



UNIVERSITÀ DEGLI STUDI DI CAGLIARI
FACOLTÀ DI SCIENZE MATEMATICHE FISICHE E NATURALI
CORSO DI LAUREA IN FISICA

**FILTRI SOFTWARE PER L'ANALISI DI SEGNALI
ACUSTICI GENERATI DA INTERAZIONI DI
NEUTRINI DI ALTISSIMA ENERGIA IN MARE**

Relatori:

Prof. Biagio Saitta

Dott. Alessandro Cardini

Tesi di laurea di:

Andrea Contu

ANNO ACCADEMICO 2004-2005

Indice

1. Introduzione	4
2. L'effetto termoacustico.....	7
3. Il rumore.....	12
4. Filtraggio del segnale: il filtro adattato.....	19
5. Implementazione del filtro adattato.....	26
6. Analisi dei risultati.....	28
<i>6.1. L'uscita del filtro.....</i>	<i>28</i>
<i>6.2. Analisi statistica: rumore gaussiano bianco.....</i>	<i>31</i>
<i>6.3. Analisi statistica: rumore gaussiano colorato.....</i>	<i>36</i>
<i>6.4. Analisi statistica: rumore da NEMO.....</i>	<i>38</i>
7. Conclusioni	41
Bibliografia	42
<i>Ringraziamenti</i>	43

Indice delle figure

Figura 1.1 – Flusso differenziale dei raggi cosmici per $E > 10^{11} eV$	5
Figura 2.1 – Schema del processo interazione-rivelazione di un segnale acustico prodotto da una cascata elettromagnetica.....	7
Figura 2.2 – Profilo della temperatura dell'acqua per la regione tropicale di AUTECH in funzione della profondità. Le due curve corrispondono a due diversi periodi dell'anno.	9
Figura 2.3 – Simulazione del segnale acustico prodotto da una cascata adronica di energia $E = 10^{20}$ eV alla distanza di 1km dalla sorgente.....	10
Figura 3.1 – Schematizzazione del processo di rumore additivo in una trasmissione.....	12
Figura 3.2 – Densità di potenza spettrale di un rumore bianco in funzione della frequenza.	14
Figura 3.3 – Particolare realizzazione del rumore bianco generato via software.	15
Figura 3.4 – Distribuzione delle ampiezze del rumore bianco generato via software in figura 3.3.	16
Figura 3.5 – Particolare realizzazione di un rumore gaussiano con potenza spettrale decrescente come $1/f$	17
Figura 3.6 – Densità di potenza spettrale del rumore registrato da NEMO. La durata della registrazione è di 15 s e la frequenza di campionamento di 96 kHz.....	18
Figura 6.1 – Segnale utilizzato nella simulazione del programma in allegato.	29
Figura 6.2 – Segnale di figura 6.1 addizionato a un rumore gaussiano bianco.	29
Figura 6.3 – Filtraggio tramite filtro adattato del segnale in figura 6.2.....	30
Figura 6.4 – Distribuzione di g_{max} ricavata su 1000 realizzazioni del rumore	31

Figura 6.5 – Distribuzione statistica dell'istante temporale di rivelazione rispetto a quello atteso ($z=0$). In ascissa si riporta z e in ordinata i conteggi. L'energia del rumore è 0.1 U.E.	32
Figura 6.6 – Andamento di g_{\max} per segnali disadattati. Ogni punto (per 100 punti) è stato calcolato effettuando una media su 100 realizzazioni del rumore.	33
Figura 6.7 – Falsi positivi al secondo in funzione della soglia. Ogni punto è stato calcolato effettuando una media su 100 realizzazioni del rumore.	34
Figura 6.8 – Efficienza del filtro relative a una soglia di 3.3 U.A.. Ogni punto è stato calcolato effettuando una media su 1000 realizzazioni del rumore.	35
Figura 6.9 - Falsi positivi al secondo in funzione della soglia nel caso di un rumore $1/f$. Ogni punto è stato calcolato effettuando una media su 100 realizzazioni del rumore.	36
Figura 6.10 – Efficienza del filtro nei due casi di rumore gaussiano bianco (in blu) e di rumore gaussiano $1/f$ (in rosso). Le due curve sono relative a una soglia corrispondente a un falso positivo al secondo; i valori di soglia sono 3.3 U.A. per il rumore bianco e 3.8 U.A. per il rumore colorato.	37
Figura 6.11 - Falsi positivi al secondo in funzione della soglia. Ogni punto è stato calcolato effettuando una media su 65 realizzazioni del rumore.	38
Figura 6.12 – Distribuzione delle ampiezze del rumore di NEMO dopo il filtraggio.	39

1. Introduzione

Lo studio di raggi cosmici di altissima energia, UHECR (ultra-high energy cosmic rays), con una energia $E > 10^{17}$ eV, e in particolare di neutrini, rappresenta uno dei più interessanti campi di ricerca nell'astrofisica moderna. Fino ad ora sono stati osservati una decina di eventi compatibili con raggi cosmici di energia di circa 10^{20} eV [1].

I meccanismi responsabili della accelerazione di particelle a energie così alte non sono tuttora chiari e inoltre i dati attualmente disponibili sembrano essere in contrasto [3] col limite GZK (Greisen-Zatsepin-Kuzmin). Esso prevede che la massima energia per protoni di origine cosmologica sia $\sim 10^{20}$ eV, a causa della loro interazione con la radiazione cosmica di fondo essi perdono la loro energia per scattering anelastico in ~ 50 Mpc. La componente di neutrini nei UHECR non è nota anche se i dati presenti sembrano indicare che essi non ne sono la componente principale, tuttavia una quantità sostanziale potrebbe provenire principalmente da esplosioni gamma e da nuclei di galassie attive, oltre a quelli prodotti per effetto GZK, ossia l'interazione di protoni altamente energetici con la radiazione cosmica di fondo [7]. Inoltre ci si attenderebbe un flusso di neutrini da annichilazione di particelle della materia oscura che dovrebbe permeare l'Universo.

L'astronomia dei neutrini presenta indubbiamente molti vantaggi, infatti queste particelle sono in grado di attraversare l'intero Universo senza interagire e poiché sono neutri non vengono deflessi, quindi è possibile risalire alla loro direzione. Inoltre, pur essendo prodotti dalle stesse sorgenti dei raggi cosmici, una volta prodotti non interagiscono più con la sorgente stessa come invece avviene per i fotoni.

Tuttavia vi sono anche degli svantaggi non trascurabili, e cioè la piccolissima sezione d'urto e la necessità di rivelatori di dimensioni enormi.

Questo è necessario perché il flusso di queste particelle è tipicamente $100 \text{ km}^{-2} \text{ anni}^{-1}$ per energie attorno ai $\sim 10^{18} \text{ eV}$, $1 \text{ km}^{-2} \text{ anni}^{-1}$ per energie di $\sim 10^{19} \text{ eV}$, $1 \text{ km}^{-2} \text{ secoli}^{-1}$ per energie di $\sim 10^{20} \text{ eV}$. La figura 1.1 mostra il flusso differenziale dei raggi cosmici in funzione della loro energia.

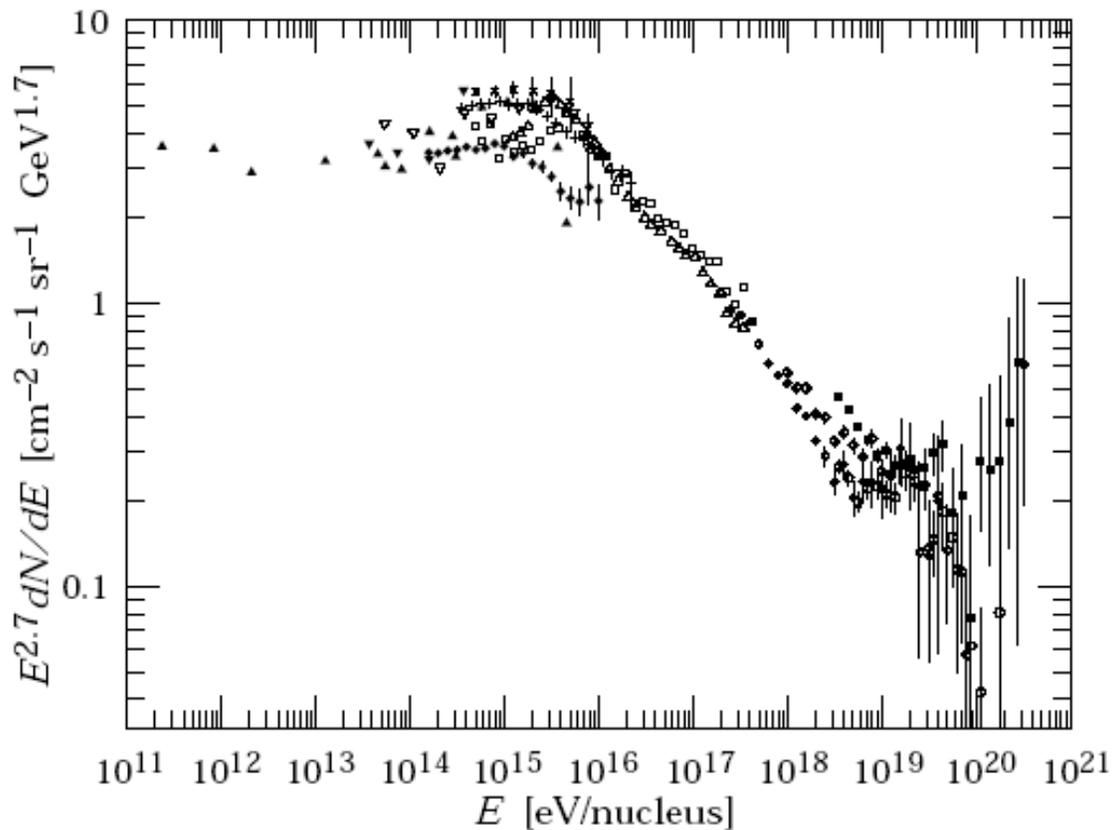


Figura 1.1 – Flusso differenziale dei raggi cosmici per $E > 10^{11} \text{ eV}$.

I metodi più importanti sono la rivelazione tramite emissione di luce Čerenkov e quella per effetto termoacustico che è anche l'oggetto di questa tesi.

Il primo metodo si basa sulla rivelazione di radiazione Čerenkov prodotta dalle particelle dello sciame provocato dall'interazione di un UHECR, in questo contesto il fenomeno può essere osservato sia in mare che nell'atmosfera. Il secondo metodo sfrutta l'effetto termoacustico, che consiste nella produzione di onde sonore in un fluido a seguito di un

repentino riscaldamento del fluido stesso causato dalla deposizione di una certa quantità di energia a opera delle particelle in questione; esso verrà presentato più in dettaglio nel prossimo paragrafo.

Attualmente esistono degli esperimenti che sfruttano l'effetto termoacustico, un esempio è AUTECH (Atlantic Undersea Test and Evaluation Center) che sfrutta idrofoni della marina esistenti in prossimità delle coste della Florida. Le profondità alle quali sono collocati gli idrofoni di AUTECH sono di 1-2 km.

2. L'effetto termoacustico

Ci si limita ora a considerare il caso in cui le particelle in questione siano neutrini e che l'interazione avvenga in acqua.

Data l'energia di tali particelle esse interagiscono per deep inelastic scattering con i quarks dei nuclei di ossigeno e di idrogeno. Questo processo di scattering produce una cascata elettromagnetica le cui caratteristiche, che ora non è necessario trattare, dipendono dal tipo di neutrino incidente.

Tipicamente una cascata di questo tipo deposita la propria energia in acqua in un cilindro di raggio di circa 20 cm e lunghezza di 20 m.

