dati in modo sincrono. Essendo poi tale protocollo largamente utilizzato nel campo delle telecomunicazioni è possibile reperire sul mercato una grande quantità di componenti commerciali prodotti in conformità a questo standard. Lo standard SDH comprende vari protocolli che differiscono tra loro per la velocità e la larghezza della banda di trasmissione. Tra le varie possibilità quello che soddisfa le specifiche del telescopio NEMO risulta lo standard STM-1. Esso, infatti, offre la possibilità di trasferire su un canale di comunicazione bidirezionale sincrono un flusso dati di circa 155 Mbps. Tale banda risulta essere

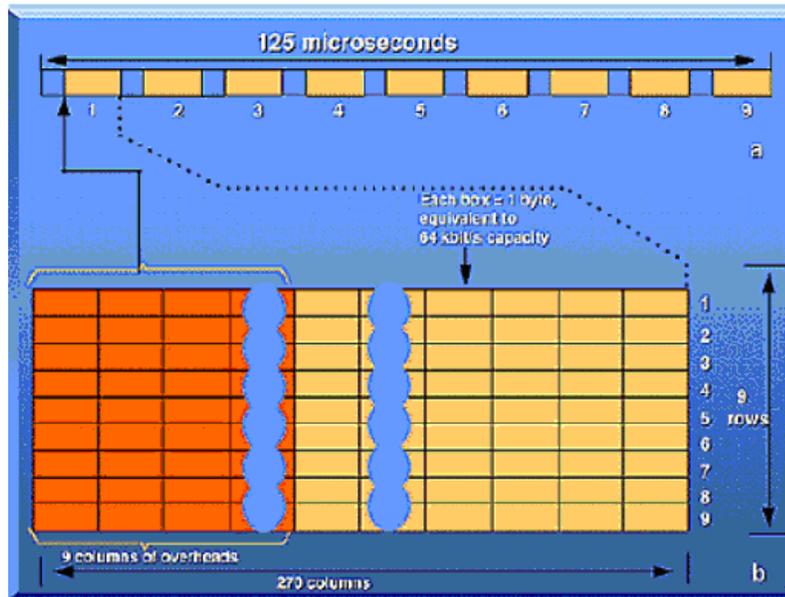


Figura 3.14: Dettaglio della trama dei dati secondo il protocollo SDH.

sufficiente per gestire sia i dati scientifici che i flussi relativi ai parametri di controllo del telescopio.

Lo schema di trasmissione di tale protocollo è identificato come una struttura ripetitiva caratterizzata da un periodo di $125 \mu s$. All'interno di un periodo la trama è suddivisa in 9 segmenti di lunghezza uguale (figura 3.14). Di essi la prima parte contiene il cosiddetto *overhead* cioè una serie di informazioni utilizzate per l'allineamento ed il controllo dell'integrità dei dati trasmessi. La restante parte di ogni riga è dedicata al *payload* e rappresenta la parte che viene riempita con le informazioni effettive da trasferire. Il sistema di comunicazione prevede che i nove segmenti vengano inviati in maniera sequenziale l'uno dopo l'altro.

3.5.2 Lo standard DWDM

Lo standard DWDM (*Dense Wavelength Division Multiplexing*) permette di inviare contemporaneamente, sulla stessa fibra ottica, più flussi di dati distinti tra loro. Questa tecnica prevede, infatti, di immettere sulla fibra le diverse

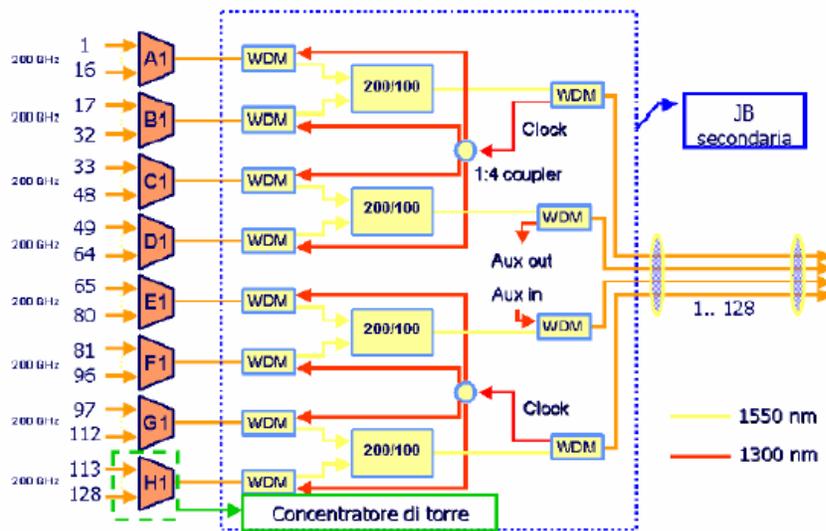


Figura 3.15: Schema della concentrazione dei segnali dai piani della torre alla SJB mediante lo standard DWDM.

trame a diverse lunghezze d'onda della luce. In questo modo ogni flusso inviato sarà caratterizzato e distinto dagli altri per mezzo di un diverso colore realizzando il cosiddetto *frequency multiplexing* ottico. Nell'architettura di trasmissione dati del telescopio NEMO si fa grande uso di questa tecnica per ridurre drasticamente il numero di fibre ottiche da utilizzare. Il primo livello di *multiplexing* ottico viene effettuato a livello della torre dove i 16 canali originatisi dai diversi piani sono inviati su un'unica fibra. Tale concentrazione di flussi di dati viene effettuata dal concentratore di torre come mostrato in figura 3.15. Le lunghezze d'onda utilizzate da ogni piano sono contenute nell'intervallo compreso tra 1520 nm e 1560 nm e sono spaziate tra loro di 200 GHz. Un secondo livello di concentrazione viene effettuato nelle SBJ. Esse convogliano i dati provenienti da 9 torri e li inviano verso la JB primaria che provvede alla connessione con il cavo elettro-ottico principale. Una volta che i dati raggiungono la stazione di raccolta a terra vengono nuovamente separati mediante una serie di *demultiplexer*. Tali dispositivi permettono di suddividere i dati provenienti dai diversi piani di tutte le torri del telescopio.

Una volta operata questa suddivisione i dati vengono raccolti da una scheda gemella dell'FCM che si trova sul piano della torre. La board FCM di terra è in grado di interconnettersi ad una rete di calcolatori che a questo punto provvedono all'immagazzinamento ed all'analisi dei dati. A questo punto è possibile iniziare la procedura di ricostruzione delle tracce.

Capitolo 4

Codici per il controllo e la correzione di errori nella trasmissione dati

4.1 Introduzione

Per quanto detto nei capitoli precedenti risulta chiaro che la riuscita di un esperimento come quello NEMO, sia basata in larga misura sull'affidabilità delle comunicazioni e del trasferimento dati. Ciò è dovuto innanzitutto alla filosofia di funzionamento di tale apparato che prevede il campionamento in acqua dei segnali acquisiti dai fotomoltiplicatori ed il loro trasferimento a terra dove avviene la ricostruzione delle tracce generate dai muoni. Maggiore è l'affidabilità del trasferimento dati maggiore sarà la probabilità di ricostruzione esatta delle tracce e quindi la probabilità di riuscita dell'esperimento. L'affidabilità delle comunicazioni è resa estremamente complessa dato il particolare ambiente in cui il telescopio si trova ad operare. Esistono, infatti, due tipi di potenziali problemi relativi alle comunicazioni in un apparato di 1 Km^3 posto in acqua a profondità di circa 3000 metri: il primo riguarda la comunicazione tra i vari dispositivi in acqua mentre il secondo coinvolge lo scambio di informazioni tra il telescopio in mare e la stazione di controllo a terra. Per quanto concerne il primo punto è necessario tenere conto che trasferire segnali su rame ad una frequenza molto elevata per distanze dell'ordine della decina di metri (distanza tra Front End e FCM su

di un piano) risulta già complesso a terra. Nel caso del telescopio NEMO tale difficoltà è accentuata dalla pressione che viene esercitata sui cavi e sugli elementi coinvolti nella trasmissione dati. Lo scambio di informazioni con il laboratorio di terra soffre ovviamente di problematiche analoghe aggravate dal fatto che le distanze in gioco sono molto più elevate (nell'ordine del centinaio di chilometri). Ovviamente, come detto nel precedente capitolo, in fase di progetto sono stati impiegati una serie di sistemi in grado di diminuire la probabilità di errori di trasmissione come, ad esempio, l'impiego di standard di comunicazione che hanno un'elevata immunità al rumore. Chiaramente per quanto molto affidabili tali tecniche non risultano infallibili. Per questo motivo il gruppo di Napoli ha pensato di progettare un sistema che possa aumentare l'affidabilità di trasmissione dati. Come vedremo il lavoro, che mi ha visto impegnato in prima persona, è stato svolto in due fasi. Dapprima c'è stata una fase di studio del problema volta all'individuazione del tipo più probabile di errori a cui le comunicazioni erano esposte. Tale fase ha condotto all'individuazione della tecnica migliore da impiegare per risolvere il particolare problema. La seconda fase ha riguardato la realizzazione ed il collaudo del dispositivo che implementasse l'algoritmo precedentemente individuato. E' importante sottolineare, infine, che ci si è approcciati al problema tentando di realizzare un dispositivo che fosse il più possibile di uso generale. Esso infatti da un lato doveva soddisfare pienamente le specifiche di progetto dettate dal particolare utilizzo nell'ambito del telescopio NEMO e dall'altro doveva essere realizzato in modo da risultare del tutto generale e poter essere utilizzato anche in ambiti differenti.

In questo capitolo verrà effettuata dapprima una breve introduzione alla teoria alla base dei codici a controllo e correzione d'errore mettendone in luce i concetti base. Tale parte introduce all'esposizione dell'algoritmo individuato per l'esperimento NEMO.

4.2 Trasmissione dati su canali rumorosi

Ogni volta che si ha a che fare con il trasferimento di informazioni di qualunque tipo si pone il problema di comprendere quanto il canale usato per

tale trasmissione sia affidabile cioè quanta parte di informazione inviata da un trasmettitore arriva al ricevitore in maniera corretta e comprensibile.

Consideriamo alcuni esempi di trasferimento di informazioni su canali rumorosi:

- una linea telefonica su cui avviene la comunicazione tra due utenti;
- il canale radio di comunicazione tra un satellite nello spazio e la stazione di controllo a terra;
- la riproduzione cellulare durante la quale le informazioni del DNA della cellula madre vengono trasferite alla cellula figlia;
- le operazioni di lettura e scrittura dal disco fisso di un computer.

Com'è facile intuire dagli esempi descritti il problema della trasmissione dati è del tutto generale ed è comune a molti ambiti completamente diversi tra loro. Esso è, infatti, indipendente dal meccanismo e dal mezzo di trasmissione nonché dalle informazioni trasferite. In particolare l'ultimo caso considerato ci fa capire come questo non risulta legato necessariamente al trasferimento di informazioni da un luogo ad un'altro. Quando vengono scritte informazioni in alcune locazioni di un disco fisso esse vengono rilette dalle stesse locazioni in un istante di tempo successivo.

I canali individuati in precedenza come esempio sono intrinsecamente rumorosi. Le linee telefoniche soffrono di rumore indotto dal trasferimento dati su linee vicine (il cosiddetto *cross-talk*). Inoltre le conversioni effettuate per la trasmissione dei segnali (cambiamento di standard o conversioni analogico-digitale) distorcono e aggiungono rumore al segnale trasmesso. Lo scambio di informazioni tra un satellite artificiale e la stazione di controllo a terra viene disturbato dalle onde elettromagnetiche generate da sorgenti galattiche e terrestri. Il DNA delle cellule è soggetto a mutazioni e danneggiamenti dovuti a diversi fattori. Il disco fisso di un computer, nel quale vengono memorizzate informazioni binarie allineando parti di materiale magnetico in una delle due direzioni, può indurre errori durante la riletture dovute ad una serie di fattori diversi: cambiamento spontaneo dello stato di magnetizzazione di una parte

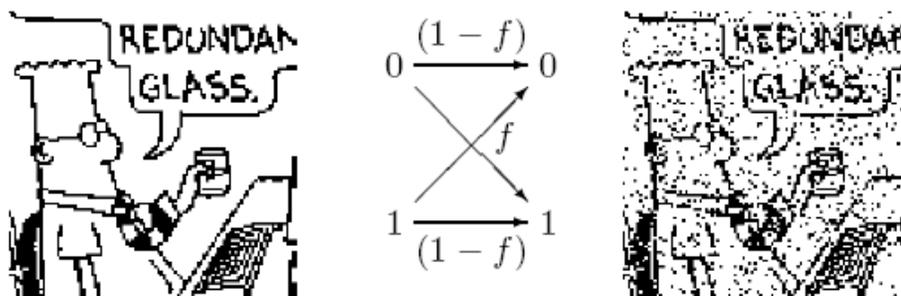


Figura 4.1: Rappresentazione di una comunicazione su un canale con livello di rumore $f = 0.1$. [23]

del materiale di cui è composto il disco, errori indotti sulla testina di lettura dall'ambiente circostante oppure errori di magnetizzazione nella scrittura di un bit dovuti alla magnetizzazione di settori vicini del disco. In tutti questi casi esiste una probabilità non nulla che il messaggio ricevuto non risulti identico a quello inviato. Ovviamente lo scopo di una trasmissione corretta è quello di avere un canale di trasmissione in cui la probabilità di errore sia zero o così vicina a zero che per gli scopi pratici sia indistinguibile da essa. Se si torna all'esempio del disco fisso di un computer si può definire come $1 - f$ la probabilità di trasferire ogni bit correttamente e, ovviamente, con f la probabilità di trasferirlo in modo errato. Tale modello di comunicazione, applicabile ad ogni trasmissione digitale, è noto come canale binario simmetrico (figura 4.1) [23].

Se, ad esempio, si immagina di avere per un ipotetico canale di trasmissione $f = 0.1$ si avrà che il 10% del totale dei bit trasmessi correttamente dalla sorgente assumerà un valore sbagliato una volta arrivato a destinazione. Da un canale di comunicazione ci si aspetta, ovviamente, che non vengano mai indotti errori di trasmissione. Se infatti si ipotizza di voler trasmettere su un canale 1 gigabyte di dati al giorno per un anno, un canale di trasmissione affidabile dovrebbe avere una probabilità di fallimento dell'ordine di 10^{-15} o inferiore. A questo punto ci sono due strade possibili per raggiungere quest'obiettivo di affidabilità: una soluzione fisica e una soluzione che può

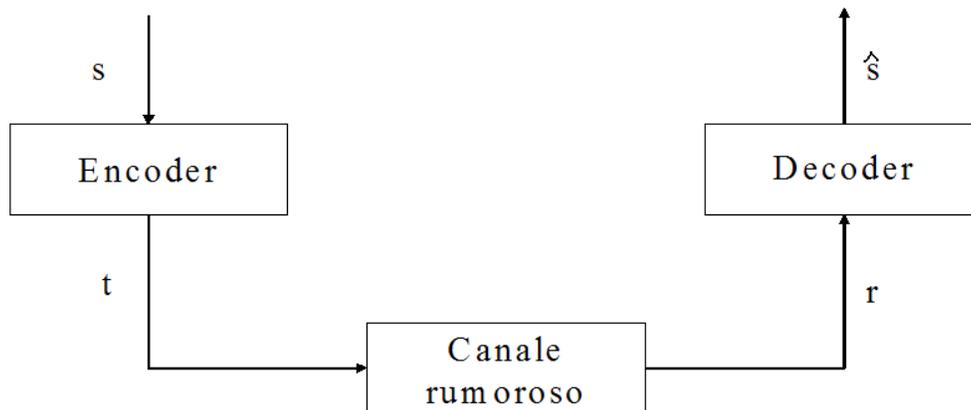


Figura 4.2: Soluzione di sistema per migliorare l'affidabilità di un canale di comunicazione.

essere definita di sistema.

4.2.1 La soluzione fisica

La soluzione fisica consiste nel migliorare le caratteristiche fisiche del canale di trasmissione in modo da ridurre la probabilità di fallimento. Se si considera l'esempio della trasmissione di dati digitali su una linea di trasmissione elettrica è possibile migliorare la comunicazione scegliendo componenti elettronici migliori, utilizzare cavi di comunicazione più affidabili, schermare gli stessi in modo da ridurre il rumore indotto, utilizzare protocolli di comunicazione con elevata immunità al rumore. Tali soluzioni generalmente fanno aumentare i costi sia in termini economici che in termini di dissipazione di potenza. Inoltre non è sempre possibile applicare queste soluzioni dato il particolare ambiente in cui si va ad operare. Se si pensa al caso di NEMO la scelta dei cavi di comunicazione, ad esempio, è soggetta, oltre che a motivi di affidabilità della comunicazione, anche a criteri di selezione dettati dalla resistenza alla pressione a cui essi saranno soggetti una volta posti in acqua.

