

DOTTORATO DI RICERCA
in
SCIENZE COMPUTAZIONALI E INFORMATICHE
Ciclo XIX

Consorzio tra Università di Catania, Università di Napoli Federico II,
Seconda Università di Napoli, Università di Palermo, Università di Salerno

SEDE AMMINISTRATIVA: UNIVERSITÀ DI NAPOLI FEDERICO II

SILVIO PARDI

ALGORITMI DI TRIGGER PER LA RIVELAZIONE DI ONDE GRAVITAZIONALI
GENERATE DA SISTEMI BINARI COALESCENTI

TESI DI DOTTORATO DI RICERCA

IL COORDINATORE
Prof. Aldo De Luca

Indice

1	Teoria delle onde gravitazionali	8
1.1	Aspetti teorici	8
1.1.1	Le onde Gravitazionali	9
1.1.2	Effetto delle onde su particelle libere	10
1.1.3	Generazione di onde gravitazionali	11
1.1.4	Le sorgenti di onde gravitazionali	14
1.1.5	Binarie coalescenti	14
1.2	Esperimenti di Rivelazione di Onde Gravitazionali	15
1.2.1	Metodi interferometrici	16
1.3	L'esperimento VIRGO	17
1.3.1	L'analisi dati in Virgo	20
2	Tecniche di rivelazione	22
2.1	La teoria della rivelazione	22
2.2	Elementi del Test delle Ipotesi	23
2.2.1	Test delle ipotesi binario	24
2.2.2	Criterio di Bayes	24
2.2.3	Il criterio di Neyman-Pearson	26
2.2.4	Test delle Ipotesi Composto	30
2.3	Rivelazione di segnali in tempo discreto	31

2.3.1	Descrizione del Problema	32
2.3.2	Rivelazione dei segnali noti in rumore indipendente	33
2.3.3	Rivelazione di segnali noti in rumore Gaussiano Colorato	34
2.3.4	Rivelazione di segnali con parametri non noti	37
3	Modello per un ricevitore di Segnali Gravitazionali prodotti da Binarie Coalescenti	39
3.1	Segnale Gravitazionale atteso da Binarie Coalescenti	39
3.2	Modello per il ricevitore di segnali gravitazionali Chirp	43
4	Il problema dell'analisi on-line	47
4.1	L'analisi on-line	47
4.2	Griglie di template e costi computazionali	48
4.3	Calcolo del tempo di esecuzione	51
5	Dynamic Matched Filtering per la rivelazione di onde gravitazionali	55
5.1	Rivelazione di segnali Chirp	55
5.1.1	Dynamic Matched Filter Technique (DMFT)	56
5.2	Sperimentazione del metodo	58
5.2.1	La statistica di decisione - Modellizzazione	58
5.2.2	Generazione dei dati e determinazione della soglia - Pianificazione	59
5.2.3	Test e risultati - Esecuzione	64
5.2.4	Rivelazione Chirp per masse comprese tra $30 - 5 M_{\odot}$	64
5.2.5	Rivelazione Chirp per masse minori di 5	74
5.3	Costi computazionali	78
6	Filtro adattivo ALE	81
6.1	Introduzione	81
6.2	Modello del segnale	82
6.3	Infinite Impulse Response Adaptive Line Enhancer (IIR -ALE)	85

6.3.1	Adattamento dei parametri	88
6.3.2	Costi computazionali	90
6.3.3	Caratterizzazione della risposta del filtro	90
6.4	Adaptive correlator detector	92
6.4.1	Statistiche di rivelazione: caso con solo rumore	93
6.4.2	Distribuzioni statistiche di rivelazione: segnale più rumore	96
6.5	Risultati simulati	97
6.5.1	Simulazioni in rumore bianco	98
7	Conclusioni	103
A	Appendice - Ottimizzazione globale ed algoritmo di Price	105
A.1	Il problema dell'ottimizzazione globale	105
A.2	Classificazione dei metodi	106
A.2.1	Metodi Euristici	106
A.2.2	Metodi di Approssimazione	108
A.2.3	Metodi Sistematici	109
A.3	Algoritmo di Price	109
A.3.1	Descrizione dell'algoritmo di Price	110
A.3.2	test di performance	111
B	Calcolo parallelo e calcolo distribuito per l'analisi off-line	114
B.1	Un approccio parallelo	114
B.1.1	Algoritmo di Price parallelo	115
B.1.2	Sperimentazioni e conclusioni	116
B.2	Approccio distribuito	116
B.2.1	Il Grid Computing	119
B.2.2	Test preliminari su piattaforma INFN-GRID	120
C	Famiglia di modelli Delta-Gamma-Normal	122

Introduzione

Il presente lavoro di tesi si inserisce nel campo delle scienze computazionali, esso si propone di fornire due algoritmi di trigger per la rivelazione di onde gravitazionali provenienti da binarie coalescenti nell'ambito dell'esperimento VIRGO. Gli algoritmi introdotti ed risultati sperimentali sono oggetto di numerose pubblicazioni su riviste internazionali del settore.

Virgo è un esperimento nato da una collaborazione-Italo francese, con l'obbiettivo di rilevare il debole segnale gravitazionale atteso da sorgenti astrofisiche, sfruttando tecniche interferometriche. L'apparato sperimentale si trova in Cascina (Pisa) ed è attualmente in fase di test. Quando l'interferometro Virgo lavora a pieno regime, la mole di dati prodotta supera il terabyte al giorno, questo rende il problema del data analysis molto oneroso in termini di costo computazionale, per cui si cercano algoritmi per diminuire tale costo.

Per fronteggiare queste esigenze, vengono definiti almeno due livelli di analisi: (1) analisi on-line o pre-analisi ed (2) analisi off-line. Nella pre-analisi si richiede che gli algoritmi utilizzati siano leggeri e rapidi, tali da reggere il flusso di dati prodotti dall'interferometro, individuando le zone di dati che possono contenere il segnale gravitazionale.

Le tecniche classiche attualmente utilizzate, si basano sulla così detto Standard Matched Filter Technique (SMFT), che integra il filtro ottimo di Wiener-Komolgorov con un ricoprimento statico di template dello spazio dei parametri, la cui densità è definita sulla base del livello di confidenza per la rivelazione del segnale. Tale strategia risulta computazionalmente onerosa e più adatta per l'analisi off-line dove i tempi sono meno stringenti.

In questo lavoro di tesi viene proposto un approccio dinamico al problema della rivelazione on-line, mediante due differenti strategie. La prima prevede l'utilizzo di algoritmi di ottimizzazione globali, in grado di abbattere i costi computazionali eseguendo un campionamento adattivo nello spazio dei parametri. La seconda strategia proposta consiste nell'impiegare un filtro adattivo per ricostruire le componenti coerenti, qualora presenti, in tranches di rumore sbiancato.

Il primo approccio viene indicato con l'acronimo DMFT che sta per Dynamic Matched Filter Technique, esso viene definito in maniera astratta nel capitolo 5 di questa tesi, prescindendo quindi dalle scelte implementative e proponendosi come metodo alternativo al SMFT.

Successivamente viene fornita una implementazione del DMFT, utilizzando l'algoritmo di Price come strumento di ottimizzazione globale. Tale algoritmo, anche conosciuto come Controlled Random Search (CRS) è basato su una potente ed efficace revisione dell'algoritmo del semplice, ed ha riscontrato risultati positivi anche in altri campi di applicazione.

Il lavoro di tesi prosegue dunque con una fase di testing svolta in ambiente simulato. I risultati ottenuti validano sia l'approccio DMFT che le specifiche scelte implementative, evidenziando ottime prestazioni sia in termini di costo computazionale che di velocità di convergenza, ottenendo risultati interessanti utilizzando semplici work station.

L'approccio alternativo, altresì dinamico, è introdotto accuratamente nel capitolo 6 dove viene definito il filtro IIR ALE del secondo ordine, vincolato e con un solo parametro adattivo. Tale strategia ha la potente proprietà di riconoscere i segnali senza conoscerne la forma apriori, ma basandosi unicamente sull'ipotesi che il segnale sia localmente sinusoidale, condizione applicabile al caso dei segnali chirp prodotti da sistemi binari in coalescenza. Le performance sono state valutate su rumore bianco, considerando che l'uscita dell'interferometro può essere comunque sbiancata con filtri di whitening. I risultati ottenuti mostrano un adeguato comportamento da parte del filtro per il riconoscimento dei segnali, viene inoltre evidenziato un basso costo computazionale (cresce linearmente con il numero

di campioni da valutare) proponendosi efficacemente come metodologia per l'analisi online.

Le due metodologie proposte in questo lavoro di tesi, rappresentano quindi delle valide soluzioni per la risoluzione del problema scottante dell'analisi on-line per l'esperimento VIRGO che sta per entrare in presa dati. Sia l'approccio DMFT, che quello basato su filtri IIR ALE, consentono di abbattere i costi computazionali richiesti dalle tecniche standard, garantendo comunque ottime performance in termini di rivelazione.

Capitolo 1

Teoria delle onde gravitazionali

In questo capitolo daremo una breve introduzione alle onde Gravitazionali, descrivendo i principali aspetti teorici, le sorgenti e gli apparati sperimentali di rivelazione. Verrà quindi introdotto Virgo, un apparato sperimentale di rivelazione basato su tecniche interferometriche, nato in da una collaborazione Italo-Francese e nell'ambito del quale si inquadra totalmente il presente lavoro di tesi.

1.1 Aspetti teorici

La rivelazione di onde Gravitazionali rappresenta certamente uno dei più interessanti campi di ricerca della fisica moderna. Essa intende fornire una dimostrazione forte della teoria della relatività generale aprendo così un nuovo canale di informazione per studiare le dinamiche e le evoluzioni degli oggetti astrofisici. In questo scenario, i rivelatori interferometrici terrestri come VIRGO[48], LIGO[49], GEO[50] e TAMA[51], assumono un ruolo importantissimo. Infatti tali rivelatori operano su una larga banda di frequenze che vanno dalle decine di Hz ai $10KHz$ con altissimi livelli di sensibilità. A questi va aggiunto il progetto LISA[52] per la costruzione di un interferometro spaziale con l'obiettivo di esplorare le bande di frequenze dai 10^{-4} ad $10^{-1}Hz$.

Tutti questi progetti attualmente attivi sottolineano il grande interesse attualmente

dedicato in ambito scientifico internazionale al problema della rivelazione delle onde Gravitazionali. Questo lavoro di tesi si inserisce nell'ambito del data analysis per il progetto Virgo.

1.1.1 Le onde Gravitazionali

Secondo le equazioni di Einstein lo spazio-tempo è un ente metrico dipendente dalla distribuzione di energia, descritto tramite il tensore metrico $g_{\mu\nu}$. Questo potenziale tensoriale è legato alla curvatura dello spazio-tempo tramite il tensore di Riemann $R_{\alpha\beta\gamma\delta}$.

La forma di queste equazioni è:

$$(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R) = \frac{8\pi G}{c^4}T_{\mu\nu}$$

in cui $T_{\mu\nu}$ è il tensore energia impulso, $R_{\mu\nu}$ il tensore di Ricci, e R lo scalare di curvatura, entrambi ricavati per contrazione del tensore di curvatura di Riemann. Nell'ipotesi di campi gravitazionali deboli e basse energie $|T_{\mu\nu}| \rightarrow 0$, è possibile eseguire una notevole semplificazione nella risoluzione analitica del problema. In questo caso lo spazio-tempo è rappresentabile come somma tra $\eta_{\mu\nu}$ spazio-tempo piatto, ed una perturbazione $h_{\mu\nu}$, in cui $|h_{\mu\nu}| \ll 1$; il tensore $g_{\mu\nu}$ in questo caso si scrive

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$$

Dalla trattazione e dalle equazioni in approssimazione di campo debole si ottiene la così detta teoria linearizzata:

$$(-\frac{\partial^2}{\partial t^2} + \nabla^2)h_{\mu\nu} = \frac{16G}{c^4}\pi T_{\mu\nu}$$

Una conseguenza importante delle equazioni di campo di Einstein, è la previsione dell'esistenza di fenomeni perturbativi alla metrica spazio-temporale di tipo ondulatorio. In assenza di materia, ossia ponendo $T_{\mu\nu} = 0$, è poi possibile ricavare l'equazione che descrive il comportamento del campo nel vuoto, da cui si evidenzia formalmente l'equazione d'onda:

$$(-\frac{\partial^2}{\partial t^2} + \nabla^2)h_{\mu\nu} = 0$$

Tramite una opportuna trasformazione detta gauge trasversa a traccia nulla, è possibile ottenere le soluzioni di questa equazione come sovrapposizione di due polarizzazioni indipendenti denotate con h_+ e h_\times ottenendo:

$$h^{\mu\nu} = h_+ e_{\mu\nu}^+ + h_\times e_{\mu\nu}^\times$$

dove $e_{\mu\nu}^+$ ed $e_{\mu\nu}^\times$ sono tensori di base. Si può verificare che i tensori di base sono legati da una rotazione di $\pi/4$.

1.1.2 Effetto delle onde su particelle libere

Andando ad analizzare il comportamento di una singola particella libera sottoposta all'azione di un'onda gravitazionale, si ha che la particella rimane ferma rispetto alla stessa. Per osservare gli effetti è necessario prendere in considerazione almeno due particelle; si può dimostrare che in questo caso si ha una variazione della distanza propria $L(t)$ tra le due. Analizzando l'equazione di deviazione geodetica risultano più chiare le interazioni dovute al fenomeno della polarizzazione, descrivibile tramite un celebre esempio. Si dispongano ad anello un certo numero di particelle (vedi figura 1.1) in un piano xy , e si ipotizzi l'arrivo di un'onda gravitazionale, in direzione perpendicolare lungo z . Nel caso in cui l'onda abbia solo polarizzazione h_+ , si osserva una deformazione dell'anello lungo gli assi principali, mentre per onda a polarizzazione h_\times si osservano deformazioni lungo le diagonali principali. Questo fenomeno è legato al fatto che le polarizzazioni, sono ruotate l'una rispetto all'altra di 45 gradi, verifica del fatto che la gravità è rappresentata da un tensore di rango due.

Per quantificare questo effetto prodotto dall'onda gravitazionale, è possibile stimare teoricamente la variazione di distanza propria prodotta dalla perturbazione sulla metrica, integrando sull'elemento di lunghezza quadri-dimensionale $ds^2 = g_{\mu\nu} dx^\mu dy^\nu = (\eta_{\mu\nu} + h_{\mu\nu}) dx^\mu dy^\nu$. Nella ipotesi di sola onda polarizzata $e_{\mu\nu}^+$, quello che si ottiene è:

$$ds^2 = -c^2 dt^2 + (1 + h_{xx}(t)) dx^2 + (1 - h_{xx}(t)) dy^2 + dz^2$$

che restituisce

$$L(t) = \sqrt{(1 + h_{xx}(t))x_0^2 + (1 - h_{xx}(t))y_0^2}$$

Da cui si verifica che la variazione delle distanze, lungo i due assi, avviene in controfase.

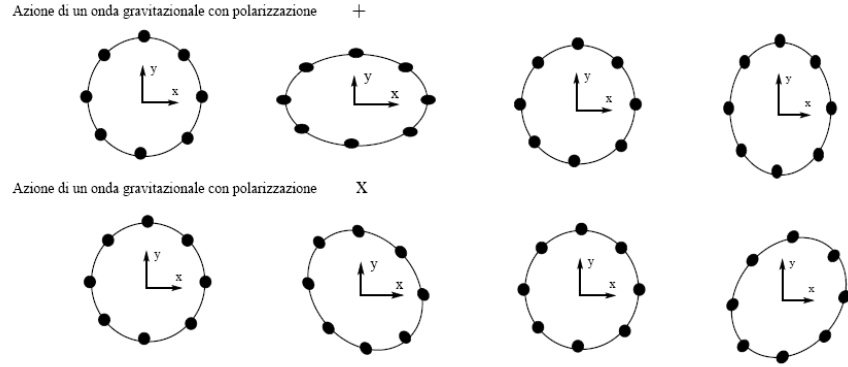


Figura 1.1: Effetto di un onda gravitazionale su un anello di masse test

1.1.3 Generazione di onde gravitazionali

Dalla teoria si ricava come le masse siano sorgenti del campo gravitazionale, il problema ora è poter dedurre teoricamente la dinamica, data l'estrema complessità delle equazioni. Un'ulteriore difficoltà di natura fisica deriva dalle ipotesi su $T_{\mu\nu}$, cioè sul modello da utilizzare per descrivere lo stato della materia. A parte pochissimi casi molto particolari, le uniche tecniche di calcolo che permettono di trattare il problema, si basano su simulazioni numeriche o su tecniche approssimate. È importante analizzare ed aver presente la natura delle sorgenti, data l'impossibilità di poter generare in laboratorio segnali utili per il rilevamento. Infatti, solamente corpi celesti estremamente densi e compatti riescono a produrre onde che abbiano una qualche probabilità di essere osservate.

Sotto ipotesi semplificative opportune, è possibile applicare lo sviluppo in multipolo per il campo gravitazionale. Si ottiene che il primo termine non nullo è quello di quadrupolo, evidenziando la natura quadrupolare delle onde gravitazionali.

L'ente matematico utilizzato per questa tipo di descrizione, è il tensore momento di quadrupolo di massa $Q^{ij}(t-r)$, legato alla distribuzione di materia all'interno delle sorgenti, che nella gauge trasversa a traccia nulla è espresso da:

$$Q_{ij} = \int_V \rho(x_i x_j - \frac{1}{3} \delta_{ij} r^2) dV$$

dove ρ è la densità del corpo in esame. E' possibile capire il motivo per cui i termini di monopolo e dipolo non contribuiscono, osservando le analogie tra la teoria dell'elettromagnetismo e quella gravitazionale. Nel caso di radiazione elettromagnetica è nota la relazione che lega il campo elettrico della radiazione alla derivata seconda nel tempo del termine di momento di dipolo elettrico:

$$\vec{E} = \frac{1}{Rc^2} ((\vec{d} \wedge \vec{n}) \wedge \vec{n}) \quad (1.1)$$

dove R è la distanza dalla sorgente, \vec{n} il versore che punta dalla sorgente all'osservatore e \vec{d} il momento di dipolo elettrico, definito come:

$$\vec{d} = \int dV \rho_q(r) \vec{r}$$

dove $\rho_q(r)$ è la densità di carica relativa alla sorgente. Oltre al termine di dipolo, è possibile utilizzare i termini di sviluppo successivo, che contribuiscono in maniera minore all'energia; mentre non è presente nello sviluppo in multipoli il termine di monopolo, in quanto rappresenta una variazione di carica totale della sorgente, effetto questo proibito dal principio di conservazione della carica¹.

Nel caso gravitazionale, la grandezza che gioca il ruolo della carica è la massa, con l'evidente differenza di aver sempre lo stesso segno. La legge di conservazione della carica è associata, nel caso gravitazionale, alla conservazione dell'energia, quindi di nuovo non è presente nello sviluppo in multipoli il termine di monopolo. Il termine successivo, il momento di dipolo gravitazionale, può essere definito in analogia con la 1.1, dove $\rho_q(r)$ è riferito alla densità massa. In questo caso il principio di conservazione della quantità di

¹Il sistema è isolato

moto proibisce anche questo termine, così come proibisce il termine costruito in analogia al momento di dipolo magnetico

$$\int dV \rho(r) \vec{r} \wedge \vec{v}(r)$$

anch'esso proibito per il principio di conservazione del momento angolare. Quindi è evidente che fino a questo ordine di approssimazione, non si hanno contributi energetici, e che il primo termine diverso da zero è quello di quadrupolo.

È possibile scrivere, in analogia con il caso elettromagnetico, la componente dominante della radiazione elettromagnetica:

$$h_{ij} = \frac{2G}{Rc^4} \ddot{Q}_{ij}$$

dove l'espressione è valutata al tempo ritardato $t - R/c$ dove R è la distanza tra sorgente ed osservatore. La costante di proporzionalità $\frac{G}{c^4}$ è il termine di accoppiamento che, dato il valore estremamente piccolo pari a circa 10^{-46} , evidenzia la debolissima perturbazione introdotta nella metrica.

Il problema peculiare dello sviluppo in multipoli nel caso relativistico, è connesso alla non linearità delle equazioni di campo, che non permette un calcolo ed una somma dei termini in modo analogo al caso elettromagnetico. Tanto più che il regime che interessa una produzione significativa di onde, è caratterizzato da campi gravitazionali estremamente intensi, portando quindi fuori dalle ipotesi utilizzate per l'approssimazione in campo debole. Il termine di quadrupolo è comunque sempre definito, la cosa che cambia è la metodologia utilizzata per eseguire l'espansione. In seguito verranno introdotte alcune tecniche a riguardo, che partono e si differenziano principalmente da ipotesi fisiche sulla sorgente.

Una proprietà importante che è possibile dimostrare, è che moti a simmetria sferica non irradiano, questa significa che data una certa distribuzione di materia, la radiazione è dovuta esclusivamente alla componente non sferica. L'emissione di onde gravitazionali da parte di un sistema comporta una perdita di energia e il flusso che ne consegue è trasportato dalle onde medesime.

1.1.4 Le sorgenti di onde gravitazionali

I segnali utili per il rilevamento effettivo sono quelli provenienti da sorgenti astrofisiche, poichè con le tecniche attuali risulta impossibile generare segnali gravitazionali che abbiano effetti apprezzabili.

I segnali che ci aspettiamo di rivelare si possono raggruppare in famiglie, a seconda delle caratteristiche temporali, associate evidentemente anche al tipo di processo astrofisico che li genera. E' possibile identificarne tre gruppi:

- *Segnali periodici* se costituiti da sovrapposizioni di segnali sinusoidali sufficientemente costanti su scale temporali lunghe.
- *Segnali transienti* con durata breve rispetto ai tempi tipici di osservazione.
- *Segnali stocastici* caratterizzati da fluttuazioni stocastiche, con durate lunghe rispetto ai tempi di osservazione.

I fenomeni periodici sono associati a stelle di neutroni o a sistemi binari, processi transienti ad esplosioni di supernovae o fenomeni di coalescenza, mentre il caso stocastico è legato alla sovrapposizione di più segnali dovuti a diversi processi catastrofici simultanei, od a effetti di gravità quantistica. Di interesse immediato sono la classe dei segnali periodici e transienti, attesi principalmente da corpi celesti con caratteristiche fisiche molto particolari come le stelle di neutroni e buchi neri.

1.1.5 Binarie coalescenti

Un sistema binario è un sistema stellare costituito da due stelle orbitanti attorno al comune centro di massa; da osservazioni astronomiche sembra che questa tali sistemi siano molto frequenti e quindi probabili, rispetto al caso di sistemi a stella singola, come il nostro. La fisica gravitazionale di questi sistemi è determinata dal tipo di stelle che lo costituiscono che possono essere stelle a neutroni, ma anche stelle comuni o addirittura buchi neri.

L'interesse predominante nei confronti di stelle a neutroni è legato alla quantità di energia in gioco, sufficiente a produrre onde gravitazionali di intensità rilevante.

Le due stelle costituenti un sistema binario, a causa della perdita di energia sotto forma anche di emissione di onde gravitazionali, tendono ad avvicinarsi progressivamente, aumentando la frequenza di rivoluzione dell'orbita per raggiungere frequenze dell'ordine di centinaia di Hz. Il processo continua fino alla collisione delle due stelle. E' ovvio che data l' elevatissima densità dei campi gravitazionali in gioco, la collisione non è modellizzabile come l'urto di due sfere rigide, ma nell'ultima fase di coalescenza(detta merging o fusione) avviene un intenso scambio di materia fra i due corpi celesti. Le velocità in gioco risultano frazioni non trascurabili della velocità della luce, fino a $c/3$ [43] e questo regime prende il nome di quasi relativistico; Il processo termina con la formazione di un oggetto unico con il raggio dell'ordine delle stelle progenitrici.

L'ultima fase di vita delle due stelle come entità fisicamente distinguibili, è detta coalescenza, ed è in questa finestra temporale che è rivolta l'attenzione di rivelatori come VIRGO e LIGO. Da stime teoriche è attesa una quantità notevole di segnale gravitazionale transiente, con intensità interessanti per il rilevamento; in termini di sensibilità relativa si parla di $(10^{-21} - 10^{-23}) \frac{1}{\sqrt{\text{Hz}}}$.

1.2 Esperimenti di Rivelazione di Onde Gravitazionali

I primi esperimenti per il rilevamento di onde gravitazionali furono ideati da Weber negli anni 60. Il cui principio di funzionamento è basato sulle proprietà elastiche intrinseche dei materiali. L'idea è di costruire barre cilindriche accordate per risuonare all'arrivo di onde gravitazionali caratterizzate da frequenze particolari. Il meccanismo è schematizzabile considerando due masse m interagenti a distanza l_0 tramite una forza elastica di costante elastica k , libere di oscillare. In presenza di un'onda gravitazionale $h(t) = h_0 \cos(\omega t)$ con polarizzazione diretta lungo la congiungente le due masse, il moto vibrazionale è descritto tramite la seguente relazione :

$$m \frac{d^2 x}{dt^2} = -k(x - L) - \gamma \frac{dx}{dt} + \frac{1}{2} m L \omega^2 h_0 \sin(\omega t)$$

dove γ è il coefficiente di smorzamento viscoso. In condizioni di risonanza, per $\omega_0 = \sqrt{\frac{k}{m}}$ il moto delle masse risulta :

$$x(t) = -\frac{1}{2} \frac{m \omega_0}{\gamma} L h_0 \cos(\omega t)$$

Questi rilevatori a barra risonante sono cilindri di materiale a basso coefficiente viscoso, isolati dall'esterno e lasciati liberi di oscillare longitudinalmente dove l'elasticità intrinseca del materiale funge da forza di richiamo elastico. Al passaggio di un'onda gravitazionale il sistema varia la propria lunghezza, ed in particolare per frequenze prossime a quella di risonanza, il segnale viene amplificato, informazione questa restituita da trasduttori capacitivi. E' evidente come il principale limite di questi sistemi sia la ristretta finestra di osservazione in frequenza.

1.2.1 Metodi interferometrici

Un metodo alternativo, che promette risultati interessanti, è quello interferometrico. Il principio di funzionamento è legato alla possibilità di misurare la variazione di distanze relative tra due punti tramite un interferometro Michelson-Morley in cui sono presenti cavità Fabry-Perot lungo i due bracci. Con questo metodo è possibile risalire al passaggio di un'onda gravitazionale, calcolando la potenza all'uscita dal separatore del fascio P_{out}

$$P_{out} = P_{in} \cos^2(2\Delta\phi) = P_{in} \cos^2(k\Delta L)$$

con $k = \frac{2\pi}{\lambda}$ il vettore d'onda della luce con lunghezza d'onda λ e $\Delta L = L_1 - L_2$ la semidifferenza tra i cammini ottici nelle due direzioni, poiché nel caso di passaggio di onda gravitazionale con polarizzazione h_+ , perpendicolarmente al piano dell' interferometro, si ha che ΔL è uguale a $L_1(1 + h(t)) - L_2(1 - h(t))$. Nel caso di $L_1 = L_2$ si ha che

$$P_{out} = P_{in} \cos^2(k2Lh(t))$$

Un limite importante all'analisi con questo tipo di approccio è dato dal tempo $\tau_0 = \frac{2L}{c}$ di percorrenza della luce nei due bracci, poiché deve risultare inferiore al periodo dell'onda gravitazionale [44] $T_{gw} = \frac{1}{f_{gw}}$, cioè

$$2\pi f_{gw}\tau_0 \ll 1$$

Facendo l'esempio di un interferometro con bracci lunghi $3km$ si ha la condizione di rilevabilità per segnali di frequenza $f_{gw} \ll 8kHz$. In queste relazioni è implicita la condizione che l'onda gravitazionale sia costante per tutto il tragitto compiuto dalla luce all'interno dell'interferometro. Sotto le ipotesi di variazioni trascurabili sulla fase della luce indotte da mutamenti nella metrica, si dimostra [44] che lo sfasamento prodotto dall'onda è dato da

$$k\Delta L = h(t) \frac{2\pi}{\lambda} \tau_0 \left(\frac{\sin(\pi f_{gw}\tau_0)}{\pi f_{gw}\tau_0} \right) e^{i\pi f_{gw}\tau_0}$$

Dalle due relazioni precedenti segue immediatamente che all'aumentare della lunghezza dei bracci, si ha una riduzione di visibilità alle frequenze alte, mentre è agevolata la rilevazione dei segnali più deboli.

Nel mondo è andato via via crescendo il fermento attorno alle tecniche interferometriche, dando vita a varie collaborazioni internazionali. Gli esperimenti più promettenti come scala, sono VIRGO (bracci di $3km$, Italia-Francia) e L.I.G.O. (due interferometri con bracci da $4km$, USA); esistono poi esperimenti in scala ridotta quali GEO (UK-Germania, $600m$) e TAMA (Giappone, $300m$).

1.3 L'esperimento VIRGO

L'esperimento VIRGO è un rivelatore di onde gravitazionali, basato su metodo interferometrico, con bracci lunghi $3km$, al cui interno viene fatto circolare un raggio laser della potenza di $20W$ e lunghezza d'onda di $1064nm$. In figura 1.2 è riportato uno schema ottico di Virgo, dove è possibile individuare varie cavità risonanti che selezionano i modi del laser in ingresso, definendone la frequenza. All'interno di ciascun braccio è inserita una cavità

Fabry - Perot che permette di immagazzinare la luce all'interno, simulando un allungamento del cammino ottico. Nel caso di VIRGO si ottiene un cammino ottico efficace pari a 120km . Una caratteristica importante di questo tipo di interferometro è la scelta del punto di lavoro, detto frangia oscura. Ciò che viene realizzato è un *null-instrument*, cioè un rivelatore che restituisce zero nel momento in cui è assente il segnale a cui siamo interessati, mentre in presenza di segnale gravitazionale permette di ottenere un segnale diverso da zero. Questo è possibile grazie alla tecnica di rivelazione utilizzata detta Pound-Drever.

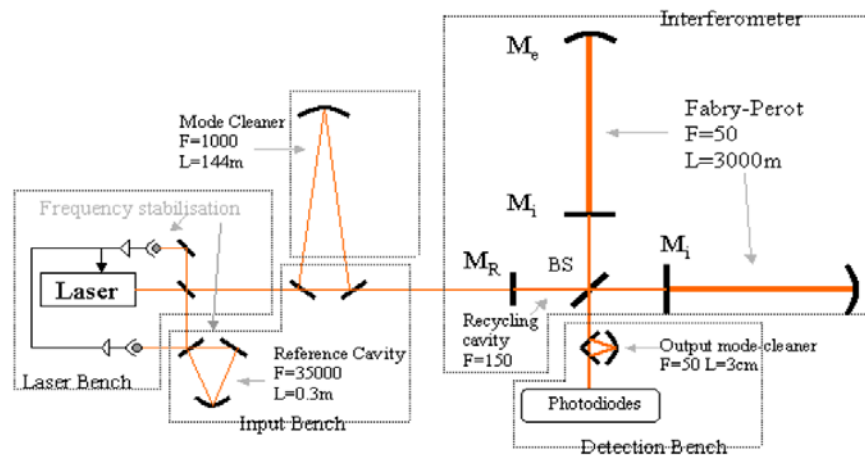


Figura 1.2: Schema ottico di Virgo

Il sito di installazione è Cascina, un comune della provincia di Pisa, dove l'apparato sperimentale è attualmente funzionante in fase di test. La gamma di segnali che l'interferometro intende osservare, nel caso di binarie coalescenti, è compresa tra 10Hz e 2000Hz, ma sono attese osservazioni anche di altri tipi di segnali gravitazionali, come segnali gravitazionali provenienti da pulsar o supernovae.

In figura 1.3 è riportata la sensibilità attesa di VIRGO secondo il progetto. Raggiungere questa tipo di sensibilità richiede l'utilizzo di tecniche raffinate e di soluzioni innovative, attraverso una analisi di ogni particolare. La più piccola fluttuazione interna dell'apparato può essere causa di rumore e può limitarne la sensibilità. I parametri fondamentali su cui

si basa la rivelazione sono sostanzialmente due (1) la distanza relativa tra gli specchi e (2) il cammino ottico della luce.

E' molto importante quindi saper determinare il contributo di ogni tipo di disturbo, cercando eventualmente tecniche di minimizzazione[45]. Le principali fonti di rumore per l'apparato sono:

- *Rumore sismico*, fonte di rumore dominante per frequenze basse inferiori ai 100Hz, dovuto ad attività geologiche ed umane. Per minimizzare questo effetto è stato realizzato il super attenuatore, a cui sono appese le ottiche di VIRGO. Questo si basa su di un sistema di pendoli invertiti che fungono da filtri.
- *Shot noise*[46]: Dovuto alle fluttuazioni statistiche nella rivelazione della luce, dovuto ai rivelatori. Questo rumore limita la sensibilità da 500Hz a 10Khz.
- *Rumore termico* è il limite importante alla sensibilità dell'interferometro VIRGO, nell'intervallo di frequenze comprese fra 5 e 500 Hz. E' legato alle fluttuazioni di un'osservabile macroscopica di un sistema all'equilibrio termico con l'ambiente, a seguito del moto di agitazione termica degli elementi microscopici che lo costituiscono. Si manifesta nel caso dell'interferometro VIRGO principalmente nelle sospensioni, causando spostamenti del centro di massa degli specchi, e negli specchi stessi, deformandone le superfici riflettenti.

Osservando la figura 1.3 è evidente come la regione a sensibilità maggiore è compresa nell'intervallo 10Hz – 1000Hz. In realtà il limite inferiore in frequenza dell'apparato, dipende dal tipo di realizzazione, poichè sono in fase di studio soluzioni innovative, che nel tempo ne permetteranno la diminuzione. Queste metodologie si basano sullo studio di rumore termico di materiali nuovi, e su metodologie avanzate per il fissaggio degli specchi ai fili di sospensione[47].