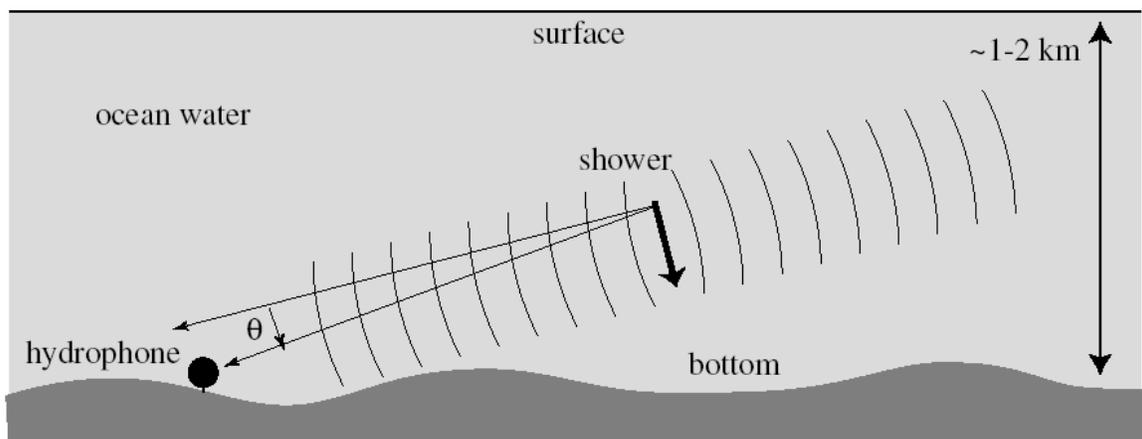


Figura 2.1 – Schema del processo interazione-rivelazione di un segnale acustico prodotto da una cascata elettromagnetica.

Il suono è prodotto essenzialmente dal riscaldamento localizzato lungo la cascata che si traduce in una espansione del fluido; lo sviluppo della cascata avviene a velocità relativistiche, lungo la direzione di moto della particella che l'ha generata (figura 2.1), e può quindi essere considerato istantaneo rispetto al fenomeno acustico.

La propagazione delle onde di pressione avviene coerentemente sui piani perpendicolari all'asse della cascata, in generale vi sarà quindi una dipendenza del segnale, soprattutto per quanto riguarda la sua ampiezza, dall'angolo di osservazione. Esperimenti agli acceleratori confermano l'esistenza del fenomeno termoacustico anche se il tutto viene realizzato con fasci a energie molto più basse.

Si introducono ora l'energia visibile dello sciame, E_0 , e l'energia depositata per unità di volume e di tempo, $E(\mathbf{r},t)$, legate dalla seguente relazione:

$$E_0 = \iiint E(\mathbf{r},t) d^3 \mathbf{r} dt$$

L'equazione d'onda per un segnale di pressione, considerando l'assorbimento da parte del mezzo, può essere scritta in questo modo [2]:

$$\nabla^2 \left(p + \frac{1}{\omega_0} \dot{p} \right) - \frac{1}{c^2} \ddot{p} = - \frac{\beta}{C_p} \frac{\partial E}{\partial t}$$

dove p rappresenta la pressione e E la densità di energia. Nel caso dell'acqua si ha che la velocità del suono $c \approx 1500$ m/s, il coefficiente di espansione termica $\beta \approx 1.2 \cdot 10^{-4}$ K⁻¹, il calore specifico a pressione costante $C_p \approx 4.2 \cdot 10^3$ J kg⁻¹ K⁻¹, e la frequenza caratteristica di attenuazione $\omega_0 \approx 2.5 \cdot 10^{12}$ s⁻¹, supposta costante nell'intervallo di frequenze considerato (10^4 - 10^5 Hz); tutte le grandezze sono relative all'acqua marina ad una temperatura di 1.5°C, tipica delle profondità a cui si trovano gli idrofoni in questo tipo di esperimenti.

In realtà β è in generale una funzione della temperatura e tende a zero all'avvicinarsi della temperatura dell'acqua ai -3°C, questo dà luogo alla necessità di effettuare questi studi in acque preferibilmente calde, quali quelle tropicali (figura 2.2).

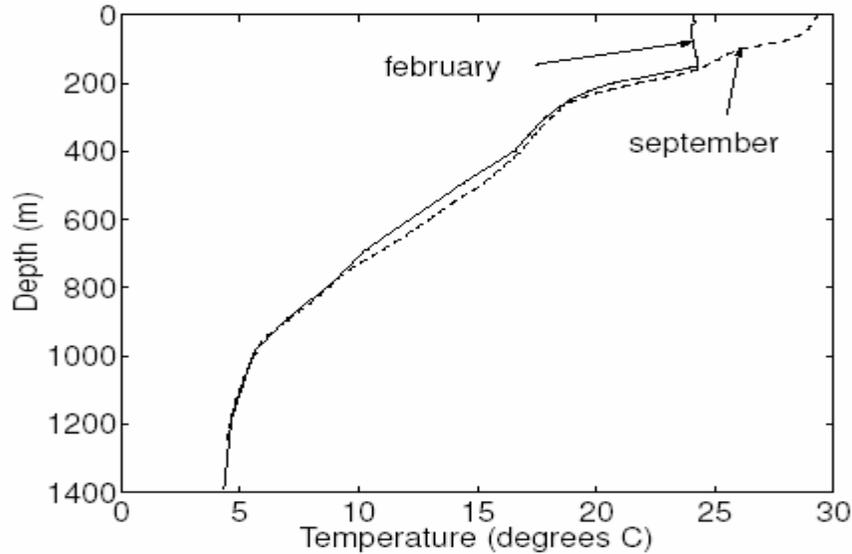


Figura 2.2 – Profilo della temperatura dell’acqua per la regione tropicale di AUTECH in funzione della profondità. Le due curve corrispondono a due diversi periodi dell’anno.

Risulta interessante cercare una soluzione dell’equazione d’onda nel caso la densità di energia depositata sia gaussiana nel tempo e localizzata spazialmente in un punto, ad esempio nell’origine, del tipo $E(0,\sigma)$. La soluzione ha una forma di questo tipo [2]:

$$p(r,z) = -\frac{AE_0}{r} \frac{z}{\sqrt{2\pi\tau^3}} e^{-\frac{z^2}{2\tau^2}}$$

dove r è la distanza dalla sorgente, t il tempo, $z = t-r/c$ con $t > 0$, $A = \beta / 4\pi C_p$ e τ definito come:

$$\tau = \sqrt{\sigma^2 + \frac{r}{\omega_0 c}}$$

nell’approssimazione di r grandi ($\gg \omega_0 c \sigma^2$), cioè dell’ordine del chilometro, si ottiene che:

$$p(r \gg \omega_0 c \sigma^2, z) = -(\omega_0 c)^{3/2} \frac{AE_0}{r^{5/2}} \frac{z}{\sqrt{2\pi}} e^{-(\omega_0 c/2r)z^2}$$

La forma del segnale di pressione è quello di una derivata di una gaussiana quindi avremo due estremi che saranno, sempre nell'approssimazione di r grande, nei punti

$$z_{12} = \pm \sqrt{\left(\frac{r}{\omega_0 c}\right)}$$

con i due valori uguali in modulo e pari a

$$|p_m| = \frac{AE_0}{r^2} \frac{\omega_0 c}{\sqrt{2\pi e}}$$

la forma del segnale è mostrata nella figura 2.3 [3].

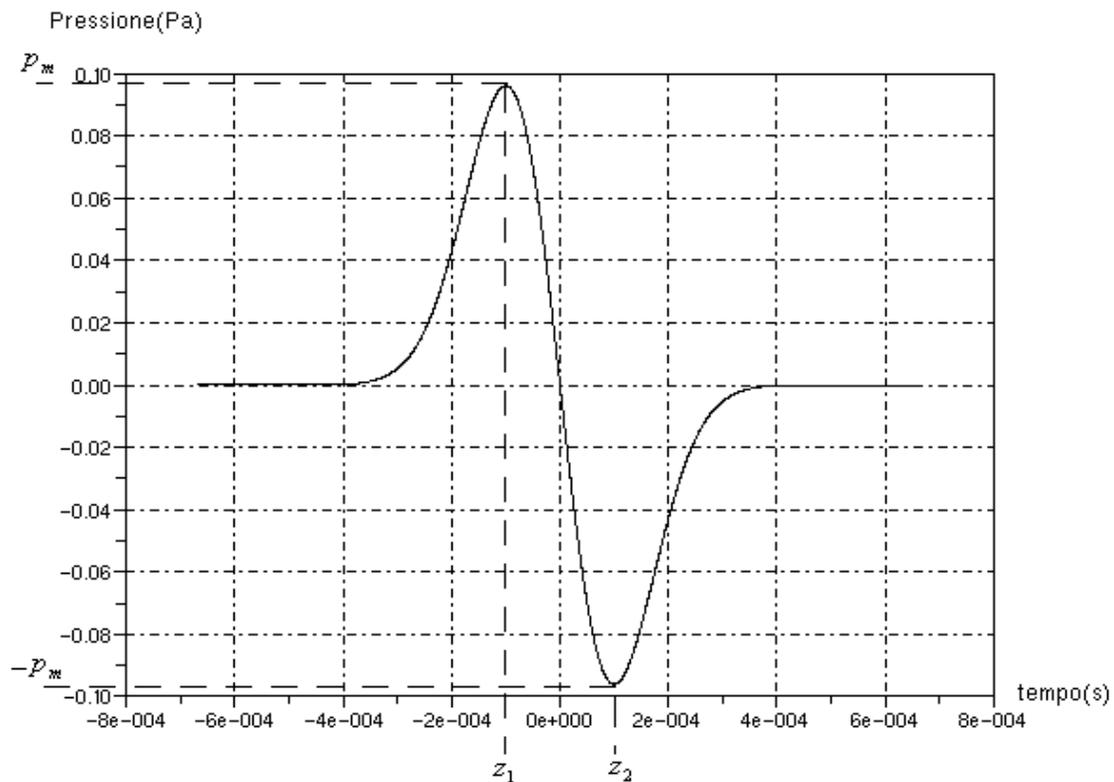


Figura 2.3 – Simulazione del segnale acustico prodotto da una cascata adronica di energia $E = 10^{20}$ eV alla distanza di 1km dalla sorgente.

L'ampiezza del segnale decresce all'aumentare della distanza come $1/r^2$ per cui il segnale si attenua allontanandosi dalla sorgente. La distanza temporale tra i picchi (Δz) cresce come \sqrt{r} , quindi se approssimativamente si considera l'inverso di questo intervallo come il doppio della frequenza caratteristica del segnale si ha che il segnale si sposta verso frequenze più basse all'aumentare della distanza di osservazione.

Un segnale di questo tipo verrà utilizzato come modello per l'applicazione del filtro software.

3. Il rumore

In ogni trasmissione di segnali è inevitabile che l'oggetto della trasmissione, che in generale è un qualche tipo di informazione, appaia al ricevitore, dopo essere passata per un canale di trasmissione (per ora astratto), come una somma di due componenti, una che costituisce il segnale che effettivamente si è voluto trasmettere, l'altra che è del tutto casuale e costituisce il cosiddetto rumore.

Quanto detto si può schematizzare come segue:

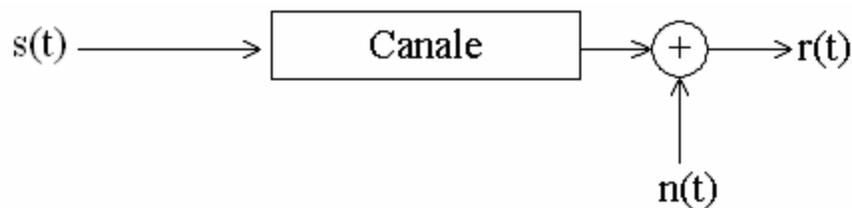


Figura 3.1 – Schematizzazione del processo di rumore additivo in una trasmissione.

che equivale a dire

$$r(t) = s(t) + n(t)$$

dove $r(t)$ rappresenta la trasmissione ricevuta, $s(t)$ il segnale e $n(t)$ il rumore.

Questa espressione vale per ogni istante sia nel caso continuo sia nel caso in cui i dati siano campionati (in effetti il caso continuo può essere considerato un campionamento a frequenza infinita), anche se va precisato che la relazione scritta si riferisce al caso più comune di rumore *additivo*, in generale la relazione potrebbe essere più complicata.