4.2.2 La soluzione di sistema

La teoria dell'informazione e dei codici offre un approccio alternativo e per certi versi più furbo. Esso consiste non tanto nel migliorare il canale fisico tentando di aumentarne l'affidabilità utilizzando componenti migliori ma adoperando delle tecniche di sistema che permettano la correzione di eventuali errori. In definitiva con quest'approccio si accetta la rumorosità del canale e si aggiungono sistemi in grado di rivelare gli errori indotti dal canale. Una volta che l'informazione è arrivata a destinazione tali errori vengono individuati e corretti. Queste soluzioni di sistema prevedono di aggiungere dispositivi di codifica e decodifica dell'informazione posti all'inizio ed alla fine dei canali di trasmissione come schematizzato in figura 4.2. In questo modo il codificatore riceve le informazioni (s) da trasmettere e le codifica, secondo un qualche algoritmo, creando così l'informazione (t) da inviare. Il canale aggiunge a tale informazione del rumore e quindi al decodificatore arriverà un'informazione potenzialmente errata (r). Il decodificatore attraverso una serie di operazioni riesce a riottenere l'informazione separando i dati inviati dall'errore indotto durante la trasmissione. Si ha in questo modo il messaggio (\hat{s}) teoricamente identico a quello (s) inviato dalla sorgente. E' possibile affermare che soluzioni di questo tipo trasformano canali rumorosi in canali affidabili aggiungendo solo codificatori e decodificatori ed eventualmente aumentando il tempo di trasmissione. Bisogna tener conto tuttavia che è impossibile effettuare operazioni che permettano l'individuazione e la correzione di errori basandosi solo sull'informazione trasmessa. E' necessario, infatti, aggiungere al dato da trasmettere una serie di dati aggiuntivi che permettano di effettuare le operazioni descritte. Tale informazione aggiuntiva viene denominata ridondanza e viene inserita nella comunicazione durante le operazioni di codifica. Il decodificatore quindi riceve il dato e le informazioni ridondanti e grazie ad esse riesce a correggere eventuali disturbi. Quando si utilizzano soluzioni di questo tipo è necessario selezionare l'algoritmo di codifica e decodifica da utilizzare in modo da aumentare il più possibile l'affidabilità del sistema cercando al contempo di contenere al massimo le informazioni ridondanti ed il tempo di elaborazione necessario per le

operazioni da effettuare.

Tornando alla comunicazione digitale è facile intuire che in questo caso la ridondanza è rappresentata da una serie di bit aggiuntivi inviati assieme ai bit di dato.

4.3 Codici a controllo e correzione d'errore

Il continuo progredire della tecnologia applicata soprattutto ai campi dell'automazione industriale e delle telecomunicazioni nonché il velocissimo sviluppo di sistemi elettronici digitali sempre più complessi ha favorito la creazione e lo sviluppo di algoritmi per l'individuazione e la correzione di errori indotti durante la trasmissione dati. Questi infatti vengono oggi usati largamente in campi molto differenti tra loro e quindi esiste in letteratura un numero elevato di codici volti al controllo ed alla correzione d'errore. Tra essi, molti sono stati sviluppati ed ottimizzati per la trasmissione di segnali digitali. Tali codici infatti sono sempre più diffusi non solo in applicazioni di nicchia ma sono presenti sui bus di comunicazione dei normali personal computers di uso domestico oppure nei dispositivi di lettura e riproduzione di CD e DVD. In quest'ultimo caso la comunicazione è spesso resa molto disturbata dal facile deterioramento dei supporti che contengono le informazioni e quindi è spesso molto elevata la probabilità di fallimento della trasmissione. In tali ambiti è spesso richiesta anche una certa affidabilità relativa alla velocità di trasmissione dati. Spesso infatti l'informazione non solo deve arrivare integra a destinazione ma il flusso di dati deve risultare costante (si pensi alla musica riprodotta da un CD). Questo enorme sviluppo e la grandissima varietà di applicazioni ha portato alla produzione di un elevato numero di codici. Di seguito ne vengono riportati alcuni semplici esempi.

4.3.1 Codici per la rivelazione di errori

Lo sviluppo dei codici per il controllo della trasmissione dati risale a molti anni fa quando si è diffusa l'automazione dei processi industriali e la nascita dei primi calcolatori progenitori degli attuali computer. Già negli anni 50 si utilizzavano algoritmi matematici in grado di controllare il trasferimento

di dati. I primi algoritmi, tuttavia, non erano molto evoluti e consentivano solo la rivelazione degli errori avvenuti durante la trasmissione ma non erano in grado di effettuare alcuna correzione. Contrariamente a quanto si può pensare tali algoritmi non risultano del tutto obsoleti ma vengono ancora oggi utilizzati in alcuni campi. Essi infatti hanno il vantaggio di essere molto semplici da sviluppare in hardware o in software e hanno una velocità di esecuzione molto elevata. Un'ulteriore vantaggio consiste nel fatto che essi non richiedono un numero elevato di bit di ridondanza e quindi l'informazione trasferita lungo il canale di comunicazione è rappresentata per la quasi totalità da dati utili. Ovviamente il difetto principale risiede nel fatto che il loro utilizzo consente solo di accorgersi che l'informazione arrivata a destinazione non è identica a quella inviata ma non è in grado in alcun modo di segnalare quale bit trasferito sia errato. Se si utilizzano codici di questo tipo, in caso di rivelazione di errori, non si può far altro che eliminare l'informazione trasferita in modo errato e richiedere alla sorgente di inviare nuovamente il messaggio. Chiaramente in questo caso i tempi di trasferimento dati possono crescere eccessivamente ed inoltre non si ha la certezza che il processo di trasferimento vada a buon fine. In caso di un canale molto rumoroso può, infatti, capitare che l'informazione non arrivi mai alla sorgente in modo corretto e l'algoritmo non potrà fare altro che richiedere una nuova trasmissione innescando un processo che non ha termine.

4.3.2 Il controllo di parità

Il più noto algoritmo per la rivelazione di errori di trasmissione è il controllo di parità. Esso consiste nell'aggiungere all'informazione da trasferire un bit che assume valore 1 o 0 a seconda che nella stringa di dati ci siano un numero pari o dispari di bit con valore uno. In maniera del tutto analoga è possibile utilizzare una delle due convenzioni che vengono dette parità pari e parità dispari. Se si suppone di utilizzare l'algoritmo di parità pari il bit di ridondanza o di parità (che di solito viene posto al termine della stringa) assume valore 1 se il numero totale di bit con valore uno nella stringa è dispari. In caso contrario, e cioè se il numero di uno nella stringa è pari, il bit di parità

assume valore 0. In definitiva dopo l'applicazione di tale algoritmo il numero totale di bit ad uno nell'informazione effettiva da trasferire (dato più bit di parità) deve essere sempre un numero pari. In questo modo una volta che l'informazione arriverà a destinazione sarà necessario contare il numero di bit con valore uno e controllare che tale numero sia pari. In caso contrario si è certi che almeno un bit ha cambiato valore durante la trasmissione.

Quanto detto può risultare più chiaro per mezzo di un esempio. Supponiamo di voler trasferire la seguente informazione binaria:

1110010

dato che nella stringa sono presenti un numero pari di bit con valore 1 (in totale sono presenti 4 bit ad 1) allora l'algoritmo di parità prevede di terminare la stringa aggiungendo in coda un bit con valore 0. In definitiva la stringa effettiva da trasferire risulterà:

11100100

Al ricevitore verrà quindi inviata una stringa di bit che contiene un bit ridondante che è appunto il bit di parità. A questo punto nella stringa inviata sarà presente un numero pari di bit con valore 1.

Supponiamo adesso che durante la trasmissione a causa del rumore sulla linea il secondo bit a partire da sinistra muti il suo valore da 1 a 0. Al ricevitore giungerà il seguente messaggio:

10100100

Il ricevitore, applicando l'algoritmo si accorgerà che nella stringa il numero di bit con valore 1 è pari a tre. Il controllo di parità sarà quindi fallito e la trasmissione verrà giudicata affetta da errore. A questo punto si potrà richiedere di inviare nuovamente l'informazione.

Questo esempio mette in luce l'estrema semplicità dell'algoritmo di parità. Le operazioni che il codificatore ed il decodificatore devono compiere sono infatti banali. Entrambi i dispositivi non devono far altro che contare il numero di bit pari ad uno nella stringa. E' chiaro quindi che sia in termini di

risorse impiegate che in termini di tempo di calcolo tale algoritmo risulta particolarmente vantaggioso. Un'analisi più approfondita di tale metodo mette in luce una serie di difetti. In particolare è facile notare che la parità risulta molto efficace quando in trasmissione si verifica un solo errore; tuttavia se si verifica il cambiamento simultaneo di due (o in generale di un numero pari) di bit il controllo effettuato dal decodificatore non porta alla rivelazione di alcun errore. La probabilità che si verifichino un numero multiplo di errori, una volta fissata la probabilità di fallimento di trasmissione su una linea, dipende solo dal numero di bit trasferiti. Se quindi si tenta di rendere più efficiente l'algoritmo di parità aumentando il numero di bit di dato controllati da un unico bit di parità si può arrivare al paradosso di aumentarne la probabilità di fallimento. Un'altro grosso limite di questo algoritmo risiede nel fatto che non è previsto alcun controllo sul bit di parità. Ciò comporta che, se durante la trasmissione, proprio l'ultimo bit subisce una variazione di stato logico, la stringa in arrivo al decodificatore viene giudicata errata pur non essendoci alcun errore nella parte di dato trasferita.

4.3.3 Codici per la rivelazione e la correzione di errori

Per tentare di superare i limiti evidenziati dagli algoritmi per la rivelazione di errori sono stati messi a punto nel tempo una serie di codici in grado non solo di individuare genericamente errori in trasmissione ma anche in grado di individuare in quale punto della stringa fosse avvenuto il cambiamento di bit. Tali informazioni nelle trasmissioni digitali sono sufficienti ad operare il cambiamento del bit errato e quindi la correzione della stringa ricevuta.

4.3.4 Codici a ripetizione d'informazione

Sotto il nome di codici a ripetizione d'informazione vengono classificati tutti quegli algoritmi che prevedono di trasferire un certo numero di copie identiche del dato da comunicare. Il più comune prevede di trasferire tre copie identiche della stessa informazione. In questo codice, generalmente indicato con R_3 , se si vuole trasferire un bit di dato ad esempio:

1

la stringa effettivamente trasferita sarà:

111

E così se si vuole trasferire la stringa

111010

si invierà

$$\begin{array}{cccccc} 1 & 1 & 1 & 0 & 1 & 0 \\ \underbrace{\hspace{1em}} & \underbrace{\hspace{1em}} & \underbrace{\hspace{1em}} & \underbrace{\hspace{1em}} & \underbrace{\hspace{1em}} & \underbrace{\hspace{1em}} \\ 111 & 111 & 111 & 000 & 111 & 000 \end{array}$$

Il ricevitore si accorgerà che sulla linea sono occorsi degli errori se i bit arrivati a destinazione non saranno a tre a tre identici.

Se infatti verrà ricevuta la stringa:

101 111 110 000 111 100

sarà immediato rilevare che essa è stata affetta sicuramente da una serie di errori.

Come detto, con tale algoritmo è possibile però anche tentare di effettuare una correzione dell'informazione utilizzando un meccanismo detto di maggioranza. Tale metodo consiste, in caso di non uguaglianza dei bit nella tripletta, di considerare corretta l'informazione contenuta nei due bit uguali.

Se si esamina la stringa precedente e si applica la maggioranza l'informazione ricostruita sarà:

$$\begin{array}{cccccc} \underbrace{101} & \underbrace{111} & \underbrace{110} & \underbrace{000} & \underbrace{111} & \underbrace{100} \\ 1 & 1 & 1 & 0 & 1 & 0 \end{array}$$

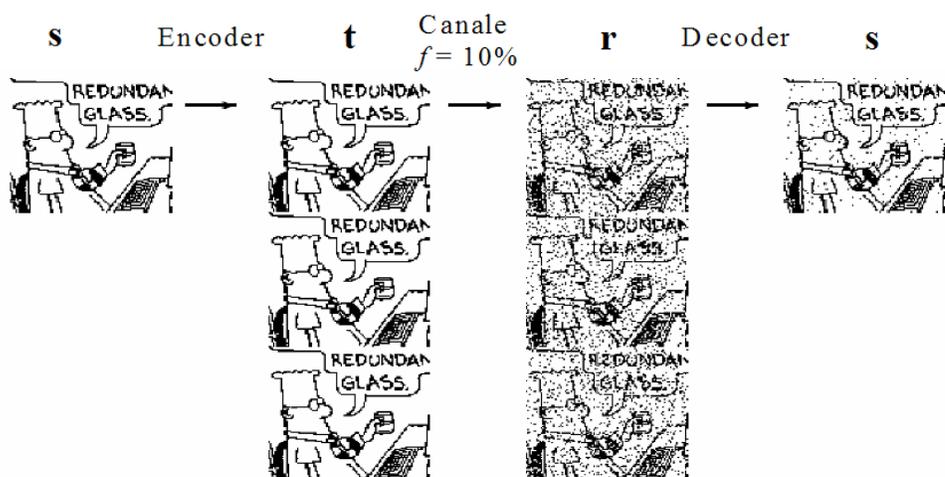


Figura 4.3: Trasmissione su un canale di comunicazione con $f=0.1$ e utilizzando il codice a correzione R_3 .

In questo caso l'informazione ricostruita risulta corretta ed identica a quella inviata.

Il meccanismo di maggioranza si basa sull'assunto che ogni tripletta di bit sia affetta al più da un errore. In altre parole si suppone che la probabilità che una tripletta sia affetta al più da un bit di errore risulta molto superiore alle altre che prevedono due o addirittura tre errori contemporanei. Chiaramente esistono una serie di casi in cui l'algoritmo di maggioranza non è in grado di ricostruire perfettamente la stringa affetta da errore.

In termini analitici si può affermare che la probabilità di fallimento del codice a ripetizione d'informazione è legata alla probabilità che due dei tre bit della tripletta cambino simultaneamente. Tale probabilità scala come f^2 e quindi se si sta trasmettendo su un canale binario simmetrico con probabilità di fallimento $f = 0.1$; il codice R_3 ha una probabilità di fallimento dopo la decodifica e l'applicazione dell'algoritmo di maggioranza pari a $p_b \simeq 0.03$ per bit. In figura 4.3 viene mostrato il risultato della trasmissione di un'immagine binaria attraverso un canale binario simmetrico usando un codice a ripetizione d'informazione R_3 .

Il principale limite dei codici a ripetizione risiede tuttavia nel grande

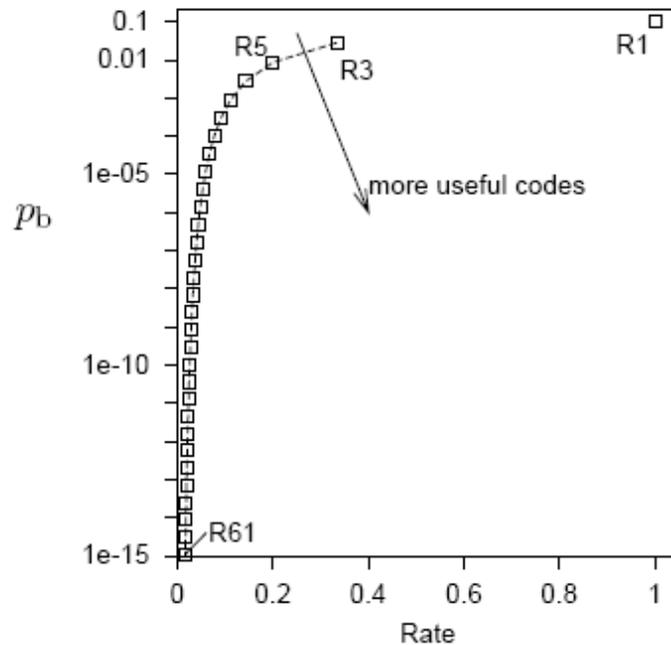


Figura 4.4: Grafico della probabilità di errore p_b dopo l'applicazione del codice a ripetizione di informazione in funzione della rate calcolati su un canale avente $f=0.1$.

numero di informazioni ridondanti da dover inviare. Nel caso in esempio la rate di dati è diminuita di un fattore 3. Se quindi si usa un codice di questo tipo si abbatta la probabilità di errore ma si paga un prezzo molto alto in termini di frequenza di trasferimento dati. Basti pensare che con tale tecnica è necessario trasferire 3 gigabyte di dati rumorosi per ottenere 1 gigabyte nel quale la probabilità di errore è pari al 3%. Chiaramente è possibile diminuire la probabilità di fallimento aumentando il numero di ripetizioni di una singola informazione ma, come risulta chiaro dal grafico di figura 4.4, il numero di ripetizioni necessarie aumenta molto velocemente rendendo di fatto impossibile la trasmissione.

