

Filtri tempo-frequenza basati su spettrogrammi come metodo per la rivelazione di onde gravitazionali

Laureando
Lorenzo Pierini

Relatori
Pia Astone
Sergio Frasca



SAPIENZA
UNIVERSITÀ DI ROMA



SAPIENZA
UNIVERSITÀ DI ROMA

Filtri tempo-frequenza basati su spettrogrammi come metodo per la rivelazione di onde gravitazionali

Facoltà di Scienze Matematiche, Fisiche e Naturali
Dipartimento di Fisica
Corso di laurea magistrale in Fisica

Lorenzo Pierini
Matricola 1177700

Relatori
Pia Astone
Sergio Frasca

A.A. 2017-2018

Indice

Introduzione	1
1 Tipologie di onde gravitazionali	3
1.1 Meccanismo di generazione delle onde gravitazionali	4
1.2 Segnali dalla coalescenza di binarie compatte	7
1.3 Segnali da stelle di neutroni rotanti	11
1.3.1 Meccanismo di emissione principale	13
1.3.2 Modi di oscillazione di stelle instabili	15
1.3.3 Generalizzazione della dinamica dello spin-down	20
1.4 Segnali di Burst da esplosioni di supernova	21
2 Attuali strategie di rivelazione	24
2.1 Segnali discreti: il problema del campionamento	24
2.2 Caratterizzazione del rumore	26
2.3 Rivelazione ottimale: filtro adattato	31
2.3.1 Applicazione ai chirp	33
2.4 Analisi dati tramite spettrogrammi	36
2.5 Ricerca di segnali continui da stelle compatte asimmetriche	39
2.5.1 Il metodo della trasformata Hough	40
2.6 Ricerca di segnali da r-mode instabili	46
3 Migliorare il SNR: filtraggio in 2D	52
3.1 Il filtro di Wiener	53
3.2 Generalizzazione alle immagini: trasformata di Fourier in 2D	55
3.2.1 Proprietà della trasformata di Fourier in 2D	56
3.2.2 Esempi di trasformate in 2D	58
3.3 Spettri discreti in 2D	64
3.3.1 Dai pixel alle linee rette	64
3.3.2 Spettri di linee curve	67
3.4 Realizzazione di filtri in 2D	70
3.4.1 Filtri di Wiener in 2D	70
3.4.2 Filtro triangolare	74
3.4.3 Finestratura gaussiana	78
3.4.4 Filtro periodico	81

3.4.5	Il limite dei filtri	84
4	Applicazione dei filtri alla ricerca di segnali da r-mode	86
4.1	Studio della risposta ai filtri triangolari tramite il SNR	86
4.1.1	Studio della risposta in funzione del taglio in frequenza	87
4.1.2	Studio della risposta in funzione dell'ampiezza del rumore	90
4.1.3	Studio della risposta in funzione delle inclinazioni di taglio	93
4.2	Effetto del filtro triangolare sulla costruzione delle peakmap	96
	Conclusioni	103
	Bibliografia	105

Introduzione

Il lavoro presentato in questo elaborato vuole inserirsi nel contesto della ricerca delle onde gravitazionali attraverso rivelatori interferometrici. I rivelatori attualmente esistenti, che fanno riferimento alla collaborazione LIGO-Virgo, hanno terminato i primi due run scientifici di osservazione dell'era "advanced", denominate Observing Run 1 e 2 (O1 ed O2), e si apprestano ad iniziarne una nuova, denominata Observing Run 3 (O3). Durante i due run sono stati osservati complessivamente 11 eventi di onda gravitazionale, tutti consistenti in transienti brevi generati dalla coalescenza di due stelle compatte [1]. Altri tipi di segnale sono attualmente oggetto di ricerca; tra questi particolare attenzione viene qui data a segnali continui generati dalla presenza di asimmetrie in stelle di neutroni in rapida rotazione e segnali transienti lunghi generati da instabilità assiali nelle stelle di neutroni. Le procedure che sono state concepite per la rivelazione di questo genere di segnali devono scontrarsi con il problema della sensibilità: le onde gravitazionali oggetto di questa ricerca sono infatti estremamente più deboli rispetto a quelle che si sono finora osservate e solo in alcune configurazioni la loro ampiezza fa sì che possano emergere dal rumore che caratterizza la misura. Scopo dell'elaborato è di presentare delle tecniche di filtraggio che possono in futuro essere implementate sugli algoritmi che ricercano questo genere di segnali al fine di migliorarne la sensibilità ed aumentare così le possibilità di rivelazione.

Nel Capitolo 1 si introdurranno brevemente i meccanismi che sono alla base della generazione delle onde gravitazionali e si presenteranno le diverse tipologie di onda che attualmente si ritiene di poter osservare. Nel Capitolo 2 si discuteranno i nodi da affrontare relativamente alla rivelazione di tali segnali, dagli accorgimenti da prendere al momento di campionare i segnali in dati discreti alle problematiche relative alla caratterizzazione del rumore che è presente in tutte le misure; si presenteranno quindi le tecniche di base per la rivelazione ottimale di segnali all'interno del rumore, evidenziandone punti di forza e limiti, e si introdurranno le procedure di ricerca gerarchiche che mirano all'osservazione dei segnali periodici da stelle di neutroni asimmetriche e dei transienti lunghi da instabilità nelle stesse, segnali per i quali non sono applicabili le tecniche di rivelazione ottimale.

Nel Capitolo 3 si svilupperanno le tecniche di filtraggio che troveranno applicazione nelle procedure di cui sopra. Queste consistono in filtri che operano in 2 dimensioni (2D), i quali dovranno essere applicati ai dati trasformati nel dominio tempo-frequenza, organizzati in database chiamati spettrogrammi. Per realizzare i filtri in 2D si sono studiate le proprietà ed il funzionamento della trasformata di Fourier generalizzata a 2D: su tale argomento la letteratura scientifica di riferimento è estremamente povera, di conseguenza su di esso è stata impiegata grande parte del lavoro complessivo. Punto di arrivo sarà la progettazione di un particolare filtro in grado di far emergere dal rumore segnali transienti lunghi che nel dominio tempo-frequenza presentano una morfologia lentamente variabile, che prenderà il nome di filtro triangolare.

Nel Capitolo 4 si inizieranno ad applicare i filtri sviluppati alla ricerca di particolari transienti lunghi

prodotti da instabilità nelle stelle di neutroni, chiamati r-mode, con la finalità di mostrare il vantaggio e la possibilità pratica di inserire tali tecniche all'interno della procedura di ricerca attualmente esistente. Si comincerà con lo studiare la configurazione ideale con la quale il filtro triangolare agisce su questo tipo di segnali ed il guadagno che la sua applicazione porta in termini di aumento del rapporto segnale/rumore; si chiuderà applicando tali filtri ad un segnale simulato per mostrare gli effetti che essi producono al primo passo della procedura di selezione di candidati eventi dell'algoritmo di ricerca.

Capitolo 1

Tipologie di onde gravitazionali

Con la rivelazione della prima onda gravitazionale avvenuta il 14 Settembre 2015 ad opera dei due interferometri di LIGO situati ad Hanford e Livingston, il cui evento è stato denominato GW150914 [2], si è ufficialmente inaugurato un nuovo modo di osservare l'universo, in grado di fornire informazioni fino ad allora inaccessibili e di osservare fenomeni altrimenti invisibili utilizzando la sola radiazione elettromagnetica. E' il caso dell'evento sopracitato e dei diversi altri che sono stati osservati successivamente: per essi il segnale è stato generato dalla coalescenza di una binaria di buchi neri di diverse masse solari, evento che per sua natura non porta all'emissione di radiazione elettromagnetica e che per questa ragione era rimasto inosservato fino ad ora. Un altro traguardo importante è stato raggiunto nell'Agosto 2017 quando l'interferometro italiano Virgo, situato a Cascina, si è unito alla rete con, tra i vari obiettivi, quello di migliorare drasticamente la precisione nel localizzare le sorgenti dei segnali osservati. L'operazione ha pagato immediatamente i dividendi: infatti il 17 Agosto 2017, con i tre interferometri attivi, è stato osservato per la prima volta un segnale che è stato attribuito alla coalescenza di due stelle di neutroni, sistemi che presentano una intensa emissione elettromagnetica. L'evento, denominato GW170817, è stato localizzato nel cielo in una regione di 28 deg^2 alla distanza di 40 Mpc : si tratta di una regione quasi 7 volte più piccola di quella che sarebbe stata individuata dai soli interferometri di LIGO. Grazie a questa precisione nella localizzazione è stato possibile verificare la coincidenza con un γ -ray burst giunto $1.7s$ dopo la coalescenza dei due oggetti, proveniente dalla stessa direzione ma in una regione più ampia, osservato dai satelliti Fermi ed INTEGRAL ed etichettato con l'evento GRB170817A [3]. Questa eccezionale coincidenza ha aperto la strada alla cosiddetta astronomia multimessaggero, la quale consiste nell'osservazione di fenomeni astrofisici attraverso più vettori di informazione.

Allo stato attuale tutti i segnali gravitazionali osservati sono riconducibili alla coalescenza di binarie di stelle compatte, la quale produce segnali estremamente energetici e la cui morfologia è predetta con grande precisione, dunque di più facile rilevazione in rapporto al livello di rumore dei detector. Con il progressivo miglioramento delle sensibilità degli strumenti e delle tecniche di elaborazione dei dati in uscita si sta aprendo alla possibilità di rilevare la presenza di altri tipi di onde gravitazionali, finora non rilevate perché troppo deboli rispetto al rumore dei detector o dalla morfologia meno prevedibile. In questo capitolo si discuteranno le differenti sorgenti di onde gravitazionali sulle quali sono attualmente focalizzate le ricerche. Come sarà evidente in seguito, a sorgenti differenti corrisponderanno segnali assai differenti per morfologia, prevedibilità e durata: tali caratteristiche concorreranno a determinare i metodi ottimali per la loro rivelazione. Come prima cosa però saranno richiamati i concetti che sono alla base della generazione delle onde gravitazionali.

1.1 Meccanismo di generazione delle onde gravitazionali

La previsione che un sistema di masse accelerate in modo da violare determinate simmetrie del sistema produca increspature nello spaziotempo che si propagano alla velocità della luce rientra in quell'insieme di effetti previsti dalla Relatività Generale che non hanno una controparte nella fisica classica pre-relativistica. La teoria della Relatività Generale, pubblicata da Albert Einstein nel 1916, risponde alla richiesta -maturata con la formulazione special relativistica dei fenomeni elettromagnetici- di avere una descrizione della gravità che fosse espressa in forma locale e non come azione a distanza e che inoltre incorporasse in sé il principio di relatività ed il principio di equivalenza.

Già dalla formulazione della Relatività Speciale era stato chiaro che spazio e tempo avrebbero dovuto perdere la loro indipendenza per andare a fondersi in una unica entità, lo spaziotempo, la cui metrica era specificata dal tensore metrico che nella versione piatta assumeva la semplice forma matriciale $\eta_{\mu\nu} = \text{diag}\{-1, 1, 1, 1\}$: attraverso esso è possibile calcolare l'intervallo tra due eventi dello spaziotempo. Nel passaggio dalla Relatività Speciale alla Relatività Generale il tensore metrico continua ad essere il protagonista, assumendo però una forma analiticamente più complessa, etichettata come $g_{\mu\nu}$, la quale rende conto dello spaziotempo curvo che rappresenta. Un'altra entità che svolge un ruolo centrale sono le connessioni affini $\Gamma_{\alpha\beta}^{\mu}$, che determinano l'evoluzione di quantità vettoriali e tensoriali quando lo spaziotempo è curvo. Consideriamo una base coordinata $\{\vec{e}_{(\mu)}\}$: un generico vettore può essere considerato nelle sue componenti su tale base: $\vec{V} = V^{\sigma}\vec{e}_{(\sigma)}$, ma mentre in uno spaziotempo piatto i vettori della base rimangono gli stessi in qualunque regione ci si trovi, in uno spaziotempo curvo questi in linea di principio cambieranno da punto a punto; dunque nel calcolare la variazione di tale vettore bisogna tenere conto della variazione sia delle sue componenti, sia dei vettori della base, essendo questi relativi ad uno spaziotempo curvo: $\frac{\partial \vec{V}}{\partial x^{\beta}} = \frac{\partial V^{\alpha}}{\partial x^{\beta}} \vec{e}_{(\alpha)} + V^{\alpha} \frac{\partial \vec{e}_{(\alpha)}}{\partial x^{\beta}}$.

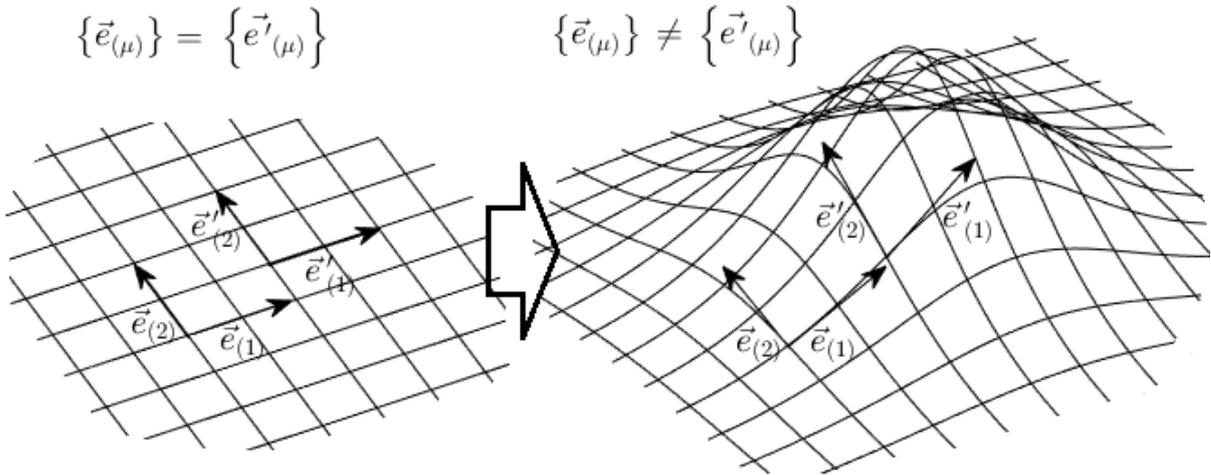


Figura 1.1: Illustrazione in 2D del passaggio da una varietà piatta ad una varietà con curvatura. Mentre in uno spaziotempo piatto un insieme di vettori di base rimane immutato indipendentemente dalla regione in cui ci si trova, in uno spaziotempo curvo i vettori di base possono cambiare da regione a regione, rendendo impossibile la definizione di un sistema di riferimento globale.

Essendo la derivata $\frac{\partial \vec{e}_{(\alpha)}}{\partial x^{\beta}}$ essa stessa un vettore, è possibile esprimerla in funzione dei vettori di base stessi

attraverso le connessioni affini: $\frac{\partial \vec{e}^{(\alpha)}}{\partial x^\beta} = \Gamma_{\alpha\beta}^\mu \vec{e}^{(\mu)}$. Queste rappresentano allora la variazione dei vettori della base coordinata dovuta alla curvatura dello spaziotempo, e sono esprimibili come combinazione lineare delle derivate prime del tensore metrico: $\Gamma_{\alpha\beta}^\sigma = \frac{1}{2} g^{\sigma\mu} (g_{\mu\alpha,\beta} + g_{\mu\beta,\alpha} - g_{\alpha\beta,\mu})$. Nel complesso, la curvatura dello spaziotempo può essere espressa in maniera compatta attraverso il tensore di curvatura di Riemann, il quale è l'unico tensore che può essere costruito utilizzando la metrica, le sue derivate prime e seconde e che risulti lineare nelle derivate seconde: $R_{\mu\beta\nu}^\alpha = -(\Gamma_{\mu\beta,\nu}^\alpha - \Gamma_{\mu\nu,\beta}^\alpha + \Gamma_{\nu\eta}^\alpha \Gamma_{\mu\beta}^\eta - \Gamma_{\beta\eta}^\alpha \Gamma_{\mu\nu}^\eta)$. L'equazione di campo di Einstein consente di determinare tale tensore a partire dalla distribuzione di massa-energia del sistema: utilizzando la versione contratta su primo e terzo indice, $R_{\mu\nu} = R_{\mu\alpha\nu}^\alpha$, chiamata tensore di Ricci, ed introdotto per una maggiore compattezza il tensore di Einstein $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu}$, l'equazione di campo di Einstein mette in diretta relazione le proprietà geometriche dello spaziotempo con la distribuzione di massa-energia [4]:

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \quad (1.1)$$

Il tensore $T_{\mu\nu}$ che compare nel membro a destra è chiamato tensore energia-impulso, ed è definito per un sistema di n particelle localizzate rispettivamente in $\vec{\xi}_n$, non interagenti, in funzione dei quadrimpulsi p_n^μ , delle energie E_n e delle coordinate $\vec{\xi}$ del sistema di riferimento scelto:

$$T^{\mu\nu} = c^2 \sum_n \frac{p_n^\mu p_n^\nu}{E_n} \delta^3(\vec{\xi} - \vec{\xi}_n(t)) \quad (1.2)$$

La 1.1 rappresenta un sistema di 10 equazioni differenziali miste del secondo ordine ed in linea di principio tramite queste il tensore metrico $g_{\mu\nu}$ è completamente determinato in tutte le sue componenti. Nella pratica tuttavia un sistema del genere è troppo complicato per essere risolto in tutta generalità: le uniche soluzioni esatte note sono relative a distribuzioni di massa energia con condizioni di simmetria molto restrittive, come ad esempio la soluzione di Schwarzschild, statica ed a simmetria sferica. La maggior parte delle soluzioni note sono invece in forma approssimata.

La soluzione di propagazione ondosa rientra per l'appunto in questa categoria: si è arrivati infatti a comprendere che qualora lo spaziotempo sia debolmente perturbato rispetto alla metrica piatta, ovvero si possa esprimere

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \quad , \quad |h_{\mu\nu}| \ll 1 \quad (1.3)$$

allora il sistema 1.1 può essere linearizzato e la perturbazione a traccia opposta, che si definisce come $\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2} \eta_{\mu\nu} h$ in modo che $\bar{h} = -h$, si propaga nello spaziotempo secondo l'equazione

$$\begin{cases} (-\partial_0^2 + \nabla^2) \bar{h}_{\mu\nu} = -\frac{16\pi G}{c^4} T_{\mu\nu} \\ \partial_\mu \bar{h}_\nu^\mu = 0 \end{cases} \quad (1.4)$$

Si può vedere facilmente che in regioni in cui non sono presenti le sorgenti, ovvero in cui $T_{\mu\nu} = 0$, la 1.4 assume la forma classica dell'equazione delle onde, con l'aggiunta della condizione di quadridivergenza nulla come specifica scelta di gauge armonica. La soluzione generale di tale sistema rimane tuttavia ancora troppo complicata, soprattutto quando le sorgenti assumono distribuzioni particolarmente irregolari nello spazio e nel tempo. La tecnica principe che si applica in queste situazione consiste nello sviluppare il tensore energia-impulso in serie di multipoli, in analogia con il modo di procedere dell'elettromagnetismo. A differenza del caso elettromagnetico, tuttavia, il primo termine dello sviluppo che produce

radiazione gravitazionale non è il dipolo: questo corrisponde fisicamente al momento totale, il quale in un sistema isolato è una quantità conservata e quindi statica. Il primo termine dello sviluppo a determinare l'emissione gravitazionale è il momento di quadrupolo, definito nelle sole componenti spaziali come:

$$q^{ij} = \frac{1}{c^2} \int_V d^3x T^{00}(t, x^k) x^i x^j \quad (1.5)$$

dove il termine temporale del tensore energia-impulso rappresenta la densità di massa-energia del sistema. Si preferisce tuttavia utilizzare il momento di quadrupolo ridotto, che è la versione a traccia nulla dello stesso, definita come $Q_{ij} = q_{ij} - \frac{1}{3} \delta_{ij} q_m^m$: questo differisce dalla 1.5 per termini costanti che si cancellano in derivazione (si vedano le 1.6, 1.7), dunque la sostituzione è legittima. Si fa questo perché la quantità q^{ij} non è quella direttamente osservabile, mentre la Q^{ij} rappresenta il termine in r^{-3} nello sviluppo in serie di potenze di r^{-1} del potenziale newtoniano in una regione dello spaziotempo abbastanza vicina alla sorgente ma in cui la gravità sia quasi newtoniana:

$$\phi \sim -\frac{M}{r} + \frac{d_j n^j}{r^2} + \frac{3Q_{jk} n^j n^k}{2r^3} + \dots$$

In questo tipo di prima approssimazione, che prende il nome di formalismo di quadrupolo e che rimane valida fino a quando la lunghezza d'onda emessa sia molto maggiore delle dimensioni del sistema, la forma d'onda emessa è non nulla nelle sole componenti spaziali ed è data da:

$$\begin{cases} \bar{h}^{ij}(t, r) = \frac{2G}{c^4 r} \frac{\partial^2}{\partial t^2} Q^{ij} \left(t - \frac{r}{c}\right) \\ \bar{h}^{\mu 0} = 0 \end{cases} \quad (1.6)$$

L'energia trasportata dall'onda, che viene quindi sottratta al sistema, è data invece da:

$$L_{GW} = \frac{dE_{GW}}{dt} = \frac{G}{5c^5} \left\langle \sum_{jk} \ddot{Q}_{jk} \left(t - \frac{r}{c}\right) \ddot{Q}_{jk} \left(t - \frac{r}{c}\right) \right\rangle \quad (1.7)$$

A seconda del tipo di sorgente e della fase dell'evoluzione in cui si trova, le equazioni 1.6, 1.7 possono risultare sufficienti a determinare l'evoluzione del sistema e la forma d'onda emessa, oppure risultare insufficienti: in quel caso occorrerà andare oltre il formalismo di quadrupolo e considerare anche i multipoli successivi. In questi casi il framework in cui ci si trova è solitamente quello del formalismo Post-Newtoniano, nel quale si considerano sviluppi successivi della 1.1 sia in $G_{\mu\nu}$ che in $T_{\mu\nu}$ [5].

Possiamo dividere i segnali che arrivano ai rivelatori in due grandi categorie:

- **Segnali transienti:** si tratta di segnali che vengono rivelati per un periodo di tempo limitato. Rientrano in questa categoria ad esempio i chirp, ovvero i segnali generati dalla coalescenza di binarie di stelle compatte, i burst generati dalle esplosioni di supernova ed i segnali cosiddetti postmerger, generati qualora dalla coalescenza di due stelle di neutroni si produca una stella supermassiva e quindi instabile.
- **Segnali continui:** sono segnali che sono sempre presenti nei dati dei rivelatori. Rientrano in questa categoria i segnali periodici, generati dalla rotazione di stelle di neutroni che presentano una deformazione rispetto alla simmetria sferica, ed il background stocastico, formato dalla moltitudine di piccole onde generate dai vari corpi celesti che vanno a sommarsi.

1.2 Segnali dalla coalescenza di binarie compatte

Come già detto, il segnale proveniente da binarie in coalescenza è stato l'unico finora rivelato. Quando due stelle compatte -come ad esempio buchi neri, stelle di neutroni, nane bianche- sono in orbita reciproca, si è nella situazione in cui il momento di quadrupolo del sistema ha derivate temporali seconde e terze non nulle. Il sistema allora emetterà radiazione gravitazionale che andrà a diminuire l'energia orbitale del sistema, il quale di conseguenza descriverà orbite sempre più strette e con periodi sempre più brevi fino ad arrivare alla coalescenza. In linea di principio il profilo d'onda $h(t)$ conterrà tutti i parametri attraverso i quali è possibile descrivere il sistema. Questi sono:

- Localizzazione: distanza di luminosità D_L ; ascensione retta α ; declinazione δ .
- Orientamento: angolo ι tra il momento angolare orbitale e la linea visiva (in alternativa angolo θ_{JN} tra il momento angolare totale e la linea visiva); angolo di polarizzazione dell'onda ψ , che rappresenta l'orientamento della proiezione del momento orbitale sul piano visivo.
- Tempo t_c e fase φ_c di coalescenza.
- Caratteristiche intrinseche: masse $\{m_1, m_2\}$ e spin $\{\vec{S}_1, \vec{S}_2\}$ (grandezze vettoriali).

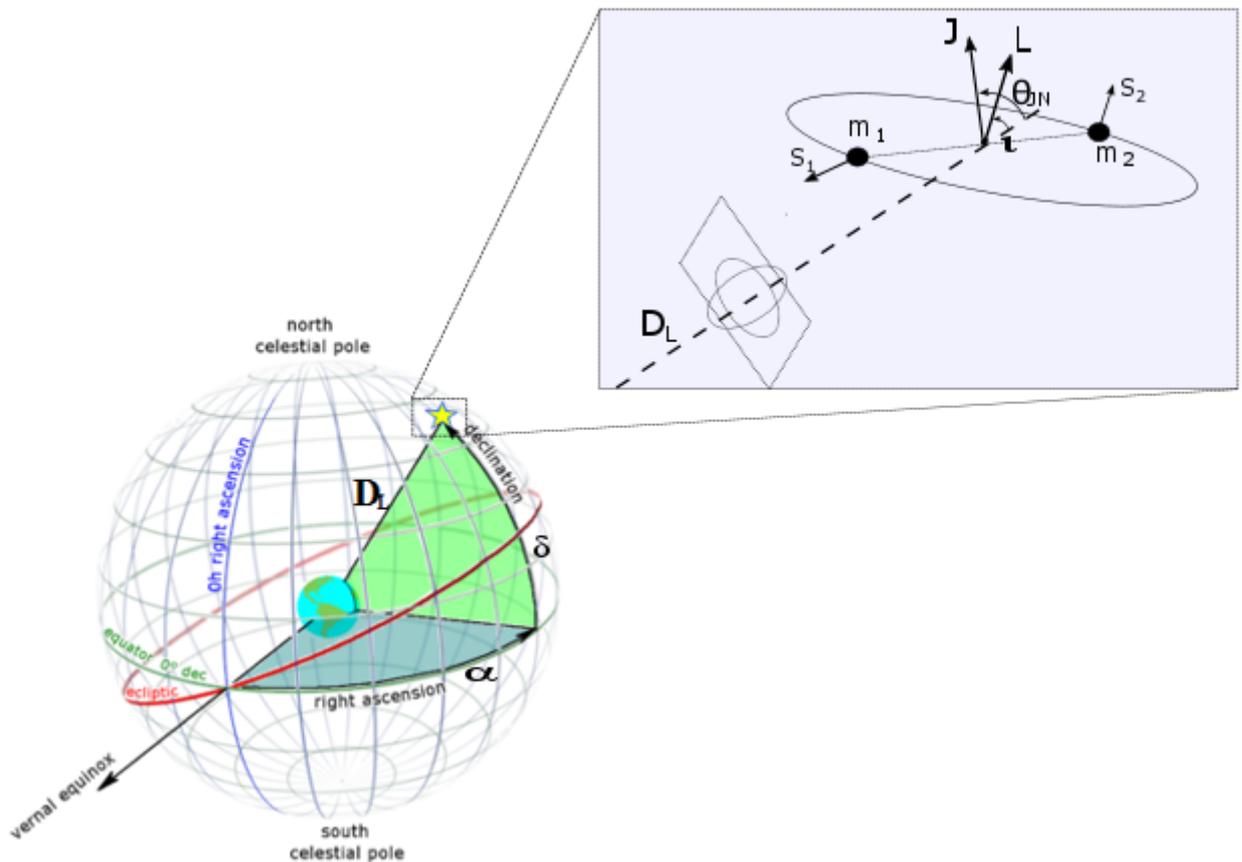


Figura 1.2: Rappresentazione dei parametri attraverso i quali è possibile descrivere un sistema binario

Va detto che non tutti questi parametri sono contenuti nella forma d'onda emessa vera e propria: bisogna infatti tener da conto che l'onda gravitazionale arriva ai detector con 2 polarizzazioni possibili, che etichettiamo con h_+ ed h_\times , alle quali gli interferometri daranno risposte in linea di principio differenti. Esprimiamo allora il segnale osservato dal detector (D) come combinazione delle due polarizzazioni [6]:

$$h^{(D)}(t) = F_+^{(D)} h_+(t) + F_\times^{(D)} h_\times(t)$$

dove $F_+^{(D)}(\alpha, \delta, \psi)$ e $F_\times^{(D)}(\alpha, \delta, \psi)$ sono le funzioni di risposta del detector alla corrispondente onda polarizzata. Queste dipendono dalla localizzazione angolare della sorgente nella sfera celeste, per la quale si utilizza il sistema di coordinate equatoriali α (ascensione retta, analoga alla longitudine) e δ (declinazione, analoga alla latitudine), e dall'angolo di polarizzazione ψ .

Ad eccezione di quei parametri la cui informazione risiede nelle funzioni di risposta $F_{+,\times}$, gli altri dovranno essere desunti dalla forma d'onda. Diventa perciò importante disporre di modelli d'onda sempre più accurati. Vediamo dunque come si compongono questi modelli.

La coalescenza di due sistemi astrofisici in orbita binaria può essere schematizzata in 3 fasi distinte, che si tradurranno in 3 soluzioni distinte della forma d'onda, le quali dovranno raccordarsi agli estremi per dare continuità al segnale. La prima fase, detta di spiraleggiamento (o inspiral), è caratterizzata da orbite che per ampiezza e velocità risultano ben descrivibili all'interno dell'approssimazione newtoniana o di sue soluzioni perturbate (approssimazione Post-Newtoniana). La forma d'onda che emerge è caratterizzata da una crescita nel tempo in ampiezza e frequenza e per questa sua caratteristica, analoga al cinguettio degli uccelli, questa classe di forme d'onda viene chiamata di chirp. Nella seconda fase, detta di fusione (o merger), si ha la fusione dei due oggetti in un unico sistema la cui natura dipende dai due iniziali. Abbiamo imparato a risolvere questa fase solo di recente, e le soluzioni d'onda dipendono pesantemente dalla natura dei due sistemi di partenza. Nella terza fase, detta di ringdown, si ha il sistema risultato della fusione, il quale tuttavia presenta una forma altamente asimmetrica: esso entrerà allora in oscillazione, secondo modi quasi-normali (QNM) per assestarsi attorno alla simmetria sferica o cilindrica. La forma d'onda risultante sarà tipicamente una sovrapposizione di sinusoidi con attenuazione esponenziale.

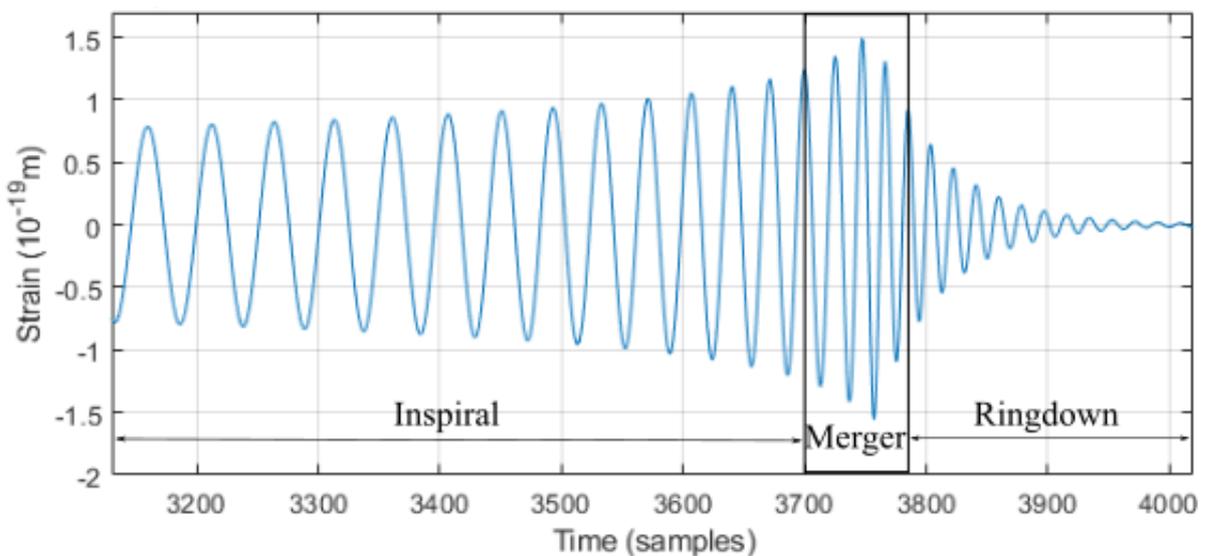


Figura 1.3: Template di una forma d'onda completa comprendente le 3 fasi. Fonte: LIGO Data Release

Fase inspiral

Come anticipato, in questa fase della coalescenza le orbite sono ancora abbastanza larghe e le velocità in gioco abbastanza basse da permettere di utilizzare l'approssimazione Newtoniana. In questo schema si calcola l'onda emessa utilizzando il formalismo di quadrupolo, che approssima i due sistemi a due punti materiali. In questa prima approssimazione la 1.6 porta alla seguente soluzione [4, 5]:

$$v_{GW}(t) = \frac{5^{3/8}}{8\pi} \left(\frac{c^3}{G\mathcal{M}} \right)^{5/8} (t_c - t)^{-3/8}$$

$$\phi(t) = -2 \left[\frac{c^3}{5G\mathcal{M}} (t_c - t) \right]^{5/8} + \phi_c$$

$$h_0(t) = -4c \left(\frac{G\mathcal{M}}{c^3} \right)^{5/3} (\pi v_{GW}(t))^{2/3}$$

$$h(t) = \frac{h_0(t)}{D_L} e^{i\phi(t)}$$

dove i parametri sono quelli introdotti precedentemente e si sono compattate nella forma esponenziale le due polarizzazioni, riottenibili separatamente prendendone la parte reale e la parte immaginaria. In questa espressione come nelle successive, le masse non compaiono distinte ma verranno parametrizzate attraverso la massa di chirp $\mathcal{M} = \mu^{3/5} M^{2/5}$ che emerge naturalmente dai calcoli del formalismo di quadrupolo. Come si vede, al fine di individuare i parametri fisici intrinseci del sistema questa approssimazione non è ancora soddisfacente: compare solamente la massa di chirp, tramite la quale possiamo stabilire dei limiti per la massa totale della binaria ma non specificare le due masse singolarmente. Inoltre non compare in alcun modo il contributo degli spin, come conseguenza dell'aver approssimato i corpi come puntiformi.

Per fare di meglio occorre considerare il contributo di effetti Post-Newtoniani, ovvero correzioni di un certo ordine alla forma d'onda Newtoniana. Più precisamente, contrassegniamo con l'ordine $P^x N$ correzioni alla formula di quadrupolo di ordine $\left(\frac{M}{r}\right)^x$. Queste correzioni includono termini di multipolo di ordine maggiore al tensore energia-impulso, che portano a correzioni nelle espressioni dell'ampiezza, dell'energia e della luminosità. Questo ultimo fatto è particolarmente importante, in quanto una correzione alla perdita di energia va a modificare il ritmo dell'inspiral e con esso la fase orbitale accumulata, che è la parte dell'onda da cui otteniamo le misure più accurate. Infatti i treni d'onda gravitazionali, nel range di frequenze di sensibilità degli interferometri, contengono un numero \mathcal{N} elevato cicli che si accumulano nell'evoluzione della fase: questo comporta misure dalla fase con un errore frazionale di un fattore $\sim \mathcal{N}^{-1}$ rispetto alle misure dall'ampiezza. Come conseguenza, nonostante in generale con gli sviluppi PN la forma d'onda si complichino parecchio, si ricorre spesso ad una espressione più com-

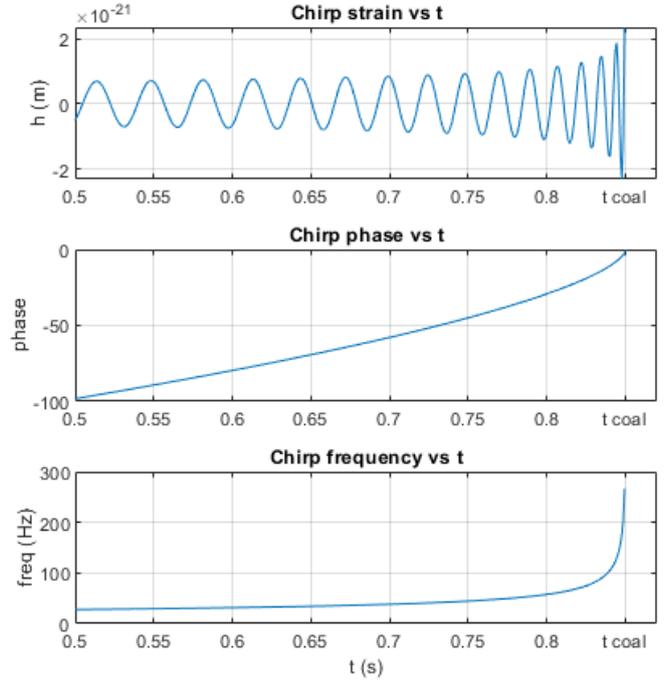


Figura 1.4: Esempio di forma d'onda di chirp.

patta in cui l'ampiezza rimane imperturbata e si introducono le correzioni solamente nella fase, ovvero esprimendola nella forma :

$$h(t) = h^{(0)}(t) e^{2i[\phi^{(0)}(t) + \phi^{(1)}(t) + \phi^{(1.5)}(t) + \phi^{(2)}(t) + \dots]}$$

dove con $h^{(0)}(t)$ si intende l'ampiezza calcolata all'ordine 0 dello sviluppo, ovvero la formula di quadrupolo. Tramite questi sviluppi vediamo che già nell'ordine P^1N compaiono i parametri di massa ridotta e totale $\{\mu, M\}$ consentendo di risolvere completamente le masse singole dei corpi, mentre il primo termine che contiene l'informazione degli spin compare nell'ordine $P^{1.5}N$:

$$\tilde{h}(f) = \frac{Q}{D_L} \mathcal{M}^{5/6} f^{-7/6} e^{i\Psi(f)} \quad (1.8)$$

$$\Psi(f) = 2\pi t_c f - \phi_c + \frac{3}{128} (\pi \mathcal{M} f)^{-5/3} \left[1 + \frac{20}{9} \left(\frac{743}{336} + \frac{11\mu}{4M} \right) x + (4\pi - \beta) x^{3/2} + O(x^2) \right] \quad (1.9)$$

$$\beta = \frac{1}{M^2} \hat{L} \cdot \left[\left(\frac{113}{12} + \frac{25}{4} \frac{m_2}{m_1} \right) \vec{S}_1 + \left(\frac{113}{12} + \frac{25}{4} \frac{m_1}{m_2} \right) \vec{S}_2 \right] \quad (1.10)$$

Per una maggiore semplicità la soluzione è stata espressa nel dominio delle frequenze, avendo l'accortezza di utilizzare lo sviluppo in termini di $x \equiv (\pi M f)^{2/3}$, che presenta il vantaggio di essere direttamente misurabile e gauge-indipendente: i suoi ordini di potenza corrispondono dunque all'ordine di sviluppo Post-Newtoniano. Come si vede, fino all'ordine $P^{1.5}N$ l'effetto dei vettori \vec{L} , \vec{S}_1 , \vec{S}_2 compare solo nella combinazione data dal parametro β . Qualora nella binaria ci siano stelle di neutroni occorrerà tenere conto anche di termini di quadrupolo e di marea. Rietichettando la fase espressa nella 1.9 come contributo point-particle $\Psi_{PP}(f)$ in cui si aggiungerà un termine σ in x^2 che conterrà termini di interazione spin-spin e spin-orbita, la fase totale che compare in 1.8 sarà allora data dalla somma $\Psi(f) = \Psi_{PP}(f) + \Psi_Q(f) + \Psi_\lambda(f)$, con [7]:

$$\Psi_{PP}(f) = 2\pi t_c f - \phi_c + \frac{3}{128} (\pi \mathcal{M} f)^{-5/3} \left\{ 1 + \frac{20}{9} \left(\frac{743}{336} + \frac{11\mu}{4M} \right) x + (4\pi - \beta) x^{3/2} + \left[\frac{15293365}{508032} + \frac{27145}{504} \frac{\mu}{M} + \frac{3085}{72} \left(\frac{\mu}{M} \right)^2 - 10\sigma \right] x^2 + O(x^{5/2}) \right\}$$

$$\Psi_Q(f) = \frac{3}{128} (\pi \mathcal{M} f)^{-5/3} \left\{ -50 \left[\left(\frac{m_1^2}{m^2} \chi_1^2 + \frac{m_2^2}{m^2} \chi_2^2 \right) (Q_S - 1) + \left(\frac{m_1^2}{m^2} \chi_1^2 - \frac{m_2^2}{m^2} \chi_2^2 \right) Q_a \right] x^2 \right\}$$

$$\Psi_\lambda(f) = -\frac{3}{128} (\pi \mathcal{M} f)^{-5/3} \left\{ 24 \left[\left(1 + 7 \frac{\mu}{M} - 31 \left(\frac{\mu}{M} \right)^2 \right) \lambda_S + \left(1 + 9 \frac{\mu}{M} - 11 \left(\frac{\mu}{M} \right)^2 \right) \lambda_a \frac{m_1 - m_2}{m} \right] x^5 \right\} \quad (1.11)$$

dove i coefficienti $Q_S = \frac{1}{2} (\bar{Q}_1 + \bar{Q}_2)$, $Q_a = \frac{1}{2} (\bar{Q}_1 - \bar{Q}_2)$ sono combinazioni dei momenti di quadrupolo delle due stelle ed i parametri $\lambda_S = \frac{1}{2} (\bar{\lambda}_1 + \bar{\lambda}_2)$, $\lambda_a = \frac{1}{2} (\bar{\lambda}_1 - \bar{\lambda}_2)$ sono i numeri di Love, che contengono informazioni riguardo l'equazione di stato della materia nel core della stella di neutroni [8]. In definitiva, tornando al segnale che arriva ai detector con due differenti polarizzazioni, possiamo schematizzare le forme d'onda nella seguente maniera [9]:

$$\begin{aligned} \tilde{h}_+(f) &= A (1 + \cos^2 \iota) f^{-7/6} e^{i\Psi(f)} \\ \tilde{h}_\times(f) &= 2A \cos \iota f^{-7/6} e^{i\Psi(f) - i\pi/2} \end{aligned} \quad (1.12)$$

dove l'espressione delle $\Psi(f)$ dipende dall'ordine di sviluppo Post-Newtoniano cui si arriva.

Fase merger

La soluzione analitica del chirp cessa di essere valida approssimativamente quando la separazione orbitale tra i due sistemi raggiunge il valore dell'ISCO (Inner-most Stable Circular Orbit) dato da $r = 6M$, la quale corrisponde ad un valore della frequenza relativa al limite Newtoniano:

$$f_{GW}^{(ISCO)} = \frac{1}{6^{3/2}\pi M}$$

Quello che avviene oltrepassato questo punto dipende in gran parte dalla natura dei sistemi in questione: nel caso di due stelle di neutroni (BNS), la separazione orbitale dell'ISCO coincide con le dimensioni fisiche delle stesse, perciò si ha subito la fusione. Nel caso di due buchi neri (BBH), questi hanno ancora un certo numero di orbite a disposizione prima di fondersi; durante questa fase ci si trova in una situazione di campo ultra-forte, nella quale non sono più valide le approssimazioni linearizzate utilizzate fino a quel momento, perciò occorre risolvere esplicitamente le equazioni di campo di Einstein in regime non lineare attraverso metodi numerici computazionali (Numerical Relativity). La parte più problematica è il momento della fusione, ovvero come trattare l'incontro di due orizzonti degli eventi e come gestire le prime oscillazioni dell'unico orizzonte risultante, quando è ancora fortemente asimmetrico. Nel 2005 si è finalmente trovato il modo tecnico di trattare l'evoluzione dello spaziotempo di un BBH in queste fasi, nel regime più generale possibile 3D non assi simmetrico [10]: da allora sappiamo risolvere la fase merger e sono stati creati, al variare delle condizioni iniziali, dei template di forme d'onda attese, per le quali non c'è espressione analitica. Questi template dovranno raccordarsi con le condizioni del sistema al raggiungimento dell'ISCO.

Fase ringdown

Nel caso in cui il risultato della fusione sia un buco nero, in questa ultima fase abbiamo un unico orizzonte degli eventi, con tipicamente un elevato spin di rotazione dato dal momento angolare totale del sistema binario, in oscillazione a causa della forma asimmetrica. In questo caso si schematizza il sistema attraverso la perturbazione di un buco nero di Kerr in presenza di fenomeni dissipativi (emissione di GW). Questi ultimi fanno sì che le frequenze di oscillazione non siano numeri reali, come nei modi normali di oscillazione, ma complessi, con la conseguente comparsa di un fattore di smorzamento nell'oscillazione: si parla in questo caso di modi quasi-normali (QNM). La forma d'onda attesa è di un'oscillazione smorzata del tipo

$$h(t) = A \sin(\omega t) e^{-\frac{t}{\tau}}$$

dove ω dipende unicamente dalla massa e dallo spin del buco nero finale.

Nel caso in cui la fusione interessi una binaria di stelle di neutroni, si apre l'interessante scenario in cui occorre capire se il risultato della fusione è un buco nero o una stella di neutroni, alternativa cui concorrono numerosi fattori e che verrà affrontata nella Sezione 1.3.2.

1.3 Segnali da stelle di neutroni rotanti

L'esistenza delle stelle di neutroni è stata ipotizzata per la prima volta da Landau (1932) e da Baade e Zwicky (1934): a loro va il merito di aver introdotto l'idea che oggetti del genere possano essere ciò che rimane delle esplosioni di supernova. Più nello specifico, secondo le correnti teorie dell'evoluzione

stellare, se una stella possiede una massa nel range $\sim 4M_{\odot} < M < 20 \div 30M_{\odot}$, una volta arrivata al termine del ciclo vitale entrerà in una fase di instabilità che culminerà in una esplosione di supernova, il cui processo sarà descritto sinteticamente nella Sezione 1.4, lasciando come residuo il nucleo della stella originaria. Dobbiamo a Pacinotti, nel 1967, il primo modello teorico di una stella di neutroni.

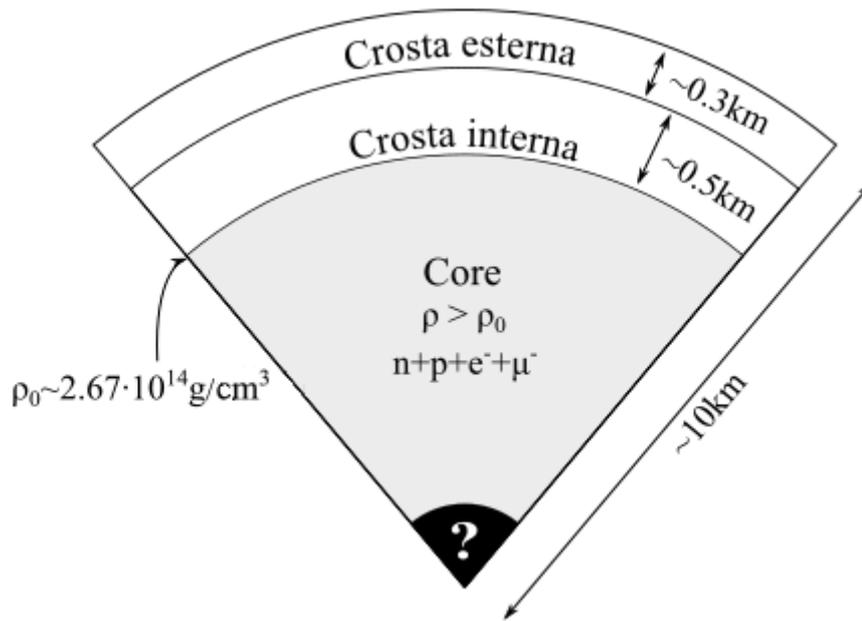


Figura 1.5: Rappresentazione fuori scala della struttura interna di una stella di neutroni

In base alla visione che ne abbiamo attualmente, possiamo descrivere la struttura stellare riconoscendo 3 fasi distinte:

- **Crosta esterna:** spessa $\sim 0.3\text{km}$, in questa fase la materia è ancora organizzata in nuclei atomici, legati tra di loro in un reticolo rigido. Avvicinandosi verso il centro, il processo β inverso $e^- + p \rightarrow n + \nu_e$ diviene sempre più efficiente andando a liberare neutroni.
- **Crosta interna:** spessa $\sim 0.5\text{km}$, in essa si ha la coesistenza di due fasi diverse della materia: gas di neutroni (NG) e materia ricca di protoni (Proton Rich Matter, PRM), dove la PRM, immersa nel NG, si organizza in maniera differente all'aumentare della densità, iniziando come gocce immerse nel gas, per poi aggregarsi prima in spaghetti (pasta facing) ed infine unirsi in piani (lasagne).
- **Core:** in questa fase la densità ha superato la densità di saturazione dei nuclei $\rho_0 = 2.67 \cdot 10^{14} \text{g/cm}^3$ e la materia è organizzata come un fluido omogeneo composto approssimativamente da un 90% di neutroni ed un 10% di protoni, con inoltre un numero di elettroni simile ai protoni, il tutto in equilibrio β , ovvero in una situazione in cui le due reazioni si equilibrano:

$$\begin{aligned} \beta : & \quad \left\{ \begin{array}{l} n \rightarrow e + p + \bar{\nu}_e \\ \beta^{-1} : \quad \left\{ \begin{array}{l} e + p \rightarrow n + \nu_e \end{array} \right. \end{array} \right. \end{aligned}$$

è presente inoltre una frazione di muoni, che si vengono a formare quando la densità raggiunge valori ai quali il potenziale chimico degli elettroni diventa più grande della massa a riposo dei muoni, rendendoli quindi energeticamente più convenienti degli elettroni nel decadimento β .

Nella parte più interna del core le densità tipiche possono arrivare a valori $\rho \gtrsim 5\rho_0$: a questi livelli le interazioni tra adroni non possono essere trascurate e vanno trattate nell'ambito della cromodinamica quantistica (QCD). Esiste tuttavia una grande varietà di modelli possibili, i quali conducono a differenti composizioni interne e conseguentemente a differenti equazioni di stato (EoS) per la stella nel suo complesso [11]. Differenti EoS comportano, a parità di massa, diversi valori per il raggio stellare: la sua stima per una stella di neutroni perciò può dirci molto su quale sia l'effettivo comportamento della materia nucleare in condizioni estreme, dato che i livelli di densità cui questi fenomeni diventano rilevanti sono irraggiungibili in esperimenti a terra. Tuttavia effettuare una misura precisa delle dimensioni di una stella di neutroni è una impresa tutt'altro che facile: queste il più delle volte vengono osservate come pulsar, ovvero sorgenti radio che emettono con una periodicità stretta ed estremamente precisa. Questo avviene a causa della loro elevata velocità di rotazione, che associata agli elevati campi magnetici in gioco ($B \sim 10^{11} - 10^{13} \text{ Gauss}$) comporta l'emissione di raggi di onde radio dai poli magnetici: dato che tipicamente l'asse magnetico non è allineato con l'asse di rotazione, tale raggio diventa visibile solamente quando punta, periodicamente, in direzione della Terra. Dall'analisi di tale segnale è possibile inferire molte proprietà della sorgente come il periodo di rotazione e lo spindown, ma non le sue dimensioni fisiche. Un'altra possibilità per ottenere informazioni sull'EoS della stella è lo studio della sua deformabilità: ad esempio, nelle ultime orbite della fase inspiral descritta nella Sezione 1.2, le stelle di neutroni risentiranno ognuna della forza di marea esercitata dalla compagna, ed il termine correttivo $\Psi_\lambda(f)$ nella 1.11 tiene conto di tale perturbazione. Una stima dei numeri di Love che compaiono nella correzione fornirà dunque informazioni sulla EoS e, a cascata, su quale sia il modello corretto delle interazioni adroniche. Altre informazioni sullo stato della materia nel core interno possono venire dal comportamento del risultato della fusione di una BNS: come sarà chiarito nella sezione 1.3.2, in base al modello di interazione corretto questo manifesterà instabilità più o meno marcate, instabilità che potranno generare un segnale gravitazionale potenzialmente osservabile.

1.3.1 Meccanismo di emissione principale

Se non operassero altri meccanismi al di fuori dell'attrazione gravitazionale, una stella compatta tenderebbe ad assumere una configurazione a simmetria sferica o cilindrica a seconda che questa sia statica od in rotazione: in uno scenario del genere non ci aspetteremmo alcuna emissione di onde gravitazionali, in quanto si annullerebbero identicamente il momento di quadrupolo del sistema e le sue derivate. Ci si aspetta tuttavia che sistemi del genere posseggano un certo livello di asimmetria nella loro configurazione: tali deformazioni possono essere generate principalmente da deformazioni elastiche nella crosta esterna o dagli intensi campi magnetici presenti nel core, dunque tali deformazioni, anche se piccole, possono produrre onde gravitazionali che in alcuni casi potremmo essere in grado di osservare [12]. Per ricavare l'emissione di onde gravitazionali da una stella compatta in rotazione attorno ad un asse è necessario effettuare alcune assunzioni al fine di semplificare i calcoli. Con buona approssimazione possiamo considerare il sistema deformato come un ellissoide a densità uniforme, descritto nel sistema di riferimento co-rotante dall'equazione: $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$, dove a,b,c sono i semiassi. Un oggetto del genere avrà un tensore di inerzia del tipo:

$$I_{ij} = \frac{M}{5} \begin{pmatrix} b^2 + c^2 & 0 & 0 \\ 0 & c^2 + a^2 & 0 \\ 0 & 0 & a^2 + b^2 \end{pmatrix}$$

Contrassegnando convenzionalmente con z l'asse attorno cui la stella ruota con velocità angolare Ω , il segnale emesso avrà frequenza doppia rispetto a quella rotazionale ($f_{GW} = 2f_{ROT}$) e la sua ampiezza dipenderà dalla componente I_{zz} del tensore di inerzia e dalla sua ellitticità, definita come

$$\varepsilon = 2 \frac{a-b}{a+b}$$

Questa quantità fornisce la misura dell'entità della deformazione, e scala con il quadrato della media sul volume del campo magnetico interno alla stella, $\varepsilon \sim \langle B \rangle^2$ [12].