1.3.1 L'analisi dati in Virgo

Per tutti i tipi di rivelatori, ma specialmente per le antenne terrestri, uno dei problemi principali da affrontare nell'ambito dell'analisi dati, è la rivelazione di segnali con un basso rapporto segnale rumore atteso (per la definizione di segnale rumore vedere capitolo 2 paragrafo 2.3). Il basso valore di questa quantità, è dovuto all'intrinseca debolezza del segnale gravitazionale atteso, rispetto al fondo di rumore strumentale. Per superare questo problema, sono stati fatti molti sforzi per lo sviluppo di tecniche di data analysis in grado di risolvere il problema.

Quando la forma d'onda attesa è in qualche modo nota, la tecnica più utilizzata è quella del *matched filtering* su griglia di template, che si riduce nel calcolo della correlazione tra l'uscita dell'interferometro ed un insieme di template di forme d'onda.

Nel caso dell'antenna Virgo, in condizioni di pieno regime, vengono prodotti ogni giorno, circa 1TB di dati, che devono essere analizzati alla ricerca del debole segnale. L'analisi on-line (ovvero in-time con il tempo di acquisizione dati per sgrossare i dati) richiede,

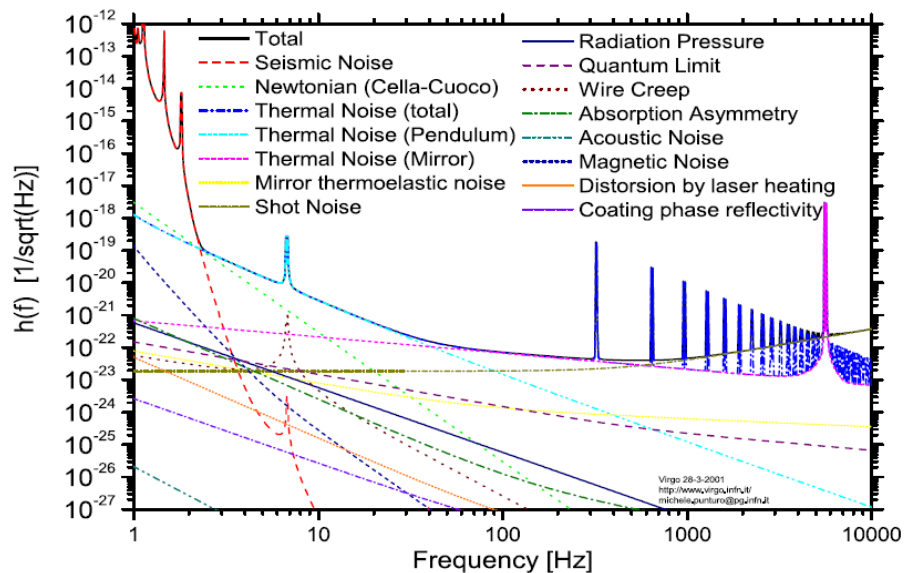


Figura 1.3: Sensibilità attesa di Virgo

utilizzando un approccio standard, circa 300 Gflops per un recupero del 90% del rapporto segnale rumore. Il problema del data analysis è ancora una questione aperta. In questo lavoro di tesi si cercherà di dare una risposta a questa problematica mediante l'introduzione di una metodologia innovativa basata su algoritmi di ottimizzazione globale.

Concettualmente si tratta di creare un metodo che effettua un campionamento nello spazio dei template, invece di procedere in serie esplorando tutto lo spazio costituito dalla griglia dei template possibili.

Capitolo 2

Tecniche di rivelazione

In questo capitolo descriveremo le più diffuse tecniche di rivelazione per l'elaborazione di segnali che verranno applicate nei capitoli successivi per la modellizzazione del problema delle onde gravitazionali. Dopo aver introdotto i differenti test di ipotesi, verrà descritto il problema della rivelazione dei segnali in tempo discreto immersi in rumore bianco ed in rumore gaussiano colorato.

2.1 La teoria della rivelazione

La teoria della rivelazione consiste nello studio delle tecniche per l'elaborazione di segnali al fine di estrarne le informazioni che essi contengono. Tale teoria trova applicazione nei campi più svariati, dalla teoria della comunicazione, del controllo automatico e della trasmissione dei dati, ma anche la radioastronomia, la sismologia e la ricerca delle onde gravitazionali. In generale le applicazioni della teoria della rivelazione consistono nel fare inferenze partendo da dati osservati, generalmente distorti o corrotti da rumore. Inoltre l'informazione che si desidera estrarre dalle osservazioni, non è a fortiori nota all'osservatore; ciò rende quindi indispensabile affrontare il problema della rivelazione con un approccio probabilistico, modellando il comportamento non noto dell'informazione da recepire, mediante variabili

casuali. Da questo punto di vista, la teoria della rivelazione può collocarsi nel campo dell'inferenza statistica.

Fondamentale nello studio della teoria della rivelazione è il concetto di osservazione casuale X che ha valori in un insieme di osservazioni Γ , che può essere un insieme di vettori, forme d'onda, numeri reali etc. Dall'osservazione di X , si desidera estrarre informazioni circa fenomeni collegati ad X ; In particolare, nei problemi di rivelazione si desidera decidere tra un certo numero finito di possibili situazioni. La connessione tra l'osservazione e l'informazione desiderata è di tipo probabilistico, nel senso che il comportamento statistico di X , è influenzato dalle possibili situazioni. Quindi un modello per questi problemi è una famiglia di distribuzioni di probabilità su Γ , i cui elementi corrispondono alle condizioni statistiche presenti nelle varie situazioni possibili. Dato un tale modello, i problemi di rivelazione consistono nel determinare un modo ottimo di elaborare l'osservazione X per estrarre l'informazione desiderata. Le caratteristiche fondamentali che distinguono tali problemi l'uno dall'altro sono la natura dell'informazione desiderata (discreta o continua), la quantità di informazioni a priori, ed i criteri di prestazioni attraverso cui si possono confrontare le varie procedure di rivelazione.

2.2 Elementi del Test delle Ipotesi

La maggior parte dei problemi di rivelazione può essere schematizzata mediante un test di M ipotesi in cui, sulla base di un'osservazione, si desidera decidere quale si è verificata tra le M possibili situazioni statistiche che possono descrivere le osservazioni. Per alcuni problemi di decisione, esistono varie strategie che sono ottime rispetto ad un certo criterio. Di seguito, dopo aver richiamato la definizione di test delle ipotesi binario, esporremo il criterio di Neyman-Pearson ed il Test delle Ipotesi Composto che ci saranno utili per la modellizzazione del problema di rivelazione di segnali gravitazionali di tipo chirp.

2.2.1 Test delle ipotesi binario

Nel test binario consideriamo due ipotesi che indicheremo con \mathcal{H}_0 e \mathcal{H}_1 e che corrispondono a due possibili distribuzioni di probabilità, P_0 e P_1 , sullo *spazio delle osservazioni* (Γ, G) , dove Γ è un insieme e G è una σ -algebra su Γ . Formalmente questo tipo di test può essere scritto nel modo seguente:

$$\mathcal{H}_0 : X \propto P_0 \tag{2.1}$$

$$\mathcal{H}_1 : X \propto P_1$$

dove la notazione $X \propto P$ sta ad indicare che la variabile casuale X che rappresenta l'osservazione, ha la distribuzione P . Le ipotesi \mathcal{H}_0 ed \mathcal{H}_1 , vengono rispettivamente chiamate, *ipotesi nulla* ed *ipotesi alternativa*.

Una *regola di decisione* δ per \mathcal{H}_0 rispetto ad \mathcal{H}_1 , è una partizione dell'insieme delle osservazioni Γ negli insiemi $\Gamma_1 \in G$ e $\Gamma_0 = \Gamma_1^C$ tale che si sceglie H_j quando $x \in \Gamma_j$ per $j = 0, 1$. Gli insiemi Γ_1 e Γ_0 sono rispettivamente detti *regione critica* e *regione di accettazione*. E' chiaro che la regola di decisione può essere pensata come una funzione su Γ definita nel seguente modo:

$$\delta(x) = \begin{cases} 1 & \text{se } x \in \Gamma_1 \\ 0 & \text{se } x \in \Gamma_0 \end{cases}$$

Così che il valore che δ assume in x è l'indice dell'ipotesi accettata. Ovviamente è proprio la scelta del partizionamento che dipende dal tipo di criterio scelto.

2.2.2 Criterio di Bayes

Il criterio di Bayes è basato sull'assegnazione di un costo alle possibili scelte, e sulla conoscenza delle probabilità apriori con cui si presentano le ipotesi \mathcal{H}_0 ed \mathcal{H}_1 . I costi sono assegnati mediante una matrice

$$C = C_{ij} \quad i, j \in 0, 1$$

il cui generico elemento C_{ij} rappresenta il costo di scegliere l'ipotesi H_i quando l'ipotesi H_j è vera, mentre con π_i indicheremo la probabilità apriori associata all'ipotesi H_i . Definiamo *rischio condizionale* per ogni ipotesi, il costo medio contratto dalla regola di decisione δ quando quella ipotesi è scelta, ovvero:

$$R_j(\delta) = C_{1j}P_j(\Gamma_1) + C_{0j}P_j(\Gamma_0), \quad j = 0, 1 \quad (2.2)$$

Indichiamo poi con il termine *rischio di Bayes* la seguente quantità:

$$r(\delta) = \pi_0 R_0(\delta) + \pi_1 R_1(\delta) \quad (2.3)$$

Allora diremo regola di Bayes per \mathcal{H}_0 rispetto ad \mathcal{H}_1 , la regola di decisione δ (vdi paragrafo precedente) che minimizza il rischio di Bayes. Se indichiamo con p_j la densità di P_j , la regola di Bayes è caratterizzata dalla seguente regione di rigetto:

$$\Gamma_1 = \{x \in P | \pi_1(C_{11} - C_{01})p_1(x) \leq \pi_0(C_{00} - C_{10})p_0(x)\} \quad (2.4)$$

Che nel caso in cui $C_{11} < C_{01}$ può essere riscritta come:

$$\Gamma_1 \{x \in \Gamma | p_1(x) \geq \tau p_0(x)\} \quad (2.5)$$

dove

$$\tau = \frac{\pi_0(C_{1,0} - C_{0,0})}{\pi_1(C_{0,1} - C_{11})}$$

. La regola di decisione descritta dalla regione di rigetto 2.5, è nota come *test del rapporto di verosimiglianza*, che ha un ruolo centrale nella teoria del test delle ipotesi. Inoltre definendo il *rapporto di verosimiglianza* tra \mathcal{H}_0 ed \mathcal{H}_1 nel modo seguente:

$$L(x) = \frac{p_1(x)}{p_0(x)}, \quad x \in \Gamma \quad (2.6)$$

con la convenzione che $(k/0) = \infty$ per qualunque $k \geq 0$ la regione di rigetto si può scrivere come

$$\Gamma_1 = \{x \in \Gamma | [p_1(x)/p_0(x)] \geq \tau\}.$$

In sintesi il criterio di Bayes consiste nel calcolare il rapporto di verosimiglianza 2.6 per il valore osservato di X e nel confrontare questo rapporto con la soglia τ .

Formalmente la regola Bayes, per la 2.1 è:

$$\delta_B(x) = \begin{cases} 1 & \text{se } L(x) \geq \tau \\ 0 & \text{se } L(x) < \tau \end{cases} \quad (2.7)$$

2.2.3 Il criterio di Neyman-Pearson

Il criterio di Neyman-Pearson fornisce una regola di decisione per il test delle ipotesi binario, quando non è possibile associare dei costi alle possibili scelte nè si conoscono le probabilità a priori con cui si verificano le ipotesi \mathcal{H}_0 e \mathcal{H}_1 . Per illustrare questo tipo di criterio, iniziamo con l'osservare che nel test binario 2.1 si possono commettere due tipi di errore:

- *errore di primo tipo o falso allarme*, quando l'ipotesi \mathcal{H}_0 è falsamente rigettata
- *errore del secondo tipo o mancanza*, quando l'ipotesi \mathcal{H}_1 è falsamente rigettata.

Quando invece è correttamente scelta l'ipotesi \mathcal{H}_1 , si parla di *rivelazione*. Vale la pena notare che questi termini derivano dai problemi di rivelazione radar, in cui le ipotesi \mathcal{H}_0 ed \mathcal{H}_1 corrispondono, rispettivamente, alla assenza o presenza di un obiettivo. Per una regola di decisione δ , la probabilità di un errore di primo tipo è nota come *probabilità di falso allarme del test* $P_F(\delta)$, mentre la probabilità di un errore di secondo tipo è detta *probabilità di mancanza* $P_M(\delta)$. Spesso al posto di $P_M(\delta)$ prendiamo in considerazione la così detta *probabilità di rivelazione*, $P_D(\delta) = 1 - P_M(\delta)$.

Progettare un test, comporta un bilanciamento tra le probabilità dei due tipi di errore, poiché si può sempre rendere piccola una a spese dell'altra. Il criterio di Neyman-Pearson, cerca di effettuare questo bilanciamento, ponendo un limite sulle probabilità di falso allarme, per poi minimizzare la probabilità di mancanza. Il criterio di progettazione di Neyman-Pearson è quindi:

$$\max_{\delta} P_D(\delta) \quad \text{con } P_F(\delta) \leq \alpha$$

Con α uguale al limite su menzionato, che è detto *livello* del test.

Definendo *regola di decisione randomizzata* una funzione $\tilde{\delta} : \Gamma \rightarrow [0, 1]$ tale che $\tilde{\delta}(x)$ è la probabilità condizionale di accettare \mathcal{H}_1 dato $X = x$ la probabilità di falso allarme per $\tilde{\delta}$ è definita dalla relazione:

$$P_F(\tilde{\delta}) = E_0\{\tilde{\delta}(X)\} = \int_{\Gamma} \tilde{\delta}(x)p_0(x)dx$$

mentre la probabilità di corretta rivelazione è:

$$P_D(\tilde{\delta}) = E_1\{\tilde{\delta}(X)\} = \int_{\Gamma} \tilde{\delta}(x)p_1(x)dx$$

Dove con E_0 ed E_1 indichiamo rispettivamente i valori attesi nelle ipotesi \mathcal{H}_0 ed \mathcal{H}_1 .

Possiamo ora enunciare il seguente:

lemma 1 (*Lemma di Neyman-Pearson*). *Supponiamo che P_j abbia densità p_j e che sia $\alpha > 0$. Allora le seguenti affermazioni sono vere:*

Ottimalità. Assegnata una qualunque regola di decisione $\tilde{\delta}$ tale che $P_F(\tilde{\delta}) \leq \alpha$ è una regola di decisione $\tilde{\delta}'$ definita da

$$\begin{cases} 1 & \text{se } p_1(x) > \eta p_0(x) \\ \gamma(x) & \text{se } p_1(x) = \eta p_0(x) \\ 0 & \text{se } p_1(x) < \eta p_0(x) \end{cases} \quad (2.8)$$

Dove $\eta \geq 0$ e $0 \leq \gamma(x) \leq 1$ con $P_F(\tilde{\delta}') = \alpha$, sia ha che $P_D(\tilde{\delta}') \geq P_D(\tilde{\delta})$.

Esistenza. Per ogni $\alpha \in (0, 1)$ esiste una regola $\tilde{\delta}_{NP}$ della formula 2.8, con $\gamma(x) = \gamma_0$ costante per cui $P_F(\tilde{\delta}_{NP}) = \alpha$.

Unicità. Sia $\tilde{\delta}''$ una regola di Neyman-Pearson di livello α . Allora $\tilde{\delta}''$ deve avere la forma 2.8 eccetto che su un sottoinsieme di Γ con probabilità nulla, rispetto ad entrambe le ipotesi.

Esempio 1

Per illustrare la teoria esposta, consideriamo il seguente test delle ipotesi:

$$\begin{aligned}\mathcal{H}_0 : X &= \epsilon \\ \mathcal{H}_1 : X &= \epsilon + \mu\end{aligned}\quad (2.9)$$

Dove ϵ è una variabile casuale gaussiana con media nulla e varianza σ^2 , la cui funzione densità di probabilità indicheremo con $N(0, \sigma^2)$, mentre $\mu \in R^+$. Poichè nell'ipotesi \mathcal{H}_1 la costante μ non fa altro che cambiare il valor medio dell'osservazione, il test 2.9 può scriversi come:

$$\begin{aligned}\mathcal{H}_0 : X &\propto N(0, \sigma^2) \\ \mathcal{H}_1 : X &\propto N(\mu, \sigma^2)\end{aligned}\quad (2.10)$$

Il rapporto di verosimiglianza è dato da:

$$L(x) = \frac{p_1(x)}{p_0(x)} = \frac{\frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}}{\frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2}} = \exp\left\{\frac{\mu}{\sigma^2}\left(x - \frac{\mu}{2}\right)\right\} \quad (2.11)$$

La regola di Bayes per il test 2.10 è data da

$$\delta_B(x) = \begin{cases} 1 & \text{se } \exp\left\{\frac{\mu}{\sigma^2}\left(x - \frac{\mu}{2}\right)\right\} \geq \tau \\ 0 & \text{altrimenti} \end{cases} \quad (2.12)$$

Dove τ è una soglia appropriata. Inoltre, poichè abbiamo supposto che $\mu \in R^+$, la 2.11 è una funzione strettamente crescente, per cui il test 2.12 può essere scritto in modo equivalente come:

$$\delta_B(x) = \begin{cases} 1 & \text{se } x \geq \tau' \\ 0 & \text{se } x \leq \tau' \end{cases} \quad (2.13)$$

dove $\tau' = \frac{\sigma^2}{\mu} \log \tau + \frac{\mu}{2}$. Il rischio di Bayes si calcola mediante la 2.3 e 2.2 conoscendo $P_j(\Gamma_1)$ con $j = 0, 1$ e sfruttando il fatto che $P_j(\Gamma_0) = 1 - P_j(\Gamma_1)$. Ovviamente poichè $\Gamma_1 = \{x \in R | x \geq \tau'\}$ risulta che

$$P_0(\Gamma_1) = \int_{\tau'}^{\infty} p_0(x) dx = 1 - \Phi\left(\frac{\tau'}{\sigma}\right) = 1 - \Phi\left(\frac{\log \tau}{d} + \frac{d}{2}\right)$$

e

$$P_1(\Gamma_1) = \int_{\tau'}^{\infty} p_1(x) dx = 1 - \Phi\left(\frac{\tau' - \mu}{\sigma}\right) = 1 - \Phi\left(\frac{\log \tau}{d} - \frac{d}{2}\right)$$

Dove $d = \mu/\sigma$ e Φ è la funzione di distribuzione di ϵ . Nel caso particolare di costi uniformi¹ e probabilità a priori uguali, risulta $\tau = 1$ e $\tau' = \frac{\mu}{2}$, e il rischio di Bayes è dato da

$$r(\delta_B) = 1 - \Phi\left(\frac{d}{2}\right)$$

funzione decrescente in d .

La probabilità di falso allarme per il test di Neyman-Pearson è data da

$$P_0(p_1(X) > \eta p_0(X)) = P(X > \eta') = 1 - \Phi\left(\frac{\eta'}{\sigma}\right),$$

dove $\eta' = \sigma^2 \frac{\log \eta}{\mu} + \frac{\mu}{2}$. La soglia η' può essere determinata mediante

$$\eta' = \sigma \Phi^{-1}(1 - \alpha) \quad (2.14)$$

Per α fissato. Inoltre poichè $P(Y = \eta') = 0$ è la randomizzazione non è necessaria e possiamo porre $\gamma_0 = 1$ per cui la regola di Neyman-Pearson è data da

$$\tilde{\delta}_{NP}(y) = \begin{cases} 1 & \text{se } y \geq \eta' \\ 0 & \text{se } y \leq \eta' \end{cases} \quad (2.15)$$

dove η è data da 2.14. La probabilità di rivelazione è

$$P_D(\tilde{\delta}_{NP}) = P_1(Y \geq \eta') = 1 - \Phi\left(\frac{\eta' - \mu}{\sigma}\right) = 1 - \Phi(\Phi^{-1}(1 - \alpha) - d) \quad (2.16)$$

dove $d = \mu/\sigma$. Per α fissato, la 2.16 come funzione di d è detta *funzione di potenza* del test. Se invece fissiamo d , la 2.16 rappresenta la probabilità di rivelazione in funzione della probabilità di falso allarme ed è detta *caratteristica operativa del test*.

¹Per costi uniformi intendiamo che nella matrice dei costi $C_{ij} = 0$ quando $i = j$, e $C_{ij} = 1$ quando $i \neq j$

2.2.4 Test delle Ipotesi Composto

Il test binario 2.1, è detto *semplice* perchè a ciascuna delle ipotesi, corrisponde una singola distribuzione per l'osservazione. Quando invece, ad almeno un' ipotesi è associata più di una distribuzione, si parla di test delle ipotesi *composto*.

Un test delle ipotesi composto può essere modellato da una famiglia di distribuzioni, di probabilità $\{P_\theta | \theta \in \Lambda\}$ dove P_θ è la distribuzione di probabilità dell'osservazione dato che θ è il vero valore del parametro, e i valori dei parametri in Λ rappresentano tutte le possibili situazioni. Per esempio nel caso del test binario 2.1, risulta che $\Lambda = \{0, 1\}$.

Nel caso in cui i parametri delle distribuzioni di probabilità possono essere modellati come variabili casuali, allora è possibile usare una strategia di tipo Bayesiano[10], mentre quando ciò non è possibile, lo sviluppo di test di ipotesi composti, che verificano precise condizioni analitiche di ottimalità, è un obiettivo illusorio.

Un modo per definire l'ottimalità in tali problemi è una generalizzazione del criterio di Neyman-Pearson.

Supponiamo che lo spazio dei parametri Λ possa esser decomposto in due insiemi disgiunti, Λ_0 e Λ_1 che rappresentano i domini dei parametri rispetto alle due ipotesi. Per una regola randomizzata, possiamo definire la probabilità di falso allarme

$$P_F(\tilde{\delta}; \theta) = E_\theta\{\tilde{\delta}(X)\}, \quad \theta \in \Lambda_0$$

e la probabilità di corretta rivelazione come:

$$P_D(\tilde{\delta}; \theta) = E_\theta\{\tilde{\delta}(X)\}, \quad \theta \in \Lambda_1.$$

Supponiamo, come nel criterio di Neyman-Pearson, che desideriamo assicurare che la probabilità di falso allarme non ecceda un certo valore α , allora un test ideale potrebbe essere quello che massimizza $P_D(\tilde{\delta}; \theta)$, per ogni $\theta \in \Lambda_1$, soggetto alla condizione $P_F(\tilde{\delta}; \theta) \leq \alpha$, $\theta \in \Lambda_0$. Tale test è noto come *test uniformemente più potente* di livello α .

Sfortunatamente i test uniformemente più potenti, esistono soltanto in condizioni molto particolari, quando si è costretti ad utilizzare altri tipi di test. Tra questi quello più usato

più di frequente è il così detto *test del rapporto di verosimiglianza generalizzato* o *strategia generalizzata a massima verosimiglianza* che è basato sul confrontare la quantità

$$\frac{\max_{\theta \in \Lambda_1} p_{\theta}(y)}{\max_{\theta \in \Lambda_0} p_{\theta}(y)}$$

con una soglia stabilita. In altre parole, il test del rapporto di verosimiglianza generalizzato, è un test del rapporto di verosimiglianza in cui ai parametri non noti, viene sostituita una loro stima, a massima verosimiglianza [11]. La difficoltà principale nell'uso di un tale tipo di test, è che non esistono risultati generali sul calcolo delle prestazioni, quindi in tal caso si ricorre spesso a simulazioni di Montecarlo, per poter stabilire una soglia.

2.3 Rivelazione di segnali in tempo discreto

In questa sezione applicheremo i criteri esposti precedentemente, con lo scopo di derivare procedure per la rivelazione di segnali immersi in rumore. Data la predominanza di implementazioni digitali di funzioni per l'elaborazione dei segnali, ci limiteremo a considerare il caso a tempo discreto.

Prima della trattazione è opportuno introdurre il concetto di rapporto segnale rumore indicato spesso con l'acronimo inglese SNR (Signal to Noise Ratio):

Definizione 1 (*Rapporto Segnale Rumore*). *Data una serie temporale $x[n] = s[n] + v[n]$, si definisce rapporto segnale rumore la quantità*

$$\rho = \frac{E[s^2]}{E[r^2]}$$

, *cioè il rapporto tra la potenza del segnale utile e la potenza del rumore.*

E' opportuno osservare che in letteratura si trovano anche altre definizioni tra loro diverse di SNR. In questo lavoro di tesi ci riferiremo sempre alla definizione 1 a meno che non venga specificato diversamente.

2.3.1 Descrizione del Problema

Il modello per l'osservazione che desideriamo considerare è quello di una serie temporale a tempo continuo che consiste di un segnale corrotto da rumore, oppure soltanto da rumore senza segnale. Il nostro obiettivo è quello di decidere se il segnale è presente o meno analizzando un numero finito N di campioni della serie temporale osservata.

Questo problema statisticamente può essere modellato con un test delle ipotesi binario sullo spazio di osservazione $(\Gamma, G) = (R^N, B^N)$, con B^N uguale alla σ -algebra di Borel su R^N .

$$\begin{aligned}\mathcal{H}_0 : x[n] &= r[n] \quad n = 1, \dots, N \\ \mathcal{H}_1 : x[n] &= s[n] + r[n] \quad n = 1, \dots, N\end{aligned}\tag{2.17}$$

Dove $x[n]$ è l'osservazione, $r[n]$ rappresenta il rumore e $s[n]$ il segnale.

Procedure ottime per il test precedente possono essere derivate utilizzando i risultati del capitolo 2, se si conoscono i comportamenti statistici del segnale e del rumore. I casi che analizzeremo sono quelli in cui:

- il segnale è completamente noto
- il segnale è noto a meno di un insieme di parametri

Assegnando la distribuzione del rumore mediante la densità p_r , il rapporto di verosimiglianza per 2.17 è dato da:

$$L(x) = \frac{p_r(x - s)}{p_r(x)}\tag{2.18}$$

dove con $x = (x[1], \dots, x[N])$ abbiamo indicato il vettore costituito dai campioni dell'osservazione, con $s = (s[1], \dots, s[N])$ il vettore dei campioni del segnale e con $r = (r[1], \dots, r[N])$, il vettore dei campioni di rumore. Le procedure ottime per 2.17 sono derivate calcolando 2.18, e sono analizzate cercando le appropriate probabilità delle regioni critiche risultanti.

Notiamo inoltre che nel corso della tesi, useremo la parola *ricevitore* per indicare il modello per la risoluzione del problema della rivelazione di segnali.

2.3.2 Rivelazione dei segnali noti in rumore indipendente

Questo problema, noto come rivelazione *coerente*, può essere risolto se si conosce la densità p_r e sfruttando il fatto che i campioni del rumore $r[1], \dots, r[N]$ sono statisticamente indipendenti. In questo caso si ha che:

$$p_r(x) = \prod_{n=1}^N p_{r[n]}(x[n]),$$

dove $p_{r[n]}$ rappresenta la densità marginale di $r[n]$, e

$$L(x) = \prod_{n=1}^N L_n(x[n])$$

con $L_n(x[n]) = p_{r[n]}(x[n] - s[n])/p_{r[n]}(x[n])$. Poichè il logaritmo di x è una funzione strettamente crescente, il test ottimo può essere scritto come:

$$\tilde{\delta} = \begin{cases} 1 & \text{se } \sum_{n=1}^N \log L_n(x[n]) > \log \tau \\ \gamma & \text{se } \sum_{n=1}^N \log L_n(x[n]) = \log \tau \\ 0 & \text{se } \sum_{n=1}^N \log L_n(x[n]) < \log \tau \end{cases} \quad (2.19)$$

Esempio 2 (*Rivelazione coerente in rumore gaussiano bianco*)

Diamo innanzitutto la definizione di rumore gaussiano bianco:

Definizione 2 (*Rumore Gaussiano Bianco*). *Un rumore si dice gaussiano bianco quando lo spettro di potenza è costante su tutta la banda di frequenze considerata, o altresì possiamo dire nel caso dei segnali discreti, quando tutti i campioni $r[1], \dots, r[N]$ sono indipendenti ed identicamente distribuiti con distribuzione marginale $N(0, \sigma^2)$.*

Partendo da questa definizione ed utilizzando i risultati dell'esempio 1, risulta che $\log L_n(x[n]) = s[n](x[n] - s[n]/2)/\sigma^2$, così che il test ottimo 2.19 diventa

$$\tilde{\delta} = \begin{cases} 1 & \text{se } \sum_{n=1}^N s[n](x[n] - s[n]/2) > \tau' \\ \gamma & \text{se } \sum_{n=1}^N s[n](x[n] - s[n]/2) = \tau' \\ 0 & \text{se } \sum_{n=1}^N s[n](x[n] - s[n]/2) < \tau' \end{cases} \quad (2.20)$$

con $\tau' = \sigma^2 \log \tau$. Notiamo che il termine $-\frac{1}{2} \sum_{n=1}^N s^2[n]$ non dipende dall'osservazione quindi può essere incorporato nella soglia in modo che il test 2.20 è equivalente al seguente

$$\tilde{\delta} = \begin{cases} 1 & \text{se } \sum_{n=1}^N s[n]x[n] > \tau'' \\ \gamma & \text{se } \sum_{n=1}^N s[n]x[n] = \tau'' \\ 0 & \text{se } \sum_{n=1}^N s[n]x[n] < \tau'' \end{cases} \quad (2.21)$$

Dove $\tau'' = \tau' + \frac{1}{2} \sum_{n=1}^N s^2[n]$. Tale ricevitore è noto come *rivelatore*.

Un'importante caratteristica di tale ricevitore è data dal fatto che esso opera confrontando l'uscita di un sistema lineare con una soglia. Infatti noi possiamo scrivere che $\sum_{n=1}^N s[n]x[n] = \frac{1}{2} \sum_{n=-\infty}^N h[N-n]x[n]$ dove

$$h[n] = \begin{cases} s[N-n] & \text{per } 0 \leq n \leq N-1 \\ 0 & \text{altrimenti} \end{cases} \quad (2.22)$$

Per cui tale ricevitore può essere immaginato come un sistema che fornisce in entrata l'osservazione x ad un filtro digitale lineare e poi confronta il campione N all'uscita del filtro con la soglia. Questa struttura è nota come *matched filter*

2.3.3 Rivelazione di segnali noti in rumore Gaussiano Colorato

Nell'esempio 2 abbiamo definito il rumore gaussiano bianco, diamo ora invece la seguente:

Definizione 3 *Rumore Gaussiano Colorato* Un rumore si dice gaussiano colorato quando lo spettro di potenza non è costante su tutta la banda di frequenze considerata, o altresì possiamo dire nel caso dei segnali discreti, quando tutti i campioni $r[1], \dots, r[N]$ sono identicamente distribuiti con distribuzione marginale $N(0, \sigma^2)$ ma non sono indipendenti.

Un rumore Gaussiano Colorato è quindi caratterizzato in tempo discerto da un vettore gaussiano r con media nulla e matrice di auto correlazione R , che assumeremo definita positiva.

In caso di rumore colorato il rapporto di verosimiglianza è dato quindi da:

$$L(x) = \frac{p_1(x)}{p_0(x)} = \exp\left\{s^T R^{-1} \left(x - \frac{s}{2}\right)\right\}. \quad (2.23)$$

Il test ottimo basato su 2.23 può essere scritto in maniera più semplice se prendiamo i logaritmi di entrambi i membri e incorporiamo il termine $\frac{1}{2}(s^T R^{-1} s)$ nella soglia, dato che non dipende da x ; infatti si ha:

$$\tilde{\delta} = \begin{cases} 1 & \text{se } s^T R^{-1} s > \tau' \\ \gamma & \text{se } s^T R^{-1} s = \tau' \\ 0 & \text{se } s^T R^{-1} s < \tau' \end{cases} \quad (2.24)$$

dove $\tau' = \log \tau + \frac{1}{2} s^T R^{-1} s$. Inoltre, posto

$$s^T R^{-1} s = \tilde{s}^T x = \sum_{n=1}^N \tilde{s}[n] x[n]$$

possiamo osservare che la struttura del ricevitore è identica a quella per il rumore gaussiano bianco, tranne per il fatto che al segnale s abbiamo sostituito \tilde{s} .

Passiamo adesso ad analizzare le prestazioni di 2.24, osservando che la quantità $T(X) = s^T R^{-1} X$ è una trasformazione lineare del vettore gaussiano X , quindi è essa stessa un vettore gaussiano caratterizzato dalla media e dalla varianza rispetto alle due ipotesi. È possibile verificare facilmente che $T(X)$ ha media $\mu \tilde{s}^T s$ rispetto all'ipotesi \mathcal{H}_1 e media nulla rispetto ad \mathcal{H}_0 , mentre la sua varianza è indipendente dall'ipotesi ed è uguale a $d^2 = s^T R^{-1} s$.

Dall'analisi effettuata sopra, si evince che la randomizzazione non è necessaria, ed inoltre la probabilità di scegliere \mathcal{H}_1 quando H_j è vera per $j = 0, 1$ è data da:

$$P_1(\Gamma_1) = \frac{1}{\sqrt{2\pi d}} \int_{\tau'}^{\infty} e^{-(x-\mu)^2/2d^2} dx = 1 - \Phi\left(\frac{\tau' - \mu}{d}\right). \quad (2.25)$$

e

$$P_0(\Gamma_1) = \frac{1}{\sqrt{2\pi d}} \int_{\tau'}^{\infty} e^{-x^2/2d^2} dx = 1 - \Phi\left(\frac{\tau'}{d}\right) \dots \quad (2.26)$$

Dove Φ e μ hanno il significato come nell'esempio 1.