4.4 Come si valuta un codice

Una volta introdotti una serie di codici per la rivelazione e la correzione di errori in trasmissione viene spontaneo cercare un metodo efficace che permetta di confrontare e valutare i codici esistenti al fine di selezionare quello che meglio si adatta alle esigenze dell'utente. Tali esigenze derivano in ultima analisi dall'affidabilità che si cerca nella trasmissione su un canale e dal numero massimo di bit di ridondanza che si è disposti a trasmettere. Una volta in possesso di queste informazioni e dopo aver studiato la rumorosità del canale, cioè la probabilità di fallimento di una trasmissione, è possibile selezionare il codice che meglio si adatta al tipo di trasmissione da realizzare.

Per effettuare una prima valutazione dei codici è possibile introdurre due parametri fondamentali:

- il rapporto di ridondanza;
- la distanza di Hamming.

Il rapporto di ridondanza permette di confrontare diversi codici in base al numero di bit di ridondanza per ogni bit d'informazione. Se infatti si indica con K il numero di bit di informazione e con N il numero totale di bit da trasferire (dato più ridondanza) si definisce rapporto di ridondanza o *rate* di un determinato codice il rapporto:

$$R = \frac{K}{N}$$

Il rapporto di ridondanza fornisce immediatamente un primo parametro di valutazione che ci permette di capire, a parità di bit trasferiti, quanta informazione è possibile immettere su un canale. E' chiaro che con questo metodo di valutazione un codice è tanto migliore quanto più il suo rapporto di ridondanza tenda a 1. In questo caso infatti si avrebbe una massimizzazione delle informazioni trasportate. Per quanto detto in precedenza è evidente, però, che per i codici a correzione d'errore è impossibile ottenere $R=1$ poichè in tal caso tutta l'informazione inviata sarebbe costituita da dati. In altre parole in un codice con $R=1$ non vengono inviati bit di ridondanza. Ciò in effetti costituisce un controsenso poichè sappiamo che senza

informazioni ridondanti è impossibile effettuare alcun tipo di algoritmo per il controllo o la correzione di errori. Detto ciò è possibile valutare i due codici precedentemente descritti in base a tale parametro.

Il codice a ripetizione d'informazione R_3 ha un rapporto di ridondanza pari a:

$$R_{R_3} = \frac{1}{3} = 0.3$$

e quindi risulta particolarmente inefficiente. Considerando poi che 3 rappresenta il numero minimo di ripetizioni da dover effettuare per permettere l'applicazione della correzione a maggioranza ci si accorge che il rapporto di ridondanza per i codici a ripetizione d'informazione può solo peggiorare.

Il codice con controllo di parità è, invece, molto particolare. Esso, infatti, prevede di mantenere fisso, e pari ad 1, il numero di bit di ridondanza indipendentemente da quanti siano i bit d'informazione da inviare. Per questo motivo si può ottenere una rate arbitrariamente vicina ad 1 facendo crescere il numero di bit di dato da inviare. Se infatti si decide di inviare parole di 4 bit si ha:

$$R_{p4} = \frac{4}{5} = 0.8$$

ma ovviamente se il numero di bit d'informazione passa ad 8 si ha:

$$R_{p8} = \frac{8}{9} = 0.88$$

Generalizzando si può affermare che per il codice a controllo di parità si ha:

$$R_p = \frac{K}{N} = \frac{K}{K+1}$$

A questo punto viene spontaneo pensare che, per aumentare l'efficienza del codice, basta aumentare il numero di bit trasportati ovvero far crescere K . Chiaramente questo discorso è errato poichè non si tiene conto della rumorosità del canale. Infatti, aumentando il numero di bit di dato, a parità

di rumorosità del canale, si aumenta anche la probabilità di ottenere errori multipli durante la trasmissione e quindi la probabilità di fallimento del codice.

Si è quindi giunti ad un paradosso: se consideriamo solo il rapporto di ridondanza il codice a controllo di parità risulta molto più efficiente di quello a ripetizione d'informazione; tuttavia sappiamo che è vero il contrario perchè il controllo di parità fallisce in caso di errori multipli e soprattutto non permette la correzione dell'informazione arrivata a destinazione. Risulta quindi evidente che la *rate* di un codice non può essere considerata l'unico parametro nella valutazione di un codice. E' necessario quindi introdurre anche un valore che misuri il potere di rivelazione e di correzione di errori in un codice. Per questo motivo si definisce la distanza di Hamming.

La distanza di Hamming individua il numero minimo di bit che devono cambiare in una stringa per passare da una parola di codice ad un'altra. In pratica la distanza di Hamming indica quanti bit devono essere affetti da errore affinché il codice non sia più in grado di rivelare l'errore. In maniera più formale è possibile definire questo parametro considerando che le stringhe di bit di un determinato codice possono essere rappresentate in uno spazio N -dimensionale dove N rappresenta il numero totale di bit che per quel determinato codice è necessario trasferire. A questo punto la distanza di Hamming può essere vista come il numero minimo di coordinate in cui differiscono due parole dello stesso codice. In questo modo la definizione di distanza assume un significato geometrico più rigoroso poichè è possibile dimostrare che in uno spazio vettoriale così definito la distanza di Hamming rappresenta la metrica dello spazio stesso.

Per il codice a ripetizione R_3 la distanza di Hamming risulta ovviamente pari a 3 poichè questo è il numero minimo di bit che devono cambiare affinché il codice non si accorga di un errore in trasmissione. Se infatti cambiassero solo uno o due bit, al ricevitore arriverebbe una stringa non consentita dal codice e quindi ci si accorgerebbe dell'errore.

La distanza di Hamming per il codice a controllo di parità è meno immediata da calcolare. Sapendo però che il codice fallisce quando si verificano almeno due errori allora per questo codice la distanza è pari a 2.

La distanza di Hamming, come è facile notare, non fornisce informazioni sulla possibilità di un codice di correggere errori ma solo di rivelarli. Tale informazione tuttavia risulta strettamente connessa alla distanza di Hamming. Se infatti indichiamo con d la distanza è possibile definire il potere di correzione di un codice come:

$$t = \frac{d - 1}{2}$$

In questo modo t rappresenta il minimo numero di bit affetti da errore che il codice riesce a correggere. A questo punto si ha che:

$$t_{R_3} = 1$$

mentre

$$t_p = \frac{1}{2}$$

Chiaramente, per come è definito, t deve essere un numero intero e quindi è chiaro che il codice a controllo di parità non è in grado di effettuare correzione di errori.

4.4.1 Il codice perfetto

Una volta definiti dei parametri di valutazione dei codici viene spontaneo chiedersi che caratteristiche deve avere il codice perfetto, ammesso che esso esista. Tale problema fu affrontato per la prima volta da Shannon nel 1948 e ancora oggi è oggetto di molte discussioni. Dalle definizioni date fino ora sembrerebbe che il codice perfetto sia quello che massimizza il numero di informazioni trasferite e che abbia una distanza di Hamming tale da rendere nulla la probabilità di errore su un canale arbitrariamente rumoroso. In effetti è possibile affermare che un codice di questo tipo non esiste. Nella selezione di un determinato algoritmo di correzione da applicare ad un canale reale è quindi necessario raggiungere un compromesso tra i parametri d e t ma bisogna anche tener conto del tempo massimo di esecuzione che si è disposti ad aspettare nonchè il tasso di errore minimo che è possibile tollerare. E' chiaro infatti che nei casi reali è inutile, nonchè praticamente impossibile,

progettare un codice che riduca a zero la probabilità di fallimento di una trasmissione.

4.5 Il codice di Hamming

Al termine di questa breve rassegna sulla teoria dei codici a correzione d'errore viene presentato il codice di Hamming. Esso, storicamente, rappresenta il primo codice efficiente in grado di correggere errori di trasmissione e, per le sue caratteristiche, tuttoggi trova ancora molti campi di applicazione. Tale codice risulta molto potente ma al contempo anche molto semplice da sviluppare in hardware. I tempo di codifica e di decodifica, data la semplicità dell'algoritmo, risultano molto contenuti e, come vedremo, esso è anche molto vantaggioso in termini di rate e di distanza.

Il codice di Hamming prende il nome dal suo inventore R. W. Hamming, un ricercatore dei Bell Laboratories che nel 1950 pubblicò un famoso articolo dal titolo *"Error Detecting and Error Correcting Codes"* nel quale per la prima volta presentò tale algoritmo.

L'obiettivo del codice di Hamming, come detto, è non solo la rivelazione di errori all'interno di una stringa di bit trasmessi su un canale ma anche la loro correzione. L'autore, per raggiungere questo scopo partì dall'analisi dei codici esistenti quali quelli di parità e ridondanza e tentò di massimizzarne le prestazioni. In particolare egli tentò di rendere più efficiente il codice a ripetizione svolgendo le stesse funzioni (cioè la correzione di un bit d'errore) ma inserendo un numero minore di bit di ridondanza. Hamming partendo dall'idea di voler individuare il bit affetto da errore per mezzo dei bit di ridondanza impose che il minimo numero di bit di controllo doveva essere tale da poter individuare con essi la posizione di tutti i bit della trama (dati e controlli). Se a questo punto indichiamo con m il numero di bit di dato e con k il numero dei bit di ridondanza possiamo imporre che:

$$2^k \geq m + k + 1$$

In questo modo con k bit di ridondanza in notazione binaria sarà possibile individuare la posizione di tutti i bit della trama. Tale condizione descrive il

limite minimo di bit di ridondanza da utilizzare per il codice di Hamming. In questo modo è possibile scegliere il numero minimo di bit di ridondanza da immettere nella trama una volta fissato il numero di bit di informazione che si vuole inviare oppure, che è la stessa cosa, il numero massimo di bit di dato che si può inviare una volta fissato il numero totale di bit da inviare. In tabella sono presentati alcuni valori di m e di k da utilizzare per la realizzazione del codice di Hamming.

$n=m+k$	m	k
3	1	2
4	1	3
5	2	3
6	3	3
7	4	3
8	4	4
9	5	4
10	6	4
11	7	4
12	8	4
13	9	4
14	10	4
15	11	4
16	11	5

Dalla tabella si può osservare che il numero di bit totali da inviare può essere scelto accuratamente in modo da massimizzare l'informazione. E' possibile, infatti, notare come inserendo nella trama 4 bit di ridondanza è possibile controllare da 4 a 11 bit di dato. Per questo motivo per massimizzare la rate d'informazione è necessario scegliere accuratamente la larghezza di banda del canale.

Una volta definito il numero di bit di ridondanza necessari per la realizzazione del codice vediamo in dettaglio il suo funzionamento. Il codice prevede che i bit di ridondanza controllino tutti i bit della trama (dato e controllo) in modo da superare il limite evidenziato da codice a controllo di parità. In particolare il codice di Hamming prevede che ogni bit di ridondanza sia funzione lineare di un certo numero di bit nella trama. Se si individuano le posizioni dei bit da inviare in notazione binaria si ha per definizione che il

primo bit di ridondanza agisce su tutti quei bit che si trovano nella posizione che ha un 1 nel bit più a destra. Ricordando che in notazione binaria:

$$\begin{aligned} 1 &= 1 \\ 3 &= 11 \\ 5 &= 101 \\ 7 &= 111 \\ 9 &= 1001 \end{aligned}$$

è chiaro che il primo bit di controllo è legato a tutti i bit che nella trama occupano posizioni dispari. Allo stesso modo è possibile individuare tutti i bit controllati dal bit di ridondanza numero 2 e così via. Resta da definire la posizione dei bit di controllo. L'algoritmo prevede di posizionare tali bit nei posti individuati dalle potenze di 2. In questo modo i bit di controllo si troveranno nella posizione 2,4,8,... Tale scelta non è casuale ma dipende dal fatto che solo così i bit di controllo sono indipendenti l'uno dall'altro e cioè che nel calcolo della combinazione lineare di bit che porta alla definizione del bit di controllo non sono presenti altri bit di controllo. La tabella seguente riassume tutte le informazioni descritte

numero del bit di controllo	posizione del bit di controllo	posizioni controllate
1	1	1,3,5,7,9,11,13,15,17,...
2	2	2,3,6,7,10,11,14,15,18,...
3	4	4,5,6,7,12,13,14,15,20,...
4	8	8,9,10,11,12,13,14,15,24,...
...

Nell'algoritmo di Hamming il calcolo del valore del bit di controllo è del tutto identico a quello utilizzato nel codice di parità. Il codice, infatti, prevede che i bit di controllo assumano valore 1 se il numero di bit controllati che assumono valore 1 sono dispari esattamente come accade per il codice a controllo di parità. Per questo motivo anche in questo caso i bit di controllo vengono denominati bit di parità. Ovviamente una volta che l'informazione

è arrivata a destinazione l'algoritmo di decodifica non deve far altro che effettuare il controllo di parità. Se, al termine delle operazioni di decodifica, si trova un'incongruenza, si segnala con un 1 il bit di parità che fallisce il controllo e con uno 0 quello che lo passa. Al termine di queste operazioni si ha che l'insieme dei bit così formati individuano, in notazione binaria, la posizione del bit affetto da errore.

La descrizione dell'algoritmo di Hamming può essere resa più chiara per mezzo di un esempio. Supponendo di voler inviare una parola di dato composta da 7 bit dalla tabella è possibile ricavare che il numero minimo di bit di controllo da impiegare è pari a 4. In tal caso si può affermare che si sta utilizzando un codice di Hamming (11,7) dove, appunto, con il primo numero si individua il numero di bit totale da inviare e con il secondo i bit d'informazione contenuti nella trama. Se si suppone ora che la parola di dato è:

0110101

è possibile calcolare i bit di parità utilizzando questa semplice notazione grafica:

	p ₁	p ₂	d ₁	p ₃	d ₂	d ₃	d ₄	p ₄	d ₅	d ₆	d ₇
dato (senza parità)			0		1	1	0		1	0	1
p ₁	1		0		1		0		1		1
p ₂		0	0			1	0			0	1
p ₃				0	1	1	0				
p ₄								0	1	0	1
parola da inviare	1	0	0	0	1	1	0	0	1	0	1

A questo punto la parola da inviare risulterà:

10001100101

Se ora si suppone che durante la trasmissione l'ultimo bit cambia valore a causa del rumore, al ricevitore giungerà la seguente parola

100011001**0**

A questo punto è possibile effettuare la decodifica descritta per il codice di Hamming utilizzando sempre lo stesso metodo:

	p ₁	p ₂	d ₁	p ₃	d ₂	d ₃	d ₄	p ₄	d ₅	d ₆	d ₇	c. di parità	bit di parità
p. ric.	1	0	0	0	1	1	0	0	1	0	0		
p ₁	1		0		1		0		1		0	errato	1
p ₂		0	0			1	0			0	0	errato	1
p ₃				0	1	1	0					corretto	0
p ₄								0	1	0	0	errato	1

Dato che almeno un bit di parità è stato posto ad uno dopo le operazioni decodifica, si può affermare che è avvenuto un errore in trasmissione. Resta ora da valutare la posizione del bit affetto da errore. Se consideriamo la stringa di bit di parità calcolati si ha:

1101

che in notazione binaria rappresenta il numero 11 che è proprio la posizione del bit affetto da errore. Non resta altro da fare che cambiare il valore di tale bit per ottenere la stringa di dati inviata dalla sorgente. Chiaramente il codice funziona anche se il bit affetto da errore è un bit di parità. In tal caso si otterrà che durante il controllo di decodifica sarà solo quello a fallire. Comunque, anche in quel caso la procedura porterà all'individuazione della posizione del bit errato. Per costruzione, in caso di trasmissione corretta il codice restituisce una sequenza di 0 per i bit di parità al termine delle operazioni di decodifica.