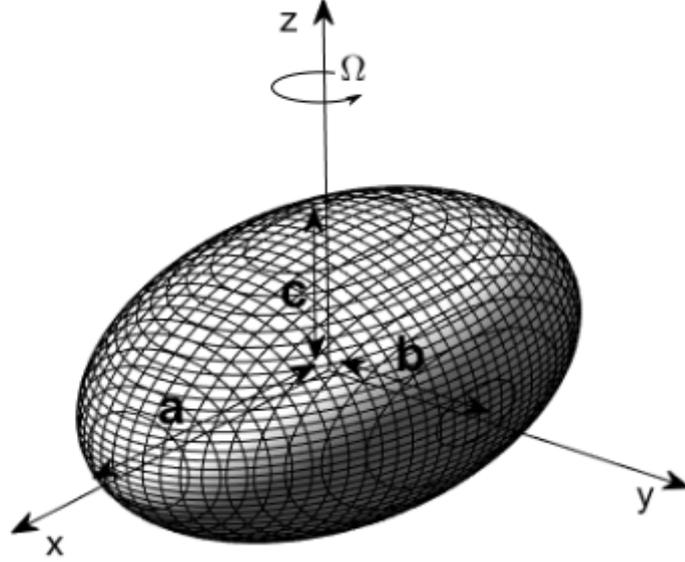


Figura 1.6: Rappresentazione schematica di una stella di neutroni come un ellissoide con semiassi a, b, c , a densità costante, in rotazione attorno ad un asse principale con velocità angolare Ω

In generale dunque il segnale emesso, compattando le due possibili polarizzazioni nella notazione esponenziale, potrà essere espresso nella forma:

$$h(t) = \frac{h_0(t)}{D_L} e^{i\phi(t)}$$

$$h_0(t) = 4\pi^2 \frac{G}{c^4} I_{zz} \varepsilon f_{GW}^2(t) \quad (1.13)$$

mente l'evoluzione della fase $\phi(t) = \int^t 2\pi f_{GW}(t) dt + \phi_{in}$ dipende dall'entità del rallentamento della velocità di rotazione della stella, il cosiddetto spindown, di cui si discuterà a breve. Possiamo subito renderci conto di come l'osservabilità di un segnale del genere dipende fortemente dall'entità della deformazione della stella: considerando infatti oggetti di massa dell'ordine della massa solare e di raggio $\sim 10\text{km}$ avremo momenti d'inerzia tipici dell'ordine di $I_{zz} \sim 10^{38} \text{kg m}^2$, mentre la vera incognita è l'ellitticità: per quanto ne sappiamo la crosta esterna della stella non dovrebbe sopportare deformazioni oltre i 20cm, il che comporterebbe un limite superiore di $\varepsilon \lesssim 10^{-5}$, ma per la presenza di forti campi magnetici tale limite potrebbe salire a $\sim 10^{-4}$. Un valore di riferimento di tipo sperimentale per il quale valutare alte o basse ellitticità è di $\varepsilon = 10^{-6}$, rispetto al quale l'ampiezza dell'onda emessa risulta più o meno

osservabile [13]:

$$h_0 = 4.21 \cdot 10^{-24} \left(\frac{I_{zz}}{10^{38} \text{kgm}^2} \right) \left(\frac{\varepsilon}{10^{-6}} \right) \left(\frac{T}{1 \text{ms}} \right)^{-2} \left(\frac{D_L}{1 \text{kpc}} \right)^{-1} \quad (1.14)$$

Risulta evidente che per avere qualche speranza di osservare segnali del genere la sorgente debba essere relativamente vicina, debba avere un basso periodo e soprattutto presentare una notevole deformazione. Rimane il problema dello spindown: questo nella maggior parte dei casi, soprattutto per stelle di neutroni datate e quindi più stabili, è molto basso, il che fa sì che il segnale osservato sia quasi monocromatico, ma su lunghe osservazioni può diventare determinante per individuarlo.

Applicando la formula 1.7 per la luminosità al sistema in questione, emerge che l'energia trasportata via dall'onda va come la sesta potenza della velocità angolare:

$$L_{GW} = \frac{32G}{5c^5} I_{zz}^2 \varepsilon^2 \Omega^6$$

Assumendo (irrealisticamente) che non ci siano altri meccanismi di dissipazione, si può dunque ottenere un limite superiore per l'ellitticità di una sorgente nota osservando il suo rallentamento:

$$\varepsilon \leq \sqrt{\frac{|\dot{\Omega}|}{\Omega^5} \frac{5c^5}{32GI_{zz}}}$$

Il problema però risiede nel fatto che i meccanismi che provocano il rallentamento sono molteplici: ad esempio oltre all'energia portata via dalle onde gravitazionali bisogna tenere in conto che le pulsar perdono energia anche per emissione elettromagnetica. Considerando un momento magnetico di intensità $|\vec{m}| = \frac{1}{2}BR^3$ ed inclinato di un angolo α rispetto all'asse di rotazione della stella, l'energia irradiata per onde elettromagnetiche sarà

$$L_{EM} = \frac{B^2 R^6}{6c^3} \sin^2 \alpha \Omega^4$$

Già fermandosi a questi due meccanismi di dissipazione si vede che il bilancio complessivo si complica:

$$\dot{\Omega} = - \left(\frac{B^2 R^6}{6c^3 I_{zz}} \sin^2 \alpha \right) \Omega^3 - \left(\frac{32G}{5c^5} I_{zz} \varepsilon^2 \right) \Omega^5 \quad (1.15)$$

Com'è evidente, allo spindown contribuiscono due dipendenze diverse: in base al momento dell'evoluzione in cui si trova il sistema i due contributi avranno un peso differente. Questo andrà di conseguenza ad accumularsi sull'evoluzione della fase. A complicazione ulteriore, oltre a questi due classici meccanismi di dissipazione, si è arrivati a concepire altri potenziali meccanismi dovuti alla presenza di fenomeni di accrescimento o di modi di oscillazione dovuti ad instabilità interne, i quali producono un rallentamento della velocità angolare della stella con una dipendenza funzionale diversa da quelle presenti nella 1.15.

1.3.2 Modi di oscillazione di stelle instabili

Come riportato ad inizio capitolo, il 17 Agosto 2017 i tre detector della collaborazione LIGO/Virgo hanno osservato l'evento GW170817, ovvero il segnale gravitazionale della coalescenza di due oggetti compatti, quasi certamente stelle di neutroni data la presenza della controparte elettromagnetica. Tale

evento offre per la prima volta l'opportunità di studiare la natura del residuo della fusione: nello scenario plausibile in cui le stelle in spiraleggiamento avevano elevati spin, la massa del risultato finale è stimata nel range $M_f \in [2.73, 3.29]M_\odot$ [3]. La natura specifica del residuo allora dipenderà da due fattori: da una parte dall'effettivo valore assunto dalla massa finale, dall'altra dalla sua reale EoS. In base a come la massa finale si colloca nell'EoS (si veda [11]) essa si troverà al di sopra o al di sotto della massa limite \bar{M} per una stella di neutroni, dunque in relazione a ciò è possibile delineare quattro scenari distinti [14]:

- (1) $M_f > \bar{M}_{rot}$: Se la massa del residuo si trova al di sopra della massa limite di una NS rotante uno scenario plausibile è che essa porti alla repentina formazione di un buco nero. In questo caso il segnale generato sarà il ringdown generato dai QNM già discusso nella sezione precedente. Noti i parametri della binaria originaria è possibile stimare che tale segnale avrebbe la frequenza dominante attorno ai $6kHz$, per la quale però i detector nella fase O2 non sono calibrati. Inoltre per un BH residuo del genere, alla distanza di $\sim 40Mpc$, il rapporto segnale-rumore (SNR) del segnale di ringdown sarebbe eccezionalmente basso.
- (2) $M_f > \bar{M}_{rot}$: Se la massa si trova al di sopra del limite c'è anche la possibilità che il residuo formi una stella di neutroni ipermassiva, che invece di collassare subito in BH è sostenuta dall'elevata rotazione e da gradienti termici. In questo scenario si verificherebbe un rapido raffreddamento dovuto all'emissione massiccia di neutrini ed un rallentamento della rotazione ad opera dei campi magnetici, così da far collassare il residuo nel giro di $\lesssim 1s$ dalla sua formazione, generando così un transiente corto [15, 16]. Simulazioni svolte su residui ipermassivi di coalescenze di BNS mostrano che l'emissione gravitazionale è dominata da modi f di oscillazione (f-mode) a $\sim 2 - 4kHz$, con picchi secondari, terziari e successivi che sono correlati con le masse e gli spin dei progenitori e con il modello corretto di EoS nucleare. Tutto questo implica che l'osservazione del segnale di una NS ipermassiva potrebbe potenzialmente permettere di porre forti limitazioni alle possibili EoS.
- (3) $\bar{M}_{nr} < M_f < \bar{M}_{rot}$: Se il residuo è una NS con massa inferiore del limite per stelle rotanti ma comunque supermassiva -ovvero la sua massa è superiore al limite per una NS non rotante- questa rallenterà la sua rotazione tramite emissione gravitazionale ed elettromagnetica, collassando eventualmente dopo un tempo più lungo, tipicamente nel range $[\sim 10, 5 \cdot 10^4]s$, generando dunque un transiente lungo. In questo scenario l'emissione gravitazionale può essere dominata da numerosi meccanismi quali le ellitticità prodotte dal campo magnetico, modi di oscillazione a barra e modi di oscillazione r instabili (r-mode): quest'ultima possibilità sarà oggetto di particolare attenzione nel presente lavoro.
- (4) $M_f < \bar{M}_{nr}$: In questo caso si genera una NS stabile. Ci si aspetta tuttavia che stelle giovani, appena nate, presentino dei forti campi magnetici toroidali al loro interno e di conseguenza delle ellitticità più marcate: dovrebbero quindi essere delle emettitrici di onde gravitazionali più forti rispetto a NS più datate. Anche in questo caso, come nel (3), è plausibile inoltre che si manifestino meccanismi di instabilità di r-mode. In questa prima fase della loro vita dunque il rallentamento di queste nuove NS dovrebbe risultare più marcato, ma a differenza del caso (3) questo non culminerebbe in un collasso in BH.

E' da sottolineare che dalle ricerche finora svolte non è emerso alcun candidato a segnali del genere, detti di post-merger, associabile all'evento GW170817. Ciò non sorprende, in quanto anche considerando i

modelli di emissione più ottimistici, il SNR di un segnale post-merger emesso a $D_L \sim 40 Mpc$ è di 1-2 ordini di grandezza troppo piccolo per essere rilevato con un filtro ottimale (si veda Capitolo 2) data la sensibilità dei detector nella fase O2. Cionondimeno questo terreno rimane di estremo interesse in quanto il miglioramento di sensibilità atteso nella prossima fase O3, in partenza nel 2019, renderebbe questo genere di segnali osservabili: diventa quindi fondamentale raffinare la comprensione teorica di tali fenomeni e concepire algoritmi di ricerca specifici più ottimali possibile in vista della nuova fase osservativa.

Oscillazioni da r-mode

In generale una stella di neutroni ha una vasta gamma di pulsazioni possibili. Riferendoci ai modelli stellari noti più semplici nonché più realistici, tra i modi più rilevanti ci sono ad esempio le oscillazioni di pressione ad alta frequenza, chiamati p-mode, delle quali il modo fondamentale, ovvero che non presenta nodi all'interno della stella, è chiamato usualmente f-mode, che è tra i meccanismi di emissione previsti nel caso (2). Nei decenni scorsi notevoli sforzi sono stati fatti nella direzione di migliorare la comprensione di come la rotazione di una stella vada ad influenzare i modi di pulsazione della stessa. Per stelle non rotanti a simmetria sferica possiamo dividere le pulsazioni in due grandi classi: sferoidali e toroidali, chiamate anche rispettivamente polari ed assiali nel gergo relativistico. Le perturbazioni sferoidali presentano componenti radiali e sono accompagnate da variazioni in pressione e densità: rientrano in questa categoria i p-modes. Le perturbazioni toroidali si sviluppano invece nelle sole componenti angolari ed in questa classe rientrano gli r-mode, che possono essere considerati come l'analogo stellare delle onde di Rossby che si osservano negli oceani terrestri, e la cui forza di richiamo è la forza di Coriolis. Generalmente per descrivere tali oscillazioni si procede a perturbare una situazione di equilibrio data dalle equazioni di Eulero (o dall'analogo relativistico) tramite un vettore di spostamento $\vec{\xi}$ in modo da provocare una variazione nel campo di velocità del fluido $\delta\vec{u} = \partial_t \vec{\xi}$. La dipendenza angolare di queste funzioni $\vec{\xi}$ è espressa in termini delle armoniche sferiche $Y_{lm}(\theta, \varphi)$ attraverso delle funzioni T_{lm} , nella forma delle coordinate sferiche [20]:

$$\frac{\vec{\xi}}{r} = \sum_{lm} \left(0, \frac{T_{lm}}{\sin \theta} \partial_\varphi, -T_{lm} \partial_\theta \right) Y_{lm} \quad (1.16)$$

Nello specifico la dipendenza dall'angolo azimutale è data da un termine oscillante $e^{im\varphi}$, mentre se si assume che tali modi abbiano una dipendenza temporale di tipo armonico si dovrà considerare anche un termine $e^{i\omega_r t}$, dove con ω_r si intende la frequenza di oscillazione misurata nel sistema di riferimento corotante, assunta positiva, data al primo ordine da:

$$\omega_r = \frac{2m\Omega}{l(l+1)} \quad (1.17)$$

allora un modo di oscillazione -classificato con $[l, m]$ - che presentasse un valore di $m > 0$ risulterà progressivo rispetto al fluido stellare, mentre un modo con $m < 0$ risulterà retrogrado [19, 20]. In figura 1.7 è possibile vedere un esempio dell'andamento del campo delle velocità sulla superficie visto da un osservatore corotante: ogni mezzo periodo di oscillazione il moto cambia direzione ed il pattern si dirige nel verso opposto rispetto al senso di rotazione. Tuttavia il solo campo delle velocità non rende una buona immagine del moto dei singoli elementi di fluido: questi, all'ordine dominante, si muovono su

una ellisse le cui dimensione sono determinate dall'ampiezza del modo, avvicinandosi a circonferenze di raggio tendente a zero quando $\theta \rightarrow 0/\pi$, mentre all'equatore ($\theta \rightarrow \pi/2$) si muoveranno solamente in alto e in basso lungo θ .

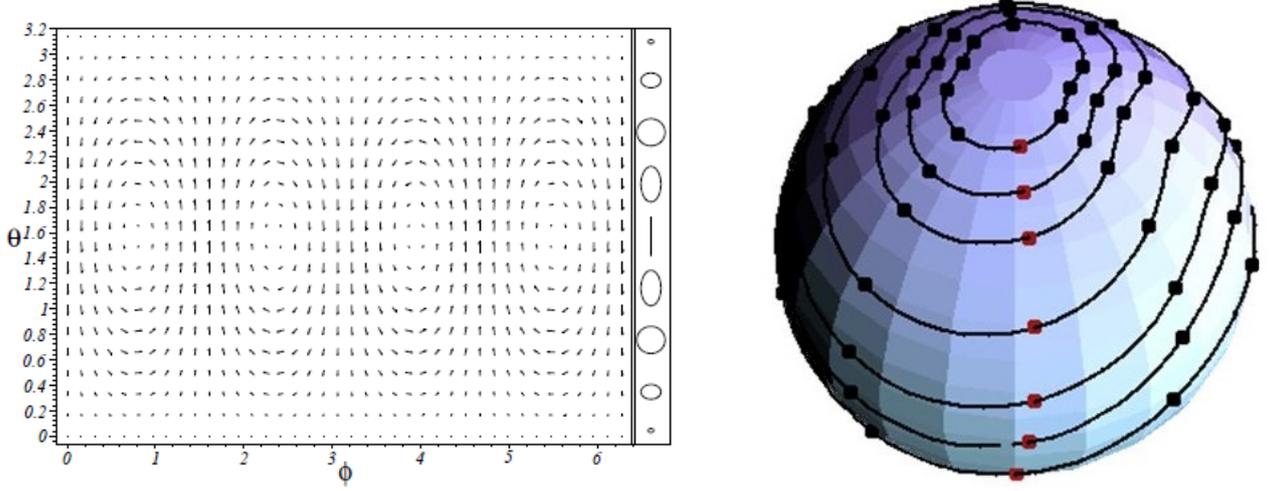


Figura 1.7: A sinistra: nel riquadro grande, rappresentazione all'ordine dominante del campo delle velocità nell'r-mode ad $m = l = 2$ in un dato istante di tempo; nel riquadro piccolo, illustrazione qualitativa del moto di elementi di fluido alla quota θ (fonte: [20]). A destra: singolo frame di una simulazione di un r-mode, comprendente ordini superiori.

Passando da un osservatore solidale ad uno inerziale, distante dal sistema, questo vedrà una frequenza corretta dalla rotazione della stella:

$$\omega_i = \omega_r - m\Omega \quad (1.18)$$

Essendo ogni modo proporzionale ad $e^{i(m\phi + \omega t)}$, le superfici a fase costante saranno date dall'equazione $m\phi + \omega t = const$, la cui differenziazione porta a definire la velocità di pattern, data da:

$$\frac{d\phi}{dt} = -\frac{\omega}{m} = \sigma$$

In questo modo si può fare un'importante osservazione. Infatti, si vede chiaramente che per un osservatore solidale alla stella la velocità di pattern di un tipico r-mode è opposto alla rotazione, mentre dall'altra parte, grazie alla 1.18, un osservatore inerziale misurerà una velocità di segno opposto [20]:

$$\sigma_r = -\frac{2\Omega}{l(l+1)} \quad \rightarrow \quad \sigma_i = \Omega \frac{(l-1)(l+2)}{l(l+1)}$$

Perciò un modo retrogrado nel sistema corotante sarà visto come progressivo in un sistema inerziale.

Il motivo per cui si ritiene che le onde gravitazionali generate dagli r-mode possano essere rilevanti a livello osservativo è dovuto ad un fenomeno noto come instabilità di Chandrasekhar-Friedman-Schutz (CFS). Consideriamo un modo di oscillazione in una NS ruotante: la rotazione induce modi sia progressivi che retrogradi, ma quelli retrogradi vengono trasportati in avanti dalla rotazione della stella. Anche se un modo retrogrado continua ad andare in verso opposto nel sistema solidale, se la rotazione della stella è sufficientemente elevata questo risulterà progressivo nel sistema inerziale. L'onda gravitazionale

emessa porterà energia via dal modo, ma porterà con se anche un momento angolare positivo, ma dato che il fluido perturbato ruota più lentamente di come farebbe in assenza della perturbazione, il momento angolare del modo sarà negativo: il momento angolare positivo sottratto dall'onda andrà allora a sottrarsi al già negativo momento del modo, il quale diventerà ulteriormente negativo aumentando così l'ampiezza dell'oscillazione. Al suo crescere, il modo emette momento angolare positivo sempre maggiore e continua a crescere sempre più velocemente. A frenare il processo si oppone la resistenza viscosa, che tende a sopprimere l'ampiezza del modo.

Un passo importante è stato capire che a tutte le velocità di rotazione stellare modi che sono retrogradi nel sistema corotante diventano progressivi nel sistema inerziale, implicando perciò che in generale gli r-mode sono sempre instabili a causa della instabilità CFS [12, 19, 20]. Questa osservazione implica che l'instabilità r potrebbe risultare rilevante a livello astrofisico, anche se gli r-mode, a differenza ad esempio degli f-mode, non portano a ingenti variazioni nella densità della NS e quindi intuitivamente non si assocerebbero a forti emissioni gravitazionali. Invece meccanismi di dissipazione come gli f-mode per diventare instabili richiedono che la rotazione della stella raggiunga un certo livello critico, mentre gli r-mode sono sempre instabili.

Al netto di tutti i meccanismi citati, si prevede che il rallentamento causato alla stella sia di entità maggiore rispetto agli altri meccanismi standard, di conseguenza ci si aspetta di osservare un andamento dello spin-down che vada come la settima potenza dello spin

$$\dot{\Omega} = -\frac{\lambda \alpha^2}{6} \Omega^7 \quad (1.19)$$

dove λ è una costante che dipende da un parametro Q dell'EoS e della densità media della stella ρ , stimata come:

$$\lambda = (2\pi)^6 \frac{12Q}{\tau(\pi G \rho)^3} \approx 10^{-20} \text{Hz}^{-5}$$

mentre la costante α , adimensionale, è l'ampiezza alla quale il modo satura. Integrando la 1.19 e passando alla frequenza dell'onda gravitazionale, si ottiene la seguente evoluzione in funzione del tempo:

$$f_{GW}(t) = f_0 [1 + \lambda \alpha^2 f_0^6 (t - t_0)]^{-1/6} \quad (1.20)$$

dove f_0 rappresenta la frequenza iniziale del segnale, al tempo iniziale t_0 . Come si vede, l'evoluzione della frequenza è tale lasciare un segno più evidente sulla fase accumulata anche su durate brevi:

$$\phi(t) = \frac{12\pi}{5} \frac{1}{\lambda \alpha^2 f_0^5} \left\{ [1 + \lambda \alpha^2 f_0^6 (t - t_0)]^{5/6} - 1 \right\} \quad (1.21)$$

Complessivamente l'ampiezza dell'onda sarà data dalla seguente relazione:

$$h(t) = \frac{h_0(t)}{D_L} e^{i\phi(t)} \quad (1.22)$$

$$h_0(t) = \sqrt{\frac{2048}{375} \frac{GJMR^3}{c^5}} \alpha f_{GW}^3(t) \quad (1.23)$$

dove $J = 1.635 \cdot 10^2$ è un momento angolare adimensionale valido per modelli stellari politropici ed M ed R sono la massa ed il raggio della NS. Come si può vedere nell'esempio in figura 1.8, l'ampiezza, la

fase e la frequenza del segnale presentano una lenta ma apprezzabile variazione.

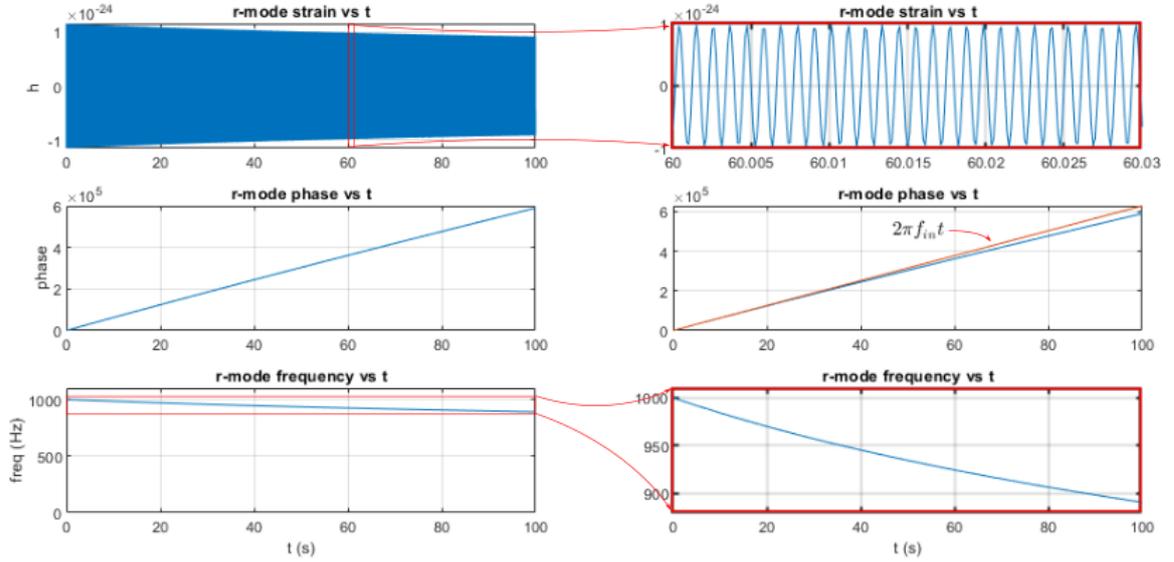


Figura 1.8: Simulazione di un segnale gravitazionale da r-mode calcolato con $f_0 = 1000\text{Hz}$, $\alpha = 1$. Per l'ampiezza e la frequenza si riporta a destra uno zoom sui dettagli, per la fase si riporta a destra il confronto con l'andamento (in rosso) che si avrebbe a frequenza costante.

Per quanto riguarda la rivelabilità del segnale, questa dipende in buona parte dalla distanza della sorgente. In generale per sorgenti agli ordini di distanza del cluster Virgo si valuta che la sensibilità dei detector nella fase osservativa appena trascorsa (O2) sia di un ordine di grandezza al di sotto del necessario per poter osservare segnali del genere, mentre nella prossima fase O3 le sensibilità raggiunte dovrebbero consentirne l'osservazione.

1.3.3 Generalizzazione della dinamica dello spin-down

Da quanto esposto finora appare evidente che al rallentamento dello spin di una NS concorreranno in linea di principio più meccanismi contemporaneamente. Restringendo l'attenzione agli r-mode ad esempio, appare evidente che allo spin-down non contribuirà solamente la legge di settima potenza espressa nella 1.19, ma si avrà certamente un contributo dato dall'emissione elettromagnetica e presumibilmente un contributo dall'emissione gravitazionale dovuta a deformazione, i quali sono caratterizzati dalla 1.15. In generale dunque ci si aspetta che il rallentamento di una stella assuma un andamento intermedio a quelli trovati, motivo per cui è opportuno generalizzare alcuni concetti.

Si assume con più generalità che l'evoluzione dello spin segua una legge di potenza del tipo [21, 22]:

$$\dot{\Omega} = -k\Omega^n \quad (1.24)$$

dove la costante k dipende dal tipo di meccanismo considerato e la costante n è chiamata indice di frenata. Diversi valori di n corrispondono a diversi meccanismi fisici di dissipazione, alcuni i quali non sono stati trattati. Ad esempio $n = 1$ corrisponde al vento di frenata, ovvero ad un flusso di particelle relativistiche che si allontanano dalla NS sotto spinta elettromagnetica il cui meccanismo di base non è ancora

Questa fa riferimento principalmente allo spettro della radiazione osservata ed in alcuni casi anche alle curve di luce, il che significa che i tipi di SN sono assegnati in base alle proprietà fisiche e chimiche degli strati più esterni della stella esplosa. I primi due macrogruppi in cui si classifica una SN sono in base alla presenza o assenza delle linee caratteristiche dell'idrogeno nello spettro luminoso: SNe con l'assenza di tali linee tradiscono la presenza dell'idrogeno e vengono classificate di tipo I, se invece non è presente idrogeno vengono classificate di tipo II. Le SNe di tipo I si possono poi dividere in Ia, Ib ed Ic: le SNIa presentano le linee di assorbimento del silicio nel loro spettro, mentre le altre no; le SNIb esibiscono le linee dell'elio negli spettri, mentre le SNIc no. Senza scendere in ulteriori dettagli si può affermare che le SN di tipo Ia sono associabili tipicamente ad esplosioni di nane bianche in accrescimento, mentre le Ib, Ic e II si associano al collasso del core di stelle massive, processo etichettato con CCSNe (Core Collapse SuperNovae). E' su quest'ultimo processo che si sono concentrati numerosi studi finalizzati a modellizzare l'emissione di onde gravitazionali e neutrini attesi.

Non tutte le stelle al termine del loro ciclo vitale danno vita ad una SN. Durante la maggior parte della sua vita una stella si mantiene in equilibrio bruciando le proprie riserve di idrogeno in reazioni di fusione nucleare producendo nuclei di elio; una volta terminate tali riserve, se la massa della stella non supera le $4M_{\odot}$, la sua gravità non sarà in grado di dare vita ad ulteriori reazioni e la stessa andrà incontro ad un progressivo raffreddamento, andando a formare nella maggior parte dei casi una nana bianca. Se invece la stella ha massa $> 4M_{\odot}$, sarà in grado di attivare reazioni di fusione che coinvolgano elementi della tavola periodica sempre più pesanti come elio, carbonio ed ossigeno, fino ad arrivare alla formazione dell'isotopo ^{56}Fe : questo nuclide rappresenta un punto limite in quanto nessun elemento più pesante di esso può essere generato tramite una reazione di fusione tra elementi più leggeri che sia esoenergetica. Unitamente a questo avviene inoltre un altro processo: all'aumentare della densità del core diventa sempre più efficiente il decadimento β inverso, nel quale un elettrone ed un protone si fondono in un neutrone liberando un neutrino elettronico secondo il processo $e^{-} + p \rightarrow n + \nu_e$; dato che i neutrini interagiscono debolmente con la materia, questi diffonderanno dal core verso la superficie esterna, portando via al netto la loro energia dal core. Tutti questi processi vanno quindi a destabilizzare il core della stella, in modo che una volta che la sua massa abbia superato il limite di Chandrasekar $M \sim 1.4M_{\odot}$ il gradiente di pressione interna diventa più piccolo dell'attrazione gravitazionale, innescandone in una frazione di secondi il collasso fino ad arrivare a densità tipiche dei nuclei atomici. In questa fase il core reagisce ad una ulteriore compressione producendo una violenta onda di shock la quale espelle gran parte del materiale esterno in una spettacolare esplosione detta appunto di supernova.

Per quanto detto risulta evidente la portata energetica di un evento del genere. Per questo motivo non sembra plausibile applicare il meccanismo di quadrupolo, che risulta eccessivamente grossolano nei suoi assunti. In questo caso si ricorre alle simulazioni numeriche di casi il più possibile realistici, come mostrato ad esempio in [24], che qui non verranno approfondite. In questa sede è opportuno sottolineare che il collasso del core di ferro di una stella, se non considerato correttamente nella sua complessità, risulterebbe un problema a simmetria assiale, il quale non comporterebbe di conseguenza alcun tipo di emissione gravitazionale. Perché si verifichi emissione bisogna introdurre meccanismi che violino la simmetria del sistema, quali: il collasso ed il successivo rimbalzo di un core che sia in rapida rotazione; instabilità non assisimmetriche nella rotazione; emissione post-bounce, valida per stelle progenitrici non rotanti, dominata da growing modes (g-mode). Nell'immagine superiore in Figura 1.10 è possibile vedere la più generica forma d'onda che si dovrebbe osservare nel primo caso, con collasso di un core in rapida rotazione e successivo rimbalzo, ottenuta dalla simulazione dell'esplosione di un modello in

3D di una stella di $27M_{\odot}$ distante $10kpc$ in [24]. Vista dall'equatore, la forma d'onda dopo il rimbalzo è caratterizzata da un forte picco seguito da una fase di ring-down a partire da $\sim 30ms$. Vista dall'asse di rotazione terrestre, l'ampiezza dell'onda è diversa da zero solamente a partire da $\sim 40ms$, quando iniziano a svilupparsi instabilità non assisimetriche.

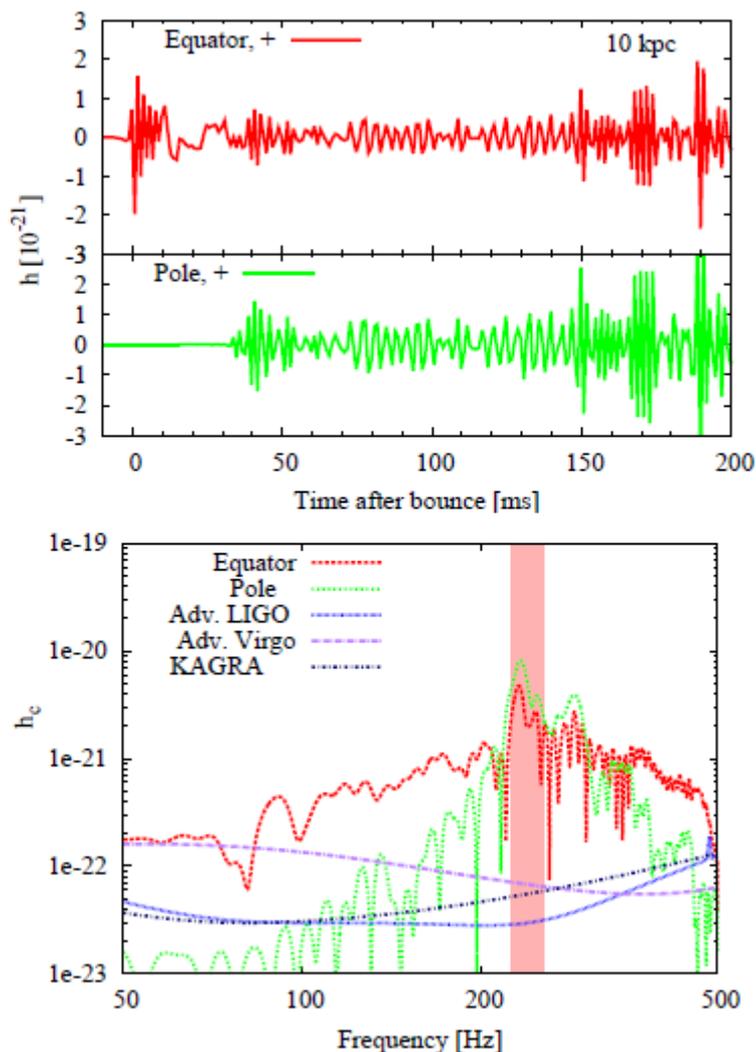


Figura 1.10: Sopra: ampiezza della componente polarizzata '+' dell'onda per osservatori posti lungo l'equatore (linea rossa) ed al polo (linea verde) per una sorgente a $10kpc$. Sotto: spettro caratteristico dell'onda posta a confronto con le curve di sensibilità dei detector. Fonte: [24].

Nell'immagine inferiore è mostrato lo spettro caratteristico dell'onda confrontato con le curve di sensibilità dei detector advanced LIGO, advanced Virgo e KAGRA per una sorgente a $10kpc$. Vista sia dall'equatore che dal polo, il picco dell'onda si colloca attorno alla frequenza di $240Hz$, suggerendo quindi che le emissioni dai modelli ipotizzati sopra potrebbero essere osservabili (si faccia riferimento alle Sezioni 2.2,2.3). Qualora si verificassero eventi del genere durante i prossimi run osservativi, l'eventuale osservazione del segnale gravitazionale sarebbe un importante banco di prova per i meccanismi di emissione ed i relativi modelli sviluppati.

Capitolo 2

Attuali strategie di rivelazione

Nel momento in cui ci si pone di osservare un'onda gravitazionale nei dati raccolti dai detector a terra, occorre necessariamente affrontare il problema del rumore in cui il segnale, se presente, è sepolto. Se da una parte abbiamo una comprensione teorica più o meno completa dei fenomeni di interesse, dal lato pratico abbiamo a che fare con una serie di dati discreti in uscita dai diversi detector in funzione: come punto di partenza abbiamo un set di dati

$$d^{(D)}(t) \in \left\{ d^{(L)}(t), d^{(H)}(t), d^{(V)}(t), d^{(K)}(t), \dots \right\}$$

per ognuno dei detector che compongono la rete, dove la variabile temporale t è discreta e con le iniziali $\{L, H, V, K\}$ si intendono i siti di Livingston, Hanford, Virgo e Kagra (quest'ultimo prossimo ad entrare in funzione). Nell'ipotesi che nei dati sia effettivamente nascosto un segnale, possiamo considerare questi dati come somma del segnale con il rumore puro, ovvero:

$$d^{(D)}(t) = h^{(D)}(t) + n^{(D)}(t)$$

dove $h^{(D)}(t)$ rappresenta i segnali giunti ad ogni detector, che ci aspettiamo essere correlati tra loro, mentre $n^{(D)}(t)$ rappresenta il rumore dello specifico detector, che ci aspettiamo essere scorrelato passando da un sito ad un altro.

2.1 Segnali discreti: il problema del campionamento

I segnali prodotti dai processi fisici descritti nel Capitolo 1 sono grandezze fisiche continue, per la comprensione che ne abbiamo attualmente. Questo significa che per poterli analizzare ed elaborare attraverso sistemi digitali, questi dovranno essere campionati: ciò comporta l'acquisizione dei valori di un segnale continuo a certi istanti temporali che costituiscono un insieme discreto. Una scelta semplice è quella di considerare istanti di tempo che siano equispaziati, scelta che prende il nome di campionamento uniforme, anche se in alcuni casi risulta più efficace usare un campionamento non uniforme (basti pensare a casi in cui il segnale presenta una alta variabilità). A partire da un segnale continuo $h(t)$ occorrerà quindi individuare un tempo di campionamento Δt_s in base al quale estrarre una successione $\{t_i\}$ di valori:

$$h(t) \rightarrow h_i = h(t_i), t_i = i\Delta t_s, i \in \mathbb{N}$$

La scelta del tempo di campionamento tuttavia non è banale: una scelta sbagliata potrebbe impedire di cogliere tutte le caratteristiche del segnale, con la conseguenza di rendere impossibile la ricostruzione del segnale originale. Per orientare la scelta si ricorre allora al teorema del campionamento, formulato da Nyquist nel 1928, che fornisce la seguente regola:

Teorema: Se un segnale reale ha spettro di potenza nullo per frequenze superiori ad una determinata ν_N , detta frequenza di Nyquist, questo può essere completamente ricostruito se è campionato ad una frequenza doppia:

$$\nu_s \geq 2 \nu_N$$

dove $\nu_s = \frac{1}{\Delta t_s}$ è la frequenza di campionamento. Dati allora i campioni $\{x_i\}$ raccolti, si ricostruisce il segnale originale secondo la:

$$x(t) = \sum_{i=-\infty}^{+\infty} x_i \frac{\sin \left[\pi \left(\frac{2t}{\Delta t_s} - i \right) \right]}{\pi \left(\frac{2t}{\Delta t_s} - i \right)}$$

Qualora la condizione del teorema non sia rispettata, ciò che succede è che le frequenze più elevate di ν_N si mascherano da frequenze inferiori a ν_N : è il fenomeno che in inglese viene chiamato aliasing. Se ne può vedere un esempio in Figura 2.1: in questo caso si ha una sinusoidale a frequenza $\nu = \frac{4}{5}Hz = \nu_N$ che si vuole campionare con frequenza $\nu_s = 1Hz$, la quale tuttavia è al di sotto del limite di Nyquist $2\nu_N = \frac{8}{5}Hz$. Il problema è evidente: infatti si vede che tale sinusoidale condivide gli stessi campioni con un'altra sinusoidale a frequenza $\nu' = \frac{1}{5}Hz$ che risulterà favorita in sede di ricostruzione, presentando 5 campioni a periodo a fronte degli 1.25 della sinusoidale iniziale (inferiori ai 2 prescritti dal teorema).

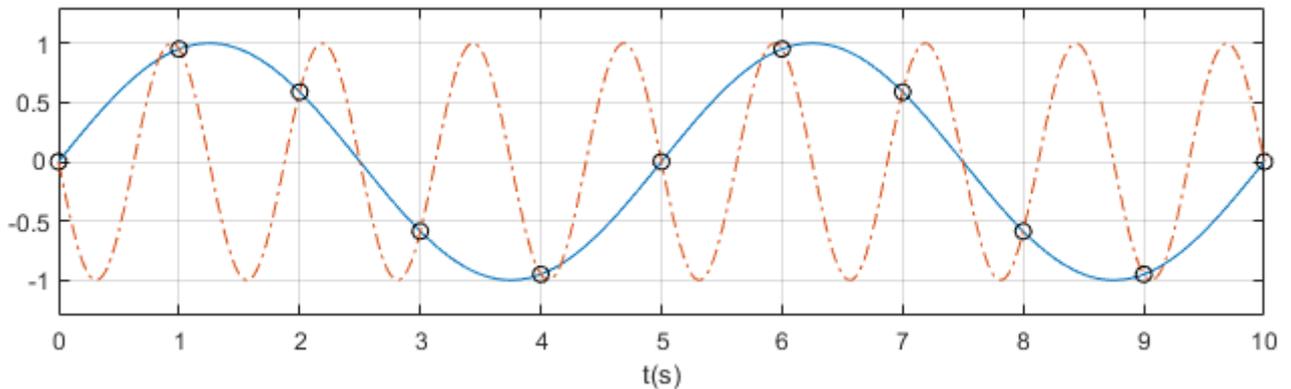


Figura 2.1: Esempificazione del fenomeno dell'aliasing. La sinusoidale tratteggiata in rosso ha equazione $h_1(t) = -\sin(2\pi\frac{4}{5}t)$ e frequenza $\frac{4}{5}Hz$, mentre quella in blu ha equazione $h_1(t) = -\sin(2\pi\frac{1}{5}t)$ e frequenza $\frac{1}{5}Hz$. I punti presi dal campionamento ad $1Hz$ sono cerchiati in nero e risultano in comune alle due funzioni.

Generalmente il fenomeno dell'aliasing risulta dannoso ai fini dell'analisi dei segnali in quanto va ad aumentare il livello del rumore. E' quindi pratica importante eliminarlo o quanto meno limitarlo, ad esempio applicando dei filtri analogici passa-basso a monte del campionamento in modo da ricadere nelle condizioni del teorema del campionamento.

2.2 Caratterizzazione del rumore

Il problema del rumore è di centrale importanza nella rivelazione, in quanto è la sua entità a rendere possibile o meno l'estrazione di un determinato segnale. La strategia di rivelazione adottata con lo schema interferometrico comporta una misura di zero: essenzialmente si misura la variazione nell'intensità di un fascio laser in uscita da un interferometro di Michelson in condizione di interferenza distruttiva o buio, condizione che viene perturbata dal passaggio dell'onda gravitazionale con conseguente aumento di intensità. Tuttavia tutti gli esperimenti che comportano misure di zero hanno come ostacolo principale tutti i disturbi dovuti alla strumentazione. Diventa dunque fondamentale una completa comprensione della natura dei disturbi rumorosi. La trattazione che segue fa riferimento a [25].

In base alla teoria dei fenomeni aleatori possiamo considerare il rumore nella famiglia dei processi stocastici discreti, ovvero un insieme di funzioni (realizzazioni) del tempo alle quali è associata una distribuzione di probabilità. Ad esempio, dato un processo stocastico $x(t)$ possiamo definire la funzione di distribuzione del primo ordine:

$$F(x;t) = P[x(t) \leq x] \rightarrow f(x;t) = \frac{\partial F(x;t)}{\partial x} \quad (\text{densità})$$

Analogamente si possono definire le funzioni di distribuzione di ordine n come:

$$F(x_1, \dots, x_n; t_1, \dots, t_n) = P[x(t_1) \leq x_1, \dots, x(t_n) \leq x_n] \rightarrow f(x_1, \dots, x_n; t_1, \dots, t_n) = \frac{\partial^n F(x_1, \dots, x_n; t_1, \dots, t_n)}{\partial x_1 \dots \partial x_n}$$

In funzione di tali distribuzioni si possono definire grandezze fondamentali per la determinazione del processo, quali la media, la varianza e l'autocorrelazione nelle versioni continua e discreta:

$$\mu(t) = E[x(t)] = \int_{\mathbb{R}} x f(x, t) dx \rightarrow \mu(t_k) = \sum_{i=-\infty}^{+\infty} x_i f(x_i, t_k)$$

$$\sigma^2(t) = E[(x(t) - \mu(t))^2] = \int_{\mathbb{R}} (x - \mu(t))^2 f(x, t) dx \rightarrow \sum_{i=-\infty}^{+\infty} (x_i - \mu(t_k))^2 f(x_i, t_k)$$

$$R_{xx}(t_1, t_2) = E[x^*(t_1)x(t_2)] = \int_{\mathbb{R}} \int_{\mathbb{R}} x_1^* x_2 f(x_1, x_2, t_1, t_2) dx_1 dx_2 \rightarrow \sum_{i_1=-\infty}^{+\infty} \sum_{i_2=-\infty}^{+\infty} x_{i_1}^* x_{i_2} f(x_{i_1}, x_{i_2}, t_1, t_2)$$

Quest'ultima grandezza (l'autocorrelazione) è di particolare importanza, in quanto rappresenta la predicibilità nel tempo di un dato processo, ovvero quanto è predicibile il valore $x(t_2)$ quando è noto $x(t_1)$. A partire da essa, esprimibile più comodamente come $R_{xx}(t, t + \tau)$, si può ricavare la densità spettrale di potenza attraverso la trasformata di Fourier:

$$S_x(t, f) = \int_{\mathbb{R}} R_{xx}(t, \tau) e^{-2\pi i f \tau} d\tau \rightarrow S_x(t_1, \nu) = \sum_{k=-\infty}^{+\infty} R_{xx}(t_1, k) e^{-i\Omega k}$$

Questa rappresenta la distribuzione della potenza del segnale nel dominio delle frequenze e viene spesso chiamata più sinteticamente spettro di potenza o semplicemente spettro. Nonostante contenga evidentemente la stessa informazione dell'autocorrelazione, essa risulta di più facile interpretazione.

Altre proprietà di importanza centrale per il rumore sono stazionarietà ed ergodicità. Un processo si dice stazionario se le funzioni dalle quali è definito sono indipendenti per traslazioni temporali generiche, ov-

vero se $x(t)$ è statisticamente indistinguibile da $x(t+T)$. Per un processo del genere l'autocorrelazione diventa funzione di una sola variabile temporale e di conseguenza lo spettro di potenza perde la dipendenza dal tempo. La proprietà di ergodicità implica invece che le statistiche di insieme siano uguali alle statistiche temporali, con la conseguenza fondamentale di poter inferire le proprietà del processo stocastico dalla realizzazione effettiva che si ha a disposizione. Stanti queste assunzioni è possibile stimare l'autocorrelazione sostituendo il valore aspettato con la media su un pezzo di dati $\{x_1, \dots, x_N\}$:

$$R_{xx}(k) = \frac{1}{N-k} \sum_{i=1}^{N-k} x_{i+k} x_i^* \quad (2.1)$$

Lo stimatore espresso dalla 2.1 ha il pregio di essere unbiased, tuttavia produce evidentemente forti fluttuazioni quando viene calcolato su valori di k vicini ad N . Per rimediare si potrebbe distorcere la stima sostituendo il coefficiente di normalizzazione $\frac{1}{N-k}$ con $\frac{1}{N}$, con il vantaggio di ridurre le fluttuazioni, ma rimane comunque il problema che il calcolo di tale stimatore richiede la ripetizione di $\frac{N^2}{2}$ operazioni, una mole di lavoro eccessivamente pesante. Per fare di meglio si può estendere la successione $\{x_1, \dots, x_N\}$ ad una successione infinita aggiungendo zeri nelle altre posizioni: in questo modo la sommatoria in 2.1 si estende a $(-\infty, +\infty)$, ma così facendo si è riprodotta la convoluzione tra la x invertita temporalmente ed il complesso coniugato della x . Sfruttando allora il teorema della convoluzione ed indicando con $\tilde{x}(v)$ la trasformata di Fourier discreta dei dati, possiamo calcolare l'autocorrelazione in maniera più efficiente:

$$R_{xx}(k) = \frac{1}{N} \mathcal{F}^{-1} \left(|\tilde{x}(v)|^2 \right) \quad (2.2)$$

dove con \mathcal{F}^{-1} si intende la trasformata inversa di Fourier. Data la definizione, la parte reale risulta essere una funzione pari, mentre la parte immaginaria è dispari. Inoltre $R_{xx}(0)$ è reale ed è il massimo assoluto di tale funzione. Infine risulta $R_{xx}(k) = R_{xx}^*(-k)$ e nel complesso la funzione è definita positiva. Queste ultime proprietà in particolare si traducono in semplici proprietà per la trasformata di Fourier dell'autocorrelazione, ovvero la densità spettrale di potenza: questa infatti risulta sempre reale e non-negativa. Se inoltre il processo è reale (come avviene solitamente per i dati raccolti) lo spettro è anche una funzione pari: in questo caso si usa considerare solamente metà spettro e si parla allora di spettro unilatero. In accordo con il teorema di Wiener-Kinchin è possibile allora calcolare lo spettro in maniera più rapida come:

$$S_x(v) = \frac{|\tilde{x}(v)|^2}{N} \quad (2.3)$$

Si noti che la stima espressa in 2.3 ha in oggetto la trasformata di Fourier dei dati discreti: per ogni v si hanno dunque due somme di N numeri casuali moltiplicati per le funzioni seno e coseno, le quali hanno media 0 e varianza $\frac{1}{2}$:

$$\tilde{x}(v) = \sum_{k=1}^N x_k e^{-2\pi i(k-1)v} = \sum_{k=1}^N x_k \cos[2\pi(k-1)v] + i \sum_{k=1}^N x_k \sin[2\pi(k-1)v]$$

In base al teorema del limite centrale il risultato della somma sono due variabili a distribuzione normale, con media nulla e varianza $\frac{N}{2}$. Dato che poi di queste quantità si sommano i quadrati, la somma dei quadrati di due variabili gaussiane segue la distribuzione del χ^2 con 2 gradi di libertà, che corrisponde alla distribuzione esponenziale. Dunque si conclude che lo stimatore dello spettro di potenza segue la

distribuzione esponenziale, a prescindere dalla distribuzione che avevano i dati in partenza. Una schematizzazione importante è quella di rumore bianco, ovvero di un processo stocastico stazionario ed ergodico in cui le funzioni di distribuzione sono gaussiane. Il fatto che le distribuzioni siano gaussiane comporta delle proprietà fondamentali: se infatti del processo sono note la media e l'autocorrelazione, queste sono sufficienti a determinare tutte le funzioni di distribuzione di qualsiasi ordine n ; inoltre una trasformazione lineare di un processo stocastico normale produce in uscita un processo stocastico normale. Per questi motivi anche quando i fenomeni in oggetto non possono essere considerati esattamente gaussiani si cerca di approssimarli con essi.

Il rumore bianco è per definizione a media nulla:

$$\mu(k) = 0$$

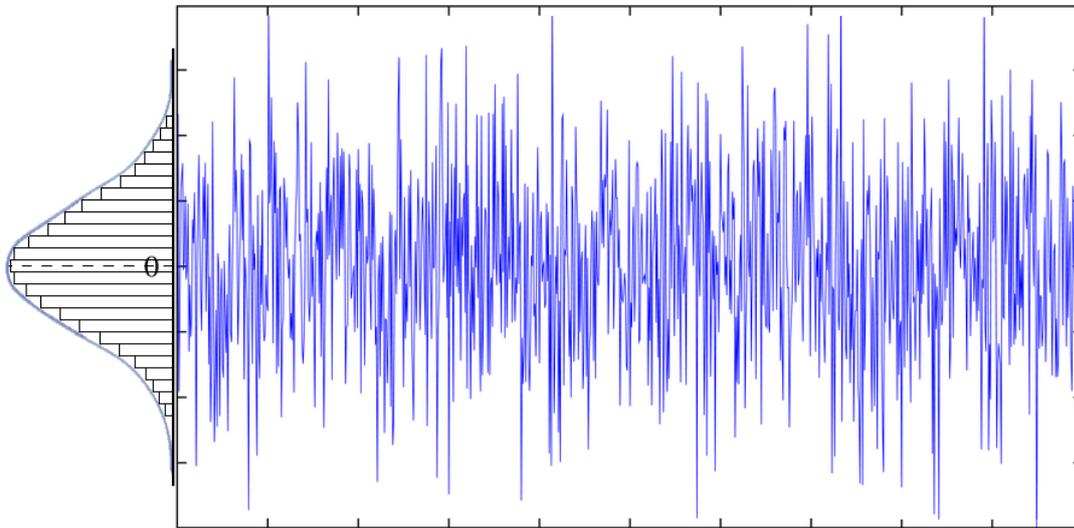


Figura 2.2: Rappresentazione illustrativa di una realizzazione di rumore bianco, la cui ampiezza segue una densità di probabilità gaussiana ed il cui valor medio è nullo.

Perciò la conoscenza dell'autocorrelazione o della densità spettrale di potenza in linea di principio esaurisce la conoscenza del processo. Queste sono esprimibili in funzione della varianza:

$$R_{xx}(k) = \sigma^2 \delta_k$$

$$S_x(\nu) = \sigma^2$$

Dunque in linea di principio lo spettro del rumore bianco è piatto. Questo è tanto più vero quanto maggiore è il numero di campioni a disposizione, dato che la finitezza del campionamento comporta inevitabilmente un certo grado di fluttuazione. Bisogna però tener conto del fatto che la risoluzione in frequenza della stima spettrale è inversamente proporzionale alla lunghezza del campione di dati: per un campione di N dati, questa è $\Delta\nu = \frac{1}{N}$ o, in unità della frequenza fisica, $\Delta f = \frac{1}{N\Delta_s}$. Questo significa che la risoluzione di un campione esteso di rumore sarà elevata, così come le oscillazioni attorno al valor medio atteso, il che renderebbe lo spettro scarsamente utilizzabile. Per ovviare a questo problema si ricorre

spesso alla seguente procedura: si spezza il treno di dati in M campioni di lunghezza uguale, si calcola lo spettro per ognuno di essi e se ne fa la media. La nuova risoluzione in questo modo sarà inversamente proporzionale non più ad N , ma al rapporto N/M . Si noti che effettuando tale scelta abbiamo modificato anche la distribuzione dei dati di cui si compone lo spettro, che adesso seguirà una distribuzione del χ^2 con $2M$ gradi di libertà. In Figura 2.3 si può vedere un esempio di come questa procedura vada a modificare l'ampiezza delle fluttuazioni dal valor medio nello spettro, ma anche la risoluzione. Evidentemente la scelta del numero di campioni M dovrà essere fatta mediando nella contingenza tra la necessità di avere una buona risoluzione da una parte e la necessità di una bassa incertezza dall'altra.

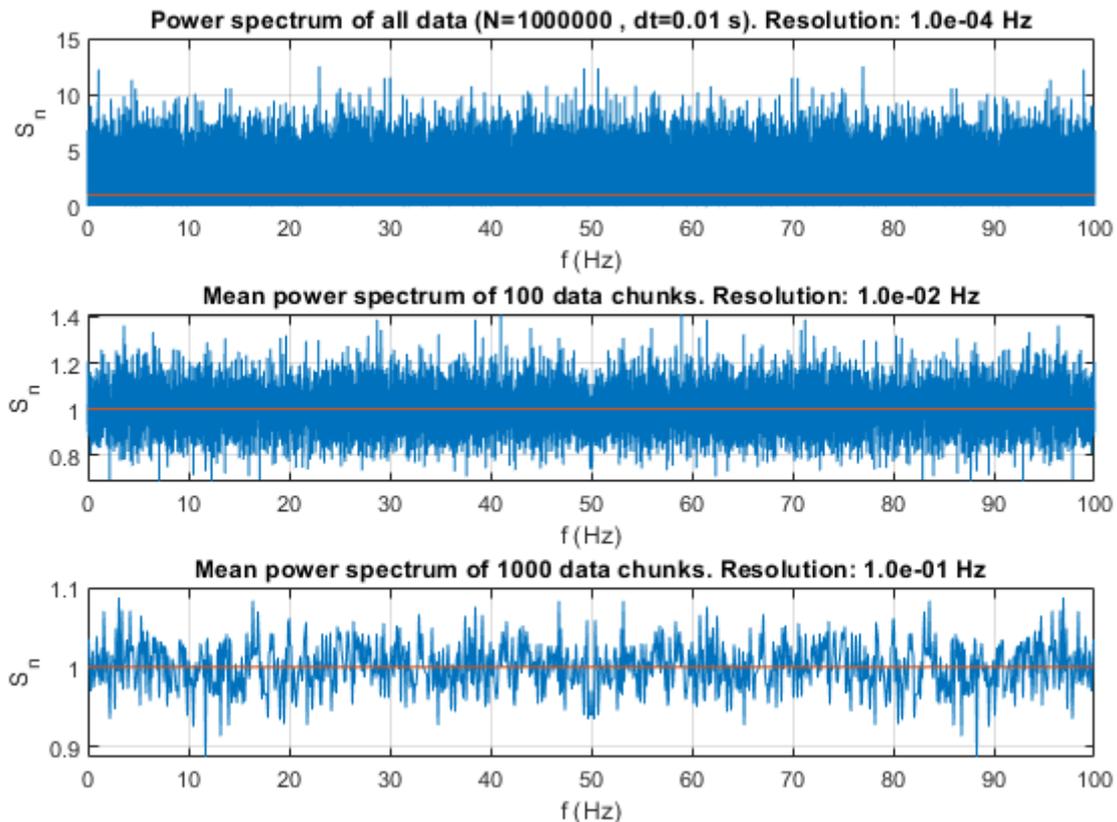


Figura 2.3: Differenti versioni dello spettro di potenza calcolate su di una stessa realizzazione di rumore composta da 1000000 dati con risoluzione temporale di 0.01s, a media nulla e varianza unitaria. Sopra, lo spettro viene calcolato sull'intero set di dati, presentando così una elevata risoluzione in frequenza ma anche una forte oscillazione dei valori. Al centro, la media degli spettri calcolati su 100 spezzoni dei dati di lunghezza 1000: in questo caso la risoluzione decresce di un fattore 100 ma le fluttuazioni dalla media sono più contenute. Sotto, gli spezzoni sono aumentati a 1000 di lunghezza 100 ognuno: la risoluzione diminuisce di un ulteriore fattore 10 e le fluttuazioni si attenuano ulteriormente.