Appare che al ricevitore sia necessario elaborare la trasmissione in modo da separare queste due componenti o almeno da poter “mettere in risalto” il segnale vero e proprio rispetto al rumore. Questo naturalmente nel caso in cui l’importanza delle due componenti sia confrontabile e non sia possibile una facile identificazione del segnale.

È già stato anticipato come il rumore sia un processo *casuale* per cui ciascun campionamento della variabile $n(t)$ è una variabile aleatoria che sarà regolata da una certa densità di probabilità, da un valor medio e una deviazione standard, che possono essere differenti per ogni istante di campionamento.

Quindi si considera ora ogni punto di campionamento come una variabile aleatoria a se stante. Tuttavia nella maggior parte dei casi si ha a che fare con processi *stazionari* ed *ergodici*. Per processo stazionario si intende sostanzialmente un processo la cui realizzazione non dipende dal particolare istante di osservazione; questo significa ad esempio che a qualsiasi istante si inizi a campionare un rumore esso mantiene sempre le stesse caratteristiche. Un processo stazionario è anche ergodico se la media temporale calcolata su di una qualunque realizzazione del processo, coincide con la media di insieme relativa ad una variabile aleatoria estratta ad un istante qualsiasi da una realizzazione qualsiasi; questo ad esempio significa che per calcolare il valor medio di qualsiasi campionamento è sufficiente eseguire una media temporale dei dati campionati, o più esattamente che ogni campionamento è statisticamente rappresentativo di tutti gli altri. Ovviamente esistono delle formulazioni matematiche molto più precise dei concetti esposti sopra tuttavia quanto detto è sufficiente per proseguire.

L’ergodicità e quindi la stazionarietà non costituiscono una forte restrizione in quanto la nella maggior parte dei casi reali, come quello oggetto di questa tesi, queste proprietà sono sempre verificate in buonissima approssimazione. D’ora in poi ci si riferirà a processi di questo tipo.

Particolarmente interessante risulta il caso di rumore additivo gaussiano bianco (AWGN, Additive White Gaussian Noise), dove additivo ha lo stesso significato di sopra; gaussiano significa che la densità di probabilità che regola ogni campionamento e il processo nel complesso è appunto una gaussiana e bianco sta a significare che la densità di potenza spettrale (il modulo quadro della trasformata di Fourier dei campionamenti) è

costante per tutte le frequenze. Il caso di AGWN rispecchia sostanzialmente un rumore provocato da un numero infinito di cause.

Verranno ora introdotte alcune quantità che riguardano i segnali in genere ma ci si riferisce specificamente a processi stocastici.

Data una variabile aleatoria $x(t)$ si definisce l'autocorrelazione, $R_{XX}(t_1, t_2)$:

$$R_{XX}(t_1, t_2) = E[x(t_1)x(t_2)] = R_X(\tau)$$

dove E rappresenta l'operatore valor medio e $\tau = t_1 - t_2$; vi è dipendenza dal solo intervallo temporale tra le due variabili in quanto ci si riferisce a processi stazionari. L'autocorrelazione dà in qualche modo una misura di quanto le due variabili sono legate statisticamente.

Dall'autocorrelazione può essere ricavata la densità di potenza spettrale, $S_X(f)$:

$$S_X(f) = F[R_X(\tau)]$$

dove F rappresenta l'operatore che esegue la trasformata di Fourier.

Per le proprietà delle trasformate di Fourier si può dire che se il rumore è bianco, cioè ha uno spettro di questo tipo:

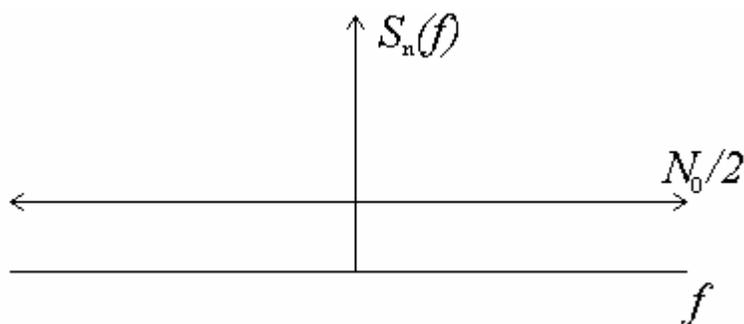


Figura 3.2 – Densità di potenza spettrale di un rumore bianco in funzione della frequenza.

dove N_0 è una costante.

In questo caso $R_{xx}(\tau)$ sarà una funzione proporzionale a una $\delta(\tau)$, che equivale a dire che le variabili $x(t_1)$ e $x(t_2)$ sono totalmente scorrelate fra loro essendo l'autocorrelazione diversa da zero solo per $\tau=0$. Questo è quanto avviene nel caso di un generico rumore bianco.

Il primo rumore simulato, scelto per essere addizionato al segnale, è un rumore gaussiano, bianco e a media zero. La scelta di questo tipo di rumore deriva dal fatto che esso è presente in molti casi di reale applicazione ed è quindi abbastanza verosimile, fermo restando che spesso volte, per ottenere un rumore bianco è necessario filtrare preventivamente un segnale prima del filtraggio vero e proprio.

Poiché si lavora con dati campionati, e quindi discreti, è chiaro che il rumore generato via software, oltre a essere una sequenza di numeri pseudo-random sarà bianco nell'intervallo che va da frequenza zero fino alla frequenza caratteristica che è pari alla metà della frequenza di campionamento. Tuttavia anche in questo intervallo il rumore è bianco solo in approssimazione, essendo comunque generato per via software.

Si ricorda infine che la potenza totale, dovuta alla somma della densità di potenza spettrale su tutte le frequenze, nel caso di un rumore gaussiano bianco a media zero è pari alla sua varianza.

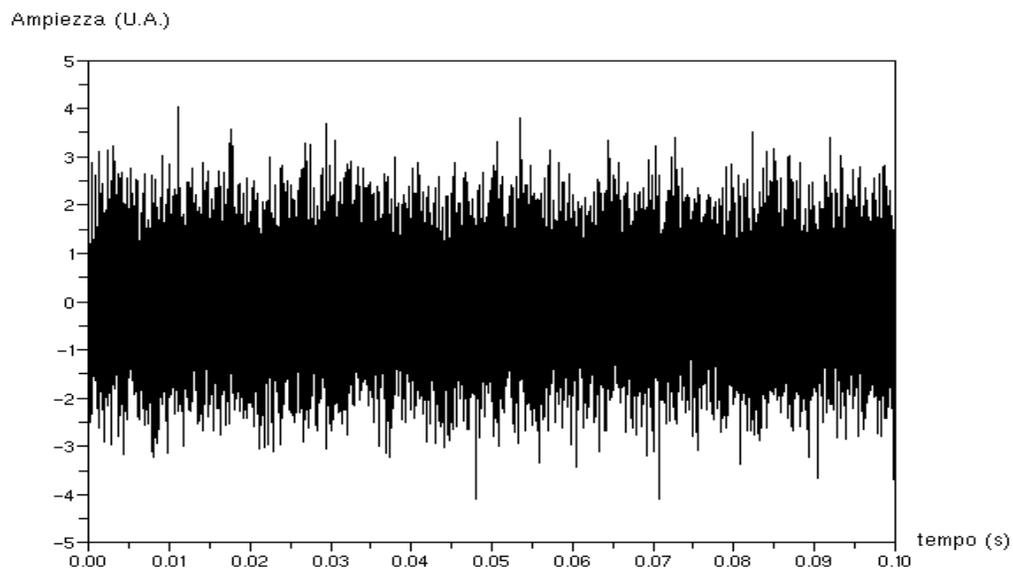


Figura 3.3 – Particolare realizzazione del rumore bianco generato via software.

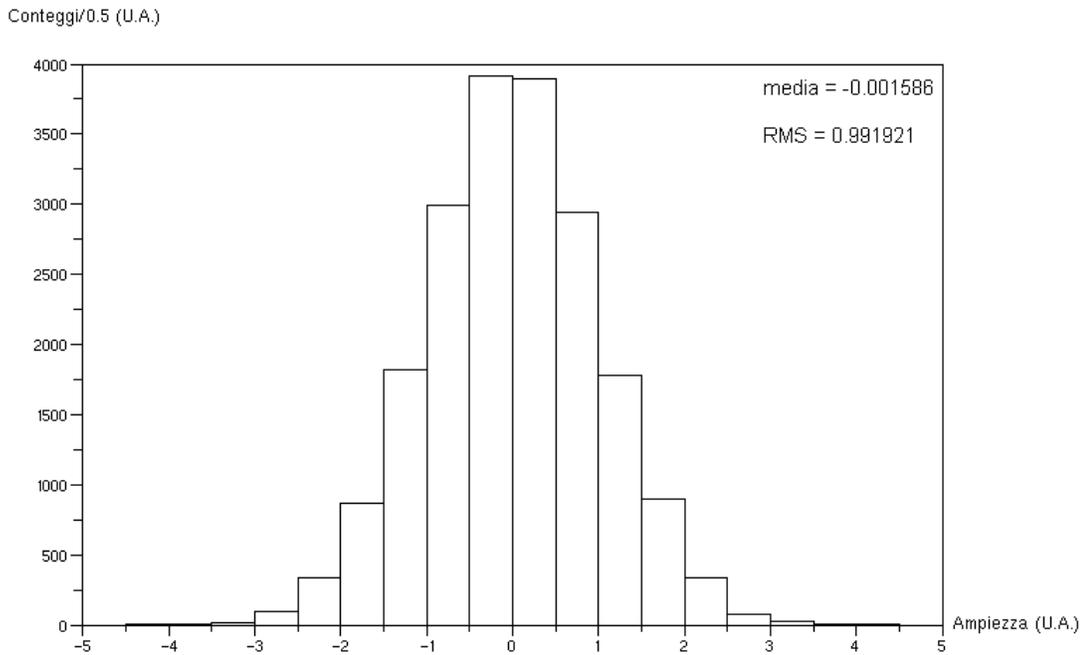


Figura 3.4 – Distribuzione delle ampiezze del rumore bianco generato via software in figura 3.3.

Nella realtà le componenti del rumore avranno due origini diverse, una dovuta al mare che comprende fattori come il passaggio di imbarcazioni, la pioggia o il vento in superficie e l'altra dovuta al rivelatore e alla sua elettronica di lettura.

Per simulare invece un rumore più realistico, più simile a quello che ci si aspetterebbe provenire dal mare, si è scelto un rumore gaussiano, colorato, a media zero. Il termine colorato sta ad indicare il fatto che in generale tutte le frequenze contribuiscono in potenza in modo differente. In particolare è stato generato un rumore la cui densità di potenza spettrale decresce come $1/f$ (figura 3.5), dove f indica la frequenza. In questo caso si avrà inoltre che l'autocorrelazione non sarà più una delta di Dirac.

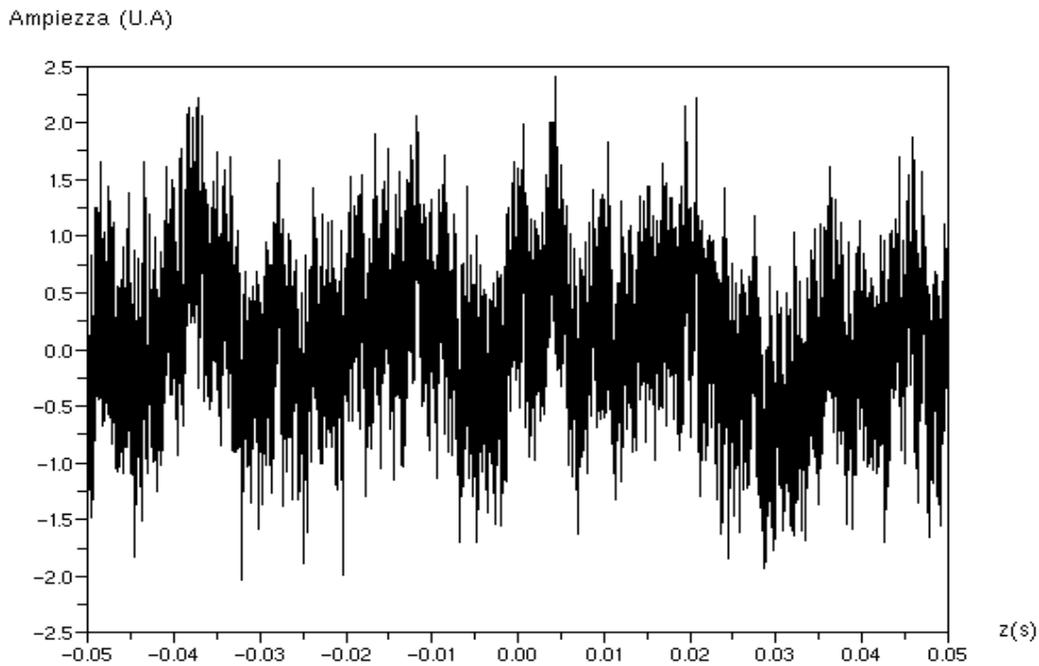


Figura 3.5 – Particolare realizzazione di un rumore gaussiano con potenza spettrale decrescente come $1/f$.