4.5.1 Codice di Hamming per la correzione di un singolo errore e rivelazione di un doppio errore

Il codice di Hamming fin qui descritto può essere reso ancora più efficace aggiungendo un ulteriore bit che esegue il controllo di parità su tutti i bit di controllo. In questo modo si può ottenere un codice in grado di individuare e correggere una parola di dati affetta da errore su un singolo bit ma anche segnalare se essa è affetta da errori su due bit. In questo caso si parla di codice a correzione di un singolo errore e rivelazione di due errori oppure SECDED (*Single Error Correcting Double Error Detecting*).

Il funzionamento di questo tipo di codice è molto semplice. In particolare, le operazioni di codifica e decodifica sono quelle descritte in precedenza e al termine ci si potrà trovare in uno dei seguenti casi:

1. **Nessun errore.** Tutti i bit di parità, compreso l'ultimo, risultano corretti dopo la decodifica.
2. **Un errore.** Il controllo sull'ultimo bit di parità fallisce in tutte le possibili situazione di errore, sia se l'errore coinvolge l'ultimo bit di parità sia se coinvolge uno dei bit di parità precedenti sia se coinvolge i bit di dato. A questo punto con la procedura standard è possibile individuare la posizione dove è avvenuto l'errore. In questo caso se la posizione individuata è la posizione 0 si intende che il bit di parità aggiuntivo è affetto da errore.
3. **Due errori.** In tutte queste situazioni il controllo effettuato sull'ultimo bit di parità è soddisfatto ma almeno uno degli altri bit di parità fallisce il controllo.

4.6 Il codice a correzione d'errore per l'esperimento NEMO

Analizzando il funzionamento dell'algoritmo di Hamming è possibile apprezzarne l'eleganza stilistica e la potenza nella rivelazione e la correzione di errori. Queste caratteristiche sono associate ad una estrema semplicità nelle operazioni di codifica e decodifica che rendono l'algoritmo estremamente efficiente nei tempi di calcolo e facile da realizzare in hardware. Dalla descrizione del funzionamento, inoltre, è possibile osservare che esso è particolarmente adatto in tutto quei casi in cui la probabilità che si verifichi un singolo errore in trasmissione risulta molto più elevata del verificarsi di due o più errori nella trama. Il codice, in questo caso, è infatti, in grado di correggere il flusso di dati nella maggior parte dei casi e solo quando si rivelano due bit errati richiedere nuovamente il trasferimento della trama. L'adozione di questo algoritmo su canali di per se non particolarmente rumorosi è quindi in grado

di aumentarne l'affidabilità. Tale vantaggio si ottiene praticamente a costo zero.

Queste caratteristiche hanno indotto a selezionare tale codice per la trasmissione dati in NEMO. Per come è concepita la trasmissione delle informazioni del telescopio e per la robustezza dei canali di comunicazione dovuta soprattutto ai componenti ed ai protocolli utilizzati, ci si aspetta che la probabilità di errori multipli (o *burst* di errori) sia molto limitata. L'estrema velocità delle operazioni richieste dall'algoritmo, inoltre, risulta particolarmente indicata per la trasmissione dei dati del telescopio NEMO. Esso infatti, prevedendo un campionamento continuo delle informazioni provenienti dai fotomoltiplicatori necessita di una trasmissione dati a tempo morto nullo. Se quindi si fosse utilizzato un algoritmo che prevedesse un tempo di calcolo elevato si sarebbe potuta raggiungere una più elevata affidabilità del canale ma paradossalmente si sarebbero perse comunque informazioni. Il codice di Hamming, invece, rende possibile effettuare le operazioni di codifica e decodifica in maniera praticamente istantanea. Come vedremo nel prossimo capitolo, utilizzando una serie di tecniche specifiche di progettazione elettronica, sarà possibile eseguire le operazioni di rivelazione e correzione di errori non introducendo alcun tempo di latenza nella trasmissione dati. In dettaglio, per gli scopi del telescopio NEMO si è deciso di sviluppare in hardware un codice Hamming SECDED per la correzione di un bit di errore e per la rivelazione di due bit di errore di tipo (22,16). In questo modo si ha la possibilità di trasferire parole di dato con lunghezza pari a 16 bit. La scelta è stata effettuata non solo considerando le specifiche del telescopio ma anche per rendere il dispositivo da realizzare il più versatile possibile. Utilizzando parole con una lunghezza pari a 16 bit si è tentato di rendere il più possibile standard il dispositivo e quindi capace di adattarlo a impieghi futuri senza la necessità di alcun cambiamento.

In questa configurazione il codice ha una rate pari a

$$R_{SECDED} = \frac{16}{22} = 0.72$$

che comunque rappresenta un ottimo valore anche confrontandolo con i codici descritti in precedenza.

Capitolo 5

Realizzazione del chip ASIC per il controllo della trasmissione dati in NEMO

5.1 Introduzione

Nel capitolo precedente è stata presentata una breve introduzione alla teoria dell'informazione ed in particolare di quella parte che riguarda la teoria dei codici a correzione di errore nella trasmissione dati su canali rumorosi. Al termine del precedente capitolo è stato presentato il codice di Hamming ed è stato dimostrato come esso soddisfi le specifiche dell'esperimento NEMO. E' stato anche messo in luce il suo carattere di generalità che quindi rende la realizzazione circuitale di tale algoritmo utile, non solo agli scopi di un esperimento per lo studio dei neutrini, ma anche ad impieghi in svariati campi di applicazione. In questo capitolo saranno presentati i dettagli della progettazione hardware del circuito elettronico digitale, da me realizzato, che effettua funzioni di codifica e decodifica dell'algoritmo di Hamming. Saranno presentate le specifiche di progetto che sono state definite affinché il dispositivo potesse essere utilizzato nel telescopio per neutrini senza comunque perdere le sue caratteristiche di generalità. Verrà anche presentata la tecnologia ASIC (*Application Specific Integrated Circuit*) con cui è stato realizzato il dispositivo e saranno esposte le motivazioni del suo impiego. Seguirà la presentazione del dettaglio dello schema a blocchi del chip ed infine verranno

descritte le sue prestazioni.

5.2 Le specifiche di progetto

Una volta stabilito che l'algoritmo di Hamming è quello che meglio si adatta all'impiego per la trasmissione dati del telescopio NEMO è stato possibile definire in dettaglio le specifiche di progetto che dovevano essere soddisfatte affinché tale algoritmo potesse essere impiegato nell'esperimento. Data le caratteristiche del telescopio ed il particolare ambiente in cui ci si trova ad operare è stato subito chiaro che il codificatore e decodificatore dell'algoritmo di Hamming dovesse essere realizzato mediante un dispositivo elettronico digitale realizzato appositamente. Questa scelta si è resa necessaria poichè, visti gli spazi ristretti e la necessità di ottenere bassi consumi ed elevata affidabilità, era impensabile integrare nel telescopio un processore di tipo *general purpose* su cui venisse realizzato via software l'algoritmo selezionato.

Scelta la tecnologia con cui produrre il dispositivo è stato necessario definire le sue specifiche tecniche. Come detto in precedenza, in questa fase si è tenuto conto della necessità di rendere il dispositivo non solo perfettamente compatibile con l'elettronica del telescopio NEMO ma anche di realizzarlo in modo che fosse il più possibile di uso generale e che quindi potesse essere usato in ambiti diversi senza la necessità di effettuare modifiche al progetto.

Il primo parametro che è stato definito è la larghezza del canale di trasmissione che con tale algoritmo si doveva controllare. E' stato deciso di progettare il codice in modo che fosse possibile inviare sul canale di trasmissione parole di 16 bit in parallelo. Questa scelta è stata effettuata soprattutto perchè tale larghezza risulta essere uno standard nel campo dell'elettronica digitale. Come è noto, infatti, in moltissime applicazioni si usano parole a 8 bit (Byte) o multipli interi di essa. Una volta definito il canale di dati la teoria alla base dell'algoritmo di Hamming stabilisce che il numero minimo di bit di ridondanza necessari alle operazioni di codifica e decodifica è pari a 5. In questo caso è possibile però realizzare solo il codice per la correzione di errori su un bit. Si è deciso invece di realizzare il codice SECDED e quindi richiedere al progetto la possibilità di riconoscere il caso in cui si verificas-

sero due errori in trasmissione. Questa richiesta rappresenta una funzione aggiuntiva che probabilmente non verrà utilizzata per gli scopi del telescopio. Come spiegato precedentemente il codice di Hamming SECDED, quando si verificano due errori simultanei sul canale di trasmissione, è solo in grado di segnalare la presenza ma non di correggerli. La procedura standard per il recupero di questi dati sarebbe quella di effettuare nuovamente la trasmissione del dato corrotto. E' noto però che le caratteristiche del telescopio NEMO non prevedono questo tipo di possibilità. Tuttavia si è deciso comunque di realizzare questa funzione in modo tale che il dispositivo progettato potesse essere di uso più generale possibile. Si è giunti quindi alla definizione dell'algoritmo da realizzare. Esso è il codice Hamming SECDED (22,16) nel quale possono essere inviati 16 bit di dato controllati da 6 bit di parità.

Un parametro importante da definire quando si tratta di dispositivi elettronici digitali è la frequenza di funzionamento. La quasi totalità dei circuiti digitali, infatti, prevede un funzionamento sincrono. Tali dispositivi svolgono tutte le operazioni seguendo un segnale periodico di cadenza temporale, denominato clock, che a questo punto ne definisce la frequenza di funzionamento. Per stabilire la frequenza di clock per il progetto si è partiti dall'analisi della massima frequenza di trasferimento dati presente sul telescopio. Tale frequenza è risultata essere di 200 MHz e viene utilizzata per la trasmissione dati tra il Front End del fotomoltiplicatore e la schede di acquisizione e smistamento dati FCM. E' stato quindi deciso che il chip da realizzare dovesse funzionare fino alla frequenza di 200 MHz per potersi adattare a quella situazione.

Per tentare di ottimizzare il progetto è stato anche richiesto che un singolo dispositivo potesse funzionare sia da codificatore che da decodificatore. E' stato quindi necessario realizzare un chip che avesse la possibilità di svolgere sia le funzioni di codifica che di decodifica cambiando modalità a seconda dei parametri funzionali con cui venisse configurato. Questa specifica tecnica è motivata da una serie di ragioni. Innanzitutto, realizzare un unico dispositivo risulta vantaggioso quando si usano tecnologie di tipo ASIC poichè viene prodotto un unico tipo chip e quindi si abbattano i costi di produzione. La scelta è anche conveniente in termini di affidabilità. Durante la fase di

assemblaggio dei componenti elettronici, infatti, non si rischia di incorrere in errori montando il codificatore al posto del decodificatore e quindi è meno probabile il ricorso a rilavorazioni delle schede elettroniche che possono spesso richiedere tempi lunghi e aumento dei costi. Inoltre la scelta di realizzare un unico chip permette di condividere una parte delle risorse logiche sia quando il dispositivo funziona in ricezione che in trasmissione. In questo modo è possibile piazzare sul silicio meno componenti e quindi ottimizzarne l'area e i consumi.

Tutte le specifiche descritte dovevano ovviamente essere ottenute in modo da garantire una elevata affidabilità e consumi estremamente contenuti.

Riassumendo, le caratteristiche richieste al progetto sono le seguenti:

- realizzazione del codice Hamming SECDED (22,16);
- massima frequenza di funzionamento pari a 200 MHz;
- realizzazione delle funzioni di codifica e decodifica su un unico chip;
- elevata affidabilità;
- consumi estremamente ridotti.

Una volta analizzate le specifiche di progetto ci si è trovati a scegliere la tecnologia da adottare per realizzare il dispositivo. Nel campo dei dispositivi elettronici digitali esistono due grandi famiglie con le quali si possono produrre circuiti di questo tipo: gli ASIC e le FPGA (*Field Programmable Gate Array*). Come vedremo nel prossimo paragrafo con esse è possibile realizzare circuiti simili. Tuttavia gli ASIC sono in grado di raggiungere prestazioni molto più elevate e possono essere ottimizzati in modo da ridurre la dissipazione di potenza. Si è quindi deciso di impiegare tale tecnologia. Nel paragrafo successivo saranno messe in luce in dettaglio le motivazioni di questa scelta.

5.3 Confronto tra ASIC e FPGA

Le FPGA (*Field Programmable Gate Array*) sono dei chip nei quali è presente una matrice di porte logiche combinatorie e una serie di flip-flop. L'utente,

una volta acquistato il dispositivo, può, per mezzo di software di sviluppo, attivare le risorse logiche presenti in modo da ottenere il progetto voluto. E' chiaro che questo tipo di approccio presenta una serie di vantaggi. Tali dispositivi, infatti, sono estremamente semplici da programmare e, soprattutto, le operazioni di configurazione possono essere effettuate dall'utente finale del dispositivo. Il progettista, infatti, viene fornito degli strumenti necessari a compiere le operazioni necessarie per far funzionare il chip secondo le proprie specifiche. Tale approccio permette quindi di realizzare in poco tempo e con costi contenuti un qualunque tipo di dispositivo digitale. I chip sono poi riutilizzabili poichè possono essere riprogrammati in modo da eseguire diverse funzioni. Ovviamente l'utilizzo delle FPGA possiede anche una serie di svantaggi. Essi sono dovuti sostanzialmente alla non ottimizzazione delle risorse. Come detto, sul chip è presente una matrice di elementi logici pre-piazzati su silicio. Per questo motivo, indipendentemente dal fatto che il progetto realizzato dall'utente impieghi o meno tutte le risorse, esse contribuiscono alla dissipazione di potenza totale. E' chiaro che in questo modo anche l'area del chip non risulta ottimizzata. Un dispositivo come l'FPGA comporta inoltre una non ottimizzazione del particolare progetto che l'utente intende realizzare. Per questo motivo di solito accade che le prestazioni ottenute con tali dispositivi non risultano le massime raggiungibili. Soprattutto in termini di frequenza di funzionamento capita spesso che, progetti realizzati in questo modo non risultano i più veloci ottenibili.

La tecnologia ASIC prevede invece che l'utente progetti totalmente in proprio il suo dispositivo e quindi possa renderlo il più possibile ottimizzato per le sue esigenze. E' allora chiaro che questo tipo di tecnologia è estremamente più flessibile e adattabile alle particolari necessità dell'utente. Anche se estremamente più complessi da progettare e costosi da realizzare, l'impiego di tali dispositivi si rende necessario qualora si vogliano raggiungere prestazioni elevate sia in termini di massima frequenza di funzionamento sia in termini di minima dissipazione di potenza e di ottimizzazione delle risorse. I vantaggi appena illustrati hanno portato a scegliere la tecnologia ASIC per la realizzazione del chip per il controllo della trasmissione dati di NEMO. Le specifiche di progetto esposte in precedenza, infatti, richie-

do un'ottimizzazione di progetto che difficilmente può essere raggiunta con dispositivi FPGA.

5.4 La tecnologia ASIC digitale

La tecnologia ASIC (*Application Specific Integrated Circuit*) come lo stesso acronimo suggerisce, permette di realizzare circuiti integrati specifici per l'applicazione a cui sono destinati. Ciò significa che il progettista, una volta definite le specifiche, può realizzare il circuito ottimizzandolo al massimo per le proprie esigenze. L'elevato livello di personalizzazione è dovuto al totale controllo che è possibile ottenere in fase di progettazione. La fase di realizzazione del chip parte infatti da una "mattonella" di silicio vuota sulla quale l'utente può, e deve, disegnare e piazzare tutti i componenti necessari alla realizzazione del circuito che ha in mente. Come sarà mostrato più avanti, ciò significa che in fase di disegno è possibile realizzare tutti i componenti logici di cui il dispositivo necessita a partire dai suoi costituenti analogici. E' inoltre necessario tracciare anche tutte le linee di interconnessione per permettere lo scambio dei segnali tra i diversi costituenti del circuito nonché tutte le linee di alimentazione che permettono di portare corrente agli elementi. E' chiaro che un così alto grado di controllo da un lato aumenta enormemente la flessibilità del progetto permettendo di realizzare praticamente qualunque tipo di circuito ma dall'altra fa crescere enormemente la complessità della fase progettuale. A differenza di altri tipi di piattaforme quali le FPGA il progetto di un ASIC richiede molto più tempo e più attenzione poichè è necessario tenere sotto controllo anche i più piccoli dettagli che possono compromettere il funzionamento del chip. Per questi motivi i passi da realizzare durante il disegno di un ASIC sono molti di più e molto più complessi di quelli necessari per il disegno di un circuito su FPGA. In particolare la fase di simulazione funzionale di un ASIC risulta estremamente approfondita poichè bisogna verificare tutti i parametri funzionali, dall'alimentazione elettrica alla distribuzione del segnale di clock al corretto piazzamento dei blocchi logici fino ai pad di ingresso e di uscita.