Dopo aver riportato tutte le caratteristiche ed i vantaggi che la schematizzazione di rumore bianco comporta, occorre purtroppo sottolineare come il rumore reale in uscita dai detector si discosti da tale comportamento ideale. In Figura 2.4 è illustrata l'ampiezza spettrale (la radice quadrata dello spettro di potenza) del rumore presente nei 3 interferometri di LIGO (Hanford e Livingston, [35]) e Virgo, attivi nell'ultimo mese di presa dati della fase O2 (Agosto 2017), a cavallo dell'evento GW170814.

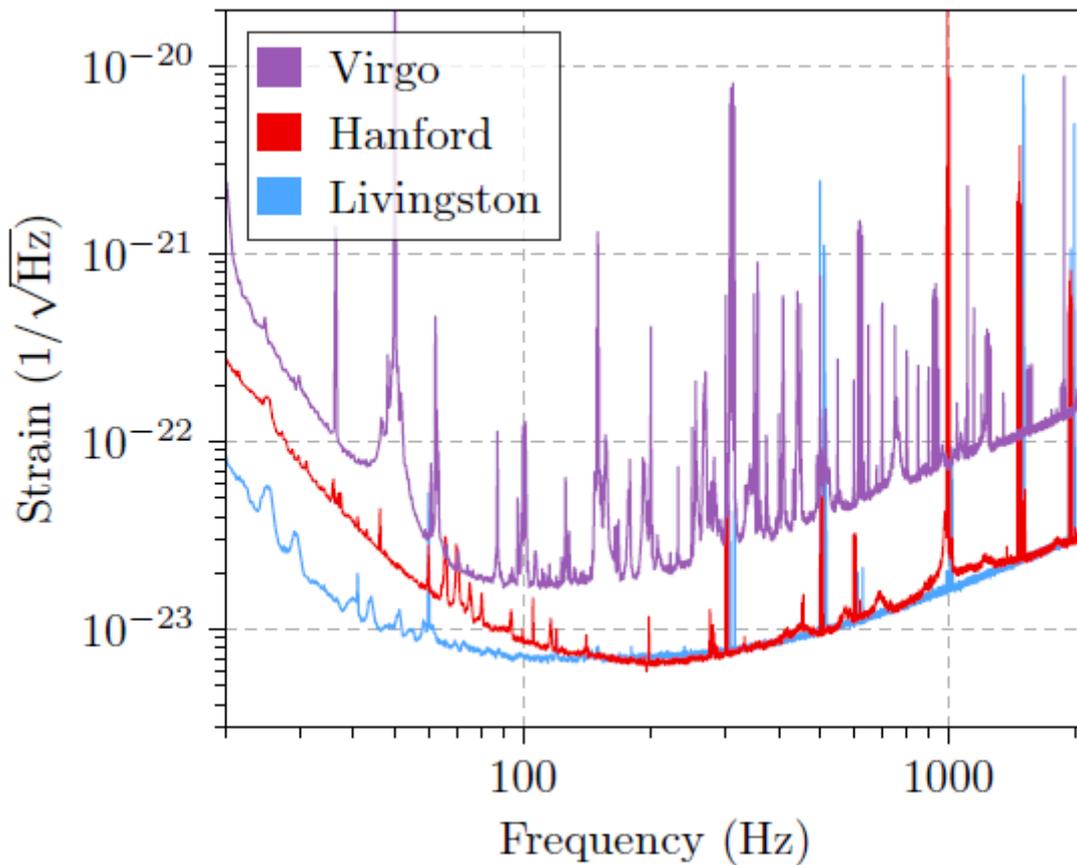


Figura 2.4: Densità spettrale di ampiezza della rete Advanced LIGO -Advanced Virgo, stimata usando 4096s di dati attorno al tempo di trigger dell'evento GW170814. Fonte: [27].

Come si vede, l'andamento è ben lontano dall'essere piatto. Questo avviene a causa delle diverse sorgenti che concorrono al rumore complessivo, ognuna con un range specifico ed una dipendenza diversa dalla frequenza. In Figura 2.5 è mostrata la curva di sensibilità del rivelatore di Virgo nella configurazione originale, prima dell'upgrade, con i diversi contributi ben visibili. Sono ben distinguibili 3 fasi principali [31]:

- Basse frequenze dell'ordine $\sim 1\text{Hz}$: il contributo dominante è dato dal rumore sismico, che arriva all'interferometro ridotto di un fattore 10^{10} grazie all'azione dei superattenuatori, cui ogni specchio è sospeso.
- Frequenze intermedie, nell'ordine dei $\sim 10\text{Hz}$ e dei $\sim 100\text{Hz}$: domina il rumore termico, dovuto principalmente alle sospensioni (nell'immagine, tnPend) ed agli specchi (tnMir). In particolare il rumore termico delle sospensioni limita la sensibilità nel range $10\text{Hz} \lesssim f \lesssim 30\text{Hz}$, mentre il rumore termico degli specchi la limita nel range $30\text{Hz} \lesssim f \lesssim 300\text{Hz}$.
- Alte frequenze, dai 500Hz ed avanzando sui 1000Hz : il contributo dominante di rumore è lo shot noise, chiamato anche rumore di conteggio, ovvero la fluttuazione della risposta in potenza dell'interferometro dovuta alla fluttuazione del numero dei fotoni, descritto dalla statistica di Poisson.

Oltre a questo, si possono notare in questa regione numerosi picchi dovuti all'eccitazione dei modi di violino dei fili attraverso i quali gli specchi sono sospesi.

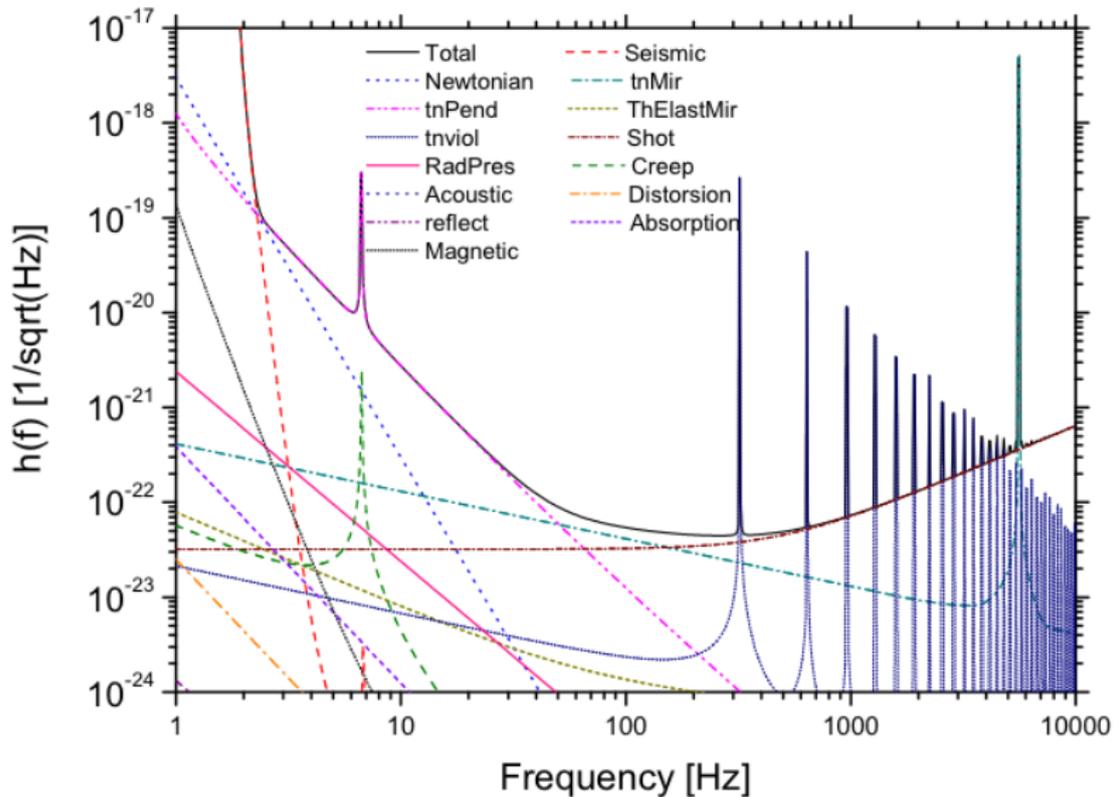


Figura 2.5: Curva di sensibilità di Virgo prima dell'upgrade, ottenuta sommando tutte le potenze spettrali dei vari contributi di rumore riportati all'ingresso dello strumento ed estraendo la radice quadrata del totale.

In situazioni del genere, quando il rumore non segue la statistica piatta del rumore bianco, si parla di rumore colorato, con colori diversi in base alla dipendenza dello spettro dalla frequenza. Dati tuttavia i vantaggi che comporta il lavorare con il rumore bianco, spesso si ricorre a dei filtri che trasformano tale rumore in rumore bianco, chiamati appunto filtri sbiancanti, tenendo però conto del fatto che così facendo si sta modificando anche il segnale immerso all'interno.

2.3 Rivelazione ottimale: filtro adattato

Il problema della rivelazione è dato dal fatto che siamo interessati a stimare un segnale $h(t)$ o la sua trasformata $\tilde{h}(f)$, ma possiamo osservare solo una sequenza di dati $d(t)$, nei quali è contenuto il segnale $h(t)$ addizionato ad un segnale rumoroso $n(t)$ indesiderato. Nel caso migliore in cui il segnale ha forma nota, bisognerà individuarne la presenza all'interno dei dati, stimarne il tempo di occorrenza e l'ampiezza. Seguendo il procedimento di [25], a tal fine si definisce come indicatore fondamentale il rapporto segnale/rumore (Signal-to-Noise-Ratio, SNR), ovvero il rapporto tra il massimo distacco del segnale dal valor medio del rumore e la deviazione standard del rumore:

$$SNR = \frac{\max[h(t)] - \bar{n}}{\sigma_n} \quad (2.4)$$

Evidentemente lo scopo che ci si prefigge è di ridurre l'effetto del rumore, in altre parole di aumentare il rapporto segnale/rumore, facendo attenzione che la distribuzione del rumore rimanga la stessa dopo la stima. La soluzione ad un problema del genere viene dal filtro adattato (matched filter). Immaginando di aver costruito il template del segnale noto in versione campionata h_i in modo da normalizzarne l'energia: $\sum_i h_i^2 = 1$ allora possiamo esprimere i dati a disposizione in forma discreta:

$$d_i = A h_{i-i_0} + n_i$$

Nel caso più semplice in cui il rumore è bianco con $S_n(f) = S_0 = \sigma_n^2$, il filtro si determina attraverso il principio di massima verosimiglianza. Si esprime la verosimiglianza in un certo istante di tempo k come

$$\mathcal{L}(A_k | \{d_i\}) = \prod_i \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\frac{(d_{i+k}-Ah_i)^2}{2\sigma_n^2}} \quad (2.5)$$

e si impongono le condizioni per la sua massimizzazione. Come risultato, la miglior stima dell'ampiezza per ogni ipotetico tempo di coincidenza k è data dalla cross-correlazione dei dati per la forma del segnale:

$$\hat{A}_k = d_k \otimes h_{-i} = \sum_{i=-\infty}^{+\infty} d_{i+k} h_i \quad (2.6)$$

Come si vede, quando le statistiche dei segnali coinvolti sono tutte gaussiane la soluzione del problema è un filtro lineare. Vista nel dominio delle frequenze, in virtù del teorema di convoluzione, la soluzione è data dal prodotto delle trasformate:

$$\tilde{A}(f) = \tilde{d}(f) \tilde{h}^*(f) \quad (2.7)$$

Nel caso in cui il rumore non sia bianco, la soluzione consiste nello sbiancare il rumore tramite un apposito filtro ed applicare all'uscita il filtro lineare trovato per il rumore bianco. In pratica per effettuare lo sbiancamento si applica un filtro $W_w(f)$ che corrisponde alla forma inversa dello spettro di rumore osservato:

$$S_n(f) = |W_w(f)|^{-2} S_0$$

Nel fare questo bisognerà tenere conto del fatto che tale sbiancamento andrà a modificare anche il segnale nascosto. Perciò, nel dominio delle frequenze, avremo:

$$\tilde{A}(f) = \tilde{d}(f) W_w(f) [\tilde{h}(f) W_w(f)]^* = \tilde{d}(f) \tilde{h}^*(f) |W_w(f)|^2 = S_0 \frac{\tilde{d}(f) \tilde{h}^*(f)}{S_n(f)}$$

In questo modo è possibile applicare il filtro adattato ad un segnale in rumore non bianco. Volendo allora dare una formulazione il più generale possibile del filtro adattato, identificheremo il SNR con la variabile $\rho(t)$ definita istante per istante dalla [28]:

$$\rho(t) = \frac{|\langle d|h \rangle(t)|}{\langle h|h \rangle^{1/2}} \quad (2.8)$$

dove il risultato è dato dalla cross-correlazione tra dati e template definita da:

$$\langle d|h \rangle(t) = 4 \int_0^\infty \frac{\tilde{d}(f) \tilde{h}^*(f)}{S_n(f)} e^{2\pi i f t} df \quad (2.9)$$

e normalizzata dall'autocorrelazione del template, nell'ipotesi che questo non sia normalizzato:

$$\langle h|h \rangle = 4 \int_0^\infty \frac{\tilde{h}(f)\tilde{h}^*(f)}{S_n(f)} df \quad (2.10)$$

Si può dimostrare [25] che il filtro adattato rappresenta la soluzione ottimale per migliorare il rapporto segnale/rumore di un segnale di forma nota nascosto in un rumore di cui sappiamo stimare le proprietà statistiche. In altre parole, il guadagno in SNR ottenuto attraverso applicazione del filtro adattato è superiore al guadagno che porterebbe qualsiasi altra tecnica di filtraggio: la sua efficienza rappresenta quindi un limite superiore di quanta informazione si possa estrarre in un dato contesto. E' bene tuttavia sottolineare i limiti di applicabilità di tale soluzione: perchè il filtro funzioni bisogna avere la piena conoscenza della forma funzionale dell'onda di cui si cerca la presenza: se c'è indeterminazione su alcuni parametri si può effettuare la ricerca con una banca di template. Questo tuttavia, come è emerso nel Capitolo 1, non è sempre possibile per molte tipologie di sorgenti di onde gravitazionali, per le quali bisognerà ricorrere a metodi differenti pur se meno efficienti.

2.3.1 Applicazione ai chirp

Una situazione della quale si ha una buona comprensione teorica è quella delle binarie compatte in coalescenza, descritte nella Sezione 1.2. Abbiamo visto come attraverso l'espansione Post-Newtoniana si sia raffinata sempre più la forma d'onda prevista, arrivando a descrivere all'interno del chirp l'effetto di precessione dato dagli spin delle due compatte ed effetti di marea dati dalla deformabilità nel caso delle stelle di neutroni. L'utilizzo del filtro adattato per individuare la presenza di segnali del genere è perciò estremamente promettente. In effetti tutti i segnali di chirp scoperti nelle fasi osservative passate O1 ed O2 sono emersi da due pipeline che lavorano in parallelo sui dati on-line, il Generic Transient Search ed il Binary Coalescence Search: il primo effettua una ricerca unbiased per individuare segnali con morfologia generica con un eccesso di potenza rispetto al rumore [29], il secondo effettua la ricerca mirata di segnali di chirp calcolando il SNR definito nella 2.8 utilizzando una banca di template costruiti per una vasta gamma di possibili masse delle binarie [28].

Senza la pretesa di riprodurre in alcun modo i raffinati meccanismi di selezione, validazione e stima della probabilità di falso allarme sviluppati nel dettaglio in [28], è stata effettuata come esercizio esemplificativo l'applicazione del filtro adattato sui dati del primo evento di onde gravitazionali osservato, GW150914: questi sono disponibili nel Data Release pubblico di LIGO ed includono sia il template selezionato dell'evento sia uno spezzone di dati centrato attorno all'evento stesso. Il campione di dati, epurato da eventuali disturbi esterni noti, è composto da $N = 131072$ valori la cui spaziatura corrisponde ad un tempo di campionamento $\Delta t_s = 0.000244s$, il che significa che l'intero campione spazia su un intervallo di tempo di $32s$. Il template, che presenta le stesse dimensioni e campionatura dei dati, corrisponde ad un modello non-processing nel quale si considerano gli spin allineati con masse individuali $m_1 = 36_{-4}^{+5}M_\odot$ e $m_2 = 29_{-4}^{+4}M_\odot$, ad una distanza di luminosità $D_L = 410_{-180}^{+160}Mpc$. Entrambi sono rappresentati in Figura 2.6. Il segnale, come si vede, ha una ampiezza direttamente confrontabile con quella dei dati dei detector ed è tuttora il più forte segnale generato da coalescenza di buchi neri rilevato, con corrispondentemente il più elevato rapporto segnale/rumore. Questo è rappresentato nella sua interezza, ma è bene sottolineare che è entrato nel range di sensibilità del detector negli ultimi 8 cicli di oscillazione, visibili nel dettaglio a destra.

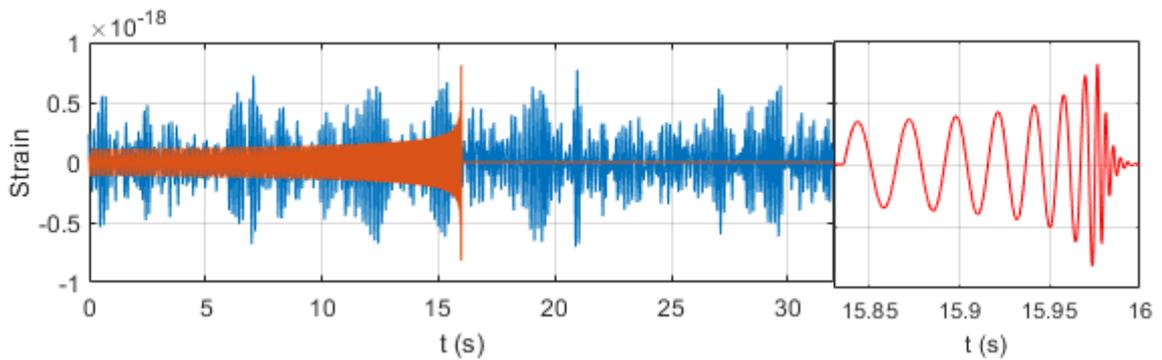


Figura 2.6: Campione di 32s di dati dall'interferometro di Hanford centrati sulla coalescenza del primo evento di fusione osservato, GW150914, posto a confronto con il template che ha ottenuto la miglior stima di SNR.

L'applicazione del filtro adattato consiste, si è visto, nel calcolo della cross-correlazione tra dati e template, eventualmente livellando le ampiezze quando lo spettro del rumore non è bianco. Il processo, effettuato nel dominio delle frequenze, è equivalente al calcolo del prodotto tra i due insiemi di dati avendoli sfasati di tutti i possibili intervalli temporali. C'è un match quando si trova lo sfasamento temporale al quale corrisponde un picco, mentre se nei dati non è presente il segnale cercato il filtro non dovrebbe restituire alcun picco: dunque dall'uscita del filtro si ricavano informazioni sia sul SNR che sul tempo di occorrenza. In Figura 2.7 sono riportati, posti a confronto, gli output che sono stati ottenuti dalla cross-correlazione definita in 2.9 rispettivamente tra template ed il campione di dati di Hanford e tra template e del rumore bianco generato casualmente con la stessa ampiezza. Nel primo caso, dove è effettivamente nascosto il segnale, si ha un match nella zona centrale del campione (il quale era in effetti centrato sull'evento), mentre nel secondo caso non si ha alcun riscontro positivo.

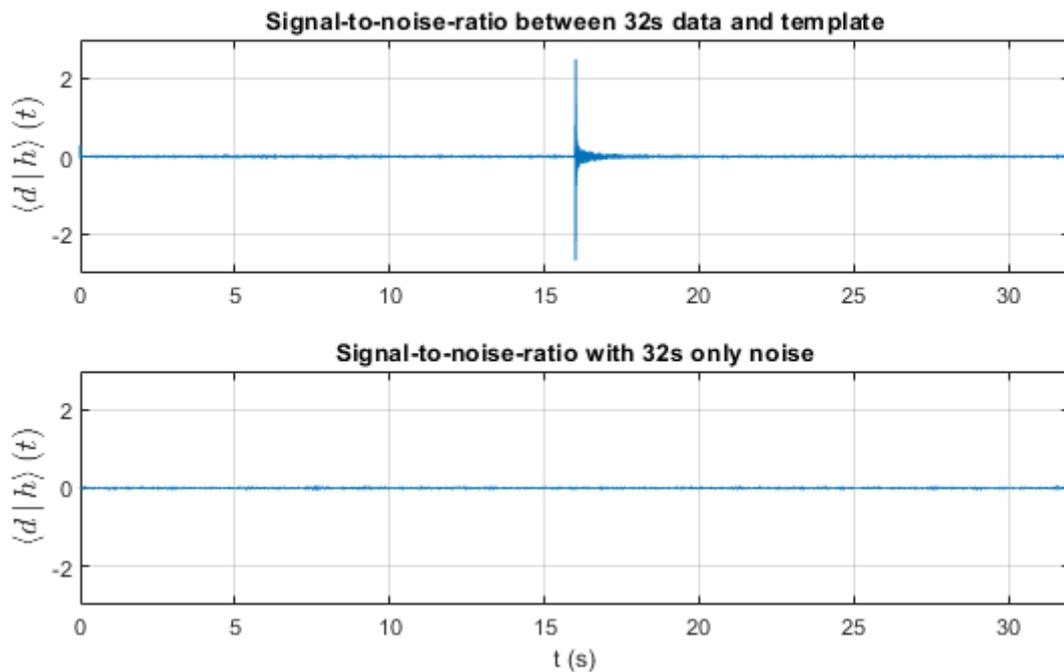


Figura 2.7: Risultato dell'applicazione della 2.9 ai dati contenenti il segnale e a del rumore generato casualmente.

In generale, a partire da un determinato segnale, si può calcolare il suo SNR atteso rispetto ad un determinato rumore di cui sia noto lo spettro $S_n(f)$: questo infatti può essere espresso come

$$\rho^2 = \int_0^\infty \frac{[2|\tilde{h}(f)|\sqrt{f}]^2}{S_n(f)} d\ln(f) \quad (2.11)$$

E' in base alla 2.11 che si sono potute fare le considerazioni riportate nel Capitolo 1 riguardo l'osservabilità delle onde generate da una determinata sorgente. Inoltre essa fornisce anche un modo di visualizzare in modo intuitivo il SNR di un segnale: questo infatti è rappresentato dall'area contenuta tra il grafico normalizzato del segnale $[2|\tilde{h}(f)|\sqrt{f}]$ ed il grafico dell'ampiezza spettrale $\sqrt{S_n(f)}$ rappresentati in scala logaritmica. In Figura 2.8 sono state riportate tali quantità calcolate con i dati ottenuti dal release di LIGO. Per quanto le dimensioni del campione siano troppo piccole per produrre una stima consistente della densità spettrale -nelle pipeline di analisi tale quantità viene ricalcolata in progressivo su spezzoni di dati di 2048s o 4096s- si possono comunque riconoscere i tre andamenti principali dovuti ai diversi contributi dominanti al rumore discussi nella sezione precedente. Se ne può apprezzare qualitativamente la corrispondenza con i grafici riportati in [30].

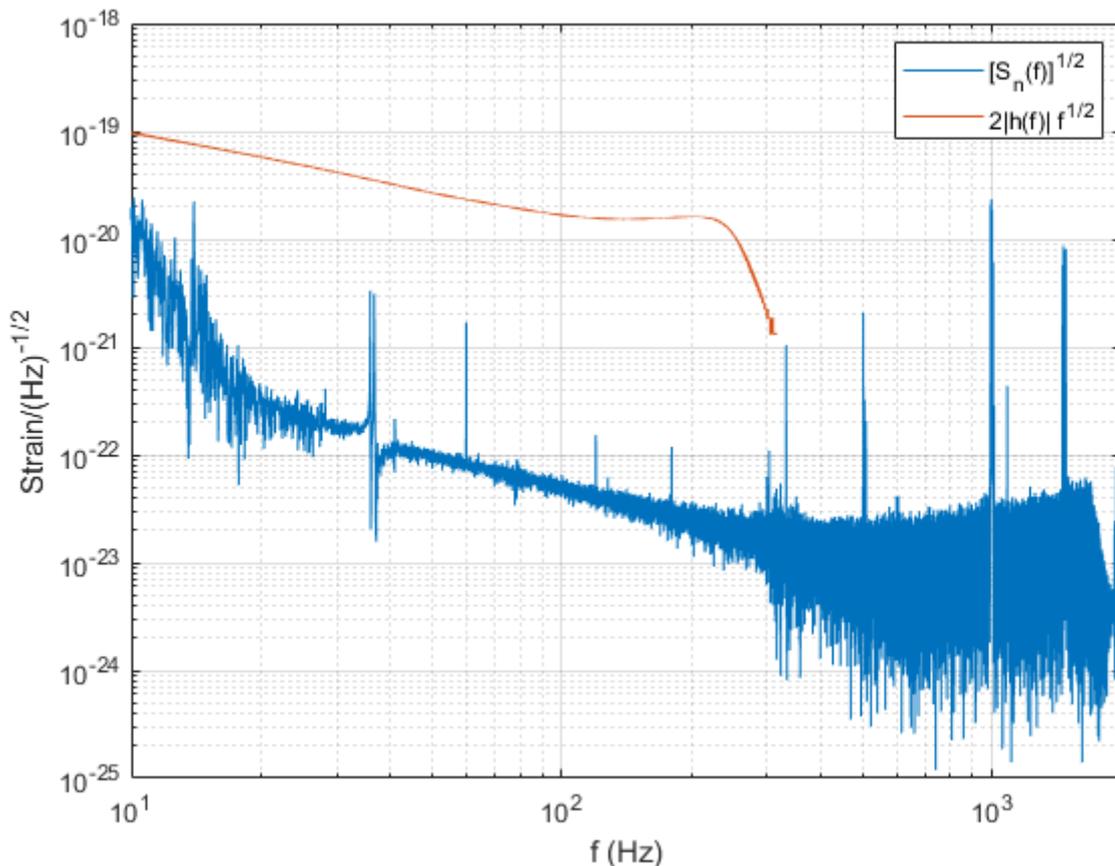


Figura 2.8: In blu, l'ampiezza spettrale calcolata sul campione di dati di Hanford. Date le sue piccole dimensioni, lo spettro monolatero è stato calcolato su tutto l'intero blocco di dati senza mediare su più spezzoni. In rosso, la forma d'onda dell'evento GW150914 normalizzata per il confronto diretto con l'ampiezza spettrale.

Per quello che riguarda la forma d'onda, nel dominio delle frequenze si ha una dinamica diversa rispetto

a quella che conosciamo nel dominio temporale: dato che il sistema effettua un maggior numero di cicli a basse frequenze, l'ampiezza decresce al crescere della frequenza durante l'inspiral, per poi crescere leggermente durante il merger ed infine crollare bruscamente durante il ringdown. Si vede altrettanto chiaramente come questa sia ben al di sopra dello spettro, cosa che si tradurrà in un elevato SNR.

2.4 Analisi dati tramite spettrogrammi

Come si vedrà successivamente, ostacoli come una indeterminazione sulla forma funzionale od i costi computazionali possono impedire di effettuare una ricerca mirata attraverso il filtro adattato: in questi casi bisogna ricorrere a strategie differenti di ricerca. Una possibilità è costituita dalle ricerche di tipo gerarchico, articolate cioè su più livelli. Al primo livello della gerarchia si seleziona un certo insieme, più o meno esteso a seconda delle soglie che si pongono, di candidati eventi. Questi verranno poi selezionati e raffinati nei parametri ai livelli successivi. Se alla fine della gerarchia qualche candidato ha superato le varie scremature e la validazione del cosiddetto follow-up, allora si può affermare che il candidato rappresenta effettivamente un evento gravitazionale.

Nelle prossime sezioni si discuteranno due pipeline di tipo gerarchico sulle quali si ispira il presente lavoro di tesi: la ricerca di segnali continui e periodici da stelle di neutroni e la ricerca di transienti lunghi da instabilità da r-mode in stelle di neutroni giovani. In entrambe le ricerche la selezione dei candidati di primo livello non avviene direttamente sui dati in uscita dagli interferometri: il primo passo consiste infatti nella costruzione di 'short FFT database' (SFDB), dai quali si ricavano degli spettrogrammi. Questi sono delle mappe in tempo-frequenza con una terza dimensione ad indicare la potenza racchiusa nello specifico bin di coordinate (t, f) e consistono in una raccolta di spettri di potenza calcolati, attraverso la FFT, su brevi spezzoni di dati interallacciati tra loro. Su di essi le due pipeline di ricerca effettuano la selezione dei candidati di primo livello, dunque è importante che vengano realizzati con le giuste modalità, pena la mancata selezione dell'eventuale segnale. A seguito i vari passaggi del procedimento, da cui dipende l'output:

- **Scelta del tempo di coerenza.**

Rappresenta il punto cruciale del procedimento [26]. Con il termine tempo di coerenza si intende la lunghezza temporale T_{FFT} che si sceglie per gli spezzoni di dati da trasformare, che si traduce in un numero discreto N di bin. Impostare la lunghezza in bin degli spezzoni di dati dei quali si calcolerà la FFT è il passo più importante di tutto il processo in quanto va a determinare la risoluzione che avrà la trasformata. Come già detto, la risoluzione in termini di bin adimensionali della FFT di uno spezzone di N dati è $\frac{1}{N}$; passando alle unità di misura fisiche, associando ad un bin nel dominio temporale una larghezza Δt_s il corrispondente bin nel dominio delle frequenze avrà larghezza $\delta f = \frac{1}{N\Delta t_s}$. Lo scopo, in questa fase, è che la potenza del segnale rimanga tutta confinata all'interno di un singolo bin di frequenza. Dato che in generale la frequenza del segnale varia nel tempo, in parte per lo spin-down, in parte per l'effetto doppler annuale e siderale (con contributi più o meno dominanti a seconda del tipo di segnale), la richiesta che guiderà la scelta sarà che durante il lasso di tempo di cui si prende lo spezzone la variazione della frequenza non ecceda la dimensione del bin corrispondente nella FFT. Questa richiesta si traduce nella condizione che la lunghezza della FFT non possa superare un determinato valore massimo, pena la diluizione della potenza del segnale in più bin di frequenza.

- **Scelta della finestra.**

Con il termine finestra si intende una determinata forma funzionale per la quale moltiplicare lo spezzone di dati prima di effettuare la FFT. Si potrebbe pensare che la scelta più naturale, unbiased, sia di utilizzare una finestra piatta, che equivale a lasciare i dati così come sono. Tuttavia, a causa del funzionamento della trasformata di Fourier, questa potrebbe non essere la scelta migliore. La trasformata di Fourier infatti lavora idealmente su funzioni (o dati) periodiche o al più assumendo dati con periodo infinito, dunque non considera lo spezzone di dati come un segmento con un inizio ed una fine, ma lo considera come un toro monodimensionale (un anello) nel quale gli estremi del segmento siano contigui, e su di esso fa la trasformazione. In Figura 2.9 è illustrata tale funzionamento. Questo però significa che un eventuale segnale presente nei dati con periodo che non sia un divisore intero del tempo di coerenza presenterà uno sfasamento tra i valori assunti all'inizio ed alla fine del segmento, sfasamento che verrà tradotto in alte frequenze che non hanno alcuna giustificazione fisica.

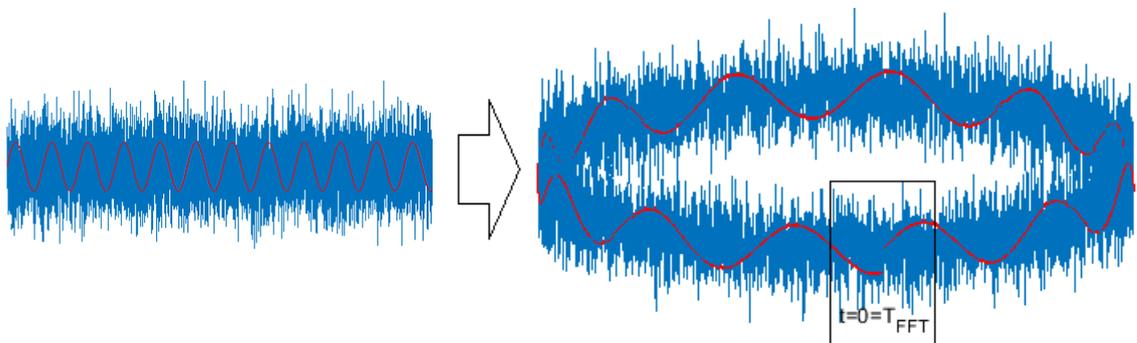
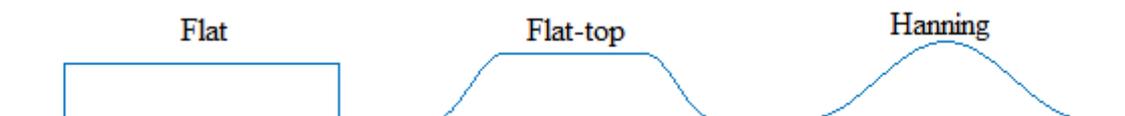


Figura 2.9: Illustrazione del problema dello sfasamento degli estremi di un segmento quando viene calcolata la FFT. Gli istanti iniziali e finali del segmento, racchiusi nel rettangolo, sono contigui, e generalmente questo comporta un salto nei dati che si traduce in componenti ad alte frequenze nella FFT.

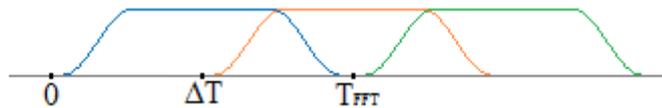
Per ovviare a tale problema si può ricorrere ad una finestra particolare, chiamata flat-top, che consiste in una finestra piatta che però agli estremi va a zero 'dolcemente' ad esempio attraverso una cosinusoide. In questo modo si evitano le fluttuazioni di bordo e si lascia immutato il segnale nella parte centrale del segmento. In base alle necessità del tipo di analisi che si effettua si può optare per diversi tipi di finestra: se ad esempio si è certi della presenza di un determinato segnale e lo si vuole mettere in evidenza si può utilizzare una finestra a campana, chiamata finestra di Hanning. In generale la scelta della finestra flat-top risulta quella più conservativa, in quanto produce una FFT con meno disturbi.



- **Scelta dell'interlacciamento.**

Impostando l'interlacciamento degli spezzoni si stabilisce dopo quanto tempo ΔT calcolare la FFT successiva. Questa scelta andrà a definire la risoluzione temporale dello spettrogramma finale.

Si può scegliere, come avviene frequentemente, di non interallacciare, ovvero semplicemente di prendere gli spezzoni contigui tra loro: così facendo però si vanno a perdere l'informazione e l'energia del segnale in prossimità degli estremi. D'altra parte, interallacciando gli spezzoni si introducono potenziali correlazioni tra le FFT, avendo questi una parte di dati in comune. La scelta più neutra in questo senso consiste nello scegliere un interlacciamento a metà spezzone, ovvero $\Delta T = \frac{1}{2}T_{FFT}$, unitamente ad una finestra flat-top: in questo modo si ha la certezza di non perdere alcuna informazione e di avere i dati dello spettrogramma scorrelati tra loro.



Bisogna comunque sottolineare che introdurre correlazioni non è necessariamente un male, ma quando si opta per un interlacciamento stretto bisogna sapere ciò che comporterà: infatti in base al genere di segnale ricercato può risultare utile talvolta enfatizzare le variazioni dello stesso tramite le correlazioni, dipende dal genere di analisi che si compie.

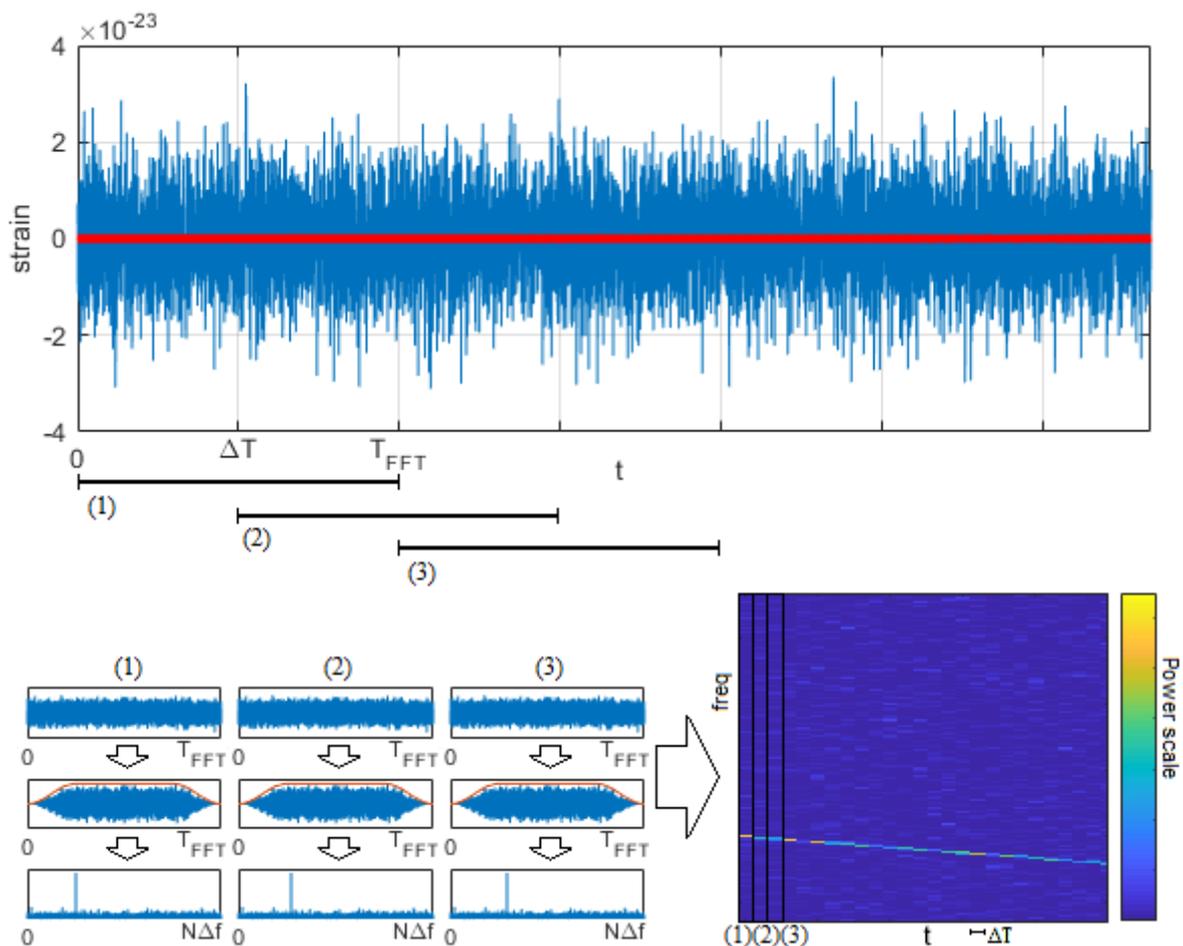


Figura 2.10: Schematizzazione del processo di costruzione di uno spettrogramma: estrazione di spezzoni di dati individuati da T_{FFT} , ΔT , finestra (nell'esempio si usa la flat-top), calcolo dello spettro, unione degli spettri nella mappa. Nell'esempio, un segnale da r-mode immerso in un rumore bianco di ampiezza 10 volte maggiore.

In sintesi, partendo da un campione di dati nel dominio temporale con tempo di campionamento Δt_s , una volta specificati il tempo di coerenza T_{FFT} , il tempo di interlacciamento ΔT e l'eventuale finestrazione si può costruire lo spettrogramma. Questo avrà una definizione temporale data non più dalla frequenza di campionamento ma da ΔT ed una definizione in frequenza data da $\delta f = \frac{1}{T_{FFT}} = \frac{1}{N\Delta t_s}$, coprendo un range di frequenze $0 \leq f \leq \frac{T_{FFT}}{\Delta t_s} \delta f = N\delta f$ che si va a dimezzare nel caso in cui i dati siano reali. Coerentemente con quanto già affermato nella Sezione 2.2, la distribuzione del rumore in questa nuova mappa sarà esponenziale. In Figura 2.10 è rappresentato il procedimento per la costruzione di tale database. In questo modo si ottiene una rappresentazione del segnale nel dominio tempo-frequenza, particolarmente utile in diverse situazioni: non solo infatti permette di individuare segnali già intrinsecamente intensi come i chirp attraverso l'eccesso di potenza sul rumore di background, ma può mettere in evidenza segnali più deboli la cui frequenza varia abbastanza lentamente da sommare l'energia molti più cicli sullo stesso bin. Dalla forma della curva descritta dal picco di un segnale si può verificare se essa segue l'andamento atteso per quel determinato segnale, dunque su questo criterio si può effettuare una prima selezione di candidati eventi.

2.5 Ricerca di segnali continui da stelle compatte asimmetriche

La ricerca di segnali continui e periodici provenienti da stelle compatte -stelle di neutroni, per quanto se ne sa- è tra gli obiettivi posti alla nuova generazione di rivelatori, animata nella sua importanza dai motivi che sono stati espressi nella Sezione 1.3. Si stima che nella nostra galassia ci siano all'incirca 10^9 stelle di neutroni, delle quali appena 2400 sono state individuate grazie all'emissione elettromagnetica, dunque come pulsar. Una parte della popolazione non nota, ad esempio perchè orientata in modo da non investirci con il cono di emissione, potrebbe però emettere onde gravitazionali che rientrino nella banda di sensibilità dei detector. Come già riportato, segnali di questo tipo si aspetta che presentino ampiezze che a parità di distanza della sorgente risultino di ordini di grandezza più piccole rispetto a quelle che caratterizzano segnali transienti come i chirp. D'altra parte, data la presenza permanente di segnali del genere all'interno dei dati e data la lenta variazione della frequenza di tali segnali, la potenza di ogni ciclo si andrebbe a sommare sullo stesso livello di frequenza, portando come conseguenza un potenzialmente elevato rapporto segnale-rumore. Questa considerazione porterebbe a ritenere la tecnica del filtro adattato la più indicata per rivelare segnali continui, tuttavia per la maggior parte dei casi questo non è vero. Infatti, possiamo distinguere due tipi di ricerca per segnali continui:

- (1) Ricerche mirate, effettuate quando la sorgente è nota in posizione, fase e frequenza di emissione: è il caso delle NS che conosciamo dettagliatamente sotto forma di pulsar.
- (2) Ricerche alla cieca, dette anche a cielo aperto: si ricercano segnali continui da tutte le direzioni, spaziando su tutti i possibili valori dello spazio dei parametri.

Per quanto riguarda le ricerche mirate, queste possono essere effettuate attraverso metodi coerenti che applicano il filtro adattato. Se ne possono vedere alcune implementazioni, ad esempio, in [32]. Il rapporto segnale-rumore atteso da una ricerca ottimale dipende ovviamente dall'ampiezza del segnale e cresce con il tempo di osservazione:

$$\text{SNR}_{\text{opt}} = \frac{h_o}{2} \sqrt{\frac{T_{\text{obs}}}{S_n(f)}} \quad (2.12)$$

Il limite delle ricerche mirate è che il numero di potenziali sorgenti sul quale possono essere condotte è limitato. Come sottolineato nella Sezione 1.3.1, l'intensità e di conseguenza l'osservabilità del segnale è in gran parte dipendente dal valore che assume l'ellitticità della NS, valore che -ripetiamo- non è deducibile attraverso le osservazioni dello spettro elettromagnetico. Il punto è che non vi è alcuna garanzia che le pulsar note abbiano una ellitticità tale da consentirne la rivelazione del segnale gravitazionale: per questo motivo è importante allargare il range della ricerca a tutte le possibili sorgenti, ovvero effettuare una ricerca a cielo aperto. Il problema però è che tali ricerche non possono essere effettuate con un metodo completamente coerente: infatti, come argomentato in [33], tale ricerca è complicata dalle enormi dimensioni dello spazio dei parametri, che unite ai lunghi tempi di osservazione richiesti dalla 2.12 portano alla necessità di una potenza di calcolo assolutamente fuori dalla portata attuale. La causa principale di tale complicazione è dovuta alla presenza dell'effetto Doppler dovuto al moto della terra rispetto alla sorgente. Per questa ragione sono state sviluppate delle procedure di tipo gerarchico che permettono una grande riduzione del costo computazionale dell'analisi al prezzo di una perdita di sensibilità relativamente piccola. Come linea generale, tali metodi fanno la scelta di dividere l'insieme di dati in pezzi corti, i quali vengono analizzati coerentemente e poi ricombinati incoerentemente, andando a perdere l'informazione della fase. I tre metodi principali sviluppati sono: lo stack slide, il PowerFlux e la trasformata Hough. In questa sede si discuterà quest'ultimo algoritmo.

2.5.1 Il metodo della trasformata Hough

Lo schema di Hough parte dai dati calibrati in uscita dal detector. In questa sede, per le finalità del presente lavoro, ci si soffermerà maggiormente sui meccanismi che compongono il primo livello di selezione.

Costruzione di un database di FFT

Il primo passo consiste nella costruzione di un database di FFT corte, ognuna calcolata su uno spezzone di dati largo un tempo di coerenza. Nel fare questo si procede anche a ripulire i dati da tutti i disturbi noti presenti, nel dominio temporale. Le logiche ed i meccanismi sono poi quelli descritti nella Sezione 2.4, con la differenza che nel database non si inseriscono le densità spettrali ma direttamente le FFT così come vengono calcolate. La scelta del tempo di coerenza T_{FFT} è fatta sempre per fare in modo che la variazione di frequenza sia contenuta all'interno di un singolo bin: in questo specifico caso, data l'esiguità dello spin-down, l'entità della variazione è principalmente dominata dall'effetto Doppler. Come dimostrato in [33], il limite superiore per il tempo di coerenza è

$$T_{FFT} \lesssim \frac{1.1 \cdot 10^5}{\sqrt{f_{max}/Hz}} s$$

dove f_{max} rappresenta la frequenza massima che tocca la FFT, misurata in Hertz. Dato che il tempo di coerenza massimo è funzione della frequenza, il database viene diviso in blocchi distinti che coprono differenti range di frequenza, ognuno calcolato con un T_{FFT} che dipende dalla massima frequenza del blocco. Nella pratica una scelta ricorrente è di utilizzare un $T_{FFT} = 8192s$ per la banda di frequenze $32 - 128Hz$ ed un tempo $T_{FFT} = 4096s$ per la banda $128 - 512Hz$. La scelta per l'interlacciamento è invece di metà tempo di coerenza, $\Delta T = \frac{1}{2}T_{FFT}$.

Creazione delle peakmap

Per ognuna delle FFT viene calcolato il modulo quadro, ottenendo così uno spettrogramma dal quale si costruirà una peakmap. Questa consiste in una collezione di punti dello spettrogramma che siano dei massimi locali e la cui ampiezza superi una determinata soglia. Per fare questo bisogna però passare per uno step intermedio: come visto nella Sezione 2.2, lo spettro di potenza nel rumore non è piatto. Occorrerà dunque normalizzare le ampiezze dello spettrogramma alla densità spettrale media dei dati, poichè se non si tiene conto di questo nel valutare se un picco supera o meno la soglia, i picchi che si trovano in regioni dello spazio delle frequenze dove lo spettro è più intenso risulteranno sfavoriti. Dunque, partendo dalle N FFT del database si calcola lo spettrogramma $S_{p;i}(f)$ e lo spettro medio autoregressivo $S_{AR;i}(f)$, dove l'indice $i \in \{1, \dots, N\}$ rappresenta il tempo discreto $t_i = i\Delta T$ e contestualmente si costruisce la matrice di rapporti calcolata per ogni bin di frequenza

$$R(i, j) = \frac{S_{p;i}(f)}{S_{AR;i}(f)}$$

dove l'indice j rappresenta i bin di frequenza discreti in modo che $f = j\delta f$. E' questa funzione $R(i, j)$ ad essere confrontata con la soglia θ : i bin di frequenza che si trovino al di sopra di tale soglia e siano anche dei massimi locali verranno selezionati. L'insieme di tutti questi picchi andrà a formare la peakmap. Nella Figura 2.11 si vede un esempio di peakmap ottenuta utilizzando una iniezione hardware durante il run VSR2 di Virgo chiamata pulsar3: si tratta di un segnale continuo di una pulsar immaginaria con frequenza $f_0 = 108.8572 \text{ Hz}$ e spin-down $\dot{f}_0 = -1.46 \cdot 10^{-17} \text{ Hz/s}$, proveniente dalle coordinate di ascensione retta $\alpha = 178.37^\circ$ e declinazione $\delta = -33.43^\circ$. La sua ampiezza, $h_0 = 8.3 \cdot 10^{-24}$, è appositamente esagerata in modo da rendere chiaramente visibile il segnale nella peakmap, della quale è mostrato il dettaglio attorno a tale segnale. L'andamento sinusoidale chiaramente visibile è dovuto alla modulazione Doppler annuale, mentre la modulazione siderale è di ampiezza molto minore ed ha periodo di un giorno terrestre, non apprezzabile alla risoluzione dell'immagine.

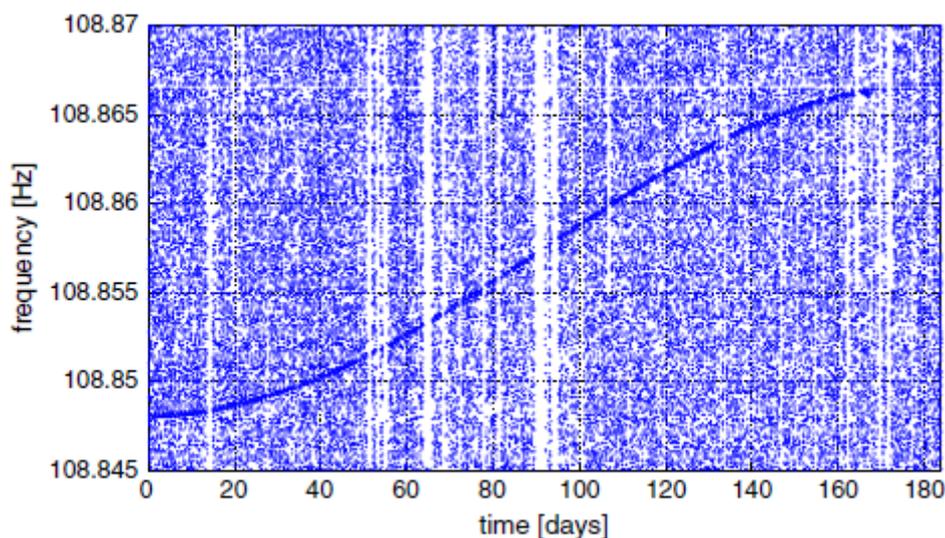


Figura 2.11: Peakmap attorno la frequenza dell'iniezione hardware pulsar3, con $f_0 = 108.8572 \text{ Hz}$, iniettata nei dati di Virgo durante il run VSR2. Fonte: [34].

E' evidente la delicatezza di questo passaggio: se infatti il segnale non dovesse venire selezionato tra i picchi, esso non entrerà mai neanche nei successivi livelli della ricerca, più raffinati, quindi non potrà essere rivelato; dall'altro lato, un picco selezionato finirà nelle successive fasi di lavorazione, che sia dovuto ad un segnale o a rumore. Chiamiamo allora con $p_0 = P(\theta; 0)$ la probabilità di selezionare un picco di rumore al di sopra della soglia θ , ovvero la probabilità di falso allarme, e con $p_\lambda = P(\theta; \lambda)$ la probabilità di selezionare un picco dovuto a segnale con ampiezza spettrale λ data da

$$\lambda = \frac{4 |\tilde{h}(f)|^2}{T_{FFT} S_n(f)}$$

La probabilità $(1 - p_\lambda)$ rappresenta allora la possibilità che un picco di segnale, pur se presente, non venga selezionato. Dato che la peakmap è ottenuta equalizzando uno spettrogramma la cui distribuzione dei dati è esponenziale, la sua distribuzione sarà ancora esponenziale, con valor medio e deviazione standard uguali ad 1. E' possibile allora calcolare la probabilità di falso allarme, la quale sarà (si veda [34]):

$$p_0 = e^{-\theta} - e^{-2\theta} + \frac{1}{3}e^{-3\theta} \quad (2.13)$$

La probabilità di avere selezionato n picchi in una peakmap a partire da una mappa con un numero totale di bin $N_{TOT} = N \cdot N_f$ è di tipo binomiale, con valore d'aspettazione $N_{TOT} p_0$ e deviazione standard $\sqrt{N_{TOT} p_0 (1 - p_0)}$.

Se nei dati è presente un segnale con ampiezza spettrale λ , allora la densità di probabilità dello spettro è quella di un χ^2 normalizzato con 2 gradi di libertà e parametro di decentramento λ . Sfruttando questo fatto, unitamente allo sviluppo dei conti nel limite di piccoli segnali, si può stimare la probabilità di selezionare il segnale come massimo locale sopra soglia come (si veda [34]):

$$p_\lambda \approx p_0 + \frac{\lambda}{2} \theta \left(e^{-\theta} - 2e^{-2\theta} + e^{-3\theta} \right) \quad (2.14)$$

A seguito di una analisi dettagliata in [34], si è arrivati ad individuare una soglia ottimale che massimizzi la probabilità di selezione rimanendo in un contesto di minime assunzioni. Tale valore è stimato come:

$$\theta_{thr} = 2.5 \quad (2.15)$$

La trasformata Frequency-Hough

La trasformata Frequency-Hough (FH) è una particolare implementazione della più generale trasformata di Hough, che è una tecnica robusta per estrarre determinati pattern da immagini digitali. Nelle ricerche delle continue si usa per mappare i punti da un piano in tempo-frequenza, nel quale si cerca un certo andamento, in uno spazio dei parametri del segnale. Nello specifico la FH effettua una mappatura lineare da punti in tempo/frequenza nel piano della frequenza intrinseca della sorgente e del suo spin-down. Prima di fare questo bisogna però correggere l'effetto Doppler: questo in linea di principio è diverso per ogni punto del cielo nel quale potrebbe essere localizzata la potenziale sorgente. A partire dalla peakmap che si è costruita, il vantaggio dell'aver scelto T_{FFT} in modo da tenere la frequenza del segnale in un solo bin si vede nel fatto che per correggere l'effetto Doppler sarà sufficiente semplicemente slittare ogni bin delle FFT su una determinata quantità. Per fare questo bisogna costruire una griglia discreta del cielo, la cui definizione deve essere ottimizzata per avere le minori perdite possibili di informazione.