La scelta di un rumore di questo tipo è stata indotta inoltre dall'analisi spettrale di un tracciato disponibile proveniente dall'esperimento NEMO, che opera nel mar Mediterraneo.

Esso consiste in una registrazione di 15 s del rumore marino grazie a un idrofono situato a 1000 m di profondità, la frequenza di campionamento è 96 kHz (figura 3.6). La potenza spettrale di questo rumore non decresce esattamente come $1/f$, o comunque lo fa solo in certi intervalli di frequenze, tuttavia questo spettro è il risultato di una “convoluzione” tra il rumore effettivo del mare e la risposta degli idrofoni, che è ignota.

In effetti si nota un picco evidente in prossimità dei 50 Hz che è dovuto all'alimentazione degli idrofoni.

Potenza spettrale (U.A.)

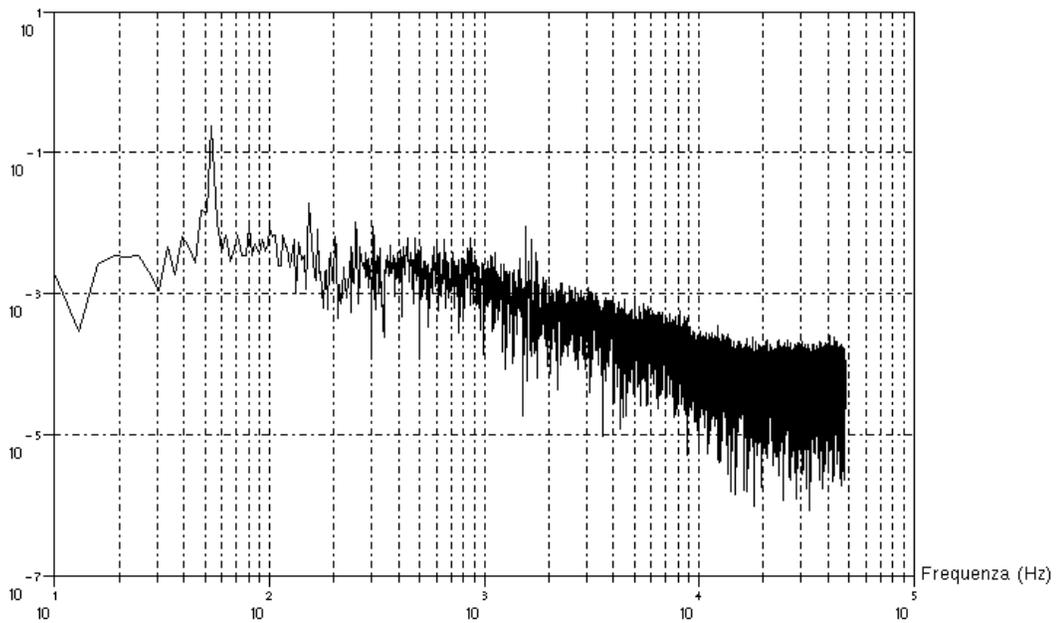


Figura 3.6 – Densità di potenza spettrale del rumore registrato da NEMO. La durata della registrazione è di 15 s e la frequenza di campionamento di 96 kHz.

4. Filtraggio del segnale: il filtro adattato

Il metodo di filtraggio sfruttato per estrarre un segnale del tipo descritto sopra è quello del filtro adattato. Esso si applica quando il segnale ha una forma nota $s(t)$ e non è pertanto necessario recuperarla attraverso il filtraggio. Infatti tale filtro si limita solo a dare un'informazione circa la presenza o meno di un segnale all'interno di una sequenza di dati.

Per semplicità si suppone di trattare segnali continui infinitamente estesi nel tempo, per cui, se il rumore additivo si può scrivere, nel caso un segnale $s(t)$ sia presente

$$r(t) = s(t) + n(t)$$

dove $r(t)$ rappresenta il segnale visto dal ricevitore ed è appunto la somma di $s(t)$ più una componente casuale $n(t)$ che costituisce il rumore e ha le caratteristiche viste in precedenza.

Nel caso opposto in cui invece non è presente nessun segnale si avrà al ricevitore

$$r(t) = n(t)$$

Eeguire una operazione di filtraggio significa sostanzialmente operare una convoluzione, data quindi una funzione $h(t)$, che si deve ricavare, si cerca una $g(t)$ data da

$$g(t) = \int_{-\infty}^{+\infty} r(\tau)h(t-\tau)d\tau$$

la funzione $g(t)$ dovrà in qualche modo dare delle informazioni sul segnale eventualmente presente.

Supponendo ora che ci si trovi in presenza di un segnale si avrà che $r(t)$ è la somma di due componenti e pertanto l'espressione di $g(t)$ potrà essere scritta come

$$g(t) = \int_{-\infty}^{+\infty} s(\tau)h(t-\tau)d\tau + \int_{-\infty}^{+\infty} n(\tau)h(t-\tau)d\tau = G_s(t) + G_n(t)$$

Si definisce ora il rapporto segnale rumore SNR (Signal to Noise Ratio) ad un istante T nella forma seguente:

$$SNR = \frac{G_s^2(T)}{E[G_n^2(T)]}$$

dove E rappresenta l'operatore valor medio. Esso intuitivamente rappresenta il rapporto tra la potenza relativa al segnale e quella dovuta al rumore. Appare infatti ovvio che l'SNR diventa tanto più piccolo quando lo diventa il segnale fino ad essere zero quando non è presente nessun segnale.

Il passo successivo è quello di scegliere $h(t)$ in maniera tale che, se è presente un segnale, questo rapporto venga massimizzato. Attraverso alcuni passaggi matematici [4] e sfruttando le proprietà del rumore e del prodotto convolutorio si giunge ad una forma dell'SNR più semplice da maneggiare

$$SNR(T) = \frac{\left[\int_{-\infty}^{+\infty} h(\tau)s(T-\tau)d\tau \right]^2}{\frac{N_0}{2} \int_{-\infty}^{+\infty} h^2(T-\tau)d\tau}$$

dove N_0 è una costante derivante dal fatto che si è supposto un rumore bianco di potenza totale

$$P_n = BN_0 = \sigma^2$$

dove B è l'ampiezza di banda del sistema di rivelazione.

Sfruttando la disuguaglianza di Schwartz si ottiene che l'SNR presenta un massimo per

$$h(\tau) \propto s(T - \tau)$$

che vale

$$SNR_{\max} = \frac{2}{N_0} \int_{-\infty}^{+\infty} (s(\tau))^2 d\tau = \frac{2E_s}{N_0}$$

dove E_s che rappresenta l'energia del segnale.

Un fatto molto importante è che il valore massimo dell'SNR non dipende dal particolare istante T scelto, infatti la variabile T non è più presente. Poiché lo stesso procedimento può essere applicato per un T qualsiasi arrivando allo stesso risultato, esso può essere scelto in modo arbitrario a seconda dell'istante in cui si vuole massimizzare l'SNR, senza modificare l'efficienza del filtro.

Il valore $g(T)$ di massimizzazione trovato può pertanto essere confrontato con una *soglia*, cioè si stabilisce un valore di riferimento per cui ogni volta che $g(T)$ supera questo valore si prende la decisione che è presente il segnale mentre se non lo è si decide che il segnale non è presente. La soglia verrà scelta, nel caso reale, a seconda dell'efficienza che si richiede da questo filtro e accettando un certo numero di “falsi positivi” per unità di tempo.

Va precisato infine che poichè si ha a che fare con processi statistici non si ha mai la certezza della presenza di un segnale o meno ma solo una certa probabilità, probabilità che naturalmente aumenta all'aumentare dell'energia del segnale relativamente al rumore.

Questo significa che a parità di segnale si può avere una risposta differente del filtro a seconda della particolare realizzazione del rumore (che tuttavia continua a rimanere gaussiano, bianco e a media zero); tale risposta fluttuerà quindi statisticamente.

Questo fatto può essere ricavato dalla espressione di $g(t)$, in cui si sostituisce l'espressione trovata di $h(t)$:

$$g(T) = \int_{-\infty}^{+\infty} s(\tau)s(\tau)d\tau + \int_{-\infty}^{+\infty} n(\tau)s(\tau)d\tau$$

Quindi vi sarà una parte dipendente dal segnale, che in questo caso è massima, mentre una parte che fluttuerà statisticamente in modo indipendente. Se ora si studia cosa succede in generale alla variabile $g(t)$ per un istante qualsiasi $t = T \pm \varepsilon$ si scopre che

$$g(T \pm \varepsilon) = \int_{-\infty}^{+\infty} s(\tau)s(\tau \mp \varepsilon)d\tau + \int_{-\infty}^{+\infty} n(\tau)s(\tau \mp \varepsilon)d\tau$$

dove in generale

$$\int_{-\infty}^{+\infty} s(\tau)s(\tau \mp \varepsilon)d\tau \leq \int_{-\infty}^{+\infty} s(\tau)s(\tau)d\tau$$

mentre la parte stocastica, nell'ipotesi di rumore stazionario, non cambia. In questo modo si nota come $g(t)$ presenti un massimo locale a $t=T$ permettendo l'identificazione dell'istante in cui il segnale è presente.

Nel caso in cui il rumore sia gaussiano ma colorato, si può ancora applicare un filtro *ottimale* (che diventa *adattato* nel caso di AGWN), la sua derivazione matematica è abbastanza laboriosa e viene pertanto riportata in maniera semplificata.

In questo caso l'espressione dell'SNR non può essere semplificata ed è della forma:

$$SNR(T) = \frac{\left[\int_a^b h(T-\tau)s(\tau)d\tau \right]^2}{\iint h(T-\tau)h(T-\vartheta)R_n(\tau-\vartheta)d\tau d\vartheta}$$

dove R_n è l'autocorrelazione del rumore. I calcoli vengono fatti inizialmente integrando tra due estremi finiti e successivamente si estende il tutto all'infinito.

Grazie alla teoria delle variazioni si arriva a

$$\int_{-\infty}^{+\infty} R_n(\tau-\vartheta)h(\vartheta)d\vartheta = s(T-\tau), \quad -\infty \leq u \leq +\infty$$

Applicando la trasformata di Fourier a entrambi i membri si può infine ricavare la funzione di trasferimento del filtro ottimale:

$$H(i\omega) = [S(-i\omega) / S_n(f)] \exp(-j\omega t)$$

dove $S_n(f)$ è la potenza spettrale del rumore colorato. Si può mostrare come nel caso di rumore bianco si ritorni al caso del filtro adattato precedentemente ricavato.

Quando si andrà ad utilizzare effettivamente un filtro di questo tipo si lavorerà con dati discreti e pertanto invece di un insieme continuo di variabili si avrà un vettore di N elementi ognuno corrispondente a un campionamento, e altrettanto sarà il filtro. Quindi si avrà a che fare con un numero finito di dati campionati ad una frequenza data.