Per il problema di Bayes è conveniente scrivere le 2.25 2.26 in funzione della soglia originale τ

$$P_j \Gamma_1 = \begin{cases} 1 - \Phi\left(\frac{\log \tau}{d} + \frac{d}{2}\right) & \text{per } j = 0 \\ 1 - \Phi\left(\frac{\log \tau}{d} - \frac{d}{2}\right) & \text{per } j = 1 \end{cases} \quad (2.27)$$

E' infine utile confrontare queste formule con quelle dell'esempio 2.24 per verificare che sono formalmente identiche. Ciò comporta che anche il calcolo della soglia delle prestazioni è identico, con la sola differenza dei valori attribuiti a d e μ .

(Interpretazione di d^2) Dalla precedente discussione è possibile vedere che le prestazioni della rivelazione ottima di segnali deterministici in rumore gaussiano, in generale migliorano al crescere del parametro d . Questo parametro, inoltre, può essere interpretato come una misura del rapporto segnale rumore già definito come rapporto tra la potenza del segnale e la potenza del rumore.

Per verificare ciò considereremo prima il caso di rumore gaussiano bianco e poi colorato. Nel caso di rumore gaussiano bianco risulta che $R = \sigma^2 I$ dove I è la matrice identica di ordine N , inoltre $R^{-1} = \sigma^{-2} I$ così che

$$d^2 = s^T R^{-1} s = \frac{s^T I s}{\sigma^2} = \frac{s^T s}{\sigma^2} = \frac{1}{\sigma^2} \sum_{n=1}^N s^2[n] = N \frac{\bar{s}^2}{\sigma^2}$$

dove $\bar{s}^2 = (1/N) \sum_{n=1}^N s^2[n]$ è la potenza media del segnale.

Notiamo che $\sigma^2 = (1/N) \sum_{n=1}^N E\{r^2[n]\}$ è la potenza media del rumore, quindi d^2 è dato dal rapporto di potenza media segnale-rumore per il numero di campioni. Quindi le prestazioni migliorano se una di queste quantità cresce.

Una simile interpretazione può essere data a d^2 anche nel caso di rumore colorato. Iniziamo con l'osservare che la quantità $\sum_{n=1}^N \tilde{s}[n]x[n]$ è uguale al campione N dell'uscita di un filtro lineare invariante alla traslazione con risposta all'impulso data da

$$\tilde{h}[n] = \begin{cases} \tilde{s}[N-n], & 0 \leq n \leq N-1 \\ 0 & \text{altrimenti} \end{cases} \quad (2.28)$$

Se l'input di tale filtro consiste solo di segnale s , allora il campione N dell'output è

$$\sum_{n=1}^N \tilde{s}[n]s[n] = d^2$$

e la sua potenza è d^4 . Se invece l'input del filtro è soltanto rumore, allora la potenza del campione N dell'output è

$$E\left(\sum_{n=1}^N \tilde{s}[n]s[n]\right) = d^2$$

quindi il rapporto delle potenze all'output dovute al solo segnale e al solo rumore è proprio d^2 .

Quindi possiamo dire che d^2 rappresenta il rapporto delle potenze segnale-rumore all'uscita del filtro usato per la rivelazione ottima al campione N . Intuitivamente è ragionevole aspettarsi che più alto è questo rapporto e meglio il segnale può essere rivelato, e ciò segue chiaramente dal fatto che le prestazioni di rivelazione sono funzioni monotone di d^2 . È interessante notare che il filtro caratterizzato dalla risposta all'impulso 2.28 è tra tutti i filtri lineari ed invarianti alla traslazione quello per cui è massimo il rapporto di potenza segnale rumore all'output del filtro. Inoltre questo è vero anche nel caso di rumore non gaussiano perchè segue soltanto dal fatto che r ha media nulla e matrice di correlazione R , mentre l'ottimalità rispetto ai criteri del capitolo precedente dipende dall'assunzione che il rumore sia gaussiano.

2.3.4 Rivelazione di segnali con parametri non noti

Per le situazioni in cui il segnale è noto tranne che per un certo numero di parametri, il test 2.17 può essere scritto nel modo seguente

$$\begin{aligned} \mathcal{H}_0 : x[n] &= r[n] \text{ con } n = 1, \dots, N \\ \mathcal{H}_1 : x[n] &= s[n, \lambda] + r[n] \end{aligned} \tag{2.29}$$

Dove $\lambda = (\lambda_1, \dots, \lambda_m)$ è il vettore costituito dagli m parametri non noti che variano negli insiemi $\Lambda_1, \dots, \Lambda_m$.

Supponendo che i parametri non noti non possono essere schematizzati come variabili casuali, possiamo usare la strategia generalizzata a massima verosimiglianza discussa in precedenza, dalla quale risulta che, nel caso di rumore gaussiano, essa consiste nel calcolare

la quantità

$$\max_{\lambda} L(x) = \max_{\lambda} \exp\{s^T R^{-1}(x - \frac{s}{2})\},$$

da cui passando ai logaritmi otteniamo:

$$L(x) = \max_{\lambda} \log L(x) = \max_{\lambda} \{s^T R^{-1}(x - \frac{s}{2})\} \quad (2.30)$$

e confrontare il risultato ottenuto con una soglia τ . Questo chiude la fase di modellizzazione del ricevitore di segnali gravitazionali di tipo chirp. I risultati appena ottenuti, verranno richiamati nel capitolo 4 dove verrà implementata una strategia di rivelazione per segnali chirp, massimizzando la 2.30 mediante algoritmi di ottimizzazione globale, scegliendo le masse come parametri per i template campione usati per la rivelazione.

Capitolo 3

Modello per un ricevitore di Segnali Gravitazionali prodotti da Binarie Coalescenti

In questo capitolo verrà introdotto il problema della rivelazione di segnali di tipo chirp. Dopo aver ricavato la forma analitica dei segnali gravitazionali prodotti da binarie coalescenti, verrà modellizzato un rivelatore basandosi sulla strategia generalizzata a massima verosimiglianza. Tale modellizzazione verrà ripresa nel capitolo 4 per l'implementazione di una strategia di rivelazione.

3.1 Segnale Gravitazionale atteso da Binarie Coalescenti

Riprendendo quanto detto nel primo capitolo, trattiamo ora il problema della rivelazione dei segnali gravitazionali prodotti da binarie coalescenti, mediante tecniche interferometriche come il caso di VIRGO. Qui di seguito ci apprestiamo quindi a ricavare la forma del segnale atteso che ci aspettiamo di rivelare attraverso l'apparato sperimentale di VIRGO. In generale l'uscita dell'interferometro è una combinazione lineare dei due campi d'onda

h_+ e h_\times :

$$\frac{\Delta L(t)}{L} = F_+ h_+(t) + F_\times h_\times \equiv h(t). \quad (3.1)$$

Dove i coefficienti F_+ ed F_\times sono dell'ordine dell'unità e dipendono, in maniera quadrupolare, dalla direzione della sorgente e dall'orientamento del rivelatore.

In questo lavoro di tesi concentreremo la nostra attenzione sul segnale gravitazionale atteso da un sistema binario compatto, costituito da stelle a neutroni o buchi neri nella fase di coalescenza.

Come su detto, all'uscita dell'interferometro, il segnale ha la forma 3.1, inoltre, modellando la sorgente come punti massa che orbitano attorno al comune centro di massa in orbite circolari, otteniamo il segnale nell'approssimazione post-Newtoniana che si scrive nel modo seguente:

$$h(t) = \mathcal{A}[\pi f(t)]^{2/3} \cos[\phi(t)] \quad (3.2)$$

dove $f(t)$ è la frequenza istantanea dell'onda gravitazionale, \mathcal{A} è una costante che dipende dalla distanza del sistema dall'interferometro, e $\phi(t)$ è la fase che può essere scritta come:

$$\phi(t) = \phi_0(t) + \phi_1(t) + \phi_{1.5}(t) + \dots,$$

dove $\phi_0(t)$ è la parte newtoniana della fase, mentre $\phi_n(t)$ rappresenta la correzione post-Newtoniana all'ordine n . Precisiamo che le correzioni post-Newtoniane sono necessarie quando non sono più trascurabili gli effetti di prossimità delle due stelle che spiraleggiano per andare a costituire un unico oggetto stellare. In altre parole, entrano in gioco le dimensioni finite delle due stelle e ci sono effetti relativistici da prendere in considerazione.

Durante questo lavoro di tesi utilizzeremo il segnale post-Newtoniano ristretto[40] al secondo ordine, dove la parola ristretto sta ad indicare che le correzioni post-Newtoniane sono incluse soltanto nella fase e non nella ampiezza. Inoltre poichè un segnale modulato in ampiezza ed in fase è detto *segnale chirp*, d'ora in avanti ci riferiremo al segnale gravitazionale come segnale chirp, o, più semplicemente, chirp.

In questo tipo di approssimazione, la fase si scrive come:

$$\phi(t) = \phi_0(t) + \phi_1(t) + \phi_{1.5}(t) + \phi_2(t),$$

Dove

$$\phi_0(t) = \frac{16\pi f_a \tau_0}{5} \left[1 - \left(\frac{f}{f_a} \right)^{-\frac{5}{3}} \right]$$

$$\phi_1(t) = 4\pi f_a \tau_1 \left[1 - \left(\frac{f}{f_a} \right)^{-1} \right]$$

$$\phi_{1.5}(t) = -5\pi f_a \tau_{1.5} \left[1 - \left(\frac{f}{f_a} \right)^{-\frac{2}{3}} \right]$$

$$\phi_2(t) = 8\pi f_a \tau_2 \left[1 - \left(\frac{f}{f_a} \right)^{-\frac{1}{3}} \right]$$

f(t) è data da

$$t - t_a = \tau_0 \left[1 - \left(\frac{f}{f_a} \right)^{-\frac{8}{3}} \right],$$

mentre le τ_i sono delle costanti, dette *chirp time*, che hanno le dimensioni di un tempo, e sono dati da:

$$\tau_0 = \frac{5}{256} \mathcal{M}^{-\frac{5}{3}} (\pi f_a)^{-\frac{8}{3}},$$

$$\tau_1 = \frac{5}{192(\pi f_a)^2} \left(\frac{743}{336} + \frac{11}{4} \eta \right),$$

$$\tau_{1.5} = \frac{1}{8\mu} \left(\frac{\mathcal{M}}{\pi^2 f_a} \right)^{\frac{1}{3}},$$

$$\tau_2 = \frac{5}{128\mu} \left(\frac{\mathcal{M}}{\pi^2 f_a} \right)^{\frac{2}{3}} \left(\frac{3058673}{1016064} + \frac{5429}{1008} \eta + \frac{617}{144} \eta^2 \right),$$

dove f_a è uguale alla frequenza dell'onda all'istante t_a , che sarà presa uguale alla frequenza di taglio inferiore dell'interferometro, $\mathcal{M} = (\mu^3 M^2)^{\frac{1}{5}}$ è il *chirp mass*, $\mu = (m_1 m_2)/(m_1 + m_2)$ massa ridotta, e $M = m_1 + m_2$ massa totale del sistema, infine $\eta = \mu/M$, per maggiori chiarimenti si rimanda alle fonti [55][56][57].

Quello che si evince dalle relazioni precedenti è che il segnale chirp può essere univocamente determinato da differenti scelte dei parametri. Infatti assegnati:

\mathcal{A} , ampiezza del segnale;

t_a , istante in cui il segnale diventa visibile nell'interferometro, *tempo di arrivo*

Φ , fase del segnale in arrivo. Il chirp può essere univocamente determinato da una delle seguenti scelte dei parametri:

- m_1, m_2 , masse del sistema binario.
- una qualunque coppia di chirp time.

In questo lavoro di tesi sceglieremo come parametri le coppie di masse, e quindi avremo, $\Theta = \mathcal{A}, t_a, \Phi, (m_1, m_2)$, e chiameremo *template* ciascun segnale chirp con assegnati valori dei parametri, e che indicheremo con $h(t, \Theta)$.

Date alcune caratteristiche del chirp, tra cui l'asintoticità (confronta con [58]) è possibile ricorrere alla approssimazione a fase stazionaria per il calcolo dello spettro (per approfondimenti vedi [59]). Da cui la trasformata di Fourier del segnale chirp per le frequenze positive è data da:

$$\tilde{h}(f) = \mathcal{A}f^{-7/6} \exp[i \sum_{k=1}^6 \psi_k(f) \lambda^k - i \frac{\pi}{4}]$$

dove λ^k con $k = 1, \dots, 6$ rappresenta i parametri post-Newtoniani

$$\lambda^k = \{t_a, \Phi, \tau_0, \tau_1, \tau_{1.5}, \tau_2\}$$

mentre

$$\psi_1 = 2\pi f$$

$$\psi_2 = -1$$

$$\psi_3 = 2\pi f - \frac{16\pi f_a}{5} + \frac{6\pi f_a}{5} \left(\frac{f}{f_a}\right)^{-\frac{5}{3}}$$

$$\psi_4 = 2\pi f - 4\pi f_a + 2\pi f_a \left(\frac{f}{f_a}\right)^{-1}$$

$$\psi_5 = -1\pi f 5\pi f_a - 3\pi f_a \left(\frac{f}{f_a}\right)^{-\frac{2}{3}}$$

$$\psi_6 = 2\pi f - 8\pi f_a + 6\pi f_a \left(\frac{f}{f_a}\right)^{-\frac{1}{3}}$$

Per $f < 0$ la trasformata di Fourier si calcola utilizzando l'identità $\tilde{h}(-f) = \tilde{h}^*(f)$, valida per le funzioni reali.

Concludiamo questa sezione introducendo due concetti che saranno utili nel seguito di questo capitolo: *le componenti in quadratura* del segnale chirp, ed il segnale chirp *normalizzato*. Le componenti in quadratura del segnale chirp, sono definite da

$$h_0(t) = [\pi f(t)]^{\frac{2}{3}} \cos[\phi(t)], h_{\pi/2}(t) = [\pi f(t)]^{\frac{2}{3}} \cos[\phi(t) + \frac{\pi}{2}] \quad (3.3)$$

dalle quali risulta che il segnale chirp 3.2 può scriversi come:

$$h(t) = \mathcal{A}h_0(t)\cos\Phi + \mathcal{A}h_{\pi/2}(t)\sin\Phi.$$

Per il concetto di segnale chirp normalizzato, introduciamo il seguente prodotto scalare nell'insieme dei segnali chirp, la cui importanza sarà chiarita in seguito:

$$\langle x(t), y(t) \rangle = \int_{-\infty}^{\infty} \frac{\tilde{x}(f)\tilde{y}^*(f)}{S_n(f)}, \quad (3.4)$$

dove $S_n(f)$ è la densità spettrale di potenza del rumore dell'interferometro. A questo punto, introdotta la norma derivante dal prodotto scalare 3.4, si può verificare, che per ogni template essa è indipendente sia dal tempo di arrivo che dalla fase iniziale nel caso in cui $\mathcal{A} = 1$. Il suo valore si indicherà con \mathcal{N} . Possiamo così definire il chirp normalizzato nel seguente modo:

$$q(t; \Theta) = \frac{h(t; \Theta)}{\mathcal{N}},$$

analogamente saranno definite le componenti in quadratura del segnale chirp normalizzate.

Infine, definiamo la quantità

$$\mathcal{S} = \mathcal{A}\mathcal{N}$$

detta intensità del segnale chirp, possiamo considerare l'insieme dei chirp normalizzati, univocamente determinati da $\Theta = (\mathcal{S}, t_a, \Phi, m_1, m_2)$. Si può verificare inoltre [41], che il parametro \mathcal{S} è proprio uguale al rapporto segnale rumore, così come definito in [42].

3.2 Modello per il ricevitore di segnali gravitazionali Chirp

Possiamo occuparci ora della determinazione del modello per il ricevitore di segnali gravitazionali chirp. Ricordiamo che il nostro problema è quello di determinare una procedura

per poter stabilire se l'uscita dell'interferometro è costituita da solo rumore o da rumore più segnale.

Dalle sezioni precedenti si evince che questo tipo di problema è equivalente ad un test delle ipotesi che in questo caso è composto, in quanto i chirp dipendono da un certo numero di parametri considerati come variabili casuali. La strategia proposta sarà dunque quella generalizzata a massima verosimiglianza (vedi capitolo 2).

Per maggiore chiarezza riscriviamo il test visto nel capitolo 2 per il caso di segnali chirp:

$$\begin{aligned} H_0 : x[n] &= r[n] \quad n = 1, \dots, N \\ H_1 : x[n] &= h[n - n_a; \lambda] + r[n] \quad n = 1, \dots, N \end{aligned} \quad (3.5)$$

Dove $x[n]$ è il segnale all'uscita dell'interferometro, che rappresenta le nostre osservazioni, $r[n]$ è il rumore dell'interferometro, e $h[n - n_a; \lambda]$ è il segnale chirp.

Quindi il ricevitore di segnali chirp prima calcola la quantità

$$\max_{\delta} \log L(x) = \max_{\delta} \left\{ h^T R^{-1} x - \frac{1}{2} (h^T R^{-1} h) \right\} \quad (3.6)$$

dove δ è l'insieme dei parametri del chirp campionato, e poi confronta questa quantità con una soglia τ . Nel seguito faremo riferimento a 3.6 come *statistica di decisione*.

Come già accennato in precedenza, i segnali chirp possono essere caratterizzati univocamente da differenti scelte di parametri. In particolare, nel caso discreto, fissati l'ampiezza A , il campione corrispondente al tempo di arrivo n_a , e la fase iniziale Φ , un chirp è univocamente determinato o dalla coppia di masse (m_1, m_2) o da una qualunque coppia di chirp time. Per motivi di semplicità, la coppia di chirp time considerata è $(\tau_{1.5}, \tau_0)$.

A questo punto per verificare qual è la scelta più conveniente per i parametri, notiamo che i segnali chirp sono simmetrici rispetto alle masse, per cui supponendo che queste varino nell'intervallo $[1,30]$, possiamo limitarci a considerare come spazio delle masse un dominio triangolare, mentre il dominio dei chirp time risulta essere più complesso.

Queste considerazioni ci suggeriscono di scegliere le masse come parametri dei segnali chirp. In conclusione quindi, i parametri dei segnali chirp sono $\delta = (A, n_a, \Phi, m_1, m_2)$.

A questo punto occorre massimizzare la funzione $L(x)$ in 3.6 rispetto a δ . Tuttavia prima di eseguire questa massimizzazione osserviamo che usando l'approssimazione asintotica della matrice di autocorrelazione R , otteniamo la versione discreta del prodotto scalare

$$\langle x, y \rangle = x^T R^{-1} y = \sum_{k=1}^N \frac{X[k]Y^*[k]}{S_n[k/N]} \quad (3.7)$$

dove $X[k], Y[k]$ sono le trasformate discrete di Fourier di $x[n], y[n]$ e $S_n[k/N]$ è la versione campionata dello spettro del rumore dell'interferometro. Mediante la 3.7, la 3.6 si scrive nel dominio delle frequenze come

$$\max_{\delta} \log L(x) = \max_{\delta} (\langle h, x \rangle - \frac{1}{2} \langle h, h \rangle) \quad (3.8)$$

La massimizzazione di 3.8 rispetto ad A e Φ , può essere eseguita analiticamente.

Infatti utilizzando le componenti in quadratura del segnale chirp e ricordando le relazioni

$$\langle h_0, h_{\pi/2} \rangle \cong 0$$

$$\langle h_0, h_0 \rangle \cong \langle h_{\pi/2}, h_{\pi/2} \rangle$$

la 3.8 diventa:

$$\max_{\delta'} \log L(x) = \max_{\delta'} \frac{(\langle x, h_0 \rangle)^2 + (\langle x, h_{\pi/2} \rangle)^2}{2 \langle h_0, h_0 \rangle} \quad (3.9)$$

dove $\delta' = (n_a, m_1, m_2)$.

Inoltre ricordando il significato di segnale chirp normalizzato, la 3.9 si scrive come

$$\max_{\delta'} \log L(x) = \max_{\delta'} \frac{(\langle x, q_0 \rangle)^2 + (\langle x, q_{\pi/2} \rangle)^2}{2} \quad (3.10)$$

Infine per attenerci alla letteratura esistente, anzichè calcolare 3.10, calcoleremo la quantità

$$\mathcal{L}(x) \max_{\delta'} Q(x) = \sqrt{(\langle x, q_0 \rangle)^2 + (\langle x, q_{\pi/2} \rangle)^2} \quad (3.11)$$

Dove abbiamo incorporato il 2 presente al denominatore nella soglia.

In sintesi, il nostro ricevitore per i segnali chirp ha la seguente forma:

$$\delta(x) = \begin{cases} 1 & \text{se } L'(x) \geq \tau \\ 0 & \text{se } L'(x) < \tau \end{cases} \quad (3.12)$$

La massimizzazione di Q , rispetto ai parametri δ' , non può essere eseguita analiticamente per cui le tecniche utilizzate sono di natura prettamente numerica come vedremo nel prossimo capitolo.

Capitolo 4

Il problema dell'analisi on-line

Come già ricordato, l'interferometro Virgo, produce oltre 1TB di dati di acquisizione la giorno. L'analisi fisica delle informazioni contenute in questa enorme mole di dati, deve essere preceduta da una fase di trigger, necessaria per setacciare le zone interessanti indicata come *analisi on-line*. In questo capitolo, dopo aver esposto il problema, verranno descritte le tecniche di analisi attualmente utilizzate ed i costi computazionali richiesti.

4.1 L'analisi on-line

L'interferometro Virgo è attualmente in fase di test e si appresta ad entrare in funzione a pieno regime. La fase di presa dati è già stata avviata durante i così detti run e minirun scientifici organizzati con frequenze mensili, in cui si testa l'intero apparato per brevi periodi di tempo, generalmente per un massimo di 7 giorni consecutivi.

Quando è in presa dati, l'interferometro Virgo produce un file di 1.6Gb ogni 3 minuti circa, contenente l'uscita dell'interferometro e i dati prodotti su tutti i canali di acquisizione, compresi quelli ambientali e di controllo. Nasce quindi il problema di sgrossare questa enorme quantità di dati al fine di selezionare le zone interessanti per la ricerca di segnali gravitazionali. Tale attività viene indicata con il termine *analisi on-line*, e rappresenta un tema cruciale e di grande attualità all'interno della Virgo Collaboration.

I problemi fondamentali sono legati ai costi computazionali richiesti dalle tecniche di analisi attualmente in uso, e dalla velocità di acquisizione dell'interferometro, che lascia poco tempo per essere perfettamente on-line o quantomeno per non accumulare un ritardo eccessivo.

Definizione 4 *Si definisce in-time factor il rapporto tra il tempo impiegato per acquisire una assegnata quantità di dati e quello necessario per analizzarli.* Tale fattore consente di caratterizzare un software per l'analisi online, fornendo un indice che rappresenta quanto siamo lontani dall'essere completamente in-time con le acquisizioni. Il raggiungimento di tale traguardo con le tecniche di analisi classiche prevede l'uso di un grande numero di processori e l'implementazione degli algoritmi di analisi su strutture di calcolo parallelo[39]

Nell'ambito di questo lavoro di tesi verrà proposta una strategia per ottenere un algoritmo di rivelazione che risulti pienamente in-time abbattendo i costi computazionali rispetto alle strategie classiche, mantenendo, nel contempo, elevate prestazioni di rivelazione.

4.2 Griglie di template e costi computazionali

Per eseguire la massimizzazione della statistica di decisione 3.11 rispetto ai parametri, la tecnica maggiormente utilizzata è quella di creare una griglia di template q ciascuno responsabile di una porzione dello spazio dei parametri. Tale strategia viene spesso indicata in letteratura come *Standard (o Static) Matched Filter Technique (SMFT)*. Le variabili indipendenti di questa strategia sono il numero di parametri che si vuole prendere in considerazione e il numero di template legato direttamente alla capacità del rivelatore.

La scelta del modo ottimale con cui piazzare i template nello spazio dei parametri è un argomento di studio, esistono tuttavia alcuni limiti sul numero di essi, legati alla capacità di rivelare il segnale, ovvero alla perdita di SNR accettabile.

In [60] viene svolta un'attenta analisi di questi argomenti, legata principalmente ai costi computazionali per l'analisi online, richiesti a seconda dei parametri scelti e del numero di template utilizzati. A questo proposito è opportuno ricordare che le operazioni dominanti

nel calcolo di un match tra un template e una trancia di dati sono il calcolo della FFT del template, poichè il match viene fatto nel dominio delle frequenze, e l'operazione di massimizzazione rispetto al campione di arrivo, che dipende tuttavia dalla lunghezza dei dati in esame. Nel caso dei segnali chirp generati da sistemi binari in coalescenza, i segnali campione non hanno tutti la stessa dimensione, ma i segnali generati da masse piccole risultano più lunghi, ciò comporta un maggior costo computazionale per il calcolo delle FFT. Ci sono inoltre altre due variabili che intervengono pesantemente sulla lunghezza del segnale, che sono la frequenza di taglio inferiore e quella superiore (in seguito indicate rispettivamente con f_l ed f_u) che indicano il range di frequenze in cui decidiamo di valutare il segnale. Le formule principali che si ottengono in [60] sono le seguenti:

1. *Numero di Template*

$$\mathcal{N}(MM) = \frac{\int d^N \sqrt{\det g}}{(2\sqrt{(1-MM)/N})^N} \quad (4.1)$$

dove MM detto *minimal match* rappresenta il minimo valore di match che la griglia deve garantire con un segnale, ovvero indica la minima perdita di segnale rumore accettabile; N è il numero di parametri mentre g rappresenta la metrica nello spazio dei parametri. E' chiaro che il numero dei template necessari per garantire un fissato minimal match, cresce con l'aumentare del numero di parametri e con il diminuire del minimal match. Nella tabella 4.1 viene mostrato un esempio di calcolo del numero di template in funzione del limite inferiore scelto per il parametro massa, per garantire $MM = 0.95$.

$m_{\min} = 0.25 M_{\odot}$	$m_{\min} = 0.5 M_{\odot}$	$m_{\min} = 1.0 M_{\odot}$
4.9×10^5	1.0×10^5	2.1×10^4

Figura 4.1: Numero di template in funzione della massa minima

2. *Complessità di Spazio*

$$\mathcal{S} = 4 \times 2 \times (2f_u) \times \mathcal{N} \times \langle \tau_0(f_l) \rangle$$

dove f_u ed f_l rappresentano rispettivamente le frequenze di taglio inferiore e superiore, mentre con $\langle \tau_0(f_l) \rangle$ indichiamo la lunghezza media di un template.

3. Complessità Computazionale

$$\mathcal{P} = (2f_u) \times \mathcal{N} \times [36(\ln_2(2f_u) + \langle \ln_2(\tau(f_l)) \rangle) + 54] \text{ flop/sec} \quad (4.2)$$

Nella tabella 4.2 sono calcolate alcune stime dei costi computazionali e di spazio necessari per garantire un recupero del 90% del rapporto segnale rumore, sulla base delle su citate formule. I valori della tabella sono stati calcolati con $f_l = 30\text{Hz}$ ed $f_u = 1000\text{Hz}$ E' pos-

$\frac{m_{\min}}{M_{\odot}}$	$2\mathcal{N}$	$\mathcal{P}(\text{flop/sec})$	$\mathcal{S}(\text{words})$	SNR recovery
0.25	9.8×10^5	3.0×10^{11}	4.7×10^{11}	90%
0.5	2.0×10^5	5.6×10^{10}	3.2×10^{10}	90%
1.0	4.2×10^4	1.1×10^{10}	2.3×10^9	90%

Figura 4.2: Complessità computazionale e di Storage

sibile aumentare le prestazioni del rivelatore abbassando la frequenza di taglio inferiore f_l . Ad esempio per $f_l = 20\text{Hz}$ possiamo arrivare fino ad un recupero del 95% del SNR , tuttavia bisogna fare i conti con il costo computazionale che cresce logicamente rispetto ad f_l . Nel caso citato, il passaggio da 90% al 95% di SNR comporterebbe un aumento dei costi computazionali di un fattore 5.

Da queste osservazioni è chiaro che un approccio di tipo $SMFT$ richiede dei costi computazionali non facilmente sostenibili, ricordiamo inoltre che i valori ottenuti riguardano esclusivamente l'analisi on-line che deve lavorare da trigger per sgrossare le parti di dati interessanti. Tuttavia in un'analisi più accurata sui dati rimanenti (analisi off-line), è opportuno utilizzare altri parametri, unitamente alle masse, per realizzare i template quali ad esempio gli spin delle singole stelle e l'orientazione del sistema binario rispetto al rivelatore. L'aumento dei parametri comporta l'aumento del numero dei template necessari per garantire le prestazioni del rivelatore e quindi una crescita ulteriore del costo computazionale.

4.3 Calcolo del tempo di esecuzione

Sulla base di quanto detto ricaviamo dunque quali sono i costi computazionali necessari per affrontare il problema della rivelazione utilizzando un approccio statico.

In riferimento alla tabella in figura 4.2, se consideriamo il caso più semplice, ovvero range di template con massa minima 1, abbiamo bisogno di 4.2×10^4 template per ricoprire lo spazio dei parametri, garantendo prestazioni accettabili del ricevitore. Se moltiplichiamo tale valore per il tempo necessario per eseguire un match tra dati e template, otteniamo una limite inferiore per la stima dei tempi necessari per eseguire una analisi di rivelazione con metodi SMFT.

E' opportuno notare che tutte le misure dei tempi di esecuzione di un match, sono altresì vincolate al fattore tecnologico. Questo fattore è stato analizzato, nell'ambito della Collaborazione VIRGO mediante un benchmark chiamato mfbench[38], che valuta esattamente il numero di match per secondo in funzione della lunghezza del template. Nelle figure 4.3 e 4.4 sono riportati alcuni risultati eseguiti su svariate piattaforme.

Sulla base dei test di benchmark, possiamo calcolare una prima stima per il tempo necessario all'algoritmo di SMFT. Se consideriamo il caso migliore, ovvero di disporre di piattaforme Opteron o Dual Itanium, abbiamo che, mediamente, riusciamo ad eseguire 50 match per secondo. Da cui per ricoprire l'intera griglia abbiamo bisogno di $(4.2 \times 10^4)/50 = 840$ secondi con una singola workstation. Tale conto, inoltre, viene fatto per una quantità di dati uguale (in termini di numero di campioni) alla lunghezza del template, e non tiene conto del processo di massimizzazione rispetto al campione di arrivo, che, sebbene trascurabile rispetto al calcolo delle FFT, per un gran numero di template può contribuire sensibilmente al tempo di esecuzione.

Se pensassimo di estendere le masse inferiori fino a $0.25M_{\odot}$ sempre dalla tabella in figura 4.2, otteniamo un tempo di esecuzione minimo pari a $(9.8 \times 10^5) / 50 = 19600$ secondi su una singola workstation, per cui per risultare in-time su tranch di dati da 1 minuto ciascuno, avremmo bisogno di 326 processori sempre trascurando il tempo di

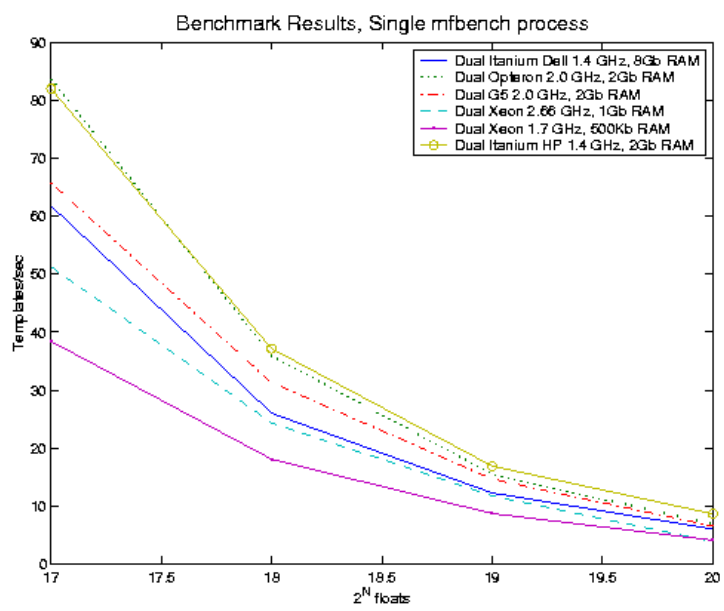


Figura 4.3: benchmark mfbench su differenti piattaforme

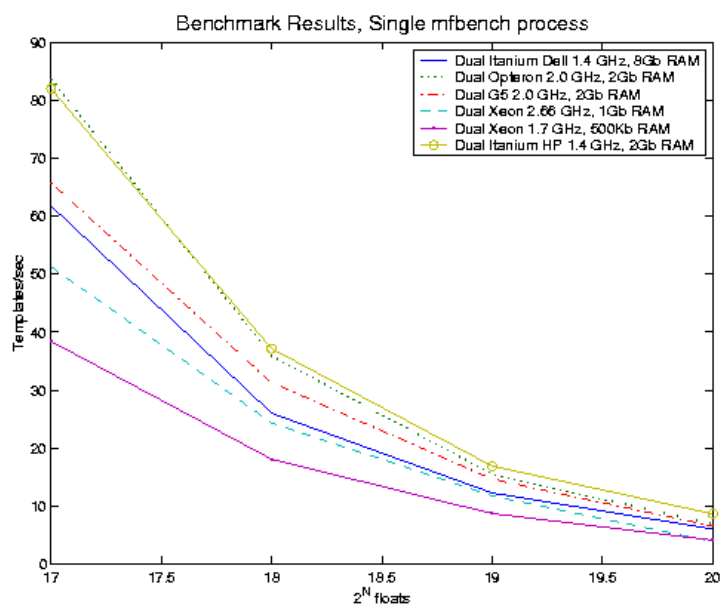


Figura 4.4: benchmark mfbench su differenti piattaforme

massimizzazione rispetto al campione di arrivo.