Considerando la velocità di movimento di un detector rispetto al baricentro del Sistema Solare $\vec{v} = \vec{v}_{orb} + \vec{v}_{rot}$, che si compone della velocità orbitale \vec{v}_{orb} della Terra attorno al Sole e della velocità di rotazione terrestre \vec{v}_{rot} nello specifico sito, che dipenderà dalle sue latitudine β e longitudine λ eclittiche, la scelta della griglia sarà dominata dalla componente orbitale:

$$f_{det}(t) = f_{source}(t) \left(1 + \frac{\vec{v} \cdot \vec{n}}{c} \right) \approx f_{source}(t) \left(1 + \frac{\Omega_{orb} R_{orb} \cos \beta \sin(\Omega_{orb} t)}{c} \right) \quad (2.16)$$

dove Ω_{orb} rappresenta la velocità angolare della Terra nella sua orbita e R_{orb} è il raggio di tale orbita. Imponendo allora che la massima variazione dovuta al Doppler corrisponda alla risoluzione della mappa, ovvero che $\Delta f_{max} = \delta f$, si ricava la risoluzione angolare:

$$\delta \lambda = \frac{1}{N_D \cos \beta} \quad (2.17)$$

$$\delta \beta = \frac{1}{N_D \sin \beta} \quad (2.18)$$

dove N_D rappresenta il numero di bin di frequenza che coprono la banda Doppler:

$$N_D = \frac{B_D}{\delta f}, \quad B_D = \frac{f_{source} \Omega_{orb} R_{orb}}{c} \quad (2.19)$$

Ovviamente i punti che vanno a comporre la griglia non sono distribuiti uniformemente, data la mappatura non lineare. Considerando un ulteriore fattore di miglioramento della risoluzione K_{sky} , il numero di punti di cui si compone la griglia saranno:

$$N_{sky} \simeq 4\pi K_{sky} N_D^2 \quad (2.20)$$

Questo significa che a partire dal database di FFT si farà un numero N_{sky} di correzioni Doppler differenti per andare poi a costruire un numero N_{sky} di trasformate FH. Ipotizzando che un segnale sia presente e che quindi una delle correzioni effettuate abbia compensato nel modo giusto la modulazione della peakmap, tale segnale sarà diventato una linea retta con una bassa inclinazione: date infatti le piccole dimensioni che hanno tipicamente gli spin-down, si può approssimare la dinamica della frequenza del segnale al primo ordine del suo sviluppo, ovvero

$$f = f_0 + \dot{f}_0 (t - t_0)$$

La trasformata FH parte proprio da questo presupposto: supponendo un andamento lineare della frequenza corretta e considerando che i dati noti sono le varie misure di tempo e frequenza della peakmap, si opera una trasformazione tra i piani

$$(t - t_0, f) \rightarrow (f_0, \dot{f}_0)$$

attraverso la trasformazione lineare

$$\dot{f}_0 = -\frac{f_0}{t - t_0} + \frac{f}{t - t_0} \quad (2.21)$$

Dunque ogni picco presente nella peakmap, di coordinate $(t - t_0, f)$, verrà mappato in una linea retta nel nuovo spazio, con inclinazione $-\frac{1}{t-t_0}$. Si può facilmente verificare che con una mappatura del genere tutti i punti che nello spazio originario formano una retta nel nuovo spazio individueranno un fascio proprio di rette con centro (f_0, \dot{f}_0) . Nell'implementazione di tale trasformazione bisogna tenere conto della larghezza δf_H associata ad un singolo bin: ogni picco viene allora trasformato in una striscia estesa, delimitata da due linee rette parallele, che ricopre un range di possibili spin-down

$$-\frac{f_0}{t-t_0} + \frac{f - \frac{1}{2}\delta f_H}{t-t_0} < \dot{f}_0 < -\frac{f_0}{t-t_0} + \frac{f + \frac{1}{2}\delta f_H}{t-t_0} \quad (2.22)$$

I punti che compongono la mappa di FH contengono un valore intero che sta ad indicare quante diverse strisce si sovrappongono su di essi. Inizialmente sono posti tutti a zero, poi in ogni punto che viene toccato da una striscia viene incrementato di 1 il conteggio; in un generico bin allora sarà presente un conteggio n che può essere visto come la somma di conteggi binari:

$$n = \sum_{i=1}^N n_i \quad , \quad n_i \in \{0, 1\}$$

La distribuzione di probabilità della mappa di Hough è di conseguenza binomiale. Allora la probabilità di avere un certo conteggio n in un dato pixel di una mappa di FH costruita a partire da N FFT è data da

$$P_n(\theta; \lambda) = \binom{N}{n} \eta^n (1 - \eta)^{N-n} \quad , \quad \eta = \begin{cases} p_0, & \text{segnale assente} \\ p_\lambda, & \text{segnale presente} \end{cases}$$

con valor medio e varianza dati, rispettivamente, da

$$\begin{aligned} \mu &= N\eta \\ \sigma^2 &= N\eta(1 - \eta) \end{aligned}$$

La selezione dei candidati eventi avviene su quei bin, nelle diverse mappe di Hough, che registrano un conteggio elevato rispetto al valor medio.

L'inclinazione delle strisce dipende dalla scelta del tempo di riferimento t_0 e va ad influenzare la concentrazione dei pixel sulla mappa; si è visto che ponendo tale tempo a metà del tempo totale di osservazione si ha il minor coinvolgimento di pixel da parte del segnale, riducendo la contaminazione dei pixel vicini. Si può inoltre scegliere di tenere conto, nel conteggio sui pixel, della presenza di disturbi non stazionari di rumore: in questo caso ai conteggi binari n_i viene associata una ampiezza w_i che viene moltiplicata ad essi nella sommatoria, che a questo punto non sarà più un numero intero. In questo caso si parla di trasformata di Hough adattiva. In Figura 2.12 si può vedere un esempio di trasformata FH adattiva, calcolata sulla peakmap riportata in Figura 2.11, relativa alla iniezione hardware pulsar3. In base al numero di candidati su cui si ritiene di poter fare la verifica ai livelli successivi, in questo caso dell'ordine di 10^9 , si imposterà la soglia di conteggi che dovrà superare un candidato della mappa FH per essere selezionato.

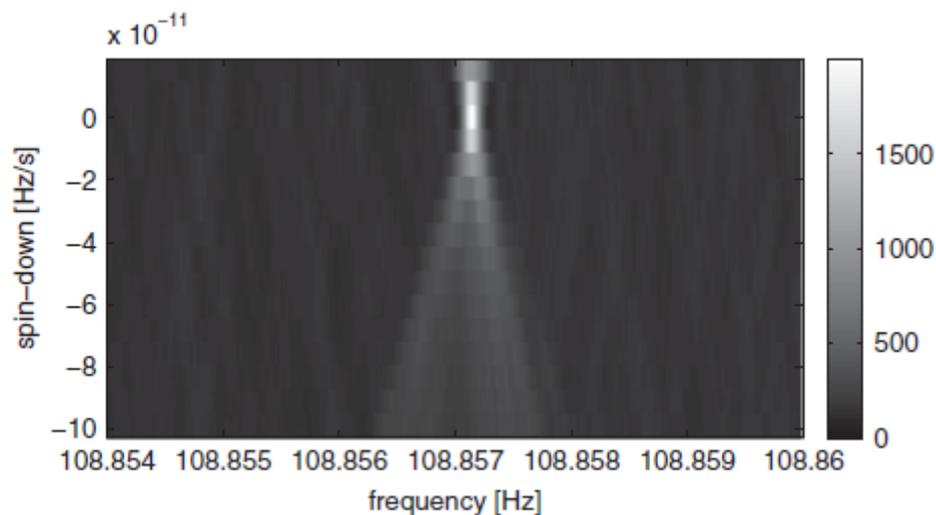


Figura 2.12: Mappa FH adattiva attorno all'iniezione hardware pulsar3, alla frequenza $f_0 = 108.8572\text{Hz}$. Il tempo di riferimento t_0 è posto a metà del tempo di osservazione. Fonte: [34].

Creazione della griglia dei parametri e livelli successivi della gerarchia

Una volta estratti i potenziali candidati di primo livello dalle varie mappe di FH si procede a costruire una grossolana griglia nello spazio dei parametri dove tali candidati verranno selezionati. Vengono costruite 3 griglie: una in frequenza, una in spin-down ed una sulla localizzazione. Con lo scopo di ridurre la probabilità di falsi positivi, viene creato un secondo insieme di candidati analizzando un diverso insieme di dati, che può provenire o meno dallo stesso detector, iterando tutti i passi finora descritti. Infatti, data la natura persistente delle onde continue, un segnale che produce un candidato con certi parametri in un data set dovrà produrre un candidato con approssimativamente gli stessi parametri nel secondo data set. Tra i candidati dei due database vengono quindi fatte le coincidenze: i candidati che sopravvivono diventano candidati di secondo livello.

Sui candidati di secondo livello viene reiterata tutta la procedura, ma andando ad aumentare la definizione attorno ai valori dei parametri emersi. Viene ricreato il database di FFT aumentando il tempo di coerenza (tipicamente di un fattore 64) e viene ricalcolata la trasformata FH in un volume più piccolo nello spazio dei parametri centrato attorno ai parametri dei candidati, andando poi a rifinire ulteriormente la griglia dei parametri. I candidati di ogni data set vengono poi organizzati in cluster, che sono sottoinsiemi di candidati la cui distanza reciproca, definita da una metrica nello spazio dei parametri, sia minore di un certo valore. Questo raggruppamento risulta particolarmente utile in quanto può dare indizi sull'origine comune dei candidati che appartengono ad uno stesso cluster. Infine vengono nuovamente fatte coincidenze tra i candidati di secondo livello attraverso le griglie rifinite.

I candidati che sopravvivono vengono sottoposti ad un processo di verifica finalizzato ad incrementare la confidenza in caso di rivelazione o a scartarli. Questa consiste nell'applicazione di vari criteri verifica non sulle coincidenze ma direttamente sui candidati che le hanno generate, oppure sui picchi delle peakmap che hanno generato tali candidati. Una verifica importante consiste nel comparare le ampiezze dei segnali associati ai candidati che hanno generato una data coincidenza, ampiezze le quali dovrebbero risultare le stesse. Una ulteriore verifica importante è basata sul pattern che segue la radiazione in corrispondenza alla posizione dei candidati coincidenti: ci si aspetta che il numero di picchi nella peakmap

che contribuiscono a questi candidati seguano il pattern di radiazione, con un minor numero di picchi quando l'orientazione del detector risulti sfavorita a causa della rotazione terrestre ed un maggior numero quando l'orientazione del detector è buona.

I candidati che passano anche questo processo di verifica vengono infine sottoposti al cosiddetto follow-up, ovvero una analisi nella quale una piccola porzione dello spazio dei parametri viene rianalizzata con un lungo tempo di coerenza. Solo a questo punto si può dire di aver effettivamente rivelato un segnale gravitazionale.

E' importante sottolineare che sia il processo di verifica che il follow-up non vanno in alcun modo ad aumentare la sensibilità della ricerca: questa infatti è data unicamente dai parametri con i quali viene effettuata la selezione di primo livello, quali la lunghezza delle FFT e le soglie poste alla selezione dei picchi nelle peakmap e dei candidati nelle mappe di Hough. Se un segnale viene mancato al primo livello, non potrà più venir recuperato. Invece la verifica ed il follow-up vanno ad incrementare notevolmente la confidenza di una eventuale rivelazione. In particolare aumentando il tempo di coerenza da T_{FFT} a T'_{FFT} si avrà un aumento del rapporto segnale-rumore di un fattore $\sqrt{\frac{T'_{FFT}}{T_{FFT}}}$ nel passaggio del follow-up.

2.6 Ricerca di segnali da r-mode instabili

Sulla ricerca presentata in questa sezione si innestano le tecniche che verranno sviluppate nel Capitolo 3. La ricerca di onde gravitazionali generate da oscillazioni r-mode in stelle di neutroni giovani prende ispirazione per alcuni aspetti dalla ricerca che si svolge per le continue, anche se con le debite differenze. Occorre sottolineare come prima cosa che per questo tipo di segnali una ricerca coerente è difficilmente applicabile: come si vede dalle equazioni 1.24, 1.25, 1.26, l'evoluzione della frequenza è determinata dal valore effettivo dell'indice di frenata n e della costante k , ma non si ha conoscenza di quale debba essere il loro valore preciso; effettuare una ricerca coerente su un segnale con indeterminazioni, sulla cui possibile durata vi sono solamente ipotesi (potrebbero durare minuti, ore o mesi), richiede nuovamente una potenza di calcolo fuori dalla portata attuale. Di contro, non si ha l'indeterminazione nella localizzazione: infatti tale ricerca viene fatta dove si hanno indizi della possibile presenza di stelle di neutroni di nuova formazione, vale a dire dove si è individuata una coalescenza tra stelle di neutroni, o dove si è registrata una esplosione di supernova o una kilonova. Il metodo di ricerca andrà quindi ad esplorare lo spazio dei parametri intrinseci della sorgente (n, k) anzichè lo spazio dei parametri di localizzazione (β, λ) .

Costruzione del database di FFT

Questa fase è del tutto simile all'analogo delle continue, la differenza principale sta nei parametri utilizzati. Trattandosi di stelle di neutroni appena formate, a causa della conservazione del momento angolare tra sistema progenitore e sistema risultante presenteranno un elevato spin, presumibilmente dell'ordine dei KiloHertz: sono valori molto più elevati di quelli con cui si effettua la ricerca delle continue. In quella situazione la scelta del tempo di coerenza è più banale, riducendosi in molti casi a due blocchi di FFT con i rispettivi tempi di coerenza. Questa scelta tuttavia funziona quando la variazione in frequenza del segnale è dovuta ad uno slittamento Doppler molto più forte dello spin-down. Quando invece è lo spin-down ad essere il fattore dominante, o quando i due effetti sono comparabili, occorre utilizzare diversi T_{FFT} per creare lo SFDB. Considerando inizialmente solo il contributo dello spin-down, scelta che

potrebbe essere realistica quando si ricercano r-mode di breve durata, questo è una funzione variabile del tempo: infatti, andando a derivare la 1.20 si vede che

$$\dot{f} = -\frac{\lambda \alpha^2}{6} f^7$$

Dato che la frequenza è una funzione decrescente del tempo, ne segue che sia la frequenza che lo spin-down raggiungeranno il loro valore massimo in t_0 . La scelta del tempo di coerenza dovrà avere allora come riferimento questo valore, poichè se il segnale rimane in un bin di frequenza a $t = t_0$ per una certa lunghezza T_{FFT} , allora vi rimarrà certamente anche nelle FFT successive. Imponendo allora la richiesta

$$|\dot{f}_0| T_{FFT} < \delta f = \frac{1}{T_{FFT}} \quad (2.23)$$

si ottiene la seguente regola:

$$T_{FFT} < \frac{1}{\sqrt{|\dot{f}_0|}} \quad (2.24)$$

dalla quale, per stelle con spin dell'ordine del KiloHertz, si ottengono tempi di coerenza dell'ordine dei secondi. Per tenere in conto anche l'effetto Doppler, che può influenzare significativamente segnali più deboli (con quindi basso spin-down) o di lunga durata, bisogna inserire nel termine a sinistra della 2.23 anche lo shift dovuto al Doppler $\Delta f = f \frac{v \cdot \hat{n}}{c}$:

$$|\Delta f_0 + \dot{f}_0 T_{FFT}| < \delta f = \frac{1}{T_{FFT}} \quad (2.25)$$

Anche la correzione Doppler cresce con la frequenza, ma date le differenti dipendenze funzionali di $\dot{f} \propto f^7$, $\Delta f \propto f$, ci si aspetta che lo spin-down sia dominante alle alte frequenze iniziali e che invece il Doppler diventi prevalente alle basse frequenze. In definitiva avremo:

$$T_{FFT} < \left| \frac{-\Delta f_0 + \sqrt{(\Delta f_0)^2 + 4 |\dot{f}_0|}}{2 \dot{f}_0} \right| \quad (2.26)$$

Creazione delle peakmap

La creazione delle peakmap avviene con le stesse modalità della ricerca delle continue, con la differenza che le regole per la determinazione della soglia cambiano. Inoltre, dato che la sorgente è localizzata, si può correggere subito l'effetto Doppler ed ottenere l'andamento della frequenza pulito.

In Figura 2.13 è riportato un esempio di peakmap ottenuta da un segnale di r-mode con indice di frenata $n = 7$ iniettato in rumore bianco: si può apprezzare lo stesso genere di andamento $f(t)$ espresso dalla 1.20 ed esemplificato in Figura 1.8.

La trasformata Frequency-Hough generalizzata

Volendo cercare segnali la cui frequenza non è più lentamente variante nel tempo, non è più possibile utilizzare la trasformata FH descritta nella sezione precedente poichè l'andamento di $f(t)$ espresso dalla 1.20 non è più approssimabile a lineare. Si ricorre allora ad un algoritmo generalizzato ad andamenti curvilinei. Il primo passo consiste nel trasformare l'equazione non lineare 1.25 per indici di frenata n

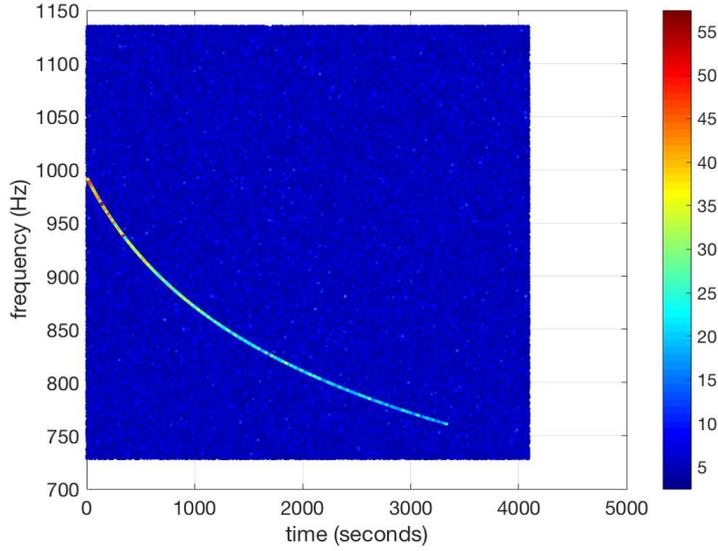


Figura 2.13: Peakmap di un r-mode ($n = 7$) iniettato in rumore bianco, con $h_0 = 1.5 \cdot 10^{-22}$, $f_0 = 993.25 \text{ Hz}$, $\dot{f}_0 = -1.99 \cdot 10^{-1} \text{ Hz/s}$. La mappa è stata realizzata con un tempo di coerenza $T_{FFT} = 2 \text{ s}$. Fonte: [22].

generici in un sistema di coordinate nel quale l'andamento del segnale sia lineare. Questa trasformazione è data da:

$$x = \frac{1}{f^{n-1}} \quad ; \quad x_0 = \frac{1}{f_0^{n-1}} \quad (2.27)$$

dove le variabili x, x_0 hanno le dimensioni fisiche di $[\text{Hz}]^{1-n}$. Attraverso tale trasformazione la 1.25 diventa l'equazione di una retta nel piano (t, x) :

$$x = x_0 + (n-1)k(t-t_0) \quad (2.28)$$

In Figura è mostrato il risultato di una trasformazione del genere applicata alla peakmap esemplificata in Figura 2.13, con $n = 7$, così come viene successivamente data in input alla trasformata FH. Si noti che a causa della mappatura non uniforme i picchi risultano più densi nella regione a bassi valori di x ed allo stesso modo la distribuzione del rumore non è più uniforme: il cambio di coordinate prende punti che sono equispaziati in frequenza e li concentra alle alte frequenze (ovvero a x bassi), diradandoli alle basse frequenze (ovvero a x alti). Data l'indeterminazione sull'effettivo valore di n , viene costruita una griglia di possibili indici di frenata nel range $n \in [2.5, 7]$ con un passo che viene calcolato per fare in modo che la variazione di frequenza passando da $n_1 = n$ a $n_2 = n + dn$ per la durata del segnale sia confinata in un bin di frequenza, ovvero:

$$f_1(t) = f_0 \left[1 + k(n_1 - 1) f_0^{n_1-1} (t-t_0) \right]^{-\frac{1}{n_1-1}} \quad ; \quad f_2(t) = f_0 \left[1 + k(n_2 - 1) f_0^{n_2-1} (t-t_0) \right]^{-\frac{1}{n_2-1}}$$

$$\Delta f_{dn} = f_2(t) - f_1(t) \leq \delta f = \frac{1}{F_{FFT}}$$

Imponendo tale condizione si ricava empiricamente il valore di dn , che dipenderà dunque dalla lunghezza della FFT, da n e dai range di frequenza e spin-down. Per ogni valore di n_i viene effettuata una trasformazione del tipo 2.27, sulla quale viene poi calcolata la trasformata FH. E' evidente che solo la

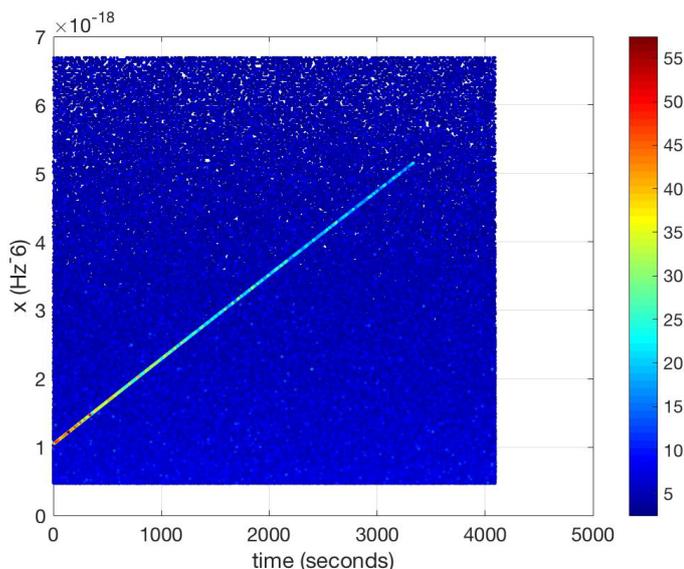


Figura 2.14: La peakmap della Figura 2.13 dopo la trasformazione data dalla 2.27, con parametri $x_0 = 4.706 \cdot 10^{-19} \text{Hz}^{-6}$ e $k = 9.458 \cdot 10^{-23} \text{Hz}^{-5}$. Segnali nel dominio tempo/frequenza vengono trasformati in linee rette nel piano tempo/x se l'indice di frenata della trasformazione è quello corretto. Fonte: [22].

trasformazione fatta con il corretto valore di n trasformerà il segnale in una linea retta.

Quando si calcola la FH bisogna tenere conto del fatto che a causa della mappatura non lineare si produrranno più linee ai valori bassi di x . Per ovviare a ciò si producono le peakmap in una banda di frequenza del 20 – 25% più larga di quella che si vuole analizzare, senza però includere nella selezione gli eventuali candidati che ricadessero nella banda extra: in questo modo si supera il problema della non linearità del rumore al prezzo di un moderato incremento nei costi computazionali. L'output della FH sarà allora un istogramma in 2D nello spazio (x_0, k) , dove i conteggi più elevati corrisponderanno ai parametri più probabili dell'eventuale segnale contenuto. Da ognuna delle mappe FH ottenute ai diversi n_i bisognerà selezionare i candidati più significativi: ogni candidato porterà con sé un valore per x_0, k, n più la localizzazione, valori in base ai quali verranno organizzati in griglie. A partire dai valori di x_0, k sarà possibile risalire, ripercorrendo all'indietro la 2.27, ai valori di f_0, \dot{f}_0 . Per selezionare tali candidati si utilizza una procedura leggermente modificata: in base al numero N_{cand} di candidati che si intende selezionare da ogni mappa, si divide ogni mappa in quadrati distinti e contigui la cui dimensione, in bin, è determinata dividendo gli assi della mappa per $\sqrt{N_{cand}}$. Da ogni quadrato si seleziona un candidato principale più un eventuale candidato secondario purchè abbastanza lontano in termini di bin del parametro k (~ 3) dal primo. Tale procedura è portata avanti con il fine di ricoprire lo spazio dei parametri il più uniformemente possibile. In Figura 2.15 è mostrato uno dei riquadri in cui è stata divisa la trasformata FH della peakmap linearizzata della Figura 2.14, nel quale è localizzato il picco del segnale.

Selezione dei candidati, coincidenze e follow-up

Come nel caso dei segnali continui, l'analisi dei punti precedenti viene ripetuta sui dati dei diversi detector. I candidati emersi vengono sottoposti a verifica delle coincidenze, ovvero si cerca se due (o più, in base ai detector) candidati risultano abbastanza vicini nello spazio dei parametri, calcolando la separazione in termini di bin in x_0 e k : in base alla risoluzione della griglia dei parametri δx_0 e δk si può

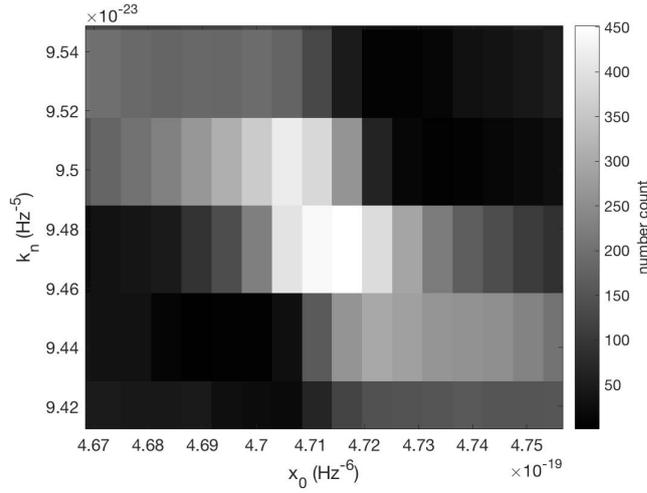


Figura 2.15: Riquadro della mappa FH contenente il segnale iniettato, con i parametri specificati nella Figura 2.14. Il colore corrisponde al numero di strisce che toccano ogni singolo bin. Fonte: [22].

calcolare la distanza tra due candidati come

$$d = \sqrt{\left(\frac{x_{0,2} - x_{0,1}}{\delta x_0}\right)^2 + \left(\frac{k_2 - k_1}{\delta k}\right)^2}$$

Se due candidati si trovano a meno di una distanza di 3 bin l'uno dall'altro allora si considerano coincidenti. Il processo di selezione tramite coincidenze elimina la maggior parte dei candidati ed i rimanenti si considerano potenziali segnali gravitazionali: su di essi viene eseguito il follow-up. Dato che per ogni candidato uscito dalle mappe FH ha con se una stima dei parametri f_0 , \dot{f}_0 , n e del tempo di occorrenza t_0 , si possono utilizzare tali stime per tornare indietro ai dati nel tempo e correggere l'evoluzione in fase del segnale. In pratica si effettua una compensazione del tipo

$$h(t) = h(t) e^{i\phi_{sd}(t)}$$

dove la fase $\phi_{sd}(t)$ è stimata dai parametri del segnale:

$$\phi_{sd}(t) = 2\pi \int_{t_0}^{t_0+T_{obs}} f_0 [1 + k(n-1) f_0^{n-1} (t-t_0)]^{-\frac{1}{n-1}} dt + \phi_{t_0}$$

Se i parametri del segnale fossero stimati perfettamente, una correzione del genere al segnale lo farebbe rimanere in un singolo bin di frequenza per tutta la sua durata nel piano tempo-frequenze, come si vede nell'esempio in Figura 2.16. Questo comporta anche la possibilità di utilizzare lunghezze maggiori delle FFT nel ricalcolare la peakmap, con conseguente miglioramento della risoluzione.

Nella pratica non ci si aspetta che tale correzione sia perfetta a causa della discretizzazione dello spazio dei parametri che limita la precisione in f_0 ; ci si aspetta quindi che il segnale in uscita nella peakmap presenti un lieve spin-up o spin-down in base all'errore commesso. Data l'esiguità di tale variazione in frequenza, questa può essere considerata approssimativamente lineare: per questa ragione nel follow-up si applicano le trasformate FH originali dalle quali si selezionano nuovamente i candidati coincidenti, i quali parametri saranno utilizzati per correggere la peakmap ottenuta in precedenza ed azzerare lo spin-

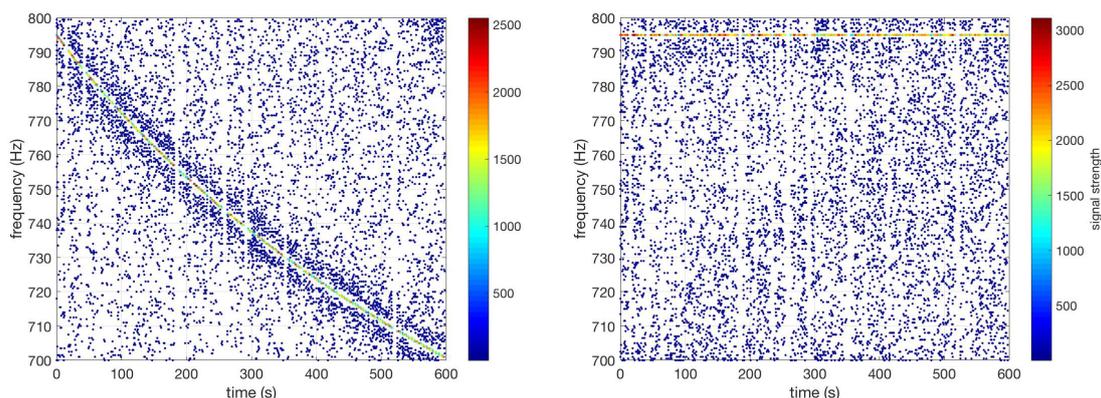


Figura 2.16: A sinistra: peakmap di un r-mode ($n = 7$) iniettato in rumore bianco, con $h_0 = 1 \cdot 10^{-22}$, $f_0 = 795 \text{ Hz}$, $\dot{f}_0 = -2.5 \cdot 10^{-1} \text{ Hz/s}$, realizzata con $T_{FFT} = 2 \text{ s}$. A destra: peakmap del segnale dopo la correzione dell'evoluzione di fase, diventato un segnale monocromatico ad $f_0 = 795 \text{ Hz}$. I tempo di coerenza è aumentato a $T_{FFT} = 4 \text{ s}$. Fonte: [22].

down.

Alla fine di tale procedura il segnale è stato effettivamente riportato in un singolo bin di frequenza: la peakmap viene allora proiettata sull'asse delle frequenze e su tale proiezione si calcola un rapporto critico tra il numero x di picchi alla frequenza stimata, il numero medio μ di picchi dovuti al rumore e la deviazione standard del rumore

$$CR = \frac{x - \mu}{\sigma}$$

Tale rapporto viene confrontato con quello del candidato prima del follow-up, facendo le dovute correzioni per rendere confrontabili le peakmap, le quali spaziano su diverse risoluzioni in frequenza. Tali CR vengono poi dilatati per la sensibilità dei detector attorno alla frequenza del candidato, calcolando un rapporto critico medio. Se tale risultato risulta maggiore di quanto calcolato con il tempo di coerenza iniziale $T_{FFT,0}$ di un fattore $T_{FFT}^{1/4}$, allora su tale candidato verrà effettuata una ulteriore analisi aumentando la coerenza, altrimenti viene scartato.

Anche per questo tipo di ricerca, così come per la ricerca dei segnali continui, la sensibilità è quella data dalla selezione dei candidati di primo livello: se un segnale manca tale selezione non potrà più essere recuperato. E' su questo anello della catena che occorre focalizzare l'attenzione se si vuole migliorare la sensibilità della ricerca: su questo si è concentrato il presente lavoro.

Capitolo 3

Migliorare il SNR: filtraggio in 2D

Nel presentare le pipeline utilizzate per la ricerca di segnali continui da NS asimmetriche rotanti o di segnali da r-mode instabili da NS giovani, si è volutamente concentrata l'attenzione sul primo livello della procedura gerarchica di analisi. E' infatti la selezione di primo livello a determinare la sensibilità della intera pipeline, come si è sottolineato nell'esposizione: se un segnale non ha ampiezza sufficiente non genererà nelle peakmap un numero adeguato di picchi tale da permettere il suo riconoscimento nella trasformata di Hough, di conseguenza non potrà entrare nei livelli successivi della gerarchia. Tuttavia i metodi di ricerca presentati sono stati sviluppati a causa della difficoltà di applicazione di algoritmi coerenti, come ad esempio il filtro adattato, al prezzo di una certa perdita di sensibilità nella ricerca. La considerazione che se ne può trarre è che in questa situazione è possibile individuare una zona grigia, nello spazio dei parametri dei segnali in oggetto, nella quale segnali che sarebbero osservabili con ricerca coerente nella pratica non lo sono a causa della perdita di coerenza. Di conseguenza uno sforzo può essere fatto nella direzione di tentare di ridurre tale zona grigia, ovvero di allargare il range di sensibilità della pipeline. In altre parole, si cerca un modo per aumentare il SNR dei dati in ingresso all'algoritmo di selezione prima di creare le peakmap. Si capisce meglio, in quest'ottica, l'attenzione riservata nella Sezione 2.4 alle modalità di costruzione degli spettrogrammi. Allo stato attuale infatti lo spettrogramma non rappresenta un vero e proprio step distinto nelle pipeline: il database che viene creato è di FFT corte dei dati, che sono collezioni di numeri complessi nei quali le informazioni di ampiezza e fase non sono distinte; da ogni FFT viene stimata la densità spettrale e su di essa vengono poi selezionati i picchi per la costruzione delle peakmap, e questo processo viene ripetuto ad ogni avanzamento sui dati. Una strada promettente consiste invece nel separare i processi, lavorando sullo spettrogramma nel suo complesso prima di effettuare la selezione dei picchi: conoscendo il pattern generale che i segnali presentano nel dominio tempo-frequenze, si può pensare di costruire dei particolari filtri che permettano di alzare il SNR in questo dominio e di conseguenza di far emergere maggiormente il segnale nella selezione dei picchi, rendendolo maggiormente visibile. In questo contesto la tecnica di filtraggio che si utilizza è chiamata filtro di Wiener, che sarà presentato nella successiva sezione. La peculiarità di questo procedimento tuttavia è che lo spettrogramma nel suo complesso non è altro che una immagine monocromatica, ovvero una matrice 2D di numeri reali positivi che definiscono la sua ampiezza pixel per pixel. Occorrerà quindi generalizzare le tecniche di analisi di Fourier, presentate in parte nella Sezione 2.2 per collezioni di dati in 1D, al caso in cui i gradi di libertà siano due, caso per il quale la letteratura scientifica a disposizione è estremamente povera.

3.1 Il filtro di Wiener

Quando si implementa una specifica tecnica di filtraggio lo scopo principale è sempre quello di eliminare o ridurre il rumore nel quale un determinato segnale è immerso. Questo si può fare facilmente qualora il segnale ed il rumore occupino differenti bande di frequenza, ma purtroppo nei casi pratici si ha sempre una sovrapposizione, almeno parziale: nel caso presente lo spettro del rumore copre tutte le frequenze. In questa situazione si utilizza il filtro di Wiener, che supera tale difficoltà attraverso un approccio statistico, riportato in [25]. Presupposto di tale tecnica è che siano note le caratteristiche spettrali del segnale ricercato e del rumore, a partire dalle quali si ricerca un filtro lineare tempo-invariante tale che l'uscita sia più vicina possibile al segnale originale.

Vista nel dominio temporale l'azione del filtro può essere considerata come una funzione $w(t)$ che, applicata in prodotto di convoluzione ai dati in entrata $d(t) = h(t) + n(t)$, restituisce una uscita:

$$\hat{h}(t) = w(t) * [h(t) + n(t)] \quad (3.1)$$

ovvero, nei dati discreti,

$$\hat{h}_k = \sum_{i=-\infty}^{+\infty} w_i d_{k-i} \quad (3.2)$$

Il segnale in uscita vuole essere il più vicino possibile al segnale puro in ingresso. Considerando allora l'errore commesso come la differenza

$$e_i = h_i - \hat{h}_i$$

si ottiene la stima ottimale con il metodo dei minimi quadrati. In altre parole, si impone che il valore atteso del quadrato dell'errore $E[e_i^2]$ sia minimizzato da una determinata scelta del filtro. Per il principio di ortogonalità tale richiesta porta alla condizione:

$$E[(h_i - \hat{h}_i) \cdot d_k] = 0 \quad (3.3)$$

La soluzione del problema si esprime in maniera elegante nel dominio delle frequenze: considerando le densità spettrali di potenza rispettivamente del segnale puro ($S_h(f)$) e del rumore ($S_n(f)$), il filtro di Wiener è dato dal rapporto

$$\tilde{w}(f) = \frac{S_h(f)}{S_d(f)} = \frac{S_h(f)}{S_n(f) + S_h(f)} \quad (3.4)$$

Indicando allora con \mathcal{F}^{-1} l'operatore di Fourier inverso si può calcolare il segnale filtrato come:

$$\hat{h}_k = \mathcal{F}^{-1} [\tilde{d}(f) \cdot \tilde{w}(f)] = \mathcal{F}^{-1} \left[\tilde{d}(f) \cdot \frac{S_h(f)}{S_d(f)} \right] \quad (3.5)$$

Si può realizzare un semplice esempio di filtraggio alla Wiener inserendo un segnale dallo spettro semplice, come una sinusoidale, all'interno di rumore bianco. Nell'esempio in Figura 3.1 si ha, nella prima immagine in alto, una sinusoidale con periodo di 1 s ed ampiezza 0.1 immersa in rumore bianco 10 volte più ampio. La prima ha la densità spettrale che si riduce ad una delta di Kronecker, mentre il secondo ha densità spettrale uniforme su tutto il dominio, come già argomentato nella Sezione 2.2; messi insieme danno il risultato visibile nella seconda immagine dall'alto. Applicando il filtro di Wiener alla trasformazione di tali dati si ha l'output $\hat{h}(t)$ visibile nella terza immagine dall'alto (curva verde) in sovrapposizione al

segnale originale (curva rossa): la fluttuazione dovuta al rumore è praticamente scomparsa ed il segnale ricostruito è vicino all'originale entro un errore di un ordine di grandezza inferiore, come si può vedere dall'ampiezza dell'errore $e(t)$ mostrato nell'ultima immagine in basso.

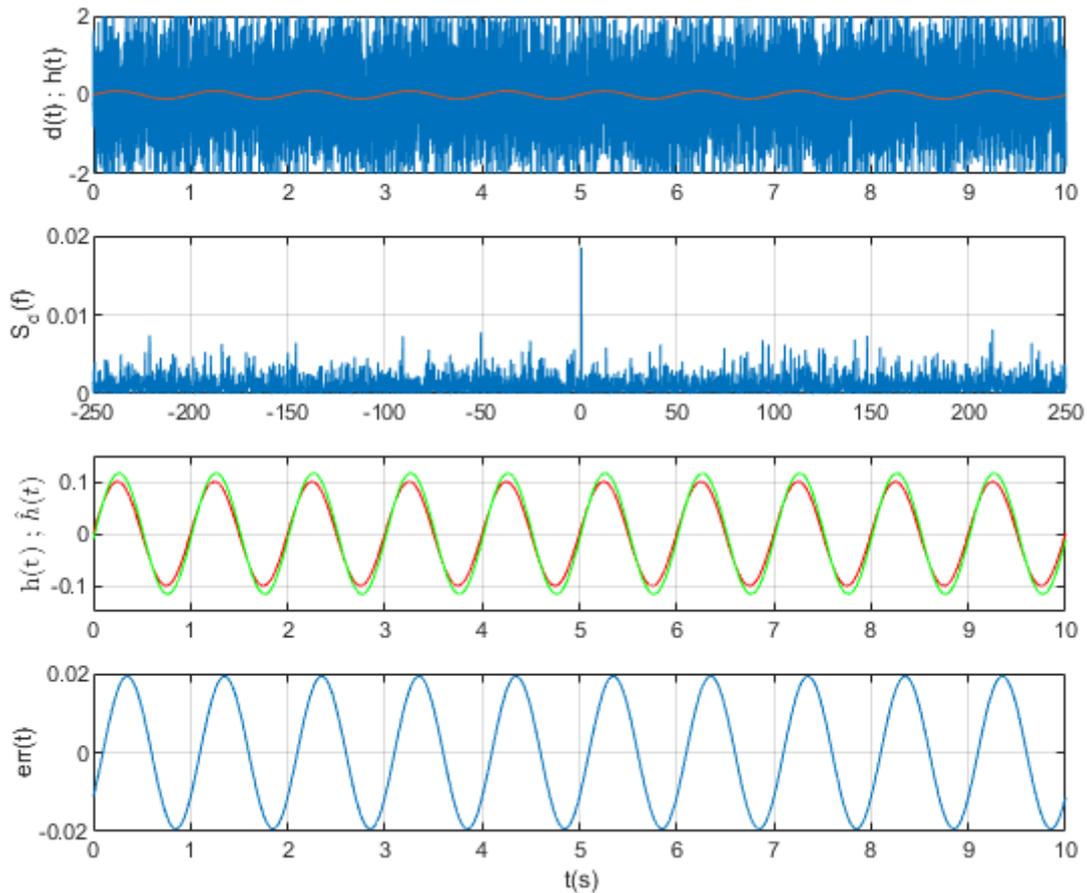


Figura 3.1: Esempio di applicazione del filtro di Wiener. Prima immagine in alto: dati (in blu) generati sommando una sinusoida con periodo 1 s ed ampiezza 0.1 a rumore bianco 10 volte più ampio (in rosso). Seconda immagine dall'alto: densità spettrale dei dati ricentrata sulle basse frequenze per agevolarne la lettura. Terza immagine dall'alto: confronto tra segnale ricostruito $\hat{h}(t)$ (in verde) e segnale originale $h(t)$ (in rosso). Ultima immagine in basso: errore associato al filtraggio, ricavato come differenza tra segnale originale e ricostruito.

In base al genere di segnale ricercato l'implementazione può essere differente: se infatti il segnale ricercato è un transiente breve, dunque è presente solo in una breve porzione dei dati, dato che con lo spettro di potenza si sta mediando l'energia sul tempo di osservazione, su lunghi tempi di osservazione il denominatore della 3.4 sarà dominato dallo spettro del rumore; si può in questo caso semplificare il calcolo utilizzando il solo spettro del rumore, posto che questo dev'essere noto. Dato però che lo spettro del rumore, nelle sue singole realizzazioni, è estremamente irregolare, si deve ricorrere ad una sua stima migliore: questo può essere fatto ad esempio tramite la media su molteplici sottoinsiemi di dati per ridurre le fluttuazioni, o attraverso il calcolo della media autoregressiva [36], oppure in alternativa si può utilizzare una sua modellizzazione teorica che parte dalla conoscenza dei contributi fisici che lo generano. Nel caso di rumore bianco questo equivale a realizzare il filtro di Wiener semplicemente dallo spettro del segnale cercato: $\tilde{w}(f) = S_h(f)$, tenendo a mente che così facendo occorrerà rinormalizzare i

dati ottenuti in uscita.

3.2 Generalizzazione alle immagini: trasformata di Fourier in 2D

Nel momento in cui si vuole applicare il concetto di filtraggio ad una varietà di dati a dimensionalità maggiore di 1 occorre generalizzare il formalismo delle trasformate di Fourier a spazi con più dimensioni. Assegnato uno spazio in \mathbb{R}^n ad n dimensioni, con coordinate spaziali $\{x_i\}_{i=1\dots n}$ ed una funzione scalare in più variabili $f(\vec{x})$, possiamo portarci nello spazio delle frequenze coniugate e viceversa attraverso le trasformazioni diretta ed inversa:

$$\tilde{f}(\vec{v}) = \mathcal{F}_n[f(\vec{x})] = \int_{\mathbb{R}^n} d^n x f(\vec{x}) e^{-i2\pi\vec{x}\cdot\vec{v}} \quad (3.6)$$

$$f(\vec{x}) = \mathcal{F}_n^{-1}[\tilde{f}(\vec{v})] = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} d^n v \tilde{f}(\vec{v}) e^{i2\pi\vec{x}\cdot\vec{v}} \quad (3.7)$$

Da un punto di vista formale una simile generalizzazione non presenta alcun problema. Ovviamente da un punto di vista implementativo bisognerà prestare attenzione perchè il costo computazionale delle trasformazioni andrà a moltiplicarsi. Dato che siamo interessati ad implementare filtri su immagini in 2D, si focalizzerà l'attenzione unicamente sulla trasformata di Fourier in 2D. Introducendo allora le coordinate spaziali $\{x, y\}$ e le frequenze coniugate $\{u, v\}$, le equazioni precedenti diventano:

$$\tilde{f}(u, v) = \mathcal{F}_2[f(x, y)] = \int_{\mathbb{R}^2} dx dy f(x, y) e^{-i2\pi(ux+vy)} \quad (3.8)$$

$$f(x, y) = \mathcal{F}_2^{-1}[\tilde{f}(u, v)] = \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} du dv \tilde{f}(u, v) e^{i2\pi(ux+vy)} \quad (3.9)$$

Nel momento in cui si vuole implementare tale operazione nel discreto, occorre prestare attenzione ad alcuni passaggi. Passando dal continuo ai dati discreti, si considera una immagine campionata in una matrice $M \times N$ descritta attraverso una funzione $f[j, k]$ con indici discreti $j = 1, \dots, M$ e $k = 1, \dots, N$. Per far lavorare la trasformata di Fourier, in analogia con il caso monodimensionale, si considera tale tabella replicata all'infinito su una tabella periodica, in modo che:

$$f[j \pm M, k] = f[j, k \pm N] = f[j \pm M, k \pm N] = f[j, k]$$

La trasformata 2D di una tabella si definisce in questo modo attraverso successive trasformate 1D applicate prima alle righe e poi alle colonne o viceversa. Dunque, tenendo ad esempio fissato l'indice k della riga, si esegue la DFT in 1D di tale riga; ripetendo il procedimento per tutte le righe si ottiene una nuova matrice intermedia $M \times N$, ancora periodica, ad indici misti u, k dove u rappresenta la frequenza coniugata dell'indice j :

$$F^R[u, k] = \sum_{j=0}^{M-1} f[j, k] e^{-i2\pi \frac{uj}{M}}$$

A questo punto per ogni colonna fissata con indice u si fa la trasformata 1D di tale colonna e reiterando per tutte le colonne, ottenendo una nuova tabella $M \times N$ periodica con indici u, v , dove v rappresenta la

frequenza coniugata dell'indice k :

$$F[u, v] = \sum_{k=0}^{N-1} F^R[u, k] e^{-i2\pi \frac{vk}{N}}$$

Quest'ultima matrice ottenuta rappresenta la trasformata discreta (DFT2) dell'immagine; data la commutatività delle operazioni è evidente l'equivalenza tra l'eseguire le trasformate prima sulle righe o sulle colonne. Unendo le due operazioni, abbiamo allora la DFT2 di una matrice e l'inversa:

$$F[u, v] = \sum_{j=0}^{M-1} \sum_{k=0}^{N-1} f[j, k] e^{-i2\pi \left(\frac{uj}{M} + \frac{vk}{N} \right)} \quad (3.10)$$

$$f[j, k] = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F[u, v] e^{i2\pi \left(\frac{uj}{M} + \frac{vk}{N} \right)} \quad (3.11)$$

Così come per la trasformata in 1D la prima componente di Fourier di una funzione rappresenta la somma (integrale o discreta) della stessa, anche in questo caso la componente $[0, 0]$ restituisce la somma di tutti gli elementi della matrice:

$$F[0, 0] = \sum_{j=0}^{M-1} \sum_{k=0}^{N-1} f[j, k]$$

Con una logica analoga si può applicare lo stesso ragionamento per calcolare l'analogo 2D della FFT, che presenta il vantaggio di ridurre notevolmente il numero di calcoli richiesti.

Precisazione sulle dimensioni fisiche

E' bene ricordare che tale formalismo viene sviluppato con l'intento di filtrare eventuali segnali nascosti all'interno di spettrogrammi: questi oggetti, per come sono stati costruiti nella Sezione 2.4, sono delle matrici in 2D dove la prima dimensione è temporale, precisamente con il passo determinato dall'interlacciamento, e la seconda è nelle frequenze. Si sarebbe portati di conseguenza ad interpretare fisicamente anche le variabili coniugate. Tuttavia, mentre quando si fa la trasformata di Fourier di una serie di dati in ampiezza nel tempo la variabile coniugata ha le dimensioni di una frequenza e rappresenta effettivamente la frequenza del segnale, nel caso in 2D le coordinate coniugate dello spazio di Fourier non hanno alcun significato fisico. Si continuerà tuttavia per comodità di nomenclatura a riferirsi a tali coordinate come frequenze coniugate. Allo stesso modo si continuerà a riferirsi alla densità spettrale come di potenza, anche se di potenza vera e propria, in senso fisico, non si tratta.

3.2.1 Proprietà della trasformata di Fourier in 2D

Date le analogie formali con la versione in 1D, si può verificare che la trasformata in 2D gode delle seguenti simmetrie:

- 1) Se $f[j, k]$ è reale, ovvero $f^*[j, k] = f[j, k]$, allora $F^*[u, v] = F[-u, -v]$
- 2) Se $f[j, k]$ è pari, ovvero $f[-j, -k] = f[j, k]$, allora anche la trasformata lo è: $F[-u, -v] = F[u, v]$
- 3) Se $f[j, k]$ è dispari, ovvero $f[-j, -k] = -f[j, k]$, allora anche la trasformata lo è:
 $F[-u, -v] = -F[u, v]$

- 4) Se $f[j, k]$ è reale e pari, allora anche $F[u, v]$ è reale e pari
- 5) Se $f[j, k]$ è reale e dispari, allora $F[u, v]$ è immaginaria pura e dispari

Per quanto riguarda le operazioni che si possono fare sulle funzioni e sulle trasformate, è possibile verificare le seguenti proprietà:

- **Linearità:** date due matrici con la stessa dimensionalità $M \times N$ $f[j, k]$ e $g[j, k]$, la trasformata della combinazione lineare delle due è la combinazione lineare delle singole trasformate con stessi pesi:

$$\mathcal{F}_2 \{ \alpha f[j, k] + \beta g[j, k] \} = \alpha F[u, v] + \beta G[u, v] \quad (3.12)$$

- **Scalatura:** data una matrice $f[j, k]$ e la sua trasformata $F[u, v]$, applicando alle variabili di partenza una contrazione/dilatazione attraverso due costanti α, β la trasformata subirà un effetto inverso:

$$\mathcal{F}_2 \{ f[\alpha j, \beta k] \} = \frac{1}{|\alpha\beta|} F \left[\frac{u}{\alpha}, \frac{v}{\beta} \right] \quad (3.13)$$

- **Traslazioni/Modulazioni di fase:** data una tabella periodica di una immagine $f[j, k]$ ed una coppia di numeri interi p, q , la tabella traslata $f[j - p, k - q]$ si tradurrà nella trasformata come una semplice modulazione di fase:

$$\mathcal{F}_2 \{ f[j - p, k - q] \} = e^{-i2\pi \left(\frac{up}{M} + \frac{vq}{N} \right)} F[u, v] \quad (3.14)$$

In maniera speculare si può dimostrare che applicando ad una matrice periodica $f[j, k]$ una modulazione di fase attraverso una coppia di numeri interi p, q , questa si tradurrà in una traslazione della sua trasformata:

$$\mathcal{F}_2 \left\{ e^{i2\pi \left(\frac{jp}{M} + \frac{kq}{N} \right)} f[j, k] \right\} = F[j - p, k - q] \quad (3.15)$$

Dato che passando al valore assoluto di una funzione se ne va a perdere la fase, questo significa che gli spettri d'ampiezza e di potenza di una matrice sono invarianti per traslazioni.

- **Convoluzione:** quando si vuole estrarre un determinato pattern da una immagine o mettere in risalto determinate caratteristiche tipicamente si va ad effettuare la convoluzione tra la matrice associata ad una immagine $f[j, k]$ ed una matrice $h[p, q]$ associata ad un filtro con la differenza, rispetto alla convoluzione vera e propria, che la seconda matrice non viene invertita nella moltiplicazione; questa produrrà una matrice di output $g[j, k]$ definita da:

$$g[j, k] = f[j, k] * h[j, k] = \sum_{p, q} h[p, q] f[j + p, k + q] \quad (3.16)$$

Tale operazione, pesante già in una dimensione, diventa decisamente rallentante nella versione in 2D. Fortunatamente, il teorema di convoluzione rimane valido anche a dimensioni superiori: è perciò possibile calcolare la convoluzione tra due matrici semplicemente moltiplicandone le trasformate nello spazio di Fourier e ritornando nello spazio di partenza.

$$f[j, k] * h[j, k] = \mathcal{F}_2^{-1} \{ F[u, v] \cdot H[u, v] \} \quad (3.17)$$

Di conseguenza si potranno applicare al filtraggio di immagini gli stessi principi già visti nel caso monodimensionale.

- **Uguaglianza di Parseval:** rimane valida a tutte le dimensionalità e garantisce la sommabilità della serie di Fourier di una funzione. In termini discreti e per matrici 2D essa afferma l'identità:

$$\sum_{j=0}^{M-1} \sum_{k=0}^{N-1} |f[j,k]|^2 = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} |F[u,v]|^2 \quad (3.18)$$

Oppure, nella versione generalizzata,

$$\sum_{j=0}^{M-1} \sum_{k=0}^{N-1} f[j,k] h^*[j,k] = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F[u,v] H^*[u,v] \quad (3.19)$$

Grazie alle proprietà sopra esposte ci sono tutte le condizioni per implementare un filtraggio di immagini nel dominio di Fourier. La costruzione dello spettro di potenza segue la stessa logica che si è vista nella Sezione 2.2: esso è dato dal modulo quadro delle matrice trasformata

$$S[u,v] = |F[u,v]|^2$$

3.2.2 Esempi di trasformate in 2D

Per comprendere meglio come opera la trasformata di Fourier quando si passa in 2D è opportuno partire da semplici casi che rappresentano la naturale estensione di come si opera in 1D. Per maggiore semplicità si partirà dalla trasformata nel caso continuo, così com'è definita nella 3.8, per poi arrivare a capire le differenze che intervengono quando si passa ad un insieme limitato e discreto. Gli esempi riportati saranno di estrema utilità per la successiva realizzazione di filtri.