La procedura utilizzata si è basata semplicemente sulla discretizzazione di quanto ricavato fino ad ora nel caso continuo. Nel caso discreto è possibile tuttavia procedere anche in un modo alternativo, col medesimo risultato [4]. Infatti è sufficiente considerare $g(t)$, $r(t)$, $s(t)$ come dei vettori, \mathbf{g} , \mathbf{r} ed \mathbf{s} , composti da N elementi che corrispondono agli N campionamenti e definire una matrice \mathbf{H} $N \times N$ che effettua l'operazione di filtraggio

$$\mathbf{g} = \mathbf{H}\mathbf{r}$$

Scegliere un istante particolare significa semplicemente moltiplicare ambo i membri per un vettore \mathbf{e}'_i composto da tutti zeri e un 1 nell' i -esima posizione corrispondente all'istante di interesse (per tutti i vettori l'apice indica che si tratta di un vettore riga, la sua assenza indica un vettore colonna).

Si definisce ora

$$\mathbf{h} = \mathbf{H}'\mathbf{e}_i$$

L'espressione dell'SNR ad un certo istante è data da

$$SNR(i) = \frac{|\mathbf{h}'\mathbf{s}|^2}{\mathbf{h}'\mathbf{C}_n\mathbf{h}}$$

dove $\mathbf{C}_n = E(\mathbf{n}'\mathbf{n})$ rappresenta la matrice di covarianza del rumore.

Il massimo dell'SNR, che vale

$$SNR_{\max} = \mathbf{s}'\mathbf{C}_n^{-1}\mathbf{s}$$

si ha per

$$\mathbf{h} = \mathbf{C}_n^{-1}\mathbf{s}$$

Infine si può dimostrare come l'SNR, nel caso di rumore bianco, all'istante i -esimo sia dato dalla seguente espressione

$$SNR(i) = |\mathbf{h}'\mathbf{s}|^2 / \sigma_n^2 \mathbf{h}'\mathbf{h}$$

che ha il massimo se $\mathbf{h}=\mathbf{s}$.

Il valore del vettore \mathbf{g} all'istante di massimizzazione sarà semplicemente dato dal prodotto scalare

$$g = \mathbf{h}'\mathbf{r}$$

la variabile g calcolata qui molto più semplicemente andrà quindi confrontata col valore di soglia.

5. Implementazione del filtro adattato

Il software utilizzato per implementare il filtro adattato è Scilab 3.1.1 [5]. Esso è stato scelto in quanto provvisto di una sezione specifica dedicata all'analisi di segnali per cui sono state disponibili molte funzioni matematiche di libreria atte a questo scopo, un esempio possono essere il calcolo di trasformate di Fourier, di convoluzioni, di correlazioni, oppure la generazione di rumori gaussiani.

Scilab permette inoltre di trattare molto semplicemente con i vettori e le matrici e inoltre è possibile inserire nei programmi delle routine scritte in Fortran o in C. In allegato a questa relazione è fornito il listato del programma che esegue il filtraggio del segnale.

Tralasciando la parte iniziale, che riguarda solamente dei settaggi di Scilab, la procedura seguita nel programma può essere schematizzata come segue:

1. si definisce il numero di punti che si intende campionare, la frequenza di campionamento, Δz ;
2. si genera la scala dei tempi nella variabile z . Questo intervallo è, per semplicità, simmetrico rispetto a $z = 0$;
3. si genera, a partire dai valori di z , un segnale della stessa forma di quello mostrato in figura 2.3 e lo si normalizza in modo da avere potenza media 1; sono inoltre inserite due righe di codice che permettono di selezionare la potenza media del segnale che è proporzionale alla sua energia;

4. dal segnale si genera il filtro e poichè si è deciso di massimizzare all'istante $z = 0$, che corrisponde alla posizione del flesso del segnale, il filtro sarà semplicemente dato dall'inversione temporale del segnale stesso;
5. si genera il rumore gaussiano bianco a varianza 1 e media 0 e lo si addiziona al segnale precedentemente generato;
6. si esegue il filtraggio tramite convoluzione, operazione che tuttavia provoca, nel caso di dati di lunghezza finita, una espansione dei dati nel tempo che è simmetrica nel passato e nel futuro. Questo è dovuto a una caratteristica dell'algoritmo di convoluzione, i dati in eccesso sono in ogni caso eliminati e non vengono visualizzati [6];
7. visualizzazione del segnale, del segnale addizionato al rumore, e questo ultimo dopo l'operazione di filtraggio.

6. Analisi dei risultati

6.1. L'uscita del filtro

Si riportano ora alcuni risultati ottenuti eseguendo il programma in allegato, i dati si riferiscono a una simulazione di campionamento a 200 kHz per 20000 punti, quindi con una finestra totale di 0.1 s, per un segnale con un Δz di 10^{-4} s, con un rapporto fra l'energia del segnale e quella del rumore pari a $3 \cdot 10^{-3}$. L'energia del segnale è stata calcolata applicando la formula:

$$E_s = \sum_{i=1}^N s(i)^2 \Delta t$$

dove Δt è l'intervallo di campionamento, mentre per quella del rumore si è moltiplicata la potenza totale, che come già detto per un rumore gaussiano bianco è pari alla varianza, per l'intervallo temporale totale dei dati che è semplicemente il numero di campionamenti diviso la frequenza di campionamento.

Il filtro è definito a meno di una costante e quello utilizzato è stato scalato in modo che la scala su cui varia $g(t)$ sia la medesima di quella di $s(t)$ e $n(t)$.

Il segnale generato, in figura 6.1, viene addizionato al rumore (figura 6.2); quindi si applica il filtro tramite convoluzione, l'uscita del filtro è mostrata in figura 6.3.

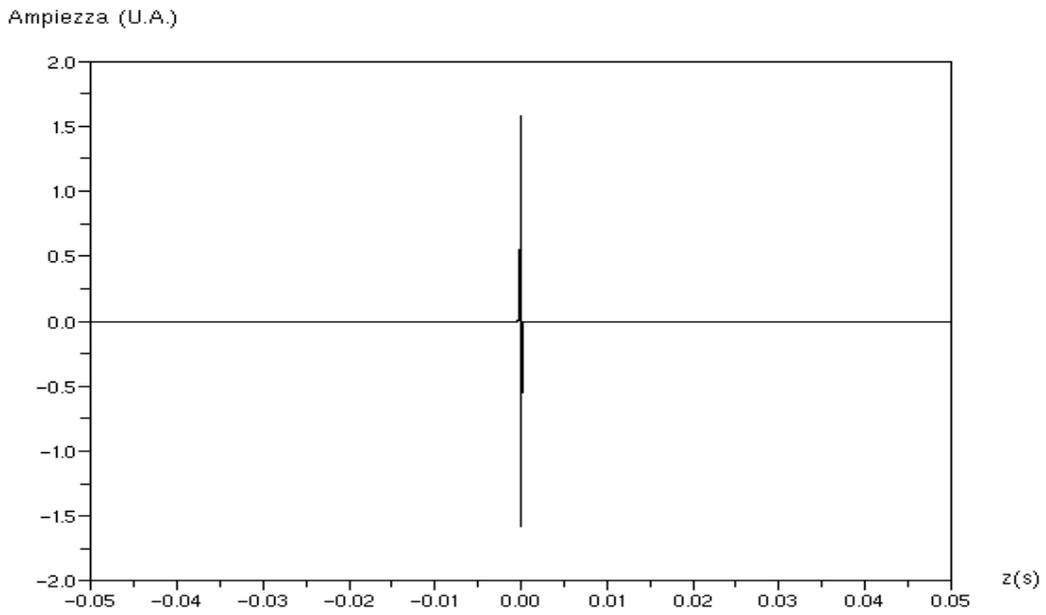


Figura 6.1 – Segnale utilizzato nella simulazione del programma in allegato.

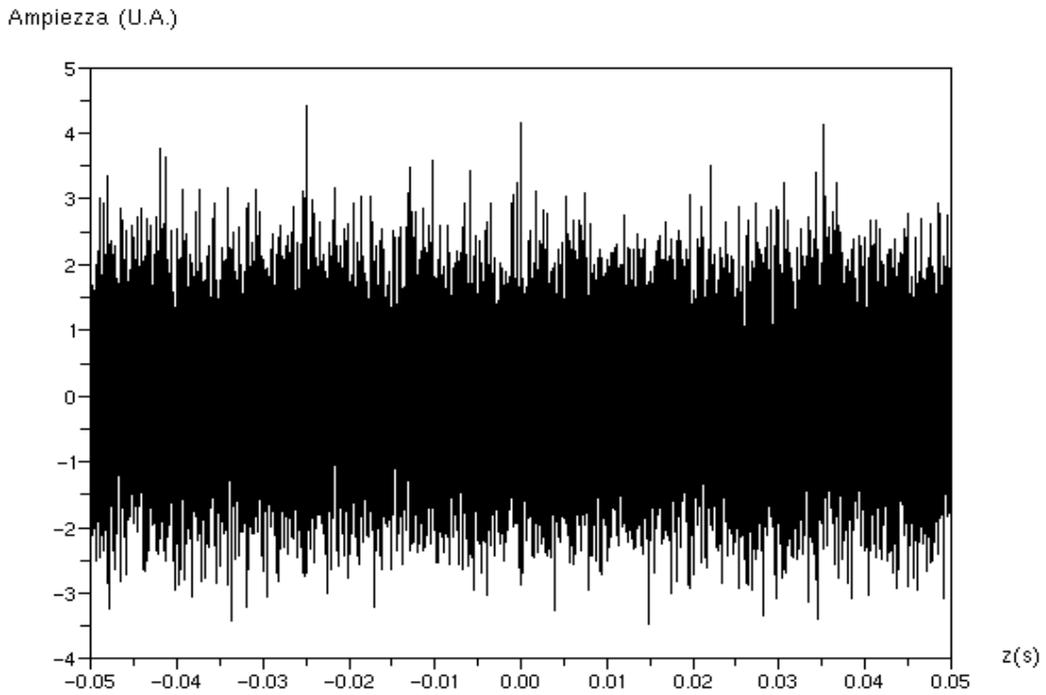


Figura 6.2 – Segnale di figura 6.1 addizionato a un rumore gaussiano bianco.

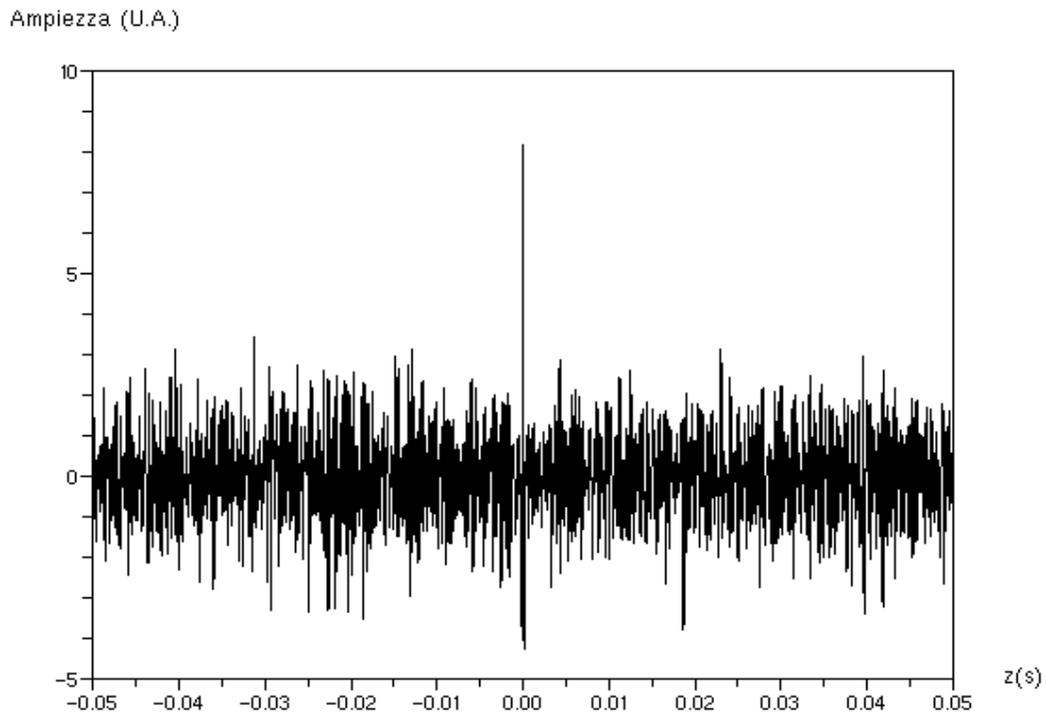


Figura 6.3 – Filtraggio tramite filtro adattato del segnale in figura 6.2.