Il progettista esegue tutta la fase di disegno per mezzo di software CAD

(*Computer Aided Design*) che forniscono un supporto fondamentale per la buona riuscita del lavoro. Viste le grandezze submicrometriche e l'enormità di parametri da tener presente sarebbe, infatti, impensabile far a meno di strumenti di calcolo estremamente sofisticati. E' necessario pensare che in un chip digitale moderno di media entità possono essere contenute molte decine di migliaia di transistor che ovviamente devono essere connessi ad uno ad uno correttamente e in modo che occupino la minima area possibile. Tale obiettivo deve essere raggiunto, ovviamente, in modo da ottimizzare al massimo le prestazioni del prodotto finale.

Una volta terminata la fase progettuale il disegno del circuito viene inviato ad apposite fabbriche dette fonderie (dall'inglese *foundary*) che provvedono alla realizzazione vera e propria del dispositivo. Tali ditte provvedono, tramite tecniche litografiche, a trasformare le indicazioni contenute nel progetto in veri e propri dispositivi elettronici di silicio. Bisogna anche considerare che le fonderie non conoscono alcun dettaglio del progetto e delle intenzioni del progettista e quindi si limitano a eseguire pedissequamente le indicazioni contenute nel disegno. E' quindi fondamentale che questo sia eseguito con estrema cura e precisione in modo da rendere minime le possibilità di errore in fase realizzativa. A tale livello di dettaglio si giunge, come vedremo in seguito, compiendo una lunghissima fase di simulazione. In fase di progetto è, infatti, possibile verificare il funzionamento virtuale del dispositivo tentando di scovarne tutti i possibili errori ed incongruenze.

5.4.1 La scelta del processo tecnologico

Quando si parla di progettazione ASIC è necessario specificare la tecnologia o il processo che si intende usare. Con questa parola si identifica il processo tecnologico che la fonderia dovrà adottare per la realizzazione del chip. In questo modo sarà possibile scegliere la dimensione minima delle tracce da poter realizzare su silicio e quindi la dimensione minima ottenibile per i componenti. I diversi processi sono identificati da grandezze spaziali, di solito espresse in micron.

Quando si definisce il processo è anche possibile definire quanti livelli di

metal si hanno a disposizione ovvero quanti strati di materiale conduttore sarà possibile sovrapporre per realizzare il dispositivo. Chiaramente un alto numero di livelli di metal può far aumentare la densità di transistor posizionabili. Tuttavia ciò comporta un aumento della complessità del progetto poichè porta ad una crescita del numero dei parametri funzionali da dover controllare. Per la realizzazione del chip per il controllo della trasmissione dati di NEMO è stata utilizzata la tecnologia $0.18 \mu m$ fornita dalla fonderia UMC che prevede 6 livelli di metal.

La progettazione di un ASIC può essere resa più agevole se si utilizzano le cosiddette celle standard. Si tratta di una serie di porte logiche e dispositivi sequenziali disegnate e messe a disposizione della fonderia per una determinata tecnologia. Tali celle possono essere usate dall'utente che in questo caso non dovrà provvedere a realizzare in proprio. In questo modo non solo si risparmia moltissimo tempo durante la fase di progettazione ma è più probabile la riuscita del chip. Esse infatti sono disegnate dalla stessa ditta che le realizzerà e quindi risultano ottimizzate per i processi che le andranno a trasferire su silicio. Tale metodo di progettazione prende il nome di flusso a celle standard (*standar cell design*) e si contrappone al flusso *full custom* (*full custom design*) che invece prevede la realizzazione da parte dell'utente di tutti i dispositivi che andrà ad impiegare.

5.4.2 Flusso di progetto a celle standard per un ASIC digitale

In questa sezione vengono riportati i passi da me compiuti nella progettazione del chip per il controllo della trasmissione dati per il telescopio NEMO in modo da rendere più chiari i risultati presentati al termine del lavoro.

La fase di progettazione di un chip ASIC digitale prevede una serie di passi descritti sinteticamente nel diagramma di flusso presentato in figura 5.1.

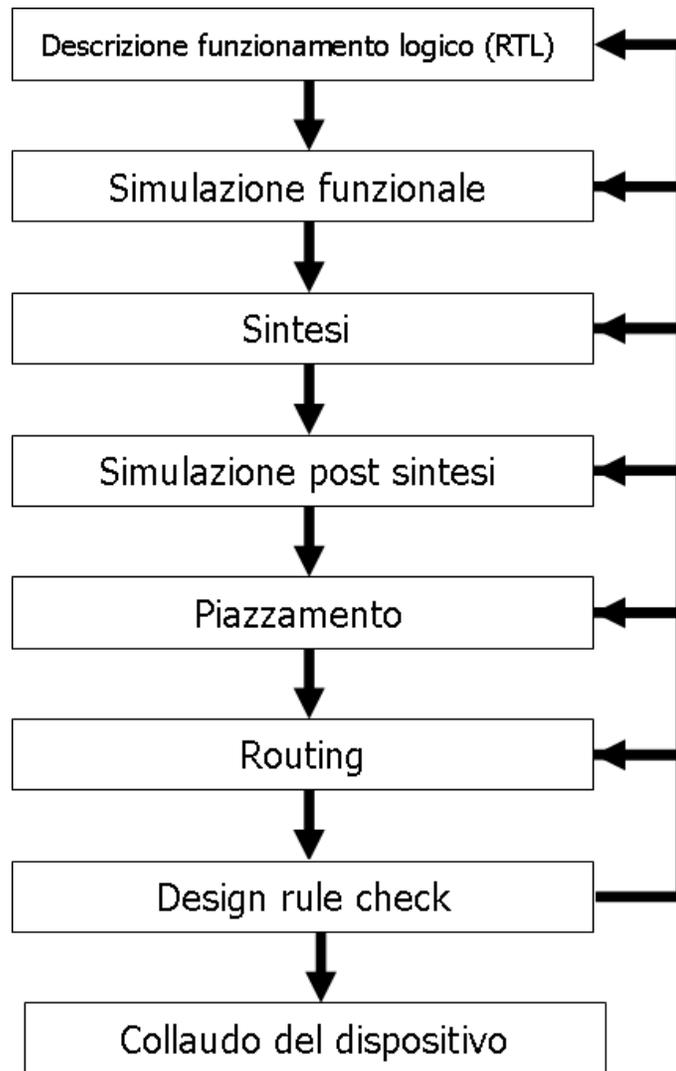


Figura 5.1: Flusso di progetto utilizzato per la realizzazione del chip per il controllo della trasmissione dati in NEMO

La descrizione RTL

Una volta definite le specifiche che il chip dovrà soddisfare e le prestazioni desiderate si procede alla descrizione del funzionamento del dispositivo mediante una serie di linguaggi specifici per la progettazione di elettronica digitale. Il linguaggio da me utilizzato è il HDL (*Hardware Description Language*) che permette di descrivere dettagliatamente il funzionamento del

dispositivo in maniera formalmente identica alla procedura per la realizzazione di un software. Questa fase prende il nome di progettazione RTL (*Register Transfer Level*).

La simulazione funzionale

Il passo successivo consiste nel provare che la descrizione RTL sia effettivamente aderente alle specifiche richieste. Con l'ausilio di una serie di software di simulazione, a partire dai file VHDL, è possibile generare una versione virtuale del dispositivo. A questo punto, con l'ausilio degli stessi software, è possibile fornire una serie di segnali simulati in ingresso e analizzare i segnali prodotti in uscita. In questo modo può essere compiuto uno studio approfondito del comportamento logico che il chip assumerà una volta realizzato. Il procedimento descritto risulta fondamentale e delicato poichè, al termine di questa fase, ci si disinteresserà della struttura logica del chip e si darà per scontata la sua correttezza. Tutti gli eventuali errori logici dovranno quindi essere corretti in questa fase.

La sintesi

Successivamente si passa alla fase cosiddetta di sintesi. Per mezzo di software dedicati, i file VHDL che descrivono il comportamento RTL del progetto vengono tradotti per formare la descrizione strutturale del circuito. Tali software denominati sintetizzatori sono in grado di associare alla descrizione comportamentale del dispositivo i blocchi logici combinatoriali e sequenziali necessari a realizzare le funzioni descritte. I blocchi necessari vengono atinti dalla libreria di celle standard fornita dalla fonderia. Al termine della procedura di sintesi viene prodotto un file che, ovviamente, ancora descrive il funzionamento del dispositivo in oggetto, ma questa volta a livello di porte logiche e flip-flop. In questo file sono anche definiti tutti i collegamenti elettrici tra i vari blocchi necessari alla propagazione dei segnali logici. Si ha quella che viene chiamata *gate level netlist*.

La simulazione post sintesi

E' possibile ora effettuare una nuova fase di simulazione per avere la sicurezza che il processo di traduzione sia andato a buon fine e che il circuito contenuto nella *gate level netlist* sia ancora perfettamente aderente alle specifiche inizialmente stabilite.

Il piazzamento

Inizia, a questo punto la fase di realizzazione vera e propria del chip. Si passa infatti alla definizione del layout del dispositivo e cioè di come saranno geograficamente piazzati tutti i suoi componenti. La prima cosa che viene “costruita” è la struttura del chip. Vengono, infatti, posizionati i pad di ingresso e di uscita e viene effettuata una stima dell'area di silicio che il dispositivo occuperà una volta realizzato. Ci si rende conto ora se il chip è *core limited* o *pad limited*. Si scopre cioè se le dimensioni finali saranno individuate dall'area occupata dai blocchi logici oppure dal perimetro definito dai pad di ingresso e di uscita. Al termine di questa operazione viene disegnata la rete di alimentazione. Vengono infatti tracciate tutte le linee necessarie a fornire corrente ai blocchi funzionali del chip. Anche questa fase risulta molto delicata poichè è necessario dimensionare perfettamente la rete di alimentazione. Connessioni troppo piccole potrebbero, infatti, comportare problemi di scarso flusso di corrente e di abbassamento di tensione che porterebbe allo spegnimento delle celle. Il dimensionamento deve essere effettuato anche considerando le massime prestazioni che il dispositivo dovrà raggiungere per poter essere sicuri che, anche alla massima velocità, sarà possibile erogare corrente sufficiente per il corretto funzionamento. Si passa poi al piazzamento delle celle logiche che costituiscono il core del dispositivo. Nel caso in cui il chip sia *pad limited* spesso accade che l'area utile per le celle sia molto più grande di quella effettivamente necessaria. In questo caso bisogna evitare che le celle vengano sparse in tutto lo spazio poichè si avrebbero grossi problemi di prestazioni. Si utilizza quindi un metodo alternativo che consiste nel definire un'area ristretta di silicio nella quale vengono piazzate in maniera compatta tutte le celle.

Il routing

L'operazione successiva è quella di creazione delle interconnessioni logiche tra le varie celle e tra i pad di ingresso e uscita: la fase di *routing*. Software dedicati, a partire dai file dove sono descritte le interconnessioni e per mezzo di algoritmi di ottimizzazione, permettono di collegare tra loro i vari blocchi in modo da ottenere non solo il corretto funzionamento ma anche il raggiungimento delle massime prestazioni possibili. Questa fase, come la precedente, è guidata dal progettista che fornisce ai software una serie di file denominati *constraints*. In essi vengono descritti dettagliatamente tutte le caratteristiche spaziali e le prestazioni che il chip dovrà raggiungere.

La verifica finale

L'ultima fase è quella di verifica finale. Una volta eseguiti tutti i passi necessari alla progettazione di un ASIC digitale è infatti possibile ottenere una stima molto affidabile delle prestazioni del dispositivo una volta prodotto. Per prima cosa è necessario estrarre tutti i componenti parassiti. Questa procedura è mirata a determinare il valore delle capacità delle resistenze e delle induttanze che si sono venute a creare effettuando il piazzamento e l'interconnessione dei diversi blocchi. Tali componenti contribuiscono in maniera sostanziale a definire i tempi reali di attraversamento dei segnali logici all'interno del dispositivo. Per questo motivo le quantità calcolate in questa fase devono essere tenute in considerazione all'atto della determinazione delle massime prestazioni. Una volta effettuata tale stima è infatti possibile fornire il risultato ai software di simulazione che sono in grado di stimare le prestazioni del dispositivo in maniera precisa. Con l'ausilio di simulazioni e delle informazioni riguardanti i componenti parassiti è anche possibile effettuare una stima della dissipazione di corrente del chip.

Il design rule check

L'ultima fase di verifica consiste nel controllare che il dispositivo sia aderente ad una serie di regole che la fonderia impone per garantirne la corretta produzione. Tale fase prende il nome di *design rule check*. Il controllo, effettuato

dal progettista e ripetuto dalla fonderia, serve a garantire che il progetto sia compatibile con i processi produttivi che porteranno alla creazione del chip.

Tutte le fasi di simulazione e di verifica forniscono importanti informazioni al progettista che in questo modo riesce a comprendere se i passi effettuati sono corretti. Chiaramente, tali procedure risultano molto delicate e, come è possibile notare, richiedono gran parte del tempo di progettazione. Bisogna anche tener presente che al termine di ogni verifica è necessario valutare con attenzione gli esiti e decidere se proseguire con il flusso oppure ripetere il passo precedente per tentare di migliorare i risultati ottenuti.

La fase progettuale termina con la generazione di file nei quali è descritto il dispositivo che vengono inviati alla fonderia per la produzione del chip.

Il collaudo del dispositivo

Una volta che il dispositivo è realizzato si procede alla campagna di misure di laboratorio che consentono di verificarne le prestazioni reali. Questa fase di test può risultare molto complessa e impegnativa. Con strumenti quali generatori di segnali e oscilloscopi è, infatti, necessario riprodurre tutte le condizioni di funzionamento e verificarne le risposte. Per cercare di rendere più agevole possibile le operazioni è possibile inserire, in fase di progetto, una serie di blocchi logici che eseguono delle funzioni speciali che permettono di provare velocemente il dispositivo e comprendere le reali prestazioni.

5.5 L'architettura del chip

Il circuito elettronico digitale, da me realizzato, per il controllo della trasmissione dati dell'esperimento NEMO, come detto, deve svolgere le funzioni di codificatore e decodificatore dell'algoritmo Hamming SECDED (22,16) secondo le specifiche descritte in precedenza. Per come è stato concepito esso deve essere posto all'inizio ed alla fine di un canale di comunicazione dati a 16 bit. Per utilizzarlo in effetti è necessario compiere una modifica al bus nel senso che è indispensabile ampliare la sua larghezza affinché sia possibile condividere anche le linee di controllo. Per usare il dispositivo il canale di

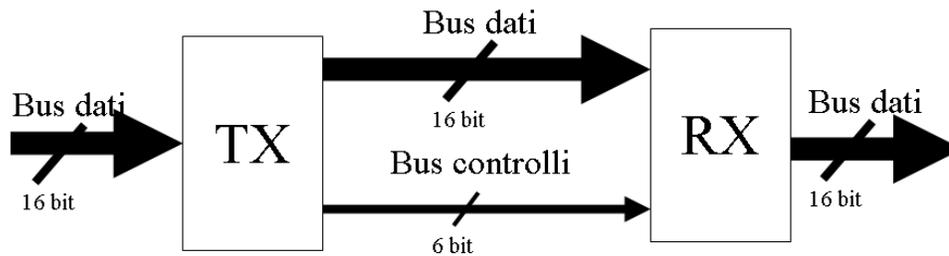


Figura 5.2: Configurazione del bus per l'utilizzo del chip per il controllo della trasmissione dati.

comunicazione deve contenere in tutto 22 linee. La configurazione necessaria è quella descritta in figura 5.2.

Dato che sul chip sono implementate sia le funzioni di codifica che di decodifica è necessario fornire al dispositivo innanzitutto un parametro di configurazione che ne definisce la modalità di funzionamento come trasmettitore oppure come ricevitore.

Per analizzare in dettaglio il comportamento del chip consideriamo le due modalità di funzionamento in maniera separata.