Nell'ambito della Virgo Collaboration, il problema dell'analisi on-line è stato affrontato mediante la piattaforma Merlino, un sistema distribuito per l'analisi dei segnali che utilizza l'approccio *SMFT*. Test svolti durante i run scientifici hanno mostrato un in-time factor 1.8 utilizzando una griglia di 10481 template, inferiore quindi alle stime sopra citate, e con l'ausilio di 34 CPU [39] in una architettura piuttosto complessa schematizzata in figura 4.5. In questo lavoro di tesi viene proposto un approccio alternativo al *SMFT* indicato come

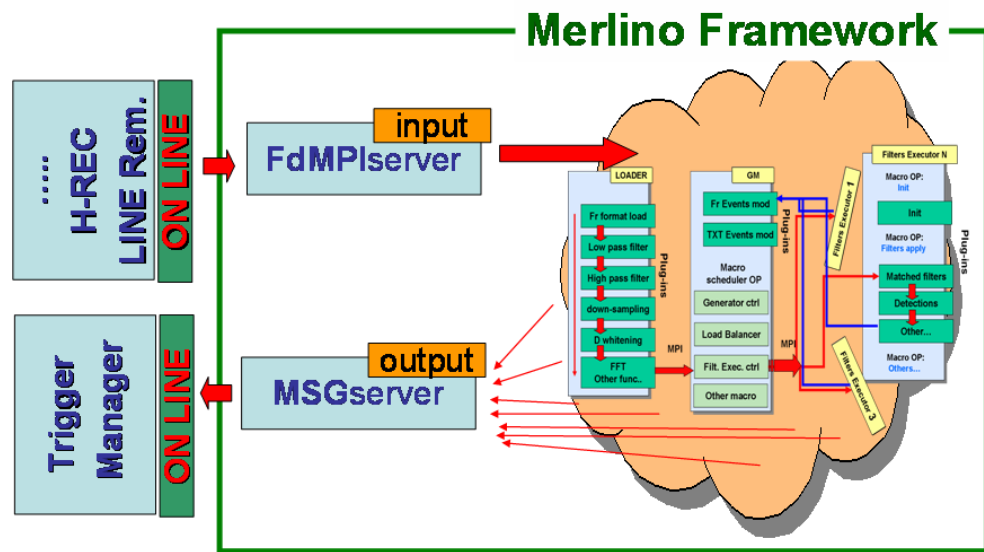


Figura 4.5: Schema on-line della piattaforma Merlino

Dynamic Matched Filter Technique (DMFT) al fine di abbattere i costi computazionali per l'analisi on-line. Infine come sviluppi futuri, verrà analizzata la possibilità di utilizzare il Grid Computing per rispondere alla ben maggiore potenza di calcolo richiesta dall'analisi off-line.

Capitolo 5

Dynamic Matched Filtering per la rivelazione di onde gravitazionali

Nei capitoli precedenti abbiamo visto come le tecniche classiche di analisi dati basate su SMFT richiedono l'utilizzo di molti processori ed un grande sforzo tecnologico, per rimanere in-time con l'aquisizione dati. In questo capitolo viene proposto un metodo innovativo, basato su una strategia dinamica, dandone altresì una prima implementazione. Una serie di test sul software prodotto, mostrano le interessanti prestazioni sia in termini di rivelazione che di tempi di esecuzione, che premiano sia la bontà del metodo che delle specifiche scelte implementative.

5.1 Rivelazione di segnali Chirp

Nei capitoli precedenti abbiamo visto che il problema della rivelazione dei segnali gravitazionali chirp, può essere schematizzato come un test delle ipotesi. Inoltre, poichè il segnale chirp dipende da parametri che non sono noti, e che non possono essere modellati come variabili casuali, ciò richiede l'utilizzo della strategia a massima verosimiglianza generalizzata.

Questa modellizzazione, ha ricondotto il nostro problema alla ricerca del massimo di

una funzione: *la funzione di verosimiglianza*, che dipende dai parametri del segnale chirp. Tuttavia abbiamo visto che questa massimizzazione, può essere eseguita analiticamente, soltanto rispetto ad alcuni dei parametri, ciò comporta la necessità di utilizzare griglie di template secondo la strategia dello Standard (Static) Matched Filter Technique (SMFT) che, tuttavia, richiede un elevato costo computazionale. In questo lavoro di tesi proponiamo quindi un approccio dinamico mediante l'utilizzo di algoritmi di ottimizzazione globale per calcolare il massimo della funzione di verosimiglianza rispetto ai restanti parametri.

5.1.1 Dynamic Matched Filter Technique (DMFT)

Il metodo DMFT [21][23][31][32][33] può essere scomposto in quattro step

1. *Modellizzazione*, in cui si stabilisce la funzione da massimizzare dato, nel nostro caso, dal rapporto a massima verosimiglianza e i parametri rispetto ai quali si vuole eseguire la massimizzazione.
2. *Pianificazione*, in questa fase si stabilisce primariamente l'algoritmo di ottimizzazione globale da utilizzare, quindi si definiscono i suoi parametri ed i loro domini, infine si stabilisce la soglia che servirà nella fase finale per decidere se c'è stata o meno rivelazione.
3. *Esecuzione*, che prevede l'esecuzione vera e propria dell'algoritmo una volta organizzati i dati di input che si intende analizzare.
4. *Rivelazione*, fase conclusiva, in cui si confronta il valore massimo restituito dall'algoritmo di ottimizzazione con il valore della soglia prestabilita, nel caso essa venga superata, si ha una rivelazione, con una certa probabilità di falso allarme, in caso di rigetto, invece, si ritiene che nei dati vi sia solo rumore.

In figura 5.1 è riportato uno schema del tutto esemplificativo usato per chiarire i vari passi di cui è composto il metodo.

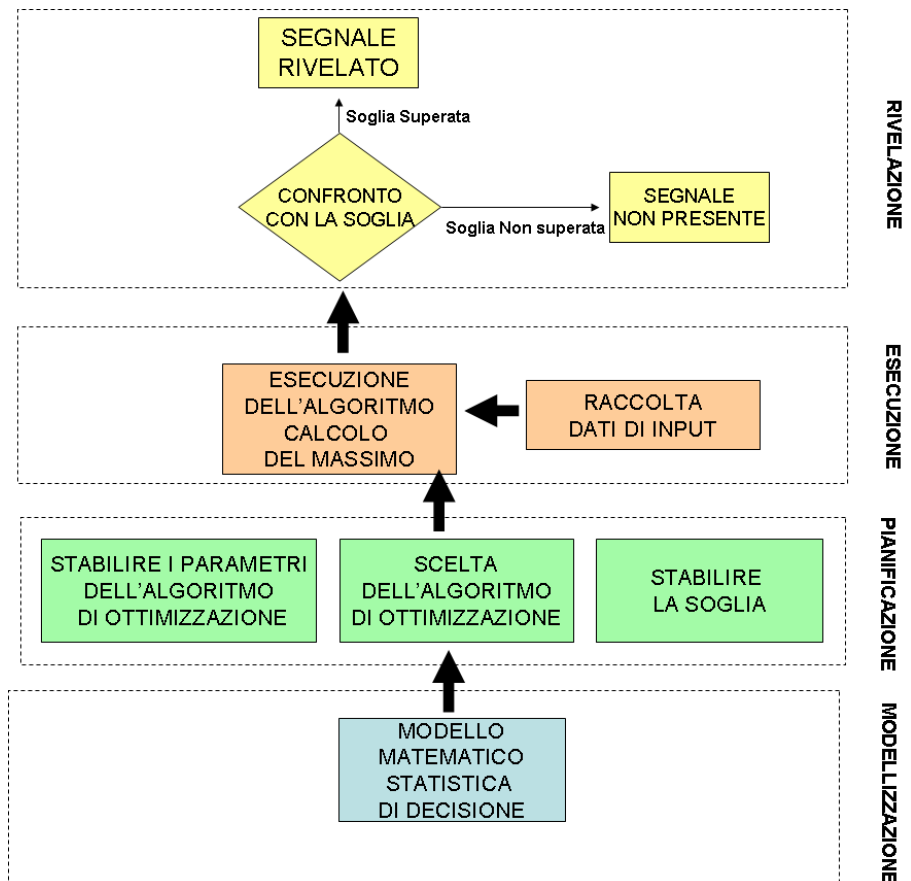


Figura 5.1: Schema del Dynamic Matched Filter Technique (DMFT)

5.2 Sperimentazione del metodo

Durante il lavoro di tesi, il metodo proposto è stato testato al fine di valutarne le capacità. Data la complessità della funzione obiettivo, dovuta soprattutto alla presenza di molti massimi locali, è stato deciso di utilizzare un algoritmo di tipo euristico, in particolare l'algoritmo di Price [18][19] analizzato dettagliatamente nella appendice A. Di seguito verranno esposti tutti i passi svolti secondo il modello descritto nella sezione precedente ovvero, il modo in cui è stata calcolata la statistica di decisione, come è stata stabilita la soglia ed i parametri dell'algoritmo, la generazione dei dati seguita da una serie di risultati sperimentali. Verranno discusse infine le proprietà di rivelazione del metodo in riferimento anche ai tempi di esecuzione.

5.2.1 La statistica di decisione - Modellizzazione

Per motivi di chiarezza, riportiamo la statistica di decisione data da

$$\mathcal{L}(x) = \max_{\delta'} Q(x) = \sqrt{(\langle x, q_0 \rangle)^2 + (\langle x, q_{\pi/2} \rangle)^2} \quad (5.1)$$

Come già notato nel capitolo 2, i prodotti scalari presenti nella formula 5.1, sono delle correlazioni che possono essere calcolate mediante trasformate discrete di Fourier [13]. Più precisamente, il calcolo di Q per fissati valori delle masse, è eseguito mediante i passi:

- Generazione delle componenti in fase ed in quadratura del template normalizzato;
- Aggiunta di zeri ai template, in modo che questi abbiano la stessa lunghezza dei dati;
- Determinazione della trasformata discreta di Fourier del template ottenuto al passo precedente;
- Esecuzione dei prodotti scalari della suddetta formula
- Determinazione della trasformata discreta di Fourier inversa dei prodotti scalari, in modo da ottenere le correlazioni nel dominio del tempo;

- Calcolo del massimo di Q rispetto a ciascun campione delle correlazioni, che è assunto essere il valore del massimo rispetto ad n_a , per il valore delle masse fissato.

Da ciò segue che la valutazione della funzione obiettivo, comporta l'esecuzione di due trasformate discrete di Fourier, per calcolare le correlazioni della formula precedente. Inoltre la stessa valutazione comprende la massimizzazione rispetto al parametro n_a , che richiede la ricerca del massimo tra i valori di Q calcolati per ogni campione delle correlazioni.

Occorre a questo punto notare che oltre all'algoritmo di Price, utilizzato per effettuare la massimizzazione rispetto alle masse. E' stato sviluppato anche un software per il calcolo della statistica di decisione, che rappresenta il massimo della funzione obiettivo. In particolare sono state utilizzate le seguenti librerie

- GSL - GNU Scientific Library per la generazione di numeri casuali[62]
- LAL - LSC Algorithm Library Una collezione di routine per l'analisi dei dati, per la rivelazione di onde gravitazionali;[63]
- FFTW , libreria per il calcolo della trasformata discreta di fourier[64]

5.2.2 Generazione dei dati e determinazione della soglia - Pianificazione

In questo paragrafo, parleremo del modo in cui sono stati generati i dati, per effettuare i test, e della determinazione della soglia. Innanzitutto diciamo che i dati utilizzati, derivano dal campionamento di una forma d'onda continua, che può essere costituita da rumore o da rumore più segnale chirp. Essi hanno in comune le seguenti caratteristiche:

- La frequenza di tagli inferiore $f_a = 40\text{Hz}$;
- La frequenza di taglio superiore $f_c = 2000\text{Hz}$;
- La frequenza di campionamento $f = 4048\text{Hz}$;
- I chirp sono generati al secondo ordine post-newotniano.

Nel caso in cui sono formati solo da rumore, vengono generati, nel dominio delle frequenze, mediante i seguenti passi:

- generazione di una sequenza di numeri casuali secondo la distribuzione gaussiana con media nulla e varianza unitaria;
- determinazione della trasformata discreta di fourier della sequenza precedente
- moltiplicazione della trasformazione discreta di fourier ottenuta, per lo spettro simulato di VIRGO.

Ovviamente i dati nel dominio del tempo (vedi figura 5.2) si ottengono eseguendo una trasformata dicreta di fourier inversa. Se invece i dati contengono sia il rumore che il segnale

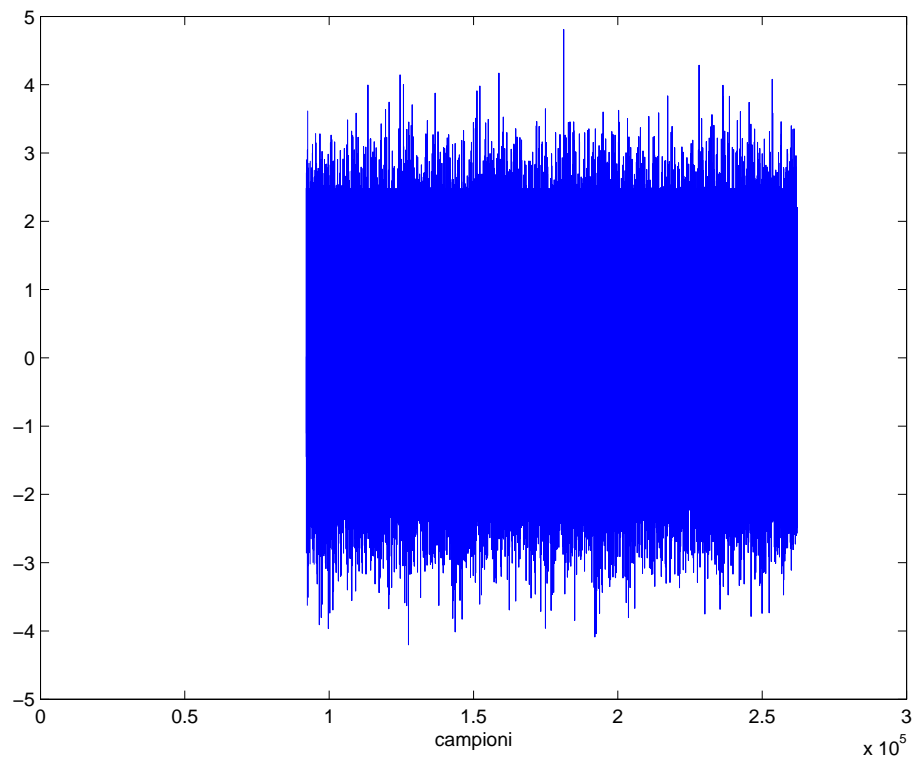


Figura 5.2: Realizzazione di rumore nel dominio del tempo con PSD di Virgo simulata

chirp, per ottenere un fissato rapporto segnale-rumore, dobbiamo eseguire i seguenti passi oltre alla generazione del rumore:

- Generazione del segnale chirp con assegnati parametri;
- Determinazione della trasformata discreta di Fourier del chirp;
- Normalizzazione del chirp;
- Prodotto del chirp normalizzato per il rapporto segnale/rumore fissato;
- Somma del rumore e del segnale chirp.

Un esempio di dati nel dominio del tempo, formati da segnale chirp, più rumore è rappresentato in figura 5.3. Il chirp iniettato nel rumore è stato evidenziato per facilitarne l'individuazione.

Per quanto riguarda la determinazione della soglia, bisogna osservare, che non essendo possibile avere un'espressione analitica delle prestazioni del ricevitore di segnali chirp in termini della probabilità di falso allarme e corretta rivelazione, il calcolo della soglia τ viene effettuato mediante simulazioni di Montecarlo.

In pratica, è stata calcolata la statistica di decisione \mathcal{L} con il nostro metodo di rivelazione un certo numero di volte con dati costruiti da diverse realizzazioni di rumore. L'insieme delle statistiche di decisione, ci ha poi permesso di ottenere una stima empirica della probabilità di falso allarme, mediante la formula.

$$P_f = \frac{N_\tau}{N} \quad (5.2)$$

dove N_τ è il numero di volte che la statistica di decisione ha superato una soglia stabilita τ , ed N è il numero di statistiche calcolate. Dopo di che la soglia è determinata in base alla probabilità di falso allarme richiesta. La figura 5.4 mostra l'andamento della probabilità di falso allarme, calcolata mediante la formula 5.2, in funzione della soglia scelta.

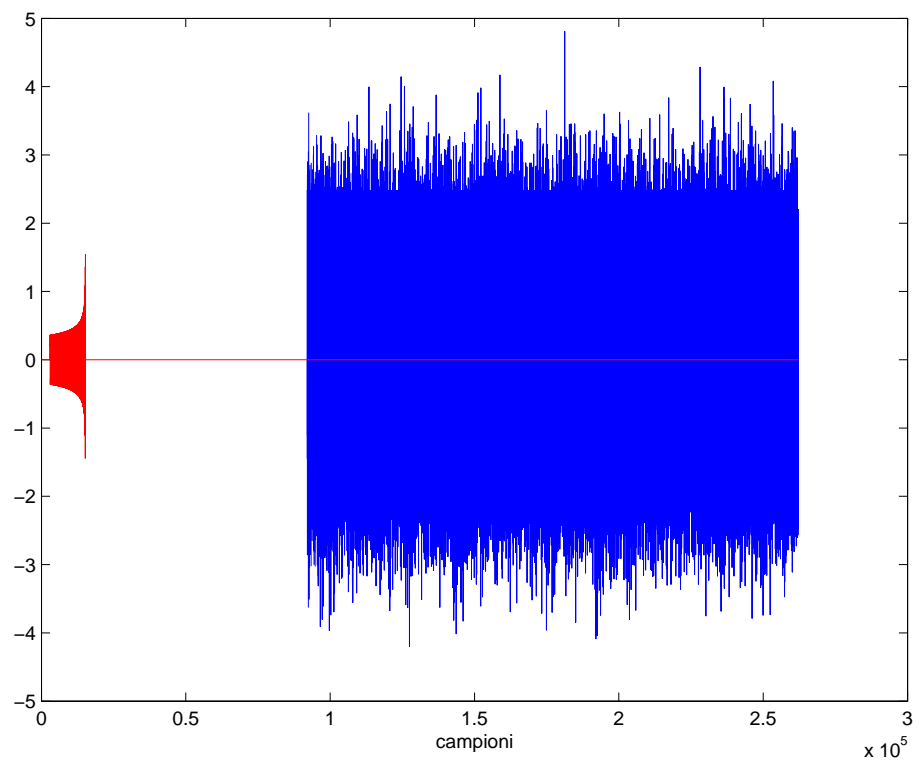


Figura 5.3: Rumore più segnale chirp nel dominio del tempo con PSD di Virgo simulata

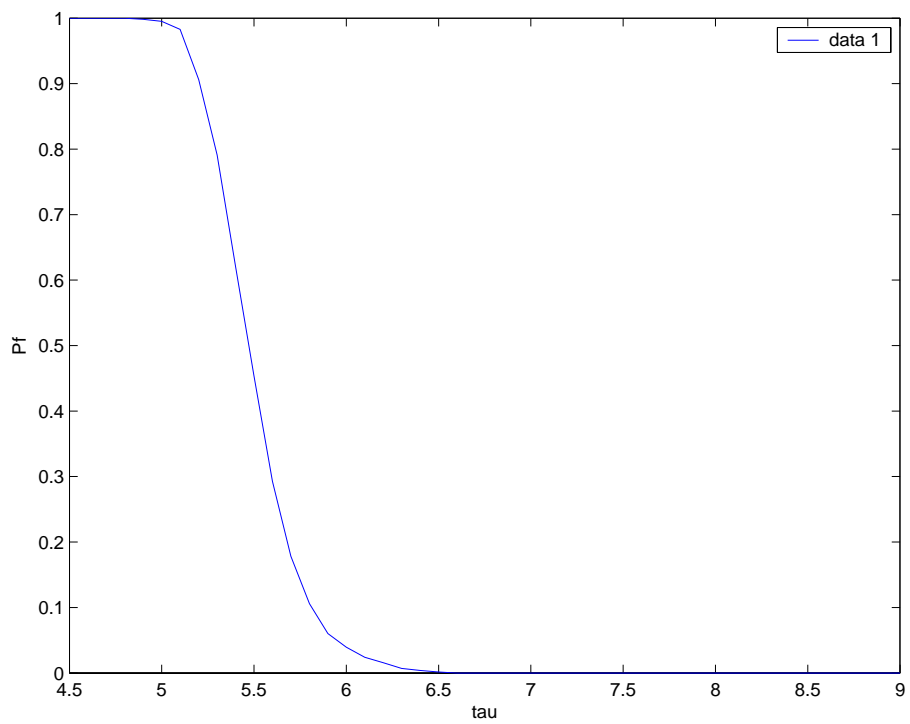


Figura 5.4: Probabilità di falso allarme empirica in funzione della soglia

Va quindi osservato che il numero di prove effettuato ci consente di avere una stima empirica della probabilità di falso allarme dell'ordine di 10^{-2} con un errore relativo di 0.01 per il 95% dei casi [12]. Questo ci permette di fissare la soglia a

$$\tau = 6.5 \tag{5.3}$$

che corrisponde ad una probabilità di falso allarme pari a 0.01.

5.2.3 Test e risultati - Esecuzione

Tutti i test sono stati svolti su 8 dei 20 nodi del Cluster Beowulf del laboratorio VIRGO di Napoli, la cui architettura è descritta ampiamente in [22][25][26]. Successivamente, il calcolo dei tempi di esecuzione è stato svolto sulla farm messa a disposizione dalla sede centrale dell'esperimento Virgo, equipaggiata con biprocessi AMD Opteron 2.1 Ghz con compilatore c, gcc 3.2.3.

Il significato dei parametri usati nei test è contenuto nei capitoli precedenti, è opportuno notare in questa sede che il massimo della funzione obiettivo non è altro che la stima del rapporto segnale-rumore espresso dal parametro S , quindi per poter parlare di rivelazione è necessario che questo massimo, sia superiore alla soglia fissata in 5.3.

5.2.4 Rivelazione Chirp per masse comprese tra 30 -5 M_{\odot}

Di seguito presentiamo i risultati ottenuti facendo eseguire il nostro algoritmo di rivelazione, su dati simulati in varie configurazioni con segnali generati di masse comprese tra 30 e 5 M_{\odot} .

I test sono stati svolti con i seguenti parametri per l'algoritmo di Price.

- 25 campioni iniziali
- 100 iterazioni
- dominio di ricerca $[1, 30] \times [1, 30]$

Il tuning dei parametri dell'algoritmo di Price è stato eseguito in maniera empirica constatando che nel caso di masse grandi, non si ottengono sostanziali miglioramenti di accuratezza aumentando la grandezza della matrice o il numero di iterazioni. Nel range di masse $30 - 5M_{\odot}$ sono state testate 5 realizzazioni di segnale più rumore di un minuto ciascuna, con chirp caratterizzati dalle seguenti configurazioni:

- Rumore + segnale 30 e $25M_{\odot}$
- Rumore + segnale 25 e $20M_{\odot}$
- Rumore + segnale 20 e $15M_{\odot}$
- Rumore + segnale 15 e $10M_{\odot}$
- Rumore + segnale 7 e $8M_{\odot}$

Per ogni una delle suddette configurazioni l'algoritmo di Price è stato eseguito 1000, volte ciascuna eseguita su un singolo nodo del cluster opteron con librerie compilate staticamente, ed utilizzando per ogni test un differente seme per l'algoritmo di generazione di numeri casuali.

In ciascuna delle prove svolte, il segnale viene correttamente rivelato con una buona precisione nell'individuazione dell'istante di arrivo (misurato in campioni). Tale rivelazione avviene con appena 125 valutazioni di funzione ed in tempi accettabili. L'algoritmo infatti, in queste batterie di test, risulta perfettamente in time con l'acquisizione dei dati, impiegando un tempo di esecuzione medio che va dai 23 ai 29 secondi con un in-time factor che varia da un minimo di 2 fino ad un massimo di 2.6 utilizzando una singola workstation. Per quanto riguarda la stima delle masse, come ci si aspettava, risulta piuttosto accurata nel caso delle masse grandi, dove la funzione obiettivo risulta più semplice da ottimizzare, perdendo tuttavia qualcosa in precisione, mano a mano che ci si avvicina verso l'estremo inferiore fissato a 5.

Nella tabella in figura 5.5 sono riassunti tutti i risultati dei test svolti, riportando per ognuno di essi i valori medi delle masse stimate, del tempo di arrivo (inteso come campione

di arrivo), del valore della funzione obiettivo e del tempo di esecuzione. Accanto ai valori medi sono riportate altresì le varianze empiriche, il cui andamento segnala un elevato grado di accuratezza nell'individuare l'istante di arrivo del segnale, parametro fondamentale per un algoritmo di rivelazione. Per ciascuno dei casi test, infine, sono stati altresì presentati 4 grafici che rappresentano in maniera chiara le prestazioni del rivelatore.

		MASSA 1	MASSA 2	Na	SNR	TEMPO
TEST1	DATI	30	25	3000	10	1min
	TEST	29.16 ± 1,32	24.93 ± 2.260	2970 ± 46	10.75 ± 0.22	25.0 ± 2.11
TEST2	DATI	25	20	3000	10	1min
	TEST	26.03 ± 2.28	19.65 ± 1.87	3002 ± 36	10.51 ± 0.12	26.4 ± 2.1
TEST3	DATI	20	15	3000	10	1min
	TEST	15.60 ± 1.61	19.64 ± 1.90	3013 ± 25	10.71 ± 0.03	23.7 ± 1.90
TEST4	DATI	15	10	3000	10	1min
	TEST	16.57 ± 2.90	9.48 ± 1.89	3018 ± 28	10.11 ± 0.07	29.8 ± 3.10
TEST5	DATI	7	8	3000	10	1min
	TEST	5.93 ± 3.480	9.31 ± 4.78	2978 ± 185	10.33 ± 1.30	25.6 ± 2.01

Figura 5.5: Riassunto dei test svolti

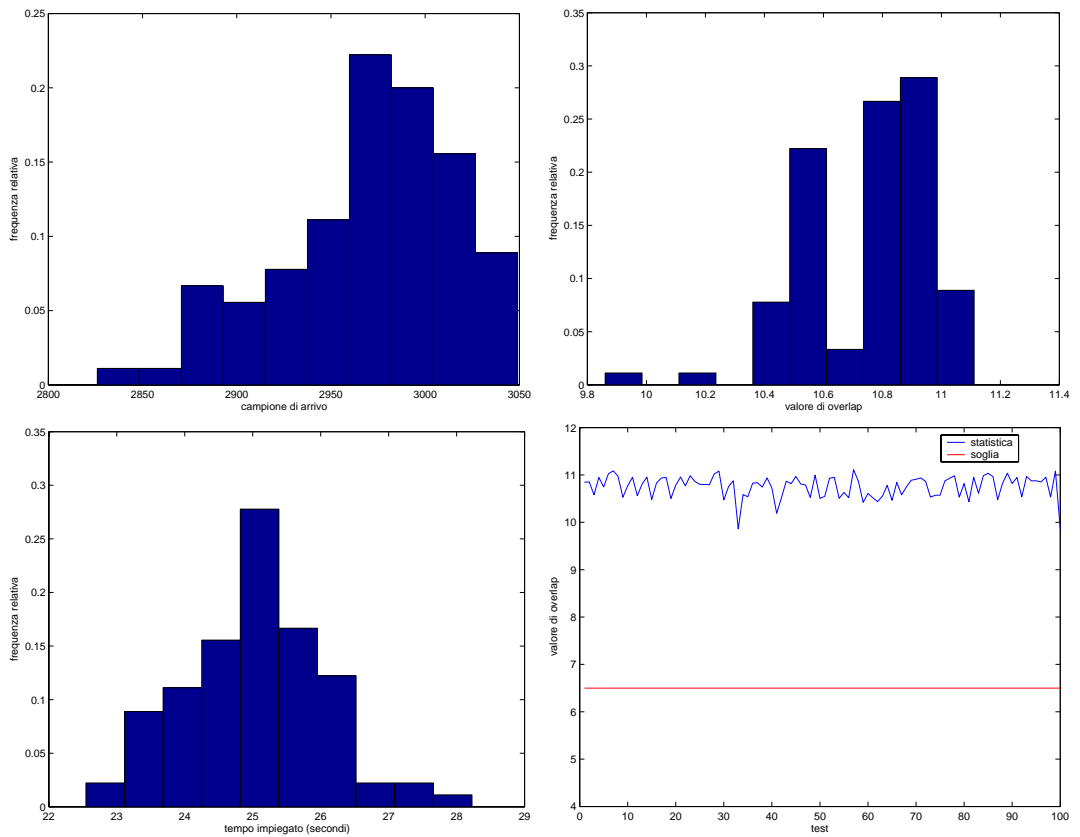


Figura 5.6: Risultati Test 1 - Masse 25 e $30M_{\odot}$ - In senso orario - La prima in alto a sinistra mostra l'istogramma del campione d'arrivo rilevato - il secondo grafico i valori di overlap verificati, il terzo grafico confronta il valore di overlap stimato rispetto alla soglia sulla finestra dei primi 100 test campione, l'ultima figura in basso a sinistra rappresenta il tempo l'istogramma (normalizzato) dei tempi di esecuzione.

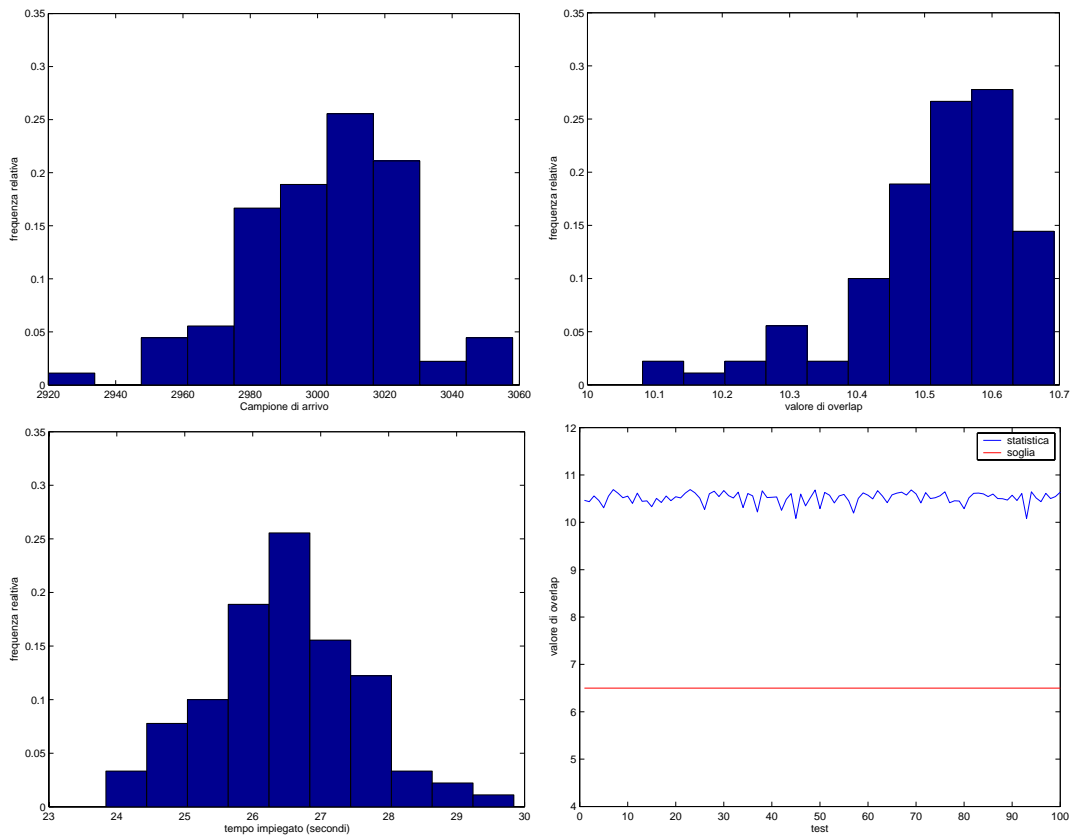


Figura 5.7: Risultati Test 2 - Masse 20 e $25M_{\odot}$ - In senso orario - La prima in alto a sinistra mostra l'istogramma del campione d'arrivo rilevato - il secondo grafico i valori di overlap verificati, il terzo grafico confronta il valore di overlap stimato rispetto alla soglia sulla finestra dei primi 100 test campione, l'ultima figura in basso a sinistra rappresenta il tempo l'istogramma (normalizzato) dei tempi di esecuzione.