Si nota la presenza di un picco all'istante in cui si è scelto di massimizzare l'SNR, $z=0$. Bisogna aggiungere che con questo stesso filtro (adattato a questo segnale) si è in grado di eseguire la stessa operazione per un segnale uguale in forma ma traslato temporalmente di un certo t' ed inoltre la massimizzazione del nuovo segnale avviene nel punto omologo a quello del segnale su cui si è adattato il filtro. Visivamente questo significa che il picco nella figura precedente risulterà spostato esattamente della quantità t' . Questo effetto è stato verificato e può essere dimostrato con le proprietà del prodotto convolutorio.

6.2. Analisi statistica: rumore gaussiano bianco

L'altezza del picco fluttua statisticamente per ogni differente realizzazione del rumore che equivale a dire ogni volta che si esegue il programma.

Questo è dimostrato dal grafico in figura 6.4 dove è riportata la distribuzione dei valori del picco (g_{max}) per diverse realizzazioni del rumore (distribuzione ricavata filtrando lo stesso segnale per 1000 realizzazioni differenti del rumore).

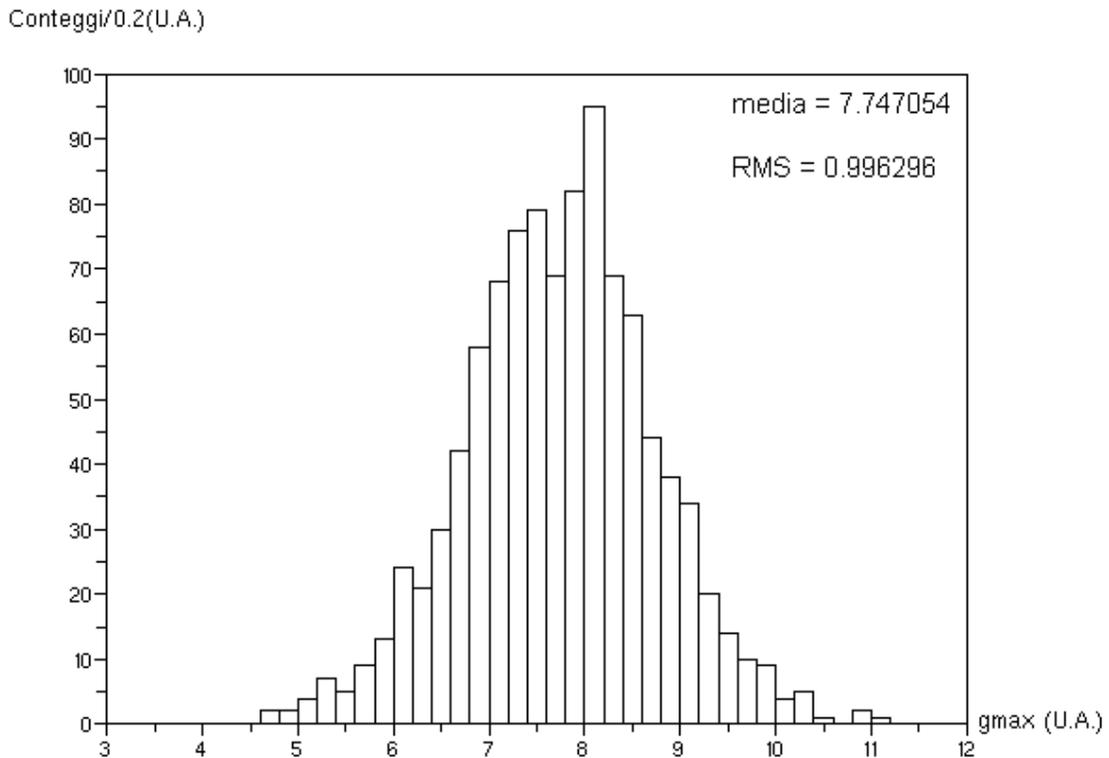


Figura 6.4 – Distribuzione di g_{max} ricavata su 1000 realizzazioni del rumore

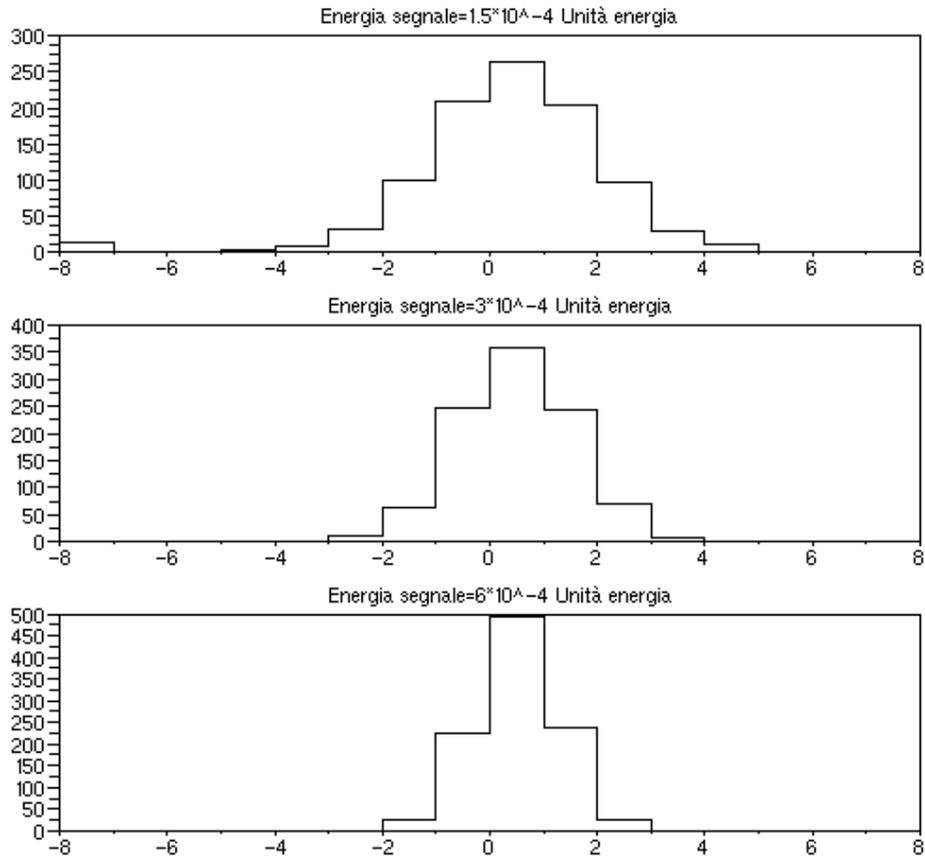


Figura 6.5 – Distribuzione statistica dell’istante temporale di rivelazione rispetto a quello atteso ($z=0$). In ascissa si riporta z e in ordinata i conteggi. L’energia del rumore è 0.1 U.E.

Nel caso reale, in cui non si conosce la posizione del segnale, essa viene dedotta dalla posizione del picco, tuttavia anche essa, a parità di segnale, varia statisticamente a seconda della particolare realizzazione del rumore, pertanto non si riesce ad individuare con esattezza l’istante temporale in cui si trova il segnale. Il tempo fornito dalla operazione di filtraggio fluttuerà quindi statisticamente attorno al valore esatto (che in questo caso è $z = 0$), come mostrato in figura 6.5. Come prevedibile, le distribuzioni si allargano al diminuire dell’energia del segnale. In ogni caso da ora in avanti il segnale verrà considerato come rivelato esattamente se il picco identificato giace in un intervallo compreso tra ± 5 campionamenti attorno all’istante esatto; questo equivale a commettere un errore massimo di $25 \mu s$.

Si è utilizzato inoltre questo stesso filtro per filtrare segnali della stessa forma e di uguale energia, ma la cui distanza temporale fra i punti di estremo è differente dal segnale di partenza. Quindi si vuole studiare la risposta del filtro a segnali “disadattati”.

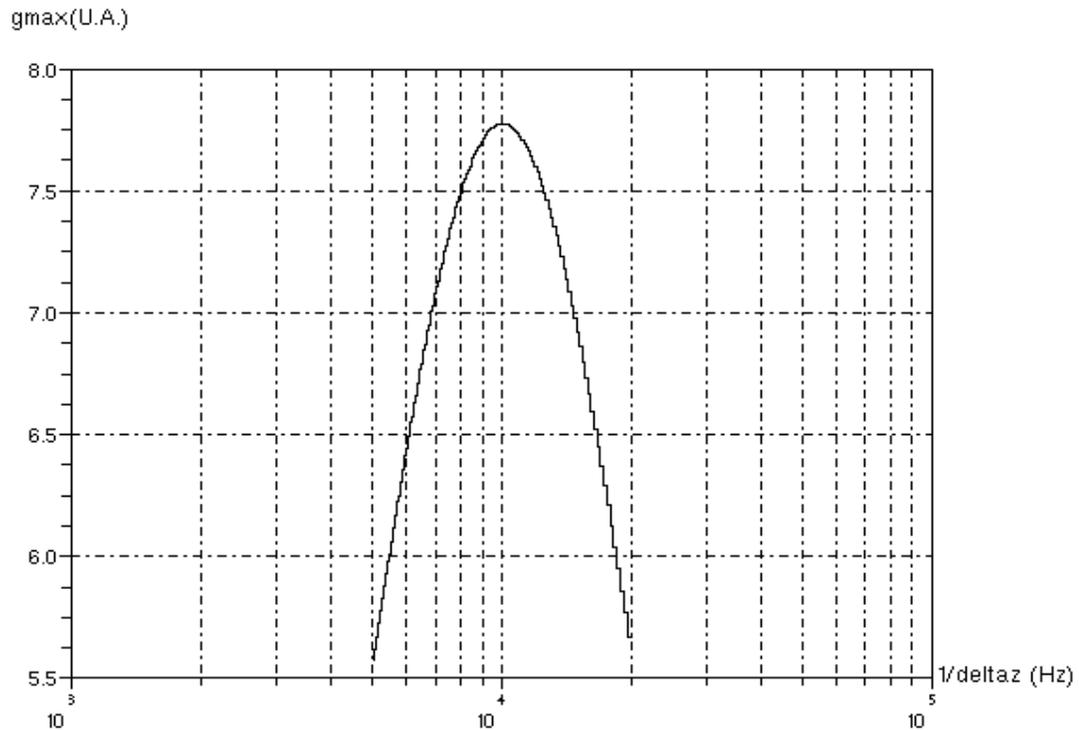


Figura 6.6 – Andamento di g_{\max} per segnali disadattati. Ogni punto (per 100 punti) è stato calcolato effettuando una media su 100 realizzazioni del rumore.

Si nota nel grafico in figura 6.6 come effettivamente la risposta del filtro abbia un massimo per $1/\Delta z = 10^4$ Hz, che corrisponde al segnale “adattato”, mentre l’efficienza diminuisce allontanandosi da questo valore. Realizzando un numero limitato di filtri centrati su valori discreti di Δz è possibile coprire quindi tutto l’intervallo d’interesse efficientemente.

Lo studio delle caratteristiche del filtro prosegue con l’identificazione di un valore di soglia, come definito prima, ragionevole. Questo valore dovrà essere non troppo basso per evitare un numero elevato di “falsi positivi” e non troppo alto per avere una efficienza sufficiente dell’algoritmo anche per segnali di bassa energia.

Quello che si è fatto è sostanzialmente effettuare l'operazione di filtraggio per il solo rumore e visualizzare (figura 6.7), in funzione della soglia scelta, il numero di falsi positivi. Questa operazione va mediata su un numero sufficiente di realizzazioni del rumore.