Trasformata di due punti

Partiamo dal caso più semplice, nel quale si hanno due punti ad una data distanza, in verticale, centrati sull'origine. Nel continuo tipicamente si schematizza un punto con una delta di Dirac, la quale ovviamente presenta le stesse proprietà che aveva in 1D:

$$f(x,y) = \delta(x,y) = \delta(x) \delta(y)$$

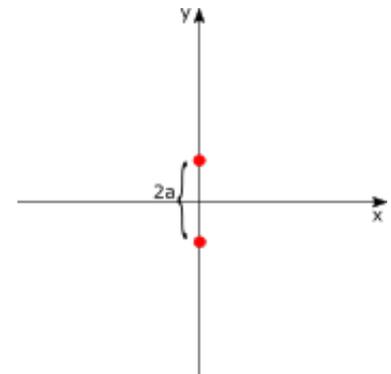
$$\rightarrow F(u,v) = \int dx \delta(x) e^{-i2\pi ux} \int dy \delta(y) e^{-i2\pi vy} = 1$$

Volendo fare la trasformata di due punti, possiamo costruire la funzione come combinazione di due funzioni delta slittate di $\pm a$:

$$f(x,y) = \frac{1}{2} [\delta(x,y-a) + \delta(x,y+a)]$$

A questo punto si può verificare facilmente che la trasformata è data dalla funzione reale:

$$F(u,v) = \cos(2\pi av)$$



Ritroviamo qui la completa analogia con il caso 1D, nel quale i parametri di Fourier di una sinusoide o cosinusoide sono due punti simmetrici. In 2D notiamo che, se i punti sono disposti verticalmente in (x, y) , la trasformata presenta delle curve di livello orizzontali nello spazio (u, v) , come si vede in Figura 3.2. Si noti inoltre che le linee di livello a $F(u, v) = 1$ hanno una spaziatura inversamente proporzionale alla distanza tra i punti.

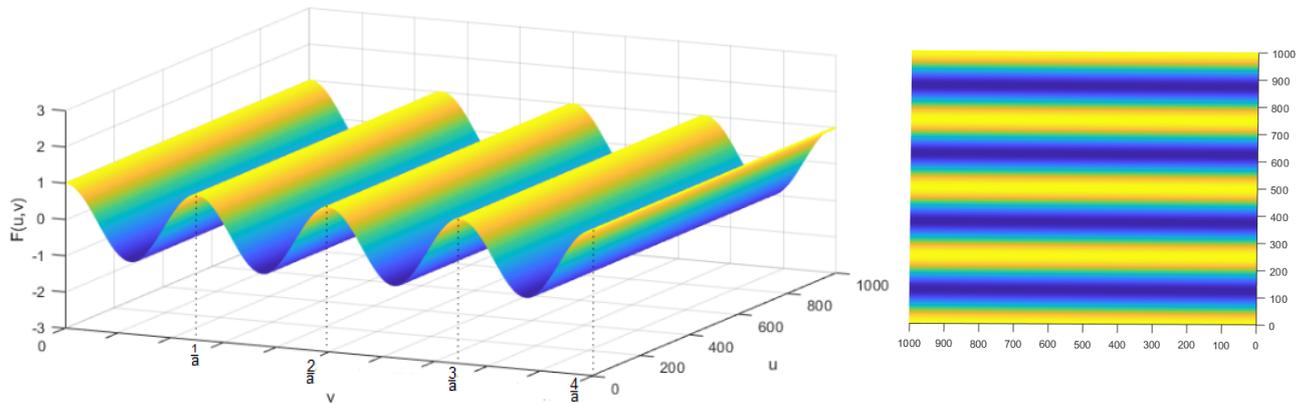


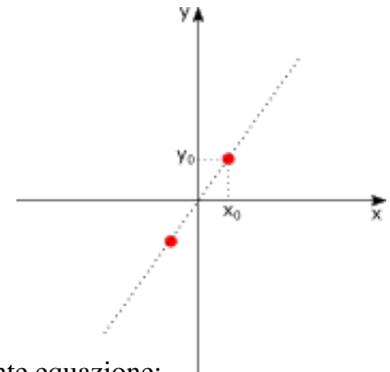
Figura 3.2: Rappresentazione della trasformata di Fourier in 2D di due punti distanti sull'asse y , calcolata analiticamente. A destra, grafico in 3D; a sinistra, curve di livello.

In maniera del tutto analoga ai punti allineati in verticale si può generalizzare calcolando la trasformata di due punti inclinati:

$$f(x, y) = \frac{1}{2} [\delta(x - x_0, y - y_0) + \delta(x + x_0, y + y_0)]$$

$$F(u, v) = \cos[2\pi(ux_0 + vy_0)]$$

Come era ovvio, la trasformata ha la stessa forma funzionale dei punti verticali. Questa volta però nell'argomento compaiono entrambe le variabili dello spazio coniugato. Il fatto importante è che andando a costruire le curve di livello, ad esempio $F(u, v) = 1$, si ottiene la seguente equazione:



$$2\pi(ux_0 + vy_0) = 2k\pi \quad \rightarrow \quad ux_0 + vy_0 - k = 0 \quad \rightarrow \quad v = -\frac{x_0}{y_0}u + \frac{k}{y_0}$$

Questa rappresenta un fascio improprio di rette con un particolare coefficiente angolare: tale inclinazione soddisfa la relazione di ortogonalità con il coefficiente angolare della retta passante per i due punti da cui siamo partiti. Si badi bene che non si può ovviamente parlare di ortogonalità vera e propria, in quanto le due funzioni vivono in spazi diversi. Tuttavia tale relazione rappresenta una indicazione numerica precisa, in quanto suggerisce che nel continuo un insieme di punti allineati con una determinata inclinazione produca nello spazio di Fourier un pattern con una certa periodicità, la cui inclinazione è data da:

$$m(x, y) = \frac{y_0}{x_0} \quad \leftrightarrow \quad m(u, v) = -\frac{x_0}{y_0} \quad (3.20)$$

Tale indicazione è di fondamentale importanza e verrà riutilizzata in seguito.

Dato che la finalità è di esprimere le trasformate di linee curve nel piano, si può pensare di affrontare il

problema considerando queste come somme di punti. Nel caso più semplice, quello di una retta, questa può essere schematizzata come un insieme di punti allineati. Per vedere che genere di output produce una configurazione del genere, a partire dai due punti si può pensare di aggiungere man mano più coppie di punti, che per semplicità di esposizione si prenderanno verticali, con la stessa spaziatura $|2a|$.

Aggiungendo due punti abbiamo il seguente risultato:

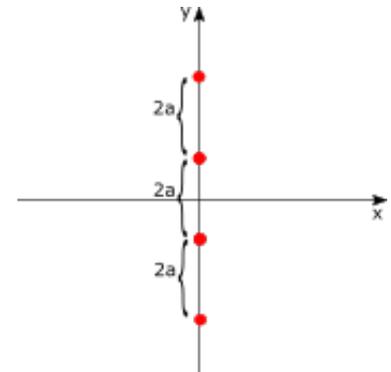
$$f(x, y) = \frac{1}{4} [\delta(x, y - 3a) + \delta(x, y - a) + \delta(x, y + a) + \delta(x, y + 3a)]$$

$$F(u, v) = \frac{1}{2} [\cos 2\pi(av) + \cos 2\pi(3av)]$$

E' immediata a questo punto la generalizzazione della procedura per $2n$ punti:

$$f(x, y) = \frac{1}{2n} \sum_{k=0}^{n-1} [\delta(x, y - a(1 + 2k)) + \delta(x, y + a(1 + 2k))]$$

$$F(u, v) = \frac{1}{n} \sum_{k=0}^{n-1} \cos [2\pi(1 + 2k)av]$$



Per capire quale sia il risultato di questa sommatoria, dobbiamo considerare che andando ad aggiungere punti successivi ai primi due stiamo sommando una serie di cosinusoidi che in generale sono sfasate tra loro: ci aspettiamo dunque che aumentando i termini della somma i vari contributi vadano a cancellarsi, salvo sommarsi nei punti comuni a tutti, ovvero in cui $\cos(2\pi av) = 1$. Iterando la procedura per diversi numeri N di coppie si può apprezzare la dinamica visibile in Figura 3.3, dove per semplicità si è riportata la sola dimensione coniugata v :

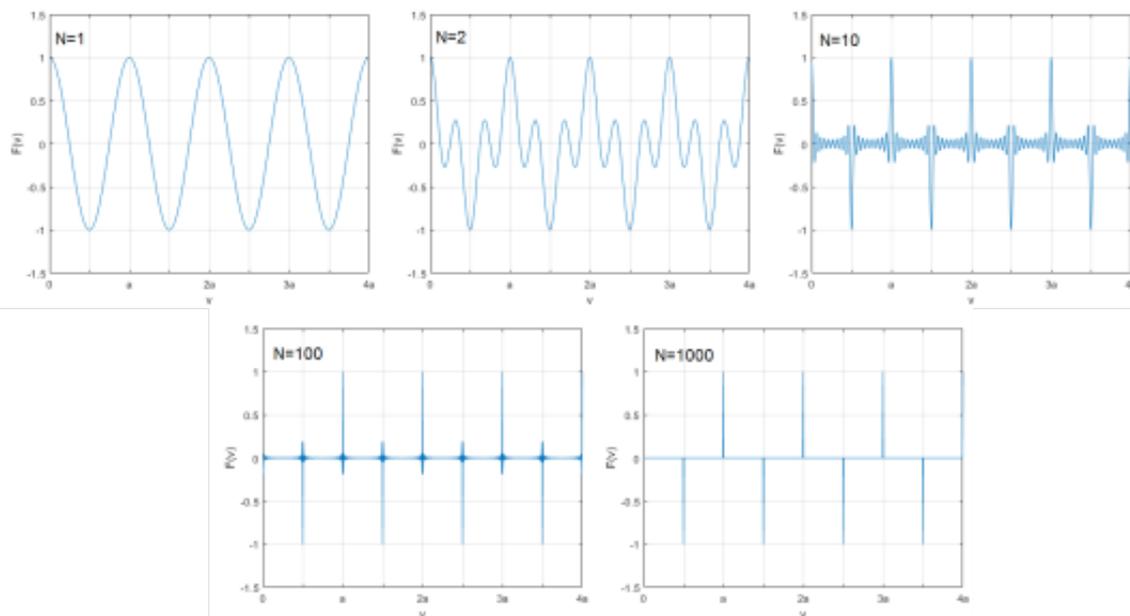


Figura 3.3: Realizzazione di differenti trasformate di $2N$ punti a diversi N , di cui è mostrata la sola dimensione v . Al crescere di N l'interferenza distruttiva tra le cosinusoidi diviene più efficace, lasciando i picchi a periodicità a sempre più stretti.

L'effetto della somma successiva di più termini oscillanti è dunque quello di stringere sempre di più i picchi della cosinusoidale originale, che è determinata dal distanziamento dei punti. L'idea in effetti è che considerando le linee rette come somme di delta di Dirac si possa trovare una certa corrispondenza con le trasformate discrete che si dovranno realizzare all'atto pratico. In Figura 3.4 è mostrata la trasformata nel caso di $N=1000$, osservata nella sua interezza nello spazio (u, v) :

Vista nella sezione in 2D, con la dimensione extra visualizzata nella tonalità del colore, tale trasformata

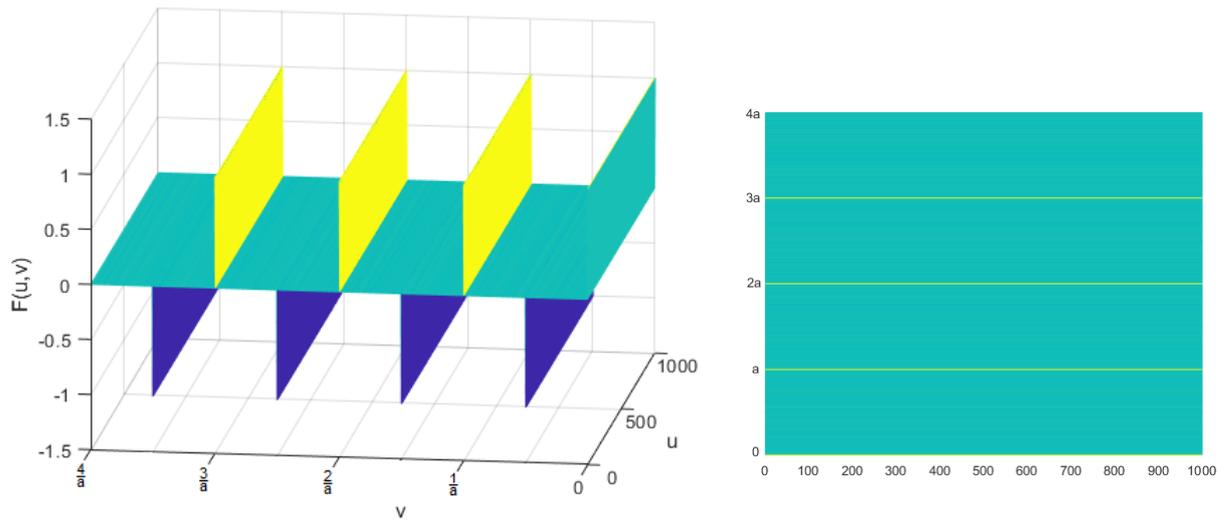


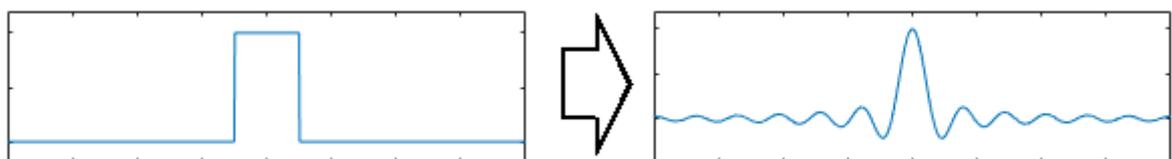
Figura 3.4: Rappresentazione della trasformata di Fourier in 2D di $2N$ punti distanti sull'asse y , con $N = 1000$, calcolata analiticamente. A destra, grafico in 3D; a sinistra, curve di livello.

assume le sembianze di un fascio improprio di rette, la cui inclinazione si può ricavare dalla 3.20 e la cui spaziatura dipende inversamente dalla distanza tra i punti. La configurazione di linea continua corrisponde al limite in cui la distanza tra i punti tende a zero ($a \rightarrow 0$) e nel quale la spaziatura tra le rette trasformate va conseguentemente come $\Delta = \frac{1}{a} \rightarrow \infty$: tra le rette trasformate rimane dunque solo la retta centrale, con inclinazione data dalla 3.20, nel finito. Come si vedrà nella sezione successiva, passando al dominio discreto e periodico bisognerà adottare alcuni accorgimenti.

Trasformata di un dominio rettangolare

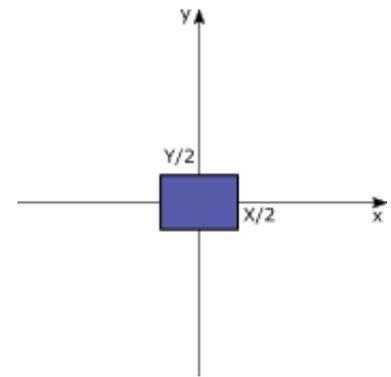
La trasformata in 2D di un dominio rettangolare rappresenta la generalizzazione di un altro caso classico ad una dimensione: in 1D sappiamo infatti che la trasformata di una funzione rettangolo è una sinusoidale con smorzamento iperbolico, chiamata anche «sinc»:

$$\mathcal{F}[\text{rect}(2\pi)] = \frac{\sin(\pi f)}{\pi f} = \text{sinc}(\pi f)$$



Passando a due dimensioni possiamo schematizzare la funzione di origine come la funzione caratteristica di un dominio rettangolare di dimensioni X, Y :

$$f(x, y) = \begin{cases} 1 & , x \in [-X/2, X/2] \ \& \ y \in [-Y/2, Y/2] \\ 0 & , x \notin [-X/2, X/2] \ | \ y \notin [-Y/2, Y/2] \end{cases}$$



Il calcolo della trasformata in 2D risulta particolarmente semplice: si tratta infatti di un dominio rettangolare per il quale è possibile separare l'integrale ed ottenere

$$F(u, v) = \int_{-X/2}^{X/2} dx e^{-i2\pi ux} \int_{-Y/2}^{Y/2} dy e^{-i2\pi vy} = XY \left(\frac{\sin \pi Xu}{\pi Xu} \right) \left(\frac{\sin \pi Yv}{\pi Yv} \right) = XY \text{sinc}(\pi Xu) \text{sinc}(\pi Yv)$$

Tale funzione presenta dei nodi in cui si azzerava ad $u = \frac{k}{X}$ e $v = \frac{k}{Y}$ con k intero, ovvero con un passo che è l'inverso dell'ampiezza del dominio. Se ne può vedere un esempio in Figura 3.5. La trasformata

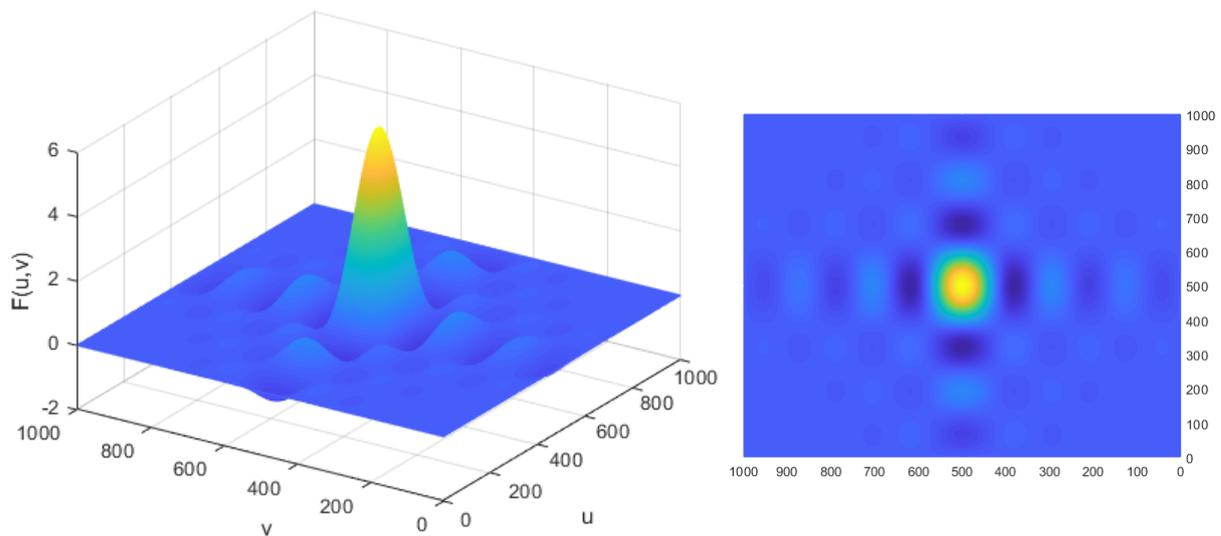


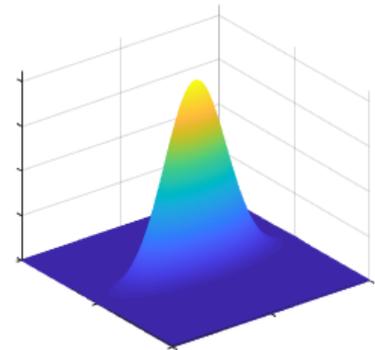
Figura 3.5: Rappresentazione della trasformata di Fourier in 2D di un dominio rettangolare, calcolata analiticamente. A destra, grafico in 3D; a sinistra, curve di livello.

di un dominio rettangolare deve la sua importanza, nell'ambito dei segnali campionati e di conseguenza discreti, al concetto di finitezza nel tempo di una data sequenza di dati. Come si è detto, nel dominio tempo-frequenze di uno spettrogramma è visibile l'andamento della frequenza di un dato segnale nel tempo; questa è descritta da una determinata da una certa funzione, la quale nei casi noti ha una descrizione analitica il cui dominio è infinitamente esteso. All'interno dello spettrogramma, tuttavia, il dominio di tale segnale è limitato al più dagli estremi dello spettrogramma stesso; si potrebbe pensare che sia sufficiente effettuare la trasformata per poi tagliarla quando questa eccede i limiti della matrice, ma tale visione non è esaustiva. Per comprendere l'effetto che tale limitatezza sortisce sulla trasformata si può allora considerare il segnale come prodotto tra la funzione, nel suo dominio infinitamente esteso, ed un dominio rettangolare dalle dimensioni minori od al più uguali alle dimensioni della matrice dell'immagine. In virtù del teorema di convoluzione inverso, la trasformata del prodotto di due domini

nello spazio di partenza è data dalla convoluzione delle trasformate dei due domini nello spazio di Fourier: alla trasformata di Fourier del segnale occorrerà applicare la convoluzione con la trasformata del dominio rettangolare, che come si è visto ha la forma di una funzione sinc in 2D. Questo si tradurrà in una modulazione in ampiezza che sarà presente nella trasformata di ogni segnale finito.

Trasformata di una campana gaussiana

Una certa importanza, legata all'ottica della rivelazione, è assunta dalla funzione normale, o gaussiana. Infatti nelle situazioni in cui si vuole mettere in evidenza un determinato andamento all'interno dei dati a discapito del resto si ricorre spesso alla finestra gaussiana, che consiste nell'effettuare una convoluzione tra i dati e una finestra gaussiana con determinate proporzioni. Tale procedimento può essere applicato anche nel caso 2D in maniera anche più intuitiva: se si vuole mettere in evidenza in uno spettrogramma un segnale con una morfologia semplice, come ad esempio un segnale a frequenza lentamente variabile, si effettua la convoluzione con una finestra gaussiana schiacciata lungo la dimensione temporale; questo, in base al teorema di convoluzione che come si è visto rimane valido anche in 2D, equivale alla semplice moltiplicazione delle rispettive trasformate nel dominio di Fourier. E' quindi di grande interesse vedere come si trasforma una campana gaussiana. Per tenere una certa generalità consideriamo una campana che rappresenta il prodotto di due gaussiane nelle due dimensioni separate, con differenti varianze:



$$f(x,y) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{x^2}{2\sigma_x^2}} \frac{1}{\sqrt{2\pi}\sigma_y} e^{-\frac{y^2}{2\sigma_y^2}} = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{\sigma_y^2 x^2 + \sigma_x^2 y^2}{2\sigma_x^2\sigma_y^2}}$$

Essendo il problema completamente separabile, vale la stessa proprietà che vale in 1D: la funzione gaussiana è autofunzione dell'operatore di Fourier, con la scala invertita:

$$F(u,v) = e^{-2\pi^2(\sigma_x^2 u^2 + \sigma_y^2 v^2)}$$

Dunque la trasformata di una finestra gaussiana è una finestra gaussiana con le larghezze invertite:

$$(\sigma_x, \sigma_y) \rightarrow (\sigma_u, \sigma_v) = \left(\frac{1}{\sigma_x}, \frac{1}{\sigma_y} \right)$$

Nell'esempio in Figura 3.6 si vede come una finestra gaussiana con $\sigma_x = 0.1$ e $\sigma_y = 1$ nel dominio di Fourier inverte completamente le proporzioni. Questo fatto è in piena sintonia con quanto visto negli esempi precedenti: se infatti si hanno all'interno di una immagine delle linee lentamente variabili con una certa pendenza, come espresso dalla 3.20 questa, nello spazio di Fourier, sarà espressa da un pattern con inclinazione ortogonale, dunque anche le proporzioni della finestra gaussiana che si applica devono rispettare tale trasformazione.

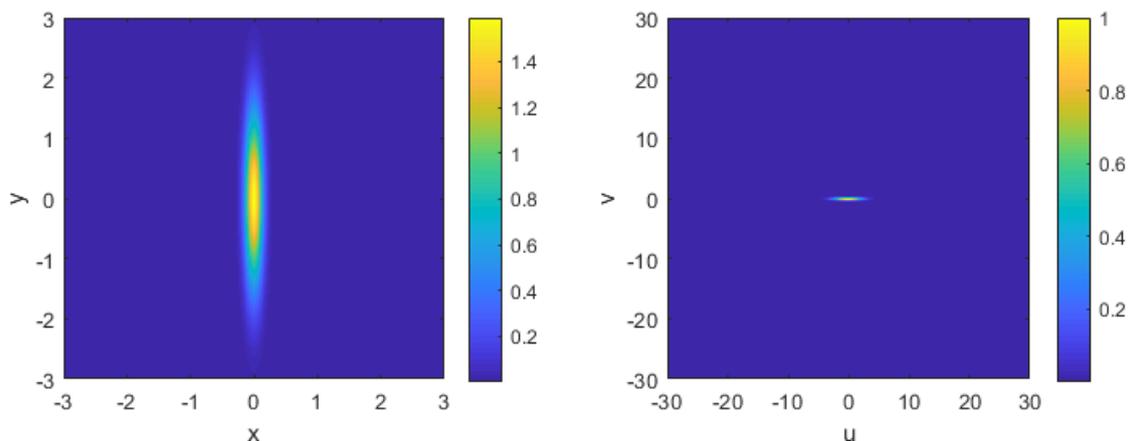


Figura 3.6: Rappresentazione in 2D, con l'ampiezza z in scala di colore. A sinistra: una finestra gaussiana schiacciata lungo la dimensione y . A destra: trasformata di Fourier 2D di tale finestra. Si noti che nello spazio coniugato la finestra trasformata è schiacciata lungo la dimensione u .

3.3 Spettri discreti in 2D

Dovendo realizzare sistemi di filtraggio applicati a matrici in 2D seguendo la logica del filtro di Wiener, nel seguito il focus sarà fissato sul solo modulo della trasformata di Fourier, dal cui quadrato si ottiene la stima della densità spettrale dei dati. Quando si passa dal continuo al discreto il cambiamento più importante riguarda la periodicità dei dati: in pratica la matrice viene considerata come un toro in 2D.

3.3.1 Dai pixel alle linee rette

Ripartendo dall'esempio dei punti nel piano, questi nel discreto diventeranno dei singoli pixel; tuttavia, mentre in \mathbb{R}^2 non c'è limite al numero di punti che si possono aggiungere, in una matrice $M \times N$ tale numero è limitato. Inoltre, trattando lo spazio come un toro in 2D, in base alla mutua distanza dei punti si avranno comportamenti differenti. Partendo da due soli punti, ad essi si associano due distanze: una diretta, individuata dalla congiungente, ed una complementare, individuata attraversando il confine della mappa e riprendendo dall'estremo opposto. Se tali distanze risultano multipli di un divisore intero, a , delle dimensioni della mappa, i punti individuano due siti di un reticolo periodico sul toro: in questo caso il passo di tale reticolo, ovvero la minima distanza tra due siti, va a determinare la distanza tra le creste della cosinusoide trasformata. Se ad esempio due punti disposti in orizzontale, dove la matrice conta M pixel, individuano un reticolo periodico con passo a tale che $M = n_x a$, con $n_x \in \mathbb{N}$, allora il modulo della FFT2 sarà una somma di cosinusoidi sulla variabile di Fourier v con creste poste a distanza, in unità di pixel, di $d = \frac{M}{a} = n_x$. Il fatto interessante è che qualsiasi altra disposizione di punti che si riconduca allo stesso passo reticolare produrrà come FFT2 una funzione oscillante con eventualmente una differente modulazione, ma con le stesse medesime creste. Andando ad aggiungere ulteriori punti sui siti scoperti del reticolo, arrivati a saturazione ci si trova nella situazione corrispondente al limite di infiniti punti del caso continuo: le oscillazioni fuori dai picchi fanno interferenza distruttiva perfetta e rimangono solo le creste isolate. Se invece le distanze tra i punti non rispettano tale condizione, non si individua nessun reticolo periodico e le oscillazioni prodotte dalla FFT2 sono irregolari. In Figura 3.7 è mostrato un esempio nel quale i punti sulla mappa sono posti a distanze tali da individuare un reticolo periodico. In questo caso il passo reticolare individuato è $a = \frac{M}{5}$, dove M è la dimensione della matrice immagine

lungo la dimensione x : questo porterebbe ad avere, in base a quanto detto, una FFT2 oscillante con creste verticali distanti tra loro $d = \frac{5a}{a} = 5$ pixel. Nei primi due casi vengono occupati due soli siti in modo da avere diverse distanze, mentre nell'ultimo vengono occupati tutti e 5 i siti. Come si vede, in tutti e tre i casi le creste dei massimi (in giallo chiaro) sono equispaziate tra loro esattamente di 5 pixel; nei primi due casi, le diverse distanze tra i due punti produce modulazioni differenti tra una cresta e l'altra, mentre nell'ultimo tali modulazioni si compensano esattamente. Si noti infine che, essendo i punti allineati in orizzontale, le linee che formano la trasformata sono continue nel toro, ovvero da dove si interrompono ad un estremo ripartono dall'estremo opposto ripercorrendo a stessa linea: tale configurazione non necessita di alcun taglio da parte di domini rettangolari ed in effetti, come si vede, tali linee non presentano nessuna modulazione nella tonalità del colore lungo la verticale. Tale condizione di idealità verrà perduta non appena si considerino punti allineati in obliquo.

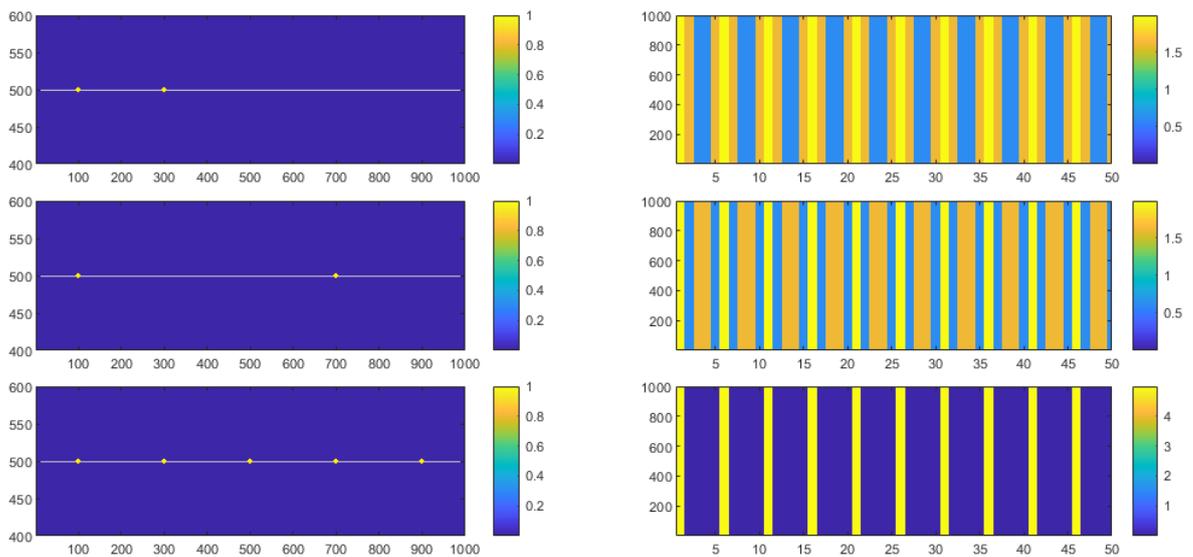


Figura 3.7: Differenti configurazioni di punti allineati in orizzontale in una mappa 1000×1000 pixel in modo da individuare un reticolo periodico di passo $a = 200$ pixel. In alto: a sinistra, coppia di punti posti in $(100, 500)$, $(300, 500)$, con distanze $a = 200$, $4a = 800$; a destra, dettaglio della rispettiva trasformata. Al centro: a sinistra, coppia di punti posti in $(100, 500)$, $(700, 500)$, con distanze $2a = 400$, $3a = 600$; a destra, dettaglio della trasformata. In basso: a sinistra, tutti i siti del reticolo sono riempiti da un punto, ognuno distanziato dall'altro di $a = 200$ pixel; a destra, dettaglio della trasformata.

Andando a disporre i punti su di una retta inclinata si vede come cambiano le cose: pur individuando ancora un reticolo periodico, questo risulta incompleto, in quanto la sequenza di punti, una volta arrivata al termine della matrice, dovrebbe continuare dalla estremità opposta il proprio percorso; bisogna allora considerare la sequenza di punti allineati come il prodotto tra una sequenza infinitamente estesa ed un dominio rettangolare delle dimensioni della matrice, il che produrrà sulle linee trasformate una modulazione in ampiezza. Inoltre introducendo l'inclinazione il passo reticolare verrà ovviamente dilatato di un fattore $\sqrt{1+m^2}$, detta m l'inclinazione della retta su cui giacciono i punti, di conseguenza la spaziatura delle linee della trasformata sarà contratta di un fattore inverso $1/\sqrt{1+m^2}$. Nell'esempio in Figura 3.8 è mostrata la trasformata di un insieme di punti allineati su di una retta di inclinazione in un piano 1000×1000 : i punti sono spazati lungo la dimensione orizzontale di $a_x = 10$ pixel, dunque ne sono presenti in numero $n_x = \frac{M}{a_x} = 100$, e la retta su cui giacciono è inclinata di 30° rispetto l'orizzontale; il

passo reticolare individuato è $a = \frac{2}{\sqrt{3}}a_x$, dunque la nella trasformata si hanno rette inclinate distanti tra loro $d = n_x \frac{\sqrt{3}}{2}$. Si noti inoltre che tali rette risultano modulate nell'ampiezza come previsto, a causa della introduzione della limitatezza.

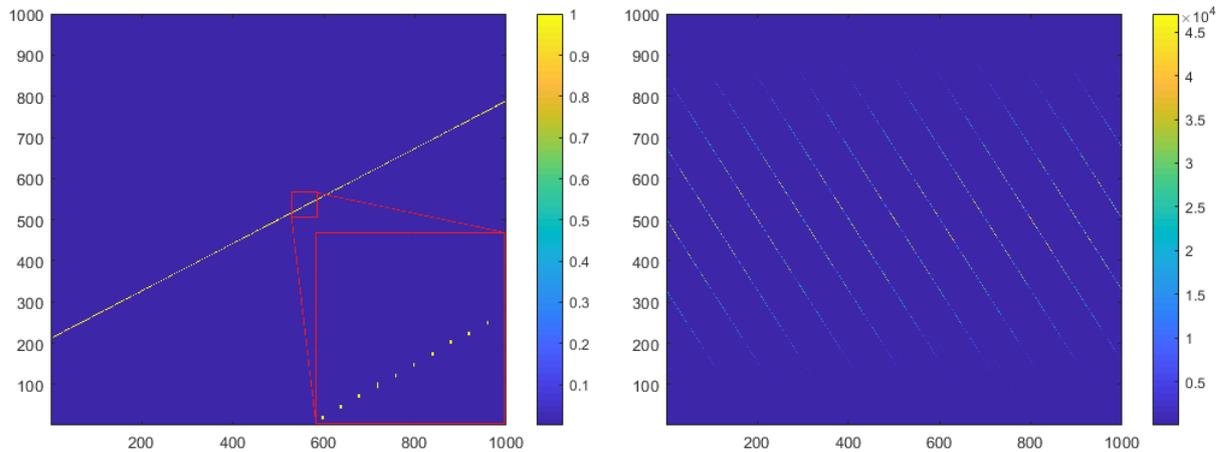


Figura 3.8: A sinistra: mappa di pixel 1000×1000 contenente $m = 100$ punti allineati lungo una retta di inclinazione 30° . A destra: densità spettrale di tale mappa.

L'ultimo passo per comprendere come lavora la trasformata su una linea continua consiste nello stringere ed aggiungere i pixel fino ad essere adiacenti. In questa configurazione, pensando ancora al caso dei punti allineati in orizzontale, il passo reticolare individuato corrisponde alla intera lunghezza della mappa, ovvero $a = 1$ ed $n_x = M$: la trasformata allora consisterà in un pattern di linee verticali con spaziatura data da $d = \frac{M}{a} = M$, ovvero con la periodicità della mappa stessa, dunque la linea visibile sarà una sola. Analoga situazione si avrà andando ad inclinare la retta.

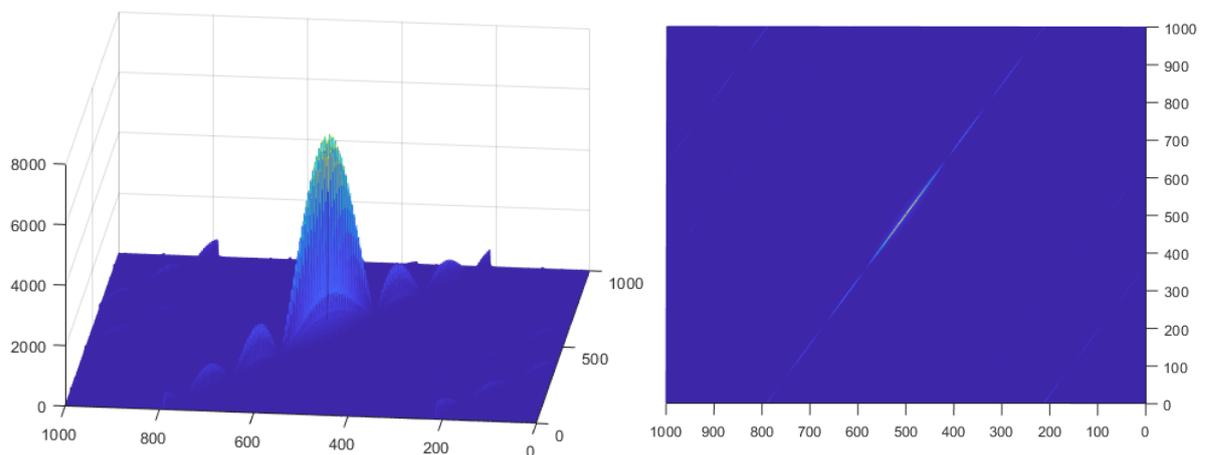


Figura 3.9: Densità spettrale di una mappa 1000×1000 contenente una linea retta di inclinazione -30° . A destra, grafico in immersione; a sinistra, curve di livello.

In Figura 3.9 si vede la configurazione che assume lo spettro di una linea retta: si può apprezzare la modulazione in ampiezza data dalla convoluzione con la funzione sinc in 2D; si noti che le linee secondarie che si vedono non sono le successive creste, ma la prosecuzione della linea principale a causa della periodicità del toro.

Quando si passa dallo spazio continuo ed illimitato ad uno spazio discreto e limitato $M \times N$, occorre infine correggere la relazione che lega le inclinazioni data dalla 3.20. Infatti tale relazione è stata ricavata in una situazione in cui le due dimensioni spaziali sono egualmente estese, mentre nel caso attuale le due dimensioni hanno limitazioni in linea di principio diverse, $M \neq N$. Dato che, come si è potuto apprezzare negli esempi della presente sezione, nello spazio di Fourier le proporzioni tra le grandezze risultano invertite rispetto allo spazio originario, bisognerà correggere la relazione precedente di un fattore $\frac{N}{M}$:

$$\boxed{m(x,y) = \frac{y_0}{x_0} \leftrightarrow m(u,v) = -\frac{x_0}{y_0} \left(\frac{N}{M} \right)} \quad (3.21)$$

Negli esempi appena visti tale correzione era inessenziale a causa delle uguali dimensioni delle matrici, tuttavia dal momento che lo scopo è di applicare le FFT2 agli spettrogrammi, questi avranno una estensione in pixel anche molto differente tra dimensione temporale e dimensione delle frequenze; diventerà allora fondamentale utilizzare tale correzione, pena la valutazione completamente errata della configurazione che assume lo spettro.

3.3.2 Spettri di linee curve

Una volta compreso il meccanismo di funzionamento della FFT2 su di un dominio contenente una linea retta, è facile comprendere come funzioni il meccanismo quando nei dati di una matrice siano presenti linee curve. Quando in 1D si guardava il modulo della trasformata di Fourier di un segnale, l'informazione che era contenuta in esso riguardava la presenza di determinate frequenze, ovvero se il segnale presentasse andamenti più o meno rapidamente variabili. Passando all'analisi in 2D il genere di informazione contenuta dal modulo della FFT2, avendo aggiunto un grado di libertà, si arricchisce notevolmente. Se da una parte infatti si può applicare la stessa logica di una sola dimensione alle due dimensioni separate, guardandole invece insieme si ottengono informazioni sul rapporto incrementale delle forme che ricorrono all'interno dell'immagine. Se ad esempio si considera una determinata curva come somma di brevi segmenti approssimativamente lineari, si può trattare ognuno di essi nella maniera presentata nel paragrafo precedente. Se ne deduce che il modulo della trasformata di una curva contiene l'informazione riguardante tutte le derivate prime presenti sulla stessa. Nelle immagini in Figura 3.10 è riportata una immagine contenente una curva di forma nota, ovvero l'andamento in tempo-frequenza di un r-mode così come ci si aspetta compaia, se presente, in uno spettrogramma; la mappa, di dimensioni 1000×1000 , è riportata nella scala dei pixel e non nella scala delle unità fisiche in quanto non è nell'interesse del momento farvi riferimento, dato che la tecnica può applicarsi in qualunque contesto. A destra è riportato il modulo della FFT2 di tale mappa, ricentrato sulle basse frequenze. Come si vede, esso è individuato da una regione triangolare simmetrica rispetto al centro la cui ampiezza, massima al centro, sfuma con l'allontanarsi verso le alte frequenze. Tale regione rappresenta un fascio proprio di rette che corrispondono alle diverse inclinazioni che assumono tutte le tangenti alla curva dall'inizio alla fine; essendo la curva una funzione monotona decrescente del tempo con derivata prima $\dot{f} \propto -f^7$, ne segue che il massimo ed il minimo dei valori che possono assumere le inclinazioni delle tangenti sono dati dalle inclinazioni agli estremi della curva. Per evidenziare questa caratteristica nelle due immagini in basso della figura sono state messe in sovrapposizione, in rosso, le rette tangenti agli estremi, mostrando nella trasformata le corrispondenti: come si vede, queste rappresentano i limiti entro i quali si concentra l'informazione del modulo della FFT2. In particolare, si può apprezzare come il lato più verticale abbia ampiezza maggiore:

questo è dovuto al fatto che nella morfologia della curva le inclinazioni più basse sono più ricorrenti di quelle più elevate, per cui la densità delle basse inclinazioni risulta maggiore. Si noti inoltre che trattandosi di un arco di curva regolare semplice monotona decrescente, il modulo della sua trasformata è localizzato nel primo e terzo quadrante della mappa di Fourier; una curva con andamento regolarmente crescente, come ad esempio è l'andamento dei chirp in tempo-frequenza, avrà, al contrario, il modulo della trasformata nel secondo e quarto quadrante.

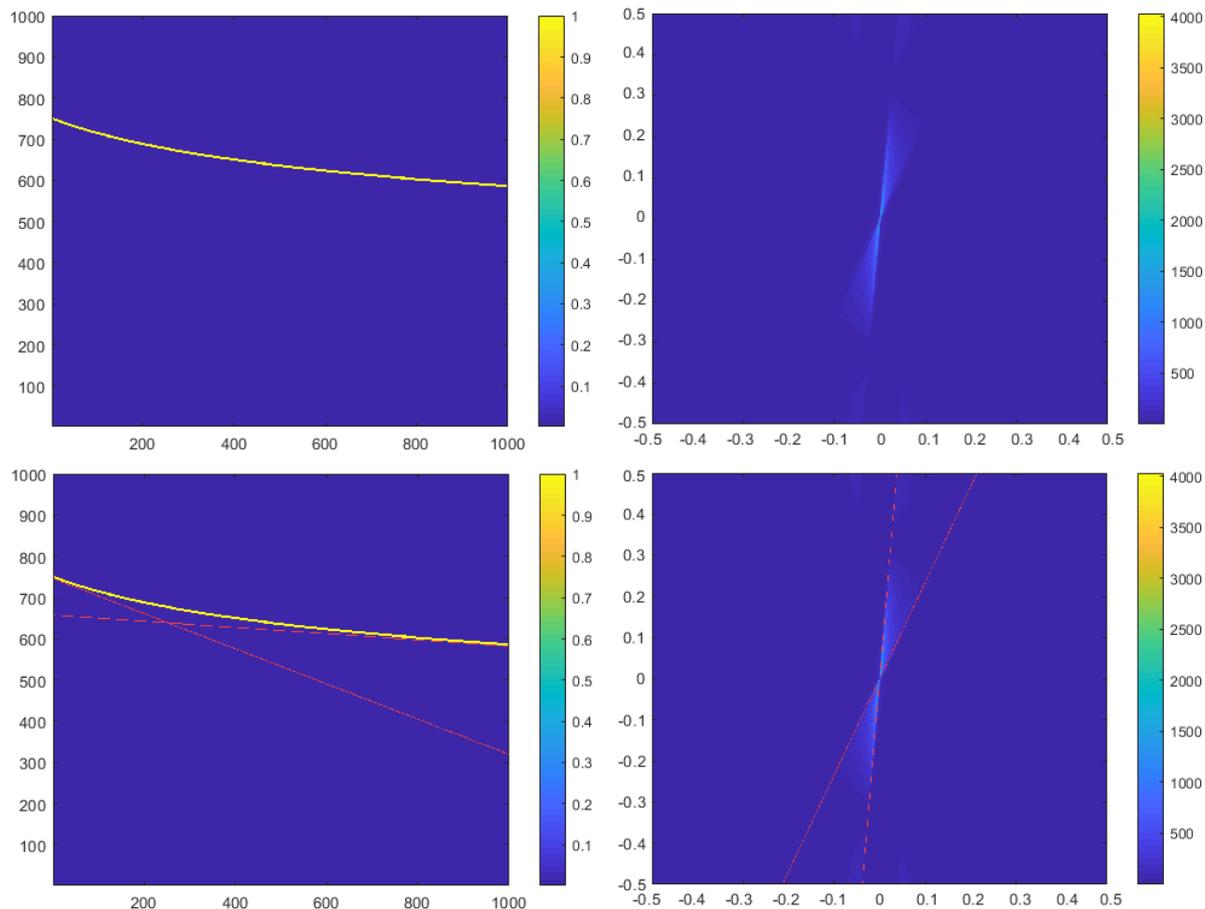


Figura 3.10: A sinistra: in alto, andamento in tempo-frequenza di un r-mode con frequenza iniziale $f_0 = 1500\text{Hz}$ con scala $f_{max} = 2000\text{Hz}$ e durata $t_{max} = 30\text{s}$, riportato in unità di pixel; in basso, la stessa curva con in sovrapposizione, in rosso, le tangenti agli estremi. A destra: in alto, modulo della FFT2 della curva, con il passo delle frequenze calcolato sui pixel e non sulle misure fisiche; in basso, lo stesso modulo con in sovrapposizione, in rosso, le rette corrispondenti alle trasformate delle tangenti.

In conclusione, si può affermare che il modulo della FFT2 di una curva regolare è rappresentata in prima istanza da un fascio proprio di rette modulato in ampiezza centrato sull'origine delle basse frequenze, i cui confini possono essere determinati come segue: detta $y = f(x)$ la rappresentazione, in termini di pixel, di una curva curva presente in una matrice $M \times N$, e detta $m(x)$ la generica inclinazione delle tangenti a tale curva in ogni suo punto, il fascio proprio di rette nello spazio di Fourieri sarà delimitato, in virtù della 3.21, da

$$m(u, v) \in \left[\frac{-1}{\max[m(x)]} \frac{N}{M}, \frac{-1}{\min[m(x)]} \frac{N}{M} \right] \quad (3.22)$$

Un altro genere di curva di grande interesse, soprattutto per la ricerca esposta nella Sezione 2.5, sono le

sinusoidi, in quanto rappresentano la modulazione che un segnale, anche a frequenza costante, subisce a causa dell'effetto Doppler. Applicando il ragionamento appena fatto, partiamo da una curva espressa nella forma $y = A \cdot \sin\left(\frac{x-x_0}{\lambda}\right)$; la derivata prima, punto per punto, è data da $y' = \frac{A}{\lambda} \cos\left(\frac{x-x_0}{\lambda}\right)$, dunque il modulo della FFT2 di tale curva consisterà in un fascio proprio di rette limitato dalle inclinazioni massime delle tangenti al grafico, raggiunte nei punti in cui il seno si annulla:

$$m(x, y) \in \frac{A}{\lambda} \cdot [-1, +1] \quad \rightarrow \quad m(u, v) \in \frac{\lambda N}{A M} \cdot [-1, +1] \quad (3.23)$$

Si evince che in base alla 3.23 il modulo della FFT2 di una senoide semplice dovrà essere, per ragioni di simmetria della funzione seno, simmetrico rispetto all'asse verticale. In Figura 3.11 sono mostrati i moduli delle trasformate di due sinusoidi, riportate in sovrapposizione in giallo (da considerarsi fuori scala di colore): la prima è stata costruita in modo da avere $A = \lambda$ e di conseguenza le linee delimitatrici dello spettro corrispondono alle bisettrici dei quadranti, mentre la seconda è stata costruita con la stessa ampiezza A ma una lunghezza $\lambda' = 2A$ in modo da raddoppiare, come in effetti avviene, l'inclinazione delle rette delimitatrici.

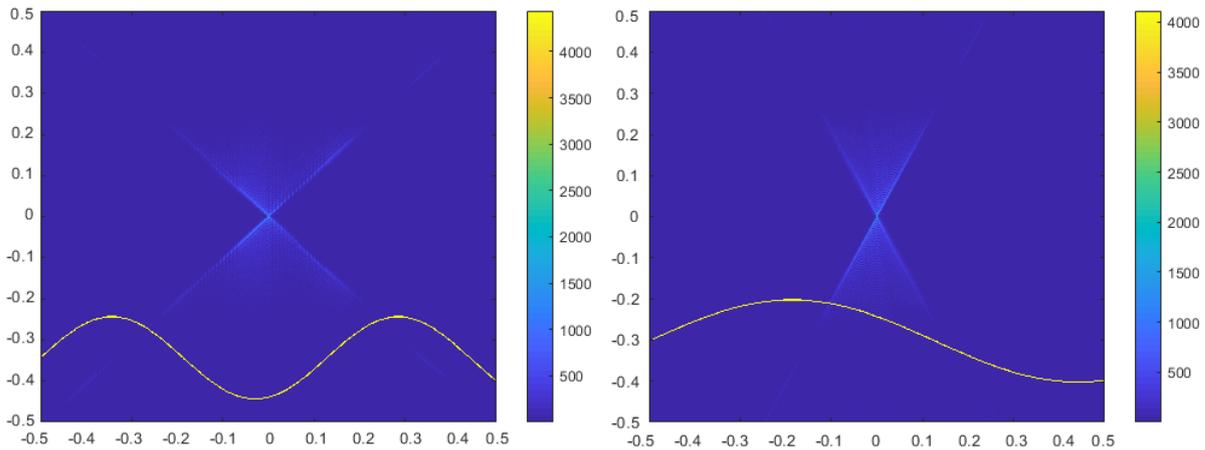


Figura 3.11: Due rappresentazioni, in mappe 1000×1000 , dei moduli delle FFT2 di due sinusoidi riportate in sovrapposizione in giallo, fuori dalla scala di colore del grafico. A sinistra: senoide di equazione $y = 100 \sin(x/100)$. A destra: senoide di equazione $y = 100 \sin(x/200)$.

Una generalizzazione interessante per le possibili applicazioni successive consiste nel ruotare la senoide di un certo angolo: questo localmente riproduce l'effetto di modulazione ulteriore del Doppler annuale su quello siderale. In base alle considerazioni già fatte, inclinando la senoide si avrà una forma data dalle equazioni:

$$y = mx + A \cdot \sin\left(\frac{x-x_0}{\lambda}\right) \quad \rightarrow \quad y' = m + \frac{A}{\lambda} \cos\left(\frac{x-x_0}{\lambda}\right)$$

di conseguenza, in virtù della 3.22, il range delle tangenti estremali si trasforma nel seguente modo:

$$m(x, y) \in \left[m - \frac{A}{\lambda}, m + \frac{A}{\lambda} \right] \quad \rightarrow \quad m(u, v) \in \frac{N}{M} \left[\frac{-1}{m + \frac{A}{\lambda}}, \frac{-1}{m - \frac{A}{\lambda}} \right] \quad (3.24)$$

Anche questa valutazione può essere verificata facilmente: in Figura 3.12 sono riportati i moduli delle trasformate di due sinusoidi, entrambe con $A = 50$ pixel e $\lambda = 50$ pixel, ruotate rispettivamente di $\pm 20^\circ$. Per facilitarne la comparazione, oltre alle sinusoidi di partenza (curve in giallo fuori scala), sono state riportate in sovrapposizione le tangenti delimitanti lo spettro calcolate attraverso la 3.24, in colore rosso:

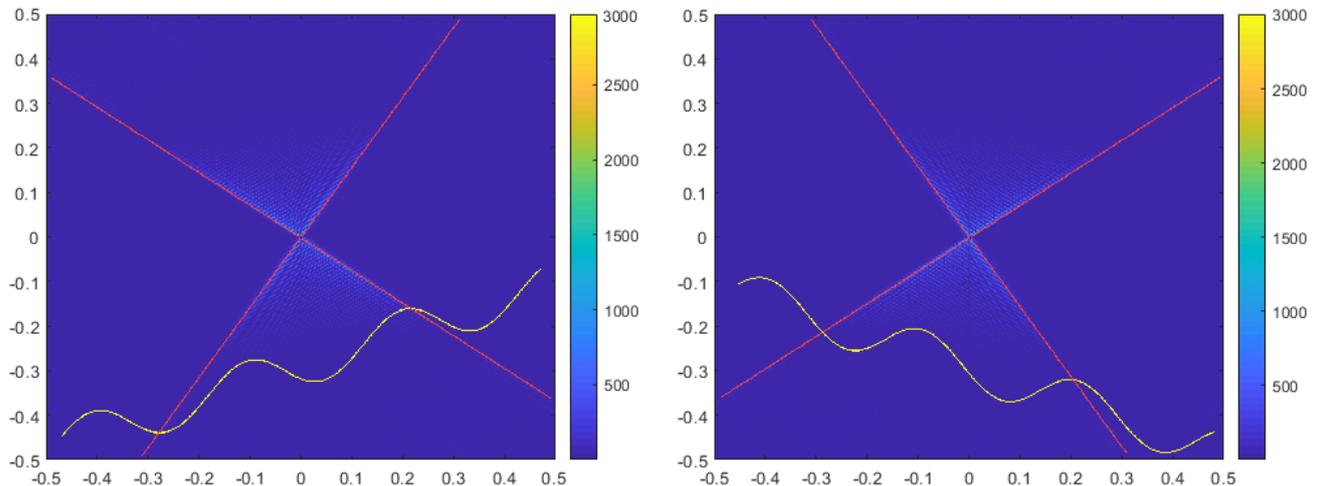


Figura 3.12: Due rappresentazioni, in mappe 1000×1000 , dei moduli delle FFT2 di due sinusoidi ruotate riportate in sovrapposizione in giallo, fuori dalla scala di colore del grafico. A sinistra: sinusoidi di equazione $y = (\tan 20^\circ)x + 50 \sin(x/50)$. A destra: sinusoidi di equazione $y = -(\tan 20^\circ)x + 50 \sin(x/50)$. In sovrapposizione in rosso, le rette delimitanti lo spettro calcolate attraverso la 3.24.

Si può dunque ritenere ragionevole la stima più generale data dalla 3.22, almeno restringendoci alle configurazioni di curve regolari e sufficientemente semplici di cui si occupano le ricerche descritte nelle Sezioni 2.5 e 2.6.