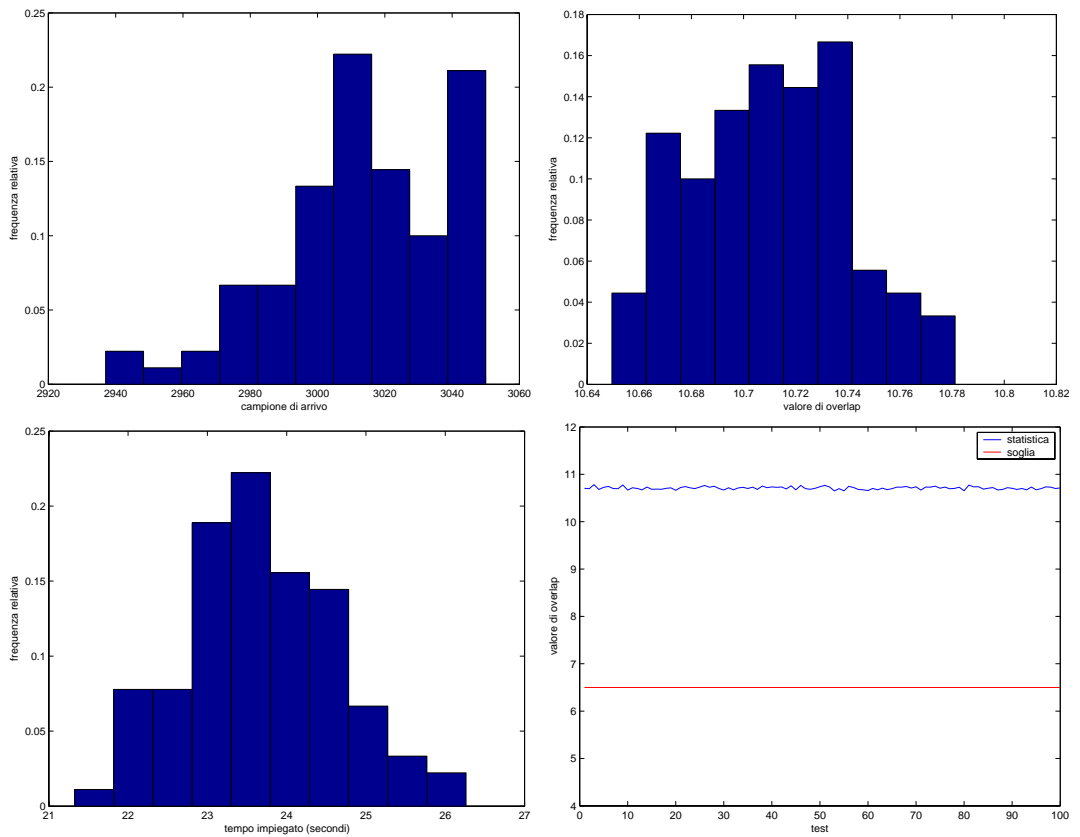


Figura 5.8: Risultati Test N 3 - Masse 15 e $20M_{\odot}$ - In senso orario - La prima in alto a sinistra mostra l'istogramma del campione d'arrivo rilevato - il secondo grafico i valori di overlap verificati, il terzo grafico confronta il valore di overlap stimato rispetto alla soglia sulla finestra dei primi 100 test campione, l'ultima figura in basso a sinistra rappresenta il tempo l'istogramma (normalizzato) dei tempi di esecuzione.

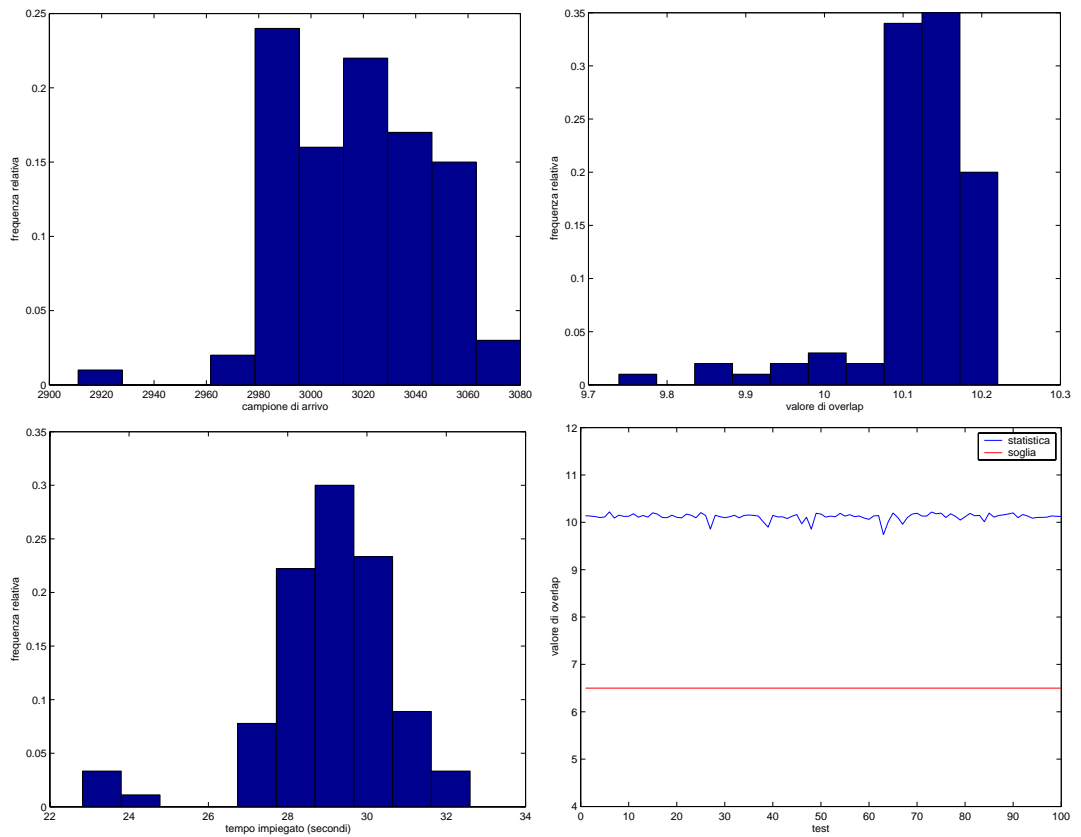


Figura 5.9: Risultati Test N 4 - Masse 10 e $15M_{\odot}$ - In senso orario - La prima in alto a sinistra mostra l'istogramma del campione d'arrivo rilevato - il secondo grafico i valori di overlap verificati, il terzo grafico confronta il valore di overlap stimato rispetto alla soglia sulla finestra dei primi 100 test campione, l'ultima figura in basso a sinistra rappresenta il tempo l'istogramma (normalizzato) dei tempi di esecuzione.

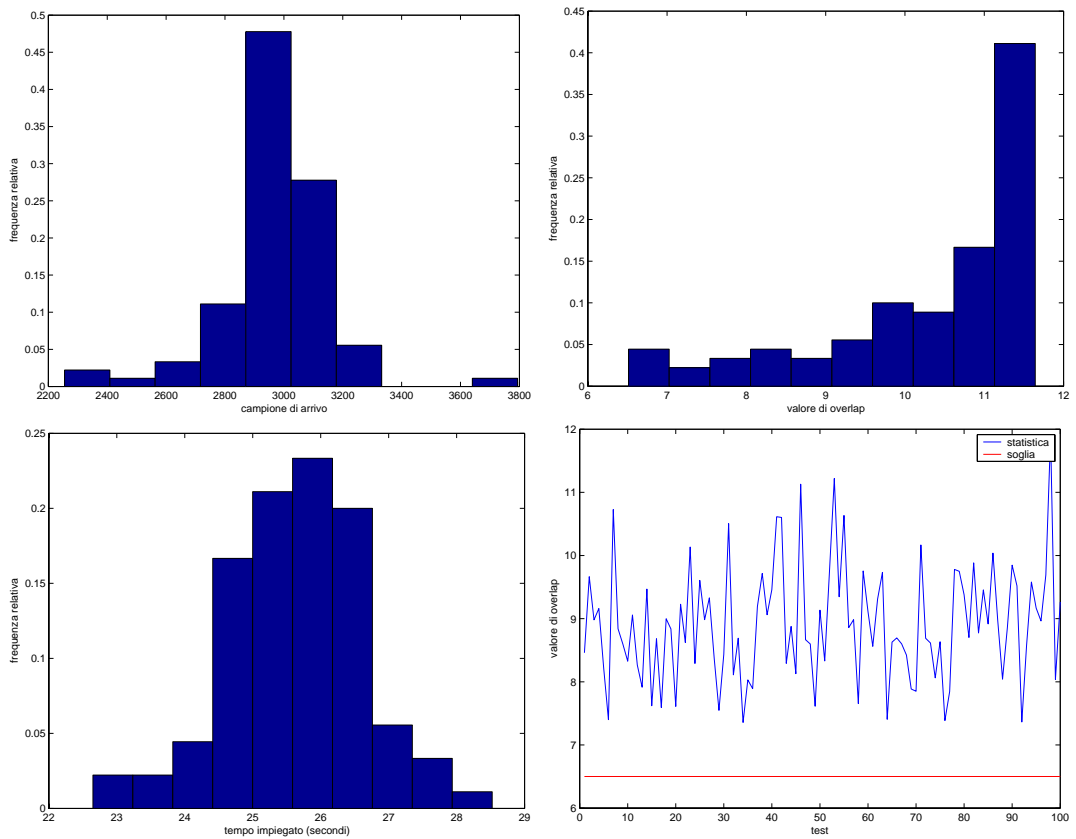


Figura 5.10: Risultati Test N 5 - Masse 7 e $8M_{\odot}$ - In senso orario - La prima in alto a sinistra mostra l'istogramma del campione d'arrivo rilevato - il secondo grafico i valori di overlap verificati, il terzo grafico confronta il valore di overlap stimato rispetto alla soglia sulla finestra dei primi 100 test campione, l'ultima figura in basso a sinistra rappresenta il tempo l'istogramma (normalizzato) dei tempi di esecuzione.

Dopo i test su masse campione, scelte come rappresentanti dei diversi reange, è stata svolta un'ulteriore batteria di 9000 test, con vettore dei dati di input generato nuovamente ad ogni prova e contenete un segnale chirp generato da masse scelte in maniera random, uniformemente nello spazio $[5, 30] \times [5, 30]$. Dai risultati ottenuti, notiamo che l'algoritmo di price riesce a rispondere bene anche a questa sorta di stress test riuscendo a rilevare il segnale nel 100% con una buona precisione nell'individuazione dell'istante di arrivo del segnale (posto sempre al 3000-esimo campione). Nei grafici rappresentati in figura 5.11, possiamo apprezzare i risultati delle statistiche svolte in analogia ai test precedenti e che mostrano le buone proprietà di rivelazione dell'algoritmo nel caso di masse grandi.

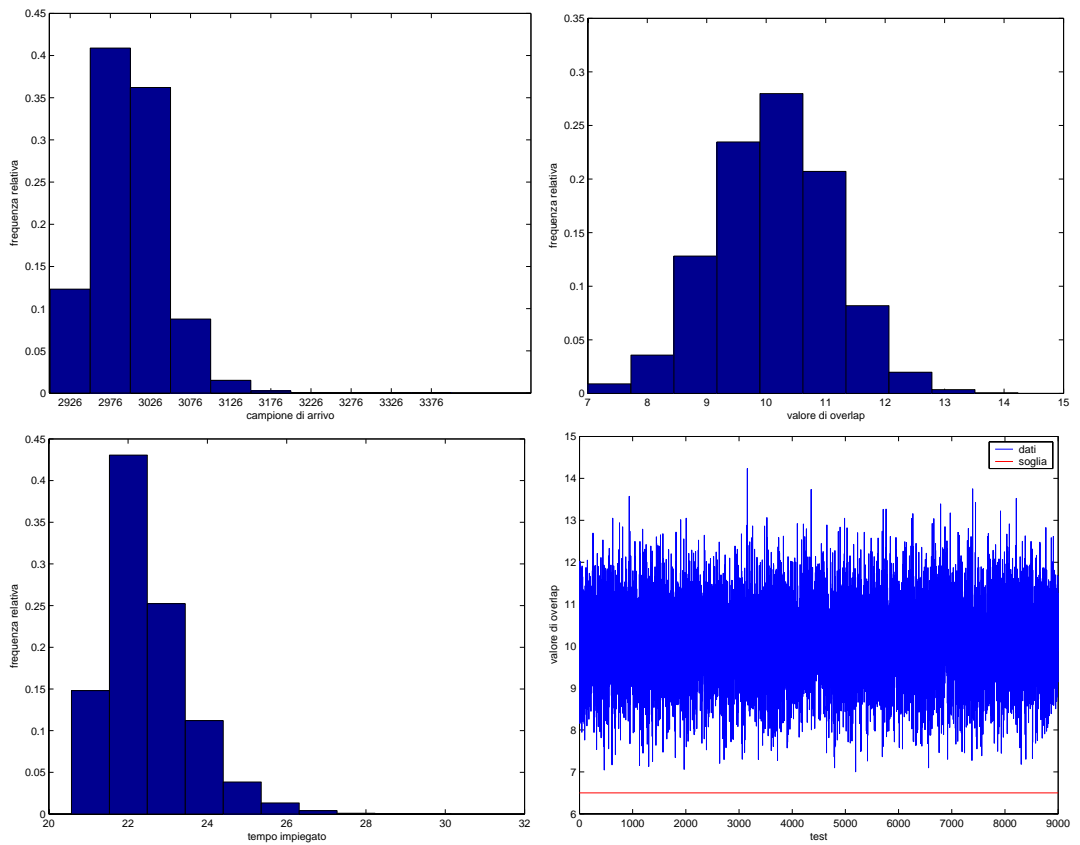


Figura 5.11: Risultati Stress Test con masse che variano in $[5, 30] \times [5, 30]$ - In senso orario - La prima in alto a sinistra mostra l'istogramma del campione d'arrivo rilevato - il secondo grafico i valori di overlap verificati, il terzo grafico confronta il valore di overlap stimato rispetto alla soglia su tutti i 9000 campioni, l'ultima figura in basso a sinistra rappresenta il tempo l'istogramma (normalizzato) dei tempi di esecuzione.

5.2.5 Rivelazione Chirp per masse minori di 5

Nel caso in cui le masse diventano molto piccole sotto la soglia di 5, l'algoritmo di Price, con i parametri finora usati, fallisce non riuscendo a incanalarsi verso il valore massimo, raggiungendo molto sporadicamente valori di overlap sopra la soglia (vedi figura 5.12). In questi tipi di configurazioni, infatti, la funzione obiettivo presenta moltissimi massimi locali, questo rende il problema dell'ottimizzazione globale molto ostico da affrontare. Per rivelare il segnale in queste circostanze, è dunque necessario variare i parametri dell'algoritmo.

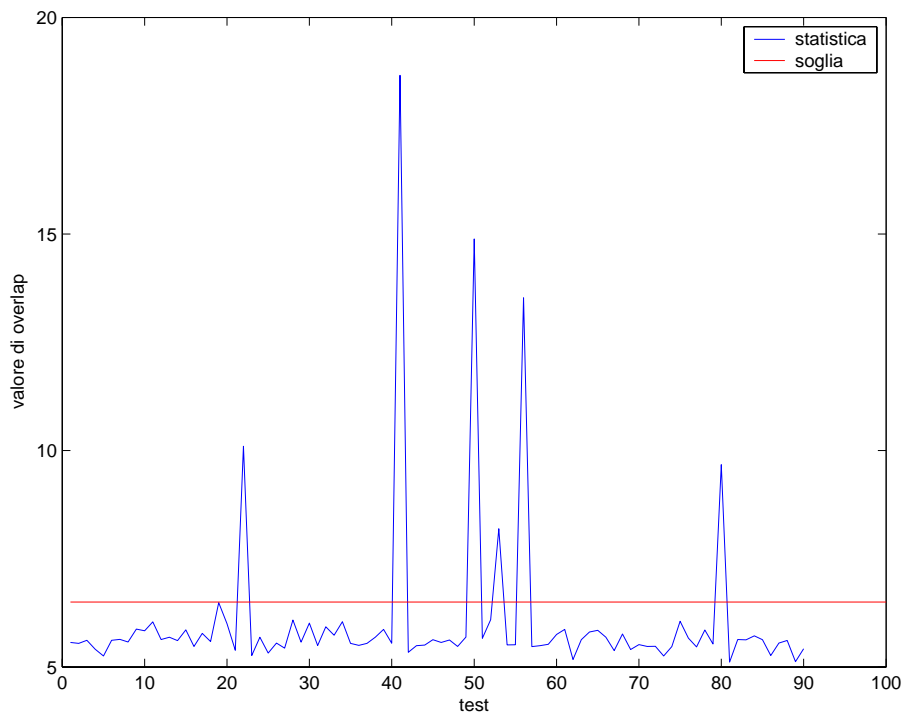


Figura 5.12: Risultati Test con parametri di Price inefficienti

Restrizione del dominio

Per aumentare le prestazioni del nostro ricevitore sono percorribili diverse strade che conducono comunque ad una qualche sorta di parallelizzazione. La soluzione più immediata, consiste nel restringere il dominio di ricerca della funzione da $[1, 30] \times [1, 30]M_{\odot}$ a

$[1, 10] \times [1, 10]M_{\odot}$. La parallelizzazione, secondo questa strategia, consiste nel far girare due istanze dell'algoritmo di Price, una che abbraccia l'intero dominio, l'altra che lavora specificatamente sul dominio ristretto.

Per dimostrare la fattibilità di questo approccio è necessario tuttavia verificare sperimentalmente, che mantenendo il dominio esteso a $[1, 30] \times [1, 30]M_{\odot}$, si riescono ad individuare segnali prodotti da coppie di masse appartenenti a $[10, 30] \times [1, 5]$. Sono stati quindi simulati 1000 test per l'esecuzione dell'algoritmo di Price su dati di input contenenti segnali chirp generati da coppie di masse del tipo (m_1, m_2) con m_1 ed m_2 scelti in maniera random rispettivamente nei range $[10, 30]$ e $[1, 5]$. I risultati ottenuti sono mostrati nei grafici in figura 5.13, dove possiamo apprezzare ancora una volta la capacità di individuare il segnale, con una leggera perdita di accuratezza, rispetto al caso delle masse grandi, nell'individuazione del campione d'arrivo.

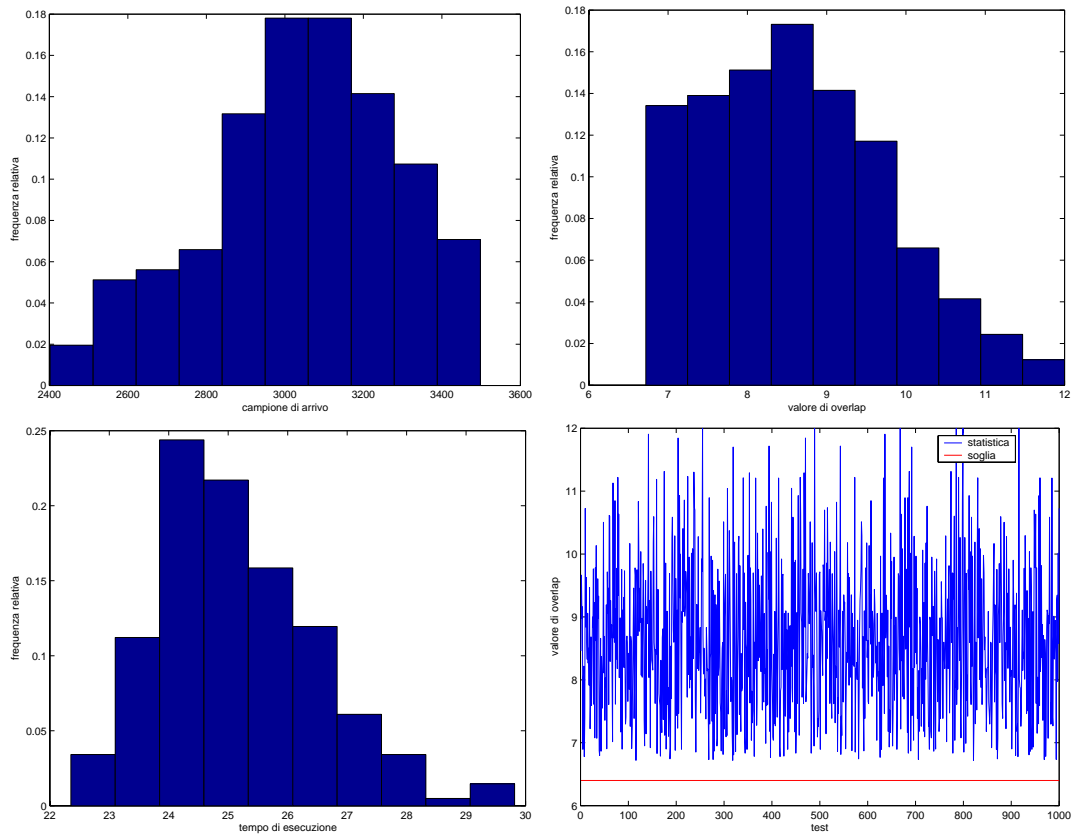


Figura 5.13: Risultati - ricerca di segnali chirp prodotti da coppie di masse (m_1, m_2) con $m_1 \in [10, 30]$ ed $m_2 \in [1, 5]$ e dominio di ricerca $[1, 30] \times [1, 30]M_\odot$ - 1000 test campione

I risultati dei test svolti restringendo il dominio, sono riportati nella tabella in figura 5.14 e nei grafici seguenti, essi mostrano come l'algoritmo risponda bene riuscendo a rivelare il segnale nella maggior parte dei casi, e mantenendo comunque una buona precisione nella determinazione del campione di arrivo. I tempi, tuttavia, si allargano leggermente mantenendosi tuttavia al di sotto dei 60 secondi, che garantisce la possibilità di rimanere in-time.

		MASSA 1	MASSA 2	Na	SNR	TEMPO
TEST6	DATI	2	2	3000	10	1min
	TEST	2.71 ± 0.7	1.66 ± 0.6	2779 ± 200	8.1 ± 1.84	55.01 ± 3.11

Figura 5.14: Risultati Test con restrizione di dominio ad $[1, 10] \times [1, 10]$

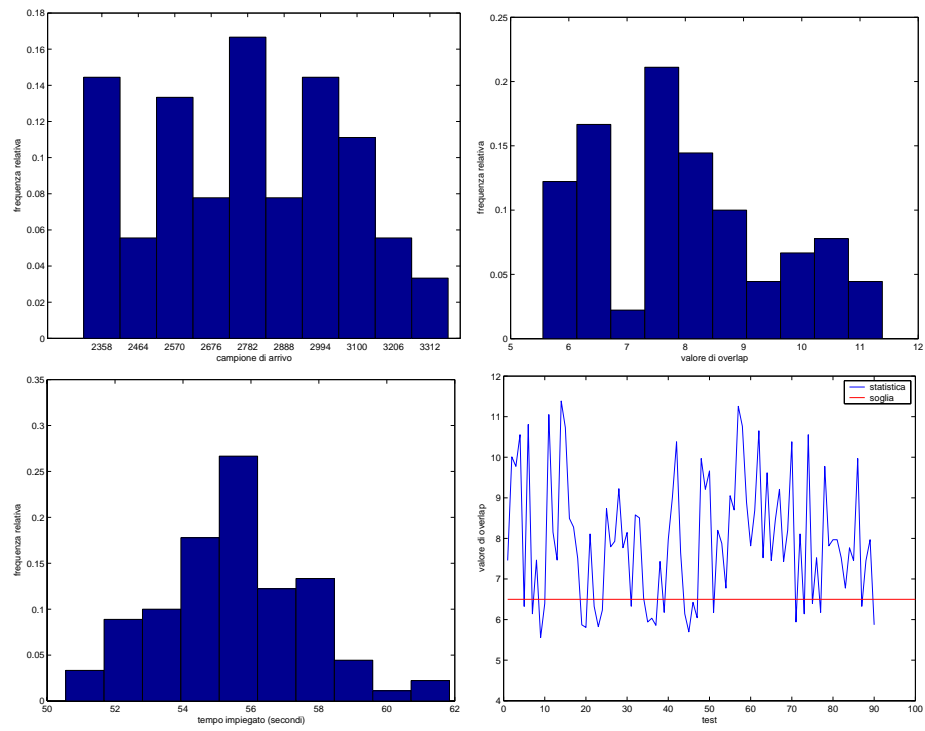


Figura 5.15: Risultati Test - Masse 2 2

Approccio Parallelo

Un'altra possibile strategia per aumentare le prestazioni del nostro ricevitore nel caso delle masse piccole, è quella di variare i parametri dell'algoritmo di Price. Per testare questa soluzione è stata dunque svolta un'ulteriore batteria di test per la rivelazione di un segnale chirp generato da due stelle con masse 2 e 2, utilizzando una matrice di ricerca composta da 100 righe e con un numero massimo di iterazioni pari a 3000. Al fine di ridurre al minimo il numero di iterazioni è stata altresì imposta una condizione di arresto aggiuntiva, che consente l'uscita dal ciclo di iterazioni quando il valore di overlap massimo contenuto nella matrice di ricerca supera la soglia fissa a 6.5.

Con questa configurazione, sono stati eseguiti 1000 test con diverse realizzazioni di rumore e differente seme per l'algoritmo di *Price*. Come mostrato dal grafico in figura ??, il rivelatore si mantiene efficiente nel rivelare il segnale, perdendo tuttavia qualcosa in precisione nel determinare il campione di arrivo del segnale rispetto al caso dei segnali generati da stelle con masse comprese tra 30 – 5. I tempi di esecuzione, ovviamente si dilatano e vanno da un minimo di 58 secondi fino a un massimo di 11 minuti con una media di 6 minuti per test. Questo abbassa notevolmente l'in-time factor precludendo la possibilità di restare in-time con l'acquisizione, utilizzando una singola workstation. Una possibile soluzione è dunque quello di utilizzare un approccio parallelo, questa volta sulla computazione. Attualmente, l'autore di questo lavoro di tesi, è impegnato nello studio di una possibile applicazione in questo senso, per approfondimenti si rimanda all'appendice B.

5.3 Costi computazionali

Per valutare i costi computazionali richiesti dall'algoritmo di Price, notiamo che le operazioni predominanti sono il calcolo della FFT dei segnali template corrispondenti alle coppie di masse campionate, così come nel caso della strategia su banche di template.

In particolare l'algoritmo esegue una sola volta la FFT dei dati in ingresso, la FFT del

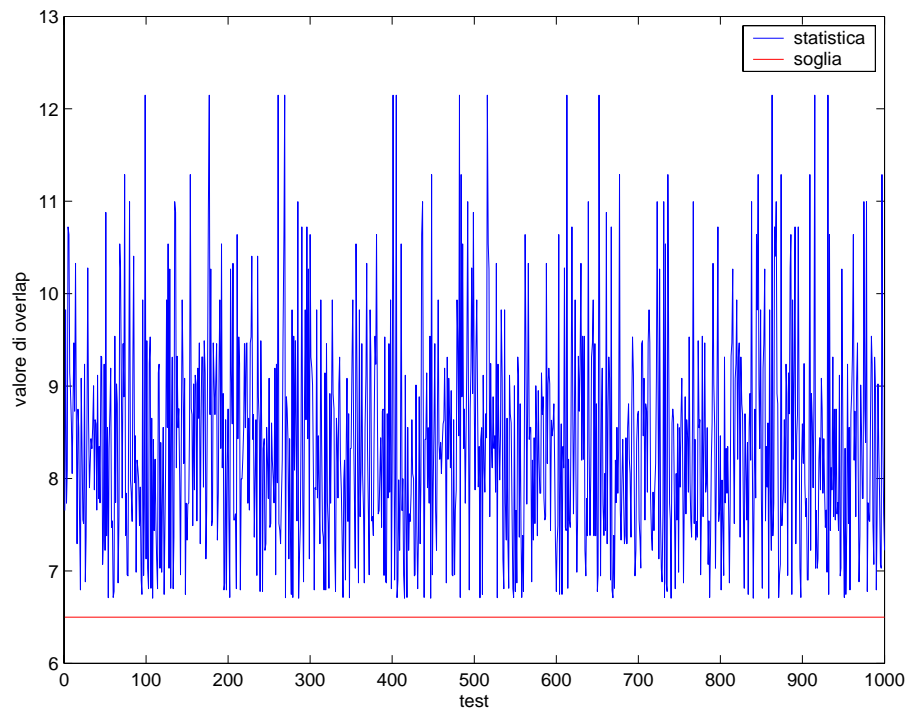


Figura 5.16: Risultati Test - Valori della statistica di decisione per la ricerca di segnali chirp prodotti da coppie di masse appartenenti al reange $[1, 5] \times [1, 5]M_{\odot}$

template da correlare man mano che viene campionato un punto nello spazio dei parametri, quindi viene eseguita l'anti-FFT per ottenere il punto in cui è stato individuato il massimo valore di overlap, nel dominio del tempo.

Per l'algoritmo di Price, il numero effettivo di FFT da calcolare è legato al numero di iterazione ed alle dimensioni della matrice di ricerca. Partendo dalla relazione 4.2 possiamo dire che il costo computazionale necessario all'algoritmo di Price settato con matrice di ricerca di N righe e con un limite di iterazioni pari ad I , è dato da:

$$\mathcal{P} = (2f_u) \times \mathcal{N} + \mathcal{I} \times [36(\ln_2(2f_u) + \langle \ln_2(\tau(f_l)) \rangle) + 54] \text{flop/sec} \quad (5.4)$$

In questo lavoro di tesi, i parametri scelti per l'algoritmo di Price sono 25 righe per la matrice di ricerca e 100 come numero di iterazioni. Riferendoci alla tabella 4.2 per l'analisi dei costi con griglie di template, utilizzando i stessi parametri di f_u, f_l , abbiamo dalla formula [?] un costo computazionale pari a $4 * 10^7$ flop/sec. Confrontando tale valore con quello ottenuto nel caso di griglie di template ($1.1 * 10^{10}$), abbiamo una abbattimento dei costi pari a 3 ordini di grandezza che ci consente di eseguire l'analisi on-line con semplici workstation.

Capitolo 6

Filtro adattivo ALE

Un altro approccio per risolvere il problema della analisi online è quello di ricercare le componenti coerenti immerse nel rumore, sfruttando le potenzialità dei filtri adattivi. In questo capitolo verrà definito il filtro IIR ALE, successivamente verranno presentati una serie di test su rumore bianco per testarne il funzionamento e dimostrarne l'effettiva applicabilità al caso delle rivelazione di onde gravitazionali.

6.1 Introduzione

Per affrontare gli elevati costi computazionali, è stato altresì esplorato un ulteriore approccio, ugualmente adattivo ma alternativo al DMFT. Nell'ambito di questo lavoro di tesi viene proposto l'utilizzo di un filtro adattivo IIR ALE[29] da usare come trigger di primo livello per selezionare i blocchi di dati interessanti che potranno essere poi analizzati con maggiore accuratezza con la tecnica del matched filter su griglia di template.

L'obbiettivo principale di questo approccio è quello di eseguire una analisi snella che mantenga le seguenti caratteristiche: bassa perdita di segnali rispetto all'uso del matched filters, robustezza nei confronti dei falsi allarme e bassi costi computazionali per una esecuzione in real-time.

Infine considerando che tale algoritmo non richiede la conoscenza dei segnali template,

può essere usato effettivamente non solo per la fase inspiral ma anche per riconoscere i processi di coalescenza. In altre parole, se vogliamo implementare un trigger che lavori in real-time, algoritmi come IIR ALE, come vedremo nel corso di questo capitolo, risultano molto utili.

L'idea principale è quella di utilizzare un filtro adattivo *IIR ALE*, per ricostruire all'uscita le componenti coerenti (se ve ne sono) presenti nell'input. Se il filtro riesce a tracciare correttamente il segnale, la ricostruzione consiste in una replica rumorosa delle componenti coerenti presenti nell'input. Se vediamo la ricostruzione come un *template modificato dal rumore*, allora possiamo creare un rivelatore di correlazione che esegue il match tra i dati in ingresso al filtro, ed il suo output.

Bisogna osservare che il filtro deve essere necessariamente adattivo, al fine di consentire la ricostruzione di segnali con differenti componenti in frequenza. Quindi il correlatore integra le informazioni locali ottenute da un filtro adattivo passo passo.

6.2 Modello del segnale

Per stabilire il modello del segnale atteso, da usare e le simulazioni, richiamiamo i risultati su studi teorici sulla fase finale di sistemi binari costituiti sia da stelle di neutroni che da due buchi neri, approfondendo quanto detto nel capitolo primo. L'investigazione teorica ha permesso di dividere l'eventi di coalescenza in tre fasi distinte:

- Spiraleggio lento (o fase di Inspiral);
- Merger;
- Ringdown;

Durante la fase di spiraleggio lento, la distanza tra i due oggetti compatti componenti il sistema binario è sufficientemente grande da poter considerare trascurabili le loro dimensioni (ipotesi del punto massa). Verso la fine della fase di spiraleggio, il sistema subisce un'instabilità dinamica, passando da un regime di radiazione ad una caduta libera tra le

due componenti, fino alla finale fusione (fase di merger). Questa fase, attualmente, costituisce una delle principali sfide per la relatività numerica e per gli sperimentali, dovuta alla mancanza di una forma accreditata da usare come template per il filtro di Wiener.

La fase non lineare del merging, gradualmente raggiunge un regime stazionario che può essere descritto come oscillazioni del nuovo oggetto astrofisico compatto (Buco nero o stella di neutrone). Le onde gravitazionali emesse, sono descritte da sovrapposizione di sinusoidi smorzate esponenzialmente e trasportano informazioni riguardo le masse e gli spin (vedi [66][67][65])del nuovo oggetto compatto (fase di ring-down).

Da questo quadro di evoluzione, risulta chiaro che le fasi per le quali esiste ad oggi una accreditata forma del segnale, sono la fase inspirale e quella ring-down, mentre per quella di merger occorrono ancora importanti sforzi di natura teorica per fornirne il corretto template.

In figura 6.1 sono mostrati alcuni esempi di forme d'onda (i così detti chirp) durante la fase inspiral, prodotti da sistemi binari di stelle di neutroni o da buchi neri nel loro processo di coalescenza.

Nella nostra analisi considereremo la fase inspiral supponendo per semplicità che il segnale sia localmente sinusoidale, modulato in ampiezza ed in frequenza. Il chirp è per l'appunto un esempio di tale segnale.

$$s(t) = a(t) \cos \psi(t) \tag{6.1}$$

dove $a(t)$ e $\psi(t)$ sono funzioni monotone lentamente crescenti.