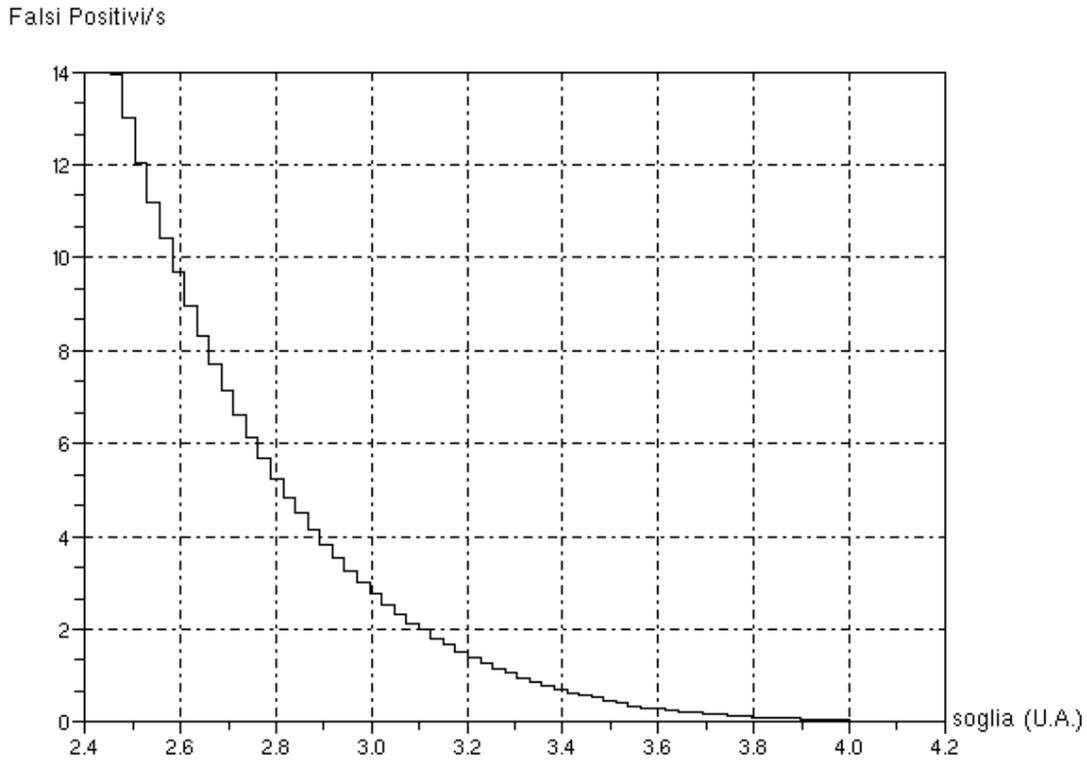


Figura 6.7 – Falsi positivi al secondo in funzione della soglia. Ogni punto è stato calcolato effettuando una media su 100 realizzazioni del rumore.

A questo punto si è scelta la soglia in modo tale che, mediamente, il numero di falsi positivi sia di 1 al secondo, cioè $soglia \approx 3.3 \text{ U.A.}$.

Riferendosi a questa soglia è possibile definire una efficienza del filtro:

$$\varepsilon(E) = \frac{\text{eventi} \left(\left| t(g_{\max}) - t_{\text{segnale}} \right| \leq 25 \mu\text{s}, g_{\max} > soglia = 3.3 \right)}{\text{n}^\circ \text{ realizzazioni del rumore}}$$

cioè si genera un certo numero di volte un rumore e si conta il numero di eventi che superano la soglia e che sono compatibili con l'istante in cui è presente il segnale, si divide quindi per il numero di rumori generati. Tutto questo in funzione del rapporto dell'energia del segnale su quella del rumore, E_s/E_n , come mostrato in figura 6.8.

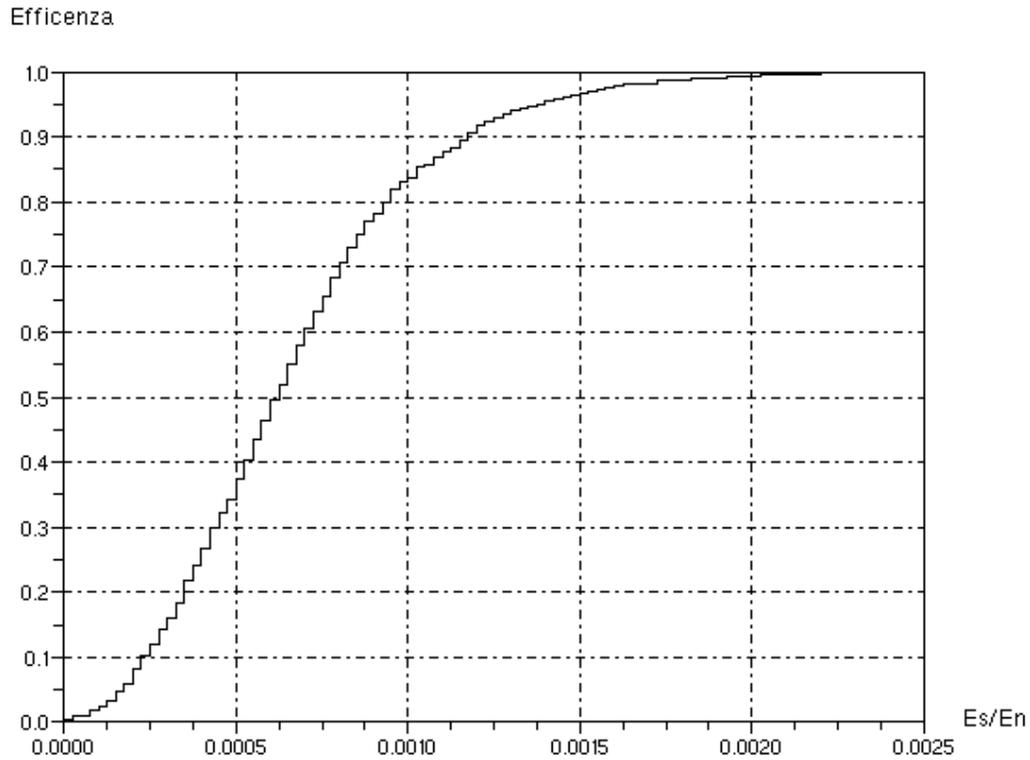


Figura 6.8 – Efficienza del filtro relative a una soglia di 3.3 U.A.. Ogni punto è stato calcolato effettuando una media su 1000 realizzazioni del rumore.

Come ci si attendeva l'efficienza aumenta all'aumentare dell'energia del segnale fino a tendere a 1.

6.3. Analisi statistica: rumore gaussiano colorato

Lo stesso tipo di ragionamento è stato fatto nel caso del rumore colorato, tuttavia la normalizzazione di questo rumore, necessaria per rendere i dati ricavati da questo confrontabili con quelli ottenuti nel caso di rumore bianco, non è stata fatta semplicemente rendendo uguale la potenza totale.

Infatti si è ritenuto opportuno effettuare la normalizzazione rendendo uguale la potenza dei due rumori per la frequenza caratteristica del segnale del segnale che si deve filtrare, che è risulta essere di poco inferiore a $1/(2\Delta z)$. Con una normalizzazione di questo tipo non si ha il problema di dover analizzare i dati per segnali di differente frequenza caratteristica ma ne basta uno, come nel caso di rumore bianco. Nel caso che si è andato a esaminare, per un segnale con $\Delta z = 10^{-4}$ s e una frequenza caratteristica di $\sim 3.2 \cdot 10^3$ Hz, la normalizzazione si traduceva in un rapporto dell'energia del rumore gaussiano colorato su l'energia di quello bianco di circa 0.36.

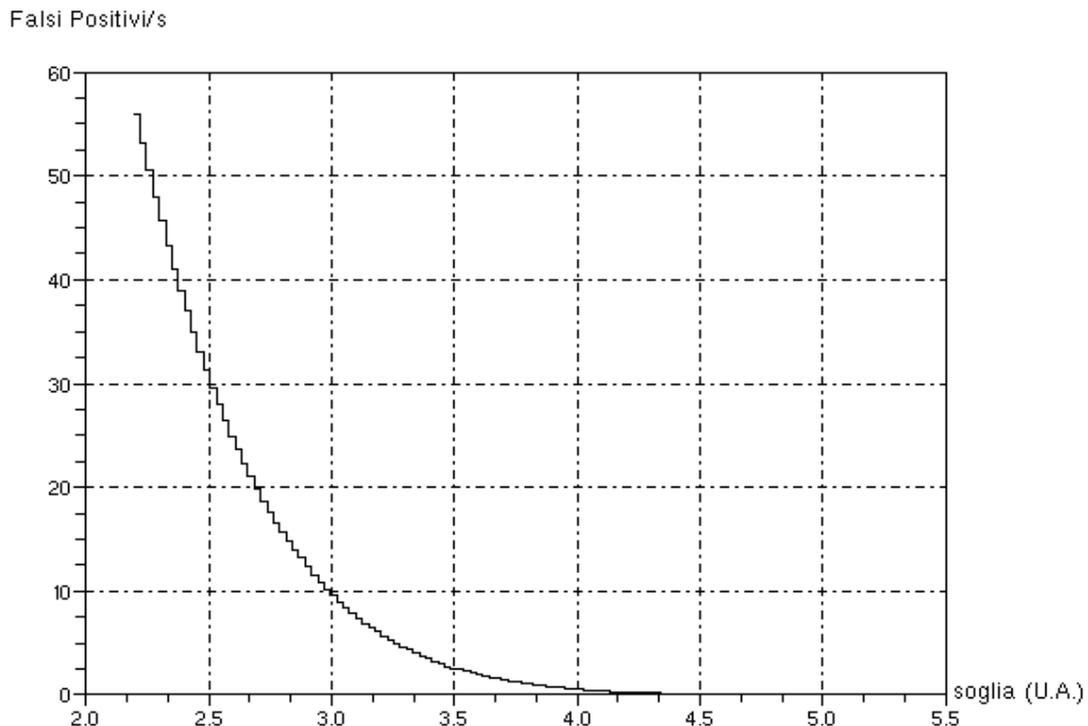


Figura 6.9 - Falsi positivi al secondo in funzione della soglia nel caso di un rumore $1/f$. Ogni punto è stato calcolato effettuando una media su 100 realizzazioni del rumore.

In questo caso, come ci si poteva facilmente attendere, per avere lo stessa frequenza di falsi positivi è necessario avere una soglia più alta (figura 6.9). Inoltre, come si può notare dalla figura sotto, la scala è differente rispetto a quella di rumore gaussiano, quindi al diminuire della soglia il numero di falsi positivi aumenta molto più rapidamente. A una frequenza di un falso positivo al secondo corrisponde qui una soglia di 3.8 U.A. Su questo valore di soglia si calcola quindi l'efficienza come definita sopra e si va a operare con quella relativa al rumore gaussiano (figura 6.10).