5.5.1 Funzionamento in trasmissione

Se il chip è configurato in modalità di trasmissione saranno abilitate le funzioni che compiono le operazioni di codifica Hamming SECDED. Il dispositivo deve quindi ricevere in ingresso i 16 bit di dato e fornire in uscita la stessa parola con, in aggiunta, i 6 bit di ridondanza come previsto dall'algoritmo. L'architettura da me realizzata prevede che i dati in ingresso ed in uscita dal chip vengano gestiti in parallelo. Per questo motivo, in modalità di trasmissione, sono abilitati 16 ingressi sui quali deve essere presentato il dato da codificare ed in uscita si pilota un bus a 22 bit che contiene il dato e i bit di parità. Essendo il funzionamento del dispositivo sincrono con un segnale di cadenza è necessario fornire anche il clock che scandisce le operazioni e che viene considerato come unico riferimento temporale di tutto il circuito. Per questo motivo i dati in ingresso devono essere presentati in fase con tale

segnale e, analogamente, i dati e i controlli in uscita saranno posti sul bus sempre in maniera sincrona con il clock.

5.5.2 Funzionamento in ricezione

Quando il dispositivo viene configurato come ricevitore esso deve svolgere le funzioni di decodifica dell'algoritmo di Hamming. In tale modalità, infatti, riceve in ingresso le 16 linee di dati e le 6 linee di controllo. Entrambe le parole devono essere fornite in parallelo e, essendo il funzionamento sincrono con il segnale di cadenza, devono risultare in fase con il segnale di clock di sistema. Una volta che il chip ha ricevuto le informazioni in ingresso effettua tutte le operazioni di decodifica e fornisce in uscita una serie di dati in base a quanto accaduto durante la trasmissione. Si possono, infatti verificare quattro casi:

1. **Nessun errore sul canale di trasmissione.** Se durante la propagazione del dato non è occorso alcun errore il chip verifica la congruità della parola di dato con quella di controllo e fornisce il dato sulle 16 linee del bus di uscita.
2. **Un errore singolo sul canale di dato.** Se durante la trasmissione un bit della parola di dato è stato affetto da errore che ne ha comportato il cambiamento di valore logico, al ricevitore arriverà una parola inattesa. Una volta effettuate le operazioni di decodifica esso sarà in grado di correggere tale bit e fornire in uscita il risultato. Il dispositivo segnala l'avvenuta correzione attivando una linea di allarme.
3. **Un errore singolo sul canale di controllo.** In questo caso il dato attraversa senza cambiamenti il canale ma sul bus di controlli avviene un errore che provoca il cambiamento di stato di un bit. Al termine delle operazioni di decodifica il chip riconosce dove si è verificata l'anomalia e quindi porta in uscita il valore del dato ricevuto in ingresso. Anche in questo caso il dispositivo attiva un allarme per segnalare l'avvenuto errore sul bus di parità.

4. **Errore doppio sui bus.** Quando si verifica il cambiamento di più bit sia sul bus dati che su quello di controlli il dispositivo è ancora in grado di riconoscere la situazione e di gestirla ma non di correggere gli errori occorsi. Per questo motivo il dato ricevuto in ingresso viene propagato in uscita anche se questa volta può non risultare corretto. Viene attivato in questo caso un terzo tipo di allarme che segnala la situazione anomala riscontrata.

5.5.3 Funzioni aggiuntive

Il chip, oltre alle funzionalità di base fin qui descritte, svolge una serie di operazioni accessorie che hanno la duplice funzione di facilitare la fase di collaudo e calcolare la rumorosità del bus oggetto del controllo. Tali funzioni si ottengono attivando la modalità di test del chip.

- **Test del trasmettitore.** Attivando la modalità di test quando il dispositivo è configurato come trasmettitore è possibile provare la funzionalità della sezione di codifica. In questo caso, infatti, il circuito ignora i dati che vengono posti sul bus d'ingresso e viene generata all'interno del chip una sequenza di numeri pseudo-random. Tale sequenza è a tutti gli effetti considerata come un dato da trasmettere e quindi su di essa vengono calcolati i bit di parità. Ad ogni colpo di clock in uscita è presente un valore di questa sequenza e i relativi bit di parità calcolati. Dato che la sequenza è nota e ripetitiva, è facile controllare che il trasmettitore stia effettuando la codifica delle informazioni in modo corretto.
- **Test del bus.** Se si attiva la modalità di test sia sul trasmettitore che sul ricevitore è possibile calcolare la rumorosità del canale di trasmissione. In questo caso il codificatore genera e pone in uscita la sequenza di valori estratti dal contatore pseudo-random con i relativi bit di controllo. Tali informazioni, dopo aver attraversato il canale arrivano al ricevitore. Questo è in grado di riconoscere la sequenza in arrivo e controllarla automaticamente. Quando parte il test, al termine della

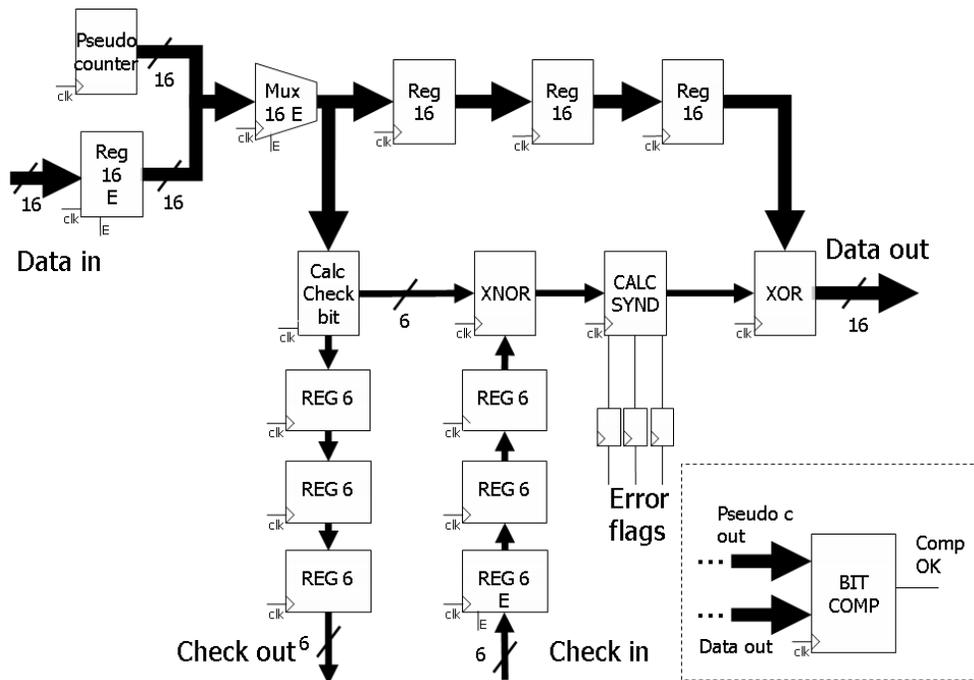


Figura 5.3: Schema a blocchi del chip SECDED per il controllo della trasmissione dati dell'esperimento NEMO.

decodifica, il chip confronta il dato ottenuto con quello atteso e, ad ogni ciclo di clock, è in grado di verificare se la sequenza in arrivo risulta corretta. Ogni volta che la verifica riesce il ricevitore attiva un segnale. Monitorando questo segnale e le linee di allarme è possibile contare il numero di errori sui bus di comunicazione.

5.6 Lo schema a blocchi del chip

Le funzionalità appena descritte sono schematizzate mediante il diagramma a blocchi riportato in figura 5.3. Per mezzo di questa rappresentazione grafica è possibile descrivere alcune dettagli della progettazione del dispositivo. La prima cosa da poter notare è che ogni blocco funzionale presente nel disegno ha in ingresso il segnale di cadenza. Ciò corrisponde al fatto che in fase di progetto è stato posizionato in uscita ad ogni blocco un livello di registri. In questo modo al termine di ogni operazione elementare il blocco

funzionale fornisce i prodotti dell'operazione svolta in fase con il segnale di cadenza. Si evitano quindi problemi di sincronizzazione dei dati che possono compromettere le prestazioni del dispositivo. Se così non fosse un blocco logico potrebbe ricevere due gruppi di dati necessari a compiere un'operazione sfasati tra loro. In questo caso, aumentando la frequenza di funzionamento si potrebbero ottenere problemi di funzionamento molto difficili da riscontrare. La tecnica utilizzata, detta *pipeline*, serve, inoltre ad ottenere un aumento di prestazioni. In questo modo è infatti possibile suddividere l'esecuzione di operazioni complesse in una serie di operazioni più semplici ognuna effettuata durante un periodo del segnale di cadenza. Tali operazioni possono essere svolte in parallelo risparmiando tempo di elaborazione. Il problema principale che si riscontra utilizzando questo tipo di tecnica è dovuto alla sincronizzazione dei dati. E' possibile infatti che le informazioni necessarie all'esecuzione di un'operazione non siano disponibili tutte nello stesso ciclo di clock. Ovviamente, una situazione di questo tipo comprometterebbe il funzionamento del circuito. Per questo motivo è necessario porre grande attenzione alla sincronizzazione del flusso di dati all'interno del chip e fare in modo che tutte le informazioni necessarie allo svolgimento di una specifica operazione siano disponibili simultaneamente. Per fare ciò è stato necessario aggiungere dei livelli di registri che permettano di sincronizzare tutti i gruppi di dati. Osservando il diagramma a blocchi presentato è possibile notare la presenza di tali registri denominati generalmente buffer. Un'architettura di questo tipo ha anche un'altra particolarità riscontrabile durante il funzionamento del dispositivo. All'accensione, il risultato dei dati posti in ingresso non è subito presentato in uscita. Esso, infatti è presentato con un certo tempo di latenza. Tale tempo, misurabile in un certo numero di cicli di clock è necessario affinché la pipeline si riempia. Solo a quel punto è possibile ottenere la sincronizzazione dei dati e l'esecuzione delle operazioni. E' importante ricordare che questo tempo di latenza non compromette in alcun modo l'utilizzo del dispositivo. E' necessario, infatti, attendere la scadenza del tempo di latenza solo all'accensione del dispositivo; una volta a regime il chip è in grado di fornire dati senza soluzione di continuità. La tecnica di pipeline utilizzata da tutti i processori ad elevate prestazioni ha

permesso di raggiungere le specifiche attese per il progetto da me realizzato. E' comunque stata posta molta cura nel diminuire al massimo il tempo di latenza iniziale portandolo a soli 6 cicli di clock. La sincronizzazione con il segnale di cadenza è stata anche imposta per tutti i segnali in uscita dal chip. In questo modo è possibile far funzionare due dispositivi (trasmettitore e ricevitore) con lo stesso segnale di cadenza evitando di dover sincronizzare i dati all'esterno dei chip.

Grande cura è stata posta nella progettazione dei blocchi in modo da rendere quanto più possibile contenute le dimensioni finali del core. Per questo motivo ho cercato di realizzare elementi funzionali che fosse possibile utilizzare sia in fase di codifica che decodifica. In questo modo è stata possibile una condivisione delle risorse che ha portato giovamenti sia in termini di prestazioni che di area occupata che di dissipazione di potenza.

Nel diagramma è possibile notare inoltre la presenza dei blocchi funzionali necessari a svolgere la modalità di test. In alto a destra è, infatti presente il generatore di numero pseudo-random. I dati in uscita da tale sezione possono essere immessi nel circuito di codifica e decodifica attivando i segnali di controlli di un multiplexer. Nel riquadro posto in basso a destra del disegno è presentato il blocco che esegue la verifica dei dati da effettuare durante la fase di test del bus dati. Esso, attivato durante la fase di ricezione dati, riceve i segnali in arrivo da trasmettitore e li confronta con quelli in uscita dal generatore pseudo-random verificandone la congruità.

Capitolo 6

Risultati sperimentali

6.1 Introduzione

L'ultima parte del presente lavoro di tesi è dedicata alla presentazione dei risultati sperimentali relativi al chip, da me realizzato, che implementa l'algoritmo di Hamming (22,16) SECDED per il controllo della trasmissione dati per l'esperimento NEMO. Tali risultati sono suddivisi in due parti principali. Verranno dapprima presentate le simulazioni più significative ottenute durante la fase di progettazione del chip e successivamente le immagini salienti ricavate durante il collaudo del dispositivo, effettuato in laboratorio, una volta realizzato dalla fonderia. In questo modo sarà possibile avere un confronto immediato delle caratteristiche del dispositivo e della sua perfetta aderenza alle specifiche di funzionamento imposte in fase di progetto.

6.2 Simulazioni del progetto

In questa sezione vengono rappresentate alcune delle simulazioni più significative ottenute durante la fase di progetto. Come descritto nel precedente capitolo la fase di simulazione riveste un ruolo fondamentale per la realizzazione di un chip ASIC. Con questo strumento è infatti possibile verificare il comportamento del dispositivo prima che esso venga prodotto e quindi avere una ragionevole certezza della corretta progettazione. In particolare le immagini di seguito riportate si riferiscono alla fase di verifica finale. In questa fase è possibile, attraverso la simulazione, studiare il comportamento del di-

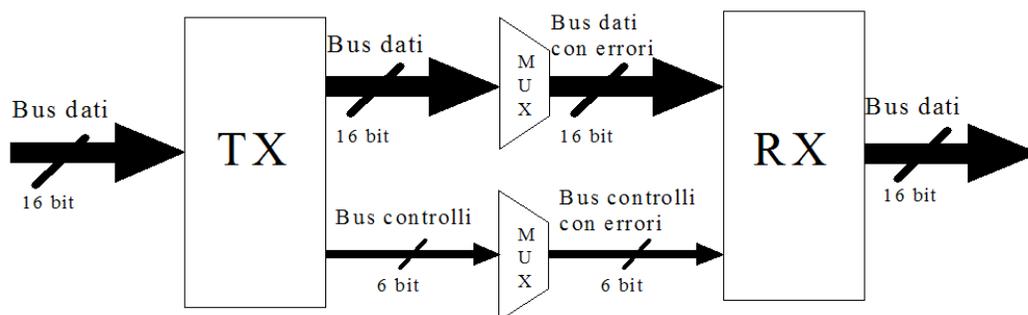


Figura 6.1: Nel grafico viene riportato lo schema utilizzato per effettuare le simulazioni in fase di progetto

spositivo in una configurazione molto simile a quella che assumerà una volta realizzato. I risultati forniscono importanti indicazioni in merito alla correttezza di tutte le fasi progettuali realizzate. Come sarà possibile notare dalle tracce riportate e come esposto nel capitolo precedente, i grafici riportano i ritardi reali di propagazione dei segnali logici. Tali informazioni temporali sono state ottenute a partire dalla valutazione dei componenti parassiti presenti nel dispositivo una volta che i blocchi logici che lo compongono sono stati piazzati e collegati tra loro. Il tempo di propagazione che occorre affinché i segnali logici posti in ingresso al chip producano dei segnali validi in uscita è costituito sia dal tempo caratteristico dei blocchi logici utilizzati per l'elaborazione sia dai componenti parassiti venutisi a creare nel collegare tali componenti tra loro. Tutte queste informazioni calcolate nella fase di verifica finale sono tenute in conto nelle simulazioni riportate.

I grafici presenti in questo lavoro sono stati ottenuti utilizzando una configurazione realizzata appositamente in simulazione per riprodurre il funzionamento reale del dispositivo nella sua configurazione operativa. In particolare in fase di simulazione sono stati testate due copie del chip progettato una configurata come trasmettitore ed una come ricevitore. I due dispositivi sono stati interconnessi tra loro mediante i due bus di dato, a 16 bit, e di controlli, a 6 bit. Ogni linea del bus passa attraverso un multiplexer pilotabile in fase di simulazione. Questo dispositivo riceve in ingresso la linea del bus e consente di produrre in uscita questo ingresso oppure uno 0 o un 1 logi-

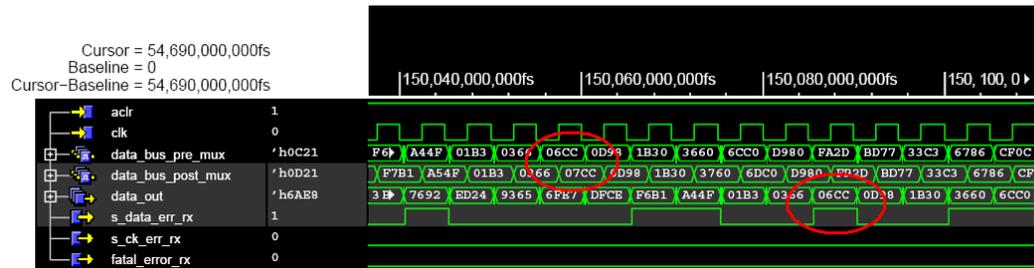


Figura 6.2: La simulazione riporta il comportamento del ricevitore nel caso in cui sul bus dati è presente un bit di errore

co. Agendo sulle linee di controllo in maniera opportuna è possibile quindi indurre artificialmente errori di comunicazione e quindi effettuare prove di tutti i modi di funzionamento del chip oggetto del test. La configurazione utilizzata è schematizzata in figura 6.1.