In seguito considereremo segnali a tempo discreto addizionato a rumore bianco gaussiano:

$$x_n = s_n + u_n \tag{6.2}$$

dove il reale rumore gaussiano u_n può essere ottenuto attraverso un operazione di sbiancamento sull'uscita dell'interferometro, mentre il segnale s_n è un chirp sbiancato della forma 6.1.

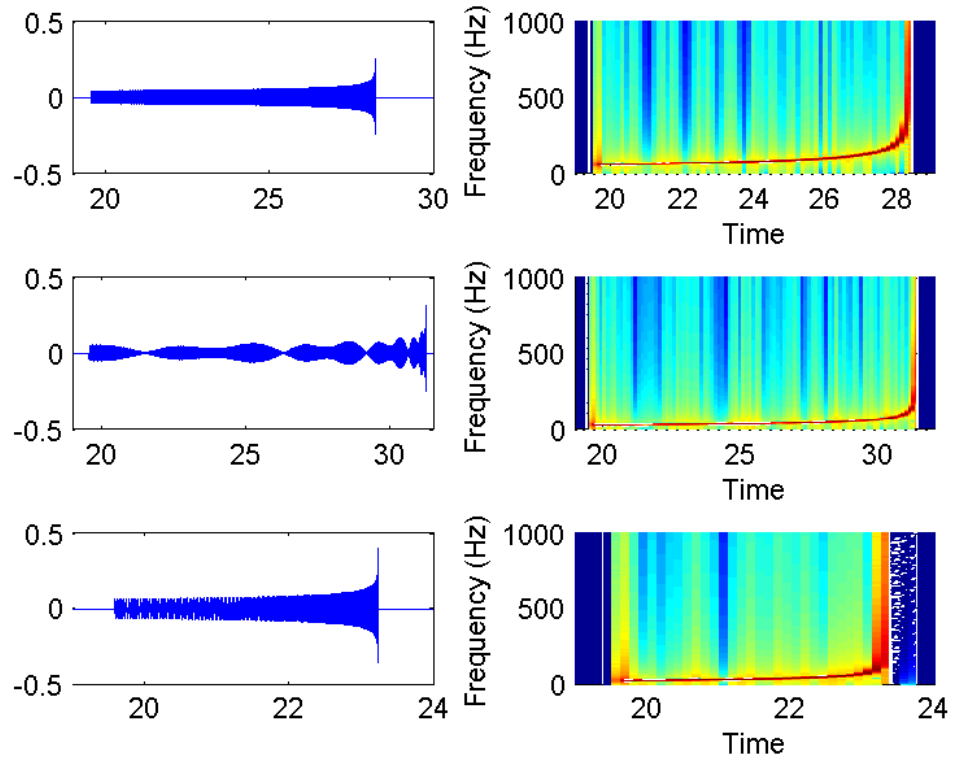


Figura 6.1: Esempi di chirps con approssimazione Post-Newtoniana di ordine 2 con frequenza di campionamento pari a $f_s = 4096Hz$. Partendo dall'alto sulla sinistra la forma d'onda sulla destra lo spettrogramma.

- 1: Due stelle di neutroni di masse $1.4M_\odot + 1.4M_\odot$; frequenza di taglio inferiore $60Hz$.
- 2: Stella di Neutroni più buco nero di masse $1.4M_\odot + 10M_\odot$ con Spin = 0.99; frequenza di taglio inferiore $30Hz$.
- 3: Effective One Body approach (EOB) di due buchi neri $5M_\odot + 10M_\odot$; frequenza di taglio inferiore $30Hz$.

6.3 Infinite Impulse Response Adaptive Line Enhancer (IIR-ALE)

Il modello di filtro adattivo indicato con Adaptive Line Enhancer filter model [76] è stato creato per approssimare il guadagno di rapporto segnale-rumore (SNR) ottenuto dalla strategia matched filter su segnali localmente sinusoidali. Il vantaggio di usare un filtro ALE è che a priori non è richiesta la conoscenza dei parametri del segnale (frequenza, ampiezza, fase). I filtri passa banda centrati in frequenza, possono essere usati quando il segnale da rilevare è a banda stretta ed immerso in rumore a banda larga. Lo svantaggio di tale filtro è che esso è capace solo di tracciare segnali localmente sinusoidali, quindi se tale assunzione non è verificata, le performance di rivelazione degradano notevolmente. Dobbiamo anche considerare problemi di pull-in time, cioè il tempo necessario al filtro per tracciare effettivamente il segnale e tunare la sua frequenza istantanea. In ambienti con basso rapporto segnale rumore, questo limita la possibilità di tracciare segnali molto brevi. Tale problema può essere usualmente aggirato eseguendo un banco di filtri ricoprendo una banda di frequenze.

La forma più nota di ALE è basata su una struttura a Finite Impulse Response (FIR), che presenta il vantaggio di essere stabile, e facile da adattare. Tuttavia esso richiede un gran numero di pesi adattivi, per fornire un adeguato riconoscimento di segnali a banda stretta, con conseguente aumento dei costi computazionali[82]. Un filtro adattivo di tipo Adaptive Infinite Impulse Response (IIR), invece, lavora tipicamente a banda stretta con un minor numero di pesi, ciò comporta un basso costo computazionale. Tuttavia ci si sono delle cause di instabilità. Questi problemi vengono superati imponendo dei vincoli sul filtro.

In questo lavoro di testi implementiamo un filtro ALE come un filtro IIR vincolato a secondo ordine con un singolo parametro adattivo. Il generale la struttura di ALE è quella nella parte sinistra in figura 6.2. L'obiettivo dell'elemento di ritardo è di decorrelare il rumore del input in input al filtro, lasciando invariata la risposta del filtro, per quanto è

possibile, sulla eccitazione sinusoidale. Questa versione modificata dal rumore della sinusoide, può essere recuperata all'uscita del filtro. Consideriamo che il filtro abbia la seguente forma che è una riparametrizzazione della struttura in [77] (per convenienza useremo gli stessi simboli per i nuovi parametri, w è il parametro da adattare mentre r è un parametro fissato determinato dalla banda del filtro):

$$H_1(z) = \frac{(1-r)(w-z^{-1})}{1-(1+r)wz^{-1}+rz^{-2}} \quad (6.3)$$

con i seguenti vincoli:

$$0 \ll r < 1, \quad -1 < w < 1 . \quad (6.4)$$

Con questa struttura del filtro, il processamento delle sequenze di input viene svolto dal filtro:

$$H(z) = z^{-1}H_1(z) = \frac{(1-r)(wz^{-1}-z^{-2})}{1-(1+r)wz^{-1}+rz^{-2}} . \quad (6.5)$$

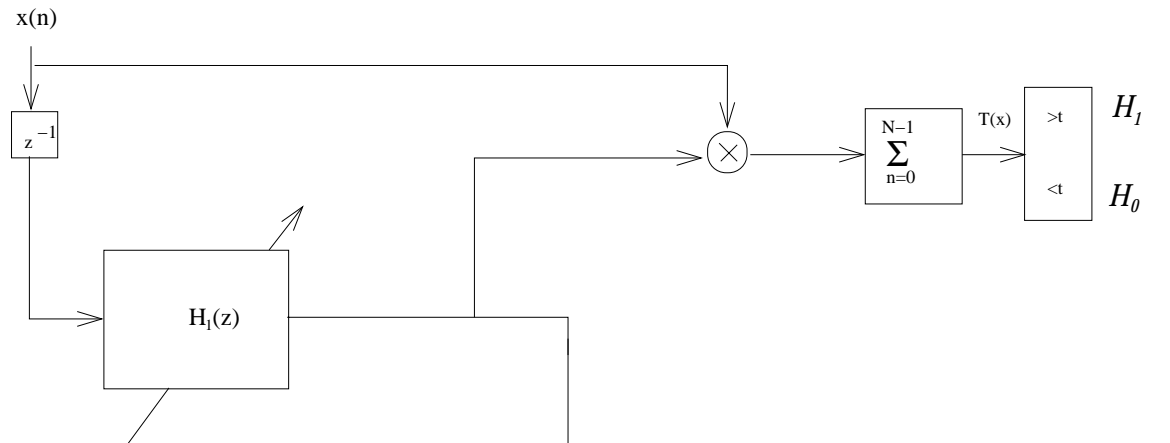


Figura 6.2: Schema del rivelatore adattivo. Nella parte sinistra della figura è mostrato uno schema del filtro ALE. Nella parte destra, la componente di rivelazione composta da un correlatore il cui output è comparato con una threshold t per scegliere se il segnale è presente o meno.

L'elemento di ritardo viene inserito per creare la struttura stessa del correlatore adattivo. Poiché il filtro $H_1(z)$ esegue una predizione un step avanti al relativo input, dobbiamo ritardare il segnale sorgente $x(n)$, così che possiamo comparare l'input e l'output $H_1(z)$ allo stesso istante di tempo per valutare il loro prodotto.

Il filtro sopra descritto, condivide alcune proprietà analitiche con i noise cancelers adattivi del secondo ordine (in particolare i poli, la banda e la locazione dei picchi) di forma generale:

$$D(z) = \frac{G(z)}{1 - 2Az^{-1} \cos \omega_0 + Bz^{-2}} \quad (6.6)$$

che hanno le seguenti caratteristiche[82]

- i poli formano una coppia di frequenze angolari ω_0 complessi coniugati, che si trovano su un cerchio di raggio approssimato $A \approx 1$ all'interno del cerchio unitario
- la banda 3-dB approssimata è $(1 - B)$;
- la magnitude response ha un picco alla frequenza angolare di ω_0

Queste caratteristiche nel caso del filtro $H(z)$, possono essere lette come interpretate nel seguente modo:

- I poli del filtro formano una coppia complessi-coniugati sopra una circonferenza di raggio approssimativamente uguale a $(1 + r)/2$ all'interno del cerchio unitario.
- la banda 3-bB approssimata è $(1 - r)$;
- la magnitude response ha un picco alla frequenza angolare:

$$\omega_0 = \arccos w \ . \quad (6.7)$$

- Inoltre, alla frequenza angolare di cui sopra, il filtro ha un guadagno unitario ed uno shift di fase nullo, che garantisce che la ampiezza locale e la fase della sinusoide di input a frequenza angolare ω_0 vengano lasciati invariati dal filtro.

In figura 6.3 mostriamo un esempio di risposta in fase ed in ampiezza del filtro $H(z)$.

L'introduzione di un meccanismo adattivo sul parametro w ha l'effetto di muovere i picchi di risposta di $H(z)$ così da tracciare la frequenza istantanea della sinusoide in input. Il parametro r è usualmente scelto vicino ad 1 per avere una banda stretta.

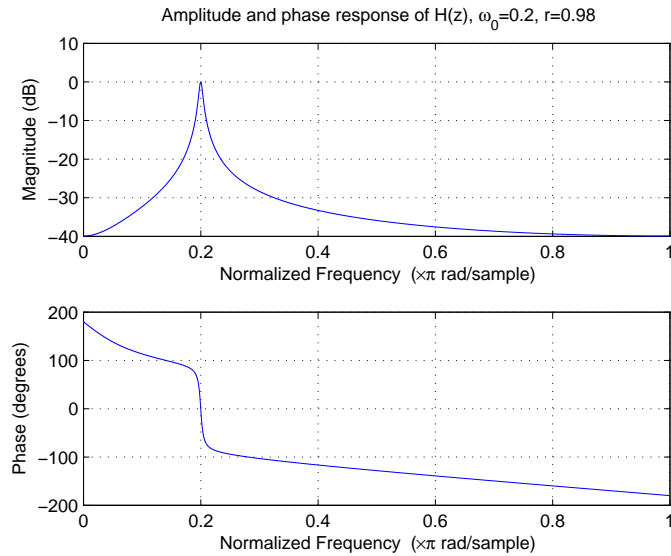


Figura 6.3: Risposta in fase ed in ampiezza di un esempio di filtro ALE. Il ritardo di fase a ω_0 è zero, e la risposta in ampiezza ha un picco a ω_0 .

6.3.1 Adattamento dei parametri

La formulazione di errore di uscita è abbastanza naturale nelle applicazioni di filtraggio [76]; tuttavia, nel caso di filtri IIR, in generale l'errore quadratico medio non è una funzione monomodale dei coefficienti del filtro, così ci possono essere minimi multipli; inoltre la stabilità del filtro durante la fase adattiva, non è garantita. Le funzione di performance usata per l'adattamento del filtro è il valore atteso del quadrato della sequenza di output del filtro y_n [77]:

$$\mathbb{E}[y_n^2] = |H(e^{i\omega_0})|^2 S + \frac{1-r}{1+r} \sigma^2 \quad (6.8)$$

dove S è la potenza (locale) del segnale di input, σ^2 è la potenza del rumore in input e ω_0 è la frequenza istantanea del segnale. Si può mostrare che il valore di w che massimizza questa funzione è dato da:

$$w^* = \cos \omega_0 . \quad (6.9)$$

La comparazione di questa equazione con 6.7, mostra che la massimizzazione della funzioni di performance corrispondono a tracciare la frequenza locale della sinusoide di input.

Un importante caratteristica di questa funzione di performance è che è monomodale per tutti i valori ammissibili di w ed r . Questo fatto suggerisce l'uso della tecnica di ricerca del gradiente per cercare in minimo. Seguendo [76], possiamo definire un algoritmo normalizzato IIR Least Mean Squares-type. L'output del filtro p dato da:

$$y_n = (1 + r)w_n y_{n-1} - r y_{n-2} + (1 - r)(w_n x_{n-1} - x_{n-2}) \quad (6.10)$$

così che la stima istantanea del gradiente è data da:

$$J = \frac{\partial y_n^2}{\partial w_n} = 2y_n \frac{\partial y_n}{\partial w_n} . \quad (6.11)$$

Il gradiente di y_n è:

$$\alpha_n = \frac{\partial y_n}{\partial w_n} = (1 + r)w_n \alpha_{n-1} - r \alpha_{n-2} + (1 + r)y_{n-1} + (1 - r)x_{n-1} . \quad (6.12)$$

Con questi elementi, possiamo costruire le equazioni dei parametri adattivi:

$$w_{n+1} = w_n + \mu y_n \alpha_n / R_{n+1} \quad (6.13)$$

where R_{n+1} è una stima adattiva della funzione di errore Hessiana[82]:

$$R_{n+1} = \nu R_n + \alpha_n^2 . \quad (6.14)$$

I parametri μ e ν , sono i coefficienti di adattamento

Nel seguito, per semplificare la notazione, useremo w per denotare il valore istantaneo della sequenza w_n

6.3.2 Costi computazionali

I costi computazionali del rivelatore adattivo in figura 6.2, possono essere quantificati contando il numero di operazioni floating points per campione. I costi rilevanti sono dovuti alle operazioni del filtro $H_1(z)$ (valutazione dell'output, adattamento dei parametri del filtro), più il prodotto dell'input e l'output dei campioni.

Dall'analisi di Eqs. 6.10–6.14, abbiamo che:

- y_n : 5 flops;
- α_n : 5 flops;
- w_{n+1} : 3 flops;
- R_{n+1} : 3 flops.

Il costo totale per processare un singolo campione (incluso il prodotto del input ed output del filtro) è pari a 17 flops. Per una stream di N campioni, il costo è pari a $17 \times N$, che può essere scritto, usando la notazione di complessità algoritmica, come $O(N)$, che vuol dire che il costo delle operazioni di filtro è *linear* sulla grandezza dei dati da processare. Per confronto, il controllare il costi per un'analisi con banco di template riportati nel capitolo 4 di questo lavoro di tesi.

6.3.3 Caratterizzazione della risposta del filtro

Applicando l'inverso della trasformata z della funzione razionale che esprime la risposta in impulso nel dominio z , otteniamo la risposta impulsiva nel dominio del tempo:

$$h_n = \frac{2^{-n}(1-r)}{A(r,w)} [w(C(r,w)^n - B(r,w)^n) + -2(C(r,w)^{n-1} - B(r,w)^{n-1})\theta_{n-1}] \quad (6.15)$$

dove $\theta()$ è la funzione gradino unitario e

$$\begin{aligned} A(r, w) &= \sqrt{-4r + (1+r)^2 w^2} \\ B(r, w) &= w + rw - A(r, w) \\ C(r, w) &= w + rw + A(r, w) \end{aligned} \quad (6.16)$$

Notiamo che il sistema è strettamente stabile, da $h_0 = 0$.

La conoscenza della risposta ad impulso del filtro è necessaria per caratterizzare l'uscita statistica del filtro. A tal fine, il primo passo è la valutazione della funzione di autocorrelazione g_n . Essa può essere scritta come [85]:

$$\begin{aligned} g_k &= \sum_{n=0}^{+\infty} h_n h_{n+k} = \\ &= \begin{cases} \frac{2^{-k-1}(1-r)}{A(r,w)(1+r)} [w(1-r)(B(r,w)^k - C(r,w)^k) + \\ - A(r,w)(B(r,w)^k + C(r,w)^k)] & k \geq 1 \\ \frac{1-r}{1+r} & k = 0 \end{cases} \end{aligned} \quad (6.17)$$

Data una sequenza di rumore bianco in input u_n , la sequenza di autocorrelazione del filtro di output y_n può essere scritta come [85]:

$$\rho_k^{yy} = g_k \sigma^2 \quad (6.18)$$

dove σ^2 è la varianza del rumore bianco. La valutazione dell'autocorrelazione al ritardo zero, fornisce la potenza attesa all'uscita del filtro:

$$\sigma_y^2 = \rho_0^{yy} = \frac{1-r}{1+r} \sigma^2 \quad (6.19)$$

dove il fattore $(1-r)/(1+r)$ è il fattore di riduzione del rumore.

La funzione di cross-correlation tra la sequenza di rumore in input al filtro u_n e l'output del filtro v_n , può essere valutata come:

$$\rho_n^{uv} = \sum_{k=0}^{+\infty} h_k \rho_{n-k}^{uu} = \sum_{k=0}^{+\infty} h_k \sigma^2 \delta_{n-k} = \sigma^2 h_n \quad (6.20)$$

dove abbiamo usato il fatto che $\rho_{n-k}^{uu} = \delta_{n-k}$ essendo u una sequenza di rumore bianco. in figura 6.4 mostriamo la risposta all'impulso e la funzione di autocorrelazione del campione in rumore gaussiano bianco, del filtro IIR ALE.

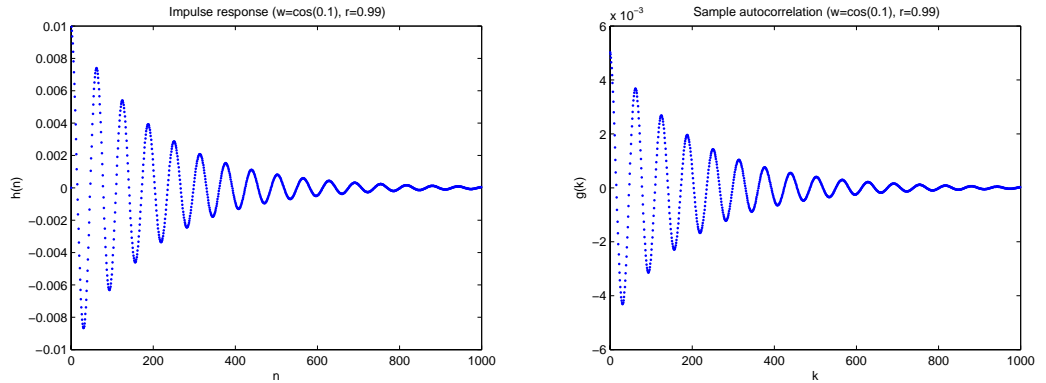


Figura 6.4: Destra: Impulse response Sinistra: Sample autocorrelation function.

6.4 Adaptive correlator detector

Abbiamo visto che, data la struttura del filtro, per rumore bianco in input, la cross-correlation tra input ed output a ritardo zero, è esattamente zero ($r_0^{uv} = \sigma^2 h_0 = 0$). In presenza di componenti coerenti tracciate dal filtro, la cross-correlation va ad essere diversa da zero.

Sembra quindi naturale usare una valutazione della funzione di cross-correlation al ritardo zero come statistica di rilevazione:

$$T_n = \mathbb{E}[xy] \approx \frac{1}{N} \sum_{k=n}^{n+N-1} x_k y_k = \frac{1}{N} \langle \mathbf{x}_n, \mathbf{y}_n \rangle \quad (6.21)$$

dove abbiamo espresso la statistica come un processo stocastico. Tuttavia, per semplicità e poiché siamo soltanto interessati nel decidere dopo l'osservazione delle sequenze non sovrapponibili dei campioni di N , ignoriamo la dipendenza di T da n .

Si può vedere [83], che nel caso di input gaussiano, T corrisponde alla stima a massima verosimiglianza della funzione di cross-correlation, e che questa stima è asintoticamente consistente

Il detector lavorerà quindi comparando il valore della statistica T con una threshold t

(precedentemente fissata in corrispondenza di una desiderata probabilità di falso allarme) e quindi decide se il segnale è presente o meno (also see Fig.6.2).

Poichè ALE produce una stima delle componenti coerenti che entrano in input, l'operazione del correlatore di cui sopra può pensarsi come un processo di matched filtering della sequenza di input con un template modificato dal rumore, dato dall'output del filtro ALE. Poiché le stime di ALE sono comunque affette da rumore, ci aspettiamo una degradazione delle performance rispetto alle convenzionali operazioni di matched filter in ipotesi di perfetta conoscenza del segnale. Il vantaggio principale di questo approccio, tuttavia, è proprio quello di prescindere dalla conoscenza del template.

Nel paragrafo seguente valuteremo le distribuzioni statistiche per i casi di solo rumore in input, ed il caso di presenza di segnale più rumore.

6.4.1 Statistiche di rivelazione: caso con solo rumore

Consideriamo il caso in cui non ci sia alcun segnale presente, così che in input al filtro entri solo una sequenza di rumore gaussiano bianco $x_n = u_n$, e l'output del filtro sia la correlazione della risposta all'impulso h_n con u_n :

$$y_n = h_n * x_n = h_n * u_n = v_n , \quad (6.22)$$

dove $*$ è l'operazione di correlazione. La sequenza v_n è rumore gaussiano correlato. Il test statistico in questo caso può essere scritto come:

$$T = \frac{1}{N} \sum_{n=0}^{N-1} x_n y_n = \frac{1}{N} \sum_{n=0}^{N-1} u_n v_n = \frac{1}{N} \langle \mathbf{u}, \mathbf{v} \rangle . \quad (6.23)$$

Per valutare la distribuzione di T , partiamo con considerare il vettore random esteso $\xi = (\mathbf{u}, \mathbf{v})'$ (dove l'apice denota l'operazione di trasposizione) la cui distribuzione congiunta può essere ottenuta dalle distribuzioni marginali delle sue componenti. Abbiamo $\mathbf{u} \sim \mathcal{N}(0, \sigma^2 I)$, $\mathbf{v} \sim \mathcal{N}(0, G)$, così:

$$\xi \sim \mathcal{N}\left(0, \frac{1}{N} \Sigma\right), \quad \Sigma = \begin{pmatrix} \sigma^2 I & K \\ K' & G \end{pmatrix} \quad (6.24)$$

dove G è la matrice di autocorrelazione su una finestra di N campioni [85], K è la matrice di crosscorrelation su una finestra di N campioni, I è la matrice identica di ordine $N \times N$, e “ \sim ” denota la distribuzione di una variabile casuale (gaussiana in questo caso). Possiamo esprimere T come una forma quadratica nella variabile ξ usando la matrice di permutazione Γ ;

$$T = \frac{1}{2}\xi'\Gamma\xi, \quad \Gamma = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}. \quad (6.25)$$

L'equazione di sopra è nella forma Delta-Gamma-Normal, che in termini di varietà indipendenti può essere scritta, dopo il cambio di variabili in questo modo (vedi eq. C.1):

$$T = \sum_{i=0}^{2N-1} \frac{1}{2}\lambda_i \zeta_i^2 \quad (6.26)$$

Dove le λ_i sono gli autovalori della matrice $\Gamma\Sigma$ (che possono essere calcolati, per esempio, per riduzione triangolare grazie alla decomposizione di Schur's e quindi prendere gli elementi diagonali) Si deve notare che poiché il filtro usato nella derivazione analitica della distribuzione di T non è adattivo, dobbiamo verificare se (e quanto) il filtro adattivo altera la derivazione teorica. Mostriamo mediante un test statistico che la derivazione analitica non è influenzata in un senso significativo.

In figura 6.5, mostriamo la distribuzione empirica e il modello di T . Il primo è ottenuto attraverso l'applicazione diretta del rivelatore adattivo su sequenze di rumore bianco, l'altro è ottenuto campionando la distribuzione modello data da Eq.6.26, che assume un filtro statico. La differenza ΔF tra le statistiche è anche mostrata, per dare una idea migliore della bontà del modello

Per valutare le prestazioni di funzionamento del ricevitore dobbiamo solamente calcolare la threshold che corrisponde alla prescritta probabilità di falso allarme α , che è il $(1 - \alpha)$ -quantile della distribuzione di T . Poiché questo è tipicamente piccolo, può essere accuratamente valutato mediante una espansione asintotica della coda di destra della dis-

tribuzione di T . In appendice D, seguendo l'approccio di [78], mostriamo un'espressione in forma chiusa per valutare la soglia di rilevazione per una fissata probabilità falso allarme.

Un elemento essenziale per l'approssimazione è la valutazione dello spettro degli autovalori della matrice $\Gamma\Sigma$; questo problema, tuttavia, per grandi valori di N , può essere computazionalmente molto intensivo. Anche se gli autovalori devono essere valutati soltanto una volta per una data struttura del filtro (definito dalla larghezza di banda e dalla frequenza iniziale, N), può essere necessario definire differenti strutture del filtro, per esempio per generare un allineamento dei filtri che funzionano parallelamente ma che partono a frequenze differenti.

In generale è difficile sapere quanto i parametri dei filtri modificano lo spettro degli autovalori (e di conseguenza la threshold). Tuttavia, utilizzando simulazioni è possibile avere una risposta empirica a questo problema. Come esempio in figura 6.6, possiamo mostrare come la scelta della banda incide sulla threshold, una volta fissata la probabilità di falso allarme (P_{fa}), la frequenza di campionamento ed N . Per una data P_{fa} , la threshold

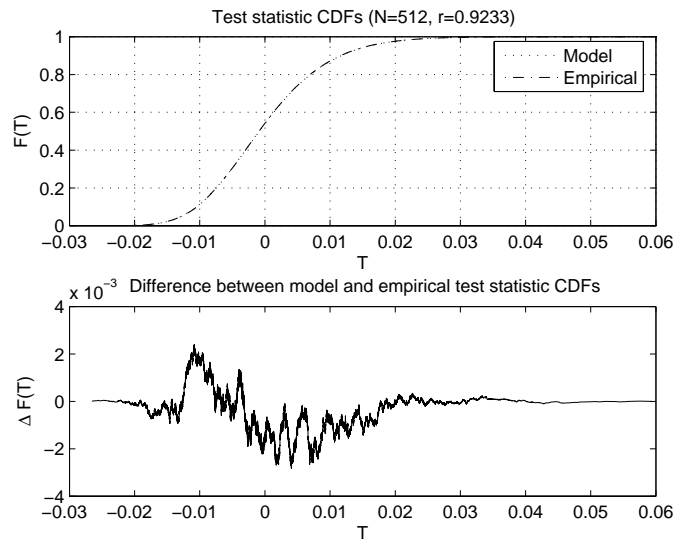


Figura 6.5: Esempio di distribuzione del output del rivelatore (modello empirico)

aumenta con la banda B . Un alta threshold, tuttavia, comporta una diminuzione della probabilità di corretta rivelazione, così che conviene avere una threshold più bassa possibile, che implica una banda bassa. Tuttavia, una banda troppo bassa riduce le capacità di rivelazione del filtro. E' dunque chiaro che occorre trovare un tradeoff tra la banda e la capacità di rivelazione.

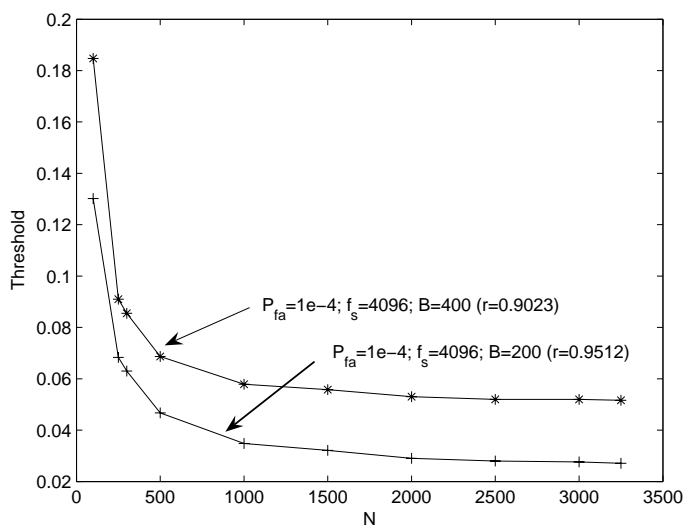


Figura 6.6: Relazione tra la banda(B) e la threshold, fissata la probabilità di falso allarme (P_{fa}), la frequenza di campionamento (f_s) ed N .

6.4.2 Distribuzioni statistiche di rivelazione: segnale più rumore

Nel caso in cui è presente un segnale $x_n = s_n + u_n$ e l'output del filtro è dato da rumore gaussiano correlato con l'aggiunta di un segnale:

$$y_n = h_n * x_n = h_n * [s_n + u_n] = h_n * s_n + v_n = \hat{s}_n + v_n . \quad (6.27)$$

Il test statistico in questo caso può essere scritto come:

$$T = \frac{1}{N} \sum_{n=0}^{N-1} x_n y_n = \frac{1}{N} \sum_{n=0}^{N-1} [s_n + u_n][\hat{s}_n + v_n] . \quad (6.28)$$

Se il filtro ALE si aggancia correttamente al segnale, così che non introduce un ritardo di fase e la sua magnitudine si mantiene approssimativamente uguale ad 1, possiamo assumere che $s_n \approx \hat{s}_n$, così che possiamo scrivere:

$$\begin{aligned}
T &= \frac{1}{N} \sum_{n=0}^{N-1} x_n y_n = \frac{1}{N} \sum_{n=0}^{N-1} [s_n + u_n][s_n + v_n] = \\
&= \frac{1}{N} \sum_{n=0}^{N-1} s_n^2 + \frac{1}{N} \sum_{n=0}^{N-1} s_n [u_n + v_n] + \frac{1}{N} \sum_{n=0}^{N-1} u_n v_n = \\
&= \mathcal{E} + \frac{1}{N} \langle \mathbf{s}, \mathbf{u} + \mathbf{v} \rangle + \frac{1}{N} \langle \mathbf{u}, \mathbf{v} \rangle \quad ,
\end{aligned} \tag{6.29}$$

dove $\mathcal{E} = \mathbf{s}'\mathbf{s}/N$ è una misura quadratica della potenza del segnale. Se introduciamo *signal vector* $\Delta = \frac{1}{\sqrt{N}}(\mathbf{s}, \mathbf{s})'$, così che $\frac{1}{N} \langle \mathbf{s}, \mathbf{u} + \mathbf{v} \rangle = \Delta' \xi$, allora abbiamo:

$$T = \mathcal{E} + \Delta' \xi + \frac{1}{2} \xi' \Gamma \xi \quad , \tag{6.30}$$

che è una forma quadratica (vedi appendice). Così la presenza del segnale si riflette nella espressione formale per il caso di solo rumore, addizionato una costante ed un termine lineare in ξ

6.5 Risultati simulati

Vediamo ora dei test simulati per studiare le performance di IIR ALE per la rivelazione di onde gravitazionali. Per far questo, tenendo conto del rumore sperimentale, testeremo il nostro algoritmo per valutarne il comportamento nel caso di rumore bianco. Per far questo sono state effettuate simulazioni realistiche utilizzando segmenti di data Virgo-like, sbiancati usando un filtro lineare autoregressivo, con 16384 coefficienti stimati con l'algoritmo di Burg [82]. La complessità del modello è stata stimata da validazioni statistiche [87]. Il numero dei parametri in un modello fornisce una misura della sua complessità. Quando il numero dei parametri cresce, un modello riesce ad approssimare meglio i dati che fitta. Tuttavia questo non garantisce che lo stesso modello reagisca bene su nuovi dati che non ha mai incontrato fin ora; il modello lavora come un interpolatore piuttosto che da regressore.

Quindi è necessario il controllo della sua complessità ed la validazione incrociata è una tecnica statistica che consente tale controllo. In pratica, un range di modelli con differente numero di parametri è fittato su un blocco di dati, e le sue performance sono testate su differenti blocchi di dati. Le performance di predizione del modello vengono valutate su quest'altro blocco, e il modello con le prestazioni migliori viene scelto come modello di riferimento.

Infine, per evitare l'influenze della non stazionarietà o di sbiancamento non corretto, valutiamo la decisione statistica per differenti sorgenti di rumore.

6.5.1 Simulazioni in rumore bianco

In questa sezione, mostreremo le performance di agganciamento simulate di IIR ALE per varie frequenze di sinusoidi di input, in rumore bianco, in modo che la corrispondente variazioni di frequenza sia uno step di frequenza alla volta. Le simulazioni sono state svolte con le seguenti specifiche dell'input e del filtro: Rapporto $SNR = 26.5$); parametri del filteo $\mu = 0.1$, $\nu = 0.97$, $r = 0.97$, $R_0 = 10000000$. Prima dell'occorrenza degli step di frequenza, bisogna assicurarsi che il filtro adattivo sia in stato costante, considerando la condizione vicina stato costante iniziale e consentendo un tempo sufficiente, cioè almeno 1000 campioni. In figura 6.7 sono mostrati i risultati della simulazione. In figura 6.7 (prima a sinistra), è mostrata la risposta in frequenza teorica lungo il tempo e la risposta media simulata (curva blu) immersa nella curva di errore standard ottenuta da quella media sommando e sottraendo la deviazione standard calcolata su 2000 realizzazioni delle simulazioni in rumore bianco di media zero e varianza = 6.25. Dalla risposta in step di frequenza è chiaro che c'è un ritardo di circa 0.1 secondi per i filtro per adattarsi come si può vedere in figura 6.7(in basso a sinistra) dove è mostrato l'ampiezza residuale, lungo il tempo, tra il segnale individuato e quello teorico.