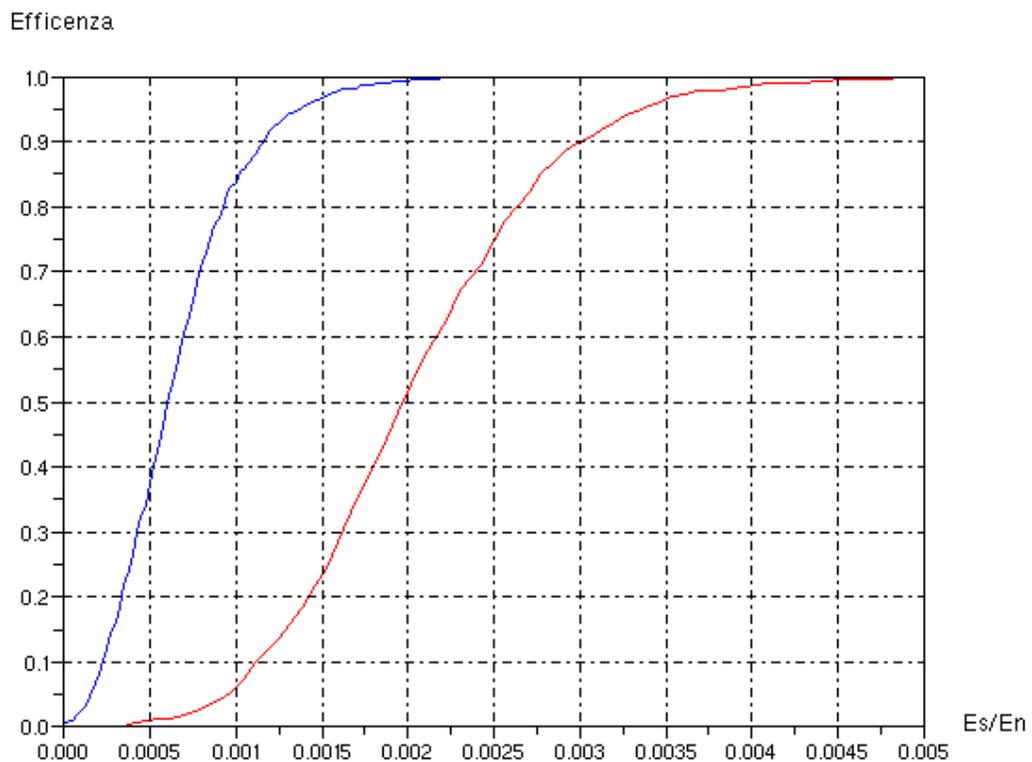


Figura 6.10 – Efficienza del filtro nei due casi di rumore gaussiano bianco (in blu) e di rumore gaussiano $1/f$ (in rosso). Le due curve sono relative a una soglia corrispondente a un falso positivo al secondo; i valori di soglia sono 3.3 U.A. per il rumore bianco e 3.8 U.A. per il rumore colorato.

L'efficienza nel caso di rumore colorato, per lo stesso numero di falsi positivi al secondo, cresce molto più lentamente al crescere dell'energia di quella relativa al rumore bianco. Se per esempio si decidesse di lavorare con una efficienza del 90% si scopre che i segnali che possono essere rivelati da un rumore bianco hanno una energia 3 volte

superiore che nel caso di rumore colorato. Questo è comunque il comportamento atteso da un filtro adattato a un rumore gaussiano bianco.

6.4. Analisi statistica: rumore da NEMO

Le stesse considerazioni sono state infine applicate al rumore di NEMO, solo che in questo caso sono state disponibili poche realizzazioni del rumore, cioè quelle estraibili da un tracciato di 15 s. In questo caso è inoltre superfluo effettuare una normalizzazione rispetto ai rumori generati via software, in effetti questo rumore è dato dalla combinazione di due fattori, il rumore del mare, quello dell'elettronica di controllo e dipende inoltre dalla risposta in frequenza di quest'ultima.

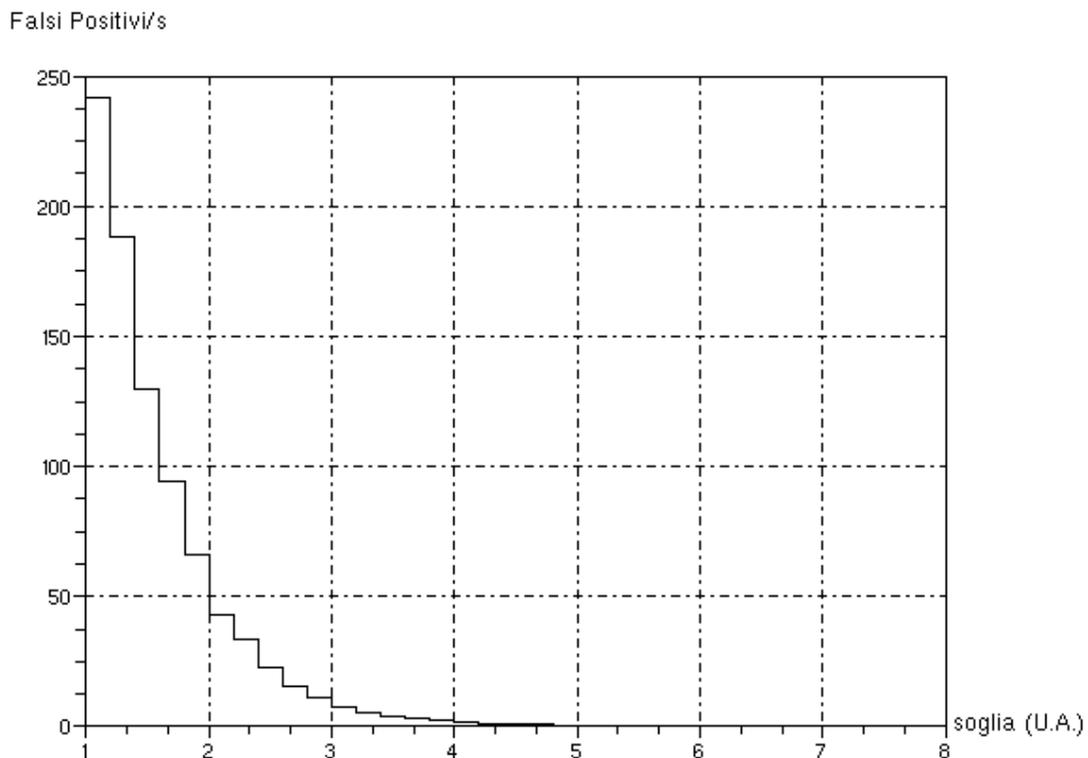


Figura 6.11 - Falsi positivi al secondo in funzione della soglia. Ogni punto è stato calcolato effettuando una media su 65 realizzazioni del rumore.

Per essere elaborato, il tracciato è stato suddiviso in 65 parti, ad ognuna delle quali è stato applicato il filtro. Successivamente si è operata una media su queste 65 realizzazioni del rumore in modo da poter fare delle considerazioni sui falsi positivi (figura 6.11). In questo caso una frequenza di un falso positivo al secondo si ha per una soglia di ~ 4.5 U.A.

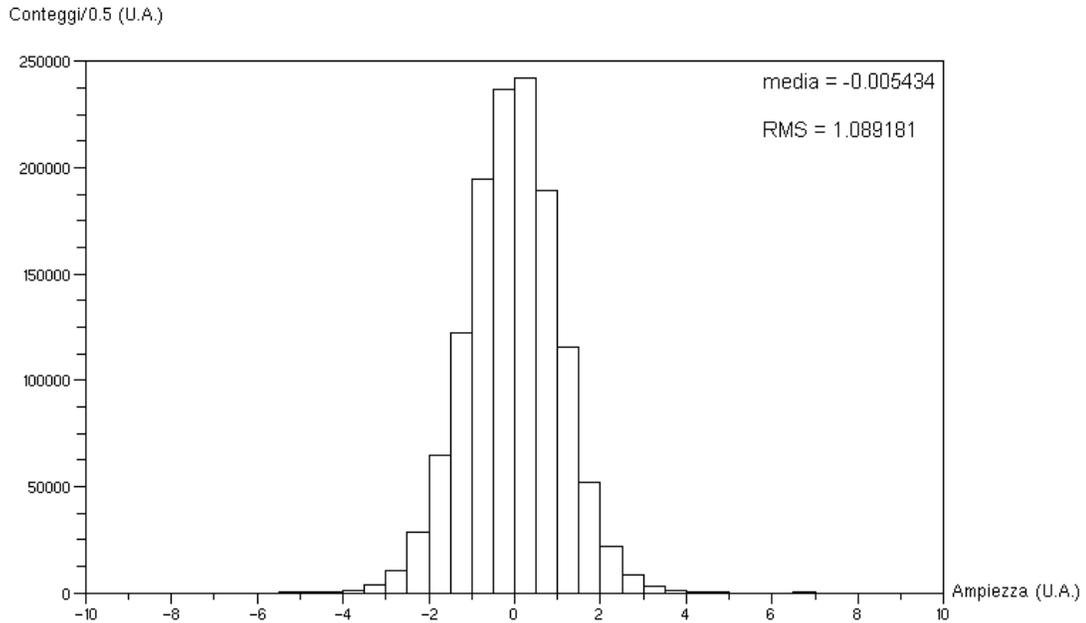


Figura 6.12 – Distribuzione delle ampezze del rumore di NEMO dopo il filtraggio.

Risulta interessante studiare la distribuzione delle ampezze dopo l'operazione di filtraggio. Infatti il numero di falsi positivi al secondo, in funzione della soglia, è proporzionale all'integrale della funzione di distribuzione dal valore di soglia a $+\infty$. Questo fatto può essere visto intuitivamente dal fatto che diminuendo la soglia si accettano ampezze minori nella distribuzione.

Si definisce ora la quantità χ , relativa ad ogni tipologia di rumore, come

$$\chi = \frac{S}{\sigma}$$

dove S rappresenta il valore di soglia corrispondente ad un falso positivo al secondo mentre σ è la deviazione standard del rumore. Questa quantità può essere vista come un indice dell'efficienza del filtro, infatti da un informazione sul numero di deviazioni standard dal quale ci si deve allontanare dal centro della distribuzione per avere un dato numero di falsi positivi. Ovviamente il filtro funziona meglio quanto più questo rapporto è piccolo. Se si calcola χ per il rumore bianco, colorato e di NEMO si ha rispettivamente

$$\chi_B = \frac{S_B}{\sigma_B} \approx \frac{3.3}{1} = 3.3$$

$$\chi_{1/f} = \frac{S_{1/f}}{\sigma_{1/f}} \approx \frac{3.8}{0.6} \approx 6.3$$

$$\chi_N = \frac{S_N}{\sigma_N} \approx \frac{4.5}{1} = 4.5$$

Come ci si aspettava il valore più basso lo si ottiene nel caso del rumore gaussiano bianco che è quello per cui è stato adattato il filtro.

7. Conclusioni

Il lavoro svolto in questa tesi è incentrato sullo studio di un filtro software per l'estrazione di un segnale acustico noto, dovuto all'interazione di neutrini di altissima energia in mare, dal rumore.

Il tipo di filtro scelto è un filtro adattato ed è stato testato su due rumori gaussiani simulati, uno bianco e uno colorato. Questi due rumori, in particolare il secondo, vogliono essere una riproduzione realistica del rumore marino.

Lo studio delle caratteristiche del filtro si è focalizzato principalmente sull'analisi statistica della sua risposta, come il numero di falsi positivi e l'efficienza, e sul confronto fra le diverse tipologie di rumore. Nella parte finale si è fatta un'analisi preliminare dei primi segnali acustici provenienti da NEMO.

Bibliografia

- [1] A.A. Watson, Phys. Rep. **333-334**, 309 (2000).
- [2] J.G. Learned, Phys. Rev. D **19**, 3293 (1979).
- [3] J. Vandenbroucke, G. Gratta, N. Lehtinen, Astroph. Journal, 621:301-312, (2005)
- [4] R.N. McDonough, A.D. Whalen, *Detection of Signals in Noise* (Academic Press, San Diego, 1995), cap. 6, 2° edizione.
- [5] Scilab 3.1.1, scilabsoft.inria.fr.
- [6] A. Falaschi, *Elementi di Trasmissione dei Segnali e Sistemi di Telecomunicazione*, edito presso <http://infocom.ing.uniroma1.it/alef/libro>.
- [7] N. Lehtinen, S. Adam, G. Gratta, T.K. Berger, M.J. Buckingham, *Astroparticle Physics* 17 (2002) 279-202.

Ringraziamenti

Ringrazio il Dott. Alessandro Cardini e il Prof. Biagio Saitta per la loro competenza, la loro disponibilità e per avermi seguito con interesse in questo lavoro.

Un grazie a tutti i colleghi, in particolare a Francesco, Riccardo, Paolo, Nicola e R.F. per avermi aiutato a raggiungere questo obiettivo con la loro amicizia, il loro supporto tecnico e i litri di birra, e a tutti gli amici che mi hanno sostenuto e aiutato.

Ringrazio infine i miei genitori senza i quali sarebbe stato impossibile il raggiungimento di questo traguardo.