Il primo grafico riportato (figura 6.2) è relativo ad una trasmissione dati tra trasmettitore e ricevitore durante la quale è indotto un errore su una linea 12 del bus dati. Per fare ciò durante la simulazione è stato pilotato il multiplexer relativo in modo che ignorasse il valore reale della linea e propagasse costantemente un 1 logico. In questo modo ogni qual volta il trasmettitore propaga un livello logico 0 al ricevitore giungerà un valore errato. Questo caso è chiaramente visibile dal grafico.

Nella simulazione sono presenti i bus prima e dopo i multiplexer (rispettivamente *data_bus_pre_mux* e *data_bus_post_mux*). Confrontandole è possibile notare l'errore indotto. Il corretto funzionamento del ricevitore è apprezzabile osservando il bus *data_out*, che rappresenta il valore prodotto in uscita al termine delle operazioni di decodifica e correzione, con i valori dei due bus precedenti. Come evidenziato in figura, il trasmettitore ad un certo istante temporale pone sul bus (*data_bus_pre_mux*) dati la parola

$$06CC_{HEX} = 0000\ 0110\ 1100\ 1100$$

per effetto dell'errore indotto dal multiplexer il ricevitore vede in ingresso (*data_bus_post_mux*)

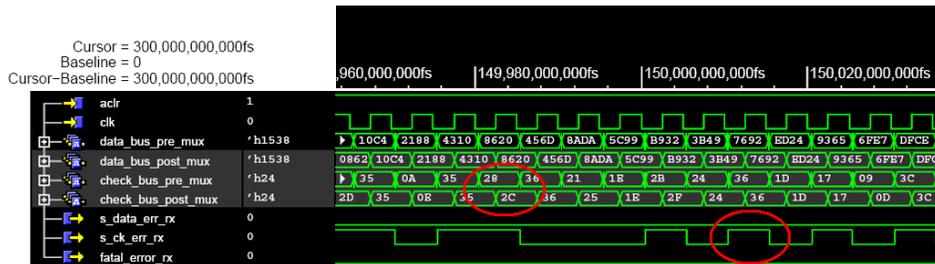


Figura 6.3: La simulazione riporta il comportamento del ricevitore nel caso in cui sul bus dei controlli è presente un bit di errore

$$07CC_{HEX} = 0000 \ 1110 \ 1100 \ 1100$$

Durante le operazioni di decodifica, che come si vede avvengono in un unico ciclo di clock, il chip si accorge dell'errore e corregge il dato propagando in uscita la parola esatta (*bus data_out*)

$$06CC_{HEX} = 0000 \ 0110 \ 1100 \ 1100$$

Contestualmente alla correzione il dispositivo attiva anche la linea di allarme (*s_data_err_rx*) che effettua una transizione $0 \rightarrow 1$. Nel ciclo di clock successivo tale allarme viene disattivato poichè la parola seguente risulta corretta.

Il secondo grafico (figura 6.3) descrive un caso analogo al precedente nel quale, questa volta, è stato indotto un errore su una linea del bus dei controlli. Sempre utilizzando i multiplexer è stata posta la terza linea del bus di controllo sempre al valore 1. In questo modo allorquando il trasmettitore, in relazione alla parola di dato, invia sul bus (*check_bus_pre_mux*) di controlli la parola

$$28_{HEX} = 10 \ 1000$$

in ingresso al ricevitore (*check_bus_post_mux*) si trova la parola

$$2C_{HEX} = 10 \ 1100$$

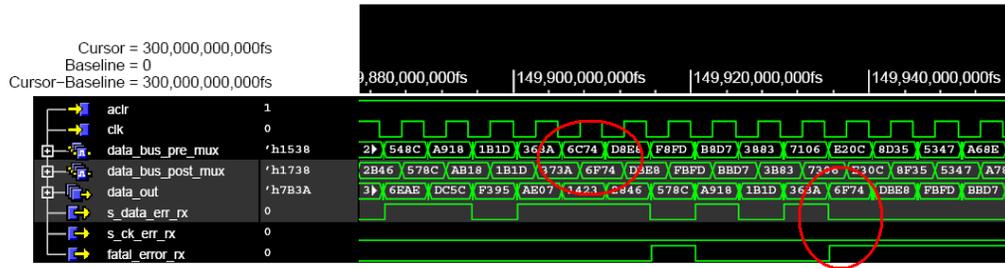


Figura 6.4: La simulazione riporta il comportamento del ricevitore nel caso in cui sono presenti 2 bit errati sul bus dati

Il ricevitore anche questa volta riconosce l'errore indotto e lo segnala mediante l'attivazione della linea d'allarme (*s_ck_err_rx*). Contestualmente pone sul bus dati in uscita la parola ricevuta in ingresso poichè essa risulta corretta.

Nel terzo grafico (figura 6.4) è invece descritto il comportamento del dispositivo allorquando sono indotti errori su due bit del bus dati. In questo caso le linee interessate sono la 9 e la 10 tenute costantemente al valore 1. Quando il trasmettitore invia

$$6C74_{HEX} = 0110\ 1100\ 0111\ 0100$$

per effetto degli errori indotti al trasmettitore arriva

$$6F74_{HEX} = 0110\ 1111\ 0111\ 0100$$

Il dispositivo effettua la decodifica e accorgendosi del doppio errore occorso attiva la linea di allarme *fatal_error_rx*. Dato che, in questo caso, l'algoritmo non è in grado di correggere gli errori presenti nel dato non può far altro che propagare in uscita la parola presente in ingresso e quindi

$$6F74_{HEX} = 0110\ 1111\ 0111\ 0100$$

L'ultimo grafico presentato è relativo al funzionamento del dispositivo in fase di test del bus di comunicazione (figura 6.5). In questo caso i chip sono entrambi configurati in modalità di test. Il trasmettitore invia i dati generati

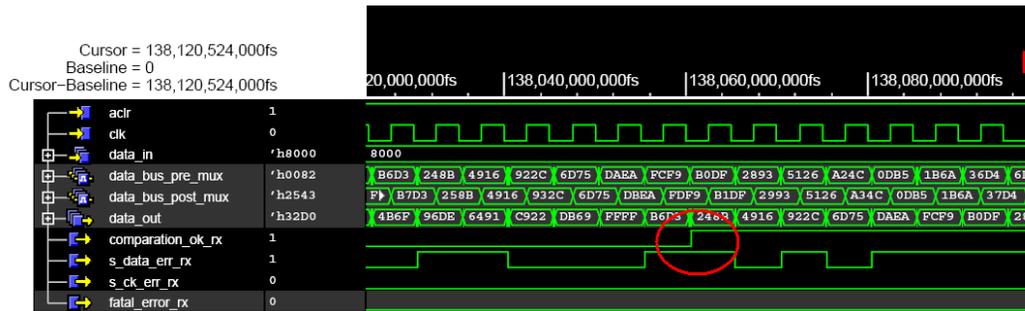


Figura 6.5: La simulazione descrive il funzionamento del chip in modalità di test del canale di trasmissione

dal contatore pseudo-random ed i bit di controlli associati. Il ricevitore effettua la decodifica dei dati e controlla la congruenza dei risultati. Una volta riconosciuta la sequenza attiva la linea *comparison_ok_rx*. Dal grafico si nota come anche in questo caso è stato indotto un errore sul bus dati. Il ricevitore però, è in grado di accorgersi della situazione, correggere tale errore e effettuare il riconoscimento a partire dal dato corretto.

6.3 Piazzamento e produzione del chip

Al termine della progettazione del dispositivo è stato prodotto il *layout* del chip. Il risultato è mostrato in figura 6.6. Sul bordo, in verde, sono raffigurati i pad di ingresso e uscita. Subito all'interno sono presenti due linee parallele che corrono per tutto il perimetro del dispositivo. Queste rappresentano le linee di distribuzione della tensione ai blocchi logici che costituiscono il chip e formano il cosiddetto *ring di alimentazione*. Su di esso sono connesse delle altre linee verticali di colore rosso denominate *strip di alimentazione*. Esse attraversano la zona del core del chip e sono necessarie affinché anche le celle logiche più lontane dal ring siano alimentate correttamente. La zona del core è tutta quella che si trova all'interno del ring di alimentazione. Essendo in questo caso il chip pad limited essa risulta molto più grande di quella necessaria al piazzamento dei blocchi logici. Per questo motivo è stata individuata una zona ristretta dove poter piazzare in maniera compatta tutti i compo-

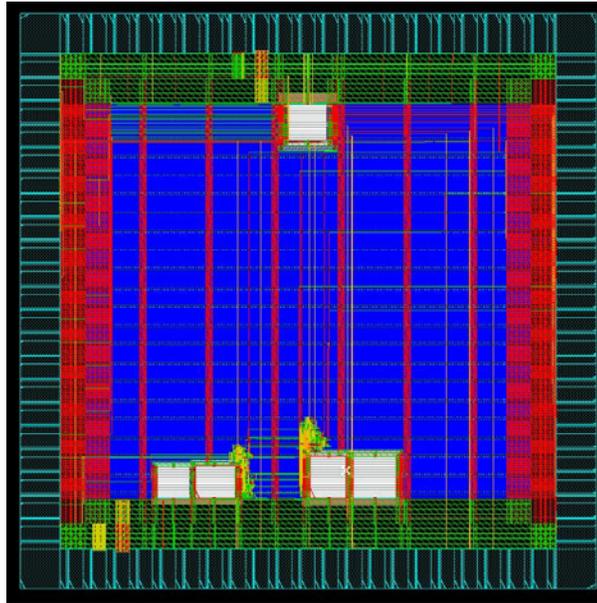


Figura 6.6: Nell'immagine viene riportato il risultato delle operazioni di piazzamento e routing del chip realizzato

nenti logici. Tale zona è colorata di bianco e posta in posizione centrale in alto.

L'immagine 6.7 descrive in dettaglio la zona del core del chip e da un'idea dell'elevato numero di interconnessioni necessarie al funzionamento del dispositivo. I colori differenti con i quali tali linee sono raffigurate è indicativo dei livelli di metal utilizzati. Una volta terminata la fase di progetto il lavoro è stato inviato alla fonderia che ha provveduto alla realizzazione del chip. Un immagine del chip realizzato su silicio è riportata in figura 6.8. La fabbrica ha anche provveduto al suo incapsulamento. Il chip è stato, infatti, posto all'interno di un involucro di materiale plastico denominato *package*. I pad di ingresso e di uscita sono stati connessi per mezzo di una speciale tecnica denominata *bonding*, a dei contatti metallici posti nella parte più esterna del contenitore. Questo incapsulamento si rende necessario per permettere l'utilizzo del chip. Utilizzando il contenitore si protegge il silicio da tutti gli agenti atmosferici che possono comprometterne l'integrità e quindi il funzionamento. Il package è anche fondamentale per la connessione ed il

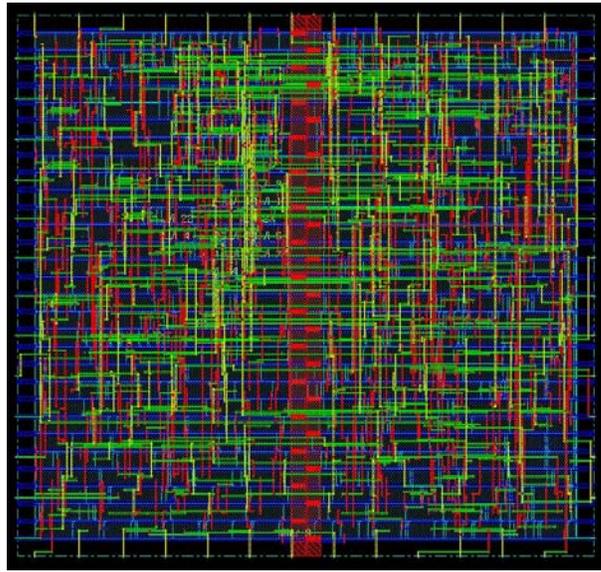


Figura 6.7: Immagine relativa al core del chip al termine della fase di progetto

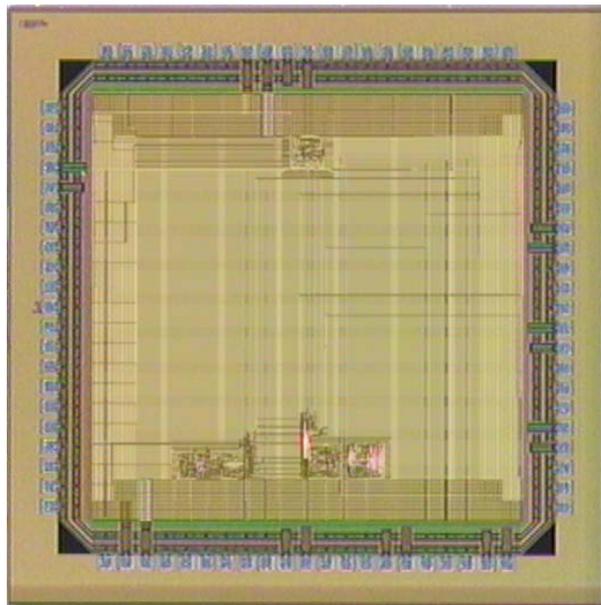


Figura 6.8: Nella foto viene riportato il chip prodotto dalla fonderia.

posizionamento del dispositivo sulla scheda elettronica che lo ospita. Dalla figura 6.8 di destra è infatti possibile rendersi conto delle dimensioni reali

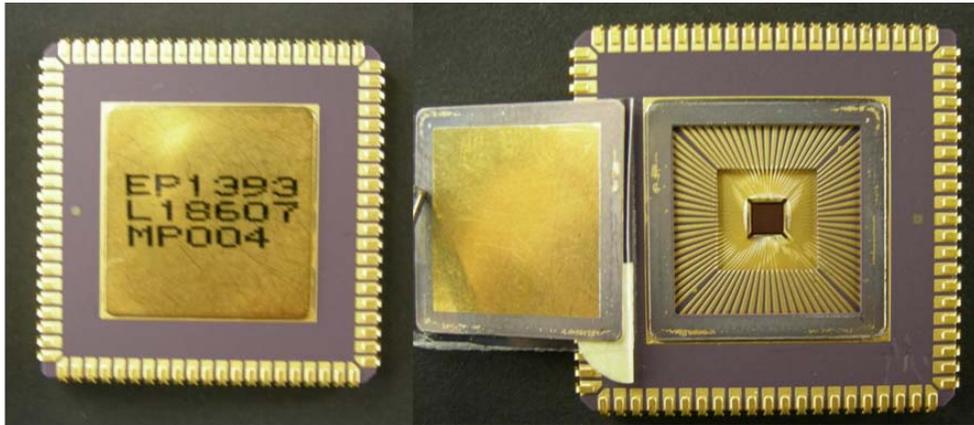


Figura 6.9: L'immagine di sinistra ritrae il chip completo di package mentre nella foto di destra è possibile notare in dettaglio il chip do silicio alloggiato nel contenitore plastico.

dell silicio. L'immagine a sinistra rappresenta il dispositivo pronto all'uso.

6.4 Collaudo del dispositivo

Una volta entrati in possesso del chip è stato necessario effettuare una serie di misure in laboratorio per costatarne il funzionamento e stabilire le reali prestazioni. Per effettuare tali prove ho personalmente progettato una scheda elettronica in grado di ospitare due chip e poterli configurare secondo tutte le modalità di funzionamento previste. La board, mostrata in figura (aggiungere figura scheda), prevede che i due chip siano interconnessi mediante il bus dati ed il bus di controllo. Su tali bus sono stati posizionati una serie di *jumper* in modo che il segnale inviato dal ricevitore possa essere interrotto. In questo modo è possibile propagare al trasmettitore un livello logico diverso simulando la condizione di errore. La progettazione di questa scheda ha necessitato, da parte mia, grande applicazione viste le alte frequenze di trasferimento dei segnali tra i due chip. E' stato necessario porre molta cura affinché i dispositivi potessero funzionare alle massime frequenze possibili senza che venissero introdotte delle limitazioni da parte della scheda. Solo in questo modo si sono potute testare le effettive caratteristiche dei chip.