Dobbiamo notare che dall'output del filtro possiamo estrarre due tipi di informazioni in tempo reale che sono l'ampiezza dell'uscita o la frequenza del segnale rintracciato se c'è ne nell'input. In figura 6.7 (in alto a destra) è mostrato il rumore in input e in rosso

l'output del filtro IIR ALE. In figura 6.7 (basso a destra) l'SNR lineare per ognuno delle 2000 realizzazioni di rumore. La linea rossa rappresenta la media di SNR (26.5) contenuta nell'errore standard dell'SNR.

Di seguito mostreremo anche tre esempi di chirp simulati in rumore bianco:

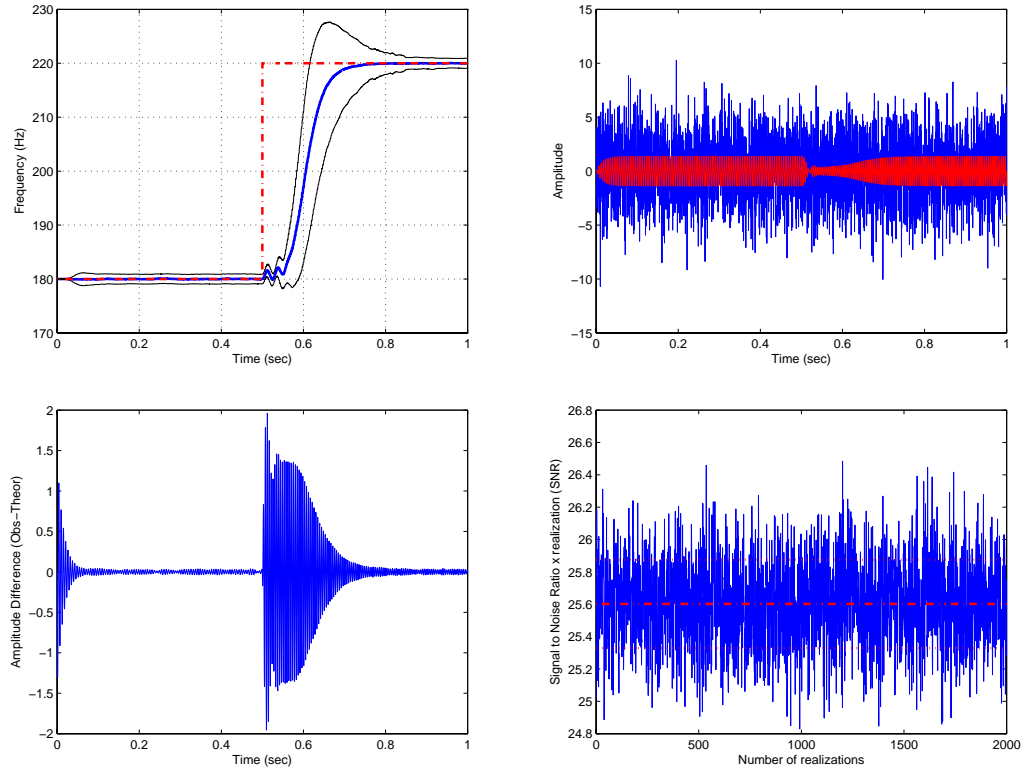


Figura 6.7: Step di frequenze tra 180 Hz e 220 Hz applicate ad un segnale sinusoidale di ampiezza $\sqrt{2}$ e durata 1 sec a frequenza di campionamento $f_s = 4096Hz$ immerso in rumore bianco (SNR = 26.5). Il parametro di IIR ALE $\mu = 0.1$, $\nu = 0.97$, $r = 0.97$, $R_0 = 10000000$

- $1.4M_{\odot} + 1.4M_{\odot}$ Sistema binario in fase inspiralling verso la fase di merge, ordine Post-Newtoniano 2.5, immerso in rumore bianco con un SNR di 24.5, frequenza di campiona-

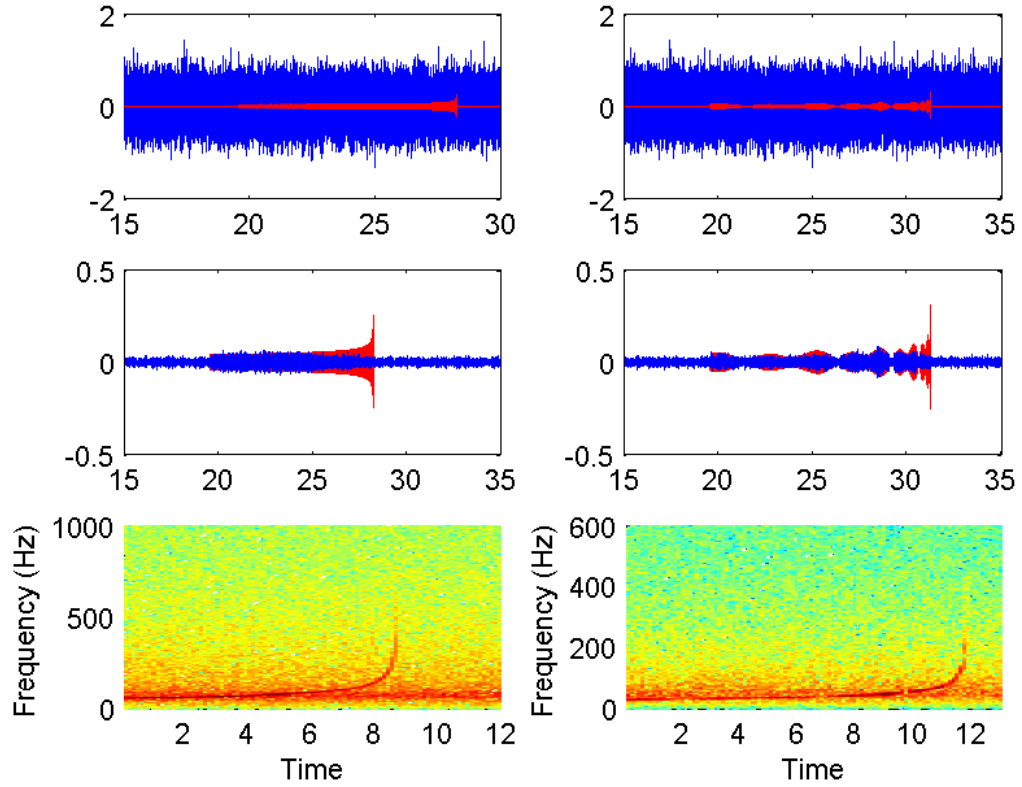


Figura 6.8: Esempio di onda gravitazionale rivelata a frequenza di campionamento $f_s = 4096Hz$. In alto a sinistra: sistema binario di stelle di neutroni $1.4M_\odot + 1.4M_\odot$ con frequenza di partenza $60Hz$, immersa in rumore bianco ($SNR = 24.5$).

Centro a sinistra: in rosso la forma d'onda teorica, in blu la forma d'onda rivelata da IIR Ale. In basso a sinistra: tempo-frequenza della forma d'onda rivelata da IIR Ale. in alto a destra: Stella di neutroni + buco nero $1.4M_\odot + 10M_\odot$; Spin = 0.99; frequenza di partenza $30Hz$, immerso in rumore bianco ($SNR = 24.5$). centro destra: forma teorica (in rosso) e la rivelata forma d'onda (blu). In basso a destra: tempo-frequenza, comportamento della forma d'onda rivelata da IIR Ale.

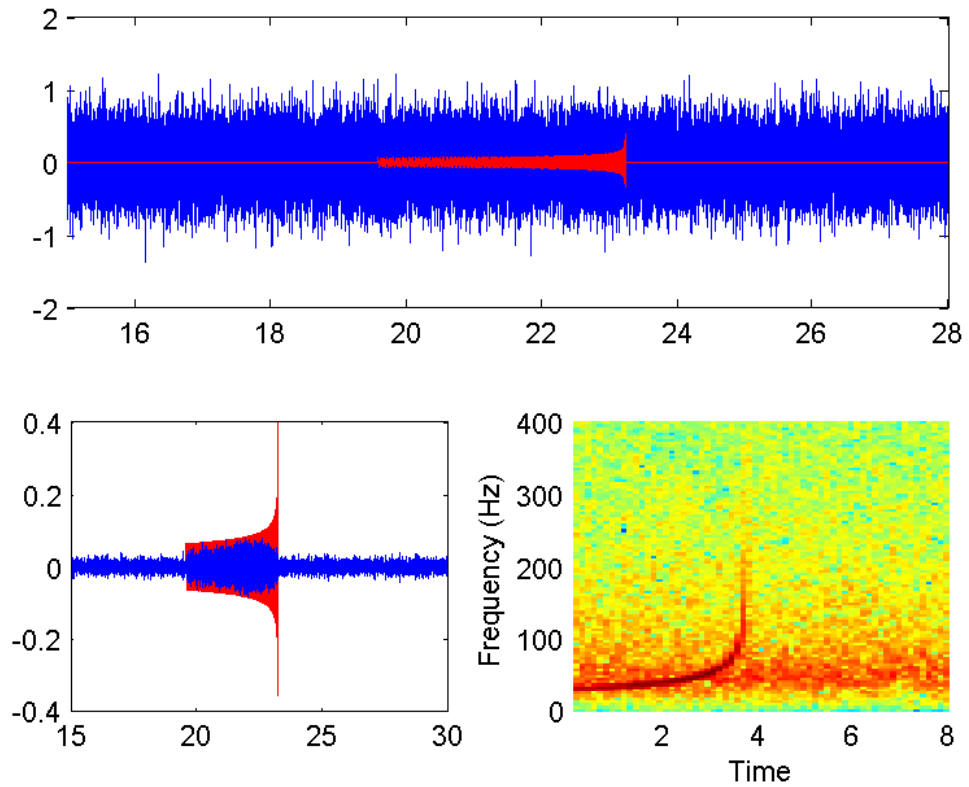


Figura 6.9: Esempio di chirps a frequenza di partenza $f_s = 4096Hz$. In alto: Stella di Neutroni + Buco nero $5M_\odot + 10M_\odot$, forme d'onda calcolata con approccio Effective One Body (EOB); con frequenza di partenza $30Hz$, immerso in rumore bianco ($SNR = 20$). In basso a sinistra: forma d'onda teorica (rosso) e forma d'onda rivelata (blu). In basso a destra: tempo-frequenza - comportamento della forma d'onda rivelata IIR Ale.

mento di $4096Hz$ e frequenza di partenza $60Hz$

- $1.4M_{\odot} + 10M_{\odot}$ Sistema binario - Buchi neri - Ordine Post-Newtoniano 2 immerso in rumore bianco con SNR di 20, frequenza di campionamento $4096Hz$, frequenza di partenza $30Hz$, Spin dei buchi neri = 0.99 (Figgura 6.9).

- $5M_{\odot} + 10M_{\odot}$ coppia di buchi neri - approccio Effective One Body (EOB), ordine Post-Newtoniano 2 immerso in rumore bianco con SNR di 20, frequenza di campionamento $4096Hz$ e frequenza di partenza di $30Hz$ (Figura 6.8)

Come è possibile vedere dalle figure 6.8 e 6.9, le caratteristiche promettenti di questo tipo di trigger ed altresì la possibilità di avere una idea dell'ampiezza di forma d'onda e del comportamento di tempo-frequenza.

Capitolo 7

Conclusioni

L'esperimento VIRGO sta per completare la fase di test ed entrare in presa dati. Alla vigilia di questo evento il problema dell'analisi on-line per sgrossare la grande mole di dati prodotta, rappresenta una tematica di ricerca di fondamentale importanza e di grande attualità.

I risultati ottenuti in questo lavoro di tesi, mostrano l'effettiva possibilità di affrontare il problema dell'analisi on-line per la ricerca di onde gravitazionali, mediante strategie dinamiche in alternativa alle tecniche standard basate su griglie di template, fornendo una risposta forte per abbattere i costi computazionali.

L'utilizzo di algoritmi di ottimizzazione globale, consente di abbattere i costi computazionali restando in-time con l'acquisizione dati utilizzando singole workstation e mantenendo adeguate proprietà di rivelazione del ricevitore. Questo consente di concentrare i cluster e le potenze di calcolo maggiori, per la più complessa analisi off-line che richiede comunque maggiore accuratezza.

L'innovatività del metodo proposto, ed i risultati ottenuti con le prime implementazioni, aumentano l'interesse per questo tipo di strategie, e la speranza di risolvere in maniera semplice e snella il problema dell'analisi on-line abbattendone i costi computazionali.

Tra gli sviluppi futuri per questo tipo di strategia, è interessante l'esplorazione di

possibili improvement dell'algoritmo di Price alla luce dei buoni risultati ottenuti con strategie miste CRS e algoritmi Genetici.

Parallelamente la soluzione per il matched filter dinamico IIR ALE + correlatore ha mostrato, di essere stabile e performante. Le sue forti basi teoriche e la caratteristica principale di rivelare segnali senza conoscerne a priori i template, ne fanno un algoritmo molto promettente e di grande interesse applicativo. Esso, basandosi solo sulla ipotesi di segnali localmente sinusoidali, risulta utilizzabile, non solo per rivelare i segnali prodotti durante la fase di spiraleggio, ma anche per riconoscere i processi di coalescenza. Infine la precisa analisi dei costi computazionali, ha mostrato l'effettiva capacità dell'algoritmo di lavorare in in-time con l'acquisizione dati, utilizzando semplici workstation.

Gli sviluppi futuri di questo lavoro di tesi prevedono l'implementazione delle due strategie durante i run di test attualmente in esecuzione presso l'esperimento Virgo. E' di fondamentale importanza infatti, testare sul campo la capacità da parte degli algoritmi di reggere un grosso carico di informazioni da processare, anche per tunare in maniera adeguata tutti i parametri in gioco, sia degli algoritmi, che quelli di natura prettamente tecnologica.

Appendice A

Appendice - Ottimizzazione globale ed algoritmo di Price

A.1 Il problema dell'ottimizzazione globale

In generale il problema dell'ottimizzazione globale, consiste nel determinare il valore di minimo globale di una funzione, posto che esista, e almeno un punto del dominio in cui tale valore viene assunto. Formalmente, assegnata una funzione f definita in C , detta *funzione obiettivo* o *funzione costo*, il problema dell'ottimizzazione globale può essere scritto come

$$\min f(x) \text{ con } x \in C$$

Dove C è detto, *dominio ammissibile*. I punti $x \in C$ sono detti *ammissibili*, e una *soluzione* del problema è un punto ammissibile $\hat{x} \in C$ tale che

$$f(\hat{x}) = \min_{x \in C} f(x)$$

.

Un minimo locale, invece, verifica la precedente proprietà soltanto in un intorno di \hat{x} . La difficoltà principale che riscontriamo nei problemi di ottimizzazione globale, consiste

nell'enorme differenza tra le condizioni necessarie e quelle sufficienti a garantire il raggiungimento dell'ottimo globale. Ad oggi, nonostante le difficoltà intrinseche del problema, sono disponibili numerose strategie ed algoritmi per la risoluzione di determinati problemi di ottimizzazione globale, in [14] è raccolto l'attuale stato dell'arte delle principali tecniche adottate e del software esistente.

Osserviamo infine che la ricerca del massimo globale, può essere ricondotta facilmente alla ricerca del minimo globale mediante la seguente relazione:

$$\min_{x \in C} f(x) = -\max_{x \in C} (-f(x))$$

A.2 Classificazione dei metodi

I metodi numerici per l'ottimizzazione globale possono essere divisi in tre classi: metodi euristici, metodi approssimativi, metodi sistematici. Diamo adesso una breve descrizione di tali metodi.

A.2.1 Metodi Euristici

A questa classe appartengono tutti quei metodi che non danno alcuna garanzia del fatto che il minimo trovato sia effettivamente un minimo globale, anche se spesso questi metodi forniscono ottimi risultati. Tra questi metodi esistono sia quelli deterministici che quelli stocastici, e per questi ultimi è possibile, in qualche caso molto particolare, provare la convergenza con probabilità arbitrariamente vicino ad uno.

Il più semplice metodo euristico è detto metodo *dell'inizio casuale multiplo*, che consiste nello scegliere punti iniziali casuali e nell'eseguire una ottimizzazione locale da questi punti, nella speranza che uno di essi sia nel bacino di attrazione del minimo globale. Da questo derivano poi tutti quei metodi che cercando di apportare miglioramenti, scegliendo per esempio i punti in maniera più opportuna, ed eseguendo l'ottimizzazione locale solo

selettivamente. Tali metodi si basano sulla analogia con processi naturali. E' importante osservare che i processi in natura non hanno la necessità di essere quelli più efficienti, quindi, queste tecniche possono essere, al più, assunte efficienti in condizioni molto particolari. In generale processi di selezione naturale sono estremamente lenti e dunque, approcci evolutivi di generale applicabilità per problemi di ottimizzazione globale non sono immaginabili in quanto è proibitivo il numero di individui da usare, o, equivalentemente di valori di funzioni obiettivo da calcolare. Tuttavia, per problemi molto particolari, per i quali si possono pensare leggi evolutive ad hoc, particolarmente efficienti, un approccio evolutivo diviene fattibile.

Tra questi metodi ricordiamo:

Metodi di smooting, basati sull'idea che in natura, le caratteristiche macroscopiche sono generalmente un effetto medio dei dettagli microscopici; il mediare *smussa* i dettagli in modo tale da rivelare le caratteristiche globali. Per esempio, un enorme valle vista da lontano, ha una forma semplice e ben definita; solo guardando più da vicino i minimi locali diventano visibili, e questo sempre di più man mano che ci si avvicina alla valle. La speranza è che, *smussando* una funzione obiettivo, il cui grafico mostra un profilo frastagliato, molti o tutti i minimi locali scompaiano e le rimanenti caratteristiche principali della funzione, mostrino così un singolo minimo. A questo punto, aggiungendo sempre più dettagli, le approssimazioni fatte diventano trascurabili e si può così riconoscere, il minimo globale;

Simulated Annealing basato sul fatto che, il riscaldamento ed il lento raffreddamento di un metallo, lo porta in uno stato uniformemente più cristallino, che si ritiene sia lo stato in cui l'energia libera di grandi quantità di materia, ha il suo minimo globale. Il ruolo della temperatura è quello di permettere alle configurazioni di raggiungere stati di energia più alti, con probabilità data dalla legge esponenziale di Boltzman, in modo che queste possano superare le barriere di energia, che le vincolano a rimanere in minimi locali;

Algoritmi genetici basati su concetti della genetica e della selezione naturale. Attualmente gli algoritmi genetici stanno emergendo come importante famiglia dei metodi di ottimiz-

zazione. Tutti i metodi di questa famiglia consistono principalmente nel cercare il minimo globale mediante un'evoluzione di una popolazione iniziale random di possibili candidati ad essere soluzioni. Le evoluzioni possono avvenire secondo tre schemi detti *Riproduzione*, *Mutazione* e *Crossover* che comportano rispettivamente la riproduzione degli individui più promettenti, variazioni sui bit di alcuni elementi o la creazione di incroci tra coppie di elementi. Gli algoritmi genetici sembrano una delle soluzioni più interessanti soprattutto per la loro totale indipendenza dalle caratteristiche della funzione obiettivo. Inoltre la presenza di operatori quali la mutazione, introducono un grado di casualità nella generazione della popolazione meccanismo che evita che l'algoritmo possa sprofondare nei minimi locali.

Algoritmi CRS (Controlled Random Search) essi rappresentano una delle famiglie di algoritmi di ottimizzazione più utilizzate [15][17][16]. Gli algoritmi basati su CRS, tendono a risolvere i problemi di ottimizzazione globale seguendo contemporaneamente due strategie. La prima è quella di esplorare uniformemente, tutta l'area di definizione della funzione in modo da individuare le zone più *promettenti* a contenere un minimo globale. Successivamente, quando viene individuata una area sufficientemente piccola tra queste zone promettenti, parte una ricerca locale che raffina la stima del minimo. Nei metodi CRS la ricerca globale viene eseguita mantenendo una popolazione di punti candidati che tendono a clusterizzarsi intorno alle aree di maggior interesse, fornendo altresì molte informazioni sul comportamento globale della funzione.

Il primo algoritmo globale basato sul metodo CRS, fu proposto da Price [18] successivamente sono state proposte alcune modifiche [19] rispetto all'algoritmo originale che ne diminuiscono significativamente il costo computazionale.

A.2.2 Metodi di Approssimazione

Questi metodi, sono basati sull'idea di trasformare, attraverso convenienti approssimazioni, il problema di ottimizzazione originario in un problema più semplice e quindi maggiormente trattabile. Una volta risolto tale problema, si ottiene una soluzione approssimata, per il problema originario; A questo punto, eseguendo un'ottimizzazione locale

da questa soluzione approssimata, si ha il minimo globale del problema originario, se l'approssimazione è sufficientemente buona, altrimenti si ottiene un buon minimo locale.

Un esempio di applicazione dei metodi di approssimazione si presenta quando, in un problema vincolato, i vincoli non possono essere espressi analiticamente. In questo caso si può ricorrere alle così dette formulazioni di: *penalità* e di *barriera*.

A.2.3 Metodi Sistemati

Fondamentalmente tutti i metodi sistematici sono basati sul *principio del branching*, che consiste nello spezzare ricorsivamente il problema originario in sotto problemi che sono più veloci e più facili da risolvere. Nei metodi di branching puro, i sottoproblemi maggiormente potenziali sono spezzati più di frequente, mentre nei *metodi di branching e bound*, per ogni sottoproblema si calcola il minimo della funzione obiettivo, nella speranza di poter eliminare molti sottoproblemi in una prima fase. Il più semplice metodo sistematico per problemi vincolati è il *metodo della griglia di ricerca*, che consiste nel generare una griglia sempre più fine di punti, e di utilizzare il miglior punto di questa griglia per eseguire un'ottimizzazione locale. L'efficienza di questo metodo è ovviamente limitata a funzioni obiettivo che dipendono da poche variabili, dato che il numero di punti della griglia cresce esponenzialmente con il numero delle dimensioni del problema.

A.3 Algoritmo di Price

Come detto nella sezione precedente, l'algoritmo di Price[18][19] è un esempio di algoritmo di ottimizzazione euristico basato sul metodo di Controlled Random Search (CRS). Più precisamente esso è ottenuto da una estensione del metodo del semplice, per l'ottimizzazione di funzioni su domini multidimensionali prescindendo da ogni informazione sul comportamento della funzione obiettivo da minimizzare.

L'algoritmo di Price è stato utilizzato nell'ambito di questo lavoro di tesi ed in numerosi

lavori pubblicati, [21][23][31][32][33] per la ricerca di segnali gravitazionali immersi nel rumore, secondo un approccio che verrà specificato dettagliatamente nei capitoli successivi.

Le performance dell’algoritmo di Price sono state messe a confronto con altri algoritmi in [61] dove, in tutti i principali test suggeriti in letteratura, mostra grandi potenzialità e velocità di convergenza anche rispetto ad algoritmi basati su tecniche genetiche. La scelta di utilizzare l’algoritmo di Price, inoltre, è anche sostenuta dagli ottimi risultati mostrati nella soluzione di difficili problemi di ottimizzazione in campo astrofisico, dove ha mostrato grande efficienza in problemi con oltre 30 gradi di libertà.[20]. Uno dei pregi di questo algoritmo consiste nella possibilità di inizzializzarlo fornendo intervalli fisici dei parametri da determinare.

A.3.1 Descrizione dell’algoritmo di Price

Questo paragrafo sarà dedicato interamente alla descrizione dettagliata dell’algoritmo utilizzato. Sia $f(x_1, x_2, \dots, x_m)$ una funzione ad m variabili, di cui si vuole cercare il punto di minimo (rispettivamente massimo) globale, in un fissato ipercubo V .

Il primo passo dell’algoritmo prevede la creazione di un array di N punti m -dimensionali di test, estratti uniformemente sul volume V , ed il calcolo del valore della funzione obiettivo in questi punti.

Si crea quindi una matrice di dimensione $N \times (m + 1)$ dove le prime m colonne di ogni riga i contengono le componenti del i -esimo punto estratto da V , mentre la colonna $m + 1$ -esima di ogni riga i , contiene il valore della funzione obiettivo calcolato nell i -esimo punto.

$$\begin{bmatrix} x_{1,1} & x_{1,2} & \dots & x_{1,m} & f(x_1) \\ x_{2,1} & x_{2,2} & \dots & x_{2,m} & f(x_2) \\ \cdot & \cdot & \dots & \cdot & \cdot \\ \cdot & \cdot & \dots & \cdot & \cdot \\ \cdot & \cdot & \dots & \cdot & \cdot \\ x_{N,1} & x_{N,2} & \dots & x_{N,m} & f(x_N) \end{bmatrix}$$

Dopo questa fase di inizializzazione, si estrae uniformemente un sottoinsieme di P_1, \dots, P_{m+1} di $m+1$ punti tra le N righe della suddetta matrice, dopodichè viene individuato un nuovo punto P nello spazio di ricerca V , ottenuto nel seguente modo:

$$P = 2G - P_{m+1}$$

Dove G è il centroide dei primi m punti. L'algoritmo verifica se il punto è ancora nello spazio di ricerca, altrimenti lo scarta e ne prende un altro con la stessa tecnica. Appurato che il punto appartiene al volume V , si calcola il valore della funzione obiettivo in esso $f(P)$, e si confronta con il valore massimo contenuto nella matrice in corrispondenza del punto che chiameremo P_{max} . Se $f(P) < f(P_{max})$, allora il nuovo punto P andrà a rimpiazzare P_{max} nella matrice, altrimenti viene generato un nuovo punto. L'algoritmo itera questo schema e l'array di N punti tende a clusterizzarsi intorno al punto di minimo assunto da f , potendo tuttavia raggiungere anche zone lontane dal minimo rilevato, proprietà che gli consente comunque di uscire all'occorrenza da minimi locali.

L'algoritmo originale non prevede un criterio di arresto che può essere definito dall'utente.

A.3.2 test di performance

In questa sezione verranno presentati alcuni risultati relativi ai test di performance svolti sull'algoritmo di Price implementato ed utilizzato in questo lavoro di tesi. I test svolti sono utili per valutarne il comportamento e per eseguire il tuning nei parametri dell'algoritmo.

Per i test sono state utilizzate alcune funzioni presenti in letteratura per la valutazione degli algoritmi di ottimizzazione globale.

Funzioni Test

1. The Rosembrock Function [34] con $n = 2$ e $D = [-1000, 1000]^n$

$$f(\bar{x}) = \sum_{i=1}^{n-1} (x_1 - 1)^2 + 100(x_i^2 - x_{i+1})^2$$

$$x_{min} = [1, 1, \dots, 1]^T, f_{min} = 0.$$

2. The McCormick Function [35] con $n = 2$ e $D = [-1.5, 4] \times [-3, 4]$

$$f(x_1, x_2) = \text{sen}(x_1 + x_2) + (x_1 + x_2)^2 - 1.5x_1 + 2.5x_2 + 1$$

$$x_{min} = (-0.54719, -1.54719), f_{min} = -1.9133.$$

3. The six hump camel back function [36] con $n = 2, D = [-2.5, 2.5] \times [-1.5, 1.5]$

$$f(x_1, x_2) = (4 - 2.1x_1^2 + x_1^4/3)x_1^2 + x_1x_2 + 4(x_2^2 - 1)x_2^2$$

$$x_{min} = \{[0.08984, -0.71266]^T, [-0.08984, 0.71266]^t\}, f_{min} = -1.0316285. \text{ 6 minimi locali}$$

4. The Shubert's Function [35] con $n = 2, D = [-10, 10]^2$

$$f(x_1, x_2) = - \sum_{j=1}^5 j[\text{sen}((j+1)x_1 + j) + \text{sen}((j+1)x_2)j]$$

$$f_{min} = -24.062499, \text{ con 400 minimi locali e 9 minimi globali nei punti } x_{min} \in \{(a, a), (a, b), (a, c), (b, a), (b, b), (b, c), (c, a), (c, b), (c, c)\} \text{ con } a = -6.774576, b = -0.491391, 5.791794.$$

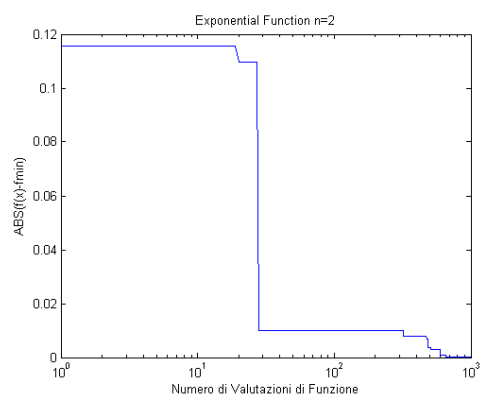
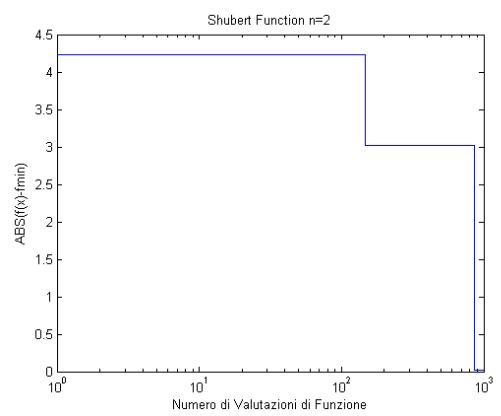
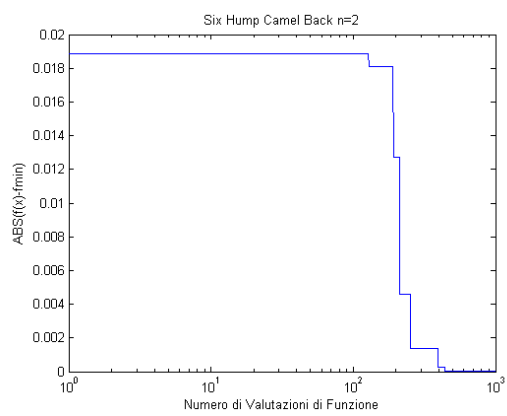
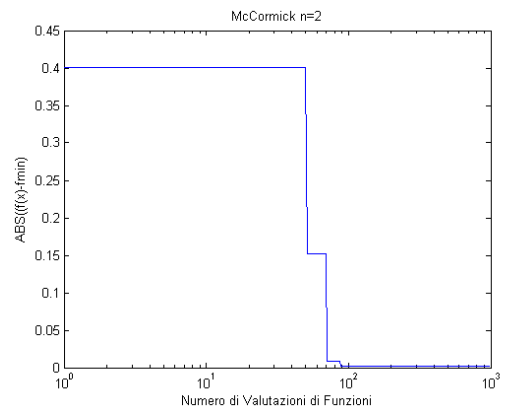
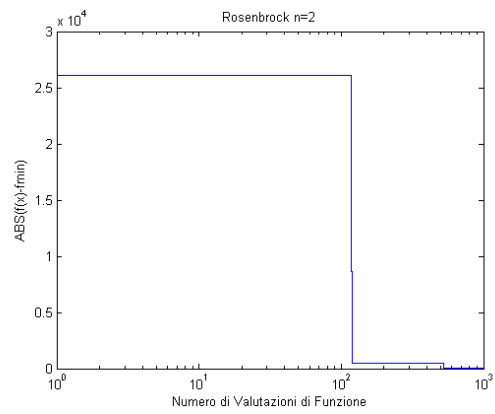
5. Funzione esponenziale [37] $n = 2, D = [-1, 1]^2$

$$f(x_1, x_2) = e^{\frac{x_1^2 + x_2^2}{2}}$$

$$x_{min} \in \{(1, 1), (-1, -1), (-1, 1), (1, -1)\}, f_{min} = 1/e.$$

Grafici dei risultati

Di seguito sono riportati una serie di grafici esemplificativi che mostrano il comportamento dell'algoritmo di Price sulle funzioni precedenti. Tutti i test sono stati svolti con i seguenti parametri: 1000 iterazioni e 25 punti di popolazione iniziale, e mostrano un corretto comportamento dell'algoritmo nella rivelazione del punto di minimo con buona approssimazione già dopo poche centinaia di iterazioni.



Appendice B

Calcolo parallelo e calcolo distribuito per l'analisi off-line

Il problema dell'analisi on-line è legato essenzialmente alla capacità di rivelare il segnale in tempi brevi. I dati sgrossati, necessitano quindi di essere rianalizzati in maniera più accurata per estrarre le informazioni di interesse fisico, e comunque per confermare l'avvenuta rivelazione del segnale. Questa appendice finale è dedicata ai possibili miglioramenti tecnologici applicabili al data analysis per Virgo, in particolare introdurremo alcuni test preliminari svolti utilizzando calcolo parallelo e piattaforme Grid per l'analisi off-line.

B.1 Un approccio parallelo

Come abbiamo visto nei test svolti, la ricerca di segnali gravitazionali prodotti da masse piccole, richiede un miglioramento dell'algoritmo di Price. La soluzione proposta è stata quella di far girare contemporaneamente due istanze di Price che lavorano su domini diversi. Questa strategia, risulta promettente grazie soprattutto alle performance mostrate dalle piattaforme di calcolo di alta fascia.