3.4 Realizzazione di filtri in 2D

Una volta chiariti i meccanismi principali che determinano il funzionamento della trasformata di Fourier in 2D di una matrice numerica, è possibile generalizzare i ragionamenti fatti finora a serie di dati in due dimensioni. Il funzionamento identico dei meccanismi che sono alla base della teoria del filtraggio, come ad esempio la convoluzione, la validità in 2D del teorema di convoluzione e l'uguaglianza di Parseval in 2D, consente infatti di trasferire l'intero impianto formale del filtraggio da oggetti ad una dimensione ad oggetti in una dimensione.

3.4.1 Filtri di Wiener in 2D

Quanto detto riguardo al filtro di Wiener su stringhe di dati può essere generalizzato per varietà in due dimensioni. Consideriamo quindi una matrice di dati in cui si suppone sia inserito un segnale all'interno del rumore:

$$d[j, k] = h[j, k] + n[j, k]$$

Per rimanere nel contesto delle ricerche in argomento, la matrice $h[j, k]$ rappresenta lo spettrogramma di un segnale puro, ovvero la sua rappresentazione in tempo-frequenza, mentre la matrice $n[j, k]$ rappresenta lo spettrogramma del rumore, che nel caso di rumore bianco, come visto nella Sezione 2.2, presenta distribuzione esponenziale. Allora il filtro di Wiener è rappresentato da una matrice $w[j, k]$ che, applicata in convoluzione ai dati, restituisce un output che si avvicina il più possibile al segnale d'ingresso:

$$\hat{h}[j, k] = w[j, k] * \{h[j, k] + n[j, k]\}$$

Come è mostrato nella Sezione 3.2.1, tale operazione non presenta alcuna difficoltà. In generale il filtro, nello spazio di Fourier, avrà espressione analoga alla 3.4:

$$W[u, v] = \frac{S_h[u, v]}{S_d[u, v]} = \frac{S_h[u, v]}{S_n[u, v] + S_h[u, v]} \quad (3.25)$$

Allo stesso modo, anche la procedura per applicare il filtro sarà la medesima:

$$\hat{h}[j, k] = \mathcal{F}_2^{-1} \{D[u, v] \cdot W[u, v]\} = \mathcal{F}_2^{-1} \left\{ D[u, v] \cdot \frac{S_h[u, v]}{S_d[u, v]} \right\} \quad (3.26)$$

Analogamente a quanto detto nella Sezione 3.1, nella 3.25 si utilizza la densità spettrale del rumore per normalizzare correttamente i dati in uscita, soprattutto se il rumore non è bianco; tuttavia noi abbiamo a disposizione solamente una stima di tale spettro deducibile dal campione limitato a disposizione. La scelta ottimale consiste allora nell'identificare il filtro di Wiener con il solo spettro del segnale che compare a numeratore, eventualmente dividendo per una stima teorica dello spettro del rumore se non si ha a che fare con rumore bianco, e rinormalizzare l'immagine a fine processo: questo a maggior ragione considerando il fatto che lo spazio di Fourier in 2D non ha un suo significato fisico. Per apprezzare la differenza tra i due approcci si può ricorrere ad un semplice esempio: in Figura 3.13 sono mostrati una circonferenza con un dato spessore con accanto il modulo della sua FFT2; questa, dovendo tenere conto delle tangenti all'immagine di partenza, le quali coprono tutto l'angolo giro, consiste in una serie di cerchi concentrici.

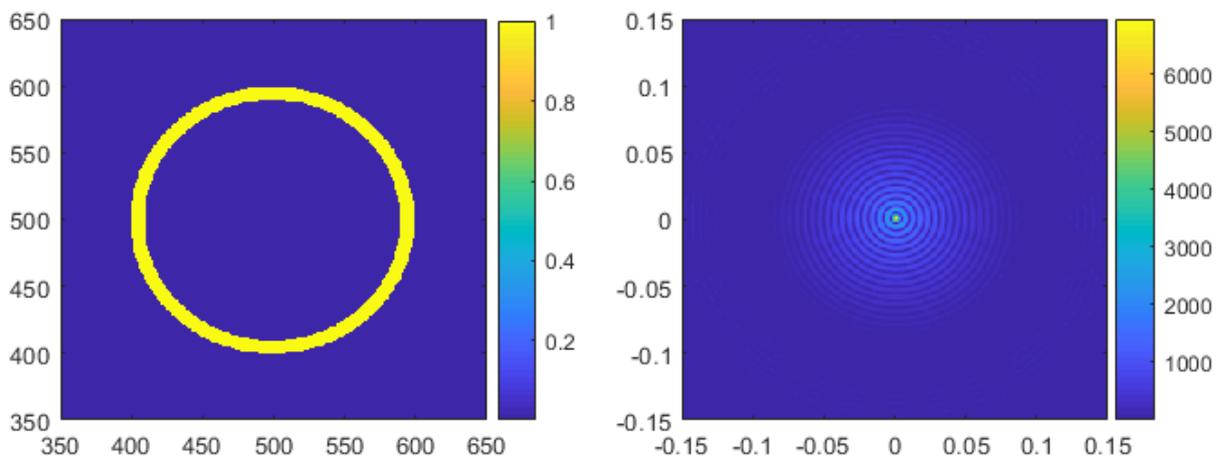


Figura 3.13: A sinistra: dettaglio ingrandito di una circonferenza di raggio 100 pixel e spessore 10 in una matrice 1000×1000 . A destra: dettaglio ingrandito del modulo della FFT2 della circonferenza.

Si consideri una immagine contenente un certo numero di tali circonferenza, come mostrato nel riquadro in alto a sinistra della Figura 3.14. A tale immagine, nel riquadro a fianco a destra, viene sommato del rumore con distribuzione esponenziale con ampiezza media pari a 5 volte quella delle circonferenze in modo da sovrastarle e renderle irriconoscibili. A partire dal modulo della FFT2 di una singola circonferenza dato in Figura 3.13 si costruisce il filtro di Wiener facendone il quadrato, ovvero calcolando lo spettro della circonferenza. In un caso, si divide tale spettro per lo spettro dell'immagine rumorosa in conformità con la 3.25; nell'altro, si lascia lo spettro così com'è e lo si utilizza come filtro. Applicando poi la procedura di filtraggio data dalla 3.26 si ottengono come risultato le due immagini nei riquadri in basso della Figura 3.14.

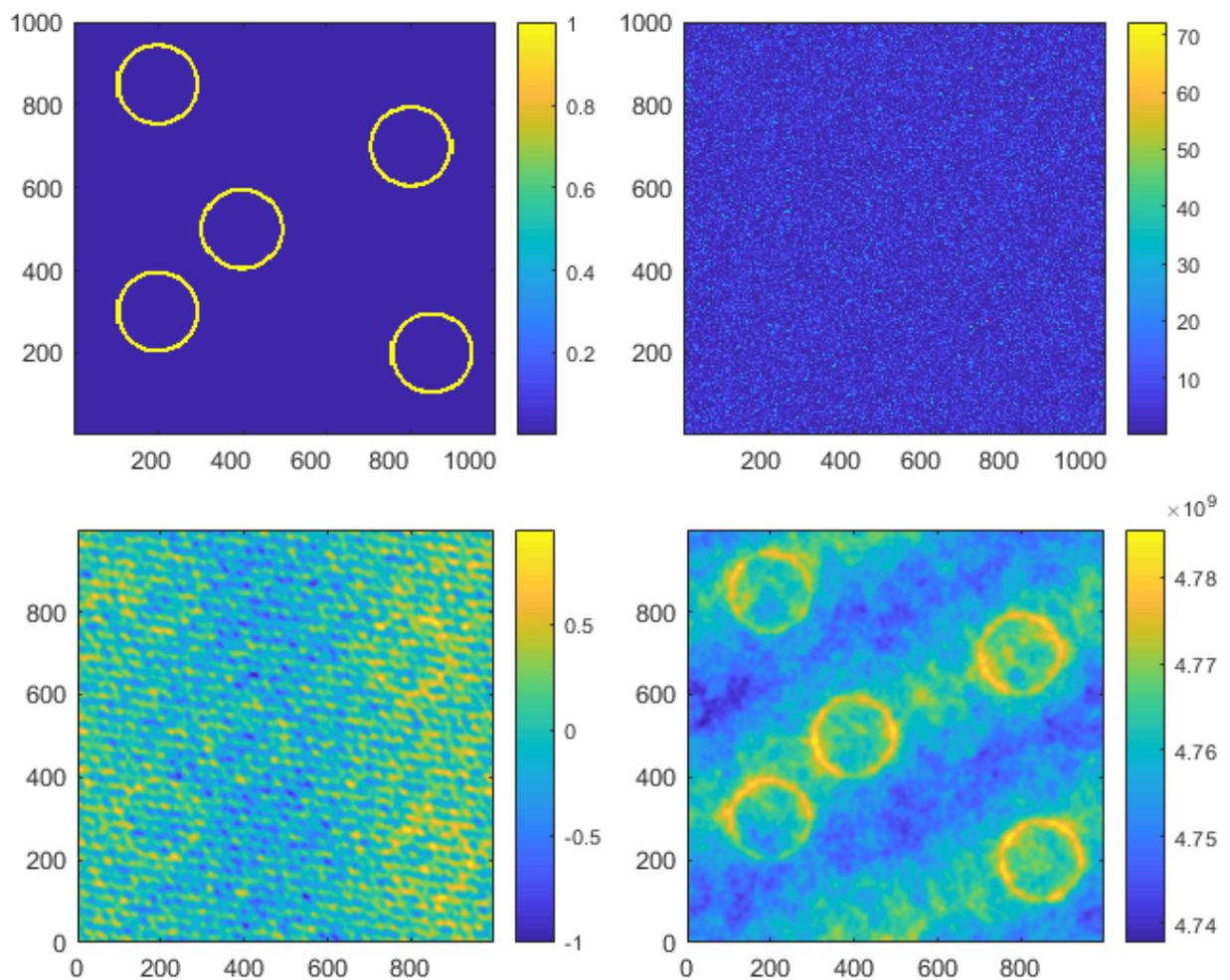


Figura 3.14: In alto a sinistra: circonferenze di raggio 100 pixel, spessore 10 pixel ed ampiezza unitaria in una immagine 1000×1000 . In alto a destra: la stessa immagine di sinistra a cui è stato sommato del rumore a distribuzione esponenziale con ampiezza media 5. In basso a sinistra: risultato dell'applicazione del filtro di Wiener utilizzando $W[u,v] = \frac{S_h[u,v]}{S_d[u,v]}$. In basso a destra: risultato dell'applicazione del filtro di Wiener utilizzando $W[u,v] = S_h[u,v]$; in quest'ultimo riquadro l'ampiezza dell'immagine risulta fuori scala.

Come si vede, l'immagine ottenuta dividendo il filtro per lo spettro rumoroso (in a sinistra) risulta molto più disturbata rispetto alla pulizia dell'immagine ottenuta senza effettuare la divisione (in basso a destra);

si può notare inoltre, guardando la scala delle ampiezze, che la prima immagine risulta correttamente normalizzata, come ci si attendeva, mentre la seconda, avendo saltato tale passaggio, è completamente fuori scala e dovrà essere normalizzata in un secondo momento. Dato che in questo contesto la priorità è far emergere una determinata forma da una immagine rumorosa con il maggior contrasto possibile nel seguito, quando si farà riferimento al filtro di Wiener, si farà riferimento sempre alla versione nella quale non si effettua la divisione per la stima dello spettro rumoroso.

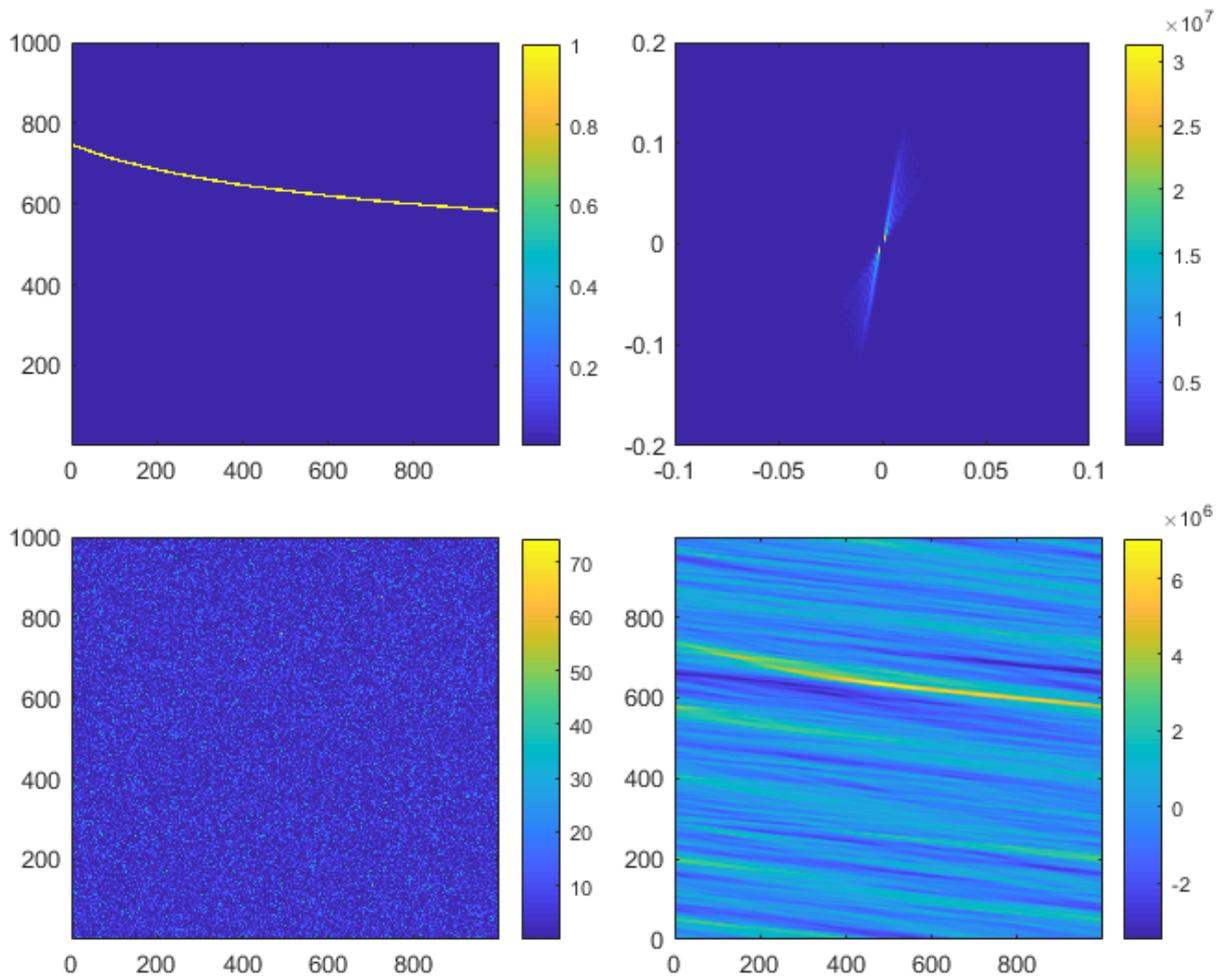


Figura 3.15: In alto a sinistra: andamento in tempo-frequenza di un r-mode con frequenza iniziale $f_0 = 1500Hz$ con scala $f_{max} = 2000Hz$, durata $t_{max} = 30s$ ed ampiezza 1, riportato in unità di pixel in una mappa 1000×1000 pixel. In alto a destra: ingrandimento della densità spettrale dell'r-mode. In basso a sinistra: allo stesso r-mode è stato sommato del rumore a distribuzione esponenziale con ampiezza media 5. In basso a destra: risultato dell'applicazione del filtro di Wiener all'immagine rumorosa.

Ritornando a forme più conformi ai segnali oggetto di ricerca, è utile vedere l'effetto del filtro di Wiener per la ricerca di un segnale dalla morfologia semplice come un r-mode. Nel riquadro in alto a sinistra della Figura 3.15 è mostrato l'andamento in tempo-frequenza di un r-mode riportato nella scala dei pixel e non delle unità fisiche; nel riquadro in alto a sinistra è mostrato un ingrandimento della densità spettrale calcolata per tale segnale, la quale verrà utilizzata per filtrare l'immagine rumorosa che si vede in basso a sinistra, nella quale è stato aggiunto del rumore esponenziale ad ampiezza media 5 volte maggiore

dell'ampiezza del segnale in modo da sovrastarlo del tutto. In basso a sinistra si vede l'immagine successivamente all'applicazione del filtro di Wiener: il segnale, prima completamente nascosto, è ritornato ben visibile. Per avere una misura quantitativa del miglioramento che c'è stato si può calcolare il rapporto segnale/rumore rispetto ad una maschera individuata dal segnale senza rumore aggiunto; in altre parole si misura la differenza tra l'ampiezza media dei pixel dell'immagine che cadono sulla maschera e quella dei pixel che cadono fuori dalla maschera in termini della deviazione standard delle ampiezze dei pixel, secondo la seguente formula, analoga per certi versi alla 2.4:

$$SNR = \frac{\text{mean}_{in}(f[j,k]) - \text{mean}_{out}(f[j,k])}{\text{std}_{out}(f[j,k])} \quad (3.27)$$

Applicando tale formula alla specifica realizzazione mostrata in Figura 3.15, si ottiene per l'immagine rumorosa prima dell'applicazione del filtro un valore di $SNR_0 = 0.1976$, mentre sull'immagine dopo l'applicazione del filtro si ottiene un $SNR_w = 4.5719$, con un miglioramento di un fattore 23. Tale valore è indicativo del fatto che il miglioramento del SNR che ci si aspetta dall'applicazione di un filtro di Wiener è in generale dell'ordine di grandezza delle decine, ma tali prestazioni devono venir studiate caso per caso per ogni differente segnale si cerchi. In generale più la morfologia di un segnale si complica e più peggioreranno le prestazioni del filtro di Wiener.

3.4.2 Filtro triangolare

Come si è visto, le potenzialità che ha l'applicazione del filtro di Wiener al fine di far emergere un determinato segnale dal rumore, ovvero aumentare il SNR , sono grandi. Tuttavia il limite più grande di tale metodo è proprio la mancanza di generalità: il filtro funziona in modo ottimale quando si ha a disposizione lo spettro di potenza del segnale che si cerca, ma nelle ricerche non coerenti che abbiamo visto non si ha la conoscenza del segnale cercato. Oltretutto, se si conoscesse effettivamente la forma di tale segnale, questa potrebbe essere utilizzata in una ricerca coerente, implementata ad esempio attraverso il filtro adattato, sicuramente più sensibile rispetto alle ricerche non coerenti. Diventa allora fondamentale costruire dei filtri più generali, controllabili, che possano avere degli effetti confrontabili con il filtro di Wiener, il quale rispetto ad essi rappresenterà l'efficienza ideale di riferimento. Tornando all'esempio in Figura 3.15, si possono fare alcune riflessioni sul rapporto che intercorre tra la particolare forma della densità spettrale del segnale, che rappresenta il filtro, ed il risultato del filtraggio. La densità spettrale, come si è visto nella Sezione 3.3, rappresenta un fascio di rette corrispondenti alle diverse inclinazioni assunte dalla curva secondo la 3.22: all'interno di tale regione i pixel assumono un valore elevato, mentre al di fuori presentano un valore prossimo allo zero. Di conseguenza, andando a moltiplicare tale spettro alla trasformata dell'immagine rumorosa l'effetto principale è di andare a deprimere le accoppiate di punti la cui inclinazione cada al di fuori della regione dello spettro, mettendo invece in evidenza quelli che vi cadono dentro. Guardando infatti all'immagine filtrata, si vede che non solo il segnale è messo in evidenza, ma il rumore al di fuori è stato cancellato lasciando solo le componenti allineate all'interno del range del segnale. Il suggerimento che se ne trae è che si possa realizzare un filtro che riproduca tale effetto senza dover tirare in ballo un dato segnale specifico: in effetti per fare ciò sarebbe sufficiente la conoscenza delle inclinazioni massima e minima di un eventuale segnale che si vuole rivelare; su questo principio sono stati sviluppati dei filtri che prendono il nome di filtri triangolari.

Un filtro triangolare consiste in una matrice delle stesse dimensioni dell'immagine che si desidera fil-

trare nella quale i pixel vengono valorizzati a 0 se risiedono al di fuori della regione determinata dalla condizione 3.22, mentre se vi cadono dentro vengono valorizzati ad 1. L'effetto di tale filtro sarà simile a quello del filtro di Wiener, con la differenza che mentre lo spettro del segnale presenta delle parti di ampiezza maggiore di altre, il filtro triangolare è piatto. La regola per la sua applicazione è la stessa del filtro di Wiener: data la matrice dell'immagine $d[j, k]$ e la sua trasformata $D[u, v]$, applicando il filtro triangolare $T[u, v]$ si avrà in uscita la matrice

$$\hat{h}[j, k] = \mathcal{F}_2^{-1} \{D[u, v] \cdot T[u, v]\}$$

Un altro parametro che si può controllare, oltre alle inclinazioni delle rette delimitanti, è il taglio in frequenza nella dimensione v : questo è un numero compreso tra 0 ed 1 che rappresenta la quota, in termini percentuali della dimensione del filtro, a cui si deve tagliare il filtro triangolare partendo dalle basse frequenze (origine) andando verso le alte frequenze (bordi). In figura 3.16 si può vedere un esempio di filtro triangolare, con il taglio in frequenza al 50%, posto a confronto con lo spettro di un r-mode dai cui estremi si sono stimate le inclinazioni con cui costruirlo: dal punto di vista della regione occupata, le due forme risultano quasi perfettamente sovrapponibili.

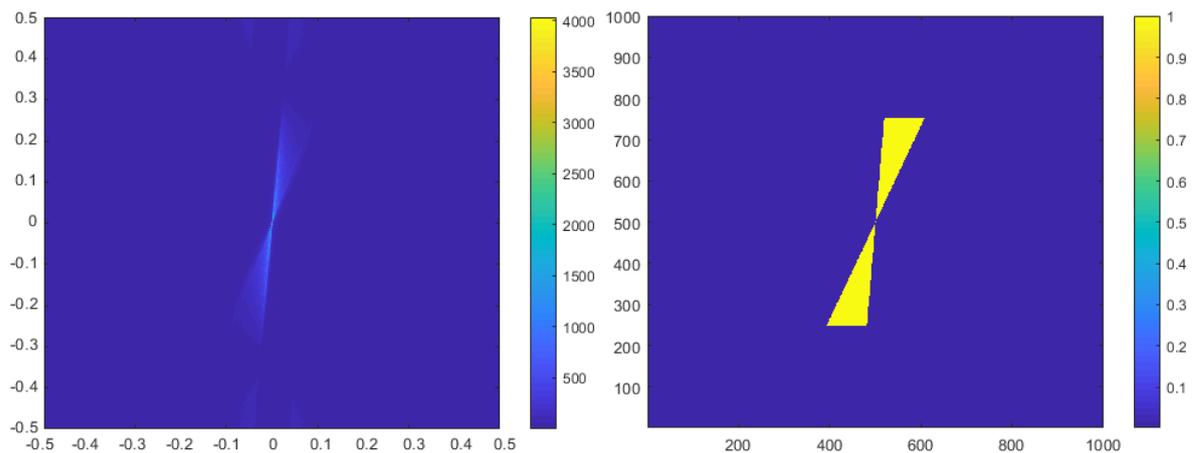


Figura 3.16: A sinistra: modulo della FFT2 di un tipico r-mode in tempo-frequenza. A destra: filtro triangolare con taglio in frequenza al 50% costruito con gli le inclinazioni degli estremi dell'r-mode.

Per rendersi conto delle similitudini e delle differenze tra il filtraggio con Wiener ed il filtraggio con un filtro triangolare è opportuno confrontare l'azione dei due filtri per due differenti forme ricorrenti e per le quali tale logica è applicabile: gli r-mode e le sinusoidi, che rappresentano l'effetto Doppler su di un segnale monocromatico. La scelta del taglio in frequenza, come si vedrà in seguito, incide in modo significativo sulla qualità dell'immagine in uscita dal filtro, in quanto l'inclusione di componenti ad alta frequenza può sporcare notevolmente il segnale. La determinazione del livello ottimale di taglio tuttavia non può essere fatta in astratto, ma va fatta relativamente ad ogni specifico segnale analizzando alcuni indicatori, come il SNR , al variare del taglio. In generale, dato che lo scopo del filtro triangolare è di simulare l'effetto del filtraggio attraverso lo spettro di potenza, tale livello dovrà corrispondere al livello di riempimento cui effettivamente arriva lo spettro di potenza del segnale. In Figura 3.17 è mostrato un esempio di applicazione del filtro triangolare ad un segnale r-mode. In alto a sinistra è riportato lo

stesso r-mode degli esempi precedenti, mentre in alto a destra è visibile il filtro triangolare generato utilizzando le inclinazioni del segnale ai suoi estremi e con una scelta del taglio in frequenza al 20%; al segnale è stato aggiunto del rumore esponenziale ad ampiezza media 5 volte maggiore, che in figura non è mostrato. In basso a sinistra è visibile il risultato dell'applicazione del filtro triangolare all'immagine rumorosa, con il risultato dell'applicazione del filtro di Wiener in basso a destra per confronto. Entrambe le immagini risultanti devono essere rinormalizzate nell'ampiezza. Il segnale in uscita dal filtro

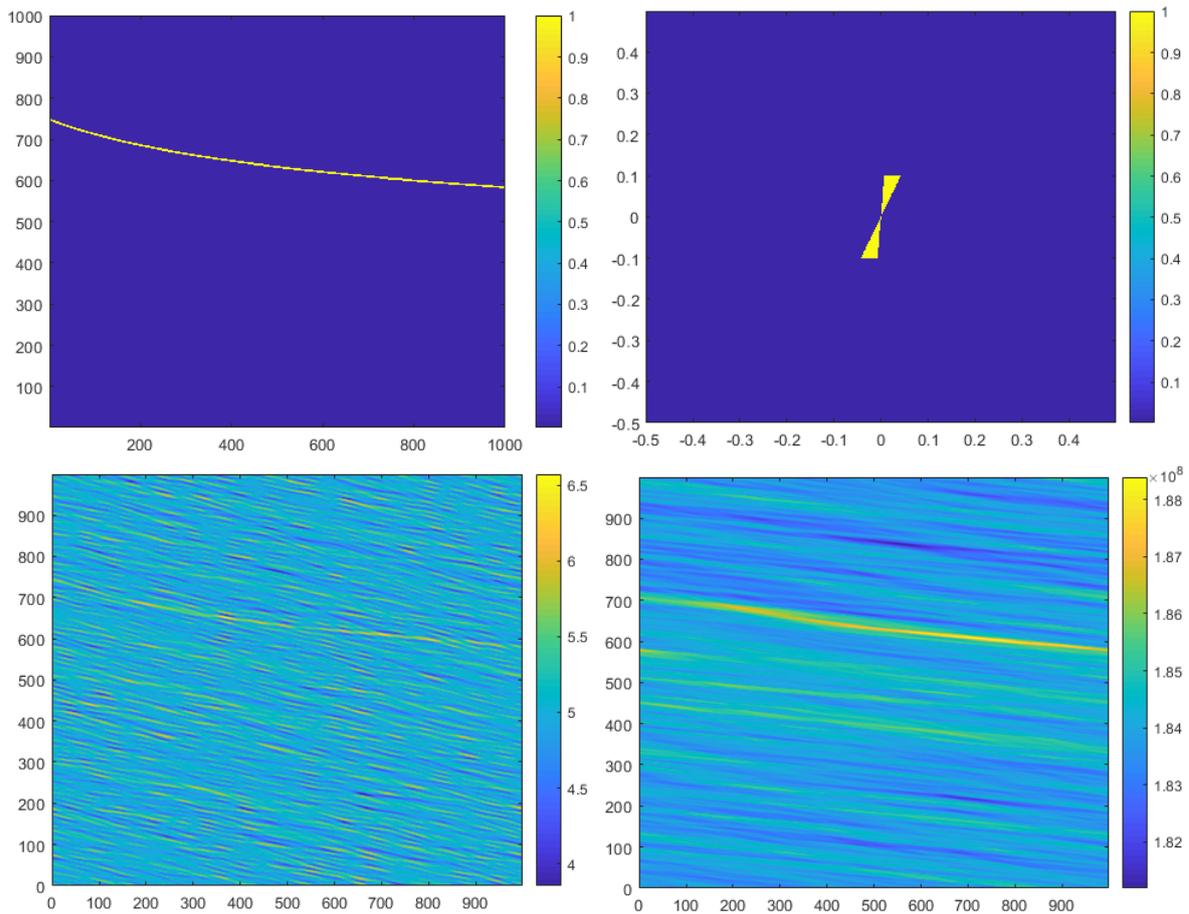


Figura 3.17: In alto a sinistra: andamento in tempo-frequenza di un r-mode con frequenza iniziale $f_0 = 1500Hz$ con scala $f_{max} = 2000Hz$, durata $t_{max} = 30s$ ed ampiezza 1, riportato in unità di pixel in una mappa 1000×1000 pixel. In alto a destra: filtro triangolare costruito sulle caratteristiche dell'r-mode, con taglio in frequenza al 20%. In basso a sinistra: risultato dell'applicazione del filtro triangolare all'immagine rumorosa. In basso a destra: risultato dell'applicazione del filtro di Wiener all'immagine rumorosa.

triangolare risulta sicuramente meno evidente rispetto a quello in uscita dal filtro di Wiener, ma rimane comunque apprezzabile nel suo decorso. Da un punto di vista quantitativo, partendo da una immagine rumorosa con un $SNR_0 = 0.1857$, l'immagine in uscita dal filtro triangolare presenta un $SNR_T = 2.0918$ con un miglioramento di un fattore 11, mentre l'immagine in uscita dal filtro di Wiener presenta un $SNR_w = 4.1131$ con un miglioramento di un fattore 22; il filtro di Wiener produce quindi un miglioramento della qualità dell'immagine doppia rispetto all'applicazione di questo specifico filtro triangolare, ma ogni diversa scelta del livello di taglio produrrà un miglioramento nel SNR differente. Ripetendo il procedimento per una forma sinusoidale con una leggera inclinazione, si deve applicare per il calcolo

delle inclinazioni massime e minime la regola espressa dalla 3.24. In Figura 3.18 è mostrato in alto a sinistra un modello di sinusoidi di periodo approssimativamente pari alle dimensioni della mappa con una inclinazione di 10° , mentre in alto a destra si può vedere il filtro triangolare costruito applicando la 3.24 e tagliando in frequenza al 20%. Anche in questo esempio non è mostrata, per sintesi, l'immagine contenente il segnale addizionato a rumore esponenziale ad ampiezza 5 volte maggiore. In basso a sinistra è mostrato il risultato dell'applicazione del filtro triangolare, mentre a destra è riportato il risultato dell'applicazione del filtro di Wiener a confronto. Anche in questo caso il segnale è tornato visibile

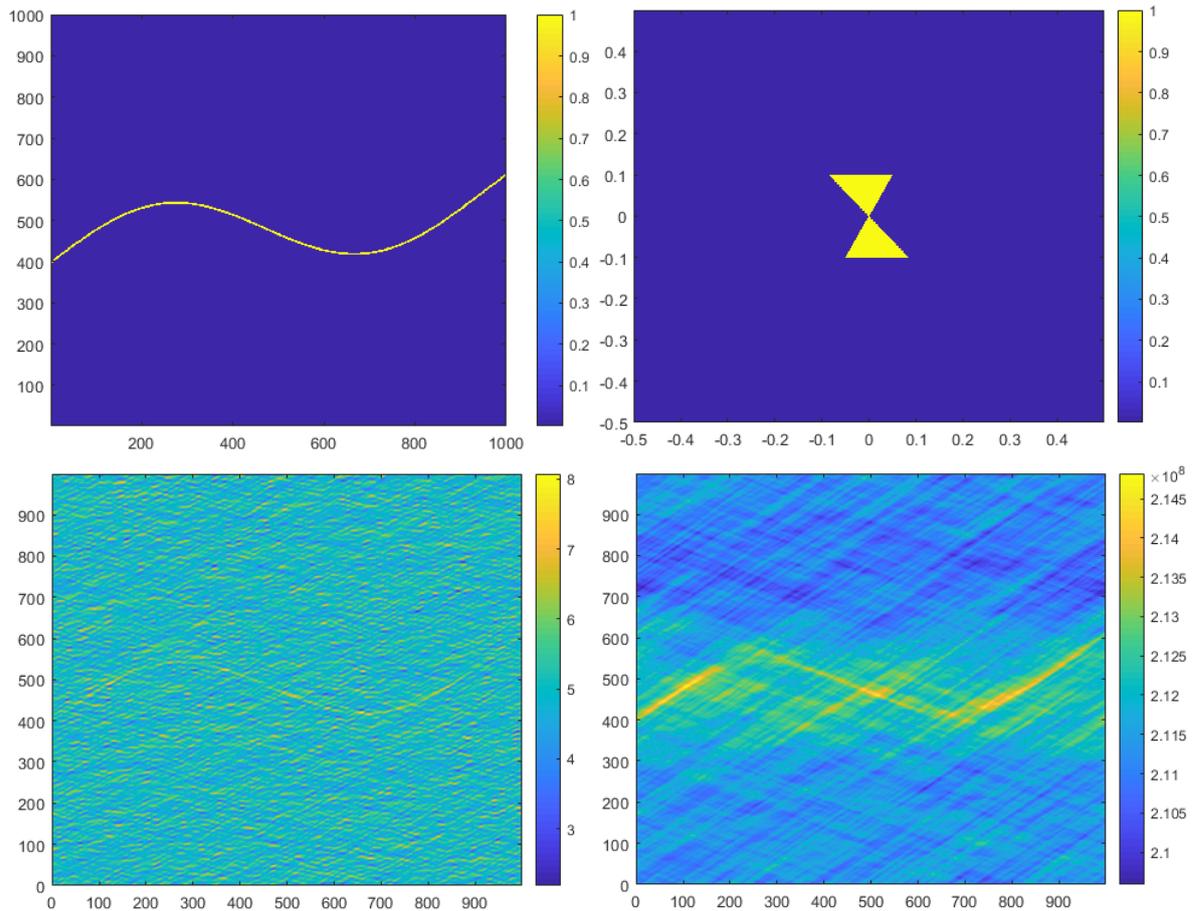


Figura 3.18: In alto a sinistra: andamento di una sinusoidi inclinata di equazione $y = (\tan 10^\circ)x + 100 \sin(x/150)$ ed ampiezza 1, riportata in unità di pixel in una mappa 1000×1000 pixel. In alto a destra: filtro triangolare costruito sui parametri della sinusoidi tramite la 3.24, con taglio in frequenza al 20%. In basso a sinistra: risultato dell'applicazione del filtro triangolare all'immagine rumorosa. In basso a destra: risultato dell'applicazione del filtro di Wiener all'immagine rumorosa.

a seguito dell'applicazione del filtro triangolare, anche se in maniera meno definita rispetto all'azione del filtro di Wiener. Andando a quantificare il miglioramento, se l'immagine con il rumore addizionato presentava un $SNR_0 = 0.2170$, a seguito dell'applicazione del filtro triangolare tale valore è passato a $SNR_T = 1.4772$ con un miglioramento di un fattore 7; l'immagine filtrata con Wiener invece presenta un $SNR_w = 3.2098$ con un miglioramento di un fattore 14. Anche in questo caso l'azione del filtro di Wiener ha efficacia doppia rispetto a quella del filtro triangolare, ma anche in questo caso lo spostamento del taglio in frequenza produrrà effetti differenti. In questo caso occorre anche verificare se l'efficacia

del filtro cambia al variare delle proporzioni e dell'inclinazione della sinusoide.

Si può ragionevolmente affermare che un filtro triangolare, in relazione ad un dato segnale regolare, può svolgere un compito analogo a quello del filtro di Wiener ma con una efficacia inferiore. A fronte di tale minor prestanza rimane comunque l'approccio più promettente al filtraggio in quanto in una ricerca reale di fatto il segnale non è a disposizione, così come non lo è il corrispondente filtro di Wiener. In alcuni casi si potrà inoltre comporre tale filtro con altri operatori per migliorare ulteriormente il risultato.

3.4.3 Finestratura gaussiana

Nella Sezione 3.2.2 si è accennato alla possibilità di utilizzo di una finestratura gaussiana per mettere in evidenza determinati segnali. Si ricorre a tale tecnica se il segnale che si ricerca è lentamente variabile in tempo-frequenza, o se è pressochè rettilineo con una determinata direzione. Effettuare la finestratura consiste nell'effettuare una convoluzione tra l'immagine ed una finestra gaussiana schiacciata lungo la direzione desiderata. Quali debbano essere le proporzioni ideali di tale finestra deve essere oggetto di studio, dato che presumibilmente questo dipenderà dalle caratteristiche e dalle proporzioni delle curva ricercata. Per vedere un effetto di tale operazione, nella Figura 3.19 è stata applicata una specifica finestratura gaussiana, con larghezze in unità di pixel $\sigma_t = 20$ e $\sigma_f = 4$, ad entrambe le immagini di r-mode e sinusoide date nelle Figure 3.17, 3.18 addizionate al rumore.

In entrambi i casi l'effetto della finestratura è di rendere in parte visibili i segnali nascosti: nel caso

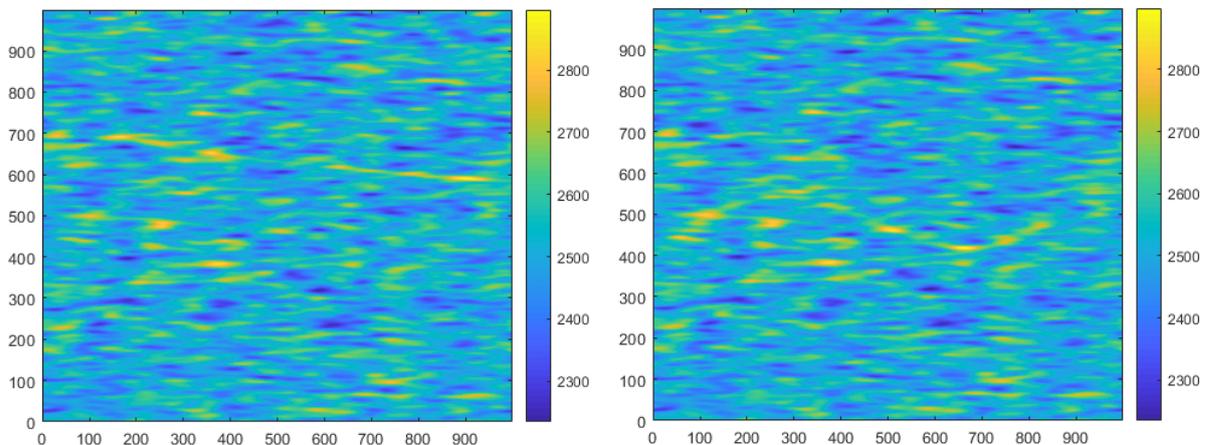


Figura 3.19: A sinistra: effetto dell'applicazione di una finestratura gaussiana con larghezze $\sigma_t = 20$ e $\sigma_f = 4$ in unità di pixel sul segnale r-mode di Figura 3.17 con addizionato il rumore. A destra: effetto dell'applicazione della stessa finestratura alla sinusoide di Figura 3.18.

dell'r-mode, l'immagine ottenuta ha un $SNR_G = 1.8291$ con un miglioramento di un fattore 10, mentre nel caso della sinusoide si ha un $SNR_G = 1.6488$ con un miglioramento di un fattore 8. Il risultato tuttavia è in generale poco soddisfacente, in quanto non permette di seguire adeguatamente le curve lungo il loro intero decorso; inoltre introduce un effetto di sgranatura che potrebbe avere come conseguenza una lieve deformazione del segnale ricostruito. Tale inconveniente potrebbe essere risolto affinando le dimensioni della finestra utilizzata per renderle più conformi alle proporzioni del segnale, ma non è detto che in generale la finestratura sia la strada corretta da intraprendere per qualunque tipo di segnale: ad esempio per le sinusoidi potrebbe rivelarsi inadeguata a causa del più ampio margine di variazione della curva stessa.

Un approccio di grande interesse pratico consiste nel combinare l'azione di una finestra gaussiana con quella di un filtro triangolare: la finestra infatti va ad amplificare gli andamenti nei dati conformi alla sua inclinazione, ma in qualunque direzione essi vadano; incrociando tale azione con quella di taglio delle inclinazioni indesiderate ad opera del filtro triangolare si potrebbero in alcuni casi ottenere dei miglioramenti complessivi. Nei due riquadri in Figura 3.20 sono mostrati gli effetti dell'applicazione concomitante della finestra gaussiana a $\sigma_r = 20$, $\sigma_f = 4$ e del filtro triangolare alle medesime immagini esemplificative precedenti. Si può apprezzare come in questo caso l'effetto sulle due immagini sia molto

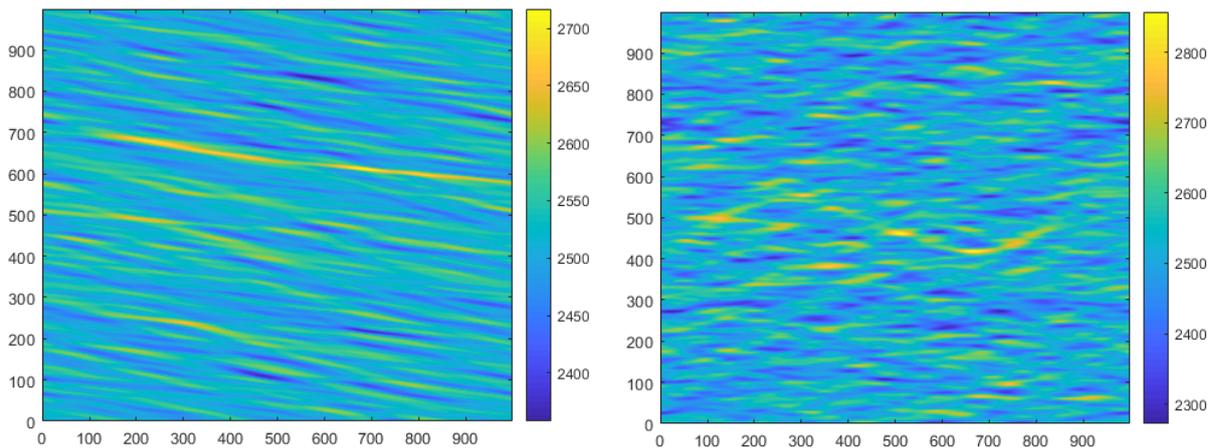


Figura 3.20: A sinistra: effetto dell'applicazione del proprio filtro triangolare e di una finestra gaussiana con larghezze $\sigma_r = 20$ e $\sigma_f = 4$ in unità di pixel sul segnale r-mode di Figura 3.17 con aggiunto il rumore. A destra: effetto dell'applicazione del proprio filtro triangolare e della stessa finestra alla sinusoide di Figura 3.18.

differente: mentre per l'r-mode si ha un netto miglioramento della qualità dell'immagine sia rispetto al solo filtro triangolare sia rispetto alla sola finestra gaussiana, per quello che riguarda la sinusoide l'effetto è praticamente indistinguibile da quello della sola finestra gaussiana. A conferma di tali impressioni qualitative si può notare, dal punto di vista quantitativo, che per l'r-mode si è ottenuto un $SNR_{T+G} = 3.2098$ con un guadagno di un fattore 16, ben superiore agli SNR ottenuti con le azioni separate dei filtri, mentre per la sinusoide si ha un $SNR_{T+G} = 1.7073$ con un guadagno di 8, praticamente lo stesso del caso della sola finestra gaussiana in azione.

Si può comprendere la motivazione che porta ad effetti tanto diversi considerando la configurazione che va ad assumere la composizione del filtro triangolare e della finestra gaussiana nei due differenti casi. Dato che nello spazio di Fourier la convoluzione tra due entità corrisponde al prodotto tra le relative trasformate, si può guardare all'azione successiva del filtro triangolare e della finestra gaussiana come all'azione di un singolo filtro che sia dato dal prodotto dei due. Confrontando il risultato di tale prodotto con lo spettro di potenza del segnale, che rappresenta il filtro ideale di Wiener, si può valutare quanto tale filtro combinato di avvicini o si discosti dall'ideale. Nell'immagine in alto a sinistra in Figura 3.21 si vede il filtro combinato triangolare gaussiano per r-mode che ha operato per ottenere l'immagine filtrata in Figura 3.20, mentre in alto a destra è mostrato il filtro di Wiener dell'r-mode. Il dato che spicca immediatamente all'occhio è che i due filtri occupano la stessa regione nello spazio di Fourier: l'aggiunta della finestra gaussiana ha l'effetto di simulare nel filtro triangolare l'andamento qualitativo che si ha nel filtro ideale di maggiore intensità attorno alle basse frequenze. Guardando invece alle immagini in basso a destra ed a sinistra, queste rappresentano rispettivamente il filtro combinato ed il filtro ideale della

sinusoide. Si vede chiaramente che le regioni dello spazio di Fourier interessate dai due filtri sono molto diverse: in particolare vi è nel filtro combinato una enfattizzazione delle regioni centrali che nel filtro ideale non vi è affatto.

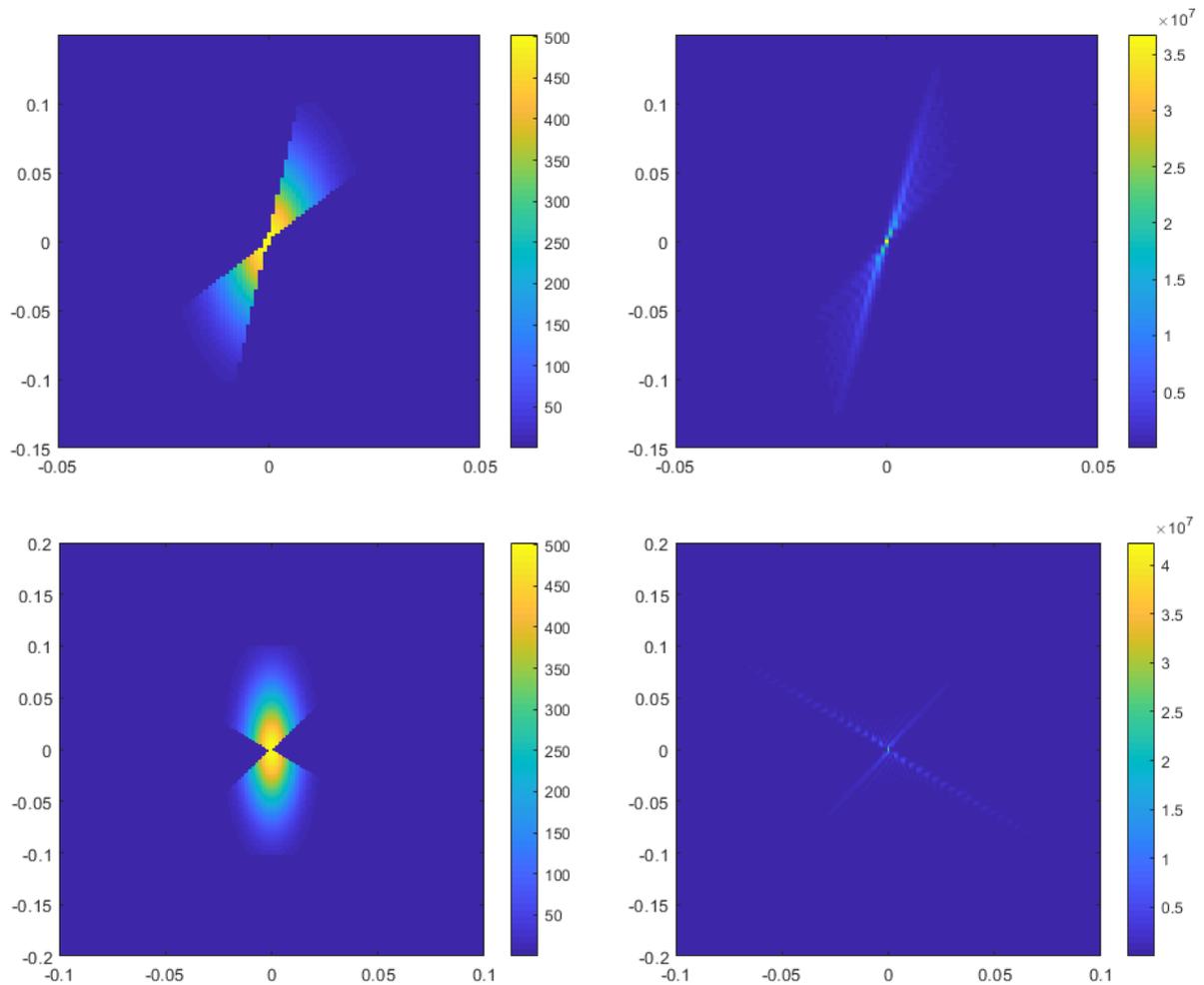


Figura 3.21: In alto a sinistra: filtro combinato triangolare gaussiano per l'r-mode utilizzato per ottenere l'immagine in Figura 3.20. In alto a destra: spettro di potenza del segnale di r-mode di cui in Figura 3.17, utilizzato come filtro di Wiener. In basso a sinistra: filtro combinato triangolare gaussiano per la sinusoide utilizzato per ottenere l'immagine in Figura 3.20. In basso a destra: spettro di potenza della sinusoide di cui in Figura 3.18.

Si può trarre dunque la seguente conclusione: mentre il filtro triangolare ha degli effetti generalmente positivi sulle immagini contenenti sia andamenti di r-mode in tempo-frequenza che curve sinusoidali, la combinazione del filtro triangolare con la finestrazione gaussiana ha degli effetti positivi solamente per la ricerca di r-mode, mentre non è assolutamente efficace per mettere in evidenza andamenti sinusoidali. Nel capitolo successivo sarà oggetto di parziale studio il suo effetto sugli r-mode.

3.4.4 Filtro periodico

Nelle sezioni precedenti si è potuto vedere come il filtro di Wiener ed il filtro triangolare risultino meno efficienti quando applicati a segnali sinusoidali. La situazione, relativamente al filtro triangolare, va ulteriormente peggiorando quando, anzichè descrivere un solo periodo di oscillazione, la sinusoida descrive un certo numero di periodi all'interno dell'immagine. Per capirne il motivo può essere utile vedere come la ripetizione di più periodi di una oscillazione vada ad influenzare il corrispondente modulo della FFT2. Nei riquadri in Figura 3.22 sono riportati i moduli delle trasformate di differenti sinusoidi, le quali sono riportate in sovrapposizione e fuori scala, costruite con ampiezza e lunghezza in modo da avere le tangenti oscillanti tra le stesse inclinazioni in base alla 3.23. Il numero di periodi di oscillazione presenti nelle differenti immagini è, rispettivamente, 1, 4, 8, 16. Si vede che, mentre la forma generale rimane la stessa, al crescere del numero di periodi si inseriscono nella trasformata delle righe di vuoto in numero sempre maggiore.

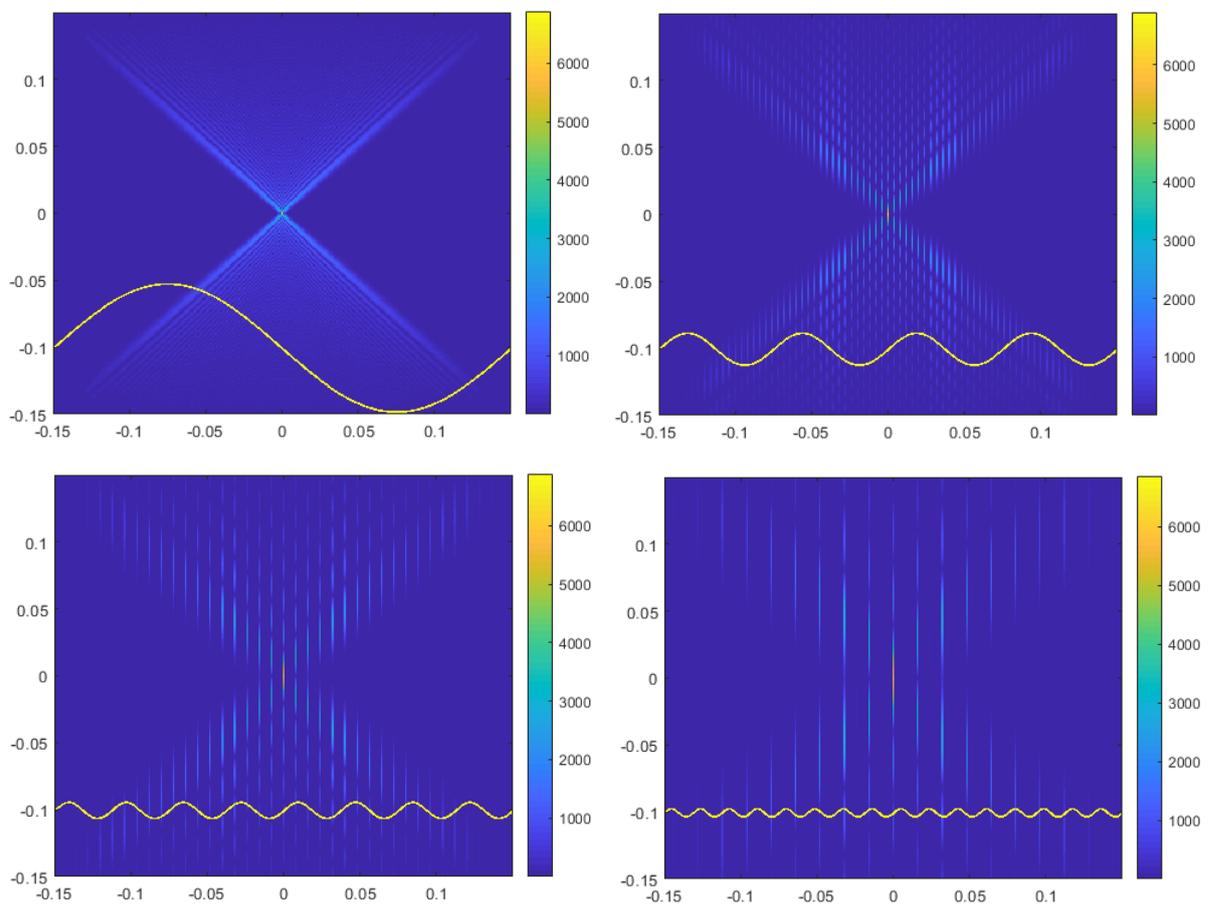


Figura 3.22: Modulo delle FFT2 di differenti sinusoidi in mappe 1000×1000 pixel, riportate in sovrapposizione (curve gialle) fuori dalla scala delle mappe. Le curve, partendo dalla posizione in alto a sinistra, hanno equazione: $y = \frac{1000}{2\pi} \sin(2\pi \frac{x}{1000})$; $y = \frac{250}{2\pi} \sin(2\pi \frac{x}{250})$; $y = \frac{125}{2\pi} \sin(2\pi \frac{x}{125})$; $y = \frac{125}{4\pi} \sin(4\pi \frac{x}{125})$.