In figura si nota anche la presenza di una serie di led che servono a dare un'informazione visiva dell'attivazione delle linee di allarme e del segnale di controllo in fase di test. Sono presenti inoltre una serie di test point sui quali prelevare agevolmente i segnali in uscita dai chip e fornire in ingresso i dati voluti.

La campagna di misura è stata effettuata in due fasi:

- collaudo del funzionamento ordinario
- collaudo in configurazione di test

Il collaudo del funzionamento ordinario è stato effettuato ponendo una serie di dati in ingresso al trasmettitore per mezzo di uno strumento capace di generare sequenze di 16 bit in parallelo ad alta frequenza. I dati propagati al ricevitore e da esso elaborati e posti in uscita sono poi stati acquisiti mediante un secondo strumento in grado di riconoscere tali sequenze. Mediante questo strumento è stato possibile inviare tali dati ad un computer con il quale i dati sono stati analizzati. Chiaramente gli strumenti utilizzati sono estremamente sofisticati e performanti poichè devono essere in grado di raggiungere prestazioni molto superiori a quelle ottenibili dai chip. Solo in questo caso si è sicuri di testare le effettive prestazioni dei chip realizzati. Anche durante la fase di collaudo sono stati indotti degli errori sui bus in modo da constatare il funzionamento degli algoritmi di codifica e decodifica.

Di seguito vengono presentate una serie di acquisizioni effettuate per mezzo di un oscilloscopio che mostrano il funzionamento nella configurazione descritta.

La figura 6.10 mostra il funzionamento del ricevitore in presenza di un errore sul bus dati. La traccia numero 4 (raffigurata in verde) mostra un bit di dato propagato dal trasmettitore una volta terminate le operazioni di codifica dell'informazione. La traccia numero 3 (in viola) rappresenta il valore del suddetto bit in ingresso al ricevitore su cui è stato indotto l'errore. Le traccia 1 riprodotta in blu mostra il dato posto in uscita dal ricevitore una volta effettuata la decodifica e la correzione mentre la traccia 2 (in azzurro) rappresenta lo stato della linea di allarme della presenza di un errore sul bus

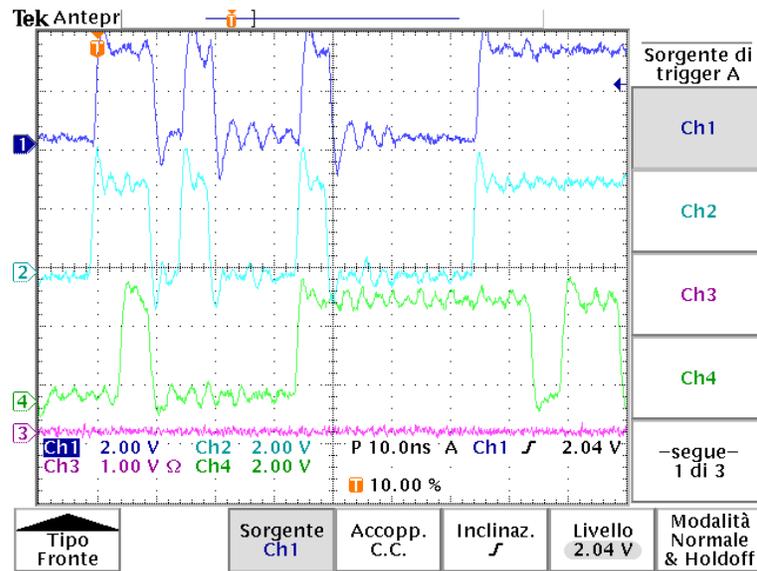


Figura 6.10: L'immagine, acquisita per mezzo di un oscilloscopio, mostra il comportamento del chip in ricezione quando sul bus dati viene scoperto un errore.

dati. Come è possibile notare le tracce 1 e 2 sono perfettamente identiche. Ciò in effetti dimostra la correttezza dell'algoritmo. Dato che per indurre l'errore il bit in esame è stato posto permanentemente a 0 ogni qual volta lo stream di dati in arrivo dal trasmettore prevede la presenza del valore 1 sulla linea osservata viene generato un errore. Per questo motivo il ricevitore corregge il dato e lo pone correttamente in uscita (traccia 1) attivando contemporaneamente l'allarme.

Una condizione analoga è stata prodotta per indurre errori sul bus dei controlli. La figura 6.10 mostra nella traccia 1 il dato corretto in uscita da ricevitore e la contemporanea attivazione della linea di allarme per i bit di controllo ogni qual volta si verifica un errore sul quel bus.

In ultimo è stato effettuato il collaudo della modalità di funzionamento dei dispositivi in configurazione di test. La figura 6.12 mostra il funzionamento corretto del ricevitore in tale modalità. In particolare la situazione descrive il risultato delle operazioni di decodifica e di comparazione del dato effettuate dal ricevitore. La traccia di colore blu (numero 1) rappresenta il segnale di reset inviato dall'esterno ai chip. Quando tale linea effettua la

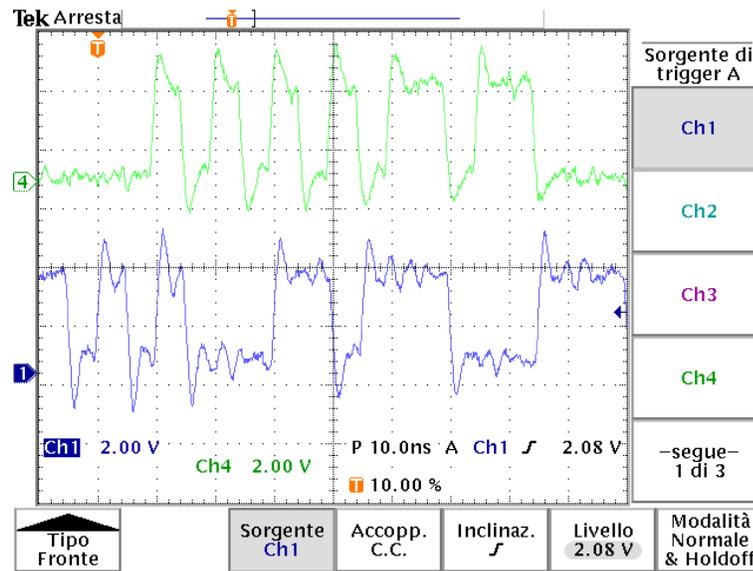


Figura 6.11: L'immagine, acquisita per mezzo di un oscilloscopio, mostra il comportamento del chip in ricezione quando sul bus dei controlli viene scoperto un errore.

transizione $0 \rightarrow 1$ il trasmettitore avvia la fase di codifica e propagazione dei dati estratti dal generatore pseudo-random. Il ricevitore avvia la decodifica e la comparazione del dato in arrivo. Allorquando i dati in arrivo vengono riconosciuti corretti viene attivato il segnale rappresentato della traccia verde (numero 4).

Come è possibile notare le acquisizioni riportate sono state effettuate facendo funzionare il chip alla massima frequenza operativa prevista e cioè fornendo al trasmettitore ed al ricevitore un segnale di clock alla frequenza di 200 MHz.

6.5 Considerazioni finali

Al termine della campagna di misure effettuate in laboratorio per eseguire il collaudo del chip realizzato è stato possibile constatare che:

- il chip progettato funziona perfettamente in tutte le modalità operative;
- le specifiche di progetto imposte sono ampiamente rispettate;

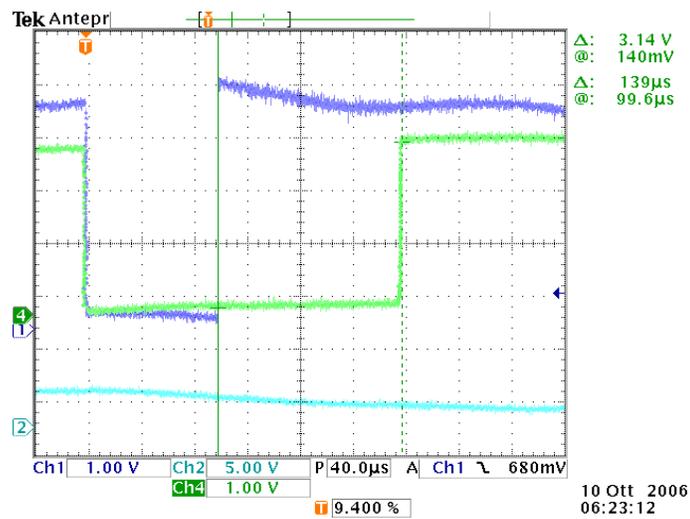


Figura 6.12: L'immagine mostra l'attivazione della linea che segnala l'avvenuto riconoscimento della sequenza inviata dal ricevitore.

- la massima frequenza di funzionamento è di 210 MHz;
- il consumo del singolo chip è pari a 23 mW.

Conclusioni

L'esperimento NEMO ha come obiettivo principale la realizzazione di un rivelatore sottomarino per neutrini cosmici. Con esso sarà possibile studiare in maniera innovativa il cosmo permettendo di fare luce su fenomeni ancora oggi misteriosi e sconosciuti. Per ottenere questi obiettivi ambiziosi tuttavia è necessario fronteggiare una serie di difficoltà derivanti dalla realizzazione di un telescopio per neutrini posto in mare ad una profondità di 3000 metri. Proprio l'installazione di apparecchiature tecnologiche tanto sofisticate in un ambiente così ostile rappresenta una sfida estremamente difficile. In particolare sarà necessario garantire che i dati acquisiti nella grandi profondità marine vengano trasferiti a terra garantendo elevati standard di affidabilità. Per questo motivo il gruppo di ricerca di Napoli, ed io in prima persona, ha studiato metodi in grado di migliorare i sistemi di comunicazioni utilizzati nel telescopio. Il mio lavoro di tesi è stato quindi suddiviso in due parti.

Nella prima parte ho analizzato i metodi di comunicazione utilizzati nell'esperimento studiando in particolare il grado di rumorosità che con essi è possibile raggiungere. Effettuando questo studio è stato possibile individuare i margini di miglioramento ottenibili e le tecniche migliori per ottenerli. Ho quindi individuato e scelto l'approccio fornito dalla teoria dell'informazione e dei codici. Esso permette il miglioramento dell'affidabilità del canale di comunicazione mediante l'impiego di algoritmi matematici in grado di rivelare errori in trasmissione e successivamente correggerli. Una volta convinto dell'approccio da utilizzare ho selezionato il codice più adatto all'impiego all'interno dell'esperimento NEMO. L'algoritmo deve infatti sia aumentare l'affidabilità del canale di comunicazione mediante un numero contenuto di informazioni ridondanti ma, al contempo, deve impiegare un tempo di

elaborazione estremamente contenuto in modo da annullare il tempo morto nella comunicazione. Queste condizioni hanno anche indotto la scelta della piattaforma tecnologica da utilizzare per la realizzazione del dispositivo che svolgesse le funzioni di codifica e decodifica dell'informazione imposte da tale algoritmo. Il dispositivo doveva infatti soddisfare una serie di condizioni tecnologiche: doveva essere compatto e affidabile ma al contempo doveva essere in grado di raggiungere elevate prestazioni in termini di velocità operativa. Tutto ciò doveva essere ottenuto mantenendo i consumi estremamente contenuti. Nel determinare tali specifiche tecniche ho posto notevole attenzione nel progettare un dispositivo che fosse ovviamente perfettamente compatibile con l'elettronica del rivelatore ma contemporaneamente fosse di utilizzo generale. Il dispositivo realizzato, infatti, risponde ad una serie di esigenze comuni ad ambiti molto diversi che lo rendono impiegabile, senza bisogno di alcuna modifica, anche in campi non strettamente legati alla ricerca.

La seconda parte del lavoro mi ha visto impegnato nella progettazione e nella realizzazione del chip ASIC digitale che sviluppa le funzioni dell'algoritmo selezionato con le caratteristiche descritte. La tecnologia ASIC risulta estremamente impegnativa da utilizzare e quindi il lavoro di progettazione è stato complesso ma mi ha permesso di raggiungere gli ambiziosi obiettivi stabiliti inizialmente. In particolare il chip realizzato è utilizzabile in tre diverse modalità operative. Esso può svolgere funzione di codifica e decodifica descritte dall'algoritmo di Hamming e quindi può essere impiegato sia come trasmettitore che come ricevitore. In questo modo è possibile utilizzare due copie del dispositivo da porre all'inizio ed alla fine del bus di comunicazione ottenendo un notevole aumento dell'affidabilità della trasmissione dati. Inoltre questo chipset può essere anche utilizzato per valutare la rumorosità del canale di comunicazione su cui è installato. Configurando i chip in una particolare modalità e monitorando le linee di allarme è possibile ottenere una stima precisa dell'occorrenza di errori sul bus di comunicazione e quindi avere una stima precisa dell'affidabilità del canale.

La campagna di misure effettuata al termine del lavoro di progettazione e realizzazione mi ha poi permesso di valutare le reali prestazioni del chip. Ho potuto constatare che esso è in grado di funzionare alle massime frequenze

di trasferimento previste dall'elettronica dell'esperimento garantendo comunque una dissipazione di corrente tanto bassa da non comportare alcun tipo di problema di alimentazione. I test effettuati in laboratorio hanno quindi permesso di constatare il pieno raggiungimento delle specifiche di progetto. Per questo motivo il chip è pronto per essere integrato con l'elettronica del telescopio NEMO. Le ottime prestazioni che il dispositivo può ottenere lo rendono tuttavia estremamente versatile. Esso può essere utilizzato, senza alcun tipo di modifica, in tutti quegli ambiti in cui è necessario aumentare l'affidabilità di trasmissione su canali rumorosi anche quando si ha a che fare con velocità elevate di trasmissione dati.

Bibliografia

- [1] Hess, *V.F. Zs. Phys* **13**, 1084 (1912).
- [2] R.A. Millikan and G. H. Cameron, *Phys. Rev.* **28** (1926).
- [3] B. Wiebel-Sooth, P. L. Biermann and H. Meyer, *Cosmic Rays. Individual element spectra: prediction and data. Astron. and Astrophs.* (Sett. 1997).
- [4] R. M. Barnett et al., *Review of particle physics, Phys. Rev* **D54(1,I):1-720** (July 1996).
- [5] Malcom S. Logair, *High energy Astrophysics vol 1 & 2* Cambridge university press, second edition. (1994)
- [6] *The european Physical Journal C*, Volume 3, 1-4 (1998).
- [7] J. A. Simpson, *Ann. Rev. Nucl. Part. Sci.* **33**, 323 (1983).
- [8] G. T. Zatsepin and V. A. Kuzmin, *JETP Lett.* **4**, 78 (1966) [*Pisma Zh. Eksp. Teor. Fiz.* **4** 114 (1966)].
- [9] K. Greisen, *Phys. Rev. Lett.* **16**, 748 (1966).
- [10] A. Levinson and E. Waxman, *Phys. Rev. Lett.* **87.171101** (2001).
- [11] M.A. Markov et. al, *Nucl. Phys.* **27, 385** (1961).
- [12] B.A. Harmon et. al, *Astrphys. Journ.* **47, L85** (1997).
- [13] F. Bohen, P. Voegel, *Physics of Massive Neutrinos.* Cambridge Press (1992).
- [14] J. Bahcall, *Neutrino Astrophysics.* Cambridge Press (1989).
- [15] R. Protheroe, *Astro-ph.* **9809144** (1998).

-
- [16] DUMAND web site: <http://www.phys.washington.edu/dumand>
- [17] Baikal web site: <http://www.ifh.de/baikal/baikalhome.html>
- [18] E. V. Bugaev, *Phys Rev.* **D58**, **054001** (1998).
- [19] A. Capone et al. Measurements of light transmission in deep sea with the ac9 transmissometer. *Nucl. Instrum. Meth A*, **487(3)**: **423-434**(July 2002).
- [20] C. Smith and K. Baker, *Applied Optics*. (1965).
- [21] F. Ameli M. Bonori and F. Massa *The European Physical Journal C*, **25(1)**: **67-75**(2002).
- [22] National Semiconductor LVDS web page:
<http://www.national.com/appinfo/lvds>
- [23] David J. C. Mac kay Information Theory, Interference, and Learning Algorithms *Cambridge University Press* (2003).
- [24] R. W. Hamming Error Detecting and Error Corretting Codes *The Bell System Technical Journal* **vol 29**, **147 - 160**(1950).