Un'altra strategia in valutazione è quella di variare i parametri dell'algoritmo di Price, aumentando la grandezza della matrice di ricerca ed il numero massimo di iterazioni

(confronta appendice A). Queste due variazioni, infatti, consentono da un lato una maggiore esplorazione dello spazio delle masse, dall'altro danno più tempo per il processo di clusterizzazione intorno alle zone interessanti, consentendo un aumento dell'accuratezza.

La variazione dei parametri, tuttavia, richiede un aumento del costo computazionale che comporta gioco forza una diminuzione drastica dell'in-time factor a parità di potenza di calcolo, questo compromette l'immediata applicazione di questa strategia al caso dell'analisi on-line.

Tale tecnica, può essere quindi utilizzata in alternativa alla griglia di template per l'analisi off-line, con l'obiettivo quindi di ottenere informazioni più accurate sulla stima dei parametri rispetto al caso semplice usato per la rivelazione. Per questo motivo nell'ambito di questo lavoro di tesi viene proposta altresì, una versione parallela dell'agoritmo di Price, valutando dei risultati preliminari contenuti in [54].

B.1.1 Algoritmo di Price parallelo

La versione parallela dell'algoritmo di Price prevede una architettura gerarchica, c'è quindi un nodo che lavora da master, gestendo la matrice di ricerca A dell'algoritmo, e tutte le funzioni di generazione di numeri casuali per estrarre dalla matrice di partenza le righe (e quindi i punti) da clusterizzare per l'avanzamento dell'algoritmo. Il ruolo dei nodi slave è quello di generare i template in base alle masse calcolate dal master, e quindi di eseguire il match restituendo la stima del SNR, che rappresenta l'operazione più costosa dal punto di vista computazionale.

All'avvio il nodo master genera k punti su cui valutare l'SNR, uno per ogni processore, quindi ogni nodo esegue la computazione e ritorna il risultato al master. Appena il risultato giunge, il master esegue l'aggiornamento della matrice e continua il processo fino a raggiungere una delle condizioni di arresto.

Prima di discutere i risultati, occorre sottolineare che tale strategia implementativa non fornisce una versione esatta dell'algoritmo di Price, poichè può accadere che alcune righe siano rimosse dalla matrice principale A , mentre è ancora in uso da qualcuno dei

nodi che eseguono la computazione. Tuttavia, nel caso di matrici molto grandi, tale effetto risulta trascurabile, i risultati sperimentali confermano in ogni caso il buon comportamento dell'algoritmo così ideato.

B.1.2 Sperimentazioni e conclusioni

L'algoritmo parallelo è stato implementato utilizzando la libreria open-source MPICH, mentre per quanto riguarda i dati di input sono stati utilizzati campioni simulati contenenti un segnale prodotto da un sistema binario di masse 5 e 7 immerso in rumore gaussiano. Il segnale è campionato a $4KHz$ con frequenza di taglio inferiore uguale a $40Hz$. L'algoritmo di Price usa una matrice A di 100 righe e cerca in un dominio di masse comprese tra 4 e 25 masse solari. Il numero di iterazioni utilizzato è 3000 mentre il criterio di stop è legato al numero di iterazioni. Il grafico in figura B.1 mostra l'evoluzione dell'algoritmo di Price durante la procedura di ottimizzazione, mentre nel grafico in figura B.2 abbiamo una rappresentazione dell'andamento del tempo di esecuzione dell'algoritmo in funzione del numero di processori utilizzato.

In conclusione i test preliminari mostrano un comportamento interessante in termini di rivelazione dell'algoritmo di Price in versione parallela. Dal grafico in figura B.2 vediamo che le prestazioni dell'algoritmo migliorano fino ad un numero di nodi pari a 6, dopo di che il fattore comunicazione prende il sopravvento. Restano comunque alcune questioni da chiarire riguardo al comportamento dell'algoritmo e delle possibilità di applicarlo sia al caso dell'analisi off-line, sia per quella on-line.

B.2 Approccio distribuito

Per completare la fase di analisi dati, come già citato nel capitolo 4 paragrafo 4.3, occorre una intensa attività di analisi off-line, che prevede l'utilizzo di modelli più sofisticati per le forme d'onda e, nel caso di strategie classiche, griglie di template molto fitte che richiedono un elevato costo computazionale. L'approccio parallelo può certamente contribuire ad una

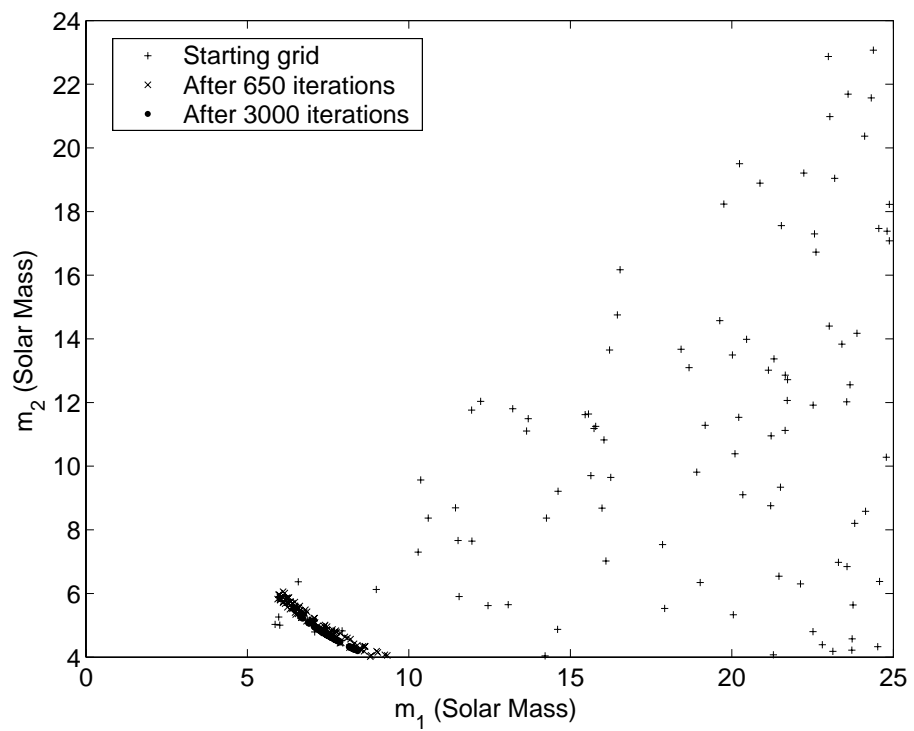


Figura B.1: Evoluzione dell'array di ricerca dell'algoritmo di Price durante la procedura di ottimizzazione

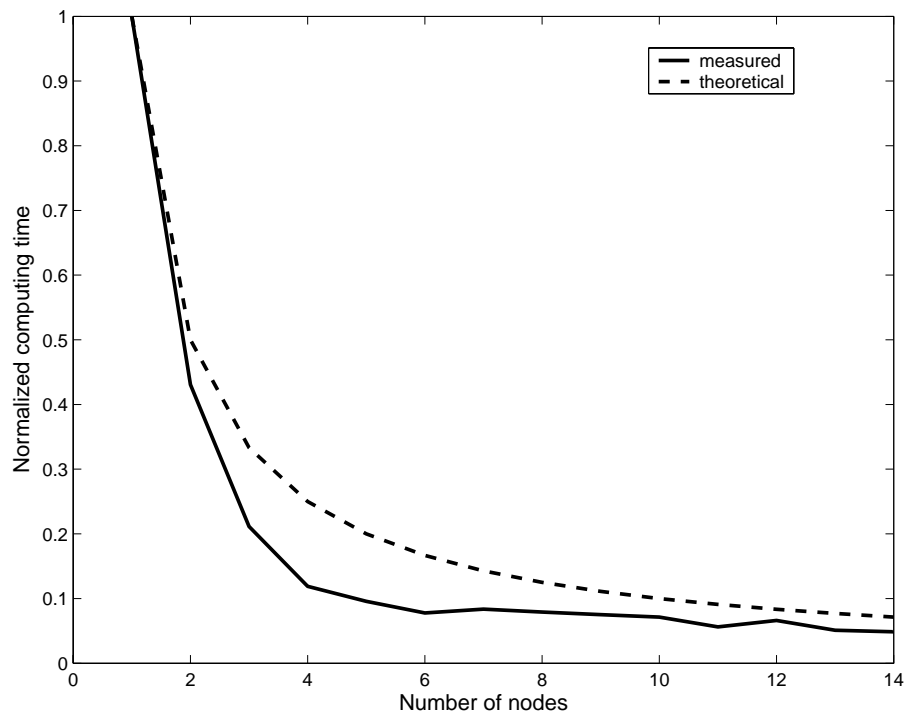


Figura B.2: Tempo di esecuzione dell'algoritmo rispetto al numero dei processori (il tempo di esecuzione è normalizzato rispetto al tempo necessario con un singolo processore)

semplificazione di questa analisi, tuttavia, data la grande mole di dati attesi, l'utilizzo di sistemi di calcolo distribuito risulta inevitabile. In questo framework si inquadra l'attività dell'esperimento Virgo legata al Grid Computing, nella quale l'autore di questo lavoro di tesi è impegnato assiduamente per studi di innovazione tecnologica a supporto del data analysis di Virgo [22][25][26][27].

B.2.1 Il Grid Computing

Il Grid Computing costituisce la nuova generazione dei sistemi distribuiti per applicazioni su scala geografica. Esso nasce per supportare i grandi esperimenti di fisica, biologia e di vari campi della scienza, che coinvolgono grandi collaborazioni internazionali. Il Grid Computing si basa sul concetto di Virtual Organization (VO) che consiste in un insieme di persone che vogliono condividere le loro risorse computazionali e di storage per applicazioni di interesse comune. Il Grid abilita altresì la condivisione di risorse tra differenti VO, abilitando gli amministratori delle risorse ad imporre policy di accesso e di utilizzo aggiuntive rispetto a quelle offerte dai sistemi operativi.

L'INFN è da sempre all'avanguardia sulle nuove tecnologie di calcolo a supporto dei suoi esperimenti, ed anche in ambito Grid si è mostrato subito uno degli enti di ricerca più attivi, portando avanti vari progetti di grid nazionali ed internazionali a partire dal 2001. Ad oggi la Grid INFN, basata sul middlewre INFN-GRID/glite[53], rappresenta la più importante griglia italiana, in grado di supportare efficacemente tutti gli esperimenti che ne fanno richiesta.

La comunità di Virgo, con i suoi ricercatori sparsi tra le differenti sedi di Italia, Francia e da poco anche Olanda, rappresenta un chiaro esempio di Virtual Organization. In questi ultimi anni molti sforzi sono stati impiegati per introdurre i siti di Virgo all'interno della principale grid italiana, in particolare da parte del gruppo di Napoli che sta promuovendo l'uso della griglia anche per le problematiche di replica dati.

E' opportuno, prima di affrontare la questione dei test, introdurre un po di nomenclatura. In ambito Grid Infn con il termine CE (Computing Element) indicheremo i front

end verso le risorse computazionali rappresentati dai WN (Worker Node). Infine useremo la sigla SE (Storage Element) per indicare i front end verso le risorse di storage.

B.2.2 Test preliminari su piattaforma INFN-GRID

Per dimostrare l'effettiva utilità dell'impiego del Grid computing in ambito Virgo, sono stati svolti alcuni test di fattibilità utilizzando la grid INFN, descriviamo quindi il tipo di test svolti ed i risultati ottenuti.

I job inviati riguardano ancora una volta istanze di matched filters per la rivelazione di binarie coalescenti, questa volta utilizzando griglie di template (SMFT). Le sedi utilizzate per la computazione sono il CNAF di Bologna, INFN Ū Napoli and INFN Ū Roma1. I test sono stati svolti su dati simulati, contenenti sia segnale che rumore, similmente a quelli prodotti per l'analisi del metodo DMFT. I dati sono stati prodotti nel sito di Napoli, mentre le risorse computazionali sono state così suddivise in: 1 Worker Node in Bologna, 4 Worker node in Napoli e 2 Worker Node in Roma1. I dati simulati sono caratterizzati dai seguenti parametri:

- Frequenza di campionamento 2 kHz;
- Frequenza di taglio inferiore 50 Hz
- Spazio di ricerca compreso tra 2 e 10 masse solari
- Minimal match per il recupero del 97% del SNR

Le condizioni poste corrispondono a circa 105 template utilizzati su dataset lunghi 15 minuti ciascuno. Il processo è stato diviso in tre

1. Primo passo consiste nel distribuire i dati dalla storage area di Napoli. Il processo prevede l'estrazione dei dati dal formato frame, usato come standard nell'ambito del progetto VIRGO. Quindi viene eseguita la replica dei dati tra gli storag element.
2. Il secondo passo consiste nel dividere lo spazio dei template in 10 sottospazi, e sottomettere dalla User Interface di Napoli, un job per ogni sottospazio. Ognuno di

questi job, legge i dati dallo Storage element più vicino e va in esecuzione sui Worker Node locali.

3. In fine tutti i risultati vengono recuperati dalla UI di Napoli

I risultati dei tests, in termini di tempo di computazione sono i seguenti: Il tempo necessario per eseguire l'intero job dalle 7.5 alle 11 ore, in dipendenza delle performance dei calcolatori e della frazione di template da calcolare(template con masse più piccole impiegano più tempo per essere calcolati). Il tempo di sottomissione alla grid è di circa 2 minuti. E' evidente che l'uso del grid aggiunge un overhead perfettamente accettabile, viste le durate medie dei job da eseguire.

Appendice C

Famiglia di modelli

Delta-Gamma-Normal

Per valutare le performance di rivelazione del rivelatore proposto, deve essere calcolata la distribuzione del test statistico T . Come abbiamo visto esso non è un compito semplice, poiché le distribuzioni che otteniamo non possono essere espresse in forma chiusa. In particolare, mostriamo che dopo alcune manipolazioni algebriche, il test statistico può essere scritta con una forma quadratica nelle variabili normali.

$$T = \theta + \Delta' \xi + \frac{1}{2} \xi' \Gamma \xi . \quad (\text{C.1})$$

Questo modello è chiamato Delta-Gamma-Normal [78], [80].

Per facilitare il trattamento analitico della statistica T , dobbiamo scrivere quest'ultimo come somma di variabili casuali indipendenti che sono funzioni di variabili casuali normali standard ζ_i . Questo può essere fatto risolvendo il problema generalizzato degli autovalori, cioè cercando la matrice C e Λ tali che:

$$CC' = \Sigma, \quad C' \Gamma C = \Lambda . \quad (\text{C.2})$$

Per definizione:

$$\xi = C\zeta, \quad \delta = C'\Delta, \quad \Lambda = \text{diag}(\lambda_0, \dots, \lambda_{N-1}) \quad (\text{C.3})$$

Possiamo scrivere:

$$T = \theta + \sum_{i=0}^{2N-1} \left[\delta_i \zeta_i + \frac{1}{2} \lambda_i \zeta_i^2 \right] \quad (\text{C.4})$$

Appendice D

Code asintotiche e quantili

Data la generalizzazione degli autovalori in T , nell'equazione 6.26 sia ordinata in modo crescente. Supponiamo che ci siano $M \leq 2N$ autovalori distinti, e sia i_j l'indice più alto del j -esimo autovalore distinto, e μ_j la sua molteplicità. Abbiamo allora che $\lambda_{i_1} < \lambda_{i_2} < \dots < \lambda_{i_M}$. Per $j = 1, \dots, M$, definiamo:

$$T_j = \frac{1}{2} \lambda_{i_j} \sum_{l=i_{j-1}+1}^{i_j} \zeta_l^2 \quad (\text{D.1})$$

Allora i T_j sono indipendenti, ed abbiamo:

$$T = \sum_{j=1}^M T_j \quad (\text{D.2})$$

L'equazione sopra essenzialmente definisce T come somma delle variabili casuali indipendenti χ^2 , con differenti gradi di libertà e scala. Specificamente se $g(\cdot; \mu_j)$ è la $\chi_{\mu_j}^2$ densità, allora la densità di T_j è data da:

$$f_j(x) = \frac{2}{|\lambda_{i_j}|} g\left(\frac{2}{\lambda_{i_j}} x; \mu_j\right) \quad (\text{D.3})$$

Poichè siamo interessati al quantile superiore della distribuzione di T , in accordo con i risultati in [78], possiamo distinguere tre casi, basato sul segno dell'autovalore più grande.

Tuttavia, nel nostro caso è possibile mostrare che il più grande autovalore di $\Gamma\Sigma$ è sempre positivo. Infatti la matrice Γ agisce su Σ come un blocco di permutazione:

$$\Gamma\Sigma = \frac{1}{N} \begin{pmatrix} K' & G \\ \sigma^2 I & K \end{pmatrix} . \quad (\text{D.4})$$

E' possibile mostrare che [85]:

$$\text{Tr}(\Gamma\Sigma) = 2\sigma^2 h_0 = 0 . \quad (\text{D.5})$$

Poichè per ogni matrice quadrata, la somma dei suoi autovalore è uguale alla traccia abbiamo:

$$\sum_{i=1}^N \lambda_i = 0 , \quad (\text{D.6})$$

dove λ_i sono gli autovalori di $\Gamma\Sigma$. Escludendo il caso degenere in cui tutti gli autovalori sono nulli (allora in questo caso per l'equazione D.2, noi prendiamo una densità degenere) allora abbiamo autovalori sia negativi, positivi che nulli, tali che la loro somma è zero. Quindi deve esistere almeno un autovalore positivo (anche con molteplicità maggiore di uno) che sia il più grande. Così possiamo considerare solo il caso in cui il più grande autovalore sia positivo.

Possiamo ora applicare il seguente risultato da [78]. Sia $\alpha \rightarrow 0$, allora il quantile superiore di T soddisfa la seguente equazione:

$$x_{1-\alpha} = \lambda_{i_M} \ln b_M + \frac{\lambda_{i_M}}{2} \chi_{1-\alpha, \mu_M}^2 + O(1/\sqrt{\chi_{1-\alpha, \mu_M}^2}) \quad (\text{D.7})$$

dove $\chi_{1-\alpha, \mu_M}^2$ è il $(1 - \alpha)$ -quantile della distribuzione $\chi_{\mu_M}^2$ e:

$$b_M = \prod_{j=1}^{M-1} \left(1 - \frac{\lambda_{i_j}}{\lambda_{i_M}} \right)^{-\mu_j/2} . \quad (\text{D.8})$$

Bisogna notare che in pratica la probabilità di falso allarme deve essere più piccola di 10^{-2} perchè valga l'approssimazione.

Bibliografia

- [1] F. Acernese et al, 2004 *Class. Quantum Grav.* **21** S385-S394.
- [2] D. Sigg, 2004 *Class. Quantum Grav.* **21** S409-S415.
- [3] B. Willke et al, 2004 *Class. Quantum Grav.* **21** S417-S423.
- [4] R. Takahashi and the TAMA Collaboration, 2004 *Class. Quantum Grav.* **21** S403-S408.
- [5] P. Bender, I. Ciufolini, J. Cornelisse, K. Danzmann, W. Folkner, F. Hechler, J. Hough, Y. Jafry, R. Reinhard, D. Robertson, A. Rüdiger, M. Sandford, R. Schilling, B. Schutz, R. Stebbins, T. Summer, P. Toubol, S. Vitale, H. Ward, W. Winkler, *LISA: Laser Interferometer Space Antenna for the detection and the observation of gravitational waves*, ESA Pre-Phase A Report, **MPQ 208**, Max Planck Institut für Quantenoptik, Munich, Germany (February 1996).
- [6] A. Einstein 'Zur allgemeinen Relativitätstheorie' Preuss Akad Wiss Berlino novembre 1915
- [7] B. S. De Witt, 'La Gravità Quantistica' *Le Scienze* No. 186 febbraio 1984 in 'Campi Forze e Particelle a cura di Luciano Maiani. Le Scienze S.p.A. Editore 1991.

- [8] Consiglio Nazionale delle Ricerche, Area della Ricerca di Frascati, Istituto della Fisica dello Spazio Interplanetario, Ricerca di Onde Gravitazionali 1981-1993
- [9]
- [10] H.V.Poor. An introduction to Signale Detection and Estimation. Springer, second edition 1994.
- [11] S.M. Key. Fundamentals of Statistical Signal Processing: Estimation Theory Prentice Hall PTR. 1998.
- [12] S.M. Key. Fundamentals of Statistical Signal Processing: Detection Theory Vol II Prentice Hall PTR. 1998.
- [13] A.V. Oppenheim, R.W. Schafer, and J.R. Buck. Discrete-Time signal processing. Prentice Hall. Second edition 1999
- [14] Arnold Neumaier Home Page -<http://www.mat.univie.ac.at/~neum/glopt.html>
- [15] Barone F, Milano L, Russo G(1990)In: Active colose binaries, Ibanoglu C (ed)(Kluwer Dordrecht)pp.161-188
- [16] Branchetti P,De Felice Ciccoli M,Di Pillo G Lucidi S (1997) J global optim:165-184
- [17] Milano L, Barone F, Milano M(1997) Phys. Review D, 55 4537-4554.
- [18] Price WL (1979) Comput. J 20:367-370
- [19] Price WL (1983) J Optimi Theory Appl 40:333-348
- [20] F. Barone, L. Di Fiore, L. Milano, G.Russo 1993 Astrophys. J 407, 237.

- [21] F. Acernese, F. Barone, R. De Rosa, A. Eleuteri, L. Milano, S. Pardi, G. Russo, Dynamic matched filter technique for gravitational wave detection from coalescing binary systems by means of optimisation algorithms, 2004, *Classical and Quantum Gravity*, 21, pp. S807-S810.
- [22] F. Acernese, F. Barone, L. Brocco, R. Esposito, S. Frasca, L. Mastroserio, L. Milano, C. Palomba, S. Pardi, K. Qipiani, F. Ricci, G. Russo, A GRID solution for Gravitational Waves Signal Analysis: performance of test algorithms and future developments, 2004, *Classical and Quantum Gravity*, 21, pp. S811-S814.
- [23] F. Acernese, F. Barone, R. De Rosa, A. Eleuteri, L. Milano, S. Pardi, G. Russo, Dynamic Matched Filters for Gravitational Waves Detection, 2004, *Classical and Quantum Gravity*, 21, pp. S1849-S1854.
- [24] F. Acernese, F. Barone, R. De Rosa, A. Eleuteri, L. Milano, S. Pardi, G. Russo, S. Spadaccini, A hierarchical Bayesian framework for nonlinearities identification in gravitational wave detector outputs, 2005, *Classical and Quantum Gravity*, 22, pp. S1123-S11232.
- [25] S. Pardi, F. Acernese, F. Barone, R. De Rosa, A. Eleuteri, L. Giordano, L. Milano, G.A. Ruele, I. Ricciardi, G. Russo, Networking in the Beowulf Clusters, 2004, *SAIt*, 5, No. 1.
- [26] S. Pardi, F. Acernese, F. Barone, R. De Rosa, A. Eleuteri, R. Esposito, L. Giordano, L. Milano, P. Mastroserio, L. Milano, G.A. Ruele, I. Ricciardi, G. Russo, A Beowulf Cluster for Matched Filters, 2004, *SAIt*, 5, No. 1.
- [27] F. Acernese, F. Barone, R. De Rosa, A. Eleuteri, L. Milano, S. Pardi, G. Russo, S. Spadaccini, A Multi-standard Farm Prototype for Gravitational Waves Signal Analysis, 2004, *Classical and Quantum Gravity*, 21, pp. S837-S842.

- [28] F. Acernese, F. Barone, R. De Rosa, A. Eleuteri, L. Milano, S. Pardi, G. Russo, et al., 'A simple line detection algorithm applied to Virgo data', (2005) *Class. Quantum Grav.* 22, pp. S1189-S1196
- [29] A. Eleuteri, L. Milano, R. De Rosa, F. Garufi, F. Acernese, F. Barone, L. Giordano, S. Pardi, Adaptive filters for detection of gravitational waves from coalescing binaries, 2006, *Physical Review D*, 73, pp. 122004.
- [30] Eleuteri, L. Milano, R. De Rosa, F. Garufi, F. Acernese, F. Barone, L. Giordano, S. Pardi, et al. 'The status of coalescing binaries search code in Virgo, and the analysis of C5 data', 2006 *Class. Quantum Grav.* 23 pp. S187-S196
- [31] F. Acernese, F. Barone, R. De Rosa, L. Milano, S. Pardi, Dynamic Matched Filter for the Detection of Gravitational Waves, 2004, in Proceedings of the SPIE International Symposium on Astronomical Telescopes Instrumentation, June 21-25, Glasgow (Scotland).
- [32] F. Acernese, F. Barone, R. De Rosa, A. Eleuteri, L. Milano, S. Pardi, Dynamic Matched Filter for Gravitational Wave Detection, 2005, in Proceedings of the 16th SIGRAV Conference on General Relativity and Gravitational Physics, AIP Conference Proceedings, n.751, G. Esposito, G. Lambiase, G. Marmo, G. Scarpetta e G. Vilasi eds., 13-16 September 2004, Vietri sul Mare (Salerno), Italy, pp. 236-238.
- [33] F. Acernese, F. Barone, R. De Rosa, A. Eleuteri, L. Milano, S. Pardi, G. Russo, Dynamic Matched Filters for Gravitational Waves Detection, 2004, in Proceedings of CHEP 2004, International Conference on Computing in High Energy and Nuclear Physics, September 27 - October 1, Interlaken, Switzerland.

- [34] Wolfe MA (1978) Numerical Methods for Unconstrained Optimization, Van Nostrand Reinhold Company, New York
- [35] <http://www.imm.dtu.dk/~km/GlobOpt/testex/testproblems.htm>
- [36] Dixon LCW, Szego GP (1978) Towards Global Optimization 2, Amsterdam
- [37] Breiman L, Culter A (1993) Math program 58: 179–199
- [38] <http://wwwcascina.virgo.infn.it/DataAnalysis/300Gf/index.html>
- [39] Virgo Collaboration- VirgoA parallel in-time analysis system for Virgo - Journal of Physics: 32 (2006) 35–43 Sixth Edoardo Amaldi Conference on Gravitational Waves
- [40] E.E. Flanagan. Astrophysical source of gravitational radiation and prospects for their detection. gr-qc/9804024 1998.
- [41] S.D. Mohanty and S.V. Dhurandhar. Hierarchical search strategy for detection of gravitational waves from coalescing binaries: Extension to post-newtonian waveforms. Physical Review D, 57(2) 1998
- [42] C. Culter and E.E. Flanagan. Gravitational waves from merging compact binaries: how accurately can one extract the binary's parameters from the inspiral waveform? Physical Review D, 49 (6) 1994.
- [43] Relativistic Gravitation and Gravitational Radiation, J.A.Marck, J.P.Lasota (eds) Cambridge U. Press (1997)- Luc Blanchet
- [44] Interferometric Gravitational wave detectors - Peter R. Saulson, World Scientific Publishing, ISBN 981-02-1820-6
- [45] Thermal noise limit to the Virgo Sensitivity P.Amico et al. Classical and Quantum Gravity Vol 18(2001)

- [46] Mechanical quality factor of large mirror substrates for Gravitational waves detectors, P.Amico at All, Rev. Sci. Instr.73 (2002) 179
- [47] Fused silica suspension for the Virgo optics: status and prespectives, Classical and Quantum Gravity, Vol.19, No 7, 7 April 2002. P.Amico at All.
- [48] C. Bradaschia et al., §The VIRGO Project, Final Design of the Italian-French large base interferometric antenna of gravitational wave detection, Proposal to INFN Italy and CNRS France, 1989, 1992, 1995.
- [49] R.E. Vogt, R.W. Drever, F.J. Raab, K.S. Thorne, §Proposal for the construction of a large interferometric detector of gravitational waves, Proposal to the National Science Foundation, California Institute of Technology, Pasadena, California, USA, 1989.
- [50] Hough et al., §Proposal for a joint german-british interferometric gravitational wave detector, MPQ 147, Max Planck Institut für Quantenoptik, Munich, Germany, 1989
- [51] unpublished, see <http://tamago.mtk.nao.ac.jp> (1996)
- [52] P. Bender, et al. §LISA: Laser Interferometer Space Antenna for the detection and the observation of gravitational waves,MPQ 208, Max Planck Institut für Quantenoptik, Munich, Germany, 1996
- [53] Vedi la pagina ufficiale <http://grid-it.cnaf.infn.it>.
- [54] F. Acernese, F. Barone, R. De Rosa, A. Eleuteri, L. Milano, S. Pardi, - Dynamic Matched Filters for Gravitational Wave Detection, Proceedings of AIP Conference, General Relativity and Gravitational Physics, vol. 751 (2005) 236-238.

- [55] L.Blanchet. Second-Post-Newtonian Generation of Gravitational Radiation. arXiv: gr-qc/9501030 24 Jan 1995
- [56] L.Blanchet. Binary Signal Modelling. Second Workshop on Gravitational Wave Data Analysis. 13-15 Novembre 1997
- [57] L.Blanchet. Post Newtonian Computation of Binary Inspiral Waveform. arXiv: gr-qc/0104084 25 Apr 2001
- [58] A.Krolak P. Trzaskoma Application of Wavelet analysis to Estimation of Parameters of the Gravitational Wave Signal from a Coalescing Binary - Classical and Quantum Gravity Vol 13 N 5 maggio 1996
- [59] S.Marano. Rivelazione di onde graviazionali di tipo transitorio - Tesi di dottorato in Ingegneria Elettronica ed Informatica - Università di Napoli, Febbraio 1997
- [60] P.Canitrot, A. Vicerè, L.Milano - Computational costs for coalescing binaries detection in VIRGO using matched filters. - Virgo note. Code: VIR-NOT-PIS-1390-149
- [61] Margherita Bresco, Giancarlo Raiconi, Fabrizio Barone, Rosario De Rosa, Leopoldo Milano: Genetic approach helps to speed classical Price algorithm for global optimization. *Soft Comput.* 9(7): 525-535 (2005)
- [62] home page <http://www.gnu.org/software/gsl/>
- [63] home page <http://www.lsc-group.phys.uwm.edu/daswg/projects/lal.html>
- [64] home page <http://www.fftw.org/>
- [65] P. Ajith, B. R. Iyer, C. A. K. Robinson, B. S. Sathyaprakash *Phys.Rev.* **D71** (2005) 044029

- [66] A. Buonanno and T. Damour, A. Buonanno and T. Damour, Phys. Rev. D 62, 064015 (2000). *Phys. Rev. D* **62**, 064015 (2000).
- [67] A. Buonanno Y. Chen Y. Pan, H. Tagoshi, M. Vallisneri, *Phys.Rev.* **D72** (2005) 084027
- [68] Eanna E. Flanagan, Scott A. Hughes,Phys. Rev. D 57, 4566-4587 (1998). *Phys. Rev. D* **57**, (1998).
- [69] P. Canitrot, L. Milano, A. Vicere. Computational costs for coalescing binaries detection in VIRGO using matched filters, VIR-NOT-PIS-1390-149, Issue 1, 5/5/2000.
- [70] F.Riccio - Un algoritmo di Ottimizzazione Globale per la Rivelazione di Onde Gravitazionali - Tesi - Seconda Università degli Studi di Napoli - Facoltà di Scienze MM.FF.NN.
- [71] C.W. Misner, K.S. Thorne, and J.A. Wheeler, *Gravitation* (Freeman & Co., San Francisco, 1973).
- [72] D.G. Blair, *The Detection of Gravitational Waves*, (Cambridge University Press, Cambridge, 1991).
- [73] P.R. Saulson, *Fundamentals of Interferometric Gravitational Wave Detectors*, (World Scientific Press, New Jersey, 1994).
- [74] L. Milano, F. Barone, M. Milano, Physical Review D **55**, 4537-4554 (1997).
- [75] F. Acernese, F. Barone, R. De Rosa, A. Eleuteri, L. Milano, Classical and Quantum Gravity **21**, S781 (2004).
- [76] B. Widrow, and S. Stearns, *Adaptive Signal Processing*, Prentice Hall, 1985.
- [77] D. R. Hush, *et al.*, IEEE Transactions on Acoustics, Speech and Signal Processing **34**, 1380–1390 (1986).

- [78] S. Jaschke, C. Klüppelberg, A. Lindner. Asymptotic Behavior of Tails and Quantiles of Quadratic Forms of Gaussian Vectors. *Journal of Multivariate Analysis* 88, 252–273 (2004).
- [79] L. Bosi - Segnali gravitazionali da binarie coalescenti: analisi dei dati per l'interferometro VIRGO - Tesi - Università degli Studi di Perugia - Facoltà di Scienze Matematiche, Fisiche e Naturali
- [80] N. N. Vakhaniya, N. P. Kandelaki, *Soobshch. Akad. Nauk Gruzin. SSR* **50**, 535–540 (1968). (In Russian)
- [81] J. Abate and W. Whitt (1992). The Fourier-series method for inverting transforms of probability distributions. *Queueing Systems. Theory and Applications*, 10:5–88.
- [82] S. Haykin. *Adaptive Filter Theory*. Prentice Hall, 1995.
- [83] S. Kay. *Fundamentals of Statistical Signal Processing, Vol I - Estimation Theory*. Prentice Hall, 1998.
- [84] S. Kay. *Fundamentals of Statistical Signal Processing, Vol II - Detection Theory*. Prentice Hall, 1998.
- [85] A. Oppenheim, R. W. Schaffer. *Digital Signal Processing*. Prentice Hall, Englewood Cliffs, New Jersey, 1975.
- [86] A. Papoulis. *Probability, Random Variables and Stochastic Processes*. 3rd edition. McGraw-Hill, 1992.
- [87] K. Fukunaga. *Introduction to Statistical Pattern Recognition*, second edition. Academic Press, 1990.
- [88] Milano L, Barone F, Milano M (1997) *Phys. Review D*, 55. 4537-4554