Tale comportamento risulta perfettamente spiegabile se si guarda alle oscillazioni aggiuntive in un'ottica diversa, ovvero considerando la sinusoida presente nell'immagine come il risultato di un prodotto di

convoluzione tra un singolo periodo di oscillazione ed un insieme di delta discrete distanti tra loro di una quantità pari alla lunghezza dell'onda. Tale schematizzazione è riassunta in Figura 3.23. Se tale ipotesi è effettivamente valida ne segue, in base al teorema di convoluzione, che è possibile costruire la FFT2 della sinusoide completa nell'immagine mediante la semplice moltiplicazione delle trasformate della singola sinusoide e dell'insieme di punti. Nella Sezione 3.3 si è visto che la trasformata di un insieme di punti che nell'immagine riempiono un reticolo periodico in orizzontale è data, nel modulo, da una serie di linee verticali con una spaziatura in pixel che è data dal numero di punti. Alla luce di questo risulta completamente spiegata la peculiare configurazione che assume la trasformata di una sinusoide che descrive più periodi: della semplice conformazione triangolare che caratterizza la trasformata di una singola sinusoide sopravvivono solamente le colonne corrispondenti alle colonne di massimo della trasformata delle delta.

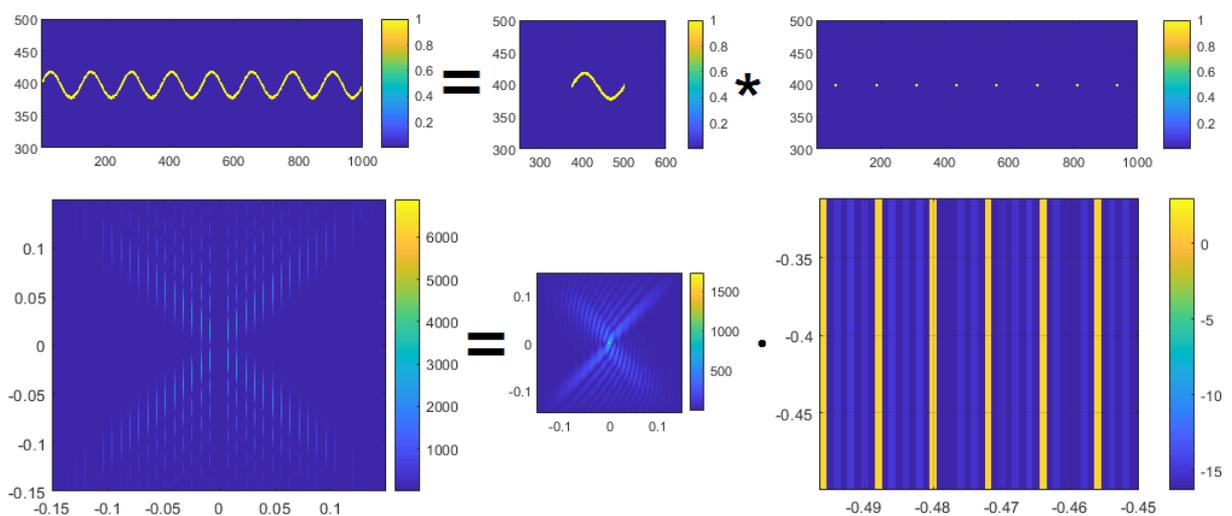


Figura 3.23: Sequenza in alto: una sinusoide che descrive 8 ripetizioni del periodo in una immagine viene schematizzata come prodotto di convoluzione (*) tra una singola oscillazione completa ed un reticolo riempito di 8 delta discrete. Sequenza in basso: I moduli delle FFT2 delle immagini corrispondenti nella sequenza in alto; nella prima e nell'ultima immagine le linee verticali hanno una separazione di 8 pixel; nell'ultimo grafico le ampiezze sono riportate in scala logaritmica.

Il funzionamento di tale meccanismo suggerisce che per filtrare in una immagine una determinata forma che si ripete periodicamente la scelta più efficace possa essere, anziché filtrare sulla forma con il filtro triangolare, filtrare sulla periodicità. In altre parole, in base al numero di periodi con cui si ripete la forma si applica un filtro che corrisponde alla trasformata delle funzioni delta in reticolo periodico: esso consisterà in un insieme di linee verticali valorizzate ad 1 che si ripetono con una spaziatura in pixel che corrisponde alla periodicità dell'immagine, con il resto della matrice valorizzato a 0. La potenza di questa tecnica risiede nel fatto che essa agisce allo stesso modo qualunque sia la forma di base, indipendentemente da essa: è sufficiente che questa venga ripetuta periodicamente. Per quanto riguarda l'efficacia del filtro, questa aumenta all'aumentare del numero di periodicità: se infatti in una immagine un segnale è effettivamente presente con una data periodicità, la sua informazione risiederà solamente nelle colonne della matrice trasformata che, partendo dall'origine, si susseguiranno ad una data distanza di pixel; conseguentemente nelle colonne della matrice non interessate dal segnale risiederà solamente informazione di rumore. Ne segue che applicando un filtro per un segnale che si ripete, ad esempio, per

due periodicità, assumendo il rumore distribuito su tutta la matrice si sta in pratica scartando il 50% di esso; filtrando un segnale che presenta 10 periodicità si scarta il 90% del rumore, mentre filtrandone uno con 100 periodicità si scarta il 99% del rumore.

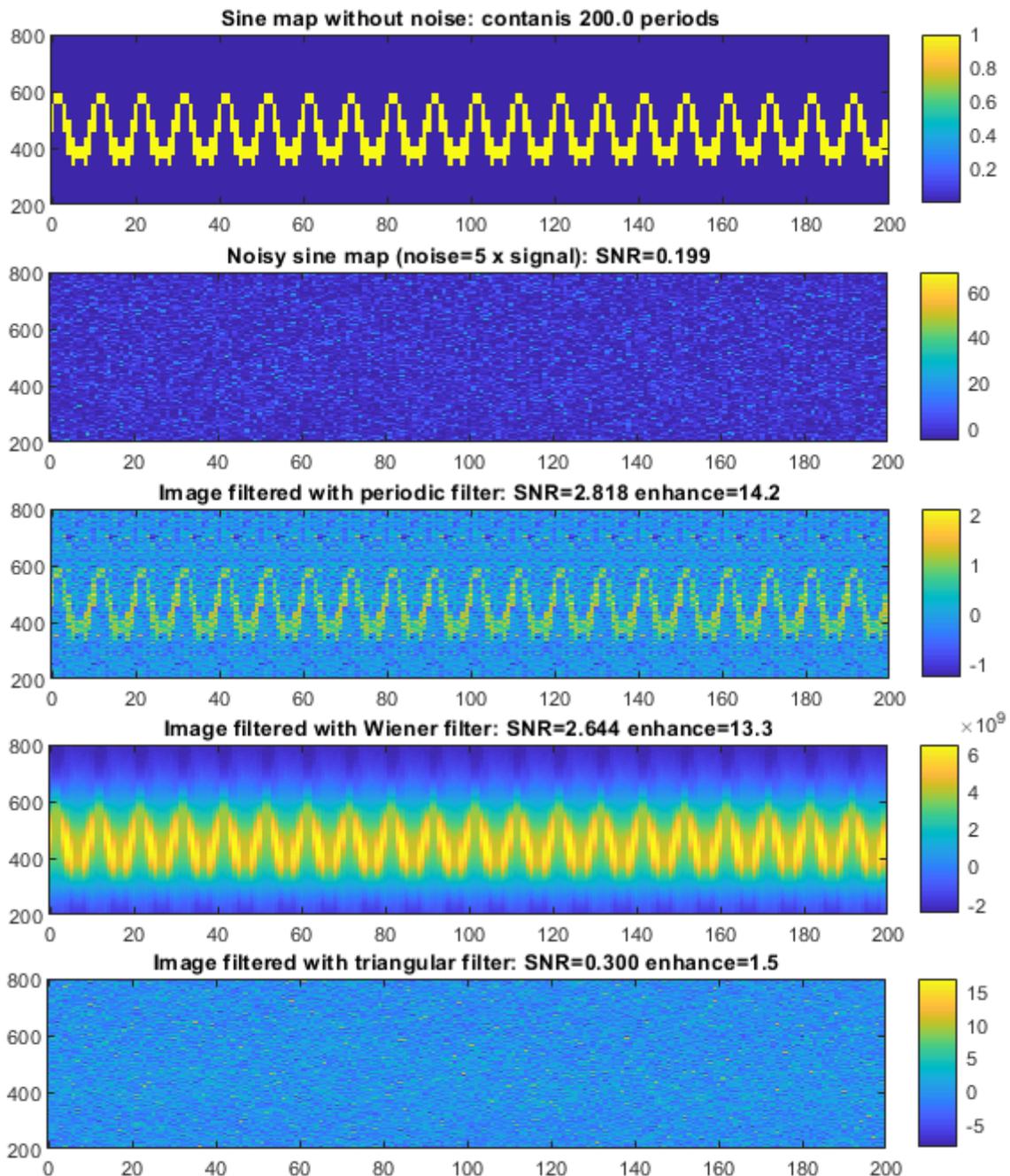


Figura 3.24: Prime due immagini in alto: ingrandimento di una sinusoida di ampiezza 1 che compie 200 cicli all'interno di una immagine 2000×1000 pixel, cui viene successivamente aggiunto rumore esponenziale di ampiezza media 5. Al centro: risultato dell'applicazione del filtro periodico. Penultima in basso: risultato dell'applicazione del filtro di Wiener. Ultima in basso: risultato dell'applicazione del filtro triangolare.

In Figura 3.24 è riportato l'esempio di una sinusoida ripetuta per 200 cicli all'interno di una immagine 2000×1000 pixel. Ad essa è aggiunto il rumore e vengono applicati il filtro periodico a 200 perio-

dicità, il filtro di Wiener ed il filtro triangolare per confrontarne le prestazioni. Si vede immediatamente che in questo contesto il filtro triangolare risulta del tutto inadeguato, producendo una immagine dal rapporto segnale/rumore cambiato di poco rispetto a quello di partenza. L'effetto del filtro periodico è invece ottimale, con un $SNR_p = 2.818$ migliorato di un fattore 14 addirittura migliore rispetto all'azione del filtro di Wiener, il quale produce una immagine con un $SNR_w = 2.644$ migliorato di un fattore 13 ed un segnale ricostruito leggermente distorto rispetto all'originale.

Si capisce bene che un limite del filtro periodico è che la periodicità cercata nel segnale deve essere proprio quella del filtro, altrimenti le linee in cui si distribuisce il segnale nella trasformata non coincidono con quelle selezionate dal filtro: l'efficienza di tale filtro decade allora rapidamente non appena le due periodicità non coincidono di poco. Ripetendo l'applicazione dello stesso filtro per 200 periodicità ad una mappa contenente stavolta 199 cicli di oscillazione si può vedere, in Figura 3.25, come il segnale venga perso quasi del tutto.

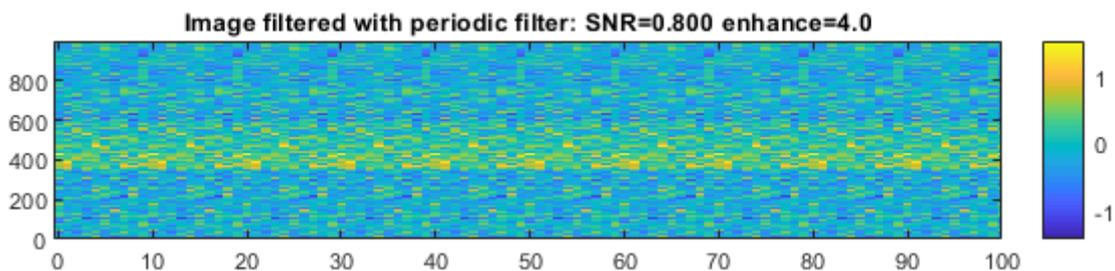


Figura 3.25: Risultato dell'applicazione di un filtro a 200 periodicità ad una immagine 2000×1000 pixel contenente un segnale sinusoidale con 199 cicli. Il filtro fa emergere un minimo eccesso di potenza nella fascia che contiene il segnale, ma non ne permette in alcun modo la ricostruzione.

A fronte di tale limite si può, in alcuni casi in cui l'informazione sia nota, costruire l'immagine in modo da farle contenere esattamente un dato numero di ripetizioni. Ad esempio, nella ricerca dei segnali continui, il segnale monocromatico subisce una modulazione siderale la cui periodicità può essere calcolata con grande precisione: in quel caso si può impostare un tempo di osservazione ad hoc che sia un multiplo intero del giorno siderale. Una successiva evoluzione nel filtro, ad ora piuttosto rozzo, potrebbe includere una leggera inclinazione dei pattern, a simulare l'effetto di un lento spin-down.

3.4.5 Il limite dei filtri

I filtri in 2D proposti nelle precedenti sezioni hanno il punto di forza nell'essere estremamente selettivi: se un segnale è presente, anche sovrastato da un livello di rumore dall'ampiezza ben al di sopra della sua, i filtri 2D sono in grado di metterlo in evidenza. Quale sia il livello minimo di SNR che l'immagine iniziale deve avere perché il segnale emerga con il filtro deve essere determinato attraverso studi specifici sui diversi segnali. Sia nel caso del filtro triangolare che del filtro periodico la cosa che deve saltare all'occhio è la grande mole di dati al di fuori della regione interessata dal filtro che viene buttata via. Questo punto di forza porta con sé anche dei limiti: è evidente che così facendo si va ad introdurre un bias alquanto pesante, in quanto nei dati buttati potrebbero essere presenti delle componenti di segnale inattese, soprattutto nelle ricerche sulle quali si è focalizzata l'attenzione, nelle quali non si conosce con certezza quale segnale sia nascosto. E' il problema dello shaping: se si guarda a tutti gli esempi fatti in queste sezioni, è vero che il segnale viene messo in risalto, ma è anche vero che pure tutto il rumore

residuo è modellato secondo le prescrizioni del filtro. In altre parole, con l'applicazione di tali filtri si riesce a vedere solo ciò che si cerca di vedere, se c'è.

Dal momento che l'applicazione dei filtri modifica pesantemente i dati originali, non si può pensare di poter fare la rivelazione vera e propria di un segnale attraverso di essi: per effettuare le verifiche ed i follow-up previsti alla fine delle procedure gerarchiche è necessario che i dati siano integri e non deformati.

Si può tuttavia ricorrere ai filtri nell'ambito della selezione dei candidati di primo livello: creando una seconda versione degli spettrogrammi a cui applicare i filtri, si può effettuare su di essi la selezione dei candidati di primo livello. Tali candidati porteranno con sé una stima, seppur rozza, dei propri parametri: questa potrà essere utilizzata nei livelli successivi della procedura gerarchica, quando si condurrà l'analisi più rifinita sui dati originali.

Capitolo 4

Applicazione dei filtri alla ricerca di segnali da r-mode

In questo capitolo si approfondirà il funzionamento dei filtri in 2D nel caso specifico di ricerca di segnali da r-mode. L'obiettivo generale, del quale il presente elaborato rappresenta solamente il punto di partenza, è di caratterizzare sufficientemente la risposta del segnale e del rumore ai filtri, nell'ottica di integrare tale tecnica nel primo livello della procedura gerarchica al fine di aumentare la sensibilità di ricerca.

Il primo passo consiste nello studio della risposta al filtro triangolare ed al filtro combinato triangolare gaussiano, in modo da individuarne la configurazione ottimale. L'indicatore tramite il quale si quantificherà la qualità del risultato sarà il *SNR* già utilizzato nel capitolo precedente, mentre il termine di paragone saranno i risultati del filtro ideale di Wiener.

Il passo successivo, in questo lavoro appena accennato, consiste poi nello studiare come l'applicazione dei filtri sullo spettrogramma di un segnale da r-mode influisca sulla selezione dei picchi nelle peakmap e di quali effetti questo produca sulla selezione dei candidati attraverso la trasformata di Hough generalizzata. L'effetto di shaping del rumore che inevitabilmente si ha con l'applicazione dei filtri potrebbe in linea di principio creare un certo numero di artefatti, ovvero di falsi candidati, dato che il filtro seleziona solamente i dati con l'andamento che si cerca di rilevare, ma tali candidati ci si aspetta vengano scartati nelle fasi successive della ricerca. L'obiettivo primario infatti, al netto degli artefatti prodotti, consiste nel verificare che la soglia minima che un segnale deve superare per essere selezionato nella peakmap venga spostata indietro grazie all'azione dei filtri. In questa ottica l'eventuale appesantimento provocato da artefatti di rumore ipoteticamente prodotti dai filtri passa in secondo piano rispetto alla possibilità di rintracciare un segnale che prima non si vedeva: tale problema, se sussiste, andrà approfondito attraverso uno studio della soglia ottimale con cui selezionare i picchi nelle peakmap od i candidati più significativi nelle mappe di Hough.

4.1 Studio della risposta ai filtri triangolari tramite il SNR

Nella Sezione 3.4.2 si è sottolineato che la risposta di una immagine ad un filtro triangolare può essere molto differente al variare del taglio in frequenza che si sceglie. Altri parametri del che si possono variare sono le inclinazioni limite: si è detto infatti che il vantaggio del filtro triangolare è di poter essere applicato facendo assunzioni più generali sul segnale; si vorrebbe vedere allora come varia la qualità del filtraggio impostando dei range più larghi, che possano includere segnali con diversi parametri.

Se si individuasse una regione di tolleranza abbastanza larga all'interno della quale l'efficacia del filtro non cambi in modo importante, si potrebbe conseguentemente individuare un insieme ristretto di filtri triangolari in grado di coprire la regione dello spazio dei parametri fisicamente interessante. Nell'effettuare le simulazioni bisogna considerare che la singola realizzazione di rumore che si utilizza potrebbe contenere particolari fluttuazioni tali da produrre, attraverso il filtraggio, degli artefatti che potrebbero sporcare l'effetto di un filtro. Dato che siamo interessati a conoscere il comportamento atteso dei filtri in situazioni medie, nelle simulazioni che seguiranno si farà riferimento al valor medio di un dato numero di stime del SNR ottenute su differenti realizzazioni del rumore con le stesse caratteristiche statistiche medie. Si può verificare facilmente in Figura 4.1 che diverse realizzazioni di rumore danno una risposta ai filtri differente ma che si comporta statisticamente bene, ovvero seguono una distribuzione a campana al crescere del numero di ripetizioni (nell'esempio in figura sono state ripetute 1000 realizzazioni).

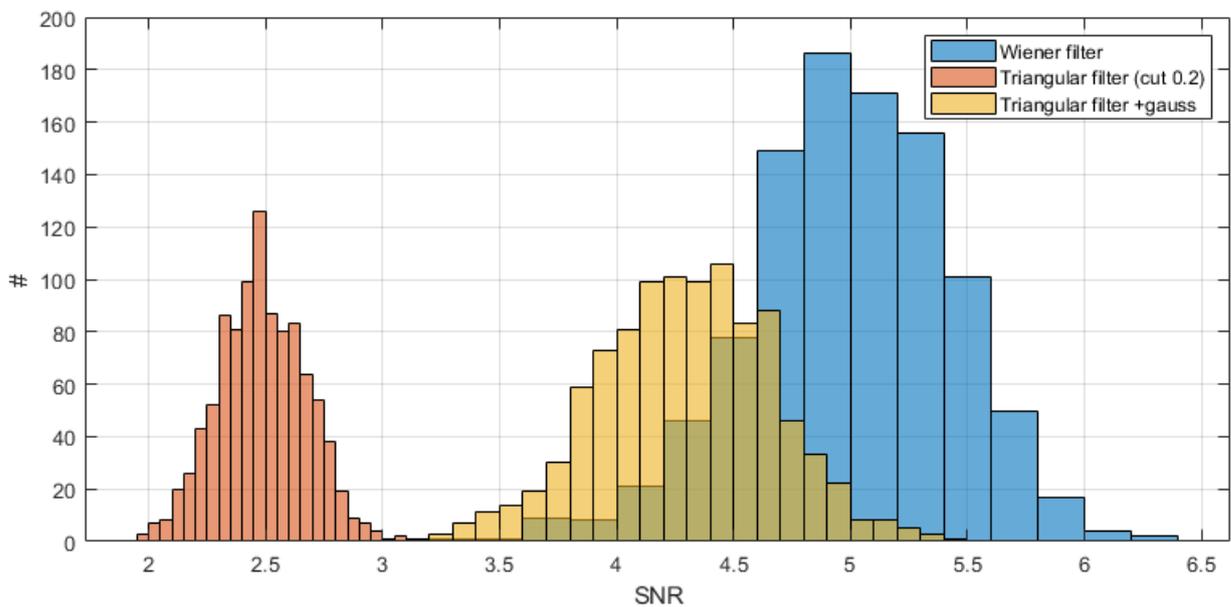


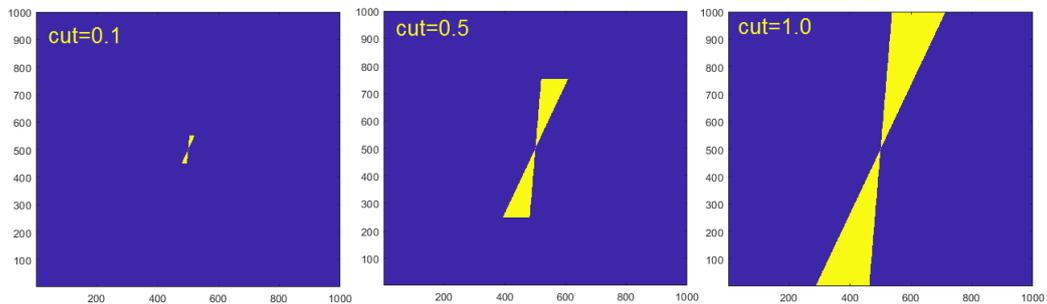
Figura 4.1: Istogrammi di 3 campioni di SNR ottenuti applicando rispettivamente un filtro di Wiener, un filtro triangolare con taglio in frequenza al 20% ed un filtro combinato triangolare gaussiano ad una immagine contenente un segnale di r-mode di ampiezza 1 addizionato a del rumore a distribuzione esponenziale con ampiezza media 5 di cui sono state prodotte 1000 realizzazioni differenti.

Nei diversi studi che seguiranno si riporteranno allora sempre dei valori medi del SNR ottenuti da un determinato numero di realizzazioni, in modo da non trarre indicazioni erronee a causa di fluttuazioni presenti in una data realizzazione.

4.1.1 Studio della risposta in funzione del taglio in frequenza

Il primo parametro di cui si vuole studiare il funzionamento nel filtro triangolare è il taglio in frequenza. Questo è l'unico parametro sul quale a priori non ci sono indizi di tipo fisico: mentre per quanto riguarda le inclinazioni si può ipotizzare che la scelta ottimale siano esattamente quelle corrispondenti al massimo ed al minimo del segnale (e sarà verificato), sul livello del taglio in frequenza non è scontato quale sia la scelta migliore. L'inclusione di componenti di Fourier ad alta frequenza sicuramente va a sporcare

il filtraggio, essendo il segnale lentamente variabile in frequenza, ma un taglio eccessivamente basso al contrario rischia di escludere una quantità eccessiva di informazione.



Per ricercare empiricamente la configurazione ottimale del filtro si è presa una forma di prova di segnale r-mode, con le inclinazioni delle tangenti agli estremi note; con esse si sono costruiti un filtro triangolare semplice ed un filtro combinato triangolare gaussiano con i quali si è proceduto ad esplorare tutto il range $(0, 1]$ del taglio con un passo di incremento $\Delta = 0.01$. Per ogni valore $n\Delta$ del taglio si applicano i filtri che ne risultano ad una immagine contenente il segnale addizionato a rumore, prendendone 100 realizzazioni differenti: per ognuna di esse si calcola il *SNR* secondo la 3.27, utilizzando come stima il valor medio. Per ogni taglio si avrà allora un valore di *SNR* rispettivamente per il filtro triangolare ed il combinato e li si porrà a confronto con il valore che restituisce il filtro di Wiener.

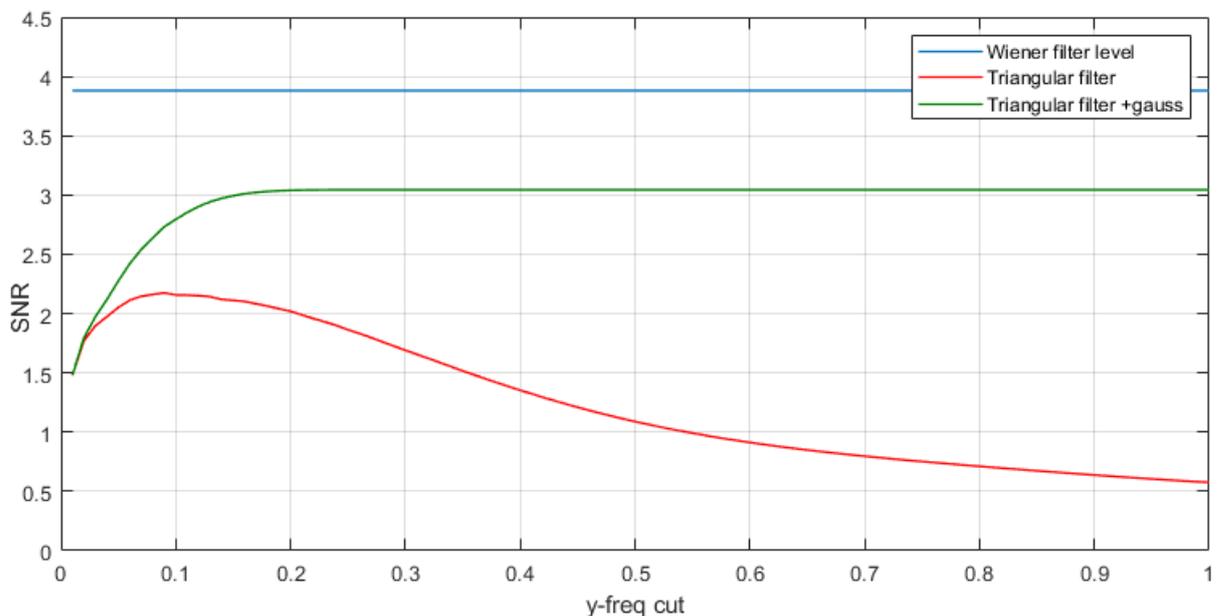


Figura 4.2: Variazione del *SNR* di una immagine filtrata, contenente un r-mode con $f_0 = 1500\text{Hz}$ con scala $f_{max} = 2000\text{Hz}$, durata $t_{max} = 30\text{s}$ ed ampiezza 1, in funzione del taglio in frequenza dei filtri triangolare (linea rossa) e combinato (linea gialla), esplorando a salti di $\Delta = 0.01$ nel range $(0, 1]$. Le stime del *SNR*, punto per punto, sono il valor medio di 100 differenti realizzazioni di rumore. Il livello di riferimento ideale costituito dal filtro di Wiener (linea blu) è ripetuto su una linea retta orizzontale in quanto indipendente dal taglio dei filtri triangolari.

In Figura 4.2 è mostrato il risultato di tale procedimento, avendo utilizzato un segnale di r-mode in tempo-frequenza con $f_0 = 1500\text{Hz}$ su una scala $f_{max} = 2000\text{Hz}$ della durata di $t_{max} = 30\text{s}$ e realizzazioni di rumore esponenziale ad ampiezza 5 volte quella del segnale. Si possono ricavare due indicazioni

fondamentali per le applicazioni future: per quanto riguarda il filtro triangolare, si riconosce un picco del SNR nell'intorno del valore di taglio in frequenza del 10%; tale picco non risulta abbastanza largo, cosicchè filtri con frequenza di taglio fino al 20% risultano ancora confrontabili, mentre dal 30% in poi le prestazioni scendono più rapidamente per finire a valori di $SNR < 1$ una volta superato il 55%. Il filtro combinato triangolare gaussiano presenta invece una dinamica completamente differente: cresce gradualmente fino a raggiungere il massimo del SNR , decisamente più vicino al livello di riferimento del filtro di Wiener rispetto al massimo del filtro triangolare semplice, attorno al taglio del 20%, per poi mantenere stabilmente tale livello per tutti i valori successivi. Tale comportamento, opposto rispetto a quello del filtro triangolare semplice, si comprende facilmente se si considera che l'applicazione della finestra gaussiana ha sì l'effetto di enfatizzare le zone interessate entro la distanza dell'ordine del σ , ma allo stesso modo attenua le zone lontane in modo esponenziale: nella pratica a distanze $> 3\sigma$ l'attenuazione è tale da essere indistinguibile dallo zero, dunque la finestra gaussiana toglie le alte frequenze a prescindere dal taglio del filtro triangolare e per alti valore del taglio lo rende di fatto superfluo. Questo da una parte risulta assai vantaggioso, in quanto in questo modo non ci si deve preoccupare di collocare il taglio del filtro su un preciso valore, che in una singola realizzazione è soggetto a fluttuazioni; d'altra parte le proporzioni che si devono dare alla larghezza della finestra gaussiana costituiscono un punto delicato, che ha a che fare con la scala del segnale rispetto alle dimensioni in pixel dell'immagine che lo contiene e che andrà approfondito successivamente. Ripetendo la simulazione con segnali differenti l'andamento non cambia da punto di vista qualitativo: si può vedere in Figura 4.3 che utilizzando un segnale della stessa durata del precedente ma che parta da $f_0 = 2000Hz$ l'efficacia dei filtri cala complessivamente, ma l'andamento qualitativo è lo stesso; in particolare il picco del filtro triangolare è raggiunto sempre nell'intorno del taglio al 10%, mentre l'efficacia del filtro combinato si stabilizza sempre dal 20% in poi.

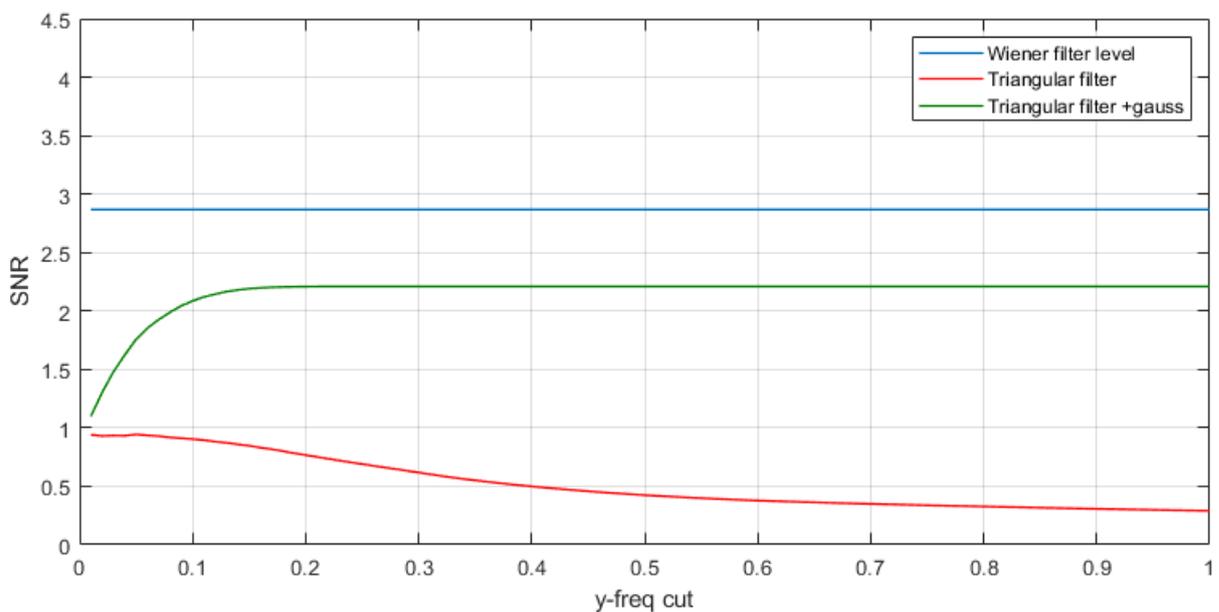


Figura 4.3: Variazione del SNR in funzione del taglio in frequenza: realizzazione analoga a quella riportata in Figura 4.2, ma con il segnale che parte da una frequenza di $f_0 = 2000Hz$.

Sono state ripetute altre prove con realizzazioni di rumore differenti, ovvero costruite imponendo una

ampiezza media maggiore. Anche in questo caso il livello generale del *SNR* si abbassa, ma confermando lo stesso andamento qualitativo. In conclusione nelle prove successive, nell'applicare filtri alle immagini, questi si utilizzeranno nella loro configurazione ottimale: si ricorrerà allora al taglio al 10% per il filtro triangolare, mentre per il filtro combinato triangolare gaussiano basterà impostare il taglio ad una certa distanza oltre il 20%, che è il livello in cui il *SNR* si stabilizza sul suo massimo.

4.1.2 Studio della risposta in funzione dell'ampiezza del rumore

Individuate le configurazioni ottimali in cui deve operare un filtro triangolare, semplice o combinato, occorre capire fino a che punto c'è la possibilità di far emergere una forma, nel nostro caso un r-mode in tempo-frequenza. Si è detto nella Sezione 3.4 che i filtri che si sono introdotti sono efficaci nel trovare una forma nascosta nel rumore, ma questa affermazione deve essere resa più quantitativa suffragandola da dati. In Figura 4.4 sono riportati i diversi andamenti di *SNR* su immagini filtrate in funzione dell'ampiezza media del rumore che si aggiunge al segnale, in scala semilogaritmica: sono visibili gli andamenti del filtro ideale di Wiener (linea rossa), dei filtri triangolari semplice (linea verde) e combinato (linea viola) nelle loro configurazioni ottimali e dell'immagine originale (linea azzurra) prima che le venisse applicato il filtro. Si vede come oltrepassato un certo livello di ampiezza del rumore, se pur continua ad esserci un miglioramento, il segnale continua ad avere *SNR* troppo basso. In particolare, in corrispondenza dell'ampiezza del rumore a 20 volte quella del segnale, il *SNR* dell'immagine filtrata con il filtro di Wiener vale circa 1, che può essere considerato lo spartiacque per ritenere un segnale osservabile; ancora peggio va ai filtri triangolari i quali, come si vede, si trovano sempre al di sotto del filtro di Wiener. La regione fisicamente più interessante risulta ovviamente quella compresa nella regione di

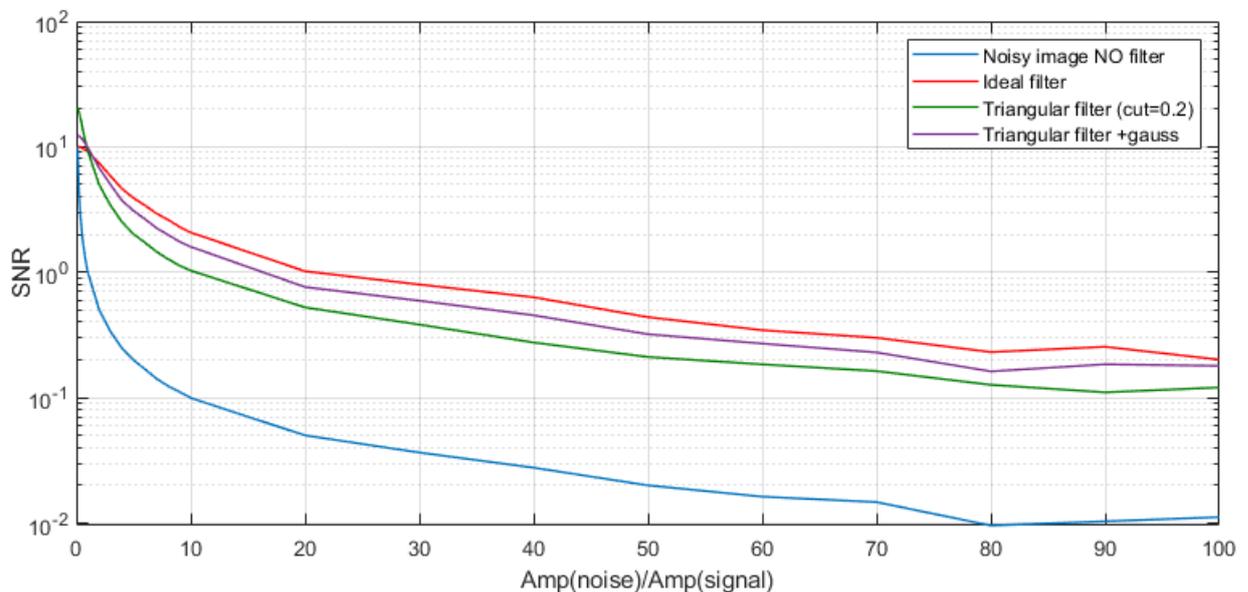


Figura 4.4: Andamento del *SNR* di un segnale da r-mode con $f_0 = 1500\text{Hz}$ su una scala con scala $f_{max} = 2000\text{Hz}$, durata $t_{max} = 30\text{s}$ ed ampiezza 1 addizionato a rumore di ampiezza media variabile, in funzione di tale ampiezza, in scala semilogaritmica.

ampiezza del rumore [1, 10]: al di sotto il segnale potrebbe essere osservabile anche senza l'ausilio dei filtri, mentre al di sopra anche i filtri iniziano a faticare a renderlo evidente. Nell'immagine in alto in Figura 4.5 è mostrato il dettaglio del grafico precedente, in scala lineare, fino ad una ampiezza relativa di

10, mentre nell'immagine in basso è riportato il fattore di guadagno relativo nella stessa regione, ovvero il rapporto tra i SNR dopo e prima del filtraggio. Si vede che il guadagno portato dal filtro, in senso relativo, risulta maggiore nelle configurazioni peggiori, ovvero quando l'ampiezza del rumore diventa 3 – 4 volte quella del segnale; in particolare, superati tali valori, il guadagno relativo si stabilizza su uno stesso livello che ne può rappresentare il limite massimo. Ovviamente, non aumentando il guadagno al crescere dell'ampiezza del rumore, quando l'ampiezza del rumore supera un certo livello il guadagno del filtro non sarà più sufficiente a portare il segnale in evidenza; per quello che riguarda il filtro di Wiener, il suo guadagno si stabilizza attorno ad un massimo di ~ 20 , dunque quando l'ampiezza media del rumore è 20 l'immagine originale avrà $SNR \sim 1/20$, dunque l'immagine filtrata avrà $SNR \sim 1$.

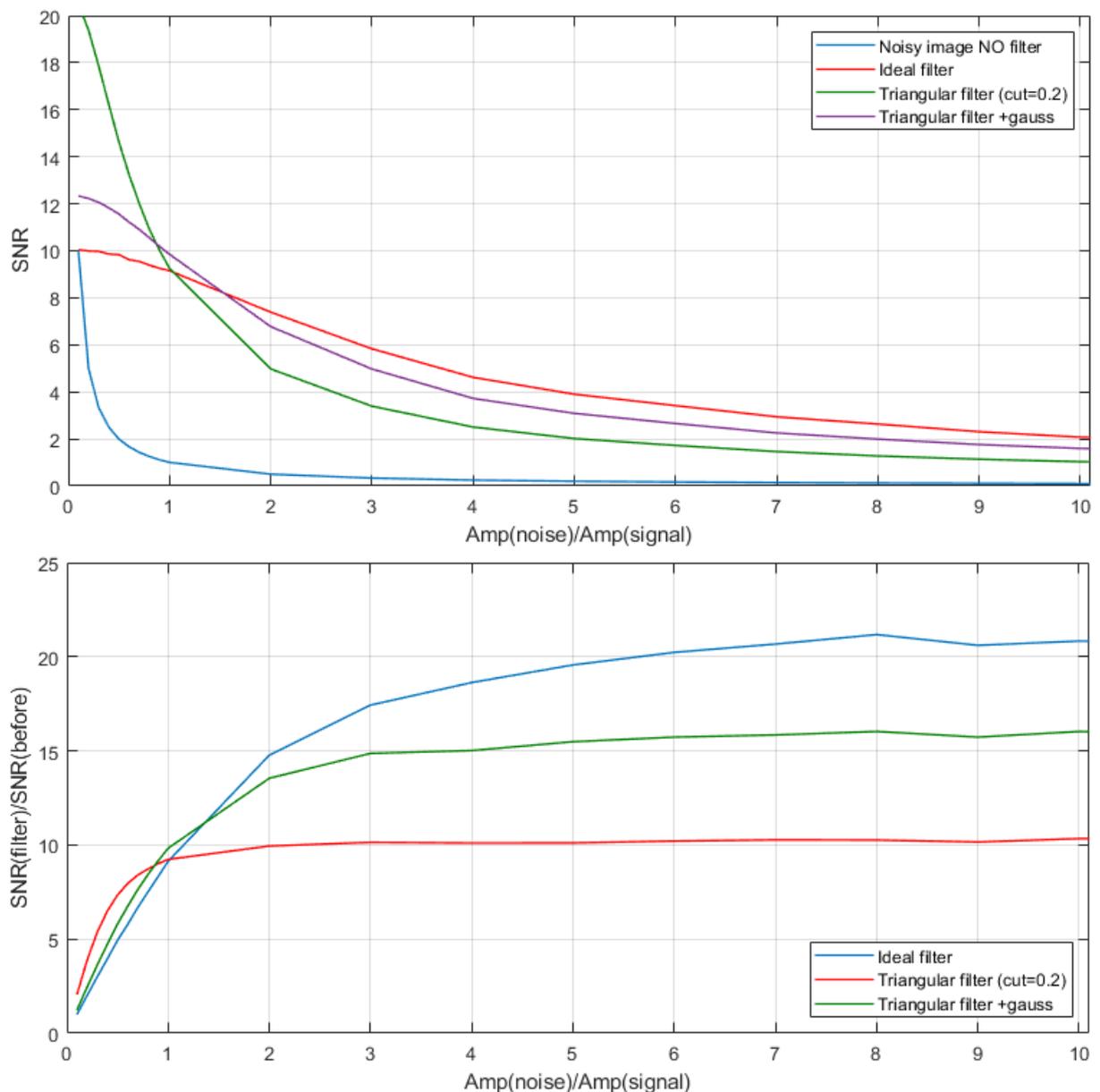


Figura 4.5: In alto: ingrandimento del grafico di Figura 4.4 nella regione $(0, 10]$ dell'ampiezza del rumore. In basso: rapporto tra SNR delle immagini filtrate e SNR delle immagini prima dell'applicazione del filtro, visualizzato nella stessa regione di cui sopra.

Tramite gli esempi riportati sopra siamo in grado di capire il limite di ampiezza di rumore dal quale i filtri sono in grado di rintracciare un segnale da r-mode. Tuttavia il parametro di riferimento è difficilmente utilizzabile in una ricerca reale o in una simulazione: fino ad ora, nella Sezione 3.4 e nel presente capitolo, si è per comodità generato un segnale di ampiezza costante già nel formato che ci si aspetta di osservare nel dominio tempo-frequenza e gli si è addizionato del rumore a distribuzione esponenziale. In una simulazione non sono questi parametri ad essere sotto controllo: la procedura standard consiste nel generare una forma d'onda -nel caso presente sempre un r-mode- in ampiezza nel dominio del tempo e nell'addizionala a rumore, anch'esso nel dominio del tempo, a distribuzione gaussiana con media nulla ed una data ampiezza, eventualmente controllabile. A partire da questi dati si costruiscono gli spettrogrammi seguendo la procedura descritta nella Sezione 2.4 e ad essi si applicano i filtri per mettere in evidenza il segnale, nel dominio tempo-frequenza, se poco evidente. E' utile allora visualizzare i dati riportati in Figura 4.4 riferendo i valori del *SNR* ottenuti dopo l'applicazione dei filtri non all'ampiezza media del rumore, che nello spettrogramma diventa esponenziale, ma direttamente al *SNR* dell'immagine prima dell'applicazione dei filtri: in questo modo si ha a disposizione un indicatore di riferimento più generale utilizzabile in sede di simulazione. In Figura 4.6 è riportato tale andamento, visualizzato per comodità in scala logaritmica.

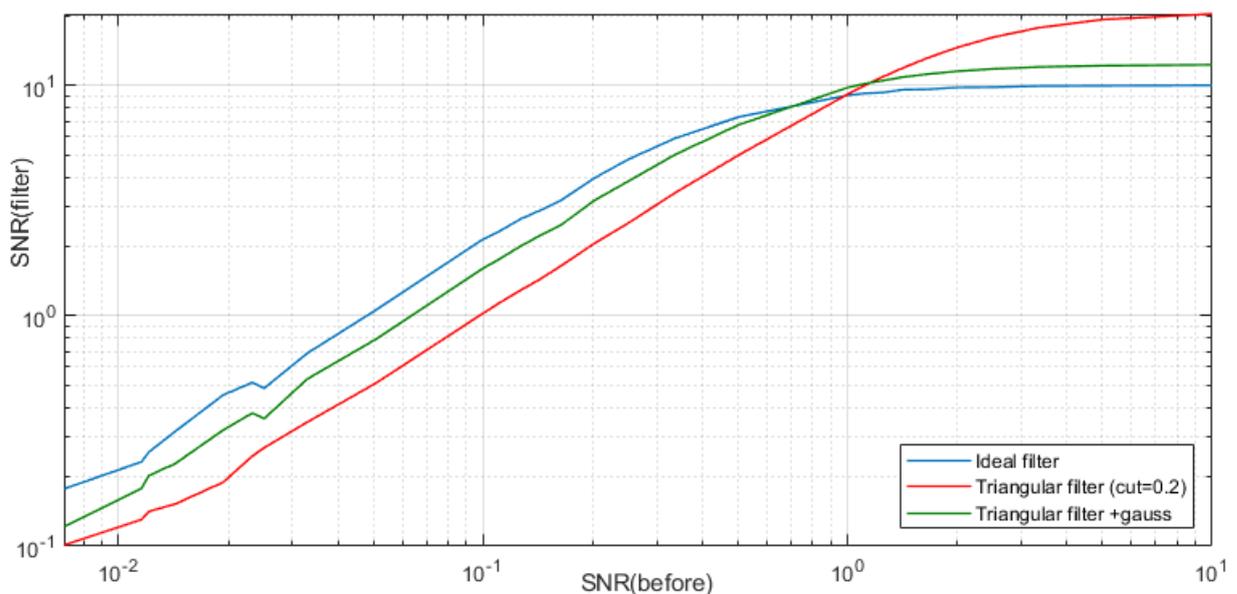


Figura 4.6: Andamento del *SNR* delle immagini filtrate in funzione del *SNR* dell'immagine prima del filtraggio, in scala logaritmica. I dati sono gli stessi utilizzati per le Figure 4.4, 4.5.

Si può notare un fatto curioso, ovvero che per valori del *SNR* di partenza maggiori dell'unità (ovvero per ampiezze medie del rumore minori di 1) i filtri triangolari risultano più efficaci di quello ideale. Questo tuttavia per la nostra analisi ha poca importanza in quanto in situazioni del genere il segnale risulta già visibile senza il bisogno di applicare alcun filtro. Si vede invece che l'applicazione dei filtri risulta effettivamente conveniente nella regione in cui il *SNR* dell'immagine originale è nel range $[0.1, 1]$: per valori minori c'è ancora una piccola regione in cui il filtro di Wiener è ancora in parte efficace, ma i filtri triangolari, che sono quelli che si pensa di poter utilizzare, non riescono più a mettere in evidenza il segnale.

4.1.3 Studio della risposta in funzione delle inclinazioni di taglio

Quando si effettua una ricerca reale di segnali da r-mode l'ostacolo maggiore è l'indeterminazione che esiste su molti parametri fondamentali. Si è già parlato dell'indice di frenata e dell'ampiezza di saturazione; altre caratteristiche difficili da prevedere sono la frequenza di partenza e la durata stessa del segnale: quest'ultima potrebbe essere dell'ordine dei secondi come dei mesi, mentre anche se si ha una stima della frequenza di rotazione di una stella di neutroni appena nata grazie ai parametri orbitali delle progenitrici, non è scontato che eventuali instabilità da r-mode si manifestino solamente dopo un certo periodo di tempo, partendo da una frequenza più bassa. Tutte le incognite citate portano a rendere poco realistica la possibilità di applicare dei filtri triangolari mirati, le cui inclinazioni limite combacino con gli effettivi inizio e fine di un segnale. Bisogna dunque applicare dei filtri, eventualmente più di uno, che abbiano un range più ampio in modo da avere più generalità nella selezione. Per capire in che proporzione tale scelta debba essere fatta occorre allora studiare quantitativamente l'ammontare della perdita in *SNR* in funzione di quanto si allarga il range dei filtri triangolari rispetto al range ideale di un dato segnale.

Chiamate con \dot{f}_{min} ed \dot{f}_{max} le inclinazioni delle tangenti agli estremi della rappresentazione in tempo-frequenza di un segnale da r-mode, queste si possono rispettivamente contrarre e dilatare di un fattore β compreso tra 0 ed 1 in modo da avere:

$$\dot{f}'_{min} = \dot{f}_{min}\beta \quad ; \quad \dot{f}'_{max} = \dot{f}_{max} \frac{1}{\beta} \quad , \quad \beta \in [0, 1] \quad (4.1)$$

In pratica si sta abbassando l'inclinazione minima della tangente fino a renderla orizzontale ed alzando l'inclinazione massima fino a renderla verticale. Si realizza allora un campionamento di possibili valori di β dividendone in 100 parti l'intervallo di definizione; per ogni valore di β si calcola il corrispondente range modificato rispetto al range ideale di un segnale fissato: da esso, tramite la 3.22, si realizzano i corrispondenti filtri triangolari, semplici e combinati, e si prende il *SNR* delle immagini risultanti dall'applicazione di tali filtri. La prima prova è stata svolta applicando i fattori ad entrambe le inclinazioni limite.



Nel caso estremo di $\beta = 0$ il filtro risultante rappresenta una selezione rettangolare su primo e terzo quadrante: esso corrisponde a selezionare un qualunque segnale di qualsivoglia inclinazione con la sola richiesta che abbia frequenza decrescente nel tempo; il caso $\beta = 1$ corrisponde invece al filtro triangolare nella sua configurazione ottimale. Gli esempi riportati nelle Figure 4.2 e 4.3 possono suggerire che un segnale che nel suo sviluppo interessa un range di inclinazioni più ampio di un altro, a parità di lunghezza, risulterà sfavorito in sede di filtraggio. In Figura 4.7 è riportato l'andamento del *SNR* delle immagini filtrate, mediato su 100 differenti realizzazioni di rumore esponenziale di ampiezza media 5, in funzione del valore del parametro β che viene utilizzato per aumentare e diminuire le inclinazioni limite del range secondo la 4.1: in essa si vede come varia l'efficienza di filtraggio dei filtri triangolari semplice e combinato all'allargarsi del range. La variazione relativa del *SNR* del filtro triangolare combinato risulta

contenuta, con una tenuta nel caso peggiore dell'80%, mentre la perdita di efficacia nel filtro triangolare semplice risulta eccessiva nel caso estremo, con una tenuta attorno al 25%.

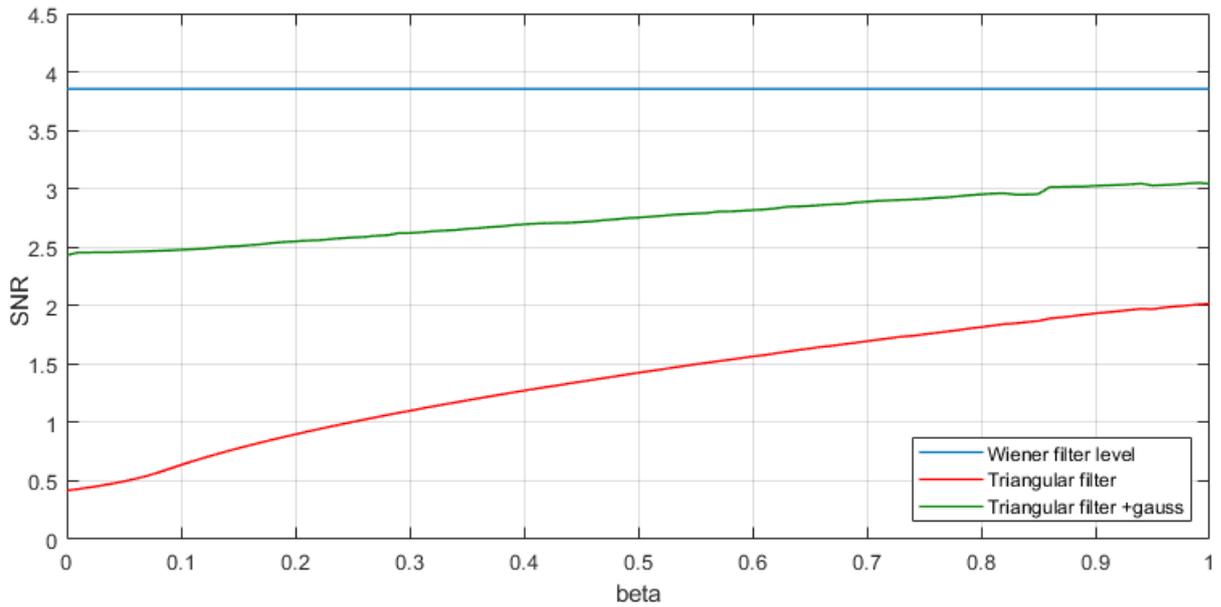


Figura 4.7: Andamento del SNR di un segnale da r-mode con $f_0 = 1500Hz$ su una scala con scala $f_{max} = 2000Hz$, durata $t_{max} = 30s$ ed ampiezza 1 addizionato a rumore di ampiezza media 5, in funzione di un fattore β tramite il quale si contraggono e dilatano le inclinazioni limite del range dei filtri triangolari secondo la 4.1.

Per rendere più evidente quale sia il contributo dominante a tale dinamica si è ripetuta la prova applicando il fattore β in modo alternato ad una sola delle due inclinazioni limite, lasciando bloccata l'altra. In Figura 4.8 è stato applicato il parametro β solamente come fattore moltiplicativo dell'inclinazione minima del segnale.

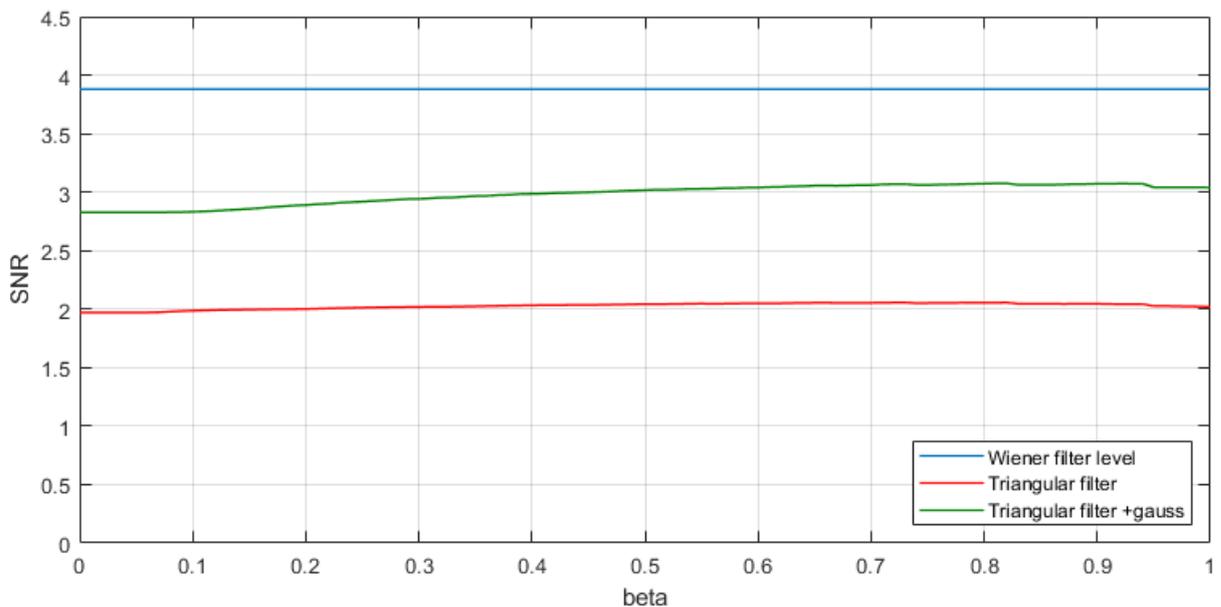


Figura 4.8: Realizzazione analoga a quella in figura 4.7, ma applicando β solamente a moltiplicare \dot{f}_{min} .

Si ricordi che l'inclinazione minima del segnale nello spazio di Fourier diventa la retta ad inclinazione maggiore, dunque nel limite $\beta \rightarrow 0$ si ha la retta più inclinata del filtro triangolare che tende a diventare verticale. La variazione indotta nel *SNR* da questa sola componente risulta estremamente piccola in entrambi i sistemi di filtraggio: se ne deduce che la dinamica riscontrata in Figura 4.7 trova il contributo principale nella dilatazione dell'inclinazione massima \hat{f}_{max} . Ripetendo allora il calcolo precedente, tenendo stavolta congelata l'inclinazione \hat{f}_{min} e applicando il parametro β a dividere l'inclinazione massima \hat{f}_{max} , si ottiene l'andamento visibile in Figura 4.9: risulta chiaramente la similitudine degli andamenti dei filtri triangolari con quelli visti in Figura 4.7. La diversità dei comportamenti delle curve dei due filtri trova una spiegazione analoga a quella per la quale l'efficacia del filtro triangolare combinato risulta indipendente dal taglio in frequenza in verticale: anche in questo caso la finestra gaussiana contrae le ampiezze lungo la direzione orizzontale, tagliando di fatto una buona fetta di alte frequenze che nel filtro triangolare rimangono incluse.

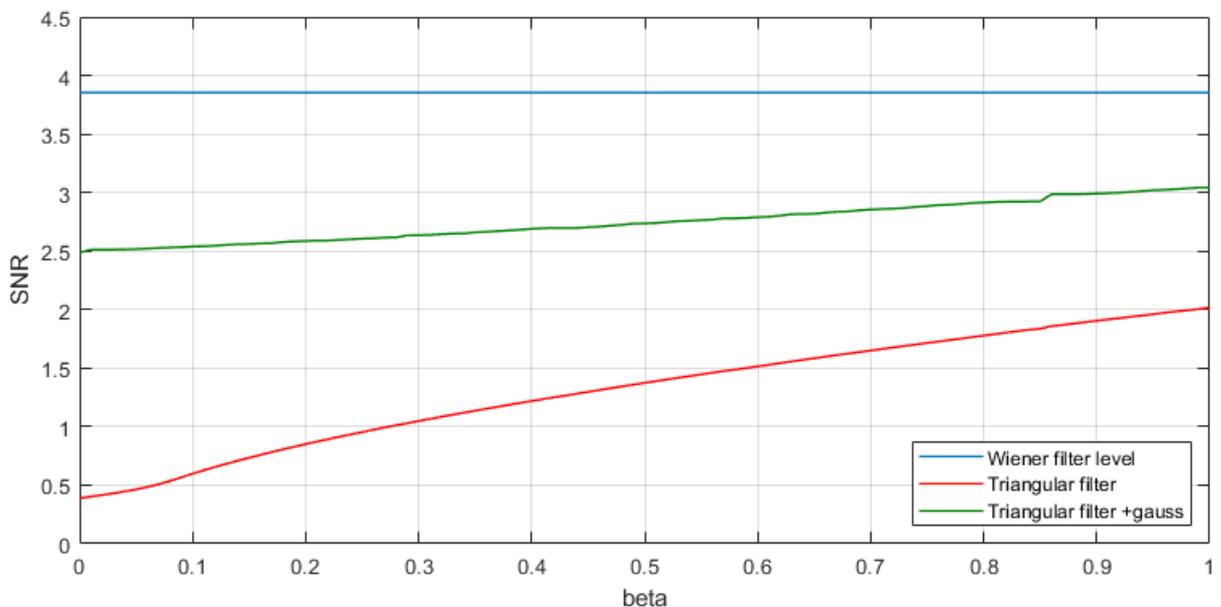


Figura 4.9: Realizzazione analoga a quella in figura 4.7, ma applicando β solamente a dividere \hat{f}_{max} .

Si può concludere che l'estensione del range di un filtro triangolare dalla parte delle basse inclinazioni non comporta un sensibile cambiamento nel *SNR* dell'immagine filtrata, di conseguenza si può in generale applicare un filtro triangolare ad un r-mode in tempo-frequenza preoccupandosi unicamente dell'inclinazione del segnale in partenza ed estendendo le basse inclinazioni a quella nulla. Se si utilizza un filtro triangolare semplice occorrerà prestare particolare attenzione al livello di partenza da cui costruire il filtro: come si è visto, una scelta sbagliata può portare ad un abbassamento eccessivo del *SNR*. Diversa è invece la situazione con il filtro triangolare combinato: data la bassa perdita relativa di *SNR* si potrebbe pensare, in alle esigenze, di applicarlo nella sua configurazione di range esteso ai 90° ; questo accorgimento consentirebbe di rilevare segnali con assunzioni minime sulla loro forma, ma bisogna ricordare che per poter utilizzare in un contesto generale il filtro combinato si dovrà essere in grado di gestire la larghezza della finestra gaussiana.

4.2 Effetto del filtro triangolare sulla costruzione delle peakmap

In questa sezione si applicheranno i concetti finora acquisiti ad una situazione più realistica: partendo da un segnale da r-mode simulato ed addizionato a rumore bianco nel dominio del tempo, si costruirà lo spettrogramma dei dati. L'obiettivo è confrontare l'effetto dell'applicazione di un filtro alla selezione delle peakmap con una situazione nella quale il filtro non viene applicato.

Il segnale simulato si considera generato da una stella di neutroni distante $40Mpc$ ed a localizzazione nota, in modo da considerare già corretto l'effetto Doppler, e presenta l'indice di frenata tipico di un r-mode puro, pari ad $n = 7$. Il segnale parte da una frequenza di $f_0 = 1000Hz$ e nella dinamica di rallentamento si considera un'ampiezza di saturazione $\alpha = 1$; gli si attribuisce una durata complessiva di $1000s$. Tenendo conto del teorema di Nyquist, il segnale viene campionato con un tempo di campionamento $\Delta t_s = \frac{1}{8f_0}$: in questo modo va a popolare un vettore contenente $8 \cdot 10^6$ elementi. Vengono poi prodotte delle realizzazioni di rumore bianco con lo stesso numero di elementi con deviazione rispettivamente 10, 50, 100 volte maggiore dell'ampiezza massima del segnale, le quali vengono poi addizionate al segnale. In figura 4.10 sono mostrati a titolo di esempio i dati generati con la realizzazione di rumore ad ampiezza 10 volte il segnale. Può essere utile valutare le ampiezze spettrali del segnale e delle 3 differenti realizzazioni

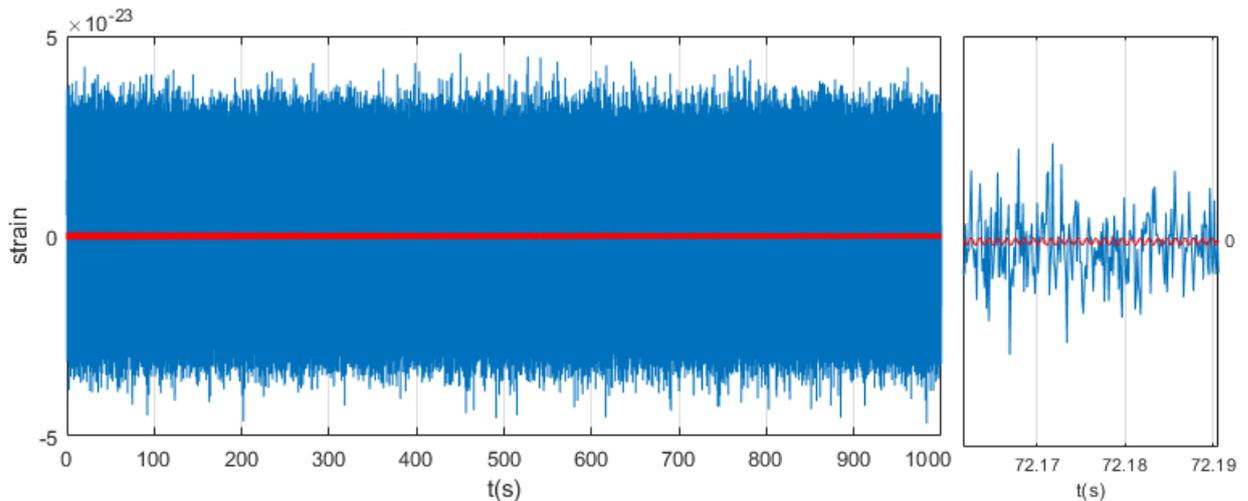


Figura 4.10: In rosso: forma d'onda nel dominio temporale da r-mode generata con $f_0 = 1000Hz$, $\alpha = 1$ ed $n = 7$, campionata con $\Delta t_s = \frac{1}{8f_0}$. In azzurro: forma d'onda addizionata a rumore bianco con deviazione pari a 10 volte l'ampiezza massima del segnale. A destra: ingrandimento di una porzione di dati alla scala dei centesimi di secondo.

di rumore per capire in quali regimi ci si trova. Per quanto riguarda il rumore bianco, per produrre una stima della densità spettrale si è scelto di dividere i dati in 1000 sottoinsiemi in modo da attenuare le fluttuazioni. In Figura 4.11 sono riportate, a confronto, le radici quadrate delle stime della densità spettrale monolatera delle 3 realizzazioni, sovrapposte all'ampiezza spettrale del segnale. Si vede che nel primo caso, corrispondente all'ampiezza del rumore a 10 volte il segnale, il segnale esce per una parte considerevole al di sopra del rumore (si ricordi che il grafico è in scala logaritmica), dunque ci si aspetta che in questa situazione negli spettrogrammi il segnale sia riconoscibile senza bisogno di filtri; nel caso del rumore a 50 volte il segnale ci si trova in una situazione intermedia, con il segnale difficilmente distinguibile ma nel range di SNR nel quale i filtri hanno la massima efficacia; l'ultimo caso, del rumore a 100 volte il segnale, potrebbe risultare troppo predominante da permettere di ritrovare il segnale. Per la

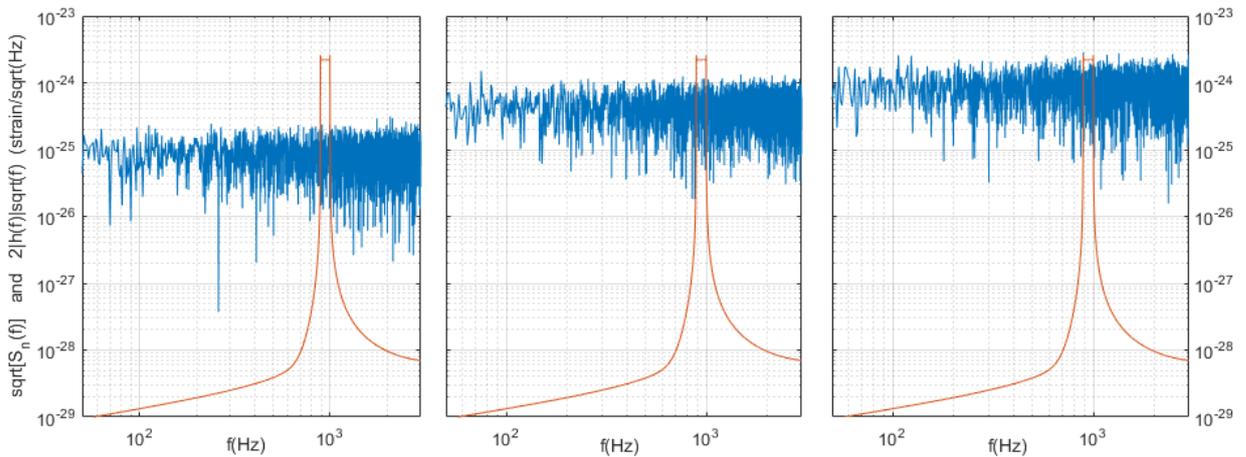


Figura 4.11: Ampiezze spettrali del segnale (in rosso) e delle 3 realizzazioni di rumore bianco (in azzurro) poste a confronto. Gli spettri del rumore sono calcolati mediando su 1000 sottoinsiemi uguali del campione. A sinistra: rumore con deviazione pari a 10 volte l'ampiezza massima del segnale. Al centro: rumore con deviazione pari a 50 volte l'ampiezza massima del segnale. A destra: rumore con deviazione pari a 100 volte l'ampiezza massima del segnale.

realizzazione degli spettrogrammi la cosa più importante è determinare il tempo di coerenza giusto. Data la frequenza iniziale del segnale, l'indice di frenata e l'ampiezza di saturazione, si stima una lunghezza ottimale di $T_{FFT} = 4s$, che corrisponde ad una lunghezza in pixel di 32000 pixel. Si imposta poi una finestrazione flat-top l'interlacciamento delle trasformate a metà del tempo di coerenza, in modo da avere ogni colonna di dati dello spettrogramma scorrelata dalla precedente e dalla successiva. In questo modo gli spettrogrammi che si ottengono sono delle matrici $M \times N$ con $M = 498$ pixel ed $N = 1601$ pixel, con definizione temporale $\Delta T = 2s$ e definizione in frequenza $\delta f = 0.25Hz$. Tutti e 3 sono mostrati in Figura 4.12: nel primo caso, a cui ci si riferirà sinteticamente come rumore a 10, il segnale è ben visibile ad occhio; nel caso del rumore a 50 questo è appena percettibile, mentre nel rumore a 100 è del tutto nascosto.

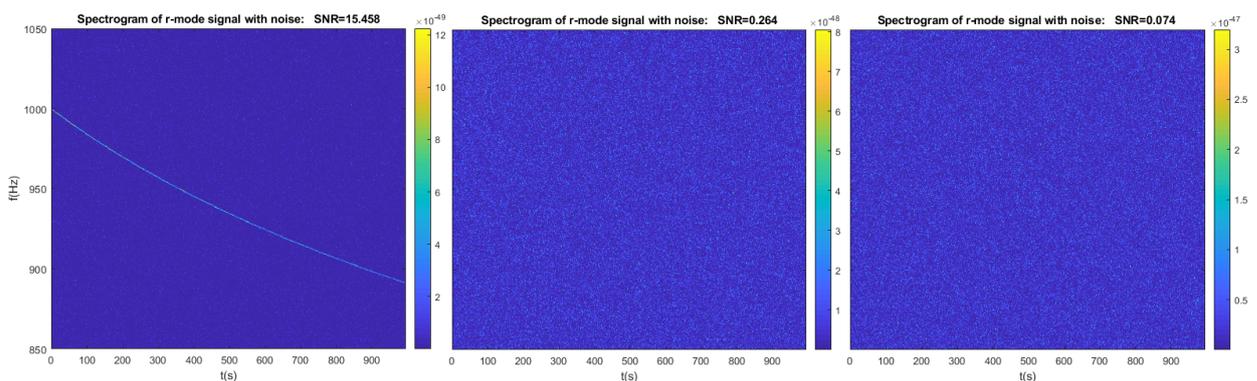


Figura 4.12: Ingrandimento degli spettrogrammi delle 3 differenti realizzazioni di rumori addizionate al segnale r-mode, costruiti con tempo di coerenza $T_{FFT} = 4s$, interlacciamento a metà $\Delta T = 2s$ e finestrazione flat-top. A sinistra: spettrogramma con rumore a 10. Al centro: spettrogramma con rumore a 50. A destra: spettrogramma con rumore a 100.

Essendo nota la forma d'onda nascosta nel rumore, si possono utilizzare le sue informazioni per costruire

il filtro di Wiener ed il filtro triangolare semplice da applicare agli spettrogrammi. A tutte e 3 le immagini originali, i cui valori di SNR sono visibili in Figura 4.12, si applicano entrambi i filtri. In Figura 4.13 sono riportate, nello stesso ordine, le immagini a seguito dell'applicazione dei filtri. Si possono ritrovare gli stessi rapporti tra SNR prima e dopo i filtri che si erano individuati in Figura 4.6 nella Sezione 4.1.2: in particolare, nel caso del rumore a 10, che presentava in partenza un SNR già elevato, ottiene un guadagno più modesto dai filtri ed inoltre l'efficacia del filtro triangolare supera quella del filtro di Wiener; l'immagine col rumore a 50, che presenta un valore del SNR nella regione di maggior guadagno da parte dei filtri, risulta effettivamente quella più interessante; invece l'immagine col rumore a 100 presenta un SNR troppo basso perchè i filtri possano rimetterlo in evidenza. In Tabella 4.1 sono riportati, per maggiore compattezza, i valori dei diversi SNR prima e dopo l'applicazione dei filtri.

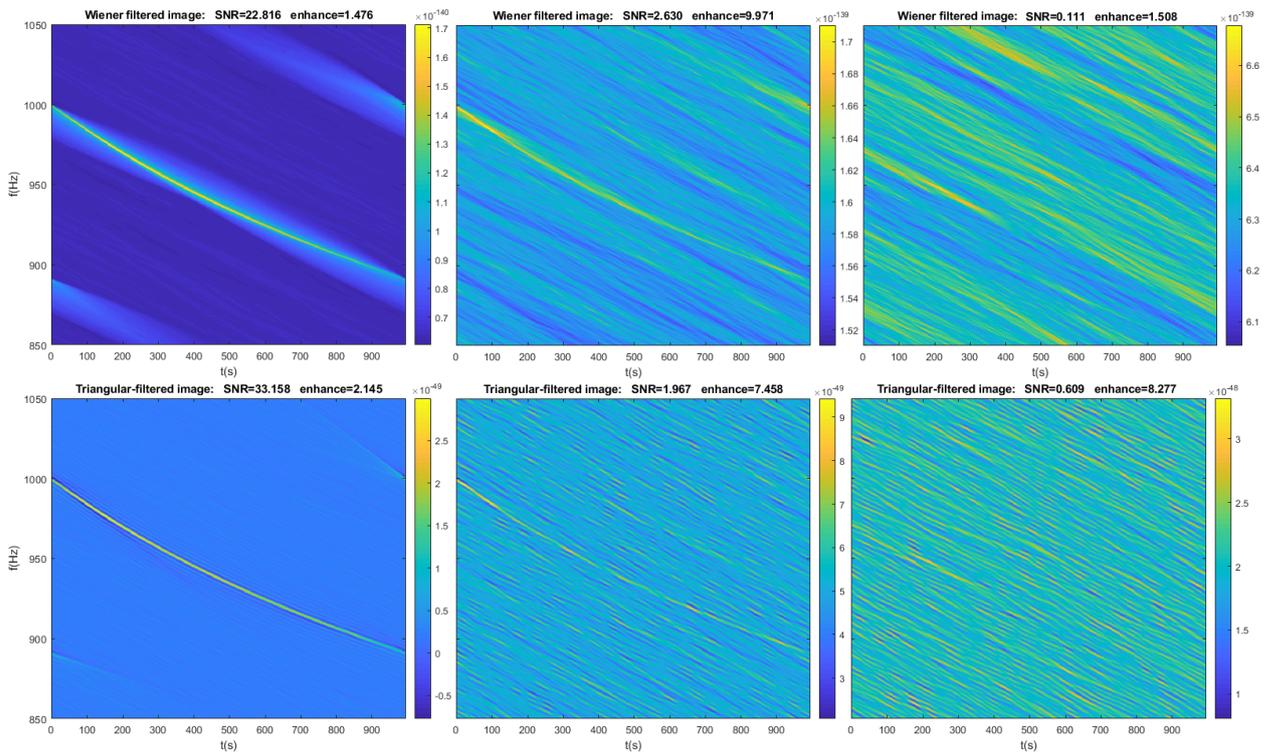


Figura 4.13: In alto: risultato dell'applicazione del filtro di Wiener alle immagini con le 3 differenti realizzazioni di rumore. In basso: risultato dell'applicazione del filtro triangolare. A sinistra: spettrogramma con rumore a 10. Al centro: spettrogramma con rumore a 50. A destra: spettrogramma con rumore a 100. Tutte le immagini sono ingrandite attorno al segnale.

	SNR iniziale	SNR Wiener	Guadagno Wiener	SNR triangolare	Guadagno triangolare
rumore a 10	15.458	22.816	1.476	33.158	2.145
rumore a 50	0.264	2.630	9.971	1.967	7.458
rumore a 100	0.074	0.111	1.508	0.609	8.277

Tabella 4.1: Valori di SNR relativi alle immagini contenenti le 3 differenti realizzazioni di rumore, prima e dopo l'applicazione dei filtri, compreso il guadagno.

Su tutte le immagini viene poi applicata la procedura per la creazione delle peakmap. Allo stato attuale non si hanno a disposizione valori di riferimento circa la soglia ottimale da porre per la selezione dei picchi: si procede allora ad esplorare i possibili livelli partendo dal valor medio delle ampiezze di tutta l'immagine fino ad arrivare, attraverso 100 passi equispaziati, al valore del massimo assoluto dell'immagine. Per ogni peakmap così costruita si effettua il conteggio di quanti picchi cadono all'interno della maschera associata al segnale e quanti invece cadono al di fuori, in modo da tenere sotto controllo fino a quale livello di soglia un segnale viene selezionato. Si ottengono in questo modo, per ognuna delle realizzazioni, delle curve che riportano il numero di picchi dentro o fuori la maschera in funzione di una soglia relativa, compresa tra 0 ed 1, che rappresenta la progressione relativa tra l'ampiezza media e l'ampiezza massima dell'immagine. Ogni filtro produrrà una curva del genere e dal confronto reciproco si cercheranno di trarre indicazioni che possano orientare le ricerche successive. In Figura 4.14 si possono vedere le curve ottenute nel caso dell'immagine con rumore a 10: il numero massimo di picchi dentro la maschera, che è dell'ordine di 500, è da considerarsi il risultato ottimale tenendo conto del fatto che gli spettrogrammi costruiti hanno la dimensione temporale di 498 pixel.

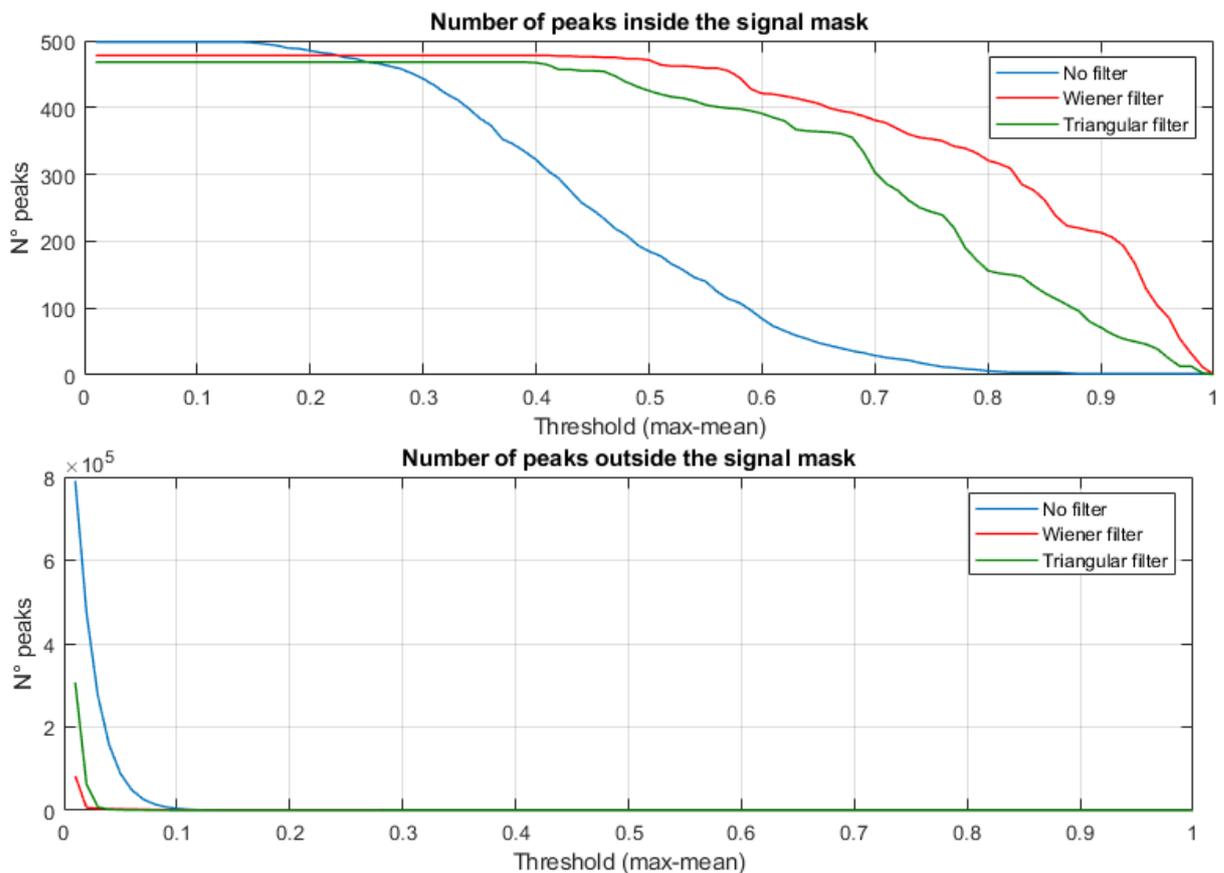


Figura 4.14: In alto: numero di picchi selezionati all'interno della maschera del segnale nelle peakmap dell'immagine a rumore 10, in funzione della soglia relativa che spazia dall'ampiezza media dell'immagine alla massima. In basso: numero di picchi selezionati all'esterno della maschera del segnale.

Si può vedere come in questo caso già l'immagine originale, da sola, consente di selezionare un numero elevato di picchi relativi al segnale; l'effetto dei filtri di Wiener e triangolare è quello di spostare più in avanti il livello di soglia oltre il quale si iniziano a perdere dei picchi; un effetto secondario è che

mentre mantengono elevati i picchi di segnale i filtri hanno l'effetto di deprimere il numero di picchi di rumore rispetto a quanti se ne selezionavano senza filtraggio. In generale in una situazione del tipo del rumore a 10 l'azione dei filtri non sembra essere determinante. In Figura 4.15 si possono vedere invece le curve ottenute con l'immagine con rumore a 50. Questa è sicuramente una situazione più interessante rispetto alla precedente, in quanto si vede chiaramente che l'efficacia nella selezione dei picchi di segnale nell'immagine non filtrata scende molto rapidamente: in questa configurazione si rischierebbe di perdere il segnale. L'azione dei filtri invece riesce a mantenere in evidenza tali picchi, facendo sì che essi continuino a venir selezionati molto più a lungo; si può notare in questa situazione una maggiore tenuta del numero di picchi al di fuori della maschera, ma questi al più andranno ad aumentare il numero di candidati, per venire poi scartati nei livelli di analisi successivi. Si intravede tuttavia anche la possibilità di ottimizzare questo aspetto: si vede infatti, relativamente al filtro triangolare, che il numero di picchi di segnale selezionati si mantiene elevato fino ad una soglia del 50%, mentre il numero di picchi di rumore inizia a calare non appena ci si allontana dal valor medio; in altre parole si è evidenziata una regione dei livelli di soglia, compresa nel range $[0, 50]\%$, all'interno della quale il numero di picchi di segnale selezionati rimane costante mentre il numero di picchi di rumore diminuisce. Tale regione potrebbe rappresentare la zona in cui ottimizzare al massimo la selezione dei picchi, ma ci sarà bisogno di svolgere un elevato numero di prove prima di poterne trarre indicazioni precise.

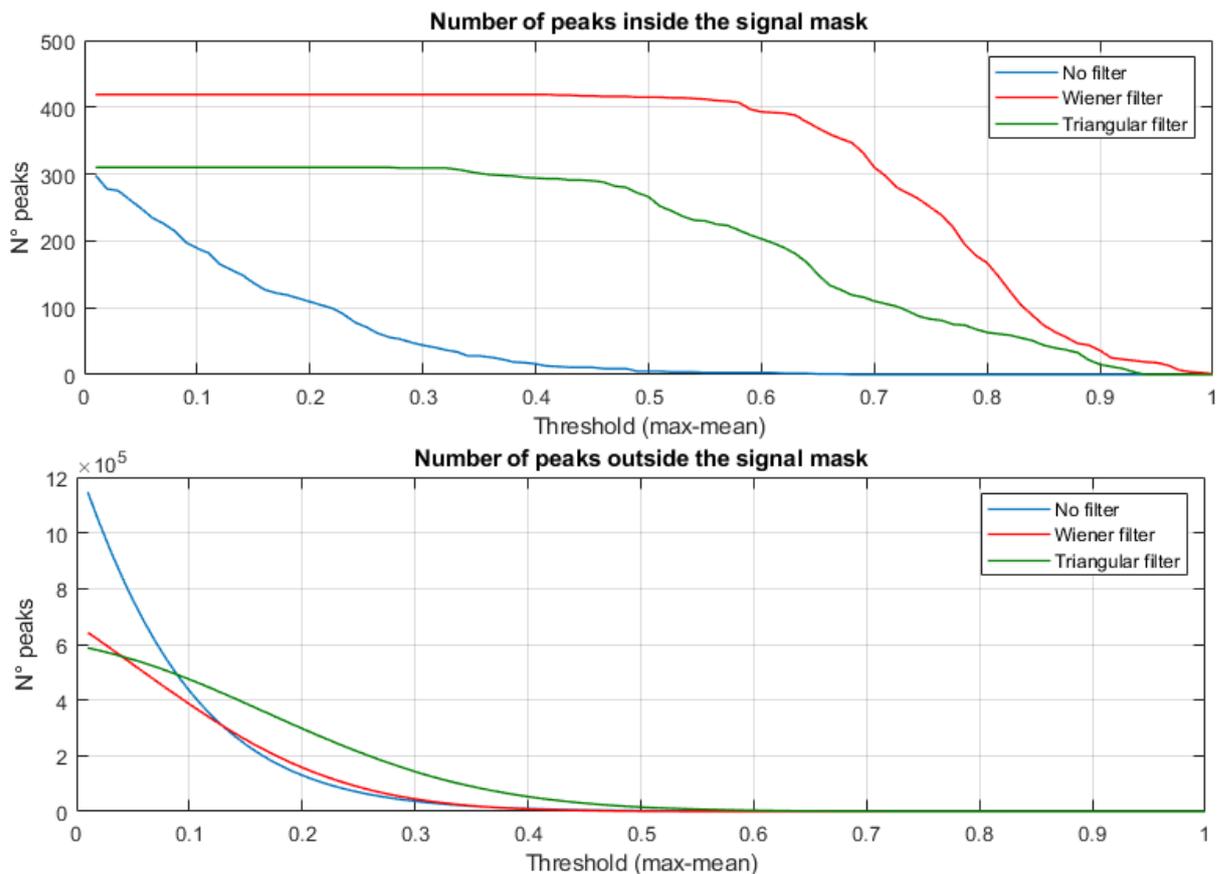


Figura 4.15: In alto: numero di picchi selezionati all'interno della maschera del segnale nelle peakmap dell'immagine a rumore 50, in funzione della soglia relativa. In basso: numero di picchi selezionati all'esterno della maschera del segnale.

Si può vedere infine in Figura 4.16 l'andamento delle curve ottenute dall'immagine con rumore a 100. Da un punto di vista qualitativo l'andamento che si osserva è simile a quello visto per il rumore a 50, ma rispetto ad esso il numero di picchi di segnale selezionati risulta sostanzialmente diminuito: i picchi selezionati grazie al filtro di Wiener raggiungono la metà del numero di pixel che compongono la dimensione temporale nel caso migliore, mentre nelle immagini filtrate con filtro triangolare tale numero supera di poco il centinaio. In questa situazione potrebbe risultare problematica l'individuazione del segnale, ma questo dipende anche dalla soglia di conteggi con cui nella fase successiva, nella mappa di Hough, si selezioneranno i candidati più significativi. Il problema è che in questa situazione il segnale reale rischia di rimanere sommerso da un mare di falsi candidati prodotti dallo shaping del filtro.

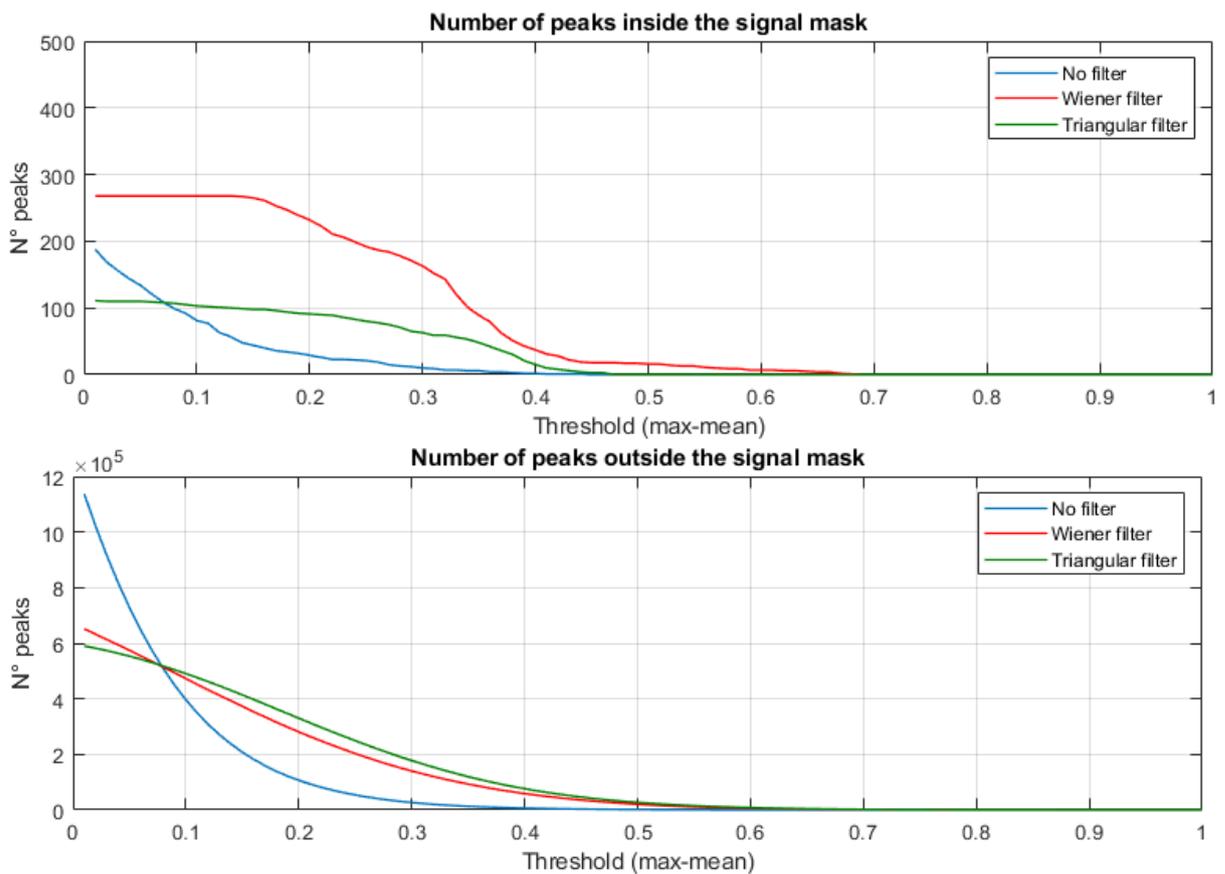


Figura 4.16: In alto: numero di picchi selezionati all'interno della maschera del segnale nelle peakmap dell'immagine a rumore 100, in funzione della soglia relativa. In basso: numero di picchi selezionati all'esterno della maschera del segnale.

Tornando al caso del rumore a 50 è utile soffermarsi maggiormente sugli effetti del filtro triangolare sulle peakmap, dato che in una ricerca reale non si ha a disposizione la forma del segnale. L'individuazione della soglia ideale oltre la quale selezionare un picco è un passo delicato, che concorre a determinare la sensibilità di ricerca dell'intero algoritmo. In Figura 4.17 sono mostrate le peakmap corrispondenti alle soglie di selezione allo 0% ed al 50%, che delimitano la zona ottimale precedentemente individuata, nella quale il numero di picchi di rumore diminuisce al crescere della soglia mentre i picchi di segnale permangono: il segnale è chiaramente visibile in entrambe le immagini. La configurazione con soglia al valor medio (a sinistra), se si guarda all'andamento delle curve in Figura 4.15, è sostanzialmente la stessa

che si ottiene anche con valori di soglia più bassi; i numerosi picchi selezionati dovuti al rumore hanno un andamento simile a quello del segnale, a causa dell'azione dei filtri; alzando il valore della soglia tali curve inizieranno a perdersi fino ad arrivare alla configurazione al 50% (a destra), dove la curva del segnale è rimasta quasi del tutto integra e la maggior parte del rumore è perduta; alzando ulteriormente la soglia anche i picchi di segnale andranno diradandosi.

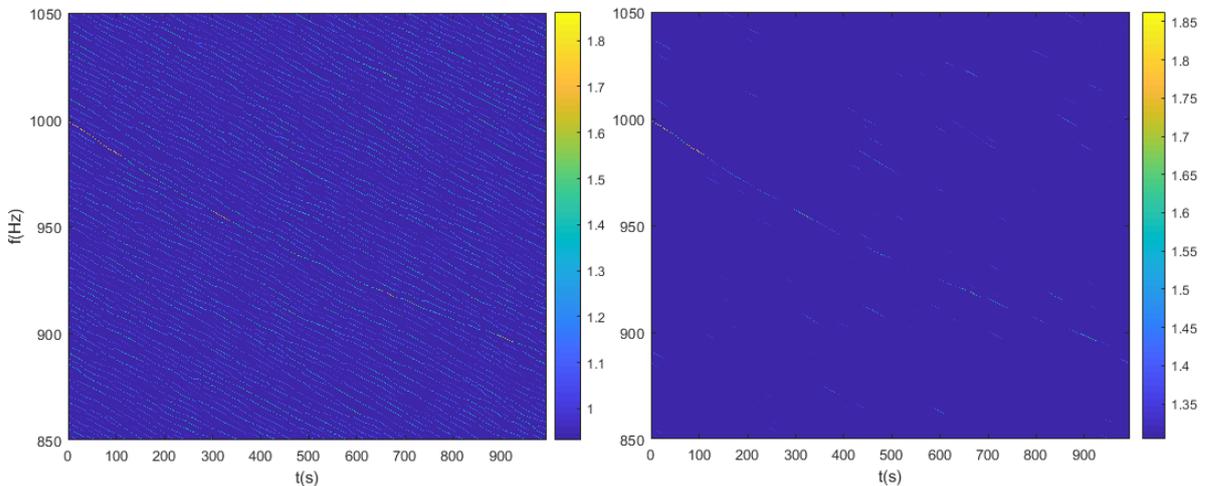


Figura 4.17: Ingrandimenti attorno al segnale di due peakmap di immagini a rumore 50 a cui è stato applicato il filtro triangolare. A sinistra: peakmap con soglia al valor medio. A destra: peakmap con soglia al 50%.

Si potrebbe pensare che la configurazione a soglia al 50% sia la migliore per rivelare il segnale ed avere al contempo il minor numero possibili di falsi candidati, ma in questo modo si rischia di perdere anche il segnale: bisogna infatti considerare che segnali diversi potrebbero avere la regione ottimale spostata rispetto a questa, di conseguenza la soglia al 50%, se per questo specifico segnale risulta ottimale, potrebbe cancellare un segnale troppo debole. Una scelta più conservativa e più robusta è invece quella opposta, di utilizzare la soglia più prossima al valor medio. La procedura attualmente esistente, infatti, tiene da conto di estrarre dal primo livello un elevato numero di candidati e di scartarli successivamente attraverso le coincidenze; in generale gli artefatti di rumore prodotti dai filtri nei dati di rivelatori diversi saranno scorrelati tra loro, ma anche qualora sopravvivessero per casuali coincidenze, scomparirebbero quando si va a ripetere nuovamente l'analisi sui dati originali con il follow-up. Inoltre molti di questi artefatti potrebbero anche non essere proprio selezionati dalla Hough generalizzata, dato che il processo di linearizzazione che si opera prima di effettuare la trasformazione vera e propria non è affatto lineare. Un segnale reale che partisse ad esempio da una frequenza maggiore rispetto a quella del segnale che è stato generato in questa sezione avrebbe anche uno spin-down maggiore, dunque una forma differente rispetto allo stesso, mentre gli artefatti di rumore prodotti hanno una forma simile al segnale anche a frequenze ben più elevate; l'effetto più probabile allora è che si producano eventualmente alcuni falsi candidati nella regione a frequenze vicine a quella del segnale sul quale si genera il filtro e meno candidati nelle regioni a frequenza lontana.

Il passo successivo di questo processo consisterà nel rendere più consistenti le considerazioni fatte nella presente sezione: queste infatti si sono ottenute su un singolo segnale di prova addizionato a 3 differenti realizzazioni di rumore bianco. Bisognerà allora effettuare numerose simulazioni di segnali al variare dei parametri caratteristici e studiare la risposta ai filtri con molte realizzazioni di rumore, nello stesso spirito

della Sezione 4.1. Occorrerà processare le peakmap ottenute nella trasformata di Hough generalizzata e valutare gli effetti dei filtri triangolari direttamente sulla loro capacità di far emergere il segnale iniettato come candidato nella mappa di Hough, studiando anche gli eventuali falsi candidati che il filtraggio può produrre. Bisognerà inoltre testare l'applicabilità del filtro triangolare combinato in condizioni di generalità nelle quali la scala del segnale può essere diversa.

Conclusioni

Nel presente elaborato è stato descritto come sfruttare le caratteristiche ed il funzionamento della trasformata di Fourier in 2D per costruire dei filtri per immagini che si applicano formalmente con le stesse logiche dei filtri in 1D. Si è visto che si possono adottare due strategie completamente opposte in base al genere di informazione che si vuole estrarre e su tale base si sono costruiti due principali tipi di filtri: il filtro triangolare, che consente di estrarre efficacemente segnali lentamente variabili nel dominio tempo-frequenza, ed il filtro periodico, che permette di mettere in evidenza dei pattern che si ripetono con una data periodicità all'interno delle immagini. L'attenzione è stata focalizzata in particolare sul filtro triangolare e sulla sua possibile combinazione con le finestre gaussiane, la quale risulta particolarmente efficace nel filtrare i segnali da processi instabili di r-mode in tempo-frequenza. Nello studiare l'effetto dei filtri su questa tipologia di segnali si sono individuate le loro caratteristiche ottimali: si è visto che definendo un rapporto segnale/rumore (SNR) in ampiezza, il maggiore guadagno lo si ha rispettivamente per il filtro triangolare semplice con il taglio in frequenza attorno al 10% e per il filtro triangolare combinato con finestra gaussiana con il taglio in frequenza oltre il 20%, risultati che rimangono validi indipendentemente dal range delle inclinazioni che delimitano il filtro. Mantenendo tale configurazione ottimale si è visto che i filtri raggiungono il massimo di efficienza quando l'immagine contenente il segnale addizionato al rumore presenta un SNR in ampiezza all'interno del range $[0.1, 1]$, che rappresenta uno dei risultati più importanti del lavoro: all'interno di questa regione il SNR dell'immagine filtrata viene aumentato di un fattore 10 per quanto riguarda il filtro triangolare semplice e di un fattore 15 per quanto riguarda il filtro triangolare combinato; tale fattore relativo di miglioramento in verità permane anche quando il SNR dell'immagine originale è al di sotto del limite 0.1, ma in quel caso il risultato perde di utilità, dato che il SNR dell'immagine risultante dal filtraggio resterebbe comunque sotto l'unità; quando invece il SNR iniziale è già al di sopra dell'unità si registra un miglioramento relativo sempre più basso, ma questa casistica non è di grande interesse perchè l'immagine non necessita di filtraggio. Si è visto inoltre che è possibile effettuare ricerche meno mirate e più generali, allo scopo di rendere la ricerca più robusta verso possibili deviazioni del segnale dai modelli, allargando il range di selezione dei filtri triangolari al prezzo di un modesto calo di efficienza.

Si è infine riprodotto il primo stadio di una ipotetica ricerca partendo da un segnale da r-mode simulato addizionato a rumore bianco e si è applicato il filtro triangolare allo spettrogramma generato a partire dai dati. Il risultato, di importanza centrale per i futuri sviluppi, mostra che l'effetto del filtraggio è di far sì che, in una configurazione in cui lo spettrogramma considerato come immagine ha $SNR \in [0.1, 1]$, i picchi relativi al segnale vengano effettivamente estratti ed inseriti nella peakmap. Senza l'applicazione del filtro tali picchi non sarebbero emersi, di conseguenza non sarebbero stati inseriti nella peakmap e quindi il segnale stesso non sarebbe stato inserito tra i candidati di primo livello.

Il lavoro qui presentato deve essere considerato come un primo, fondamentale passo di un percorso di

studio ed approfondimento che porterà ad una futura integrazione di tali tecniche di filtraggio all'interno della procedura di analisi attualmente esistente. Le tecniche di filtraggio sviluppate si sono rivelate estremamente promettenti: esse risultano facilmente maneggiabili nei parametri che le definiscono e consentono, nelle condizioni riportate sopra, di ottenere miglioramenti nel SNR di un ordine di grandezza in un'immagine contenente un segnale.

I punti aperti da affrontare nell'immediato futuro riguardano l'individuazione di una soglia ottimale di selezione dei picchi che tenga conto degli effetti sortiti dal filtraggio ed una valutazione, a monte di ciò, del criterio di selezione dei candidati più significativi all'interno delle mappe di Hough. Questi due livelli, assieme al tempo di coerenza T_{FFT} con cui si costruisce il database di FFT, vanno a determinare la sensibilità dell'algoritmo di ricerca, che ci si attende possa trarre guadagno dall'innesto di tale tecnica di filtraggio.

Merita inoltre un prossimo approfondimento la possibilità di applicare sia filtri triangolari che filtri periodici alla ricerca di segnali continui: in questo caso l'obiettivo sarebbe quello di individuare tramite i filtri le modulazioni Doppler che deformano il segnale altrimenti monocromatico e che attualmente sono causa di un grande impiego di risorse computazionali nella procedura gerarchica descritta.

Per il 2019 è prevista la partenza della nuova fase osservativa O3 dei rivelatori LIGO-Virgo; durante il periodo di spegnimento questi sono stati soggetti a numerosi upgrade che porteranno ad un miglioramento della sensibilità su tutto lo spettro di frequenze. Se si riuscisse, in concomitanza con l'aumentata sensibilità strumentale, a far corrispondere un aumento della sensibilità degli algoritmi di ricerca, questo significherebbe un aumento della probabilità di rivelare segnali gravitazionali finora rimasti inosservati.

Bibliografia

- [1] LIGO Scientific Collaboration, & Virgo Collaboration. (2018). GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs. arXiv preprint arXiv:1811.12907.
- [2] Abbott, B. P., Abbott, R., Abbott, T. D., Abernathy, M. R., Acernese, F., Ackley, K., ... & Adya, V. B. (2016). Observation of gravitational waves from a binary black hole merger. *Physical review letters*, 116(6), 061102.
- [3] Abbott, B. P., Abbott, R., Abbott, T. D., Acernese, F., Ackley, K., Adams, C., ... & Affeldt, C. (2017). GW170817: observation of gravitational waves from a binary neutron star inspiral. *Physical Review Letters*, 119(16), 161101.
- [4] Schutz, B. (2009). *A first course in general relativity*. Cambridge university press.
- [5] Straumann, N. (2004). *General relativity. With applications to astrophysics*. Springer.
- [6] Abbott, B. P., Abbott, R., Abbott, T. D., Abernathy, M. R., Acernese, F., Ackley, K., ... & Adya, V. B. (2016). Properties of the binary black hole merger GW150914. *Physical review letters*, 116(24), 241102.
- [7] Flanagan, É. É., & Hinderer, T. (2008). Constraining neutron-star tidal Love numbers with gravitational-wave detectors. *Physical Review D*, 77(2), 021502.
- [8] Hinderer, T. (2008). Tidal Love numbers of neutron stars. *The Astrophysical Journal*, 677(2), 1216.
- [9] Veitch, J., & Vecchio, A. (2010). Bayesian coherent analysis of in-spiral gravitational wave signals with a detector network. *Physical Review D*, 81(6), 062003.
- [10] Pretorius, F. (2005). Evolution of binary black-hole spacetimes. *Physical review letters*, 95(12), 121101.
- [11] Lattimer, J. M., & Prakash, M. (2001). Neutron star structure and the equation of state. *The Astrophysical Journal*, 550(1), 426.
- [12] Lasky, P. D. (2015). *Gravitational waves from neutron stars: a review*. Publications of the Astronomical Society of Australia, 32.
- [13] Giazotto, A., Bonazzola, S., & Gourgoulhon, E. (1997). Gravitational waves emitted by an ensemble of rotating neutron stars. *Physical Review D*, 55(4), 2014.

- [14] Abbott, B. P., Abbott, R., Abbott, T. D., Acernese, F., Ackley, K., Adams, C., ... & Affeldt, C. (2017). Search for post-merger gravitational waves from the remnant of the binary neutron star merger GW170817. *The Astrophysical Journal Letters*, 851(1), L16.
- [15] Shapiro, S. L. (2000). Differential rotation in neutron stars: Magnetic braking and viscous damping. *The Astrophysical Journal*, 544(1), 397.
- [16] Hotokezaka, K. (2013). K. Hotokezaka, K. Kiuchi, K. Kyutoku, T. Muranushi, Y.-i. Sekiguchi, M. Shibata, and K. Taniguchi, *Phys. Rev. D* 88, 044026 (2013). *Phys. Rev. D*, 88, 044026.
- [17] Ruffert, M., Janka, H. T., & Schäfer, G. (1995). Coalescing Neutron Stars—a Step Towards Physical Models. I. Hydrodynamic Evolution and Gravitational-Wave Emission. arXiv preprint astro-ph/9509006.
- [18] Shibata, M., & Uryū, K. (2000). Simulation of merging binary neutron stars in full general relativity: $\Gamma = 2$ case. *Physical Review D*, 61(6), 064001.
- [19] Andersson, N. (1998). A new class of unstable modes of rotating relativistic stars. *The Astrophysical Journal*, 502(2), 708.
- [20] Andersson, N., & Kokkotas, K. D. (2001). The r-mode instability in rotating neutron stars. *International Journal of Modern Physics D*, 10(04), 381-441.
- [21] Lasky, P. D., Sarin, N., & Sammut, L. (2017). Long-duration waveform models for millisecond magnetars born in binary neutron star mergers. Tech. Rep. LIGO-T1700408, <https://dcc.ligo.org/LIGO-T1700408/public>.
- [22] Miller, A., Astone, P., D'Antonio, S., Frasca, S., Intini, G., La Rosa, I., ... & Piccinni, O. J. (2018). A method to search for long duration gravitational wave transients from isolated neutron stars using the generalized FrequencyHough. arXiv preprint arXiv:1810.09784.
- [23] Turatto, M. (2003). Classification of supernovae. In *Supernovae and Gamma-Ray Bursters* (pp. 21-36). Springer, Berlin, Heidelberg.
- [24] Takiwaki, T., & Kotake, K. (2018). Anisotropic emission of neutrino and gravitational-wave signals from rapidly rotating core-collapse supernovae. *Monthly Notices of the Royal Astronomical Society: Letters*, 475(1), L91-L95.
- [25] Papoulis, A. (1965). *Probability, Random Variables, and Stochastic Processes: Solutions to the Problems in Probability, Random Variables and Stochastic Processes*. McGraw-Hill.
- [26] Astone, P., Bassan, M., Bonifazi, P., Carelli, P., Coccia, E., Cosmelli, C., ... & Modena, I. (2001). Search for periodic gravitational wave sources with the Explorer detector. *Physical Review D*, 65(2), 022001.
- [27] Abbott, B. P., Abbott, R., Abbott, T. D., Acernese, F., Ackley, K., Adams, C., ... & Affeldt, C. (2017). GW170814: a three-detector observation of gravitational waves from a binary black hole coalescence. *Physical review letters*, 119(14), 141101.

- [28] Abbott, B. P., Abbott, R., Abbott, T. D., Abernathy, M. R., Acernese, F., Ackley, K., ... & Adya, V. B. (2016). GW150914: First results from the search for binary black hole coalescence with Advanced LIGO. *Physical Review D*, 93(12), 122003.
- [29] Abbott, B. P., Abbott, R., Abbott, T. D., Abernathy, M. R., Acernese, F., Ackley, K., ... & Adya, V. B. (2016). Observing gravitational-wave transient GW150914 with minimal assumptions. *Physical Review D*, 93(12), 122004.
- [30] Abbott, B. P., Abbott, R., Abbott, T. D., Abernathy, M. R., Acernese, F., Ackley, K., ... & Adya, V. B. (2016). Binary black hole mergers in the first advanced LIGO observing run. *Physical Review X*, 6(4), 041015.
- [31] Martynov, D. V., Hall, E. D., Abbott, B. P., Abbott, R., Abbott, T. D., Adams, C., ... & Arain, M. A. (2016). Sensitivity of the Advanced LIGO detectors at the beginning of gravitational wave astronomy. *Physical Review D*, 93(11), 112004.
- [32] Abadie, J., Abbott, B. P., Abbott, R., Abernathy, M., Accadia, T., Acernese, F., ... & Allen, G. S. (2011). Beating the spin-down limit on gravitational wave emission from the Vela pulsar. *The Astrophysical Journal*, 737(2), 93.
- [33] Frasca, S., Astone, P., & Palomba, C. (2005). Evaluation of sensitivity and computing power for the Virgo hierarchical search for periodic sources. *Classical and Quantum Gravity*, 22(18), S1013.
- [34] Astone, P., Colla, A., D'Antonio, S., Frasca, S., & Palomba, C. (2014). Method for all-sky searches of continuous gravitational wave signals using the frequency-Hough transform. *Physical Review D*, 90(4), 042002.
- [35] Abbott, B. P., Abbott, R., Abbott, T. D., Abernathy, M. R., Acernese, F., Ackley, K., ... & Adya, V. B. (2016). GW150914: The Advanced LIGO detectors in the era of first discoveries. *Physical review letters*, 116(13), 131103.
- [36] Astone, P., Frasca, S., & Palomba, C. (2005). The short FFT database and the peak map for the hierarchical search of periodic sources. *Classical and Quantum Gravity*, 22(18), S1197.